JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 9 (2023) ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Original Papers

Viewpoints

Reviews

Protocol

Corrigenda and Addenda

Short Paper

Original Paper

Identifying Hot Spots of Tuberculosis in Nigeria Using an Early Warning Outbreak Recognition System: Retrospective Analysis of Implications for Active Case Finding Interventions

Chidubem Ogbudebe[1], MPH; Dohyo Jeong[2], MA; Bethrand Odume[1], MD, MPH; Ogoamaka Chukwuogo[1], MD, MA; Cyril Dim[3], MD, PhD; Sani Useni[1], MD, MPH; Okey Okuzu[4], MBA; Chenchita Malolan[5], MPH; Dohyeong Kim[2], PhD; Fiemu Nwariaku[5], MD, PhD; Nkiru Nwokoye[1], PhD; Stephanie Gande[1], MPH; Debby Nongo[6], MA; Rupert Eneogu[6], MD, MSc; Temitayo Odusote[6], MD, MPH; Salewa Oyelaran[6], MHSA; Obioma Chijioke-Akaniro[7], MD, MPH; Nrip Nihalani[8], MSc; Chukwuma Anyaike[7], MD, MPH; Mustapha Gidado[9], MD, PhD

[1]Technical Division, KNCV Tuberculosis Foundation, Abuja, Nigeria
[2]University of Texas, Dallas, TX, United States
[3]College of Medicine, University of Nigeria Teaching Hospital, Enugu, Nigeria
[4]InStrat Global Health Solutions, Abuja, Nigeria
[5]University of Texas Southwestern Medical Center, Dallas, TX, United States
[6]Office of Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome and Tuberculosis, United States Agency for International Development, Abuja, Nigeria
[7]National Tuberculosis, Leprosy and Buruli Ulcer Control Program, Federal Ministry of Health, Abuja, Nigeria
[8]Plus91 Technologies, Delhi, India
[9]Program Management Unit, KNCV Tuberculosis Foundation, Hague, Netherlands

Corresponding Author:
Chidubem Ogbudebe, MPH
Technical Division
KNCV Tuberculosis Foundation
Plot 564/565 Independence Avenue
Central Business District
Abuja
Nigeria
Phone: 234 8038926544
Email: cogbudebe@kncvnigeria.org

Abstract

Background: Undiagnosed tuberculosis (TB) cases are the major challenge to TB control in Nigeria. An early warning outbreak recognition system (EWORS) is a system that is primarily used to detect infectious disease outbreaks; this system can be used as a case-based geospatial tool for the real-time identification of hot spot areas with clusters of TB patients. TB screening targeted at such hot spots should yield more TB cases than screening targeted at non–hot spots.

Objective: We aimed to demonstrate the effectiveness of an EWORS for TB hot spot mapping as a tool for detecting areas with increased TB case yields in high TB-burden states of Nigeria.

Methods: KNCV Tuberculosis Foundation Nigeria deployed an EWORS to 14 high-burden states in Nigeria. The system used an advanced surveillance mechanism to identify TB patients' residences in clusters, enabling it to predict areas with elevated disease spread (ie, hot spots) at the ward level. TB screening outreach using the World Health Organization 4-symptom screening method was conducted in 121 hot spot wards and 213 non–hot spot wards selected from the same communities. Presumptive cases identified were evaluated for TB using the GeneXpert instrument or chest X-ray. Confirmed TB cases from both areas were linked to treatment. Data from the hot spot and non–hot spot wards were analyzed retrospectively for this study.

Results: During the 16-month intervention, a total of 1,962,042 persons (n=734,384, 37.4% male, n=1,227,658, 62.6% female) and 2,025,286 persons (n=701,103, 34.6% male, n=1,324,183, 65.4% female) participated in the community TB screening outreaches in the hot spot and non–hot spot areas, respectively. Presumptive cases among all patients screened were 268,264

(N=3,987,328, 6.7%) and confirmed TB cases were 22,618 (N=222,270, 10.1%). The number needed to screen to diagnose a TB case in the hot spot and non–hot spot areas was 146 and 193 per 10,000 people, respectively.

Conclusions: Active TB case finding in EWORS-mapped hot spot areas yielded higher TB cases than the non–hot spot areas in the 14 high-burden states of Nigeria. With the application of EWORS, the precision of diagnosing TB among presumptive cases increased from 0.077 to 0.103, and the number of presumptive cases needed to diagnose a TB case decreased from 14.047 to 10.255 per 10,000 people.

(JMIR Public Health Surveill 2023;9:e40311) doi:10.2196/40311

KEYWORDS

early warning outbreak recognition system; active case finding; WHO-four-symptom screen; GeneXpert; active case; cluster; early warning; hot spot; mapping; disease spread; infection spread; retrospective study; retrospective analysis; surveillance; outbreak; TB; tuberculosis; infectious disease; case finding; communicable disease

Introduction

Tuberculosis (TB) continues to be a global health concern, which explains why ending the disease epidemic by 2030 is a target of sustainable development goals [1]. Regarding the burden of TB at the country level, Nigeria is sixth globally and the first in Africa [2]; therefore, controlling this chronic infectious disease is a top priority for the national TB program [3]. Globally, the number of annual case notifications is usually far below the estimated incident cases. Unfortunately, this gap (ie, missing TB cases), which had consistently narrowed over the years, widened in 2020 due to a large drop in the number of newly diagnosed TB cases due to the impact of the COVID-19 pandemic [2,4]. In 2020, there were about 9.9 million estimated incident TB cases worldwide, but only 5.8 million were notified, leaving a large gap of about 4.1 million missing cases; notably, Nigeria contributed significantly to this global gap, ranking first in Africa [2].

These missing TB cases could either be due to underreporting or underdiagnosis. In Nigeria, however, underdiagnosis contributes to most missing cases; this calls on the country to strengthen access to high-quality screening, diagnostic, and treatment services [5]. The KNCV Tuberculosis Foundation Nigeria (KNCV Nigeria) continues to assist the Nigerian TB control program in this regard through several projects, including the successful demonstration of the excellent capacity of TB loop-mediated isothermal amplification technology for TB diagnosis among adults in Nigeria [6], as well as the need for X-ray–based mass TB screening in Nigerian prisons [7].

Because undiagnosed TB cases are the major challenge to TB control in Nigeria, any strategy that will identify ongoing TB transmission hot spots in the community is crucial for finding and treating missing TB cases. Electronic recording and reporting of TB are becoming adopted widely by the national TB control programs [8]. The electronic data can be imported into a geographic information system (GIS) for mapping and spatial analysis; thus, the data can be used to identify TB transmission hot spots in the community [8]. Such an electronic GIS strategy has been used in early warning systems (EWSs) for infectious disease surveillance [9], such as the China Infectious Disease Automated Alert and Response System (CIDARS) [10] and an early warning outbreak recognition system (EWORS) implemented by the Indonesian Ministry of Health [11]. Both systems serve as complementary country-wide disease surveillance tools; however, while the CIDARS is a case-based EWS, the EWORS in Indonesia is a syndromic surveillance-based EWS [9]. In Nigeria, an EWORS is operational at the subnational level for routine electronic disease surveillance—the system automatically sends elevated disease activity messages to appropriate disease surveillance officers, which triggers the investigation of the affected areas to prevent outbreaks [12].

Though the EWORS is an early-warning system for infectious diseases to detect outbreaks and guide control practice, it can also predict TB hot spots at the community level by identifying TB patients’ residences in clusters. Because TB is infectious, it is expected that community-based active case finding (ACF) in such EWORS-identified hot spot areas will yield more TB cases when compared to non–hot spot areas. To demonstrate this expectation, KNCV Nigeria applied an EWORS as a geospatial tool for real-time identification of hot spot areas with high TB prevalence for targeted community-based ACF interventions. This innovative project aimed to demonstrate the effectiveness of EWORS TB hot spot mapping for increased TB case yield in 14 states of Nigeria with a high TB burden.

Methods

Overview

This is a retrospective review of data from community-based active TB screening interventions by KNCV Nigeria in EWORS-identified TB hot spot areas in 14 states of Nigeria with a high TB burden. This public health intervention was part of the US Agency for International Development–funded TB Local Organization Network Regions 1 and 2 Project. The 14 states involved include 8 northern Nigerian states (Bauchi, Benue, Kaduna, Katsina, Kano, Nasarawa, Plateau, and Taraba) and 6 southern Nigerian states (Anambra, Akwa Ibom, Cross River, Delta, Imo, and Rivers). The 16-month intervention lasted from October 2020 to January 2022.

KNCV Nigeria instituted TB hot spot analysis using InStrat Global Health Solutions’ EWORS to inform community mass TB screening interventions in the identified communities in the 14 project states. The EWORS used an advanced surveillance mechanism to identify TB patients’ residences in clusters, enabling it to predict hot spots at the ward and community level. The EWORS hot spot analytic dashboard portal was developed per state. State hot spot data visualization included heat maps that allowed users to drill down to the ward or community level to pinpoint locations of active cases. Every day, EWORS pulled real-time data from a Commcare-based mobile app—an open-source data capture tool running on Android smartphones that records routine TB surveillance data at health facilities providing TB services in the project states. The Commcare cloud server sent data to the EWORS via a custom data transfer application programming interface. Information was synchronized that predicted the likelihood of the occurrence of disease clusters at the ward or community level and included patient-level clinical data (ie, TB symptoms, TB investigations, TB diagnoses, and treatment status) and nonclinical data (ie, state of residence, local government area [LGA], and ward).

The EWORS triggered an alarm when a hot spot was identified. For the project, a hot spot was identified when any ward or community had a confirmed TB case count of 5 or more, and the presumptive to confirmed case count was less than 10. Similar alarm thresholds were set for individual states. The thresholds that drove the state algorithms were informed by the historical state-level TB notification and epidemiological data. The EWORS applied the incoming ward level data against these thresholds to determine wards and communities that met the thresholds and automatically generated alarms for affected communities. The algorithms analyzed data on a rolling 7-day basis, and alarm notification emails were sent to field officers to institute community outreach activities. Following the notification, the state and local teams would review the alarms in concert with the heat maps, then mobilize and conduct mass TB screening outreach activities at the affected alarm locations. We identified 121 hot spot areas—an average of 10 (range 6-25) hot spots per state. Field officers used the World Health Organization (WHO) 4-symptom screen (W4SS) to offer TB screening to consenting individuals during community outreach activities. Symptomatic outreach participants (presumptive TB patients) were further evaluated for TB using the GeneXpert instrument or chest X-rays (for patients that could not produce sputum). Confirmed TB cases were linked to treatment and to the national TB program.

To demonstrate the effectiveness of the EWORS analytic system in identifying missing TB cases, this study compared the TB case yield from the hot spot wards and areas with the yield obtained from data from health outreach activities in 213 non–hot spot wards and areas in the same 14 project states. The non–hot spot areas were selected purposely from all the LGAs where one or more community hot spot areas were identified. The methods of TB screening, evaluation, notification, and treatment for consenting patients in the non–hot spot areas were the same as described for the hot spot areas.

Deidentified project data relevant for this report from the hot spots and non–hot spots were retrieved from the Commcare-based mobile app and analyzed retrospectively using SPSS (IBM Corp). The data included the participants' age, sex, TB case yield, and the number needed to screen (NNS) for each location area (hot spot and non–hot spot). Proportions were compared with the chi-square test, and a *P* value of less than .05 was considered significant. TB case yield was operationally defined as TB cases detected directly through the mass TB screening activities, calculated as a percentage of persons screened versus presumptive TB cases within the defined geographical area.

To assess the effectiveness of EWORS, this study used hot spot analyses to verify whether there were statistically significant spatial clusters of confirmed TB cases. Getis-Ord Gi* and optimized hot spot analyses were conducted to reveal the spatial heterogeneity of hot spots and non–hot spots [13]. Additionally, we examined spatial-stratified heterogeneity, which indicated the spatial differences between multiple layers or domains [14]. Ignoring the characteristics of spatial-stratified heterogeneity in the analysis could have led to model misspecification and estimation errors [15]. In this study, we used the Q statistic to measure the degree of stratified heterogeneity for hot spot analysis. The Q statistic is a value of the intensity of spatially stratified heterogeneity. It ranges between 0 and 1 and is calculated by comparing the stratified heterogeneities within the attribute and those between the strata. The higher the Q value, the stronger the stratification heterogeneity effect [14]. The Q statistic values from the Getis-Ord Gi* and optimized hot spot analyses were compared to each other to select the best hot spot model, which was then used to estimate the effectiveness of the EWORS. The Q statistic was calculated using the R package *geodetector*.

Ethical Considerations

The study was determined to be a nonresearch program evaluation. It required no direct contact with human subjects (ie, there were no interviews and no sample collection). Also, only deidentified, pooled program data that formed part of the standard of care were used; thus, informed consent was not required.

Results

To identify TB hot spots in the 14 study states, we first mapped them based on the number of diagnosed TB cases per 10,000 people. The result is shown in Figure 1. As shown on the map, Plateau, Nasarawa, and Akwa Ibom states had the highest numbers of confirmed cases. On the other hand, the states of Imo, Kaduna, and Anambra showed relatively low numbers.

Table 1 shows the estimated numbers of cases, TB tests, and treatments in the 14 states based on the numbers of diagnosed TB cases per 10,000 people. These results confirmed that to accurately identify hot spots and ultimately evaluate the effectiveness of the EWORS, it would be necessary to consider various factors, such as the number of presumptive TB cases and disease density, as well as the number of TB diagnoses in each state.

TB is a representative infectious disease, so to find precise hot spots, we considered the number of occurrences and the density of diagnosed cases. Therefore, we performed a kernel density analysis based on the location where TB was diagnosed. This analysis differentiated between regions with a high number of diagnosed TB cases but low disease density and areas with a low number of diagnosed cases but high disease density. Using a threshold for kernel density of 46,242.3 meters, we applied planar kernel density estimation for mapping. In planar kernel density estimation, an area is characterized as a 2D, uniform

Euclidean space, and its density is estimated at many regularly spaced positions [16,17]. Figure 2 shows the results of the kernel density analysis. The density of TB was relatively high in the northern and southern regions but with different characteristics. The northern region showed a very high density only in a specific area centered on Kano state. On the other hand, TB cases were widely distributed in the southern regions with a relatively low density. We defined TB hot spots by considering the number and density of hot spots using the optimal hot spot analysis technique with LGAs as units.

The Moran I test revealed a statistically significant level of spatial autocorrelation among confirmed tuberculosis cases (Moran I=0.773, *z* score=31.54), which confirms the validity of the hot spot analysis. Table 2 shows the results of the Getis-Ord Gi* and optimized hot spot analyses, indicating that the Getis-Ord Gi* method had a higher spatial stratified heterogeneity measured by the Q statistic than the optimized method. Therefore, in this study, the effectiveness of the EWORS was estimated using the Getis-Ord Gi* method.

Figure 1. Tuberculosis cases diagnosed per 10,000 people. Blue dots indicate diagnosed patients. Colors closer to red indicate a higher number of tuberculosis cases diagnosed per 10,000 people. Created using ArcGIS® software by Esri and are used under license. Copyright © Esri.

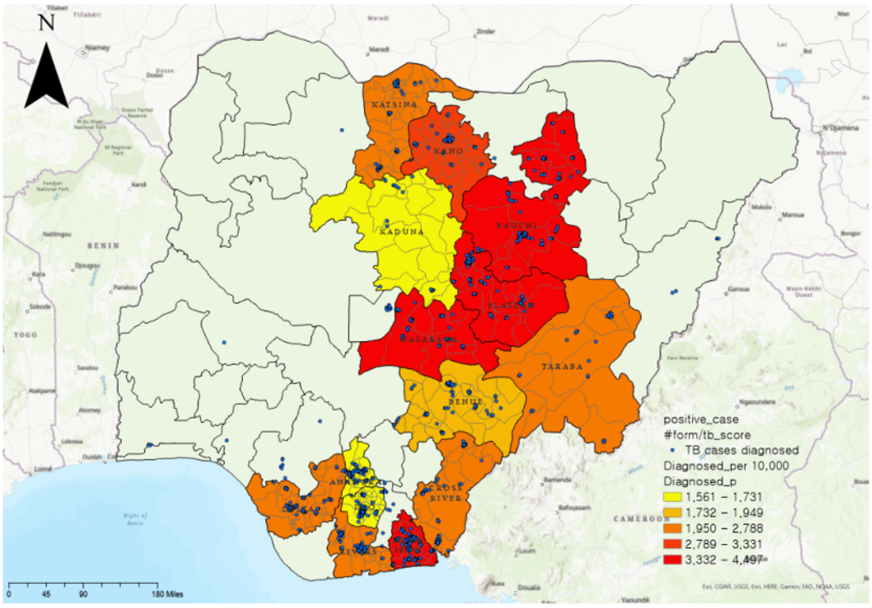


Table 1. Ranking of the 14 study states for number of screened tuberculosis cases. All values represent cases per 10,000 people.

	State	Screened, n	Presumptive, n	Tested, n	Diagnosed, n	Treated, n
1	Plateau	957.29	50.16	38.71	4.25	4.08
2	Taraba	696.20	55.08	41.38	5.23	4.62
3	Bauchi	667.17	38.10	25.90	3.91	3.86
4	Nasarawa	598.21	44.80	34.50	3.47	3.31
5	Delta	585.78	34.05	28.40	2.63	2.34
6	Katsina	487.82	29.69	26.68	2.36	2.15
7	Cross River	460.68	32.18	25.52	2.25	2.06
8	Kano	442.55	33.01	27.61	3.01	2.90
9	Akwa Ibom	394.03	27.07	22.32	3.96	3.42
10	Anambra	360.45	38.01	30.04	1.61	1.52
11	Rivers	347.40	25.67	22.17	2.16	1.90
12	Kaduna	343.31	21.20	15.97	1.56	1.52
13	Imo	325.84	19.89	15.15	1.48	1.43
14	Benue	315.41	22.31	17.33	1.57	1.46

Figure 2. Kernel density based on number of diagnosed tuberculosis cases. Darker shades of purple indicate a higher density of occurrences. Created using ArcGIS® software by Esri and are used under license. Copyright © Esri.

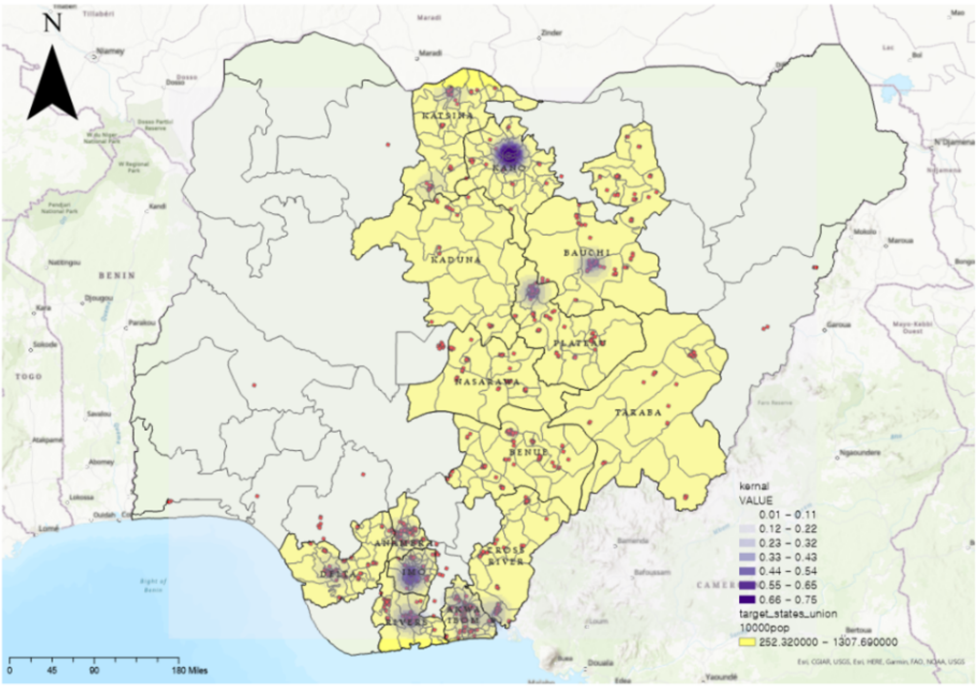


Table 2. Results for the Q statistic.

Hot spot model	Value	*P* value
Optimized method		
F	404.88	<.001
Q statistic	0.603	<.001
Getis-Ord Gi* method		
F	251.61	<.001
Q statistic	0.709	<.001

Figure 3 shows the hot spot identification results. The hot spot areas had a higher TB incidence with a higher density than the surrounding area. On the other hand, the cold spots were areas with a lower occurrence frequency and density than the surrounding area. The unmarked areas indicate that there might have been confirmed cases, but the hot spot clustering was not statistically significant.

With the EWORS, we were able to identify 121 hot spots in LGAs, 113 cold spots, and 100 other regions, but cold spots and other areas were combined as "non–hot spot" areas for analysis in this study. Therefore, we identified 121 hot spot regions and 213 non–hot spot regions to confirm the effectiveness of the EWORS. We first examined the characteristics of the TB outbreaks in hot spots and non–hot spots. The results are shown in Table 3. In the TB case-finding intervention, 1,962,042 persons (n=734,384, 37.4% male, n=1,227,658, 62.6% female) versus 2,025,286 persons (n=701,103, 34.6% male, n=1,324,183, 65.4% female) participated in the community TB screening outreaches in the hot spot and non–hot spot areas, respectively. The participants' most and least common age groups for both areas were 25 to 34 years (hot spots: 447,911/1,962,042, 22.8%; non–hot spots: 506,913/2,025,286, 25%) and 65 years or older (hot spots: 117,212/1,962,042, 6%; non–hot spots: 133,360/2,025,286, 6.6%). Only 127,733 of 1,962,042 (6.5%) patients in the hot spot areas were presumptive TB cases, against 140,531 of 2,025,286 (6.9%) patients in the non–hot spot areas. Details of the distribution of participants in both hot spot and non–hot spot regions according to sex, age group, and number of presumptive TB cases identified and evaluated for TB are shown in Table 3.

As shown in Table 4, of the 3,987,328 patients screened, 222,270 (ie, 5.5% of all patients screened during the project) were further evaluated for TB in the hot spot and non–hot spot areas. The patients were predominantly female (n=2,551,841, 64%) and the modal age group for presumptive TB cases tested was 25 to 34 years (n=954,824, 23.9%). A total of 22,618 patients were confirmed TB cases, giving a TB yield of 10.2% (n=222,270) among presumptive TB cases evaluated and a yield of 0.6% (3,987,328) among all patients screened for TB. Among the TB cases, males predominated (n=13,607, 60.2%) while the main age group was 25 to 34 years (n=5,325, 23.5%). Out of all TB cases diagnosed, 21,071 (93.1%) were placed on treatment.

Figure 3. Hot spot analysis of diagnosed tuberculosis cases per 10,000 people. Red areas on the map indicate hot spots discovered through the early warning outbreak recognition system and blue areas indicate cold spots. Created using ArcGIS® software by Esri and are used under license. Copyright © Esri.

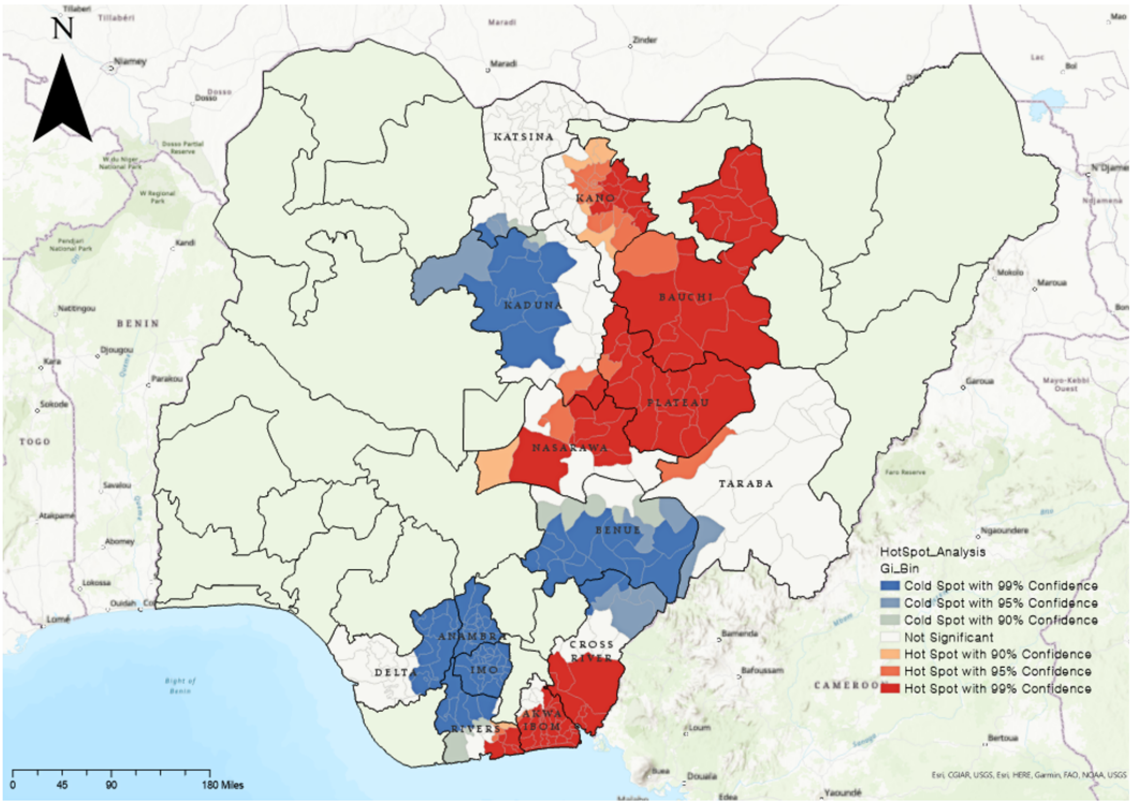


Table 3. Distribution of participants by sex, age group, and number of presumptive tuberculosis cases in hot spot and non–hot spot areas.

Characteristics	Hot spot area patients (n=1,962,042)	Non–hot spot area patients (n=2,025,286)
Sex, n (%)		
Male	734,384 (37.4)	701,103 (34.6)
Female	1,227,658 (62.6)	1,324,183 (65.4)
Age group (years), n (%)		
0-4	211,792 (10.8)	258,942 (12.8)
5-14	169,125 (8.6)	130,761 (6.5)
15-24	329,470 (16.8)	282,060 (13.9)
25-34	447,911 (22.8)	506,913 (25)
35-44	326,720 (16.7)	350,993 (17.3)
45-54	214,122 (10.9)	214,544 (10.6)
55-64	145,690 (7.4)	147,713 (7.3)
≥65	117,212 (6)	133,360 (6.6)
Presumptive TB[a] cases, n (%)		
Identified TB	127,733 (6.5)	140,531 (6.9)
Tested for TB	108,281 (5.5)	113,989 (5.6)

[a] TB: tuberculosis.

Table 4. Patients screened for tuberculosis, presumptive tuberculosis cases evaluated, and diagnostic outcomes in the hot spot and non–hot spot areas.

Characteristic	Patients screened (N=3,987,328)	Presumptive TB[a] cases evaluated (222,270/3,987,328, 5.5%)	All TB cases diagnosed (22,618/222,270, 10.1%)	All TB cases treated (21,071/22,618, 93.1%)
Sex, n (%)				
Male	1,435,487 (36)	97,171 (43.7)	13,607 (60.2)	12,762 (60.6)
Female	2,551,841 (64)	125,099 (56.3)	9011 (39.8)	8309 (39.4)
Age groups (years), n (%)				
0-4	470,734 (11.8)	16,366 (7.4)	608 (2.7)	552 (2.6)
5-14	299,886 (7.5)	18,309 (8.2)	1142 (5)	1062 (5)
15-24	611,530 (15.3)	31,962 (14.4)	3472 (15.4)	3286 (15.6)
25-34	954,824 (23.9)	47,202 (21.2)	5325 (23.5)	4945 (23.5)
35-44	677,713 (17)	40,612 (18.3)	4884 (21.6)	4487 (21.3)
45-54	428,666 (10.8)	27,281 (12.3)	3142 (13.9)	2941 (14)
55-64	293,403 (7.4)	20,094 (9)	2102 (9.3)	1982 (9.4)
≥65	250,572 (6.3)	20,444 (9.2)	1943 (8.6)	1816 (8.6)

[a] TB: tuberculosis.

Table 5 compares the TB diagnostic outcomes per 10,000 people between the hot spot and non–hot spot areas based on the EWORS. We used precision, number needed to screen (NNS), and number needed to test (NNT) to examine the effectiveness of TB diagnosis in the EWORS hot spot and non–hot spot areas. Precision is the ratio of true diagnosed cases to total presumptive cases; the higher the value, the higher the diagnostic accuracy. On the other hand, NNS and NNT refer to the ratios required for screening and testing for accurate diagnoses. Therefore, the lower the value, the lower the social cost. The prevalence of TB among all patients in the hot spot areas (3.76 per 10,000) was significantly higher than that in non–hot spot areas (2.14 per 10,000, $P<.001$; 95% CI 0.96-2.45). The precision was 10.5% for hot spots and 7.5% for the non–hot spots. The NNS for the hot spot areas was lower than the non–hot spot areas (146.22 versus 193.44). The NNT required fewer cases in hot spots, with an NNT in hot spots and non–hot spots of 9.51 and 13.43, respectively.

We thus confirmed that the EWORS could effectively identify TB hot spots and cold spots. According to the EWORS, the number of confirmed cases of TB per 10,000 people was higher in hot spot areas. Higher accuracy and a lower NNS and NNT were needed for hot spot areas with a higher risk of TB transmission. Based on these results, we compared the changes before and after application of the EWORS to measure its effectiveness, with particular attention to its effectiveness based on an analysis of the changes in precision and NNT before and after EWORS was applied in 14 major states. A 1-tailed, paired *t* test was used as an analysis method. Figure 4 shows the precision before and after EWORS application. According to the map results, before EWORS was applied, the average precision was 7.7% and was less than 10% in all regions. On the other hand, the precision after applying EWORS increased to an average of 10.3%. In other words, in the region where EWORS was applied, the precision to identify actual TB cases increased significantly.

Table 5. Comparison of tuberculosis diagnostic outcomes between the hot spot and non–hot spot areas.

	Hot spot area	Non–hot spot area	*P* value
Patients screened per local government area (per 10,000 people), n	549.80	413.97	N/A[a]
Presumptive cases evaluated for tuberculosis (per 10,000 people), n	35.79	28.75	N/A
All tuberculosis cases diagnosed (per 10,000 people), n	3.76	2.14 (95% CI 0.96-2.45)	<.001
Precision	0.105	0.075	N/A
Number needed to screen (per 10,000 people), n	146.22	193.44	N/A
Number needed to test (per 10,000 people), n	9.51	13.43	N/A

[a] N/A: not applicable.

Figure 4. Comparison of diagnostic precision before and after application of an early warning outbreak recognition system. Light blue indicates precision less than 10%, and dark blue indicates precision higher than 10%. A: Before the system was applied (average precision: 7.7%); B: after the system was applied (average precision: 10.3%). Created using ArcGIS® software by Esri and are used under license. Copyright © Esri.

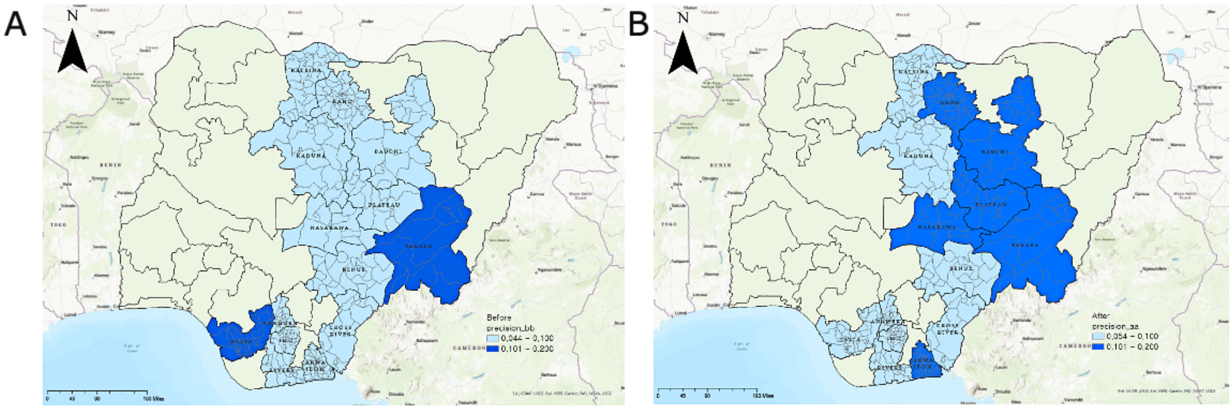


Table 6 shows the paired *t* test results for changes in precision before and after the EWORS was applied. Comparing the overall scores shows that the precision before the EWORS was applied was 7.7%. However, after applying the EWORS, the precision was 10.3%. This was a statistically significant improvement of about 2.7%. We also compared the precision changes by dividing the hot spot and non–hot spot regions. Prior to the application of the EWORS, the precision of the hot spot areas was 6.4%. However, after applying the system, the accuracy of the actual TB diagnoses increased to 11.9%. Similarly, in non–hot spot regions, the precision increased to 9.2% after the EWORS was applied. In other words, it was possible to improve the precision of diagnosing cases of TB among presumptive cases with the EWORS.

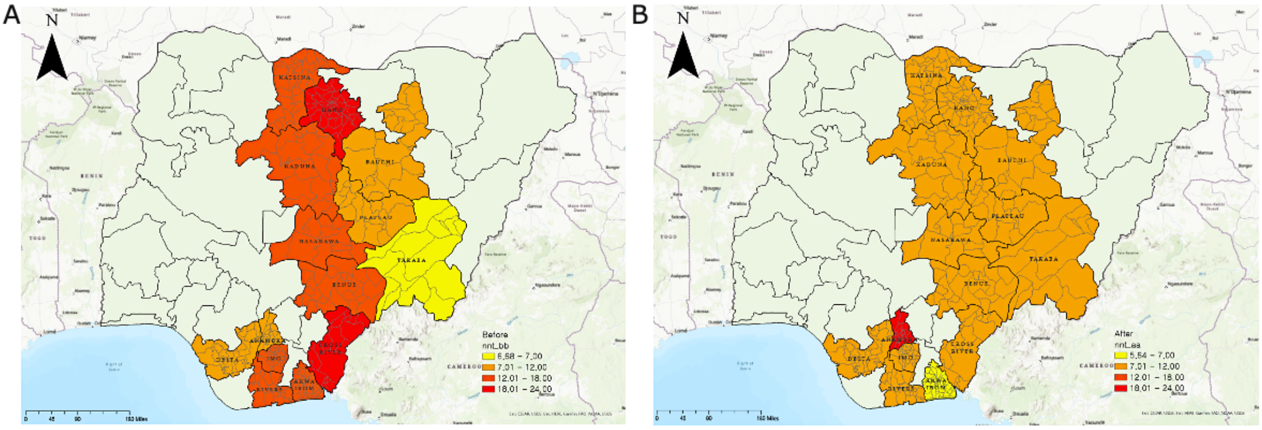
Next, we compared the NNT for diagnosing TB among presumptive cases before and after the EWORS was applied. Figure 5 shows the NNT before and after the EWORS was applied. According to the figure, a high NNT was required in most states before the EWORS was applied. However, after the EWORS was applied, a very low NNT was needed in all but one state. In other words, NNT, representing the social cost of diagnosed cases, was found to be substantially reduced after the implementation of EWORS.

Table 6. The results of the paired *t* test for precision. Precision is defined as true positive cases / (correct presumptive cases + incorrect presumptive cases).

	Before the EWORS[a]	After the EWORS	Difference	$t_{(668)}$	*P* value
Total precision (SE)	7.7% (0.12%)	10.3% (0.15%)	2.7% (0.19%)	13.764	<.001
Hot spot precision (SE)	6.4% (0.14%)	11.9% (0.26%)	5.5% (0.29%)	18.613	<.001
Non–hot spot precision (SE)	8.6% (0.15%)	9.2% (0.11%)	0.5% (0.19%)	2.747	.003

[a] EWORS: early warning outbreak recognition system.

Figure 5. Comparison of numbers needed to test for tuberculosis diagnosis before and after the early warning outbreak recognition system was applied. Deeper shades of red indicate a higher number. A: Before the system was applied (average number needed to test: 14.047); B: after the system was applied (average number needed to test: 10.255). Created using ArcGIS® software by Esri and are used under license. Copyright © Esri.



Finally, we statistically compared the change in NNT before and after the EWORS was applied. The results are shown in Table 7. NNT before the EWORS was applied was 14.047 and decreased to 10.255 after it was applied, a statistically significant difference. Also, when the hot spot regions were classified, the NNT in the hot spot regions before the EWORS was applied was very high, at 16.788. However, after the system was applied, the NNT in the hot spot areas decreased to 8.829. Similarly, NNT was significantly reduced after using the EWORS in the non–hot spot regions. In other words, by applying the EWORS, NNT in the high-risk hot spot areas, as well as the entire area, was statistically significantly reduced.

Table 7. The results of paired *t* tests for the number needed to test to diagnose cases of tuberculosis among presumptive cases.

	Before the EWORS[a]	After the EWORS	Difference	$t_{(668)}$	*P* value
Total number needed to test (SE)	14.047 (0.2229)	10.255 (0.1457)	3.792 (0.2663)	14.237	<.001
Hot spot number needed to test (SE)	16.788 (0.3756)	8.829 (0.1522)	7.958 (0.4053)	19.633	<.001
Non–hot spot number needed to test (SE)	12.006 (0.1498)	11.316 (0.1953)	0.690 (0.2462)	2.802	.004

[a] EWORS: early warning outbreak recognition system.

Discussion

KNCV Nigeria proposed that an EWORS, which is primarily used to detect and control infectious-disease outbreaks, could predict TB hot spots at the ward and community levels by identifying TB patients' residences in clusters. This project made innovative use of an EWORS for identifying TB hot spot areas in 14 states of Nigeria and found that the TB case yield in hot spot areas was significantly higher than in non–hot spot areas. Such large-scale, community-based ACF was necessary to counter the negative effect of the COVID-19 pandemic on TB notification and care at both the global and local levels [2,4]. Despite the predominance of female patients during outreach, there was a higher predominance of males among the TB cases, as expected because of the known epidemiology of TB in Nigeria [18-20]. The predominant age group also agreed with the reported pattern in the Nigerian population [7,18,21].

The WHO recommends systematic screening for TB among the general population in areas with an estimated TB prevalence of 0.5% or higher [22]; therefore, with an average TB case yield per LGA of 3.8% in the hot spot areas and 2.14% in the non–hot spot areas, the ACF interventions were justified in all areas of the project states. Moreover, this project demonstrated that because the EWORS tracked clusters of households with TB cases, it offered the added advantage of predicting communities with higher than usual TB prevalence (ie, hot spot areas), which will help prioritize community TB screening activities in underresourced settings like Nigeria. Despite the relatively lower efficiency of the W4SS TB screening method in the general population, it was suitable for this large-scale project because it is easy to implement, cheap, highly acceptable to patients, and accurate [22]. However, with a sensitivity of about 71%, a few false-negative TB cases were likely missed [20]. The same algorithm was used for both hot spot and non-spot areas, so the effect of this limitation on the study's estimates would have been nondirectional. Nevertheless, with a reported sensitivity of 85% to 94%, mobile chest X-rays for community ACF, where affordable, would have reduced the missed cases and increased the TB yield for both areas [22].

The average precision and diagnostic accuracy for identifying a patient as a TB case following W4SS TB screening and further testing with GeneXpert instrument/chest X-ray in the EWORS-identified hot spot areas was about twice that of the non–hot spot areas in this project (Table 5). This finding suggests that the EWORS can be used to enhance the identification of missed TB cases in Nigerian communities. Most importantly, in underresourced settings such as Nigeria, it can be used to prioritize the siting of community health outreaches for TB to save cost. The EWORS was designed as an early warning system to detect infectious disease outbreaks and guide control practices, so its adaptation in this project as a warning system for areas or wards with a high cluster of TB cases, marking them for follow-up with community ACF, was innovative and novel. The possible effect of the W4SS TB screening regarding missed cases discussed above was a limitation. Also, the project did not evaluate patients without TB symptoms for TB preventive treatment eligibility.

The analysis showed that the EWORS effectively identified tuberculosis-risk hot spots in the 14 Nigerian states. In particular, it offered high precision, a low NNS, and a low NNT for hot spot areas. We also measured the effectiveness of applying the EWORS and found that the precision increased significantly after it was applied. In particular, the precision of hot spots was dramatically improved, and the number of presumptive cases required for a diagnosis of TB saw a statistically significant decrease after the EWORS was applied. Taken together, the EWORS not only effectively identified hot spots but also improved precision after it was applied. Prior to the introduction of the EWORS, the TB surveillance process to identify hot spots in Nigeria involved the manual collation of TB notification data by a designated local government TB and leprosy supervisor (LGTBLS) at the LGA level. Each LGTBLS collated data from each health facility within the LGA on case findings at the end of the quarter and submitted the data to the state TB program; thereafter, the data were reviewed to determine areas of high or low TB notification [23].

The EWORS applies a syndromic surveillance-based EWS and uses kernel density estimation to create a smooth, continuous surface for the density of observations. The application of Bayesian techniques in EWORS hot spot identification helps to reduce the rates of false-negative and false-positive hot spots in comparison to classical methods of hot spot identification, such as confidence intervals [24]. This is in line with other recent tools, such as Gettis-Ord Gi*, StatsCan, and Local Indicators of Spatial Association (LISA), which are used to improve the identification of hot spots through the "mining" of spatial patterns and applying hot-spot related factors, which can help distinguish between hot spots and normal areas [25]. However, while the Gettis-Ord Gi* and StatsCan add definition to maps by estimating the density distribution of events at the local level, which allows assessment of the spatial association of a particular observation or in a study area and identifies statistically significant hot spots and cold spots, the LISA technique measures the extent to which points that are close to a given point have similar values based on a measure of contiguity among these units within a specified radius, and thus is useful for identifying local spatial autocorrelation [26]. These methods serve as complementary techniques to address the limitations of mapping methods that use geographic boundary areas, such as uniform grid cells (or quadrants) or census blocks as the unit of spatial analysis (eg, thematic maps) to depict patterns of spatial clustering [24].

However, our approach has some limitations. First, we used planar kernel density estimation to show the density of tuberculosis outbreaks by region. Since the planar method assumes spatial homogeneity, it is impossible to confirm the movement of patients with confirmed tuberculosis over time [16]. Also, we used the optimal threshold for kernel density mapping, but there may be differences in the density map depending on the choice of reference point. In addition, we analyzed hot spots by LGA unit, but aggregated them at the state level when comparing efficiencies. This approach may find hot spots more accurately by considering internal deviations within each state, but the efficiency comparison is restricted to only the state level. Additionally, our hot spot analysis results cannot be used to determine causation [17].

We conclude that active TB case finding in EWORS-mapped hot spot areas yielded a higher number of TB cases than in the non–hot spot areas in 14 high-burden states of Nigeria. As the national TB program scales up active TB case finding in Nigeria, the EWORS should be used to identify and prioritize wards or communities for TB screening health outreach to optimize the use of available resources.

Acknowledgments

We acknowledge the important contributions of Nigerian community leaders in supporting local mobilization and awareness efforts. We thank the national and state tuberculosis and leprosy control programs and the United State Agency for International Development for their support. We extend our warm gratitude to all the study participants and the staff of KNCV Nigeria for their excellent work on this study.

Authors' Contributions

BO, CO, SU, OC, N Nkiru, and OO planned the EWORS intervention. CA, BO, SO, TO, DN, MG, OO, and OC-A conducted advocacy and engagement of strategic partners. OC, CO, US, N Nkiru, and BO supervised field engagement, screening, and testing. BO, SU, N Nkiru, OC, CO, OO, DN, and RE contributed to intervention design, algorithms, and field manuals. CO, SG, OC-A, OO, and N Nihalani designed the monitoring and evaluation system and field manuals. OO, CO, DJ, CM, DK, and FN harmonized data sets, DJ and CO conducted data analyses, and CD and OC wrote the draft manuscript. OO, DJ, CM, DK, FN, and GM contributed to data interpretation and writing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Transforming Our World: the 2030 Agenda for Sustainable Development. United Nations. 2015. URL: https://tinyurl.com/ju52b25t [accessed 2021-11-12]
2. Global tuberculosis report 2021. World Health Organization. URL: https://www.who.int/publications/i/item/9789240037021 [accessed 2022-03-07]
3. Nigeria Federal Ministry of Health. National Tuberculosis, Leprosy and Buruli Ulcer Management and Control Guidelines. Sixth Edition. Abuja, Nigeria: Nigeria Federal Ministry of Health; 2021.
4. Odume B, Falokun V, Chukwuogo O, Ogbudebe C, Useni S, Nwokoye N, et al. Impact of COVID-19 on TB active case finding in Nigeria. Public Health Action 2020 Dec 21;10(4):157-162 [FREE Full text] [doi: 10.5588/pha.20.0037] [Medline: 33437681]
5. Global tuberculosis report 2020. World Health Organization. URL: https://www.who.int/publications/i/item/9789240013131 [accessed 2022-01-21]
6. Odume B, Nwokoye N, Spruijt A. Diagnostic accuracy of TB-LAMP for diagnosis of pulmonary tuberculosis among adult presumptive TB in Nigeria. Greener J Med Sci 2021;11(2):122-129 [FREE Full text]
7. Tukur M, Odume B, Bajehson M, Dimpka C, Useni S, Ogbudebe C, et al. Outcome of tuberculosis case surveillance at Kano Central Correctional Center, North-west Nigeria: a need for routine active case findings for TB in Nigerian correctional centers. Int J Trop Dis Health 2021 Dec 13:37-45 [FREE Full text] [doi: 10.9734/IJTDH/2021/v42i1830536]
8. Electronic recording and reporting for tuberculosis care and control. World Health Organization. 2012. URL: https://apps.who.int/iris/handle/10665/44840 [accessed 2022-01-05]
9. Yang W. Introduction. In: Yang W, editor. Early Warning for Infectious Disease Outbreak: Theory and Practice. London, UK: Elsevier; 2017:E-22.

10. Wang R, Jiang Y, Michael E, Zhao G. How to select a proper early warning threshold to detect infectious disease outbreaks based on the China infectious disease automated alert and response system (CIDARS). BMC Public Health 2017 Jun 12;17(1):570 [FREE Full text] [doi: 10.1186/s12889-017-4488-0] [Medline: 28606078]
11. Siswoyo H, Permana M, Larasati RP, Farid J, Suryadi A, Sedyaningsih ER. EWORS: using a syndromic-based surveillance tool for disease outbreak detection in Indonesia. BMC Proc 2008 Nov 14;2 Suppl 3(Suppl 3):S3 [FREE Full text] [doi: 10.1186/1753-6561-2-s3-s3] [Medline: 19025680]
12. EWORS - Disease Outbreak Surveillance. InStrat Global Health Solutions. URL: https://instratghs.com/ewors-disease-outbreak-surveillance [accessed 2022-03-15]
13. Sánchez-Martín J, Rengifo-Gallego J, Blas-Morato R. Hot spot analysis versus cluster and outlier analysis: an enquiry into the grouping of rural accommodation in Extremadura (Spain). Int J Geoinf 2019 Apr 04;8(4):176 [FREE Full text] [doi: 10.3390/ijgi8040176]
14. Wang J, Zhang T, Fu B. A measure of spatial stratified heterogeneity. Ecol Indic 2016 Aug;67(1):250-256. [doi: 10.1016/j.ecolind.2016.02.052]
15. Zhang Y, Lu H, Qu W. Geographical detection of traffic accidents spatial stratified heterogeneity and influence factors. Int J Environ Res Public Health 2020 Jan 16;17(2):572 [FREE Full text] [doi: 10.3390/ijerph17020572] [Medline: 31963135]
16. Xie Z, Yan J. Kernel density estimation of traffic accidents in a network space. Comput Environ Urban Syst 2008 Sep;32(5):396-406. [doi: 10.1016/j.compenvurbsys.2008.05.001]
17. Anselin L. A Local Indicator of Multivariate Spatial Association: Extending Geary's c. Geogr Anal 2018 Aug 13;51(2):133-150. [doi: 10.1111/gean.12164]
18. National Tuberculosis and Leprosy Control Program: Annual report. Nigeria Federal Ministry of Health. 2019. URL: https://www.health.gov.ng/doc/Draft-2019-NTBLCP-Annual-report-22032020.pdf [accessed 2021-11-05]
19. Dim C, Dim N. Trends of tuberculosis prevalence and treatment outcome in an under-resourced setting: The case of Enugu state, South East Nigeria. Niger Med J 2013 Nov;54(6):392-397 [FREE Full text] [doi: 10.4103/0300-1652.126292] [Medline: 24665153]
20. Ogbudebe CL, Chukwu JN, Nwafor CC, Meka AO, Ekeke N, Madichie NO, et al. Reaching the underserved: Active tuberculosis case finding in urban slums in southeastern Nigeria. Int J Mycobacteriol 2015 Mar;4(1):18-24 [FREE Full text] [doi: 10.1016/j.ijmyco.2014.12.007] [Medline: 26655193]
21. Ogbudebe CL, Izuogu S, Abu CE. Magnitude and treatment outcomes of pulmonary tuberculosis patients in a poor urban slum of Abia State, Nigeria. Int J Mycobacteriol 2016 Jun;5(2):205-210 [FREE Full text] [doi: 10.1016/j.ijmyco.2016.03.003] [Medline: 27242233]
22. WHO operational handbook on tuberculosis. Module 2: screening - systematic screening for tuberculosis disease. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/340256/9789240022614-eng.pdf [accessed 2022-02-18]
23. Onyeonoro U, Chukwu J, Nwafor C, Meka A, Omotowo B, Ogbudebe C, et al. Assessment of the quality of tuberculosis surveillance data in six selected states in Southern Nigeria. Ann Trop Med Public Health 2015;8(4):94 [FREE Full text] [doi: 10.4103/1755-6783.162318]
24. Cheng W, Washington SP. Experimental evaluation of hotspot identification methods. Accid Anal Prev 2005 Sep;37(5):870-881. [doi: 10.1016/j.aap.2005.04.015] [Medline: 15963449]
25. Wang D, Ding W, Lo H, Stepinski T, Salazar J, Morabito M. Crime hotspot mapping using the crime related factors—a spatial data mining approach. Appl Intell 2012 Dec 8;39(4):772-781. [doi: 10.1007/s10489-012-0400-x]
26. Anselin L. Local Indicators of Spatial Association—LISA. Geogr Anal 1995;27(2):115. [doi: 10.1111/j.1538-4632.1995.tb00338.x]

Abbreviations

ACF: active case finding
CIDARS: China Infectious Disease Automated Alert and Response System
EWORS: early warning outbreak recognition system
EWS: early warning systems
LGA: local government area
LISA: Local Indicator of Spatial Autocorrelation
NNS: number needed to screen
NNT: number needed to test
TB: tuberculosis
W4SS: WHO 4-symptom screen
WHO: World Health Organization

Edited by Y Khader; submitted 15.06.22; peer-reviewed by YL Cheong, J Wang; comments to author 26.08.22; revised version received 07.10.22; accepted 16.12.22; published 08.02.23.

Please cite as:
Ogbudebe C, Jeong D, Odume B, Chukwuogo O, Dim C, Useni S, Okuzu O, Malolan C, Kim D, Nwariaku F, Nwokoye N, Gande S, Nongo D, Eneogu R, Odusote T, Oyelaran S, Chijioke-Akaniro O, Nihalani N, Anyaike C, Gidado M
Identifying Hot Spots of Tuberculosis in Nigeria Using an Early Warning Outbreak Recognition System: Retrospective Analysis of Implications for Active Case Finding Interventions
JMIR Public Health Surveill 2023;9:e40311
URL: https://publichealth.jmir.org/2023/1/e40311
doi:10.2196/40311
PMID:36753328

Original Paper

A Pay-It-Forward Approach to Improve Chlamydia and Gonorrhea Testing Uptake Among Female Sex Workers in China: Venue-Based Superiority Cluster Randomized Controlled Trial

Weiming Tang[1,2,3], PhD; Yewei Xie[3,4], MSc; Mingzhou Xiong[1,2], MD; Dan Wu[3,5,6], PhD; Jason J Ong[5,7], MD, PhD; Teodora Elvira Wi[8], MD; Bin Yang[1,2], PhD; Joseph D Tucker[3,5], PhD; Cheng Wang[1,2], PhD

[1] Dermatology Hospital of Southern Medical University, Guangzhou, China

[2] Southern Medical University Institute for Global Health, Guangzhou, China

[3] University of North Carolina at Chapel Hill Project-China, Guangzhou, China

[4] Health Service & System Research programme, Duke-NUS Medical School, Singapore, Singapore

[5] Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[6] West China School of Public Health, Sichuan University, Chengdu, China

[7] Central Clinical School, Faculty of Medicine, Monash University, Melbourne, Australia

[8] Department of Global HIV, Hepatitis and STI Programmes, World Health Organization Headquarters, Geneva, Switzerland

Corresponding Author:
Cheng Wang, PhD
Dermatology Hospital of Southern Medical University
Lujing Road, No 2
Yuexiu District
Guangzhou, 510000
China
Phone: 86 1 360 246 3723
Email: wangcheng090705@gmail.com

Abstract

Background: Regular chlamydia and gonorrhea testing are essential for key populations, such as female sex workers (FSWs). However, testing cost, stigma, and lack of access prevent FSWs in low- and middle-income countries from receiving chlamydia and gonorrhea testing. A social innovation to address these problems is “pay it forward,” where an individual receives a gift (free testing) and then asks whether they would like to give a gift to another person in the community.

Objective: This cluster randomized controlled trial examined the effectiveness and cost of the pay-it-forward strategy in increasing access to chlamydia and gonorrhea testing among FSWs in China.

Methods: This trial integrated a pay-it-forward approach into a community-based HIV outreach service. FSWs (aged 18 years or older) were invited by an outreach team from 4 Chinese cities (clusters) to receive free HIV testing. The 4 clusters were randomized into 2 study arms in a 1:1 ratio: a pay-it-forward arm (offered chlamydia and gonorrhea testing as a gift) and a standard-of-care arm (out-of-pocket cost for testing: US $11). The primary outcome was chlamydia and gonorrhea test uptake, as ascertained by administrative records. We conducted an economic evaluation using a microcosting approach from a health provider perspective, reporting our results in US dollars (at 2021 exchange rates).

Results: Overall, 480 FSWs were recruited from 4 cities (120 per city). Most FSWs were aged ≥30 years (313/480, 65.2%), were married (283/480, 59%), had an annual income <US $9000 (301/480, 62.7%), and had never been tested for chlamydia (401/480, 83.5%) or gonorrhea (397/480, 82.7%). Chlamydia and gonorrhea test uptake in the pay-it-forward and standard-of-care arms were 82% (197/240) and 4% (10/240), respectively, with an adjusted proportion difference of 76.7% (lower bound 95% CI 70.8%). All those who tested positive were referred to and received treatment at local sexually transmitted infection clinics. This finding was consistent when adjusting for marital status, income, inconsistent condom use during commercial sex in the last 3 months, and HIV testing history. Among 197 women who received tests in the pay-it-forward arm, 99 (50.3%) donated money, with a median donation of US $1.54 (IQR 0.77-1.54). The economic cost per person tested was US $568.71 for standard of care and US $43.20 for pay it forward.

Conclusions: The pay-it-forward strategy has the potential to enhance chlamydia and gonorrhea testing for Chinese FSWs and may be useful for scaling up preventive services. Further implementation research is needed to inform the transition of pay-it-forward research into practice.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000037653; https://www.chictr.org.cn/showprojen.aspx?proj=57233

(JMIR Public Health Surveill 2023;9:e43772) doi:10.2196/43772

KEYWORDS

pay-it-forward; chlamydia; gonorrhea; testing; female sex workers; women; China; cost; stigma; prevention; community; HIV; care; STD; implementation; research

Introduction

Chlamydia and gonorrhea are two of the most common bacterial sexually transmitted infections (STIs). Each year, there are an estimated 129 million new cases of chlamydia and 82 million new cases of gonorrhea globally, and the majority are asymptomatic [1]. Female sex workers (FSWs) face a high burden of chlamydia and gonorrhea, with an estimated prevalence of 4% to 15% and 1% to 11%, respectively, while FSWs in low- and middle-income countries (LMICs) have a higher prevalence of chlamydia and gonorrhea [2-6].

Testing, especially for high-risk women, is essential for chlamydia and gonorrhea control and the reduction of adverse sexual and reproductive complications [1,7]. The World Health Organization guidelines recommend regular chlamydia and gonorrhea molecular assay testing for sex workers in LMICs [8]. However, despite a high disease burden, few FSWs receive chlamydia and gonorrhea testing in LMICs. For example, less than one-tenth of Chinese FSWs have ever received chlamydia and gonorrhea testing [9]. First, although chlamydia and gonorrhea testing are available in many Chinese hospitals, the out-of-pocket cost for testing is relatively high (US $11) and not covered by most public or private health insurance programs [10]. Given that more than half of Chinese FSWs have an average monthly income of less than US $1000 [11], financial limitations may remain a significant barrier preventing FSWs from accessing chlamydia and gonorrhea testing. Second, current STI prevention services in many LMICs, including China, mainly focus on HIV and syphilis [1]. These well-established HIV and syphilis testing systems do not typically integrate other STI testing [12]. China does not have a screening strategy for chlamydia or gonorrhea for high-risk and other vulnerable populations, nor does it have widespread programs supporting chlamydia and gonorrhea testing. Third, peer factors relating to mistrust in services may decrease interest in routine service delivery [13]. This mistrust in STD services might dramatically reduce access to STD care services, including chlamydia and gonorrhea testing in clinical settings [14]. A lack of community engagement might be an essential reason for mistrust [15].

A pay-it-forward health approach may help promote test uptake among FSWs. In a pay-it-forward approach, one person receives a gift (eg, a free STI test, alongside community-engaged messages) and then is asked whether they would like to donate money to another person [16]. This social innovation has the potential of not only reducing financial barriers but also increasing trust and community engagement in health services by increasing community ownership. Two studies among Chinese men who have sex with men (MSM) found that the pay-it-forward approach could substantially increase chlamydia and gonorrhea testing uptake [10,17]. However, these studies only focused on MSM in 2 cities, limiting the generalizability of the findings.

This multisite cluster randomized controlled trial (RCT) aimed to evaluate the effectiveness and cost of a pay-it-forward strategy in increasing dual chlamydia and gonorrhea testing among FSWs in 4 cities of Guangdong Province, China, compared with a standard fee-based system.

Methods

Study Design

This study was a superiority cluster RCT to test whether the pay-it-forward intervention was superior to the standard of care in increasing chlamydia and gonorrhea testing. The study was conducted in 4 cities (Maoming, Yunfu, Yangjiang, and Yingde), with 1 cluster per city, in Guangdong Province, China. These cities were chosen based on local capacity and the availability of ongoing public health community-based outreach programs for FSWs. The reason for choosing a cluster RCT design was that many FSWs in the same city would be recruited from the same venues, and an individual RCT design may have caused contamination between different arms.

The local Centre for Disease Control and Prevention outreach team included a nurse, an STI physician from a local STI clinic, and a public health practitioner (as a health educator). The outreach team routinely provided venue-based sexual health care services to local FSWs. These outreach teams have long-standing relations with FSWs and provide comprehensive health services, including condom promotion, reproductive health services, STI counseling, symptomatic STI treatment, and on-site HIV testing.

We estimated we needed 480 participants from the 4 clusters (120 from each city). Participants were recruited after randomization of the selected cluster. Our previous studies among MSM indicated that the pay-it-forward approach is superior to the standard of care by a 20% margin. We kept the same margin for this superiority trial. The sample size was determined using the following assumptions: the testing rate of the null hypothesis for the pay-it-forward group was set at 30% (the null hypothesis was that the pay-it-forward group would increase the testing rate by at least 10%), the testing rate of the alternative hypothesis was set at 50%, and the testing rate of the null hypothesis for the standard of care group was set at 20%, with an interclass correlation of 0.1.

Ethical Considerations

Study approval was gained from the ethics review committees of the Southern Medical University Dermatology Hospital (#2020018). The CONSORT (Consolidated Standards Of Reporting Trials) cluster-extension guidelines were used to report the findings from this study (Multimedia Appendix 1). This study was registered at the Chinese Clinical Trial Registry (ChiCTR2000037653). All the participants who participated in the survey signed an informed consent form. The informed consent form included a statement that study data were anonymous or would be deidentified.

Participants

Many sex workers in China do not routinely seek clinic-based services, so we integrated this study within an ongoing HIV- and syphilis-testing outreach program. The study participants were recruited by having the outreach team, including a study staff member, go to their workplaces during working hours (usually in the evening). Before this study, the local outreach team mapped sex-work venues in each study site according to geographic area and type of venue. Convenience sampling was used to enroll FSWs at venues in each city. We categorized the sex work venues into high and low tiers based on the clients' socioeconomic status and the cost for clients of commercial sex at these venues [18]. Low-tier venues included foot-bath shops, hair salons, massage parlors, roadside restaurants, roadside shops, guesthouses, streets, and public outdoor places. High-tier venues included karaoke bars, hotels, saunas, and nightclubs.

The inclusion criteria were as follows: sex at birth was female, age was 18 years or older, the participant self-reported as having engaged in transactional sex to obtain money or property in the past month in the sampling city, the participant had not previously participated in a research project related to gonorrhea or chlamydia testing, and the participant had not been tested for gonorrhea or chlamydia in the past 12 months.

The study was anonymous, but the eligible participants had to provide a cell-phone number for testing results notification. The phone number was also used for deduplication. If the participants had positive test results for as either chlamydia or gonorrhea, they were provided free treatment at a local STI clinic, regardless of whether they had symptoms.

Randomization and Masking

Before participant recruitment, the 4 clusters were randomized into the 2 study arms in a 1:1 ratio. Participants recruited from Yunfu and Yingde cities were randomized into the pay-it-forward intervention group, and participants recruited from Maoming and Yangjiang cities were randomized into the control group. The reason for adopting a cluster randomized allocation method was the strong correlation between FSW test uptake in local areas of a city [19]. In addition, many of the FSWs were recruited from the same study sites. Individual-level randomization would have induced strong contamination between the study groups. We did not randomize the participants based on venue, as many street-based FSWs work alone.

Procedures

The intervention was adapted from our previous trials, and a pilot intervention was implemented among 40 FSWs in Guangzhou, China, before the study was carried out.

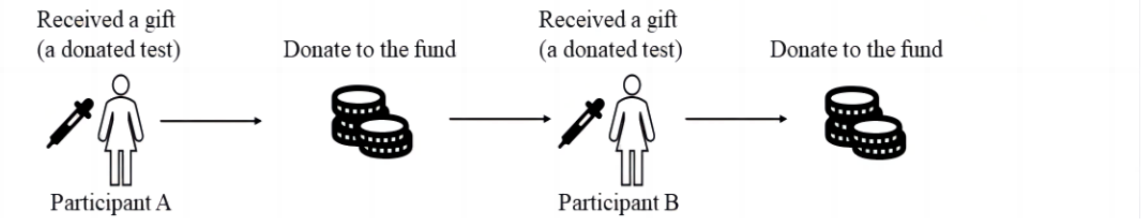
The pay-it-forward strategy was implemented in the intervention group. The study staff visited FSW workplaces in the selected cities in the pay-it-forward arm. First, after providing HIV and syphilis testing services, the study staff briefly introduced the reasons for chlamydia and gonorrhea testing, explained the pay-it-forward strategy, informed the participants of the market price of chlamydia and gonorrhea testing (approximately US $11 for both the intervention and control groups), and communicated that the testing cost had been covered by donations from other FSWs who care about their health. Therefore, each FSW could decide whether to receive donated chlamydia and gonorrhea testing services. If they chose to receive the test, they self-collected urine samples in a private room at their workplace. After that, they were asked whether they would voluntarily donate any amount to other FSWs for testing. A combination of study funds and donations from previous participants covered all testing costs for the intervention group. The donations were voluntary and were given after the testing. The key concepts of the pay-it-forward model are depicted in Figure 1.

The researchers also introduced chlamydia and gonorrhea testing to the standard-of-care arm without mentioning the pay-it-forward intervention. Participants who wished to be tested for chlamydia and gonorrhea in the standard-of-care arm self-collected their samples for testing in outreach settings. They were only required to pay a reduced price for their testing (approximately US $11). The participants in both arms were offered an incentive of US $7.50 for their participation and time spent in the research.

Participants in each group were invited to complete a short baseline questionnaire to collect information on their sexual history, testing history, attitudes toward the test, and the psychosocial environment in which they lived. We asked participants whether they used a condom during their last commercial sex act and whether they consistently used a condom during commercial sex in the previous months (consistent use meant using a condom every time).

All samples were transported to the laboratory at the Dermatology Hospital of Southern Medical University in Guangzhou for nucleic acid amplification testing (Cobas 4800 CT/NG Test Kits; Roche Molecular Systems).

Figure 1. The pay-it-forward approach.



Outcomes

The primary outcome was the uptake of dual chlamydia and gonorrhea testing immediately after the intervention, measured during the study visit. The secondary outcome was the cost per person tested and the incremental cost-effectiveness ratio (ICER) between the 2 study arms. We calculated the financial and economic costs from a health care provider perspective, including direct medical expenses related to testing and personnel time. Costs are presented in US dollars (at 2021 exchange rates). We calculated the cost per person tested by dividing each arm's total cost by the number of FSWs tested. The ICER was also calculated as the incremental cost of the intervention arm compared to the control arm divided by the incremental effectiveness (per person tested).

Statistical Analysis

All statistical analyses were conducted at the cluster level. We first compared the sociodemographic characteristics of the participants in the 2 study arms. The generalized estimating equations (GEEs) with a binomial distribution and an identity link function were used to compare the proportion of chlamydia and gonorrhea tests in the different arms, reported as the crude probability difference. The equal correlation structure was specified as the correlation structure within groupings. Additionally, due to the small number of clusters in this study, the Mancl and DeRouen bias-corrected covariance estimator was applied to our generalized estimating equations to minimize the risk of a type I error [20,21]. A sensitivity analysis was also conducted, exploring the effects when considering FSW workplaces as clusters in the GEE model corrected by the Kauermann-Carroll method (Multimedia Appendix 2, Table S1) [21,22]. In the adjusted model, we further adjusted for the following variables: marital status and past experience of testing for HIV, chlamydia, or gonorrhea. These covariates were considered potential confounders, and some were not balanced between the 2 study groups because of the small cluster size. As this study used a superiority trial design, only one side of the CI is reported [17].

A subgroup analysis was performed to investigate the probability difference categorized by age, consistent condom use during commercial sex in the past 3 months, history of ever having been tested for HIV, and whether the participant worked in a high-tier workplace. These analyses were conducted with Stata (version 16; Stata Corp). The economic evaluation was performed using TreeAge Pro 2021 (TreeAge Software Inc).

Results

Study Participants

This study was implemented from August 12, 2020, to November 15, 2020. We screened 530 female sex workers and recruited 480 who met the screening criteria. Fifty women were excluded from the study (2 were younger than 18 years, and 48 self-reported not engaging in commercial sex; Figure 2); 240 participants from 2 clusters were included in each study group. All enrolled women had data on the primary outcome.

Table 1 shows the participants' sociodemographic characteristics, sexual behavior, and past STI testing history. Most women were aged 30 years or older (313/480, 65.2%), were married (283/480, 59%), and had a middle-school or lower education level (395/480, 82.3%). The majority worked at high-tier venues (311/480, 64.8%) and had an annual income of less than US $9000 (301/480, 63.6%). Nearly half of the women (235/480, 49%) had worked for longer than one year at their workplace. Around one-third of the participants (179/480, 37.3%) had used a condom consistently during commercial sex in the past 3 months. Many (290/480, 60.4%) women had never been tested for HIV, and the majority had never been tested for gonorrhea (397/480, 82.7%) or chlamydia (401/480, 83.5%).

Figure 2. Flow diagram of participant inclusion.

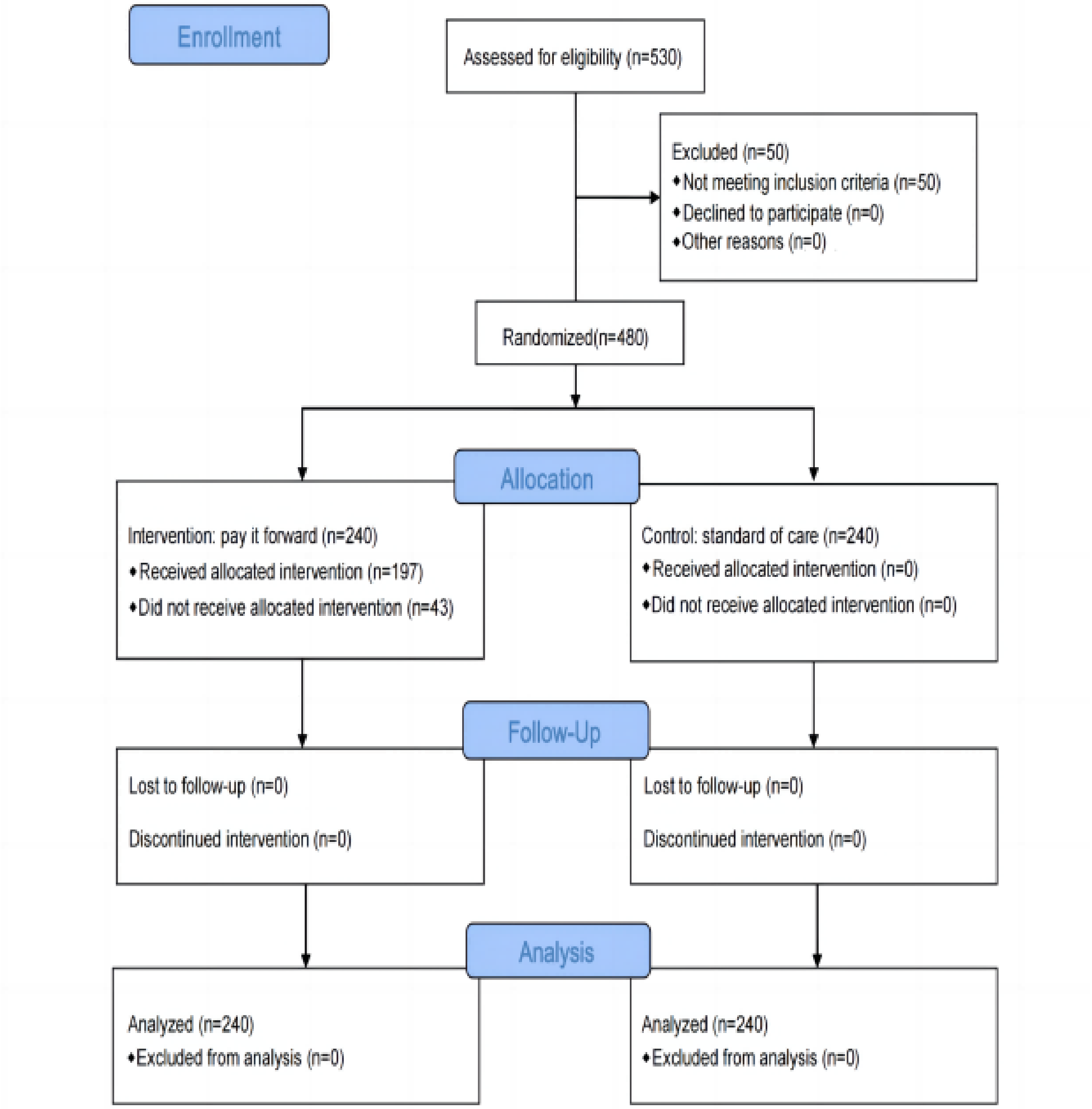


Table 1. Participant characteristics by testing scenario among female sex workers in Guangdong, China, in 2020.

Characteristics	Total (N=480)	Pay it forward (n=240)	Standard of care (n=240)
Age (years), n (%)			
<30	167 (34.8)	86 (35.8)	81 (33.8)
>30	313 (65.2)	154 (64.2)	159 (66.3)
Age (years), mean (SD)	34.7 (9.0)	34.6 (9.7)	34.75 (8.2)
Marital status, n (%)			
Never married	114 (23.8)	62 (25.8)	52 (21.7)
Married	283 (59)	119 (49.6)	164 (68.3)
Divorced or widowed	83 (17.3)	59 (24.6)	24 (10)
Highest education, n (%)			
Middle school or below	395 (82.3)	191 (79.6)	204 (85)
High or vocational school	76 (15.8)	45 (18.8)	31 (12.9)
College or above	9 (1.9)	4 (1.7)	5 (2.1)
Annual income (US $), n (%)			
<1800	2 (0.4)	1 (0.4)	1 (0.4)
1800-6000	299 (62.3)	158 (65.8)	141 (58.7)
>9000	179 (37.3)	81 (33.8)	98 (40.9)
Type of workplace, n (%)			
High-tier venues[a]	311 (64.8)	156 (64.2)	157 (65.4)
Low-tier venues[b]	169 (35.2)	86 (35.8)	83 (34.6)
Time working at current workplace (months), n (%)			
<1	32 (6.7)	22 (9.2)	10 (4.2)
1-6	121 (25.2)	75 (31.3)	46 (19.2)
7-12	92 (19.2)	50 (20.8)	42 (17.5)
>12	235 (49)	93 (38.8)	142 (59.2)
Used condom during last commercial sex act, n (%)			
Yes	358 (74.6)	148 (61.7)	210 (87.5)
No	122 (25.4)	92 (38.3)	30 (12.5)
Frequency of condom use during commercial sex in the past 3 months, n (%)			
Nonuse	67 (14)	57 (23.8)	10 (4.2)
Sometimes (less than half the time)	120 (25)	48 (20)	72 (30)
Often (more than half the time)	114 (23.8)	29 (12.1)	85 (35.4)
Consistent use	179 (37.3)	106 (44.2)	73 (30.4)
HIV testing frequency in the past 2 years, n (%)			
Never tested	290 (60.4)	195 (81.3)	95 (39.6)
<Once every 2 years	70 (14.6)	19 (7.9)	51 (21.3)
Once a year or more frequently	120 (25)	26 (10.9)	94 (39.2)
Gonorrhea testing frequency in the past 2 years, n (%)			
Never tested	397 (82.7)	219 (91.3)	178 (74.2)
<Once every 2 years	53 (11)	8 (3.3)	45 (18.8)
Once a year or more frequently	30 (6.2)	13 (5.4)	17 (7.1)
Chlamydia testing frequency in the past 2 years, n (%)			

Characteristics	Total (N=480)	Pay it forward (n=240)	Standard of care (n=240)
Never tested	401 (83.5)	226 (94.2)	175 (72.9)
<Once every 2 years	52 (10.8)	4 (1.7)	48 (20)
Once a year or more frequently	27 (5.6)	10 (4.2)	17 (7.1)
Size of donation (US $), median (IQR)	N/A[c]	10 (5-10)	N/A
Gonorrhea test result[d], n (%)			
Positive	3 (1.5)	3 (1.5)	0 (0)
Negative	200 (98.5)	194 (98.5)	10 (100)
Chlamydia test result[d], n (%)			
Positive	34 (16.7)	34 (17.3)	0 (0)
Negative	169 (83.3)	163 (82.7)	10 (100)

[a] High-tier venues included karaoke bars, hotels, saunas, and night clubs.

[b] Low-tier venues included foot-bath shops, hair salons or barbershops, massage parlors, roadside restaurants, guesthouses, streets, and public outdoor places.

[c] N/A: not applicable.

[d] For these values, n=203, n=197, and n=10 for the total, pay-it-forward, and standard-of-care groups, respectively.

Testing Uptake

Overall, 197 of 240 (82.1%) women in the pay-it-forward arm and 10 of 240 (4.2%) women in the standard-of-care arm received dual chlamydia and gonorrhea testing (Table 2). In the pay-it-forward arm, 34 of 197 (17.3%) women were diagnosed with chlamydia, and 3 of 197 (1.5%) were diagnosed with gonorrhea. No chlamydia or gonorrhea was detected among women who received a standard-of-care test. All women who received a test were informed of the test results via text message (for women who tested negative) or phone call (for women who tested positive for a chlamydia or gonorrhea infection). Test-positive women were referred to local STI clinics for free treatment. The reasons for testing or not testing in the pay-it-forward arm and the program's perceived benefits are summarized in Multimedia Appendix 3, Table S4.

The GEE model showed that the pay-it-forward intervention was associated with a 77.9% (lower-bound 95% CI 70.3%) increase in test uptake probability compared to the standard-of-care arm. After adjusting for marital status and past testing experience for HIV, chlamydia, and gonorrhea, the outputs were consistent with the crude results (adjusted probability difference 76.7%, lower-bound 95% CI 70.8%; Table 2).

Table 3 shows findings from prespecified subgroup analyses. The difference between the control and pay-it-forward arms was most significant among FSWs who had used condoms consistently during commercial sex in the previous 3 months (86%, 95% CI 78.6%-93.5%) and women older than 30 years (82%, 95% CI 75.7%-88.2%). The difference between the control and pay-it-forward arms was 4% (95% CI –8.6% to 16.6%) for a past history of HIV testing (no vs yes) and –5.1% (95% CI –16.3% to 6%) for workplace tier.

Table 2. The proportion of participating female sex workers tested for chlamydia and gonorrhea in Guangdong, China, in 2020 (N=480). The model was adjusted for marital status and past testing experience for HIV, chlamydia, and gonorrhea.

	Subjects, n (%)	Probability difference[a]	One-sided 95% CI[b]	Intraclass correlation	Adjusted probability difference[a]	One-sided 95% CI[b]
Pay-it-forward group (n=240)	197 (82.1)	77.9%	70.3%	<0.001	76.7%	70.8%
Standard-of-care group (n=240)	10 (4.2)	N/A[c]	N/A	<0.001	N/A	N/A

[a] Indicates probability of difference between the intervention arm (pay-it-forward) and the standard-of-care arm.

[b] The lower-bound 95% CI is reported.

[c] N/A: not applicable.

Table 3. Subgroup analysis of 480 participants in China in 2020.

	SOC[a] (n=240), n/N (%)	PIF[b] (n=240), n/N (%)	Probability difference (PIF minus SOC), % (95% CI)	Difference in probability difference, % (95% CI)	*P* value
Age (years)				11.2 (–0.6 to 23.1)[c]	.07
≤30	3/81 (4)	64/86 (74)	70.7 (60.6-80.8)		
>30	7/159 (4)	133/154 (86)	82 (75.7-88.2)		
Condom use during commercial sex in the past 3 months				–15 (–25.8 to –4.1)[d]	.007
Inconsistent	6/167 (4)	100/134 (75)	71 (63.1-78.9)		
Consistent	4/73 (5)	97/106 (92)	86 (78.6-93.5)		
Tested for HIV in the past				4 (–8.6 to 16.6)[e]	.54
No	0/95 (0)	159/195 (82)	81.5 (76.1-87)		
Yes	10/145 (7)	38/45 (84)	77.5 (66.2-88.9)		
Workplace				–5.1 (–16.3 to 6)[f]	.37
High tier	5/157 (3)	122/154 (79)	76 (69.1-83)		
Low tier	5/83 (6)	75/86 (87)	81 (72.5-89.9)		

[a] SOC: standard of care.

[b] PIF: pay it forward.

[c] ≤30 years vs >30 years.

[d] Inconsistent vs consistent.

[e] No vs yes.

[f] High tier vs low tier.

Economic Evaluation

Among 197 women who received chlamydia and gonorrhea tests in the pay-it-forward arm, 99 (50.3%) donated to the pooled funds. The total donation amount was US $326, and the median donation amount per donor was US $1.54 (IQR 0.77-1.54). The largest donation was US $119.54, and the lowest was US $0.15. Complete cost and cost-effectiveness analyses are provided in Multimedia Appendix 4, Tables S2 to S3 and Figures S1 to S3. The economic cost per person tested was US $42.24 for the standard-of-care arm and US $41.80 for the pay-it-forward arm. The ICER for the pay-it-forward arm compared to standard of care was minus US $2.79 per additional person tested.

Discussion

Principal Findings

Upstream reciprocity theory suggests that people who are cared for by someone are more likely to help others. Our study confirmed this theory by promoting chlamydia and gonorrhea testing among FSWs in China through a pay-it-forward approach. Our study extends the existing literature by testing the effectiveness of the pay-it-forward approach in increasing STI testing uptake among FSWs, conducting the study in an outreach setting outside of a formal clinic, and integrating the approach with HIV services. As a social innovation approach, we found that the pay-it-forward intervention dramatically increased STI testing uptake among Chinese FSWs and that the participants covered a substantial portion of the costs associated with testing.

Women in the pay-it-forward arm had higher chlamydia and gonorrhea testing uptake than women in the standard-of-care arm. This finding is consistent with previous studies among Chinese MSM, which suggest that the pay-it-forward approach can motivate more people to use STI testing services [10,17]. In addition, the pay-it-forward strategy is more effective than other STI testing–uptake intervention approaches, such as social marketing [23], digital health [24], and point of care [25]. Several factors may have led to the strong testing uptake among individuals in the pay-it-forward arm. The first was generosity and trust. Due to stigma and other barriers, few Chinese FSWs are willing to undergo or pay for chlamydia and gonorrhea testing [26]. The low testing rates in our standard-of-care arm demonstrate this. Our pay-it-forward strategy revealed substantial generosity and promoted responsibility among the FSWs to change their testing behaviors. Second, free testing may have driven the increased test uptake rates because of the zero-price effect [27]. Our previous studies indicate that generosity (ie, voluntary donations) integrated with free testing strongly motivates behavior change [10,17].

Lack of community engagement and low social trust are 2 critical barriers to STI testing uptake among FSWs [28]. In our study, about half of the FSWs who were tested in the pay-it-forward arm donated some money for future testing. Although the donation rate in our study was lower than in our previous studies among MSM in China [10,17], around half of the FSWs chose to contribute. This is remarkable given the low socioeconomic status of FSWs and the lack of a culture of donating in China. Pay-it-forward approaches may increase community ownership and social trust, supporting STI testing [16,29]. Two systematic reviews further support this explanation, reporting that being kind is associated with well-being, especially psychological functioning [30] and subjective well-being [31]. Qualitative studies that aim to further confirm these psychological impulses are needed.

Our findings indicate that the pay-it-forward strategy increased STI testing uptake and reduced barriers to sustaining STI testing. STI testing among FSWs reduces STI transmission and adverse sexual and reproductive health consequences. The pay-it-forward approach provides an alternative to using limited public funding to maintain or expand testing. Our findings need to be extended beyond this initial research study, as they have several public health implications. First, a pay-it-forward strategy could reduce costs for STI testing for funding bodies and dramatically remove the financial barrier to STI testing among participants. Second, unlike the free-testing approach, the pay-it-forward method could generate strong community ownership and trust. Third, as an approach to increase testing uptake and the regularity of testing, the pay-it-forward intervention has a robust public health benefit. Improving testing coverage and the regularity of testing could reduce the high chlamydia and gonorrhea prevalence among Chinese FSWs (17.3% for chlamydia in our study) by identifying and treating asymptomatic infections earlier, thus reducing morbidity and onward transmission to the clients of FSWs and their other sexual partners [32].

Limitations and Implications

Our study has several limitations. First, although this was a cluster RCT, the number of clusters in our study was relatively small. Therefore, the characteristics of the participants in the 2 study arms were not completely balanced, which may have resulted in a selection bias. However, our main findings were robust when we adjusted for marital status, condom use during the last commercial sex act, and past HIV testing experience. In addition, the data were analyzed with GEE models for a small number of clusters. At the same time, a sensitivity analysis was also conducted by treating FSWs recruited from the same venue as a cluster. Second, the test uptake rate in the standard-of-care arm was low, further complicating data analysis. However, FSW test uptake rates are generally low in China [11] and many other LMICs [33]. Third, the level of community engagement in our study was suboptimal. Unlike our previous studies evaluating the pay-it-forward strategy in improving STI care, the FSWs in this study did not engage in community activities, such as passing postcards or messages to other FSWs in their social network; this needs to be improved in future studies.

Our study has several implications. First, qualitative research is needed to evaluate how and why generosity works among FSWs in promoting STI testing. Second, this study did not collect data on implementing the evidence-based pay-it-forward intervention and scaling it up. Additional implementation-science research is needed to expand the pay-it-forward intervention into other settings, including integrating it with other STI services. Third, strategies to increase demand and the informational messages on the need for chlamydia and gonorrhea screening (ie, the introduction to chlamydia and gonorrhea and the pay-it-forward intervention model) were not developed together with the FSWs; the study was carried out solely by study staff from public health authorities. Future research could examine how the engagement of local FSWs at the formative research stage could further improve testing uptake and program sustainability. Fourth, this study was implemented during the period from August to November 2020, a time when the COVID-19 pandemic may have impacted the implementation of the study itself and reduced the responsiveness of the FSWs, as many FSWs had temporarily ceased working during the pandemic period. We anticipated that more low-income FSWs would continue their work even while high-tier venues were shut down due to COVID-19 measures.

Conclusion

Innovative and effective intervention strategies are urgently needed to fill the gap between the high global burden of chlamydia and gonorrhea disease and suboptimal STI testing services. A pay-it-forward approach has the potential to enhance chlamydia and gonorrhea testing uptake among Chinese FSWs. It may be a valuable tool for scaling up STI services. Further implementation research is needed to inform the transition of pay-it-forward research into practice.

Acknowledgments

We are grateful to all participants in this study, without whom this study could not have been completed. We thank the study implementation team at the Dermatology Hospital Guangzhou site for all their dedication and efforts, which made this research possible. We appreciate the collaborating chronic-disease control centers in Maoming, Yunfu, Yangjiang, and Yingde for their help in facilitating and coordinating this study. Funding for this study came from the National Natural Science Foundation of China (grant 81903371) and the National Institutes of Health (grants NIAID K24AI143471 and R34MH119963).

Data Availability

The deidentified data have been uploaded as supplemental data.

Authors' Contributions

WT and CW conceived the study. YX and MX led the piloting and implementation of the study. MX, CW, and BY assisted with recruitment. YX cleaned and analyzed the data and generated figures and tables to present the paper’s findings. JJO advised on modeling approaches. WT, YX, and CW wrote the initial draft of the paper. DW, JJO, TEW, and JDT provided constructive comments and edited the manuscript. All authors read and authorized the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT 2010 checklist.
[DOCX File , 24 KB - publichealth_v9i1e43772_app1.docx]

Multimedia Appendix 2
Sensitivity analysis of number of cluster size.
[DOCX File , 13 KB - publichealth_v9i1e43772_app2.docx]

Multimedia Appendix 3
Cost effectiveness analysis.
[DOCX File , 13 KB - publichealth_v9i1e43772_app3.docx]

Multimedia Appendix 4
The reason for a test or not in the PIF program and the program's perceived benefits.
[DOCX File , 305 KB - publichealth_v9i1e43772_app4.docx]

References

1. Global progress report on HIV, viral hepatitis and sexually transmitted infections, 2021: accountability for the global health sector strategies 2016–2021: actions for impact. World Health Organization. 2016. URL: https://apps.who.int/iris/handle/10665/341412 [accessed 2021-10-13]
2. Decker MR, Wirtz AL, Baral SD, Peryshkina A, Mogilnyi V, Weber RA, et al. Injection drug use, sexual risk, violence and STI/HIV among Moscow female sex workers. Sex Transm Infect 2012 Jun;88(4):278-283 [FREE Full text] [doi: 10.1136/sextrans-2011-050171] [Medline: 22287530]
3. Kaul R, Kimani J, Nagelkerke NJ, Fonck K, Ngugi EN, Keli F, Kibera HIV Study Group. Monthly antibiotic chemoprophylaxis and incidence of sexually transmitted infections and HIV-1 infection in Kenyan sex workers: a randomized controlled trial. JAMA 2004 Jun 02;291(21):2555-2562. [doi: 10.1001/jama.291.21.2555] [Medline: 15173146]
4. Beattie TS, Mohan HL, Bhattacharjee P, Chandrashekar S, Isac S, Wheeler T, et al. Community mobilization and empowerment of female sex workers in Karnataka State, South India: associations with HIV and sexually transmitted infection risk. Am J Public Health 2014 Aug;104(8):1516-1525. [doi: 10.2105/ajph.2014.301911]
5. Znazen A, Frikha-Gargouri O, Berrajah L, Bellalouna S, Hakim H, Gueddana N, et al. Sexually transmitted infections among female sex workers in Tunisia: high prevalence of Chlamydia trachomatis. Sex Transm Infect 2010 Dec;86(7):500-505. [doi: 10.1136/sti.2010.042770] [Medline: 20656718]
6. Platt L, Grenfell P, Bonell C, Creighton S, Wellings K, Parry J, et al. Risk of sexually transmitted infections and violence among indoor-working female sex workers in London: the effect of migration from Eastern Europe. Sex Transm Infect 2011 Aug;87(5):377-384. [doi: 10.1136/sti.2011.049544] [Medline: 21572111]
7. Wu X, Zhou X, Chen Y, Zhai K, Sun R, Luo G, et al. The impact of COVID-19 lockdown on cases and deaths of AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China: an interrupted time series analysis. JMIR Public Health Surveill 2023 Jan 12:2023 [FREE Full text] [doi: 10.2196/40591] [Medline: 36634257]
8. Prevention and Treatment of HIV and Other Sexually Transmitted Infections for Sex Workers in Low- and Middle-Income Countries: Recommendations for a Public Health Approach. World Health Organization. 2012. URL: https://www.who.int/publications/i/item/9789241504744 [accessed 2012-12-01]
9. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. Bull World Health Organ 2019 Aug 01;97(8):548-562 [FREE Full text] [doi: 10.2471/BLT.18.228486] [Medline: 31384073]
10. Li KT, Tang W, Wu D, Huang W, Wu F, Lee A, et al. Pay-it-forward strategy to enhance uptake of dual gonorrhea and chlamydia testing among men who have sex with men in China: a pragmatic, quasi-experimental study. Lancet Infect Dis 2019 Jan;19(1):76-82 [FREE Full text] [doi: 10.1016/S1473-3099(18)30556-5] [Medline: 30587296]
11. Zhao PZ, Wang YJ, Cheng HH, Zhang Y, Tang WM, Yang F, et al. Uptake and correlates of chlamydia and gonorrhea testing among female sex workers in Southern China: a cross-sectional study. BMC Public Health 2021 Jul 28;21(1):1477 [FREE Full text] [doi: 10.1186/s12889-021-11526-w] [Medline: 34320978]
12. Wu D, Huang W, Zhao P, Li C, Cao B, Wang Y, et al. A crowdsourced physician finder prototype platform for men who have sex with men in China: qualitative study of acceptability and feasibility. JMIR Public Health Surveill 2019 Oct 08;5(4):e13027 [FREE Full text] [doi: 10.2196/13027] [Medline: 31596245]

13. Zhu X, Zhang W, Operario D, Zhao Y, Shi A, Zhang Z, et al. Effects of a mobile health intervention to promote HIV self-testing with MSM in China: a randomized controlled trial. AIDS Behav 2019 Nov;23(11):3129-3139 [FREE Full text] [doi: 10.1007/s10461-019-02452-5] [Medline: 30852728]
14. Cai G, Liu Y, Zhuang J, Lu Y, Wu J, Hu Z, et al. Differences in Socio-demographics Status, Risk Behaviours, Healthcare Uptake and HIV/STIs Between Brothel-based and Street-based Female Sex Workers in Yunnan, China. Research Square. Preprint posted online May 25, 2021. [doi: 10.21203/rs.3.rs-529460/v1]
15. Purnat TD, Vacca P, Czerniak C, Ball S, Burzo S, Zecchin T, et al. Infodemic signal detection during the COVID-19 pandemic: development of a methodology for identifying potential information voids in online conversations. JMIR Infodemiology 2021;1(1):e30971 [FREE Full text] [doi: 10.2196/30971] [Medline: 34447926]
16. Tang W, Wu D, Yang F, Wang C, Gong W, Gray K, et al. How kindness can be contagious in healthcare. Nat Med 2021 Jul;27(7):1142-1144. [doi: 10.1038/s41591-021-01401-x] [Medline: 34127853]
17. Yang F, Zhang TP, Tang W, Ong JJ, Alexander M, Forastiere L, et al. Pay-it-forward gonorrhoea and chlamydia testing among men who have sex with men in China: a randomised controlled trial. Lancet Infect Dis 2020 Aug;20(8):976-982 [FREE Full text] [doi: 10.1016/S1473-3099(20)30172-9] [Medline: 32530426]
18. Chen X, Liang G, Wang Q, Yin Y, Jiang N, Zhou Y, et al. HIV prevalence varies between female sex workers from different types of venues in southern China. Sex Transm Dis 2012 Nov;39(11):868-870. [doi: 10.1097/OLQ.0b013e318264c3ba] [Medline: 23064536]
19. Tucker JD, Peng H, Wang K, Chang H, Zhang S, Yang L, et al. Female sex worker social networks and STI/HIV prevention in South China. PLoS One 2011;6(9):e24816 [FREE Full text] [doi: 10.1371/journal.pone.0024816] [Medline: 21931856]
20. Mancl LA, DeRouen TA. A covariance estimator for GEE with improved small-sample properties. Biometrics 2001 Mar;57(1):126-134. [doi: 10.1111/j.0006-341x.2001.00126.x] [Medline: 11252587]
21. Gallis JA, Li F, Turner EL. xtgeebcv: A command for bias-corrected sandwich variance estimation for GEE analyses of cluster randomized trials. Stata J 2020 Jun;20(2):363-381 [FREE Full text] [doi: 10.1177/1536867x20931001] [Medline: 35330784]
22. Kauermann G, Carroll RJ. A Note on the efficiency of sandwich covariance matrix estimation. J Am Stat Assoc 2001 Dec;96(456):1387-1396. [doi: 10.1198/016214501753382309]
23. Wei C, Herrick A, Raymond H, Anglemyer A, Gerbase A, Noar SM. Social marketing interventions to increase HIV/STI testing uptake among men who have sex with men and male-to-female transgender women. Cochrane Database Syst Rev 2011 Sep 07(9):CD009337. [doi: 10.1002/14651858.CD009337] [Medline: 21901734]
24. Cao B, Bao H, Oppong E, Feng S, Smith KM, Tucker JD, et al. Digital health for sexually transmitted infection and HIV services: a global scoping review. Curr Opin Infect Dis 2020 Feb;33(1):44-50 [FREE Full text] [doi: 10.1097/QCO.0000000000000619] [Medline: 31789695]
25. Ong JJ, Fu H, Smith MK, Tucker JD. Expanding syphilis testing: a scoping review of syphilis testing interventions among key populations. Expert Rev Anti Infect Ther 2018 May;16(5):423-432 [FREE Full text] [doi: 10.1080/14787210.2018.1463846] [Medline: 29633888]
26. Zhang J, Li C, Xu J, Hu Z, Rutstein SE, Tucker JD, et al. Discontinuation, suboptimal adherence, and reinitiation of oral HIV pre-exposure prophylaxis: a global systematic review and meta-analysis. Lancet HIV 2022 Apr;9(4):e254-e268. [doi: 10.1016/S2352-3018(22)00030-3] [Medline: 35364026]
27. Shampanier K, Mazar N, Ariely D. Zero as a Special Price: The True Value of Free Products. Mark Sci 2007 Nov;26(6):742-757. [doi: 10.1287/mksc.1060.0254]
28. Beyrer C, Crago A, Bekker L, Butler J, Shannon K, Kerrigan D, et al. An action agenda for HIV and sex workers. Lancet 2015 Jan 17;385(9964):287-301 [FREE Full text] [doi: 10.1016/S0140-6736(14)60933-8] [Medline: 25059950]
29. Li KT, Huang W, Tang W, Wu F, Zhao Y, Wu D, et al. A secondary mixed methods analysis of a pay-it-forward gonorrhea/chlamydia testing program among men who have sex with men in China. Sex Transm Dis 2020 Jun;47(6):395-401 [FREE Full text] [doi: 10.1097/OLQ.0000000000001161] [Medline: 32149952]
30. Hui BPH, Ng JCK, Berzaghi E, Cunningham-Amos LA, Kogan A. Rewards of kindness? A meta-analysis of the link between prosociality and well-being. Psychol Bull 2020 Dec;146(12):1084-1116. [doi: 10.1037/bul0000298] [Medline: 32881540]
31. Curry OS, Rowland LA, Van Lissa CJ, Zlotowitz S, McAlaney J, Whitehouse H. Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor. J Exp Soc Psychol 2018 May;76:320-329 [FREE Full text] [doi: 10.1016/j.jesp.2018.02.014]
32. Tang W, Mao J, Li KT, Walker JS, Chou R, Fu R, et al. Pregnancy and fertility-related adverse outcomes associated with infection: a global systematic review and meta-analysis. Sex Transm Infect 2020 Aug;96(5):322-329 [FREE Full text] [doi: 10.1136/sextrans-2019-053999] [Medline: 31836678]
33. Chemaitelly H, Weiss H, Smolak A, Majed E, Abu-Raddad LJ. Epidemiology of Treponema pallidum, Chlamydia trachomatis, Neisseria gonorrhoeae, Trichomonas vaginalis, and herpes simplex virus type 2 among female sex workers in the Middle East and North Africa: systematic review and meta-analytics. J Glob Health 2019 Dec;9(2):020408 [FREE Full text] [doi: 10.7189/jogh.09.020408] [Medline: 31360448]

Abbreviations

FSW: female sex worker
GEE: generalized estimating equations
ICER: incremental cost-effectiveness ratio
LMICs: low- and middle-income countries
MSM: men who have sex with men
RCT: randomized controlled trial
STI: sexually transmitted infection

Edited by A Mavragani, T Sanchez; submitted 24.10.22; peer-reviewed by J Jiang, U Roy Paladhi; comments to author 18.01.23; revised version received 08.02.23; accepted 14.02.23; published 02.03.23.

Please cite as:
Tang W, Xie Y, Xiong M, Wu D, Ong JJ, Wi TE, Yang B, Tucker JD, Wang C
A Pay-It-Forward Approach to Improve Chlamydia and Gonorrhea Testing Uptake Among Female Sex Workers in China: Venue-Based Superiority Cluster Randomized Controlled Trial
JMIR Public Health Surveill 2023;9:e43772
URL: https://publichealth.jmir.org/2023/1/e43772
doi:10.2196/43772
PMID:36862485

Original Paper

Understanding Public Attitudes and Willingness to Share Commercial Data for Health Research: Survey Study in the United Kingdom

Yasemin Hirst[1,2,3], BA, MSc, PhD; Sandro T Stoffel[2,4], BA, MSc, PhD; Hannah R Brewer[3], BSc, MSc, PhD; Lada Timotijevic[5,6,7], BSc, PhD; Monique M Raats[5,6,7], MSc, PhD; James M Flanagan[3], BSc, PhD

[1]Lancaster Medical School, Lancaster University, Lancaster, United Kingdom
[2]Department of Behavioural Science and Health, University College London, London, United Kingdom
[3]Department of Surgery and Cancer, Imperial College London, London, United Kingdom
[4]Institute of Pharmaceutical Medicine, University of Basel, Basel, Switzerland
[5]School of Psychology, University of Surrey, Guildford, United Kingdom
[6]Surrey Institute for People-Centred Artificial Intelligence, University of Surrey, Guildford, United Kingdom
[7]Institute for Sustainability, University of Surrey, Guildford, United Kingdom

Corresponding Author:
Yasemin Hirst, BA, MSc, PhD
Lancaster Medical School
Lancaster University
Sir John Fisher Drive, Bailrigg
Lancaster, LA1 4YW
United Kingdom
Phone: 44 1524 65201
Email: y.hirst@lancaster.ac.uk

Abstract

Background: Health research using commercial data is increasing. The evidence on public acceptability and sociodemographic characteristics of individuals willing to share commercial data for health research is scarce.

Objective: This survey study investigates the willingness to share commercial data for health research in the United Kingdom with 3 different organizations (government, private, and academic institutions), 5 different data types (internet, shopping, wearable devices, smartphones, and social media), and 10 different invitation methods to recruit participants for research studies with a focus on sociodemographic characteristics and psychological predictors.

Methods: We conducted a web-based survey using quota sampling based on age distribution in the United Kingdom in July 2020 (N=1534). Chi-squared tests tested differences by sociodemographic characteristics, and adjusted ordered logistic regressions tested associations with trust, perceived importance of privacy, worry about data misuse and perceived risks, and perceived benefits of data sharing. The results are shown as percentages, adjusted odds ratios, and 95% CIs.

Results: Overall, 61.1% (937/1534) of participants were willing to share their data with the government and 61% (936/1534) of participants were willing to share their data with academic research institutions compared with 43.1% (661/1534) who were willing to share their data with private organizations. The willingness to share varied between specific types of data—51.8% (794/1534) for loyalty cards, 35.2% (540/1534) for internet search history, 32% (491/1534) for smartphone data, 31.8% (488/1534) for wearable device data, and 30.4% (467/1534) for social media data. Increasing age was consistently and negatively associated with all the outcomes. Trust was positively associated with willingness to share commercial data, whereas worry about data misuse and the perceived importance of privacy were negatively associated with willingness to share commercial data. The perceived risk of sharing data was positively associated with willingness to share when the participants considered all the specific data types but not with the organizations. The participants favored postal research invitations over digital research invitations.

Conclusions: This UK-based survey study shows that willingness to share commercial data for health research varies; however, researchers should focus on effectively communicating their data practices to minimize concerns about data misuse and improve public trust in data science. The results of this study can be further used as a guide to consider methods to improve recruitment strategies in health-related research and to improve response rates and participant retention.

(JMIR Public Health Surveill 2023;9:e40814) doi:10.2196/40814

KEYWORDS

commercial data; data sharing; participant recruitment; loyalty cards; sociodemographic factors; data donation; data; health; public; acceptability; digital; mobile phone

Introduction

Health researchers are increasingly aiming to include accurate personal information collected outside of health care settings, including commercial data collected or processed for businesses relating to their customers (eg, internet searches, social media, loyalty cards, wearable devices, and mobile phone apps), to enhance our understanding of individuals' health-related behaviors and health outcomes. With the rise of different data sources to track, monitor, and forecast disease and health outcomes, interest in carrying out research using individual commercial data has grown substantially [1-3]. However, much of this valuable research is often criticized for its representativeness and the low participation rates associated with public attitudes toward data sharing.

Evidence on the public acceptability of sharing health-related data is vast and suggests that improving the transparency of data collection and processing practices across institutions, creating trustworthy data ecosystems, and providing agency and data stewardship for data participants can improve the willingness to take part in research and share data [4-12]. The evidence for willingness to share data across different contexts varies depending on the type, purpose, and use of data. These often show that the public has some understanding of how data are being used and equally suggest that raising awareness about data practices does not increase willingness to share data [9]. A recent report highlights that further research is needed to improve public trust in light of the Cambridge Analytica scandal and other reported data misuse incidences [12].

Furthermore, with the implementation of the General Data Protection Regulation (GDPR) in 2018 [13], all individuals were given the right to carry out a subject access request from any organization that holds any information about them, thus allowing researchers to start analyzing new types of data sets with individual consent to understand behaviors such as diet [14], self-medication [15], and cancer risk [16] using purchase history recorded on loyalty cards; that is, an identity card issued by retailers to its customers to collect information on buyer behavior and generate reward schemes. However, a common limitation is the small sample size and biased population of individuals who are more willing to share their data [17]. A number of qualitative studies investigated willingness to share commercial data, specifically loyalty cards for health research, echoing the principal evidence shared across disciplines, as discussed earlier [15,18,19]. In contrast, for mobile and biosensor data sharing, there is a growing body of literature on the importance of understanding nonparticipation and willingness to share mobile phone apps and biosensor data [20-24]. A study that took place in England before the GDPR highlighted that, in the context of mobile data sharing, user behavior is also associated with willingness to share passive or actively collected mobile data [21]. Further experimental studies have highlighted the behavior of sharing mobile data, which requires capabilities from the users to fulfill the task and the characteristics of the individuals, framing of the request, emphasis on control over data, and assurances of privacy and confidentiality [22]. The implications of the willingness to share smartphone and sensor data with researchers are further understood in studies where the response rate for data sharing is less than 15%, and the representativeness of the population that shares the data is less than optimal [23]. This highlights the importance of understanding the characteristics of the population who are willing to share commercial data sets before data collection so that strategies can be developed to improve response rates and minimize bias.

Therefore, this study aims to investigate GDPR awareness and sociodemographic and psychological factors associated with the willingness to share commercial data for health research purposes after the implementation of the GDPR in the United Kingdom. Furthermore, it aimed to provide a summary of the public's awareness of GDPR in 2020, 2 years after the GDPR and Data Protection Act 2018 were enacted in the United Kingdom. The GDPR has been kept in UK law as the UK GDPR [24]. For epidemiological research to advance using commercial data sets and effectively recruit participants, it is important to investigate the factors associated with the willingness to share commercial data for health research.

Methods

Setting and Design

A 10-minute web-based survey was conducted in the United Kingdom in August 2020 via Survey Monkey using Dynata International Limited. Nonprobability quota sampling was used for an adequate representation of different age groups in the United Kingdom, with the aim of recruiting 1500 participants to achieve a 1:10 participant-to-item ratio [25]. The distribution of the sample, respectively, based on age and sex distribution in the UK population was 18 to 29 years (20%), 30 to 39 years (17%), 40 to 49 years (18.5%), 50 to 59 years (15.5%), 60 to 69 years (14%), >70 years (15%), male (49.4%), and female (50.6%) [26].

Ethics Approval

The project was reviewed by the University of College London Research Ethics Committee and received a favorable opinion (ref: 18095/001) and reported using the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for cross-sectional research [27]. Information regarding ethics approval can be obtained directly from the University of College London Research Ethics Committee. The survey study only included anonymized data from the participants; therefore, the research team had no contact with the participants following their participation in the study. If individuals dropped out of the survey before its completion, this was considered a withdrawal from the study, and no data were included. Participants were paid a small monetary incentive through Dynata International Limited in line with their participant payment policies.

Survey Measures

All measures with their item heritage are included in Multimedia Appendix 1. There were 3 primary outcome measures. These were the willingness to share commercial data for health research with different institutions (government, private, and academic), the willingness to share different types of commercial data (internet searches, social media, shopping data on loyalty cards, wearable devices, and mobile phone apps) with academic institutions, and the willingness to participate in health research based on different invitation sources. The rationale for these outcomes is as follows. In comparison to government and private organizations, which are often the primary data controllers for health, administrative, and commercial data, researchers at academic institutions often need to request access to data collected and controlled by government and private organizations. The differences between institutions were considered to understand the potential baseline response rate for potential research projects that aim to use commercial data in health research at academic institutions. The second primary outcome was then focused on willingness by data type and the extent to which this differentiates from the baseline willingness to share commercial data with academic institutions. The last outcome is included to consider how much willingness varied depending on the source of the invite to better understand the best ways to recruit participants who are more willing to share their data. All these outcomes were used to inform the communication strategies of a much larger academic project that aimed to recruit individual participants with informed consent requesting access to their commercial data, specifically loyalty card data from 2 UK-based high-street retailers, investigating self-care behaviors before ovarian cancer diagnosis [16].

The independent variables were included under four sections: (1) sociodemographic factors, including age, sex, marital status, education, ethnicity, and location in the United Kingdom; (2) the participants' GDPR awareness; (3) psychological factors including trust in institutions, trust in data practices in academia, worry about data misuse, perceived risk in data sharing, perceived importance of privacy, and perceived benefit of data sharing; and (4) past experience taking part in health research and past experiences of data misuse.

Statistical Analysis

We reported all measures and exclusions in this study and used complete case analysis without imputing missing data, as all questions were mandatory.

Factor analyses using principal component analysis (PCA) and reliability tests were carried out to ensure that the items included in various other studies measured the intended outcomes. Once the factors were identified, the scales were computed using total scores. A Cronbach alpha coefficient was calculated for each scale for internal consistency, and the interitem and interscale correlations were checked for internal consistency of items and scales. Each computed scale was reported using range, mean, SD, and Cronbach alpha coefficient.

Participant characteristics, self-reported GDPR awareness, and people's awareness of personal data and GDPR law were reported using descriptive statistics. Some categorical items were recoded for ease of presentation and understanding of the differences in each category. Responses to items including "prefer not to say" and "other" were coded as missing because of the low cell count (<5 observations in each category) in sociodemographic items which would not have been coded negatively and subsequently excluded from the main analyses (40/1594, 2.5%). Primary outcome variables were recoded into "definitely yes" or "probably yes"=1 and "probably no" or "definitely no"=0 to compare 2 distinct intentions to share data for the comparison between sociodemographic characteristics of the participants [28]. Differences in the proportions of willingness to share commercial data were tested using chi-square statistics and reported in percentages. Ordered logistic regression was used to test for psychological factors associated with willingness to share commercial data for health research with different organizations and different types of data adjusted for the sociodemographic characteristics of the participants, previous research participation, and GDPR awareness. The variance explained by each model is included in Multimedia Appendix 1. Further ordered regression analyses were carried out for the different types of research invitations to identify whether there were sociodemographic factors associated with willingness to participate in health research (Multimedia Appendix 1). All results reported using adjusted odds ratios (aORs) and 95% CIs were reported using a P value of <.05.

Results

Factor Analysis Results

The 32 items measured in this study were subjected to PCA using SPSS (version 27; IBM Corp). Before performing PCA, the suitability for performing PCA was assessed. Inspection of the correlation matrix revealed the presence of many coefficients >0.3. The Kaiser-Meyer-Olkin value was 0.93, above the recommended value of 0.6, and the Bartlett Test of Sphericity reached statistical significance, supporting the factorability of the correlation matrix. The PCA revealed the presence of 5 components with eigenvalues exceeding 1, explaining 28.6%, 19.7%, 7.5%, 6.9%, and 5.3% variance (Multimedia Appendix 1). A total of 8 items were recoded, and 2 items were deleted, as they measured trust in 2 different organizations. On the basis of these results, 5 scales were computed. These are, namely, perceived importance of privacy (mean 8.46, SD 1.38; range 2-10; Cronbach α=.65), worry about data misuse (6 items; mean 21.28, SD 5.69; range 5-30; Cronbach α=.95), trust in data practices in academic institutions (9 items; mean 31.96, SD 7.14; range 5-45; Cronbach α=.95), perceived risk of data sharing for health research (3 items; mean 9.40, SD 2.73; range 3-15; Cronbach α=.88) and perceived benefits of data sharing (5 items; mean 17.76, SD 4.31; range 5-25; Cronbach α=.93). Factor correlations as separate scales suggest that the scales have weak to moderate correlations, indicating that they measure separate scales (Multimedia Appendix 1).

Participant Characteristics

Out of the 1897 responses, 1534 participants gave their consent and completed the survey (Table 1). Approximately 49.1% (753/1534) of participants were male, 50.7% (777/1534) were female, and 0.2% (4/1534) indicated other. The age distribution of participants was consistent with the quota sample for the distribution of age in England. Most respondents self-identified with a White ethnic background (1325/1534, 86.4%), compared with only 12.9% (198/1534) who identified themselves with other ethnicities. Approximately 53.1% (814/1534) of the participants were married or had a legal partnership. Approximately 45.6% (699/1534) of participants had higher education (degree and above) qualifications, almost half of them (765/1534, 49.9%) had less than higher education qualifications, and only 4.5% (69/1534) did not have any educational qualification.

Table 1. Participant characteristics (N=1534).

Characteristics	Values, n (%)	Population composition of England and Wales based on 2011 Census (excludes Scotland and Northern Ireland) [21], %
Sex		
Male	753 (49.1)	49.4
Female	777 (50.7)	50.6
Other	4 (0.2)	0
Age (years)		
18-29	284 (18.5)	20
30-39	255 (16.6)	17
40-49	250 (16.3)	18.5
50-59	266 (17.3)	15.5
60-69	213 (13.9)	14
≥70	266 (17.3)	15
Ethnicity		
White British	1326 (87.0)	85.4
Black	45 (3.0)	3.5
Asian	111 (7.3)	7.1
Mixed	32 (2.1)	2.3
Other	10 (0.7)	1.0
Marital status		
Single	492 (32.4)	34.5
Married or legal partnership	814 (53.1)	50.8
Widowed, divorced, or separated	212 (13.8)	14.6
Educational level		
Higher education	697 (45.6)	27.1
Higher education with qualification	762 (49.9)	49.9
No qualification	68 (4.5)	23
Location in the United Kingdom		
London	247 (16.2)	14.6
East of England	149 (9.8)	10.4
South East	205 (13.4)	15.4
South West	136 (8.9)	9.4
West and East Midlands	218 (14.3)	18.1
Yorkshire and the Humber and North East	197 (12.9)	14
North West	184 (12)	12.6
Scotland	129 (8.4)	N/A[a]
Wales	62 (4.1)	5.5
General Data Protection Regulation awareness		
Not aware	179 (11.7)	—[b]
Yes, I have heard but do not know much about it	428 (27.9)	—
Yes, I have heard and know a little about it	658 (42.9)	—
Yes, I have heard and I know a lot about it	269 (17.5)	—
Previous health research participation		

Characteristics	Values, n (%)	Population composition of England and Wales based on 2011 Census (excludes Scotland and Northern Ireland) [21], %
Yes	402 (26.2)	—
No	1132 (73.8)	—

[a] N/A: not applicable.

[b] Data are not available for the distribution of the General Data Protection Regulation Awareness and previous health research participation in England.

GDPR and Personal Data Awareness

At the time of the survey, 11.7% (179/1534) of the participants indicated that they were not aware of GDPR, 27.9% (428/1534) had heard of GDPR but did not know much about it, 42.9% (658/1534) had heard and knew a little about GDPR, and 17.5% (269/1534) of respondents had heard and knew a lot about GDPR.

The results of participants' expectations of what is considered personal data under GDPR showed that >80% (1227/1534) of the participants were able to correctly state common information that was classified as personal information, such as name, age, gender, marital status, and home address and email address. Less than 75% (1150/1534) of participants considered sensitive personal information, such as sexual orientation (1121/1534, 73.1%), religion (1067/1534, 69.6%), criminal records (1100/1534, 71.7%), and health or medical records (1136/1534, 74.1%) as personal data.

Less than two-thirds of the participants expected the various types of information collected on the internet to count as personal data. A quarter of the participants incorrectly stated that web-based purchases (383/1534, 25.0%), location data based on General Packet Radio Service recorded on mobile phones (353/1534, 23%), tracking information on websites (cookies; 383/1534, 25%), social media information (424/1534, 27.6%), and device IDs (353/1534, 23%) were not personal data.

More than 80% (1227/1534) of the participants correctly identified what GDPR law should cover most of the rights that protect personal data. Most of the remaining participants stated that they did not know the right answer ranging from 6.6% (101/1534) to 13.4% (205/1534). Approximately 13.4% (205/1534) did not know that they had the right to erase their data, and 12.3% (189/1534) did not know that they had the right to be informed about the use of their data. Additional details are provided in Multimedia Appendix 1.

Willingness to Share Commercial Data With Different Institutions

Descriptive Results

Figure 1 shows that nearly two-thirds of the participants indicated that they would be willing to share their commercial data for health research if their data are shared with a government institution (937/1534, 61.1%) or an academic research institution (936/1534, 61.0%). In contrast, less than half were happy to share their commercial data with private organizations for health research (658/1534, 42.9%). Across all participants, only 4.8% (73/1534) of the participants stated "definitely yes" to share with all types of institutions. In comparison, 7.4% (114/1534) of the participants stated "definitely no" to share commercial data for health research with all institutions.

Figure 1. Willingness to share commercial data with different organizations for health research.

Sociodemographic Factors

In Table 2, the analysis shows significant differences in the willingness to share commercial data with government institutions by age, sex, education, and GDPR awareness. Specifically, less than two-thirds of participants were happy to share their data among those aged 40 to 49 years (841/1534, 54.8%) and 50 to 59 years (871/1534, 56.8%), compared with the 18 to 29 (923/1534, 60.2%), 30 to 39 (1040/1534, 67.8%), 60 to 69 (922/1534, 60.1%), and ≥70 years (1020/1534, 66.5%) groups, respectively (χ^2_5=14.6, *P*=.12). Male participants were more likely to share their commercial data for health research than female participants with the government (994/1534, 64.8% vs 887/1534, 57.8%; χ^2_1=7.9, *P*=.005). There was a 14.4% difference between willingness to share among those who were unaware of GDPR (822/1534, 53.6%) and those who stated that they knew a lot about GDPR (1043/1534, 68.0%; χ^2_3=9.7; *P*=.02). No differences were found in marital status, ethnicity, previous research participation, and personal experience of data misuse in the past.

There were differences in the willingness to share commercial data with private organizations based on most factors, except for education and previous research participation. The largest differences were observed for age between those who were aged 30 to 39 years (147/255, 57.6%), and 60 to 69 years (64/213, 30%), and ≥70 years (80/266, 30.1%), as well as by ethnicity among those identified as Black (28/45, 62.2%) and White (555/1326, 41.9%; χ^2_5=74.1, *P*<.001). Similarly, only one-third of those who reported not being aware of GDPR were willing to share their data (54/179, 30.2%) compared with those who were aware but did not know much (177/428, 41.4%), a little (292/658, 44.4%), and knew a lot about GDPR (138/269, 51.3%; χ^2_3=20.5, *P*<.001). Ever experienced a negative event of data misuse was also positively associated with willingness to share commercial data for health research with private organizations compared with never experiencing a negative event (287/596, 48.2% vs 374/938, 39.9%; χ^2_1=10.1, *P*=.001).

There were no significant differences in the proportion of people who indicated “Definitely and Probably yes” to sharing with academic institutions by marital status, age, sex, ethnicity, and previous experience. However, educational level (above degree: 459/697, 65.9% vs below degree: 474/830, 57.1%; χ^2_1=12.1, *P*<.001), previous participation in health research (yes: 266/402, 66.2% vs no: 670/1132, 59.2%; χ^2_1=6.0, *P*=.01), and greater GDPR awareness (not aware 84/179, 46.9% vs know a lot about it 179/269, 66.5%; χ^2_3=20.8, *P*<.001) were positively associated with sharing data with academic institutions.

Table 2. Willingness to share commercial data with different organizations by sociodemographic factors (N=1534).

Willingness to share	Government organizations	Private organizations	Academic institutions
Marital status			
Single, n (%)	284 (57.7)	230 (46.7)	306 (62.2)
Married or legal partnership, n (%)	512 (62.9)	354 (43.5)	489 (60.1)
Widowed, divorced, or separated, n (%)	134 (63.2)	73 (34.4)	134 (63.2)
Chi-square (*df*)	3.8 (2)	9.1 (2)	1.0 (2)
P value	.14	.01	.61
Age (years)			
18-29, n (%)	171 (60.2)	153 (53.9)	177 (62.3)
30-39, n (%)	173 (67.8)	147 (57.6)	168 (65.9)
40-49, n (%)	137 (54.8)	118 (47.2)	144 (57.6)
50-59, n (%)	151 (56.8)	99 (37.2)	154 (57.9)
60-69, n (%)	128 (60.1)	64 (30.0)	129 (60.6)
≥70, n (%)	177 (66.5)	80 (30.1)	164 (61.7)
Chi-square (*df*)	14.6 (5)	74.1 (5)	5.1 (5)
P value	.01	<.001	.40
Ethnicity			
White, n (%)	808 (60.9)	555 (41.9)	821 (61.9)
Black, n (%)	33 (73.3)	28 (62.2)	26 (57.8)
Asian, n (%)	66 (59.5)	55 (49.5)	64 (57.7)
Mixed, n (%)	18 (56.3)	15 (46.9)	15 (46.9)
Other ethnicities, n (%)	6	<5	5
Chi-square (*df*)	3.2 (4)	9.6 (4)	4.3 (4)
P value	.51	.04	.35
Sex			
Male, n (%)	488 (64.8)	359 (47.7)	473 (62.8)
Female, n (%)	449 (57.8)	300 (38.6)	462 (59.5)
Chi-square (*df*)	7.9 (1)	12.8 (1)	1.8 (1)
P value	.005	<.001	.17
Education			
<Degree and no formal education, n (%)	481 (58.0)	367 (44.2)	474 (57.1)
≥Degree, n (%)	451 (64.7)	290 (41.6)	459 (65.9)
Chi-square (*df*)	7.2 (1)	1.0 (1)	12.1 (1)
P value	.007	.31	<.001
General Data Protection Regulation awareness			
Not aware, n (%)	96 (53.6)	54 (30.2)	84 (46.9)
Yes, I have heard but do not know much about it, n (%)	258 (60.3)	177 (41.4)	254 (59.3)
Yes, I have heard but know little about it, n (%)	400 (60.8)	292 (44.4)	419 (63.7)
Yes, I have heard and I know a lot about it, n (%)	183 (68.0)	138 (51.3)	179 (66.5)
Chi-square (*df*)	9.7 (3)	20.5 (3)	20.8 (3)
P value	.02	<.001	<.001
Previous research participation			
Yes, n (%)	257 (63.9)	171 (42.5)	266 (66.2)

Willingness to share	Government organizations	Private organizations	Academic institutions
No, n (%)	680 (60.1)	490 (43.3)	670 (59.2)
Chi-square (*df*)	1.8 (1)	0.1 (1)	6.0 (1)
P value	.17	.79	.01
Personal experience of data misuse			
Never, n (%)	559 (59.7)	374 (39.9)	559 (59.6)
Ever, n (%)	378 (63.4)	287 (48.2)	377 (63.3)
Chi-square (*df*)	2.2 (1)	10.1 (1)	2.0 (1)
P value	.13	.001	.15

Psychological Predictors

Adjusted ordered regression analyses for willingness to share data with each institution in Table 3 show that greater trust is positively associated with sharing commercial data with government (aOR 2.499, 95% CI 2.228-2.802; *P*<.001), private (aOR 2.513, 95% CI 2.221-2.842; *P*<.001), and academic institutions (aOR 2.283, 95% CI 2.011-2.59; *P*<.001). Greater worry about data misuse was negatively associated with willingness to share with government (aOR 0.94, 95% CI 0.918-0.961; *P*<.001), private (aOR 0.951, 95% CI 0.930-0.973; *P*<.001), and academic institutions (aOR 0.947, 95% CI 0.926-0.969; *P*<.001).

Participants' perceived importance of privacy was negatively associated with willingness to share with the government (aOR 0.909, 95% CI 0.833-0.992; *P*=.03), private institutions (aOR 0.833, 95% CI 0.763-0.909; *P*<.001), and academic institutions (aOR 0.869, 95% CI 0.797-0.948; *P*=.002). Participants' perceived risk of data sharing was not associated with their willingness to share their data with any organization. The perceived benefits of sharing data were positively associated with government institutions (aOR 1.111, 95% CI 1.083-1.14; *P*<.001), private institutions (aOR 1.081, 95% CI 1.054-1.109; *P*<.001), and academic institutions (aOR 1.116, 95% CI 1.087-1.146; *P*<.001).

Table 3. Psychological predictors of willingness to share commercial data for health research with different organizations[a].

	Government institutes, aOR[b] (95% CI)	Private institutes, aOR (95% CI)	Academic institutes, aOR (95% CI)
Trust in organizations	2.499 (2.228-2.802[c])	2.513 (2.221-2.842[c])	2.283 (2.011-2.590[c])
Worry about data misuse	0.940 (0.918-0.961[c])	0.951 (0.930-0.973[c])	0.947 (0.926-0.969[c])
Perceived risk of data sharing	1.041 (0.997-1.086)	1.042 (0.997-1.089)	1.016 (0.974-1.060)
Perceived importance of privacy	0.909 (0.833-0.992[d])	0.833 (0.763-0.909[c])	0.869 (0.797-0.948[c])
Perceived benefits of sharing data and participation	1.111 (1.083-1.140[c])	1.081 (1.054-1.109[c])	1.116 (1.087-1.146[c])

[a] Adjusted for age, sex, location, ethnicity, education, General Data Protection Regulation awareness, and past health research participation. The full model with *P* values is reported in Multimedia Appendix 1.

[b] aOR: adjusted odds ratio.

[c] *P*<.001.

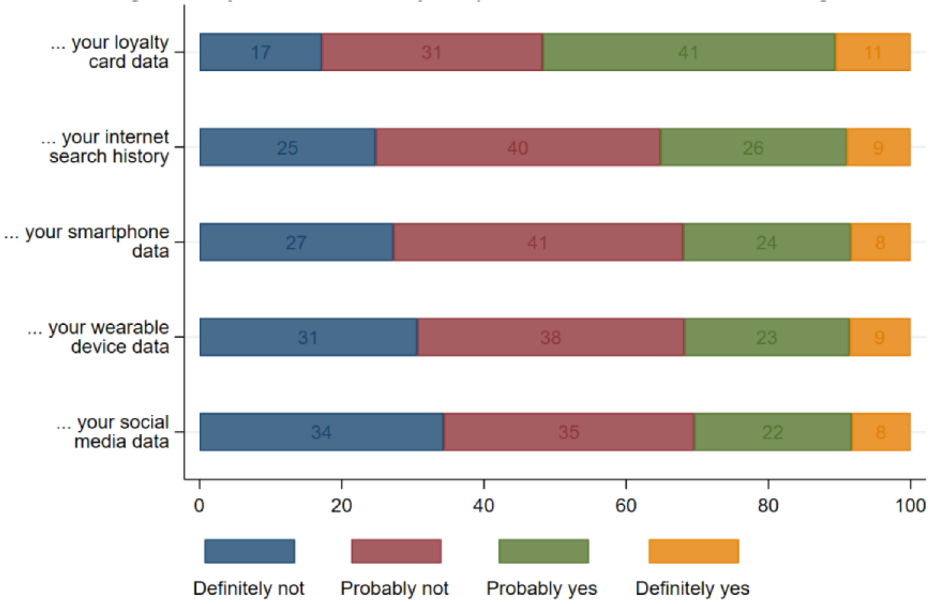
[d] *P*<.05.

Willingness to Share Different Types of Commercial Data

Descriptive Results

Figure 2 shows that the participants' willingness to share commercial data varied across all data types. The willingness to share loyalty card data had the highest proportion of participants at 51.8% (795/1534) stating that "Definitely or Probably yes." In comparison, the proportion was much lower at 35% (540/1534) for internet search history, 32% (491/1534) for smartphone data, 32% (488/1534) for sharing wearable device data, and 30% (467/1534) for social media data. Across all participants, only about 3.2% (49/1534) of the participants stated "definitely yes" to share all types of commercial data sets. In comparison, 13.3% (204/1534) of the participants stated "definitely no" to share all types of commercial data sets.

Figure 2. Willingness to share different types of commercial data with academic institutions for health research.



Sociodemographic Factors

Table 4 shows the proportion of people who stated “Definitely and Probably yes” for willingness to share different types of commercial data for health research with academic institutions based on the sociodemographic characteristics of the participants. There were significant differences across all types of health research data according to marital status, age, and past experience of data misuse. In contrast, no associations were found between the participants’ educational level and previous participation in the research. Greater GDPR awareness was positively associated with willingness to share all types of data, except for internet searches.

Among these characteristics, notable differences were observed for marital status, where a larger proportion of people who were single reported willingness to share commercial data sets compared with those who were married or in a legal partnership, or widowed, divorced, or separated. Furthermore, an increase in the age of participants was negatively associated with their willingness to share. Less than a fifth of the participants in the 60 to 69 years and above age groups were willing to share smartphone, social media, and wearable device data. Across all types of commercial data, those aged 18 to 29 years had the highest proportion of individuals willing to share at 65.1% (185/284) for loyalty card data, 48.6% (138/284) for smartphone data and wearable devices, 47.2% (134/284) for social media, and 46.8% (133/284) for internet data.

Female participants were less likely to share smartphone data (male: 282/753, 37.5% vs female: 207/777, 26.6%; χ^2_1=20.5 *P*<.001), wearable devices (male: 280/753, 37.2% vs female: 206/777, 26.5%; χ^2_1=20.0 *P*<.001), and social media data (male: 271/753, 36% vs female: 196/777, 25.2%; χ^2_1=20.8 *P*<.001).

In comparison, the differences in proportions were smaller for internet searches (male: 286/753, 38% vs female: 253/777, 32.6%; χ^2_1=4.9, *P*=.03) and loyalty card data (male: 399/753, 53% vs female: 393/777, 50.6%; χ^2_1=0.8, *P*=.35). The proportion of people willing to share loyalty card data did not differ by ethnicity or sex. In contrast, the proportion was lower among those from White ethnic backgrounds for internet searches, social media, wearable devices, and smartphone data compared with those who were identified from Black and Asian ethnic backgrounds.

Those who had ever experienced a data misuse event were more likely to share loyalty card data (ever: 344/596, 57.7% vs never: 450/938, 48%; χ^2_1=13.8, *P*<.001), internet search data (ever: 263/596, 44.1% vs never: 277/938, 29.5%; χ^2_1=34.0, *P*<.001), smartphone (ever: 245/596, 41.1% vs never: 246/938, 26.2%; χ^2_1=37.0, *P*<.001), social media (ever: 240/596, 40.3% vs never: 227/938, 24.2%; χ^2_1=44.4, *P*<.001), and wearable devices (ever: 255/596, 42.8% vs never: 233/938, 24.8%; χ^2_1=54.1, *P*<.001) than those with no previous data misuse experience.

Table 4. Willingness to share different types of commercial data by sociodemographic factors.

Willingness to share	Internet searches	Loyalty card	Smartphone	Social media	Wearable devices
Marital status					
Single, n (%)	205 (41.7)	297 (60.4)	190 (38.6)	180 (36.6)	194 (39.4)
Married or legal partnership, n (%)	273 (33.5)	406 (49.9)	263 (32.3)	235 (28.9)	250 (30.7)
Widowed, divorced, or separated, n (%)	57 (26.9)	87 (41.0)	34 (16.0)	48 (22.6)	41 (19.3)
Chi-square (*df*)	16.4 (2)	25.4 (2)	34.7 (2)	15.7 (2)	28.7 (2)
P value	<.001	<.001	<.001	<.001	<.001
Age (years)					
18-29, n (%)	133 (46.8)	185 (65.1)	138 (48.6)	134 (47.2)	138 (48.6)
30-39, n (%)	132 (51.8)	156 (61.2)	122 (47.8)	129 (50.6)	141 (55.3)
40-49, n (%)	98 (39.2)	134 (53.6)	94 (37.6)	81 (32.4)	87 (34.8)
50-59, n (%)	70 (26.3)	129 (48.5)	66 (24.8)	60 (22.6)	57 (21.4)
60-69, n (%)	45 (21.1)	91 (42.7)	34 (16.0)	29 (13.6)	29 (13.6)
≥70 , n (%)	62 (23.3)	99 (37.2)	37 (13.9)	34 (12.8)	36 (13.5)
Chi-square (*df*)	93.4 (5)	60.3 (5)	140.4 (5)	162.3 (5)	189.4 (5)
P value	<.001	<.001	<.001	<.001	<.001
Ethnicity					
White, n (%)	447 (33.7)	674 (50.8)	400 (30.2)	377 (28.4)	397 (29.9)
Black, n (%)	25 (55.6)	26 (57.8)	22 (48.9)	25 (55.6)	24 (53.3)
Asian, n (%)	53 (47.7)	66 (59.5)	50 (45.0)	47 (42.3)	49 (44.1)
Mixed, n (%)	9 (28.1)	20 (62.5)	13 (40.6)	13 (40.6)	15 (46.9)
Other ethnicities, n (%)	5 (50.0)	6 (60.0)	<5	<5	<5
Chi-square (*df*)	18.7 (4)	5.4 (4)	17.9 (4)	24.8 (4)	23.4 (4)
P value	.001	.24	.001	<.001	<.001
Sex					
Male, n (%)	286 (38.0)	399 (53.0)	282 (37.5)	271 (36.0)	280 (37.2)
Female, n (%)	253 (32.6)	393 (50.6)	207 (26.6)	196 (25.2)	206 (26.5)
Chi-square (*df*)	4.9 (1)	0.8 (1)	20.5 (1)	20.8 (1)	20.0 (1)
P value	.02	.34	<.001	<.001	<.001
Education					
<Degree and no formal education, n (%)	290 (34.9)	435 (52.4)	252 (30.4)	244 (29.4)	248 (29.9)
≥Degree, n (%)	247 (35.4)	355 (50.9)	235 (33.7)	219 (31.4)	238 (34.1)
Chi-square (*df*)	0.4 (1)	0.3 (1)	1.9 (1)	0.7 (1)	3.1 (1)
P value	.83	.56	.16	.39	.07
General Data Protection Regulation awareness					
Not aware, n (%)	52 (29.1)	74 (41.3)	49 (27.4)	39 (21.8)	48 (26.8)
Yes, I have heard but do not know much about it, n (%)	159 (37.1)	216 (50.5)	129 (30.1)	136 (31.8)	129 (30.1)
Yes, I have heard but know little about it, n (%)	221 (33.6)	351 (53.3)	196 (29.8)	192 (29.2)	197 (29.9)
Yes, I have heard and I know a lot about it, n (%)	108 (40.1)	153 (56.9)	117 (43.5)	100 (37.2)	114 (42.4)
Chi-square (*df*)	7.3 (3)	11.5 (3)	20.2 (3)	12.9 (3)	17.5 (3)

Willingness to share	Internet searches	Loyalty card	Smartphone	Social media	Wearable devices
P value	.06	.009	<.001	.005	.001
Previous research participation					
Yes, n (%)	131 (32.6)	195 (48.5)	122 (30.3)	115 (28.6)	114 (28.4)
No, n (%)	409 (36.1)	599 (52.9)	369 (32.6)	352 (31.1)	374 (33.0)
Chi-square (*df*)	1.6 (1)	2.3 (1)	0.6 (1)	0.8 (1)	2.9 (1)
P value	.20	.12	.40	.35	.08
Personal experience of data misuse					
Never, n (%)	277 (29.5)	450 (48.0)	246 (26.2)	227 (24.2)	233 (24.8)
Ever, n (%)	263 (44.1)	344 (57.7)	245 (41.1)	240 (40.3)	255 (42.8)
Chi-square (*df*)	34.0 (1)	13.8 (1)	37.0 (1)	44.4 (1)	54.1 (1)
P value	<.001	<.001	<.001	<.001	<.001

Psychological Predictors

The results from the ordered logistic regression analyses adjusted for the sociodemographic characteristics of the participants in Table 5 show that each point increase in trust in data practices in academia, perceived benefits in participation, and perceived risks of data sharing are positively associated with willingness to share all types of commercial data. In contrast, each point increase in the perceived importance of privacy and worry about data misuse was negatively associated with the willingness to share all types of commercial data.

Table 5. Psychological predictors of willingness to share different types of commercial data for health research with academic institutions[a].

	Internet search data, aOR[b] (95% CI)	Loyalty card data, aOR (95% CI)	Smartphone data, aOR (95% CI)	Social media data, aOR (95% CI)	Wearable devices data, aOR (95% CI)
Trust in data practices in academic institutions	1.097 (1.078-1.117[c])	1.103 (1.083-1.123[c])	1.097 (1.077-1.117[c])	1.087 (1.067-1.107[c])	1.075 (1.056-1.095[c])
Perceived importance of privacy	0.682 (0.625-0.744[c])	0.707 (0.648-0.772[c])	0.685 (0.628-0.748[c])	0.716 (0.656-0.781[c])	0.732 (0.670-0.798[c])
Worry about data misuse	0.960 (0.938-0.982[c])	0.975 (0.953-0.997[d])	0.960 (0.938-0.982[c])	0.963 (0.941-0.985[c])	0.940 (0.919-0.962[c])
Perceived benefit in data sharing and research participation	1.057 (1.029-1.086[c])	1.102 (1.072-1.132[c])	1.070 (1.040-1.100[c])	1.049 (1.020-1.078[c])	1.086 (1.056-1.116[c])
Perceived risks of data sharing	1.222 (1.169-1.276[c])	1.114 (1.067-1.163[c])	1.224 (1.171-1.279[c])	1.208 (1.156-1.263[c])	1.232 (1.179-1.287[c])

[a] Adjusted for age, sex, location, ethnicity, education, General Data Protection Regulation awareness, and past health research participation. The full model with *P* values is reported in Multimedia Appendix 1.

[b] aOR: adjusted odds ratio.

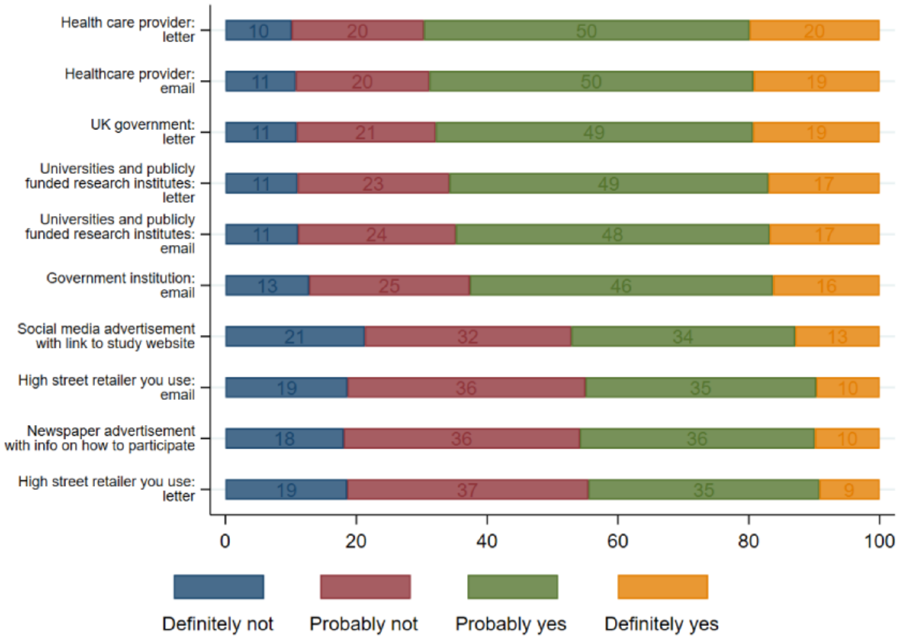
[c] *P*<.001.

[d] *P*<.05.

Willingness to Take Part in Research Based on Invitation Sources

Figure 3 shows that the most preferred ways of being invited to health research using commercial data were receiving a letter (1069/1534, 69.6%) or an email invitation (1056/1534, 68.6%) from the health care provider, followed by a letter invitation from the government (1041/1534, 67.9%) and universities or publicly funded research institutes (1009/1534, 65.8%). Digital invitations had a much lower preference compared with letter-based invitations, except for letter invitations from high-street retailers to participate in the research (683/1534, 44.5%). Research advertisements on social media or newspapers were preferable for less than half of the participants, with 47.1% (723/1534) and 45.8% (703/1534), respectively. The ordered logistic regression analysis in Multimedia Appendix 1 shows that across all invitation types, greater GDPR awareness was the only predictor of invitation source, and there were some nuanced differences in sociodemographic characteristics.

Figure 3. Preferences for the source of research participation invitation.



Discussion

Principal Findings

With an interest to develop a better understanding of psychological factors and sociodemographic characteristics of individuals who would be willing to participate in health research using individual and commercially collected data sets, this study investigates the willingness to share commercial data for health research with different institutions and different types of commercial data in an age-stratified population-based sample in the United Kingdom. Our results showed that two-thirds of the participants were willing to share their commercial data on health research with academic institutions. In contrast, when participants were specifically asked about sharing different types of commercial data for health research with a focus on academic institutions, only about half of them were willing to share their shopping data, and less than one-third were happy to share internet search history, wearable devices, social media, and smartphone apps. Only a small minority of the participants across all outcomes were willing to share their data, highlighting the potential barrier in participant recruitment that needs to be addressed in health research using complex data sets.

Comparison With Prior Work

A key outcome of this study is the validation of the previous evidence that, irrespective of individuals' GDPR awareness and their sociodemographic characteristics, greater trust is consistently associated with greater willingness to share commercial data for health research [12]. This study also adds to the evidence that greater perceived importance of data privacy and greater worry about data misuse are negatively associated with the willingness to share commercial data for health research. Interestingly, the perceived risk of sharing data was not associated with the willingness to share data with institutions, but it played a role in all types of data that they were willing to share. This is in line with the Nissenbaum [29] contextual integrity framework for privacy, suggesting that individuals' information-sharing principles are context dependent and socially constructed. A previous focus group study also found similar results in that participants were more concerned about the use of subjective data sets such as social media posts compared with objective shopping data collected on loyalty cards as more important than the organizations in which they share their data with [15]. Furthermore, we found a positive univariate association between the participants who had ever experienced a negative data-related event, for example, they might have had a data breach and personal information stolen on the internet and used for other people's gain, were more likely to share their data. Although we do not have sufficient information to assess why this may be the case, it could be interpreted that those who have experienced a negative event are more likely to be risk-aware than risk-averse. An interesting explanation for this result could be studied further based on "the privacy paradox," which suggests that intentions to share data may not be directly associated with actual behavior, and other mediators should be better understood [30]. Future experimental studies with a factorial design can further explore how these experiences impact people's perceptions and use of technologies.

We identified various sociodemographic factors associated with all outcome variables with the participants' age particularly being an important factor to be considered for participant recruitment. Our results showed that the participants' willingness reduced with an increase in age, except for sharing data with government institutions where we observed an inverse association. This is an important outcome to consider when implementing pilot health interventions for the general population [17]. The participants who identified themselves as Black consistently had a higher proportion of willingness not just across all institutions but also for all data types. Health researchers should identify resources to improve the visibility of health research opportunities to improve participation and diversity in research. Recruitment through social media advertisements has been shown to be effective in targeting minority populations [31]. However, there is a lack of ethical and methodological guidance for recruitment via paid social media advertisements to be carried out effectively [32].

Limitations

A key limitation of this survey was that participants were not provided with examples of how commercial data could be used for specific health studies, such as facilitating earlier cancer diagnosis, identifying mental health conditions, and not including nonprofit organizations. A previous study showed that people were more willing to donate their data to Cancer Research UK compared with nonspecific health research organizations, which was found to be associated with individuals' level of altruism and prosocial tendencies [33]. Owing to the exploratory nature of this study, a priori hypotheses were not included in the statistical analysis plans, and the following warnings are warranted. Although we adjusted for past health research participation, participants from the recruitment panel were subject to desirability bias. The social desirability bias in perceived privacy, intention to share data, and actual behaviors has previously been demonstrated in the privacy paradox [30]. Future studies arising from this study could investigate mediators between intention and behavior gaps. Similarly, the use of nonprobability sampling of the participants could also lead to greater bias in this study. Although accumulating evidence suggests that the willingness to share commercial data is a complex behavior and cannot be reduced to one-off intention measures, we believe that the outcomes of this study can be used as a guide for identifying populations with the least likelihood of sharing data when recruitment methods for studies requesting data from its participants are operationalized. However, it should be acknowledged that the measures included in this study, as well as sociodemographic characteristics, explain less than 20% of the variance in the willingness to share commercial data (Multimedia Appendix 1). This highlights the complexity of the evidence surrounding data sharing and slow progress in health-related research in relation to building a better understanding of the mechanisms that hinder and facilitate data-sharing principles. Thus, future studies could benefit from the use of a theoretical model, such as the capability, opportunity, motivation, and behavior model [34], to understand how physical and psychological capabilities, such as participants' engagement with existing technologies, could potentially moderate their willingness to share data. Similarly, social opportunities could be an important factor for the willingness to participate in research advertised through social media. Notwithstanding, this survey included a UK age-representative cohort based on the 2011 Census [26], had a larger proportion of individuals from non-White ethnic backgrounds in comparison with other panel-based survey studies with a UK population representative sample [35,36], and also adjusted for the participants' geographic region in the United Kingdom to improve the external validity and generalizability of its outcomes for the wider UK population.

Conclusions

This survey study demonstrated the public acceptability of sharing commercial data for health research in the United Kingdom with an extensive exploration of people's knowledge and understanding of what constitutes personal data and GDPR, their willingness to share with different organizations, their willingness to share various types of commercial data, and their willingness to consent and share data if invited through different methods. The outcomes of this study are of interest to be considered in the guidelines and recommendations for public acceptability of data sharing beyond electronic health records and will be useful for developing data stewardship frameworks and initiatives to improve the use of data in the United Kingdom. Where possible, these outcomes can also be used to develop recruitment strategies for research using stratified sampling techniques where it is expected to have low response rates. Future studies using experimental methods are warranted to identify the effectiveness of behavioral science techniques and communication methods to improve the public acceptability of sharing commercial data for health research.

Acknowledgments

This study was funded by Cancer Research UK Early Detection and Diagnosis project grant (C38463/A26726), with support from the Peter Sowerby Foundation. YH and HRB were funded by Cancer Research UK (C38463/A26726). The authors would like to acknowledge Wenjia Wang, MSc, who worked on this project as part of her MSc Health Psychology dissertation project at University College London.

Data Availability

The anonymized data collected as part of this survey will be made available to researchers upon publication in the Open Science Framework.

Authors' Contributions

YH developed the study concept. YH, LT, and MMR designed and developed the questionnaire. STS and YH performed the data analysis, and YH, LT, MMR, JMF, and HRB informed the interpretation. YH drafted the manuscript, and all authors provided critical revisions. All authors approved the final version of the manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

(1) The results of the factor analysis; (2) a detailed description of the survey measures including the original questionnaire with item heritage; (3) descriptive data tables for the General Data Protection Regulation awareness in the United Kingdom; (4) ordered logistic regression results on willingness to share commercial data based on invitation sources; and (5) full ordered regression model tables for the primary outcomes.

[DOCX File , 72 KB - publichealth_v9i1e40814_app1.docx]

References

1. Mavragani A. Infodemiology and infoveillance: scoping review. J Med Internet Res 2020 Apr 28;22(4):e16206 [FREE Full text] [doi: 10.2196/16206] [Medline: 32310818]
2. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. Per Med 2018 Sep;15(5):429-448 [FREE Full text] [doi: 10.2217/pme-2018-0044] [Medline: 30259801]
3. Ru B, Yao L. A literature review of social media-based data mining for health outcomes research. Social Web Health Res 2019:1-4. [doi: 10.1007/978-3-030-14714-3_1]
4. Aitken M, de St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Med Ethics 2016 Nov 10;17(1):73 [FREE Full text] [doi: 10.1186/s12910-016-0153-x] [Medline: 27832780]
5. Davidson S, McLean C, Treanor S, Aitken M, Cunningham-Burley S, Laurie G, et al. Public Acceptability of Data Sharing Between the Public, Private and Third Sectors for Research Purposes. Edinburgh: Scottish Government Social Research; 2013.
6. Keusch F, Struminskaya B, Antoun C, Couper MP, Kreuter F. Willingness to participate in passive mobile data collection. Public Opin Q 2019 Jul;83(Suppl 1):210-235 [FREE Full text] [doi: 10.1093/poq/nfz007] [Medline: 31337924]
7. Paprica PA, McGrail KM, Schull MJ. Notches on the dial: a call to action to develop plain language communication with the public about users and uses of health data. Int J Popul Data Sci 2019 Aug 05;4(1):1106 [FREE Full text] [doi: 10.23889/ijpds.v4i1.1106] [Medline: 32935032]
8. Ghafur S, Van Dael J, Leis M, Darzi A, Sheikh A. Public perceptions on data sharing: comparing attitudes in the US and UK. SSRN J 2020. [doi: 10.2139/ssrn.3586711]
9. Stockdale J, Cassell J, Ford E. "Giving something back": a systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. Wellcome Open Res 2018 Jan 17;3:6 [FREE Full text] [doi: 10.12688/wellcomeopenres.13531.2] [Medline: 30854470]
10. Aitken M, Tully M, Porteous C, Denegri S, Cunningham-Burley S, Banner N, et al. Consensus statement on public involvement and engagement with data intensive health research. Int J Popul Data Sci 2019 Feb 12;4(1):586 [FREE Full text] [doi: 10.23889/ijpds.v4i1.586] [Medline: 34095528]
11. Participatory data stewardship A framework for involving people in the use of data. Ada Lovelace Institute. URL: https://www.adalovelaceinstitute.org/report/participatory-data-stewardship/ [accessed 2023-01-17]
12. Peppin A. Who cares what the public think? The Ada Lovelace Institute. 2022 May 5. URL: https://www.adalovelaceinstitute.org/evidence-review/public-attitudes-data-regulation/ [accessed 2023-01-17]
13. Hoofnagle CJ, van der Sloot B, Borgesius FZ. The European Union general data protection regulation: what it is and what it means. Inform Commun Technol Law 2019 Feb 10;28(1):65-98. [doi: 10.1080/13600834.2019.1573501]
14. Green MA, Watson AW, Brunstrom JM, Corfe BM, Johnstone AM, Williams EA, et al. Comparing supermarket loyalty card data with traditional diet survey data for understanding how protein is purchased and consumed in older adults for the UK, 2014-16. Nutr J 2020 Aug 13;19(1):83 [FREE Full text] [doi: 10.1186/s12937-020-00602-3] [Medline: 32791968]
15. Flanagan JM, Skrobanski H, Shi X, Hirst Y. Self-care behaviors of ovarian cancer patients before their diagnosis: proof-of-concept study. JMIR Cancer 2019 Jan 17;5(1):e10447 [FREE Full text] [doi: 10.2196/10447] [Medline: 30664464]
16. Brewer HR, Hirst Y, Sundar S, Chadeau-Hyam M, Flanagan JM. Cancer Loyalty Card Study (CLOCS): protocol for an observational case-control study focusing on the patient interval in ovarian cancer diagnosis. BMJ Open 2020 Sep 08;10(9):e037459 [FREE Full text] [doi: 10.1136/bmjopen-2020-037459] [Medline: 32900761]
17. Huang F, Blaschke S, Lucas H. Beyond pilotitis: taking digital health interventions to the national level in China and Uganda. Global Health 2017 Jul 31;13(1):49 [FREE Full text] [doi: 10.1186/s12992-017-0275-z] [Medline: 28756767]
18. Dolan EH, Shiells K, Goulding J, Skatova A. Public attitudes towards sharing loyalty card data for academic health research: a qualitative study. BMC Med Ethics 2022 Jun 07;23(1):58 [FREE Full text] [doi: 10.1186/s12910-022-00795-8] [Medline: 35672737]
19. Shiells K, Di Cara N, Skatova A, Davis O, Haworth C, Skinner A, et al. Participant acceptability of digital footprint data collection strategies: an exemplar approach to participant engagement and involvement in the ALSPAC birth cohort study. IJPDS Special Issue Public Involve Engage 2020;5(3). [doi: 10.23889/ijpds.v5i3.1728]
20. Wenz A, Jäckle A, Couper M. Willingness to use mobile technologies for data collection in a probability household panel. Survey Res Method 2019;13(1):1-22 [FREE Full text]
21. Struminskaya B, Toepoel V, Lugtig P, Haan M, Luiten A, Schouten B. Understanding willingness to share smartphone-sensor data. Public Opin Q 2020;84(3):725-759 [FREE Full text] [doi: 10.1093/poq/nfaa044] [Medline: 34025296]

22. Ságvári B, Gulyás A, Koltai J. Attitudes towards participation in a passive data collection experiment. Sensors (Basel) 2021 Sep 10;21(18):6085 [FREE Full text] [doi: 10.3390/s21186085] [Medline: 34577291]
23. Keusch F, Bähr S, Haas G, Kreuter F, Trappmann M, Eckman S. Non‐participation in smartphone data collection using research apps. Royal Stats Soc Series A 2022 Apr 12;185(S2) [FREE Full text] [doi: 10.1111/rssa.12827]
24. Overview – Data Protection and the EU. Information Commissioner's Office. URL: https://ico.org.uk/for-organisations/dp-at-the-end-of-the-transition-period/overview-data-protection-and-the-eu/ [accessed 2023-01-17]
25. Osborne JW, Costello AB. Sample size and subject to item ratio in principal components analysis. Practical Assess Res Eval 2019;9(11) [FREE Full text] [doi: 10.7275/ktzq-jq66]
26. Crofts S. Population statistics research update: June 2018. Office of National Statistics. 2018 Jun 22. URL: https://tinyurl.com/2pc6fdej [accessed 2023-01-17]
27. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457 [FREE Full text] [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
28. Stoffel ST, Hirst Y, Ghanouni A, McGregor LM, Kerrison R, Verstraete W, et al. Testing active choice for screening practitioner's gender in endoscopy among disinclined women: an online experiment. J Med Screen 2019 Jun 14;26(2):98-103 [FREE Full text] [doi: 10.1177/0969141318806322] [Medline: 30428750]
29. Nissenbaum H. Privacy as contextual integrity. Wash Law Rev 2004;79(1).
30. Norberg P, Horne DR, Horne. The privacy paradox: personal information disclosure intentions versus behaviors. J Consum Affair 2007;41(1):100-126 [FREE Full text] [doi: 10.1111/j.1745-6606.2006.00070.x]
31. Myers KJ, Jaffe T, Kanda DA, Pankratz VS, Tawfik B, Wu E, et al. Reaching the "Hard-to-reach" sexual and gender diverse communities for population-based research in cancer prevention and control: methods for online survey data collection and management. Front Oncol 2022;12:841951 [FREE Full text] [doi: 10.3389/fonc.2022.841951] [Medline: 35756657]
32. Russomanno J, Patterson JG, Jabson Tree JM. Social media recruitment of marginalized, hard-to-reach populations: development of recruitment and monitoring guidelines. JMIR Public Health Surveill 2019 Dec 02;5(4):e14886 [FREE Full text] [doi: 10.2196/14886] [Medline: 31789598]
33. Skatova A, Goulding J. Psychology of personal data donation. PLoS One 2019 Nov 20;14(11):e0224240 [FREE Full text] [doi: 10.1371/journal.pone.0224240] [Medline: 31747408]
34. West R, Michie S. A brief introduction to the COM-B Model of behaviour and the PRIME Theory of motivation. Qeios 2020 Apr 07. [doi: 10.32388/ww04e6]
35. Chorley AJ, Hirst Y, Vrinten C, von Wagner C, Wardle J, Waller J. Public understanding of the purpose of cancer screening: a population-based survey. J Med Screen 2018 Jun 22;25(2):64-69 [FREE Full text] [doi: 10.1177/0969141317699440] [Medline: 28530514]
36. Connor K, Hudson B, Power E. Awareness of the signs, symptoms, and risk factors of cancer and the barriers to seeking help in the UK: comparison of survey data collected online and face-to-face. JMIR Cancer 2020 Jan 17;6(1):e14539 [FREE Full text] [doi: 10.2196/14539] [Medline: 31951219]

Abbreviations

aOR: adjusted odds ratio
GDPR: General Data Protection Regulation
PCA: principal component analysis

Edited by A Mavragani, T Sanchez; submitted 07.07.22; peer-reviewed by H Huang, R Bach; comments to author 18.10.22; revised version received 16.01.23; accepted 19.01.23; published 23.03.23.

Please cite as:
Hirst Y, Stoffel ST, Brewer HR, Timotijevic L, Raats MM, Flanagan JM
Understanding Public Attitudes and Willingness to Share Commercial Data for Health Research: Survey Study in the United Kingdom
JMIR Public Health Surveill 2023;9:e40814
URL: https://publichealth.jmir.org/2023/1/e40814
doi:10.2196/40814
PMID:36951929

Original Paper

Evidencing the Impact of Web-Based Coproduction With Youth on Mental Health Research: Qualitative Findings From the MindKind Study

Blossom Fernandes[1*], PhD; Lakshmi Neelakantan[2*], PhD; Himani Shah[3], MA; Sushmita Sumant[3], MA; Pamela Y Collins[4,5], MPH, MD; Jennifer Velloza[5,6], MPH, PhD; Emily Bampton[1]; Swetha Ranganathan[3], MSc; Refiloe Sibisi[7], BA; Toiba Bashir[8]; Joshua Bowes[9]; Esther Larisa David[8]; Harsimar Kaur[8], MBBS; Umairah Malik[9], MSc; Issy Shannon[9]; Suvlaxmi Gurumayum[8], MA; Anne-Marie Burn[10], PhD; Solveig K Sieberts[11], PhD; The MindKind Consortium[11]; Mina Fazel[1], DM

[1]Department of Psychiatry, University of Oxford, Oxford, United Kingdom
[2]Centre for Mental Health, School of Population and Global Health, University of Melbourne, Melbourne, Australia
[3]Centre for Mental Health Law & Policy, Indian Law Society, Pune, India
[4]Department of Psychiatry & Behavioral Sciences, University of Washington, Seattle, WA, United States
[5]Department of Global Health, University of Washington, Seattle, WA, United States
[6]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, United States
[7]Higher Health, Higher Education and Training: Health, Wellness, and Development Centre, Pretoria, South Africa
[8]Mindkind Young People's Advisory Group, Centre for Mental Health Law & Policy, Indian Law Society, Pune, India
[9]Mindkind Young People's Advisory Group, Department of Psychiatry, University of Oxford, Oxford, United Kingdom
[10]Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom
[11]Sage Bionetworks, Seattle, WA, United States
[*]these authors contributed equally

Corresponding Author:
Blossom Fernandes, PhD
Department of Psychiatry
University of Oxford
Warneford Hospital
Oxford, OX3 7JX
United Kingdom
Phone: 44 7985089247
Email: blossom.fernandes@psych.ox.ac.uk

Abstract

Background: Public involvement in research is a growing phenomenon as well as a condition of research funding, and it is often referred to as coproduction. Coproduction involves stakeholder contributions at every stage of research, but different processes exist. However, the impact of coproduction on research is not well understood. Web-based young people's advisory groups (YPAGs) were established as part of the MindKind study at 3 sites (India, South Africa, and the United Kingdom) to coproduce the wider research study. Each group site, led by a professional youth advisor, conducted all youth coproduction activities collaboratively with other research staff.

Objective: This study aimed to evaluate the impact of youth coproduction in the MindKind study.

Methods: To measure the impact of web-based youth coproduction on all stakeholders, the following methods were used: analysis of project documents, capturing the views of stakeholders using the *Most Significant Change* technique, and impact frameworks to assess the impact of youth coproduction on specific stakeholder outcomes. Data were analyzed in collaboration with researchers, advisors, and YPAG members to explore the impact of youth coproduction on research.

Results: The impact was recorded on 5 levels. First, at the paradigmatic level, a novel method of conducting research allowed for a widely diverse group of YPAG representations, influencing study priorities, conceptualization, and design. Second, at the infrastructural level, the YPAG and youth advisors meaningfully contributed to the dissemination of materials; infrastructural constraints of undertaking coproduction were also identified. Third, at the organizational level, coproduction necessitated implementing new communication practices, such as a web-based shared platform. This meant that materials were easily accessible to the whole team and communication streams remained consistent. Fourth, at the group level, authentic relationships developed between the YPAG members, advisors, and the rest of the team, facilitated by regular web-based contact. Finally, at the individual level, participants reported enhanced insights into mental well-being and appreciation for the opportunity to engage in research.

Conclusions: This study revealed several factors that shape the creation of web-based coproduction, with clear positive outcomes for advisors, YPAG members, researchers, and other project staff. However, several challenges of coproduced research were also encountered in multiple contexts and amid pressing timelines. For systematic reporting of the impact of youth coproduction, we propose that monitoring, evaluation, and learning systems be designed and implemented early.

(JMIR Public Health Surveill 2023;9:e42963) doi:10.2196/42963

KEYWORDS

web-based youth coproduction; mental health; public involvement; young people; advisory groups

Introduction

Background

Increasingly, participatory and collaborative research methods have been used to reduce the gap between evidence and implementation [1-3]. One way to do this is to involve target groups in research, using coproduction processes, particularly the involvement of youth, to reflect their needs in research [4,5]. Coproduction is “an approach in which researchers, practitioners and the public work together, sharing power and responsibility from the start to the end of the project, including the generation of knowledge” [6]. Coproduction aims to ensure that the knowledge generated is informed by the needs of the target group, is relevant to them, and has greater applicative value in their local setting [7]. This is somewhat different from what Baum et al [8] describe as participatory action research, where participants are involved in a reflective manner, by collecting and analyzing the data and then determining the outcome. Thus, the active involvement of knowledge users (eg, young people with mental health difficulties) in research, who are experts in their experience, aims to enhance the quality and utility of research.

Coproduction in Mental Health Research With Youth

Different forms of mental health researcher and knowledge user partnerships have grown in popularity, primarily in high-income settings, particularly among young people [9]. With most lifelong mental health conditions having an onset by the mid-20s [10], there is a vital focus on better understanding youth mental health needs and curating services tailored to their needs. Several benefits of such youth-researcher collaborations have been identified. For example, coproduction with youth increases the relevance of research aims [11] and facilitation of recruitment [12,13] and helps the production of richer and more reliable data [14]. Researcher-intensive tasks such as data analysis, the presentation of findings, and dissemination also benefit from youth involvement [15]. Coproduction also enables incorporating lived experiences effectively; with regard to user outcomes, coproduction helps young people develop transferable skills [9,16] and research skills [11] and to understand their own mental health better [17].

Complexities and Challenges of Undertaking Coproduction With Young People

Though coproduction of knowledge may seem like an elegant solution to ensure the implementation of evidence-based practices, it is a highly complex, context-specific experimental process that is often time-consuming and resource intensive [18]. Effective coproduction requires collaborative effort at every stage of the research process [19]. The environment is important, as are researcher traits, such as openness, tolerance, and flexibility [20], and organizational qualities, such as building trustworthy relationships and innovative methodologies [21,22].

The complex process of coproduction with youth comes with specific challenges, as complicated power dynamics may need to be identified, acknowledged, and addressed. Researchers may need to address youth-adult hierarchies [23], which may disrupt effective collaborative work. If such traditional patterns are not disrupted, tokenism, difficulties in conflict resolution and shared decision-making between youth and researchers can ensue [24]. Research processes that ensure coproduction with youth require additional time and resources owing to a number of cultural, ethical, and safety considerations, such as assent-consent procedures for legal minors [25] and ensuring safe practices while also recognizing their agency [5].

Measuring the Impact of Coproduction

Understanding and evidencing the impact of coproduction with youth remains an area in need of further development. The current literature on the monitoring and evaluation of young people’s involvement in research is diffuse and dispersed and does not have a standard taxonomy or methodology [26]. Moreover, much of the focus has been on “measurable, economic and quantifiable impacts,” which does not take into account the numerous dynamic processes of coproduction [27]. The diversity of approaches, evaluative frameworks, and depth of discourse have made it difficult to measure the impact of coproduction on research processes and outcomes.

Few studies report the impacts of youth coproduction [28], complex processes of collaboration, and nonlinear impacts for stakeholders [29]. The MindKind study used mixed methods to elicit young people’s views on data governance and to assess the feasibility of setting up a user-controlled global mental health databank [30]. This study was conducted in India, South Africa, and the United Kingdom. Participants aged 18 to 24 years (16 to 24 years in the United Kingdom) were enrolled on a separate website and randomized into 1 of the 4 data governance options [30]. Using an app designed for the study, participants were invited to share data on key active ingredients of mental health, such as sleep, body movement, social connections, and positive experiences over a 12-week period. Study participants in the qualitative arm took part in country-specific and multinational group deliberative democracy sessions [31] focused on building a consensus pertaining to the governance of a future mental health databank. Each arm of the study had a strong focus on coproduction.

Given the gaps that exist in the literature, in this study, we aimed to evaluate the impact of coproduction in the MindKind study, and the processes involved in achieving such an impact. We defined the impact of coproduction as a powerful or major influence on 4 stakeholder groups: youth (professional youth advisors and young people's advisory group [YPAG] members), researchers (staff who contributed to research design, data collection, and analysis), decision makers (principal investigators and funders), and support and administrative staff (broadly defined to include anyone who undertook substantial administrative and support responsibilities as part of their role). We used the impact framework proposed by Beckett et al [27], which applies a multifaceted lens to understand the breadth of coproduction at different levels, the interactions between these levels, and emergent mechanisms. These levels include paradigmatic, structural, organizational, group, and individual levels. Paradigmatic impacts are those impacts of coproduction with the potential to modify ways of understanding and shift frames of reference. Infrastructural impacts include broader social, economic, policy, and political impacts. Organizational impacts include rules, norms (culture), practices, and organizational structures. Group impacts include interpersonal and stakeholder relationships, whereas individual impacts include personal changes such as improved mental or physical health and improved practice and skills. The emphasis within this framework is to "consider longer term developments, wider social changes, any unintended consequences and how coproduced research might affect and be affected by different power dynamics" [27].

We note that the findings discussed in this manuscript do not include any data from the main MindKind study or any of the randomized arms. The objective of this manuscript was to describe the impact of coproduction with youth on different stakeholder groups [32]. Coproduction with youth was designed to influence both the qualitative and quantitative study arms (ie, youth were coresearchers), and we evaluated such impact of youth acting as coresearchers on key stakeholder groups. Advisors in the project did not participate in the qualitative or quantitative study conducted as part of the MindKind study.

Methods

Ethics Approval

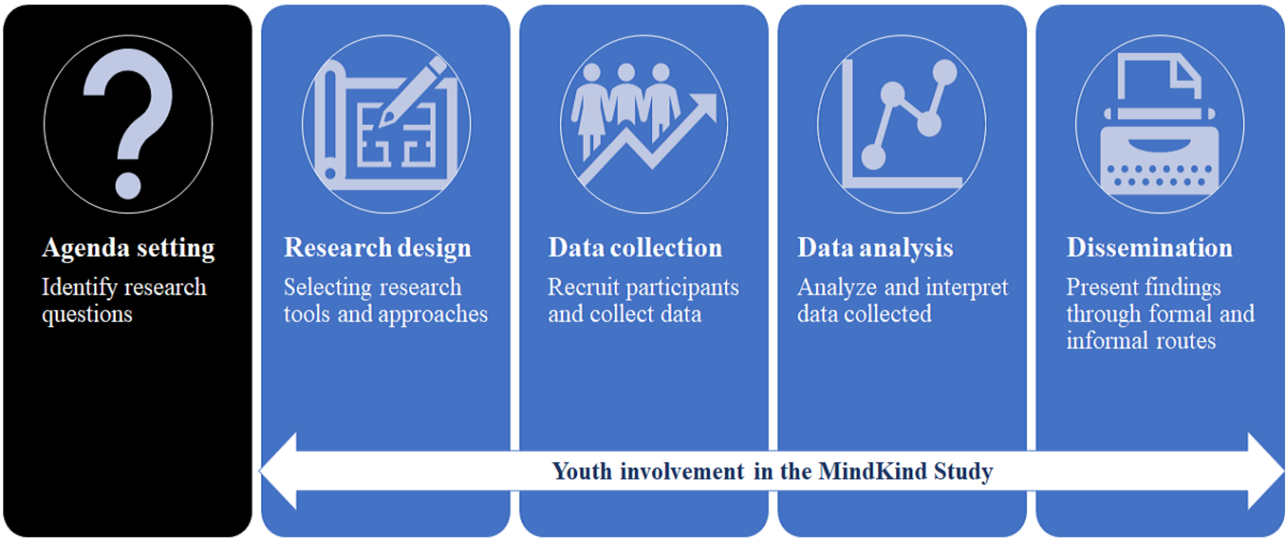
Ethics approval for the MindKind study was obtained from each of the country sites, namely the United States (WIRB #20212067), the United Kingdom (University of Cambridge, Department of Psychology Research Ethics Committee: Ref. PRE.2021.031 and University of Oxford: Ref R73366/RE00), South Africa (Walter Sisulu University #029/2021 and the Department of Higher Education and Training), and India (India Law Society #ILS/242/2021 and Health Ministry Screening Committee).

Overview of Youth Coproduction

Each MindKind study site employed a full-time youth lead or professional youth advisor (youth advisor), to convene and run a YPAG in each country site (India, South Africa, and the United Kingdom) and advise on all aspects of the study. The youth advisor was a young person aged 18-23 years who had lived experience of mental health challenges. The in-country YPAGs in each site comprised of young people with lived experience of mental health challenges aged 18-24 years (16-24 years in the United Kingdom), known as YPAG members. In addition to the in-country YPAGs, around 3 to 4 YPAG members from each site joined an international YPAG to advise on broader study questions.

Each youth advisor had 1 vote on the study's Steering Committee and, therefore, actively participated in study decision-making. Sessions conducted with the in-country and international YPAGs were planned collaboratively by the research teams and led by the respective youth advisors. Sessions were recorded, and feedback shared by the YPAGs was reported by the youth advisors to the research teams via a web-based platform called Airtable. Researchers were then required to respond in Airtable to clarify whether the feedback had been actioned, and if not, why. Youth involvement in MindKind is described in Figure 1.

Figure 1. Youth involvement in the MindKind study.



Data Sources

Coproduction with youth occurred during the study period between September 2020 and July 2022. Data collection to evaluate the impact of coproduction was conducted sequentially in 2 stages with 4 sets of stakeholder groups: youth (advisors and YPAG members), researchers (staff who contributed to research design, data collection, and analysis), decision makers (principal investigators and funders), and support and administrative staff (defined broadly to include anyone who undertook substantial administrative and support responsibilities as part of their role).

First, we collated various project documents generated between September 2020 and January 2022, such as meeting recordings and minutes, project descriptions, and work plans, to determine the coproduction outcomes relevant to all stakeholder groups (youth, researchers, decision makers, and administrative staff). A total of 17 documents were downloaded from the project database for analysis. A summary of these data sources is presented in Textbox 1. The outcomes identified at this stage allowed us to refine subsequent data collection on significant changes for the 4 stakeholder groups.

Second, building on the identified outcomes, we generated data based on the *Most Significant Change* (MSC) technique on the main changes that resulted from coproduction with youth for the 4 stakeholder groups, that is, moving from the outcomes that were intended to take place to the impacts that actually took place [33]. This technique involved gathering stories of the MSCs from stakeholders via open questions to distill key learning about the study and recommendations for the future. In this study, MSC technique meant gathering views from the 4 stakeholder groups on what had changed or had been impacted as a result of youth coproduction. Researchers and youth advisors implemented the MSC technique via web-based in-depth group discussions with in-country and international YPAGs. In these group discussions with YPAG members, the youth advisor in that study site and the lead authors gathered youth views on what impacts they felt had taken place as a result of youth coproduction.

For other stakeholders (ie, researchers, decision makers, and administrative staff), an open-ended web-based questionnaire solicited views on significant changes across 4 to 5 specified outcomes for each stakeholder group. This web-based questionnaire was emailed to all stakeholders in the study and completed by 16 (52%) out of 31 participants. Although the final sample of respondents was not very large, it involved representation from each of the stakeholder groups and therefore, represented a wide cross section of views. All questionnaire responses were deidentified before analysis.

Textbox 1. Summary of documentary data sources.

Initial project discussion notes

- Youth panel decisions between August and November 2020
- Global Mental Health Databank (initial name of MindKind) feedback documents
- Global Mental Health Databank project structure and governance
- Youth panel engagement
- Youth panel objectives

Steering committee notes

- Agenda and meeting minutes
- Presentation on project goals
- Presentation on data usability group feedback and results

Global youth panel meetings

- Meeting notes
- Presentation

Youth advisor onboarding materials

- Project governance plan
- Proposal for youth involvement
- Youth panel concept note

Youth advisory planning spreadsheet

- Youth panel deliverable tracker

Youth advisory planning presentation

- Airtable feedback loop

Funder materials on youth mental health

- Mental health priority log

Data Analysis

The analysis focused on building a picture of key changes that occurred for the 4 stakeholder groups as a result of implementing youth coproduction in the MindKind study and understanding how stakeholders perceived its functioning and progress. We used a qualitative approach to collate and synthesize the collected data. Analysis took place at 2 time points: from December to February 2022 and from March to May 2022.

The first stage consisted of a document analysis performed by on of the lead authors (LN). Outcomes aimed to be achieved due to youth coproduction were derived from study documents. These outcomes were then mapped to the impact framework by Beckett et al [27] (Textbox 2). Following this, inputs were gathered from youth advisors, researchers, and YPAG members in each site, and outcomes were revised as a result of such inputs.

The second stage consisted of analyzing the MSC data to identify the MSCs identified by the 4 stakeholder groups that took place as a result of youth coproduction. This analysis was performed collaboratively by 3 researchers, 3 youth advisors, and 7 YPAG members from the United Kingdom and India, who expressed their interest in being involved in data analysis. Web-based sessions with the YPAGs were transcribed, and open-ended survey responses were summarized. These transcripts and summaries were then scrutinized for impacts on the 4 stakeholder groups, which were categorized under the impact framework [27].

In the final stage of analysis, MSC data were coanalyzed by all analysts using an iterative voting approach to increase the reliability of the findings [33]. This involved all analysts discussing the key impacts, identifying the most impactful quotations shared by participants from each of the 4 stakeholder groups, discussing reasons for their choices, and then voting again. A preliminary version of the findings was developed and revised after discussion with advisors and YPAG members.

Textbox 2. Outcomes of coproduction intended to take place, derived from study documents.

Outcomes at the paradigmatic level

- Better conceptualization of study questions, measures, and design
- Reorientating common research practices and assumptions

Outcomes at the infrastructural level

- Research that values youth voices
- Exploration of youth preferences and attitudes on a range of topics pertinent to data governance and mental health
- Youth-informed dissemination strategies
- Empowerment of young people as data users

Outcomes at the organizational level

- Established communication channels between all stakeholders
- Accountability mechanisms for receipt of feedback
- Increased connection between youth and research institutions
- Differentiated information outputs based on different information needs of stakeholders
- Time and effort invested in team building to fully integrate youth

Outcomes at the group level

- Engaging and authentic relationships between stakeholders, especially between youth and other stakeholders
- Youth-driven rules of engagement for advisors and young people advisory group members

Outcomes at the individual level

- Increased knowledge and capacity of research concepts and processes (for youth) and contexts of youth lives (for other stakeholder groups)
- Adequate compensation for time and effort; networking and skill development opportunities; insights into own mental health experiences; feeling heard (for youth)

Results

Overview

A total of 31 stakeholders working across all MindKind study sites, belonging to the 4 stakeholder groups, participated in group discussions and questionnaires: 11 youths (3 advisors; 7 members across YPAGs in India, South Africa, the United Kingdom and an international YPAG; and 1 lived experience consultant at the funding organization); 11 researchers in India, South Africa, the United Kingdom, and the United States; 6 decision makers (individuals responsible for key study decisions, such as principal investigators and youth); and 3 support and administrative staff. All participants were employed in the MindKind study and had <1 to 25 years of health research experience. Where stakeholders belonged to more than 1 group, the primary group affiliation was used for the analysis (eg, a professional youth advisor could belong to the youth, researcher, and decision maker stakeholder groups). The participants were based in both high- and middle-income settings.

We found that impact occurred at 5 levels detailed in the framework developed by Beckett et al [27], namely, the paradigmatic, infrastructural, organizational, group, and individual levels. Detailed findings are available in Multimedia Appendix 1.

Paradigmatic Impact: Big-Picture Learnings for Coproduction

Paradigmatic impacts are those impacts of coproduction with the potential to modify ways of understanding and shift frames of reference [27]. We identified 2 paradigmatic impacts of coproduction in the MindKind study: first, coproduction enabled new and unexpected ways of conducting research, and second, coproduction modified research priorities and influenced study conceptualization and design, although this was not always successfully implemented.

With regard to the first impact of new ways of conducting research, we observed that youth advisors prompted researchers and decision makers to engage with diverse groups of youth to advise as part of YPAGs and use innovative methods to recruit study participants (eg, using Instagram stories and posts to reach young people). For YPAGs, advisors ensured that recruited members represented diverse backgrounds in research experience, ethnicity, geographic location, gender, or disability status. Previous research shows that YPAG recruitment is limited in diversity and geographic reach (eg, more advisors who are women with prior research experience, based locally) [11]. However, improved digital and technological flexibility in web-based coproduction enabled greater diversity.

Key impacts of youth coproduction were the introduction of Airtable, weekly youth-focused meetings, and relationship building among individual site teams. These immediate impacts then helped improve the process of coproduction by creating separate spaces for youth to share feedback, enabling practices for all stakeholders to engage with one another, and setting up multiple avenues for examining how coproduction goals were being met. Accountability mechanisms are also important for tracking the implementation of generated ideas:

> *We have to create spaces to explicitly solicit young people's feedback; just "having them in the room" may be too overwhelming an environment to engage.* [Researcher]

Second, coproduction with youth transformed research priorities and activities and influenced study conceptualization and design. Through coproduction, we identified points of dissonance between different stakeholder groups and used them to improve study activities, key research decisions, and research knowledge generated to align more with youth priorities, life experiences, and contexts. One example of this was to introduce more capacity building activities for advisors and YPAGs as a priority to effectively advise and participate in research implementation. Without such capacity building activities, youth stakeholders can find it difficult to provide timely and effective advice, and researchers and decision makers find it challenging to incorporate youth perspectives in research. A shared mission and common goals were important for guiding the overall direction of coproduction.

YPAG members and advisors advised on the structure and content of the MindKind app, such as preferred app engagement features, which were a core component of the quantitative study. They also reviewed and shared feedback on questions designed to be used in qualitative group discussions with the study participants. However, some of their feedback on the MindKind app could not be successfully incorporated into the app development process because of the research and funding timelines. Timing was critical, as a number of early decisions had to be made (owing to logistical and funding constraints) before the coproduction systems had been set up. This highlights the importance of timing and early engagement; otherwise, there was a risk of disengagement among youth stakeholders:

> *If one is going to be doing research around youth, then involve youth from the beginning right through to the end. I think MindKind didn't quite do this as we had already developed the App and the areas to focus on before we started the youth advisory groups...we did a bit of "roughshodding" and hoped they would agree with us!* [Decision maker]

Capacity building activities for youth to advise on research implementation and dissemination needed significant additional time and resources, especially from advisors and early career researchers, which were exacerbated by restricted funding timelines. These constraints meant that there were contexts in which coproduction was limited, and the study had to consider meaningful ways of course correction or acknowledge which parts of the study could not be meaningfully coproduced. However, systems for reflection in multiple teams, forums, and different groups were critical in ensuring that meaningful engagement was always possible, and attempted, even if not successful in every instance:

> *Many discussions around involvement have revolved around equality. However, I thought it was important to emphasise equity. This meant taking a unique approach which focused on youth capacity development and ongoing reflexivity rather than assuming involvement just meant a seat at the table.* [Youth stakeholder]

Infrastructural Impact: Constraints and Opportunities in Existing Systems

We describe infrastructural impacts of coproduction on dissemination activities, as well as infrastructural constraints within research, administrative, and funding structures that governed the larger MindKind study.

Coproduction had positive impacts on research dissemination for both formal (the MindKind final report, journal publications, and conference presentations) and informal (blog posts, seminar discussions, and internal presentations) outputs. For example, we coanalyzed and cowrote our youth coproduction publications with advisors and YPAG members, which introduced views and perspectives that might not have otherwise been incorporated. Advisors at each site (India, South Africa, and the United Kingdom) shared their views on the topics to be addressed in publications arising from the study. They also contributed to preparing an outline for each publication, participated in documentary analysis, and shared key learnings from youth involvement via blog posts. Advisors and YPAG members also coanalyzed qualitative data by assisting with theme and subtheme selection, voting on the most important findings to be highlighted, and commenting on drafts of the manuscript.

We were regularly invited to share our learnings with external research groups who had either begun or were contemplating undertaking coproduction in the form of informal discussions or extended presentations. We presented the challenges and opportunities of coproduction in posters at academic conferences, internal seminars, and university meetings (eg, within the Child and Adolescent Psychiatry research group at the Department of Psychiatry, University of Oxford). One such presentation at an internal seminar was led entirely by the advisor and YPAG members. The presence of a full-time advisor was a positive catalyst for learning in other research projects.

Factors that facilitated positive infrastructural impact included greater funding and administrative support to ensure that these tensions were addressed appropriately. Flexible timelines and the timely involvement of youth were also critical in ensuring that youth stakeholders could be fully integrated and included in the project as equal partners in research. This also meant that research funding was critical to timing, as coproduction needed resources available at the funding application stage to be done appropriately.

Infrastructural impacts varied across contexts. For example, the youth advisor and YPAG members in South Africa frequently faced planned electricity outages or "load shedding," which hampered their ability to meet via web-based portals on a regular basis. YPAG members in South Africa also did not consistently have access to the internet, and their agreed remuneration was delayed owing to wider issues with the financial administration, both of which negatively impacted their ability to engage. In contrast, advisors in India and the United Kingdom faced fewer infrastructural issues, which aided regular engagement and advice from YPAG members. These findings underscore the importance of infrastructure in a wider sense, including access to the internet and responsive administrative systems that were essential for facilitating effective coproduction.

An additional infrastructural impact within this study was that we experienced a constant tension between project timelines, deliverables, and coproduction aims. To undertake coproduction with youth fully, we needed to invest significant time and effort on capacity building activities, tailoring existing systems, and mentorship, but these were not always compatible with the study timelines. Pressure to produce reports or updates in line with project milestones were also often at odds with the time that was needed to effectively coproduce a specific aspect of the study, whether that was about the design, data collection, or analysis.

Organizational Impact: Differentiated Communication Practices and the Feedback Loop

Organizational impacts include impacts on rules, norms (culture), practices, and organizational structures [27]. We found that the most critical organizational impact involved changes in communication. Coproduction prompted organizations to engage appropriately with diverse stakeholder groups, including (1) ensuring that a feedback loop was in place to communicate with stakeholders, especially youth, and that this loop was closed in a way that stakeholders felt heard and (2) implementing differentiated communication practices for different stakeholders.

With regard to the feedback loop, we set up an Airtable system to record feedback from YPAG members and advisors. Research and project teams then acted on the feedback. If implementing feedback was not possible, they explained why it could not be done. Although this was an encouraging start, we found that the critical component was ensuring that feedback loops needed to be closed in a way that the youth felt they had been addressed. Data collection on Airtable also had to be made more conversational to engage with youth, so formal data collection platforms were supplemented by meetings where youth could share their feedback verbally, in addition to written feedback:

> *Communication must be done continuously and through multiple channels. Because everyone learns and contributes differently.* [Decision maker]

Given that effective communication meant something different for each stakeholder group, our communication practices needed different frequencies and levels of detail as well as a more inclusive information-sharing strategy to include youth. This sometimes meant that given the large number of stakeholders to consult, decisions needed more time to be taken. We implemented other changes to address these issues, for example, a weekly digest email for important announcements; quick links to project documentation; and a list of the upcoming week's meetings, including any scheduled YPAGs. Importantly, the digest was designed to ensure that everyone, including the advisors, was informed about the decisions to be made, who was making them, and when and in what format feedback would be welcomed. These digests became important tools for increasing transparency and inclusivity, directly impacting communication between stakeholders.

Group Impact and Interpersonal Relationships: Authenticity and Cycles of Engagement

Group-level impacts constitute interpersonal and stakeholder relationships within a system [27]. We observed that due to coproduction, authentic relationships developed between researchers and advisors and between advisors and YPAG members, and YPAG members experienced cycles of engagement and disengagement.

Authentic relationships among the advisors were aided by creating multiple channels of communication with them and among them. The advisors and study team met every week during the course of the study to discuss updates, youth integration, challenges, and learnings. The youth advisors also met each other, together with an external lived experience consultant and advisor (employed by the funder), for regular check-in meetings. They also had a monthly scheduled check-in with each other.

Advisors were responsible for leading their own YPAG and were encouraged to foster authentic relationships with the members. They met the YPAG on a fortnightly basis and conducted individual check-ins with the members to understand any concerns or feedback from them. Overall, the project team treating the advisors and YPAG members with curiosity and professionalism was helpful in fostering these relationships.

Providing the YPAG members with opportunities to upskill and spending time and resources on engagement were important in creating positive group relationships. For example, actively soliciting youth feedback, trying innovative ideas, and offering multiple avenues for youth to engage and contribute to the study beyond the role of an advisor (involvement in capacity building activities, manuscript ideation, and writing) helped increase engagement. Furthermore, closing the feedback loop in a way felt by the youth was helpful, but when this was not possible, it could lead to disengagement if not properly communicated by researchers and decision makers:

> *Some key things to consider when engaging and retaining youth in the MindKind Study—Providing due credit to young people where necessary and allowing them to be key decision makers who hold equal power as adult staff on the team; Providing compensation/reimbursement/honorarium for the time that youth engage with us; Providing young people with constant opportunities to upskill (through workshops, manuscript involvement, consultancies on projects aside from MindKind); Ensuring that young people are given support when needed (emotional/support with work etc.)* [Youth]

Collaboration between researchers and advisors was established over time and was beneficial to overall youth engagement. This was done through multiple channels, such as organizing check-ins between researchers and advisors, as well as directly involving advisors in all research-related activities.

Individual Impact: Better Skills, Knowledge, and Capacities

According to Beckett et al [27], individual impact constitutes the characteristics of stakeholders, including biological and psychological aspects such as improved mental or physical health and improved practice and skills for practitioners. For advisors and YPAG members, major individual impacts included greater insight into one's own mental health; increased knowledge of research concepts and processes; opportunities to harness lived experiences of mental health challenges; remuneration for time and skills; and new academic, presentation, and study skills. The advisors and YPAG members reported developing better insights into their own mental health by participating in project activities and having discussions about a future global mental health databank. Furthermore, they reported increased skills and knowledge development in research-related concepts:

> *One of the most significant changes for me is probably the importance of research and studies which, which I mean, are now involved in but just in general, for the understanding of mental health...it's helped me look at mental health as more collective thing, rather than being more focused on the individual just because, like we, we convene, we talk on a weekly basis, but also the moving parts and the systems which have to be in place for some sort of change to happen. And some sort of, like, insight into mental health and research to actually come to fruition.* [Youth]

These impacts were mediated by the following factors: having site-specific and international YPAGs led by advisors, capacity building activities for advisors to independently coordinate and lead several study components, regular capacity building activities for YPAG members to advise on study, and dedicated spaces for youth to interact with each other. These capacity building activities involved dedicating sessions to demonstrate research in practice. Customized support for advisors to suit their backgrounds and interests was also critical for enhancing outcomes:

> *Young people's social and cultural contexts affect their knowledge and opinions on mental health. While this was something that I had some idea about prior to the study, having regular conversations with advisors and researchers across the sites led to interesting learnings on how social and cultural contexts play a role in shaping a young person's perspective and knowledge on mental health (for example, what mental health means to a young person in a HIC [high-income country] vs what it means to a young person in an LMIC [low- and middle-income country]).* [Youth]

For researchers and decision makers, enhanced coproduction with youth resulted in a much richer understanding of young people's life experiences and contexts:

> *I think we learned a lot about how youth wanted to be engaged...the decision making on the side of the research team needed to happen at a much faster pace than the youth were comfortable (or able) to make decisions. As a result, we sometimes just "told" rather than "asked." I think the youth forced us to slow down and demanded inclusion, which was a very powerful change to the study.* [Decision maker]

This was aided by ongoing reflexivity by the research team and decision makers as well as course correction based on youth feedback.

Discussion

Summary of Findings

Our findings suggest that coproduction with youth in the MindKind study had significant positive impacts for all stakeholder groups (youth, researchers, decision makers, and administrative and support staff). In addition, having a web-based YPAG allowed engagement from a diverse group of panel members, facilitating wider reach and convenience by conducting sessions at suitable times. We found that coproduction impacted, to varying degrees, all 5 domains of practice: paradigms (new or innovative ways of doing research and influence on research priorities, conceptualization, and design), infrastructure (improved dissemination of learning and constant tensions between study timelines and coproduction owing to infrastructural stressors), organizational (differentiated communication needs, practices, and outputs; greater emphasis on closing the feedback loop; and better access between youth and research institutions), group (more authentic relationships among youth and between youth and other stakeholders, and ebbs and flows in engagement), and individual (increased skills and capacities for all, opportunity to use lived experience of mental health for youth, and increased knowledge of youth life experiences for researchers and decision makers). Some impacts, such as societal or infrastructural impacts, occurring beyond the project can be challenging to measure and document but cannot be ignored.

Implications and Recommendations

Youth coproduction generally led to better outcomes overall, not only for youth but also for researchers, decision makers, administrative staff, and the research as a whole. Individual- and group-level impacts included increased knowledge, skills, and capacities for youth and other stakeholders; youth being able to use their lived experience of mental health and feeling heard; authentic relationships between the YPAG and advisors; and ebbs and flows in engagement over the project life cycle. Many of these findings align with the wider literature on the impact of coproduction with youth [4,28,34]. Having a dedicated youth advisor in each site was critical in achieving these impacts. There was no full-time youth advisor in the US study team (ie, not a study site, but the coordination hub); if present, this might have resulted in improved communication and greater youth input into some of the technological decisions.

Coproduction with youth consistently challenged nonyouth stakeholders on their assumptions of what the research could look like. This aligns with the wider literature, which emphasizes that the involvement of stakeholders in research is linked with greater effectiveness, if there is articulation of a shared mission and goals [35]. Our findings further confirm that although a shared mission is critical, establishing clear expectations with funders and organizations in terms of the practical costs of undertaking coproduction and implementing necessary changes in organizational communication practices are critical in achieving better outcomes for all.

Our experience of coproduction with youth as an active process of valuing all types of knowledge and experiences and creating a balance between different stakeholders aligns with others' experiences [36]. However, we encountered several barriers in implementing coproduction with youth according to its tenets and in line with our shared mission [37]. Our systems for reflection and discussion were critical in addressing these barriers effectively, which is a characteristic of effective partnerships [35]. Our findings point to the need for ongoing monitoring and evaluation of impact and to ensure that this is set up as early as practicable in the research cycle. Such monitoring and evaluation should take the form of first articulating outcomes and measurable indicators of what coproduction is meant to achieve, for whom, and by when. Establishing feedback forms and recurring surveys that capture periodic progress toward these goals is a good starting point.

Coproduction was enacted, implemented, and received differently by stakeholders in India, South Africa, and the United Kingdom (MindKind study sites). Equally, public and institutional infrastructure (eg, access to the internet) also played a significant role in the way coproduction was implemented and received in different contexts. Although such challenges broadly map to the literature on coproduction in low- and middle-income countries versus high-income countries [28], we found variations between South Africa and India in this context. Our findings underscore the concept that coproduction is highly "place based" and occurs in particular social, economic, and ecological contexts [38].

An important organizational impact detected was the change in communication practices and outputs needed to undertake coproduced research and ensuring that there was a feedback loop in place and stakeholders (especially youth) felt that this was closed in an appropriate way. This begins with the recruitment and integration of youth advisors, understanding their varied communication needs, and changing communication and decision-making practices. These findings are consistent with previous literature that highlights that effective coproduction is achieved when there are positive working relationships within and between teams, stakeholders feel heard, and tailored communication helps create shared meanings of concepts [35]. Our key mechanisms of achieving this impact included establishing feedback loops and testing them to ensure that they were fit for practice and supporting administrative staff to implement these communication practices. These practices are characterized as "maintenance tasks" in the wider literature that supports the functioning of partnerships by addressing core administrative and support needs [35].

The relatively recent completion of the study limits attempts to describe the wider infrastructural or societal impact as intended in the framework by Beckett et al [27]. Instead, we discuss modest impacts on dissemination and infrastructural constraints while undertaking coproduction with youth. However, to determine the wider infrastructural impact of youth involvement at the societal level, coproduction with youth would need to be assessed at the national scale, where such methods are adopted to measure its public value. Our experiences broadly map the practical costs and challenges of coproduction, such as large administrative burdens, increased researcher time and resources, the lack of training on the implementation of coproduction, and insufficient funds and resources to effectively undertake coproduced research [39]. Although we found significant benefits of undertaking coproduction, we also found that grappling with timelines and costs could lead to a diluted version of coproduction to *coproduction lite*. Despite the focus and funding support, our experience highlights the entrenched structural challenges in undertaking coproduced scientific research more broadly [40]. Our findings highlight the need to cocreate the infrastructure needed to undertake coproduction, resource it appropriately, and recognize the time and skills needed to coproduce research [39].

Youth Involvement

We were committed to ensuring the meaningful involvement of young people, but funding and timeline constraints meant that such involvement needed to be proportional and pragmatic. Therefore, we involved YPAG members in areas where they had particular knowledge, interests, or skills, specifically data collection, analysis, and the interpretation of findings. We analyzed the data in collaboration with advisors and YPAG members who were interested in doing so and jointly interpreted the findings and decided on specific findings to be presented in the paper. In relation to dissemination, we also met YPAG members and advisors who had participated in data analysis and interpretation to understand how they would like to be acknowledged in the paper (in this study, as coauthors).

Conclusions

Although the resources involved in coproduction are significant, they yield several benefits for various stakeholder groups and across several domains. We have attempted to capture these impacts and their enabling or disabling factors in further detail in the context of the MindKind study, which focused on young people's views of mental health data governance in India, South Africa, and the United Kingdom. We found that "effective coproduction emerges in practice" [41], and numerous opportunities and challenges arise in the implementation of coproduction.

There are no one-size-fits-all coproduction efforts. In sharing the impact of coproduction on various stakeholders and outcomes, we hope to have additionally demonstrated the value of measuring and reporting outcomes associated with youth coproduction [28]. We recommend that monitoring, evaluation, and learning systems be designed and implemented early in coproduction studies to enable more systematic reporting of coproduction with youth and its impacts. Future research should also examine how we can standardize the reporting of youth involvement in research; the Guidance for Reporting Involvement of Patients and the Public (GRIPP) checklist provide a useful starting point [42]. More frequent and detailed reporting of the impact of coproduction with youth will likely challenge the current funding and research infrastructure constraints and enable better health research outputs.

Acknowledgments

The authors would like to thank all members of the MindKind Youth Group Advisory Panels in India, South Africa, and the United Kingdom; the International Youth Panel; and the Global Youth Panel for their thoughtful and generous contributions to this project.

The MindKind Consortium includes the following members: Faith Oluwasemilore Adeyemi, Patricia A Areán, Emily Bampton, Elin A Björling, Ljubomir Bradic, Anne-Marie Burn, Emma Grace Carey, Sonia P Carlson, Pamela Y Collins, Tessa Concepcion, Meera Damji, Megan Doerr, Julia C Dunbar, Mina Fazel, Blossom Fernandes, Gillian Finchilescu, Tamsin Jane Ford, Melvyn Freeman, Isabell R Griffith Fillipo, Jay Hodgson, Jasmine Kalha, Minal Karani, Michael R Kellen, Christopher G Kemp, Simthembile Lindani, Lara M Mangravite, Carly Marten, Hedwick Masomera, Felicia Mata-Greve, Emily Moore, Erin Mounts, Lakshmi Neelakantan, Larsson Omberg, Lisa Pasquale, Soumitra Pathare, Swetha Ranganathan, Nichole Sams, Erin Joy Scanlan, Himani Shah, Sotirios Short, Refiloe Sibisi, Solveig K Sieberts, Stockard Simon, Sushmita Sumant, Christine Suver, Yanga Thungana, Meghasyam Tummalacherla, Chandre Van Vught, Jennifer Velloza and Zukiswa Zingela.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Impact, processes, and mechanisms of coproduction with youth in the MindKind study.
[DOCX File , 30 KB - publichealth_v9i1e42963_app1.docx]

References

1. Elwyn G, Nelson E, Hager A, Price A. Coproduction: when users define quality. BMJ Qual Saf 2020 Sep;29(9):711-716 [FREE Full text] [doi: 10.1136/bmjqs-2019-009830] [Medline: 31488570]
2. Smith H, Budworth L, Grindey C, Hague I, Hamer N, Kislov R, et al. Co-production practice and future research priorities in United Kingdom-funded applied health research: a scoping review. Health Res Policy Syst 2022 Apr 02;20(1):36 [FREE Full text] [doi: 10.1186/s12961-022-00838-x] [Medline: 35366898]
3. Bandola-Gill J, Arthur M, Leng RI. What is co-production? Conceptualising and understanding co-production of knowledge and policy across different theoretical perspectives. Evid Policy 2023;19(2):275-298. [doi: 10.1332/174426421X16420955772641]
4. Boswell N, Douglas-Osborn E, Halkyard T, Woods K. Listening to children and young people: an educational psychology service co-production journey. Educ Psychol Pract 2021 Sep 24;37(4):396-412. [doi: 10.1080/02667363.2021.1975097]
5. Taylor A, McMellon C, French T, MacLachlan A, Evans R, Lewis R, et al. Defining research priorities for youth public mental health: reflections on a coproduction approach to transdisciplinary working. Health Res Policy Syst 2022 Jun 20;20(1):72 [FREE Full text] [doi: 10.1186/s12961-022-00871-w] [Medline: 35725482]
6. Hickey G. The potential for coproduction to add value to research. Health Expect 2018 Aug 31;21(4):693-694 [FREE Full text] [doi: 10.1111/hex.12821] [Medline: 30168263]
7. Heaton J, Day J, Britten N. Collaborative research and the co-production of knowledge for practice: an illustrative case study. Implement Sci 2016 Feb 20;11(1):20 [FREE Full text] [doi: 10.1186/s13012-016-0383-9] [Medline: 26897169]
8. Baum F, MacDougall C, Smith D. Participatory action research. J Epidemiol Community Health 2006 Oct;60(10):854-857 [FREE Full text] [doi: 10.1136/jech.2004.028662] [Medline: 16973531]
9. Anyon Y, Bender K, Kennedy H, Dechants J. A systematic review of youth participatory action research (YPAR) in the United States: methodologies, youth outcomes, and future directions. Health Educ Behav 2018 Dec 11;45(6):865-878. [doi: 10.1177/1090198118769357] [Medline: 29749267]
10. Kessler RC, Angermeyer M, Anthony JC, DE Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176 [FREE Full text] [Medline: 18188442]
11. Pavarini G, Lorimer J, Manzini A, Goundrey-Smith E, Singh I. Co-producing research with youth: the NeurOx young people's advisory group model. Health Expect 2019 Aug;22(4):743-751 [FREE Full text] [doi: 10.1111/hex.12911] [Medline: 31095837]
12. Gaillard S, Malik S, Preston J, Escalera BN, Dicks P, Touil N, et al. Involving children and young people in clinical research through the forum of a European Young Persons' Advisory Group: needs and challenges. Fundam Clin Pharmacol 2018 Aug 08;32(4):357-362. [doi: 10.1111/fcp.12360] [Medline: 29457267]

13. Comfort M, Raymond-Flesch M, Auerswald C, McGlone L, Chavez M, Minnis A. Community-engaged research with rural Latino adolescents: design and implementation strategies to study the social determinants of health. Gateways 2018 May 31;11(1):90-108 [FREE Full text] [doi: 10.5130/ijcre.v11i1.5721] [Medline: 31236142]
14. Kelly B, Friel S, McShane T, Pinkerton J, Gilligan E. “I haven't read it, I've lived it!” the benefits and challenges of peer research with young people leaving care. Qualitative Soc Work 2018 Sep 19;19(1):108-124. [doi: 10.1177/1473325018800370]
15. Delman J. Participatory action research and young adults with psychiatric disabilities. Psychiatr Rehabil J 2012;35(3):231-234. [doi: 10.2975/35.3.2012.231.234] [Medline: 22246121]
16. de Jong G, Meijer E, Schout G, Abma T. Involving undergraduate nursing students in participatory health research: implications from the Netherlands. J Prof Nurs 2018 Nov;34(6):507-513. [doi: 10.1016/j.profnurs.2018.04.001] [Medline: 30527701]
17. Valdez ES, Skobic I, Valdez L, O Garcia D, Korchmaros J, Stevens S, et al. Youth participatory action research for youth substance use prevention: a systematic review. Subst Use Misuse 2020 Oct 09;55(2):314-328 [FREE Full text] [doi: 10.1080/10826084.2019.1668014] [Medline: 31596160]
18. Filipe A, Renedo A, Marston C. The co-production of what? Knowledge, values, and social relations in health care. PLoS Biol 2017 May 3;15(5):e2001403 [FREE Full text] [doi: 10.1371/journal.pbio.2001403] [Medline: 28467412]
19. Dewa LH, Lawrence-Jones A, Crandell C, Jaques J, Pickles K, Lavelle M, et al. Reflections, impact and recommendations of a co-produced qualitative study with young people who have experience of mental health difficulties. Health Expect 2021 May;24 Suppl 1(Suppl 1):134-146 [FREE Full text] [doi: 10.1111/hex.13088] [Medline: 32515538]
20. Rycroft-Malone J, Burton CR, Bucknall T, Graham ID, Hutchinson AM, Stacey D. Collaboration and co-production of knowledge in healthcare: opportunities and challenges. Int J Health Policy Manag 2016 Jan 28;5(4):221-223 [FREE Full text] [doi: 10.15171/ijhpm.2016.08] [Medline: 27239867]
21. Kothari A, MacLean L, Edwards N, Hobbs A. Indicators at the interface: managing policymaker-researcher collaboration. Knowledge Manag Res Pract 2017 Dec 19;9(3):203-214. [doi: 10.1057/kmrp.2011.16]
22. Cooke J, Langley J, Wolstenholme D, Hampshaw S. "Seeing" the difference: the importance of visibility and action as a mark of "Authenticity" in co-production comment on "Collaboration and co-production of knowledge in healthcare: opportunities and challenges". Int J Health Policy Manag 2017 Jun 01;6(6):345-348 [FREE Full text] [doi: 10.15171/ijhpm.2016.136] [Medline: 28812827]
23. Camino LA. Youth-adult partnerships: entering new territory in community work and research. Applied Develop Sci 2000 Jun;4(sup1):11-20. [doi: 10.1207/S1532480XADS04Suppl_2]
24. Heffernan OS, Herzog TM, Schiralli JE, Hawke LD, Chaim G, Henderson JL. Implementation of a youth-adult partnership model in youth mental health systems research: challenges and successes. Health Expect 2017 Dec 14;20(6):1183-1188 [FREE Full text] [doi: 10.1111/hex.12554] [Medline: 28295940]
25. Huang X, O'Connor M, Ke L, Lee S. Ethical and methodological issues in qualitative health research involving children: a systematic review. Nurs Ethics 2016 May;23(3):339-356. [doi: 10.1177/0969733014564102] [Medline: 25552586]
26. Wilson O, Daxenberger L, Dieudonne L, Eustace J, Hanard A, Krishnamurthi A, et al. A rapid evidence review of young people’s involvement in health research. Wellcome. 2020. URL: https://cms.wellcome.org/sites/default/files/2021-02/a-rapid-evidence-review-of-young-peoples-involvement-in-health-research.pdf [accessed 2022-08-04]
27. Beckett K, Farr M, Kothari A, Wye L, le May A. Embracing complexity and uncertainty to create impact: exploring the processes and transformative potential of co-produced research through development of a social impact model. Health Res Policy Syst 2018 Dec 11;16(1):118 [FREE Full text] [doi: 10.1186/s12961-018-0375-0] [Medline: 30537975]
28. Sellars E, Pavarini G, Michelson D, Creswell C, Fazel M. Young people's advisory groups in health research: scoping review and mapping of practices. Arch Dis Child 2021 Jul 18;106(7):698-704. [doi: 10.1136/archdischild-2020-320452] [Medline: 33208398]
29. May CR, Johnson M, Finch T. Implementation, context and complexity. Implement Sci 2016 Oct 19;11(1):141 [FREE Full text] [doi: 10.1186/s13012-016-0506-3] [Medline: 27756414]
30. The MindKind Consortium. MindKind: a mixed-methods protocol for the feasibility of global digital mental health studies in young people. Wellcome Open Res 2021 May 12;6:275 [FREE Full text] [doi: 10.12688/wellcomeopenres.17167.2] [Medline: 35686088]
31. Burchardt T. Deliberative research as a tool to make value judgements. Qual Res 2013 Jan 11;14(3):353-370. [doi: 10.1177/1468794112469624]
32. Sieberts SK, Marten C, Bampton E, Björling EA, Burn AM, Carey EG, MindKind Consortium, et al. Young people's data governance preferences for their mental health data: MindKind Study findings from India, South Africa, and the United Kingdom. PLoS One 2023;18(4):e0279857 [FREE Full text] [doi: 10.1371/journal.pone.0279857] [Medline: 37074995]
33. Davies R, Dart J. The ‘most significant change’ (MSC) technique: a guide to its use”. CARE International. 2005. URL: https://www.researchgate.net/publication/275409002_The_%27Most_Significant_Change%27_MSC_Technique_A_Guide_to_Its_Use?channel=doi&linkId=553bd3b60cf29b5ee4b87d86&showFulltext=true [accessed 2022-08-04]

34. Dixon J, Ward J, Blower S. “They sat and actually listened to what we think about the care system”: the use of participation, consultation, peer research and co-production to raise the voices of young people in and leaving care in England. Child Care Pract 2018 Oct 17;25(1):6-21 [FREE Full text] [doi: 10.1080/13575279.2018.1521380]
35. Corbin JH, Jones J, Barry MM. What makes intersectoral partnerships for health promotion work? A review of the international literature. Health Promot Int 2018 Feb 01;33(1):4-26 [FREE Full text] [doi: 10.1093/heapro/daw061] [Medline: 27506627]
36. Turnhout E, Metze T, Wyborn C, Klenk N, Louder E. The politics of co-production: participation, power, and transformation. Curr Opinion Environ Sustainability 2020 Feb;42:15-21. [doi: 10.1016/j.cosust.2019.11.009]
37. Corbin JH, Mittelmark MB. Partnership lessons from the global programme for health promotion effectiveness: a case study. Health Promot Int 2008 Dec 28;23(4):365-371. [doi: 10.1093/heapro/dan029] [Medline: 18835888]
38. Norström AV, Cvitanovic C, Löf MF, West S, Wyborn C, Balvanera P, et al. Principles for knowledge co-production in sustainability research. Nat Sustain 2020 Jan 20;3(3):182-190. [doi: 10.1038/s41893-019-0448-2]
39. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? Health Res Policy Syst 2019 Mar 28;17(1):33 [FREE Full text] [doi: 10.1186/s12961-019-0432-3] [Medline: 30922339]
40. Williams O, Sarre S, Papoulias SC, Knowles S, Robert G, Beresford P, et al. Lost in the shadows: reflections on the dark side of co-production. Health Res Policy Syst 2020 May 07;18(1):43 [FREE Full text] [doi: 10.1186/s12961-020-00558-0] [Medline: 32380998]
41. Bovaird T, Flemig S, Loeffler E, Osborne SP. How far have we come with co-production—and what’s next? Public Money Manage 2019 May 20;39(4):229-232. [doi: 10.1080/09540962.2019.1592903]
42. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ 2017 Aug 02;358:j3453 [FREE Full text] [doi: 10.1136/bmj.j3453] [Medline: 28768629]

Abbreviations

GRIPP: Guidance for Reporting Involvement of Patients and the Public
MSC: Most Significant Change
YPAG: young people’s advisory group

Edited by Y Khader; submitted 26.09.22; peer-reviewed by E Rozendaal, S Acquilano; comments to author 17.02.23; revised version received 09.04.23; accepted 14.04.23; published 19.06.23.

Please cite as:
Fernandes B, Neelakantan L, Shah H, Sumant S, Collins PY, Velloza J, Bampton E, Ranganathan S, Sibisi R, Bashir T, Bowes J, David EL, Kaur H, Malik U, Shannon I, Gurumayum S, Burn AM, Sieberts SK, The MindKind Consortium, Fazel M
Evidencing the Impact of Web-Based Coproduction With Youth on Mental Health Research: Qualitative Findings From the MindKind Study
JMIR Public Health Surveill 2023;9:e42963
URL: https://publichealth.jmir.org/2023/1/e42963
doi:10.2196/42963
PMID:37335609

Original Paper

Association Between the Loss of Gait Harmony and Cognitive Impairment: Cross-Sectional Study

Ju-Young Choi[1], MPH; Sang-Won Ha[2], MD; Da-Eun Jeong[2], MD; Jaeho Lee[1,3], MPH; Donghoon Kim[1,3], MPH; Jin-Young Min[4*], PhD; Kyoung-Bok Min[1,3,5*], PhD

[1]Department of Preventive Medicine, College of Medicine, Seoul National University, Seoul, Republic of Korea
[2]Department of Neurology, Veterans Health Service Medical Center, Seoul, Republic of Korea
[3]Integrated Major in Innovative Medical Science, Seoul National University Graduate School, Seoul, Republic of Korea
[4]Veterans Medical Research Institute, Veterans Health Service Medical Center, Seoul, Republic of Korea
[5]Institute of Health Policy and Management, Seoul National University Medical Research Center, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Kyoung-Bok Min, PhD
Department of Preventive Medicine
College of Medicine
Seoul National University
103 Daehak-ro, Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 1036954330
Email: minkb@snu.ac.kr

Abstract

Background: Functional limitations and disabilities have been associated with a decrease in cognitive function due to increasing age. Gait performance and cognitive function have been associated with gait variability in executive function, the phase domain in memory, and gait abnormalities in cognitive decline.

Objective: Our study aimed to investigate whether gait harmony was associated with cognitive function in the older adult population. Moreover, we aimed to investigate whether gait harmony was associated with cognitive function and explore each cognitive function in a specific harmonic state.

Methods: The study population included 510 adults aged ≥60 years who visited the Department of Neurology at the Veterans Health Service Medical Center, Seoul, South Korea. Gait data were collected using a 3D motion capture device with a wireless inertial measurement unit system. For cognitive function assessments, we used the Seoul Neuropsychological Screening Battery-Core test, which evaluates the level of cognitive function or impairment in 5 cognitive domains.

Results: In general, the association between the Seoul Neuropsychological Screening Battery-Core tests and the stance-to-swing ratio in the >1.63 ratio group yielded lower β coefficients than those in the 1.50-1.63 ratio group. After adjustment for confounders, the odds ratio (OR) for the Digit Symbol Coding test (adjusted OR 0.42, 95% CI 0.20-0.88) and the Korean version of the Color Word Stroop Test: 60 seconds (adjusted OR 0.51, 95% CI 0.29-0.89) for frontal and executive function were significantly lower for the >1.63 ratio group than the reference group.

Conclusions: Our findings suggest that the gait phase ratio is a valuable indicator of walking deficits and may also be associated with cognitive impairment in older adults.

(JMIR Public Health Surveill 2023;9:e46264) doi:10.2196/46264

KEYWORDS

cognitive function; gait phase; physical performance; dementia; older adult; aging; asymmetric; balance; gait analysis; cognition; cognitive impairment; gait; gait pattern; aging

Introduction

The global population is rapidly aging, and the number of people aged 65 years or older is expected to increase from 703 million in 2019 to 1.5 billion in 2050 [1]. With the increasing older adult population, age-related cognitive impairment has increased and will become a public health problem unless prevention and intervention are implemented [2]. The development of many neurodegenerative disorders is usually a feature of the older adult population in the context of aging [3], and age is known as the primary risk factor for many diseases such as Alzheimer disease (AD) and dementia [4,5]. Approximately 50 million people have been diagnosed with dementia, which is one of the leading causes of disability and mortality among aging adults [6]. Moreover, dementia resulting from cognitive decline is an irreversible process, and there are no effective treatments or drugs for dementia [7]. Therefore, it is important to identify factors that can delay the onset of cognitive impairment or impact cognitive outcomes.

Human gait is complicated and divided into 5 primary modal domains: rhythm, phase, variability, pace, and base of support [8]. Safe and effective gaits are important markers throughout life [9]. In older individuals, gait is used as a predictor of health status [10], falls, activities of daily living [11,12], risk of dementia [13], and even risk of early mortality [14]. Functional limitations and disabilities have been associated with a decrease in cognitive function due to increasing age [15]; motor performance may be related to cognitive impairment and dementia [16,17]. Furthermore, impairment of the motor system, such as gait abnormalities, can lead to cognitive decline [16,17] or early stages of dementia with aging [18]. Gait performance and cognitive function have been associated with gait variability in executive function [19], the phase domain in memory [20], and gait abnormalities in cognitive decline [21]. An abnormal gait, which increases the risk of dementia in older adults, suggests low physical fitness in areas such as mobility [22] and balance [23,24]. Therefore, gait cannot be simply explained as an autonomic movement.

A principal issue is that older adults who demonstrate an imbalance in gait are likely to be at a higher risk of cognitive impairment [25]. Gait harmony, which indicates the proportion of stance to swing in kinematic gait, is defined as the capacity to transform the symmetrical human body into alternated, synchronized, symmetric, and rhythmic movements [26]. Previous studies have shown that the proportion of stance to swing is reduced during fast walking and increases under pathological conditions such as Parkinson disease (PD) [27], stroke [28], and spastic cerebral palsy [29]. A ratio from stance to swing of 70% to 30% was found in patients with stroke [28], and a longer stance phase increased gait stability [30]. It has been suggested that a stance between 59% and 70% is a good compromise between fast and stable walking by the harmonic properties of the stance-to-swing ratio [30]. The harmonic ratio property is founded not only in various disparate physical and biological systems, including leaf disposition [31], nucleotide frequencies [32], and cell [33], but also in human sciences, such as body proportions [34] and aesthetic preferences [35]. However, there is a lack of evidence supporting the association between gait harmony and cognitive function.

Therefore, the objective of this study was to investigate whether gait harmony was associated with cognitive function in the older adult population. In addition, we aimed to explore each cognitive function in a specific harmonic state.

Methods

Setting and Study Design

A cross-sectional study was conducted at the Veterans Health Service Medical Center, Seoul, South Korea. Individuals aged ≥60 years who visited the Department of Neurology between March and December 2021 were recruited in this study. The inclusion criteria were as follows: (1) those concerned about cognitive decline, (2) those who could independently undergo all the clinical tests and respond to questionnaires, and (3) those who agreed to participate in this study. Individuals were excluded based on the following criteria: (1) a diagnosis of dementia (International Classification of Diseases, 10th Revision, codes F00-F09 and G30); (2) diagnoses of brain infarction, cerebral hemorrhage, or PD; and (3) a diagnosis of another serious disease (eg, cancer or mental illness). Experienced neurological clinicians evaluated the inclusion and exclusion criteria.

The participants were subjected to a health survey consisting of gait measurements, cognitive examinations, and questionnaires. The survey was conducted at the Veterans Medical Research Institute of the Veterans Health Service Medical Center. A total of 575 individuals volunteered to participate in this study and provided informed consent at enrollment. Of these, 65 participants subsequently dropped out of the study because they were younger than 60 years (n=5) or had missing data on variables (ie, sex, education level, BMI, gait measurements, and cognitive parameters; n=60). After these exclusions, 510 people were eligible for this study.

Ethics Approval

Study protocols were approved by the Institutional Ethics Review Board of the Veterans Health Service Medical Center (BOHUN 2021-02-024 and BOHUN 2021-01-066). All participants provided signed informed consent prior to study enrollment.

Gait Analysis

Kinematic data of gait parameters were collected using the NORAXON myoMOTION sensor, which is a 3D motion capture device used with a wireless inertial measurement unit (IMU) system. The IMU sensor plays a role in transmitting human movement data directly to the myoMOTION receiver to compute the angular changes of the selected body segments. In a particular space, the 3D rotation angles of each sensor on selected body segments were measured from the 3D accelerometer, gyroscope, and magnetometer using *fusion algorithms*; these angles are also known as orientation or navigation angles. The sensor for angular orientation uses a robust fusion algorithm combining the elemental sensor component axes' readings as quaternion element values. This technology mathematically combines and filters incoming source signals at the sensor level and transmits the 4 quaternions of each sensor. Then, using NORAXON's software, the quaternion data are automatically converted into anatomical angles using a rigid body model with joint segments. NORAXON's software also records each sensor's orientation angles and linear acceleration. The system is intended to quantify angular changes of the involved joint and can mathematically derive the x, y, and z displacements in space from linear acceleration data and inverse kinematics modeling. Using this process, the algorithm uses a gyroscope and acceleration data from body-mounted sensors to identify the stance and swing phases in the gait cycle through the signal recorded at a sampling rate of 100-200 Hz. The recoded IMU data were added, filtered, and processed using the myoMOTION software to quantify angular changes in the joints, and the output was exported to Microsoft Excel files [36-38].

For this study, 7 IMU sensors were attached to the participants' feet, frontal tibia bones, quadriceps, and the sacrum on the left and right sides symmetrically (ie, 2 on the feet, 2 on the frontal tibia bones, 2 on the quadriceps, and 1 on the sacrum, respectively). Calibration was performed using an upright posture to determine the value of the 0° angle in the joints. The participants were instructed to walk at their usual pace for the measured distance (10 m), and walking start or stop points for the measured distance were marked on the floor. The participants started walking at the start signal and stopped walking by themselves at the stop point. However, owing to an acceleration and deceleration phase during the 10-m gait test, the first and last steps were removed.

The data transmitted to the myoMOTION software comprised the kinematic parameters of gait variables including the automatically processed average signals for each stance and swing phase. The stance-to-swing ratio was calculated using the average signal values of the stance and swing phases (*stance phase [%] / swing phase [%]*) [30]. The total gait cycle comprises the stance and swing phases. The literature suggests that during stable human walking, with the total gait cycle set to 100%, the stance phase comprises 60% to 62% and the swing phase comprises 40% to 38%, equating to a ratio of 1.50-1.63 [39-41]. These values were then used to determine the harmonic gait group in this study. Therefore, gait ratios were stratified into 3 groups (<1.50, 1.50-1.63, or >1.63). The harmonic group included individuals in the 1.50-1.63 ratio range, whereas the nonharmonic group included those with the other ratios (<1.50 and >1.63).

Cognition Evaluation

Participants performed a brief version of the Seoul Neuropsychological Screening Battery, named the Seoul Neuropsychological Screening Battery-Core (SNSB-C), which evaluates the level of cognitive function or impairment in the following 5 cognitive domains: attention, language and related functions, visuospatial functions, memory, and frontal and executive functions [42]. SNSB-C is made up of 14 subtests, including the Digit Span Test, a short version of the Korean-Boston Naming Test (S-K-BNT), Rey Complex Figure Test (RCFT), Seoul Verbal Learning Test: Delayed Recall (SVLT: DR), the Korean version of Color Word Stroop Test: 60 seconds (K-CWST: 60sec), Controlled Oral Word Association Test (COWAT), Trail Making Test-Elderly: Part B (TMT-E: B), and Digit Symbol Coding [43]. Each SNSB-C score is expressed as standardized percentile, stratified for age, sex, and education.

Other Variables

The questionnaire information included age (≥60 years) and sex (male or female). Health behavior variables included smoking status (current, former, or never), alcohol consumption (drinker or nondrinker), vigorous exercises (yes or no), and BMI. BMI was calculated by dividing the individual's weight (kg) by height squared (m^2).

Statistical Analyses

Statistical differences between the characteristics of the study population were analyzed according to the stance-to-swing ratio (<1.50, 1.50-1.63, or >1.63 ratio group) using the proportion of physiologically comfortable human gait [30,40,41,44,45]. In the analysis, cognitive function and gait ratio were used as the dependent and independent variables, respectively. For each variable, the chi-square test and 2-tailed *t* test were performed for each group of participants. Linear regression analysis was used to evaluate the association between cognitive function tests and the stance-to-swing ratio in each group of gait phase ratio and to provide β coefficients and SEs. The logistic regression model provided odds ratios (ORs) and 95% CIs for each cognitive function test percentiles as the stance-to-swing ratio increased or decreased. The 1.50-1.63 ratio group was used as the reference. Regression models were adjusted for age, sex, education level, smoking status, alcohol consumption, vigorous exercises, and BMI. All statistical analyses were performed with SAS (version 9.4; SAS Institute), and the statistical significance was established at α=.05.

Results

Participant Characteristics

Table 1 summarizes the characteristics of the study population according to the 3 stance-to-swing ratio groups. A total of 510 participants with a mean age of 74.1 (SD 5.6) years were included in the study. Of the participants, 34 (6.7%) were in the <1.50 ratio group, 122 (23.9%) were in the 1.50-1.63 ratio group, and 354 (69.4%) were in the >1.63 ratio group.

Table 1. Characteristics of the study population according to the 3 stance-to-swing ratio groups.

Characteristic	Total population (n=510)	Stance-to-swing ratio group[a]			*P* value[b]
		<1.50	1.50-1.63	>1.63	
Participant, n (%)	510 (100)	34 (6.7)	122 (23.9)	354 (69.4)	
Age (year), mean (SD)	74.1 (5.6)	74.1 (5.2)	73.7 (5.0)	74.3 (5.8)	.56[c]
Sex, n (%)					.57[d]
Male	272 (53.3)	17 (6.3)	70 (25.7)	185 (68)	
Female	238 (46.7)	17 (7.1)	52 (21.9)	169 (71)	
Education, n (%)					.06[d]
Less than high school	259 (50.8)	13 (5)	51 (19.7)	195 (75.3)	
High school	118 (23.1)	9 (7.6)	34 (28.8)	75 (63.6)	
College or more	133 (26.1)	12 (9)	37 (27.8)	84 (63.2)	
Smoking status, n (%)					.66[d]
Current or former smoker	196 (38.4)	12 (6.1)	51 (26)	133 (67.9)	
Never smoker	314 (61.6)	22 (7)	71 (22.6)	221 (70.4)	
Alcohol drinking, n (%)					.12[d]
Drinker	308 (60.4)	18 (5.8)	83 (27)	207 (67.2)	
Nondrinker	202 (39.6)	16 (7.9)	39 (19.3)	147 (72.8)	
Vigorous intensity exercises for a week, n (%)					.045[d]
Yes	101 (19.8)	12 (11.9)	26 (25.7)	63 (62.4)	
No	409 (80.2)	22 (5.4)	96 (23.5)	291 (71.2)	
BMI (kg/m^2), mean (SD)	24.9 (3.2)	24.1 (2.5)	23.9 (2.9)	25.4 (3.2)	<.001
Gait phase (%), mean (SD)					
Stance	63.0 (2.1)	59.0 (1.2)	61.2 (0.5)	64.0 (1.6)	<.001[c]
Swing	37.0 (2.1)	41.0 (1.2)	38.8 (0.5)	36.0 (1.6)	<.001[c]
Stance-to-swing	1.7 (0.2)	1.4 (1.2)	1.6 (0.03)	1.8 (0.1)	<.001[c]

[a] The denominators of the percentages in these columns correspond to the n values in the “Total population” column of the same row.

[b] The *P* values are based on chi-square test for categorical variables and one-way ANOVA for numerical variables in the stance-to-swing ratio groups.

[c] ANOVA.

[d] Chi-square test.

Cognitive Performance and Gait Balance

Table 2 shows the mean (SE) of each SNSB-C test (percentile) data according to the stance-to-swing ratios. The Digit Symbol Coding (56.31 vs 61.65 vs 52.88; *P*=.02), TMT-E: B (61.77 vs 57.76 vs 51.49; *P*=.02), and K-CWST: 60sec (45.67 vs 48.72 vs 40.87; *P*=.03) tests for frontal and executive functions showed statistically significant differences among the stance-to-swing ratio groups. However, there were no differences in the other cognitive functions.

Table 2. Mean (SE) of each Seoul Neuropsychological Screening Battery-Core (SNSB-C) test (percentile) according to the stance-to-swing ratio groups.

Function and test	Stance-to-swing ratio group, mean (SE)			*P* value[a]
	<1.50	1.50-1.63	>1.63	
Attention				
Digit Span Test	44.25 (4.88)	50.22 (2.83)	43.71 (1.54)	.11
Visuospatial				
RCFT[b]: copy	65.16 (4.37)	55.97 (2.44)	55.98 (1.53)	.19
Language				
S-K-BNT[c]	53.41 (5.11)	58.57 (2.59)	53.35 (1.47)	.20
Memory				
SVLT: DR[d]	38.64 (5.58)	32.95 (2.78)	36.77 (1.63)	.44
Frontal and executive				
Digit Symbol Coding	56.31 (4.97)	61.65 (2.46)	52.88 (1.61)	.02
COWAT[e]: animal + ㄱ	49.21 (4.88)	46.45 (2.61)	41.96 (1.50)	.16
TMT-E: B[f]	61.77 (3.27)	57.76 (2.52)	51.49 (1.49)	.02
K-CWST: 60sec[g]	45.67 (4.92)	48.72 (2.64)	40.87 (1.56)	.03

[a] The *P* values are based on one-way ANOVA for numerical variables in the stance-to-swing ratio groups.

[b] RCFT: Rey Complex Figure Test.

[c] S-K-BNT: short version of the Korean-Boston Naming Test.

[d] SVLT: DR: Seoul Verbal Learning Test: Delayed Recall.

[e] COWAT: Controlled Oral Word Association Test.

[f] TMT-E: B: Trail Making Test-Elderly: Part B.

[g] K-CWST: 60sec: the Korean version of Color Word Stroop Test: 60 seconds.

Association Between Cognitive Performance and Gait Balance

Table 3 shows the estimated β coefficients (SE) of each SNSB-C test based on the stance-to-swing ratio groups. In the Digit Span Test for attention functions, a significant association was observed in the >1.63 ratio group (β=–6.514, SE 3.098; *P*=.04) before adjustment. The unadjusted β coefficients were significant in the >1.63 ratio group for the Digit Symbol Coding (β=–8.773, SE 3.098; *P*=.005), TMT-E: B (β=–6.269, SE 2.890; *P*=.03), and K-CWST: 60sec (β=–7.855, SE 3.075; *P*=.01) tests for frontal and executive functions. In the fully adjusted model, a significant β coefficient was observed in the >1.63 ratio group for the S-K-BNT for language function (β=–6.573, SE 2.813; *P*=.02). Compared with the reference, there were significant β coefficients in the >1.63 ratio group for the Digit Symbol Coding (β=–7.991, SE 3.028; *P*=.009) and K-CWST: 60sec (β=–8.083, SE 3.076; *P*=.009) tests for frontal and executive functions. In general, the association between the SNSB-C test data and the stance-to-swing ratio in the >1.63 ratio group yielded lower β coefficients than those in the 1.50-1.63 ratio group. Most β coefficients showed a decreasing relationship for the SNSB-C test data and the stance-to-swing ratio, whereas there were no significant associations in the <1.50 ratio group (all *P*>.05).

Table 3. The estimated β coefficient (SE) of each Seoul Neuropsychological Screening Battery-Core (SNSB-C) test based on the stance-to-swing ratio groups.

Function and test	Unadjusted model (stance-to-swing ratio group)					Adjusted[a] model (stance-to-swing ratio group)				
	<1.50		1.50-1.63	>1.63		<1.50		1.50-1.63	>1.63	
	Value, β (SE)	*P* value		Value, β (SE)	*P* value	Value, β (SE)	*P* value		Value, β (SE)	*P* value
Attention										
Digit Span Test	–5.967 (5.723)	.30	Ref[b]	–6.514 (3.098)	.04	–5.474 (5.739)	.34	Ref	–4.141 (3.184)	.19
Visuospatial										
RCFT[c]: copy	9.193 (5.461)	.09	Ref	0.014 (2.956)	>.99	8.913 (5.460)	.10	Ref	0.981 (3.029)	.75
Language										
S-K-BNT[d]	–5.160 (5.444)	.34	Ref	–5.215 (2.947)	.08	–3.471 (5.071)	.49	Ref	–6.573 (2.813)	.02
Memory										
SVLT: DR[e]	5.691 (5.979)	.34	Ref	3.817 (3.237)	.24	5.696 (5.724)	.32	Ref	3.479 (3.175)	.27
Frontal and executive										
Digit Symbol Coding	–5.344 (5.706)	.35	Ref	–8.773 (3.089)	.005	–4.092 (5.458)	.45	Ref	–7.991 (3.028)	.009
COWAT[f]: animal + ㄱ	2.766 (5.514)	.62	Ref	–4.486 (2.985)	.13	2.904 (5.538)	.60	Ref	–5.175 (3.072)	.09
TMT-E: B[g]	4.016 (5.338)	.45	Ref	–6.269 (2.890)	.03	5.552 (5.290)	.29	Ref	–5.206 (2.934)	.08
K-CWST: 60sec[h]	–3.049 (5.680)	.59	Ref	–7.855 (3.075)	.01	–2.318 (5.545)	.68	Ref	–8.083 (3.076)	0=.009

[a] This result was adjusted for age, sex, education, smoking status, alcohol consumption, vigorous exercises, and BMI.

[b] Ref: reference.

[c] RCFT: Rey Complex Figure Test.

[d] S-K-BNT: short version of the Korean-Boston Naming Test.

[e] SVLT: DR: Seoul Verbal Learning Test: Delayed Recall.

[f] COWAT: Controlled Oral Word Association Test.

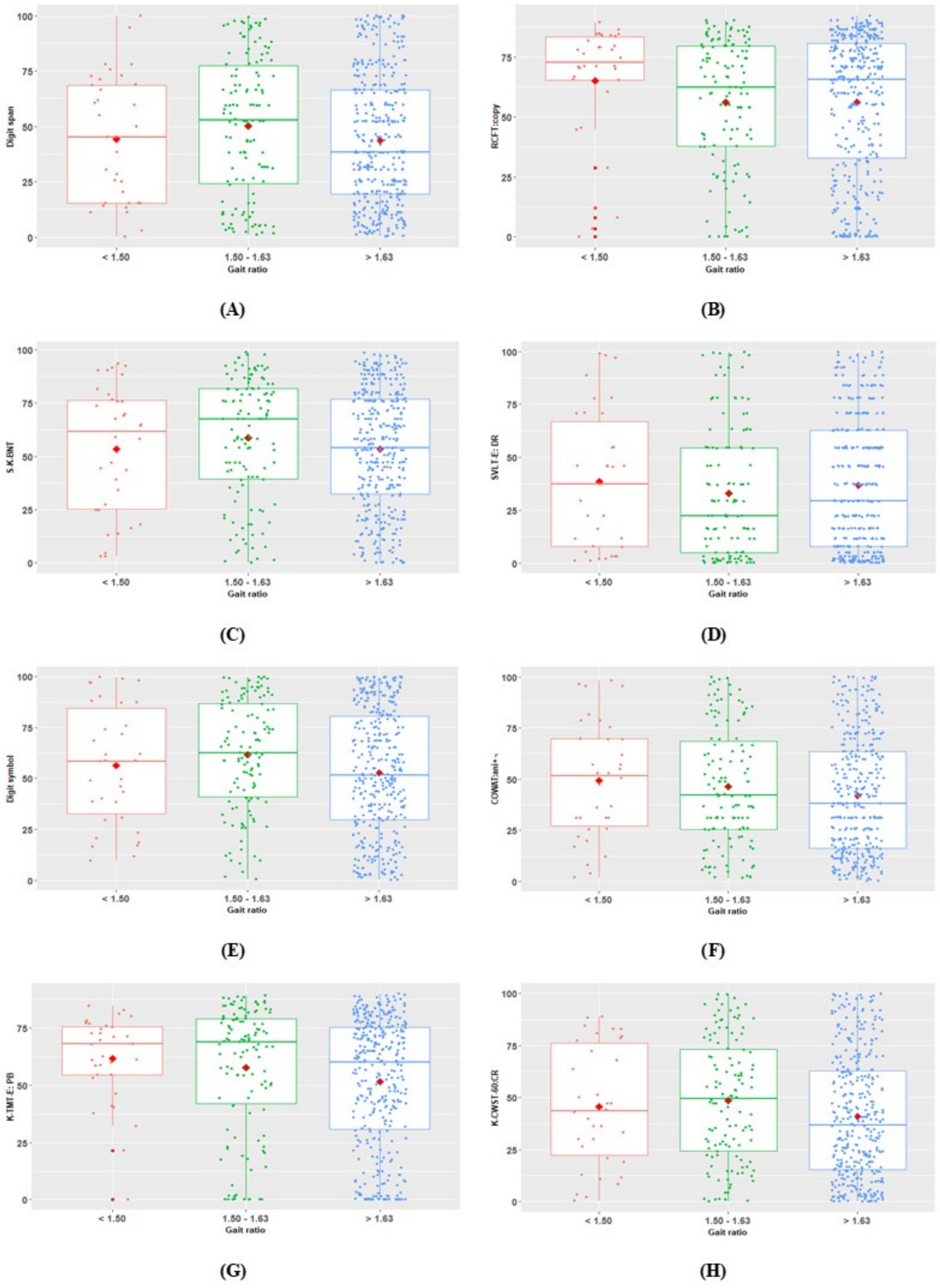
[g] TMT-E: B: Trail Making Test-Elderly: Part B.

[h] K-CWST: 60sec: the Korean version of Color Word Stroop Test: 60seconds.

Figure 1 presents the distributions between SNSB-C tests and stance-to-swing ratio groups using box and scatter plots. Overall, the study participants in the >1.63 gait ratio group were distributed across all the percentiles for each SNSB-C test.

The adjusted OR (95% CI) for each SNSB-C test according to the stance-to-swing ratio group is shown in Figure 2. After adjustment for age, sex, education, smoking status, alcohol consumption, vigorous exercises, and BMI, the adjusted OR for the Digit Symbol Coding (adjusted OR 0.42, 95% CI 0.20-0.88) and K-CWST: 60sec (adjusted OR 0.51, 95% CI 0.29-0.89) tests for frontal and executive function was significantly lower for the >1.63 ratio group than for the reference group (the 1.50-1.63 stance-to-swing ratio group). However, most cognitive function tests showed no differences according to the stance-to-swing ratio group.

Figure 1. The box and scatter plots of the gait ratio groups for each SNSB-C test. The red box and scatter plots represent the <1.50 ratio group, the green box and scatter plots represent the 1.50-1.63 ratio group, and the blue box and scatter plots represent the >1.63 ratio group. The red dots are the mean of each SNSB-C test percentile according to each gait ratio group: (A) Digit Span Test, (B) RCFT: copy, (C) S-K-BNT, (D) SVLT: DR, (E) Digit Symbol Xoding, (F) COWAT: animal + ㄱ, (G) TMT-E: B, and (H) K-CWST: 60 sec. COWAT: Controlled Oral Word Association Test; K-CWST: 60sec: the Korean version of Color Word Stroop Test: 60 seconds; RCFT: Rey Complex Figure Test; S-K-BNT: short version of the Korean-Boston Naming Test; SNSB-C: Seoul Neuropsychological Screening Battery-Core; SVLT: DR: Seoul Verbal Learning Test: Delayed Recall; and TMT-E: B: Trail Making Test-Elderly: Part B.



(A)

(B)

(C)

(D)

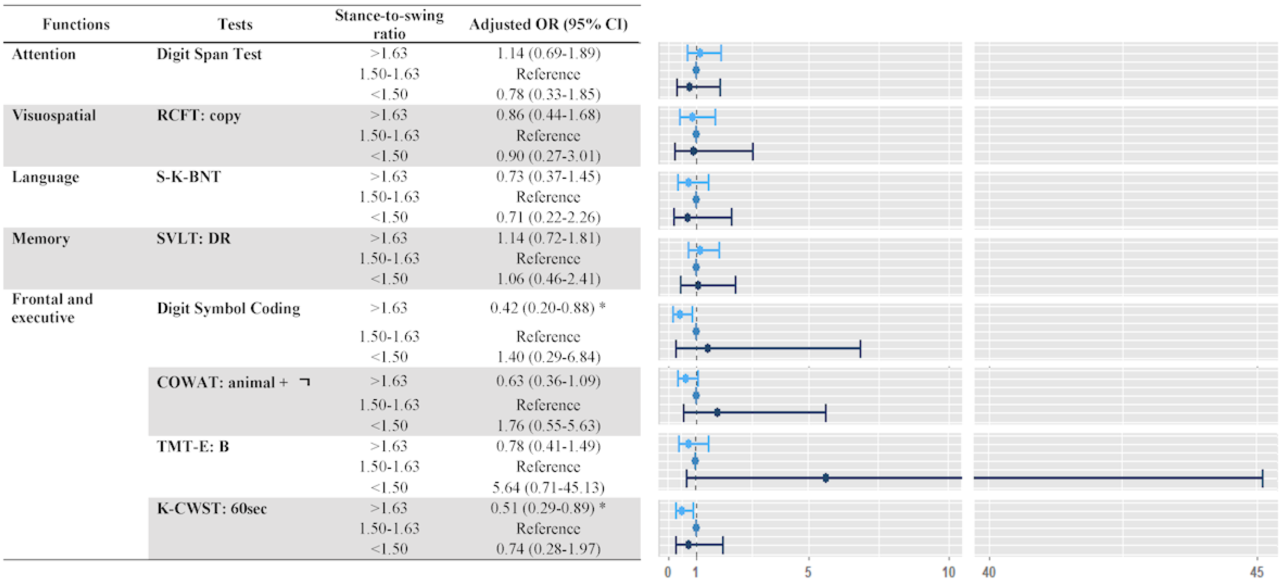
(E)

(F)

(G)

(H)

Figure 2. Adjusted odds ratio (OR; 95% CI) of each Seoul Neuropsychological Screening Battery-Core (SNSB-C) test according to the stance-to-swing ratio group. Adjusted for age, sex, education, smoking status, alcohol consumption, vigorous exercises, and BMI. **P*<.001. COWAT: Controlled Oral Word Association Test; K-CWST: 60sec: the Korean version of Color Word Stroop Test: 60seconds; RCFT: Rey Complex Figure Test; S-K-BNT: short version of the Korean-Boston Naming Test; SVLT: DR: Seoul Verbal Learning Test: Delayed Recall; TMT-E: B: Trail Making Test-Elderly: Part B.



Functions	Tests	Stance-to-swing ratio	Adjusted OR (95% CI)
Attention	Digit Span Test	>1.63	1.14 (0.69-1.89)
		1.50-1.63	Reference
		<1.50	0.78 (0.33-1.85)
Visuospatial	RCFT: copy	>1.63	0.86 (0.44-1.68)
		1.50-1.63	Reference
		<1.50	0.90 (0.27-3.01)
Language	S-K-BNT	>1.63	0.73 (0.37-1.45)
		1.50-1.63	Reference
		<1.50	0.71 (0.22-2.26)
Memory	SVLT: DR	>1.63	1.14 (0.72-1.81)
		1.50-1.63	Reference
		<1.50	1.06 (0.46-2.41)
Frontal and executive	Digit Symbol Coding	>1.63	0.42 (0.20-0.88) *
		1.50-1.63	Reference
		<1.50	1.40 (0.29-6.84)
	COWAT: animal + ㄱ	>1.63	0.63 (0.36-1.09)
		1.50-1.63	Reference
		<1.50	1.76 (0.55-5.63)
	TMT-E: B	>1.63	0.78 (0.41-1.49)
		1.50-1.63	Reference
		<1.50	5.64 (0.71-45.13)
	K-CWST: 60sec	>1.63	0.51 (0.29-0.89) *
		1.50-1.63	Reference
		<1.50	0.74 (0.28-1.97)

Discussion

Principal Findings

This study investigated the association between cognitive function and gait harmony in older Korean adults. We found a significantly higher cognitive performance percentile in frontal and executive functions among participants in the harmonic stance-to-swing ratio group. Compared with participants that exhibited a harmonic gait ratio, the cognitive assessment percentile decreased in those who showed the nonharmonic gait ratio, especially for visuospatial, frontal, and executive functions. Specifically, the relationship between the Digit Symbol Coding and K-CWST: 60sec test data on frontal and executive functions and the gait phase ratio remained significant after controlling for covariates. This information supports our understanding that cognitive function is interrelated with gait harmony in the older adult population.

Comparison With Prior Work and Clinical Implications

Our findings on gait balance and cognitive function were partially consistent with the results of previous studies. According to a recent cross-sectional study by Noh et al [21] of 735 community-based individuals (aged 65-89 years), cognitive function was associated with the stance phase at a slower walking speed (β=0.088; *P*=.02). Based on a previous study in Italy on people with and without PD, alterations in gait ratio were expected; Peppe et al [27] suggested that compared with healthy participants, patients with PD have prolonged stance phases (patients with PD: 68.1% vs patients without PD: 63.6%; *P*<.001) [27]. Another study in 31 patients older than 55 years in China [46] investigated the relationship between gait characteristics and dementia, such as poststroke dementia (PSD) and AD. Ni et al [46] showed that in all gait tests, the percentage of time spent in the stance phase was longer (patients with PSD: 63.95% vs patients with AD: 63.09% vs healthy adults: 62.15%; *P*=.002) and in the swing phase was shorter (patients with PSD: 36.04% vs patients with AD: 36.91% vs healthy adults: 37.86%; *P*=.002) among patients with dementia than among healthy controls. Taken together, these studies highlight that older individuals with gait imbalance are relatively susceptible to a decrease in cognitive function.

The mechanism underlying the relationship between proper gait proportion and cognition remains unclear. One possible explanation is that a well-balanced gait is related not only to the motor system, such as muscle strength, but also to cognition, such as memory, attention, executive function, and visuospatial capacity [47,48]. Attention and executive functions are significantly associated with gait speed in older adults with and without cognitive decline [49]. Poor attention and executive processing in the brain are correlated with white matter hyperintensity, which negatively affects gait pace, spatiality, and variability [50]. The variability of gait is related to the stance and swing phases, and the double support phase is more variable in the presence of poor balance. In memory and spatial functions, beyond the primary role of the hippocampus and parahippocampal gyrus, the induction of memory by hippocampal atrophy also influences rhythm, variability, and human balance control [51-53]. Gait reflects the health of individuals in compensating for changes in postural balance and is controlled by well-balanced neural circuits and specific brain structures, including the frontal lobes, basal ganglia, cerebellum, and sensory and motor systems [48,54,55]. Generally, during walking, the complexity of controlling redundant degrees of freedom of bilateral multijoint limbs is reduced by the nervous system [56]. Human gait is structured for specific phases of the gait leg movements [57] that can lead to the maintenance of the basic walking rhythm, which generates alternating activity of the flexor and extensor motoneurons [58]. Simple alternation of flexor and extensor activity is converted into more complex and adaptable walking patterns by hippocampal neurons that receive serotonergic projections from the median raphe nuclei. The projections of serotonergic neurons of the brain stem that innervate the spinal cord stepping generator play a role in activating and influencing the walking rhythm [59]. A decrease in serotonin levels in cerebrospinal fluid and severe gait and posture disorders have been observed in patients with PD [60]. Hence, the stance-to-swing ratios under pathological conditions may be far from the value of the harmonic proportion value.

Strengths and Limitations

To our knowledge, this study demonstrated an association between gait harmony and cognitive function, which has not been reported previously. Additionally, a strength of this study is the use of noninvasive measurement as an index for predicting decreased cognitive function. Moreover, we showed association between gait and cognitive performance using the SNSB-C tests, which includes 5 cognitive functions (ie, attention, visuospatial, language, memory, and frontal and executive). However, the study had some limitations. First, as this was a cross-sectional study, it is difficult to confirm the causal relationship between gait harmony and cognitive function. Therefore, we cannot indicate trends or changes from normal cognitive function to cognitive dysfunction. Second, the study was not free from bias due to self-reported data. As this study was based on an observational investigation, the questions asked to collect data may concern private or sensitive topics, such as education level, smoking status, and alcohol consumption. Thus, self-reporting data can be affected by social desirability. Moreover, we cannot rule out that some variables affecting cognition, such as dietary intake, medication, and occupation, were not analyzed in the statistical model. These unmeasured variables may be residual confounders.

Conclusions

Dementia is one of the most important health issues associated with aging. However, the diagnosis of dementia is expensive and involves complex measurements. Gait, on the other hand, can be measured noninvasively, comparatively conveniently, and rapidly. Therefore, this study can be important for public health management, in that gait performance assessment can be used to screen for potential cognitive impairment.

Our study found that low cognitive function was related to a nonharmonic gait ratio in participants aged 60 years or older in South Korea. In contrast, harmonic gait ratio was associated with good cognitive performance. These findings suggest that the gait ratio may be a valuable indicator of cognitive impairment. The presence of a nonharmonic gait ratio that is different from the general gait ratio may be an indirect marker of cognitive decline, independent of confounding factors in the older adult population. However, more research is required to replicate our results and establish possible mechanisms of a relationship between proper gait ratio and cognitive function.

Acknowledgments

The authors would like to acknowledge all of the participants and study staffs. This research was supported by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) and Korea Dementia Research Center (KDRC), funded by the Ministry of Health and Welfare and Ministry of Science and ICT, South Korea (grant HU20C0487). This work was supported by Veterans Health Service Medical Center Research Grant, Republic of Korea (VHSMC 22043).

Authors' Contributions

JYM and KBM are co-corresponding authors. JYC, SWH, DEJ, JL, DK, JYM, and KBM conceived the study. JYC, JYM, and KBM conducted all statistical analyses. JL and DK reviewed the results. SWH and DEJ provided expertise on information governance and data management. JYC wrote the first draft of the manuscript. SWH, DEJ, JYM, and KBM commented on and reviewed the final draft of the manuscript. All authors agree on the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. United Nations Department of Economic and Social Affairs, Population Division. World Population Ageing 2019. New York, NY: United Nations; Oct 2020:73.
2. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology 2013 May 07;80(19):1778-1783 [FREE Full text] [doi: 10.1212/WNL.0b013e31828726f5] [Medline: 23390181]
3. Juan SMA, Adlard PA. Ageing and cognition. Subcell Biochem 2019 Mar 20;91:107-122. [doi: 10.1007/978-981-13-3681-2_5] [Medline: 30888651]
4. Trevisan K, Cristina-Pereira R, Silva-Amaral D, Aversi-Ferreira TA. Theories of aging and the prevalence of Alzheimer's disease. Biomed Res Int 2019 Jun 16;2019:9171424-9171429 [FREE Full text] [doi: 10.1155/2019/9171424] [Medline: 31317043]

5. Wu JW, Yaqub A, Ma Y, Koudstaal W, Hofman A, Ikram MA, et al. Biological age in healthy elderly predicts aging-related diseases including dementia. Sci Rep 2021 Aug 05;11(1):15929 [FREE Full text] [doi: 10.1038/s41598-021-95425-5] [Medline: 34354164]
6. The global dementia observatory reference guide. World Health Organization. 2018. URL: https://apps.who.int/iris/handle/10665/272669 [accessed 2023-06-23]
7. Dong X, Li S, Sun J, Li Y, Zhang D. Association of coffee, decaffeinated coffee and caffeine intake from coffee with cognitive performance in older adults: National Health and Nutrition Examination Survey (NHANES) 2011-2014. Nutrients 2020 Mar 20;12(3):840 [FREE Full text] [doi: 10.3390/nu12030840] [Medline: 32245123]
8. Hollman JH, McDade EM, Petersen RC. Normative spatiotemporal gait parameters in older adults. Gait Posture 2011 May;34(1):111-118 [FREE Full text] [doi: 10.1016/j.gaitpost.2011.03.024] [Medline: 21531139]
9. Lord S, Galna B, Verghese J, Coleman S, Burn D, Rochester L. Independent domains of gait in older adults and associated motor and nonmotor attributes: validation of a factor analysis approach. J Gerontol A Biol Sci Med Sci 2013 Jul 18;68(7):820-827. [doi: 10.1093/gerona/gls255] [Medline: 23250001]
10. Cesari M, Kritchevsky SB, Penninx BWHJ, Nicklas BJ, Simonsick EM, Newman AB, et al. Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. J Am Geriatr Soc 2005 Oct;53(10):1675-1680. [doi: 10.1111/j.1532-5415.2005.53501.x] [Medline: 16181165]
11. Abellan van Kan G, Rolland Y, Andrieu S, Bauer J, Beauchet O, Bonnefoy M, et al. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. J Nutr Health Aging 2009 Dec 3;13(10):881-889. [doi: 10.1007/s12603-009-0246-z] [Medline: 19924348]
12. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
13. Verghese J, Wang C, Lipton RB, Holtzer R, Xue X. Quantitative gait dysfunction and risk of cognitive decline and dementia. J Neurol Neurosurg Psychiatry 2007 Sep 01;78(9):929-935 [FREE Full text] [doi: 10.1136/jnnp.2006.106914] [Medline: 17237140]
14. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. JAMA 2011 Jan 05;305(1):50-58 [FREE Full text] [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
15. Tait JL, Duckham RL, Milte CM, Main LC, Daly RM. Influence of sequential vs. simultaneous dual-task exercise training on cognitive function in older adults. Front Aging Neurosci 2017 Nov 07;9:368 [FREE Full text] [doi: 10.3389/fnagi.2017.00368] [Medline: 29163146]
16. Verghese J, Lipton RB, Hall CB, Kuslansky G, Katz MJ, Buschke H. Abnormality of gait as a predictor of non-Alzheimer's dementia. N Engl J Med 2002 Nov 28;347(22):1761-1768. [doi: 10.1056/NEJMoa020441] [Medline: 12456852]
17. Wang L, Larson EB, Bowen JD, van Belle G. Performance-based physical function and future dementia in older people. Arch Intern Med 2006 May 22;166(10):1115-1120. [doi: 10.1001/archinte.166.10.1115] [Medline: 16717174]
18. Albers MW, Gilmore GC, Kaye J, Murphy C, Wingfield A, Bennett DA, et al. At the interface of sensory and motor dysfunctions and Alzheimer's disease. Alzheimers Dement 2015 Jan;11(1):70-98 [FREE Full text] [doi: 10.1016/j.jalz.2014.04.514] [Medline: 25022540]
19. Hausdorff JM, Yogev G, Springer S, Simon ES, Giladi N. Walking is more like catching than tapping: gait in the elderly as a complex cognitive task. Exp Brain Res 2005 Aug 28;164(4):541-548. [doi: 10.1007/s00221-005-2280-3] [Medline: 15864565]
20. Verlinden VJ, van der Geest JN, Hofman A, Ikram MA. Cognition and gait show a distinct pattern of association in the general population. Alzheimers Dement 2014 May 12;10(3):328-335 [FREE Full text] [doi: 10.1016/j.jalz.2013.03.009] [Medline: 23849591]
21. Noh B, Youm C, Lee M, Park H. Associating gait phase and physical fitness with global cognitive function in the aged. Int J Environ Res Public Health 2020 Jul 03;17(13):4786 [FREE Full text] [doi: 10.3390/ijerph17134786] [Medline: 32635202]
22. Ayers E, Verghese J. Gait dysfunction in motoric cognitive risk syndrome. J Alzheimers Dis 2019 Sep 19;71(s1):S95-S103 [FREE Full text] [doi: 10.3233/JAD-181227] [Medline: 30909242]
23. Mazoteras Muñoz V, Abellan van Kan G, Cantet C, Cortes F, Ousset PJ, Rolland Y, et al. Gait and balance impairments in Alzheimer disease patients. Alzheimer Dis Assoc Disord 2010 Jan;24(1):79-84. [doi: 10.1097/WAD.0b013e3181c78a20] [Medline: 20220324]
24. Kearney FC, Harwood RH, Gladman JR, Lincoln N, Masud T. The relationship between executive function and falls and gait abnormalities in older adults: a systematic review. Dement Geriatr Cogn Disord 2013 May 23;36(1-2):20-35 [FREE Full text] [doi: 10.1159/000350031] [Medline: 23712088]
25. Lindh-Rengifo M, Jonasson SB, Ullén S, Stomrud E, Palmqvist S, Mattsson-Carlgren N, et al. Components of gait in people with and without mild cognitive impairment. Gait Posture 2022 Mar;93:83-89 [FREE Full text] [doi: 10.1016/j.gaitpost.2022.01.012] [Medline: 35101749]
26. Iosa M, Marro T, Paolucci S, Morelli D. Stability and harmony of gait in children with cerebral palsy. Res Dev Disabil 2012 Jan;33(1):129-135. [doi: 10.1016/j.ridd.2011.08.031] [Medline: 22093657]

27. Peppe A, Chiavalon C, Pasqualetti P, Crovato D, Caltagirone C. Does gait analysis quantify motor rehabilitation efficacy in Parkinson's disease patients? Gait Posture 2007 Sep;26(3):452-462. [doi: 10.1016/j.gaitpost.2006.11.207] [Medline: 17240143]
28. Kuan T, Tsou J, Su F. Hemiplegic gait of stroke patients: the effect of using a cane. Arch Phys Med Rehabil 1999 Jul;80(7):777-784. [doi: 10.1016/s0003-9993(99)90227-7] [Medline: 10414762]
29. Wang X, Wang Y. Gait analysis of children with spastic hemiplegic cerebral palsy. Neural Regen Res 2012 Jul 15;7(20):1578-1584 [FREE Full text] [doi: 10.3969/j.issn.1673-5374.2012.20.008] [Medline: 25657696]
30. Iosa M, Fusco A, Marchetti F, Morone G, Caltagirone C, Paolucci S, et al. The golden ratio of gait harmony: repetitive proportions of repetitive gait phases. Biomed Res Int 2013 Jun 4;2013:918642-918647 [FREE Full text] [doi: 10.1155/2013/918642] [Medline: 23862161]
31. Okabe T. Physical phenomenology of phyllotaxis. J Theor Biol 2011 Jul 07;280(1):63-75. [doi: 10.1016/j.jtbi.2011.03.037] [Medline: 21497163]
32. Yamagishi MEB, Shimabukuro AI. Nucleotide frequencies in human genome and fibonacci numbers. Bull Math Biol 2008 Apr 10;70(3):643-653. [doi: 10.1007/s11538-007-9261-6] [Medline: 17994268]
33. Staff L, Hurd P, Reale L, Seoighe C, Rockwood A, Gehring C. The hidden geometries of the Arabidopsis thaliana epidermis. PLoS One 2012 Sep 11;7(9):e43546 [FREE Full text] [doi: 10.1371/journal.pone.0043546] [Medline: 22984433]
34. Ferring V, Pancherz H. Divine proportions in the growing face. Am J Orthod Dentofacial Orthop 2008 Oct;134(4):472-479. [doi: 10.1016/j.ajodo.2007.03.027] [Medline: 18929263]
35. Russell PA. The aesthetics of rectangle proportion: effects of judgment scale and context. Am J Psychol 2000;113(1):27-42. [Medline: 10742842]
36. Alothmany N, Khan A, Alnowaimi M, Morfeq AH, Hafez EA. Accuracy of joint angles tracking using markerless motion system. Journal of LaTeX Templates. Preprint posted online on September 21, 2014 2023 [FREE Full text]
37. Struzik A, Konieczny G, Stawarz M, Grzesik K, Winiarski S, Rokita A. Relationship between lower limb angular kinematic variables and the effectiveness of sprinting during the acceleration phase. Appl Bionics Biomech 2016 Jul 19;2016:7480709-7480709 [FREE Full text] [doi: 10.1155/2016/7480709] [Medline: 27516724]
38. Szpala A, Kołodziej M, Struzik A, Jasiński R, Bałchanowski KJ, Pietraszewski B, et al. Selected spatiotemporal and joint angle parameters in normal gait and nordic walking with classical and mechatronic poles in aspects of sex differences. Biomed Res Int 2022 Dec 13;2022:7905120-7905129 [FREE Full text] [doi: 10.1155/2022/7905120] [Medline: 36560964]
39. Perry J, Slack T, Davids JR. Gait analysis: normal and pathological function. J Pediatr Orthop 1992 Nov;12(6):815. [doi: 10.1097/01241398-199211000-00023]
40. Winter D, Patla AE, Frank JS, Walt SE. Biomechanical walking pattern changes in the fit and healthy elderly. Phys Ther 1990 Jun;70(6):340-347. [doi: 10.1093/ptj/70.6.340] [Medline: 2345777]
41. Kirtley C. Clinical Gait Analysis: Theory and Practice. London, UK: Elsevier Health Sciences; 2006.
42. Kang Y, Na DL, Hahn S. Seoul Neuropsychological Screening Battery-Core (SNSB-C). Incheon, Republic of Korea: Human Brain Research & Consulting Co; 2015.
43. Jahng S, Na DL, Kang Y. Constructing a composite score for the Seoul Neuropsychological Screening Battery-Core. Dement Neurocognitive Disord 2015 Dec 31;14(4):137-142. [doi: 10.12779/dnd.2015.14.4.137]
44. Iosa M, Mazzà C, Frusciante R, Zok M, Aprile I, Ricci E, et al. Mobility assessment of patients with facioscapulohumeral dystrophy. Clin Biomech (Bristol, Avon) 2007 Dec;22(10):1074-1082. [doi: 10.1016/j.clinbiomech.2007.07.013] [Medline: 17850940]
45. Shemmell J, Johansson J, Portra V, Gottlieb GL, Thomas JS, Corcos DM. Control of interjoint coordination during the swing phase of normal gait at different speeds. J Neuroeng Rehabil 2007 Apr 27;4(1):10 [FREE Full text] [doi: 10.1186/1743-0003-4-10] [Medline: 17466065]
46. Ni L, Lv W, Sun D, Sun Y, Sun Y, Xu X, et al. Pathological gait signatures of post-stroke dementia with toe-off and heel-to-ground angles discriminate from Alzheimer's disease. Front Aging Neurosci 2021 Nov 18;13:766884 [FREE Full text] [doi: 10.3389/fnagi.2021.766884] [Medline: 34867293]
47. Mc Ardle R, Morris R, Wilson J, Galna B, Thomas AJ, Rochester L. What can quantitative gait analysis tell us about dementia and its subtypes? a structured review. J Alzheimers Dis 2017 Nov 7;60(4):1295-1312. [doi: 10.3233/JAD-170541] [Medline: 29036826]
48. Allali G, Montembeault M, Brambati SM, Bherer L, Blumen HM, Launay CP, et al. Brain structure covariance associated with gait control in aging. J Gerontol A Biol Sci Med Sci 2019 Apr 23;74(5):705-713. [doi: 10.1093/gerona/gly123] [Medline: 29846517]
49. Verghese J, Robbins M, Holtzer R, Zimmerman M, Wang C, Xue X, et al. Gait dysfunction in mild cognitive impairment syndromes. J Am Geriatr Soc 2008 Jul;56(7):1244-1251 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01758.x] [Medline: 18482293]
50. Wilson J, Allcock L, Mc Ardle R, Taylor J, Rochester L. The neural correlates of discrete gait characteristics in ageing: a structured review. Neurosci Biobehav Rev 2019 May;100:344-369 [FREE Full text] [doi: 10.1016/j.neubiorev.2018.12.017] [Medline: 30552912]

51. Scherder E, Eggermont L, Swaab D, van Heuvelen M, Kamsma Y, de Greef M, et al. Gait in ageing and associated dementias; its relationship with cognition. Neurosci Biobehav Rev 2007 Jan;31(4):485-497. [doi: 10.1016/j.neubiorev.2006.11.007] [Medline: 17306372]
52. Zimmerman ME, Lipton RB, Pan JW, Hetherington HP, Verghese J. MRI- and MRS-derived hippocampal correlates of quantitative locomotor function in older adults. Brain Res 2009 Sep 29;1291:73-81 [FREE Full text] [doi: 10.1016/j.brainres.2009.07.043] [Medline: 19631621]
53. Beauchet O, Launay CP, Sekhon H, Montembeault M, Allali G. Association of hippocampal volume with gait variability in pre-dementia and dementia stages of Alzheimer disease: results from a cross-sectional study. Exp Gerontol 2019 Jan;115:55-61. [doi: 10.1016/j.exger.2018.11.010] [Medline: 30447261]
54. Takakusaki K. Neurophysiology of gait: from the spinal cord to the frontal lobe. Mov Disord 2013 Sep 15;28(11):1483-1491. [doi: 10.1002/mds.25669] [Medline: 24132836]
55. Tian Q, Chastan N, Bair W, Resnick SM, Ferrucci L, Studenski SA. The brain map of gait variability in aging, cognitive impairment and dementia-a systematic review. Neurosci Biobehav Rev 2017 Mar;74(Pt A):149-162 [FREE Full text] [doi: 10.1016/j.neubiorev.2017.01.020] [Medline: 28115194]
56. Ivanenko YP, Cappellini G, Dominici N, Poppele RE, Lacquaniti F. Modular control of limb movements during human locomotion. J Neurosci 2007 Oct 10;27(41):11149-11161 [FREE Full text] [doi: 10.1523/JNEUROSCI.2644-07.2007] [Medline: 17928457]
57. Zehr EP. Neural control of rhythmic human movement: the common core hypothesis. Exerc Sport Sci Rev 2005 Jan;33(1):54-60. [Medline: 15640722]
58. Ivanenko Y, Poppele R, Lacquaniti F. Distributed neural networks for controlling human locomotion: lessons from normal and SCI subjects. Brain Res Bull 2009 Jan 15;78(1):13-21. [doi: 10.1016/j.brainresbull.2008.03.018] [Medline: 19070781]
59. Takahashi H, Takada Y, Nagai N, Urano T, Takada A. Serotonergic neurons projecting to hippocampus activate locomotion. Brain Res 2000 Jun 30;869(1-2):194-202. [doi: 10.1016/s0006-8993(00)02385-4] [Medline: 10865074]
60. Iacono RP, Kuniyoshi SM, Ahlman JR, Zimmerman GJ, Maeda G, Pearlstein RD. Concentrations of indoleamine metabolic intermediates in the ventricular cerebrospinal fluid of advanced Parkinson's patients with severe postural instability and gait disorders. J Neural Transm (Vienna) 1997 Apr;104(4-5):451-459. [doi: 10.1007/BF01277663] [Medline: 9295177]

Abbreviations

AD: Alzheimer disease
COWAT: Controlled Oral Word Association Test
IMU: inertial measurement unit
K-CWST: 60sec: the Korean version of Color Word Stroop Test: 60 seconds
OR: odds ratio
PD: Parkinson disease
PSD: poststroke dementia
RCFT: Rey Complex Figure Test
S-K-BNT: short version of the Korean-Boston Naming Test
SNSB-C: Seoul Neuropsychological Screening Battery-Core
SVLT DR:: Seoul Verbal Learning Test: Delayed Recall
TMT-E B:: Trail Making Test-Elderly: Part B

Edited by A Mavragani, T Sanchez; submitted 04.02.23; peer-reviewed by M Iosa, J Curcic; comments to author 01.03.23; revised version received 20.03.23; accepted 15.06.23; published 10.07.23.

Please cite as:
Choi JY, Ha SW, Jeong DE, Lee J, Kim D, Min JY, Min KB
Association Between the Loss of Gait Harmony and Cognitive Impairment: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e46264
URL: https://publichealth.jmir.org/2023/1/e46264
doi:10.2196/46264
PMID:37428538

Original Paper

Comorbidity Patterns in Patients Newly Diagnosed With Colorectal Cancer: Network-Based Study

Hang Qiu[1,2*], PhD; Liya Wang[1*], MM; Li Zhou[3], MPA; Xiaodong Wang[4], MD

[1]Big Data Research Center, University of Electronic Science and Technology of China, Chengdu, China

[2]School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China

[3]Health Information Center of Sichuan Province, Chengdu, China

[4]Department of Gastrointestinal Surgery, West China Hospital, Sichuan University, Chengdu, China

[*]these authors contributed equally

Corresponding Author:
Hang Qiu, PhD
School of Computer Science and Engineering
University of Electronic Science and Technology of China
No. 2006, Xiyuan Ave, West Hi-Tech Zone
Chengdu, 611731
China
Phone: 86 28 61830278
Fax: 86 28 61830278
Email: qiuhang@uestc.edu.cn

Abstract

Background: Patients with colorectal cancer (CRC) often present with multiple comorbidities, and many of these can affect treatment and survival. However, previous comorbidity studies primarily focused on diseases in commonly used comorbidity indices. The comorbid status of CRC patients with respect to the entire spectrum of chronic diseases has not yet been investigated.

Objective: This study aimed to systematically analyze all chronic diagnoses and diseases co-occurring, using a network-based approach and large-scale administrative health data, and provide a complete picture of the comorbidity pattern in patients newly diagnosed with CRC from southwest China.

Methods: In this retrospective observational study, the hospital discharge records of 678 hospitals from 2015 to 2020 in Sichuan Province, China were used to identify new CRC cases in 2020 and their history of diseases. We examined all chronic diagnoses using ICD-10 (International Classification of Diseases, 10th Revision) codes at 3 digits and focused on chronic diseases with >1% prevalence in at least one subgroup (1-sided test, *P*<.025), which resulted in a total of 66 chronic diseases. Phenotypic comorbidity networks were constructed across all CRC patients and different subgroups by sex, age (18-59, 60-69, 70-79, and ≥80 years), area (urban and rural), and cancer site (colon and rectum), with comorbidity as a node and linkages representing significant correlations between multiple comorbidities.

Results: A total of 29,610 new CRC cases occurred in Sichuan, China in 2020. The mean patient age at diagnosis was 65.6 (SD 12.9) years, and 75.5% (22,369/29,610) had at least one comorbidity. The most prevalent comorbidities were hypertension (8581/29,610, 29.0%; 95% CI 28.5%-29.5%), hyperplasia of the prostate (3816/17,426, 21.9%; 95% CI 21.3%-22.5%), and chronic obstructive pulmonary disease (COPD; 4199/29,610, 14.2%; 95% CI 13.8%-14.6%). The prevalence of single comorbidities was different in each subgroup in most cases. Comorbidities were closely associated, with disorders of lipoprotein metabolism and hyperplasia of the prostate mediating correlations between other comorbidities. Males and females shared 58.3% (141/242) of disease pairs, whereas male-female disparities occurred primarily in diseases coexisting with COPD, cerebrovascular diseases, atherosclerosis, heart failure, or renal failure among males and with osteoporosis or gonarthrosis among females. Urban patients generally had more comorbidities with higher prevalence and more complex disease coexistence relationships, whereas rural patients were more likely to have co-existing severe diseases, such as heart failure comorbid with the sequelae of cerebrovascular disease or COPD.

Conclusions: Male-female and urban-rural disparities in the prevalence of single comorbidities and their complex coexistence relationships in new CRC cases were not due to simple coincidence. The results reflect clinical practice in CRC patients and emphasize the importance of measuring comorbidity patterns in terms of individual and coexisting diseases in order to better understand comorbidity patterns.

(JMIR Public Health Surveill 2023;9:e41999) doi:10.2196/41999

KEYWORDS

colorectal cancer; comorbidity patterns; prevalence; health status disparities; network analysis; routinely collected health data

Introduction

Colorectal cancer (CRC) is the third most common malignancy and the second most deadly cancer globally, with 1.93 million new cases and 0.93 million deaths estimated in 2020 [1]. China has become one of the countries with the highest incidences of CRC cases and CRC-related deaths worldwide, as a consequence of lifestyle changes and accelerated aging of the population [2]. In 2020, newly diagnosed CRC cases in China accounted for 28.8% of new global cases and CRC-related deaths in China accounted for 30.6% of all CRC-related deaths worldwide [3].

Patients with CRC commonly present at an older age and have a relatively high proportion of other chronic diseases (ie, comorbidities). In a retrospective study in Korea, 49.6% of colon cancer patients had ≥3 comorbidities [4]. In an observational cohort study in the United States, 40% of CRC patients had 1 to 3 comorbidities and 19% had ≥4 comorbidities [5]. In addition, evidence is increasing that the coexistence of various chronic diseases has a substantial effect on the treatment, outcomes, and survival of CRC patients [6-9]. In a retrospective cohort study of 29,733 CRC patients, approximately 9% of deaths were attributable to congestive heart failure (CHF), more than 5% were attributable to chronic obstructive pulmonary disease (COPD), and nearly 4% were attributable to diabetes mellitus (DM) [10]. Several studies have also suggested that CRC patients burdened by comorbidities have higher mortality than those without coexisting diseases [8,11,12]. Thus, increasing the understanding about and attention to patterns of diseases that coexist with CRC is important for disease management and more personalized medicine.

To date, investigations on comorbidities associated with CRC have primarily focused on commonly used comorbidity indices, such as the Charlson comorbidity index and Elixhauser comorbidity index [5,13-15], or defined a comorbid condition as 1 of the 14 selected health conditions [16]. Although some comorbid conditions, such as hypertension and DM, are well recognized [5,15,17,18], many others are likely unidentified. A significant gap in knowledge remains regarding comorbidity patterns with CRC, especially among Chinese populations whose lifestyles and diets differ from those of Western populations.

In recent years, the development of network theory has provided new approaches to understand the complex interrelations between diseases [19-21]. The phenotypic comorbidity network (PCN) also known as the disease co-occurrence network has been widely used to study the comorbidity patterns of various chronic conditions, such as depression [22], ischemic heart disease (IHD) [23], pediatric pulmonary hypertension [24], migraine [25], COPD [26], hip fracture [27], and hepatocellular carcinoma [28]. Additionally, rapid growth in administrative health data (eg, hospital discharge data) offers opportunities to simultaneously assess the entire spectrum of diagnoses in comorbidity studies [29,30]. Administrative health data include a substantial amount of health information in the form of ICD-10 (International Classification of Diseases, 10th Revision) codes, which have tremendous potential to increase the understanding of the nature of comorbidities [30]. Nevertheless, to our knowledge, no study has yet applied network theory with regional administrative health data to systematically exploit the hidden information of comorbidity patterns in Chinese patients with CRC.

The aim of this study was to apply a network-based approach using routinely collected hospital discharge records (HDRs) to identify comorbidity patterns in a general population of newly diagnosed CRC patients. Comorbidity patterns were assessed by measuring the prevalence of individual diseases and the comorbid strength of coexisting diseases and by determining differences in sex, age at diagnosis, region, and cancer site.

Methods

Study Design and Participants

A retrospective cohort study was performed using longitudinal data from a provincial health administrative data set. The data set was mandatorily collected, and it included routinely collected anonymized HDRs from all secondary hospitals and tertiary hospitals in Sichuan Province, China, since 2015 [21]. HDRs are medical record documents with legal effect and cover detailed discharge diagnosis information, which are coded using standardized ICD-10 codes [31]. Each inpatient's longitudinal clinical data were available, including anonymized identity, age, sex, residential address, admission date, discharge date, principal discharge diagnosis, and up to 15 secondary diagnoses. Sichuan Province, one of the most populous provinces in China, is separated into 2 cultural/topographic sections by the Hengduan Mountains, thus providing a scaled-down model of the broader situation in China [32,33].

A previous validation study integrated HDRs in Sichuan Province to identify new cancer cases, and indicated no statistical significance with findings from population-based cancer registry data, but it had important strengths, including a large sample, high accuracy of diagnosis coding, and time effectiveness [34]. In this study, provincial HDRs were used to identify new cases of CRC, where a new case was defined as a patient without any CRC diagnosis code (ICD-10 codes: C18 to C21) for at least 5 years before the first CRC diagnosis [17]. To minimize bias, all inpatients living in Sichuan Province during the study period were considered, with the exception of those with incomplete information or incorrect codes. The inclusion criteria were permanent residence in Sichuan Province, age ≥18 years at diagnosis, and first CRC occurrence between January 1, 2020, and December 31, 2020. Patients with incomplete information, conflicting reports of sex among multiple hospitalizations, conflicts between diagnoses and sex, or a history of a malignant neoplasm but with an unknown or not reported cancer site were excluded. Based on the longitudinal health care data, 29,610 patients diagnosed with first CRC occurrence were identified between January 1, 2020, and December 31, 2020 (Figure 1).

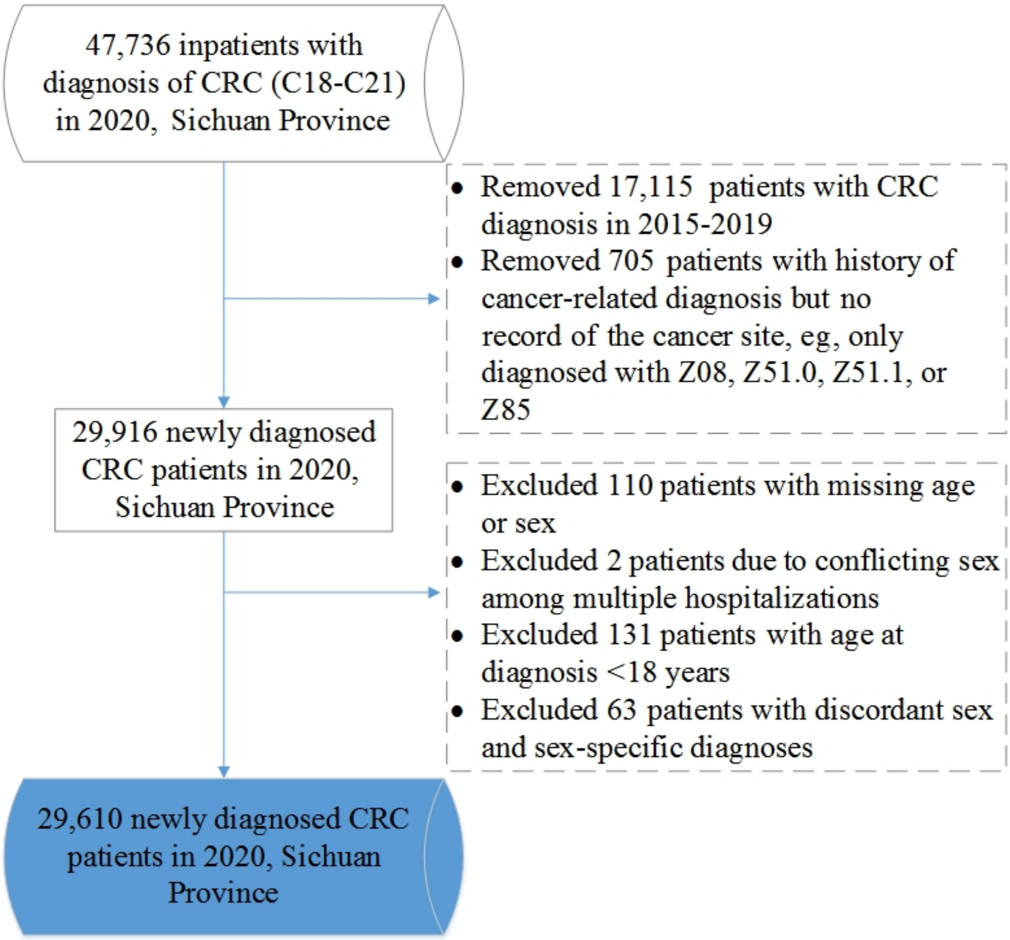


Figure 1. Flowchart of the study participants. CRC: colorectal cancer; Z08: follow-up examination after treatment for malignant neoplasms; Z51.0: radiotherapy session; Z51.1: chemotherapy session for neoplasm; Z85: personal history of malignant neoplasm.

Ethical Considerations

This study was approved by the Ethics Committee of the Health Information Center of Sichuan Province (ZX-EC202100301). The requirement of obtaining informed consent was waived because of the secondary nature of the deidentified data and the retrospective study design.

Definition of Comorbidity

Because chronic conditions are not expected to end in a single hospitalization period, a 5-year look-back period from the first diagnosis was used to identify comorbidities [17]. First, for each CRC patient, an index hospitalization was identified, which was the first hospitalization that occurred or was within 4 weeks after the date of the first diagnosis of CRC [17]. Second, the index hospitalization and hospitalizations during the 5 years before the first diagnosis of CRC were scanned to determine the history of diseases. Lastly, the Chronic Condition Indicator, defined by the Healthcare Cost and Utilization Project, was used to differentiate between acute and chronic ICD-10 codes at 3 digits [20,35]. Diagnostic codes from chapters XV to XXII, which are certain conditions originating in pregnancy, childbirth, puerperium, and the perinatal period or are general symptoms, were excluded, as the study population included mainly middle-aged or older patients. Notably, to exclude the complications of CRC or its treatment, some chronic conditions that occurred in the index hospitalization were excluded [17], including secondary malignant neoplasm, anemia, nutritional marasmus, agranulocytosis, metabolism disorders, and intestinal obstruction. To estimate the effects of different look-back periods on comorbidity patterns, we performed a complementary analysis using a 3-year look-back period.

Comorbidity Prevalence

The prevalence of all chronic diagnoses with ICD-10 codes and the corresponding 95% CIs were calculated. The prevalence cutoff was set to a statistically significant value of >1% (1-sided test, $P<.025$) in at least one subgroup to avoid diagnoses used very rarely or never [36]. The final list of 66 distinct comorbidities and their ICD-10 codes are provided in Multimedia Appendix 1. To assess differences in comorbidity prevalence by sex, region, and cancer site, relative differences in prevalence were calculated and then tested for significance using a Z-test with Bonferroni correction (α level of .05). With relative difference by sex as an example, the calculation was according to the following equation [20]:

where P_{female} and P_{male} are the prevalences of comorbidities in females and males, respectively.

If the relative difference was greater than 0.1 [28] and the absolute difference was significant after Bonferroni correction, there was a significant difference in comorbidity prevalence. Among comorbidities with a significant difference in prevalence, enriched comorbidities were assessed further. When a comorbidity had a ≥0.5-fold increase in prevalence, it was defined as enriched. For example, a comorbidity was enriched in females when there was a significant prevalence difference by sex and a ≥0.5-fold increase in prevalence in females compared with males. A Cochran-Armitage trend test was applied to assess whether the prevalence of comorbidities increased with age. Furthermore, a K-means clustering algorithm was applied to cluster comorbidities in terms of prevalence by age. As the prevalence of comorbidities increased with age, we further adjusted the influence of age in estimating the enrichment comorbidities. The total number of cases in each age group (18-49, 50-59, 60-69, 70-79, and ≥80 years) by sex, region, and cancer site was regarded as the standard population.

PCN Construction

A PCN was constructed to capture the coexistence of all comorbidities (also named connections) as recorded through HDRs. In a PCN, nodes represent chronic disease codes (ICD-10 codes at 3 digits) that are connected through edges. Node size is proportional to disease prevalence, and node color identifies the ICD-10 category. The cosine index was applied to quantify the comorbid strength of coexisting diseases, which considers the co-occurrence and prevalence of comorbidities and thus is not affected by the sample size [37]. To detect comorbidity coexistence measured by the cosine index, a cutoff was defined by assessing the relationship between the Pearson correlation and cosine index, in which the number of significant coexisting comorbidities was equal in both networks measured by the cosine index and Pearson correlation [21,38]. The cosine index and Pearson correlation coefficient were defined as in equations (2) and (3), respectively, and the significance of $\varphi_{ab} \neq 0$ was determined by performing a t test, which was calculated according to equation (4).

In the equations, N is the total number of newly diagnosed CRC patients, and n_a, n_b, and n_{ab} denote the number of patients with disease a, disease b, and both diseases, respectively.

Network Properties

The structural properties of a PCN are measured using network indices, such as network density, degree, average degree of neighbors, and betweenness centrality [39]. Network density is the proportion of significant connections to all possible connections in a network, which measures how compact a network is. A higher network density is associated with a higher number of connections between comorbidities. The number of correlations with other comorbidities is degree (k), which provides a general idea of how involved a comorbidity is with other comorbidities. Median and IQR were calculated to describe the distribution of degree (k). A Kolmogorov-Smirnov test was applied to estimate whether the degree (k) distribution followed a power law. When one comorbidity was directly connected with other comorbidities, the others were named as neighbors. The average degree of neighbors was calculated to measure neighbor connectivity. Betweenness centrality, which is the number of shortest paths between any 2 nodes of which a node is a part divided by all possible paths, indicates how central one comorbidity is relative to all other comorbidities. A high betweenness centrality indicates a high likelihood of forming bridges between other comorbidities or the end points of many comorbidities [25]. To identify the most important comorbidity in a PCN, the PageRank algorithm that considers the edge weight was applied [36]. A higher PageRank value of a comorbidity is associated with a greater effect on the network. Comorbidities with the top 10 percentiles of PageRank values were defined as the most important comorbidities in PCNs.

Abundant Comorbidity Connections in Each Subgroup of CRC Patients

PCNs were constructed separately for males and females, for rural and urban patients, and for rectal and colon cancer patients. Then, comorbid strengths were compared to measure disparities by sex, region, and cancer site. When a coexisting disease was unique to or enriched (with 0.05 higher comorbid strength than another) in a given subgroup, the difference was notable and thus was defined as an abundant connection. With sex as an example, when a comorbidity connection occurred only in males or occurred in both subgroups but the comorbid strength was 0.05 (approximately half of the minimum comorbid strength in the PCN) higher in males than in females, the comorbidity connection was abundant in males.

All analyses and visualizations were conducted using R 4.0.3 (R Foundation for Statistical Computing).

Results

Characteristics of Newly Diagnosed CRC Patients

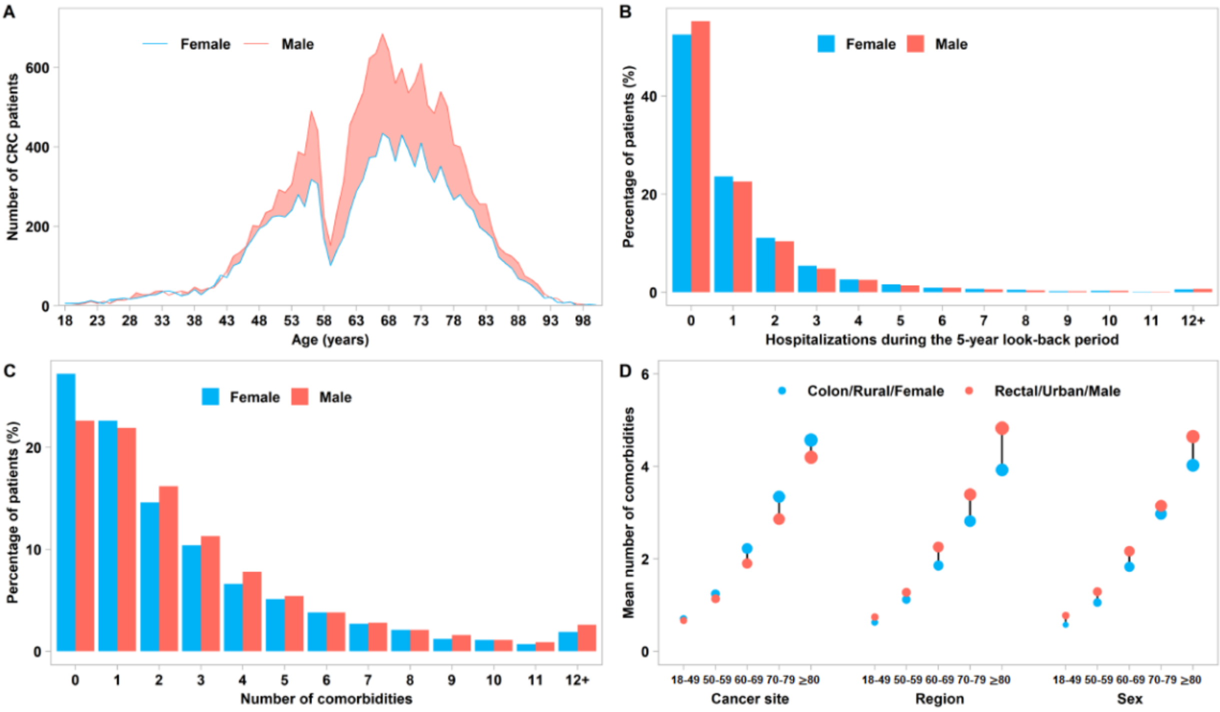
Table 1 shows the characteristics of the 29,610 newly diagnosed CRC patients. The mean patient age at diagnosis was 65.6 years. Among the patients, 30.1% (8922/29,610) were diagnosed at <60 years of age, 58.9% (17,426/29,610) were male, 45.7% (13,522/29,610) lived in urban regions, and 52.5% (15,543/29,610) had rectal cancer. Overall, approximately one-fourth (7241/29,610, 24.5%) of patients were without any comorbidity, 22.2% (6574/29,610) had 1 comorbidity, and more than half (15,795/29,610, 53.3%) had ≥2 comorbidities. The sex-specified distributions of age at diagnosis, number of hospitalizations during the 5-year look-back period, and number of comorbidities are presented in Figure 2A-C. The frequency of at least one comorbidity was significantly higher in male patients than in female patients (13,495/17,426, 77.4% vs 8874/12,184, 72.8%; $P<.001$), in urban patients than in rural patients (10,542/13,522, 78.0% vs 11,746/15,988, 73.5%;

P<.001), and in colon cancer patients than in rectal cancer patients (10,115/13,064, 77.4% vs 11,528/15,543, 74.2%; *P*<.001). The proportion of patients with at least one comorbidity increased with age (Cochran-Armitage trend test, *P*<.001), and the increasing trend was similar in each subgroup. The mean number of comorbidities by cancer site, region, and sex increased with age (Figure 2D).

Table 1. Descriptive characteristics of patients newly diagnosed with colorectal cancer in 2020 in Sichuan Province, China.

Demographic and clinical factors	Value (N=29,610)
Age at diagnosis (years), mean (SD)	65.6 (12.9)
Age group at diagnosis (years), n (%)	
18-49	3384 (11.4)
50-59	5538 (18.7)
60-69	8308 (28.1)
70-79	8582 (29.0)
≥80	3798 (12.8)
Sex, n (%)	
Female	12,184 (41.1)
Male	17,426 (58.9)
Region, n (%)	
Rural	15,988 (54.0)
Urban	13,522 (45.7)
Unknown	100 (0.3)
Cancer site, n (%)	
Colon	13,064 (44.1)
Rectum	15,543 (52.5)
Anus	813 (2.7)
Multiple sites	190 (0.6)
Number of comorbidities, n (%)	
0	7241 (24.5)
1	6574 (22.2)
2	4595 (15.5)
3	3241 (10.9)
4	2172 (7.3)
5	1563 (5.3)
≥6	4224 (14.3)

Figure 2. Number of comorbidities and hospitalizations for newly diagnosed colorectal cancer (CRC) patients in 2020 in Sichuan Province, China. (A) Age at diagnosis by sex. (B) Hospitalizations during the 5-year look-back period before diagnosis. (C) Frequency of patients per distinct number of comorbidities per patient among male and female patients. (D) Age-specific mean number of comorbidities by cancer site, region, and sex.



Comorbidity Prevalence and Differences Between Subgroups

Among the 66 distinct comorbidities (Multimedia Appendix 1), the most prevalent comorbidities were hypertension (8581/29,610, 29.0%; 95% CI 28.5%-29.5%), hyperplasia of the prostate (3816/17,426, 21.9%; 95% CI 21.3%-22.5%), COPD (4199/29,610, 14.2%; 95% CI 13.8%-14.6%), DM (3715/29,610, 12.5%; 95% CI 12.2%-12.9%), and IHD (3200/29,610, 10.8%; 95% CI 10.5%-11.2%). Comorbidity prevalence using a 3-year look-back period is shown in Multimedia Appendix 2, and a scatter plot of comorbidity prevalence using 3-year and 5-year look-back periods is shown in Multimedia Appendix 3. As for subgroup difference, most comorbidities were common in each subgroup, whereas a few comorbidities were only detected in a subgroup and with relatively low prevalence. For example, anxiety disorder (ICD-10: F41) was only detected in females, with a prevalence of 1.3% (156/12,184; 95% CI 1.1%-1.5%). Among comorbidities with significant differences in prevalence (Multimedia Appendices 4-6), differences were large and some comorbidities were enriched (prevalence increased by ≥50%) in a subgroup (Table 2). For example, COPD (ICD-10: J43 and J44), renal failure (N18 and N19), conduction disorders (I44 and I45), and gout (M10) were enriched in males, whereas osteoporosis (M81), gonarthrosis (M17), other nontoxic goiter (E04), and anxiety disorder (F41) were enriched in females. Most comorbidities had higher prevalence and were even enriched (prevalence increased by ≥50%) in urban patients, for example, atherosclerosis (ICD-10: I70), cerebrovascular diseases (I65 and I67), hypertensive heart disease (I11), diverticular disease of the intestine (K57), and spondylosis (M47). Comorbidities had somewhat higher prevalence in colon cancer patients than in rectal cancer patients, and disorders of glycoprotein metabolism (ICD-10: E77), other coagulation defects (D68), angina pectoris (I20), other diseases of the biliary tract (K83), and malignant neoplasm of the liver and intrahepatic bile ducts (C22) were enriched in colon cancer patients. After adjustment with age, osteoporosis (ICD-10: M81) was not enriched in urban patients, and the other comorbidities still showed enrichment (Multimedia Appendix 7).

For age-specific comorbidities, 62 comorbidities with a prevalence significantly greater than 1% in at least one age group were divided into 4 clusters (Figure 3). Comorbidity prevalence increased with age, but the rate of increase varied by cluster. Cluster 1 contained 36 comorbidities with low prevalence (the average prevalence among the 5 age groups ranged from 0.8% to 2.2%), and most had a low growth rate with age, but a few decreased with age (eg, endometriosis) or were stable across ages (eg, chronic sinusitis). Cluster 2 contained 17 comorbidities with moderate prevalence (the average prevalence among the 5 age groups ranged from 1.3% to 8.1%), and most had a relatively high growth rate with age. Cluster 3 contained 7 highly prevalent comorbidities, most of which had a high growth rate with age (the average prevalence among the 5 age groups ranged from 1.4% to 21.5%), including COPD (ICD-10: J44 and J43), IHD (I25), cerebral infarction (I63), heart failure (I50), atherosclerosis (I70), and DM (E11). Cluster 4 contained the 2 comorbidities with the highest prevalence (hypertension [ICD-10: I10] and hyperplasia of the prostate [N40]).

Table 2. Absolute prevalence differences of enrichment comorbidities in subgroups of colorectal cancer patients in Sichuan Province, China.

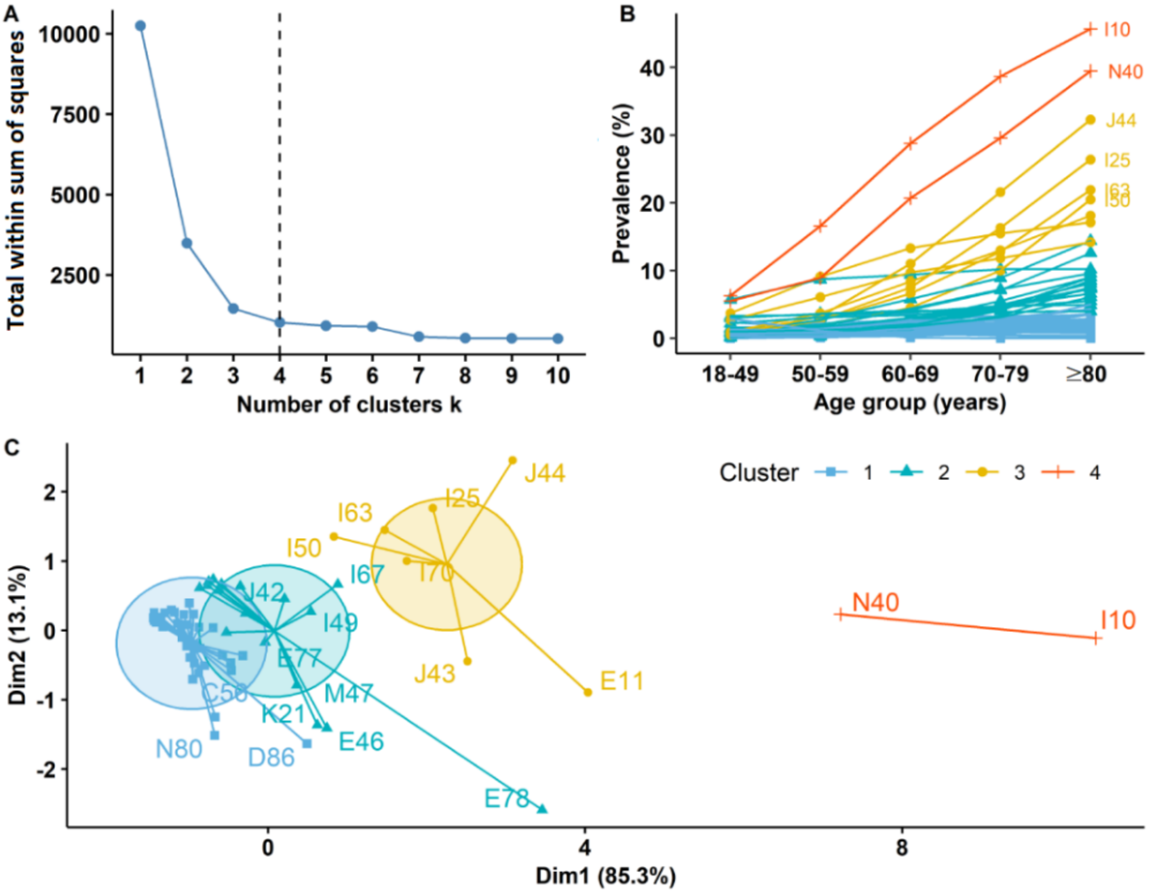
ICD-10 code	Enrichment comorbidity	Prevalence, value (95% CI)	Absolute prevalence difference, value (95% CI)[a]		
			Male-female	Urban-rural	Colon-rectal
J44	COPD[b]	14.2 (13.8 to 14.6)	7.9 (7.2 to 8.7)	—[c]	—
J43	Emphysema	9.4 (9.1 to 9.8)	6.9 (6.3 to 7.5)	—	—
I70	Atherosclerosis	9.2 (8.9 to 9.6)	—	4.2 (3.5 to 4.9)	—
I67	Other cerebrovascular diseases	6.7 (6.4 to 6.9)	—	3.6 (3.0 to 4.2)	—
K21	Gastroesophageal reflux disease	3.8 (3.6 to 4.0)	—	2.3 (1.9 to 2.8)	—
M47	Spondylosis	3.8 (3.6 to 4.0)	—	1.7 (1.2 to 2.1)	—
E77	Disorders of glycoprotein metabolism	3.6 (3.4 to 3.8)	—	—	1.8 (1.3 to 2.2)
M81	Osteoporosis without pathological fracture	3.1 (2.9 to 3.3)	−3.5 (−3.9 to −3.0)	—	—
I11	Hypertensive heart disease	3.0 (2.8 to 3.2)	—	1.6 (1.2 to 2.0)	—
N19	Unspecified renal failure	2.5 (2.3 to 2.7)	1.2 (0.9 to 1.6)	—	—
E04	Other nontoxic goiter	2.2 (2.0 to 2.3)	−2.0 (−2.3 to −1.6)	1.1 (0.7 to 1.4)	—
D68	Other coagulation defects	2.1 (1.9 to 2.2)	—	—	1.0 (0.7 to 1.4)
K83	Other diseases of the biliary tract	2.1 (2.0 to 2.3)	—	—	0.9 (0.6 to 1.3)
I65	Occlusion and stenosis of precerebral arteries	1.9 (1.8 to 2.1)	—	1.1 (0.8 to 1.4)	—
N18	Chronic renal failure	1.7 (1.6 to 1.9)	0.7 (0.4 to 1.0)	1.1 (0.8 to 1.4)	—
K57	Diverticular disease of the intestine	1.5 (1.4 to 1.6)	—	1.5 (1.4 to 1.6)	—
M17	Gonarthrosis	1.5 (1.4 to 1.7)	−1.3 (−1.6 to −1.0)	0.8 (0.5 to 1.1)	
I20	Angina pectoris	1.3 (1.2 to 1.4)	—	0.7 (0.5 to 1.0)	0.6 (0.4 to 0.9)
I44	Atrioventricular and left bundle-branch block	1.3 (1.2 to 1.4)	0.8 (0.5 to 1.0)	—	—
I45	Other conduction disorders	1.2 (1.1 to 1.4)	0.7 (0.5 to 1.0)	—	—
M10	Gout	1.2 (1.1 to 1.3)	1.5 (1.3 to 1.7)	0.6 (0.3 to 0.8)	—
E27	Other disorders of the adrenal gland	1.1 (1.0 to 1.3)	—	0.7 (0.5 to 0.9)	—
C22	Malignant neoplasm of the liver and intrahepatic bile ducts	1.1 (1.0 to 1.3)	—	—	0.5 (0.2 to 0.7)
F41	Other anxiety disorders	0.8 (0.7 to 0.9)	−0.8 (−1.1 to −0.6)	—	—

[a] Absolute prevalence differences by sex, region, and cancer site were statistically significant after Bonferroni correction.

[b] COPD: chronic obstructive pulmonary disease.

[c] The comorbidity was not enriched in this subgroup.

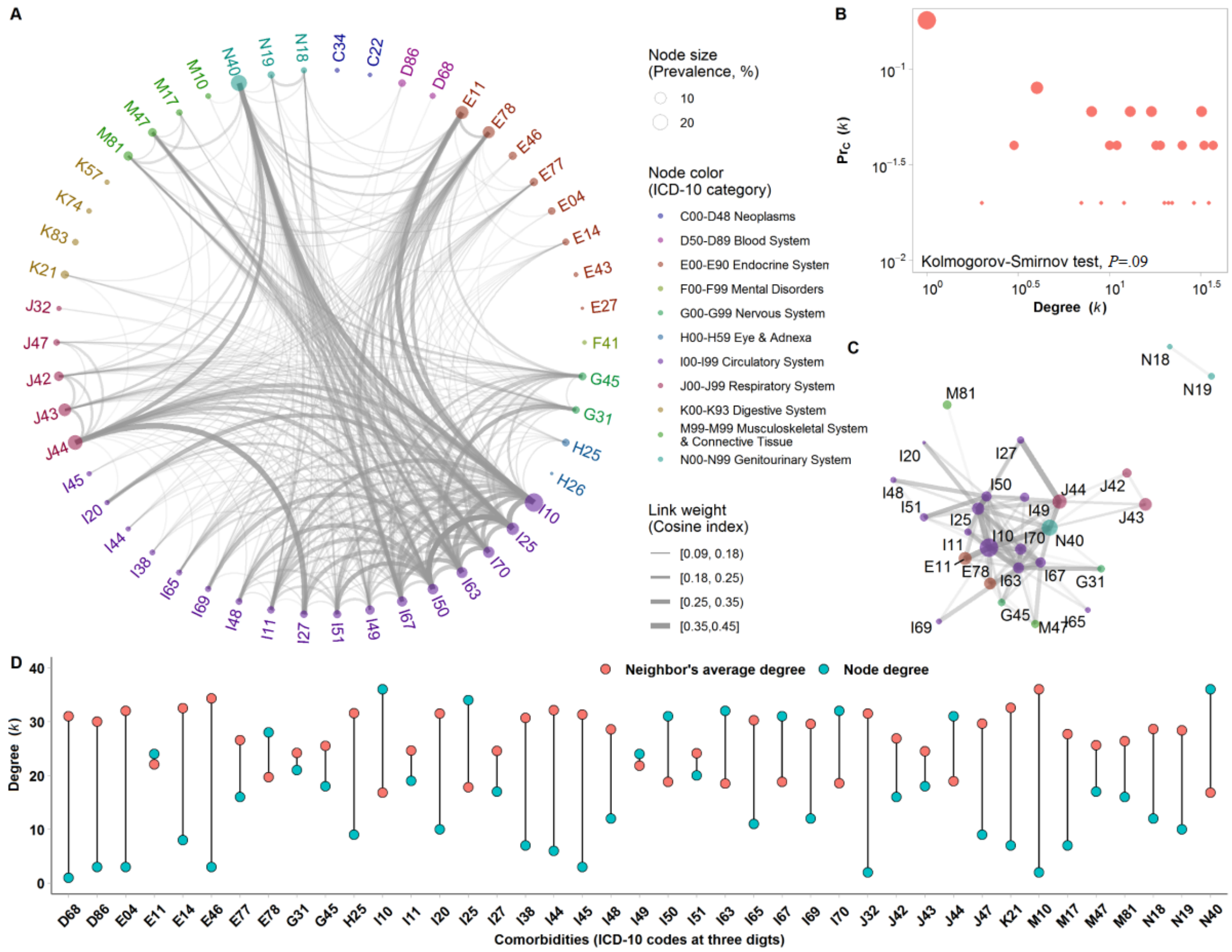
Figure 3. Clusters of comorbidities in colorectal cancer patients based on age-specific prevalence in Sichuan Province, China. (A) Optimal number of clusters using K-means clustering algorithm. (B) Age-specific prevalence of comorbidities in each cluster. Here, comorbidities with a prevalence of >20% in the ≥80 years age group were labeled with ICD-10 (International Classification of Diseases, 10th Revision) codes at 3 digits. (C) Cluster plot. Based on the principal component analysis, comorbidity prevalence in 5 dimensions (18-49, 50-59, 60-69, 70-79, and ≥80 years) was reduced to 2 dimensions (x-lab and y-lab). The ICD-10 codes are clarified in Multimedia Appendix 1.



PCN in Newly Diagnosed CRC Patients

The PCN of newly diagnosed CRC patients contained 50 diseases and 327 significant comorbid disease pairs (Figure 4). The number of significant correlations shared with other comorbidities per diagnosis code was 10.5 (IQR 3.0-19.8). Hypertension (ICD-10: I10), IHD (I25), hyperplasia of the prostate (N40), cerebral infarction (I63), and heart failure (I50) were the most important comorbidities in the PCN, indicating that they often coexisted with one another, had the highest number of coexisting comorbidities, and thus largely increased the complexity of coexisting diseases. Among the top 100 significant comorbid disease pairs with a cosine index ≥0.18, 51 connections occurred within the same disease system (eg, 45 connections occurred within circulatory system diseases) and 49 connections involved 2 disease systems. Across betweenness centrality measures, comorbidities showed the highest values in disorders of lipoprotein metabolism (ICD-10: E78), hyperplasia of the prostate (N40), COPD (J44), other cerebrovascular diseases (I67), and heart failure (I50). The results indicated that those comorbidities mediated correlations between others or were the end points of many comorbidities. When those 5 comorbidities (10% of comorbidities in the PCN) were removed, the scale of the PCN decreased considerably, with the number of significant comorbid disease pairs decreasing by 45%. Notably, hypertensive heart disease (ICD-10: I11), one of the 30 comorbidities with fewer coexisting diseases than their neighbors, connected with all 10 of the most important diseases and thus ranked as the 11th most important disease in the network.

Figure 4. Phenotypic comorbidity network (PCN) in newly diagnosed colorectal cancer (CRC) patients in Sichuan Province, China. (A) The PCN in newly diagnosed CRC patients. Nodes represent comorbidities (ICD-10 [International Classification of Diseases, 10th Revision] codes at 3 digits), such that the node size is proportional to the comorbidity prevalence in CRC patients and its color identifies the ICD-10 category. Link weights are proportional to the magnitudes of the cosine index. (B) Cumulative degree (k) distribution. The degree distribution showed an exponential decay when the degree was ≥12 (Kolmogorov-Smirnov test, *P*=.09). (C) The top 100 edges in the PCN. Here, the top 100 links where cosine index values were ≥0.18 are shown. (D) The degree distribution of the node and its neighbors. Here, the connected nodes (degree >0) in the PCN are shown. The ICD-10 codes are clarified in Multimedia Appendix 1.



Differences in PCNs Between Subgroups

Table 3 shows the network structures of PCNs constructed separately for males and females, for different age groups, for rural and urban patients, and for rectal cancer patients and colon cancer patients. The number of comorbidities with a prevalence of >1% in early onset CRC patients (diagnosis age of 18-49 years) was 15, and comorbidities were not significantly connected. Comorbidity coexistence relationships were more complex in males than females, older patients than those aged 50-59 years, and urban patients than rural patients. Male and female patients had 46 and 43 comorbidities, respectively, with median numbers of connections for comorbidities of 10.5 (IQR 2.0-17.5) and 6 (IQR 0.5-12.5), respectively, and median numbers of connections for node's neighbors of 23.2 (IQR 19.0-25.5) and 17.9 (IQR 15.0-19.0), respectively. Except in the subgroup of CRC patients with a diagnosis age of 50-59 years, the most important comorbidities in the PCNs were similar, with hypertension (ICD-10: I10), IHD (I25), and cerebral infarction (I63) common to each subgroup. In early onset CRC patients (diagnosis age of 18-59 years), the most important comorbidities in the PCN were hypertension (ICD-10: I10) and disorders of lipoprotein metabolism (E78).

Figure 5 shows the abundant connections in a given subgroup of newly diagnosed CRC patients. Without considering sex-specific diseases, 141 connections were common to both sexes, but 9 and 7 connections were separately enriched in males and females (cosine index increased more than 0.05), respectively, and 71 and 14 connections were significant only in males and females, respectively. Abundant connections in females were primarily associated with osteoporosis (ICD-10: M81) or gonarthrosis (M17), whereas abundant connections in males were primarily associated with COPD (J42, J43, J44, and J47), cerebrovascular diseases (I65 and I69), atherosclerosis (I70), heart failure (I50), or renal failure (N18 and N19). Urban-rural disparities in comorbidity connections were very large, with 13 (13/260, 5%) and 117 (117/260, 45%) abundant connections in rural and urban patients, respectively. Abundant connections in urban patients were primarily associated with disorders of lipoprotein metabolism (ICD-10: E78), DM (E11), hypertension (I10), hypertensive heart disease (I11), IHD (I20 and I25), cardiac arrhythmias (I44, I48, and I49), cerebrovascular diseases (I63, I65, I67, and I69), atherosclerosis (I70), gonarthrosis (M17), spondylosis (M47), osteoporosis (M81), renal failure (N18), or hyperplasia of the prostate (N40). Differences in connections according to cancer site were relatively small, with 181 (181/250, 72.4%) connections common to both colon cancer and rectal cancer patients.

Table 3. Phenotypic comorbidity network structures in subgroups of newly diagnosed colorectal cancer patients in Sichuan Province, China.

Subgroup	Nodes, n	Density	Degree, median (IQR)	Degree of neighbors, median (IQR)	Most important diseases in the PCN[a,b]
Sex					
Female	43	0.189	6.0 (0.5-12.5)	17.9 (15.0-19.0)	I10, I25, I63, and I50
Male	46	0.251	10.5 (2.0-17.5)	23.2 (19.0-25.5)[c]	I10, I63, N40, and I25
Age group (years)					
18-49	15	0	0	0	—[d]
50-59	26	0.043	0 (0-1.1)	3.0 (2.8-3.4)	E11 and I10
60-69	42	0.099	1.5 (0-7.8)[c]	11.2 (8.5-13.5)[c]	I10, I63, and I50
70-79	52	0.122	3.0 (0-11.0)[c]	17.1 (13.3-19.2)[c]	I10, I63, I25, and N40
≥80	58	0.099	1.0 (0-9.8)[c]	18.6 (13.5-21.0)[c]	I10, N40, I63, and I25
Region					
Rural	46	0.178	6.0 (0-12.8)	18.5 (15.1-20.0)	I10, I25, I50, and I63
Urban	50	0.202	8.0 (1.0-15.8)	21.7 (18.4-24.7)[c]	I10, I63, N40, and I25
Cancer site					
Rectal	44	0.218	8.5 (1.0-14.0)	19.1 (16.6-20.3)	I10, I25, I63, and I50
Colon	52	0.179	7.0 (0-16.0)	20.3 (17.6-23.0)	I10, I25, N40, and I63

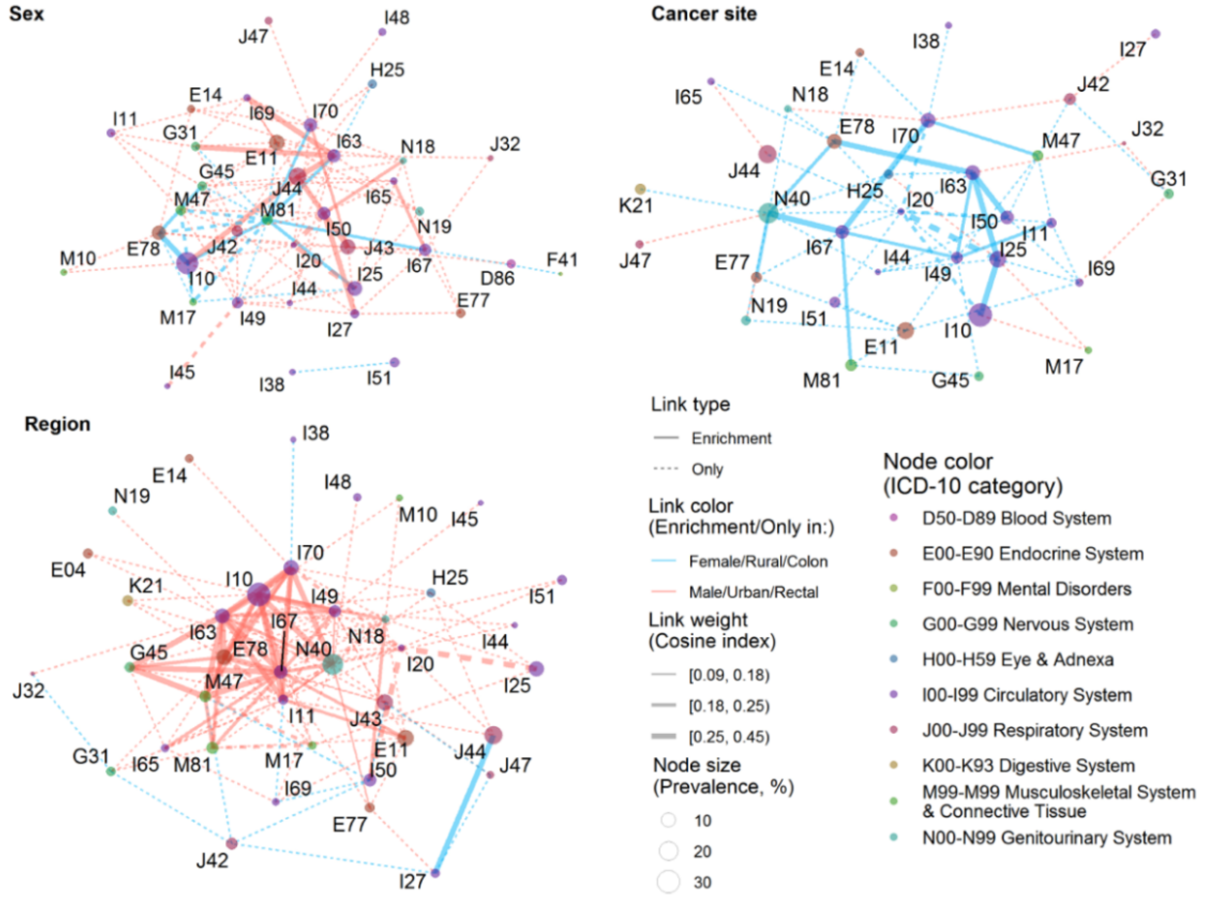
[a] PCN: phenotypic comorbidity network.

[b] Comorbidities with the top 10 percentiles of PageRank values were defined as the most important comorbidities in the PCN, and they were sorted by descending PageRank values. The ICD-10 codes and corresponding comorbidities were are follows: I10, hypertension; I25, chronic ischemic heart disease; I63, cerebral infarction; I50, heart failure; N40, hyperplasia of the prostate; E78, disorders of lipoprotein metabolism.

[c] Greater average degree of nodes or average degree of neighbors in males, older patients, and urban patients than in females, patients aged 50-59 years, and rural patients, respectively (1-sided *t* test with $P<.025$).

[d] Comorbidities were not connected to colorectal cancer patients aged 18-49 years at diagnosis.

Figure 5. Abundant connections by sex, region, and cancer site in newly diagnosed colorectal cancer patients in Sichuan Province, China. Except for those common to both subgroups (link weight difference <0.05), disease pairs were identified as abundant edges in one subgroup, including enrichment (solid lines) or only occurrence (dotted lines). For example, if the link weight difference by sex (female-male), region (rural-urban), and cancer site (rectal-colon) was ≥0.05, the disease pair was enriched in females, rural patients, and colon cancer patients, and the link was colored blue. On the other hand, if the difference was ≤−0.05, the disease pair was enriched in males, urban patients, and rectal cancer patients, and the link was colored red. The ICD-10 (International Classification of Diseases, 10th Revision) codes are clarified in Multimedia Appendix 1.



Discussion

Principal Findings

In this population-based cohort study, the entire spectrum of comorbidities in patients newly diagnosed with CRC was examined by scanning all hospitalizations during a 5-year look-back period before the first CRC diagnosis. Comorbidities were common in CRC patients at diagnosis, and prevalence and complex coexistence relationships varied by sex, age at diagnosis, region, and cancer site. The most prevalent comorbidities were hypertension, hyperplasia of the prostate, COPD, DM, and IHD. There were large disparities in disease co-occurrences in the male-female and urban-rural subgroups. Prevalence, diversity of diseases, and strength of disease coexistence were higher in urban patients, whereas severe diseases were more likely to co-occur in rural patients, such as the sequelae of cerebrovascular disease or COPD co-occurring with heart failure.

It was not surprising that comorbidities were common among newly diagnosed CRC patients, because both CRC and chronic diseases generally occur in older people [40-42]. CRC and comorbid conditions, such as DM and hypertension, may share common risk factors, such as smoking, poor diet, hyperlipidemia, lack of physical activity, and obesity [43-45]. In this population-based study, the most prevalent comorbidities were hypertension, hyperplasia of the prostate, COPD, DM, and IHD. Similar results were found in previous studies in New Zealand and the United States, although the prevalence of comorbidities differed [8,17,46]. Variations of prevalence might be related to differences in the study population (eg, some studies focused only on elderly patients among whom comorbidities are more common) and the definition used for chronic conditions (eg, a 1-year look-back period for hospitalization data resulted in a lower yield of comorbid conditions). Accumulating evidence suggests that sedentary lifestyle, metabolic syndrome (presence of ≥3 conditions among obesity, hypertension, hyperlipidemia, and DM), and antibiotic consumption may contribute to the development of early onset CRC [45,47-49]. In our study, we found 15 comorbidities with prevalences greater than 1% in early onset CRC patients, where hypertension, disorders of lipoprotein metabolism, hyperplasia of the prostate, and DM were the most prevalent comorbidities. Comorbidity is a strong and independent factor in the treatment and survival of CRC patients [5,50,51], for example, CRC patients comorbid with DM continued to receive chemotherapy less frequently than those without DM. Given the survival disadvantage in CRC patients with comorbidities [52], further evidence on the risks and benefits of therapy according to the types of co-occurring comorbidities and the severity of the diseases is needed to foster personalized medical care for those patients.

The complex coexistence relationships that occur among comorbidities are not due to simple coincidence, and therefore, to understand those relations, comorbidities must be studied collectively and not in isolation. Coexisting diseases within the same disease system were primarily within the circulatory system. The highest number of comorbidities was within the circulatory system, and those comorbidities included the highly prevalent diseases of hypertension, IHD, heart failure, and atherosclerosis. Coexisting diseases involved 2 disease systems, and they primarily included diseases of the circulatory system and those of the endocrine, nervous, respiratory, genitourinary, musculoskeletal, or connective tissue system. In a population-based cohort study in Spain, the most frequent comorbidity was DM, and the most frequent disease pair was CHF with DM, followed by CHF with COPD and peripheral vascular disease with DM [53]. In that study, the modified Charlson comorbidity score (only including 12 comorbidities) used to measure comorbidity might have underestimated the comorbidity burden and the complex coexistence relationships, because it did not include some important comorbidities, such as highly prevalent hypertension (I10). According to the network approach in this study, disorders of lipoprotein metabolism and hyperplasia of the prostate were part of some of the shortest paths between 2 comorbidities (mediated correlations between others), and heart failure was the end point of many comorbidities. Those comorbidities had a high likelihood to form bridges between other comorbidities, suggesting possibilities for the management of target comorbidities in CRC patients.

Many disease pairs were unique to or enriched in a sex group. In this regard, coexistence with COPD, cerebrovascular diseases, atherosclerosis, heart failure, or renal failure was predominant in males, whereas coexistence with osteoporosis or gonarthrosis was predominant in females. The prevalence of musculoskeletal disorders was higher in women than in men, resulting in an increased likelihood of co-existence with other highly prevalent chronic diseases (eg, hypertension, IHD, cerebrovascular diseases, or atherosclerosis) in female CRC patients. Notably, the prevalence of some severe diseases (eg, cerebral infarction and heart failure) was similar in males and females, but there were sex differences in coexistence with other chronic diseases (eg, a higher likelihood of cerebral infarction comorbid with the sequelae of cerebrovascular disease, COPD, or degenerative diseases of the nervous system, and heart failure comorbid with other pulmonary heart diseases or renal failure in males). Previous studies have found a prevalence difference in comorbidities by sex among CRC patients [54]. A population-based study using administrative data to describe comorbidities in CRC patients at diagnosis in Spain showed that the prevalence of COPD in males was 23.0%, which was 2.4 times higher than that in females (prevalence of 9.5%) [54]. In China, the prevalence of COPD in males was 19.0% (95% CI 16.9%-21.2%), which was approximately 2.3 times higher than that in females (prevalence of 8.1%, 95% CI 6.8%-9.3%), mainly because of a significant difference in smoking status between males and females (current smokers: 58.2% vs 4.0%) [55]. Such differences might indicate a disparity in lifestyle-related diseases, which may be due to physical, hormonal, or even genetic differences [56]. It is becoming increasingly standard practice to report sex-specific estimates because of the vital importance of gender consideration in precision medicine [56-58].

Notably, sex differences were observed in the prevalence of anxiety and the comorbid strength of anxiety with other diseases, for example, anxiety was enriched in females and had a higher likelihood of co-existence with cerebrovascular disease. Previous observational studies indicated that cardiocerebrovascular modulation and dynamic cerebral autoregulation are compromised in patients with anxiety [59-61]. In addition, evidence from a prospective cohort study indicated that greater anxiety or depression can not only impede adherence to healthy habits in CRC patients, but also be a marker for accelerated CRC progression [62]. Furthermore, in a previous population-based study in New Zealand, the prevalence of anxiety was 1.0% and 0.6% in colon cancer and rectal cancer patients, respectively [17], which were values similar to those in this study. The use of regional hospitalizations to identify the prevalence of clinically significant anxiety as well as other mental disorders might lead to underestimates, because estimated prevalence is generally higher when questionnaires are used [63,64]. Considering the well-known underestimated prevalence of mental disorders, especially in low- and middle-income countries [65], and their negative effects on quality of life, scanning for symptoms of anxiety and depression, implementation of dietary and physical activity interventions, and implementation of social support are of utmost importance among CRC patients at diagnosis and even years after treatment [66-68].

Rural-urban disparities were detected in the prevalence of comorbidities and the comorbid strength among comorbidities. Socioeconomic differences in the prevalence of comorbidities in CRC patients were also reported in England, where at least 11 of the 14 chronic conditions showed increasing prevalence along with the level of deprivation in CRC patients [16]. For example, the most deprived groups had approximately twice the odds of having COPD compared with the least deprived groups, which might be related to the higher prevalence of smoking in the more deprived population, but the association between smoking status and deprivation was not quantifiable as there was no information on smoking prevalence [16]. In our study, the prevalence of most comorbidities among newly diagnosed CRC patients was higher in urban patients than in rural patients, including some highly prevalent comorbidities, such as DM, disorders of lipoprotein metabolism, cerebrovascular diseases, atherosclerosis, and hyperplasia of the prostate, as well as some comorbidities with relatively low prevalence, such as nontoxic goiter, degenerative diseases of the nervous system, hypertensive heart disease, gout, and renal failure. Thus, comorbidities were more closely connected in urban patients than in rural patients, and some disease pairs were unique to or enriched in urban patients. Such differences might be associated with Western lifestyles in urban areas. Compared with the rural lifestyle, the urban lifestyle is more sedentary, with manual labor replaced by computer-based work and walking replaced by driving automobiles. Differences were also likely due to disparities in accessibility to diagnostic and treatment services between urban and rural residents [69,70]. Notably, the prevalences of hypertension, heart failure, and COPD were somewhat higher but without clinical significance in urban patients than in rural patients; whereas the sequelae of cerebrovascular disease comorbid with hypertensive heart disease or heart failure, heart failure comorbid with COPD or spondylosis, and COPD comorbid with pulmonary heart diseases were over-presented in rural patients. The co-occurrence of severe diseases in rural patients might be associated with relatively low adherence to colonoscopy screening among rural residents. Thus, CRC at diagnosis was generally more severe, with a significantly higher risk of diagnosis at a late stage [71]. Given the relatively low participation rate in colonoscopy screening in high-risk populations in China [72] and the coexistence of severe diseases at diagnosis, health promotion campaigns and adoption of noninvasive screening tests should be priorities to improve adherence and diagnosis in population-based programs.

We found that approximately 50% of comorbidities had significantly higher prevalence in colon cancer patients than in rectal cancer patients. A literature review indicated that diet, smoking, and physical activity might have different effects on colon cancer compared with the effects on rectal cancer, for example, physical activity decreased the risk of colon cancer but not of rectal cancer [73]. A population-based register of all primary cancers diagnosed in New Zealand was linked to routine hospital discharge data, and the comorbidity prevalence in colon and rectal cancer patients was identified, with a prevalence difference by cancer site similar to that in our study [17]. A previous study identified differences in patient characteristics and surgical outcomes between colon cancer and rectal cancer [74], and there was a significantly higher proportion of comorbidities in colon cancer patients than in rectal cancer patients (76.3% vs 68.8%; *P*=.02). Among CRC patients in the Central Region of Denmark in 2000-2011, 38% of colon cancer patients and 32% of rectal cancer patients had at least one Charlson comorbidity [75]. These previous studies reported the comorbidity difference by cancer site, but they did not provide a clear definition of comorbidity [74] or just focused on limited chronic diseases [75]. Further studies are needed to systematically explore the impacts of comorbidity patterns or comorbidity trajectories on the treatment choice and survival prediction in patients with colon and rectal cancers.

Limitations

This study had some limitations. First, the results should be interpreted in the context of an inpatient population in a developing country, since the HDRs did not include outpatient records and mental conditions might not be consistently diagnosed and recorded in HDRs [22]. Moreover, Berkson bias is unavoidable when only using an inpatient population to estimate morbidity [76]. Second, stage information is not yet included in HDRs because of the inherent limitation of the administrative database [31,34]. Further studies are needed to explore differences in comorbidity patterns according to TNM stages. Third, we systematically analyzed all chronic conditions and comorbidity patterns in newly diagnosed CRC patients in different subgroups, but did not explore CRC-specific comorbidities. As the emergence of comorbidities increases with age, further studies are needed to explore comorbidities that can be attributed to CRC rather than age, which may be helpful to provide etiological hypotheses or to construct surveillance programs [77]. Lastly, the common comorbidity profile identified in the study appears to be associated with treatment decision-making; however, limited by the available data, the relationships between comorbidities with different disease severities and different outcomes could not be identified in our study. The provincial data set was implemented in 2015 and did not contain proactive patient follow-up. Population-based cancer registries in China mainly report cancer incidence, mortality, expected survival, and trends in incidence and mortality, and comorbidity information is missing [78,79]. If routinely collected health data can be linked to the regional mortality database and even to local cancer registries, future studies may assess the relationships of comorbidities, comorbid disease pairs, and their severities with different outcomes (eg, short-term and long-term mortality, and cancer-specific mortality).

Conclusions

This study applied network analysis to routinely collected health data in order to examine the full spectrum of comorbidities and the complex coexisting relationships in newly diagnosed CRC patients. Some comorbidities, which were enriched in a given subgroup, largely increased the complexity of comorbid relationships, mediated correlations between other comorbidities, and might be the end points of many comorbidities. In addition, some co-occurring disease pairs predominated in a given subgroup. This study systematically analyzed the prevalence of individual diseases as well as the comorbid strength of coexisting diseases to provide a complete picture of comorbidities in CRC patients at diagnosis and also provide new insights into patterns of comorbidity. The data-driven discovery of comorbidity patterns may help improve the understanding of complex diseases, supplement traditional approaches in clinical studies, and help formulate appropriate preventive health measures to address high-risk comorbidities.

Acknowledgments

This study was supported by the Key Research and Development Program of Sichuan Province (2021YFS0112 and 2022YFS0163) and the Technological Innovation Research and Development Project of Chengdu (2021-YF05-01214-SN).

Data Availability

The data sets generated or analyzed during this study are restricted by the Ethics Committee of the Health Information Center of Sichuan Province and are not publicly available.

Authors' Contributions

HQ and XW designed the study, obtained funding, and directed the study. HQ, LW, and LZ contributed to data collection and statistical analysis. HQ and LW created the figures and drafted the manuscript. XW and HQ interpreted the results and revised the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The frequency and prevalence of comorbidities in new colorectal cancer cases.
[DOC File , 88 KB - publichealth_v9i1e41999_app1.doc]

Multimedia Appendix 2
Comorbidity prevalence using a 3-year look-back period and its difference compared with using a 5-year look-back period.
[DOC File , 74 KB - publichealth_v9i1e41999_app2.doc]

Multimedia Appendix 3
Scatter plot of comorbidity prevalence using 3-year and 5-year look-back periods.
[DOC File , 83 KB - publichealth_v9i1e41999_app3.doc]

Multimedia Appendix 4
Comorbidity prevalence difference by sex using a 5-year look-back period.
[DOC File , 95 KB - publichealth_v9i1e41999_app4.doc]

Multimedia Appendix 5
Comorbidity prevalence difference by region using a 5-year look-back period.
[DOC File , 81 KB - publichealth_v9i1e41999_app5.doc]

Multimedia Appendix 6
Comorbidity prevalence difference by cancer site using a 5-year look-back period.
[DOC File , 81 KB - publichealth_v9i1e41999_app6.doc]

Multimedia Appendix 7
Prevalence differences of enrichment comorbidities in subgroups after adjustment with age.
[DOC File , 54 KB - publichealth_v9i1e41999_app7.doc]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol 2021 Oct;14(10):101174 [FREE Full text] [doi: 10.1016/j.tranon.2021.101174] [Medline: 34243011]
3. Li N, Lu B, Luo C, Cai J, Lu M, Zhang Y, et al. Incidence, mortality, survival, risk factor and screening of colorectal cancer: A comparison among China, Europe, and northern America. Cancer Lett 2021 Dec 01;522:255-268. [doi: 10.1016/j.canlet.2021.09.034] [Medline: 34563640]
4. Yoon S, Kim E, Seo H, Oh I. The Association between Charlson Comorbidity Index and the Medical Care Cost of Cancer: A Retrospective Study. Biomed Res Int 2015;2015:259341 [FREE Full text] [doi: 10.1155/2015/259341] [Medline: 26347086]
5. Hahn EE, Gould MK, Munoz-Plaza CE, Lee JS, Parry C, Shen E. Understanding Comorbidity Profiles and Their Effect on Treatment and Survival in Patients With Colorectal Cancer. J Natl Compr Canc Netw 2018 Jan;16(1):23-34. [doi: 10.6004/jnccn.2017.7026] [Medline: 29295878]

6. Sarfati D, Hill S, Blakely T, Robson B, Purdie G, Dennett E, et al. The effect of comorbidity on the use of adjuvant chemotherapy and survival from colon cancer: a retrospective cohort study. BMC Cancer 2009 Apr 20;9(1):116 [FREE Full text] [doi: 10.1186/1471-2407-9-116] [Medline: 19379520]
7. Iversen L, Nørgaard M, Jacobsen J, Laurberg S, Sørensen H. The impact of comorbidity on survival of Danish colorectal cancer patients from 1995 to 2006--a population-based cohort study. Dis Colon Rectum 2009 Jan;52(1):71-78. [doi: 10.1007/DCR.0b013e3181974384] [Medline: 19273959]
8. Sarfati D, Tan L, Blakely T, Pearce N. Comorbidity among patients with colon cancer in New Zealand. N Z Med J 2011 Jul 08;124(1338):76-88. [Medline: 21946965]
9. Baretti M, Rimassa L, Personeni N, Giordano L, Tronconi MC, Pressiani T, et al. Effect of Comorbidities in Stage II/III Colorectal Cancer Patients Treated With Surgery and Neoadjuvant/Adjuvant Chemotherapy: A Single-Center, Observational Study. Clin Colorectal Cancer 2018 Sep;17(3):e489-e498. [doi: 10.1016/j.clcc.2018.03.010] [Medline: 29650416]
10. Gross C, Guo Z, McAvay G, Allore H, Young M, Tinetti M. Multimorbidity and survival in older persons with colorectal cancer. J Am Geriatr Soc 2006 Dec;54(12):1898-1904. [doi: 10.1111/j.1532-5415.2006.00973.x] [Medline: 17198496]
11. Jørgensen TL, Hallas J, Friis S, Herrstedt J. Comorbidity in elderly cancer patients in relation to overall and cancer-specific mortality. Br J Cancer 2012 Mar 27;106(7):1353-1360 [FREE Full text] [doi: 10.1038/bjc.2012.46] [Medline: 22353805]
12. Pule ML, Buckley E, Niyonsenga T, Roder D. The effects of comorbidity on colorectal cancer mortality in an Australian cancer population. Sci Rep 2019 Jun 12;9(1):8580 [FREE Full text] [doi: 10.1038/s41598-019-44969-8] [Medline: 31189947]
13. Mehta HB, Sura SD, Adhikari D, Andersen CR, Williams SB, Senagore AJ, et al. Adapting the Elixhauser comorbidity index for cancer patients. Cancer 2018 May 01;124(9):2018-2025 [FREE Full text] [doi: 10.1002/cncr.31269] [Medline: 29390174]
14. Strombom P, Widmar M, Keskin M, Gennarelli RL, Lynn P, Smith JJ, et al. Assessment of the Value of Comorbidity Indices for Risk Adjustment in Colorectal Surgery Patients. Ann Surg Oncol 2019 Sep 17;26(9):2797-2804 [FREE Full text] [doi: 10.1245/s10434-019-07502-9] [Medline: 31209671]
15. Pule L, Buckley E, Niyonsenga T, Roder D. Optimizing the measurement of comorbidity for a South Australian colorectal cancer population using administrative data. J Eval Clin Pract 2020 Aug 13;26(4):1250-1258. [doi: 10.1111/jep.13305] [Medline: 31721394]
16. Fowler H, Belot A, Ellis L, Maringe C, Luque-Fernandez MA, Njagi EN, et al. Comorbidity prevalence among cancer patients: a population-based cohort study of four cancers. BMC Cancer 2020 Jan 28;20(1):2 [FREE Full text] [doi: 10.1186/s12885-019-6472-9] [Medline: 31987032]
17. Sarfati D, Gurney J, Lim BT, Bagheri N, Simpson A, Koea J, et al. Identifying important comorbidity among cancer populations using administrative data: Prevalence and impact on survival. Asia Pac J Clin Oncol 2016 Mar 19;12(1):e47-e56. [doi: 10.1111/ajco.12130] [Medline: 24354451]
18. Chen Y, Li L, Xu R. Disease Comorbidity Network Guides the Detection of Molecular Evidence for the Link Between Colorectal Cancer and Obesity. AMIA Jt Summits Transl Sci Proc 2015;2015:201-206 [FREE Full text] [Medline: 26306270]
19. Hidalgo CA, Blumm N, Barabási A, Christakis NA. A dynamic network approach for the study of human phenotypes. PLoS Comput Biol 2009 Apr 10;5(4):e1000353 [FREE Full text] [doi: 10.1371/journal.pcbi.1000353] [Medline: 19360091]
20. Westergaard D, Moseley P, Sørup F, Baldi P, Brunak S. Population-wide analysis of differences in disease progression patterns in men and women. Nat Commun 2019 Feb 08;10(1):666 [FREE Full text] [doi: 10.1038/s41467-019-08475-9] [Medline: 30737381]
21. Wang L, Qiu H, Luo L, Zhou L. Age- and Sex-Specific Differences in Multimorbidity Patterns and Temporal Trends on Assessing Hospital Discharge Records in Southwest China: Network-Based Study. J Med Internet Res 2022 Feb 25;24(2):e27146 [FREE Full text] [doi: 10.2196/27146] [Medline: 35212632]
22. Qiu H, Wang L, Zeng X, Pan J. Comorbidity patterns in depression: A disease network analysis using regional hospital discharge records. J Affect Disord 2022 Jan 01;296:418-427. [doi: 10.1016/j.jad.2021.09.100] [Medline: 34606805]
23. Zhou D, Wang L, Ding S, Shen M, Qiu H. Phenotypic Disease Network Analysis to Identify Comorbidity Patterns in Hospitalized Patients with Ischemic Heart Disease Using Large-Scale Administrative Data. Healthcare (Basel) 2022 Jan 01;10(1):80 [FREE Full text] [doi: 10.3390/healthcare10010080] [Medline: 35052244]
24. Ong M, Mullen MP, Austin ED, Szolovits P, Natter MD, Geva A, et al. Learning a Comorbidity-Driven Taxonomy of Pediatric Pulmonary Hypertension. Circ Res 2017 Aug 04;121(4):341-353. [doi: 10.1161/circresaha.117.310804]
25. Korolainen MA, Tuominen S, Kurki S, Lassenius MI, Toppila I, Purmonen T, et al. Burden of migraine in Finland: multimorbidity and phenotypic disease networks in occupational healthcare. J Headache Pain 2020 Jan 31;21(1):8 [FREE Full text] [doi: 10.1186/s10194-020-1077-x] [Medline: 32005102]
26. Divo MJ, Casanova C, Marin JM, Pinto-Plata VM, de-Torres JP, Zulueta JJ, BODE Collaborative Group. COPD comorbidities network. Eur Respir J 2015 Sep 09;46(3):640-650 [FREE Full text] [doi: 10.1183/09031936.00171614] [Medline: 26160874]
27. Bhavnani SK, Dang B, Penton R, Visweswaran S, Bassler KE, Chen T, et al. How High-Risk Comorbidities Co-Occur in Readmitted Patients With Hip Fracture: Big Data Visual Analytical Approach. JMIR Med Inform 2020 Oct 26;8(10):e13567 [FREE Full text] [doi: 10.2196/13567] [Medline: 33103657]
28. Mu X, Wang W, Jiang Y, Feng J. Patterns of Comorbidity in Hepatocellular Carcinoma: A Network Perspective. Int J Environ Res Public Health 2020 Apr 29;17(9):3108 [FREE Full text] [doi: 10.3390/ijerph17093108] [Medline: 32365650]

29. Kuan V, Denaxas S, Gonzalez-Izquierdo A, Direk K, Bhatti O, Husain S, et al. A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. The Lancet Digital Health 2019 Jun;1(2):e63-e77. [doi: 10.1016/s2589-7500(19)30012-3]
30. Guo M, Yu Y, Wen T, Zhang X, Liu B, Zhang J, et al. Analysis of disease comorbidity patterns in a large-scale China population. BMC Med Genomics 2019 Dec 12;12(Suppl 12):177 [FREE Full text] [doi: 10.1186/s12920-019-0629-x] [Medline: 31829182]
31. Duan Z, Gao L, Xun Y, Zhang Z, Hang X, Guo X, et al. Design and application of the quality evaluation platform for home page of inpatient medical records. Chinese Health Quality Management 2020;27(3):40-42.
32. Wang X, Yang H, Duan Z, Pan J. Spatial accessibility of primary health care in China: A case study in Sichuan Province. Soc Sci Med 2018 Jul;209:14-24. [doi: 10.1016/j.socscimed.2018.05.023] [Medline: 29778934]
33. Pan J, Zhao H, Wang X, Shi X. Assessing spatial access to public and private hospitals in Sichuan, China: The influence of the private sector on the healthcare geography in China. Soc Sci Med 2016 Dec;170:35-45. [doi: 10.1016/j.socscimed.2016.09.042] [Medline: 27750066]
34. Liu X, Qiao L, Luo Y, Ma J, Li L, Jiang K, et al. Analysis of Hospital-based Cancer Incidence in Sichuan Province, 2018. China Cancer 2021;30(6):429-436 [FREE Full text]
35. Chronic Condition Indicator (CCI) for ICD-10-CM (beta version V2019.1) [updated January 07, 2020]. HCUP. 2019. URL: https://hcup-us.ahrq.gov/toolssoftware/chronic_icd10/chronic_icd10.jsp [accessed 2023-08-18]
36. Amell A, Roso-Llorach A, Palomero L, Cuadras D, Galván-Femenía I, Serra-Musach J, et al. Disease networks identify specific conditions and pleiotropy influencing multimorbidity in the general population. Sci Rep 2018 Oct 29;8(1):15970 [FREE Full text] [doi: 10.1038/s41598-018-34361-3] [Medline: 30374096]
37. Ahlgren P, Jarneving B, Rousseau R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. J. Am. Soc. Inf. Sci 2003 Apr;54(6):550-560. [doi: 10.1002/asi.10242]
38. Egghe L, Leydesdorff L. The relation between Pearson's correlation coefficient and Salton's cosine measure. J. Am. Soc. Inf. Sci 2009 May;60(5):1027-1036. [doi: 10.1002/asi.21009]
39. Freeman LC. Centrality in social networks conceptual clarification. Social Networks 1978 Jan;1(3):215-239. [doi: 10.1016/0378-8733(78)90021-7]
40. Movsas B, Extermann M. Introduction: Cancer, aging, and comorbidities. Semin Radiat Oncol 2012 Oct;22(4):263-264. [doi: 10.1016/j.semradonc.2012.05.001] [Medline: 22985807]
41. Sanoff HK, Bleiberg H, Goldberg RM. Managing Older Patients With Colorectal Cancer. JCO 2007 May 10;25(14):1891-1897. [doi: 10.1200/jco.2006.10.1220]
42. The Lancet Oncology. Not old, just older: considering age in cancer care. The Lancet Oncology 2019 Jul;20(7):887. [doi: 10.1016/s1470-2045(19)30427-9]
43. O'Sullivan D, Sutherland RL, Town S, Chow K, Fan J, Forbes N, et al. Risk Factors for Early-Onset Colorectal Cancer: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2022 Jun;20(6):1229-1240.e5. [doi: 10.1016/j.cgh.2021.01.037] [Medline: 33524598]
44. Thanikachalam K, Khan G. Colorectal Cancer and Nutrition. Nutrients 2019 Jan 14;11(1):164 [FREE Full text] [doi: 10.3390/nu11010164] [Medline: 30646512]
45. Chen H, Zheng X, Zong X, Li Z, Li N, Hur J, et al. Metabolic syndrome, metabolic comorbid conditions and risk of early-onset colorectal cancer. Gut 2021 Jun 09;70(6):1147-1154 [FREE Full text] [doi: 10.1136/gutjnl-2020-321661] [Medline: 33037055]
46. Ljubic B, Pavlovski M, Alshehri J, Roychoudhury S, Bajic V, Van Neste C, et al. Comorbidity network analysis and genetics of colorectal cancer. Informatics in Medicine Unlocked 2020;21:100492. [doi: 10.1016/j.imu.2020.100492]
47. Burnett-Hartman AN, Lee JK, Demb J, Gupta S. An Update on the Epidemiology, Molecular Characterization, Diagnosis, and Screening Strategies for Early-Onset Colorectal Cancer. Gastroenterology 2021 Mar;160(4):1041-1049 [FREE Full text] [doi: 10.1053/j.gastro.2020.12.068] [Medline: 33417940]
48. Patel SG, Karlitz JJ, Yen T, Lieu CH, Boland CR. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. The Lancet Gastroenterology & Hepatology 2022 Mar;7(3):262-274. [doi: 10.1016/s2468-1253(21)00426-x]
49. McDowell R, Perrott S, Murchie P, Cardwell C, Hughes C, Samuel L. Oral antibiotic use and early-onset colorectal cancer: findings from a case-control study using a national clinical database. Br J Cancer 2022 Apr 17;126(6):957-967 [FREE Full text] [doi: 10.1038/s41416-021-01665-7] [Medline: 34921228]
50. Boakye D, Rillmann B, Walter V, Jansen L, Hoffmeister M, Brenner H. Impact of comorbidity and frailty on prognosis in colorectal cancer patients: A systematic review and meta-analysis. Cancer Treat Rev 2018 Mar;64:30-39. [doi: 10.1016/j.ctrv.2018.02.003] [Medline: 29459248]
51. Zanders MMJ, van Steenbergen LN, Haak HR, Rutten HJT, Pruijt JFM, Poortmans PMP, et al. Diminishing differences in treatment between patients with colorectal cancer with and without diabetes: a population-based study. Diabet Med 2013 Oct 28;30(10):1181-1188. [doi: 10.1111/dme.12253] [Medline: 23758334]

52. Boakye D, Nagrini R, Ahrens W, Haug U, Günther K. The association of comorbidities with administration of adjuvant chemotherapy in stage III colon cancer patients: a systematic review and meta-analysis. Ther Adv Med Oncol 2021 Jan 21;13:1758835920986520 [FREE Full text] [doi: 10.1177/1758835920986520] [Medline: 33613694]
53. Luque-Fernandez MA, Gonçalves K, Salamanca-Fernández E, Redondo-Sanchez D, Lee SF, Rodríguez-Barranco M, et al. Multimorbidity and short-term overall mortality among colorectal cancer patients in Spain: A population-based cohort study. Eur J Cancer 2020 Apr;129:4-14 [FREE Full text] [doi: 10.1016/j.ejca.2020.01.021] [Medline: 32114366]
54. Luque-Fernandez MA, Redondo-Sanchez D, Lee SF, Rodríguez-Barranco M, Carmona-García M, Marcos-Gragera R, et al. Multimorbidity by Patient and Tumor Factors and Time-to-Surgery Among Colorectal Cancer Patients in Spain: A Population-Based Study. Clin Epidemiol 2020;12:31-40. [doi: 10.2147/CLEP.S229935] [Medline: 32021469]
55. Fang L, Gao P, Bao H, Tang X, Wang B, Feng Y, et al. Chronic obstructive pulmonary disease in China: a nationwide prevalence study. Lancet Respir Med 2018 Jun;6(6):421-430 [FREE Full text] [doi: 10.1016/S2213-2600(18)30103-6] [Medline: 29650407]
56. Lee JM, Lee K, Kim DB, Ko S, Park YG. Colorectal Cancer Risks According to Sex Differences in Patients With Type II Diabetes Mellitus. Clinical and Translational Gastroenterology 2019;10(10):e00090. [doi: 10.14309/ctg.0000000000000090]
57. Hases L, Ibrahim A, Chen X, Liu Y, Hartman J, Williams C. The Importance of Sex in the Discovery of Colorectal Cancer Prognostic Biomarkers. IJMS 2021 Jan 29;22(3):1354. [doi: 10.3390/ijms22031354]
58. Legato MJ, Johnson PA, Manson JE. Consideration of Sex Differences in Medicine to Improve Health Care and Patient Outcomes. JAMA 2016 Nov 08;316(18):1865-1866. [doi: 10.1001/jama.2016.13995] [Medline: 27802499]
59. Zhang H, Guo Z, Yang G, Yang L, Han K, Wu J, et al. Compromised cerebrovascular modulation in chronic anxiety: evidence from cerebral blood flow velocity measured by transcranial Doppler sonography. Neurosci Bull 2012 Dec 14;28(6):723-728 [FREE Full text] [doi: 10.1007/s12264-012-1282-y] [Medline: 23152329]
60. Guo Z, Feng L, Yan X, Yang L, Huang S, Xing Y, et al. Characteristics of cardio-Cerebrovascular modulation in patients with generalized anxiety disorder: an observational study. BMC Psychiatry 2017 Jul 18;17(1):259 [FREE Full text] [doi: 10.1186/s12888-017-1428-6] [Medline: 28720078]
61. Guo Z, Lv S, Liu J, Wang Z, Jin H, Qiu Q, et al. Compromised dynamic cerebral autoregulation in patients with generalized anxiety disorder: a study using transfer function analysis. BMC Psychiatry 2018 Jun 01;18(1):164 [FREE Full text] [doi: 10.1186/s12888-018-1713-z] [Medline: 29859053]
62. Trudel-Fitzgerald C, Tworoger SS, Zhang X, Giovannucci EL, Meyerhardt JA, Kubzansky LD. Anxiety, Depression, and Colorectal Cancer Survival: Results from Two Prospective Cohorts. J Clin Med 2020 Sep 30;9(10):3174 [FREE Full text] [doi: 10.3390/jcm9103174] [Medline: 33007946]
63. Mols F, Schoormans D, de Hingh I, Oerlemans S, Husson O. Symptoms of anxiety and depression among colorectal cancer survivors from the population-based, longitudinal PROFILES Registry: Prevalence, predictors, and impact on quality of life. Cancer 2018 Jun 15;124(12):2621-2628 [FREE Full text] [doi: 10.1002/cncr.31369] [Medline: 29624635]
64. Peng Y, Huang M, Kao C. Prevalence of Depression and Anxiety in Colorectal Cancer Patients: A Literature Review. Int J Environ Res Public Health 2019 Jan 31;16(3):411 [FREE Full text] [doi: 10.3390/ijerph16030411] [Medline: 30709020]
65. Arokiasamy P, Uttamacharya, Kowal P, Capistrant BD, Gildner TE, Thiele E, et al. Chronic Noncommunicable Diseases in 6 Low- and Middle-Income Countries: Findings From Wave 1 of the World Health Organization's Study on Global Ageing and Adult Health (SAGE). Am J Epidemiol 2017 Mar 15;185(6):414-428 [FREE Full text] [doi: 10.1093/aje/kww125] [Medline: 28399566]
66. Ho M, Ho JWC, Fong DYT, Lee CF, Macfarlane DJ, Cerin E, et al. Effects of dietary and physical activity interventions on generic and cancer-specific health-related quality of life, anxiety, and depression in colorectal cancer survivors: a randomized controlled trial. J Cancer Surviv 2020 Aug 18;14(4):424-433 [FREE Full text] [doi: 10.1007/s11764-020-00864-0] [Medline: 32072434]
67. Gonzalez-Saenz de Tejada M, Bilbao A, Baré M, Briones E, Sarasqueta C, Quintana J, CARESS-CCR Group. Association between social support, functional status, and change in health-related quality of life and changes in anxiety and depression in colorectal cancer patients. Psychooncology 2017 Sep 19;26(9):1263-1269. [doi: 10.1002/pon.4303] [Medline: 28872742]
68. Chen L, Cheng C, Huang I, Chang J, Chang N, Hsiao F. The multiple mediating effects of cancer threat appraisal and quality of life on the association between mindfulness and depression for colorectal cancer survivors. Psychooncology 2021 Jun 20;30(6):853-862. [doi: 10.1002/pon.5644] [Medline: 33523559]
69. Yuan S, Xie S. Urban-rural disparity in cancer incidence in China, 2008-2012: a cross-sectional analysis of data from 36 cancer registers. BMJ Open 2021 Apr 30;11(4):e042762 [FREE Full text] [doi: 10.1136/bmjopen-2020-042762] [Medline: 33931407]
70. Wen D, Zou W, Wen X, Yang Y, Chen Y, He Y, et al. Urban-rural disparity in colorectal cancer incidence and increasing trend in relation to socioeconomic development and urbanization in China. J Int Med Res 2018 Oct 20;46(10):4181-4196 [FREE Full text] [doi: 10.1177/0300060518791090] [Medline: 30124349]
71. Zeng H, Ran X, An L, Zheng R, Zhang S, Ji JS, HBCR Working Group. Disparities in stage at diagnosis for five common cancers in China: a multicentre, hospital-based, observational study. Lancet Public Health 2021 Dec;6(12):e877-e887 [FREE Full text] [doi: 10.1016/S2468-2667(21)00157-2] [Medline: 34838194]

72. Chen H, Li N, Ren J, Feng X, Lyu Z, Wei L, group of Cancer Screening Program in Urban China (CanSPUC). Participation and yield of a population-based colorectal cancer screening programme in China. Gut 2019 Aug 30;68(8):1450-1457. [doi: 10.1136/gutjnl-2018-317124] [Medline: 30377193]
73. Tamas K, Walenkamp A, de Vries E, van Vugt M, Beets-Tan R, van Etten B, et al. Rectal and colon cancer: Not just a different anatomic site. Cancer Treat Rev 2015 Sep;41(8):671-679. [doi: 10.1016/j.ctrv.2015.06.007] [Medline: 26145760]
74. van der Sijp MP, Bastiaannet E, Mesker WE, van der Geest LGM, Breugom AJ, Steup WH, et al. Differences between colon and rectal cancer in complications, short-term survival and recurrences. Int J Colorectal Dis 2016 Oct 6;31(10):1683-1691 [FREE Full text] [doi: 10.1007/s00384-016-2633-3] [Medline: 27497831]
75. Ostenfeld E, Nørgaard M, Thomsen R, Iversen L, Jacobsen J, Søgaard M. Comorbidity and survival of Danish patients with colon and rectal cancer from 2000-2011: a population-based cohort study. Clin Epidemiol 2013 Nov;5(Suppl 1):65-74 [FREE Full text] [doi: 10.2147/CLEP.S47154] [Medline: 24227924]
76. Pearce N, Richiardi L. Commentary: Three worlds collide: Berkson's bias, selection bias and collider bias. Int J Epidemiol 2014 Apr 28;43(2):521-524 [FREE Full text] [doi: 10.1093/ije/dyu025] [Medline: 24585732]
77. Thong M, Boakye D, Jansen L, Martens U, Chang-Claude J, Hoffmeister M, et al. Comorbidities, Rather Than Older Age, Are Strongly Associated With Higher Utilization of Healthcare in Colorectal Cancer Survivors. J Natl Compr Canc Netw 2021 Dec 07;20(5):468-478.e7. [doi: 10.6004/jnccn.2021.7030] [Medline: 34875626]
78. Wei W, Zeng H, Zheng R, Zhang S, An L, Chen R, et al. Cancer registration in China and its role in cancer prevention and control. Lancet Oncol 2020 Jul;21(7):e342-e349. [doi: 10.1016/S1470-2045(20)30073-5] [Medline: 32615118]
79. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016;66(2):115-132 [FREE Full text] [doi: 10.3322/caac.21338] [Medline: 26808342]

Abbreviations

CHF: congestive heart failure
COPD: chronic obstructive pulmonary disease
CRC: colorectal cancer
DM: diabetes mellitus
HDR: hospital discharge record
ICD-10: International Classification of Diseases, 10th Revision
IHD: ischemic heart disease
PCN: phenotypic comorbidity network

Edited by A Mavragani, T Sanchez; submitted 18.08.22; peer-reviewed by T Wu, F Chen; comments to author 18.04.23; revised version received 18.05.23; accepted 25.07.23; published 05.09.23.

Please cite as:
Qiu H, Wang L, Zhou L, Wang X
Comorbidity Patterns in Patients Newly Diagnosed With Colorectal Cancer: Network-Based Study
JMIR Public Health Surveill 2023;9:e41999
URL: https://publichealth.jmir.org/2023/1/e41999
doi:10.2196/41999
PMID:37669093

Original Paper

Trial of the Pluslife SARS-CoV-2 Nucleic Acid Rapid Test Kit: Prospective Cohort Study

Dandan Zhu[1*], MM; Jing Huang[2*], MM; Bei Hu[3*], MD; Donglin Cao[4*], MD; Dingqiang Chen[5*], MD; Xinqiang Song[6*], DBA; Jialing Chen[1], MM; Hao Zhou[3], BMed; Aiqun Cen[2], BMed; Tieying Hou[7], MD

[1]Guangdong Center for Clinical Laboratory, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China

[2]Fever Clinic, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China

[3]Department of Emergency Medicine, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China

[4]Department of Laboratory Medicine, Guangdong Second Provincial General Hospital, Guangzhou, China

[5]Microbiome Medicine Center, Department of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, China

[6]Party Committee Office and President Office, Guangdong University of Finance and Economics, Guangzhou, China

[7]Hospital Office, Huazhong University of Science and Technology Union Shenzhen Hospital/Shenzhen Nanshan People's Hospital, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Tieying Hou, MD
Hospital Office
Huazhong University of Science and Technology Union Shenzhen Hospital/Shenzhen Nanshan People's Hospital
89 Taoyuan Road, Nanshan District
Shenzhen, 518052
China
Phone: 86 075526553111
Email: sz_houtieying@yeah.net

Abstract

Background: In response to the SARS-CoV-2 epidemic, a convenient, rapid, and sensitive diagnostic method for detecting COVID-19 is crucial for patient control and timely treatment.

Objective: This study aimed to validate the detection of SARS-CoV-2 with the Pluslife SARS-CoV-2 rapid test kit developed based on a novel thermostatic amplification technique called RNase hybridization-assisted amplification.

Methods: From November 25 to December 8, 2022, patients with suspected or confirmed COVID-19, close contacts, and health care workers at high risk of exposure were recruited from 3 hospitals and 1 university. Respiratory specimens were collected for testing with the Pluslife SARS-CoV-2 rapid test kit and compared with reverse transcription-quantitative polymerase chain reaction (RT-qPCR) and a commercial antigen assay kit. Samples from 1447 cases were obtained from 3 "ready-to-test" scenarios in which samples were collected on site and tested immediately, and samples from 503 cases were obtained from a "freeze-thaw test" scenario in which samples were collected, frozen, and thawed for testing.

Results: Pluslife SARS-CoV-2 rapid testing of samples from the "ready-to-test" scenario was found to be accurate (overall sensitivity and specificity of 98.3% and 99.3%, respectively) and diagnostically useful (positive and negative likelihood ratios of 145.45 and 0.02, respectively). Pluslife SARS-CoV-2 rapid testing of samples from the "freeze-thaw test" scenario was also found to be accurate (overall sensitivity and specificity of 71.2% and 98.6%, respectively) and diagnostically useful (positive and negative likelihood ratios of 51.01 and 0.67, respectively). Our findings demonstrated that the time efficiency and accuracy of the results in a "ready-to-test" scenario were better. The time required from sample preparation to the seeing the result of the Pluslife SARS-CoV-2 rapid test was 10 to 38 minutes, which was substantially shorter than that of RT-qPCR (at least 90 minutes). In addition, the diagnostic efficacy of the Pluslife SARS-CoV-2 rapid test was better than that of a commercial antigen assay kit.

Conclusions: The developed RNase hybridization-assisted amplification assay provided rapid, sensitive, and convenient detection of SARS-CoV-2 infection and may be useful for enhanced detection of COVID-19 in homes, high-risk industries, and hospitals.

(JMIR Public Health Surveill 2023;9:e48107) doi:10.2196/48107

KEYWORDS

SARS-CoV-2; COVID-19; RHAM; RNase hybridization-assisted amplification; field trial; diagnosis; screening; rapid test; cohort study; detection; recruitment; infection; Pluslife

Introduction

According to the World Health Organization, cases of COVID-19 caused by SARS-CoV-2 were reported in several regions in late 2019 and early 2020 [1]. The rapid spread of SARS-CoV-2 posed a major public health and economic hazard. Currently, the primary test used by most laboratories to diagnose COVID-19 is a nucleic acid amplification test using reverse transcription quantitative polymerase chain reaction (RT-qPCR) [2,3]. Although the World Health Organization declared on May 5, 2023, that the COVID-19 outbreak no longer constituted a "public health emergency of international concern", this merely demonstrates that the hazard can be effectively controlled with current human capacity—it does not mean that the pandemic is over. The COVID-19 epidemic remains a global health threat, and continued efforts to prevent and control the COVID-19 are still necessary.

For some low-resource primary health care conditions, the application scenarios of traditional nucleic acid testing are limited, and an assay with detection performance comparable to RT-qPCR and less demanding on the testing environment is needed. For high-traffic areas, such as customs and airports, the time required to obtain RT-qPCR results is too long, so a faster and more accurate test is needed to increase the speed of population screening. Therefore, it is essential to develop a rapid, convenient, and accurate method for detecting SARS-CoV-2 with detection performance comparable to that of RT-qPCR for areas where medical resources are scarce and for scenarios with high foot traffic.

Compared with RT-qPCR, thermostatic amplification technology does not require a temperature rise or fall to achieve nucleic acid amplification, which helps to substantially mitigate the cost of the instrument and speed up the reaction process [4-7]. However, while most of the existing traditional thermostatic amplification techniques can solve the problem of instrument cost and fast amplification speed, they suffer from inadequate stability and sensitivity when directly benchmarked against RT-qPCR. Thermostatic amplification technology has been developing for decades, during which a variety of nucleic acid detection methods have emerged [8]. RNase hybridization-assisted amplification (RHAM) is a thermostatic amplification technology developed independently by Guangzhou Pluslife Technology with underlying intellectual property rights. The actual performance of this technology is comparable to that of RT-qPCR, and it is more tolerant and compatible. RHAM can extract, amplify, and detect in 1 step without the opening of the cap after amplification, and it requires a simple professional environment and hardware support. The Pluslife SARS-CoV-2 rapid test kit based on this technology is in lyophilized form, which is easy to transport at room temperature and can achieve extraction-free molecular detection.

In this study, a field trial of the Pluslife SARS-CoV-2 rapid test kit was conducted at 4 centers, including Guangdong Provincial People's Hospital, Zhujiang Hospital at Southern Medical University, Guangdong Second Provincial General Hospital, and Guangdong University of Finance and Economics. Based on the gold standard RT-qPCR assay, we evaluated the diagnostic performance of the Pluslife SARS-CoV-2 rapid test kit and compared it with a commercial antigen assay kit by Wondfo Biotech for the diagnosis of SARS-CoV-2.

Methods

Ethical Considerations

The study protocol was approved by the Ethics Review Committee of Guangdong Provincial People's Hospital (KY2023-067-02).

Candidates

The study population comprised patients with suspected or confirmed COVID-19, close contacts, and high-risk exposed health care workers from 4 centers, including Guangdong Provincial People's Hospital, Zhujiang Hospital at Southern Medical University, Guangdong Second Provincial General Hospital, and Guangdong University of Finance and Economics, from November 25 to December 8, 2022. Nasal swabs or oropharyngeal swabs were collected for SARS-CoV-2 testing. The flow chart of the study protocol is shown in Figure 1. The characteristics of the 2028 candidates who received the Pluslife SARS-CoV-2 rapid test are shown in Table 1.

Figure 1. Study flow chart. We recruited candidates with suspected or confirmed COVID-19, close contacts, and health care workers at high risk of exposure at 4 centers. All participants provided written informed consent. Parallel nasal and oropharyngeal swabs were collected and tested with the Pluslife rapid test, antigen test, and RT-qPCR from November 25 to December 8, 2022. RT-qPCR: real-time reverse transcription quantitative polymerase chain reaction.

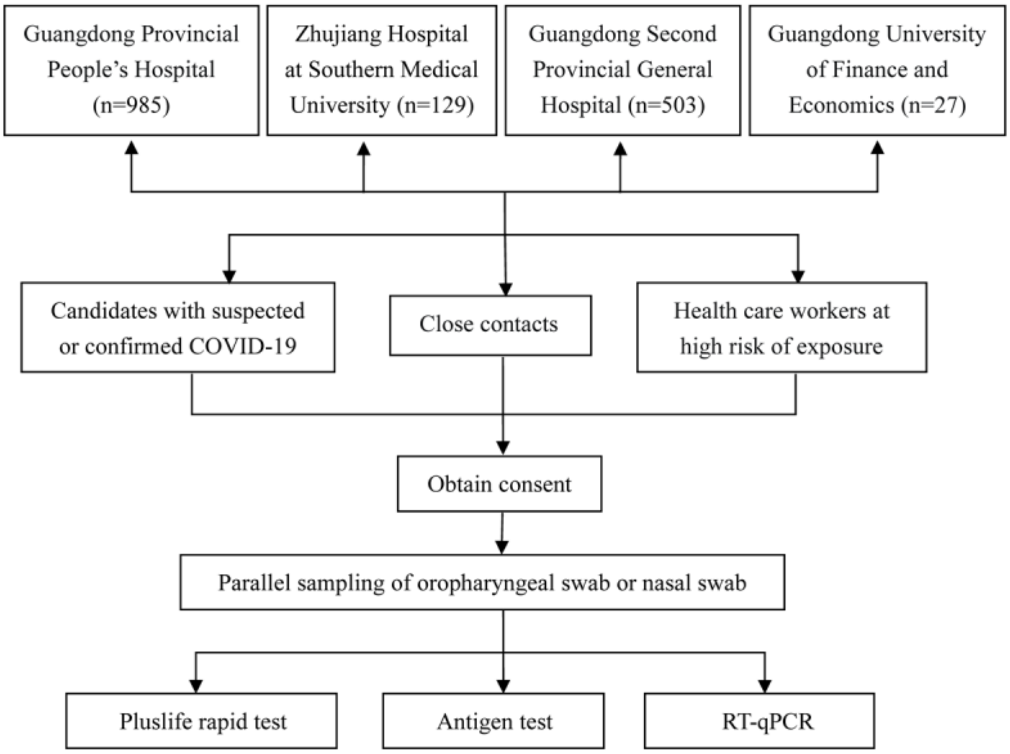


Table 1. Characteristics of the 2028 candidates whose samples were tested with the Pluslife SARS-CoV-2 rapid test from November 25 to December 8, 2022.

Characteristic, n	Guangdong Provincial People's Hospital (n=1333)	Zhujiang Hospital at Southern Medical University (n=147)	Guangdong Second Provincial General Hospital (n=518)	Guangdong University of Finance and Economics (n=30)
Population				
Medical staff (n=1206)	985	121	100	0
Teachers (n=11)	0	0	0	11
Students (n=8)	0	0	0	8
Logistics staff (n=8)	0	0	0	8
Fever outpatients (n=213)	205	8	0	0
Emergency patients (n=101)	101	0	0	0
Confirmed patients with COVID-19 (n=403)	0	0	403	0
Collection workers				
Grassroots workers (n=1)	0	0	0	1
Nurses (n=34)	21	3	10	0
Medical self-collection (n=131)	10	121	0	0
Testing methods				
RT-qPCR[a] (n=1950)	1291	129	503	27
Pluslife rapid test (n=2028)	1333	147	518	30
Antigen test (n=283)	283	0	0	0
Presence of symptoms				
Yes (n=213)	205	8	0	0
No (n=1737)	1086	121	503	27

[a] RT-qPCR: real-time reverse transcription quantitative polymerase chain reaction.

The RT-qPCR Assay

Viral RNA was extracted from oropharyngeal swabs using the liquid proteinase K magnetic bead method. The BioGerm 2019-nCoV assay (BioGerm) was used to detect SARS-CoV-2 RNA with SARS-CoV-2 RNA-dependent RNA polymerase and nucleocapsid genes according to the product instructions. Briefly, 12 μL of the RT-qPCR mixture, 4 μL of the RT-qPCR enzyme mixture, and 4 μL of the ORFlab/N reaction mixture were combined into a 20 μL master mix for each sample well. The master mix was pipetted into each well of a 96-well optical reaction plate. Then, 5 μL each of specimen nucleic acid, positive control, and negative control were added to the master mix. The tubes were closed securely and centrifuged briefly. Amplification was performed with a fluorescent qPCR instrument for 1 cycle at 50 °C for 10 minutes and 95 °C for 5 minutes, followed by 45 cycles at 95 °C for 10 seconds and 55 °C for 40 seconds. The results of the assay were evaluated with a cycling threshold value, where a value <40 for both target genes was defined as a positive result.

Pluslife SARS-CoV-2 Card Test Procedure

First, the comprehensive nucleic acid test was preheated. The entire absorbent tip of the disposable sampling swab was gently inserted into one nostril approximately 1.5 to 2 cm deep and rotated 3 times on the inner wall or rubbed 3 times on the posterior pharyngeal wall and both sides of the pharyngeal-palatal arch. For the "ready-to-test" method, the tip of the disposable sampling swab was twisted against the bottom and side of the nucleic acid releaser tube 10 times while pinching the tip of the disposable sampling swab. For the "freeze-thaw test" method, the collected pharyngeal swabs were placed in a 500 μL sample tube of Youkang virus preservation solution, shaken and mixed, and frozen in a –20°C refrigerator for 24 to 240 hours. The sample tubes were thawed and 150 μL of Youkang virus preservation solution was added to 2 mL of nucleic acid release agent, the caps were screwed on, and the tubes were vortexed and mixed for 5 to 10 seconds or mixed by inverting 10 to 15 times. After that, the procedure was the same for both methods. The nucleic acid–releasing agent solution was poured into the SARS-CoV-2 reaction card sample tube between the 2 injection lines. The protruding curved air pocket on the SARS-CoV-2 reaction card sample tube was pressed down, and the card was held and shaken up and down 10 times. The SARS-CoV-2 response card was inserted into the device, and the "Start Test" button was pressed. The red indicator light indicated a positive SARS-CoV-2 nucleic acid result, and the blue indicator light indicated a negative

SARS-CoV-2 nucleic acid result. The specific operation process is illustrated in Multimedia Appendix 1.

Antigen Test Procedure

An antigen detection kit (Wondfo Biotech) was used to detect the SARS-CoV-2 nucleocapsid antigen in respiratory specimens. This assay device contained 2 precoated antibody lines: a “C” (control) line and a “T” (test) line. The control area was precoated with chicken IgY antibody, and the test area was precoated with SARS-CoV-2 antibody. Color particle–coupled anti–SARS-CoV-2 antibody was used to detect SARS-CoV-2 antigen. In the assay, SARS-CoV-2 antigen in the respiratory specimen interacted with the color particle–coupled anti–SARS-CoV-2 monoclonal antibody to form a colored antigen-antibody complex. The complex migrated across the membrane by capillary action up to the test line, where it was captured by the pre-encapsulated anti–SARS-CoV-2 antibody. If SARS-CoV-2 antigen was present in the respiratory specimen, a colored test line was visible in the result window, and the intensity of the stained test line varied with the amount of SARS-CoV-2 antigen in the specimen. Colored particles conjugated to chicken IgY were used as the control line.

Statistical Analysis

We estimated the minimum sample size using the single group target value method with the following formula:

where n is the sample size; $Z_{1-\alpha/2}$ and $Z_{1-\beta}$ are the significance level and the fractional positions of the standard normal distribution of the degree of certainty, respectively; P_0 is the clinically acceptable standard of the evaluation index; and P_T is the expected value of the evaluation index of the assessment reagent. The positive and negative compliance rates of the assessment reagents and control reagents should reach 80% and 85% clinically, respectively, and the expected positive and negative compliance rates of the assessment reagents and control reagents can reach 90%. The minimum positive and negative sample sizes were estimated to be 137 and 470 cases, respectively, when the significance level α=.05 and the degree of certainty β=.20. Therefore, the minimum total sample size was estimated to be 607 cases. Sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios were calculated using Microsoft Excel 2021 (Microsoft Corp). Sensitivity was calculated as (true positives) / (true positives + false negatives) × 100. Specificity was calculated as (true negatives) / (true negatives + false positives) × 100. Positive predictive value was calculated as (true positives) / (true positives + false positives) × 100. Negative predictive value was calculated as (true negatives) / (true negatives + false negatives) ×100. The positive likelihood ratio was calculated as sensitivity / (1 – specificity). The negative likelihood ratio was calculated as specificity / (1 – sensitivity). $P<.05$ was considered statistically significant. Figures were constructed using GraphPad Prism 8.0 (GraphPad) and R software (R Foundation for Statistical Computing).

Results

Comparison of the Pluslife Rapid SARS-CoV-2 Detection Kit With RT-qPCR

The results of the Pluslife rapid SARS-CoV-2 test and RT-qPCR in the 4 centers are shown in Table 2. Of 1950 patients tested using RT-qPCR, a total of 1547 (79.3%) were negative and 403 (20.7%) were positive. Of the 1547 negative patients, 1535 (99.2%, 95% CI 98.8%-99.7%) were also negative when tested with the Pluslife rapid SARS-CoV-2 test. Of the 403 positive patients, 318 (78.9%, 95% CI 76.8%-81%) were also positive when tested with the Pluslife rapid SARS-CoV-2 test. The overall compliance rate was 95% (1853/1950; 95% CI 93.9%-96.2%), with a κ coefficient of 0.84 (95% CI 0.82-0.86). The positive predictive value (PPV) was 96.4% (318/330; 95% CI 95.4%-97.3%), and the negative predictive value (NPV) was 94.8% (1535/1620; 95% CI 93.6%-95.9%). Next, we adopted 2 study protocols, namely, the “ready-to-test” and “freeze-thaw test” protocols, for the field trial scenarios, in which Guangdong Provincial People’s Hospital, Zhujiang Hospital at Southern Medical University, and Guangdong University of Finance and Economics were used in the “ready-to-test” scenarios and Guangdong Second Provincial General Hospital in the “freeze-thaw test” scenario. The results of the Pluslife rapid SARS-CoV-2 test and RT-qPCR are shown in Table 3. RT-qPCR results were negative for 1332 (92%) and positive for 115 (8%) of 1447 patients. Of the 1332 negative patients, 1323 (99.3%, 95% CI 98.9%-99.8%) were also negative when tested with the Pluslife rapid SARS-CoV-2 test. Of the 115 positive patients, 113 (98.3%, 95% CI 98%-100%) were also positive when tested with the Pluslife rapid SARS-CoV-2 test. The overall compliance rate was 99.2% (1436/1447; 95% CI 98.8%-99.7%), with a κ coefficient of 0.95 (95% CI 0.94-0.96). The PPV was 92.3% (113/122; 95% CI 91.3%-94%), and the NPV was 99.8% (1323/1325; 95% CI 99.6%-100%). The results of the “freeze-thaw test” for the Pluslife rapid SARS-CoV-2 test and RT-qPCR are shown in Table 4. A total of 215 (42.7%) of 503 patients received negative RT-qPCR results, and 288 (57.3%) received positive results. Of the 215 negative patients, 212 (98.6%, 95% CI 98%-99.2%) were also negative when tested with the Pluslife rapid SARS-CoV-2 test. Of the 288 patients positive, 205 (71.2%, 95% CI 68.8%-73.5%) were also positive when tested with the Pluslife rapid SARS-CoV-2 test. The overall compliance rate was 82.9% (417/503; 95% CI 81%-84.8%), and the κ coefficient was 0.67 (95% CI 0.64-0.69). The PPV was 98.6% (205/208; 95% CI 97.9%-99.2%), and the NPV was 71.9% (212/295; 95% CI 69.6%-74.2%). These data suggest that the “ready-to-test” scenario was better than the “freeze-thaw test” scenario.

Table 2. Performance of the Pluslife SARS-CoV-2 rapid test compared to real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) for 1950 samples from all trial scenarios.

Pluslife SARS-CoV-2 rapid test results	RT-qPCR results		Performance
	Positive (n=403)	Negative (n=1547)	
Positive (n=330), n	318	12	—[a]
Negative (n=1620), n	85	1535	—
Sensitivity (n=403), n (%, 95% CI)	—	—	318 (78.9, 76.8-81)
Specificity (n=1547), n (%, 95% CI)	—	—	1535 (99.2, 98.8-99.9)
Overall compliance rate (n=1950), n (%, 95% CI)	—	—	1853 (95, 93.9-96.2)
κ (95% CI)	—	—	0.84 (0.82-0.86)
Positive predictive value (n=330), n (%, 95% CI)	—	—	318 (96.4, 95.4-97.3)
Negative predictive value (n=1620), n (%, 95% CI)	—	—	1535 (94.8, 93.6-95.9)
Positive likelihood ratio	—	—	101.73
Negative likelihood ratio	—	—	0.21

[a] Not applicable.

Table 3. Performance of the Pluslife SARS-CoV-2 rapid test compared to real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) for 1447 samples from all “ready-to-test” scenarios.

Pluslife SARS-CoV-2 rapid test results	RT-qPCR results		Performance
	Positive (n=115)	Negative (n=1332)	
Positive (n=122), n	113	9	—[a]
Negative (n=1325), n	2	1323	—
Sensitivity (n=115), n (%, 95% CI)	—	—	113 (98.3, 98-100)
Specificity (n=1332), n (%, 95% CI)	—	—	1323 (99.3, 98.9-99.8)
Overall compliance rate (n=1447), n (%, 95% CI)	—	—	1436 (99.2, 98.8-99.7)
κ (95% CI)	—	—	0.95 (0.94-0.96)
Positive predictive value (n=122), n (%, 95% CI)	—	—	113 (92.3, 91.3-94)
Negative predictive value (n=1325), n (%, 95% CI)	—	—	1323 (99.8, 99.6-100)
Positive likelihood ratio	—	—	145.45
Negative likelihood ratio	—	—	0.02

[a] Not applicable.

Table 4. Performance of the Pluslife SARS-CoV-2 rapid test compared to real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) for 503 samples from all “freeze-thaw test” scenarios.

Pluslife SARS-CoV-2 rapid test results	RT-qPCR results		Performance
	Positive (n=288)	Negative (n=215)	
Positive (n=208), n	205	3	—[a]
Negative (n=295), n	83	212	—
Sensitivity (n=288), n (%, 95% CI)	—	—	205 (71.2, 68.8-73.5)
Specificity (n=215), n (%, 95% CI)	—	—	212 (98.6, 98-99.2)
Overall compliance rate (n=503), n (%, 95% CI)	—	—	417 (82.9, 81-84.8)
κ (95% CI)	—	—	0.67 (0.64-0.69)
Positive predictive value (n=208), n (%, 95% CI)	—	—	205 (98.6, 97.9-99.2)
Negative predictive value (n=295), n (%, 95% CI)	—	—	212 (71.9, 69.6-74.2)
Positive likelihood ratio	—	—	51.01
Negative likelihood ratio	—	—	0.67

[a] Not applicable.

In addition, we compared the testing time of RT-qPCR and the Pluslife rapid SARS-CoV-2 test in the 4 centers, and the results are shown in Table 5. Among participants with negative nucleic acid test results, the mean testing time for RT-qPCR (534.70, SD 466.94 minutes) was substantially longer than that for the Pluslife rapid SARS-CoV-2 test (35.00, SD 0 minutes). Similarly, among participants with positive nucleic acid test results, the testing time for RT-qPCR (882.20, SD 517.97 minutes) was substantially longer than that for the Pluslife rapid SARS-CoV-2 test (16.49, SD 5.51 minutes). Moreover, similar results were obtained when comparing the testing times obtained using the 2 testing methods for the 4 field trial sites (Multimedia Appendix 2).

Table 5. Comparison of the testing times of the Pluslife SARS-CoV-2 rapid test and real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) at the 4 centers.

Testing result	Testing time (min), mean (SD)	*P* value
Negative		<.001
RT-qPCR	534.70 (466.94)	
Pluslife SARS-CoV-2 rapid test	35.00 (0)	
Positive		<.001
RT-qPCR	882.20 (517.97)	
Pluslife SARS-CoV-2 rapid test	16.49 (5.51)	

Comparison of the Pluslife Rapid SARS-CoV-2 Test and Commercial Antigen Test With the RT-qPCR Nucleic Acid Test

A total of 283 samples that underwent testing using both the commercial antigen test and Pluslife rapid SARS-CoV-2 test were included. RT-qPCR confirmed that there were 93 positive samples and 190 negative samples. The results of the 2 methodological assays compared with the RT-qPCR results are shown in Table 6. The data showed that, compared to the RT-qPCR results, the sensitivity of the commercial antigen test was 88.2% (82/93; 95% CI 86.5%-89.8%) with a specificity of 100%, and the sensitivity of the Pluslife rapid SARS-CoV-2 test was 98.9% (92/93; 95% CI 98.4%-99.4%) with a specificity of 100%.

Table 6. Comparison of the results of testing 283 samples using the 2 assays (commercial antigen kit and Pluslife SARS-CoV-2 rapid test kit) with real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) results.

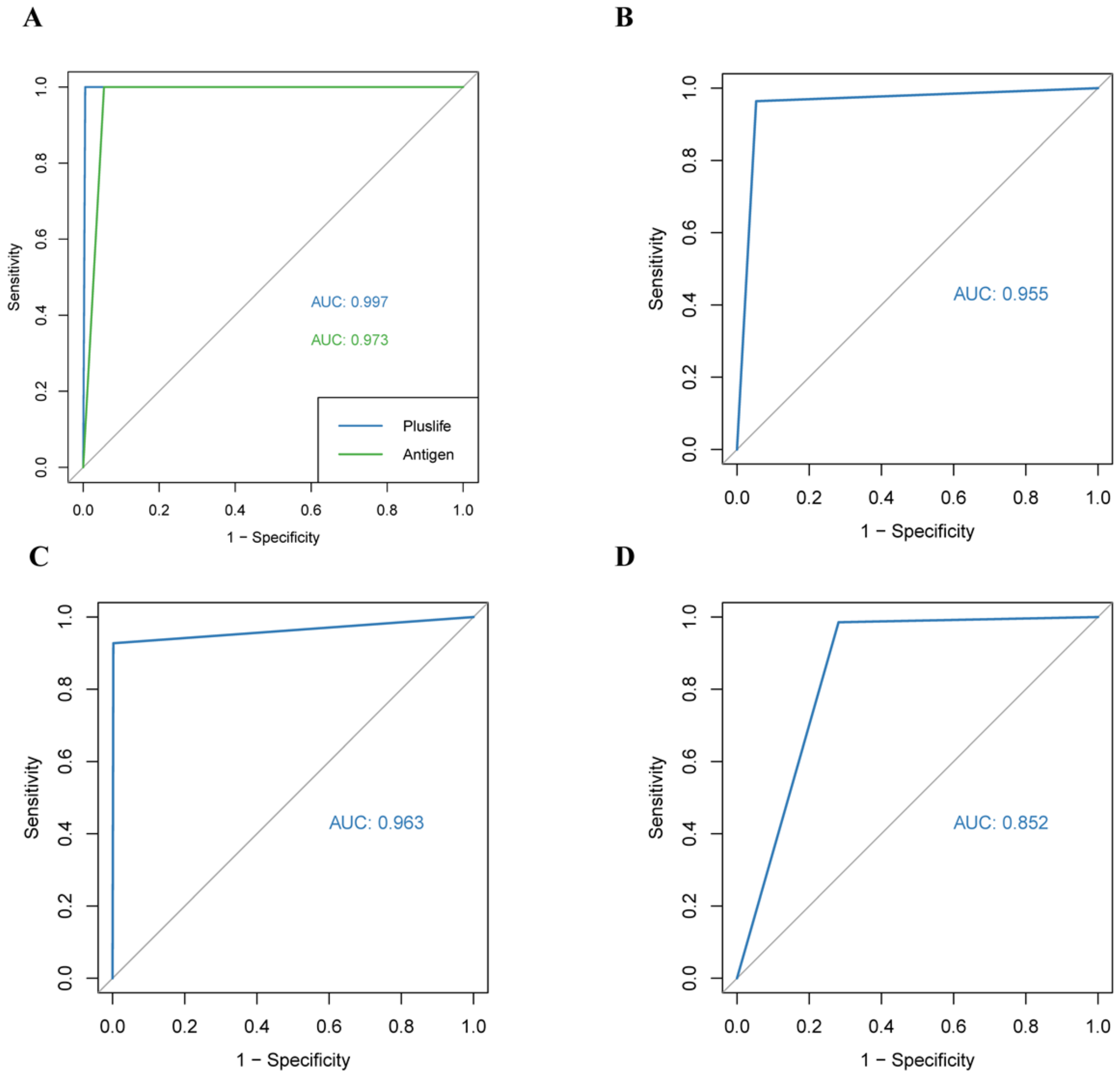
Assay results	RT-qPCR results		Performance
	Positive (n=93)	Negative (n=190)	
Commercial antigen test			
Positive (n=82), n	82	0	—[a]
Negative (n=201), n	11	190	—
Sensitivity (n=93), n (%, 95% CI)	—	—	82 (88.2, 86.5-89.8)
Specificity (n=190), n (%)	—	—	190 (100)
Pluslife SARS-CoV-2 rapid test			
Positive (n=92), n	92	0	—
Negative (n=191), n	1	190	—
Sensitivity (n=93), n (%, 95% CI)	—	—	92 (98.9, 98.4-99.4)
Specificity (n=190), n (%)	—	—	190 (100)

[a] Not applicable.

Comparison of the Diagnostic Value of Different Testing Methods and Different Trial Sites

We performed a receiver operating characteristic (ROC) curve analysis to assess the predictive value of the diagnostic variables. The optimal cutoff values and the area under the receiver operating characteristic curve (AUROC) were first calculated for the commercial antigen test and the Pluslife rapid SARS-CoV-2 test. The AUROCs were 0.973 and 0.997 for the commercial antigen test and the Pluslife rapid SARS-CoV-2 test, respectively (Figure 2A), indicating the higher diagnostic value of the Pluslife rapid SARS-CoV-2 test. After that, the optimal cutoff values and AUROC of the Pluslife rapid SARS-CoV-2 test were calculated for the different study protocols and for each field trial site. The AUROCs were 0.955, 0.963, and 0.852 for the 4 centers combined, the "ready-to-test" scenario, and the "freeze-thaw test" scenario, respectively (Figure 2B-D), indicating that the diagnostic value of testing immediately after field sample collection was higher than that after freezing and thawing samples. The AUROCs of the Pluslife rapid SARS-CoV-2 test for each individual site were 0.959, 1.000, 0.852, and 1.000 for Guangdong Provincial People's Hospital, Zhujiang Hospital at Southern Medical University, Guangdong Second Provincial General Hospital, and Guangdong University of Finance and Economics, respectively (Multimedia Appendix 3).

Figure 2. Analysis of the diagnostic value of different assay methodologies and different study protocols. (A) ROC curve analysis comparing the diagnostic value of antigen testing with that of the Pluslife SARS-CoV-2 rapid test. ROC curves predicting the diagnostic value of the Pluslife SARS-CoV-2 rapid test for (B) the 4 centers, (C) the "ready-to-test" study protocol, and (D) the "freeze-thaw test" study protocol. AUC: area under the curve; ROC: receiver operating characteristic.



Discussion

Principal Findings

Molecular testing is currently the standard diagnostic method for confirming SARS-CoV-2 infection [9], and RT-qPCR for SARS-CoV-2 in respiratory specimens is widely used in COVID-19 diagnostic laboratories [2,10-12]. The easy-to-use and mass-producible nature of the Pluslife rapid SARS-CoV-2 test allows for timely nucleic acid test results and reduces the burden on laboratories [10]. Our objective was to evaluate the performance of the Pluslife rapid SARS-CoV-2 test by comparing it with other tests, such as RT-qPCR and antigen tests, through field trials of the Pluslife rapid SARS-CoV-2 test kit at 4 centers.

First, our results from the 4 field trial sites revealed that RT-qPCR showed 95% (1853/1950; 95% CI 93.9%-96.2%) compliance with the results of the Pluslife rapid SARS-CoV-2 test. The specificity, sensitivity, concordance, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of the Pluslife rapid SARS-CoV-2 test were high (Table 2). To test the optimal conditions for the Pluslife rapid SARS-CoV-2 test, we adopted the "ready-to-test" and "freeze-thaw test" protocols in the field trial sites. Guangdong Provincial People's Hospital, Zhujiang Hospital at Southern Medical University, and Guangdong University of Finance and Economics were used as the test sites for the "ready-to-test" scenario, and Guangdong Second Provincial General Hospital was used as the test site for the "freeze-thaw test" scenario. Our results showed that the RT-qPCR results of the "ready-to-test" and "freeze-thaw test" protocols were 99.2% (1436/1447; 95% CI 98.8%-99.7%) and 82.9% (417/503; 95% CI 81%-84.8%), respectively, and the specificity, sensitivity, concordance, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of the SARS-CoV-2 assay were higher with the Pluslife rapid SARS-CoV-2 test. The above results indicated that freezing and thawing samples after collection and then testing with the Pluslife rapid SARS-CoV-2 test kit is likely to cause false-negative test results, while the diagnostic performance of the test is better when the samples are processed in the field immediately, which may be related to the easy degradation of viral RNA after freezing and thawing. In addition, we validated the diagnostic value of the Pluslife rapid SARS-CoV-2 test by ROC curve analysis for the different study protocols and various field trial sites, and the results also showed that the diagnostic value of the test immediately after field collection was higher than that of the test after freezing and thawing the samples. It is well known that RT-qPCR assay results take a longer time to become available [11]. Our comparison of RT-qPCR and Pluslife rapid SARS-CoV-2 testing times in the 4 centers revealed that the Pluslife rapid SARS-CoV-2 testing time was considerably shorter than the RT-qPCR testing time for both negative and positive results.

Rapid antigen testing is now a valuable alternative to RT-qPCR for the diagnosis of SARS-CoV-2 infection because of its simplicity, speed, low cost, and lack of a need for special equipment or skills [13,14]. However, since viral antigens are expressed when the virus is in active replication, the antigen test is more suitable for the acute infection period, and the accuracy of the COVID-19 antigen test is relatively high for the suspected population within 7 days of the onset of symptoms [15,16]. Therefore, there is a need to develop a test that covers a long period and is fast, accurate, and convenient, which is exactly what the Pluslife rapid SARS-CoV-2 test offers. To compare the diagnostic performance of the Pluslife rapid SARS-CoV-2 test with a commercial antigen test, we compared the 2 assays with RT-qPCR in a field trial scenario at Guangdong Provincial People's Hospital, and the results showed that the sensitivity of the antigen and Pluslife rapid SARS-CoV-2 tests were 88.2% (82/93; 95% CI 86.5%-89.8%) and 98.9% (92/93; 95% CI 98.4%-99.4%), respectively. Thus, the positive detection rate of the Pluslife rapid SARS-CoV-2 test was increased by 10.8% (10/93; 95% CI 9.2%-12.4%) compared to the commercial antigen test. In addition, we validated the diagnostic value of these 2 tests by ROC curve analysis, and the results showed a higher diagnostic value for SARS-CoV-2 with the Pluslife rapid SARS-CoV-2 test than with the antigen test. Therefore, although the detection time of the Pluslife rapid SARS-CoV-2 test is slightly longer than that of the antigen test, the diagnostic performance of the Pluslife rapid SARS-CoV-2 test for SARS-CoV-2 is superior to that of the antigen test.

However, there are some limitations to our study. First, because most of the samples were collected by health care workers in the 4 centers, it was not possible to compare the effects of the 2 sampling methods, namely, nasal swabs and oropharyngeal swabs, on the performance of the Pluslife rapid SARS-CoV-2 test. Second, because antigen testing was only performed in a subset of patients in the fever clinic of Guangdong Provincial People's Hospital, the number of samples for which both the Pluslife rapid SARS-CoV-2 test and antigen test were performed was small.

Conclusion

In conclusion, our multicenter field trial showed that, based on the gold standard RT-qPCR assay, the specificity, sensitivity, concordance, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of the Pluslife rapid SARS-CoV-2 kit were high, and the "ready-to-test" assay conditions were optimal. In addition, the detection time of the Pluslife rapid SARS-CoV-2 test was substantially shorter than that of RT-qPCR, which helps to reduce laboratory pressure and time costs with better performance in the detection of SARS-CoV-2. Furthermore, when compared with the antigen test, we found that the diagnostic performance of the Pluslife rapid SARS-CoV-2 test was better for SARS-CoV-2 with similar detection times. Therefore, under "ready-to-test" assay conditions, the Pluslife rapid SARS-CoV-2 test can be used as a fast, convenient, and accurate method for detecting SARS-CoV-2 with a detection performance comparable to that of RT-qPCR, offering significant application value for grassroots level facilities with poor medical conditions, special application scenarios (eg, traffic checkpoints), and easy home testing, thus reducing pressure on medical institutions.

Acknowledgments

This research was funded by the 2023 Guangdong Provincial Special Project on Prevention and Treatment of Novel Coronavirus Infection, China (grant 2023B1111010003).

Data Availability

The data sets generated and analyzed during this study are not publicly available due to participant privacy concerns but are available from the corresponding author on reasonable request.

Authors' Contributions

DZ conceptualized the study and contributed to the methodology, validation, investigation, and writing and revising of the manuscript. JH contributed to the methodology, formal analysis, and revising of the manuscript. BH, D Chen, and XS contributed to the methodology, validation, investigation, and revising of the manuscript. JC, HZ, and AC contributed to the validation and investigation. D Cao and TH supervised the study and contributed to the writing and revising of the manuscript. TH procured resources.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Pluslife SARS-CoV-2 card test procedure.
[PDF File (Adobe PDF File), 463 KB - publichealth_v9i1e48107_app1.pdf]

Multimedia Appendix 2
Comparison of the testing times of the Pluslife SARS-CoV-2 rapid test and real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) at Guangdong Provincial People's Hospital, Zhujiang Hospital at Southern Medical University, Guangdong Second Provincial General Hospital, and Guangdong University of Finance and Economics.
[DOCX File , 17 KB - publichealth_v9i1e48107_app2.docx]

Multimedia Appendix 3
Receiver operating characteristic curves predicting the diagnostic value of the Pluslife SARS-CoV-2 rapid test at (A) Guangdong Provincial People's Hospital, (B) Zhujiang Hospital at Southern Medical University, (C) Guangdong Second Provincial General Hospital, and (D) Guangdong University of Finance and Economics. AUC: area under the curve.
[PDF File (Adobe PDF File), 282 KB - publichealth_v9i1e48107_app3.pdf]

References

1. Zhou P, Yang X, Wang X, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020 Mar;579(7798):270-273 [FREE Full text] [doi: 10.1038/s41586-020-2012-7] [Medline: 32015507]
2. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 2020 Jan;25(3):2000045 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.3.2000045] [Medline: 31992387]
3. Sharfstein JM, Becker SJ, Mello MM. Diagnostic testing for the novel coronavirus. JAMA 2020 Apr 21;323(15):1437-1438. [doi: 10.1001/jama.2020.3864] [Medline: 32150622]
4. Baek YH, Um J, Antigua KJC, Park J, Kim Y, Oh S, et al. Development of a reverse transcription-loop-mediated isothermal amplification as a rapid early-detection method for novel SARS-CoV-2. Emerg Microbes Infect 2020 Dec;9(1):998-1007 [FREE Full text] [doi: 10.1080/22221751.2020.1756698] [Medline: 32306853]
5. Jiang M, Pan W, Arasthfer A, Fang W, Ling L, Fang H, et al. Development and validation of a rapid, single-step reverse transcriptase loop-mediated isothermal amplification (RT-LAMP) system potentially to be used for reliable and high-throughput screening of COVID-19. Front Cell Infect Microbiol 2020;10:331 [FREE Full text] [doi: 10.3389/fcimb.2020.00331] [Medline: 32626666]
6. Yan C, Cui J, Huang L, Du B, Chen L, Xue G, et al. Rapid and visual detection of 2019 novel coronavirus (SARS-CoV-2) by a reverse transcription loop-mediated isothermal amplification assay. Clin Microbiol Infect 2020 Jun;26(6):773-779 [FREE Full text] [doi: 10.1016/j.cmi.2020.04.001] [Medline: 32276116]
7. Broughton JP, Deng X, Yu G, Fasching CL, Servellita V, Singh J, et al. CRISPR-Cas12-based detection of SARS-CoV-2. Nat Biotechnol 2020 Jul;38(7):870-874 [FREE Full text] [doi: 10.1038/s41587-020-0513-4] [Medline: 32300245]
8. Wei Z, Wang X, Feng H, Ji F, Bai D, Dong X, et al. Isothermal nucleic acid amplification technology for rapid detection of virus. Crit Rev Biotechnol 2023 May;43(3):415-432. [doi: 10.1080/07388551.2022.2030295] [Medline: 35156471]
9. Tahamtan A, Ardebili A. Real-time RT-PCR in COVID-19 detection: issues affecting the results. Expert Rev Mol Diagn 2020 May;20(5):453-454 [FREE Full text] [doi: 10.1080/14737159.2020.1757437] [Medline: 32297805]
10. LeBlanc JJ, Gubbay JB, Li Y, Needle R, Arneson SR, Marcino D, et al. Real-time PCR-based SARS-CoV-2 detection in Canadian laboratories. J Clin Virol 2020 Jul;128:104433 [FREE Full text] [doi: 10.1016/j.jcv.2020.104433] [Medline: 32405254]
11. Tang Y, Schmitz JE, Persing DH, Stratton CW. Laboratory diagnosis of COVID-19: current issues and challenges. J Clin Microbiol 2020 May 26;58(6):e00512-20 [FREE Full text] [doi: 10.1128/JCM.00512-20] [Medline: 32245835]
12. van Kasteren PB, van der Veer B, van den Brink S, Wijsman L, de Jonge J, van den Brandt A, et al. Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. J Clin Virol 2020 Jul;128:104412 [FREE Full text] [doi: 10.1016/j.jcv.2020.104412] [Medline: 32416600]
13. Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. N Engl J Med 2020 Dec 17;383(25):2451-2460. [doi: 10.1056/NEJMcp2009575] [Medline: 32412710]
14. Lambert-Niclot S, Cuffel A, Le Pape S, Vauloup-Fellous C, Morand-Joubert L, Roque-Afonso A, et al. Evaluation of a rapid diagnostic assay for detection of SARS-CoV-2 antigen in nasopharyngeal swabs. J Clin Microbiol 2020 Jul 23;58(8):e00977-20 [FREE Full text] [doi: 10.1128/JCM.00977-20] [Medline: 32404480]

15. Sheridan C. Fast, portable tests come online to curb coronavirus pandemic. Nat Biotechnol 2020 May;38(5):515-518. [doi: 10.1038/d41587-020-00010-2] [Medline: 32203294]
16. Scohy A, Anantharajah A, Bodéus M, Kabamba-Mukadi B, Verroken A, Rodriguez-Villalobos H. Low performance of rapid antigen detection test as frontline testing for COVID-19 diagnosis. J Clin Virol 2020 Aug;129:104455 [FREE Full text] [doi: 10.1016/j.jcv.2020.104455] [Medline: 32485618]

Abbreviations

AUROC: area under the receiver operating characteristic curve
RHAM: RNase hybridization-assisted amplification
ROC: receiver operating characteristic
RT-qPCR: real-time reverse transcription quantitative polymerase chain reaction

---

Edited by A Mavragani, T Sanchez; submitted 11.04.23; peer-reviewed by KH Chan, N Xia; comments to author 08.06.23; revised version received 28.06.23; accepted 17.10.23; published 14.11.23.

Please cite as:
Zhu D, Huang J, Hu B, Cao D, Chen D, Song X, Chen J, Zhou H, Cen A, Hou T
Trial of the Pluslife SARS-CoV-2 Nucleic Acid Rapid Test Kit: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e48107
URL: https://publichealth.jmir.org/2023/1/e48107
doi:10.2196/48107
PMID:37962934

Viewpoint

Trusted Data Spaces as a Viable and Sustainable Solution for Networks of Population-Based Patient Registries

Nicholas Nicholson[1], PhD; Sandra Caldeira[1], PhD; Artur Furtado[2], MSc; Ciaran Nicholl[1], PhD

[1]European Commission, Joint Research Centre, Ispra, Italy

[2]European Commission, Directorate General for Health and Food Safety, Luxembourg, Luxembourg

Corresponding Author:
Nicholas Nicholson, PhD
European Commission
Joint Research Centre
via E. Fermi 2749
Ispra, 21027
Italy
Phone: 39 0332 78 ext 9365
Email: Nicholas.Nicholson@ec.europa.eu

Abstract

Harmonization and integration of health data remain as the focus of many ongoing efforts toward the goal of optimizing health and health care policies. Population-based patient registries constitute a critical element of these endeavors. Although their main function is monitoring and surveillance of a particular disease within a given population, they are also an important data source for epidemiology. Comparing indicators across national boundaries brings an extra dimension to the use of registry data, especially in regions where supranational initiatives are or could be coordinated to leverage good practices; this is particularly relevant for the European Union. However, strict data protection laws can unintentionally hamper the efforts of data harmonization to ensure the removal of statistical bias in the individual data sets, thereby compromising the integrated value of registries' data. Consequently, there is the motivation for creating a new paradigm to ensure that registries can operate in an environment that is not unnecessarily restrictive and to allow accurate comparison of data to better ascertain the measures and practices that are most conducive to the public health of societies. The pan-European organizational model of cancer registries, owing to its long and successful establishment, was considered as a sound basis from which to proceed toward such a paradigm. However, it has certain drawbacks, particularly regarding governance, scalability, and resourcing, which are essential elements to consider for a generic patient registry model. These issues are addressed in a proposal of an adapted model that promises a valuable pan-European data resource for epidemiological research, while providing a closely regulated environment for the processing of pseudonymized patient summary data on a broader scale than has hitherto been possible.

(JMIR Public Health Surveill 2023;9:e34123) doi:10.2196/34123

KEYWORDS

population-based patient registries; trusted research environments; registry network model; data federation; cancer registries; noncommunicable diseases

Context

The World Health Organization (WHO) defines public health as "the art and science of preventing disease, prolonging life and promoting health through the organized efforts of society." Surveillance of population health and well-being is listed as the first point in WHO's list of 10 essential public health operations [1].

In practice, the task of monitoring health status is fraught with difficulty. Health data need to be collected from many different points, raising challenges in data harmonization, data contextualization, and data quality. Depending on the type of public health infrastructure, there may additionally be different regional practices that compound these types of problems. Moreover, the introduction of strict data protection laws adds further layers of complexity in linking data, which is an essential part of the process. A wide range of endeavors has been undertaken to align health data, and a certain level of progress has been made through international collaboration on many different aspects of the underlying challenges. However, within Europe, fragmentation of data still persists owing in part to a rich tapestry of different types of national health care infrastructure in place.

In the field of population health monitoring, population-based patient registries are valuable assets in the surveillance and control of disease. Registries expend substantial effort to ensure accuracy and completeness of data in a given population to limit statistical bias [2]. Moreover, registry data are structured and generally well described using standardized coding classifications, which have positioned registries as key epidemiological resources [3,4]. Although registries are relatively autonomous in the population they cover, an extra dimension in their usefulness and value is created by the integration and comparison of data across registries. Combined registry data use allows studies with large cohorts of patients, thus leading to deep analyses and enhanced knowledge [5,6]. Despite the potential considerable savings in resources and benefits to patients that integrating registry data can bring [7,8], there have been few concerted efforts to support or facilitate such initiatives in a sustainable manner. Moreover, increasingly strict data protection regimes can lead to significant administrative overheads in the collection and processing of pseudonymized data, further exacerbating the challenges.

Epidemiological research is particularly affected by strict data protection paradigms that do not generally differentiate the needs of accessing personal data for data linkage purposes from that for studies on specific individuals. Epidemiology requires data linkage solely to derive the patient cohorts of interest (rather than to conduct studies on specific individuals). Moreover, to ensure unbiased intercomparisons of national data, the necessary data validation techniques require access at least to pseudonymized data, which are also considered as sensitive data by the strict data protection paradigms. In view of the current difficulties and barriers confronting initiatives to monitor disease at a supranational level for the purposes of informing and coordinating strategies for control and prevention of disease at the European Union (EU) level, new mechanisms are required to facilitate the necessary access to data without conflicting with the core intent of EU's general data protection regulation (GDPR) legislation.

EU Member States have committed to the United Nations' sustainable development goals to reduce premature mortality owing to noncommunicable diseases (NCDs) [9], and the European Commission (EC) is addressing the rise in numbers of European citizens experiencing the more prevalent chronic diseases [10]. In particular, EC continues to provide support to the work of cancer registration through the European Network of Cancer Registries (ENCR) to derive statistical indicators used to steer EU policies and coordinate Member States' activities for tackling the cancer burden. EC has also funded several projects among the Member States to further the availability of indicators for other diseases [11-13] and supported registry networks via specific project-related activities [14]. However, these initiatives are of relatively short duration, and EU funding cannot replace national support for recurrent activities within country remits. What is missing is a long-term perspective about patient registry networks leading to a sustainable solution. After the conclusion of the joint action initiative on cross-border patient registries in 2014 [14], the European Medicines Agency launched several follow-on initiatives including the creation of an inventory of patient registries and a draft guideline on registry-based studies, which were open to public consultation [15]. Although the aim is to create an EU-wide framework for patient registries, the framework is with reference to facilitating general, open collaboration between the various stakeholders rather than specifically addressing models for integrating registries' data. Although patient registries in Europe are generally willing to establish supranational networks [16,17], for the most part, they struggle to find the resources and instruments to do so, and an EU-coordinated sustainable and scalable model of patient registry networks would surmount many difficulties. In addition, such a model could scale to other regions of the world in which there are agreements between countries for tackling common public health issues, thus further increasing the potential of the approach.

A Missing Infrastructure

Despite the acknowledgment that policy and decision-making processes should be based on evidence and supported by adequate health information systems, Bogaert and Van Oyen [18] observed that “there is no single comprehensive EU-wide public health monitoring system or health information system that allows policy-oriented research or advice.” Although indicators of NCDs have been, and continue to be, estimated from various sources including hospital discharge records, death certificates, health insurance claims data, and health surveys, the quality of such indicators can be extremely variable. This is particularly the case for health insurance data that can lead to biased coverage of the population where there is no universal national health insurance system. For these reasons, insurance data tend to be used for evaluating and assisting health care planning rather than for generating health statistics [19]. Other difficulties arise from the different disease coding systems used and even different interpretations of codes within the same coding system. Moreover, health surveys have limited sample sizes and tend to miss illnesses with high fatality rates [19]. Although in the absence of any systematic collection of data, these estimated processes play an important role, they also place onerous demands on the institutions involved when conducted on a periodic basis, and there will be a trade-off between the resources required and the cost-benefit of the indicators depending on their quality and veracity. Many challenges facing such initiatives are summarized in the Eurostat report on pilot studies performed for collecting morbidity statistics in EU [19].

Regarding the establishment of a dedicated EU health information system, Rosenkötter et al [20] identified the key features expected from an EU health information system based on a series of interactions with stakeholders from international institutions, national ministries, public health authorities, universities, and EU cofunded projects. In a ranking of 10 features, the most important one was considered to be a permanent and sustainable system (in terms of content, infrastructure, and resources), followed by quality of data and information. The authors note that in spite of the acknowledged relevance of public health surveillance and the numerous requests for a comprehensive EU health information system, there seem to have been considerable barriers to hinder its realization. They cite various possible factors such as the data collection processes through different and fragmented national health information systems and the concerns of oversimplified rankings in data comparisons; the specific types of issues related to both these factors have been described by Tijhuis et al [21]. An even more fundamental problem identified by WHO is that “measuring health is conceptually and technically complex, requiring statistical, public health and biomedical knowledge and expertise unique to each disease or programme area” [22], and therefore dependent upon coordination and integrated strategies across many different entities.

Although population-based patient registries may have variable coverage and comparability [19], they nevertheless provide a well-established and systematic data collection process, with data often stretching back many years. Additional value is obtained through their collaboration within supranational networks, as these provide the means to harmonize indicators via agreement on data alignment, validation, and cleaning processes. This is essentially the approach followed by ENCR in the collection and derivation of European cancer indicators.

Stakeholder discussions have since gained momentum toward the establishment of the European Health Data Space (EHDS) [23], which foresees an infrastructure that facilitates access to health data across EU. It is not yet clear whether the envisaged solutions underlying the implementation of EHDS will be sufficient to provide the level of service needed for an accurate comparison of health indicators across Member States. Many challenges still need to be addressed, including the timely provision of data, ensuring the prerequisite quality and harmonization of data, and issues related to data linkage. The extent of these difficulties can be appreciated from the processes followed by ENCR to ensure the harmonization and quality of cancer indicators for meaningful comparison across national boundaries. In view of the underlying verification and validation processes to ensure harmonized quality indicators, it is likely that ENCR will need to furnish EHDS with precomputed cancer indicators rather than the converse scenario, whereby European cancer indicators can be computed from data made available by EHDS. In this regard, EHDS will not necessarily circumvent the need for a network of cancer registries (CRs) or any other supranational registry network providing harmonized indicators at the EU level. Therefore, agreeing with a sustainable approach toward the operation and resourcing of such networks remains as an important objective.

The advantage of a registry network is precisely that it consists of a network of experts that allows it to function as more than a mere data collection and validation point. The WHO’s Health Metrics Network categorizes a health information system into 6 components (resources, indicators, data sources, data management, information products, and dissemination or use) [22], and these components stand to be addressed more holistically in entities such as registry networks, which also allow for supranational coordination and harmonization. Moreover, given that registries collect hospital discharge records and outpatient records, many of them either do or could collect comorbidity data that are considered to be necessary for a comprehensive picture of public health [19].

Building on the Model of ENCR

Overview

To gauge the practical feasibility of sustainable models enabling the collation of accurate and harmonized indicators for intercomparison at the supranational level, experiences and lessons can be drawn from existing, well-established networks. In this regard, ENCR serves as a prime example. We present the network in terms of how it is organized and how it currently operates, especially in view of the associated advantages and disadvantages.

ENCR Organizational Structure

The members of ENCR are national or local CRs. They are not always public authorities. The participation of registries in ENCR is voluntary and has been promoted by EU initiatives since the 1990s. The funding for participation in the network is obtained from the budgets of each registry and from EU cofinancing.

The organizational structure of ENCR consists of a steering committee, with a 3-year rolling mandate. Members are either elected by the member registries or nominated by some of the participant bodies. The steering committee is responsible for prioritizing the issues facing the network and establishing working groups comprising individuals with specific skills or experience from the individual registries. Examples include review and revision of coding recommendations, agreement of data validation rules, and harmonization of variables extending the common data set. The steering committee is assisted by a secretariat that provides administrative and technical support.

The secretariat was initially provided by WHO’s International Agency for Research on Cancer, which played an instrumental role in the formation of the network in 1990 under EC’s Europe Against Cancer action plan. In 2012, the secretariat was transferred to EC’s Joint Research Centre. ENCR has been essential for coordinating the periodic collection of data for deriving the statistical indicators used at the EU level to compare cancer incidence, mortality, and survival and their trends over time for different types of cancer. The indicators are publicly available on the European Cancer Information System (ECIS) website [24], which is the reference point for monitoring and projecting the burden of cancer in Europe.

Challenges Facing the ENCR Model

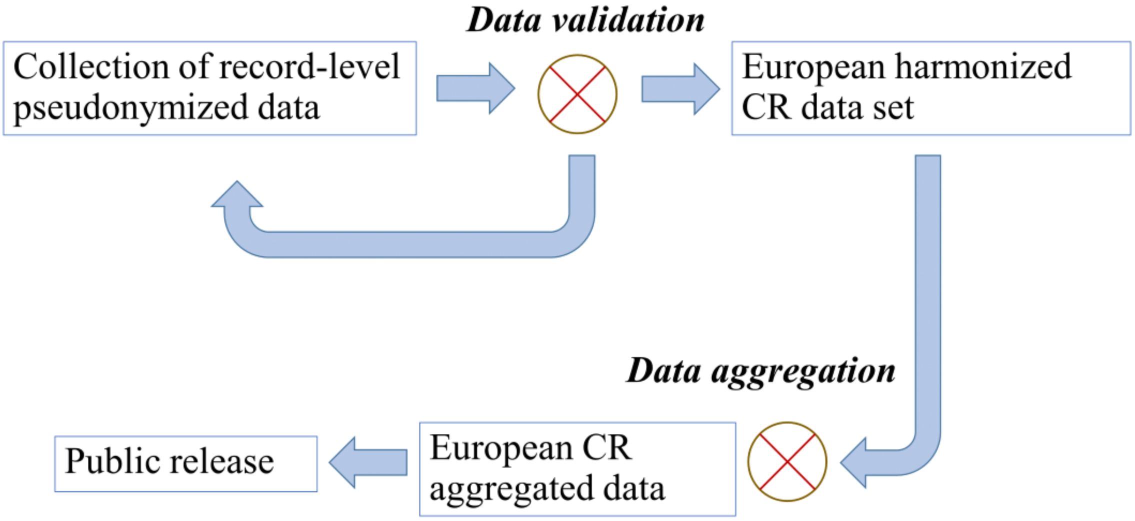
Overview

Ensuring the accuracy of comparable cancer data from many different registries is a nontrivial process. The organizational landscape of CRs is extremely heterogeneous across Europe, and the complexity of the situation can be gauged by the sheer number of registries constituting ENCR (>150 individual CRs). Registries can be nationally based, regionally based, or even metropolitan, and Member State coverage is also quite variable [25]. Challenges relate to data curation, data interfaces, and lack of legal entity status for an informal network.

Data Curation

The collection of data from all CRs relies not only on the good will of the registries to provide their data within a given time frame but also on the legal contracts between the data collector and the data provider. First, the data are provided as pseudonymized patient summary records conforming to a template describing the ENCR common data set, which consists of harmonized variables fundamental for deriving the major statistical indicators. Then, the data are cleaned according to a set of predefined rules; the more basic checks relate to file formatting errors and ensuring that all the mandatory variables have been provided. The more intricate checks verify the accuracy of the variables' data ranges, particularly regarding the values of other variables (intervariable checks). The data cleaning step may require several iterations with the registries until the data are considered to meet the required level of conformity with the data rules, which is necessary for national and international comparisons. Finally, the indicators are derived from the data on a per-registry basis and displayed at an aggregated level that removes the possibility of reidentification. Following a round of final checks with the registries, they are uploaded to the ECIS website. The process is summarized in Figure 1.

Figure 1. Summary of the centralized European cancer-registry data collection and cleaning process.



The processes for collecting, cleaning, and final verification add time delays beyond those already incurred by the CRs in collecting and verifying the data according to their own local processes and requirements. Once the aggregated data are made available at the European level, they are generally 4 to 5 years out of date. Forecasts, based on the historic data, are calculated to cover the intervening time lag; nevertheless, it stands to reason that any possible means of reducing the delays while not compromising the veracity of the data should be promoted. Delayed availability of data compromises their usefulness in feeding back in real time to health care policies, especially in monitoring the immediate effects of the introduction of cancer control initiatives, such as screening programs.

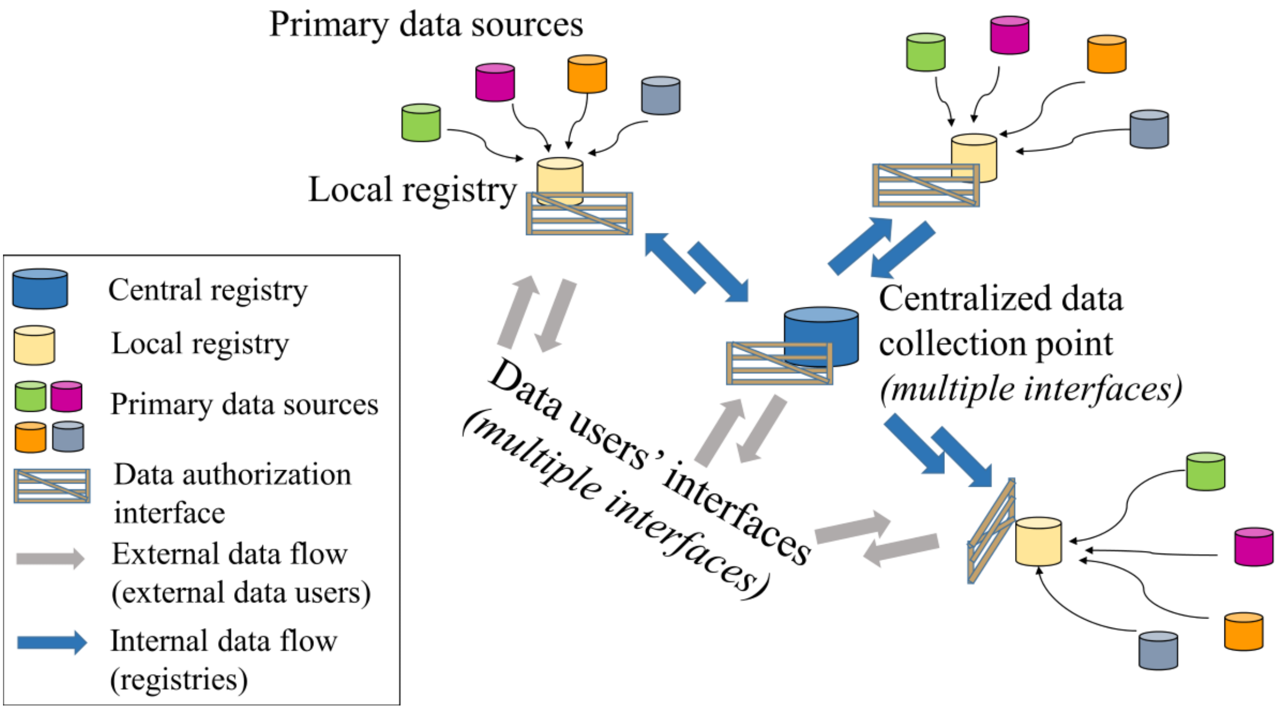
Data Interfaces

Another major difficulty lies with the interfaces between the various independent entities involved in the data chain, especially in terms of data encoding and anonymization.

Encoding of Data

In primary health care settings, clinicians are busy and often may not have the time necessary to complete or verify the summary data that are forwarded to the registry. Medical secretaries entrusted with the work may themselves not have the necessary experience to verify the data or determine the correct codes to use from the case reports. If the data are not encoded accurately, the resulting inconsistencies tend to amplify the inefficiencies as they propagate through the downstream processes. Improved focus and attention given to ensuring the accuracy of the summary case reports at the initial point of encoding could save considerable expense and delay in correcting the data later. However, the same is true for all the interfaces, with each interface adding to the overall delay and inefficiency, as data have to be checked, verified, and corrected via dialogue between both sides of the interface. Figure 2 illustrates the multiple data interfaces in the current registry network scenario. Although these processes are in the domain of national health infrastructure, there are motives for establishing good practices in view of saving cumulative costs, both in terms of time and resources. Data reporting is often viewed as an unnecessary burden in the busy clinical environment, and further efforts could be undertaken to provide the results of data analyses back into the hands of the clinicians to help change such attitudes and demonstrate the practical advantages of accurate data reporting. In addition, measures could be implemented, such as training, data quality audits, or provision of standard templates encapsulating data quality and data semantic contexts following, for example, the model of Findable, Accessible, Interoperable, and Reusable data digital objects [26].

Figure 2. Multiple data interfaces exist in the common pan-European registry model. Each interface adds barriers to the type of data interaction allowed. In particular, the interface presented to general data users is complex on account of the lack of a single access point.



Data Anonymization

Moreover, where the interface bridges different health care authorities, the transfer of pseudonymized data also falls under the data protection requirements of GDPR. Data requesters are required to state explicitly the intended purposes for the data, and these requests have to pass through various overseeing processes to ensure that the purposes are appropriate. The interfaces are further complicated between national boundaries where different legal interpretations and contexts may apply. The way in which GDPR has been enforced across different EU Member States since its introduction in May 2018 means that some registries are no longer able to provide even pseudonymized summary data. However, anonymized data cannot be used for verifying data consistency and bias. This development has profound consequences for the reliable intercomparability of the indicators currently provided in ECIS. Although initiatives are underway to federate the data cleaning processes, which would eventually remove the need to collect record-level data, it is not a straightforward task because there are many exceptional cases to handle and apply through a consistent approach. Moreover, merely federating the centralized entity's processes does not tackle the broad and more important issues of leveraging the full value of the registries' data (in a centralized data collection process, the focus tends to be on only the reduced subset of the core data variables).

The Legal Entity Issue

The ease of acquiring data, at least from the perspective of recognized legitimacy of the data request, can be facilitated via a legal mandate. However, a legal mandate requires a data-requesting entity to be a legal entity, and although this does not pose any problems for individual registries, an informal network of registries is itself not a legal entity. The lack of legal entity status also confounds funding of the network or any financial transaction. This means that the network is fully dependent on the legal entity or entities that perform the administrative affairs of the network.

Proposal for an Adapted Model

Overview

The considerations mentioned in the previous section can be used to frame a proposal toward a more robust and efficient model by streamlining the processes, increasing coordination among the various entities, and providing the capacity to scale across other registry domains. Apart from cancer, other chronic disease domains receiving attention for more coordinated European initiatives include diabetes, cardiovascular disease, and chronic obstructive pulmonary disease. The need to increase the availability and access to NCD data for both policy and research uses has long been argued within Member State public health institutes and by experts and stakeholders in the field. However, important challenges remain before the provision of accurate, comparable data on key aspects can be realized—even of fundamental information such as the prevalence of diabetes, cardiovascular disease, or chronic obstructive pulmonary disease. Moreover, it would be particularly valuable to be able to track and derive geographical and temporal trends of epidemiological indicators (such as prevalence, incidence, mortality, and survival) for all the major NCDs and any comorbidities.

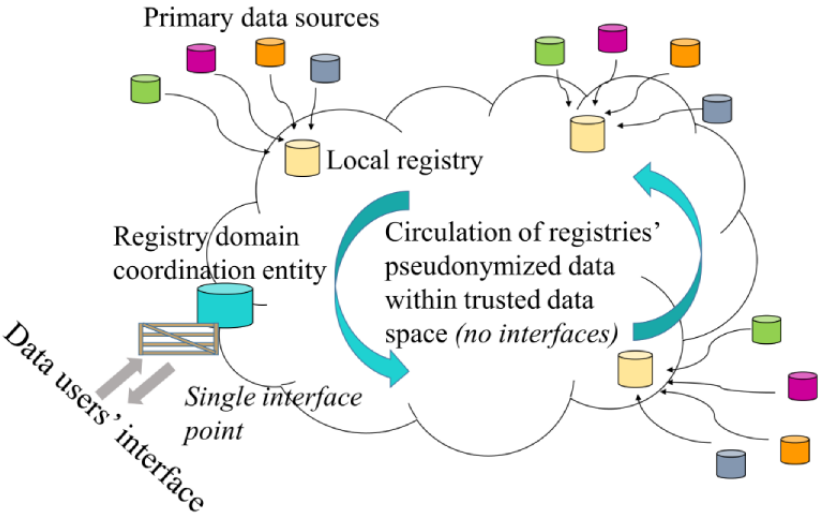
We argue that a paradigm shift with the potential for a significant increase in efficiency is achievable if all the superfluous external interfaces are removed. This could be realized by creating a trusted registry space, for example, similar to trusted research environments (TREs) being established in the United Kingdom for data-driven health research [27]. An environment of this nature would create a secure and trusted working space in which registries within a given disease domain could operate across regional and national boundaries.

Pan-EU TREs

In TREs, researchers can only access the types of data agreed upon within their data sharing agreements. They are also prevented from downloading any record-level data and can only access the data within the TRE. Creating such an EU-wide environment on a registry domain basis would effectively unleash the full potential of registry data for the benefit of all citizens. Given that all registries have the same objectives, albeit within their own specific geographic boundaries, and share the same ethical principles, there is, in principle, no reason why such a pan-EU space could not be set up, especially when safeguarded by data sharing agreements. Moreover, the data space would generally only require access to pseudonymized data, consistent with the principles laid out by GDPR to reduce the risks of patient identification. In any case, the fundamental aim of population-based registries is to provide a monitoring and surveillance function for a particular disease; identification or reidentification of patients is not the purpose, and personal identifiers are only used when it is necessary for linking data to obtain the relevant information for selecting cohorts of patients sharing the specific commonalities to which the epidemiological research is addressed. Furthermore, pseudonymized data need never leave the trusted registry space. The processing would be performed entirely within the data space, and only anonymized data would be made available outside the space.

A means for formalizing the network is also needed such that it is recognized as a legal entity. This could be achieved via an administrative and technical entity within the network (and operating in the space, thereby removing a further unnecessary interface and dependence on an external organization) that could manage the budgeting and financial aspects of the network. A registry network with legal status would resolve various difficulties that are becoming increasingly apparent, especially regarding the interface with EHDS, which foresees data nodes as legal entities either at the national level or formally recognized European agencies or bodies (such as the European Centre for Disease Prevention and Control, European Reference Networks, and European Research Infrastructure Consortia). In contrast, registry networks are informal networks of registries that have no legal recognition even though each individual member registry may be a legal entity in its own right. However, with a legal entity status, the administrative entity of the registry network could be formalized as the network's contact office and serve as the registry domain node in EHDS. Such a contact office may also be able to rationalize the many current access points vis-à-vis the national statistical offices and those offices responsible for providing information to registries. Figure 3 illustrates the benefits of replacing the multiple access points shown in Figure 2 with a single data access point.

Figure 3. A trusted registry environment or data space for a specific disease domain. The registries within the trusted data space are free to exchange pseudonymized data on the condition that no data leave the space and that the purpose of the exchange is legally and ethically compliant. A single interface point is provided by the coordination entity that works in coordination with all the registries in the data space.



Benefits of the Model

A trusted data space on a disease domain basis, in which registries are able to share pseudonymized data, would transform the current processes of working and considerably facilitate supranational research (such as at the EU level) that can currently only be performed with great difficulty and effort. It would also accelerate the provision of statistical indicators by removing many administrative and bureaucratic hurdles impeding the processes for collating data. An additional potential benefit is that it could provide a safe environment for the data, because the data need never leave the trusted registry space, as is the case in centralized collection points, and thereby motivate the sharing of a rich data set with more variables than are currently included in the minimum data set.

The legal status of the contact office forms a critical element in the network's business case, as funding for the activities of the network could be provided via administrative and data processing fees to ensure the long-term sustainability of the network. Regarding this aspect, the benefits should be emphasized on the positive feedback mechanism that secondary data use is likely to provide on the quality of data. The more the data are used, the more resources would be available to the network to reinvest in developing more robust processes and tools for verifying the quality of data and extending the set of harmonized variables. Moreover, these resources could in turn be used to support and improve the quality of the primary data feed processes.

A further strength of the model would be to focus the access point to the registry data on a single entity, not only helping to ensure that the data requested from different registries within the same domain are harmonized to the degree necessary for secondary analyses but also allowing a shared view within the community about the ethical aspects and appropriate use of registry data in the incoming data requests. Individual registries within the network would always retain the right to refuse the use of their data in any specific study. Table 1 summarizes the characteristics and advantages of the proposed model over those of the current ENCR model for some of the main data processes.

Table 1. Characteristics and advantages of the proposed adapted model compared with the characteristics of the current ENCR[a] model for some of the main data processes.

Process	ENCR model	Proposed adapted model	Advantages
Data collection	Centralized data collection point	Data remain federated but accessible in data space	• All registries have access to other registries' data • Parallelization of centralized tasks, thus avoiding serial data pipeline congestion • Improved transparency and verification of data processing • Patients' record-level data remain within the secure data space
Registry interface	Multiple data interfaces between the central data point and registries	No interfaces between registries (data sharing agreements are formalized via initial set of contracts)	• Fast availability of data • Few administrative processes • More rapid derivation of statistical indicators • Avoids the need for duplicating data sets and the consequent data integrity issues
Data access	Only the central data point has access to full set of pseudonymized data	All registries can access other registries' pseudonymized data	• More extensive data use, which in turn improves data quality • Increased collaborative analyses between registries • Deep data pool
Data variables	Data variables are limited to those described in common data set	All data variables are potentially available	• High-resolution analyses • Facilitation of metadata harmonization (improved visibility of nonstandardized metadata)
Users' interface	Multiple data interfaces between external data users and registry sources	Single data interface between external data users and registry network	• Improved control over data collections and versions, thus avoiding discrepancies in data analyses arising from release of different collections or versions of data
EHDS[b] interface	No clear interface with EHDS	Interface with EHDS via legal entity representative of the registry network	• Well-defined point for data query submissions • Fast response times to data queries • Improved coordination between registries for handling external data requests

[a] ENCR: European Network of Cancer Registries.

[b] EHDS: European Health Data Space.

There are other benefits that such spaces could bring to the registries' data feeding processes from the initiating sources, whether from primary sources such as hospitals and clinics or from linkage data (eg, from statistical offices). Registries spend significant amounts of time ensuring consistency and completeness of data [28], both of which are critical processes to remove bias from the data. The time spent in ensuring these particular dimensions of data quality is one of the major contributors to the time delay associated with the provision of validated registry data. Agreeing the common data sets and metadata standards at the European level stand to influence clinical data recording practices across EU to ensure that the common variables are transferred consistently and accurately and to lead to economies of scale by nurturing approaches toward common data quality tools and methodologies. Having 1 point of access to the registry domain would itself facilitate the transfer of death certificate information, which is a piece of critical information required by registries for ensuring data completeness. Therefore, the network could act as a type of data-clearing house, providing a single interface between national statistical offices rather than multiple interfaces that exist currently.

Within the EU at least, there are few reasons to argue against a model that goes in the direction of a trusted data space in which registries are free to exchange pseudonymized data under appropriate and formalized safeguards. It is based on similar arguments as those used to frame the development goals of EHDS but with more justified reason given the very specific aims. The model would provide registries with the autonomy to agree the common data models, metadata, common data sets, and data quality criteria. The latter is essential in any cross-comparison of data and could easily be audited and graded in such a common space, with important ramifications on the trustability of the data in downstream analyses. The model is adaptable to countries at either end of the country income spectrum because it builds on the local registry infrastructure. A study undertaken on the overall costs of CRs for 2010 [29] revealed that costs per inhabitant ranged between €0.03 (US $0.04) and €0.97 (US $1.35), with an average of €0.27 (US $0.37) in comparison with cancer health care of €102 [US $142.8] per EU citizen in 2009 [30]), whereas the costs per registered cancer case ranged between €6 (US $8.4) and €213 (US $298.2), with an average of €50.71 (US $70.9). These figures were comparable with those of registries in the United States. Costs for registries covering large populations were less than those for registries covering small populations, and correspondingly, costs of national registries fared better than regional registries. Increasing expenditure for CRs was associated with the increasing economic wealth of the country, with the greatest proportion of costs being spent on personnel. In view of the health expenditure savings that registry data can bring, the costs are not disproportionate to the total health care costs—at least in the domain of cancer.

Practical Considerations

A framework allowing such seamless pooling and processing of data, even if only pseudonymized data, would require agreement of national authorities at the EU Member State level despite the provisions laid out in GDPR for processing of sensitive data for reasons relating to the public interest. This would be the first major task to accomplish. However, several aspects will also need to be addressed at a practical level. The first aspect is to identify the means by which a network of legal entities could itself become a legal entity in its own right; having the network agree jointly on terms and implementation would not be sufficient. The solution to this could perhaps be bound together with the legal issues defining a trusted patient registry domain data space. The second aspect relates to the technical aspects of creating a secure EU data space accessible to registries across EU, which may be more difficult to accomplish than a trusted data space housed in and limited to a particular country. The legal and technical aspects may further be compounded by the distinction between EU Member States and European non-EU countries, which are also members of patient registry networks such as ENCR. The third aspect relates to the need for funding, at least initially, before a patient registry domain space could become self-sustaining. A more ambitious endeavor could be to realize a single cross–disease domain patient registry data space, in which resources could be combined to realize a large synergistic technical facility for the development of common processes and tools for data cleaning and validation, without reinventing similar processes and tools on a per patient registry domain basis. There is a risk that a multiplicity of TREs without strong coordination can itself lead to duplicated effort, monopolies on access, and obstructive divergence around data curation [31].

Conclusions

To overcome the restrictions severely limiting the benefits to public health that could otherwise be achieved by facilitating interregistry integration, a new model is urgently required. The urgency is precipitated by the introduction of GDPR that has changed the context of the former model of collecting national and regional registry data for EU-coordinated actions. A model has been presented that promises to overcome many current hurdles and provide an improved solution to unlock the full potential of integrated registry data by allowing access to the full set of data variables rather than a restricted set. The model builds on the concept of ENCR that has proved to be a successful and viable means for deriving harmonized cancer indicators at the European level but introduces elements critical to a more sustainable and generic framework that could be applied to other NCD domains. The 2 elements that are especially important in this regard are the provision of a trusted data space to facilitate access to pseudonymized data in a secure environment and the creation of a legal entity as an integral part of the registry network serving as the formal single representative of and interface to the network. Regarding the legal entity acting as the interface to the registry network, there is a certain degree of trust required by the registries that can be ensured via appropriate contracts; it would also be critical to ensure stringent technical measures against unauthorized access to the data space and against any possibility of downloading sensitive data. Although both these aspects would require support and formalization at the national level and interregistry contractual agreements regarding the sharing of pseudonymized record-level data, there is, in principle, no reason why this cannot be achieved given the similar aims of local or national registries and explicit agreements at the EU level to tackle the NCD burden.

However, regardless of the specificities of any particular model, the fundamental requirements are to remove the inefficiencies imposed by the various interfaces within the current registry network models and allow a degree of autonomy within a registry domain to provide timely, reliable, and accurate indicators for steering policies to tackle the societal burden of disease and improve the outcomes of patients. In particular, accelerating the availability of data could save considerable costs by feeding back the results of health care measures and programs early in the implementation cycle. The arguments presented are intended to show how a solution for unlocking the resource-rich assets of interconnected patient registries is within reach and how a single interface to these data can most optimally be organized in supranational frameworks such as those existing in Europe. This interface would also provide a controlled means for ensuring access to the harmonized set of European health data in a given disease domain and could serve as an access node within EHDS.

Ethical Considerations

No ethics approval was applied for since the work described here does not involve research on any of the following: human embryos and fetuses, human cells or tissues, animals, non-EU countries, the environment or health and safety, dual use, or exclusive focus on civil applications. Nor does it contain reference to any personal data and no personal data were collected in the course of the study. Finally, there is no potential for the adverse misuse of the research results here presented. European Commission Horizon 2020 ethical self-assessment [32].

Acknowledgments

All authors are employed by governmental or supranational entities and report no additional funding for the development of this manuscript.

Disclaimer

The scientific opinions expressed here do not imply a policy position of the European Commission.

Conflicts of Interest

None declared.

References

1. The 10 essential public health operations. World Health Organization. URL: https://www.euro.who.int/__data/assets/pdf_file/0007/172762/Strengthening-public-health-services-and-capacity-an-action-plan-for-Europe-Eng.pdf [accessed 2022-12-19]
2. Tucker TC, Durbin EB, McDowell JK, Huang B. Unlocking the potential of population-based cancer registries. Cancer 2019 Nov 01;125(21):3729-3737 [FREE Full text] [doi: 10.1002/cncr.32355] [Medline: 31381143]
3. Blandin G, Béroud C. Utility of patients' registries to gather clinical, epidemiological and molecular information. Orphanet J Rare Dis 2015 Nov 11;10(S2):O29. [doi: 10.1186/1750-1172-10-s2-o29]
4. Grosclaude P, Zadnik V. Population-based cancer registries: a data stream to help build an evidence-based cancer policy for Europe and for European countries. In: Launoy G, Zadnik V, Coleman MP, editors. Social Environment and Cancer in Europe. Cham: Springer; 2021. URL: https://doi.org/10.1007/978-3-030-69329-9_2
5. Maret-Ouda J, Tao W, Wahlin K, Lagergren J. Nordic registry-based cohort studies: possibilities and pitfalls when combining Nordic registry data. Scand J Public Health 2017 Jul 06;45(17_suppl):14-19. [doi: 10.1177/1403494817702336] [Medline: 28683665]
6. Söderström HK, Räsänen J, Saarnio J, Toikkanen V, Tyrväinen T, Rantanen T, et al. Cohort profile: a nationwide population-based retrospective assessment of oesophageal cancer in the Finnish National Esophago-Gastric Cancer Cohort (FINEGO). BMJ Open 2020 Oct 14;10(10):e039575 [FREE Full text] [doi: 10.1136/bmjopen-2020-039575] [Medline: 33055119]
7. Main K. Saving lives with patient data registries. AMRC. 2018 Nov 28. URL: https://www.amrc.org.uk/Blog/saving-lives-with-patient-data-registries [accessed 2022-10-10]
8. Larsson S, Lawyer P, Garellick G, Lindahl B, Lundström M. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value. Health Aff (Millwood) 2012 Jan;31(1):220-227. [doi: 10.1377/hlthaff.2011.0762] [Medline: 22155485]
9. SDG Target 3.4 | Noncommunicable diseases and mental health: by 2030, reduce by one third premature mortality from non-communicable diseases through prevention and treatment and promote mental health and well-being. World Health Statistics. URL: https://apps.who.int/iris/handle/10665/208282 [accessed 2022-10-10]
10. Non-communicable diseases. European Commission. URL: https://ec.europa.eu/health/non_communicable_diseases/overview_en [accessed 2022-10-10]
11. EUROCHIP European Cancer Health Indicator Project. European Commission. 2003 Jul 15. URL: https://ec.europa.eu/health/ph_projects/2001/monitoring/fp_monitoring_2001_exs_07_en.pdf [accessed 2022-10-10]
12. The BRIDGE Health project. BRIDGE Health. URL: https://www.bridge-health.eu/ [accessed 2022-10-10]
13. InfAct (Information for Action!). INFACT Joint Action on Health Information. URL: https://www.inf-act.eu/ [accessed 2022-10-10]
14. PARENT. EUnetHTA. URL: https://www.eunethta.eu/parent/ [accessed 2022-10-10]
15. Patient registries. European Medicines Agency. URL: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries [accessed 2022-10-10]
16. Zaletel M, Kralj M, Magajne M, Doupi P. Methodological Guidelines and Recommendations for Efficient and Rational Governance of Patient Registries. Ljubljana, Slovenia: European Journal of Public Health; Oct 2015. URL: https://doi.org/10.1093/eurpub/ckv169.006

17. Patient Registries Workshop, 28 October 2016 Observations and recommendations arising from the workshop. European Medicines Agency. 2017. URL: https://www.ema.europa.eu/en/documents/report/report-patient-registries-workshop_en.pdf [accessed 2022-12-22]
18. Bogaert P, Van Oyen H, for BRIDGE Health. An integrated and sustainable EU health information system: national public health institutes' needs and possible benefits. Arch Public Health 2017 Jan 18;75(1):3 [FREE Full text] [doi: 10.1186/s13690-016-0171-7] [Medline: 28116076]
19. Morbidity Statistics in the EU Report on Pilot Studies. eurostat. 2014. URL: https://ec.europa.eu/eurostat/documents/3888793/5858521/KS-TC-14-003-EN-N.pdf/bd959e6e-10ed-4078-915e-308941c02811?t=1422361722000 [accessed 2022-10-10]
20. Rosenkötter N, Achterberg PW, van Bon-Martens MJ, Michelsen K, van Oers HA, Brand H. Key features of an EU health information system: a concept mapping study. Eur J Public Health 2016 Feb 04;26(1):65-70. [doi: 10.1093/eurpub/ckv075] [Medline: 25944870]
21. Tijhuis M, Finger JD, Slobbe L, Sund R, Tolonen H. Data collection. In: Verschuuren M, van Oers H, editors. Population Health Monitoring. Cham: Springer; 2019.
22. World Health Organization. Health Metrics Network & World Health Organization. In: Framework and standards for country health information systems, 2nd ed. World Health Organization. Geneva: World Health Organization; 2008. URL: https://apps.who.int/iris/handle/10665/43872 [accessed 2022-12-29]
23. European health data space. European Commission. URL: https://ec.europa.eu/health/ehealth/dataspace_en [accessed 2022-10-10]
24. ECIS - European Cancer Information System. European Commission. URL: https://ecis.jrc.ec.europa.eu/ [accessed 2022-10-10]
25. Welcome to ENCR. European Network of Cancer Registries. URL: https://encr.eu/ [accessed 2022-10-10]
26. FAIR digital object framework documentation, working draft. FAIR Working Draft. 2022 Oct 27. URL: https://fairdigitalobjectframework.org/ [accessed 2022-10-10]
27. Trusted Research Environment service for England. NHS Digital. URL: https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england [accessed 2022-10-10]
28. Krumholz HM. Registries and selection bias. Circ Cardiovascular Qual Outcome 2009 Nov;2(6):517-518. [doi: 10.1161/circoutcomes.109.916601]
29. Zanetti R, Sacchetto L, Calvia M, Bordoni A, Hakulinen T, Znaor A, Eurocourse WP3 Working Group. Economic evaluation of cancer registration in Europe. J Registry Manag 2014;41(1):31-37. [Medline: 24893186]
30. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. Lancet Oncol 2013 Nov;14(12):1165-1174. [doi: 10.1016/s1470-2045(13)70442-x]
31. Better, broader, safer: using health data for research and analysis. GOV.UK. 2022 Apr 7. URL: https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis [accessed 2022-10-10]
32. Horizon 2020 Programme Guidance: How to complete your ethics self-assessment. Version 6.1. European Commission. 2019 Feb 4. URL: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf [accessed 2022-12-27]

Abbreviations

CR: cancer registry
EC: European Commission
ECIS: European Cancer Information System
EHDS: European Health Data Space
ENCR: European Network of Cancer Registries
EU: European Union
GDPR: general data protection regulation
NCD: noncommunicable disease
TRE: trusted research environment
WHO: World Health Organization

Edited by H Bradley; submitted 07.10.21; peer-reviewed by S Lalla-Edward, N Nekliudov; comments to author 11.05.22; revised version received 31.05.22; accepted 11.10.22; published 13.01.23.

Please cite as:
Nicholson N, Caldeira S, Furtado A, Nicholl C
Trusted Data Spaces as a Viable and Sustainable Solution for Networks of Population-Based Patient Registries
JMIR Public Health Surveill 2023;9:e34123
URL: https://publichealth.jmir.org/2023/1/e34123
doi:10.2196/34123
PMID:36637894

Viewpoint

Technology, Training, and Task Shifting at the World's Largest Mass Gathering in 2025: An Opportunity for Antibiotic Stewardship in India

Isaac H Y Chan[1], MD, MPH; Miriam Gofine[2,3], MPH; Shitij Arora[4], MD; Ahmed Shaikh[5], MD; Satchit Balsari[1,6], MD, MPH

[1]Lakshmi Mittal and Family South Asia Institute, Harvard University, Cambridge, MA, United States
[2]Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States
[3]Vilcek Institute of Graduate Biomedical Sciences, NYU Langone Health, New York, NY, United States
[4]Division of Hospital Medicine, Department of Internal Medicine, Montefiore Medical Center, New York, NY, United States
[5]Institute for Critical Care Medicine, Mount Sinai Hospital, New York, NY, United States
[6]Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Satchit Balsari, MD, MPH
Department of Emergency Medicine
Beth Israel Deaconess Medical Center
Harvard Medical School
1 Deaconess Road
Boston, MA, 02215
United States
Phone: 1 6177542400
Email: sbalsari@bidmc.harvard.edu

Abstract

The role of antibiotic overuse in intensifying selection pressures and contributing to the emergence of antimicrobial resistance is well established. The Kumbh Mela, a religious festival that occurs in 4 Indian cities of spiritual significance, is the world's largest mass gathering, attracting over 80 million pilgrims in 2013. Digital syndromic surveillance from the 2013 and 2015 Melas demonstrated a consistent pattern of antibiotic overuse, with an antibiotic prescribing rate of up to 31% for all patient encounters. As preparations for the 2025 Kumbh Mela begin, task shifting, point-of-care diagnostic and digital tools, robust clinician training, and community awareness can promote the restrained and evidence-based use of antibiotics, minimizing the potential for the emergence of antimicrobial resistance at the world's largest mass gathering.

(JMIR Public Health Surveill 2023;9:e45121) doi:10.2196/45121

KEYWORDS

digital tools; mass gathering; Kumbh Mela; antibiotics; antimicrobial; stewardship; surveillance; public health; informatics; India

Introduction

The World Health Organization's global action plan on antimicrobial resistance (AMR) recognizes that AMR "threatens the very core of modern medicine and the sustainability of an effective, global public health response to the enduring threat of infectious diseases" [1]. The contribution of antimicrobial use and overuse in intensifying selection pressures that lead to the emergence of AMR is well established [2]. India has the highest rate of human antibiotic use in the world and has a substantial burden of antimicrobial-resistant organisms as well [3]. Both the overuse of antimicrobial agents and the emergence of resistant organisms are exacerbated during stressors to population health, as seen during the COVID-19 pandemic [4-6]. These phenomena are also seen in certain regularly scheduled events like large mass gatherings where millions congregate over several days and where wanton prescription of antibiotics has been observed [7].

The Kumbh Mela is one such religious festival held periodically in 4 cities of spiritual significance in India [8]. Millions from across the country visit the weeks-long festival for a religious "dip" in the holy river waters on the banks of which the Melas are held. The Melas were last held in Allahabad in 2013, in Nashik in 2015, in Ujjain in 2016, and in Haridwar in 2021.

Our digital surveillance at 2 of these events in the past, including at the 2013 Kumbh Mela (the world's largest mass gathering to date), shows that antibiotic stewardship at the publicly funded health care facilities at these events is nearly nonexistent [9,10].

Planning for the Melas begins nearly 1 to 2 years in advance, presenting a unique opportunity to strengthen public health systems and improve antibiotic stewardship in time for the 2025 Kumbh Mela in Allahabad, which is again expected to set the record for the largest mass gathering in history. In this viewpoint, we argue for the restrained use of antibiotics; the use of digital tools to monitor, nudge, and support evidence-based prescribing; and for robust clinician training and community awareness to modify expectations and change practice in anticipation of, and during, the 2025 Kumbh Mela.

The 2013 and 2015 Kumbh Melas

Clinical encounters in the 2013 and 2015 Melas in Allahabad and Nashik, respectively, were digitally documented via a syndromic surveillance system deployed at both Melas [8-10]. Some of the authors (SB and AS) were directly involved in extensively studying various aspects of public health preparedness at these Melas and were responsible for the design and implementation of the digital systems, the first of cloud-based syndromic surveillance systems deployed at scale. It is now well documented that significant resources are directed toward public health preparedness and public safety at the Melas, in response to cholera outbreaks and stampedes that have marked previous iterations of the gathering [11]. At each festival, the state and central governments cooperate to provide a range of services for the visiting pilgrims, including temporary shelters, free meals, subsidized transport, and health care. Temporary health clinics are staffed by physicians seconded from other government-run primary clinics across the host state. The clinics are characterized by very high patient volumes, where each provider may see hundreds of patients a day. The doctor-patient encounter is cursory, and patients expect to be prescribed medications for their ailments, in concordance with known health-seeking behavior in India [12,13]. Protocols established by the state health ministry allow the provision of only a 3-day supply of antibiotics, recommending a follow-up visit [9]. With some exceptions, the vast majority of pilgrims only make a day trip to the Mela and return home.

The 2013 syndromic surveillance system was deployed at the 4 busiest of 13 clinics at the Allahabad Mela and at the larger secondary care hospital on site [8,9]. The clinics provided free services to the 80 million pilgrims who visited the Mela over a 55-day period. Roughly 280,000 patients were seen in the clinics during the Mela, and data were captured from 49,131 of these encounters [8,9]. We found that as the number of visitors rose on certain auspicious days, clinic visits spiked with some sites seeing over 2000 patients a day [9]. On February 9, the busiest of all days at the 2013 Mela, over 30 million pilgrims visited the festival. The median age in 49,131 patient encounters recorded at the 5 study sites was 46 years, the majority being men (70%) [9]. The most common recorded presenting complaints were musculoskeletal pain (19%), fever (17%), cough (17%), coryza (16%), and diarrhea (5%) [9]. A total of 91% of patients at the 4 sector clinics and nearly 70% at the central hospital received at least 1 or more drug prescriptions. Of all patient encounters, 31% received an antimicrobial agent [9]. Of patients who presented with a complaint of subjective fever (n=8490), nearly half received a 3-day prescription for antibiotics, empirically, without any confirmatory tests [9].

The syndromic surveillance exercise was repeated 2 years later at the Nashik Kumbh Mela, which was attended by nearly 30 million people [10]. Data were collected from over 40 clinics servicing the visiting pilgrims in the host cities of Nashik and Trimbakeshwar [10]. The clinics mainly ran on the most auspicious days, totaling four 3-day runs of operation, sometimes including both sites. A total of 33,305 unique patient encounters were captured over 9 days. The most common presenting complaints were upper respiratory tract symptoms (n=10,866), joint pain (n=5730), and fever (n=4861). Prescription data were available for 14,725 patient encounters from a subgroup of randomly selected clinics. Within this group, of the 2067 patients who presented with subjective fever, half (n=1111, 53.7%) were prescribed antibiotics empirically (as they were at the 2013 Mela). Among 3921 patients presenting with upper respiratory tract symptoms, 2692 (68.7%) received antibiotics. Of the 1429 patients with gastrointestinal complaints, 668 (46.7%) were prescribed antibiotics (Figure 1).

These findings are concerning for several reasons. Data from these Melas show that antibiotic prescribing rates can be as high as 31% for all encounters. Almost 69% of patients with upper respiratory tract symptoms received antibiotics, an alarmingly high rate, given the overwhelming evidence that the vast majority of upper respiratory tract infections are caused by viral pathogens [14]. These rates are concordant with previously published estimates of antibiotic prescription rates from undifferentiated outpatient settings in India, which range from 39% to 66% [15-17]. There is also substantial interprovider variation in prescribing practices (Figure 2 and Multimedia Appendices 1 and 2). When antibiotics are used, there is little consistency or evidence-based practice to guide decision-making when choosing antibiotics. Across all clinics at the 2015 Kumbh Mela, the rate of fluoroquinolone prescription for patients with upper respiratory tract symptoms ranged from 0% to over 50%, with similar variation in the prescription of other antibiotics. Such use of broad-spectrum antibiotics as empiric therapy is consistent with other observations from India and indicates the need for ongoing education for providers [18,19].

Figure 1. Percentage of patients prescribed antibiotics based on their syndromic presentation.

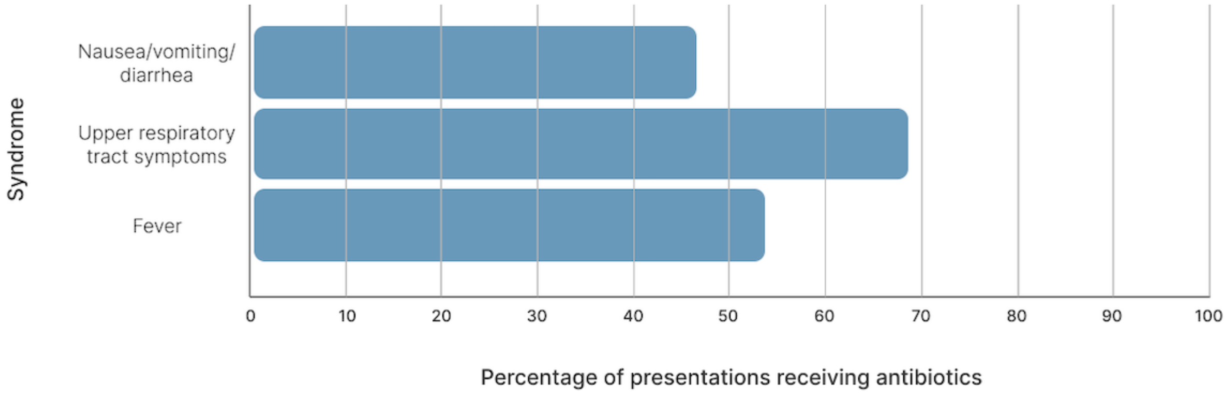
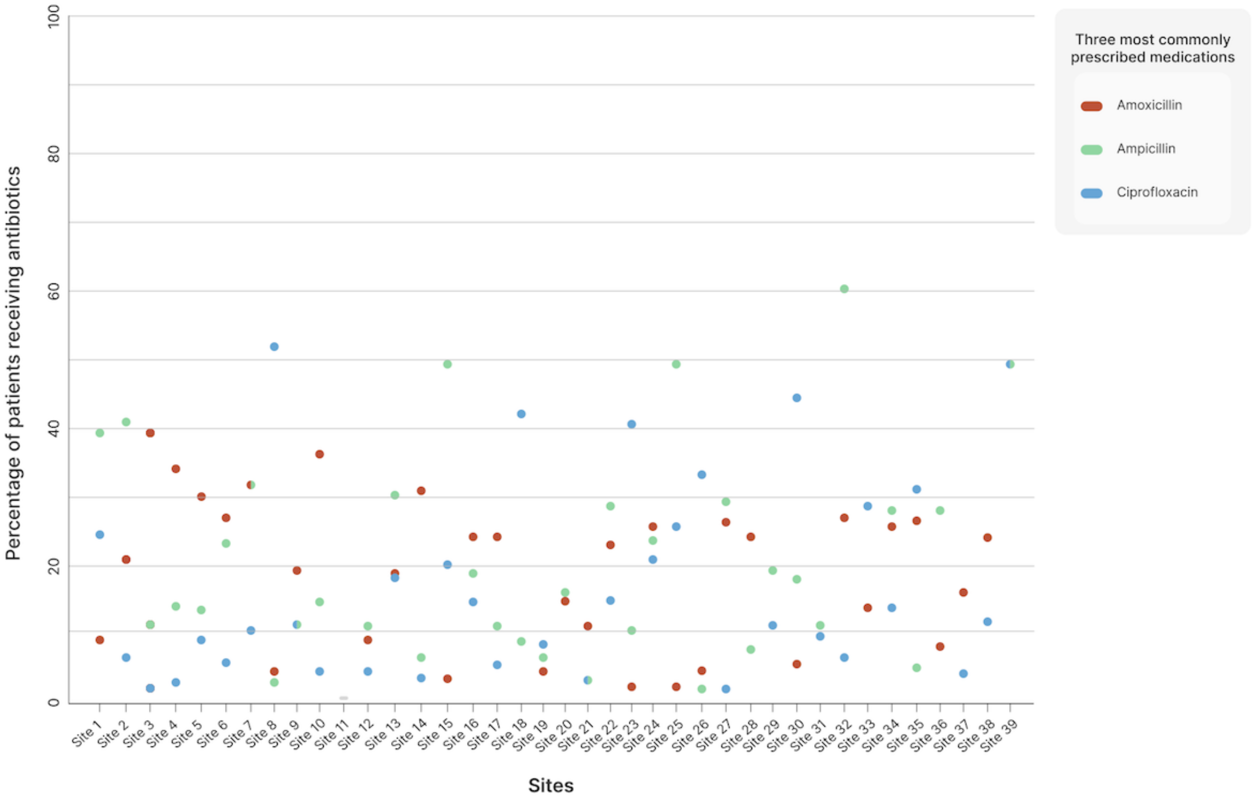


Figure 2. Variation in antibiotic prescription practices across sites for patients presenting with upper respiratory tract symptoms at the 2015 Kumbh Mela.



Policy Recommendations

While well intentioned, the provision of health services for low-acuity illnesses at the Kumbh Mela needs urgent reform. Preparation for the 2025 Allahabad Kumbh Mela commences now as the state assembles its administrative team and various contractors begin to reactivate their supply chains. The Mela administration will have a dedicated team dealing with public health preparedness and health services. Now is the time to reimagine what these should look like in 2025, for what is likely to become the world's largest mass gathering.

Evidence from the prior 2 Melas we studied has shown the following: the vast majority of presentations are of low acuity; there are little to no diagnostics available except at the district hospital; there is near universal overprescription of antibiotics; and when prescribed, the choice and dosage of antibiotics chosen appears arbitrary. Each of these observations warrants attention as planning for the Kumbh begins.

Our recommendations to promote antimicrobial stewardship at the Kumbh Mela are summarized in Table 1.

Table 1. Recommendations to promote antimicrobial stewardship at the Kumbh Mela, the world's largest mass gathering.

Intervention	Impact on antimicrobial stewardship
Highly protocolized triage by mid-level providers (task sharing)	Clinically stable patients not requiring medications are discharged with symptomatic management and patient education
Point-of-care diagnostics (technology)	Targeted, rapid diagnostics may prevent unwarranted prescriptions of antibiotics or help better target therapy
Physician education (training)	Simple, user-centered clinical protocols can guide appropriate antibiotic prescription
Technological assistance (technology)	Benchmarking of one's antibiotic prescription rate against peers acts as a behavioral nudge that rationalizes prescribing practice
Patient education (training)	Patient-facing education initiatives address the patient expectations that drive the demand for antibiotics in the Indian context

1. Task Sharing

All clinics are staffed by trained allopathic physicians. As epidemiological data from these 70,000 patient encounters have consistently shown, patients at the Mela often present with low-acuity symptoms, many of which likely warrant no diagnostics or therapeutics, and most likely do not need the attention of a physician—especially one that has been drawn away from a busy clinic elsewhere in the state in order to participate in a fleeting patient encounter.

In recent years, the Ministry of Health and Family Welfare has been promoting task sharing with mid-level practitioners and with trained nonallopathic physicians. Given what is known about the epidemiology of patient presentations at the Mela, highly protocolized triage by mid-level providers, community health workers, and senior medical students can identify patients who do not require any intervention, as well as those that require a detailed physician review. Patients who meet predefined criteria for clinical stability can be discharged. Such task sharing allows for a more effective allocation of limited physician resources. Fewer patient encounters per physician may reduce decision fatigue, which also contributes to antibiotic initiation [20].

2. Point-of-Care Diagnostics

The proposed task shifting will enable richer doctor-patient interactions for those who have been identified by first-line providers as needing additional attention. Physicians at previous Melas spent less than 3 minutes per patient encounter, often prescribing antibiotics without examining the patient [9]. Previous Melas have also been characterized by a lack of diagnostic capacity, with no laboratory or radiology services outside of the central referral hospital [9]. As a result, only 0.3% of patient encounters at the 2013 Mela used x-rays and almost none received blood tests [9]. The absence or underutilization of diagnostic testing may lead to well-intentioned but inappropriate initiation of empirical antibiotic therapy. Many diagnostic tests, including radiography and rapid antigen testing, for example, can now be performed at the point of care. Such diagnostics must be considered for an appropriately triaged subset of patients in whom empiric therapy cannot be justified.

3. Regulating Antibiotic Prescription

Promoting the appropriate use of antibiotics remains a significant challenge, and multiple approaches have been tried around the world. A multipronged approach will be necessary to promote antibiotic stewardship at the Kumbh Mela.

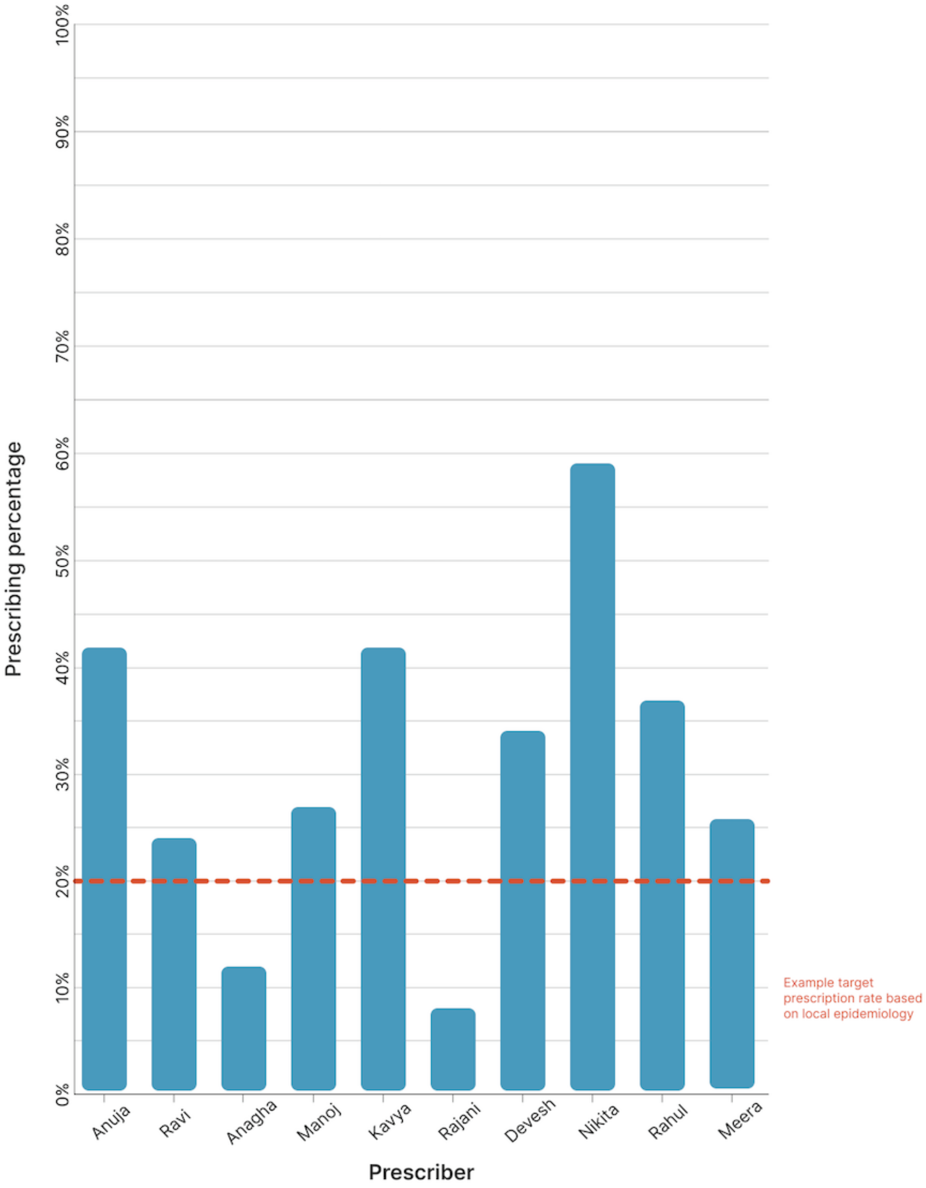
Physician Education

Sustained provider education, streamlined clinical decision support systems, standardized treatment algorithms, and behavioral nudges are core components of antimicrobial stewardship programs [21]. The Kumbh Mela, which runs for several weeks, provides unique access to large numbers of physicians visiting from around the state. The education of prescribers involved in antibiotic use is a core pillar of antibiotic stewardship [21]. Training them in simple, user-centered clinical treatment protocols with standardized thresholds for antibiotic initiation in the weeks leading up to the Mela may help streamline clinical decision-making and promote the rational use of antimicrobial agents. Resources permitting, following a cohort of these physicians over time as they return to their hometowns may offer insights into the sustainability of these interventions and their generalizability to antimicrobial stewardship programs in India at large. Of note, the current policy of providing a 3-day supply of antibiotics for ailments diagnosed at the Mela should be reexamined, as it risks both overtreating patients without bacterial infection and undertreating those that do. In the case of pneumonia, for example, society guidelines in India and abroad recommend at least 5 days of antibiotic therapy [22,23].

Technological Assistance

The Mela provides an excellent opportunity to build on prior successes with lightweight, mobile digital health tools to promote sustainable antimicrobial use. An electronic medical record, which has already been successfully deployed at previous Melas in the form of a tablet-based app, will allow for performance scores and quality metrics related to rates of antibiotic prescribing. Anonymized benchmarking of providers against their peers may serve as an effective behavioral nudge to rationalize prescribing practice (Figure 3). The use of electronic prescribing, as envisioned by the Ayushman Bharat Digital Health Mission (ABDHM), and clinical decision support tools embedded within the electronic medical record could further support good prescribing practice [24]. These initiatives can inform broader efforts to regulate antibiotic sales in India, which is characterized by high levels of over-the-counter dispensing of antimicrobial agents [25].

Figure 3. Proposed end-of-the-day feedback to prescribers at the Kumbh Mela demonstrating their prescribing practices compared to their peers and to a validated benchmark of expected antibiotic utilization, according to local epidemiology.



Patient Education

A particular challenge in the Indian context is that of patient expectations that drive demand for antibiotics, and risk aversion among providers, leading to inappropriate initiation of empiric antibiotic therapy [26]. The Mela provides an ideal opportunity for initiating a conversation about antibiotic use with the broader public, as has been done with other themes at previous Melas (eg, the 2015 Kumbh Mela was designated as an eco-friendly “Harit Kumbh” [Green Kumbh], with a focus on sustainability [27]). Awareness campaigns at congregation spaces and in the vicinity of clinics are likely to be noticed.

Limitations and Ethical Considerations

Digital tools may facilitate responsible antibiotic prescribing through education, monitoring, and surveillance, but they must be combined with other core elements of antimicrobial stewardship, including a strong commitment to stewardship by local and national health authorities, as well as ongoing reporting and feedback [21]. The proposed digital tools will work best as part of the local and national commitment to antimicrobial stewardship.

Ethical issues surrounding digital tools and data privacy in health care records must be considered [28]. In India, there has been a robust debate regarding the governance framework for health data exchange in light of the ABDHM [28]. Similar digital tools to the ones proposed above have been deployed at previous iterations of the Mela without undermining individual privacy [8-10]. Peer-based benchmarking of antibiotic prescription rates will require anonymized and aggregated prescribing data to allow for supportive and nonpunitive learning opportunities [28]. Such anonymization is at least theoretically possible within the digital architecture proposed by the ABDHM.

Future Directions

Despite the scale of the threat posed by AMR, stewardship activities have been difficult to implement not just in India but across the world. India continues to be the world's largest consumer of antibiotics, but there is limited evidence to guide antimicrobial stewardship activities in the Indian context [3]. The government's support for task shifting and focus on technology and public angst over the state of health care in the aftermath of the COVID-19 pandemic provide an impetus for change. Given the tens of millions that are expected to converge at the Kumbh Mela in Allahabad in 2025, bold measures to combat antibiotic overuse and AMR are urgently required. The Mela offers a unique opportunity to deploy and validate digital interventions in an environment with multiple stressors and can act as a catalyst for sustained, larger-scale interventions across the Global South. Lessons learned from the Kumbh Mela may also inform antimicrobial stewardship activities at other organized mass gatherings, such as the Hajj in Saudi Arabia, which face similar challenges related to antibiotic overuse [29,30].

Acknowledgments

We thank the medical student volunteers from the Rotaract Club of the Caduceus, Mumbai, for their role in data collection at the 2013 and 2015 Kumbh Melas, as well as Sujata Saunik, principal secretary of the Public Health Department at the 2015 Mela. Joseph Nallen provided graphical assistance.

The 2013 and 2015 Kumbh Mela studies were supported by grants from the Lakshmi Mittal and Family South Asia Institute at Harvard University and UNICEF (United Nations Children's Fund), Mumbai.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Variation in antibiotic prescription practices across sites for patients presenting with self-reported fever at the 2015 Kumbh Mela.

[PNG File , 123 KB - publichealth_v9i1e45121_app1.png]

Multimedia Appendix 2

Variation in antibiotic prescription practices across sites for patients presenting with gastrointestinal symptoms at the 2015 Kumbh Mela.

[PNG File , 122 KB - publichealth_v9i1e45121_app2.png]

References

1. Global action plan on antimicrobial resistance. World Health Organization. 2016 Jan 01. URL: https://www.who.int/publications/i/item/9789241509763 [accessed 2023-02-24]
2. Holmes AH, Moore LSP, Sundsfjord A, Steinbakk M, Regmi S, Karkey A, et al. Understanding the mechanisms and drivers of antimicrobial resistance. Lancet 2016 Jan 09;387(10014):176-187 [FREE Full text] [doi: 10.1016/S0140-6736(15)00473-0] [Medline: 26603922]
3. Manesh A, Varghese GM. Rising antimicrobial resistance: an evolving epidemic in a pandemic. Lancet Microbe 2021 Sep;2(9):e419-e420. [doi: 10.1016/s2666-5247(21)00173-7]
4. Tsay SV, Bartoces M, Gouin K, Kabbani S, Hicks LA. Antibiotic Antibiotic prescriptions associated with COVID-19 outpatient visits among Medicare beneficiaries, April 2020 to April 2021. JAMA 2022 May 24;327(20):2018-2019 [FREE Full text] [doi: 10.1001/jama.2022.5471] [Medline: 35394497]
5. Russell CD, Fairfield CJ, Drake TM, Turtle L, Seaton RA, Wootton DG, ISARIC4C investigators. Co-infections, secondary infections, and antimicrobial use in patients hospitalised with COVID-19 during the first pandemic wave from the ISARIC WHO CCP-UK study: a multicentre, prospective cohort study. Lancet Microbe 2021 Aug;2(8):e354-e365 [FREE Full text] [doi: 10.1016/S2666-5247(21)00090-2] [Medline: 34100002]
6. Patel J, Sridhar D. The pandemic legacy of antimicrobial resistance in the USA. Lancet Microbe 2022 Oct;3(10):e726-e727. [doi: 10.1016/s2666-5247(22)00227-0]
7. Memish ZA, Steffen R, White P, Dar O, Azhar EI, Sharma A, et al. Mass gatherings medicine: public health issues arising from mass gathering religious and sporting events. Lancet 2019 May;393(10185):2073-2084. [doi: 10.1016/s0140-6736(19)30501-x]
8. Balsari S, Greenough PG, Kazi D, Heerboth A, Dwivedi S, Leaning J. Public health aspects of the world's largest mass gathering: the 2013 Kumbh Mela in Allahabad, India. J Public Health Policy 2016 Dec 8;37(4):411-427. [doi: 10.1057/s41271-016-0034-z] [Medline: 28202925]

9. Kazi DS, Greenough PG, Madhok R, Heerboth A, Shaikh A, Leaning J, et al. Using mobile technology to optimize disease surveillance and healthcare delivery at mass gatherings: a case study from India's Kumbh Mela. J Public Health (Oxf) 2017 Sep 01;39(3):616-624 [FREE Full text] [doi: 10.1093/pubmed/fdw091] [Medline: 27694349]
10. Shaikh A, Bhatia A, Yadav G, Hora S, Won C, Shankar M, et al. Applying human-centered design principles to digital syndromic surveillance at a mass gathering in India: viewpoint. J Med Internet Res 2022 Jan 10;24(1):e27952 [FREE Full text] [doi: 10.2196/27952] [Medline: 35006088]
11. Mehrotra R, Vera F, Eck D, Leaning J, Mehta D, Mehta D. Kumbh Mela: Mapping the Ephemeral Megacity. Berlin, Germany: Hatje Cantz; 2015.
12. Sulis G, Daniels B, Kwan A, Gandra S, Daftary A, Das J, et al. Antibiotic overuse in the primary health care setting: a secondary data analysis of standardised patient studies from India, China and Kenya. BMJ Glob Health 2020 Sep 15;5(9):e003393 [FREE Full text] [doi: 10.1136/bmjgh-2020-003393] [Medline: 32938614]
13. Chandy SJ, Mathai E, Thomas K, Faruqui AA, Holloway K, Lundborg CS. Antibiotic use and resistance: perceptions and ethical challenges among doctors, pharmacists and the public in Vellore, South India. IJME 2013 Jan(1). [doi: 10.20529/ijme.2013.005]
14. Milucky J, Pondo T, Gregory CJ, Iuliano D, Chaves SS, McCracken J, Adult TAC Working Group. The epidemiology and estimated etiology of pathogens detected from the upper respiratory tract of adults with severe acute respiratory infections in multiple countries, 2014-2015. PLoS One 2020;15(10):e0240309 [FREE Full text] [doi: 10.1371/journal.pone.0240309] [Medline: 33075098]
15. Sudarsan M, Sitikantha B, Aparajita D. Audit and quality assessment of prescriptions in an urban health centre of Kolkata. IJMEDPH 2016 Sep 01;6(3):136-139. [doi: 10.5530/ijmedph.2016.3.8]
16. Kasabi GS, Subramanian T, Allam RR, Grace CA, Reddy S, Murhekar MV. Prescription practices & use of essential medicines in the primary health care system, Shimoga district, Karnataka, India. Indian J Med Res 2015 Aug;142(2):216-219 [FREE Full text] [doi: 10.4103/0971-5916.164261] [Medline: 26354220]
17. Saurabh M, Biswas N, Yadav A, Singhai A, Saurabh A. Study of prescribing habits and assessment of rational use of drugs among doctors of primary health care facilities. Asian J Pharm Clin Res 2011:102-105 [FREE Full text]
18. Trikha S, Dalpath SK, Sharma M, Shafiq N. Antibiotic prescribing patterns and knowledge of antibiotic resistance amongst the doctors working at public health facilities of a state in northern India: a cross sectional study. J Family Med Prim Care 2020 Aug;9(8):3937-3943 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_367_20] [Medline: 33110790]
19. Sharma M, Eriksson B, Marrone G, Dhaneria S, Lundborg CS. Antibiotic prescribing in two private sector hospitals; one teaching and one non-teaching: a cross-sectional study in Ujjain, India. BMC Infect Dis 2012 Jul 12;12:155 [FREE Full text] [doi: 10.1186/1471-2334-12-155] [Medline: 22788873]
20. Linder JA, Doctor JN, Friedberg MW, Reyes Nieva H, Birks C, Meeker D, et al. Time of day and the decision to prescribe antibiotics. JAMA Intern Med 2014 Dec 01;174(12):2029-2031 [FREE Full text] [doi: 10.1001/jamainternmed.2014.5225] [Medline: 25286067]
21. Antimicrobial stewardship programmes in health-care facilities in low- and middle-income countries. A WHO practical toolkit. World Health Organization. 2019. URL: https://apps.who.int/iris/handle/10665/329404 [accessed 2023-02-24]
22. Metlay JP, Waterer GW, Long AC, Anzueto A, Brozek J, Crothers K, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Respir Crit Care Med 2019 Oct 01;200(7):e45-e67 [FREE Full text] [doi: 10.1164/rccm.201908-1581ST] [Medline: 31573350]
23. Gupta D, Agarwal R, Aggarwal AN, Singh N, Mishra N, Khilnani GC, Pneumonia Guidelines Working Group. Guidelines for diagnosis and management of community- and hospital-acquired pneumonia in adults: joint ICS/NCCP(I) recommendations. Lung India 2012 Jul;29(Suppl 2):S27-S62 [FREE Full text] [doi: 10.4103/0970-2113.99248] [Medline: 23019384]
24. Ayushman Bharat Digital Mission. National Health Authority. 2022. URL: https://abdm.gov.in/ [accessed 2023-02-24]
25. Koya SF, Ganesh S, Selvaraj S, Wirtz VJ, Galea S, Rockers PC. Consumption of systemic antibiotics in India in 2019. Lancet Reg Health - Southeast Asia 2022 Sep;4:100025. [doi: 10.1016/j.lansea.2022.100025]
26. Kotwani A, Wattal C, Katewa S, Joshi PC, Holloway K. Factors influencing primary care physicians to prescribe antibiotics in Delhi India. Fam Pract 2010 Dec;27(6):684-690. [doi: 10.1093/fampra/cmq059] [Medline: 20660529]
27. Saunik S, Gedam P, Nadkarni VC, Sengupta A. Green Kumbh. Deconstructing the Kumbh Mela—Nashik-Trimbakeshwar 2015: A Public Health Perspective. 2022. URL: https://drive.google.com/file/d/1mvXEFrHPdo5-1ZLahqUVYNueETjJMhOt/view [accessed 2023-03-02]
28. Saksena N, Matthan R, Bhan A, Balsari S. Rebooting consent in the digital age: a governance framework for health data exchange. BMJ Glob Health 2021 Jul 22;6(Suppl 5):e005057 [FREE Full text] [doi: 10.1136/bmjgh-2021-005057] [Medline: 34301754]
29. Bokhary H, Barasheed O, Abd El Ghany M, Khatami A, Hill-Cawthorne GA, Rashid H, Hajj Research Team. Pilot survey of knowledge, attitudes and perceptions of Hajj deployed health care workers on antibiotics and antibiotic prescriptions for upper respiratory tract infections: results from two Hajj seasons. Trop Med Infect Dis 2020 Jan 29;5(1) [FREE Full text] [doi: 10.3390/tropicalmed5010018] [Medline: 32013238]

30. Bokhary H, Research Team H, Barasheed O, Othman H, Saha B, Rashid H, et al. Evaluation of the rate, pattern and appropriateness of antibiotic prescription in a cohort of pilgrims suffering from upper respiratory tract infection during the 2018 Hajj season. Access Microbiol 2022;4(4) [FREE Full text] [doi: 10.1099/acmi.0.000338] [Medline: 35812707]

Abbreviations

ABDHM: Ayushman Bharat Digital Health Mission
AMR: antimicrobial resistance
UNICEF: United Nations Children's Fund

Edited by A Mavragani, T Sanchez; submitted 16.12.22; peer-reviewed by G Mboowa, K Arteaga-Livias; comments to author 23.01.23; revised version received 07.02.23; accepted 14.02.23; published 08.03.23.

Please cite as:
Chan IHY, Gofine M, Arora S, Shaikh A, Balsari S
Technology, Training, and Task Shifting at the World's Largest Mass Gathering in 2025: An Opportunity for Antibiotic Stewardship in India
JMIR Public Health Surveill 2023;9:e45121
URL: https://publichealth.jmir.org/2023/1/e45121
doi:10.2196/45121
PMID:36805363

Viewpoint

Monitoring and Combating Waterpipe Tobacco Smoking Through Surveillance and Taxation

Randa K Saad[1], MD; Adna Maiteh[1], BSc; Rima Nakkash[2,3], DrPH; Ramzi G Salloum[4], PhD; Ali Chalak[5], PhD; Niveen M E Abu-Rmeileh[6], PhD; Yousef Khader[7], SCD; Mohannad Al Nsour[1], PhD

[1]Global Health Development | Eastern Mediterranean Public Health Network, Amman, Jordan
[2]Department of Health Promotion and Community Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon
[3]Department of Global and Community Health, College of Health and Human Services, George Mason University, Fairfax, VA, United States
[4]Department of Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Florida, FL, United States
[5]Department of Agriculture, Faculty of Agricultural and Food Sciences, American University of Beirut, Beirut, Lebanon
[6]Institute of Community and Public Health, Birzeit University, Birzeit, Occupied Palestinian Territory
[7]Public Health, Faculty of Medicine, Jordan University of Science and Technology, Irbid, Jordan

Corresponding Author:
Randa K Saad, MD
Global Health Development | Eastern Mediterranean Public Health Network
Abdallah Ben Abbas St 42
Amman, 11196
Jordan
Phone: 962 781665060
Email: randaksaad@gmail.com

Abstract

Waterpipe tobacco smoking (WTS) is a traditional tobacco use method that originated in the Eastern Mediterranean Region (EMR) and has resurged in recent decades. WTS rates in the EMR are the highest worldwide, especially among youth, exceeding cigarette-smoking rates in select jurisdictions. Despite its documented harm, the growing prevalence of WTS has been met with a poor regulatory response globally. At the epicenter of the WTS epidemic, countries in the EMR are in urgent need of effective tobacco control strategies that consider the particularities of WTS. A roundtable session, titled “Monitoring and Combating WTS Through Taxation and the Global Tobacco Surveillance System (GTSS),” was held as part of the 7th Eastern Mediterranean Public Health Network’s regional conference. The session provided an overview of evidence to date about WTS policy control, the taxation of WTS, volumetric choice experiments for tobacco control research, and monitoring WTS patterns and control policies among adults and youth through the GTSS. The session highlighted the need to update the regulation of WTS in the current global tobacco control policy frameworks and the need for developing tailored, evidence-based, and WTS-specific regulations to complement current tobacco control policy frameworks. Raising taxes to increase the price of tobacco products is the single most effective tobacco control measure, and these taxes can fund expanded government health programs. The effectiveness of taxation can be measured via volumetric choice experiments, which allow for the estimation of a complete set of own-price and cross-price elasticities that are instrumental for fiscal policy simulations. Finally, the surveillance of WTS (for example, through the GTSS) is critical to informing policy and decision makers. The Global Youth Tobacco Survey (GYTS) and Global Adult Tobacco Survey (GATS) are 2 GTSS products that provide nationally representative data among students aged 13-15 years and persons ≥15 years, respectively.

(JMIR Public Health Surveill 2023;9:e40177) doi:10.2196/40177

KEYWORDS

waterpipe tobacco; smoking; tobacco taxation; Global Tobacco Surveillance System; GTSS; Eastern Mediterranean Region; tobacco; public health; surveillance; taxation

Introduction

The Eastern Mediterranean Region (EMR) has the lowest average prices of tobacco products among all World Health Organization (WHO) regions [1] and is the only region in which smoking prevalence has been projected to increase by 2025 [2]. However, there is little research on the economics of tobacco, and the majority of demand elasticities estimates for tobacco products are for cigarettes [3]. In a recent report on the performance of cigarette tax policies, the EMR as a whole had the second lowest score [4].

Waterpipe tobacco smoking (WTS) is a traditional tobacco use method that originated in the EMR [5], and its use continues to increase globally, especially among youth and young adults [6-8], exceeding cigarette-smoking rates [9]. Despite its documented harm, including its link with lung cancer [9], respiratory illnesses, periodontal diseases, and low birth weight [10], the growing prevalence of WTS has been met with a poor regulatory response globally [9]. At the epicenter of the WTS epidemic, countries in the EMR need effective tobacco control strategies that consider the specific particularities of WTS.

Article 6 of the WHO Framework Convention on Tobacco Control (FCTC) supports taxation, and where appropriate, pricing policies, to curb the use of tobacco products [11]. Additionally, the MPOWER policy package of effective tobacco control policies stresses that raising the price of tobacco products through taxation is the most effective way to reduce smoking [12]. However, evidence to date has been largely limited to cigarettes, and thus, evidence to support fiscal measures to curb WTS is scarce. This evidence gap was acknowledged in a 2015 WHO advisory note and more recently in a National Cancer Institute monograph on the economics of tobacco control [13]. A limited number of published studies exist on the economics of tobacco control in the EMR. For the most part, these studies have generally focused on examining cigarette smoking. Given that WTS prevalence exceeds that of cigarette smoking among certain populations in the region, there is a need for models that more accurately capture WTS particularities. For example, the WHO recommends considering scenarios where taxation is on the individual user (at the consumer level) and possibly taxing the waterpipe parts and accessories [14].

In addition to the importance of raising taxes on tobacco products, the MPOWER policy package stresses the importance of monitoring tobacco use, as it is a cross-cutting activity and involves periodically collecting nationally representative population-based youth and adult data on key indicators of tobacco use [12]. To that end, and to assist countries in establishing tobacco surveillance and control programs, the WHO, US Centers for Disease Control and Prevention (CDC), and Canadian Public Health Association initiated the Global Tobacco Surveillance System (GTSS) [15]. The GTSS aims to build the capacity of countries to plan, implement, monitor, and evaluate tobacco control interventions, providing central targets and indicators for the WHO FCTC and the WHO MPOWER technical package, including exposure to mass media campaigns and cost-related indicators [15]. Currently, the GTSS collects data through the Global Youth Tobacco Survey (GYTS), the Global School Personnel Survey, the Tobacco Questions for Surveys (TQS), and the Tobacco Questions for Surveys of Youth (TQS-Youth) [15]. With the Global Adult Tobacco Survey (GATS) being the nationally representative household survey among individuals 15 years of age or older [16], GTSS data can be used to inform policy makers about the tobacco problem in their country, leading to new policy decisions on tobacco prevention and control.

This viewpoint aims to report on the learnings of a roundtable session that focused on the various evidence-based WTS policies; describe the methods used to assess the price elasticity of demand for WTS and the cross-price elasticity between cigarettes and WTS in Lebanon, Jordan, Palestine, and Egypt; and explain how the GTSS can provide important data for the EMR countries on WTS use and related WHO MPOWER policy measures, including exposure to mass media campaigns and cost-related indicators.

Roundtable Description

A roundtable session was held on November 17, 2021, as part of the 7th Eastern Mediterranean Public Health Network's regional conference to discuss the most recent evidence on WTS in the EMR, including monitoring and controlling demand through taxation. The roundtable included oral presentations and an interactive discussion of questions and comments from participants. The following topics were presented to address the roundtable objectives: WTS policy control, taxation of WTS, volumetric choice experiments (VCEs) for tobacco control research, and monitoring WTS and control policies among adults and youth through the GTSS.

Waterpipe Tobacco Control Policy: Evidence to Date

Mainly driven by youth uptake, WTS has increased over the past 2 decades in many countries around the world, including in the EMR [17]. Factors that may have contributed to the increased use and prevalence of WTS include the introduction of flavoring with reduced harshness, the misperception of it being healthier than other tobacco products, its affordability, the quick lit charcoal, the allure of WTS on social media, the social acceptance of waterpipe cafes, the lack of waterpipe-specific policies and regulations, and the immigration patterns from countries with a high prevalence of use to low-prevalence countries [17]. Some unique features of WTS include the fact that it is a stationary and time-consuming tobacco use method, with the usual duration of each session lasting 30 minutes or more [18-20], often used in dedicated cafes and restaurants; flavor is the main product focus; and sharing it with others is a dominant product feature. WTS involves several accessories, including charcoal, the hose, and the device itself, and is self-assembled.

A systematic review of interventions for WTS concluded that there is a lack of evidence of the effectiveness for most waterpipe control interventions, with few showing promising results, and recommended that higher-quality interventions are needed [21].

The FCTC is an international health treaty adopted by the World Health Assembly in 2003. The WHO FCTC is largely based on evidence collected from cigarette policy effectiveness [22]. Although this framework may help in the regulation of other tobacco products, the global rise in their prevalence, particularly WTS [6], has introduced a number of issues [23]. The expansion of producers, exporters, and manufacturers of waterpipe tobacco and accessories poses a significant challenge to WTS legislation.

This is in stark contrast to the tobacco industry, which is dominated by a few global corporations [24]. WTS also has distinct characteristics from cigarette smoking, such as the use of charcoal briquettes, a large apparatus and hose available in a variety of sizes, and a diverse range of tobacco flavors and packaging modes, all of which may necessitate an alternative regulatory mechanism to supplement the current WHO FCTC–recommended framework [25].

Existing regulations must be adapted to effectively address WTS, in addition to the need for waterpipe-specific legislation to meet these unique difficulties. For example, current tobacco control policy frameworks in many jurisdictions do not define whether health warning labels should be applied to waterpipe apparatuses and other accessories, and if they do, there is no guidance on how to use them in practice [26]. This is essential because persons who smoke at waterpipe-serving establishments are rarely exposed to waterpipe tobacco packages; instead, they are only offered a waterpipe apparatus [26]. Despite this, the health warning labels on present waterpipe tobacco packets include a number of deceptive characteristics [27], including incorrect ingredients labeling [28], and do not follow the WHO FCTC's guidelines [29]. The introduction of dangerous waterpipe tobacco alternatives (called "herbal" or "nontobacco"), which may be excluded from tobacco control legislation but are sold and consumed alongside waterpipe tobacco and may be indistinguishable from it, further complicates matters [30].

There are regulatory gaps in place to regulate WTS across the world, and these gaps could jeopardize existing tobacco control strategies. There is a chance to assess current and prospective policy alternatives for reducing WTS and use that information to establish a new tobacco control policy framework. Tobacco control researchers play a vital role in providing policy makers with the necessary evidence to propose effective legislation. The WHO FCTC secretariat should consider establishing a scientific working group to explore WTS regulatory concerns and propose a complementary framework to the WHO FCTC. Countries that bear the brunt of the load of tobacco-related diseases and have experience in regulating WTS should be involved [31].

When it comes to reducing the uptake of smoking among young people, tobacco product costs are an important aspect to consider because decreasing affordability is the most effective strategy [32]. To minimize tobacco consumption, the WHO FCTC proposes that taxation policies consider the price elasticity of demand, and that all tobacco products be taxed equally to avoid unexpected consequences such as product substitution [2].

Taxation of WTS

There are direct costs associated with tobacco use, including health care cost to treat tobacco-related diseases, and non–health care costs, such as transportation to a clinic and the time of family members providing care [33]. There are also indirect costs of tobacco use, associated with the reduction in potential economic productivity due to morbidity and premature mortality [33,34]. Additionally, there are 3 types of societal costs: external, internal, and indirect. External societal costs are the costs that tobacco users impose on others such as through secondhand smoking. Internal societal costs result from the information failures in the market that can be thought of as external costs. Indirect societal costs are costs paid by tobacco users and their families, incurred as a result of tobacco (eg, out-of-pocket costs for health care to treat diseases caused by smoking) [33]. The latest estimates of the total worldwide economic cost of tobacco smoking are above US $1.4 trillion (2012), which is 1.8% of the world's annual GDP [34]. Direct health care costs alone are estimated at US $443 billion dollars, whereas indirect costs, incurred due to the loss of productivity as a result of morbidity and mortality, are estimated at US $357 billion and US $657 billion, respectively [34].

Taxation is an evidence-based tool for tobacco control. First, taxation promotes public health. Depending on the tax structure, taxation can have a higher impact on demand among vulnerable individuals such as the young population, whose demand is more elastic than the demand of older adults. This is also true for less educated and low-income individuals who are disproportionately affected by the burden of tobacco smoking. Second, taxation is an efficient revenue generation strategy for governments as smoking is relatively inelastic, where consumers have a lower behavioral response when the price is raised compared to products with higher elasticity [35]. Third, from a societal perspective, taxation corrects for the external costs of tobacco borne by members of society other than the smoker.

There is a substantial body of research over many decades and from many countries, showing that significantly increasing the excise tax and price of tobacco products to reduce their affordability is the single most consistently effective tool for reducing tobacco use [13]. There are a variety of tobacco taxes used: some are based on sales taxes or value-added taxes; some are in the form of customs duties on tobacco leaf and product imports or exports; and some are implicit taxes when governments monopolize production and distribution. There are also excise taxes, which are of the most interest given their specificity to the tobacco products. Excise taxes are of 2 types. The first type is the specific tax, which is based on the weight or volume of the tobacco used. The benefit of this type of tax is that it reduces price gaps, deters tax avoidance, is easy to administer, and stabilizes tax revenue. The second type of excise tax is ad valorem, which is based on the commercial value of the product. The benefits of the ad valorem tax are that it adjusts with inflation and is progressive, imposing higher taxes on higher-priced products (ie, taxes higher-income individuals). However, it leads to greater price gaps and incentivizes tax avoidance.

Early efforts to examine the price elasticity of demand for noncigarette products have provided sufficient evidence of the effectiveness of tax and price increase in decreasing demand. A systematic review concluded that there is positive substitution between cigarette and noncigarette products, suggesting that tax and price increase should be simultaneous and comparable across all tobacco products [36].

The Eastern Mediterranean Consortium on the Economics of WTS is an ongoing project examining the economics of WTS. This is a collaboration between the American University of

Beirut, Jordan University of Science and Technology, Birzeit University, and Ain Shams University and is funded by Cancer Research UK and the International Development Research Centre (IDRC). In its most recent study, the Consortium examined the price elasticities of waterpipe tobacco products and cigarette products [3]. The WHO FCTC recommends that the price elasticity of demand be taken into account in taxation policy to reduce tobacco use [2]. The price elasticity of a product measures how sensitive the demand for it responds to a change in its own price (own-price elasticity) or in the price of other related products (cross-price elasticity) [37]. Overall, the price elasticity observed for WTS and cigarette smoking varied by product and across countries, with the demand for premium cigarettes being price elastic (range −1.0 to −1.2) and that for premium waterpipe tobacco being highly elastic in Lebanon (−1.9), moderately elastic in Jordan (−0.6), and inelastic in Palestine (0.2) [3].

VCEs for Tobacco Control Research

The methodology that was used to guide the research of the Eastern Mediterranean Consortium on the Economics of WTS includes VCEs for tobacco control research. All economic problems are rooted in the issue of choices that need to be made as resources are scare (opportunity cost). Choices reveal information about preferences of consumers, which is instrumental for understanding the costs and benefits attendant to health policies and interventions. This informs how responsive consumers would be reacting to changes in prices. The most obvious avenue to observe choices, and thus preferences, is to observe market choices, using the revealed preference methods. The methods involve collecting market data and inferring measures of demand and preference intensities mainly through the vehicle of elasticities. The advantage of these methods is that they have a high face validity as they rely on real-life data [38]. However, in some situations, it is limited by the lack of availability of products, that is, it cannot measure choice preference for products that do not exist on the market. Therefore, economists sometimes rely on observing hypothetical choices by means of stated preference (SP) methods. These methods present consumers with hypothetical products and collect what they state they will buy at experimentally varied prices [39]. However, these methods lack consequentiality, often resulting in respondents inflating their willingness to purchase and pay for the products under study (hypothetical bias) [39].

The most common SP approach used in marketing and economic research is discrete choice experiments (DCEs), which are an attribute-based, hypothetical, and survey-based tool for measuring preference and value [40,41]. DCEs incorporate the notion of opportunity cost and choice. DCEs can be disadvantageous in the demand analysis of certain products such as WTS. Therefore, VCEs were developed as an alternative to the SP methods. DCEs force respondents to choose one and only one option, which is often restrictive compared to real market situations. VCEs, in contrast, accommodate simultaneous choices of multiple options and accommodate “quantitative” or “volumetric” choices (ie, the number of units selected). It does this by simultaneously offering multiple distinct volumetric choices across competing products. Data from Jordan, Palestine, and Lebanon show the elasticity of demand for waterpipe tobacco products and the existence of substitution between tobacco products, which should be considered when formulating tobacco taxation strategies [42]. Thus, VCEs allow for the estimation of a complete set of own-price and cross-price elasticities that are instrumental for the purpose of fiscal policy simulations [3]. This technique is useful in forecasting how consumers will respond to policies that increase the prices of products such as tobacco and by how much such policies may increase revenue to governments. This is illustrated in a study where the elasticity estimates from the VCE were fed to a simulation model that was adapted to forecast the effects of taxation policies. Specifically, the model used country-specific and market share–specific price and consumption data from the WHO and United Nations (UN) Comtrade, in addition to the VCE’s elasticity estimates obtained from nationally representative surveys conducted in Lebanon, Jordan, and Palestine. It then forecasted the effects of specific excise taxes, which met a 35.9% tax burden on waterpipe tobacco in Lebanon and Jordan (in line with the global average) and doubled government revenues from excise duties in Palestine [43].

Monitoring WTS and Control Policies Among Adults and Youth Through the GTSS

As COVID-19 spread in the EMR, there was a renewed focus on curbing smoking as 17 countries/sites temporarily banned WTS in indoor and outdoor public places during the pandemic [44]. Additionally, the CDC, in collaboration with the Eastern Mediterranean Public Health Network and Vital Strategies, launched the “United Against Tobacco and COVID” campaign in Jordan, Egypt, Iraq, and Palestine. This project aimed to warn on the harms of tobacco, in accordance with the MPOWER policy package recommendations [12]. The project intended to raise awareness of the harms of smoking, especially during the COVID-19 pandemic, through media coverage. This project used evidence-based information and data to develop country-tailored and culturally appropriate messages to disseminate, educate, and promote smoking cessation to reduce COVID-19 morbidity and mortality.

This campaign, similar to all antitobacco interventions, programs, and policies, relied on accurate measures of tobacco use to determine the target audience as well as the target messages that need to be addressed within each county. Indeed, to plan any tobacco control strategy or intervention effectively, counties need accurate data to inform implementation and increase the likelihood of success.

One approach of obtaining such data is through globally standardized surveys such as the GTSS, which systematically monitors youth and adult tobacco use as well as key tobacco control indicators (the WHO FCTC and MPOWER) [15]. It also monitors and enhances the capacity to design, implement, and evaluate tobacco control policies. The GYTS, a nationally representative school-based survey of students aged 13-15 years, is the largest public health surveillance system ever developed and maintained [22]. The survey is self-administered in the classroom setting using a paper-and-pencil questionnaire [22]. It contains a core questionnaire with optional modules for countries to add additional questions. Between 1999-2020, the GYTS was implemented in 188 countries/sites [45]. This standard, systematic, and consistent process generates comparable data within and across countries. The GYTS *Shisha (or Waterpipe)* module is an optional module that can be used [46].

These collected surveillance data are helpful in designing, guiding, monitoring, and evaluating tobacco control interventions, programs, and policies at the country level. One key goal for such data is to highlight the extent of the tobacco epidemic in countries, including WTS, and thus generate interest and support for tobacco control policy, and those for WTS, among key policy makers and stakeholders. For example, if policy makers are considering a smoke-free ban, data will serve to educate stakeholders about the extent of exposure to secondhand smoke within the country and thus support the ban.

Additionally, using surveillance data to understand the tobacco market within a country is one of the strategies to ensure that tobacco taxation, including taxation on WTS, is effective. Surveys enable the identification of changes in consumption patterns so that existing policies can be adjusted accordingly. Authorities might ask for more information on certain aspects such as consumption patterns and attitudes toward WTS among youth. This information is crucial to assess the tobacco industry’s marketing power and can inform discussions on elasticities.

In the EMR, GYTS WTS data are available from 20 countries/sites from 2005 to 2016. Table 1 provides the prevalence rates of current WTS among students aged 13-15 years in these 20 countries/sites. The overall prevalence rates ranged from 34.8% in Lebanon (2011) to 3.8% in Morocco (2016). The prevalence rates were significantly higher among boys than girls in all but 2 countries/sites.

The GATS is a nationally representative household-based face-to-face interview survey of individuals aged 15 years or older, using a global standard protocol and electronic data collection methodology [16]. It similarly includes a core questionnaire with optional questions and options for countries to add additional questions [16]. Since 2008, the GATS has been implemented in 36 countries representing more than 70% of the world’s adult population [22]. It contains WTS core questions including *Current WTS* and *Frequency of WTS*, as well as an optional WTS module [16].

In the EMR, GATS WTS data are available from 4 countries from 2009 to 2019. Table 2 provides the prevalence rates of current WTS among individuals aged 15 years and older in these 4 countries. The prevalence rates were as follows: 6.7% in Saudi Arabia in 2019 (9.7% males and 2.3% females); 3.4% in Qatar in 2013 (4.9% males and 1.6% females); 3.3% in Egypt in 2009 (6.2% males and 0.3% females); and 3.0% in Pakistan in 2014 (4.7% males and 1.1% females).

Unfortunately, monitoring rates through standardized national population-based surveillance surveys vary by country income group, with better coverage achieved in higher-income countries. As per the WHO’s most recent global report on trends in the prevalence of tobacco use from 2000-2025, the lowest population coverage by such surveys is in the EMR, with only 88% of the population living in 76% of the EMR having sufficient available data to calculate tobacco use trends [2].

To that end, the TQS is beneficial. The TQS includes a subset of 22 key questions from the GATS for adults [47], and the TQS-Youth contains a subset of 21 key questions from the GYTS for youth [48]. These questions are integrated into ongoing surveys (health or otherwise, such as the WHO Stepwise Approach to Surveillance surveys) for sustainable monitoring and global consistency. Between 2008 and 2021, 100 countries completed surveys with TQS integration.

However, in addition to the data that are collected by the standardized surveillance surveys, policy makers may require data about the composition of tobacco and WTS products, including flavors and their production and marketing costs by brand. This additional information can provide a better understanding of the market power, market share, and behavior of the tobacco industry within a country. It can also inform discussions on elasticities, including how consumers choose specific brands or flavors.

Table 1. Prevalence of current waterpipe smoking among students 13-15 years old, by sex—Global Youth Tobacco Survey, Eastern Mediterranean Region countries/sites, 2005-2016.

Country/site (year)	Current waterpipe smoking			
	Overall (%)	Boys (%)	Girls (%)	*P* value (boys vs girls)
Lebanon (2011)	34.8	39.3	31.0	.03
Jordan (2014)	26.7	34.5	18.4	.001
Syria (2010)	19.2	24.7	13.8	.04
Bahrain (2015)	18.3	26.1	10.2	.001
Palestine—West Bank (2016)	17.9	23.4	12.4	.001
Qatar (2013)	17.1	24.6	10.1	.001
Palestine—Gaza Strip (2013)	16.2	22.0	11.1	.01
Yemen (2014)	14.1	17.1	9.3	.001
Kuwait (2016)	13.9	19.5	9.0	.001
United Arab Emirates (2005)	12.4	16.3	7.6	.001
Djibouti (2013)	11.6	10.2	12.8	.20
Iran (2016)	11.2	14.3	8.5	.03
Iraq (2014)	10.6	15.1	5.3	.01
Sudan (2014)	9.9	13.0	5.4	.01
Saudi Arabia (2010)	9.5	13.3	6.1	.04
Oman (2016)	7.4	11.6	3.7	.03
Egypt (2014)	5.9	7.2	4.1	.09
Tunisia (2010)	5.8	10.1	2.1	.001
Libya (2010)	4.3	6.0	2.6	.001
Morocco (2016)	3.8	5.0	2.6	.04

Table 2. Prevalence of current waterpipe smoking among individuals aged 15+ years, by sex—Global Adult Tobacco Survey, Eastern Mediterranean Region countries, 2009-2019.

Country (year)	Current waterpipe smoking			
	Overall (%)	Male (%)	Female (%)	*P* value[a] (males vs females)
Saudi Arabia (2019)	6.7	9.7	2.3	.001
Qatar (2013)	3.4	4.9	1.6	.001
Egypt (2009)	3.3	6.2	0.3	.001
Pakistan (2014)	3.0	4.7	1.1	.001

[a] Significance at $P<.05$

Conclusion

This paper provides an overview of WTS regulatory and surveillance issues in the EMR and discusses examples of success in taxation at the level of the consumer as well as the methods used to assess the feasibility and effectiveness of taxation on demand. The immense need to collect data on WTS using the GTSS was also stressed, to enable policy makers to make better informed decisions.

Recommendations and Key Areas for Improvement

- There is a need to update the regulation of WTS in the current global tobacco control policy frameworks.
- There is a need to develop tailored and evaluated WTS-specific regulatory frameworks to complement current tobacco control policy frameworks.
- Raising taxes to increase the price of tobacco products is the single most effective tobacco control measure, and this also applies to WTS.
- Increased taxes from WTS can fund expanded government health programs.

- VCEs allow for the estimation of a complete set of own-price elasticities and cross-price elasticities for WTS products, which is instrumental for the purpose of fiscal policy simulations.
- Gaining political buy-in is key to adopting key tax reforms relating to tobacco and WTS, and to this end, data from standardized validated surveillance systems are crucial.
- Data collection on WTS through the GTSS is critical to inform and sensitize policy makers and decision makers from the EMR about the public health and socioeconomic burdens caused by WTS, as well as the growing use among females and youth.
- The GYTS and GATS can provide nationally representative data among students aged 13-15 years and ≥15 years, respectively, on WTS, and countries can include an optional waterpipe module.
- As a cost-effective measure, and to ensure some data on WTS are being collected, it is suggested to include the TQS on WTS within other surveys that countries are already implementing.
- Data on WTS that are generated from the GYTS and GATS are the most valuable when disseminated in a way that will gain the attention of key stakeholders, especially policy makers and the media, as they are the most important audience if these data are to cause a ripple effect. Effective data dissemination is crucial, and so is engaging partners to help translate important information to key decision makers, who have the authority to change tobacco control policy.

Acknowledgments

The authors would like to thank Mr Jeremy Morton (Global Tobacco Control Branch, Office on Smoking and Health, US Centers for Disease Control and Prevention [CDC]) for his valuable insight and input during the roundtable discussion.

Data Availability

Data sharing not applicable to this paper as no data sets were generated or analyzed during the current study.

Conflicts of Interest

None declared.

References

1. El-Awa F, Bettcher D, Al-Lawati JA, Alebshehy R, Gouda H, Fraser CP. The status of tobacco control in the Eastern Mediterranean Region: progress in the implementation of the MPOWER measures. East Mediterr Health J 2020 Jan 30;26(1):102-109 [FREE Full text] [doi: 10.26719/2020.26.1.102] [Medline: 32043552]
2. WHO global report on trends in prevalence of tobacco use 2000-2025, fourth edition. World Health Organization. 2021 Nov 16. URL: https://www.who.int/publications/i/item/9789240039322 [accessed 2023-03-16]
3. Chalak A, Nakkash R, Abu-Rmeileh NME, Khader YS, Jawad M, Mostafa A, et al. Own-price and cross-price elasticities of demand for cigarettes and waterpipe tobacco in three Eastern Mediterranean countries: a volumetric choice experiment. Tob Control 2023 Jan 30;32(1):86-92 [FREE Full text] [doi: 10.1136/tobaccocontrol-2021-056616] [Medline: 34193608]
4. Chaloupka FJ, Drope J, Vulovic V, Stoklosa M, Mirza M, Rodriguez-Iglesias G, et al. Cigarette tax scorecard. Tobacconomics. 2020 Dec. URL: https://tobacconomics.org/research/cigarette-tax-scorecard/ [accessed 2023-03-16]
5. Maziak W, Taleb ZB, Bahelah R, Islam F, Jaber R, Auf R, et al. The global epidemiology of waterpipe smoking. Tob Control 2015 Mar 08;24 Suppl 1(Suppl 1):i3-i12 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051903] [Medline: 25298368]
6. Akl EA, Gunukula SK, Aleem S, Obeid R, Jaoude PA, Honeine R, et al. The prevalence of waterpipe tobacco smoking among the general and specific populations: a systematic review. BMC Public Health 2011 Apr 19;11(1):244 [FREE Full text] [doi: 10.1186/1471-2458-11-244] [Medline: 21504559]
7. Tucktuck M, Ghandour R, Abu-Rmeileh NME. Waterpipe and cigarette tobacco smoking among Palestinian university students: a cross-sectional study. BMC Public Health 2017 Jul 10;18(1):1 [FREE Full text] [doi: 10.1186/s12889-017-4524-0] [Medline: 28693529]
8. Maziak W, Ben Taleb Z, Jawad M, Afifi R, Nakkash R, Akl EA, Expert Panel on Waterpipe Assessment in Epidemiological Studies. Consensus statement on assessment of waterpipe smoking in epidemiological studies. Tob Control 2017 May;26(3):338-343 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-052958] [Medline: 27165995]
9. Maziak W. The waterpipe: an emerging global risk for cancer. Cancer Epidemiol 2013 Feb;37(1):1-4 [FREE Full text] [doi: 10.1016/j.canep.2012.10.013] [Medline: 23196170]
10. Waziry R, Jawad M, Ballout RA, Al Akel M, Akl EA. The effects of waterpipe tobacco smoking on health outcomes: an updated systematic review and meta-analysis. Int J Epidemiol 2017 Feb 01;46(1):32-43 [FREE Full text] [doi: 10.1093/ije/dyw021] [Medline: 27075769]

11. van Walbeek C, Filby S. Analysis of Article 6 (tax and price measures to reduce the demand for tobacco products) of the WHO Framework Convention on Tobacco Control. Tob Control 2019 Jun 25;28(Suppl 2):s97-s103 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054462] [Medline: 30045973]
12. MPOWER: a policy package to reverse the tobacco epidemic. World Health Organization. 2008. URL: https://www.afro.who.int/publications/mpower-policy-package-reverse-tobacco-epidemic [accessed 2023-03-16]
13. National Cancer Institute, World Health Organization. NCI tobacco control monograph series 21 - the economics of tobacco and tobacco control. World Health Organization. 2017 Jan 10. URL: https://www.who.int/publications/m/item/nci-tobacco-control-monograph-series-21-the-economics-of-tobacco-and-tobacco-control [accessed 2023-03-16]
14. World Health Organization, WHO Study Group on Tobacco Product Regulation. Advisory note: waterpipe tobacco smoking: health effects, research needs and recommended actions by regulators, 2nd ed. World Health Organization. 2015. URL: https://apps.who.int/iris/handle/10665/161991 [accessed 2023-03-16]
15. Centers for Disease Control and Prevention, Office on Smoking and Health, National Center for Chronic Disease Prevention and Health Promotion. Smoking and tobacco use: global tobacco control. Centers for Disease Control and Prevention. 2021 Nov 30. URL: https://www.cdc.gov/tobacco/global/gtss/index.htm [accessed 2022-01-02]
16. Global Adult Tobacco Survey Collaborative Group. Global Adult Tobacco Survey (GATS): fact sheet templates. Centers for Disease Control and Prevention. 2020 Sep. URL: https://nccd.cdc.gov/GTSSDataSurveyResources/Ancillary/DownloadAttachment.aspx?ID=53 [accessed 2023-03-16]
17. Jawad M, Charide R, Waziry R, Darzi A, Ballout RA, Akl EA. The prevalence and trends of waterpipe tobacco smoking: a systematic review. PLoS One 2018 Feb 9;13(2):e0192191 [FREE Full text] [doi: 10.1371/journal.pone.0192191] [Medline: 29425207]
18. Smith-Simone S, Maziak W, Ward KD, Eissenberg T. Waterpipe tobacco smoking: knowledge, attitudes, beliefs, and behavior in two U.S. samples. Nicotine Tob Res 2008 Feb;10(2):393-398 [FREE Full text] [doi: 10.1080/14622200701825023] [Medline: 18236304]
19. Maziak W, Rastam S, Ibrahim I, Ward KD, Shihadeh A, Eissenberg T. CO exposure, puff topography, and subjective effects in waterpipe tobacco smokers. Nicotine Tob Res 2009 Jul;11(7):806-811. [doi: 10.1093/ntr/ntp066] [Medline: 19420278]
20. Eissenberg T, Shihadeh A. Waterpipe tobacco and cigarette smoking: direct comparison of toxicant exposure. Am J Prev Med 2009 Dec;37(6):518-523 [FREE Full text] [doi: 10.1016/j.amepre.2009.07.014] [Medline: 19944918]
21. Jawad M, Jawad S, Waziry RK, Ballout RA, Akl EA. Interventions for waterpipe tobacco smoking prevention and cessation: a systematic review. Sci Rep 2016 May 11;6(1):25872 [FREE Full text] [doi: 10.1038/srep25872] [Medline: 27167891]
22. Asma SM, Mackay J, Song SY, Zhao L, Morton J, Palipudi KM, et al. The GATS Atlas: Global Adult Tobacco Survey. World Health Organization. 2015. URL: https://www.cdc.gov/tobacco/global/gtss/tobacco_atlas/index.htm [accessed 2023-03-16]
23. Maziak W, Nakkash R, Bahelah R, Husseini A, Fanous N, Eissenberg T. Tobacco in the Arab world: old and new epidemics amidst policy paralysis. Health Policy Plan 2014 Sep 19;29(6):784-794 [FREE Full text] [doi: 10.1093/heapol/czt055] [Medline: 23958628]
24. Saloojee Y, Dagli E. Tobacco industry tactics for resisting public policy on health. Bull World Health Organ 2000;78(7):902-910 [FREE Full text] [Medline: 10994263]
25. WHO Framework Convention on Tobacco Control. World Health Organization. 2003 May 25. URL: https://fctc.who.int/who-fctc/overview [accessed 2023-03-16]
26. Jawad M. Legislation enforcement of the waterpipe tobacco industry: a qualitative analysis of the London experience. Nicotine Tob Res 2014 Jul;16(7):1000-1008. [doi: 10.1093/ntr/ntu022] [Medline: 24642591]
27. Khalil J, Heath RL, Nakkash RT, Afifi RA. The tobacco health nexus? health messages in narghile advertisements. Tob Control 2009 Oct 24;18(5):420-421. [doi: 10.1136/tc.2009.030148] [Medline: 19779065]
28. Vansickel AR, Shihadeh A, Eissenberg T. Waterpipe tobacco products: nicotine labelling versus nicotine delivery. Tob Control 2012 May;21(3):377-379 [FREE Full text] [doi: 10.1136/tc.2010.042416] [Medline: 21636612]
29. Nakkash R, Khalil J. Health warning labelling practices on narghile (shisha, hookah) waterpipe tobacco products and related accessories. Tob Control 2010 Jun 25;19(3):235-239 [FREE Full text] [doi: 10.1136/tc.2009.031773] [Medline: 20501497]
30. Hammal F, Chappell A, Wild TC, Kindzierski W, Shihadeh A, Vanderhoek A, et al. 'Herbal' but potentially hazardous: an analysis of the constituents and smoke emissions of tobacco-free waterpipe products and the air quality in the cafés where they are served. Tob Control 2015 May;24(3):290-297. [doi: 10.1136/tobaccocontrol-2013-051169] [Medline: 24128428]
31. Jawad M, El Kadi L, Mugharbil S, Nakkash R. Waterpipe tobacco smoking legislation and policy enactment: a global analysis. Tob Control 2015 Mar;24 Suppl 1(Suppl 1):i60-i65 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051911] [Medline: 25550418]
32. Jha P, Chaloupka FJ, Corrao M, Jacob B. Reducing the burden of smoking world-wide: effectiveness of interventions and their coverage. Drug Alcohol Rev 2006 Nov;25(6):597-609. [doi: 10.1080/09595230600944511] [Medline: 17132576]
33. Economics of tobacco toolkit: assessment of the economic costs of smoking. World Health Organization. 2011. URL: https://apps.who.int/iris/handle/10665/44596 [accessed 2023-03-16]

34. Goodchild M, Nargis N, Tursan d'Espaignet E. Global economic cost of smoking-attributable diseases. Tob Control 2018 Jan;27(1):58-64 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053305] [Medline: 28138063]
35. Vulovic V. Economic costs of tobacco use: a Tobacconomics policy brief. Tobacconomics. 2019 Apr. URL: https://tobacconomics.org/files/research/523/UIC_Economic-Costs-of-Tobacco-Use-Policy-Brief_v1.3.pdf [accessed 2023-03-16]
36. Jawad M, Lee JT, Glantz S, Millett C. Price elasticity of demand of non-cigarette tobacco products: a systematic review and meta-analysis. Tob Control 2018 Nov 23;27(6):689-695 [FREE Full text] [doi: 10.1136/tobaccocontrol-2017-054056] [Medline: 29363611]
37. Huang J, Gwarnicki C, Xu X, Caraballo RS, Wada R, Chaloupka FJ. A comprehensive examination of own- and cross-price elasticities of tobacco and nicotine replacement products in the U.S. Prev Med 2018 Dec;117:107-114 [FREE Full text] [doi: 10.1016/j.ypmed.2018.04.024] [Medline: 29684418]
38. OECD. Revealed preference methods. In: Cost-Benefit Analysis and the Environment: Further Developments and Policy Use. Paris, France: OCED Publishing; Jun 25, 2018.
39. Zawojska E, Czajkowski M. Re-examining empirical evidence on stated preferences: importance of incentive compatibility. Journal of Environmental Economics and Policy 2017 May 05;6(4):374-403. [doi: 10.1080/21606544.2017.1322537]
40. Salloum RG, Abbyad CW, Kohler RE, Kratka AK, Oh L, Wood KA. Assessing preferences for a university-based smoking cessation program in Lebanon: a discrete choice experiment. Nicotine Tob Res 2015 May;17(5):580-585 [FREE Full text] [doi: 10.1093/ntr/ntu188] [Medline: 25239962]
41. Salloum RG, Maziak W, Hammond D, Nakkash R, Islam F, Cheng X, et al. Eliciting preferences for waterpipe tobacco smoking using a discrete choice experiment: implications for product regulation. BMJ Open 2015 Sep 09;5(9):e009497 [FREE Full text] [doi: 10.1136/bmjopen-2015-009497] [Medline: 26353876]
42. Awawda S, Chalak A, Khader Y, Mostafa A, Abla R, Nakkash R, et al. Gender differences in the price elasticity of demand for waterpipe and cigarette smoking in Lebanon, Jordan and Palestine: a volumetric choice experiment. BMJ Open 2022 Jul 18;12(7):e058495 [FREE Full text] [doi: 10.1136/bmjopen-2021-058495] [Medline: 35851023]
43. Jawad M, Awawda S, Abla R, Chalak A, Khader YS, Nakkash RT, et al. Impact of waterpipe tobacco taxation on consumption, government revenue and premature deaths averted in Jordan, Lebanon and Palestine: a simulation study. Tob Control 2022 Dec 13:tobaccocontrol-2022-057284 [FREE Full text] [doi: 10.1136/tc-2022-057284] [Medline: 36601792]
44. WHO report on the global tobacco epidemic 2021: addressing new and emerging products. World Health Organization. 2021 Jul 27. URL: https://www.who.int/publications/i/item/9789240032095 [accessed 2023-03-16]
45. Pan American Health Organization, Regional Office for the Americas of the World Health Organization. Global Youth Tobacco Survey (GYTS). Pan American Health Organization. URL: https://www.paho.org/en/gyts [accessed 2022-01-03]
46. Global Youth Tobacco Survey Collaborative Group. Global Youth Tobacco Survey (GYTS): core questionnaire with optional questions, version 1.2. Pan American Health Organization. 2014 Nov. URL: https://www.paho.org/hq/dmdocuments/2017/1-GYTS-CoreQuestionnairewithOptionalQuestions-v1.2-Nov2014.pdf [accessed 2023-03-16]
47. Global Adult Tobacco Survey Collaborative Group. Tobacco questions for surveys: a subset of key questions from the Global Adult Tobacco Survey GATS , 2nd Edition. World Health Organization. 2011. URL: https://www.who.int/publications/i/item/9789241500951 [accessed 2023-03-16]
48. World Health Organization, Centers for Disease Control and Prevention. Tobacco questions for surveys of youth (TQS-Youth) : a subset of key questions from the Global Youth Tobacco Survey (GYTS) . World Health Organization. 2019. URL: https://apps.who.int/iris/handle/10665/339158 [accessed 2023-03-16]

Abbreviations

CDC: Centers for Disease Control and Prevention
DCE: discrete choice experiment
EMR: Eastern Mediterranean Region
FCTC: Framework Convention on Tobacco Control
GATS: Global Adult Tobacco Survey
GTSS: Global Tobacco Surveillance System
GYTS: Global Youth Tobacco Survey
IDRC: International Development Research Centre
SP: stated preference
TQS: Tobacco Questions for Surveys
TQS-Youth: Tobacco Questions for Surveys of Youth
UN: United Nations
VCE: volumetric choice experiment
WHO: World Health Organization
WTS: waterpipe tobacco smoking

Edited by A Mavragani, T Sanchez; submitted 09.06.22; peer-reviewed by T Tabuchi, T Wu, N Baharom; comments to author 24.01.23; revised version received 31.01.23; accepted 14.02.23; published 23.03.23.

Please cite as:
Saad RK, Maiteh A, Nakkash R, Salloum RG, Chalak A, Abu-Rmeileh NME, Khader Y, Al Nsour M
Monitoring and Combating Waterpipe Tobacco Smoking Through Surveillance and Taxation
JMIR Public Health Surveill 2023;9:e40177
URL: https://publichealth.jmir.org/2023/1/e40177
doi:10.2196/40177
PMID:36951907

Viewpoint

Convergence in Mobility Data Sets From Apple, Google, and Meta

Gustavo Sganzerla Martinez[1,2*], PhD; David J Kelvin[1,2*], PhD

[1]Department of Microbiology and Immunology, Dalhousie University, Halifax, NS, Canada

[2]Department of Pediatrics, Izaak Walton Killam Health Center, Canadian Center for Vaccinology, Halifax, NS, Canada

[*]all authors contributed equally

Corresponding Author:
David J Kelvin, PhD
Department of Microbiology and Immunology
Dalhousie University
5850 College street
Halifax, NS, B3H4H7
Canada
Phone: 1 647 529 3556
Email: david.kelvin@dal.ca

Abstract

Background: The higher movement of people was one of the variables that contributed to the spread of the infectious agent SARS-CoV-2 during the COVID-19 pandemic. Governments worldwide responded to the virus by implementing measures that would restrict people's movements, and consequently, the spread of the disease. During the onset of the pandemic, the technology companies Apple, Google, and Meta used their infrastructure to anonymously gather mobility reports from their users.

Objective: This study aims to compare mobility data reports collected by Apple, Google, and Meta (formerly Facebook) during the COVID-19 pandemic and a major winter storm in Texas in 2021. We aim to explore the hypothesis that different people exhibit similar mobility trends during dramatic events and to emphasize the importance of this type of data for public health measures. The study also aims to promote evidence for companies to continue releasing mobility trends data, given that all 3 companies have discontinued these services.

Methods: In this study, we collected mobility data spanning from 2020 to 2022 from 3 major tech companies: Apple, Google, and Meta. Our analysis focused on 58 countries that are common to all 3 databases, enabling us to conduct a comprehensive global-scale analysis. By using the winter storm that occurred in Texas in 20201 as a benchmark, we were able to assess the robustness of the mobility data obtained from the 3 companies and ensure the integrity of our findings.

Results: Our study revealed convergence in the mobility trends observed across different companies during the onset of significant disasters, such as the first year of the COVID-19 pandemic and the winter storm that impacted Texas in 2021. Specifically, we observed strong positive correlations (r=0.96) in the mobility data collected from different tech companies during the first year of the pandemic. Furthermore, our analysis of mobility data during the 2021 winter storm in Texas showed a similar convergence of trends. Additionally, we found that periods of stay-at-home orders were reflected in the data, with record-low mobility and record-high stay-at-home figures.

Conclusions: Our findings provide valuable insights into the ways in which major disruptive events can impact patterns of human mobility; moreover, the convergence of data across distinct methodologies highlights the potential value of leveraging mobility data from multiple sources for informing public health decision-making. Therefore, we conclude that the use of mobility data is an asset for health authorities to consider during natural disasters, as we determined that the data sets from 3 companies yielded convergent mobility patterns. Comparatively, data obtained from a single source would be limited, and therefore, more difficult to interpret, requiring careful analysis.

(JMIR Public Health Surveill 2023;9:e44286) doi:10.2196/44286

KEYWORDS

Google; Apple; Meta; COVID-19 mobility; COVID-19; mobility; data set; data; pattern; pandemic; mobile; operating system; system; validation; tool; asset

Introduction

The COVID-19 pandemic has affected people worldwide and presented major public health challenges for authorities. Early in 2020, shortly after the emergence of SARS-CoV-2, many countries imposed restrictive measures to limit the movement of populations, and consequently, limit the spread of SARS-CoV-2 infections and the COVID-19 disease [1]. To aid in evaluating the effect of restrictive measures, the tech industry rapidly responded by offering tools to quantify the movement of people using anonymized mobility data. In fact, higher movement patterns have been directly associated with COVID-19 transmission [2-4], rendering this measurement a powerful tool for public health authorities to understand the dynamics of populations when facing a disaster.

With the onset of COVID-19, tech companies, including Apple, Google, and Meta, provided data sets throughout the pandemic, with each company using different methodologies for obtaining mobility data. Such data can be used by scientists and health authorities to understand transmission patterns, identify hotspots, and potentially effectively control outbreaks. The importance of this data is seen as it provides a real-time snapshot of population dynamics. For instance, it was suggested that the early introduction of the B.1.1.529 (ie, Omicron) SARS-CoV-2 variant in Mexico was facilitated by higher movement observed in the end-of-year celebrations [5]. Moreover, in Canada, a so-called successful pandemic management was attributed to the fact that the movement patterns of people were highly reduced during the initial waves of SARS-CoV-2 [6]. Higher transmission of SARS-CoV-2 in India was attributed to an urban exodus of city dwellers that had to return to their villages after lockdown policies were implemented [7]. Apart from COVID-19, crisis management in other natural disasters can benefit from mobility data. For example, in the aftermath of the cyclone Gaja in south India (November 2018), mobility data were used to visualize where displaced people were relocating, so humanitarian efforts could be directed to those areas [8]. Importantly, the in- and outbound flow of people in particular areas during significant events can be quantified by services that capture mobility trends [9].

Comparing distinct data can provide a more comprehensive understanding of mobility trends, which can be crucial for public health measures. Data are commonly seen as abstractions of the real world. For instance, in applications that benefit from data, there is generally a need to promote validation with external data [10]; this reality is widely explored in clinical settings [11], genomic studies [12], among others. Consequently, a model that can react similarly upon different inputs strengthens the conclusions that are derived from such data.

This study aims to explore the hypothesis that different people exhibit similar mobility trends during dramatic events, such as the outbreak of an infectious disease or climatic disasters, and that these trends can be captured by companies such as Apple, Google, and Meta using distinct collection methodologies. To test this hypothesis, the study will compare mobility data reports collected by these companies during the first year of the COVID-19 pandemic and a major winter storm that hit Texas in 2021. Additionally, the study will observe mobility trends in countries with varying shares of the mobile operating system (MOS) penetration market. The ultimate goal of the study is to emphasize the importance of these types of data for public health measures and to provide evidence supporting the continued release of mobility trends data by companies, given that all 3 companies have discontinued these services.

Methods

Apple Mobility Trends Reports

Apple collected daily data from January 13, 2020, until April 14, 2022, in 63 countries or regions [13]. The company obtained the number of anonymous requests for directions (walking, driving, or using public transportation) made to the Apple Maps application, native to Apple's MOS. The collected information was then compared against itself to reflect a change in the volume of people moving, in some cases including a minimum threshold for directions made per day. As of today, the data are not available any longer within the company website.

Google Community Mobility Reports

Google made use of insights obtained from processing data from Google Maps, the company's location service, to provide a response to health authorities regarding COVID-19 [14]. The data were collected from 15 February 15, 2020, to October 13, 2022. The mobility reports encompass data from 135 countries or regions with major metropolitan areas also included. A baseline was established using the data accumulated during the 5-week pre-pandemic period from January 3, 2020, to February 6, 2020. Whenever a user who chose to enable their location services visited a categorized location on Google Maps on a specific date, an anonymous record was assigned. The data were split into the following 6 categories: grocery and pharmacy, retail and recreation, parks, workplace, residential, and public transportation. Even though the company is no longer releasing updated data, their previously collected data set is publicly available within the company's website.

Movement Data From Meta

Non-public data were obtained from Meta [15] as part of an agreement with Dalhousie University. Meta collected data from 198 countries and territories from February 22, 2020, until May 22, 2022. Major regions targeted by Meta were divided into defined units (600 m^2 600 m^2). Whenever a user had location services enabled and moved between 2 units, a movement record was assigned. Regions with 300 or less qualifying users in a 24-hour period were disregarded. A pre-pandemic baseline was established for each country; however, no details of the individual calculation of the baseline are present in Meta's documentation.

2021 US Winter Storm

In 2021, a major winter storm struck North America. In the state of Texas, which hardly ever records snowfall, weather stations in the Dallas-Forth Worth airport recorded at least 139 hours of freezing or below freezing temperatures, characterizing an uncommon phenomena in the region [16]. During the storm, officially, a death toll of 151 was reported. The estimated cost of the storm in Texas was around US $295 billion dollars. For control purposes, we selected the state of Georgia as a representative state that recorded no snowfall in the period of February-March 2021. We inquired snowfall precipitation to data collected from the National Center for Environmental Information, which enables precipitation records of several weather stations in the Southern United States.

MOS Penetration Data

We obtained data from the service StatCounter Global Stats [17] to select the market penetration of the MOS for Apple's iOS and Google's Android. The service is composed of bots installed in over 1.5 million websites. The bots gather stats on operating systems that are accessing a particular website. In this work, we selected the month of April 2022 to analyze the market share of different MOSs, which is when the first tech company (ie, Apple) stopped their mobility data collection. A total of 58 countries were analyzed, which were the countries present in all 3 mobility data sets. A breakdown of the countries is available in Multimedia Appendix 1.

Data Processing and Statistical Analyses

Each one of the 3 companies provided their data in daily batches, which we treated in 7-day periods. As the data from the 3 companies followed different magnitudes, a scaling process was performed by using the scale function in R (version 4.1.2; R Foundation for Statistical Computing). The function calculates the mean and standard deviation of a vector then removes the mean from each element and divides it by the standard deviation. For comparison purposes, we selected the countries that were shared between the 3 data sets.

The statistical analyses performed in this study included Shapiro-Wilk test for data normality, Pearson correlation, as well as *t* test and Wilcoxon rank test for mean comparison. These analyses were performed using the R programming language (version 4.1.2; R Core Team) under the package rstatisx (version 0.7.0). All figures generated in this study were created using the R package ggplot2 (version 3.3.6).

Ethical Considerations

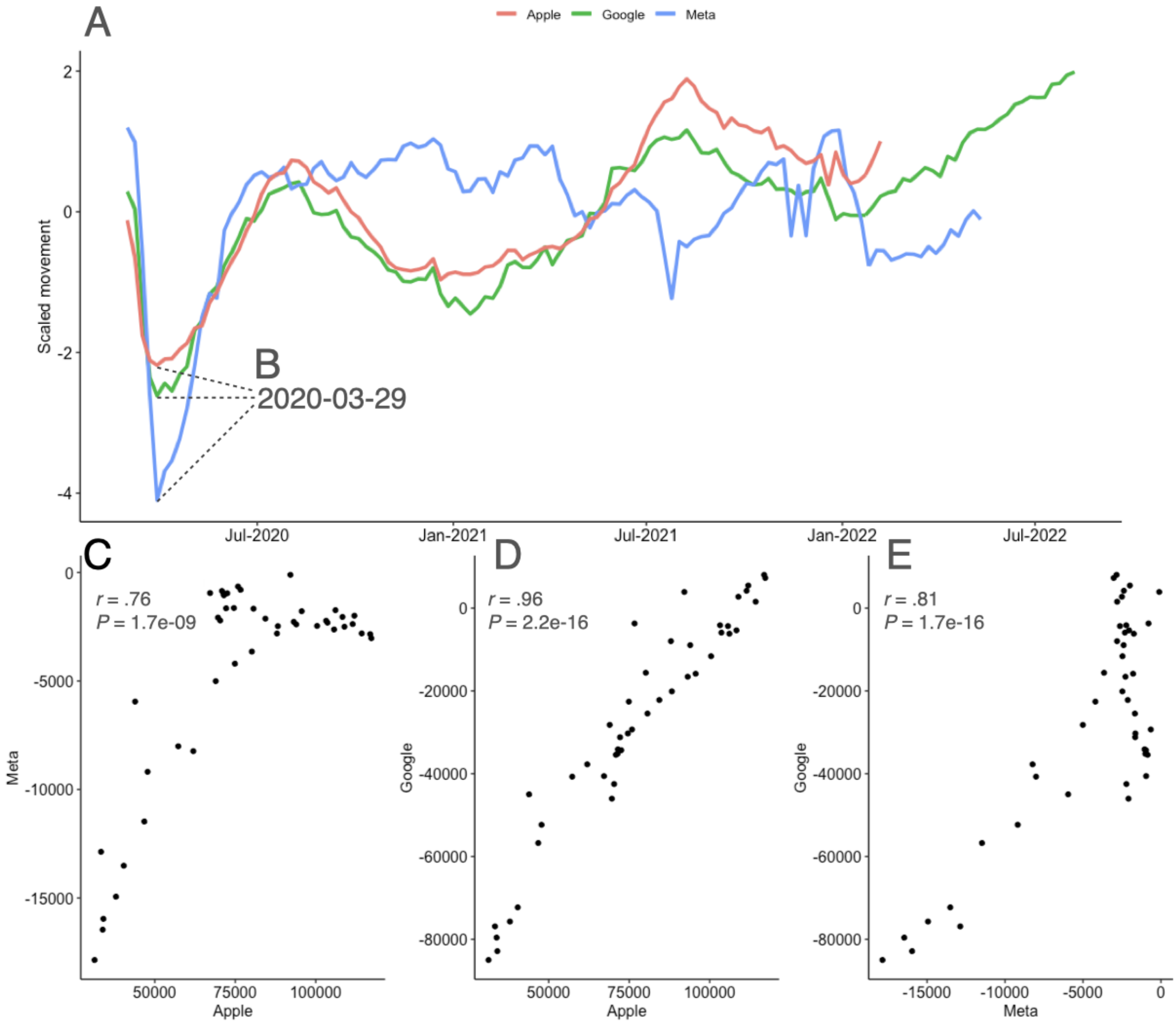
The data collected from Apple, Google, and Meta users were already sampled, anonymized, and depersonalized. Data sources from Apple and Google are publicly obtainable; on the other hand, data from Meta are not publicly available. We, as research partners with Meta, have access only to data formatted to prevent reidentification, which we obtained through a partnership between Meta and author DJK (a licensee at Dalhousie University). Moreover, as users, we commit to using these data only for the purposes outlined in our agreement with Meta and we ensure that any conclusions drawn from these data are based on solely scientific practices.

Results

Mobility Data Well Represent the Initial Onset of the SARS-COV-2 Pandemic

We used the data from the mobility reports of 58 countries collected by Apple, Google, and Meta to perform a comparative analysis to analyze people's movements throughout the first year of the pandemic (ie, 2020; Figure 1). First, we noted the time series of each data collection methodology (ie, company) following the same trend (Figure 1A). Next, we analyzed the point at which each data set presented the lowest mobility score in the first year of the pandemic (Figure 1B). In 2020, the lowest mobility point for the 3 data sets (Figure 1A) occurred on March 29, 2020 (Figure 1B). Next, we performed a correlation analysis on the 3 data sets to depict their convergence or disparity (Figure 1C; Figure 1D, and Figure 1E); we report positive correlations for Apple versus Meta (r=0.76), Apple versus Google (r=0.95), and Google versus Meta (r=0.87), when comparing the data.

Figure 1. The 2020 mobility patterns from Apple, Google, and Meta across 58 countries. We selected the 58 countries that were present in the data from Apple, Google, and Meta. (A) From these countries, the year 2020 was isolated, and a mobility score for each company was assigned and later plotted. (B) The lowest mobility trend indicated in 2020 occurred in the same week for all 3 data sets, on March 29, 2020. (C) Next, we performed a Pearson correlation analysis of the data from Apple versus Meta, (D) Apple versus Google, and (E) Google versus Meta.



The 2021 US Winter Storm Reflects in the Convergence of Mobility Data

We analyzed the movement data from Apple, Google, and Meta to determine the rate of individuals staying at home and not moving during the 2021 winter storm in the state of Texas. Figure 2A shows the period of the storm (depicted by the red-dashed rectangle), indicating low records in movement recorded by the 3 companies. We found the lowest mobility point occurred with no more than a week between each data set, specifically on February 21, 2021 (Apple), February 19, 2021 (Google), and February 14, 2021 (Meta). Additionally, we determined stay-at-home data from Google residential and Meta single tile users (represented by the grey dotted lines in Figure 2); these data were not available for Apple. Both companies noted record highs for stay-at-home data during the same days when record-low mobility was recorded. For control purposes, we selected the dates with the lowest mobility from each data set and compared them with data from the state of Georgia, which presented no snowfall precipitation during the 3 aforementioned days (Multimedia Appendix 2). In Figure 2B, we show that Georgia, unlike Texas, did not record a downfall in data on the dates of February 14, 19, and 21. Moreover, in Table 1, through a 2-sampled statistical test comparing the averages of mobility data from Apple, Google, and Meta over Texas and Georgia, we found low *P* values (Apple and Google: *P*<.001; Google and Meta: *P*=.005).

In Table 1, we present a comparison of the average movement during the months of January and February 2021 in the states of Texas and Georgia. The time series were found to be predominantly normally distributed through Shapiro-Wilk tests (Texas: *P*=.71 for Apple; *P*=.38 for Google; and *P*=.78 for Meta; Georgia: *P*=.71 for Apple; *P*=.04 for Google; and *P*=.51 for Meta). Moreover, the 2-sampled statistical comparison of the averages achieved by a 1-tailed *t* test resulted in statistically significant differences among all comparisons.

Figure 2. Record-low mobility trends recorded during a major winter storm in Texas in 2021. (A) Mobility data in the US state of Texas during the year of 2021. The negative area of the y-axis, indicating movement of data from 3 different tech companies—Apple (green solid line), Google (blue solid line), and Meta (orange solid line)—is considered. The data have been scaled to accommodate different magnitudes. The positive area of the y-axis indicates stay-at-home data, enabled only by Google (light dotted line) and Meta (dark dotted line). We also indicate the period in which a major winter storm in Texas took place (red dashed rectangle), from February 13 to 17, 2021. (B) A control version of part A, for the US state of Georgia. The state of Georgia presented no snow precipitation in the period from February 13 to 17, 2021.

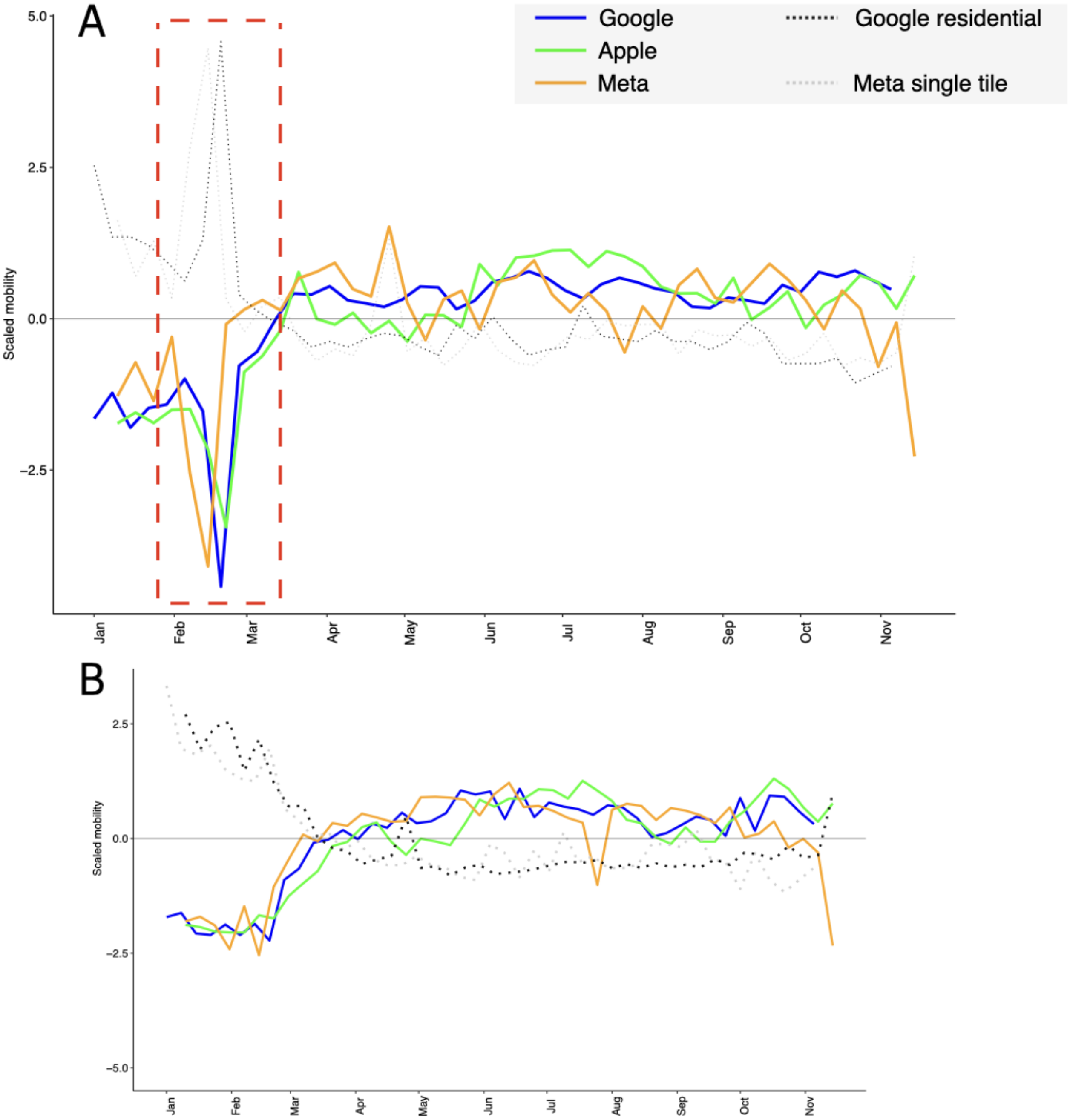


Table 1. Mean comparison of mobility trends during the Texas 2021 winter storm.

Data sets for 2 US states	Mobility average (January to February 2021)	*P* value
Apple		<.001
Texas	215147	
Georgia	135228	
Google		<.001
Texas	–497.4	
Georgia	–250.8	
Meta		.005
Texas	–3.15	
Georgia	–0.81	

Countries With Varied Market Penetration of MOS Have Convergent Mobility Data

To determine if the predominance of an MOS would affect how mobility was recorded, we assessed the market share of mobile users regarding iOS (Apple) and Android (Google) from 54 countries that were common in the mobility data of Apple, Google, and Meta. We found that Android was the predominant platform in India, Argentina, Indonesia, Colombia, and Brazil, with greater than 90% of mobile internet access in April 2022, whereas iOS (Apple's MOS) was the most popular platform in 5 countries—Japan, Denmark, Norway, United States, and Australia (67.6%, 60.3%, 59.3%, 57.8%, and 57.6%, respectively). To assess convergence or disparity between data collection methodologies, we individually correlated the 3 data sets in each country. In Figure 3, we present the time series data from the 5 iOS-predominant countries. The correlation table of these countries' data is found in Table 2. The lowest averaged correlation was found in the United States (0.73, impacted by reduced correlation in the data of Apple and Meta), while the highest averaged correlation was found in Australia (80.66).

Also, in Figure 3, countries with Android predominance are shown and their correlation is presented in Table 2. First, we report an overall convergence in the mobility data collected by Apple and Google (ie, 0.85 average correlation across all the 10 countries considered). Second, we report a higher convergence between the data of Google and Meta (compared to Apple and Meta) in Android-dominated markets (except for Indonesia). This relationship may be attributed to the fact that Facebook is a third-party app that runs on an MOS, which in these countries is Android, thereby boosting the number of devices from which Google collects data. Finally, data from Apple was not as correlated with Meta as Google's data in the iOS-prevalent countries, potentially due to the still higher proportion of Android devices being used in these countries.

Figure 3. Comparison of mobility time series data from Apple, Google, and Meta in distinct countries. We selected the first 5 countries where Android is the most predominant mobile operating system (alphabetically, Argentina, Brazil, Colombia, India, and Indonesia) and the first 5 countries where iOS is the most predominant mobile operating system (alphabetically, Australia, Denmark, Japan, Norway, and the United States) according to data obtained from StatCounter Global Stats [17].

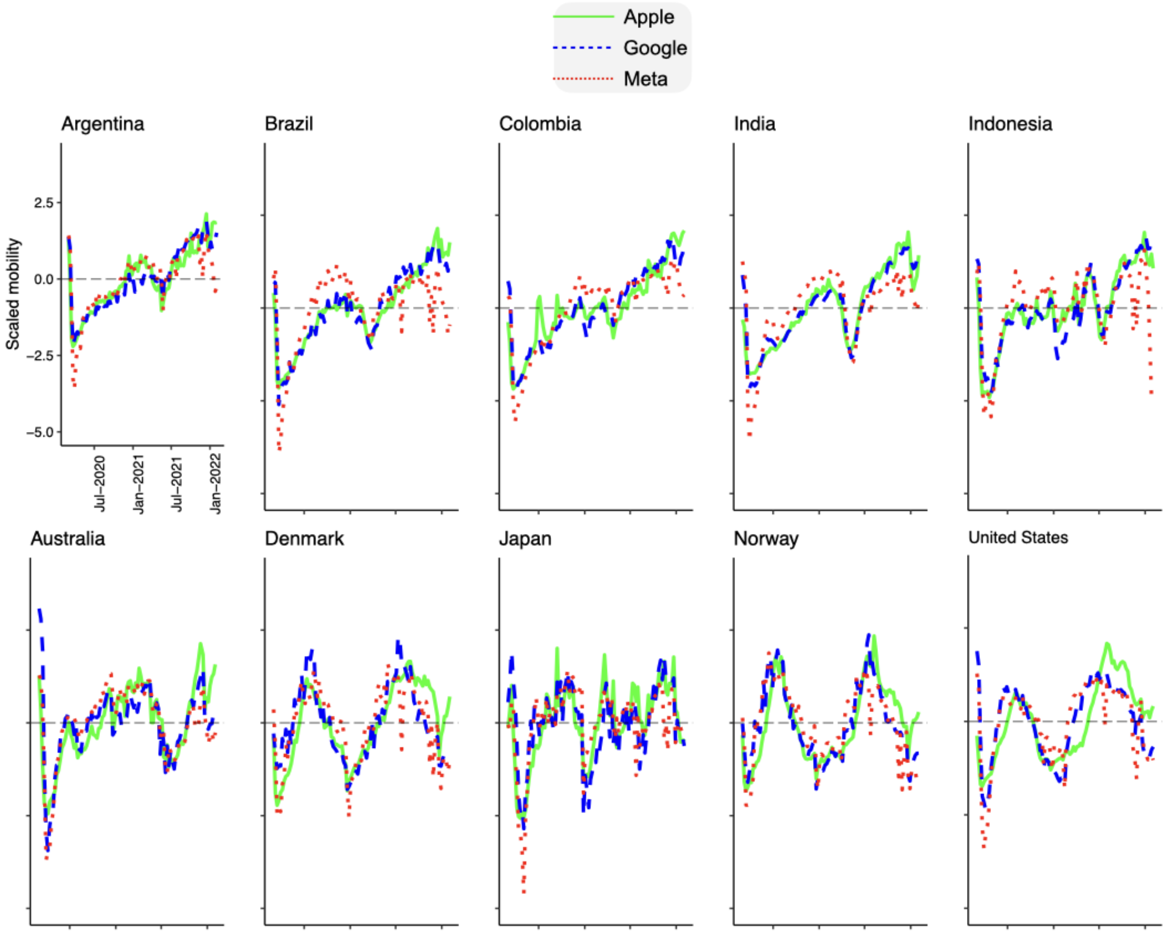


Table 2. Correlation coefficients between data sets in different countries. To compare the time series between each data set, we performed Pearson correlation analyses for the iOS-prevalent countries and Android-prevalent countries.

Country	Comparison			Mobile operating system predominance
	Apple vs Google	Apple vs Meta	Google vs Meta	
Argentina	.91	.77	.84	Android
Brazil	.96	.68	.75	Android
Colombia	.91	.81	.86	Android
India	.95	.77	.84	Android
Indonesia	.93	.69	.64	Android
Australia	.78	.76	.88	iOS
Denmark	.76	.72	.79	iOS
Japan	.79	.76	.75	iOS
Norway	.84	.68	.87	iOS
United States	.71	.59	.89	iOS

Discussion

Principal Findings

In our work, we depicted a convergent behavior in mobility data gathered by distinct data collection methodologies independently of the popularity of MOSs. First, we showed that even over a longer 2-year time span, there is still a similar trend in mobility patterns, as evidenced by the moderate-to-strong correlation analyses in data collection from 3 distinct companies. Moreover, we observed that the mobility data converged more during events that required people to change their movement trends (ie, lockdowns, to avoid the spread of SARS-CoV-2; and a major winter storm that forced people to stay at home), granting this type of measure to be an asset for public health decisions.

We attribute the fact that we observed an alike behavior on data collected by different sources to 2 potential phenomena. First, when people are affected by a catastrophic event, such as the pandemic and the 2021 winter storm in Texas, their mobility and lifestyle will drastically change [10,18]. Considering a large number of people are experiencing the same stressor, the limited ways they develop to cope with stressful events, such as loneliness and social isolation, as reported by Hwang et al [19], will eventually converge. Moreover, it was reported in India and in the United Kingdom that a massive urban exodus took place during the pandemic, forcing many city-working young adults to have to return to their hometowns [20,21], despite being from different backgrounds, their mobility patterns might have converged once they were forced out of their workplaces. In this viewpoint, we argue that this convergence results from people developing ways to cope with their stress and will eventually be reflected in people moving in similar patterns. Next, users of distinct MOSs, whose data are being collected by distinct companies, might not significantly differ in terms of their movement patterns. In fact, behavioral similarities were previously reported among Android and iOS users, adding consistency to our analysis [22].

As a limitation of our study, we acknowledge that we did not have access to general mobility data from periods outside the pandemic, as these data were not being collected. Without these data, it was challenging to establish a solid baseline for comparison during and outside pandemic periods. It would have allowed us to gain a better understanding of the extent to which the observed behavior changes were driven by the pandemic itself or if they represented more general changes in human behavior over time. Furthermore, although smartphones have become increasingly prevalent in recent years, they do not represent the entire population. Therefore, we acknowledge that our sample may not have been a representative of the entire population. Specifically, individuals who do not own smartphones or have limited access to the internet might have different mobility patterns than those who do. Therefore, it is essential to consider the limitations of our sample when interpreting the results of our study. We believe these changes would lead to more robust policy and decision-making in public health.

The Data for Good portal [15], maintained by Meta with infrastructure from the company, provides a valuable data source for public health offices, as it collects data to support and inform efforts to tackle humanitarian and environmental challenges. As such, it is an asset for public health decision makers. For example, the company has been actively releasing their data at times of natural disasters to help inform decision makers, including during the flooding in Samar Island, Eastern Visayas, Philippines; the flooding in Southern California on January 5, 2023; the flooding in Makassar City, South Sulawesi, Indonesia; and the wildfire in Viña del Mar, Valparaíso, Chile, among others. As shown by our results, during these catastrophic events, the behavior of people might converge. Thus, the analysis of people's movement might help answer important questions of public interest to reduce the impact of these events when they happen again.

Conclusions

In our study, we observed that movement data collected under distinct methodologies by different tech companies presented a convergent aspect in depicting catastrophic events in which the movement of people was majorly impacted. The similar behaviors exhibited by different people facing similar stressors add reliability for these types of data, which would be an important asset upon which public health measures could rely. Thus, based on our results, we encourage Apple, Google, and Meta to continue releasing these types of data.

Acknowledgments

The corresponding author, DJK, is the Canada Research Chair in Translational Vaccinology and Inflammation. We are grateful to Dr Nikki Kelvin for her assistance with the editing of this manuscript.

Data Availability

The data sets generated and analyzed during this study are publicly available for Google [23] and Apple [24]. The Meta movement patterns were included in a private data set and obtained as part of a private agreement between Meta and Dalhousie University.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Intersection of countries that are shared between Apple, Google, and Meta mobility data sets.
[PDF File (Adobe PDF File), 94 KB - publichealth_v9i1e44286_app1.pdf]

Multimedia Appendix 2
The 2021 winter storm in southern United States.

[PNG File , 677 KB - publichealth_v9i1e44286_app2.png]

References

1. Wilder-Smith A, Freedman DO. Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. J Travel Med 2020 Mar 13;27(2):taaa020 [FREE Full text] [doi: 10.1093/jtm/taaa020] [Medline: 32052841]
2. Nouvellet P, Bhatia S, Cori A, Ainslie KEC, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission. Nat Commun 2021 Feb 17;12(1):1090 [FREE Full text] [doi: 10.1038/s41467-021-21358-2] [Medline: 33597546]
3. Her PH, Saeed S, Tram KH, Bhatnagar SR. Novel mobility index tracks COVID-19 transmission following stay-at-home orders. Sci Rep 2022 May 10;12(1):7654 [FREE Full text] [doi: 10.1038/s41598-022-10941-2] [Medline: 35538129]
4. Hu T, Wang S, She B, Zhang M, Huang X, Cui Y, et al. Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges. Int J Digit Earth 2021 Jul 14;14(9):1126-1147 [FREE Full text] [doi: 10.1080/17538947.2021.1952324]
5. Castelán-Sánchez HG, Martínez-Castilla LP, Sganzerla-Martínez G, Torres-Flores J, López-Leal G. Genome evolution and early introductions of the SARS-CoV-2 omicron variant in Mexico. Virus Evol 2022;8(2):veac109 [FREE Full text] [doi: 10.1093/ve/veac109] [Medline: 36582501]
6. Sganzerla Martinez G, Hewins B, LeBlanc JJ, Ndishimye P, Toloue Ostadgavahi A, Kelvin DJ. Evaluating the effectiveness of lockdowns and restrictions during SARS-CoV-2 variant waves in the Canadian province of Nova Scotia. Front Public Health 2023 Apr 27;11:1142602 [FREE Full text] [doi: 10.3389/fpubh.2023.1142602] [Medline: 37181684]
7. Pal SC, Saha A, Chowdhuri I, Roy P, Chakrabortty R, Shit M. Threats of unplanned movement of migrant workers for sudden spurt of COVID-19 pandemic in India. Cities 2021 Feb;109:103035. [doi: 10.1016/j.cities.2020.103035]
8. Maas P. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019 Presented at: KDD '19; August 4-8; Anchorage AK p. 3173. [doi: 10.1145/3292500.3340412]
9. Changruenngam S, Bicout DJ, Modchang C. How the individual human mobility spatio-temporally shapes the disease transmission dynamics. Sci Rep 2020 Jul 09;10(1):11325 [FREE Full text] [doi: 10.1038/s41598-020-68230-9] [Medline: 32647225]
10. Rahman R, Ventz S, McDunn J, Louv B, Reyes-Rivera I, Polley MC, et al. Leveraging external data in the design and analysis of clinical trials in neuro-oncology. The Lancet Oncology 2021 Oct;22(10):e456-e465. [doi: 10.1016/s1470-2045(21)00488-5]
11. Martinez G, Garduno A, Mahmud-Al-Rafat A, Ostadgavahi AT, Avery A, de Avila E Silva S, et al. An artificial neural network classification method employing longitudinally monitored immune biomarkers to predict the clinical outcome of critically ill COVID-19 patients. PeerJ 2022;10:e14487 [FREE Full text] [doi: 10.7717/peerj.14487] [Medline: 36530391]
12. Sganzerla Martinez G, Perez-Rueda E, Kumar A, Sarkar S, de Avila E Silva S. Explainable artificial intelligence as a reliable annotator of archaeal promoter regions. Sci Rep 2023 Jan 31;13(1):1763 [FREE Full text] [doi: 10.1038/s41598-023-28571-7] [Medline: 36720898]
13. Apple. Mobility Trends Reports. 2022 Apr 14. URL: https://covid19.apple.com/mobility [accessed 2022-04-14]
14. See how your community moved differently due to COVID-19. 2022 Oct 15. URL: https://www.google.com/covid19/mobility/ [accessed 2022-11-25]
15. Data for Good. Meta. 2022 May 22. URL: https://dataforgood.facebook.com/ [accessed 2022-06-01]
16. Donald J. Winter Storm Uri 2021. The economic impact of the storm. Fiscal Notes. A review of the Texas Economy. URL: https://comptroller.texas.gov/economy/fiscal-notes/2021/oct/winter-storm-impact.php [accessed 2023-05-03]
17. Statcounter global stats. URL: https://gs.statcounter.com/os-market-share/mobile [accessed 2022-09-15]
18. de Palma A, Vosough S, Liao F. An overview of effects of COVID-19 on mobility and lifestyle: 18 months since the outbreak. Transp Res Part A Policy Pract 2022 May;159:372-397 [FREE Full text] [doi: 10.1016/j.tra.2022.03.024] [Medline: 35350704]
19. Hwang T, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 Oct 26;32(10):1217-1220 [FREE Full text] [doi: 10.1017/S1041610220000988] [Medline: 32450943]
20. Pal SC, Saha A, Chowdhuri I, Roy P, Chakrabortty R, Shit M. Threats of unplanned movement of migrant workers for sudden spurt of COVID-19 pandemic in India. Cities 2021 Feb;109:103035. [doi: 10.1016/j.cities.2020.103035]
21. Rowe F, Calafiore A, Arribas-Bel D, Samardzhiev K, Fleischmann M. Urban exodus? Understanding human mobility in Britain during the COVID-19 pandemic using Meta-Facebook data. Popul Space Place 2023 Jan;29(1):e2637 [FREE Full text] [doi: 10.1002/psp.2637] [Medline: 36718419]
22. Götz FM, Stieger S, Reips U. Users of the main smartphone operating systems (iOS, Android) differ only little in personality. PLoS One 2017 May;12(5):e0176921 [FREE Full text] [doi: 10.1371/journal.pone.0176921] [Medline: 28467473]
23. COVID-19 community mobility reports. Google. URL: https://www.google.com/covid19/mobility/ [accessed 2023-06-14]
24. Gustavsganzerla / covid-in-ns. Github. URL: https://github.com/gustavsganzerla/covid-in-ns/blob/main/applemobilitytrends-2022-02-13.csv [accessed 2023-06-14]

Abbreviations

MOS: mobile operating system

Edited by A Mavragani, T Sanchez; submitted 14.11.22; peer-reviewed by C Gissel, M Mileski; comments to author 20.12.22; revised version received 09.02.23; accepted 26.05.23; published 22.06.23.

Please cite as:
Sganzerla Martinez G, Kelvin DJ
Convergence in Mobility Data Sets From Apple, Google, and Meta
JMIR Public Health Surveill 2023;9:e44286
URL: https://publichealth.jmir.org/2023/1/e44286
doi:10.2196/44286
PMID:37347516

Viewpoint

Implementing and Maintaining a SARS-CoV-2 Exposure Notification Application for Mobile Phones: The Finnish Experience

Mika Pihlajamäki[1], MSc; Sara Wickström[1], MA; Kaija Puranen[2], MSc; Otto Helve[1], MD, PhD; Aleksi Yrttiaho[1], MSc; Lotta Siira[1], PhD

[1]Finnish Institute for Health and Welfare, Helsinki, Finland
[2]Gofore Oyj, Tampere, Finland

Corresponding Author:
Lotta Siira, PhD
Finnish Institute for Health and Welfare
PO Box 30
Helsinki, 00271
Finland
Phone: 358 29 524 8755
Email: lotta.siira@thl.fi

Abstract

Exposure notification applications (ENAs) or digital proximity tracing apps were used in several countries during the COVID-19 pandemic. In this viewpoint, we share our experience of implementing and running the Finnish ENA (Koronavilkku), one of the national ENAs with the highest proportion of users during the pandemic. With the aim of strengthening public trust and increasing app uptake, there was a strong prioritization of privacy and data security for the end user throughout the ENA development. This, in turn, limited the use of the app as a tool for health care professionals and deeper insight into its potential effectiveness. The ENA was designed to supplement conventional contact tracing, rather than replace it, and to serve as an early warning system and a trigger for action for the user in case of potential exposure. The predefined target of 40% uptake in the population was achieved within 3 months of the ENA launch. We consider easy-to-understand information produced together with communication experts crucial during the changing pandemic situation. This information educated people about the app as one component in mitigating the pandemic. As the pandemic and its mitigation evolved, the ENA also needed adapting and updating. A few months after its launch, Finland joined European interoperability, which allowed the ENA to share information with ENAs of other countries. We added automatic token issuing to the ENA as of mid-2021. If added earlier and more comprehensively, automatization could have more effectively saved resources in health care services and prevented overburdening contact tracing teams, while also notifying potentially exposed individuals quicker and more reliably. In the spring of 2021, the number of active apps started to gradually decline. Quarantine and testing practices for asymptomatic vaccinated individuals following exposure to the virus were eased and home tests became more common, eventually replacing laboratory testing for much of the population. Taken together, this led to decreased token issuance, which weakened the potential public health usefulness of the app. A self-service option for token issuance would likely have prolonged the lifespan of the app. The ENA was discontinued in mid-2022. Regularly conducted surveys would have helped gain timely knowledge on the use and effectiveness of the app for better responding to the changing needs during the pandemic.

(JMIR Public Health Surveill 2023;9:e46563) doi:10.2196/46563

KEYWORDS

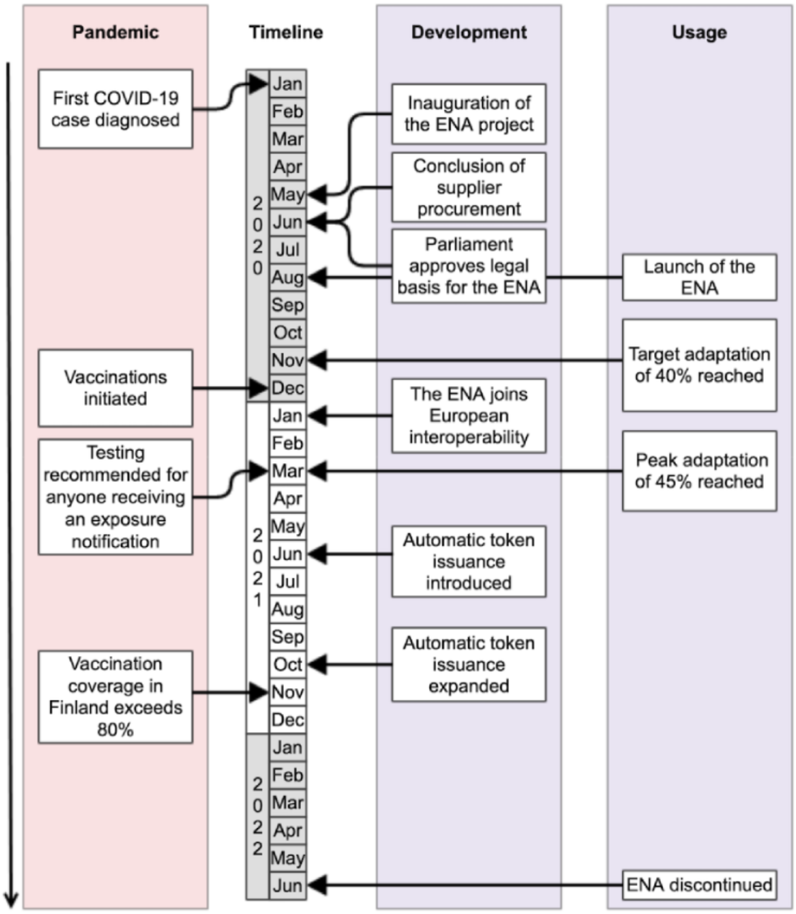
digital proximity tracing; DPT; exposure notification application; ENA; COVID-19; Finland; digital health; mobile health; mHealth; contact tracing; user; data privacy; effectiveness; mobile app; technology; public health

Introduction

Large-scale rollouts of digital proximity tracing (DPT) and exposure notification apps (ENAs) were undertaken hastily in several countries during the COVID-19 pandemic. It is important to gather and document the lessons learned in the process for possible similar situations in the future. In this viewpoint, we share our experiences of implementing and running the Finnish ENA (Koronavilkku). We base our account on the first-hand experiences of developing and maintaining the national ENA that were discussed at the final project workshop. We refer to feedback received from the public, health care professionals, and decision makers during the app's lifetime.

SARS-CoV-2 started to spread in the Finnish population at the end of February 2020 [1]. Like in many countries, in mid-March 2020, the government declared a state of emergency and introduced restrictions due to the coronavirus outbreak [2]. In early April, the European Commission adopted a recommendation on the use of mobile technology to control the COVID-19 pandemic [3], and the Finnish government initiated legislation work related to a national app later that month (Figure 1). The aim of the mobile app would be to promote public health by empowering everyone to participate in breaking SARS-CoV-2 transmission chains, as part of the government's strategy to combat the COVID-19 pandemic [4].

Figure 1. Timeline of key events in Finland during January 2020 to June 2022, related to the COVID-19 pandemic mitigation and the development and use of the Finnish COVID-19 exposure notification application (ENA).



Defining the Aims and Functionality of the ENA

According to the temporary law, the Finnish app would be voluntary and free for the public to use [5]. These principles are in line with findings in a review on best guidance for DPT apps [6]. The national legislation and the European guidelines defined the handling of personal data and directed the Finnish app to use a decentralized model and the Google and Apple Exposure Notification system without collecting location data [5,7]. Privacy and data security concerns have been raised by the public in studies on acceptability [8,9]. Local legislation is needed to ensure suitable and specific safeguards for privacy and data protection to gain public trust [10,11].

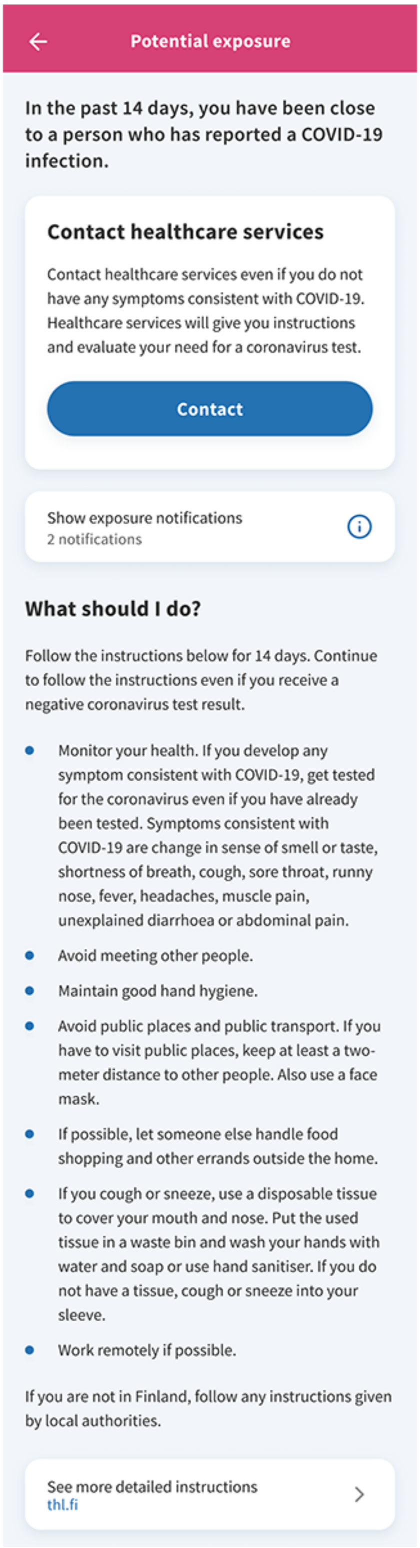
The Finnish app was divided into a mobile app and a backend system [12]. The mobile app calculated the exposure risk based on encounter duration, distance, and the presumed infectivity of the case. If a defined risk threshold was exceeded, the app notified the user of a potential exposure and provided further instructions. This functionality for sending and retrieving identifiers and matching them was similar in all decentralized ENAs in the European Union (EU) [13].

Based on early simulated models on potential for infection control [14,15], a high app download rate by the public was deemed more important than making the app a tool for health care professionals. In a simulation, a 56% app coverage would end the pandemic in the modelled country [14]. At that time, app coverage in other countries was up to 38% [16]. The project steering committee set the objective to 40% of the population in Finland, to have a concrete and realistic target coverage to mitigate the pandemic.

The app was designed to supplement conventional contact tracing, rather than replace it, and to serve as an early warning system and a trigger for end user action. The aim was that upon receiving an exposure notification, the user would change his or her behavior according to the guidance given by the app, and thus, break transmission chains (Figure 2). The instructions were kept in synchronization with the national guidelines.

Already in the planning stage, it was decided that we should emphasize privacy and maximize app uptake, which imposed limits on the development but gave a clear focus to the work. After the launch, several independent data security experts, including the National Cybersecurity Center, evaluated the app against the privacy concerns that had been discussed in the media [17,18]. We believe that the focus on privacy and data security, as well as the open communication about this aim, was instrumental in gaining public trust. The importance of upholding public trust to increase app uptake is consistent with other studies [19]. Distrust and a perceived lack of transparency can increase the belief in inadequate design, and in turn, decrease belief in app effectivity and app adoption [20]. Privacy concerns among the public are understandable, as many COVID-19–related apps required numerous rights, like access to photos, location, or microphone [21].

Figure 2. The guidance shown to users of the Finnish COVID-19 exposure notification app upon receipt of an exposure notification during March 4, 2021, to June 16, 2021. The instructions were updated in line with the national guidelines.



In the past 14 days, you have been close to a person who has reported a COVID-19 infection.

Contact healthcare services

Contact healthcare services even if you do not have any symptoms consistent with COVID-19. Healthcare services will give you instructions and evaluate your need for a coronavirus test.

Show exposure notifications
2 notifications

What should I do?

Follow the instructions below for 14 days. Continue to follow the instructions even if you receive a negative coronavirus test result.

- Monitor your health. If you develop any symptom consistent with COVID-19, get tested for the coronavirus even if you have already been tested. Symptoms consistent with COVID-19 are change in sense of smell or taste, shortness of breath, cough, sore throat, runny nose, fever, headaches, muscle pain, unexplained diarrhoea or abdominal pain.
- Avoid meeting other people.
- Maintain good hand hygiene.
- Avoid public places and public transport. If you have to visit public places, keep at least a two-meter distance to other people. Also use a face mask.
- If possible, let someone else handle food shopping and other errands outside the home.
- If you cough or sneeze, use a disposable tissue to cover your mouth and nose. Put the used tissue in a waste bin and wash your hands with water and soap or use hand sanitiser. If you do not have a tissue, cough or sneeze into your sleeve.
- Work remotely if possible.

If you are not in Finland, follow any instructions given by local authorities.

See more detailed instructions
thl.fi

Development

In most EU countries, ENAs were developed in collaboration between governmental institutions and private software developers [13]. This was also true for Finland, where establishing an ENA required several actors to join forces. The project group consisted of some 2 dozen individuals from the involved organizations. Because of the pandemic-related recommendations to work remotely, meetings were web based, and most team members met in person for the first time only at the final project workshop. One of the key lessons learned was the importance of seamless teamwork. Home organizations were less visible in the web-based meetings and discussions were informal yet structured. This fostered open communication, an excellent team spirit, a low level of hierarchy, and avoided silos, while also engaging project members and their home organizations. In addition to the core project group, nearly 200 individuals worked intermittently on different aspects of the app and related issues.

To support the project group, we set up an expert group of communicable disease physicians representing different municipalities and hospital districts. The group lacked decision-making power, but it was an important organ for keeping in touch with the health care services, where contact tracing was conducted and app tokens issued. The group was informed about the development and consulted on guidance and the calculation of risk factor parameters. The aim was to keep the parameters that defined possible exposure and recommendations in the exposure notifications in line with the national guidelines.

Communication and Deployment

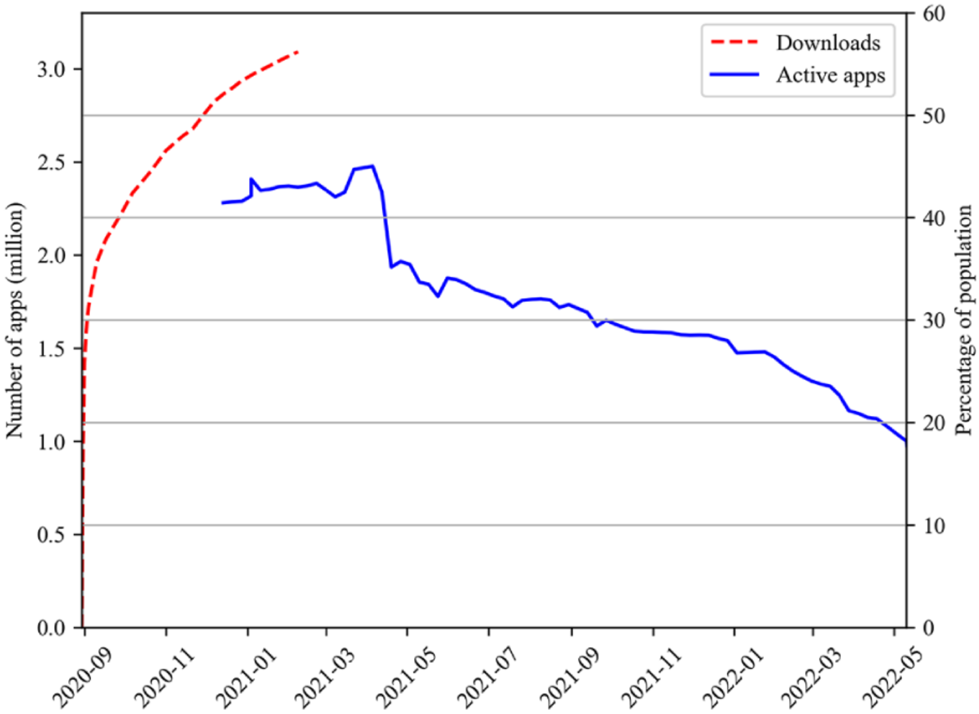
From the start, communication experts planned how to share easy-to-understand information about the ENA with the public. As identified elsewhere, end user benefits of an app and clarifications on how individual privacy will be protected are important to communicate [22]. Communication efforts included a series of public webinars during development and a marketing campaign in connection with the launch, and communication continued throughout the lifetime of the app. The responsibilities of the national COVID-19 helpline were expanded to give advice on the app and its use.

Finland approached advertising the app as part of a bigger strategy against the pandemic and started informing the public about the app at an early stage [13]. However, some other European countries included the public in the development by inviting feedback or carrying out public consultations to generate trust and increase acceptance [13]. Because of time and resource constraints, we were unable to accommodate similar public consultations.

The Finnish ENA was published in late summer 2020 (Figure 1), simultaneously or later than that of many other countries and states [13,23]. Probably largely due to the media attention that the Finnish ENA gained, especially around the launch, the app was downloaded by 1 million users within the first 24 hours. This accounts for nearly a fifth of the Finnish population. In the next 2 weeks, the app reached 2 million users, and the objective of 40% uptake was achieved in early November 2020 (Figure 3) [24]. Over the next few months, the number of active users continued to grow, and by spring 2021, the app had become one of the national or state ENAs with the highest proportion of users [13,25].

The ENA could be used in the 2 official languages of Finland, Finnish and Swedish, and additionally in English. However, some groups were not as frequent users of the app as the general population [26], and making the app available in further languages could have allowed it to have a wider and more inclusive public health impact.

Figure 3. Downloads or active app counts for the Finnish SARS-CoV-2 exposure notification application (ENA) in the period of August 31, 2020, to May 10, 2022. The number of Finnish ENA downloads was collected from mobile app stores. After February 8, 2021, the number of downloaded apps was not followed, as it no longer reflected the number of active apps. Active apps were defined as apps contacting the back end at least once per 24 hours.



Adaptation of the App in an Evolving Pandemic

The ENA was developed on the premise that it would be a short-term app requiring no further releases. In Finland, the legal mandates of several control and mitigation measures related to the COVID-19 pandemic were temporary. If needed, the validity of these laws was extended; during the app's lifetime, its legal basis was extended twice. In the changing pandemic situation, several changes were required in the app.

The project received numerous requests for minor changes and ideas for implementation, and we needed an unbiased method for prioritization and approval of them. Through the use of a structured form, we identified which parts of the app, its backend system or linked information, would be affected; the time and effort required to implement the change; and the estimated public health impact. This helped us reach informed decisions on what to implement.

In the fall of 2020, the national recommendation on quarantine duration was shortened from 14 days to 10 days. In line with this, changes in the app were required, so the duration an exposure notification was visible to the user corresponded with the updated guidelines. Similar changes were implemented 3 times during the lifetime of the app. Other updates were also made, such as a modification to exposure calculations to consider all exposures over the course of a day and changes to the interface to make it more user-friendly; these changes included displaying the number of potential exposures and adding a button to manually launch exposure checking. The motivation for all these updates was to keep the ENA current, attractive, and easy to use, thereby enabling its intended public health impact.

Several European countries launched their own mobile proximity contact tracing and warning apps, and an effort to explore sharing exposure notifications across apps was pursued by the European Commission and the member states [27]. In April 2020, a common approach for contact tracing apps was defined [7], and the first European countries joined the interoperability of apps in October 2020 [28]. Finland joined in early January 2021 (Figure 1) [13].

In accordance with increased laboratory testing capacity for SARS-CoV-2 and an updated national testing strategy [29], testing was extended to asymptomatic users who received an exposure notification, if the local testing capacity allowed (Figure 1). The web-based symptom assessment tool [30] that was linked to the ENA was eventually available for 72% of the population by August 2021. This enabled a majority of the users to book a test using various digital systems or platforms accessible through the internet.

Automatization of Token Issuance

Manual token issuance to laboratory-confirmed COVID-19 cases became an increasing burden on the contact tracing teams in the health care services as the pandemic grew. From the beginning, we recognized that automatization would have been faster, cheaper, and less demanding on health care professionals. Similar observations have also been made elsewhere [31]. We considered automatization unfeasible to implement, due to the decentralized health care and the multitude of information management systems in use. Further, early in the pandemic, there was fear and stigma surrounding contracting COVID-19

[32], and talking to a health care professional was preferred to an automatic text message communicating the positive result of a SARS-CoV-2 test.

We routinely followed the average timespan or delay from symptom onset until app token use, which seemed to depend on the workload of laboratories and contact tracing teams and increased when case numbers grew. This raised concerns about the usefulness of the app in times when fast notification of potential contacts would have been of most importance and made us revisit automatization of token issuance.

Automatic token issuance was pursued as of early 2021 in cooperation with 2 of the largest public health care operators, City of Helsinki and the Helsinki and Uusimaa hospital region, and it was taken into use in June and October 2021 (Figure 1). Tokens could be automatically issued based on digital laboratory results and patient contact information collected during the booking of a laboratory test. This made the process of allowing cases to notify potentially exposed individuals through the app quicker and more reliable, while saving resources in the health care services. Its public health impact would have been even greater if it had started early on. Notifying contacts quickly after exposure substantially affects the effect of quarantine or self-quarantine as a control measure for the spread of SARS-CoV-2 [15,33,34].

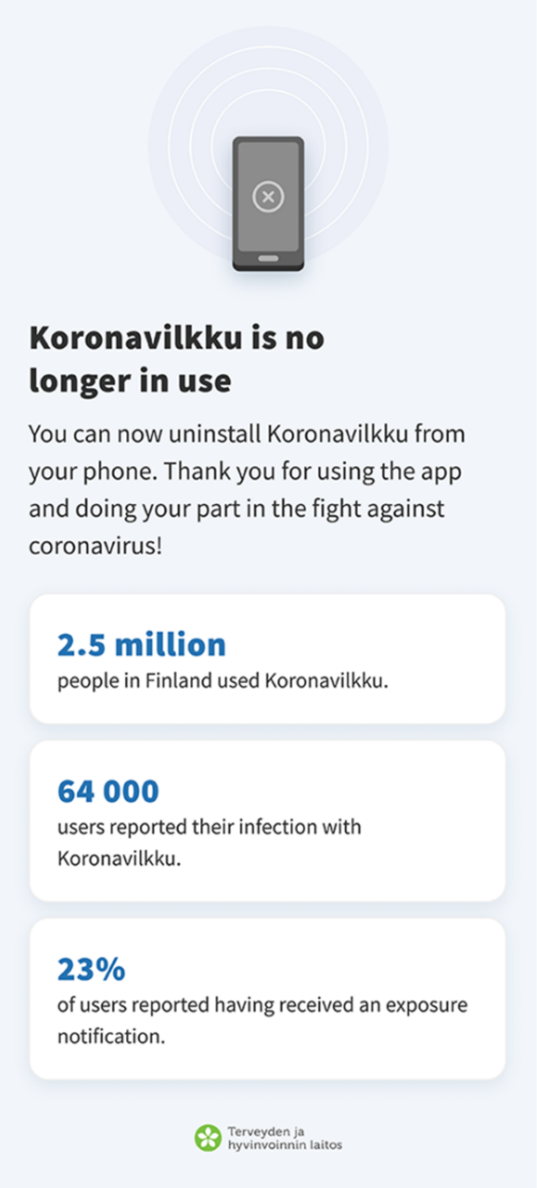
Shutdown

The use of the app slowly dwindled over the year 2021 (Figure 3); the likely reasons were many. Vaccinated individuals were no longer required to observe full quarantine after exposure to SARS-CoV-2 and were not automatically referred to testing if asymptomatic [35]. To many vaccinated individuals, the ENA may have seemed redundant, as the guidance also applied to those who received an app exposure notification [36].

With new SARS-CoV-2 variants causing more transmission, several municipalities and hospital regions stopped issuing app tokens in late 2021 to reduce the workload of contact tracing teams. Home tests for SARS-CoV-2 became more common and gradually replaced laboratory testing for much of the population. Positive home test results could not be registered in the app, as tokens were only issued through health care providers. Implementing a self-service option for token issuance was discussed in Finland and elsewhere [37]. Although it was not feasible for this ENA, self-service options should be considered from the start for any future similar projects to safeguard the longer-term public health use of the app.

We designed the Finnish ENA to gracefully retire by creating a screen that thanked the users and informed them about the shutdown (Figure 4). The ENA was discontinued on June 1, 2022 (Figure 1).

Figure 4. The "Thank you" screen of the Finnish SARS-CoV-2 exposure notification application—Koronavilkku. This screen was shown to users after Finland discontinued the use of the app.



Koronavilkku is no longer in use

You can now uninstall Koronavilkku from your phone. Thank you for using the app and doing your part in the fight against coronavirus!

2.5 million
people in Finland used Koronavilkku.

64 000
users reported their infection with Koronavilkku.

23%
of users reported having received an exposure notification.

Reflections and Recommendations

The development and launch of the Finnish ENA, Koronavilkku, illustrates how a widely used public health app, requiring new legislation, using new technology, and demanding new ways of cooperation, could be successful within a strict time frame. We consider the cornerstones of this success an excellent team spirit, low level of hierarchy within the project, open communication, avoiding silos, and a clearly defined goal. In addition, it required commitment from the project members and their home organizations.

The app was developed with a focus on the end users. Communication and marketing throughout played a key role in ensuring that potential users were familiar with the app, and we believe it contributed to the high number of downloads. Despite this, we received feedback that some health care professionals were disappointed that the app did not provide information to support their contact tracing work. In the literature, mobile apps are considered valuable tools for citizens, health care professionals, and decision makers [38], but there were difficulties in measuring and communicating these potential benefits clearly enough. In retrospect, more resources should have been allocated in the run-up to the app launch to clarify the goal to health care professionals and elaborate on the envisioned role of the app in the larger effort to mitigate the pandemic.

We believe that focusing on privacy and data security in communication and implementation strengthened public trust and, in turn, increased app uptake. However, compared to many other countries, public trust in the national government and public institutions is high in Finland [39]. Building on this trust, it may have been possible to offer the public an ENA that would have collected more information than the current interpretation of the national legislation and EU guidance allowed. In line with this, some user feedback indicated willingness to share more information with the app in exchange for receiving more information about potential exposures. This could have allowed the development of an app using the persuasive design model rather than the control design model that relies on self-monitoring. In another study [19], the persuasive design model, which uses social learning, was found to increase user intention to use the app. Enabling collection of data on exposures could have also benefited health care but would have required transparency and clear communication about the goals of data collection and its intended use. For future similar apps, we recommend enabling the collection of at least aggregated data on the number of exposure notifications generated.

Only some countries have published outcomes on their ENAs [40]. The difficulty of assessing the effectiveness of privacy-preserving DPT apps has been previously raised [23] and was also felt in relation to the Finnish ENA. The restrictions on what information we could collect through the app limited the insight into its potential effectiveness. Evidence of effectiveness may have maintained the number of users for a longer duration by swaying both health care professionals and end users to continue supporting the app. For future similar endeavors, we recommend investing in monitoring and communicating the potential benefits early in the process. If statutory limits constrain the collection of important data, regular surveys can be considered as an alternative to gain some knowledge on the use and effectiveness of such apps.

Other important lessons that we learned along the way include the fact that we did not initially foresee how many updates the app would require as the pandemic and mitigation efforts evolved. Forced updates and feeding text through a dedicated website into the app would have ensured that the guidance to users remained up-to-date, thereby enhancing the potential public health impact of the app. Another lesson learned is how a self-service option or earlier and more comprehensive automatization for token issuing for COVID-19 cases would have eased the burden on contact tracing teams and prolonged the useful lifespan of the app, while also notifying potentially exposed persons quicker and more reliably. Even with lessons learned and reflections on how the app could have been improved, the wide use of the Finnish ENA gave it potential to deliver public health and pandemic mitigation results.

Acknowledgments

We thank all our colleagues who worked on any aspect of the exposure notification application (ENA) or activities related to it for their cooperation and dedication. We acknowledge the users of the app for joining the effort to control the pandemic and for the feedback we received over the course of the lifetime of the app.

The project was funded by the Finnish Government (LTAE III/2020) and an European Union (EU) grant (FINLAND-LC-01590941). Funders had no role in writing or approving this manuscript.

Authors' Contributions

MP, SW, KP, and LS conceptualized the manuscript; MP wrote the first draft; all authors contributed to the manuscript.

Conflicts of Interest

None declared.

References

1. Coronavirus cases, tests, and deaths, daily and by region. Finnish Institute for Health and Welfare. URL: https://sampo.thl.fi/pivot/prod/en/epirapo/covid19case/summary_tshcddaily [accessed 2022-12-20]
2. Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021 Apr;5(4):529-538. [doi: 10.1038/s41562-021-01079-8] [Medline: 33686204]
3. European Commission. Communication from the Commission Guidance on Apps supporting the fight against COVID 19 pandemic in relation to data protection 2020/C 124 I/01. Official Journal of the European Union 2020 Apr 17;63((2020/C 124 I/01)):C 124 I/1-C 124 I/9 [FREE Full text]
4. Koronakriisin vaikutukset ja suunnitelma epidemian hallinnan hybridistrategiaksi (The effects of the corona crisis and a plan for a hybrid strategy for epidemic control) Last updated May 4, 2020. Valtioneuvosto. URL: https://julkaisut.valtioneuvosto.fi/handle/10024/162214 [accessed 2022-12-19]
5. Laki tartuntatautilain väliaikaisesta muuttamisesta 582/2020 (Act on the temporary amendment of the Communicable Diseases Act 582/2020) Last updated Jul 9, 2020. Finlex. URL: https://www.finlex.fi/fi/laki/alkup/2020/20200582 [accessed 2022-12-23]
6. O'Connell J, Abbas M, Beecham S, Buckley J, Chochlov M, Fitzgerald B, et al. Best practice guidance for digital contact tracing apps: a cross-disciplinary review of the literature. JMIR Mhealth Uhealth 2021 Jun 07;9(6):e27753 [FREE Full text] [doi: 10.2196/27753] [Medline: 34003764]
7. Mobile applications to support contact tracing in the EU's fight against COVID-19 -- Common EU Toolbox for Member States Version 1. eHealth Network. 2020 Apr 15. URL: https://health.ec.europa.eu/system/files/2020-04/covid-19_apps_en_0.pdf [accessed 2022-12-20]

8. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: Cross-country survey evidence. JMIR Mhealth Uhealth 2020 Jul 24:1-44 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
9. Osmanlliu E, Paquette J, Rodriguez Duarte MA, Bédard S, de Marcellis-Warin N, Zhegu M, et al. Public perspectives on exposure notification apps: a patient and citizen co-designed study. J Pers Med 2022 Apr 30;12(5):729 [FREE Full text] [doi: 10.3390/jpm12050729] [Medline: 35629150]
10. Lintvedt MN. COVID-19 tracing apps as a legal problem: an investigation of the Norwegian 'Smittestopp' app. OLR 2021 Dec 16;8(2):69-87. [doi: 10.18261/issn.2387-3299-2021-02-01]
11. Bagchi K, Bannan C, Franklin S. Digital tools for COVID-19 contact tracing: identifying and mitigating the equity, privacy, and civil liberties concerns. Harvard University Center for Ethics. 2020 Jun 02. URL: https://ethics.harvard.edu/digital-tools-for-contact-tracing [accessed 2023-04-11]
12. European proximity tracing: an interoperability architecture for contact tracing and warning apps. eHealth Network. 2020 Sep 02. URL: https://health.ec.europa.eu/system/files/2020-09/mobileapps_interop_architecture_en_0.pdf [accessed 2023-05-24]
13. European Commission. Digital contact tracing study on lessons learned, best practices and epidemiological impact of the common European approach on digital contact tracing to combat and exit the COVID-19 pandemic. Publications Office of the European Union. 2022 Dec 12. URL: https://digital-strategy.ec.europa.eu/en/library/covid-19-digital-contact-tracing-study [accessed 2022-12-19]
14. Hinch R, Probert W, Nurtay A. Effective configurations of a digital contact tracing app: a report to NHSX Last updated Aug 10, 2020. GitHub. 2020 Aug 10. URL: https://tinyurl.com/mrkxanh5 [accessed 2022-12-23]
15. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 Mar 31:eabb6936. [doi: 10.1126/science.abb6936] [Medline: 32234805]
16. Johnson B. Nearly 40% of Icelanders are using a covid app—and it hasn't helped much. MIT Technology Review. 2020 May 11. URL: https://www.technologyreview.com/2020/05/11/1001541/iceland-rakning-c19-covid-contact-tracing/ [accessed 2023-04-03]
17. Finnish Transport and Communications Agency Traficom Awards Cybersecurity Label to contact tracing app Koronavilkku. Traficom. 2020 Dec 09. URL: https://www.traficom.fi/en/news/finnish-transport-and-communications-agency-traficom-awards-cybersecurity-label-contact [accessed 2022-12-16]
18. Mäntylä J. Pelkäätkö Koronavilkkua? Vielä keväällä ammattihakkeri Benjamin Särkkä sanoi, ettei asentaisi koronasovellusta – 5 syytä miksi mieli on nyt muuttunut [(Are you afraid of Koronavilkku? In the spring, professional hacker Benjamin Särkkä said he would not install the corona application - 5 reasons why he has changed his mind)]. Yle Uutiset. 2020 Sep 02. URL: https://yle.fi/a/3-11523504 [accessed 2022-12-20]
19. Oyibo K, Morita PP. COVID alert: factors influencing the adoption of exposure notification apps among Canadian residents. Front Digit Health 2022;4:842661 [FREE Full text] [doi: 10.3389/fdgth.2022.842661] [Medline: 35360366]
20. Rowe F, Ngwenyama O, Richet J. Contact-tracing apps and alienation in the age of COVID-19. Eur J Inf Syst 2020 Sep 13;29(5):545-562. [doi: 10.1080/0960085x.2020.1803155]
21. Sharma T, Bashir M. Use of apps in the COVID-19 response and the loss of privacy protection. Nat Med 2020 May 26:1165-1167. [doi: 10.1038/s41591-020-0928-y] [Medline: 32457443]
22. Walrave M, Waeterloos C, Ponnet K. Tracing the COVID-19 virus: a health belief model approach to the adoption of a contact tracing app. JMIR Public Health Surveill 2020 Aug 05:e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
23. Bair H, Wanger JD, Shah NR. A brief history of exposure notification during the COVID-19 pandemic in the United States, 2020-2021. Public Health Rep 2022;137(2_suppl):23S-28S [FREE Full text] [doi: 10.1177/00333549221099533] [Medline: 35674242]
24. Väestön tieto- ja viestintätekniikan käyttö verkkojulkaisu. Official Statistics of Finland. URL: http://www.stat.fi/til/sutivi/2021/sutivi_2021_2021-11-30_tau_019_fi.html [accessed 2023-03-03]
25. Ladyzhets B. We investigated whether digital contact tracing actually worked in the US. MIT Technology Review. 2021 Jun 16. URL: https://tinyurl.com/4rwyd38r [accessed 2023-02-15]
26. Skogberg N, Koponen P, Lilja E. Access to information, preventive measures and working conditions during the coronavirus epidemic: findings of the population-based MigCOVID Survey among persons who have migrated to Finland. Finnish Institute for Health and Welfare. 2021. URL: https://www.julkari.fi/handle/10024/141067 [accessed 2022-12-19]
27. National COVID-19 contact tracing apps. European Parliment. 2020 May. URL: https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/652711/IPOL_BRI(2020)652711_EN.pdf [accessed 2023-01-11]
28. Coronavirus: commission starts testing interoperability gateway service for national contact tracing and warning apps. European Commission. 2020 Sep 14. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1606 [accessed 2022-12-20]

29. Kansallisen COVID-19-testausstrategian päivitys (Update of the national COVID-19 testing strategy). Sosiaali- ja terveysministeriö (Ministry of Social Affairs and Health). 2021 Feb 10. URL: https://tinyurl.com/5tau365y [accessed 2022-12-19]
30. Jormanainen V, Soininen L. Use and users of the web-based Omaolo Covid-19 symptom self-assesment tool in Finland since March 16, 2020. Stud Health Technol Inform 2021 May 27;281:739-743. [doi: 10.3233/SHTI210270] [Medline: 34042674]
31. Daniore P, Ballouz T, Menges D, von Wyl V. The SwissCovid Digital Proximity Tracing App after one year: were expectations fulfilled? Swiss Med Wkly 2021 Sep 08;151(35-36):35-36 [FREE Full text] [doi: 10.4414/smw.2021.w30031] [Medline: 34495624]
32. Lohiniva A, Dub T, Hagberg L, Nohynek H. Learning about COVID-19-related stigma, quarantine and isolation experiences in Finland. PLoS One 2021;16(4):e0247962 [FREE Full text] [doi: 10.1371/journal.pone.0247962] [Medline: 33852581]
33. Auranen K, Shubin M, Erra E, Isosomppi S, Kontto J, Leino T, et al. Efficacy and effectiveness of case isolation and quarantine during a growing phase of the COVID-19 epidemic in Finland. Sci Rep 2023 Jan 06;13(1):298 [FREE Full text] [doi: 10.1038/s41598-022-27227-2] [Medline: 36609431]
34. Quilty BJ, Clifford S, Hellewell J, Russell TW, Kucharski AJ, Flasche S, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 working group. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. Lancet Public Health 2021 Mar;6(3):e175-e183 [FREE Full text] [doi: 10.1016/S2468-2667(20)30308-X] [Medline: 33484644]
35. Quarantine practices revised for those who have been vaccinated against COVID-19 and those who have already had the disease Last updated May 19, 2021. Finnish Institute for Health and Welfare. URL: https://tinyurl.com/3huz25e6 [accessed 2023-01-11]
36. Latest version of the Koronavilkku app was released 17 June - with instructions for vaccinated people who receive an exposure notification. Finnish Institute for Health and Welfare. 2021 Jun 22. URL: https://tinyurl.com/3c9a5zc6 [accessed 2023-01-11]
37. Considerations on the use of self-tests for COVID-19 in the EU/EEA. European Centre for Disease Prevention and Control. 2021 Mar 17. URL: https://tinyurl.com/4cbjtyhr [accessed 2023-05-25]
38. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. J Med Internet Res 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: 10.2196/23170] [Medline: 33197234]
39. Building trust to reinforce democracy. OECD iLibrary. 2022 Jun 13. URL: https://www.oecd-ilibrary.org/content/publication/b407f99c-en [accessed 2023-01-11]
40. Pandit JA, Radin JM, Quer G, Topol EJ. Smartphone apps in the COVID-19 pandemic. Nat Biotechnol 2022 Jul;40(7):1013-1022. [doi: 10.1038/s41587-022-01350-x] [Medline: 35726090]

Abbreviations

DPT: digital proximity tracing
ENA: exposure notification application
EU: European Union

Edited by A Mavragani, T Sanchez; submitted 16.02.23; peer-reviewed by V von Wyl, V Jormanainen; comments to author 28.02.23; revised version received 17.04.23; accepted 16.06.23; published 13.07.23.

Please cite as:
Pihlajamäki M, Wickström S, Puranen K, Helve O, Yrttiaho A, Siira L
Implementing and Maintaining a SARS-CoV-2 Exposure Notification Application for Mobile Phones: The Finnish Experience
JMIR Public Health Surveill 2023;9:e46563
URL: https://publichealth.jmir.org/2023/1/e46563
doi:10.2196/46563
PMID:37440286

Viewpoint

Co-Constructing a Community-Based Telemedicine Program for People With Opioid Use Disorder During the COVID-19 Pandemic: Lessons Learned and Implications for Future Service Delivery

Stine Bordier Høj[1*], BCom, MBus, PhD; Catherine de Montigny[2*], LLB, MD; Sofiane Chougar[2], BSN; Robert Léandre[1], BA; Marie-Ève Beauchemin-Nadeau[2], MD; Geneviève Boyer-Legault[3], BA, MSc; Amélie Goyette[3], DCS; Sara-Kim Lamont[2], BSN; Julie Bruneau[1,2,4], MD, MSc

[1]Centre de recherche du Centre hospitalier de l'Université de Montréal, Montreal, QC, Canada
[2]Service de médecine des toxicomanies, Centre hospitalier de l'Université de Montréal, Montreal, QC, Canada
[3]CACTUS Montréal, Montreal, QC, Canada
[4]Department of family medicine and emergency medicine, Université de Montréal, Montreal, QC, Canada
[*]these authors contributed equally

Corresponding Author:
Julie Bruneau, MD, MSc
Centre de recherche du Centre hospitalier de l'Université de Montréal
900 Rue Saint-Denis
Montreal, QC, H2X 0A9
Canada
Phone: 1 514 890 8000 ext 35713
Email: julie.bruneau@umontreal.ca

Abstract

The COVID-19 pandemic triggered unprecedented expansion of telemedicine, including in the delivery of opioid agonist treatment (OAT) for people with opioid use disorder (OUD). However, many people with OUD lack the technological resources necessary for remote care, have complex needs, and are underserved, with precarious access to mainstream services. To address the needs of these individuals, we devised a unique program to deliver OAT via telemedicine with the support of community outreach workers in Montreal (Quebec, Canada). The program was co-constructed by the service de médecine des toxicomanies of the Centre hospitalier de l'Université de Montréal (CHUM-SMT)—a hospital-based addiction medicine service—and CACTUS Montréal—a community-based harm reduction organization known and trusted by its clientele. All procedures were jointly developed to enable flexible and rapid appointment scheduling. CACTUS Montréal workers promoted the program, facilitated private on-site telemedicine connections to the CHUM-SMT, accompanied patients during web-based appointments if requested, and provided ongoing holistic support and follow-up. The CHUM-SMT offered individualized OAT regimens and other health services as needed. Overall, our experience as clinicians and community-based workers intimately involved in establishing and running this initiative suggests that participants found it to be convenient, nonjudgmental, and responsive to their needs, and that the implication of CACTUS Montréal was highly valued and integral to patient engagement and retention. Beyond the context of the COVID-19 pandemic, similar programs may present a flexible and accessible means to deliver alternative treatment options for people with OUD disengaged from traditional care, bridge gaps between communities and health providers, and improve access to care in rural or remote settings.

(JMIR Public Health Surveill 2023;9:e39236) doi:10.2196/39236

KEYWORDS

opioid agonist treatment; opioid use disorder; medications for opioid use disorder; harm reduction; access to care; retention; telemedicine; telehealth; community-based services; opioid use; remote care; healthcare service; health care service; COVID-19; substance abuse; opioid disorder

Background

The field of telemedicine has a long-standing interest in improving the accessibility and quality of health care for remote and otherwise underresourced communities [1,2]. In response to the COVID-19 pandemic and associated physical distancing measures, the use of telemedicine modalities expanded exponentially to facilitate clinical services for a wide spectrum of health problems [3-5]. Continued integration into regular clinical care is likely, particularly in high-income countries [4,6,7], and could greatly improve service delivery for persons whose circumstances render in-person visits inconvenient or undesirable [7]. A lack of attention to the "digital divide" may, however, exclude groups that have been historically marginalized from experiencing these benefits and could even exacerbate health inequities [7-12]. Additionally, engaging such populations calls for approaches that reduce the social as well as physical barriers to accessing health providers [13-16]. This viewpoint paper describes how integrating a trusted community service into the framework of a hospital-based telemedicine service in Montreal (Quebec, Canada) created a program capable of engaging and retaining vulnerable people with opioid use disorder (OUD) in opioid agonist treatment (OAT) during the COVID-19 pandemic. We recount the success of this program as clinicians and community-based workers involved from its inception and by drawing upon patient testimonials, with a view to inform future service delivery for such populations.

People With OUD, the COVID-19 Pandemic, and OAT Delivery via Telemedicine

People using illicit opioids experience a high burden of disease [17,18] and may be disproportionately impacted by major disruptive events such as a pandemic [19,20], particularly when facing other challenges such as housing instability [21,22]. Apart from risks related to COVID-19 exposure [23], public health responses to the pandemic directly impacted the risk environment shaping drug-related harms [24]: many low-threshold services that people with OUD rely on to meet essential needs or reduce harms associated with drug consumption, such as day shelters and supervised consumption sites, were forced to temporarily close or reduce their capacity, making already scarce resources even less accessible, while pandemic-related prevention measures also affected people's ability to generate income, secure shelter, and obtain safe and secure supplies of drugs [25-29]. In Canada, these combined challenges were reflected in a 95% increase in opioid-related deaths in the first year of the pandemic (April 2020 to March 2021) compared with the year before [30].

Treatment with opioid agonists such as methadone or buprenorphine/naloxone is the first-line treatment for opioid use disorder [31,32] and can be highly effective in reducing illicit opioid use, preventing overdose, and improving multiple other health and social outcomes [18]. Nevertheless, inadequate treatment coverage and poor retention persist worldwide [18,33], with evidence for suboptimal and restrictive delivery practices (eg, inadequate medication dosing, poor access to unsupervised or "take-home" doses, and frequent urine drug screening) in many settings [34,35]. In Canada, as elsewhere, OAT is highly regulated [36] and typically delivered within health institution settings [37]. Prior to the pandemic, an estimated 66% of people who inject drugs nationwide, and 44% of those in Quebec, were receiving OAT [38]. People with OUD in Montreal had called for greater service outreach in places in which they spend their time, while emphasizing increased patient autonomy, flexible scheduling, diversified treatment options, and the implication of peer workers as ways to improve OAT services [39]. Indeed, participants in our program often recounted unsatisfactory experiences with conventional OAT services, describing a lack of responsiveness to their needs and experiences, ineffective communication, an absence of participatory decision-making, and stigmatization, alongside difficulties with medications and dosing, withdrawal symptoms and side effects, and rigid service rules.

OAT delivery in Canada is bound by federal regulations under the Controlled Drugs and Substances Act (CDSA), while provincial and territorial governments regulate the scope of practice for health professionals [40]. Following the declaration of the COVID-19 pandemic in March 2020, federal health authorities issued emergency legal exemptions to the CDSA, enabling verbal prescription of controlled substances alongside expanded authorizations for pharmacists to extend, renew, and transfer such prescriptions [41]. Meanwhile, to support the implementation of best practices, federal health authorities commissioned the Canadian Research Initiative in Substance Misuse (CRISM) to develop a series of national guidance documents, including recommendations for the use of telemedicine in addiction services [42] and prescribing of psychoactive substances to support people with OUD who need to self-isolate [43]. The medical regulatory authority of Quebec rapidly updated their OUD treatment and telemedicine guidelines to integrate these changes, relax restrictions on unsupervised dosing of long-acting opioid agonists, and endorse the use of telemedicine for OAT initiation and continuation [44-46]. Quebec-specific expert recommendations for OAT delivery, including guidance for "safer supply" prescriptions of short-acting opioids, were subsequently released in October 2020 [47]. Together, these changes represent a seismic shift in the regulatory structures and guidance shaping treatment for people with OUD [48]. Evidence suggests a broad uptake of telemedicine prescribing in Canada, with over half of OAT clinics surveyed offering this service by June 2020 [49].

Recognizing a Need for Further Adaptation

The Centre hospitalier de l'Université de Montréal (CHUM) is an academic hospital in downtown Montreal, whose addiction medicine service (Service de médecine des toxicomanies [CHUM-SMT]) offers an integrated model of care for people with substance use disorders and complex social and medical needs. This model spans a wide range of interventions, from harm reduction to specialized multidisciplinary medical and psychiatric treatment, and includes a low-barrier OAT program. From March 2020, OAT practices were rapidly amended in line with the aforementioned recommendations to shift outpatient consultations to web-based or telephone visits and allow initiation of new patients without requiring that they visit the outpatient clinic. Like clinicians in other settings [50-52], however, we noted that access to OAT remained inequitable under a telemedicine model, particularly for individuals unable to access a phone or computer or are otherwise unwilling to access hospital-based services for fear of stigmatization [8,53].

Recognizing the need to actively reach out and build trust with this population that has been socially marginalized, we created a new partnership with a community-based harm reduction organization, CACTUS Montréal. Located in the core of downtown Montreal, CACTUS Montréal is home to one of North America’s oldest needle-syringe programs and takes a pragmatic and humanistic approach to serving people who use drugs through prevention, education, and leisure activities. It was also one of the few such services to remain open in Montreal during the initial months of the COVID-19 pandemic. Together, we co-designed a novel model of care premised on offering easily accessible, high-quality telecare for people with OUD with the support of the CACTUS Montréal community workers they trust.

Procedures were developed jointly to enable a flexible and rapid appointment scheduling system that would meet the needs of this population while ensuring confidential flows of patient information. CACTUS Montréal community workers informed their clients of the program, facilitated initial eligibility screening, and helped interested clients book an on-site web-based appointment with the CHUM-SMT team, generally within the same week. Participants thus remained physically within the CACTUS Montréal facilities, where they could not only access the technological resources and assistance necessary to engage in care but also draw on the support of community workers during and after their appointment. This enabled warm hand off from a trusted person to the CHUM-SMT and equally created an environment in which patients could easily pose questions, reschedule appointments, and follow up with staff about the program throughout their participation. Clinics ran on Tuesday and Wednesday afternoons, allowing some walk-in appointments as needed.

During the web-based appointments, the CHUM-SMT team (comprising a physician and a nurse) evaluated each patient and initiated an individualized treatment regimen. To foster engagement in care, treatment was tailored to the participants’ particular needs and their prior experiences with OAT. An initial regimen typically included a long-acting opioid (eg, methadone, buprenorphine/naloxone, or extended-release morphine) combined with a short-acting opioid (eg, hydromorphone) to manage withdrawal, increase comfort, and reduce the risk of illicit opioid use and overdose during initiation. Prescriptions were sent by fax directly to the participant’s local pharmacy and a follow-up appointment was immediately scheduled. Importantly, the CHUM-SMT team could also assess and treat patients’ other medical needs, with CACTUS Montréal nurses available to draw blood samples (eg, for HIV or hepatitis C testing) as necessary. Participant consent was obtained from the outset to define the nature of information exchange between CHUM-SMT and CACTUS Montréal, who communicated frequently to schedule appointments based on the participants’ needs.

To support implementation, both partners entered a formal collaboration agreement and devised a tool kit including a shared appointment calendar, patient evaluation and referral forms, user guides for connecting to various telemedicine tools, and trainings in the web-based clinic and on OAT for CACTUS Montréal staff. Patient flows were managed with assistance from the CHUM Network Flow Optimization Center (Centre d’optimisation des flux reseau), a service promoting accessibility and continuity of care through various measures including coordination of linkages to the broader health network and telehealth services.

Successes, Challenges, and the Importance of Building Institutional Relationships

A major challenge at the time of implementation was to rapidly organize the aforementioned trainings, tools, and procedures so that CACTUS workers could appropriately support clients initiating OAT through the service. This effort was facilitated by a previously well-established working relationship and institutional support from both partners, with our first patient enrolled just 30 days after the first planning meeting and 45 days after declaration of the COVID-19 public health emergency in Quebec.

The initiative was viewed as transformative by the staff and enthusiastically received by the participants, who, we believe, found the program convenient, flexible, and responsive to their needs despite some reservations about the “impersonal” nature of telemedicine. In our view, the implication of CACTUS Montréal and its community workers was critical to creating an environment of trust, confidence, and nonjudgment; maintaining open lines of communication between appointments; and enabling access to care within a setting that participants were already frequenting. Many participants had long-standing ties and positive relationships with CACTUS Montréal, which contrasted starkly with testimonies of their unsatisfactory experiences with mainstream health services. This appeared to facilitate engagement and lend legitimacy to the program, with telemedicine being a tool enabling us to quickly and easily reach beyond the confines of the conventional health system. Unlike programs delivering OAT through telemedicine alone, our hybrid model also enabled easy access to laboratory testing and subsequent initiation of care for other chronic health problems (eg, HIV and hepatitis C virus) affecting this population.

Partnerships between institutional health services and independent community organizations can be complex to develop because the community approach to health, particularly in the area of harm reduction and prevention of sexually transmitted and blood-borne infections, has developed in response to the inability of traditional services to reach and maintain ties with groups that have been historically marginalized. In this case, a history of constructive engagement between the CHUM and CACTUS Montréal—also in the context of various prior research initiatives—was the critical foundation that allowed us to respond quickly and positively to the disruption caused by the COVID-19 pandemic. Our telemedicine initiative has further strengthened these ties, marking a significant step toward further innovations in service delivery. Indeed, the program proved so popular, it has now expanded to a second harm reduction organization in Montreal. Mobilizing the technical and human resources needed to establish new partnerships and meet client demand has been taxing and confirms a need to develop longer-term plans and infrastructure so that the model, once fully evaluated, can be deployed more broadly.

Conclusion

Facilitated by the COVID-19 pandemic and the ensuing amendments to federal and provincial regulations and recommendations for best practices, our community-based telemedicine program succeeded in bringing OAT to people with OUD within a space where they feel comfortable and accepted. The implication of CACTUS Montréal, an ally who accompanies clients through difficulties at their own pace and based on their own needs and priorities, was fundamental to engaging our patients and responded to priorities previously expressed by local people with OUD.

It is imperative that such initiatives and the emergency-led legal and structural contexts underpinning them persist, evolve, and expand well beyond the pandemic. Emerging evidence suggests that both patients and providers support the continuation of regulatory reforms and telemedicine delivery of OAT and perceive these as facilitating individualized patient-centered care [54,55]. More programs enabled by these innovations must be quickly and thoroughly evaluated to build an evidence base for their safety and effectiveness and guide further adaptation. Investments must then allow the science to move alongside implementation, scale-up, and formalization of such novel care pathways toward a more agile and adaptive response to the continually evolving overdose crisis [56,57]. The success of vaccine development and scale-up initiatives amid the pressing threat of the COVID-19 pandemic, as well as the rapid deployment of telemedicine and other sweeping societal interventions, has demonstrated our capacity to innovate and implement simultaneously. Reducing overdoses among people with OUD deserves the same sense of emergency and commitment to innovative thinking. We firmly believe that programs such as ours can reduce health inequalities and save lives by meeting people “where they are,” bridging gaps to institutional services and shifting patient-provider relationships toward a place of mutual understanding and holistic engagement [58].

Acknowledgments

This work was supported by the Canadian Institutes of Health Research through the CRISM Quebec-Atlantic Node (#SMN-139149). Further support was obtained from the Canada Research Chair in Addiction Medicine, JB. SBH receives salary support from a CIHR project grant (#PJT-165880).

Conflicts of Interest

JB has served on the advisory boards of AbbVie and Gilead Sciences and has received research funding from Gilead Sciences, outside of the current work. SC has served as a consultant to AbbVie, outside of the current work.

References

1. Bashshur RL. Telemedicine and health care. Telemed J E Health 2002;8(1):5-12 [FREE Full text] [doi: 10.1089/15305620252933365] [Medline: 12020402]
2. Colbert GB, Venegas-Vera AV, Lerma EV. Utility of telemedicine in the COVID-19 era. Rev Cardiovasc Med 2020 Dec 30;21(4):583-587 [FREE Full text] [doi: 10.31083/j.rcm.2020.04.188] [Medline: 33388003]
3. Webster P. Virtual health care in the era of COVID-19. Lancet 2020 Apr;395(10231):1180-1181. [doi: 10.1016/s0140-6736(20)30818-7]
4. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of telehealth during the COVID-19 pandemic: scoping review. J Med Internet Res 2020 Dec 01;22(12):e24087 [FREE Full text] [doi: 10.2196/24087] [Medline: 33147166]
5. Hincapié MA, Gallego JC, Gempeler A, Piñeros JA, Nasner D, Escobar MF. Implementation and usefulness of telemedicine during the COVID-19 pandemic: a scoping review. J Prim Care Community Health 2020 Dec 10;11:2150132720980612 [FREE Full text] [doi: 10.1177/2150132720980612] [Medline: 33300414]
6. Stokel-Walker C. Why telemedicine is here to stay. BMJ 2020 Oct 06;371:m3603. [doi: 10.1136/bmj.m3603] [Medline: 33023876]
7. Dorsey ER, Topol EJ. State of telehealth. N Engl J Med 2016 Jul 14;375(2):154-161. [doi: 10.1056/NEJMra1601705] [Medline: 27410924]
8. Alami H, Lehoux P, Attieh R, Fortin J, Fleet R, Niang M, et al. A “Not So Quiet” revolution: systemic benefits and challenges of telehealth in the context of COVID-19 in Quebec (Canada). Front. Digit. Health 2021 Oct 5;3. [doi: 10.3389/fdgth.2021.721898]
9. Katzow MW, Steinway C, Jan S. Telemedicine and health disparities during COVID-19. Pediatrics 2020 Aug;146(2). [doi: 10.1542/peds.2020-1586] [Medline: 32747592]

10. Annaswamy TM, Verduzco-Gutierrez M, Frieden L. Telemedicine barriers and challenges for persons with disabilities: COVID-19 and beyond. Disabil Health J 2020 Oct;13(4):100973 [FREE Full text] [doi: 10.1016/j.dhjo.2020.100973] [Medline: 32703737]
11. Seifert A, Batsis JA, Smith AC. Telemedicine in long-term care facilities during and beyond COVID-19: challenges caused by the digital divide. Front Public Health 2020 Oct 26;8:601595 [FREE Full text] [doi: 10.3389/fpubh.2020.601595] [Medline: 33194999]
12. Contreras CM, Metzger GA, Beane JD, Dedhia PH, Ejaz A, Pawlik TM. Telemedicine: patient-provider clinical engagement during the COVID-19 pandemic and beyond. J Gastrointest Surg 2020 Jul;24(7):1692-1697 [FREE Full text] [doi: 10.1007/s11605-020-04623-5] [Medline: 32385614]
13. O'Donnell P, Tierney E, O'Carroll A, Nurse D, MacFarlane A. Exploring levers and barriers to accessing primary care for marginalised groups and identifying their priorities for primary care provision: a participatory learning and action research study. Int J Equity Health 2016 Dec 03;15(1):197 [FREE Full text] [doi: 10.1186/s12939-016-0487-5] [Medline: 27912783]
14. Stone EM, Kennedy-Hendricks A, Barry CL, Bachhuber MA, McGinty EE. The role of stigma in U.S. primary care physicians' treatment of opioid use disorder. Drug Alcohol Depend 2021 Apr 01;221:108627 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108627] [Medline: 33621805]
15. Werremeyer A, Mosher S, Eukel H, Skoy E, Steig J, Frenzel O, et al. Pharmacists' stigma toward patients engaged in opioid misuse: When "social distance" does not mean disease prevention. Subst Abus 2021 Mar 22;42(4):919-926. [doi: 10.1080/08897077.2021.1900988] [Medline: 33750283]
16. Lang K, Neil J, Wright J, Dell CA, Berenbaum S, El-Aneed A. Qualitative investigation of barriers to accessing care by people who inject drugs in Saskatoon, Canada: perspectives of service providers. Subst Abuse Treat Prev Policy 2013 Oct 01;8:35 [FREE Full text] [doi: 10.1186/1747-597X-8-35] [Medline: 24079946]
17. Degenhardt L, Bucello C, Mathers B, Briegleb C, Ali H, Hickman M, et al. Mortality among regular or dependent users of heroin and other opioids: a systematic review and meta-analysis of cohort studies. Addiction 2011 Jan;106(1):32-51. [doi: 10.1111/j.1360-0443.2010.03140.x] [Medline: 21054613]
18. Degenhardt L, Grebely J, Stone J, Hickman M, Vickerman P, Marshall BDL, et al. Global patterns of opioid use and dependence: harms to populations, interventions, and future action. Lancet 2019 Oct 26;394(10208):1560-1579 [FREE Full text] [doi: 10.1016/S0140-6736(19)32229-9] [Medline: 31657732]
19. Zolopa C, Hoj S, Bruneau J, Meeson J, Minoyan N, Raynault M, et al. A rapid review of the impacts of "Big Events" on risks, harms, and service delivery among people who use drugs: Implications for responding to COVID-19. Int J Drug Policy 2021 Jun;92:103127 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103127] [Medline: 33549464]
20. Nguyen T, Buxton JA. Pathways between COVID-19 public health responses and increasing overdose risks: a rapid review and conceptual framework. Int J Drug Policy 2021 Jul;93:103236 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103236] [Medline: 33838990]
21. Hyshka E, Dong K, Meador K, Speed K, Abele B, LeBlanc S, et al. Supporting people who use substances in shelter settings during the COVID-19 pandemic. Canadian Research Initiative in Substance Misuse (Edmonton, Alberta). 2020. URL: https://crism.ca/wp-content/uploads/2020/06/CRISM-Guidance-Supporting-People-Who-Use-Substances-in-Emergency-Shelter-Settings-V1.pdf [accessed 2023-06-19]
22. Liu M, Richard L, Campitelli MA, Nisenbaum R, Dosani N, Dhalla IA, et al. Drug overdoses during the COVID-19 pandemic among recently homeless individuals. Addiction 2022 Jun 14;117(6):1692-1701 [FREE Full text] [doi: 10.1111/add.15823] [Medline: 35129239]
23. Patel J, Nielsen F, Badiani A, Assi S, Unadkat V, Patel B, et al. Poverty, inequality and COVID-19: the forgotten vulnerable. Public Health 2020 Jun;183:110-111 [FREE Full text] [doi: 10.1016/j.puhe.2020.05.006] [Medline: 32502699]
24. Rhodes T. The ‘risk environment’: a framework for understanding and reducing drug-related harm. Int J Drug Policy 2002;13(2):85-94. [doi: 10.1016/S0955-3959(02)00007-5]
25. Radfar SR, De Jong CAJ, Farhoudian A, Ebrahimi M, Rafei P, Vahidi M, ISAM-PPIG Global Survey Consortium, et al. Reorganization of substance use treatment and harm reduction services during the COVID-19 pandemic: a global survey. Front Psychiatry 2021 Apr 29;12:639393 [FREE Full text] [doi: 10.3389/fpsyt.2021.639393] [Medline: 34025471]
26. Morin KA, Acharya S, Eibl JK, Marsh DC. Evidence of increased Fentanyl use during the COVID-19 pandemic among opioid agonist treatment patients in Ontario, Canada. Int J Drug Policy 2021 Apr;90:103088 [FREE Full text] [doi: 10.1016/j.drugpo.2020.103088] [Medline: 33385974]
27. Russell C, Ali F, Nafeh F, Rehm J, LeBlanc S, Elton-Marshall T. Identifying the impacts of the COVID-19 pandemic on service access for people who use drugs (PWUD): a national qualitative study. J Subst Abuse Treat 2021 Oct;129:108374 [FREE Full text] [doi: 10.1016/j.jsat.2021.108374] [Medline: 34080545]
28. Whitfield M, Reed H, Webster J, Hope V. The impact of COVID-19 restrictions on needle and syringe programme provision and coverage in England. Int J Drug Policy 2020 Sep;83:102851 [FREE Full text] [doi: 10.1016/j.drugpo.2020.102851] [Medline: 32736959]
29. Ali F, Russell C, Nafeh F, Rehm J, LeBlanc S, Elton-Marshall T. Changes in substance supply and use characteristics among people who use drugs (PWUD) during the COVID-19 global pandemic: a national qualitative assessment in Canada. Int J Drug Policy 2021 Jul;93:103237 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103237] [Medline: 33893026]

30. Apparent opioid and stimulant toxicity deaths: Surveillance of opioid- and stimulant-related harms in Canada, January 2016 to June 2021. Government of Canada. Ottawa URL: https://publications.gc.ca/collections/collection_2021/aspc-phac/HP33-7-2021-2-eng.pdf [accessed 2023-06-19]
31. Department of Mental Health and Substance Abuse, World Health Organization. Guidelines for the Psychosocially Assisted Pharmacological Treatment of Opioid Dependence. Geneva: World Health Organization; 2009.
32. Bruneau J, Ahamad K, Goyer M, Poulin G, Selby P, Fischer B, CIHR Canadian Research Initiative in Substance Misuse. Management of opioid use disorders: a national clinical practice guideline. CMAJ 2018 Mar 05;190(9):E247-E257 [FREE Full text] [doi: 10.1503/cmaj.170958] [Medline: 29507156]
33. Larney S, Peacock A, Leung J, Colledge S, Hickman M, Vickerman P, et al. Global, regional, and country-level coverage of interventions to prevent and manage HIV and hepatitis C among people who inject drugs: a systematic review. Lancet Glob Health 2017 Dec;5(12):e1208-e1220 [FREE Full text] [doi: 10.1016/S2214-109X(17)30373-X] [Medline: 29074410]
34. Jin H, Marshall BDL, Degenhardt L, Strang J, Hickman M, Fiellin DA, et al. Global opioid agonist treatment: a review of clinical practices by country. Addiction 2020 Dec 19;115(12):2243-2254 [FREE Full text] [doi: 10.1111/add.15087] [Medline: 32289189]
35. Artenie AA, Minoyan N, Jacka B, Høj S, Jutras-Aswad D, Roy É, et al. Opioid agonist treatment dosage and patient-perceived dosage adequacy, and risk of hepatitis C infection among people who inject drugs. CMAJ 2019 Apr 29;191(17):E462-E468 [FREE Full text] [doi: 10.1503/cmaj.181506] [Medline: 31036608]
36. Priest KC, Gorfinkel L, Klimas J, Jones AA, Fairbairn N, McCarty D. Comparing Canadian and United States opioid agonist therapy policies. Int J Drug Policy 2019 Dec;74:257-265 [FREE Full text] [doi: 10.1016/j.drugpo.2019.01.020] [Medline: 30765118]
37. Eibl JK, Morin K, Leinonen E, Marsh DC. The state of opioid agonist therapy in Canada 20 years after federal oversight. Can J Psychiatry 2017 Jul 19;62(7):444-450 [FREE Full text] [doi: 10.1177/0706743717711167] [Medline: 28525291]
38. Jacka B, Larney S, Degenhardt L, Janjua N, Høj S, Krajden M, et al. Prevalence of Injecting Drug Use and Coverage of Interventions to Prevent HIV and Hepatitis C Virus Infection Among People Who Inject Drugs in Canada. Am J Public Health 2020 Jan;110(1):45-50 [FREE Full text] [doi: 10.2105/AJPH.2019.305379] [Medline: 31725310]
39. Lachapelle E, Archambault L, Blouin C, Perreault M. Perspectives of people with opioid use disorder on improving addiction treatments and services. Drugs: Education, Prevention and Policy 2020 Oct 02;28(4):316-327. [doi: 10.1080/09687637.2020.1833837]
40. Lussier-Hoskyn S, Skinner B. Pharmacy scope of practice and access to Opioid Agonist Therapies (OAT) in Canada after COVID-19. Can Health Policy 2022 Jan 31. [doi: 10.54194/DIGR6414]
41. Saxe J. Subsection 56(1) class exemption for patients, practitioners and pharmacists prescribing and providing controlled substances in Canada. Government of Canada. 2020. URL: https://www.canada.ca/en/health-canada/services/health-concerns/controlled-substances-precursor-chemicals/policy-regulations/policy-documents/section-56-1-class-exemption-patients-pharmacists-practitioners-controlled-substances-covid-19-pandemic.html [accessed 2022-03-28]
42. Bruneau J, Rehm J, Wild T, Wood E, Sako A, Swansburg J, et al. Telemedicine Support for Addiction Services: National Rapid Guidance Document. Canadian Research Initiative in Substance Misuse (Montreal, Quebec). 2020. URL: https://crism.ca/wp-content/uploads/2020/05/CRISM-National-Rapid-Guidance-Telemedicine-V1.pdf [accessed 2023-06-19]
43. Brar R, Bruneau J, Butt P, Goyer M, Lim R, Poulin G, et al. Medications and other clinical approaches to support physical distancing for people who use substances during the COVID-19 pandemic: National Rapid Guidance Document. Canadian Research Initiative in Substance Misuse (Vancouver, British Columbia). 2020. URL: https://crism.ca/wp-content/uploads/2020/06/CRISM-Guidance-Medications-and-other-clinical-approaches-22062020-final.pdf [accessed 2023-06-19]
44. Les téléconsultations réalisées par les médecins durant la pandémie de COVID-19: Guide à l'intention des médecins. Collège des médecins du Québec (Montreal, QC). 2020. URL: https://numerique.banq.qc.ca/patrimoine/details/52327/4073726 [accessed 2023-06-19]
45. Trouble lié à l'utilisation d'opioïdes (TUO): prescription d'un traitement par agonistes opioïdes (TAO) durant la pandémie. Collège des médecins du Québec (Montreal, Quebec). URL: http://www.cmq.org/page/fr/covid-19-trouble-lie-a-l-utilisation-d-opioides-tuo-prescription-d-un-traitement-par-agonistes-opioides-tao-durant-la-pandemie.aspx [accessed 2023-06-19]
46. Le traitement du trouble lié à l'utilisation d'opioïdes: Lignes directrices 03/2020. Collège des médecins du Québec (Montreal, Quebec). 2020. URL: http://www.cmq.org/publications-pdf/p-1-2020-03-20-fr-le-traitement-du-trouble-lie-a-l-utilisation-d-opioides-tuo.pdf [accessed 2023-06-19]
47. Goyer M, Hudon K, Plessis-Bélair M, Ferguson Y. Substance Replacement Therapy in the Context of the COVID-19 Pandemic in Québec: Clinical Guidance for Prescribers. Institut universitaire sur les dépendances (Montreal, Quebec). 2020. URL: http://dependanceitinerance.ca/wp-content/uploads/2020/10/Guide-Pharmaco-COVID_ANG-VF.19.10.20.pdf [accessed 2023-06-19]
48. Potenza MN, El-Guebaly N. Addiction medicine in the time of COVID-19: an overview of the 2020 Joint Scientific Annual Conference of the International Society of Addiction Medicine and Canadian Society of Addiction Medicine. Subst Abus 2021 May 16;42(3):261-263. [doi: 10.1080/08897077.2021.1949668] [Medline: 34283688]

49. Joudrey PJ, Adams ZM, Bach P, Van Buren S, Chaiton JA, Ehrenfeld L, et al. Methadone access for opioid use disorder during the COVID-19 pandemic within the United States and Canada. JAMA Netw Open 2021 Jul 01;4(7):e2118223 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.18223] [Medline: 34297070]
50. Hew A, Arunogiri S, Lubman DI. Challenges in delivering telemedicine to vulnerable populations: experiences of an addiction medical service during COVID-19. Med J Aust 2021 Sep 06;215(5):237-237.e1 [FREE Full text] [doi: 10.5694/mja2.51213] [Medline: 34365643]
51. Tofighi B, McNeely J, Walzer D, Fansiwala K, Demner A, Chaudhury C, et al. A telemedicine buprenorphine clinic to serve New York City: initial evaluation of the NYC Public Hospital System's initiative to expand treatment access during the COVID-19 pandemic. J Addict Med 2022;16(1):e40-e43 [FREE Full text] [doi: 10.1097/ADM.0000000000000809] [Medline: 33560696]
52. Aronowitz SV, Engel-Rebitzer E, Dolan A, Oyekanmi K, Mandell D, Meisel Z, et al. Telehealth for opioid use disorder treatment in low-barrier clinic settings: an exploration of clinician and staff perspectives. Harm Reduct J 2021 Nov 25;18(1):119 [FREE Full text] [doi: 10.1186/s12954-021-00572-7] [Medline: 34823538]
53. Tsai AC, Kiang MV, Barnett ML, Beletsky L, Keyes KM, McGinty EE, et al. Stigma as a fundamental hindrance to the United States opioid overdose crisis response. PLoS Med 2019 Nov 26;16(11):e1002969 [FREE Full text] [doi: 10.1371/journal.pmed.1002969] [Medline: 31770387]
54. Levander XA, Hoffman KA, McIlveen JW, McCarty D, Terashima JP, Korthuis PT. Rural opioid treatment program patient perspectives on take-home methadone policy changes during COVID-19: a qualitative thematic analysis. Addict Sci Clin Pract 2021 Dec 11;16(1):72 [FREE Full text] [doi: 10.1186/s13722-021-00281-3] [Medline: 34895346]
55. Treitler PC, Bowden CF, Lloyd J, Enich M, Nyaku AN, Crystal S. Perspectives of opioid use disorder treatment providers during COVID-19: adapting to flexibilities and sustaining reforms. J Subst Abuse Treat 2022 Jan;132:108514 [FREE Full text] [doi: 10.1016/j.jsat.2021.108514] [Medline: 34098210]
56. Ciccarone D. The triple wave epidemic: supply and demand drivers of the US opioid overdose crisis. Int J Drug Policy 2019 Sep;71:183-188 [FREE Full text] [doi: 10.1016/j.drugpo.2019.01.010] [Medline: 30718120]
57. Strike C, Watson TM. Losing the uphill battle? Emergent harm reduction interventions and barriers during the opioid overdose crisis in Canada. Int J Drug Policy 2019 Sep;71:178-182. [doi: 10.1016/j.drugpo.2019.02.005] [Medline: 30975595]
58. Høj SB, Jacka B, Minoyan N, Artenie AA, Bruneau J. Conceptualising access in the direct-acting antiviral era: an integrated framework to inform research and practice in HCV care for people who inject drugs. Int J Drug Policy 2019 Oct;72:11-23. [doi: 10.1016/j.drugpo.2019.04.001] [Medline: 31003825]

Abbreviations

CDSA: Controlled Drugs and Substances Act
CHUM: Centre hospitalier de l’Université de Montréal
CHUM-SMT: service de médecine des toxicomanies of the Centre hospitalier de l’Université de Montréal
CRISM: Canadian Research Initiative in Substance Misuse
OAT: opioid agonist treatment
OUD: opioid use disorder

Edited by A Mavragani, T Sanchez; submitted 04.05.22; peer-reviewed by J Kennel, L Narbey, S Xu, K Williams; comments to author 09.02.23; revised version received 12.05.23; accepted 06.06.23; published 26.07.23.

Please cite as:
Høj SB, de Montigny C, Chougar S, Léandre R, Beauchemin-Nadeau MÈ, Boyer-Legault G, Goyette A, Lamont SK, Bruneau J
Co-Constructing a Community-Based Telemedicine Program for People With Opioid Use Disorder During the COVID-19 Pandemic: Lessons Learned and Implications for Future Service Delivery
JMIR Public Health Surveill 2023;9:e39236
URL: https://publichealth.jmir.org/2023/1/e39236
doi:10.2196/39236
PMID:37494097

Viewpoint

Enabling Multicentric Participatory Disease Surveillance for Global Health Enhancement: Viewpoint on Global Flu View

Onicio Leal Neto[1,2], PhD; Daniela Paolotti[3], PhD; Craig Dalton[4], MMS; Sandra Carlson[4], BSc; Patipat Susumpow[5], BSc; Matt Parker[5], BSc, MPH; Polowat Phetra[5], BSc; Eric H Y Lau[6], PhD; Vittoria Colizza[7], PhD; Albert Jan van Hoek[8], PhD; Charlotte Kjelsø[9], MSc; John S Brownstein[10], PhD; Mark S Smolinski[1], MD

[1]Ending Pandemics, San Francisco, CA, United States
[2]Department of Computer Science, ETH Zurich, Zurich, Switzerland
[3]Institute for Scientific Interchange Foundation, Turin, Italy
[4]FluTracking, Newcastle, Australia
[5]Opendream, Bangkok, Thailand
[6]School of Public Health, University of Hong Kong, Hong Kong, China
[7]Pierre Louis Institute of Epidemiology and Public Health, INSERM, Sorbonne Université, Paris, France
[8]National Institute for Public Health and the Environment, Bilthoven, Netherlands
[9]Statens Serum Institute, Copenhagen, Denmark
[10]Boston Children Hospital, Harvard University, Boston, MA, United States

Corresponding Author:
Onicio Leal Neto, PhD
Ending Pandemics
870 Market Street
Suite 528
San Francisco, CA, 94102
United States
Phone: 1 5106938359
Email: onicio@gmail.com

Abstract

Participatory surveillance (PS) has been defined as the bidirectional process of transmitting and receiving data for action by directly engaging the target population. Often represented as self-reported symptoms directly from the public, PS can provide evidence of an emerging disease or concentration of symptoms in certain areas, potentially identifying signs of an early outbreak. The construction of sets of symptoms to represent various disease syndromes provides a mechanism for the early detection of multiple health threats. Global Flu View (GFV) is the first-ever system that merges influenza-like illness (ILI) data from more than 8 countries plus 1 region (Hong Kong) on 4 continents for global monitoring of this annual health threat. GFV provides a digital ecosystem for spatial and temporal visualization of syndromic aggregates compatible with ILI from the various systems currently participating in GFV in near real time, updated weekly. In 2018, the first prototype of a digital platform to combine data from several ILI PS programs was created. At that time, the priority was to have a digital environment that brought together different programs through an application program interface, providing a real time map of syndromic trends that could demonstrate where and when ILI was spreading in various regions of the globe. After 2 years running as an experimental model and incorporating feedback from partner programs, GFV was restructured to empower the community of public health practitioners, data scientists, and researchers by providing an open data channel among these contributors for sharing experiences across the network. GFV was redesigned to serve not only as a data hub but also as a dynamic knowledge network around participatory ILI surveillance by providing knowledge exchange among programs. Connectivity between existing PS systems enables a network of cooperation and collaboration with great potential for continuous public health impact. The exchange of knowledge within this network is not limited only to health professionals and researchers but also provides an opportunity for the general public to have an active voice in the collective construction of health settings. The focus on preparing the next generation of epidemiologists will be of great importance to scale innovative approaches like PS. GFV provides a useful example of the value of globally integrated PS data to help reduce the risks and damages of the next pandemic.

(JMIR Public Health Surveill 2023;9:e46644) doi:10.2196/46644

KEYWORDS

participatory surveillance; digital epidemiology; influenza-like illness; data transfer; surveillance; digital platform; Global Flu View; program; data sharing; public health; innovative; flu

Overview

Digital epidemiology has expanded significantly in recent years [1,2], encompassing not only passive digital disease detection [3], but also active crowdsourcing programs known as participatory surveillance (PS) [4-7]. PS has been defined as the bidirectional process of transmitting and receiving data for action by directly engaging the target population [8]. Often represented as self-reported symptoms directly from the public, PS can provide evidence of an emerging disease or concentration of symptoms in certain areas, potentially identifying signs of an early outbreak. The construction of sets of symptoms to represent various disease syndromes provides a mechanism for the early detection of multiple health threats [9-12]. In addition to sets of syndromes, PS benefits from the acquisition of data on nonsymptomatic individuals, providing some measure of a denominator to make calculations on prevalence, incidence, attack rate, among other metrics [13]. It is also possible to include questions that act as proxies for the identification of risk factors that may be associated with a health threat within a community. For example, questions about seeking health care after reporting symptoms may reveal the severity of the event for those individuals [11-13]. Or asking if a participant reporting into the system has recently traveled may indicate that symptoms are from a disease acquired from another location. Inquiring whether the reporter had recent contact with someone with similar symptoms might reveal a possible contagion network [13]. Additional questions about sociodemographics (eg, age, gender, income, and occupation) or even vaccination status can complement the syndromic information provided by users who participate in such programs to help inform at-risk groups [14]. Another layer that may be emphasized around the benefits of PS is its stability when something goes epidemic. While other surveillance systems rely on the health care system and are not scalable, web-based surveillance remain standing, even when the health care system is flooded. It is very important to have reasonable stable time trends to interpret what is going on. Furthermore, not so much the exact questions of the surveillance is important, but that the system is reasonably flexible including different questions, and that you can survey whatever is necessary, whether it is the uptake of a COVID-19–tracking app or self-tests, or vaccine-sentiment.

By these evidences it seems that PS is an affordable and effective proxy for public health departments to estimate, for example, influenza-like illness (ILI) and COVID-19–like illness numbers, despite the fact that clinical surveillance efforts are receiving far less data.

After the H1N1 pandemic emerged in 2009, PS for ILI became vital for monitoring this health challenge [6,15,16]. Since then, several platforms have expanded around the world seeking to use the participation of citizens in the systematic and periodic collection of symptoms related to influenza [9,17-26]. The numerous programs monitoring ILI have provided evidence regarding the acceptability, reliability, promptness, accuracy, and usefulness of this approach [6,20,22]. On the other hand, the aspects of data integration and interoperability between these systems for the purpose of comparing ILI activity across countries and regions remain a challenge. While several programs have similar objectives, the different governance and data compliance procedures placed operational hurdles to integrate and promote interoperability between the databases. A solution to realize the value of PS in monitoring global ILI activity was desperately needed.

In 2018, the first prototype of a digital platform to combine data from several ILI PS programs was created [27]. At that time, the priority was to have a digital environment that brought together different programs through an application program interface, providing a real-time map of syndromic trends that could demonstrate where and when ILI was spreading in various regions of the globe. After 2 years running as an experimental model and incorporating feedback from partner programs, Global Flu View (GFV) was restructured to empower the community of public health practitioners, data scientists, and researchers by providing an open data channel among these contributors for sharing experiences across the network. GFV [28] is redesigned to serve not only as a data hub, but also as a dynamic knowledge network around participatory ILI surveillance by providing knowledge exchange among programs [28]. The goal of this study is to serve as a viewpoint exploration, dissecting the challenges, solutions, and potential of PS applied in digital ecosystems like GFV.

Connecting ILI Programs Across the Globe in GFV

GFV is the first ever system that merges ILI data from more than 8 countries plus 1 region (Hong Kong) on 4 continents for global monitoring of this annual health threat. GFV provides a digital ecosystem for spatial and temporal visualization of syndromic aggregates compatible with ILI from the various systems currently participating in GFV in near real time (Table 1), updated weekly. In addition, GFV presents a list of features that support using PS as a data layer for enhanced ILI surveillance (Table 2). It is important to emphasize, however, that PS does not aim to replace traditional surveillance systems but rather to complement the information obtained from such systems and to generate actionable data.

Table 1. List of programs that are participating in GFV[a] as of December 2022.

Program's name	Territory coverage	Year of foundation	Description	Outcomes	Host institution
Outbreaks Near Me (formerly Flu Near You) [29]	United States and Canada	2011	A system that allows its users across North America to securely and anonymously self-report ILI[b] symptoms or to report feeling healthy.	• Since users generally report before visiting a health care provider, trends on the spread of ILI are often detected before local and national public health agencies do.	Boston Children's Hospital—Harvard University
FluTracking Australia [30]	Australia	2006	A web-based community-based acute respiratory illness surveillance system that is not biased by health seeking behavior, clinician testing practices, or differences in jurisdictional surveillance method.	• To contribute to community-level acute respiratory illness surveillance in Australia • To provide consistent surveillance of acute respiratory illness attack rates and testing across all jurisdictions and over time. • Provide year-to-year comparison of the timing, attack rates, and seriousness of acute respiratory illness in the community.	Hunter New England Local Health District, Hunter Medical Research Institute—The University of Newcastle
FluTracking Hong Kong [31]	Hong Kong	2021	A web-based health surveillance system used to detect the potential spread of influenza and COVID-19.	• To compare ILI syndrome proportions according vaccination status and seasonality, providing consistent monitoring of influenza surveillance activity.	School of Public Health—the University of Hong Kong
SickSense [32]	Thailand	2020	A network of digital volunteers who participate through anonymous reporting of their own health.	• This system can identify early signs of an emerging epidemic threat, be an assistant to find real time signs of illness, and engage through digital tools.	The Health Promotion Foundation and Open dream
InfluenzaNet[c] [33]	Italy, France, the Netherlands, and Denmark	2009	A web-based survey tool to conduct syndromic surveillance through self-reported symptoms volunteered by participants residing in the InfluenzaNet countries.	• The platforms collect demographic and risk-factor data from participants upon enrollment, capture weekly symptoms, and report analyzed surveillance results. This provides insights about the ILI activity.	Influweb.org, Covidnet.fr, Influmeter.dk, Infecieradar.nl

---

[a] GFV: Global Flu View.

[b] ILI: influenza-like illness.

[c] InfluenzaNet integration is still going on. This is a partial list of countries which enrolled in the platform.

Table 2. Overview of functionalities and main features of GFV.[a]

Feature	Description	Technical specification	Availability
Time-series report (Figure 1)	Temporal distribution of participant's reports and ILI[b] proxy defined by each program. In this feature,[c] it is also possible to filter by specific symptoms, vaccination status, gender, and age range.	Apex charts.JS is an open-source library for AngularJS that enables a smooth user experience loading time-series charts.	Publicly available. For registered partners, download of the line listing data is also possible in .CSV format.
Map (Figure 2)	Spatial distribution of participant's reports and ILI proxy defined by each program. In this view, it is possible to visualize in a (1) cluster view, (2) disaggregated reports by the minimum allowed level (eg, postcode centroids), or (3) heatmap with ILI cases.	For the use of clustering and heatmap view, Mapbox has been used.	Publicly available.
Resources	Library of recent peer-reviewed publications about PS.[d]	Automated scrapper that is connected with PubMed database, looking for keywords associated with PS. Before a paper is displayed on the resource page, a human input is necessary to accept or decline the suggestion. Once accepted, it populates a resource section.	Publicly available.
API[e] documentation (Figure 3)	The documentation for API requests of post-processed data points in the platform. The validation is based on a token generated by each program partner and the authorization is given by passing the key via header. The available requests for postprocessed data are: GET /api_map (returning data pinned to the map); GET /map_dates (returning the dates from the registered data); and GET /surveys (returning chart data from Reports section).	For the map data: Lat (number $double); Lng (number ($double)); Report_count (integer); ILI_count (integer); No_symptoms_count (integer); Some_symptoms_count; (integer); Postcode (string); Country (string); ILI_percentage (integer); No_symptoms_percentage (integer); Some_symptoms_percentage (integer); For the survey data: Match_count (integer); Percentage_count (number); Total_count (integer); Week (string)	Restricted to program partners with approved access to this feature.
Codelab (Figure 4)	Digital space for sharing codes, notes, and snippets across the community of program partners. In addition, this brings the possibility to specify the permission (whether creative commons or GPL[f]) and the language used. The codes are visible for all program partners that may use them for their projects or data analysis pipelines.	Besides the title, description, and tags, one can define the permission that the code has as well as the language the snippet was coded in.	Restricted to program partners with approved access to this feature.

[a] GFV: Global Flu View.

[b] ILI: influenza-like illness.

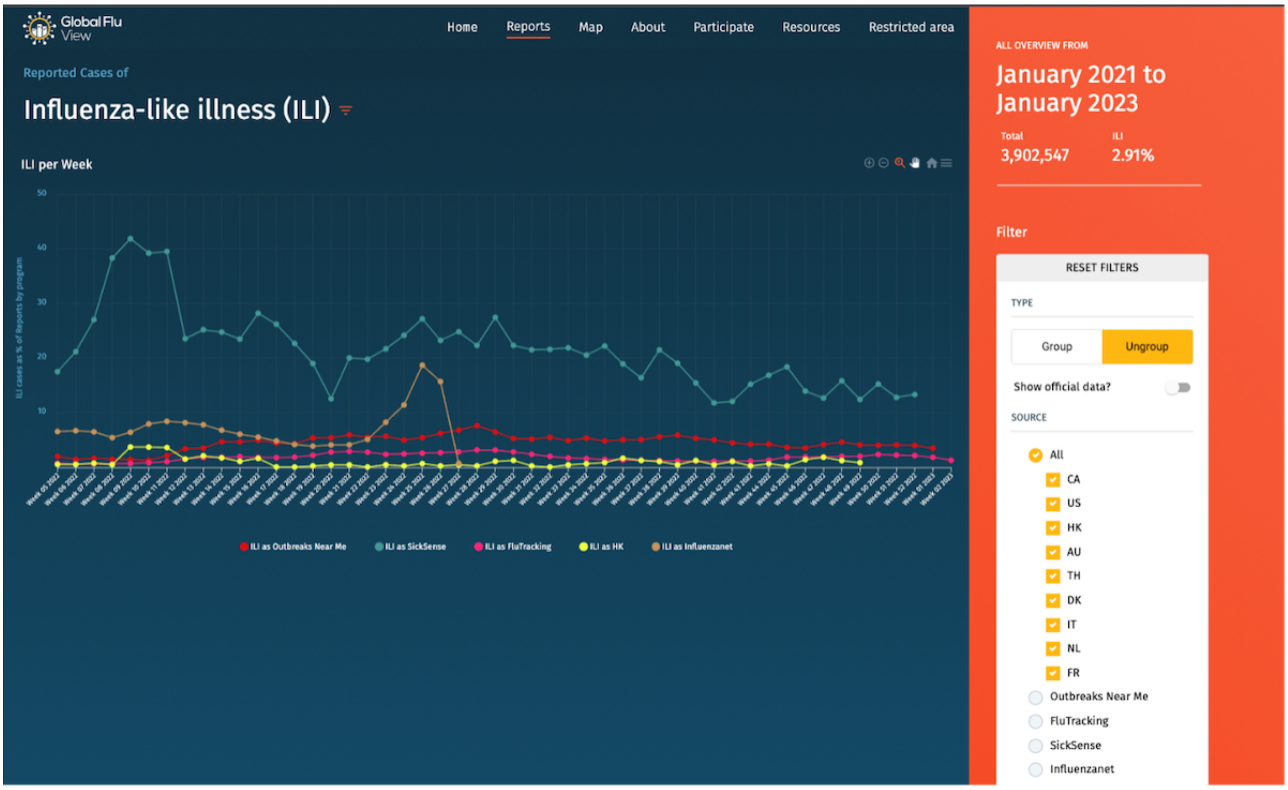
[c] When available by the program partner.

[d] PS: participatory surveillance.

[e] API: application program interface.

[f] GPL: General Public License.

Figure 1. Time-series reports view, showing the aggregated number of reports and proportion related to ILI reports. The filter functionality selects specific countries or programs as well as vaccination status, age range, and gender of participants. ILI: influenza-like illness.



ALL OVERVIEW FROM
January 2021 to January 2023
Total 3,902,547
ILI 2.91%

Show official data? ○

● All
☑ CA
☑ US
☑ HK
☑ AU
☑ TH
☑ DK
☑ IT
☑ NL
☑ FR
○ Outbreaks Near Me
○ FluTracking
○ SickSense
○ Influenzanet

Figure 2. Screenshot from the map view. It shows three different views that are available for users: (A) clustering view (aggregated mode based on the zoom level); (B) disaggregated view using the centroid of a zip or postal code; and (C) heatmap view for ILI reports from the filtered period. The map view also provides visualization of clustering view, aggregating reports based on the zoom level. Disaggregated level view is supported if the partner program has given permission. ILI: influenza-like illness.

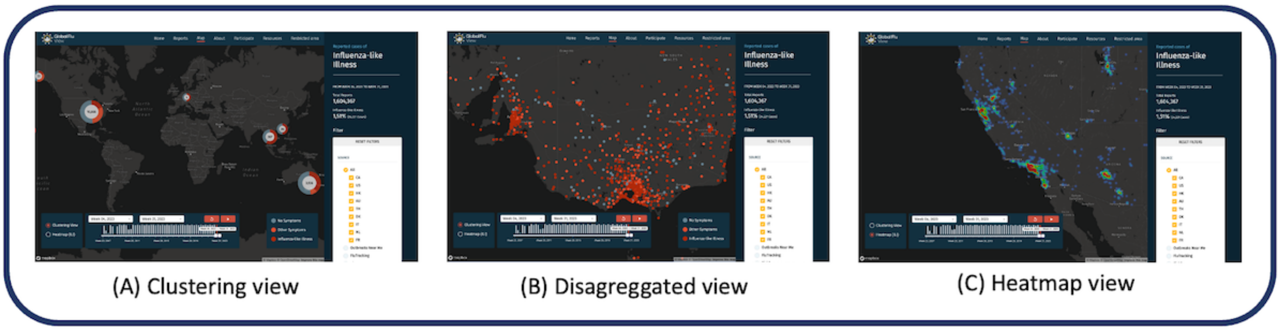
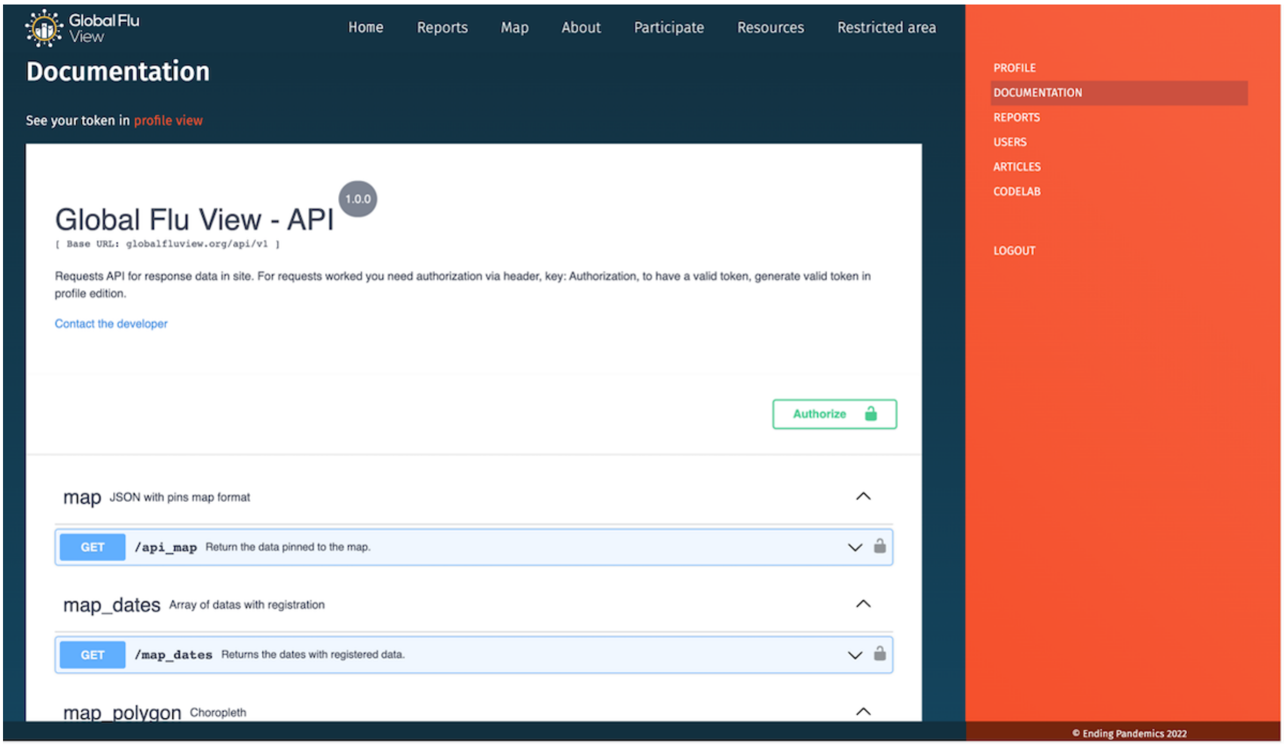

(A) Clustering view (B) Disaggregated view (C) Heatmap view

Figure 3. Documentation view of GFV API, providing program partners with all information needed to integrate GFV data into their systems, respecting the governance of each program and data sharing agreement. API: application program interface; GFV: Global Flu View.



Documentation

See your token in profile view

Global Flu View - API

[Base URL: globalfluview.org/api/v1]

Requests API for response data in site. For requests worked you need authorization via header, key: Authorization, to have a valid token, generate valid token in profile edition.

Contact the developer

map JSON with pins map format

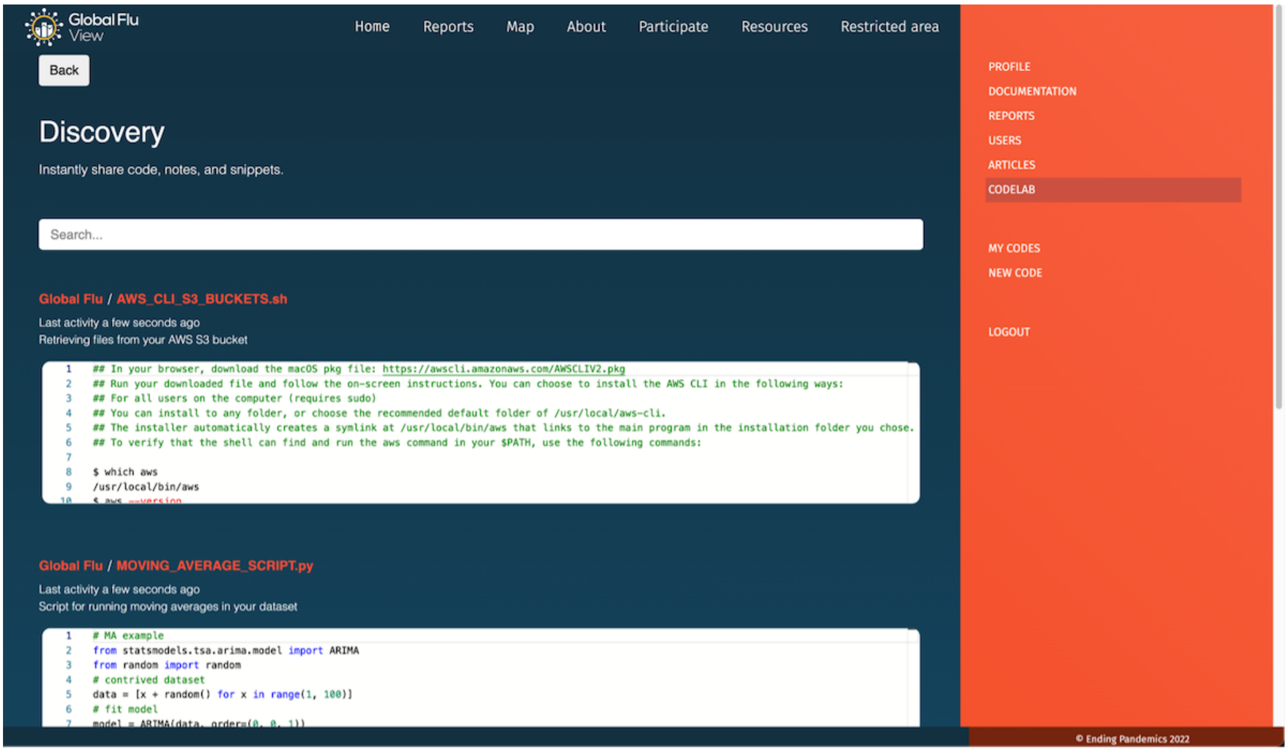
/api_map Return the data pinned to the map.

map_dates Array of datas with registration

/map_dates Returns the dates with registered data.

map_polygon Choropleth

Figure 4. Codelab is the tool into GFV that allows program partners to share among themselves notes, scripts, and codes about techniques and methods for PS data analysis. GFV: Global Flu View; PS: participatory surveillance.



Discovery

Instantly share code, notes, and snippets.

Global Flu / AWS_CLI_S3_BUCKETS.sh

Last activity a few seconds ago

Retrieving files from your AWS S3 bucket

```
## In your browser, download the macOS pkg file: https://awscli.amazonaws.com/AWSCLIV2.pkg
## Run your downloaded file and follow the on-screen instructions. You can choose to install the AWS CLI in the following ways:
## For all users on the computer (requires sudo)
## You can install to any folder, or choose the recommended default folder of /usr/local/aws-cli.
## The installer automatically creates a symlink at /usr/local/bin/aws that links to the main program in the installation folder you chose.
## To verify that the shell can find and run the aws command in your $PATH, use the following commands:

$ which aws
/usr/local/bin/aws
$ aws --version
```

Global Flu / MOVING_AVERAGE_SCRIPT.py

Last activity a few seconds ago

Script for running moving averages in your dataset

```
# MA example
from statsmodels.tsa.arima.model import ARIMA
from random import random
# contrived dataset
data = [x + random() for x in range(1, 100)]
# fit model
model = ARIMA(data, order=(0, 0, 1))
```

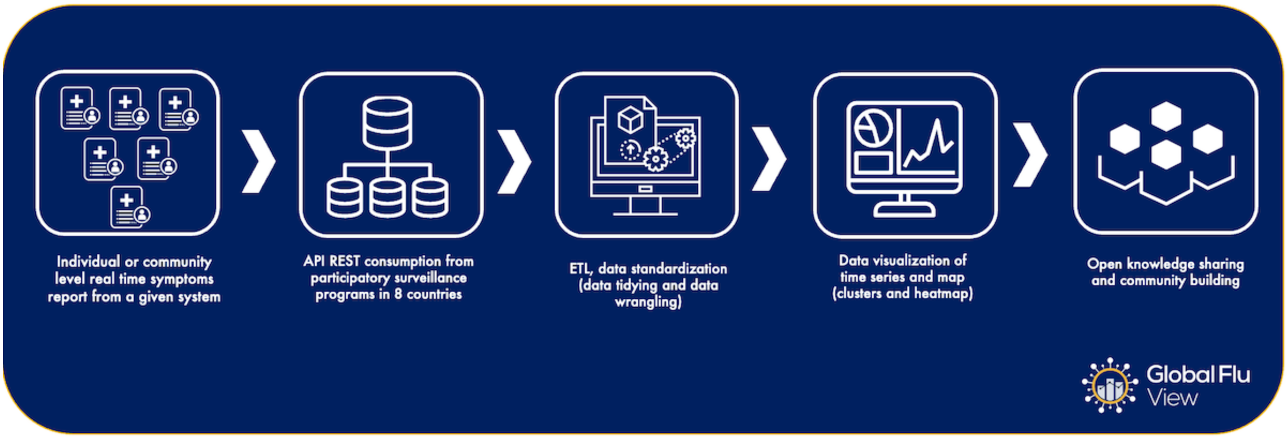
Understanding the GFV Ecosystem

GFV may be best described as an ecosystem with three different layers: (1) program partners, (2) community, and (3) outcomes. PS programs for ILI were informed about the purpose of GFV, and after accepting the data sharing agreement, became part of the core layer by contributing to the availability of real-time data. In the second layer, the "community" of potential consumers of this data can be described in three categories: (1) public health practitioners focused on disease surveillance, (2) data consumers focused on the use and integration of the data for third-party systems and platforms, and (3) researchers, who may use the data for comparative, predictive, or characterization of global epidemiological scenarios about ILI outbreaks as well as the development of forecasting and modeling techniques. Finally, the outcomes layer demonstrates the impact that these data can generate on the knowledge of the current and future ILI landscapes; and the potential of information sharing for constructing and strengthening of the ILI PS community. Pooling data from different regions and integrating them into GFV generates a pivotal tool for participatory surveillance. The collaborative network formed from this international data integration not only fosters shared learning but also strengthens ties between researchers and public health practitioners globally. This community aspect provides a robust response to ILI on an international scale.

The global perspective offered by GFV is especially valuable considering the nature of influenza epidemics. Tracking these patterns across various geographical regions provides critical insights into disease spread dynamics, contributing significantly to our understanding of these epidemics. Further, the collective data from diverse sources amplify the statistical power of analyses conducted using this platform. With a larger, more diverse data set at their disposal, researchers can draw more robust conclusions and gain key insights into influenza trends and variations across different settings. This enriched perspective provides a significant contribution to the global understanding and management of ILI.

The list of program partners is periodically updated, giving opportunities for ILI programs from anywhere to be part of this ecosystem. GFV is also open to the general public, who can use the potential of this tool to understand the importance of their contribution in reports to program partners in the collective construction of information on global health. Figure 5 describes the GFV’s workflow for a better understanding. We will further delve into the intricacies of community building, interoperability challenges, and data governance specific to participatory surveillance in subsequent sections. These important aspects help paint a comprehensive picture of our work and elucidate the significant contribution of GFV to the field of public health surveillance.

Figure 5. GFV’s workflow, describing the steps for data acquisition, API integration, data processing, data visualization, and knowledge sharing among program partners. API: application program interface; AWS: Amazon Web Services; ETL: Extract, Transform and Load ; GFV: Global Flu View.



Community Building and Knowledge Sharing

A community of PS enthusiasts can be connected with knowledge sharing in several ways. First, this community can provide a forum for practitioners to share their knowledge and expertise with each other. This can take the form of discussions and collaboration on specific projects or challenges, as well as the sharing of resources, such as research papers, technical manuals, and training materials. By sharing knowledge within the community, practitioners can learn from each other's experiences and can help to advance the field as a whole. Additionally, a community of health technology practitioners can also facilitate the sharing of knowledge with other stakeholders, such as policy makers, health care organizations, and the general public. For example, a community can organize conferences, workshops, and seminars where practitioners can present their research and findings to a broader audience. This can help to raise awareness of the latest developments in health technology and can encourage the dissemination of new ideas and best practices.

The GFV ecosystem aims to strengthen the use of surveillance data among professionals in various fields. Professionals such as public health practitioners, epidemiologists, data scientists, and software developers may use the GFV data in the routine of disease surveillance. The platform encourages contribution of data analysis techniques, methods, and processes to promote knowledge sharing among partners and ensures that data are used in a reproducible manner.

Interoperability Challenges: Overcoming Hurdles Based on Data Standards

Interoperability for participatory disease surveillance data can help to ensure that data from different systems can be easily exchanged and integrated, leading to more effective and efficient processes. However, several challenges must be considered to achieve this goal, including the following: (1) lack of standardization: different PS programs may collect different types of data, at different granularities, use different data formats and structures, which can make it difficult to exchange data. Moreover, without standardized data definitions and formats, it can be difficult to ensure that data from different PS programs are comparable; (2) complexity of data integration: integrating data from diverse PS programs systems can be a complex and time-consuming process, especially if the systems are not designed to interoperate; (3) limited technical capacity: some PS systems may not have the technical capacity or expertise to integrate data from other systems; and (4) privacy and security concerns: ensuring the privacy and security of sensitive health or disease data is a key concern when integrating data from distinct surveillance systems and are subject to different national or international regulatory frameworks (eg, European Union General Data Protection Regulation).

To overcome these challenges, it is necessary to adopt data standards that provide a common framework for the exchange and integration of data from different surveillance systems. Data standards offer a set of rules and guidelines for defining data elements and formats, which can help to ensure that data from different systems are comparable and can be easily integrated.

In developing GFV, the number of global platforms that use secondary epidemiologic data was considered. Documenting the processes and end points for accessing data on the platform is fundamental to the interoperability of this data with other data streams to allow connections for integration with GFV. Using an API to connect the various systems with GFV was deemed the most efficient mechanism.

Data Governance: Getting Legal Agreements for a Multicentric PS Ecosystem

A data sharing agreement is required for any partner that provides its data to GFV (except for those programs which make the data publicly available). This agreement covers core principles such as (1) preserving data privacy by sharing deidentified data, (2) maintaining data integrity, and (3) respecting the governance structure and compliance required by each program partner.

Data governance is an essential aspect of PS systems to safeguard the integrity, quality, and security of data collected. It includes policies, processes, and practices for data collection, management, and use, with defined roles and responsibilities for data quality assurance and security. Clear and transparent policies and practices can ensure that participants feel confident that their data will be used responsibly.

Looking to the Future

Enabling PS for Digital Public Health Intelligence

The popularization of social media, added to the wave of "big data," creates the opportunity for crowdsourcing participatory disease surveillance in GFV to expand the understanding of ILI and its impact across the globe [19,20].

Several initiatives are integrating diverse data streams to increasingly provide timely information to detect public health events. Low- and middle-income countries, in particular, are greatly benefited by these approaches. Examples of these initiatives include the Integrated Disease Surveillance and Response framework in sub-Saharan Africa. For 20 years it has demonstrated the importance of improving the detection and response of public health threats by combining data from diverse systems and laboratories [34]. In China, another model framework integrates multichannel surveillance, data analysis for early warning, and a specific strategy for sustainability of the surveillance system [35]. Within the scope of PS, it is possible to see the integration of laboratory data for the monitoring of SARS-CoV-2 antibodies seroconversion [36-38]. PS has historically also used data from social media to leverage its sensitivity and detection capability [39,40]. In Netherlands, a strategy to combine PS reports plus self-tests given free of charge and centralized testing of nose or throat samples has been performed, showing another angle of empowering public health intelligence by these approaches [41]. The previous examples show how different ecosystems across the globe are ready to integrate diverse data streams, where GFV may be one of these, enabling a feasible view for PS.

During the first years of the COVID-19 pandemic, the value of other types of data, from wearables, wastewater, and mobility through cellular data for early identification of risks, became evident [42-45]. Approaches like this also show the ability of omnichannel surveillance for other infectious diseases with pandemic potential, such as a novel strain of influenza. Indeed, challenges of integration and interoperability between these data sources will exist and demand that the next generation of epidemiologists be familiar with the format and structure of each data source. The assessment for data standards for collecting comparable data will be crucial to guarantee the unity of such different systems.

Cross-Engagement

Cross-engagement is a term inherited from the marketing industry, which describes the use of multiple marketing channels to seamlessly leverage retention of a target audience [46]. Collaborative platforms can benefit from this concept, as the retention and engagement of users in PS strategies has always been a point of friction. Maintaining engagement and participation in these programs can be challenging, as individuals and communities may have other priorities or may lose interest over time. With cross engagement, the user is motivated on multiple fronts for multiple purposes. The GFV platform provides cross-engagement as user data are seen at a local or national level by the partner platform and at an international level displaying trends for multicentric comparison between countries and continents. The empirical evidence is that people become further incentivized by the knowledge that their data matter both locally and globally.

Programs for monitoring and combating chronic diseases [47] can serve as valuable models for infectious disease PS programs, particularly for influenza, which is a key focus of the GFV platform. These programs can provide valuable insights and lessons learned that can be applied to the design and implementation of infectious disease surveillance programs, helping to ensure that these programs are effective and efficient. By leveraging the experience and expertise of chronic disease surveillance programs, it is possible to learn how to improve the engagement and retention performance of participatory infectious disease surveillance programs [48-50].

Smart Health Communities

Smart Health Communities has been another important trend to consider in the future of disease surveillance [51]. The core features of this approach are (1) empowering individuals to proactively manage one's own health and well-being, (2) fostering a sense of community and belonging, (3) using digital technologies and behavioral sciences, (4) using data for substantial improvements in health outcomes, and (5) creating innovative ecosystems [52].

In this scenario, studies demonstrate how these programs reach the status of social contagion to achieve a collective good [53-55]. This type of social influence can be one of the ways to increase user participation in PS platforms, considerably increasing the potential for using the data that circulates in this environment [53,54]. The parallelism between geographic smart health communities and digital smart health communities enables the possibilities of using the metaverse for PS [55]. By these examples, having a web-based hub that serve as stage for smart health communities look at the data seems realistic and needed. With the open access that GFV promotes through temporal and spatial data, this becomes a concrete example for future interactions in smart health communities.

Opportunities for the Next Generation of Digital Epidemiologists

Given these perspectives, the importance of the next-generation of digital epidemiologists actively participating in the ongoing digital transformation of public health surveillance cannot be overstated. Their role in this rapidly evolving landscape goes beyond merely applying existing tools and systems; they must also acquire a comprehensive understanding of this emerging field. In order to cultivate a new breed of professionals capable of steering this digital revolution, the inclusion of data science, software engineering, mathematical modeling, and data design in academic and technical training programs is vital. This multidisciplinary approach equips future digital epidemiologists with the broad spectrum of skills they require to be self-reliant and innovative in their roles. While collaboration and cross-pollination of ideas are indeed essential in this multifaceted field, it is equally important for these professionals to have the capability to make independent, informed decisions. The interdependence inherent in such a diverse field should not lead to an overreliance on others. These professionals will need to be well-versed in all aspects of digital epidemiology, from disease surveillance to the intricacies of digital public health. GFV is pioneering in this educational journey by providing a platform that fosters collaboration, knowledge exchange, and the cultivation of skills. The GFV Codelab, for instance, is an exemplary feature that demonstrates how digital platforms can be harnessed as digital learning environments. As we navigate this new digital landscape, it is incumbent upon us to ensure that the next generation of professionals is well-equipped to lead the field of digital epidemiology into the future.

Further, the endorsement of participatory surveillance strategies by major health organizations, such as the World Health Organization [56], underscores the pivotal role that this approach plays in global health. It signals an expectation for the next generation of digital epidemiologists to not only understand these approaches but to implement them effectively and innovatively.

Limitations

GFV has several limitations. First, the data in GFV are dependent on the quality of its partners' data. Since PS programs rely on the active involvement of individuals and communities, GFV data are impacted by this reality that not all individuals and communities may be willing or able to participate in these programs. This can lead to incomplete or biased data, which can impact the accuracy and usefulness of the information collected through these programs. Second, data quality and integrity are an issue due to the characteristics of self-reported data, which can be subject to errors, biases, and inconsistencies. Third, the various programs contributing data to GFV often require significant resources, including funding, staff time, and technology. These resources may be limited, which can impact the ability of these programs to effectively collect data on an ongoing basis and share it with GFV.

Conclusions

Connectivity between existing PS systems enables a network of cooperation and collaboration with great potential for continuous public health impact. Thus, it is crucial for professionals engaged in ILI studies, educators, students, and public health practitioners to engage with digital ecosystems. It is an opportunity to make use of these open data channels and contribute to knowledge, and help to enhance the depth and breadth of ILI surveillance. Moreover, it offers a platform for learning, interaction, and collaboration, fostering a more integrated and proactive approach toward public health challenges. The exchange of knowledge within this network is not limited only to health professionals and researchers, but also provides an opportunity for the general public to have an active voice in the collective construction of health settings. The focus on preparing the next generation of epidemiologists will be of great importance to scale innovative approaches like PS. GFV provides a useful example of the value of globally integrated PS data to help reduce the risks and damages of the next pandemic.

Acknowledgments

We would like to acknowledge all participants or users from all programs listed above. We acknowledge Nomita Divi and Jay Atanda for the early contributions in the coordination of Global Flu View program. The authors would also like to thank Rodrigo

Paiva and Bruno Barros, who are the software engineers for Global Flu View. The authors acknowledge all FluTracking team members and the tech leader Stephen Clarke.

Conflicts of Interest

None declared.

References

1. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. PLoS Comput Biol 2012;8(7):e1002616 [FREE Full text] [doi: 10.1371/journal.pcbi.1002616] [Medline: 22844241]
2. Salathé M. Digital epidemiology: what is it, and where is it going? Life Sci Soc Policy 2018;14(1):1 [FREE Full text] [doi: 10.1186/s40504-017-0065-7] [Medline: 29302758]
3. Denecke K. An ethical assessment model for digital disease detection technologies. Life Sci Soc Policy 2017;13(1):16 [FREE Full text] [doi: 10.1186/s40504-017-0062-x] [Medline: 28929347]
4. Kostkova P. Disease surveillance data sharing for public health: the next ethical frontiers. Life Sci Soc Policy 2018;14(1):16 [FREE Full text] [doi: 10.1186/s40504-018-0078-x] [Medline: 29971516]
5. Jost CC, Mariner JC, Roeder PL, Sawitri E, Macgregor-Skinner GJ. Participatory epidemiology in disease surveillance and research. Rev Sci Tech 2007;26(3):537-549 [FREE Full text] [Medline: 18293603]
6. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. Clin Microbiol Infect 2014;20(1):17-21 [FREE Full text] [doi: 10.1111/1469-0691.12477] [Medline: 24350723]
7. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg Themes Epidemiol 2014;11:7 [FREE Full text] [doi: 10.1186/1742-7622-11-7] [Medline: 24991229]
8. McNeil C, Verlander S, Divi N, Smolinski M. The landscape of participatory surveillance systems across the one health spectrum: systematic review. JMIR Public Health Surveill 2022;8(8):e38551 [FREE Full text] [doi: 10.2196/38551] [Medline: 35930345]
9. Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, et al. Participatory syndromic surveillance of influenza in Europe. J Infect Dis 2016;214(Suppl 4):S386-S392 [FREE Full text] [doi: 10.1093/infdis/jiw280] [Medline: 28830105]
10. Kalimeri K, Delfino M, Cattuto C, Perrotta D, Colizza V, Guerrisi C, et al. Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms. PLoS Comput Biol 2019;15(4):e1006173 [FREE Full text] [doi: 10.1371/journal.pcbi.1006173] [Medline: 30958817]
11. Neto OL, Cruz O, Albuquerque J, de Sousa MN, Smolinski M, Cesse E, et al. Participatory surveillance based on crowdsourcing during the Rio 2016 Olympic Games using the guardians of health platform: descriptive study. JMIR Public Health Surveill 2020;6(2):e16119 [FREE Full text] [doi: 10.2196/16119] [Medline: 32254042]
12. Patterson-Lomba O, Van Noort S, Cowling BJ, Wallinga J, Gomes MGM, Lipsitch M, et al. Utilizing syndromic surveillance data for estimating levels of influenza circulation. Am J Epidemiol 2014;179(11):1394-1401 [FREE Full text] [doi: 10.1093/aje/kwu061] [Medline: 24748609]
13. Neto OL, Dimech GS, Libel M, de Souza WV, Cesse E, Smolinski M, et al. Saúde na Copa: the world's first application of participatory surveillance for a mass gathering at FIFA World Cup 2014, Brazil. JMIR Public Health Surveill 2017;3(2):e26 [FREE Full text] [doi: 10.2196/publichealth.7313] [Medline: 28473308]
14. Cantarelli P, Debin M, Turbelin C, Poletto C, Blanchon T, Falchi A, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. BMC Public Health 2014;14:984 [FREE Full text] [doi: 10.1186/1471-2458-14-984] [Medline: 25240865]
15. Carlson SJ, Dalton CB, Durrheim DN, Fejsa J. Online flutracking survey of influenza-like illness during pandemic (H1N1) 2009, Australia. Emerg Infect Dis 2010;16(12):1960-1962 [FREE Full text] [doi: 10.3201/eid1612.100935] [Medline: 21122231]
16. Dawood FS, Dalton CB, Durrheim DN, Hope KG. Rates of hospitalisation for acute respiratory illness and the emergence of pandemic (H1N1) 2009 virus in the Hunter New England Area Health Service. Med J Aust 2009;191(10):573-574. [doi: 10.5694/j.1326-5377.2009.tb03318.x] [Medline: 19912093]
17. Chunara R, Aman S, Smolinski M, Brownstein JS. Flu near you: an online self-reported influenza surveillance system in the USA. Online J Public Health Inform 2013:e133 [FREE Full text] [doi: 10.5210/ojphi.v5i1.4456]
18. Dalton CB, Carlson SJ, Butler MT, Elvidge E, Durrheim DN. Building influenza surveillance pyramids in near real time, Australia. Emerg Infect Dis 2013;19(11):1863-1865 [FREE Full text] [doi: 10.3201/eid1911.121878] [Medline: 24207165]
19. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats. JMIR Public Health Surveill 2017;3(4):e62 [FREE Full text] [doi: 10.2196/publichealth.7540] [Medline: 29021131]

20. Dalton C, Carlson S, Butler M, Cassano D, Clarke S, Fejsa J, et al. Insights from flutracking: thirteen tips to growing a web-based participatory surveillance system. JMIR Public Health Surveill 2017;3(3):e48 [FREE Full text] [doi: 10.2196/publichealth.7333] [Medline: 28818817]
21. Kim M, Yune S, Chang S, Jung Y, Sa SO, Han HW. The fever coach mobile app for participatory influenza surveillance in children: usability study. JMIR mHealth uHealth 2019;7(10):e14276 [FREE Full text] [doi: 10.2196/14276] [Medline: 31625946]
22. Lee L, Desroches M, Mukhi S, Bancej C. FluWatchers: evaluation of a crowdsourced influenza-like illness surveillance application for Canadian influenza seasons 2015-2016 to 2018-2019. Can Commun Dis Rep 2021;47(9):357-363 [FREE Full text] [doi: 10.14745/ccdr.v47i09a02] [Medline: 34650332]
23. Lwin MO, Lu J, Sheldenkar A, Panchapakesan C, Tan YR, Yap P, et al. Effectiveness of a mobile-based influenza-like illness surveillance system (FluMob) among health care workers: longitudinal study. JMIR mHealth uHealth 2020;8(12):e19712 [FREE Full text] [doi: 10.2196/19712] [Medline: 33284126]
24. Perrotta D, Bella A, Rizzo C, Paolotti D. Participatory online surveillance as a supplementary tool to sentinel doctors for influenza-like illness surveillance in Italy. PLoS One 2017;12(1):e0169801 [FREE Full text] [doi: 10.1371/journal.pone.0169801] [Medline: 28076411]
25. Pini A, Merk H, Carnahan A, Galanis I, Van Straten E, Danis K, et al. High added value of a population-based participatory surveillance system for community acute gastrointestinal, respiratory and influenza-like illnesses in Sweden, 2013-2014 using the web. Epidemiol Infect 2017;145(6):1193-1202 [FREE Full text] [doi: 10.1017/S0950268816003290] [Medline: 28137317]
26. Prieto JT, Jara JH, Alvis JP, Furlan LR, Murray CT, Garcia J, et al. Will participatory syndromic surveillance work in Latin America? Piloting a mobile approach to crowdsource influenza-like illness data in Guatemala. JMIR Public Health Surveill 2017;3(4):e87 [FREE Full text] [doi: 10.2196/publichealth.8610] [Medline: 29138128]
27. Crawley AW, Paolotti D, Dalton C, Brownstein J, Smolinski M. Global flu view: a platform to connect crowdsourced disease surveillance around the world. Int J Infect Dis 2019;79:7 [FREE Full text] [doi: 10.1016/j.ijid.2018.11.036]
28. Global Flu View. URL: https://globalfluview.org [accessed 2023-08-03]
29. Outbreaks Near Me. URL: https://outbreaksnearme.org/us/en-US/ [accessed 2023-08-03]
30. Flutracking Australia. URL: https://info.flutracking.net/ [accessed 2023-08-03]
31. Flutracking Hong Kong. URL: https://www.flutracking.sph.hku.hk/ [accessed 2023-08-03]
32. Sicksense—Digital Disease Detection. URL: https://www.sicksense.org/ [accessed 2023-08-03]
33. InfluenzaNet analytics. Influenzanet.Info. URL: https://influenzanet.info/#page/info [accessed 2023-08-03]
34. Mremi IR, George J, Rumisha SF, Sindato C, Kimera SI, Mboera LEG. Twenty years of integrated disease surveillance and response in Sub-Saharan Africa: challenges and opportunities for effective management of infectious disease epidemics. One Health Outlook 2021;3(1):22 [FREE Full text] [doi: 10.1186/s42522-021-00052-9] [Medline: 34749835]
35. Xu L, Zhou C, Luo S, Chan DK, McLaws ML, Liang W. Modernising infectious disease surveillance and an early-warning system: the need for China's action. Lancet Reg Health West Pac 2022;23:100485 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100485] [Medline: 35685717]
36. Kohler P, Güsewell S, Seneghini M, Egger T, Leal O, Brucher A, et al. Impact of baseline SARS-CoV-2 antibody status on syndromic surveillance and the risk of subsequent COVID-19-a prospective multicenter cohort study. BMC Med 2021;19(1):270 [FREE Full text] [doi: 10.1186/s12916-021-02144-9] [Medline: 34649585]
37. Leal-Neto O, Egger T, Schlegel M, Flury D, Sumer J, Albrich W, et al. Digital SARS-CoV-2 detection among hospital employees: participatory surveillance study. JMIR Public Health Surveill 2021;7(11):e33576 [FREE Full text] [doi: 10.2196/33576] [Medline: 34727046]
38. Strahm C, Seneghini M, Güsewell S, Egger T, Leal-Neto O, Brucher A, et al. Symptoms compatible with long coronavirus disease (COVID) in healthcare workers with and without Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infection-results of a prospective multicenter cohort. Clin Infect Dis 2022;75(1):e1011-e1019 [FREE Full text] [doi: 10.1093/cid/ciac054] [Medline: 35090015]
39. Brownstein JS, Chu S, Marathe A, Marathe MV, Nguyen AT, Paolotti D, et al. Combining participatory influenza surveillance with modeling and forecasting: three alternative approaches. JMIR Public Health Surveill 2017;3(4):e83 [FREE Full text] [doi: 10.2196/publichealth.7344] [Medline: 29092812]
40. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. PLoS Comput Biol 2015;11(10):e1004513 [FREE Full text] [doi: 10.1371/journal.pcbi.1004513] [Medline: 26513245]
41. RIVM De Zorg Voor Morgen Begint Vandaag. URL: https://www.rivm.nl/griep-griepprik/feiten-en-cijfers [accessed 2023-08-03]
42. Risch M, Grossmann K, Aeschbacher S, Weideli OC, Kovac M, Pereira F, COVID-19 remote early detection (COVID-RED) consortium. Investigation of the use of a sensor bracelet for the presymptomatic detection of changes in physiological parameters related to COVID-19: an interim analysis of a prospective cohort study (COVI-GAPP). BMJ Open 2022;12(6):e058274 [FREE Full text] [doi: 10.1136/bmjopen-2021-058274] [Medline: 35728900]

43. Wolfe MK, Duong D, Bakker KM, Ammerman M, Mortenson L, Hughes B, et al. Wastewater-based detection of two influenza outbreaks. Environ Sci Technol Lett 2022;9(8):687-692 [FREE Full text] [doi: 10.1021/acs.estlett.2c00350]
44. Bagutti C, Hug MA, Heim P, Pekerman LM, Hampe EI, Hübner P, et al. Wastewater monitoring of SARS-CoV-2 shows high correlation with COVID-19 case numbers and allowed early detection of the first confirmed B.1.1.529 infection in Switzerland: results of an observational surveillance study. Swiss Med Wkly 2022;152:w30202 [FREE Full text] [doi: 10.4414/smw.2022.w30202] [Medline: 35822578]
45. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv 2020;6(23):eabc0764 [FREE Full text] [doi: 10.1126/sciadv.abc0764] [Medline: 32548274]
46. Why cross-channel engagement matters for your business. Braze. 2020. URL: https://www.braze.com/resources/articles/cross-channel-engagement-matters [accessed 2023-08-03]
47. NDPP policy and billing guidelines. New York State Department of Health. URL: https://health.ny.gov/health_care/medicaid/redesign/ndpp/policy-billing_guide.htm [accessed 2023-08-03]
48. Liu S, Han J, Puyal EL, Kontaxis S, Sun S, Locatelli P, et al. Fitbeat: COVID-19 estimation based on wristband heart rate using a contrastive convolutional auto-encoder. Pattern Recognit 2022;123:108403 [FREE Full text] [doi: 10.1016/j.patcog.2021.108403] [Medline: 34720200]
49. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. Nat Biomed Eng 2020;4(12):1208-1220 [FREE Full text] [doi: 10.1038/s41551-020-00640-6] [Medline: 33208926]
50. Natarajan A, Su HW, Heneghan C. Assessment of physiological signs associated with COVID-19 measured using wearable devices. NPJ Digit Med 2020;3(1):156 [FREE Full text] [doi: 10.1038/s41746-020-00363-7] [Medline: 33299095]
51. Ng F, Briggs D, Liu Y. Smart health communities: from sick care to health care. In: Fong BYF, Wong MCS, editors. The Routledge Handbook of Public Health and the Community. Abingdon, Oxon: Routledge; 2021:302-312.
52. Smart health communities. Deloitte. 2022. URL: https://www2.deloitte.com/global/en/pages/public-sector/articles/urban-future-with-a-purpose/smart-health-communities.html [accessed 2023-08-03]
53. Coviello L, Sohn Y, Kramer ADI, Marlow C, Franceschetti M, Christakis NA, et al. Detecting emotional contagion in massive social networks. PLoS One 2014;9(3):e90315 [FREE Full text] [doi: 10.1371/journal.pone.0090315] [Medline: 24621792]
54. Aral S, Nicolaides C. Exercise contagion in a global social network. Nat Commun 2017;8:14753 [FREE Full text] [doi: 10.1038/ncomms14753] [Medline: 28418379]
55. Garavand A, Aslani N. Metaverse phenomenon and its impact on health: a scoping review. Inform Med Unlocked 2022;32:101029 [FREE Full text] [doi: 10.1016/j.imu.2022.101029]
56. Best practices for the design, implementation, analysis and reporting of participatory surveillance for influenza-like illness. World Health Organization. URL: https://cdn.who.int/media/docs/default-source/influenza/participatory_surveillance_guidance_2022_10_18.pdf?sfvrsn=34905d39_1 [accessed 2023-08-03]

Abbreviations

GFV: Global Flu View
ILI: influenza-like illness
PS: participatory surveillance

Edited by A Mavragani, T Sanchez; submitted 20.02.23; peer-reviewed by K Baltrusaitis, Y Katayama, V von Wyl; comments to author 01.06.23; revised version received 21.06.23; accepted 25.07.23; published 01.09.23.

Please cite as:
Leal Neto O, Paolotti D, Dalton C, Carlson S, Susumpow P, Parker M, Phetra P, Lau EHY, Colizza V, Jan van Hoek A, Kjelsø C, Brownstein JS, Smolinski MS
Enabling Multicentric Participatory Disease Surveillance for Global Health Enhancement: Viewpoint on Global Flu View
JMIR Public Health Surveill 2023;9:e46644
URL: https://publichealth.jmir.org/2023/1/e46644
doi:10.2196/46644
PMID:37490846

Viewpoint

Reconsidering the Use of Population Health Surveys for Monitoring of Mental Health

Jorge Arias de la Torre[1,2,3], PhD; Gemma Vilagut[2,4], PhD; Amy Ronaldson[5], PhD; Ioannis Bakolis[5], PhD; Alex Dregan[5], PhD; Fernando Navarro-Mateu[2], PhD; Katherine Pérez[2,6,7], PhD; Anna Szücs[8], MD; Xavier Bartoll-Roca[6,7], PhD; Antonio J Molina[2,3], PhD; Matilde Elices[9], PhD; Víctor Pérez-Solá[9,10], Prof Dr, PhD; Vicente Martín[2,3], Prof Dr, PhD; Antoni Serrano-Blanco[2,11], PhD; Jose M Valderas[12*], Prof Dr, PhD; Jordi Alonso[2,4,13*], Prof Dr, PhD

[1]Care in Long Term Conditions Research Division, King's College London, London, United Kingdom
[2]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain
[3]Department of Biomedical Sciences, Universidad de León, León, Spain
[4]Health Services Research Group, Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain
[5]Institute of Psychiatry, Psychology, and Neuroscience (IoPPN), King's College London, London, United Kingdom
[6]Agència de Salut Pública de Barcelona (ASPB), Barcelona, Spain
[7]Institut d'Investigació Biomèdica Sant Pau (IIB SANT PAU), Barcelona, Spain
[8]Department of Medicine, National University of Singapore (NUS), Singapore, Singapore
[9]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain
[10]Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain
[11]Institut de Recerca Sant Joan de Déu, Parc Sanitari Sant Joan de Déu, Barcelona, Spain
[12]University of Exeter, Exeter, United Kingdom
[13]Department of Medical and Life Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain
[*]these authors contributed equally

Corresponding Author:
Jose M Valderas, Prof Dr, PhD
University of Exeter
Magdalen Road, Smeall Building
Exeter, EX12LU
United Kingdom
Phone: 44 6772 3874
Email: jmvalderas@nus.edu.sg

Abstract

Monitoring of the mental health status of the population and assessment of its determinants are 2 of the most relevant pillars of public mental health, and data from population health surveys could be instrumental to support them. Although these surveys could be an important and suitable resource for these purposes, due to different limitations and challenges, they are often relegated to the background behind other data sources, such as electronic health records. These limitations and challenges include those related to measurement properties and cross-cultural validity of the tools used for the assessment of mental disorders, their degree of representativeness, and possible difficulties in the linkage with other data sources. Successfully addressing these limitations could significantly increase the potential of health surveys in the monitoring of mental disorders and ultimately maximize the impact of the relevant policies to reduce their burden at the population level. The widespread use of data from population health surveys, ideally linked to electronic health records data, would enhance the quality of the information available for research, public mental health decision-making, and ultimately addressing the growing burden of mental disorders.

(JMIR Public Health Surveill 2023;9:e48138) doi:10.2196/48138

KEYWORDS

mental health; public heath; burden; health surveys; monitoring; status; data collection; electronic health record; challenges; assessment tool; population; population health survey

Introduction

In line with the motto of World Mental Health Day on October 10, 2022, “Make Mental Health and Well-Being for All a Global Priority,” mental health has become a public health and research priority in different countries worldwide. In the United States and the United Kingdom, for example, the White House and the UK Health Security Agency have included mental health as one of its corporate priority areas on their health and research strategies [1,2]. Such prioritization processes depend on the availability of data and on the quality of the evidence for a specific context and period. However, not all the data are considered equally relevant for this prioritization, and the evidence derived from population health surveys is often considered to be of lower quality than the evidence from other data sources, such as electronic health records (EHRs).

Population health surveys are observational studies designed for research and public health purposes, aimed to obtain a representation of the distribution of different health conditions and the association of this distribution with sociodemographic, lifestyle, and other factors in well-defined populations. Since the second half of the previous century, these surveys have been carried out in different countries worldwide (eg, the National Health Interview Survey [NHIS] in the United States since 1957 [3] and the Health Survey for England in the United Kingdom since 1991 [4]). Due to their population-based nature (ie, based on representative as opposed to convenience sampling), data from these surveys may constitute a helpful resource to determine the burden of different mental health problems in the community, to identify vulnerable populations or population groups, and to inform the planning of health care resources and the development of health policies and targeted preventive measures [5]. Additionally, it must be highlighted that individual-level data from multiple health surveys worldwide, particularly from middle- and high-income countries, are publicly available for researchers to access and use (eg, data from the NHIS [3] or the National Health and Nutrition Examination Survey in the United States [6] and data from the European Health Interview Survey [EHIS] [7] or the European Social survey in Europe [8]). Some examples of the use of data from population health surveys for these purposes include the use of the data from the World Mental Health Surveys as a primary source to calculate the prevalence of major depression internationally [9], the identification of differences in depression among religious groups in Europe using data from the European Social Survey [10], and monitoring of the mental health and suicide prevention reforms in Australia [11]. While the use of data from population health surveys has been successfully implemented for these purposes, these data need to meet certain requirements, including the use of valid and reliable measures for the assessment of mental disorders and the selection of representative samples.

In this context, we propose that extending the use of data from population health surveys would have a considerable impact on improving the evidence base of public mental health decision-making and, hence, could be instrumental in reducing the growing burden of mental disorders in the community.

Public Mental Health and the Use of Population Health Surveys

Monitoring of the mental health status of the population and assessment of its determinants (eg, sociodemographic factors, lifestyle habits, and use of health services) are 2 of the pillars of public mental health [12-14], and the use of harmonized data from multiple sources could be instrumental to support them.

If the quality of the data could be guaranteed, the monitoring of mental health at the population level should be based on all the available data. For example, the Global Burden of Disease study integrates data from several sources, including data from more than 65,000 household health surveys and 87,000 medical records [15]. Although this can be considered the ideal scenario and there is growing interest on the data from population health surveys, particularly due to their potential to reach different population groups with low accessibility by using digital surveys, these data have been used less frequently than data from EHRs for monitoring and informing public mental health decision-making [16,17]. This is the case in several countries, such as Italy and the Netherlands, where EHRs have been used, for example, to monitor the impact of the COVID-19 pandemic on the mental health of the general population [18] and to monitor mental health conditions in children [19], respectively.

One controversial aspect related to the use of data from population health surveys for the monitoring of mental health is the type of mental health measures included within them. Population health surveys commonly include self-reported measures (eg, the 8-item version of the Patient Health Questionnaire) that are considered less valid and reliable than clinical interviews. However, it must be highlighted that, in contrast with other medical conditions, and because of the subjective and self-reported nature of most of the symptoms of mental disorders [20,21], self-reported tools (eg, the 8-item version of the Patient Health Questionnaire) represent valid and reliable measures at the population level (group or aggregate level) and, hence, can be used for timely monitoring and identification of (vulnerable) groups of interest.

Timely assessment and identification of factors related to mental disorders is another strength of the data captured by population health surveys. These surveys usually include within their questionnaires very rich information about sociodemographic and environmental determinants (eg, education and air pollution) and lifestyle habits (eg, diet or physical activity), constituting a valuable resource to assess their potential relationship with mental health. One example of the use of data from population health surveys to assess sociodemographic determinants of mental health is the use of data from the second and third waves of the EHIS (EHIS-2 and EHIS-3, respectively) to identify differences by country in the prevalence of depression across 27 European countries (with a higher prevalence: 1.8 times higher in Germany and 1.5 times higher in Luxembourg relative to the rest of Europe) [22,23]. Additionally, some population health surveys capture data about the use of mental health services (eg, the number of primary care consultations during a specific time), which is a potentially useful resource to inform the planning of public mental health resources [13,24,25]. This information has been previously used to investigate the unmet need of mental health care [9,24] and is particularly relevant for assessing vulnerable population groups that usually have higher rates of mental disorders, lower access to these services or, due to their circumstances, do not use these services when needed (eg, people experiencing homelessness) [26].

Synergistic Relevance of the Data From Population Health Surveys and EHRs

While the use of population health surveys is suitable for different public mental health purposes, in some cases, data from EHRs could be a better option due to their potentially higher completeness and representativeness, the inclusion of richer relevant clinical data, and the possibility of following up with their participants over time (Table 1).

Table 1. General characteristics of population health surveys and electronic health records (EHRs).

	Population health surveys	EHRs
Population included	General population	Clinical population
Representativeness	Household samples to quota samples	Usually public providers
Type of design	Usually cross-sectional	Longitudinal
Assessment of mental disorders	Usually self-reported	Usually diagnostic codes
Range of mental disorders usually considered	Limited	Wide
Other relevant variables	Sociodemographic factors, lifestyle habits, and use of health services	Only clinical variables
Cost	Low	High

One advantage that EHRs have over population health surveys is their potentially higher external validity. In contrast with surveys, which are usually carried out in samples with different degrees of representativeness, data from EHRs are more likely to include entire populations, such as the Hospital Episodes Statistics in England, in which more than 99% of attendees of mental health services from the National Health Service of the United Kingdom are captured [27]. Besides, data from EHRs usually capture high-quality clinical information, such as information about diagnostics (eg, International Classification of Diseases [ICD] codes), treatments (eg, prescribed medication), and other clinical aspects of care (eg, the number of consultations in a specific service), positioning them as the most suitable alternative for the assessment of the use of mental health services and to inform their planning [16,17.24]. Paradoxically, it should be noted that despite the quality of the data from EHRs (particularly to capture severe mental disorders), medical records are not designed for research purposes and focus on clinical populations (ie, not in the general population), thus limiting their generalizability to the general population and even more so to vulnerable population groups less likely to access these services [26]. Additionally, data from EHRs are limited by data capturing systems. However, given the potential differences in codification systems (eg, mental disorders codified using different ICD versions [21]) and across studies, achieving semantic interoperability is a key aspect to consider when leveraging data across multiple data sets [28]. As a consequence, certain relevant outcomes (eg, drug abuse) may be poorly documented and need to be validated for research, or they reach only public mental health service users, thus limiting their generalizability to the growing population groups using private health care services.

Another advantage of EHRs over population health surveys is the possibility of long-term follow-up of participants. While population health surveys are often cross-sectional and anonymous (or anonymized) and participants cannot be followed up over time, the inclusion within EHRs of variables that facilitate the identification of participants and date variables (eg, dates of admission or follow-up consultations) make it possible to follow up with them over time. This is particularly important for correct estimation of the incidence of different health problems, the rates of recovery or relapse of some disorders over time, or the impact on mental health over time on different factors and the establishment inferences about potential causal associations [29]. However, there is an increasing interest in longitudinal surveys, in panel data with repeated assessments of the same individuals over time (eg, the British Cohort Study in the United Kingdom) [30,31], and in the inclusion of variables within the survey questionnaires that allow the follow-up of their participants over their life course (eg, participant identification codes). Thus, while EHRs could be considered currently more suitable for the timely assessment of clinical aspects of mental disorders, complementing their data with those from longitudinal surveys could enhance and enrich such assessment [32,33].

Despite the abovementioned differences between the data from population health surveys and those from EHRs for monitoring of mental health, the complementary nature of these data sources must be highlighted. One successful example of their complementarity for the study of mental disorders has been reported in Denmark [32], where a shared identification number assigned to the individuals included in their information systems allows the linkage between population health surveys and EHRs [33]. The extensive linkage of individual data in the Danish population (including population health surveys; national records from hospitals, clinics, pharmacies, and death registries; and various other public and private data sources) synergistically enhanced the data's overall quality for the assessment and monitoring of mental health conditions.

Moving Forward Toward the Generalization of the Use of Population Health Surveys in Public Mental Health

Given all the potential benefits and applications indicated herein, widespread use of population health survey data (ideally linked with data from high-quality EHRs) could be helpful both for research purposes and to enhance decision-making in public mental health. Their widespread use might be achieved through not only the deployment of large, address-based probability-sampled surveys, similar to the deployment of the NHIS in the United States [3], but also encouraging individual research groups to join forces and carry out large-scale population health surveys or leverage the ones already implemented. However, in order to guarantee adequate and appropriate use of data from population health surveys, the development of a common usage framework and enhance their interoperability with other data sources remains essential.

The development of a common usage framework of data from population health surveys must be in line with widely recognized initiatives such as the GRADE (Grading of Recommendations Assessment, Development and Evaluation) Evidence to Decision (EtD) framework [34], and will need to include unified protocols and guidelines for data collection, data analysis, and interpretation. For data collection, guidelines must consider both the tools used for monitoring of mental health and the minimum set of general domains (eg, socioeconomic and health service use) and specific factors (eg, educational level attained and number of primary care consultations during the last year) included within any population health survey.

It should be also mentioned that the data collection process and the potential biases related to the representativeness of the data and their quality are also key challenges. The use of multistage sample designs and data verification techniques, for instance, would enhance the validity of data from population health surveys for both research and public mental health purposes. Additionally, for the correct interpretation of their results, a set of valid and reliable standard measures for the assessment of a broad range of mental disorders needs to be established, and their cross-cultural equivalence in different contexts should be ensured [10]. Therefore, the development of a common usage framework for the use of population health surveys for monitoring mental health would improve the suitability of their data for this purpose, enhance the comparability of their results and, consequently, increase their impact on public mental health decision-making.

Finally, given the complementing advantages of the data from population health surveys and those from EHRs, another key step forward is to promote the interoperability of these 2 sources. However, to achieve interoperability, it will be necessary to systematically include variables within their data sets, which facilitate the identification and linkage of individuals across them while also ensuring anonymity. This linkage will be helpful to bring together their strengths and potentially improve the accuracy and relevance of both EHRs and health surveys as tools for monitoring mental health at the population level [35]. However, due to the substantial variation between countries in data protection laws, systematically including these types of variables could be challenging, particularly in the case of international surveys, such as the World Mental Health Surveys or EHIS. Hence, promoting the implementation of shared international regulations about data linkage and anonymization could be a helpful way to enhance the quality of administrative health data for monitoring the health status of the population.

Opportunities and Implications of the Use of Population Health Surveys

Monitoring of mental disorders at the population level, identifying their determinants, and determining their association with health care usage are key components for effective mental health prevention. Balancing their advantages and limitations, data from population health surveys are instrumental in addressing these challenges. However, to guarantee the adequate use of these data, there are different aspects that should be considered, including their specific limitations and challenges, the potential adoption of a common and shared framework for their use, and their interoperability with other data sources, such as EHRs, using shared identification variables. Promoting the use of linked data from population health surveys, EHRs, and other public and private data sources could enhance the quality of the information available for public mental health decision-making, particularly in middle- and high-income countries and ultimately improve the planning of mental health resources and maximize the impact of relevant policies to reduce the burden of mental disorders.

Acknowledgments

This work is partially funded by the University of Exeter, and CIBER (Centro de Investigación Biomédica en Red) Epidemiology and Public Health (grant ESP21PI05). This paper represents independent research partly funded by the National Institute for Health Research's (NIHR's) Biomedical Research Centre at South London and Maudsley NHS (National Health Service) Foundation Trust, and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Conflicts of Interest

None declared.

References

1. Prabhakar A, Rice S. White House Report on Mental Health Research Priorities. The White House. 2023. URL: https://www.whitehouse.gov/wp-content/uploads/2023/02/White-House-Report-on-Mental-Health-Research-Priorities.pdf [accessed 2023-03-18]
2. PHE Strategy 2020 to 2025. Government of the United Kingdom. 2019. URL: https://www.gov.uk/government/publications/phe-strategy-2020-to-2025 [accessed 2023-10-25]
3. Blewett LA, Dahlen HM, Spencer D, Rivera Drew JA, Lukanen E. Changes to the design of the National Health Interview Survey to support enhanced monitoring of health reform impacts at the state level. Am J Public Health 2016 Nov;106(11):1961-1966. [doi: 10.2105/ajph.2016.303430]
4. Mindell J, Biddulph JP, Hirani V, Stamatakis E, Craig R, Nunn S, et al. Cohort profile: the health survey for England. Int J Epidemiol 2012 Dec 09;41(6):1585-1593. [doi: 10.1093/ije/dyr199] [Medline: 22253315]
5. Aromaa A, Koponen P, Tafforeau J, Vermeire C, HIS/HES Core Group. Evaluation of health interview surveys and health examination surveys in the European Union. Eur J Public Health 2003 Sep;13(3 Suppl):67-72. [doi: 10.1093/eurpub/13.suppl_1.67] [Medline: 14533752]
6. Seeman T, Merkin SS, Crimmins E, Koretz B, Charette S, Karlamangla A. Education, income and ethnic differences in cumulative biological risk profiles in a national sample of US adults: NHANES III (1988-1994). Soc Sci Med 2008 Jan;66(1):72-87 [FREE Full text] [doi: 10.1016/j.socscimed.2007.08.027] [Medline: 17920177]
7. Arias de la Torre J, Vilagut G, Ronaldson A, Valderas JM, Bakolis I, Dregan A, et al. Reliability and cross-country equivalence of the 8-item version of the Patient Health Questionnaire (PHQ-8) for the assessment of depression: results from 27 countries in Europe. Lancet Reg Health Eur 2023 Aug;31:100659 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100659] [Medline: 37332385]
8. Huijts T, Stornes P, Eikemo T, Bambra C, HiNews Consortium. Prevalence of physical and mental non-communicable diseases in Europe: findings from the European Social Survey (2014) special module on the social determinants of health. Eur J Public Health 2017 Feb 01;27(suppl_1):8-13. [doi: 10.1093/eurpub/ckw232] [Medline: 28355647]
9. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. JAMA 2004 Jun 02;291(21):2581-2590. [doi: 10.1001/jama.291.21.2581] [Medline: 15173149]
10. Schnittker J. Religion, social integration, and depression in Europe: Evidence from the European Social Survey. Soc Sci Med 2020 Dec;267:112376. [doi: 10.1016/j.socscimed.2019.112376] [Medline: 31255360]
11. Monitoring mental health and suicide prevention reform: national report 2019. National Mental Health Commission (Australia). 2019. URL: https://apo.org.au/sites/default/files/resource-files/2019-09/apo-nid260121.pdf [accessed 2022-10-04]
12. Eaton WW. Public Mental Health. New York, NY: Oxford University Press; 2012.
13. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. Br J Psychiatry 2004 Jun;184:526-533. [doi: 10.1192/bjp.184.6.526] [Medline: 15172947]
14. Wahlbeck K. Public mental health: the time is ripe for translation of evidence into practice. World Psychiatry 2015 Feb;14(1):36-42 [FREE Full text] [doi: 10.1002/wps.20178] [Medline: 25655149]
15. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150 [FREE Full text] [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
16. Knapp M, Wong G. Economics and mental health: the current scenario. World Psychiatry 2020 Feb 10;19(1):3-14 [FREE Full text] [doi: 10.1002/wps.20692] [Medline: 31922693]
17. Jha AK. The promise of electronic records: around the corner or down the road? JAMA 2011 Aug 24;306(8):880-881. [doi: 10.1001/jama.2011.1219] [Medline: 21862751]
18. Gualano MR, Lo Moro G, Voglino G, Bert F, Siliquini R. Monitoring the impact of COVID-19 pandemic on mental health: a public health challenge? Reflection on Italian data. Soc Psychiatry Psychiatr Epidemiol 2021 Jan;56(1):165-167 [FREE Full text] [doi: 10.1007/s00127-020-01971-0] [Medline: 33034669]
19. Koning NR, Büchner FL, van den Berg AW, Choi SYA, Leeuwenburgh NA, Paijmans IJM, et al. The usefulness of electronic health records from preventive youth healthcare in the recognition of child mental health problems. Front Public Health 2021 May 31;9:658240 [FREE Full text] [doi: 10.3389/fpubh.2021.658240] [Medline: 34136452]
20. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Virginia: American Psychiatric Association; 2013.
21. Reed GM, First MB, Kogan CS, Hyman SE, Gureje O, Gaebel W, et al. Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. World Psychiatry 2019 Feb 02;18(1):3-19 [FREE Full text] [doi: 10.1002/wps.20611] [Medline: 30600616]
22. Arias-de la Torre J, Vilagut G, Ronaldson A, Serrano-Blanco A, Martín V, Peters M, et al. Prevalence and variability of current depressive disorder in 27 European countries: a population-based study. Lancet Public Health 2021 Oct;6(10):e729-e738 [FREE Full text] [doi: 10.1016/S2468-2667(21)00047-5] [Medline: 33961802]

23. Arias-de la Torre J, Vilagut G, Ronaldson A, Bakolis I, Dregan A, Martín V, et al. Prevalence and variability of depressive symptoms in Europe: update using representative data from the second and third waves of the European Health Interview Survey (EHIS-2 and EHIS-3). Lancet Public Health 2023 Nov;8(11):e889-e898 [FREE Full text] [doi: 10.1016/S2468-2667(23)00220-7] [Medline: 37898521]
24. Moran V, Suhrcke M, Ruiz-Castell M, Barré J, Huiart L. Investigating unmet need for healthcare using the European Health Interview Survey: a cross-sectional survey study of Luxembourg. BMJ Open 2021 Aug 03;11(8):e048860 [FREE Full text] [doi: 10.1136/bmjopen-2021-048860] [Medline: 34344682]
25. Kilbourne AM, Beck K, Spaeth-Rublee B, Ramanuj P, O'Brien RW, Tomoyasu N, et al. Measuring and improving the quality of mental health care: a global perspective. World Psychiatry 2018 Feb 19;17(1):30-38 [FREE Full text] [doi: 10.1002/wps.20482] [Medline: 29352529]
26. Arias-de la Torre J, Valderas JM, Benavides FG, Alonso J. Cardboard floor: about the barriers for social progression and their impact on the representativeness of epidemiological studies. J Epidemiol Community Health 2021 Feb 22;75(2):105-106. [doi: 10.1136/jech-2020-214978] [Medline: 32963014]
27. Davis KAS, Bashford O, Jewell A, Shetty H, Stewart RJ, Sudlow CLM, et al. Using data linkage to electronic patient records to assess the validity of selected mental health diagnoses in English Hospital Episode Statistics (HES). PLoS One 2018 Mar 26;13(3):e0195002 [FREE Full text] [doi: 10.1371/journal.pone.0195002] [Medline: 29579109]
28. de Mello BH, Rigo SJ, da Costa CA, da Rosa Righi R, Donida B, Bez MR, et al. Semantic interoperability in health records standards: a systematic literature review. Health Technol (Berl) 2022;12(2):255-272 [FREE Full text] [doi: 10.1007/s12553-022-00639-w] [Medline: 35103230]
29. de Zwart PL, Jeronimus BF, de Jonge P. Empirical evidence for definitions of episode, remission, recovery, relapse and recurrence in depression: a systematic review. Epidemiol Psychiatr Sci 2018 May 17;28(5):544-562. [doi: 10.1017/s2045796018000227]
30. Mcelroy E, Villadsen A, Patalay P, Goodman A, Richards M, Northstone K, et al. Socioeconomic and sex inequalities in parent-reported adolescent mental ill-health: Time trends in four British birth cohorts. OSF Preprints Preprint posted online January 13, 2022. [doi: 10.31219/osf.io/3zn2h]
31. Arias-de la Torre J, Ronaldson A, Prina M, Matcham F, Pinto Pereira SM, Hatch SL, et al. Depressive symptoms during early adulthood and the development of physical multimorbidity in the UK: an observational cohort study. Lancet Healthy Longev 2021 Dec;2(12):e801-e810. [doi: 10.1016/s2666-7568(21)00259-2]
32. Momen NC, Plana-Ripoll O, Agerbo E, Benros ME, Børglum AD, Christensen MK, et al. Association between mental disorders and subsequent medical conditions. N Engl J Med 2020 Apr 30;382(18):1721-1731. [doi: 10.1056/nejmoa1915784]
33. Momen NC, Lasgaard M, Weye N, Edwards J, McGrath J, Plana-Ripoll O. Representativeness of survey participants in relation to mental disorders: a linkage between national registers and a population-representative survey. Int J Popul Data Sci 2022 Oct 19;7(4):1759 [FREE Full text] [doi: 10.23889/ijpds.v7i4.1759] [Medline: 37152406]
34. Moberg J, Oxman AD, Rosenbaum S, Schünemann HJ, Guyatt G, Flottorp S, GRADE Working Group. The GRADE Evidence to Decision (EtD) framework for health system and public health decisions. Health Res Policy Syst 2018 May 29;16(1):45 [FREE Full text] [doi: 10.1186/s12961-018-0320-2] [Medline: 29843743]
35. Arias de la Torre J, Ronaldson A, Valderas JM, Vilagut G, Serrano-Blanco A, Hatch SL, et al. Diagnostic promiscuity: the use of real-world data to study multimorbidity in mental health. Br J Psychiatry 2021 May;218(5):237-239. [doi: 10.1192/bjp.2020.257] [Medline: 33436108]

Abbreviations

EHIS: European Health Interview Survey
EHIS-2: second wave of the European Health Interview Survey
EHIS-3: third wave of the European Health Interview Survey
EHR: electronic health record
EtD: Evidence to Decision
GRADE: Grading of Recommendations Assessment, Development and Evaluation
ICD: International Classification of Diseases
NHIS: National Health Interview Survey

Edited by A Mavragani, T Sanchez; submitted 12.04.23; peer-reviewed by N Robles, J Black, E Sükei; comments to author 07.08.23; revised version received 18.08.23; accepted 26.09.23; published 23.11.23.

Please cite as:
Arias de la Torre J, Vilagut G, Ronaldson A, Bakolis I, Dregan A, Navarro-Mateu F, Pérez K, Szücs A, Bartoll-Roca X, Molina AJ, Elices M, Pérez-Solá V, Martín V, Serrano-Blanco A, Valderas JM, Alonso J
Reconsidering the Use of Population Health Surveys for Monitoring of Mental Health
JMIR Public Health Surveill 2023;9:e48138
URL: https://publichealth.jmir.org/2023/1/e48138
doi:10.2196/48138
PMID:37995112

Viewpoint

Combatting SARS-CoV-2 With Digital Contact Tracing and Notification: Navigating Six Points of Failure

Joanna Masel[1], BSc, DPhil; James Ian Mackie Petrie[2,3,4], BASc, MMath, PhD; Jason Bay[5], BS, MSc; Wolfgang Ebbers[6], PhD; Aalekh Sharan[7], BEng, MBA; Scott Michael Leibrand[8], BS; Andreas Gebhard[9]; Samuel Zimmerman[10], BA

[1]Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ, United States
[2]Department of Applied Mathematics, University of Waterloo, Waterloo, ON, Canada
[3]Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[4]Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[5]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore
[6]Erasmus School of Social and Behavioural Sciences, Department of Public Administration and Sociology, Erasmus University Rotterdam, Rotterdam, Netherlands
[7]Aarogya Setu, Government of India, New Delhi, India
[8]Community Epidemiology in Action (coEpi), Seattle, WA, United States
[9]Temporary Contact Number Protocol (TCN) Coalition, New York, NY, United States
[10]PathCheck, Cambridge, MA, United States

Corresponding Author:
Joanna Masel, BSc, DPhil
Department of Ecology & Evolutionary Biology
University of Arizona
1041 E Lowell St
Tucson, AZ, 85721
United States
Phone: 1 5206269888
Email: masel@arizona.edu

Abstract

Digital contact tracing and notification were initially hailed as promising strategies to combat SARS-CoV-2; however, in most jurisdictions, they did not live up to their promise. To avert a given transmission event, both parties must have adopted the technology, it must detect the contact, the primary case must be promptly diagnosed, notifications must be triggered, and the secondary case must change their behavior to avoid the focal tertiary transmission event. If we approximate these as independent events, achieving a 26% reduction in the effective reproduction number R_t would require an 80% success rate at each of these 6 points of failure. Here, we review the 6 failure rates experienced by a variety of digital contact tracing and contact notification schemes, including Singapore's TraceTogether, India's Aarogya Setu, and leading implementations of the Google Apple Exposure Notification system. This leads to a number of recommendations, for example, that the narrative be framed in terms of user autonomy rather than user privacy, and that tracing/notification apps be multifunctional and integrated with testing, manual contact tracing, and the gathering of critical scientific data.

(JMIR Public Health Surveill 2023;9:e49560) doi:10.2196/49560

KEYWORDS

COVID-19; SARS-CoV-2; pandemic preparedness; decentralized protocols; smartphone; mobile phone; contact tracing

Introduction

Contact tracing is a time-tested tool to fight an emerging outbreak of infectious diseases such as COVID-19, caused by the virus SARS-CoV-2. If x% of infections are identified and y% of those in contact with a known case are traced in time and persuaded to stay home, then in a well-mixed population, the effective reproduction number R_t will decrease by a factor of x×y. This seems sufficient to reach the level $R_t<1$ needed to quash an outbreak of many emerging pathogens (which have not yet evolved a basic reproduction number R_0>>1) before the outbreak develops too far, especially in combination with modest social distancing. However, SARS-CoV-2 proved to be a particular challenge for contact tracing. With short incubation periods and presymptomatic transmission making it difficult to trace contacts in time, digital contact tracing held tremendous promise, especially when leveraging smartphones that were already in consumers' hands. Here, we reflect on why digital alternatives to traditional contact tracing instead had limited impact on SARS-CoV-2 transmission. We do so with a focus on Exposure Notification (EN) and related protocols, in part by drawing on first-hand experiences from our various involvements with this technology, including material that has not previously been public as well as publicly available documents that are not indexed for literature searches. Our intent is to better inform those who might wish to prepare for and fight a new pandemic with a similar technological approach, allowing them to learn from what happened during the COVID-19 pandemic.

A Brief Note on Terminology

Some digital protocols make it impossible to identify pairs of interacting individuals, even in cases where one transmitted disease to the other—in this case, we refer to *notification* instead of *tracing*. We refer to *proximity* versus *presence* notification and tracing based on whether exposure is assessed on the basis of proximity to an individual versus presence at a shared-air venue. We reserve the term EN for the specific protocol implemented by Apple and Google; this protocol is a form of proximity notification. EN is one of several protocols used to assess proximity based on signals sent and received between pairs of devices using low-energy Bluetooth. Other protocols may use ultra wideband or ultrasound for proximity detection, or detect presence from other information such as QR code scan histories, GPS coordinates, or logs of Wi-Fi access points. These technologies can also be used to warn of future infection risk [1], a use case that is beyond the scope of this study.

Overview of the Six Failure Points

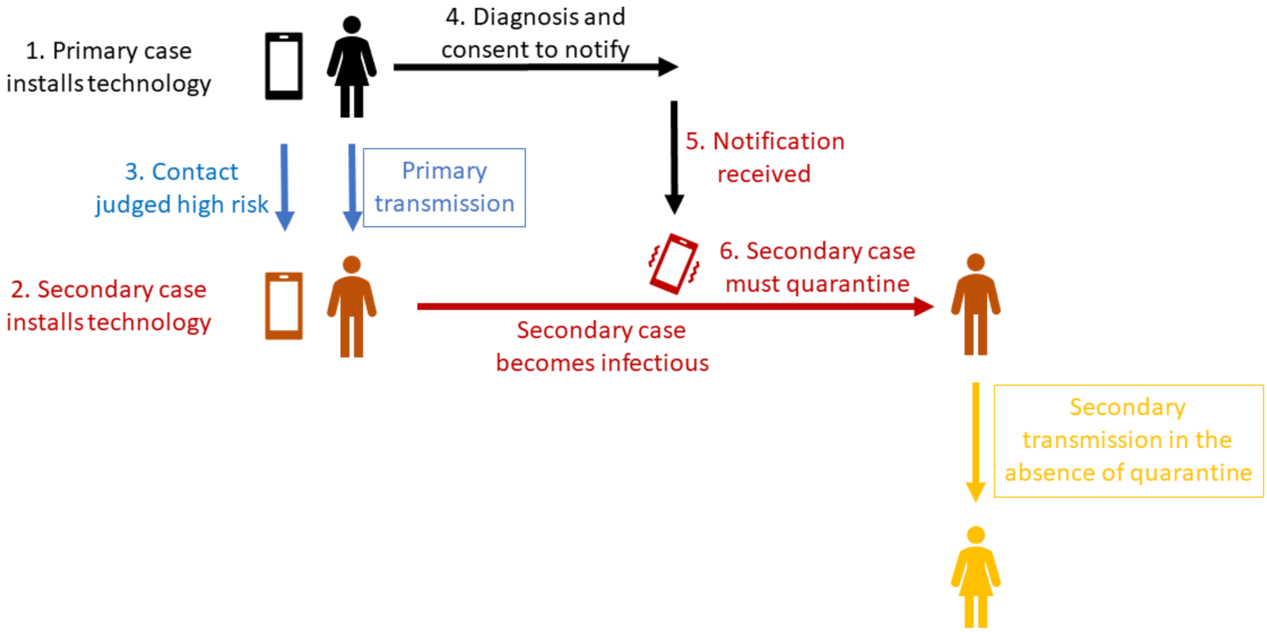
To be effective in stemming a transmission, a notification must navigate six potential points of failure (Figure 1):

1. The primary case must have the technology in place at the time of transmission.
2. The secondary case must have the technology in place at the time of transmission.
3. The exposure that resulted in transmission must be judged to be high risk.
4. The primary case must obtain a positive diagnosis in a timely manner.
5. Notifications stemming from the primary case must be rapidly triggered following a positive diagnosis.
6. After receiving a notification, the secondary case must change their behavior in a manner that prevents onward transmission to tertiary cases.

If each of these steps were successful 80% of the time, and we approximate the 6 steps as independent events, then transmission (R_t) would be reduced by 0.8^6=26%. Although not transformational on its own, this would be a significant contribution to quashing an outbreak, and nonindependence will make this figure somewhat higher. However, if each step were successful a still-respectable 40% of the time, again assuming independent events, R_t is reduced by only 0.4%. Although this can make a valuable contribution to flattening a curve [2] or reducing the stringency of indiscriminate social distancing [3], it falls far short of containing a pandemic. Given this simple mathematical consideration, if the aim is containment such that life is relatively normal while waiting for a vaccine, then we clearly need to achieve low failure rates at all of the 6 failure points.

Next, we discuss each of the failure points. We provide some history of how they were handled, together with speculation about how they might have been handled better.

Figure 1. Chain of viral transmission from primary to secondary to tertiary case. Digital contact tracing and notification must succeed at all 6 critical points to avert transmission to the tertiary case.



Failure Points 1 and 2: App Adoption

We jointly discuss the first 2 points of failure because they are similar; both involve technology adoption among the population destined to become infected. Following media coverage of the model by Ferretti et al [4], more attention was paid to app adoption than to other points of failure, with an odd obsession with 60% adoption in the general population as a magic number [5]. Note that the adoption rate that matters is not that of the general population but that of the population likely to become infected. For SARS-CoV-2, this meant that adoption among essential workers was what mattered during stay-at-home mandates; when staying at home was common but optional, young adults constituted a disproportionate share of cases. However, app adoption among the general population can serve as a rough approximation.

The most straightforward way to achieve low failure rates at points #1 and #2, while still maintaining user autonomy, is to make the broadcasting and reception of appropriate signals opt-out. Opt-in consent of the primary case would still be required at failure point #5 and that of the secondary case at #6. Apple and Google were the only entities with the power to implement an opt-out protocol, with competitors such as Huawei having little market share. Despite a secure EN design that ensures that no information about Bluetooth signals sent or received leaves the phone without subsequent user consent, they chose to make Bluetooth broadcasting and receiving opt-in. We note that Apple's "Find my..." service uses Bluetooth in a similar fashion and is at the time of writing opt-out rather than opt-in [6].

An even lower failure rate could have been achieved by a protocol proposed by one of us (JIMP) to use the Wi-Fi logs already collected by Android and by Apple's iOS operating systems instead of an additional purpose-built Bluetooth layer. Had Google and Apple chosen to allow an app post hoc access to these logs or to build operating system functionality around them, this would have completely eliminated failure point #1 for individuals who regularly carry smartphones. Individuals would still need to have the technology switched on to bypass failure point #2 to receive a notification. Similar to Bluetooth solutions, those who do not regularly carry smartphones (mostly young children and older adults) would need to be provided with other devices that could record proximity or presence; in Singapore, small digital tokens resembling keychain fobs covered smartphone nonusers who would not otherwise have been able to participate.

To understand why Apple and Google did not make the technology opt-out or Wi-Fi based, it is important to consider the incentive structure from Apple's and Google's point of view. Cooperation requires time from their employees and has no apparent commercial upside (beyond the economic gains to the company should efforts succeed in containing a pandemic). From their perspective, the primary consideration is the effect of their actions on brand perception. The upside impact on their brand is gained primarily from doing *something* (eg, working together), but there is no proportionate impact from doing something more rather than less effective. The potential for downside impact is substantial and focused mostly on privacy concerns, exacerbated by any ceding of control over the technology and its marketing to public health authorities. They are thus incentivized to do *something*, to carefully manage perceptions of that something, and to avoid brand risk, but not to increase effectiveness against disease transmission. In future, if companies such as Apple and Google are to be gatekeepers of such technology, governments should devote urgent attention to aligning their incentive structures.

Achieving high opt-in adoption was difficult. Singapore's non-EN solution succeeded, with opt-in adoption >90% even before the app became required for entry into public spaces [7]. In contrast, the highest EN adoption rates are of Germany's

Corona-Warn-App, used by approximately 34% of the population [8], and the National Health Service (NHS) COVID-19 app, with 17% to 25% of the population of England and Wales having activated Bluetooth exchanges in 2020 and 2021 [9]. Lower rates were reported elsewhere: 8% in Canada, 13% in New Zealand, 10% to 16% in the Netherlands, and 19% in Switzerland [10].

A positive caveat to this fairly bleak assessment of uptake is that if one individual has adopted the app, it is more likely that other members of their social network have also adopted it. In other words, failures at points 1 and 2 are correlated, making the probability of overall failure lower than it would be if each failure were an independent event [11]. Both for this reason and as part of good marketing practices [12], it can make sense to look at adoption rates within smaller communities. Adoption by 46% of the cases was achieved in a campus setting with an intensive marketing push [13] and adoption by >33% of the population was achieved in an island setting [14].

Key to the relatively high adoption of the NHS COVID-19 app and the Corona-Warn-App were their additional functionalities. When entering a venue, scanning a QR code with the NHS COVID-19 app was offered as an alternative to writing down contact details. Although subsequent presence tracing on the basis of either form of information was conducted relatively rarely (refer to *Failure Point 5: Triggering Notifications* section), the requirement to check in to a venue did prompt app installation [15], especially in the subset of the population most likely to become infected through contact with strangers. The NHS COVID-19 app was also useful for ordering tests, and both it and the Corona-Warn-App offered a rapid and secure system for receiving test results. Later, the integration of European Union digital COVID-19 certificates (for vaccination, recent negative test result, or recovery from infection) triggered another wave of app adoption. Near-universal adoption in Singapore was achieved only after the proximity tracing app became the only means for compulsory venue check-in [16]. Similarly, adoption rates in India skyrocketed after the government made the app mandatory for smartphone owners to move freely in public areas. We believe that to achieve high adoption, future pandemic apps need to seamlessly integrate services that are useful and convenient for users.

In late 2020, Apple and Google launched the “Exposure Notification Express” (ENX) system. In Androids, this is a simple app that public health authorities can autogenerate from a small set of choices. In Apple’s iOS, this is part of the Settings. Apple claimed that placing the scheme in the Settings created less friction for opt-in than was present for an app, and hence, ENX should be preferred to a custom app. However, ENX never reached use levels comparable with those of more successful European EN apps. This is despite the fact that the phones of residents of adopting US states were pinged to promote activation and installation. Such pings were used to promote ENX but not custom EN apps. When this situation was pointed out, Google immediately offered the same adoption promotion pings for custom apps as for ENX, whereas Apple declined to do so.

A number of players, including but not limited to Apple, argued that the key to persuading people to adopt lay in ensuring the privacy of the system. Fortunately, decentralized protocol designs offer powerful solutions to the privacy problem. Centralized approaches send information about who went where with whom to a central database. This breach of privacy, although clearly facilitating contact tracing, also poses risks of abuse as part of a mass surveillance state. Under a decentralized protocol, information about a user is stored on that user’s device and, to some degree, the devices with which it was in contact. Users can be given autonomy over the use of data on their own devices so that, for example, the data cannot be accessed without consent and can be deleted at any time. Far from threatening privacy, apps that use decentralized protocols are among the safest apps on users’ smartphones.

Interestingly, privacy-invasive schemes, such as QR code check-in presence tracing in Australia, were well accepted and widely used (self-reports of 61.9% always checking in either digitally or on paper, 26.3% mostly, 4.7% sometimes, 3.3% occasionally, and 3.8% never [17]). This achieved acceptably low failure rates far superior to any EN implementation, even with enforcement left to the venue and peer pressure. Indeed, supported by celebrity-driven public relations pushes, 40% had downloaded or were willing to download even the ineffective but convenient Luca QR code check-in app in Germany [18,19]. Perhaps, these much higher adoption rates for presence tracing than for proximity notification are because what was recorded was considered public information that a user was in a public space, rather than also capturing who had interacted with whom even in private. A simpler explanation is that the QR code check-in system is easy to understand as equivalent to writing down your name and number; decentralized privacy-preserving schemes, by being harder for the public to understand, run the paradoxical risk of decreasing rather than increasing trust and hence adoption.

Trust in government predicts adoption more strongly and consistently than privacy [15,20-23]. Although 20% to 25% of respondents to cross-sectional surveys cited privacy concerns as a reason for not installing a contact tracing app [23,24], a longitudinal study found no causal relationship (although it did for health concerns and social norming) [25], and a focus group found that health concerns were more important than privacy considerations [26].

High trust in the government in general [27] as well as in the government’s response to the COVID-19 pandemic [28] was critical to Singapore’s high adoption. When trust in governments is low, it might help if more trusted entities, such as primary care providers [29], were used to promote adoption. Singapore’s explicit policy to relax certain pandemic restrictions once adoption increased from 50% to 70% might also have helped drive adoption, together with the scheme’s high profile as a clear government priority.

Privacy is not the only aspect of trust salient to potential users; many also wanted to know that an app was effective before they downloaded it. Even if they decided to take a chance at first and install it, they wanted subsequent reassurance that it had proved effective to keep it active on their phones. Multifunctional apps, such as the NHS COVID-19 app and the Corona-Warn-App, can also provide reassurance via interaction that there is a point to maintaining the app on the user's phone. Although it is clearly important to hire good marketing and public relations professionals to promote adoption, it is just as important that scientists and engineers do not make product decisions in isolation from their impact on marketing. This is because achieving high adoption is part of the science of making these apps effective.

Distinct from the scientific study of app effectiveness in stemming transmission [30] is the study of viral transmission itself to learn its incubation period, infectious period, and mode of transmission [31]. An obstacle to such a study was that some *privacy-first* rhetoric rejected making the study of SARS-CoV-2 transmission an aim of the scheme, even if the science could be done in a manner that preserved privacy.

The only reliable way to learn about transmission to and from humans is to observe it, either in human challenge trials (directly observing incubation periods and also the infectious period if the study includes transmission to cohoused animals) or in natural circumstances (ie, via contact tracing). Given ethical concerns about the former, an unacknowledged corollary of rejecting the use of digital schemes to study transmission is that studies must be performed via manual contact tracing (ie, in a more privacy-invasive way) or not at all. Unfortunately, even the more basic functions of manual contact tracing were quickly overwhelmed, leaving little capacity for the more intensive investigations required to study transmission dynamics. Not using digital approaches to fill this void was a lost opportunity. This left early manual contact tracing studies as the only source for basic transmission parameters, even after the incubation time and infectious period of SARS-CoV-2 were suspected to have shifted because of viral evolution and immunity.

Even an app whose failure rates across the 6 points were too high to substantially reduce transmission could have been enough to generate valuable information about evolving incubation periods and infectious periods in close to real time [31]. Apps can also help with epidemiological surveillance; for example, the Indian government combined location (with a random error added for privacy; EN apps were banned from accessing location) and symptom information to identify emerging hot spots before seeing spikes in test positivity. Singapore used the number of detected contacts per case to inform changes in social distancing policies. Support for the scheme, both from the public and from public health authorities, might have been shored up by evidence that it was at least doing good for science or disease surveillance, and this might have created a virtuous circle of adoption.

A prevailing narrative described the technological choices in terms of a trade-off between privacy and effectiveness, in which any collection of more data can increase perceived privacy risk and hence reduce public acceptance [30]. We believe that this framing is harmful, especially when it comes to promoting app adoption; for example, the public might conclude that because the technology is private, it cannot be effective. Regarding substantive risks beyond those of perception, with flexibility and creativity, effectiveness can be achieved by a decentralized, privacy-preserving design that does not risk expanding mass surveillance. With sufficient ingenuity, there may be no need for compromise at all, which makes the privacy versus effectiveness trade-off narrative misleading. *Privacy first* should therefore not be used as grounds to support the status quo to the point of refusing to engage in ongoing dialog regarding how to safely proceed with iterative improvements to effectiveness, its measurement, and the gathering of broadly valuable scientific or epidemiological surveillance data. Given that harms from technology are not limited to privacy concerns [32], a privacy-preserving design does not relieve governments of their duty to monitor effectiveness [33].

Furthermore, privacy is merely a component and a means to the more important end goals of autonomy and protection from abuse. The latter need not be a serious issue because a well-designed, decentralized solution can offer both effectiveness and protection from abuse. However, trade-offs between effectiveness and autonomy are real, that is, coercion can increase effectiveness, whereas autonomy can weaken it. The option to maintain privacy while extracting benefit from the system is an important but not the only aspect of autonomy.

Although the privacy approach of EN provides excellent protection from abuse, it actually limits autonomy by denying users the right to share their data with public health authorities if they wish to. This became clear when one of us (JIMP) proposed a small change to the EN protocol that could convert it from proximity notification to proximity tracing, without significantly raising the danger of misuse. As described in the *Failure Point 3: Detecting and Evaluating Exposure* section, this modified protocol would allow certain users to share more data with public health authorities. This would apply to users who gave explicit consent to sharing after both receiving a notification and testing positive. By helping manual contact tracers do their job of pairing cases and identifying superspreading events, this change would presumably have generated more buy-in from public health authorities (discussed in the *Failure Point 5: Triggering Notifications* section). More buy-in from the local public health authorities responsible for manual contact tracing could have set up a virtuous circle by which they, in turn, promote adoption in the community. This small modification was rejected by Apple as being a threat to privacy, despite the fact that any privacy loss is triggered by users exercising autonomy over the use of their data. Contributing to this dynamic is the fact that privacy is part of Apple's brand, aligned with the broader deployment of antitracking as a business strategy [34-36]. A preference for privacy over autonomy, by disallowing the opt-in sharing of more information with public health authorities, has also been expressed by some academia-based protocol designers [37].

The slogan *privacy first* is an odd one. We believe that autonomy and protection from abuse are better-framed goals than privacy. Furthermore, the best way to put privacy first is self-evidently to have no digital system at all, raising the possibility that privacy first schemes were designed primarily to displace more invasive options, with stemming transmission almost an afterthought. For each proposed scheme, one should assess the nature and scope of any vulnerabilities versus the anticipated impact on disease transmission.

Failure Point 3: Detecting and Evaluating Exposure

Point of failure #3 concerns the need to record an interaction at which viral transmission took place, but viewing it this way is not sufficient. Total success could trivially be achieved by a general stay-at-home order, which is equivalent to telling everyone that they are potentially exposed. This illustrates the importance of avoiding false positives, that is, alerts received by people who are not infected. The primary purpose of proximity or presence notification is best seen as identifying and alerting people at a sufficiently high risk of infecting other people such that they should change their behavior in response to that information. This benefit applies only to secondary cases who would not otherwise know of their exposure in a timely manner, that is, it excludes household contacts and other social networks that perform rapid do-it-yourself tracing. A major advantage of digital schemes is their complementary ability to notify strangers.

One aspect of this is to detect any contact at all. With core (ie, regular) Bluetooth, iPhone-to-iPhone communication does not work in the background. Apple's cooperation in setting up EN was needed to overcome this serious obstacle. Apple could have chosen to enable iPhone-to-iPhone communication in the background in the same way that they already allowed background communication between iPhones and Androids, which would have allowed app developers to implement their own protocols. Indeed, India launched the Aarogya Setu app using core Bluetooth rather than EN, given that iPhones are rare in India compared with Androids. Singapore stuck to its own BlueTrace protocol (based on core Bluetooth), given the sacrifice of epidemiological utility that switching to EN would have implied [38], eventually leveraging Android-iPhone background communication into a gossip protocol that effectively allowed iPhones to see each other in the background whenever an app-using Android was also present [39]. Other non-EN jurisdictions such as France pivoted away from proximity notification and toward presence tracing. In Australia, state-specific centralized presence tracing apps were the dominant response, whereas a federal app based on TraceTogether, despite a respectable early download rate, languished once it became apparent that it failed at other steps.

The lack of interoperability among jurisdictions can also interfere with contact detection. For example, if each person is far from home for 5% of their interactions (eg, being away for 18 d/y), the corresponding contact detection failure rate will be approximately 10%. Thus, when viewed quantitatively, the considerable attention devoted to this issue seems disproportionate relative to other sources of failure. As an alternative to immediate global standardization, metropolitan or other highly interconnected areas spanning multiple jurisdictions (eg, Washington, District of Columbia combined with adjacent counties in other US states or the Navajo Nation spanning multiple US states) could choose to roll out a joint product. Multiple apps need not interfere with each other, making this an option for some commuters. A common server was set up to enable communication among the EN implementations of different US states and another common server to communicate among implementations of different European nations [40], but despite the attention to this issue, global interoperability was never achieved. Perhaps the best model for a new pandemic is to allow divergence in the early stages as part of a process of innovation at the cost of interoperability, to learn from the different experiments performed by different jurisdictions, and then to use the desire for interoperability as the catalyst for switching from lower-performing systems to higher-performing systems through app updates. An advantage of postponing interoperability is to ensure that premature standardization does not suppress innovation.

Beyond detecting exposure at all, doing a better job in assessing the infection risk posed by a given exposure can be seen as an ethical imperative, given the harms caused by unnecessary quarantine [32]. Better risk analysis means more precise resolution as to which individuals pose how much statistical risk to others on a given day and using a risk threshold to trigger notification. In other words, good risk analysis means improving the receiver operating characteristic (ROC) curve with respect to infection and then choosing a socially optimal point along the ROC curve.

Under reasonable assumptions, the socially optimal approach is to alert those who are significantly more likely to infect others than is the average member of the population not already in quarantine or isolation [3]; for the NHS COVID-19 app, notified individuals were 2- to 20-fold more likely to subsequently report a positive test [41]. Under the assumption that some form of regulation maintains a geometric mean of approximately R_t=1 while waiting for a vaccine, the total value of both manual contact tracing and digital notification schemes is highest at low case prevalence [3] when fewer individuals need to quarantine to achieve the same population benefit. Digital contact tracing also becomes more important when social restrictions are few, and interactions among strangers are thus more common. Unfortunately, in many jurisdictions, attention to all pandemic measures tended to rise and fall with case counts rather than fluctuating between a focus on contact tracing during lulls and on population-wide measures when case counts were high. This interfered with effective planned implementation.

One obstacle to good risk analysis was public health guidance that ignored evidence for airborne transmission in favor of the droplet theory that physical distancing of 2 m (or 1.5 m or 6 feet) was effective protection. As a result, significant technological research focused on calibrating EN Bluetooth settings to a threshold of 2 m or similar [42-45]. A number of unpublished analyses suggest close to superimposable ROC curves, regardless of Bluetooth settings. Superior distance assessment via Bluetooth could likely have been achieved using all 3 Bluetooth channels instead of 1 or by using the median attenuation instead of the mean (decisions made within the EN application programming interface [API] and hence not available to app developers). Ultra wideband or ultrasound would assess the distance more precisely than Bluetooth. However, none of these tweaks change the fact that distance only moderately predicts infection risk, and proportionate effort was not invested in integration with predictors such as local carbon dioxide levels. Nor were users provided with information such as the time or location of exposure, although this information was present on their phone. This prevented users from integrating information about masking and ventilation into their personal assessment of the risk of infection. Not allowing this, even with the permission of the primary case, is another example of privacy being put before autonomy.

After coming up with *good enough* Bluetooth settings, a better approach would have been to deploy an app that collected data both on the parameters of exposure and on whether it was followed by infection. These data could have informed the risk settings better than any experiment on Bluetooth-distance relationships. We note that a protocol might flag an individual as high risk on the basis of an exposure other than the one that infected them; real-world calibration will include such cases, potentially increasing impact beyond that of causal connections.

An ongoing process of data collection and risk calibration would have taught us much about transmission, for example, any changes to the timing of infectiousness as new strains appeared and as individuals acquired prior immunity. This is important because exposure dose, and hence infection risk, depends not only on the physical characteristics and duration of an encounter but also on how infectious the primary case is at that moment. The main data used to assess this were the date of symptom onset or, if asymptomatic, the date of the first positive test. The infectiousness window was initially estimated to run from approximately 2 days before to ≥5 days after symptom onset [46]. Unfortunately, a bug in the code of the original analysis concealed the earlier onset of infectiousness. Although this was rapidly identified and corrected [47,48], public health guidelines, describing the days for which the primary case should be considered infectious for contact tracing purposes, did not change accordingly. Nor did they modify the contact tracing window following the emergence of new SARS-CoV-2 variants, nor for reinfection or breakthrough cases. To re-estimate this window, they would need to repeat the same type of intensive manual contact tracing studies; this did not occur. Other information about infectiousness could, in principle, come from the Ct count (which was not reported in standard laboratory SARS-CoV-2 polymerase chain reaction testing protocols) or connection to a superspreading event, but neither was used in practice.

EN version 1 allowed the app that is installed on the smartphone of the primary case to assign 1 of 8 levels of infectiousness to each day during which others might have been exposed. This integer is the only metadata regarding the primary case that is available to the exposed contact. EN version 2 enabled some improvements to risk analysis by lifting the previous 30-minute cap on exposure reported by the contact's phone but reduced the number of levels of infectiousness to 2 (although Germany repurposed other bits of information to reclaim the use of 6 [49]). When asked to make the infectiousness metadata easier to use for both risk analysis and scientific study, specifically to propagate the timing of exposure relative to symptom onset or test date, from the server to the exposed contact, Apple replied that they could not do so because of a vulnerability whose details had been worked out by Google but which they could not remember.

Beyond EN, some jurisdictions acknowledged airborne transmission to the point of performing not just proximity tracing but also presence tracing. In corresponding presence tracing schemes, people going to venues such as restaurants would either write down their name and contact details or use an app to check in by scanning a QR code. For Singapore's SafeEntry, they could also use a Bluetooth-based sensor (to detect a TraceTogether app or token) or scan a barcode. With a low-risk threshold, for example, a jurisdiction with a zero COVID-19 policy, all individuals who were present at the same venue as an infectious individual could be alerted.

A better strategy for a higher risk threshold is to focus on *cluster busting*, that is, prioritizing those venues at which at least 1 transmission event is already known to have occurred. These generally correspond to gatherings at which the primary case was highly infectious, ventilation was poor, and there were other risk factors such as speaking, singing, and exercising and lack of mask use. There are compelling arguments for the effectiveness of such *backward tracing* for pathogens such as SARS-CoV-2 [50-52]. This requires cases to be linked.

A decentralized protocol such as EN could have contributed to cluster busting, had the app of a user who was first exposed then tested positive been allowed to upload the cryptographic key responsible for triggering the exposure. EN chose to keep this key sequestered within the operating system, with no option to notify the server. If a key associated with a transmission event were uploaded, the infectiousness score of that key could then be significantly upgraded on the server, triggering follow-up notifications with stronger wording as well as additional notifications following briefer exposures to the possible superspreading event. If users were additionally given the option of sharing their identity with public health authorities, paired with the cryptographic key that served to link them, then both members of the pair could be prioritized for interviews and the location of the superspreading event identified to assist manual contact tracers.

Failure Point 4: Test Access

No form of contact tracing, digital or manual, will work unless we are able to identify who is likely infected and, thus, whose contacts should be notified and traced. Although some countries such as South Korea did better, in many countries, SARS-CoV-2 tests were slow to roll out, testing capacity was quickly overwhelmed, test shortages persisted for a surprisingly long time, and not all symptomatic individuals sought testing [53-55]. It is obviously critical to improve pandemic preparedness in this regard, independent of its relationship with digital notification or tracing.

Improving the allocation of scarce testing resources by identifying which individuals were most likely to test positive was a primary use case for India's Aarogya Setu, with the highest-risk individuals having positive predictive values >40%. To achieve this, Aarogya Setu used a sophisticated risk algorithm to trace not just direct contacts but also contacts of contacts and so on (multiple-degree tracing) in a risk-consistent manner, faster than testing could keep up [56]. In EN, functionality for contacts of contacts was limited, relied on self-reporting, and was never used. Another important factor for allocating scarce testing resources is utility for onward contact tracing and notification, for example, by abandoning samples more than a certain number of days past symptom onset to rapidly turn around tests whose results will be more actionable.

For some pandemics, there may be one or more hallmark symptoms (such as loss of taste and smell) that are sufficiently distinctive to warrant presumptive diagnosis. In other cases, the symptoms might be suggestive but less definitive. Strategies for triggering notifications given contact with an unverified case, in the absence of test results, overlap with the next point of failure, namely, how to manage the verification of positive test results.

The CoEpi app [57] attempted to launch on the basis of symptoms alone using core Bluetooth (it was designed before EN). It was rejected by the Apple Store on the basis of a lack of public health authority involvement, with the suggestion to seek public health authority support from a “city or county.” Its resubmission with explicit support and endorsement from the public health department of a county in Washington state (from a different nonprofit account because Apple suspended the primary developer’s account without explanation) was also rejected. Apple representatives told developers that unwritten and nonpublic rules stood in the way and that it was the desire of Apple to not have “competing” approaches.

When test turnarounds take several days, but each day is critical, it might be helpful to issue *preliminary* notifications on the basis of exposure to an exposed, symptomatic individual awaiting test results to be converted later to a confirmed exposure. The same principle can be expanded deeper into a social network. We note that recursive protocols can incidentally achieve some of the functionality of backward contact tracing, showing up as multiple social network paths via all attendees to the index case.

Failure Point 5: Triggering Notifications

Following a positive diagnosis for SARS-CoV-2 in a primary case, speed is essential to notify secondary cases before they transmit onward to others. The simplest option is to allow any app user to self-attest that they are infected. An issue with this is the potential for abuse. For example, before an election, a coordinated set of individuals could intentionally socialize widely in a setting that tends to vote in a particular way, for example, a college campus, and then falsely report positive diagnoses timed to trigger quarantine on the day of the election. This was a specific concern for the US November 2020 elections. Disruption of essential workers at core infrastructure might also be achieved through targeted attack.

In this light, and to ensure trust in the system, all EN jurisdictions launched with systems that ensured that only verified positive test results could be used to trigger notifications. Many jurisdictions initially relied on manual contact tracers to issue a time-sensitive secure code over the phone to individuals who tested positive. This led to extremely high failure rates, given that case investigation was overwhelmed. EN imposed extra work on case investigators in issuing verification codes but provided nothing back to them in return, for example, by helping link cases into transmission chains, as could be done by the protocol tweak described earlier. Even when the code was successfully delivered and entered into an app, it tended to occur with a significant delay. This defeated much of the purpose of a digital scheme, whose motivation was to make contact tracing faster [58]. Primary care can similarly become overwhelmed, and in a controlled, nonoverwhelmed setting, voluntary adherence even once a code is provided was only 64% [14]. In contrast, Singapore’s solution was fully integrated with both testing and manual contact tracing, speeding up the latter from 4 days to <2 days [59] and switching to primarily automated alerts only after case counts increased during the Delta wave in late 2021.

Google and Apple insisted on having only 1 EN app per country or, in the case of the United States, 1 per state. However, test and trace programs are run by counties and tribal nations in the United States, by provinces in Canada, by cantons in Switzerland, and so on. This disconnect did not encourage buy-in by the local public health authorities charged with distributing secure codes; app design choices were made by a different level of government than the level responsible for implementing the test and trace policies.

Presumably, Google and Apple preferred to limit the number of relationships they needed to maintain and of EN apps whose code they needed to review. However, this could also have been solved via a flexible global app (or an ecosystem of several at least partially interoperable apps); Google and Apple would only need to deal with 1 or several app developers, who in turn would deal with the customizations requested by the various jurisdictions. Such an ecosystem began to spontaneously appear through players such as NearForm, PathCheck, and WeHealth. A bottleneck in many US states was the slow process of government procurement to pay such players. One reason many US states opted for ENX was that because no payment was required, it bypassed delays in the procurement process, despite the fact that running ENX and associated verification code distribution still generally required a state public health authority to internally dedicate full-time staff. With the advent of ENX, Google and Apple ended up needing to maintain more relationships than they would have with private-sector middlemen.

Given the difficulties in providing verification codes by phone, most US implementations, beginning with the state of Colorado, shifted to a system based on SMS text messages. Positive test results reported to the state were collated, and cases were then, in batches, each sent an SMS text message with a deep link that acted as a verification code. Vague language and a few decoy SMS text messages were used to comply with the Health Insurance Portability and Accountability Act. There was still some posttesting delay associated with reporting and collating, some SMS text messages were caught by spam filters, some were lost to the SMS delivery network during peak use periods, and the system was confusing for recipients who had never heard of EN. However, it was faster and had a higher success rate than having a manual contact tracer issue a code over the phone. The ratio of #claims/#cases (measuring a combination of failure points 1 and 5) went up, for example, from 1.8% to 9.6% in the state of Washington (with the caveat that both are upper bounds on success because the total neglects delays in receiving codes and includes codes claimed by individuals who installed EN only after receiving the deep link) [60].

An alternative way to issue verification codes was to integrate them with testing, that is, with the health care system rather than the public health system. The best solution is to make the app into a test result delivery system, that is, to link test samples to QR codes, with the app knowing the ID of its user's test and checking a server for matches. This shortens the time from a sample testing positive in the laboratory to notifications being triggered. It still requires the user to check their notifications for a positive test result and to consent to have their contacts notified. In Singapore, the median time from SMS text message notification of test results to consent to upload data was <30 minutes, with a consent rate of 70% to 80%. EN was eventually modified to allow an even better preauthorization workflow, whereby consent could be given at the time of testing, and notifications were sent as soon as the app became aware of the positive test result, with no further user input required. This system was used in Germany. Germany also offered free supervised antigen tests; the unvaccinated required a recently verified negative result to permit access to some venues, and positive results could be verified by the same system for use in EN.

Integration with testing was easier for the NHS COVID-19 app than for many others, given an already centralized system. However, it was also possible elsewhere; Germany integrated its app with >10,000 different test providers. The speed and convenience of obtaining a polymerase chain reaction test result back through the app helped prompt download, although some users opted only to use this feature and not to trigger notifications. It is important to investigate the cause of this refusal, for example, did it stem from distrust regarding anonymity, were people trying to avoid pushing friends and family into uncompensated quarantine, or was it simply an instinctive "no" under time pressure in a stressful situation, for which a different workflow might have elicited a different decision?

Two EN pilots at US universities explored integration with campus testing programs, which provided rapid test turnaround before adequate testing became available to the general public. The University of Alabama linked test results to phone numbers and then allowed individuals to use their phone number as their verification code through the use of a 1-way hash. The University of Arizona provided verification codes as part of the web-based portal from which the test results were distributed, with 25% of cases claiming a verification code [13]. As part of an (abortive) expansion to the state at large, API integrations were created not only with the state's largest test providers (LabCorp, SonoraQuest, and others) but also with Doximity, an app believed to be used by 70% of the state's physicians. Surveys show higher trust in health care providers than in public health authorities [29,61], making physicians' involvement potentially useful.

In retrospect, all these schemes were too complex to work reliably from the outset, at least in most jurisdictions. Better might have been to allow self-attestation of positive test results or even self-attestation on the basis of symptoms alone at times during which tests are in short supply (refer to the *Failure Point 5: Triggering Notifications* section). To avoid the potential for malicious use, unverified reports could trigger notifications with different messaging: not requiring quarantine, merely warning the recipient, and including the fact that the person who exposed them had self-attested to their infected or symptomatic status. Once postexposure quarantine was relaxed and home testing became the norm, many EN jurisdictions switched to self-attestation.

All the same difficulties arise with presence tracing. Many QR code check-in schemes required manual contact tracers to identify locations of interest to trigger app notifications. Similar to other forms of manual contact tracing, including pen-and-paper check-in at venues, this process was often overwhelmed. Although fully automated systems would have been faster and more reliable, they were not permitted for apps that also ran EN. New Zealand implemented this functionality early on, that is, users who tested positive were able to upload the set of QR codes they had scanned to public health authorities. However, when the NHS COVID-19 app attempted to follow suit, its app update was rejected [62], and New Zealand was also forced to make changes. Google and Apple allowed EN app users to store a local copy of their check-in history on their phone and to read it over the phone to manual contact tracers, but they did not allow an opt-in upload button, which is again a rejection of user autonomy on the grounds of privacy. They did permit an alternative automated QR code check-in system in the German EN app that was designed to never be linked to identities; however, because this system did not satisfy laws in the many German states requiring the collection of identity information, this limited its adoption.

Finally, many systems assumed that once an individual tested positive, they would enter isolation such that only past contacts and not future contacts would need to be notified. This turned out to be overly optimistic. We recommend that future systems anticipate nonadherence, with daily prompts of "Did you succeed in isolating today?" leading to the option to upload new data to anonymously notify that day's contacts, in addition to contacts from the initial upload.

Failure Point 6: Behavior Change

To stem transmission, notification needs to result in behavioral change by secondary cases. Rates of quarantine can be low, estimated at 28% for asymptomatic individuals in Norway adhering for at least 1 day [63], and 11% in the United Kingdom for quarantine adherence on all recommended days [64] (albeit much higher in the app-using subset [65]). Quarantine adherence was 40% among the app-using subpopulation in the Netherlands [61]. Adherence increases with trust in the government's response to a pandemic [66].

Short of outright coercion, quarantine adherence might be improved by paying people, whether paid directly by the government or via employer mandates. Adherence might also be higher if, out of respect for their time, rigorous risk analysis (including the use of negative tests) were used to reduce quarantine duration to a minimum, with a risk threshold set on a rational basis. Some European countries gave quarantine pay to individuals traced by contact tracers but not those notified by EN. It is important to make such choices rationally, that is, to estimate the positive predictive value of an EN with given characteristics and to treat individuals with the same risk of infecting others similarly, regardless of the mode of risk detection.

There can still be substantial utility to issuing low-exposure warnings to individuals whose positive predictive value does not warrant quarantine. Short of full quarantine, more modest behavior changes also help stem transmission, for example, testing and isolating if positive [67], being more alert for symptoms and isolating if they appear, mask wearing, or avoidance of large groups and individuals who are vulnerable to severe disease. However, adherence to even these more modest requests is also far from universal, for example, in the state of Washington, only 40% of the subset of EN recipients who responded to a survey intended to get tested and 67% intended to watch for symptoms; rates were 58% and 84% on the smaller subset that also replied to a follow-up survey about actual behaviors [68]. With Paxlovid or other antivirals most effective when taken early, we note that even if notification fails to stem transmission, it can be of direct benefit to an individual who makes no changes other than monitoring for symptoms, then testing and treating if positive. Adoption rates (points of failure 1 and 2) might be higher if this direct benefit were stressed rather than relying on altruistic motives.

Even without complexities such as recommending quarantine for some but not all exposures, effective communication of *next steps* at a low reading grade level is surely important for adherence. In contrast, some EN implementations launched with quarantine recommendations that did not specify an end date, leaving users to assume that it began with the date of notification receipt rather than the unknown date of exposure. Other recommendations gave the date of exposure and left the user to perform the calculation of the quarantine end date.

Although the purpose of this piece is to present lessons learned, the better to inform responses to the next airborne pandemic, we note that the SARS-CoV-2 pandemic remains unpredictable in its evolution and that EN has not yet been sunsetted in all jurisdictions. In that light, we note that the end of quarantine does not mean the end of EN. If notification caused individuals to mask, avoid those most vulnerable to severe disease, or test then isolate if positive, then the technology would still be doing something useful, should sufficiently low rates of failure at stages 1 to 5 be achieved.

Governance for Effectiveness

To contain a pandemic, that is, to achieve $R_t<1$ through digital tracing and notification, an implementation must keep all 6 failure rates low, for example, <20%. This will require good governance, for which it can be instructive to learn from the best implementations thus far. The NHS and Germany had among the best EN implementations, even though their failure rates remained far higher than this benchmark. Singapore's non-EN implementation was more successful, with a failure rate of <5% for failure points 1 and 2 and acceptably low failure rates for 4 and 6. We hope that more scientific data might eventually emerge from it.

It is notable that both the German and NHS COVID-19 app projects engaged not only software developers, privacy experts, and applied public health practitioners but also well-respected academic epidemiologists with significant track records of directly related research and that these epidemiologists had substantial (but not sole) influence on decision-making. They might be in the best position to balance ambition of scope (to achieve $R_t<1$ not "we did something") with realistic expectations about how things will play out on the ground. This follows a more general pattern during the COVID-19 pandemic in which applied public health institutions such as the Centers for Disease Control and Prevention performed poorly, but many universities and research institutions such as the Robert Koch Institut performed relatively well, as did biosecurity-run initiatives such as Operation Warp Speed, and the pharmaceutical industry when given a good incentive structure.

The fact that iPhone-to-iPhone low-energy Bluetooth communication did not initially run in the background meant that some response from Apple was needed for effective Bluetooth-based proximity tracing and notification using iPhones. Once that response took the form of EN, countries that did not capitulate to Apple's extensive conditions for access often ended up focusing on presence tracing and notification systems instead (eg, France and Australia), unless they used a gossip protocol (Singapore) or issued alternative hardware (Singapore) or had few iPhones in their country (India). Apple's ability to exert power over EN protocol design and over API access substantially restricted the scope for innovation among individual apps. Whatever the best form of governance is, few would argue that it is best done behind the closed doors of a technology company. Given that Apple exerted its power to dictate terms and limit innovation, it is a tragedy that Apple did not use its power to enact an opt-out system. For example, the NHS COVID-19 app is estimated to have saved 10,000 lives in its first year [41], and it is estimated that for every 1% increase in uptake, it could have reduced cases by a further 1% to 2% [2].

Future success will require the measurement of failure rates at all 6 points and rapid on-the-fly adaptations to improve them. Ideally, we would invest now in these technologies, perhaps within island nations or other close-knit communities as test cases, to iteratively improve systems as part of pandemic preparedness while at the same time attempting to reduce SARS-CoV-2 and potentially influenza transmission in the short term. An iteratively improved technology is more likely to be successfully deployed should a new pathogen begin transmitting between humans, one that combines Severe Acute Respiratory Syndrome–1 or influenza A virus subtype H5N1 mortality with SARS-CoV-2 presymptomatic transmission. In the absence of current investment, we hope that this document will help kickstart the design of effective strategies at such a time.

Acknowledgments

The authors thank Ryan Barrett, Justus Benzler, Dana Lewis, Ron Rivest, Curran Schiefelbein, and Tina White for their helpful comments on this manuscript.

Conflicts of Interest

JM and JIMP have an equity interest in WeHealth Solutions PBC, which distributed an Exposure Notification app to Arizona and Bermuda. AG is the chief executive officer of Forward Momentum, LLC, an industry-agnostic consulting firm that may at times serve clients in the software, hardware, or public health space. All other authors declare no conflicts of interest.

References

1. Loh PS. Flipping the perspective in contact tracing. arXiv Preprint posted online October 8, 2020. [FREE Full text]
2. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D, et al. The epidemiological impact of the NHS COVID-19 app. Nature 2021 Jun 12;594(7863):408-412. [doi: 10.1038/s41586-021-03606-z] [Medline: 33979832]
3. Petrie J, Masel J. The economic value of quarantine is higher at lower case prevalence, with quarantine justified at lower risk of infection. J R Soc Interface 2021 Sep 08;18(182):20210459 [FREE Full text] [doi: 10.1098/rsif.2021.0459] [Medline: 34493093]
4. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
5. O'Neill PH. No, coronavirus apps don’t need 60% adoption to be effective. MIT Technology Review. 2020 Jun 5. URL: https://www.technologyreview.com/2020/06/05/1002775/covid-apps-effective-at-less-than-60-percent-download/ [accessed 2023-11-20]
6. Clover J. Find my app: everything to know. MacRumors. 2022 Feb 16. URL: https://www.macrumors.com/guide/find-my/ [accessed 2022-08-10]
7. Singapore G. TraceTogether-only SafeEntry to start from 1 June 2021; only TraceTogether app or token will be accepted for SafeEntry check-in. Smart Nation Singapore. 2021 Apr 22. URL: https://www.smartnation.gov.sg/media-hub/press-releases/tos-to-start-1-june-2021 [accessed 2023-11-20]
8. SB05: How many active users does the Corona-Warn-App have? Corona-Warn-App. 2022. URL: https://www.coronawarn.app/en/science/2022-03-03-science-blog-5/ [accessed 2023-03-02]
9. Kendall M. NHS COVID-19 app statistics. Oxford Big Data Institute: Pathogen Dynamics Group. URL: https://bdi-pathogens.shinyapps.io/NHS-COVID-19-app-statistics/ [accessed 2023-07-01]
10. Sun S, Shaw M, Moodie EE, Ruths D. COVID alert application effectiveness. GitHub. URL: https://github.com/druths/published-resources/raw/main/reports/covid-alert-assessment/COVIDAlertReport-Content-Final.pdf [accessed 2023-11-20]
11. Burdinski A, Brockmann D, Maier BF. Understanding the impact of digital contact tracing during the COVID-19 pandemic. PLOS Digit Health 2022 Dec 6;1(12):e0000149 [FREE Full text] [doi: 10.1371/journal.pdig.0000149] [Medline: 36812611]
12. Farronato C, Iansiti M, Bartosiak M, Denicolai S, Ferretti L, Fontana R. How to get people to actually use contact-tracing apps. Harvard Business Review. 2020 Jul 15. URL: https://hbr.org/2020/07/how-to-get-people-to-actually-use-contact-tracing-apps [accessed 2023-11-20]
13. Masel J, Shilen A, Helming B, Rutschman J, Windham G, Pogreba-Brown K, et al. Quantifying meaningful usage of a SARS-CoV-2 exposure notification app on the campus of the University of Arizona. medRxiv Preprint posted online October 11, 2021. [FREE Full text] [doi: 10.1101/2021.02.02.21251022]
14. Rodríguez P, Graña S, Alvarez-León EE, Battaglini M, Darias FJ, Hernán MA, RadarCovidPilot Group, et al. A population-based controlled experiment assessing the epidemiological impact of digital contact tracing. Nat Commun 2021 Jan 26;12(1):587 [FREE Full text] [doi: 10.1038/s41467-020-20817-6] [Medline: 33500407]
15. Al-Haboubi M, Exley J, Allel K, Erens B, Mays N. One year of digital contact tracing: who was more likely to install the NHS COVID-19 app? Results from a tracker survey in England and Wales. Digit Health 2023 Feb 28;9. [doi: 10.1177/20552076231159449]
16. White paper on Singapore’s response to COVID-19: lessons for the next pandemic. Government of Singapore. 2023 Mar 8. URL: https://www.gov.sg/-/media/gov/covid-19-white-paper/publication/white_paper_on_singapore_response_to_covid19_080323.pdf [accessed 2023-11-20]
17. Biddle N, Gray M, Sollis K. The use of QR codes to identify COVID-19 contacts and the role of data trust and data privacy. Australian National University. 2021. URL: https://csrm.cass.anu.edu.au/research/publications/use-qr-codes-identify-covid-19-contacts-and-role-data-trust-and-data-privacy [accessed 2023-11-20]
18. Willingness to install the contact tracing "Luca app" in Germany 2021. Statista Research Department. 2022 Nov 10. URL: https://www.statista.com/statistics/1241849/luca-app-contact-tracing-smartphone-installation-willingness-germany/ [accessed 2023-02-16]

19. Galaski J. Why was Germany's Covid contact tracing app barely used by health authorities? Liberties. 2022 Feb 22. URL: https://www.liberties.eu/en/stories/luca-app/44032 [accessed 2023-11-20]
20. Jones K, Thompson R. To use or not to use a COVID-19 contact tracing app: mixed methods survey in Wales. JMIR Mhealth Uhealth 2021 Nov 22;9(11):e29181 [FREE Full text] [doi: 10.2196/29181] [Medline: 34698645]
21. Prakash AV, Das S. Explaining citizens' resistance to use digital contact tracing apps: a mixed-methods study. Int J Inf Manage 2022 Apr;63:102468 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2021.102468] [Medline: 36540570]
22. Trkman M, Popovič A, Trkman P. The roles of privacy concerns and trust in voluntary use of governmental proximity tracing applications. Gov Inf Q 2023 Jan;40(1):101787. [doi: 10.1016/j.giq.2022.101787]
23. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of acceptance of COVID-19 proximity tracing apps in Switzerland: panel survey analysis. JMIR Public Health Surveill 2021 Jan 06;7(1):e25701 [FREE Full text] [doi: 10.2196/25701] [Medline: 33326411]
24. Thomas R, Michaleff ZA, Greenwood H, Abukmail E, Glasziou P. Concerns and misconceptions about the Australian Government's COVIDSafe app: cross-sectional survey study. JMIR Public Health Surveill 2020 Nov 04;6(4):e23081 [FREE Full text] [doi: 10.2196/23081] [Medline: 33048826]
25. Fox G, Clohessy T, van der Werff L, Rosati P, Lynn T. Exploring the competing influences of privacy concerns and positive beliefs on citizen acceptance of contact tracing mobile applications. Comput Human Behav 2021 Aug;121:106806 [FREE Full text] [doi: 10.1016/j.chb.2021.106806]
26. Verbeek PP. Ethische analyse van de COVID-19 notificatie-app ter aanvulling op bron en contactonderzoek GGD deel 2: analyse door burgers en professionals. ECP Platform voor de InformatieSamenleving. 2020 Sep 10. URL: https://ecp.nl/wp-content/uploads/2020/11/rapport-begeleidingsethiek-coronamelder.pdf [accessed 2023-11-20]
27. Min CH. Government remains most trusted institution in Singapore, amid global trend of societal polarisation: survey. Channel News Asia. 2023 Mar 15. URL: https://www.channelnewsasia.com/singapore/edelman-trust-barometer-2023-government-business-polarisation-inequality-3345571 [accessed 2023-11-20]
28. Hamzah F. Most Singaporeans feel more united than before COVID-19 pandemic; other countries see increased division: survey. Channel News Asia. 2022 Aug 12. URL: https://www.channelnewsasia.com/singapore/covid-19-united-divided-pandemic-handling-management-vaccines-pew-2876451 [accessed 2023-11-20]
29. Liccardi I, Alekseyev J, Woltz VL, McLean JE, Zurko ME. Public willingness to engage with COVID-19 contact tracing, quarantine, and exposure notification. Public Health Rep 2022;137(2_suppl):90S-95S [FREE Full text] [doi: 10.1177/00333549221125891] [Medline: 36255241]
30. Lueks W, Benzler J, Bogdanov D, Kirchner G, Lucas R, Oliveira R, et al. Toward a common performance and effectiveness terminology for digital proximity tracing applications. Front Digit Health 2021 Aug 5;3:677929 [FREE Full text] [doi: 10.3389/fdgth.2021.677929] [Medline: 34713149]
31. Kraemer MU, Pybus OG, Fraser C, Cauchemez S, Rambaut A, Cowling BJ. Monitoring key epidemiological parameters of SARS-CoV-2 transmission. Nat Med 2021 Nov 08;27(11):1854-1855. [doi: 10.1038/s41591-021-01545-w] [Medline: 34750555]
32. Singer P, Masel J. How (not) to fight COVID-19. Project Syndicate. 2020 Sep 3. URL: https://www.project-syndicate.org/commentary/tech-solutions-to-targeting-covid19-quarantine-by-peter-singer-and-joanna-masel-2020-08 [accessed 2023-11-20]
33. States must assess the effectiveness of tracing applications, while complying with data protection standards. Parliamentary Assembly. 2023 Jan 25. URL: https://pace.coe.int/en/news/8954/states-must-assess-the-effectiveness-of-tracing-applications-while-complying-with-data-protection-standards [accessed 2023-03-18]
34. Vigliarolo B. Apple app transparency changes bring in the ad bucks... for Apple. The Register. 2022 Sep 6. URL: https://www.theregister.com/2022/09/06/apple_app_transparency_ad_revenues/ [accessed 2023-11-20]
35. McGee P. Apple’s privacy changes create windfall for its own advertising business. Financial Times. 2021 Oct 16. URL: https://www.ft.com/content/074b881f-a931-4986-888e-2ac53e286b9d [accessed 2023-11-20]
36. Miller C. Apple takes on data brokers and auctions with new ‘privacy on iPhone’ ad campaign. 9To5Mac. 2022 May 18. URL: https://9to5mac.com/2022/05/18/apple-privacy-on-iphone-new-ad/ [accessed 2023-11-20]
37. Troncoso C, Bogdanov D, Bugnion E, Chatel S, Cremers C, Gürses S, et al. Deploying decentralized, privacy-preserving proximity tracing. Commun ACM 2022 Aug 19;65(9):48-57. [doi: 10.1145/3524107]
38. Two reasons why Singapore is sticking with TraceTogether’s protocol. Government Technology Agency Singapore. 2020 Jun 29. URL: https://www.tech.gov.sg/media/technews/two-reasons-why-singapore-sticking-with-tracetogether-protocol [accessed 2023-04-13]
39. OpenTT-Android. GitHub. URL: https://github.com/GovTechSG/OpenTT-Android/blob/master/app/src/main/java/sg/gov/tech/bluetrace/streetpass/StreetPassScanner.kt [accessed 2023-04-13]
40. Coronavirus: EU interoperability gateway goes live, first contact tracing and warning apps linked to the system. EU Monitor. 2020 Oct 19. URL: https://www.eumonitor.eu/9353000/1/j9vvik7m1c3gyxp/vld1i7ydtgyg [accessed 2023-03-04]
41. Kendall M, Tsallis D, Wymant C, Di Francia A, Balogun Y, Didelot X, et al. Epidemiological impacts of the NHS COVID-19 app in England and Wales throughout its first year. Nat Commun 2023 Feb 22;14(1):858 [FREE Full text] [doi: 10.1038/s41467-023-36495-z] [Medline: 36813770]

42. Wilson AM, Aviles N, Petrie JI, Beamer PI, Szabo Z, Xie M, et al. Quantifying SARS-CoV-2 infection risk within the Google/Apple exposure notification framework to inform quarantine recommendations. Risk Anal 2022 Jan 21;42(1):162-176 [FREE Full text] [doi: 10.1111/risa.13768] [Medline: 34155669]
43. Meyer S, Windisch T, Perl A, Dzibela D, Marzilger R, Witt N. Contact tracing with the exposure notification framework in the German Corona-Warn-App. In: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN). 2021 Presented at: International Conference on Indoor Positioning and Indoor Navigation (IPIN); November 29-December 02, 2021; Lloret de Mar, Spain. [doi: 10.1109/ipin51156.2021.9662591]
44. Schiefelbein CN, St-Germain J, Gervin RCJ, Mazzola SL. Bluetooth low energy (BLE) data collection for COVID-19 exposure notification. Massachusetts Institute of Technology Lincoln Laboratory. 2022 Apr 13. URL: https://www.ll.mit.edu/r-d/publications/bluetooth-low-energy-ble-data-collection-covid-19-exposure-notification [accessed 2023-11-20]
45. Lovett T, Briers M, Charalambides M, Jersakova R, Lomax J, Holmes C. Inferring proximity from Bluetooth low energy RSSI with unscented Kalman smoothers. arXiv Preprint posted online July 9, 2020. [FREE Full text]
46. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020 May 15;26(5):672-675. [doi: 10.1038/s41591-020-0869-5] [Medline: 32296168]
47. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Author correction: temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020 Sep;26(9):1491-1493 [FREE Full text] [doi: 10.1038/s41591-020-1016-z] [Medline: 32770170]
48. Ashcroft P, Huisman JS, Lehtinen S, Bouman JA, Althaus CL, Regoes RR, et al. COVID-19 infectivity profile correction. Swiss Med Wkly 2020 Aug 5;150(3132):w20336 [FREE Full text] [doi: 10.4414/smw.2020.20336] [Medline: 32757177]
49. Klingbeil T. cwa-documentation. GitHub. URL: https://github.com/corona-warn-app/cwa-documentation/blob/master/images/risk_calculation/client_interpretation.pdf [accessed 2023-11-20]
50. Endo A, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Leclerc QJ, Knight GM, Medley GF, Atkins KE, et al. Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks. Wellcome Open Res 2020 Mar 31;5:239 [FREE Full text] [doi: 10.12688/wellcomeopenres.16344.3] [Medline: 33154980]
51. Bradshaw WJ, Alley EC, Huggins JH, Lloyd AL, Esvelt KM. Bidirectional contact tracing could dramatically improve COVID-19 control. Nat Commun 2021 Jan 11;12(1):232 [FREE Full text] [doi: 10.1038/s41467-020-20325-7] [Medline: 33431829]
52. Raymenants J, Geenen C, Thibaut J, Nelissen K, Gorissen S, Andre E. Empirical evidence on the efficiency of backward contact tracing in COVID-19. Nat Commun 2022 Aug 13;13(1):4750 [FREE Full text] [doi: 10.1038/s41467-022-32531-6] [Medline: 35963872]
53. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020 Aug 22;396(10250):535-544 [FREE Full text] [doi: 10.1016/S0140-6736(20)31483-5] [Medline: 32645347]
54. Pullano G, Di Domenico L, Sabbatini CE, Valdano E, Turbelin C, Debin M, et al. Underdetection of cases of COVID-19 in France threatens epidemic control. Nature 2021 Feb 21;590(7844):134-139. [doi: 10.1038/s41586-020-03095-6] [Medline: 33348340]
55. Villalobos C. SARS-CoV-2 infections in the world: an estimation of the infected population and a measure of how higher detection rates save lives. Front Public Health 2020 Sep 25;8:489 [FREE Full text] [doi: 10.3389/fpubh.2020.00489] [Medline: 33102412]
56. Guttal V, Krishna S, Siddharthan R. Risk assessment via layered mobile contact tracing for epidemiological intervention. medRxiv Preprint posted online May 01, 2020. [FREE Full text] [doi: 10.1101/2020.04.26.20080648]
57. CoEpi. GitHub. URL: https://github.com/co-epi [accessed 2023-11-20]
58. Kretzschmar ME, Rozhnova G, Bootsma MC, van Boven M, van de Wijgert JH, Bonten MJ. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. Lancet Public Health 2020 Aug;5(8):e452-e459. [doi: 10.1016/s2468-2667(20)30157-2]
59. Chee K. Time for contact tracing cut by more than half with TraceTogether, SafeEntry. Strait Times. URL: https://www.straitstimes.com/singapore/health/time-for-contact-tracing-cut-by-more-than-half-with-tracetogether-safeentry [accessed 2023-11-20]
60. Segal C, Zhang Z, Karras BT, Revere D, Zane G, Baseman JG. Early epidemiological evidence of public health value of WA notify, a smartphone-based exposure notification tool: modeling COVID-19 cases averted in Washington State. medRxiv Preprint posted online June 18, 2021. [FREE Full text] [doi: 10.1101/2021.06.04.21257951]
61. Ebbers W. Wat volgt er na een melding van de CoronaMelder? Eindrapportage van de evaluatie van de CoronaMelder-app van een kans op besmetting, een kwantitatief onderzoek. Erasmus School of Social and Behavioural Sciences. 2021. URL: https://www.eur.nl/media/94469 [accessed 2023-11-20]
62. Kelion L. NHS Covid-19 app update blocked for breaking Apple and Google's rules. BBC News. 2021 Apr 12. URL: https://www.bbc.com/news/technology-56713017 [accessed 2023-11-20]

63. Steens A, Freiesleben de Blasio B, Veneti L, Gimma A, Edmunds W, Van Zandvoort K, et al. Poor self-reported adherence to COVID-19-related quarantine/isolation requests, Norway, April to July 2020. Euro Surveill 2020 Sep;25(37):2001607 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.37.2001607] [Medline: 32945254]
64. Smith LE, Potts HW, Amlôt R, Fear NT, Michie S, Rubin GJ. Adherence to the test, trace, and isolate system in the UK: results from 37 nationally representative surveys. BMJ 2021 Mar 31;372:n608 [FREE Full text] [doi: 10.1136/bmj.n608] [Medline: 33789843]
65. Coronavirus and self-isolation after being in contact with a positive case in England. Office for National Statistics. 2021 Sep 8. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/coronavirusandselfisolationafterbeingincontactwithapositivecaseinengland [accessed 2023-11-20]
66. Wright L, Steptoe A, Fancourt D. Predictors of self-reported adherence to COVID-19 guidelines. A longitudinal observational study of 51,600 UK adults. Lancet Reg Health Eur 2021 May;4:100061 [FREE Full text] [doi: 10.1016/j.lanepe.2021.100061] [Medline: 33997831]
67. Ferretti L, Wymant C, Nurtay A, Zhao L, Hinch R, Bonsall D, et al. Modelling the effectiveness and social costs of daily lateral flow antigen tests versus quarantine in preventing onward transmission of COVID-19 from traced contacts. medRxiv Preprint posted online August 08, 2021. [doi: 10.1101/2021.08.06.21261725]
68. Baseman JG, Karras BT, Revere D. Engagement in protective behaviors by digital exposure notification users during the COVID-19 pandemic, Washington State, January-June 2021. Public Health Rep 2022 Aug 02;137(2_suppl):96S-100S [FREE Full text] [doi: 10.1177/00333549221110301] [Medline: 35915982]

Abbreviations

API: application programming interface
EN: Exposure Notification
ENX: Exposure Notification Express
NHS: National Health Service
ROC: receiver operating characteristic

Edited by A Mavragani, T Sanchez; submitted 01.06.23; peer-reviewed by M Gasmi, Y Pachankis, V von Wyl, M Kendall; comments to author 30.06.23; revised version received 06.07.23; accepted 24.10.23; published 04.12.23.

Please cite as:
Masel J, Petrie JIM, Bay J, Ebbers W, Sharan A, Leibrand SM, Gebhard A, Zimmerman S
Combatting SARS-CoV-2 With Digital Contact Tracing and Notification: Navigating Six Points of Failure
JMIR Public Health Surveill 2023;9:e49560
URL: https://publichealth.jmir.org/2023/1/e49560
doi:10.2196/49560
PMID:38048155

Viewpoint

Social Media, Public Health Research, and Vulnerability: Considerations to Advance Ethical Guidelines and Strengthen Future Research

Philip M Massey[1], MPH, PhD; Regan M Murray[2], MPH; Shawn C Chiang[3], MPH, PhD; Alex M Russell[4], PhD; Michael A Yudell[5], MPH, PhD

[1]Department of Community Health Sciences, Fielding School of Public Health, UCLA, Los Angeles, CA, United States
[2]College of Education and Health Professions, University of Arkansas, Fayetteville, AR, United States
[3]Department of Health Behavior, Texas A&M University, College Station, TX, United States
[4]Recovery Research Institute, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States
[5]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

Corresponding Author:
Philip M Massey, MPH, PhD
Department of Community Health Sciences
Fielding School of Public Health
UCLA
Box 951772, Suite 16-035 CHS
650 Charles E. Young Dr. South
Los Angeles, CA, 90095
United States
Phone: 1 310 825 5308
Email: pmassey@ucla.edu

Abstract

The purpose of this article is to build upon prior work in social media research and ethics by highlighting an important and as yet underdeveloped research consideration: how should we consider vulnerability when conducting public health research in the social media environment? The use of social media in public health, both platforms and their data, has advanced the field dramatically over the past 2 decades. Applied public health research in the social media space has led to more robust surveillance tools and analytic strategies, more targeted recruitment activities, and more tailored health education. Ethical guidelines when using social media for public health research must also expand alongside these increasing capabilities and uses. Privacy, consent, and confidentiality have been hallmarks for ethical frameworks both in public health and social media research. To date, public health ethics scholarship has focused largely on practical guidelines and considerations for writing and reviewing social media research protocols. Such ethical guidelines have included collecting public data, reporting anonymized or aggregate results, and obtaining informed consent virtually. Our pursuit of the question related to vulnerability and public health research in the social media environment extends this foundational work in ethical guidelines and seeks to advance research in this field and to provide a solid ethical footing on which future research can thrive.

(JMIR Public Health Surveill 2023;9:e49881) doi:10.2196/49881

KEYWORDS
research ethics; social media; vulnerable populations; public health; ethical guidelines; algorithms; manipulation

Introduction

In October 2021, the Senate Subcommittee on Consumer Protection, Product Safety, and Data Security convened a hearing titled "Protecting kids online: testimony from a Facebook whistleblower" [1]. While this hearing focused on the protection of children, highlighting the amplification of content related to eating disorders targeting teenagers as well as the platform's "blind eye" toward age verification, broader takeaways included the platform's ability to create and cultivate a manipulative environment on social media. In part, this senate hearing, which examined questions about the inner workings of the social media ecosystem, was spurred by the 2016 election scandals [2], the spread of misinformation during the COVID-19 pandemic [3], and the mental health crisis that has most notably affected our young people [4]. While public health efforts have played a critical role in combatting misinformation on social media as well as addressing the mental health crisis, little has been done to examine the fundamental question that prompted congress' interest in social media: can social media create a manipulative environment that makes us vulnerable to undue influence? The short answer is yes, as documented by congressional hearings [1,5-7], independent research [8-11], and investigative journalism [12-15].

Algorithms and Vulnerability

The driver behind the discussion of social media manipulation appears to hinge on one key idea: algorithms. While research has examined how algorithms create an inescapable environment and thus an extensive network primed for digital discrimination, systematic bias, unethical targeting, and misinformation and disinformation campaigns [16-18], we still do not know enough about how algorithms function, what goals inform these functions, or the impact of algorithms on public health research and practice.

While algorithms and their ethical concerns have entered various dialogues, from congressional hearings to investigative journalism, little discussion has taken place in the field of public health. To date, public health ethics has focused largely on practical guidelines for writing and reviewing social media research protocols. Such ethical guidelines have included collecting public data, reporting anonymized or aggregate results, and obtaining informed consent virtually, to name a few. Susser [19] extends these considerations by discussing the role of manipulation, autonomy, and bias for digitally targeted public health interventions. Ethical considerations when using social media for public health research must expand alongside our increased understanding of how the social media environment functions, specifically concerning the presence of algorithms and how these may contribute to issues related to vulnerability. Privacy, consent, and confidentiality are important hallmarks for ethical frameworks in social media research [20], but we must move beyond these foundational questions and begin unpacking how our research and practice may or may not contribute to and benefit from the manipulative environments that many experience on social media.

The use of social media in public health has advanced the field dramatically over the last 2 decades. Traditional public health methods in surveillance and outbreak investigation [21], approaches in health education and promotion [22], and strategies in policy advocacy [23] and community organizing [24] have all been applied, refined, and adapted for the social media environment. Social media research, or the process of using social media data to conduct quantitative or qualitative research, ranging from observational data collection to experimental designs, is an invaluable tool in public health research and practice and continues to expand both in terms of how it is conducted and where it takes place [20]. Decades of applied public health research in the social media space have led to more robust surveillance tools and analytic strategies, more targeted recruitment activities, and more tailored health education [21]. However, more must be done to advance our understanding of algorithms and how they may ultimately compromise data for public health research and practice.

The potential use of compromised information in public health research is relevant to both observational and intervention research on social media. For instance, when conducting observational research that collects public data from individual accounts related to a specific topic (eg, vaccine safety), how do we disentangle the extent to which content was shared due to behind-the-scenes platform manipulation (ie, due to algorithms that place content in a user's thread with the goal of increasing interaction and engagement and with little regard to the content itself)? Adding to this complexity is the presence of social media bots, or automated programs, that artificially amplify or spread content based on an array of goals (eg, to spread disinformation, notify of emergencies, share advertisements, or aggregate news articles) [25].

Similarly, for intervention research, do we know how interacting with social media content produced for a research study may influence the platform's tailoring of future content for that same individual (eg, joining a vaccine-related research study may place the individual at greater risk of being exposed to future highly engaging vaccine information, which is more often than not misinformation)? Intervention research often takes place in closed or "private" groups on social media, making it easier to moderate and monitor the content directly administered by the study; however, the closed group exists within the larger ecosystem of the platform, and we do not yet know how participation in a research study may impact content exposure outside of that closed environment. Furthermore, when using social media to recruit study participants, we must also consider the potential collection and use of compromised information. For example, how much do targeted recruitment ads rely on interactions by users with content that was manipulated or artificially placed in a user's thread to solicit interaction? These questions highlight very practical ways in which our seemingly innocuous research activities (eg, public data collection or targeted recruitment) may in fact be interacting with and relying on compromised information, thus creating a scenario where researchers are relying on information that is, to an undetermined extent, artificially manipulated by opaque algorithmic intervention, contributing to vulnerabilities that have yet to be considered by public health research taking place on social media.

Moving Forward

We wish to move the field of public health and social media research forward by posing the following question: how should we consider vulnerability when conducting public health research in the social media environment? We pose this question not to limit or stifle public health research in the social media environment; in fact, quite the contrary—we pose this question to activate our collective understanding and consciousness to strengthen research in this environment in part due to the ever-changing social media landscape. At its core, the primary goal of the social media ecosystem is to keep users on the platform, interacting and engaging with content, for as long as possible, often at any cost [16]. Our hope is that the issues raised here will do the following: (1) contribute to frameworks that more clearly describe how vulnerability, much like privacy, consent, and confidentiality, is an essential concept for conducting ethical social media research; and (2) establish the need for partnerships with social media companies, supported through federal resources, that will facilitate collaborative yet independent research led by academic partners. While algorithms themselves are not nefarious, it is the intent behind the use of algorithms, and the goals and parameters set forth to use the algorithms in particular ways, that evokes concerns surrounding manipulation that may contribute to and enhance various vulnerabilities.

Conflicts of Interest

None declared.

References

1. Subcommittee: protecting kids online: testimony from a Facebook whistleblower. U.S. Senate Committee on Commerce, Science, and Transportation. URL: https://www.commerce.senate.gov/2021/10/protecting%20kids%20online:%20testimony%20from%20a%20facebook%20whistleblower [accessed 2023-09-13]
2. Rosenberg M, Confessore N, Cadwalladr C. How Trump consultants exploited the Facebook data of millions. The New York Times. URL: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html [accessed 2023-09-21]
3. Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol Sci 2020 Jul;31(7):770-780 [FREE Full text] [doi: 10.1177/0956797620939054] [Medline: 32603243]
4. United S, Public HS, Office OTSG. Social media and youth mental health. The U.S. Surgeon General's Advisory. URL: https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf [accessed 2023-12-18]
5. Algorithms and amplification: how social media platforms' design choices shape our discourse and our minds. U.S. Senate Committee on the Judiciary. URL: https://www.judiciary.senate.gov/committee-activity/hearings/algorithms-and-amplification-how-social-media-platforms-design-choices-shape-our-discourse-and-our-minds [accessed 2023-09-13]
6. Facebook, social media privacy, and the use and abuse of data. U.S. Senate Committee on the Judiciary. URL: https://www.judiciary.senate.gov/committee-activity/hearings/facebook-social-media-privacy-and-the-use-and-abuse-of-data [accessed 2023-09-13]
7. Bond S, Allyn B. Whistleblower tells Congress that Facebook products harm kids and democracy. NPR. URL: https://www.npr.org/2021/10/05/1043207218/whistleblower-to-congress-facebook-products-harm-children-and-weaken-democracy [accessed 2023-09-13]
8. González-Bailón S, Lazer D, Barberá P, Zhang M, Allcott H, Brown T, et al. Asymmetric ideological segregation in exposure to political news on Facebook. Science 2023 Jul 28;381(6656):392-398. [doi: 10.1126/science.ade7138] [Medline: 37499003]
9. Nyhan B, Settle J, Thorson E, Wojcieszak M, Barberá P, Chen AY, et al. Like-minded sources on Facebook are prevalent but not polarizing. Nature 2023 Aug;620(7972):137-144 [FREE Full text] [doi: 10.1038/s41586-023-06297-w] [Medline: 37500978]
10. Guess AM, Malhotra N, Pan J, Barberá P, Allcott H, Brown T, et al. How do social media feed algorithms affect attitudes and behavior in an election campaign? Science 2023 Jul 28;381(6656):398-404. [doi: 10.1126/science.abp9364] [Medline: 37498999]
11. Guess AM, Malhotra N, Pan J, Barberá P, Allcott H, Brown T, et al. Reshares on social media amplify political news but do not detectably affect beliefs or opinions. Science 2023 Jul 28;381(6656):404-408. [doi: 10.1126/science.add8424] [Medline: 37499012]
12. Horwitz J. Facebook says its rules apply to all. Company documents reveal a secret elite that's exempt. The Wall Street Journal. URL: https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353?mod=article_inline [accessed 2023-09-13]
13. Wells G, Horwitz J, Seetharaman D. Facebook knows Instagram is toxic for teen girls, company documents show. The Wall Street Journal. URL: https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=hp_lead_pos7&mod=article_inline [accessed 2023-09-13]
14. Hagey K, Horwitz J. Facebook tried to make its platform a healthier place. It got angrier instead. The Wall Street Journal. URL: https://www.wsj.com/articles/facebook-algorithm-change-zuckerberg-11631654215?mod=article_inline [accessed 2023-09-13]
15. Scheck J, Purnell N, Horwitz J. Facebook employees Ffag drug cartels and human traffickers. The company's response is weak, documents show. The Wall Street Journal. URL: https://www.wsj.com/articles/facebook-drug-cartels-human-traffickers-response-is-weak-documents-11631812953?mod=article_inline [accessed 2023-09-13]
16. Fisher M. The chaos machine: the inside story of how social media rewired our minds and our world. Tijdschr voor Communicatiewetenschap 2022;51(4):427-430. [doi: 10.5117/tcw2023.4.006.damm]

17. Milan S, Treré E. The rise of the data poor: the COVID-19 pandemic seen from the margins. Soc Media Soc 2020 Jul;6(3):2056305120948233 [FREE Full text] [doi: 10.1177/2056305120948233] [Medline: 34192035]
18. Gillett R, Stardust Z, Burgess J. Safety for whom? Investigating how platforms frame and perform safety and harm interventions. Soc Media Soc 2022 Dec 15;8(4):205630512211443 [FREE Full text] [doi: 10.1177/20563051221144315]
19. Susser D. Ethical considerations for digitally targeted public health interventions. Am J Public Health 2020 Oct;110(S3):S290-S291. [doi: 10.2105/ajph.2020.305758]
20. Moreno MA, Goniu N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):708-713 [FREE Full text] [doi: 10.1089/cyber.2012.0334] [Medline: 23679571]
21. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EHY, Olsen JM, et al. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. PLoS One 2015;10(10):e0139701 [FREE Full text] [doi: 10.1371/journal.pone.0139701] [Medline: 26437454]
22. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
23. Jackson M, Brennan L, Parker L. The public health community's use of social media for policy advocacy: a scoping review and suggestions to advance the field. Public Health 2021 Sep;198:146-155. [doi: 10.1016/j.puhe.2021.07.015] [Medline: 34428607]
24. Brady SR, Young JA, McLeod DA. Utilizing digital advocacy in community organizing: lessons learned from organizing in virtual spaces to promote worker rights and economic justice. J Community Pract 2015 May 18;23(2):255-273. [doi: 10.1080/10705422.2015.1027803]
25. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: twitter bots and Russian trolls amplify the vaccine debate. Am J Public Health 2018 Oct;108(10):1378-1384. [doi: 10.2105/AJPH.2018.304567] [Medline: 30138075]

Edited by A Mavragani; submitted 12.06.23; peer-reviewed by J Cappella, J Alvarez-Galvez, E Miyagi; comments to author 03.08.23; revised version received 23.09.23; accepted 17.11.23; published 29.12.23.

Please cite as:
Massey PM, Murray RM, Chiang SC, Russell AM, Yudell MA
Social Media, Public Health Research, and Vulnerability: Considerations to Advance Ethical Guidelines and Strengthen Future Research
JMIR Public Health Surveill 2023;9:e49881
URL: https://publichealth.jmir.org/2023/1/e49881
doi:10.2196/49881
PMID:38157235

Review

Monitoring School Absenteeism for Influenza-Like Illness Surveillance: Systematic Review and Meta-analysis

Tim K Tsang[1], PhD; Xiaotong Huang[1], MPH; Yiyang Guo[1], MPH; Eric H Y Lau[1,2], PhD; Benjamin J Cowling[1,2], PhD; Dennis K M Ip[1], MD

[1]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Laboratory of Data Discovery for Health Limited, Hong Kong Science and Technology Park, Hong Kong, China (Hong Kong)

Corresponding Author:
Tim K Tsang, PhD
WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health
Li Ka Shing Faculty of Medicine
The University of Hong Kong
7 Sassoon Road
Pokfulam
Hong Kong
China (Hong Kong)
Phone: 852 39179715
Email: timtsang@connect.hku.hk

Abstract

Background: Influenza causes considerable disease burden each year, particularly in children. Monitoring school absenteeism has long been proposed as a surveillance tool of influenza activity in the community, but the practice of school absenteeism could be varying, and the potential of such usage remains unclear.

Objective: The aim of this paper is to determine the potential of monitoring school absenteeism as a surveillance tool of influenza.

Methods: We conducted a systematic review of the published literature on the relationship between school absenteeism and influenza activity in the community. We categorized the types of school absenteeism and influenza activity in the community to determine the correlation between these data streams. We also extracted this correlation with different lags in community surveillance to determine the potential of using school absenteeism as a leading indicator of influenza activity.

Results: Among the 35 identified studies, 22 (63%), 12 (34%), and 8 (23%) studies monitored all-cause, illness-specific, and influenza-like illness (ILI)–specific absents, respectively, and 16 (46%) used quantitative approaches and provided 33 estimates on the temporal correlation between school absenteeism and influenza activity in the community. The pooled estimate of correlation between school absenteeism and community surveillance without lag, with 1-week lag, and with 2-week lag were 0.44 (95% CI 0.34, 0.53), 0.29 (95% CI 0.15, 0.42), and 0.21 (95% CI 0.11, 0.31), respectively. The correlation between influenza activity in the community and ILI-specific absenteeism was higher than that between influenza activity in community all-cause absenteeism. Among the 19 studies that used qualitative approaches, 15 (79%) concluded that school absenteeism was in concordance with, coincided with, or was associated with community surveillance. Of the 35 identified studies, only 6 (17%) attempted to predict influenza activity in the community from school absenteeism surveillance.

Conclusions: There was a moderate correlation between school absenteeism and influenza activity in the community. The smaller correlation between school absenteeism and community surveillance with lag, compared to without lag, suggested that careful application was required to use school absenteeism as a leading indicator of influenza epidemics. ILI-specific absenteeism could monitor influenza activity more closely, but the required resource or school participation willingness may require careful consideration to weight against the associated costs. Further development is required to use and optimize the use of school absenteeism to predict influenza activity. In particular, the potential of using more advanced statistical models and validation of the predictions should be explored.

(JMIR Public Health Surveill 2023;9:e41329) doi:10.2196/41329

KEYWORDS

influenza; surveillance; school absenteeism; monitoring; school attendance; influenza-like illness; correlation; trend; pattern; predict; prediction; influenza activity; infection; surveillance tolls

Introduction

Influenza virus causes substantial morbidity and mortality in humans each year on average [1,2]. Influenza activity and hence infections decreased due to public health and social measures of COVID-19 [3]. However, rebound of influenza virus activity is expected [4,5], given the relaxation of public health and social measures [6,7] and low population immunity of influenza [8]. Therefore, monitoring influenza activity is important in the post–COVID-19 era. Common modes of surveillance for influenza activity include the following: (1) sentinel surveillance, in which the consultation rates or the number of influenza-like illness (ILI) in outpatient and private medical practitioner clinics is recorded and (2) laboratory surveillance, in which respiratory specimens are collected and the proportion of positive tests of the influenza virus are recorded.

Children are believed to be a major driver of influenza virus transmission since they have more frequent person-to-person close contacts with low preexisting immunity [9-11]. Therefore, school is a high-risk setting for the transmission of influenza and other respiratory viruses. Preventing transmission in schools could block transmissions to family members and further reduce community epidemics [12]. School surveillance offers an opportunity for the early detection of these viruses [13]. In some regions, student absenteeism surveillance could be in real time, and hence it could have minimal reporting delay, with relative low cost [14,15].

Some studies explore the potential use of school absenteeism to monitoring the influenza activity in the community. In some regions, school absenteeism has been integrated as a part of routine disease surveillance, particularly for respiratory virus infections [16-21], although there are also usages for other diseases [22,23]. However, the component of school absenteeism surveillance could be different, such as the use of all-cause and illness-specific absenteeism. Moreover, their performances are rarely assessed. Therefore, validating and optimizing the performance of the use of school absenteeism on the surveillance of influenza activity would be critical.

In this paper, we conducted a systematic review to collect information on studies that described both school absenteeism surveillance and community surveillance. We summarized the types of school absenteeism surveillance and community surveillance methods to assess the relationship between these 2 surveillances, and the potential use of school absenteeism surveillance to predict influenza activity in the community.

Methods

Definition of School Absenteeism Surveillance and Community Surveillance

School absenteeism surveillance was defined as the time series recording the absent rate per day or week in a school. Absenteeism could be all-cause, specific to ILI, or specific to respiratory illness but not limited to ILI.

Community surveillance was defined as the time series monitoring the influenza activity in a region. There were the following two main types of surveillance: (1) ILI rate or count per day or week and (2) the proportion of laboratory specimens testing positive for influenza virus.

Search Strategy and Selection Criteria

This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Review and Meta-analysis) statement [24]. A standardized search was done in PubMed, Embase, and Web of Science, using the search term “((school AND (absent OR absence OR absenteeism) AND (flu OR influenza))”. The search was done on May 12, 2022, with no language restrictions. Additional relevant articles from the reference sections were also reviewed.

Two authors (XH and YG) independently screened the titles and extracted data from the included studies. Disagreements were resolved by consensus with a third author (TKT). Studies identified from different databases were deduplicated.

Eligible articles were those reporting daily or weekly student absenteeism data and community ILI surveillance in the same region. There was no restriction on the methods used for finding the relationship between school absenteeism and influenza surveillance in the community. Studies without both school absenteeism and community surveillance data were excluded. Articles that met the following were also excluded: (1) the study summarized the findings published elsewhere; (2) the study used data generated from simulation or prediction; (3) the study analyzed the combination of data from studies published elsewhere; or (4) the full text was not available.

Data were extracted from included studies using a standard form, with the following information in following 3 major components: school absenteeism, influenza surveillance in the community, and comparison between school absenteeism and influenza surveillance in community. Each mentioned variable below was a column in the standard form. For school absenteeism, the number of schools, the school size, the grade of students, methods used for school absenteeism surveillance, the types of absenteeism (all-cause, ILI-related, or illness-specific), illness ascertainment methods for illness-related absenteeism, and the treatment of holiday in the analysis were extracted. For influenza surveillance in the community, the extracted data included the types of surveillance (influenza-like illness with or without laboratory test) and the information of laboratory surveillance (influenza strain and the use of test-positive number or proportion). For comparison between school absenteeism and influenza surveillance in the community, information included methods, period and lag time (day/week) of comparison, and estimated correlation coefficients.

Data Analysis

For studies reporting regression coefficients, we used the approach by Rodgers and Nicewander [25] to transform them to Pearson correlation coefficient. We conducted random effects meta-analyses using the inverse variance method and restricted maximum likelihood estimator for heterogeneity to obtain the pooled correlation between school absenteeism and community surveillance [26-29]. Cochran Q test and the I^2 statistic were used to identify and quantify heterogeneity among included studies [30]. An I^2 value of more than 75% indicated high heterogeneity [31]. We conducted subgroup meta-analyses by the types of school absenteeism surveillance (all-cause vs ILI-related vs illness-specific), by the type of community surveillance (ILI vs lab-confirmed), with or without considering the delay effect of school absenteeism on community influenza. Meta-analysis on subgroups was only performed when there were at least 5 estimates. However, it should be noted that as I^2 could be biased in small studies, we only quoted the I^2 estimate when the number of studies was at least 10 [32]. We also conducted metaregressions to explore the impact of the following factors: types of school absenteeism, types of community surveillance, primary school or lower grade in school absenteeism surveillance, use of count data in school absenteeism surveillance, use of weekly data, and study type (prospective vs retrospective) in the analysis.

For studies using qualitative methods, in which only plots or tables of time series of school absenteeism and influenza activity in the community were provided without using any statistical comparison such as estimating correlations, a qualitative description of the relationship between these 2 surveillance systems was extracted. The used terms for the relationship in those articles was summarized. For the studies attempting to predict community surveillance by using school absenteeism, the following information about prediction was extracted: the types of data and its time period used for training for the models for prediction, data type and prediction methods, the type of data for validating the developed models and its time period, and the evaluation method.

Results

Overview

In the systematic review, we identified 3579 studies in our search, and 1144 duplicated articles were excluded (Figure 1). After screening the titles and the abstracts of the remaining articles, we identified 171 studies that may contain relevant information for full-text screening. Among these, 35 articles (Table S1 in Multimedia Appendix 1) met the inclusion criteria and were included in this review [15-19,33-62], of which 9 [19,33,37,41,46,48,54,61,62] and 6 [17,18,40,42,51,53] were prospective and retrospective studies, respectively. In terms of school absenteeism surveillance, 22 [15,16,34-39,41,47-49,51,53-57,59-62], 12 [15,17-19,35-37,44,45,50,54,61], and 8 [33,40,42,43,46,52,58,61] of the studies provided counts on all-cause, ILI-related, and illness-specific absenteeism, respectively. In terms of community surveillance, 25 [16-19,33-37,39-45,47,49-51,55,58-60,62] and 26 [15-17,19,33,34,36-43,46-49,52-54,56,57,59-61] studies provided the number of ILI cases and laboratory-confirmed cases, respectively. Moreover, 24 [15,17,19,33,35-38,40-45,47-49,51,54,56-58,60,62] and 11 [16,18,34,39,46,50,52,53,55,59,61] studies used weekly and daily scale of time series for analysis, respectively. In total, 7060 (we excluded 5 studies that did not provide this information) schools were included in our study. Of these, 2 articles [16,56] did not provide any comparison between school absenteeism and community surveillance; 16 articles [17-19,33,37,40-42,46,48,51,53-55,61,62] with 7703 school years performed quantitative analysis by using statistical method to compute the correlation between school absenteeism and community surveillance; and 19 articles [15,16,34-36,38,39,43-45,47,49,50,52,56-60] with 2123 school years performed qualitative analysis comparing their trends visually by figures or plots. In addition, 6 [17,18,42,46,54,61], 5 [17,18,42,46,61], and 1 [37] articles considered time lag and reported comparisons of school absenteeism and community surveillance with 1-week, 2-week, and 3-week lags, respectively. Moreover, 25 studies excluded holidays in the comparison, while 10 studies did not provide any information. Among the 25 studies that excluded holidays in the comparison, 15 of them [16-18,33,37,45,48-53,55,61,62] used weekly absolute count in school absenteeism surveillance, which may not have accounted for the reduction in denominator due to holidays within the week. Overall, among 16 and 19 studies using quantitative and qualitative analysis, 15 and 17 studies supported the association between school absenteeism and community surveillance, respectively (Table 1).

Figure 1. Process of systematic review. ILI: influenza-like illness.

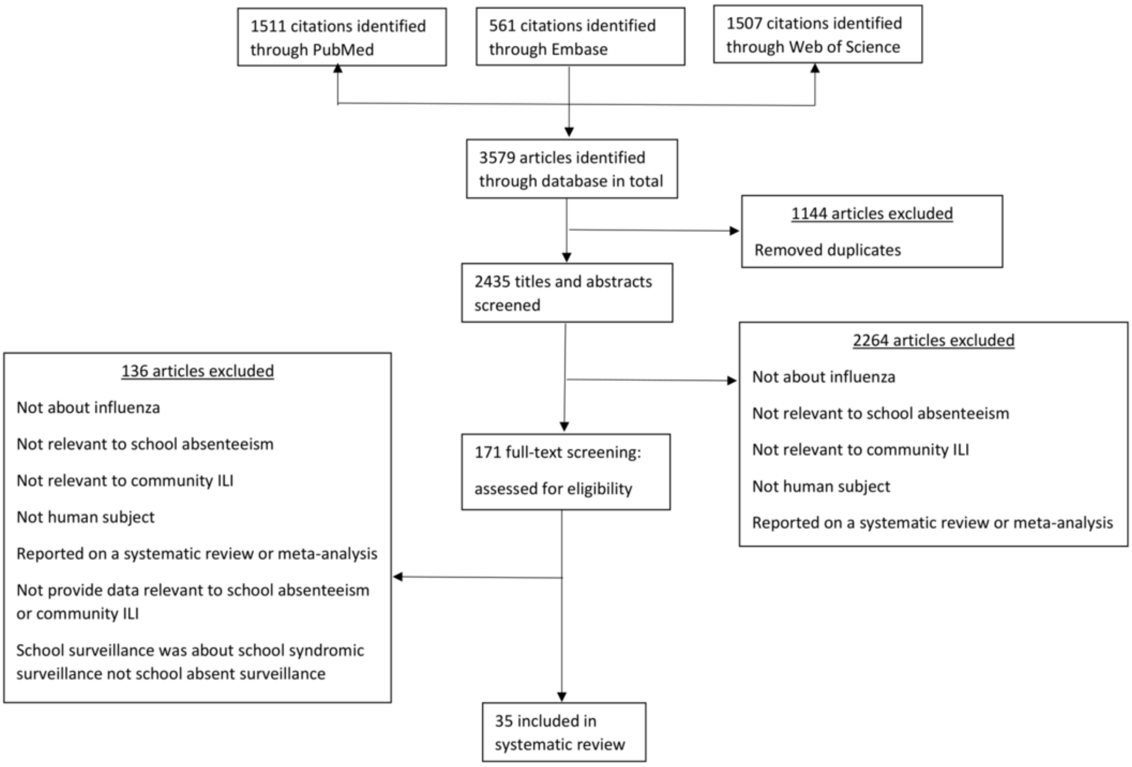


Table 1. Summary of included studies for pooled analysis (n=35).

Characteristics	Studies, n	Reference
Methods to assess correlation		
Quantitative analysis	16	N/A[a]
Regression	4	[33,41,51,62]
Correlation	11	[17-19,37,40,42,46,48,53,54,61]
Other measures	1	[55]
Qualitative analysis by comparing of trends in plots of time series	19	[15,16,34-36,38,39,43-45,47,49,50,52,56-60]
Time lag between school absent and surveillance		
0 week	14	[17-19,33,40-42,46,48,51,53,54,61,62]
1 week	6	[17,18,42,46,54,61]
Others	1	[37]
Association between school absent and surveillance		
By quantitative analysis		
Significance for 0-week lag	6	[17-19,33,40,46]
Significance for 1-week lag	2	[17,18]
Significance for 2-week lag	2	[17,18]
Significance for 3-week lag	1	[37]
Unknown significance	8	[41,42,51,53-55,61,62]
No significant association	1	[48]
By quantitative analysis		
With association	17	[16,34-36,38,39,43-45,47,49,50,56-60]
No detected association	2	[15,52]

[a] N/A: not applicable.

Correlation Between School Absenteeism and Community Surveillance

For quantitative analysis, all studies used regression coefficient or Pearson or Spearman correlation coefficient for evaluating the relationship between school absenteeism and community surveillance, except 1 [63], which used receiver operating characteristic to determine the accuracy of using school absenteeism to predict the occurrence of influenza outbreak, which was not included in our calculation of pooled correlation. Overall, 14 studies [17-19,33,40-42,46,48,51,53,54] provided 33 correlation estimates (Figure 2 [17-19,33,40-42,46,48,51,53,54,61,62] and Table S2 in Multimedia Appendix 1), and 30 estimates supported a positive correlation between school absenteeism and community surveillance [18,42]. The correlation (without lag) between school absences and community surveillance was 0.44 (95% CI 0.34, 0.53), with high heterogeneity (Figure 3). In terms of community surveillance, the correlation between school absences and community surveillance using laboratory-confirmed cases (correlation: 0.45; 95% CI 0.33, 0.57) was higher than using the number of ILI (correlation: 0.41; 95% CI 0.23, 0.59), which was less specific. In terms of school absenteeism, the correlation between community and ILI-specific absenteeism (correlation: 0.51; 95% CI 0.30, 0.73) and illness-specific absenteeism (correlation: 0.43; 95% CI 0.30, 0.57) was higher than all-cause absenteeism (correlation: 0.36; 95% CI 0.11, 0.62). While there was a positive relationship between school absenteeism and the community surveillance with 1-week (correlation: 0.29; 95% CI 0.15, 0.42) and 2-week lag (correlation: 0.21; 95% CI 0.11, 0.31), the positive correlation was smaller compared with those without lag.

In the metaregression (Table 2), we estimated that the correlation between school absenteeism and the community surveillance with 1-week and 2-week lag was 0.19 (95% CI 0.04, 0.34) and 0.26 (95% CI 0.10, 0.41) lower than without lag. We found that the correlation between community surveillance and ILI-related absenteeism was 0.25 (95% CI 0.04, 0.45) higher than that between community surveillance and all-cause absenteeism. The correlation between school absenteeism and laboratory-confirmed community surveillance was 0.17 (95% CI 0.04, 0.30) higher than that between school absenteeism and ILI surveillance. We found that the correlation between school absenteeism and community surveillance from prospective studies was 0.19 (95% CI 0.05, 0.33) higher than from retrospective studies.

Figure 2. The temporal correlation between school absenteeism and influenza activity in community from identified studies. ILI: influenza-like illness; PCR: Polymerase Chain Reaction.

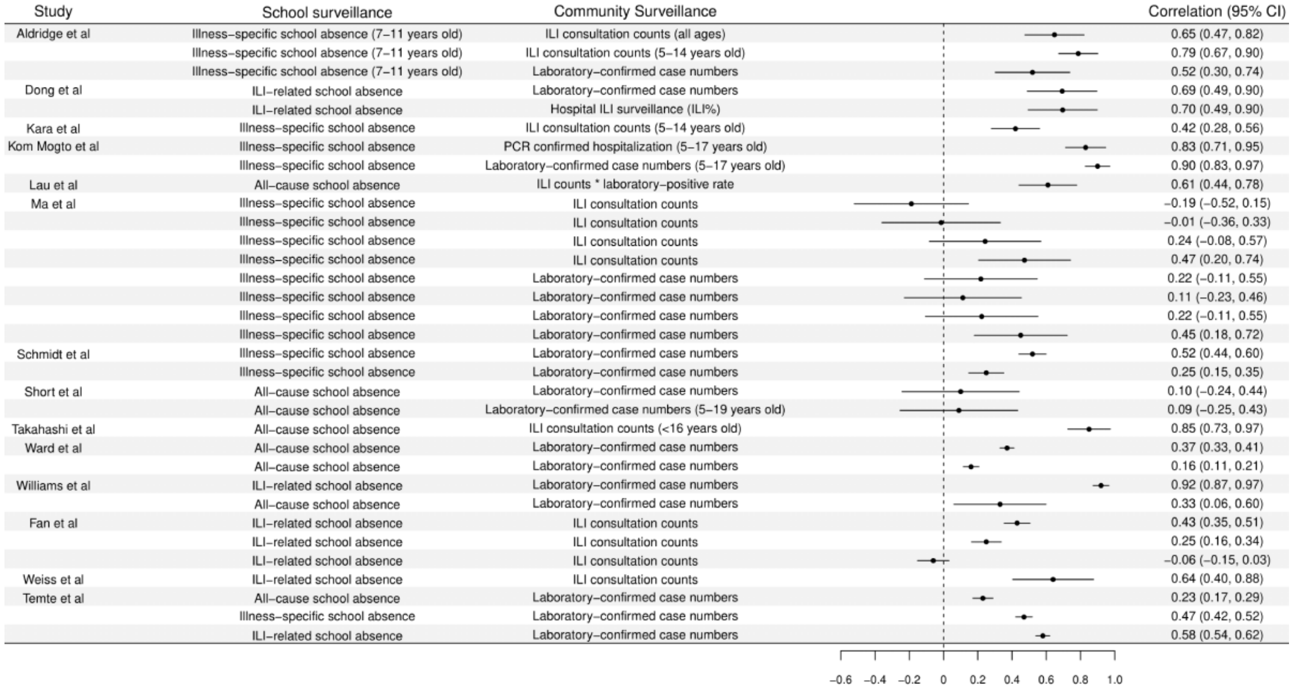
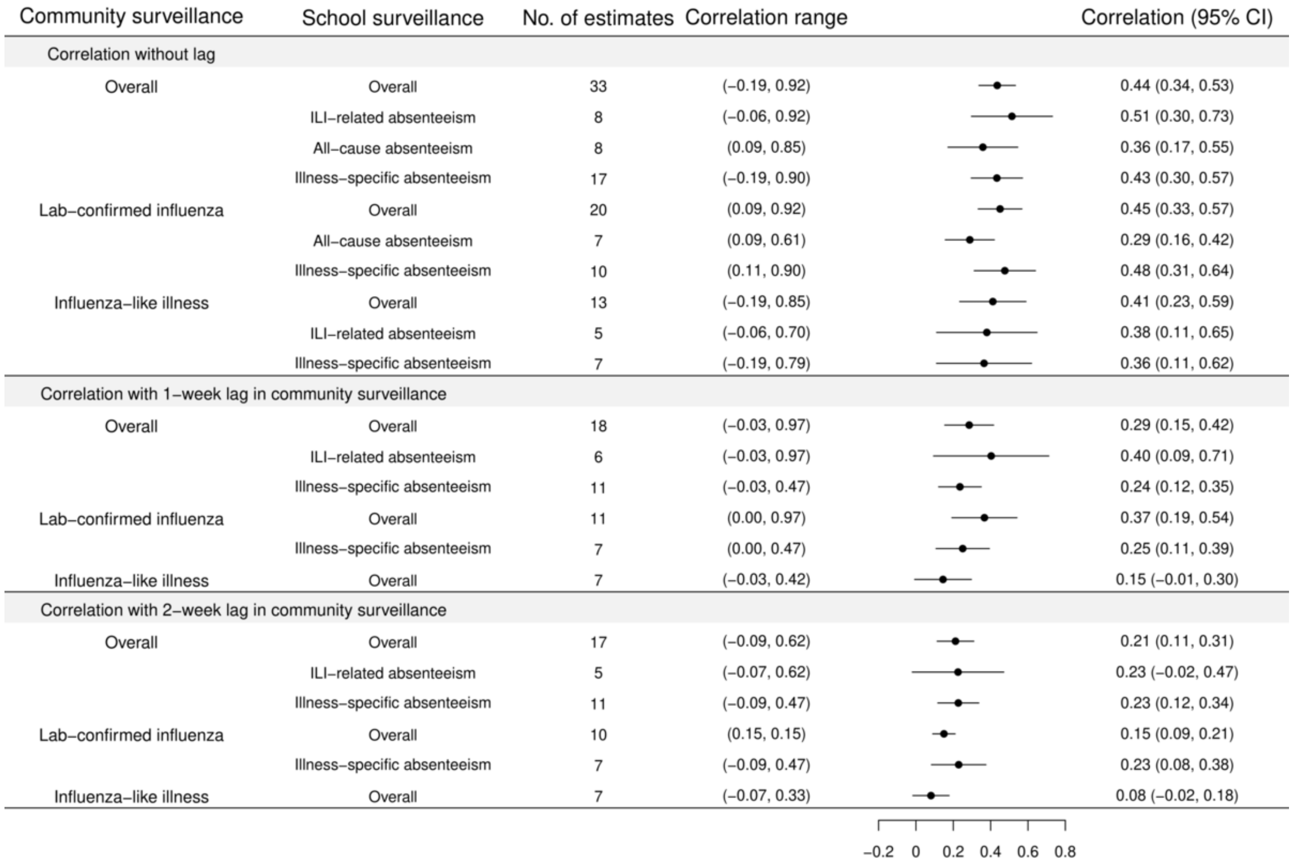


Figure 3. The pooled estimate of temporal correlations between school absenteeism and influenza activity in the community by types of school absenteeism surveillance and types of surveillance of influenza activity in community. ILI: influenza-like illness.



Community surveillance	School surveillance	No. of estimates	Correlation range	Correlation (95% CI)
Correlation without lag				
Overall	Overall	33	(−0.19, 0.92)	0.44 (0.34, 0.53)
	ILI−related absenteeism	8	(−0.06, 0.92)	0.51 (0.30, 0.73)
	All−cause absenteeism	8	(0.09, 0.85)	0.36 (0.17, 0.55)
	Illness−specific absenteeism	17	(−0.19, 0.90)	0.43 (0.30, 0.57)
Lab−confirmed influenza	Overall	20	(0.09, 0.92)	0.45 (0.33, 0.57)
	All−cause absenteeism	7	(0.09, 0.61)	0.29 (0.16, 0.42)
	Illness−specific absenteeism	10	(0.11, 0.90)	0.48 (0.31, 0.64)
Influenza−like illness	Overall	13	(−0.19, 0.85)	0.41 (0.23, 0.59)
	ILI−related absenteeism	5	(−0.06, 0.70)	0.38 (0.11, 0.65)
	Illness−specific absenteeism	7	(−0.19, 0.79)	0.36 (0.11, 0.62)
Correlation with 1−week lag in community surveillance				
Overall	Overall	18	(−0.03, 0.97)	0.29 (0.15, 0.42)
	ILI−related absenteeism	6	(−0.03, 0.97)	0.40 (0.09, 0.71)
	Illness−specific absenteeism	11	(−0.03, 0.47)	0.24 (0.12, 0.35)
Lab−confirmed influenza	Overall	11	(0.00, 0.97)	0.37 (0.19, 0.54)
	Illness−specific absenteeism	7	(0.00, 0.47)	0.25 (0.11, 0.39)
Influenza−like illness	Overall	7	(−0.03, 0.42)	0.15 (−0.01, 0.30)
Correlation with 2−week lag in community surveillance				
Overall	Overall	17	(−0.09, 0.62)	0.21 (0.11, 0.31)
	ILI−related absenteeism	5	(−0.07, 0.62)	0.23 (−0.02, 0.47)
	Illness−specific absenteeism	11	(−0.09, 0.47)	0.23 (0.12, 0.34)
Lab−confirmed influenza	Overall	10	(0.15, 0.15)	0.15 (0.09, 0.21)
	Illness−specific absenteeism	7	(−0.09, 0.47)	0.23 (0.08, 0.38)
Influenza−like illness	Overall	7	(−0.07, 0.33)	0.08 (−0.02, 0.18)

Table 2. Factors affecting the correlation between school absenteeism and community surveillance by metaregression.

Variables	No lag (95% CI)	1-week lag (95% CI)	2-week lag (95% CI)	Overall (0- to 2-week lag; 95% CI)
School absence type				
All-cause absenteeism	Reference	Reference	Reference	Reference
ILI[a]-related absenteeism	0.20 (–0.11, 0.50)	0.39 (–0.11, 0.89)	0.21 (–0.16, 0.57)	0.25 (0.04, 0.45)
Illness-specific absenteeism	0.10 (–0.16, 0.35)	0.16 (–0.32, 0.64)	0.14 (–0.20, 0.47)	0.12 (–0.07, 0.30)
Community surveillance type				
ILI	Reference	Reference	Reference	Reference
Laboratory confirmed	0.08 (–0.14, 0.30)	0.30 (0.05, 0.54)	0.23 (0.03, 0.43)	0.17 (0.04, 0.30)
Study type				
Retrospective study	Reference	Reference	Reference	Reference
Prospective study	0.25 (0.04, 0.45)	0.12 (–0.20, 0.44)	0.07 (–0.19, 0.33)	0.19 (0.05, 0.33)
Lag[b]				
No lag	N/A[c]	N/A	N/A	Reference
1-week lag	N/A	N/A	N/A	–0.19 (–0.34, –0.04)
2-week lag	N/A	N/A	N/A	–0.26 (–0.41, –0.10)
Only primary school and lower grade in school surveillance	–0.10 (–0.53, 0.34)	–0.21 (–0.59, 0.16)	–0.15 (–0.41, 0.11)	–0.14 (–0.36, 0.08)
Using count in school absenteeism (Reference: using rate or proportion)	0.03 (–0.23, 0.29)	–0.16 (–0.50, 0.18)	0.22 (–0.10, 0.54)	0.01 (–0.16, 0.18)
Using weekly scale (Reference: daily scale)	0.11 (–0.15, 0.38)	0.09 (–0.22, 0.41)	–0.07 (–0.33, 0.19)	0.06 (–0.10, 0.22)

[a] ILI: influenza-like illness.

[b] Number of weeks that school absenteeism data were leading community surveillance.

[c] N/A: not applicable.

Qualitative Analysis for the Relationship School Absenteeism and Community Surveillance

In the qualitative analysis, there were 19 articles that reported the relationship between school absenteeism and community surveillance by presenting the time series of these surveillances (Table 3). Of these 19 studies, 2 [16,56] did not compare the relationship in plots of time series. Overall, a similar pattern was reported within school absenteeism and community surveillance in all studies (Table 3), except 2, which reported no detected association [15,52]. Moreover, 6 studies [35,36,43,47,49,50] compared the peak of school absent and peak of epidemics. Among these, 4 [35,36,43,50], 1 [49], and 1 [47] studies found that school absenteeism peaked before, at the same time, and after the peak of community surveillance, respectively. In addition, 15 studies concluded that school absenteeism was in concordance with, coincided with, or was associated with community surveillance, using terms including “trends are coincident” [34,47,59], “associated with” [16,34,36,56,57,60], “concordance with” [38,43,58], “similar trend” [39,44,50], or ‘’mirror” [45].

Table 3. Summary of term count for qualitative results.

Term combined and term used	Study	Comparison	Count
School absenteeism is in concordance with, coincided with, or is associated with community ILI[a]			16
Increases in absences coincided with community-wide ILI outbreaks	Besculides [34], 2005	All-cause absenteeism vs community influenza isolates	1
The second wave of school absenteeism coincided with the second round of community ILI	Schoub [47], 1994	All-cause absenteeism vs laboratory isolations	1
The drop in school absenteeism coincided with the epidemic	McCormick [59], 2010	All-cause absenteeism vs community ILI surveillance or virus-confirmed influenza	1
School absenteeism is associated with community ILI			6
The extent of school absenteeism is associated with the extent of community ILI peak	Besculides [34], 2005	All-cause absenteeism vs community influenza isolates	
School absenteeism is associated with community ILI	Cheng [36], 2012	ILI-related absenteeism vs laboratory isolations	
Influenza activity was reflected by the school absenteeism rates	Chin [57], 1974	All-cause absenteeism vs laboratory-confirmed influenza	
Influenza activity was associated with or reflected by school absenteeism	Rubin [60], 1975	All-cause absenteeism vs community ILI surveillance	
School absenteeism was less timely than laboratory data	Chu [16], 2013	All-cause absenteeism vs laboratory-confirmed influenza	
Close temporal correlation between school absenteeism and the isolation of strains of influenza virus	Olson [56], 1980	All-cause absenteeism vs laboratory-confirmed influenza	
School absenteeism is concordant with community ILI			3
	Jaeger [38], 2011	All-cause absenteeism vs laboratory-confirmed influenza	
	Mook [43], 2007	Illness-defined absenteeism vs community ILI surveillance	
	Lenaway [58], 1995	ILI-related absenteeism vs community ILI surveillance	
Similar trend			3
	Janusz [39], 2011	All-cause absenteeism vs community ILI surveillance	
	Nasrullah [44], 2012	ILI-related absenteeism vs community ILI surveillance	
	Suzue [50], 2012	ILI-related absenteeism vs community ILI surveillance	
School absenteeism mirrored or corresponded to community ILI	Read [45], 2021	Influenza-confirmed absenteeism vs community virus–confirmed influenza	1
Comparison by peak time			6
Peak at the same time	Sigmundsdottir [49], 2010	All-cause absenteeism vs laboratory confirmed cases	1
School absenteeism peak preceded or was ahead of the epidemic peak			4
	Bollaerts [35], 2010	All-cause absenteeism vs community ILI surveillance	
	Cheng [36], 2012	ILI-related absenteeism vs laboratory isolations	
	Mook [43], 2007	Illness-defined absenteeism vs community ILI surveillance or laboratory isolations	
	Suzue [50], 2012	ILI-related absenteeism vs community ILI surveillance	

Term combined and term used	Study	Comparison	Count
Virus isolation commence and peak before the school absenteeism	Schoub [47], 1994	All-cause absenteeism vs laboratory isolations	1
No detected association			2
No related outbreaks were detected, and no peaks were found	Tan [52], 2014	ILI-related absenteeism vs laboratory-confirmed cases	1
Within seasons, cases peaked in winter, whereas county-level absences varied throughout the year	Quandelacy [15], 2021	ILI-related absenteeism or all-cause absenteeism vs community virus–confirmed influenza	1

[a] ILI: influenza-like illness.

Prediction of Community Surveillance From School Surveillance

Six studies attempted to predict community ILI from school absenteeism (Table S3 in Multimedia Appendix 1), 3 of which [15,41,62] predicted the influenza rate (case number), and 3 [18,53,55] predicted the occurrence of outbreaks. All the studies agreed that school absenteeism surveillance was of good use for influenza outbreak detection [15,18,41,53,62], except 1 using all-cause absenteeism [55]. Among these, 4 studies [15,53,55,62] and 2 studies [18,41] used the all-cause absenteeism and ILI-related absenteeism for prediction, respectively. Moreover, 1 [18] and 5 [15,41,53,55,62] studies used mechanistic models (SEIR) and statistical models, respectively, 3 of which [41,53,62] considered the autocorrelation of the time series. In term of evaluation, those studies that predicted the occurrence of outbreaks used receiver operating characteristic curves [18,55] and false-positive rate (1 minus sensitivity) [53]. For those studies that predicted the number of cases, 1 used the model fit [41], 1 used mean absolute error [15], and 1 used Akaike Information Criterion value [62]. Two articles [15,53] conducted out-sample testing to evaluate the performance.

Discussion

Principal Findings

In this study, we summarized the practice of school absenteeism surveillance and their potential use on monitoring and predicting influenza activity in the community. Studies could broadly be classified as quantitative and qualitative. In quantitative studies, most studies used temporal correlation to quantify the relationship between school absenteeism and influenza activity in the community. In qualitative studies, time series between these 2 data streams were compared visually. Overall, we found a moderate correlation between school absenteeism and influenza activity in the community based on quantitative studies. This suggested the potential to use school absenteeism data to monitor or predict influenza activity in the community.

The measure of school absence could be classified to nonspecific (all-cause absent) and specific (illness- or ILI-specific absence). We found that ILI-specific absenteeism had a higher correlation with the community surveillance, compared with all-cause absenteeism, which was consistent with the findings from another review [64]. While using specific measure of school absenteeism could be slightly more accurate [54], the implementation of using more specific absenteeism surveillance in schools should be considered jointly with the associated costs. Those specific measures may require more resources to obtain, such as follow-up of the reasons of absence, which may also jeopardize the timeliness of the school absence data. Moreover, the willingness of school participation may be lower due to the higher requirement of resource; hence the sample size may decrease. Such factors should be considered when deciding the use of all-cause absence or ILI-specific absence to monitor influenza activity. The actual cost of implementation of using more specific absenteeism surveillance in schools was likely different by regions and countries; therefore, our studies could not provide recommendation on this.

A number of studies proposed that one of the values of school absenteeism surveillance was its lead on the traditional surveillance [35,36,43,50]. However, we found that the correlation between school absenteeism and community surveillance without lag was the highest. In addition, the correlation between school absenteeism and 1-week lag of community surveillance was only marginally smaller than community surveillance without lag. Hence, the school absenteeism surveillance could at most lead influenza outbreaks by 1 week. Therefore, despite the moderate correlation between school absenteeism and community surveillance, the use of school absenteeism as a surveillance tool may require further exploration and development of methodology. In particular, we found that the correlations between school absenteeism and community surveillance from prospective studies were higher than from retrospective studies, suggesting that prospective data collection could improve the accuracy by carefully checking the data with timely correction.

More than half of the studies only use qualitative approaches to explore the relationship between school absenteeism and influenza activity in the community, in which only figures of time series of these 2 surveillances were compared visually. Such approaches were relatively subjective, and further quantitative comparison should be performed. Furthermore, almost all quantitative comparisons were based on temporal correlation, which required further steps to determine the usefulness of school absenteeism for monitoring or predicting influenza activity in the community, such as the development of prediction or forecasting tools.

Only 6 out of 35 studies attempted to develop methods to use the school absenteeism data to predicted influenza activity in the community. However, those applied approaches were suboptimal compared with other influenza-forecasting approaches [65,66]; particularly, the potential of more advanced statistical approaches was less explored, as well as the lack of validation. In terms of modeling fitting, only 2 out of 6 studies conducted out-sample evaluation on their forecasting approach to avoid overfitting. In terms of model evaluation, all studies evaluated the point forecast, but none of them evaluated the prediction intervals from those models. The evaluation of interval forecast by some proper score rules would also be important to use the prediction performance [66,67] and to support the accuracy of the predictions. In future studies, using school absenteeism data in more well-developed models should be explored, including mechanistic models (ie, SIR-type) [65,66], and statistical models, such as generalized additive model and random forecast regression, which could particularly handle nonlinear relationship [68].

Limitations

Our study may have some limitations. First, we did not summarize the change of rule on the school attending in the identified studies. For example, many regions may not allow students to attend school when they have fever, particularly during or after the 2009 pandemic influenza outbreaks. Therefore, our study may not be able to evaluate this impact on the relationship between school absenteeism and community surveillance. Second, we did not stratify the analysis by type of reported correlation coefficients (Pearson or Spearman). However, we expected that the direction of correlation should be the same. Third, most studies did not report information of burden or costs of implementing school surveillance. Therefore, it was impossible for us to determine the real-life impact or cost-effectiveness of different types of surveillance. Finally, we could not rule out other potential cofounders in the identified relationship between school absenteeism and influenza activity in the community. For example, information of school or class size was not available.

Conclusion

We found there was a moderate correlation between school absenteeism and influenza activity in the community. We found that the correlation between influenza activity and ILI-specific absenteeism was higher than all-cause absenteeism. However, implementing more specific surveillance in school may require careful consideration, since more resources may be required, and it may have a negative impact on the willingness of school participation. There was potential for using school absenteeism as a surveillance and prediction tool of influenza activity. A further development of methodology was required to use and optimize such usage. In particular, more statistical models should be explored, and the validation of prediction performance is missed in most studies.

Acknowledgments

We thank Can Wang for technical assistance. This study was supported by the Theme-based Research Scheme project no. T11-712/19N from the Hong Kong Government, the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region (grant no. 21200292 and 11101092).

Conflicts of Interest

BJC reports honoraria from Sanofi Pasteur, GSK, Moderna, and Roche. The authors report no other potential conflicts of interest.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 36 KB - publichealth_v9i1e41329_app1.docx]

References

1. Dawood FS, Iuliano AD, Reed C, Meltzer MI, Shay DK, Cheng P, et al. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. Lancet Infect Dis 2012 Sep;12(9):687-695. [doi: 10.1016/S1473-3099(12)70121-4] [Medline: 22738893]
2. Iuliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, Global Seasonal Influenza-associated Mortality Collaborator Network. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. Lancet 2018 Mar 31;391(10127):1285-1300 [FREE Full text] [doi: 10.1016/S0140-6736(17)33293-2] [Medline: 29248255]
3. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Health 2020 May;5(5):e279-e288 [FREE Full text] [doi: 10.1016/S2468-2667(20)30090-6] [Medline: 32311320]
4. Wu K, Wu X, Wang W, Hong L. Epidemiology of influenza under the coronavirus disease 2019 pandemic in Nanjing, China. J Med Virol 2022 May;94(5):1959-1966 [FREE Full text] [doi: 10.1002/jmv.27553] [Medline: 34964514]
5. Ali ST, Lau YC, Shan S, Ryu S, Du Z, Wang L, et al. Prediction of upcoming global infection burden of influenza seasons after relaxation of public health and social measures during the COVID-19 pandemic: a modelling study. Lancet Glob Health 2022 Nov;10(11):e1612-e1622 [FREE Full text] [doi: 10.1016/S2214-109X(22)00358-8] [Medline: 36240828]
6. Sovann LY, Sar B, Kab V, Yann S, Kinzer M, Raftery P, et al. An influenza A (H3N2) virus outbreak in the Kingdom of Cambodia during the COVID-19 pandemic of 2020. Int J Infect Dis 2021 Feb;103:352-357 [FREE Full text] [doi: 10.1016/j.ijid.2020.11.178] [Medline: 33249287]

7. Zheng L, Qi J, Wu J, Zheng M. Changes in Influenza Activity and Circulating Subtypes During the COVID-19 Outbreak in China. Front Med (Lausanne) 2022;9:829799 [FREE Full text] [doi: 10.3389/fmed.2022.829799] [Medline: 35391894]
8. Lee K, Jalal H, Raviotta JM, Krauland MG, Zimmerman RK, Burke DS, et al. Estimating the Impact of Low Influenza Activity in 2020 on Population Immunity and Future Influenza Seasons in the United States. Open Forum Infect Dis 2022 Jan;9(1):ofab607 [FREE Full text] [doi: 10.1093/ofid/ofab607] [Medline: 35024374]
9. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 2008 Mar 25;5(3):e74 [FREE Full text] [doi: 10.1371/journal.pmed.0050074] [Medline: 18366252]
10. Ranjeva S, Subramanian R, Fang VJ, Leung GM, Ip DKM, Perera RAPM, et al. Age-specific differences in the dynamics of protective immunity to influenza. Nat Commun 2019 Apr 10;10(1):1660 [FREE Full text] [doi: 10.1038/s41467-019-09652-6] [Medline: 30971703]
11. Kwok KO, Cowling BJ, Wei VWI, Wu KM, Read JM, Lessler J, et al. Social contacts and the locations in which they occur as risk factors for influenza infection. Proc Biol Sci 2014 Aug 22;281(1789):20140709 [FREE Full text] [doi: 10.1098/rspb.2014.0709] [Medline: 25009062]
12. Tsang TK, Lau LLH, Cauchemez S, Cowling BJ. Household Transmission of Influenza Virus. Trends Microbiol 2016 Feb;24(2):123-133 [FREE Full text] [doi: 10.1016/j.tim.2015.10.012] [Medline: 26612500]
13. Yeng PK, Woldaregay AZ, Solvoll T, Hartvigsen G. Cluster Detection Mechanisms for Syndromic Surveillance Systems: Systematic Review and Framework Development. JMIR Public Health Surveill 2020 May 26;6(2):e11512 [FREE Full text] [doi: 10.2196/11512] [Medline: 32357126]
14. Worby CJ, Chaves SS, Wallinga J, Lipsitch M, Finelli L, Goldstein E. On the relative role of different age groups in influenza epidemics. Epidemics 2015 Dec;13:10-16 [FREE Full text] [doi: 10.1016/j.epidem.2015.04.003] [Medline: 26097505]
15. Quandelacy TM, Zimmer S, Lessler J, Vukotich C, Bieltz R, Grantz KH, et al. Predicting virologically confirmed influenza using school absences in Allegheny County, Pennsylvania, USA during the 2007-2015 influenza seasons. Influenza Other Respir Viruses 2021 Nov;15(6):757-766 [FREE Full text] [doi: 10.1111/irv.12865] [Medline: 34477304]
16. Chu A, Savage R, Whelan M, Rosella LC, Crowcroft NS, Willison D, et al. Assessing the relative timeliness of Ontario's syndromic surveillance systems for early detection of the 2009 influenza H1N1 pandemic waves. Can J Public Health 2013 May 14;104(4):340-347 [FREE Full text] [doi: 10.17269/cjph.104.3721] [Medline: 24044464]
17. Dong X, Boulton ML, Carlson B, Montgomery JP, Wells EV. Syndromic surveillance for influenza in Tianjin, China: 2013-14. J Public Health (Oxf) 2017 Jun 01;39(2):274-281. [doi: 10.1093/pubmed/fdw022] [Medline: 26968483]
18. Fan Y, Wang Y, Jiang H, Yang W, Yu M, Yan W, et al. Evaluation of outbreak detection performance using multi-stream syndromic surveillance for influenza-like illness in rural Hubei Province, China: a temporal simulation model based on healthcare-seeking behaviors. PLoS One 2014 Nov 19;9(11):e112255 [FREE Full text] [doi: 10.1371/journal.pone.0112255] [Medline: 25409025]
19. Kom Mogto CA, De Serres G, Douville Fradet M, Lebel G, Toutant S, Gilca R, et al. School absenteeism as an adjunct surveillance indicator: experience during the second wave of the 2009 H1N1 pandemic in Quebec, Canada. PLoS One 2012;7(3):e34084 [FREE Full text] [doi: 10.1371/journal.pone.0034084] [Medline: 22479531]
20. Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, et al. Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe. JMIR Public Health Surveill 2017 Sep 19;3(3):e66 [FREE Full text] [doi: 10.2196/publichealth.7429] [Medline: 28928112]
21. Donaldson AL, Harris JP, Vivancos R, Hungerford D, Hall I, O'Brien SJ. School Attendance Registers for the Syndromic Surveillance of Infectious Intestinal Disease in UK Children: Protocol for a Retrospective Analysis. JMIR Res Protoc 2022 Jan 20;11(1):e30078 [FREE Full text] [doi: 10.2196/30078] [Medline: 35049509]
22. Adnan M, Gao X, Bai X, Newbern E, Sherwood J, Jones N, et al. Potential Early Identification of a Large Campylobacter Outbreak Using Alternative Surveillance Data Sources: Autoregressive Modelling and Spatiotemporal Clustering. JMIR Public Health Surveill 2020 Sep 17;6(3):e18281 [FREE Full text] [doi: 10.2196/18281] [Medline: 32940617]
23. Lawpoolsri S, Khamsiriwatchara A, Liulark W, Taweeseneepitch K, Sangvichean A, Thongprarong W, et al. Real-time monitoring of school absenteeism to enhance disease surveillance: a pilot study of a mobile electronic reporting system. JMIR Mhealth Uhealth 2014 May 12;2(2):e22 [FREE Full text] [doi: 10.2196/mhealth.3114] [Medline: 25099501]
24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339:b2535 [FREE Full text] [doi: 10.1136/bmj.b2535] [Medline: 19622551]
25. Lee Rodgers J, Nicewander WA. Thirteen Ways to Look at the Correlation Coefficient. The American Statistician 1988 Feb;42(1):59-66. [doi: 10.1080/00031305.1988.10475524]
26. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychological Methods 1998 Dec;3(4):486-504. [doi: 10.1037/1082-989x.3.4.486]
27. Veroniki AA, Jackson D, Bender R, Kuss O, Langan D, Higgins JPT, et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. Res Synth Methods 2019 Mar;10(1):23-43 [FREE Full text] [doi: 10.1002/jrsm.1319] [Medline: 30129707]

28. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999 Oct 30;18(20):2693-2708. [doi: 10.1002/(sici)1097-0258(19991030)18:20<2693::aid-sim235>3.0.co;2-v] [Medline: 10521860]
29. Langan D, Higgins JPT, Jackson D, Bowden J, Veroniki AA, Kontopantelis E, et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Res Synth Methods 2019 Mar;10(1):83-98 [FREE Full text] [doi: 10.1002/jrsm.1316] [Medline: 30067315]
30. Cochran WG. The Combination of Estimates from Different Experiments. Biometrics 1954 Mar;10(1):101. [doi: 10.2307/3001666]
31. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 06;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
32. von Hippel PT. The heterogeneity statistic I(2) can be biased in small meta-analyses. BMC Med Res Methodol 2015 Apr 14;15(1):35 [FREE Full text] [doi: 10.1186/s12874-015-0024-z] [Medline: 25880989]
33. Aldridge RW, Hayward AC, Field N, Warren-Gash C, Smith C, Pebody R, Decipher my Data projectschools. Are School Absences Correlated with Influenza Surveillance Data in England? Results from Decipher My Data-A Research Project Conducted through Scientific Engagement with Schools. PLoS One 2016 Mar 2;11(3):e0146964 [FREE Full text] [doi: 10.1371/journal.pone.0146964] [Medline: 26933880]
34. Besculides M, Heffernan R, Mostashari F, Weiss D. Evaluation of school absenteeism data for early outbreak detection, New York City. BMC Public Health 2005 Oct 07;5(1):105 [FREE Full text] [doi: 10.1186/1471-2458-5-105] [Medline: 16212669]
35. Bollaerts K, Antoine J, Robesyn E, Van Proeyen L, Vomberg J, Feys E, et al. Timeliness of syndromic influenza surveillance through work and school absenteeism. Arch Public Health 2010 Sep 29;68(3):115-120. [doi: 10.1186/0778-7367-68-3-115]
36. Cheng CKY, Cowling BJ, Lau EHY, Ho LM, Leung GM, Ip DKM. Electronic school absenteeism monitoring and influenza surveillance, Hong Kong. Emerg Infect Dis 2012 May;18(5):885-887 [FREE Full text] [doi: 10.3201/eid1805.111796] [Medline: 22516519]
37. Ip DK, Lau EH, So HC, Xiao J, Lam CK, Fang VJ, et al. A Smart Card-Based Electronic School Absenteeism System for Influenza-Like Illness Surveillance in Hong Kong: Design, Implementation, and Feasibility Assessment. JMIR Public Health Surveill 2017 Oct 06;3(4):e67 [FREE Full text] [doi: 10.2196/publichealth.6810] [Medline: 28986338]
38. Jaeger V, Shick-Porter M, Moore D, Grant D, Wolfe V. GotFlu channel: an online syndromic surveillance tool supporting college health practice and public health work. J Am Coll Health 2011 Apr 13;59(5):415-418. [doi: 10.1080/07448481.2010.521961] [Medline: 21500061]
39. Janusz KB, Cortes JE, Serdarevic F, Jones RC, Jones JD, Ritger KA, Chicago H1N1 Epidemiology Team. Influenza-like illness in a community surrounding a school-based outbreak of 2009 pandemic influenza A (H1N1) virus-Chicago, Illinois, 2009. Clin Infect Dis 2011 Jan 01;52(Suppl 1):S94-101 [FREE Full text] [doi: 10.1093/cid/ciq025] [Medline: 21342907]
40. Kar EO, Elliot AJ, Bagnall H, Foord DGF, Pnaiser R, Osman H, et al. Absenteeism in schools during the 2009 influenza A(H1N1) pandemic: a useful tool for early detection of influenza activity in the community? Epidemiol. Infect 2011 Oct 21;140(7):1328-1336. [doi: 10.1017/s0950268811002093]
41. Lau EHY, Cheng CKY, Ip DKM, Cowling BJ. Situational awareness of influenza activity based on multiple streams of surveillance data using multivariate dynamic linear model. PLoS One 2012;7(5):e38346 [FREE Full text] [doi: 10.1371/journal.pone.0038346] [Medline: 22675456]
42. Ma T, Englund H, Bjelkmar P, Wallensten A, Hulth A. Syndromic surveillance of influenza activity in Sweden: an evaluation of three tools. Epidemiol. Infect 2014 Dec 04;143(11):2390-2398. [doi: 10.1017/s0950268814003240]
43. Mook P, Joseph C, Gates P, Phin N. Pilot scheme for monitoring sickness absence in schools during the 2006/07 winter in England: can these data be used as a proxy for influenza activity? Euro Surveill 2007 Dec 01;12(12):E11-E12. [doi: 10.2807/esm.12.12.00755-en] [Medline: 18076857]
44. Nasrullah M, Breiding MJ, Smith W, McCullum I, Soetebier K, Liang JL, et al. Response to 2009 pandemic influenza A H1N1 among public schools of Georgia, United States--fall 2009. Int J Infect Dis 2012 May;16(5):e382-e390 [FREE Full text] [doi: 10.1016/j.ijid.2012.01.010] [Medline: 22424896]
45. Read JM, Zimmer S, Vukotich C, Schweizer ML, Galloway D, Lingle C, et al. Influenza and other respiratory viral infections associated with absence from school among schoolchildren in Pittsburgh, Pennsylvania, USA: a cohort study. BMC Infect Dis 2021 Mar 22;21(1):291 [FREE Full text] [doi: 10.1186/s12879-021-05922-1] [Medline: 33752625]
46. Schmidt W, Pebody R, Mangtani P. School absence data for influenza surveillance: A pilot study in the United Kingdom. Eurosurveillance 2010;15(3):1-6. [doi: 10.2807/ese.15.03.19467-en]
47. Schoub BD, Johnson S, McAnerney J, Blackburn NK. Benefits and limitations of the Witwatersrand influenza and acute respiratory infections surveillance programme. S Afr Med J 1994 Oct;84(10):674-678. [Medline: 7839255]
48. Short VL, Marriott CK, Ostroff S, Waller K. Description and Evaluation of the 2009–2010 Pennsylvania Influenza Sentinel School Monitoring System. Am J Public Health 2011 Nov;101(11):2178-2183. [doi: 10.2105/ajph.2011.300132]
49. Sigmundsdottir G, Gudnason T, Ólafsson Ö, Baldvinsdottir G, Atladottir A, Löve A, et al. Surveillance of influenza in Iceland during the 2009 pandemic. Euro Surveill 2010 Dec 09;15(49):3 [FREE Full text] [doi: 10.2807/ese.15.49.19742-en] [Medline: 21163181]

50. Suzue T, Hoshikawa Y, Nishihara S, Fujikawa A, Miyatake N, Sakano N, et al. The new school absentees reporting system for pandemic influenza A/H1N1 2009 infection in Japan. PLoS One 2012;7(2):e30639 [FREE Full text] [doi: 10.1371/journal.pone.0030639] [Medline: 22363458]
51. Dai Y, Luo K, Wang X. Thermoelectric properties of graphene-like nanoribbon studied from the perspective of symmetry. Sci Rep 2020 Jun 04;10(1):9105 [FREE Full text] [doi: 10.1038/s41598-020-66073-y] [Medline: 32499513]
52. Tan L, Cheng L, Yan W, Zhang J, Xu B, Diwan V, et al. Using daily syndrome-specific absence data for early detection of school outbreaks: a pilot study in rural China. Public Health 2014 Sep;128(9):792-798. [doi: 10.1016/j.puhe.2014.06.004] [Medline: 25200208]
53. Ward MA, Stanley A, Deeth LE, Deardon R, Feng Z, Trotz-Williams LA. Methods for detecting seasonal influenza epidemics using a school absenteeism surveillance system. BMC Public Health 2019 Sep 05;19(1):1232 [FREE Full text] [doi: 10.1186/s12889-019-7521-7] [Medline: 31488092]
54. Williams NJ, Ghosh TS, Bisgard KM, Vogt RL. Comparison of 3 school-based influenza surveillance indicators: lessons learned from 2009 pandemic influenza A (H1N1)--Denver Metropolitan Region, Colorado. J Public Health Manag Pract 2013;19(2):119-125. [doi: 10.1097/PHH.0b013e318252f005] [Medline: 23358289]
55. Egger JR, Hoen AG, Brownstein JS, Buckeridge DL, Olson DR, Konty KJ. Usefulness of school absenteeism data for predicting influenza outbreaks, United States. Emerg Infect Dis 2012 Aug;18(8):1375-1377 [FREE Full text] [doi: 10.3201/eid1808.111538] [Medline: 22840354]
56. Olson JG. School absenteeism during an outbreak of B/Hong Kong/5/72-like influenza virus in Taipei, Taiwan. Southeast Asian Journal of Tropical Medicine and Public Health 1980;11(4):429-434.
57. Chin J, Magoffin RL, Lennette EH. The epidemiology of influenza in California, 1968-1973. West J Med 1974 Aug;121(2):94-99 [FREE Full text] [Medline: 4847436]
58. Lenaway DD, Ambler A. Evaluation of a school-based influenza surveillance system. Public Health Rep 1995;110(3):333-337 [FREE Full text] [Medline: 7610226]
59. McCormick JB, Yan C, Ballou J, Salinas Y, Reininger B, Gay J, et al. Response to H1N1 in a U.S.-Mexico border community. Biosecur Bioterror 2010 Sep;8(3):233-242 [FREE Full text] [doi: 10.1089/bsp.2010.0014] [Medline: 20825334]
60. Rubin RJ, Gregg MB. Influenza surveillance in the United States 1972-1974. Am J Epidemiol 1975 Sep;102(3):225-232. [doi: 10.1093/oxfordjournals.aje.a112151] [Medline: 1163528]
61. Temte JL, Barlow S, Goss M, Temte E, Schemmel A, Bell C, et al. Cause-specific student absenteeism monitoring in K-12 schools for detection of increased influenza activity in the surrounding community-Dane County, Wisconsin, 2014-2020. PLoS One 2022;17(4):e0267111 [FREE Full text] [doi: 10.1371/journal.pone.0267111] [Medline: 35439269]
62. Weiss Z. Detecting the Onset of Infectious Disease Outbreaks Using School Sign-out Logs. Epidemiology 2019 May;30(3):e18-e19. [doi: 10.1097/EDE.0000000000000969] [Medline: 30640731]
63. Egger J, Konty K, Wilson E, Karpati A, Matte T, Weiss D, et al. The effect of school dismissal on rates of influenza-like illness in New York City schools during the spring 2009 novel H1N1 outbreak. J Sch Health 2012 Mar;82(3):123-130. [doi: 10.1111/j.1746-1561.2011.00675.x] [Medline: 22320336]
64. Donaldson AL, Hardstaff JL, Harris JP, Vivancos R, O'Brien SJ. School-based surveillance of acute infectious disease in children: a systematic review. BMC Infect Dis 2021 Aug 03;21(1):744 [FREE Full text] [doi: 10.1186/s12879-021-06444-6] [Medline: 34344304]
65. Reich NG, McGowan CJ, Yamana TK, Tushar A, Ray EL, Osthus D, et al. Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the U.S. PLoS Comput Biol 2019 Nov;15(11):e1007486 [FREE Full text] [doi: 10.1371/journal.pcbi.1007486] [Medline: 31756193]
66. Ray EL, Reich NG. Prediction of infectious disease epidemics via weighted density ensembles. PLoS Comput Biol 2018 Feb;14(2):e1005910 [FREE Full text] [doi: 10.1371/journal.pcbi.1005910] [Medline: 29462167]
67. Bracher J, Ray EL, Gneiting T, Reich NG. Evaluating epidemic forecasts in an interval format. PLoS Comput Biol 2021 Feb;17(2):e1008618 [FREE Full text] [doi: 10.1371/journal.pcbi.1008618] [Medline: 33577550]
68. Paireau J, Andronico A, Hozé N, Layan M, Crépey P, Roumagnac A, et al. An ensemble model based on early predictors to forecast COVID-19 health care demand in France. Proc Natl Acad Sci U S A 2022 May 03;119(18):e2103302119 [FREE Full text] [doi: 10.1073/pnas.2103302119] [Medline: 35476520]

Abbreviations

ILI: influenza-like illness
PRISMA: Preferred Reporting Items for Systematic Review and Meta-analysis

Edited by A Mavragani; submitted 22.07.22; peer-reviewed by Y Katayama, F Mallor, B Smarr; comments to author 21.10.22; revised version received 04.11.22; accepted 25.11.22; published 11.01.23.

Please cite as:
Tsang TK, Huang X, Guo Y, Lau EHY, Cowling BJ, Ip DKM
Monitoring School Absenteeism for Influenza-Like Illness Surveillance: Systematic Review and Meta-analysis
JMIR Public Health Surveill 2023;9:e41329
URL: https://publichealth.jmir.org/2023/1/e41329
doi:10.2196/41329
PMID:36630159

Review

The Usability, Feasibility, Acceptability, and Efficacy of Digital Mental Health Services in the COVID-19 Pandemic: Scoping Review, Systematic Review, and Meta-analysis

Shaoling Zhong[1], MD; Xinhu Yang[1], MSc; Zihua Pan[1], MSc; Yu Fan[1], MSc; Yanan Chen[1], MSW; Xin Yu[1], BM; Liang Zhou[1], MD

[1] The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou, China

Corresponding Author:
Liang Zhou, MD
The Affiliated Brain Hospital of Guangzhou Medical University
36 Mingxin Road
Liwan District
Guangzhou, 510370
China
Phone: 86 81268045
Email: liangzhou_csu@vip.163.com

Abstract

Background: After the rapid spread of the novel SARS-CoV-2, the short-term and long-term mental health impacts of the pandemic on the public, in particular on susceptible individuals, have been reported worldwide. Although digital mental health services expand accessibility while removing many barriers to in-person therapy, their usability, feasibility, acceptability, and efficacy require continued monitoring during the initial phase of the pandemic and its aftermath.

Objective: In this study, we aimed to understand what mental health services are offered, whether they are practical or acceptable, and to what extent digital mental health services are effective in response to the COVID-19 pandemic across high-income and low- and middle-income countries.

Methods: This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and the PRISMA Extension for Scoping Reviews (PRISMA-ScR) guideline. We implemented searches in PubMed (MEDLINE), Embase, PsycINFO, and Cochrane databases for studies that were published between December 2019 and November 2021 and that involved the use of digital mental health services. Two review authors screened, assessed, and extracted studies independently. The protocol was registered on the International Prospective Register of Systematic Reviews.

Results: This review identified 7506 articles through database searching. In total, 65 (0.9%) studies from 18 countries with 67,884 participants were eligible for the scoping review. Of the 65 studies, 16 (24.6%) were included in the meta-analysis. A total of 15 (23.1%) studies measured the usability; 31 (47.7%) studies evaluated the feasibility; 29 (44.6%) studies assessed the acceptability; and 51 (78.5%) studies assessed the efficacy. Web-based programs (21/65, 32.3%), videoconferencing platforms (16/65, 24.6%), smartphone apps (14/65, 21.5%), and SMS text messaging (5/65, 7.7%) were the main techniques. Psychotherapy (44/65, 67.7%) followed by psychoeducation (6/65, 9.2%) and psychological support (5/65, 7.7%) were commonly used. The results of the meta-analysis showed that digital mental health interventions were associated with a small reduction in depressive symptoms (standardized mean difference=−0.49; 95% CI −0.74 to −0.24; *P*<.001) and a moderate reduction in anxiety symptoms (standardized mean difference=−0.66; 95% CI −1.23 to −1.0; *P*=.02) significantly.

Conclusions: The findings suggest that digital mental health interventions may be practical and helpful for the general population, at-risk individuals, and patients with preexisting mental disorders across high-income and middle-income countries. An expanded research agenda is needed to apply different strategies for addressing diverse psychological needs and develop integrated mental health services in the post–COVID-19 era.

Trial Registration: PROSPERO CRD42022307695; https://tinyurl.com/2jcuwjym

(JMIR Public Health Surveill 2023;9:e43730) doi:10.2196/43730

KEYWORDS

digital medicine; COVID-19; mental health services; psychological well-being; COVID-19 pandemic

Introduction

Background

As of October 5, 2022, there were 624.4 million confirmed cases of COVID-19 spread over 228 countries and territories, and COVID-19 has claimed the lives of 6.6 million individuals [1]. After the rapid spread of the novel SARS-CoV-2, known risk factors for mental health impacts have been reported worldwide. Fear of the virus and containment strategies might challenge psychological well-being [2]. Social isolation, loneliness, unemployment, and loss of income after the incidence of COVID-19 have become common. These risk factors might result in mental health problems, such as anxiety, depression, and insomnia, particularly in susceptible populations that include patients with COVID-19, first-line health professionals, and older adults [3]. In a cross-sectional study by Lai et al in China [4], 50% (634/1257) of health care workers reported experiencing depressive symptoms; 45% (560/1257) had anxiety; and 34% (427/1257) reported experiencing insomnia. Increased symptoms of mental health disorders and limited access to mental health care services and social support have also been reported in people with preexisting mental health disorders [5]. Furthermore, emerging reports suggested the possibility of the long-term effects of pandemics on psychological well-being [6,7].

Although most evidence has been on the negative mental health impacts of COVID-19, a parallel area of research interest has explored how mental health services change when exposed to such stressors [8]. Web-based digital mental health services expand accessibility while removing many barriers to in-person therapy; thus, digital mental health services might be a solution in response to the challenge. Recent efforts in implementing digital mental health services have shown promising applications based on efficacy results from randomized controlled trials (RCTs). In a systematic review [9], internet-based cognitive behavioral therapy (CBT) was linked to a higher reduction in depressive symptoms at posttreatment than treatment as usual and the waiting list. Usability, feasibility, acceptability, and efficacy results at the time of pandemics, however, require continued monitoring. Knowing what mental health care service is available, whether it is feasible and acceptable, and to what extent digital mental health services take effect is crucial to inform policy decisions, such as the delivery, implementation, and target areas for applying mental health resources during the pandemic and its aftermath.

Objective

Against this background, the objective of this scoping review and systematic review was to qualify and quantify the usability, feasibility, acceptability, and efficacy of digital mental health interventions applied for infectious disease outbreaks in the initial phase of the pandemic. We focused on the assessment of whether each of the applied techniques was usable, feasible, acceptable, and effective. We also separately synthesized data from high-income countries and low- and middle-income countries (LMICs). As there is no standard definition of usability, feasibility, acceptability, and efficacy, we considered comprehensive and broad definitions of these terms. We defined usability as program use, user engagement, and the ease of use of the services being tested, as followed in prior research [10,11]. We measured feasibility by attrition, attendance, adherence, retention, and qualitative feedback. We defined acceptability as user satisfaction, intent to continue use, and the perceived appropriateness of the intervention [12]. We defined efficacy as the intended effects and effect size estimation of the services [12].

Methods

Overview

This report includes a scoping review of studies on digital mental health interventions applied for infectious disease outbreaks and a systematic review of studies assessing commonly reported mental distresses identified in the scoping review. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and the PRISMA Extension for Scoping Reviews (PRISMA-ScR) guideline when appropriate (Multimedia Appendix 1).

Data Search

We sought to include original studies focusing on digital mental health services applied during or in response to COVID-19. We implemented searches in PubMed (MEDLINE), Embase, PsycINFO, and Cochrane between December 2019 and November 2021 using a combination of search terms relating to both digital mental health services and COVID-19. The complete search strategy is provided in Multimedia Appendix 2. We performed a manual search of the reference lists of relevant reviews, eligible studies, and relevant conference abstracts.

Study Selection

We removed duplicate citations across electronic databases before the screening. Initially, pairs of 2 review authors (SZ, XY, ZP, and YF) independently screened the titles and abstracts for eligibility. We referred to full texts if titles and abstracts did not provide information for eligibility based on the inclusion and exclusion criteria. The full texts of the potentially included studies were screened by 1 review author and checked by another author for agreement. In case of disagreements, the opinion of a third reviewer was sought, and a consensus was reached through discussion.

We included qualitative, quantitative, and mixed methods studies that involved the use of digital mental health services. Eligible interventions included health education, evaluation, consulting, helplines, psychological interventions, and telemedicine. The intervention was delivered by any means of digital technology, including SMS text messaging, smartphones, websites, social media, videoconferencing, wearable devices, or mobile apps. We excluded comments, protocol papers, systematic reviews, preprints, conference papers, and studies that did not provide any qualitative or quantitative data for defined outcomes. We excluded studies that were conducted before December 2019. There was no exclusion for age, gender, ethnicity, socioeconomic status, professions (health workers or not), or any type of mental disorder (clinical or nonclinical). For the meta-analysis, we included parallel RCTs and crossover RCTs that provided adequate data for analyses, using the same aforementioned criteria.

Data Extraction

We developed a data extraction form that could be used by all review authors, as recommended by the Joanna Briggs Institute Scoping Review methodological guidance [13]. The data extraction form included information on study characteristics (authors, publication year, study design, study period, geographic regions, and study aim), participant characteristics (age, gender, sample size, and type of mental illness), digital mental health services characteristics (types of techniques and type of mental health services), and outcome characteristics (type of outcomes and their measurements). Owing to the diverse outcomes identified across eligible studies, we extracted the study results in 2 phases. In phase 1, we extracted the key findings for all reported outcomes. For effectiveness data, we extracted the results for significant differences between digital mental health services and comparison groups or trends between postintervention and preintervention outcome measurements for digital mental health services. In phase 2, we identified RCTs that reported similar digital mental health services and outcomes for the systematic review and meta-analysis. Here, effectiveness data from the comparison groups and intervention arms were retrieved using separate extraction forms. The data were extracted by 2 reviewers (XY and ZP) independently. Disagreements were resolved by discussion. A third reviewer (SZ) was involved when necessary.

Risk of Bias Assessment

Two reviewers (XY and ZP) independently assessed the risk of bias using the Cochrane Risk of Bias (RoB 2) tool for all included RCTs. Disagreements were resolved by discussions. In case of uncertainties, a third reviewer (SZ) was consulted. The 5 domains included bias owing to the assignment of intervention, adherence to the intervention, missing outcome data, outcome measurement, and selection of reported outcome that were measured. Studies were rated as “low” risk, “some concerns,” or “high” risk.

Data Synthesis

Owing to the nature of the research question and the heterogeneity of the studies, we synthesized qualitative and quantitative results separately. For the scoping review, we descriptively summarized the included studies with a narrative synthesis of individual studies. We summarized the overall number of included studies, the total number of study populations and their basic characteristics, countries, study design, type of digital mental health services, and the objectives and findings of the included studies. We clustered the studies by the type of techniques used and location (high-income countries vs LMICs). For each type of technique, the findings on usability, feasibility, acceptability, and effectiveness data were narratively synthesized by a textual approach.

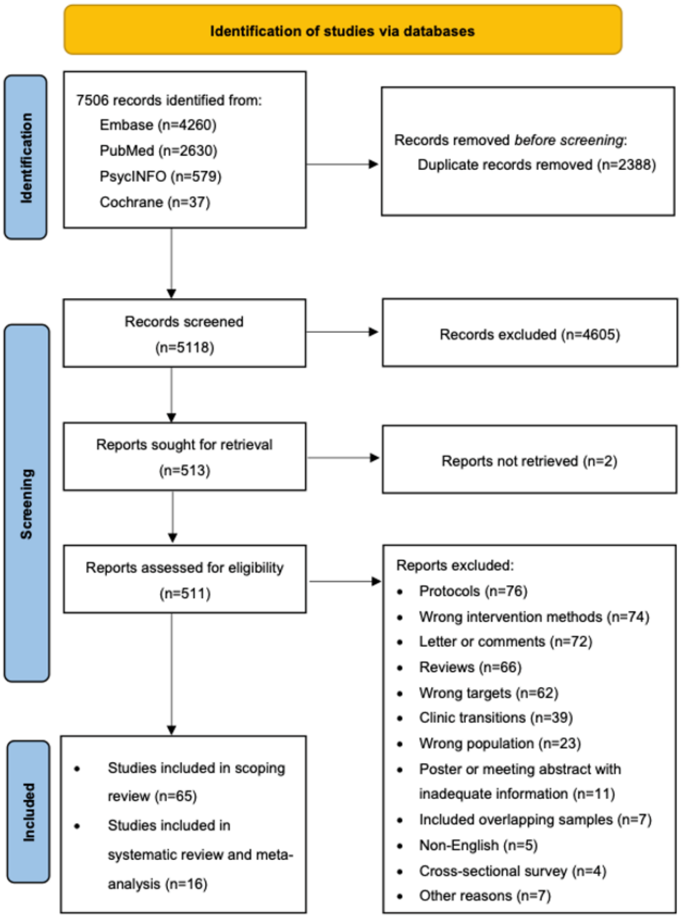
For meta-analyses, we synthesized findings on similar efficacy outcomes that were measured by >3 RCTs. Outcomes, such as loneliness, burnout, or psychological well-being, that were measured by 1 or 2 RCTs were not synthesized. We pooled continuous outcomes as standardized mean differences (SMD) with 95% of CIs using a random-effects meta-analysis. The values of 0.2, 0.5, and 0.8 in SMDs denote minor, medium, and large effects, respectively [14]. Cochran Q and the I^2 statistic were used to assess heterogeneity. An I^2 value of ≥60% can be considered a substantial level of inconsistency across studies [15]. The analyses were stratified according to the reported outcomes. For each of these, we pooled the evidence from all eligible RCTs, regardless of the techniques applied, measurements used, and population studied. We conducted subgroup analyses to examine whether there were differences in outcomes on the basis of sample size (as a dichotomous variable), measurement of the outcome, location, duration of follow-up, and duration of intervention. We also performed meta-regression analyses to explore the effect size by sample size (as a continuous variable), age, and proportion of females in the population studied. We used funnel plots and ran the Egger test to look for asymmetry to detect publication bias that might impair the validity of our findings. We performed all analyses with R statistical software (version 4.2.2; R Foundation for Statistical Computing) [16].

Results

Study Selection

The PRISMA flowchart shows the study selection process (Figure 1). This review identified 7506 articles that were identified through database searching. After removing duplicate records, we screened 5118 (68.19%) records and identified 65 (1.27%) eligible studies for the scoping review, and the results were synthesized narratively. Of the 19 RCTs, 17 (89%) studies reported depressive symptoms and 18 (95%) studies reported anxiety symptoms as the outcome. In total, 3 (16%) studies were not included in the meta-analyses; 2 (67%) studies [17,18] that reported on depression and anxiety disorders did not provide adequate data, and 1 (33%) study that reported anxiety scores using state anxiety was with a very large estimate [19]. Accordingly, of the total 65 eligible studies, 16 (25%) studies that reported depressive symptoms or anxiety symptoms were included in the meta-analysis.

Figure 1. Flowchart of study selection.



Study Characteristics

Table 1 presents the characteristics of participants and studies included in the review. The 65 included studies were conducted in the following 18 different countries: 24 (37%) studies from the United States (participants: 54,245/67,884, 79.9%), 8 (12%) studies from China (1114/67,884, 1.6%), 6 (9%) studies from Canada (5490/67,884, 8.1%), 5 (8%) studies from Italy (315/67,884, 0.5%), 3 (5%) studies from Australia (1004/67,884, 1.5%), 3 (5%) studies from France (2235/67,884, 3.3%), 2 (3%) studies from India (654/67,884, 1%), 2 (3%) studies from Malaysia (123/67,884, 0.2%), 2 (3%) studies from Greece (164/67,884, 0.2%), 2 (3%) studies from Turkey (121/67,884, 0.2%), and further 9 (14%) studies were conducted in other countries (ie, Dominican Republic, Israel, Japan, Oman, South Africa, Spain, Sweden, and the United Kingdom). Of these, 49 (75%) studies were from high-income countries, and 16 (25%) studies were conducted in middle-income countries. None of the included studies reported data from low-income countries.

The participant pool consisted of 67,884 participants; most (44,125/67,884, 65%) were female. There were 2663 (3.9%) people with a mental illness or mental distress, 713 (1.1%) people with or suspected of COVID-19, and 748 (1.1%) hospital workers.

Of the 65 studies, a total of 15 (23.1%) studies measured usability; 31 (47.7%) evaluated feasibility; 29 (44.6%) assessed acceptability; and 51 (78.5%) assessed efficacy.

Web-based programs (21/65, 32.3%), videoconferencing platforms (16/65, 24.6%), smartphone apps (14/65, 21.5%), and SMS text messaging (5/65, 7.7%) were the main techniques applied. There were other studies that used social media (3/65, 4.6%), hotlines (3/65, 4.6%), a chatbot (1/65, 1.5%), virtual reality (VR; 1/65, 1.5%), and robotic telemedicine (1/65, 1.5%). Most studies provided psychotherapy (44/65, 67.7%), followed by psychoeducation (6/65, 9.2%), psychological support (5/65, 7.7%), psychological assessments (4, 6.1%), and psychiatric clinical services (4/65, 6.1%).

Table 1. Characteristics of participants and studies included in the qualitative systematic review.

Author and year	Design and study type	Sample size, n	Age (years), mean (SD)[a]	Female, %	Intervention vs control group
Aguilera et al [20]	Pre- and postintervention evaluation; quantitative	303	33.3 (11.0)	78	A texting intervention (StayWell at Home)
Agyapong et al [21]	Pre- and postintervention evaluation; quantitative	2767	76.7% were aged between 26 and 60	88	Text4Hope subscribers who received once-daily supportive SMS text messages vs Text4Hope subscribers who did not receive messages
Al-Alawi et al [18]	RCT[b]; quantitative	46	28.51 (8.7)	78	Therapist-guided web-based therapy vs newsletter received via email containing self-help information and tips to cope with distress
Al-Refae et al [22]	RCT; quantitative	165	25.24 (8.74)	78.8	Smartphone intervention (Serene) vs waiting list
Bantjes et al [23]	Pre and postintervention evaluation; quantitative	175	22.2 (4.9)	85.4	A web-based group CBT[c] intervention
Ben-Zeev et al [24]	RCT; quantitative	315	37.89 (11.64)	83.8	Smartphone app (CORE) vs waiting list
Brouzos et al [25]	Pre- and postintervention evaluation; quantitative	82	33.07 (9.55)	78	Web-based positive psychology intervention ("Staying Home—Feeling Positive") vs no intervention
Casas [26]	Pre- and postintervention evaluation; mixed methods	15	39.52 (13.57)	73.3	Written exposure therapy delivered via telehealth
Chandra et al [27]	Pre- and postintervention evaluation; mixed methods	643	49.57 (15.23)	37.2	Mental health support telecounseling
Coifman et al [28]	Pre- and postintervention evaluation; quantitative	28	45.33 (9.6)	75	Daily coping toolkit intervention
Comer et al [29]	RCT; quantitative	40	Children=6.2 (1.8); parents=38.39 (4.5)	Children: 72.5; parents: 75	iCALM Telehealth Program vs waiting list
Craig et al [30]	Pre- and postintervention evaluation; mixed methods	96	22.3 (4.1)	17.7	AFFIRM online group vs waiting list
Detweiler Guarino et al [31]	Pre- and postintervention evaluation; quantitative	2484	44 (15.2)	53.2	The PATH program
Dincer and Inangil [32]	RCT; quantitative	72	33.46 (9.6)	88.9	Received emotional freedom techniques vs stayed comfortable in a calm and tranquil environment
Esentürk and Yarımkaya [33]	Pre- and postintervention evaluation; mixed methods	14	Children=12.07; parents=51.4	Children=42.9; parents=64.2	WhatsApp-based physical activity intervention
Fiol-DeRoque et al [34]	RCT; quantitative	482	41.37 (10.4)	83.2	PsyCovidApp vs a general control app
Gabrielli et al [35]	Pre- and postintervention evaluation; mixed methods	71	20.6 (2.4)	68	Healthy-coping intervention via chatbot
Geoffroy et al [36]	Postintervention evaluation;	149	32.7 (11.0)	86	Covid-Psy hotline
Gordon et al [37]	Pre- and postintervention evaluation; mixed methods	99	51.4 (14.7)	86	A guided imagery mobile app
Graziano et al [38]	Pre- and postintervention evaluation; quantitative	30	Patients=22.5 (6.9); parents=37 (6.3)	Patient=56.3; parents=92.9	Telehealth psychological interventions
Gromatsky et al [39]	Pre- and postintervention evaluation; mixed methods	20	54.2 (11.95)	15	VA CONNECT

Author and year	Design and study type	Sample size, n	Age (years), mean (SD)[a]	Female, %	Intervention vs control group
Guan et al [19]	RCT; quantitative	95	18.63 (0.75)	28.4	Self-compassion writing induction condition vs a writing control condition (to think about a negative event)
Guo et al [40]	Pre- and postintervention evaluation; quantitative	508	Intervention=47.0 (4.3), control=46.6 (4.2)	84.1	Web-based education program vs no intervention
Held et al [41]	Pre- and postintervention evaluation; mixed methods	2	—[d]	0	Intensive cognitive processing therapy–based program
Hom et al [42]	Postintervention evaluation; mixed methods	23	39.04 (16.15)	52.2	Virtual partial hospital program
Hosseinzadeh Asl [43]	RCT; quantitative	49	33.06 (6.02)	55.1	Mindfulness-based intervention vs waiting list
Hu et al [44]	Pilot RCT; quantitative	48	42.6 (15.8)	54.2	WeChat-based psychological interventions vs conventional nursing
Ibrahim et al [45]	Pre- and postintervention evaluation; quantitative	43	63.41	67.4	Virtual group exercise
Jaworski et al [46]	Postintervention evaluation; quantitative	49,287	—	—	COVID coach
Kahlon et al [47]	RCT; quantitative	240	69.1 (12.1)	79	Receive calls vs no calls
Karagiozi et al [48]	RCT; quantitative	82	44.5 (10.6)	—	Web-based intervention group vs onsite intervention
Kim [49]	RCT; quantitative	80	41.7 (13.8)	56.25	Teleacupressure self-practice group vs web-based communication group
Kolbe et al [50]	Postintervention evaluation; mixed methods	13 patients and 11 staff	—	—	Virtual reality
Lazzaroni et al [51]	Pre- and postintervention evaluation; quantitative	50	Between 13 and 24	84	Eye movement desensitization and reprocessing intervention group
Lima et al [52]	Postintervention evaluation; mixed methods	Healthy volunteers=10; older adults and people with dementia=1	Healthy volunteers=21-59; older adults and people with dementia=not mentioned	Healthy volunteers=50; older adults and people with dementia=0	Robotic telemedicine
Maldonado [53]	Pre- and postintervention evaluation; quantitative	40	44.1 (13)	82.5	Online intervention group
McKeon et al [54]	Pre- and postintervention evaluation; quantitative	11	64.4 (4.3)	82	Mental health–informed lifestyle intervention
Nauphal et al [55]	Pre- and postintervention evaluation; quantitative	5	36.2 (9.5)	40	Web-based CBT group
Ortiz et al [56]	Pre- and postintervention evaluation; mixed methods	383	46.9 (16.8)	70.1	Texting intervention group
Parolin et al [57]	Pre- and postintervention evaluation; mixed methods	134	33.20 (10.61)	86	Italia Ti Ascolto intervention group
Philip et al [58]	Pre- and postintervention evaluation; quantitative	2069	43.52 (13.94)	65.1	KANOPEE group
Puspitasari et al [59]	Pre- and postintervention evaluation; quantitative	76	36.6 (13.4)	86	Adult Transitions Program teletherapy group
Rojas et al [60]	Pre- and postintervention evaluation; mixed methods	1	—	0	Brief CBT for suicide prevention
Shalaby et al [61]	Postintervention evaluation; quantitative	2032	44.58 (13.45)	88	Text4Hope intervention group
Shapira et al [62]	Pilot RCT; quantitative	82	72 (5.63)	—	A short-term digital group vs a waitlist control group

Author and year	Design and study type	Sample size, n	Age (years), mean (SD)[a]	Female, %	Intervention vs control group
Sharrock et al [63]	Pre- and postintervention evaluation; quantitative	904	37.83 (12.64)	67.1	iCBT[e] group during COVID-19 vs iCBT group before COVID-19
Song et al [64]	Pre- and postintervention evaluation; quantitative	129	34.64 (9.11)	69	Mobile internet CBT group vs a waitlist control group
Sosa Lovera et al [65]	Postintervention evaluation; quantitative	497	32	73	COVID-19 helpline group
Sturgill et al [66]	Pre- and postintervention evaluation; quantitative	99	19.9 (1.94)	69	Ajivar activities intervention + routine mental wellness instruction group vs group that received routine mental wellness instruction
Suffoletto et al [67]	Pilot RCT; quantitative	52	18.7 (0.46)	86.5	MoST-MH[f] Intervention vs enhanced usual care group
Summers, Wu, and Taylor [68]	Pre- and postintervention evaluation; quantitative	273	49.6 (9.24)	59.3	Gro Health app users
Sun et al [69]	Pilot RCT; mixed methods	114	22.21 (2.67)	73.7	Mindfulness-based mHealth[g] group vs social support–based mHealth group
Tarquinio et al [70]	Pre- and postintervention evaluation; quantitative	17	33.2 (4.10)	100	Eye Movement Desensitization and Reprocessing intervention group
Vallefuoco et al [71]	Postintervention evaluation; quantitative	30	—	—	18 therapists and 12 parents of children with ASD[h]
van Agteren et al [72]	Pre- and postintervention evaluation; quantitative	89	38.67 (13.06)	66	Internet-based group mental health intervention
Wagner [73]	Postintervention evaluation; mixed methods	Children: 204; clinicians: 9	Children=27.54 months (5.36 months)	23	ASD-PEDS intervention group
Wahlund et al [74]	RCT; mixed methods	670	46 (13.50)	81.6	Web-based psychological intervention vs waiting list
Wasil et al [75]	Pre- and postintervention evaluation; quantitative	189	31.04 (8.91)	72.99	COMET intervention
Wei et al [17]	RCT; quantitative	26	44.65 (12.09)	38.5	An internet-based integrated intervention vs support care
Wood et al [76]	Postintervention evaluation; mixed methods	7	26.9 (4.8)	—	Group teletherapy
Ying et al [77]	Pre- and postintervention evaluation; quantitative	137	73.39 (7.37)	68.5	iCBT group
Zepeda et al [78]	Pre- and postintervention evaluation; mixed methods	27	9.56	67	iCOPE
Zhang et al [79]	Pilot RCT; quantitative	57	50.12 (6.79)	47.1	WeChat-based psychological interventions vs waitlist group
Zimmerman et al [80]	Postintervention evaluation; quantitative	480	37.1 (14)	67.5	Telehealth program vs partial in-person program before the COVID-19 outbreak
Van Lieshout et al [81]	RCT; quantitative	403	31.8 (4.40)	100	Web-based CBT + treatment as usual vs treatment as usual

[a] Where mean (SD) was not available in the original study, mean or ranges were used where applicable.

[b] RCT: randomized controlled trial.

[c] CBT: cognitive behavioral therapy.

[d] Not applicable.

[e] iCBT: internet-based congnitive behavioral therapy.

[f] MoST-MH: Mobile Support Tool for Mental Health.

[g] mHealth: mobile health.

[h] ASD: autism spectrum disorder.

Qualitative Synthesis of Results

Overview

This section summarizes the findings from 65 studies that reported the feasibility, acceptability, and effectiveness of digital mental health services during the COVID-19 pandemic. The synthesis was organized according to the type of digital technique used. Table 2 presents a summary of usability, feasibility, acceptability, and efficacy by the types of techniques in high-income countries and LMICs. Table S1 (Multimedia Appendix 3 [17-81]) presents summaries of individual studies on mental health interventions and major findings and their measurements.

Table 2. Summaries of usability, feasibility, acceptability, and efficacy by type of techniques in high-income countries (HICs) and low- and middle-income countries (LMICs).

Techniques and groups	Usability	Feasibility	Acceptability	Efficacy
Videoconferencing platforms				
HICs (n=11)	—[a]	• Three studies reported a completion rate between 92% and 100%. • Three studies reported qualitative feedback of highly perceived benefits. • Three studies reported technology-related challenges.	• Five studies reported high levels of overall satisfaction.	• Nine studies reported psychotherapy improved depressive levels and 5 studies found improved anxiety symptoms. • One study found no effects on decreasing anxiety symptoms.
LMICs (n=5)	—	• Two studies reported the mean attendance ranges from 6.4 (SD 2.8) to 10.26 (SD 7.02).	• One study reported high levels of satisfaction.	• Three studies reported improvements in depressive symptoms or negative effects. • Two studies reported no effects on decreasing depression levels.
Web-based programs				
HICs (n=17)	• Two studies reported a significant increase in registered users. • One study reported a high rating in usability.	• Five studies reported the completed rate ranged from 30.5% to 85%. • Three studies reported retention rate, which ranged from 30.5% to 85%. • One study reported high attrition rate of 13%. • One study reported strong attendance. • Two studies reported that retentions were strong. • Two studies reported high ratings of feasibility, and one study reported technical difficulty.	• Five studies reported high satisfaction. • One study reported that acceptability scores were significantly better than average. • Three studies reported high acceptability and rated the interventions as acceptable, helpful, appropriate, and positive.	• Eight studies reported improvement in anxiety symptoms and anxiety-related social impairment neither with nor without controls. • Two studies indicated no effects on reducing anxiety. • Five studies indicated improvement in depression symptoms and depression severity. • One study reported no difference in depressive symptoms when compared with onsite groups. • Three studies reported improved stress-related symptoms. • Two studies reported reductions in COVID-19–related worry. • One study indicated improvements in insomnia.
LMICs (n=4)	• One study reported the mean time spent was 35.63 (SD 25.41) min.	• One study reported a completion rate of 87.4%. • One study reported high attendance and high retention rates (91.2%).	• One study reported high ratings of acceptability and the other a high level of satisfaction.	• Four studies indicated improvement in anxiety symptoms neither with no controls nor control groups. • Two studies reported improvement in depressive symptoms with no controls. • One study reported a reduction in stress and burnout.
Smartphone apps				
HICs (n=11)	• Two studies reported high usability scores. Two studies reported user engagement. The time ranged from 36.7 min over 12 weeks to 1424 min over 14 weeks.	• One study reported the number of days retained was 42.44 (SD 44.40). • Three studies reported a retention rate that ranged from 28.3% to 85.7%.	• Two studies rated the intervention as acceptable, and 2 studies rated the intervention as satisfied. • One rated the intervention as good. • One rated the intervention as helpful.	• Four studies reported reduced depressive symptoms compared with the preintervention and waitlist groups. • Five studies reported similar improvements in anxiety symptoms compared with the preintervention, waitlist group, and a control app with limited access to psychoeducational content.

Techniques and groups	Usability	Feasibility	Acceptability	Efficacy
LMICs (n=3)	—	—	• One study indicated high satisfaction and would recommend the service. One study reported helpful and enjoyable.	• Two studies reported improvements in depressive symptoms, insomnia, psychological flexibility, and self-compassion, either with no controls or in comparison with the waitlist group.
Texting				
HICs (n=5)	• One study reported high usability.	• Two studies reported a completion rate of 16% and 78%, respectively.	• One study reported a high level of overall satisfaction.	• Four studies reported an improved mood rating was observed at posttreatment assessment either with or without the control group.
LMICs (n=0)	—	—	—	—
Social media				
HICs (n=1)	—	• One study reported high retention.	• One study reported high acceptability.	• One study reported an effect on psychological distress, quality of life, functioning, loneliness, and physical activity without a control group.
LMICs (n=2)	• One study reported an average user time of 18.7 hours.	—	—	• Two studies reported improvements in anxiety and depressive symptoms when compared with either the usual care or the waitlist group.
Hotline and telephone calls				
HICs (n=2)	• One study reported 5.73 (SD 3.22) calls/day.	• One study reported a dropout rate of 7.5%.	—	• One study reported improvements in loneliness, depression, anxiety, and general mental health.
LMICs (n=1)	—	—	• One study reported high satisfaction, and the participants would use again, and would recommend it to others.	—
Robotic telemedicine and VR[b]				
HICs (n=2)	• One study reported an average of 78 (SD 24.8) times.	• One reported a completion rate of 58%.	• One study reported high satisfaction.	• One study reported a decrease in anxiety symptoms and stress symptoms.
LMICs (n=1)	• One study reported an overall positive impression of the multimodal robotic system.	• One study reported the need to adjust some features.	—	—

[a] No evidence found.

[b] VR: virtual reality.

Web-Based Programs

Overall, 21 (32.3%; 5480/67,884, 8.1% out of total participants) out of the 65 studies included in this scoping review used web-based programs. In 17 studies that were conducted in high-income countries, a significant increase in registered users (n=2) and a high rating in usability (n=1) were reported. The completion rate ranged from 30.5% to 85% (n=5); the retention rate ranged from 30.5% to 85% (n=3); and the attrition rate was 13% (n=1). Quality checks of feasibility reported strong attendance (n=1), strong retentions (n=2), and high ratings of feasibility (n=2). Studies presented preliminary support for the feasibility of web-based programs yet with possible technical difficulty (n=1). In terms of acceptability, there was high satisfaction (n=5); the interventions were perceived to be acceptable, helpful, appropriate, and positive (n=3); and the acceptability scores were significantly better than average (n=1). Most studies, either when no control group is involved or when compared with the control group, reported improvements in anxiety symptoms and anxiety-related social impairment (n=8), while no effects on reducing anxiety were reported as well (n=2). The same patterns were found in outcomes measuring depression. Most studies reported improvements in depressive symptoms and depression severity (n=5), and there was no difference in depressive symptoms when the intervention group was compared with onsite groups (n=1). Improvements in stress-related symptoms (n=3), COVID-19–related worry (n=2), and insomnia (n=1) were identified.

In 4 studies conducted in LMICs, the mean time spent was 35.6 (SD 25.4) minutes (n=1) and the completion rate was 87.4% (n=1). High attendance and high retention rates (91.2%) were reported (n=1). With regard to acceptability, high ratings of acceptability and a high level of satisfaction (n=2) were reported. Studies, either when no control group is involved or when compared with the waitlist group, reported improvements in anxiety symptoms (n=4); some studies reported improvements in depressive symptoms (n=2), and 1 study reported reduction in stress and burnout.

Videoconferencing Platforms

Among the 65 studies, 16 (24.6%; participants: 1973/67,884, 2.9%) used teleconferencing platforms, including Zoom [30,39,49,55,59,60,62,72,73,81], Skype [25], Microsoft Teams [1,49], Google Meet [45], Tencent Meeting [19], and other platforms [41,42]. In 11 studies that were conducted in high-income countries, a completion rate between 92% and 100% (n=3), qualitative feedback of high perceived benefits (n=3), and technology-related challenges (n=3) were reported in terms of feasibility. High levels of satisfaction were reported (n=5). Studies reported improvements in depressive levels (n=9) and anxiety symptoms (n=5). No effects on decreasing anxiety symptoms were reported in one study. In 5 studies conducted in LMICs, the mean attendance ranged from 6.4 (SD 2.8) to 10.26 (SD 7.02; n=2). High levels of satisfaction (n=1) were reported. There were improvements in depressive symptoms and negative affect (n=3), yet no effects on decreasing depression levels were identified (n=2). No evidence was found in usability for LMICs or high-income countries.

Smartphone Apps

Of the 65 studies, 14 (21.5%; participants: 53,786/67,884; 79.2%) developed smartphone apps to deliver mental health services. In 11 studies conducted in high-income countries, studies reported high usability scores (n=2), and user engagement time ranged from 36.7 minutes over 12 weeks to 1424 minutes over 14 weeks (n=2). Feasibility was reported in terms of the average number of days retained (mean 42.4, SD 44.4; n=1), and the retention rate ranged from 28.3% to 85.7% (n=3). In terms of acceptability, the interventions were rated as “acceptable” (n=2), “satisfied” (n=2), “good” (n=1), and “helpful” (n=1). Studies reported a reduction in depressive symptoms compared with the preintervention and waitlist group (n=4). Similar improvements were reported in anxiety symptoms compared with preintervention, waitlist group, and a control app with limited access to psychoeducational contents (n=5). In 4 studies in LMICs, no evidence was found for usability and feasibility. Acceptability was reported as high satisfaction (n=1) and helpful and enjoyable (n=1). There were improvements in depressive symptoms, insomnia, psychological flexibility, and self-compassion, either when no control group is involved or when compared with the waitlist group (n=2).

SMS Text Messaging

Of the 65 studies, 5 (7.7%; participants: 5537/67,884, 8.1%) used the SMS text messaging technique. All studies were conducted in high-income countries. Studies reported high usability of ecological momentary intervention (n=1), and the completion rate ranged from 16% to 78%. An overall high satisfaction was reported (n=1). Studies suggested that an improved mood rating was observed at posttreatment assessments either with or without the control group (n=4). No evidence was found of studies conducted in LMICs for the technique of SMS text messaging.

Social Media

The use of social media as a technique to deliver mental health services was described in 3 (4.6%; participants: 116/67,884, 0.2%) out of 65 studies [44,54,79]. In one study conducted in high-income countries, high retention, high acceptability, and an effect on multiple outcomes (psychological distress, quality of life, functioning, loneliness, and physical activity) were reported (n=1). In the 2 studies conducted in LMICs, an average user time of 18.7 hours (n=1) was reported. Evidence suggested improvements in anxiety and depressive symptoms when compared with either usual care or the waitlist group (n=2). No evidence was found for other outcomes.

Hotline and Telephone Calls

Of the 65 studies, 3 (4.6%; participants: 886/67,884, 1.3%) reported using hotlines and telephone calls to deliver mental health services during the COVID-19 pandemic [36,47,65]. In 2 studies conducted in high-income countries, an average of 5.73 (SD 3.22) calls per day was reported (n=1). A dropout rate of 7.5% was reported (n=1), and there were improvements in loneliness, depression, anxiety, and general mental health, as reported in the same study (n=1). In one study that was conducted in LMICs, participants reported high satisfaction and mentioned that they would use the service again and would recommend it to others (n=1). No evidence was found in terms of usability, feasibility, or efficacy for the techniques of hotline and telephone calls by the studies conducted in LMICs.

Robotic Telemedicine and VR

Of the 65 studies, 3 (4.6%; participants: 106,67,884, 0.2%) used robotic telemedicine and VR. Two studies conducted in high-income countries reported an average use of 78 (SD 24.8) times (n=1) and a completion rate of 58% (n=1). High satisfaction was reported (n=1). Evidence suggested a decrease in anxiety symptoms and stress symptoms for all participants after the intervention (n=1). In 1 study conducted in LMICs, an overall positive impression of the multimodal robotic system was reported (n=1), yet the need to adjust some features was suggested (n=1). No evidence for acceptability or efficacy was found by studies conducted in LMICs.

Meta-analyses

Of the 65 studies, 15 (23.1%) that reported depressive symptoms were included in the meta-analysis. Digital mental health interventions were associated with a small significant reduction in depressive symptoms (SMD −0.49, 95% CI −.74 to −.24; *P*<.001; Figure 2) with substantial heterogeneity (I^2=87%, 95% CI 80%-91%; *Q*=107.57; *P*<.001).

The effectiveness of digital mental health interventions on anxiety symptoms was reported in 15 RCTs. The pooled SMD showed a moderate and significant effect of these interventions in reducing anxiety symptoms (SMD=−0.66, 95% CI −1.23 to −.10; *P*=.02; Figure 3) with substantial heterogeneity (I^2=93%, 95% CI 90%-95%; *Q*=207.22; *P*<.001).

We examined the sources of heterogeneity by studying covariates across studies (Tables 3 and 4). Subgroup analyses revealed that reduced depressive symptoms were associated with the measurement of depression (using Depression Anxiety Stress Scales-21: SMD=−0.13, 95% CI −0.31 to 0.06 vs other measurements: SMD=−0.61, 95% CI −0.91 to −0.30; *Q*=7.04; *P*=.008). Reduced anxiety symptoms were not associated with sample size, location, measurement of anxiety, duration of follow-up, or the duration of intervention.

Meta-regression analyses are reported in Multimedia Appendix 4. There was no significant association between SMDs of depression and sample size, age, and proportion of females in the study population. Meta-regression analyses did not reveal differences in the SMDs of anxiety depending on sample size, age, and proportion of females in the study population.

For the main analyses of depression and anxiety, we observed evidence of funnel plot asymmetry, which could suggest no publication bias (Egger test: *P*=.57 and *P*=.50, respectively; Multimedia Appendix 5).

Figure 2. Effectiveness of digital mental health interventions in reducing depressive symptoms. BDI-II: Beck Depression Inventory-II; BSI: Brief Symptom Inventory; CCAPS: College Counseling Center Assessment of Psychological Symptoms; DASS-21: Depression Anxiety Stress Scales-21; EPDS: Edinburgh Postnatal Depression Scale; HAMD: Hamilton Rating Scale for Depression; MADRS-S: Montgomery Åsberg Depression Rating Scale-Self rated; PHQ: Patient Health Questionnaire; and SDS: Self-rating Depression Scale; SMD: standardized mean difference.

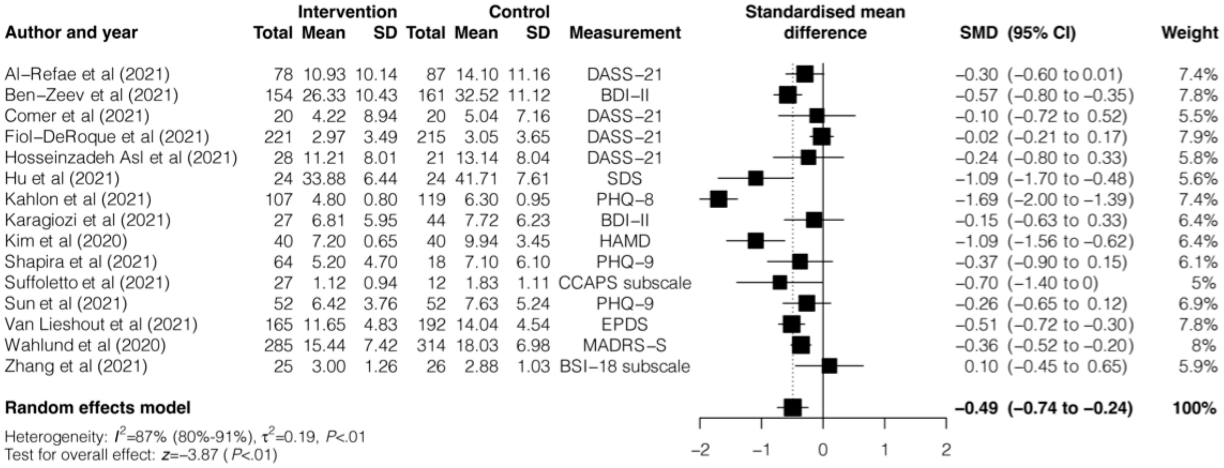


Figure 3. Effectiveness of digital mental health interventions in reducing anxiety symptoms. BDI-II: Beck Depression Inventory-II; BSI: Brief Symptom Inventory; CBCL: Child Behavior Checklist; CCAPS: College Counseling Center Assessment of Psychological Symptoms; DASS-21: Depression Anxiety Stress Scales-21; GAD-7: Generalized Anxiety Disorder-7; HAMA: Hamilton Anxiety Rating Scale; SAS: Self-rating Anxiety Scale; SMD: standardized mean difference; and STAI: State-Trait Anxiety Inventory.

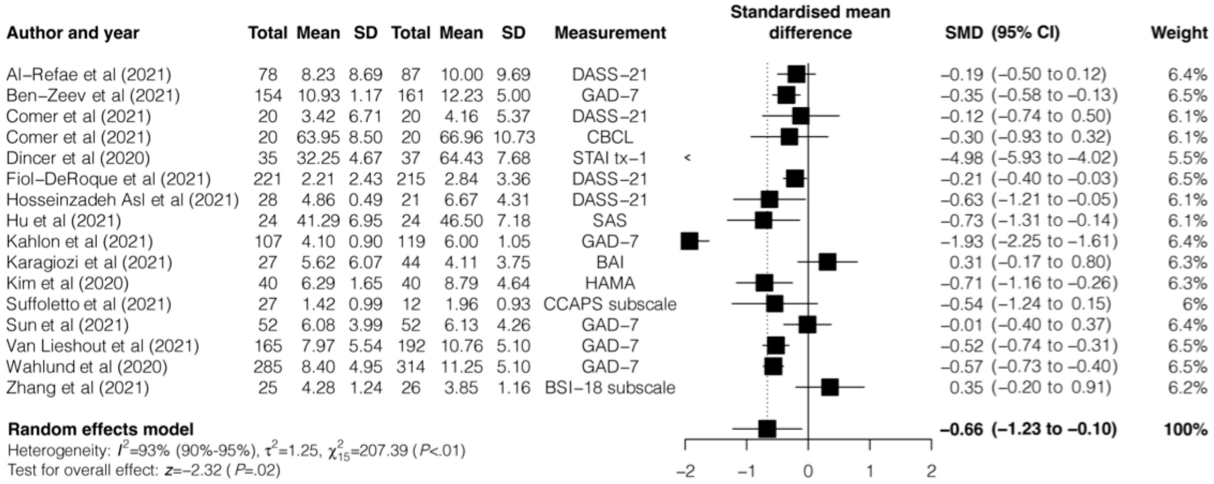


Table 3. Effect size of digital mental health services for reducing depressive symptoms during the COVID-19 pandemic according to study characteristics.

Group	Studies, n	SMD[a] (95% CI)	*Q*	*P* value
Sample size			0.03	.86
≥81	8	−0.51 (−0.86 to −0.16)		
<81	7	−0.46 (−0.84 to −0.09)		
Measurement of depression			7.04	.008
DASS-21[b]	4	−0.13 (−0.31 to 0.06)		
Other measurements[c]	11	−0.61 (−0.91 to −0.30)		
Location			0.41	.52
High-income countries	11	−0.53 (−0.83 to −0.24)		
Low- and middle-income countries	4	−0.35 (−0.82 to 0.10)		
Duration of follow-up			1.23	.27
Short-term	12	−0.55 (−0.84 to −0.26)		
Long-term	3	−0.28 (−0.66 to −0.09)		
Duration of intervention			0.84	.36
≥4 weeks	10	−0.56 (−0.90 to −0.23)		
<4 weeks	5	−0.34 (−0.68 to −0.02)		

[a] SMD: standardized mean difference.

[b] DASS−21: Depression Anxiety Stress Scales-21.

[c] Other measurements include Beck Depression Inventory-II, Brief Symptom Inventory, College Counseling Center Assessment of Psychological Symptoms, Edinburgh Postnatal Depression Scale, Hamilton Rating Scale for Depression, Montgomery Åsberg Depression Rating Scale—Self rated, Patient Health Questionnaire, and Self-rating Depression Scale.

Table 4. Effect size of digital mental health services for reducing anxiety symptoms during the COVID-19 pandemic according to study characteristics.

Group	Studies, n	SMD[a] (95% CI)	*Q*	*P* value
Sample size			0.24	.62
≥81	7	−0.54 (−1.01 to −0.07)		
<81	8	−0.85 (−1.99 to 0.29)		
Location			0.49	.48
High-income countries	10	−0.49 (−0.66 to −0.13)		
LMICs[b]	5	−1.17 (−3.02 to 0.68)		
Measurement of anxiety1			0	.95
GAD-7[c]	5	−0.68 (−1.31 to −0.04)		
Other measurements[d]	10	−0.71 (−1.61 to 0.19)		
Measurement of anxiety2			2.06	.15
DASS-21[e]	4	−0.23 (−0.38 to −0.08)		
Other measurements	11	−0.85 (−1.68 to −0.02)		
Duration of follow-up			1.14	.29
Short-term	12	−0.80 (−1.55 to −0.06)		
Long-term	3	−0.29 (−0.86 to 0.29)		
Duration of intervention			0.58	.45
≥4 weeks	9	−0.47 (−0.91 to −0.04)		
<4 weeks	6	−1.07 (−2.56 to 0.41)		

[a] SMD: standardized mean difference.

[b] LMICs: low- and middle-income countries.

[c] GAD-7: Generalized Anxiety Disorder-7.

[d] Other measurements including Beck Anxiety Inventory, Brief Symptom Inventory, Child Behavior Checklist, College Counseling Center Assessment of Psychological Symptoms, Hamilton Anxiety Rating Scale, Self-rating Anxiety Scale, and State-Trait Anxiety Inventory.

[e] DASS−21: Depression Anxiety Stress Scales-21.

Risk of Bias Assessments

For the meta-analysis, we evaluated the risk of bias for the included RCTs, and the results are presented in Multimedia Appendix 6 [22,24,29,32,34,43,44,47-49,62,67,69,74,79,81]. Overall, the included studies showed low risk of bias. Of the 16 RCTs included, 6 (38%) studies were rated "low" risk; 7 (44%) studies were rated "some concerns"; and 3 (19%) studies were rated "high" risk. Owing to the absence of previously published analysis plans, bias in reporting was assessed as "some concerns" for all 16 studies.

Discussion

Principal Findings

To our knowledge, this is the most comprehensive review of the evidence examining the feasibility, acceptability, and efficacy of digital mental health services during the COVID-19 pandemic. The 65 studies identified in the scoping review reported using videoconferencing platforms, web-based programs, smartphone apps, SMS text messaging, social media, hotline, robotic telemedicine, and VR to deliver mental health services among 67,884 participants across 18 countries. Psychotherapy, psychoeducation, and psychological support were among the most common mental health services delivered during the infectious disease outbreak. Overall, digital mental health services delivered during the initial phase of COVID-19 were usable, feasible, and acceptable and improved psychological well-being in participants across the included studies in high-income countries and LMICs. Sixteen RCTs that reported comparable outcomes, including for depression or anxiety, were systematically reviewed and meta-analyzed. The results indicated that digital mental health interventions were significantly associated with a small reduction in depressive symptoms and a moderate reduction in anxiety symptoms. The significance and effect sizes of the interventions differed among the measurements of the depression.

Our study suggested that the feasibility, acceptability, and effectiveness of digital mental health services during the pandemic were similar to those before the pandemic [82,83], indicating the potential for digital intervention to respond to the mental health needs of the people during the pandemic. These needs include addressing mental health problems among people experiencing COVID-19–related symptoms of anxiety, depression, and stress; patients with suspected or confirmed COVID-19; health workers; and patients with preexisting mental disorders. The advantages of digital mental health services in promoting psychological well-being depend on the ability to overcome the physical distance between people with mental health needs and providers and increasing the accessibility of mental health services. Although the use of digital mental health services has long been advocated, it has not yet been widely adopted because of a lack of support from both service users and health care professionals. In most nations, the situation has changed after the COVID-19 pandemic [84]. In China, for example, increased use of web-based mental health assessments, psychoeducation, and psychological counseling were observed during the COVID-19 pandemic [85].

Different strategies should be used for applying suitable techniques to digital mental health services. In our review, we found that smartphone app outreach was used by most service users (79.2%). Individuals with mild to moderate mental health symptoms, which may be addressed by psychosocial support and psychological interventions based on a stepped-care approach, are the target populations for the expanded use of smartphone apps to improve the delivery of mental health services. The technology with broad reach could also be used for psychoeducation [46], which is necessary for the support and prevention of mental disorders. Previous research has found that using apps or other web-based treatments for the older adult population can be challenging. Our study found that older adults have positive attitudes toward and may benefit from the electronic device–based mental health services, including videoconferencing platforms [45,62], social media [54], and web-based programs [77]. We suggest that smartphone apps and other web-based interventions should be considered for the older age population if technical assistance could be provided for people with difficulties with these devices. Telephone calls that are easy to use should also be considered for older adults [47].

The use of digital mental health services was reported across high-income countries and middle-income countries, but there were no studies from low-income countries. Digital mental health services have great potential in LMICs because of the limited availability of mental health services to cover the population needs [86] along with diminished health system capacity during the pandemic. Promising findings were reported for digital mental health services (videoconferencing platforms, web-based programs, and smartphone apps) in most high-income countries and some middle-income countries. However, these techniques rely on the internet and may be difficult to replicate in low-income areas. In addition, limited studies have examined the use of SMS text messaging and telephone calls in middle-income countries. Techniques that do not rely on the internet are a priority in low-resource areas. For instance, owing to their great accessibility and low cost of use, SMS text messaging and hotline might be considered for use in resource-limited or isolated locations. The benefits of delivering digital mental health services in remote and low-resource areas, in terms of potential impact and resource savings, must be considered alongside the feasibility of achieving high coverage of mental health services.

The finding that digital mental health services are associated with the amelioration of anxiety and depressive symptoms is consistent with previous evidence examining their effects [9,87]. Subgroup and meta-regression analyses allowed us to identify the effects of the study and population characteristics. For instance, there was no evidence of differences in SMD based on sample size. We found that depression measures contributed to heterogeneity. The posttreatment effects of digital mental health interventions differed from Depression Anxiety Stress Scales-21 (DASS-21) and other measurements. Although the current evidence supports the short-term effectiveness of most digital mental health services, a few studies have evaluated long-term outcomes after the interventions [43,69,79]. The long-term benefits of receiving digital mental health services should be assessed through practical evaluations. Along with the viability of reaching high coverage, it is important to weigh the potential impacts and resource savings of providing digital mental health care in rural and underresourced regions.

Our review sheds light on future research. First, further efforts should be focused on how digital mental health services can be integrated with existing traditional treatment modalities to form a complementary and integrated service model. In particular, the integration of medication and digital services remains to be further validated for safety and feasibility. Second, most studies included in our review were digital psychotherapy interventions, suggesting that digital technologies may have facilitated the accessibility of psychotherapy during the pandemic. We found evidence supporting the efficacious delivery of psychotherapies, either guided [62,77] or self-help therapy [31,34,64,74]. In a study by Al-Alawi et al [18], therapist-guided web-based therapy was associated with greater effectiveness than self-help therapy for people with anxiety and depressive symptoms during the COVID-19 pandemic. Future studies should focus on determining whether people require a guided web-based psychotherapy approach and promote individualized treatment to ensure the optimal allocation of treatment resources. Third, despite the diversity of mental disorders, only a few are covered in this review. Depression and anxiety were the most common outcomes measured. Thus, the role of digital services for many severe mental illnesses, including schizophrenia and bipolar disorder, needs to be further examined.

Our study has implications for policy development in health care. During a pandemic, digital mental health technology can be used as a backup to address psychological demands and broaden service coverage. We urge policy makers to create future industry regulatory laws so that the digital health sector can grow in a way that provides security and effectiveness testing. The development of digital health infrastructure, including information transmission facilities and technical testing methods for data security issues, is also a crucial prerequisite for the future development of digital mental health. If digital mental health services are to continue, there is a need to specify the minimum level of privacy and security that is acceptable. For example, in the United States, digital mental health services have rapidly made changes in response to policies and regulations concerning confidentiality and privacy to promote telehealth delivery of care [88].

Limitations

Our study has several limitations that should be noted. First, owing to the nature of our research question, the scoping review demonstrated substantial heterogeneity. In the pool of included studies, there was a high heterogeneity in terms of study design, type of techniques, type of mental health services, intervention components, target conditions, populations, and outcomes. Most studies did not have comparisons, and their effectiveness remains to be examined in future studies. For meta-analyses, we found only a few factors that contribute to heterogeneity between studies. The subgroup analyses were based on outcomes and limited by the variations in the type of techniques and type of mental health services. As a result, the findings should be interpreted with caution and read in terms of implications for future research.

Second, most participants in our scoping review were from high-income nations, primarily the United States and European countries. There is an urgent need for more independent research that examines the viability, acceptability, and efficacy of digital mental health as a response to pandemics in LMICs, particularly in low-income countries.

Third, although the current evidence supports the short-term effectiveness of most digital mental health services, only 2 studies have evaluated the long-term outcomes. A practical evaluation of the long-term benefits of digital mental health services should be undertaken, particularly in urban contexts.

Fourth, this study did not include preprint servers in the search, which may lead to the overestimation of our findings. However, a methodological research study of systematic reviews and meta-analyses revealed that including unpublished data (such as preprints) may add new sources of bias, although it does not change the results of reviews [89]. In addition, there is a possibility that the findings in preprints will be changed further, which leads to the results of the review being amended as well.

Finally, the target of this study was to evaluate the usability, feasibility, acceptability, and initial efficacy of digital mental health services in the initial phase of the pandemic. Therefore, we limited the data search to the first 2 years of the pandemic. We recommend that future studies should assess whether these services will remain the same or be further refined during other phases of the pandemic.

Conclusions

In summary, we have reported a range of digital mental health services delivered during the initial phase of the COVID-19 pandemic. We found that digital mental health services were usable, feasible, acceptable, and effective in response to mental health needs in the initial phase of pandemic across high-income countries and middle-income countries. Our findings also highlight the amelioration of depressive and anxiety symptoms based on the pooled results of the RCTs. To date, the strategies to lessen the COVID-19 pandemic's negative impacts on mental health remain a subject of much debate. We suggest that policy makers develop, implement, and assess different strategies for addressing the long-term mental health needs of the pandemic and its aftermath. Digital mental health services as novel strategies for promoting psychological well-being can purposively address physical barriers and provide a promising element for future integrated mental health service models in the post–COVID-19 era.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (grant 82101575) and Science and Technology Program of Guangzhou (grant 202102020702).

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklists.
[DOCX File , 54 KB - publichealth_v9i1e43730_app1.docx]

Multimedia Appendix 2
Search databases and strategy.
[DOCX File , 34 KB - publichealth_v9i1e43730_app2.docx]

Multimedia Appendix 3
Summary of the findings of individual studies included in the qualitative systematic review.
[DOCX File , 145 KB - publichealth_v9i1e43730_app3.docx]

Multimedia Appendix 4
Results of the meta-regression for depression and anxiety.
[DOCX File , 324 KB - publichealth_v9i1e43730_app4.docx]

Multimedia Appendix 5
Publication bias of meta-analyses.

[DOCX File , 114 KB - publichealth_v9i1e43730_app5.docx]

Multimedia Appendix 6
Risk of Bias across randomized studies for meta-analyses.
[DOCX File , 534 KB - publichealth_v9i1e43730_app6.docx]

References

1. COVID-19 Coronavirus pandemic. World-O-Meter. URL: https://www.worldometers.info/coronavirus/ [accessed 2022-10-05]
2. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
3. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. The Lancet Psychiatry 2020 Sep;7(9):813-824 [FREE Full text] [doi: 10.1016/S2215-0366(20)30307-2] [Medline: 32682460]
4. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors Associated With Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019. JAMA Netw Open 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
5. Yao H, Chen J, Xu Y. Patients with mental health disorders in the COVID-19 epidemic. The Lancet Psychiatry 2020 Apr;7(4):e21 [FREE Full text] [doi: 10.1016/S2215-0366(20)30090-0] [Medline: 32199510]
6. Lu Z, Shi L, Que J, Zheng Y, Wang Q, Liu W, et al. Long-term psychological profile of general population following COVID-19 outbreak: symptom trajectories and evolution of psychopathological network. Epidemiol Psychiatr Sci 2022 Sep 27;31:e69 [FREE Full text] [doi: 10.1017/S2045796022000518] [Medline: 36165185]
7. Mak IW, Chu CM, Pan PC, Yiu MG, Chan VL. Long-term psychiatric morbidities among SARS survivors. Gen Hosp Psychiatry 2009;31(4):318-326 [FREE Full text] [doi: 10.1016/j.genhosppsych.2009.03.001] [Medline: 19555791]
8. Yue J, Yan W, Sun Y, Yuan K, Su S, Han Y, et al. Mental health services for infectious disease outbreaks including COVID-19: a rapid systematic review. Psychol. Med 2020 Nov 05;50(15):2498-2513. [doi: 10.1017/s0033291720003888]
9. Karyotaki E, Efthimiou O, Miguel C, Bermpohl FM, Furukawa TA, Cuijpers P, Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration, et al. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. JAMA Psychiatry 2021 Apr 01;78(4):361-371 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
10. 11. Measure performance. Australian Government Digital Transformation Agency. URL: https://www.dta.gov.au/help-and-advice/digital-service-standard/digital-service-standard-criteria/11-measure-performance [accessed 2023-01-04]
11. Heynsbergh N, Heckel L, Botti M, Livingston PM. Feasibility, useability and acceptability of technology-based interventions for informal cancer carers: a systematic review. BMC Cancer 2018 Mar 02;18(1):244 [FREE Full text] [doi: 10.1186/s12885-018-4160-9] [Medline: 29499663]
12. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
13. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
14. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Cambridge, Massachusetts, United States: Academic Press; 1969.
15. Deeks J, Higgins J, Altman D. Analysing data and undertaking meta-analyses. In: Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Hoboken, New Jersey: Wiley; 2022.
16. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2019.
17. Wei N, Huang B, Lu S, Hu J, Zhou X, Hu C, et al. Efficacy of internet-based integrated intervention on depression and anxiety symptoms in patients with COVID-19. Journal of Zhejiang University 2020 May;21(5):400-404 [FREE Full text] [doi: 10.1631/jzus.B2010013] [Medline: 32425006]
18. Al-Alawi M, McCall RK, Sultan A, Al Balushi N, Al-Mahrouqi T, Al Ghailani A, et al. Efficacy of a six-week-long therapist-guided online therapy versus self-help internet-based therapy for COVID-19-induced anxiety and depression: open-label, pragmatic, randomized controlled trial. JMIR Ment Health 2021 Feb 12;8(2):e26683 [FREE Full text] [doi: 10.2196/26683] [Medline: 33512323]
19. Guan F, Wu Y, Ren W, Zhang P, Jing B, Xu Z, et al. Self-compassion and the mitigation of negative affect in the era of social distancing. Mindfulness 2021 Jun 28;12(9):2184-2195 [FREE Full text] [doi: 10.1007/s12671-021-01674-w] [Medline: 34221182]
20. Aguilera A, Hernandez-Ramos R, Haro-Ramos AY, Boone CE, Luo TC, Xu J, et al. A text messaging intervention (StayWell at home) to counteract depression and anxiety during COVID-19 social distancing: pre-post study. JMIR Ment Health 2021 Nov 01;8(11):e25298 [FREE Full text] [doi: 10.2196/25298] [Medline: 34543230]

21. Agyapong VI, Shalaby R, Hrabok M, Vuong W, Noble JM, Gusnowski A, et al. Mental health outreach via supportive text messages during the COVID-19 pandemic: improved mental health and reduced suicidal ideation after six weeks in subscribers of Text4Hope compared to a control population. Int J Environ Res Public Health 2021 Feb 23;18(4):2157 [FREE Full text] [doi: 10.3390/ijerph18042157] [Medline: 33672120]
22. Al-Refae M, Al-Refae A, Munroe M, Sardella NA, Ferrari M. A self-compassion and mindfulness-based cognitive mobile intervention (Serene) for depression, anxiety, and stress: promoting adaptive emotional regulation and wisdom. Front Psychol 2021 Mar 22;12:648087 [FREE Full text] [doi: 10.3389/fpsyg.2021.648087] [Medline: 33828514]
23. Bantjes J, Kazdin AE, Cuijpers P, Breet E, Dunn-Coetzee M, Davids C, et al. A web-based group cognitive behavioral therapy intervention for symptoms of anxiety and depression among university students: open-label, pragmatic trial. JMIR Ment Health 2021 May 27;8(5):e27400 [FREE Full text] [doi: 10.2196/27400] [Medline: 34042598]
24. Ben-Zeev D, Chander A, Tauscher J, Buck B, Nepal S, Campbell A, et al. A smartphone intervention for people with serious mental illness: fully remote randomized controlled trial of CORE. J Med Internet Res 2021 Nov 12;23(11):e29201 [FREE Full text] [doi: 10.2196/29201] [Medline: 34766913]
25. Brouzos A, Vassilopoulos SP, Baourda VC, Tassi C, Stavrou V, Moschou K, et al. "Staying Home - Feeling Positive": Effectiveness of an on-line positive psychology group intervention during the COVID-19 pandemic. Curr Psychol 2021 Mar 20:1-13 [FREE Full text] [doi: 10.1007/s12144-021-01613-x] [Medline: 33776381]
26. Casas J. Evaluating the feasibility and acceptability of written exposure therapy delivered via telehealth for the treatment of post-traumatic stress disorder. University of Nevada, Reno. 2021. URL: https://www.proquest.com/openview/22c4a4bc7c2d0a10c02d7b3270b566a2/1?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2022-09-05]
27. Chandra M, Rai CB, Sandhu VK, Kumari N, Vishnoi S, Aman S, et al. Mhealth based mental health support counselling service for covid-19 suspect and positive patients in isolation facilities. J Indian Academy Clin Med 2021;22(1-2):6-11.
28. Coifman KG, Disabato DD, Seah TH, Ostrowski-Delahanty S, Palmieri PA, Delahanty DL, et al. Boosting positive mood in medical and emergency personnel during the COVID-19 pandemic: preliminary evidence of efficacy, feasibility and acceptability of a novel online ambulatory intervention. Occup Environ Med 2021 Apr 26;78(8):555-557. [doi: 10.1136/oemed-2021-107427] [Medline: 33903280]
29. Comer JS, Furr JM, Del Busto C, Silva K, Hong N, Poznanski B, et al. Therapist-led, internet-delivered treatment for early child social anxiety: a waitlist-controlled evaluation of the iCALM telehealth program. Behav Ther 2021 Sep;52(5):1171-1187. [doi: 10.1016/j.beth.2021.01.004] [Medline: 34452671]
30. Craig SL, Leung VW, Pascoe R, Pang N, Iacono G, Austin A, et al. AFFIRM Online: utilising an affirmative cognitive-behavioural digital intervention to improve mental health, access, and engagement among LGBTQA+ youth and young adults. Int J Environ Res Public Health 2021 Feb 05;18(4):1541 [FREE Full text] [doi: 10.3390/ijerph18041541] [Medline: 33562876]
31. Detweiler Guarino I, Cowan DR, Fellows AM, Buckey JC. Use of a self-guided computerized cognitive behavioral tool during COVID-19: evaluation study. JMIR Form Res 2021 May 31;5(5):e26989 [FREE Full text] [doi: 10.2196/26989] [Medline: 33973856]
32. Dincer B, Inangil D. The effect of Emotional Freedom Techniques on nurses' stress, anxiety, and burnout levels during the COVID-19 pandemic: a randomized controlled trial. Explore 2021 Mar;17(2):109-114 [FREE Full text] [doi: 10.1016/j.explore.2020.11.012] [Medline: 33293201]
33. Esentürk OK, Yarımkaya E. Whatsapp-based physical activity intervention for children with autism spectrum disorder during the novel coronavirus (COVID-19) pandemic: a feasibility trial. Adapt Phys Activ Q 2021 Oct 01;38(4):569-584. [doi: 10.1123/apaq.2020-0109] [Medline: 34193630]
34. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, Zamanillo-Campos R, Yáñez-Juan AM, Bennasar-Veny M, et al. A mobile phone-based intervention to reduce mental health problems in health care workers during the COVID-19 pandemic (PsyCovidApp): randomized controlled trial. JMIR Mhealth Uhealth 2021 May 18;9(5):e27039 [FREE Full text] [doi: 10.2196/27039] [Medline: 33909587]
35. Gabrielli S, Rizzi S, Bassi G, Carbone S, Maimone R, Marchesoni M, et al. Engagement and effectiveness of a healthy-coping intervention via chatbot for university students during the COVID-19 pandemic: mixed methods proof-of-concept study. JMIR Mhealth Uhealth 2021 May 28;9(5):e27965 [FREE Full text] [doi: 10.2196/27965] [Medline: 33950849]
36. Geoffroy PA, Le Goanvic V, Sabbagh O, Richoux C, Weinstein A, Dufayet G, et al. Psychological support system for hospital workers during the Covid-19 outbreak: rapid design and implementation of the Covid-Psy hotline. Front Psychiatry 2020 May 28;11:511 [FREE Full text] [doi: 10.3389/fpsyt.2020.00511] [Medline: 32670100]
37. Gordon JS, Sbarra D, Armin J, Pace TW, Gniady C, Barraza Y. Use of a guided imagery mobile app (See Me Serene) to reduce COVID-19-related stress: pilot feasibility study. JMIR Form Res 2021 Oct 04;5(10):e32353 [FREE Full text] [doi: 10.2196/32353] [Medline: 34546941]
38. Graziano S, Boldrini F, Righelli D, Milo F, Lucidi V, Quittner A, et al. Psychological interventions during COVID pandemic: telehealth for individuals with cystic fibrosis and caregivers. Pediatr Pulmonol 2021 Jul 27;56(7):1976-1984 [FREE Full text] [doi: 10.1002/ppul.25413] [Medline: 33905614]

39. Gromatsky M, Sullivan SR, Mitchell EL, Spears AP, Edwards ER, Goodman M. Feasibility and acceptability of VA CONNECT: caring for our nation's needs electronically during the COVID-19 transition. Psychiatry Res 2021 Feb;296:113700 [FREE Full text] [doi: 10.1016/j.psychres.2020.113700] [Medline: 33422845]
40. Guo L, Wu M, Zhu Z, Zhang L, Peng S, Li W, et al. Effectiveness and influencing factors of online education for caregivers of patients with eating disorders during COVID-19 pandemic in China. Eur Eat Disord Rev 2020 Nov 27;28(6):816-825 [FREE Full text] [doi: 10.1002/erv.2783] [Medline: 32852142]
41. Held P, Klassen BJ, Coleman JA, Thompson K, Rydberg TS, Van Horn R. Delivering intensive PTSD treatment virtually: the development of a 2-week intensive cognitive processing therapy-based program in response to COVID-19. Cogn Behav Pract 2021 Nov;28(4):543-554 [FREE Full text] [doi: 10.1016/j.cbpra.2020.09.002] [Medline: 34629839]
42. Hom MA, Weiss RB, Millman ZB, Christensen K, Lewis EJ, Cho S, et al. Development of a virtual partial hospital program for an acute psychiatric population: lessons learned and future directions for telepsychotherapy. J Psychother Integration 2020 Jun;30(2):366-382. [doi: 10.1037/int0000212]
43. Hosseinzadeh Asl NR. A randomized controlled trial of a mindfulness-based intervention in social workers working during the COVID-19 crisis. Curr Psychol 2022 Aug 09;41(11):8192-8199 [FREE Full text] [doi: 10.1007/s12144-021-02150-3] [Medline: 34393464]
44. Hu J, Cai Z, Ma X. Effects of WeChat-based psychological interventions on the mental health of patients with suspected new coronavirus pneumonia: a pilot study. Jpn J Nurs Sci 2021 Oct 17;18(4):e12429 [FREE Full text] [doi: 10.1111/jjns.12429] [Medline: 33999501]
45. Ibrahim A, Chong MC, Khoo S, Wong LP, Chung I, Tan MP. Virtual group exercises and psychological status among community-dwelling older adults during the COVID-19 pandemic-a feasibility study. Geriatrics (Basel) 2021 Mar 22;6(1):31 [FREE Full text] [doi: 10.3390/geriatrics6010031] [Medline: 33810155]
46. Jaworski BK, Taylor K, Ramsey KM, Heinz A, Steinmetz S, Pagano I, et al. Exploring usage of COVID coach, a public mental health app designed for the COVID-19 pandemic: evaluation of analytics data. J Med Internet Res 2021 Mar 01;23(3):e26559 [FREE Full text] [doi: 10.2196/26559] [Medline: 33606656]
47. Kahlon MK, Aksan N, Aubrey R, Clark N, Cowley-Morillo M, Jacobs EA, et al. Effect of layperson-delivered, empathy-focused program of telephone calls on loneliness, depression, and anxiety among adults during the COVID-19 pandemic: a randomized clinical trial. JAMA Psychiatry 2021 Jun 01;78(6):616-622 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.0113] [Medline: 33620417]
48. Karagiozi K, Margaritidou P, Tsatali M, Marina M, Dimitriou T, Apostolidis H, et al. Comparison of on site versus online psycho education groups and reducing caregiver burden. Clin Gerontol 2022 Jul 05;45(5):1330-1340. [doi: 10.1080/07317115.2021.1940409] [Medline: 34219617]
49. Kim YJ. The effect of tele-acupressure self-practice for mental health and wellbeing in the community during COVID-19. Current Psychiatry Res Rev 2021 Jan 15;16(4):267-274. [doi: 10.2174/2666082216999201117124408]
50. Kolbe L, Jaywant A, Gupta A, Vanderlind WM, Jabbour G. Use of virtual reality in the inpatient rehabilitation of COVID-19 patients. Gen Hosp Psychiatry 2021 Jul;71:76-81 [FREE Full text] [doi: 10.1016/j.genhosppsych.2021.04.008] [Medline: 33964789]
51. Lazzaroni E, Invernizzi R, Fogliato E, Pagani M, Maslovaric G. Coronavirus disease 2019 emergency and remote eye movement desensitization and reprocessing group therapy with adolescents and young adults: overcoming lockdown with the butterfly hug. Front Psychol 2021;12:701381 [FREE Full text] [doi: 10.3389/fpsyg.2021.701381] [Medline: 34512458]
52. Lima MR, Wairagkar M, Natarajan N, Vaitheswaran S, Vaidyanathan R. Robotic telemedicine for mental health: a multimodal approach to improve human-robot engagement. Front Robot AI 2021 Mar 18;8:618866 [FREE Full text] [doi: 10.3389/frobt.2021.618866] [Medline: 33816568]
53. Maldonado J. The effect of an online mental health promotion program during COVID-19 doctor of nursing practice projects. Andrews University. 2020 Oct. URL: https://digitalcommons.andrews.edu/dnp/10/ [accessed 2022-09-10]
54. McKeon G, Tiedemann A, Sherrington C, Teasdale S, Mastrogiovanni C, Wells R, et al. Feasibility of an online, mental health-informed lifestyle program for people aged 60+ years during the COVID-19 pandemic. Health Promot J Austr 2022 Jul 19;33(3):545-552. [doi: 10.1002/hpja.538] [Medline: 34496102]
55. Nauphal M, Swetlitz C, Smith L, Rosellini AJ. A preliminary examination of the acceptability, feasibility, and effectiveness of a telehealth cognitive-behavioral therapy group for social anxiety disorder. Cognit Behavioral Pract 2021 Nov;28(4):730-742. [doi: 10.1016/j.cbpra.2021.04.011]
56. Ortiz R, Southwick L, Schneider R, Klinger EV, Pelullo A, Guntuku SC, et al. Improving mood through community connection and resources using an interactive digital platform: development and usability study. JMIR Ment Health 2021 Feb 26;8(2):e25834 [FREE Full text] [doi: 10.2196/25834] [Medline: 33635280]
57. Parolin L, Benzi I, Fanti E, Milesi A, Cipresso P, Preti E. Italia Ti Ascolto [Italy, I am listening]: an app-based group psychological intervention during the COVID-19 pandemic. OSF Preprints. 2021. URL: https://doi.org/10.4081/ripppo.2021.517 [accessed 2022-08-30]
58. Philip P, Dupuy L, Morin CM, de Sevin E, Bioulac S, Taillard J, et al. Smartphone-based virtual agents to help individuals with sleep concerns during COVID-19 confinement: feasibility study. J Med Internet Res 2020 Dec 18;22(12):e24268 [FREE Full text] [doi: 10.2196/24268] [Medline: 33264099]

59. Puspitasari AJ, Heredia D, Gentry M, Sawchuk C, Theobald B, Moore W, et al. Rapid adoption and implementation of telehealth group psychotherapy during COVID 19: practical strategies and recommendations. Cogn Behav Pract 2021 Nov;28(4):492-506 [FREE Full text] [doi: 10.1016/j.cbpra.2021.05.002] [Medline: 34188434]
60. Rojas SM, Gold SD, Bryan CJ, Pruitt LD, Felker BL, Reger MA. Brief cognitive-behavioral therapy for suicide prevention (BCBT-SP) via video telehealth: a case example during the COVID-19 outbreak. Cogn Behav Pract 2022 May;29(2):446-453 [FREE Full text] [doi: 10.1016/j.cbpra.2020.12.001] [Medline: 34602808]
61. Shalaby R, Vuong W, Hrabok M, Gusnowski A, Mrklas K, Li D, et al. Gender differences in satisfaction with a text messaging program (Text4Hope) and anticipated receptivity to technology-based health support during the COVID-19 pandemic: cross-sectional survey study. JMIR Mhealth Uhealth 2021 Apr 15;9(4):e24184 [FREE Full text] [doi: 10.2196/24184] [Medline: 33750738]
62. Shapira S, Yeshua-Katz D, Cohn-Schwartz E, Aharonson-Daniel L, Sarid O, Clarfield AM. A pilot randomized controlled trial of a group intervention via Zoom to relieve loneliness and depressive symptoms among older persons during the COVID-19 outbreak. Internet Interv 2021 Apr;24:100368 [FREE Full text] [doi: 10.1016/j.invent.2021.100368] [Medline: 33527072]
63. Sharrock MJ, Mahoney AE, Haskelberg H, Millard M, Newby JM. The uptake and outcomes of internet-based cognitive behavioural therapy for health anxiety symptoms during the COVID-19 pandemic. J Anxiety Disord 2021 Dec;84:102494 [FREE Full text] [doi: 10.1016/j.janxdis.2021.102494] [Medline: 34763219]
64. Song J, Jiang R, Chen N, Qu W, Liu D, Zhang M, et al. Self-help cognitive behavioral therapy application for COVID-19-related mental health problems: a longitudinal trial. Asian J Psychiatr 2021 Jun;60:102656 [FREE Full text] [doi: 10.1016/j.ajp.2021.102656] [Medline: 33866282]
65. Sosa Lovera A, Ureña AJ, Arias J, Araujo Rodríguez A, Canario Guzmán JA. Psychological helpline in response to the COVID-19 pandemic in the Dominican Republic. Couns Psychother Res 2022 Jun 29;22(2):534-541 [FREE Full text] [doi: 10.1002/capr.12482] [Medline: 34899067]
66. Sturgill R, Martinasek M, Schmidt T, Goyal R. A novel artificial intelligence-powered emotional intelligence and mindfulness app (Ajivar) for the college student population during the COVID-19 pandemic: quantitative questionnaire study. JMIR Form Res 2021 Jan 05;5(1):e25372 [FREE Full text] [doi: 10.2196/25372] [Medline: 33320822]
67. Suffoletto B, Goldstein T, Gotkiewicz D, Gotkiewicz E, George B, Brent D. Acceptability, engagement, and effects of a mobile digital intervention to support mental health for young adults transitioning to college: pilot randomized controlled trial. JMIR Form Res 2021 Oct 14;5(10):e32271 [FREE Full text] [doi: 10.2196/32271] [Medline: 34647893]
68. Summers C, Wu P, Taylor AJ. Supporting mental health during the COVID-19 pandemic using a digital behavior change intervention: an open-label, single-arm, pre-post intervention study. JMIR Form Res 2021 Oct 06;5(10):e31273 [FREE Full text] [doi: 10.2196/31273] [Medline: 34459740]
69. Sun S, Lin D, Goldberg S, Shen Z, Chen P, Qiao S, et al. A mindfulness-based mobile health (mHealth) intervention among psychologically distressed university students in quarantine during the COVID-19 pandemic: a randomized controlled trial. J Couns Psychol 2022 Mar;69(2):157-171. [doi: 10.1037/cou0000568] [Medline: 34264696]
70. Tarquinio C, Brennstuhl M, Rydberg JA, Bassan F, Peter L, Tarquinio CL, et al. EMDR in telemental health counseling for healthcare workers caring for COVID-19 patients: a pilot study. Issues Ment Health Nurs 2021 Jan 14;42(1):3-14. [doi: 10.1080/01612840.2020.1818014] [Medline: 33052727]
71. Vallefuoco E, Purpura G, Gison G, Bonifacio A, Tagliabue L, Broggi F, et al. A multidisciplinary telerehabilitation approach for supporting social interaction in autism spectrum disorder families: an Italian digital platform in response to COVID-19. Brain Sci 2021 Oct 25;11(11):1404 [FREE Full text] [doi: 10.3390/brainsci11111404] [Medline: 34827403]
72. van Agteren J, Ali K, Fassnacht DB, Iasiello M, Furber G, Howard A, et al. Testing the differential impact of an internet-based mental health intervention on outcomes of well-being and psychological distress during COVID-19: uncontrolled intervention study. JMIR Ment Health 2021 Sep 15;8(9):e28044 [FREE Full text] [doi: 10.2196/28044] [Medline: 34357876]
73. Wagner L, Corona LL, Weitlauf AS, Marsh KL, Berman AF, Broderick NA, et al. Use of the TELE-ASD-PEDS for autism evaluations in response to COVID-19: preliminary outcomes and clinician acceptability. J Autism Dev Disord 2021 Sep 30;51(9):3063-3072 [FREE Full text] [doi: 10.1007/s10803-020-04767-y] [Medline: 33125622]
74. Wahlund T, Mataix-Cols D, Olofsdotter Lauri K, de Schipper E, Ljótsson B, Aspvall K, et al. Brief online cognitive behavioural intervention for dysfunctional worry related to the COVID-19 pandemic: a randomised controlled trial. Psychother Psychosom 2021;90(3):191-199 [FREE Full text] [doi: 10.1159/000512843] [Medline: 33212440]
75. Wasil AR, Taylor ME, Franzen RE, Steinberg JS, DeRubeis RJ. Promoting graduate student mental health during COVID-19: acceptability, feasibility, and perceived utility of an online single-session intervention. Front Psychol 2021 Apr 7;12:569785 [FREE Full text] [doi: 10.3389/fpsyg.2021.569785] [Medline: 33897511]
76. Wood HJ, Gannon JM, Chengappa KR, Sarpal DK. Group teletherapy for first-episode psychosis: piloting its integration with coordinated specialty care during the COVID-19 pandemic. Psychol Psychother 2021 Jun 20;94(2):382-389. [doi: 10.1111/papt.12310] [Medline: 33078496]
77. Ying Y, Ji Y, Kong F, Chen Q, Lv Y, Hou Y, et al. Internet-based cognitive behavioral therapy for psychological distress in older adults without cognitive impairment living in nursing homes during the COVID-19 pandemic: a feasibility study. Internet Interv 2021 Dec;26:100461 [FREE Full text] [doi: 10.1016/j.invent.2021.100461] [Medline: 34631432]

78. Zepeda MS, Deighton S, Markova V, Madsen JW, Racine N. iCOPE with COVID-19: a brief telemental health intervention for children and adolescents during the COVID-19 pandemic. Cogn Behav Pract 2021 Oct 22 [FREE Full text] [doi: 10.1016/j.cbpra.2021.10.001] [Medline: 34703196]
79. Zhang H, Zhang A, Liu C, Xiao J, Wang K. A brief online mindfulness-based group intervention for psychological distress among chinese residents during COVID-19: a pilot randomized controlled trial. Mindfulness (N Y) 2021;12(6):1502-1512 [FREE Full text] [doi: 10.1007/s12671-021-01618-4] [Medline: 33758629]
80. Zimmerman M, Benjamin I, Tirpak JW, D'Avanzato C. Patient satisfaction with partial hospital telehealth treatment during the COVID-19 pandemic: comparison to in-person treatment. Psychiatry Res 2021 Jul;301:113966 [FREE Full text] [doi: 10.1016/j.psychres.2021.113966] [Medline: 33990071]
81. Van Lieshout RJ, Layton H, Savoy CD, Brown JS, Ferro MA, Streiner DL, et al. Effect of online 1-day cognitive behavioral therapy-based workshops plus usual care vs usual care alone for postpartum depression: a randomized clinical trial. JAMA Psychiatry 2021 Nov 01;78(11):1200-1207 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.2488] [Medline: 34495285]
82. Zhang X, Lewis S, Firth J, Chen X, Bucci S. Digital mental health in China: a systematic review. Psychol. Med 2021 Sep 28;51(15):2552-2570. [doi: 10.1017/s0033291721003731]
83. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500. [doi: 10.1016/s2215-0366(17)30096-2]
84. Shore JH, Schneck CD, Mishkind MC. Telepsychiatry and the coronavirus disease 2019 pandemic-current and future outcomes of the rapid virtualization of psychiatric care. JAMA Psychiatry 2020 Dec 01;77(12):1211-1212. [doi: 10.1001/jamapsychiatry.2020.1643] [Medline: 32391861]
85. Liu S, Yang L, Zhang C, Xiang Y, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e17-e18. [doi: 10.1016/S2215-0366(20)30077-8] [Medline: 32085841]
86. Merchant R, Torous J, Rodriguez-Villa E, Naslund JA. Digital technology for management of severe mental disorders in low-income and middle-income countries. Curr Opin Psychiatry 2020 Sep;33(5):501-507 [FREE Full text] [doi: 10.1097/YCO.0000000000000626] [Medline: 32520747]
87. Sin J, Galeazzi G, McGregor E, Collom J, Taylor A, Barrett B, et al. Digital interventions for screening and treating common mental disorders or symptoms of common mental illness in adults: systematic review and meta-analysis. J Med Internet Res 2020 Sep 02;22(9):e20581 [FREE Full text] [doi: 10.2196/20581] [Medline: 32876577]
88. Bartels SJ, Baggett TP, Freudenreich O, Bird BL. COVID-19 emergency reforms in massachusetts to support behavioral health care and reduce mortality of people with serious mental illness. Psychiatr Serv 2020 Oct 01;71(10):1078-1081. [doi: 10.1176/appi.ps.202000244] [Medline: 32487009]
89. Schmucker CM, Blümle A, Schell LK, Schwarzer G, Oeller P, Cabrera L, OPEN consortium. Systematic review finds that study data not published in full text articles have unclear impact on meta-analyses results in medical research. PLoS One 2017;12(4):e0176210 [FREE Full text] [doi: 10.1371/journal.pone.0176210] [Medline: 28441452]

Abbreviations

CBT: cognitive behavioral therapy
DASS-21: Depression Anxiety Stress Scales-21
LMICs: low- and middle-income countries
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: PRISMA Extension for Scoping Reviews
RCT: randomized controlled trial
SMD: standardized mean difference
VR: virtual reality

Edited by Y Khader; submitted 22.10.22; peer-reviewed by T Sathish, M Musker, C Leung; comments to author 02.12.22; revised version received 09.01.23; accepted 11.01.23; published 13.02.23.

Please cite as:
Zhong S, Yang X, Pan Z, Fan Y, Chen Y, Yu X, Zhou L
The Usability, Feasibility, Acceptability, and Efficacy of Digital Mental Health Services in the COVID-19 Pandemic: Scoping Review, Systematic Review, and Meta-analysis
JMIR Public Health Surveill 2023;9:e43730
URL: https://publichealth.jmir.org/2023/1/e43730
doi:10.2196/43730
PMID:36634261

Original Paper

The Impact of Digital Transformation on Inpatient Care: Mixed Methods Study

Philipp Koebe[1], BSc, MSc, LLM; Sabine Bohnet-Joschko[1], Prof Dr

[1]Faculty of Management, Economics and Society, Witten/Herdecke University, Witten, Germany

Corresponding Author:
Philipp Koebe, BSc, MSc, LLM
Faculty of Management, Economics and Society
Witten/Herdecke University
Alfred-Herrhausen-Str. 50
Witten, 58455
Germany
Phone: 49 17621951063
Email: philipp.koebe@uni-wh.de

Abstract

Background: In the context of the digital transformation of all areas of society, health care providers are also under pressure to change. New technologies and a change in patients' self-perception and health awareness require rethinking the provision of health care services. New technologies and the extensive use of data can change provision processes, optimize them, or replace them with new services. The inpatient sector, which accounts for a particularly large share of health care spending, plays a major role in this regard.

Objective: This study examined the influences of current trends in digitization on inpatient service delivery.

Methods: We conducted a scoping review. This was applied to identify the international trends in digital transformation as they relate to hospitals. Future trends were considered from different perspectives. Using the defined inclusion criteria, international peer-reviewed articles published between 2016 and 2021 were selected. The extracted core trends were then contextualized for the German hospital sector with 12 experts.

Results: We included 44 articles in the literature analysis. From these, 8 core trends could be deduced. A heuristic impact model of the trends was derived from the data obtained and the experts' assessments. This model provides a development corridor for the interaction of the trends with regard to technological intensity and supply quality. Trend accelerators and barriers were identified.

Conclusions: The impact analysis showed the dependencies of a successful digital transformation in the hospital sector. Although data interoperability is of particular importance for technological intensity, the changed self-image of patients was shown to be decisive with regard to the quality of care. We show that hospitals must find their role in new digitally driven ecosystems, adapt their business models to customer expectations, and use up-to-date information and communications technologies.

(JMIR Public Health Surveill 2023;9:e40622) doi:10.2196/40622

KEYWORDS

digital transformation; digitization; health care provision; hospital; trends

Introduction

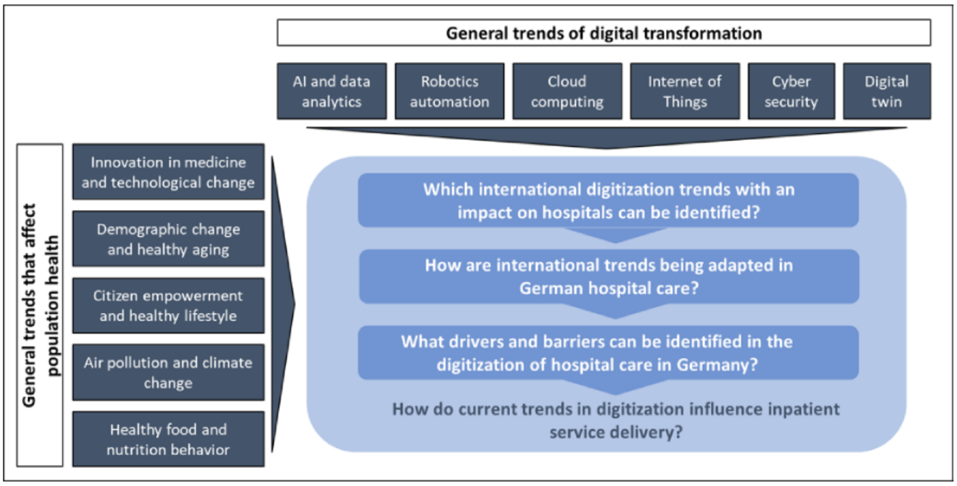
Background

Digital transformation is permeating all areas of society [1,2]. Scientific circles also refer to it as the Fourth Industrial Revolution [3]. On the basis of the latest advancements in medicine and in the context of general technological progress, digital trends also influence the provision of services in health care [4-7]. Health systems have become the focus of political and social debates in industrialized Western countries for a number of reasons [8,9]. Demographic developments have led to an increase in life expectancy [10]. This increase in life expectancy is primarily due to the medical-technical progress in the past as well as the comprehensive access to health care services in most high-income Western countries [11]. The result is an expansion in the use of services with an upward trend [12]. Furthermore, high life expectancy and the associated decline in the birth rate have led to a shift in the overall population structure in recent decades [10,11]. Fewer younger people have to care for a growing proportion of older people. This is already resulting in numerous bottlenecks [13]. Certain constraints include the financing of increasing health and nursing care services, the provision of comprehensive care even in less populated regions, and the staffing of health and nursing care facilities [14,15]. Owing to the personnel intensity of health services, many industrialized Western countries are already dependent on immigration for medical professionals. The use of new technologies and digital applications can reduce the financial and personnel burden on health care systems, especially on inpatient care [4-6,16].

On the one hand, there are numerous current trends resulting from digital transformation [17]. Artificial intelligence (AI), data analytics, further automation and robotization, and cloud computing or the Internet of Things (IoT) are driving digital permeation in society and the economy [17-20]. This is also accompanied by visionary technologies such as digital twins for use in virtual worlds such as the metaverse [18,21]. These trends are also creating a greater need for cybersecurity.

In contrast, general trends related to population health are leading to changing framework conditions [12,13]. Leap innovations in medicine and technological change, the aforementioned aging population with a need for healthy aging, the empowerment of the population, changing lifestyles, and healthy nutrition are influencing current developments in health care [22-24]. Environmental influences such as air pollution and climate change are also creating pressure for action in some health care fields [25-27]. As a result of these trends, various questions are being asked to better assess and control future developments. We make a special reference to German health care delivery as the German hospital and health care system can definitely be seen as a reference model for many other European and international health care settings [28,29]. This study considered the emerging trends for the future provision of care in hospitals and beyond in a patient-centered or outcome-oriented health care system that incorporates digital technologies [30-32]. Figure 1 shows the relationships between the general trends that have developed, with the purpose of deriving the research question.

Figure 1. Derivation of the research objective. AI: artificial intelligence.



Objectives

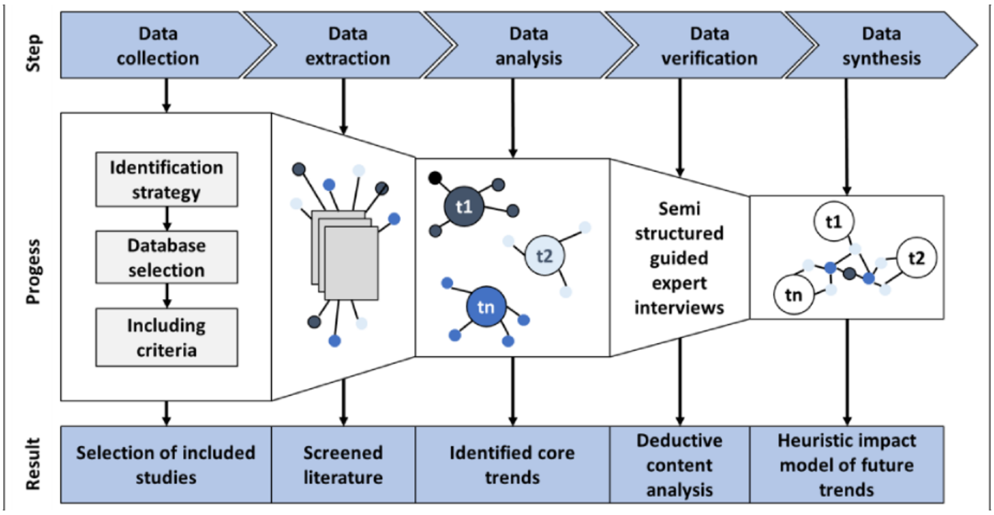
Our central research question addressed how current trends in digitization influence inpatient service delivery. We approached this question in 3 steps. First, we identified the international trends in digitization that have an impact on the hospital sector. Second, we addressed the adaptation of these trends with regard to the German hospital sector. Finally, we derived drivers and barriers and concluded which developments result from these trends.

Methods

Overview

We conducted a sequential study. Our mixed methods approach consisted of a scoping review and a qualitative survey. Our research design was based on 5 steps, as shown in Figure 2. The first step involved fundamental data collection. This resulted in the inclusion of studies, which was the starting point of our analysis. In the second step, data extraction took place. The available studies were compared, suitable classification features were determined, and their contents were exploited. The third step was data analysis and synthesis using a narrative approach, resulting in core trends. In the fourth step, a classification, verification, and discussion of the selected trends took place. The data obtained were subjected to a deductive content analysis. In the fifth step, a higher-level synthesis of both data collection points was performed. A heuristic logic model was used as an instrument.

Figure 2. Research methodology.



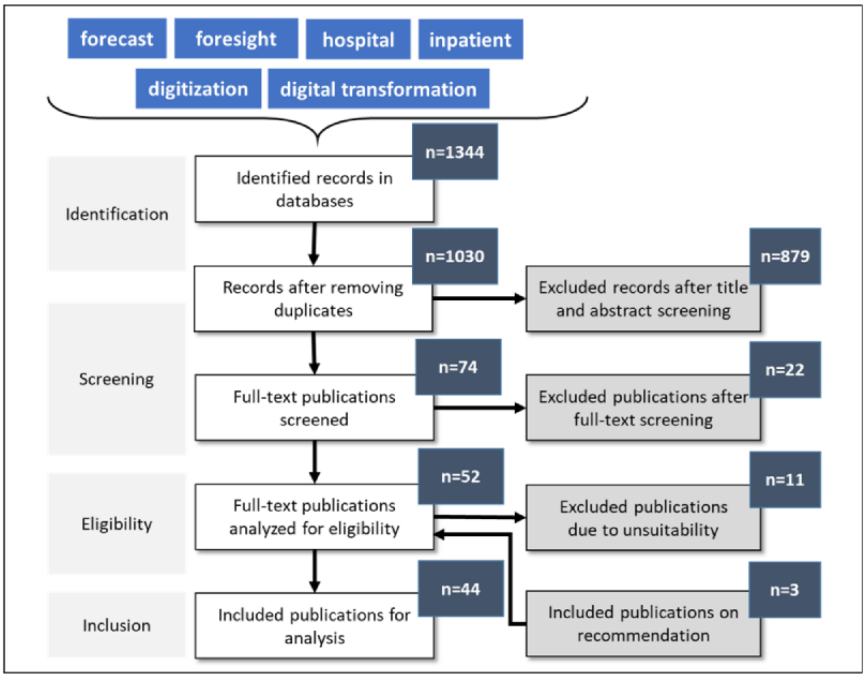
Scoping Review

Search Strategy

The scoping review is reported according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [33]. Accordingly, the following work stages were fulfilled: (1) definition of the research question, (2) systematic search, (3) selection of relevant studies, (4) data extraction, (5) data analysis and synthesis, and (6) interpretation of the results. We conducted our scoping review based on the extended current recommendations [34,35] and presented the results of our search strategy in a flowchart (Figure 3). Between October 2020 and February 2021, we conducted a systematic search of the following databases: PubMed, MEDLINE, EconLit, EconBiz, and ScienceDirect. An initial search was conducted in September 2020 to select an appropriate search strategy. Our search was limited in time to the last 5 years. Papers published between 2016 and 2021 were included. The inclusion of earlier papers would lead to bias in the results as the consideration of trends is only meaningful over a manageable period in the past [36,37]. The language was restricted to English. After screening the titles and abstracts, all articles dealing with future trends as well as those dealing with inpatient care or the service provision of the health care system and digitalization were included in the full-text search. In total, 2 reviewers performed the full-text screening according to these inclusion and exclusion criteria. Disagreements were discussed between the 2 reviewers to reach a consensus. In the case of a possible disagreement, a third author was appointed.

Our scoping review referred to current and future trends related to digital transformation on the one hand and service delivery in hospitals on the other. The words “forecast” and “foresight” were searched in combination with other terms (eg, “digital transformation,” “digitization,” and “hospital”). The search terms used in the various databases are shown in Multimedia Appendix 1. In a further step, the reference lists were searched backward for other relevant publications not listed in the aforementioned databases. In our review, we also included manuscripts that were recommended by authors or experts. Some search terms were discarded in the process of developing a search strategy. For example, the search terms “trend” or “future” were not included as they were too nonspecific and triggered too many irrelevant hits. The concept of service provision also had to be differentiated more precisely. The search strategy was ultimately refined to focus on hospitals.

Figure 3. Flowchart of literature search strategy.



Inclusion Criteria

Articles were included if the following 3 factors applied: a reference was made to the future or trends, a reference was made to the terms “digital” or “technology,” and a reference was made to inpatient service provision or hospitals. The documentation is shown in Multimedia Appendix 2 [20,38-80]. Articles were excluded if the terms “trend” or “future” were used to refer to a possible new treatment or diagnostic procedure; the prediction of the spread of disease; or specific developments in the fields of brain, genetic, or biotechnology research. Very frequently, the overall search showed the use of AI and machine learning (117 studies); IoT, wearables, and mobile health apps (86 studies); or big data (70 studies). Publications from a medical discipline or on medical indications were not included. The most common indications were related to cardiovascular disease (79 studies); oncology (96 studies); or imaging, particularly radiology (89 studies). However, of these, 1 study was included that was related to the organization of radiology departments in hospitals. Psychiatric indications and mental health were also not included in the study. Case studies and feasibility studies related to individual countries or organizations were also excluded, as were studies on trends in education and training in clinical settings (eg, the future of medical staff training). Publications related to the COVID-19 pandemic were not included. In total, the search provided 77 studies linked to COVID-19. The particular use of technology, such as blockchain or cybersecurity, was only included when there was genuinely a direct link to inpatient care. This occurred in only 1 case.

Study Inclusion

The identified articles from all databases were uploaded to the EndNote software (Clarivate Analytics), and duplicates were removed. In total, 2 researchers independently performed the first step of the screening process based on titles and abstracts, as reported in Multimedia Appendix 3. In the second step, 2 independent researchers carefully read the articles with available full texts to decide which would ultimately be included in the systematic review. Articles that met the eligibility criteria were selected for inclusion in the systematic review. The reasons for the exclusion of full-text articles were recorded and included in the PRISMA flow diagram [33]. The reference lists of the included articles were carefully searched by hand to identify other appropriate articles. In total, 2 researchers independently extracted the following data from each article: (1) study identification (study author and year of publication), (2) main trend of the study (technology, innovation, or process), and (3) key statements of the study (fields of application, opportunities, risks, and recommendations for action).

Any discrepancies that arose during the data extraction process were clarified through discussion among the authors. Subsequently, the extracted contextual factors were labeled according to their topic, and the studies were summed up so as to show the respective factors. In addition, the studies were classified using 2D scales. The first dimension described the inclusion of the user within the trend. The second dimension related to the intersectoral component. Furthermore, the trend factors were grouped into the 3 application fields: planning and steering, core clinical processes, and communication and data connection. For each of these areas, individual trend factors were described as having an impact on inpatient care. Each factor was double-checked by the 2 researchers, and incongruities were resolved through discussion. We used a narrative approach to synthesize the data [81]. Finally, the results of the synthesis were presented in a trend chart to contextualize the trends’ influences on other trends and illustrate the factors affecting each trend. In addition, barriers that contributed to the deceleration of the trends were extracted from the studies.

Expert Interviews

The identified trends were further processed for in-depth analysis and verification. The transfer of the trends that were extracted based on international scientific literature resulted in a qualitative survey of experts [82]. In this methodological component of our study, we focused on the German setting. Later, we will put these results back into a generalizable context so that the findings can also be applied to a broad spectrum of other countries. The expert interviews were conducted in the form of semistructured, guide-based interviews [83]. The interview guide was largely based on the identified core trends. The experts were selected based on the 2 decisive criteria of digitization experience in the health care sector and experience in or knowledge of the area of service provision.

Experts from 4 specific groups were interviewed in equal numbers: first, the group of inpatient care suppliers; second, the group of digital service providers in the health care industry; third, overarching institutions in the health care sector; and fourth, the group of relevant scientists. The number of interviews that needed to be conducted was not determined in advance. After 12 interviews, information saturation was reached [84]. The guide was tested for suitability in advance in a pretest with 3 participants and adjusted before final use. The data generated over the course of the interviews were analyzed deductively [85,86] in a sense-reconstructive, interpretative manner. The statements to be analyzed were compiled using the MAXQDA software (VERBI GmbH). The data analysis was performed by 2 independent researchers following a theory-driven definition of structuring dimensions to form the main categories and subcategories. An iterative process was used to establish definitions, identify anchor examples, and generate coding rules. The generated codes were subjected to a reliability test in several steps until the summary test of all the quality criteria was completed.

Data Synthesis

The extracted data from the literature and the obtained research data from the interviews were synthesized into a heuristic logic model [87,88]. Heuristic logic models are used in designing systematic review questions and collecting research results and as analytical tools [89]. Analytical logic models demonstrate a logical chain between inputs and outcomes. When the influences of the overarching relationship are demonstrated, it may not be necessary to examine all the individual components and their connections. We show a logical sequence of effects, including possible adverse effects, that influence inpatient care in a chain of identified trend characteristics. The model was used to illustrate the authoritative impact of the core trends on inpatient service delivery. This enabled the sketching of a trend development corridor and a more detailed description of the impact relationships [90]. Finally, the decelerating factors were also shown based on the data material.

Ethical Considerations, Informed Consent, and Participation

The study did not meet the need for ethical review as it did not involve special categories of personal data. All procedures performed in this study were in accordance with the ethical conducts of the General Data Protection Regulation and *Bundesdatenschutzgesetz* (German implementation of European Union regulations). Email invitations were sent to prospective candidates. The invitation included an overview of the project and the research scope. The voluntary nature and anonymity of participation were pointed out, together with complete information on the European Union General Data Protection Regulation. We requested written informed consent from all participants to record the interviews and use the information provided in our research.

Results

Results of the Scoping Review

The search strategy identified 1344 articles from the 5 databases, as shown in detail in Figure 3. After duplicates were removed, 76.64% (1030/1344) of the articles were retained for title and abstract screening. A total of 74 articulated full-text articles were read. After screening, 59% (44/74) of the studies [20,38-80] met the eligibility criteria for this systematic review.

The main characteristics of the included studies are listed in Multimedia Appendix 4 [20,38-80]. The classification into the topic field was performed to categorize a rough content assignment of the article. The core statements reflected the main contribution of the article and were later used to categorize the trends. The topic of AI and machine learning was addressed in 14% (6/44) of the studies [47,49,57,58,61,69] in different contexts. A total of 30% (13/44) of the studies [42,43,48,51,52,60,64,66,67,72,76,79,80] were originally related to data-driven topics (eg, data-driven resource use). The change in processes used by digitization was the subject of 25% (11/44) of the studies [40,44,46,50,55,63,68,70,74,75,78]. Among other things, processes in inpatient care that use electronic health records played a role. Mobile apps and IoT were the subject of 14% (6/44) of the studies [20,39,41,59,62,71]. These studies also showed close connections to data use and process redesign. The use of robots in clinical settings was addressed in 7% (3/44) of the studies [38,45,54]. The establishment of platforms or the development of a digitally connected ecosystem was considered more closely in 9% (4/44) of the studies [53,65,73,77]. A total of 2% (1/44) of the studies dealt explicitly with personalized medicine [56]. However, this topic was also addressed in numerous other studies (eg, with the use of AI or IoT). In all the studies, trends were described, and future developments were derived from them. In addition, they identified problem areas that will be compared at a later stage in this paper as internal and external trend barriers. In some cases, recommendations for action were derived or inferred from the barriers.

From the included studies, 8 trends were identified that affect hospitals as a result of digital transformations. For this purpose, 44 studies [20,38-80] were analyzed, the core topics were classified, the most important technologies were pinpointed, and the fields of application and hurdles were extracted, as reported in Multimedia Appendix 5 [20,38-80] (Textbox 1).

The core trends are presented in Table 1. In addition to naming the trends, the essence of each trend is described, as well as the most important and relevant technologies and the affected application fields within the hospital setting.

To illustrate the interrelationships between the individual trends, we presented the connecting factors in a matrix according to the number of times they were mentioned in the studies. Figure

4 illustrates which trends have a close linkage and shows the intensity with increasing red coloring. The closest links are between trends 3 and 4, with 15 common factors. Dense relationships were also evident between trends 4 and 7. Trend 1, in contrast, shows the fewest connections with other trends. The matrix represents an initial basis for further processing in the synthesis of the results, which will later also include the results of the interviews.

Textbox 1. Trends identified from the systematic review.

Trend 1—changing the patient role

- Patients are taking on a new role in the entire care process from prevention to diagnosis and therapy. They participate in decision-making, generate and use data, and have transparent information on diseases and available health care services. The most important technologies are the internet, smart devices, and mobile health apps.

Trend 2—connected, integrated delivery of care

- Connected technologies and cross-provider data exchanges promote intersectoral collaboration. Providers exchange and share data across the care process. Important technologies include electronic health records, the Internet of Things (IoT), and e-prescriptions.

Trend 3—data-driven resource allocation

- A data-based allocation of resources by means of predictive models enables the precise planning and control of care capacities. As a result, available resources are better allocated and used across an organization. Consequently, waiting times can be reduced, and patient throughput can be optimized. The most important technologies are data mining, machine learning, and predictive analytics.

Trend 4—performance optimization in primary processes

- The performance of core clinical processes in diagnostics and therapy is optimized. This involves using available data in real time to deploy equipment, rooms, or personnel efficiently. Relieving systems are used for the staff. In addition, services are patient-centric to meet the individual needs of patients within an institution. The most important technologies are electronic health records, machine learning, and assistive robots.

Trend 5—new information and communications media

- Emerging information and communications technologies are transforming the interactions between all stakeholders in the care process. They promote information transfer and facilitate the understanding of complex interventions through visual media. The most important technologies are avatars, mobile health, and IoT.

Trend 6—increased technological intensity

- The intensity of technology is increasing rapidly at all levels. The hospital is becoming a high-technology digital institution. Intensive data use leads to intensified collaboration with digital service providers and the extensive use of software or algorithms. The most important technologies are smart devices, robots, and cloud services.

Trend 7—outcome improvement and personalized treatments

- The quality of outcomes is rising in the context of increasing data use and the associated optimization of treatment procedures focusing on predictive forecasting models for the early detection of diseases. The perspective of personalized diagnostics and therapy is at the heart of the treatment process. The most important technologies are data analytics, machine learning, and predictive medicine.

Trend 8—emergence of new ecosystems

- Digital networking is leading to new providers and new business areas so that health care is evolving into an ecosystem in which established service providers have to maintain their competitive position and find their role in providing care. Platforms for exchanging data and mediating health care services play an important role in these ecosystems. The most important technologies are mobile health, machine learning, and predictive analytics.

Table 1. Identified trends in the scoping review.

Trend number	Core trend	Description	Key technologies	Fields of application
1	• Changing the patient role	• Patients have an active role in the entire care process. They participate in decision-making processes and have comprehensive transparency in their own data and available services.	• Internet, smart devices, and mobile health	• Decision support [78] • Increase in patient compliance [63,71] • Integration into the therapy and monitoring process [68]
2	• Connected and integrated delivery of care	• Service providers collaborate without sector barriers. They are connected to each other and provide all necessary diagnostic and therapeutic information to one another.	• EHRs[a], smart devices, IoT[b], and e-prescriptions	• Integrated data exchange between providers [77] • Full use of the EHR [46] • Transmission of upstream and downstream treatment information [51,66]
3	• Data-driven resource allocation	• The capacities of service providers are optimally used. They analyze real-time data and make decisions based on currently available resources beyond their own institution.	• Big data, machine learning, data mining, EHR, HISs[c], and predictive analytics	• Process optimization [42,48] • Planning and control of resources by data-driven decision models [44,48,52,55] • Predictive forecasting for capacity planning and the identification of bottlenecks and potential risks [48,52,69,80]
4	• Performance optimization in primary processes	• Hospitals manage their internal processes efficiently and in a patient-centric manner. They use technology to support staff in diagnostics and therapy decisions.	• EHRs, robotics, IoT, machine learning, and data platforms	• Development of personalized treatment pathways [42,44] • Optimization of diagnostics and outcomes across the care continuum [50,58,76,79] • Robotics in several clinical areas [38,45,54]
5	• New information and communications media	• Service providers are broadening the range of information and communications channels. They create shared and low-threshold ways to convey information and meet patients' communication needs.	• Avatars, virtual reality, augmented reality, EHRs, mobile health, and IoT	• Transfer media for data exchange between providers and patients [41,77] • Enhancing patient engagement [68,78] • Creation of new barrier-free communication channels [63]
6	• Increased technological intensity	• Hospitals are deploying an increasing number of technological solutions. They are exploiting the potential of innovative technical applications and raising the level of technological complexity.	• Smart devices, IoT, robotics, machine learning, and cloud services	• Expansion of the inventory of hardware and technical equipment [38,54] • Building partnerships with digital service providers [39,77] • Connection to external systems for permanent data exchange in real time [48,57,77]

Trend number	Core trend	Description	Key technologies	Fields of application
7	• Outcome improvement and personalized treatments	• The outcomes of patient care are steadily improving. Early detection and disease prediction systems enable an optimal treatment sequence according to the personalized needs of the patient.	• Big data, machine learning, predictive analytics, personalized medicine, and 3D printing	• Evaluation of incoming monitoring data from patients [20,50,62] • Connection to databases and data platforms for the training of AI[d] [58,61,79] • Individual therapies [53,65,76]
8	• Emergence of new ecosystems	• Service providers are linked together in an ecosystem to provide care. They use platforms to exchange data and optimally allocate services to the right place at the right time.	• Digital platforms, mobile health, machine learning, and predictive analytics	• Appointment management [66] • Reduction in waiting times [46,79] • Assignment of appropriate experts [65,69] • Distribution of available materials [69,73] • Switching on of suitable gatekeepers [56,70]

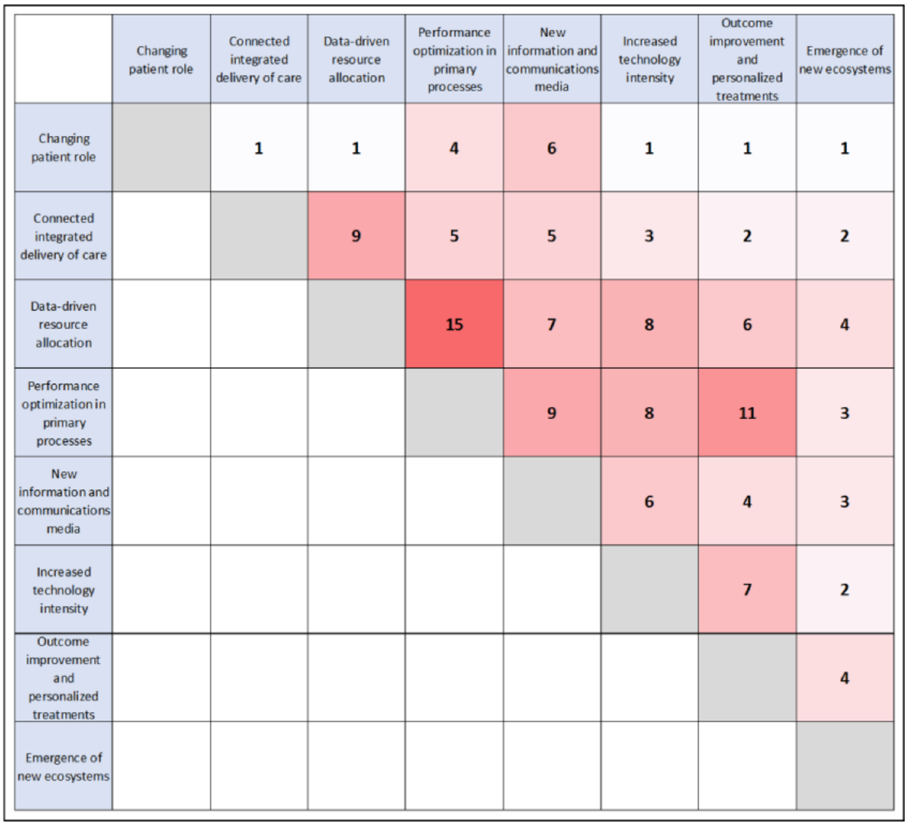
[a] EHR: electronic health record.

[b] IoT: Internet of Things.

[c] HIS: health information system.

[d] AI: artificial intelligence.

Figure 4. Connections between the trends.



	Changing patient role	Connected integrated delivery of care	Data-driven resource allocation	Performance optimization in primary processes	New information and communications media	Increased technology intensity	Outcome improvement and personalized treatments	Emergence of new ecosystems
Changing patient role		1	1	4	6	1	1	1
Connected integrated delivery of care			9	5	5	3	2	2
Data-driven resource allocation				15	7	8	6	4
Performance optimization in primary processes					9	8	11	3
New information and communications media						6	4	3
Increased technology intensity							7	2
Outcome improvement and personalized treatments								4
Emergence of new ecosystems								

Results of the Expert Interviews

The qualitative approach to analyzing the impacts of trends brought to light additional aspects aside from the findings from the literature. On the basis of the core trends shown, Table 2 presents the positive and negative trend effects, conditions for trend realization, and trend drivers. This made it possible to describe additional characteristics for each individual trend, enabling a holistic approach to the impact mechanisms of the trends resulting from the digitization of inpatient care.

The changed patient role leads to an increase in transparency, more patient autonomy, better opportunities for participation in the treatment process, and an expansion of self-competence. However, in the case of an overload of information, this can also lead to excessive demands or the communication of incorrect information. Therefore, the decisive conditions are to strengthen trust in acting institutions and make information available through public authorities. The driving factors are good health literacy and a high level of digital competence. Connected and integrated cooperation between providers leads to better access to health care services and faster data exchange between them. However, this may lead to increased competitive pressure, resulting in the exit of market players. The key conditions for this trend are incentives to cooperate so that individual providers are not disadvantaged. To achieve this, compensation systems must be adapted so that network-integrated collaboration is worthwhile and the necessary interfaces for data exchange are available. Driving factors include easy market access and existing supply gaps. Data-driven resource allocation leads to increased efficiency, a better use of capacity, and offsetting of skill shortages. However, there are strong dependencies among service providers regarding the extent to which they are willing to collaborate and share information. The intensive exchange of data at various levels between institutions can lead to data protection and data security risks. The governing conditions include the availability of data, interfaces, and technical expertise. The driving factors include scarce capacities, high data density, and public data platforms.

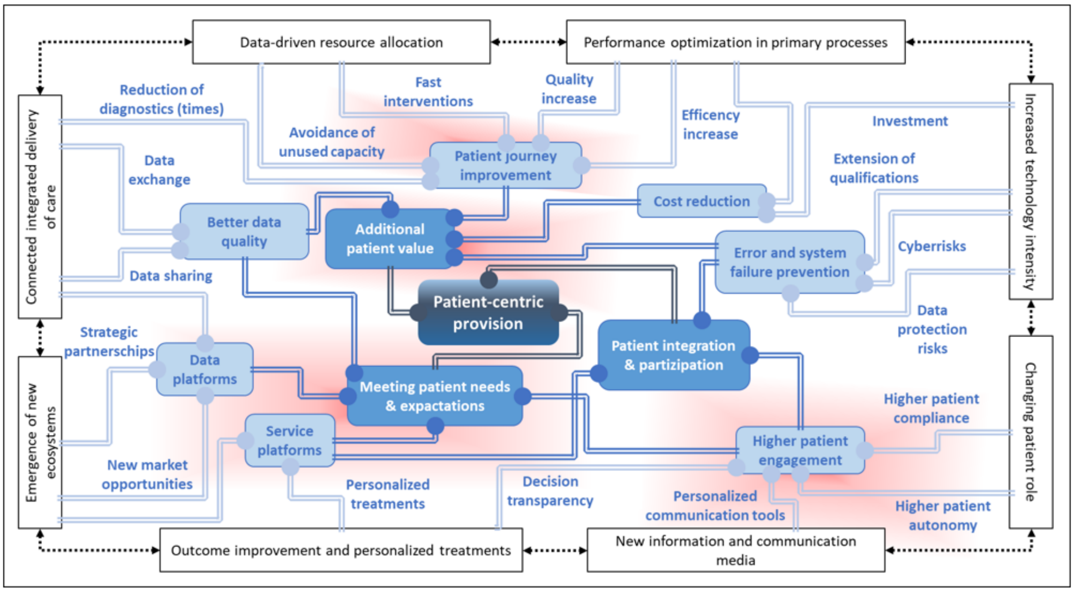
The performance optimization of core clinical processes leads to higher productivity, increases in quality, and higher patient and employee satisfaction. In contrast, there is a high internal dependency on a wide range of factors and the need for a continuous process of adaptation and optimization. The decisive conditions are investments in new systems and the associated qualifications of the workforce. Acceptance and openness are essential for success. The driving factors are medical progress, falling costs of innovations, and new job profiles. New information and communications technologies are accompanied by a change in patients' need to interact with providers. They lower communication barriers and lead to better information delivery. For clinics, the wide range of media may result in higher costs for their operation, and additional qualifications may be necessary. The key conditions are an understanding of communication needs and an acceptance of media use by all stakeholders. The driving factors are the digital competencies of all stakeholders and the current general societal trends in media preferences. Increasing technology intensity ensures higher productivity and compensates for a lack of skilled workers. This is accompanied by greater complexity and the associated requirements for cybersecurity and employee qualifications. Comprehensive long-term investments and the acquisition of new knowledge are required as decisive conditions. The driving factors are technological progress and rising patient expectations. The application of personalized medicine leads to an improvement in outcomes as well as the possibility of specialization and the opening up of new fields of treatment. In contrast, there is an adjustment of care-oriented business models, which, among other things, also include prevention to a greater extent. The key conditions are high data quality and availability, the market maturity of innovative technologies, and confidence in new treatment procedures. The driving factors are medical and technical progress as well as the service offerings of competitors.

The emergence of new ecosystems has led to new customer-centric health care services, an exploitation of the potential of the platform economy, and the associated opportunities to increase customer value. This creates new competitive situations that can put pressure on providers and lead to inappropriate power relationships among individual market participants. The decisive conditions are an understanding of digital ecosystems also on the part of the legislator, transparency, and fair conditions on platforms for all service providers. The driving factors are low barriers to market entry, little regulation, and a strong focus on customer needs. Figure 5 shows the trend correlations in a heuristic impact model.

Table 2. Trend characteristics from expert interviews in Germany.

Trend number	Core trend	Positive effects	Negative effects	Prerequisites	Trend drivers
1	• Changing the patient role	• Increasing transparency • Growing patient autonomy • Better participation opportunities • Growth of self-competence	• Risk of disinformation • Overwhelming media diversity	• Trust in institutions • Public provision of information • Comprehensive education	• Health literacy • Digital competencies • Information access
2	• Connected and integrated delivery of care	• Better access to services • Faster data exchange • Reduction in interface problems	• Additional competitive pressure between providers	• Performance incentives • Adjustment of reimbursement systems • Data exchange interfaces	• Easy market access • Existing supply gaps • Integrated planning tools
3	• Data-driven resource allocation	• Increasing efficiency • Optimal use of resources • Compensation for the shortages of skilled workers	• Dependence on service providers • Data protection risks • Increasing complexity	• Data availability • Interface availability • Specialist availability	• Scarce resources • High data density • Public data platforms
4	• Performance optimization in primary processes	• Efficient provisioning • Increasing quality • Higher satisfaction (patients and staff)	• Dependencies of many factors • Constant change process	• Investments in new systems • Acceptance of new processes • Permanent qualification	• Medical progress • Decreasing costs • New professions
5	• New information and communications media	• Meeting the needs of patients • Reduction in communication barriers • Improving the availability of information • Higher satisfaction	• High expenditure for the maintenance and provision of the media • Additional need for qualification	• Acceptance of the parties involved • Knowing and understanding needs • Quality assurance	• Digital skills (patients and staff) • Demand for contemporary media from everyday social life
6	• Increased technological intensity	• Achieving advances in productivity • Compensation for the lack of skilled workers • Relief of the workforce	• High effort for cybersecurity • Increasing complexity • Need for technical knowledge	• Extensive investments • Comprehensive qualification • Acquisition of new areas of knowledge	• Technical progress • Patient expectations
7	• Outcome improvement and personalized treatments	• Increasing outcomes • Development of new service areas • Opportunities for specialization	• Adjustment of the business model	• Data quality and availability • Market maturity of technologies • Confidence in new treatment procedures	• Medical progress • Technical progress • Competitors’ offers
8	• Emergence of new ecosystems	• Creation of customer-centric proposals • Exploiting the potential of the platform economy • Increasing customer value	• Highly competitive pressure • Inappropriate market power of some players • Inadequate regulation	• Understanding digitally controlled ecosystems • Transparency of the service offering • Fair conditions on platforms	• Low barriers to market entry • Minimal regulation • High customer interest and demand

Figure 5. Heuristic impact model of digital transformation trends in hospitals.



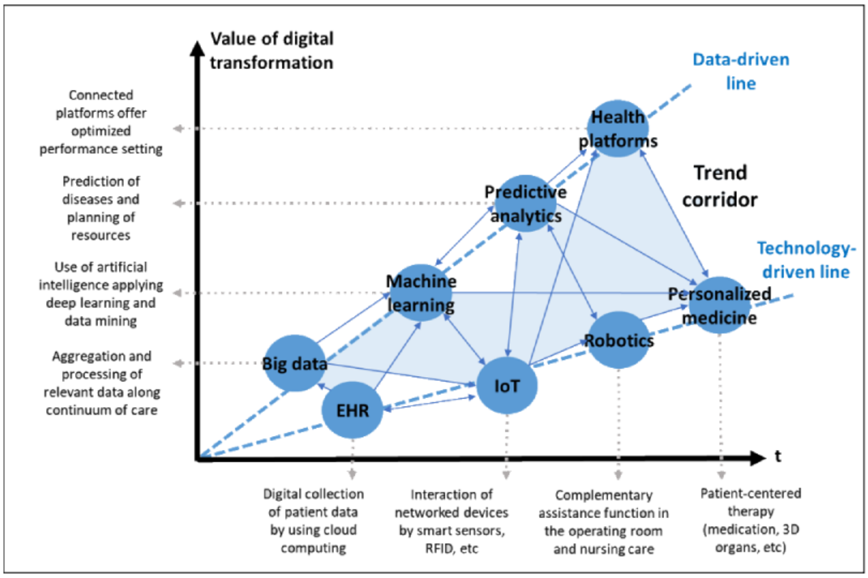
Results of the Data Synthesis

The interrelationships between the trends are very complex and require an adequate form of representation to be able to understand the trend effects. We show these relationships in a heuristic impact model in Figure 5. The 8 core trends are plotted on the outer level. They form the starting point of the interaction. At the next level inward, the first-degree effects are derived. The links there are underpinned with the help of the extracted factors from the literature and the expert interviews. At the next level, in medium blue, the second-degree effects are shown. The characteristics of additional patient value, meeting patient needs and expectations, and patient integration and participation emerge. At the last level, the third-degree effect is derived, which provides patient-centric care. This representation illustrates the positive foresight of future development predicted in the literature and in the expert interviews. The data synthesis shows that, as a result of the digital transformation, the value of care for patients will increase considerably. The nodes highlighted in red show a concentration of related parameters. They are based on the trend correlations shown in Figure 4 and the qualitative responses from the interviews.

For a more intuitive illustration of the trend development perspective, we placed the value of care as a result of digital transformation in a temporal context. Figure 6 shows the development horizon with increasing digital transformation progress and the associated added value. On the one hand, this is driven by technological intensification [38,39,41,45,54,59,71]. The primary manifestation of this is the use of digitally controlled and networked devices as well as software and algorithms. In contrast, a rise in the quality of care is to be expected [20,43,44,67,72,76]. The availability of large amounts of data enables better predictability of diseases and possible therapies and better control of the care itself. Machine learning leads to a continuous improvement in outcomes and decision-making [47,57,58,61,69].

The convergence of personalized medicine as a result of better diagnostics and individualized therapy decisions, incorporating the spectrum of technological innovations of today and the future, results in an adapted range of care offered by hospitals [56,58,61]. Thus, hospitals are increasingly becoming players on a health care platform on which various providers offer health care services depending on their competencies [53,65,73,77]. The schematic diagram illustrates the successive stages and draws 2 elementary trend lines along which the progress runs. Within this is the trend corridor. Along the time axis, both lines contribute to a higher value of care, with the contribution of technologies being lower than the value contribution of data. This is because technologies can provide additional benefits on their own. However, the higher overall benefit comes from the networked use of the data.

Figure 6. Schematic trend development perspective. EHR: electronic health record; IoT: Internet of Things; RFID: radio frequency identification.



Detected Trend Decelerators and Barriers

The analyses described numerous blockades and barriers that slow down the trends. Textbox 2 classifies these decelerators according to internal and external hurdles. External hurdles are defined as factors that are not under the hospitals' control and usually have to be removed by a higher-level institution. This primarily refers to state institutions or self-governance. These aspects include the regulatory framework for health care services [60,65] or data protection regulations [40,46]; access to and the interoperability of data [41,46,49,55]; and an insufficient investment in hospital infrastructure [68,70,74], especially in technical equipment.

Internal hurdles are factors that occur within the hospital setting and can largely be solved there. The most frequently cited obstacle is a lack of knowledge of the use of digital applications [41,50,68,74]. The changes resulting from digital transformation are also accompanied by a shift in tasks and processes, some of which are not widely accepted by the workforce [44,63,67,75,76]. Furthermore, ensuring cybersecurity to avoid risks is an elementary component that clinics must accommodate to an increasing extent [20,42]. Finally, there is often a lack of perspective on how hospitals can position themselves in a new digital ecosystem of service providers as part of the digital transformation and how they can develop new business models and exploit the potential of digital provision [56,65,79].

Textbox 2. Type of hurdle and description of the results.

External hurdles

- Regulation of service delivery [60,74]
- Data protection regulations [40,46]
- Uniformity of data structures [41,46,49,55,75]
- Lack of interfaces for data exchange [50,71,76]
- Insufficient investment in technical infrastructure [68,70,74]
- Lack of political understanding of appropriate frameworks [64,74]

Internal hurdles

- Inadequately qualified personnel [41,50,68]
- Low acceptance of substantially changing processes [44,63,67,75]
- Unwillingness to make large investments [38,55,63,68]
- Failure to ensure cybersecurity [20,42]
- Continuous training needs [49,55,59,71]
- Availability of suitable business models [56,65,79]

Discussion

Principal Findings

We included 44 studies in our analysis, drawing a picture of hospital trends. Subsequently, the trends in the German hospital care setting were discussed, validated, and refined through expert interviews. A wide range of aspects were considered, covering the use of new technologies, the implementation of process innovations, and behavior changes or role shifts among patients and hospital staff. The trends shown provide a framework for better assessing current and future developments from the perspective of clinics and other health care stakeholders. This allows for a snapshot of one's own point of view or an assessment by higher-level authorities as to which changes in the framework conditions can accelerate or slow down the trends. For this purpose, trend barriers and developmental hurdles were also extracted from the studies. These can be used to derive policy options.

The trends described were not examined in terms of their positive or negative impact. Nevertheless, a positive trend is discernible in the development horizon, which is improving inpatient care and steering it in a patient-centered direction. Therefore, it is conceivable that not all trends will lead to a positive development for all stakeholders involved. The increased use of technology, automation, and AI systems will in some cases entail job losses or a restructuring of work [91]. Therefore, for hospital staff (employee representatives or trade unions), change through digital transformation is not necessarily associated with improvements. These aspects must also be considered. In contrast, digitization leads to the extensive aggregation of data, which can restrict people's freedom of data self-determination. In addition, one-sided power positions of companies that have very large amounts of data at their disposal can lead to imbalances [92]. This can also preclude better patient care or unfair competition in the market [93]. The constant change in the course of the digital transformation of hospitals will in the future involve staff and decision makers in the qualification and further development of the range of services as the speed of change is increasing and the demands on service providers are continually rising [94]. The positive aspects of digital transformation should also be emphasized as they have the potential to prevent illnesses in the future, improve outcomes, and relieve staff of nonmedical tasks. In view of the impending shortage of skilled workers in health care professions around the hospital and beyond, this can unleash tied-up resources. Technical systems and algorithms can take over assisting tasks and, thus, optimize internal clinical processes as well as improve the entire spectrum of care from prevention to individualized therapy [95]. We are already seeing how technical innovations and digital applications are changing the professions in hospitals [96-98]. With the help of these tools, the high future demand for skilled workers can be partially compensated or substituted. In addition, dealing with innovations can lead to a higher attractiveness of the professions. The included studies tended to show a positive overall picture in which the opportunities dominate over the risks. The trend development was partly based on specific findings from Germany. However, overall, the trend perspectives can also be applied to other countries. In Europe, there are countries such as Switzerland, Austria, and the Netherlands, among others, that have similar health care system conditions. The analogs of the perspectives of digitally controlled hospital care can also be transferred to almost all other countries.

Implications for Public Health and Future Medicine

Our study shows that there is enormous potential for public health to make care in hospitals more patient-centric as a result of digital transformation. This means that diseases can be detected and treated earlier. At the end of a successful transformation, there is a healthier population that takes advantage of the opportunities offered by preventative and precautionary health care. Clinics contribute to realizing this goal and must be prepared by decision makers for their changed role. As a result, comprehensive savings are possible, which can relieve the pressure on the health care system [99]. Strained health care systems such as those in Germany and other European countries could be relieved as a result. The shortage of specialists in inpatient care can be slowed down. For medical professions, this transformation means, on the one hand, relief through new technologies or support in diagnostics and therapy. Therefore, more time can be spent on advising and caring for patients, and the quality of care can be improved. In contrast, increasing digitization requires the learning of new skills and a change in working methods from physical to technical or technology-supporting activities. In the continuing education of physicians, it is essential that these competencies are sharpened. The increased interaction of patients, who, in the future, will collect their own data, monitor the progress of their illnesses with data support, and use mobile devices to contact hospitals, requires a high level of health literacy and digital competencies. School education should prepare children for this and also offer programs for older people to acquire the necessary knowledge. A low-threshold access to information and guidance should also be established.

Limitations

The included studies already cover a comprehensive range of current and future trends. Nevertheless, the individual aspects may not have been considered. The analysis was primarily conducted from the perspective of inpatient service provision. The patient, as an important participant in the care process, was not explicitly examined. For example, sociological trends with regard to the use of digital applications could lead to findings that were not reflected [100]. These include the acceptance of new technologies and the willingness to use them. This aspect can also be extended to personnel in clinics. The distinction between trends, future forecasts, and expectations is not entirely clear-cut. Trends comprise medium-term developments in a relatively delimited subject area. We tried to apply this reference position as best as possible in our study. Some developments also refer to a longer time horizon as the initial situation varies greatly among countries. Thus, this study only showed the general developments that are possible in the next 5 to 10 years. The extent to which these developments can be realized or which other influencing factors accelerate or slow down individual aspects must be assessed individually for each health care system [49]. Therefore, a generalization to all health care systems is not conceivable. The German perspective could lead to minor distortions here in some cases. However, Germany's reference role in many respects supports our argument. The reference period of the past 5 years covers a relatively short time horizon. We decided to carry out our study in this way as, in our view, studies conducted before 2016 with data from previous years are not informative about future trends [36,37]. The screening of databases also revealed that most studies on trends in the digitization of health care were published in the past 3 years (2018 to 2020) in particular. Owing to the COVID-19 pandemic, there may have been additional trends in 2020 that we did not explicitly include in our analysis [101]. As little, or at best rudimentary, research was available in the fall of 2020 when we conducted our database search, we excluded this option from our analysis. Nonetheless, we cannot rule out that new trends for clinics may have emerged as a result of the pandemic that have not yet been adequately considered in this work.

Conclusions

In our study, we showed the influence of digital transformation on inpatient care. In the process, the high complexity of the interacting players became clear. We extracted 8 core trends from the international literature and validated them with experts in relation to the change in hospital care provision in Germany. We presented the interrelationships in a heuristic impact model. Our study highlights the barriers to successful transformation based on internal and external factors. We provide recommendations for action for decision makers on how public health can be improved by a digitally transformed hospital landscape and which hurdles need to be overcome.

Acknowledgments

This research was supported as part of the ATLAS project "Innovation and digital transformation in healthcare" by the state of North Rhine-Westphalia, Germany (grant ITG-1-1).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms.
[PDF File (Adobe PDF File), 383 KB - publichealth_v9i1e40622_app1.pdf]

Multimedia Appendix 2
Inclusion and exclusion criteria.
[PDF File (Adobe PDF File), 518 KB - publichealth_v9i1e40622_app2.pdf]

Multimedia Appendix 3
Data extraction form.
[PDF File (Adobe PDF File), 388 KB - publichealth_v9i1e40622_app3.pdf]

Multimedia Appendix 4
Literature overview.
[PDF File (Adobe PDF File), 420 KB - publichealth_v9i1e40622_app4.pdf]

Multimedia Appendix 5
Documentation of the scoping review.
[PDF File (Adobe PDF File), 410 KB - publichealth_v9i1e40622_app5.pdf]

References

1. Van Veldhoven Z, Vanthienen J. Digital transformation as an interaction-driven perspective between business, society, and technology. Electron Mark 2022;32(2):629-644 [FREE Full text] [doi: 10.1007/s12525-021-00464-5] [Medline: 35602117]
2. Hilbert M. Digital technology and social change: the digital transformation of society from a historical perspective. Dialogues Clin Neurosci 2020 Jun;22(2):189-194 [FREE Full text] [doi: 10.31887/DCNS.2020.22.2/mhilbert] [Medline: 32699519]
3. Zhong RY, Newman ST, Huang GQ, Lan S. Big data for supply chain management in the service and manufacturing sectors: challenges, opportunities, and future perspectives. Comput Ind Eng 2016 Nov;101:572-591 [FREE Full text] [doi: 10.1016/j.cie.2016.07.013]
4. Ford G, Compton M, Millett G, Tzortzis A. The role of digital disruption in healthcare service innovation. In: Pfannstiel MA, Rasche C, editors. Service Business Model Innovation in Healthcare and Hospital Management: Models, Strategies, Tools. Cham, Switzerland: Springer; 2017:57-70.

5. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
6. Benis A, Tamburis O, Chronaki C, Moen A. One digital health: a unified framework for future health ecosystems. J Med Internet Res 2021 Mar 05;23(2):e22189 [FREE Full text] [doi: 10.2196/22189] [Medline: 33492240]
7. Herrmann M, Boehme P, Mondritzki T, Ehlers JP, Kavadias S, Truebel H. Digital transformation and disruption of the health care sector: internet-based observational study. J Med Internet Res 2018 Mar 27;20(3):e104 [FREE Full text] [doi: 10.2196/jmir.9498] [Medline: 29588274]
8. Gabutti I, Mascia D, Cicchetti A. Exploring "patient-centered" hospitals: a systematic review to understand change. BMC Health Serv Res 2017 May 22;17(1):364 [FREE Full text] [doi: 10.1186/s12913-017-2306-0] [Medline: 28532463]
9. Woiceshyn J, Blades K, Pendharkar SR. Integrated versus fragmented implementation of complex innovations in acute health care. Health Care Manage Rev 2017 Jan;42(1):76-86 [FREE Full text] [doi: 10.1097/HMR.0000000000000092] [Medline: 26469705]
10. Ho JY, Hendi AS. Recent trends in life expectancy across high income countries: retrospective observational study. BMJ 2018 Aug 15;362:k2562 [FREE Full text] [doi: 10.1136/bmj.k2562] [Medline: 30111634]
11. Ma L, Luo N, Wan T, Hu C, Peng M. An improved healthcare accessibility measure considering the temporal dimension and population demand of different ages. Int J Environ Res Public Health 2018 Oct 31;15(11):2421 [FREE Full text] [doi: 10.3390/ijerph15112421] [Medline: 30384482]
12. Rowley WR. Healthcare 2025: alternative views and underlying values. Foresight 2003 Oct;5(5):16-21 [FREE Full text] [doi: 10.1108/14636680310507271]
13. Saritas O, Keenan M. Broken promises and/or techno dreams? The future of health and social services in Europe. Foresight 2004;6(5):281-291 [FREE Full text] [doi: 10.1108/14636680410562981]
14. Rees GH, Crampton P, Gauld R, MacDonell S. Rethinking health workforce planning: capturing health system social and power interactions through actor analysis. Futures 2018 May;99:16-27 [FREE Full text] [doi: 10.1016/j.futures.2018.03.009]
15. Grah B, Dimovski V, Colnar S, Bogataj D. Modelling the nurses employment dynamics in the ageing society. IFAC-PapersOnLine 2019;52(25):219-224 [FREE Full text] [doi: 10.1016/j.ifacol.2019.12.476]
16. Cucciniello M, Lapsley I, Nasi G, Pagliari C. Understanding key factors affecting electronic medical record implementation: a sociotechnical approach. BMC Health Serv Res 2015 Jul 17;15:268 [FREE Full text] [doi: 10.1186/s12913-015-0928-7] [Medline: 26184405]
17. Ashmarina S, Mesquita A, Vochozka M. Digital Transformation of the Economy: Challenges, Trends and New Opportunities. Cham, Switzerland: Springer; 2020.
18. Abdallah YO, Shehab E, Al-Ashaab A. Understanding digital transformation in the manufacturing industry: a systematic literature review and future trends. Prod Manag Dev 2021;19(1):e20200021 [FREE Full text] [doi: 10.4322/pmd.2021.001]
19. Pihir I, Tomičić-Pupek K, Tomičić Furjan M. Digital transformation playground - literature review and framework of concepts. J Inf Organ Sci 2019 Jun 21;43(1):33-48 [FREE Full text] [doi: 10.31341/jios.43.1.3]
20. Kelly JT, Campbell KL, Gong E, Scuffham P. The internet of things: impact and implications for health care delivery. J Med Internet Res 2020 Nov 10;22(11):e20135 [FREE Full text] [doi: 10.2196/20135] [Medline: 33170132]
21. Far SB, Rad AI. Applying digital twins in metaverse: user interface, security and privacy challenges. J Metaverse 2022 Mar 31;2(1):8-15 [FREE Full text]
22. Wepner B, Giesecke S. Drivers, trends and scenarios for the future of health in Europe. Impressions from the FRESHER project. Eur J Futures Res 2018 Jan 10;6(1):2 [FREE Full text] [doi: 10.1007/s40309-017-0118-4]
23. Freyn SL, Farley F. Competitive intelligence: a prescription for US health-care? Foresight 2020 Jul 16;22(5/6):617-632 [FREE Full text] [doi: 10.1108/fs-02-2020-0011]
24. Fatoum H, Hanna S, Halamka JD, Sicker DC, Spangenberg P, Hashmi SK. Blockchain integration with digital technology and the future of health care ecosystems: systematic review. J Med Internet Res 2021 Nov 02;23(11):e19846 [FREE Full text] [doi: 10.2196/19846] [Medline: 34726603]
25. Lauwers L, Bastiaens H, Remmen R, Keune H. The integration of interlinkages between nature and human health in primary health care: protocol for a scoping review. JMIR Res Protoc 2019 Jan 18;8(1):e12510 [FREE Full text] [doi: 10.2196/12510] [Medline: 30664484]
26. Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Beagley J, Belesova K, et al. The 2020 report of the Lancet countdown on health and climate change: responding to converging crises. Lancet 2021 Jan 09;397(10269):129-170. [doi: 10.1016/S0140-6736(20)32290-X] [Medline: 33278353]
27. Rahimi-Ardabili H, Magrabi F, Coiera E. Digital health for climate change mitigation and response: a scoping review. J Am Med Inform Assoc 2022 Nov 14;29(12):2140-2152. [doi: 10.1093/jamia/ocac134] [Medline: 35960171]
28. De Santis KK, Jahnel T, Sina E, Wienert J, Zeeb H. Digitization and health in Germany: cross-sectional nationwide survey. JMIR Public Health Surveill 2021 Nov 22;7(11):e32951 [FREE Full text] [doi: 10.2196/32951] [Medline: 34813493]
29. Duncan R, Eden R, Woods L, Wong I, Sullivan C. Synthesizing dimensions of digital maturity in hospitals: systematic review. J Med Internet Res 2022 Mar 30;24(3):e32994 [FREE Full text] [doi: 10.2196/32994] [Medline: 35353050]

30. Tortorella GL, Saurin TA, Fogliatto FS, Rosa VM, Tonetto LM, Magrabi F. Impacts of Healthcare 4.0 digital technologies on the resilience of hospitals. Technol Forecast Soc Change 2021 May;166:120666 [FREE Full text] [doi: 10.1016/j.techfore.2021.120666]
31. Krasuska M, Williams R, Sheikh A, Franklin BD, Heeney C, Lane W, et al. Technological capabilities to assess digital excellence in hospitals in high performing health care systems: international eDelphi exercise. J Med Internet Res 2020 Aug 18;22(8):e17022 [FREE Full text] [doi: 10.2196/17022] [Medline: 32808938]
32. Williams PA, Lovelock B, Cabarrus T, Harvey M. Improving digital hospital transformation: development of an outcomes-based infrastructure maturity assessment framework. JMIR Med Inform 2019 Jan 11;7(1):e12465 [FREE Full text] [doi: 10.2196/12465] [Medline: 30632973]
33. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J Clin Epidemiol 2021 Jun;134:103-112. [doi: 10.1016/j.jclinepi.2021.02.003] [Medline: 33577987]
34. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
35. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
36. Xie Y, Zhang J, Wang H, Liu P, Liu S, Huo T, et al. Applications of blockchain in the medical field: narrative review. J Med Internet Res 2021 Oct 28;23(10):e28613 [FREE Full text] [doi: 10.2196/28613] [Medline: 34533470]
37. Pieper D, Antoine SL, Neugebauer EA, Eikermann M. Up-to-dateness of reviews is often neglected in overviews: a systematic review. J Clin Epidemiol 2014 Dec;67(12):1302-1308 [FREE Full text] [doi: 10.1016/j.jclinepi.2014.08.008] [Medline: 25281222]
38. Abitbol J, Munir A, How J, Lau S, Salvador S, Kogan L, et al. The shifting trends towards a robotically-assisted surgical interface: clinical and financial implications. Health Policy Technol 2020 Jun;9(2):157-165 [FREE Full text] [doi: 10.1016/j.hlpt.2020.03.003]
39. Aceto G, Persico V, Pescapé A. Industry 4.0 and health: internet of things, big data, and cloud computing for Healthcare 4.0. J Ind Inf Integr 2020 Jun;18:100129 [FREE Full text] [doi: 10.1016/j.jii.2020.100129]
40. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. J Am Med Inform Assoc 2017 Nov 01;24(6):1142-1148 [FREE Full text] [doi: 10.1093/jamia/ocx080] [Medline: 29016973]
41. Aghdam ZN, Rahmani AM, Hosseinzadeh M. The role of the internet of things in healthcare: future trends and challenges. Comput Methods Programs Biomed 2021 Feb;199:105903. [doi: 10.1016/j.cmpb.2020.105903] [Medline: 33348073]
42. Baslyman M, Almoaber B, Amyot D, Bouattane EM. Using goals and indicators for activity-based process integration in healthcare. Procedia Comput Sci 2017;113:318-325 [FREE Full text] [doi: 10.1016/j.procs.2017.08.340]
43. Bergey MR, Goldsack JC, Robinson EJ. Invisible work and changing roles: health information technology implementation and reorganization of work practices for the inpatient nursing team. Soc Sci Med 2019 Aug;235:112387. [doi: 10.1016/j.socscimed.2019.112387] [Medline: 31272078]
44. Berntsen G, Strisland F, Malm-Nicolaisen K, Smaradottir B, Fensli R, Røhne M. The evidence base for an ideal care pathway for frail multimorbid elderly: combined scoping and systematic intervention review. J Med Internet Res 2019 Apr 22;21(4):e12517 [FREE Full text] [doi: 10.2196/12517] [Medline: 31008706]
45. Bhandari M, Zeffiro T, Reddiboina M. Artificial intelligence and robotic surgery: current perspective and future directions. Curr Opin Urol 2020 Jan;30(1):48-54. [doi: 10.1097/MOU.0000000000000692] [Medline: 31724999]
46. Blaser J. Challenges of digital medicine. Praxis (Bern 1994) 2018 Jun;107(13):712-716. [doi: 10.1024/1661-8157/a003008] [Medline: 29921177]
47. Blease C, Kharko A, Locher C, DesRoches CM, Mandl KD. US primary care in 2029: a Delphi survey on the impact of machine learning. PLoS One 2020 Oct 08;15(10):e0239947 [FREE Full text] [doi: 10.1371/journal.pone.0239947] [Medline: 33031411]
48. Brink JA, Arenson RL, Grist TM, Lewin JS, Enzmann D. Bits and bytes: the future of radiology lies in informatics and information technology. Eur Radiol 2017 Sep;27(9):3647-3651. [doi: 10.1007/s00330-016-4688-5] [Medline: 28280932]
49. Bukowski M, Farkas R, Beyan O, Moll L, Hahn H, Kiessling F, et al. Implementation of eHealth and AI integrated diagnostics with multidisciplinary digitized data: are we ready from an international perspective? Eur Radiol 2020 Oct;30(10):5510-5524 [FREE Full text] [doi: 10.1007/s00330-020-06874-x] [Medline: 32377810]
50. Burkoski V, Yoon J, Hutchinson D, Solomon S, Collins BE. Experiences of nurses working in a fully digital hospital: a phenomenological study. Nurs Leadersh (Tor Ont) 2019 May;32(SP):72-85. [doi: 10.12927/cjnl.2019.25813] [Medline: 31099748]
51. Cano I, Tenyi A, Vela E, Miralles F, Roca J. Perspectives on Big Data applications of health information. Curr Opin Syst Biol 2017 Jun;3:36-42 [FREE Full text] [doi: 10.1016/j.coisb.2017.04.012]

52. Chen PT. Medical big data applications: intertwined effects and effective resource allocation strategies identified through IRA-NRM analysis. Technol Forecast Soc Change 2018 May;130:150-164 [FREE Full text] [doi: 10.1016/j.techfore.2018.01.033]
53. Chong NK, Chu Shan Elaine C, de Korne DF. Creating a learning televillage and automated digital child health ecosystem. Pediatr Clin North Am 2020 Aug;67(4):707-724. [doi: 10.1016/j.pcl.2020.04.016] [Medline: 32650868]
54. Daneshmand M, Bilici O, Bolotnikova A, Anbarjafari G. Medical robots with potential applications in participatory and opportunistic remote sensing: a review. Rob Auton Syst 2017 Sep;95:160-180 [FREE Full text] [doi: 10.1016/j.robot.2017.06.009]
55. Deiters W, Burmann A, Meister S. Strategies for digitalizing the hospital of the future. Urologe A 2018 Sep;57(9):1031-1039. [doi: 10.1007/s00120-018-0731-2] [Medline: 30088025]
56. Denicolai S, Previtali P. Precision medicine: implications for value chains and business models in life sciences. Technol Forecast Soc Change 2020 Feb;151:119767 [FREE Full text] [doi: 10.1016/j.techfore.2019.119767]
57. Desai AN, Kraemer MU, Bhatia S, Cori A, Nouvellet P, Herringer M, et al. Real-time epidemic forecasting: challenges and opportunities. Health Secur 2019 Jul;17(4):268-275 [FREE Full text] [doi: 10.1089/hs.2019.0022] [Medline: 31433279]
58. Dreyer KJ, Geis JR. When machines think: radiology's next frontier. Radiology 2017 Dec;285(3):713-718. [doi: 10.1148/radiol.2017171183] [Medline: 29155639]
59. Eckert M, Volmerg JS, Friedrich CM. Augmented reality in medicine: systematic and bibliographic review. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e10967 [FREE Full text] [doi: 10.2196/10967] [Medline: 31025950]
60. Eleftheriou I, Embury SM, Moden R, Dobinson P, Brass A. Data journeys: identifying social and technical barriers to data movement in large, complex organisations. J Biomed Inform 2018 Mar;78:102-122 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.001] [Medline: 29223464]
61. Fang CH, Chen K, Zhang P. The current situation and prospects of the application of intelligent diagnosis and treatment technology in the general surgery. Zhonghua Wai Ke Za Zhi 2019 Jan 01;57(1):1-5. [doi: 10.3760/cma.j.issn.0529-5815.2019.01.001] [Medline: 30612385]
62. Farahani B, Barzegari M, Shams Aliee F, Shaik KA. Towards collaborative intelligent IoT eHealth: from device to fog, and cloud. Microprocess Microsyst 2020 Feb;72:102938 [FREE Full text] [doi: 10.1016/j.micpro.2019.102938]
63. Fuller TE, Pong DD, Piniella N, Pardo M, Bessa N, Yoon C, et al. Interactive digital health tools to engage patients and caregivers in discharge preparation: implementation study. J Med Internet Res 2020 Apr 28;22(4):e15573 [FREE Full text] [doi: 10.2196/15573] [Medline: 32343248]
64. Galetsi P, Katsaliaki K, Kumar S. Values, challenges and future directions of big data analytics in healthcare: a systematic review. Soc Sci Med 2019 Nov;241:112533. [doi: 10.1016/j.socscimed.2019.112533] [Medline: 31585681]
65. Gryson O. Who will kick off the digital transformation of health? Med Sci (Paris) 2018 Jun;34(6-7):587-589 [FREE Full text] [doi: 10.1051/medsci/20183406019] [Medline: 30067202]
66. Guntuku SC, Schwartz HA, Kashyap A, Gaulton JS, Stokes DC, Asch DA, et al. Variability in language used on social media prior to hospital visits. Sci Rep 2020 Mar 12;10(1):4346 [FREE Full text] [doi: 10.1038/s41598-020-60750-8] [Medline: 32165648]
67. Ker JI, Wang Y, Hajli N. Examining the impact of health information systems on healthcare service improvement: the case of reducing in patient-flow delays in a U.S. hospital. Technol Forecast Soc Change 2018 Feb;127:188-198 [FREE Full text] [doi: 10.1016/j.techfore.2017.07.013]
68. Lee D. Strategies for technology-driven service encounters for patient experience satisfaction in hospitals. Technol Forecast Soc Change 2018 Dec;137:118-127 [FREE Full text] [doi: 10.1016/j.techfore.2018.06.050]
69. Mitterecker A, Hofmann A, Trentino KM, Lloyd A, Leahy MF, Schwarzbauer K, et al. Machine learning-based prediction of transfusion. Transfusion 2020 Sep;60(9):1977-1986 [FREE Full text] [doi: 10.1111/trf.15935] [Medline: 32596877]
70. Neumann K, Reichl V, Rong O. Urgent need of action for the future of digital hospitals. HNO 2019 May;67(5):350-355. [doi: 10.1007/s00106-019-0655-1] [Medline: 30963220]
71. Palanica A, Docktor MJ, Lee A, Fossat Y. Using mobile virtual reality to enhance medical comprehension and satisfaction in patients and their families. Perspect Med Educ 2019 Apr;8(2):123-127 [FREE Full text] [doi: 10.1007/s40037-019-0504-7] [Medline: 30912006]
72. Perez I, Brown M, Pinchin J, Martindale S, Sharples S, Shaw D, et al. Out of hours workload management: Bayesian inference for decision support in secondary care. Artif Intell Med 2016 Oct;73:34-44. [doi: 10.1016/j.artmed.2016.09.005] [Medline: 27926380]
73. Reda M, Kanga D, Fatima T, Azouazi M. Blockchain in health supply chain management: state of art challenges and opportunities. Procedia Comput Sci 2020;175:706-709 [FREE Full text] [doi: 10.1016/j.procs.2020.07.104]
74. Ricciardi W, Pita Barros P, Bourek A, Brouwer W, Kelsey T, Lehtonen L, Expert Panel on Effective Ways of Investing in Health (EXPH). How to govern the digital transformation of health services. Eur J Public Health 2019 Oct 01;29(Supplement_3):7-12 [FREE Full text] [doi: 10.1093/eurpub/ckz165] [Medline: 31738442]
75. Serbanati LD. Health digital state and smart EHR systems. Inform Med Unlocked 2020;21:100494 [FREE Full text] [doi: 10.1016/j.imu.2020.100494]

76. Staib A, Sullivan C, Jones M, Griffin B, Bell A, Scott I. The ED-inpatient dashboard: uniting emergency and inpatient clinicians to improve the efficiency and quality of care for patients requiring emergency admission to hospital. Emerg Med Australas 2017 Jun;29(3):363-366. [doi: 10.1111/1742-6723.12661] [Medline: 27592365]
77. Tian S, Yang W, Le Grange JM, Wang P, Huang W, Ye Z. Smart healthcare: making medical care more intelligent. Glob Health J 2019 Sep;3(3):62-65 [FREE Full text] [doi: 10.1016/j.glohj.2019.07.001]
78. Vallo Hult H, Hansson A, Svensson L, Gellerstedt M. Flipped healthcare for better or worse. Health Informatics J 2019 Sep;25(3):587-597 [FREE Full text] [doi: 10.1177/1460458219833099] [Medline: 30887867]
79. Wang Y, Hajli N. Exploring the path to big data analytics success in healthcare. J Bus Res 2017 Jan;70:287-299 [FREE Full text] [doi: 10.1016/j.jbusres.2016.08.002]
80. Zolbanin HM, Davazdahemami B, Delen D, Zadeh AH. Data analytics for the sustainable use of resources in hospitals: predicting the length of stay for patients with chronic diseases. Inf Manag 2022 Jul;59(5):103282 [FREE Full text] [doi: 10.1016/j.im.2020.103282]
81. Safi Z, Abd-Alrazaq A, Khalifa M, Househ M. Technical aspects of developing chatbots for medical applications: scoping review. J Med Internet Res 2020 Dec 18;22(12):e19127 [FREE Full text] [doi: 10.2196/19127] [Medline: 33337337]
82. Polkinghorne DE. Language and meaning: data collection in qualitative research. J Couns Psychol 2005;52(2):137-145 [FREE Full text] [doi: 10.1037/0022-0167.52.2.137]
83. Kallio H, Pietilä AM, Johnson M, Kangasniemi M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. J Adv Nurs 2016 Dec;72(12):2954-2965. [doi: 10.1111/jan.13031] [Medline: 27221824]
84. Nelson J. Using conceptual depth criteria: addressing the challenge of reaching saturation in qualitative research. Qual Res 2017 Oct;17(5):554-570 [FREE Full text] [doi: 10.1177/1468794116679873]
85. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
86. Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. Nurse Educ Today 2017 Sep;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
87. Rohwer A, Pfadenhauer L, Burns J, Brereton L, Gerhardus A, Booth A, et al. Series: Clinical Epidemiology in South Africa. Paper 3: logic models help make sense of complexity in systematic reviews and health technology assessments. J Clin Epidemiol 2017 Mar;83:37-47 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.06.012] [Medline: 27498377]
88. Carroll C, Booth A, Leaviss J, Rick J. "Best fit" framework synthesis: refining the method. BMC Med Res Methodol 2013 Mar 13;13:37 [FREE Full text] [doi: 10.1186/1471-2288-13-37] [Medline: 23497061]
89. Anderson LM, Petticrew M, Rehfuess E, Armstrong R, Ueffing E, Baker P. Using logic models to capture complexity in systematic reviews. Research Synthesis Methods 2011 Mar;2(1):33-42 [FREE Full text] [doi: 10.1002/jrsm.32] [Medline: 26061598]
90. Pfadenhauer L, Rohwer A, Burns J, Booth A, Lysdahl KB, Hofmann B, et al. Guidance for the assessment of context and implementation in health technology assessments (HTA) and systematic reviews of complex interventions: the context and implementation of complex interventions (CICI) framework. INTEGRATE-HTA Consortium. 2016 Feb 01. URL: https://www.researchgate.net/publication/298340571 [accessed 2023-04-07]
91. Hofmarcher MM, Festl E, Bishop-Tarver L. Health sector employment growth calls for improvements in labor productivity. Health Policy 2016 Aug;120(8):894-902 [FREE Full text] [doi: 10.1016/j.healthpol.2016.06.001]
92. Safavi K, Mathews SC, Bates DW, Dorsey ER, Cohen AB. Top-funded digital health companies and their impact on high-burden, high-cost conditions. Health Aff (Millwood) 2019 Jan;38(1):115-123. [doi: 10.1377/hlthaff.2018.05081] [Medline: 30615535]
93. Kelley LT, Fujioka J, Liang K, Cooper M, Jamieson T, Desveaux L. Barriers to creating scalable business models for digital health innovation in public systems: qualitative case study. JMIR Public Health Surveill 2020 Dec 10;6(4):e20579 [FREE Full text] [doi: 10.2196/20579] [Medline: 33300882]
94. Leone D, Schiavone F, Appio FP, Chiao B. How does artificial intelligence enable and enhance value co-creation in industrial markets? An exploratory case study in the healthcare ecosystem. J Bus Res 2021 May;129:849-859 [FREE Full text] [doi: 10.1016/j.jbusres.2020.11.008]
95. Seibert K, Domhoff D, Bruch D, Schulte-Althoff M, Fürstenau D, Biessmann F, et al. Application scenarios for artificial intelligence in nursing care: rapid review. J Med Internet Res 2021 Nov 29;23(11):e26522 [FREE Full text] [doi: 10.2196/26522] [Medline: 34847057]
96. Kraushaar J, Bohnet-Joschko S. Smartphone use and security challenges in hospitals: a survey among resident physicians in Germany. Int J Environ Res Public Health 2022 Dec 09;19(24):16546 [FREE Full text] [doi: 10.3390/ijerph192416546] [Medline: 36554426]
97. Korte L, Bohnet-Joschko S. Digitization in everyday nursing care: a vignette study in German hospitals. Int J Environ Res Public Health 2022 Aug 30;19(17):10775 [FREE Full text] [doi: 10.3390/ijerph191710775] [Medline: 36078491]
98. Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ 2021 Jun 14;373(1190):n1190 [FREE Full text] [doi: 10.1136/bmj.n1190]

99. Persson J, Rydenfält C. Why are digital health care systems still poorly designed, and why is health care practice not asking for more? Three paths toward a sustainable digital work environment. J Med Internet Res 2021 Jun 22;23(6):e26694 [FREE Full text] [doi: 10.2196/26694] [Medline: 34156336]
100. Mahajan S, Lu Y, Spatz ES, Nasir K, Krumholz HM. Trends and predictors of use of digital health technology in the United States. Am J Med 2021 Jan;134(1):129-134. [doi: 10.1016/j.amjmed.2020.06.033] [Medline: 32717188]
101. Kapoor A, Guha S, Kanti Das M, Goswami KC, Yadav R. Digital healthcare: the only solution for better healthcare during COVID-19 pandemic? Indian Heart J 2020 Mar;72(2):61-64 [FREE Full text] [doi: 10.1016/j.ihj.2020.04.001] [Medline: 32534691]

Abbreviations

AI: artificial intelligence
IoT: Internet of Things
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

---

Edited by Y Khader; submitted 29.06.22; peer-reviewed by S Sarejloo, O Tamburis, A Benis; comments to author 18.11.22; revised version received 13.01.23; accepted 07.02.23; published 21.04.23.

Please cite as:
Koebe P, Bohnet-Joschko S
The Impact of Digital Transformation on Inpatient Care: Mixed Methods Study
JMIR Public Health Surveill 2023;9:e40622
URL: https://publichealth.jmir.org/2023/1/e40622
doi:10.2196/40622
PMID:37083473

Review

Profiles and Findings of Population-Based Esophageal Cancer Screening With Endoscopy in China: Systematic Review and Meta-analysis

He Li[1*], PhD; Yi Teng[1*], BSc; Xinxin Yan[1*], BSc; Maomao Cao[1], MPhil; Fan Yang[1], MPhil; Siyi He[1], BSc; Shaoli Zhang[1], BSc; Qianru Li[1], BSc; Changfa Xia[1], PhD; Kai Li[2*], MD; Wanqing Chen[1*], PhD

[1]Office of Cancer Screening, National Cancer Center of China/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]Department of Surgical Oncology and General Surgery, Key Laboratory of Precision Diagnosis and Treatment of Gastrointestinal Tumors, Ministry of Education, The First Affiliated Hospital of China Medical University, Shenyang, China

[*]these authors contributed equally

Corresponding Author:
Wanqing Chen, PhD
Office of Cancer Screening, National Cancer Center of China/Cancer Hospital
Chinese Academy of Medical Sciences and Peking Union Medical College
No 17 Panjia-yuan South Lane, Chaoyang District
Beijing, 100021
China
Phone: 86 13241914190
Fax: 86 87787039
Email: chenwq@cicams.ac.cn

Abstract

Background: Population-based esophageal cancer (EC) screening trials and programs have been conducted in China for decades; however, screening strategies have been adopted in different regions and screening profiles are unclear.

Objective: We performed a meta-analysis to profile EC screening in China by positivity rate, compliance rate, and endoscopy findings, aiming to provide explicit evidence and recommendations for EC screening programs.

Methods: English (PubMed, Embase) and Chinese (China National Knowledge Infrastructure, Wanfang) language databases were systematically searched for population-based EC screening studies in the Chinese population until December 31, 2022. A meta-analysis was performed by standard methodology using a random-effects model. Pooled prevalence rates were calculated for three groups: high-risk areas with a universal endoscopy strategy, rural China with a risk-stratified endoscopic screening (RSES) strategy, and urban China with an RSES strategy. Positive cases included lesions of severe dysplasia, carcinoma in situ, intramucosal carcinoma, submucosal carcinoma, and invasive carcinoma.

Results: The pooled positivity rate of the high-risk population was higher in rural China (44.12%) than in urban China (23.11%). The compliance rate of endoscopic examinations was the highest in rural China (52.40%), followed by high-risk areas (50.11%), and was the lowest in urban China (23.67%). The pooled detection rate of positive cases decreased from 1.03% (95% CI 0.82%-1.30%) in high-risk areas to 0.48% (95% CI 0.25%-0.93%) in rural China and 0.12% (95% CI 0.07%-0.21%) in urban China. The pooled detection rate of low-grade intraepithelial neoplasia (LGIN) was also in the same order, being the highest in high-risk areas (3.99%, 95% CI 2.78%-5.69%), followed by rural China (2.55%, 95% CI 1.03%-6.19%) and urban China (0.34%, 95% CI 0.14%-0.81%). Higher detection rates of positive cases and LGIN were observed among males than among females and at older ages. The pooled early detection rate was 81.90% (95% CI 75.58%-86.88%), which was similar to the rates in high-risk areas (82.09%), in rural China (80.76%), and in urban China (80.08%).

Conclusions: Under the current screening framework, a higher screening benefit was observed in high-risk areas than in other regions. To promote EC screening and reduce the current inequality of screening in China, more focus should be given to optimizing strategies of high-risk individual assessment and surveillance management to improve compliance with endoscopic examination.

Trial Registration: PROSPERO CRD42022375720; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=375720

(JMIR Public Health Surveill 2023;9:e45360) doi:10.2196/45360

KEYWORDS

esophageal cancer; screening; high-risk individuals; detection rates; China

Introduction

Esophageal cancer (EC) represents a global public health burden, with the primary pathological type being esophageal squamous cell carcinoma (ESCC) [1]. More than 50% of global ESCC cases occur in China, causing severe disease and an economic burden for the country [2,3]. A series of prevention and control strategies for EC have been implemented in China since the 1960s, especially in well-known high-risk areas such as Linzhou of Henan Province, where the incidence and mortality of EC were estimated to be approximately five times higher than the national average levels [4]. These strategies and actions have brought success, contributing to an average 4.5% reduction in the incidence and mortality of EC in China since 2000 [5]. In addition, the results of a screening program in high-risk areas in China demonstrated the effectiveness of EC screening with endoscopy examination in reducing EC incidence and mortality [6], which provided high-quality evidence and strong application recommendations for international EC screening and early detection and treatment.

Three organized screening programs have been launched based on the National Key Public Health Project since 2005 to conduct population-based EC screening in the high-risk areas of Huaihe River and urban China [7]. The primary aim of these programs is to reduce the incidence and mortality of EC and to explore a suitable and feasible EC screening strategy. In practice, two main EC screening strategies have been adopted and implemented in China. One is the universal endoscopy screening strategy, which is used in high-risk areas. The other is the risk-stratified endoscopic screening (RSES) strategy, which provides endoscopies for a limited group of individuals at high risk of EC and is currently widely implemented in nonhigh-risk areas [7]. These programs and other population-based ESCC screening studies have covered more than 150 counties or cities in China [7], providing and accumulating a great deal of experience and real-world data for ESCC prevention and control in China and internationally.

Clear and comprehensive knowledge of the profile of EC screening, including the positivity rate of high-risk individuals for EC, compliance with endoscopy screening, and endoscopy findings in real-world EC screening programs, has potential public health value. For instance, such knowledge would help health policy makers understand the actual acceptability of EC screening, direct benefits from EC screening, and burden of surveillance endoscopy. In addition, this crucial real-world evidence would promote the further implementation of EC screening programs in China and other countries that face a high EC burden. To the best of our knowledge, EC screening profiles in China are unclear based on existing studies and the literature. To obtain a comprehensive profile of EC screening in China, we performed this systematic review and meta-analysis to estimate the positivity rate of EC among high-risk individuals, compliance with endoscopy screening, and endoscopy findings by high-risk and nonhigh-risk areas (rural China and urban China). By summarizing these indices, we hope to provide more explicit evidence and recommendations for Chinese EC prevention and control and to promote further EC screening in other countries facing the threat of a high EC burden, which has great importance for public health.

Methods

Design

The protocol for this systematic review was published in the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD42022375720) and the review was designed following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines 2020 [8].

Data Sources and Search Strategy

The English-language (PubMed, Embase) and Chinese (China National Knowledge Infrastructure, Wanfang) databases were systematically searched for population-based EC screening studies in the Chinese population until December 31, 2022, in any language. The following relevant Medical Subject Heading (MeSH) terms and key words were used for the search: esophageal neoplasms, screening, and China (see Table S1 in Multimedia Appendix 1 for the full search strategy). Two researchers (HL, YT) performed the literature search independently and discrepancies were resolved by consultation with a third researcher (XY).

Study Selection

Inclusion Criteria

Studies were selected based on the following inclusion criteria: (1) studies that reported compliance with endoscopy examination and/or endoscopy findings from population-based EC screening in the Chinese population without restrictions on age; and (2) cross-sectional studies, cohort studies, and randomized controlled trials that reported baseline findings.

Exclusion Criteria

Studies were excluded if they (1) were nonpopulation-based screening studies, (2) lacked information on compliance rates with endoscopy exam and endoscopy findings of positive cases or low-grade intraepithelial neoplasia (LGIN), and (3) were studies with duplicated data for all outcomes of interest published elsewhere or parts of another study from the same geographical location.

Data Extraction and Quality Assessment

For each included study, two reviewers (HL, YT) independently assessed the studies and extracted the data for analysis. The extracted information included the first author and year of publication, study periods, study region, number of study centers, screening strategy (universal endoscopy screening/RSES), sex (proportion of males), age, available data of sex-specific and age-specific subgroups, and available data of early detection among patients with positive screening results.

The study quality and risk of bias were assessed using an instrument developed by Hoy and colleagues [9] for population-based prevalence studies. The tool has 10 questions, and a score of 1 (yes) or 0 (no) was assigned for each item (see Table S2 in Multimedia Appendix 1). Scores were summed across items to generate an overall quality score that ranged from 0 to 10, and studies were then classified as having a low (scores≥8), moderate (scores of 6-7), or high (scores≤5) risk of bias. This method has been used in previous systematic reviews. This assessment tool was tested in previous studies; it is easy to use and deals with risk of bias well.

Outcomes Assessed

The pooled outcomes assessed included the positivity rate of high-risk individuals for EC/upper gastrointestinal (UGI) cancer, compliance rates to endoscopy examinations, prevalence of endoscopy findings (positive cases, LGIN, and negative endoscopy), and early detection rate. The definition of each index is described in Table S3 of Multimedia Appendix 1. Positive cases included lesions of severe dysplasia, carcinoma in situ, intramucosal carcinoma, submucosal carcinoma, and invasive carcinoma. LGIN included lesions of mild dysplasia and moderate dysplasia. Negative endoscopy findings were defined as a baseline endoscopic examination that did not reveal any of the dysplasia lesions mentioned above.

Statistical Analysis

We used meta-analysis techniques with logit transformation to calculate the pooled proportions along with the corresponding 95% CIs under a random-effects model. We calculated the pooled prevalence in three populations with different screening strategies: (1) high-risk areas with a universal endoscopy strategy, (2) rural China with an RSES strategy, and (3) urban China with an RSES strategy. Analyses of the pooled prevalence of endoscopy findings by sex and age group (40-49, 50-59, and 60-69 years) were also performed. We performed sensitivity analyses by publication year, sample size, and the number of study centers to evaluate the robustness of the study.

Heterogeneity between studies was assessed with the I^2 statistic, which estimates the percentage of the total variation across studies due to true between-study differences. Generally, I^2values greater than 60%-70% indicate the presence of substantial heterogeneity. Publication bias was inspected visually on a funnel plot and by the Egger test; *P*<.05 was considered statistically significant. All meta-analyses were carried out using R software (version 4.1.0; R Foundation for Statistical Computing, Vienna, Austria).

Ethical Considerations

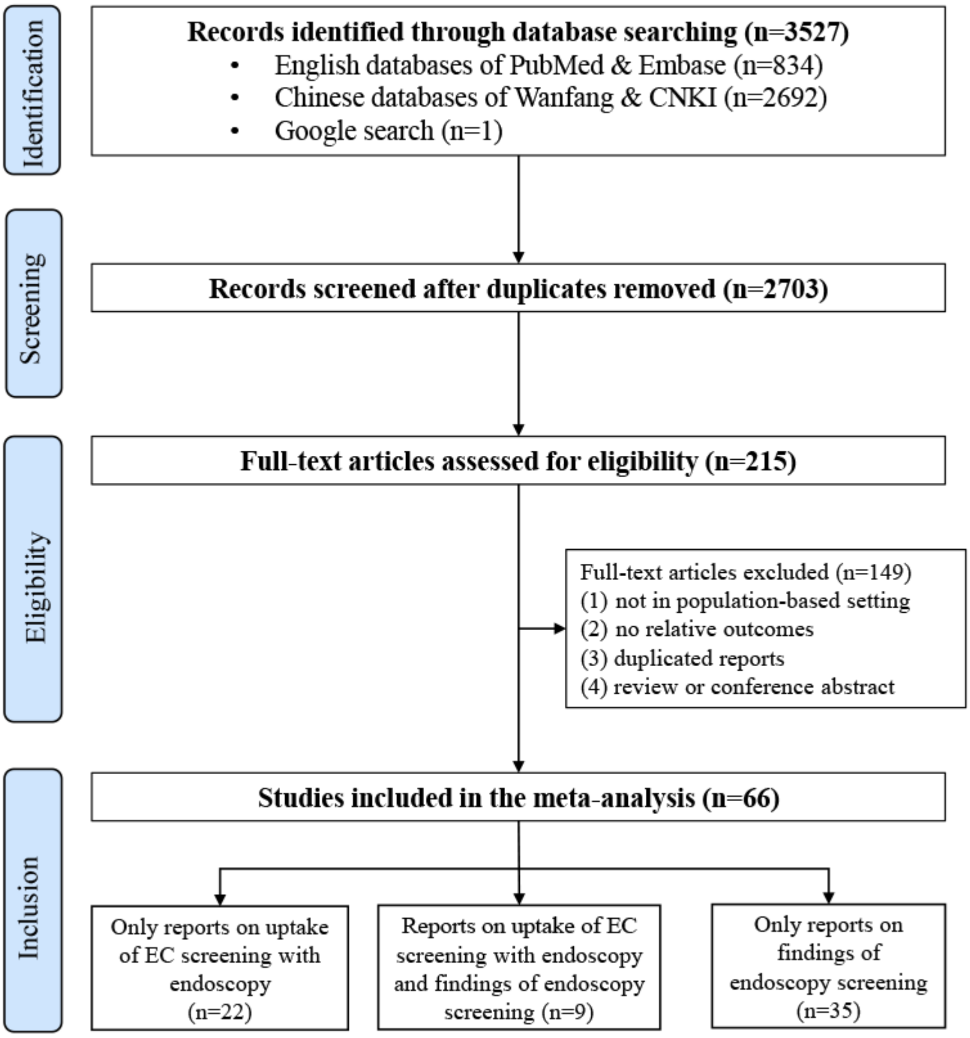
This is a meta-analysis of published studies and data, and did not involve active human participants and/or animals. The present study used only publicly available summary-level statistics and did not involve individual information. Therefore, formal consent, informed consent, institutional review board approval, and ethics approval are not applicable and/or not needed.

Results

Search Results

A schematic diagram of study selection is provided in Figure 1. The initial search identified 3527 records, and 2703 titles and abstracts were screened after removing duplicates. A total of 215 full-length articles were evaluated in detail and 66 studies were included in the final analysis. Among these studies, 22 studies only reported the compliance of EC screening with endoscopy, 35 studies only reported the findings of endoscopy examinations, and 9 studies reported both outcomes.

Figure 1. Flowchart presenting the selection of studies for inclusion in the systematic review and meta-analysis. CNKI: China National Knowledge Infrastructure; EC: esophageal cancer.



Study Characteristics and Quality Assessment

The characteristics of the 66 included studies are outlined in Table S4 of Multimedia Appendix 1, and the summary characteristics about reporting compliance with endoscopy examinations and endoscopy findings are summarized in Table 1 and Table 2, respectively. Eligible ages for EC screening were 40-69 years in all studies conducted in high-risk areas and rural China, except for one study performed in Hua County with an age range of 25-65 years [10]. Eligibility ages in urban China were primarily 40-74 years, and the remaining studies recommended 40-69 years.

A total of 2,472,920 asymptomatic individuals were included in 31 studies that reported compliance with endoscopy examinations, one of which reported data from both rural and urban China [11]. Therefore, 2,006,235 individuals were from high-risk areas during 1999-2017 (n=12) [6,10,12-21], 125,473 individuals were from rural China during 2010-2017 (n=2) [11,22], and 341,212 individuals were from urban China during 2012-2019 (n=18) [11,23-39], as shown in Table 1 and Table S4 of Multimedia Appendix 1. The details of the high-risk assessment strategy to select targets to undergo further endoscopy screening are summarized in Table S5 of Multimedia Appendix 1.

A total of 997,004 individuals had reports of endoscopy findings in 44 studies, among which one study reported endoscopy findings in high-risk areas, rural China, and urban China [11]. Therefore, a total of 876,170 individuals aged 40-69 years (n=36 studies) [6,11,12,14-17,40-68] from high-risk areas underwent endoscopic examinations, where a universal endoscopy strategy was adopted. A total of 102,413 individuals with a high risk of EC (n=4 studies) [11,69-71] aged 40-69 years in rural China and 18,421 individuals with a high risk of EC/UGI cancer (n=6 studies) [11,29,33,72-74] aged 40-74 years in urban China underwent endoscopic examinations where the RSES strategy was adopted (Table 2). The number of individual study samples ranged from 1151 [74] to 116,630 [67] across all included studies. Study participants were enrolled from 1999 to 2019 in high-risk areas, followed by 2007 to 2017 in rural China and 2015 to 2019 in urban China. Study regions also presented disparities: study regions in high-risk areas mainly consisted of north China (22%) and southwest China (22%), followed by eastern China (19%), central China (17%), and northwest China (11%). Study regions in rural China were mostly concentrated in east China (50%) and central China (25%). Study regions in urban China were relatively evenly distributed across the country, except in central and northeast China (Table 2, Table S4 in Multimedia Appendix 1).

Of the 31 studies reporting compliance with EC screening using endoscopy, 12 (39%) and 19 (61%) were classified as having a low and moderate risk of bias, respectively (Table S6 of Multimedia Appendix 1). Of the 44 studies that reported findings of endoscopy examinations, 39 (89%) and 5 (11%) were classified as having a low and moderate risk of bias, respectively (Table S7 in Multimedia Appendix 1).

Table 1. Main characteristics of the 31 included studies that reported compliance with endoscopy exams in population-based esophageal cancer screening in China by region.

Characteristics	High-risk areas	Rural China	Urban China
Age range (years)	40-69[a]	40-69	40-74[b]
Screening strategy	Universal endoscopy screening	Risk-stratified endoscopy screening	Risk-stratified endoscopy screening
Initial screening	No	Risk assessment	Risk assessment
Examination technique	Endoscopy with pathology	Endoscopy with pathology	Endoscopy with pathology
Number of studies[c]	12	2	18
Year of enrollment	1999-2017	2010-2017	2012-2019
Year of publication	2003-2021	2020-2021	2016-2021
Number of participants	2,006,235	125,473	341,212
Representativeness, n (%)			
Single center	10 (83.33)	0 (0.00)	12 (66.67)
Multiple centers within one province	0 (0.00)	0 (0.00)	5 (27.78)
Multiple centers across two provinces or more	2 (16.67)	2 (100.00)	1 (5.55)
Regions, n (%)			
Central China	1 (8.33)	0 (0.00)	3 (16.67)
East China	1 (8.33)	0 (0.00)	6 (33.32)
North China	4 (33.33)	0 (0.00)	2 (11.11)
South China	0 (0.00)	0 (0.00)	2 (11.11)
Northeast China	0 (0.00)	0 (0.00)	1 (5.56)
Southwest China	3 (25.00)	0 (0.00)	1 (5.56)
Northwest China	1 (8.33)	0 (0.00)	2 (11.11)
Multiple regions	2 (16.68)	2 (100.00)	1 (5.56)

[a] The age of eligibility of one study in high-risk areas was 25-65 years and that of the others was 40-69 years.

[b] The age of eligibility of one study in urban China was >35 years and that of the others was 40-69 years.

[c] One study reported compliance with endoscopy examinations in both rural China and urban China.

Table 2. Main characteristics of the 44 included studies that reported endoscopy findings in population-based esophageal cancer screening in China by region.

Characteristics	High-risk areas	Rural China	Urban China
Number of studies[a]	36	4	6
Year of enrollment	1999-2019	2007-2017	2015-2019
Year of publication	2003-2022	2019-2020	2017-2021
Number of participants	876,170	102,413	18,421
Age range (years)	40-69[b]	40-69	40-74
Representativeness, n (%)			
Single center	20 (55.56)	0 (0.00)	3 (50.00)
Multiple centers within one province	13 (36.11)	3 (75.00)	2 (33.33)
Multiple centers across two provinces or above	3 (8.33)	1 (25.00)	1 (16.67)
Regions, n (%)			
Central China	6 (16.67)	1 (25.00)	0 (0.00)
East China	7 (19.44)	2 (50.00)	1 (16.67)
North China	8 (22.22)	0 (0.00)	1 (16.67)
South China	0 (0.00)	0 (0.00)	1 (16.67)
Northeast China	0 (0.00)	0 (0.00)	0 (0.00)
Southwest China	8 (22.22)	0 (0.00)	1 (16.67)
Northwest China	4 (11.11)	0 (0.00)	1 (16.67)
Multiple regions	3 (8.34)	1 (25.00)	1 (16.67)

[a] One study reported endoscopy findings in high-risk areas, rural China, and urban China.

[b] The age of eligibility of one study in high-risk areas was 25-65 years and that of the others was 40-69 years.

Pooled Outcomes

Pooled Positivity Rates of the High-Risk Population and Compliance With Endoscopic Examinations

All pooled results are summarized in Table 3. A total of 373,516 (in rural China) and 1,449,536 (in urban China) individuals of the asymptomatic population completed individualized risk assessment by professional staff from EC screening programs or trials. The pooled positivity rate of the high-risk population in rural China was higher than that in urban China; the details are shown in Figure S1 of Multimedia Appendix 1. The overall compliance rate with endoscopic examinations among the 31 included studies was 35.39% (95% CI 27.96%-42.82%), covering 2,472,920 eligible individuals for endoscopic examinations. The compliance rate was the highest in rural China, followed by high-risk areas, and was the lowest in urban China, as shown in Figure 2 and Table 3.

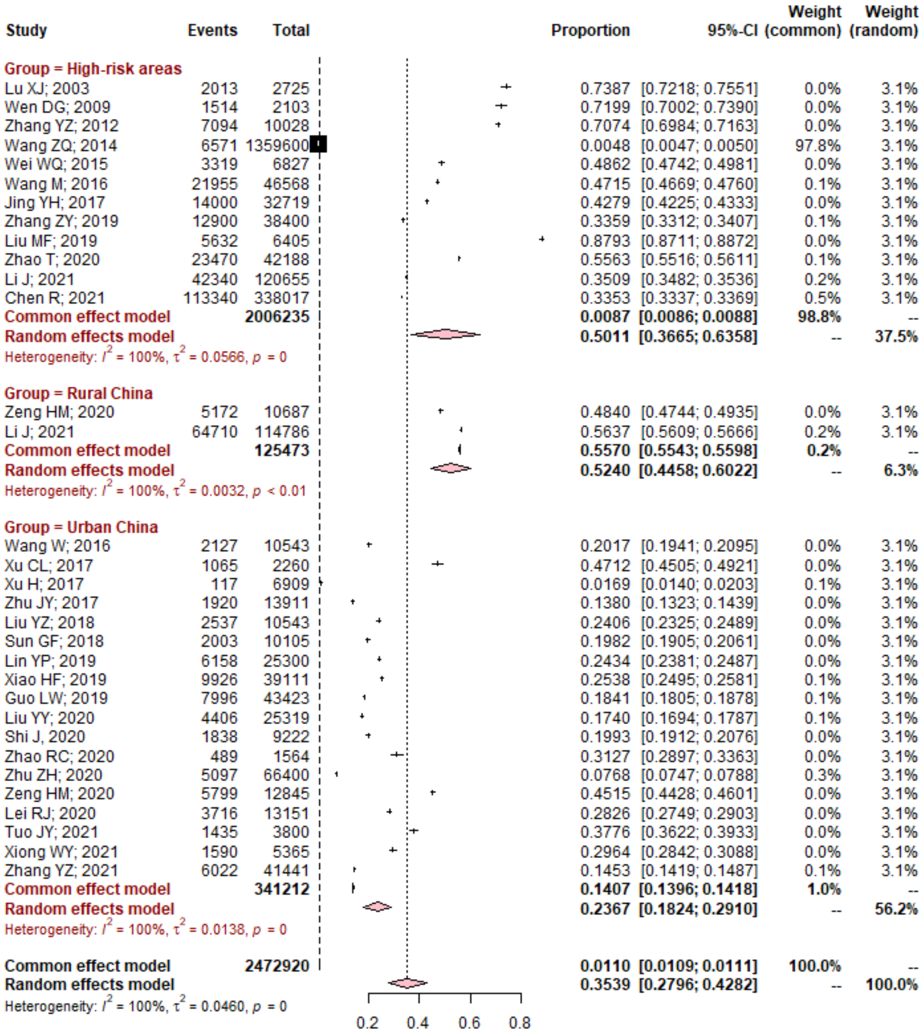
Table 3. Summary of pooled positivity rates of high-risk individuals, compliance rates of endoscopy examinations, and prevalence of endoscopy findings in population-based esophageal cancer (EC) screening in China by region.

Assessed metrics in population-based EC screening	High-risk areas		Rural China		Urban China	
	Pooled rate, % (95% CI)	I^2	Pooled rate, % (95% CI)	I^2	Pooled rate, % (95% CI)	I^2
Positivity rates of high-risk individuals in EC screening	—[a]	—	44.12 (28.76-60.69)	100%	23.11 (19.52-27.13)	100%
Compliance rates of endoscopy examinations	50.11 (36.65-63.58)	100%	52.40 (44.58-60.22）	100%	23.67 (18.24-29.10)	100%
Endoscopy findings						
Negative endoscopy findings	94.72 (92.82-96.14)	100%	96.93 (92.99-98.68)	99%	99.41 (99.07-99.62)	82%
LGIN[b]	3.99 (2.78-5.69)	100%	2.55 (1.03-6.19)	99%	0.34 (0.14-0.81)	84%
Positive cases	1.03 (0.82-1.30)	98%	0.48 (0.25-0.93)	98%	0.12 (0.07-0.21)	46%
Early detection rates	82.09 (74.85-87.59)	93%	80.76 (69.24-88.68)	93%	80.08 (27.21-97.74)	85%

[a] Not applicable.

[b] LGIN: low-grade intraepithelial neoplasia.

Figure 2. Pooled compliance rates of endoscopy examinations in population-based esophageal cancer screening by region (high-risk areas, rural China, and urban China).



Study	Events	Total	Proportion	95%-CI	Weight (common)	Weight (random)
Group = High-risk areas						
Lu XJ; 2003	2013	2725	0.7387	[0.7218; 0.7551]	0.0%	3.1%
Wen DG; 2009	1514	2103	0.7199	[0.7002; 0.7390]	0.0%	3.1%
Zhang YZ; 2012	7094	10028	0.7074	[0.6984; 0.7163]	0.0%	3.1%
Wang ZQ; 2014	6571	1359600	0.0048	[0.0047; 0.0050]	97.8%	3.1%
Wei WQ; 2015	3319	6827	0.4862	[0.4742; 0.4981]	0.0%	3.1%
Wang M; 2016	21955	46568	0.4715	[0.4669; 0.4760]	0.1%	3.1%
Jing YH; 2017	14000	32719	0.4279	[0.4225; 0.4333]	0.0%	3.1%
Zhang ZY; 2019	12900	38400	0.3359	[0.3312; 0.3407]	0.1%	3.1%
Liu MF; 2019	5632	6405	0.8793	[0.8711; 0.8872]	0.0%	3.1%
Zhao T; 2020	23470	42188	0.5563	[0.5516; 0.5611]	0.1%	3.1%
Li J; 2021	42340	120655	0.3509	[0.3482; 0.3536]	0.2%	3.1%
Chen R; 2021	113340	338017	0.3353	[0.3337; 0.3369]	0.5%	3.1%
Common effect model		**2006235**	**0.0087**	**[0.0086; 0.0088]**	**98.8%**	--
Random effects model			**0.5011**	**[0.3665; 0.6358]**	--	**37.5%**
Heterogeneity: $I^2 = 100\%$, $\tau^2 = 0.0566$, $p = 0$						
Group = Rural China						
Zeng HM; 2020	5172	10687	0.4840	[0.4744; 0.4935]	0.0%	3.1%
Li J; 2021	64710	114786	0.5637	[0.5609; 0.5666]	0.2%	3.1%
Common effect model		**125473**	**0.5570**	**[0.5543; 0.5598]**	**0.2%**	--
Random effects model			**0.5240**	**[0.4458; 0.6022]**	--	**6.3%**
Heterogeneity: $I^2 = 100\%$, $\tau^2 = 0.0032$, $p < 0.01$						
Group = Urban China						
Wang W; 2016	2127	10543	0.2017	[0.1941; 0.2095]	0.0%	3.1%
Xu CL; 2017	1065	2260	0.4712	[0.4505; 0.4921]	0.0%	3.1%
Xu H; 2017	117	6909	0.0169	[0.0140; 0.0203]	0.1%	3.1%
Zhu JY; 2017	1920	13911	0.1380	[0.1323; 0.1439]	0.0%	3.1%
Liu YZ; 2018	2537	10543	0.2406	[0.2325; 0.2489]	0.0%	3.1%
Sun GF; 2018	2003	10105	0.1982	[0.1905; 0.2061]	0.0%	3.1%
Lin YP; 2019	6158	25300	0.2434	[0.2381; 0.2487]	0.0%	3.1%
Xiao HF; 2019	9926	39111	0.2538	[0.2495; 0.2581]	0.1%	3.1%
Guo LW; 2019	7996	43423	0.1841	[0.1805; 0.1878]	0.1%	3.1%
Liu YY; 2020	4406	25319	0.1740	[0.1694; 0.1787]	0.1%	3.1%
Shi J, 2020	1838	9222	0.1993	[0.1912; 0.2076]	0.0%	3.1%
Zhao RC; 2020	489	1564	0.3127	[0.2897; 0.3363]	0.0%	3.1%
Zhu ZH; 2020	5097	66400	0.0768	[0.0747; 0.0788]	0.3%	3.1%
Zeng HM; 2020	5799	12845	0.4515	[0.4428; 0.4601]	0.0%	3.1%
Lei RJ; 2020	3716	13151	0.2826	[0.2749; 0.2903]	0.0%	3.1%
Tuo JY; 2021	1435	3800	0.3776	[0.3622; 0.3933]	0.0%	3.1%
Xiong WY; 2021	1590	5365	0.2964	[0.2842; 0.3088]	0.0%	3.1%
Zhang YZ; 2021	6022	41441	0.1453	[0.1419; 0.1487]	0.1%	3.1%
Common effect model		**341212**	**0.1407**	**[0.1396; 0.1418]**	**1.0%**	--
Random effects model			**0.2367**	**[0.1824; 0.2910]**	--	**56.2%**
Heterogeneity: $I^2 = 100\%$, $\tau^2 = 0.0138$, $p = 0$						
Common effect model		**2472920**	**0.0110**	**[0.0109; 0.0111]**	**100.0%**	--
Random effects model			**0.3539**	**[0.2796; 0.4282]**	--	**100.0%**
Heterogeneity: $I^2 = 100\%$, $\tau^2 = 0.0460$, $p = 0$						

Pooled Rates of Endoscopy Findings

The prevalence of endoscopy findings varied widely by region. Specifically, the pooled detection rates of positive cases decreased from high-risk areas to rural China and urban China (Table 3, Figure S2 in Multimedia Appendix 1). The pooled detection rate of LGIN was also in the same order, being the highest in high-risk areas, followed by rural China and urban China, as shown in Table 3 and Figure S3 in Multimedia Appendix 1. Consequently, the detection rates of negative endoscopy findings increased from high-risk areas to rural China and urban China (Table 3, Figure S4 in Multimedia Appendix 1).

Pooled Rates of Early Detection Among Positive Cases

A total of 34 of the 44 included studies reported early detection cases, covering 8445 patients diagnosed with positive cases at endoscopy screening (Figure S5 in Multimedia Appendix 1). The pooled early detection rate was approximately 81.90% (95% CI 75.58%-86.88%), which was similar to the rates in high-risk areas, rural China, and urban China (Table 3).

Pooled Detection Rates of Endoscopy Findings by Subgroups

Table 4 summarizes the pooled rates of endoscopy findings of sex-specific and age-specific subgroup analyses. Overall, males had higher detection rates of positive cases (1.65% vs 0.82%) and LGIN (5.71% vs 3.66%) and a lower detection rate of negative endoscopy findings (92.33% vs 95.38%) than females. The detection rates of positive cases and LGIN increased with age group, as the rates of positive cases were 0.52% in the 40-49 years group, 1.26% in the 50-59 years group, and 2.73% in the 60-69 years group, and the corresponding rates of LGIN were 2.62%, 4.24%, and 7.13%, respectively. The sex-specific and age-group detection rates of endoscopy screening in the three specific areas are also shown in Table 4, demonstrating the highest detection rates of positive cases and LGIN in high-risk areas, followed by rural China and urban China in each subgroup.

Table 4. Prevalence of negative endoscopy findings, low-grade intraepithelial neoplasia (LGIN), and positive cases in different subgroups in population-based esophageal cancer screening in China.

Characteristics	Studies, n	Study participants, n	Negative endoscopy findings		LGIN		Positive cases	
			Pooled prevalence rate, % (95% CI)	I^2	Pooled prevalence rate, % (95% CI)	I^2	Pooled prevalence rate, % (95% CI)	I^2
All								
Sex								
Male	17	165,522	92.33 (89.03-94.69)	100%	5.71 (3.71-8.69)	99%	1.65 (1.22-2.24)	97%
Female	17	214,180	95.38 (92.82-97.06)	100%	3.66 (2.22-5.98)	100%	0.82 (0.55-1.20)	97%
Age group (years)								
40-49	12	104,222	96.57 (93.33-98.26)	100%	2.62 (1.42-4.79)	99%	0.52 (0.22-1.23)	100%
50-59	12	105,033	94.04 (89.40-96.72)	100%	4.24 (2.17-8.11)	100%	1.26 (0.71-2.21)	98%
60-69	12	90,413	89.10 (82.14-93.56)	100%	7.13 (4.00-12.40)	99%	2.73 (1.57-4.71)	99%
High-risk areas								
Sex								
Male	13	124,774	91.04 (86.81-94.00)	100%	6.60 (4.08-10.49)	100%	1.99 (1.46-2.71)	96%
Female	13	157,326	94.40 (91.13-96.51)	100%	4.39 (2.56-7.41)	100%	0.95 (0.64-1.42)	97%
Age group (years)								
40-49	8	73,745	94.79 (88.78-97.66)	100%	3.80 (1.92-7.40)	99%	0.80 (0.26-2.48)	100%
50-59	8	68,611	91.81 (84.34-95.89)	100%	5.46 (2.41-11.89)	100%	2.05 (1.19-3.50)	98%
60-69	8	59,762	84.81 (75.09-91.18)	100%	9.17 (4.62-17.40)	100%	4.30 (2.50-7.30)	99%
Rural China								
Sex								
Male	3	58,943	93.63 (92.09-94.88)	98%	5.42 (4.58-6.42)	96%	0.92 (0.56-1.50)	97%
Female	3	81,694	96.05 (95.10-96.82)	98%	3.47 (2.96-4.06)	95%	0.45 (0.24-0.83)	98%
Age group (years)								
40-49	3	46,472	97.93 (97.70-98.14)	63%	1.86 (1.69-2.04)	47%	0.21 (0.15-0.32)	74%
50-59	3	52,811	95.22 (94.54-95.81)	92%	4.26 (3.93-4.61)	74%	0.50 (0.29-0.85)	95%
60-69	3	41,339	92.02 (89.52-93.96)	99%	6.69 (5.34-8.36)	98%	1.24 (0.71-2.15)	98%
Urban China								
Sex								
Male	1	479	98.75 (97.24-99.44)	—[a]	0.63 (0.20-1.92)	—	0.63 (0.20-1.92)	—
Female	1	672	99.70 (98.82-99.93)	—	0.15 (0.02-1.05)	—	0.15 (0.02-1.05)	—
Age group (years)								
40-49	1	347	99.71 (97.98-99.96)	—	0	—	0.29 (0.04-2.02)	—
50-59	1	394	99.49 (97.99-99.87)	—	0.25 (0.04-1.78)	—	0.25 (0.04-1.78)	—
60-69	1	409	98.78 (97.10-99.49)	—	0.73 (0.24-2.25)	—	0.49 (0.12-1.93)	—

[a] Not applicable.

Validation of the Meta-analysis Results

Sensitivity Analysis

To assess whether publication year, sample size, and the number of study centers had a dominant effect on the meta-analysis, we additionally analyzed their effect on the detection of positive cases (see Table S8 of Multimedia Appendix 1). The findings in this analysis showed that studies with publication years before 2009 and sample sizes less than 5000 had a significantly higher detection rate of positive cases in high-risk areas. In addition, multicenter studies presented lower heterogeneity in urban China.

Heterogeneity

High heterogeneity (I^2) among studies was observed for all estimated indices, and these findings remained unchanged in the subgroups (Table 4; Table S8 in Multimedia Appendix 1).

Publication Bias

Publication bias assessment was performed based on the prevalence of positive cases. Visual inspection of the funnel plots (Figure S6 in Multimedia Appendix 1) and the two-tailed Egger test (P=.10) demonstrated no evidence of publication bias.

Discussion

Based on this systematic review and meta-analysis of 66 studies covering over 2 million asymptomatic Chinese populations participating in population-based EC screening, we summarized a series of crucial indices to profile the status of EC screening in China. Under the current EC screening practice, the positivity rates of the high-risk population and compliance rates with endoscopic examinations, yields from EC screening, and burden of surveillance endoscopy varied greatly in populations from high-risk areas and other nonhigh-risk areas (rural China and urban China). These findings not only show the actual situation of current population-based ESCC screening in China but also reflect corresponding challenges and potential future directions in scientific research and policy-making to promote EC prevention in China.

Identifying high-risk individuals for EC to undergo further endoscopic examinations has been one of the most critical factors for the success of ESCC screening programs. Selection of assessment items and the threshold to define endoscopic examination eligibility are essential but troublesome tasks in this field [75]. In nonhigh-risk areas of EC where the RSES strategy was adopted, the pooled positivity rate of high-risk individuals for endoscopic examinations in the rural Chinese population was approximately twice that in the urban Chinese population (44.12% vs 23.11%), fully reflecting the rural-urban inequality of exposure to EC risk factors. However, the existing individualized assessment tool in EC screening programs or trials in China differed in the assessment items, thresholds, and definition of outcomes (EC or UGI cancer), as summarized in Table S5 of Multimedia Appendix 1. Few assessment tools have estimated the accuracy and effectiveness of EC screening. In addition, all the existing prediction models for EC presented a high risk of bias and none of these models was estimated in diverse populations, which limited the implementation values in large-scale population-based EC screening programs or even the national EC screening program [75]. Before implementing the national EC screening program, development of a specific high-risk assessment strategy to define eligible individuals for endoscopy screening should be prioritized. Through comprehensive estimation and validation in diverse populations with national representatives, updated individualized EC risk assessment strategies could be developed in the future, serving EC screening in China and other countries.

Optimal compliance for endoscopic examinations in eligible asymptomatic individuals is essential to confer benefits from EC screening, as shown in cohort and modeling studies [6,76,77]. However, the overall compliance of endoscopy screening in existing population-based EC screening programs or trials in China was found to be suboptimal, especially in the urban Chinese population, with a compliance rate of less than 24%. Therefore, improving the endoscopic compliance rate by adopting comprehensive strategies is an urgent need for EC screening programs. Several available measures could be taken in the future. First, from the aspect of providers of EC screening programs, there is an urgent need to conduct multifaceted actions to promote and improve the core knowledge of cancer prevention and control in Chinese residents and help them correctly understand cancer and prevent it [78]. A previous EC screening study among 28,543 high-risk individuals for EC showed a positive association between cancer prevention awareness and compliance with endoscopic examinations [79]. This highlights the importance and feasibility of raising eligible individuals' awareness of cancer prevention in improving compliance rates in population-based EC screening programs. Second, for community servers in population-based EC screening, more attention and interventions should be paid to several vulnerable populations in health literacy to improve their compliance rate with endoscopic examinations. These populations included males, cigarette smokers, and those with low socioeconomic status, who had a higher risk of developing EC but lower compliance rates of endoscopic examinations, as observed in previous population-based EC screening studies [31,33,79]. In addition, optimization of the screening procedure, such as shortening the waiting time for endoscopic examinations, may be considered another effective measure to improve compliance, as indicated in population-based cancer screening programs [80].

The endoscopy findings exhibited apparent disparities among studies, mainly due to the heterogeneous study populations in the disease burden of EC, different EC screening strategies, and varying compliance rate for endoscopic examination, which were mentioned above. Specifically, a higher screening benefit was obtained in high-risk areas, followed by rural China and urban China, with pooled prevalence rates of positive cases being 1.03%, 0.48%, and 0.12%, respectively. In addition, early detection among positive cases was more than 80% in all studies, which emphasizes the importance of EC screening in improving survival and quality of life. In terms of the surveillance endoscopy burden, a higher burden of LGIN was also observed in high-risk areas, followed by rural China and urban China, with pooled rates of 3.99%, 2.55%, and 0.34%, respectively. A much larger fraction of the baseline endoscopy screening population comprised those diagnosed with negative endoscopy findings, with prevalence higher than 90% in all previous reports. These results will help policy makers grasp the profile of population-based EC screening under the framework of current strategies in different regions and provide essential benchmarks for health economic evaluation models or other studies. Moreover, there is a considerable need for managing the vast population identified with negative endoscopy findings and LGIN [7]. The current Chinese guidelines for EC screening recommend surveillance intervals of 1-3 years for patients with LGIN and 5 years for individuals with negative endoscopy findings [81]. However, these recommendations were developed based on limited studies with different findings on observational or modeling studies from ESCC high-risk areas in China [76,77,82-84]. More high-quality studies are urgently needed to estimate and update these recommendations.

To the best of our knowledge, this is the first study to provide a comprehensive profile of the available research on population-based EC screening in the Chinese population, which offers comprehensive and objective evidence for policy makers. Findings in this study have important implications for better conducting EC screening in the future. First, development of high-risk assessment criteria for selecting eligible individuals for endoscopic examinations for diverse populations should be prioritized. Second, multidimensional attempts are urgently needed to improve compliance with endoscopic examinations, including improving residents' cancer literacy and optimizing and simplifying screening procedures. Third, more robust and high-quality cohort and modeling studies are urgently needed to estimate or optimize the current surveillance strategy, which will essentially compensate for the limited evidence in the Chinese population.

Our study also has several limitations. First, significant heterogeneity was observed in all pooled indices. Understandably, high heterogeneity is unavoidable given that study-related factors (screening strategy and population-based disease burden of EC) and individual-related factors (individualized risk of EC, compliance with endoscopic examinations, age, and sex) could not be controlled for in all studies. In this meta-analysis, we performed several subgroup analyses with consideration of these significant factors contributing to heterogeneity. Second, a detailed comparison of exposure factors related to developing EC and determinations of compliance with endoscopic examinations in rural-urban Chinese populations were not performed due to data limitations in the existing studies. Third, subgroup analysis of endoscopy findings was limited to sex-specific and age-specific findings. Other potential essential factors, including individual-related (such as smoking status, family history of cancer, and disease history of the digestive system) and procedure-related (such as endoscopist and pathologist performance and quality of endoscopic examinations) factors, were limited due to the lack of data. In addition, available studies and data under different EC screening strategies need to be more balanced. Data on endoscopy findings in high-risk areas were sufficient; however, data to profile EC screening in nonhigh-risk areas were limited. There were only four studies in rural China that reported compliance with endoscopy, and only 18,421 individuals with endoscopy screening were reported in the urban Chinese population. These limited data could not comprehensively profile actual population-based EC screening in this population, and an updated meta-analysis will be performed when more evidence is available.

In summary, in this meta-analysis, the current profile of population-based EC screening in China and potential challenges were summarized with a series of indices in screening procedures (positivity rates of high-risk individuals, compliance rates, and endoscopic findings) by different regions. Under the current screening framework, a higher screening benefit was observed in high-risk areas than in other regions. To promote EC screening and reduce the current inequality of screening in China, more focus should be given to optimizing strategies of high-risk individual assessment and surveillance management to improve compliance with endoscopic examination.

Acknowledgments

We are sincerely grateful for funding from the Fundamental Research Funds for the Central Universities (grant number 3332022136); National Natural Science Foundation of China (grant number 82273721); Cooperation Project in Beijing, Tianjin, and Hebei of China (J200017); and Sanming Project of Medicine in Shenzhen (grant number SZSM201911015). We thank all of the individuals who contributed to this study for their important and greatly appreciated contributions to the preparation of this report.

Data Availability

The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HL, KL, and WC contributed to the study concept and design. HL, YT, and XY contributed to the screening and data abstraction and quality assessment of the included studies. HL and XY collaborated in drafting the manuscript and revising it critically for important intellectual content. All authors contributed to the data interpretation and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supporting methods and results.
[DOCX File , 12883 KB - publichealth_v9i1e45360_app1.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Morgan E, Soerjomataram I, Rumgay H, Coleman H, Thrift A, Vignat J, et al. The global landscape of esophageal squamous cell carcinoma and esophageal adenocarcinoma incidence and mortality in 2020 and projections to 2040: new estimates from GLOBOCAN 2020. Gastroenterology 2022 Sep;163(3):649-658. [doi: 10.1053/j.gastro.2022.05.054] [Medline: 35671803]
3. Li Y, Xu J, Gu Y, Sun X, Dong H, Chen C. The disease and economic burdens of esophageal cancer in China from 2013 to 2030: dynamic cohort modeling study. JMIR Public Health Surveill 2022 Mar 02;8(3):e33191 [FREE Full text] [doi: 10.2196/33191] [Medline: 34963658]
4. He Z, Ke Y. Precision screening for esophageal squamous cell carcinoma in China. Chin J Cancer Res 2020 Dec 31;32(6):673-682 [FREE Full text] [doi: 10.21147/j.issn.1000-9604.2020.06.01] [Medline: 33446991]
5. Zheng R, Zhang S, Zeng H, Wang S, Sun K, Chen R, et al. Cancer incidence and mortality in China, 2016. J Natl Cancer Center 2022 Mar 23;2(1):1-9. [doi: 10.1016/j.jncc.2022.02.002] [Medline: 36944542]
6. Chen R, Liu Y, Song G, Li B, Zhao D, Hua Z, et al. Effectiveness of one-time endoscopic screening programme in prevention of upper gastrointestinal cancer in China: a multicentre population-based cohort study. Gut 2021 Feb 02;70(2):251-260 [FREE Full text] [doi: 10.1136/gutjnl-2019-320200] [Medline: 32241902]
7. Cao M, Li H, Sun D, He S, Yu Y, Li J, et al. Cancer screening in China: The current status, challenges, and suggestions. Cancer Lett 2021 May 28;506:120-127. [doi: 10.1016/j.canlet.2021.02.017] [Medline: 33684533]
8. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
9. Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. J Clin Epidemiol 2012 Sep;65(9):934-939. [doi: 10.1016/j.jclinepi.2011.11.014] [Medline: 22742910]
10. Liu M, He Z, Guo C, Xu R, Li F, Ning T, et al. Effectiveness of intensive endoscopic screening for esophageal cancer in China: a community-based study. Am J Epidemiol 2019 Apr 01;188(4):776-784. [doi: 10.1093/aje/kwy291] [Medline: 30608546]
11. Zeng H, Sun K, Cao M, Zheng R, Sun X, Liu S, et al. Initial results from a multi-center population-based cluster randomized trial of esophageal and gastric cancer screening in China. BMC Gastroenterol 2020 Nov 24;20(1):398 [FREE Full text] [doi: 10.1186/s12876-020-01517-3] [Medline: 33228549]
12. Lu X, Hou J, Chen Z, Guo C, Li S, Bai W. Endoscopic screening of esophageal cancer in high risk areas. China J Cancer Prev Treat 2003;10(9):900-903. [doi: 10.3969/j.issn.1673-5269.2003.09.002]
13. Wen DG, Wang SJ, Zhang LW, Zhou W, Yu WF, Wang XL. Natural history of esophageal and gastric cardia precursor by repetitive endoscope screening with 425 adults in a high-risk area in China. Cancer Epidemiol 2009 Aug;33(2):108-112. [doi: 10.1016/j.canep.2009.06.002] [Medline: 19679056]
14. Zhang Y, Wang X, Zhang M, Ma Z, Cao L, Shi Y. Analysis of 5-year results of endoscopic iodine staining screening for esophageal/cardia cancer in areas with high incidence of esophageal cancer. China Cancer 2012;21(1):32-34 [FREE Full text]
15. Wang Z, Zhu D, Tan R, Song Y, Zhang Y, Yang F. Analysis of the screening results of esophageal cancer in Bazhong, Sichuan from 2010 to 2013. J Cancer Control Treat 2014;27(2):81-84 [FREE Full text] [doi: 10.12677/qrb.2014.11004]
16. Wei W, Chen Z, He Y, Feng H, Hou J, Lin D, et al. Long-term follow-up of a community assignment, one-time endoscopic screening study of esophageal cancer in China. J Clin Oncol 2015 Jun 10;33(17):1951-1957 [FREE Full text] [doi: 10.1200/JCO.2014.58.0423] [Medline: 25940715]
17. Wang M. A prospective cohort study of the distribution and the progress rate of precancerous lesions of esophageal squamous cell carcinoma in high-risk areas of China. Master's thesis, Tsinghua University School of Medicine Peking Union Medical College Chinese Academy of Medical Sciences. Wanfang Data. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkzMDc1MzQ5GghhYmQzMnUxbA%3D%3D [accessed 2022-05-31]
18. Jing Y, Yao Y, Shuai P, Hou W, He X, He Q. Analysis of the esophageal cancer screening for a high risk population of 14,000 40–69 years-old residents in Nanbu County, Sichuan Province. J Cancer Control Treat 2017;30:49-52. [doi: 10.3969/j.issn.1674-0904.2017.01.010]
19. Zhang Z, Wu Z, Lu L, Fan P, Zhao G, Liu J. Analysis of the upper gastrointestinal cancer screening and follow-up results in Liangzhou District of Wuwei City from 2009 to 2017. Chin J Cancer Prev Treat 2019;26(23):1750-1755. [doi: 10.16073/j.cnki.cjcpt.2019.23.02]
20. Zhao T, Li Y, Chen R, Zhang N, Wang J, Wei W. Long-term effects of endoscopic screening for upper gastrointestinal cancer in Feicheng. Chin J Cancer Prev Treat 2020;27(18):1470-1475. [doi: 10.16073/j.cnki.cjcpt.2020.18.06]
21. Li J, Gu J, He Y, Yuan Z, He Q, Xie S. Incidence of upper gastrointestinal cancer and endoscopy screening results in Yanting County, Sichuan Province. China Cancer 2021;30(5):346-351. [doi: 10.11735/j.issn.1004-0242.2021.05.A005]

22. Li J, Li H, Zeng H, Zheng R, Cao M, Sun D, et al. Trends in high-risk rates and screening rates for the population-based cancer screening program on esophageal, stomach and liver cancer in China, 2010-2016. J Natl Cancer Cent 2021 Sep;1(3):101-107. [doi: 10.1016/j.jncc.2021.05.001]
23. Wang W, Li H, Zhu C, Sun X, Zhang M, Wang Y. Cancer risk assessment and screening results analysis of residents in Zhejiang Province from 2014 to 2015. Zhe Jiang Yi Xue 2016;38(22):1795-1798 [FREE Full text]
24. Xu C, Zhou S, Li M, Cao H. Evaluation of the screening effect of two-step screening of upper gastrointestinal cancer in with low incidence areas. Jiangsu J Prev Med 2017;28(01):54-55. [doi: 10.12677/qrb.2014.11004]
25. Xu H, Wang Q. Analysis of cancer risk assessment and screening results of urban residents in Nantong City from 2015 to 2016. Modern Prevent Med 2017;44(20):3724-3726 [FREE Full text]
26. Zhu J, Gu X, Song S, Yao F, Zhou T, Zhu L. Analysis of 2015-2016 population screen of high cancer risk urban residents in Urumqi city. J Practical Oncol 2017;32(04):367-370. [doi: 10.13267/j.cnki.syzlzz.2017.04.019]
27. Liu Y, Kong W, Luo X, Dong D, Chen H, Li N. Analysis of early diagnosis and treatment of cancer in urban populations of Xuzhou city from 2014 to 2016. Jiangsu J Prev Med 2018;4:382-385. [doi: 10.12677/qrb.2014.11004]
28. Sun G, Zhao E, Xie H, Zhu J, Wan L. Analysis on the result of early detection and treatment of cancer in 2014 Urumqi. Chin Primary Health Care 2018;32(3):57-60. [doi: 10.3969/j.issn.1001-568X.2018.03.0021]
29. Lin Y, Ma J, Zhang Q, Lu Y, Zhang L, Zhang X. Analysis of upper gastrointestinal cancer screening results in Kunming from 2015 to 2018. China Cancer 2019;28(6):411-416. [doi: 10.11735/j.issn.1004-0242.2019.06.A003]
30. Xiao H, Yan S, Xu K, Zou Y, Shi Z, Zhu S. Analysis of cancer screening program in Changsha urban area from 2012 to 2018. China Cancer 2019;28(11):807-815. [doi: 10.11735/j.issn.1004-0242.2019.11.A001]
31. Guo L, Zhang S, Liu S, Zheng L, Chen Q, Cao X, et al. Determinants of participation and detection rate of upper gastrointestinal cancer from population-based screening program in China. Cancer Med 2019 Nov;8(16):7098-7107 [FREE Full text] [doi: 10.1002/cam4.2578] [Medline: 31560836]
32. Liu Y, Zuo T, Yu L, Wu Y, Dong S, Luo Y. Analysis of the results on cancer risk assessment and screening of urban residents in Shenyang. Chin J Public Health 2020;36(1):1-4. [doi: 10.11847/zgggws1127286]
33. Shi J, Liang D, Xia C, Chen S, Gao W, Wang J. Analysis of upper gastrointestinal cancer screening results in urban areas of Hebei Province, 2018-2019. China Cancer 2020;29(6):419-424. [doi: 10.11735/j.issn.1004-0242.2020.06.A003]
34. Zhao R, Lin Z, Yu W, Lei L, Liu Z, Xu Y. Analysis of the prevalence of five cancer risks among permanent residents aged 40-74 years old in Baoan District,Shenz-hen,2017-2019. Pract Oncol J 2020;34(4):315-320. [doi: 10.11904/j.issn.1002-3070.2020.04.005]
35. Zhu Z, Cao J, Yu J, Li Q, Yang Z, Ge L. Assessment of cancer risk and analysis of screening results of cancer screening program in urban Guangxi from 2014 to 2017. Chin J of Oncol Prev and Treat 2020;12(6):681-685. [doi: 10.3969/j.issn.1674-5671.2020.06.15]
36. Lei R, Zhang M, Wang L, Fang Z, Wang Y, Li H. Analysis of screening results of urban cancer early diagnosis and treatment program in Quzhou city, Zhejiang province, 2014?2018. Chin J of Oncol Prev and Treat 2021;13(6):618-623. [doi: 10.3969/j.issn.1674-5671.2021.06.08]
37. Tuo J, Zhang Z, He X, Jiang W, Feng J, Jin X. Analysis of the screening results of the urban cancer early diagnosis and treatment project in Hubei province from 2018 to 2019. Pract Oncol J 2021;35(4):297-301. [doi: 10.11904/j.issn.1002-3070.2021.04.002]
38. Xiong W, Chen H, Feng X, Fan Y, Sun Y, Zhu X. Assessment and screening results of risk of upper gastrointestinal cancer in urban inhabitants located in Nanchang from 2018 to 2019. Chin J Clin Oncol Rehabil 2021;28(11):1316-1320. [doi: 10.13455/j.cnki.cjcor.2021.11.10]
39. Zhang Y, Gao Q, Cui W, Cao L, Guo X, Ma Z. Analysis of cancer screening program in Shanxi urban area from 2014 to 2018. China Cancer 2021;30(02):131-136. [doi: 10.11735/j.issn.1004-0242.2021.02.A005]
40. Wang X, Zhang Y, Yuan F, Wang F, Liu K. Analysis on 5073 persons of upper gastrointestinal endoscopic screening in Yangcheng county Shanxi province. Cancer Research and Clin 2009;21(9):648-642. [doi: 10.3760/cma.j.issn.1006-9801.2009.09.028]
41. Yin C. The assessment of efficacy of endoscopic staining examination with iodine solution voluntary in high incidence areas of esophageal cancer in Feicheng. Master's thesis, Shandong Academy of Medical Sciences. Wanfang Data. 2009. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkxNTc1NTc4GghybHNwajZycQ%3D%3D [accessed 2022-05-31]
42. Chen L, Zhang S, Yang L, Liu Y. Preliminary report on screening of esophagus/cardia cancer in Gansu province. China Cancer 2011;20(11):798-800 [FREE Full text]
43. Zou W. Endoscopic screening for esophageal and cardia cancer in high-incidence areas and a cohort study of natural history of precancerous lesions. Master's thesis, Hebei Medical University. Wanfang Data. 2011. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkxOTAwMzk5GghtaHVvdmV2dw%3D%3D [accessed 2022-05-31]
44. Guo S. Analysis of population and endoscopic characteristics of synchronous multiple primary early cancers and precancerous lesions of esophagus and cardia in high-incidence areas.Hebei Medical University Master's thesis,. Wanfang Data. 2013. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkyMzM3MzAzGgg0Mjc2eWs2Zw%3D%3D [accessed 2022-05-31]

45. Yang L, Ma L, Yu S. Investigation and analysis of early diagnosis and treatment project of upper gastrointestinal cancer in Tang County, Hebei Province. Ke Ji Feng 2014;18:176. [doi: 10.3969/j.issn.1671-7341.2014.18.156]
46. Zhang Z, Wu Z, Lu L, Fan P, Zhao G, Liu J. Analysis on screening for upper gastrointestinal cancer in Liangzhou District, Wuwei City, Gansu Province, 2009-2012. China Cancer 2014;23(9):743-747. [doi: 10.11735/j.issn.1004-0242.2014.09.A008]
47. Du L, Wang M, Wu W, Li Y, Xu S, Mao W. Analysis of screening results of early diagnosis and treatment of upper gastrointestinal cancer in rural areas of Zhejiang Province. Zhejiang Clin Med J 2015;17(1):143-144 [FREE Full text]
48. Liang S, Li K, Gong J, Wang J, Ma H, Wang G. Results of the endoscopic screening program of esophageal and gastric cardia cancers using iodine staining in Feicheng, Shandong Province, from 2006 to 2012. Chin J Oncol 2015;7:549-553. [doi: 10.3760/cma.j.issn.0253-3766.2015.07.015]
49. Lv CH. Screening results of early esophageal cancer in Wuning County, Jiangxi Province in recent 4 years. Contemp Med 2015;21(26):157-158. [doi: 10.1353/pbm.2015.0001]
50. Tuo J, Wei S, Li G, Zhang Y, Zhang M, Xu F. An analysis the results of early detection and early treatment of esophageal/cardia cancer (upper gastrointestinal cancer) in different regions of Hubei Province. China Cancer 2015;24(8):653-656. [doi: 10.11735/j.issn.1004-0242.2015.08.A006]
51. Zhang M, Li X, Zhang S, Chen Q, Wang F, Zhang Y. Analysis of effect of screening of esophageal cancer in 12 cities and counties of Henan province. Chin J Prev Med 2015;49(10):879-882. [doi: 10.3760/cma.j.issn.0253-9624.2015.10.008]
52. Zheng X, Mao X, Xu K, Lü L, Peng X, Wang M, et al. Massive endoscopic screening for esophageal and gastric cancers in a high-risk area of China. PLoS One 2015;10(12):e0145097 [FREE Full text] [doi: 10.1371/journal.pone.0145097] [Medline: 26699332]
53. Li Z, Guo H. Endoscopic screening for early carcinoma of the upper gastrointestinal tract in 6,000 cases of natural population. Modern Digest Interv 2016;21(1):32-34. [doi: 10.3969/j.issn.1672-2159.2016.01.010]
54. Qian F, Wang H, Ge X, Li J, Liu X, Jia J. Analysis of 23513 cases of precancerous lesions of esophagus screened by gastroscopy in Changyuan County, Henan Province. Contemp Med 2016;22(21):160-161. [doi: 10.3969/j.issn.1009-4393.2016.21.109]
55. He X, Dong L, Zhao Q, Cao L, Tang X, Li S. Analysis of effect of screening of esophageal cancer of 40-69 year old residents in Langzhong city of Sichuan province. Yi Yao Qian Yan 2017;7(35):362-363. [doi: 10.3969/j.issn.2095-1752.2017.35.322]
56. He Y, Zhang W, Liu Y. Analysis of endoscopic screening for upper digestive tract cancer in Gaotai County, 2011-2015. China Cancer 2017;26(6):447-451. [doi: 10.11735/j.issn.1004-0242.2017.06.A007]
57. Wang X, Li B, Bao Y, Wang Y, Wang A, Qiao L. Efficacy of esophageal cancer screening in high risk population: results of 105 561 subjects in Sichuan province. Zhonghua Zhong Liu Za Zhi 2017 Jan 23;39(1):67-71. [doi: 10.3760/cma.j.issn.0253-3766.2017.01.013] [Medline: 28104037]
58. Zhang N, Ma H, Sun Y, Xu C, Gu J, Gong J. Analysis of social screening results for esophageal cancer among rural residents in Shandong Province from 2013 to 2016. Chin J Cancer Prevent Treat 2017;24(5):287-290 [FREE Full text]
59. Ding G. Analysis of early diagnosis and treatment screening and its influencing factors of upper gastrointestinal cancer among residents in Wuwei City, Gansu Province. Master's thesis, Gansu University of Traditional Chinese Medicine. Wanfang Data. 2018. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCUQwMTU1ODA1OBoINHRvcDR0eGM%3D [accessed 2022-05-31]
60. Tang X, Cao L, Long S, Bai Y, He X, Dong L. Results of the endoscopic screening program of esophageal using iodine staining in Langzhong, Sichuan Province, from 2012 to 2016. J Cancer Control Treat 2018;31(1):41-45. [doi: 10.3969/j.issn.1674-0904.2018.01.007]
61. Liang F, Qiao X, Wang C, Han C, Yang B, Yang S. Endoscopic iodine staining for clinical investigation of esophageal disease screening in Huai'an of Jiangsu Province with a survey of 1983 cases. Chin J Digest Med Imageol 2019;9(1):17-21. [doi: 10.3877/cma.j.issn.2095-2015.2019.01.005]
62. Wang W, Chen S, Chen J, Wei J, Li A, Li X. Screening results of upper gastrointestinal cancer in 27,994 normal people. J Canc Control Treat 2019;32(11):988-992. [doi: 10.3969/j.issn.1674-0904.2019.11.007]
63. He Z, Liu Z, Liu M, Guo C, Xu R, Li F, et al. Efficacy of endoscopic screening for esophageal cancer in China (ESECC): design and preliminary results of a population-based randomised controlled trial. Gut 2019 Feb 06;68(2):198-206. [doi: 10.1136/gutjnl-2017-315520] [Medline: 29306867]
64. Feng X, Hua Z, Qian D, Zhou Q, Shi A, Wei W, et al. Efficacy of esophageal cancer screening program on population at high risk: a survey carried out in people aged 40-69 years in Yangzhong, Jiangsu province. Zhonghua Liu Xing Bing Xue Za Zhi 2020 Jun 10;41(6):908-912. [doi: 10.3760/cma.j.cn112338-20190606-00407] [Medline: 32564558]
65. Jia S, Xu W, Shen X, Liu R, Zha Z, Wang D. Analysis of esophagus cancer screening results by using idone staining and endoscopy in Anhui Province, China. Chin J Cancer Prev Treat 2020;27(3):165-171. [doi: 10.16073/j.cnki.cjcpt.2020.03.01]
66. Zhuang Y, Zhu L, Zhao P, Tang L, Xu L. Analysis of screening results of upper gastrointestinal cancer early diagnosis and early treatment project sites among rural residents in Guizhou Province in 2018. Trace Elements Health Res 2020;37(6):55-56 [FREE Full text]
67. Cao X, Zhang S, Wang F, Chen Q, Guo L, Liu S, et al. Analysis of the effects of esophageal cancer screening in Henan rural areas with cancer screening program, 2014-2018. Zhonghua Yu Fang Yi Xue Za Zhi 2021 Feb 06;55(2):184-188. [doi: 10.3760/cma.j.cn112150-20200320-00398] [Medline: 34645177]

68. Zhu L, Hu Y, Zhao F, Xu L, Yu Z, Liu T. Screening of upper digestive tract cancer in rural areas of Guizhou Province, 2009-2019. China Cancer 2022;31(11):903-908. [doi: 10.11735/j.issn.1004-0242.2022.11.A009]
69. Cao X, Guo L, Liu S, Li X, Chen Q, Quan P. Pathological distribution in esophageal cancer screening population in non-high incidence areas. China Cancer 2019;28(10):731-737. [doi: 10.11735/j.issn.1004-0242.2019.10.A003]
70. Ma H, Shi S, Chen W, Lu P, Li B, Xu Y. Esophageal cancer screening in five counties of southwest Shandong Province from 2007 to 2015. China Cancer 2019;28(10):738-742. [doi: 10.11735/j.issn.1004-0242.2019.10.A004]
71. Sun Q, Pan E, Sun Z, Wen J, Liao D, Yang H. Analysis of the endoscopic screening results in high incidence areas of esophageal cancer in Huai'an from 2009 to 2017. Chin J Cancer Prev Treat 2020;27(4):251-255. [doi: 10.16073/j.cnki.cjcpt.2020.04.01]
72. Zhou T, Gu X, Zhu J, Zhu L. Endoscopic screening of upper digestive tract cancer among high-risk population in Urumqi. China Cancer 2017;26(10):781-785 [FREE Full text] [doi: 10.12677/qrb.2014.11004]
73. Huang Y. Preliminary analysis of early diagnosis and treatment of upper gastrointestinal cancer in Guangxi and clinical application of ME-NBI in screening. Master's thesis, Guangxi Medical University. WanFang Data. 2019. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkzNTU2MDU1Gghwa2VseDhzbQ%3D%3D [accessed 2022-05-31]
74. Zhang J, Wei D, Yin H, Ma Y. Analysis of esophageal cancer and precancerous lesions in high-risk groups of esophageal cancer. J Taishan Med College 2021;42(1):10-12. [doi: 10.3969/j.issn.1004-7115.2021.01.003]
75. Li H, Sun D, Cao M, He S, Zheng Y, Yu X, et al. Risk prediction models for esophageal cancer: a systematic review and critical appraisal. Cancer Med 2021 Oct 20;10(20):7265-7276 [FREE Full text] [doi: 10.1002/cam4.4226] [Medline: 34414682]
76. Xia R, Zeng H, Liu W, Xie L, Shen M, Li P, et al. Estimated cost-effectiveness of endoscopic screening for upper gastrointestinal tract cancer in high-risk areas in China. JAMA Netw Open 2021 Aug 02;4(8):e2121403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.21403] [Medline: 34402889]
77. Xia R, Li H, Shi J, Liu W, Cao M, Sun D, et al. Cost-effectiveness of risk-stratified endoscopic screening for esophageal cancer in high-risk areas of China: a modeling study. Gastrointest Endosc 2022 Feb;95(2):225-235. [doi: 10.1016/j.gie.2021.08.008] [Medline: 34418461]
78. He S, Li H, Cao M, Sun D, Yang F, Yan X, et al. Geographic, demographic, and socioeconomic disparities and factors associated with cancer literacy in China: National Cross-sectional Study. JMIR Public Health Surveill 2023 Feb 17;9:e43541 [FREE Full text] [doi: 10.2196/43541] [Medline: 36800218]
79. Jia S, Li H, Zeng H, Zheng R, Li J, Shi J, et al. Association of cancer prevention awareness with esophageal cancer screening participation rates: results from a population-based cancer screening program in rural China. Chin J Cancer Res 2019 Aug;31(4):601-608 [FREE Full text] [doi: 10.21147/j.issn.1000-9604.2019.04.04] [Medline: 31564803]
80. Cao W, Tan F, Liu K, Wu Z, Wang F, Yu Y, et al. Uptake of lung cancer screening with low-dose computed tomography in China: a multi-centre population-based study. EClinicalMedicine 2022 Oct;52:101594 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101594] [Medline: 35923428]
81. He J, Chen W, Li Z, Li N, Ren J, Tian J. China guideline for the screening, early detection and early treatment of esophageal cancer (2022, Beijing). Chin J Oncol 2022;44(6):491-522. [doi: 10.3760/cma.j.cn115610-20220517-00282]
82. Wang G, Abnet CC, Shen Q, Lewin KJ, Sun XD, Roth MJ, et al. Histological precursors of oesophageal squamous cell carcinoma: results from a 13 year prospective follow up study in a high risk population. Gut 2005 Feb 01;54(2):187-192 [FREE Full text] [doi: 10.1136/gut.2004.046631] [Medline: 15647178]
83. Wei W, Hao C, Guan C, Song G, Wang M, Zhao D, et al. Esophageal histological precursor lesions and subsequent 8.5-year cancer risk in a population-based prospective study in China. Am J Gastroenterol 2020 Jul 28;115(7):1036-1044 [FREE Full text] [doi: 10.14309/ajg.0000000000000640] [Medline: 32618654]
84. Yang J, Wei WQ, Niu J, Liu ZC, Yang CX, Qiao YL. Cost-benefit analysis of esophageal cancer endoscopic screening in high-risk areas of China. World J Gastroenterol 2012 May 28;18(20):2493-2501 [FREE Full text] [doi: 10.3748/wjg.v18.i20.2493] [Medline: 22654446]

Abbreviations

EC: esophageal cancer
ESCC: esophageal squamous cell carcinoma
LGIN: low-grade intraepithelial neoplasia
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RSES: risk-stratified endoscopic screening
UGI: upper gastrointestinal

Edited by A Mavragani, T Sanchez; submitted 30.01.23; peer-reviewed by S Zhang, Z Su; comments to author 23.02.23; revised version received 08.03.23; accepted 18.04.23; published 01.06.23.

Please cite as:
Li H, Teng Y, Yan X, Cao M, Yang F, He S, Zhang S, Li Q, Xia C, Li K, Chen W
Profiles and Findings of Population-Based Esophageal Cancer Screening With Endoscopy in China: Systematic Review and Meta-analysis
JMIR Public Health Surveill 2023;9:e45360
URL: https://publichealth.jmir.org/2023/1/e45360
doi:10.2196/45360
PMID:37261899

Review

COVID-19 Contact Tracing Strategies During the First Wave of the Pandemic: Systematic Review of Published Studies

Anna Maria Vincenza Amicosante[1*], MSc, PhD; Annalisa Rosso[1,2*], MSc, MD, PhD; Fabio Bernardini[1], MA; Elisa Guglielmi[1], MA; Erica Eugeni[1], MA, PhD; Filippo Da Re[3], MA, MSc; Giovanni Baglio[1], MSc, MD

[1]Research and International Relations Unit, Italian National Agency for Regional Healthcare Services, Rome, Italy

[2]Department of Environmental and Prevention Sciences, University of Ferrara, Ferrara, Italy

[3]Regional Directorate of Prevention, Food Safety, Veterinary, Public Health-Veneto Region, Venice, Italy

[*]these authors contributed equally

Corresponding Author:
Giovanni Baglio, MSc, MD
Research and International Relations Unit
Italian National Agency for Regional Healthcare Services
Via Piemonte, 60
Rome, 00187
Italy
Phone: 39 06 427491
Email: baglio@agenas.it

Abstract

Background: Contact tracing (CT) represented one of the core activities for the prevention and control of COVID-19 in the early phase of the pandemic. Several guidance documents were developed by international public health agencies and national authorities on the organization of COVID-19 CT activities. While most research on CT focused on the use digital tools or relied on modelling techniques to estimate the efficacy of interventions, poor evidence is available on the real-world implementation of CT strategies and on the organizational models adopted during the initial phase of the emergency to set up CT activities.

Objective: We aimed to provide a comprehensive picture of the organizational aspects of CT activities during the first wave of the pandemic through the systematic identification and description of CT strategies used in different settings during the period from March to June 2020.

Methods: A systematic review of published studies describing organizational models of COVID-19 CT strategies developed in real-world settings was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement. PubMed, Embase, and Cochrane Library were searched. Studies not providing a description of the organizational aspects of CT strategies and studies reporting or modelling theoretical strategies or focusing on the description of digital technologies' properties were excluded. Quality of reporting was assessed by using the Template for Intervention Description and Replication Checklist for Population Health and Policy. We developed a narrative synthesis, using a conceptual framework to map the extracted studies broken down by target population.

Results: We retrieved a total of 1638 studies, of which 17 were included in the narrative synthesis; 7 studies targeted the general population and 10 studies described CT activities carried out in specific population subgroups. Our review identified some common elements across studies used to develop CT activities, including decentralization of CT activities, involvement of trained nonpublic health resources (eg, university students or civil servants), use of informatics tools for CT management, interagency collaboration, and community engagement. CT strategies implemented in the workplace envisaged a strong collaboration with occupational health services. Outreach activities were shown to increase CT efficiency in susceptible groups, such as people experiencing homelessness. Data on the effectiveness of CT strategies are scarce, with only few studies reporting on key performance indicators.

Conclusions: Despite the lack of systematically collected data on CT effectiveness, our findings can provide some indication for the future planning and development of CT strategies for infectious disease control, mainly in terms of coordination mechanisms and the use of human and technical resources needed for the rapid development of CT activities. Further research on the organizational models of CT strategies during the COVID-19 pandemic would be required to contribute to a more robust evidence-making process.

(JMIR Public Health Surveill 2023;9:e42678) doi:10.2196/42678

KEYWORDS

COVID-19; SARS-CoV-2; contact tracing; public health; infectious disease; disease control; community engagement; digital tool

Introduction

Background

Contact tracing (CT) has historically been one of the key public health response actions to control the outbreak of a novel virus, particularly in the absence of a vaccine [1]. As with other person-to-person infectious diseases, early case detection, identification, and management of contacts through CT was one of the top priorities for interrupting the chain of infection and controlling the spread of SARS-CoV-2 [2,3]. Before effective vaccines against SARS-CoV-2 became available in December 2020, CT was one of the few tools globally applied to prevent the spread of infection, combined with physical isolation of infected persons and their close contacts (the so-called quarantine), social distancing, and the use of protective devices in public places.

Several guidance documents were produced and disseminated by international public health agencies and national authorities on the organization of CT activities for COVID-19 control, including indications on the type of human and technical resources needed for the different steps of CT (case notification, contact identification, information, management, and surveillance of contacts) [3-10]. Reports on CT activities implemented in several countries suggest marked differences in the organizational models adopted in different settings based on the characteristics of the local health systems and structures, as well as the diagnostic and tracking capabilities [8,11]. Although adapting CT strategies according to the local epidemiological situation and available resources has been emphasized [1,12-14], little evidence is available on the actual implementation of CT activities in real-world settings. In fact, most published studies on CT for COVID-19 have focused on the combination of traditional CT and digital technologies and on cost-effectiveness, ethical concerns, and governance issues related to the use of digital tools [15-20]. Several studies, including systematic reviews, were aimed at estimating the effectiveness of CT strategies for SARS-CoV-2 control, but they mainly relied on modeling techniques [15,21], given the difficulties in measuring real-world effectiveness [22].

Aim of the Study

Further understanding the real-world implementation of CT strategies under different conditions, including measures adopted to scale-up activities, would be relevant to support future planning of CT activities for infectious disease control. Therefore, we conducted a systematic review of the literature to provide a comprehensive picture of the organizational aspects of CT activities during the first wave of the pandemic through the systematic identification and description of CT strategies used in different settings from March to June 2020. We decided to focus on the first wave of the pandemic, when CT represented one of the core public health activities for COVID-19 containment, to describe actions taken for the rapid set up of CT strategies and scale up of resources.

Methods

Study Protocol

We conducted a systematic review of studies describing organizational models of CT strategies for the surveillance and control of SARS-CoV-2 infection. This systematic review was registered in PROSPERO (International Prospective Register of Systematic Reviews; CRD42021279172). The review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [23] and the Template for Intervention Description and Replication Checklist for Population Health and Policy (TIDieR-PHP) [24].

Literature Search Strategy

A preliminary exploratory search was carried out restricting the research field to systematic reviews and meta-analyses only, which did not return any noteworthy results. Subsequently, a more specific research strategy was developed to identify primary studies, adapted for each database, using both Medical Subject Headings (MeSH) terms and free text keywords in the title and abstract fields. We searched relevant databases including PubMed, Embase and Cochrane Library from January 1, 2020, to July 31, 2021, for published studies in Italian and English with the terms "contact tracing," "contact investigation," "case finding," "case detect*," "contact examin*," "contact screen*," "COVID-19," "coronavirus," and "SARS-COV-2," with no limitations on study design. The complete search strategy is detailed in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

Studies were included if they provided a description of the organizational aspects of real-world CT strategies (eg, resources involved and activities conducted in each step of the CT process) applied during the first pandemic wave, with no restrictions on study type, setting, or population. The eligibility criteria (Textbox 1) for this review are described according to the PICOS (Population or Problem, Intervention, Comparison, Outcome, and Study Type) framework. Studies that did not provide a description of the organizational aspects of CT strategies or reported or modeled theoretical CT strategies were excluded. Comments, opinions, editorials, and news reports in which no original information was reported were excluded. In the initial screening phase, we classified studies describing CT strategies focusing on digital application tools such as exposure notification, Bluetooth, GPS, or big data management technologies. Because of their peculiarity, studies focusing on the description of the technological features of digital applications were evaluated separately and were therefore excluded from this review.

Textbox 1. Inclusion criteria for this review.

Population or problem

- Organization of contact tracing (CT) activities in populations hit by COVID-19 during the first pandemic wave.

Intervention

- Any real-world CT strategies for the control of COVID-19 during first wave not centered on exposure notification, Bluetooth, GPS, or big data management technologies.

Comparison

- Either no or any type of real-world CT strategy for the control of COVID-19, depending on whether comparative analyses are available in the included studies.

Outcome

- Identification of the main elements characterizing the organization of CT activities for COVID-19 control.
- Description of the main types of CT strategies for COVID-19 control.

Study type

- All types of studies.
- Papers published in peer-reviewed journals, in Italian or English, available in full-text.

Data Extraction and Study Quality

Search results were imported into a reference management database (EndNote 7.8 [Clarivate Plc]). Duplicate articles were removed, and the titles and abstracts of all the collected records were screened by 2 reviewers (AMVA and AR). Studies that clearly did not meet the inclusion criteria were excluded. Full texts of potentially relevant articles were retrieved and independently examined by the 2 researchers. The reference lists of retrieved articles were also searched to identify other potentially relevant studies. All excluded articles and reasons for exclusion were recorded (Multimedia Appendix 2) and any disputes between the 2 researchers were resolved through discussion.

A standardized data extraction file was developed, including the following information: main author, year of publication, country, study design, study period, epidemic phase, type of CT program (institutional level or local level), study population, study setting, activities carried out during the various steps of the CT process (case notification, contact identification, information, management, and surveillance), human and technical resources used, main features of the CT model, and quantitative results.

The quality of reporting was assessed by 2 reviewers (AMVA and AR) using the TIDieR-PHP checklist [24]. The checklist enables clear and comprehensive reporting of population health and policy interventions, providing 11 items to capture pertinent features of these interventions. Adherence to these 11 items was assessed in each of the included studies.

We developed a narrative synthesis using a conceptual framework to map the extracted studies broken down by the target population.

Results

Overview

We retrieved a total of 1638 studies. After duplicate removal and title abstract selection, 130 full texts were assessed, and 17 of them were included in the narrative synthesis [25-41] (Figure 1).

Studies focusing on the organizational models of CT apps and other digital tools were not included in this study and will be addressed in a different review. The main features of the included studies are summarized in Table 1.

All included papers were descriptive accounts of CT strategies implemented in various settings, except for a qualitative study [25]. A total of 8 studies were conducted in the United States [26,27,29-32,36,41], 3 in Asia [34,35,39], 3 in Europe [33,37,40], 2 in Africa [25,28], and 1 in Australia [38].

One study reported information on implemented CT strategies disaggregated for the first and second waves [33], and 1 study did not clearly report the timing when the CT model was implemented [35]. A total of 10 studies described organizational models implemented by national or local governments or public health agencies [25,26,28,29,31,37-41]. The remaining studies described strategies implemented locally in specific contexts [27,30,32-36]. In addition, 7 studies targeted the general population [25-31], and the remaining 10 studies described CT activities carried out in specific population subgroups (workers, travelers, and vulnerable populations) [32-41]. Further details are provided in Table S1 in Multimedia Appendix 3 [25-41].

Quantitative results of CT activities were seldom reported, and the available data were not comparable across the studies (Table S2 in Multimedia Appendix 3).

The results of the assessment of reporting quality are included in Multimedia Appendix 4 [25-41].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of included studies [23]. CT: contact tracing.

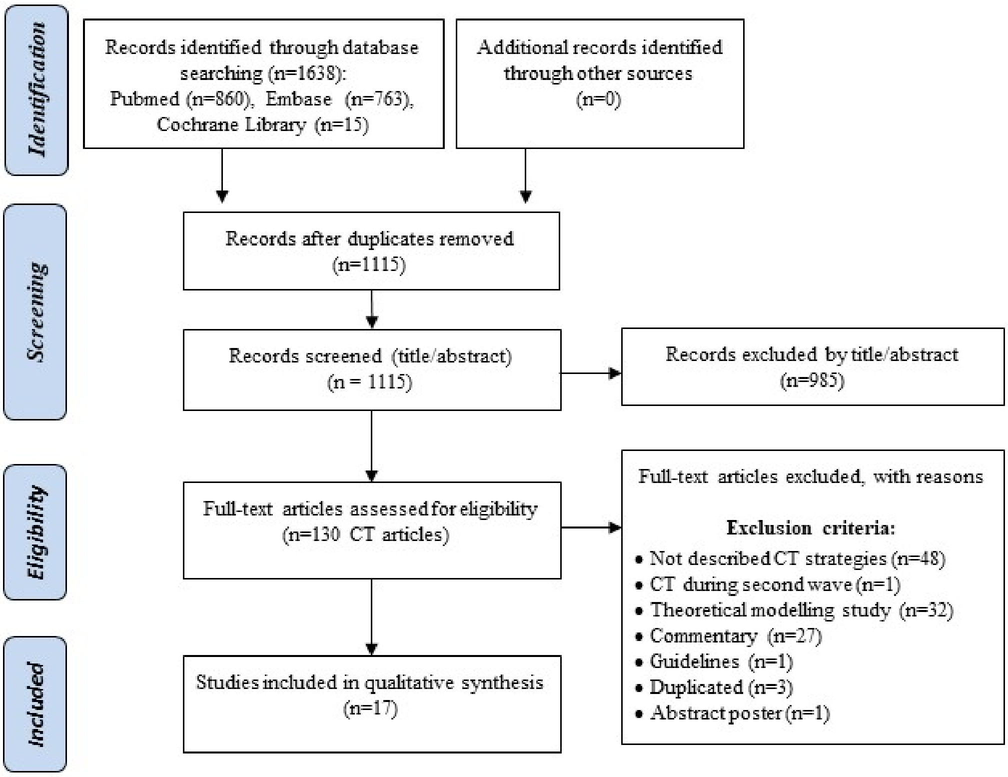


Table 1. Characteristics of included studies.

Study	Country	Epidemiological scenario	Target	Setting	Institutional level
Asiimwe et al [25], 2021	Africa–Ghana (Greater Accra Region)	First wave	General population	Regional territory	Government
Breeher et al [32], 2020	United States–Minnesota, Florida, and Arizona	First wave	Health care workers and patients	Hospital (Mayo Clinic)	Specific context
Clarke et al [40], 2020	Europe–Ireland	First wave	Detainees	Prison	Government
de Laval et al [37], 2021	Europe–France	First wave	Workers	Working environment (Creil Air Base–MSFAC[a])	Government
Draper et al [38], 2021	Australia–Northern Territory	First wave	Travelers entering Australia and general population	Airports, ports, and communities	Government
Fields et al [41], 2021	United States–Utah (Salt Lake County)	First wave	General population and PEH[b]	County territory and accommodation for PEH	Government
Hall et al [36], 2021	United States–Virginia	First wave	Workers and their external contacts	Work environment (US Navy Medicine, the Bureau of Medicine and Surgery, Falls Church)	Specific context
Zirbes et al [33], 2021	Europe–Germany	First and second wave	Patients, employees, and health care workers	Marburg University Hospital	Specific context
Kalyanaraman and Fraser [26], 2021	United States–County in Maryland	First wave	General population	County territory	Government
Koetter et al [27], 2021[c]	United States–Central Pennsylvania	First wave	General population, with CT[d] starting from hospital-diagnosed cases—Penn State College of Medicine	Territory of the region	Specific context
Mak et al [35], 2021	China–Hong Kong	Not clearly stated	Patients	Department of Ophthalmology—United Christian Hospital Hong Kong	Specific context
Mueller et al [28], 2020	Africa–Lagos State	First wave	General population	Territory of 5 subareas with the highest number of cases	Government
Niccolai [29], 2020	United States–Connecticut (New Haven)	First wave	General population and university community	City territory and university (Yale University)	Government
Pelton et al [30], 2021[c]	United States–central Pennsylvania	First wave	General population, with CT starting from hospital-diagnosed cases—Penn State College of Medicine	Territory of the region	Specific context
Quach et al [39], 2021	Vietnam–Hanoi	First wave	Travelers entering Vietnam	Flight	Government
Reid et al [31], 2021	United States–City of San Francisco	First wave	General population	City territory	Government
Wong et al [34], 2020	China–Hong Kong	First wave	Patients and health care workers	Queen Elizabeth Hospital	Specific context

[a] MSFAC: Military Support Facility.

[b] PEH: people experiencing homelessness.

[c] The articles refer to the same study but with 2 different focuses: Koetter et al [27] described the organizational model of contact tracing (CT), whereas Pelton et al [30] focused on key performance indicators.

[d] CT: contact tracing.

General Population

All studies targeting the general population have described CT models implemented at the local level (regional territories, counties or cities). Two studies focused on cases diagnosed within a university hospital, the Penn State College of Medicine [27,30], while 1 study described CT activities carried out within Yale University's community, in addition to those targeting the city of New Haven's general population [29].

Not all studies provided details on case notifications. Where described, case notification envisaged the communication of new COVID-19 cases from test laboratories or test centers to CT teams [27,28,30] or the use of infectious disease notification systems, either with automatic notification [26] or through an active search of new cases conducted by an epidemiologist [29].

All studies reported the identification of close contacts by telephone interviews with cases, except for the model described by Mueller et al [28] in the Lagos area, where contacts were line-listed for possible follow-up at the point of sample collection. A total of 3 studies specified that the interviews with cases were conducted using defined questionnaires or forms [25,27,29] developed by epidemiologists. In all cases, the contact list was entered into an electronic database or software.

Contacts were notified of their exposure to a COVID-19 case via telephone in all models. In 2 cases [26,31], the phone call was accompanied by a letter, email, or SMS text messaging notifying the exposure. Koetter et al [27] reported that the CT team from the Penn State College of Medicine used a premade phone script designed with assistance from epidemiologists treating infectious diseases to provide information on exposure and quarantine.

The management and monitoring of contacts during quarantine (14 days in all studies) was carried out mainly using telephone [25-28]. As an alternative to telephone monitoring, the model implemented in the Lagos area also provides home visits by nurses to assess symptoms and take swabs of symptomatic contacts [28]. The students of Penn State College of Medicine [27,30] monitored the symptoms of the cases contacted through a questionnaire sent by email from the CT management software (REDCap [Research Electronic Data Capture], Vanderbilt University) with automatic feedback to the CT team. In the model described by Reid et al [31], screening of symptoms during quarantine occurred via SMS text messaging sent automatically from the COVID-19 tracking application CommCare (Dimagi Inc) and subsequent feedback to the CT team.

The CT models retrieved from the literature used different types of human resources. In some cases, only health workers were involved [25,26], providing different roles depending on their professional background. In the model implemented in the Accra Region [25], clinicians based at the regional level informed contacts of exposure and supervised the work of community nurses engaged in contact monitoring, whereas case investigations were conducted by epidemiologists and disease control officers at the local level. The model described by Kalyanaraman et al [26] involved CT teams composed of nurses responsible for investigating cases and informing contacts, as well as health assistants and runners responsible for monitoring contacts, under the supervision of an epidemiologist. Some studies have reported the involvement of medical and health profession students in CT activities. At the Penn State College of Medicine, medical students were involved in the CT of cases diagnosed within the hospital, being employed either in the management of cases (case teams) or contacts (contact teams) [27]. Yale University health care professional students were involved on a voluntary basis in a CT program aimed at both campus and off-campus populations in collaboration with the local health department, which provided web-based training for volunteers [29]. The University of California San Francisco also involved medical students in a CT program developed with the San Francisco Department of Public Health, and public health experts from the 2 bodies coordinated the activities and trained the CT team, which also included retired doctors, librarians, and other civil servants [31].

All the models described used software or web platforms for case and contact management. Most studies used open-source platforms (eg, Surveillance Outbreak Response Management and Analysis System, SORMAS Foundation [25], and Open Data Kit, Jekyll and Minimal Mistakes [28]) or relied on existing platforms such as such as RedCap (Vanderblit University) [27,30] or Veoci Inc (Virtual Emergency Operations Center Software) [29]. The CT model developed by the students of University of California San Francisco and the experts from San Francisco Department of Health was initially partnered with Dimagi Inc to make use of their web-based COVID-19 tracking application, CommCare, and then transitioned to the digital platform, CalConnnect, to align San Francisco with most other jurisdictions in California [31].

Only 2 models provided the development of indicators that allowed for the measurement of the effectiveness of the interventions [30,31]. In both cases, the development of a CT model based on the rapid mobilization of human resources (including university students), team organization, and the use of digital platforms led to a reduction in the time required to complete CT activities. The study conducted at the Penn State College of Medicine showed a reduction in the test turnaround time from 21.8 to 2.3 days, also due to improvements in the testing capacity [30]. The study by Reid et al [31] showed a reduction in the time between contact registration and the first attempted contact from 5 days to 1 day over a 2-month period.

Special Populations

Studies describing CT models not addressed in the general population targeted the following population groups: workers, including health care workers (HCWs), military workers, and civilians working in military areas [32-37]; travelers [38,39]; and vulnerable populations, including inmates and homeless people [40,41]

Workers

Among studies carried out on workers, 4 studies focused on CT models developed in the health care setting [32-35] and 2 studies reported CT strategies addressed to civilians and military workers.

HCWs and Patients

Studies describing CT models addressing HCWs also reported CT activities in patients and contacts; 1 study focused on hospital patients only [35]. One study was carried out in the Mayo Clinic campuses, United States [32]; 1 was carried out in the University Hospital of Marburg, Germany [33]; and all the other studies were conducted in Hong Kong public health service hospitals.

Where indicated, the structure responsible for notifying cases among HCWs was the Occupational Health Services (OHS) [32] or the Infection Control and Prevention (ICP) service [33]. The procedure for case notification was not described in all studies. In the model described by Wong et al [34], a positive health worker was notified by the hospital's in-house laboratory, while in the model described by Breeher et al [32], case notification was made by multiple parties (infectious disease control teams, public health bodies, or self-reporting).

Contacts identification was carried out mainly by hospital internal teams: in the model described by Breeher et al [32] it was carried out by a designated exposure triage provider (ETP) composed of doctor and nurses, in collaboration with the OHS [32], while in other models, contact identification was conducted by the ICP service [33,35], sometimes in collaboration with the local health department [35]. Different approaches have been described for contact listing: some models relied on cases filling standard forms available on the web on the hospital intranet [32,33] and telephone interviews [32]; other models relied on information provided by the ward manager [34] or collected through a hospital software tracking the movement of patients within the hospital (UQ Web) [35]. Patients' contacts were identified through different digital information systems (the electronic medical record, EMR [32]; health information systems [33]; patient administration CT systems [34]; and UQ Web software [35]. All studies classified contacts according to the levels of exposure, risk, and subsequent definition of isolation measures.

Information on contacts was managed internally in all models as well: it was carried out either by hospital staff (eg, by the ETP, the exposure investigation team, and the nursing exposure investigation team in the model described by Breeher et al [32] or by the ICP in the model described by Zirbes et al [33]) or directly by cases using an intranet platform [33].

In the Mayo Clinic model, cases self-filled an internet form and were then interviewed by telephone to assess the level of risk, define isolation measures and prescribe diagnostic tests [32]. In the Marburg Hospital model, the ICP automatically notified HCWs identified as close contacts to define work restrictions and patients to begin isolation through the intranet platform. HCWs were sent to the clinic's testing centers, and results were available on the intranet, while the patients were cared by the local health authorities [33]. In the Hong Kong model described by Wong et al [34], quarantine for close contacts was arranged at a designated camp (staff) or in isolation rooms for patients with airway-transmitted infections, whereas casual contacts were subjected to medical surveillance.

Management of contacts included monitoring of symptoms, in some cases self-reported by the HCWs [32], and testing. Surveillance of symptoms duration ranged from 14 [33] to 28 [34] days from the last contact with the index case. The Marburg University Hospital model envisaged repeating antigenic tests (for health professionals) and polymerase chain reaction (PCR) tests (for patients) 3 times every 48 hours since the last contact with the index case [33], while Wong et al [34] reported that all contacts were monitored daily for temperature and symptoms for a 28 days (including the quarantine of 14 days).

The human resources used and their functions were not always clearly reported. The Mayo Clinic model detailed the inclusion of different professionals at the central level in Richmond and at the local campuses: at the central level, nurses and doctors—as part of the ETP—conducted case investigation, the exposure investigation team (composed of clinicians reallocated from other departments) conducted contacts' risk assessment, and administrative or relocated laboratory staff supported data collection; at the local level, nurses exposure investigation team established work restrictions and arranged testing for symptomatic cases [32]. In the model described by Zirbes et al [33], the ICP worker was the key figure in all the CT workflow (in December 2020, 3 ICP workers managed to monitor up to 1201 contacts).

Data on cases and contacts were collected and managed using ad hoc software [32,33,35]. In addition, software solutions were used to track cases and' movement of contacts within the hospital [35] or to allow cases and contacts to fill in standard forms, for contact listing, collecting information on exposure and symptoms [32,33].

Military Area Workers

Two studies addressing CT in the workplace described activities conducted in military areas: 1 at the headquarters of the US Navy Medicine in Virginia [36] and the other at the Creil Air Base in France [37]. Both the studies describe CT activities implemented further to notification of a COVID-19 case within the workplace: in the model described by Hall et al [36], the hospitalized index case self-notified SARS-CoV-2 infection to the office manager, whereas in the French study, the index cases (whose number is not reported) were diagnosed within the military base. In both studies, cases were notified to the local health authorities.

Military personnel [37] or the public health officer (PHO) embedded within the US Navy Medicine headquarters [36] were responsible for investigating the contact, drawing up the contact list, and identifying the possible source of infection. The local health department investigated contacts outside the workplace [36].

Information to contacts was based on their risk level [36,37]. In the study by de Laval [37], all personnel at the air base were classified as close contacts and the model focused on prompt testing of all new symptomatic people. In the US Navy study, the PHO sent information to contacts via mass email with confirmation of receipt, prescribing self-observation, or home quarantine for 14 days for low- and medium-risk contacts, respectively [36]. All personnel were instructed to contact the base health center [37] or their doctor upon onset of symptoms [36]. In the study performed by de Laval et al [37], tests were carried out directly in the base, in a specially set up field-sampling unit.

Surveillance of all identified contacts was performed daily via telephone [36,37], and in case of development of symptoms, PCR testing was required [37].

Human resources used were not specified in the study by de Laval et al [37]. In the US Navy Medicine model, all CT activities within the office were conducted by the PHO, who arranged and coordinated workspace sanitation and was responsible for investigating the case, informing and monitoring contacts, and for campaigning internal information. Administrative staff were responsible for the internal information campaign and infection control policies.

In all studies, CT activities were conducted via telephone, email, and interviews with standardized questionnaires; the PHO also made use of invitation lists, meeting attendance, and carpooling data [36].

Travelers

Two studies described CT strategies used for travelers: one focused on an index case identified on a cruise ship arriving in the Northern Territory of Australia [38] and the second focused on a CT of a flight arriving in Vietnam [39].

In both studies, the index cases were confirmed by a PCR test, but the notification procedure was not described.

In the case of a confirmed positive traveler, the identification of contacts started from the passenger list of the same flight or cruise. In the study performed by Draper et al [38], the passenger list was provided by the airline or the Australian Government Department of Health National Incident Room. In Vietnam, the passenger list was provided by the immigration office and the Civil Aviation Administration to the competent Provincial Center for Disease Control. At the provincial level, local health personnel worked with local authorities, social security departments, and local volunteers to contact passengers and identify their contacts [39].

Draper et al [38] specified that information to contacts was provided via telephone using a standard questionnaire by the contact tracer team (CTT), which also collected data on the time, place, and duration of contact and on COVID-19 symptoms [38]. In the study performed by Quach et al [39], local health personnel interviewed, tracked, tested, and arranged quarantine (for 14 days) in centralized structures for primary and secondary contacts or suggested self-quarantine at home [39].

Contact monitoring was performed randomly by compliance officers by sending daily SMS text messages to monitor the development of fever or respiratory symptoms and compliance with quarantine measures [38]. Contacts in centralized quarantine had their symptoms and temperature checked twice daily; swabs were collected after 3 to 5 days and on day 13 before exiting quarantine [39]. Accommodation, meals, and basic hygiene needs were provided by the Ministry of Health. Any contact who tested positive during centralized or home quarantine was transferred to a referral hospital for isolation and monitoring [39]. Both studies did not provide information on the composition of the CT teams.

Tools used for CT included the Telstra Integrated Messaging platform [39] for sending SMS text messages and a web-based epidemiology database (NetEpi) to collect information on close contacts [38].

Vulnerable Populations

Two studies described CT strategies in vulnerable populations. One study addressed CT of inmates within the Irish prison system [40], and the study performed by Fields et al [41] described CT strategies of people experiencing homelessness (PEH) hosted in quarantine or isolation facilities in Salt Lake County.

The notification of index cases was only described in the prison context, where prison staff notified the internal CTT of positive cases or symptomatic (experiencing cough and fever) inmates [40].

Contacts were identified through interviews with notified inmates and analysis of closed-circuit television footage [40]. For PEH, case investigations were conducted by dedicated staff visiting isolation and quarantine facilities using a standard form, initially in person and later using prepaid mobile phones or walkie talkies at the facilities.

Only Clarke et al [40] reported details on information to contacts. In the prison system model, the CTT informed inmates' contacts, arranged for their isolation, provided indications on home self-quarantine to family members and staff, and informed the public health agency.

Inmates' contacts quarantined inside prisons received clinical monitoring daily, whereas prison staff's contacts in home quarantine were monitored by the community contact management program; however, casual contacts were instructed to self-monitor in case of any COVID-19 symptoms [40]. The PEH were monitored by nurses who filled out a spreadsheet with medical and epidemiological information for each person housed in quarantine or isolation facilities [41].

The human resources involved in CT were described in the prison model. Each prison had a CTT made up of at least 4 people, including security chiefs, assistant chief officers, prison officers, psychologists, or clerical staff, whereas doctors and nurses were responsible for informing positive results and for clinical monitoring. The involved staff followed a training program developed by the National Infection Control Team, public health agency, and National Quality Improvement team [40]. In PHE facilities, nurses were responsible for collecting information on cases and contacts and for monitoring symptoms, and no information was provided on the profile of contact tracers [41].

All models envisaged the collection of case and contact information in local and sometimes centralized databases. Irish prisons collected information on cases and contacts using an Excel (Microsoft Corporation) spreadsheet and saved data in the penitentiary's IT system, which was then sent by secure email to the National Infection Control Team and the public health agency to be archived in the Health Service Executive central database [40]. Information on PEH was collected using the existing Utah National Electronic Disease Surveillance System or EpiTrax software [41].

Discussion

Principal Findings

A systematic review of the published literature on the organizational models of CT implemented during the first wave of the COVID-19 pandemic identified a limited number of studies. Despite the fact that published literature on the topic is scarce, some elements characterizing the setup of different CT programs and some recommendations to increase the efficacy of CT activities can be drawn.

A common feature of all studies was the decentralization of CT activities at the local level: CT was delegated to regions, counties, metropolitan areas, or specific settings such as hospitals, prisons, and communities, with the involvement of local call centers or human resources, as opposed to a centralized approach where CT is usually conducted in a national or central center [11]. However, in most models all steps of CT (case identification, identification, and monitoring of contacts) were implemented locally; in others, the identification of contacts and the overall management of CT activities took place at a central level [25,32,40]. Centrally managed models were mainly described in specific settings characterized by an internal CT tracing system (eg, the Mayo Clinic model in the study by Breeher et al [32] and the Irish Penitentiary Institutes model described by Clarke et al [40]). The model described by Asiimwe et al [25], implemented in the Great Accra Region of Ghana, followed the structure of the country's health system, with activities carried out in 3 tiers: national, regional, and district tiers with a strong focus on community care [42]. Evidence from European case studies suggests that the governance and organization of CT systems follow the structure of health systems, with a greater decentralization of activities in countries with regional management of health services [8]. According to the European Centre for Disease Control and Prevention (ECDC), the decentralization of CT systems represents a challenge for the collection of comprehensive and harmonized data on the volume and effectiveness of the interventions carried out that need to be addressed [8].

Several studies have described the involvement of human resources not belonging to public health agencies as contact tracers, such as university students from health faculties [27,29-31], hospital health workers, staff of CT programs implemented within health facilities [32], or representatives of other organizations, such as United Nations agencies (eg, study conducted by Mueller et [28] in Nigeria, with the involvement of the WHO [World Health Organization] and UNFPA [United Nations Population Fund] staff). As the number of cases increased, the main international public health agencies (Centers for Disease Control and Prevention [CDC], ECDC, and WHO) recommended the mobilization of nonpublic health staff, such as students, community health workers, volunteers, and civil servants, provided that they are adequately trained and supervised by public health bodies responsible for epidemic control [3,4,43,44]. Many of the documented experiences complied with this recommendation, showing its effectiveness in reducing the time needed to complete CT activities and in increasing the number of people contacted [27-32].

A common element emerging from all studies is the need to use IT software to support CT activities, as the number of cases and contacts to be monitored increased. The use of digital technology can overcome challenges related to incomplete contact identification, delays in the identification and isolation of cases, and notification and quarantine of contacts. Available evidence has been synthetized in some reviews, suggesting the effectiveness of digital technologies in supporting the control of the epidemic, but also underlying several normative, technical, and acceptance barriers to be addressed [45-47]. The WHO stressed the need to integrate such tools into comprehensive and adequately resourced CT strategies [4].

Different CT models were based on multiagency collaboration, with partnerships between public health agencies and other actors (eg, universities, United Nations agencies, community organizations, companies, and other institutions). Collaboration across different actors was not only aimed at the mobilization of human resources but also at the exchange of information; in the CT models aimed at travelers [38,39], collaboration between public health bodies and airline companies, port or airport management, and flight control bodies was essential to help trace persons who may have been exposed to SARS-CoV-2 during flights, as also recommended by international guidelines [48].

Three studies reported the use of outreach strategies to carry CT activities through home visits to closed settings (eg, prisons, shelters for PHE) [25,38,41]. This approach was used to address susceptible communities, which are difficult to reach through usual communication channels (telephone, email). The outreach approach was also used to reach out to the general population in the 2 models developed in African Countries (Ghana and Nigeria), where community worker programs are widely implemented [25,38]. The positive role of outreach activities conducted by community health workers in improving the effectiveness of CT interventions was previously highlighted in a systematic review conducted in the context of tuberculosis control [43]. The review indicated the potential value of outreach and community health workers in conducting case investigations in specific populations, such as drug addicts or homeless people. For these populations, it also indicated that location-based strategies of CT might lead to identification of an increased number of contacts, as also described in the model developed by Fields et al [41] for PHE.

CT programs aimed at health care professionals were characterized by the direct involvement of in-hospital OHS [32] and ICPs in CT activities [32,33]. In the Marburg University Hospital model described by Zirbes et al [33], ICP workers were the pivot of the CT process, accessing all information related to cases and contacts through a web-based platform. As also suggested by ECDC guidance on infection prevention and control and preparedness for COVID-19 in health care settings, potential mitigation measures, including CT, need to be addressed in collaboration with the existing OHS or health and safety committees [49]. As stressed by Breeher et al [32], CT in hospital settings may be extremely resource intensive. The use of electronic tools and organization into functional teams were shown to have improved efficiency and integration of standardized processes and made CT scale up feasible.

The 2 studies describing CT in the workplace, both conducted in the military setting, were based on an integration of the work conducted by public health agencies and staff responsible for occupational health and safety (in the US Navy model described by Hall et al [36]), with the presence of PHOs embedded in the workplace. CT activities in the workplace and collaboration between the different bodies responsible for safety at work, also by virtue of the existing legislation in various countries, are strongly recommended to limit the spread of the SARS-CoV-2, helping to reduce the need for close work activities [50-52]. International bodies also recommend promptly testing symptomatic workers and providing for their isolation, as described in the strategy developed by de Laval et al [37].

Finally, attention was paid to some of the models for the development of strategies aimed at promoting community engagement, which is recognized as an essential element for the success of CT programs [53]. Two studies have emphasized the importance of addressing the social needs of individuals placed in isolation and quarantine (eg, food, drugs, and connection to other services), also to increase compliance with isolation measures [26,27]. Another element facilitating community engagement was the involvement of contact tracers speaking other languages to ensure compliance of linguistic minorities with isolation and quarantine [27,31].

In general, one of the main findings of our literature review was the scarcity of published studies specifically aimed at describing the organizational aspects of CT activities. This could be because of difficulties in measuring and describing the effectiveness of CT interventions, making the topic of little interest in the scientific community [15]. Conducting a proper evaluation of the effectiveness of CT, testing, and isolation interventions is a complex task, as the type of evidence required is difficult to obtain (ie, randomization of interventions is not ethically acceptable); hence, available evidence is mainly based on modeling techniques or proxy data [2,22,54]. Most published studies on CT have focused on the development of mathematical models aimed at describing the factors determining the effectiveness of CT [55]. Though these studies can be useful in estimating the effectiveness of CT under different assumptions, they cannot provide indications on the organizational aspects of CT activities. Disseminating information on how CT strategies have been developed in different contexts, including data on human and technical resources, organization, coordination, and governance of activities, could provide useful elements for future planning and contribute to "evidence making" in this field. The absence of empirical data during the COVID-19 pandemic has challenged the traditional evidence-based approach, requiring other ways of generating and synthesizing evidence, where narrative studies on CT organizational models represent examples of evidence-making interventions [56,57].

Limitations

Only a few studies included in our review attempted to provide data on the effectiveness of the intervention, by identifying and calculating some key performance indicators to detect improvements in the effectiveness of CT interventions (mainly measured by a reduction in the time needed to complete CT activities or an increase in the number of cases and contacts reached) following some adjustments in the organization of activities (eg, increase in human resources and adoption of a digital tool). As also suggested in recent systematic reviews on the effectiveness of CT strategies for infectious disease control, more evidence is needed to understand how to optimize the effectiveness of CT across a range of settings and contexts, including large-scale comparative studies [15], informing "how, where, and when" to deploy CT most effectively [15]. Few data collected were not fully comparable across different studies. As suggested by Vogt et al [22], a universally agreed set of indicators is needed to allow for cross-system comparisons and to improve the performance of CT systems.

The main limitation of this systematic review was the focus on the first phase of the epidemic (first wave, March to June 2020) only. Our primary aim was to describe how CT strategies were first developed in different settings to respond to a novel virus outbreak. Nevertheless, collecting and synthesizing information on how CT strategies were adapted to changes in the epidemiological situation would be relevant to support the future planning of CT activities in response to viruses characterized by different modes of transmission, incubation time, and virulence. Furthermore, we did not consider changes that occurred with the availability of COVID-19 vaccines at the end of 2020, when most public health efforts were directed toward the organization of vaccination campaigns, and research mainly focused on measuring vaccine effectiveness.

Therefore, further evidence is needed to evaluate the adaptation of CT models to the different phases of the epidemic, including strategies adopted to scale up interventions, integration of traditional and digital methods, and adjustments of CT with regard to vaccination status.

Conclusions

In conclusion, our systematic review provides some preliminary evidence on organizational models of CT developed across various settings and contexts during the first wave of the COVID-19 pandemic. We identified some common elements in all strategies that allowed for the effective development of CT activities in the early phase of a novel virus epidemic, including decentralization of activities (case notification, identification, and management of contacts), involvement of nonpublic health trained resources, use of digital tools for CT management, interagency collaboration, adoption of strategies to increase community engagement, and aspects peculiar to each setting (eg, outreach, involvement of OHS and ICPs). Despite the lack of data on CT effectiveness, these findings can provide some indications for future planning and development of CT strategies for infectious disease control. Further research on the organizational models of CT strategies during the COVID-19 pandemic, including data on real-world effectiveness and on strategies developed during the following phases of the epidemic, would be needed to contribute to a more robust evidence-making process.

Acknowledgments

The review was conducted under the scope of the CCM (National Centre for Disease Prevention and Control) 2020 Projects "Modelli organizzativi per l'efficientamento delle attività di contact tracing" (Organizational models to improve the efficiency of contact tracing activities) and "Potenziamento del contact tracing attraverso il rafforzamento del ruolo dei Dipartimenti di Prevenzione sul territorio, anche in relazione alle condizioni di maggiore suscettibilità alle conseguenze sfavorevoli dell'infezione (fattori di rischio comportamentali, malattie croniche e multi-morbidità, età avanzata, ecc)" (Enhancement of contact tracing by strengthening the role of local Prevention Departments) funded by the Italian Ministry of Health.

The authors wish to thank all project partners for their contributions: Dr Pio Pellegrini, Local Health Authority of Frosinone; Dr Claudia Celebrini, Local Health Authority of Frosinone; Dr Alessandro Perrella, Hospital A Cardarelli, Napoli; Dr Massimo Majolo, Hospital A Cardarelli, Napoli; Dr Federica Michieletto, Veneto Region; Dr Elena Narne, Azienda Zero, Veneto Region; Dr Gisella Pitterm, Azienda Zero, Veneto Region; Dr Giulia Capodaglio, Azienda Zero, Veneto Region; and Dr Stefania Bellio, Azienda Zero, Veneto Region. They also thank Dr Stella Lanzi (Italian Agency for Regional Healthcare services) for her support in reviewing of the English language.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 19 KB - publichealth_v9i1e42678_app1.docx]

Multimedia Appendix 2
List of excluded studies.
[DOCX File , 38 KB - publichealth_v9i1e42678_app2.docx]

Multimedia Appendix 3
Summary of findings of included studies [25-41].
[DOCX File , 56 KB - publichealth_v9i1e42678_app3.docx]

Multimedia Appendix 4
Template for Intervention Description and Replication Checklist for Population Health and Policy items for the quality of studies [25-41].
[DOCX File , 23 KB - publichealth_v9i1e42678_app4.docx]

References

1. Dhillon RS, Srikrishna D. When is contact tracing not enough to stop an outbreak? Lancet Infect Dis 2018 Dec;18(12):1302-1304. [doi: 10.1016/s1473-3099(18)30656-x]
2. Thomas Craig KJ, Rizvi R, Willis V, Kassler WJ, Jackson GP. Effectiveness of contact tracing for viral disease mitigation and suppression: evidence-based review. JMIR Public Health Surveill 2021 Oct 06;7(10):e32468 [FREE Full text] [doi: 10.2196/32468] [Medline: 34612841]
3. Contact tracing for COVID-19: current evidence, options for scale-up and an assessment of resources needed. European Centre for Disease Prevention and Control. 2020 May 5. URL: https://www.ecdc.europa.eu/en/publications-data/contact-tracing-covid-19-evidence-scale-up-assessment-resources [accessed 2023-05-24]
4. World Health Organization. Contact tracing in the context of COVID-19: interim guidance, 1 February 2021. World Health Organization. 2021. URL: https://apps.who.int/iris/handle/10665/339128 [accessed 2023-05-24]
5. Interim guidance on developing a COVID-19 case investigation and contact tracing plan: overview. Centers for Disease Control and Prevention. 2022 Feb 28. URL: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/overview.html [accessed 2023-05-24]
6. Contact tracing: public health management of persons, including healthcare workers, who have had contact with COVID-19 cases in the European Union – third update. European Centre for Disease Prevention and Control (ECDC). 2020 Nov 18.

URL: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-contact-tracing-public-health-management-third-update.pdf [accessed 2023-05-24]

7. Resource estimation for contact tracing, quarantine and monitoring activities for COVID-19 cases in the EU/EEA. European Centre for Disease Prevention and Control (ECDC). 2020 Mar 2. URL: https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-resources-for-contact-tracing-2-March-2020_0.pdf [accessed 2023-05-24]
8. Analysis of COVID-19 contact tracing data from Ireland, Italy and Spain – 2020 data. European Center for Disease Prevention and Control (ECDC). 2022 Mar 15. URL: https://www.ecdc.europa.eu/en/publications-data/analysis-covid-19-contact-tracing-data-ireland-italy-and-spain-2020-data [accessed 2023-05-24]
9. Guidance on contact tracing for COVID-19 pandemic. Africa Centre for Disease Control and Prevention (CDC). 2020 Apr 9. URL: https://africacdc.org/download/guidance-on-contact-tracing-for-covid-19-pandemic/ [accessed 2023-03-24]
10. NHS Test and Trace: what to do if you are contacted. UK Health Security Agency. 2020 May 27. URL: https://www.gov.uk/guidance/nhs-test-and-trace-how-it-works#:~:text=If%20you%20are%20a%20close,you%20have%20a%20negative%20result [accessed 2023-05-24]
11. COVID-19 Contact tracing: country experiences and way forward. European Centre for Disease Prevention and Control. 2022 Jul 12. URL: https://www.ecdc.europa.eu/en/publications-data/covid-19-contact-tracing-country-experiences-and-way-forward [accessed 2023-05-24]
12. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health 2020 Apr;8(4):e488-e496 [FREE Full text] [doi: 10.1016/S2214-109X(20)30074-7] [Medline: 32119825]
13. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). J Epidemiol Community Health 2020 Oct 23;74(10):861-866 [FREE Full text] [doi: 10.1136/jech-2020-214051] [Medline: 32576605]
14. Hatami H, Qaderi S, Shah J, Rezaeian A, Farsi Y, Alinasab F, et al. COVID-19: national pandemic management strategies and their efficacies and impacts on the number of secondary cases and prognosis: A systematic review. Int J Prev Med 2022;13(1):100. [doi: 10.4103/ijpvm.ijpvm_464_20]
15. Hossain AD, Jarolimova J, Elnaiem A, Huang CX, Richterman A, Ivers LC. Effectiveness of contact tracing in the control of infectious diseases: a systematic review. Lancet Public Health 2022 Mar;7(3):e259-e273 [FREE Full text] [doi: 10.1016/S2468-2667(22)00001-9] [Medline: 35180434]
16. Anglemyer A, Moore TH, Parker L, Chambers T, Grady A, Chiu K, et al. Digital contact tracing technologies in epidemics: a rapid review. Cochrane Database Syst Rev 2020 Aug 18;8(8):CD013699 [FREE Full text] [doi: 10.1002/14651858.CD013699] [Medline: 33502000]
17. Skarp JE, Downey LE, Ohrnberger JW, Cilloni L, Hogan AB, Sykes AL, et al. A systematic review of the costs relating to non-pharmaceutical interventions against infectious disease outbreaks. Appl Health Econ Health Policy 2021 Sep 11;19(5):673-697 [FREE Full text] [doi: 10.1007/s40258-021-00659-z] [Medline: 34114184]
18. Akinbi A, Forshaw M, Blinkhorn V. Contact tracing apps for the COVID-19 pandemic: a systematic literature review of challenges and future directions for neo-liberal societies. Health Inf Sci Syst 2021 Dec 13;9(1):18 [FREE Full text] [doi: 10.1007/s13755-021-00147-7] [Medline: 33868671]
19. Seto E, Challa P, Ware P. Adoption of COVID-19 contact tracing apps: a balance between privacy and effectiveness. J Med Internet Res 2021 Mar 04;23(3):e25726 [FREE Full text] [doi: 10.2196/25726] [Medline: 33617459]
20. Hogan K, Macedo B, Macha V, Barman A, Jiang X. Contact tracing apps: lessons learned on privacy, autonomy, and the need for detailed and thoughtful implementation. JMIR Med Inform 2021 Jul 19;9(7):e27449 [FREE Full text] [doi: 10.2196/27449] [Medline: 34254937]
21. Mbwogge M. Mass testing with contact tracing compared to test and trace for the effective suppression of COVID-19 in the United Kingdom: systematic review. JMIRx Med 2021;2(2):e27254 [FREE Full text] [doi: 10.2196/27254] [Medline: 33857269]
22. Vogt F, Kurup KK, Mussleman P, Habrun C, Crowe M, Woodward A, et al. Contact tracing indicators for COVID-19: rapid scoping review and conceptual framework. PLoS One 2022 Feb 28;17(2):e0264433 [FREE Full text] [doi: 10.1371/journal.pone.0264433] [Medline: 35226699]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339(jul21 1):b2535 [FREE Full text] [doi: 10.1136/bmj.b2535] [Medline: 19622551]
24. Campbell M, Katikireddi SV, Hoffmann T, Armstrong R, Waters E, Craig P. TIDieR-PHP: a reporting guideline for population health and policy interventions. BMJ 2018 May 16;361:k1079 [FREE Full text] [doi: 10.1136/bmj.k1079] [Medline: 29769210]
25. Asiimwe N, Tabong PT, Iro SA, Noora CL, Opoku-Mensah K, Asampong E. Stakeholders perspective of, and experience with contact tracing for COVID-19 in Ghana: a qualitative study among contact tracers, supervisors, and contacts. PLoS One 2021 Feb 11;16(2):e0247038 [FREE Full text] [doi: 10.1371/journal.pone.0247038] [Medline: 33571296]
26. Kalyanaraman N, Fraser MR. Containing COVID-19 through contact tracing : a local health agency approach. Public Health Rep 2021 Nov 10;136(1):32-38 [FREE Full text] [doi: 10.1177/0033354920967910] [Medline: 33170094]

27. Koetter P, Pelton M, Gonzalo J, Du P, Exten C, Bogale K, et al. Implementation and process of a COVID-19 contact tracing initiative: leveraging health professional students to extend the workforce during a pandemic. Am J Infect Control 2020 Dec;48(12):1451-1456 [FREE Full text] [doi: 10.1016/j.ajic.2020.08.012] [Medline: 32798633]
28. Mueller U, Omosehin O, Akinkunmi A, Ayanbadejo J, Somefun E, Momah-Haruna A. Contact tracing in an African Megacity during COVID 19: lessons learned. Afr J Reprod Health 2020 Jun;24(s1):27-31. [doi: 10.29063/ajrh2020/v24i2s.4] [Medline: 34077049]
29. Niccolai L, Shelby T, Weeks B, Schenck C, Goodwin J, Hennein R, et al. Community trace: rapid establishment of a volunteer contact tracing program for COVID-19. Am J Public Health 2021 Jan;111(1):54-57. [doi: 10.2105/ajph.2020.305959]
30. Pelton M, Medina D, Sood N, Bogale K, Buzzelli L, Blaker J, et al. Efficacy of a student-led community contact tracing program partnered with an academic medical center during the coronavirus disease 2019 pandemic. Ann Epidemiol 2021 Apr;56:26-33.e1 [FREE Full text] [doi: 10.1016/j.annepidem.2020.10.004] [Medline: 33775279]
31. Reid M, Enanoria W, Stoltey J, Philip S, Fuchs J, Lockwood A, et al. The SARS-CoV-2 pandemic: the race to trace: contact tracing scale-up in San Francisco-early lessons learned. J Public Health Policy 2021 Jun 04;42(2):211-221 [FREE Full text] [doi: 10.1057/s41271-021-00285-y] [Medline: 34088978]
32. Breeher L, Boon A, Hainy C, Murad MH, Wittich C, Swift M. A framework for sustainable contact tracing and exposure investigation for large health systems. Mayo Clin Proc 2020 Jul;95(7):1432-1444 [FREE Full text] [doi: 10.1016/j.mayocp.2020.05.008] [Medline: 32561146]
33. Zirbes J, Sterr CM, Steller M, Dapper L, Nonnenmacher-Winter C, Günther F. Development of a web-based contact tracing and point-of-care-testing workflow for SARS-CoV-2 at a German University Hospital. Antimicrob Resist Infect Control 2021 Jul 02;10(1):102 [FREE Full text] [doi: 10.1186/s13756-021-00971-2] [Medline: 34215330]
34. Wong S, Kwong R, Wu T, Chan J, Chu M, Lee S, et al. Risk of nosocomial transmission of coronavirus disease 2019: an experience in a general ward setting in Hong Kong. J Hosp Infect 2020 Jun;105(2):119-127 [FREE Full text] [doi: 10.1016/j.jhin.2020.03.036] [Medline: 32259546]
35. Mak S, Fung KS, Li KK. Formulation of a departmental COVID-19 contingency plan for contact tracing and facilities management. Hong Kong Med J 2021 Apr 7;27(2):148-149 [FREE Full text] [doi: 10.12809/hkmj208670] [Medline: 33824217]
36. Hall M, Bui H, Rowe J, Do T. COVID-19 case and contact investigation in an office workspace. Mil Med 2020 Dec 30;185(11-12):e2162-e2165 [FREE Full text] [doi: 10.1093/milmed/usaa194] [Medline: 32676657]
37. de Laval F, Grosset-Janin A, Delon F, Allonneau A, Tong C, Letois F, et al. Lessons learned from the investigation of a COVID-19 cluster in Creil, France: effectiveness of targeting symptomatic cases and conducting contact tracing around them. BMC Infect Dis 2021 May 19;21(1):457 [FREE Full text] [doi: 10.1186/s12879-021-06166-9] [Medline: 34011278]
38. Draper A, Dempsey K, Boyd R, Childs E, Black H, Francis L, et al. The first 2 months of COVID-19 contact tracing in the Northern Territory of Australia, March-April 2020. Commun Dis Intell (2018) 2020 Jul 02;44 [FREE Full text] [doi: 10.33321/cdi.2020.44.53] [Medline: 32615916]
39. Quach H, Hoang NT, Nguyen CK, Pham QT, Phung CD, Tran ND, et al. Successful containment of a flight-imported COVID-19 outbreak through extensive contact tracing, systematic testing and mandatory quarantine: lessons from Vietnam. Travel Med Infect Dis 2021 Jul;42:102084 [FREE Full text] [doi: 10.1016/j.tmaid.2021.102084] [Medline: 34048935]
40. Clarke M, Devlin J, Conroy E, Kelly E, Sturup-Toft S. Establishing prison-led contact tracing to prevent outbreaks of COVID-19 in prisons in Ireland. J Public Health (Oxf) 2020 Aug 18;42(3):519-524 [FREE Full text] [doi: 10.1093/pubmed/fdaa092] [Medline: 32567653]
41. Fields VL, Kiphibane T, Eason JT, Hafoka SF, Lopez AS, Schwartz A, et al. Assessment of contact tracing for COVID-19 among people experiencing homelessness, Salt Lake County Health Department, March-May 2020. Ann Epidemiol 2021 Jul;59:50-55 [FREE Full text] [doi: 10.1016/j.annepidem.2021.04.002] [Medline: 33894384]
42. Utilizing primary health care for COVID-19 testing and contact tracing in Ghana. The Primary Health Care Performance Initiative (PHCPI). URL: https://improvingphc.org/sites/default/files/Ghana_Testing%20and%20contact%20tracing.pdf [accessed 2023-05-24]
43. Baxter S, Goyder E, Chambers D, Johnson M, Preston L, Booth A. Interventions to improve contact tracing for tuberculosis in specific groups and in wider populations: an evidence synthesis. Health Serv Deliv Res 2017 Jan;5(1):1-102. [doi: 10.3310/hsdr05010] [Medline: 28121090]
44. Operational considerations for adapting a contact tracing program to respond to the COVID-19 pandemic in non-US settings. Centers for Disease Control and Prevention. 2021. URL: https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/grc-747141 [accessed 2023-04-24]
45. Mazza C, Girardi D, Gentile L, Gaeta M, Signorelli C, Odone A. Public health effectiveness of digital contact tracing in the COVID-19 pandemic: a systematic review of available data. Acta Biomed 2021 Dec 10;92(S6):e2021439 [FREE Full text] [doi: 10.23750/abm.v92iS6.12237] [Medline: 34889315]
46. Grekousis G, Liu Y. Digital contact tracing, community uptake, and proximity awareness technology to fight COVID-19: a systematic review. Sustain Cities Soc 2021 Aug;71:102995 [FREE Full text] [doi: 10.1016/j.scs.2021.102995] [Medline: 34002124]

47. Jenniskens K, Bootsma MC, Damen JA, Oerbekke MS, Vernooij RW, Spijker R, et al. Effectiveness of contact tracing apps for SARS-CoV-2: a rapid systematic review. BMJ Open 2021 Jul 12;11(7):e050519 [FREE Full text] [doi: 10.1136/bmjopen-2021-050519] [Medline: 34253676]
48. Guidance on ending the isolation period for people with COVID-19, third update. European Centre for Disease Prevention and Control. 2022 Jan 28. URL: https://www.ecdc.europa.eu/en/publications-data/covid-19-guidance-discharge-and-ending-isolation [accessed 2023-05-24]
49. Infection prevention and control and preparedness for COVID-19 in healthcare settings - sixth update. European Centre for Disease Prevention and Control. 2021 Feb 9. URL: https://www.ecdc.europa.eu/en/publications-data/infection-prevention-and-control-and-preparedness-covid-19-healthcare-settings [accessed 2023-05-24]
50. COVID-19 clusters and outbreaks in occupational settings in the EU/EEA and the UK. European Centre for Disease Prevention and Control. 2020 Aug 11. URL: https://tinyurl.com/mrxav6b7 [accessed 2023-05-24]
51. Preventing and mitigating COVID-19 at work: policy brief, 19 May 2021. World Health Organization. 2021 May 19. URL: https://apps.who.int/iris/handle/10665/341328 [accessed 2023-05-24]
52. COVID-19 case investigation and contact tracing in non-healthcare workplaces : how employers can assist health departments. Centers for Disease Control and Prevention (U.S.). 2020 Sep 30. URL: https://stacks.cdc.gov/view/cdc/99214 [accessed 2023-05-24]
53. World Health Organization. Regional Office for Europe. Risk communication and community engagement for COVID-19 contact tracing: interim guidance. World Health Organization. 2021. URL: https://apps.who.int/iris/handle/10665/339100 [accessed 2023-05-24]
54. Fetzer T, Graeber T. Measuring the scientific effectiveness of contact tracing: evidence from a natural experiment. Proc Natl Acad Sci U S A 2021 Aug 17;118(33):e2100814118 [FREE Full text] [doi: 10.1073/pnas.2100814118] [Medline: 34385318]
55. Koçak C. COVID-19 isolation and contact tracing with country samples: a systematic review. Iran J Public Health 2021 Aug 28;50(8):1547-1554 [FREE Full text] [doi: 10.18502/ijph.v50i8.6800] [Medline: 34917525]
56. Rhodes T, Lancaster K. Evidence-making interventions in health: a conceptual framing. Soc Sci Med 2019 Oct;238:112488. [doi: 10.1016/j.socscimed.2019.112488] [Medline: 31422173]
57. Lancaster K, Rhodes T, Rosengarten M. Making evidence and policy in public health emergencies: lessons from COVID-19 for adaptive evidence-making and intervention. Evidence Policy 2020;16(3):477-490 [FREE Full text] [doi: 10.1332/174426420X15913559981103]

Abbreviations

CDC: Centers for Disease Control and Prevention
CT: contact tracing
CTT: contact tracer team
ECDC: European Centre for Disease Control and Prevention
ETP: exposure triage provider
HCW: health care worker
ICP: Infection Control and Prevention
MeSH: Medical Subject Headings
OHS: Occupational Health Services
PCR: polymerase chain reaction
PEH: people experiencing homelessness
PHO: public health officer
PICOS: Population or Problem, Intervention, Comparison, Outcome, and Study Type
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
REDCap: Research Electronic Data Capture
TIDieR-PHP: Template for Intervention Description and Replication Checklist for Population Health and Policy
UNFPA: United Nations Population Fund
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 11.10.22; peer-reviewed by M Sane Schepisi, CY Lin; comments to author 16.12.22; revised version received 27.01.23; accepted 14.02.23; published 23.06.23.

Please cite as:
Amicosante AMV, Rosso A, Bernardini F, Guglielmi E, Eugeni E, Da Re F, Baglio G
COVID-19 Contact Tracing Strategies During the First Wave of the Pandemic: Systematic Review of Published Studies
JMIR Public Health Surveill 2023;9:e42678
URL: https://publichealth.jmir.org/2023/1/e42678
doi:10.2196/42678
PMID:37351939

Review

Safety and Efficacy of Long-Acting Injectable Agents for HIV-1: Systematic Review and Meta-Analysis

Wenjing Wang[1*], BMed; Shengnan Zhao[1*], MD; Yaxin Wu[1*], BMed; Wenshan Duan[1], BMed; Sibo Li[1], BMed; Zhen Li[2], PhD; Caiping Guo[1], PhD; Wen Wang[1], PhD; Tong Zhang[1], PhD; Hao Wu[1], PhD; Xiaojie Huang[1], MD, PhD

[1]Clinical and Research Center for Infectious Diseases, Beijing Youan Hospital, Capital Medical University, Beijing, China

[2]Beijing Key Laboratory for HIV/AIDS Research, Clinical and Research Center for Infectious Diseases, Beijing Youan Hospital, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Xiaojie Huang, MD, PhD
Clinical and Research Center for Infectious Diseases
Beijing Youan Hospital, Capital Medical University
No.8 Xitoutiao, Youanmenwai, Feng Tai District
Beijing, 100069
China
Phone: 86 13681144481
Email: huangxiaojie78@ccmu.edu.cn

Abstract

Background: HIV-1 infection continues to affect global health. Although antiretrovirals can reduce the viral load or prevent HIV-1 infection, current drugs require daily oral use with a high adherence level. Long-acting antiretrovirals (LA-ARVs) significantly improve medication adherence and are essential for HIV-1 prophylaxis and therapy.

Objective: This study aimed to investigate the safety and efficacy of long-acting cabotegravir (CAB-LA) and long-acting rilpivirine (RPV-LA) in the prevention and treatment of HIV-1 infection.

Methods: PubMed, Embase, and the Cochrane Library were searched for studies from database inception to November 12, 2022. We included studies that reported efficacy and safety data on LA-ARV intervention in people living with HIV and excluded reviews, animal studies, and articles with missing or duplicate data. Virological suppression was defined as plasma viral load <50 copies/mL 6 months after antiviral therapy initiation. We extracted outcomes for analysis and expressed dichotomous data as risk ratios (RRs) and continuous data as mean differences. Depending on the heterogeneity assessment, a fixed- or random-effects model was used for data synthesis. We performed subgroup analyses of the partial safety and efficacy outcomes of CAB-LA+RPV-LA. The protocol was registered with the Open Science Framework.

Results: We included 12 trials comprising 10,957 individuals, of which 7 were prevention trials and 5 were treatment trials. CAB-LA and RPV-LA demonstrated safety profiles comparable with those of the placebo in terms of adverse event–related withdrawal. Moreover, the efficacy data showed that CAB-LA had a better effect on HIV-1 prevention than tenofovir disoproxil fumarate–emtricitabine (17/5161, 0.33% vs 75/5129, 1.46%; RR 0.21, 95% CI 0.07-0.61; I^2=70%). Although CAB-LA+RPV-LA had more drug-related adverse events (556/681, 81.6% vs 37/598, 6.2%; RR 12.50, 95% CI 3.98-39.23; I^2=85%), a mild or moderate injection site reaction was the most common reaction, and its frequency decreased over time. The efficacy of CAB-LA+RPV-LA was comparable with that of daily oral drugs at 48 and 96 weeks (1302/1424, 91.43% vs 915/993, 92.2%; RR 0.99, 95% CI 0.97-1.02; I^2=0%), and a high level of virological suppression of 80.9% (186/230) was maintained even after 5 years of LA-ARV use. Similar efficacy outcomes were observed in both treatment-naive and treatment-experienced patients (849/911, 93.2% vs 615/654, 94%; RR 0.99, 95% CI 0.96-1.02; I^2=0%). According to the questionnaires, more than 85% of people living with HIV favored LA-ARVs.

Conclusions: LA-ARVs showed favorable safety profiles for both the prevention and treatment of HIV-1 infection and were well tolerated. CAB-LA has more satisfactory efficacy than tenofovir disoproxil fumarate–emtricitabine, significantly reducing the rate of HIV-1 infection. CAB-LA+RPV-LA maintains virological suppression for a long time and may be a viable switching strategy with enhanced public health benefits by reducing transmission. However, further trials are required to confirm the efficacy of these drugs.

(JMIR Public Health Surveill 2023;9:e46767) doi:10.2196/46767

KEYWORDS

long-acting cabotegravir; CAB-LA; long-acting rilpivirine; RPV-LA; pre-exposure prophylaxis; PrEP; treatment; long-term suppression

Introduction

Background

HIV-1 infection continues to affect global human health, with an estimated 38.4 million individuals living with HIV-1 by the end of 2021 and 28.7 million accessing antiretroviral therapy (ART) by the end of December 2021 [1]. Despite the enormous progress made in ART, over 5000 people are newly infected with HIV-1 worldwide every day [2]. Studies have shown that ART has demonstrated significant efficacy in limiting HIV-1 viral replication, reducing plasma viral load (VL), and strengthening the immune system. This provides people living with HIV with long-lasting virological suppression [3,4]. The use of ART has the potential to considerably reduce HIV-related morbidity and death, thereby improving the overall health status of people living with HIV and prolonging their life expectancy [5]. Pre-exposure prophylaxis (PrEP) with ART before HIV-1 exposure is a major HIV-1 prevention innovation that can effectively reduce HIV-1 infection rates and provide benefits to populations at high risk for HIV-1 infection [6]. Oral regimens comprising tenofovir disoproxil fumarate–emtricitabine (TDF-FTC) have proven to be effective in preventing HIV-1 infection in high-risk individuals in several clinical trials of HIV-1 PrEP [7,8].

Although antiretrovirals can control the symptoms of people living with HIV, there is no cure for HIV-1 infections. Currently, ART requires lifelong administration, and most antiretrovirals must be taken daily to suppress HIV-1 infection [9,10], which requires a high level of adherence to ART by users [11,12]. people living with HIV or people at high risk of HIV-1 infection face considerable challenges in maintaining the efficacy of ART in the face of the pressure to take daily drugs. In addition, prolonged daily oral medications can lead to treatment fatigue [13,14] coupled with changes in daily lifestyle, stigma associated with long-term medication use, burden of drug use, and multiple social factors [15], all of which can lead to suboptimal adherence and reduce the efficacy of medications or even lead to the emergence of drug-resistant viral variants. Long-acting antiretrovirals (LA-ARVs) have the potential to increase ART adherence, reduce HIV-1 transmission, and achieve public health benefits by decreasing the number of new HIV-1 infections.

The advent of LA-ARVs may overcome several problems associated with the daily use of oral pills and privacy of patients taking drugs. They are used once a month or every 2 months, greatly reducing the frequency and burden of daily medication, eliminating the need to take daily pills [16,17] and improving the convenience of taking medicine while protecting patients' privacy and reducing the social stigma associated with HIV-1, which is important for people living with HIV and those at high risk of HIV-1 infection. Cabotegravir is a novel integrase inhibitor with acceptable safety and tolerability when administered orally once daily [18]. The nonnucleoside reverse transcriptase inhibitor rilpivirine is administered orally once daily for HIV-1 treatment. LA-ARVs have currently been approved for the treatment and prevention of HIV-1 infection in several countries. In March 2020, ViiV Healthcare announced that Health Canada approved Cabenuva (cabotegravir and rilpivirine extended-release injectable suspensions) for the market. This was the first global approval of Cabenuva as an alternative to existing antiviral regimens for virologically suppressed people living with HIV [19]. The United States Food and Drug Administration has approved Cabenuva as a complete regimen for the treatment of HIV-1 infection in adults with virological suppression on a stable antiretroviral regimen [19] and Apretude (cabotegravir extended-release injectable suspension) for HIV-1 PrEP in high-risk adults and adolescents weighing at least 35 kg [20]. Apretude is also approved as a prophylactic drug in Europe [21], Australia, South Africa [22], Zimbabwe [23], and other countries. In light of the benefits associated with LA-ARVs, there is a growing interest in their development, as well as concerns regarding their potential adverse effects, safety, and efficacy.

Objectives

Therefore, this meta-analysis aimed to summarize existing trials on LA-ARVs, particularly long-acting cabotegravir (CAB-LA), long-acting rilpivirine (RPV-LA), and CAB-LA+RPV-LA; assess their safety and efficacy; and provide evidence for their widespread use.

Methods

Overview

We conducted a meta-analysis using version 6.2 of the Cochrane Handbook for Systematic Reviews of Intervention [24]. We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines were followed [25] (Multimedia Appendix 1). The protocol was deposited and registered in the Open Science Framework [26].

Search Strategy

We thoroughly searched the PubMed, Embase, and Cochrane Library databases for eligible articles published from their inception to November 12, 2022. Only articles published in English were included in this study. Search terms included "HIV," "AIDS," "long-acting formulations," "cabotegravir," "gsk1265744," and "rilpivirine" (Multimedia Appendix 2). After deleting duplicates, 2 writers independently screened the titles, abstracts, and full-text papers. Disagreements were resolved through dialogue and reaching a consensus.

Study Eligibility Criteria

The inclusion criteria were as follows: (1) study design, randomized trials; (2) study population, adults aged ≥18 years; (3) intervention, LA-ARVs; and (4) end points, which included safety and efficacy data. For the exclusion criteria, we excluded pregnant or lactating women. We excluded reviews, conference abstracts, case reports, letters, animal studies, irretrievable full-text articles, and studies that shared the same data set. We only included the latest or most detailed data when repeated data were encountered.

Currently, different guidelines have different definitions of HIV-1 virological suppression. According to the World Health Organization guideline [27], we defined virological suppression as a plasma VL <50 copies/mL 6 months after the start of ART.

Drug-related adverse event (AE) and AE-related withdrawals were the primary safety outcomes. The major efficacy outcomes were confirmed with HIV-1 infection for prevention and the percentage of individuals with plasma HIV-1 RNA <50 copies/mL for treatment. Secondary safety outcomes included the proportion of individuals with (1) any AE, (2) AE of grade 3 or higher, (3) injection site reaction (ISR), (4) serious AE (SAE), and (5) death, whereas secondary efficacy outcomes of treatment included (1) changes in CD4$^+$T cell counts from baseline, (2) the incidence of confirmed virological failure (VF), and (3) the proportion of resistance-associated mutations (RAMs) in patients who acquired HIV-1 infection or with confirmed virologic failure.

Data Extraction and Risk of Bias Assessment

A Microsoft Excel spreadsheet was used to extract data from each study. We extracted all relevant data: (1) general information, including the name of the first author and trial, trial number, publication year, and the number of patients; (2) study parameters, including interventions, controls, and study design; and (3) participant characteristics, including age and the proportion of the male population; and (4) outcomes, including safety and efficacy data.

As a means of assessing bias risk, the Cochrane risk-of-bias tool was used to analyze the 5 domains of random sequence generation, allocation concealment, blinding of participants and personnel and outcome assessors, incomplete outcome data, and selective outcome reporting. Findings were classified as low, unclear, or high bias. The 2 reviewers worked separately to extract the data and assess the possibility of bias. Multimedia Appendix 3 [28-42] shows details of the risk of bias assessment.

Statistical Analysis

RevMan statistical software (version 5.3; The Cochrane Collaboration) was used for statistical analyses and Adobe Illustrator (Adobe) was used for graphical editing and presentation. We used proportion and risk ratio (RR) values to express dichotomous data, whereas for continuous data, we used mean differences. A forest plot was used to estimate cumulative effects. The I^2 test revealed statistically significant heterogeneity. Four degrees of heterogeneity were distinguished: 0% to 40%, presumably insignificant; 30% to 60%, medium; 50% to 90%, substantial; and 75% to 100%, high [24]. Subgroup analysis was used to investigate the potential heterogeneity. If applicable, leave-one-out sensitivity analysis was performed to investigate the consistency of the results. If sufficient publications were available, publication bias was investigated using funnel plots and the Egger test.

Results

Study Selection and Characteristics

We obtained 5662 records through an initial literature search, including 1321 in PubMed, 3857 in Embase, and 484 in the Cochrane Library. After removal of duplicates, 4108 studies remained. A total of 72 publications remained after screening the titles and abstracts for full-text inspection. Furthermore, 15 papers comprising 10,957 individuals were ultimately enrolled for data analysis, which contained 12 trials, including 7 trials for prevention (5 on CAB-LA [28-32], 2 on RPV-LA [33,34]), and 5 trials on CAB-LA+RPV-LA [35-42] for treatment. The LATTE-2, ATLAS, and FLAIR treatment trials included one maintenance phase and one extension phase study. Figure 1 illustrates the procedure for conducting the literature search.

CAB-LA was injected intramuscularly and subcutaneously at intervals between 4 and 12 weeks, at dosages ranging from 100 to 800 mg. RPV-LA was injected intramuscularly at doses between 300 and 1200 mg every 4 to 8 weeks. Two doses of CAB-LAs and RPV-LAs were administered intramuscularly. In one case, 400 mg CAB-LA and 600 mg RPV-LA were administered every 4 weeks; in the other case, 600 mg CAB-LA and 900 mg RPV-LA were administered every 8 weeks. Only one CAB-LA trial used both injection methods (intramuscular and subcutaneous). Only intramuscular injection was selected for combined analysis of the data. Table 1 summarizes the characteristics of the included studies.

The bias risk assessment showed that 8 studies [28,29,31-34,36,40] had a low bias risk, whereas 7 open-label trials [30,35,37-39,41,42] with unclear allocation concealment had a high risk of bias (Figures S1 and S2 in Multimedia Appendix 3).

Figure 1. Diagram flow illustrating literature search process. *We searched the database until November 12, 2022.

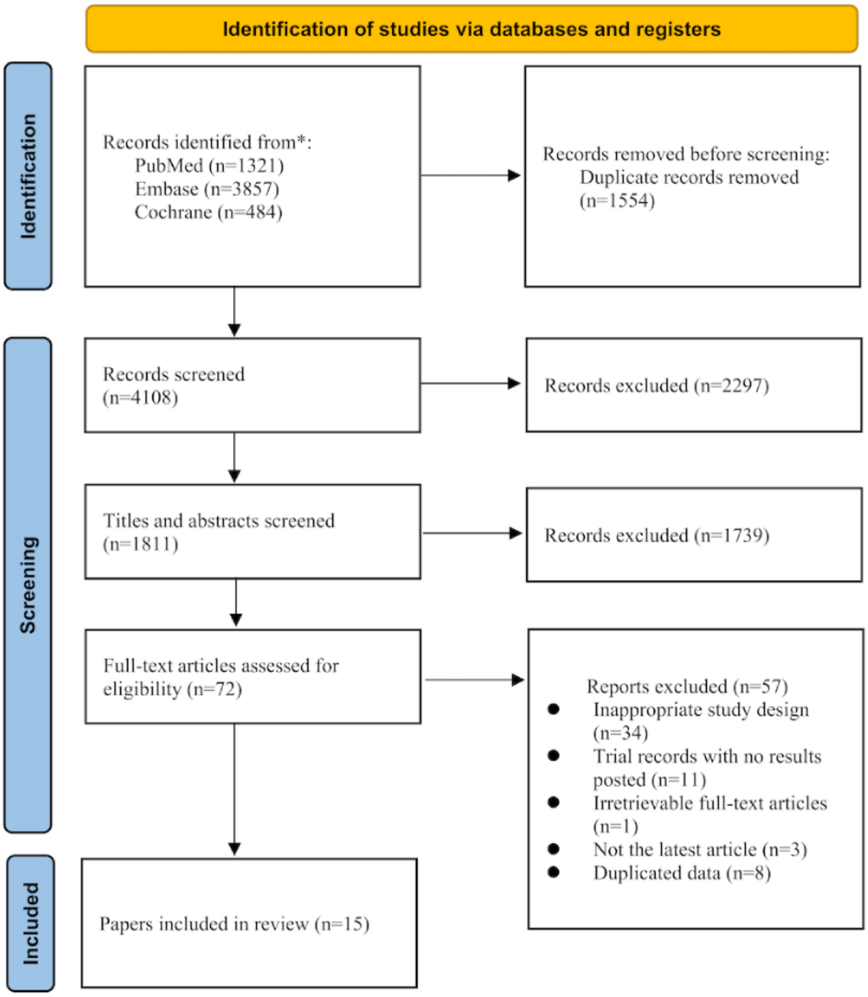


Table 1. Characteristics of included studies.

Study (author, year [trial name; NCT ID or Clinicaltrials.gov identifier])	Phase	Masking	Location	Population characteristics	Design	Sample size, n	Age (years)[a]	Male, n (%)
Prophylaxis								
Landovitz et al [28], 2018 (HPTN 077; NCT02178800)	II a	Double-blind	Multicenter	Participants were HIV uninfected at screening and at low risk for HIV infection	CAB[b]-LA[c] IM[d] 600 mg Q8W[e] vs placebo	89 (69I[f], 20C[g])	31 (24-37)	29 (33)
Landovitz et al [28], 2018 (HPTN 077; NCT02178800)	II a	Double-blind	Multicenter	Participants were HIV uninfected at screening and at low risk for HIV infection	CAB-LA IM 800 mg Q12W[h] vs placebo	110 (82I, 28C)	33 (25-42)	38 (35)
Markowitz et al [29], 2017 (ECLAIR; NCT02076178)	II a	Double-blind	Multicenter	Participants were male at birth, HIV uninfected, and reported having at least one casual se partner in the past 24 mo	CAB-LA IM 800 mg Q12W vs placebo	127 (106I, 21C)	31 (20-61)	127 (100)
Spreen et al [30], 2014 (NCT01756131)	I	Open-label	Single-center	Healthy volunteers	CAB-LA IM 100, 200, 200×2, 400, 400×2 mg single-dose vs placebo	72 (58I, 14C)	35.1 (10.4)	39 (54.2)
Landovitz et al [31], 2021 (HPTN 083; NCT02720094)	II b-III	Double-blind	Multicenter	Adults had a negative HIV serological test at enrollment and had an undetectable blood HIV RNA viral load within 14 days before trial entry	CAB-LA IM 600 mg Q8w with TDF[i]-FTC[j] placebo QD[k] vs TDF-FTC QD with CAB-LA placebo IM Q8w	4570 (2283I, 2287C)	26 (22-32)	3992 (87.4)
Delany-Moretlwe et al [32], 2022 (HPTN 084; NCT03164564)	III	Double-blind	Multicenter	Female reported at least 2 episodes of vaginal intercourse in the previous 30 days were at risk of HIV infection based on an HIV risk score	CAB-LA IM 600 mg Q8w with TDF-FTC placebo QD vs TDF-FTC QD with CAB-LA placebo IM Q8w	3224 (1614I, 1610C)	25 (22-30)	0 (0)
Verloes et al [33] 2015, (NCT01031589)	I	Open-label; double-blind	Single-center	Healthy volunteers	RPV-LA IM 300,600 mg single-dose; RPV-LA IM 1200 and 600 and 600 mg Q4w[m] vs placebo	11; 8 (6I, 2C)	47 (31-58); 47 (31-58)	6 (31.6)
Bekker et al [34], 2020 (HPTN 076; NTC02165202)	II	Double-blind	Multicenter	Healthy, sexually active, low-risk, HIV-uninfected women	RPV[l]-LA IM 1200 mg Q8W vs placebo	136 (91I, 45C)	31 (25-38)	0 (0)
Treatment								

Study (author, year [trial name; NCT ID or Clinicaltrials.gov identifier])	Phase	Masking	Location	Population characteristics	Design	Sample size, n	Age (years)[a]	Male, n (%)
Margolis et al [35] 2017, (LATTE-2; NCT02120352)	II b	Open-label	Multicenter	Treatment-naive adults with HIV-1 who were given either CAB 30 mg PO[n]+ABC[o]/3TC[p] 600/300 mg PO QD for 20 weeks and who had a VL[q]<50 copies/mL	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W or CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W vs CAB PO 30 mg + ABC PO 600 mg + 3TC PO 300 mg QD	286 (115I,115I, 56C)	35 (19-64)	262 (92)
Smith et al [36], 2021 (LATTE-2 extension phase; NCT02120352)	II b	Open-label	Multicenter	Adults completing 96 weeks of LATTE-2 enter an extension phase	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W. Optimized loading dose (100 weeks) followed by CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; Optimized loading dose (100 weeks and 104 weeks) followed by CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W	115; 115; 10; 34	36 (19-62); 34 (20-64); 41 (21-56); 36 (19-56)	109 (95); 107 (93); 8 (80); 28 (82)

Study (author, year [trial name; NCT ID or Clinical-trials.gov identifier])	Phase	Masking	Location	Population characteristics	Design	Sample size, n	Age (years)[a]	Male, n (%)
Orkin et al [37], 2021 (FLAIR; NCT02938520)	III	Open-label	Multicenter	Treatment-naive adults with HIV-1 who were given DTG[q]/ABC/3TC PO 50/600/300 mg QD for 20 weeks and who had a VL<50 copies/mL	CAB PO 30 mg+RPV PO 25 mg QD for 4 weeks, followed by CAB-LA; IM 600 mg+RPV-LA IM 900 mg, then CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W for 100 weeks vs DTG PO 50 mg + ABC PO 600 mg + 3TC PO 300 mg QD for 100 weeks	566 (283I, 283C)	34 (29-43)	439 (78)
Orkin et al [38], 2021 (FLAIR extension phase; NCT02938520)	III	Open-label	Multicenter	Adults completing 100 weeks of ATLAS enter an extension phase	Switched from CAB 30 mg+RPV 25 mg QD to CAB-LA+RPV-LA (direct-to-injection group); switched from CAB 30 mg+RPV 25 mg QD to CAB-LA+RPV-LA (oral lead-in group); continued the long-acting regimen	111; 121; 283	36 (30-45); 38 (31-46); 34 (29-42)	24 (22); 27 (22); 63 (22)

Study (author, year [trial name; NCT ID or Clinicaltrials.gov identifier])	Phase	Masking	Location	Population characteristics	Design	Sample size, n	Age (years)[a]	Male, n (%)
Swindells et al [39], 2020 (ATLAS; NCT02951052)	Ⅲ	Open-label	Multicenter	Adults with HIV-1 and had a VL<50 copies/mL for ≧6 months while taking PI-, NNRTI-, or INSTI-based regimen with a two-NRTI backbone	CAB 30 mg+RPV 25 mg QD for 4 weeks, followed by CAB-LA IM 600 mg+RPV-LA IM 900 mg, then CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W for 52 weeks vs PI-, NNRTI-, or INSTI-based QD for 52 weeks	616 (308I, 308C)	42 (18-82)	413 (67)
Swindells et al [40], 2022 (ATLAS extension phase; NCT02951052)	Ⅲ	Open-label	Multicenter	Adults completing 52 weeks of ATLAS enter an extension phase	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W vs switched from CAB 30 mg+RPV 25 mg QD to CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W	52 (23, 29)	—[r]	—
Jaeger et al [41], 2021 (ATLAS-2M; NCT03299049)	Ⅲb	Open-label	Multicenter	Parents from ATLAS with a VL<50 copies/mL at screening and additional adults with HIV-1 and a VL<50 copies/mL for ≥6 months while taking standard oral ART	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W	523; 522	42 (34-50)	765 (73)

Study (author, year [trial name; NCT ID or Clinicaltrials.gov identifier])	Phase	Masking	Location	Population characteristics	Design	Sample size, n	Age (years)[a]	Male, n (%)
Mills et al [42], 2021 (POLAR; NCT03639311)	II b	Open-label	Multicenter	Adults with HIV-1 who had a VL<50 copies/mL and completed at least 300 weeks of the LATTE study	CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W for 48 weeks vs DTG PO 50 mg + RPV PO 25 mg QD for 48 weeks	97 (90I, 7C)	41 (25-63)	92 (94.8)

[a] Unless otherwise stated, age is presented as mean (SD) or median (IQR).

[b] CAB: cabotegravir.

[c] LA: long-acting.

[d] IM: intramuscular.

[e] Q8W: every 8 weeks.

[f] I: intervention group.

[g] C: control group.

[h] Q12W: every 12 weeks.

[i] TDF: tenofovir.

[j] FTC: emtricitabine.

[k] QD: daily.

[l] RPV: rilpivirine.

[m] Q4W: every 4 week.

[n] PO: per os.

[q] DTG: dolutegravir.

[o] ABC: abacavir.

[p] 3TC: lamivudine.

[q] VL: viral load.

[r] Not available.

PrEP Medication

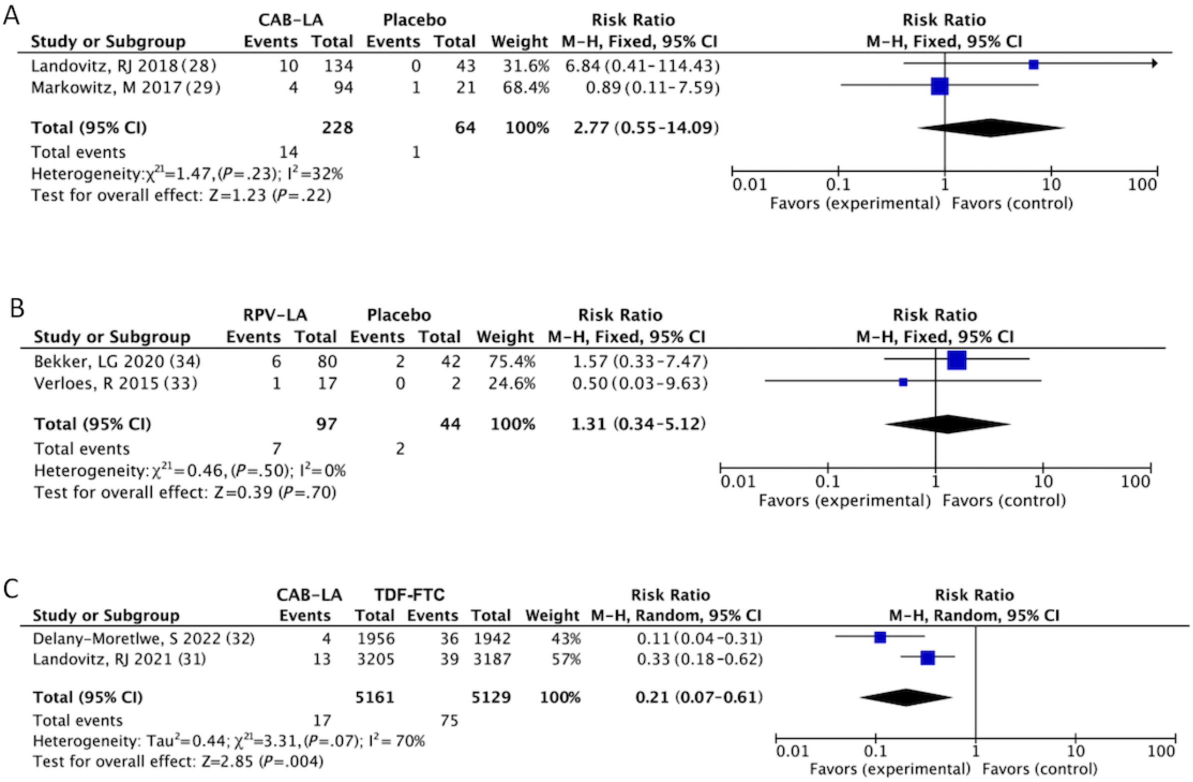
Safety

For CAB-LA, 3 of the 5 trials were placebo controls and 2 were TDF-FTC controls. Compared with the placebo, the frequency of AE-related withdrawals (14/228, 6.1% vs 1/64, 2%; RR 2.77, 95% CI 0.55-14.09; I^2=32%; Figure 2A) was similar, as was any AE or SAE (Figures S1A and S1B in Multimedia Appendix 4 [28-40,42]). In contrast, there was high heterogeneity for ISR (165/268, 61.6% vs 15/72, 21%; RR 3.38, 95% CI 0.79-14.44; I^2=80%; Figure S1C in Multimedia Appendix 4). Because of the lack of studies, we could not determine the origin of heterogeneity.

Compared with daily oral TDF-FTC, there was no apparent difference in AE of grade 3 or higher (1003/3894, 25.76% vs 1047/3892, 26.9%; RR 0.96, 95% CI 0.89-1.03; I^2=0%; Figure S1D in Multimedia Appendix 4) and SAE (153/3894, 3.93% vs 154/3892, 3.96%; RR 0.99, 95% CI 0.80-1.24; I^2=0%; Figure S1E in Multimedia Appendix 4). In comparison, CAB-LA had more ISRs (2301/3799, 60.57% vs 815/3798, 21.46%; RR 3.03, 95% CI 2.27-4.04; I^2=91%; Figure S1F in Multimedia Appendix 4) than TDF-FTC. Although ISR was more frequent in the intervention group, the most frequently reported side effect was injection site pain. ISR is often mild to moderate in severity and its frequency gradually diminishes.

Similarly, RPV-LA was well tolerated. For the primary outcome, the occurrence of AE-related withdrawal (7/97, 7% vs 2/44, 5%; RR 1.31, 95% CI 0.34-5.12; I^2=0%; Figure 2B) did not differ between the 2 groups. Analyses of the SAE and ISR (Figures S1G and S1H in Multimedia Appendix 4) showed no obvious differences.

Figure 2. Meta-analyses on safety and efficacy profiles of CAB-LA and RPV-LA: (A) AE-related withdrawals: CAB-LA versus placebo [28,29], (B) AE-related withdrawal: RPV-LA versus placebo [33,34], and (C) confirmed HIV-1 infection: CAB-LA versus TDF-FTC [31,32]. AE: adverse event; CAB-LA: long-acting cabotegravir; RPV-LA: long-acting rilpivirine; TDF-FTC: tenofovir disoproxil fumarate–emtricitabine.



Efficacy

The ECLAIR trial [29] revealed 1 proven HIV-1 infection in the cabotegravir group throughout the follow-up period and 1 confirmed HIV-1 infection in the placebo group within the injection phase. In addition to TDF-FTC's reduction in HIV-1 infection by TDF-FTC, CAB-LA also reduced the rate by 76% (17/5161, 0.33% vs 75/5129, 1.46%; RR 0.21, 95% CI 0.07-0.61; I^2=70%; Figure 2C). One of the 13 participants who acquired HIV-1 infection after enrollment in the CAB-LA group of the HIV Prevention Trials Network (HPTN) 083 trial [31] was reclassified as having a baseline infection. Only 4 infections occurred with regular CAB-LA injections and 2 had drug-resistant mutations.

In HPTN 083 trial [31], 86% of the TDF-FTC group participants maintained plasma concentrations over the lower limit of quantification (0.31 ng/mL), which showed high adherence to oral drugs. In the HPTN 084 trial [32], 2 of the 4 patients with infections in the CAB-LA group had undetectable plasma cabotegravir concentrations, one had delayed cabotegravir injections, and one had baseline infection. None of the infections occurred with on-time injections. The TDF-FTC group documented poor or nonadherence in 35 of the 36 infections, and TDF concentrations greater than 0.31 ng per mL were observed in 55.91% (1084/1939) of participants.

For RPV-LA, the HPTN 076 trial [34] reported that one placebo participant developed HIV-1 infection during the injection period. A total of 67.9% (76/112) of the participants strongly agreed to use RPV-LA at the time of their last injection visit, and 88% stated that they would use it in the future. The efficacy and safety results of the prophylactic drugs are detailed in Multimedia Appendix 5 [28-34]. The RAMs in the people living with HIV are detailed in Multimedia Appendix 6 [28-42].

Treatment

Safety

Table 2 presents the results for the CAB-LA+RPV-LA. Data analysis revealed that the AE rate of CAB-LA+RPV-LA was higher than that of the daily oral drugs. The incidence of drug-related AE (556/681, 81.6% vs 37/598, 6.2%; RR 12.50, 95% CI 3.98-39.23; I^2=85%; Figure 3A) at 48 weeks in CAB-LA+RPV-LA group was higher than that in the daily oral group, so were the AE-related withdrawal (48/1194, 4.02% vs 14/937, 1.5%; RR 2.65, 95% CI 1.48-4.74; I^2=0%; Figure 3B) at 48 weeks and 96 weeks, any AE (647/681, 95% vs 448/598, 74.9%; RR 1.27, 95% CI 1.12-1.44; I^2=74%; Figure S2A in Multimedia Appendix 4), and AE of grade 3 or higher (75/681, 11% vs 34/598, 5.7%; RR 1.93, 95% CI 1.30-2.87; I^2=5%; Figure S2B in Multimedia Appendix 4).

When ISR was excluded, drug-related AE after excluding ISR (167/591, 28.3% vs 36/591, 6.1%; RR 5.41, 95% CI 1.35-21.64; I^2=91%; Figure S2C in Multimedia Appendix 4) and AE of grade 3 or higher after excluding ISR (47/591, 8% vs 34/591, 5.8%; RR 1.38, 95% CI 0.90-2.12; I^2=45%; Figure S2D in Multimedia Appendix 4) were reanalyzed. There was no obvious difference in the SAE (Figure S2E in Multimedia Appendix 4). No deaths occurred because of the drugs used.

When patients in the LATTE-2, ATLAS, and FLAIR trials entered the extension phase [31,32,35], people on daily oral medication switched to CAB-LA+RPV-LA, and the long-acting group continued their previous treatment. No statistically significant differences were observed between the groups. The incidence of AE-related withdrawal (14/513, 2.7% vs 5/276, 1.8%; RR 0.76, 95% CI 0.24-2.41; I^2=19%; Figure 3C), any AE (506/513, 98.6% vs 246/276, 89.1%; RR 1.06, 95% CI 0.89-1.25; I^2=97%; Figure S3A in Multimedia Appendix 4), and SAE (39/513, 7.6% vs 16/276, 5.8%; RR 0.82, 95% CI 0.46-1.47; I^2=0%; Figure S3B in Multimedia Appendix 4) in extension phase did not differ significantly between the long-acting arm and the switch arm.

Table 2. Safety and efficacy results of CAB[a]-LA[b]+RPV[c]-LA (all results are expressed in terms of frequency (n/N) unless otherwise stated).

Trial name	Design	Safety							Efficacy		
		Any AE[d]	Drug-related AE	AE of grade 3 or higher	SAE[e]	ISR[f]	AE withdrawal	Death	HIV-1 RNA level<50 copies/mL	Median change from baseline in $CD4^+$ lymphocyte count-per mm^3 (n)	Confirmed VF[g]
LATTE-2 (96 weeks)	CAB-LA IM[h] 400 mg+RPV-LA IM 600 mg Q4W[i]; CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W[j]; CAB PO[k] 30 mg+ABC[l] PO 600 mg + 3TC[m] PO 300 mg QD[n]	N/A[o]	N/A	N/A	11/115; 11/115; 7/56	112/115; 110/115	8/115; 2/115; 1/56	N/A	100/115; 108/115; 47/56	226 (IQR 145 to 393) (100); 239 (IQR 111 to 359) (109); 317 (IQR 214 to 505) (47)	0; 2/115; 1/56
LATTE-2 extension phase (256 weeks)	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W; Optimized loading dose (100 weeks) followed by CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; Optimized loading dose (100 weeks and 104 weeks) followed by CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W	115/115; 115/115; 10/10; 34/34	NR[p]	38/115; 39/115; 3/10; 7/34	27/115; 25/115; 1/10; 6/34	N/A	20/115; 3/115; 1/10; 1/34	3/115; 0; 0; 0	85/115; 101/115; 9/10; 32/34	396 (SD 294) (85); 326 (SD 218) (102); 211 (SD 318) (9); −14 (SD 319) (32)	NR

Trial name	Design	Safety							Efficacy		
		Any AE[d]	Drug-related AE	AE of grade 3 or higher	SAE[e]	ISR[f]	AE withdrawal	Death	HIV-1 RNA level<50 copies/mL	Median change from baseline in CD4$^+$ lymphocyte count-per mm^3 (n)	Confirmed VF[g]
FLAIR (48 weeks)	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; DTG[q] PO 50 mg + ABC PO 600 mg + 3TC PO 300 mg QD	267/283; 225/283	236/283; 28/283	31/283; 11/283	18/283; 12/283	N/A	9/283; 4/283	0; 0	265/283; 264/283	NR	4/283; 3/283
FLAIR (96 weeks)	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; DTG PO 50 mg + ABC PO 600 mg + 3TC PO 300 mg QD	274/283; 242/283	246/283; 33/283	40/283; 16/283	24/283; 22/283	N/A	14/283; 4/283	0; 0	245/283; 253/283	57 (IQR −43 to 181) (246); 109.5 (IQR 18 to 228) (254)	4/283; 4/283
FLAIR (124 weeks)	DTI group (after 24 weeks of CAB+RPV); OLI group (after 24 weeks of CAB+RPV); Randomly assigned long-acting arm (after 124 weeks of CAB+RPV)	102/111; 100/121; 276/283	86/111; 79/121; 248/283	5/111; 9/121; 49/283	4/111; 5/121; 33/283	NR	1/111; 2/121; 15/283	0; 0; 0	110/111; 113/121; 227/283	NR	N/A
ATLAS (48 weeks)	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; PI-, NNRTI-, or INSTI-based QD	294/308; 220/308	255/308; 8/308	35/308; 23/308	13/308; 14/308	N/A	14/308; 5/308	0; 1/308	285/308; 294/308	4.0 (IQR −536 to 801) (308); 13.5 (IQR −1043 to 521) (308)	3/308; 4/308

Trial name	Design	Safety							Efficacy		
		Any AE[d]	Drug-related AE	AE of grade 3 or higher	SAE[e]	ISR[f]	AE withdrawal	Death	HIV-1 RNA level<50 copies/mL	Median change from baseline in CD4$^+$ lymphocyte count-per mm^3 (n)	Confirmed VF[g]
ATLAS (96 weeks)	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; Switched from CAB 30 mg+RPV 25 mg QD to CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W	N/A	N/A	N/A	N/A	N/A	N/A	0; 0	23/23; 28/29	−5.7 (SD 167.6) (23); −33.6 (SD 145.3) (29)	0; 0
ATLAS-2M (96 weeks)	CAB-LA IM 400 mg+RPV-LA IM 600 mg Q4W; CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W	499/523; 488/522	413/523; 415/522	NR	28/523; 33/522	400/517; 412/516	19/523; 18/522	1/523; 1/522	472/523; 475/522	NR	2/523; 9/522
POLAR (52 weeks)	CAB-LA IM 600 mg+RPV-LA IM 900 mg Q8W; DTG PO 50 mg + RPV PO 25 mg QD	86/90; 3/7	65/90; 1/7	9/90; 0/7	5/90; 1/7	N/A	1/90; 0/7	NR	88/90; 7/7	−12.5 (IQR −138 to 71) (90); −68 (IQR −152 to 152) (7)	0; 0

[a] CAB: cabotegravir.

[b] LA: long-acting.

[c] RPV: rilpivirine.

[d] AE: adverse event.

[e] SAE: serious adverse event.

[f] ISR: injection-site reaction.

[g] VF: virological failure.

[h] IM: intramuscular.

[i] Q4W: every 4 weeks.

[j] Q8W: every 8 weeks.

[k] PO: per os.

[l] ABC: abacavir.

[m] 3TC: lamivudine.

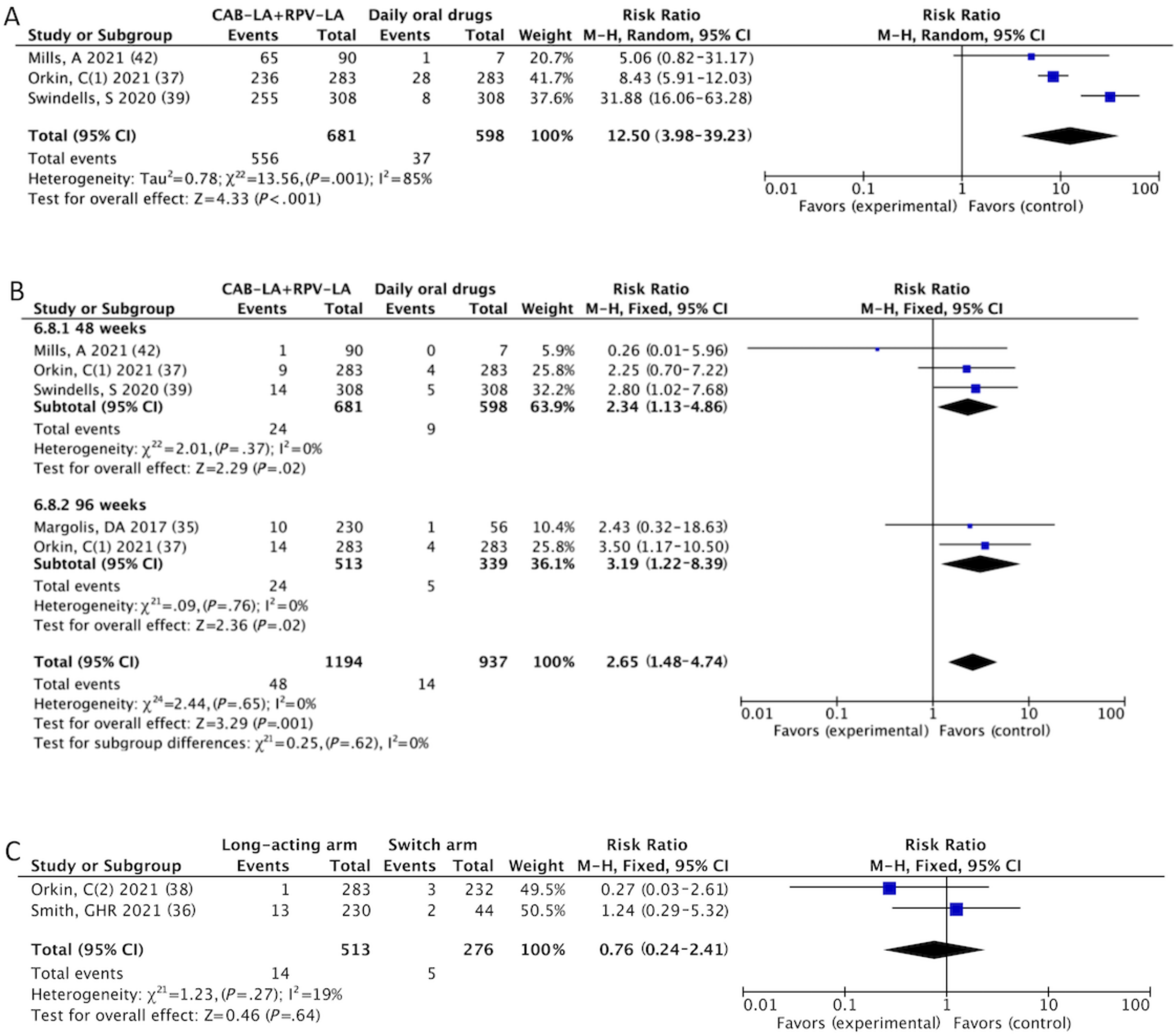
[n] QD: daily.

[o] N/A: not available.

[p] NR: not reported.

[q] DTG: dolutegravir.

Figure 3. Meta-analyses on safety profiles of CAB-LA+RPV-LA: (A) drug-related AE: CAB-LA+RPV-LA versus daily oral drugs [37,39,42], (B) AE-related withdrawal: CAB-LA+RPV-LA versus daily oral drugs [35,37,39,42], and (C) AE-related withdrawal: long-acting arm versus switch arm [36,38]. AE: adverse event; CAB-LA: long-acting cabotegravir; RPV-LA: long-acting rilpivirine.



Efficacy

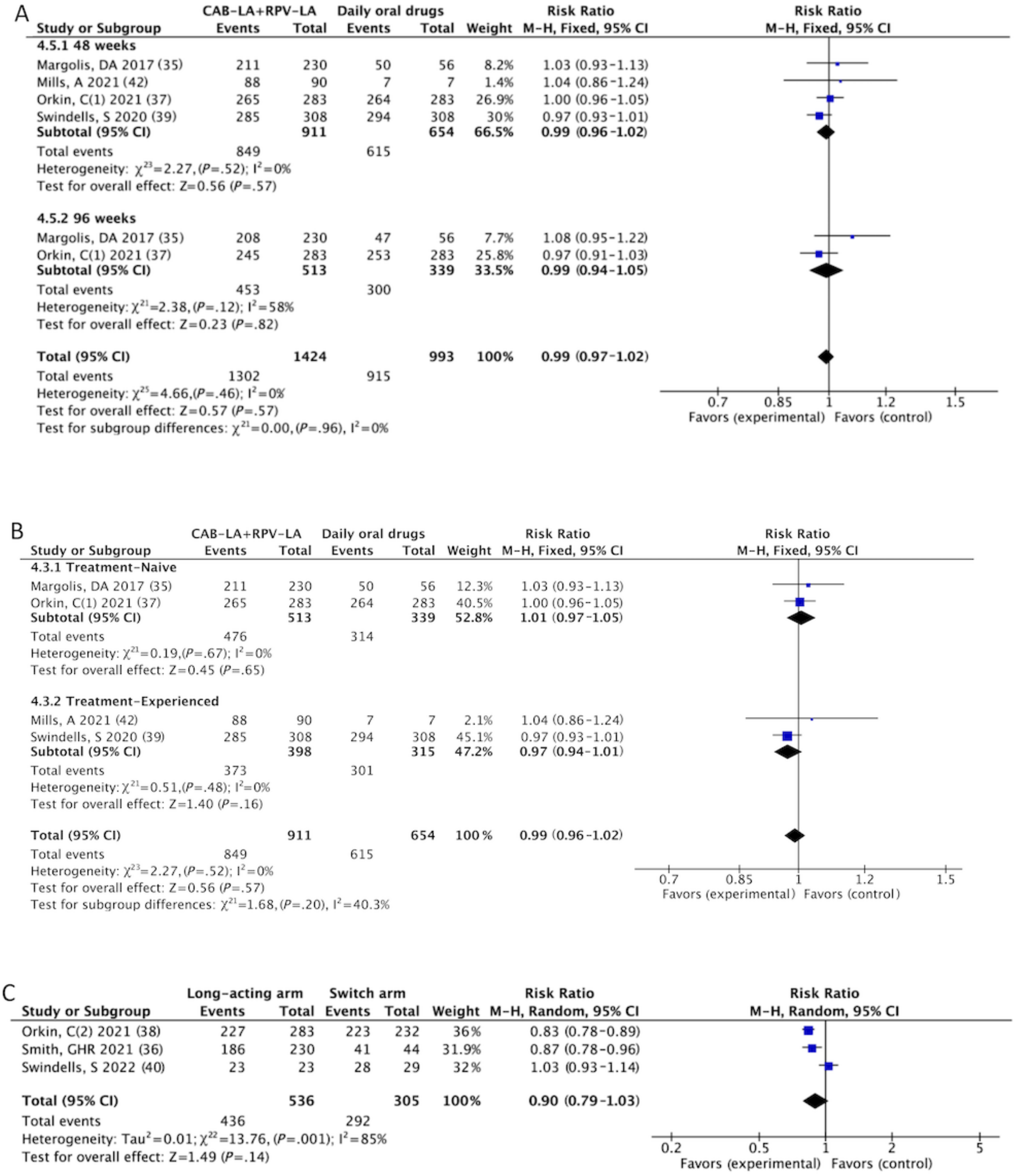
In terms of efficacy, we performed a subgroup analysis of the percentage of individuals with plasma HIV-1 RNA levels <50 copies/mL (1302/1424, 91.43% vs 915/993, 92.2%; RR 0.99, 95% CI 0.97-1.02; I^2=0%; Figure 4A) and the incidence of confirmed VF (13/1104, 1.18% vs 12/930, 1.3%; RR 0.93, 95% CI 0.43-2.04; I^2=0%; Figure S2F in Multimedia Appendix 4) at 48 weeks and 96 weeks. Similarly, when we grouped the treatment-naive patients in the study with treatment-experienced patients (849/911, 93.2% vs 615/654, 94%; RR 0.99, 95% CI 0.96-1.02; I^2=0%; Figure 4B), there were no statistically significant differences in the efficacy outcomes.

We summarized and analyzed the results of RAMs in patients with VF during treatment. The results showed no statistical difference between the CAB-LA+RPV-LA group and the daily oral drug group in terms of RAMs associated with integrase inhibitors (6/9, 67% vs 0/9, 0%; RR 4.64, 95% CI 1.00-21.62; I^2=0%; Figure S1A in Multimedia Appendix 6) or nonnucleoside reverse transcriptase inhibitors (7/9, 78% vs 3/9, 34%; RR 2.12, 95% CI 0.90-4.97; I^2=0%; Figure S2B in Multimedia Appendix 6). In addition, a bar chart of the number of RAMs occurring between groups is presented in Figure S3 in Multimedia Appendix 6.

For patients in the extension phase, there was no significant difference in the percentage of infected people with plasma HIV-1 RNA less than 50 copies/mL (436/536, 81.3% vs 292/305, 95.7%; RR 0.90, 95% CI 0.79-1.03; I^2=85%; Figure 4C) between those who continued LA-ARVs and those who switched to LA-ARVs. This was the change from the baseline in CD4$^+$ T cell counts (Figure S3C in Multimedia Appendix 4). Only one participant in the switch arm reported confirmed VF during the extension phase of the FLAIR trial. The 96-week results of the ATLAS-2M [41] corroborated that 90.62% (947/1045) of individuals maintained plasma HIV-1 RNA levels of less than 50 copies/mL. Compared with daily oral ART, patient satisfaction levels with CAB-LA+RPV-LA were high.

Most participants preferred the long-acting regimen; after all injections, 99% (206/208) of the long-acting group in the LATTE-2 trial [35] expressed satisfaction with their current ongoing therapy. In the ATLAS and POLAR trials [39,42], 97.4% (266/273) and 87.5% (77/88) of respondents in the long-acting group, who filled out the questionnaires, said they were inclined to use CAB-LA+RPV-LA instead of daily oral drugs. In the FLAIR extension phase [38], we found that 4 people living with HIV using CAB-LA+RPV-LA for 124 weeks had RAMs, 3 of which had been detected at 96 weeks of the FLAIR trial. The detailed resistance data are provided in Multimedia Appendix 6.

Figure 4. Meta-analyses on efficacy profiles of CAB-LA+RPV-LA: (A) plasma HIV-1 RNA less than 50 copies/mL: CAB-LA+RPV-LA versus daily oral drugs [35,37,39,42], (B) plasma HIV-1 RNA less than 50 copies/mL: CAB-LA+RPV-LA versus daily oral drugs [35,37,39,42], and (C) plasma HIV-1 RNA less than 50 copies/mL: long-acting arm versus switch arm [36,38,40]. CAB-LA: long-acting cabotegravir; RPV-LA: long-acting rilpivirine.



A

Study or Subgroup	CAB-LA+RPV-LA Events	CAB-LA+RPV-LA Total	Daily oral drugs Events	Daily oral drugs Total	Weight	Risk Ratio M-H, Fixed, 95% CI
4.5.1 48 weeks						
Margolis, DA 2017 (35)	211	230	50	56	8.2%	1.03 (0.93-1.13)
Mills, A 2021 (42)	88	90	7	7	1.4%	1.04 (0.86-1.24)
Orkin, C(1) 2021 (37)	265	283	264	283	26.9%	1.00 (0.96-1.05)
Swindells, S 2020 (39)	285	308	294	308	30%	0.97 (0.93-1.01)
Subtotal (95% CI)		**911**		**654**	**66.5%**	**0.99 (0.96-1.02)**
Total events	849		615			
Heterogeneity: $\chi^{23}=2.27$, (P=.52); $I^2=0\%$						
Test for overall effect: Z=0.56 (P=.57)						
4.5.2 96 weeks						
Margolis, DA 2017 (35)	208	230	47	56	7.7%	1.08 (0.95-1.22)
Orkin, C(1) 2021 (37)	245	283	253	283	25.8%	0.97 (0.91-1.03)
Subtotal (95% CI)		**513**		**339**	**33.5%**	**0.99 (0.94-1.05)**
Total events	453		300			
Heterogeneity: $\chi^{21}=2.38$, (P=.12); $I^2=58\%$						
Test for overall effect: Z=0.23 (P=.82)						
Total (95% CI)		**1424**		**993**	**100%**	**0.99 (0.97-1.02)**
Total events	1302		915			
Heterogeneity: $\chi^{25}=4.66$, (P=.46); $I^2=0\%$						
Test for overall effect: Z=0.57 (P=.57)						
Test for subgroup differences: $\chi^{21}=0.00$, (P=.96), $I^2=0\%$						

B

Study or Subgroup	CAB-LA+RPV-LA Events	CAB-LA+RPV-LA Total	Daily oral drugs Events	Daily oral drugs Total	Weight	Risk Ratio M-H, Fixed, 95% CI
4.3.1 Treatment-Naive						
Margolis, DA 2017 (35)	211	230	50	56	12.3%	1.03 (0.93-1.13)
Orkin, C(1) 2021 (37)	265	283	264	283	40.5%	1.00 (0.96-1.05)
Subtotal (95% CI)		**513**		**339**	**52.8%**	**1.01 (0.97-1.05)**
Total events	476		314			
Heterogeneity: $\chi^{21}=0.19$, (P=.67); $I^2=0\%$						
Test for overall effect: Z=0.45 (P=.65)						
4.3.2 Treatment-Experienced						
Mills, A 2021 (42)	88	90	7	7	2.1%	1.04 (0.86-1.24)
Swindells, S 2020 (39)	285	308	294	308	45.1%	0.97 (0.93-1.01)
Subtotal (95% CI)		**398**		**315**	**47.2%**	**0.97 (0.94-1.01)**
Total events	373		301			
Heterogeneity: $\chi^{21}=0.51$, (P=.48); $I^2=0\%$						
Test for overall effect: Z=1.40 (P=.16)						
Total (95% CI)		**911**		**654**	**100 %**	**0.99 (0.96-1.02)**
Total events	849		615			
Heterogeneity: $\chi^{23}=2.27$, (P=.52); $I^2=0\%$						
Test for overall effect: Z=0.56 (P=.57)						
Test for subgroup differences: $\chi^{21}=1.68$, (P=.20), $I^2=40.3\%$						

C

Study or Subgroup	Long-acting arm Events	Long-acting arm Total	Switch arm Events	Switch arm Total	Weight	Risk Ratio M-H, Random, 95% CI
Orkin, C(2) 2021 (38)	227	283	223	232	36%	0.83 (0.78-0.89)
Smith, GHR 2021 (36)	186	230	41	44	31.9%	0.87 (0.78-0.96)
Swindells, S 2022 (40)	23	23	28	29	32%	1.03 (0.93-1.14)
Total (95% CI)		**536**		**305**	**100%**	**0.90 (0.79-1.03)**
Total events	436		292			
Heterogeneity: $Tau^2=0.01$; $\chi^{22}=13.76$, (P=.001); $I^2=85\%$						
Test for overall effect: Z=1.49 (P=.14)						

Discussion

Principal Findings

To the best of our knowledge, this meta-analysis is the first to assess the safety and efficacy of LA-ARVs in people living with HIV, as well as the advantages of LA-ARVs as a switching strategy. We found a previously published article on LA-ARVs that focused on exploring the safety and pharmacokinetic profiles of 2 drugs, CAB-LA and RPV-LA, for use as PrEP in the prevention of HIV-1 infection [43]. We added newly updated articles on this premise, including data from the HPTN 083 [31] and HPTN 084 [32] trials. The CAB-LA was examined and described in greater detail. In addition, we not only focused on prophylactic drugs but also compared the existing daily antiretrovirals with all existing trials of the combination of LA-ARVs for treating HIV-1 infection. We collected study data with a long follow-up period, and in addition to the results from the maintenance phase of treatment, we collected data within the extension phase for CAB-LA+RPV-LA, providing information on the efficacy beyond 96 weeks and even up to 5 years. To provide more comprehensive evidence for the use of LA-ARVs, we extracted and analyzed the safety and efficacy data, adherence, and patient satisfaction. The availability of LA-ARVs is important for HIV-1 prevention and treatment and for reducing HIV-1 transmission. LA-ARVs with longer dose intervals significantly improve medication adherence by reducing the burden of daily dosing [44], protecting patient privacy, and reducing the stigma associated with HIV-1 infection.

This analysis suggests that participants have a favorable safety profile for LA-ARVs in terms of either prevention or treatment. LA-ARVs have also shown better efficacy than daily oral drugs; however, fewer trials and more studies are required to validate their efficacy. First of all, all participants tolerated the prophylactic drugs, CAB-LA and RPV-LA. Although we observed differences in drug-related AE between the 2 groups in the ECLAIR trial [29], the majority were classified as grade 1 or 2, had little effect on the participant, and were not life-threatening. CAB-LA had more ISRs than oral TDF-FTC owing to the injection of CAB-LA placebo in the control group, which had less of an effect on the organism. Most ISRs in the long-acting group were mild or moderate with a low frequency, and a few patients dropped out of the trial due to ISR. Therefore, CAB-LAs have a good safety profile as a prophylactic agent. CAB-LA is more effective than TDF-FTC in preventing HIV-1 infection and can effectively lower the incidence of HIV-1 infection. Adherence has been shown to have a substantial effect on drug efficacy. The efficacy results from the HPTN 084 trial were comparable with those of the HPTN 083 trial; however, participants in the HPTN 084 trial had lower adherence to the drug, resulting in a higher infection rate in the oral drug group. Second, most infections that occurred in the CAB-LA group in both trials [31,32] were due to low or undetectable plasma concentrations, which were the result of participants not injecting the dose on time or the infection was occurring during the oral induction phase. A combined analysis of the efficacy results from the 2 trials revealed a high degree of heterogeneity, which may be attributable to the variance in the enrolled populations. The statistically significant difference in efficacy in the HPTN 083 trial also demonstrated the superiority of CAB-LA over TDF-FTC. Therefore, cabotegravir is superior to TDF-FTC in preventing HIV-1 infection and has adherence advantages.

However, we observed 4 breakthrough infections in the CAB-LA group in the HPTN 083 trial. Researchers believe that this could be due to several factors, such as low concentrations of cabotegravir in the plasma or rectal tissue, delayed detection of HIV-1 infection, and drug resistance [45]. As mentioned in the package inserts of Apretude, drugs that might significantly reduce cabotegravir's plasma concentration should be avoided. Therefore, to conduct a more thorough investigation of the efficacy of CAB-LA as PrEP, additional data, including pharmacokinetics and other characteristics, are needed, and in the future, additional trials will be done in multiple enrollment populations.

CAB-LA+RPV-LA also had a better safety profile in treating people living with HIV, although patients experienced more AEs (any AE, drug-related AE, AE-related withdrawals, and AE of grade 3 or higher) at the primary end point time of 48 or 96 weeks than those on daily oral drugs. Firstly, similar to CAB-LA, ISR was the most prevalent AE, predominantly mild or moderate, and its frequency decreased as the study progressed. A few patients dropped out of the trials because of ISR, and most participants generally accepted the overall ISR. We reanalyzed the results of drug-related AE and AE of grade ≥3 after excluding ISR. The results demonstrated no statistically significant difference between the 2 groups in terms of AE of grade 3 or higher after excluding ISR. In contrast, most drug-related AEs, excluding ISR, were fever, which is a subjective symptom of patients and may affect the experimental results. In addition, we observed a lower incidence of ISR in the extension phase than in the maintenance phase, which could be attributed to tolerance to long-term injections.

Second, most control groups might have continued their current daily oral drugs, whereas the intervention groups underwent the oral induction phase or switching therapy, which could have led to more AE in the long-acting group. This possibility was consistent with the findings of a previous switch study [46] that the consequences of starting a new treatment instead of continuing the same treatment may lead to an increase in AE. Although patients who received LA-ARVs experienced more AEs, cabotegravir plus rilpivirine was the most popular regimen. Even in participants who switched from daily oral drugs to CAB+RPV, the safety results were not significantly different from those who used CAB-LA+RPV-LA for a long time. Most participants preferred the cabotegravir plus rilpivirine therapy and presented higher levels of satisfaction.

There was no significant difference between CAB-LA+RPV-LA and daily oral drugs in terms of the primary efficacy end point, and CAB-LA+RPV-LA maintained virological suppression in patients well, even in the extension phase, with the advantage of a long injection interval, which proved that CAB-LA+RPV-LA was superior to daily oral agents. In addition, analysis of data from patients treated with CAB-LA+RPV-LA for more than 2 years revealed that the

HIV-1 viral suppression rate was maintained at a high level, even in patients who received LA-ARVs for 5 years [36], indicating that CAB-LA+RPV-LA had the capacity to suppress VL for a long time. This finding lays the foundation for the long-lasting effects of LA-ARVs. However, subgroup analysis of the efficacy outcomes of treatment-naive and treatment-experienced patients who participated in the study revealed heterogeneity between subgroups, indicating that variances in the enrolled population may have influenced the trial results. We observed that the enrolled studies maintained virological suppression before receiving LA-ARVs or after the oral induction phase. According to the guidelines [47], if patients achieve virological suppression of their current ART, arbitrary regimen adjustments are not recommended but may be considered in certain specific situations, such as simplifying the regimen by reducing the number of pills and frequency of administration to facilitate a smooth drug switch. In these trials [36,38,40], patients who smoothly switched from daily oral drugs to LA-ARVs were able to maintain high levels of virological suppression, with no statistically significant differences in efficacy compared with patients on LA-ARVs, which provides support for individuals who are currently receiving oral therapy and wish to switch to LA-ARVs. If LA-ARVs are popularized in the future, CAB-LA + RPV-LA may be a good switching strategy for individuals who use daily oral drugs for long periods and maintain virological suppression.

We observed VFs in both the long-acting arm and daily oral drug arm and analyzed their RAMs. There were no statistically significant differences between the groups; however, because of the limited number of studies and participants, no definitive conclusions could be drawn. Most patients in the long-acting group continued to receive LA-ARVs and had a higher predilection for LA-ARVs than daily oral drugs. Thus, CAB-LA+RPV-LA, which greatly improved patient adherence, provided a more convenient dosing regimen and optimized the daily oral regimen in people living with HIV, who achieved and effectively maintained virological suppression. It performed well in terms of safety and tolerance of drug use in infected patients and could be an effective and promising treatment or switching strategy. However, additional trials are required to confirm the results of our study.

Our study is significant for the promotion of LA-ARVs. The emergence of LA-ARVs has made the prevention and treatment of AIDS easier. It reduces the burden on people living with HIV of different genders, avoids stigmatization associated with daily oral drug treatment, and improves the correct use and compliance of doses. This is particularly important to ensure the effectiveness of pre-exposure prevention measures. Improving the prevention and treatment of HIV-1 will reduce its transmission and contribute to reducing the prevalence of new HIV-1 infections. The United Nations Programme on HIV/AIDS has called for these life-changing injectable drugs to be quickly available, affordable, and fairly distributed to those who need them the most worldwide [48]. This requires more regional and national regulators to quickly approve and adopt a series of measures to reduce sales prices. National HIV-1 prevention programs need to develop LA-ARVs promotion plans and help health systems and communities to prepare for the use of this new drug.

Limitations

This review had several limitations. First, because the results are presented in several formats, it is impossible to integrate data for analysis such as the median form. Second, owing to the relatively low number of trials, we could not perform a publication bias analysis, and the lack of trials on RPV-LA and efficacy data may lead to potential bias. Third, so far, LA-ARVs have been studied mainly in high-income countries, and the cost of cabotegravir and rilpivirine may also be a major obstacle to their supply in low- and middle-income countries. Therefore, more research is needed to determine the efficacy, acceptability, and economic burden of LA-ARVs in different income groups, particularly in low-income and middle-income countries. Finally, we retained only the original articles and trials, discarded conferences, and the resulting data that appeared at the conference. This could have led to the omission of valid data and weakened the persuasiveness of the results.

Conclusions

In conclusion, our findings support the safety and efficacy of CAB-LA, RPV-LA, and CAB-LA+RPV-LA, which are effective and well-tolerated monthly or bimonthly injections that can provide PrEP to people at risk of HIV-1 infection, reduce the rate of HIV-1 infection, and provide a more convenient treatment option for people living with HIV. CAB-LA+RPV-LA maintained virological suppression of HIV-1 for approximately 2 years, with good efficacy even after 5 years of treatment. Similarly, effective virological suppression was maintained after switching from daily oral drugs to LA-ARVs in people living with HIV who successfully achieved virological suppression. In summary, LA-ARVs have shown good safety and tolerability in dosing, can effectively improve patient adherence, protect patient privacy, and can be a promising treatment or an alternative to oral ART. Currently, multiple LA-ARVs have been evaluated in clinical trials, including injectables, rings, and implants, such as PrEP and subcutaneous injections, for treating HIV-1 infection [49]. For example, a phase 2/3 CAPELLA study on lenacapavir confirmed a high rate of virological suppression in patients with multidrug resistance [50], and islatravir has also been confirmed for drug safety in a phase I trial [51]. Therefore, research and development of LA-ARVs is tremendously advantageous for HIV-1 prevention and treatment. Given the lack of such studies, further investigation of the efficacy of RPV-LAs is required.

Acknowledgments

This systematic review was supported by the Fund for Public Health Talent Grant by Beijing Municipal Health Commission (Global Health Governance-02-12; 2022-1-007), Capital Health Development Research (2022-2-2185; 2022-1G-3011), Beijing

Talent Project in the New Millennium (2020A35), and National Science and Technology Major Project of China during the 13th Five-Year Plan Period (grant 2017ZX10201101).

Data Availability

All data generated or analyzed in this study are publicly available and are included in this published article.

Authors' Contributions

WW, SZ, and YW conceived the study. WW, SZ and YW designed the search strategy and performed the literature search. WW and SZ screened studies for eligibility. WW and YW performed data extraction. WD and SL assessed the risk of bias. WW and SZ performed the data analysis. WW and YW interpreted the data analysis and assessed the certainty of evidence. WW and YW wrote the first draft of the manuscript, and all other authors revised the manuscript. All authors have contributed to the manuscript and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklists.
[DOCX File , 41 KB - publichealth_v9i1e46767_app1.docx]

Multimedia Appendix 2
Search strategy.
[DOCX File , 19 KB - publichealth_v9i1e46767_app2.docx]

Multimedia Appendix 3
Risk of Bias for meta-analyses.
[DOCX File , 383 KB - publichealth_v9i1e46767_app3.docx]

Multimedia Appendix 4
Forest plots for the outcomes.
[DOCX File , 1472 KB - publichealth_v9i1e46767_app4.docx]

Multimedia Appendix 5
Safety and efficacy profiles of long-acting antiretroviral drugs for prophylaxis.
[DOCX File , 39 KB - publichealth_v9i1e46767_app5.docx]

Multimedia Appendix 6
Drug resistance data.
[DOCX File , 736 KB - publichealth_v9i1e46767_app6.docx]

References

1. Global HIV and AIDS statistics — Fact sheet. Joint United Nations Programme on HIV/AIDS. 2021. URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2023-05-01]
2. Global AIDS Update—Seizing the moment: tackling entrenched inequalities to end epidemics. Joint United Nations Programme on HIV/AIDS. 2020. URL: https://www.unaids.org/sites/default/files/media_asset/2020_global-aids-report_en.pdf [accessed 2023-05-01]
3. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim EY, Archer J, Pond SL, et al. Persistent HIV-1 replication maintains the tissue reservoir during therapy. Nature 2016 Feb 04;530(7588):51-56 [FREE Full text] [doi: 10.1038/nature16933] [Medline: 26814962]
4. Coelho L, Grinsztejn B, Castilho JL, De Boni R, Quintana MS, Campos DP, et al. Mortality in HIV-infected women, heterosexual men, and men who have sex with men in Rio de Janeiro, Brazil: an observational cohort study. Lancet HIV 2016 Oct;3(10):e490-e498 [FREE Full text] [doi: 10.1016/S2352-3018(16)30052-2] [Medline: 27658875]
5. Ruelas DS, Greene WC. An integrated overview of HIV-1 latency. Cell 2013 Oct 24;155(3):519-529 [FREE Full text] [doi: 10.1016/j.cell.2013.09.044] [Medline: 24243012]
6. Lundgren J, Phillips A. Prevention of HIV transmission by antiretroviral therapy. Lancet HIV 2018 Mar;5(3):e108-e109. [doi: 10.1016/S2352-3018(17)30204-7] [Medline: 29199099]

7. Riddell J4, Amico KR, Mayer KH. HIV preexposure prophylaxis: a review. JAMA 2018 Mar 27;319(12):1261-1268. [doi: 10.1001/jama.2018.1917] [Medline: 29584848]
8. Desai M, Field N, Grant R, McCormack S. Recent advances in pre-exposure prophylaxis for HIV. BMJ 2017 Dec 11;359:j5011 [FREE Full text] [doi: 10.1136/bmj.j5011] [Medline: 29229609]
9. Gardner MR. Promise and progress of an HIV-1 cure by adeno-associated virus vector delivery of anti-HIV-1 biologics. Front Cell Infect Microbiol 2020 Apr 23;10:176 [FREE Full text] [doi: 10.3389/fcimb.2020.00176] [Medline: 32391289]
10. Shapiro MB, Cheever T, Malherbe DC, Pandey S, Reed J, Yang ES, et al. Single-dose bNAb cocktail or abbreviated ART post-exposure regimens achieve tight SHIV control without adaptive immunity. Nat Commun 2020 Jan 07;11(1):70 [FREE Full text] [doi: 10.1038/s41467-019-13972-y] [Medline: 31911610]
11. Sidebottom D, Ekström AM, Strömdahl S. A systematic review of adherence to oral pre-exposure prophylaxis for HIV - how can we improve uptake and adherence? BMC Infect Dis 2018 Nov 16;18(1):581 [FREE Full text] [doi: 10.1186/s12879-018-3463-4] [Medline: 30445925]
12. Altice F, Evuarherhe O, Shina S, Carter G, Beaubrun AC. Adherence to HIV treatment regimens: systematic literature review and meta-analysis. Patient Prefer Adherence 2019 Apr 3;13:475-490 [FREE Full text] [doi: 10.2147/PPA.S192735] [Medline: 31040651]
13. Claborn KR, Meier E, Miller MB, Leffingwell TR. A systematic review of treatment fatigue among HIV-infected patients prescribed antiretroviral therapy. Psychol Health Med 2015;20(3):255-265 [FREE Full text] [doi: 10.1080/13548506.2014.945601] [Medline: 25110152]
14. Nyaku AN, Kelly SG, Taiwo BO. Long-acting antiretrovirals: where are we now? Curr HIV/AIDS Rep 2017 Apr;14(2):63-71. [doi: 10.1007/s11904-017-0353-0] [Medline: 28303548]
15. Viswanathan S, Detels R, Mehta SH, Macatangay BJC, Kirk GD, Jacobson LP. Level of adherence and HIV RNA suppression in the current era of highly active antiretroviral therapy (HAART). AIDS Behav 2015 Apr;19(4):601-611 [FREE Full text] [doi: 10.1007/s10461-014-0927-4] [Medline: 25342151]
16. Benning L, Mantsios A, Kerrigan D, Coleman JS, Golub E, Blackstock O, et al. Examining adherence barriers among women with HIV to tailor outreach for long-acting injectable antiretroviral therapy. BMC Womens Health 2020 Jul 25;20(1):152 [FREE Full text] [doi: 10.1186/s12905-020-01011-8] [Medline: 32711509]
17. Clement ME, Kofron R, Landovitz RJ. Long-acting injectable cabotegravir for the prevention of HIV infection. Curr Opin HIV AIDS 2020 Jan;15(1):19-26 [FREE Full text] [doi: 10.1097/COH.0000000000000597] [Medline: 31644481]
18. Patel P, Xue Z, King KS, Parham L, Ford S, Lou Y, et al. Evaluation of the effect of UGT1A1 polymorphisms on the pharmacokinetics of oral and long-acting injectable cabotegravir. J Antimicrob Chemother 2020 Aug 01;75(8):2240-2248 [FREE Full text] [doi: 10.1093/jac/dkaa147] [Medline: 32361755]
19. FDA approves first injectable treatment for HIV pre-exposure prevention. U.S. Food & Drug Administration. 2021 Dec 20. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-first-injectable-treatment-hiv-pre-exposure-prevention [accessed 2023-05-01]
20. FDA approves first extended-release, injectable drug regimen for adults living with HIV. U.S. Food & Drug Administration. 2021 Jan 21. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-first-extended-release-injectable-drug-regimen-adults-living-hiv [accessed 2023-05-01]
21. European medicines agency validates ViiV healthcare's marketing authorisation application for cabotegravir long-acting injectable for HIV prevention. ViiV Healthcare. 2022 Oct 28. URL: https://viivhealthcare.com/hiv-news-and-media/news/press-releases/2022/october/european-medicines-agency-validates-viiv-healthcare/#5 [accessed 2023-05-01]
22. Significant milestone for HIV prevention in sub-Saharan Africa as South Africa grants regulatory approval for apretude. ViiV Healthcare. URL: https://viivhealthcare.com/hiv-news-and-media/news/company-statements/significant-milestone-for-hiv-prevention/ [accessed 2023-05-01]
23. Major milestone in public health as Zimbabwe approves HIV prevention drug. HealthCare Middle East & Africa. 2022 Oct 21. URL: https://www.healthcareafrica.info/major-milestone-in-public-health-as-zimbabwe-approves-hiv-prevention-drug/ [accessed 2023-05-01]
24. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions. London, UK: The Cochrane Collaboration; Sep 20, 2019.
25. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
26. Wang W. Safety and efficacy of long-acting antiretroviral drugs for HIV-1: a protocol for meta-analysis. Open Science Framework. 2021 Oct 12. URL: https://osf.io/znu6e [accessed 2023-07-10]
27. Consolidated guidelines on HIV prevention, testing, treatment, service delivery and monitoring: recommendations for a public health approach. World Health Organization. Geneva: World Health Organization; 2021 Jul 16. URL: https://www.who.int/publications/i/item/9789240031593 [accessed 2023-05-01]
28. Landovitz RJ, Li S, Grinsztejn B, Dawood H, Liu AY, Magnus M, et al. Safety, tolerability, and pharmacokinetics of long-acting injectable cabotegravir in low-risk HIV-uninfected individuals: HPTN 077, a phase 2a randomized controlled trial. PLoS Med 2018 Nov;15(11):e1002690 [FREE Full text] [doi: 10.1371/journal.pmed.1002690] [Medline: 30408115]

29. Markowitz M, Frank I, Grant RM, Mayer KH, Elion R, Goldstein D, et al. Safety and tolerability of long-acting cabotegravir injections in HIV-uninfected men (ECLAIR): a multicentre, double-blind, randomised, placebo-controlled, phase 2a trial. Lancet HIV 2017 Aug;4(8):e331-e340. [doi: 10.1016/S2352-3018(17)30068-1] [Medline: 28546090]
30. Spreen W, Ford SL, Chen S, Wilfret D, Margolis D, Gould E, et al. GSK1265744 pharmacokinetics in plasma and tissue after single-dose long-acting injectable administration in healthy subjects. J Acquir Immune Defic Syndr 2014 Dec 15;67(5):481-486. [doi: 10.1097/QAI.0000000000000301] [Medline: 25140909]
31. Landovitz RJ, Donnell D, Clement ME, Hanscom B, Cottle L, Coelho L, HPTN 083 Study Team. Cabotegravir for HIV prevention in cisgender men and transgender women. N Engl J Med 2021 Aug 12;385(7):595-608 [FREE Full text] [doi: 10.1056/NEJMoa2101016] [Medline: 34379922]
32. Delany-Moretlwe S, Hughes JP, Bock P, Ouma SG, Hunidzarira P, Kalonji D, HPTN 084 study group. Cabotegravir for the prevention of HIV-1 in women: results from HPTN 084, a phase 3, randomised clinical trial. Lancet 2022 May 07;399(10337):1779-1789 [FREE Full text] [doi: 10.1016/S0140-6736(22)00538-4] [Medline: 35378077]
33. Verloes R, Deleu S, Niemeijer N, Crauwels H, Meyvisch P, Williams P. Safety, tolerability and pharmacokinetics of rilpivirine following administration of a long-acting formulation in healthy volunteers. HIV Med 2015 Sep;16(8):477-484 [FREE Full text] [doi: 10.1111/hiv.12247] [Medline: 25988676]
34. Bekker LG, Li S, Pathak S, Tolley EE, Marzinke MA, Justman JE, HPTN 076 Study Team. Safety and tolerability of injectable Rilpivirine LA in HPTN 076: a phase 2 HIV pre-exposure prophylaxis study in women. EClinicalMedicine 2020 Apr;21:100303 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100303] [Medline: 32280940]
35. Margolis DA, Gonzalez-Garcia J, Stellbrink H, Eron JJ, Yazdanpanah Y, Podzamczer D, et al. Long-acting intramuscular cabotegravir and rilpivirine in adults with HIV-1 infection (LATTE-2): 96-week results of a randomised, open-label, phase 2b, non-inferiority trial. Lancet 2017 Sep 23;390(10101):1499-1510. [doi: 10.1016/S0140-6736(17)31917-7] [Medline: 28750935]
36. Smith GH, Henry WK, Podzamczer D, Masiá MD, Bettacchi CJ, Arasteh K, et al. Efficacy, safety, and durability of long-acting Cabotegravir and Rilpivirine in adults with human immunodeficiency virus type 1 infection: 5-year results from the LATTE-2 study. Open Forum Infect Dis 2021 Sep;8(9):ofab439 [FREE Full text] [doi: 10.1093/ofid/ofab439] [Medline: 34557563]
37. Orkin C, Oka S, Philibert P, Brinson C, Bassa A, Gusev D, et al. Long-acting cabotegravir plus rilpivirine for treatment in adults with HIV-1 infection: 96-week results of the randomised, open-label, phase 3 FLAIR study. Lancet HIV 2021 Apr;8(4):e185-e196. [doi: 10.1016/S2352-3018(20)30340-4] [Medline: 33794181]
38. Orkin C, Bernal Morell E, Tan DH, Katner H, Stellbrink H, Belonosova E, et al. Initiation of long-acting cabotegravir plus rilpivirine as direct-to-injection or with an oral lead-in in adults with HIV-1 infection: week 124 results of the open-label phase 3 FLAIR study. Lancet HIV 2021 Nov;8(11):e668-e678. [doi: 10.1016/S2352-3018(21)00184-3] [Medline: 34656207]
39. Swindells S, Andrade-Villanueva JF, Richmond GJ, Rizzardini G, Baumgarten A, Masiá M, et al. Long-acting cabotegravir and Rilpivirine for maintenance of HIV-1 suppression. N Engl J Med 2020 Mar 19;382(12):1112-1123. [doi: 10.1056/NEJMoa1904398] [Medline: 32130809]
40. Swindells S, Lutz T, Van Zyl L, Porteiro N, Stoll M, Mitha E, et al. Week 96 extension results of a Phase 3 study evaluating long-acting cabotegravir with rilpivirine for HIV-1 treatment. AIDS 2022 Feb 01;36(2):185-194 [FREE Full text] [doi: 10.1097/QAD.0000000000003025] [Medline: 34261093]
41. Jaeger H, Overton ET, Richmond G, Rizzardini G, Andrade-Villanueva JF, Mngqibisa R, et al. Long-acting cabotegravir and rilpivirine dosed every 2 months in adults with HIV-1 infection (ATLAS-2M), 96-week results: a randomised, multicentre, open-label, phase 3b, non-inferiority study. Lancet HIV 2021 Nov;8(11):e679-e689. [doi: 10.1016/S2352-3018(21)00185-5] [Medline: 34648734]
42. Mills A, Richmond GJ, Newman C, Osiyemi O, Cade J, Brinson C, et al. Long-acting cabotegravir and rilpivirine for HIV-1 suppression: switch to 2-monthly dosing after 5 years of daily oral therapy. AIDS 2022 Feb 01;36(2):195-203 [FREE Full text] [doi: 10.1097/QAD.0000000000003085] [Medline: 34652287]
43. Lazarus G, Wangsaputra VK, Christianto, Louisa M, Soetikno V, Hamers RL. Safety and pharmacokinetic profiles of long-acting injectable antiretroviral drugs for HIV-1 pre-exposure prophylaxis: a systematic review and meta-analysis of randomized trials. Front Pharmacol 2021;12:664875 [FREE Full text] [doi: 10.3389/fphar.2021.664875] [Medline: 34305587]
44. Thoueille P, Choong E, Cavassini M, Buclin T, Decosterd LA. Long-acting antiretrovirals: a new era for the management and prevention of HIV infection. J Antimicrob Chemother 2022 Feb 02;77(2):290-302 [FREE Full text] [doi: 10.1093/jac/dkab324] [Medline: 34499731]
45. Marzinke MA, Grinsztejn B, Fogel JM, Piwowar-Manning E, Li M, Weng L, et al. Characterization of human immunodeficiency virus (HIV) infection in cisgender men and transgender women who have sex with men receiving injectable cabotegravir for HIV prevention: HPTN 083. J Infect Dis 2021 Nov 16;224(9):1581-1592 [FREE Full text] [doi: 10.1093/infdis/jiab152] [Medline: 33740057]
46. Llibre JM, Hung CC, Brinson C, Castelli F, Girard PM, Kahl LP, et al. Efficacy, safety, and tolerability of dolutegravir-rilpivirine for the maintenance of virological suppression in adults with HIV-1: phase 3, randomised, non-inferiority SWORD-1 and SWORD-2 studies. Lancet 2018 Mar 03;391(10123):839-849. [doi: 10.1016/S0140-6736(17)33095-7] [Medline: 29310899]

47. U.S. Department of Health and Human Services. Guidelines for the use of antiretroviral agents in adults and adolescents with HIV. Clinical Info HIV.gov. URL: https://clinicalinfo.hiv.gov/en/guidelines/hiv-clinical-guidelines-adult-and-adolescent-arv/whats-new-guidelines [accessed 2023-05-01]
48. New milestone in HIV response: long-acting injectable drug significantly prevents HIV in women. United Nations. 2020 Nov 10. URL: https://news.un.org/zh/story/2020/11/1071372 [accessed 2023-05-01]
49. Philbin MM, Perez-Brumer A. Promise, perils and cautious optimism: the next frontier in long-acting modalities for the treatment and prevention of HIV. Curr Opin HIV AIDS 2022 Mar 01;17(2):72-88 [FREE Full text] [doi: 10.1097/COH.0000000000000723] [Medline: 35225248]
50. Segal-Maurer S, DeJesus E, Stellbrink HJ, Castagna A, Richmond GJ, Sinclair GI, CAPELLA Study Investigators. Capsid inhibition with lenacapavir in multidrug-resistant HIV-1 infection. N Engl J Med 2022 May 12;386(19):1793-1803. [doi: 10.1056/NEJMoa2115542] [Medline: 35544387]
51. Matthews RP, Patel M, Barrett SE, Haspeslagh L, Reynders T, Zhang S, et al. Safety and pharmacokinetics of islatravir subdermal implant for HIV-1 pre-exposure prophylaxis: a randomized, placebo-controlled phase 1 trial. Nat Med 2021 Oct;27(10):1712-1717. [doi: 10.1038/s41591-021-01479-3] [Medline: 34608329]

Abbreviations

AE: adverse event
ART: antiretroviral therapy
CAB-LA: long-acting cabotegravir
ISR: injection site reaction
LA-ARV: long-acting antiretroviral
PrEP: pre-exposure prophylaxis
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RAM: resistance-associated mutation
RPV-LA: long-acting rilpivirine
RR: risk ratio
SAE: serious adverse event
TDF-FTC: tenofovir disoproxil fumarate–emtricitabine
VF: virological failure
VL: viral load

Edited by A Mavragani, T Sanchez; submitted 13.03.23; peer-reviewed by S Hosek, T Crowell; comments to author 18.04.23; accepted 13.06.23; published 27.07.23.

Please cite as:
Wang W, Zhao S, Wu Y, Duan W, Li S, Li Z, Guo C, Wang W, Zhang T, Wu H, Huang X
Safety and Efficacy of Long-Acting Injectable Agents for HIV-1: Systematic Review and Meta-Analysis
JMIR Public Health Surveill 2023;9:e46767
URL: https://publichealth.jmir.org/2023/1/e46767
doi:10.2196/46767
PMID:37498645

Review

Association Between Socioeconomic Status and Adherence to Fecal Occult Blood Tests in Colorectal Cancer Screening Programs: Systematic Review and Meta-Analysis of Observational Studies

Zilin Luo[1,2*], BCh; Xuesi Dong[1,2*], PhD; Chenran Wang[1,2*], MM; Wei Cao[1,2], PhD; Yadi Zheng[1,2], BCh; Zheng Wu[1,2], BCh; Yongjie Xu[1,2], PhD; Liang Zhao[1,2], BCh; Fei Wang[1,2], PhD; Jibin Li[1], MPH; Jiansong Ren[1,2], PhD; Jufang Shi[1,2], PhD; Wanqing Chen[1,2], PhD; Ni Li[1,2,3], PhD

[1]Office of Cancer Screening, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]Chinese Academy of Medical Sciences Key Laboratory for National Cancer Big Data Analysis and Implement, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[3]Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Ni Li, PhD
Office of Cancer Screening
National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College
No. 17 Panjiayuan Nanli, Chaoyang District
Beijing, 100021
China
Phone: 86 18201116334
Fax: 86 010 8663546
Email: nli@cicams.ac.cn

Abstract

Background: Screening adherence is important in reducing colorectal cancer (CRC) incidence and mortality. Disparity in CRC screening adherence was observed in populations of different socioeconomic status (SES), but the direction and strength of the association remained unclear.

Objective: We aimed to systematically review all the observational studies that have analyzed the association between SES and adherence to organized CRC screening based on fecal occult blood tests.

Methods: We systematically reviewed the studies in PubMed, Embase, and Web of Science and reference lists of relevant reviews from the inception of the database up until June 7, 2023. Individual SES, neighborhood SES, and small-area SES were included, while any SES aggregated by geographic areas larger than neighbors were excluded. Studies assessing SES with any index or score combining indicators of income, education, deprivation, poverty, occupation, employment, marital status, cohabitation, and others were included. A random effect model meta-analysis was carried out for pooled odds ratios (ORs) and relative risks for adherence related to SES.

Results: Overall, 10 studies, with a total of 3,542,379 participants and an overall adherence rate of 64.9%, were included. Compared with low SES, high SES was associated with higher adherence (unadjusted OR 1.73, 95% CI 1.42-2.10; adjusted OR 1.53, 95% CI 1.28-1.82). In the subgroup of nonindividual-level SES, the adjusted association was significant (OR 1.57, 95% CI 1.26-1.95). However, the adjusted association was insignificant in the subgroup of individual-level SES (OR 1.46, 95% CI 0.98-2.17). As for subgroups of the year of print, not only was the unadjusted association significantly stronger in the subgroup of early studies (OR 1.97, 95% CI 1.59-2.44) than in the subgroup of late studies (OR 1.43, 95% CI 1.31-1.56), but also the adjusted one was significantly stronger in the early group (OR 1.86, 95% CI 1.43-2.42) than in the late group (OR 1.26, 95% CI 1.14-1.39), which was consistent and robust. Despite being statistically insignificant, the strength of the association seemed lower in studies that did not adjust for race and ethnicity (OR 1.31, 95% CI 1.21-1.43) than the overall estimate (OR 1.53, 95% CI 1.28-1.82).

Conclusions: The higher-SES population had higher adherence to fecal occult blood test–based organized CRC screening. Neighborhood SES, or small-area SES, was more competent than individual SES to be used to assess the association between SES and adherence. The disparity in adherence between the high SES and the low SES narrowed along with the development of interventions and the improvement of organized programs. Race and ethnicity were probably important confounding factors for the association.

(JMIR Public Health Surveill 2023;9:e48150) doi:10.2196/48150

KEYWORDS

adherence; colorectal cancer; fecal occult blood test; screening; socioeconomic status

Introduction

Worldwide, colorectal cancer (CRC) ranks third in cancer incidence and second in cancer mortality [1]. It has been confirmed that screening with fecal occult blood tests (FOBT) is effective in lowering CRC mortality in randomized controlled trials [2-4] and in the real world [5-7]. Unlike opportunistic screening programs, organized projects actively invite all individuals within a given age range in a specific geographic area to participate in protocol-based screening programs [8,9]. International academic organizations highly recommend the implementation of organized screening programs to reduce the burden of CRC [10-13].

The benefits of screening depend highly on screening compliance [14]. What is more, extensive research showed that the lower–socioeconomic status (SES) population had a higher incidence of CRC and even higher mortality due to poorer treatment [15-18], which meant there was a greater need for high screening adherence in the low SES to reduce health inequity. Although previous evidence suggested that inequities of income observed in opportunistic programs could largely be eliminated in organized programs [19], the association between SES and organized CRC screening adherence was inconsistent across studies [9].

A review conducted by de Klerk et al [9] qualitatively assessed the association of SES and individuals' adherence to FOBT-based organized CRC screening programs. This review found that among 11 programs, 90% (28/31) of publications reported lower adherence in the lower socioeconomic population. However, there were distinct differences in how SES was measured. Most used some types of indexes of deprivation, while some studies used more than one indicator to assess SES [9]. Currently, there were no data on the direction and strength of the association between SES and FOBT-based organized CRC screening adherence. In addition, although differences in adherence by SES and race and ethnicity had received much attention and racial disparities were partially explained by differences in SES [20], the role of race in the association of SES with adherence was controversial [21-23]. What is more, with socioeconomic disparity coming under the spotlight, efforts should be made to narrow it, and its changing trend deserves continued attention [9].

There is no conclusive evidence on the association between SES and adherence in FOBT-based organized screening. It is warranted to examine the association between SES and adherence, contributing to the improvement of screening adherence and equity to promote health equity in screening programs. Thus, the study aimed to perform a meta-analysis focusing on the association between SES and FOBT-based organized CRC screening adherence.

Methods

Search Strategy and Selection

We conducted the systematic review and meta-analysis under the direction of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [24]. We searched PubMed, Embase, and Web of Science from the inception date to June 7, 2023, using the search strategy: (socioeconomic OR "socioeconomic status" OR SES OR income OR education OR occupation OR insurance OR "social status") AND (colorecta* OR bowel) AND cancer* AND (screen* OR "early diagnosis" OR "early detection") AND (organized OR program*) AND (uptake OR adherence OR nonadherence OR compliance OR noncompliance OR participation OR participating OR nonparticipation OR attendance OR nonattendance OR engagement OR determine* OR factor* OR associate*). The integrated search strategy is shown in Table S1 in Multimedia Appendix 1, and references cited in review articles that were on similar topics were also checked.

Inclusion and Exclusion Criteria

First, titles and abstracts, and then full-text articles, were screened for inclusion by 2 investigators (ZL and CW) independently. We included studies that measure the association between SES and adherence to organized CRC screening programs. Inclusion and exclusion criteria are shown in Textbox 1.

Textbox 1. Inclusion and exclusion criteria of studies.

Inclusion criteria

- Original articles
- Observational studies conducted in organized colorectal cancer (CRC) screening programs
- Used fecal occult blood tests (FOBT) as the only primary screening test
- Average-risk population
- Assessed socioeconomic status (SES), including individual SES, neighborhood SES, and small-area SES
- Assessed SES with any index or score combining indicators of income, education, deprivation, poverty, occupation, employment, marital status, cohabitation, and others, and the SES index or score was defined as categorical variables
- Provided data to obtain related either unadjusted effect sizes (including odds ratios and relative risks) or adjusted ones
- Assessed adherence or nonadherence to FOBT-based CRC screening, which was objectively recorded
- Assessed adherence rate with the percentage of individuals invited to participate in screening that participate
- English language

Exclusion criteria

- Lacked enough data to calculate either unadjusted or adjusted effect sizes, including qualitative studies, reviews, commentaries, letters, and guidelines
- Conference abstracts that did not contain enough information for quality assessment
- Assessed adherence to colonoscopy or other screening tests
- Focused on only low-SES population or high-SES population, for example, focusing on only low-income population
- If several studies were conducted in the same program, participants should be counted only once. Duplicated samples were excluded, and the study with the biggest sample that offered the most information was included.
- Assessed country-level SES, province-level SES, or any SES aggregated by geographic areas larger than neighbors
- Assessed self-reported adherence
- Any language other than English

Data Extraction and Statistical Analysis

Sample data were extracted by 2 investigators (ZL and CW) independently, with disagreements solved by discussion and consensus. The standardized form consists of (1) the first author; (2) publication year; (3) study design; (4) country; (5) program name; (6) sample size; (7) adherence rate; (8) name of the SES index or score; (9) SES measures involved; (10) focusing on SES or not; (11) SES level; (12) original adjusted effect size; (13) available unadjusted and adjusted effect sizes (relative risks were transformed into odds ratios [ORs]) [25]; (14) covariates of adjusted effect sizes; (15) single-level or multilevel multivariate model; and (16) age range of screening. If effect sizes of an SES index or score with more than 2 levels were reported (usually), the effect sizes between the 2 extremes were extracted (eg, most vs least deprivation). The lowest SES was used as the reference group, and if the highest SES was used as the reference group, the reference group of the OR would be converted. When the unadjusted effect sizes were absent, they were calculated with the present cross-table data.

The Newcastle-Ottawa Quality Assessment Scale (NOS) for cohort studies and the Agency for Healthcare Research and Quality (AHRQ) methodology checklist were separately used to assess the quality of cohort and cross-sectional studies, which categorized the studies as high, middle, and low-quality [26-28]. The pair of investigators rated the quality independently, and discrepancies were solved through discussion and consensus. The Kappa coefficient was used to evaluate the consistency of the results.

Data Synthesis

ORs with corresponding 95% CIs were used for analysis, and an inverse variance method was used. We used Stata (version 17.0; StataCorp) to analyze data. The “flocci” command was used to convert the reference group of original estimates. “mean,” “meta bias,” and “meaning” were used to perform analysis. The Cochrane Q statistics and I^2 estimations were used to assess study heterogeneity. With considerable heterogeneity, a random effect model using the Daimonian and Laird method was used to conduct the meta-analysis.

We first assessed the unadjusted bivariate correlations between SES and adherence. Then we used the effect sizes of multivariate models with the most adjustment factors to assess the adjusted association. Tables and forest plots were used to display the results.

We examined heterogeneity using the Cochrane Q statistic and the I^2 statistic. Subgroup analysis was conducted to probe the potential source of heterogeneity across studies according to some characteristics of the studies: year of print (before or after the median year of print of the studies), adherence rate (lower or higher than 45%, which was the acceptable threshold set by the European guidelines [14]), SES focus (whether focusing on SES or other factors), SES level (individual level or nonindividual level, which includes neighborhood and small-area SES), and adjustment (adjusting for race and ethnicity or not, only for adjusted effect sizes) with at least 3 studies in a group. Publication bias was assessed using Egger and Begg tests [29,30]. Sensitivity analysis investigated the influence of each study (including the only cross-sectional study) on the overall meta-analysis summary estimate through a random effect model using the method of Daimonian and Laird.

Results

Systematic Review

The initial search yielded 1819 articles, and 9 articles were identified from reference lists. Of the 1819 articles, 910 remained after removing 909 duplicates by EndNote X8.1 (Thompson Reuters) automatically (Figure 1) [24,31]. In total, 102 full-text articles were assessed, and 10 met the inclusion and exclusion criteria.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram. CRC: colorectal cancer; FOBT: fecal occult blood test; SES: socioeconomic status.

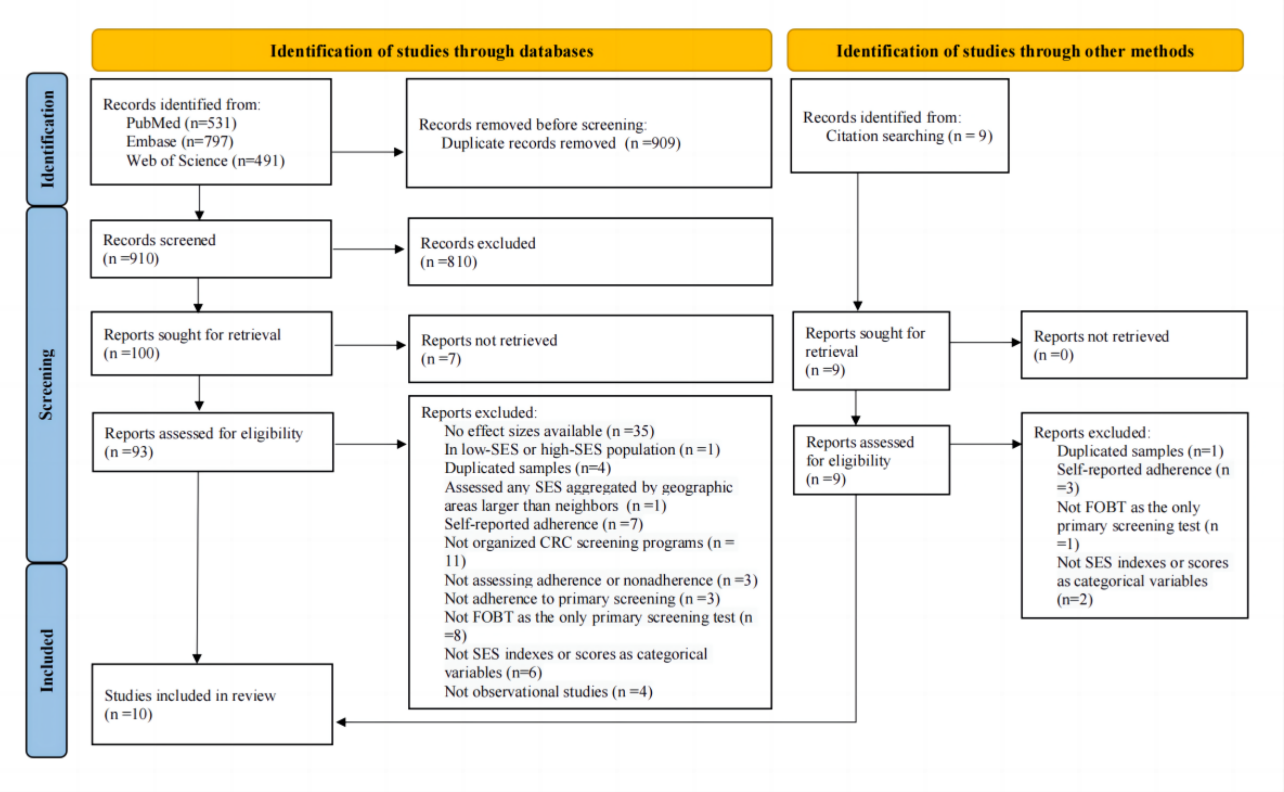


Table 1 shows an overview of the included studies, and Table S2 in Multimedia Appendix 1 lists other characteristics. A total of 9 of them were cohort studies, while 1 (515,388 participants) was cross-sectional. All studies included were carried out in developed European countries. All of the studies were of high quality, and the consistency of the literature quality evaluation results was 1.00 (Tables S3 and S4 in Multimedia Appendix 1). There was no uniform adjustment, and the potential confounders, including age, sex, race, ethnicity, cohabitation, marriage, and employment, were included and adjusted in different ways. Surprisingly, all ORs included were significantly larger than 1.00; in other words, all the studies agreed that SES was positively associated with adherence. Therefore, the results of the comparison of unadjusted and adjusted pooled estimates between subgroups were more important.

Table 1. Overview of the studies included in the meta-analysis.

Reference	Country	Program	SES[a] index or score	SES measures involved	SES focus[b]	SES level	Multivariate model
Poncet et al (2013) [32]	France	—[c]	Townsend Index	Employment, car ownership, home ownership, and overcrowded housing units	No	Nonindividual	Multilevel
Pornet et al (2010) [33]	France	—	Townsend Index	Employment, car ownership, home ownership, and overcrowded housing units	Yes	Nonindividual	Multilevel
Solís-Ibinagagoitia et al (2020) [34]	Spain	Bowel Cancer Screening Programme of the Basque Country	Deprivation Index	Education, occupation, and employment	No	Individual	Single-level
Steele et al (2010) [35]	United Kingdom	—	Scottish Index of Multiple Deprivation	Education, income, employment, housing, health, and access to facilities	Yes	Nonindividual	—
Szczepura et al (2008) [36]	United Kingdom	English Bowel Cancer Screening Pilot	The Carstairs Index of Deprivation	Employment, car ownership, overcrowding, and social class	No	Individual	Single-level
van der Meulen et al (2022) [37]	Netherlands	Dutch national CRC[d] screening programme	SES scores based on income, education, and employment	Education, income, and employment	Yes	Nonindividual	Single-level
van der Vlugt et al (2017) [38]	Netherlands	—	SES scores based on income, education, and employment	Income, education, employment, and position on the labor market	No	Nonindividual	—
Ward et al (2011) [39]	Australia	National Bowel Cancer Screening Program	Index of Relative Social Disadvantage	Education, income, and employment	No	Nonindividual	Single-level
Weller et al (2007) [40]	United Kingdom	UK Colorectal Cancer Screening Pilot	Index of Multiple Deprivation	Education, income, employment, health and disability, barriers to housing and services, crime, and living environment	No	Nonindividual	Single-level
Buron et al (2017) [41]	Spain	Barcelona colorectal cancer screening programme	Medea Deprivation Index	Education, occupation, and employment	Yes	Individual	Single-level

[a] SES: socioeconomic status.

[b] SES focus: whether focusing on socioeconomic status or other factors.

[c] Not available.

[d] CRC: colorectal cancer.

Meta-Analysis With Unadjusted and Adjusted Effect Sizes

A total of 9 studies provided unadjusted effect sizes, with 1,927,075 (63.7%) participants in 3,026,991 targeted individuals. And 8 studies reported adjusted effect sizes, involving 3,220,822 targeted individuals and 2,117,839 (65.8%) participants.

A positive and strong association was observed between SES and adherence (unadjusted OR 1.73, 95% CI 1.42-2.10; adjusted OR 1.53, 95% CI 1.28-1.82). The forest plots are shown in Figures 2 [32,33,35-41] and 3 [32-34,36,37,39-41]. There was no evidence of publication bias. Sensitivity analysis showed robust results (Figures S1 and S2 in Multimedia Appendix 1). A fair degree of heterogeneity was found.

Figure 2. Forest plot of total unadjusted effect sizes for the association between socioeconomic status and adherence. OR: odds ratio.

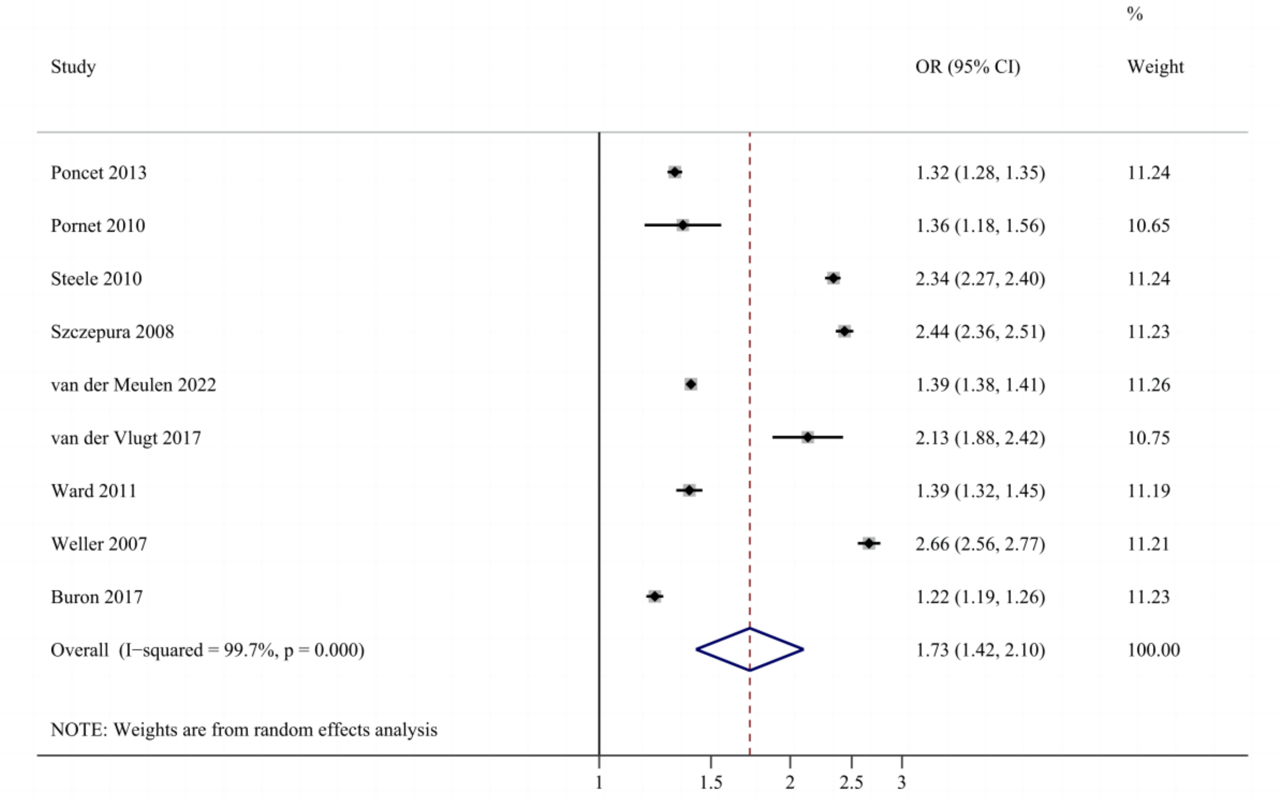
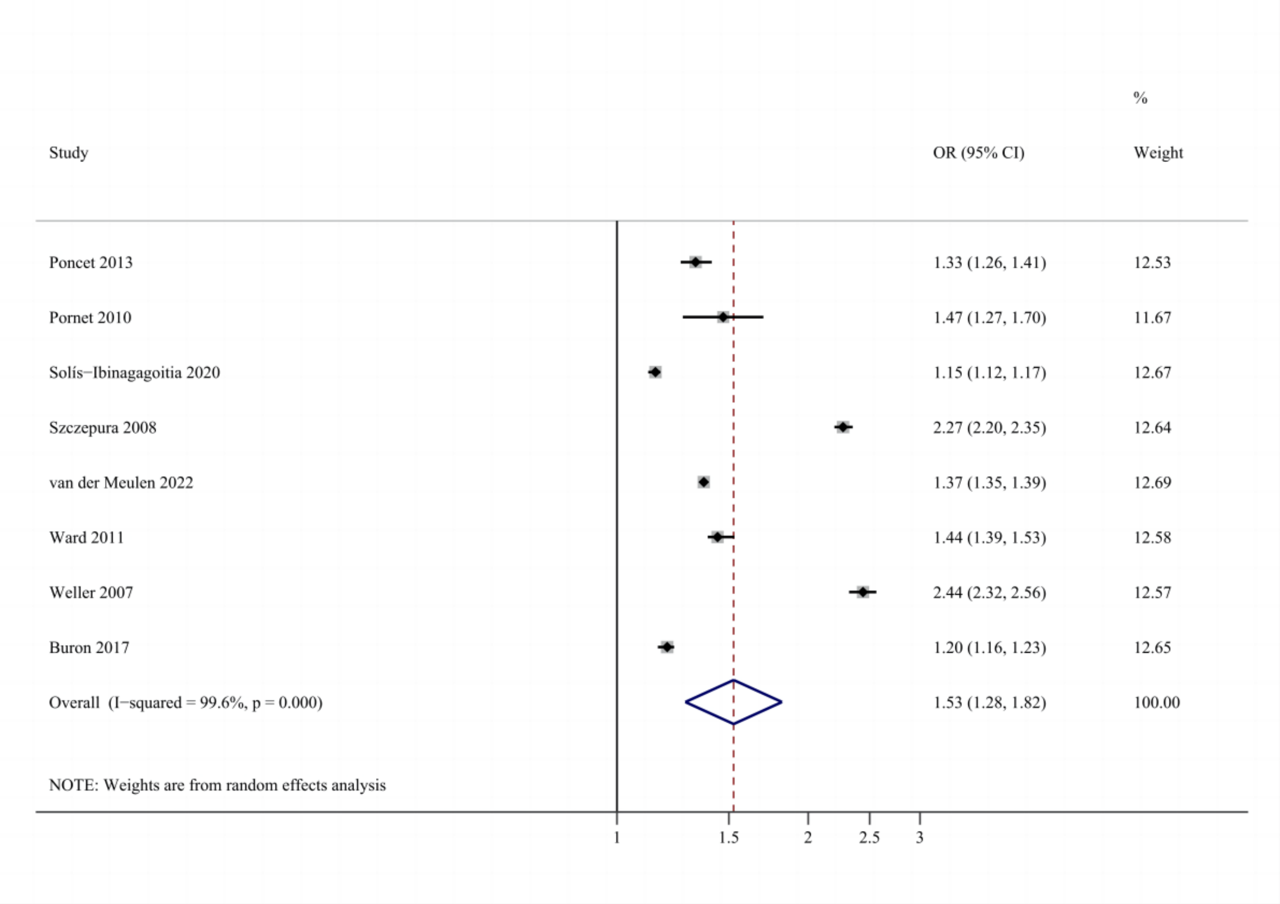


Figure 3. Forest plot of total adjusted effect sizes for the association between socioeconomic status and adherence. OR: odds ratio.



Subgroup Analyses With Unadjusted and Adjusted Effect Sizes

The positive association between SES and adherence remained significant in all feasible subgroup analyses (with at least three studies in a group), except for the adjusted association in the subgroup of individual-level SES (adjusted OR 1.46, 95% CI 0.98-2.17; Figures 4 and 5; Figure S3 in Multimedia Appendix 1). On the contrary, in subgroups of nonindividual-level SES, the adjusted association was significant. In other words, high nonindividual SES increased adherence, but individual SES did not influence adherence.

Figure 4. Forest plot of subgroup analysis using unadjusted effect sizes. There were only 2 studies in the subgroup of individual-level socioeconomic status (SES), which were not included in the subgroup analysis. *P* values less than .05 were considered statistically significant. OR: odds ratio; SES focus: whether focusing on socioeconomic status or other factors.

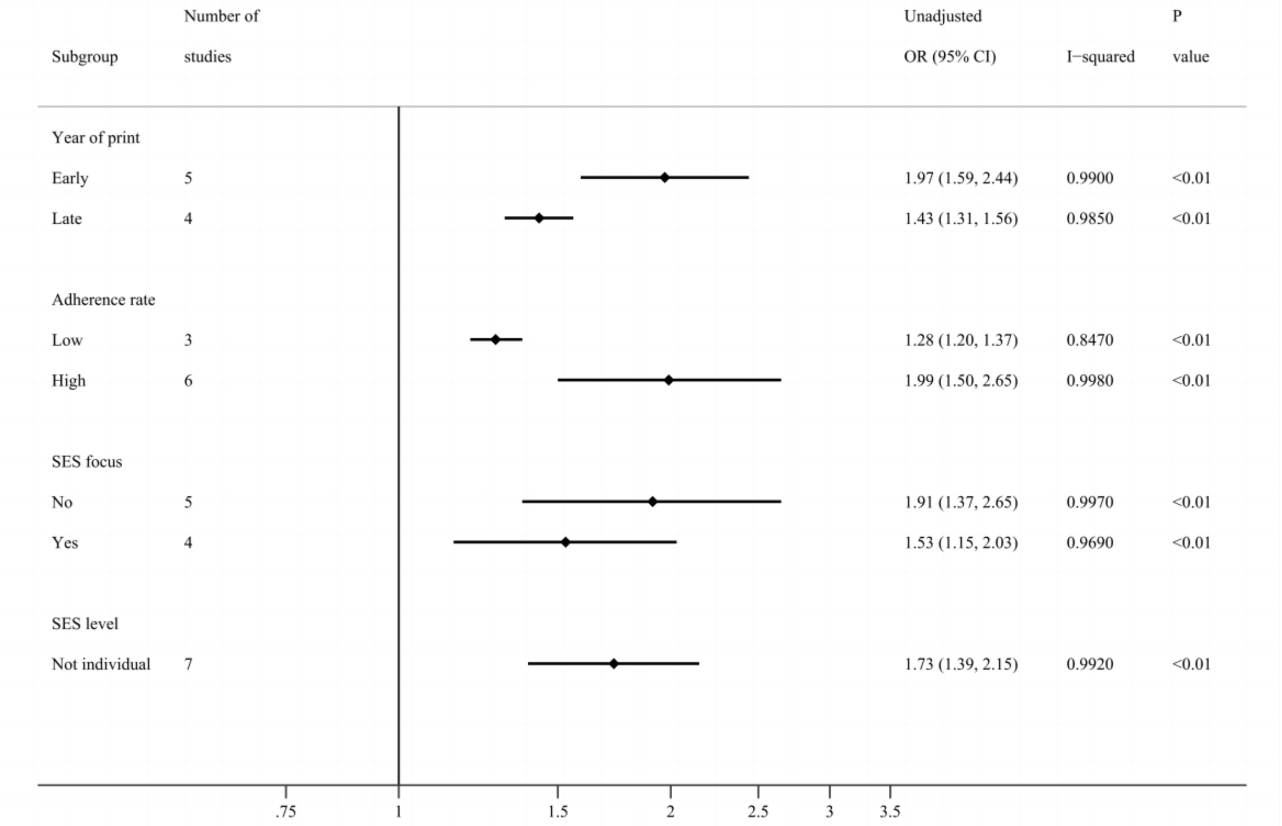
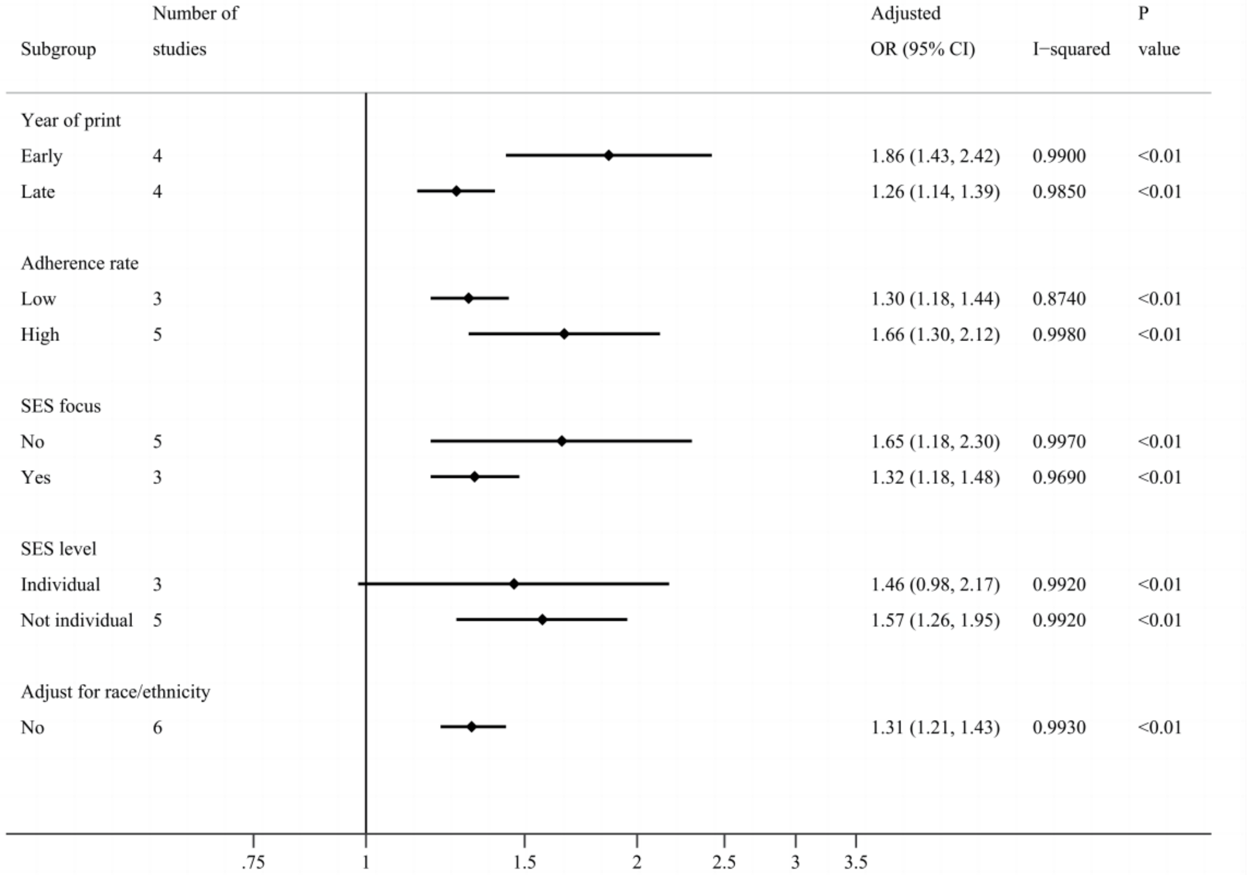


Figure 5. Forest plot of subgroup analysis using adjusted effect sizes. There were only 2 studies that adjusted for race and ethnicity, which were not included in the subgroup analysis. OR: odds ratio; SES focus: whether focusing on socioeconomic status or other factors.



As for subgroups of the year of print, not only was the unadjusted association significantly stronger in the subgroup of early studies (OR 1.97, 95% CI 1.59-2.44) than in late studies (OR 1.43, 95% CI 1.31-1.56; Figure S4 in Multimedia Appendix 1), but also the adjusted one was significantly stronger in the early group (OR 1.86, 95% CI 1.43-2.42) than the late group (OR 1.26, 95% CI 1.14-1.39; Figure S5 in Multimedia Appendix 1), which was consistent and robust. In addition, the overall adherence rate of the late studies was higher than the early ones, which were 67.2% versus 52.1% in unadjusted associations and 68.1% versus 50.1% in adjusted associations.

The other results were not statistically significant, but they provided some insights. For multivariate analysis, the pooled estimate of the subgroup of studies that did not adjust for race and ethnicity (OR 1.31, 95% CI 1.21-1.43) seemed smaller than the overall pooled estimate (OR 1.53, 95% CI 1.28-1.82). The adjustment for race and ethnicity better showed a positive association. In addition, the strength of the association between SES and adherence seemed lower when the adherence rate was low or when SES was the focus of the studies (unadjusted OR 1.53, 95% CI 1.15-2.03; adjusted OR 1.32, 95% CI 1.18-1.48) than not (unadjusted OR 1.91, 95% CI 1.37-2.65; adjusted OR 1.65, 95% CI 1.18-1.48).

Discussion

Meta-Analytic Findings

To our knowledge, this is the first systematic review and meta-analysis assessing the strength of the association between SES and FOBT-based organized CRC screening adherence. It indicated considerably higher adherence among the highest SES compared with the lowest. The positive association was significant in both unadjusted and adjusted analyses, while being consistent across most possible sets of pooled estimates and all possible subgroup analyses, including the year of print, adherence rate, SES focus, SES level, and adjustment for race and ethnicity.

The association between high SES and high adherence was similar to those reported in 4 previous reviews investigating the participation of multiple modalities of CRC screening in both organized and opportunistic screening programs [42-45]. Furthermore, the disparity in adherence would lead to disparity in CRC incidence and mortality. In a modeling study estimating the disparity between African American and White individuals in the United States, the disparities in uptake of screening were able to account for 42% of the disparity in incidence and 19% of the disparity in mortality [46]. It thus seemed urgent to decrease the disparity in adherence to reduce health inequity driven by SES.

Differences in Subgroup Analysis

Subgroup analysis of different SES levels indicated that individual-level SES was not associated with adherence, which suggested that different SES levels could be the source of heterogeneity between studies. A systematic review by Mosquera et al [47] indicated that different indicators of individual SES showed various associations with adherence, separately. While individual educational level was positively associated with screening participation and higher deprivation was associated with lower participation, studies differed in the results of the association between individual income and participation [47]. In this meta-analysis that assessed SES index and scores rather than single indicators, individual SES was not associated with adherence after adjustment.

Nevertheless, country-level SES, which was not involved in this meta-analysis, was confirmed to not be associated with adherence either. In 2009, Pruitt et al [48] conducted a systematic review of the association between area SES and adherence. They found that 3 of 5 studies had at least one positive association and noted that country-level SES variables were not associated with any type of CRC screening adherence. Therefore, adherence might be associated with neighborhood or small area–level SES rather than individual- or country-level [48].

Due to the difficulty and trouble of collecting individual-level data, commonly used measures of SES were based on geography, at different levels of aggregation from countries to neighborhoods [49]. And it was proven that the agreement between area- and individual-level SES was low [50]. As a result, neighborhood SES or small-area SES was more competent to be used to assess the association between SES and adherence, and in this case, SES was significantly associated with adherence.

Although socioeconomic and racial and ethnic disparities in organized CRC screening were noted [9,51], the effect of race and ethnicity on the association has not been concluded yet. Another important finding of this analysis was that race and ethnicity might be important confounding factors for the association between SES and adherence. In accordance with the present results, a previous study in the United States using income level as the measure of SES reported that in the hierarchical analysis by race and ethnicity, there was no significant association between SES and adherence among the non-White population. Especially for adherence to a screening colonoscopy, a significant interaction between race and ethnicity and income level was identified [52]. Although O'Malley et al [21] declared that racial differences in adherence could be fully explained by differences in SES, most studies of other conditions found independent associations with both SES and race and ethnicity [22,23]. Consequently, in order to improve adherence and address inequity, future research should consider strategies to improve the adherence of the low-SES population regarding ethnic composition.

The results of this meta-analysis showed that the disparity between the high and low-SES populations in adherence decreased. Since the adherence rate of the subgroup with a later year of print was higher than that of the other subgroup, a probable explanation was that adherence of the low-SES population increased more than the high as organizational screening programs developed with more and more interventions to promote adherence. Actually, a large-sample randomized controlled trial proved that certain interventions could not only improve overall participation but also particularly enhance the adherence of the low-SES population [53]. In this context, future studies should focus on cost-effective strategies that are suitable to be integrated into large-scale organized programs and further address the disparity in the near future. Our ideal goal is not equal adherence but to eliminate the health inequity of the CRC.

Strengths and Limitations

This is the first meta-analysis assessing the association between SES and adherence, especially with the first focus on FOBT-based organized CRC screening adherence. Previous reviews merely focused on the qualitative association between SES and adherence and did not deal with the influence of confounding factors. However, this study confirmed, quantified, and further explored the positive association between SES and adherence by performing meta-analysis and subgroup analysis and comprehensively interpreting the results of both the unadjusted association and the adjusted one.

A few limitations should be considered when interpreting the results. First, considerable heterogeneity existed and although we performed subgroup analysis and sensitivity analysis, it remained. The probable explanation of the heterogeneity was that in the absence of a widely recognized SES index, we observed different kinds of indexes and scores, which combined different indicators at different levels and used different weights to indicate SES. However, most indexes and scores incorporate the 3 SES indicators of income, education, and employment. The different start ages and stop ages used by the different programs resulted in differences in the mean age of the population, leading to differences in sample characteristics. Second, we recognized that all the studies included were carried out in the high-income setting even though we did not make a restriction on the inclusion criteria and actively supplement the literature collection from the reference lists. It might be caused by a limited number of studies that quantitatively assessed the association between SES and adherence to organized screening programs, especially in low-income countries where organized screening was not widespread. Therefore, whether the results can be generalized to low- and middle-income countries remains unknown. Finally, different adjustments might cause residual confounders in the included studies. To minimize this effect, we took both unadjusted and adjusted effect sizes into account and were more confident to draw a conclusion when the results of them were consistent.

Conclusions

To conclude, we found the high-SES population had higher adherence to FOBT-based organized CRC screening. Neighborhood SES or small-area SES was more competent than individual- and country-level SES to be used to assess the association between SES and adherence. Race and ethnicity were probably important confounding factors for the association. The good news was that the disparity of adherence between the high SES and the low SES narrowed with the development of interventions and the improvement of organized programs. Future research should focus on digging into the causation of the association and creating targeted interventions and strategies to improve the adherence of the low-SES population, which aims at overcoming the inequality in the chances of benefiting from organized screening programs.

Acknowledgments

This study is funded by the CAMS Innovation Fund for Medical Science (grants 2021-I2M-1-067 and 2021-I2M-1-011), Sanming Project of Medicine in Shenzhen (grant SZSM201911015), and the Hope Star of Talent Program of National Cancer Center/ Cancer Hospital of the Chinese Academy of Medical Sciences.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

ZL, XD, CW, and NL conceived and designed the study; WC, YZ, and ZW provided study materials and tools; ZL, XD, and CW were responsible for the collection and assembly of data; LZ and FW were responsible for data analysis; JL, JR, and JS were responsible for interpretation; ZL and CW were involved in writing the manuscript; YX, WC, and NL revised the manuscript. All the work was performed under NL’s instructions. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Tables and Figures.
[DOCX File , 955 KB - publichealth_v9i1e48150_app1.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota colon cancer control study. N Engl J Med 1993;328(19):1365-1371 [FREE Full text] [doi: 10.1056/NEJM199305133281901] [Medline: 8474513]
3. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348(9040):1472-1477 [FREE Full text] [doi: 10.1016/S0140-6736(96)03386-7] [Medline: 8942775]
4. Kronborg O, Fenger C, Olsen J, Jørgensen OD, Søndergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348(9040):1467-1471 [FREE Full text] [doi: 10.1016/S0140-6736(96)03430-7] [Medline: 8942774]
5. Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut 2015;64(5):784-790 [FREE Full text] [doi: 10.1136/gutjnl-2014-307508] [Medline: 25179811]
6. Libby G, Brewster DH, McClements PL, Carey FA, Black RJ, Birrell J, et al. The impact of population-based faecal occult blood test screening on colorectal cancer mortality: a matched cohort study. Br J Cancer 2012;107(2):255-259 [FREE Full text] [doi: 10.1038/bjc.2012.277] [Medline: 22735907]
7. Faivre J, Dancourt V, Lejeune C, Tazi MA, Lamour J, Gerard D, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. Gastroenterology 2004;126(7):1674-1680 [FREE Full text] [doi: 10.1053/j.gastro.2004.02.018] [Medline: 15188160]
8. Zhang L, Carvalho AL, Mosquera I, Wen T, Lucas E, Sauvaget C, et al. An international consensus on the essential and desirable criteria for an 'organized' cancer screening programme. BMC Med 2022;20(1):101 [FREE Full text] [doi: 10.1186/s12916-022-02291-7] [Medline: 35317783]
9. de Klerk CM, Gupta S, Dekker E, Essink-Bot ML, Expert Working Group ‘Coalition to reduce inequities in colorectal cancer screening’ of the World Endoscopy Organization. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. Gut 2018;67(4):679-687 [FREE Full text] [doi: 10.1136/gutjnl-2016-313311] [Medline: 28073892]

10. Levin B, Lieberman DA, McFarland B, Smith RA, Brooks D, Andrews KS, American Cancer Society Colorectal Cancer Advisory Group, US Multi-Society Task Force, American College of Radiology Colon Cancer Committee. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on colorectal cancer, and the American college of radiology. CA Cancer J Clin 2008;58(3):130-160 [FREE Full text] [doi: 10.3322/CA.2007.0018] [Medline: 18322143]
11. European Colorectal Cancer Screening Guidelines Working Group, von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy 2013;45(1):51-59 [FREE Full text] [doi: 10.1055/s-0032-1325997] [Medline: 23212726]
12. Sung JJY, Ng SC, Chan FKL, Chiu HM, Kim HS, Matsuda T, Asia Pacific Working Group. An updated Asia pacific consensus recommendations on colorectal cancer screening. Gut 2015;64(1):121-132 [FREE Full text] [doi: 10.1136/gutjnl-2013-306503] [Medline: 24647008]
13. Canadian Task Force on Preventive Health Care. Colorectal cancer screening. Recommendation statement from the Canadian Task Force on preventive health care. CMAJ 2001;165(2):206-208 [FREE Full text] [Medline: 11501466]
14. Moss S, Ancelle-Park R, Brenner H, International Agency for Research on Cancer. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First edition—evaluation and interpretation of screening outcomes. Endoscopy 2012;44(Suppl 3):SE49-SE64 [FREE Full text] [doi: 10.1055/s-0032-1309788] [Medline: 23012122]
15. Ladabaum U, Clarke CA, Press DJ, Mannalithara A, Myer PA, Cheng I, et al. Colorectal cancer incidence in Asian populations in California: effect of nativity and neighborhood-level factors. Am J Gastroenterol 2014;109(4):579-588 [FREE Full text] [doi: 10.1038/ajg.2013.488] [Medline: 24492754]
16. Tao L, Ladabaum U, Gomez SL, Cheng I. Colorectal cancer mortality among Hispanics in California: differences by neighborhood socioeconomic status and nativity. Cancer 2014;120(22):3510-3518 [FREE Full text] [doi: 10.1002/cncr.28837] [Medline: 25042119]
17. Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. Br J Cancer 2002;87(6):585-590 [FREE Full text] [doi: 10.1038/sj.bjc.6600515] [Medline: 12237766]
18. McLeod A. Variation in the provision of chemotherapy for colorectal cancer. J Epidemiol Community Health 1999;53(12):775-781 [FREE Full text] [doi: 10.1136/jech.53.12.775] [Medline: 10656086]
19. Decker KM, Demers AA, Nugent Z, Biswanger N, Singh H. Reducing income-related inequities in colorectal cancer screening: lessons learned from a retrospective analysis of organised programme and non-programme screening delivery in Winnipeg, Manitoba. BMJ Open 2016;6(2):e009470 [FREE Full text] [doi: 10.1136/bmjopen-2015-009470] [Medline: 26908517]
20. Liss DT, Baker DW. Understanding current racial/ethnic disparities in colorectal cancer screening in the United States: the contribution of socioeconomic status and access to care. Am J Prev Med 2014;46(3):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.10.023] [Medline: 24512861]
21. O'Malley AS, Forrest CB, Feng S, Mandelblatt J. Disparities despite coverage: gaps in colorectal cancer screening among medicare beneficiaries. Arch Intern Med 2005;165(18):2129-2135 [FREE Full text] [doi: 10.1001/archinte.165.18.2129] [Medline: 16217003]
22. Dharni N, Armstrong D, Chung-Faye G, Wright AJ. Factors influencing participation in colorectal cancer screening-a qualitative study in an ethnic and socio-economically diverse inner city population. Health Expect 2017;20(4):608-617 [FREE Full text] [doi: 10.1111/hex.12489] [Medline: 27550367]
23. Orbell S, Szczepura A, Weller D, Gumber A, Hagger MS. South Asian ethnicity, socioeconomic status, and psychological mediators of faecal occult blood colorectal screening participation: a prospective test of a process model. Health Psychol 2017;36(12):1161-1172. [doi: 10.1037/hea0000525] [Medline: 28726477]
24. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
25. Viera AJ. Odds ratios and risk ratios: what's the difference and why does it matter? South Med J 2008;101(7):730-734 [FREE Full text] [doi: 10.1097/SMJ.0b013e31817a7ee4] [Medline: 18580722]
26. Wells GA, Shea B, O'Connell D, Robertson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomized Studies in MetaAnalysis. Ottawa Hospital Research Institute. 2012. URL: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2012-06-15]
27. Wells GS, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. NewCastle-Ottawa Quality Assessment Scale: cohort studies. Ottawa Hospital Research Institute. 2012. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2012-06-15]
28. Rostom A, Dube C, Cranney A, Saloojee N, Sy R, Garritty C, et al. 104: Appendix D. quality assessment forms. In: AHRQ Evidence Report Summaries. Rockville, MD: Agency for Healthcare Research and Quality (US); 2004.
29. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]

30. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50(4):1088-1101 [FREE Full text] [Medline: 7786990]
31. PRISMA. URL: http://www.prisma-statement.org/?AspxAutoDetectCookieSupport=1 [accessed 2023-09-13]
32. Poncet F, Delafosse P, Seigneurin A, Exbrayat C, Colonna M. Determinants of participation in organized colorectal cancer screening in Isère (France). Clin Res Hepatol Gastroenterol 2013;37(2):193-199 [FREE Full text] [doi: 10.1016/j.clinre.2012.04.011] [Medline: 22704817]
33. Pornet C, Dejardin O, Morlais F, Bouvier V, Launoy G. Socioeconomic determinants for compliance to colorectal cancer screening. a multilevel analysis. J Epidemiol Community Health 2010;64(4):318-324 [FREE Full text] [doi: 10.1136/jech.2008.081117] [Medline: 19740776]
34. Solís-Ibinagagoitia M, Unanue-Arza S, Díaz-Seoane M, Martínez-Indart L, Lebeña-Maluf A, Idigoras I, et al. Factors related to non-participation in the basque country colorectal cancer screening programme. Front Public Health 2020;8:604385 [FREE Full text] [doi: 10.3389/fpubh.2020.604385] [Medline: 33363095]
35. Steele RJC, Kostourou I, McClements P, Watling C, Libby G, Weller D, et al. Effect of gender, age and deprivation on key performance indicators in a FOBT-based colorectal screening programme. J Med Screen 2010;17(2):68-74 [FREE Full text] [doi: 10.1258/jms.2010.009120] [Medline: 20660434]
36. Szczepura A, Price C, Gumber A. Breast and bowel cancer screening uptake patterns over 15 years for UK south Asian ethnic minority populations, corrected for differences in socio-demographic characteristics. BMC Public Health 2008;8:346 [FREE Full text] [doi: 10.1186/1471-2458-8-346] [Medline: 18831751]
37. van der Meulen MP, Toes-Zoutendijk E, Spaander MCW, Dekker E, Bonfrer JMG, van Vuuren AJ, et al. Socioeconomic differences in participation and diagnostic yield within the Dutch national colorectal cancer screening programme with faecal immunochemical testing. PLoS One 2022;17(2):e0264067 [FREE Full text] [doi: 10.1371/journal.pone.0264067] [Medline: 35176116]
38. van der Vlugt M, Grobbee EJ, Bossuyt PM, Bongers E, Spijker W, Kuipers EJ, et al. Adherence to colorectal cancer screening: four rounds of faecal immunochemical test-based screening. Br J Cancer 2017;116(1):44-49 [FREE Full text] [doi: 10.1038/bjc.2016.399] [Medline: 27923037]
39. Ward PR, Javanparast S, Matt MA, Martini A, Tsourtos G, Cole S, et al. Equity of colorectal cancer screening: cross-sectional analysis of national bowel cancer screening program data for South Australia. Aust N Z J Public Health 2011;35(1):61-65 [FREE Full text] [doi: 10.1111/j.1753-6405.2010.00637.x] [Medline: 21299702]
40. Weller D, Coleman D, Robertson R, Butler P, Melia J, Campbell C, et al. The UK colorectal cancer screening pilot: results of the second round of screening in England. Br J Cancer 2007;97(12):1601-1605 [FREE Full text] [doi: 10.1038/sj.bjc.6604089] [Medline: 18026197]
41. Buron A, Auge JM, Sala M, Román M, Castells A, Macià F, PROCOLON research group. Association between socioeconomic deprivation and colorectal cancer screening outcomes: low uptake rates among the most and least deprived people. PLoS One 2017;12(6):e0179864 [FREE Full text] [doi: 10.1371/journal.pone.0179864] [Medline: 28622365]
42. Moons L, Mariman A, Vermeir P, Colemont L, Clays E, Van Vlierberghe H, et al. Sociodemographic factors and strategies in colorectal cancer screening: a narrative review and practical recommendations. Acta Clin Belg 2020;75(1):33-41 [FREE Full text] [doi: 10.1080/17843286.2018.1563736] [Medline: 30609904]
43. Wools A, Dapper EA, de Leeuw JRJ. Colorectal cancer screening participation: a systematic review. Eur J Public Health 2016;26(1):158-168 [FREE Full text] [doi: 10.1093/eurpub/ckv148] [Medline: 26370437]
44. Javanparast S, Ward P, Young G, Wilson C, Carter S, Misan G, et al. How equitable are colorectal cancer screening programs which include FOBTs? A review of qualitative and quantitative studies. Prev Med 2010;50(4):165-172 [FREE Full text] [doi: 10.1016/j.ypmed.2010.02.003] [Medline: 20153359]
45. Gimeno García AZ. Factors influencing colorectal cancer screening participation. Gastroenterol Res Pract 2012;2012:483417 [FREE Full text] [doi: 10.1155/2012/483417] [Medline: 22190913]
46. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, van Ballegooijen M, Zauber AG, Jemal A. Contribution of screening and survival differences to racial disparities in colorectal cancer rates. Cancer Epidemiol Biomarkers Prev 2012;21(5):728-736 [FREE Full text] [doi: 10.1158/1055-9965.EPI-12-0023] [Medline: 22514249]
47. Mosquera I, Mendizabal N, Martín U, Bacigalupe A, Aldasoro E, Portillo I, Desberdinak Group. Inequalities in participation in colorectal cancer screening programmes: a systematic review. Eur J Public Health 2020;30(3):416-425 [FREE Full text] [doi: 10.1093/eurpub/ckz236] [Medline: 32361732]
48. Pruitt SL, Shim MJ, Mullen PD, Vernon SW, Amick BC. Association of area socioeconomic status and breast, cervical, and colorectal cancer screening: a systematic review. Cancer Epidemiol Biomarkers Prev 2009;18(10):2579-2599 [FREE Full text] [doi: 10.1158/1055-9965.EPI-09-0135] [Medline: 19815634]
49. Carethers JM, Doubeni CA. Causes of socioeconomic disparities in colorectal cancer and intervention framework and strategies. Gastroenterology 2020;158(2):354-367 [FREE Full text] [doi: 10.1053/j.gastro.2019.10.029] [Medline: 31682851]
50. Buajitti E, Chiodo S, Rosella LC. Agreement between area- and individual-level income measures in a population-based cohort: implications for population health research. SSM Popul Health 2020;10:100553 [FREE Full text] [doi: 10.1016/j.ssmph.2020.100553] [Medline: 32072008]

51. Deutekom M, van Rijn AF, Dekker E, Blaauwgeers H, Stronks K, Fockens P, et al. Uptake of faecal occult blood test colorectal cancer screening by different ethnic groups in the Netherlands. Eur J Public Health 2009;19(4):400-402 [FREE Full text] [doi: 10.1093/eurpub/ckp051] [Medline: 19372193]
52. Ananthakrishnan AN, Schellhase KG, Sparapani RA, Laud PW, Neuner JM. Disparities in colon cancer screening in the medicare population. Arch Intern Med 2007;167(3):258-264 [FREE Full text] [doi: 10.1001/archinte.167.3.258] [Medline: 17296881]
53. Wardle J, von Wagner C, Kralj-Hans I, Halloran SP, Smith SG, McGregor LM, et al. Effects of evidence-based strategies to reduce the socioeconomic gradient of uptake in the english NHS bowel cancer screening programme (ASCEND): four cluster-randomised controlled trials. Lancet 2016;387(10020):751-759 [FREE Full text] [doi: 10.1016/S0140-6736(15)01154-X] [Medline: 26680217]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
CRC: colorectal cancer
FOBT: fecal occult blood test
NOS: Newcastle-Ottawa Quality Assessment Scale
OR: odds ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SES: socioeconomic status

Edited by A Mavragani, T Sanchez; submitted 15.04.23; peer-reviewed by Z Su, X Zhou; comments to author 01.06.23; revised version received 21.07.23; accepted 31.07.23; published 31.10.23.

Please cite as:
Luo Z, Dong X, Wang C, Cao W, Zheng Y, Wu Z, Xu Y, Zhao L, Wang F, Li J, Ren J, Shi J, Chen W, Li N
Association Between Socioeconomic Status and Adherence to Fecal Occult Blood Tests in Colorectal Cancer Screening Programs: Systematic Review and Meta-Analysis of Observational Studies
JMIR Public Health Surveill 2023;9:e48150
URL: https://publichealth.jmir.org/2023/1/e48150
doi:10.2196/48150
PMID:37906212

Review

The Use of Machine Translation for Outreach and Health Communication in Epidemiology and Public Health: Scoping Review

Paula Sofia Herrera-Espejel[1,2], MPP; Stefan Rach[1,2], PhD

[1]Department Epidemiological Methods and Etiological Research, Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany
[2]Leibniz ScienceCampus Digital Public Health, Bremen, Germany

Corresponding Author:
Stefan Rach, PhD
Department Epidemiological Methods and Etiological Research
Leibniz Institute for Prevention Research and Epidemiology - BIPS
Achterstr 30
Bremen, 28359
Germany
Phone: 49 421 218 56 841
Email: rach@leibniz-bips.de

Abstract

Background: Culturally and linguistically diverse groups are often underrepresented in population-based research and surveillance efforts, leading to biased study results and limited generalizability. These groups, often termed “hard-to-reach,” commonly encounter language barriers in the public health (PH) outreach material and information campaigns, reducing their involvement with the information. As a result, these groups are challenged by 2 effects: the medical and health knowledge is less tailored to their needs, and at the same time, it is less accessible for to them. Modern machine translation (MT) tools might offer a cost-effective solution to PH material language accessibility problems.

Objective: This scoping review aims to systematically investigate current use cases of MT specific to the fields of PH and epidemiology, with a particular interest in its use for population-based recruitment methods.

Methods: PubMed, PubMed Central, Scopus, ACM Digital Library, and IEEE Xplore were searched to identify articles reporting on the use of MT in PH and epidemiological research for this PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews)–compliant scoping review. Information on communication scenarios, study designs and the principal findings of each article were mapped according to a settings approach, the *World Health Organization monitoring and evaluation* framework and the *service readiness level* framework, respectively.

Results: Of the 7186 articles identified, 46 (0.64%) were included in this review, with the earliest study dating from 2009. Most of the studies (17/46, 37%) discussed the application of MT to existing PH materials, limited to one-way communication between PH officials and addressed audiences. No specific article investigated the use of MT for recruiting linguistically diverse participants to population-based studies. Regarding study designs, nearly three-quarters (34/46, 74%) of the articles provided technical assessments of MT from 1 language (mainly English) to a few others (eg, Spanish, Chinese, or French). Only a few (12/46, 26%) explored end-user attitudes (mainly of PH employees), whereas none examined the legal or ethical implications of using MT. The experiments primarily involved PH experts with language proficiencies. Overall, more than half (38/70, 54% statements) of the summarizing results presented mixed and inconclusive views on the technical readiness of MT for PH information.

Conclusions: Using MT in epidemiology and PH can enhance outreach to linguistically diverse populations. The translation quality of current commercial MT solutions (eg, Google Translate and DeepL Translator) is sufficient if postediting is a mandatory step in the translation workflow. Postediting of legally or ethically sensitive material requires staff with adequate content knowledge in addition to sufficient language skills. Unsupervised MT is generally not recommended. Research on whether machine-translated texts are received differently by addressees is lacking, as well as research on MT in communication scenarios that warrant a response from the addressees.

(JMIR Public Health Surveill 2023;9:e50814) doi:10.2196/50814

KEYWORDS

machine translation; public health; epidemiology; population-based; recruitment; outreach; multilingual; culturally and linguistically diverse communities

Introduction

Background

Public health (PH) and epidemiology are increasingly challenged by decreasing response proportions, in general, and an underrepresentation of culturally and linguistically diverse (CALD) communities, in particular [1]. Such underrepresentation increases the risk of biased estimates and, therefore, might limit the generalizability of findings in population-based research [2-7]. Ultimately, it might hinder the inclusion and involvement of these communities in disease prevention, surveillance efforts, and emergency response. In PH outreach and information campaigns, reaching CALD populations often poses greater difficulties than reaching other groups. As a result, these groups are challenged by 2 effects: the medical and health knowledge is less tailored to their needs, and at the same time, it is less accessible to them. These effects will only increase in importance as migration owing to globalization, global conflict, and economic inequalities increasingly shapes our societies toward multiculturality.

Using personalized recruitment material is an effective approach to engage individuals from CALD groups in population-based studies [8]. The choice of language matters because language barriers often result in their disengagement with PH initiatives [9-11]. If recipients are not able to comprehend transmitted information in the first place, they cannot react to it or provide an informed response [12]. Inclusive outreach approaches in PH study material, such as simplifying technical language or using multilingual cover letters, have been proven to improve access to information, foster meaningful participation, and reduce study nonresponse [13-15].

PH officials and researchers often struggle to effectively reach and engage all target audiences evenly. Although sufficient knowledge about the cultural composition of the target populations may be available (ie, the necessity to use particular languages), budget limitations usually restrict how many professional translations can be prepared and used for PH communication and outreach efforts to start with. A further complication with printed outreach material is that the number of different language versions that can be sent out in a single letter is physically limited, but the preferred language of an individual often is not known; therefore, it is difficult to conduct targeted outreach with specific language versions tailored to each recipient.

The use of machine translation (MT) technology poses a potential solution to overcome language hurdles in multilingual populations and improve effective material dissemination. As a computerized system, MT is able to automatically translate text or speech from 1 source language to multiple output languages [16]. In clinical settings, the technology has already been used to lower language barriers and facilitate services independently of the spoken language of the physician [17]. In the context of PH and epidemiology, MT could also be used to increase outreach by providing cross-lingual access to information and supporting PH staff to optimize material translation workflows.

Prior Work

To our knowledge, there are 4 recent systematic reviews that cover aspects of the use of translation technologies in medical and clinical settings. In 2018, Dew et al [18] published a review on how the development of MT technology could be useful to assist one-way communication among individual stakeholders. In 2020, Frampton et al [19] systematically mapped digital tools for the recruitment and retention of participants in randomized controlled trials. Although the authors did not specifically address MT or similar language technologies, one of their main takeaways was that few studies address its use to support underserved groups. A year later, Thonon et al [20] published a review on the use of mobile apps to facilitate dialogue between health care professionals and CALD individuals with low language proficiency levels. In 2022, Vieira [21] published a review with a focus on the use of MT in medical and legal settings as 2 separate cases of translations of highly specialized vocabulary. The paragraphs devoted to medical settings mostly focused on one-to-one communication examples, mainly corroborating the findings of Dew et al [18].

In addition to these systematic reviews, other studies have assessed the use of MT in different health settings. Panayiotou et al [22] provided a methodical evaluation of 15 Apple iPad-compatible language translation apps to facilitate conversations between health care providers and patients in Australia; aside from its geographically bounded context, the study centers on native mobile apps for one-to-one communication. Nurminen and Koponen [23] outlined several applications of MT for increasing information accessibility in humanitarian settings (eg, an armed conflict, a natural disaster, or an epidemic), including a paragraph devoted to discussing community-based health, as well as safety and security information. Although relevant to PH, the overview neither specifically reviews other contexts nor identifies patterns in the literature regarding the state of readiness of MT for PH settings.

These earlier publications are mostly confined to reporting literature on the use of MT for real-time bilingual person-to-person communication. The technology is mainly studied as an on-premise solution to support medical service provision in spoken interactions between specific groups of patients (eg, tourists, refugees, or expatriates) and health care staff (eg, general practitioners, caregivers, or paramedics) [24-28]. Only a few of the articles explore the use of multilingual translation tools for disseminating PH information to specific target audiences [29] or for population-wide health initiatives [30].

The Goal of This Study

The objective of this scoping review was to systematically map the use of MT for conducting PH outreach, with a particular focus on population-based recruitment methods. As a first step, we identify the information exchange scenarios in which MT technology is used to facilitate essential PH operations in different health and care settings. Second, we provide an overview of the types of study designs and research instruments for monitoring and evaluating the use of MT in these cases. Third and last, we synthesize the reported findings, benefits, and risks in relation to technical, socioeconomic, and ethicolegal technology readiness levels.

Methods

Search Strategy and Selection Criteria

This scoping review was preregistered on the Open Science Framework on February 11, 2022 [31], and conducted in accordance with the updated guidance on scoping reviews of the JBI Manual for Evidence Synthesis [32] as well as the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Multimedia Appendix 1 [33]) [34,35].

This scoping review exclusively includes peer-reviewed original research describing and assessing the use and suitability of MT for written texts for the purpose of improving collective outreach as well as the response and involvement of participants in the fields of epidemiology and PH, regardless of the specific target interventions or health areas involved. Given the technical nature of the research topic, peer-reviewed conference papers were also included. In addition, articles reporting guidelines or consensus statements concerning the use of MT in PH settings were included. The scoping review considers only studies written in English and published from 2007 onward, a year after the launch of the first fully web-based MT system and the publication of the first reference framework for MT quality assurance (the EN15038 standard) [36]. Studies of individual care or counseling settings (eg, practitioner and patient) were excluded because, in these settings, MT is used for spoken two-way communication. Textbox 1 presents the eligibility criteria.

Textbox 1. Eligibility criteria for the scoping review.

Inclusion criteria

- Article type
  - Peer-reviewed original research
  - Peer-reviewed conference papers
  - Consensus statements concerning the use of machine translation in public health settings
- Language
  - Studies published in English
- Time span
  - Studies published after January 2007
- Study design
  - Empirical studies

Exclusion criteria

- Article type
  - Non–peer-reviewed research or gray literature
- Language
  - Studies not published in English
- Time span
  - Studies published before 2007
- Study design
  - Nonempirical studies
  - Studies of individual care or counseling settings

Search Strategy

Searches were conducted in PubMed (MEDLINE), PubMed Central, Scopus, ACM Digital Library, and IEEE Xplore.

As recommended by JBI, the search string was constructed according to the population or participants, concept, and context (PPC) framework [32]. No specific restrictions were used to define the study populations. The concept was defined by terms related to automatic translation technology and the context by defining settings for population-based communication in PH, epidemiology, and community-based health care (Multimedia Appendix 2).

Initially, the search string was created to query the PubMed search engine and thereafter adapted to PubMed Central, Scopus, IEEE Xplore, and ACM Digital Library (the search terms are listed in Multimedia Appendix 2). Where available, database-specific index terms were added (eg, Medical Subject Headings [MeSH] terms for PubMed). The search was restricted to abstracts and titles. The search strategy was refined with the assistance of a professional librarian. All searches were executed on January 31, 2022, and updated on March 3, 2023.

After deduplication and the application of the exclusion criteria (ie, language not English, publication before 2007, and non–peer-reviewed articles), both authors (PSH-E and SR) independently screened the titles and abstracts of all remaining records using the R packages *revtools* [37] and *metagear* [38], which provide tools for semiautomatic deduplication and title or abstract screening. Disagreements were discussed and resolved by reaching a consensus. If necessary, full texts were consulted.

Data Extraction, Synthesis, and Analysis

Data extraction was conducted using a standardized data extraction template to extract bibliographic characteristics, health information exchange scenarios, research objectives and corresponding study designs, and technical characteristics of the MT tools used, as well as to identify the principal findings in the selected articles.

Health information exchange scenarios were assessed using a settings approach to health promotion [39]. We extracted and classified data regarding the (1) transmitters and recipients of translated materials, (2) types of translated materials, (3) types of MT systems and the source and target languages studied, and (4) nature of the use of MT in PH procedures as unsupervised (ie, without editing efforts) or supervised (ie, combined with editing efforts).

Research objectives were assessed according to the World Health Organization (WHO) monitoring and evaluation (M&E) framework [40], which is useful to map the research and development of digital health technologies according to their stage in the innovation maturity life cycle. We then classified the articles as either monitoring studies or evaluation studies. We considered monitoring studies to be those involving research on the technical quality and stability of MT (eg, technology assessments and comparative experiments) and evaluation studies to be those reporting on the appraisals of the technology-based interventions over time (eg, usability, affordability, and economic cost-effectiveness studies), as well as implementation research for integrating developed systems within broader PH workflows.

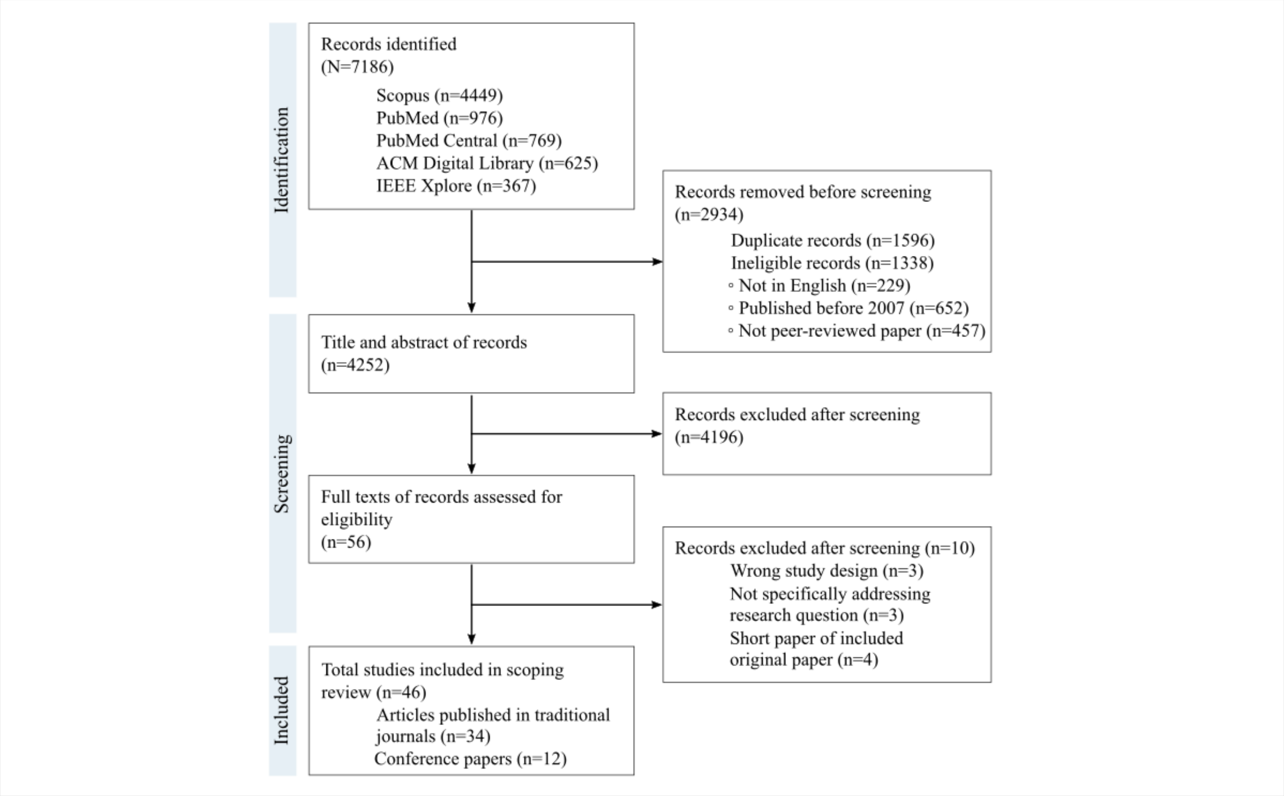
To assess the principal findings, we extracted sentences reporting quantitative and qualitative outcomes from the results sections. Following the *service readiness level* framework of evidence proposed by Hughes et al [41], we organized the statements as concerning technical, socioeconomic, or ethicolegal readiness levels of MT technology. On the basis of a manual sentiment analysis, we then detected the tonality of each text and classified them as positive, negative, or neutral.

Results

Search Outcomes

Conducted on January 31, 2022, and updated on March 3, 2023, the search yielded a total of 7186 records, of which 2934 (40.83%) were removed (1596/2934, 54.4% duplicates and 1338/2934, 45.6% not meeting the eligibility criteria). A review of the titles and abstracts of the remaining 4252 records resulted in 56 (1.32%) being selected for a full-text screening. From these 56 articles, 10 (18%) were removed for not meeting the study design criteria, not specifically addressing the research question, or for providing duplicate information from another included paper (Multimedia Appendix 3), and 46 (82%) were included in the systematic scoping review (Figure 1; Multimedia Appendix 4 [29,42-86]).

Figure 1. Flow diagram of the search and study selection process following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines.



The records were published between 2009 and 2023 as either conference papers (12/46, 26%) [48,53,55,60,61,63,65,71,78,80,82,84] or articles in traditional journals (34/46, 74%) [29,42-47,49-52,54,56-59,62,64,66-70,72-77,79,81,83,85,86] (Multimedia Appendix 4).

PH Information Exchange Scenarios

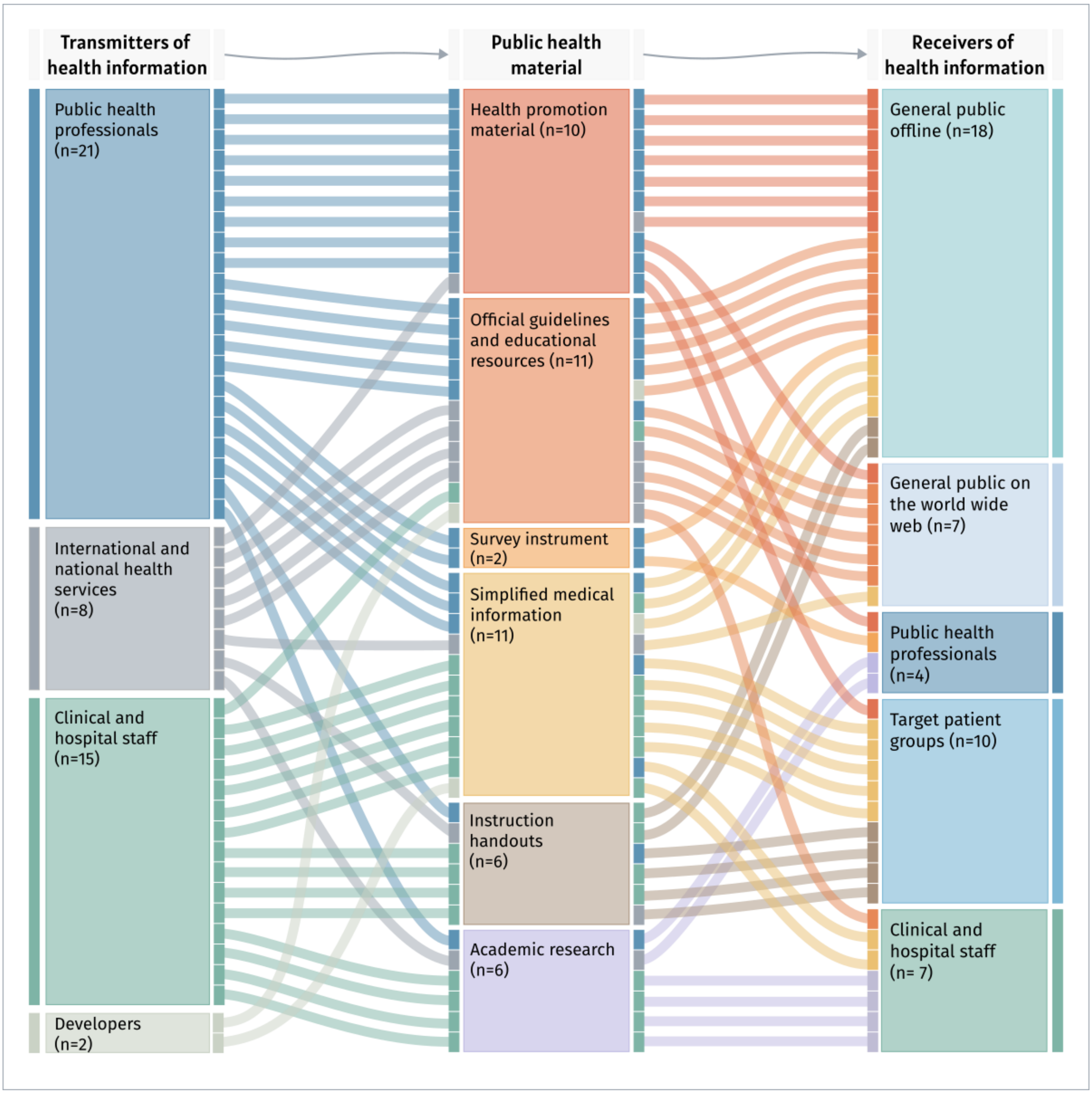
Four types of information transmitters (ie, the end users of MT) could be identified: PH departments and research institutions (21/46, 46%) [29,49,50,52-54,56,59,60,62,63,65,70,73-78,80,86]; clinical and hospital staff (15/46, 33%) [44-47,51,58,61,66,68,69,71,72,79,81,85]; international and national health organizations, such as the WHO, the US Centers for Disease Control and Prevention, and the UK National Health Service (8/46, 17%) [42,43,55,57,64,82-84]; and developers of web-based health information platforms (eg, Cochrane) or social media outlets (eg, Facebook; 2/46, 4%) [48,67].

The types of PH materials translated with MT fell into 6 broad categories: official guidelines and educational resources (11/46, 24%) [42,43,53,62,67,70,75,79,80,82,83], simplified medical information and lexica (11/46, 24%) [48,54,55,61,63,68,69,71,81,85,86], PH promotional material (10/46, 22%) [29,49,50,59,60,64,65,76-78], instruction handouts (6/46, 13%) [51,52,57,58,66,72], academic research (6/46, 13%) [44-47,73,84], and survey instruments (2/46, 4%) [56,74].

The information receivers (ie, the end users of translated material) could be categorized into 5 types: the wider population as targets of PH material offline (18/46, 39%) [29,48,51,53,55,58,59,62,64,65,67,68,74,76-78,80,86] or on the world wide web (7/46, 15%) [42,43,49,70,75,79,82], patient groups or communities (10/46, 22%) [50,52,57,61,63,66,69,72,81,85], clinical and hospital staff (7/46, 15%) [44-47,54,71,83], and PH professionals (4/46, 9%) [56,60,73,84].

Figure 2 provides a Sankey diagram visualizing PH information exchanges supported with the use of MT technology between groups of transmitters and receivers across the selected articles (Multimedia Appendix 5 [29,42-86]).

Figure 2. Public health information exchange scenarios: transmitters and receivers of public health information and the types of public health materials.



Overall, the most frequent use case of MT for collective communication in a PH setting is between PH staff and the public receiving paper-based health information (ie, offline; 14/46, 30%) [29,53,56,59,60,62,65,73,74,76-78,80,86] such as health promotion material (6/14, 43%) [29,59,65,76-78] or educational resources (3/14, 21%) [53,62,80]. The second most frequent use cases were exchanges involving clinical and hospital staff (10/46, 22%) [51,58,61,66,68,69,72,79,81,85] such as nurses or emergency wards disseminating simplified medical information (5/10, 50%) [61,68,69,81,85] or preparing instruction handouts for targeted audiences (4/10, 40%) [51,58,66,72].

Most of the articles (40/46, 87%) [29,42-48,50-63,66-69,71-74,77,79-86] specified the type of procedure for the use of MT as either unsupervised (30/40, 75%) [42-48,50-52,54-58,63,67-69,71-73,79-85] or supervised (10/39, 26%) [29,53,59-62,66,74,77,86] (Multimedia Appendix 5). Unsupervised MT was used for enabling translation plug-ins on PH websites (9/30, 30%) [42,43,48,67,69,79,81-83], translating PH material via web-based MT services (8/30, 27%) [50-52,56-58,68,72], and translating English content into the researchers' language (7/30, 23%) [44-47,73,80,84], as well as to investigate the risks of mistranslation and translation quality (5/30, 17%) [54,55,63,71,85]. Supervised MT procedures included postediting (6/10, 60%) [29,53,59,60,77,86] or pre-editing of source language (2/10, 20%) [61,62], as well as back translations (2/10, 20%) [66,74] of sample texts on the web and paper-based material.

The tested MT software tools were either freely available on the web from commercial technology vendors or were in-house built systems created by the research teams themselves (Multimedia Appendix 5). Regarding commercial vendors, Google Translate was the most used translation engine (28/46, 61%) [29,43-45,48,50-52,54-59,62,63,66,69,72,74,76-83,86], followed by Microsoft Bing (5/46, 11%) [44,47,61,69,79] and DeepL Translator (2/46, 4%) [73,86], among others. All these systems were used as domain-agnostic systems and not pretrained on specific language corpora. All articles regarding in-house built systems (9/46, 20%) [61,63,69,71,76,79,80,82,84] presented a prototype demonstration of domain-specific MT systems specifically trained on PH-related and medical vocabulary. The studies comparing these systems against each other (4/9, 44%) [71,79,80,84] advocate for using in-house built systems for shorter text with medical terminologies in long-term projects, whereas off-the-shelf systems may be used for more general information. In relation to each other, the evidence does not clearly favor 1 translation engine over another. Instead, it suggests that the choice among systems depends on the language pairs and the vocabulary domain used in the material. Provided that the texts are not exclusively reliant on specific terminologies, domain-agnostic solutions are equally suited for handling short-text translations.

Of the 46 articles, 40 (87%) [29,42-48,50-63,66-69,71-74,77,79-86] studied the use of MT to translate from 1 source language into 1 or several target languages (Multimedia Appendix 5). Of these 40 articles, 37 (93%) [29,42-45,48,50-63,66-69,71-74,77,79-85] specified the source language, whereas 35 (88%) [29,42-45,48,50-52,54-63,66,67,69,71-74,77,79-86] specified the target language. Of the 12 source languages, English was the most commonly evaluated (32/37, 86%) [29,42-45,48,50-60,62,66-68,71,72,74,77,79-85], followed by French (4/37, 11%) [48,69,79,80,84], German (3/37, 8%) [69,79,80], and Spanish (3/37, 8%) [79,80,84]. MT was tested in translating texts from English into at least 44 other languages, with Spanish (17/35, 49%) [29,48,50-52,57-60,66,72,74,79,80,83-85], Chinese (13/35, 37%) [50-52,56,58,62,72,74,77,81-83,85], and French (7/35, 20%) [52,71,74,79,80,83,84] being the most frequent target languages. Within this subset of 35 articles, 25 (71%) [29,42-45,50-52,54,56,57,59-63,66,69,72-74,77,79,80,85] conducted studies in specific geographies targeting populations and communities with limited language proficiency. Of these 25 articles, 19 (76%) [29,50-52,56,57,59-63,66,69,72,74,77,79,80,85] targeted individuals with limited English proficiency, mainly residing in the United States (17/19, 89%) [29,50-52,56,57,59-61,63,66,69,72,77,79,80,85].

Study Designs According to the WHO M&E Framework

In accordance with the WHO M&E framework [40], we identified 6 types of research designs across the selected articles (Table 1; Multimedia Appendix 6 [29,42-86]): MT technology assessments (study type 1; 23/46, 50%) [29,43-46,48,50-52,54-59,62,68,72-74,77,85,86], technology stability standards (study type 2; 3/46, 7%) [66,71,81], prototype demonstrations (study type 3; 8/46, 17%) [61,63,69,79,80,82-84], usability studies (study type 4; 4/46, 9%) [42,47,64,65], economic evaluations (study type 5; 4/46, 9%) [49,70,75,78], and implementation research (study type 6; 4/46, 9%) [53,60,67,76]. Nearly three-quarters (34/46, 74%) of the articles [43-46,48,50-52,54-59,61-63,66,68,69,71-74,76,79-86] conducted monitoring studies (ie, study types 1, 2, and 3), whereas more than a quarter (12/46, 26%) [42,47,49,53,60,64,65,67,70,75,76,78] conducted evaluation studies (ie, study types 4, 5, and 6).

Table 1. Categorization of studies according to the World Health Organization monitoring and evaluation framework (n=46).

Study type and research design	Studies, n (%)
Monitoring studies: functionality and stability of MT[a] at predefined levels of quality	
1. MT technology assessments: studies assessing MT quality, functionality, and performance	23 (50)
2. Technology stability standards: studies proposing standards or criteria for MT quality assurance	3 (7)
3. Prototype demonstrations: studies reporting on the development and design of an in-house built MT-based system	8 (17)
Evaluation studies: MT technology in health-related settings	
4. Usability studies: studies addressing end-user attitudes, perceptions, and responses when using the prototype system and assessing how easily end users can interact with the system	4 (9)
5. Economic evaluations: studies addressing accessibility, availability, or affordability of the system	4 (9)
6. Implementation research: studies around the implementation of MT technology within a broader (public) health system architecture	4 (9)

[a] MT: machine translation.

The monitoring studies adopted standard MT evaluation methods to measure the quality of MT output across various samples of health information material. Most of these studies focused on studying MT quality in terms of structural accuracy (28/34, 82%) [29,43,45,46,48,50-52,54-58,61,63,68,69,72-74,77,79-84,86] and fluency in unsupervised MT procedures (17/34, 50%) [44-46,50,51,54,56,57,59,62,66,71,72,77,80,81,85]. A quarter (8/28, 29%) of the articles [61,62,69,71,79,80,82,84] assessing structural accuracy supplemented their findings with standard automatic evaluation methods to verify the quality of MT output in comparison with the output of professional human translators. Flesch-Kincaid grade level scores and content analysis techniques were used to measure the readability levels and meaning preservation of the translated sentences. In a few of the articles (6/34, 18%) [50,51,57,58,66,72], MT was also evaluated in terms of the risk severity of mistranslation (ie, the degree of negative impact on the patient's health outcome because of a wrong translation). Studies investigating postediting (4/34, 12%) [29,59,77,86] or back translation (2/34, 6%) [66,74] focused on identifying error patterns or measuring the amount of time saved, whereas pre-editing (2/34, 6%) [61,62] was investigated to understand the ability of MT to handle PH jargon and medical terminologies.

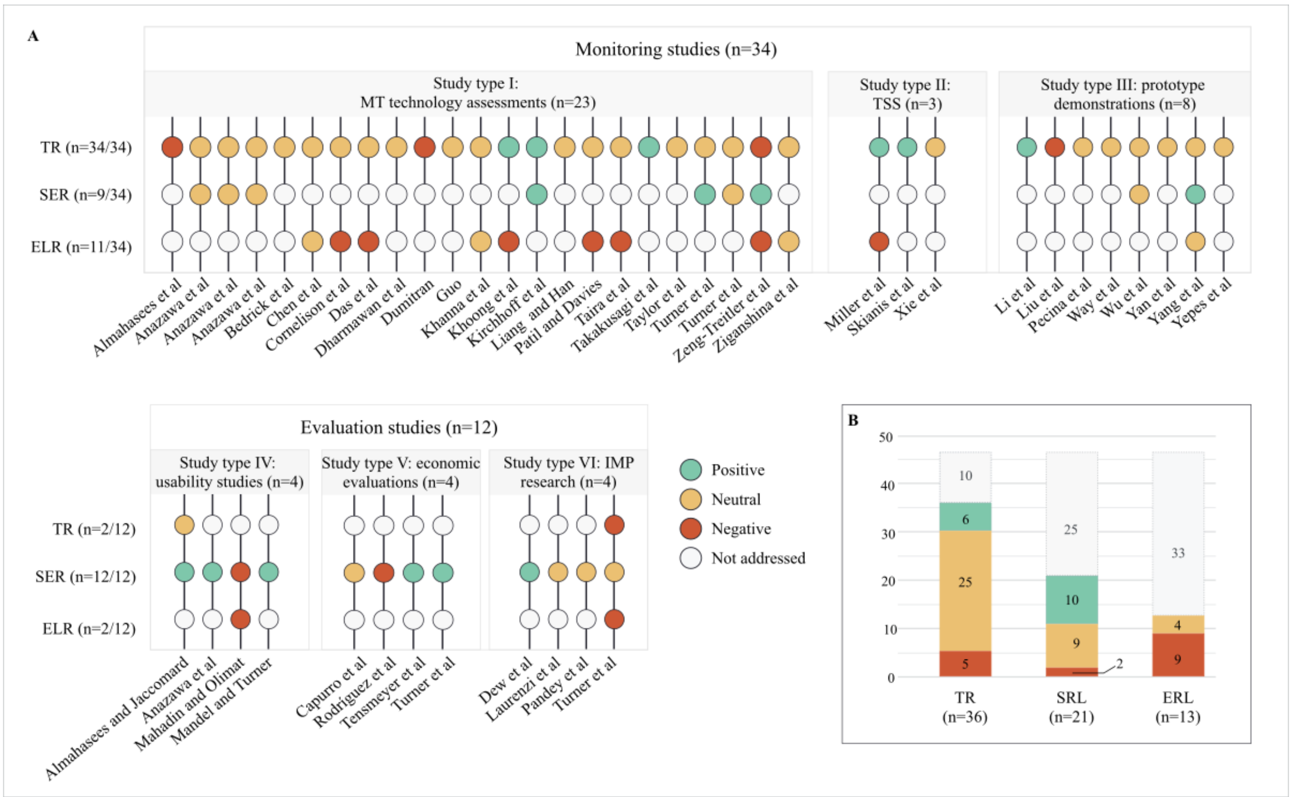
In their experiments, some of the studies (4/34, 12%) [29,43,59,81] extracted sentences from global PH (2/4, 50%) [43,81] and local PH promotion documents (2/4, 50%) [29,59]. Others (7/34, 21%) [48,51,58,61,66,68,72] carried out their experiments with general patient care instructions (2/7, 29%) [66,68], with side effects lists and directions for the use of prescribed drugs (3/7, 43%) [48,51,61], and from free-text or commonly used sentences in discharge instructions (2/7, 29%) [58,72]. A few of the articles (5/34, 15%) [44-47,54] used sentences from nursing abstracts (4/5, 80%) [44-47] and technical glossaries and dictionaries (1/5, 20%) [54]. Most of these experiments (27/34, 79%) [29,43-46,48,50-52, 54-59,61,63,66,68,74,77,80-82,85,86] recruited participants among PH professionals and certified translators with high proficiency in target and source languages or some experience with PH vocabulary. In general, discussions on ethical issues and quality inefficiencies across different languages did not address the impact of MT on possible information divides.

The evaluation studies deployed qualitative research instruments to understand how different types of end users view the adoption of MT technology in their information communication processes. Along with semistructured interviews, cognitive workflow analyses were used to understand current practices and the actual use of MT for multilingual document production workflows by PH departments. In addition, many of the studies (8/12, 67%) [42,47,49,53,64,65,67,78] used structured questionnaires and semistructured interviews to assess the perceived usefulness of MT (4/8, 50%) [42,47,64,67] and attitudes toward its adoption (4/8, 50%) [49,53,65,78] in such cases. Some of the articles (3/12, 25%) [53,60,76] sought to determine the practicality of implementing MT in combination with postediting efforts into local PH department workflows by timing and measuring the translation error rate of different translation procedures. Of the 12 studies, 4 (33%) [47,60,65,78] interviewed PH personnel and experts, and 2 (17%) [49,70] analyzed the availability of languages in PH web pages. Only 1 (8%) [42] of the 12 studies surveyed individuals in a real-world setting (ie, Facebook posts) to understand intelligibility or comprehension problems produced by MT in daily life situations on the web.

Technical, Socioeconomic, and Ethicolegal Readiness Levels

A total of 70 statements were identified as principal findings in the discussion and conclusions sections within the 46 articles (Multimedia Appendix 7 [29,42-86]). The majority (36/46, 78%) of the articles under review [29,42-46,48, 50-52,54-59,61-63,66,68,69,71-74,76,77,79-84,86] drew conclusions regarding the level of technical readiness of MT, nearly half (21/46, 46%) [29,42,44-47,49,53,59,60, 64,65,67,70,75-78,80,83,85] considered MT's socioeconomic readiness, and more than a quarter (13/46, 28%) [50-52,57,58,64,66,68,72,76,83,85,86] discussed the ethicolegal readiness of the translation outcomes for PH operations. Overall, one-fifth (16/70, 23%) of the statements within the articles expressed optimism about the use of MT for PH purposes [29,42,47,53,58,59,61,65,66,71,73,75,78,83,85], whereas another one-fifth (16/70, 23%) was pessimistic [43,51,52,55,58,63,64,66,68,70,72,76,85], and the remainder (38/70, 54%) presented mixed or inconclusive results [29,42,44-52,54,56,57,60,62,67-69,72,74,76,77,79-84,86]. Optimistic, pessimistic, and neutral statements are accounted for by green, red, and yellow circles, respectively, in Figure 3 [29,42-86].

Figure 3. Appraisal of study results. (A) Positive, negative, and mixed findings on the use of machine translation (MT) in public health settings by type of study and technology readiness dimensions. (B) Aggregate of final statements (N=70) by technological readiness levels across the 46 selected articles. IMP: implementation research; ELR: ethicolegal readiness; SER: socioeconomic readiness; TR: technical readiness; TSS: technology stability standards.



Concerning final statements on the technical readiness of MT, three-quarters (25/36, 69%) [29,42,44-46,48,50-52,54,56,57,62,68,69,72,74,77,79-84,86] of the articles were inconclusive and expressed cautionary optimism for translating PH material. Some of these articles (5/25, 20%) [29,62,74,77,86] recommended using a combination of MT and postediting efforts to improve translation quality and productivity. Of the 25 articles, 2 (8%) [61,62] also recommended using pre-editing efforts, such as controlled language and vocabularies, to reduce the need for postediting efforts. Articles addressing the socioeconomic readiness of automatic translations (21/46, 46%) [29,42,44-47,49,53,59,60,64,65,67,70,75-78,80,83,85] concluded either with optimistic (10/21, 48%) [29,42,47,53,59,65,75,78,83,85] or mixed results (9/21, 43%) [44-46,49,60,67,76,77,80], whereas a couple presented pessimistic results (2/21, 10%) [64,70]. On the one hand, these articles confirmed the enthusiasm of PH workers to adopt MT to increase cost-effectiveness as well as provide diverse material to wider audiences. On the other hand, they also stressed the importance of preparing the workforce to use the technology and ensuring that standard processes are created in light of PH equity goals. Regarding ethical and legal readiness, none of the articles mentioning the topics (n=13) concluded with optimistic statements; the findings were mostly negative (9/13, 69%) [51,52,58,64,66,68,72,76,85] and a few were mixed (4/13, 31%) [50,57,83,86]. In general, the articles stressed that the technology represents noteworthy communication risks, namely owing to a varying translation accuracy across languages. A few studies (4/13, 31%) [52,66,72,86] also pointed out that the commercial vendors' algorithms are not verifiable by the researchers or staff, resulting in a loss of control when not combined with editing efforts.

Discussion

Principal Findings

In our scoping review, we sought to systematically identify and map existing peer-reviewed literature on the use of MT for population-based outreach, with a particular interest in its use for recruiting participants for PH and epidemiological research. None of the included articles (n=46), published between 2009 and 2023, tested MT for recruiting participants to population-based studies or in scenarios where a response from addressees is expected. Research on the use of MT for PH activities is still in its early stages, primarily concentrating on assessing the technical readiness for one-way written communication between PH officials and addressed audiences. The majority of information transmitters (ie, the end users of MT) were PH professionals in PH departments and research, clinical and hospital staff, or staff at international and national health organizations. PH materials translated with MT were predominantly official guidelines and educational resources, simplified medical information, or PH promotional material. The intended target audiences (ie, the receivers of translated material) were the wider population (both offline and seeking information on the world wide web), patient groups, or professionals in PH and clinical settings. Nearly three-quarters (34/46, 74%) of the articles reported monitoring studies, with the remaining quarter (12/46, 26%) reporting evaluation studies.

Research on the Use of MT for PH Activities Is Still Nascent

The current focus of research is mostly concentrated on understanding the extent to which machine-translated output is reliable and stable enough for translating specific sample texts, while placing less emphasis on the feasibility of its use in real-world settings. Published study types mostly provided technical maturity assessments of MT (eg, in exploratory research, experimental proofs of concept, and implementation research studies).

The majority of the studies (28/46, 61%) [29,43,45,46,48,50-52,54-58,61,63,68,69,72-74,77,79-84,86] solely focused on MT accuracy errors and how to drive error rates down. Most of the articles (23/46, 50%) [29,43-46,48,50-52,54-59,62,68,72-74,77,85,86] provided technical assessments, and in most of the cases (14/23, 61%) [29,43-46,48,54-56,59,62,73,74,77], they studied neither the reliability of the technology for specific target audiences nor the potential risks of mistranslation. Although articles often specified the type of MT algorithm as either statistical MT or neural MT (Multimedia Appendix 5), none systematically compared the algorithms or reported on specific advantages or disadvantages. Therefore, it is not clear whether the type of algorithm has any relevance for using MT in PH scenarios.

A handful of studies (9/46, 20%) [61,63,69,71,76,79,80,82,84] reported ongoing research in the development of in-house software, pretrained on specific vocabulary. These systems were reported to outperform off-the-shelf models (eg, Google Translate and DeepL Translator), namely when translating shorter text with specialized terminologies, such as those used in medical guidelines or prescriptions. The fact that the technology is evolving and can now be trained in PH and biomedical vocabulary sheds light on future possibilities to meet the needs of staff working with more complex PH material. However, the current state of evaluations on the advantages and disadvantages of the off-the-shelf systems over internally developed models does not yet allow PH researchers to model the best use of both systems during specific stages of material production. Provided that PH material does not heavily rely on domain-specific vocabulary, off-the-shelf MT solutions are sufficiently reliable in terms of translating shorter text. Given that these systems are predominantly free to use and easily adaptable to a translation workflow, proprietary models are relatively costly to develop and maintain, as well as scale to new vocabularies.

The literature tends to focus on evaluating the accuracy of supervised translations from the language of the working staff or researchers (typically English) to 1 or a few languages (in most cases, Spanish, Chinese, or French). The observed inclination to study English as a source can be attributed to the origin of the selected articles in this review. For most of the studies (19/46, 41%) [29,50-52,56,57,59-63,66,69,72,74,77,79,80,85], the target audiences of interest were large linguistically diverse communities residing in predominantly English-speaking countries (eg, the United States, the United Kingdom, and Australia). Future studies could also aim to cover underrepresented languages beyond that of the largest linguistically diverse groups and continue exploring cases to support linguistically diverse PH staff. For now, a few of these studies (6/46, 13%) [29,59,66,74,77,86] tested MT in light of postediting efforts. As user-friendly MT applications become more accessible to the public and professionals, we can reasonably assume that the focus of MT research in PH might shift from generating texts with MT to generating texts that are optimized for MT, that is, the emphasis might shift from technical accuracy and postediting efforts to pre-editing of texts.

A limited number of articles (21/46, 46%) [29,42,44-47,49,53,59,60,64,65,67,70,75-78,80,83,85] investigated the societal acceptance of MT, mainly by surveying the attitudes of PH staff toward its adoption, formulating new concepts, and studying current practices and standards. The selected studies point to the conclusion that PH staff are enthusiastic and open to adopting MT in their workflows. Almost half (10/21, 48%) of the studies held positive attitudes toward the potential cost-effectiveness of using MT to increase public access to PH information. However, the technology has not been routinely adopted by PH departments owing to safety concerns, the loss of control over content, and the unquantified variability of the quality of translation between languages. There is a need to further identify relevant stakeholders for implementing and deploying MT, as well as to test proposed solutions in controlled environments with the end users of translated material.

Most of the experiments (31/46, 67%) were based on expert focus groups and surveying PH professionals, whereas only a few (3/46, 7%) explored end-user interactions, preferences, and perspectives in real-world settings. However, without real-world studies conducted outside laboratory settings and in field experiments, the user experience of the technology remains largely unknown. Only a few studies (8/46, 17%) [42,47,49,53,64,65,67,78] tested the usability and acceptability of MT in community settings. Future studies could explore, for example, end-user interactions with machine-translated text in daily life settings, while also continuing to survey PH professionals in digital environments and capturing their attitudes toward use and adoption, as well as measuring the actual information uptake by groups targeted with machine-translated materials compared with nontranslated materials alone.

Moreover, no article focused solely on the legal or ethical aspects of the use of MT for PH purposes. However, some of the studies (13/46, 28%) [50-52,57,58,64,66,68,72,76,83,85,86] did provide a generic consideration of ethical compliance aspects as part of their discussions. To the extent that these concerns were addressed, 2 (4%) of the 46 studies called attention to the fact that the commercial vendors' algorithms are not transparent to researchers and staff. Investigating MT from an ethical perspective, such as its impact on the digital divide, and establishing standards for its adoption also remain pending in light of PH equity goals and the risk of harmful errors.

No Current Research on MT in Two-Way Communication Scenarios

None of the reviewed studies specifically tested MT for the recruitment of participants in population-based research. The literature only covers the use of MT for communicating in PH settings that do not warrant a response from addressees. Most of the studies (27/46, 59%) [29,42,43,49-53,57-60,62,64-67,70,72,75-80,82,83] focused on the use of MT for translating simple text in flyers, instructions, and general information sheets from 1 language into a selected few. Hardly any of the articles (44/46, 96%) [29,42-74,80-86] discussed cases where the technology was used to communicate with several linguistically diverse populations at once. Only 2 (4%) [75,79] of the 46 studies introduced the use of MT for emergency preparedness and outreach prompted by the COVID-19 pandemic. These cases remain examples of unidirectional communication between PH staff and addressed audiences who are not expected to provide a response in return.

One possible reason why MT has not been used for recruitment in population-based research may be that there is limited utility in providing translations of PH material into languages that are not spoken or read by researchers or field staff or in recruiting participants who cannot interact with the languages in which the study is offered. On the contrary, if studies are offered in multiple languages, they are usually prepared with research instruments and personnel pre-equipped with the skills to meet the language diversity of the study population. It is therefore rather unlikely that MT would be necessary for translating recruitment materials in the first place.

However, there are scenarios in which MT may prove beneficial in population-based recruitment; for example, in studies on children and adolescents, the actual study participants often speak the language of the country fluently, but their legal guardians, who have to consent to their children's participation, might not be proficient in the language. Providing them with study information and consent forms in their preferred language might help them to understand what is asked from them and their children and, therefore, increase the probability that they will provide consent. However, for such purposes, ensuring a certain translation quality is crucial to meet ethical and legal requirements, but, as mentioned before, this review did not find much evidence of research regarding this problem.

Furthermore, providing multilingual invitations could also help PH employees to understand the demand for different languages at the population level. If addressees could be enabled to report their preferred languages back to PH staff, the collected data might be used to adapt ongoing or future studies to provide additional language support. Alternatively, addressees could be informed that participation is possible, contingent on being accompanied by a translator.

Finally, even if it is not possible to add each language preferred by potential study participants, using MT tools for PH study invitations would ensure that more addressees understand the content of the invitation letters, which, given their official appearance, might otherwise leave them uncertain regarding missing out on something important or even undermine trust in PH departments and reduce participation in future studies or initiatives.

Limitations

Our findings should be considered with limitations. First, this review is limited to publications addressing the use of MT either as part of the research question or as a key point of discussion in the publications. It cannot be ruled out that MT might already be used as a routine tool, and therefore, its use is not reported in peer-reviewed papers. Second, we used an interpretative sentiment analysis to classify the principal findings for each article based on the extraction of selected statements. This exercise, although systematic and with the intention of objectivity, is prone to the authors' interpretation of enthusiasm regarding the specific dimensions of digital technology maturity. Finally, the search was limited to articles published only in English, which might bias the results toward studies examining MT from or into English. There is also a possibility that articles published before 2007 could contain information relevant to the research question. However, because the technology has evolved exponentially in the last 2 decades, prior information is likely to be outdated and no longer applicable to current standards.

Conclusions

Using MT in epidemiology and PH can enhance outreach to linguistically diverse populations. The translation quality of current off-the-shelf systems, such as Google Translate or DeepL Translator, is sufficient if postediting is a mandatory step in the translation workflow. Postediting of legally or ethically sensitive material requires staff with adequate content knowledge in addition to sufficient language skills. When preparing texts for translation, it is advisable to use shorter sentences and specifically mark domain-specific vocabulary for possible postediting. Unsupervised MT is generally not recommended. Research on whether machine-translated texts are received differently by addressees is lacking, as well as research on MT in communication scenarios that warrant a response from the addressees.

Acknowledgments

The authors gratefully acknowledge Lara Christianson's counsel and her contribution to refining the search strategy, as well as Sara E. Cooper's final proofreading and language editing. This study was partially supported by Leibniz ScienceCampus Bremen Digital Public Health [87], which is jointly funded by the Leibniz Association (W4/2018), the federal state of Bremen, and the Leibniz Institute for Prevention Research and Epidemiology–BIPS. The funding source had no role in the design, data collection, data analysis, or decision to submit the manuscript for publication.

Data Availability

The complete list of included studies and all data generated or analyzed during this study are included in this published article and the multimedia appendices.

Authors' Contributions

PSH-E and SR conceptualized and guided the research project, established methodology, conducted the literature search, and contributed to the data interpretation. PSH-E curated and analyzed the data and designed and visualized the figures. PSH-E wrote the original draft of the manuscript. PSH-E and SR revised and edited the manuscript. SR was in charge of supervision and the initial funding acquisition. Both authors reviewed and validated the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 138 KB - publichealth_v9i1e50814_app1.pdf]

Multimedia Appendix 2
Search terms for PubMed, PubMed Central, Scopus, IEEE Xplore, and ACM Digital Library.
[PDF File (Adobe PDF File), 198 KB - publichealth_v9i1e50814_app2.pdf]

Multimedia Appendix 3
List of studies excluded during full-text screening.
[PDF File (Adobe PDF File), 118 KB - publichealth_v9i1e50814_app3.pdf]

Multimedia Appendix 4
Study characteristics.
[PDF File (Adobe PDF File), 252 KB - publichealth_v9i1e50814_app4.pdf]

Multimedia Appendix 5
Transmitters, receivers, public health material, and machine translation engines.
[PDF File (Adobe PDF File), 186 KB - publichealth_v9i1e50814_app5.pdf]

Multimedia Appendix 6
Research study designs.
[PDF File (Adobe PDF File), 190 KB - publichealth_v9i1e50814_app6.pdf]

Multimedia Appendix 7
Authors’ sentiment on the principal findings in the discussion and conclusions sections.
[PDF File (Adobe PDF File), 137 KB - publichealth_v9i1e50814_app7.pdf]

References

1. Rindfuss RR, Choe MK, Tsuya NO, Bumpass LL, Tamaki E. Do low survey response rates bias results? evidence from Japan. Demographic Res 2015 Mar 25;32:797-828 [FREE Full text] [doi: 10.4054/demres.2015.32.26]
2. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
3. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q 2008 May 07;72(2):167-189 [FREE Full text] [doi: 10.1093/poq/nfn011]
4. Mindell JS, Giampaoli S, Goesswald A, Kamtsiuris P, Mann C, Männistö S, HES Response Rate Group. Sample selection, recruitment and participation rates in health examination surveys in Europe--experience from seven national surveys. BMC Med Res Methodol 2015 Oct 05;15:78 [FREE Full text] [doi: 10.1186/s12874-015-0072-4] [Medline: 26438235]
5. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006 Feb 01;163(3):197-203. [doi: 10.1093/aje/kwj036] [Medline: 16339049]
6. Stang A. Nonresponse research--an underdeveloped field in epidemiology. Eur J Epidemiol 2003;18(10):929-931. [doi: 10.1023/a:1025877501423] [Medline: 14598921]

7. Xu M, Richardson L, Campbell S, Pintos J, Siemiatycki J. Response rates in case-control studies of cancer by era of fieldwork and by characteristics of study design. Ann Epidemiol 2018 Jun;28(6):385-391. [doi: 10.1016/j.annepidem.2018.04.001] [Medline: 29703520]
8. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009 Jul 08;2009(3):MR000008 [FREE Full text] [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]
9. Brand T, Samkange-Zeeb F, Dragano N, Keil T, Krist L, Yesil-Jürgens R, et al. Participation of Turkish migrants in an epidemiological study: does the recruitment strategy affect the sample characteristics? J Immigr Minor Health 2019 Aug;21(4):811-819. [doi: 10.1007/s10903-018-0788-4] [Medline: 30027505]
10. Samkange-Zeeb F, Samerski S, Doos L, Humphris R, Padilla B, Bradby H. "It's the First Barrier" - lack of common language a major obstacle when accessing/providing healthcare services across Europe. Front Sociol 2020 Nov 5;5:557563 [FREE Full text] [doi: 10.3389/fsoc.2020.557563] [Medline: 33869495]
11. Feskens R. Difficult groups in survey research and the development of Tailor-Made Approach strategies. Centraal Bureau voor de Statistiek. 2009. URL: https://www.cbs.nl/-/media/imported/documents/2009/01/2009-x11-feskens-pub.pdf [accessed 2021-02-18]
12. Guo Y, Kopec JA, Cibere J, Li LC, Goldsmith CH. Population survey features and response rates: a randomized experiment. Am J Public Health 2016 Aug;106(8):1422-1426. [doi: 10.2105/AJPH.2016.303198] [Medline: 27196650]
13. Askin N. Providing access to multilingual patient education materials. J Consum Health Internet 2017;21(3):305-312 [FREE Full text] [doi: 10.1080/15398285.2017.1353289]
14. Morales M. Health literacy research methods: a pilot study investigation with adult beginning readers. Int J Qual Methods 2017 Jan;16(1):160940691770991. [doi: 10.1177/1609406917709913]
15. Siegel JS. Language as a factor in the quality of demographic data and translation issues in the management of surveys. In: Siegel JS, editor. Demographic and Socioeconomic Basis of Ethnolinguistics. Cham, Switzerland: Springer; 2017:153-183.
16. Liu Q, Zhang X. Machine translation: General. In: Sin-wai C, editor. Routledge Encyclopedia of Translation Technology. Oxfordshire, UK: Routledge; 2014.
17. Khoong EC, Rodriguez JA. A research agenda for using machine translation in clinical medicine. J Gen Intern Med 2022 Apr;37(5):1275-1277 [FREE Full text] [doi: 10.1007/s11606-021-07164-y] [Medline: 35132559]
18. Dew KN, Turner AM, Choi YK, Bosold A, Kirchhoff K. Development of machine translation technology for assisting health communication: a systematic review. J Biomed Inform 2018 Sep;85:56-67 [FREE Full text] [doi: 10.1016/j.jbi.2018.07.018] [Medline: 30031857]
19. Frampton GK, Shepherd J, Pickett K, Griffiths G, Wyatt JC. Digital tools for the recruitment and retention of participants in randomised controlled trials: a systematic map. Trials 2020 Jun 05;21(1):478 [FREE Full text] [doi: 10.1186/s13063-020-04358-3] [Medline: 32498690]
20. Thonon F, Perrot S, Yergolkar AV, Rousset-Torrente O, Griffith JW, Chassany O, et al. Electronic tools to bridge the language gap in health care for people who have migrated: systematic review. J Med Internet Res 2021 May 06;23(5):e25131 [FREE Full text] [doi: 10.2196/25131] [Medline: 33955837]
21. Vieira LN. Translation. In: Burns J, Duncan D, editors. Transnational Modern Languages: A Handbook. Liverpool, UK: Liverpool University Press; 2022:297-304.
22. Panayiotou A, Hwang K, Williams S, Chong TW, LoGiudice D, Haralambous B, et al. The perceptions of translation apps for everyday health care in healthcare workers and older people: a multi-method study. J Clin Nurs 2020 Sep;29(17-18):3516-3526. [doi: 10.1111/jocn.15390] [Medline: 32558965]
23. Nurminen M, Koponen M. Machine translation and fair access to information. Transl Spaces 2020 Aug;9(1):150-169 [FREE Full text] [doi: 10.1075/ts.00025.nur]
24. Berbyuk Lindström N, Rodríguez Pozo R. Perspectives of nurses and doulas on the use of information and communication technology in intercultural pediatric care: qualitative pilot study. JMIR Pediatr Parent 2020 Mar 17;3(1):e16545 [FREE Full text] [doi: 10.2196/16545] [Medline: 32181748]
25. Furaijat G, Kleinert E, Simmenroth A, Müller F. Implementing a digital communication assistance tool to collect the medical history of refugee patients: DICTUM Friedland - an action-oriented mixed methods study protocol. BMC Health Serv Res 2019 Feb 06;19(1):103 [FREE Full text] [doi: 10.1186/s12913-019-3928-1] [Medline: 30728030]
26. Kaliyadan F, Gopinathan Pillai S. The use of Google language tools as an interpretation aid in cross-cultural doctor-patient interaction: a pilot study. Inform Prim Care 2010;18(2):141-143 [FREE Full text] [doi: 10.14236/jhi.v18i2.764] [Medline: 21078237]
27. Noack EM, Kleinert E, Müller F. Overcoming language barriers in paramedic care: a study protocol of the interventional trial 'DICTUM rescue' evaluating an app designed to improve communication between paramedics and foreign-language patients. BMC Health Serv Res 2020 Mar 18;20(1):223 [FREE Full text] [doi: 10.1186/s12913-020-05098-5] [Medline: 32183775]
28. Marais L, Louw JA, Badenhorst J, Calteaux K, Wilken I, van Niekerk N, et al. AwezaMed: a multilingual, multimodal speech-to-speech translation application for maternal health care. In: Proceedings of the 23rd International Conference on

Information Fusion. 2020 Presented at: FUSION '20; July 6-9, 2020; Rustenburg, South Africa p. 1-8 URL: https://ieeexplore.ieee.org/document/9190240 [doi: 10.23919/fusion45008.2020.9190240]

29. Turner AM, Bergman M, Brownstein M, Cole K, Kirchhoff K. A comparison of human and machine translation of health promotion materials for public health practice: time, costs, and quality. J Public Health Manag Pract 2014 Sep;20(5):523-529 [FREE Full text] [doi: 10.1097/PHH.0b013e3182a95c87] [Medline: 24084391]
30. Mawudeku A, Blench M. Global public health intelligence network (GPHIN). In: Proceedings of the 8th Conference of the Association for Machine Translation in the Americas. 2005 Presented at: AMTA '05; September 13-15, 2005; Phuket, Thailand p. 7-11 URL: https://aclanthology.org/2005.mtsummit-invited.4.pdf
31. Herrera-Espejel PS, Rach S. Machine translation in epidemiology and public health: scoping review. Open Science Framework. 2022 Feb 11. URL: https://osf.io/avt73 [accessed 2022-02-11]
32. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
33. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
34. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
35. Pollock D, Davies EL, Peters MD, Tricco AC, Alexander L, McInerney P, et al. Undertaking a scoping review: a practical guide for nursing and midwifery students, clinicians, researchers, and academics. J Adv Nurs 2021 Apr;77(4):2102-2113 [FREE Full text] [doi: 10.1111/jan.14743] [Medline: 33543511]
36. DeCamp J, Zetsche J. A history of translation technology in the United States. In: Sin-Wai C, editor. The Routledge Encyclopedia of Translation Technology. Oxfordshire, UK: Routledge; 2014.
37. Westgate MJ. revtools: an R package to support article screening for evidence synthesis. Res Synth Methods 2019 Dec;10(4):606-614. [doi: 10.1002/jrsm.1374] [Medline: 31355546]
38. Lajeunesse MJ. Facilitating systematic reviews, data extraction and meta-analysis with the metagear package for R. Methods Ecol Evol 2015 Oct 13;7(3):323-330 [FREE Full text] [doi: 10.1111/2041-210x.12472]
39. Poland B, Krupa G, McCall D. Settings for health promotion: an analytic framework to guide intervention design and implementation. Health Promot Pract 2009 Oct;10(4):505-516. [doi: 10.1177/1524839909341025] [Medline: 19809004]
40. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. World Health Organization. 2016. URL: https://www.who.int/publications/i/item/9789241511766 [accessed 2022-06-01]
41. Hughes J, Lennon M, Rogerson RJ, Crooks G. Scaling digital health innovation: developing a new 'service readiness level' framework of evidence. Int J Environ Res Public Health 2021 Nov 29;18(23):12575 [FREE Full text] [doi: 10.3390/ijerph182312575] [Medline: 34886302]
42. Almahasees Z, Jaccomard H. Facebook Translation Service (FTS) usage among Jordanians during COVID-19 lockdown. Adv Sci Technol Eng Syst J 2020 Nov;5(6):514-519 [FREE Full text] [doi: 10.25046/aj050661]
43. Almahasees Z, Meqdadi S, Albudairi Y. Evaluation of Google Translate in rendering English COVID-19 texts into Arabic. J Lang Linguist Stud 2021 Oct 28;17(4):2065-2080 [FREE Full text] [doi: 10.52462/jlls.149]
44. Anazawa R, Ishikawa H, Park MJ, Kiuchi T. Preliminary study of online machine translation use of nursing literature: quality evaluation and perceived usability. BMC Res Notes 2012 Nov 14;5(1):635 [FREE Full text] [doi: 10.1186/1756-0500-5-635] [Medline: 23151362]
45. Anazawa R, Ishikawa H, Park MJ, Kiuchi T. Online machine translation use with nursing literature: evaluation method and usability. Comput Inform Nurs 2013 Feb;31(2):59-65. [doi: 10.1097/NXN.0b013e3182701056] [Medline: 23095720]
46. Anazawa R, Ishikawa H, Takahiro K. Evaluation of online machine translation by nursing users. Comput Inform Nurs 2013 Aug;31(8):382-387. [doi: 10.1097/NXN.0b013e3182999dc2] [Medline: 23851708]
47. Anazawa R, Ishikawa H, Takahiro K. Use of online machine translation for nursing literature: a questionnaire-based survey. Open Nurs J 2013;7:22-28 [FREE Full text] [doi: 10.2174/1874434601307010022] [Medline: 23459140]
48. Bedrick SD, Mauro A. A multi-lingual web service for drug side-effect data. AMIA Annu Symp Proc 2009 Nov 14;2009:34-38 [FREE Full text] [Medline: 20351818]
49. Capurro D, Chaudhuri S, Turner AM. The online availability of multilingual health promotion materials produced by local health departments: an information assessment. Stud Health Technol Inform 2015;216:380-385. [Medline: 26262076]
50. Chen X, Acosta S, Barry AE. Evaluating the accuracy of Google Translate for diabetes education material. JMIR Diabetes 2016 Jun 28;1(1):e3 [FREE Full text] [doi: 10.2196/diabetes.5848] [Medline: 30291056]
51. Cornelison BR, Al-Mohaish S, Sun Y, Edwards CJ. Accuracy of Google Translate in translating the directions and counseling points for top-selling drugs from English to Arabic, Chinese, and Spanish. Am J Health Syst Pharm 2021 Nov 09;78(22):2053-2058. [doi: 10.1093/ajhp/zxab224] [Medline: 34048533]
52. Das P, Kuznetsova A, Zhu M, Milanaik R. Dangers of machine translation: the need for professionally translated anticipatory guidance resources for limited English proficiency caregivers. Clin Pediatr (Phila) 2019 Feb;58(2):247-249. [doi: 10.1177/0009922818809494] [Medline: 30392383]

53. Dew K, Turner AM, Desai L, Martin N, Laurenzi A, Kirchhoff K. PHAST: a collaborative machine translation and post-editing tool for public health. AMIA Annu Symp Proc 2015 Nov 5;2015:492-501 [FREE Full text] [Medline: 26958182]
54. Dharmawan R, Nababan MR, Tarjana MS, Djatmika D. Mistranslation and maltranslation in a medical website: evidences from Dorland’s medical dictionary. J Soc Stud Educ Res 2019;10(1):210-240 [FREE Full text]
55. Dumitran A. Translation error in Google Translate from English into Romanian in texts related to coronavirus. eLearn Softw Educ 2021;2:37-43 [FREE Full text] [doi: 10.12753/2066-026x-21-078]
56. Guo JW. Is Google Translate adequate for facilitating instrument translation from English to Mandarin? Comput Inform Nurs 2016 Sep;34(9):377-383. [doi: 10.1097/CIN.0000000000000287] [Medline: 27606576]
57. Khanna RR, Karliner LS, Eck M, Vittinghoff E, Koenig CJ, Fang MC. Performance of an online translation tool when applied to patient educational material. J Hosp Med 2011 Nov;6(9):519-525. [doi: 10.1002/jhm.898] [Medline: 22034196]
58. Khoong EC, Steinbrook E, Brown C, Fernandez A. Assessing the use of Google Translate for Spanish and Chinese translations of emergency department discharge instructions. JAMA Intern Med 2019 Apr 01;179(4):580-582 [FREE Full text] [doi: 10.1001/jamainternmed.2018.7653] [Medline: 30801626]
59. Kirchhoff K, Turner AM, Axelrod A, Saavedra F. Application of statistical machine translation to public health information: a feasibility study. J Am Med Inform Assoc 2011 Jul;18(4):473-478 [FREE Full text] [doi: 10.1136/amiajnl-2011-000176] [Medline: 21498805]
60. Laurenzi A, Brownstein MK, Turner AM, Kientz JA, Kirchhoff K. A web-based collaborative translation management system for public health workers. In: Proceedings of the 2013 CHI Extended Abstracts on Human Factors in Computing Systems. 2013 Presented at: CHI EA '13; April 27-May 2, 2013; Paris, France p. 511-516 URL: https://dl.acm.org/doi/10.1145/2468356.2468446 [doi: 10.1145/2468356.2468446]
61. Li J, Lester C, Zhao X, Ding Y, Jiang Y, Vydiswaran V. PharmMT: a neural machine translation approach to simplify prescription directions. In: Proceedings of the 2020 Findings of the Association for Computational Linguistics. 2020 Presented at: EMNLP '20; November 16-20, 2020; Virtual Event p. 2785-2796 URL: https://aclanthology.org/2020.findings-emnlp.251.pdf [doi: 10.18653/v1/2020.findings-emnlp.251]
62. Liang Y, Han W. Source text pre-editing versus target text post-editing in using Google Translate to provide health services to culturally and linguistically diverse clients. Sci Eng Health Stud 2022 Sep:1-5 [FREE Full text]
63. Liu W, Cai S. Translating electronic health record notes from English to Spanish: a preliminary study. In: Proceedings of the 2015 Workshop on Biomedical Natural Language Processing. 2015 Presented at: BioNLP '15; July 30, 2015; Beijing, China p. 134-140 URL: https://aclanthology.org/W15-3816/ [doi: 10.18653/v1/w15-3816]
64. Mahadin D, Olimat S. Jordanian translators' use of machine translation and glossary of COVID-19 terminology with reference to Arabic. New Voices Transl Stud 2022 Jun:26-54 [FREE Full text]
65. Mandel H, Turner AM. Exploring local public health work in the context of novel translation technologies. Stud Health Technol Inform 2013;192:1209. [Medline: 23920983]
66. Miller JM, Harvey EM, Bedrick S, Mohan P, Calhoun E. Simple patient care instructions translate best: safety guidelines for physician use of Google translate. J Clin Outcomes Manag 2018;1(1):25 [FREE Full text]
67. Pandey R, Gautam V, Pal R, Bandhey H, Dhingra LS, Misra V, Arushi, et al. A machine learning application for raising WASH awareness in the times of COVID-19 pandemic. Sci Rep 2022 Jan 17;12(1):810 [FREE Full text] [doi: 10.1038/s41598-021-03869-6] [Medline: 35039533]
68. Patil S, Davies P. Use of Google Translate in medical communication: evaluation of accuracy. BMJ 2014 Dec 15;349:g7392 [FREE Full text] [doi: 10.1136/bmj.g7392] [Medline: 25512386]
69. Pecina P, Dušek O, Goeuriot L, Hajič J, Hlaváčová J, Jones GJ, et al. Adaptation of machine translation for multilingual information retrieval in the medical domain. Artif Intell Med 2014 Jul;61(3):165-185. [doi: 10.1016/j.artmed.2014.01.004] [Medline: 24680188]
70. Rodriguez JA, Davis RB, Percac-Lima S. Non-English language availability of community health center websites. Med Care 2019 Jun;57 Suppl 6 Suppl 2:S121-S126. [doi: 10.1097/MLR.0000000000001027] [Medline: 31095050]
71. Skianis K, Briand Y, Desgrippes F. Evaluation of machine translation methods applied to medical terminologies. In: Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis. 2020 Presented at: HTMIA '20; November 20, 2020; Virtual Event p. 59-69 URL: https://aclanthology.org/2020.louhi-1.7.pdf [doi: 10.18653/v1/2020.louhi-1.7]
72. Taira BR, Kreger V, Orue A, Diamond LC. A pragmatic assessment of Google Translate for emergency department instructions. J Gen Intern Med 2021 Nov;36(11):3361-3365 [FREE Full text] [doi: 10.1007/s11606-021-06666-z] [Medline: 33674922]
73. Takakusagi Y, Oike T, Shirai K, Sato H, Kano K, Shima S, et al. Validation of the reliability of machine translation for a medical article from Japanese to English using DeepL translator. Cureus 2021 Sep;13(9):e17778 [FREE Full text] [doi: 10.7759/cureus.17778]
74. Taylor RM, Crichton N, Moult B, Gibson F. A prospective observational study of machine translation software to overcome the challenge of including ethnic diversity in healthcare research. Nurs Open 2015 Apr;2(1):14-23 [FREE Full text] [doi: 10.1002/nop2.13] [Medline: 27708797]

75. Tensmeyer NC, Dinh NN, Sun LT, Meyer CB. Analysis of language translations of state governments' coronavirus disease 2019 vaccine websites. Health Equity 2022 Sep 01;6(1):738-749 [FREE Full text] [doi: 10.1089/heq.2021.0189] [Medline: 36225668]
76. Turner AM, Brownstein MK, Cole K, Karasz H, Kirchhoff K. Modeling workflow to design machine translation applications for public health practice. J Biomed Inform 2015 Feb;53:136-146 [FREE Full text] [doi: 10.1016/j.jbi.2014.10.005] [Medline: 25445922]
77. Turner AM, Dew KN, Desai L, Martin N, Kirchhoff K. Machine translation of public health materials from English to Chinese: a feasibility study. JMIR Public Health Surveill 2015 Nov 17;1(2):e17 [FREE Full text] [doi: 10.2196/publichealth.4779] [Medline: 27227135]
78. Turner AM, Mandel H, Capurro D. Local health department translation processes: potential of machine translation technologies to help meet needs. AMIA Annu Symp Proc 2013 Nov 16;2013:1378-1385 [FREE Full text] [Medline: 24551414]
79. Way A, Haque R, Xie G, Gaspari F, Popović M, Poncelas A. Rapid development of competitive translation engines for access to multilingual COVID-19 information. Informatics 2020;7(2):19 [FREE Full text] [doi: 10.3390/informatics7020019]
80. Wu C, Xia F, Deleger L, Solti I. Statistical machine translation for biomedical text: are we there yet? AMIA Annu Symp Proc 2011;2011:1290-1299 [FREE Full text] [Medline: 22195190]
81. Xie W, Ji M, Huang R, Hao T, Chow CY. Predicting risks of machine translations of public health resources by developing interpretable machine learning classifiers. Int J Environ Res Public Health 2021 Aug 20;18(16):8789 [FREE Full text] [doi: 10.3390/ijerph18168789] [Medline: 34444538]
82. Yan R, Liao W, Cui J, Zhang H, Hu Y, Zhao D. Multilingual COVID-QA: learning towards global information sharing via web question answering in multiple languages. In: Proceedings of the 2021 Web Conference. 2021 Presented at: WWW '21; April 19-23, 2021; Ljubljana, Slovenia p. 2590-2600 URL: https://dl.acm.org/doi/10.1145/3442381.3449991 [doi: 10.1145/3442381.3449991]
83. Yang LW, Ng WY, Lei X, Tan SC, Wang Z, Yan M, et al. Development and testing of a multi-lingual natural language processing-based deep learning system in 10 languages for COVID-19 pandemic crisis: a multi-center study. Front Public Health 2023 Feb 13;11:1063466 [FREE Full text] [doi: 10.3389/fpubh.2023.1063466] [Medline: 36860378]
84. Yepes AJ, Névéol A, Neves M, Verspoor K, Bojar O, Boyer A, et al. Findings of the WMT 2017 biomedical translation shared task. In: Proceedings of the 2017 Conference on Machine Translation. 2017 Presented at: WMT '17; September 7-8, 2017; Copenhagen, Denmark p. 234-247 URL: https://aclanthology.org/W17-4719.pdf [doi: 10.18653/v1/w17-4719]
85. Zeng-Treitler Q, Kim H, Rosemblat G, Keselman A. Can multilingual machine translation help make medical record content more comprehensible to patients? Stud Health Technol Inform 2010;160(Pt 1):73-77. [Medline: 20841653]
86. Ziganshina LE, Yudina EV, Gabdrakhmanov AI, Ried J. Assessing human post-editing efforts to compare the performance of three machine translation engines for English to Russian translation of Cochrane plain language health information: results of a randomised comparison. Informatics 2021;8(1):9 [FREE Full text] [doi: 10.3390/informatics8010009]
87. Home page. Leibniz ScienceCampus - DiPH. URL: https://www.lsc-digital-public-health.de/en/ [accessed 2023-11-06]

Abbreviations

CALD: culturally and linguistically diverse
M&E: monitoring and evaluation
MeSH: Medical Subject Headings
MT: machine translation
PH: public health
PPC: population or participants, concept, and context
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 13.07.23; peer-reviewed by W Han, M Xu; comments to author 23.08.23; revised version received 11.09.23; accepted 06.10.23; published 20.11.23.

Please cite as:
Herrera-Espejel PS, Rach S
The Use of Machine Translation for Outreach and Health Communication in Epidemiology and Public Health: Scoping Review
JMIR Public Health Surveill 2023;9:e50814
URL: https://publichealth.jmir.org/2023/1/e50814
doi:10.2196/50814
PMID:37983078

Original Paper

Projections of Lung Cancer Incidence by 2035 in 40 Countries Worldwide: Population-Based Study

Ganfeng Luo[1*], PhD; Yanting Zhang[2,3*], PhD; Jaione Etxeberria[4,5], PhD; Melina Arnold[6], PhD; Xiuyu Cai[7], PhD; Yuantao Hao[2,8], PhD; Huachun Zou[1,9], PhD

[1]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

[2]Department of Medical Statistics, School of Public Health, Sun Yat-sen University, Guangzhou, China

[3]Department of Epidemiology and Health Statistics, School of Public Health, Guangdong Medical University, Dongguan, China

[4]Department of Statistics, Computer Science and Mathematics, Public University of Navarre, Navarre, Spain

[5]Institute for Advanced Materials and Mathematics (INAMAT2), Public University of Navarre, Navarre, Spain

[6]Cancer Surveillance Branch, International Agency for Research on Cancer, Lyon, France

[7]Department of VIP Inpatient, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

[8]Peking University Center for Public Health and Epidemic Preparedness & Response, Peking University, Beijing, China

[9]Kirby Institute, University of New South Wales, Sydney, Australia

[*]these authors contributed equally

Corresponding Author:
Huachun Zou, PhD
School of Public Health (Shenzhen), Sun Yat-sen University
No. 66, Gongchang Road, Guangming District
Shenzhen, 518107
China
Phone: 86 20 87335651
Email: zouhuachun@mail.sysu.edu.cn

Abstract

Background: The global burden of lung cancer (LC) is increasing. Quantitative projections of the future LC burden in different world regions could help optimize the allocation of resources and provide a benchmark for evaluating LC prevention and control interventions.

Objective: We aimed to predict the future incidence of LC in 40 countries by 2035, with an emphasis on country- and sex-specific disparities.

Methods: Data on LC incidence from 1978 to 2012 were extracted from 126 cancer registries of 40 countries in *Cancer Incidence in Five Continents Volumes V-XI* and used for the projection. Age-standardized incidence rates (ASRs) per 100,000 person-years and the number of incident cases were predicted through 2035, using the NORDPRED age-period-cohort model.

Results: Global ASRs of the 40 studied countries were predicted to decrease by 23% (8.2/35.8) among males, from 35.8 per 100,000 person-years in 2010 to 27.6 in 2035, and increase by 2% (0.3/16.8) among females, from 16.8 in 2010 to 17.1 in 2035. The ASRs of LC among females are projected to continue increasing dramatically in most countries by 2035, with peaks after the 2020s in most European, Eastern Asian, and Oceanian countries, whereas the ASRs among males will continue to decline in almost all countries. The ASRs among females are predicted to almost reach those among males in Ireland, Norway, the United Kingdom, the Netherlands, Canada, the United States, and New Zealand in 2025 and in Slovenia in 2035 and even surpass those among males in Denmark in 2020 and in Brazil and Colombia in 2025. In 2035, the highest ASRs are projected to occur among males in Belarus (49.3) and among females in Denmark (36.8). The number of new cases in 40 countries is predicted to increase by 65.32% (858,000/1,314,000), from 1.31 million in 2010 to 2.17 million in 2035. China will have the largest number of new cases.

Conclusions: LC incidence is expected to continue to increase through 2035 in most countries, making LC a major public health challenge worldwide. The ongoing transition in the epidemiology of LC highlights the need for resource redistribution and improved LC control measures to reduce future LC burden worldwide.

(JMIR Public Health Surveill 2023;9:e43651) doi:10.2196/43651

KEYWORDS

lung cancer; incidence; projections; temporal trends; worldwide

Introduction

Background

Globally, lung cancer (LC) represents the first and third most commonly diagnosed cancer among males and females in 2020, respectively [1]. In 2020, an estimated 2.2 million new LC cases occurred worldwide, representing 14.3% (1.4 million) and 8.4% (0.8 million) of all new cancer diagnoses among males and females, respectively [1]. LC incidence rates have been steadily decreasing among males in most high-income countries (eg, European and Oceanian countries) in recent decades, whereas they have been rapidly increasing among females [2-4].

Research Significance

However, the global LC burden continues to increase owing to population growth, population aging, and changes in the prevalence of risk factors (eg, smoking and air pollution), which will further exert tremendous strain on populations and health systems worldwide [2-4]. Quantitative projections of the future LC burden in different world regions, in terms of expected number of new cases and incidence rates in both absolute size and trends, can help optimize the allocation of resources for screening, diagnosis, and therapy and provide a benchmark for evaluating LC prevention and control interventions worldwide, with the aim of reducing the LC burden and promoting health equity [5].

Objective

Historical LC incidence trends have been comprehensively evaluated worldwide [2-4,6]. However, to our knowledge, no previous study has predicted the future burden of LC incidence on a global scale. We aimed to quantify the future LC incidence in 40 countries by 2035 using long-standing and high-quality population-based data, with an emphasis on country- and sex-specific disparities.

Methods

Data Sources

Data on LC incidence (International Classification of Diseases 10th Revision C33-34) stratified by a 5-year period of diagnosis (1978-1982, ... , 2008-2012), sex, and 5-year age group were extracted from national and regional population-based cancer registries available in *Cancer Incidence in Five Continents* volumes V to XI [7,8]. The specific requirements for the inclusion of a registry were at least 15 consecutive years of data and compilation in the latest volume (XI) of the *Cancer Incidence in Five Continents* series [2,9,10]. These criteria are indicative of each registry's data quality over time, given that the editorial process involves a detailed assessment of the comparability, completeness, and validity of the incidence data [2,9,10]. A total of 126 cancer registries from 40 countries in 10 world regions (Northern Europe, Western Europe, Southern Europe, Central and Eastern Europe, North America, Central and South America, Eastern Asia, Southeastern Asia, Southwestern Asia, and Oceania) were included. Of these, 19 countries contributed national data for the analysis (Denmark, Estonia, Iceland, Ireland, Lithuania, Norway, Austria, the Netherlands, Croatia, Cyprus, Malta, Slovenia, Belarus, Bulgaria, the Czech Republic, Slovakia, Costa Rica, Israel, and New Zealand). For the remaining 21 countries (the United Kingdom, France, Germany, Switzerland, Italy, Spain, Poland, Canada, the United States, Brazil, Chile, Colombia, Ecuador, China, Japan, Republic of Korea, Philippines, Thailand, India, Turkey, and Australia), data from multiple regional registries in a given country were aggregated to obtain a proxy for the national incidence. Population predictions (medium-fertility variant) were obtained from the United Nations World Population Prospects 2022 Revision up to 2035 by country, year, sex, and age [11].

Statistical Analysis

Incidence rates per 5-year period of diagnosis and 5-year age group by sex were calculated based on the corresponding incident cases and population data from population-based cancer registries. Age-standardized incidence rates (ASRs) per 100,000 person-years were calculated based on the world standard population [12]. The number of cases by age group at the national scale were estimated by applying regional age-specific incidence rates to the national populations [9,10]. Lifetime cumulative risks were calculated for those aged 0 to 74 years and are expressed as percentages, which can be interpreted as the probability that an individual will develop LC before the age of 75 years in the absence of competing causes of death [13].

In this study, the well-known NORDPRED age-period-cohort model [9,10,14,15] was used to predict the number of new LC cases and incidence rates from 2013 to 2017 to 2033 to 2037 by country, sex, and age. Let us assume that, conditional on rates, the number of incident LC cases follows a Poisson distribution:

Where y_{ap} and n_{ap} represent the number of incident cases and person-years at risk, respectively, in the age group a=0−4, ... , 85+ and in the period P=1978-1982, ... , 2008-2012.

Then, the NORDPRED age-period-cohort model using a power-5 link function is defined as follows:

In this expression, A_a is the age component of age group a; D is the common linear drift parameter of both calendar period p and birth cohort c; is the nonlinear period component of period ; and is the nonlinear cohort component of cohort c. Birth cohort c was extracted by subtracting the midpoint of the 5-year age group a from the corresponding midpoint of the 5-year calendar period P [2]. The 3 to 7 recent 5-year observed periods (depending on data availability) were extrapolated using a power-5 link function to level off growth, with a projection of drift parameter *D* that was attenuated by 25%, 50%, 75%, and 75% in the second, third, fourth, and fifth prediction periods, respectively [14]. Furthermore, we assessed the prediction performance of the NORDPRED age-period-cohort model, with detailed information and the corresponding findings presented in Multimedia Appendix 1 [15-20].

The projected number of new cases in 2035 were calculated by applying the predicted age-specific rates for 2033 to 2037 using the United Nations national population forecasts for 2035 for each country [14]. Mean annual differences in the number of predicted cases in 2035 (the midpoint of calendar period 2033-2037) relative to 2010 (the midpoint of calendar period 2008-2012) are partitioned into changes in risk (rates) and changes in demographics (population size and age structure) [15]; the detailed information is presented in Multimedia Appendix 1. The number of predicted cases and ASR in 2035 are presented and compared with those in 2010. A percentage change of ≥5% was defined as an increase or decrease; otherwise, the number and rates were considered stable [21]. Throughout the paper, the number of cases has been rounded to avoid spurious precision. In some cases, this creates small discrepancies with the displayed totals and percentages, which are based on the data before rounding.

Data management and modeling analyses were performed using R software (version 4.0.2; R Foundation for Statistical Computing) [22] and the NORDPRED package.

Ethical Considerations

This study does not involve human participants and animal subjects. Ethics approval was not required for this study as the study used existing nonidentifiable data that were aggregated and population level.

Results

Changes in LC Incidence Rates: 2010 to 2035

Global ASRs of the 40 studied countries are predicted to decrease by 23% (8.2/35.8) among males, from 35.8 per 100,000 person-years in 2010 to 27.6 in 2035, and increase by 2% (0.3/16.8) among females, from 16.8 in 2010 to 17.1 in 2035. Among the 40 studied countries, the ASRs are predicted to decline among males in 37 countries and among females in 17 countries between 2010 and 2035 (Figures 1-3; Multimedia Appendices 2 and 3). The largest decreases in ASRs are projected in the Philippines (−21.5/36.9, −58%; ASR 36.9 per 100,000 person-years in 2010 vs 15.4 in 2035), Colombia (−7.7/13.9, −55%; 13.9 vs 6.2), and Slovenia (−25.3/48.2, −53%; 48.2 vs 22.9) among males and in the Philippines (−5.3/12.2, −44%; 12.2 vs 6.9), Israel (−5/11.4, −43%; 11.4 vs 6.4), and Turkey (−3.3/8.1, −40%; 8.1 vs 4.8) among females. For each region, the largest reductions in ASRs are projected in Estonia (−26.3/57, −46%; ASR 57.0 per 100,000 person-years in 2010 vs 30.7 in 2035) for males and in Iceland (−9.2/33.0, −28%; 33.0 vs 23.8) for females in Northern Europe; in Switzerland for both sexes (−13.5/37.2, −36%; 37.2 vs 23.7 for males; −2.4/20.4, −12%; 20.4 vs 18.0 for females) in Western Europe; in Slovenia (−25.3/48.2, −53%; 48.2 vs 22.9) for males in Southern Europe; in the Czech Republic (−25.7/51.2, −50%; 51.2 vs 25.5) for males and in Poland (−2.7/13.4, −20%; 13.4 vs 10.7) for females in Central and Eastern Europe; in the United States for both males (−13.7/39.4, −35% 39.4 vs 25.7) and females (−10.6/30.2, −35%; 30.2 vs 19.6) in North America; in Colombia (−7.7/13.9, −55%; 13.9 vs 6.2) for males and in Costa Rica (−1.0/4.3, −24%; 4.3 vs 3.3) for females in Central and South America; in the Republic of Korea (−8.0/42.8, −19%; 42.8 vs 34.8) for males in Eastern Asia; in the Philippines for both sexes (−21.5/36.9, −58%; 36.9 vs 15.4 for males; −5.3/12.2, −44%; 12.2 vs 6.9 for females) in Southeastern Asia; in Turkey (−33.0/69.3, −48%; 69.3 vs 36.3) for males and in Israel (−5.0/11.4, −43%; 11.4 vs 6.4) for females in Southwestern Asia; and in New Zealand for both sexes (−6.8/29.5, −23%; 29.5 vs 22.7 for males; −3.4/24.2, −14%; 24.2 vs 20.8 for females) in Oceania.

However, the ASRs will increase in 2 countries for males and in 20 countries for females from 2010 to 2035.

Among males, percentage increases in ASRs are predicted in Cyprus (14.6/33.4, 44%; ASR 33.4 per 100,000 person-years in 2010 vs 48.0 in 2035) and Ecuador (0.7/8.0, 9%; 8.0 vs 8.7). among females, the largest increases in ASRs were predicted in Estonia (7.3/10.7, 68%; ASR 10.7 per 100,000 person-years in 2010 vs 18.0 in 2035), Malta (6.5/10.9, 60%; 10.9 vs 17.4), and Cyprus (4.8/8.7, 54%; 8.7 vs 13.5). As for each region, the steepest increases in ASRs are predicted in Estonia (7.3/10.7, 68%; ASR 10.7 per 100,000 person-years in 2010 vs 18.0 in 2035) in Northern Europe; France (3.3/15.8, 20%; 15.8 vs 19.1) in Western Europe; Malta (6.5/10.9, 60%; 10.9 vs 17.4) in Southern Europe; Slovakia (4.8/13.1, 37%; 13.1 vs 17.9) in Central and Eastern Europe; Chile (1.6/4.7, 33%; 4.7 vs 6.3) in Central and South America; Japan (6.5/16.8, 38%; 16.8 vs 23.3) in Eastern Asia; and India (0.8/4.1, 19%; 4.1 vs 4.9) in Southwestern Asia. The ASRs among females are predicted to peak after the 2020s in 27 countries (most European, Eastern Asian, and Oceanian countries), whereas those among males have already peaked in the historical period in almost all countries (except for Cyprus, Ecuador, and Japan).

Although the ASRs among males have been historically higher than those among females in all countries (except for Iceland), the sex gaps are gradually narrowing (Figures 4 and 5). The male to female ASR ratios are projected to decline in 33 countries from 2010 to 2035. The largest reductions in male to female ratios are predicted for Estonia (−3.6/5.3, −68%; sex ratio 5.3 in 2010 vs 1.7 in 2035), Brazil (−1.1/1.8, −62%; 1.8 vs 0.7), Slovenia (−1.6/2.8, −58%; 2.8 vs 1.2), Spain (−2.2/4.4, −50%; 4.4 vs 2.2), and Slovakia (−2.0/4.1, 48%; 4.1 vs 2.1). ASRs among females are predicted to almost reach those among males in Ireland, Norway, the United Kingdom, the Netherlands, Canada, the United States, and New Zealand in 2025; in Slovenia in 2035; and even surpass those among males in Denmark in 2020 and in Brazil and Colombia in 2025.

In 2035, LC incidence is projected to vary 8-fold among males and 10-fold among females across countries (Figure 3; Multimedia Appendices 2 and 3). The highest ASRs were projected in Belarus (49.3 per 100,000 person-years), Cyprus (48.0), Japan (45.3), France (42.0), and Bulgaria (40.1) among males and in Denmark (32.9), the Netherlands (28.7), the United Kingdom (28.0), Ireland (27.1), and China (24.7) among females. The male to female ratio is projected to vary 11-fold across countries, ranging from 0.7 in Brazil to 7.5 in Turkey.

Figure 1. Trends in age-standardized (world) incidence rates of lung cancer in European countries, 1980 to 2035: observed rates (solid line) and projected rates (dashed line). The blue line stands for males, and the red line stands for females. Age-standardized (world) rates per 100,000 person-years.

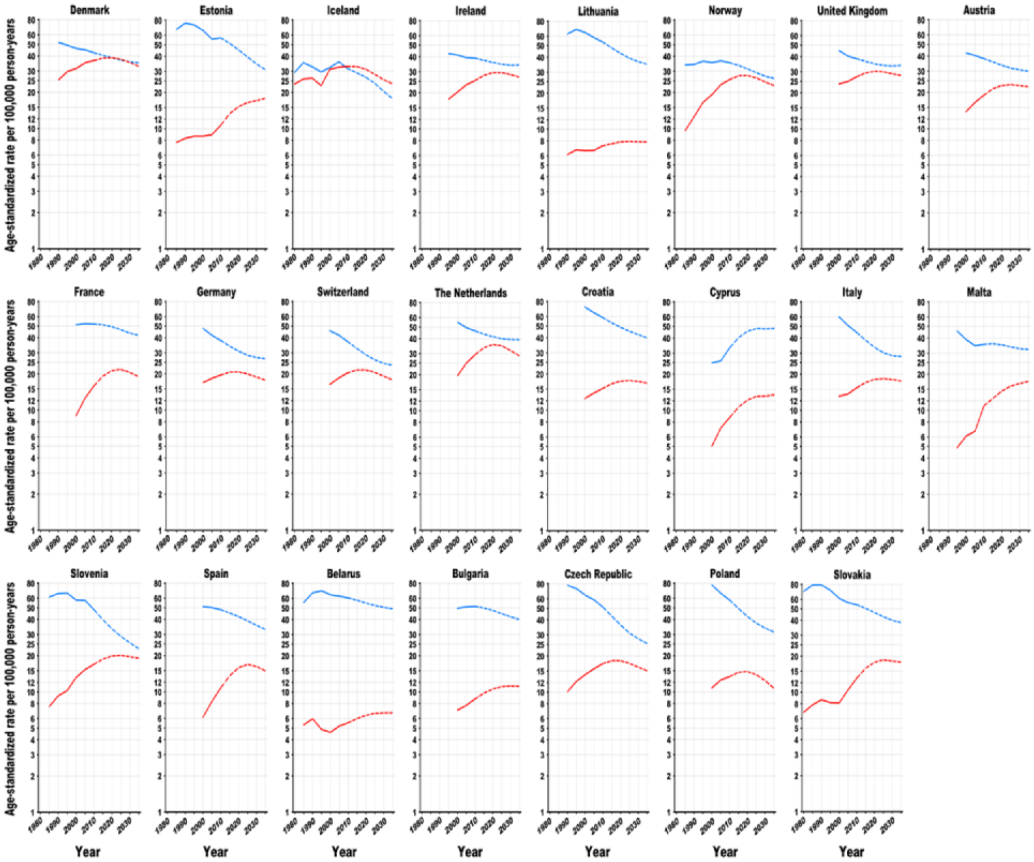


Figure 2. Trends in age-standardized (world) incidence rates of lung cancer in American, Asian, and Oceanian countries, 1980 to 2035: observed rates (solid line) and projected rates (dashed line). The blue line stands for males, and the red line stands for females. Age-standardized (world) rates per 100,000 person-years.

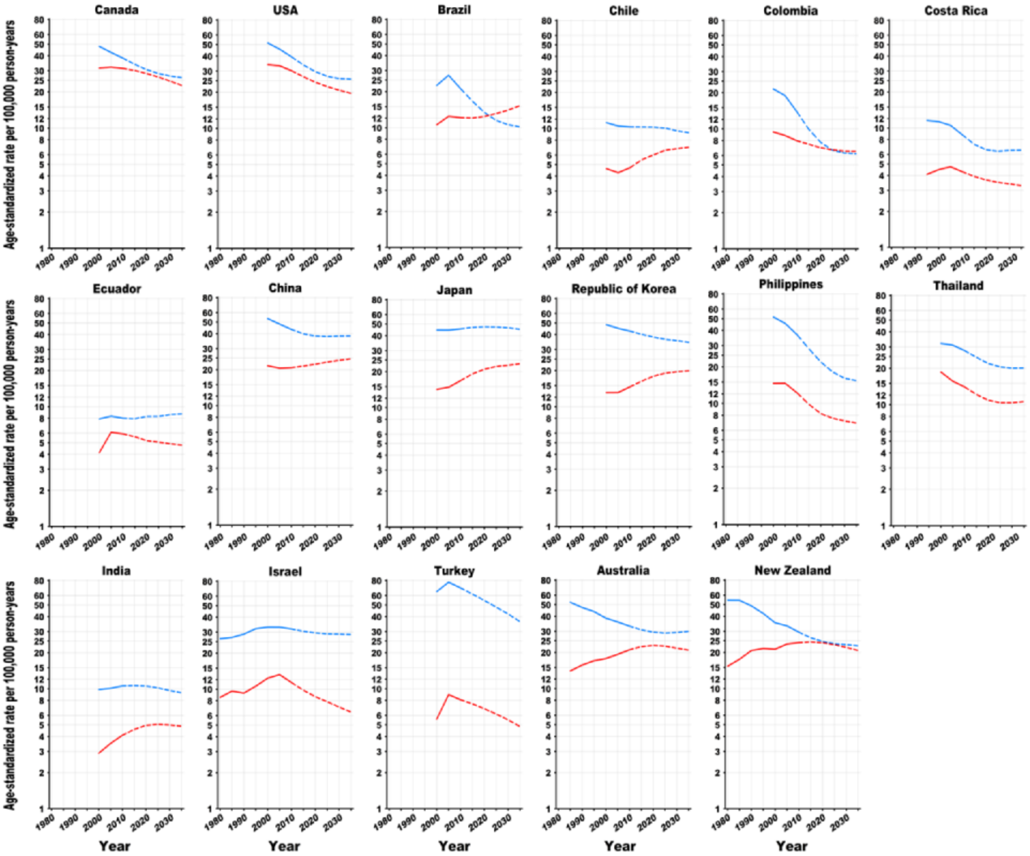


Figure 3. Lung cancer age-standardized incidence rates in 2010 and projected rates in 2035 in 40 countries. Rates for 2010 and 2035 represent average rates for the 5-year period centered on the respective year. Blue color stands for males, and red color stands for females. Age-standardized (world) rates per 100,000 person-years.

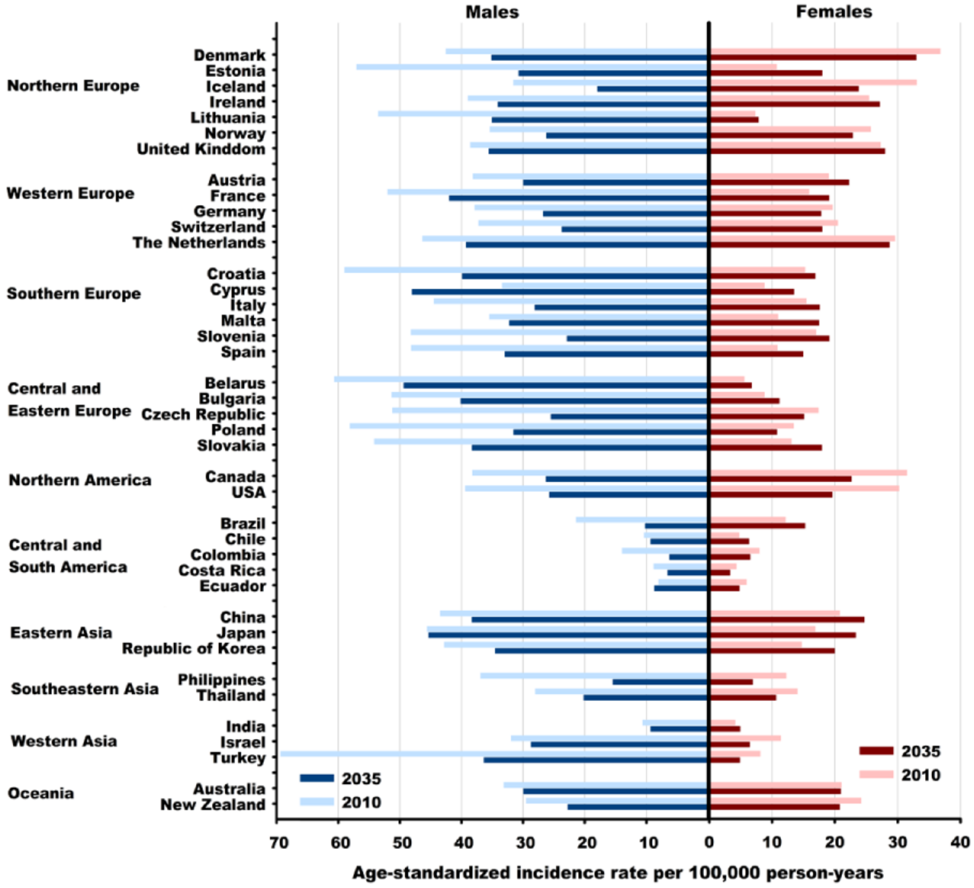


Figure 4. Trends in the male to female age-standardized (world) incidence rate ratios of lung cancer in European countries, 1980 to 2035: observed rate ratios (solid line) and projected rate ratios (dashed line).

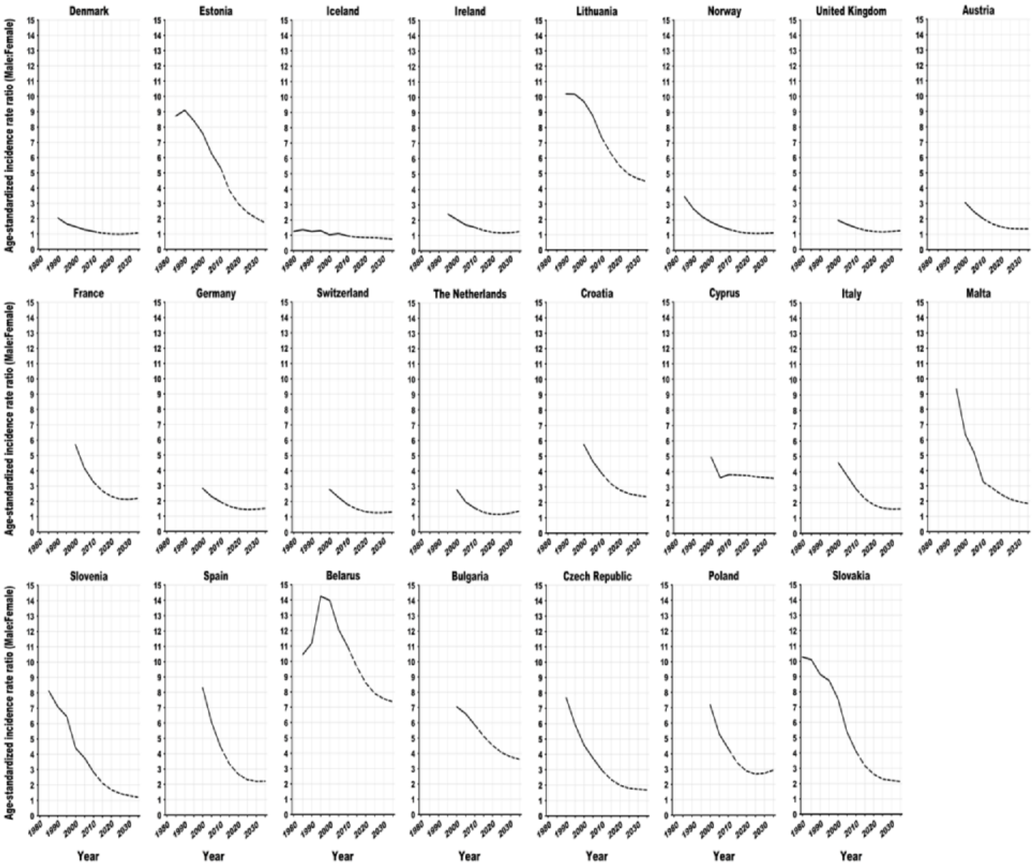
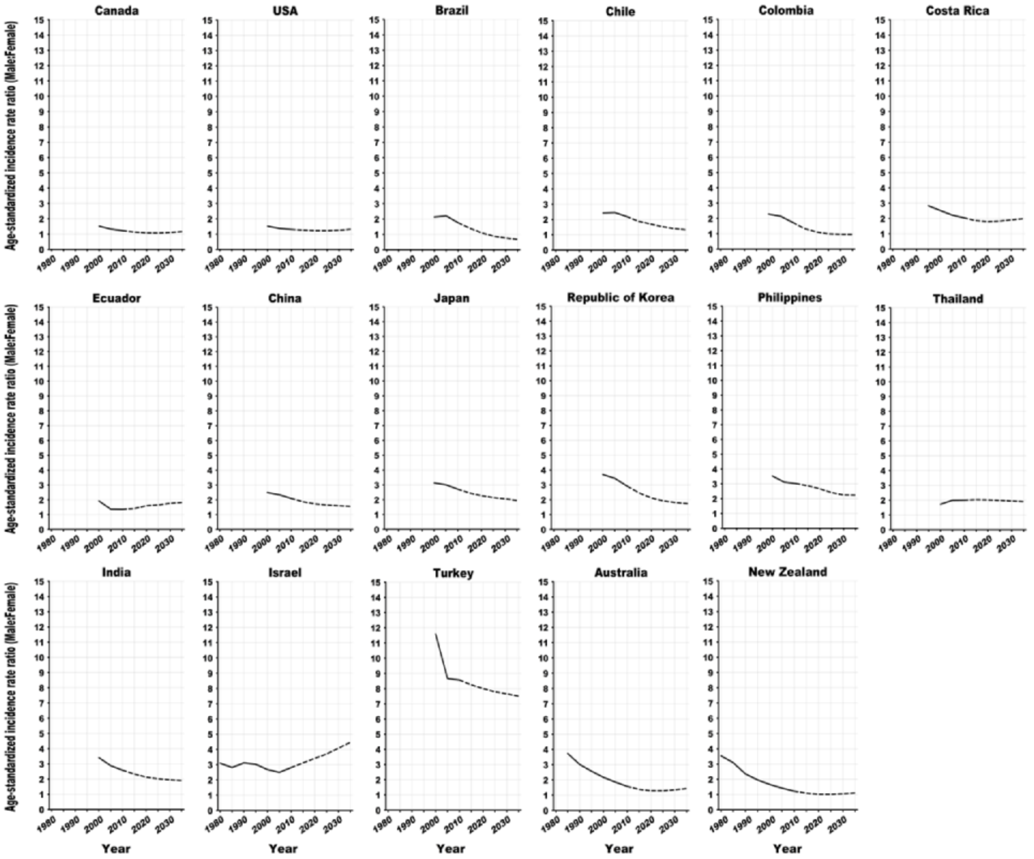


Figure 5. Trends in the male to female age-standardized (world) incidence rate ratios of lung cancer in American, Asian, and Oceanian countries, 1980 to 2035: observed rate ratios (solid line) and projected rate ratios (dashed line).



Changes in the Number of New LC Cases: 2010 to 2035

Overall, the number of new cases in the 40 included countries is predicted to increase by 50.19% (429,000/854,000), from approximately 0.85 million in 2010 to approximately 1.28 in 2035, among males and is set to nearly double, from approximately 0.46 million to approximately 0.89, among females (Multimedia Appendices 2 and 3). When stratified by geographic region, the most rapid increases were predicted in Eastern Asia (342,000/433,000, 79.00% for males and 303,000/216,000, 140.05% for females).

When analyzed by country, the number of new cases is expected to increase in 32 countries among males and in all 40 countries among females between 2010 and 2035 (Multimedia Appendices 2 and 3). The predicted increases among males in most of these 32 countries are mainly due to population growth and aging, regardless of marked decreases in LC risk, whereas those among females in most countries in Southern, Western, and Central and Eastern Europe are largely due to marked increases in LC risk (Figures 6 and 7). The most rapid increases were predicted in Cyprus (459/257, 178.3%) among males and Chile (1253/620, 202.2%) among females (Multimedia Appendices 2 and 3). In 2035, the largest number of new cases was projected in China (655,000), Japan (112,000), and the United States (105,000) among males and China (451,000), the United States (95,000), and Japan (61,000) among females.

Figure 6. Projected changes in the number of new lung cancer cases between 2010 and 2035 in 40 countries among males, partitioned into population and risk change. Numbers for 2010 and 2035 represent average numbers for the 5-year period centered on the respective year. Due to population change: projected change in the number of cases due to population growth and aging. Due to risk change: projected change in the number of cases due to changes in incidence rates. Blue color stands for population change, and red color stands for risk change.

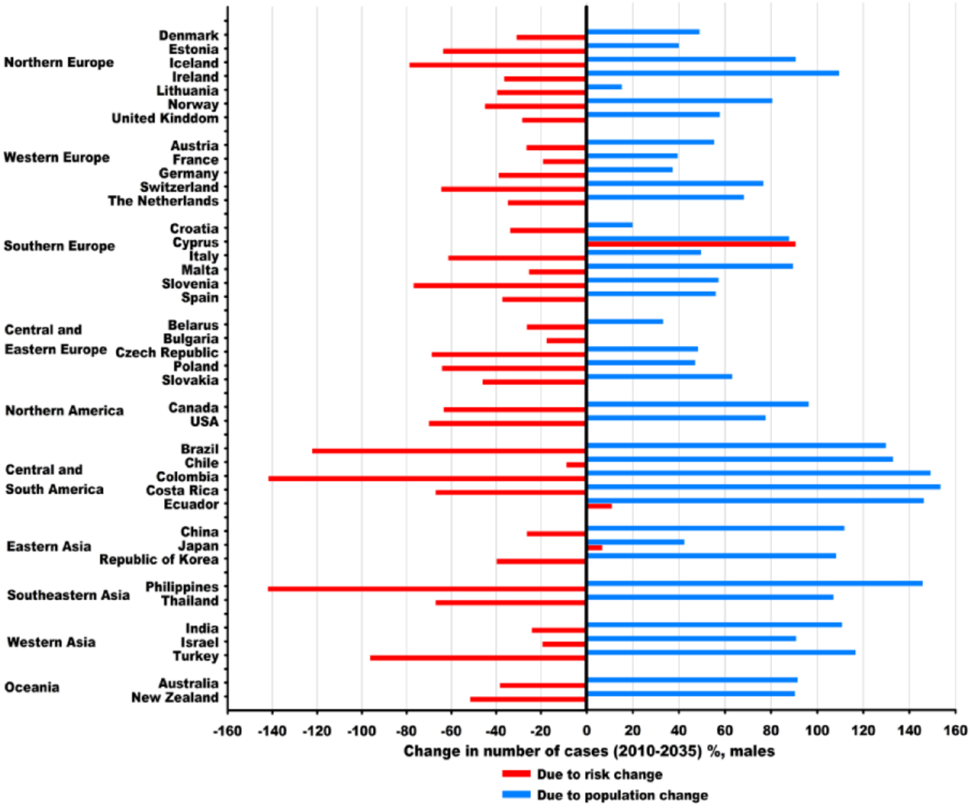
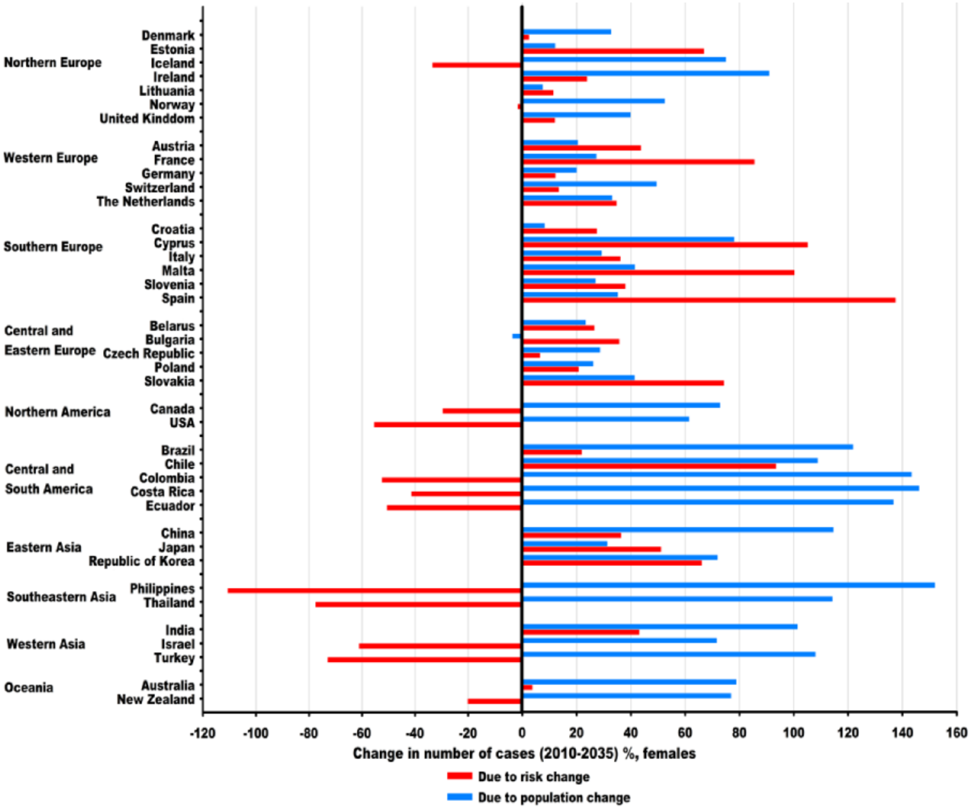


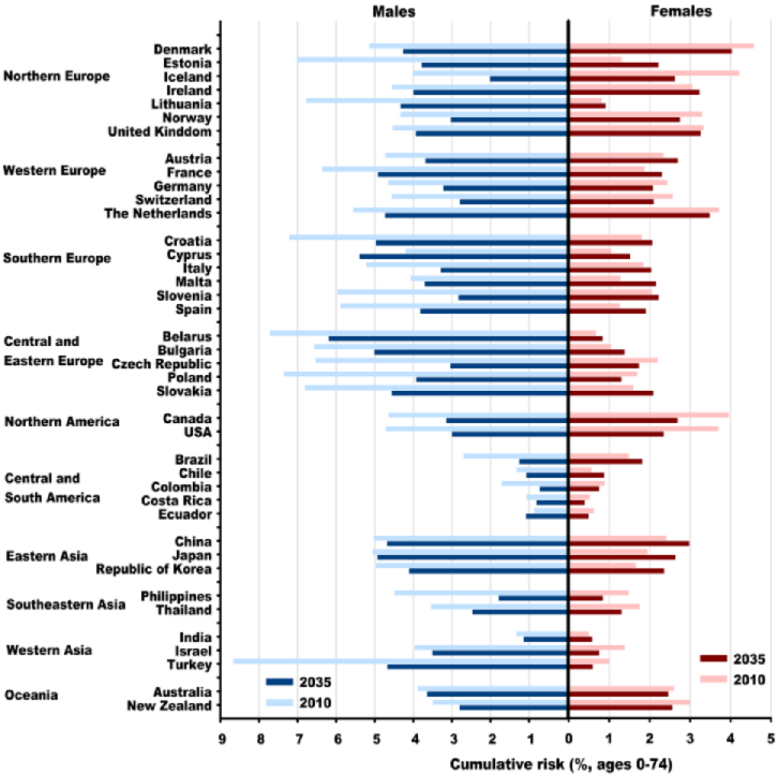
Figure 7. Projected changes in the number of new lung cancer cases between 2010 and 2035 in 40 countries among females, partitioned into population and risk change. Numbers for 2010 and 2035 represent average numbers for the 5-year period centered on the respective year. Due to population change: projected change in the number of cases due to population growth and aging. Due to risk change: projected change in the number of cases due to changes in incidence rates. Blue color stands for population change, and red color stands for risk change.



Cumulative Risk of LC Incidence in 2035

The cumulative risks among males are expected to decrease in almost all countries (except Cyprus and Ecuador) between 2010 and 2035, whereas those among females are expected to increase in 20 countries (especially in Eastern Asian and Southern European countries; Figure 8; Multimedia Appendices 2 and 3). In 2035, the highest cumulative risk among males is expected in Belarus (6.2/100, 6.2%), Cyprus (5.4/100, 5.4%), Bulgaria (5/100, 5%), Croatia (5/100, 5%), and Japan (4.9/100, 4.9%), which indicates that approximately 5 to 6 of 100 males will be diagnosed with LC before the age of 75 years in the aforementioned countries. Among females, the highest cumulative risks are expected in Denmark (4/100, 4%), the Netherlands (3.5/100, 3.5%), the United Kingdom (3.3/100, 3.3%), Ireland (3.2/100, 3.2%), and China (3/100, 3%).

Figure 8. Cumulative risk (in percent, ages 0-74 years) of lung cancer in 2010 and 2035 in 40 countries. Blue color stands for males, and red color stands for females.



Discussion

Principal Findings

In our study, the ASRs of LC among females are predicted to increase in half of the countries between 2010 and 2035, while those among males are expected to decrease in almost all studied countries. The ASRs among females are expected to peak after the 2020s in most European, Eastern Asian, and Oceanian countries and will reach and surpass those among males in 11 countries after the 2020s. In 2035, high rates are predicted to mainly occur among males in Central and Eastern European countries and among females in Northern European countries. The largest increases in the number of new cases in both sexes between 2010 and 2035 are predicted in Eastern Asia. China will have the largest number of new cases in both sexes by 2035.

Interpretation of Changes in LC Incidence Rates

Tobacco smoking is the leading risk factor for LC, with 76.2% (74.6 to 77.8) of LC-related deaths among males and 38.9% (36.7 to 40.9) of LC-related deaths among females attributed to smoking in 2019 [3]. Temporal trends in LC incidence are largely affected by smoking patterns in the past 30 to 40 years [23,24], and the future LC burden will depend on historic smoking patterns at the population level [25]. Smoking prevalence among males in most European countries (eg, the United Kingdom and Spain) peaked during the 1950s and 1960s, whereas that among females peaked in the 1980s and 1990s [23,26-32]. For example, the smoking prevalence among Spanish females increased from 5.8% in 1970 to 26.5% in 1990 and then decreased after the 2000s [23,30]; this could largely explain the predicted increase in the LC incidence rate and the peak LC rate in 2025 among females. Similarly, the lag duration of 30 to 40 years combined with the maximum smoking prevalence during the 1980s and 1990s [26,27,31,32] could indicate that LC incidence rates among females will peak after the 2020s in many European countries. In addition, a recent study showed that the smoking-attributable fractions of LC mortality among males are predicted to decline in Denmark and the Netherlands from 2009 to 2050, whereas those among females will increase until it peaks in 2028 in Denmark and 2033 in the Netherlands [33], similar to our projections of changes in LC rates from 2010 to 2035. Notably, smoking patterns in younger birth cohorts may directly indicate future LC incidence rates [32,34,35]. The predicted increase in LC incidence rates among females in most European countries could be largely explained by the cohort effect because of the increased smoking prevalence in cohorts born after the 1950s [32,36]. The predicted decreases in LC incidence rates in the United States in both sexes might largely result from changes in smoking habits in cohorts born after the 1940s, who mainly smoked low-tar and filter-tipped cigarettes [2,37,38].

Interpretation of the Discrepancy of Changes in LC Incidence Rates Between Sexes

The convergence of decreasing male and increasing female incidence rates of LC in our study could be largely explained by the changes in smoking prevalence. Globally, females started smoking later than males (the 1950s), but the smoking prevalence increased rapidly until the 1990s and then remained relatively stable [23,26-32,39]. For example, in France, tobacco smoking increased from 0.4 daily smoked cigarettes per woman in 1953 to 3.7 in 1991 and thereafter remained stable, whereas males smoked approximately 8.6 cigarettes per day in 1953 until a peak in 1980 (9.6 cigarettes per day) and declined steadily

(5.2 cigarettes per day in 2003) [39]. Notably, smoking behavior and smoking prevalence among males and females have become increasingly similar in the younger generation worldwide [40,41], indicating the convergence of decreasing male and increasing female incidence rates of LC.

Interpretation of Geographic Distributions in Predicted LC Incidence Rates and Cumulative Risks

In 2035, the incidence rates of LC are projected to vary significantly across regions for both sexes. The highest predicted rate and cumulative risk among males in Central and Eastern European countries could be largely attributable to the historically higher smoking prevalence than that of Western countries and delays in the implementation of smoking prevention and cessation measures [42-44]. LC mortality rates attributable to smoking were also estimated to be the highest in Central and Eastern European countries in 2019 [3]. Among females, the highest predicted LC rate and cumulative risk in Northern European countries could be largely explained by the historically high LC rate and the fact that smoking prevalence among females only began to decrease after the late 1980s [31,32].

Interpretation of Changes in New LC Cases

Worldwide, the reduction in LC risk among males will only partially offset the expected increase in the number of new LC cases because of population growth and aging. However, the increasing future LC burden among females will be driven by increasing LC risks, and thus mainly tobacco smoking patterns and trends, in most countries in Southern, Western, and Central and Eastern Europe. In addition, the largest number of new LC cases is predicted in China, accounting for nearly half (45%) of the total number of new cases among all studied countries in 2035, and more than half (60%) of the total increase in the number of new cases between 2010 and 2035. A recent study showed that the LC mortality rates attributed to smoking and ambient particulate matter pollution in China were estimated to be 80.8% and 20.8% among males and 23.4% and 19.7% among females in 2017, respectively [45]. China has the largest smoking population in the world, with an estimated 301 million smokers, and the smoking proportions of males and females in the Chinese population are 52.9% and 2.4% in 2010 [46]. More than half of the LC cases attributed to exposure to ambient particulate matter pollution have been reported in China and other East Asian countries [47,48]. In addition, China has experienced rapid population growth and a transition to an aging society in recent decades [49]. Thus, the substantial future burden of LC in China might largely result from the large smoking population, deteriorating air pollution, and changing demographics, and will, to some extent, hinder the realization of global noncommunicable disease control goals with a reduction of one-third in premature mortality from noncommunicable diseases by 2030 [50]. Tobacco control programs and clean air legislation should be strengthened and additional resources should be allocated to reduce the LC burden in China.

Comparison With Prior Work

The Global Cancer Observatory of the International Agency for Research on Cancer provides the future LC incidence worldwide by 2040, based on demographic projections estimated by the United Nations Development Programme and scenarios of stable rates from the baseline year 2020 [51]. However, projections of future LC incidence by 2040 in the Global Cancer Observatory did not consider changes in the incidence rates of LC. Our predictions have accounted for historical changes in LC incidence rates using the NORDPRED age-period-cohort model, which could largely reflect the real situation of LC incidence. To the best of our knowledge, this is the first study to predict the future burden of LC incidence from a global perspective based on incidence data from high-quality population-based cancer registries, enabling direct comparisons of findings between countries. The well-known modern method based on the NORDPRED age-period-cohort model was applied to provide estimates of future LC burden in our study.

Limitations

However, this study has some limitations. First, the trend-based predictions in our study might have a certain degree of uncertainty because we assumed that the historical trends in LC incidence will continue into the future in the model [5]. However, when comparing our predicted ASRs with those already observed in, for example, Australia, Norway, and the Netherlands, these figures seem fairly close (eg, observed ASR 30.7 in Australian males in 2015 [52] vs 30.9 predicted in 2015, 21.8 in Australian females [52] vs 22.4; observed ASR 31.3 in Norwegian males in 2015 [53] vs 33.6 predicted in 2015, 28.0 in Norwegian females [53] vs 27.6; and observed ASR 38.6 in Dutch males in 2019 [54] vs 41.4 predicted in 2020, 33.1 in Dutch females [54] vs 35.7); these might largely indicate the accuracy of our projections. Second, for lower-risk and small populations (eg, Malta and Cyprus), random variation might exist. Third, we were unable to predict the future LC burden in many countries in Africa, Latin America, and Asia because of the lack of high-quality data. Fourth, the aggregated data from multiple regional registries in a given country might not fully represent the national level, although this method has been applied to cover the largest geographic area in many previously published papers [2,55,56]. Figures from regional cancer registries with limited resources could be underestimated and those from registries with abundant resources could be overestimated. Caution should be exercised when interpreting these findings.

Conclusions

The LC incidence rates among females are predicted to continue increasing dramatically in most countries by 2035, with peaks after the 2020s in most European, Eastern Asian, and Oceanian countries, whereas incidence rates among males are predicted to continue declining in almost all countries. The sex difference in the worldwide incidence of LC will further narrow down in the future. Geographic variations in LC incidence rates will remain high, with the highest rate in Central and Eastern Europe among males and in Northern Europe among females. The number of new LC cases are expected to continue to increase through 2035 in most countries because of population aging and growth, as well as changes in the prevalence of risk factors, making LC a major challenge to public health worldwide. These ongoing transitions in LC epidemiology are highly relevant to future cancer control and clinical practice. Clinicians can expect to observe more cases of LC in the future. The findings highlight the need for the redistribution of resources and improved LC control measures (eg, smoking prevention and cessation programs and air quality management programs) to reduce the future LC burden worldwide.

Acknowledgments

The authors gratefully acknowledge the contribution of all the cancer registry staff in generating the cancer registry data used in our study. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization. This study was supported by the Natural Science Foundation of China Excellent Young Scientists Fund (82022064), Natural Science Foundation of China International/Regional Research Collaboration Project (72061137001), the National Natural Science Foundation of China (81973150), the Natural Science Foundation of China Young Scientist Fund (81703278), the National Science and Technology Major Project of China (2018ZX10721102), the National Science and Technology Major Project of China (2018ZX10715004), the Sanming Project of Medicine in Shenzhen (SZSM201811071), the High Level Project of Medicine in Longhua, Shenzhen (HLPM201907020105), the National Key Research and Development Program of China (2020YFC0840900), the Shenzhen Science and Technology Innovation Commission Basic Research Program (JCYJ20190807155409373), Special Support Plan for High-Level Talents of Guangdong Province (2019TQ05Y230), and the Fundamental Research Funds for the Central Universities (58000-31620005). None of the funding parties had any role in the study design or explanation of the data.

Data Availability

The Cancer Incidence in Five Continents data are publicly available on the web [7]. The processed data sets used in the analysis are available from the corresponding author upon reasonable request.

Authors' Contributions

GL, YZ, HZ, and YH conceived and designed the study. GL and YZ collected data, conducted the statistical analysis, prepared the figures and tables, interpreted the results, and drafted the manuscript. GL is responsible for the overall content. All authors critically reviewed and substantively revised the manuscript. All authors approved the final version of the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Statistical analysis of the assessment of the prediction performance of the NORDPRED age-period-cohort model and the corresponding findings.

[DOCX File , 36 KB - publichealth_v9i1e43651_app1.docx]

Multimedia Appendix 2

Number of new lung cancer cases and age-standardized incidence rates in lung cancer incidence in 40 countries in 2010 and 2035 among males and the corresponding percentage change between 2010 and 2035.

[DOCX File , 34 KB - publichealth_v9i1e43651_app2.docx]

Multimedia Appendix 3

Number of new lung cancer cases and age-standardized incidence rates in lung cancer incidence in 40 countries in 2010 and 2035 among females and the corresponding percentage change between 2010 and 2035.

[DOCX File , 34 KB - publichealth_v9i1e43651_app3.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Zhang Y, Luo G, Etxeberria J, Hao Y. Global patterns and trends in lung cancer incidence: a population-based study. J Thorac Oncol 2021 Jun;16(6):933-944 [FREE Full text] [doi: 10.1016/j.jtho.2021.01.1626] [Medline: 33607309]

3. GBD 2019 Respiratory Tract Cancers Collaborators. Global, regional, and national burden of respiratory tract cancers and associated risk factors from 1990 to 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Respir Med 2021 Sep;9(9):1030-1049 [FREE Full text] [doi: 10.1016/S2213-2600(21)00164-8] [Medline: 34411511]
4. Deng Y, Zhao P, Zhou L, Xiang D, Hu J, Liu Y, et al. Epidemiological trends of tracheal, bronchus, and lung cancer at the global, regional, and national levels: a population-based study. J Hematol Oncol 2020 Jul 20;13(1):98 [FREE Full text] [doi: 10.1186/s13045-020-00915-0] [Medline: 32690044]
5. Bray F, Møller B. Predicting the future burden of cancer. Nat Rev Cancer 2006 Jan;6(1):63-74. [doi: 10.1038/nrc1781] [Medline: 16372017]
6. Fidler-Benaoudia MM, Torre LA, Bray F, Ferlay J, Jemal A. Lung cancer incidence in young females vs. young men: a systematic analysis in 40 countries. Int J Cancer 2020 Aug 01;147(3):811-819 [FREE Full text] [doi: 10.1002/ijc.32809] [Medline: 32020598]
7. Bray F, Colombet M, Mery L, Piñeros M, Znaor A, Zanetti RJ, et al. Cancer Incidence in Five Continents, Vol. XI (electronic version). Lyon: International Agency for Research on Cancer; 2017.
8. International Agency for Research on Cancer, International Association of Cancer Registries. In: Bray F, Colombet M, Mery L, Piñeros M, Znaor A, Zanetti R, et al, editors. Cancer Incidence in Five Continents, Vol. XI IARC CancerBase No. 14. Geneva: World Health Organization; 2017.
9. Arnold M, Park JY, Camargo MC, Lunet N, Forman D, Soerjomataram I. Is gastric cancer becoming a rare disease? A global assessment of predicted incidence trends to 2035. Gut 2020 May;69(5):823-829 [FREE Full text] [doi: 10.1136/gutjnl-2019-320234] [Medline: 32001553]
10. Znaor A, Skakkebaek NE, Rajpert-De Meyts E, Laversanne M, Kuliš T, Gurney J, et al. Testicular cancer incidence predictions in Europe 2010-2035: a rising burden despite population ageing. Int J Cancer 2020 Aug 01;147(3):820-828 [FREE Full text] [doi: 10.1002/ijc.32810] [Medline: 31773729]
11. World Population Prospects 2022. United Nations Department of Economic and Social Affairs Population Division. URL: https://population.un.org/wpp/ [accessed 2023-02-06]
12. Doll R, Cook P. Summarizing indices for comparison of cancer incidence data. Int J Cancer 1967 May 15;2(3):269-279. [doi: 10.1002/ijc.2910020310] [Medline: 6041994]
13. Day NE. Cancer Incidence in Five Continents. Cumulative rate and cumulative risk. IARC Sci Publ 1992(120):862-864. [Medline: 1284607]
14. Møller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talbäck M, et al. Prediction of cancer incidence in the Nordic countries: empirical comparison of different approaches. Stat Med 2003 Sep 15;22(17):2751-2766. [doi: 10.1002/sim.1481] [Medline: 12939784]
15. Møller B, Fekjaer H, Hakulinen T, Tryggvadóttir L, Storm HH, Talbäck M, et al. Prediction of cancer incidence in the Nordic countries up to the year 2020. Eur J Cancer Prev 2002 Jun;11 Suppl 1:S1-96. [Medline: 12442806]
16. Yu J, Yang X, He W, Ye W. Burden of pancreatic cancer along with attributable risk factors in Europe between 1990 and 2019, and projections until 2039. Int J Cancer 2021 Sep 01;149(5):993-1001 [FREE Full text] [doi: 10.1002/ijc.33617] [Medline: 33937984]
17. Liu Z, Xu K, Jiang Y, Cai N, Fan J, Mao X, et al. Global trend of aetiology-based primary liver cancer incidence from 1990 to 2030: a modelling study. Int J Epidemiol 2021 Mar 03;50(1):128-142. [doi: 10.1093/ije/dyaa196] [Medline: 33349860]
18. Cheng F, Xiao J, Shao C, Huang F, Wang L, Ju Y, et al. Burden of thyroid cancer from 1990 to 2019 and projections of incidence and mortality until 2039 in China: findings from global burden of disease study. Front Endocrinol (Lausanne) 2021;12:738213 [FREE Full text] [doi: 10.3389/fendo.2021.738213] [Medline: 34690931]
19. Du Z, Chen W, Xia Q, Shi O, Chen Q. Trends and projections of kidney cancer incidence at the global and national levels, 1990-2030: a Bayesian age-period-cohort modeling study. Biomark Res 2020;8:16 [FREE Full text] [doi: 10.1186/s40364-020-00195-3] [Medline: 32435498]
20. Lee TC, Dean CB, Semenciw R. Short-term cancer mortality projections: a comparative study of prediction methods. Stat Med 2011 Dec 20;30(29):3387-3402. [doi: 10.1002/sim.4373] [Medline: 21965149]
21. Weir HK, Thompson TD, Soman A, Møller B, Leadbetter S. The past, present, and future of cancer incidence in the United States: 1975 through 2020. Cancer 2015 Jun 01;121(11):1827-1837 [FREE Full text] [doi: 10.1002/cncr.29258] [Medline: 25649671]
22. The R Project for Statistical Computing. R Project. URL: https://www.R-project.org/ [accessed 2023-02-06]
23. Bilal U, Fernández E, Beltran P, Navas-Acien A, Bolumar F, Franco M. Validation of a method for reconstructing historical rates of smoking prevalence. Am J Epidemiol 2014 Jan 01;179(1):15-19. [doi: 10.1093/aje/kwt224] [Medline: 24077091]
24. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. The Health Consequences of Smoking—50 Years of Progress A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2014.
25. Jha P. Avoidable global cancer deaths and total deaths from smoking. Nat Rev Cancer 2009 Sep;9(9):655-664. [doi: 10.1038/nrc2703] [Medline: 19693096]
26. Zhang Y, Ren J, Huang H, Shi J, Li N, Zhang Y, et al. International trends in lung cancer incidence from 1973 to 2007. Cancer Med 2018 Apr;7(4):1479-1489 [FREE Full text] [doi: 10.1002/cam4.1359] [Medline: 29542259]

27. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev 2010 Aug;19(8):1893-1907. [doi: 10.1158/1055-9965.EPI-10-0437] [Medline: 20647400]
28. Malvezzi M, Carioli G, Bertuccio P, Boffetta P, Levi F, La Vecchia C, et al. European cancer mortality predictions for the year 2017, with focus on lung cancer. Ann Oncol 2017 May 01;28(5):1117-1123 [FREE Full text] [doi: 10.1093/annonc/mdx033] [Medline: 28327906]
29. Levi F, Bosetti C, Fernandez E, Hill C, Lucchini F, Negri E, et al. Trends in lung cancer among young European females: the rising epidemic in France and Spain. Int J Cancer 2007 Jul 15;121(2):462-465 [FREE Full text] [doi: 10.1002/ijc.22694] [Medline: 17373662]
30. Fernández E, Schiaffino A, García M, Saltó E, Villalbí JR, Borràs JM. [Smoking in Spain, 1945-1995. A retrospective analysis based on the Spanish National Health Interview Surveys]. Med Clin (Barc) 2003 Jan 18;120(1):14-16. [doi: 10.1016/s0025-7753(03)73589-1] [Medline: 12525299]
31. Graham H. Smoking prevalence among females in the European community 1950-1990. Soc Sci Med 1996 Jul;43(2):243-254. [doi: 10.1016/0277-9536(95)00369-x] [Medline: 8844928]
32. Bosetti C, Malvezzi M, Rosso T, Bertuccio P, Gallus S, Chatenoud L, et al. Lung cancer mortality in European females: trends and predictions. Lung Cancer 2012 Dec;78(3):171-178. [doi: 10.1016/j.lungcan.2012.08.008] [Medline: 22999082]
33. Stoeldraijer L, Bonneux L, van Duin C, van Wissen L, Janssen F. Addiction 2015 Feb;110(2):336-345. [doi: 10.1111/add.12775] [Medline: 25331556]
34. Cayuela A, Rodríguez-Dominguez S, López-Campos JL, Jara-Palomares L, Otero R, Vigil E. Lung cancer mortality in Spain: estimating the future burden to the year 2028. Int J Tuberc Lung Dis 2011 Aug;15(8):1117-1121. [doi: 10.5588/ijtld.10.0577] [Medline: 21740678]
35. Jemal A, Cokkinides VE, Shafey O, Thun MJ. Lung cancer trends in young adults: an early indicator of progress in tobacco control (United States). Cancer Causes Control 2003 Aug;14(6):579-585. [doi: 10.1023/a:1024891201329] [Medline: 12948289]
36. Forey B, Lee P, Hamling J, Wald N. Estimation of sex-specific smoking statistics by standardized age groups and time periods. Research Gate. 2002 Jan. URL: https://www.researchgate.net/publication/228497499_Estimation_of_sex-specific_smoking_statistics_by_standardized_age_groups_and_time_periods [accessed 2023-02-06]
37. Jemal A, Travis WD, Tarone RE, Travis L, Devesa SS. Lung cancer rates convergence in young males and females in the United States: analysis by birth cohort and histologic type. Int J Cancer 2003 May 20;105(1):101-107 [FREE Full text] [doi: 10.1002/ijc.11020] [Medline: 12672038]
38. Report of "Tar," Nicotine, and Carbon Monoxide of the Smoke of 1294 Varieties of Domestic Cigarettes For the Year 1998. Federal Trade Commission. 2000. URL: https://www.ftc.gov/sites/default/files/documents/reports/2000-report-tar-nicotine-and-carbon-monoxide-covering-1998/1998tarnicotinereport_0.pdf [accessed 2023-02-06]
39. Defossez G, Uhry Z, Delafosse P, Dantony E, d'Almeida T, Plouvier S, French Network of Cancer Registries (FRANCIM). Cancer incidence and mortality trends in France over 1990-2018 for solid tumors: the sex gap is narrowing. BMC Cancer 2021 Jun 24;21(1):726 [FREE Full text] [doi: 10.1186/s12885-021-08261-1] [Medline: 34167516]
40. Anderson CM, Burns DM, Dodd KW, Feuer EJ. Chapter 2: birth-cohort-specific estimates of smoking behaviors for the U.S. population. Risk Anal 2012 Jul;32 Suppl 1(0 1):S14-S24 [FREE Full text] [doi: 10.1111/j.1539-6924.2011.01703.x] [Medline: 22882884]
41. Holford TR, Levy DT, McKay LA, Clarke L, Racine B, Meza R, et al. Patterns of birth cohort-specific smoking histories, 1965-2009. Am J Prev Med 2014 Feb;46(2):e31-e37 [FREE Full text] [doi: 10.1016/j.amepre.2013.10.022] [Medline: 24439359]
42. Malvezzi M, Bosetti C, Rosso T, Bertuccio P, Chatenoud L, Levi F, et al. Lung cancer mortality in European men: trends and predictions. Lung Cancer 2013 May;80(2):138-145. [doi: 10.1016/j.lungcan.2013.01.020] [Medline: 23434350]
43. Tyczynski JE, Bray F, Aareleid T, Dalmas M, Kurtinaitis J, Plesko I, et al. Lung cancer mortality patterns in selected Central, Eastern and Southern European countries. Int J Cancer 2004 Apr 20;109(4):598-610 [FREE Full text] [doi: 10.1002/ijc.20019] [Medline: 14991583]
44. Brennan P, Crispo A, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, et al. High cumulative risk of lung cancer death among smokers and nonsmokers in Central and Eastern Europe. Am J Epidemiol 2006 Dec 15;164(12):1233-1241. [doi: 10.1093/aje/kwj340] [Medline: 17032696]
45. Wang N, Mengersen K, Tong S, Kimlin M, Zhou M, Hu W. Global, regional, and national burden of lung cancer and its attributable risk factors, 1990 to 2017. Cancer 2020 Sep 15;126(18):4220-4234 [FREE Full text] [doi: 10.1002/cncr.33078] [Medline: 32648980]
46. Li Q, Hsia J, Yang G. Prevalence of smoking in China in 2010. N Engl J Med 2011 Jun 23;364(25):2469-2470. [doi: 10.1056/NEJMc1102459] [Medline: 21696322]
47. -. Air Pollution and Cancer IARC Scientific Publication No. 161. Lyon, France: IARC Scientific Publication; 2013.
48. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]

49. Tsoi KK, Hirai HW, Chan FC, Griffiths S, Sung JJ. Cancer burden with ageing population in urban regions in China: projection on cancer registry data from World Health Organization. Br Med Bull 2017 Jan 01;121(1):83-94. [doi: 10.1093/bmb/ldw050] [Medline: 27913398]
50. Transforming our world: the 2030 Agenda for Sustainable Development. United Nations Department of Economic and Social Affairs Sustainable Development. URL: https://sdgs.un.org/2030agenda [accessed 2021-10-30]
51. Cancer tomorrow. World Health Organization International Agency for Research on Cancer. URL: https://gco.iarc.fr/tomorrow/en [accessed 2022-12-05]
52. Cancer data in Australia. Australian Institute of Health and Welfare (AIHW). URL: https://www.aihw.gov.au/reports/cancer/cancer-data-in-australia/data [accessed 2021-12-04]
53. NORDCAN: cancer incidence, mortality, prevalence and survival in the Nordic Countries. World Health Organization International Agency for Research on Cancer. URL: https://iarc.spherical.horse/research-project-websites/nordcan-cancer-incidence-mortality-prevalence-and-survival-in-the-nordic-countries/ [accessed 2021-12-04]
54. Hendriks LE, Dingemans AC, De Ruysscher DK, Aarts MJ, Barberio L, Cornelissen R, et al. Lung cancer in the Netherlands. J Thorac Oncol 2021 Mar;16(3):355-365 [FREE Full text] [doi: 10.1016/j.jtho.2020.10.012] [Medline: 33641718]
55. Luo G, Zhang Y, Guo P, Wang L, Huang Y, Li K. Global patterns and trends in stomach cancer incidence: age, period and birth cohort analysis. Int J Cancer 2017 Oct 01;141(7):1333-1344 [FREE Full text] [doi: 10.1002/ijc.30835] [Medline: 28614909]
56. Luo G, Zhang Y, Guo P, Ji H, Xiao Y, Li K. Global patterns and trends in pancreatic cancer incidence: age, period, and birth cohort analysis. Pancreas 2019 Feb;48(2):199-208. [doi: 10.1097/MPA.0000000000001230] [Medline: 30589831]

Abbreviations

ASR: age-standardized incidence rate
LC: lung cancer

Edited by Y Khader; submitted 20.10.22; peer-reviewed by M Dai, S Sang; comments to author 02.12.22; revised version received 16.12.22; accepted 11.01.23; published 17.02.23.

Please cite as:
Luo G, Zhang Y, Etxeberria J, Arnold M, Cai X, Hao Y, Zou H
Projections of Lung Cancer Incidence by 2035 in 40 Countries Worldwide: Population-Based Study
JMIR Public Health Surveill 2023;9:e43651
URL: https://publichealth.jmir.org/2023/1/e43651
doi:10.2196/43651
PMID:36800235

Original Paper

Surveillance of Symptom Burden Using the Patient-Reported Outcome Version of the Common Terminology Criteria for Adverse Events in Patients With Various Types of Cancers During Chemoradiation Therapy: Real-World Study

Danbee Kang[1,2], PhD; Sooyeon Kim[1,2], BSc; Hyunsoo Kim[2], MPH; Mangyeong Lee[1,2], PhD; Sun-Young Kong[3,4,5], MD, PhD; Yoon Jung Chang[6], MD, PhD; Sung Hoon Sim[7], MD, PhD; Yeon-Joo Kim[8], MD, PhD; Juhee Cho[1,2], MA, PhD

[1]Department of Clinical Research Design and Evaluation, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Seoul, Republic of Korea

[2]Center for Clinical Epidemiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[3]Targeted Therapy Branch and Department of Laboratory Medicine, National Cancer Center, Goyang, Republic of Korea

[4]Department of Laboratory Medicine, National Cancer Center, Goyang, Republic of Korea

[5]Cancer Biomedical Science, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, Republic of Korea

[6]Division of Cancer Control & Policy, National Cancer Control Institute, National Cancer Center, Goyang, Republic of Korea

[7]Center for Breast Cancer, National Cancer Center, Goyang, Republic of Korea

[8]Department of Radiation Oncology, National Cancer Center, Goyang, Republic of Korea

Corresponding Author:
Juhee Cho, MA, PhD
Department of Clinical Research Design and Evaluation
Samsung Advanced Institute for Health Sciences & Technology
Sungkyunkwan University
115 Irwon-ro, Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 2 3410 1448
Fax: 82 2 3410 6639
Email: jcho@skku.edu

Abstract

Background: Over 90% of patients with cancer experience 1 or more symptoms caused directly by cancer or its treatment. These symptoms negatively impact on the completion of planned treatment as well as patients' health-related quality of life (HRQoL). It often results in serious complications and even life-threatening outcomes. Thus, it has been recommended that surveillance of symptom burden should be performed and managed during cancer treatment. However, differences in symptom profiles in various patients with cancer have not been fully elucidated for use in performing surveillance in the real world.

Objective: This study aims to evaluate the burden of symptoms in patients with various types of cancers during chemotherapy or radiation therapy using the PRO-CTCAE (Patient-Reported Outcome Version of the Common Terminology Criteria for Adverse Events) and its impact on quality of life.

Methods: We performed a cross-sectional study of patients undergoing outpatient-based chemotherapy, radiation therapy, or both at the National Cancer Center at Goyang or at the Samsung Medical Center in Seoul, Korea between December 2017 and January 2018. To evaluate cancer-specific symptom burden, we developed 10 subsets for using the PRO-CTCAE-Korean. To measure HRQoL, we used the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire Core 30 (EORTC QLQ-C30). Participants answered questions prior to their clinic appointments on tablets. Multivariable linear regression was used to analyze symptoms based on cancer type and to evaluate the association between the PRO-CTCAE items and the EORTC QLQ-C30 summary score.

Results: The mean age (SD) of the patients was 55.0 (11.9) years, and 39.94% (540/1352) were male. Overall, symptoms in the gastrointestinal category were the most dominant in all cancers. Fatigue (1034/1352, 76.48%), decreased appetite (884/1352, 65.38%), and numbness and tingling (778/1352, 57.54%) were the most frequently reported. Patients reported more local symptoms caused by a specific cancer. In terms of nonsite-specific symptoms, patients commonly reported concentration (587/1352, 43.42%), anxiety (647/1352, 47.86%), and general pain (605/1352, 44.75%). More than 50% of patients with colorectal (69/127, 54.3%), gynecologic (63/112, 56.3%), breast (252/411, 61.3%), and lung cancers (121/234, 51.7%) experienced decreased libido, whereas 67/112 (59.8%) patients with gynecologic cancer and lymphoma/myeloma reported pain during sexual intercourse. Patients with breast, gastric, and liver cancers were more likely to have the hand-foot syndrome. Worsening PRO-CTCAE scores were associated with poor HRQoL (eg, fatigue: coefficient –8.15; 95% CI –9.32 to –6.97), difficulty in achieving and maintaining erection (coefficient –8.07; 95% CI –14.52 to –1.61), poor concentration (coefficient –7.54; 95% CI –9.06 to –6.01), and dizziness (coefficient –7.24; 95% CI –8.92 to –5.55).

Conclusions: The frequency and severity of symptoms differed by cancer types. Higher symptom burden was associated with poor HRQoL, which suggests the importance of appropriate surveillance of PRO symptoms during cancer treatment. Considering patients had comprehensive symptoms, it is necessary to include a holistic approach in the symptom monitoring and management strategies based on comprehensive patient-reported outcome measurements.

(JMIR Public Health Surveill 2023;9:e44105) doi:10.2196/44105

KEYWORDS

surveillance; patient-reported outcome; symptoms; cancer

Introduction

Over 90% of patients with cancer experience 1 or more symptoms caused directly by cancer or its treatment [1,2]. These symptoms negatively impact on the completion of planned treatment as well as their health-related quality of life (HRQoL). It often results in serious complications and even life-threatening outcomes [3]. Thus, it has been recommended that surveillance of symptom burden should be performed and managed during cancer treatment [4].

However, barriers to symptom monitoring are medical jargon and lack of trust [5]. In addition, there have been only few reliable and valid tools for this purpose [5]. Thus, there have been growing calls for developing and implementing standardized patient-reported outcome measurements for symptom surveillance in patients with cancers for both clinical care and research purposes [6]. A major advancement in this direction is the US National Cancer Institute's PRO-CTCAE (Patient-Reported Outcome Version of the Common Terminology Criteria for Adverse Events), which comprises 124 items of 15 categories, based on 78 CTCAE toxicities considered appropriate for patient reporting [7,8]. According to a qualitative study [5], although many participants appreciated the personalized approach with symptom monitoring, they do not want to use a symptom checker that asks too many questions or that takes too long to complete. As frequent administration of the complete library of the PRO-CTCAE is impractical and burdensome, the Food and Drug Administration recommends selecting a relevant item set that can provide insights into the most important toxicities for the treatments being evaluated [9].

So far, most guidelines developed for selecting symptom measures were designed for clinical trials, and there is a lack of guidance for practitioners and for performing surveillance in the real-world clinical setting [10]. In fact, when the adverse events (AEs) were collected from the real world, new side effects related to the medication use not listed on the drug label have been reported [11]. Although a subset for practice was recently suggested based on a Delphi study with stakeholder panels [12,13], they only covered common cancers in Western countries, such as breast, lung, and colorectal cancers; further, the guidelines are limited to uncommon cancers, such as lymphoma, stomach, or liver cancers. In addition, differences in symptom profiles among various patients with cancer have not been fully elucidated for use in performing surveillance in the real world [14]. Therefore, we aimed to evaluate the burden of symptoms in patients with various types of cancers during chemotherapy or radiation therapy using the PRO-CTCAE in the real world and its impact on quality of life.

Methods

Study Population

We performed a cross-sectional study of patients undergoing outpatient-based chemotherapy, radiation therapy, or both at the National Cancer Center (NCC) at Goyang or at the Samsung Medical Center (SMC) in Seoul, Korea between December 2017 and January 2018. Eligible participants were (1) older than age 18; (2) diagnosed with cancer; (3) currently receiving chemotherapy or radiation therapy or both; and (4) those who can read, speak, and comprehend Korean. To include a more diverse sample of patients with cancer who have relatively little information about symptom burden, we aimed to recruit at least 50 patients with lymphoma, gastric, gynecologic, head and neck, and liver cancers. To simultaneously evaluate the measurement properties of all items of the PRO-CTCAE-Korean (n=124) within a single study, we aimed to recruit 1300 patients with cancer. Based on the site investigator's assessment, patients with clinically significant cognitive impairment were excluded from the study. The sampling frame was monitored to ensure that a minimum of 15% of participants had an impaired performance status (PS), defined as an Eastern Cooperative Oncology Group (ECOG) PS of 2 or higher [7].

Participants answered questions prior to their clinic appointments on tablets without assistance but could request technical assistance from the study staff if required.

Measures

For cancer-specific symptom burden surveillance, we developed 10 subsets of the PRO-CTCAE-Korean (9 for specific cancers and 1 for general purpose). The PRO-CTCAE item library has been previously translated and validated in Korean [15,16]. The PRO-CTCAE-Korean instrument has been linguistically validated for use in Korean-speaking populations [15]. In addition, the instrument has shown high construct validity (correlation [*r*] for all items >0.30 with the anchor items) and high test-retest reliability (range of intraclass correlation coefficient 0.33-0.83) [16]. To generate a subset for each type of cancer, we included common symptomatic AEs recommended by the National Cancer Institute based on their high prevalence across different cancer treatment types [17]. In addition, we included symptomatic AEs that were prevalent in specific subgroups for different cancer sites based on a literature review and recommendation by a panel of 15 medical and radiation oncologists. We performed a Delphi survey 2 times by providing the panel with the entire list of the PRO-CTCAE-Korean instrument and then asking them to choose prevalent and important symptoms to generate subsets for each cancer type. The subsets contained a minimum of 28 to a maximum of 58 symptomatic AEs.

To measure HRQoL, we used the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire Core 30 (EORTC QLQ-C30), previously validated in Korean [18], and scored the QLQ-C30 summary score according to the scoring manual [19]. Higher scores on these indicate better function. Further, demographic information was gathered by self-reporting, and clinical variables were obtained from electronic health records.

Statistical Analysis

A composite-grade scoring algorithm was used to obtain single numerical grades for AEs based on multiple PRO-CTCAE items [20]. Following the composite grading algorithm, a single composite grade was given to the PRO-CTCAE item combinations within the range of 0-3. A higher composite grade indicates the worse symptom experience. One rule of thumb for interpreting the difference in PRO scores is 10% of the instrument range [21]. As the score ranged between 0 and 3, we considered that differences of 0.4 points (ie, 10% of the score range) were clinically meaningful.

The symptom prevalence by composite grades of the PRO-CTCAE items is shown using a tree map. The size of the rectangles in the tree map indicates the proportion of patients with symptoms, with the darker colors indicating a higher prevalence of patients who reported symptoms as severe. Linear regression was used to analyze symptoms based on cancer type. Covariates adjusted were patient's age, sex, ECOG, and treatment types. Linear regression was also performed to evaluate the association between the PRO-CTCAE items and the EORTC QLQ-C30 summary score [22].

All analyses were performed using STATA version 16 (StataCorp LP) and R 3.6.1 (R Foundation for Statistical Computing). *P* values <.05 were considered statistically significant.

Ethical Considerations

Study participants provided written informed consent. We gave the participants a US $5 gift card to thank them for their participation. The Institutional Review Board of the Samsung Medical Center (SMC 2020-04-157) and the National Cancer Center (NCC2017-0249) approved this study. All the research data were encrypted.

Results

A total of 1352 patients (breast, n=411; colorectal, n=127; gastric, n=123; gynecologic, n=112; head and neck, n=56; liver, n=67; lung, n=234; lymphoma, n=112; prostate, n=57; and others, n=53) participated in this study. The mean age (SD) of the patients was 55.0 (11.9) years; 39.94% (540/1352) were male and 79.29% (1072/1352) received chemotherapy (Table 1).

Gastrointestinal cancer (purple) was the most dominant among all cancers (Figure 1). Patients commonly reported decreased appetite (884/1352, 65.38%), taste change (764/1352, 56.51%), nausea (588/1352, 43.49%), constipation (607/1352, 44.90%), and diarrhea (459/1352, 33.95%; Figure 1 and Multimedia Appendix 1). Fatigue (1034/1352, 76.48%) was the most frequently reported symptom followed by decreased appetite (884/1352, 65.38%), numbness and tingling (778/1352, 57.54%), insomnia (773/1352, 57.17%), taste change (764/1352, 56.51%), and hair loss (691/1352, 51.11%; Multimedia Appendix 1).

When we compared the symptoms across different types of cancer, patients reported more local symptoms caused by a specific cancer (Multimedia Appendix 1). After adjusting for age, sex, ECOG status, and treatment type, patients with gastric cancer were more likely to have taste changes, decreased appetite, vomiting, diarrhea, and abdominal pain compared with those with other types of cancer. Patients with head and neck cancer had more dry mouth and difficulty in swallowing compared with those with other types of cancer (Multimedia Appendix 2).

In terms of nonsite-specific symptoms, patients commonly reported concentration (587/1352, 43.42%), anxiety (647/1352, 47.86%), sadness (638/1352, 47.19%), and general pain (605/1352, 44.75%; Multimedia Appendix 1). After adjusting for confounders, patients with liver cancer experienced relatively more fatigue and anxiety than those with other types of cancer (Multimedia Appendix 2).

More than 50% of patients with colorectal, gynecologic, breast, and lung cancers and lymphoma/myeloma experienced sexual symptoms, such as decreased libido or pain during sexual intercourse (Multimedia Appendix 1). For example, 49.6% (63/127) of patients with colorectal cancer reported problems in achieving and maintaining erection and 51.7% (121/234) of patients with lung cancer reported decreased libido. Patients with breast (265/411, 64.4%), colorectal (63/127, 49.6%), gastric (68/123, 55.3%), gynecologic (65/112, 58.0%), and lung cancers (103/234, 44.0%) and lymphoma (52/112, 46.4%) more frequently reported skin dryness (Multimedia Appendix 1). Patients with breast (coefficient 0.19; 95% CI 0.09 to 0.28), gastric (coefficient 0.17; 95% CI 0.04 to 0.30), and liver (coefficient 0.21; 95% CI 0.04 to 0.38) cancers were more likely to have the hand-foot syndrome even after adjusting for potential confounders (Multimedia Appendix 2).

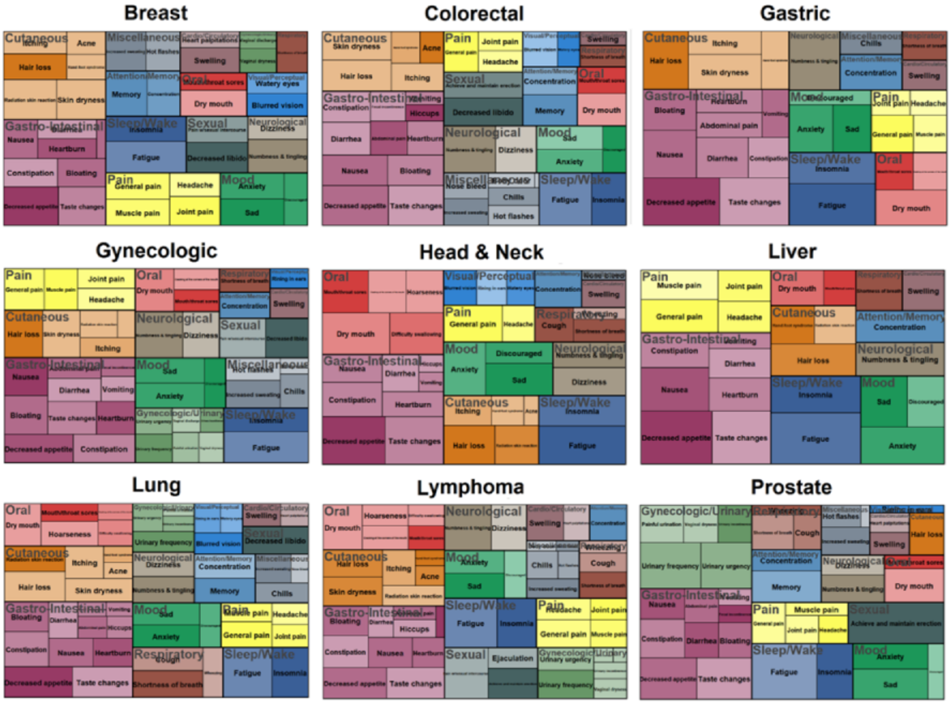
We observed a significant decrease in the mean QLQ-C30 summary scores across worsening PRO-CTCAE scores (Multimedia Appendix 3). In particular, memory, all the symptoms in the mood category, fatigue, difficulty in achieving and maintaining erection, body odor, concentration, and dizziness were associated with decreasing of QLQ-C30 summary scores (Multimedia Appendix 3).

Table 1. Characteristics of study participants (N=1352).

Characteristics	Breast (n=411)	Colorectal (n=127)	Gastric (n=123)	Gynecologic (n=112)	Head and neck (n=56)	Liver (n=67)	Lung (n=234)	Lymphoma (n=112)	Prostate (n=57)	Others (n=53)	*P* value
Age group, mean (SD)	49.1 (9.4)	56.4 (9.6)	56 (11.4)	52.8 (11)	58.9 (11.3)	59.7 (9.4)	61 (9.4)	52.9 (16.5)	67.8 (10.2)	55.9 (13.2)	<.01
Gender, n (%)											<.01
Female	409 (99.5)	59 (46.5)	40 (32.5)	112 (100)	18 (32.1)	23 (34.3)	82 (35.0)	40 (35.7)	8 (14.0)	21 (39.6)	
Male	2 (0.5)	68 (53.5)	83 (67.5)	0 (0)	38 (67.9)	44 (65.7)	152 (65.0)	72 (64.3)	49 (86.0)	32 (60.4)	
Education, n (%)											<.01
Less than middle school	36 (8.8)	27 (21.3)	28 (22.8)	22 (19.6)	17 (30.4)	16 (23.8)	79 (33.8)	22 (19.6)	19 (33.3)	9 (17.0)	
High school	170 (41.4)	51 (40.2)	51 (41.5)	51 (45.5)	16 (28.6)	27 (40.3)	102 (43.6)	46 (41.1)	19 (33.3)	19 (35.8)	
More than college	205 (49.9)	49 (38.6)	44 (35.8)	39 (34.8)	23 (41.1)	24 (35.8)	53 (22.6)	44 (39.3)	19 (33.3)	25 (47.2)	
Employment status, n (%)											<.01
Employed	110 (26.8)	40 (31.5)	44 (35.8)	11 (9.8)	15 (26.8)	15 (22.4)	58 (24.8)	37 (33.0)	8 (14.0)	15 (28.3)	
Unemployed	301 (73.2)	87 (68.5)	79 (64.2)	101 (90.2)	41 (73.2)	52 (77.6)	176 (75.2)	75 (67.0)	49 (86.0)	38 (71.7)	
Monthly family income, n (%)											<.01
<US $1990	75 (18.2)	28 (22.0)	44 (35.8)	27 (24.1)	14 (25.0)	24 (35.8)	84 (35.9)	29 (25.9)	27 (47.4)	9 (17.0)	
US $2000-US $3990	153 (37.2)	62 (48.8)	45 (36.6)	55 (49.1)	26 (46.4)	26 (38.8)	97 (41.5)	41 (36.6)	21 (36.8)	26 (49.1)	
≥US $4000	183 (44.5)	37 (29.1)	34 (27.6)	30 (26.8)	16 (28.6)	17 (25.4)	53 (22.6)	42 (37.5)	9 (15.8)	18 (34.0)	
ECOG[a] performance status at the first visit, n (%)											.01
0-1	347 (84.4)	111 (87.4)	87 (70.7)	84 (75.0)	47 (83.9)	57 (85.1)	190 (81.2)	89 (79.5)	50 (87.7)	42 (79.2)	
2-4	64 (15.6)	16 (12.6)	36 (29.3)	28 (25.0)	9 (16.1)	10 (14.9)	44 (18.8)	23 (20.5)	7 (12.3)	11 (20.8)	
Current treatment, n (%)											<.01
Chemotherapy	306 (74.5)	105 (82.7)	123 (100)	94 (83.9)	25 (44.6)	60 (89.6)	182 (77.8)	109 (97.3)	32 (56.1)	36 (67.9)	
Radiation	81 (19.7)	3 (2.4)	0 (0)	6 (5.4)	15 (26.8)	4 (6.0)	13 (5.6)	1 (0.9)	19 (33.3)	15 (28.3)	
Both	24 (5.8)	19 (15)	0 (0)	12 (10.7)	16 (28.6)	3 (4.5)	39 (16.7)	2 (1.8)	6 (10.5)	2 (3.8)	

[a] ECOG: Eastern Cooperative Oncology Group.

Figure 1. Tree map representing symptom severity and prevalence by type of cancer. Colors indicate the symptom summary score. Sizes of squares are proportion. Thus, the darker the color and greater the size of a square, the more severe and more prevalent the symptom.



Discussion

In this large real-world surveillance study of burden of symptoms among patients with various types of cancers undergoing chemoradiation therapy, frequency and severity of symptoms differed by cancer types. A higher symptom burden score was associated with poor quality of life, suggesting the importance of appropriate surveillance of PRO symptoms during cancer treatment.

Symptoms in the gastrointestinal category were most commonly reported among patients with different cancer types than those in other categories, and this finding is consistent with the results of a previous study [23]. Gastrointestinal cancer symptoms reportedly plagued many patients regardless of their cancer types [23] because chemotherapy can damage healthy cells in the lining of the digestive system from the mouth to the anus [24]. Therefore, chemotherapy can affect these areas and cause vomiting, diarrhea, constipation, and mouth sores. The most common side effects of chemotherapy associated with the gastrointestinal tract are taste changes, chemotherapy-induced nausea and vomiting, constipation, and diarrhea [25]. The prevention of chemotherapy-induced nausea and vomiting has been revolutionized over the past few years. Vomiting can be prevented in most cases [25]. However, other factors such as advanced age, decreased mobility, dietary errors, psychological alterations, and cancer-related complications may increase its occurrence [26]. In clinical practice, the implementation of nonpharmacological strategies plays an important role as an adjunct to pharmacological agents in alleviating chemotherapy-induced gastrointestinal symptoms.

In this study, we found that patients experienced more frequent and more severe symptoms caused by a specific cancer. For example, head and neck as well as lung and prostate cancers had relatively more oral, respiratory, and urinary tract symptoms, respectively, compared with other types of cancer. Our study findings are somewhat similar to the results of previous studies which reported that most patients (>80%) experienced symptoms related to their cancer site [27]. Although chemotherapy and related treatment regimens affect the whole body, the origin site of cancer seems to have a greater impact on patient's symptom burden. As the different types of cancer had different symptoms even when patients received the same treatment, it is necessary to provide specific care to manage the complexity of the symptoms by the cancer type [28]. Therefore, it is important to develop and use a specific subset of the assessment when evaluating chemotherapy- or radiation-induced symptoms for patients with different types of cancers or undergoing treatment options.

Using the PRO-CTCAE will help detect symptoms that were often underreported by health professionals [29], as we have shown in this study, where many patients reported problems with sexual dysfunction, which was not often assessed in previous studies [17]. In our study, more than 50% of patients with colorectal (69/127, 54.3%), gynecologic (63/112, 56.3%), breast (252/411, 61.3%), and lung cancers (121/234, 51.7%) experienced decreased libido, whereas 67/112 (59.8%) patients with gynecologic cancer and lymphoma/myeloma reported pain during sexual intercourse. However, these sexual dysfunctions are often overlooked in evaluating patients with cancers [30]. The causes of sexual dysfunction are psychological distress and endocrine dysfunction caused by the cancer itself or side effects of anticancer treatments such as surgery, radiotherapy, chemotherapy, and hormonal therapy. For example, among patients with colorectal cancer, the rates of sexual dysfunction can be even higher given the physiological changes that can result from surgery, chemotherapy, and radiation therapy [31]. Similarly, surgery for lung cancer may adversely affect the psychogenic status and sexual function in patients with lung cancer due to its invasive nature [32]. As cancer treatments and the emotional suffering from a cancer diagnosis can affect all aspects of sexuality [33], sexuality should be a routine assessment, and communication should begin early on [34].

The PRO-CTCAE includes 6 symptoms of cutaneous toxicity, which were frequently reported by our patients with lymphoma/myeloma, colorectal, gynecologic, gastric, breast, and lung cancers. Although cutaneous toxicities are the common side effects reported by patients with cancer receiving chemotherapy [35], they are often considered minor complaints compared with other side effects such as nausea or vomiting [36]. However, cutaneous toxicity is strongly associated with psychological well-being and HRQoL. Most patients with cancer reported that the impact of chemotherapy on skin irritation and dry skin was worse than they had anticipated [37]. More than two-thirds of patients that developed cutaneous side effects due to chemotherapy were significantly distressed by their altered appearance [38]. It disturbs their daily activities and personal relationships and negatively impacts their HRQoL [39]. Considering the burden of cutaneous symptoms on patients' daily activities and HRQoL [40], a multidisciplinary cancer care team, including dermatologists, oncologists, and nurses, should perform regular surveillance and appropriate interventions. Furthermore, cutaneous toxicity needs to be monitored more actively in patients with targeted agents, as dermatological toxicities are among the most common complications of targeted agents [41].

Regarding the impact of symptoms burden on HRQoL, most symptoms were associated with lower HRQoL. In particular, memory, mood, fatigue, erection, body odor, concentration, and dizziness were associated with clinically noticeable declines in HRQoL. As these symptoms were associated with daily life, they might have a greater impact on HRQoL due to their burden. In particular, fatigue was reported as one of the most common side effects of cancer that was associated with poor HRQoL, which is similar to the finding reported in a previous study [42]. Fatigue negatively impacts work, social relationships, mood, and daily activities and causes significant impairment in the overall HRQoL during and after treatment [42]. Regarding mood, all patients with cancer had a certain level of anxiety, which could be related to the fear of treatment or fear of cancer recurrence. Considering that anxiety as a symptom is dynamic and can change over time in response to cancer-related events [43], surveillance and treatment planning should incorporate factors contributing to anxiety and patient preference for psychiatric care [43]. Pain is also a frequent, nonsite-specific symptom in all patients. A more active approach may be necessary to manage pain, as it could affect many other symptoms. Patients experienced loss of confidence and restriction in physical and social activities due to chronic dizziness [44]. The burden of shortness of breath might be further compounded by fatigue, anxiety, and depression, resulting in functional limitations and compromised HRQoL [45]. Patients with cancer reported more distress when they experienced a symptom they did not anticipate [37]. Therefore, it is important to provide pretreatment information, counseling, and management resources concerning possible treatment-related side effects in these patients.

There are several limitations to our study. First, the reporting of the symptom scores was voluntary. Thus, the receipt of symptom screening may itself bias the estimates of symptom burden because the routine collection of PROs is associated with improved clinical outcomes and increased patient satisfaction. In addition, patient factors, including male sex and advanced age, were associated with lower rates of PRO-CTCAE reporting, which could reflect differential rates of participation among patient subgroups. Second, as symptom assessments are only recorded at outpatient visits, we did not capture the symptoms of patients who are admitted to the hospital or hospice, or who are otherwise too unwell to visit clinics and may probably be the most symptomatic. In addition, because of the heterogeneity in our cohort, we did not assess the influence of treatment modalities, which will differ substantially among stages, on symptom burden. Finally, although we compared symptom burden based on disease site groups, we did not describe the symptoms of unique cancers, which may mask heterogeneity in the symptom profiles of distinct cancers within larger categories, such as lymphoma/myeloma, colorectal, head and neck/esophageal, prostate/bladder, and gynecologic cancers. Despite these limitations, this study provides guidance on symptoms that should be asked about to patients in the real-world clinical setting. This study also illustrated the feasibility of linking routinely collected PROs to large population-based health care databases.

In conclusion, the frequency and severity of symptoms differed according to the type of cancer, and the symptoms were associated with poor HRQoL. Recently, there has been an emphasis on the appropriate assessment of PRO symptoms during cancer treatment [5,46]. Considering that patients would have various symptoms at different periods, and that symptoms could affect not only disease but also HRQoL, it is necessary to take a holistic approach when implementing a symptom monitoring and management strategy. The results of this study would help physicians to improve their understanding of the variations in cancer treatment–related symptoms and to use proper symptoms list by types of cancer in routine care for developing management plans and guidelines.

Acknowledgments

This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1I1A2074210), the National R&D Program for Cancer Control, Ministry of Health and Welfare, Republic of Korea (grant number 1520240), a grant (number 19182MFDS426) from the Ministry of Food and Drug Safety in 2019, and Future Medicine 20 × 30 Project of the Samsung Medical Center (SMX1210831).

Data Sharing

The data sets generated or analyzed during this study are available from the corresponding author (JC) on reasonable request.

Authors' Contributions

DK and JC were involved in the study design. ML, SYK, YJC, SS, and YJK were involved in data collection. DK, SK, and HK were involved in data acquisition and management. DK, SK, and HK were involved in data analysis. All authors were involved in results interpretation and manuscript writing. JC made the decision to submit this manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Symptom frequency by composite grades of PRO-CTCAE items. PRO-CTCAE: Patient-Reported Outcome Version of the Common Terminology Criteria for Adverse Events.

[DOCX File , 33 KB - publichealth_v9i1e44105_app1.docx]

Multimedia Appendix 2

Difference (95% CI) of cancer-specific symptoms by types of cancer compared with other cancers.

[DOCX File , 32 KB - publichealth_v9i1e44105_app2.docx]

Multimedia Appendix 3

Association between symptoms and health-related quality of life in different types of cancer.

[DOCX File , 44 KB - publichealth_v9i1e44105_app3.docx]

References

1. Shi Q, Smith TG, Michonski JD, Stein KD, Kaw C, Cleeland CS. Symptom burden in cancer survivors 1 year after diagnosis: a report from the American Cancer Society's Studies of Cancer Survivors. Cancer 2011 Jun 15;117(12):2779-2790 [FREE Full text] [doi: 10.1002/cncr.26146] [Medline: 21495026]
2. Deshields TL, Potter P, Olsen S, Liu J. The persistence of symptom burden: symptom experience and quality of life of cancer patients across one year. Support Care Cancer 2013 Dec 3;22(4):1089-1096. [doi: 10.1007/s00520-013-2049-3]
3. National Chemotherapy Advisory Group. Chemotherapy Services in England: Ensuring Quality and Safety. England, UK: NHS National Cancer Action Team; 2009.
4. Trotti A, Colevas AD, Setser A, Basch E. Patient-reported outcomes and the evolution of adverse event reporting in oncology. JCO 2007 Nov 10;25(32):5121-5127. [doi: 10.1200/jco.2007.12.4784]
5. Aboueid S, Meyer S, Wallace JR, Mahajan S, Chaurasia A. Young adults' perspectives on the use of symptom checkers for self-triage and self-diagnosis: qualitative study. JMIR Public Health Surveill 2021 Jan 06;7(1):e22637 [FREE Full text] [doi: 10.2196/22637] [Medline: 33404515]
6. Basch E, Wood WA, Schrag D, Sima CS, Shaw M, Rogak LJ, et al. Feasibility and clinical impact of sharing patient-reported symptom toxicities and performance status with clinical investigators during a phase 2 cancer treatment trial. Clin Trials 2016 Jun;13(3):331-337 [FREE Full text] [doi: 10.1177/1740774515615540] [Medline: 26542025]
7. Kim D, Chang H. Key functional characteristics in designing and operating health information websites for user satisfaction: an application of the extended technology acceptance model. Int J Med Inform 2007 Dec;76(11-12):790-800. [doi: 10.1016/j.ijmedinf.2006.09.001] [Medline: 17049917]
8. Basch E, Reeve BB, Mitchell SA, Clauser SB, Minasian LM, Dueck AC, et al. Development of the National Cancer Institute's patient-reported outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). J Natl Cancer Inst 2014 Sep;106(9) [FREE Full text] [doi: 10.1093/jnci/dju244] [Medline: 25265940]
9. Kluetz PG, Slagle A, Papadopoulos EJ, Johnson LL, Donoghue M, Kwitkowski VE, et al. Focusing on core patient-reported outcomes in cancer clinical trials: symptomatic adverse events, physical function, and disease-related symptoms. Clin Cancer Res 2016 Apr 01;22(7):1553-1558 [FREE Full text] [doi: 10.1158/1078-0432.CCR-15-2035] [Medline: 26758559]
10. Montgomery N, Howell D, Ismail Z, Bartlett SJ, Brundage M, Bryant-Lukosius D, et al. Selecting, implementing and evaluating patient-reported outcome measures for routine clinical use in cancer: the Cancer Care Ontario approach. J Patient Rep Outcomes 2020 Nov 26;4(1). [doi: 10.1186/s41687-020-00270-1]
11. Park S, Choi SH, Song Y, Kwon J. Comparison of online patient reviews and national pharmacovigilance data for tramadol-related adverse events: comparative observational study. JMIR Public Health Surveill 2022 Jan 04;8(1):e33311 [FREE Full text] [doi: 10.2196/33311] [Medline: 34982723]

12. Geerse O, Wynia K, Kruijer M, Schotsman M, Hiltermann T, Berendsen A. Health-related problems in adult cancer survivors: development and validation of the Cancer Survivor Core Set. Support Care Cancer 2016 Oct 14;25(2):567-574. [doi: 10.1007/s00520-016-3439-0]
13. Ramsey I, Corsini N, Hutchinson AD, Marker J, Eckert M. A core set of patient-reported outcomes for population-based cancer survivorship research: a consensus study. J Cancer Surviv 2020 Aug 31;15(2):201-212. [doi: 10.1007/s11764-020-00924-5]
14. Zhang C, Zhang C, Wang Q, Li Z, Lin J, Wang H. Differences in stage of cancer at diagnosis, treatment, and survival by race and ethnicity among leading cancer types. JAMA Netw Open 2020 Apr 08;3(4):e202950. [doi: 10.1001/jamanetworkopen.2020.2950]
15. Cho J, Yoon J, Kim Y, Oh D, Kim SJ, Ahn J, et al. Linguistic Validation of the US National Cancer Institute’s Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events in Korean. JGO 2019 Dec(5):1-10. [doi: 10.1200/jgo.18.00193]
16. Yoon J, Sim SH, Kang D, Han G, Kim Y, Ahn J, et al. Reliability and Validity of the Korean Language Version of the U.S. National Cancer Institute's Patient-Reported Outcomes Common Terminology Criteria for Adverse Events. J Pain Symptom Manage 2020 May;59(5):1082-1088.e6. [doi: 10.1016/j.jpainsymman.2020.01.015] [Medline: 32044423]
17. Reeve BB, Mitchell SA, Dueck AC, Basch E, Cella D, Reilly CM, et al. Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials. J Natl Cancer Inst 2014 Jul 08;106(7):dju129-dju129 [FREE Full text] [doi: 10.1093/jnci/dju129] [Medline: 25006191]
18. Yun Y, Park Y, Lee E, Bang S, Heo D, Park S, et al. Validation of the Korean version of the EORTC QLQ-C30. Qual Life Res 2004 May;13(4):863-868. [doi: 10.1023/b:qure.0000021692.81214.70]
19. Fayers P, Aaronson N, Bjordal K, Groenvold M, Curran D, Bottomley A. The EORTC QLQ-C30 Scoring Manual, 3rd Edition. Brussels, Belgium: European Organization for Research and Treatment of Cancer; 2001.
20. Basch E, Becker C, Rogak LJ, Schrag D, Reeve BB, Spears P, et al. Composite grading algorithm for the National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Clin Trials 2021 Feb 01;18(1):104-114 [FREE Full text] [doi: 10.1177/1740774520975120] [Medline: 33258687]
21. Barrett B, Brown D, Mundt M, Brown R. Sufficiently important difference: expanding the framework of clinical significance. Med Decis Making 2005 Jun 10;25(3):250-261. [doi: 10.1177/0272989x05276863]
22. Cocks K, King MT, Velikova G, Fayers PM, Brown JM. Quality, interpretation and presentation of European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 data in randomised controlled trials. Eur J Cancer 2008 Sep;44(13):1793-1798. [doi: 10.1016/j.ejca.2008.05.008] [Medline: 18599286]
23. Abernethy A, Wheeler J, Zafar S. Management of gastrointestinal symptoms in advanced cancer patients: the rapid learning cancer clinic model. Curr Opin Support Palliat Care 2010 Mar;4(1):36-45 [FREE Full text] [doi: 10.1097/SPC.0b013e32833575fd] [Medline: 19952928]
24. Sougiannis AT, VanderVeen BN, Davis JM, Fan D, Murphy EA. Understanding chemotherapy-induced intestinal mucositis and strategies to improve gut resilience. Am J Physiol Gastrointest Liver Physiol 2021 May 01;320(5):G712-G719 [FREE Full text] [doi: 10.1152/ajpgi.00380.2020] [Medline: 33471628]
25. Jordan K, Jahn F, Aapro M. Recent developments in the prevention of chemotherapy-induced nausea and vomiting (CINV): a comprehensive review. Ann Oncol 2015 Jun;26(6):1081-1090 [FREE Full text] [doi: 10.1093/annonc/mdv138] [Medline: 25755107]
26. Jordan K, Gralla R, Jahn F, Molassiotis A. International antiemetic guidelines on chemotherapy induced nausea and vomiting (CINV): content and implementation in daily routine practice. Eur J Pharmacol 2014 Jan 05;722:197-202. [doi: 10.1016/j.ejphar.2013.09.073] [Medline: 24157984]
27. Koo MM, Swann R, McPhail S, Abel GA, Elliss-Brookes L, Rubin GP, et al. Presenting symptoms of cancer and stage at diagnosis: evidence from a cross-sectional, population-based study. Lancet Oncol 2020 Jan;21(1):73-79 [FREE Full text] [doi: 10.1016/S1470-2045(19)30595-9] [Medline: 31704137]
28. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discovery 2022 Dec 1;12(1):31-46. [doi: 10.1158/2159-8290.cd-21-1059]
29. Dueck AC, Scher HI, Bennett AV, Mazza GL, Thanarajasingam G, Schwab G, et al. Assessment of adverse events from the patient perspective in a phase 3 metastatic castration-resistant prostate cancer clinical trial. JAMA Oncol 2020 Feb 01;6(2):e193332 [FREE Full text] [doi: 10.1001/jamaoncol.2019.3332] [Medline: 31556911]
30. Furlow B. Sexual dysfunction in patients with lung disease. Lancet Respir Med 2014 Jun;2(6):439. [doi: 10.1016/s2213-2600(14)70081-5] [Medline: 25028763]
31. Averyt JC, Nishimoto PW. Addressing sexual dysfunction in colorectal cancer survivorship care. J Gastrointest Oncol 2014 Oct;5(5):388-394 [FREE Full text] [doi: 10.3978/j.issn.2078-6891.2014.059] [Medline: 25276411]
32. Hung M, Chen Y, Huang T, Ho D, Lee C, Chen P, et al. Erectile dysfunction after surgical treatment of lung cancer: real-world evidence. CLEP 2020 Sep;Volume 12:977-987. [doi: 10.2147/clep.s264439]
33. Williams AC, Reckamp K, Freeman B, Sidhu R, Grant M. Sexuality, lung cancer, and the older adult: an unlikely trio? J Adv Pract Oncol 2013 Sep 1;4(5):331-340 [FREE Full text] [doi: 10.6004/jadpro.2013.4.5.5] [Medline: 25032012]

34. Yi J, Syrjala K. Sexuality after hematopoietic stem cell transplantation. Cancer J 2009;15(1):57-64 [FREE Full text] [doi: 10.1097/PPO.0b013e318198c758] [Medline: 19197175]
35. Heidary N, Naik H, Burgin S. Chemotherapeutic agents and the skin: an update. J Am Acad Dermatol 2008 Apr;58(4):545-570. [doi: 10.1016/j.jaad.2008.01.001] [Medline: 18342708]
36. Ra H, Shin S, Kim J, Lim H, Cho B, Roh M. The impact of dermatological toxicities of anti-cancer therapy on the dermatological quality of life of cancer patients. J Eur Acad Dermatol Venereol 2013 Jan 13;27(1):e53-e59. [doi: 10.1111/j.1468-3083.2012.04466.x] [Medline: 22329482]
37. Gandhi M, Oishi K, Zubal B, Lacouture ME. Unanticipated toxicities from anticancer therapies: survivors' perspectives. Support Care Cancer 2010 Nov 3;18(11):1461-1468. [doi: 10.1007/s00520-009-0769-1] [Medline: 19956983]
38. Zabora J, BrintzenhofeSzoc K, Curbow B, Hooker C, Piantadosi S. The prevalence of psychological distress by cancer site. Psychooncology 2001 Jan;10(1):19-28. [doi: 10.1002/1099-1611(200101/02)10:1<19::aid-pon501>3.0.co;2-6] [Medline: 11180574]
39. Haley AC, Calahan C, Gandhi M, West DP, Rademaker A, Lacouture ME. Skin care management in cancer patients: an evaluation of quality of life and tolerability. Support Care Cancer 2011 Apr 25;19(4):545-554 [FREE Full text] [doi: 10.1007/s00520-010-0851-8] [Medline: 20336328]
40. Pérez-Soler R. Can rash associated with HER1/EGFR inhibition be used as a marker of treatment outcome? Oncology (Williston Park) 2003 Nov;17(11 Suppl 12):23-28 [FREE Full text] [Medline: 14682120]
41. Huynh Dagher S, Blom A, Chabanol H, Funck-Brentano E. Cutaneous toxicities from targeted therapies used in oncology: literature review of clinical presentation and management. Int J Womens Dermatol 2021 Dec 28;7(5Part A):615-624 [FREE Full text] [doi: 10.1016/j.ijwd.2021.09.009] [Medline: 35024416]
42. Fabi A, Bhargava R, Fatigoni S, Guglielmo M, Horneber M, Roila F, ESMO Guidelines Committee. Cancer-related fatigue: ESMO Clinical Practice Guidelines for diagnosis and treatment. Ann Oncol 2020 Jun;31(6):713-723 [FREE Full text] [doi: 10.1016/j.annonc.2020.02.016] [Medline: 32173483]
43. Traeger L, Greer JA, Fernandez-Robles C, Temel JS, Pirl WF. Evidence-based treatment of anxiety in patients with cancer. JCO 2012 Apr 10;30(11):1197-1205. [doi: 10.1200/jco.2011.39.5632]
44. Lindell E, Kollén L, Johansson M, Karlsson T, Rydén L, Fässberg MM, et al. Dizziness and health-related quality of life among older adults in an urban population: a cross-sectional study. Health Qual Life Outcomes 2021 Oct 02;19(1):231 [FREE Full text] [doi: 10.1186/s12955-021-01864-z] [Medline: 34600557]
45. Reddy SK, Parsons HA, Elsayem A, Palmer JL, Bruera E. Characteristics and correlates of dyspnea in patients with advanced cancer. J Palliat Med 2009 Jan;12(1):29-36. [doi: 10.1089/jpm.2008.0158] [Medline: 19284260]
46. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]

Abbreviations

AE: adverse event
ECOG: Eastern Cooperative Oncology Group
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire Core 30
HRQoL: health-related quality of life
NCC: National Cancer Center
PRO-CTCAE: Patient-Reported Outcome Version of the Common Terminology Criteria for Adverse Events
PS: performance status
SMC: Samsung Medical Center

Edited by A Mavragani, T Sanchez; submitted 07.11.22; peer-reviewed by M Schoen, J Drott; comments to author 06.01.23; revised version received 10.01.23; accepted 31.01.23; published 08.03.23.

Please cite as:
Kang D, Kim S, Kim H, Lee M, Kong SY, Chang YJ, Sim SH, Kim YJ, Cho J
Surveillance of Symptom Burden Using the Patient-Reported Outcome Version of the Common Terminology Criteria for Adverse Events in Patients With Various Types of Cancers During Chemoradiation Therapy: Real-World Study
JMIR Public Health Surveill 2023;9:e44105
URL: https://publichealth.jmir.org/2023/1/e44105
doi:10.2196/44105
PMID:36884274

Original Paper

Changes in the Demographic Distribution of Chicago Gun-Homicide Decedents From 2015-2021: Violent Death Surveillance Cross-sectional Study

Maryann Mason[1], PhD; Rushmin Khazanchi[1], BSc; Audrey Brewer[2], MD, MPH; Karen Sheehan[2], MD, MPH; Yingxuan Liu[1], BA, MA; Lori Post[1], PhD

[1]Department of Emergency Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States
[2]Department of Pediatrics, Lurie Children's Hospital of Chicago, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:
Maryann Mason, PhD
Department of Emergency Medicine
Feinberg School of Medicine
Northwestern University
Chicago, IL
United States
Phone: 1 312 503 5142
Email: maryann-mason@northwestern.edu

Abstract

Background: Homicide is one of the 5 leading causes of death in the United States for persons aged 1 to 44 years. In 2019, 75% of US homicides were by gun. Chicago has a gun-homicide rate 4 times the national average, and 90% of all homicides are by gun. The public health approach to violence prevention calls for a 4-step process, beginning with defining and monitoring the problem. Insight into the characteristics of gun-homicide decedents can help frame next steps, including identifying risk and protective factors, developing prevention and intervention strategies, and scaling effective responses. Although much is known about gun homicide because it is a long-standing, entrenched public health problem, it is useful to monitor trends to update ongoing prevention efforts.

Objective: This study aimed to use public health surveillance data and methods to describe changes in the race/ethnicity, sex, and age of Chicago gun-homicide decedents from 2015-2021, in the context of year-to-year variation and an overall increase in the city’s gun-homicide rate.

Methods: We calculated the distribution of gun-related homicide deaths by 6 race/ethnicity and sex groups (non-Hispanic Black female, non-Hispanic White female, Hispanic female, non-Hispanic Black male, non-Hispanic White male, and Hispanic male), age in years, and age by age group. We used counts, percentages, and rates per 100,000 persons to describe the distribution of deaths among these demographic groups. Comparisons of means and column proportions with tests of significance set at $P \leq .05$ were used to describe changes in the distribution of gun-homicide decedents over time by race-ethnicity-sex and age groups. The comparison of mean age by race-ethnicity-sex group is done using 1-way ANOVA with significance set at $P \leq .05$.

Results: The distribution of gun-homicide decedents in Chicago by race/ethnicity and sex groups had been relatively stable from 2015 to 2021 with 2 notable exceptions: a more than doubling of the proportion of gun-homicide decedents who were non-Hispanic Black female (3.6% in 2015 to 8.2% in 2021) and an increase of 3.27 years in the mean age of gun-homicide decedents. The increase in mean age coincided with a decrease in the proportion of non-Hispanic Black male gun-homicide decedents between the ages of 15-19 and 20-24 years and, conversely, an increase in the proportion of non-Hispanic Black male gun-homicide decedents aged 25-34 years.

Conclusions: The annual gun-homicide rate in Chicago had been increasing since 2015 with year-to-year variation. Continued monitoring of trends in the demographic makeup of gun-homicide decedents is necessary to provide the most relevant and timely information to help shape violence prevention efforts. We detected several changes that suggest a need for increased outreach and engagement marketed toward non-Hispanic Black female and non-Hispanic Black male individuals between the ages of 25-34 years.

(JMIR Public Health Surveill 2023;9:e43723) doi:10.2196/43723

KEYWORDS

gun-homicide surveillance; gun-homicide decedents; demographics; age, gun violence; firearm

Introduction

The United States is an outlier among the world's 28 most populous high-income countries, with a homicide by gun rate 11.4 times that of all these countries combined [1]. Within the United States, gun homicides constitute 75% of all homicides [2] and is one of the 5 leading causes of death for persons aged 1 to 44 years [3]. The US gun homicide rate increased by 34% from 2019 to 2020 [4]. In 2020, Chicago, the third largest US city, had a gun homicide rate of 22.65 per 100,000 persons, nearly 4 times the overall US rate of 5.9 per 100,000 persons [2].

The public health approach to violence prevention calls for a 4-step process, beginning with defining and monitoring the problem [5]. Insight into the characteristics of gun-homicide decedents can help frame next steps, including identifying risk and protective factors, developing prevention and intervention strategies, and scaling effective responses. Although much is known about gun homicide because it is a long-standing, entrenched public health problem, it is useful to monitor trends to update ongoing prevention efforts.

In Chicago, as is the case nationally, the risk for gun-related homicide is not evenly distributed. Adolescents and young adults, males, and Black persons have elevated gun-related homicide rates compared to other population segments [6-8]. Historically, persons between the ages of 15 and 24 years have had the highest gun-related homicide rates [7].

Chicago's annual rate of gun-related homicides has varied year to year [5,8], but the overall trend has increased since 2015 [9]. In 2021, the number of gun-related homicides in Chicago hit a 40-year high, surpassing the most recent high in 2020 [9]. We wondered if the distribution of gun-homicide decedent demographic characteristics in Chicago had changed with year-to-year variation in Chicago's gun-homicide rates and in the context of the overall increase in gun-homicide rates.

Understanding changes in decedent characteristics can help advance knowledge of drivers of increases in gun-homicide rates and inform prevention and intervention efforts, including the allocation of resources to prevent future gun homicides. Since prevention settings and content are usually based on intent to reach the populations most at risk, knowledge of changes in the demographics of at-risk populations may help shape more efficacious responses. Consideration of demographic factors such as age and sex in relationship to race/ethnicity is paramount because, as Feld and Bauldry [10] note, gun homicides are unevenly distributed by race/ethnicity and overall summary conclusions may obfuscate important variations.

Our objective was to use public health surveillance data and methods to describe changes in the race/ethnicity, sex, and age of Chicago gun-homicide decedents from 2015-2021, in the context of year-to-year variation and an overall increase in the city's gun-homicide rate.

Methods

Ethical Considerations

This study was determined to be exempt from institutional review board review by the Northwestern University institutional review board because research involving deceased persons is not considered human subjects research, and the data are publicly available and deidentified. The study data source was deidentified retrospective case-level surveillance data from the Cook County Medical Examiner Case Archive [9]. These data were collected after death was declared; therefore, no informed consent was possible or required. To protect against possible identification of individuals in the data set, we suppress counts for cells with fewer than 5 in table presentations.

Data Sources

Gun-homicide decedent data were from Cook County Medical Examiner Case Archive data [9]. The archive includes data on decedent characteristics and fatal incident date and location. Population data came from American Community Survey [11].

Inclusion Criteria

This study included all homicides in which the fatal injury occurred in the city of Chicago with the death occurring between January 1, 2015, and December 31, 2021, where the manner of death was ruled "homicide" and the "gun-related" variable was endorsed by the Cook County Coroner's Office. The start date is January 1, 2015, because that is the first year that electronic surveillance of these deaths had complete case ascertainment.

For our analyses, we used a continuous variable for age in years, as well as a categorical variable where we collapsed the decedent age in years into the following age groups: <15, 15-19, 20-24, 25-34, 35-44, 45-54, 55-64, and 65+ years. Age groups were based on standard age groups generated from the US Census. We also created a race-ethnicity-sex group variable by assigning decedents into 1 of the following 6 race/ethnicity and sex groups: non-Hispanic Black female, non-Hispanic White female, Hispanic female, non-Hispanic Black male, non-Hispanic White male, and Hispanic male. We did not include data on Asian, Native Hawaiian or Other Pacific Islander, or American Indian or Alaska Native racial groups due to the low numbers of decedents, which makes comparisons of distributions unreliable.

Statistical Analysis

We calculate the overall Chicago annual gun-homicide rate per 100,000 persons as the annual number of gun-homicides divided by the 2020 Chicago population. For analyses, we assigned decedents to 1 of 6 race-ethnicity-sex groups (non-Hispanic Black female, non-Hispanic White female, Hispanic female, non-Hispanic Black male, non-Hispanic White male, and Hispanic male). We graphed the mean age of gun homicide decedents by the year of death and used Pearson *r* to evaluate the association between mean age and the year of death. We described the proportion of decedents by race-ethnicity-sex and age groups over time and tested for statistically significant change over time using comparisons of column proportions with significance set at $P \leq .05$. The comparison of mean age by race-ethnicity-sex group was done using 1-way ANOVA with significance set at $P \leq .05$. We also examined changes in annual gun-homicide rates per 100,000 persons by age group using the annual number of gun-homicide deaths by age group divided by the annual Chicago population for that age. Analyses were done in IBM Statistics 27. We used graphing procedures in Microsoft Excel 2016, IBM Statistics 27, and Tableau (Salesforce Inc) for data visualizations.

Results

Between 2015 and 2021, there were 3723 gun-homicides among Chicago residents recorded by the Cook County Medical Examiner's Office. Decedents were 92% (n=3425/3723) male, 82% (n=3054/3723) Black, and 15.2% (n=566/3723) Hispanic with a mean age of 28.99 (SD 11.18) years. The overall Chicago gun homicide rate increased 58% from 14.46 to 22.92 per 100,000 persons during this time frame (see Figure 1).

Figure 1. Chicago gun-homicide rate per 100,000 persons: 2015-2021.

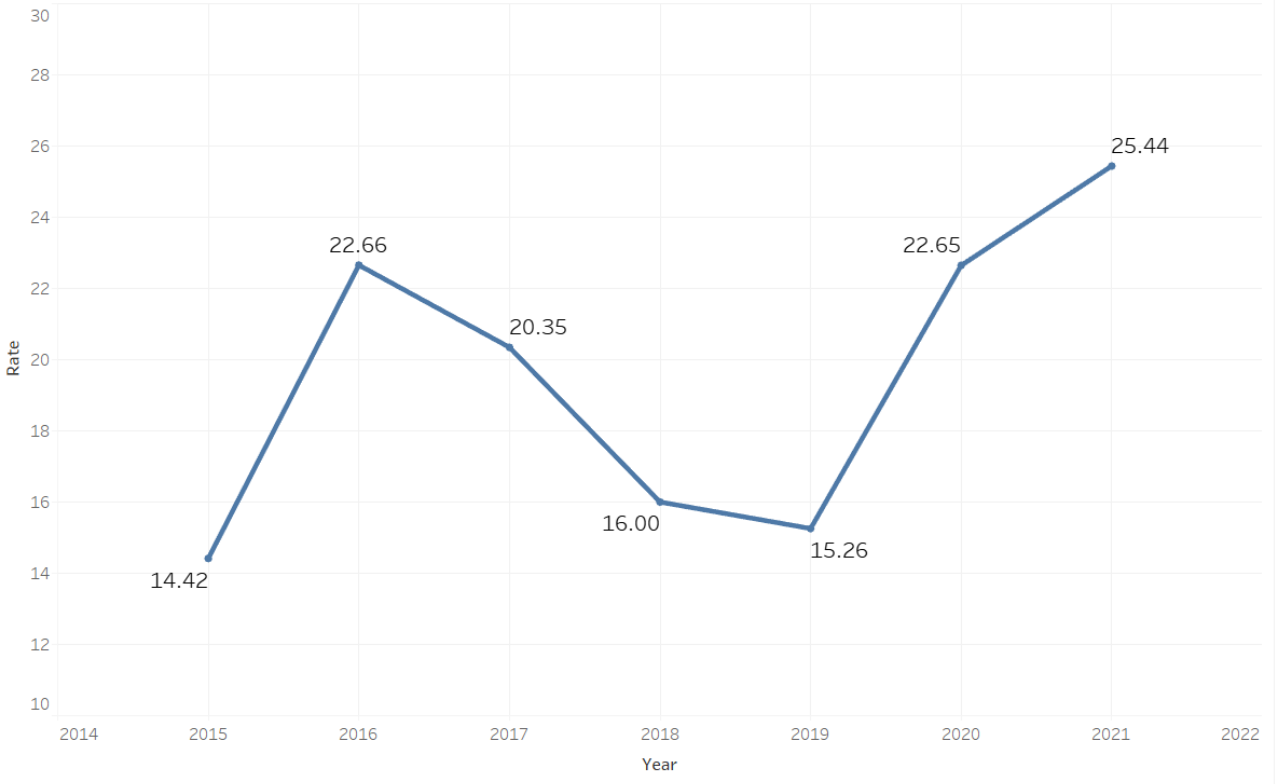


Table 1 reports the annual percentage of gun-homicide deaths by 6 race-ethnicity-sex group combinations. Each row represents the proportion of gun-homicide decedents by race-ethnicity-sex group. Across all years, non-Hispanic Black male individuals consistently made up the largest proportion of gun-homicide decedents in Chicago (≥72.5%). The proportion of gun-homicide deaths among non-Hispanic Black female individuals had more than doubled from 2015 (n=14/391, 3.6%) to 2021 (n=56/685, 8.2%); however, this increase was not statistically significant ($P > .05$).

From 2015 to 2021, the mean age of gun-related homicide decedents increased by 3.27 years (Figure 2). The largest increase (+1.94 years) occurred between 2017 and 2018.

The mean age among Chicago gun-homicide decedents differed by race-ethnicity-sex group ($F_5 = 9.39$; $P < .001$; Table 2). Non-Hispanic White male and female groups had higher mean ages compared to other groups.

Table 1. Proportion of Chicago gun-homicide deaths by race-ethnicity-sex groups from 2015-2021.

Race-ethnicity-sex group	Year of death and the proportion of gun-homicide death[a]						
	2015, n (%)	2016, n (%)	2017, n (%)	2018, n (%)	2019, n (%)	2020, n (%)	2021, n (%)
Non-Hispanic Black female	14 (3.6)	27 (4.4)	33 (6)	30 (6.9)	33 (8)	45 (7.3)	56 (8.2)
Non-Hispanic White female	0 (0)	<5 (0.5)	<5 (0.4)	<5 (0.7)	<5 (0.7)	<5 (0.3)	<5 (0.6)
Hispanic female	<5 (0.8)	<5 (1)	6 (1.1)	6 (1.4)	6 (1.5)	<5 (0.6)	10 (1.5)
Non-Hispanic Black male	313 (80.1)	467 (75.9)	409 (74.4)	336 (77.6)	308 (75.1)	448 (72.5)	521 (76.1)
Non-Hispanic White male	8 (2)	14 (2.3)	11 (2)	10 (2.3)	9 (2.2)	16 (2.6)	11 (1.6)
Hispanic male	53 (13.6)	98 (15.9)	89 (16.2)	48 (11.1)	51 (12.4)	103 (16.7)	83 (12.1)

[a] Denominators are not reported due to the suppression of cell counts to protect against possible identification of individuals.

Figure 2. Chicago gun-homicide decedent mean age by year of death: 2015-2021.

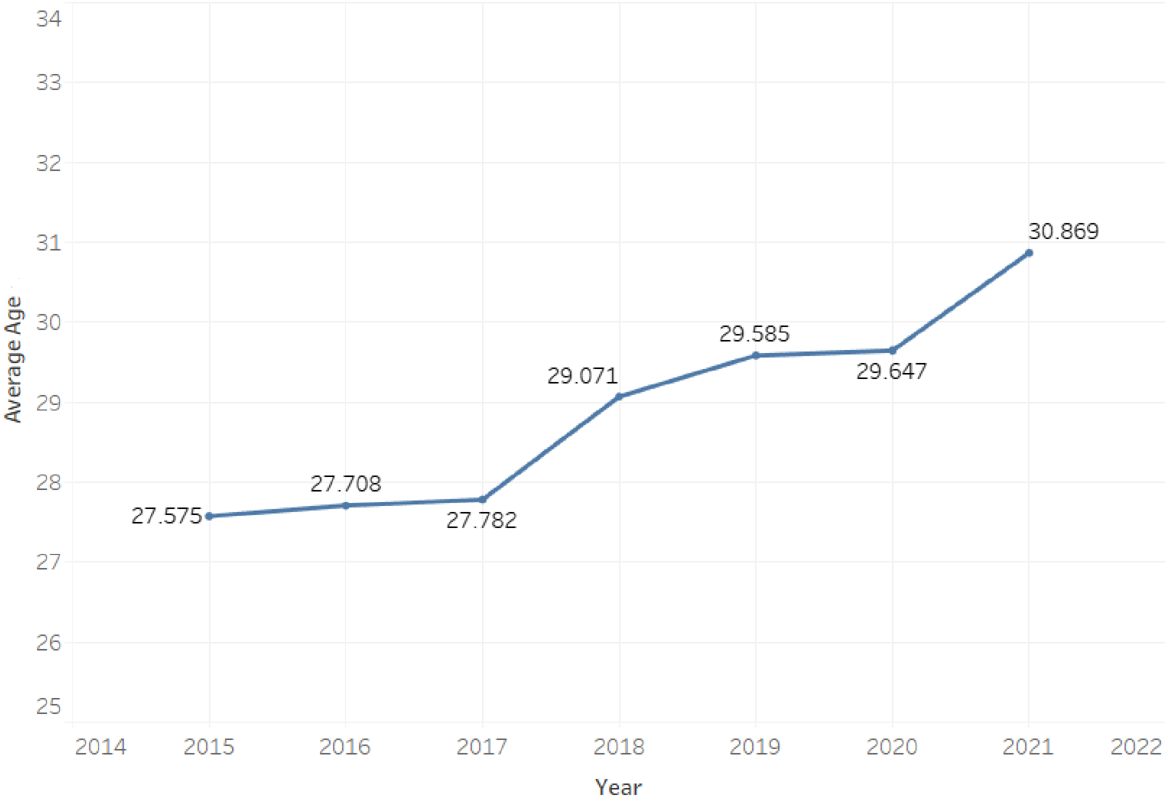


Table 2. Chicago gun-homicide decedents' mean age in years by race-ethnicity-sex group from 2015-2021.

Race-ethnicity-sex group	Year of death and mean age (years)						
	2015	2016	2017	2018	2019	2020	2021
Non-Hispanic Black female	27	28	24	38	28	27	31
Non-Hispanic White female	—[a]	53	60	28	48	31	37
Hispanic female	29	22	24	28	34	24	28
Non-Hispanic Black male	28	28	28	29	29	30	31
Non-Hispanic White male	31	34	34	41	34	30	39
Hispanic male	27	27	27	27	29	30	30

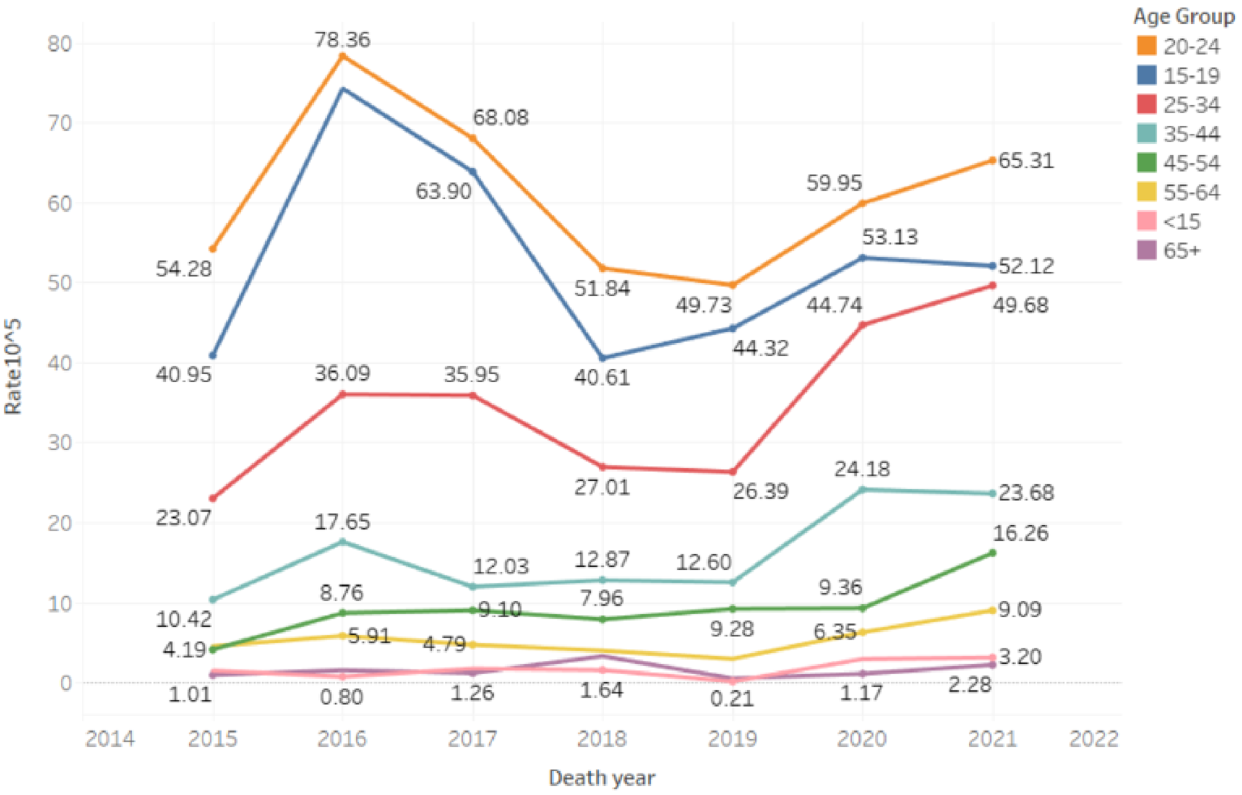
[a] Not applicable.

We also examined the annual distribution of decedents by age group across each of the 6 race-ethnicity-sex groups using tests of column proportions to evaluate statistical significance of differences (data not shown). Only 2 of the 6 race-ethnicity-sex groups examined had statistically significant changes in the annual distribution of decedents. The non-Hispanic Black female group had a statistically significant decrease in the proportion of gun-homicide deaths among those aged 65+ years, dropping from 16.7% (n=5/30) of gun-homicide deaths among non-Hispanic Black female individuals in 2018 to 2.2% (n<5/45) in 2020 (*P*=.02). However, the total number of non-Hispanic Black female deaths among those aged 65+ years during the study period was low (n=6/3702).

Among the non-Hispanic Black male group, there were statistically significant drops in the proportion of 15–19-year old gun-homicide decedents from 2016 (n=94/615, 20.1%) to 2020 (n=53/618, 11.8%; *P*=.01), from 2016 (n=94/615, 20.1%) to 2021 (n=54/685, 10.4%; *P*<.001), and from 2017 (n=78/550, 19.1%) to 2021 (n=54/685, 10.4%; *P*=.003). Among the non-Hispanic Black male group, there were statistically significant drops in the proportion of 20–24-year-old gun-homicide decedents from 2015 (n=95/391, 30.4%) to 2020 (n=81/618, 18.1%; *P*=.002), from 2015 (n=95/391, 30.4%) to 2021 (n=104/685, 20%; *P*=.01), and from 2016 (n=125/615, 26.8%) to 2020 (n=81/618, 18.1%; *P*=.04). Meanwhile, there were statistically significant increases in the proportion of 25–34-year-old non-Hispanic Black male gun-homicide decedents from 2015 (n=99/391, 31.6%) to 2020 (n=191/618, 42.6%; *P*=.04) and from 2016 (n=146/615, 31.3%) to 2020 (n=191/618, 42.6%; *P*=.008).

We also examined gun-homicide rates per 100,000 persons for each age group (see Figure 3). Rates offer important information about the relative risk to groups based on their proportion of the total population. The highest gun-homicide rates were consistently among the 15-19, 20-24, and 25-34 years age groups. From 2020 to 2021, the 15-19 and 35-44 years age groups saw slight declines. Most other groups increased, although the magnitude of increase differed by age group.

Figure 3. Chicago gun-homicide rate by age group from 2015-2021.



Discussion

Principal Findings

The annual distribution of gun-homicide decedents by race-ethnicity-sex group had been relatively stable over time, even in the context of year-to-year changes in gun-homicide rates and the 2015-2021 increase in gun-homicide rates in Chicago. There was a large but nonstatistically significant increase in the proportion of non-Hispanic Black female gun-homicide decedents from 2015 (3.6%) to 2021 (8.2%). The lack of statistical significance may be due to a relatively small (238/3723, 6.4%) number of non-Hispanic Black female individuals among gun-homicide decedents. The mean age of Chicago gun-homicide decedents increased by 3.27 years from 2015-2021. This increase in mean age coincided with a decrease in the proportion of non-Hispanic Black male gun-homicide decedents between the ages of 15-19 and 20-24 years and, conversely, an increase in the proportion of non-Hispanic Black male gun-homicide decedents between the ages of 25-34 years.

Although there is much discussion of firearm violence in Chicago in news publications [12] and, less often, research literature [13,14], there is little detailed information available on the demographics of gun-homicide decedents in the face of changes in violence other than to note broad populations at the highest risk at one point in time. This study examined the intersection of race/ethnicity, sex, and age among of Chicago gun-homicide decedents over a 7-year period in which there were year-to-year changes in overall gun-homicide rates. It provides key information on group representation in gun homicide during periods of escalation and periods of retraction. The changes in the makeup of gun-homicide decedents we detected signal a need for expanded prevention outreach and intervention engagement marketed toward non-Hispanic Black female individuals of all ages and non-Hispanic Black male individuals between the ages of 25-34 years, while at the same time, efforts are needed to expand outreach to and engagement of younger non-Hispanic Black male individuals, who consistently have the highest gun-homicide rates in Chicago.

This study has several strengths and limitations. In terms of strengths, the study incorporated very recent and reliable surveillance data from the Cook County Medical Examiner's Office, a nonpartisan office led by an appointed chief medical examiner. However, it is well known that death certificate reporting of race and ethnicity underreports certain subgroups [15], and so, it is likely that certain race/ethnicity groups, especially Alaska Native or American Indian individuals, those with more than one race/ethnicity affiliation, and Hispanic individuals may be underreported in the Medical Examiner data used in this study. Regarding other limitations, our reliance on US Census age groups as denominators for calculating homicide rates prevents a more nuanced analysis of single-year age risks. Further, although we report on demographics of gun-homicide decedents, ideally, more information on the circumstances surrounding these gun-related homicide deaths is needed to inform intervention approaches.

Conclusions

The findings from this modest study provide a basis for informing firearm violence prevention and intervention strategies and can inform a research agenda going forward. Further research could address the following questions: Will this shift toward increased non-Hispanic Black female and non-Hispanic Black male gun-homicide decedents between the ages of 25-34 year endure over time? What are the drivers associated with this increase in the proportion of non-Hispanic Black female gun-homicide decedents? and Are these age and gender shifts accompanied by changes in the precipitators and features of gun-homicide events? Because gun-homicide is a leading public health problem in the United States and, even more so, locally in Chicago, understanding the stability of and changes within decedent demographic subgroups is part of a required first step to inform targeted prevention responses.

Acknowledgments

We would like to acknowledge the Cook County Medical Examiner's Office for making archive data publicly available for this project.

Data Availability

Gun-homicide decedent data are available from the Cook County Medical Examiner Case Archive [9]. Population data came from the 2020 US Census [11].

Conflicts of Interest

None declared.

References

1. Hemenway D, Nelson E. The scope of the problem: gun violence in the USA. Curr Trauma Rep 2020 Jan 11;6(1):29-35. [doi: 10.1007/s40719-020-00182-x]
2. National Center for Health Statistics. FastStats - assault or homicide. Centers for Disease Control and Prevention. 2022 Sep 6. URL: https://www.cdc.gov/nchs/fastats/homicide.htm [accessed 2022-10-13]
3. National Center for Injury Prevention and Control. 10 leading causes of death by age group, United States - 2018. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/injury/images/lc-charts/leading_causes_of_death_by_age_group_2018_1100w850h.jpg [accessed 2021-10-21]
4. About underlying cause of death, 1999-2020. Centers for Disease Control and Prevention. 2019. URL: http://wonder.cdc.gov/ucd-icd10.html [accessed 2021-06-30]
5. National Center for Injury Prevention and Control. The public health approach to violence prevention. Centers for Disease Control and Prevention. 2022 Jan 18. URL: https://www.cdc.gov/violenceprevention/pdf/PH_App_Violence-a.pdf [accessed 2023-02-01]
6. Bentle K, Berlin J, Marx R, Rumore K. 40,000 homicides: retracing 63 years of murder in Chicago. Chicago Tribune. 2021 Apr 27. URL: https://www.chicagotribune.com/news/breaking/ct-history-of-chicago-homicides-htmlstory.html [accessed 2023-03-13]
7. Chadha S, McLone S, Mason M, Sheehan K. Adolescent firearm homicides in Chicago, 2013-2017. J Adolesc Health 2020 Sep;67(3):438-443. [doi: 10.1016/j.jadohealth.2020.02.025] [Medline: 32295721]

8. Walker GN, McLone S, Mason M, Sheehan K. Rates of firearm homicide by Chicago region, age, sex, and race/ethnicity, 2005-2010. J Trauma Acute Care Surg 2016 Oct;81(4 Suppl 1):S48-S53. [doi: 10.1097/TA.0000000000001176] [Medline: 27488485]
9. Medical Examiner Case Archive. Cook Country Government. 2021. URL: https://datacatalog.cookcountyil.gov/Public-Safety/Medical-Examiner-Case-Archive/cjeq-bs86 [accessed 2021-06-30]
10. Feld S, Bauldry S. Separate, unequal, and uncorrelated: why we need to consider race-specific homicide rates in US metropolitan areas. Socius 2018 May 03;4:237802311877395. [doi: 10.1177/2378023118773959]
11. American Community Survey, 5-year estimates subject tables: S0101 AGE AND SEX. United States Census Bureau. 2021. URL: https://data.census.gov/table?q=CHICAGO+AGE&tid=ACSST5Y2021.S0101 [accessed 2023-03-13]
12. Hendershot S. All big cities have a violence problem: Chicago's is different. Crain's Chicago Business. October 24. 2022 Oct 24. URL: https://www.chicagobusiness.com/crains-forum-safer-chicago/chicago-violence-problem-debate-safety-inequality [accessed 2023-03-13]
13. Sutherland M, McKenney M, Elkbuli A. Gun violence during COVID-19 pandemic: paradoxical trends in New York City, Chicago, Los Angeles and Baltimore. Am J Emerg Med 2021 Jan;39:225-226 [FREE Full text] [doi: 10.1016/j.ajem.2020.05.006] [Medline: 32444297]
14. Fontanarosa PB, Bibbins-Domingo K. The unrelenting epidemic of firearm violence. JAMA 2022 Sep 27;328(12):1201-1203. [doi: 10.1001/jama.2022.17293] [Medline: 36166046]
15. Wang K, Grossetta Nardini H, Post L, Edwards T, Nunez-Smith M, Brandt C. Information loss in harmonizing granular race and ethnicity data: descriptive study of standards. J Med Internet Res 2020 Jul 20;22(7):e14591 [FREE Full text] [doi: 10.2196/14591] [Medline: 32706693]

Edited by A Mavragani, T Sanchez; submitted 21.10.22; peer-reviewed by T Alcorn, D Bowen; comments to author 12.12.22; revised version received 13.02.23; accepted 23.02.23; published 07.04.23.

Please cite as:
Mason M, Khazanchi R, Brewer A, Sheehan K, Liu Y, Post L
Changes in the Demographic Distribution of Chicago Gun-Homicide Decedents From 2015-2021: Violent Death Surveillance Cross-sectional Study
JMIR Public Health Surveill 2023;9:e43723
URL: https://publichealth.jmir.org/2023/1/e43723
doi:10.2196/43723
PMID:37027193

Original Paper

COVID-19 Testing Practices, Preventive Behaviors, and Factors Associated With Test Positivity: Population-Based Statewide Survey Study

Mufaro Kanyangarara[1], PhD; Virginie Daguise[1,2], PhD; Lídia Gual-Gonzalez[1], MSc; Alain H Litwin[3,4,5], MPH, MD; Jeffrey Korte[6], PhD; Connor Ross[1], BA; Melissa S Nolan[1,2], PhD; SC STRONG Project Team[7]

[1]Department of Epidemiology and Biostatistics, University of South Carolina Arnold School of Public Health, Columbia, SC, United States
[2]South Carolina Department of Health and Environmental Control, Columbia, SC, United States
[3]University of South Carolina School of Medicine, Greenville, SC, United States
[4]Prisma Health, Greenville, SC, United States
[5]Clemson University, Clemson, SC, United States
[6]Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC, United States
[7]See Acknowledgements, Columbia, SC, United States

Corresponding Author:
Mufaro Kanyangarara, PhD
Department of Epidemiology and Biostatistics
University of South Carolina Arnold School of Public Health
915 Greene Street
Columbia, SC, 29208
United States
Phone: 1 8037777563
Email: mufaro@mailbox.sc.edu

Abstract

Background: The COVID-19 pandemic has challenged public health efforts globally. Timely population-based surveillance is crucial to support public health programs and policies to limit the spread of COVID-19. The South Carolina (SC) Sampling and Testing Representative Outreach for Novel Coronavirus Guidance (SC STRONG) statewide initiative was established to estimate population-level prevalence and immunity and characterize the transmission dynamics of SARS-CoV-2 using community testing and online surveys.

Objective: This paper aimed to leverage the survey data collected as part of the initiative to understand risk perceptions, testing practices, and preventive behaviors and identify risk factors for COVID-19 test positivity in SC over time.

Methods: Probability proportionate to size cluster random sampling was used to select SC residents to participate in testing for COVID-19 infection and antibodies and to complete an online survey. This paper focuses on data from the online surveys completed between November 2020 and June 2021. Descriptive statistics were used to describe risk perceptions, attitudes and behaviors, and associated changes over time. Univariate and multivariate logistic regression models were used to identify factors associated with self-reported COVID-19 test positivity.

Results: Among the 7170 online survey respondents, 58.7% (4213/7170) self-reported ever testing for COVID-19. The most commonly cited barriers to testing were inconvenient dates, time, and location, as well as discomfort. Overall, 18.7% (790/7170) of respondents reported a history of COVID-19 test positivity. Multivariate logistic regression results indicated that individuals who were aged 50 years or older, self-identified as Black/African American, were obese, and were employed as frontline health care workers or nursing home staff were more likely to self-report COVID-19 test positivity. By contrast, there was a decreased likelihood of test positivity among respondents who were concerned about the burden of COVID-19 in their community and about being infected.

Conclusions: Strategies to remove testing barriers should be implemented to improve access. Our findings provide insights on statewide testing patterns, adoption of prevention behaviors, and risk factors for infection and may inform public health strategies to curb transmission.

(JMIR Public Health Surveill 2023;9:e34579) doi:10.2196/34579

KEYWORDS

SARS-CoV-2; South Carolina; surveillance; attitude; behavior; COVID-19; testing; prevention; United States; population; survey; risk; perception; risk factor

Introduction

The year 2020 brought new public health challenges and highlighted the need for data to inform ongoing public health response efforts. In South Carolina (SC), the first case of COVID-19 was detected on March 4, 2020 [1]. As of July 2021, the cumulative numbers of confirmed COVID-19 cases and associated deaths in SC were over 700,000 and 11,500, respectively [1]. Mitigation measures in the state have included face mask ordinances, closure of nonessential businesses, isolation of COVID-19 cases, tracing and quarantining of close contacts of cases, and promotion of personal protective behaviors, such as routine testing, face coverings, hand hygiene, and social distancing. These measures were implemented to reduce disease spread and remain crucial to prevent overcrowding hospitals and emergency rooms with patients with severe disease during local outbreaks [2]. Even as COVID-19 vaccines are rolled out, preventive behaviors remain crucial components in the arsenal to address the pandemic. To effectively minimize COVID-19 transmission risk and prevent the loss of life, achieving and sustaining high uptake of preventive behaviors needs to be prioritized [3,4].

The adoption of preventive behaviors in part depends on the perceived severity of disease, perceived susceptibility to disease, benefits of compliance, and removal of barriers to the adoption of protective behaviors [5]. Knowledge and attitudes partially influence the adoption of these behaviors [6]. While several studies have assessed COVID-19 attitudes and behaviors, most were at the national scale or based on convenience sampling [7-11]. Some of these studies discuss the role of information exposure and knowledge in health behavior, and national-level estimates are likely to differ from state-level estimates. There is a paucity of research at the state level, particularly in SC, a state primarily defined by conservative ideology; thus, a description of the population’s perceptions is warranted for a locally tailored public health response [12-14]. Given the continuing incidence of COVID-19, the overwhelming public health burden, and the rapidly changing epidemiology of COVID-19, there is an urgent need for timely population-based surveillance data to support state-specific, targeted policies and public health efforts to address the pandemic.

The objectives of this analysis were 2-fold: first, to understand SC respondents’ testing patterns, risk perceptions, and preventive behaviors and associated temporal changes and second, to identify risk factors for COVID-19 test positivity to inform public health policies and programs in SC. Understanding the population’s perceived risk allows for the targeting of public health interventions to reduce the burden of COVID-19 while addressing health inequities [15]. This paper reports on results from 3 rounds of surveys conducted from November 2020 to June 2021 among 7170 SC residents participating in the SC Sampling and Testing Representative Outreach for Novel Coronavirus Guidance (SC STRONG) project. The SC STRONG project was established in October 2020 across all 4 public health regions in the state with the goal of estimating population-level seropositivity and immunity and characterizing population-specific transmission dynamics [16]. This ongoing initiative is led by the SC Department of Health and Environmental Control (SC DHEC) together with multidisciplinary academic collaborators, local health clinics, and health care providers across the state. The project consists of multi-round cross-sectional community testing and surveys, and to date, 3 rounds of community testing and surveys have been completed. The project complements reportable disease surveillance for COVID-19 by characterizing attitudes and behaviors toward public health measures to limit COVID-19 transmission among a large random sample of SC residents.

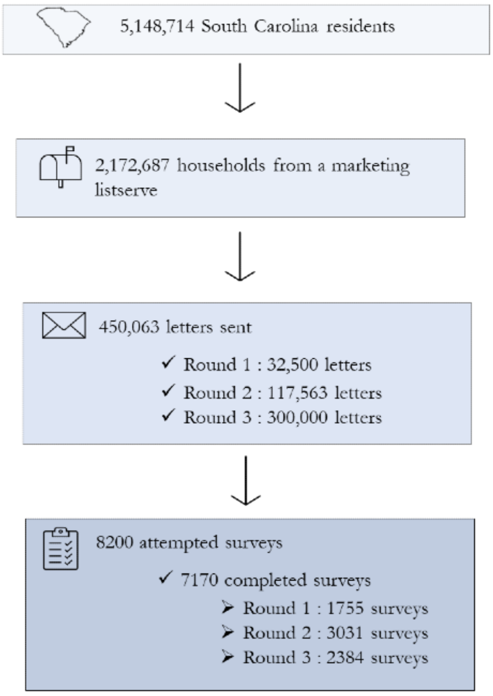
Methods

Setting and Study Design

SC has a population of approximately 4.86 million inhabitants aged 5 years and older [17]. The state has 46 counties distributed across 4 public health regions, all of which were included in the SC STRONG project. A description of the SC STRONG project has been published elsewhere [18]. Briefly, a direct mailer marketing list with 2,172,687 unique physical addresses was purchased from Mailers Haven and used to establish a sampling frame (Figure 1). Multi-stage cluster sampling with probability proportionate to size was used to select clusters and 30 residents from each selected cluster. Invitation letters to participate in testing for active COVID-19 infection and antibodies and to complete an online survey were mailed to selected residents. The invitation letters included an explanation of the project, instructions on how to participate in free community testing, and a QR code and URL for the online electronic survey. The number of invitations mailed each round varied depending on the anticipated response rate. To improve response rates, at least two reminders were sent to selected residents. Additionally, to accommodate participants with no internet access and allow them to complete the online survey, the testing site provided the option to fill out the survey on site.

The online survey was administered via REDCap (Research Electronic Data Capture) and hosted by Health Sciences South Carolina. REDCap is a secure web-based software platform designed to support data capture for research studies [19]. A standardized survey questionnaire available in English and Spanish was used to collect data on sociodemographic characteristics, testing behaviors, daily behaviors, and risk perceptions. The questionnaire consisted of questions from existing COVID-19 surveys and new questions developed by the SC STRONG team [20]. Additional survey questions regarding acceptance, motivations, and barriers to getting the COVID-19 vaccine were added in January 2021.

Figure 1. Flow diagram of study participation.



Statistical Analysis

Descriptive statistics and visualizations were used to describe the sample population and chi-square statistics were used to compare sociodemographic characteristics of the SC general population, survey respondents, and respondents who reported ever testing for COVID-19. Data on the SC general population were obtained from the 2019 American Community Survey [17]. Testing behaviors and barriers, attitudes and barriers toward the vaccine, and uptake of preventive behaviors were summarized. The preventive behaviors considered included practicing social distancing, wearing a mask, and self-isolation or quarantining. Respondents were considered adherent to preventive behaviors if they reported practicing these behaviors “sometimes,” “most of the time,” or “all the time” during the previous 2 weeks. Univariate logistic regression models with “week” as an independent variable were used to assess temporal changes in behaviors and attitudes. Changes in behaviors and attitudes per week were expressed as maximum likelihood estimates with 95% CIs.

History of testing was derived from 2 survey questions: “Have you ever been tested for active coronavirus infection (nasal or saliva test)?” and “Have you ever been tested for coronavirus antibodies (blood test)?” As the goal was to identify risk factors for a positive COVID-19 test among those with history of testing, individuals with no prior testing for active coronavirus infection or antibodies were excluded from this analysis. The outcome of interest was based on self-reported diagnosis of coronavirus antibodies or active coronavirus infection derived from the following survey questions: “Have you ever tested positive for active coronavirus infection?” and “Have you ever tested positive for coronavirus antibodies?” Respondents who answered yes to either question were categorized as having ever tested positive for COVID-19. The likelihood of COVID-19 test positivity was assessed with univariate multivariable logistic regression. Explanatory variables of interest were selected based on previous studies assessing risk factors for COVID-19. To account for likely racial inequalities, “Black” and “Hispanic” were added as indicator variables. Household income was added as a proxy for socioeconomic status. Other explanatory variables considered were preexisting medical conditions, number of household members, living or working in congregate settings, testing behaviors, risk perceptions, and knowledge of COVID-19 preventive measures. To account for differences in risk of COVID-19 over time, models adjusted for time in weeks since recruitment for SC STRONG began. In addition to age, gender, income, and race or ethnicity, which were specified a priori, variables with a *P* value less than 0.1 in the univariate logistic regression models were included in the initial multivariable regression model. A stepwise backwards elimination procedure was used to obtain the final multivariable model. Crude and adjusted odds ratios (aORs) from the logistic regression analyses are reported. Checks for multicollinearity were performed using variance inflation factors (VIFs). Covariates with a VIF greater than 8 were dropped. A *P* value less than .05 was considered statistically significant. Analyses were conducted using STATA/SE (version 16.1; StataCorp) and SAS Studio (version 3.8; SAS Institute).

Ethical Considerations

Human-subjects ethics approval was sought for the SC STRONG project. The institutional review boards at the University of South Carolina and the SC DHEC both determined that SC STRONG project activities did not constitute human-subjects research under the auspices of public health surveillance (102072). Participation in testing and survey completion was voluntary.

Results

Overview

From October 26, 2020, to June 16, 2021, 3 rounds of data collection were conducted, and 450,063 invitation letters were mailed out (Figure 1). The first round of data collection took place from October 26, 2020, to January 22, 2021, and 32,500 invitation letters were mailed. This round coincided with the second wave, when daily COVID-19 cases peaked to over 5500 (Figure 2). The second round of data collection took place from January 25, 2021, to March 31, 2021, and 117,563 letters were mailed (Figure 1). This round was toward the end of the second wave of COVID-19, when the daily number of cases fell to less than 1200 per day and COVID-19 vaccines became available under phase 1b to populations living and working in shared settings with increased risk and to frontline essential workers (Figure 2). The third round of data collection ran from April 2, 2021, to June 16, 2021, and 300,000 letters were mailed (Figure 1). At the time of this round, COVID-19 vaccines were available to individuals aged 16 years or older (Figure 3).

A total of 8200 surveys were attempted across the 3 rounds, and 1030 surveys were excluded due to incomplete or inconsistent responses; therefore, a total of 7170 completed surveys was used for the present analysis (Figure 1). The response rate for all 3 rounds was 1.6% (7170/450,063). The response rate decreased with each round, from 5.4% (1755/32,500) in the first round to 0.8% (2384/300,000) in the third round. There was also variation in the percentage of responses by county of residence (Figure 3). More survey respondents lived in the upstate region (eg, Greenville County at 950/7091, 13.4%), in coastal areas (eg, Charleston County at 726/7091, 10.3%, and Horry County at 639/7091, 9%), and in areas surrounding the capital city (eg, Lexington County at 500/7091, 7.1%, and Richland County at 667/7091, 9.4%).

Most survey respondents were aged 60 years or older (4087/7170, 57%), female (4151/7170, 57.9%), and Caucasian (6087/7170, 84.9%; Table 1). Survey respondents were fairly representative of the general SC population in terms of gender and income. However, Hispanic/Latino Americans and African Americans were underrepresented in the survey. Whereas 5.8% and 26.3% of the general SC population are Hispanic/Latino Americans and African Americans, only 2.2% (158/7170) and 8.9% (683/7170) of survey respondents identified as Hispanic/Latino Americans and African Americans, respectively. Moreover, older age groups were overrepresented in the study population. There were also differences among survey respondents by COVID-19 testing status. Respondents who were previously tested for COVID-19 tended to be younger, female, and have higher income than the overall sample of survey respondents (Table 1).

Figure 2. Trends in number of COVID-19 cases and deaths in South Carolina, public health measures, and data collection rounds. SC: South Carolina.

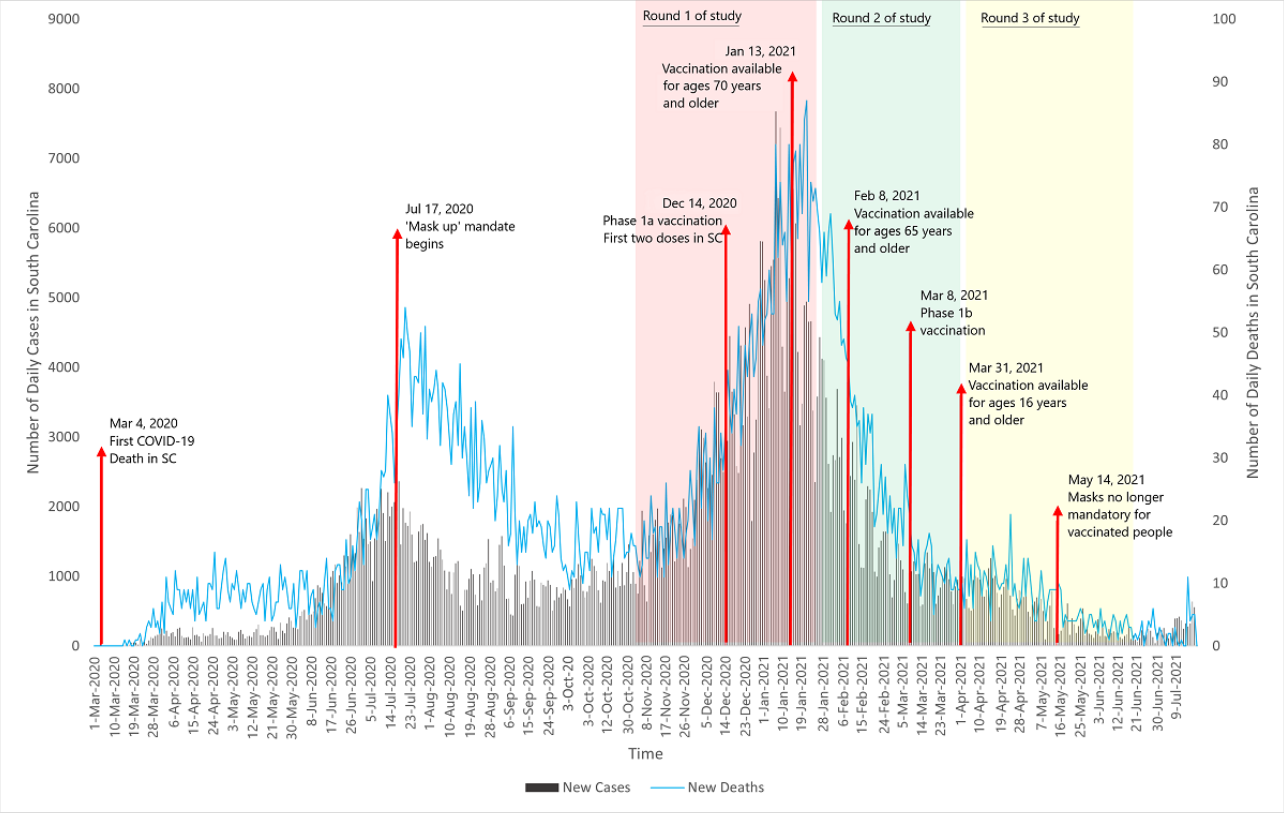


Figure 3. Percentage of responses by county.

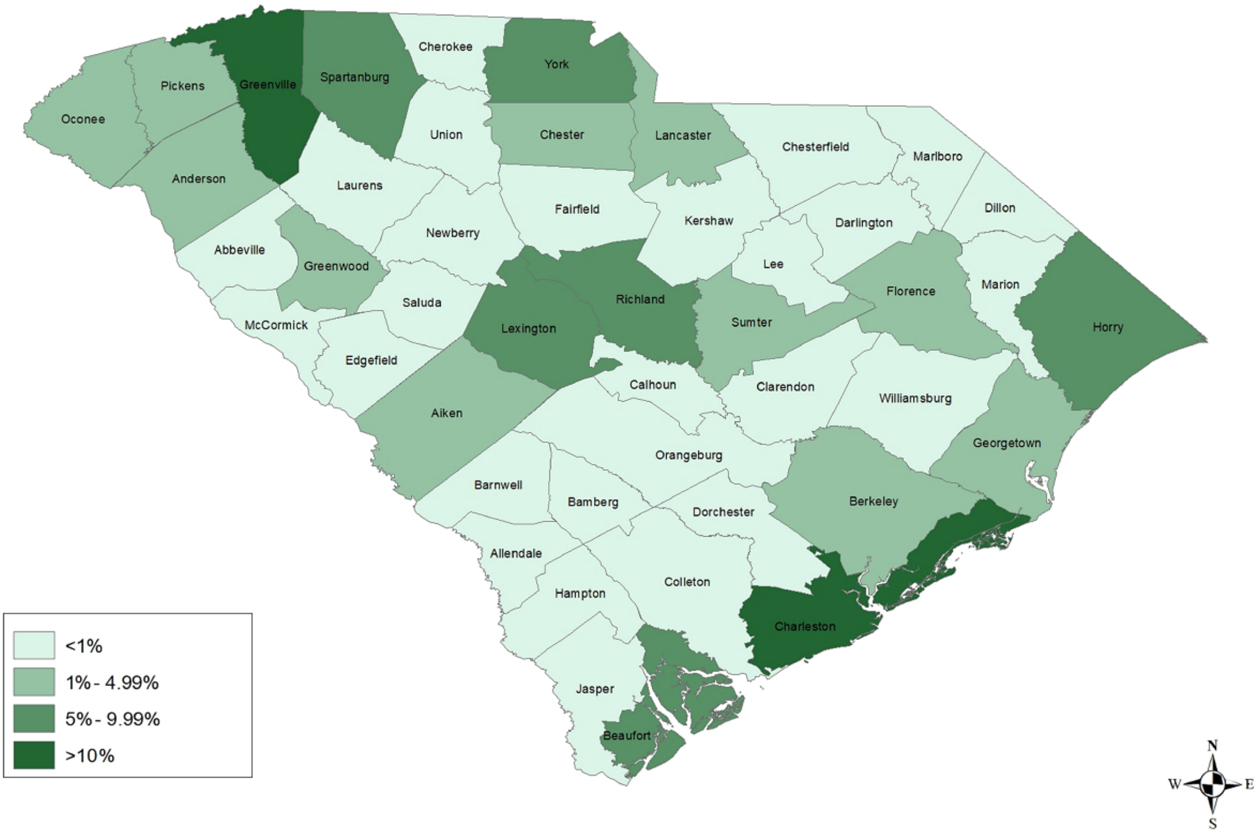


Table 1. Sociodemographic characteristics of survey respondents (n=7170) and respondents ever tested for COVID-19 (n=4197).

Characteristics	General South Carolina population[a] (N=5,148,714), %	Survey respondents (n=7170), n (%)	*P* value[b]	Ever tested for COVID-19 (n=4197), n (%)	*P* value[c]
Age (years)			<.001		<.001
<18	21.6	122 (1.7)		0 (0)	
18-29	15.9	194 (2.7)		201 (4.8)	
30-39	12.6	445 (6.2)		290 (6.9)	
40-49	12	882 (12.3)		546 (13)	
50-59	13	1441 (20.1)		894 (21.3)	
60-69	12.6	2251 (31.4)		1267 (30.2)	
≥70	12.3	1836 (25.6)		999 (23.8)	
Gender			.40		<.001
Female	51.7	4151 (57.9)		2514 (59.9)	
Male	48.3	2976 (41.5)		1683 (40.1)	
Other	0	43 (0.6)		0 (0)	
Race/ethnicity[d]			.007		
Hispanic/Latino American	5.8	158 (2.2)		92 (2.2)	.80
African American	26.3	638 (8.9)		436 (10.4)	.10
Caucasian	63.6	6087 (84.9)		3542 (84.4)	.30
Asian	1.7	129 (1.8)		55 (1.3)	<.001
Native American	0.3	50 (0.7)		29 (0.7)	.90
Other/prefer not to answer	0.2	222 (3.1)		25 (0.6)	.20
Income (US $)			.20		<.001
≤14,999	11.5	208 (2.9)		109 (2.6)	
15,000-34,999	19.1	595 (8.3)		327 (7.8)	
35,000-49,999	13.9	638 (8.9)		378 (9)	
50,000-74,999	18.2	1097 (15.3)		646 (15.4)	
75,000-99,999	12.6	939 (13.1)		571 (13.6)	
100,000-149,000	13.6	1111 (15.5)		638 (15.2)	
≥150,000	11	968 (13.5)		613 (14.6)	

[a] Data obtained from South Carolina Department of Health and Environmental Control [16].

[b] *P* values (chi square test) for differences between the South Carolina general population and the survey respondents.

[c] *P* values (chi square test) for differences between those ever tested for COVID-19 and those with no history of testing.

[d] Percentages may add to more than 100%, as survey respondents could give multiple answers.

Impact of the Pandemic

The COVID-19 pandemic impacted the lives of SC residents in multiple ways. In terms of employment, 3.8% (271/7170) reported becoming unemployed during the pandemic, 10.6% (760/7170) reported reductions in income or pay, and 7.8% (562/7170) reported working fewer hours than normal (Table 2). In terms of mental and emotional health, 72.1% (5173/7170) reported often feeling stressed, nervous, or on the edge, 56.7% (4063/7170) reported often feeling sad or depressed in the previous 2 weeks, and 19% (1365/7170) reported having physical reactions when thinking about the pandemic. However, only 2.6% (191/7170) reported being diagnosed with anxiety or depression during the pandemic. In terms of food security, very few respondents reported experiencing lack of food (73/7170, 1%) or going without eating (282/7170, 4%) in the previous 2 weeks due to lack of money or resources.

Table 2. Trends in behaviors and attitudes among South Carolina residents toward testing, vaccination, and preventive measures for the COVID-19 pandemic (n=7170).

Behaviors and attitudes	Respondents, n (%)	Change per week[a]	*P* value
Ever been tested	4213 (58.7)	0.03	<.001
Active infection[b]	3988 (94.6)	0.03	<.001
Antibodies[b]	862 (20.5)	0.02	<.001
Ever tested positive[b]	790 (18.8)	0.03	<.001
Reasons for testing[b,c]			<.001
Close contact tested positive	1397 (33.2)	0.02	<.001
Coronavirus-like symptoms	1314 (31.2)	0.01	<.001
Concerned about high number of cases in community	789 (18.7)	0.02	<.001
Medical, travel, or employment requirement	834 (19.7)	0.02	<.001
Curiosity	721 (17.1)	0.02	<.001
Testing location[b]			<.001
Doctor’s office	943 (22.3)	0.01	<.001
Drive through or community pop-up	2746 (65.2)	0.02	<.001
Frequency of COVID–19 testing[b]			
1 time	1841 (43.7)	–0.03	<.001
2-5 times	2167 (51.4)	0.02	<.001
6 or more times	205 (4.9)	0.03	<.001
Barriers to testing[c,d]			
Inconvenient dates and times	149 (20.3)	–0.03	.002
Discomfort of the test	141 (19.2)	–0.03	<.001
Did not know where to go	118 (16.1)	–0.07	<.001
Economic	98 (13.4)	–0.03	.007
Inconvenient location	87 (11.9)	–0.03	.01
Attitudes toward vaccines[c]			
Vaccinated	2498 (45.8)	0.20	<.001
Think vaccines are safe	3982 (72.9)	0.02	<.001
Plan to get vaccinated	2270 (41.6)	–0.16	<.001
No plan to get vaccinated	516 (9.5)	N/A[e]	
Barriers to vaccination[c]			
Fear of needles	24 (4.2)	—[f]	.9
Uncomfortable being one of the first vaccinated	282 (49.4)	—[f]	.06
Practice of preventive behaviors			
Social distancing	7068 (98.6)	–0.1	<.001
Mask wearing	7035 (98.1)	–0.07	<.001
Self–isolation/quarantine	676 (9.4)	–0.04	<.001
Impact of pandemic on work			
Became unemployed	271 (3.8)	–0.06	<.001
Pay was reduced	760 (10.6)	–0.07	<.001
Worked with children at home	455 (6.3)	–0.07	<.001

Behaviors and attitudes	Respondents, n (%)	Change per week[a]	*P* value
Worked less hours	562 (7.8)	–0.08	<.001
Worked from home more	1469 (20.5)	–0.08	<.001
Impact of pandemic on mental health			
Being sad/depressed in the past 2 weeks	4063 (56.7)	–0.02	<.001
Being diagnosed with depression or anxiety	191 (2.6)	N/A[e]	
Having felt physical reactions when thinking about the pandemic in the past 2 weeks	1365 (19)	–0.02	<.001
Having felt some level of stress, anxiety, or being on the edge in the past 2 weeks	5173 (72.1)	–0.03	<.001
Impact of pandemic on food insecurity			
Lack of food in the past 2 weeks	73 (1)	–[f]	.1
Gone without eating in the past 2 weeks	282 (4)	–0.02	<.001

[a] Adjusted logistic regressions including time (in weeks) as a predictor to assess for trends over time.

[b] Percentages represent percentages of those ever tested.

[c] Percentage may add more to than 100%, as survey respondents could give multiple answers.

[d] Percentage of those with no history of testing (n=593).

[e] N/A: not applicable, as the variable violated the assumption for logistic regression.

[f] Point estimate for the change per week was smaller than 0.001; the – symbol indicates the direction of the change (ie, negative).

COVID-19 Testing Practices and Barriers

Of the 7170 survey respondents, 58.7% (4213/7170) reported ever testing for COVID-19 active infection or antibodies, and there was a 3% increased likelihood of testing for COVID-19 with every week that passed (Table 2). The most common barriers to testing were inconvenient times and dates (143/593, 20.3%), the discomfort of the test (141/593, 19.2%), and lack of knowledge of where to go to be tested (118/593, 16.1%). Nevertheless, the likelihood of citing these testing barriers decreased with time. On the other hand, the most common reasons for getting tested were having a close contact who had tested positive (1397/4213, 33.2%); having coronavirus-like symptoms (1314/4213, 31.2%); medical, travel, or employment requirements (834/4213, 19.8%); and being concerned about the high number of COVID-19 cases in the community (789/4213, 18.7%). A less common reason for testing was curiosity (721/4213, 17.1%). More than half of respondents received testing at a drive-through or community pop-up testing site (2746/4213, 65.2%) and reported being tested more than once (2372/4213, 56.3%). Over time, respondents were more likely to get tested and reported getting tested in a drive-through or community pop-up testing site. Of those who had ever been tested, 18.8% (790/4213) reported testing positive for antibodies or active infection at some point.

COVID-19 Preventive Behaviors

Most respondents reported social distancing (7068/7170, 98.6%) and wearing face masks (7035/7170, 98.1%) in the previous 2 weeks, and fewer reported self-isolation (676/7170, 9.4%) or quarantining (Table 2). The likelihood of engaging in these practices decreased with time (*P*<.001). Overall, 45.8% (2498/5460) of respondents reported being vaccinated against COVID-19. An assessment of attitudes toward the vaccine indicated that 72.9% (3982/5460) of respondents thought the COVID-19 vaccines were safe, and 41.6% (2270/5460) planned to get vaccinated at some point. Among those who reported not planning to get the vaccine, the main reason was discomfort with being one of the first to be vaccinated (282/571, 49.4%). Overall, there was an increased likelihood of vaccination (20%) and positive attitudes toward the vaccines per week.

Factors Associated With COVID-19 Test Positivity

In the multivariable logistic regression analysis, several factors were significantly associated with COVID-19 test positivity (Table 2). The likelihood of COVID-19 test positivity was increased among individuals who identified as Black/African American, were employed in a nursing home or as a frontline health care worker, were obese, and reported testing at the doctor's office. Furthermore, the aOR of a positive COVID-19 test was highest for individuals who decided to test because of coronavirus-like symptoms (aOR 7.37, 95% CI 6.08-8.94) and those with a family member or close friend who currently or previously had COVID-19 (aOR 4.61, 95% CI 3.41-6.24). On the other hand, the likelihood of COVID-19 test positivity was decreased among individuals who expressed concern about being infected (aOR 0.38, aOR 0.30-0.47) and who chose to test out of concern for the perceived high burden of COVID-19 in the community (aOR 0.28, 95% CI 0.20-0.40). There was no statistically significant difference in the likelihood of test positivity by gender or income.

Discussion

Principal Findings

Using serial population-based surveys conducted between November 2020 and June 2021, this paper sought to describe COVID-19 risk perceptions, testing patterns, and preventive behaviors and identify factors associated with test positivity among a large random sample of SC residents. Overall, 58.7% (4213/7170) of respondents reported ever testing for COVID-19, and over time, there was a decrease in barriers to testing. This finding is consistent with the increased availability of testing locations across the state over time. Moreover, our results suggested an increased likelihood of testing over time, an increased likelihood of positivity based on symptoms and exposure, and a decreased likelihood of positivity based on individual concern. Despite increased testing, there were decreases in practicing preventive behaviors such as mask-wearing and social distancing, indicating fatigue. Previous research has shown that states with no mask policy and low mask adherence reported high COVID-19 infection rates, while states with high mask adherence and strict mask-wearing policies showed decreased infection rates [21]. It is important to note that politics has had a drastic effect on COVID-19 epidemiology, and there has been contradictory messaging regarding infection information and uptake of mitigation policies. This is reflected in misinformation driven by aligned political beliefs that conflicts with the scientific evidence and recommendations from public health agencies [22].

Comparison With Prior Work

In this paper, 18.7% (790/7170) of respondents self-reported ever testing positive for COVID-19, which is comparable to the estimated seroprevalence of 16.4% for SC residents aged 5 years or older [18]. Our estimate could be influenced by self-selection, as our data for ever having tested positive were based on self-reports. Moreover, availability of the vaccine began in mid-January for individuals aged 70 or older, which likely also impacted testing demand. At the time of the survey, 58.7% (4213/7170) of respondents had ever been tested for COVID-19, whereas it has previously been reported that 66.7% received the vaccine, suggesting the possibility that a number of respondents did not actively seek testing due to fear of exposure [23]. Our results show the likelihood of test positivity was highest among Black/African Americans and older individuals (aged ≥50 years). Although the state does not report race or ethnicity data for testing, count data for COVID-19 deaths in SC indicate that Black/African Americans and older people are more likely to be unvaccinated, to be hospitalized for severe disease, and to die from COVID-19 [18]. The likelihood of test positivity decreased 38% in individuals who expressed concern about becoming infected, which aligns with the findings of Yildirim and Güler [24], who highlighted the direct association between perceived risk and positivity. When developing strategies to increase the adoption of preventive behaviors, understanding perceived individual risk is essential. Moreover, testing sites are often inconveniently located or open during times that are less available for people living in deep poverty due to the low flexibility of low-wage jobs and difficulty obtaining transportation. Therefore, considerable improvements should be made to accommodate the lowest socioeconomic strata, which cannot appropriately use these public health services.

Limitations

The findings presented in this paper should be interpreted in light of several limitations. First, given the cross-sectional survey design, the temporal relationship between outcome and exposures cannot be ascertained and causality cannot be determined. Second, the representativeness of the sample is limited due to the low response rate; the lack of partner clinics in some areas of the state may have hindered participation in the community testing and online survey. Third, the sampling frame was based on a purchased marketing listserve that may not be up-to-date. Fourth, COVID-19 test positivity was based on self-reports, which are prone to bias. Fifth, comparison of sociodemographic characteristics by history of testing indicated that the respondents who sought testing were more likely to be older, female, Caucasian, and have a higher income than the general population, potentially limiting the generalizability of our results. Lastly, although an electronic survey was selected to ensure respondent safety and minimize costs of survey implementation, only about 81.8% of households in SC have access to the internet [25]. Staff were available to assist with completion of the survey online, and paper surveys were also available at the pop-up events. The analysis was restricted to responses included in the survey; this approach did not represent nonrespondents and did not capture several socioeconomic and education variables that were not included in the survey. Similarly, the results need to be considered cautiously, as the distribution of respondents might only represent individuals with broadband internet access, which would underrepresent those populations that commonly suffer health disparities and inequities related to socioeconomic, educational, and racial background.

Conclusions

Population-based surveys are useful tools in understanding attitudes, behaviors, and practices and can inform public health responses and the adaptation of strategies to reach high-risk populations. Despite its limitations, the SC STRONG initiative represents a statewide outreach and partnership program with local clinics and health care providers. The findings presented here provide insights to understand risk perceptions and testing behaviors and have important implications given that Southern states have historically struggled with compliance and implementation of preventive public health measures.

Acknowledgments

We would like to thank the South Carolina residents and clinical partners who participated in this surveillance initiative. We would also like to thank our additional partners from the South Carolina Sampling and Testing Representative Outreach for Novel Coronavirus Guidance (SC STRONG) project team (listed in alphabetical order): Graham Adams (South Carolina Office of Rural Health), Roxanne Ancheta (South Carolina Department of Health and Environmental Control; SC DHEC), O’Neal Best (University of South Carolina; U of SC), Sarah Cockrell (South Carolina Primary Health Care Association), Madison Collins (U of SC), Hanson Cowan (U of SC), Megan Davis (SC DHEC), Kenneth R Deans, Jr (Health Sciences South Carolina; HSSC), Abbie Drake (College of Charleston), Joan M. Duwve (SC DHEC), Jan M. Eberth (U of SC), Katrina Fryar (HSSC), Scott Hagins (SC DHEC), Sophie Henry (U of SC), Moonseong Heo (Clemson University), Sadie Hildebrandt (U of SC), Amelia Hornaday (U of SC), Rongjie Huang (U of SC), Danielle Johnson (SC DHEC), Sammy Khalil (U of SC), Darin King (U of SC), Mary K. Lynn (U of SC), Josephine A Morrisey (U of SC), Kateryna Parkhomenko (U of SC), Nitin Patel (Prisma Health), Stella Self (U of SC), Windsor Westbrook Sherrill (Prisma Health, Greenville; Clemson University), Callie Shirley (U of SC), Jessica Siebenschuh (HSSC), Darin Thomas (Prisma Health), Madlyn Wills (U of SC), and Michael Yan (U of SC). We would like to thank Dr Ann Blair-Kennedy and the Patient Engagement Studio members for their contributions. This work was supported by the US Department of Health and Human Services (grant NU50CK000542) and US Centers for Disease Control and Prevention (grant #OT21-2103). The funders had no role in the design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. South Carolina County-Level Data for COVID-19. South Carolina Department of Health and Environmental Control. URL: https://scdhec.gov/covid19/south-carolina-county-level-data-covid-19 [accessed 2023-03-23]
2. Kwon S, Joshi AD, Lo C, Drew DA, Nguyen LH, Guo C, et al. Association of social distancing and face mask use with risk of COVID-19. Nat Commun 2021 Jun 18;12(1):3737 [FREE Full text] [doi: 10.1038/s41467-021-24115-7] [Medline: 34145289]
3. Mohammed I, Nauman A, Paul P, Ganesan S, Chen K, Jalil SMS, et al. The efficacy and effectiveness of the COVID-19 vaccines in reducing infection, severity, hospitalization, and mortality: a systematic review. Hum Vaccin Immunother 2022 Dec 31;18(1):2027160 [FREE Full text] [doi: 10.1080/21645515.2022.2027160] [Medline: 35113777]
4. Bruinen de Bruin Y, Lequarre A, McCourt J, Clevestig P, Pigazzani F, Zare Jeddi M, et al. Initial impacts of global risk mitigation measures taken during the combatting of the COVID-19 pandemic. Saf Sci 2020 Aug;128:104773 [FREE Full text] [doi: 10.1016/j.ssci.2020.104773] [Medline: 32296266]
5. Rosenstock IM. Historical Origins of the Health Belief Model. Health Educ Monogr 1974 Dec 01;2(4):328-335. [doi: 10.1177/109019817400200403]
6. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
7. Rakotoarisoa F, Praptapa A, Suyono E, Usuman I. Factors influencing preventive intention behavior towards COVID-19 in Indonesia. J Behav Sci 2021 Jan 24;16(1):14-27 [FREE Full text]
8. Masoud AT, Zaazouee MS, Elsayed SM, Ragab KM, Kamal EM, Alnasser YT, KAP-COVIDGLOBAL Investigators. KAP-COVID: a multinational survey of the levels and determinants of public knowledge, attitudes and practices towards COVID-19. BMJ Open 2021 Feb 23;11(2):e043971 [FREE Full text] [doi: 10.1136/bmjopen-2020-043971] [Medline: 33622949]
9. Geldsetzer P. Knowledge and perceptions of COVID-19 among the general public in the United States and the United Kingdom: A cross-sectional online survey. Ann Intern Med 2020 Jul 21;173(2):157-160 [FREE Full text] [doi: 10.7326/M20-0912] [Medline: 32196071]
10. Clements JM. Knowledge and behaviors toward COVID-19 among US residents during the early days of the pandemic: Cross-sectional online questionnaire. JMIR Public Health Surveill 2020 May 08;6(2):e19161 [FREE Full text] [doi: 10.2196/19161] [Medline: 32369759]
11. Fisher KA, Barile JP, Guerin RJ, Vanden Esschert KL, Jeffers A, Tian LH, et al. Factors associated with cloth face covering use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):933-937 [FREE Full text] [doi: 10.15585/mmwr.mm6928e3] [Medline: 32673303]
12. Geana MV. Kansans in the middle of the pandemic: risk perception, knowledge, compliance with preventive measures, and primary sources of information about COVID-19. Kans J Med 2020;13:160-164 [FREE Full text] [Medline: 32612749]
13. Tam CC, Qiao S, Li X. Factors associated with decision making on COVID-19 vaccine acceptance among college students in South Carolina. Psychol Health Med 2022 Jan;27(1):150-161 [FREE Full text] [doi: 10.1080/13548506.2021.1983185] [Medline: 34555955]
14. Huang Q, Jackson S, Derakhshan S, Lee L, Pham E, Jackson A, et al. Urban-rural differences in COVID-19 exposures and outcomes in the South: A preliminary analysis of South Carolina. PLoS One 2021;16(2):e0246548 [FREE Full text] [doi: 10.1371/journal.pone.0246548] [Medline: 33534870]
15. Keller LO, Strohschein S, Lia-Hoagberg B, Schaffer MA. Population-based public health interventions: practice-based and evidence-supported. Part I. Public Health Nurs 2004;21(5):453-468. [doi: 10.1111/j.0737-1209.2004.21509.x] [Medline: 15363026]

16. DHEC, UofSC, to Lead Year-long ‘SC STRONG’ Initiative to Evaluate COVID-19 Infection and Immunity Over Time. South Carolina Department of Health and Environmental Control. URL: https://scdhec.gov/news-releases/dhec-uofsc-lead-year-long-sc-strong-initiative-evaluate-covid-19-infection-immunity [accessed 2023-03-23]
17. South Carolina: Profile Data. Census Reporter. URL: http://censusreporter.org/profiles/04000US45-south-carolina/ [accessed 2023-03-23]
18. Nolan MS, Daguise V, Davis M, Duwve JM, Sherrill WW, Heo M, SC STRONG Project Team. SARS-CoV-2 viral incidence, antibody point prevalence, associated population characteristics, and vaccine attitudes, South Carolina, February 2021. Public Health Rep 2022 Mar 10;137(3):457-462 [FREE Full text] [doi: 10.1177/00333549221081128] [Medline: 35264040]
19. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
20. COVID-19 Protocol Library. PhenX Toolkit. URL: https://www.phenxtoolkit.org/covid19/ [accessed 2023-03-23]
21. Fischer CB, Adrien N, Silguero JJ, Hopper JJ, Chowdhury AI, Werler MM. Mask adherence and rate of COVID-19 across the United States. PLoS One 2021;16(4):e0249891 [FREE Full text] [doi: 10.1371/journal.pone.0249891] [Medline: 33852626]
22. Havey NF. Partisan public health: how does political ideology influence support for COVID-19 related misinformation? J Comput Soc Sci 2020;3(2):319-342 [FREE Full text] [doi: 10.1007/s42001-020-00089-2] [Medline: 33163686]
23. Gual-Gonzalez L, McCarter MSJ, Dye-Braumuller K, Self S, Ross CH, Rodriguez-Ramos C, et al. Determinants of COVID-19 vaccinations among a state-wide year-long surveillance initiative in a conservative Southern state. Vaccines (Basel) 2022 Mar 09;10(3):412 [FREE Full text] [doi: 10.3390/vaccines10030412] [Medline: 35335044]
24. Yıldırım M, Güler A. Positivity explains how COVID-19 perceived risk increases death distress and reduces happiness. Pers Individ Dif 2021 Jan 01;168:110347 [FREE Full text] [doi: 10.1016/j.paid.2020.110347] [Medline: 32843780]
25. Number and percentage of households with computer and internet access, by state: 2018. National Center for Education Statistics. URL: https://nces.ed.gov/programs/digest/d19/tables/dt19_702.60.asp [accessed 2023-03-23]

Abbreviations

aOR: adjusted odds ratio
DHEC: Department of Health and Environmental Control
HSSC: Health Sciences South Carolina
SC: South Carolina
SC STRONG: South Carolina Sampling and Testing Representative Outreach for Novel Coronavirus Guidance
VIF: variance inflation factor

Edited by T Sanchez, A Mavragani; submitted 01.11.21; peer-reviewed by S Clipman, A Ren, A Sheon; comments to author 30.04.22; revised version received 30.09.22; accepted 30.01.23; published 19.04.23.

Please cite as:
Kanyangarara M, Daguise V, Gual-Gonzalez L, Litwin AH, Korte J, Ross C, Nolan MS, SC STRONG Project Team
COVID-19 Testing Practices, Preventive Behaviors, and Factors Associated With Test Positivity: Population-Based Statewide Survey Study
JMIR Public Health Surveill 2023;9:e34579
URL: https://publichealth.jmir.org/2023/1/e34579
doi:10.2196/34579
PMID:36720159

Original Paper

Epidemiological Characteristics of Respiratory Syncytial Virus Infection Among Hospitalized Children With Acute Respiratory Tract Infections From 2014 to 2022 in a Hospital in Hubei Province, China: Longitudinal Surveillance Study

Xing-Wen Hu[1*], MBBS; Yiguo Zhou[2*], MSc; Song Yi[3*], MBBS; Wan-Xue Zhang[4], MBBS; Xin-Rui Wang[5], BSc; Juan Du[6], MM; Qing-Bin Lu[2,4,5,6,7], MD

[1]Department of Clinical Laboratory, Maternal and Child Health Hospital of Hubei Province, Wuhan, China
[2]Department of Health Policy and Management, School of Public Health, Peking University, Beijing, China
[3]Department of Medical Genetic Center, Maternal and Child Health Hospital of Hubei Province, Wuhan, China
[4]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[5]Department of Laboratorial Science and Technology & Vaccine Research Center, School of Public Health, Peking University, Beijing, China
[6]Global Center for Infectious Disease and Policy Research & Global Health and Infectious Diseases Group, Peking University, Beijing, China
[7]Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Qing-Bin Lu, MD
Department of Laboratorial Science and Technology & Vaccine Research Center
School of Public Health
Peking University
38th Xueyuan Road, Haidian District
Beijing
China
Phone: 86 10 82805327
Email: qingbinlu@bjmu.edu.cn

Abstract

Background: Longitudinal studies characterizing the epidemic trend of respiratory syncytial virus (RSV) in Hubei Province are scarce.

Objective: We aimed to depict the dynamics of the RSV epidemic among hospitalized children with acute respiratory tract infections (ARTIs) during 2014 to 2022 in the Maternal and Child Health Hospital of Hubei Province and investigate the influence of the 2-child policy and the COVID-19 pandemic on RSV prevalence.

Methods: The medical records and testing results of hospitalized children with ARTI from January 2014 to June 2022 were extracted. Nasopharyngeal samples were tested with direct immunofluorescence assay. Detection rates of RSV were categorized according to the diagnosis of patients: (1) overall, (2) upper respiratory tract infection (URTI), and (3) lower respiratory tract infection (LRTI). Poisson regression models were used to investigate the association between RSV detection rate and age, gender, or diagnosis. The detection rates of RSV before and after the implementation of the universal 2-child policy were compared using a Poisson regression model. Multiple comparisons of RSV detection rates were conducted among 3 stages of the COVID-19 pandemic using chi-square tests. Seasonal autoregressive integrated moving average was performed to predict RSV behaviors from February 2020 to June 2020 under the assumption of a non-COVID-19 scenario.

Results: Among 75,128 hospitalized children with ARTI, 11.1% (8336/75,128) were RSV-positive. Children aged <1 year had higher detection rates than older children (4204/26,498, 15.9% vs 74/5504, 1.3%; *P*<.001), and children with LRTI had higher detection rates than children with URTI (7733/53,145, 14.6% vs 603/21,983, 2.7%; *P*<.001). Among all the children, a clear seasonal pattern of the RSV epidemic was observed before 2021. Most of the highest detection rates were concentrated between December and February. The yearly detection rate of RSV remained at a relatively low level (about 8%) from 2014 to 2017, then increased to 12% and above from 2018. The highest monthly detection rate was in December 2018 (539/1493, 36.1%), and the highest yearly rate was in 2021 (1372/9328, 14.7%). There was a moderate increase in the RSV detection rate after the 2-child policy was implemented (before: 860/10,446, 8.2% vs after: 4920/43,916, 11.2%; *P*<.001). The largest increase, by 5.83%, occurred in children aged <1 year. The RSV epidemic level decreased sharply in the short term after the COVID-19 outbreak (detection rate before: 1600/17,010, 9.4% vs after: 32/1135, 2.8%; *P*<.001). The largest decrease, by 12.0%, occurred in children aged <1 year, but a rebounding epidemic occurred after 2020 (680/5744, 11.8%; *P*<.001).

Conclusions: Children have been experiencing increased prevalence of RSV since 2018 based on surveillance from a hospital in Hubei Province with a large sample size. The 2-child policy might have increased the RSV prevalence, and the COVID-19 epidemic had a temporary inhibitory effect on RSV transmission. Vaccines against RSV are urgently needed.

(JMIR Public Health Surveill 2023;9:e43941) doi:10.2196/43941

KEYWORDS

respiratory syncytial virus; acute respiratory tract infection; epidemiological characteristics; China; COVID-19

Introduction

Respiratory syncytial virus (RSV) is an enveloped, single-stranded, negative-strand RNA virus belonging to the genus *Orthopneumovirus* within the family *Pneumoviridae* [1]. RSV is one of the most common pathogens leading to acute lower respiratory tract infections (LRTIs) among infants [2]. It has been reported that approximately 70% of children are infected in the first year after birth, and almost 100% are infected in their first 2 years [3]. In adults and older children, RSV usually causes mild clinical cold-like signs and symptoms such as congested or runny nose, cough, fever, sore throat, sneezing, and headache, which can disappear in 1 week to 2 weeks [4]. In contrast, RSV can cause much more severe outcomes including pneumonia, bronchiolitis, and even death in young children, especially in infants with underlying diseases or a weak immune system [4]. Recent research estimated that 33.0 million episodes, 3.6 million hospital admissions, 26,300 in-hospital deaths, and 101,400 overall deaths were attributed to RSV-associated acute LRTI in children under 5 years old in 2019 globally [5], and more than 95% of episodes and 97% of deaths happened in low and middle-income countries. RSV-attributable deaths were responsible for 1 in 50 and 1 in 28 deaths in children aged under 5 years old and children aged 28 days to 6 months, respectively. In addition, RSV infection in early life can cause an increased risk of multiple long-term adverse effects including asthma and decreased lung function [6]. RSV infection can also lower immunity and contribute to the risk of COVID-19 and the severity of COVID-19 illnesses [4]. Nevertheless, no drug has been successfully developed to specifically treat RSV-related illnesses to date.

Similar to other respiratory viruses, RSV is spread easily on infected respiratory droplets by air to people's eyes, noses, and mouths if people with RSV cough or sneeze near susceptible individuals. Common measures to prevent RSV infection include washing hands frequently, limiting children's contact with people who have cold-like symptoms, keeping a hygienic environment, and not sharing drinking glasses with others [4]. Dozens of promising vaccine candidates against RSV based on multiple technologies are being developed; however, no candidate has yet completed a phase 3 trial [1]. Some of these vaccine candidates may receive regulatory approval in the near future [1].

Previous studies revealed that the RSV epidemic has obvious seasonal patterns worldwide [7-9]. In the northern hemisphere, it usually occurs from October/November to April/May, while it occurs from May to September in the southern hemisphere [10]. In light of the regionality and seasonality of the epidemic, investigation of conducting continuous surveillance of RSV activities widely in different regions is warranted, especially in metropolitan cities with larger population movement, to enable early warning and medical preparations. During the early stage of the COVID-19 pandemic, marked decreases in non-COVID-19 respiratory infections including RSV that were observed in previous epidemic seasons as multiple mitigation measures were adopted to cope with COVID-19 [11-17]. The strictest nonpharmaceutical interventions (NPIs) including city lockdown, social distancing, nucleic acid testing, wearing masks, and suspension of work and school were implemented in Wuhan to contain the COVID-19 outbreak in the first half of 2020. After the control of the first wave of COVID-19, most NPIs were lifted or relaxed, and nucleic acid testing was regarded as the leading component of measures to control COVID-19. It was a typical example of assessing the impact of interventions against COVID-19 on the epidemic regularity of RSV. In addition, the universal 2-child policy was implemented in Hubei Province in 2016 and lifted the 1-child limit and allowed all couples, regardless of residence in urban or rural areas and regions and ethnic groups, to have 2 children. The potential increased number of newborns in Hubei Province after the policy shift might have influenced the epidemiological characteristics of diseases in young children. Based on data from the largest maternal and child health hospital in Hubei Province of China from 2014 to 2022, we aimed to depict the dynamics of the RSV epidemic in the past 8 years and investigate the influence of the 2-child policy and the COVID-19 pandemic on RSV prevalence.

Methods

Design, Setting, and Population

This study retrospectively reviewed RSV detection results among hospitalized children with acute respiratory tract infection (ARTI) from January 2014 to June 2022 in the Maternal and Child Health Hospital of Hubei Province. The Maternal and Child Health Hospital of Hubei Province is the largest tertiary hospital specializing in maternal and child health service and academic research in Hubei Province, with 1900 inpatient beds and more than 2.77 million outpatient visits in 2019 [18]. Hospitalized children were included if they met the following criteria: (1) diagnosed with ARTI and (2) aged less than 18 years. Specifically, ARTI consisted of upper respiratory tract infection (URTI) and LRTI. LRTI included bronchiolitis and pneumonia [19]. Children were divided into 4 age groups: <1 year, 1-2 years, 3-5 years, and 3-17 years. Basic information about the patients including demographics, symptoms, diagnosis, and examination results was extracted from the hospital case information database.

Samples and Laboratory Testing

Respiratory tract samples were collected within 24 hours after admission using nasopharyngeal swabs, which were stored in centrifuge tubes containing 3 mL saline. RSV was detected through direct immunofluorescence assay with a D^3 Ultra DFA Respiratory Virus Screening and ID Kit (Diagnostic Hybrids Inc). Detailed laboratory operations under the guidance of the kit instructions were as follows: (1) fluid containing samples was blended fully using a vortex oscillator; (2) tubes were centrifuged at 500×g for 10 minutes, then only 150 μL supernatant and sediment were kept; (3) supernatant and sediment were mixed for cell suspension, and 150 μL cell suspension was added to a glass slide, dried, immersed in cold acetone solution for 10 minutes, and dried again; (4) a drop of fluorescein isothiocyanate–labeled monoclonal antibody (about 25 μL) was added to the glass slide to cover the sample and incubated for 30 minutes in a wet box at 37 °C and rinsed and covered by a glass cover; (5) the glass slide containing samples was observed under a fluorescent microscope (Olympus BX53). A sample was regarded as positive for RSV when at least 2 positive cells with apple green fluorescence were observed in the field of view at a 200-fold magnification of the microscope.

Statistical Analysis

The detection rate was calculated as the number of RSV-positive samples divided by the total number of samples, with the 95% CI of the detection rate determined using the Clopper-Pearson method. In all analyses, detection rates are presented in 3 categories according to the diagnosis of patients: overall, URTI, and LRTI. In each category, detection rates were further described according to age group and gender. An adjusted risk ratio (RR) and 95% CI were obtained using a Poisson regression model to show the relative risk of getting RSV at 1 level compared with the reference level in 1 variable, after adjusting for the other 2 variables. For example, the risk of RSV infection in male children versus female children was assessed by adjusting the impact of age and diagnosis. Line plots and bar plots were generated to delineate the detection rate and the number of samples, respectively. A heat map was generated to compare the seasonal distribution of the RSV epidemic across different years, in which a detection rate of 8% was assumed to be an RSV epidemic. Considering the 2-child policy was implemented in 2016 in Hubei Province, the difference in the pooled RSV detection rates between the 2014-2015 and 2017-2019 periods was illustrated by calculating an adjusted RR in the Poisson regression model. These statistical comparisons were considered significant at *P*<.05. According to the COVID-19 status in Hubei Province, 3 stages were determined: stage 1 (prior to the COVID-19 pandemic: February to June in each year from 2017 to 2019), stage 2 (early COVID-19 outbreak: February 2020 to June 2020), and stage 3 (COVID-19 normalization period: February to June in each year from 2021 and 2022). The change in the RSV epidemic in the 2 periods after the COVID-19 outbreak in contrast to the period prior to COVID-19 was investigated. The chi-square test was used in multiple comparisons with statistical significance set at *P*<.025 in the Bonferroni method. Seasonal autoregressive integrated moving average (SARIMA) was used to present the difference between the predicted epidemic level of RSV from February 2020 to June 2020 in a non-COVID-19 scenario and the true detection rate of RSV in Hubei Province. Time series data of URTI cases were not analyzed with a SARIMA model as they were not stationary in the unit root test. Microsoft Excel, Stata 17 (StataCorp), and R 4.1.2 (R Core Team) were used for data processing and visualization.

Ethical Considerations

This study was approved by the clinical research ethics committee of the Maternal and Child Health Hospital of Hubei Province (2022IEC052). Informed consent was exempted due to the retrospective nature of our study. Data used in this study were anonymized and deidentified.

Results

Basic Characteristics of the Overall Sample

A total of 75,128 children from January 2014 to June 2022 met the inclusion criteria of this study (Table 1). These participants primarily consisted of children aged <2 years (50,417/75,128, 67.1%), who were male (44,959/75,128, 59.8%), and who were diagnosed with LRTI (53,145/75,128, 70.7%). Among all the children, 11.1% (8336/75,128) were RSV-positive. The children aged <1 year had the highest RSV detection rate (4204/26,498, 15.9%), followed by children aged 1 year to 2 years (2777/23,919, 11.6%), and rates in both age groups were significantly higher than that for older children (*P*<.001). No significant difference in the RSV detection rate existed between male children and female children (*P*=.54). In addition, the children with LRTI were more likely to be RSV-positive than those with URTI (7773/53,145, 14.6% vs 603/21,983, 2.7%; adjusted RR 4.88, 95% CI 4.50-5.30; *P*<.001). Among 986 children requiring care in the intensive care unit, 256 (26.0%) were positive for RSV.

Table 1. Basic characteristics of the overall sample (n=75,128).

Variables	Positive tests, n	Tests, n	Detection rate, % (95% CI)	Adjusted RR[a] (95% CI)	*P* value
Overall	8336	75,128	11.1 (10.9-11.3)	—[b]	—
Age (years)					
<1	4204	26,498	15.9 (15.4-16.3)	9.84 (7.83-12.36)	<.001[c]
1-2	2777	23,919	11.6 (11.2–12.0)	8.15 (6.48-10.25)	<.001[c]
3-5	1281	19,207	6.7 (6.3-7.0)	4.61 (3.65-5.82)	<.001[c]
6-17	74	5504	1.3 (1.1-1.7)	Reference	
Gender					.54[d]
Male	5123	44,959	11.4 (11.1-11.7)	0.99 (0.95-1.03)	
Female	3213	30,169	10.7 (10.3-11.0)	Reference	
Diagnosis					<.001[e]
URTI[f]	603	21,983	2.7 (2.5-3.0)	Reference	
LRTI[g]	7733	53,145	14.6 (14.3-14.9)	4.88 (4.50-5.30)	

[a] RR: risk ratio.

[b] Not calculable in the overall population.

[c] Adjusted for gender and diagnosis.

[d] Adjusted for age and diagnosis.

[e] Adjusted for age and gender.

[f] URTI: upper respiratory tract infection.

[g] LRTI: lower respiratory tract infection.

Time Trend and Seasonal Pattern of RSV

Among all the children, a clear seasonal pattern of the RSV epidemic was observed before 2021. Most top detection rates were concentrated between December and February, except for March in 2018 (Figure 1). The yearly detection rate of RSV remained at a relatively low level (about 8%) from 2014 to 2017, then increased to 12% and above from 2018. The highest monthly detection rate was in December 2018 (539/1493, 36.1%), and the highest yearly rate was in 2021 (1372/9328, 14.7%). Notably, from May to September in the years before 2021, the RSV detection rate was maintained at an extremely low level, close to zero, while a wave of a moderate epidemic (about 10%) arose in the same period in 2021. From the perspective of age, obvious differences in the RSV detection rate across 3 age groups were seen. Younger children had a higher RSV detection rate. However, RSV detection rates were hard to distinguish between children <1 year old and children 1 year to 2 years old after January 2021. The detection rates between the male and female children almost coincided.

Among the children with URTI, a seasonal peak also existed in December or January (Figure 2). Monthly detection rates of more than 10% were only seen in 2021 and 2022. From 2015 to 2021, the yearly RSV detection rate in URTI cases increased continuously from 0.5% (8/1521) to 5.6% (220/3929). Children aged 1 year to 2 years had a higher positive RSV detection rate than those aged 6 years to 17 years (258/8017, 3.2% vs 26/2136, 1.2%, *P*<.001; Table S1 in Multimedia Appendix 1). No significant difference in RSV detection rate was found between genders (*P*=.37). Among the children with LRTI, the trend and seasonal pattern of RSV detection rate were similar to that in all the children (Figure 3). The monthly RSV detection rate peaked in December 2018 (521/1294, 40.3%), and the yearly detection rate peaked in 2021 (1152/5399, 21.3%). Children aged <1 year (4055//20,849, 19.5%) or 1 year to 2 years (2519/15,902, 15.8%) had obviously higher RSV-positive rates than the oldest age group (48/3368, 1.4%; all *P*<.001; Table S2 in Multimedia Appendix 1). No difference was observed between genders (*P*=.57).

The heat maps show the beginning and end weeks of the RSV epidemic in each year (Figure 4 and Multimedia Appendix 2). A similar distribution of the RSV epidemic was observed between all the children and those with LRTI. The earliest beginning of the epidemic (about the 40th week) happened in 2017 and 2021, and the latest end of the epidemic (the 15th and 20th weeks) happened in 2018 and 2021, respectively. Particularly in 2021, RSV had the longest epidemic period lasting 40 weeks (about 77% of the whole year). In contrast, the shortest epidemic period of only 15 weeks occurred in 2020. Things were quite different in children with URTI. The RSV epidemic lasted a short number of weeks (<5 weeks) in the years before 2020. However, a long RSV epidemic period occurred in 2021 (16 weeks). Heat maps of the 2 younger age groups and the 2 genders did not provide different information from that of all the children. A much shorter epidemic period was seen for children aged 3 years to 5 years or 6 years to 17 years.

Figure 1. Monthly distribution of tests for (bars) and detection rate of (lines) respiratory syncytial virus from 75,128 hospitalized children with acute respiratory tract infection during 2014 to 2022: (A) overall sample, (B) by age group, (C) by gender.

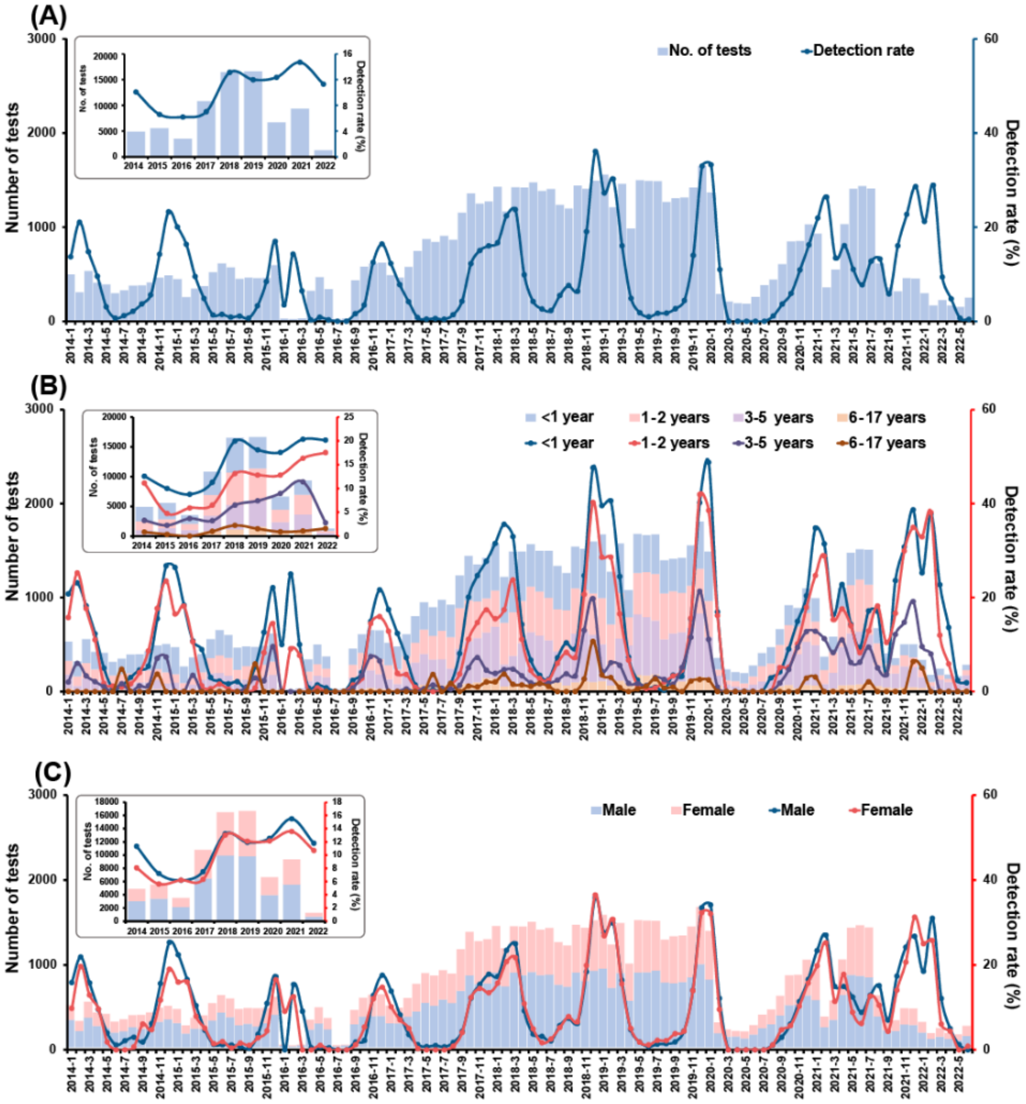


Figure 2. Monthly distribution of tests for (bars) and detection rates of (lines) respiratory syncytial virus from 21,983 hospitalized children with acute upper respiratory tract infection during 2014 to 2022: (A) overall sample, (B) by age group, (C) by gender.

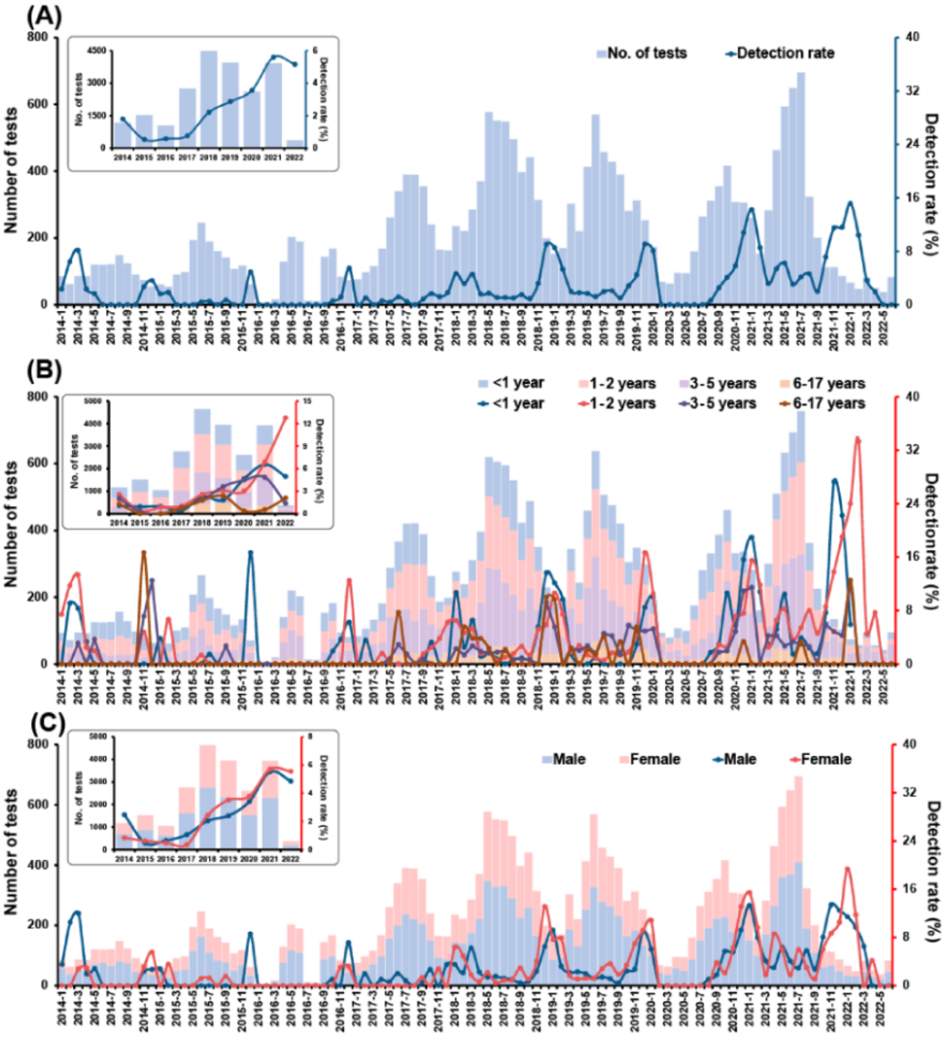


Figure 3. Monthly distribution of tests for (bars) and detection rate of (lines) respiratory syncytial virus from 53,145 hospitalized children with acute lower respiratory tract infection during 2014 to 2022: (A) overall sample, (B) by age group, (C) by gender.

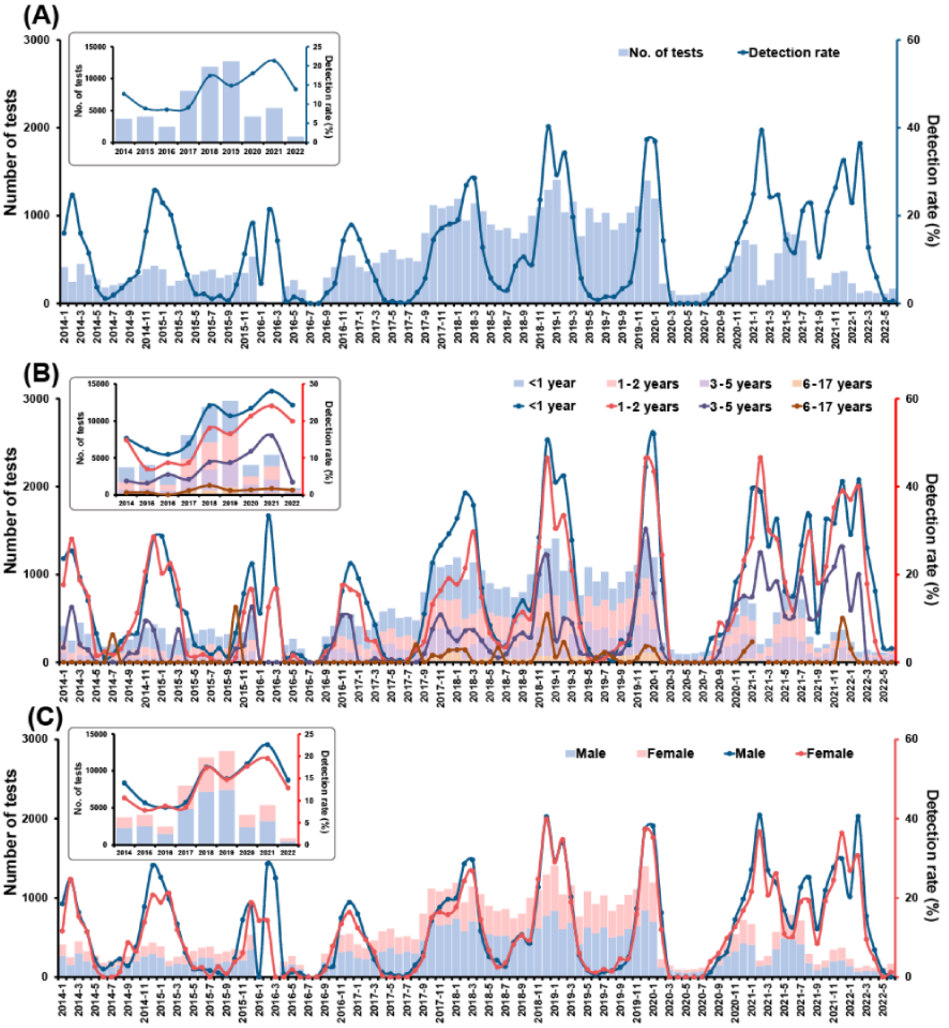
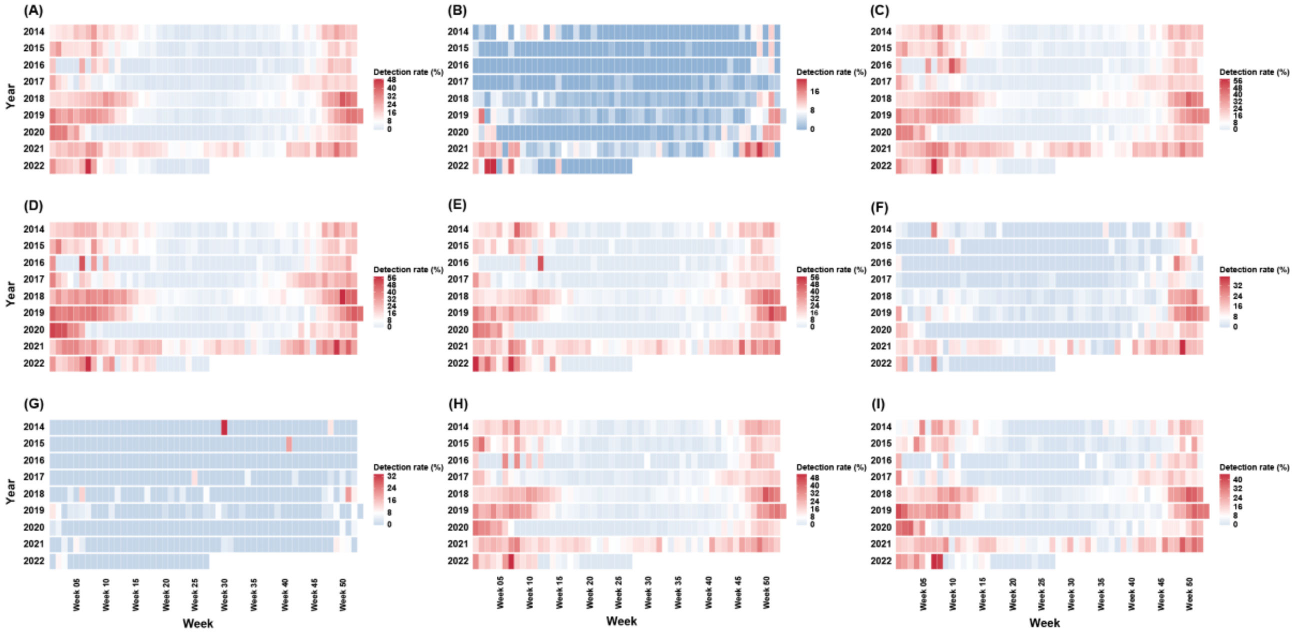


Figure 4. Heat maps of the respiratory syncytial virus detection rate by week during 2014 to 2022: (A) overall sample, (B) in those with upper respiratory tract infection, (C) in those with lower respiratory tract infection, (D) in those <1 year old, (E) in those 1 year to 2 years old, (F) in those 3 years to 5 years old, (G) in those 6 years to 17 years old, (H) in male children, (I) in female children.



Impact of the 2-Child Policy on the RSV Epidemic

Among all the children, there was a moderate increase in the RSV detection rate in most subgroups, except for children aged 6 years to 17 years, after the 2-child policy was implemented (all *P*<.001; Table 2). The largest increase, by 5.83%, occurred in children aged <1 year. Children aged 6 years to 17 years had the highest risk of being infected after the implementation of the policy (adjusted RR 2.80, 95% CI 0.88-8.90). Similar results were found with children with LRTI. Among the children with URTI, the increased RSV detection rates in only the entire group of children and female children were significant (*P*=.001 and *P*=.002, respectively). Though female children had a high risk of RSV infection after the implementation of the 2-child policy (adjusted RR 3.12, 95% CI 1.52-6.39), the prevalence of RSV increased slightly from 0.7% (8/1133) to 2.3% (107/4658).

Table 2. Comparison of the detection rate of respiratory syncytial virus before and after the implementation of the universal 2-child policy.

Group	Before (2014-2015)			After (2017-2019)			Adjusted RR[a] (95% CI)	*P* value
	Positive tests	Number of tests	Detection rate, % (95% CI)	Positive tests	Number of tests	Detection rate, % (95% CI)		
Total population	860	10,446	8.23 (7.71-8.78)	4920	43,916	11.20 (10.91-11.50)	1.52 (1.42-1.63)	<.001[b]
Total population								
Age (years)								
<1	573	5105	11.22 (10.37-12.12)	2554	14,977	17.05 (16.45-17.66)	1.50 (1.38-1.63)	<.001[c]
1-2	234	2984	7.84 (6.9-8.87)	1626	14,248	11.41 (10.89-11.95)	1.41 (1.24-1.60)	<.001[c]
3-5	50	1850	2.70 (2.05-3.54)	683	11,224	6.09 (5.66-6.55)	2.26 (1.70-2.99)	<.001[c]
6-17	3	507	0.59 (0.20-1.72)	57	3467	1.64 (1.27-2.12)	2.80 (0.88-8.90)	.08[c]
Gender								
Male	589	6436	9.15 (8.46-9.88)	2966	26,183	11.33 (10.95-11.72)	1.41 (1.30-1.53)	<.001[d]
Female	271	4010	6.76 (6-7.58)	1954	17,733	11.02 (10.56-11.49)	1.77 (1.57-1.99)	<.001[d]
URTI[e]	29	2691	1.08 (0.72-1.54)	236	11,340	2.08 (1.83-2.36)	1.88 (1.28-2.76)	.001[f]
URTI								
Age (years)								
<1	10	1001	1.00 (0.48-1.83)	44	2713	1.62 (1.18-2.17)	1.62 (0.82-3.21)	.17[g]
1-2	13	956	1.36 (0.73-2.31)	99	4164	2.38 (1.94-2.89)	1.75 (0.98-3.11)	.06[g]
3-5	5	536	0.93 (0.40-2.16)	72	3282	2.19 (1.74-2.75)	2.36 (0.96-5.81)	.06[g]
6-17	1	198	0.51 (0.09-2.81)	21	1181	1.78 (1.17-2.71)	3.53 (0.48-25.80)	.22[g]
Gender								
Male	21	1558	1.35 (0.84–2.05)	129	6682	1.93 (1.61-2.29)	1.41 (0.89-2.23)	.15[h]
Female	8	1133	0.71 (0.31–1.39)	107	4658	2.30 (1.89-2.77)	3.12 (1.52-6.39)	.002[h]
LRTI[i]	831	7755	10.72 (10.04-11.43)	4684	32,576	14.38 (14-14.76)	1.51 (1.41-1.62)	<.001[f]
LRTI								
Age (years)								
<1	563	4104	13.72 (12.68-14.81)	2510	12,264	20.47 (19.76-21.19)	1.49 (1.37-1.63)	<.001[g]
1-2	221	2028	10.90 (9.57-12.34)	1527	10,084	15.14 (14.45-15.86)	1.39 (1.22-1.59)	<.001[g]
3-5	45	1314	3.42 (2.57-4.55)	611	7942	7.69 (7.12-8.30)	2.24 (1.67-3.02)	<.001[g]
6-17	2	309	0.65 (0.18-2.33)	36	2286	1.57 (1.14-2.17)	2.43 (0.59-10.04)	.22[g]
Gender								
Male	568	4878	11.64 (10.76-12.58)	2837	19,501	14.55 (14.06-15.05)	1.41 (1.30-1.53)	<.001[h]
Female	263	2877	9.14 (8.11-10.25)	1847	13,075	14.13 (13.53-14.74)	1.72 (1.52-1.95)	<.001[h]

[a] RR: rate ratio.

[b] Adjusted for age, gender, and diagnosis.

[c] Adjusted for gender and diagnosis.

[d] Adjusted for age and diagnosis.

[e] URTI: upper respiratory tract infection.

[f] Adjusted for age and gender.

[g] Adjusted for gender.

[h] Adjusted for age.

[i] LRTI: lower respiratory tract infection.

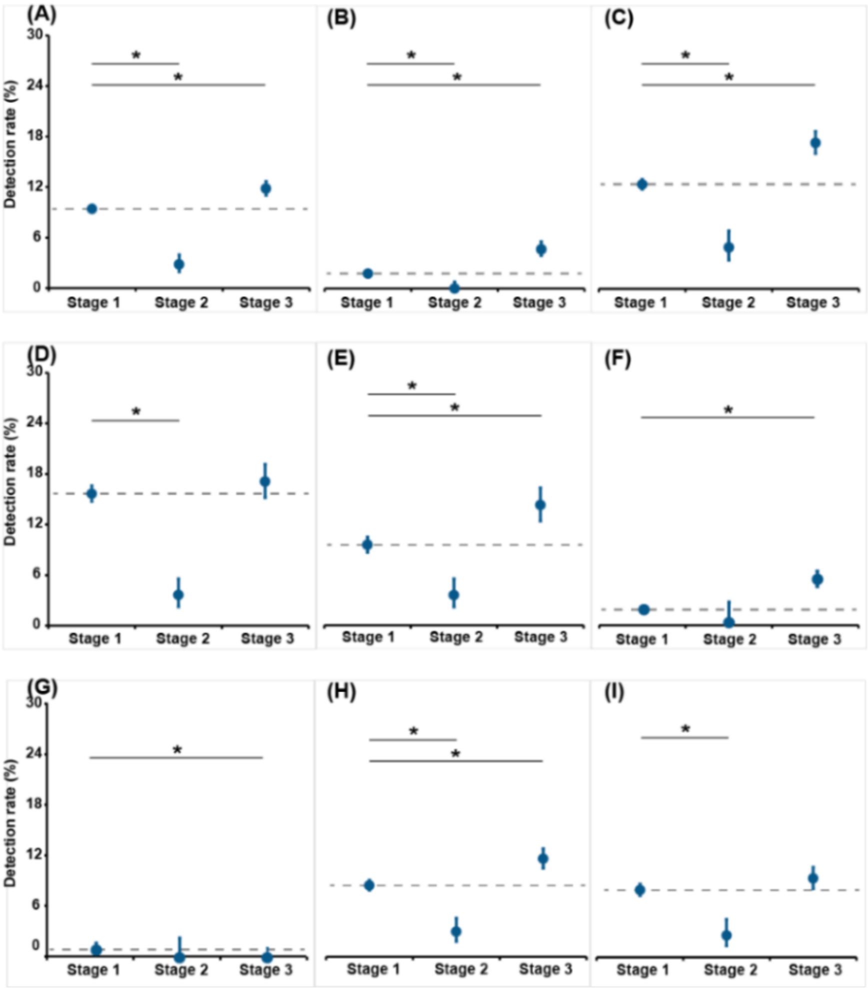
Impact of the COVID-19 Outbreak on the RSV Epidemic

Among all the children, compared with the pooled levels from February to June in each year from 2017 to 2019, all the stratifications except for the 2 older groups achieved significant reductions in the RSV detection rate during the earliest months of the COVID-19 pandemic (February 2020 to June 2020; Figure 5, Multimedia Appendix 3, and Table S3 in Multimedia Appendix 1). Of these, the sharpest decline, by 12.0%, occurred in children aged <1 year. Among the children with URTI, only the overall sample and male children showed declines of statistical significance. Among the children with LRTI, only the children aged 1 year to 2 years, 3 years to 5 years, or 6 years to 17 years did not have significantly decreased RSV detection rates. The largest decrease (13.8%) was observed in children aged <1 year.

Among all the children, compared with the pooled levels during February to June in each of 2017 to 2019, the overall RSV detection rate increased from 9.4% (1600/17,010) to 11.8% (680/5744) during the same periods in 2021 and 2022 (Figure 5, Multimedia Appendix 3, and Table S3 in Multimedia Appendix 1). All the stratifications, except for children aged <1 year and female children, had significant increases in the RSV detection rate. Of these, children aged 1 year to 2 years had the largest increase, by 4.7%. Among the children with URTI, all the stratifications, except for children aged 6 years to 17 years, had significant increases. The children aged 1 year to 2 years had the largest increase, by 4.3%. Among the children with LRTI, all the stratifications, except for children aged 6 years to 17 years, had significant increases, with the largest increase in children aged 1 year to 2 years, by 7.7%.

Multimedia Appendix 4 compares the observed rates with the predicted rates of RSV detection from February 2020 to June 2020 based on the SARIMA model. Among all the children, the forecasted RSV detection rates were much higher than the actual values. The largest decrease, by 28.5%, in the RSV detection rate happened in March 2020, denoting the most significant impact of the COVID-19 outbreak. In all the stratified analyses, the observed rates were significantly lower than the lower limit of the 95% CI of the forecasted rates representing the counterfactual scenario without COVID-19.

Figure 5. Comparison of the respiratory syncytial virus detection rate among periods according to COVID-19 epidemic status (stage 1: February to June in each of 2017, 2018, and 2019; stage 2: February 2020 to June 2020; stage 3: February to June in each of 2021 and 2022): (A) overall sample, (B) those with upper respiratory tract infection, (C) those with lower respiratory tract infection, (D) in those <1 year old, (E) in those 1 year to 2 years old, (F) in those 3 years to 5 years old, (G) in those 6 years to 17 years old, (H) in male children, (I) in female children. *P<.05.



Discussion

In this study, more than 8 years of RSV surveillance including 75,128 hospitalized children with ARTI was carried out to better understand the epidemiological characteristics of RSV based on a sentinel hospital in Hubei Province. Yearly RSV epidemic peaks concentrated in the period from December to February were observed before 2021. The RSV detection rate in 2018 and later was significantly higher than that in 2017 and before. RSV was more likely to be detected among younger children and children with an LRTI. The 2-child policy implemented in 2016 was related to an increased RSV detection rate in the following 3 years. The RSV epidemic level decreased sharply from February 2020 to June 2020, but a rebounding epidemic peak in the subsequent winters and increased epidemic levels in the summer of 2021 occurred. The results about the influence of the 2-child policy and COVID-19–related NPIs on the RSV epidemic were obtained using lengthy surveillance data from the Maternal and Child Health Hospital of Hubei Province and will need more multicenter data to further confirm.

RSV was the most frequently detected in the winter and the least frequently detected in the summer in our study, which was similar to the seasonal pattern in some temperate countries [20]. The area in which the Maternal and Child Health Hospital of Hubei Province is located is in a subtropical monsoon climate, with a high temperature in the summer and a low temperature and high relative humidity in the winter. A low temperature is likely to cause vasoconstriction of the nasal mucosa and weakens the immunity of the respiratory tract against viruses [21,22]. In winter, more frequent gathering indoors and a lack of air circulation can also contribute to the transmission of RSV. In addition, relative humidity plays an essential role in respiratory virus transmission. It has been reported that RSV epidemics tend to occur in the rainy season. Wet conditions may increase the number of viruses deposited on surfaces, promote the survival of viruses in droplets on surfaces, and, in turn, encourage the contact transmission of RSV [20,23]. Therefore, prior to rainy or cold weather, knowledge about prevention measures against RSV such as hand hygiene, environmental cleaning, and indoor ventilation should be conveyed to the public, in particular to families with young children.

Limited literature has focused on the longitudinal trends of the RSV epidemic over the past decade in Hubei Province. Most studies devoted to the epidemiological characteristics of RSV in China analyzed the positive rate of RSV among ARTI cases collected during 1 or more years in a cross-sectional manner. A study including pediatric and adult patients with ARTI from 2011 to 2016 in Guangdong Province reported a less clear and unstable seasonal pattern of RSV with epidemics almost year-round [24]. The highest detection rates of RSV usually appeared from March to May and ranged from 15% to 30% in that study. The highest detection rates in that period in our study were similar to those reported in other studies, but RSV broke out earlier. Research conducted in Beijing reported that the peak months of the RSV epidemic were November to January [25]. The results of these studies suggest that earlier RSV epidemics happen with increased latitude and decreased annual average temperatures. Another surveillance study that enrolled 19,898 hospitalized children with ARTI showed an annual RSV detection rate in the range of 5.4% to 7.0% from 2014 to 2018 in China, which was significantly lower than that in our study [26]. Considering the limited sample size and representativeness in that study, we were unable to conclude that the RSV epidemic in our study was more severe than the average countrywide level. More efforts should be taken to improve the surveillance and analyses of the RSV epidemic in different areas in China to formulate tailored prevention measures. Additionally, RSV rates and time trends are totally different across countries. Our study documented an increased RSV detection rate starting in 2018 according to hospital data. In contrast, the RSV infection rate in children with LRTI remained stable, at around 50%, during 2014 to 2018 in the València region in Spain [27,28]. Compared with our results, Germany witnessed similar RSV positivity rates for children aged 0 to 4 years from 2011 to 2021, but the yearly epidemic level was more stable [29]. RSV positivity rates of children fluctuated from 7.5% to 23% during 2007 to 2016 in the Philippines [30]. However, RSV in southwest Finland (2008-2012) and Russia (2013-2018) was rarely detected, with positivity rates of 6% and 4.4%, respectively [31,32]. In general, the area served by the hospital in our study is facing higher pressure from RSV infections in recent years.

The significant increase in the RSV-positive rate since 2018 among young children in our study might be caused by multiple reasons. A potential factor associated with the RSV epidemic among young children was the implementation of the 2-child policy in 2016 in Hubei Province. Since the beginning of the policy, the total number of births rose moderately, while the number of second child births significantly increased [33]. On one hand, the increased number of susceptible infants would contribute to a higher RSV epidemic size. This is in line with a study revealing that there was a sharp increase in inpatient admission rates for infants after China's universal 2-child policy [34]. In contrast, a modeling study showed that the 1-child policy would lower the average annual influenza attack rate in China [35]. On the other hand, the younger, second child in a family could be infected by other family members, especially their elder siblings. A study in Shanghai elucidated that elder siblings might be the main route of hand, foot, and mouth disease transmission to neonates [36]. Keeping young children from people with flu-like symptoms is a key measure to protect them from RSV infection when a vaccine against RSV is unavailable. In addition, we noted that, although children aged 3 years or older were not directly influenced by the change of the policy after 2017, the RSV epidemic in this age group also increased. There might be 2 reasons. First, RSV is a communicable virus, and the increased scope of RSV infection in infants can also lead to more infections in elder children than before. Second, in addition to the impact of the policy shift, other factors including climate change and more spread of viruses from other provinces could cause an increase in RSV infection in elder children. However, the impact of other factors needs further investigation.

It is universally acknowledged that the COVID-19 pandemic effectively decreased the occurrence of an RSV epidemic due to various measures against the transmission of respiratory viruses [17,37-43]. This phenomenon was even more obvious in the several months under the strictest interventions after the COVID-19 outbreak in our study. However, RSV rebounded to a higher level in the years after strict NPIs were loosened. A delayed outbreak or a severe rebound of RSV was also found in multiple countries [44-48]. NPIs in the context of COVID-19 could prevent the transmission of respiratory viruses but also lead to a greater number of susceptible persons, which might have resulted in larger outbreaks once the transmission routes were reopened [46]. Children aged 1 year to 2 years had a larger rebound than the other age groups in our study. That might be due to the relatively weak immune system and the expanded social activities in this age group. Additionally, the continuous increase in the RSV detection rate among young children with URTI should be noted. Though URTI symptoms are mild, a URTI caused by RSV could progress to LRTI, which has a much higher disease burden [49].

There were some limitations in this study. First, though we used data from a large sample size from the largest maternal and child health hospital in Hubei Province, the conclusions drawn from only 1 hospital might not represent the overall situation at the provincial level. More surveillance studies of RSV from multiple sites are needed to better clarify the epidemiological characteristics of RSV in Hubei Province and perform comparisons among different cities. Genotyping of RSV was not performed; thus, we were unable to elucidate the detailed epidemiological characteristics of different genotypes. Moreover, missing medical records existing for several months in 2016 potentially caused an unprecise estimate of indicators of interest. Nevertheless, despite the limited sample size in February 2016, an epidemic peak was seen, which reflected the stable seasonal pattern of RSV. Only hospitalized children were included in our study, and our findings might overestimate the general level of the RSV epidemic among children. We could only identify the longitudinal trend of RSV and compare the relative levels across different years and different populations. Additionally, preterm birth is a known risk factor for RSV infection [50]. However, due to the lack of this variable in this study, we were unable to further confirm this association or manage the potential confounding effect of preterm birth in the models.

In summary, RSV was frequently detected among hospitalized children with ARTI based on surveillance of data from a large sample size at a hospital in Hubei Province. RSV positivity was more likely to happen in young children, cases diagnosed as LRTI, and wet and low-temperature seasons. Early warning of the risk of RSV infection among high-risk populations and seasons should be considered in local public health practices. A sudden rise in the RSV epidemic after 2018 might be affected by the 2-child policy implementation in 2016. Observed in our study, the RSV epidemic significantly decreased in the short term after the COVID-19 outbreak, followed by a sharp rebound in the years that followed. This implies that healthy lifestyle habits to prevent respiratory viruses should be maintained even though the COVID-19 situation has eased. A specific vaccine against RSV is urgently needed to lower the disease burden attributable to RSV.

Acknowledgments

This work was supported by the Joint Research Fund for Beijing Natural Science Foundation and Haidian Original Innovation (L202007) and Fundamental Research Funds for the Central Universities and Peking University Health Science Center (BMU2021YJ041), Peking University Medicine Fund of Fostering Young Scholars' Scientific and Technological Innovation (BMU2021PY005).

Data Availability

The data sets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

QBL and XWH conceived and designed the experiments, analyzed the data, reviewed drafts of the paper, and approved the final draft. SY and YZ collected and analyzed the data, prepared the figures and tables, authored drafts of the paper, and approved the final draft. WXZ, XRW, and JD collected the data and reviewed the draft. All authors approved the final draft and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File , 634 KB - publichealth_v9i1e43941_app1.docx]

Multimedia Appendix 2
Heat maps of respiratory syncytial virus detection rates by week from 2014 to 2022.
[DOCX File , 285 KB - publichealth_v9i1e43941_app2.docx]

Multimedia Appendix 3
Comparison of the respiratory syncytial virus detection rate among periods according to COVID-19 epidemic status (stage 1: February to June in each of 2017, 2018, and 2019; stage 2: February 2020 to June 2020; stage 3: February to June in each of 2021 and 2022).
[DOCX File , 136 KB - publichealth_v9i1e43941_app3.docx]

Multimedia Appendix 4
Comparison of the forecasted respiratory syncytial virus detection rate with the actual status from February 2020 to June 2020 using the seasonal autoregressive integrated moving average.
[DOCX File , 185 KB - publichealth_v9i1e43941_app4.docx]

References

1. Respiratory Syncytial Virus (RSV) disease. World Health Organization. URL: https://www.who.int/teams/health-product-policy-and-standards/standards-and-specifications/vaccine-standardization/respiratory-syncytial-virus-disease [accessed 2023-04-08]
2. Borchers AT, Chang C, Gershwin ME, Gershwin LJ. Respiratory syncytial virus--a comprehensive review. Clin Rev Allergy Immunol 2013 Dec;45(3):331-379 [FREE Full text] [doi: 10.1007/s12016-013-8368-9] [Medline: 23575961]
3. Hu M, Bogoyevitch MA, Jans DA. Impact of respiratory syncytial virus infection on host functions: implications for antiviral strategies. Physiol Rev 2020 Oct 01;100(4):1527-1594 [FREE Full text] [doi: 10.1152/physrev.00030.2019] [Medline: 32216549]
4. Respiratory syncytial virus (RSV). Mayo Clinic. URL: https://www.mayoclinic.org/diseases-conditions/respiratory-syncytial-virus/symptoms-causes/syc-20353098 [accessed 2023-04-08]
5. Li Y, Wang X, Blau DM, Caballero MT, Feikin DR, Gill CJ, Respiratory Virus Global Epidemiology Network, RESCEU investigators. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. Lancet 2022 May 28;399(10340):2047-2064 [FREE Full text] [doi: 10.1016/S0140-6736(22)00478-0] [Medline: 35598608]
6. Bacharier LB, Cohen R, Schweiger T, Yin-Declue H, Christie C, Zheng J, et al. Determinants of asthma after severe respiratory syncytial virus bronchiolitis. J Allergy Clin Immunol 2012 Jul;130(1):91-100.e3 [FREE Full text] [doi: 10.1016/j.jaci.2012.02.010] [Medline: 22444510]
7. Hall CB, Weinberg GA, Iwane MK, Blumkin AK, Edwards KM, Staat MA, et al. The burden of respiratory syncytial virus infection in young children. N Engl J Med 2009 Feb 05;360(6):588-598 [FREE Full text] [doi: 10.1056/NEJMoa0804877] [Medline: 19196675]
8. Rose EB, Wheatley A, Langley G, Gerber S, Haynes A. Respiratory syncytial virus seasonality - United States, 2014-2017. MMWR Morb Mortal Wkly Rep 2018 Jan 19;67(2):71-76 [FREE Full text] [doi: 10.15585/mmwr.mm6702a4] [Medline: 29346336]
9. Obando-Pacheco P, Justicia-Grande AJ, Rivero-Calle I, Rodríguez-Tenreiro C, Sly P, Ramilo O, et al. Respiratory syncytial virus seasonality: a global overview. J Infect Dis 2018 Apr 11;217(9):1356-1364 [FREE Full text] [doi: 10.1093/infdis/jiy056] [Medline: 29390105]
10. Respiratory syncytial virus infection: Clinical features and diagnosis. UpToDate. URL: https://www.uptodate.com/contents/respiratory-syncytial-virus-infection-clinical-features-and-diagnosis [accessed 2023-04-08]
11. Agha R, Avner JR. Delayed seasonal RSV surge observed during the COVID-19 pandemic. Pediatrics 2021 Sep;148(3):1. [doi: 10.1542/peds.2021-052089] [Medline: 34108234]
12. Van Brusselen D, De Troeyer K, Ter Haar E, Vander Auwera A, Poschet K, Van Nuijs S, et al. Bronchiolitis in COVID-19 times: a nearly absent disease? Eur J Pediatr 2021 Jun;180(6):1969-1973 [FREE Full text] [doi: 10.1007/s00431-021-03968-6] [Medline: 33517482]
13. Yeoh DK, Foley DA, Minney-Smith CA, Martin AC, Mace AO, Sikazwe CT, et al. Impact of coronavirus disease 2019 public health measures on detections of influenza and respiratory syncytial virus in children during the 2020 Australian winter. Clin Infect Dis 2021 Jun 15;72(12):2199-2202 [FREE Full text] [doi: 10.1093/cid/ciaa1475] [Medline: 32986804]
14. Friedrich F, Ongaratto R, Scotta MC, Veras TN, Stein RT, Lumertz MS, et al. Early impact of social distancing in response to coronavirus disease 2019 on hospitalizations for acute bronchiolitis in infants in Brazil. Clin Infect Dis 2021 Jun 15;72(12):2071-2075 [FREE Full text] [doi: 10.1093/cid/ciaa1458] [Medline: 32986818]
15. Nolen LD, Seeman S, Bruden D, Klejka J, Desnoyers C, Tiesinga J, et al. Impact of social distancing and travel restrictions on non-coronavirus disease 2019 (non-COVID-19) respiratory hospital admissions in young children in rural Alaska. Clin Infect Dis 2021 Jun 15;72(12):2196-2198 [FREE Full text] [doi: 10.1093/cid/ciaa1328] [Medline: 32888007]
16. Haddadin Z, Schuster JE, Spieker AJ, Rahman H, Blozinski A, Stewart L, et al. Acute respiratory illnesses in children in the SARS-CoV-2 pandemic: prospective multicenter study. Pediatrics 2021 Aug;148(2):1 [FREE Full text] [doi: 10.1542/peds.2021-051462] [Medline: 33986150]

17. Olsen SJ, Winn AK, Budd AP, Prill MM, Steel J, Midgley CM, et al. Changes in influenza and other respiratory virus activity during the COVID-19 pandemic - United States, 2020-2021. MMWR Morb Mortal Wkly Rep 2021 Jul 23;70(29):1013-1019 [FREE Full text] [doi: 10.15585/mmwr.mm7029a1] [Medline: 34292924]
18. Hospital profile. Hubei Provincial Maternal and Child Health Hospital. URL: https://www.hbfy.com/yygk/cont/about-1.html [accessed 2023-04-08]
19. Acute respiratory infections (ARI): among the leading causes of death in the world. FIFARMA. URL: https://fifarma.org/en/acute-respiratory-infections-ari-among-the-leading-causes-of-death-in-the-world/ [accessed 2023-04-08]
20. Staadegaard L, Caini S, Wangchuk S, Thapa B, de Almeida WAF, de Carvalho FC, et al. Defining the seasonality of respiratory syncytial virus around the world: National and subnational surveillance data from 12 countries. Influenza Other Respir Viruses 2021 Nov;15(6):732-741 [FREE Full text] [doi: 10.1111/irv.12885] [Medline: 34255934]
21. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of respiratory viral infections. Annu Rev Virol 2020 Sep 29;7(1):83-101 [FREE Full text] [doi: 10.1146/annurev-virology-012420-022445] [Medline: 32196426]
22. Pica N, Bouvier NM. Ambient temperature and respiratory virus infection. Pediatr Infect Dis J 2014 Mar;33(3):311-313. [doi: 10.1097/INF.0000000000000235] [Medline: 24378933]
23. Liu W, Chen D, Tan W, Xu D, Qiu S, Zeng Z, et al. Epidemiology and clinical presentations of respiratory syncytial virus subgroups A and B detected with multiplex real-time PCR. PLoS One 2016;11(10):e0165108 [FREE Full text] [doi: 10.1371/journal.pone.0165108] [Medline: 27764220]
24. Luo H, Huang X, Zhong H, Ye C, Tan X, Zhou K, et al. Epidemiological characteristics and phylogenic analysis of human respiratory syncytial virus in patients with respiratory infections during 2011-2016 in southern China. Int J Infect Dis 2020 Jan;90:5-17 [FREE Full text] [doi: 10.1016/j.ijid.2019.10.009] [Medline: 31634614]
25. Luo M, Gong C, Zhang Y, Wang X, Liu Y, Luo Q, et al. Comparison of infections with respiratory syncytial virus between children and adults: a multicenter surveillance from 2015 to 2019 in Beijing, China. Eur J Clin Microbiol Infect Dis 2022 Dec;41(12):1387-1397 [FREE Full text] [doi: 10.1007/s10096-022-04492-7] [Medline: 36197575]
26. Zhang H, Zhu A, Gao GF, Li Z. Epidemiological and Clinical Characteristics of Respiratory Syncytial Virus Infections in Children Aged <5 Years in China, from 2014–2018. Vaccines (Basel) 2022 May 20;10(5):1 [FREE Full text] [doi: 10.3390/vaccines10050810] [Medline: 35632566]
27. Mira-Iglesias A, Demont C, López-Labrador FX, Mengual-Chuliá B, García-Rubio J, Carballido-Fernández M, Valencia Hospital Network for the Study of Influenzaother Respiratory Viruses. Role of age and birth month in infants hospitalized with RSV-confirmed disease in the Valencia Region, Spain. Influenza Other Respir Viruses 2022 Mar;16(2):328-339 [FREE Full text] [doi: 10.1111/irv.12937] [Medline: 34821055]
28. Heppe-Montero M, Walter S, Hernández-Barrera V, Gil-Prieto R, Gil-de-Miguel Á. Burden of respiratory syncytial virus-associated lower respiratory infections in children in Spain from 2012 to 2018. BMC Infect Dis 2022 Mar 31;22(1):315 [FREE Full text] [doi: 10.1186/s12879-022-07261-1] [Medline: 35361139]
29. Cai W, Dürrwald R, Biere B, Schweiger B, Haas W, Wolff T, et al. Determination of respiratory syncytial virus epidemic seasons by using 95% confidence interval of positivity rates, 2011-2021, Germany. Influenza Other Respir Viruses 2022 Sep;16(5):854-857 [FREE Full text] [doi: 10.1111/irv.12996] [Medline: 35485999]
30. Calaor-Morin J, Arguelles VL, Foronda JL, Tan A, Lagamayo E, Dapat C, et al. Genotyping of respiratory syncytial virus among influenza-like illness and severe acute respiratory infection cases of children in the Philippines from 2006 to 2016. Influenza Other Respir Viruses 2022 Sep;16(5):942-951 [FREE Full text] [doi: 10.1111/irv.12986] [Medline: 35582932]
31. Toivonen L, Karppinen S, Schuez-Havupalo L, Teros-Jaakkola T, Mertsola J, Waris M, et al. Respiratory syncytial virus infections in children 0-24 months of age in the community. J Infect 2020 Jan;80(1):69-75. [doi: 10.1016/j.jinf.2019.09.002] [Medline: 31521741]
32. Caini S, Stolyarov K, Sominina A, Smorodintseva E, Staadegaard L, Paget J, et al. A comparative analysis of the epidemiology of influenza and respiratory syncytial virus in Russia, 2013/14 to 2018/19. J Glob Health 2022 Feb 05;12:04009 [FREE Full text] [doi: 10.7189/jogh.12.04009] [Medline: 35136600]
33. Xiang H, Zhang Y, Li B. The influence of the overall two-child policy on the urbanization of Hubei Province. Population and Society 2018;34(04):99-110. [doi: 10.14132/j.2095-7963.2018.04.010]
34. Shen M, Liang X, Wu Y, Fang S. Estimates of inpatient admission rates and associated costs for infants before and after China's universal two-child policy. BMC Health Serv Res 2022 Feb 08;22(1):163 [FREE Full text] [doi: 10.1186/s12913-022-07571-9] [Medline: 35135539]
35. Liu F, Enanoria WTA, Ray KJ, Coffee MP, Gordon A, Aragón TJ, et al. Effect of the one-child policy on influenza transmission in China: a stochastic transmission model. PLoS One 2014;9(2):e84961 [FREE Full text] [doi: 10.1371/journal.pone.0084961] [Medline: 24516519]
36. Xu S, Li H, Qiao P, Xu G, Zhao D, Lin X, et al. Neonatal hand, foot, and mouth disease due to coxsackievirus A6 in Shanghai. BMC Pediatr 2020 Aug 03;20(1):364 [FREE Full text] [doi: 10.1186/s12887-020-02262-y] [Medline: 32741368]
37. Torres-Fernandez D, Casellas A, Mellado MJ, Calvo C, Bassat Q. Acute bronchiolitis and respiratory syncytial virus seasonal transmission during the COVID-19 pandemic in Spain: A national perspective from the pediatric Spanish Society (AEP). J Clin Virol 2021 Dec;145:105027 [FREE Full text] [doi: 10.1016/j.jcv.2021.105027] [Medline: 34781241]

38. Doroshenko A, Lee N, MacDonald C, Zelyas N, Asadi L, Kanji JN. Decline of influenza and respiratory viruses with COVID-19 public health measures: Alberta, Canada. Mayo Clin Proc 2021 Dec;96(12):3042-3052 [FREE Full text] [doi: 10.1016/j.mayocp.2021.09.004] [Medline: 34863395]
39. Wagatsuma K, Koolhof IS, Shobugawa Y, Saito R. Decreased human respiratory syncytial virus activity during the COVID-19 pandemic in Japan: an ecological time-series analysis. BMC Infect Dis 2021 Aug 03;21(1):734 [FREE Full text] [doi: 10.1186/s12879-021-06461-5] [Medline: 34344351]
40. Di Mattia G, Nenna R, Mancino E, Rizzo V, Pierangeli A, Villani A, et al. During the COVID-19 pandemic where has respiratory syncytial virus gone? Pediatr Pulmonol 2021 Oct;56(10):3106-3109 [FREE Full text] [doi: 10.1002/ppul.25582] [Medline: 34273135]
41. Garg I, Shekhar R, Sheikh AB, Pal S. Impact of COVID-19 on the changing patterns of respiratory syncytial virus infections. Infect Dis Rep 2022 Jul 24;14(4):558-568 [FREE Full text] [doi: 10.3390/idr14040059] [Medline: 35893478]
42. Fourgeaud J, Toubiana J, Chappuy H, Delacourt C, Moulin F, Parize P, et al. Impact of public health measures on the post-COVID-19 respiratory syncytial virus epidemics in France. Eur J Clin Microbiol Infect Dis 2021 Nov;40(11):2389-2395 [FREE Full text] [doi: 10.1007/s10096-021-04323-1] [Medline: 34347190]
43. Kim J, Roh YH, Ahn JG, Kim MY, Huh K, Jung J, et al. Respiratory syncytial virus and influenza epidemics disappearance in Korea during the 2020-2021 season of COVID-19. Int J Infect Dis 2021 Sep;110:29-35 [FREE Full text] [doi: 10.1016/j.ijid.2021.07.005] [Medline: 34245886]
44. Weinberger Opek M, Yeshayahu Y, Glatman-Freedman A, Kaufman Z, Sorek N, Brosh-Nissimov T. Delayed respiratory syncytial virus epidemic in children after relaxation of COVID-19 physical distancing measures, Ashdod, Israel, 2021. Euro Surveill 2021 Jul;26(29):1 [FREE Full text] [doi: 10.2807/1560-7917.ES.2021.26.29.2100706] [Medline: 34296678]
45. Vittucci AC, Piccioni L, Coltella L, Ciarlitto C, Antilici L, Bozzola E, et al. The disappearance of respiratory viruses in children during the COVID-19 pandemic. Int J Environ Res Public Health 2021 Sep 10;18(18):1 [FREE Full text] [doi: 10.3390/ijerph18189550] [Medline: 34574472]
46. Baker RE, Park SW, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. Proc Natl Acad Sci U S A 2020 Dec 01;117(48):30547-30553 [FREE Full text] [doi: 10.1073/pnas.2013182117] [Medline: 33168723]
47. Jia R, Lu L, Su L, Lin Z, Gao D, Lv H, et al. Resurgence of respiratory syncytial virus infection during COVID-19 pandemic among children in Shanghai, China. Front Microbiol 2022;13:938372 [FREE Full text] [doi: 10.3389/fmicb.2022.938372] [Medline: 35875547]
48. Ujiie M, Tsuzuki S, Nakamoto T, Iwamoto N. Resurgence of respiratory syncytial virus infections during COVID-19 pandemic, Tokyo, Japan. Emerg Infect Dis 2021 Nov;27(11):2969-2970 [FREE Full text] [doi: 10.3201/eid2711.211565] [Medline: 34388086]
49. Nam HH, Ison MG. Respiratory syncytial virus. Semin Respir Crit Care Med 2021 Dec;42(6):788-799. [doi: 10.1055/s-0041-1736182] [Medline: 34918321]
50. Shi N, Palmer L, Chu B, Katkin JP, Hall CB, Masaquel AS, et al. Association of RSV lower respiratory tract infection and subsequent healthcare use and costs: a Medicaid claims analysis in early-preterm, late-preterm, and full-term infants. J Med Econ 2011;14(3):335-340 [FREE Full text] [doi: 10.3111/13696998.2011.578188] [Medline: 21524154]

Abbreviations

ARTI: acute respiratory tract infection
LRTI: lower respiratory tract infection
NPI: nonpharmaceutical intervention
RR: risk ratio
RSV: respiratory syncytial virus
SARIMA: seasonal autoregressive integrated moving average
URTI: upper respiratory tract infection

Edited by T Sanchez, A Mavragani; submitted 31.10.22; peer-reviewed by CY Lin, S Alghamdi; comments to author 23.01.23; revised version received 11.02.23; accepted 23.03.23; published 27.04.23.

Please cite as:
Hu XW, Zhou Y, Yi S, Zhang WX, Wang XR, Du J, Lu QB
Epidemiological Characteristics of Respiratory Syncytial Virus Infection Among Hospitalized Children With Acute Respiratory Tract Infections From 2014 to 2022 in a Hospital in Hubei Province, China: Longitudinal Surveillance Study
JMIR Public Health Surveill 2023;9:e43941
URL: https://publichealth.jmir.org/2023/1/e43941
doi:10.2196/43941
PMID:36975172

Original Paper

Cause of Death in Patients with Oropharyngeal Carcinoma by Human Papillomavirus Status: Comparative Data Analysis

Dong-Dong Zhang[1], MD; Min Lei[1], MD; Yue Wang[1], MD; Pei-Ji Zeng[1], MD; Yong-Jun Hong[1], MD; Cheng-Fu Cai[1,2,3,4,5,6], MD

[1]Department of Otolaryngology-Head and Neck Surgery, School of Medicine, Zhongshan Hospital, Xiamen University, Xiamen, China
[2]School of Medicine, Xiamen University, Xiamen, China
[3]Xiamen Key Laboratory of Otolaryngology Head and Neck Surgery, Xiamen, Xiamen, China
[4]Department of Otorhinolaryngology, Head and Neck Surgery, Xiamen University, Xiamen, China
[5]Department of Otolaryngology Head and Neck Surgery, Fujian Medical University, Xiamen, China
[6]College of Otorhinolaryngology Head and Neck Surgery, Xiamen Haicang Hospital, Xiamen, China

Corresponding Author:
Cheng-Fu Cai, MD
Department of Otolaryngology-Head and Neck Surgery
School of Medicine, Zhongshan Hospital
Xiamen University
Hubin South Road 201-209
Xiamen, 361003
China
Phone: 86 13616011099
Email: caichengfuxmu@126.com

Abstract

Background: The incidence of oropharyngeal squamous cell carcinomas (OPSCC) has increased in recent decades, and human papillomavirus (HPV) infection is the main cause of OPSCC. The data regarding causes of death (CODs) are vitally important in informing follow-up strategies and revising treatment strategies to deal with any possible preventable treatment-related COD. However, limited studies have assessed the competing COD by HPV status in patients with OPSCC.

Objective: We aimed to analyze the distribution of the competing COD according to HPV status in OPSCC.

Methods: We retrospectively included stage I-IVB patients with OPSCC from the Surveillance, Epidemiology, and End Results database between 2010 and 2015. The association between HPV status and head and neck cancer–specific mortality (HNCSM), second primary cancer mortality (SPCM), and noncancer-caused mortality (NCCM) were analyzed. The chi-square test, Kaplan-Meier analysis, and Fine and Gray model were used for statistical analysis.

Results: We included 5852 patients in this study and 73.2% (n=4283) of them had HPV-related tumors. A total of 1537 (26.3%) patients died, including 789 (51.3%), 333 (21.7%), and 415 (27%) patients who died from head and neck cancer, second cancer, and noncancer causes, respectively. The 5-year HNCSM, SPCM, NCCM, and overall mortality were 14.7%, 6.5%, 7.7%, and 26.4%, respectively. Those with HPV-positive disease had a lower cumulative incidence of HNCSM (subdistribution hazard ratio [sHR] 0.362, 95% CI 0.315-0.417; $P<.001$), SPCM (sHR 0.400, 95% CI 0.321-0.496; $P<.001$), and NCCM (sHR 0.460, 95% CI 0.378-0.560; $P<.001$) than those with HPV-negative disease. The 5-year risk of HNCSM was 26.9% and 10.7% in those with HPV-negative and HPV-positive disease, respectively ($P<.001$). The 5-year risk of SPCM was 12.4% and 4.6% in those with HPV-negative and HPV-positive disease, respectively ($P<.001$). The 5-year risk of NCCM of death was 13.7% and 5.8% in those with HPV-negative and HPV-positive disease, respectively ($P<.001$). Using the Fine and Gray competing-risks model, our results show that those with HPV-negative tumors had a significantly higher risk of HNCSM ($P<.001$), SPCM ($P<.001$), and NCCM ($P<.001$) than those with HPV-negative tumors.

Conclusions: HPV-positive OPSCC has a lower NCSM, SPCM, and NCCM as compared to those with HPV-negative OPSCC. HPV positivity is a favorable prognostic factor in the context of overcoming cancer as well as in terms of reducing the risk of other CODs in OPSCC. Our finding supports the need to tailor patient follow-up based on the HPV status of patients with OPSCC.

(JMIR Public Health Surveill 2023;9:e47579) doi:10.2196/47579

KEYWORDS

oropharyngeal carcinomas; human papillomavirus; causes of death; survival; competing mortality; human papillomavirus; comparative data; analysis; cell death; carcinoma; oropharyngeal carcinoma; surveillance; database; human papillomavirus-related tumor

Introduction

Human papillomavirus (HPV) is the most prevalent sexually transmitted infection globally and is associated with both anogenital and oropharyngeal cancers. It is estimated that over 80% of individuals will be affected by HPV at some point in their lifetime [1]. Like other types of head and neck squamous cell carcinomas (HNSCCs), oropharyngeal squamous cell carcinoma (OPSCC) has been historically associated with alcohol and tobacco consumption. However, infection with carcinogenic HPV infection has emerged as an important risk factor that has driven an increase in the incidence of OPSCC. In recent decades, OPSCC has been one of the malignant tumors with rapidly rising incidences in Western countries [2-5]. In Italy, the incidence of HPV-related OPSCC increased from 16.7% in 2000-2006 to 46.1% in 2013-2018 [6]. Moreover, HPV now accounts for 71% and 51.8% of all OPSCCs in the United States and the United Kingdom, respectively [7-9]. The etiological relation with OPSCC by HPV status was also found in the Asian population [10,11]. The existing evidence shows that HPV status impacts the prognosis and treatment decision-making of OPSCC [12,13]. Patients with HPV-positive disease had a significantly better prognosis than those with HPV-negative disease [12,13]. Moreover, several studies also explored the treatment de-escalation for HPV-positive OPSCC [14,15]. These differences have a direct effect on clinical trials that will determine treatment therapeutic standards in the future clinical practice of OPSCC.

As survivorship from OPSCC continues to increase in recent decades [16-18], the main participants in the medical field should identify the patients with OPSCC with the highest risk of dying as well as the risk of their specific cause of death (COD). In addition to HPV, smoking and alcohol use are also closely related to the development of OPSCC, which may cause competing mortality in this population. In 2 studies from the Netherlands and the US cancer registry showed that the survival outcomes of head and neck cancer (HNC) were inferior to that of a healthy population of similar gender and age [19,20]. Smoking and alcohol use are the causative factors for cardiopulmonary diseases and second primary tumors. A previous Surveillance, Epidemiology, and End Result (SEER) study showed that life expectancy for patients with HNC was significantly shorter than expected after excluding the COD related to HNC [21]. The other COD may contribute to the differences in the survival of patients with HNC [22,23]. This information is vitally important in informing follow-up strategies and revising treatment strategies to deal with any possible preventable treatment-related COD.

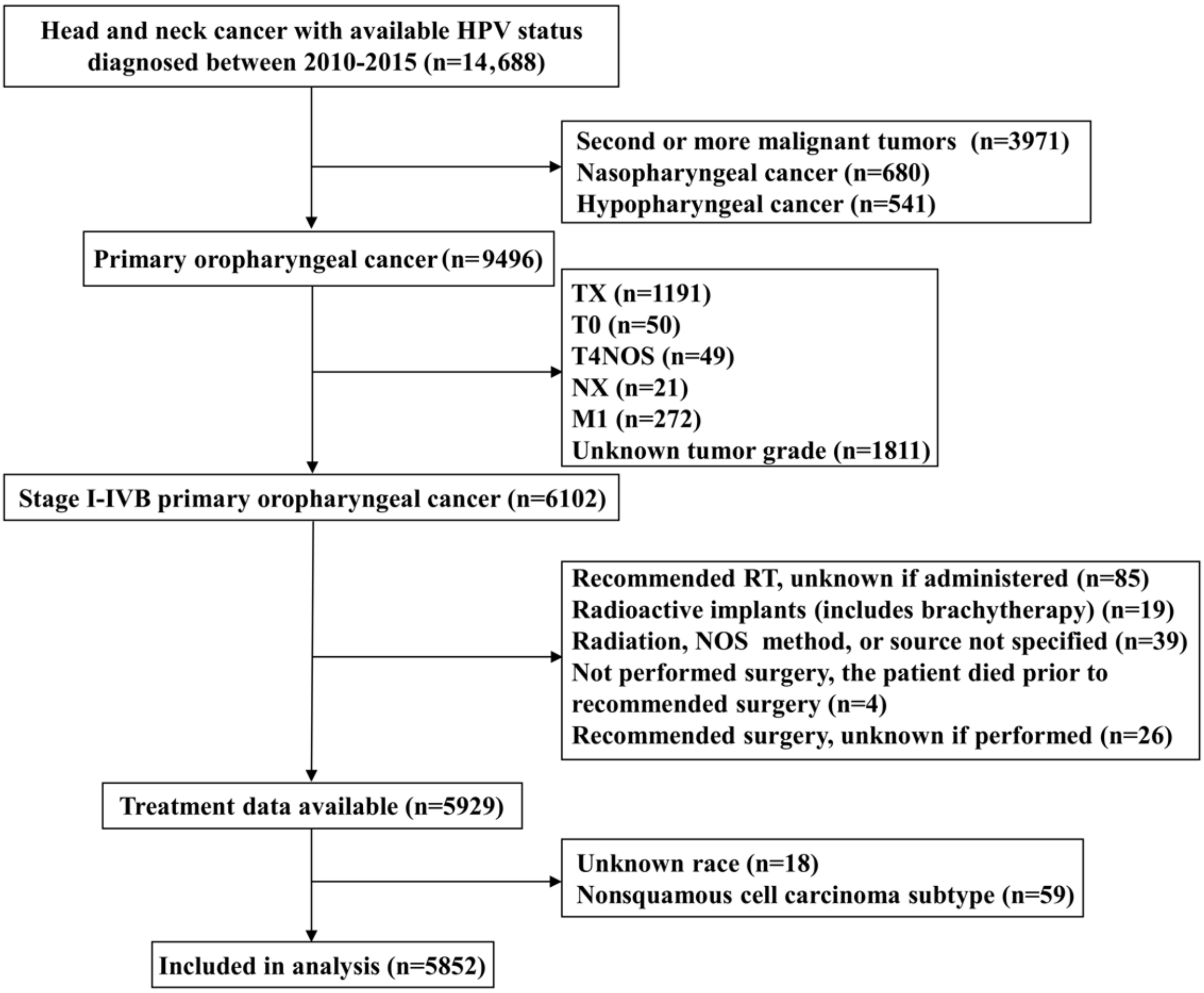
Because HPV-positive patients with OPSCC are usually younger and consume less alcohol and tobacco, the risk of cardiopulmonary disease and secondary primary tumor in this population may be lower [24,25]. The differences in survival rates for patients with HNSCC can primarily be attributed to an unequal burden of comorbidities associated with tobacco and alcohol use, treatment-related sequelae, and a high risk for the development of second primary cancer. Moreover, the surveillance for HPV-related tumors is similar to HPV-negative tumors in the current guidelines. Given the rapid rise in the incidence of potentially HPV-related OPSCC, it is crucial to provide a more comprehensive description of the COD in OPSCC. Understanding the COD by different HPV status in OPSCC could bring even more positive impact on public health, potentially impacting the follow-up strategies and revising treatment strategies while reducing treatment-related morbidity. However, limited studies have assessed the competing COD by HPV status in patients with OPSCC [26]. The objective of this study was to analyze the differences in the competing COD depending on HPV status in OPSCC.

Methods

Patient Selection

The National Cancer Institute SEER database collects population-level data on approximately 28% of the US population from 18 registries with the morbidity, mortality, and disease status of millions of patients with cancer [27]. We used the SEER databases to study patients diagnosed with OPSCC between 2010 and 2015 (International Classification of Disease for Oncology ICD-O C019, C090, C091, C098-C104, C108, C109, C142; histology codes 8073-8079 and 8083-8084). Registrars for the SEER database were required to enter HPV status beginning in 2010 and we have obtained HPV-related variables through the use of a data user agreement. The data of the retrospective cohort in the present study were obtained from the SEER HPV database, which includes the demographic data, clinicopathological characteristics, treatment, HPV status, and follow-up of patients with HNSCC [27]. We included patients who met the following criterion: (1) diagnosed between 2010 and 2015; (2) diagnosed with stage I-IVB OPSCC according to the 7th American Joint Committee on Cancer (AJCC) staging system; (3) available data regarding HPV status, race, tumor grade, AJCC staging, tumor stage, nodal stage, and metastatic stage; and (4) available data on surgery, radiotherapy, and chemotherapy. Patients with unknown HPV status, second or more malignant tumors, distant stage, and treated with radioactive implants or radioisotopes were excluded. The patient selection flowchart is listed in Figure 1.

Figure 1. Flow diagram of this study's cohort. HPV: human papillomavirus; M: metastatic stage; N: nodal stage; NOS: not otherwise specified; NX: unknown nodal stage; RT: radiotherapy; T: tumor stage; TX: unknown tumor stage.



Ethics Approval

The SEER database is a public database, and we have obtained permission to access the database (approval number 18275-Nov2021). The present study did not require ethical approval by the institutional review board of the Zhongshan Hospital, Xiamen University, because of the nonidentifying information in the SEER database. The requirement for informed consent was waived because all data used in this study were deidentified.

Data Collection

The following variables were included: age, gender, race, tumor grade, HPV status, AJCC 7th staging, and the receipt of treatment. HPV status was determined by p16 immunohistochemistry, tissue polymerase chain reaction, in situ hybridization, real-time polymerase chain reaction for E6/7 RNA, or in situ hybridization for E6/7 RNA [28]. The primary outcomes were HNC-specific mortality (HNCSM), second primary cancer mortality (SPCM), and noncancer-caused mortality (NCCM). Any COD due to HNC was encoded as HNCSM. Events for SPCM were death from the second primary cancer. Events for NCCM were death from noncancer diseases. Our secondary outcome was overall mortality (OM). OM was defined as death from all causes.

Statistical Analysis

Categorical variables were represented as absolute values with percentages, and the chi-square test was performed for the comparison between patients who were HPV-negative and HPV-positive. Kaplan-Meier analysis was used to generate the survival curves to compare differences in survival between groups over time, and the equality of these curves was compared by the log-rank test. Due to the competitive relationship among multiple outcomes in clinical survival data, the traditional Kaplan-Meier analysis will misjudge the cumulative mortality rate [29]. Therefore, competing-risks analyses using the Fine and Gray model were used to estimate the effects of covariates on HNCSM, SPCM, and NCCM. We conducted stratum-specific analyses to estimate the effects of HPV status on HNCSM, SPCM, and NCCM in accordance with gender, age, race, AJCC staging, grade, surgery, radiotherapy, and chemotherapy. All statistical analyses were conducted using SPSS (version 22.0; IBM Corp) and Stata/SE (version 14; StataCorp). A *P* value <.05 was considered statistically significant.

Results

Patient Baseline Characteristics

From 2010 to 2015, there were 14,688 cases of HNC recorded with the HPV status in the SEER HPV database; 9496 of these cases were oropharyngeal cancer and 5852 of these cases were stage I-IVB OPSCC and had treatment data available. In this cohort (n=5852; Table 1), 4283 patients (73.2%) had HPV-positive OPSCC and 1569 patients (26.8%) had HPV-negative OPSCC. Of these patients, 84.1% (n=4922), 89.2% (n=5220), and 69.9% (n=4088) were male, White, and had stage T4 disease, respectively. Regarding treatment, 89.1% (n=5216), 71.3% (n=4174), and 55% (n=2626) of patients received radiotherapy, chemotherapy, and surgery, respectively.

Table 1. Patient baseline characteristics.

Variables	Patients, n	HPV[a]-negative, n (%)	HPV-positive, n (%)	*P* value
Gender				<.001
Male	4922	1213 (77.3)	3709 (86.6)	
Female	930	356 (22.7)	574 (13.4)	
Age (years)				<.001
<50	840	207 (13.2)	633 (14.8)	
50-64	3425	847 (54)	2578 (60.2)	
≥65	1587	515 (32.8)	1072 (25)	
Race				<.001
White	5220	1288 (82.1)	3932 (91.8)	
Black	414	212 (13.5)	202 (4.7)	
Other	218	69 (4.4)	149 (3.5)	
Grade				<.001
Well differentiated	233	113 (7.2)	120 (2.8)	
Moderately differentiated	2240	754 (48.1)	486 (34.7)	
Poorly or undifferentiated	3379	702 (44.7)	2677 (62.5)	
AJCC[b] stage				<.001
I	218	93 (5.9)	125 (2.9)	
II	384	108 (6.9)	276 (6.4)	
III	1162	342 (21.8)	820 (19.1)	
IVA and IVB	4088	1026 (65.4)	3062 (71.5)	
T[c] stage				<.001
T1	1630	360 (22.9)	1270 (29.7)	
T2	2322	527 (33.6)	1795 (41.9)	
T3	1057	342 (21.8)	715 (16.7)	
T4	843	340 (21.7)	503 (11.7)	
N[d] stage				<.001
N0	935	338 (21.5)	597 (13.9)	
N1	1096	326 (20.8)	770 (18)	
N2	3562	836 (53.3)	2726 (63.6)	
N3	259	69 (4.4)	190 (4.4)	
Surgery				<.001
No	3226	964 (61.4)	2262 (52.8)	
Yes	2626	605 (38.6)	2021 (47.2)	
Radiotherapy				<.001
No	636	228 (14.5)	408 (9.5)	
Yes	5216	1341 (85.5)	3875 (90.5)	
Chemotherapy				.02
No	1678	487 (31)	1191 (27.8)	
Yes	4174	1082 (69)	3092 (72.2)	
Causes of death (N=1537)				.50
Head and neck cancer	789	378（52.6）	411 (50.2)	

Variables	Patients, n	HPV[a]-negative, n (%)	HPV-positive, n (%)	*P* value
Second cancer	333	156 (21.7)	177 (21.6)	
Noncancer cause disease	415	184 (25.6)	231 (28.2)	

[a] HPV: human papillomavirus.

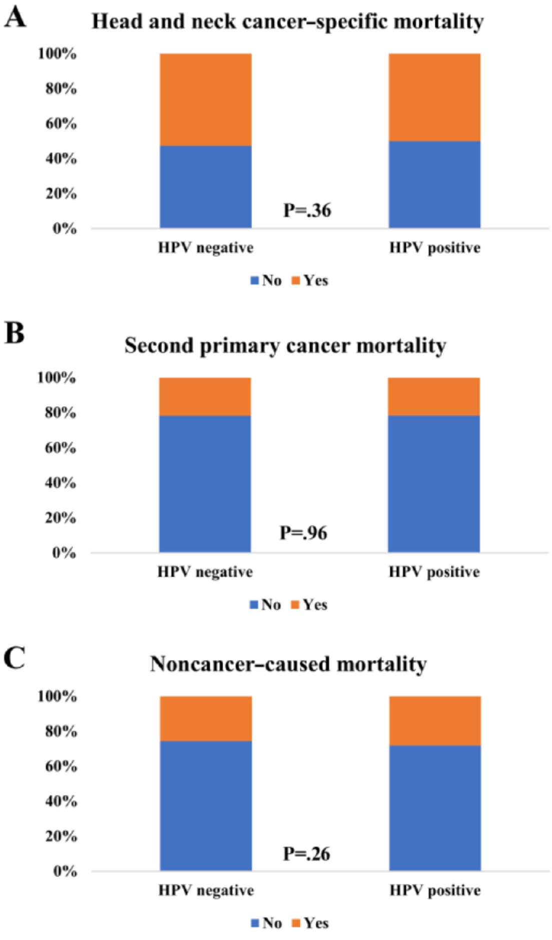
[b] AJCC: American Joint Committee on Cancer.

[c] T: tumor stage.

[d] N: nodal stage.

Patients with HPV-positive disease were more likely to be male, aged 50-64 years, White, have higher tumor grade, have early tumor stage, have advanced nodal stage, and have stage IV disease (all *P*<.001). Moreover, those with HPV-positive disease were more likely to receive surgery (*P*<.001) and radiotherapy (*P*=.015; Table 1). Similar distributions in HNCSM (*P*=.36), SPCM (*P*=.96), and NCCM (*P*=.26) were found according to HPV status (Figure 2).

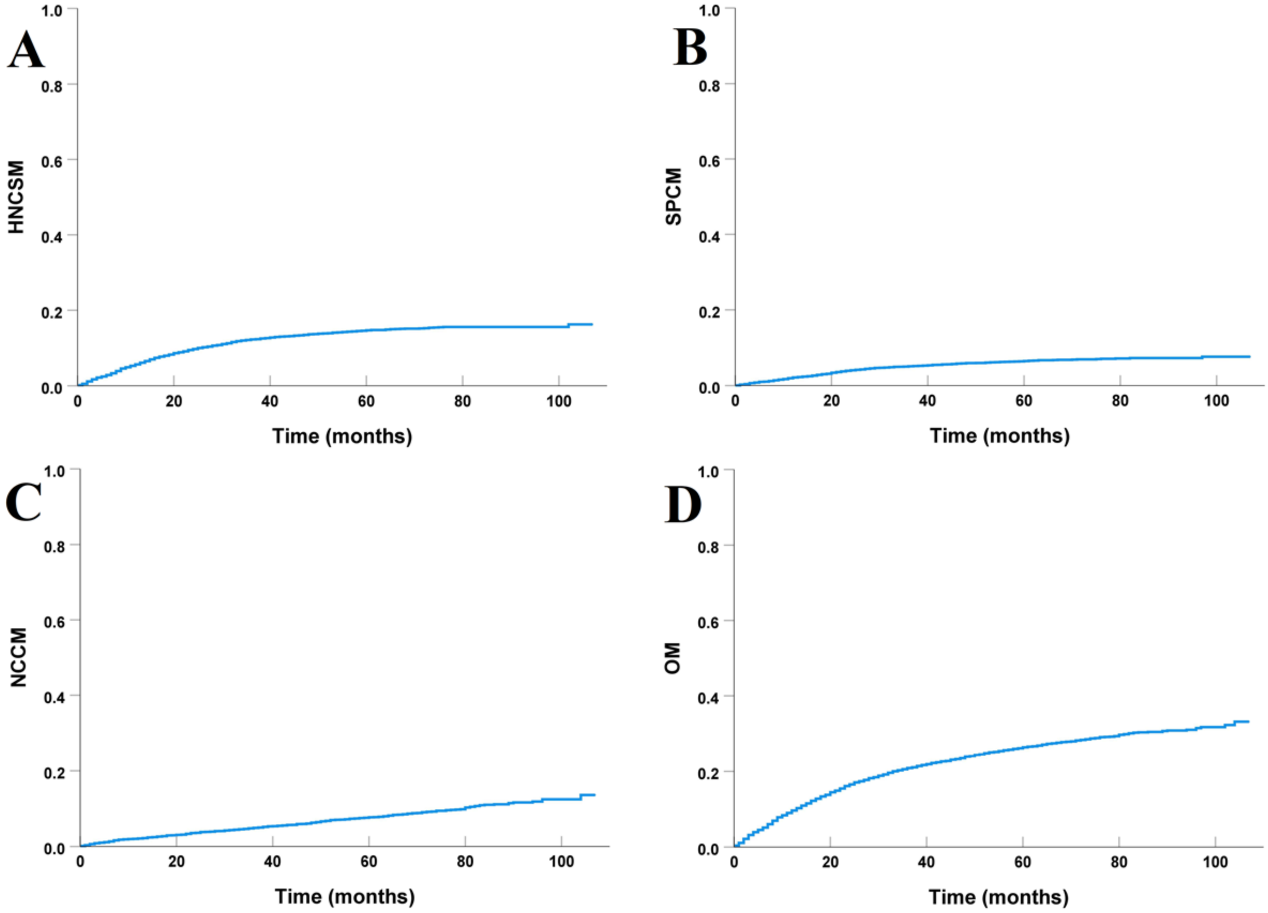
Figure 2. The distributions in (A) head and neck cancer–specific mortality, (B) second primary cancer mortality, and (C) noncancer-caused mortality according to human papillomavirus status. HPV: human papillomavirus.



Survival

With a median follow-up time of 52 months (range, 0-107 months), a total of 1537 (26.3%) patients died. There were 789 (51.3%), 333 (21.7%), and 415 (27%) patients who died from HNC, second cancer, and noncancer causes, respectively. The survival curves have been listed in Figure 3. The 5-year HNCSM, SPCM, NCCM, and OM were 14.7%, 6.5%, 7.7%, and 26.4%, respectively.

Figure 3. The curves of (A) head and neck cancer–specific mortality, (B) second primary cancer mortality, (C) noncancer-caused mortality, and (D) overall mortality in the entire cohort. HNCSM: head and neck cancer–specific mortality; NCCM: noncancer-caused mortality; OM: overall mortality; SPCM: second primary cancer mortality.



Competing Mortality

The cumulative incidence of competing CODs based on HPV status regarding HNCSM, SPCM, and NCCM are listed in Figure 4. Those with HPV-positive disease had a lower cumulative incidence of HNCSM (subdistribution hazard ratio [sHR] 0.362, 95% CI 0.315-0.417; *P*<.001; Figure 4A), SPCM (sHR 0.400; 95% CI 0.321-0.496; *P*<.001; Figure 4B), and NCCM (sHR 0.460; 95% CI 0.378-0.560; *P*<.001; Figure 4C) than those with HPV-negative disease. Using Kaplan-Meier analysis, the results showed that patients who were HPV-negative were more likely to gain experience from second cancer and noncancer causes. The 5-year risk of HNCSM was 26.9% and 10.7% in those with HPV-negative and HPV-positive disease, respectively (*P*<.001, Figure 5A). The 5-year risk of SPCM was 12.4% and 4.6% in those with HPV-negative and HPV-positive disease, respectively (*P*<.001, Figure 5B). The 5-year risk of NCCM of death was 13.7% and 5.8% in those with HPV-negative and HPV-positive disease, respectively (*P*<.001, Figure 5C). The 5-year OM was 26.1% and 10.7% in those with HPV-negative and HPV-positive, respectively (*P*<.001; Figure 5D).

Figure 4. The cumulative incidence of (A) head and neck cancer–specific mortality, (B) second primary cancer mortality, and (C) noncancer-caused mortality by human papillomavirus status using competing-risks regression. HNCSM: head and neck cancer–specific mortality; HPV: human papillomavirus; NCCM: noncancer-caused mortality; SPCM: second primary cancer mortality.

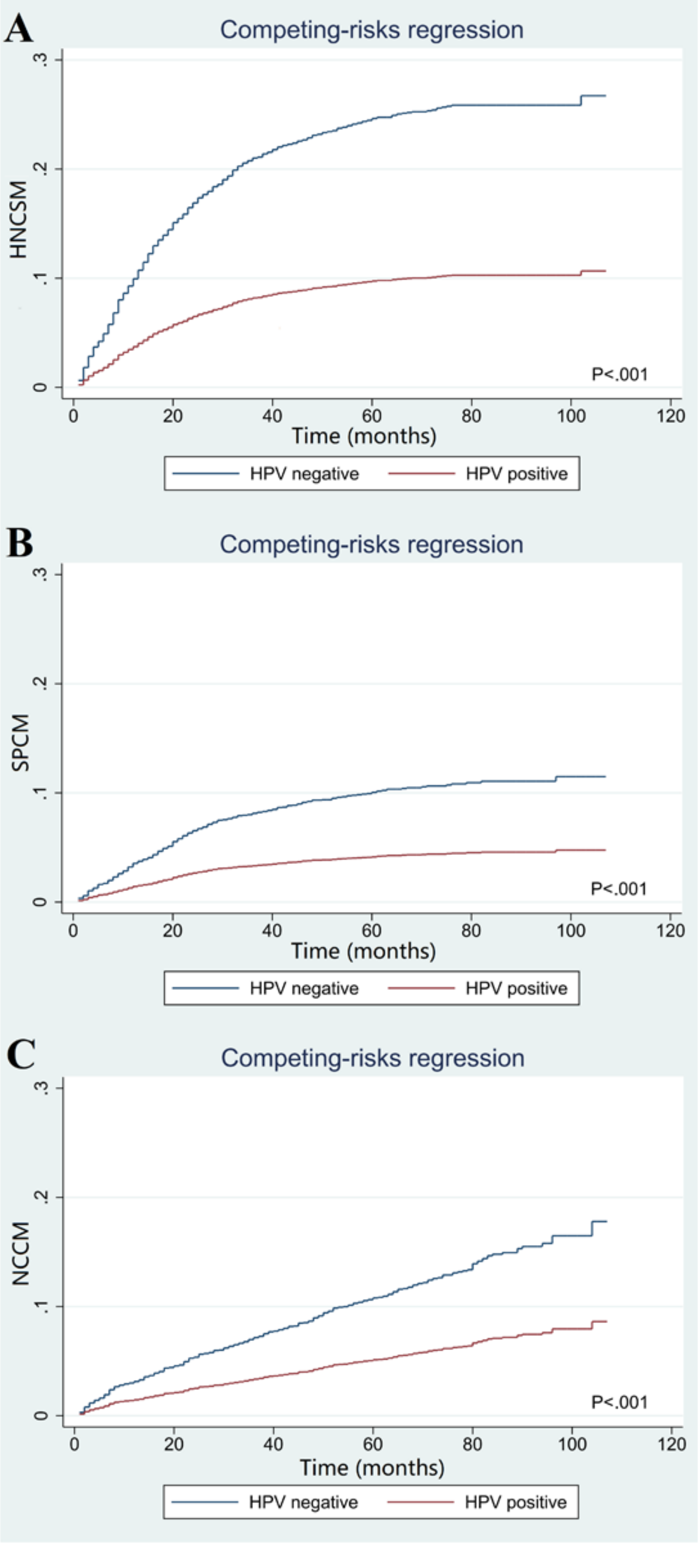
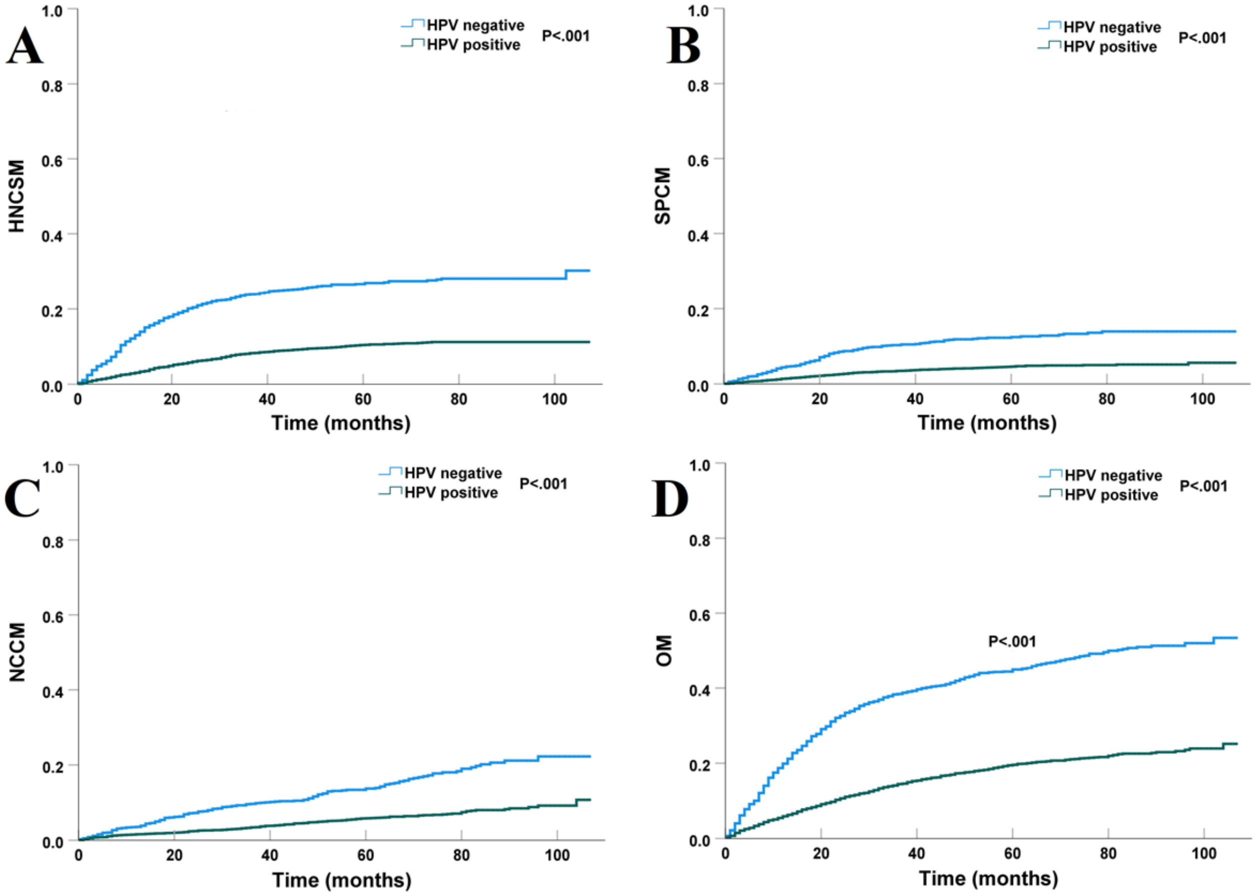


Figure 5. The effect of human papillomavirus status on (A) head and neck cancer–specific survival, (B) overall survival, (C) second primary cancer survival, and (D) noncancer-caused survival. HNCSM: head and neck cancer–specific mortality; HPV: human papillomavirus; NCCM: noncancer-caused mortality; OM: overall mortality; SPCM: second primary cancer mortality.



Fine and Gray Competing-Risks Model

Table 2 shows the sHR and 95% CI for selected covariates using competing-risks regression for different COD. The results of the multivariate analysis showed that the variables related to HNCSM were female, older age, Black, advanced stage, HPV-negative, receipt of surgery, and receipt of radiotherapy (all *P*<.05). Advanced stage, HPV-negative, receipt of surgery, and receipt of radiotherapy were the independent risk factors associated with SPCM (all *P*<.05). Older age, Black race, HPV-negative, receipt of surgery, and receipt of chemotherapy were the independent risk factors associated with NCCM (all *P*<.05). Stratum-specific analyses also showed that patients with HPV-positive tumor had lower HNCSM, SPCM, and NCCM than those with HPV-negative tumors after stratification by different demographic, clinicopathologic, and treatment characteristics (Table S1 in Multimedia Appendix 1).

Table 2. Multivariate prognostic analyses using the competing-risks model.

Variables	HNC[a]-specific mortality			Second primary cancer mortality			Noncancer-caused mortality		
	sHR[b]	95% CI	*P* value	sHR	95% CI	*P* value	sHR	95% CI	*P* value
Gender									
Male (reference)	1	N/A[c]	N/A	1	N/A	N/A	1	N/A	N/A
Female	1.218	1.012-1.467	.04	0.729	0.528-1.005	.05	0.993	0.771-1.279	.96
Age (years)									
<50 (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
50-64	1.497	1.160-1.932	.002	1.231	0.872-1.737	.24	1.251	0.861-1.819	.24
≥65	2.176	1.665-2.844	<.001	1.26	0.868-1.830	.22	3.139	2.164-4.554	<.001
Race									
White (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
Black	1.525	1.215-1.913	<.001	1.167	0.811-1.677	.41	1.548	1.131-2.120	.006
Other	1.049	0.721-1.525	.08	0.931	0.523-1.658	.81	1.272	0.787-2.055	.33
AJCC[d] stage									
I (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
II	2.147	1.117-4.126	.02	1.124	0.435-2.901	.81	0.986	0.575-1.691	.96
III	2.969	1.637-5.384	<.001	0.403	0.615-3.201	.42	0.895	0.545-1.469	.66
IVA and IVB	4.515	2.534-8.043	<.001	2.583	1.168-5.713	.02	0.875	0.542-1.412	.59
Grade									
Well differentiated (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
Moderately differentiated	0.797	0.581-1.094	.16	1.083	0.638-1.839	.77	1.059	0.669-1.676	.81
Poorly or undifferentiated	0.590	0.430-0.810	.001	0.814	0.480-1.381	.45	0.847	0.532-1.351	.49
HPV[e] status									
Negative (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
Positive	0.439	0.378-0.508	<.001	0.437	0.349-0.547	<.001	0.584	0.472-0.721	<.001
Surgery									
No (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
Yes	0.463	0.389-0.552	<.001	0.524	0.400-0.689	<.001	0.572	0.456-0.716	<.001
Radiotherapy									
No (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
Yes	0.407	0.318-0.524	<.001	0.387	0.267-0.561	<.001	0.932	0.666-1.304	.68
Chemotherapy									
No (reference)	1	N/A	N/A	1	N/A	N/A	1	N/A	N/A
Yes	0.987	0.801-1.215	.90	1.152	0.834-1.592	.39	0.688	0.525-0.877	.003

[a] HNC: head and neck cancer.

[b] sHR: subdistribution hazard ratio.

[c] N/A: not applicable.

[d] AJCC: American Joint Committee on Cancer.

[e] HPV: human papillomavirus.

Discussion

Principal Findings

In this study, we investigated the competing mortality in OPSCC by HPV status using a large cohort. Our results indicated that patients with HPV-negative tumors had a significantly higher risk of HNCSM, SPCM, and NCCM than those with HPV-positive tumors. The main finding of our study was that HPV-positive disease was an independent factor for better survival outcomes across all COD.

HPV status has been confirmed to be an important prognostic and predictive biomarker in OPSCC [12,13]. In recent years, several studies also have explored the relationship between HPV status and the second tumor in OPSCC. Gorphe et al [30] analyzed 888 patients with OPSCC between 2011 and 2020, they found that the 5-year competing risks of a second cancer were 16.1% and 49.9% in those patients who were HPV-positive and HPV-negative, respectively (*P*<.001). In addition, a study from Bosshart et al [31] that included 91 patients also found that HPV-positive had a lower risk of second cancer than those with HPV-negative tumors (14% vs 38.2%). Since the time of COD was only recorded in the SEER database, we could not obtain the exact time of diagnosis of the second primary tumor. However, Culié et al [32] included 340 patients with OPSCC from 14 French hospitals, and they found that patients who were HPV-positive developed low second tumors than patients who were HPV-negative (3.3% vs 11.7%, *P*=.02). However, the COD was similar in the 2 groups.

In our study, we found HNC-related death was the main COD in OPSCC (14.7%), followed by noncancer causes (7.7%), and second cancer (6.5%). Alcohol and tobacco use have been confirmed as risk factors for HNSCC. HPV-related OPSCC cannot be considered in the same field of cancerization as HPV-negative OPSCC and other HNSCCs [2-5]. However, we observed a quite similar distribution of CODs between the 2 groups, with primary cancer as the major COD. Our study suggested that the overall survival pattern of patients with OPSCC depends not only on their clinicopathological characteristics, but also significantly varies based on their age and the presence of comorbidities.

Several previous studies have assessed CODs in patients with HNSCC. They found that age, gender, race or ethnicity, presence of comorbidities, BMI, or type of treatment were the risk factors in the competing COD in patients with HNSCC [22,33-37]. In our study, we analyzed whether the HPV status would influence the COD of patients with OPSCC. A population-based cohort from Denmark was the first to investigate COD by HPV status in 1521 patients with OPSCC [26]. There were 54.3% of patients were HPV-positive and they found that HPV-positive was associated with better survival across all CODs [26]. Lop et al [38] included 423 patients with OPSCC and 12.5% were HPV-positive. They found that patients with HPV-positive disease had better OPSCC-specific survival (*P*<.001) and second primary cancer survival (*P*<.001) than those with HPV-negative tumors. However, noncancer-related causes survival was similar between the 2 groups (*P*=.13). The difference in patient characteristics may affect the distribution of CODs in the previous studies. Those with HPV-positive disease had a significantly lower cumulative incidence of death from all CODs than those with HPV-negative disease. Therefore, the inferior survival in patients with HPV-negative tumors was reflected in the higher risk of HNCSM, SPCM, and NCCM. Patients with HPV-negative OPSCC may benefit from screening for the risk to die of noncancer death or early detection of second cancers. Our study suggests that OPSCC should undergo more different effective therapies, more prevention methods to avoid noncancer deaths, and more screening tools for early detection of the second tumor according to different HPV status. They may benefit from increased research funding.

There were several studies indicating that those with HPV-positive OPSCC have a favorable survival than those with HPV-negative disease regardless of the treatment received [39,40]. In our study, 89.8% of patients were locally advanced-stage OPSCC, we also found that the receipt of treatment was associated with all CODs. The probability of treatment compliance was not recorded in the SEER database. Hess et al [41] found that treatment compliance in patients who were HPV-negative was significantly lower than in those with patients who were HPV-positive. Therefore, treatment compliance by HPV status might also have a negative effect on cancer-specific survival in OPSCC.

Screening for secondary cancer has an important effect in reducing mortality, especially for the HPV-negative population with a higher level of tobacco use [42]. Our data indicate that there were 51.3% and 21.7% patients who died from HNC and second cancer. Therefore, more frequent follow-ups should be performed to increase the detection of recurrence and secondary cancers. Moreover, our study shows an increased risk of SPCM, and NCCM in patients who were HPV-positive than in those with HPV-negative tumors. Our finding is important because it adds to the current knowledge regarding the differences in CODs among HPV status and supports the need to individualize follow-ups according to the HPV status.

In our study, 415 patients died of noncancer-related diseases, and 127 (30.6%) died from heart disease. In our study, patients with HPV-positive OPSCC tend to be younger, White, and male. Previous studies also showed that women of the Black race had a higher risk of cardiovascular disease in the general population and were cancer survivors [43,44]. According to the differences regarding epidemiological characteristics by HPV status may affect the burden of comorbidity, thus affecting the NCCM.

HPV-related tumors have a significant burden on health insurance [45]. While the HPV vaccination has been shown to prevent cervical cancer, HPV vaccination also has the potential to prevent a significant proportion of HPV-related OPSCC. We should note that 84.1% of patients in our study were male, and 85.6% were aged 50 years or older. The vaccine only prevents new HPV infections and does not treat existing infections [46]. HPV vaccination is generally not recommended for those older than age 26 years who are unvaccinated, recent guidelines have a permissive recommendation for HPV vaccination in unvaccinated adults aged 27-45 years [46]. Additionally, 47.7% of patients continued to smoke after their OPSCC diagnosis [47]. Therefore, current smoking status should be monitored throughout the treatment course of OPSCC, and smoking-cessation assistance should be provided as necessary, especially for patients with negative HPV status and a history of smoking. Therefore, further studies are needed to verify the efficacy of HPV vaccination in reducing the incidence of OPSCC and to investigate the effect of smoking cessation on various COD in patients with OPSCC.

We acknowledge certain limitations in our study. First, retrospective studies have inherent weaknesses of selection bias. Second, the lack of data regarding other sexually transmitted infections limits the interpretation of our data because behaviors that put people at risk for HPV also increase their risk for other sexually transmitted infections. Third, the lack of comorbidity data also limits the interpretation of our data. Of note, HNCSM, SPCM, and NCCM data were available and were used in this study. Fourth, the consumption of smoking and alcohol was not recorded in the SEER database. The high rates of HPV-positive OPSCC in our study could potentially be attributed to the low prevalence of smoking habits and differences in sexual behavior when compared with other geographic regions with a low incidence of HPV-positive OPSCC [37,48]. This difference in the incidence of HPV-positive tumors could limit the generalization of our findings to regions with a low incidence of HPV-positive tumors and high levels of tobacco and alcohol use. Moreover, the median follow-up time was 52 months in our study and second cancer or noncancer death is likely to develop after this delay. Finally, the SEER does not have specific records on HPV status detection methods. However, a previous study showed that the overall HPV status (positive or negative) had high degrees of accuracy in the SEER database [49]. Further research addressing these limitations could provide a more comprehensive understanding of the relationship between HPV status and competing mortality in patients with OPSCC.

Conclusions

In conclusion, our study suggests that HPV-positive OPSCC has a lower NCSM, SPCM, and NCCM as compared to those with HPV-negative OPSCC. HPV positivity is a favorable prognostic factor in the context of overcoming cancer as well as in terms of reducing the risk of other CODs in OPSCC. These disparities have direct implications on clinical trials aiming to define future therapeutic standards in clinical practice, including treatment intensification or de-escalation according to response rates, investigation of adjuvant treatment after complete response, or follow-up strategies involving imaging modalities and duration. Moreover, medical interventions tailored to the intercurrent CODs may independently provide survival benefits in addition to cancer treatments.

Acknowledgments

This project was supported by the Innovation Project in Fujian Province (2019-CXB-29), the Natural Science Foundation of Fujian Science and Technology Department (2020J02060), the Key Medical and Health Project of Xiamen Science and Technology Bureau (3502Z20204009), the Natural Science Foundation of Fujian Province of China (2022J05006), and the National Natural Science Foundation of China (32070761).

Data Availability

The data set used in the data analysis is publicly available.

Authors' Contributions

DDZ and ML drafted the paper. CFC acquired the data sets. CFC conceived this study. YW, PJZ, and YJH conducted the statistical analyses. DDZ, ML, and CFC participated in this study's design. All authors read and approved the final paper.

Conflicts of Interest

None declared

Multimedia Appendix 1
Table S1. Stratum-specific analyses to determine the effect of HPV status according to different demography and clinicopathological characteristics using the Fine and Gray competing-risks model.
[DOCX File , 27 KB - publichealth_v9i1e47579_app1.docx]

References

1. Scott-Wittenborn N, Fakhry C. Epidemiology of HPV related malignancies. Semin Radiat Oncol 2021;31(4):286-296 [FREE Full text] [doi: 10.1016/j.semradonc.2021.04.001] [Medline: 34455984]
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127(12):2893-2917 [FREE Full text] [doi: 10.1002/ijc.25516] [Medline: 21351269]
3. Blomberg M, Nielsen A, Munk C, Kjaer SK. Trends in head and neck cancer incidence in Denmark, 1978-2007: focus on human papillomavirus associated sites. Int J Cancer 2011;129(3):733-741 [FREE Full text] [doi: 10.1002/ijc.25699] [Medline: 20878955]

4. Carlander ALF, Larsen CG, Jensen DH, Garnæs E, Kiss K, Andersen L, et al. Continuing rise in oropharyngeal cancer in a high HPV prevalence area: a Danish population-based study from 2011 to 2014. Eur J Cancer 2017;70:75-82. [doi: 10.1016/j.ejca.2016.10.015] [Medline: 27888679]
5. Tardini E, Zhang X, Canahuate G, Wentzel A, Mohamed ASR, Van Dijk L, et al. Optimal treatment selection in sequential systemic and locoregional therapy of oropharyngeal squamous carcinomas: deep Q-learning with a patient-physician digital twin dyad. J Med Internet Res 2022;24(4):e29455 [FREE Full text] [doi: 10.2196/29455] [Medline: 35442211]
6. Del Mistro A, Frayle H, Menegaldo A, Favaretto N, Gori S, Nicolai P, et al. Age-independent increasing prevalence of human papillomavirus-driven oropharyngeal carcinomas in North-East Italy. Sci Rep 2020;10(1):9320 [FREE Full text] [doi: 10.1038/s41598-020-66323-z] [Medline: 32518378]
7. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol 2011;29(32):4294-4301 [FREE Full text] [doi: 10.1200/JCO.2011.36.4596] [Medline: 21969503]
8. Senkomago V, Henley SJ, Thomas CC, Mix JM, Markowitz LE, Saraiya M. Human papillomavirus-attributable cancers—United States, 2012-2016. MMWR Morb Mortal Wkly Rep 2019;68(33):724-728 [FREE Full text] [doi: 10.15585/mmwr.mm6833a3] [Medline: 31437140]
9. Schache AG, Powell NG, Cuschieri KS, Robinson M, Leary S, Mehanna H, et al. HPV-related oropharynx cancer in the United Kingdom: an evolution in the understanding of disease etiology. Cancer Res 2016;76(22):6598-6606 [FREE Full text] [doi: 10.1158/0008-5472.CAN-16-0633] [Medline: 27569214]
10. Maruyama H, Yasui T, Ishikawa-Fujiwara T, Morii E, Yamamoto Y, Yoshii T, et al. Human papillomavirus and p53 mutations in head and neck squamous cell carcinoma among Japanese population. Cancer Sci 2014;105(4):409-417 [FREE Full text] [doi: 10.1111/cas.12369] [Medline: 24521534]
11. Chor JSY, Vlantis AC, Chow TL, Fung SC, Ng FYH, Lau CH, et al. The role of human papillomavirus in head and neck squamous cell carcinoma: a case control study on a Southern Chinese population. J Med Virol 2016;88(5):877-887. [doi: 10.1002/jmv.24405] [Medline: 26467027]
12. Wittekindt C, Klussmann JP. Tumor staging and HPV-related oropharyngeal cancer. Recent Results Cancer Res 2017;206:123-133. [doi: 10.1007/978-3-319-43580-0_9] [Medline: 27699534]
13. Wagner S, Reder H, Sharma SJ, Würdemann N, Wittekindt C, Klußmann JP. HPV-driven oropharyngeal carcinoma - incidence, trends, diagnosis and therapy. Urologe 2018;57(12):1457-1463 Das HPV-getriebene Oropharynxkarzinom – Inzidenz, Trends, Diagnose und Therapie. [doi: 10.1007/s00120-018-0810-4] [Medline: 30421134]
14. Petrelli F, Luciani A, Ghidini A, Cherri S, Gamba P, Maddalo M, et al. Treatment de-escalation for HPV+ oropharyngeal cancer: a systematic review and meta-analysis. Head Neck 2022;44(5):1255-1266. [doi: 10.1002/hed.27019] [Medline: 35238114]
15. Patel RR, Ludmir EB, Augustyn A, Zaorsky NG, Lehrer EJ, Ryali R, et al. De-intensification of therapy in human papillomavirus associated oropharyngeal cancer: a systematic review of prospective trials. Oral Oncol 2020;103:104608. [doi: 10.1016/j.oraloncology.2020.104608] [Medline: 32163853]
16. Fakhry C, Andersen KK, Eisele DW, Gillison ML. Oropharyngeal cancer survivorship in Denmark, 1977-2012. Oral Oncol 2015;51(11):982-984. [doi: 10.1016/j.oraloncology.2015.08.006] [Medline: 26319353]
17. Duman-Lubberding S, van Uden-Kraan CF, Peek N, Cuijpers P, Leemans CR, de Leeuw IMV. An eHealth application in head and neck cancer survivorship care: health care professionals' perspectives. J Med Internet Res 2015;17(10):e235 [FREE Full text] [doi: 10.2196/jmir.4870] [Medline: 26489918]
18. Mujcic A, Blankers M, Boon B, de Leeuw IMV, Smit F, van Laar M, et al. Effectiveness, cost-effectiveness, and cost-utility of a digital smoking cessation intervention for cancer survivors: health economic evaluation and outcomes of a pragmatic randomized controlled trial. J Med Internet Res 2022;24(3):e27588 [FREE Full text] [doi: 10.2196/27588] [Medline: 35297777]
19. van der Schroeff MP, van de Schans SAM, Piccirillo JF, Langeveld TPM, de Jong RJB, Janssen-Heijnen MLG. Conditional relative survival in head and neck squamous cell carcinoma: permanent excess mortality risk for long-term survivors. Head Neck 2010;32(12):1613-1618. [doi: 10.1002/hed.21369] [Medline: 20310043]
20. Massa ST, Osazuwa-Peters N, Christopher KM, Arnold LD, Schootman M, Walker RJ, et al. Competing causes of death in the head and neck cancer population. Oral Oncol 2017;65:8-15. [doi: 10.1016/j.oraloncology.2016.12.006] [Medline: 28109473]
21. Massa ST, Cass LM, Osazuwa-Peters N, Christopher KM, Walker RJ, Varvares MA. Decreased cancer-independent life expectancy in the head and neck cancer population. Head Neck 2017;39(9):1845-1853. [doi: 10.1002/hed.24850] [Medline: 28640483]
22. Mell LK, Dignam JJ, Salama JK, Cohen EEW, Polite BN, Dandekar V, et al. Predictors of competing mortality in advanced head and neck cancer. J Clin Oncol 2010;28(1):15-20 [FREE Full text] [doi: 10.1200/JCO.2008.20.9288] [Medline: 19933920]
23. Baxi SS, Pinheiro LC, Patil SM, Pfister DG, Oeffinger KC, Elkin EB. Causes of death in long-term survivors of head and neck cancer. Cancer 2014;120(10):1507-1513 [FREE Full text] [doi: 10.1002/cncr.28588] [Medline: 24863390]

24. Larsen CG, Jensen DH, Carlander ALF, Kiss K, Andersen L, Olsen CH, et al. Novel nomograms for survival and progression in HPV+ and HPV- oropharyngeal cancer: a population-based study of 1,542 consecutive patients. Oncotarget 2016;7(44):71761-71772 [FREE Full text] [doi: 10.18632/oncotarget.12335] [Medline: 27708214]
25. Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 2008;100(4):261-269 [FREE Full text] [doi: 10.1093/jnci/djn011] [Medline: 18270337]
26. Nørregaard C, Grønhøj C, Jensen D, Friborg J, Andersen E, von Buchwald C. Cause-specific mortality in HPV+ and HPV- oropharyngeal cancer patients: insights from a population-based cohort. Cancer Med 2018;7(1):87-94 [FREE Full text] [doi: 10.1002/cam4.1264] [Medline: 29171183]
27. SEER*Stat database: incidence—SEER research plus data (specialized head and neck fields), 18 registries, nov 2020 sub (2000-2018)—linked to county attributes—total U.S., 1969-2019 counties, National Cancer Institute, DCCPS, Surveillance Research Program, released april 2021, based on the november 2020 submission. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov). URL: https://seer.cancer.gov/data-software/documentation/seerstat/nov2020/ [accessed 2023-08-08]
28. Fullerton ZH, Butler SS, Mahal BA, Muralidhar V, Schoenfeld JD, Tishler RB, et al. Short-term mortality risks among patients with oropharynx cancer by human papillomavirus status. Cancer 2020;126(7):1424-1433 [FREE Full text] [doi: 10.1002/cncr.32652] [Medline: 31930488]
29. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007;26(11):2389-2430. [doi: 10.1002/sim.2712] [Medline: 17031868]
30. Gorphe P, Classe M, Ammari S, Garcia G, Even C, Casiraghi O, et al. Patterns of disease events and causes of death in patients with HPV-positive versus HPV-negative oropharyngeal carcinoma. Radiother Oncol 2022;168:40-45 [FREE Full text] [doi: 10.1016/j.radonc.2022.01.021] [Medline: 35093406]
31. Bosshart SL, Morand GB, Broglie MA. Frequency and localization of second primary tumors in patients with oropharyngeal carcinoma-the influence of the human papilloma virus. Cancers (Basel) 2021;13(8):1755 [FREE Full text] [doi: 10.3390/cancers13081755] [Medline: 33916999]
32. Culié D, Rousseau A, Pretet JL, Guily JLS. HPV status and therapeutic initial strategy impact on survival and oncologic outcomes: 5-year results from the multicentric prospective cohort of oropharyngeal cancers papillophar. Eur Arch Otorhinolaryngol 2022;279(6):3071-3078. [doi: 10.1007/s00405-021-07117-5] [Medline: 34661716]
33. Väisänen JA, Alho OP, Koivunen PT, Läärä E. Cause-specific mortality in patients with head and neck cancer: long-term follow-up of a population-based cohort from 1986 to 2012 accounting for competing risks. Oral Oncol 2018;79:20-26. [doi: 10.1016/j.oraloncology.2018.02.008] [Medline: 29598946]
34. Läärä E, Korpi JT, Pitkänen H, Alho OP, Kantola S. Competing risks analysis of cause-specific mortality in patients with oral squamous cell carcinoma. Head Neck 2017;39(1):56-62. [doi: 10.1002/hed.24536] [Medline: 27437667]
35. Rose BS, Jeong JH, Nath SK, Lu SM, Mell LK. Population-based study of competing mortality in head and neck cancer. J Clin Oncol 2011;29(26):3503-3509 [FREE Full text] [doi: 10.1200/JCO.2011.35.7301] [Medline: 21844503]
36. Shen W, Sakamoto N, Yang L. Cancer-specific mortality and competing mortality in patients with head and neck squamous cell carcinoma: a competing risk analysis. Ann Surg Oncol 2015;22(1):264-271. [doi: 10.1245/s10434-014-3951-8] [Medline: 25074664]
37. Kwon M, Roh JL, Song J, Lee SW, Kim SB, Choi SH, et al. Noncancer health events as a leading cause of competing mortality in advanced head and neck cancer. Ann Oncol 2014;25(6):1208-1214 [FREE Full text] [doi: 10.1093/annonc/mdu128] [Medline: 24669018]
38. Lop J, García J, López M, Taberna M, Mena M, Alemany L, et al. Competing mortality in oropharyngeal carcinoma according to human papillomavirus status. Head Neck 2019;41(5):1328-1334. [doi: 10.1002/hed.25559] [Medline: 30549389]
39. Petrelli F, Sarti E, Barni S. Predictive value of human papillomavirus in oropharyngeal carcinoma treated with radiotherapy: an updated systematic review and meta-analysis of 30 trials. Head Neck 2014;36(5):750-759. [doi: 10.1002/hed.23351] [Medline: 23606404]
40. Liu H, Li J, Zhou Y, Hu Q, Zeng Y, Mohammadreza MM. Human papillomavirus as a favorable prognostic factor in a subset of head and neck squamous cell carcinomas: a meta-analysis. J Med Virol 2017;89(4):710-725. [doi: 10.1002/jmv.24670] [Medline: 27575972]
41. Hess CB, Rash DL, Daly ME, Farwell DG, Bishop J, Vaughan AT, et al. Competing causes of death and medical comorbidities among patients with human papillomavirus-positive vs human papillomavirus-negative oropharyngeal carcinoma and impact on adherence to radiotherapy. JAMA Otolaryngol Head Neck Surg 2014;140(4):312-316 [FREE Full text] [doi: 10.1001/jamaoto.2013.6732] [Medline: 24526276]
42. Young L, Cho L. Unique cardiovascular risk factors in women. Heart 2019;105(21):1656-1660. [doi: 10.1136/heartjnl-2018-314268] [Medline: 31315936]
43. He J, Zhu Z, Bundy JD, Dorans KS, Chen J, Hamm LL. Trends in cardiovascular risk factors in US adults by race and ethnicity and socioeconomic status, 1999-2018. JAMA 2021;326(13):1286-1298 [FREE Full text] [doi: 10.1001/jama.2021.15187] [Medline: 34609450]

44. Sturgeon KM, Deng L, Bluethmann SM, Zhou S, Trifiletti DM, Jiang C, et al. A population-based study of cardiovascular disease mortality risk in US cancer patients. Eur Heart J 2019;40(48):3889-3897 [FREE Full text] [doi: 10.1093/eurheartj/ehz766] [Medline: 31761945]
45. Abramowitz L, Guily JLS, Moyal-Barracco M, Bergeron C, Borne H, Dahlab A, et al. Epidemiological and economic burden of potentially HPV-related cancers in France. PLoS One 2018;13(9):e0202564 [FREE Full text] [doi: 10.1371/journal.pone.0202564] [Medline: 30235216]
46. Daniel CL, Atkins A. HPV vaccination disparities and strategies to prevent oropharyngeal cancer in males. Prev Med Rep 2021;24:101521 [FREE Full text] [doi: 10.1016/j.pmedr.2021.101521] [Medline: 34466373]
47. Gray JL, Maghlouth AA, Hussain HA, Sheef MA. Impact of oral and oropharyngeal cancer diagnosis on smoking cessation patients and cohabiting smokers. Tob Induc Dis 2019;17:75 [FREE Full text] [doi: 10.18332/tid/109413] [Medline: 31768167]
48. Toman J, Von Larson S, Umeno H, Kurita T, Furusaka T, Hasegawa H, et al. HPV-positive oropharyngeal cancer via p16 immunohistochemistry in Japan. Ann Otol Rhinol Laryngol 2017;126(2):152-158. [doi: 10.1177/0003489416681582] [Medline: 27913709]
49. Husain S, Lohia S, Petkov V, Blackwell T, Swisher-McClure S, Mizrachi A, et al. Disparities and guideline adherence for HPV testing among patients with oropharyngeal squamous cell carcinoma, NCDB, and SEER. Head Neck 2021;43(7):2110-2123 [FREE Full text] [doi: 10.1002/hed.26679] [Medline: 33851469]

Abbreviations

AJCC: American Joint Committee on Cancer
COD: cause of death
HNC: head and neck cancer
HNCSM: head and neck cancer–specific mortality
HNSCC: head and neck squamous cell carcinoma
HPV: human papillomavirus
NCCM: noncancer-caused mortality
OM: overall mortality
OPSCC: oropharyngeal squamous cell carcinoma
SEER: Surveillance, Epidemiology, and End Result
sHR: subdistribution hazard ratio
SPCM: second primary cancer mortality

Edited by A Mavragani, T Sanchez; submitted 25.03.23; peer-reviewed by J Wang, M Khokhar; comments to author 19.04.23; revised version received 14.05.23; accepted 25.07.23; published 29.08.23.

Please cite as:
Zhang DD, Lei M, Wang Y, Zeng PJ, Hong YJ, Cai CF
Cause of Death in Patients with Oropharyngeal Carcinoma by Human Papillomavirus Status: Comparative Data Analysis
JMIR Public Health Surveill 2023;9:e47579
URL: https://publichealth.jmir.org/2023/1/e47579
doi:10.2196/47579
PMID:37642982

Original Paper

Reported Global Avian Influenza Detections Among Humans and Animals During 2013-2022: Comprehensive Review and Analysis of Available Surveillance Data

Christine M Szablewski[1], MPH, DVM; Chelsea Iwamoto[1], MPH; Sonja J Olsen[1], PhD; Carolyn M Greene[1], MD; Lindsey M Duca[1], PhD; C Todd Davis[1], PhD; Kira C Coggeshall[2], DVM; William W Davis[1], DrPH; Gideon O Emukule[1], PhD; Philip L Gould[1], MD; Alicia M Fry[1], MD; David E Wentworth[1], PhD; Vivien G Dugan[1], PhD; James C Kile[1], DVM; Eduardo Azziz-Baumgartner[1], MD

[1]Influenza Division, National Center for Immunization and Respiratory Diseases, United States Centers for Disease Control and Prevention, Atlanta, GA, United States

[2]Division of Global Health Protection, Global Health Center, United States Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Christine M Szablewski, MPH, DVM
Influenza Division
National Center for Immunization and Respiratory Diseases
United States Centers for Disease Control and Prevention
1600 Clifton Road NE
Atlanta, GA, 30333
United States
Phone: 1 470 998 1990
Email: LQZ9@cdc.gov

Abstract

Background: Avian influenza (AI) virus detections occurred frequently in 2022 and continue to pose a health, economic, and food security risk. The most recent global analysis of official reports of animal outbreaks and human infections with all reportable AI viruses was published almost a decade ago. Increased or renewed reports of AI viruses, especially high pathogenicity H5N8 and H5N1 in birds and H5N1, H5N8, and H5N6 in humans globally, have established the need for a comprehensive review of current global AI virus surveillance data to assess the pandemic risk of AI viruses.

Objective: This study aims to provide an analysis of global AI animal outbreak and human case surveillance information from the last decade by describing the circulating virus subtypes, regions and temporal trends in reporting, and country characteristics associated with AI virus outbreak reporting in animals; surveillance and reporting gaps for animals and humans are identified.

Methods: We analyzed AI virus infection reports among animals and humans submitted to animal and public health authorities from January 2013 to June 2022 and compared them with reports from January 2005 to December 2012. A multivariable regression analysis was used to evaluate associations between variables of interest and reported AI virus animal outbreaks.

Results: From 2013 to 2022, 52.2% (95/182) of World Organisation for Animal Health (WOAH) Member Countries identified 34 AI virus subtypes during 21,249 outbreaks. The most frequently reported subtypes were high pathogenicity AI H5N1 (10,079/21,249, 47.43%) and H5N8 (6722/21,249, 31.63%). A total of 10 high pathogenicity AI and 6 low pathogenicity AI virus subtypes were reported to the WOAH for the first time during 2013-2022. AI outbreaks in animals occurred in 26 more Member Countries than reported in the previous 8 years. Decreasing World Bank income classification was significantly associated with decreases in reported AI outbreaks (P<.001-.02). Between January 2013 and June 2022, 17/194 (8.8%) World Health Organization (WHO) Member States reported 2000 human AI virus infections of 10 virus subtypes. H7N9 (1568/2000, 78.40%) and H5N1 (254/2000, 12.70%) viruses accounted for the most human infections. As many as 8 of these 17 Member States did not report a human case prior to 2013. Of 1953 human cases with available information, 74.81% (n=1461) had a known animal exposure before onset of illness. The median time from illness onset to the notification posted on the WHO event information site was 15 days (IQR 9-30 days; mean 24 days). Seasonality patterns of animal outbreaks and human infections with AI viruses were very similar, occurred year-round, and peaked during November through May.

Conclusions: Our analysis suggests that AI outbreaks are more frequently reported and geographically widespread than in the past. Global surveillance gaps include inconsistent reporting from all regions and human infection reporting delays. Continued monitoring for AI virus outbreaks in animals and human infections with AI viruses is crucial for pandemic preparedness.

(JMIR Public Health Surveill 2023;9:e46383) doi:10.2196/46383

KEYWORDS

avian influenza; novel influenza; pandemic influenza; One Health; zoonotic influenza; surveillance

Introduction

Avian influenza (AI) viruses are a health, economic, and food security risk. While substantial H5, H7, and H9 virus detections in birds have occurred during the past several years, little is known about whether these detections are occurring more frequently or threatening a greater proportion of the world’s poultry, wild birds (especially threatened species), and human populations than in previous years.

AI viruses primarily affect avian species and are grouped as low pathogenicity AI (LPAI) or high pathogenicity AI (HPAI). The difference is based on the ability of the virus to cause mortality and morbidity in poultry or as assessed through genetic sequencing [1]. LPAI causes mild morbidity in poultry but lowers production and may circulate undetected in poultry populations. However, the virus can mutate (specifically H5 and H7 viruses) and become HPAI, which can cause severe disease and mortality in poultry [2]. Migration patterns of asymptomatic reservoirs such as wild waterfowl and shorebirds can lead to the transmission of AI virus to other wild birds and poultry, ultimately contributing to the global spread of AI viruses [3]. Global movement of poultry through trade can further lead to increased AI virus spread [2,4].

AI outbreaks in birds, especially HPAI in poultry, can have large economic implications. A World Bank report from 2008 estimated that global costs attributable to HPAI outbreaks could be as high as 0.7% of global gross domestic product, if AI viruses were to become enzootic throughout the world [5]. Even with control measures, outbreaks of AI viruses continue to occur globally, leading to millions of dead and culled birds, as well as increased chances of virus reassortment events in birds and nonhuman mammals such as pigs and mink [6-8]. HPAI outbreaks can lead to trade restrictions, expensive poultry products, food insecurity, and loss of livelihood [9].

Zoonotically transmitted AI viruses can cause high mortality and may recombine with human influenza A viruses to become more transmissible between humans [10,11]. Efficient and sustained transmission among humans could result in an influenza pandemic with high rates of illness and death [10]. Since 1997, there have been thousands of reports of AI virus infections in humans, most of which have been associated with the H5, H7, and H9 virus lineages [12-14].

A comprehensive review of current global AI virus surveillance information is key to assessing the pandemic risk of AI viruses. This includes reviewing the subtypes of AI viruses circulating during the past decade, their locations, what species are hosting them, and to determine if and why they are increasing in number and spread [15]. The most recent global analysis of official reports of AI outbreaks in birds and human infections with AI viruses was almost a decade ago [16]. In this investigation, contemporary AI virus reports were analyzed among animals and humans from January 2013 to June 2022 and compared with those from 2005 to 2012. AI outbreaks in animals and human AI virus infections were described, detailing the circulating virus subtypes, regions, and temporal trends in reporting animal outbreaks and human cases, and Member Country/Nation characteristics associated with animal AI outbreak reporting; additionally, surveillance and reporting gaps for animals and humans were identified.

Methods

Annual and Weekly Records of AI Outbreaks Among Animals

The World Organisation for Animal Health (WOAH) is responsible for improving animal health worldwide. During our timeframe of interest, the WOAH’s Terrestrial Animal Health Code required Member Countries to submit outbreak notifications of all HPAI viruses detected in domestic and wild birds and unusual species (such as mammals), all LPAI viruses of subtypes H5 and H7 detected in poultry, and unusual mortality events among wild birds [1]. The WOAH defines an outbreak as an occurrence of 1 or more cases in a group of animals with a defined epidemiologic relationship, such as a shared environment or common management practices. Each Member Country is responsible for reporting outbreaks to the WOAH. Immediate WOAH notification is not necessary for AI virus subtypes that Member Countries judge endemic [1]. The Food and Agriculture Organization of the United Nations (FAO) publishes monthly reports on all AI virus subtypes including outbreak information for HPAI H5N1 in endemic Member Nations and all LPAI H7N9 viruses [17].

The US Centers for Disease Control and Prevention (CDC) has monitored reports of AI virus outbreaks in animals since 2005. Information about animal outbreaks (in all nonhuman species) reported to the WOAH, FAO, and other official sources (eg, publications and local government websites) was compiled into an AI virus outbreak database that included the number of new outbreaks, by subtype, reported by each Member Country/Nation by report publication year. Starting in 2016, the weekly (based on report publication date) number of newly reported outbreaks, by subtype, from each Member Country/Nation and all animal categories in which that subtype had been reported since 2005 were also included. Animal categories were defined as wild birds (including aquatic and terrestrial waterfowl and shorebirds), zoo/captive birds (including exotic birds), birds in live bird markets, backyard/village birds (poultry such as chickens, ducks, turkeys, and geese), farm/commercial birds (poultry), and nonhuman mammals. All outbreak data are based on the number of outbreak events reported and not on the quantity of animals affected.

Outbreak data from January 1, 2013, to June 30, 2022, were used to determine annual totals for reported outbreaks; the number of Member Countries reporting, by subtype; and the predominant subtype. Subtypes that accounted for the most global outbreaks each year were considered the yearly predominant subtype. Weekly data from January 1, 2016, to June 30, 2022, were used to calculate the total number of outbreaks per subtype per month, as well as the number of newly reported animal categories. A new animal category was defined as a specific animal category, per subtype, reported by a Member Country for the first time since 2005. Results were compared with database information from 2005 to 2012. World Bank data were used to assign region categorization and income classifications to maintain consistency between animal and human health authorities [18]. WOAH data were used to assign the number of individual birds infected and culled [19]. The FAOSTAT was used to assign each Member Nation the percent of global poultry production and total number of annual live-bird imports [20]. The Wetlands International global flyway map was used to determine the number of major migratory flyways passing through each Member Country/Nation [21].

SAS (version 9.4; SAS Institute) was used to conduct univariate analyses of the relationship between independent variables of interest (World Bank income classification, yearly poultry production, total number of flyways through WOAH Member Countries, and total live-bird imports by Member Nations) and the dependent variables of interest (reported animal outbreaks). The regression analysis was restricted to 2013-2020, the years for which FAOSTAT poultry production data were available. Outbreaks from subtypes reported as endemic by a Member Country were excluded from the regression analyses. Variables of interest were assessed in the univariate model and covariates significant at an α level of .10 were included in the multivariable regression model. Relative rates were calculated using a Poisson regression and human population offset. A generalized estimating equation was used to account for correlations within Member Countries during the evaluation period. The final multivariable model was adjusted for World Bank income classification, yearly poultry production, total number of flyways, outbreak year, and total live-bird imports by WOAH Member Countries. Significance for the final model was assigned at $P \leq .05$.

AI Virus Infections Among Humans

The International Health Regulations (IHR) 2005 require Member States to report events of international public health importance to the World Health Organization (WHO) [22]. Human infections with influenza A viruses caused by subtypes that do not circulate in humans (novel) are notifiable [23]. The CDC maintains a line list of human infections caused by novel influenza virus subtypes as reported by the WHO per the IHR (2005) or other official sources (eg, ministries of health, FAO, publications); data were collected by weekly checks of these sources and include Member State of residence or diagnosis, symptom onset and report dates, outcome, exposure information, and reported human-to-human transmission (HTHT, which is defined as infection reported to have been caused by, or to have caused, an infection in another person). We analyzed all human AI virus infection notifications from January 1, 2013 to June 30, 2022.

The number of human cases was calculated by subtype, geographic location, animal exposure, age, reported HTHT, and reported case fatality proportion (rCFP). HTHT ratios (number of cases associated with HTHT divided by all reported infections) and rCFP (deaths noted at report time or in subsequent notifications divided by all reported infections) were calculated by subtype. Reporting lag was calculated as the difference in reported onset date and report date by the WHO in the event information site or other official source. United Nations data were used for population size calculations [24].

Ethical Considerations

This study was reviewed by the CDC and is consistent with federal law and CDC policy. The study was determined exempt from Institutional Review Board and Institutional Animal Care and Use Committee approval as it uses deidentified data.

Results

Annual AI Outbreaks Among Animals

Between 2013 and 2022, 95 of 182 WOAH Member Countries (52.2%), which currently produce 87.1% of the world's poultry [20], reported 21,249 AI virus outbreaks among animals associated with 34 AI virus subtypes (15 HPAI and 19 LPAI); HPAI accounted for 93.39% (n=19,844) of the reported outbreaks (Figure 1 and Table 1); 66% (39/59) of high-income and 31% (9/29) of low-income Member Countries reported an AI virus outbreak during 2013-2022 (Table 2). On average, 42 Member Countries reported AI virus outbreaks each year (65 in 2017 to 20 in 2013); 1 in 4 (26/95) Member Countries first reported AI virus outbreaks among animals in 2013-2022 (Figure 2). All but 6 reporting Member Countries were in the northern hemisphere (Table 2). High-income Member Countries reported the greatest proportion of outbreaks (14,336/21,249, 67.47%; Table 2). Overall, the European and Central Asian regions reported the greatest number of outbreaks (10,550/21,249, 49.65%), whereas Latin American and the Caribbean region reported the fewest number of outbreaks (188/21,249, 0.88%; Table 2 and Figure 2).

Figure 1. Global animal outbreaks (A) and human infections (B) with avian influenza (AI) virus by subtype, January 2013-June 2022.

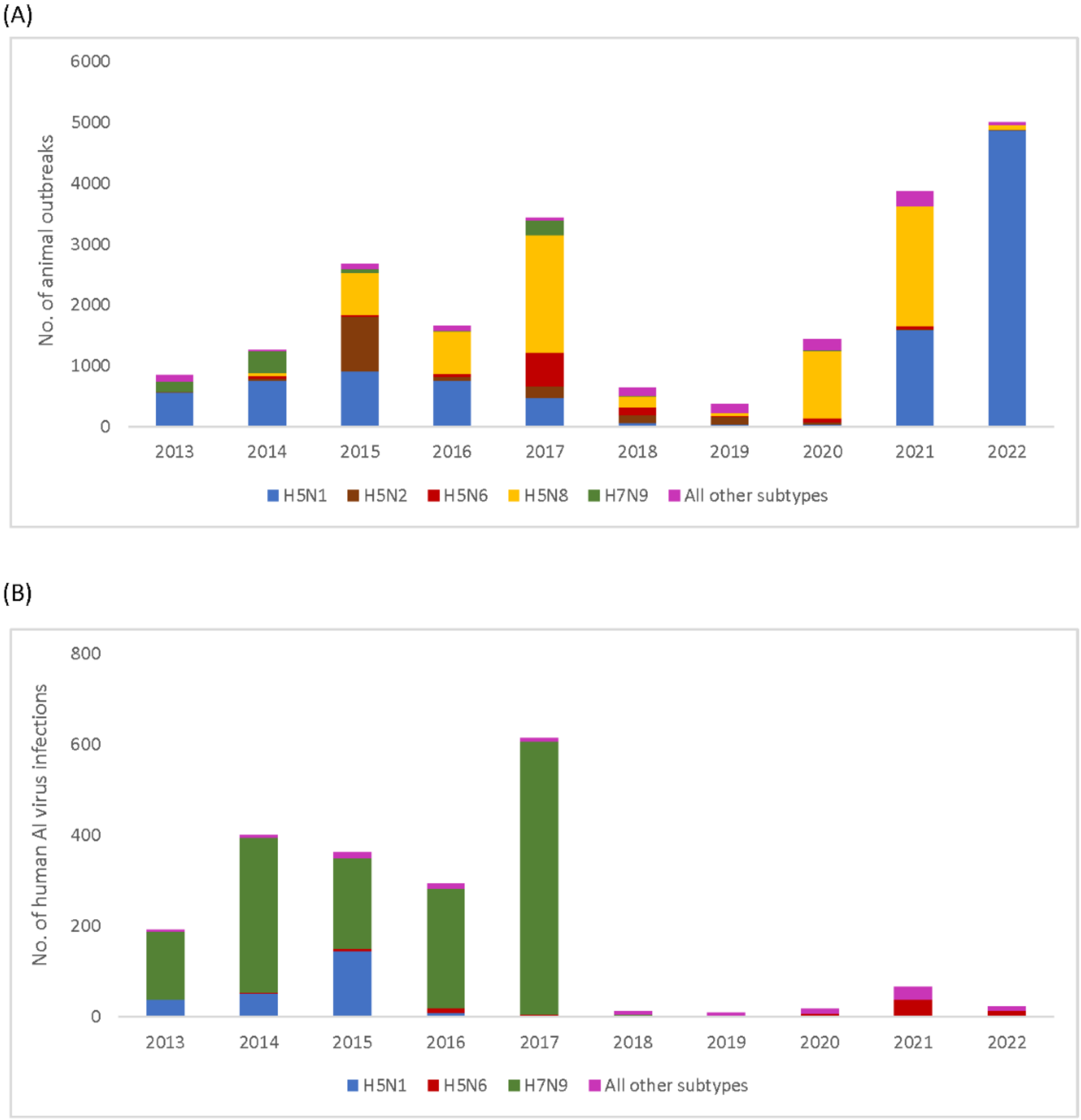


Table 1. Global avian influenza A virus subtypes reported in animal outbreaks from January 2013 to June 2022.

Subtype	Subtype reported in animals (N=21,249), n (%)
High pathogenicity avian influenza	19,844 (93.39)
H5N1	10,079 (47.43)
H5N2	1341 (6.31)
H5N3	41 (0.19)
H5N4	10 (0.05)
H5N5	179 (0.84)
H5N6	926 (4.36)
H5N8	6722 (31.63)
H5N9	25 (0.12)
H5Nx	333 (1.57)
H7N1	1 (0.00)
H7N2	2 (0.01)
H7N3	137 (0.64)
H7N7	15 (0.07)
H7N8	1 (0.00)
H7N9	32 (0.15)
Low pathogenicity avian influenza	1405 (6.61)
H3N1	74 (0.35)
H5N1	34 (0.16)
H5N2	211 (0.99)
H5N3	54 (0.25)
H5N5	1 (0.00)
H5N6	5 (0.02)
H5N8	2 (0.01)
H5N9	6 (0.03)
H5Nx	23 (0.11)
H6Nx	1 (0.00)
H7N1	11 (0.05)
H7N2	4 (0.02)
H7N3	25 (0.12)
H7N4	3 (0.01)
H7N6	7 (0.03)
H7N7	32 (0.15)
H7N8	8 (0.04)
H7N9	865 (4.07)
H7Nx	39 (0.18)

Table 2. WOAH[a] Member Countries reporting avian influenza outbreaks by region and income classification from January 2013 to June 2022.

Region	All WOAH Member Countries[b] (n=182)	Proportion of all WOAH Member Countries reporting outbreaks[c] (n=95)	Distribution of reporting Member Countries[d] (n=95)	Reported outbreaks (n=21,249)
Hemisphere, n (%)				
Northern	154 (84.6)	89 (57.8)	89 (93.7)	20,680 (97.3)
Southern	28 (15.4)	6 (21.4)	6 (6.3)	569 (2.7)
World Bank region, n (%)				
East Asia and Pacific	24 (13.2)	13 (54.2)	13 (13.7)	5815 (27.4)
Europe and Central Asia	51 (28.0)	40 (78.4)	40 (42.1)	10,550 (49.6)
Latin America and the Caribbean	29 (15.9)	4 (13.8)	4 (4.2)	188 (0.9)
Middle East and North Africa	20 (11.0)	10 (50.0)	10 (10.5)	1484 (7.0)
North America	2 (1.1)	2 (100.0)	2 (2.1)	1057 (5.0)
South Asia	8 (4.4)	6 (75.0)	6 (6.3)	537 (2.5)
Sub-Saharan Africa	48 (26.4)	20 (41.7)	20 (21.1)	1618 (7.6)
World Bank income classification, n (%)				
High income	59 (32.4)	39 (66.1)	39 (41.1)	14,336 (67.5)
Upper middle	47 (25.8)	23 (48.9)	23 (24.2)	3564 (16.8)
Lower middle	47 (25.8)	24 (51.1)	24 (25.3)	3233 (15.2)
Low income	29 (15.9)	9 (31.0)	9 (9.5)	116 (0.5)
Poultry production[e],%	99.9	87.1	N/A[f]	N/A
Population[g]	7,770,202,146	6,452,817,271 (83.0)	N/A	N/A

[a] WOAH: World Organisation for Animal Health.

[b] Denominator is all WOAH Member Countries.

[c] Denominator is all WOAH Member Countries in that hemisphere/region/income classification.

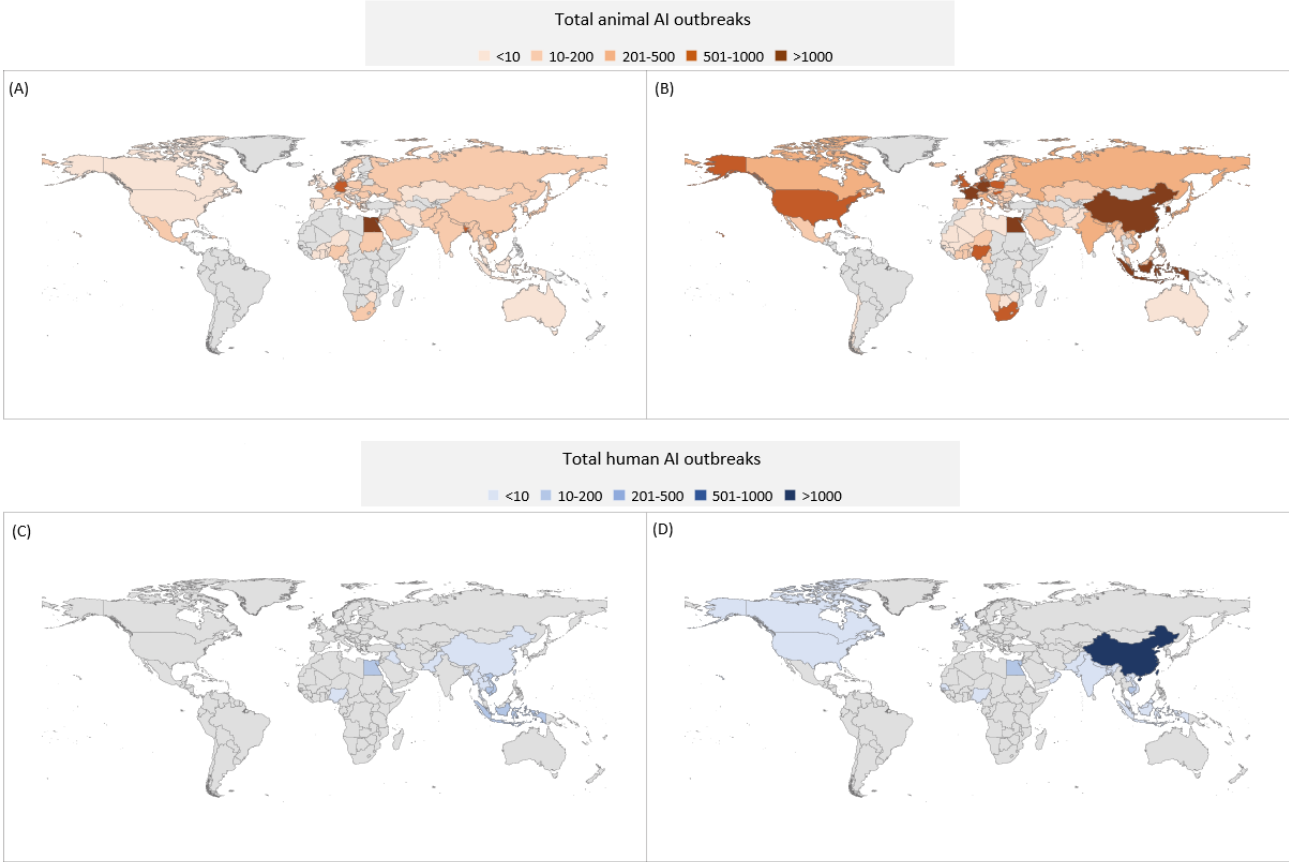
[d] Denominator is all WOAH Member Countries that reported an outbreak.

[e] Proportion of world poultry production based on most recently reported data from FAOSTAT for 2020.

[f] N/A: not applicable.

[g] Based on most recently reported data from 2021, Republic of Niue and Cook Islands are excluded from human population totals as data were not available.

Figure 2. World Organisation for Animal Health Member Countries reporting an animal avian influenza (AI) outbreak from January 2005 to December 2012 (A) and January 2013 to June 2022 (B) and World Health Organization Member States that reported human AI virus infections from January 2005 to December 2012 (C) and January 2013 to June 2022 (D).



A total of 10 HPAI and 6 LPAI virus subtypes were reported for the first time during 2013-2022. As many as 4 HPAI virus subtypes accounted for nearly 90% (n=19,068) of reported animal outbreaks: H5N1 (10,079/21,249, 47.43%), H5N8 (6722/21,249, 31.63%), H5N2 (1341/21,249, 6.31%), and H5N6 (926/21,249, 4.36%); 2 LPAI virus subtypes accounted for 5% (n=1076) of outbreaks: H7N9 (865/21,249, 4.07%) and H5N2 (211/21,249, 0.99%; Table 1). The predominant virus subtype shifted from H5N1 (767/1659, 46.23%, in 2016 to 480/3439, 13.96%, in 2017) to H5N8 (686/1659, 41.35%, in 2016 to 1924/3439, 55.95%, in 2017; Figure 1). H5N8 continued to be the predominant virus subtype in animal outbreaks each year since 2017, except during 2019 and 2022 (Figure 1). Overall, there were approximately 79,709,668 AI virus infections in birds and 323,884,584 birds were culled from 2013 to 2022 per the WOAH.

Results from the multivariable analysis for WOAH Member Countries (Table 3) that reported at least one outbreak from 2013 to 2022 indicated that decreased AI virus outbreak reporting among animals was associated with decreasing country income classification. Compared with high-income Member Countries, on average, upper-middle income countries reported 56% (adjusted relative rate 0.44, 95% CI 0.21-0.90) fewer animal outbreaks, lower-middle income countries reported 82% (adjusted relative rate 0.18, 95% CI 0.09-0.36) fewer animal outbreaks, and low-income countries reported 87% (adjusted relative rate 0.13, 95% CI 0.04-0.43) fewer animal outbreaks. Between January 2013 and June 2022, for every annual poultry production increase of 50 million head, there was a 2% decrease in reported AI outbreaks.

Table 3. Multivariable regression model[a] results for variables of interest associated with reported animal avian influenza outbreaks, January 2013-June 2022.

Independent variable	Crude relative rate	95% CI	*P* value	Adjusted relative rate	95% CI	*P* value
World Bank income classification						
High income	Reference[b]	Reference	Reference	Reference	Reference	Reference
Upper-middle income	0.24	0.12-0.45	<.001	0.44	0.21-0.90	.02
Lower-middle income	0.12	0.06-0.23	<.001	0.18	0.09-0.36	<.001
Low income	0.08	0.03-0.24	<.001	0.13	0.04-0.43	.001
Annual poultry import (per 50 million head)	1.09	1.00-1.18	.054	0.94	0.85-1.04	.23
Annual poultry production (per 50 million head)	0.98	0.97-0.99	<.001	0.98	0.98-0.99	<.001
Flyways						
1	Reference	Reference	Reference	Reference	Reference	Reference
2	0.72	0.35-1.47	.36	1.34	0.72-2.51	.36
≥3	0.18	0.09-0.36	<.001	0.67	0.32-1.38	.28

[a] Poisson distribution with a generalized estimating equation and human population offset adjusted for outbreak year.

[b] No data to be presented here because this is the reference group for the variable.

Weekly Reports of AI Outbreaks Among Animals

The WOAH received 2485 weekly reports from 2016 to 2022, comprising 15,240 AI virus outbreaks among animals from 91 Member Countries. On average, a median of 13 reports were submitted per week. Taiwan (208/2485, 8.37%) and Germany (176/2485, 7.08%) submitted the most weekly reports. The greatest number of reported animal outbreaks occurred during November 2021-June 2022 (6032/15,240, 39.58%) followed by November 2016-June 2017 (3423/15,240, 22.46%). While outbreaks occurred throughout the year, 87.17% (13,285/15,240) began during November-May (13,024/14,755, 88.27%, in the northern hemisphere Member Countries and 261/485, 53.8%, in the southern hemisphere Member Countries), and onset peaked in January and April in the northern hemisphere and January and November in the southern hemisphere Member Countries (Multimedia Appendix 1). The mean percentage of all outbreaks reported each month that are in a new animal category, for any given member-state, was 14% (range 0%-62%). Of the reports in new animal categories, wild bird (162/439, 36.9%), farm/commercial bird (112/439, 25.5%), and backyard/village bird (105/439, 23.9%) categories comprised most reports. Nonhuman mammals made up the least number of new reports at 1.6% (7/439) of new animal category reports.

AI Virus Infections Among Humans

During 2013-2022, 17 of WHO’s 194 Member States, which contain more than one-half (4,469,153,194/7,890,000,000, 56.64%) of the world’s population, reported 2000 AI virus infections in humans, which were associated with 10 AI virus subtypes (Figure 1). Both North American region Member States and 40% (4/10) of South Asian region Member States reported a human AI virus infection during 2013-2022 (Table 4). On average, 4 Member States reported human AI virus infections each year (range 2 in 2018 to 7 in 2014). Oman is the only Member State, with a reported human infection, that has never reported an AI outbreak among animals; 8/17 (47%) Member States first reported a human AI virus infection in 2013-2022 (Figure 2).

Table 4. WHO[a] Member States reporting human infection with avian influenza virus by region and income classification (period: January 2013-June 2022).

Region and income classification	All WHO Member States[b] (n=194)	Proportion of all WHO Member States reporting infections[c] (n=17)	Distribution of reporting Member States[d] (n=17)	Reported infections (n=2000)
Hemisphere, n (%)				
Northern	163 (84.0)	17 (10.4)	17 (100)	17 (100)
Southern	31 (16.0)	0 (0)	0 (0)	0 (0)
World Bank region, n (%)				
East Asia and Pacific	32 (16.5)	6 (18.8)	6 (35.3)	1786 (89.3)
Europe and Central Asia	50 (25.8)	1 (2.0)	1 (5.9)	1 (0.1)
Latin America and the Caribbean	33 (17.0)	0 (0)	0 (0)	0 (0)
Middle East and North Africa	20 (10.3)	2 (10.5)	2 (11.8)	195 (9.8)
North America	2 (1.0)	2 (100.0)	2 (11.8)	6 (0.3)
South Asia	9 (4.6)	4 (50.0)	4 (23.5)	8 (0.4)
Sub-Saharan Africa	48 (24.7)	2 (4.2)	2 (11.8)	4 (0.2)
World Bank income classification, n (%)				
High income	61 (31.4)	4 (6.6)	4 (23.5)	9 (0.5)
Upper middle	29 (14.9)	3 (10.3)	3 (17.6)	1743 (87.2)
Lower middle	50 (25.8)	10 (20.4)	10 (58.8)	248 (12.4)
Low income	55 (28.4)	0 (0)	0 (0)	0 (0)
Poultry production[e], %	96.9	67.8	N/A[f]	N/A
Population[g], n	7,853,022,689	4,469,153,194	N/A	N/A

[a] WHO: World Health Organization.

[b] Denominator is all WHO Member States.

[c] Denominator is all WHO Member States in that hemisphere/region/income classification.

[d] Denominator is all WHO Member States that reported an infection.

[e] Based on most recently reported data from 2020.

[f] N/A: not applicable.

[g] Based on most recently reported data from 2021, Republic of Niue and Cook Islands are excluded from human population totals as data were not available.

H7N9 (1568/2000, 78.40%) and H5N1 (254/2000, 12.70%) viruses accounted for the most human infections (Table 5). China reported the most H7N9 virus infections (1565/1568, 99.81%). Egypt reported the majority of non-H7N9 human AI virus infection cases (194/432, 44.9%), followed by China (170/432, 39.4%). The remaining 15 Member States were responsible for 15.7% (68/432) of cases (Figure 2). Overall, the East Asia and Pacific region reported the greatest number of human cases (1786/2000, 89.30%), followed by the Middle East and North Africa region (195/2000, 9.75%; Table 4). All Member States that reported human infections were in the northern hemisphere; 24% (4/17) were classified as high income and none were classified as low income (Table 4).

Table 5. Reported case fatality proportion and known animal exposure for human infections with avian influenza virus (January 2013-June 2022).

Subtype	Cases reported (n=2000), n (%)	Deaths reported (n=446), n (%)	Reported case fatality proportion[a], (%)	Known animal exposure[b] (n=1461), n (%)
H3N8	2 (0.1)	0 (0)	0	2 (100)
H5N1	254 (12.7)	89 (20.0)	35.0	219 (86.2)
H5N6	80 (4.0)	27 (6.1)	33.8	70 (87.5)
H5Nx	3 (0.2)	0 (0)	0	3 (100.0)
H7N2	2 (0.1)	0 (0)	0	0 (0)
H7N4	1 (0.1)	0 (0)	0	1 (100.0)
H7N9	1568 (78.4)	326 (73.1)	20.8	1106 (70.5)
H9N2	86 (4.3)	2 (0.4)	2.3	58 (67.4)
H10N3	1 (0.1)	0 (0)	0	0 (0)
H10N8	3 (0.2)	2 (0.4)	66.7	2 (66.7)

[a] Deaths noted at report time or in subsequent notifications divided by all reported infections.

[b] A total of 438 cases had a missing value for the animal exposure variable.

While human AI virus infections occurred throughout the year, 89.65% (1793/2000) had illness onset dates during December-May, with the frequency of onset typically peaking in January (Multimedia Appendix 1). Of the 1953 human infections with available exposure information, 1461 (74.81%) had a known animal exposure prior to illness onset and 101 (5.17%) had no known animal exposure (Table 5). Age and sex information was available for 1989/2000 cases (99.45%); 64.30% (1286/2000) of cases were in males. Most cases were in persons aged 18-64 years (1253/2000, 62.7%). The rCFP varied by virus subtype from 0% (H3N8; 0/2) to 67% (H10N8; 2/3). The overall rCFP was 22.30% (446/2000; Table 5). There were 58/2000 reports of confirmed or possible HTHT (2.90%). The median time from illness onset to the notification posted on the IHR (2005) event information site or other official source was 15 (IQR 9-31; mean 24) days.

Discussion

Principal Results and Comparison With Prior Work

Between 2013 and 2022, more AI virus outbreaks in animals were reported by WOAH Member Countries than in the previous 8 years, with 26 Member Countries reporting their first animal AI virus outbreak. From January 2013 to June 2022, there were reports of more than 21,000 AI virus outbreaks in animals and 2000 human infections with AI viruses globally. As many as 10 new HPAI virus subtypes (new hemagglutinin/neuraminidase [HA/NA] combinations) were identified in animals, representing an almost 2-fold increase in the number of subtypes identified during our study period compared with the previous 8 years. This likely reflects broadened viral diversity, increased reporting, and enhanced surveillance. Although there were 34 subtypes reported, only 4 accounted for 89.74% (19,068/21,249) of animal outbreaks; all 4 were of the HPAI H5 lineage. During the study period, the predominant subtype causing animal outbreaks changed from H5N1 to H5N8, for all years but 1. However, in 2022, H5N1 reemerged as the predominant subtype. From 2016 to 2022, about 14% of outbreaks each month occurred in a new animal category with wild birds representing the highest percentage (36.9%, 162/439). This finding suggests that in addition to spreading geographically, AI viruses are also moving into more animal categories. Increases in geographic and animal category reporting could be attributed to increased poultry production and commercial trade [25], increased exposure to wild birds through repeated annual migrations of infected birds, changes in migration patterns because of climate change or land conversion to agricultural production [4,26,27], viral genetic evolution, and improved AI virus awareness and surveillance capacity [28].

Seasonality patterns of animal AI virus outbreaks and human infections with AI viruses were very similar and occurred year-round; the frequency of both peaked during November through May. This finding likely reflects the dominance of reporting from the northern hemisphere countries. The seasonal concurrence of animal and human infections indicates the potential for animal-to-human transmission of AI and highlights the importance of improved monitoring of AI virus subtypes in humans and animals. Animal outbreaks might rise during periods when commercial bird distribution surges in preparation for the winter holidays and lunar new year [16]. This is consistent with our quantitative results which suggest that the risk of human infection increased for those with a known exposure to animals. These epidemic patterns suggest that animal and human health authorities might consider intensifying AI virus infection prevention and control measures immediately before the onset of the November-May increase in AI animal outbreaks [16].

While AI virus outbreaks among animals and human AI virus infections had similar epidemic periods, less than one-half (921/2000, 46.1%) of human cases occurred during years when the same HPAI virus subtype was reported among animals in the same country. In our data set, H9N2 was the third most reported AI virus infection in humans. However, LPAI H9 subtypes are not required to be reported to the WOAH. As human surveillance complements animal surveillance, AI virus detections in humans may serve as signals of AI virus activity among animals. The IHR (2005) stipulates reporting of human AI virus infection within 48-72 hours of identification. Although the time from laboratory detection to reporting was not evaluated (laboratory confirmation date was not routinely reported), the median time of onset to IHR (2005) notification was 15 days with about 26.1% (517/1976) of reports lagging by 30 or more days. Delayed reporting is especially concerning for AI viruses that become capable of efficient HTHT, as significant community transmission may occur by the time of reporting. Timely identification and notification of novel influenza viruses are vital for global pandemic influenza preparedness. The most effective AI virus surveillance would include comprehensive surveillance, inclusive of all subtypes, and timely reporting in domestic and wild birds, humans and, ideally, nonhuman mammals especially swine, as pigs can play a crucial role in novel influenza virus reassortment [6,28].

In our analysis, Member Countries with World Bank income classifications other than high income were less likely to report animal AI virus outbreaks. Latin America and the Caribbean region reported the fewest number of outbreaks. There are published reports of AI virus outbreaks in the Americas, especially Latin America [29]; however, these outbreaks were not reported to the WOAH, and therefore, global surveillance systems that rely on the WOAH reporting mechanism did not capture these outbreaks. A 2022 report from the WOAH noted that HPAI is nationally notifiable in only 73.4% of Member Countries and only 67.5% of Member Countries report having surveillance in place [30]. The lack of established surveillance systems or lack of reporting from these and other Member Countries exposes a gap in global AI virus surveillance. Additionally, the WOAH does not require Member Countries to report enzootic AI [1]. This can lead to routine undercounting of AI virus activity in those Member Countries and shows that while surveillance and reporting may have improved since 2012, there is still room for improvement.

Limitations

Our study is subject to at least four limitations. First, not all Member Countries/States report AI virus outbreaks among animals or human infections with AI viruses to the WOAH or WHO, respectively, because of inadequate surveillance systems or political and economic concerns about the impact of reporting on poultry trade [28,31]. Further, our multisource data acquisition process may not have identified all cases. Therefore, these reports are likely underestimates. Second, sequencing and clade data were not consistently reported or recorded for animal outbreaks; therefore, genetic relatedness of virus subtypes was not available for this analysis. Third, data collection for the yearly and weekly outbreak and human infection databases occurred manually over the years and is therefore subject to human error. Finally, observed increases in outbreak numbers and reporting Member States/Countries may be an artifact of enhanced surveillance and not necessarily wholly because of an increase in outbreak frequency over time.

Conclusions

AI virus surveillance from 2013 to 2022 identified 10 new HA/NA combinations of HPAI virus, AI virus in novel animal categories, and 26 more Member Countries reporting animal outbreaks than in the preceding 8 years. Although we are unable to fully quantify the magnitude of each animal outbreak from our data, outbreaks from January 2013 to June 2022 have resulted in the cumulative loss of over 325 million birds. Zoonotic viral transmission from animals to humans poses a continued public health threat because efficient, sustained transmission of AI viruses in humans could trigger the next pandemic [10]. Public and animal health leaders should encourage Member Countries to adhere to international standards put forth by the WOAH and report AI virus events as indicated [1]. Timely notifications from all Member States/Countries of novel influenza virus infections in humans and AI virus outbreaks in animals are central to pandemic preparedness and prevention. It is essential for global AI virus surveillance that WHO Member State and WOAH Member Country health officials ensure that AI virus reports are submitted within the required timeline by international standards. Commonly reported AI virus subtypes in humans were not required to be reported in animals, making One Health collaboration between animal and human surveillance systems even more crucial. Further research to explore causes for AI virus increases in animals might prove beneficial, including evaluation of commercial poultry and swine farm biosecurity, mechanisms of domestic poultry and swine exposure to wild migratory bird populations, laboratory studies to investigate phenotypic characteristics of currently circulating viruses, and the effects of climate and ecological degradation on migration patterns.

Acknowledgments

This analysis was funded by the US Centers for Disease Control and Prevention (US CDC). This project uses deidentified data and is Institutional Review Board/Institutional Animal Care and Use Committee (IACUC) exempt. All authors have met all 4 International Committee of Medical Journal Editors (ICMJE) authorship criteria. We thank Ray Arthur, Joshua Mott, Young M. Yoo, Jazmin Duque, Diane Gross, Natalie Olson, Francisco Palomeque, and Katherine Roguski for helping the Division track avian influenza outbreaks during the past decade. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the US CDC.

Data Availability

The data sets generated during or analyzed in this study may be made available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Seasonality of global avian influenza animal outbreaks and human avian influenza virus infections, January 2016 to June 2022.
[PNG File , 49 KB - publichealth_v9i1e46383_app1.png]

References

1. Thiermann AB. International standards: the World Organisation for Animal Health Terrestrial Animal Health Code. Rev Sci Tech 2015 Apr 01;34(1):277-281 [FREE Full text] [doi: 10.20506/rst.34.1.2340] [Medline: 26470463]
2. Aiello SE, Moses MA. The Merck Veterinary Manual, 11th Edition. Hoboken, NJ: Wiley; Jul 2016.
3. Olsen B, Munster VJ, Wallensten A, Waldenström J, Osterhaus AD, Fouchier RA. Global patterns of influenza a virus in wild birds. Science 2006 Apr 21;312(5772):384-388. [doi: 10.1126/science.1122438] [Medline: 16627734]
4. Kilpatrick AM, Chmura AA, Gibbons DW, Fleischer RC, Marra PP, Daszak P. Predicting the global spread of H5N1 avian influenza. Proc Natl Acad Sci U S A 2006 Dec 19;103(51):19368-19373 [FREE Full text] [doi: 10.1073/pnas.0609227103] [Medline: 17158217]
5. Burns A, van der Mensbrugghe D, Timmer H. Evaluating the economic consequences of avian influenza. World Bank. Washington DC: World Bank Group; 2006 Jun 01. URL: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/977141468158986545/evaluating-the-economic-consequences-of-avian-influenza [accessed 2021-04-07]
6. Chauhan RP, Gordon ML. Deciphering transmission dynamics and spillover of avian influenza viruses from avian species to swine populations globally. Virus Genes 2021 Dec 08;57(6):541-555 [FREE Full text] [doi: 10.1007/s11262-021-01873-6] [Medline: 34625868]
7. Floyd T, Banyard AC, Lean FZX, Byrne AMP, Fullick E, Whittard E, et al. Encephalitis and Death in Wild Mammals at a Rehabilitation Center after Infection with Highly Pathogenic Avian Influenza A(H5N8) Virus, United Kingdom. Emerg Infect Dis 2021 Nov;27(11):2856-2863 [FREE Full text] [doi: 10.3201/eid2711.211225] [Medline: 34670647]
8. Agüero M, Monne I, Sánchez A, Zecchin B, Fusaro A, Ruano MJ, et al. Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. Euro Surveill 2023 Jan;28(3):5-10 [FREE Full text] [doi: 10.2807/1560-7917.ES.2023.28.3.2300001] [Medline: 36695488]
9. Sonaiya E. Family poultry, food security and the impact of HPAI. World's Poultry Science Journal 2007 Mar 14;63(1):132-138. [doi: 10.1017/s0043933907001353]
10. Richard M, de Graaf M, Herfst S. Avian influenza A viruses: from zoonosis to pandemic. Future Virol 2014 May 01;9(5):513-524 [FREE Full text] [doi: 10.2217/fvl.14.30] [Medline: 25214882]
11. Kalthoff D, Globig A, Beer M. (Highly pathogenic) avian influenza as a zoonotic agent. Vet Microbiol 2010 Jan 27;140(3-4):237-245. [doi: 10.1016/j.vetmic.2009.08.022] [Medline: 19782482]
12. Lai S, Qin Y, Cowling BJ, Ren X, Wardrop NA, Gilbert M, et al. Global epidemiology of avian influenza A H5N1 virus infection in humans, 1997–2015: a systematic review of individual case data. The Lancet Infectious Diseases 2016 Jul;16(7):e108-e118. [doi: 10.1016/s1473-3099(16)00153-5] [Medline: 27211899]
13. Wang X, Jiang H, Wu P, Uyeki TM, Feng L, Lai S, et al. Epidemiology of avian influenza A H7N9 virus in human beings across five epidemics in mainland China, 2013–17: an epidemiological study of laboratory-confirmed case series. The Lancet Infectious Diseases 2017 Aug;17(8):822-832. [doi: 10.1016/s1473-3099(17)30323-7] [Medline: 28583578]
14. Song W, Qin K. Human-infecting influenza A (H9N2) virus: A forgotten potential pandemic strain? Zoonoses Public Health 2020 May 12;67(3):203-212. [doi: 10.1111/zph.12685] [Medline: 31930694]
15. Centers for Disease Control and Prevention (CDC). Influenza Risk Assessment Tool (IRAT):National Pandemic Strategy. CDC. 2020. URL: https://www.cdc.gov/flu/pandemic-resources/national-strategy/risk-assessment.htm [accessed 2020-06-01]
16. Durand LO, Glew P, Gross D, Kasper M, Trock S, Kim IK, et al. Timing of influenza A(H5N1) in poultry and humans and seasonal influenza activity worldwide, 2004-2013. Emerg Infect Dis 2015 Feb;21(2):202-208 [FREE Full text] [doi: 10.3201/eid2102.140877] [Medline: 25625302]
17. Food and Agriculture Organization of the United Nations (FAO). Empres-Plus. EMPRES-i. 2021. URL: https://empres-i.review.fao.org/#/ [accessed 2021-04-05]
18. World B. World Bank Country and Lending Group Income and Region Data. World Bank. 2021. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2021-04-10]
19. World OFAH. WOAH-WAHIS Quanitative Data. WOAH-WAHIS. URL: https://wahis.woah.org/#/dashboards/qd-dashboard [accessed 2021-04-07]
20. Food and Agriculture Organization of the United Nations. FAOstat Crops and Livestock Products. FAOstat. 2021. URL: http://www.fao.org/faostat/en/?#data/QA [accessed 2021-06-11]
21. Galbraith CA, Jones T, Kirby J, Mundkur T. A Review of Migratory Bird Flyaways and Priorities for Management (CMS Technical Series Publication No. 27). UNEP/CMS Secretariat. Bonn, Germany: Secretariat of the Convention on the

Conservation of Migratory Species of Wild Animals; 2014. URL: http://www.cms.int/sites/default/files/publication/CMS_Flyways_Reviews_Web.pdf [accessed 2023-07-31]

22. World Health Organization (WHO). International Health Regulations (2005) Third Edition. In: WHO. Geneva, Switzerland: WHO; 2016.
23. World Health Organization. Case definitions for the four diseases requiring notification in all circumstances under the international health regulations. WHO. Geneva, Switzerland: WHO; 2009. URL: https://www.who.int/publications/m/item/case-definitions-for-the-four-diseases-requiring-notification-to-who-in-all-circumstances-under-the-ihr-(2005) [accessed 2022-06-21]
24. United N. World Population Prospects-Population Division-United Nations. UN Department of Economic and Social Affairs Population Division. 2021. URL: https://population.un.org/wpp/ [accessed 2021-06-11]
25. Radin JM, Shaffer RA, Lindsay SP, Araneta MRG, Raman R, Fowler JH. International chicken trade and increased risk for introducing or reintroducing highly pathogenic avian influenza A (H5N1) to uninfected countries. Infect Dis Model 2017 Nov;2(4):412-418 [FREE Full text] [doi: 10.1016/j.idm.2017.09.001] [Medline: 30137719]
26. Rohr JR, Barrett CB, Civitello DJ, Craft ME, Delius B, DeLeo GA, et al. Emerging human infectious diseases and the links to global food production. Nat Sustain 2019 Jun 11;2(6):445-456 [FREE Full text] [doi: 10.1038/s41893-019-0293-3] [Medline: 32219187]
27. Carey C. The impacts of climate change on the annual cycles of birds. Philos Trans R Soc Lond B Biol Sci 2009 Nov 27;364(1534):3321-3330 [FREE Full text] [doi: 10.1098/rstb.2009.0182] [Medline: 19833644]
28. Machalaba CC, Elwood SE, Forcella S, Smith KM, Hamilton K, Jebara KB, et al. Global avian influenza surveillance in wild birds: a strategy to capture viral diversity. Emerg Infect Dis 2015 Apr;21(4):e1-e7 [FREE Full text] [doi: 10.3201/eid2104.141415] [Medline: 25811221]
29. Afanador-Villamizar A, Gomez-Romero C, Diaz A, Ruiz-Saenz J. Avian influenza in Latin America: A systematic review of serological and molecular studies from 2000-2015. PLoS One 2017 Jun 20;12(6):e0179573 [FREE Full text] [doi: 10.1371/journal.pone.0179573] [Medline: 28632771]
30. World Organisation for Animal Health. Implementation of WOAH standards: the Observatory Annual Report First edition 2022. World Organisation for Animal Health. Paris, France: World Organisation for Animal Health; 2022. URL: https://www.woah.org/en/what-we-do/standards/observatory/implementation-of-standards-the-observatory-annual-report/ [accessed 2023-07-31]
31. Vallat B, Pinto J, Schudel A. International organisations and their role in helping to protect the worldwide community against natural and intentional biological disasters. Rev Sci Tech 2006 Apr 01;25(1):163-172 [FREE Full text] [doi: 10.20506/rst.25.1.1645] [Medline: 16796046]

Abbreviations

AI: avian influenza
CDC: Centers for Disease Control and Prevention
FAO: Food and Agriculture Organization of the United Nations
HA: hemagglutinin
HPAI: high pathogenicity avian influenza
HTHT: human-to-human transmission
IHR (2005): International Health Regulations 2005
LPAI: low pathogenicity avian influenza
NA: neuraminidase
rCFP: reported case fatality proportion
WHO: World Health Organization
WOAH: World Organisation for Animal Health

Edited by A Mavragani, T Sanchez; submitted 17.02.23; peer-reviewed by A Hammond, S von Dobschuetz; comments to author 15.03.23; revised version received 25.04.23; accepted 26.06.23; published 31.08.23.

Please cite as:
Szablewski CM, Iwamoto C, Olsen SJ, Greene CM, Duca LM, Davis CT, Coggeshall KC, Davis WW, Emukule GO, Gould PL, Fry AM, Wentworth DE, Dugan VG, Kile JC, Azziz-Baumgartner E
Reported Global Avian Influenza Detections Among Humans and Animals During 2013-2022: Comprehensive Review and Analysis of Available Surveillance Data
JMIR Public Health Surveill 2023;9:e46383
URL: https://publichealth.jmir.org/2023/1/e46383
doi:10.2196/46383
PMID:37651182

Original Paper

Cancer and Atrial Fibrillation Comorbidities Among 25 Million Citizens in Shanghai, China: Medical Insurance Database Study

Mu Chen[1*], MD; Cheng Li[1*], MD; Jiwei Zhang[2*], PhD; Xin Cui[3], MS; Wenqi Tian[3], MSc; Peng Liao[4], PhD; Qunshan Wang[1], MD; Jian Sun[1], MD; Li Luo[4], PhD; Hong Wu[5], MD; Yi-Gang Li[1], MD

[1]Department of Cardiology, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

[2]Shanghai Key Laboratory of Compound Chinese Medicines, The Ministry of Education Key Laboratory for Standardization of Chinese Medicines, Institute of Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[3]Shanghai Health Statistics Center, Shanghai, China

[4]School of Public Health, Fudan University, Shanghai, China

[5]Shanghai Municipal Health Commission, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Yi-Gang Li, MD
Department of Cardiology, Xinhua Hospital
School of Medicine
Shanghai Jiao Tong University
1665 Kongjiang Rd, Yangpu District
Shanghai, 200092
China
Phone: 86 21 25077275
Email: liyigang@xinhuamed.com.cn

Abstract

Background: With population aging, the prevalence of both cancer and atrial fibrillation (AF) have increased. However, there is scarce epidemiological data concerning the comorbid state of cancer and AF in low- and middle-income countries, including China.

Objective: We aimed to evaluate the site-, sex-, and age-specific profiles of cancer and AF comorbidities in Chinese populations.

Methods: Data from the Shanghai Municipal Health Commission database between 2015 and 2020 were screened, covering all medical records of Shanghai residents with medical insurance. Site-specific cancer profiles were evaluated for the population with AF relative to the age- and sex-adjusted population of residents without AF. The sex distribution and peak age of cancer diagnosis were also assessed.

Results: A total of 25,964,447 adult patients were screened. Among them, 22,185 patients presented cancers comorbid with AF (median 77, IQR 67-82 years of age; men: n=13,631, 61.44%), while 839,864 presented cancers without AF (median 67, IQR 57-72 years of age; men: n=419,020, 49.89%), thus yielding a higher cancer prevalence among residents with AF (8.27%) than among those without AF (6.05%; $P<.001$). In the population with AF, the most prevalent cancer type was lung cancer, followed by colorectal, male genital organ, stomach, breast, liver, bladder, thyroid, leukemia, and esophageal cancers. AF was associated with an average of nearly 1.4-fold (prevalence ratio [PR] 1.37, 95% CI 1.35-1.38) increased prevalence of cancer after adjusting for age and sex. For site-specific analyses, an increased prevalence of cancer in the population with AF was observed in 20 of 21 cancer sites. This increased prevalence was most prominent for nonsolid tumors, including multiple myeloma (PR 2.56, 95% CI 2.28-2.87), leukemia (PR 1.73, 95% CI 1.57-1.90), and non-Hodgkin lymphoma (PR 1.59, 95% CI 1.43-1.77); intrathoracic malignancies, including mediastinum (PR 2.34, 95% CI 1.89-2.90), lung (PR 1.64, 95% CI 1.59-1.69), and esophageal cancers (PR 1.41, 95% CI 1.28-1.56); bone and soft tissue neoplasms (PR 1.56, 95% CI 1.37-1.77); and kidney cancer (PR 1.53, 95% CI 1.36-1.72). Cancer prevalence in the population with AF relative to that in the population without AF was higher in men than in women in 14 of 18 cancer sites, and female predominance was only observed for thyroid cancer. The peak age of index cancer diagnosis was lower in the population with AF (age group: 70-74 years) than in that without AF (age group: 75-79 years), especially for specific cancer types, including thyroid, central nervous system, mediastinum, esophageal, bladder, and biliary cancers.

Conclusions: Patients with AF are associated with increased prevalence, heightened male predominance, and younger peak age of cancer. Further studies are needed to determine whether early screening of specific cancers is cost-effective and beneficial for patients with AF.

(JMIR Public Health Surveill 2023;9:e40149) doi:10.2196/40149

KEYWORDS

cardiovascular; malignancy; arrhythmia; cardio-oncology; prevalence; epidemiology; atrial fibrillation

Introduction

Atrial fibrillation (AF) is the most prevalent arrhythmia and affects approximately 1% of the general population [1,2]. Besides its relation to cardiovascular complications and mortality, increased risk of various noncardiovascular diseases, including cancers, was evident in a population with AF [3]. Cancer is the leading cause of noncardiovascular deaths and is associated with increased major bleeding, intracranial hemorrhage, and arguably thromboembolic events in patients with AF [4-6]. On the other hand, new-onset AF was related to a poorer prognosis in patients with cancer [7-9]. With rapid population aging and improved survival of both diseases, AF and cancer comorbidities are becoming increasingly clinically important and prevalent and should gain more attention from both cardiologists and oncologists.

Due to the asymptomatic nature of early-stage cancer and the intermittent onset of paroxysmal AF, diagnoses of both cancer and AF may experience years of latency, especially in low- and middle-income countries (LMICs) [10,11]. Wearable devices and mobile health might largely improve the screening efficiency and early detection of both cancer and AF but are not widely adopted currently [12-14]. Consequently, the time of diagnosis of AF and cancer does not reflect the true onset time of either disease, and preclinical stages, which may involve carcinoma in situ or atrial high-rate episodes, can last years. In addition, common risk factors (eg, smoking, obesity, and aging) and related molecular pathways (eg, chronic inflammation and autonomous dysregulation) might also contribute to the causality of the co-occurrence of AF and cancer [15]. Taken together, index diagnosis of AF in patients with cancer or cancer in patients with AF based on the temporal relationship between the two disorders might contain nonnegligible selection or lead-time bias [3,5,7,8]. Such potential bias was indicated by the fact that the highest diagnosis rate of new-onset AF occurred in the first 60 days after cancer diagnosis [9]. This was attributed, at least partially, to the intensive cardiovascular examinations of newly diagnosed patients with cancer.

The prevalence of both AF and cancer, as well as site-specific cancer types, varied across different regions, races, and ethnicities [16]. Consequently, the prevalence of AF-cancer comorbidities should also present with regional, racial, and ethnic differences. There is scarce epidemiological data concerning the comorbid state of cancer and AF, especially in LMICs, including China. Taking advantage of the government-issued database containing all sources of medical records from 25 million Shanghai citizens [1], this study investigated the site-, sex-, and age-specific profiles of comorbid cancer in the population with AF in Shanghai, thus providing insights regarding cancer and AF comorbidities in the Chinese population.

Methods

Database and Study Population

Data were retrieved from the Shanghai Municipal Health Commission database from 2015 to 2020 [1]. The database was administered by the Shanghai municipal government and covered all sources of medical records in Shanghai from 2346 medical institutions, including information on inpatient and outpatient visits at all hospitals and clinics and routine primary and family care. With health insurance coverage of over 96%, the local population with medical insurance largely represents the residents of Shanghai [1]. The government-instituted universal health insurance included the Urban Residents' Basic Medical Insurance scheme, the Urban Employee Basic Medical Insurance scheme, and the New Rural Cooperative Medical scheme. Patients without local medical insurance were excluded to avoid the inclusion of nonlocal patients who sought temporary medical services in Shanghai. Pediatric patients were also excluded. Adult patients (aged ≥20 years) in the database were screened for cancer and AF. Patients with cancer and AF comorbidities were defined as the patients with index diagnoses of cancer and AF between 2015 and 2020, regardless of the temporal order of the index date between the two diagnoses.

Diagnoses were recorded using the *International Statistical Classification of Diseases, Tenth Revision* (*ICD-10*). AF was identified by *ICD-10* codes I48.01-03 and O99.418. Cancers in general (*ICD-10* codes: C00-C97) and of individual systems were screened in patients with and without AF. A total of 21 types of cancer were listed as follows: esophageal cancer (C15), stomach cancer (C16), colorectal cancer (C18 to C21), liver cancer (C22), biliary cancer (C23 and C24), pancreatic cancer (C25), head and neck cancer (C00 to C14 and C30 to C32), lung cancer (C34), malignant neoplasm of mediastinum (C37 and C38), malignant neoplasm of bone and soft tissue (C40, C41, and C45 to C49), melanoma (C43), breast cancer (C50), gynecologic cancer (C51 to C58), male genital cancer (C60 to C63), renal cancer (C64), bladder cancer (C67), central nervous system (CNS) cancer (C69 to C72), thyroid cancer (C73), non-Hodgkin lymphoma (C82 to C85), multiple myeloma (C90), and leukemia (C91 to C95). Carcinoma in situ (D00 to D09) and borderline tumors (D37 to D48) were not included. The detailed *ICD-10* codes are listed in Table S1 in Multimedia Appendix 1. For analyses of sex disparities, gynecologic, breast, and male genital organ cancers were not included. China has released a nationwide regulation for cancer registration, which requires health care systems to report cancer cases since 2015 [17]. Therefore, cancer diagnosis data in this study was assumed to be reliable because the Chinese government provides financial support for patients with cancer-related *ICD-10* codes on the basis of clinical pathologic assessments.

Statistical Analysis

Data are presented as median (IQR) or as absolute values and percentages. The prevalence of comorbid cancer in the population with AF was calculated from the number of patients with cancer and AF divided by the total number of patients with AF. Due to different age and sex distributions between populations with and without AF, standardization was conducted for patients without AF according to the age and sex distributions of patients with AF (the numbers of patients before and after the adjustment are shown in Table S2 in Multimedia Appendix 1). The prevalence of cancer in patients without AF was subsequently calculated from the number of sex- and age-adjusted patients without AF who had cancer divided by the total number of sex- and age-adjusted patients without AF. The sex-specific cancer prevalence in patients with and without AF was also analyzed and compared. Pearson chi-square tests or Fisher precision probability tests were used to analyze the differences in the prevalence of various cancer types between populations with and without AF. The relative prevalence of cancer was presented as the prevalence ratio (PR) and 95% CI. *P*<.05 was considered statistically significant. All analyses were performed with SPSS 22.0 (IBM Corp).

Ethical Considerations

This study was reviewed and approved by the ethics committee of Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University (XHEC-D-2022-043). Informed consent was exempted due to the retrospective design and the anonymized and deidentified patient records.

Results

Baseline Characteristics

A total of 25,964,447 adult patients (median 47, IQR 32-62 years of age; men: n=12,649,586, 48.72%) were included in the analyses (Table S3 in Multimedia Appendix 1). Among them, 22,185 patients presented with comorbid AF and cancer (median 77, IQR 67-82 years of age; men: n=13,631, 61.44%). Cancers were found in 839,864 individuals without AF (median 67, IQR 57-72 years of age; men: n=419,020, 49.89%), while AF was observed in 246,216 patients without cancer (median 77, IQR 67-87 years of age; men: n=118,063, 47.95%). As shown in Table 1, despite a uniformly higher prevalence of cancer in patients with AF, the top 10 cancers in the populations with and without AF were essentially identical, with the exceptions of leukemia in the population with AF and pancreatic cancer in the population without AF. The most prevalent cancer in the population with AF was lung cancer, followed by colorectal, male genital organ, stomach, breast, liver, bladder, thyroid, leukemia, and esophageal cancers.

Table 1. Prevalence of different cancers in patients with and without atrial fibrillation (AF).

Cancer site	Patients with AF (n=268,401), n (%)	Patients without AF (n=3,954,796[a]), n (%)
Lung	4572 (1.70)	41,168 (1.04)
Colorectum	2389 (0.89)	27,976 (0.71)
Male genital organs	1588 (0.59)	21,093 (0.53)
Stomach	1529 (0.57)	17,982 (0.45)
Breast	1032 (0.38)	12,427 (0.31)
Liver	774 (0.29)	8247 (0.21)
Bladder	593 (0.22)	7313 (0.18)
Thyroid	565 (0.21)	6927 (0.18)
Leukemia	473 (0.18)	4037 (0.10)
Esophagus	438 (0.16)	4570 (0.12)
Pancreas	424 (0.16)	5359 (0.14)

[a] Data from patients without AF were age- and sex-adjusted.

Increased Prevalence of Cancer in Patients With AF

As shown in Table 2, the prevalence of cancer in patients with AF (8.27% in total; 10.35% in men and 6.26% in women) was higher than that in patients without AF (6.05% in total; 7.21% in men and 4.89% in women; *P*<.001 for all comparisons). AF was associated with a nearly 1.4-fold higher prevalence of cancer after adjusting for age and sex (PR 1.37, 95% CI 1.35-1.38; Figure 1). A consistently increased prevalence of comorbidity with AF was observed in 20 of 21 cancer sites, with the only exception being CNS cancers (PR 0.94, 95% CI 0.71-1.25). The strongest association was observed in multiple myeloma (PR 2.56, 95% CI 2.28-2.87), followed by mediastinum cancer, leukemia, lung cancer, bone and soft tissue neoplasms, kidney cancer, and esophageal cancer.

Table 2. Prevalence of different cancers in men and women with and without atrial fibrillation (AF).

Cancer site	All patients				Men				Women			
	With AF, %	Without AF[a], %	Chi-square	*P* value[b]	With AF, %	Without AF, %	Chi-square	*P* value	With AF, %	Without AF, %	Chi-square	*P* value
All sites	8.27	6.05	2118.27	<.001	10.35	7.21	1785.11	<.001	6.26	4.89	506.75	<.001
Head and neck	0.23	0.20	10.21	.001	0.32	0.28	6.99	.008	0.14	0.12	3.53	.06
Esophagus	0.16	0.12	48.15	<.001	0.23	0.15	42.32	<.001	0.10	0.08	8.27	.004
Stomach	0.57	0.45	72.26	<.001	0.78	0.60	66.45	<.001	0.36	0.31	11.07	.001
Colorectum	0.89	0.71	117.52	<.001	1.10	0.82	114.96	<.001	0.69	0.60	17.76	<.001
Liver	0.29	0.21	75.17	<.001	0.38	0.27	61.96	<.001	0.20	0.15	16.01	.001
Biliary tract	0.12	0.09	16.39	.001	0.11	0.09	9.43	.002	0.12	0.10	7.15	.008
Pancreas	0.16	0.14	9.28	.002	0.16	0.14	2.04	.15	0.16	0.13	8.92	.003
Lung	1.70	1.04	1029.57	<.001	2.33	1.28	1010.04	<.001	1.10	0.81	138.10	<.001
Mediastinum	0.04	0.02	64.44	<.001	0.05	0.02	47.38	<.001	0.03	0.01	18.55	<.001
Bone and soft tissue	0.09	0.06	45.43	<.001	0.11	0.07	29.22	<.001	0.08	0.05	17.45	<.001
Skin	0.03	0.02	7.39	.007	0.03	0.02	6.08	.01	0.03	0.02	1.99	.16
Breast	0.38	0.31	39.07	<.001	0.02	0.05	17.51	<.001	0.73	0.57	58.82	<.001
Kidney	0.12	0.08	52.26	<.001	0.16	0.10	38.95	<.001	0.07	0.05	14.61	.001
Bladder	0.22	0.18	17.46	<.001	0.36	0.29	22.26	<.001	0.09	0.09	0.06	.81
Gynecology	0.14	0.10	25.16	<.001	N/A[c]	N/A	N/A	N/A	0.28	0.21	25.16	<.001
Male genital organs	0.59	0.53	15.99	.001	1.21	1.02	40.54	<.001	N/A	N/A	N/A	N/A
Thyroid	0.21	0.18	17.74	<.001	0.13	0.11	2.23	.14	0.29	0.24	16.55	<.001
Non-Hodgkin lymphoma	0.14	0.09	76.72	<.001	0.17	0.11	44.94	<.001	0.12	0.07	32.93	<.001
Multiple myeloma	0.13	0.05	277.42	<.001	0.16	0.06	202.11	<.001	0.09	0.04	82.89	<.001
Leukemia	0.18	0.10	129.55	<.001	0.21	0.12	84.64	<.001	0.14	0.08	45.90	<.001
CNS[d]	0.02	0.02	0.18	.67	0.03	0.02	1.00	.32	0.01	0.02	2.73	.10

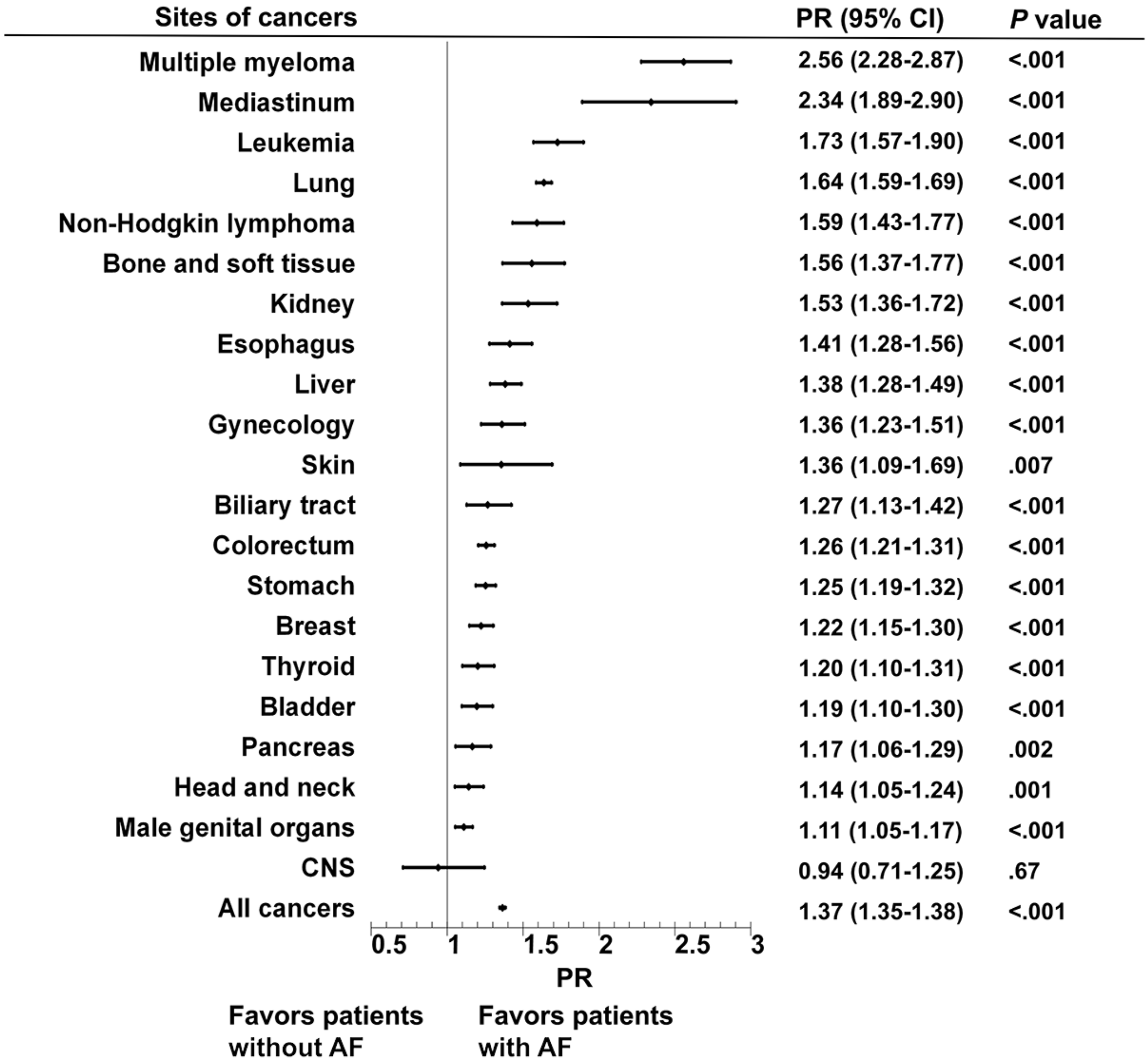
[a] Data from patients without AF were age-adjusted.

[b] *P*<.05 was considered statistically significant.

[c] N/A: not applicable.

[d] CNS: central nervous system.

Figure 1. Increased cancer prevalence in patients with AF. Plot showing relative prevalence of site-specific cancers in patients with and without AF, ordered from largest to smallest relative prevalence. AF: atrial fibrillation; CNS: central nervous system; PR: prevalence ratio.



Sites of cancers	PR (95% CI)	*P* value
Multiple myeloma	2.56 (2.28-2.87)	<.001
Mediastinum	2.34 (1.89-2.90)	<.001
Leukemia	1.73 (1.57-1.90)	<.001
Lung	1.64 (1.59-1.69)	<.001
Non-Hodgkin lymphoma	1.59 (1.43-1.77)	<.001
Bone and soft tissue	1.56 (1.37-1.77)	<.001
Kidney	1.53 (1.36-1.72)	<.001
Esophagus	1.41 (1.28-1.56)	<.001
Liver	1.38 (1.28-1.49)	<.001
Gynecology	1.36 (1.23-1.51)	<.001
Skin	1.36 (1.09-1.69)	.007
Biliary tract	1.27 (1.13-1.42)	<.001
Colorectum	1.26 (1.21-1.31)	<.001
Stomach	1.25 (1.19-1.32)	<.001
Breast	1.22 (1.15-1.30)	<.001
Thyroid	1.20 (1.10-1.31)	<.001
Bladder	1.19 (1.10-1.30)	<.001
Pancreas	1.17 (1.06-1.29)	.002
Head and neck	1.14 (1.05-1.24)	.001
Male genital organs	1.11 (1.05-1.17)	<.001
CNS	0.94 (0.71-1.25)	.67
All cancers	1.37 (1.35-1.38)	<.001

Sex Disparities in Cancer Prevalence

The increased prevalence of various cancer types in the population with AF was further analyzed regarding sex. As shown in Table 3, the strongest association between AF and cancer was found in multiple myeloma (in men: PR 2.73, 95% CI 2.36-3.16; in women: PR 2.32, 95% CI 1.93-2.80) and mediastinum cancer (in men: PR 2.50, 95% CI 1.91-3.28; in women: PR 2.12, 95% CI 1.49-3.01) for both sexes. The increased AF prevalence in patients with lung cancer ranked third in men and seventh in women. The increased AF prevalence in patients with gynecologic cancer ranked eighth in women.

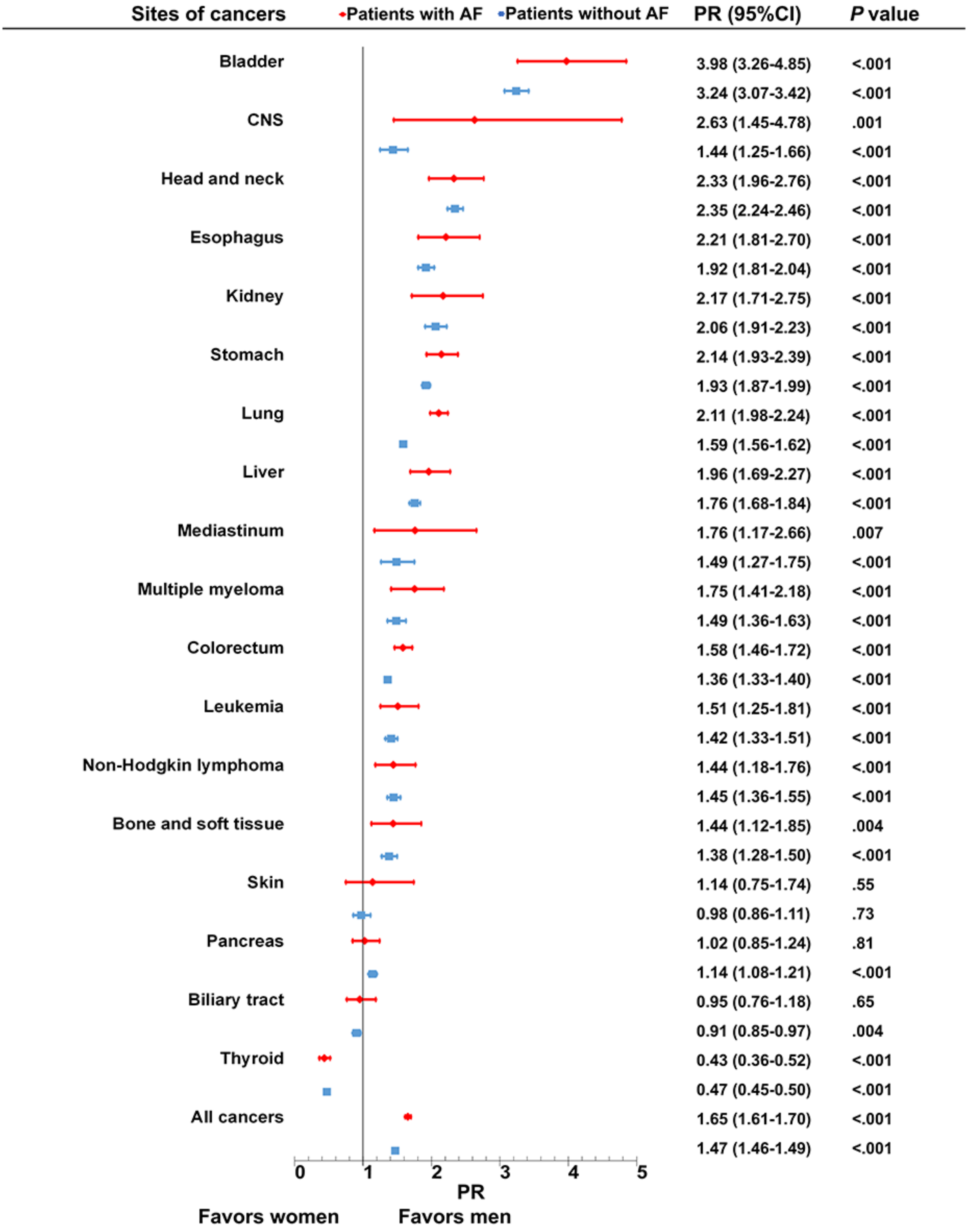
Table 3. Atrial fibrillation (AF)–related cancer prevalence in men and women.

Cancer site	Men		Women	
	PR[a] (95% CI)	*P* value	PR (95% CI)	*P* value
Multiple myeloma	2.73 (2.36-3.16)	<.001	2.32 (1.93-2.80)	<.001
Mediastinum	2.50 (1.91-3.28)	<.001	2.12 (1.49-3.01)	<.001
Lung	1.81 (1.75-1.88)	<.001	1.37 (1.30-1.44)	<.001
Leukemia	1.77 (1.57-2.00)	<.001	1.66 (1.43-1.93)	<.001
Non-Hodgkin lymphoma	1.59 (1.39-1.83)	<.001	1.60 (1.36-1.88)	<.001
Bone and soft tissue	1.59 (1.34-1.88)	<.001	1.53 (1.25-1.87)	<.001
Kidney	1.56 (1.36-1.80)	<.001	1.49 (1.21-1.83)	<.001
Esophagus	1.48 (1.31-1.67)	<.001	1.29 (1.08-1.53)	.004
Skin	1.47 (1.08-1.99)	.01	1.26 (0.91-1.73)	.16
Liver	1.44 (1.31-1.57)	<.001	1.29 (1.14-1.46)	<.001

[a] PR: prevalence ratio.

As shown in Figure 2, cancer prevalence showed a male predominance in the population without AF (PR 1.47, 95% CI 1.46-1.49), which was further strengthened in the population with AF (PR 1.65, 95% CI 1.61-1.70). In the population with AF, the prevalence of cancer was higher in men than in women at 14 of 18 cancer sites (gynecologic, breast, and male genital organ cancers were excluded). Among them, a male predominance of over a 2:1 male to female ratio was found in bladder, CNS, head and neck, esophageal, kidney, stomach, and lung cancers. Cancer risk in the population with AF was similar for skin, pancreas, and biliary cancers between sexes, while a female predominance was found in thyroid cancer.

Figure 2. Sex disparities in cancer prevalence among different cancer types in patients with and without AF. Plot showing relative prevalence of site-specific cancers comparing men to women among patients with (red) and without (blue) AF. The cancer types are ranked from largest to smallest relative risk in patients with AF. A total of 18 cancer sites were analyzed. Gynecologic cancer, breast cancer, and cancer of male genital organs (mainly prostate cancer) were not included in the analyses for sex disparities. AF: atrial fibrillation; CNS: central nervous system; PR: prevalence ratio.



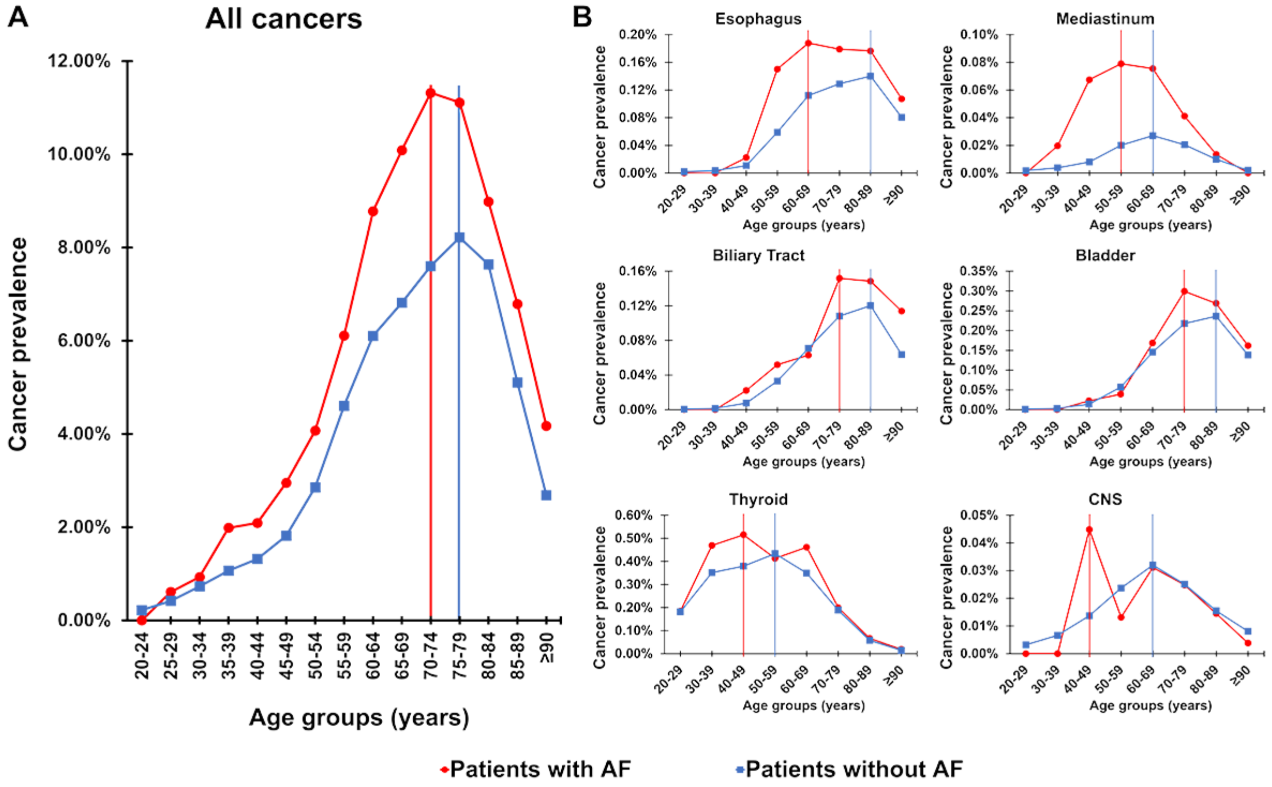
Sites of cancers	PR (95%CI)	*P* value
Bladder	3.98 (3.26-4.85)	<.001
	3.24 (3.07-3.42)	<.001
CNS	2.63 (1.45-4.78)	.001
	1.44 (1.25-1.66)	<.001
Head and neck	2.33 (1.96-2.76)	<.001
	2.35 (2.24-2.46)	<.001
Esophagus	2.21 (1.81-2.70)	<.001
	1.92 (1.81-2.04)	<.001
Kidney	2.17 (1.71-2.75)	<.001
	2.06 (1.91-2.23)	<.001
Stomach	2.14 (1.93-2.39)	<.001
	1.93 (1.87-1.99)	<.001
Lung	2.11 (1.98-2.24)	<.001
	1.59 (1.56-1.62)	<.001
Liver	1.96 (1.69-2.27)	<.001
	1.76 (1.68-1.84)	<.001
Mediastinum	1.76 (1.17-2.66)	.007
	1.49 (1.27-1.75)	<.001
Multiple myeloma	1.75 (1.41-2.18)	<.001
	1.49 (1.36-1.63)	<.001
Colorectum	1.58 (1.46-1.72)	<.001
	1.36 (1.33-1.40)	<.001
Leukemia	1.51 (1.25-1.81)	<.001
	1.42 (1.33-1.51)	<.001
Non-Hodgkin lymphoma	1.44 (1.18-1.76)	<.001
	1.45 (1.36-1.55)	<.001
Bone and soft tissue	1.44 (1.12-1.85)	.004
	1.38 (1.28-1.50)	<.001
Skin	1.14 (0.75-1.74)	.55
	0.98 (0.86-1.11)	.73
Pancreas	1.02 (0.85-1.24)	.81
	1.14 (1.08-1.21)	<.001
Biliary tract	0.95 (0.76-1.18)	.65
	0.91 (0.85-0.97)	.004
Thyroid	0.43 (0.36-0.52)	<.001
	0.47 (0.45-0.50)	<.001
All cancers	1.65 (1.61-1.70)	<.001
	1.47 (1.46-1.49)	<.001

Peak Age of Cancer Diagnosis in Patients With AF

Overall, there was a younger peak age of cancer diagnosis in the population with AF (range 70-74 years) than in that without AF (range 75-79 years; Figure 3A). For specific cancer types, a younger peak age of cancer diagnosis was observed in the population with AF than in that without AF, including thyroid (with AF: range 40-49 years; without AF: range 50-59 years), CNS (with AF: range 40-49 years; without AF: range 60-69 years), mediastinum (with AF: range 50-59 years; without AF: range 60-69 years), esophageal (with AF: range 60-69 years; without AF: range 80-89 years), bladder (with AF: range 70-79 years; without AF: range 80-89 years), and biliary (with AF: range 70-79 years; without AF: range 80-89 years) cancers

(Figure 3B). In other types of cancers, the peak age of cancer diagnosis was similar between the populations with and without AF (Figure S1 in Multimedia Appendix 1).

Figure 3. Peak age of cancer in general (A) and specific cancer types (B) in patients with and without AF. AF: atrial fibrillation; CNS: central nervous system.



Discussion

Main Findings

In this whole population–based medical insurance database study, we reported the prevalence of the AF-cancer comorbid state in Shanghai, the largest city in China. The main findings were as follows: (1) the most prevalent cancer types were essentially identical between the populations with and without AF; (2) cancer prevalence was higher in 20 of 21 cancer types in patients with AF compared to those without AF; (3) prevalence of nonsolid tumors, intrathoracic malignancies, bone and soft tissue cancers, and kidney cancer were ≥1.4-fold higher in the population with AF; (4) male predominance of cancer was more prominent in the population with AF; and (5) a younger peak age of the index cancer diagnosis was observed in the population with AF, especially for esophageal, mediastinum, thyroid, bladder, biliary, and CNS cancers.

Increased Prevalence of Cancer With Comorbid AF

Increasingly, AF has been recognized as an important comorbidity in cancer. Previous studies reported that patients with cancer have an increased risk of new-onset AF compared with the general population or individuals without cancer [7,8,18]. In addition, patients with cancer and AF showed increased risks of thromboembolism, bleeding, heart failure, and mortality, which highlighted the importance of detecting and managing AF in patients with cancer [19,20]. Notably, elevated breast cancer risk was reported in patients with pre-existing AF [3,9]; however, it remained unknown if the results of these studies could be extended to other cancer types among patients with AF. This study has contributed to filling in this knowledge gap, especially in the Chinese population. Significantly increased cancer prevalence was found in the population with AF, most prominently for hematological tumors (including multiple myeloma, leukemia, and non-Hodgkin lymphoma), intrathoracic malignancies (including mediastinum, lung, and esophageal cancers), and bone, soft tissue, and kidney cancers. It is important to note that the prominent association between AF and CNS cancer, which was previously reported [8], was not observed in our study. The reason for this disparity was not clear. It is possible that the previously reported new-onset AF after CNS cancer diagnosis was mainly attributed to cardiotoxicity following cancer treatment, rather than the common pathway between oncogenesis and arrhythmogenesis.

Challenges in Establishing Temporal Relationships Between AF and Cancer

In this study, the temporal relationship between the two diseases was not differentiated for 4 reasons. First, the diagnoses of both AF and cancer commonly present latency following the development of the disease [10,11]. Years of delay in AF diagnoses have been noticed in more than 30% of patients with prior cryptogenic stroke, even under frequent electrocardiographic screening [10]. Likewise, early diagnosis of cancer relies on screening at its asymptomatic stage, which is more challenging in LMICs due to weaker health care and economic infrastructure [21]. Therefore, it would be challenging to determine the temporal order of the index diagnoses of AF and cancer.

Second, both diseases at their preclinical stage may already predispose similar pathological effects prior to their clinical stage. Atrial high-rate episodes, the precursor to AF associated with subsequent clinically documented AF, could already present AF-prone thromboembolic, cardiovascular, and mortality risks [22]. Carcinoma in situ, the precursor to invasive cancer, may have already created a tumorous microenvironment and triggered a systemic immune response during its transition [23]. Therefore, the index diagnoses might be unable to determine the potential interconnection of the two entities at their subclinical stages.

Third, various common risk factors suggested that temporal sequencing did not necessarily imply causation. Common risk factors included aging, hypertension, diabetes, obesity, smoking, consumption of alcohol, air pollution, mental and physical stress, sleep disturbances, and westernized diet and lifestyle, among others. These risk factors led to pathological processes that precipitated both diseases, such as chronic inflammation, oxidative stress, electrolyte-handling abnormalities, and autonomous dysregulation [15]. Therefore, the AF-cancer comorbid state may comprehensively reflect the true nature of the development of the two diseases.

Finally, the temporal association may result in selection bias and lead-time bias [3,5,7,8]. The highest rate of new-onset AF in the first 60 or 90 days after cancer diagnosis may partially be attributed to frequent clinical visits, electrophysiological monitoring, and comprehensive cardiovascular evaluation that may affect cancer therapy [3,9]. On the other hand, the initiation of anticoagulation after AF diagnosis may reveal occult cancers by the alerting sign of bleeding, especially at gastrointestinal and genitourinary sites [24]. Nevertheless, a definite timeline between the onset of cancer and AF did exist in several circumstances; these conditions were commonly correlated with antineoplastic therapies, including chemotherapy, targeted therapy, surgery, and radiation, as well as cancer-related pain and stress [25]. However, the effects of arrhythmogenic cardiotoxicity differed among individuals. Those with proarrhythmic atrial substrate may be more vulnerable to AF during cancer therapy.

Heightened Male Predominance of Cancer Among Patients With AF

Robust sex differences exist in almost all aspects of cancer and AF, including epidemiology, progression, treatment response, and survival [26-28]. A striking male predominance in cancers was found in the general population from over 60 countries [26]. Similarly, a 3:2 male to female ratio in AF prevalence was also present, despite a female predominance in the older population [27]. Due to the higher prevalence of both cancer and AF in men, it seemed reasonable to expect higher rates of the comorbid conditions in men than in women. Our results in the population without AF demonstrated a male predominance of cancer at 15 of 18 sites, which was consistent with a previous study [26]. Of note, the male to female ratio in cancer was further intensified in the population with AF. The extents to which sex differences among cancer types were strengthened by AF were not uniform among the different cancer types and were most prominent in lung and colorectal cancers. The varied extents suggested that mechanisms other than simple pile-up effects were at play. The response of sex differentiated cardiotoxicity to cancer therapy may also contribute to a heightened sex disparity in AF development [28].

Importance of Cancer Screening in Patients With AF

As one of the most important topics in oncology, cancer screening aims to achieve early cancer detection, improve clinical outcomes, and reduce the burden of health economics. The benefits of screening for a particular cancer must outweigh the associated costs and risks of screening. The success of a screening program also depends on the disease burden, the reliability of the test, population participation, adequate resources, and staffing. Therefore, selection of the high-risk population and the cut-off age is of high importance in cancer screening, especially in LMICs compared to countries with ample resources [29]. The current cancer screening guidelines highlighted several high-risk conditions that benefited from early screening and intensive precursor surveillance for specific cancers, such as smoking for lung cancer, human papillomavirus infection for cervix cancer, family history of malignancies, and so on [30]. The results of our study suggest not only a higher prevalence but also a younger peak age of cancer diagnosis in the population with AF. Specifically, nonsolid tumors (including multiple myeloma, leukemia, and non-Hodgkin lymphoma) and intrathoracic malignancies (including lung, esophageal, and mediastinum cancers) were characterized by higher absolute prevalence, younger peak age, and stronger associations with AF than other cancer types. Therefore, early screening of nonsolid and intrathoracic cancers via methods such as complete blood count, endoscopy, and chest computed tomography may be beneficial for patients with AF. Accordingly, more intensive heart rhythm monitoring for potential AF should also be considered in these types of cancers to optimize the risk assessment and standardize anticoagulation management.

Limitations

First, only adult patients were included in this study. AF is rare in children, and pediatric malignancies may need to be seen as a distinct subset from adult cancers due to different cancer genetics, distributions, risk factors, comorbidities, and consequently, pathophysiological mechanisms. Second, the generalizability of our findings may be limited with regard to population and time, despite the large volume of data. The profiles of both cancer and AF were undergoing a transition in China due to the constantly changing demographic structure and health economic background and were affected by other diseases and different lifestyles [31]. Shanghai, as the most industrialized region in China with the oldest population [32], may have heavier burdens of both cancer and AF than other regions in China and other LMICs [16,33]. However, the epidemiological features, medical resources, and disease management provisions in Shanghai may share a mixture of characteristics with those in LMICs and high-income countries. This may represent a future trend in other parts of China and other LMICs around the world. Third, asymptomatic AF may have gone underdiagnosed. However, thanks to the popularity of contracted family doctor services, residents in Shanghai undergo routine cardiac auscultation and intermittent handheld electrocardiogram recording, which may have largely unveiled silent AF during the 6-year inclusion period. Lastly, endogeneity bias might exist since other diseases and lifestyles also participate in the AF-cancer association.

Conclusions

This large medical insurance database study demonstrates that the comorbid state of AF and cancer is substantial in Shanghai, China. Patients with AF face increased prevalence, heightened male predominance, and a younger peak age of cancer. Physicians and patients should be aware of the cancer risks when AF is diagnosed. Screening for AF-associated cancers, such as by complete blood count for hematological tumors or thoracic computer tomography for intrathoracic malignancies, might be considered in patients with AF after a consultation with a cardio-oncologist. Likewise, more intensive heart rhythm monitoring for potential AF might be performed in specific cancers to optimize the prognosis, assessment, and prevention of cardiovascular complications. Further studies are needed to determine whether early screening of specific cancers in patients with AF is cost-effective and beneficial. Investigations regarding the value of wearable devices and mobile health in the screening and early detection of AF-cancer comorbidity are also warranted.

Acknowledgments

This work was supported by the Shanghai Hospital Development Center (grant SHDC2020CR2026B), the Clinical Research Plan of Shanghai Municipal Health Commission (grant 20224Y0078), the National Natural Science Foundation of China (grants 81900288, 82130009, and 82070515), and the Shanghai Municipal Science and Technology Commission (grant 19YF1431900).

Data Availability

The data that support the findings of this study were obtained from the Shanghai Municipal Health Commission. Data sets are available from the corresponding author with the permission of the Shanghai Municipal Health Commission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 240 KB - publichealth_v9i1e40149_app1.docx]

References

1. Chen M, Li C, Liao P, Cui X, Tian W, Wang Q, et al. Epidemiology, management, and outcomes of atrial fibrillation among 30 million citizens in Shanghai, China from 2015 to 2020: a medical insurance database study. Lancet Reg Health West Pac 2022 Jun;23:100470 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100470] [Medline: 35542895]
2. Rahman F, Kwan GF, Benjamin EJ. Global epidemiology of atrial fibrillation. Nat Rev Cardiol 2014 Nov;11(11):639-654. [doi: 10.1038/nrcardio.2014.118] [Medline: 25113750]
3. Conen D, Wong JA, Sandhu RK, Cook NR, Lee I, Buring JE, et al. Risk of malignant cancer among women with new-onset atrial fibrillation. JAMA Cardiol 2016 Jul 01;1(4):389-396 [FREE Full text] [doi: 10.1001/jamacardio.2016.0280] [Medline: 27438314]
4. Pastori D, Marang A, Bisson A, Menichelli D, Herbert J, Lip GYH, et al. Thromboembolism, mortality, and bleeding in 2,435,541 atrial fibrillation patients with and without cancer: a nationwide cohort study. Cancer 2021 Jun 15;127(12):2122-2129 [FREE Full text] [doi: 10.1002/cncr.33470] [Medline: 33631041]
5. Marijon E, Le Heuzey J, Connolly S, Yang S, Pogue J, Brueckmann M, et al. Causes of death and influencing factors in patients with atrial fibrillation. Circulation 2013 Nov 12;128(20):2192-2201. [doi: 10.1161/circulationaha.112.000491]
6. Bassand J, Accetta G, Camm AJ, Cools F, Fitzmaurice DA, Fox KAA, GARFIELD-AF Investigators. Two-year outcomes of patients with newly diagnosed atrial fibrillation: results from GARFIELD-AF. Eur Heart J 2016 Oct 07;37(38):2882-2889 [FREE Full text] [doi: 10.1093/eurheartj/ehw233] [Medline: 27357359]
7. Jakobsen CB, Lamberts M, Carlson N, Lock-Hansen M, Torp-Pedersen C, Gislason GH, et al. Incidence of atrial fibrillation in different major cancer subtypes: a nationwide population-based 12 year follow up study. BMC Cancer 2019 Nov 14;19(1):1105 [FREE Full text] [doi: 10.1186/s12885-019-6314-9] [Medline: 31726997]
8. Yun JP, Choi E, Han K, Park SH, Jung J, Park SH, et al. Risk of atrial fibrillation according to cancer type: a nationwide population-based study. JACC CardioOncol 2021 Jun;3(2):221-232 [FREE Full text] [doi: 10.1016/j.jaccao.2021.03.006] [Medline: 34396327]
9. Guha A, Fradley MG, Dent SF, Weintraub NL, Lustberg MB, Alonso A, et al. Incidence, risk factors, and mortality of atrial fibrillation in breast cancer: a SEER-Medicare analysis. Eur Heart J 2022 Jan 31;43(4):300-312 [FREE Full text] [doi: 10.1093/eurheartj/ehab745] [Medline: 34791123]

10. Brachmann J, Morillo CA, Sanna T, Di Lazzaro V, Diener H, Bernstein RA, et al. Uncovering atrial fibrillation beyond short-term monitoring in cryptogenic stroke patients: three-year results from the cryptogenic stroke and underlying atrial fibrillation trial. Circ Arrhythm Electrophysiol 2016 Jan;9(1):e003333. [doi: 10.1161/CIRCEP.115.003333] [Medline: 26763225]
11. Brand NR, Qu LG, Chao A, Ilbawi AM. Delays and barriers to cancer care in low- and middle-income countries: a systematic review. Oncologist 2019 Dec;24(12):e1371-e1380 [FREE Full text] [doi: 10.1634/theoncologist.2019-0057] [Medline: 31387949]
12. Schliemann D, Tan MM, Hoe WMK, Mohan D, Taib NA, Donnelly M, et al. mHealth interventions to improve cancer screening and early detection: scoping review of reviews. J Med Internet Res 2022 Aug 15;24(8):e36316 [FREE Full text] [doi: 10.2196/36316] [Medline: 35969450]
13. Kwon S, Lee S, Choi E, Ahn H, Song H, Lee Y, et al. Comparison between the 24-hour Holter test and 72-hour single-lead electrocardiogram monitoring with an adhesive patch-type device for atrial fibrillation detection: prospective cohort study. J Med Internet Res 2022 May 09;24(5):e37970 [FREE Full text] [doi: 10.2196/37970] [Medline: 35532989]
14. Nazarian S, Lam K, Darzi A, Ashrafian H. Diagnostic accuracy of smartwatches for the detection of cardiac arrhythmia: systematic review and meta-analysis. J Med Internet Res 2021 Aug 27;23(8):e28974 [FREE Full text] [doi: 10.2196/28974] [Medline: 34448706]
15. Chu G, Versteeg HH, Verschoor AJ, Trines SA, Hemels MEW, Ay C, et al. Atrial fibrillation and cancer - an unexplored field in cardiovascular oncology. Blood Rev 2019 May;35:59-67. [doi: 10.1016/j.blre.2019.03.005] [Medline: 30928168]
16. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021 Jan;71(1):7-33 [FREE Full text] [doi: 10.3322/caac.21654] [Medline: 33433946]
17. Wei W, Zeng H, Zheng R, Zhang S, An L, Chen R, et al. Cancer registration in China and its role in cancer prevention and control. Lancet Oncol 2020 Jul;21(7):e342-e349. [doi: 10.1016/S1470-2045(20)30073-5] [Medline: 32615118]
18. O'Neal WT, Lakoski SG, Qureshi W, Judd SE, Howard G, Howard VJ, et al. Relation between cancer and atrial fibrillation (from the REasons for Geographic And Racial Differences in Stroke Study). Am J Cardiol 2015 Apr 15;115(8):1090-1094. [doi: 10.1016/j.amjcard.2015.01.540] [Medline: 25711434]
19. Hu Y, Liu C, Chang PM, Tsao H, Lin Y, Chang S, et al. Incident thromboembolism and heart failure associated with new-onset atrial fibrillation in cancer patients. Int J Cardiol 2013 May 10;165(2):355-357. [doi: 10.1016/j.ijcard.2012.08.036] [Medline: 22989607]
20. Hussain M, Misbah R, Donnellan E, Alkharabsheh S, Hou Y, Cheng F, et al. Impact of timing of atrial fibrillation, CHADS-VASc score and cancer therapeutics on mortality in oncology patients. Open Heart 2020 Nov;7(2):e001412 [FREE Full text] [doi: 10.1136/openhrt-2020-001412] [Medline: 33243931]
21. Shah SC, Kayamba V, Peek RM, Heimburger D. Cancer control in low- and middle-income countries: is it time to consider screening? J Glob Oncol 2019 Dec(5):1-8 [FREE Full text] [doi: 10.1200/jgo.18.00200]
22. Khan AA, Boriani G, Lip GYH. Are atrial high rate episodes (AHREs) a precursor to atrial fibrillation? Clin Res Cardiol 2020 Apr;109(4):409-416 [FREE Full text] [doi: 10.1007/s00392-019-01545-4] [Medline: 31522249]
23. Risom T, Glass DR, Averbukh I, Liu CC, Baranski A, Kagel A, et al. Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. Cell 2022 Jan 20;185(2):299-310.e18 [FREE Full text] [doi: 10.1016/j.cell.2021.12.023] [Medline: 35063072]
24. Hu Y, Chang S, Chern C, Wu C, Lin Y, Lo L, et al. Bleeding and new-onset cancers in patients with atrial fibrillation receiving nonvitamin K antagonist oral anticoagulants. Am J Cardiol 2019 Mar 01;123(5):782-786. [doi: 10.1016/j.amjcard.2018.11.046] [Medline: 30551841]
25. Herrmann J. Adverse cardiac effects of cancer therapies: cardiotoxicity and arrhythmia. Nat Rev Cardiol 2020 Aug;17(8):474-502 [FREE Full text] [doi: 10.1038/s41569-020-0348-1] [Medline: 32231332]
26. Edgren G, Liang L, Adami H, Chang ET. Enigmatic sex disparities in cancer incidence. Eur J Epidemiol 2012 Mar;27(3):187-196. [doi: 10.1007/s10654-011-9647-5] [Medline: 22212865]
27. Volgman AS, Bairey Merz CN, Benjamin EJ, Curtis AB, Fang MC, Lindley KJ, et al. Sex and race/ethnicity differences in atrial fibrillation. J Am Coll Cardiol 2019 Dec 03;74(22):2812-2815 [FREE Full text] [doi: 10.1016/j.jacc.2019.09.045] [Medline: 31779796]
28. Wilcox NS, Rotz SJ, Mullen M, Song EJ, Ky Hamilton B, Moslehi J, et al. Sex-specific cardiovascular risks of cancer and its therapies. Circ Res 2022 Feb 18;130(4):632-651 [FREE Full text] [doi: 10.1161/CIRCRESAHA.121.319901] [Medline: 35175846]
29. Sankaranarayanan R. Screening for cancer in low- and middle-income countries. Ann Glob Health 2014;80(5):412-417 [FREE Full text] [doi: 10.1016/j.aogh.2014.09.014] [Medline: 25512156]
30. Smith RA, Andrews KS, Brooks D, Fedewa SA, Manassaram-Baptiste D, Saslow D, et al. Cancer screening in the United States, 2019: a review of current American Cancer Society guidelines and current issues in cancer screening. CA Cancer J Clin 2019 May;69(3):184-210 [FREE Full text] [doi: 10.3322/caac.21557] [Medline: 30875085]
31. Cao W, Chen H, Yu Y, Li N, Chen W. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. Chin Med J (Engl) 2021 Mar 17;134(7):783-791 [FREE Full text] [doi: 10.1097/CM9.0000000000001474] [Medline: 33734139]

32. 2020 Shanghai statistical yearbook. Shanghai Municipal People's Government. URL: https://tjj.sh.gov.cn/tjnj/zgsh/tjnj2020en.html [accessed 2023-09-28]
33. Dein E, Sharfman W, Kim J, Gellad F, Shah AA, Bingham CO, et al. Two cases of sinusitis induced by immune checkpoint inhibition. J Immunother 2017 Oct;40(8):312-314 [FREE Full text] [doi: 10.1097/CJI.0000000000000174] [Medline: 28614096]

Abbreviations

AF: atrial fibrillation
CNS: central nervous system
ICD-10: International Statistical Classification of Diseases, Tenth Revision
LMIC: low- and middle-income countries
PR: prevalence ratio

Edited by A Mavragani, T Sanchez; submitted 08.06.22; peer-reviewed by D Xu, M Ghaderzadeh, X Lu, K Uludag; comments to author 09.11.22; revised version received 13.11.22; accepted 19.09.23; published 17.10.23.

Please cite as:
Chen M, Li C, Zhang J, Cui X, Tian W, Liao P, Wang Q, Sun J, Luo L, Wu H, Li YG
Cancer and Atrial Fibrillation Comorbidities Among 25 Million Citizens in Shanghai, China: Medical Insurance Database Study
JMIR Public Health Surveill 2023;9:e40149
URL: https://publichealth.jmir.org/2023/1/e40149
doi:10.2196/40149
PMID:37847541

Original Paper

Residential Segregation and County-Level COVID-19 Booster Coverage in the Deep South: Surveillance Report and Ecological Study

Chengbo Zeng[1,2,3], PhD; Jiajia Zhang[2,3,4], PhD; Zhenlong Li[2,3,5], PhD; Xiaowen Sun[2,3,4], MSc; Huan Ning[3,5], MSc; Xueying Yang[1,2,3], PhD; Sharon Weissman[2,6], MD; Bankole Olatosi[2,3,7], PhD; Xiaoming Li[1,2,3], PhD

[1]Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[2]South Carolina SmartState Center for Healthcare Quality, University of South Carolina, Columbia, SC, United States

[3]Big Data Health Science Center, University of South Carolina, Columbia, SC, United States

[4]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[5]Geoinformation and Big Data Research Laboratory, Department of Geography, College of Arts and Sciences, University of South Carolina, Columbia, SC, United States

[6]Department of Internal Medicine, School of Medicine, University of South Carolina, Columbia, SC, United States

[7]Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

Corresponding Author:
Chengbo Zeng, PhD
Department of Health Promotion, Education, and Behavior
Arnold School of Public Health
University of South Carolina
4th Floor
915 Greene Street
Columbia, SC, 29208
United States
Email: czeng@email.sc.edu

Abstract

Background: COVID-19 had a greater impact in the Deep South compared with other regions in the United States. While vaccination remains a top priority for all eligible individuals, data regarding the progress of booster coverage in the Deep South and how the coverage varies by county and age are sparse. Despite existing evidence of racial and ethnic disparities in COVID-19 vaccinations at the individual level, there is an urgent need for evidence at the population level. Such information could highlight vulnerable communities and guide future health care policy-making and resource allocation.

Objective: We aimed to evaluate county-level COVID-19 booster coverage by age group in the Deep South and explore its association with residential segregation.

Methods: An ecological study was conducted at the population level by integrating COVID-19 vaccine surveillance data, residential segregation index, and county-level factors across the 418 counties of 5 Deep South states from December 15, 2021, to October 19, 2022. We analyzed the cumulative percentages of county-level COVID-19 booster uptake by age group (eg, 12 to 17 years, 18 to 64 years, and at least 65 years) by the end of the study period. The longitudinal relationships were examined between residential segregation, the interaction of time and residential segregation, and COVID-19 booster coverage using the Poisson model.

Results: As of October 19, 2022, among the 418 counties, the median of booster uptake was 40% (IQR 37.8%-43%). Compared with older adults (ie, at least 65 years; median 63.1%, IQR 59.5%-66.5%), youth (ie, 12 to 17 years; median 14.1%, IQR 11.3%-17.4%) and adults (ie, 18 to 64 years; median 33.4%, IQR 30.5%-36.5%) had lower percentages of booster uptake. There was geospatial heterogeneity in the county-level COVID-19 booster coverage. We found that higher segregated counties had lower percentages of booster coverage. Such relationships attenuated as time increased. The findings were consistent across the age groups.

Conclusions: The progress of county-level COVID-19 booster coverage in the Deep South was slow and varied by age group. Residential segregation precluded the county-level COVID-19 booster coverage across age groups. Future efforts regarding vaccination strategies should focus on youth and adults. Health care facilities and resources are needed in racial and ethnic minority communities.

(JMIR Public Health Surveill 2023;9:e44257) doi:10.2196/44257

KEYWORDS

Deep South; COVID-19; vaccine; booster; residential segregation

Introduction

Following its approval in September 2021, the booster dose has contributed to more effective prevention and lower mortality for COVID-19 [1,2]. In the United States, the first COVID-19 booster was initially recommended to be administered at least 6 months after the completion of the primary series for individuals aged 18 years and older [2-4]. This guidance later expanded to individuals aged between 16 and 17 years [2-4]. By January 2022, the US Food and Drug Administration [5,6] extended the booster recommendation to those aged 12 to 15 years and shortened the interval to 5 months after the primary vaccination. Despite the encouraging evidence on the efficacy of the COVID-19 vaccine, as of October 19, 2022, only 50.7% of Americans aged 12 years and older had both completed the primary series and received a booster [7,8]. In the Deep South, which includes Alabama, Georgia, Louisiana, Mississippi, and South Carolina, this percentage was even lower (40%-44%) [7,8]. Although around 75% of the individuals aged 12 years and older in the United States (and approximately 63% in the Deep South) are fully vaccinated, the relatively low percentage of booster uptake is concerning as a booster dose can effectively decrease the risk of infection and severe illness [7,8]. Given that the Deep South experienced a disproportionately more severe impact from COVID-19 compared with other regions, vaccination remains a top priority for all eligible individuals [9-11]. However, there are limited data delineating booster coverage in the Deep South and how this rate varied by county and age, which is of critical importance for future vaccine planning, especially for vaccination of updated booster dose.

Racial and ethnic minority communities, including Black and Hispanic communities, are vulnerable to COVID-19 but lack a variety of health care resources. In the prevaccination period of the pandemic, these communities displayed higher incidences of infections, hospitalizations, and unfavorable treatment outcomes [9,12-14]. In the postvaccination period, both Black and Hispanic communities had a lower percentage of completing at least 1 COVID-19 vaccine dose when compared to their White counterparts [15]. This racial and ethnic disparity is also evident among children and adolescents. While the child and adolescent COVID-19 vaccination rates were low overall, Black and Hispanic children and adolescents (aged 5 to 17 years) had lower COVID-19 vaccination rates than their White and Asian peers [16]. One of the major barriers is residential segregation that restricts access to health care resources in minority communities [17]. Empirical research demonstrated that more segregated counties had more pronounced differences in COVID-19 vaccine coverage between Black and White residents [18,19]. However, the vast majority of evidence regarding racial and ethnic disparities in COVID-19 vaccination was generated from the studies at the individual level with limited population-level analyses. Of the few population-based studies, most were cross-sectional designs and did not examine the difference in COVID-19 vaccination by age groups [18,19]. Knowledge gleaned from these age-specific disparities can illuminate the progress of vaccination in racial and ethnic minority communities and facilitate the process of health care policy-making.

Well-designed ecological studies hold the potential to yield compelling evidence on racial and ethnic disparities in vaccination at the population level. Such insights can facilitate the identification of vulnerable communities and guide the optimal allocation of resources for disease control and prevention. A notable strength of ecological studies is their ability to harness comprehensive surveillance data [20,21]. Compared with individual patient data that often require rigorous ethical approvals, surveillance data are publicly accessible. Moreover, by examining racial and ethnic disparities in vaccination at aggregate levels such as communities or counties, ecological studies generate findings that are more generalizable [21]. This is a remarkable strength compared with results generated from analyses that focused on limited individuals and small geographic areas.

In this study, we evaluated the COVID-19 booster coverage by age group among the 418 counties from the 5 Deep South states and examined its relationship with racial and ethnic residential segregation using an ecological design, vaccine surveillance data, and spatiotemporal analysis.

Methods

Design, Setting, and Study Period

We conducted an ecological study at the population level by integrating COVID-19 vaccine surveillance data, residential segregation index, and county-level factors from multiple public data sets across the 418 counties of 5 Deep South states (ie, Alabama, Georgia, Louisiana, Mississippi, and South Carolina) from December 15, 2021, to October 19, 2022.

County-Level COVID-19 Booster Coverage

We retrieved variables regarding vaccine uptake from the Centers for Disease Control and Prevention (CDC) [22] COVID-19 vaccine surveillance data. This data set is representative as it reflects all vaccine partners including jurisdictional partner clinics, retail pharmacies, long-term care facilities, dialysis centers, Federal Emergency Management Agency and Health Resources and Services Administration partner sites, and federal entity facilities. The data regarding the percentage of adults who completed a primary series and have received a booster within a county (henceforth, “county-level booster coverage”) were released on December 15, 2021. The CDC also reported county-level booster coverage for people aged at least 12 years since January 27, 2022. The CDC vaccine surveillance data were updated daily until June 29, 2022, and were then updated weekly.

We retrieved the biweekly cumulative county-level booster uptake for people aged at least 18 years and those aged at least 65 years between December 15, 2021, and October 19, 2022. A total of 23 time points were included in the analyses. Since there was no existing variable for people aged between 18 and 64 years, we manually calculated the biweekly cumulative county-level booster uptake for this group by subtracting the booster uptake of people aged at least 65 years from the overall adults for each county.

For people aged between 12 and 17 years, we retrieved the biweekly cumulative booster uptake for people aged at least 12 years and at least 18 years from January 27 to October 19, 2022. A total of 20 time points were included in the analyses. We used the same method to calculate the biweekly booster uptake for this group in each county.

Residential Segregation

We defined residential segregation using the index of concentration that measures the relative amount of physical space occupied by minority groups [23,24]. Minority groups of the same relative size occupying less space were considered more concentrated and thus more segregated [23,24]. Since there is a large proportion of the Black population in the Deep South, we considered it as the main minority group and calculated the residential segregation for each county using equation (1):

where n is the number of tracts in a county, x_i is the size of the Black population in tract i, X is the size of the Black population in a county, a_i is the land area of tract i, and A is the total land area in a county [24]. The residential segregation ranged from 0 to 1, with a higher value indicating a higher degree of segregation. In each county, the higher the residential segregation score, the fewer spaces the Black population of the same relative size occupy.

Potential Confounders

Given the ecological design at the county level, we identified a list of potential county-level confounders based on prevailing literature addressing the structural and social determinants of racial and ethnic disparities of the COVID-19 pandemic in the United States [25]. Thakur et al [25] underscored that racism, social class, and social stratification shaped the risk of exposure to COVID-19 and the access to health care resources through (1) income and occupation, (2) housing and crowding, and (3) health insurance and resource distribution. Therefore, we organized the potential county-level confounders into four dimensions: (1) regional socioeconomic status (ie, the Gini index, the proportion of households with public assistance income [%], the proportion of people in low working class [%], the proportion of people with low education [%], and the proportion of noncitizen [%]); (2) housing and neighborhood environment (ie, household size); (3) health care access and susceptibility (ie, primary care provider rate [per 100,000 people] and proportion of adults who report fair or poor health [%]); and (4) transportation accessibility (ie, proportion of occupied housing units without car access [%]). These variables were validated in prior research to reflect the structural barriers to health care access and delivery [26,27]. We retrieved these variables from multiple public data sets and linked to the county level. Multimedia Appendix 1 shows the definitions and sources of all potential county-level confounders by each dimension.

Statistical Analysis

Using the geospatial mapping technique, we mapped the cumulative percentages of county-level COVID-19 booster uptake for all ages, people between 12 and 17 years, at least 18 years, between 18 and 64 years, and at least 65 years, respectively. The county-level residential segregation was also mapped. We described the median and IQR for the cumulative percentages of COVID-19 booster coverage at the county level for each of the 5 Deep South states on October 19, 2022.

We used generalized estimating equation with Poisson distribution to examine the relationship between residential segregation and COVID-19 booster coverage rate, adjusting for the repeated measures in each county and potential confounders. Since the US Food and Drug Administration [5,6] advises the first COVID-19 booster dose be administered 5 months after the completion of the primary vaccination series, to generate a robust estimate of the county-level booster coverage rate, we used the 7-day moving average of the total individuals who completed the primary series 5 months before each study time point as an offset in the Poisson model. The model can be presented using equation (2):

where $\mu_{ijt} = E(y_{ijt}|X_{ij}, V_{ijt}, t)$ is the marginal mean at time t given the covariates, and y_{ijt} denotes the total number of patients who took a booster at jth county within ith state during time t. The response (county-level COVID-19 booster coverage rate at time t) is assumed to be independent across the state but correlated within each county over time. β is a vector of regression coefficients, X_{ij} denotes the vector of county-level variables including residential segregation and other county-level covariates, γ denotes the coefficient of time, and V_{ijt} is the 7-day moving average of the total number of people who completed the primary series of COVID-19 vaccine at county j within state i at time t.

To avoid collinearity, county-level factors were standardized into the same scale with a mean of 0 and an SD of 1 before the analysis. First, we tested the main effects of time and residential segregation on the cumulative percentage of COVID-19 booster coverage. Second, given the temporal trend of COVID-19 vaccination in empirical research, we also examined the interaction between time and residential segregation and sought to understand whether the impact of residential segregation on booster coverage rate changes over time [28]. We used simple slope analysis to interpret interaction [29]. The analysis was replicated for the overall population and by age group. Finally, besides using 5 months as the interval between the primary vaccination series and booster dose as suggested by the US Food and Drug Administration, we did the sensitivity analyses, in which 6-month was used as a cutoff to calculate the offset for the Poisson model. All analyses were conducted using SAS (version 9.4; SAS Institute, Inc).

Ethical Considerations

The institutional review boards at the University of South Carolina approved the study protocol (PRO00100854). This study was an ecological analysis based on the CDC vaccine surveillance data and multiple public access data sets. No personal identification information was involved in this analysis.

Results

Overview

A total of 418 counties across 5 Deep South states were included in this study. There were 67, 159, 64, 82, and 46 counties in Alabama, Georgia, Louisiana, Mississippi, and South Carolina, respectively.

As of October 19, 2022, among the 418 counties, the median of booster uptake was 40% (IQR 37.8%-43%). In the individual states, the median of county-level booster uptake ranged from 38.4% (IQR 36.2%-40.1%) in Alabama to 43.4% (IQR 40.5%-45.4%) in South Carolina (Table 1).

Table 1. Cumulative percentages of county-level COVID-19 booster uptake for the overall sample and by age group across the 5 Deep South states, as of October 19, 2022.

States and age groups	Values (%), median (IQR)
Deep South (N=418)	
Overall population	40 (37.8-43)
12 to 17 years[a]	14.1 (11.3-17.4)
Overall adults	
≥18 years	42.2 (39.8-45.1)
18 to 64 years	33.4 (30.5-36.5)
At least 65 years	63.1 (59.5-66.5)
Alabama (n=67)	
Overall population	38.4 (36.2-40.1)
12 to 17 years[a]	13 (10.5-15.6)
Overall adults	
≥18 years	40 (37.8-41.8)
18 to 64 years	30.2 (28.1-33.8)
At least 65 years	59.5 (57.5-61.7)
Georgia (n=159)	
Overall population	40.2 (37.5-43.2)
12 to 17 years[a]	14.9 (11.8-17.8)
Overall adults	
≥18 years	42 (39.6-45.4)
18 to 64 years	33.8 (31.2-36.6)
At least 65 years	61.6 (58.3-64.8)
Louisiana (n=64)	
Overall population	40.7 (38.5-43.6)
12 to 17 years[a]	13.5 (11.2-16.8)
Overall adults	
≥18 years	42.7 (40.5-45.9)
18 to 64 years	33.3 (30.3-36.1)
At least 65 years	67.4 (63.4-71.8)
Mississippi (n=82)	
Overall population	39.9 (37.7-41.1)
12 to 17 years[a]	12.6 (10.6-16)
Overall adults	
≥18 years	42.2 (39.9-44.1)
18 to 64 years	32.5 (30.4-35.7)
At least 65 years	64 (61.1-67.3)
South Carolina (n=46)	
Overall population	43.4 (40.5-45.4)
12 to 17 years[a]	16.9 (14-20.3)
Overall adults	
≥18 years	46 (43.2-47.9)

States and age groups	Values (%), median (IQR)
18 to 64 years	36.2 (33.9-37.6)
At least 65 years	66.4 (64.4-69.4)

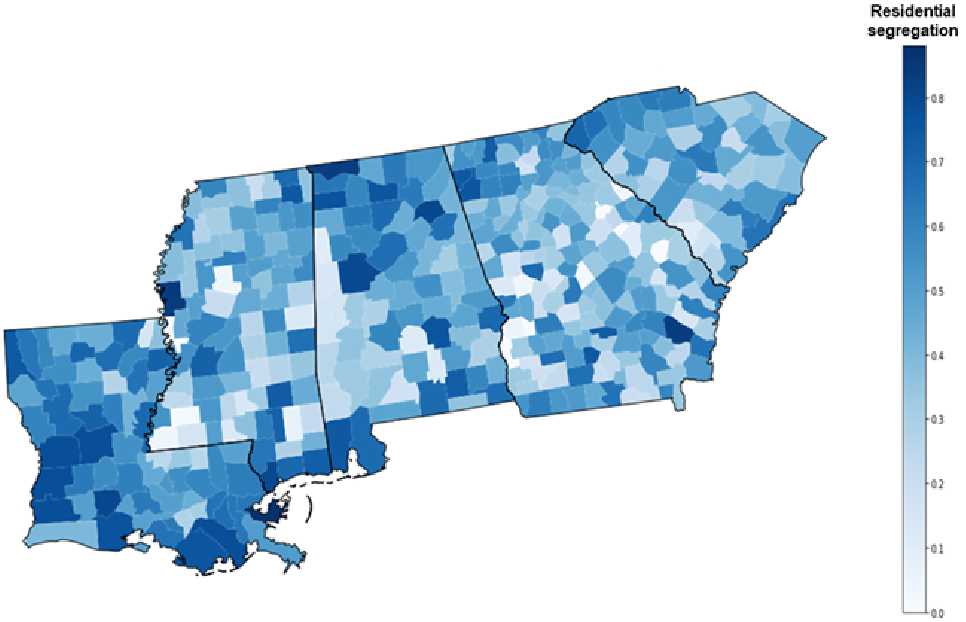
[a] The first record for the group of 12 to 17 years was available on January 27, 2022. For other groups, the first record was available on December 15, 2021.

The percentage of booster uptake for people aged at least 65 years was high with a median value of 63.1% (IQR 59.5%-66.5%) across the 418 counties. The percentage of booster uptake for people aged between 12 and 17 years was low with a median value of 14.1% (IQR 11.3%-17.4%). For people aged between 18 and 64 years, the percentage of booster coverage was 33.4% (IQR 30.5%-36.5%). Table 1 depicts the cumulative county-level percentage of COVID-19 booster uptake by age group across the 5 Deep South states, as of October 19, 2022.

Geospatial Heterogeneities in Residential Segregation and County-Level Booster Coverage

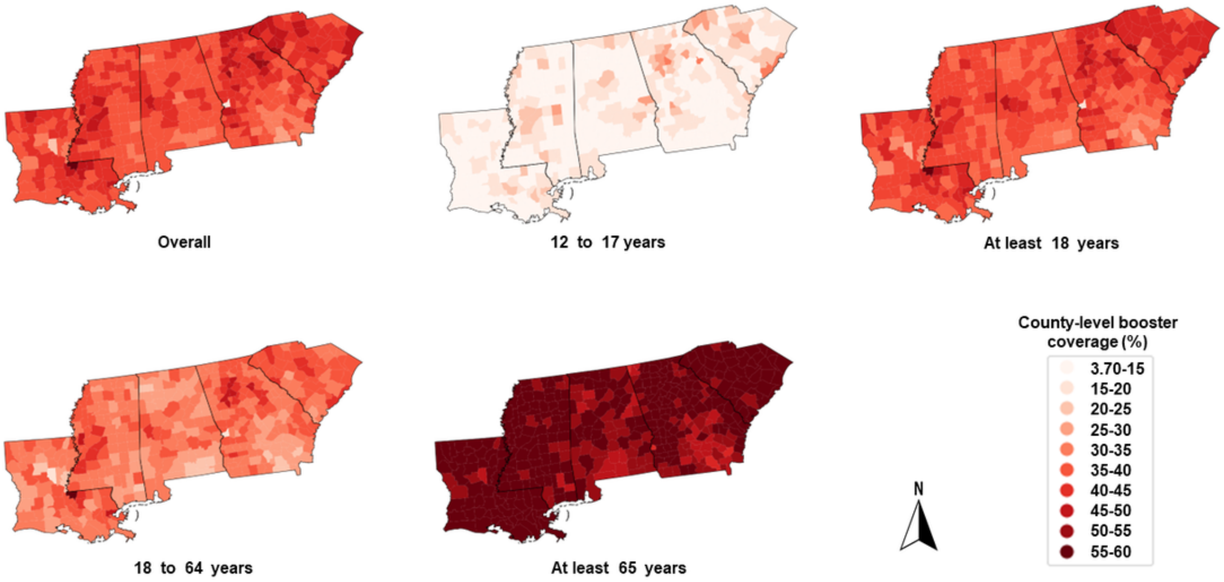
Figures 1 and 2 show the distribution of county-level residential segregation score and booster uptake, respectively. In Figure 1, counties in dark blue had higher residential segregation scores than those in light blue. There are some geospatial clusters with high levels of residential segregation within each state. For instance, Louisiana had more counties in dark blue than the other 4 states, indicating that counties in Louisiana were more segregated. These counties had high levels of residential segregation and were mainly located in the southwestern areas. In South Carolina, counties located in the northwestern and southeastern regions are more segregated compared with others.

Figure 1. County-level residential segregation in the 5 Deep South states. From left to right, the states were Louisiana, Mississippi, Alabama, Georgia, and South Carolina, respectively.



In Figure 2, counties in dark red had higher percentages of booster uptake than those in light red. In general, the percentage of completing primary series with 1 booster among people aged at least 65 years was higher than that in other age groups. In each individual state, there were some counties with low booster uptake. These counties were mainly located in southern Alabama, southeastern Georgia, southwestern and central Louisiana, southwestern and central Mississippi, and northern South Carolina. Figure 2 also shows the county-level booster uptake for other age groups.

Figure 2. County-level COVID-19 booster coverage for the overall sample and by age group in the 5 Deep South states on October 19, 2022. From left to right, the states were Louisiana, Mississippi, Alabama, Georgia, and South Carolina, respectively.



Taken Figures 1 and 2 together, there were some counties in which there was high residential segregation but low booster uptake. Some of these counties were Franklin and Coffee in Alabama, Bacon and Colquitt in Georgia, Vernon and Acadia in Louisiana, Hancock and Lauderdale in Mississippi, and Cherokee and Dorchester in South Carolina.

Relationship Between Residential Segregation and County-Level Booster Coverage

In general, as time increased, the change rate of county-level booster coverage decreased for people aged at least 18 years (β=–.051. 95% CI –0.072 to –0.031), between 18 and 64 years (β=–.053, 95% CI –0.079 to –0.027), and at least 65 years (β=–.028, 95% CI –0.044 to –0.012). For people aged between 12 and 17 years, with time increased, the change rate increased (β=.055, 95% CI 0.037 to 0.074).

Residential segregation was negatively associated with the county-level booster coverage in people aged at least 18 years (β=–.067, 95% CI –0.088 to –0.047), between 18 and 64 years (β=–.074, 95% CI –0.099 to –0.049), and at least 65 years (β=–.044, 95% CI –0.060 to –0.028). However, this negative association was not found in the analysis in people aged between 12 and 17 years (β=–.009, 95% CI –0.039 to 0.020).

In the analyses that examined the interaction between time and residential segregation on the county-level booster coverage, we consistently found a significant interaction for the overall sample (β=.054, 95% CI 0.040 to 0.069) and by age group (between 12 and 17 years: β=.031, 95% CI 0.018 to 0.044; at least 18 years: β=.057, 95% CI 0.042 to 0.072; between 18 and 64 years: β=.065, 95% CI 0.049 to 0.082; and at least 65 years: β=.047, 95% CI 0.032 to 0.062; Table 2). Simple slope analyses indicated that at a given time point, counties characterized by residential segregation above the mean (high residential segregation) experienced lower COVID-19 county-level booster coverage compared to counties where residential segregation was at or below the mean (low residential segregation). However, such difference attenuated as time increased (Figure 3).

Table 2. Residential segregation and COVID-19 booster coverage for the overall sample and by age group in the 418 counties across the 5 Deep South states from December 15, 2021, to October 19, 2022[a].

Model and variable	Overall[b], β (95% CI)	12-17 years[c], β (95% CI)	≥18 years, β (95% CI)	18-64 years, β (95% CI)	≥65 years, β (95% CI)
Model 1: main effects[d]					
Time	–.058 (–0.078 to –0.038)[e]	.055 (0.037 to 0.074)[e]	–.051 (–0.072 to –0.031)[e]	–.053 (–0.079 to –0.027)[e]	–.028 (–0.044 to –0.012)[e]
Residential segregation	–.066 (–0.086 to –0.045)[e]	–.009 (–0.039 to 0.020)	–.067 (–0.088 to –0.047)[e]	–.074 (–0.099 to –0.049)[e]	–.044 (–0.060 to –0.028)[e]
Model 2: main effects and interaction[d]					
Time	–.082 (–0.100 to –0.064)[e]	.044 (0.026 to 0.061)[e]	–.077 (–0.095 to –0.058)[e]	–.080 (–0.102 to –0.058)[e]	–.051 (–0.068 to –0.034)[e]
Residential segregation	–.073 (–0.094 to –0.052)[e]	–.016 (–0.046 to 0.014)	–.075 (–0.096 to –0.053)[e]	–.083 (–0.109 to –0.058)[e]	–.049 (–0.065 to –0.032)[e]
Time×residential segregation	.054 (0.040 to 0.069)[e]	.031 (0.018 to 0.044)[e]	.057 (0.042 to 0.072)[e]	.065 (0.049 to 0.082)[e]	.047 (0.032 to 0.062)[e]

[a] Unless otherwise noted, the study period was from December 15, 2021, to October 19, 2022.

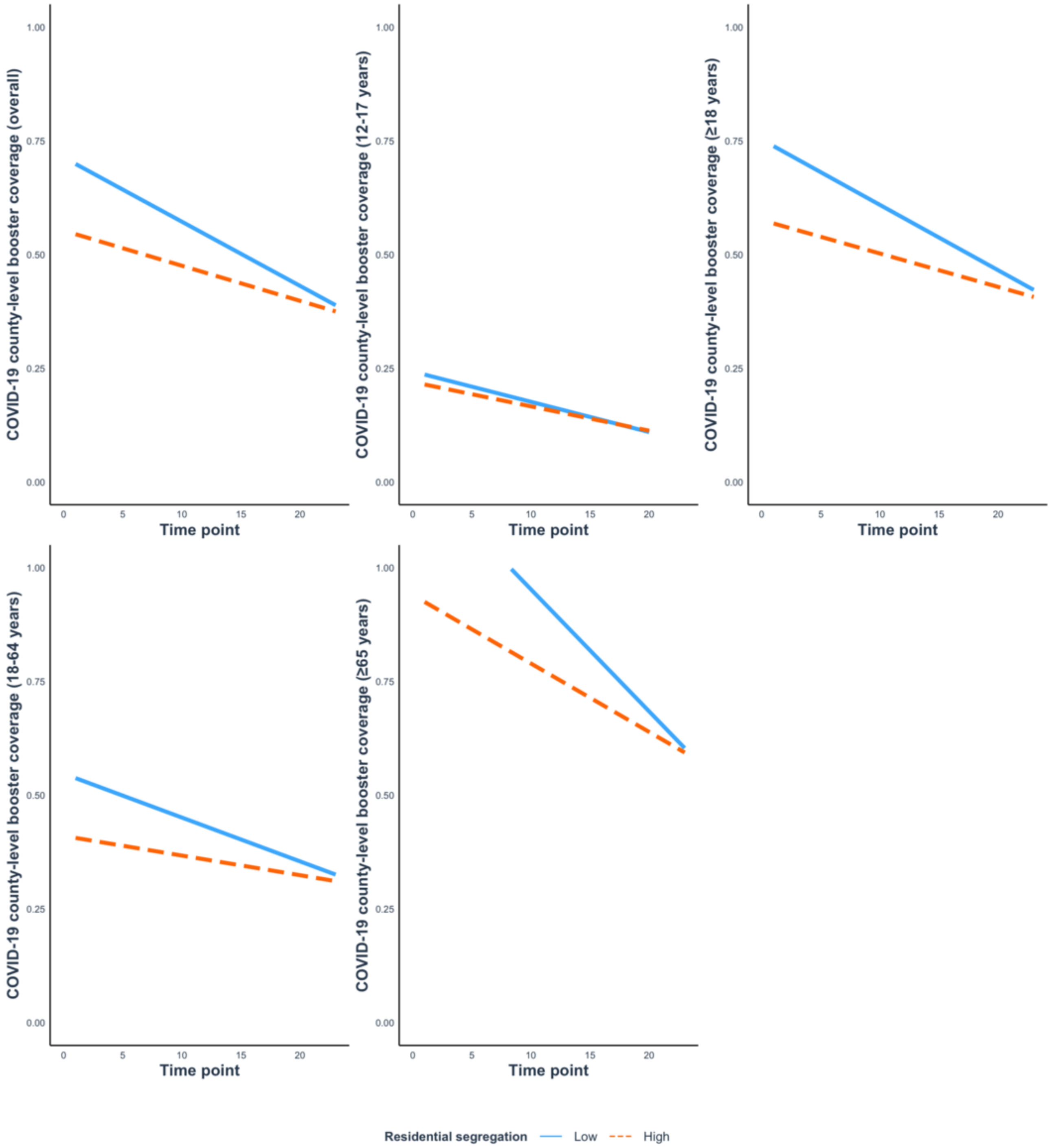
[b] From December 15, 2021 to January 26, 2022, the overall population referred to people aged at least 18 years. Since January 27, 2022, it referred to people aged at least 12 years.

[c] The first record for the group of 12 to 17 years was available on January 27, 2022. For other groups, the first record was available on December 15, 2021.

[d] Confounders: Gini index, proportion of households with public assistance income, proportion of people in low working class, proportion of people with low education, proportion of noncitizen, household size, primary care provider rate, proportion of adults who report fair or poor health, proportion of occupied housing units without access to a vehicle.

[e] $P \leq .05$.

Figure 3. Simple slope plots for interaction between time and residential segregation for the overall sample and by age group.



The sensitivity analyses using 6 months as a cutoff to derive the offset in Poisson model found that the regression coefficients for the overall population and specific age groups closely aligned with those from the primary analyses (Multimedia Appendix 2). Residential segregation was not significantly associated with county-level COVID-19 booster coverage in the subgroup analysis for people aged 12 to 17 years. The sensitivity analyses corroborated our main findings and did not change the interpretation.

Discussion

Principal Findings

In a longitudinal analysis across the 418 counties in the Deep South, we found county-level COVID-19 booster coverage was generally low and exhibited variations across different age groups. Compared with the older adult population, booster uptake was lower in youth (ie, 12 to 17 years) and adults (ie, 18 to 64 years). There was a geospatial heterogeneity in the county-level COVID-19 booster coverage, which was negatively associated with residential segregation. Specifically, counties with higher levels of segregation experienced lower booster coverage. However, this disparity decreased over time. This study was innovative as we evaluated the county-level booster coverage by age group from a longitudinal perspective, which was rare in previous research. Our findings were in line with those from the analyses at the individual level and CDC reports, which found that vaccination varied by age and racial and ethnic groups [8,22,30].

The progress of county-level COVID-19 booster coverage in the Deep South was slow and insufficient to curb the transmission of new COVID-19 variants from person to person. This slow progress was found in youth and adults. While the slow progress in these 2 age groups might be due to the low percentage of people who were fully vaccinated and eligible for a booster dose, it might also be due to the fear of long-term effects and serious side effects among the parents and the adults themselves [31,32]. Additionally, some young adults hesitated to get vaccinated against COVID-19 because they did not think it is necessary or beneficial [31,32]. Denford et al [31] found that people were unvaccinated because they considered themselves to be young, healthy, and at low risk of getting sick. However, to effectively control the pandemic in the United States, at least 70% of the population needs to get vaccinated although it is challenging.

Our findings indicated that residential segregation reflects racial and ethnic disparities at the county level and negatively impacts on the COVID-19 booster coverage, which was consistent across all age groups. Defined as the relative amount of physical space occupied by the Black population, the residential segregation in this study had consistent distributions with data from the National Institute on Minority Health and Health Disparities, demonstrating the validity of our residential segregation index [33]. There are more and more research conducted at the population level that examine the association of residential segregation with health outcomes [18,19]. Our findings add value to the existing literature by reinforcing the validity of county-level residential segregation as an indicator of racial and ethnic disparities. The negative impact of residential segregation on COVID-19 booster coverage was consistent with the empirical findings [17-19]. Racial and ethnic minority communities often experienced a scarcity of vaccine distribution sites and vaccine doses, resulting in low vaccination rates in these areas [17]. Additionally, the unequal distribution of educational resources and opportunities in highly segregated communities may preclude health education regarding the safety and effectiveness of the COVID-19 vaccine and serve as a roadblock to county-level booster coverage [34].

Our findings provide critical insights into the design of vaccine surveillance studies and the development of public health interventions to enhance booster coverage. From the research perspective, our findings demonstrated the strength of ecological study and confirmed that well-designed ecological research can yield consistent results with analyses based on individual patient data. Importantly, our conclusions had strong external validity and can be generalized to racial and ethnic minority communities in other parts of the United States. From a public health perspective, there is a pressing need to enhance vaccination initiatives in the Deep South, with a particular focus on youth, parents, and adults. Community-based health education campaigns should emphasize the safety and effectiveness of the COVID-19 vaccine. Given the disparities in health care access within racial and ethnic minority communities, to promote booster coverage, there is an urgent need to prioritize the allocation of health care facilities and resources in these areas.

Limitations

There are several limitations in this study. First, this was an ecological study focusing on county-level COVID-19 booster coverage and its contextual factors. Our findings might suffer from the ecological fallacy [20]. Caution may be needed when interpreting our findings at the individual level. Second, residential segregation was calculated based on the Black population, given the large proportion of the Black population in the Deep South. Consequently, our findings only reveal the relationship between residential segregation and booster coverage in Black communities. Third, although our analysis adjusted for a list of important county-level confounders that were selected based on a strong conceptual model and validated in prior research, other unadjusted confounders, such as political ideology and religious culture, might attenuate or enlarge the effect of residential segregation on booster coverage [15,25]. Our findings may be affected by residual confounding. We recommended future studies examine our findings with the consideration of more validated confounders. Fourth, given the nature of ecological design at the county level, we did not include individual factors in this study. Future studies incorporating both county- and individual-level factors can explore the interactions between them in predicting booster coverage. Finally, we used a proxy number of people who completed the primary series 5-month before each selected time point as an offset to model the booster coverage rate in the Poisson model, since this study was based on vaccine surveillance data, and no personal identifiers were involved. We did not have the information regarding vaccine types (eg, Pfizer-BioNTech, Moderna, and Johnson & Johnson) and the actual time interval between primary series and booster for each individual, which precluded us to model booster coverage uptake precisely. To counteract this limitation, we used the 7-day moving average of individuals who completed the primary series in our analyses. We also did sensitivity analyses using 6 months as an interval. The results were consistent with those from the analyses using 5 months. We suggest that future studies use health care administrative data to further validate our findings.

Conclusions

The progress of county-level COVID-19 booster coverage in the Deep South was slow and varied by age group. The progress was even slower in youths and adults as compared with older adults. The residential segregation precluded booster coverage across the age groups. Future efforts regarding vaccine planning and implementation should focus on the youths and adults. Health care facilities and resources are needed in racial and ethnic minority communities.

Acknowledgments

The authors wish to thank the funders for their support on this study and the reviewers for their insightful comments.

Data Availability

Vaccine uptake data and county-level factors are included in this published paper. Residential segregation score is available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions and sources of the proposed county-level confounders by each dimension.
[DOCX File , 21 KB - publichealth_v9i1e44257_app1.docx]

Multimedia Appendix 2
Residential segregation and COVID-19 booster coverage by age group in the 418 counties across the 5 Deep South states from December 15, 2021, to October 19, 2022.
[DOCX File , 20 KB - publichealth_v9i1e44257_app2.docx]

References

1. Arbel R, Hammerman A, Sergienko R, Friger M, Peretz A, Netzer D, et al. BNT162b2 vaccine booster and mortality due to Covid-19. N Engl J Med 2021 Dec 23;385(26):2413-2420 [FREE Full text] [doi: 10.1056/NEJMoa2115624] [Medline: 34879190]
2. FDA authorizes booster dose of Pfizer-BioNTech COVID-19 vaccine for certain populations. U.S. Food and Drug Administration. 2021. URL: https://tinyurl.com/mtbxydsv [accessed 2022-10-29]
3. Coronavirus (COVID-19) update: FDA takes additional actions on the use of a booster dose for COVID-19 vaccines. U.S. Food and Drug Administration. 2021. URL: https://tinyurl.com/2myx234a [accessed 2022-10-29]
4. Coronavirus (COVID-19) update: FDA expands eligibility for Pfizer-BioNTech COVID-19 booster dose to 16- and 17-year-olds. U.S. Food and Drug Administration. 2021. URL: https://tinyurl.com/33hh98hb [accessed 2022-10-29]
5. Coronavirus (COVID-19) update: FDA shortens interval for booster dose of Moderna COVID-19 vaccine to five months. U.S. Food and Drug Administration. 2022. URL: https://tinyurl.com/ypjt72my [accessed 2022-10-29]
6. Coronavirus (COVID-19) update: FDA takes multiple actions to expand use of Pfizer-BioNTech COVID-19 vaccine. U.S. Food and Drug Administration. 2022. URL: https://tinyurl.com/54e8dwwa [accessed 2022-10-29]
7. Vaccines for COVID-19. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/index.html [accessed 2022-10-24]
8. COVID-19 vaccinations in the United States. Centers for Disease Control and Prevention. 2022. URL: https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-people-booster-percent-pop5 [accessed 2022-10-24]
9. Eichenbaum A, Tate AD. Health inequity in Georgia during the COVID-19 pandemic: an ecological analysis assessing the relationship between county-level racial/ethnic and economic polarization using the ICE and SARS-CoV-2 cases, hospitalizations, and deaths in Georgia as of October 2020. Health Equity 2022;6(1):230-239 [FREE Full text] [doi: 10.1089/heq.2021.0118] [Medline: 35402766]
10. Zeng C, Zhang J, Li Z, Sun X, Olatosi B, Weissman S, et al. Spatial-temporal relationship between population mobility and COVID-19 outbreaks in South Carolina: time series forecasting analysis. J Med Internet Res 2021 Apr 13;23(4):e27045 [FREE Full text] [doi: 10.2196/27045] [Medline: 33784239]
11. Zeng C, Zhang J, Li Z, Sun X, Yang X, Olatosi B, et al. Population mobility and aging accelerate the transmission of coronavirus disease 2019 in the Deep South: a county-level longitudinal analysis. Clin Infect Dis 2022 May 15;74(Suppl_3):e1-e3 [FREE Full text] [doi: 10.1093/cid/ciac050] [Medline: 35568472]
12. Brown KM, Lewis JY, Davis SK. An ecological study of the association between neighborhood racial and economic residential segregation with COVID-19 vulnerability in the United States' capital city. Ann Epidemiol 2021 Jul;59:33-36 [FREE Full text] [doi: 10.1016/j.annepidem.2021.04.003] [Medline: 33895243]
13. Egede LE, Walker RJ, Garacci E, Raymond JR. Racial/ethnic differences in COVID-19 screening, hospitalization, and mortality in Southeast Wisconsin. Health Aff (Millwood) 2020 Nov;39(11):1926-1934 [FREE Full text] [doi: 10.1377/hlthaff.2020.01081] [Medline: 33136498]
14. Franz B, Parker B, Milner A, Braddock JH. The relationship between systemic racism, residential segregation, and racial/ethnic disparities in COVID-19 deaths in the United States. Ethn Dis 2022;32(1):31-38 [FREE Full text] [doi: 10.18865/ed.32.1.31] [Medline: 35106042]

15. Wu C. Racial concentration and dynamics of COVID-19 vaccination in the United States. SSM Popul Health 2022 Sep;19:101198 [FREE Full text] [doi: 10.1016/j.ssmph.2022.101198] [Medline: 35996681]
16. Valier MR, Elam-Evans LD, Mu Y, Santibanez TA, Yankey D, Zhou T, et al. Racial and ethnic differences in COVID-19 vaccination coverage among children and adolescents aged 5-17 years and parental intent to vaccinate their children—National Immunization Survey-Child COVID Module, United States, December 2020-September 2022. MMWR Morb Mortal Wkly Rep 2023 Jan 06;72(1):1-8 [FREE Full text] [doi: 10.15585/mmwr.mm7201a1] [Medline: 36602930]
17. Anderson KF, Ray-Warren D. Racial-ethnic residential clustering and early COVID-19 vaccine allocations in five urban Texas counties. J Health Soc Behav 2022 Dec;63(4):472-490 [FREE Full text] [doi: 10.1177/00221465221074915] [Medline: 35164599]
18. Medcalfe SK, Slade CP. Racial residential segregation and COVID-19 vaccine uptake: an analysis of Georgia USA county-level data. BMC Public Health 2023 Jul 20;23(1):1392 [FREE Full text] [doi: 10.1186/s12889-023-16235-0] [Medline: 37468835]
19. Swietek K, Gianattasio KZ, Henderson S, Khanna S, Ubri P, Douglas M, et al. Association between racial segregation and COVID-19 vaccination rates. J Public Health Manag Pract 2023;29(4):572-579. [doi: 10.1097/PHH.0000000000001738] [Medline: 36943401]
20. Sedgwick P. Ecological studies: advantages and disadvantages. BMJ 2014 May 02;348:g2979. [doi: 10.1136/bmj.g2979] [Medline: 25134102]
21. Smith KF, Dobson AP, McKenzie FE, Real LA, Smith DL, Wilson ML. Ecological theory to enhance infectious disease control and public health policy. Front Ecol Environ 2005 Feb 01;3(1):29-37 [FREE Full text] [doi: 10.1890/1540-9295(2005)003[0029:ETTEID]2.0.CO;2] [Medline: 19838319]
22. COVID-19 vaccinations in the United States, county. Centers for Disease Control and Prevention. 2022. URL: https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh [accessed 2022-10-25]
23. White MJ. Segregation and diversity measures in population distribution. Popul Index 1986;52(2):198-221. [doi: 10.2307/3644339]
24. Massey DS, Denton NA. The dimensions of residential segregation. Soc Forces 1988 Dec;67(2):281-315 [FREE Full text] [doi: 10.2307/2579183]
25. Thakur N, Lovinsky-Desir S, Bime C, Wisnivesky JP, Celedón JC. The structural and social determinants of the racial/ethnic disparities in the U.S. COVID-19 pandemic. What's our role? Am J Respir Crit Care Med 2020 Oct 01;202(7):943-949 [FREE Full text] [doi: 10.1164/rccm.202005-1523PP] [Medline: 32677842]
26. Zeng C, Zhang J, Sun X, Li Z, Weissman S, Olatosi B, et al. County-level predictors of retention in care status among people living with HIV in South Carolina from 2010 to 2016: a data-driven approach. AIDS 2021 May 01;35(Suppl 1):S53-S64 [FREE Full text] [doi: 10.1097/QAD.0000000000002832] [Medline: 33867489]
27. Shi F, Zhang J, Yang X, Sun X, Li Z, Zeng C, et al. Moderation effect of community health on the relationship between racial/ethnic residential segregation and HIV viral suppression in South Carolina: a county-level longitudinal study from 2013 to 2018. Front Public Health 2022;10:1013967 [FREE Full text] [doi: 10.3389/fpubh.2022.1013967] [Medline: 36699939]
28. Larsen SL, Shin I, Joseph J, West H, Anorga R, Mena GE, et al. Quantifying the impact of SARS-CoV-2 temporal vaccination trends and disparities on disease control. Sci Adv 2023 Aug 02;9(31):eadh9920 [FREE Full text] [doi: 10.1126/sciadv.adh9920] [Medline: 37531439]
29. Analyzing and visualizing interactions in SAS. UCLA: Statistical Consulting Group. URL: https://stats.oarc.ucla.edu/sas/seminars/analyzing-and-visualizing-interactions/ [accessed 2023-11-20]
30. Zhang W, Wu YY, Wu B. Racial/ethnic disparities in getting COVID-19 vaccine: do age, gender, and education matter? Health Equity 2022;6(1):500-507 [FREE Full text] [doi: 10.1089/heq.2022.0025] [Medline: 36186620]
31. Denford S, Mowbray F, Towler L, Wehling H, Lasseter G, Amlôt R, et al. Exploration of attitudes regarding uptake of COVID-19 vaccines among vaccine hesitant adults in the UK: a qualitative analysis. BMC Infect Dis 2022 Apr 26;22(1):407 [FREE Full text] [doi: 10.1186/s12879-022-07380-9] [Medline: 35473486]
32. Suran M. Why parents still hesitate to vaccinate their children against COVID-19. JAMA 2022 Jan 04;327(1):23-25. [doi: 10.1001/jama.2021.21625] [Medline: 34910088]
33. An ecosystem of health disparities and minority health resources. National Institute on Minority Health and Health Disparities. HDPulse. 2023. URL: https://hdpulse.nimhd.nih.gov/ [accessed 2023-09-16]
34. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. Public Health Rep 2001;116(5):404-416 [FREE Full text] [doi: 10.1093/phr/116.5.404] [Medline: 12042604]

Abbreviations

CDC: Centers for Disease Control and Prevention

Edited by A Mavragani, T Sanchez; submitted 17.11.22; peer-reviewed by N Yiannakoulias, Y Guo; comments to author 18.08.23; revised version received 17.09.23; accepted 27.10.23; published 05.12.23.

Please cite as:
Zeng C, Zhang J, Li Z, Sun X, Ning H, Yang X, Weissman S, Olatosi B, Li X
Residential Segregation and County-Level COVID-19 Booster Coverage in the Deep South: Surveillance Report and Ecological Study
JMIR Public Health Surveill 2023;9:e44257
URL: https://publichealth.jmir.org/2023/1/e44257
doi:10.2196/44257
PMID:38051568

Original Paper

Independent and Joint Associations of BMI and Waist Circumference With the Onset of Type 2 Diabetes Mellitus in Chinese Adults: Prospective Data Linkage Study

Sixuan Li[1], MPH; Yong Wang[1], MPhil; Yanyan Ying[1], MPhil; Qinghai Gong[1], MPhil; Ge Lou[2], MPhil; Yang Liu[3], MPH; Shiwei Liu[4*], PhD; Hui Li[1*], MPH

[1]Ningbo Municipal Center for Disease Control and Prevention, Ningbo, China
[2]Center for Disease Control and Prevention of Chaoyang District, Beijing, China
[3]Shanghai Pudong New Area Center for Disease Control and Prevention, Pudong Institute of Preventive Medicine, Fudan University, Shanghai, China
[4]Chinese Center for Disease Control and Prevention, Bejing, China
[*]these authors contributed equally

Corresponding Author:
Hui Li, MPH
Ningbo Municipal Center for Disease Control and Prevention
237 Yongfeng Road
Haishu District
Ningbo, 315010
China
Phone: 86 057487274542
Email: lihui4329@163.com

Abstract

Background: General obesity and abdominal obesity, typically measured by BMI and waist circumference (WC), respectively, are associated with an increased risk of type 2 diabetes mellitus (T2DM). However, the magnitude of the association of these two obesity indicators and their joint association with the onset of T2DM remain controversial.

Objective: The aim of this study was to investigate the associations between these two obesity indicators and T2DM among the Chinese population to contribute scientific evidence for appropriate T2DM interventions.

Methods: A cohort of 3001 eligible participants was selected from the Ningbo Adult Chronic Disease Surveillance Project running since 2015. Based on BMI, individuals were categorized into groups of underweight or normal, overweight, and obesity. Based on WC, individuals were categorized in groups of normal, precentral obesity, and central obesity. Follow-up was performed by linking data of the baseline data set with the diabetes registry data set and the vital registry data set (both from the Ningbo Municipal Integrated Noncommunicable Disease Collaborative Management System), mainly using the participants' identity numbers. Follow-up was completed when a participant was diagnosed with T2DM. The associations were estimated with multivariate Cox proportional hazard regression.

Results: In the cohort, 90 of 3001 participants developed T2DM (incidence density: 6.483/1000 person-years) with a median 4.72 years of follow-up. After controlling for age, sex, hypertension, dyslipidemia, smoking status, and family history of diabetes, the multivariate adjusted hazard ratios (HRs) across underweight/normal, overweight, and obesity BMI categories were 1.000, 1.653 (95% CI 1.030-2.654), and 2.375 (95% CI 1.261-4.473), respectively. The multivariate adjusted HRs across the normal, precentral obesity, and central obesity WC categories were 1.000, 1.215 (95% CI 0.689-2.142), and 1.663 (95% CI 1.016-2.723), respectively. Compared with the reference group (normal WC with an underweight/normal BMI), the multivariate adjusted HR for participants with both central obesity according to WC and obesity according to BMI was 2.489 (95% CI 1.284-4.825).

Conclusions: Both elevated BMI and WC at baseline increased the risk of T2DM. Compared with WC, BMI as an obesity indicator was more strongly associated with the onset of T2DM.

(2023;9:e39459) doi:10.2196/39459

KEYWORDS

body mass index; waist circumference; type 2 diabetes mellitus; prospective study; data linkage; epidemiology; hazard ratio

Introduction

Diabetes mellitus is a group of metabolic disorders characterized by elevated blood glucose levels, which can lead to various health problems and severe complications with disease progression [1]. A nationwide survey in 2015-2017 reported that the prevalence of diabetes was 12.8% in China [2]. Obesity is as an important modifiable risk factor for type 2 diabetes mellitus (T2DM), and the prevalence of overweight, obesity, and abdominal obesity among the Chinese population was estimated at 28.1%, 5.2%, and 29.1%, respectively, in 2012-2015 [3,4].

BMI and waist circumference (WC) are two of the key indicators used to measure general obesity and abdominal obesity, respectively. Some studies found that compared with that of people without T2DM, the proportion of body fat distributed in the abdominal area was higher among patients who were not classified as obese according to BMI criteria, whereas some people with obesity were metabolically healthy [5,6]. Since BMI does not take the exact fat mass into account, especially the visceral adipose tissue, this indicator has a limitation in reflecting the body fat distribution [7]. It has been reported that WC might better reflect visceral obesity than BMI [8]. Recently, a growing body of studies indicated that WC was strongly associated with T2DM and is a better predictor for the onset of T2DM than BMI [9-11].

Although both general obesity and abdominal obesity have been confirmed to be associated with the risk of T2DM, the magnitude of these two obesity indicators and their joint association with T2DM remain controversial [12,13]. Moreover, the number of prospective studies addressing this issue is limited for the Chinese population. Therefore, we performed this prospective data-linkage study to investigate the independent and joint associations of these two obesity indicators with the risk of T2DM to provide better scientific evidence for T2DM interventions in Ningbo, China.

Methods

Study Design and Participants

Baseline Survey

We conducted the baseline survey (the Ningbo Adult Chronic Disease Surveillance Project) on October 26 and December 31, 2015. This was a population-based project with the aims of understanding the prevalence, awareness, treatment, and control situation, and to explore the relevant behavior risk factors of major chronic diseases (hypertension, diabetes mellitus, and dyslipidemia) among Chinese adult residents aged 15-79 years in Ningbo city.

A total of 5280 residents were selected using a multistage cluster random sampling method covering 11 districts and counties in Ningbo. In the first stage, three towns or streets were selected from each county or district using a proportionate-to-population size sampling scheme. In the second stage, two administrative villages or neighborhood communities were selected from each town or street using the same method as used in the first stage. In the third stage, one village containing at least 105 households was selected from each neighborhood community or administrative village. In the final stage, 105 randomly selected households from each village and one resident aged 15-74 years in each household whose date of birth was closest to the 15th was selected as the survey participant. Finally, a total of 5160 residents completed the survey with a response rate of 97.73% (5160/5280).

Inclusion and Exclusion Criteria for Baseline Participants

Inclusion criteria were as follows: (1) aged ≥40 years, (2) living in the current district or county for at least 6 months, and (3) capable of communicating in Mandarin or local dialects with reading skills. The exclusion criterion was having been diagnosed with diabetes by clinicians in hospitals or community health service centers at the time of the baseline survey. A total of 3001 eligible participants aged above 40 years were selected from the pool of 5160 participants. All participants signed the written informed consent form.

Ethics Approval

The research protocol was approved by the ethics committee of Ningbo Center for Disease Control and Prevention (approval number 201702).

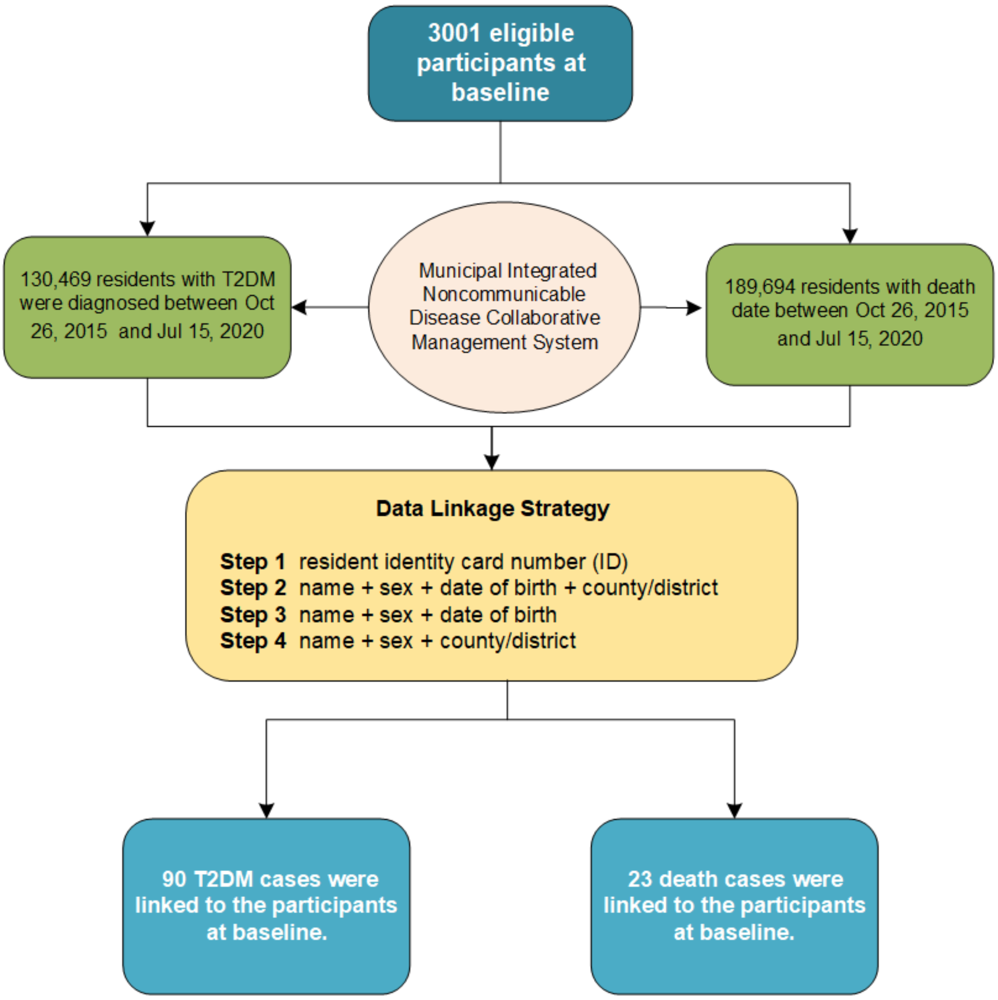
Data Collection

The data collection methods included a questionnaire survey, physical measurements, and laboratory tests. The questionnaire consisted of five components, including physical activities, environment and facilities, signs and symptoms, dietary and drinking habits, and demographic characteristics. All questionnaires were filled in face-to-face by trained staff. The completeness and accuracy of the questionnaire were double-checked by quality control personnel following the quality control requirements. We measured the height, weight, WC, blood pressure, fasting plasma glucose (FPG), glycated hemoglobin (HbA_{1c}), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) for all participants. The physical examinations were carried out by the investigators using uniform devices and methods. Participants fasted the night before measurements were taken. The height, weight, and WC were measured while the participant was wearing light clothing (shorts and T-shirt) and without shoes. The participants were required to avoid vigorous exercises, eating, and drinking beverages (especially caffeinated drinks) within 30 minutes before measuring blood pressure and to rest quietly for 5 minutes before the first measurement. The blood pressure was tested three times and the average was taken. Five milliliters of the fasting venous blood was collected from the participants to detect FPG, HbA_{1c}, TC, TG, HDL-C, and LDL-C.

Follow-up by Data Linkage

The data sources for the follow-up were obtained from the Ningbo Municipal Integrated Noncommunicable Disease Collaborative Management System (NCDCMS), which has continuously collected surveillance data since 2009 and covered the whole population of Ningbo. This system, which has been proven to obtain high-quality data and to have a low rate of underreporting (<5%), collects four types of noncommunicable disease data (diabetes mellitus, ischemic heart disease and cardiac arrest, cerebrovascular disease, malignant neoplasms) and cause of death data in a real-time manner [14]. We obtained the data from the diabetes registry with a diagnosis date from October 26, 2015, to July 15, 2020; the data of vital registration were collected according to a death date from October 26, 2015, to July 15, 2020. A total of 130,469 cases of T2DM and 189,694 cases of death were obtained from NCDCMS. The follow-up was completed by linking data of the baseline data set with the diabetes registry data set and the vital registration data set using the personal identification (ID) number. The participants who failed to link with these two data sets by ID number were linked via three combinations of identifiers in turn (ie, name+sex+date of birth+county/district; name+sex+date of birth; name+sex+county/district). A double check was performed to confirm whether the participants who successfully linked with two data sets via three identifiers were correct. Finally, 90 cases of T2DM and 23 deaths were linked to the participants at baseline through the data-linkage procedure. The flowchart of data linkage is provided in Figure 1.

Figure 1. Flowchart of data linkage. T2DM: type 2 diabetes mellitus.



Definition of Outcome (Endpoint of Follow-up)

The endpoint of follow-up was a diagnosis of T2DM. The participants who failed to be linked with the diabetes registry data set were regarded as censored. According to the Chinese Type 2 Diabetes Prevention and Control 2017 Guidelines, T2DM was defined as typical clinical manifestations (thirst, frequent urination, overeating and unexplained weight loss), and (1) random plasma glucose≥11.1 mmol/L (200 mg/dL), (2) FPG≥7.0 mmol/L (126 mg/dL), or (3) oral glucose tolerance test 2-hour plasma glucose≥11.1 mmol/L (200 mg/dL) [15].

Follow-up Time Calculation

The follow-up time was calculated as the date of diagnosis minus the start date of the baseline survey (October 26, 2015) for participants who linked with the diabetes registry data set, as the date of death minus the start date of the baseline survey (October 26, 2015) for participants who linked with the vital registration data set, and as the set end time of follow-up (July 15, 2020) minus the start date of the baseline survey (October 26, 2015) for participants who failed to link with these two data sets.

Exposure

According to the 2013 Health Industry Standards of the People's Republic of China "Criteria of Weight for Adults" and the World Health Organization report of WC and waist-to-hip ratio (WHR), BMI was categorized as underweight or normal (≤23.9 kg/m^2), overweight (24-27.9 kg/m^2), and obesity (≥28 kg/m^2). WC was divided into three groups as follows: (1) normal: WC<85 cm for men and WC<80 cm for women; (2) precentral obesity: WC of 85-89.9 cm for men, WC of 80-84.9 cm for women; and (3) central obesity: WC≥90 cm for men WC≥85 cm for women [16,17].

Covariates

Demographic and Lifestyle Factors

Demographic factors included age, sex, education level, marital status, residence place, and occupation. Lifestyle factors included smoking status, alcohol consumption, and physical activity level. Alcohol consumption was derived from the question: "Have you had any kind of alcoholic drink in the past 12 months?" Participants were divided into three groups: No; Yes, within 30 days; Yes, 30 days ago. Participants who currently smoke cigarettes or quit smoking were defined as ever smokers. Participants who never smoke cigarettes were defined as never smokers. Physical activity was classified into the following three levels in strict accordance with the Chinese guidelines for data processing and analysis of the International Physical Activity Questionnaire: low, moderate, and high. A moderate physical level was defined as (1) 3 or more days of vigorous-intensity activity of at least 20 minutes per day; (2) 5 or more days of moderate-intensity activity and/or walking at least 30 minutes per day; or (3) 5 or more days of any combination of walking, moderate-intensity activities, or vigorous-intensity activities achieving a minimum total physical activity of at least 600 metabolic equivalent (MET)-minutes/week. A high physical activity level was defined as: (1) vigorous-intensity activity at least 3 days per week achieving a minimum total physical activity of at least 1500 MET-minutes/week or (2) 7 or more days of any combination of walking, moderate-intensity activities, or vigorous-intensity activities achieving a minimum total physical activity of at least 3000 MET-minutes/week [18].

Medical Conditions

Hypertension was defined as systolic blood pressure≥140 mmHg, diastolic blood pressure≥90 mmHg, current use of antihypertensive medication, or self-reported history of hypertension [19]. Dyslipidemia was defined as TC≥5.2 mmol/L, TG≥1.7 mmol/L, HDL-C≤1.0 mmol/L, LDL-C≥3.4 mmol/L, current use of lipid-lowering medication, or self-reported history of dyslipidemia.

Statistical Analyses

The incidence density of T2DM was obtained by dividing the number of T2DM cases by the person-years of follow-up. Normally distributed continuous data are described as mean (SD) and skewed-distributed continuous data are described as median (IQR). Categorical data are described as count (percentage). Differences in the distribution of baseline characteristics between participants with and without T2DM were evaluated using the Mann-Whitney *U* test for continuous variables and with the Pearson χ^2 test for categorical variables. The strength of the association was measured using Cramer V for the χ^2 test. Tests for linear trend across categories of BMI and WC were performed by treating the categories as a continuous variable with the median value for each category.

Univariate Cox regressions were performed for each exposure (BMI categories, central obesity categories) and covariates (residence place, sex, age, education level, marital status, occupation, family history of diabetes, alcohol consumption, smoking status, physical activity, hypertension, and dyslipidemia) to calculate the unadjusted hazard ratios (HRs) with 95% CIs. Age and sex–adjusted HRs with 95% CIs were estimated with Cox proportional hazard models stratified by BMI and central obesity categories adjusting for age and sex. Multivariate Cox proportional hazard models were performed to calculate HRs with 95% CIs based on the different categories of BMI and central obesity by adjusting for potential confounders (age, sex, hypertension, dyslipidemia, smoking status, family history of diabetes).

To estimate the joint effect of general obesity according to BMI and central obesity according to WC on the risk of T2DM, the participants were classified into six groups based on the BMI and WC categories. The adjusted HRs along with potential confounders were estimated using Cox models. BMI and WC were also included in the Cox regression model as continuous variables, either separately or together. The proportionality assumption of the hazards in the Cox regression analyses was tested by Schoenfeld residuals and the assumption was satisfied for all analyses. Forest plots were used to compare the independent and joint associations of BMI and WC. All data management and statistical analyses were performed using the statistical software R version 3.4.3. *P*<.05 was considered statistically significant.

Results

Baseline Characteristics of the Participants

The final cohort included 3001 participants at baseline with a median age of 54 years. These participants were followed up for a median of 4.72 years, comprising 13,880.8 person-years of follow-up. Overall, 90 of the 3001 participants developed T2DM, with an incidence density of 6.483 per 1000 person-years by the end of follow-up. As shown in Table 1, the participants who developed T2DM were older with higher baseline BMI, WC, FBG, and HbA_{1c}, and with a greater proportion reporting a family history of diabetes compared with those who did not develop T2DM during the study period.

At baseline, 1570 (52.32%) of the participants were underweight or normal, 1159 (38.62%) were overweight, and 272 (9.06%) had obesity. Compared to the underweight or normal group, participants in the overweight and obesity groups tended to have higher WC, SBP, DBP, HbA_{1c}, LDL-C, TC, and TG; lower HDL-C; and a higher proportion of hypertension and dyslipidemia. There were 1397 (46.55%) participants with normal WC, 721 (24.03%) with precentral obesity, and 865 (28.82%) with central obesity, demonstrating similar patterns as observed for BMI categories.

Table 1. Baseline characteristics of the 3001 participants by diabetes status.

Variables	Total (N=3001)	T2DM[a] (n=90)	No T2DM (n=2911)	*P* value[b]	Cramer V
Residence place, n (%)				.50	0.01
Urban	1573 (52.4)	44 (48.9)	1529 (52.5)		
Rural	1428 (47.6)	46 (51.1)	1382 (47.5)		
Sex, n (%)				.56	0.009
Male	1277 (42.6)	41 (45.6)	1236 (42.5)		
Female	1724 (57.4)	49 (54.4)	1675 (57.5)		
Education level, n (%)				.37	0.026
Primary school and above	1423 (47.4)	44 (48.9)	1379 (47.4)		
Middle school	1048 (34.9)	35 (38.9)	1013 (34.8)		
High school and above	530 (17.7)	11 (12.2)	519 (17.8)		
Marital status, n (%)				.52	0.028
Single	46 (1.5)	2 (2.2)	44 (1.5)		
Married	2734 (91.1)	78 (86.7)	2656 (91.2)		
Widowed	151 (5.0)	7 (7.8)	144 (4.9)		
Divorced/separated/ cohabitated	70 (2.3)	3 (3.3)	67 (2.3)		
Occupation, n (%)				.20	0.039
Agriculture/husbandry/transportation	803 (26.8)	24 (26.7)	779 (26.8)		
Government agency/enterprises	324 (10.8)	6 (6.7)	318 (10.9)		
Other industries	589 (19.6)	13 (14.4)	576 (19.8)		
Unemployed/housework/retired	1285 (42.8)	47 (52.2)	1238 (42.5)		
Family history of diabetes, n (%)				<.001	0.074
No	2347 (78.2)	57 (63.3)	2290 (78.7)		
Yes	407 (13.6)	25 (27.8)	382 (13.1)		
I don’t know	247 (8.2)	8 (8.9)	239 (8.2)		
Alcohol consumption, n (%)				.67	0.016
No	2062 (68.7)	62 (68.9)	2000 (68.7)		
Yes, within 30 days	780 (26.0)	25 (27.8)	755 (25.9)		
Yes, 30 days ago	159 (5.3)	3 (3.3)	156 (5.4)		
Smoking status, n (%)				.38	0.013
Never smoker	2158 (71.9)	61 (67.8)	2097 (72.0)		
Ever smoker	843 (28.1)	29 (32.2)	814 (28.0)		
Physical activity, n (%)				.53	0.021
Low level	418 (13.9)	16 (17.8)	402 (13.8)		
Median level	2455 (81.8)	71 (78.9)	2384 (81.9)		
High level	128 (4.3)	3 (3.3)	125 (4.3)		
Hypertension, n (%)				<.001	0.079
No	1492 (49.7)	24 (26.7)	1468 (50.4)		
Yes	1509 (50.3)	66 (73.3)	1443 (49.6)		
Dyslipidemia, n (%)				.01	0.045
No	1231 (41.0)	25 (27.8)	1206 (41.4)		
Yes	1770 (59.0)	65 (72.2)	1705 (58.6)		
BMI categories, n (%)				<.001	0.072

Variables	Total (N=3001)	T2DM[a] (n=90)	No T2DM (n=2911)	*P* value[b]	Cramer V
Underweight or normal	1570 (52.3)	30 (33.3)	1540 (52.9)		
Overweight	1159 (38.6)	45 (50.0)	1114 (38.3)		
Obesity	272 (9.1)	15 (16.7)	257 (8.8)		
WC[c] categories, n (%)				.002	0.064
Normal	1415 (47.2)	29 (32.2)	1386 (47.6)		
Precentral obesity	721 (24.0)	21 (23.3)	700 (24.0)		
Central obesity	865 (28.8)	40 (44.4)	825 (28.3)		
BMI (kg/m^2), median (IQR)	23.81 (21.94-25.89)	25.14 (23.12-26.95)	23.77 (21.90-25.84)	<.001	N/A[d]
WC (cm), median (IQR)	82.25 (76.50-88.25)	86.00 (80.30-90.84)	82.20 (76.50-88.10)	<.001	N/A
FBG[e] (mmol/L), median (IQR)	4.94 (4.60-5.38)	5.90 (5.23-6.99)	4.93 (4.59-5.36)	<.001	N/A
HbA_{1c}[f] (%), median (IQR)	4.70 (4.30-5.10)	5.40 (4.80-6.18)	4.70 (4.30-5.10)	<.001	N/A
Age (years), median (IQR)	54.00 (48.00-63.00)	58.00 (52.00-65.50)	54.00 (48.00-63.00)	<.001	N/A

[a] T2DM: type 2 diabetes mellitus.

[b] Differences in the distribution of participants with and without T2DM were evaluated using the Mann-Whitney *U* test for continuous variables and with the Pearson χ^2 test for categorical variables.

[c] WC: waist circumference.

[d] N/A: not applicable.

[e] FBG: fasting blood glucose.

[f] HbA_{1c}: glycated hemoglobin.

Univariate Cox Regression Analysis

BMI categories, WC categories, age, occupation, family history of diabetes, hypertension, and dyslipidemia were significantly associated with the incidence of T2DM. Compared with participants of normal weight, the overweight (HR 2.052, 95% CI 1.293-3.258) and obesity (HR 2.924, 95% CI 1.573-5.434) categories were significantly associated with an increased risk of the development of T2DM. There was a dose-response relationship between the BMI categories and the onset of T2DM (*P* for trend <.001). Compared with participants in the normal WC group as reference, central obesity (HR 2.289, 95% CI 1.419-3.691) was significantly associated with an increased risk of the development of T2DM (*P* for trend <.001). The risk of T2DM onset was significantly increased with the increase of age. The HRs for age groups of 40-49 years, 50-59 years, 60-69 years, and ≥70 years were 1.000, 3.670, 3.423, and 5.279, respectively. The risk of T2DM in participants with hypertension and dyslipidemia was respectively 2.754 (95% CI 1.726-4.393) and 1.823 (95% CI 1.149-2.891) higher that of participants free of diseases.

Independent and Joint Associations of BMI and WC on the Risk of T2DM

The independent and joint associations of BMI and WC with T2DM development are displayed in Table 2 and Figure 2. The age and sex–adjusted HRs increased across categories of BMI and the same pattern was found after adjustment for potential confounders with a slight decrease, although the trend remained significant (*P* for trend <.001). The age and sex–adjusted HRs also increased across categories of WC (*P* for trend .003) and the trend remained significant after adjustment of potential confounders (*P* for trend .04).

Among the participants who were categorized in the central obesity group, the age and sex–adjusted HRs increased successively for those categorized in the underweight or normal, overweight, and obesity groups according to BMI compared with the reference group (normal WC with an underweight or normal BMI). Compared to the reference group, the multivariate adjusted HR for participants with both central obesity according to WC and obesity according to BMI was much higher than that for the independent association of central obesity or obesity (Table 2).

When BMI and WC were treated as continuous variables, the HR of BMI was still slightly higher than that of WC. When the BMI and WC were considered together in the regression model using mutual adjustment, BMI, but not WC, was significantly associated with the risk of T2DM (Table 3).

Table 2. Risk of type 2 diabetes mellitus stratified by categories of BMI and waist circumference (WC) in univariate, age and sex–adjusted, and multivariate-adjusted regressions.

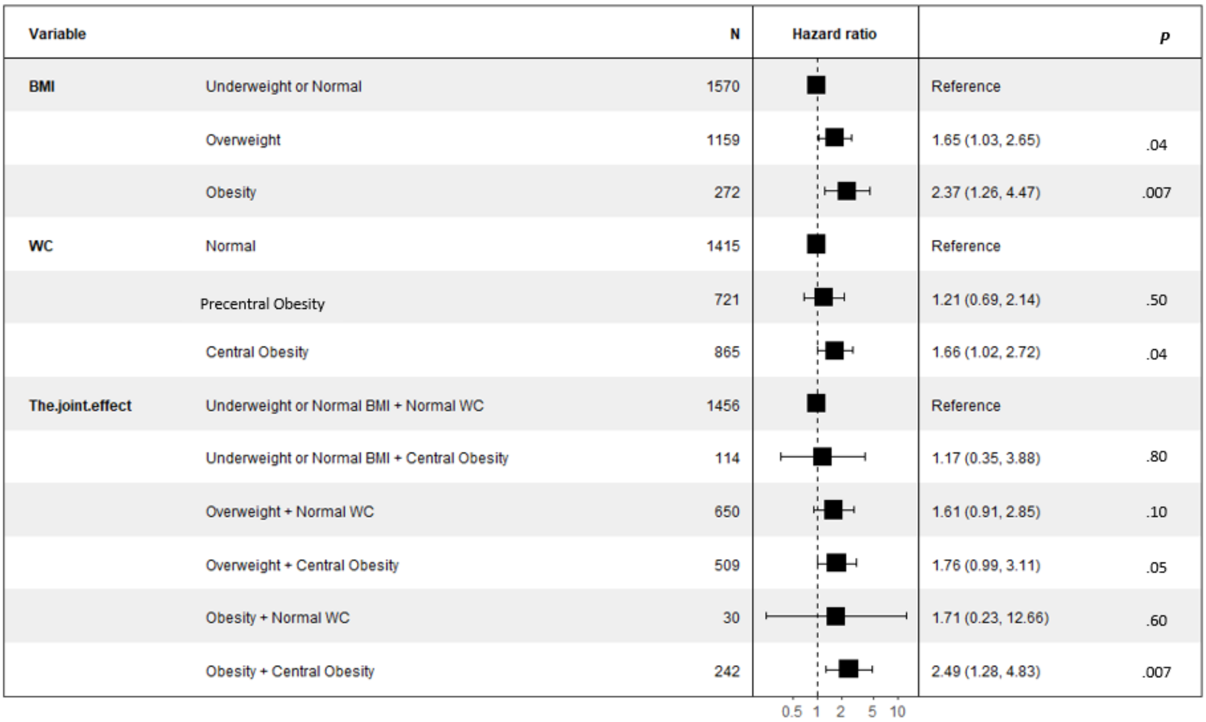
Categories	Univariate, HR[a] (95% CI)	Age and sex–adjusted[b], HR (95% CI)	Multivariate[c], HR (95% CI)
BMI categories			
Underweight or normal (n=1570)	1	1	1
Overweight (n=1159)	2.052 (1.293-3.258)	1.941 (1.221-3.087)	1.653 (1.030-2.654)
Obesity (n=272）	2.924 (1.573-5.434)	2.918 (1.569-5.425)	2.375 (1.261-4.473)
P for trend	<.001	<.001	.004
WC categories			
Normal WC (n=1415)	1	1	1
Precentral obesity (n=721)	1.425 (0.813-2.499)	1.374 (0.767-2.364)	1.215 (0.689-2.142)
Central obesity (n=865)	2.289 (1.419-3.691)	2.059 (1.273-2.059)	1.663 (1.016-2.723)
P for trend	<.001	.003	.04
Joint effect of BMI and WC			
Underweight or normal BMI+normal WC (n=1456)	1	1	1
Underweight or normal BMI+central obesity (n=114)	1.418 (0.430-4.676)	1.316 (0.398-4.352)	1.170 (0.353-3.878)
Overweight+normal WC (n=650)	1.833 (1.044-3.218)	1.819 (1.035-3.198)	1.608 (0.908-2.847)
Overweight+central obesity (n=509)	2.479 (1.421-4.323)	2.185 (1.249-3.823)	1.756 (0.991-3.110)
Obesity+normal WC (n=30)	1.777 (0.241-13.07)	1.913 (0.259-14.08)	1.713 (0.231-12.65)
Obesity+central obesity (n=242)	3.169 (1.662-6.044)	3.116 (1.633-5.945)	2.489 (1.284-4.825)

[a] HR: hazard ratio.

[b] Adjusted for age and sex.

[c] Adjusted for age, sex, hypertension, dyslipidemia, smoking, and family history of diabetes.

Figure 2. Forest plot of the association between categories of BMI and/or waist circumference (WC) and the onset of type 2 diabetes mellitus adjusting for age, gender, hypertension, dyslipidemia, smoking, and family history of diabetes.



Variable		N	Hazard ratio		*P*
BMI	Underweight or Normal	1570		Reference	
	Overweight	1159		1.65 (1.03, 2.65)	.04
	Obesity	272		2.37 (1.26, 4.47)	.007
WC	Normal	1415		Reference	
	Precentral Obesity	721		1.21 (0.69, 2.14)	.50
	Central Obesity	865		1.66 (1.02, 2.72)	.04
The.joint.effect	Underweight or Normal BMI + Normal WC	1456		Reference	
	Underweight or Normal BMI + Central Obesity	114		1.17 (0.35, 3.88)	.80
	Overweight + Normal WC	650		1.61 (0.91, 2.85)	.10
	Overweight + Central Obesity	509		1.76 (0.99, 3.11)	.05
	Obesity + Normal WC	30		1.71 (0.23, 12.66)	.60
	Obesity + Central Obesity	242		2.49 (1.28, 4.83)	.007

Table 3. Risk of type 2 diabetes based on BMI and waist circumference (WC) as continuous variables.

Models	Univariate, HR[a] (95% CI)	Age and sex–adjusted[b], HR (95% CI)	Multivariate[c], HR (95% CI)
Independent effects			
BMI	1.142 (1.078-1.210)	1.141 (1.076-1.210)	1.118 (1.049-1.191)
WC	1.041 (1.019-1.063)	1.038 (1.015-1.061)	1.029 (1.004-1.054)
Joint effects[d]			
BMI	1.116 (1.021-1.221)	1.143 (1.041-1.254)	1.130 (1.027-1.243)
WC	1.011 (0.979-1.045)	1.000 (0.965-1.035)	0.995 (0.960-1.031)

[a] HR: hazard ratio.

[b] Adjusted for age and sex.

[c] Adjusted for age, sex, hypertension, dyslipidemia, smoking, and family history of diabetes.

[d] BMI and WC were considered together in the regression model.

Discussion

Principal Findings and Comparison to Previous Work

In this study, we found that both elevated BMI and WC at baseline increased the risk of T2DM. Moreover, the magnitude of association of BMI with T2DM was stronger than that of WC, regardless of treating the exposures as continuous or categorical variables, which was consistent with previous studies conducted in a Chinese population [20-22]. A prospective study performed over 7.26 years of follow-up in a southeast province of China indicated that BMI may be superior to WC in predicting the incidence of T2DM in the Chinese population [20]. A cohort study including 27,009 retirees in a central province concluded that BMI was the strongest predictor among various anthropometric indices with 4.6 years of follow-up data [22]. However, other studies reported inconsistent conclusions [23-26]. Kodama et al [23] reported that the pooled relative risk of T2DM per 1-SD increase in BMI was significantly lower than that for WC, indicating that WC was more strongly associated with diabetes risk than BMI. Lee et al [24] concluded that although WC had a stronger association with the risk of T2DM in comparison with BMI, there was no overall advantage for WC at discriminating the risk of developing T2DM compared with BMI.

The discrepancies in conclusions among studies were likely due to the differences in ethnicity of the populations. Since there is an assumption that the superiority of WC to BMI for T2DM prediction was based on the characteristics of the study population, previous studies mainly conducted in white-dominant populations or in the western sociocultural context might lead to discrepancies [23-25,27]. This study adds to this evidence for the Chinese population. Another reason for the inconsistent conclusions might be the length of follow-up. Aging is associated with progressive changes in the total and regional fat distribution [22]. As age increases, the body fat distribution changes with a decrease in lower subcutaneous fat and an increase in abdominal fat [28]. The results of previous studies might also be distorted by the increasing follow-up time as more participants tended to have abdominal obesity, which diluted the overall impact of baseline body fat distribution [23]. Differences in research methods and populations can also lead to different conclusions. Although a large-population prospective study in Chinese adults showed that WC was a better predictor than BMI for the onset of T2DM in women, this cohort study used a logistic regression model rather than a Cox regression model to predict the magnitude of associations [29]. The effect of time of follow-up was not considered in the model, which might have also affected the results. Moreover, as a cohort study carried out in an employee population, the “healthy worker effect” might have impacted these results [29]. Furthermore, the inconsistent conclusions might also be related to the different classification criteria of BMI and WC in different studies. Some studies obtained conclusions using categories of BMI or WC, whereas others used the SD.

In our study, after 4.72 years of follow-up, 90 participants developed T2DM with a total incidence density of 6.483 per 1000 person-years, similar to the study of Yang et al [21] (0.89/100 person-years). However, this incidence density was much lower than that reported in some previous longitudinal studies [30,31]. In a 10-year population-based Asian Indian cohort study, the incidence density of diabetes was 22.2 per 1000 person-years [32]. The cumulative incidence of T2DM was 14.4 and 13.7 per 1000 person-years for men and women, respectively, in an 18-year follow-up cohort study performed in Mexico [30]. The relatively low incidence density in our study may be attributed to the small number of participants at baseline, relatively young baseline population, and short follow-up time.

Limitations and Future Work

There were some limitations of this study. First, the number of participants at baseline was relatively small and the follow-up time was short. Referring to the joint associations of BMI and WC adjusting for age and sex, the HRs revealed a dose-response relationship in the overweight category. However, due to the short time of follow-up and relatively young baseline participants in this study, only 90 participants were diagnosed with T2DM. The number of newly diagnosed T2DM participants in each BMI category might not have been sufficient to detect the dose-response relationship with the increase of WC. Second, the distribution and composition of body fat might gradually change during follow-up, along with the corresponding BMI and WC. In this study, the change values of BMI and WC during follow-up were not adjusted in the models owing to the study design, which may have influenced the results. Data linkage, the key technique used in this study, has become a popular research method for generating population-level electronic cohorts, and is used as a supplement to conventional cohort studies [14,29]. Data linkage provides a convenient and cost-effective approach for collecting data from multichannel sources and investigating the causes and development in the variation of outcomes across the individual's life course [29,33]. Nevertheless, compared with a conventional cohort study, data linkage has some limitations and might introduce bias into results due to linkage errors (false matches or missed matches). Despite NCDCMS being confirmed as a platform with high data quality and a very low underreporting rate, some linkage errors are still inevitable [14]. In this study, we tried to avoid linkage errors by using three combinations of identifiers as a complement to the ID and double checking whether the cases linked via three combinations of identifiers were correct. Moreover, we controlled the proven and possible confounding factors in the models collected from the baseline survey and strictly controlled the quality of data during the investigation. Additionally, since the follow-up was completed by data linkage based on NCDCMS in Ningbo, the results might not be generalizable to other cities.

The Chinese population generally has a higher percentage of body fat and a greater propensity to central obesity compared with those of western counterparts under the same BMI [34]. Although BMI as an obesity indicator was more strongly associated with the risk of T2DM than WC in this study, we recommend that the general population and people at high risk of T2DM should focus on control of body weight as well as WC, which is consistent with the Chinese Type 2 Diabetes Prevention and Control Guideline [15]. In addition to obesity, the risk of developing T2DM was significantly higher in people who showed comorbid hypertension or dyslipidemia in comparison with participants who were free of diseases, indicating the importance of blood pressure control and maintaining a normal lipid level for the general population and especially for people at high risk of T2DM. In addition, a healthy lifestyle, including a balanced diet, regular physical activities, tobacco cessation, and alcohol control, is highly recommended, especially in the elderly population.

A study with longer-term follow-up is needed to further investigate the joint association of BMI and WC with the risk of T2DM. Although BMI, WC, and the WHR are inexpensive and easy measures to assess the adipose tissue, there is growing interest in using magnetic resonance techniques for evaluating total body fat and different fat compartments in western countries. Magnetic resonance imaging provides high accuracy in quantitatively detecting the subcutaneous and visceral adipose tissue, which could be an option to further explore the association between obesity and T2DM accurately [35]. Additionally, performing a meta-analysis for studies in the Chinese population is suggested to better detect the pooled relative risk by removing the influence of ethnicity.

Conclusion

In conclusion, both elevated BMI and WC at baseline increased the risk of T2DM. Compared with WC, BMI as an obesity indicator was more strongly associated with the onset of T2DM among Chinese adults in Ningbo. We recommend that residents of this population, especially the elderly, maintain a normal weight and a healthy lifestyle.

Acknowledgments

The authors wish to thank Dr Stacy Davlin and Dr Charlalynn Harris for providing their professional advice for this study, as well as Miss Chensheng Zhang for the abstract revision. In addition, the authors thank all investigators and study participants for their support and contributions to this study. This study was supported by the National Natural Science Foundation of China (81872721), National Key R&D Program of China (2017YFC1310902 and 2018YFC1315305), and Ningbo Medical and Health Brand Discipline (PPXK2018-10).

Conflicts of Interest

None declared.

References

1. Diabetes. World Health Organization. URL: https://www.who.int/health-topics/diabetes#tab=tab_1 [accessed 2021-03-29]
2. Li Y, Teng D, Shi X, Qin G, Qin Y, Quan H, et al. Prevalence of diabetes recorded in mainland China using 2018 diagnostic criteria from the American Diabetes Association: national cross sectional study. BMJ 2020 Apr 28;369:m997 [FREE Full text] [doi: 10.1136/bmj.m997] [Medline: 32345662]
3. Zhang L, Wang Z, Wang X, Chen Z, Shao L, Tian Y, China Hypertension Survey investigators. Prevalence of abdominal obesity in China: results from a cross-sectional study of nearly half a million participants. Obesity 2019 Nov;27(11):1898-1905. [doi: 10.1002/oby.22620] [Medline: 31549787]
4. Zhang L, Wang Z, Wang X, Chen Z, Shao L, Tian Y, China Hypertension Survey investigators. Prevalence of overweight and obesity in China: results from a cross-sectional study of 441 thousand adults, 2012-2015. Obes Res Clin Pract 2020;14(2):119-126. [doi: 10.1016/j.orcp.2020.02.005] [Medline: 32139330]
5. HEARTS-D: diagnosis and management of type 2 diabetes. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/331710 [accessed 2021-03-29]

6. Wildman R, Muntner P, Reynolds K, McGinn A, Rajpathak S, Wylie-Rosett J, et al. The obese without cardiometabolic risk factor clustering and the normal weight with cardiometabolic risk factor clustering: prevalence and correlates of 2 phenotypes among the US population (NHANES 1999-2004). Arch Intern Med 2008 Aug 11;168(15):1617-1624. [doi: 10.1001/archinte.168.15.1617] [Medline: 18695075]
7. Kim Y, Kim S, Han K, Jung J, Lee S, Oh S, et al. Waist circumference and all-cause mortality independent of body mass index in Korean population from the National Health Insurance Health Checkup 2009 2015. J Clin Med 2019 Jan 10;8(1):72 [FREE Full text] [doi: 10.3390/jcm8010072] [Medline: 30634601]
8. Nazare JA, Smith J, Borel AL, Aschner P, Barter P, Van Gaal L, INSPIRE ME IAA Investigators. Usefulness of measuring both body mass index and waist circumference for the estimation of visceral adiposity and related cardiometabolic risk profile (from the INSPIRE ME IAA Study). Am J Cardiol Internet 2015;115(3):307-315. [doi: 10.1016/j.amjcard.2015.02.045] [Medline: 25499404]
9. Rosenthal AD, Jin F, Shu X, Yang G, Elasy TA, Chow W, et al. Body fat distribution and risk of diabetes among Chinese women. Int J Obes Relat Metab Disord 2004 Apr 10;28(4):594-599. [doi: 10.1038/sj.ijo.0802597] [Medline: 14770196]
10. Hou X, Chen S, Hu G, Chen P, Wu J, Ma X, China National Diabetes, Metabolic Disorders Study Group. Stronger associations of waist circumference and waist-to-height ratio with diabetes than BMI in Chinese adults. Diabetes Res Clin Pract 2019 Jan;147:9-18. [doi: 10.1016/j.diabres.2018.07.029] [Medline: 30144478]
11. Wang Y, Rimm E, Stampfer M, Willett W, Hu F. Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. Am J Clin Nutr 2005 Mar;81(3):555-563 [FREE Full text] [doi: 10.1093/ajcn/81.3.555] [Medline: 15755822]
12. Fan Y, Wang R, Ding L, Meng Z, Zhang Q, Shen Y, et al. Waist circumference and its changes are more strongly associated with the risk of type 2 diabetes than body mass index and changes in body weight in Chinese adults. J Nutr 2020 May 01;150(5):1259-1265. [doi: 10.1093/jn/nxaa014] [Medline: 32006008]
13. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, UCLEB Consortium; METASTROKE Consortium, et al. Causal associations of adiposity and body fat distribution with coronary heart disease, stroke subtypes, and type 2 diabetes mellitus: a Mendelian randomization analysis. Circulation 2017 Jun 13;135(24):2373-2388 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.116.026560] [Medline: 28500271]
14. Li S, Zhang L, Liu S, Hubbard R, Li H. Surveillance of noncommunicable disease epidemic through the integrated noncommunicable disease collaborative management system: feasibility pilot study conducted in the city of Ningbo, China. J Med Internet Res 2020 Jul 23;22(7):e17340 [FREE Full text] [doi: 10.2196/17340] [Medline: 32706706]
15. Chinese Diabetes Society. Guidelines for the prevention and control of type 2 diabetes in China (2017 Edition). Chinese J Pract Intern Med 2018;4:344. [doi: 10.19538/j.nk2018040108]
16. World Health Organization. Waist circumference and waist-hip ratio: Report of a WHO Expert Consultation. Geneva: World Health Organization; 2011.
17. People's Republic of China National Health and Family Planning Commission. WS/T 428-2013 Criteria of weight for adults. Chinese Standard. 2013. URL: https://www.chinesestandard.net/PDF/English.aspx/WST428-2013 [accessed 2021-03-29]
18. Fan M, Lyu J, He P. Chinese guidelines for data processing and analysis concerning the International Physical Activity Questionnaire. Zhonghua Liu Xing Bing Xue Za Zhi 2014 Aug;35(8):961-964. [Medline: 25376692]
19. Chinese Hypertension League. Guidelines for the prevention and control of hypertension in China (2018 revised Edition). Chinese J Cardiovasc Med 2019;24(1):56.
20. Wang H, Hu R, Qian Y, Wang C, Xie K. Prospective study on the effect of BMI and waist circumference on diabetes of adults in Zhejiang province. Chin J Epidemiol 2018;39(6):5. [doi: 10.3760/cma.j.issn.0254-6450.2018.06.022]
21. Yang X, Zhang M, Luo X, Jinjin W, Yin L. Body mass index, waist circumference and waist-to-height ratio associated with the incidence of type 2 diabetes mellitus: a cohort study. Chin J Epidemiol 2016;50(4):33. [doi: 10.3760/cma.j.issn.0253-9624.2016.04.009]
22. Yang J, Wang F, Wang J, Han X, Hu H, Yu C, et al. Using different anthropometric indices to assess prediction ability of type 2 diabetes in elderly population: a 5 year prospective study. BMC Geriatr 2018 Sep 17;18(1):218-215 [FREE Full text] [doi: 10.1186/s12877-018-0912-2] [Medline: 30223783]
23. Kodama S, Horikawa C, Fujihara K, Heianza Y, Hirasawa R, Yachi Y, et al. Comparisons of the strength of associations with future type 2 diabetes risk among anthropometric obesity indicators, including waist-to-height ratio: a meta-analysis. Am J Epidemiol 2012 Dec 01;176(11):959-969 [FREE Full text] [doi: 10.1093/aje/kws172] [Medline: 23144362]
24. Lee CMY, Woodward M, Pandeya N, Adams R, Barrett-Connor E, Boyko EJ, Obesity, DiabetesCardiovascular Disease Collaboration. Comparison of relationships between four common anthropometric measures and incident diabetes. Diabetes Res Clin Pract 2017 Oct;132:36-44 [FREE Full text] [doi: 10.1016/j.diabres.2017.07.022] [Medline: 28783531]
25. Taylor A, Ebrahim S, Ben-Shlomo Y, Martin R, Whincup P, Yarnell J, et al. Comparison of the associations of body mass index and measures of central adiposity and fat mass with coronary heart disease, diabetes, and all-cause mortality: a study using data from 4 UK cohorts. Am J Clin Nutr 2010 Mar;91(3):547-556 [FREE Full text] [doi: 10.3945/ajcn.2009.28757] [Medline: 20089729]

26. Jia Z, Zhou Y, Liu X, Wang Y, Zhao X, Wang Y, et al. Comparison of different anthropometric measures as predictors of diabetes incidence in a Chinese population. Diabetes Res Clin Pract 2011 May;92(2):265-271. [doi: 10.1016/j.diabres.2011.01.021] [Medline: 21334088]
27. Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? Am J Epidemiol 1996 Feb 01;143(3):228-239 [FREE Full text] [doi: 10.1093/oxfordjournals.aje.a008733] [Medline: 8561156]
28. Kuk JL, Saunders TJ, Davidson LE, Ross R. Age-related changes in total and regional fat distribution. Ageing Res Rev 2009 Oct;8(4):339-348. [doi: 10.1016/j.arr.2009.06.001] [Medline: 19576300]
29. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011 Apr 21;32(1):91-108. [doi: 10.1146/annurev-publhealth-031210-100700] [Medline: 21219160]
30. González-Villalpando C, Dávila-Cervantes CA, Zamora-Macorra M, Trejo-Valdivia B, González-Villalpando ME. Incidence of type 2 diabetes in Mexico: results of the Mexico City Diabetes Study after 18 years of follow-up. Salud Publica Mex 2014;56(1):11-17. [doi: 10.21149/spm.v56i1.7318] [Medline: 24912516]
31. Man RE, Charumathi S, Gan ATL, Fenwick EK, Tey CS, Chua J, et al. Cumulative incidence and risk factors of prediabetes and type 2 diabetes in a Singaporean Malay cohort. Diabetes Res Clin Pract 2017 May;127:163-171. [doi: 10.1016/j.diabres.2017.03.007] [Medline: 28371687]
32. Anjana R, Shanthi Rani CS, Deepa M, Pradeepa R, Sudha V, Divya Nair H, et al. Incidence of diabetes and prediabetes and predictors of progression among Asian Indians: 10-year follow-up of the Chennai Urban Rural Epidemiology Study (CURES). Diabetes Care 2015 Aug;38(8):1441-1448. [doi: 10.2337/dc14-2814] [Medline: 25906786]
33. Hyler SE, Skodol AE, Oldham JM, Kellman HD, Doidge N. Validity of the Personality Diagnostic Questionnaire-Revised: a replication in an outpatient sample. Compr Psychiatry 1992;33(2):73-77. [doi: 10.1016/0010-440x(92)90001-7] [Medline: 1544299]
34. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
35. Schwenzer NF, Machann J, Schraml C, Springer F, Ludescher B, Stefan N, et al. Quantitative analysis of adipose tissue in single transverse slices for estimation of volumes of relevant fat tissue compartments: a study in a large cohort of subjects at risk for type 2 diabetes by MRI with comparison to anthropometric data. Invest Radiol 2010 Dec;45(12):788-794. [doi: 10.1097/RLI.0b013e3181f10fe1] [Medline: 20829704]

Abbreviations

FPG: fasting plasma glucose
HbA_{1c}: glycated hemoglobin
HDL-C: high-density lipoprotein cholesterol
HR: hazard ratio
ID: identification
LDL-C: low-density lipoprotein cholesterol
MET: metabolic equivalent
NCDCMS: Ningbo Municipal Integrated Noncommunicable Disease Collaborative Management System
T2DM: type 2 diabetes mellitus
TC: total cholesterol
TG: triglycerides
WC: waist circumference
WHR: waist-to-hip ratio

Edited by Y Khader; submitted 11.05.22; peer-reviewed by R Kelishadi, J Vilchis; comments to author 22.09.22; revised version received 19.11.22; accepted 26.11.22; published 11.01.23.

Please cite as:
Li S, Wang Y, Ying Y, Gong Q, Lou G, Liu Y, Liu S, Li H
Independent and Joint Associations of BMI and Waist Circumference With the Onset of Type 2 Diabetes Mellitus in Chinese Adults: Prospective Data Linkage Study
2023;9:e39459
URL: https://publichealth.jmir.org/2023/1/e39459
doi:10.2196/39459
PMID:36630180

Original Paper

Gastroenteritis Forecasting Assessing the Use of Web and Electronic Health Record Data With a Linear and a Nonlinear Approach: Comparison Study

Canelle Poirier[1,2,3,4,5], PhD; Guillaume Bouzillé[3,4,5], MD, PhD; Valérie Bertaud[3,4,5], MD, PhD; Marc Cuggia[3,4,5], MD, PhD; Mauricio Santillana[1,2,6,7*], MSc, PhD; Audrey Lavenu[8,9,10*], MD, PhD

[1]Department of Pediatrics, Harvard Medical School, Boston, MA, United States
[2]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, United States
[3]Institut national de la santé et de la recherche médicale U1099, Rennes, France
[4]Laboratoire Traitement du Signal et de l'Image, Université de Rennes 1, Rennes, France
[5]Centre de Données Cliniques, Centre Hospitalier Universitaire Rennes, Rennes, France
[6]Harvard Tseng-Hsi Chan School of Public Health, Boston, MA, United States
[7]Machine Intelligence Group for the Betterment of Health and the Environment, Network Science Institute, Northeastern University, Boston, MA, United States
[8]Faculté de médecine, Université de Rennes 1, Rennes, France
[9]Institut de Recherche Mathématique de Rennes, Rennes, France
[10]Institut national de la santé et de la recherche médicale CIC 1414, Université de Rennes 1, Rennes, France
[*]these authors contributed equally

Corresponding Author:
Canelle Poirier, PhD
Computational Health Informatics Program
Boston Children's Hospital
300 Longwood Avenue
Boston, MA, 02115
United States
Phone: 1 617 355 6000
Email: canelle.poirier@outlook.fr

Abstract

Background: Disease surveillance systems capable of producing accurate real-time and short-term forecasts can help public health officials design timely public health interventions to mitigate the effects of disease outbreaks in affected populations. In France, existing clinic-based disease surveillance systems produce gastroenteritis activity information that lags real time by 1 to 3 weeks. This temporal data gap prevents public health officials from having a timely epidemiological characterization of this disease at any point in time and thus leads to the design of interventions that do not take into consideration the most recent changes in dynamics.

Objective: The goal of this study was to evaluate the feasibility of using internet search query trends and electronic health records to predict acute gastroenteritis (AG) incidence rates in near real time, at the national and regional scales, and for long-term forecasts (up to 10 weeks).

Methods: We present 2 different approaches (linear and nonlinear) that produce real-time estimates, short-term forecasts, and long-term forecasts of AG activity at 2 different spatial scales in France (national and regional). Both approaches leverage disparate data sources that include disease-related internet search activity, electronic health record data, and historical disease activity.

Results: Our results suggest that all data sources contribute to improving gastroenteritis surveillance for long-term forecasts with the prominent predictive power of historical data owing to the strong seasonal dynamics of this disease.

Conclusions: The methods we developed could help reduce the impact of the AG peak by making it possible to anticipate increased activity by up to 10 weeks.

(JMIR Public Health Surveill 2023;9:e34982) doi:10.2196/34982

KEYWORDS

infectious disease; acute gastroenteritis; modeling; modeling disease outbreaks; machine learning; public health; machine learning in public health; forecasting; digital data

Introduction

Background

Acute gastroenteritis (AG) is a major public health problem worldwide [1]. Commonly defined as diarrhea or vomiting in the past 24 hours [2], AG is one of the main causes of morbidity and mortality among young people and causes up to 2.5 million deaths per year in children aged <5 years around the world [3]. Although it is generally a mild disease, its morbidity and economic burden are high [4]. In France, there are >21 million episodes of AG each year [5]. Although AG episodes occur throughout the year, there is a winter peak, mainly owing to norovirus and rotavirus [6,7]. During these peaks, the increase of visits to general practitioners and emergency or pediatric departments causes health care system disruptions [8].

Disease surveillance systems capable of producing accurate real-time and short-term forecasts can help public health officials design timely public health interventions to mitigate the effects of disease outbreaks in affected populations. In France, all acute diarrhea cases seen during medical appointments are reported weekly by volunteer outpatient health care providers. An estimation of AG incidence rate is then computed, at the national or regional scale, by considering the number of sentinel physicians and the medical density of the area of interest [9]. However, data collection, processing, aggregation, and distribution processes introduce up to 3 weeks of delay in the availability of AG activity information. This temporal data gap prevents public health officials from having a timely perspective about AG activity and thus leads to the design of interventions that do not take into consideration the most recent changes in disease dynamics. Therefore, there is a growing interest in finding new ways to mitigate this information gap [10,11].

To alleviate this time lag, several studies have proposed approaches to produce accurate and reliable real-time disease activity estimates, for example, to monitor influenza [11-14]. For AG, studies have been focused on identifying the clinical characteristics of the disease. Norovirus and rotavirus are the viruses responsible for most gastroenteritis outbreaks [6,7,15-18]. This disease has a strong wintertime seasonality, but this seasonality could be affected by the climate change, which would affect norovirus transmission, host's susceptibility to norovirus infection, and resistance of norovirus to environmental conditions. This may cause large oscillations in the number of cases per year [6,7]. AG remains as a major cause of hospitalizations, especially for children, and the use of a vaccine could help to decrease the impact of the disease [16,18]. Some research teams have assessed the correlation between data sources (eg, drug reimbursement data and emergency department visits) and general practitioner visits for AG [3,19]. Other studies have shown a significant correlation between internet search query trends and AG incidence rates in different locations such as the United States, Mexico, the United Kingdom, and France [20,21]. However, none, to the best of our knowledge [22], have proposed a feasible methodology to forecast AG activity. Through this study, we investigated the challenges of achieving this and proposed a reliable forecasting approach.

State of the Art

Existing forecasting systems for other disease outbreaks, such as influenza, include statistical models that leverage information available in near real time [11-14]. One of the first and prominent studies is Google Flu Trends [23], a web-based service operated by Google. Created in 2009, the platform used the volume of selected Google search terms to estimate influenza activity in real time. However, the web service was stopped following several prediction errors owing to changes in people's search behavior as a result of the exceptional nature of the pandemic or owing to the announcement of a pandemic that finally did not appear [24]. Following this, some authors updated the Google Flu Trends algorithm to improve influenza forecasting, by including data from Google Correlate and Google Trends web services and other sources, for instance, historical influenza information [11]. Internet is not the only data source that can be used to produce information in real time. With the widespread adoption of patient electronic health records (EHRs), hospitals also generate a huge amount of data. Bouzillé et al [25] showed that EHRs are strongly correlated with influenza incidence rates. Some authors proposed statistical models using EHRs to predict influenza incidence rates in real time [12,26]. In addition, other studies showed that internet users' searches were strongly correlated with influenza epidemics and other diseases, including AG [8,21].

In this study, we evaluated the feasibility of using internet search query trends and EHR to predict AG incidence rates in near real time, at the national and regional scales, and for long-term forecasts (up to 10 weeks). We used 2 different methods—a linear approach using Elastic Net and a nonlinear approach using random forest (RF). In addition, as AG outbreaks cause disruptions in hospitals and emergency departments, we estimated AG incidence rates at the level of emergency departments and hospital stays.

Methods

Variables to Be Predicted

National Level

We obtained the national (Metropolitan France) acute diarrhea weekly incidence rates (per 100,000 inhabitants) from the French Sentinel network [27], from January 2008 to March 2018. We retrieved these data in April 2018.

Regional Level

We obtained the regional (Brittany region) acute diarrhea incidence rates (per 100,000 inhabitants) from the French Sentinel network [27], from January 2008 to March 2018. We chose the Brittany region as we used her data from a hospital in Brittany. We retrieved these data in April 2018.

Predictive Variables

Web Data

We obtained the frequency per week of the 100 most correlated French queries from Google Correlate [28]. For each signal to be predicted (national and regional levels), we retrieved Google Correlate data for the period from January 2008 to March 2018. As our prediction period is from May 2014 to February 2018, the correlation was calculated from January 2008 to April 2014. All signals were normalized to obtain mean 0 and SD 1 before calculating the correlation. The reason to correlate was to choose the most appropriate queries to predict the outbreak without previous knowledge [29]. The most correlated queries obtained for national and regional levels can differ because the weekly incidence rates for France and Brittany are different.

Clinical Data

We used data from the clinical data warehouse (CDW) of Rennes University Hospital (France), called entrepôt de données de l'HÔPital (eHOP). This CDW includes structured (laboratory test results, prescriptions, and International Statistical Classification of Diseases and Related Health Problems 10th Revision diagnoses) and unstructured (discharge letter, pathology reports, and operative reports) patients' data from 1.2 million inpatients and outpatients and 45 million documents. To identify patients with specific criteria, eHOP has its own search engine system that allows to query unstructured data with keywords or structured data with codes based on terminologies.

First, to retrieve clinical data connected with AG, we performed different full-text queries (related to gastroenteritis, its symptoms, virus, or treatments). These queries allowed to obtain all documents matching with the search criteria (often, several documents for 1 patient and 1 stay). Then, for each week, we kept the oldest document for 1 patient and 1 hospital stay, and we calculated the number of hospital stays with at least one document mentioning the keyword contained in the query. As we used 19 keywords, we obtained 19 variables from CDW eHOP.

Then, we built a database containing the time series constructed from the structured data (total n=1,335,347 time series). Regrading Google Correlate, we calculated the Pearson correlation between both national and regional incidence rates and the time series from the database. We retrieved the 100 most correlated signals. As our prediction period is from May 2014 to February 2018, we calculated the correlation between January 2008 and April 2014.

Overall, we obtained 119 variables (n=19, 15.9% of variables from the full-text queries and n=100, 84% of the most correlated variables from the structured data). The 100 most correlated variables can be different for national and regional levels. We retrieved EHR data for the period from January 2008 to March 2018 in April 2018. All these data could be extracted in real time if needed.

Historical Data

We used the incidence rates for the previous 52 weeks as predictive variables, for both national and regional levels.

Ethics Approval

This study was approved by the local ethics committee of the Rennes Academic Hospital (approval number 16.69).

Statistical Models

Linear Approach

To minimize the negative effects of using a large number of input variables, potentially including redundant information, we used Elastic Net, a regularized multivariate regression methodology that can identify parsimonious models [30]. Elastic Net combines the power of Lasso and Ridge regressions, allowing to perform a variable selection on variables that are highly correlated [31,32]. We performed the Elastic Net regression analysis using the *caret* package in R (R Foundation for Statistical Computing) and the associated function fit with the *glmnet* method [33,34]. We fixed a coefficient λ=0.5 to give the same importance to Ridge and Lasso methods.

The formulation of our model is the following:

Here, yT denotes AG incidence rate at time T=t, t+1, t+2, t+3 (for the different levels of prediction), denotes historical variables, denotes Google data, denotes EHR data, and denotes residuals.

For a given week, we needed to find the parameters, $\alpha=(\alpha_1,..\alpha_{52})$, $\beta=(\beta_1,..\beta_{100})$, and $\gamma=(\gamma_1,..\gamma_{119})$, that minimize the following:

Here, are hyperparameters of the Elastic Net regression. We used 10-block cross-validation to optimize the parameters. All parameters ($\alpha=[\alpha_1,..\alpha_{52}]$, $\beta=[\beta_1,..\beta_{100}]$, and $\gamma=[\gamma_1,..\gamma_{119}]$) were dynamically trained every week with a rolling window using all data available. In this way, the size of our training data set increased every week. For example, for the first week of January 2015, our training data set ranged from January 2008 to the last week of December 2014. To predict the first week of January 2016, our training data set ranged from January 2008 to the last week of December 2015. We obtained estimates from May 2014 to February 2018.

Nonlinear Approach

RF is a nonlinear machine learning approach based on the construction of multiple decision trees using the general bootstrap aggregating technique (known as bagging) [35]. We used this method as it showed good performance in short-term forecasting even when it is compared with other machine learning approaches such as support vector machine or neural network or a traditional approach such as autoregressive integrated moving average [36,37].

With RF, the AG incidence rates are obtained with the following:

Here, y_T denotes AG incidence rate at time T=t, t+1, t+2, t+3 (for the different levels of prediction) and denotes AG incidence rates estimate obtained with the decision tree b. We used the R package, *randomForest* [38], to create our RF models. The hyperparameters corresponding to the number of decision trees and the number of variables randomly sampled at each split were optimized on a training data set from January 2008 to May 2014. Then, regarding the Elastic Net model, RF was dynamically recalibrated for every new week of prediction by incorporating all the data available. We obtained estimates from May 2014 to February 2018.

Contribution of Each Data Source

In addition, to assess the contribution of each individual data sources or their combinations, we built Elastic Net and RF models using the following predictive variables:

1. AG incidence rates—baseline model called autoregressive model of order 52 (AR(52)) in the following sections—for the previous 52 weeks
2. Google data
3. EHR data
4. Google data and AR(52)
5. EHR data and AR(52)
6. Google data and EHR data

Evaluation

To assess the performance of our models, we compared our estimates with the real incidence rates from the Sentinel network. We calculated the root mean squared error and the Pearson correlation coefficient for our test period starting from May 2014 to February 2018. The model allowing to obtain the most accurate estimates is the one having the highest correlation and the lowest error:

1.
2.

Here, is the predicted value for the week t, is the mean of predicted values, y_t is the real value for the week t, and is the mean of real values.

Comparison With Influenza

As we used a method developed for influenza outbreaks, we compared the results obtained for AG with those obtained for influenza. The aim was to determine whether external data sources are as relevant for AG as for influenza. We started by comparing the stationarity and the seasonality of both time series by calculating the following:

1. The autocorrelation function (ACF), allowing to determine the autocorrelation between y_t and $y_{t–h}$:

where $\gamma(h)=\mathrm{cov}(y_t, y_{t–h})$

2. The partial ACF (PACF), allowing to determine the autocorrelation between y_t and $y_{t–h}$ after removing the autocorrelation between the intermediate variables $y_{t–1},\dots,y_{t–h+1}$:

$r(h)=\mathrm{corr}(y_t,y_{t–h}|y_{t–1},\dots,y_{t–h+1})$

Then, we compared the accuracy of estimates for forecast up to 10 weeks with Elastic Net and RF models using only historical data or combining Google, EHR, and historical data.

Results

Overview

First, we studied the impact of each data source for short-term forecasts with the 2 different approaches already used to predict influenza outbreaks—a linear approach with the Elastic Net model and a nonlinear approach with an RF model.

Then, we analyzed the AG and influenza time series, especially the seasonality, to better understand the differences between the 2 diseases.

Finally, we compared AG and influenza results obtained for long-term forecasts with the 2 approaches, and we assessed the impact of external data sources to increase the accuracy of our estimates.

Linear Approach

Overview

At the national and regional levels, in terms of error, the lowest values are obtained with models using historical data and external data sources (Table 1). At the national level, in terms of error, both data sources, Google and EHR produce the most accurate estimates compared with the model using only historical data—AR (52). At the regional level, the model using only historical data and EHR allows to obtain lower errors than the model using historical data and both Google and EHR data.

In terms of correlation, in most cases, at the national and regional levels, the model using only historical data allows to obtain the highest values.

Table 1. PCC[a] and RMSE[b] values obtained for the entire prediction period (May 2014 to March 2018) at the national and regional levels, with all the combinations of data sources.

Levels and data sources	Real time		1-week forecast		2-week forecast		3-week forecast	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
National								
AR(52)[c]	*0.946* [d]	*16.16*	*0.910*	22.69	*0.898*	26.95	*0.884*	30.69
Google	0.830	42.75	0.803	44.99	0.801	41.27	0.770	38.96
EHR[e]	0.477	48.35	0.512	45.59	0.489	47.37	0.519	44.65
AR(52) and Google	*0.941*	18.10	0.896	24.17	0.871	26.98	0.847	28.24
AR(52) and EHR	0.932	*16.41*	0.880	21.58	0.820	26.15	0.823	*25.93*
Google and EHR	0.836	36.09	0.846	34.48	0.779	34.23	0.795	32.32
AR(52), Google, and EHR	0.936	21.26	0.903	*20.94*	0.856	*24.16*	0.845	*25.33*
Regional								
AR(52)	0.725	*40.75*	*0.705*	44.18	*0.670*	47.65	*0.681*	49.12
Google	0.652	65.84	0.603	64.79	0.594	60.33	0.596	61.67
EHR	0.462	59.83	0.538	55.62	0.546	55.87	0.582	52.90
AR(52) and Google	*0.738*	42.07	0.665	46.44	0.616	47.82	0.619	47.74
AR(52) and EHR	0.697	*40.99*	0.685	*42.38*	0.637	*46.48*	0.634	*46.31*
Google and EHR	0.608	60.70	0.610	60.97	0.615	57.50	0.628	59.72
AR(52), Google, and EHR	0.724	42.12	0.689	45.24	0.646	47.37	0.620	52.19

[a] PCC: Pearson correlation coefficient.

[b] RMSE: root mean squared error.

[c] AR(52): autoregressive model of order 52.

[d] Italicization highlights the 2 highest correlations and lowest errors obtained with the models for real time and 1-week, 2-week, and 3-week forecasts.

[e] EHR: electronic health record.

National Analysis

For real-time estimates, the error values range from 48.4 to 16.2 and the correlation values range from 0.83 to 0.95, with the lowest error and the highest correlation obtained with the model using only historical data—AR(52). For 1-week estimates, the error values range from 45.6 to 20, with the lowest error and the highest correlation obtained with the model using historical data and both external data sources, Google and EHR. In terms of correlation, the correlation values range from 0.51 to 0.91, with the highest value obtained with the model using only historical data. For 2-week and 3-week estimates, we have similar results, with error values ranging from 47.4 to 24.2 and 44.6 to 25.3, respectively, obtained with the model using historical data and both external data sources, Google and EHR. In terms of correlation, the values range from 0.49 to 0.90 and from 0.52 to 0.88, respectively, with the highest correlation obtained with AR(52) model.

Figure 1 illustrates the estimates obtained at the national level for forecasts up to 3 weeks with the model using only historical data and the model using historical data and both data sources, Google and EHR. For real-time estimates, the results obtained with the 2 models are comparable, but for long-term forecasts (1, 2, and 3 weeks), the estimates obtained with the AR(52) model are delayed. In addition, the model using only historical data tends to smooth estimates and overestimate between peaks.

Figure 2 is a visualization of the values of the coefficients for the model using historical data and both data sources, Google and EHR. For real-time estimates, the heat map shows that the model uses multiple variables from all data sources, such as historical data, Google data, and EHR data. Similar plots are presented in Multimedia Appendix 1 for long-term estimates.

Figure 1. National level. Predictions up to 3 weeks obtained at the national level with the model using only historical data and the model using historical data and both data sources, Google and EHR. Gold standard, French Sentinel network data. EHR: electronic health record.

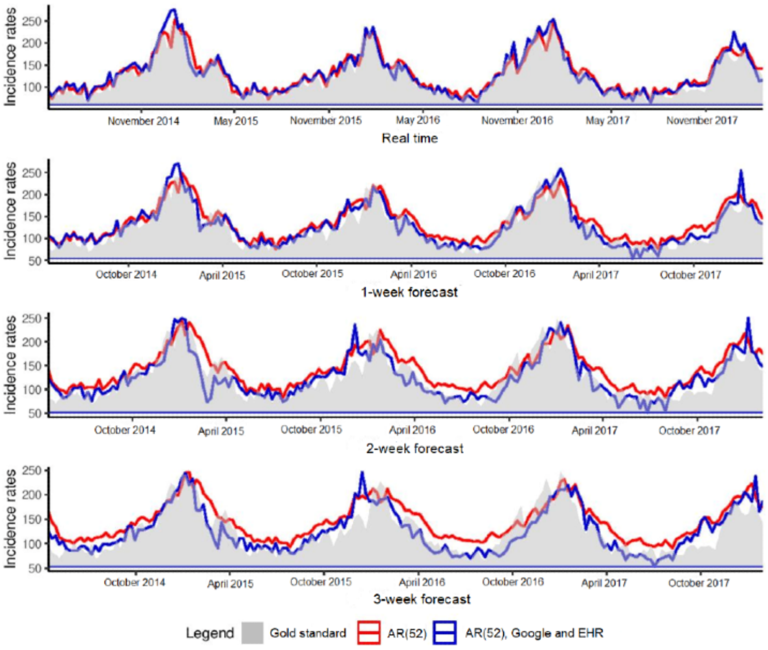
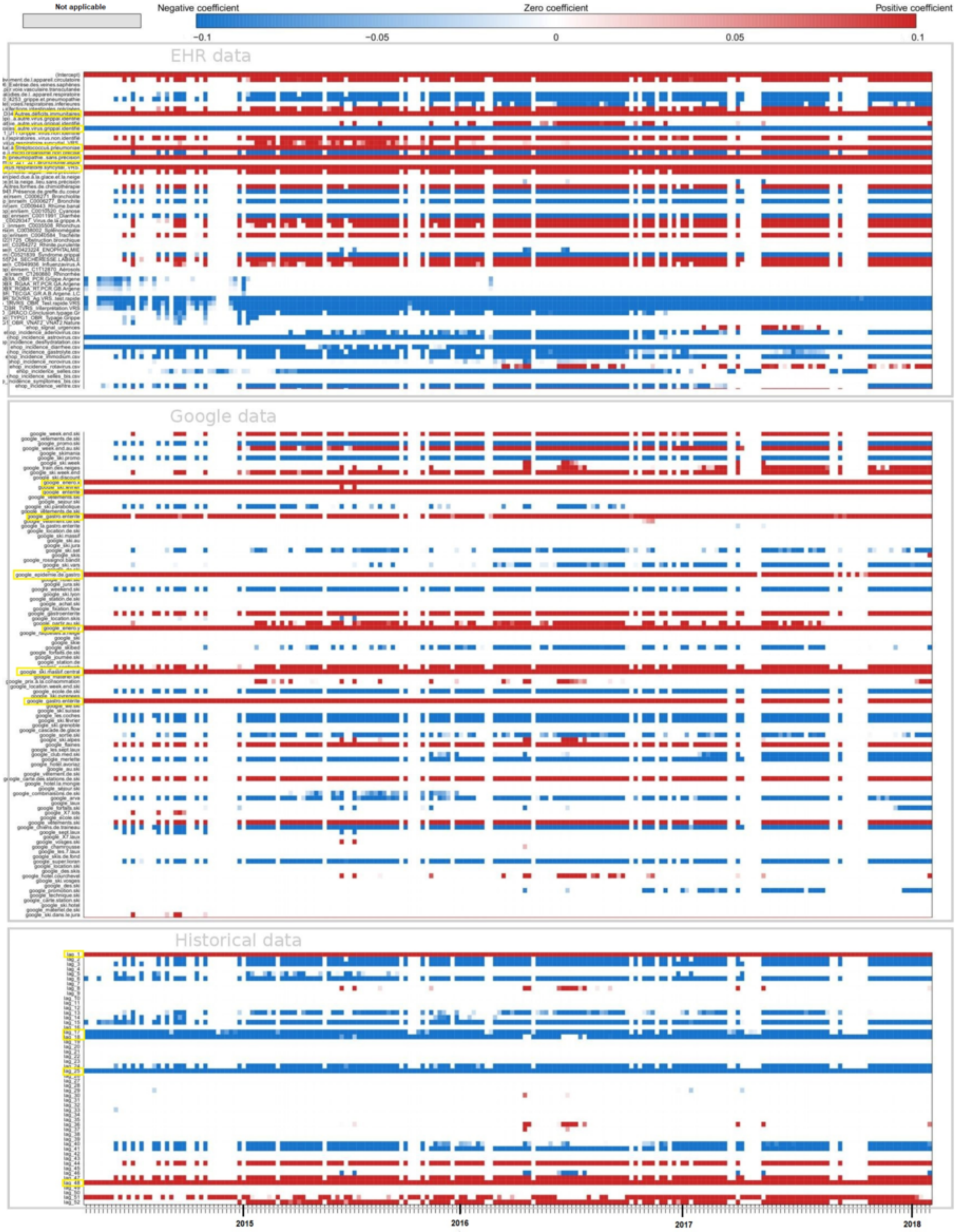


Figure 2. National level. Heatmap of the coefficients. Each line of the heatmap corresponds to one predictive variable used in the model and each point of the line corresponds to 1 week predicted. The first block of variables corresponds to electronic health record (EHR) data, the second one corresponds to Google data, and the third one to historical data. In blue, a negative coefficient is associated with the variable, whereas in red, it is a positive coefficient. The white color means that the predictive variable is not selected by the model and does not participate in forecasting the corresponding week. In yellow, highlighted variables that are kept by the model almost all the time. For EHR data, it corresponds to the predictive variables for the keywords “Autres deficits immunitaires,” “Autre virus grippal identifié,” “Streptococcus pneumoniae,” “Pneumopathie,” “Virus respiratoire syncytial.” For Google data, it is the keywords: “enero,” “enterite,” “epidemie de gastro,” “gastro entérite,” “ski massif central.” For historical data, it corresponds to the previous week as well as week 17, week 18, week 25, and week 48 before the one we want to predict.



Regional Analysis

For real-time estimates, the error values range from 65.8 to 40.8 and the correlation values range from 0.46 to 0.74, with the lowest value for the error obtained with the model using only historical data and the highest value for the correlation obtained with the model using historical data and Google data. For 1-week, 2-week, and 3-week estimates, the error values range from 64.8 to 42.4, from 60.3 to 46.5, and from 61.7 to 46.3, respectively. The lowest errors values for long-term forecasts are all obtained with the model using historical data and EHR data. In terms of 1-week, 2-week, and 3-week correlation, the values range from 0.54 to 0.71, from 0.55 to 0.67, and from 0.58 to 0.68, respectively. The highest correlations for long-term forecasts are all obtained with the model using only historical data—AR(52).

Figure 3 illustrates the estimates obtained at the regional level for forecasts up to 3 weeks with the model using only historical data and the model using historical data and both data sources, Google and EHR. At the national level, for real-time estimates, the results obtained with the 2 models are comparable, but for long-term forecasts, the estimates obtained with the AR(52) model are delayed and tend to be smoothed and overestimated between peaks.

The heat map (Figure 4) shows that for real-time estimates at the regional level, the model uses multiple variables from historical data (approximately 11 variables) and low number of variables from Google data (approximately 10 variables) and EHR data (approximately 9 variables) compared with those at the national level. Similar plots are presented in Multimedia Appendix 1 for long-term estimates.

Figure 3. Regional level. Predictions up to 3 weeks obtained at the regional level with the model using only historical data and the model using historical data and both data sources, Google and EHR. Gold standard, French Sentinel network data. EHR: electronic health record.

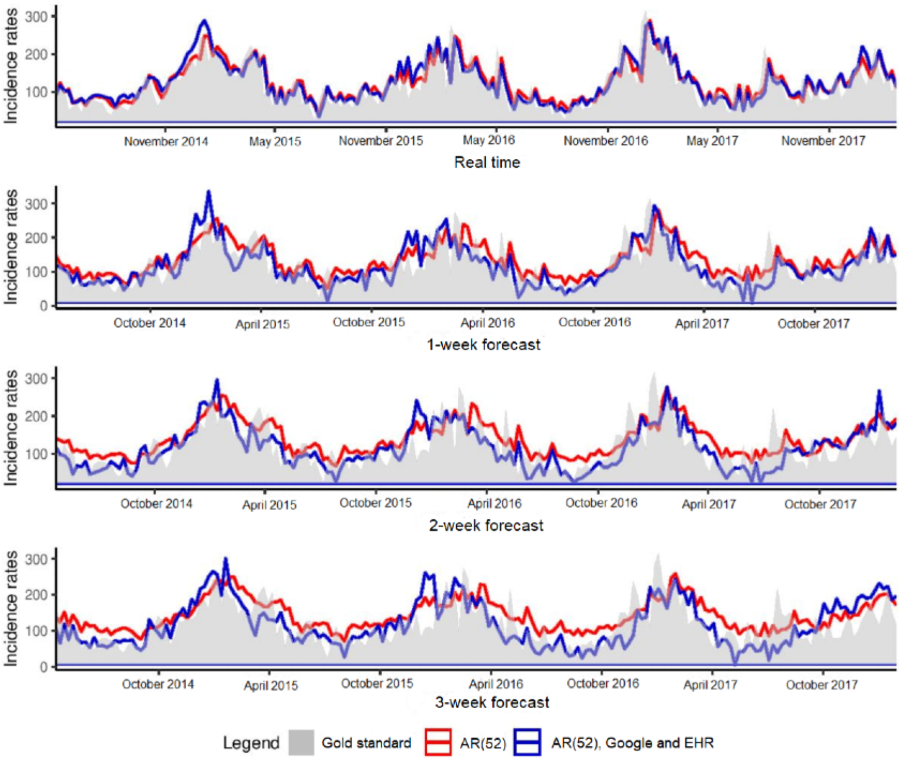
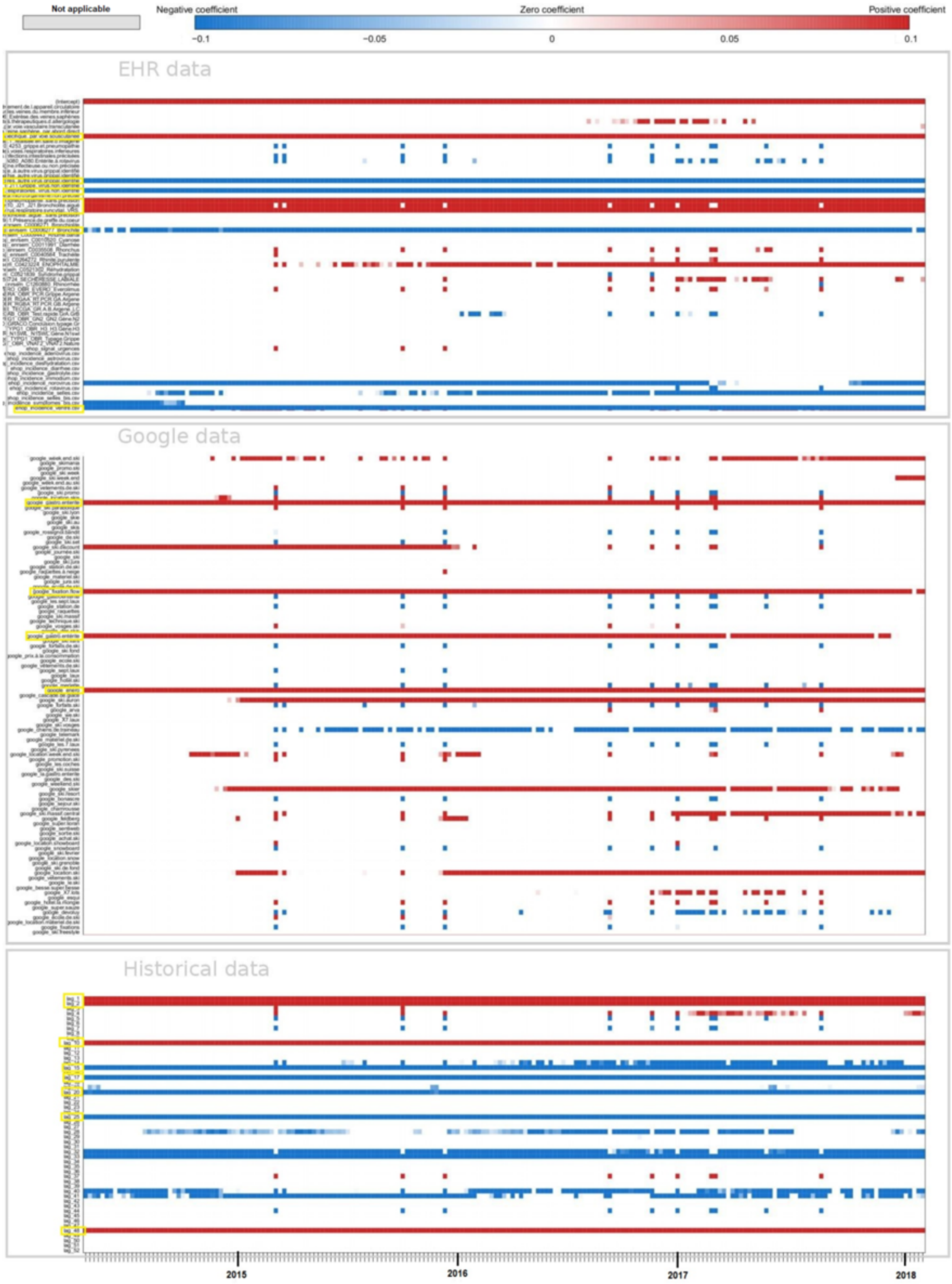


Figure 4. Regional level. Heatmap of the coefficients. Each line of the heatmap corresponds to one predictive variable used in the model and each point of the line corresponds to 1 week predicted. The first block of variables corresponds to electronic health record (EHR) data, the second one corresponds to Google data, and the third one to historical data. In blue, a negative coefficient is associated with the variable, whereas in red, it is a positive coefficient. The white color means that the predictive variable is not selected by the model and does not participate in forecasting the corresponding week. In yellow, highlighted variables that are kept by the model almost all the time. For EHR data it corresponds to the predictive variables for the keywords “Par voie sous cutannée,” “Autre virus grippal identifié,” “Voies respiratoires. Virus non identifié,” “Pneumopathie,” “Bronchiolite aigüe,” “Virus respiratoire syncytial,” “Bronchite,” “Ventre.” For Google data, it is the keywords: “enero,” “gastro enterite,” “gastro entérite,” “fixations.” For historical data, it corresponds to the two previous weeks as well as week 10, week 15, week 17, week 20, week 25, and week 48 before the one we want to predict.



Nonlinear Approach

Overview

For the nonlinear approach, at the national level, in terms of error and correlation, results are comparable between the model using only historical data—AR(52)—and the models combining historical data and external data sources (Table 2). At the regional level, in terms of error, the lowest errors are mostly obtained with the model including historical and EHR data. In terms of correlation, the highest values are mostly obtained with the model combining historical data and both data sources, Google and EHR. For the nonlinear approach, the values for correlation are higher and the values for errors are lower than the values obtained with the linear approach.

Table 2. PCC[a] and RMSE[b] values obtained for the entire prediction period (May 2014 to March 2018) for all levels and models.

Levels and data sources	Real time		1-week forecast		2-week forecast		3-week forecast	
	PCC	RMSE	PCC	RMSE	PCC	RMSE	PCC	RMSE
National								
AR(52)[c]	*0.942* [d]	*15.47*	*0.913*	*19.71*	*0.892*	*22.19*	*0.903*	*22.30*
Google	0.884	45.59	0.876	45.72	0.858	42.63	0.830	40.52
EHR[e]	0.795	32.93	0.615	50.68	0.739	37.84	0.692	41.30
AR(52) and Google	*0.946*	*15.87*	*0.913*	21.68	*0.892*	23.63	*0.909*	*22.98*
AR(52) and EHR	0.938	*15.93*	0.906	*20.21*	0.887	*22.85*	0.890	23.31
Google and EHR	0.833	43.26	0.780	49.50	0.849	37.70	0.790	41.88
AR(52), Google, and EHR	*0.946*	*15.72*	0.909	21.76	*0.895*	23.87	0.886	24.11
Regional								
AR(52)	0.745	*38.47*	0.699	42.68	0.685	*44.11*	0.677	45.05
Google	0.708	62.90	0.658	61.58	0.671	57.02	0.689	54.55
EHR	0.651	47.76	0.531	66.99	0.562	60.51	0.526	63.26
AR(52) and Google	0.757	39.71	0.700	46.91	0.694	47.38	*0.703*	47.87
AR(52) and EHR	0.743	*38.37*	*0.720*	*41.05*	0.694	*43.83*	0.694	*44.09*
Google and EHR	0.542	76.87	0.584	69.17	0.663	55.48	0.658	56.25
AR(52), Google, and EHR	*0.759*	*38.88*	*0.718*	44.63	*0.702*	46.25	*0.701*	47.17

[a] PCC: Pearson correlation coefficient.

[b] RMSE: root mean squared error.

[c] AR(52): autoregressive model of order 52.

[d] Italicization highlights the 2 highest correlations and lowest errors obtained with the models for real time and 1-week, 2-week, and 3-week forecasts.

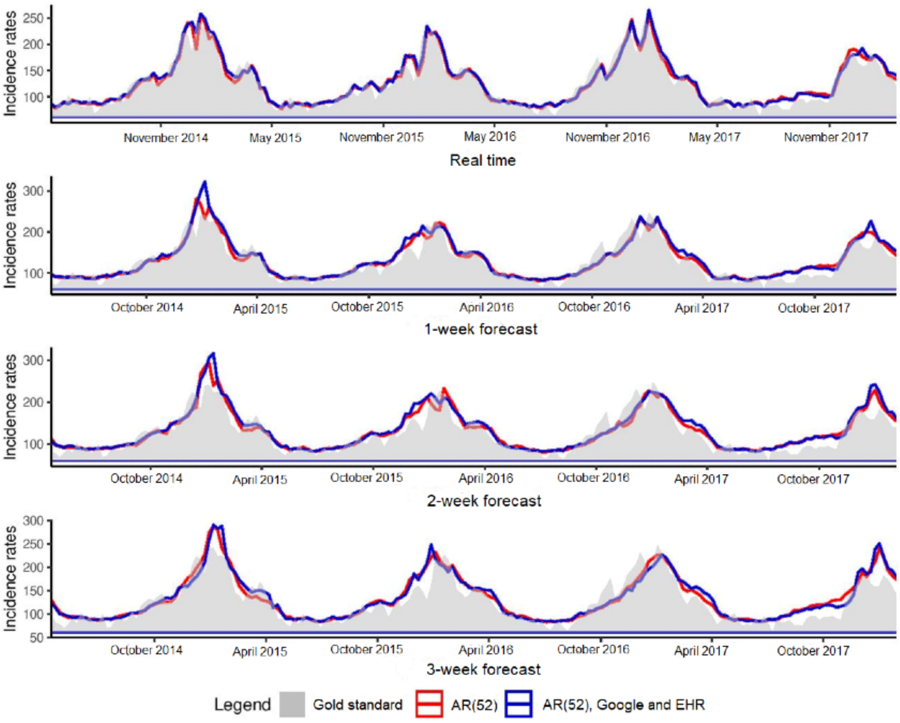
[e] EHR: electronic health record.

National Analysis

For real-time estimates, the error values range from 45.6 to 15.5 and the correlation values range from 0.80 to 0.95, with the lowest error and the highest correlation obtained with the model using only historical data—AR(52)—or the models combining historical data and external data sources. The results are similar for long-term forecasts, with error values ranging from 50.7 to 19.7 and correlation values ranging from 0.62 to 0.91 for 1-week estimates. For 2-week and 3-week estimates, the error values range from 42.6 to 22.8 and 41.9 to 22.3, respectively. In terms of 2-week and 3-week correlation, the values range from 0.74 to 0.90 and from 0.69 to 0.91, respectively.

Figure 5 illustrates the estimates obtained at the national level for forecasts up to 3 weeks with the model using only historical data and the model using historical data and both data sources, Google and EHR. For real-time estimates and long-term forecasts, the results obtained with the 2 models are comparable. In comparison with the linear approach, the nonlinear approach tends to smooth estimates.

Figure 5. National level. Predictions up to 3 weeks obtained at the national level with the model using only historical data and the model using historical data and both data sources, Google and EHR. Gold standard, French Sentinel network data. EHR: electronic health record.

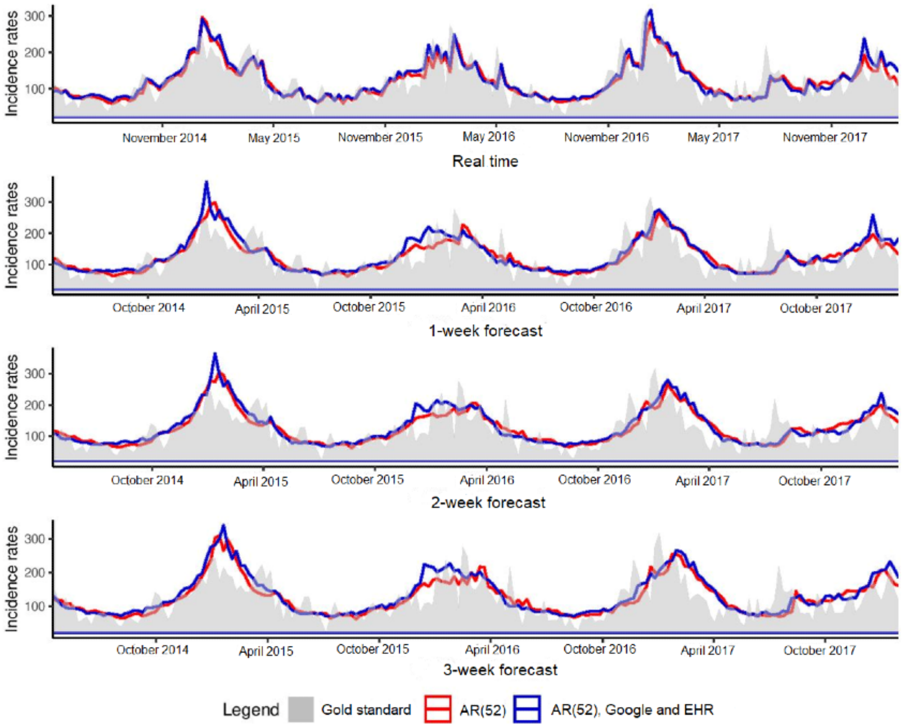


Regional Analysis

For real-time estimates, the error values range from 76.9 to 38.4 and the correlation values range from 0.54 to 0.76, with the lowest error and the highest correlation values obtained with AR(52) model and the models combining historical data and external data sources. For 1-week, 2-week, and 3-week estimates, the error values range from 69.2 to 41.1, from 60.5 to 43.8, and from 63.3 to 44.1, respectively. The lowest errors values for long-term forecasts are all obtained with the model using historical and EHR data. In terms of 1-week, 2-week, and 3-week correlation, the values range from 0.53 to 0.72, from 0.56 to 0.70, and from 0.53 to 0.70, respectively. The highest correlations for long-term forecasts are all obtained with the model using historical data and both data sources, Google and EHR.

Figure 6 illustrates the estimates obtained at the regional level for forecasts up to 3 weeks with the model using only historical data and the model using historical data and both data sources, Google and EHR. At the national level, results are comparable between the 2 models, and the nonlinear approach tends to smooth the estimates.

Figure 6. Regional level. Predictions up to 3 weeks obtained at the regional level with the model using only historical data and the model using historical data and both data sources, Google and EHR. Gold standard, French Sentinel network data. EHR: electronic health record.



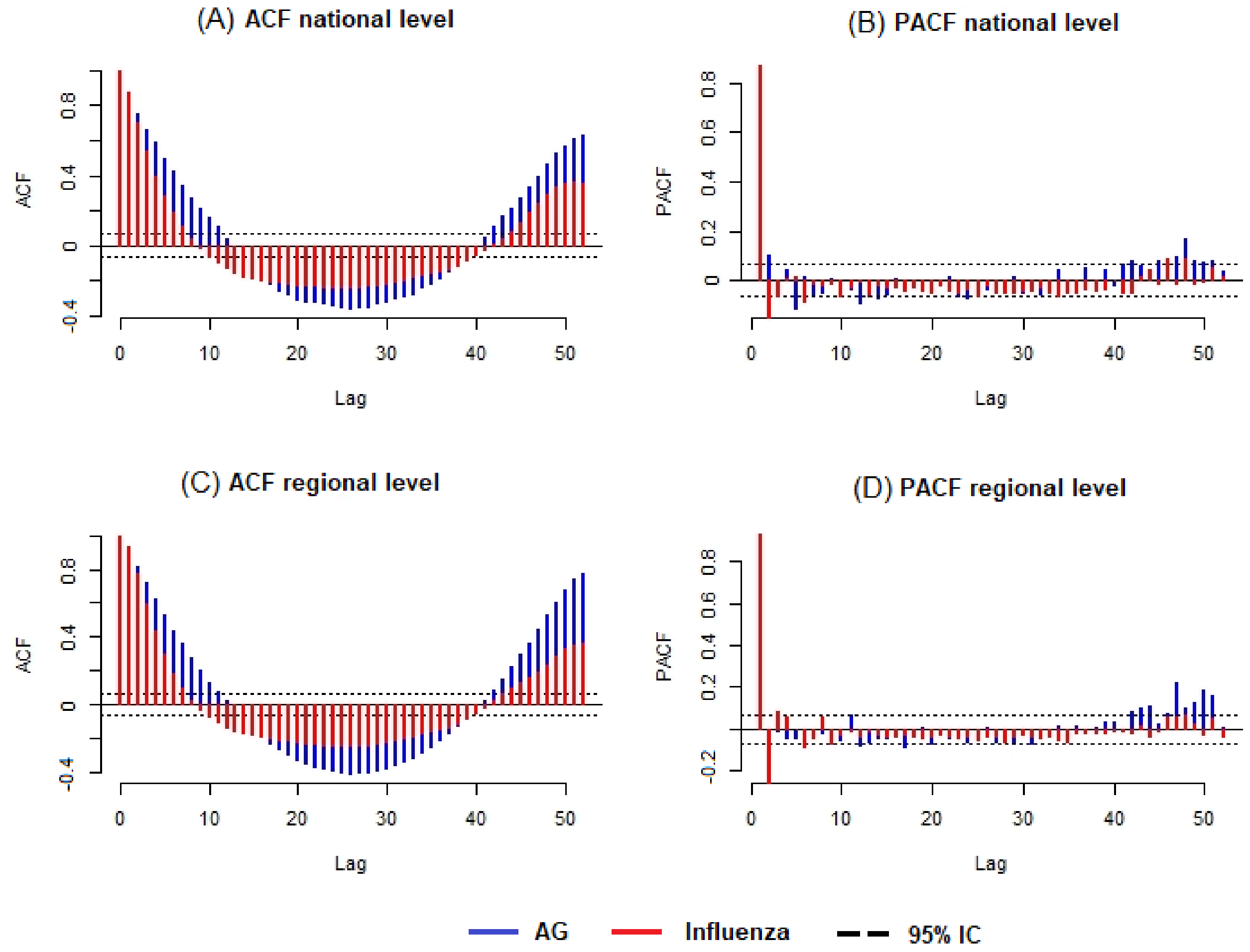
Comparison of AG and Influenza

To assess the role of external data sources in AG forecasting in comparison with influenza forecasting, we studied both time series, at the national and regional levels. As both series were stationary, we compared the seasonality. Figure 7 corresponds to ACF and PACF obtained for AG and influenza.

The ACF plot provides the correlation coefficients between a time series and its lagged values. The PACF plot provides the correlation coefficients between a time series and its lagged values after removing the effects that are already explained by the previous lags.

The ACF plots at the national and regional levels (Figures 7A and 7C) show that both time series, AG and influenza, are seasonal, but with autocorrelation more important for AG than for influenza. This result can explain why historical data are able to provide more information for AG than for influenza. We have similar results for PACF plots (Figures 7B and 7D), at the national and regional levels, where the coefficients of partial autocorrelation are larger for AG than for influenza.

Figure 7. ACF and PACF. Autocorrelation obtained for flu and AG at the national level (Figures A and B) and regional level (Figures C and D). ACF: autocorrelation function; AG: acute gastroenteritis; PACF: partial autocorrelation function.



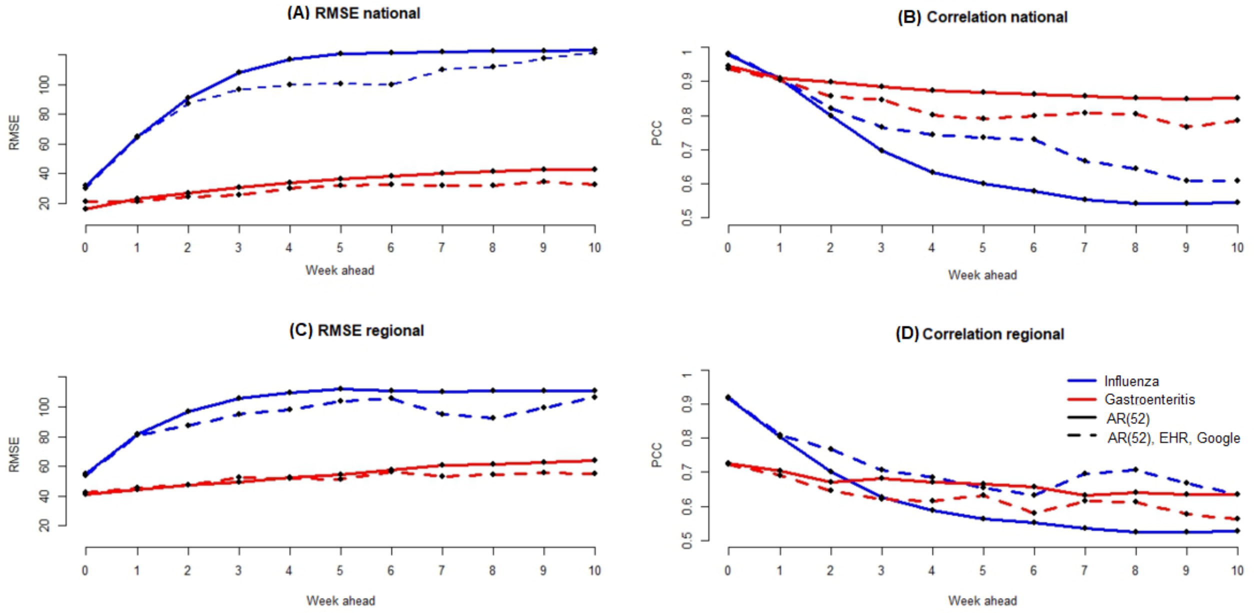
Analysis of Forecast up to 10 Weeks

Linear Approach

Figure 8 and Table S1 in Multimedia Appendix 1 show, for the linear approach, errors and correlation for AG at the national and regional levels, for forecasts up to 10 weeks. At the national level, the lowest error for real-time estimates is obtained with the linear approach using only historical data—AR(52). For long-term forecasts, from up to 1 week to up to 10 weeks, the lowest errors are obtained by using historical data and both data sources, Google and EHR. In terms of correlation, in all cases, the highest values are obtained by using only historical data. At the regional level, in terms of errors, both data sources, Google and EHR, allow to improve accuracy for forecasts from up to 4 weeks to up to 10 weeks. In terms of correlation, results are similar to those at the national level, with high values obtained by using only historical data.

Figure 8 and Table S2 in Multimedia Appendix 1 show, for the linear approach, errors and correlation for influenza at the national and regional levels, for forecasts up to 10 weeks. In contrast to AG at the national and regional levels, in terms of errors and correlation, the most accurate results are obtained by using historical data, Google data, and EHR data.

Figure 8. (A) Error values obtained at the national level for the flu and gastroenteritis for forecasts up to 10 weeks with the Elastic Net model. The solid line corresponds to the results obtained with the Elastic Net model using only historical data. The dotted line corresponds to the results obtained with the Elastic Net model using historical data and both Google and EHR data. The red color is the results for gastroenteritis disease, whereas the blue color is the results for the flu. This style line and color code are used for the 4 panels of this figure. (B) Correlation values obtained at the national level. (C) Error values obtained at the regional level. (D) Correlation values obtained at the regional level. EHR: electronic health record; RMSE: root mean squared error.

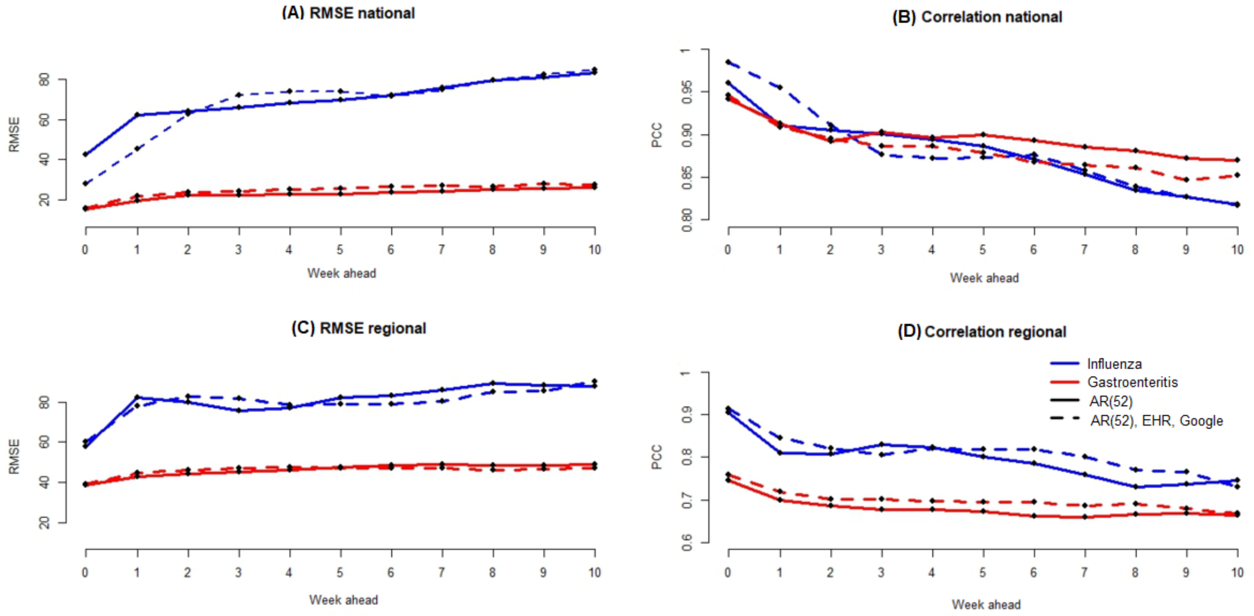


Nonlinear Approach

Figure 9 and Table S3 in Multimedia Appendix 1 show, for the nonlinear approach, errors and correlation for AG at the national and regional levels, for forecasts up to 10 weeks. At the national level, in terms of errors, the lowest values are obtained by using only historical data—AR(52). In terms of correlation, for long-term forecasts, the highest values are obtained by using only historical data. At the regional level, in terms of errors, for forecast up to 4 weeks, the lowest values are obtained by using only historical data. However, for long-term forecasts, the most accurate results are obtained by using historical data and both data sources, Google and EHR.

Figure 9 and Table S4 in Multimedia Appendix 1 show, for the nonlinear approach, errors and correlation for influenza at the national and regional levels, for forecasts up to 10 weeks. At the national level, in terms of errors and correlation, the most accurate values for forecasts up to 2 weeks are obtained by using historical data and both Google and EHR data. For forecasts from up to 3 weeks to up to 5 weeks, most accurate estimates are obtained by using only historical data. For long-term forecasts, results are similar for both models, the one using only historical data and the one using historical data and Google and EHR data. At the regional level, for forecasts up to 4 weeks, in terms of errors, the lowest values are obtained, in most cases, by using only historical data. For long-term forecasts, the most accurate estimates are obtained with the model using historical data and both Google and EHR data. In terms of correlation, in most cases, the highest values are obtained by using historical data and both Google and EHR data.

Figure 9. (A) Error values obtained at the national level for the flu and gastroenteritis for forecasts up to 10 weeks with the RF model. The solid line corresponds to the results obtained with the random forest (RF) model using only historical data. The dotted line corresponds to the results obtained with the RF model using historical data and both Google and EHR data. The red color is the results for gastroenteritis disease, whereas the blue color is the results for the flu. This style line and color code are used for the 4 panels of this figure. (B) Correlation values obtained at the national level. (C) Error values obtained at the regional level. (D) Correlation values obtained at the regional level. EHR: electronic health record.



Discussion

Principal Findings

We adjusted a methodology developed for influenza, to accurately track AG activity. Our method is able to provide forecasts up to 10 weeks for national and regional levels and for emergency and hospitalization stays (Multimedia Appendix 1). To the best of our knowledge, this is a disease and a spatial resolution (French regions and hospitals) for which no forecasting approaches have been explored previously.

In this study, we show that external data sources, EHR and Google, contribute to improving AG surveillance, in particular for long-term forecasts, with more important contribution from historical data. Specifically, when we use the linear approach (Elastic Net), in terms of errors at the national level, the lowest values are obtained by using historical data and both Google and EHR data. These results are consistent for forecasts from up to 1 week to up to 10 weeks (Table S1 in Multimedia Appendix 1). At the regional level, the model using only historical data is the model producing the lowest errors for short-term forecasts (Table S1 in Multimedia Appendix 1). However, for long-term forecasts, the inclusion of external data sources (Google and EHR) improves the estimates. We conducted a Diebold Mariano test [39] to assess if the forecasts are statistically different when using only historical data or the combination of historical data, Google data, and EHR data (Table S5 in Multimedia Appendix 1). We can see that at the national level, the estimates are statistically more accurate when using historical data and both Google and EHR data for 3-week and long-term forecasts. At the regional level, the use of external data sources produces estimates that are statistically more accurate for 7-week and long-term forecasts.

As we used a method developed for influenza outbreaks, we compared the results obtained for AG with those obtained for influenza. At the national and regional levels, with the linear approach, for both short-term and long-term forecasts, the most accurate estimates are obtained with the model using historical data and external data sources (Google and EHR data). An understanding of these results can emerge from the time series analysis (Figure 7). We show that the seasonality is more important for AG epidemics than for influenza, resulting in historical data capable of providing more information for AG than for influenza. Nonetheless, for long-term forecasts, historical data are not sufficient and external data sources can be used to supplement them. Thus, it is important to integrate external data to improve long-term estimates.

In addition to the linear approach, we conducted the same analysis with a nonlinear approach (RF). At the national level, the results differ slightly from those obtained using the linear approach. In terms of error and correlation, the model using only historical data provides more accurate estimates than the model using historical data, Google data, and EHR data. These results are consistent for real-time estimates and long-term forecasts (Table S3 in Multimedia Appendix 1). At the regional level, regarding the linear approach, in terms of error for short-term forecasts, the model using only historical data allows to produce the most accurate estimates. For long-term forecasts, the model including external data sources, Google and EHR, decreases the error. In terms of correlation, for both short-term and long-term forecasts, the model producing the highest values is the model using historical data, Google data, and EHR data. In all cases, the nonlinear approach allows us to obtain high values in terms of correlation and low values in terms of error when compared with those obtained using the linear approach. However, as seen in Figures 5 and 6, the nonlinear approach tends to smooth the estimates compared with those obtained using the linear approach. This can result in decrease in error and increase in correlation.

The fact that we could only access EHR data from Rennes University Hospital, and thus from the Brittany region, prevented us from being able to quantify the added value of nation-wide EHR information. This should be evaluated in future studies by integrating EHR data from different hospitals from all the French regions. However, it is interesting that data from a hospital in Rennes can improve AG forecasting at the national level, even if, as we described previously, EHR data seem more important for the regional level.

Data retrieved from Google Correlate are normalized by Google in a (frequently) distinct sample and over different time periods depending on the data request. This prenormalization can affect our results, but as shown in the study by Arena et al [15], the process of dynamic training minimizes the impact of this instability.

It would be interesting to test other approaches that gave good results for influenza, for example, an ensemble method that combines the power of the linear and the nonlinear approaches [14] or other machine learning methods such as Support Vector Machine or neural networks. We tested a long short-term memory model to forecast gastroenteritis up to 10 weeks. We obtained root mean squared error=2.96 for real-time forecasting. We believe that these results are really promising and could be further studied in the future by developing a neural network combining long short-term memory for historical data and another neural network for external data sources such as Google data or EHR data. In addition, other methods could be tested to obtain more information from external data sources as transformations of the input variables. Variable transformations could be tested on external data sources to check whether we could get more information. Finally, it could be meaningful to first remove the multicollinearity of our predictive variables with traditional methods such as the Variance Inflation Factor and then select the most important variables with a stepwise regression to run a linear regression on the remaining variables.

Conclusions

We show that hospital data and internet search data significantly contribute to predict AG outbreaks, in particular for long-term forecasts. The use of these external data sources in combination with historical data could supplement traditional surveillance systems. The methods we developed could help to reduce the impact of the AG peak, particularly in hospitals, by making it possible to anticipate increased activity by up to 10 weeks.

We acknowledge that there is still scope for improvement. Future studies could explore the incorporation of more information from external data sources as a way to yield more robust results.

Acknowledgments

The authors would like to thank the French Agence Nationale de Recherche for funding this study through the Integrating and Sharing Health Data for Research project (grant ANR-15-CE19-0024). The authors also thank the French Sentinel network and Google search engine for making their data publicly available. MS and CP were partially funded by the National Institute of General Medical Sciences of the National Institutes of Health, under award number R01GM130668. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

CP, AL, and GB conceived the study, and CP and GB obtained the data sets. CP and MS proposed the forecasting methodology. CP conducted the statistical experiments. CP and MS analyzed and interpreted the results. CP wrote the manuscript with support from MS, AL, and GB. All authors reviewed and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Heat maps obtained at both, national and regional levels, for the linear approach at 1-week, 2-week and 3-week forecasts. We also added the correlation and errors obtained up to 10-week forecast, for the linear and nonlinear approaches for both, influenza and gastroenteritis diseases.
[DOCX File , 3952 KB - publichealth_v9i1e34982_app1.docx]

References

1. Farthing M. Diarrhoea: a significant worldwide problem. Int JAntimicrobial Agent 2000 Feb;14(1):65-69. [doi: 10.1016/s0924-8579(99)00149-1]
2. Majowicz SE, Hall G, Scallan E, Adak GK, Gauci C, Jones TF, et al. A common, symptom-based case definition for gastroenteritis. Epidemiol Infect 2008 Jul;136(7):886-894. [doi: 10.1017/S0950268807009375] [Medline: 17686196]
3. Kosek M, Bern C, Guerrant RL. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bull World Health Organ 2003;81(3):197-204 [FREE Full text] [Medline: 12764516]

4. Rivière M, Baroux N, Bousquet V, Ambert-Balay K, Beaudeau P, Jourdan-Da Silva N, et al. Secular trends in incidence of acute gastroenteritis in general practice, France, 1991 to 2015. Euro Surveill 2017 Dec;22(50):17-00121 [FREE Full text] [doi: 10.2807/1560-7917.ES.2017.22.50.17-00121] [Medline: 29258648]
5. VAN CAUTEREN D, De VALK H, VAUX S, Le STRAT Y, VAILLANT V. Burden of acute gastroenteritis and healthcare-seeking behaviour in France: a population-based study. Epidemiol Infect 2011 Jun 07;140(4):697-705. [doi: 10.1017/s0950268811000999]
6. Rohayem J. Norovirus seasonality and the potential impact of climate change. Clin Microbiol Infect 2009 Jun;15(6):524-527 [FREE Full text] [doi: 10.1111/j.1469-0691.2009.02846.x] [Medline: 19604277]
7. Greer AL, Drews SJ, Fisman DN. Why "winter" vomiting disease? Seasonality, hydrology, and Norovirus epidemiology in Toronto, Canada. Ecohealth 2009 Jun 12;6(2):192-199. [doi: 10.1007/s10393-009-0247-8] [Medline: 20151172]
8. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564. [doi: 10.1086/630200] [Medline: 19845471]
9. Situation observed in metropolitan France for week 51 of the year 2022, from 19/12/2022 to 25/12/2022. Sentiweb - the site of the Sentinels Network. URL: https://websenti.u707.jussieu.fr/sentiweb/ [accessed 2022-07-01]
10. Shah MP, Wikswo ME, Barclay L, Kambhampati A, Shioda K, Parashar UD, et al. Near real-time surveillance of U.S. Norovirus outbreaks by the norovirus sentinel testing and tracking network - United States, August 2009-July 2015. MMWR Morb Mortal Wkly Rep 2017 Feb 24;66(7):185-189 [FREE Full text] [doi: 10.15585/mmwr.mm6607a1] [Medline: 28231235]
11. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc Natl Acad Sci U S A 2015 Nov 24;112(47):14473-14478 [FREE Full text] [doi: 10.1073/pnas.1515373112] [Medline: 26553980]
12. Santillana M, Nguyen AT, Louie T, Zink A, Gray J, Sung I, et al. Cloud-based electronic health records for real-time, region-specific influenza surveillance. Sci Rep 2016 May 11;6:25732 [FREE Full text] [doi: 10.1038/srep25732] [Medline: 27165494]
13. Yang S, Santillana M, Brownstein JS, Gray J, Richardson S, Kou SC. Using electronic health records and internet search information for accurate influenza forecasting. BMC Infect Dis 2017 May 08;17(1):332 [FREE Full text] [doi: 10.1186/s12879-017-2424-7] [Medline: 28482810]
14. Lu FS, Hattab MW, Clemente CL, Biggerstaff M, Santillana M. Improved state-level influenza nowcasting in the United States leveraging Internet-based data and network approaches. Nat Commun 2019 Jan 11;10(1):147 [FREE Full text] [doi: 10.1038/s41467-018-08082-0] [Medline: 30635558]
15. Arena C, Amoros JP, Vaillant V, Ambert-Balay K, Chikhi-Brachet R, Jourdan-Da Silva N, et al. Acute diarrhea in adults consulting a general practitioner in France during winter: incidence, clinical characteristics, management and risk factors. BMC Infect Dis 2014 Oct 30;14(1):574 [FREE Full text] [doi: 10.1186/s12879-014-0574-4] [Medline: 25358721]
16. Charles M, Holman R, Curns A, Parashar U, Glass R, Bresee J. Hospitalizations associated with rotavirus gastroenteritis in the United States, 1993-2002. Pediatr Infect Dis J 2006;25(6):489-493. [doi: 10.1097/01.inf.0000215234.91997.21]
17. Hall AJ, Wikswo ME, Manikonda K, Roberts VA, Yoder JS, Gould LH. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. Emerg Infect Dis 2013 Aug;19(8):1305-1309 [FREE Full text] [doi: 10.3201/eid1908.130482] [Medline: 23876187]
18. Amador JJ, Vicari A, Turcios-Ruiz RM, Melendez DC, Malek M, Michel F, et al. Outbreak of rotavirus gastroenteritis with high mortality, Nicaragua, 2005. Rev Panam Salud Publica 2008 Apr;23(4):277-284. [doi: 10.1590/s1020-49892008000400008] [Medline: 18505609]
19. Kirian ML, Weintraub JM. Prediction of gastrointestinal disease with over-the-counter diarrheal remedy sales records in the San Francisco Bay Area. BMC Med Inform Decis Mak 2010 Jul 20;10(1):39 [FREE Full text] [doi: 10.1186/1472-6947-10-39] [Medline: 20646311]
20. Shah M, Lopman B, Tate J, Harris J, Esparza-Aguilar M, Sanchez-Uribe E. Use of internet search data to monitor rotavirus vaccine impact in the United States, United Kingdom, and Mexico. J Pediatric Infect Dis Soc 2016;3(suppl_1):771. [doi: 10.1093/ofid/ofw172.634]
21. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron A. More diseases tracked by using Google Trends. Emerg Infect Dis 2009 Aug;15(8):1327-1328 [FREE Full text] [doi: 10.3201/eid1508.090299] [Medline: 19751610]
22. Adadi A, Adadi S, Berrada M. Gastroenterology meets machine learning: status quo and quo vadis. Adv Bioinformatics 2019 Apr 02;2019:1870975-1870924 [FREE Full text] [doi: 10.1155/2019/1870975] [Medline: 31065266]
23. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
24. Butler D. When Google got flu wrong. Nature 2013 Feb 14;494(7436):155-156. [doi: 10.1038/494155a] [Medline: 23407515]
25. Bouzillé G, Poirier C, Campillo-Gimenez B, Aubert M, Chabot M, Chazard E, et al. Leveraging hospital big data to monitor flu epidemics. Comput Methods Programs Biomed 2018 Feb;154:153-160 [FREE Full text] [doi: 10.1016/j.cmpb.2017.11.012] [Medline: 29249339]
26. Poirier C, Lavenu A, Bertaud V, Campillo-Gimenez B, Chazard E, Cuggia M, et al. Real time influenza monitoring using hospital big data in combination with machine learning methods: comparison study. JMIR Public Health Surveill 2018 Dec 21;4(4):e11361 [FREE Full text] [doi: 10.2196/11361] [Medline: 30578212]

27. Situation observed in metropolitan France for week 01 of the year 2023, from 02/01/2023 to 08/01/2023. Sentinelles. 2022 Nov 1. URL: http://websenti.u707.jussieu.fr/sentiweb [accessed 2023-01-12]
28. https://trends.google.fr/trends/?geo=FR. URL: https://trends.google.fr/trends/?geo=FR [accessed 2023-01-11]
29. Mohebbi M, Vanderkam D, Kodysh J, Schonberger R, Choi H, Kumar S. Google correlate whitepaper. Google. 2011. URL: https://research.google/pubs/pub41695/ [accessed 2018-03-05]
30. Zou H, Hastie T. Regularization and variable selection via the elastic net. J Royal Statistical Soc B 2005 Apr;67(2):301-320. [doi: 10.1111/j.1467-9868.2005.00503.x]
31. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Statistical Soc B (Methodological) 2018 Dec 05;58(1):267-288. [doi: 10.1111/j.2517-6161.1996.tb02080.x]
32. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 1970 Feb;12(1):55-67. [doi: 10.2307/1271436]
33. caret (Classification And Regression Training) R package that contains misc functions for training and plotting classification and regression models. GitHub. URL: https://github.com/topepo/caret/ [accessed 2020-05-10]
34. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2015.
35. Breiman L. Random forests. In: Machine Learning. Cham: Springer; 2001.
36. Kane MJ, Price N, Scotch M, Rabinowitz P. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. BMC Bioinformatics 2014 Aug 13;15(1):276 [FREE Full text] [doi: 10.1186/1471-2105-15-276] [Medline: 25123979]
37. Dudek G. Short-term load forecasting using random forests. In: Intelligent Systems'2014. Cham: Springer; 2015.
38. Classification and regression by randomForest. R News. 2002. URL: https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf [accessed 2021-07-03]
39. Diebold FX, Mariano RS. Comparing predictive accuracy. J Business Econ Stat 1995 Jul;13(3):253-263. [doi: 10.1080/07350015.1995.10524599]

Abbreviations

ACF: autocorrelation function
AG: acute gastroenteritis
AR(52): autoregressive model of order 52
CDW: clinical data warehouse
eHOP: entrepôt de données de l'HÔPital
EHR: electronic health record
PACF: partial autocorrelation function
RF: random forest

Edited by G Eysenbach, H Bradley; submitted 15.11.21; peer-reviewed by A Staffini, YL Cheong, E Sükei; comments to author 24.02.22; revised version received 19.07.22; accepted 28.11.22; published 31.01.23.

Please cite as:
Poirier C, Bouzillé G, Bertaud V, Cuggia M, Santillana M, Lavenu A
Gastroenteritis Forecasting Assessing the Use of Web and Electronic Health Record Data With a Linear and a Nonlinear Approach: Comparison Study
JMIR Public Health Surveill 2023;9:e34982
URL: https://publichealth.jmir.org/2023/1/e34982
doi:10.2196/34982
PMID:36719726

Original Paper

Evaluating the Discriminatory Ability of the Sickle Cell Data Collection Program's Administrative Claims Case Definition in Identifying Adults With Sickle Cell Disease: Validation Study

Ashima Singh[1], MS, PhD; Marci K Sontag[2], PhD; Mei Zhou[3], MS; Mahua Dasgupta[1], MS; Tessa Crume[4], MSPH, PhD; Morgan McLemore[5], MD; Najibah Galadanci[6], MD; Eldrida Randall[5], MS; Nicole Steiner[1], BS; Amanda M Brandow[1], DO, MS; Kathryn Koch[7], APNP; Joshua J Field[7,8], MD; Kathryn Hassell[9], MD; Angela B Snyder[3*], MPH, PhD; Julie Kanter[6*], MD

[1]Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI, United States
[2]Center for Public Health Innovation, CI International, Littleton, CO, United States
[3]Georgia Health Policy Center, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA, United States
[4]Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Center, Aurora, CO, United States
[5]Department of Hematology and Oncology, Winship Cancer Institute, Emory School of Medicine, Atlanta, GA, United States
[6]Department of Medicine, University of Alabama Birmingham, Birmingham, AL, United States
[7]Department of Medicine, Medical College of Wisconsin, Milwaukee, WI, United States
[8]Versiti Blood Center of Wisconsin, Milwaukee, WI, United States
[9]University of Colorado, Aurora, CO, United States
*these authors contributed equally

Corresponding Author:
Julie Kanter, MD
Department of Medicine
University of Alabama Birmingham
1720 2nd street South, NP2510
Birmingham, AL, 35209
United States
Phone: 1 205 801 9034
Email: jkanter@uabmc.edu

Abstract

Background: Sickle cell disease (SCD) was first recognized in 1910 and identified as a genetic condition in 1949. However, there is not a universal clinical registry that can be used currently to estimate its prevalence. The Sickle Cell Data Collection (SCDC) program, funded by the Centers for Disease Control and Prevention, funds state-level grantees to compile data within their states from various sources including administrative claims to identify individuals with SCD. The performance of the SCDC administrative claims case definition has been validated in a pediatric population with SCD, but it has not been tested in adults.

Objective: The objective of our study is to evaluate the discriminatory ability of the SCDC administrative claims case definition to accurately identify adults with SCD using Medicaid insurance claims data.

Methods: Our study used Medicaid claims data in combination with hospital-based medical record data from the Alabama, Georgia, and Wisconsin SCDC programs to identify individuals aged 18 years or older meeting the SCDC administrative claims case definition. In order to validate this definition, our study included only those individuals who were identified in both Medicaid's and the partnering clinical institution's records. We used clinical laboratory tests and diagnostic algorithms to determine the true SCD status of this subset of patients. Positive predictive values (PPV) are reported overall and by state under several scenarios.

Results: There were 1219 individuals (354 from Alabama and 865 from Georgia) who were identified through a 5-year time period. The 5-year time period yielded a PPV of 88.4% (91% for data from Alabama and 87% for data from Georgia), when only using data with laboratory-confirmed (gold standard) cases as true positives. With a narrower time period (3-year period) and data from 3 states (Alabama, Georgia, and Wisconsin), a total of 1432 individuals from these states were included in our study. The overall 3-year PPV was 89.4% (92%, 93%, and 81% for data from Alabama, Georgia, and Wisconsin, respectively) when only considering laboratory-confirmed cases as true cases.

Conclusions: Adults identified as having SCD from administrative claims data based on the SCDC case definition have a high probability of truly having the disease, especially if those hospitals have active SCD programs. Administrative claims are thus a valuable data source to identify adults with SCD in a state and understand their epidemiology and health care service usage.

(JMIR Public Health Surveill 2023;9:e42816) doi:10.2196/42816

KEYWORDS

surveillance using administrative data; rare conditions; sickle cell disease; disease; surveillance; genetic; prevention; data; adults; epidemiology; utilization

Introduction

Sickle cell disease (SCD) is a rare genetic condition, which has consistently suffered from disparity in terms of research funding, availability of registries, and surveillance programs [1]. Recently, the National Academies of Sciences, Engineering, and Medicine (NASEM) published a strategic plan and blueprint for action addressing SCD [2]. One of the recommendations of the NASEM report includes establishing a nationwide surveillance program for SCD in addition to a longitudinal clinical registry. A clinical registry is an organized system of data collection using observational study methods to evaluate specified outcomes for a population defined by a particular disease or condition (in this case, SCD). In most cases, a registry requires patient consent to participate, which may limit the inclusion of all individuals with the specific condition. In contrast, surveillance data are defined as the ongoing, systematic collection, analysis, and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice [2]. While a registry can inform on clinically specific information including disease trajectory, a surveillance program provides a more population-based assessment and is often more inclusive of the entire population. The Centers for Disease Control and Prevention has invested in establishing the Sickle Cell Data Collection (SCDC) [3] program to establish a population-based surveillance program in multiple US states using a combination of data sources. A population-based surveillance program for SCD both helps to reveal the distribution of individuals living with SCD and inform stakeholders about the health care usage patterns to improve outcomes. Understanding the patterns of care for people with SCD can increasingly ensure sufficient staffing, and expertise is available to treat the patient population.

Administrative claims data are one of the data sources used in the SCDC population-based surveillance program. Currently, all states participating in the SCDC program include or plan to include data from different administrative sources as they establish their state-specific programs in addition to data from newborn screening programs (NBSs), SCD center–specific databases, electronic health records data warehouses, and vital health records. In contrast to data from NBS programs or clinical databases, claims data are only based on codes of the International Classification of Diseases, Ninth and Tenth Revisions (ICD-9 and -10, respectively), and not clinically verified data. Despite this limitation, claims data are a valuable source of information since they include comprehensive information on health care service usage for their beneficiaries beyond specific hospitals and clinics. This is specifically important for the SCDC surveillance efforts, which aim to include all individuals living with SCD in participating states. Prior work has validated the definition of ≥3 SCD-coded encounters within 5 years to identify a pediatric patient cohort at a large urban children’s hospital [4]. This definition, however, has not been tested in the adult population. As the SCDC program expands to multiple states, the generalizability and validity of the SCDC administrative case definition needs to be determined. Also, some of the recently participating states have limited years of available data. The objective of our study is for 3 of the SCDC states to evaluate the discriminatory ability of the SCDC administrative case definition to identify adults with SCD (≥18 years of age) who are Medicaid beneficiaries, using Medicaid insurance claims data from their state, considering confirmatory laboratory assessment of SCD as the gold standard. We hypothesize that the SCDC case definition applied to Medicaid claims data will have a high positive predictive value (PPV) for identifying SCD cases among people with SCD aged ≥18 years, but that it will be lower than that observed among pediatric cases.

Methods

Administrative Claims Data Sources to Identify Adults With SCD

Our study includes Medicaid claims data from the Alabama, Georgia, and Wisconsin SCDC programs.

Inclusion Criteria

Adult individuals (aged 18 years or older at the beginning of the study period) with 3 or more claims with an SCD ICD code (ICD-9: 282.41, 282.42, and 282.6*; ICD-10: D57.0*, D57.1, D57.2*, D57.4*, and D57.8*) in their state-specific Medicaid data were eligible for inclusion. In addition, eligible individuals also needed evidence of at least 1 visit (with any diagnosis) at the specific clinical partner institutions. Our 5-year time frame analysis included data from Alabama and Georgia. Eligible individuals identified within Alabama Medicaid claims data from 2015 to 2019 were cross-referenced with electronic health record data from the University of Alabama at Birmingham’s hospital system. Eligible individuals identified within Georgia’s Medicaid claims data from 2015 to 2019 were included if any claim contained a Grady Memorial Hospital facility code. Our 3-year time frame analysis included data from Alabama, Georgia, and Wisconsin. The inclusion criteria for Alabama and Georgia were identical, as described above, with a restricted time frame of 2015-2017. The Wisconsin data included eligible individuals identified within Medicaid claims data from 2018 to 2020 who were linked and matched with electronic health record data from Froedtert Hospital.

Ethics Approval

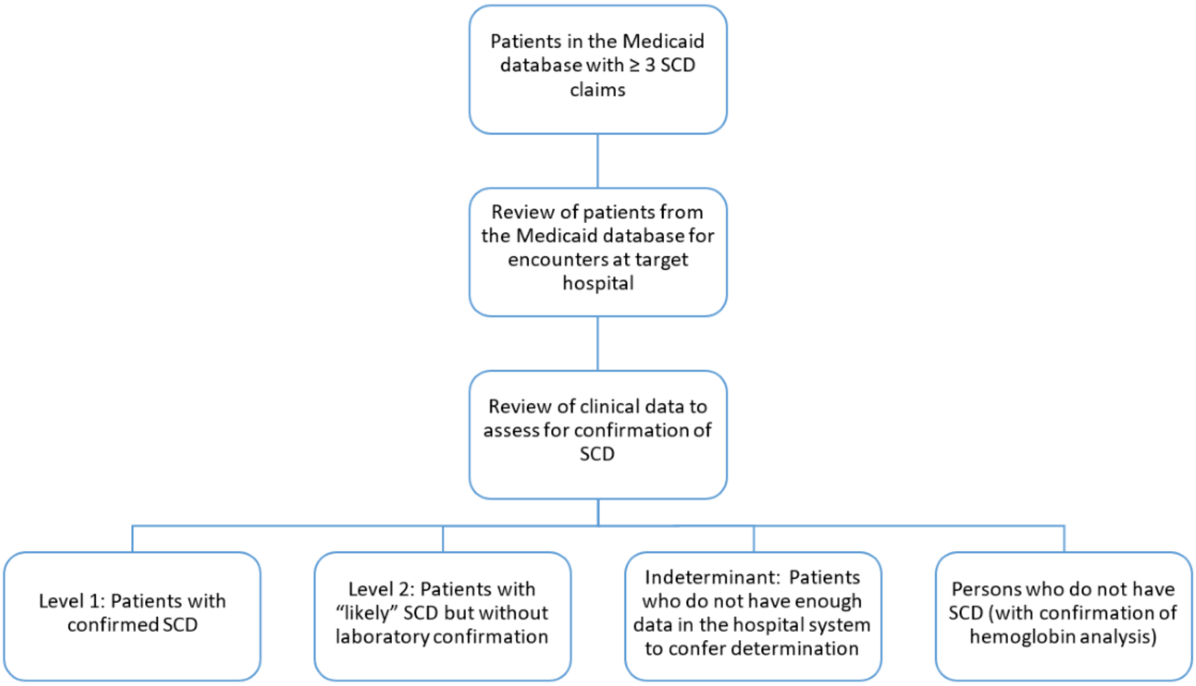
All states providing data had approvals or exemption to conduct the project. Specifically, the University of Alabama Institutional Review Board (IRB) approved the project as Non-human subjects research (IRB-300004733), the Georgia State University's IRB approved the state's SCDC program under a public health exemption (protocol #H11142), and the Medical College of Wisconsin IRB granted an exemption for the study in accordance with 45 CFR 46.104(d)(4) (PRO00043293).

Validation Data

SCD centers from the University of Alabama at Birmingham, Grady Memorial Hospital, and Froedtert Hospital contributed their patient-level databases, which included people with “confirmed” SCD. Hematologists at each site also reviewed the laboratory reports of all the individuals identified at their institutions as part of this study. An individual had confirmatory evidence of SCD if he or she had a confirmed Clinical Laboratory Improvement Amendments of 1988–certified laboratory result from a hemoglobin electrophoresis or high-performance liquid chromatography test in their medical records. All patients with confirmed SCD had these laboratory tests in their medical records (even if they were not active patients within the SCD center). For individuals without evidence of confirmatory laboratory testing, a clinical algorithm was used to evaluate nondiagnostic laboratory data available in the electronic medical record system to classify them as being “likely” to have SCD. The clinical algorithm considered individuals “likely” to have SCD if they had at least 15 laboratory values that included a total bilirubin level of >1.1 mg/dL and any one of the following: reticulocyte proportion of >2%, lactate dehydrogenase level of >250 units/L, or hemoglobin level of <11 g/dL. Individuals with a hemoglobin analysis in their medical records, confirming that they did not have SCD (including those with the sickle cell trait), were categorized as “do not have SCD.” Individuals who did not have confirmatory laboratory evidence regarding their SCD status or those who were not classified as likely cases were labeled as “indeterminant” cases, implying that there was not enough evidence in their medical record to make a diagnosis.

The stepwise method of identification of the eligible population and the classification of cases into various categories is illustrated in Figure 1.

Figure 1. Overview of study methodology. SCD: sickle cell disease.



Analysis

Summary statistics were calculated to describe the demographic characteristics of age, sex, and race and ethnicity based on information available in each state’s Medicaid data. Analysis was conducted for 2 time periods based on available data: 5 years (Alabama and Georgia) and 3 years (Alabama, Georgia, and Wisconsin). We determined the PPV, overall and by state, under various scenarios for the respective time frames. In the first scenario, true positives (TPs) included individuals “confirmed” to have SCD. In the second scenario, TPs also included individuals “confirmed” and “likely” to have SCD. The PPV was calculated using the TP as the numerator and the total number of SCD cases identified per the definition as the denominator (TPs+false positives [FPs]). As an alternate strategy, we also calculated a best-case PPV excluding individuals with an “indeterminant” SCD status for both the scenarios described above. This was done because “indeterminant” cases did not have sufficient information to be either classified as TP or FP. Exact binomial 95% CIs were reported for all proportions. For the group of patients recognized as FP cases based on confirmatory evidence of not having SCD in their clinical data, we provide a count of those with the sickle cell trait.

Results

Overview

There were a total of 3425 (Alabama, n=935; Georgia, n=2490) and 3365 (Alabama, n=803; Georgia, n=2055; Wisconsin, n=507) individuals who were aged ≥18 years and had at least ≥3 Medicaid SCD claims in the 5-year and 3-year time frame, respectively. A total of 1219 unique individuals (354 from Alabama and 865 from Georgia) in the 5-year time period and 1432 (328 from Alabama, 684 from Georgia, and 420 from Wisconsin) in the 3-year time period had at least 3 Medicaid SCD claims and an encounter with the respective partner clinical institution. Importantly, not all of those individuals with SCD claims were seen by the sickle cell center at their respective institutions (ie, they only had to be seen within the hospital system). The demographic characteristics and the total number of SCD-related Medicaid encounters are described in Table 1.

Table 1. Demographic characteristics of individuals included in the study.

Characteristics	Values							
	Total		Alabama		Georgia		Wisconsin	
	5-year time frame (n=1219)	3-year time frame (n=1432)	5-year time frame (n=354)	3-year time frame (328)	5-year time frame (n=865)	3-year time frame (n=684)	5-year time frame	3-year time frame (n=420)
Age (years), mean (SD)	30.1 (10.7)	32.1 (11.3)	29.0 (9.3)	28.8 (9.4)	30.5 (11.2)	30.03 (10.8)	N/A[a]	38.0 (11.4)
Age group (years), n (%)								
18-29	735 (60)	731 (51)	237 (67)	221 (67)	489 (57)	397 (58)	N/A	113 (27)
30-39	299 (25)	396 (28)	76 (21)	68 (21)	223 (26)	175 (26)	N/A	153 (36)
40-49	119 (10)	184 (13)	26 (7)	24 (7)	93 (11)	70 (10)	N/A	90 (21)
50-59	49 (4)	88 (6)	12 (3)	12 (4)	37 (4)	27 (4)	N/A	49 (12)
≥60	26 (2)	36 (2)	3 (1)	3 (1)	23 (3)	15 (2)	N/A	18 (4)
Sex (female), n (%)	743 (61)	847 (59)	224 (63)	209 (64)	519 (60)	396 (58)	N/A	242 (57)
SCD[b] encounters per patient during the study period, mean (SD)	64.5 (80.5)	42.3 (51.4)	47.8 (33.5)	31.4 (20.2)	71.4 (92.3)	51.1 (61.7)	N/A	36.5 (47.2)

[a] N/A: not applicable.

[b] SCD: sickle cell disease.

PPV Based on Data With a 5-Year Time Period

Table 2 shows the demographic characteristics of individuals in each validation category by state. The mean age of individuals with confirmed SCD from Alabama and Georgia was 27.9 (SD 8.6) years and 29.8 (SD 10.7) years, respectively, with the majority of them being in the 18-29–year age group. Overall, 69% of individuals with confirmed SCD across Alabama and Georgia had Hemoglobin SS/Sβ0 type of disease. PPV values by state are shown in Table 3. The overall PPV was 88% when considering only those with a confirmatory clinical laboratory test indicating SCD as TPs. For the scenario that included both likely and confirmed cases as TP, the PPV was slightly higher (89%).

Overall, across Alabama and Georgia, there were 58 cases (14 from Alabama and 44 from Georgia) that were FP (cases with confirmatory evidence of not having SCD). Of them, 36% (5 from Alabama and 16 from Georgia) had the sickle cell trait. There were 73 cases that were indeterminate (8 from Alabama and 65 from Georgia). When excluding the indeterminate cases from the denominator (data not shown in tables), the PPVs for the 2 scenarios of TPs were 94% and 95%, respectively. The PPVs (when excluding indeterminate cases from the denominator; data not shown in tables) remained at >83% for both the TP scenarios irrespective of age group and state.

Table 2. Demographic characteristics of individuals with confirmed sickle cell disease (SCD), likely SCD, those not having SCD, and those with an indeterminate status identified by the clinical validation process by state over a 5-year period.

	Alabama[a] (n=354)				Georgia[b] (n=865)			
	Confirmed SCD (n=323)	Likely SCD (n=9)	Confirmed as not having SCD (n=14)	Indeterminant status (n=8)	Confirmed SCD (n=755)	Likely SCD (n=1)	Confirmed as not having SCD (n=44)	Indeterminant status (n=65)
Age (years), mean (SD)	27.9 (8.6)	32.9 (4.8)	27.3 (5.6)	39.1 (11.7)	29.8 (10.7)	—[c]	30.3 (11.1)	38.7 (13.3)
Age groups (years), n (%)								
18-29	224 (69)	1 (11)	10 (71)	2 (25)	448 (59)	—	26 (59)	15 (23)
30-39	62 (19)	8 (89)	4 (29)	2 (25)	189 (25)	—	10 (23)	24 (37)
40-49	23 (7)	—	—	3 (38)	70 (9)	1 (100)	7 (16)	15 (23)
50-59	11 (3)	—	—	1 (12)	33 (4)	—	—	4 (6)
≥60	3 (1)	—	—	—	15 (2)	—	1 (2)	7 (11)
Sex (female), n (%)	204 (63)	6 (67)	9 (64)	4 (50)	430 (57)	—	39 (89)	50 (77)
SCD encounters per patient during the study period, mean (SD)	50.1 (33.5)	30 (37)	19.1 (11.3)	17.4 (18.4)	80.2 (96.0)	27 (N/A[d])	14.0 (15.9)	9.2 (17.6)

[a] SCD types in Alabama: hemoglobin SS/Sβ0, n=205 (63%); hemoglobin SC, n=68 (21%); others, n=50 (15%).

[b] SCD types in Georgia: hemoglobin SS/Sβ0, n=536 (71%); hemoglobin SC, n=176 (23%); others, n=43 (6%).

[c] Not available.

[d] N/A: not applicable.

Table 3. The positive predictive value (PPV) of the Sickle Cell Data Collection (SCDC) case definition by state over a 5- and 3-year period.

	5-year time period		3-year time period	
	PPV including only confirmatory cases as TP[a] (%), PPV (95% CI)	PPV including confirmatory and likely cases as TP (%), PPV (95% CI)	PPV including only confirmatory cases as TP (%), PPV (95% CI)	PPV including confirmatory and likely cases as TP (%), PPV (95% CI)
Alabama	91 (87-94)	94 (91-96)	92 (89-95)	95 (92-97)
Georgia	87 (85-90)	87 (85-90)	93 (91-95)	93 (91-95)
Wisconsin	N/A[b]	N/A	81 (77-85)	82 (78-86)

[a] TP: true positive.

[b] N/A: not applicable.

PPV Based on Data With a 3-Year Time Period

Table 4 shows the demographic characteristics of individuals in each validation category by state. Similar to the 5-year time frame, the mean age of individuals confirmed with SCD was 28.0 (SD 9.5) years and 29.6 (SD 10.5) years, respectively, with the majority being in the 18-29–year age group. The individuals confirmed with SCD from the Wisconsin records had a higher mean age (36.7, SD 10.5 years), with 30% and 38% of them being in the 18-29–year and 30-39–year age group, respectively. The PPVs for all scenarios by state are shown in Table 3. The overall PPV was 89% when considering only those with a confirmatory clinical laboratory test indicating SCD as TP. For the scenario that included both likely and confirmed cases as TP, the PPV was slightly higher (91%).

Overall, across Alabama, Georgia, and Wisconsin, there were 60 cases (11 from Alabama, 21 from Georgia, and 28 from Wisconsin) that were FP (cases with confirmatory evidence of not having SCD), of whom 50% (5 from Alabama, 9 from Georgia, and 16 from Wisconsin) had the sickle cell trait. There were 79 (6 from Alabama, 25 from Georgia, and 48 from Wisconsin) indeterminate cases. When excluding the indeterminate cases from the denominator, the PPVs for the 2 scenarios of TPs were 95% and 96%, respectively. The PPVs (when excluding indeterminate cases from the denominator; data not shown in tables) remained at >83% for both the TP scenarios irrespective of age group and state.

Table 4. Demographic characteristics of individuals with confirmed sickle cell disease (SCD), likely SCD, those not having SCD, and those with an indeterminate status identified by the clinical validation process by state over a 3-year period.

	Alabama[a] (n=328)				Georgia[b] (n=684)				Wisconsin[c] (n=420)			
	Confirmed SCD (n=302)	Likely SCD (n=8)	Confirmed as not having SCD (n=12)	Indeterminant status (n=6)	Confirmed SCD (n=638)	Likely SCD (n=0)	Confirmed as not having SCD (n=21)	Indeterminant status (n=25)	Confirmed SCD (n=340)	Likely SCD (n=4)	Confirmed as not having SCD (n=28)	Indeterminant status (n=48)
Age (years), mean (SD)	28.0 (9.5)	32.9 (5.1)	28.58 (6.3)	36.8 (12.2)	29.6 (10.5)	N/A[d]	31.1 (9.6)	39.2 (14.5)	36.7(10.5)	43.8 (11.5)	37.4(8.6)	46.4 (14.9)
Age groups (years), n (%)												
18-29	210 (70)	1 (12)	8 (67)	3 (50)	380 (60)	N/A	10 (48)	7 (28)	101 (30)	0 (0)	5 (18)	7 (14)
30-39	55 (18)	7 (88)	4 (33)	2 (33)	160 (25)	N/A	6 (29)	9 (36)	129 (38)	1 (25)	12 (43)	9 (19)
40-49	23 (8)	—[e]	—	1 (17)	63 (10)	N/A	5 (24)	2 (8)	64 (19)	2 (50)	9 (32)	15 (31)
50-59	11 (4)	—	—	1 (17)	23 (4)	N/A	—	4 (16)	37 (11)	1 (25)	2 (7)	8 (17)
≥60	3 (1)	—	—	—	12 (2)	N/A	—	3 (12)	9 (3)	0 (0)	0 (0)	9 (19)
Sex (female), n (%)	192 (64)	7 (88)	7 (58)	3 (50)	358 (56)	N/A	19 (90)	19 (76)	183 (54)	3 (75)	21 (75)	35 (73)
SCD encounters per patient during the study period, mean (SD)	32.4 (20.2)	22 (17)	15.1 (7.1)	16.2 (13.5)	53.9 (62.7)	N/A	12.6 (14.9)	11.7 (26.3)	40.3 (43.7)	7.5 (2.6)	11.3 (20.7)	26.5 (72.0)

[a] SCD types in Alabama: hemoglobin SS/Sβ0, n=199 (66%); hemoglobin SC, n=57 (19%); others, n=46 (15%).

[b] SCD types in Georgia: hemoglobin SS/Sβ0, n=458 (72%); hemoglobin SC, n=146 (23%); others, n=34 (5%).

[c] SCD types in Wisconsin: hemoglobin SS/Sβ0, n=220 (65%); hemoglobin SC, n=86 (25%); others, n=35 (10%).

[d] N/A: not applicable.

[e] Not available.

Discussion

Principal Findings

This study supports the use of a standardized surveillance case definition within administrative claims data, specifically using Medicaid claims (≥3 ICD-9 or -10 codes), across multiple states to identify adults living with SCD. The PPVs achieved through systematic application of the case definitions among adults who receive care at hospitals with an SCD program has resulted in PPVs of ~90%, which is laudable. When excluding the indeterminate cases, the PPV of the SCDC administrative case definition for adults is similar to the PPV of >95% demonstrated in pediatric populations [4,5].

The lower PPV in adults may have resulted from inaccurate coding in Medicaid claims. A sizeable proportion of the FPs identified within Medicaid data included individuals with the sickle cell trait. This represents an opportunity to better educate providers and the general population about the distinction between the 2 conditions. The sickle cell trait is a mostly benign condition and individuals do not experience vaso-occlusive pain episodes. In contrast, individuals with SCD are at risk for significant organ damage as well as acute pain episodes. Thus, research that incorporates individuals with the sickle cell trait into the SCD cohort may lead to substantial underestimation of the disease burden. This emphasizes the need for research specific to SCD based on the combination of information from multiple data sets to minimize the inclusion of individuals with the sickle cell trait. This also supports algorithms developed for other types of data sources such as electronic health record data that exclude individuals with the sickle cell trait when defining those with SCD [6,7]. Specifically, the lower predictive value in Wisconsin, as compared to that in Alabama and Georgia, was mainly driven by the number of indeterminant cases. The indeterminant cases did not have enough information to be classified as having or not having SCD. Interestingly, Wisconsin is also the newest state to join the SCDC program, which may be affecting their number of indeterminate cases (and it may change over time). While better education of providers managing people with SCD may result in improved coding practices, these inaccuracies may also occur with ancillary providers, such as those providing radiology or laboratory testing services, who may not directly interact with the patient. These ancillary providers could contribute to both FPs if they assume that a patient has SCD because he or she was referred by a SCD specialist or had been seen at a hematology clinic, as well as to false negatives since they must often assume a patient's diagnosis based on conditions inaccurately listed on a problem list. Further, many physicians caring for people with SCD also see individuals with other hematologic conditions. Moreover, there are no codes for "hemoglobinopathy screening" to be used by providers assessing these individuals. Some individuals may be miscategorized as having “SCD” simply because they are undergoing hemoglobinopathy testing, which can be carried forward in the medical record. Fragmentation of care for adult patients with SCD [8,9] may also result in less consistent diagnosis and coding.

As efforts continue to improve access to care for adults with SCD, the SCDC program remains an important source of data to characterize SCD including evaluations of SCD management and acute care use on a population level. Administrative data are paramount to creating the surveillance database used by state SCDC programs. While many sickle cell centers have local patient databases, a single current national registry for SCD does not exist, although efforts are underway through the National Alliance of Sickle Cell Centers [10], the National Institutes of Health’s Sickle Cell Disease Implementation Consortium [11], and the American Society of Hematology’s Research Collaborative [12]. However, adults living with SCD may be excluded from these efforts if they do not receive care at one of the participating comprehensive SCD centers. Thus, it is not possible to include the entire population living with SCD in a registry, and performance metrics including sensitivity become impossible to calculate. While efforts are underway to improve access to care for those with SCD [13], it is vital that we have a surveillance program to better understand the health care usage pattern of those living with SCD to determine where additional SCD centers should be established.

Georgia and California were the first 2 states to develop robust SCD surveillance programs. Current SCD surveillance efforts are expanding as the SCDC program receives additional federal funding to extend to additional states. As new states join the SCDC program, it becomes increasingly important to ensure that epidemiologic studies resulting from these data provide an accurate portrayal of SCD in terms of its morbidity and mortality across all age groups. Notably, PPVs consistently over 80%, even among older age groups, support the use of these data to track survival and other outcomes in SCD over the individuals’ lifespans with reasonable accuracy. These findings add to the data supporting state programs in equitable resource allocation and in determining the best strategies to improve outcomes and quality of life for those living with SCD. Administrative data are especially important in identifying adults who are too old to have been screened for SCD at birth (ie, missed the NBS window) as well as those adults who have little access to SCD centers. Additionally, these findings demonstrate the importance of continued data collection and support for the SCDC program, which will allow for additional refinement of surveillance-based case definitions for SCD.

Limitations

There are several limitations to this study. First, individuals who do not have SCD are not included in the SCDC program. Thus, we are unable to comment on the number of false negative codes within the data sets. As a result, it is not feasible to assess the specificity of the data definition on a population level. In addition, these studies were conducted at hospitals that treat a large number of individuals living with SCD (and have National Alliance of Sickle Cell Centers–recognized SCD centers). Thus, it is possible that hospitals that. Although we report a conservative estimate of the definition’s PPVs, the performance of the definition might differ at nonacademic community-based hospitals or those in rural settings or both. Finally, there are several indeterminate cases where insufficient laboratory data were available to confirm an SCD diagnosis. This limitation highlights the importance of obtaining baseline data for persons with SCD newly seen at any hospital system. Confirming an SCD diagnosis is clinically important for medical management and should not be assumed based on a previously notated problem in the medical history or a previous administrative code. Instead, providers need to ensure that all individuals who present with a diagnosis of SCD (or having a chief complaint of an SCD-related symptom) have a hemoglobin analysis performed if never previously evaluated.

Conclusions

Overall, these data validate the use of the administrative case definition identified by the SCDC programs to identify adults with SCD. While there are limitations to the use of this definition, it can be highly valuable for SCD surveillance to improve the understanding of the patterns of health care usage in the SCD population over time.

Acknowledgments

The work was supported by the Sickle Cell Data Collection Program grant funded by the Centers for Disease Control and Prevention.

Conflicts of Interest

JJF receives research funding for studies on SCD through Vifor, Shire, and Forma Therapeutics. JK gives consultancy to Guide point Global, GLG, Novartis, Bluebird Bio, Ecor1, and receives honorarium from Novartis, Bausch, Glycomimetics, Oric, Sanofi

References

1. Farooq F, Mogayzel PJ, Lanzkron S, Haywood C, Strouse JJ. Comparison of US federal and foundation funding of research for sickle cell disease and cystic fibrosis and factors associated with research productivity. JAMA Netw Open 2020 Mar 02;3(3):e201737 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.1737] [Medline: 32219405]
2. National Academies of Sciences, Engineering, and Medicine. Addressing Sickle Cell Disease: A Strategic Plan and Blueprint for Action. Washington, DC: The National Academies Press; 2020.

3. Sickle Cell Data Collection (SCDC) Program. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ncbddd/hemoglobinopathies/scdc.html [accessed 2020-08-24]
4. Snyder AB, Zhou M, Theodore R, Quarmyne M, Eckman J, Lane PA. Improving an administrative case definition for longitudinal surveillance of sickle cell disease. Public Health Rep 2019 Apr 10;134(3):274-281 [FREE Full text] [doi: 10.1177/0033354919839072] [Medline: 30970223]
5. Grosse SD, Green NS, Reeves SL. Administrative data identify sickle cell disease: a critical review of approaches in U.S. health services research. Pediatr Blood Cancer 2020 Dec 17;67(12):e28703 [FREE Full text] [doi: 10.1002/pbc.28703] [Medline: 32939942]
6. Singh A, Mora J, Panepinto JA. Identification of patients with hemoglobin SS/Sβ thalassemia disease and pain crises within electronic health records. Blood Adv 2018 Jun 12;2(11):1172-1179 [FREE Full text] [doi: 10.1182/bloodadvances.2018017541] [Medline: 29792312]
7. Michalik DE, Taylor BW, Panepinto JA. Identification and validation of a sickle cell disease cohort within electronic health records. Acad Pediatr 2017 Apr;17(3):283-287. [doi: 10.1016/j.acap.2016.12.005] [Medline: 27979750]
8. Ogu U, Billett H. Comorbidities in sickle cell disease: adult providers needed!. Indian J Med Res 2018;147(6):527. [doi: 10.4103/ijmr.ijmr_1019_18]
9. Panepinto J, Owens P, Mosso A, Steiner C, Brousseau D. Concentration of hospital care for acute sickle cell disease-related visits. Pediatr Blood Cancer 2012 Oct 16;59(4):685-689 [FREE Full text] [doi: 10.1002/pbc.24028] [Medline: 22180290]
10. National Alliance for Sickle Cell Centers. URL: https://sicklecellcenters.org/ [accessed 2023-05-12]
11. DiMartino LD, Baumann AA, Hsu LL, Kanter J, Gordeuk VR, Glassberg J, Sickle Cell Disease Implementation Consortium. The sickle cell disease implementation consortium: translating evidence-based guidelines into practice for sickle cell disease. Am J Hematol 2018 Dec 17;93(12):E391-E395 [FREE Full text] [doi: 10.1002/ajh.25282] [Medline: 30203558]
12. ASH Research Collaborative. URL: https://www.ashresearchcollaborative.org/s/ [accessed 2023-05-12]
13. Kanter J, Smith W, Desai P, Treadwell M, Andemariam B, Little J, et al. Building access to care in adult sickle cell disease: defining models of care, essential components, and economic aspects. Blood Adv 2020 Aug 25;4(16):3804-3813 [FREE Full text] [doi: 10.1182/bloodadvances.2020001743] [Medline: 32785684]

Abbreviations

FP: false positive
ICD: International Classification of Diseases
NASEM: National Academies of Sciences, Engineering, and Medicine
NBS: newborn screening program
SCD: sickle cell disease
SCDC: Sickle Cell Data Collection
TP: true positive

Edited by Y Khader; submitted 20.09.22; peer-reviewed by X Li, S Gopal, F Coyne, C Edwards; comments to author 18.03.23; revised version received 03.04.23; accepted 11.04.23; published 28.06.23.

Please cite as:
Singh A, Sontag MK, Zhou M, Dasgupta M, Crume T, McLemore M, Galadanci N, Randall E, Steiner N, Brandow AM, Koch K, Field JJ, Hassell K, Snyder AB, Kanter J
Evaluating the Discriminatory Ability of the Sickle Cell Data Collection Program's Administrative Claims Case Definition in Identifying Adults With Sickle Cell Disease: Validation Study
JMIR Public Health Surveill 2023;9:e42816
URL: https://publichealth.jmir.org/2023/1/e42816
doi:10.2196/42816
PMID:37379070

Original Paper

A Developmental Surveillance Score for Quantitative Monitoring of Early Childhood Milestone Attainment: Algorithm Development and Validation

Yonatan Bilu[1*], PhD; Guy Amit[1*], PhD; Tamar Sudry[1,2], MD; Pinchas Akiva[1], PhD; Meytal Avgil Tsadok[3], PhD; Deena R Zimmerman[4], PhD, MPH; Ravit Baruch[4], MA, RN; Yair Sadaka[1,2,5], MD, MHA, PhD

[1]KI Research Institute, Kfar Malal, Israel

[2]Neuro-Developmental Research Center, Mental Health Institute, Be'er-Sheva, Israel

[3]TIMNA Inititative, Big Data Platform, Israel Ministry of Health, Jerusalem, Israel

[4]Public Health Services, Israel Ministry of Health, Jerusalem, Israel

[5]Faculty of Health Sciences, Ben-Gurion University of the Negev, Be'er-Sheva, Israel

[*]these authors contributed equally

Corresponding Author:
Yonatan Bilu, PhD
KI Research Institute
Hazayit 11
Kfar Malal, 4592000
Israel
Phone: 972 99797788
Email: yonatan@kinstitute.org.il

Abstract

Background: Developmental surveillance, conducted routinely worldwide, is fundamental for timely identification of children at risk of developmental delays. It is typically executed by assessing age-appropriate milestone attainment and applying clinical judgment during health supervision visits. Unlike developmental screening and evaluation tools, surveillance typically lacks standardized quantitative measures, and consequently, its interpretation is often qualitative and subjective.

Objective: Herein, we suggested a novel method for aggregating developmental surveillance assessments into a single score that coherently depicts and monitors child development. We described the procedure for calculating the score and demonstrated its ability to effectively capture known population-level associations. Additionally, we showed that the score can be used to describe longitudinal patterns of development that may facilitate tracking and classifying developmental trajectories of children.

Methods: We described the Developmental Surveillance Score (DSS), a simple-to-use tool that quantifies the age-dependent severity level of a failure at attaining developmental milestones based on the recently introduced Israeli developmental surveillance program. We evaluated the DSS using a nationwide cohort of >1 million Israeli children from birth to 36 months of age, assessed between July 1, 2014, and September 1, 2021. We measured the score's ability to capture known associations between developmental delays and characteristics of the mother and child. Additionally, we computed series of the DSS in consecutive visits to describe a child's longitudinal development and applied cluster analysis to identify distinct patterns of these developmental trajectories.

Results: The analyzed cohort included 1,130,005 children. The evaluation of the DSS on subpopulations of the cohort, stratified by known risk factors of developmental delays, revealed expected relations between developmental delay and characteristics of the child and mother, including demographics and obstetrics-related variables. On average, the score was worse for preterm children compared to full-term children and for male children compared to female children, and it was correspondingly worse for lower levels of maternal education. The trajectories of scores in 6 consecutive visits were available for 294,000 children. The clustering of these trajectories revealed 3 main types of developmental patterns that are consistent with clinical experience: children who successfully attain milestones, children who initially tend to fail but improve over time, and children whose failures tend to increase over time.

Conclusions: The suggested score is straightforward to compute in its basic form and can be easily implemented as a web-based tool in its more elaborate form. It highlights known and novel relations between developmental delay and characteristics of the mother and child, demonstrating its potential usefulness for surveillance and research. Additionally, it can monitor the developmental trajectory of a child and characterize it. Future work is needed to calibrate the score vis-a-vis other screening tools, validate it worldwide, and integrate it into the clinical workflow of developmental surveillance.

(JMIR Public Health Surveill 2023;9:e47315) doi:10.2196/47315

KEYWORDS

child development; risk scores; scoring methods; language delay; motor skills delay; developmental; surveillance; developmental delays; developmental milestones; young children; intervention; child

Introduction

With growing awareness to the high prevalence of developmental, behavioral, or social delay among young children, and the importance of early intervention to mitigate this risk [1-5], many international organizations have recommended routine developmental surveillance for all children [2,6,7]. This process is typically conducted by evaluating the children's ability to attain a battery of age-appropriate milestones at routine clinic visits during the first few years of their life [2]. Interpreting the results of such evaluations is not straightforward. For a specific milestone, one can establish the population's age-dependent norms of attaining the milestone and use them to assess the level of concern in case a child fails to attain it, similar to the way physical growth measures are monitored [8-10]. However, unlike physical growth norms, which are continuous and whose trajectories over time are readily understood, success or failure at attaining a developmental milestone is a binary measure, and it is not obvious how to integrate the results of multiple different milestones across several developmental domains to quantitatively monitor and assess a child's development over time.

The assessment of child development can be done at varying level of details using 3 different types of tools: surveillance (or monitoring), screening, and evaluation. Developmental surveillance is based on milestone attainment checklists and is used worldwide by pediatricians and health care providers at routine encounters, as well as by educators and parents. Screening requires a more formal and elaborated assessment, typically done by caregivers or health care professionals at specific ages. Finally, developmental evaluation is an in-depth examination, typically done by a trained specialist, which aims to provide a formal diagnosis of the child. Importantly, surveillance is based on developmental norms, whereas screening tools are validated against a "gold standard" obtained from evaluation.

A commonly used screening tool is the Denver II Screening Tool [8,11], where the outcome is either "normal" or "suspicious," based on how many milestones were failed and the general rate of failure for them. A common alternative is the Ages & Stages Questionnaires (ASQ-3) [12] screening tool, where caregivers select 1 of 3 answers for an array of questions, and the total score identifies the child's development as being "on schedule," requiring "learning activities and monitor," or needing "further assessment." Both of these screening tools take about 20 minutes to administer, depending on the age of the child and the experience of the person administering them. A widely used developmental evaluation tool is the Bayley Scales of Infant and Toddler Development [13], which typically takes 30-70 minutes to complete and yields a numerical score for each developmental category, as well as an estimate for a child's developmental age—that is, at what age do neurotypical children exhibit a similar level of milestone attainment.

Previous work [14] has attempted to combine and standardize the results of 12 commonly used screening and evaluation tools into a single metric. However, doing so for surveillance tools is more challenging. There is a lack of standardization at this level of assessment, and the quantification of developmental surveillance assessments has not been previously suggested. At best, surveillance tools are calibrated using real-world data to determine the rate of milestone attainment at different ages [9] and then administered accordingly.

In this work, we suggested a relatively simple new methodology for translating a milestone-based developmental surveillance scale into a single score, denoted as the Developmental Surveillance Score (DSS), that conveys a child's developmental status during a specified time period. Based on data from a national developmental surveillance program in Israel, we demonstrated that this score consistently captures known associations between the development and characteristics of the mother and child. Moreover, the score can be used to reveal and explore new associations, which may further improve our understanding of the factors that impact developmental delay. Finally, the score can be used to track individual children longitudinally, by describing the trajectory of their development over time. We showed that by clustering these trajectories, we can identify several typical patterns of development.

The focus of this work was on defining a straightforward surveillance score (in the sense that computing it as part of the surveillance workflow adds essentially no overhead over the current practice) and establishing its coherence and potential usefulness. Further work is required to refine this score, validate it using various data sets internationally, and derive from it explicit protocols.

Methods

Developmental Surveillance in Israel

Developmental surveillance (from birth to 6 years of age) in Israel is performed routinely (and free of charge) according to national standards by trained public health nurses in approximately 1000 maternal child health clinics (MCHCs). The collected data of approximately 70% of the Israeli population of this age group are documented in a single common database managed by the Israeli Ministry of Health. The developmental assessments include 59 milestones across 4 domains: personal-social, language, fine motor, and gross motor [9].

Parents are instructed to visit the MCHC after hospital discharge and then at ages of 1, 2, 4, 6, 9, 12, 18, 24, 36, 48, and 60 months. At each visit, a predefined group of age-related milestones is evaluated, according to the expected development at that age (denoted "age step"). Children may also be evaluated on milestones of a previous age step, in cases of a missed visit or a failure to attain milestones at the preceding visit.

The child's ability to attain each milestone is reported as observed in the clinic; although in cases of difficult attainments, this may by documented according to a parent's report. If the evaluated milestone was not attained by neither observation nor parental report, it is documented as unattained.

Study Cohort

This study included all children born between July 1, 2014, and September 1, 2021, who were followed at the MCHCs and had at least one developmental evaluation recorded during the study period. In most of the analyses, we excluded children born preterm (gestational age of <37 weeks)—the one exception is the analysis of gestational age. Additionally, children with missing gestational age were excluded, as well as visits without developmental data or without the child's age. The final cohort included 1,130,005 children in total, with 1,052,905 of them born on-term.

DSS Definition

Sudry et al [9] have recently introduced the Tipat Halav Israel Surveillance (THIS) developmental scale, a data-driven developmental scale comprising curves of attainment rate by age for each of the 59 milestones evaluated in the Israeli developmental surveillance program (the scale can be downloaded from [15]). Broadly, when a child fails to attain a milestone, the THIS developmental scale categorizes the severity of this failure into 1 of 4 categories, depending on how often children of the same age fail to attain this milestone. Accordingly, in this study, we defined the Discrete Milestone Attainment Score (DMAS) for a failed milestone as the numerical order of the failure severity: a score of 1, 2, 3, or 4 is assigned for failure occurring when <75%, 75% to 90%, 90% to 95%, or >95% of the children at the same age attain this milestone, respectively. For an attained milestone, the DMAS value is 0. If an milestones is attempted multiple times, it will be scored separately each time it is attempted. The total score for a set of milestones is the average DMAS over all milestones of all developmental domains.

More formally, for each milestone, the age thresholds for attainment by 75%, 90%, and 95% of the children were calculated [9]; we denoted these age thresholds for milestone t by t_{75}, t_{90}, and t_{95}, respectively, and considered the 4 consecutive age brackets they define:

$$b_1 = [t_0, t_{75}],\ b_2 = (t_{75}, t_{90}],\ b_3 = (t_{90}, t_{95}],\ b_4 = (t_{95}, t_{100}]$$

where t_0 and t_{100} are the minimal and maximal ages at which the milestone t is assessed, respectively.

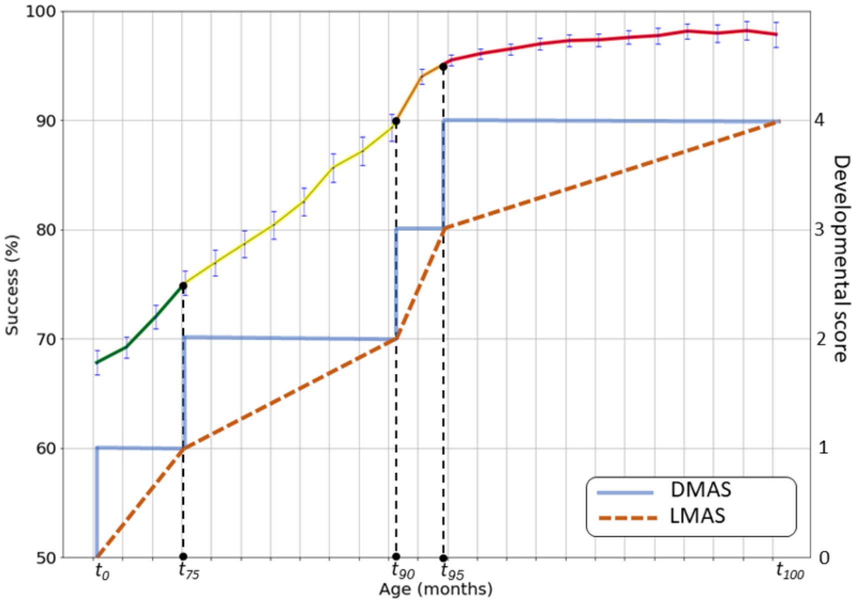
For a milestone t evaluated at age a, we defined i such that a is in the bracket b_i (i indicates the severity of failure):

To avoid noncontinuity, we extended the above definition into a Linearized Milestone Attainment Score (LMAS), using a c function as follows:

where a_{min} and a_{max} are the low and high ends of b_i, respectively.

The definitions of DMAS and LMAS are graphically illustrated in Figure 1. In the remainder of this paper, we used the LMAS version of the score, unless otherwise noted. In practice, deciding which of the 2 to use depends on the use case. DMAS is straightforward to compute from the THIS scale, whereas LMAS offers finer resolution.

Figure 1. A schematic depiction of the Discrete Milestone Attainment Score (DMAS) and the Linearized Milestone Attainment Score (LMAS) computed from the trend of milestone attainment rate versus the child's age.



For a set of milestones T, we defined the developmental surveillance score $DSS(T)$ as the average of the individual milestone attainment scores:

where a_t is the age at which milestone t was assessed. See Multimedia Appendix 1 for a concrete example of computing the score.

The set of milestones used for calculating the DSS can be defined by the evaluation period and by the types of developmental domains. For example, when computing the fine-motor score for a child during the first year of life, we computed the score for each fine-motor milestone attempted by the child during this period and then the average of the scores. In particular, if a milestone was attempted multiple times during this period, all attempts were used for the calculation of the score. Determining the evaluation period is a delicate point, which depends on the DSS application. Herein, we considered a broad period of 1 year in the subpopulation analysis and visits during each MCHC-determined age bracket (typically, a single visit) when analyzing developmental trajectories.

In this study, we aggregated personal-social milestones with language milestones, denoting them as "language-social" milestones. This was motivated by the relatively small number of milestones in the social domain and the interdependence of development in these 2 domains.

In Multimedia Appendix 1, we described an alternative score definition, the q-score, which is motivated by the notion of developmental quotient and is based on a more formal statistical approach. As described there, these 2 approaches lead to a similar ranking of children according to the quantified developmental delay, that is, when asking for which of 2 given children there is a greater concern for developmental delay, the 2 approaches tend to give the same answer.

Associations Between Mother and Child Characteristics and the Developmental Score

We examined the relations between the DSS and the characteristics of the mother and child. The children's characteristics included sex, gestational age at birth, birth weight, birth order, and records of an existing developmental tracking.

When analyzing gestational age, we partitioned preterm births to extremely preterm (less than 27 weeks), very preterm (27-31 weeks), and late preterm (32-36 weeks) [16]. This was the only analysis that included preterm children.

Characteristics of mothers included age at delivery; level of education; and the result of postpartum depression (PPD) evaluation, using the Edinburgh Postnatal Depression Scale (EPDS). For the purpose of the analysis, mothers were considered as having symptoms of PPD if their EPDS score was ≥10 or if their score in question number 10 (self-harm) was other than 0 [17].

To test whether differences between score averages were significant, we used the Mann-Whitney *U* test [18].

Developmental Trajectory Vectors

We described the developmental trajectory of a child by the series of its DSS values at each age step from birth to 36 months of age. Each age step s has an associated set of milestones $T(s)$. We further partitioned the milestones by their developmental domains, denoting by $T(s, d)$ the subset of $T(s)$ from domain d (where d can be either "language-social" or "motor"—an aggregation of fine-motor and gross-motor milestones). This allowed us to describe the trajectory per domain as the Developmental Trajectory Vectors (DTVs):

DTV(d) = DSS(T(s1,d),…, DSS(T(s7d)

where s_i goes over the steps of 1-3 months, 3-6 months, 6-9 months, 9-12 months, 12-18 months, 18-24 months, and 24-36 months.

This representation yielded DTVs of length 7 for each child that was assessed at all age steps. For this analysis, we excluded children whose data was missing for 1 or more age steps, analyzing the remaining groups of 294,624 and 294,066 children in the motor and language-social domains, respectively.

DTVs Clustering

We used the k-means clustering [19] to identify distinct patterns of DTVs. In addition, for sensitivity analysis of the clustering method (see Multimedia Appendix 1), we examined an alternative clustering method using a Gaussian Mixture Model [20]. Cluster validity was assessed using the Calinski-Harabasz score [21] (see Multimedia Appendix 1).

The clustering was done using only 6 of the 7 DTV entries. This is because for each domain, there is one step that included only a single milestone (for motor milestones, the 12-18 months step; for language-social milestones, the 6-9 months step), which may reduce the reliability of the results. Nonetheless, when computing cluster centroids, all entries were taken into account.

Analyses were done using Python (version 3.6.7; Python Software Foundation) with the *scikit-learn* package (version 0.23.2).

Ethics Approval

The study protocol was approved by the Soroka University Medical Center institutional ethical committee (MHC-0014-19) and was conducted in accordance with the principles of the Declaration of Helsinki. The need for informed consent was waived owing to the use of deidentified data.

Results

DSS of Different Population Subgroups

Table 1 shows the main characteristics of the children in the study cohort, grouped by their age at the time of the visit at the MCHC. It was evident that the number of children who visit the MCHC decreased with the child's age (880,688/1,052,905, 83.6% of the cohort visited at 0-12 months of age, whereas only 635,009/1,052,905, 60.3% visited at 24-36 months of age).

Table 1. Number of children with recorded developmental surveillance from the Israeli Ministry of Health, between July 2014 and September 2021, according to age group and stratified by child and mother characteristics. Some categories do not sum up to the total number due to missing values. Children for which the value of some characteristic is missing are not counted toward the tallies of that characteristic. Preterm children were not included in analysis, except for the analysis on gestational age.

Characteristic	Children aged 0-12 months (n=880,688), n (%)	Children aged 12-24 months (n=805,231), n (%)	Children aged 24-36 months (n=635,009), n (%)
Developmental tracking			
Tracked	8595 (1)	13,711 (1.7)	13,508 (2.1)
Not tracked	842,307 (95.6)	765,075 (95)	597,559 (94.1)
Sex			
Female	429,695 (48.8)	393,115 (48.8)	309,995 (48.8)
Male	450,993 (51.2)	412,116 (51.2)	325,014 (51.2)
Postpartum depression			
Positive	27,791 (3.2)	23,719 (2.9)	17,520 (2.8)
Negative	685,335 (77.8)	567,633 (70.5)	399,514 (62.9)
Mother's age (years)			
18-39	739,073 (83.9)	686,705 (85.3)	549,114 (86.5)
40-50	130,270 (14.8)	110,398 (13.7)	81,249 (12.8)
Birth weight (kg)			
1-2.5	30,036 (3.4)	27,761 (3.4)	22,298 (3.5)
2.5-3	193,291 (21.9)	178,017 (22.1)	142,000 (22.4)
3-3.5	393,690 (44.7)	360,481 (44.8)	284,226 (44.8)
3.5-4	226,507 (25.7)	205,010 (25.5)	159,857 (25.2)
4-4.5	44,687 (5.1)	40,366 (5)	31,412 (4.9)
4.5-6	3929 (0.4)	3580 (0.4)	2778 (0.4)
Child number			
1	316,738 (36)	310,973 (38.6)	274,206 (43.2)
2	289,212 (32.8)	274,439 (34.1)	221,029 (34.8)
3	167,647 (19)	142,394 (17.7)	97,590 (15.4)
Mothers' education			
Academic	267,199 (30.3)	246,190 (30.6)	198,099 (31.2)
Tertiary education	91,060 (10.3)	79,979 (9.9)	59,102 (9.3)
High school	233,613 (26.5)	217,952 (27.1)	176,831 (27.8)
Elementary	18,651 (2.1)	17,273 (2.1)	13,974 (2.2)
Gestational age (weeks; total includes preterm children: aged 0-12 months, n=943,354; aged 12-24 months, n=864,421; and aged 24-36 months, n=682,999)			
23-27	1410 (0.1)	1387 (0.2)	1134 (0.2)
28-31	5135 (0.5)	4912 (0.6)	4090 (0.6)
32-36	56,934 (6)	53,634 (6.2)	43,346 (6.3)
37-38	228,343 (24.2)	209,255 (24.2)	166,043 (24.3)
39-42	651,532 (69)	595,233 (68.8)	468,386 (68.5)

To assess the relations between the DSS and characteristics of the children or their mothers, we compared, for each domain, the average DSS of several subgroups during the first, second, and third years of life. Figure 2 shows that the average DSS was higher (worse) for children that were under designated developmental tracking, compared to the complementary group (Figure 2A). Higher DSS was evident in the following subgroups: male children (Figure 2B), children whose mothers reported symptoms of PPD (Figure 2C), and children of older mothers (Figure 2D).

Figure 2. Developmental surveillance score (DSS) of binary variables. (A) Children under developmental (dev.) tracking compared to those who are not; (B) female children compared to male children; (C) children whose mothers reported postpartum depression (PPD) symptoms compared to those who did not; and (D) children of younger mothers (aged 18-39 years) compared to older mothers (aged 40-50 years). Asterisks denote a statistically significant difference between a pair of bars (**P*<.05; ***P*<.01; ****P*<.001). Based on developmental surveillance data from the Israeli Ministry of Health, between July 2014 and September 2021.

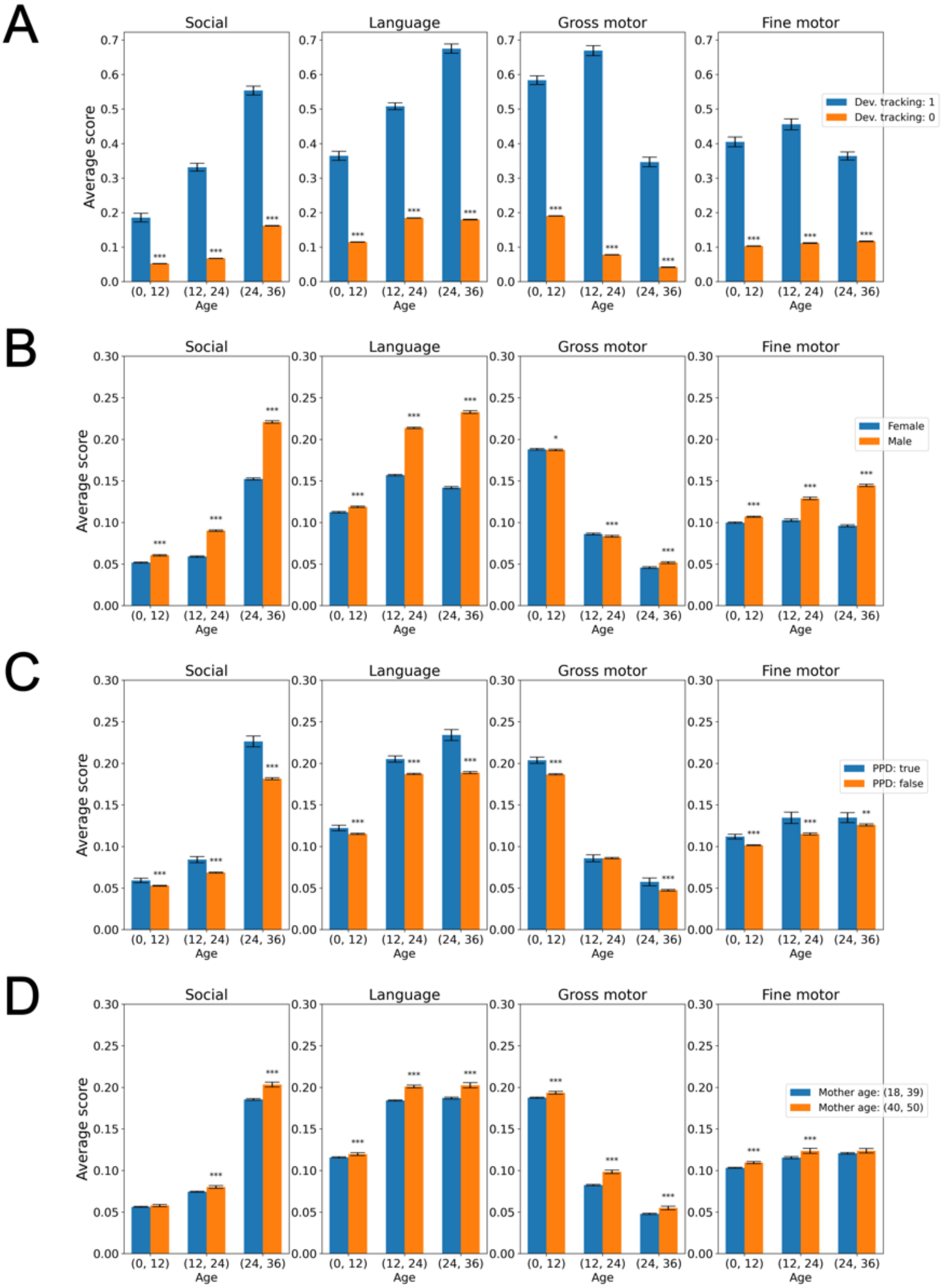


Figure 3A demonstrates the relation between the DSS and birth weight: children with birth weight of <2.5 kg or >4.5 kg had higher average DSS than children with normative birth weight (2.5-4.5 kg). Figure 3B shows that the DSS was negatively correlated to the gestational age at birth (eg, in the first year of life, Pearson *r*=–0.2 for gross motor milestones, –0.25 for fine motor milestones, and –0.18 for language-social milestones; *P*<.001). There were marked differences between preterm and on-term children, as well as between subgroups of extremely preterm, very preterm, moderate preterm, early term, and full-term children.

Figure 3. Relation between Developmental Surveillance Score (DSS) and numeric birth variables. (A) The child's birth weight and (B) gestational (gest.) age at birth. Asterisks denote a statistically significant difference between a pair of consecutive bars (*P<.05; **P<.01; ***P<.001). Based on developmental surveillance data from the Israeli Ministry of Health, between July 2014 and September 2021.

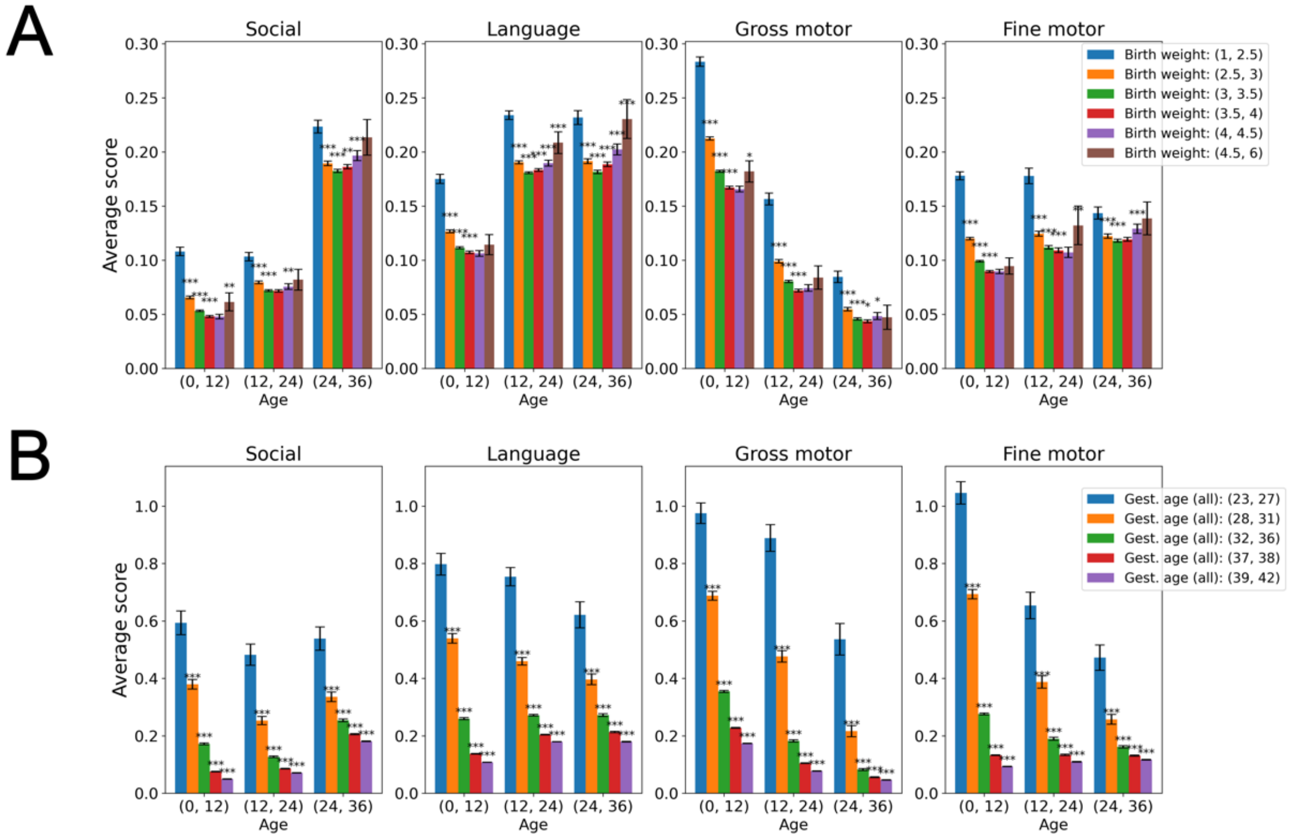
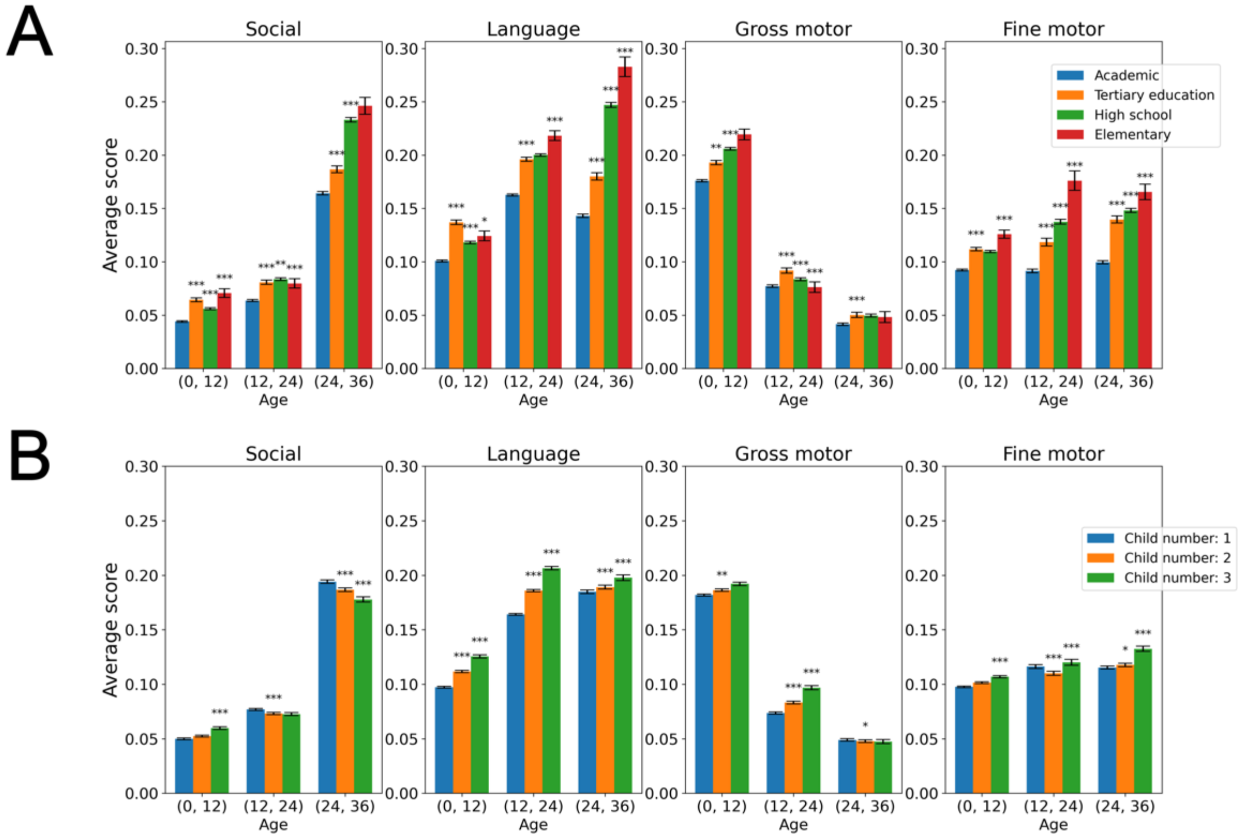


Figure 4A shows the association between the DSS and the mothers' level of education. The DSS tended to be higher among mothers with less formal education. In addition, the score appeared to be positively correlated with the child's birth order during the first year of life (Figure 4B; Pearson *r*=0.02 for gross motor milestones, 0.03 for fine motor milestones, and 0.08 for language-social milestones; *P*<.001), with firstborn children having the least average score. This trend was maintained for the gross motor and language-social scores during the second year of life (Pearson *r*=0.03 for gross motor milestones, 0.01 for fine motor milestones, and 0.07 for language-social milestones; *P*<.001). Conversely, this correlation was evident during the third year of life only for fine-motor tasks (*r*=0.02; *P*<.001). Importantly, these correlations should be considered as affirmation for the trends suggested by the graphs—their relatively low values on these large cohorts certainly do not imply that the DSS "explains" in any way the measured characteristics.

Figure 4. Relation between Developmental Surveillance Score (DSS) and categorical variables. (A) Maternal education level and (B) child's birth order. Asterisks denote a statistically significant difference between a pair of consecutive bars (*P<.05; **P<.01; ***P<.001). Based on developmental surveillance data from the Israeli Ministry of Health, between July 2014 and September 2021.

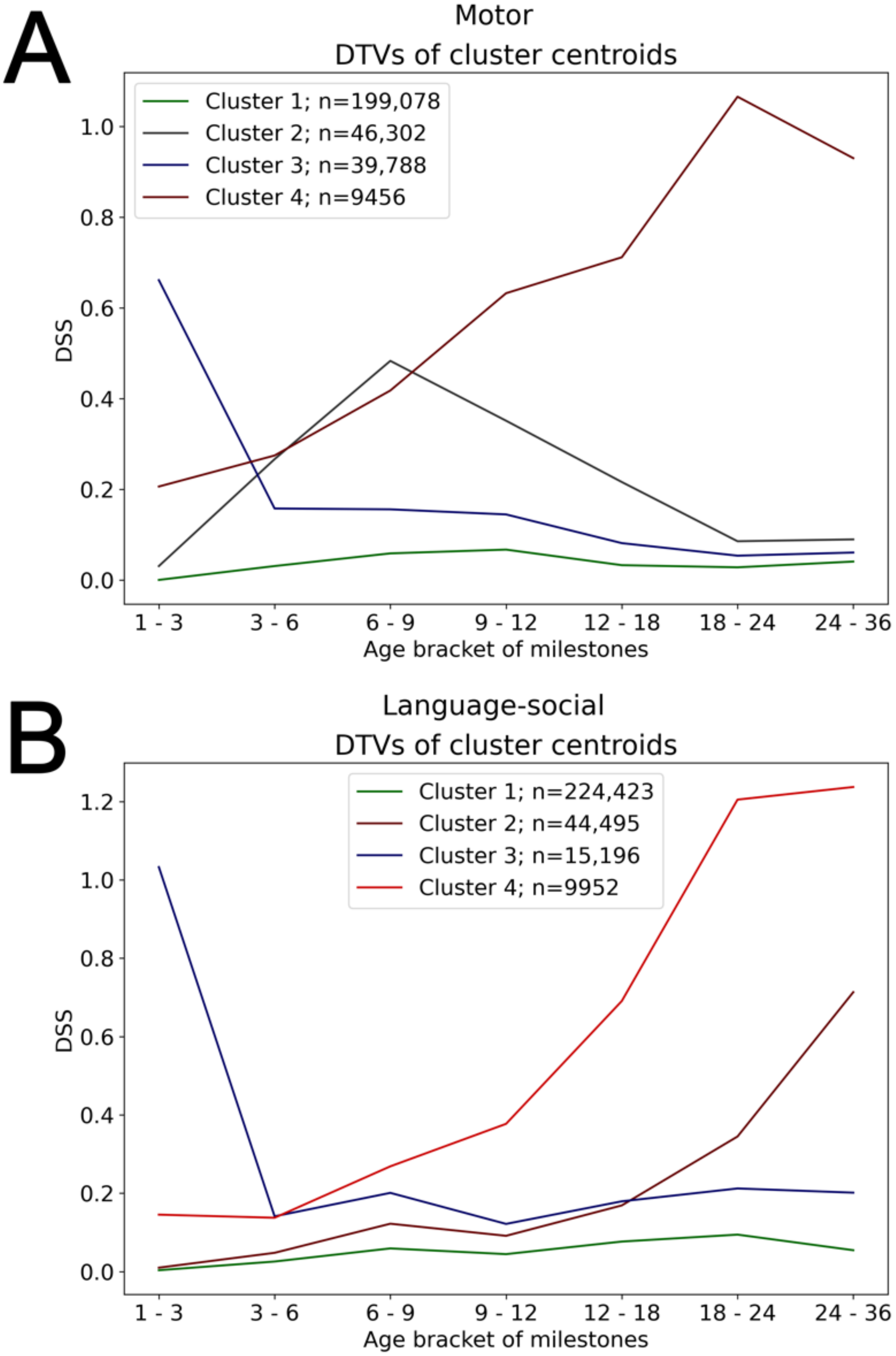


Note that all these graphs depict average values. For the most part, children attained the assessed milestones and received a score of 0. See Table S1 in Multimedia Appendix 1 for the median and IQR values of the DSS and Figures S1-S3 in Multimedia Appendix 1 for the same analysis using DMAS instead of LMAS.

Using the DSS to Describe Children's Developmental Trajectories

Figure 5 depicts the centroids derived from clustering of all children's DTVs into 4 clusters. Both motor DTVs and language-social DTVs exhibited similar patterns. There was a single cluster of children with near-zero DSS at all age steps. This cluster included the majority of children ("adequate"; motor DTVs: 199,078/294,624, 67.6%; language-social DTVs: 224,423/294,066, 76.3%). There was a single cluster of children who were "catching up"—their DSS was initially high but tended to decrease over time. There were clusters of "worsening" children whose scores tended to increase over time (2 clusters for language-social milestones and 1 for motor milestones). For motor milestones, there was also a cluster of children whose DSS increased at an early age but then decreased back to normal values and, so, did not conform to any of these 3 patterns.

Figure 5. Centroids of motor Developmental Trajectory Vectors (DTV) clusters. (A) Scores derived from fine motor and gross motor milestones and (B) scores derived from language and social milestones. Centroids of clusters are (A): 1. [0.01, 0.03, 0.06, 0.07, 0.03, 0.03, 0.04]; 2. [0.03, 0.27, 0.48, 0.35, 0.22, 0.08, 0.09]; 3. [0.66, 0.16, 0.16, 0.14, 0.05, 0.05, 0.06]; 4. [0.21, 0.27, 0.42, 0.63, 0.71, 1.07, 0.93]; and (B): 1. [0, 0.03, 0.04, 0.04, 0.08, 0.09, 0.05]; 2. [0.01, 0.05, 0.11, 0.09, 0.17, 0.35, 0.71]; 3. [1.03, 0.14, 0.15, 0.12, 0.18, 0.21, 0.20]; 4. [0.15, 0.14, 0.26, 0.38, 0.69, 1.20, 1.24]. Based on developmental surveillance data from the Israeli Ministry of Health, between July 2014 and September 2021. DSS: Developmental Surveillance Score.



Tables 2 and 3 show the distributions of child and mother's characteristics among the 4 different clusters in each domain. In the clusters depicting an increasing trajectory, there was an overrepresentation of male children relative to the "adequate" cluster. Specifically, in the motor domain, 50.7% (100,901/199,078) of the children in the "adequate" cluster were male, compared to 57.8% (5466/9456) in the "worsening" cluster. In the language-social domain, male children were 48.1% (107,970/224,423) of those in the "adequate" cluster, compared to 71.3% (7093/9952) and 62% (27,566/44,495) in the rapidly “worsening” and moderately “worsening” clusters, respectively. In addition, the “worsening” clusters had larger proportions of children that were born by cesarean section, had low birth weight, or were under developmental tracking.

Table 2. Distribution of child and mother characteristics in motor milestones clusters. Based on developmental surveillance data (N=294,624) from the Israeli Ministry of Health, between July 2014 and September 2021.

Characteristic	Cluster 1 (“adequate”; n=199,078, 67.6%), n (%)	Cluster 2 (n=46,302, 15.7%)	Cluster 3 (“catching up”; n=39,788, 13.5%), n (%)	Cluster 4 (“worsening”; n=9456, 3.2%), n (%)
Developmental tracking				
Tracked	4310 (2.2)	3200 (6.9)	1853 (4.7)	2471 (26.1)
Not tracked	194,768 (97.8)	43,102 (93.1)	37,935 (95.3)	6985 (73.9)
Sex				
Female	98,177 (49.3)	22,783 (49.2)	19,108 (48)	3990 (42.2)
Male	100,901 (50.7)	23,519 (50.8)	20,680 (52)	5466 (57.8)
Postpartum depression				
Positive	7638 (3.8)	1943 (4.2)	1647 (4.1)	436 (4.6)
Negative	171,034 (85.9)	38,357 (82.8)	33,626 (84.5)	7887 (83.4)
Mother’s age (years)				
18-39	170,833 (85.8)	39,633 (85.6)	33,539 (84.3)	7837 (82.9)
40-50	26,985 (13.6)	6483 (14)	6065 (15.2)	1561 (16.5)
Birth weight (kg)				
1-2.5	6826 (3.4)	2977 (6.4)	2252 (5.7)	816 (8.6)
2.5-3	43,905 (22.1)	12,114 (26.2)	10,301 (25.9)	2637 (27.9)
3-3.5	89,570 (45)	19,910 (43)	17,283 (43.4)	3797 (40.2)
3.5-4	50,903 (25.6)	9864 (21.3)	8625 (21.7)	1912 (20.2)
4-4.5	9610 (4.8)	1845 (4)	1659 (4.2)	375 (4)
4.5-6	790 (0.4)	173 (0.4)	139 (0.3)	43 (0.5)
Child number				
1	88,383 (44.4)	19,530 (42.2)	17,986 (45.2)	4003 (42.3)
2	70,331 (35.3)	16,363 (35.3)	13,379 (33.6)	3184 (33.7)
3	30,518 (15.3)	7640 (16.5)	6129 (15.4)	1601 (16.9)
Mother’s education				
Academic	72,371 (36.4)	15,419 (33.3)	13,242 (33.3)	2659 (28.1)
Tertiary education	15,538 (7.8)	3995 (8.6)	3213 (8.1)	807 (8.5)
High school	57,749 (29)	14,980 (32.4)	12,127 (30.5)	3206 (33.9)
Elementary	4124 (2.1)	1194 (2.6)	1023 (2.6)	239 (2.5)

Table 3. Distribution of child and mother characteristics in language-social milestones clusters. Based on developmental surveillance data (N=294,066) from the Israeli Ministry of Health, between July 2014 and September 2021.

Characteristic	Cluster 1 ("adequate"; n=224,423, 76.3%), n (%)	Cluster 2 ("worsening"; n=44,495, 15.1%), n (%)	Cluster 3 ("catching up"; n=15,196, 5.2%), n (%)	Cluster 4 ("worsening"; n=9952, 3.4%), n (%)
Developmental tracking				
Tracked	5225 (2.3)	3349 (7.5)	678 (4.5)	2536 (25.5)
Not tracked	219,198 (97.7)	41,146 (92.5)	14,518 (95.5)	7416 (74.5)
Sex				
Female	116,453 (51.9)	16,929 (38)	7540 (49.6)	2859 (28.7)
Male	107,970 (48.1)	27,566 (62)	7656 (50.4)	7093 (71.3)
Postpartum depression				
Positive	8368 (3.7)	2048 (4.6)	688 (4.5)	563 (5.7)
Negative	191,397 (85.3)	37,869 (85.1)	12,814 (84.3)	8341 (83.8)
Mother's age (years)				
18-39	192,242 (85.7)	37,812 (85)	13,115 (86.3)	8199 (82.4)
40-50	30,837 (13.7)	6473 (14.5)	2011 (13.2)	1690 (17)
Birth weight (kg)				
1-2.5	8866 (4)	1922 (4.3)	1382 (9.1)	690 (6.9)
2.5-3	51,919 (23.1)	10,117 (22.7)	4345 (28.6)	2483 (24.9)
3-3.5	100,692 (44.9)	19,246 (43.3)	6252 (41.1)	4109 (41.3)
3.5-4	54,809 (24.4)	11,185 (25.1)	2854 (18.8)	2258 (22.7)
4-4.5	10,176 (4.5)	2295 (5.2)	497 (3.3)	489 (4.9)
4.5-6	828 (0.4)	219 (0.5)	48 (0.3)	50 (0.5)
Child number				
1	101,311 (45.1)	18,400 (41.4)	5972 (39.3)	4245 (42.7)
2	78,671 (35.1)	16,053 (36.1)	5059 (33.3)	3318 (33.3)
3	33,612 (15)	7590 (17.1)	2735 (18)	1695 (17)
Mother's education				
Academic	84,575 (37.7)	12,518 (28.1)	4378 (28.8)	2062 (20.7)
Tertiary education	17,411 (7.8)	3788 (8.5)	1411 (9.3)	849 (8.5)
High school	62,599 (27.9)	16,805 (37.8)	4574 (30.1)	4092 (41.1)
Elementary	4426 (2)	1359 (3.1)	450 (3)	371 (3.7)

In Multimedia Appendix 1, we demonstrate that qualitatively, these results were consistent over different range of clusters number, as well as when using an alternative clustering method.

Discussion

The goal of this study was to construct a DSS that can be used for comparative tracking of children's development, quantifying milestones attainment in a concise and straightforward way. We presented a simple methodology for calculating the DSS, a quantitative developmental surveillance score that aggregates age-dependent milestones results over a chosen time frame and domain into a single score. To demonstrate its coherence, we explored 2 main use cases for this score: comparing its value among subpopulations and using it to depict the developmental trajectory of individuals. We demonstrated that the DSS reflects known associations between developmental status and characteristics of the child and mother and its potential for suggesting possible new associations and insights, which may be a stepping stone for further research.

Children who have been referred to developmental tracking, indicating that they are likely to exhibit a developmental delay, had on average a much higher score than their counterparts, at all 3 examined age groups and for all 3 developmental domains. In addition, the score was shown to reflect previously reported associations between developmental status and the child's sex [22-28], birth weight [27,29-32], gestational age [30-34], birth order [32,35-37], maternal age [31,38,39], maternal education [28,31,32,38], and maternal symptoms of PPD [28,40-43].

For some of these variables, the DSS suggests a possible association with developmental delays, depicting different score distributions among subgroups stratified by the variable, even within the normal range. For example, it is well established that low birth weight is associated with developmental delays [27,29-32], yet the results herein suggest that this may also be true for birth weight within the lower normal range (2.5-3 kg) and for birth weight above the normal range (more than 4.5 kg). Similarly, although the scores of preterm children are higher than full-term children, there is a gradual decrease in the average score by the level of prematurity (extreme preterm, very preterm, and late preterm children), as well as a difference between early term and full-term children.

At the same time, some characteristics show a more complex behavior; for example, the DSS tends to be positively correlated with the child's order, yet for language-social tasks evaluated at 24-36 months of age, the correlation becomes negative. Indeed, although previous work generally associate primipara with lower risk for development delay [32,35,36], Oshima-Takane et al [37], who focused on language development at 21 and 24 months of age, observed higher language skills among second-born children.

Cluster analysis consistently identified 3 types of developmental trajectories: 1 cluster of children who succeed in attaining nearly all milestones, containing most of the children; 1 cluster of children who tend to fail early-age milestones but show improvement over time and succeed in attaining later milestones; and 1 or more clusters of children whose performance grow worse over time, with different clusters depicting different severities of failures. These clusters correspond to common types of developmental patterns observed in clinical practice; although, importantly, not all clusters can be categorized as 1 of these 3 types. Future work may use these clusters as class labels, in an attempt to predict the developmental trajectory type of a child at an early age and, accordingly, consider timely intervention when needed.

This work has several limitations. Importantly, the main goal was to present the DSS and show that it is consistent with current knowledge on risk factors for developmental delay such as low birth weight, preterm birth, older maternal age, symptoms of PPD, or lower level of maternal education, as well as to suggest interesting new observations. It is not proposed as a screening tool, and although we demonstrated its rationale and coherence, we lacked a "ground truth" of developmental delay for validating the score against. Future work should aim to assess the score's potential contribution to the clinical workflow of developmental assessment, for example, by comparing it to developmental screening tools such as the Bayley [13] and Denver [8,11] scales, as well as to developmental outcomes beyond those in the current data set, such as a diagnosis of autism.

Such a comparison is also needed for the calibration of the method with respect to milestones and age windows used to derive the score. For example, deriving the score by averaging milestone attainment during a full year implicitly assumes that a single number can represent the developmental delay over this entire period. Conversely, calculating a new score per visit does not take into account valuable information from past evaluations.

Another limitation stems from the use of slightly different cohorts for each age group. As depicted in Table 1, the cohorts differ in size and some of the characteristics, which may introduce some bias to the comparisons of age groups. However, as most of the presented results compare stratified population groups, the existence of similar differences in each age group strengthens the derived observations.

The results described herein pertain to the milestones used in Israeli MCHCs and the age thresholds computed in the THIS developmental scale [9]. Generalizing these results to other settings can be done by adopting the same methodology but would require having, or constructing, a developmental scale that is suitable for that setting. With such a scale at hand, one can compute a DSS from milestone attainment data by comparing them to the age thresholds and defining the score accordingly.

Taken together, our results suggest the potential usefulness of incorporating the DSS into the developmental surveillance workflow. We envision it as being computed automatically once a child's electronic health record is updated with new milestone attainment results and compared to the child's trajectory of past achievements, as well as to the population's norm. In cases where the score deviates significantly on either count, the system would notify the nurse, possibly leading to a more thorough evaluation. When calibrated correctly, such a system could identify developmental delays in a timely manner and foster interventions for improving the prospective outcomes.

Acknowledgments

The data were extracted with the help of TIMNA—a national research platform established by the Israeli government to enable big-data studies combining deidentified health data from multiple organizations.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The deidentified patient-level data used for this study contain sensitive information and therefore are not available outside the secured research environment of the Israeli Ministry of Health. Summary aggregate-level data and analysis code for this study can be made available upon reasonable request to the corresponding author.

Conflicts of Interest

YS is a shareholder in LinkCaring LTD. LinkCaring develops intelligent decision-support systems that gather medical information using smartphones, tablets, computers, wearables, and sensors; analyze this information; and generate smart decision support reports for children. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Supplemental materials and results.
[DOCX File , 309 KB - publichealth_v9i1e47315_app1.docx]

References

1. Guevara JP, Gerdes M, Localio R, Huang YV, Pinto-Martin J, Minkovitz CS, et al. Effectiveness of developmental screening in an urban setting. Pediatrics 2013 Jan;131(1):30-37. [doi: 10.1542/peds.2012-0765] [Medline: 23248223]
2. Lipkin PH, Macias MM, Council on Children With Disabilities, Section on Developmental and Behavioral Pediatrics. Promoting optimal development: identifying infants and young children with developmental disorders through developmental surveillance and screening. Pediatrics 2020 Jan 1;145(1). [doi: 10.1542/peds.2019-3449] [Medline: 31843861]
3. Hirai AH, Kogan MD, Kandasamy V, Reuland C, Bethell C. Prevalence and variation of developmental screening and surveillance in early childhood. JAMA Pediatrics 2018 Sep 01;172(9):857-866 [FREE Full text] [doi: 10.1001/jamapediatrics.2018.1524] [Medline: 29987317]
4. Barger B, Rice C, Wolf R, Roach A. Better together: developmental screening and monitoring best identify children who need early intervention. Disability Health Journal 2018 Jul;11(3):420-426 [FREE Full text] [doi: 10.1016/j.dhjo.2018.01.002] [Medline: 29459217]
5. Reichow B, Hume K, Barton EE, Boyd BA. Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). Cochrane Database Syst Rev 2018 May 09;5(5):CD009260 [FREE Full text] [doi: 10.1002/14651858.CD009260.pub3] [Medline: 29742275]
6. Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics 2006 Jul;118(1):405-420. [doi: 10.1542/peds.2006-1231] [Medline: 16818591]
7. Hagan JF, Shaw JS, Duncan RM. Bright futures: guidelines for health supervision of infants, children, and adolescents. Vanderbilt University. Elk Grove Village, IL: American Academy of Pediatrics; 2017. URL: https://catalog.library.vanderbilt.edu/discovery/fulldisplay/alma991005178939703276/01VAN_INST:BIOMED [accessed 2022-03-28]
8. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics 1992 Jan;89(1):91-97. [Medline: 1370185]
9. Sudry T, Zimmerman DR, Yardeni H, Joseph A, Baruch R, Grotto I, et al. Standardization of a Developmental Milestone Scale using data from children in Israel. JAMA Network Open 2022 Mar 01;5(3):e222184 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.2184] [Medline: 35285917]
10. Sheldrick RC, Perrin EC. Evidence-based milestones for surveillance of cognitive, language, and motor development. Acad Pediatr 2013 Nov;13(6):577-586 [FREE Full text] [doi: 10.1016/j.acap.2013.07.001] [Medline: 24238685]
11. Frankenburg WK, Dodds JB. The Denver Developmental Screening Test. J Pediatr 1967 Aug;71(2):181-191. [doi: 10.1016/s0022-3476(67)80070-2] [Medline: 6029467]
12. Squires JK, Bricker D. Ages & Stages Questionnaires. Questionnaires Set. 3rd ed. Baltimore, MD: Brookes; 2009.
13. Aylward GP, Zhu J. The Bayley Scales: clarification for clinicians and researchers. Pearson Assessments. 2019. URL: https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/bayley-4/bayley-4-technical-report.pdf [accessed 2023-08-08]
14. Weber AM, Rubio-Codina M, Walker SP, van Buuren S, Eekhout I, Grantham-McGregor SM, Working group membersdata contributors. The D-score: a metric for interpreting the early development of infants and toddlers across global settings. BMJ Glob Health 2019 Nov 19;4(6):e001724 [FREE Full text] [doi: 10.1136/bmjgh-2019-001724] [Medline: 31803508]
15. The Israeli development scale for ages 0-5 years. KI Institute. 2022. URL: https://kinstitute.org.il/publication_files/the-israeli-developmental-scale-this-scale-for-ages-0-5-years/ [accessed 2023-08-08]
16. Quinn JA, Munoz FM, Gonik B, Frau L, Cutland C, Mallett-Moore T, Brighton Collaboration Preterm Birth Working Group. Vaccine 2016 Dec 01;34(49):6047-6056 [FREE Full text] [doi: 10.1016/j.vaccine.2016.03.045] [Medline: 27743648]
17. Murray L, Carothers AD. The validation of the Edinburgh Post-natal Depression Scale on a community sample. Br J Psychiatry 1990 Aug 2;157(2):288-290. [doi: 10.1192/bjp.157.2.288] [Medline: 2224383]
18. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Statist 1947 Mar;18(1):50-60. [doi: 10.1214/aoms/1177730491]
19. Lloyd S. Least squares quantization in PCM. IEEE Trans Inform Theory 1982 Mar;28(2):129-137. [doi: 10.1109/tit.1982.1056489]

20. Bishop CM. attern Recognition and Machine Learning. Softcover reprint of the original 1st edition 2006 (corrected at 8th printing 2009). New York, NY: Springer New York; 2016.
21. Calinski T, Harabasz J. A dendrite method for cluster analysis. Comm in Stats - Simulation & Comp 1974;3(1):1-27. [doi: 10.1080/03610917408548446]
22. Hyde JS, Linn MC. Gender differences in verbal ability: a meta-analysis. Psychol Bull 1988 Jul;104(1):53-69. [doi: 10.1037/0033-2909.104.1.53]
23. Rinaldi P, Pasqualetti P, Volterra V, Caselli MC. Gender differences in early stages of language development. some evidence and possible explanations. J Neurosci Res 2023 May;101(5):643-653. [doi: 10.1002/jnr.24914] [Medline: 34240751]
24. Dinkel D, Snyder K. Exploring gender differences in infant motor development related to parent's promotion of play. Infant Behav Dev 2020 May;59:101440. [doi: 10.1016/j.infbeh.2020.101440] [Medline: 32240875]
25. Escolano-Pérez E, Sánchez-López CR, Herrero-Nivela ML. Early environmental and biological influences on preschool motor skills: implications for early childhood care and education. Front Psychol 2021 Aug 13;12:725832 [FREE Full text] [doi: 10.3389/fpsyg.2021.725832] [Medline: 34484085]
26. WHO Multicentre Growth Reference Study Group. Assessment of sex differences and heterogeneity in motor milestone attainment among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl 2006 Apr;450:66-75. [doi: 10.1111/j.1651-2227.2006.tb02377.x] [Medline: 16817680]
27. To T, Guttmann A, Dick PT, Rosenfield JD, Parkin PC, Cao H, et al. What factors are associated with poor developmental attainment in young Canadian children? Can J Public Health 2004 Jul;95(4):258-263 [FREE Full text] [doi: 10.1007/BF03405127] [Medline: 15362466]
28. To T, Guttmann A, Dick PT, Rosenfield JD, Parkin PC, Tassoudji M, et al. Risk markers for poor developmental attainment in young children: results from a longitudinal national survey. Arch Pediatr Adolesc Med 2004 Jul 01;158(7):643-649. [doi: 10.1001/archpedi.158.7.643] [Medline: 15237063]
29. Lima MC, Eickmann SH, Lima ACV, Guerra MQ, Lira PIC, Huttly SRA, et al. Determinants of mental and motor development at 12 months in a low income population: a cohort study in northeast Brazil. Acta Paediatr 2004 Jul;93(7):969-975. [doi: 10.1111/j.1651-2227.2004.tb18257.x] [Medline: 15303815]
30. Drozd-Dąbrowska M, Trusewicz R, Ganczak M. Selected risk factors of developmental delay in Polish infants: a case-control study. Int J Environ Res Public Health 2018 Dec 02;15(12):2715 [FREE Full text] [doi: 10.3390/ijerph15122715] [Medline: 30513831]
31. Ozkan M, Senel S, Arslan EA, Karacan CD. The socioeconomic and biological risk factors for developmental delay in early childhood. Eur J Pediatr 2012 Dec 15;171(12):1815-1821. [doi: 10.1007/s00431-012-1826-1] [Medline: 22983025]
32. Hediger ML, Overpeck MD, Ruan WJ, Troendle JF. Birthweight and gestational age effects on motor and social development. Paediatr Perinat Epidemiol 2002 Jan;16(1):33-46. [doi: 10.1046/j.1365-3016.2002.00393.x] [Medline: 11856453]
33. Hochstedler KA, Bell G, Park H, Ghassabian A, Bell EM, Sundaram R, et al. Gestational age at birth and risk of developmental delay: the Upstate KIDS study. Am J Perinatol 2021 Aug 06;38(10):1088-1095 [FREE Full text] [doi: 10.1055/s-0040-1702937] [Medline: 32143225]
34. Schonhaut L, Armijo I, Pérez M. Gestational age and developmental risk in moderately and late preterm and early term infants. Pediatrics 2015 Apr;135(4):e835-e841. [doi: 10.1542/peds.2014-1957] [Medline: 25733752]
35. Gayraud F, Kern S. Influence of preterm birth on early lexical and grammatical acquisition. First Language 2007 May;27(2):159-173. [doi: 10.1177/0142723706075790]
36. Alvares GA, Licari MK, Stevenson PG, Bebbington K, Cooper MN, Glasson EJ, et al. Investigating associations between birth order and autism diagnostic phenotypes. J Child Psychol Psychiatry 2021 Aug 08;62(8):961-970. [doi: 10.1111/jcpp.13349] [Medline: 33164221]
37. Oshima-Takane Y, Goodz E, Derevensky JL. Birth order effects on early language development: do secondborn children learn from overheard speech? Child Development 1996 Apr;67(2):621-634. [doi: 10.2307/1131836]
38. Demirci A, Kartal M. Sociocultural risk factors for developmental delay in children aged 3-60 months: a nested case-control study. Eur J Pediatr 2018 May 6;177(5):691-697. [doi: 10.1007/s00431-018-3109-y] [Medline: 29404718]
39. Geronimus AT, Korenman S, Hillemeier MM. Does young maternal age adversely affect child development? evidence from cousin comparisons in the United States. Popul Dev Rev 1994 Sep;20(3):585-609. [doi: 10.2307/2137602]
40. Murray L, Cooper PJ. Effects of postnatal depression on infant development. Arch Dis Child 1997 Aug 01;77(2):99-101 [FREE Full text] [doi: 10.1136/adc.77.2.99] [Medline: 9301345]
41. Lubotzky-Gete S, Ornoy A, Grotto I, Calderon-Margalit R. Postpartum depression and infant development up to 24 months: a nationwide population-based study. J Affect Disord 2021 Apr 15;285:136-143. [doi: 10.1016/j.jad.2021.02.042] [Medline: 33647581]
42. Grace SL, Evindar A, Stewart DE. The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature. Arch Womens Ment Health 2003 Nov;6(4):263-274. [doi: 10.1007/s00737-003-0024-6] [Medline: 14628179]
43. Deave T, Heron J, Evans J, Emond A. The impact of maternal depression in pregnancy on early child development. BJOG 2008 Jul;115(8):1043-1051. [doi: 10.1111/j.1471-0528.2008.01752.x] [Medline: 18651886]

Abbreviations

ASQ-3: Ages & Stages Questionnaires
DMAS: Discrete Milestone Attainment Score
DSS: Developmental Surveillance Score
DTV: Developmental Trajectory Vector
EPDS: Edinburgh Postnatal Depression Scale
LMAS: Linearized Milestone Attainment Score
MCHC: maternal child health clinic
PPD: postpartum depression
THIS: Tipat Halav Israel Surveillance

Edited by A Mavragani, T Sanchez; submitted 15.03.23; peer-reviewed by T Whitaker, C Dunst; comments to author 18.04.23; revised version received 07.05.23; accepted 25.07.23; published 18.08.23.

Please cite as:
Bilu Y, Amit G, Sudry T, Akiva P, Avgil Tsadok M, Zimmerman DR, Baruch R, Sadaka Y
A Developmental Surveillance Score for Quantitative Monitoring of Early Childhood Milestone Attainment: Algorithm Development and Validation
JMIR Public Health Surveill 2023;9:e47315
URL: https://publichealth.jmir.org/2023/1/e47315
doi:10.2196/47315
PMID:37489583

Original Paper

Utility of the Comprehensive Health and Stringency Indexes in Evaluating Government Responses for Containing the Spread of COVID-19 in India: Ecological Time-Series Study

Kamal Kishore[1], PhD; Vidushi Jaswal[2], PhD; Anuj Kumar Pandey[3], MPH; Madhur Verma[4], MD; Vipin Koushal[1], MD

[1]Department of Biostatistics, Post Graduate Institute of Medical Education and Research, Chandigarh, India
[2]Department of Psychology, Mehr Chand Mahajan DAV College, Chandigarh, India
[3]Department of Health Research, International Institute of Health Management Research, New Delhi, India
[4]Department of Community & Family Medicine, All India Institute of Medical Sciences, Bhatinda, India

Corresponding Author:
Kamal Kishore, PhD
Department of Biostatistics
Post Graduate Institute of Medical Education and Research
Madhya Marg, Sector 12,
Chandigarh, 160012
India
Phone: 91 9591349768
Email: kkishore.pgi@gmail.com

Abstract

Background: Many nations swiftly designed and executed government policies to contain the rapid rise in COVID-19 cases. Government actions can be broadly segmented as movement and mass gathering restrictions (such as travel restrictions and lockdown), public awareness (such as face covering and hand washing), emergency health care investment, and social welfare provisions (such as poor welfare schemes to distribute food and shelter). The Blavatnik School of Government, University of Oxford, tracked various policy initiatives by governments across the globe and released them as composite indices. We assessed the overall government response using the Oxford Comprehensive Health Index (CHI) and Stringency Index (SI) to combat the COVID-19 pandemic.

Objective: This study aims to demonstrate the utility of CHI and SI to gauge and evaluate the government responses for containing the spread of COVID-19. We expect a significant inverse relationship between policy indices (CHI and SI) and COVID-19 severity indices (morbidity and mortality).

Methods: In this ecological study, we analyzed data from 2 publicly available data sources released between March 2020 and October 2021: the Oxford Covid-19 Government Response Tracker and the World Health Organization. We used autoregressive integrated moving average (ARIMA) and seasonal ARIMA to model the data. The performance of different models was assessed using a combination of evaluation criteria: adjusted R^2, root mean square error, and Bayesian information criteria.

Results: implementation of policies by the government to contain the COVID-19 crises resulted in higher CHI and SI in the beginning. Although the value of CHI and SI gradually fell, they were consistently higher at values of >80% points. During the initial investigation, we found that cases per million (CPM) and deaths per million (DPM) followed the same trend. However, the final CPM and DPM models were seasonal ARIMA (3,2,1)(1,0,1) and ARIMA (1,1,1), respectively. This study does not support the hypothesis that COVID-19 severity (CPM and DPM) is associated with stringent policy measures (CHI and SI).

Conclusions: Our study concludes that the policy measures (CHI and SI) do not explain the change in epidemiological indicators (CPM and DPM). The study reiterates our understanding that strict policies do not necessarily lead to better compliance but may overwhelm the overstretched physical health systems. Twenty-first–century problems thus demand 21st-century solutions. The digital ecosystem was instrumental in the timely collection, curation, cloud storage, and data communication. Thus, digital epidemiology can and should be successfully integrated into existing surveillance systems for better disease monitoring, management, and evaluation.

(JMIR Public Health Surveill 2023;9:e38371) doi:10.2196/38371

KEYWORDS

COVID-19; government response; nonpharmaceutical interventions; lockdown; Comprehensive Health Index; Stringency Index; time-series modeling; ARIMA; SARIMA; Oxford COVID-19 Government Response Tracker; public health; surveillance; Oxford tracker; ecological study; health data; health policy; Bayesian information criteria

Introduction

SARS-CoV-2 is contagious, capricious, and hysteric—it adversely affected the social, mental, and physical well-being of individuals and societies—from the time the first case of COVID-19 was reported from Wuhan, China, in December 2019 [1]. The World Health Organization (WHO) declared COVID-19 a pandemic in March 2020 [2]. The viral disease occurred almost a century later than the deadliest viral disease in history—the influenza of 1918-1919 [3]. Globally, COVID-19 is associated with morbidity and mortality of 458 and 6.1 million people, respectively, till March 15, 2022 [4]. However, the official numbers underestimate actual morbidity and mortality estimates—a significant reason for concern among health experts [5]. The literature points to various reasons for deflating figures: differences in disease monitoring and reporting, testing strategy, asymptomatic cases, medically unattended cases, and deaths across nations [6-8].

Many nations swiftly designed and executed government policies to contain the rapid rise of COVID-19 cases [9]. Government actions can be broadly categorized as movement and mass gathering restrictions (such as travel restrictions and lockdown), public awareness (such as face masking and hand washing), emergency health care investment, and social welfare provisions (such as poor welfare scheme and providing food and shelter to migrant workers) [10]. However, the intermix of culture, communication, society, religion, politics, socioeconomic disparities, and education plays a dominant role in successfully implementing government policies [11,12]. Despite aggressive interventions by the Government of India (GOI), mass migration and verbal and physical abuse against health care providers were perplexing during the COVID-19 crisis [12-14]. These problems indicate that implementing policies to contain the public health crisis is one part of the puzzle, and understanding, adherence, and participation by citizens to make it a success is the other. Thus, a mismatch and mistrust between the public and policy makers will determine how the crises unfolds [15,16]. Despite these constraints, the principal priority was to contain the pandemic and mitigate future waves.

Traditional surveillance systems are robust and well-developed in many countries, but these were insufficient during the COVID-19 pandemic. Prominent challenges are that physical networks are extremely slow to upgrade (such as surveillance and physical structure), resource-intensive (in terms of finances and personnel), and involve bureaucratic hurdles (such as permissions and administrative authority). An emergency or crisis demands agile and innovative solutions. Internet technology can supplement the already existing traditional structure. The technological innovation of the internet and digital devices has fundamentally reshaped how people seek information and adhere to guidelines [17,18]. Therefore, blending existing surveillance systems with a digital ecosystem is ideal for harnessing and optimizing their untapped potential. Technology, medical development, health infrastructure, and lifestyle are pivotal in combating public health crises.

Many studies have highlighted the importance of nonpharmaceutical Interventions (NPIs) in containing the COVID-19 crisis [19-22]. Internet and digital devices played a critical role in screening, surveillance, monitoring spread, sharing, and managing data [23]. However, most of the studies focused on mobility during the current crisis. Mobility is vital, but many other essential indicators of disease spread exist. The Oxford tracker keeps track of various policy initiatives to mitigate COVID-19 and releases them as Comprehensive Health Index (CHI) and Stringency Index (SI) data [24]. CHI and SI are composite indicators covering the various aspects of NPIs; these indicators provide a systematic approach to understanding how the government has responded to mitigate disease spread over a period. We aimed to assess the utility of CHI and SI—innovative measures of the University of Oxford—to gauge and evaluate government responses for containing the spread of COVID-19. We hypothesize that there is a significant inverse relationship between policy indices (CHI and SI) and COVID-19 severity indices (morbidity and mortality). The analyses hope to inform and better prepare the experts to face and refine epidemiological indicators for future contagious disease crises.

Methods

Study Design

An ecological time-series study design was adopted, using secondary data.

Study Period

The first case of COVID-19 in India was reported from Kerala on January 27, 2020 [25]. The GOI responded to the COVID-19 threat by issuing travel advisories and screening at airports in January and later restricting international travel from March 2020 onward [26,27]. As the cases in India started escalating, the GOI called for a *Janata* curfew on March 23, 2020, and implemented a national lockdown from March 24 to June 7, 2020, for 75 days [28]. The GOI started phased unlocking from June 8, 2020, onward [28]. COVID-19 and SI data for India are available from January 21, 2020, to March 2022. This study used data from March 4, 2020 (when India crossed 10 cumulative cases), to October 24, 2021.

Data Sources

We used 2 publicly available data sources from the Oxford Covid-19 Government Response Tracker (OxCGRT) [29] and the WHO [4] licensed under the Creative Commons Attribution (CC BY) standard.

COVID-19 Data

The WHO collects the numbers of confirmed COVID-19 cases and deaths through official communications under the

International Health Regulations of 2005. WHO experts monitor the official ministries of health's websites and social media accounts for data validation. The data set, after curation, is made available as a .csv file and can be downloaded from the website [24].

CHI and SI Data

The OxCGRT for containment and closure and health data set evaluates the government response to containing the COVID-19 crisis. The OxCGRT collected 23 indicators to capture government policies related to closure and containment, health, and economics for more than 180 countries since January 1, 2020. The policy indicators are captured on an ordinal or numerical scale to describe each category's degree or the strength of the government response. Each ordinal indicator is transformed to a value between 0 and 100 per the formula given in the codebook [25]. Finally, to gauge government performance, individual indicators are aggregated and published in 5 composite indices: Government Response Index, Containment and Health Index, SI, Economic Support Index, and Legacy Stringency Index. The CHI uses 14 indicators and SI uses 9 indicators out of 23. The value for composite indices on any given day represents the aggregated average of each indicator for the day. The composite indices report a number between 0 and 100 that reflects the overall stringency of the government's response. A higher index indicates a higher overall response level. Table 1 shows details regarding individual indicators used to calculate the CHI and SI, both of which are updated regularly. The indicators are reported for each day that a policy was implemented (not on the day it was announced). A continuously updated OxCGRT index provides comparable information regarding various countries' policy measures. Many indicators have another flag variable to imply whether they are targeted (applying only to a subregion of a jurisdiction or to a specific sector) or general. The OxCGRT uses simple, additive, and unweighted indices that are easier to interpret. The missing value contributes a zero to the index. Multimedia Appendix 1 demonstrates the steps to download the OxCGRT data sets. Details of all the indicators and calculations for the composite index are provided in the working paper entitled, "Variation in government responses to COVID-19" [30].

Table 1. Details regarding individual and composite Oxford Covid-19 Government Response Tracker indicators.

ID	Policy measures	Scale	Score, maximum (range)	Stringency Index	Comprehensive Health Index	Flag[a]
Containment and closure						
C1	School closing	Ordinal	3 (0-3)	✓	✓	Yes=1
C2	Workplace closing	Ordinal	3 (0-3)	✓	✓	Yes=1
C3	Cancel public event	Ordinal	2 (0-2)	✓	✓	Yes=1
C4	Restriction on gathering size	Ordinal	4 (0-4)	✓	✓	Yes=1
C5	Close public transport	Ordinal	2 (0-2)	✓	✓	Yes=1
C6	Stay-at-home requirements	Ordinal	3 (0-3)	✓	✓	Yes=1
C7	Restriction on internal movement	Ordinal	2 (0-2)	✓	✓	Yes=1
C8	Restriction on international travel	Ordinal	4 (0-4)	✓	✓	No=0
Health systems						
H1	Public information campaign	Ordinal	2 (0-2)	✓	✓	Yes=1
H2	Testing policy	Ordinal	3 (0-3)		✓	No=0
H3	Contact tracing	Ordinal	2 (0-2)		✓	No=0
H6	Facial covering	Ordinal	4 (0-4)		✓	Yes=1
H7	Vaccination policy	Ordinal	5 (0-5)		✓	Yes=1
H8	Protection of older individuals	Ordinal	3 (0-3)		✓	Yes=1

[a] A value of 0 indicates that the policy is targeted to a subregion.

Study Variables

Dependent Variables

The data about morbidity and mortality were downloaded in the .csv file and reported using standardized metrics such as cases per million (CPM) and deaths per million (DPM). CPM and DPM are valuable indicators for intergeographical comparisons—the same metrics were calculated and reported using standard formulae provided by the WHO. Multimedia Appendix 2 provides details about the 7-day moving average, CPM, and DPM.

Independent Variables

Government policies and timing are vital indicators of morbidity and mortality related to disease during and after the crisis. Therefore, we used time, CHI, and SI, which are cumulative nonweighted indices of government policies as predictor variables. Table 1 provides detail of both CHI and SI.

Data Analysis

Data Processing and Summarization

We retained originally downloaded data sets of COVID-19 and OxCGRT for record and referral. Subsequently, a copy of each original data set was generated to clean, code, and analyze. Data cleaning in both files involved renaming, relabeling, and removing of undesired variables. Subsequently, we used the *date* variable to merge and prepare a final data set for analysis. The raw day–wise data were smoothed using a 7-day moving average that was subsequently used to calculate and report CPM and DPM. The morbidity indicator in the country was summarized using frequency and percentage increase. Initially, time-series plots were used to gauge the disease burden. Subsequently, a dual-axis chart was used to visualize the pattern between disease burden and policy indicators. The descriptive tables and graphs were prepared using Excel (Microsoft Corp).

Time-Series Modeling

The daily SARS-CoV-2 data follow the characteristics of the time series. As an initial step, we carefully inspected the daily incidence (morbidity and mortality), descriptive statistics (mean and variance), and seasonality (weekly and periodicity). We subsequently used second-order differencing for CPM and first-order differencing for DPM to make them stationary, which is an essential requirement to apply time-series analysis [31]. The twin advantage of simple structure and immediate applicability of autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) made them lucrative for analyzing the current study data set [32]. ARIMA considers only the past values for prediction, whereas SARIMA also considers the seasonality patterns, making it a more robust algorithm for prediction. We plotted autocorrelation function (ACF) and partial ACF (PACF) values to assess autoregressive and moving average components of the ARIMA and SARIMA.

Model Construction and Comparison

We could not finalize the order of autoregressive and moving average components from ACF and PACF plots. Therefore, considering the Expert Modeler and Box-Jenkins 5-step methodology (describe data, identify model, estimate parameters, diagnosis check, and forecasting), we relied on an expert modeler for an initial model from multiple candidate models. Subsequently, we used the Box-Jenkins technique that gives the flexibility to customize and attain the final model [33]. We used the date (time) as an explanatory variable in the initial model. Subsequently, CHI or SI and time were added to an expanded model to build and evaluate the models. The performance of different models was assessed using a combination of evaluation criteria: adjusted R^2, root mean square error, and Bayesian information criteria (BIC). After attaining the final model, we performed the diagnosis check to validate model assumptions. Multimedia Appendix 3 elucidates the step-by-step approach to attain the final model. We used a 2-tailed P value of <.05 to declare statistical significance. The time-series analysis of data was performed using SPSS (version 23; IBM Corp).

Ethical Considerations

Ethical clearance for the study was obtained from the institutional review board of the Postgraduate Institute of Medical Education and Research, Chandigarh, India (vide letter INT/IEC/2020/SPL–1594).

Results

Disease Burden

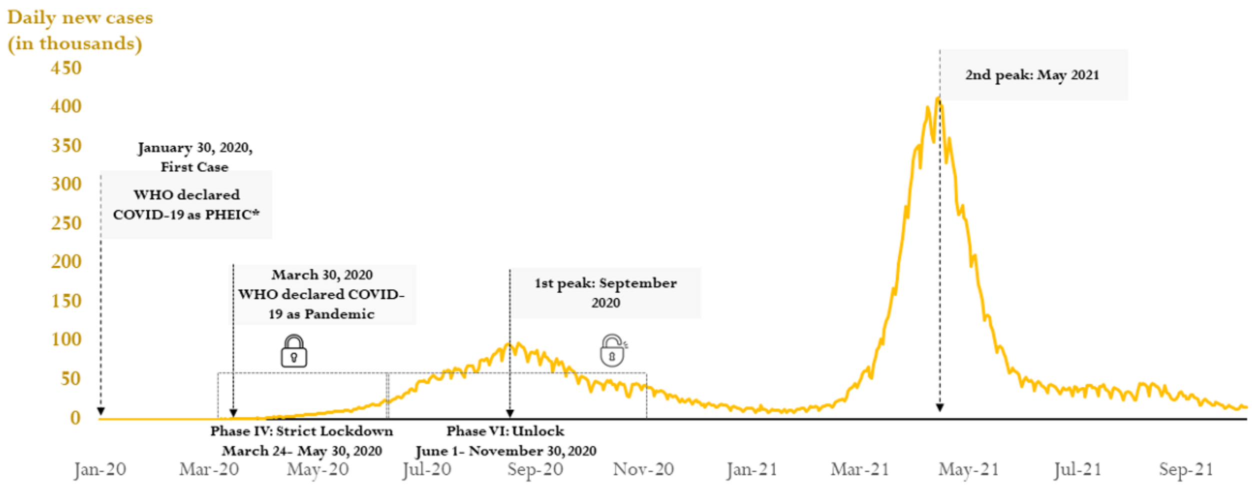
The first case of COVID-19 was detected in India on January 27, 2020. It took 46 days for the infection to spread among 100 people. India had 60 COVID-19–infected individuals on March 11, 2020, when the WHO declared COVID-19 a pandemic. True to its nature, COVID-19 started spreading rapidly, and it took only 15 days to reach 1000 cases from 100. Public health experts and policy makers were on high guard after GOI notification—a fact also reflected by the SI and CHI. Table 2 provides details about different milestones regarding the rise in disease burden. The increase in cases followed a linear trajectory during the lockdown stage. After unlocking, India achieved its first peak in September 2020, which was almost 100,000 cases per day. Subsequently, the second peak in May 2021 was characterized by a maximum caseload of ~400,000 cases per day. Figure 1 shows the change in India's disease burden from January 2020 to October 2021.

Table 2. Different milestones in the rise of the COVID-19 burden in India.

Date	Lockdown phases	Cumulative COVID-19 cases in India, n	Time from the report of the first COVID-19 case in India	Stringency index, median (range)	Comprehensive Health Index, median (range)
January 27, 2020	No	1	Frist case	—[a]	—
March 4, 2020	No	10	n=35 (1 month)	10.2 (10.2-10.2)	17.3 (13.7-17.3)
March 15, 2020	No	100	n=46 (1.5 months)	29.4 (26.9-38.9)	29.6 (27.9-35.7)
March 30, 2020	Phase 1: March 25-April 14, 2020	1000	n=61 (2 months)	81.7 (48.2-100)	66.8 (41.7-80.9)
April 14, 2020	Phase 2: April 15-May 3, 2020	10,000	n=76 (2.5 months)	100 (100-100)	86.2 (80.9-91.9)
May 19, 2020	Phase 4: May 18-May 31, 2020	100,000	n=111 (4 months)	90.3 (81.9-100)	84.1 (76.8-91.9)
July 17, 2020	Unlock 2.0: July 1-July 31, 2020	1,000,000	n=170 (5.5 months)	86.4 (81.9-87.5)	74.9 (74.4-76.8)
December 19, 2020	Partial	10,000,000	n=325 (10.5 months)	75.9 (61.6-87.5)	68.8 (61.3-74.4)
May 4, 2021	Partial	20,000,000	n=461 (15 months)	65.6 (57.9-74.5)	71.3 (66.1-81.1)
June 23, 2021	Partial	30,000,000	n=547 (18 months)	81.1 (73.6-81.9)	79.4 (74.5-79.8)

[a] Not determined.

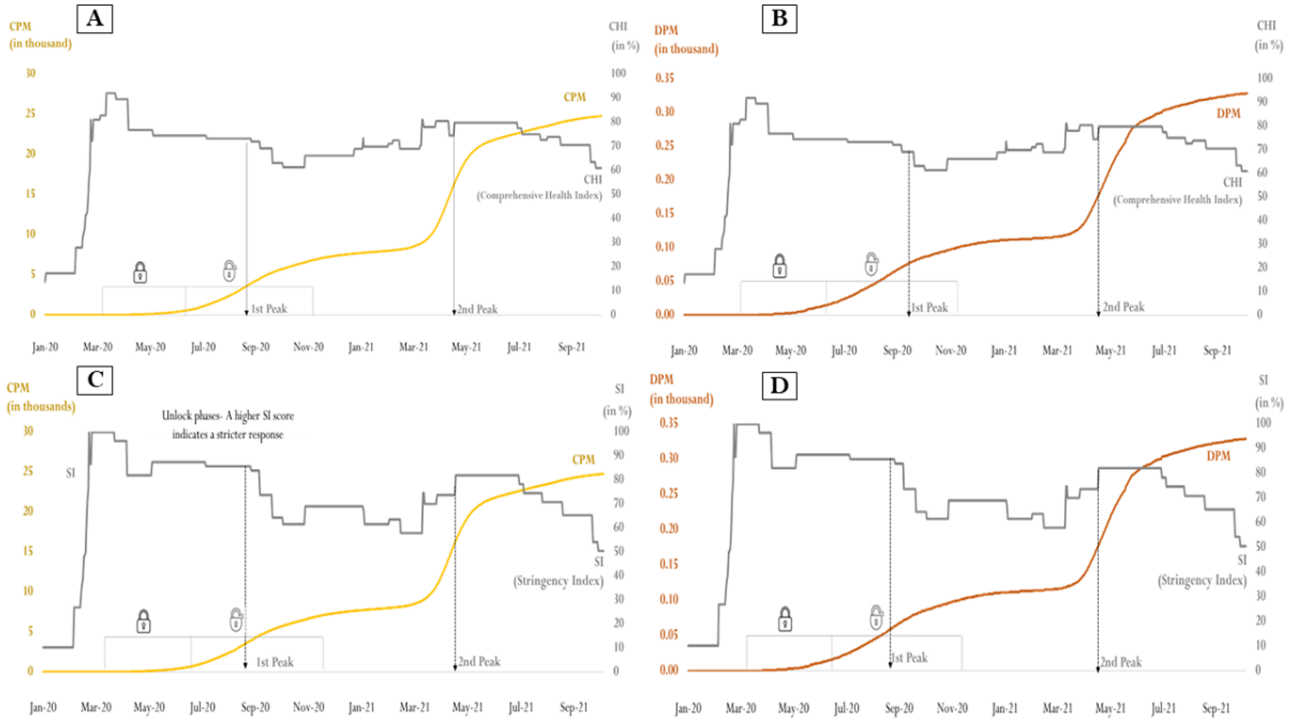
Figure 1. A line graph (epidemic curve) showing daily new COVID-19 cases in India. Adapted from the World Health Organization's (WHO's) Coronavirus Dashboard [4]. PHEIC: public health emergency of international concern.



Epidemiological and Policy Indicators

The strict implementation of policies by the government to contain the COVID-19 crisis led to a very high value of CHI and SI in the beginning. Figure 2A shows that the CPM was near the x-axis from the beginning till the end of the fourth lockdown. Subsequently, CHI reduced gradually; in contrast, the CPM swiftly increased after the GOI initiated unlocking from July 2020. Figure 2B shows that the DPM also increased steadily with time—the highest during the unlock phase. Although the CHI and SI reduced marginally, there were spikes during the first and second peaks. The CPM and DPM also displayed similar increasing trends with the SI. Figures 2C and 2D depict CPM and DPM changes with SI over time.

Figure 2. Line graphs depicting changes in the Comprehensive Health Index or Stringency Index (SI) versus cases per million (CPM) or deaths per million (DPM) for SARS-CoV-2 in India. (A) Change in the Comprehensive Health Index and CPM, (B) change in the Comprehensive Health Index and DPM, (C) change in the Stringency Index (SI) and CPM, and (D) change in the SI and DPM. Data obtained from the World Health Organization's Coronavirus Dashboard [4] and from GitHub [28].



Time-Series Modeling

Figures 1 and 2A-2D show that both positive (symptomatic) and deaths (mortality) were nonstationary time series. On the surface, the CPM and DPM followed the same trend. However, the final CPM and DPM model were SARIMA (3,2,1)(1,0,1) and ARIMA (1,1,1), respectively. Multimedia Appendix 3 provides details regarding the steps involved in obtaining the final models. There was no improvement in model performance for CPM: the BIC values for both time (BIC 2.42) and the expanded model (BIC 2.46) were almost the same. Similarly, the BIC values for both time (BIC –1.35) and the expanded model (BIC –1.34) models for DPM were the same. Despite the appearance of an inverse relationship between COVID-19 and policy indicators, neither CHI nor SI were significant predictors of CPM and DPM. Table 3 displays the final model's output and performance indicators: BIC, stationary R^2, and root mean square error.

Table 3. The output of the final model for COVID-19 cases per million and deaths per million in India from March 2020 to October 2021.

Model and parameter	Estimate	*P* value	Bayesian information criteria	Stationary R^2	Root mean square error
Seasonal autoregressive integrated moving average (3,2,1)(1,0,1)			2.42	0.62	3.22
AR1	0.67	<.001			
AR2	0.12	.02			
AR3	0.10	.04			
MA1	0.73	<.001			
SAR1	0.86	<.001			
SMA1	0.25	<.001			
Autoregressive integrated moving average (1,1,1)			–1.35	0.64	0.50
AR1	0.99	<.001			
MA1	0.82	<.001			

Discussion

Principal Findings

We used the CHI and SI data from the Blavatnik School of Government, University of Oxford, response tracker to evaluate and model COVID-19 morbidity and mortality. There are specific critical findings in our study. COVID-19 displayed the typical characteristic of a highly contagious disease that spread rapidly after a slow start. At the beginning of the pandemic in India, cases were geographically scattered and scarce; thus, the CHI and SI were low as neither the GOI issued any guidelines nor did the people adhere to preventive behaviors. There was a sharp increase in cases after unlocking, and India witnessed 2-clear and visible peaks in September 2020 and May 2021; this was in line with a temporary increase in CHI and SI during peaks. The value of CHI and SI gradually decreased; similarly, however, these were consistently higher at values of >80% points.

For intercountry comparisons, the CPM and DPM were calculated and reported, since these are better measures than the raw number of cases. The CPM and DPM linearly increased when stringency was high during the lockdowns and rapidly increased after the government lifted the restrictions. Despite initial indication, both CHI and SI were (rather surprisingly) not significant predictors of COVID-19 morbidity and mortality.

People's participation is vital to the success of public health interventions. Thus, the GOI tested people's moods and sentiments with the *Janata Curfew* (self-restricted curfew) before implementing strict public health measures. After the success of the Janata curfew, the GOI enforced a nationwide lockdown to close schools, parks, transport, offices, and borders to contain the COVID-19 crisis. The daily incidence was relatively flat during the entire lockdown period; the strict response to epidemic control measures reduced daily incidence [26]. However, social control measures mitigate contagious diseases; they do not eradicate them [27,28]. The rate of disease spread increased rapidly after the unlocking process began. Despite the nationwide lockdown, significant intra- and interstate variation was a cause of concern—the majority (~76%) of cases were reported from 10 out of 36 states and union territories in India [29]; the disparity can be partially attributed to constitutional provisions besides factors such as mobility, urbanization, and population density. Moreover, while the central government has the power to make public health laws, the state government develops infrastructure and executes public health policy because health care and hospitals are subject matters in the state list [4].

NPIs such as tests, contact tracing, masking, social distancing, mobility, hygiene, and vaccination form the backbone for controlling contagious diseases. However, in the absence of vaccination, strategic implementation of NPIs helps mitigate the rapid spread of COVID-19 [13-16]. The CHI and SI are more comprehensive than specific mobility indicators—these cover various aspects of NPIs. The CHI and SI attain very high values, and the results are in parallel with those of Ma et al [34], which indicate that countries that implemented nationwide lockdowns in March took strict measures [26]. The high value of CHI and SI during the lockdown reflects public health experts' and policy makers' comprehensive and dynamic responses to contain the COVID-19 crisis. Government-enforced lockdowns are significant contributors to restrictions on activities. People also respond to contagious threats by restricting socializing and travelling [35]. Overall, CHI and SI reflected high stringency; this further increased during the peak. Initial investigation reveals a distinct inverse relationship between restriction indicators (CHI and SI indices) and disease burden (morbidity and mortality); this relationship was explicit as the unlocking phase began.

We fitted a time-series model to investigate the relationship between policy and disease indicators. The time-series model with dates explained more than 60% of variability for both CPM and DPM models. The initial investigation indicated the relationship between predictors (CHI and SI) and disease burden (CPM and DPM). However, the expanded inferential models do not increase the models' performances—neither CHI nor SI contributed significantly explaining the change in CPM and DPM. In other words, stringency measures such as CHI and SI do not explain the change in both COVID-19 morbidity and mortality. The results of our study are in contrast to those of other studies [26,30,31]. However, it may be crucial to note that strict implementation of policies does not necessarily lead to strict compliance [32,33]. A thorough inspection of CHI and SI indicates the government's proactiveness on different indicators compared to participation. India saw the mass movement of migrant workers [34,36] and multiple attacks on health care workers during the lockdowns [34,37]. Further, COVID-19 was a relatively urban phenomenon; most of India is rural.

Strengths and Limitations

A significant strength of our study is that this is the first study to investigate the relationship between policy indicators CHI and SI and epidemiological indicators CPM and DPM. It is the first pan-India study that has used WHO and OxCGRT data to quantify and model COVID-19 transmission. Our study is different from other studies focused on NPIs including mobility or SI but not CHI.

However, a significant limitation of this study, and perhaps most digital epidemiological studies, is the validity and reliability of the data. As already highlighted, CHI and SI are proxy measures that do not reflect strict compliance with the implementation of policies. There may be an inherent bias or reporting error in composite CHI and SI indices at the country level due to within and between heterogeneity at the state and union territory levels. The unweighted indices (CHI and SI) are easy to interpret but make strong assumptions, and a user may obtain different results when using weighted indices. Lastly, correct reporting of the daily incidence of COVID-19 cases depends on a country's accuracy and testing capacity.

Future Work

Countries with vast geographical and administrative regions may differ significantly in implementing policies. Therefore, further research is required to develop and validate the metrics to identify whether country- or county-level (state and union territory) metrics are needed. The capacity to test, trace, and treat varies from country to country. Do disease morbidity and mortality reporting reflect the actual scenario? What logical steps and weightage must be given to each region is also a perplexing problem vouching for researchers' attention. The media plays a vital role in containing public health crises. However, did it aid in signaling by motivating people to adhere to guidelines, or has it added to noise by creating panic among the public? Lastly, are OxCGRT indicators enough to capture people's participation and policy makers' execution or do they require appropriate weightage? The aforementioned issues invite attention from health experts and policy makers.

Conclusions and Recommendations

Our study concludes that the policy measures (CHI and SI) do not explain the change in epidemiological indicators (CPM and DPM). The study reiterates our understanding that strict policies do not necessarily lead to better compliance but may overwhelm the overstretched physical health systems. Twenty-first–century problems thus demand 21st-century solutions. The digital ecosystem was instrumental in the timely collection, curation, cloud storage, and data communication. Thus, digital epidemiology can and should be successfully integrated into existing surveillance systems for better disease monitoring, management, and evaluation. An OxCGRT policy metric is a novel innovation to assess government actions during the epidemic, which have the potential for future use and refinement. Therefore, policy makers, public health experts, and programmers must start collaborating to design a hybrid health system that can borrow from the strengths of the existing physical surveillance system and the ever-expanding digital ecosystem.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of Steps to download the OxCGRT datasets.
[DOCX File , 623 KB - publichealth_v9i1e38371_app1.docx]

Multimedia Appendix 2 [DOCX File , 18 KB - publichealth_v9i1e38371_app2.docx]

Multimedia Appendix 3
Details of the steps to obtain final models.
[DOCX File , 265 KB - publichealth_v9i1e38371_app3.docx]

References

1. Zhu H, Wei L, Niu P. The novel coronavirus outbreak in Wuhan, China. Glob Health Res Policy 2020 Mar 02;5(1):6 [FREE Full text] [doi: 10.1186/s41256-020-00135-6] [Medline: 32226823]
2. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. 2020 Mar 11. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2022-01-16]
3. Bloom D, Cadarette D. Infectious disease threats in the twenty-first century: strengthening the global response. Front Immunol 2019;10:549 [FREE Full text] [doi: 10.3389/fimmu.2019.00549] [Medline: 30984169]
4. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int [accessed 2022-01-16]
5. Jha P, Deshmukh Y, Tumbe C, Suraweera W, Bhowmick A, Sharma S, et al. COVID mortality in India: national survey data and health facility deaths. Science 2022 Feb 11;375(6581):667-671 [FREE Full text] [doi: 10.1126/science.abm5154] [Medline: 34990216]
6. Havers FP, Reed C, Lim T, Montgomery JM, Klena JD, Hall AJ, et al. Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020. JAMA Intern Med 2020 Jul 21 [FREE Full text] [doi: 10.1001/jamainternmed.2020.4130] [Medline: 32692365]
7. Nakamoto I, Zhang J. Modeling the underestimation of COVID-19 infection. Results Phys 2021 Jun;25:104271 [FREE Full text] [doi: 10.1016/j.rinp.2021.104271] [Medline: 34002127]
8. Pullano G, Di Domenico L, Sabbatini CE, Valdano E, Turbelin C, Debin M, et al. Underdetection of cases of COVID-19 in France threatens epidemic control. Nature 2021 Feb 21;590(7844):134-139. [doi: 10.1038/s41586-020-03095-6] [Medline: 33348340]
9. Zeng K, Bernardo SN, Havins WE. The use of digital tools to mitigate the COVID-19 pandemic: comparative retrospective study of six countries. JMIR Public Health Surveill 2020 Dec 23;6(4):e24598 [FREE Full text] [doi: 10.2196/24598] [Medline: 33302255]
10. GRID COVID-19 Study Group. Combating the COVID-19 pandemic in a resource-constrained setting: insights from initial response in India. BMJ Glob Health 2020 Nov;5(11) [FREE Full text] [doi: 10.1136/bmjgh-2020-003416] [Medline: 33187963]

11. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. Nat Hum Behav 2020 May 30;4(5):460-471. [doi: 10.1038/s41562-020-0884-z] [Medline: 32355299]
12. Chavan BS, Sidana A, Arun P, Rohilla R, Singh GP, Solanki RK, et al. Factors leading to reverse migration among migrant workers during the COVID-19 pandemic. Prim Care Companion CNS Disord 2021 Mar 25;23(2). [doi: 10.4088/pcc.20m02810]
13. Larkin H. Navigating attacks against health care workers in the COVID-19 era. JAMA 2021 May 11;325(18):1822-1824 [FREE Full text] [doi: 10.1001/jama.2021.2701] [Medline: 33881489]
14. Devi S. COVID-19 exacerbates violence against health workers. Lancet 2020 Sep;396(10252):658. [doi: 10.1016/s0140-6736(20)31858-4]
15. Fridman I, Lucas N, Henke D, Zigler CK. Association between public knowledge about COVID-19, trust in information sources, and adherence to social distancing: cross-sectional survey. JMIR Public Health Surveill 2020 Sep 15;6(3):e22060 [FREE Full text] [doi: 10.2196/22060] [Medline: 32930670]
16. Al-Hasan A, Yim D, Khuntia J. Citizens' adherence to COVID-19 mitigation recommendations by the government: a 3-country comparative evaluation using web-based cross-sectional survey data. J Med Internet Res 2020 Aug 11;22(8):e20634 [FREE Full text] [doi: 10.2196/20634] [Medline: 32716896]
17. Kelly JT, Campbell KL, Gong E, Scuffham P. The Internet of Things: impact and implications for health care delivery. J Med Internet Res 2020 Nov 10;22(11):e20135 [FREE Full text] [doi: 10.2196/20135] [Medline: 33170132]
18. Yeung AWK, Kulnik ST, Parvanov ED, Fassl A, Eibensteiner F, Völkl-Kernstock S, et al. Research on digital technology use in cardiology: bibliometric analysis. J Med Internet Res 2022 May 11;24(5):e36086 [FREE Full text] [doi: 10.2196/36086] [Medline: 35544307]
19. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nat Hum Behav 2020 Dec 16;4(12):1303-1312. [doi: 10.1038/s41562-020-01009-0] [Medline: 33199859]
20. Perra N. Non-pharmaceutical interventions during the COVID-19 pandemic: a review. Phys Rep 2021 May 23;913:1-52 [FREE Full text] [doi: 10.1016/j.physrep.2021.02.001] [Medline: 33612922]
21. Seale H, Dyer CEF, Abdi I, Rahman KM, Sun Y, Qureshi MO, et al. Improving the impact of non-pharmaceutical interventions during COVID-19: examining the factors that influence engagement and the impact on individuals. BMC Infect Dis 2020 Aug 17;20(1):607 [FREE Full text] [doi: 10.1186/s12879-020-05340-9] [Medline: 32807087]
22. Chan LYH, Yuan B, Convertino M. COVID-19 non-pharmaceutical intervention portfolio effectiveness and risk communication predominance. Sci Rep 2021 May 19;11(1):10605 [FREE Full text] [doi: 10.1038/s41598-021-88309-1] [Medline: 34012040]
23. Kishore K, Jaswal V, Verma M, Koushal V. Exploring the utility of Google mobility data during the COVID-19 pandemic in India: digital epidemiological analysis. JMIR Public Health Surveill 2021 Aug 30;7(8):e29957 [FREE Full text] [doi: 10.2196/29957] [Medline: 34174780]
24. Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021 Apr 08;5(4):529-538. [doi: 10.1038/s41562-021-01079-8] [Medline: 33686204]
25. Andrews M, Areekal B, Rajesh K, Krishnan J, Suryakala R, Krishnan B, et al. First confirmed case of COVID-19 infection in India: a case report. Indian J Med Res 2020 May;151(5):490-492 [FREE Full text] [doi: 10.4103/ijmr.IJMR_2131_20] [Medline: 32611918]
26. Text of PM’s address to the nation on vital aspects relating to the menace of COVID-19. Press Information Bureau. Government of India. Prime Minister's Office. 2020 Mar 24. URL: https://pib.gov.in/newsite/PrintRelease.aspx?relid=200654 [accessed 2022-01-15]
27. Order No. 40-3/2020-DM-I(A). Government of India. Ministry of Home Affairs. 2020 Mar 24. URL: https://www.mha.gov.in/sites/default/files/MHAorder%20copy_0.pdf [accessed 2022-01-15]
28. Order No. 40-3/2020-DM-I(A). Government of India. Ministry of Home Affairs. 2020 May 30. URL: https://www.mha.gov.in/sites/default/files/MHAOrderDt_30052020_0.pdf [accessed 2022-01-15]
29. Covid–policy–tracker/data at master · OxCGRT/covid–policy–tracker. GitHub. URL: https://github.com/OxCGRT/covid?policy?tracker [accessed 2022-01-15]
30. Wade A, Petherick A, Kira B, Cameron-Blake E, Tatlow H, Elms J, et al. Variation in government responses to COVID-19 BSG-WP-2020/032. Blavatnik School of Government. University of Oxford. URL: https://www.bsg.ox.ac.uk/sites/default/files/2022-08/BSG-WP-2020-032-v14.1.pdf [accessed 2023-01-23]
31. Hyndman R, Athanasopoulos G. Forecasting: Principles and Practice (2nd ed). New York, NY: Springer; 2018.
32. Wang Y, Yan Z, Wang D, Yang M, Li Z, Gong X, et al. Prediction and analysis of COVID-19 daily new cases and cumulative cases: times series forecasting and machine learning models. BMC Infect Dis 2022 May 25;22(1):495 [FREE Full text] [doi: 10.1186/s12879-022-07472-6] [Medline: 35614387]

33. Guleryuz D. Forecasting outbreak of COVID-19 in Turkey; comparison of Box-Jenkins, Brown's exponential smoothing and long short-term memory models. Process Saf Environ Prot 2021 May;149:927-935 [FREE Full text] [doi: 10.1016/j.psep.2021.03.032] [Medline: 33776248]
34. Ma Y, Mishra SR, Han X, Zhu D. The relationship between time to a high COVID-19 response level and timing of peak daily incidence: an analysis of governments' Stringency Index from 148 countries. Infect Dis Poverty 2021 Jul 05;10(1):96 [FREE Full text] [doi: 10.1186/s40249-021-00880-x] [Medline: 34225774]
35. Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021 Apr;5(4):529-538 [FREE Full text] [doi: 10.1038/s41562-021-01079-8] [Medline: 33686204]
36. Huang H, Chen Y, Yan Z. Impacts of social distancing on the spread of infectious diseases with asymptomatic infection: a mathematical model. Appl Math Comput 2021 Jun 01;398:125983 [FREE Full text] [doi: 10.1016/j.amc.2021.125983] [Medline: 33518834]
37. Heywood AE, Macintyre CR. Elimination of COVID-19: what would it look like and is it possible? Lancet Infect Dis 2020 Sep;20(9):1005-1007. [doi: 10.1016/s1473-3099(20)30633-2]

Abbreviations

ACF: autocorrelation function
ARIMA: autoregressive integrated moving average
BIC: Bayesian information criteria
CHI: Comprehensive Health Index
CPM: cases per million
DPM: deaths per million
GOI: Government of India
NPI: nonpharmaceutical intervention
OxCGRT: Oxford Covid-19 Government Response Tracker
PACF: partial autocorrelation function
SARIMA: seasonal autoregressive integrated moving average
SI: Stringency Index
WHO: World Health Organization

Edited by T Sanchez, A Mavragani; submitted 30.03.22; peer-reviewed by K Goel, M Panda, B Gummidi, T Kiran, M Mirza; comments to author 05.10.22; revised version received 25.10.22; accepted 18.01.23; published 10.02.23.

Please cite as:
Kishore K, Jaswal V, Pandey AK, Verma M, Koushal V
Utility of the Comprehensive Health and Stringency Indexes in Evaluating Government Responses for Containing the Spread of COVID-19 in India: Ecological Time-Series Study
JMIR Public Health Surveill 2023;9:e38371
URL: https://publichealth.jmir.org/2023/1/e38371
doi:10.2196/38371
PMID:36395334

Original Paper

Artificial Intelligence in Community-Based Diabetic Retinopathy Telemedicine Screening in Urban China: Cost-effectiveness and Cost-Utility Analyses With Real-world Data

Senlin Lin[1,2], MSc; Yingyan Ma[1,2*], PhD; Yi Xu[1,2], PhD; Lina Lu[1,2], MPH; Jiangnan He[1,2], PhD; Jianfeng Zhu[1,2], MD; Yajun Peng[1,2], BA; Tao Yu[1,2], BA; Nathan Congdon[3,4,5*], MD; Haidong Zou[1,2*], MD

[1]Department of Eye Disease Prevention and Control, Shanghai Eye Disease Prevention and Treatment Center/Shanghai Eye Hospital, Shanghai, China

[2]Shanghai Key Laboratory of Ocular Fundus Diseases, Shanghai General Hospital, Shanghai Engineering Center for Visual Science and Photomedicine, Shanghai, China

[3]Centre for Public Health, Queen's University Belfast, Belfast, United Kingdom

[4]Orbis International, New York, NY, United States

[5]Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Haidong Zou, MD
Department of Eye Disease Prevention and Control, Shanghai Eye Disease Prevention and Treatment Center/Shanghai Eye Hospital
No. 1440, Hongqiao Road
Shanghai, 200336
China
Phone: 86 02162539696
Email: zouhaidong@sjtu.edu.cn

Abstract

Background: Community-based telemedicine screening for diabetic retinopathy (DR) has been highly recommended worldwide. However, evidence from low- and middle-income countries (LMICs) on the choice between artificial intelligence (AI)–based and manual grading–based telemedicine screening is inadequate for policy making.

Objective: The aim of this study was to test whether the AI model is more worthwhile than manual grading in community-based telemedicine screening for DR in the context of labor costs in urban China.

Methods: We conducted cost-effectiveness and cost-utility analyses by using decision-analytic Markov models with 30 one-year cycles from a societal perspective to compare the cost, effectiveness, and utility of 2 scenarios in telemedicine screening for DR: manual grading and an AI model. Sensitivity analyses were performed. Real-world data were obtained mainly from the Shanghai Digital Eye Disease Screening Program. The main outcomes were the incremental cost-effectiveness ratio (ICER) and the incremental cost-utility ratio (ICUR). The ICUR thresholds were set as 1 and 3 times the local gross domestic product per capita.

Results: The total expected costs for a 65-year-old resident were US $3182.50 and US $3265.40, while the total expected years without blindness were 9.80 years and 9.83 years, and the utilities were 6.748 quality-adjusted life years (QALYs) and 6.753 QALYs in the AI model and manual grading, respectively. The ICER for the AI-assisted model was US $2553.39 per year without blindness, and the ICUR was US $15,216.96 per QALY, which indicated that AI-assisted model was not cost-effective. The sensitivity analysis suggested that if there is an increase in compliance with referrals after the adoption of AI by 7.5%, an increase in on-site screening costs in manual grading by 50%, or a decrease in on-site screening costs in the AI model by 50%, then the AI model could be the dominant strategy.

Conclusions: Our study may provide a reference for policy making in planning community-based telemedicine screening for DR in LMICs. Our findings indicate that unless the referral compliance of patients with suspected DR increases, the adoption of the AI model may not improve the value of telemedicine screening compared to that of manual grading in LMICs. The main reason is that in the context of the low labor costs in LMICs, the direct health care costs saved by replacing manual grading with AI are less, and the screening effectiveness (QALYs and years without blindness) decreases. Our study suggests that the magnitude of the value generated by this technology replacement depends primarily on 2 aspects. The first is the extent of direct health care costs reduced by AI, and the second is the change in health care service utilization caused by AI. Therefore, our research can also provide analytical ideas for other health care sectors in their decision to use AI.

(JMIR Public Health Surveill 2023;9:e41624) doi:10.2196/41624

KEYWORDS

artificial intelligence; cost; diabetic retinopathy; utility; low- and middle-income countries; screening

Introduction

Diabetic retinopathy (DR) is a leading cause of blindness worldwide. It often develops 10-15 years after the onset of diabetes and can take several forms—all potentially causing vision loss or blindness, including diabetic macular edema (DME) due to increased retinal vascular permeability and central retinal thickening, retinal ischemia resulting in the damage or death of light-sensing retinal photoreceptors, and proliferative DR, where the growth of fragile new blood vessels causes vitreous hemorrhage and retinal detachment [1].

In 2020, the global burden of DR and sight-threatening DR (STDR) was estimated to be 103 million and 29 million people, respectively, which will increase to 161 million and 45 million by 2045 due to the increasing prevalence of diabetes mellitus (DM) [1,2]. From 1990 to 2010, visual impairment due to DR increased by 64% and blindness by 27% [1], both of which were due to the rising DM prevalence in low- and middle-income countries (LMICs). Several studies have confirmed the benefits of telemedicine screening for DR [3-6]. Compared to no screening and traditional face-to-face screening, telemedicine screening is highly cost-effective in the long term. As a result, telemedicine screening will become the main form of community-based eye disease screening [5]. Some recent studies have suggested that using artificial intelligence (AI) can further reduce the costs of telemedicine screening [3,7-10]. Studies in high-income countries such as the United Kingdom and Singapore have shown that when AI is used in DR screening programs, screening costs can be reduced by up to approximately 20% compared with the costs incurred in manual grading [7-9]. This can be easily understood: from the perspective of a health economic evaluation, the main difference between the AI model and manual grading is that technology costs replace labor costs. Therefore, in settings where labor costs are high, such as in high-income countries, using AI instead of manual grading would save a lot of screening costs, making the screening more cost-effective [10]. However, because labor costs are low in low-income countries, conclusions from high-income countries may not be equally suitable, and evidence from LMICs is inadequate. Therefore, the objective of our community-based telemedicine screening for DR was to examine whether the AI model can be more cost-effective than manual grading in LMICs. We conducted a health economic evaluation by using real-world data from a large community-based telemedicine screening program for DR in Shanghai, China. We expect that this study will provide a reference for policy making with regard to DR screening in the context of low labor costs.

Methods

Study Setting

This study was conducted in Shanghai, China, wherein the prevalence of type 2 diabetes among adults was 6.25% between 2016 and 2019 [11]. Since 2010, Shanghai has conducted a teleophthalmology-based DR screening program under which residents can undergo fundus photography at community health service centers. After retinal experts at designated DR diagnosis centers have made a diagnosis based on these images, screening results are fed into the community health service center, where patients are counselled by general practitioners, and medical advice is offered. By 2017, all 250 community health service centers in Shanghai were equipped to participate in this program, and plans had begun to build an AI-assisted DR screening system [12,13]. A convolutional neural network, a type of deep learning model [14], was applied to the problem of diagnosing DR from fundus images, with the aim of replacing retinal experts in DR diagnosis centers with the AI algorithm on a cloud-based server. Since 2020, 56 community health service centers have shifted to an AI-assisted DR screening system. In 2021, these centers screened approximately 40,000 community residents for DR by using AI. To maximize the efficiency of the Shanghai program, community health service centers coordinated voluntary screening for residents of a given community at a particular place and time. Those who were diagnosed with DR at hospitals could still participate in the free annual community screening to monitor disease progression.

Model Overview

TreeAge Pro (TreeAge Software) was used to build a decision-analytic Markov model to compare the actual cost, effectiveness, and utility of manual grading telemedicine screening and AI-based assessment for DR (Multimedia Appendices 1-4). The incremental cost-effectiveness ratio (ICER) and incremental cost-utility ratio (ICUR) were calculated as the primary results. The effectiveness was defined as years without blindness per 100,000 people with DM, and the utility was evaluated by quality-adjusted life years (QALYs). Although all residents with DM could participate in our community-based screening, the majority were older people [15,16]; therefore, a hypothetical cohort of community residents with DM was followed in the model from the age of 65 years through a total of 30 one-year Markov cycles [5]. The characteristics of the simulated cohort were extracted using the Shanghai Digital Eye Disease Screening Program (Table 1).

Table 1. Characteristics of the simulated cohort and comparisons between community health centers using artificial intelligence and those using manual grading.

	All CHCs[a]	CHCs using artificial intelligence	CHCs using manual grading	*P* value
Community health center characteristics[b], mean (SD)				
Number of full-time or part-time ophthalmologists	1.1 (0.7)	1.1 (0.5)	1.2 (0.7)	.62
Annual numbers of ophthalmology outpatients	6248.6 (6019.0)	6774.8 (6238.2)	6016.2 (5946.4)	.54
Screened residents' characteristics[c]				
Age (years), mean (SD)	69.3 (7.2)	69.8 (6.9)	68.8 (7.4)	<.001
Sex (male), n (%)	15,032 (46.0)	7225 (46.1)	7807 (45.8)	.60
Duration of diabetes (years), mean (SD)	10.3 (7.7)	9.8 (8.0)	10.7 (7.3)	<.001
Visual acuity of right eye (logMAR), mean (SD)	0.4 (0.4)	0.4 (0.4)	0.3 (0.4)	<.001
Visual acuity of left eye (logMAR), mean (SD)	0.4 (0.4)	0.4 (0.4)	0.3 (0.4)	<.001

[a] CHC: community health center.

[b] In 2019, we conducted an investigation of the ophthalmic resources in Shanghai. Out of 250 community health centers, 111 were randomly selected. After matching with the Shanghai Digital Eye Disease Screening program data, 34 community health centers using artificial intelligence and 77 community health centers using manual grading–based telemedicine screening were investigated. We compared both these groups to see whether the community health centers' characteristics were different between those choosing new technology and those who did not.

[c] A total of 32,695 residents with diabetes were screened in the Shanghai Digital Eye Disease Screening program. Among them, 15,663 residents were in the community health centers with artificial intelligence, and the rest 17,032 were in the community health centers with manual grading. We compared the residents' characteristics in both groups of community health centers. Although some residents' characteristics were significantly different owing to the large sample size in the 2 groups, the differences were not clinically meaningful. Therefore, it can be assumed that there is no practical difference between the 2 groups.

Individuals were enrolled as healthy (free from DR) or unhealthy (experiencing DR) and could die due to any reason. According to the English National Screening Program for Diabetic Retinopathy, a Markov model was constructed that included non-STDR, STDR, and DME [5,15-17]. The category was assigned based on the DR grade in the worse eye. During each 1-year cycle, an individual had a risk of progressing to the more severe stage or staying in the same stage. However, the model does not allow returning to an earlier stage even with treatment because of the nature of the disease. Moreover, the treatment can only decrease the probability of progression to the next stage. The prevalence of DR, the incidence of DR (including STDR and DME), transition probabilities, characteristics of DR screening tests, referral and treatment compliance, utility, mortality, and other relevant parameters were collected from published studies specific to Shanghai, other cities in China, and other Asian regions, as well as unpublished data sources (eg, Shanghai Digital Eye Disease Screening Program). The costs of screening, ocular examinations, and treatment were all derived from a real-world eye disease screening program in Shanghai and the unified health care service pricing of the Shanghai Municipal Health Commission. The parameters used in the basic analysis and the ranges used in the sensitivity analyses are listed in detail in Multimedia Appendices 1-4.

Overview of the Screening Strategies

Manual Grading–Based Telemedicine Screening

We invited the entire population with DM living in communities to participate in the DR screening program at local community health centers. All the participants underwent a series of screening tests conducted by trained general practitioners, ophthalmic technicians, optometrists, and ophthalmologists. The screening included a vision acuity test, refraction measurement by an autorefractor, and fundus photography using a non–mydriatic fundus camera. The data were transferred to the corresponding designated diagnosis center through a telemedicine platform after the completion of all the tests. After all the participants in 1 community health center completed the annual screening, the community health center contacted the designated diagnosis center, and 2 retinal experts (ophthalmologists) began to make the diagnosis based on retinal photography. In 2 weeks, screening results were provided as feedback to the community health center, where residents could receive medical advice from the general practitioners. Finally, patients with suspected STDR were referred to specialized ophthalmic hospitals or tertiary hospitals for a detailed re-examination to confirm the diagnosis (Multimedia Appendix 5 shows the screening and referral pathway). Those who were confirmed to have STDR were assumed to receive appropriate treatment and routine clinical care according to the severity of DR.

AI-Assisted DR Telemedicine Screening

We invited the entire population with DM living in the community to participate in the DR screening program at the local community health center. The screening process was the same as that described for the manual grading–based telemedicine screening. However, after all the screening tests were completed, the data were transmitted to the AI algorithm on a cloud-based server center through the telemedicine platform. The screening results were provided as feedback immediately. Further management of patients with suspected STDR was the same as that described for manual grading–based telemedicine screening.

Prevalence and Transition Probabilities

Data on the prevalence and incidence of DR, DME, and STDR were collected from published studies in Shanghai [15,18]. Because Jin et al's [18] study only reported the 5-year incidence of STDR and DME, the 1-year incidence was calculated based on the formula: $r = -\log(1 - p)/t$, where r represents the 1-year incidence and p represents the cumulative incidence over time interval t [19]. Other transition probabilities were obtained from published studies specific to China, and if few data were available for Chinese patients, data from other Asian regions were used. We searched PubMed and China National Knowledge Infrastructure by using the following combinations of terms: "diabetic retinopathy" AND "progression" OR "transition" AND "Chinese" OR "China."

Screening and Intervention Costs

Our study included both direct and indirect costs and analyzed them from a societal perspective. Direct medical costs comprised the charges of screening, examination, and treatment. Direct nonmedical costs consisted of transportation costs related to hospital visits, and indirect costs consisted of family members' time associated with the visits and their wage loss. All costs were collected in Chinese yuan and then converted into US dollars at an exchange rate of CNY 6.90 per dollar [20]. All cost data are listed in Multimedia Appendices 4 and 6-8.

The screening costs were determined based on the Shanghai Digital Eye Disease Screening Program. The screening costs consisted of the purchase and maintenance costs of equipment, labor costs of medical personnel, transportation fees, and income loss for residents. We calculate the annualized cost for fixed assets by assuming a life span of 5 years and no salvage value. The construction and maintenance costs of the telemedicine platform were based on the Shanghai Digital Eye Disease Screening Program. Based on our field observations, it took 6.2, 3, 3.3, and 4.8 minutes on average for 1 participant to complete registration, visual acuity test autorefraction, and retinal photography, respectively. Theoretically, a team with 4 optometrists could screen approximately 100 participants per day, but under real-world working conditions, this is nearly 30 per day. As the participants in our model were older than 65 years, we assumed that they did not incur wage loss. Moreover, we did not include wage loss for the accompanying family members in the screening costs. Therefore, the total costs per person for manual grading–based and AI-based telemedicine screening were US $10.10 and US $9.60, respectively. Multimedia Appendix 6 shows the detailed composition of the screening costs.

To calculate the costs of the detailed re-examinations after referral, direct medical costs consisted of the costs of ocular examinations and equipment and wages for medical personnel; direct nonmedical costs comprised transportation fees related to the visits; and indirect costs included 1 accompanying family member's wage loss for time spent and per capita daily income in Shanghai in 2020. The examination costs were the unified pricing of the Shanghai Municipal Health Commission. Because public hospitals are nonprofit institutions, the money from these fees is mainly used to subsidize the cost of health care services. Hence, prices in public hospitals can be used to estimate the direct medical costs. Detailed information on the hospital-based examination costs is provided in Multimedia Appendix 7. It was assumed that the wage loss of the accompanying family member for referral was 0 because the majority of them were older than 65 years.

For treatment costs, direct medical costs included the costs of treatment, equipment, and wages for medical personnel; direct nonmedical costs consisted of costs of transportation related to the visits; and indirect costs included 1 accompanying family member's wage loss based on time spent and per capita daily income in Shanghai in 2020. In the first year, patients with DME were assumed to have received 3 antivascular endothelial growth factor injections. Photocoagulation or vitrectomy was administered to patients with severe nonproliferative DR or proliferative DR. In the follow-up years, an average of 1 antivascular endothelial growth factor injection was administered, and an annual outpatient review was required for patients with STDR. Direct medical costs were estimated using the prices of health care services in public hospitals. The total economic burden for blind patients in the first year was estimated to be US $8920, which included 53.2% direct medical costs, 6.4% direct nonmedical costs, and 40.4% indirect costs (loss of labor resources for family members and low-vision services costs), and there were only indirect costs in the follow-up years [5,21]. Detailed information on the treatment costs is provided in Multimedia Appendix 8.

Utility and QALYs

We estimated the utility values for each DR stage (seen in Multimedia Appendix 9) to calculate QALYs. Utility values were based on published studies from China and other Asian countries [21,22]. Because the residents who participated in the screening should have diabetes, utility was assumed to be 0.87 but not 1.0 for people without DR, 0.79 for those with non-STDR, and 0.7 for those with STDR (including severe nonproliferative DR, proliferative DR, and DME). The utility value for people with blindness was assumed to be 0.55 [22]. All the values for the base case and sensitivity analyses are listed in Multimedia Appendix 3.

Compliance

Compliance with referral to specialized ophthalmic hospitals or tertiary hospitals for a full examination among patients screened for signs of STDR was assumed to be 50.4% for manual grading–based telemedicine screening, according to our investigation in Shanghai [23]. However, compliance with AI-based telemedicine screening was unclear. Because only 1 published study suggested that adopting an AI-assisted diagnosis model in DR screening may impact the participants' adherence to ophthalmic care [24], the evidence is insufficient. Therefore, we assumed that compliance with referral in AI-based telemedicine screening was the same as that in manual grading–based telemedicine screening, while we set a wide range (±25%) for sensitivity analysis (Multimedia Appendix 3).

Screening Accuracy

The accuracy of AI-based telemedicine screening was extracted from published studies specific to the AI-assisted screening model conducted in Shanghai based on the current dominant architecture of convolutional neural networks (Multimedia Appendix 10) [25]. Briefly, the sensitivity was 80.47% (95% CI 75.07%-85.14%) and the specificity was 97.96% (95% CI 96.75%-98.81%) for STDR [25]. In our screening program, 2 experienced ophthalmologists were employed to make the diagnoses based on the retinal images. Furthermore, the accuracy of the manual grading–based telemedicine screening was assumed to be 100%, which was in accordance with the DR diagnosis criteria [8,9,26,27]. However, as described in some other studies, since trained graders instead of ophthalmologists performed the grading and diagnosis [5,7] in the sensitivity analysis, we adopted the accuracy range of the manual grading based on the Singaporean study (Multimedia Appendix 3) [7].

Other Parameters

The natural age-specific mortality rates of the general Chinese population reported by Zhang and Wei [28] were used in this study. Increased odds of mortality were assumed for people without DR but with DM, non-STDR, STDR, and blindness (Multimedia Appendix 3) [29,30]. Both costs and health state utilities were discounted at a 3.5% annual rate in the base analysis, following the National Institute for Health and Care Excellence recommendations [31]. For the cost-effectiveness threshold, 2 thresholds representing cost-effectiveness and high cost-effectiveness were used according to the World Health Organization recommendations [5,21]. Among the interventions improving the patients' utilities, those that cost less than the gross domestic product (GDP) per capita are defined as highly cost-effective, those that cost 1-3 times the GDP per capita are defined as cost-effective, and those that cost more than 3 times the GDP per capita are determined as not cost-effective [32]. On the contrary, among the interventions reducing the participants' utilities, among those saving costs higher than 3 times, the GDP per capita was defined as highly cost-effective; among those saving between 1 and 3 times, the GDP per capita was defined as cost-effective; and those costing less than the GDP per capita were determined as not cost-effective [32]. As the GDP per capita in Shanghai in 2020 was reported to be US $22,600, the thresholds in this study were defined as US $22,600 and US $67,800 [33].

Outcomes

The ICER and ICUR were calculated as the difference in the total costs between the AI-assisted and manual grading telemedicine screening divided by the difference in the total years without blindness and the QALYs between the 2 conditions, respectively. Values for the AI-assisted screening cohort minus those for the manual grading screening cohort, which were set as the baseline, were calculated as the differences.

Sensitivity Analysis

Extensive 1-way deterministic and probabilistic sensitivity analyses were performed to calculate the uncertainties of the base-case results. A variation of 10% was adopted because probability-related statistics (ie, utility, prevalence, sensitivity, specificity, transition probability, and compliance) were mainly derived from previously published studies. For the influence of AI use on compliance with referral, a range of 25% was used. A large floating range of 50% was adopted for these costs. In addition, we adopted the accuracy range for manual grading according to the Singaporean study (Multimedia Appendix 3) to account for the influence of trained graders performing the grading and diagnosis instead of retinal experts [7]. A probabilistic sensitivity analysis was conducted using Monte Carlo simulation for 10,000 simulations to assess the robustness of the base case analysis. Beta distributions were adopted for probability-related data and utility values, gamma distributions were used for costs, and log-normal distributions were used for odds ratios. The methods and results conformed to the Consolidated Health Economic Evaluation Reporting Standards (2022) (Multimedia Appendix 11).

Ethics Approval

This study was mainly based on the secondary analyses of published data. Written informed consent was obtained from all the participants. All the study data were anonymous. There were no compensation fees for the participants. This study was approved by the Institutional Review Board of the Shanghai General Hospital (2022SQ272) and Shanghai Eye Diseases Prevention and Treatment Center (2022SQ007).

Results

The cost-effectiveness and cost-utility analyses showed that AI-based telemedicine screening was dominated by manual grading–based telemedicine screening in Shanghai (Table 2). In the manual grading–based telemedicine screening, a community resident with DM would incur a total cost of US $3265.40, including screening, hospital referral for confirmation, and treatment as needed, with 9.83 years without blindness and 6.753 QALYs. In the AI-based telemedicine screening, a community resident with DM would incur a total cost of US $3182.50, with 9.80 years without blindness and 6.748 QALYs. Therefore, compared with the cost of manual grading–based telemedicine screening, that of the AI-based telemedicine screening model was 2.5% lower, while the years without blindness was 0.3% less, and the QALYs were 0.1% less.

Table 2. Base-case cost-effectiveness and cost-utility results[a].

	Costs per person (USD)	Incremental costs per 100,000 people screened (USD)	Years without blindness per person	Incremental years without blindness per 100,000 people screened	QALY[b] per person	Incremental quality-adjusted life years per 100,000 people screened	ICER[c] (USD)	ICUR[d] (USD)
AI[e]-assisted model	3182.47	–8,289,840.65	9.80	–3121.32	6.748	–544.78	2553.39	15,216.96
Manual grading	3265.37	N/A[f]	9.83	N/A	6.753	N/A	N/A	N/A

[a] Costs, years without blindness, and quality-adjusted life years are lifetime values per person, whereas incremental costs, incremental years without blindness, incremental cost-effectiveness ratio, incremental quality-adjusted life years, and incremental cost-utility ratio are calculated against the manual grading–based telemedicine screening scenario per 100,000 people screened.

[b] QALY: quality-adjusted life year.

[c] ICER: incremental cost-effectiveness ratio.

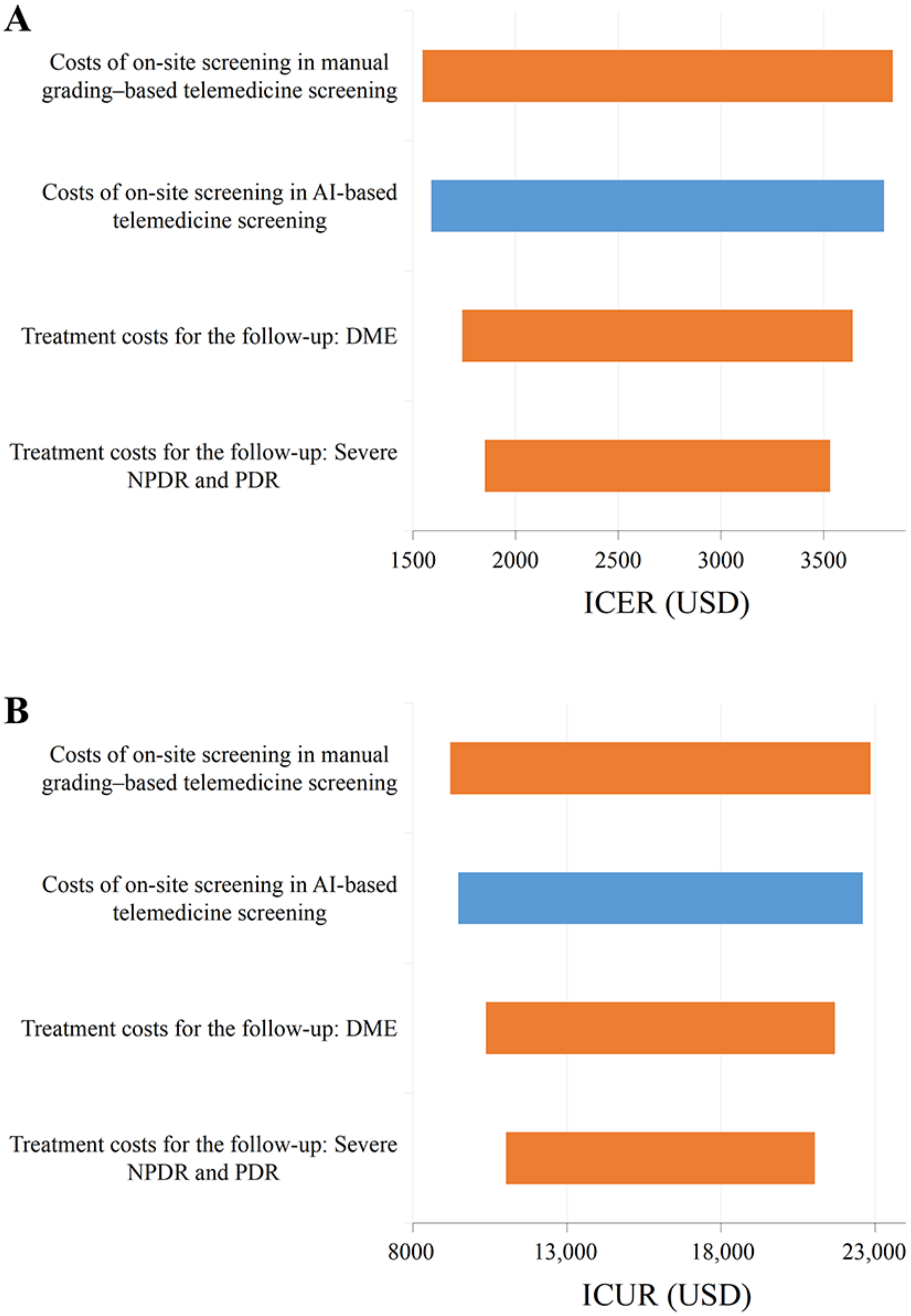
[d] ICUR: incremental cost-utility ratio.

[e] AI: artificial intelligence.

[f] N/A: not applicable.

Our results showed that by replacing manual grading–based telemedicine screening with AI-based telemedicine screening, 1 participant could save US $15,216.96 but needed to lose 1 more QALY (ICUR=US $15,216.96), indicating that AI-based telemedicine screening was not cost-effective as in Shanghai in 2020; at least US $22,600 (GDP per capita) should be saved if 1 more QALY is lost due to the shift in interventions. A 1-way deterministic sensitivity analysis of cost-effectiveness and cost-utility analyses indicated that the impact of the adoption of AI on compliance with referral, costs of on-site screening in manual grading–based telemedicine screening, costs of on-site screening in AI-based telemedicine screening, treatment costs for the follow-up of patients with DME, and treatment costs for the follow-up of patients with severe nonproliferative DR and proliferative DR were the 5 most influential variables. In particular, according to the cost-utility analysis, if the adoption of AI could improve compliance with referrals by 7.5%, the AI-assisted model might be cost-effective; if compliance was improved by 17.5%, the AI-assisted model might be highly cost-effective; and if compliance was improved by 25%, the AI-assisted model might be the absolutely dominant strategy, as it could save costs and increase the years without blindness and QALYs (Multimedia Appendix 12). Moreover, the increase in the costs of on-site screening in manual grading–based telemedicine screening and the decrease in the costs of on-site screening in AI-based telemedicine screening might help the AI-based telemedicine screening to be cost-effective (Figure 1). The detailed sensitivity analysis results of the other parameters are shown in Multimedia Appendices 13 and 14.

Figure 1. One-way deterministic sensitivity analysis (Tornado diagram). A. One-way sensitivity analysis of cost-effectiveness. B. One-way sensitivity analysis of cost-utility. Since negative values of incremental cost-effectiveness ratio or incremental cost-utility ratio might occur due to the change of compliance with referral after the adoption of artificial intelligence (multiplier), detailed results have been shown in Multimedia Appendix 12 separately. Therefore, in this Tornado diagram, the impact of the change of compliance with referral after the adoption of artificial intelligence (multiplier) is not shown. AI: artificial intelligence; DME: diabetic macular edema; ICER: incremental cost-effectiveness ratio; ICUR: incremental cost-utility ratio; NPDR: nonproliferative diabetic retinopathy; PDR: proliferative diabetic retinopathy.



Probabilistic sensitivity analysis showed that the base-case ICER and ICUR were robust to randomly distributed parameters (Figure 2). We obtained cost-effectiveness acceptability curves by taking 10,000 random draws (Figure 3). This means that when both AI-based and manual grading–based telemedicine screening were available, manual grading–based telemedicine screening was the dominant strategy in 60.6% of the simulations under the threshold of GDP per capita (US $22,600) and in 84.5% of the simulations under the threshold of 3 times the GDP per capita (US $67,800).

Figure 2. Probabilistic sensitivity analysis. A. Probabilistic sensitivity analysis of cost-effectiveness. B. Probabilistic sensitivity analysis of cost-utility. GDP: gross domestic product.

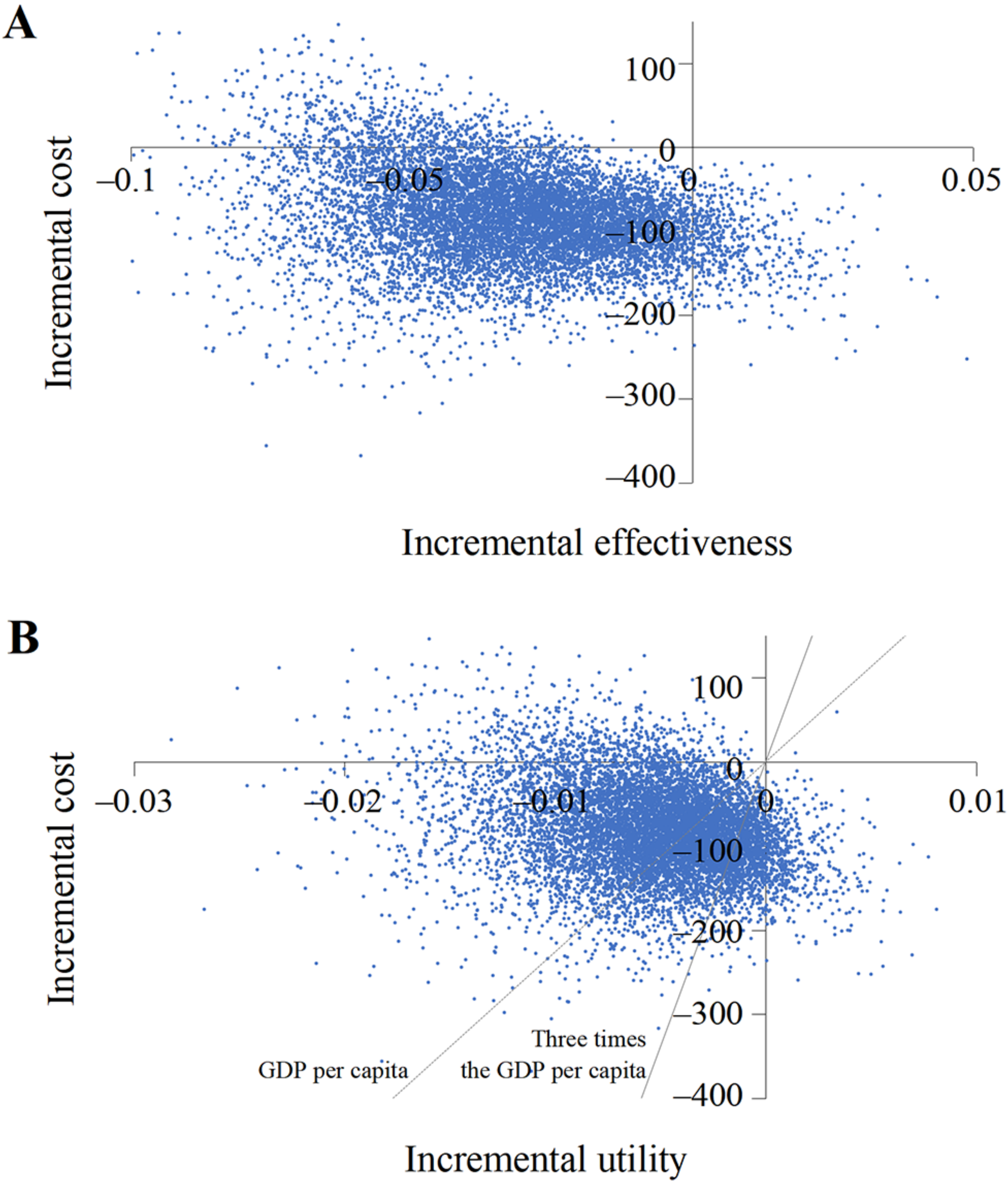
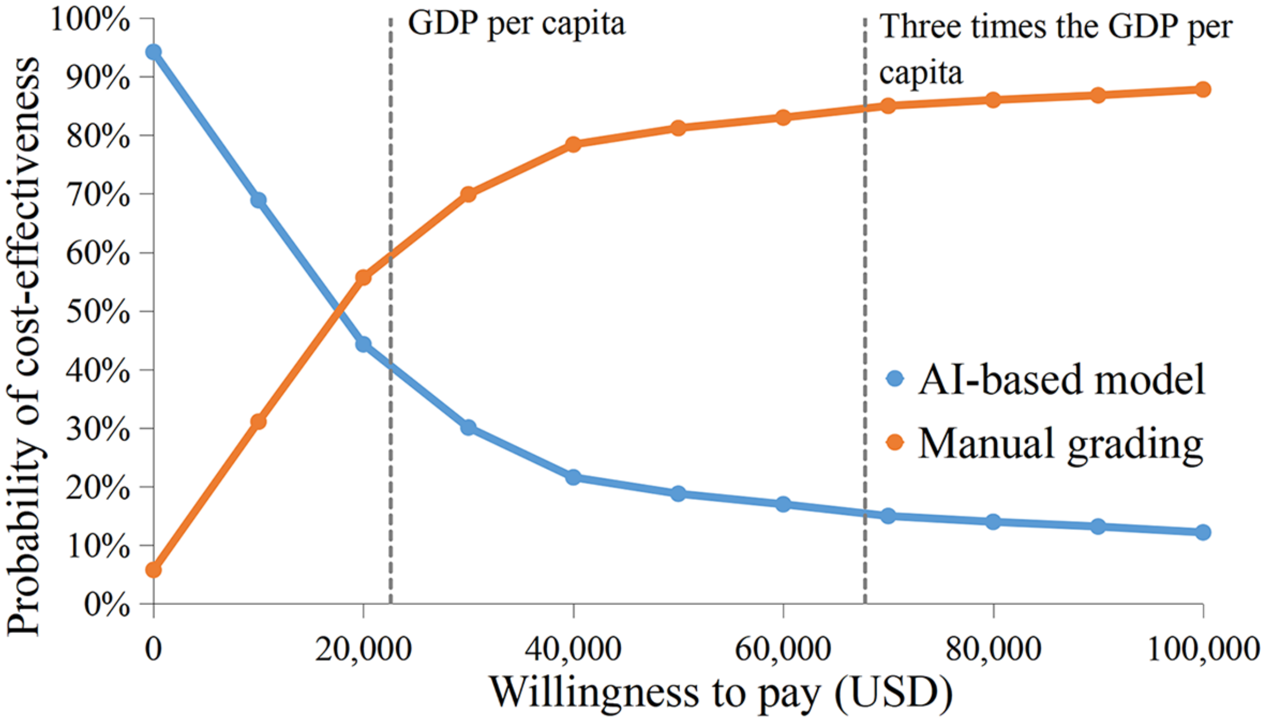


Figure 3. Cost-effectiveness acceptability curves. AI: artificial intelligence; GDP: gross domestic product.



Discussion

Principal Findings

This study presents one of the first health economic evaluations in the context of low labor costs in an LMIC setting of competing telemedicine models for community-based DR screening using manual grading and AI models. In line with previous studies [7,8], our analysis was based on an established telemedicine screening program. We showed that, in this context, the value of AI-based telemedicine DR screening depended heavily on the referral compliance of patients with suspected STDR. If this compliance did not increase, AI-based telemedicine DR screening would not be more cost-effective than manual grading–based telemedicine DR screening because it would decrease the long-term screening effectiveness and individuals' health utility and not save enough costs.

Prior studies in Singapore and the United Kingdom [7,8] showed that replacing manual grading with an AI model for DR screening led to a 12%-20% cost reduction. A study in Scotland reported an even greater cost reduction of 46.7% [9]. These studies were conducted in high-income countries in the context of high labor costs. However, China's national conditions differ from those in high-income countries. One of the most important differences is that the labor costs of the medical staff are much lower in China. For example, in Singapore, the labor cost for DR grading is US $26 per participant [7], which is over 20 times that of the labor cost for manual grading in Shanghai. Consequently, in the context of low labor costs, a reduction in screening costs resulting from the use of AI solutions is limited in China.

In the Shanghai program, the on-site screening cost for 1 participant, including screening examinations and diagnosis, was US $10.10 in the manual grading model and US $9.60 in the AI-assisted model—a reduction of only 5%. Moreover, the labor costs of the medical staff in our screening program were among the highest in China. For example, according to the Wenzhou ophthalmologic screening program, the labor cost for on-site screening examinations and diagnosis of glaucoma was US $1.70 per participant, which was about a quarter of the labor costs in our Shanghai program (US $7.20 per participant for manual grading–based telemedicine screening, Multimedia Appendix 6) [21], and according to the Finance Department and Procurement Center of Beijing Tongren Hospital, Beijing Tongren Eye Centre Ocular Reading Centre, and China Intelligent Ophthalmology Big Data Research Center, the labor costs for on-site screening examinations and diagnosis were US $1.75 per participant for traditional face-to-face screening and US $0.80 per participant for manual grading–based telemedicine screening [5], which is only one-ninth of the labor costs in our Shanghai program (US $7.20 per participant for manual grading–based telemedicine screening, Multimedia Appendix 6). Because the main difference between the cost components of the AI-assisted model and those of the manual grading model is that equipment and telemedicine platform costs replace labor costs, in settings where labor costs are extremely low, cost reduction via the adoption of AI is expected to be even lesser. Therefore, AI-assisted DR screening is less cost-effective in other urban areas of China.

Our sensitivity analysis confirmed these results. The on-site screening costs of both manual grading and AI-assisted models are among the most influential variables. An increase in the on-site screening costs of manual grading by 50% or a decrease in the on-site screening costs of the AI-assisted model by 50% may help AI-based telemedicine screening be cost-effective. In other words, the gap between the on-site screening costs of the manual grading model and the AI-assisted model is the key point. Moreover, the cost of AI software was only 7% of the on-site costs in AI-based telemedicine screening (Multimedia Appendix 6). Therefore, even if the AI software were completely free, a 50% reduction in the on-site screening costs of the AI-assisted model would not be achieved. As a result, unless the labor costs of medical staff increase dramatically in the future, the AI-assisted model will be hardly cost-effective in Shanghai, holding all the other conditions constant.

However, there was one exception to this. Our sensitivity analysis shows that if the referral compliance of patients with suspected STDR increased after the adoption of AI even to a small extent, then the AI-assisted model would be cost-effective. A study in Missouri [24] suggested that after the adoption of the automated retinal image assessment system, which is based on AI, the rates of completed referral eye examinations at 3, 6, and 12 months after screening increased from 9.4% to 32.6%, from 13.4% to 46.7%, and from 18.7% to 55.4%, respectively [24]. However, relevant evidence is still inadequate; therefore, it is difficult to determine whether this improvement is an isolated case. Previously, we implemented a discrete choice experiment in Shanghai to measure individuals' preferences for AI-based screening [34]. The results suggested that the impact of the adoption of AI on individuals' preferences may be bidirectional. On the one hand, algorithm aversion should be noted, which means that compared to manual grading, the residents were in disfavor of the AI-assisted screening technology [34,35]. On the other hand, the immediate feedback of retinal screening results by the adoption of AI could increase the individuals' preferences and have profound effects on participants' follow-up behavior [24,34]. Nevertheless, there is still a lack of empirical studies on the association between the results of feedback efficiency and residents' referral compliance.

This study has several strengths. This study may provide a reference for policy making in planning community-based DR screening in LMICs by modelling 2 practical telemedicine screening models for DR by using real-world data from an ongoing program in urban China. In addition, we conducted a sensitivity analysis of our models within wide ranges and identified the most influential variables affecting the decision to use AI and manual grading in telemedicine screening. Therefore, our conclusions provide practical value in the policy-making process regarding when to deploy AI-assisted diagnostic technology.

Limitations

Our study had several limitations despite its numerous strengths. Most notably, we only compared the models for centralized screening. Other models must also be considered going forward. For example, in Shanghai, some community health service centers are beginning to provide DR screening as part of their outpatient services for patients with diabetes. This change in the model may impact both the costs and patients' compliance, thus altering the results of health economic evaluations such as ours. Second, our comparison is based on the premise that both human- and AI-based models are available and affordable. However, in some remote regions, due to the lack of human resources, manual screening for eye disease may be impractical, and AI-based screening, if available, may be the only option. Third, our study is mainly based on empirical data from Shanghai; therefore, it cannot be representative of the whole of China because of the huge regional and medical care differences between urban and rural areas. Therefore, there is an urgent need for more extensive and in-depth studies. However, as we have discussed above, the labor costs of medical staff in Shanghai are among the highest in China, and AI-based telemedicine screening will become even more less cost-effective if the labor cost of medical staff is further reduced. Therefore, our findings can be extrapolated within the Chinese context.

Conclusion

Our study may provide a reference for policy making in planning community-based telemedicine screening for DR in LMICs. Our findings indicate that unless the referral compliance of patients with suspected STDR increases, the adoption of the AI model may not further improve the value of telemedicine screening compared to that of manual grading in LMICs. The main reason is that in the context of low labor costs, the direct health care costs saved by replacing manual grading with AI are limited, and screening effectiveness will decrease. In conclusion, our study suggests that the magnitude of the value generated by this technology replacement depends mainly on 2 aspects. The first is the extent of direct health care costs reduced by using AI, and the second is the change in health care service utilization caused by using AI. Therefore, our research also provides analytical ideas for other health care sectors in addition to eye care when deciding whether to use AI.

Acknowledgments

This study was funded by the Chinese National Key Research and Development Program (project 2021YFC2702100), Chinese National Nature Science Foundation (Project 82071012), Shanghai Engineering Research Center of Precise Diagnosis and Treatment of Eye Diseases, Shanghai, China (project 19DZ2250100), Shanghai Hospital Development Center (SHDC12021613), Shanghai Municipal Health Commission (2022HP61, 2022YQ051), and Shanghai Eye Diseases Control and Treatment Center (20LC01002).

Authors' Contributions

SL, YM, and HZ conceptualized this study and the methodology. SL and YM developed the software and prepared the original draft of this manuscript. SL, YX, LL, JH, JZ, YP, and TY conducted the investigations. SL, JH, YP, and TY performed the data curation. YX, LL, NC, and HZ reviewed and edited the paper. LL, JZ, and HZ supervised this study. SL, LL, and HZ acquired the funds. All authors have read and agreed to the published version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prevalence and uncertainty ranges of different stages of diabetic retinopathy.
[DOCX File , 13 KB - publichealth_v9i1e41624_app1.docx]

Multimedia Appendix 2
Variation range and distributions assumed for the transitional probabilities before and after treatment for the different diabetic retinopathy stages.
[DOCX File , 14 KB - publichealth_v9i1e41624_app2.docx]

Multimedia Appendix 3
Variation range and distributions assumed for compliance, utilization, mortality, and other parameters.
[DOCX File , 20 KB - publichealth_v9i1e41624_app3.docx]

Multimedia Appendix 4
Variation range and distributions assumed for screening cost and medical cost of treating diabetic retinopathy at different stages.
[DOCX File , 13 KB - publichealth_v9i1e41624_app4.docx]

Multimedia Appendix 5
Care pathways for both artificial intelligence–based and manual grading–based telemedicine screening.
[DOCX File , 53 KB - publichealth_v9i1e41624_app5.docx]

Multimedia Appendix 6
Cost composition of manual grading–based and artificial intelligence–based telemedicine screening.
[DOCX File , 19 KB - publichealth_v9i1e41624_app6.docx]

Multimedia Appendix 7
Cost composition of full examination.
[DOCX File , 13 KB - publichealth_v9i1e41624_app7.docx]

Multimedia Appendix 8
Cost composition of treatment for patients with sight-threatening diabetic retinopathy.
[DOCX File , 15 KB - publichealth_v9i1e41624_app8.docx]

Multimedia Appendix 9
Markov model for diabetic retinopathy.
[DOCX File , 87 KB - publichealth_v9i1e41624_app9.docx]

Multimedia Appendix 10
Accuracy of the artificial intelligence–assisted model used in Shanghai.
[DOCX File , 13 KB - publichealth_v9i1e41624_app10.docx]

Multimedia Appendix 11
The Consolidated Health Economic Evaluation Reporting Standards 2022 checklist.
[DOCX File , 17 KB - publichealth_v9i1e41624_app11.docx]

Multimedia Appendix 12
Deterministic sensitivity analysis results of change of compliance with referral after the adoption of artificial intelligence (multiplier).
[DOCX File , 32 KB - publichealth_v9i1e41624_app12.docx]

Multimedia Appendix 13
Sensitivity analysis results for incremental cost-effectiveness ratio.
[DOCX File , 16 KB - publichealth_v9i1e41624_app13.docx]

Multimedia Appendix 14
Sensitivity analysis results for incremental cost-utility ratio.
[DOCX File , 16 KB - publichealth_v9i1e41624_app14.docx]

References

1. Jampol LM, Glassman AR, Sun J. Evaluation and Care of Patients with Diabetic Retinopathy. N Engl J Med 2020 Apr 23;382(17):1629-1637. [doi: 10.1056/nejmra1909637]
2. Teo ZL, Tham Y, Yu M, Chee ML, Rim TH, Cheung N, et al. Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045: Systematic Review and Meta-analysis. Ophthalmology 2021 Nov;128(11):1580-1591. [doi: 10.1016/j.ophtha.2021.04.027] [Medline: 33940045]
3. Vujosevic S, Aldington SJ, Silva P, Hernández C, Scanlon P, Peto T, et al. Screening for diabetic retinopathy: new perspectives and challenges. Lancet Diabetes Endocrinol 2020 Apr;8(4):337-347 [FREE Full text] [doi: 10.1016/S2213-8587(19)30411-5] [Medline: 32113513]
4. Siegel KR, Ali MK, Zhou X, Ng BP, Jawanda S, Proia K, et al. Cost-effectiveness of Interventions to Manage Diabetes: Has the Evidence Changed Since 2008? Diabetes Care 2020 Jul;43(7):1557-1592. [doi: 10.2337/dci20-0017] [Medline: 33534729]
5. Li R, Yang Z, Zhang Y, Bai W, Du Y, Sun R, et al. Cost-effectiveness and cost-utility of traditional and telemedicine combined population-based age-related macular degeneration and diabetic retinopathy screening in rural and urban China. Lancet Reg Health West Pac 2022 Jun;23:100435 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100435] [Medline: 35355615]
6. Avidor D, Loewenstein A, Waisbourd M, Nutman A. Cost-effectiveness of diabetic retinopathy screening programs using telemedicine: a systematic review. Cost Eff Resour Alloc 2020;18:16 [FREE Full text] [doi: 10.1186/s12962-020-00211-1] [Medline: 32280309]
7. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. The Lancet Digital Health 2020 May;2(5):e240-e249. [doi: 10.1016/s2589-7500(20)30060-1]
8. Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, et al. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. Ophthalmology 2017 Mar;124(3):343-351 [FREE Full text] [doi: 10.1016/j.ophtha.2016.11.014] [Medline: 28024825]
9. Scotland GS, McNamee P, Philip S, Fleming AD, Goatman KA, Prescott GJ, et al. Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland. Br J Ophthalmol 2007 Nov;91(11):1518-1523 [FREE Full text] [doi: 10.1136/bjo.2007.120972] [Medline: 17585001]
10. Dismuke C. Progress in examining cost-effectiveness of AI in diabetic retinopathy screening. The Lancet Digital Health 2020 May;2(5):e212-e213. [doi: 10.1016/s2589-7500(20)30077-7]
11. Ping L, Yuanyuan W. Investigation on the prevalence and influencing factors of type 2 diabetes among adults in Shanghai and research on prevention and treatment strategies. Chinese Primary Health Care. 2020. URL: https://caod.oriprobe.com/articles/59098826/Investigation_on_the_Prevalence_and_Influencing_Fa.htm [accessed 2023-01-26]
12. Xu Y, Wang Y, Liu B, Tang L, Lv L, Ke X, et al. The diagnostic accuracy of an intelligent and automated fundus disease image assessment system with lesion quantitative function (SmartEye) in diabetic patients. BMC Ophthalmol 2019 Aug 14;19(1):184 [FREE Full text] [doi: 10.1186/s12886-019-1196-9] [Medline: 31412800]
13. Dai L, Wu L, Li H, Cai C, Wu Q, Kong H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun 2021 May 28;12(1):3242 [FREE Full text] [doi: 10.1038/s41467-021-23458-5] [Medline: 34050158]
14. Li B, Chen H, Zhang B, Yuan M, Jin X, Lei B, et al. Development and evaluation of a deep learning model for the detection of multiple fundus diseases based on colour fundus photography. Br J Ophthalmol 2022 Aug;106(8):1079-1086. [doi: 10.1136/bjophthalmol-2020-316290] [Medline: 33785508]
15. He J, Xu X, Zhu J, Zhu B, Zhang B, Lu L, et al. Lens Power, Axial Length-to-Corneal Radius Ratio, and Association with Diabetic Retinopathy in the Adult Population with Type 2 Diabetes. Ophthalmology 2017 Mar;124(3):326-335. [doi: 10.1016/j.ophtha.2016.10.041] [Medline: 27993483]
16. Lian JX, Gangwani RA, McGhee SM, Chan CKW, Lam CLK, Primary Health Care Group,, et al. Systematic screening for diabetic retinopathy (DR) in Hong Kong: prevalence of DR and visual impairment among diabetic population. Br J Ophthalmol 2016 Feb;100(2):151-155. [doi: 10.1136/bjophthalmol-2015-307382] [Medline: 26271268]
17. Harding S, Greenwood R, Aldington S, Gibson J, Owens D, Taylor R, Diabetic Retinopathy GradingDisease Management Working Party. Grading and disease management in national screening for diabetic retinopathy in England and Wales. Diabet Med 2003 Dec;20(12):965-971. [doi: 10.1111/j.1464-5491.2003.01077.x] [Medline: 14632697]

18. Jin P, Peng J, Zou H, wang W, Fu J, Bai X, et al. A 5-year prospective study of type 2 diabetes patients in Shanghai Xinjing Community, the incidence and risk factors of diabetic retinopathy and diabetic macular edema in Chinese type 2 diabetes residents. Chin J Exp Ophthalmol (in Chinese). 2016. URL: https://tinyurl.com/6ywjj3w9 [accessed 2023-01-26]
19. Rothman K. Epidemiology: An Introduction, Second Edition. New York: Oxford University Press; 2012.
20. Central parity rate of CNY (historical data). State Administration of Foreign Exchange. URL: https://www.safe.gov.cn/safe/2020/1218/17833.html [accessed 2022-05-08]
21. Tang J, Liang Y, O'Neill C, Kee F, Jiang J, Congdon N. Cost-effectiveness and cost-utility of population-based glaucoma screening in China: a decision-analytic Markov model. The Lancet Global Health 2019 Jul;7(7):e968-e978. [doi: 10.1016/s2214-109x(19)30201-3]
22. Nguyen HV, Tan GSW, Tapp RJ, Mital S, Ting DSW, Wong HT, et al. Cost-effectiveness of a National Telemedicine Diabetic Retinopathy Screening Program in Singapore. Ophthalmology 2016 Dec;123(12):2571-2580. [doi: 10.1016/j.ophtha.2016.08.021] [Medline: 27726962]
23. Zhu X, Xu Y, Lu L, Zou H. Patients' perspectives on the barriers to referral after telescreening for diabetic retinopathy in communities. BMJ Open Diabetes Res Care 2020 Mar;8(1):e000970 [FREE Full text] [doi: 10.1136/bmjdrc-2019-000970] [Medline: 32193199]
24. Liu J, Gibson E, Ramchal S, Shankar V, Piggott K, Sychev Y, et al. Diabetic Retinopathy Screening with Automated Retinal Image Analysis in a Primary Care Setting Improves Adherence to Ophthalmic Care. Ophthalmol Retina 2021 Jan;5(1):71-77 [FREE Full text] [doi: 10.1016/j.oret.2020.06.016] [Medline: 32562885]
25. Zhang W, Li D, Wei Q, Ding D, Meng L, Wang Y, et al. The Validation of Deep Learning-Based Grading Model for Diabetic Retinopathy. Front Med (Lausanne) 2022;9:839088 [FREE Full text] [doi: 10.3389/fmed.2022.839088] [Medline: 35652075]
26. Lee AY, Yanagihara RT, Lee CS, Blazes M, Jung HC, Chee YE, et al. Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. Diabetes Care 2021 May;44(5):1168-1175 [FREE Full text] [doi: 10.2337/dc20-1877] [Medline: 33402366]
27. Lin D, Xiong J, Liu C, Zhao L, Li Z, Yu S, et al. Application of Comprehensive Artificial intelligence Retinal Expert (CARE) system: a national real-world evidence study. The Lancet Digital Health 2021 Aug;3(8):e486-e495. [doi: 10.1016/s2589-7500(21)00086-8]
28. Zhang W, Wei M. The Evaluation of the Mortality and Life Expectancy of Chinese Population. Population Journal. (in Chinese) 2016;38(3):18-28 [FREE Full text] [doi: 10.16405/j.cnki.1004-129X.2016.03.002]
29. Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. JAMA 2000 Feb 16;283(7):889-896. [doi: 10.1001/jama.283.7.889] [Medline: 10685713]
30. Bragg F, Holmes MV, Iona A, Guo Y, Du H, Chen Y, China Kadoorie Biobank Collaborative Group. Association Between Diabetes and Cause-Specific Mortality in Rural and Urban Areas of China. JAMA 2017 Jan 17;317(3):280-289 [FREE Full text] [doi: 10.1001/jama.2016.19720] [Medline: 28114552]
31. Guide to the methods of technology appraisal 2013. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/process/pmg9/chapter/developing-the-scope [accessed 2022-05-07]
32. Al MJ. Cost-effectiveness acceptability curves revisited. Pharmacoeconomics 2013 Feb;31(2):93-100. [doi: 10.1007/s40273-012-0011-8] [Medline: 23329426]
33. GDP per capita in Shanghai (1978-2020). Shanghai Municipal Statistics Bureau. URL: https://tjj.sh.gov.cn/tjnj/nj21.htm?d1=2021tjnj/C0305.htm [accessed 2022-05-01]
34. Lin S, Li L, Zou H, Xu Y, Lu L. Medical Staff and Resident Preferences for Using Deep Learning in Eye Disease Screening: Discrete Choice Experiment. J Med Internet Res 2022 Sep 20;24(9):e40249 [FREE Full text] [doi: 10.2196/40249] [Medline: 36125854]
35. Dietvorst BJ, Simmons JP, Massey C. Algorithm aversion: people erroneously avoid algorithms after seeing them err. J Exp Psychol Gen 2015 Feb;144(1):114-126. [doi: 10.1037/xge0000033] [Medline: 25401381]

Abbreviations

AI: artificial intelligence
DM: diabetes mellitus
DME: diabetic macular edema
DR: diabetic retinopathy
GDP: gross domestic product
ICER: incremental cost-effectiveness ratio
ICUR: incremental cost-utility ratio
LMICs: low- and middle-income countries
QALY: quality-adjusted life year
STDR: sight-threatening diabetic retinopathy

Edited by A Mavragani; submitted 02.08.22; peer-reviewed by Z Hou, N Wang; comments to author 24.10.22; revised version received 13.11.22; accepted 12.01.23; published 23.02.23.

Please cite as:
Lin S, Ma Y, Xu Y, Lu L, He J, Zhu J, Peng Y, Yu T, Congdon N, Zou H
Artificial Intelligence in Community-Based Diabetic Retinopathy Telemedicine Screening in Urban China: Cost-effectiveness and Cost-Utility Analyses With Real-world Data
JMIR Public Health Surveill 2023;9:e41624
URL: https://publichealth.jmir.org/2023/1/e41624
doi:10.2196/41624
PMID:36821353

Original Paper

Automating Case Reporting of Chlamydia and Gonorrhea to Public Health Authorities in Illinois Clinics: Implementation and Evaluation of Findings

Ninad Mishra[1*], MD, MS; Reynaldo Grant[1,2*], BA; Megan Toth Patel[2*], MPH; Siva Guntupalli[2*], MS; Andrew Hamilton[3*], RN, BSN, MS; Jeremy Carr[3*], BSc; Elizabeth McKnight[3*], MMI; Wendy Wise[4], MPH; David deRoode[4], MHA; Jim Jellison[5*], MPH; Natalie Viator Collins[5*], MPH; Alejandro Pérez[1*], MPH; Saugat Karki[1*], MD, MS

[1]Division of STD Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States
[2]Division of Infectious Diseases, Office of Health Protection, Illinois Department of Public Health, Springfield, IL, United States
[3]AllianceChicago, Chicago, IL, United States
[4]Lantana Consulting Group, East Thetford, VT, United States
[5]Public Health Informatics Institute, Atlanta, GA, United States
[*]these authors contributed equally

Corresponding Author:
Saugat Karki, MD, MS
Division of STD Prevention
Centers for Disease Control and Prevention
1600 Clifton Road
Atlanta, GA, 30329
United States
Phone: 1 4047187483
Email: oql7@cdc.gov

Abstract

Background: Chlamydia and gonorrhea cases continue to rise in Illinois, increasing by 16.4% and 70.9% in 2019, respectively, compared with 2015. Providers are required to report both chlamydia and gonorrhea, as mandated by public health laws. Manual reporting remains a huge burden; 90%-93% of cases were reported to Illinois Department of Public Health (IDPH) via electronic laboratory reporting (ELR), and the remaining were reported through web-based data entry platforms, faxes, and phone calls. However, cases reported via ELRs only contain information available to a laboratory facility and do not contain additional data needed for public health. Such data are typically found in an electronic health record (EHR). Electronic case reports (eCRs) were developed and automated the generation of case reports from EHRs to be reported to public health agencies.

Objective: Prior studies consolidated *trigger* criteria for eCRs, and compared with manual reporting, found it to be more complete. The goal of this project is to pilot standards-based eCR for chlamydia and gonorrhea. We evaluated the throughput, completeness, and timeliness of eCR compared to ELR, as well as the implementation experience at a large health center–controlled network in Illinois.

Methods: For this study, we selected 8 clinics located on the north, west, and south sides of Chicago to implement the eCRs; these cases were reported to IDPH. The study period was 52 days. The centralized EHR used by these clinics leveraged 2 of the 3 case detection scenarios, which were previously defined as the *trigger*, to generate an eCR. These messages were successfully transmitted via Health Level 7 electronic initial case report standard. Upon receipt by IDPH, these eCRs were parsed and housed in a staging database.

Results: During the study period, 183 eCRs representing 135 unique patients were received by IDPH. eCR reported 95% (n=113 cases) of all the chlamydia cases and 97% (n=70 cases) of all the gonorrhea cases reported from the participating clinical sites. eCR found an additional 14 (19%) cases of gonorrhea that were not reported via ELR. However, ELR reported an additional 6 cases of chlamydia and 2 cases of gonorrhea, which were not reported via eCR. ELR reported 100% of chlamydia cases but only 81% of gonorrhea cases. While key elements such as patient and provider names were complete in both eCR and ELR, eCR was found to report additional clinical data, including history of present illness, reason for visit, symptoms, diagnosis, and medications.

Conclusions: eCR successfully identified and created automated reports for chlamydia and gonorrhea cases in the implementing clinics in Illinois. eCR demonstrated a more complete case report and represents a promising future of reducing provider burden for reporting cases while achieving greater semantic interoperability between health care systems and public health.

(JMIR Public Health Surveill 2023;9:e38868) doi:10.2196/38868

KEYWORDS

public health surveillance; sexually transmitted diseases; gonorrhoea; chlamydia; electronic case reporting; eCR; health information interoperability; electronic health records; EHR; case reporting; automated; reporting; recording; patient records; cases; health care system; semantic; interoperability; implementation

Introduction

Background

Burden of Sexually Transmitted Infections in Illinois

Based on 2019 surveillance data, the disease burden for chlamydia and gonorrhea has increased by 16.4% and 70.9%, respectively, in Illinois since 2015 [1]. Public health laws mandate that these sexually transmitted infections be reported to Illinois Department of Public Health (IDPH) by health care providers and clinical laboratories within 7 business days [2]. Health care providers in Illinois generally report sexually transmitted infection cases via a web-based data entry platform; health systems and laboratories generally report their case results by electronic laboratory reporting (ELR). For many health care providers, reporting via the web-based platform poses an enormous workforce burden, and many report barriers to reporting such as lack of capacity, difficulty extracting relevant information, confusing vocabulary standards, etc [3]. Consequently, case reporting via electronic health records (EHRs) is a welcome prospect for all parties—reporting entities and public health agencies.

Public Health Surveillance and Informatics

Public health surveillance as a distinct discipline of public health was described by Langmuir [4] with the emphasis of ongoing and systemic collection of relevant data, consolidation, and analysis, followed by regular dissemination to all who need to know and can take public health action [5]. The role of public health informatics in facilitating surveillance is: “1) to improve timelines and completeness of data collection and analysis and 2) to free human resources to focus on the areas that require the most creative thought and to do the work that technology cannot” [6]. The Centers for Disease Control and Prevention (CDC) published the agency’s vision for public health surveillance in the 21st century [7] and recognized an opportunity to improve data quality and timeliness by accessing EHRs [8].

ELR and electronic case reporting (eCR) are creating a paradigm shift in public health surveillance. ELR is the automated messaging of laboratory reports of notifiable cases and has been widely adopted in the United States [9]. eCR is the automated generation of case reports from EHRs and subsequent reporting to public health agencies [10,11]. Both eCR and ELR were found to be more complete and timelier than paper-based reporting [12,13]. These approaches represent an advancement toward better semantic interoperability to support public health surveillance [14], greatly reducing the burden of reporting from clinical providers and improving the completeness and timeliness of those reports [15]. However, further evaluation and consideration is required to achieve a greater level of success and widespread implementation.

Prior Work in eCR Architecture for Chlamydia and Gonorrhea

The Council of State and Territorial Epidemiologists determines the case definition of notifiable conditions and maintains position statements for these conditions, including for chlamydia and gonorrhea [16-18]. An earlier study by Mishra et al [12] leveraged these position statements to develop the case detection logic for chlamydia and gonorrhea in EHRs. As an outcome of this study, value sets were created using national health care data standards (eg, International Classification of Diseases, Tenth Revision, Clinical Modifications; Logical Observation Identifiers Names and Codes [LOINC]; and Systematized Nomenclature of Medicine – Clinical Terms [SNOMED-CT]) [19] and Health Level 7 (HL7) [20]. Mishra et al [12] also compared eCR reporting with manual reporting and found that eCR increased provider reporting and improved the completeness of those case reports [12]. This study further consolidated the *trigger* for eCRs (when the case-detection logic is met in the EHR) based on the following three scenarios: (1) when an individual is diagnosed, (2) when a confirmatory laboratory result (named organism) is returned, or (3) a combination of laboratory test performed and result indicating the presence of infection is found (without the organism named in the result) [21]. Subsequently, this eCR architecture was implemented in Oregon simultaneously to our implementation in Illinois. In both the Oregon implementation and this study reporting the Illinois findings, the eCRs were generated via the HL7 electronic initial case report (eICR) standard [20,22]. The Oregon implementation found that this eCR architecture successfully reported cases of chlamydia and gonorrhea to public health while at the same time improving on the completeness compared to ELRs for the same cases [22].

Objective

The primary goal of this project was to pilot standards-based eCR for chlamydia and gonorrhea in 2 participating jurisdictions (public health agencies of the states of Illinois and Oregon), study the completeness of data between ELRs and eCRs, and disseminate the findings to promote adoption. In this process, we not only furthered the previous work by Mishra et al [12] but also learned about local variations while implementing the same architecture of eCR in 2 separate state health departments [12,22]. We recognize that local codes in EHRs are very common, and mapping them is a challenge [14]. Due to the challenges posed by variations in nonstandard codes used in EHRs, every implementation of eCR becomes an exercise of local implementation with some degree of customization. We further evaluated this implementation experience at a large health center–controlled network in Illinois. Additionally, we evaluated the completeness of eCR and compared it to ELR.

Methods

Ethical Review and Study Duration

For this evaluation, case-related information was sent from the EHR to IDPH, the State of Illinois's public health agency. An internal review committee of the participating clinics deemed that an Institutional Review Board review was not necessary for this project. The study period was a little over 7 weeks (or 52 days), which began on August 12, 2020, and ended on October 2, 2020.

Implementation Partners and Clinical Setting

Clinical Site and EHR Platform

We sought a network of clinics that shared a common platform of centrally hosted EHR. This would allow for scaling up across multiple clinical sites using a single implementation of the eCR architecture. AllianceChicago is a health center–controlled network that supports 45 ambulatory primary care practices in 19 states. For this evaluation, we selected 8 clinics operated by Near North Health, a federally qualified health center, located on the north, west, and south sides of Chicago [23]. In 2019, the selected clinical sites served 37,223 patients for a total of 122,277 visits. These clinical sites used AthenaHealth as the EHR [24] and QVERA as the interface engine [25].

Public Health Department

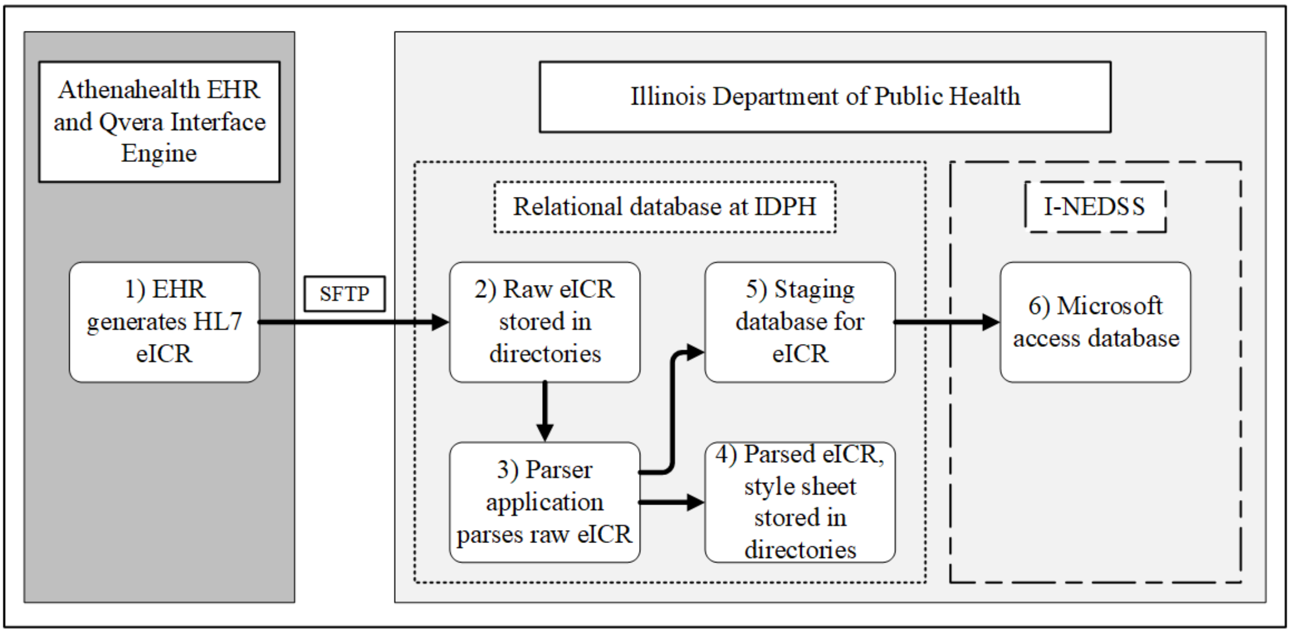
IDPH has a centralized IT infrastructure that uses the Illinois National Electronic Disease Surveillance System (I-NEDSS), a home-grown person-based and event-based system [26]. Health care providers and clinical laboratories are required to report all cases of chlamydia and gonorrhea within 7 working days [2,27]. Prior to this project, IDPH accepted reports via ELR, electronic provider report (a custom flat file generated by the health care provider), provider reports via a web-based data entry platform, faxes, and in rare instances, phone calls. An internal review (unpublished) of IDPH's I-NEDSS surveillance system revealed that in 2020, 90%-93% of the chlamydia and gonorrhea cases originated from ELRs.

Implementation of eCR Architecture and Method of Receiving Case Reports

Implementing eCR Architecture at Clinical Sites

The reference eCR implementations were defined in the previous study [12]. Some degree of customization was necessary to accommodate existing public health workflows, available code terminologies, and data parsing. The EHR leveraged 2 of the 3 case-detection scenarios [21]—scenario 1 (encounter diagnoses) and scenario 3 (name of the laboratory test performed and abnormal result flag), but not scenario 2 (name of the microorganism identified). While the case detection value sets include standard code concepts and terminologies, a given EHR instance may not have equivalent codes that reflect all 3 scenarios. This was the case for the Illinois implementation where the EHR in use contained codes aligned with Scenarios 1 and 3 only. The case detection logic is designed to work in the background of the EHR to detect chlamydia and gonorrhea cases using a predefined case detection algorithm consisted of value sets using industry standard terminologies (eg, International Classification of Diseases, Tenth Revision, Clinical Modifications; LOINC; and SNOMED-CT) [21], with no need for additional action on the part of the patient care team. Once the case detection logic was met, the EHR created XML documents in accordance with the HL7 eICR [28] and delivered them to IDPH via Secure File Transfer Protocol (Step 1, Figure 1).

Figure 1. Workflow depicting electronic case report (eCR) generation, transportation, and ingestion. EHR: electronic health record; eICR: electronic initial case report; HL7: Health Level 7; I-NEDSS: Illinois National Electronic Disease Surveillance System; IDPH: Illinois Department of Public Health; SFTB: secure file transfer protocol.



Receiving and Ingesting eCR at Public Health Agency

Directly ingesting eCRs into I-NEDSS was deemed out of scope due to project timeline and the extent of invasive intervention required to make modifications in the surveillance system. The raw eCRs were, instead, stored in a directory at IDPH (Step 2, Figure 1) and parsed by XML Harvester [29], an open-source XML parser application (Step 3, Figure 1). The parser application was housed in a relational database and passed the reformatted records to the staging database (Step 5, Figure 1). Relevant tables were identified by mapping data elements to the I-NEDSS and ingested into a Microsoft Access database for analysis (Step 6, Figure 1).

Data Collection and Analysis Methods

AllianceChicago tallied a count of eCRs generated daily throughout the pilot period and extracted data directly from the EHR. In some instances, manual inspection was performed on the captured data. Once the eCRs were received at IDPH, they were evaluated for the following: (1) compliance with the trigger logic for confirmed chlamydia and gonorrhea laboratory results or diagnosis, (2) completeness of key data elements, (3) timeliness of reporting, and (4) the degree to which eCR laboratory results matched ELR reporting during the study period. We compared the timeliness of eCR to ELR by calculating the median time from documented patient encounter to the time the public health agency received eCR and ELR for an individual.

Results

Case Report Throughput and Case Detection Metrics

eCR Throughput

During the study period, 11,192 encounters were logged in the EHR across all 8 clinical sites for various types of clinical care. During the study period, IDPH received 202 eCRs which were valid as per the HL7 requirements and stored in the staging database. Of these 202 eCRs, 19 (9.4%) did not meet the Council of State and Territorial Epidemiologists case definition and were not tallied. They were rejected because a manual review determined that they were not positive chlamydia or gonorrhea eCR tests or diagnosis. Therefore, a total of 183 eCRs representing 135 unique patients (some patients had both chlamydia and gonorrhea) were deemed valid and met the case definitions.

Case Detection Metrics

There were 113 instances where the case detection logic was satisfied for chlamydia; 80% (90/113) were based on scenario 3 (name of the laboratory test performed and abnormal result flag) and 20% (23/113) were based on both scenario 1 (encounter diagnoses) and scenario 3 (Table 1). Similarly, for the 70 instances where the case detection logic was satisfied for gonorrhea, 61% (43/70) were based on scenario 3, and 39% (27/70) were based on scenarios 1 and 3. Of note, there were no instances when scenario 1 (encounter diagnoses) alone was met, but instead, all of them had a corresponding confirmatory laboratory test (Table 1). eCR reported 95% (n=113 cases) of all the chlamydia cases that were reported from the participating clinical sites (Table 1). However, ELR reported 6 additional cases of chlamydia, which were not reported via eCR. eCR reported 97% (n=70 cases) of all the gonorrhea cases found by ELR, finding an additional 14 cases (19%) that were not reported via ELR (Table 1). ELR reported 2 additional cases of gonorrhea not detected by eCR. ELR reported 100% (119/119) of chlamydia cases but only 81% (58/72) of gonorrhea cases.

Table 1. Case detection metrics for electronic case reporting (eCR) and electronic laboratory reporting (ELR).

Variables	Chlamydia (n=119), n (%)		Gonorrhea (n=72), n (%)	
	Reported by ELR		Reported by ELR	
	Yes	No	Yes	No
Reported by eCR				
Yes				
Scenario 1 (encounter diagnoses) only	0	0	0	0
Scenario 3 (name of the laboratory test performed and abnormal result flag) only	90 (76)	0	33 (57)	10 (71)
Both scenarios 1 and 3	23 (19)	0	23 (40)	4 (29)
No				
No eCR was generated	6 (5)	0	2 (3)	0
Total, n	119	0	58	14

eCR Data Completeness and Timeliness

Key elements for public health reporting such as patient and provider names were complete in all eCRs and ELRs (Table 2). Race and ethnicity were slightly lower in eCR (n=113, 95%) than ELR (n=119, 100%; Table 2). However, eCR was found to report additional clinical data such as social history information, which is important for portraying a more complete epidemiologic picture but was beyond the scope of this evaluation. History of present illness and reason for visit were both reported on 82% (97/119) of eCR in this study; symptoms, diagnosis, and medications were reported on 33% (39/119), 23% (27/119), and 22% (26/119) eCRs, respectively (Table 2). While pregnancy is supported in the HL7 eICR standard [21], total absence of pregnancy information is due to the configuration of the EHR and is specific to this implementation.

The median time from the documented patient encounter to the time IDPH received ELR was 4 days. Similarly, the median time for scenario 1 (encounter diagnosis) was 11 days, and for scenario 3 (laboratory result) was 6 days.

Table 2. Comparison of the completeness in electronic case reporting (eCR) and electronic laboratory reporting (ELR).

eCR data element	Complete, n (%)	ELR data element	Complete, n (%)
Provider name	119 (100)	Ordering provider	119 (100)
Provider phone	119 (100)	Oder callback phone number	119 (100)
Provider fax	0 (0)	N/A[a]	N/A
Provider email	90 (76)	N/A	N/A
Patient name	119 (100)	Patient name	119 (100)
Patient phone	119 (100)	Phone number (home); phone number (business)	119 (100)
Patient email	29 (24)	N/A	N/A
Street address	119 (100)	Patient address	119 (100)
Birth date	119 (100)	Date or time of birth	119 (100)
Patient sex	119 (100)	Administrative sex	119 (100)
Race	113 (95)	Race	119 (100)
Ethnicity	113 (95)	Ethnic group	119 (100)
Preferred language	119 (100)	N/A	N/A
Pregnant	0 (0)	N/A	N/A
History of present illness	97 (82)	N/A	N/A
Reason for visit	97 (82)	Reason for study	0 (0)
Date of onset	39 (33)	N/A	N/A
Symptoms (list)	39 (33)	N/A	N/A
Diagnoses	27 (23)	N/A	N/A
Date of diagnosis	28 (24)	Date or time of the analysis	119 (100)
Medication administered (list)	26 (22)	N/A	N/A

[a] N/A: not applicable.

Discussion

Principal Findings

We implemented an approach to case reporting that automated and provided more complete, timely, and relevant information to public health authorities for chlamydia and gonorrhea in Illinois. This approach furthered data exchange between health care providers and public health by using health care data standards (SNOMED-CT, LOINC) to meet case detection logic in the EHR [12,21] and messaging standards (HL7 eICR) as the transport mechanism [20]. This method could facilitate better public health surveillance and inform public health practice. While ELRs comprised 90% of all cases of notifiable conditions reported to IDPH in 2020, eCR can provide better clinical information than the ELR is capable of reporting. Additionally, we demonstrated that eCR can be configured to retrigger at a later date to capture information not available at the time the initial case definition was met.

In this study, ELR detected 6 additional chlamydia cases and 2 additional gonorrhea cases, and eCR detected 14 additional cases of gonorrhea. These 14 cases were detected via scenarios where laboratory results were present in the EHR. We did not investigate why these were missed by ELRs, but some hypotheses include as follows: (1) specimen was collected toward the end of the study period (median time for receipt of ELR by IDPH was 4 days), (2) laboratory where the test was done could have had problems leading to delayed reporting, (3) specimens could have been tested by out-of-jurisdiction labs, and (4) other unknown reasons. Similarly, eCR detection logic missed 6 chlamydia and 2 gonorrhea cases. Potential reasons include the following: (1) faulty local code mapping; (2) if patient was seen toward the end of the study period, eCR may not have triggered yet (could take up to 11 days for IDPH to receive eCR, compared to ELR, which took 4 days); and (3) other unknown reasons. Of note, among all the cases reported by the clinical sites, eCR reported 95% and 97% of all chlamydia and gonorrhea cases, respectively. We also found that eCR contains more complete information compared to ELR and allowed for the collection of additional data such as diagnoses, treatment, and other clinical information that are typically not available in ELR. However, these additional data elements were not available on all eCRs. Some records reflect clients who were tested or diagnosed at external locations whose records were subsequently scanned into the EHR system as PDF files with no accompanying lab records. More work needs to be done to explore whether these were not available at the time of the case report generation or if this information was present in some other unstructured format not represented as a standard vocabulary such as SNOMED-CT.

An earlier study by Mishra et al [12] had developed the case detection logic, curated the value sets needed for this eCR architecture, and collaborated with a clinical partner to demonstrate the case detection and eCR efficacy. Another study in Oregon, simultaneously conducted with this evaluation, also implemented this approach with clinical sites reporting to the Oregon Health Authority. They reported on the local customization required for that implementation and differed from this study in that the Oregon Health Authority had existing infrastructure to consume eCR via its interface engine into its proprietary surveillance system [21]. A broader adoption of eCR will be a step toward bridging the information gap between health systems and public health. Such steps toward improved semantic interoperability will allow a timely and better understanding of epidemiology, will inform the development of effective policy and most relevant updates to treatment guidelines, and will reduce burdens for both health systems and public health agencies by automating case reporting and investigating cases for more information.

Limitations

The EHR had limited Fast Healthcare Interoperability Resources capability. The health centers had to collaborate with the EHR vendor to develop, test, and release a beta version of eCR that could be specifically used for this project. eCR standards allow for data elements that are above and beyond those that have been received through traditional reporting. This information may be present in an unstructured format without any semantic representation. An eCR triggered by laboratory results (scenario 3) would contain information in a structured format with semantic representation, while for scenario 1, diagnoses may not be available in a structured format but may be represented in an unstructured, narrative format, resulting in a lower number of triggers. Some additional examples of unstructured data elements include those found in the “Problems,” “History of present illness,” “Reason for visit,” “Social history,” and so on, and could explain why symptoms, diagnosis, and medications were reported only on 33%, 23%, and 22% of the eCRs, respectively.

I-NEDSS, the surveillance system at IDPH, required additional resources to consume eCR. Any addition to the surveillance system’s capability would require extensive modification. Due to the constraint of project timeline and lack of dedicated resources for this activity, we elected not to reconfigure the surveillance information system to receive eCR. Instead, we used a parser application and created a staging database that could be mapped to the data elements of the surveillance information system. This challenge of ingesting eCRs into surveillance information systems is likely to be far more common across health departments in the United States and needs additional work to create a standardized and scalable approach. We did not quantify the burden of manual reporting, and how much of that burden was potentially alleviated via automated case reporting via eCR. Future studies could provide insight into the resources saved via such automated reporting methods.

Conclusions

The eCR approach to public health surveillance successfully identified and created automated case reports for chlamydia and gonorrhea cases in the selected Illinois clinics. This approach also demonstrated a more complete case report with additional demographic, clinical, and treatment information. eCR reduced the burden of reporting cases on clinical providers and represents a promising future of greater semantic interoperability between health care systems and public health by automating the case report using health care data standards in a scalable manner.

Acknowledgments

This project was supported by the Centers for Disease Control and Prevention funding to the Task Force for Global Health under cooperative agreement number U38OT000216-3.

Authors' Contributions

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention (CDC). Mention of company names or products does not imply endorsement by the CDC.

Conflicts of Interest

None declared.

References

1. Sexually Transmitted Disease Surveillance 2019. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/std/statistics/2019/default.htm [accessed 2022-04-11]
2. Control of Sexually Transmissible Infections. State of Illinois Joint Committee on Administrative Rules. URL: https://www.ilga.gov/commission/jcar/admincode/077/077006930000300R.html [accessed 2021-08-06]
3. Challenges to Public Health Reporting Experienced by Non-Federal Acute Care Hospitals. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/data/data-briefs/challenges-public-health-reporting-experienced-non-federal-acute-care-hospitals [accessed 2022-02-15]
4. Langmuir AD. The surveillance of communicable diseases of national importance. N Engl J Med 1963 Jan 24;268(4):182-192. [doi: 10.1056/nejm196301242680405]

5. Thacker S, Berkelman R. Public health surveillance in the United States. Epidemiol Rev 1988;10:164-190. [doi: 10.1093/oxfordjournals.epirev.a036021] [Medline: 3066626]
6. Lee LM, Thacker SB. The cornerstone of public health practice: public health surveillance--2011. MMWR Suppl 2011;60(4):15-21. [doi: 10.1007/978-1-4020-5614-7_343]
7. Buehler JW, Centers for Disease Control Prevention. CDC's vision for public health surveillance in the 21st century. Introduction. MMWR Suppl 2012 Jul 27;61(3):1-2. [Medline: 22832989]
8. Thacker S, Qualters J, Lee LM, Centers for Disease ControlPrevention. Public health surveillance in the United States: evolution and challenges. MMWR Suppl 2012 Jul 27;61(3):3-9. [Medline: 22832990]
9. Lamb E, Satre J, Hurd-Kundeti G, Liscek B, Hall C, Pinner R, Centers for Disease ControlPrevention (CDC). Update on progress in electronic reporting of laboratory results to public health agencies - United States, 2014. MMWR Morb Mortal Wkly Rep 2015 Apr 03;64(12):328-330 [FREE Full text] [Medline: 25837244]
10. Mac Kenzie WR, Davidson AJ, Wiesenthal A, Engel JP, Turner K, Conn L, et al. The Promise of Electronic Case Reporting. Public Health Rep 2016 Nov 14;131(6):742-746 [FREE Full text] [doi: 10.1177/0033354916670871] [Medline: 28123218]
11. Birkhead GS, Klompas M, Shah NR. Uses of electronic health records for public health surveillance to advance public health. Annu Rev Public Health 2015 Mar 18;36(1):345-359. [doi: 10.1146/annurev-publhealth-031914-122747] [Medline: 25581157]
12. Mishra N, Jellison J, Hamilton A, Carr J, Padilla R, Viator N. Leveraging Informatics to Identify Reportable Cases: Pilot Findings on Electronic Case Reporting of Chlamydia and Gonorrhea. J Public Health Manag Pract 2019;25(6):595-597. [doi: 10.1097/PHH.0000000000000954] [Medline: 30789599]
13. Overhage JM, Grannis S, McDonald CJ. A Comparison of the Completeness and Timeliness of Automated Electronic Laboratory Reporting and Spontaneous Reporting of Notifiable Conditions. Am J Public Health 2008 Feb;98(2):344-350. [doi: 10.2105/ajph.2006.092700]
14. Dixon BE, Vreeman DJ, Grannis SJ. The long road to semantic interoperability in support of public health: experiences from two states. J Biomed Inform 2014 Jun;49:3-8 [FREE Full text] [doi: 10.1016/j.jbi.2014.03.011] [Medline: 24680985]
15. Revere D, Hills RH, Dixon BE, Gibson PJ, Grannis SJ. Notifiable condition reporting practices: implications for public health agency participation in a health information exchange. BMC Public Health 2017 Mar 11;17(1):247 [FREE Full text] [doi: 10.1186/s12889-017-4156-4] [Medline: 28284190]
16. Position Statement Archive. Council of State and Territorial Epidemiologists. URL: https://www.cste.org/page/PositionStatements [accessed 2019-08-20]
17. Public Health Reporting and National Notification for Infection Caused by Chlamydia trachomatis. Council of State and Territorial Epidemiologists. URL: https://cdn.ymaws.com/www.cste.org/resource/resmgr/PS/09-ID-08.pdf [accessed 2020-11-13]
18. Update to Public Health Reporting and National Notification for Gonorrhea. Council of State and Territorial Epidemiologists. URL: https://cdn.ymaws.com/www.cste.org/resource/resmgr/PS/13-ID-03.pdf [accessed 2020-11-13]
19. eCR for STIs Toolkit. Public Health Informatics Institute. URL: https://phii.org/module-6/electronic-case-reporting-for-sexually-transmitted-infections-downloads/ [accessed 2021-01-06]
20. HL7 CDA® R2 Implementation Guide: Public Health Case Report, Release 2: the Electronic Initial Case Report (eICR), Release 1, STU Release 1 - US Realm. Health Level 7. URL: https://www.hl7.org/dstucomments/showdetail.cfm?dstuid=199 [accessed 2020-11-02]
21. Case Detection Logic and Trigger Code Value Sets for eCR. Public Health Informatics Institute. URL: https://phii.org/download/ecr-sti-case-detection/ [accessed 2020-11-02]
22. Todd JV, Collins NV, Oakley J, Menza T, Barber M, Kasarskis I, et al. Automating Case Reporting of Chlamydia and Gonorrhea to Public Health Authorities in Oregon Clinics. Sexual Trans Dis 2021 Jul 14;49(1):38-42. [doi: 10.1097/olq.0000000000001507]
23. About Near North Health. Near North Health. URL: https://www.nearnorthhealth.org/about [accessed 2022-01-06]
24. Electronic Health Records. Athenahealth. URL: https://www.athenahealth.com/solutions/electronic-health-records [accessed 2021-03-19]
25. HL7 Interface Engine. QVERA. URL: https://www.qvera.com/hl7-interface-engine/ [accessed 2021-03-19]
26. Infectious Disease Reporting. Illinois Department of Public Health. URL: https://dph.illinois.gov/topics-services/diseases-and-conditions/infectious-diseases/infectious-disease-reporting [accessed 2021-03-09]
27. Illinois Sexually Transmissible Disease Control Act. State of Illinois General Assembly. URL: https://tinyurl.com/bdf5k5rb [accessed 2021-03-09]
28. HL7 CDA® R2 Implementation Guide: Public Health Case Report, Release 2 - US Realm - the Electronic Initial Case Report (eICR). Health Level 7 International. URL: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=436 [accessed 2021-03-09]
29. XmlHarvester (Lantana Group). GitHub. URL: https://github.com/lantanagroup/XmlHarvester [accessed 2022-01-06]

Abbreviations

CDC: Centers for Disease Control and Prevention
eCR: electronic case reporting
EHR: electronic health record
eICR: electronic initial case report
ELR: electronic laboratory reporting
HL7: Health Level 7
I-NEDSS: Illinois National Electronic Disease Surveillance System
IDPH: Illinois Department of Public Health
LOINC: Logical Observation Identifiers Names and Codes
SNOMED-CT: Systematized Nomenclature of Medicine – Clinical Terms

Edited by G Eysenbach; submitted 19.04.22; peer-reviewed by M Barber, H Sun; comments to author 16.07.22; revised version received 16.08.22; accepted 31.10.22; published 14.03.23.

Please cite as:
Mishra N, Grant R, Patel MT, Guntupalli S, Hamilton A, Carr J, McKnight E, Wise W, deRoode D, Jellison J, Collins NV, Pérez A, Karki S
Automating Case Reporting of Chlamydia and Gonorrhea to Public Health Authorities in Illinois Clinics: Implementation and Evaluation of Findings
JMIR Public Health Surveill 2023;9:e38868
URL: https://publichealth.jmir.org/2023/1/e38868
doi:10.2196/38868
PMID:36917153

Original Paper

The Association Between Social Determinants of Health and Population Health Outcomes: Ecological Analysis

Ace Vo[1], PhD; Youyou Tao[1], PhD; Yan Li[2], PhD; Abdulaziz Albarrak[3], PhD

[1]Information Systems and Business Analytics Department, Loyola Marymount University, Los Angeles, CA, United States

[2]Center for Information Systems and Technology, Claremont Graduate University, Claremont, CA, United States

[3]Information Systems Department, King Faisal University, Al-Ahsa, Saudi Arabia

Corresponding Author:
Ace Vo, PhD
Information Systems and Business Analytics Department
Loyola Marymount University
1 Loyola Marymount University Dr
Los Angeles, CA, 90045
United States
Phone: 1 3103384522
Email: ace.vo@lmu.edu

Abstract

Background: With the increased availability of data, a growing number of studies have been conducted to address the impact of social determinants of health (SDOH) factors on population health outcomes. However, such an impact is either examined at the county level or the state level in the United States. The results of analysis at lower administrative levels would be useful for local policy makers to make informed health policy decisions.

Objective: This study aimed to investigate the ecological association between SDOH factors and population health outcomes at the census tract level and the city level. The findings of this study can be applied to support local policy makers in efforts to improve population health, enhance the quality of care, and reduce health inequity.

Methods: This ecological analysis was conducted based on 29,126 census tracts in 499 cities across all 50 states in the United States. These cities were grouped into 5 categories based on their population density and political affiliation. Feature selection was applied to reduce the number of SDOH variables from 148 to 9. A linear mixed-effects model was then applied to account for the fixed effect and random effects of SDOH variables at both the census tract level and the city level.

Results: The finding reveals that all 9 selected SDOH variables had a statistically significant impact on population health outcomes for ≥2 city groups classified by population density and political affiliation; however, the magnitude of the impact varied among the different groups. The results also show that 4 SDOH risk factors, namely, asthma, kidney disease, smoking, and food stamps, significantly affect population health outcomes in all groups ($P<.01$ or $P<.001$). The group differences in health outcomes for the 4 factors were further assessed using a predictive margin analysis.

Conclusions: The analysis reveals that population density and political affiliation are effective delineations for separating how the SDOH affects health outcomes. In addition, different SDOH risk factors have varied effects on health outcomes among different city groups but similar effects within city groups. Our study has 2 policy implications. First, cities in different groups should prioritize different resources for SDOH risk mitigation to maximize health outcomes. Second, cities in the same group can share knowledge and enable more effective SDOH-enabled policy transfers for population health.

(JMIR Public Health Surveill 2023;9:e44070) doi:10.2196/44070

KEYWORDS

social determinants of health; public policy; health outcomes; policy recommendation; cities

Introduction

Overview

Social determinants of health (SDOH), defined by the World Health Organization, encompass economic policies, social and physical environments, and access to health services and shapes the conditions "in which people are born, grow, work, live, and age" [1].

It is well established that SDOH factors, such as health behaviors, clinical care, social and economic status, and physical environment, account for 30% to 55% of health outcomes [2]. There is considerable literature examining the association between SDOH factors and various health outcomes, among which most studies used patient-level and hospital-level data. For example, prior studies have found that SDOH factors are associated with medication adherence [3], care use [4], readmission risk [5,6], length of stay [7,8], postoperative surgical outcomes [9], mortality risk [10,11], and risk of exposure to and subsequent health outcomes after contracting the SARS-CoV-2 virus [12,13].

Understanding the ecological association between SDOH factors and population-level health outcomes is vital. Such an understanding is particularly relevant for researchers, clinicians, and policy makers in assessing new SDOH-enabled programs or policies to improve population health, enhance quality of care, and reduce health inequities [14,15]. Several frameworks have been developed to understand the impact of SDOH risk factors on population health outcomes [16-19]. Among these frameworks, the County Health Ranking (CHR) model has been widely applied to explicate the relationship between SDOH factors and population health outcomes in the United States [17,20].

The CHR model uses >30 ranking criteria to measure the impact of SDOH factors on the current and future population health outcomes at the county level [17]. However, this model is not always accurate at the state level or lower-than-state level (eg, county, city, and census tract level) because it uses a predetermined set of weights for the SDOH factors. Two recent empirical studies of the CHR found that the influence of SDOH factors on health outcomes varies among different states in the United States [21,22]. In addition, existing CHR studies are limited to either the county level or the state level. Analysis at a lower level is desired by local policy makers because local governments need to derive insights into the SDOH and health outcomes within their administrative delineation [23]. For instance, Corburn et al [23] showed substantial results in reducing health inequities by examining all policies adopted by the city of Richmond, California, at the zip code level.

This study investigates the ecological associations between the SDOH and population health outcomes at the city level to support health policy decisions using the CHR as the foundation. For this purpose, we curated data from 5 different sources and integrated them at the census tract level from 29,126 census tracts within 499 cities across all 50 states in the United States. With such smaller geographical delineations, researchers, policy makers, and other relevant parties can aggregate data into larger administrative divisions to make more impactful and effective decisions.

After grouping cities through 2 factors, namely, population density and political affiliation, we formalized measures for SDOH factors and population health outcomes and used a voting-based feature selection approach to reduce the original 148 sociodemographic and SDOH-related variables to 9. We then used a linear mixed-effects model to examine the ecological associations between the SDOH factors and population health outcomes. At the census tract level, the goodness of fit ranges from 0.65 to 0.75. At the city level, the total variance explained by the model was high, ranging from 0.86 to 0.90. The effect size of variables is different across groups. Noticeably, asthma, kidney disease, smoking, and food stamp variables were significant in all groups. Post hoc analysis was later conducted using predictive margin to assess group differences in health outcomes for the 4 behavioral health indicators that majorly affect health outcomes across all groups.

Background

Social Determinants of Health

SDOH encompass a wide set of dimensions such as socioeconomics, education, physical environment, food access, health care system condition, health behaviors, community status, and politics [1,24,25]. The importance of the SDOH is apparent to both academia and policy makers. Research on SDOH is impactful and prominent, especially in the context of this research, that is, linking health outcomes and disparities [2,3,12,26,27].

In the United States, the CHR model is a widely applied model that describes how SDOH factors contribute to population health [17,20]. The CHR model uses >30 ranking measures to understand the current population health outcomes and health factors (ie, health behaviors, clinical care, social and economic status, and physical environment) that would affect future health in the United States [17]. Both the health outcomes and health factors are weighted by a panel of national experts on population health in the CHR model [17].

Although the CHR model makes health outcomes and factors easy to calculate and be understood by the general public and policy makers, it has 2 major drawbacks. First, the weights proposed by experts in the CHR may not be applicable to the entire population because different locations may have various population characteristics and dissimilar social and policy environments. Two recent studies that empirically tested the CHR [21,22] revealed that the influence of health factors on health outcomes varies among the US states. In other words, the CHR model performed better in some states than in others. Second, previous studies used SDOH through CHR at either the county level or the state level. A more granular level of data is needed to comprehensively understand the SDOH factors that affect population health and to provide insights for local policy makers, such as governments at the city level. In this study, we use census tract data, a geographical delineation of a county that encompasses approximately 4000 population, with relatively well-maintained and updated demographic information [28]. Using census tracts as the unit of analysis, researchers and policy makers can aggregate the data into larger administrative divisions.

Generally, SDOH research has both breadth and depth. Research can be found using study sites worldwide, with different takes on what SDOH are, and using a diverse range of research methods. The same perspective can be gleamed when focusing on the relationship between health outcomes and the SDOH. Owing to the divergent and seemingly agglomeration research that explores the relationship between health outcomes and SDOH factors, researchers need to be careful when defining (1) the health outcomes measurement, (2) the SDOH factors, (3) the population and study sites, and (4) the research methods. These definitions set up the phenomenon for research while providing concrete findings to help promote population health, which in turn affects policy changes.

SDOH Factors and Health Outcomes

Health outcomes and SDOH factors are interlinked. For instance, poverty, education, and income are known to be closely related to health outcomes [12,26,29-32]. Specifically, Washington et al [30] identified low income to be related to the higher unmet needs for health care, whereas Bauer et al [33] found a major association between income and health literacy. Several studies have associated income and insurance status with risk of readmission [6,34]. Similarly, studies have found that income, age, and vulnerable populations have worse postoperative surgical outcomes or higher mortality risk [9-11]. Furthermore, when present alongside SDOH, disadvantaged racial and ethnic groups are negatively affected in several health outcomes areas, including COVID-19 [35], stroke [36], and kidney transplant [37].

Health behavior is another essential component of SDOH [38,39]. It is related to people's health practices such as smoking, diet, exercise, and alcohol habits. According to Healthy People 2020, individuals' healthy behaviors can substantially influence their health needs and outcomes [40]. Likewise, it was estimated that approximately 443,000 Americans die annually of smoking-related diseases, such as cancer, stroke, lung disease, and heart disease [41]. Lack of physical activities and high-calorie food intake can lead to obesity, which majorly affects people's health conditions.

Population Density

At any location, the population can be delineated into living in 3 areas: urban, suburban, and rural. Although the general public could easily discern the differences between the 3, in research, this has been proven as not as clear-cut as it may seem [42]. Generally, there is a decision framework, along with several decision criteria, to designate a location to be either urban, suburban, or rural. This designation separates the way in which people live. Unfortunately, owing to the nature of our research inquiry that examines census tracts at the city level, the abovementioned delineation cannot be definitively justified. However, based on the simple criteria of population density, we were able to dissect cities into different segments, each of which had different health behaviors and outcomes.

Indeed, population density is closely related to health service delivery and further influences health outcomes. On the one hand, health care institutions located in low population–density places may need to manage small-scale operations and handle financial losses owing to low volume, whereas people who live in low population–density locations may have difficulties accessing health care facilities and services because of human service and resource deficiency. For example, in less populated areas, ambulance response time is likely to be longer [43].

Furthermore, these low population–density locations usually have an increasingly aging population, demanding more health services and resources [44,45]. Government-supported health programs are often limited in low population–density areas because of the high rate of poverty and limited tax in such areas [45], and high population–density areas may cause congestion in major hospitals [46], suggesting a complex relationship between population density and health outcomes. The interplay between the supply and demand of health care is greatly influenced by the availability of services and the needs of the population. Naturally, the denser the population, the more health care is needed. In contrast, low population–density areas require fewer health care services but suffer from scarcity.

Previous studies have found that population density is associated with various health outcomes, such as mortality rate [47,48], survival outcome [49], and morbidity in certain diseases [48]. To account for the complex relationship between population density and health outcomes, we examined the relationship between SDOH and health outcomes by grouping cities based on their population density.

Political Affiliation

Local governments and their political leaning greatly affect health policy. The US political system is a constant wrestle between 2 major political parties: the Democratic Party (often colorized as blue) and the Republican Party (often colorized as red). Contention exists throughout the United States and at various levels of the government. At the state level, the differences between the parties encompass both health policy and social issues (eg, attitude and policy regarding abortion and substance use) and the preferred role of the government (eg, big vs small government) in addressing health-related problems [50]. For example, Pagel et al [51] painted a stark contrast in the priorities of health care policy between Democratic and Republican state legislators. Republicans prioritize reducing health costs and smaller government, whereas Democrats prioritize improving health and equity and reducing disparities over other goals [51].

At the individual level, partisan polarization in public attitudes shapes individuals' health behaviors. For example, studies found that there are diverging attitudes between Republicans and Democrats toward influenza and the COVID-19 vaccine, in which Republicans displayed a negative attitude and intention toward vaccine, whereas attitudes and intentions of Democrats remained largely stable [52,53]. Given the importance of how political affiliation can affect both policies and attitudes toward health outcomes, we surmise that the identification of political parties would also influence the relationship between health outcomes and SDOH.

Despite the research outcomes suggesting linkages between health outcomes and SDOH, policy makers do not view SDOH as a priority, as is evident in the absence of SDOH in the general government policy agenda, despite earlier emphasis [54,55]. Instead, current health policy in the United States still focuses on "medicalizing" health problems, assuming that the solution to health is medical care [55,56]. For example, policy makers often emphasized health access policies to increase geographic and financial access to health services for vulnerable populations [56] but neglected other important social and economic factors that are associated with health disparities [56]. Prior studies also pointed out that even with universal health care access, rich populations are more likely to have healthier lives because they can afford advanced health care; thus, the most important policy issues in health care are largely dependent on the overall allocation of resources to health care rather than merely on the distributive justice within health care [57,58].

Furthermore, Embrett and Randall [54] suggested that extant SDOH policy studies appear to be focused on advocacy rather than analysis. To promote "healthy public policies" based on SDOH factors, health disparities need to be examined so that the government can develop specific policy responses with their tools at hand such as regulation, legislation, taxation, and financing [59]. Health care professionals, researchers, and governments can work together to empirically examine the extent to which SDOH factors contribute to disparate health outcomes so that policies can be developed to solve health disparities in a more effective and efficient manner [13].

This study makes the following contributions to research and practice. First, the study divulges additional insights into how different SDOH factors affect health outcomes by curating related data in census tracts. In contrast to the county level, targeted but limited SDOH data are available at the census tract level. Our study presents a novel data curation process that creates additional SDOH variables that are otherwise not readily available. At the census tract level, data could be aggregated to the city level, allowing policy makers to devise policy more effectively. Second, city is a living and emergent ecosystem, and each city presents its own opportunities and challenges in terms of population health [60]. Despite these differences, our study sheds unique insights through 5 distinct groups with 2 prevailing properties: the city's population density and its political affiliation. The analysis results showed that the health effectiveness intragroup was similar, whereas the effects diverged intergroup. Therefore, it is prudent for cities from different groups to prioritize resource allocations to address SDOH factors based on their group properties so that population health outcomes can be maximized.

Methods

Study Design

The study design includes 4 considerations. First, to perform the ecological analysis at the city level, we need to identify data sources for SDOH factors and health outcomes at the group level. Second, we need to determine the unit of analysis, that is, how to define the study population and the method of grouping. Third, we need to formalize the measures for SDOH factors and population health outcomes. Finally, we need to determine the appropriate data analysis method. In the subsequent sections, we describe each consideration in detail.

Data Collection

We curated and integrated a data set from 5 different sources: the PLACES program [61], the National Center for Health Statistics, Census Data portal, Simply Analytics, and the Massachusetts Institute of Technology Election Laboratory [62]. The data obtained from the PLACES program included population health outcomes, such as mental and physical health days, as well as population health behaviors, such as the percentage of the population with asthma or kidney disease. Additional data were collected at the census tract level for a selected 500 cities in the United States, which serves as the basis of the data for analysis. Life expectancy data for each census tract were obtained from the National Center for Health Statistics [63]. These data were later used to measure the health outcomes. Population-related variables such as age, income, ethnicities, and education levels were obtained through the Census Data portal, supplemented by data from the Simply Analytics platform. We used data from the Massachusetts Institute of Technology Election Laboratory to determine the political affiliation to which each census tract belongs. Political affiliation is proxied by the 2020 presidential election. In the United States, voting for a presidential candidate happens in the general election every 4 years, in which most of the population votes, and it solidifies the general political direction of the country, whether it will lean toward Democratic or Republican policies. Election data are reported at the precinct level, which is a larger spatial delineation than census tracts, is smaller than US counties, and overlaps with US cities. When precincts have more votes for the Democrats presidential candidate, we code the census tracts to reside within it as blue. Similarly, we coded red for census tracts that reside in precincts that have more votes for the Republican presidential candidate. census tract data from Washington, District of Columbia, were dropped because they did not have a life expectancy measure. The final data set included 29,126 census tracts within 499 cities across all 50 states in the United States, all obtained in 2021.

Determining Unit of Analysis

The results of our data collection process yielded the census tract as the default unit of analysis. We further examined whether this unit of analysis was sufficient or whether an additional grouping mechanism was warranted. As the census tracts resided within the city boundary, a natural grouping was to coalesce tracts based on the city itself. However, this type of grouping did not help explain the relationship between cities. Rather, additional grouping of cities was required. Therefore, we used political affiliation and population density, as discussed in the previous sections with the same name.

To determine the political affiliation of a city, we revisited the political affiliation of each census tract, which was determined using the abovementioned description. With a simple majority rule, if a city has more census tracts that are red, we assign red as the city's political affiliation. Similarly, a city will be coded blue as its political affiliation if a majority of the census tracts residing within it are blue. Correspondingly, the population density of a tract was calculated by dividing the total population by its area in square miles [44,49]. Each city was then sorted and classified into 3 quartiles: sparsely populated cities were below the 25th percentile (denoted low), medium-populated cities were between the 25th and 75th percentiles (denoted mid), and highly populated cities were above the 75th percentile (denoted high). The final grouping included 73 cities in blue-low, 198 cities in blue-mid, 120 cities in blue-high, 53 cities in red-low, 51 cities in red-mid, and 4 cities in red-high. As red-high had a small number of cities, it was combined with the red-mid category, making the final number of cities in the combined category, red-high, 55. Multimedia Appendix 1 includes the list of cities in each classification.

Formalizing Measurements

Formalizing Health Outcome Measurement

After integrating and cleaning the data, we then formulated the health outcomes measurements for our research based on the CHR model [17]. In the CHR model, health outcome measures the current state of population health and can be further divided into 2 categories: length of life and quality of life. Length of life can be measured by life expectancy and quality of life by poor physical health days and poor mental health days. Specifically, the health outcome measure was based on the CHR model, as shown in the following equation:

$$\text{Health outcome} = 50\% \times (\text{life expectancy}) - 25\% \times (\text{poor physical health days}) - 25\% \times (\text{poor mental health days}) \quad (1)$$

Similar to the CHR model, our proposed health outcome measure assigns equal weightage to the length of life (through life expectancy) and quality of life (through poor physical health days and poor mental health days). The subtraction signs signify the negative effects of having poor physical health days and poor mental health days on health outcomes.

Formalizing SDOH Measurements

To formalize the measurements for SDOH factors, we first standardized the different scales such as rates, percentages, and averages of the survey responses. We also followed the CHR model by standardizing all measures with *z* scores, in which the standardized variables had a mean of 0 and a SD of 1 [17,22]. The initial SDOH data comprised 148 variables (Multimedia Appendix 2), including some highly correlated features. Thus, a feature reduction is required. As different feature selection techniques may result in a different set of features and no standardized rules on which technique might be better than the others, we experimented with the following 5 feature selection methods: Pearson correlation, recursive feature elimination, backward stepwise regression, XGBoost, and random forest. Independent variables were then ranked into 4 quartiles from highest to lowest, based on their feature importance or effect size against the dependent variable health outcome. We then applied a majority voting method, retaining variables that were placed in the top quartile based on at least 3 methods. The final data set included 9 variables for SDOH risk factors (Table 1).

Table 1. Metadata of the final data set[a].

Variable name	Type	Description
Health outcome	DV[b]	Health outcome of the population (refer to equation 1)
Asthma	IV[c]	Percentage of the population aged >18 years and with asthma
Kidney disease	IV	Percentage of the population aged >18 years and with chronic kidney disease
Smoking	IV	Percentage of the population aged >18 years who has smoked >100 cigarettes in their lifetime and currently smoke
Teeth lost	IV	Percentage of the population aged ≥65 years who has lost all their natural teeth
Annual checkup	IV	Percentage of the population aged >18 years who has visited a physician for a routine checkup
Lack of sleep	IV	Percentage of the population aged >18 years who sleeps <7 hours over a 24-hour period
Lack of health insurance	IV	Percentage of the population aged between 18 and 64 years who does not have health insurance
Below poverty	IV	Percentage of the population below the federal poverty level
Food stamps	IV	Percentage of households using food stamps or other cash public assistance programs
Group	Grouping	Grouping of tracts based on the city's population density and political affiliation (5 groups in total)

[a] The final data set included 1 dependent variable, 9 independent variables, and 1 grouping variable. Metadata applies for all 29,126 census tracts within 499 cities in the United States in 2021.

[b] DV: dependent variable.

[c] IV: independent variable.

Data Analysis

For data analysis, a linear mixed-effects model [64] was used to analyze the statistical parameters that varied at the census tract level and the city level. Observations in census tracts were hypothesized to have a systematic and predictable influence on health outcomes. Thus, the census tract data were modeled as a fixed effect, nonrandom, or nonindependent. As different cities are expected to have unpredictable and nonsystematic effects, they were modeled as random effects.

Ethical Considerations

The data used in this study are secondary data, which do not involve human participants, and are collected and aggregated by the respective organizations mentioned in the *Data Collection* section, most of which are curated from the US Census Bureau as the primary source. No additional data were collected specifically for this study.

Results

All the linear mixed-effects models converged. We present the results of the linear mixed-effect models for 5 groups: blue-low (model 1), blue-mid (model 2), blue-high (model 3), red-low (model 4), and red-high (model 5) in Tables 2-4. The R^2 results from the 5 models indicated that the overall fit of each model was adequate. Specifically, the conditional R^2 represents the variance of the variables explained by both fixed and random effects within the model. The conditional R^2 value ranging from 0.86 to 0.90 in models 1 to 5 indicated that the grouping at the city level exemplifies the intertwined relationship between population density and political affiliation grouping against health outcomes. Furthermore, the marginal R^2 represents the variance affected by the fixed effects. The marginal R^2 value ranging from 0.60 to 0.75 in models 1 to 5 indicates that the set of variables at the census tract level also has high goodness of fit to health outcomes.

Our results suggest that all the groups examined in the study have an equipotential baseline, as evidenced by the comparable intercepts across all 5 city groups classified by population density and political affiliation. Furthermore, the results showed that all SDOH variables had a statistically significant impact on population health outcomes for ≥2 or city groups, but the magnitude of the impact varied among the different groups. For example, an increase in the proportion of the population with asthma, kidney disease, smoking, or using cash assistance programs (such as food stamps) is associated with a decline in population health outcomes across all groups, consistent with the SDOH literature [65-68]. However, the magnitude of the impact of these variables on population health outcomes varied among different groups. The impact of asthma on population health outcomes is highest for cities classified as red-high (model 5: red-high) and lowest for cities classified as blue-low (model 1: blue-low), whereas the impact of kidney disease or smoking is higher for blue cities (as observed in models 1-3). The impact of food stamps on population health outcomes is higher for red cities with high population density (model 5: red-high) and blue cities with medium population density (model 2: blue-mid).

Some SDOH variables were statistically significant for a subset of the city groups. For instance, the percentage of the population having an annual checkup does not have a statistically significant impact on population health outcomes for red cities with low population density (model 4: red-low), whereas the impact of the percentage of the population below the federal poverty level is not statistically significant only for red cities with high population density (model 5: red-high). Similarly, the impact of the percentage of the population without health insurance is not statistically significant for blue cities with low density (model 1: blue-low). Finally, the impact of lack of sleep is only statistically significant for rural cities, and the impact of teeth loss is statistically significant for cities with a high population density (model 3: blue-high and model 5: red-high), regardless of their political affiliation.

To further investigate these findings, a post hoc analysis was conducted using predictive margins to assess group differences in health outcomes for the 4 SDOH variables, namely, asthma, kidney disease, smoking, and food stamps, that substantially affected health outcomes across all groups (as shown in Figure 1). Figure 1 illustrates that each group had a better health outcome with low asthma, low smoking, low kidney disease, or low food stamps and a worse health outcome with high asthma, high smoking, high kidney disease, or high food stamps. One exception was the red-low group, which had a slightly better health outcome with a high food stamp. Overall, the blue-low and red-low groups had better health outcomes than the other 3 groups, whereas blue-high had the lowest health outcome compared with the other 4 groups. Furthermore, asthma, kidney disease, smoking, and food stamps had larger impacts on the red-high and blue-high groups than the other 3 groups.

Table 2. Census tract level–fixed effect results of the linear mixed-effect model.

	Model 1 (blue-low)		Model 2 (blue-mid)		Model 3 (blue-high)		Model 4 (red-low)		Model 5 (red-high)	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
Lack of health insurance	−0.124 (0.095)	.19	−0.284 (0.042)	<.001	−0.697 (0.038)	<.001	−0.792 (0.093)	<.001	−0.808 (0.138)	<.001
Asthma[a]	−0.848 (0.158)	<.001	−1.006 (0.109)	<.001	−1.248 (0.075)	<.001	−1.266 (0.228)	<.001	−1.731 (0.256)	<.001
Kidney disease[a]	−1.247 (0.130)	<.001	−1.491 (0.096)	<.001	−1.973 (0.092)	<.001	−0.738 (0.193)	<.001	−1.348 (0.284)	<.001
Smoking[a]	−3.203 (0.123)	<.001	−1.925 (0.094)	<.001	−2.259 (0.098)	<.001	−1.955 (0.167)	<.001	−1.136 (0.187)	<.001
Teeth lost	0.011 (0.112)	.92	−0.049 (0.081)	.54	0.342 (0.081)	<.001	−0.301 (0.158)	.06	−0.656 (0.228)	.004
Below poverty	−0.158 (0.064)	.01	−0.256 (0.044)	<.001	−0.328 (0.042)	<.001	−0.184 (0.088)	.04	−0.079 (0.113)	.48
Food stamps[a]	−0.281 (0.072)	<.001	−0.346 (0.047)	<.001	−0.156 (0.039)	<.001	−0.279 (0.103)	.007	−0.377 (0.121)	.002
Annual checkup	0.292 (0.110)	.008	0.520 (0.079)	<.001	1.022 (0.065)	<.001	−0.056 (0.162)	.73	0.563 (0.222)	.01
Lack of sleep	0.955 (0.158)	<.001	0.193 (0.108)	.07	0.048 (0.078)	.07	1.346 (0.232)	<.001	0.565 (0.305)	.01
Intercept	1.344 (0.044)	<.001	1.261 (0.025)	<.001	1.248 (0.037)	<.001	1.314 (0.034)	<.001	1.383 (0.042)	<.001

[a] Variables are statistically significant in all groups.

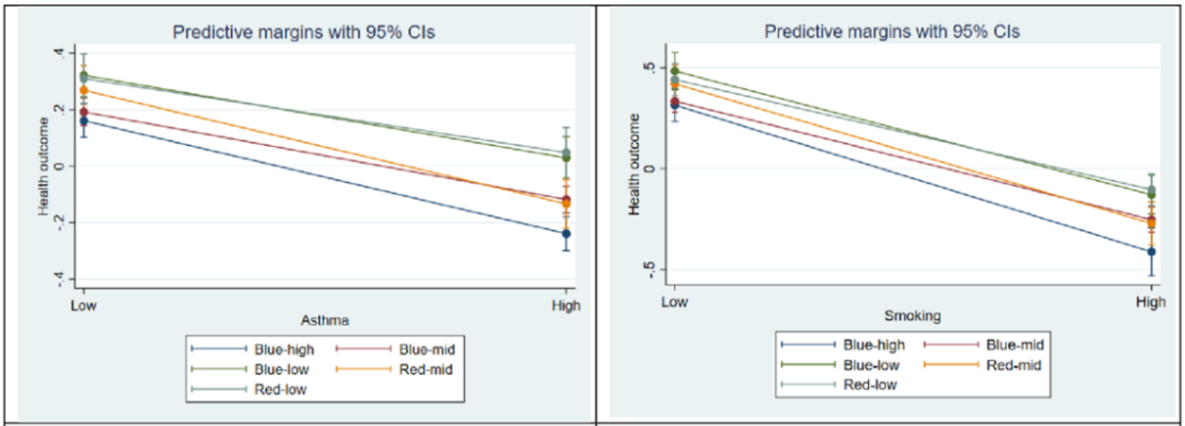
Table 3. City level–random effect results of the linear mixed-effect model.

	Model 1 (blue-low)	Model 2 (blue-mid)	Model 3 (blue-high)	Model 4 (red-low)	Model 5 (red-high)
City groups–intercept, variance (SD)	0.116 (0.340)	0.094 (0.307)	0.143 (0.378)	0.358 (0.189)	0.068 (0.260)
Residual, variance (SD)	0.050 (0.223)	0.064 (0.252)	0.055 (0.235)	0.054 (0.231)	0.059 (0.243)
Random effects	0.2225	0.1936	0.2861	0.0926	0.1416

Table 4. Overall model statistics and performance of the linear mixed-effect model.

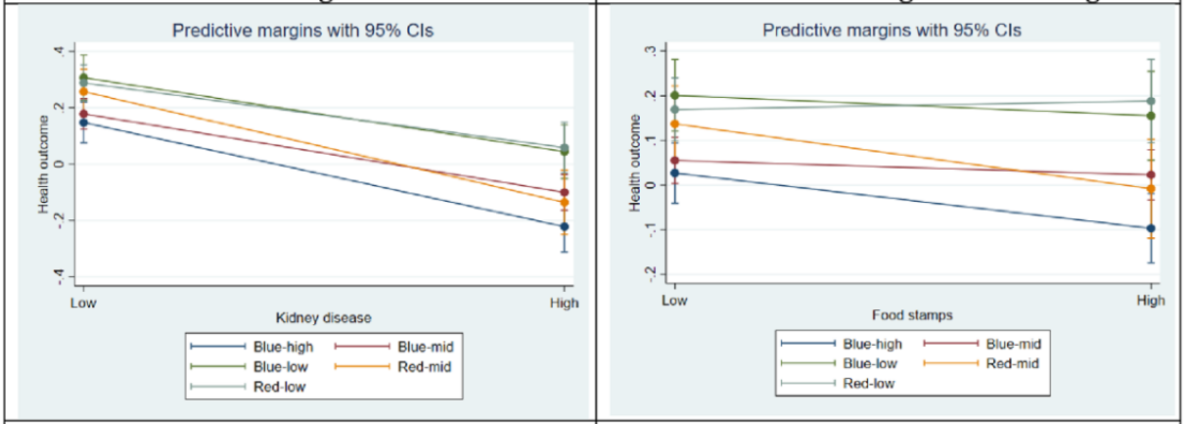
	Model 1 (blue-low)	Model 2 (blue-mid)	Model 3 (blue-high)	Model 4 (red-low)	Model 5 (red-high)
Number of census tracts	3350	9803	9144	1991	1728
Number of city	73	198	120	53	55
R^2 marginal—census tract level	0.6816	0.6759	0.6034	0.7690	0.7351
R^2 conditional—city level	0.9041	0.8695	0.8895	0.8616	0.8767

Figure 1. Post hoc analysis via predictive margins for (A) asthma, (B) smoking, (C) kidney disease, and (D) food stamps. Each colored line represents each city grouping. The x-axis shows whether the variable is high or low, where low means 1 SD below the mean and high means 1 SD above the mean. The y-axis displays the value of health outcome. This figure applies to all 29,126 census tracts within 499 cities in the United States in 2021.



A. Predictive margins of asthma

B. Predictive margins of smoking



C. Predictive margins of kidney disease

D. Predictive margins of food stamps

Discussion

Principal Findings

In this study, we have demonstrated that, with the use of census tracts as the unit of analysis, more insights are extracted regarding how SDOH factors would affect the health outcomes of a diverse population. Furthermore, aggregating data into city jurisdictions enables a more targeted examination of how each city operates and influences health outcomes with an eye on policy generation and implementation at the city level.

Using a unique grouping method for cities based on population density and political affiliation, we reveal that cities are similar in many ways yet exhibit remarkable differences in public health. Specifically, we observed a major divide in public health and the different impacts of SDOH factors on population health outcomes between different political support in the United States, which contains only 2 main political parties. The results reinforce the current political climate in the United States, namely, the polarization between the 2 parties and how it affects health care outcome [69-71]. These findings add complexity for city officials who are monitoring and maintaining the quality of public health while upholding the political ideals and expectations through policy enforcement. Similarly, population density, which substitutes urban-rural designations, adds to the dichotomy of health. When cities are in the same grouping, they share many aspects of SDOH effects on health outcomes, although there is a distinct difference among members of different groups.

Our research yielded several interesting findings related to different SDOH factors. For example, we found that the number of adults who lost all their natural teeth before age 65 contributed to health outcomes in an oscillating manner. Although there is an established link between oral health and quality of life, there seems to be a disconnect between oral health and population health outcomes in the literature [72,73]. In addition, we observed the interrelationship between access to health services and annual checkups, highlighting the importance of health insurance for improving health outcomes. The health care system in the United States is complex, and obtaining health insurance for the disadvantaged group has been a major roadblock. Having health insurance is not enough; with it alone, there is only a negligible uptick in annual checkups [74]. These findings strongly support the need to establish incentives for having more individuals perform annual checkups. For example, the Preventive health Evidence-based Recommendation Form is a program that can potentially fulfill such a need [75].

This study has several implications for public health. First, we highlight that public health policies should differ among cities with different population densities and political affiliations. For instance, policies targeting the population with asthma, such as promoting environmental cleanliness, reducing pollution particles, and reducing the costs of asthma treatment and medication, would have a much stronger effect in red-leaning cities with a high population density than in cities in other groups. Likewise, smoking affects the blue-leaning cities the most, so a policy to discourage smoking and promote quitting would be more beneficial to population health outcomes in those cities. As resources are scarce, cities should allocate their resources according to the effectiveness of the proposed solutions. Therefore, each devised policy could target different SDOH factors for policy interventions to optimize health outcomes based on the categorization of each city. Second, our findings provide additional variables of interest for invigorating public health policy transfer possibilities among cities. Policy transfer is a well-studied phenomenon worldwide, especially in the European Union, but it has been much less studied in the United States. This study indicates that policy transfer between cities in the same group is possible. For example, a blue-leaning city with low population density might try to focus more on reducing smoking by perusing policies from health care offices residing in other blue-leaning rural cities. Similarly, cities may be able to pool resources together and procure a repository containing all related health policies. This repository could help facilitate faster and more efficient knowledge transfer between cities across the United States.

Limitations

This study has several limitations. First, the curated data were not equivalent to those of the CHR. Only a handful of equivalent variables are available at the census tract level. Second, data normalization in the CHR encompasses all counties in the United States, although this study limits the data set to 499 cities, potentially skewing the results. Third, although the cities may represent many populations, sparse suburban areas and rural areas are largely neglected owing to data availability. Fourth, as an ecological study, there are possible confounders that existed outside our data set. Finally, the ecological nature of the study prohibits generalization of conclusions such as giving more food stamps to an individual, which could result in a change in personal health outcomes. Future research should continue to explore the relationship between SDOH and health outcomes in other delineations suitable for policy decision making.

Conclusions

This study aimed to investigate the impact of SDOH factors on population health outcomes using a large data set comprising 29,126 census tracts within 499 cities across all 50 states in the United States. Our results identified 4 SDOH factors, namely, asthma, kidney disease, smoking, and food stamps, that have major effects across cities with different population densities and political affiliations. In addition, this study highlights the need for differentiated public health policies among cities with different population densities and political affiliations. The analysis of data at the city level, in which policies and decisions directly affect its citizens, promotes an understanding of how SDOH factors affect population health outcomes. The grouping mechanism, based on the combination of population density and political affiliation, provides a useful framework for separating and comparing different census tracts in different cities. To that end, this study adds to the existing literature on various ways to improve health equity among geographic areas or demographic and socioeconomic groups.

Acknowledgments

This work was partially funded by the Le Family Endowed Faculty Opportunity Fund at the College of Business Administration, Loyola Marymount University, California, United States. The Le Family Endowed Faculty Opportunity Fund had no direct or indirect influence on the results or direction of this research. AV and YT thank the Le Family for their support.

Data Availability

The data set will be available upon reasonable request to the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of 499 cities, breaking down by 5 groups: blue-low, blue-mid, blue-high, red-low, and red-high.
[DOCX File , 31 KB - publichealth_v9i1e44070_app1.docx]

Multimedia Appendix 2
List of the initial 148 independent variables.
[DOCX File , 22 KB - publichealth_v9i1e44070_app2.docx]

References

1. Social determinants of health. World Health Organization. URL: https://www.who.int/health-topics/social-determinants-of-health [accessed 2022-04-13]

2. Marmot M, Friel S, Bell R, Houweling TA, Taylor S, Commission on Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health. Lancet 2008 Nov 08;372(9650):1661-1669. [doi: 10.1016/S0140-6736(08)61690-6] [Medline: 18994664]
3. Wilder ME, Kulie P, Jensen C, Levett P, Blanchard J, Dominguez LW, et al. The impact of social determinants of health on medication adherence: a systematic review and meta-analysis. J Gen Intern Med 2021 May;36(5):1359-1370 [FREE Full text] [doi: 10.1007/s11606-020-06447-0] [Medline: 33515188]
4. Gadson A, Akpovi E, Mehta PK. Exploring the social determinants of racial/ethnic disparities in prenatal care utilization and maternal outcome. Semin Perinatol 2017 Aug;41(5):308-317. [doi: 10.1053/j.semperi.2017.04.008] [Medline: 28625554]
5. Evans WN, Kroeger S, Munnich EL, Ortuzar G, Wagner KL. Reducing readmissions by addressing the social determinants of health. Am J Health Econ 2021;7(1):1-40 [FREE Full text] [doi: 10.1086/711725]
6. Meddings J, Reichert H, Smith SN, Iwashyna TJ, Langa KM, Hofer TP, et al. The impact of disability and social determinants of health on condition-specific readmissions beyond Medicare risk adjustments: a cohort study. J Gen Intern Med 2017 Jan;32(1):71-80 [FREE Full text] [doi: 10.1007/s11606-016-3869-x] [Medline: 27848189]
7. Delanois RE, Tarazi JM, Wilkie WA, Remily E, Salem HS, Mohamed NS, et al. Social determinants of health in total knee arthroplasty : are social factors associated with increased 30-day post-discharge cost of care and length of stay? Bone Joint J 2021 Jun;103-B(6 Supple A):113-118. [doi: 10.1302/0301-620X.103B6.BJJ-2020-2430.R1] [Medline: 34053276]
8. Hedinger D, Hämmig O, Bopp M, Swiss National Cohort Study Group. Social determinants of duration of last nursing home stay at the end of life in Switzerland: a retrospective cohort study. BMC Geriatr 2015 Oct 01;15:114 [FREE Full text] [doi: 10.1186/s12877-015-0111-3] [Medline: 26429705]
9. Paro A, Hyer JM, Diaz A, Tsilimigras DI, Pawlik TM. Profiles in social vulnerability: the association of social determinants of health with postoperative surgical outcomes. Surgery 2021 Dec;170(6):1777-1784. [doi: 10.1016/j.surg.2021.06.001] [Medline: 34183179]
10. Chan KS, Roberts E, McCleary R, Buttorff C, Gaskin DJ. Community characteristics and mortality: the relative strength of association of different community characteristics. Am J Public Health 2014 Sep;104(9):1751-1758. [doi: 10.2105/AJPH.2014.301944] [Medline: 25033152]
11. Gerend MA, Pai M. Social determinants of Black-White disparities in breast cancer mortality: a review. Cancer Epidemiol Biomarkers Prev 2008 Nov;17(11):2913-2923. [doi: 10.1158/1055-9965.EPI-07-0633] [Medline: 18990731]
12. Burström B, Tao W. Social determinants of health and inequalities in COVID-19. Eur J Public Health 2020 Aug 01;30(4):617-618 [FREE Full text] [doi: 10.1093/eurpub/ckaa095] [Medline: 32638998]
13. Singu S, Acharya A, Challagundla K, Byrareddy SN. Impact of social determinants of health on the emerging COVID-19 pandemic in the United States. Front Public Health 2020 Jul 21;8:406 [FREE Full text] [doi: 10.3389/fpubh.2020.00406] [Medline: 32793544]
14. Brakefield WS, Ammar N, Olusanya OA, Shaban-Nejad A. An urban population health observatory system to support COVID-19 pandemic preparedness, response, and management: design and development study. JMIR Public Health Surveill 2021 Jun 16;7(6):e28269 [FREE Full text] [doi: 10.2196/28269] [Medline: 34081605]
15. Brakefield WS, Olusanya OA, Shaban-Nejad A. Association between neighborhood factors and adult obesity in Shelby county, Tennessee: geospatial machine learning approach. JMIR Public Health Surveill 2022 Aug 09;8(8):e37039 [FREE Full text] [doi: 10.2196/37039] [Medline: 35943795]
16. Frieden TR. A framework for public health action: the health impact pyramid. Am J Public Health 2010 Apr;100(4):590-595. [doi: 10.2105/AJPH.2009.185652] [Medline: 20167880]
17. Remington PL, Catlin BB, Gennuso KP. The County Health Rankings: rationale and methods. Popul Health Metr 2015 Apr 17;13:11 [FREE Full text] [doi: 10.1186/s12963-015-0044-2] [Medline: 25931988]
18. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. World Health Organization. 2010. URL: https://apps.who.int/iris/bitstream/handle/10665/44489/9789241500852_eng.pdf [accessed 2023-03-01]
19. Danaher A. Reducing disparities and improving population health: the role of a vibrant community sector flip sheet. Wellesley Institute. 2011 Aug 17. URL: https://www.wellesleyinstitute.com/publications/reducing-disparities-and-improving-population-health-the-role-of-a-vibrant-community-sector/ [accessed 2022-10-31]
20. County Health Rankings Model. County Health Rankings and Roadmaps, Robert Wood Johnson Foundation Program. 2016 Mar 29. URL: https://www.countyhealthrankings.org/resources/county-health-rankings-model [accessed 2020-11-12]
21. Davis W, Gordan A, Tchernis R. Measuring the spatial distribution of health rankings in the United States. Health Econ 2021 Nov;30(11):2921-2936. [doi: 10.1002/hec.4416] [Medline: 34476867]
22. Hood CM, Gennuso KP, Swain GR, Catlin BB. County health rankings: relationships between determinant factors and health outcomes. Am J Prev Med 2016 Feb;50(2):129-135. [doi: 10.1016/j.amepre.2015.08.024] [Medline: 26526164]
23. Corburn J, Curl S, Arredondo G, Malagon J. Health in All Urban Policy: city services through the prism of health. J Urban Health 2014 Aug;91(4):623-636 [FREE Full text] [doi: 10.1007/s11524-014-9886-3] [Medline: 25047156]
24. Social Determinants of Health at CDC. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/socialdeterminants/index.htm [accessed 2022-04-13]
25. Healthy people 2030: building a healthier future for all. US Department of Health and Human Services. URL: https://health.gov/healthypeople [accessed 2020-11-12]

26. Braveman P, Gruskin S. Defining equity in health. J Epidemiol Community Health 2003 Apr;57(4):254-258 [FREE Full text] [doi: 10.1136/jech.57.4.254] [Medline: 12646539]
27. Marmot M. Social determinants of health inequalities. Lancet 2005 Mar;365(9464):1099-1104 [FREE Full text] [doi: 10.1016/S0140-6736(05)71146-6] [Medline: 15781105]
28. Census Tracts. United States Census Bureau. 2015. URL: https://www2.census.gov/geo/pdfs/education/CensusTracts.pdf [accessed 2022-05-16]
29. Elixhauser A, Machlin SR, Zodet MW, Chevarley FM, Patel N, McCormick MC, et al. Health care for children and youth in the United States: 2001 annual report on access, utilization, quality, and expenditures. Ambul Pediatr 2002;2(6):419-437. [doi: 10.1367/1539-4409(2002)002<0419:hcfcay>2.0.co;2] [Medline: 12437388]
30. Washington DL, Bean-Mayberry B, Riopelle D, Yano EM. Access to care for women veterans: delayed healthcare and unmet need. J Gen Intern Med 2011 Nov;26 Suppl 2(Suppl 2):655-661 [FREE Full text] [doi: 10.1007/s11606-011-1772-z] [Medline: 21989618]
31. Bryant T, Raphael D, Schrecker T, Labonte R. Canada: a land of missed opportunity for addressing the social determinants of health. Health Policy 2011 Jun;101(1):44-58. [doi: 10.1016/j.healthpol.2010.08.022] [Medline: 20888059]
32. Fang P, Dong S, Xiao J, Liu C, Feng X, Wang Y. Regional inequality in health and its determinants: evidence from China. Health Policy 2010 Jan;94(1):14-25. [doi: 10.1016/j.healthpol.2009.08.002] [Medline: 19735959]
33. Bauer AM, Schillinger D, Parker MM, Katon W, Adler N, Adams AS, et al. Health literacy and antidepressant medication adherence among adults with diabetes: the diabetes study of Northern California (DISTANCE). J Gen Intern Med 2013 Sep;28(9):1181-1187 [FREE Full text] [doi: 10.1007/s11606-013-2402-8] [Medline: 23512335]
34. Park HS, White RS, Ma X, Lui B, Pryor KO. Social determinants of health and their impact on postcolectomy surgery readmissions: a multistate analysis, 2009-2014. J Comp Eff Res 2019 Dec;8(16):1365-1379 [FREE Full text] [doi: 10.2217/cer-2019-0114] [Medline: 31797679]
35. Howell CR, Zhang L, Yi N, Mehta T, Garvey WT, Cherrington AL. Race versus social determinants of health in COVID-19 hospitalization prediction. Am J Prev Med 2022 Jul;63(1 Suppl 1):S103-S108 [FREE Full text] [doi: 10.1016/j.amepre.2022.01.034] [Medline: 35725136]
36. Skolarus LE, Sharrief A, Gardener H, Jenkins C, Boden-Albala B. Considerations in addressing social determinants of health to reduce racial/ethnic disparities in stroke outcomes in the United States. Stroke 2020 Nov;51(11):3433-3439 [FREE Full text] [doi: 10.1161/STROKEAHA.120.030426] [Medline: 33104471]
37. Wesselman H, Ford CG, Leyva Y, Li X, Chang CC, Dew MA, et al. Social determinants of health and race disparities in kidney transplant. Clin J Am Soc Nephrol 2021 Feb 08;16(2):262-274 [FREE Full text] [doi: 10.2215/CJN.04860420] [Medline: 33509963]
38. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
39. Andersen RM, Davidson PL, Baumeister SE. Improving access to care in America: individual and contextual indicators. In: Andersen RM, Rice TH, Kominski GF, editors. Changing the U.S. Health Care System: Key Issues in Health Services Policy and Management. 3rd edition. Hoboken, NJ, USA: Jossey-Bass; 2007:3-31.
40. Reports | Healthy People 2020. U.S. Department of Health and Human Services. URL: https://www.healthypeople.gov/2020/about/advisory/Reports [accessed 2022-04-13]
41. Tobacco-Related Mortality. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/tobacco/data_statistics/fact_sheets/health_effects/tobacco_related_mortality/index.htm [accessed 2022-04-13]
42. Classification and Delineation of Urban Areas in a Census. United States Census Bureau. 2020 Jun. URL: https://www.census.gov/content/dam/Census/library/working-papers/2020/demo/urban-stic.pdf [accessed 2022-05-16]
43. Strömsöe A, Svensson L, Claesson A, Lindkvist J, Lundström A, Herlitz J. Association between population density and reported incidence, characteristics and outcome after out-of-hospital cardiac arrest in Sweden. Resuscitation 2011 Oct;82(10):1307-1313. [doi: 10.1016/j.resuscitation.2011.04.025] [Medline: 21628082]
44. Holzer III CE, Goldsmith HF, Ciarlo JA. The availability of health and mental health providers by population density. J Wash Acad Sci 2000 Dec;86(3):25-33.
45. Nayar P, Yu F, Apenteng BA. Frontier America's health system challenges and population health outcomes. J Rural Health 2013 Jun;29(3):258-265. [doi: 10.1111/j.1748-0361.2012.00451.x] [Medline: 23802928]
46. Nong S, Chen Z. Whither the roads lead to? Estimating association between urbanization and primary healthcare service use with Chinese prefecture-level data in 2014. PLoS One 2020 Jun 3;15(6):e0234081 [FREE Full text] [doi: 10.1371/journal.pone.0234081] [Medline: 32492048]
47. Sharif N, Ahmed SN, Opu RR, Daullah MU, Khan S, Talukder AA, et al. Impact of meteorological parameters and population density on variants of SARS-CoV-2 and outcome of COVID-19 pandemic in Japan. Epidemiol Infect 2021 Apr 28;149:e103 [FREE Full text] [doi: 10.1017/S095026882100100X] [Medline: 33908339]
48. Carnegie ER, Inglis G, Taylor A, Bak-Klimek A, Okoye O. Is population density associated with non-communicable disease in western developed countries? A systematic review. Int J Environ Res Public Health 2022 Feb 24;19(5):2638 [FREE Full text] [doi: 10.3390/ijerph19052638] [Medline: 35270337]

49. Nehme Z, Andrew E, Cameron PA, Bray JE, Bernard SA, Meredith IT, et al. Population density predicts outcome from out-of-hospital cardiac arrest in Victoria, Australia. Med J Aust 2014 May 05;200(8):471-475. [doi: 10.5694/mja13.10856] [Medline: 24794610]
50. Blendon RJ, Benson JM. Implications of the 2020 election for U.S. health policy. N Engl J Med 2020 Oct 29;383(18):e105. [doi: 10.1056/NEJMsr2031592] [Medline: 33113302]
51. Pagel C, Bates DW, Goldmann D, Koller CF. A way forward for bipartisan health reform? Democrat and republican state legislator priorities for the goals of health policy. Am J Public Health 2017 Oct;107(10):1601-1603. [doi: 10.2105/AJPH.2017.304023] [Medline: 28817318]
52. Baum MA. Red state, blue state, flu state: media self-selection and partisan gaps in swine flu vaccinations. J Health Polit Policy Law 2011 Dec;36(6):1021-1059. [doi: 10.1215/03616878-1460569] [Medline: 21948819]
53. Fridman A, Gershon R, Gneezy A. COVID-19 and vaccine hesitancy: a longitudinal study. PLoS One 2021 Apr 16;16(4):e0250123 [FREE Full text] [doi: 10.1371/journal.pone.0250123] [Medline: 33861765]
54. Embrett MG, Randall GE. Social determinants of health and health equity policy research: exploring the use, misuse, and nonuse of policy analysis theory. Soc Sci Med 2014 May;108:147-155. [doi: 10.1016/j.socscimed.2014.03.004] [Medline: 24641879]
55. Knight EK, Benjamin GD, Yanich D. Framing social determinants of health within the professional public health community: research translation and implications for policy change. J Appl Commun Res 2016 Jun 16;44(3):256-274. [doi: 10.1080/00909882.2016.1192291]
56. Lantz PM, Lichtenstein RL, Pollack HA. Health policy approaches to population health: the limits of medicalization. Health Aff (Millwood) 2007;26(5):1253-1257. [doi: 10.1377/hlthaff.26.5.1253] [Medline: 17848434]
57. Qureshi S. Creating a better world with information and communication technologies: health equity. Inf Technol Dev 2016 Jan 22;22(1):1-14. [doi: 10.1080/02681102.2015.1121585]
58. Sen A. Why health equity? In: Anand S, Peter F, Sen A, editors. Public Health, Ethics, and Equity. Oxford, UK: Oxford University Press; 2004:21-33.
59. Gamble VN, Stone D. U.S. policy on health inequities: the interplay of politics and research. J Health Polit Policy Law 2006 Feb;31(1):93-126. [doi: 10.1215/03616878-31-1-93] [Medline: 16484670]
60. Rydin Y, Bleahu A, Davies M, Dávila JD, Friel S, De Grandis G, et al. Shaping cities for health: complexity and the planning of urban environments in the 21st century. Lancet 2012 Jun 02;379(9831):2079-2108 [FREE Full text] [doi: 10.1016/S0140-6736(12)60435-8] [Medline: 22651973]
61. PLACES: Local Data for Better Health. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/places/index.html [accessed 2022-04-13]
62. MIT Election Data and Science Lab - County Presidential Election Returns 2000-2020. Harvard Dataverse. 2022. URL: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ [accessed 2023-03-01]
63. Life Expectancy Data Viz. National Center for Health Statistics. 2022. URL: https://www.cdc.gov/nchs/data-visualization/life-expectancy/index.html [accessed 2022-04-13]
64. Bates D, Maechler M, Bolker B, Walker S. lme4: Mixed-effects modeling with R. The Comprehensive R Archive Network. 2022 Nov 1. URL: https://cran.r-project.org/web/packages/lme4/lme4.pdf [accessed 2023-03-01]
65. Shahidi FV, Ramraj C, Sod-Erdene O, Hildebrand V, Siddiqi A. The impact of social assistance programs on population health: a systematic review of research in high-income countries. BMC Public Health 2019 Jan 03;19(1):2 [FREE Full text] [doi: 10.1186/s12889-018-6337-1] [Medline: 30606263]
66. Quiñones J, Hammad Z. Social determinants of health and chronic kidney disease. Cureus 2020 Sep 05;12(9):e10266 [FREE Full text] [doi: 10.7759/cureus.10266] [Medline: 33042704]
67. Trivedi M, Pappalardo AA, Udoko M, Garg A, Phipatanakul W, Szefler SJ, et al. Social determinants of health in asthma through the life course. J Allergy Clin Immunol Pract 2022 Apr;10(4):953-961. [doi: 10.1016/j.jaip.2021.12.032] [Medline: 35033702]
68. Wang G, Wu L. Healthy people 2020: social determinants of cigarette smoking and electronic cigarette smoking among youth in the United States 2010-2018. Int J Environ Res Public Health 2020 Oct 15;17(20):E7503 [FREE Full text] [doi: 10.3390/ijerph17207503] [Medline: 33076356]
69. Casola WR, Beall JM, Nils Peterson MN, Larson LR, Brent Jackson SB, Stevenson KT. Political polarization of conservation issues in the era of COVID-19: an examination of partisan perspectives and priorities in the United States. J Nat Conserv 2022 Jun;67:126176 [FREE Full text] [doi: 10.1016/j.jnc.2022.126176] [Medline: 35370533]
70. Liu N. Political polarization in the United States: the influences of exceptionalism and religion. Int Soc Sci Rev 2020;96(2):4 [FREE Full text]
71. Padala A, Conrad L. Political polarization during 2017-2021. J Stud Res 2021 Nov 25;10(3):1817. [doi: 10.47611/jsrhs.v10i3.1817]
72. Bersell CH. Access to oral health care: a national crisis and call for reform. J Dent Hyg 2017 Feb;91(1):6-14. [Medline: 29118145]

73. Lahti S, Suominen-Taipale L, Hausen H. Oral health impacts among adults in Finland: competing effects of age, number of teeth, and removable dentures. Eur J Oral Sci 2008 Jun;116(3):260-266. [doi: 10.1111/j.1600-0722.2008.00540.x] [Medline: 18471245]
74. Hong YR, Jo A, Mainous 3rd AG. Up-to-date on preventive care services under affordable care act: a trend analysis from MEPS 2007-2014. Med Care 2017 Aug;55(8):771-780. [doi: 10.1097/MLR.0000000000000763] [Medline: 28671929]
75. Dubey V, Mathew R, Iglar K, Moineddin R, Glazier R. Improving preventive service delivery at adult complete health check-ups: the Preventive health Evidence-based Recommendation Form (PERFORM) cluster randomized controlled trial. BMC Fam Pract 2006 Jul 12;7:44 [FREE Full text] [doi: 10.1186/1471-2296-7-44] [Medline: 16836761]

Abbreviations

CHR: County Health Ranking
SDOH: social determinants of health

Edited by A Mavragani; submitted 04.11.22; peer-reviewed by E Carnegie, R Gold, N Mathews, O Okoye; comments to author 01.12.22; revised version received 21.12.22; accepted 23.02.23; published 29.03.23.

Please cite as:
Vo A, Tao Y, Li Y, Albarrak A
The Association Between Social Determinants of Health and Population Health Outcomes: Ecological Analysis
JMIR Public Health Surveill 2023;9:e44070
URL: https://publichealth.jmir.org/2023/1/e44070
doi:10.2196/44070
PMID:36989028

Original Paper

Contribution of Frailty to Multimorbidity Patterns and Trajectories: Longitudinal Dynamic Cohort Study of Aging People

Lucía A Carrasco-Ribelles[1,2,3,4,5], MSc; Margarita Cabrera-Bean[2], PhD; Marc Danés-Castells[3], MD; Edurne Zabaleta-del-Olmo[1,4,5,6,7], PhD; Albert Roso-Llorach[1,4,5,8], PhD; Concepción Violán[3,4,5,8,9], MD, PhD

[1]Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol I Gurina (IDIAPJGol), Barcelona, Spain

[2]Signal Processing and Communications Group (SPCOM), Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

[3]Unitat de Suport a la Recerca Metropolitana Nord, Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol I Gurina (IDIAPJGol), Mataró, Spain

[4]Grup de REcerca en Impacte de les Malalties Cròniques i les seves Trajectòries (GRIMTRA) (2021 SGR 01537), Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol I Gurina (IDIAPJGol), Barcelona, Spain

[5]Network for Research on Chronicity, Primary Care, and Health Promotion (RICAPPS) (RD21/0016/0029), Instituto de Salud Carlos III, Madrid, Spain

[6]Gerència Territorial de Barcelona, Institut Català de la Salut, Barcelona, Spain

[7]Nursing Department, Faculty of Nursing, Universitat de Girona, Girona, Spain

[8]Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

[9]Fundació Institut d'Investigació en ciències de la Salut Germans Trias i Pujol (IGTP), Badalona, Spain

Corresponding Author:
Concepción Violán, MD, PhD
Unitat de Suport a la Recerca Metropolitana Nord
Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol I Gurina (IDIAPJGol)
CAP La Llàntia
C/ Mare de Déu de Guadalupe 2
Mataró, 08303
Spain
Phone: 34 937415338
Email: cviolanf.mn.ics@gencat.cat

Abstract

Background: Multimorbidity and frailty are characteristics of aging that need individualized evaluation, and there is a 2-way causal relationship between them. Thus, considering frailty in analyses of multimorbidity is important for tailoring social and health care to the specific needs of older people.

Objective: This study aimed to assess how the inclusion of frailty contributes to identifying and characterizing multimorbidity patterns in people aged 65 years or older.

Methods: Longitudinal data were drawn from electronic health records through the SIDIAP (Sistema d’Informació pel Desenvolupament de la Investigació a l’Atenció Primària) primary care database for the population aged 65 years or older from 2010 to 2019 in Catalonia, Spain. Frailty and multimorbidity were measured annually using validated tools (eFRAGICAP, a cumulative deficit model; and Swedish National Study of Aging and Care in Kungsholmen [SNAC-K], respectively). Two sets of 11 multimorbidity patterns were obtained using fuzzy c-means. Both considered the chronic conditions of the participants. In addition, one set included age, and the other included frailty. Cox models were used to test their associations with death, nursing home admission, and home care need. Trajectories were defined as the evolution of the patterns over the follow-up period.

Results: The study included 1,456,052 unique participants (mean follow-up of 7.0 years). Most patterns were similar in both sets in terms of the most prevalent conditions. However, the patterns that considered frailty were better for identifying the population whose main conditions imposed limitations on daily life, with a higher prevalence of frail individuals in patterns like *chronic ulcers &peripheral vascular*. This set also included a dementia-specific pattern and showed a better fit with the risk of nursing home admission and home care need. On the other hand, the risk of death had a better fit with the set of patterns that did not include frailty. The change in patterns when considering frailty also led to a change in trajectories. On average, participants were in 1.8 patterns during their follow-up, while 45.1% (656,778/1,456,052) remained in the same pattern.

Conclusions: Our results suggest that frailty should be considered in addition to chronic diseases when studying multimorbidity patterns in older adults. Multimorbidity patterns and trajectories can help to identify patients with specific needs. The patterns that considered frailty were better for identifying the risk of certain age-related outcomes, such as nursing home admission or home care need, while those considering age were better for identifying the risk of death. Clinical and social intervention guidelines and resource planning can be tailored based on the prevalence of these patterns and trajectories.

(JMIR Public Health Surveill 2023;9:e45848) doi:10.2196/45848

KEYWORDS

multimorbidity; frailty; clustering; electronic health record; primary care; trajectory

Introduction

Aging is associated with the development of complex conditions, such as multimorbidity and frailty [1,2], which need to be assessed at the individual level. Frailty is a holistic state defined by the biological age-related loss of homeostasis and resistance to stressors, not by particular conditions, and it increases vulnerability to adverse outcomes [2-5]. Frailty can be measured either through the frailty phenotype [5] or cumulative deficit models that consider physical, psychological, and social domains [2,6,7]. On the other hand, multimorbidity is defined as the simultaneous presence of two or more chronic diseases [8]. There is a 2-way causal relationship between multimorbidity and frailty [4,9], and both are associated with higher health care utilization and expenditure [10-14]. Health systems thus need to characterize and monitor the older population to estimate health care and social resource demand.

Electronic health records (EHRs) are essential nowadays to monitor and evaluate patients [15]. Real-world studies use EHRs to obtain a large quantity of observational data from diverse populations. These data sources allow studies to be conducted at a lower cost than traditional epidemiological studies or randomized clinical trials [16,17] and can achieve similar results to randomized clinical trials [18]. Most multimorbidity studies in recent years have used EHRs [19], while EHR-based tools have also been recently developed to measure frailty [20,21].

Clustering is an unsupervised exploratory data analysis technique used for identifying and characterizing population groups. It has already been extensively used to find subgroups of people based on the similarity, in terms of co-occurrence, of their concurrent chronic disease [22-28]. Several systematic reviews [27-29] describe different clustering techniques used to group patients based on multimorbidity, including hierarchical clustering, exploratory factor analysis, multiple correspondence analysis, network analysis, and k-means. K-means and fuzzy c-means are the most common approaches [22-26]. K-means is a hard clustering algorithm that forces each record to belong to a single cluster, while fuzzy c-means is a soft clustering technique that allows records to be simultaneously assigned to multiple clusters through membership probability [30]. In our study, this fuzziness allowed individuals to belong to several clusters, thus creating clusters characterized by broader disease combinations. These techniques have also been applied to identify subgroups of people based on their frailty [31,32]. Although multimorbidity and frailty are strongly associated [4], our review identified only 1 study that considers both simultaneously to build clustering-based patterns [33]. Other authors have related multimorbidity clusters to frailty-related outcomes using regression models [10,26]. However, to our knowledge, no study has assessed whether frailty may influence well-established multimorbidity patterns.

Considering frailty in analyses of multimorbidity is important for tailoring health and social care to the specific needs of the ever-expanding population of elderly people [34]. Moreover, frailty and multimorbidity evolve as people age, and these patterns can change over time, defining a trajectory. Only a few studies have been found that explored these trajectories [19], using hidden Markov models [23,35], latent class growth analysis [36], and descriptive statistics [37]. Furthermore, identifying changes in multimorbidity patterns and trajectories when considering frailty can enrich our understanding of patients' complex care needs and inform social and health care service strategies [38]. Therefore, our primary aim was to assess how the inclusion of frailty contributes to identifying and characterizing multimorbidity patterns in people aged 65 years or older. Moreover, we described the trajectory of the multimorbidity patterns of individuals as they aged.

Methods

Study Design, Setting, Data Source, and Participants

This observational study followed a dynamic cohort from primary care services in Catalonia (Spain) from January 1, 2010, to December 31, 2019. The cohort was drawn from the Information System for Performing Primary Care Research (SIDIAP [Sistema d'Informació pel Desenvolupament de la Investigació a l'Atenció Primària]) database [39]. The SIDIAP database collects pseudoanonymized EHRs from 328 primary care centers in Catalonia managed by the Catalan Health Institute (CHI) since 2005, and it currently has EHRs on more than 8 million patients. This represents almost 80% of the Catalan population and is a reliable representation of the region in terms of age, sex, and geographic distribution [40].

Participants were included at baseline if they were aged 65 years or older in 2010, or were added over the study period as they turned 65 years or arrived in the catchment area (if already aged 65 years or older). They were followed until death, transfer out of the catchment area (lost to follow-up), or end of the study (December 31, 2019). Individuals with no available information, those who did not attend a primary care center over the study period, and those who were aged 100 years or older in 2010 were excluded. Of the initial sample of 1,702,062 individuals, 1,456,052 were finally included.

The CHI linked primary care data with hospital admission data from public health care providers to maintain the pseudoanonymization of the data for researchers. Data included (1) sociodemographic information (ie, sex, age, and socioeconomic status [41]), (2) visits to primary care (ie, date of visit, health professional, and institution visited), (3) clinical measures (eg, BMI, blood pressure, frailty, and dependency questionnaires), (4) all diagnoses made in primary care (using International Classification of Diseases, 10th revision [ICD-10]), (5) laboratory results (eg, cholesterol and glycated hemoglobin), (6) emergency admission episodes (ie, date, number of diagnoses at admission, and length of stay), (7) medications dispensed in pharmacies (using Anatomical Therapeutic Classification [ATC] 5th level), and (8) inclusion in social assistance programs. Socioeconomic status was analyzed by census tract according to a 5-category classification, which considers 22 indicators, for instance, the proportion of the population with a manual occupation or dependency, households without internet access, and single-parent households.

This study complied with the RECORD (Reporting of Studies Conducted using Observational Routinely-collected Data) statement [42] (Multimedia Appendix 1).

Measurement of Multimorbidity and Frailty

Multimorbidity was measured using the operational definition of the Swedish National Study of Aging and Care in Kungsholmen (SNAC-K), which defined 60 categories of chronic conditions using more than 900 ICD-10 codes, along with clinical, laboratory, and drug-related parameters for assessing certain conditions [43]. The SNAC-K definition of multimorbidity is widely used in studies on older populations, so our results are amenable to comparisons with other studies. Frailty was measured using eFRAGICAP, a validated tool that uses EHRs from Catalan primary care centers [21]. This index considers 36 possible deficits that can be extracted from the EHRs, with 20 related to diseases and 16 related to signs, symptoms, laboratory results, and disabilities. According to the proportion of deficits a person has, their frailty status can be obtained using the cutoff points proposed by Clegg et al [20] (ie, fit, <0.12; mild, 0.12-0.24; moderate, 0.24-0.36; and severe ≥0.36). The complete list of codes considered in both multimorbidity and frailty definitions can be found in [43] and [21], respectively.

Statistical Analysis

Following study approval, data were obtained from SIDIAP. All authors had access to the database. Diagnoses with inconsistent dates and wrong sex-specific diagnoses were excluded. Duplicated diagnoses and clinical measures (same person, same day, and same code) were also excluded. The presence of each of the 60 disease groups and 36 deficits was calculated annually for each participant, according to which conditions were active and which laboratory results or clinical measures were out of range in the participants' EHRs [21,43]. There were no missing values related to diagnoses or frailty, as a lack of information was interpreted as the absence of the condition or frailty deficit, not as a loss of information. Continuous variables were described using medians and IQRs, as testing for normality showed a nonparametric distribution in all cases, and categorical variables were expressed as absolute and relative frequencies. Clustering and Cox regressions were performed on R v4.1 (R Project for Statistical Computing). Statistical significance was defined as $P<.05$ (2-sided).

Clustering Analysis

In this study, the information for each included person and year was used to group people based on the similarity of their combined concurrent chronic diseases. Each individual in the clustering analysis contributed records for each year they were included in the study. These groups represented multimorbidity patterns and were found using fuzzy c-means and 2 sets of data. Fuzzy c-means is a fuzzy form of clustering in which records for each individual can be assigned to more than one cluster, or multimorbidity pattern, through fuzzy membership, allowing the pattern definition to be more diverse. Both sets considered chronic conditions, as defined by SNAC-K; however, *multimorbidity & age* also included the age associated with the record, while *multimorbidity & frailty* considered the number of frailty deficits. A detailed description of the clustering analysis can be found in Multimedia Appendix 2 [21,43-47].

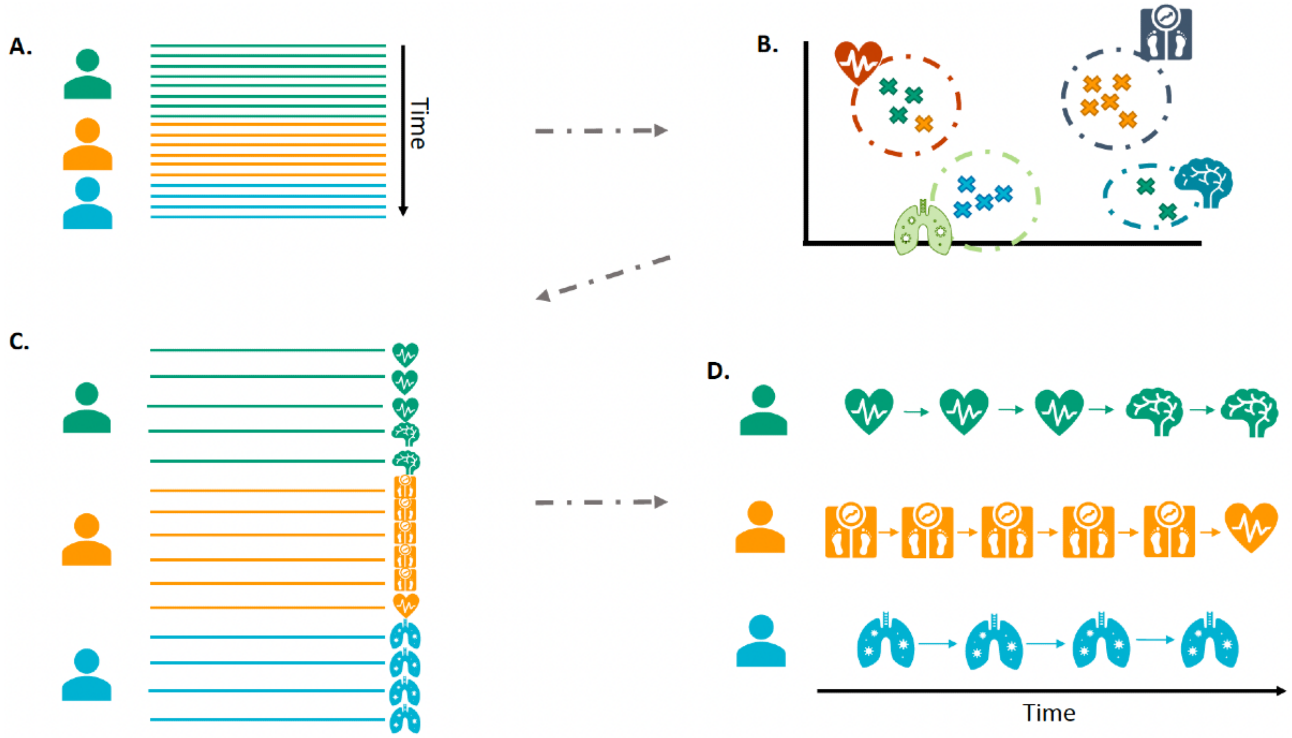
Dimensionality was reduced before clustering to simultaneously reduce computational cost and obtain more meaningful variables. First, chronic conditions with a low mean annual prevalence (<2%) were removed. Second, a PCAmix transformation [44], which is a mixture of the well-known Principal Component Analysis (PCA) and Multiple Correspondence Analysis (MCA), was applied, and a dimension reduction was performed using the Karlis-Saporta-Spinaki rule [45]. The choice of both the number of clusters (k) and the degree of fuzziness (m) was validated between $k \in (2, 15)$ and $m \in (1.1, 1.2, 1.4, 1.8)$, calculating analytical indexes using a subset of 100,000 randomly selected participants and 100 repetitions to account for random initialization of the cluster centroids. In addition to the analytical indexes, the opinion on the clinical usefulness and validity of the different sets of patterns of the research team was also considered to select the final k. This approach has been used in other studies [22-24,26].

Description of the Identified Patterns and Trajectories

To characterize the patterns, each person's annual record was assigned to the pattern with the highest membership probability. The observed/expected (OE) ratio, that is, the ratio between the condition prevalence in the pattern and the condition prevalence in the overall population, and exclusivity, that is, the ratio between the number of individuals in the pattern with the condition and the total number of individuals in the population with the condition, were calculated (see Multimedia Appendix 2 [21,43-47]). Conditions were considered associated with a specific pattern when the exclusivity was ≥25% or the OE ratio was ≥2. The patterns were named in line with these conditions by consensus within the research team (2 general practitioners, 1 nurse, and 2 statisticians), aiming to maximize their clinical utility and consistency with previous literature. In addition, each pattern was described in terms of age, sex, socioeconomic status, multimorbidity and frailty prevalence, smoking and alcohol intake, and health care service use.

The clustering model demonstrated the probability of every record and person belonging to each pattern, showing which pattern was most likely for each person each year. Therefore, the evolution among patterns over the study period could be followed as shown in Figure 1. Alluvial and chord plots were used to describe the trajectories, focusing on their evolution with aging, and a transition matrix showed the probability of change from the pattern assigned in the first year of inclusion to that in the last year.

Figure 1. Visual summary of the process of obtaining multimorbidity patterns and trajectories. A: Electronic Health Records (EHR) for three subjects, represented as a line per year of follow-up. B: Clustering gathers each year of each patient. C: Multimorbidity Pattern Assignment. D: Multimorbidity Trajectory.



Cox Regression

The association between the patterns identified for each data set and the following outcomes was calculated: all-cause mortality, nursing home admission, and home care need. The last 2 outcomes were measured using ICD-10 codes (Z59.3 and a Z74 starting code, respectively). The time to event was defined as the interval between cohort entry and the event. Patients were followed until censored (event, lost to follow-up, or end of observation). Cox proportional hazard regression models were fitted to test the association between the patterns and mortality. Similarly, cause-specific Cox models were calculated for nursing home admission and home care need, considering the competing risk of death through a multistate definition. For each model, the Akaike Information Criterion (AIC), R^2, and c-statistic (area under the receiver operating characteristic curve [AUC]) were calculated to assess the goodness of fit, the explained variation, and the predictive capacity of each set of patterns. The only covariate for building these models was the assigned pattern, and was considered time-varying, as each person was assigned to a pattern every year. The proportional hazard assumption was assured in all cases by checking the distribution of the Schoenfeld residuals.

Ethical Considerations

This study was approved by the Scientific and Ethical Committees of IDIAP (19/518-P) on December 18, 2019. The SIDIAP database is based on optout presumed consent. If a patient decides to opt out, their routine data are excluded from the database. Regarding the hospital admission data, the CHI acts as a trusted third party to execute the linkage and provide the pseudoanonymized data set, without needing informed consent. More information about the management of the SIDIAP database can be found in a previous report [40].

Results

Description of the Population

During the follow-up period, 1,456,052 unique participants were included in the study population (Figure 2 and Multimedia Appendix 3), with a mean follow-up of 7.04 (SD 3.15) years. The median age at cohort entry was 69.0 years, and 55.8% (813,074/1,456,052) were women. Most (1,297,810/1,456,052, 89.1%) joined the study with at least two chronic conditions; by the end of follow-up, this proportion was 94.5% (1,376,367/1,456,052). Frailty prevalence increased from 33.4% (486,320/1,456,052) to 60.3% (877,861/1,456,052) (Table 1). The prevalence of each chronic condition is presented in Table S2 in Multimedia Appendix 2.

Figure 2. Flow chart of the study population. The figure reports the number of individuals who met each exclusion criterion, as well as the number of individuals who met all the criteria simultaneously (unique IDs). CHI: Catalan Health Institute; SIDIAP: Sistema d'Informació pel Desenvolupament de la Investigació a l'Atenció Primària.

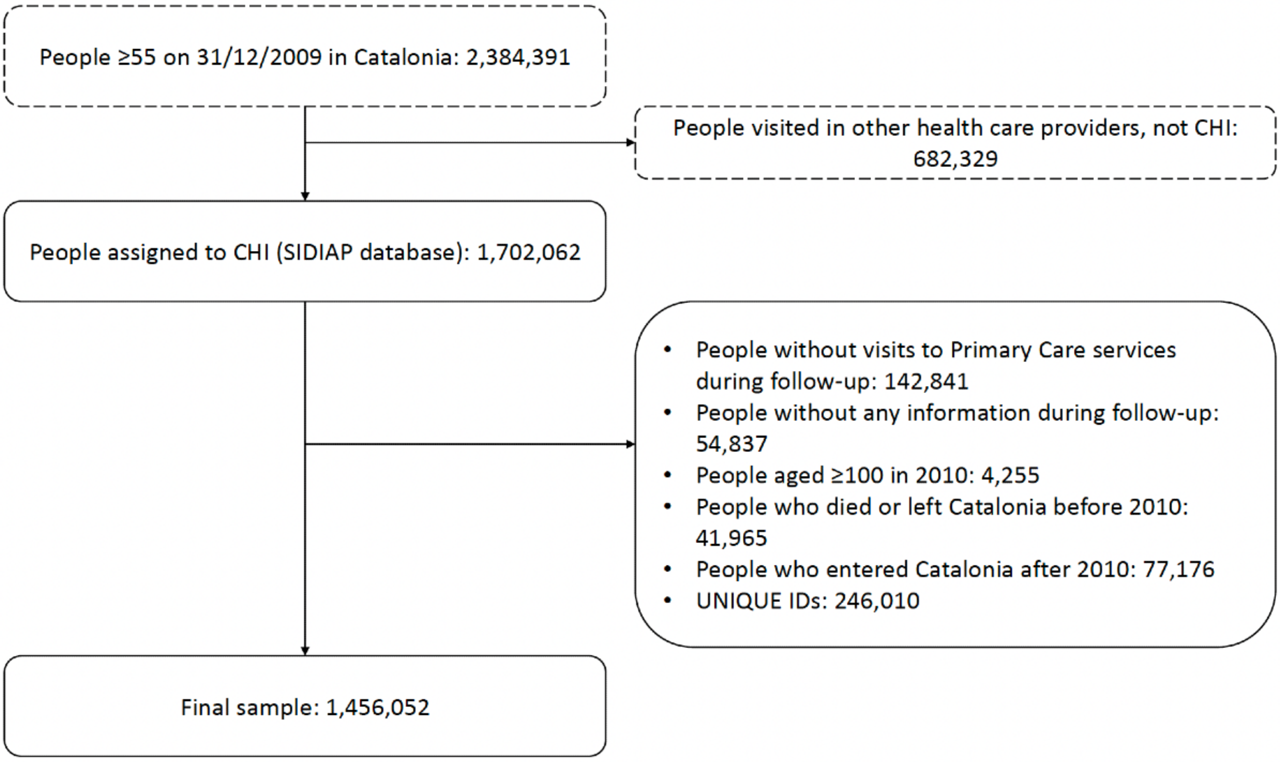


Table 1. Characteristics of the included individuals in their first and last years of follow-up (N=1,456,052).

Variable	First year of follow-up[a,b]	Last year of follow-up[a,b]
Age (years), median (IQR)	69.0 (65.0-77.0)	77.0 (70.0-85.0)
Sex, n (%)		
Women	813,074 (55.8)	813,074 (55.8)
Men	642,978 (44.2)	642,978 (44.2)
Deprivation index, n (%)		
1 (less deprived)	163,452 (13.2)	163,452 (13.2)
2	373,120 (30.2)	373,120 (30.2)
3	411,920 (33.4)	411,920 (33.4)
4	233,724 (18.9)	233,724 (18.9)
5 (more deprived)	51,594 (4.2)	51,594 (4.2)
Missing	222,242	222,242
Multimorbidity, median (IQR)		
SNAC-K[c] groups of chronic conditions	5.0 (3.0-7.0)	7.0 (5.0-10.0)
Type of multimorbidity, n (%)		
No multimorbidity	158,242 (10.9)	79,685 (5.5)
2-5 diseases	692,530 (47.6)	401,306 (27.6)
6-10 diseases	528,865 (36.3)	675,798 (46.4)
>10 diseases	76,415 (5.3)	299,263 (20.6)
Frailty, median (IQR)		
Deficits	3.0 (2.0-5.0)	6.0 (3.0-9.0)
Disease-related deficits	1.0 (1.0-2.0)	3.0 (1.0-4.0)
SSLD[d] deficits	2.0 (1.0-3.0)	3.0 (1.0-4.0)
Type of frailty, n (%)		
Fit	969,732 (66.6)	578,191 (39.7)
Mild	404,665 (27.8)	511,505 (35.1)
Moderate	71,950 (4.9)	263,893 (18.1)
Severe	9705 (0.7)	102,463 (7.0)
Smoking status, n (%)		
Nonsmoker	845,869 (65.6)	854,407 (61.1)
Exsmoker	284,398 (22.1)	418,086 (29.9)
Smoker	159,178 (12.3)	125,421 (9.0)
Missing	166,607	58,138
Alcohol intake, n (%)		
Nondrinker	412,938 (66.6)	445,245 (66.6)
Low-risk drinker	195,815 (31.6)	217,360 (32.5)
High-risk drinker	10,916 (1.8)	6,002 (0.9)
Missing	836,383	787,445
Health care service use, median (IQR)		
Visits to primary care	9.0 (4.0-16.0)	9.0 (4.0-17.0)
Distinct drugs[e]	7.0 (3.0-12.0)	7.0 (3.0-11.0)
Clinical measurements	5.0 (0.0-11.0)	5.0 (0.0-14.0)

Variable	First year of follow-up[a,b]	Last year of follow-up[a,b]
Laboratory measurements	16.0 (0.0-20.0)	16.0 (0.0-24.0)
Receiving home care, n (%)	44,473 (3.1)	194,052 (13.3)
Living in a nursing home, n (%)	31,480 (2.2)	122,844 (8.4)
Death, n (%)	25,254 (1.7)	355,901 (24.4)

[a] Follow-up began in 2010 for individuals aged 65 years or older who were in the catchment area, or the year when they turned 65 years or arrived in the catchment area. The end of follow-up was 2019, the year of death, or the year they left the catchment area. The characteristics were calculated at the end of each year, using the records from that year.

[b] For categorical variables, missing values, if any, are excluded from the calculation of the percentage.

[c] SNAC-K: Swedish National Study of Aging and Care in Kungsholmen.

[d] SSLD: signs, symptoms, laboratory results, and disabilities.

[e] The number of distinct drugs was calculated using the first 5 digits of the Anatomical Therapeutic Classification (ATC) code.

Description of the Identified Multimorbidity Patterns

Two sets of 11 multimorbidity patterns were identified (*multimorbidity & age* and *multimorbidity & frailty*). The most prevalent chronic conditions in each pattern were similar between sets in most patterns, regardless of whether frailty was included (see Multimedia Appendix 4 for the complete description; a demonstration is shown in Table 2) and the characteristics of their members (see Multimedia Appendix 5 for the complete description; a demonstration is shown in Table 3). The following patterns were identified in both data sets: *allergy & migraine*, *chronic ulcers & peripheral vascular*, *diabetes & obesity*, *genitourinary & respiratory*, *heart & circulatory*, *mental & neurodegenerative*, *neuromusculoskeletal*, *nonspecific*, *peripheral vascular & respiratory*, and *respiratory*. The prevalence of all chronic conditions in *nonspecific* was lower than in the general population. On the other hand, *dementia & motility digestive* appeared only in *multimorbidity & frailty*, while *autoimmune & metabolic* appeared only in *multimorbidity & age*.

Some patterns were female-dominant, such as *allergy & migraine*, *neuromusculoskeletal*, *chronic ulcers & peripheral vascular*, *dementia & motility digestive*, and *mental & neurodegenerative*. On the other hand, *autoimmune & metabolic*, *respiratory*, *genitourinary & respiratory*, and *peripheral vascular & respiratory* were male-dominant, with the latter 2 patterns having the highest rates of smokers and high-risk drinkers. *Allergy & migraine*, *diabetes & obesity*, *genitourinary & respiratory*, *neuromusculoskeletal*, and *nonspecific* included younger individuals, while *chronic ulcers & peripheral vascular*, *dementia & motility digestive*, *heart & circulatory*, and *mental & neurodegenerative* were more common in older individuals. These, together with *peripheral vascular & respiratory*, were the patterns with the highest prevalence of frailty. All patterns, except *nonspecific*, had a high prevalence of multimorbidity.

The emergence of the *dementia & motility digestive* pattern in *multimorbidity & frailty* significantly changed the definition of *mental & neurodegenerative* between sets. As shown in Table 3, in *multimorbidity & frailty*, it comprised younger, less frail individuals with more chronic conditions, such as Parkinson disease and other neurological diseases (ie, Huntington disease or myasthenia), rather than dementia (see Multimedia Appendix 4), while in *multimorbidity & age*, this pattern comprised older and frailer individuals. On the other hand, the *dementia & motility digestive* pattern was made up of older individuals, mostly women, with moderate and severe frailty and a high prevalence of dementia.

Table 2. Top 10 conditions in terms of the observed/expected ratio and exclusivity for the heart & circulatory and mental & neurodegenerative patterns.

Pattern and disease	Multimorbidity & frailty		Multimorbidity & age	
	OE[a] ratio	Exclusivity, %	OE ratio	Exclusivity, %
Heart & circulatory				
Heart failure	7.86[b]	54.38[b]	7.67[b]	55.57[b]
Cardiac valve diseases	7.37[b]	50.98[b]	7.24[b]	52.47[b]
Atrial fibrillation	6.54[b]	45.29[b]	6.46[b]	46.81[b]
Bradycardia and conduction diseases	5.75[b]	39.83[b]	5.87[b]	42.57[b]
Ischemic heart disease	3.69[b]	25.51[b]	3.62[b]	26.22[b]
Chronic kidney diseases	2.30[b]	15.95	2.23[b]	16.17
Anemia	2.29[b]	15.85	2.21[b]	16.05
Cerebrovascular disease	2.16[b]	14.93	2.13[b]	15.41
COPD[c], emphysema, and chronic bronchitis	2.07[b]	14.33	1.99	14.43
Inflammatory arthropaties	1.86	12.85	N/A[d]	N/A
Chronic pancreas diseases, and biliary tract and gallbladder diseases	N/A	N/A	1.75	12.67
Mental & neurodegenerative				
Parkinson disease and parkisonism	19.98[b]	90.67[b]	9.24[b]	73.46[b]
Other neurological diseases	18.54[b]	84.14[b]	4.68[b]	37.19[b]
Dementia	2.24[b]	10.18	6.95[b]	55.24[b]
Cerebrovascular disease	2.10[b]	9.53	3.39[b]	26.92[b]
Depression and mood diseases	1.76	8.00	2.20[b]	17.50[b]
Colitis and related diseases	1.48	6.71	1.83	14.59
Anemia	1.38	6.28	2.17[b]	17.26
Sleep disorders	1.37	6.22	N/A	N/A
Other digestive diseases	1.35	6.11	4.57[b]	36.36[b]
Dorsopathies	1.33	6.05	N/A	N/A
Chronic kidney diseases	N/A	N/A	1.65	13.10
Deafness and hearing impairment	N/A	N/A	1.47	11.69

[a] OE: observed/expected.

[b] The conditions used to name the pattern. The conditions for the rest of the patterns can be found in Multimedia Appendix 4.

[c] COPD: chronic obstructive pulmonary disease.

[d] N/A: not applicable.

Table 3. Description of the participants in the dementia & motility digestive, heart & circulatory, and mental & neurodegenerative patterns.

Variable	Dementia & motility digestive[a,b]	Heart & circulatory[a,b]		Mental & neurodegenerative[a,b]	
	Multimorbidity & frailty (n=783,590)	Multimorbidity & age (n=743,257)	Multimorbidity & frailty (n=709,909)	Multimorbidity & age (n=815,597)	Multimorbidity & frailty (n=465,417)
Age (years), median (IQR)	83.0 (78.0-88.0)	82.0 (77.0-87.0)	82.0 (76.0-87.0)	84.0 (80.0-89.0)	78.0 (72.0-84.0)
Sex, n (%)					
Women	588,734 (75.1)	398,282 (53.6)	382,768 (53.9)	541,039 (66.3)	264,225 (56.8)
Men	194,856 (24.9)	344,975 (46.4)	327,141 (46.1)	274,558 (33.7)	201,192 (43.2)
Deprivation index, n (%)					
1 (less deprived)	109,882 (23.3)	93,677 (20.9)	87,347 (20.3)	119,722 (25.5)	66,020 (21.9)
2	95,209 (20.2)	89,449 (20.0)	85,288 (19.8)	97,525 (20.7)	61,623 (20.4)
3	99,621 (21.1)	91,379 (20.4)	87,580 (20.3)	99,451 (21.1)	63,731 (21.1)
4	87,895 (18.6)	89,469 (20.0)	87,256 (20.3)	83,397 (17.7)	59,089 (19.6)
5 (more deprived)	78,990 (16.7)	83,854 (18.7)	82,908 (19.3)	70,225 (14.9)	51,528 (17.1)
Missing	311,993	295,429	279,530	345,577	163,426
Multimorbidity, median (IQR)					
SNAC-K[c] groups of chronic conditions	9.00 (7.0-11.0)	10.0 (8.0-12.0)	10.0 (8.0-2.0)	8.0 (6.0-10.0)	8.0 (6.0-11.0)
Type of multimorbidity, n (%)					
No multimorbidity	0 (0.0)	0 (0.0)	0 (0.0)	445 (0.1)	2025 (0.4)
2-5 diseases	71,676 (9.2)	45,354 (6.1)	32,702 (4.6)	145,376 (17.8)	80,560 (17.3)
6-10 diseases	507,806 (64.8)	395,802 (53.3)	368,236 (51.9)	500,503 (61.4)	254,128 (54.6)
>10 diseases	204,108 (26.0)	302,101 (40.6)	308,971 (43.5)	169,273 (20.8)	128,704 (27.7)
Frailty, median (IQR)					
Deficits	9.0 (7.0-11.0)	9.0 (7.0-11.0)	9.0 (7.0-12.0)	8.0 (6.0-10.0)	7.0 (5.0-10.0)
Disease-related deficits	5.0 (3.0-6.0)	3.0 (2.0-5.0)	4.0 (2.0-5.0)	5.0 (3.0-6.0)	3.0 (2.0-5.0)
SSLD[d] deficits	4.0 (3.0-5.0)	5.0 (4.0-7.0)	5.0 (4.0-7.0)	3.0 (2.0-5.0)	4.0 (2.0-5.0)
Type of frailty, n (%)					
Fit	40,827 (5.21)	49,158 (6.6)	31,624 (4.5)	90,181 (11.1)	96,397 (20.7)
Mild	349,212 (44.6)	295,094 (39.7)	268,474 (37.8)	364,281 (44.7)	207,835 (44.7)
Moderate	312,751 (39.9)	269,356 (36.2)	276,323 (38.9)	279,797 (34.3)	122,853 (26.4)
Severe	80,800 (10.3)	129,649 (17.4)	133,488 (18.8)	81,338 (10.0)	38,332 (8.2)
Smoking status, n (%)					
Nonsmoker	541,491 (71.8)	454,844 (62.4)	430,558 (61.7)	544,408 (70.5)	299,351 (66.8)
Exsmoker	182,092 (24.1)	246,506 (33.8)	240,234 (34.4)	196,319 (25.4)	120,789 (27.0)
Smoker	30,528 (4.1)	27,744 (3.8)	27,585 (4.0)	31,545 (4.1)	28,002 (6.2)
Missing	29,480	14,163	11,532	43,325	17,275
Alcohol intake, n (%)					
Nondrinker	315,999 (83.8)	342,552 (75.7)	333,548 (75.6)	297,365 (83.6)	178,703 (76.0)
Low-risk drinker	59,668 (15.8)	107,637 (23.8)	104,929 (23.8)	56,921 (16.0)	54,962 (23.4)
High-risk drinker	1,596 (0.4)	2,475 (0.6)	2,600 (0.6)	1,218 (0.3)	1,434 (0.6)
Missing	406,327	290,593	268,832	460,093	230,318
Health care service use, median (IQR)					
Visits to primary care	14.0 (8.0-23.0)	20.0 (10.0-33.0)	21.0 (11.0-34.0)	12.0 (6.0-21.0)	13.0 (7.0-21.0)

Variable	Dementia & motility digestive[a,b]	Heart & circulatory[a,b]		Mental & neurodegenerative[a,b]	
	Multimorbidity & frailty (n=783,590)	Multimorbidity & age (n=743,257)	Multimorbidity & frailty (n=709,909)	Multimorbidity & age (n=815,597)	Multimorbidity & frailty (n=465,417)
Distinct drugs[e]	10.0 (7.0-14.0)	11.0 (8.0-15.0)	12.0 (8.0-16.0)	9.0 (6.0-13.0)	10.0 (6.0-14.0)
Clinical measures	7.00 (2.0-14.0)	11.0 (5.0-19.0)	12.0 (5.0-20.0)	6.0 (1.0-12.0)	7.0 (2.0-14.0)
Laboratory results	17.0 (10.0-30.0)	18.0 (10.0-32.0)	18.0 (11.0-33.0)	17.0 (0.0-24.0)	17.0 (3.0-25.0)

[a] Each column reports the information of all individuals and years included in each pattern. The description of the rest of the patterns can be found in Multimedia Appendix 5.

[b] For categorical variables, missing values, if any, are excluded from the calculation of the percentage.

[c] SNAC-K: Swedish National Study of Aging and Care in Kungsholmen.

[d] SSLD: signs, symptoms, laboratory results, and disabilities.

[e] The number of distinct drugs was calculated using the first 5 digits of the Anatomical Therapeutic Classification (ATC) code.

Effect of Considering Frailty

Compared to the *multimorbidity & age* patterns, the *multimorbidity & frailty* grouping assigned more frail individuals to the patterns defined by chronic conditions imposing greater limitations on daily life. For example, severe frailty was more prevalent in the *heart & circulatory* pattern and a lack of frailty was more common in the *nonspecific* pattern in *multimorbidity & frailty* than in *multimorbidity & age* (Table 3). In addition, the *dementia & motility digestive* pattern appeared only in *multimorbidity & frailty*, and the definition of *mental & neurodegenerative* changed considerably, as described above.

Each set of patterns behaved differently in terms of the associated outcomes (Table 4). *Multimorbidity & age* patterns had a better goodness of fit (AIC) with mortality than *multimorbidity & frailty*, while *multimorbidity & frailty* patterns had a better or similar goodness of fit with nursing home admission or home care need. A similar behavior was observed for R^2, while both sets of patterns achieved similar results in all outcomes in terms of AUC. Regarding the hazard ratios, they were very similar between *multimorbidity & age* and *multimorbidity & frailty* in all patterns and outcomes, except *mental & neurodegenerative*, which were lower in *multimorbidity & frailty*. *Chronic ulcers & peripheral vascular*, *dementia & motility digestive*, *heart & circulatory*, *mental & neurodegenerative*, and *peripheral vascular & respiratory* showed the highest risk of death. All patterns had a higher risk of nursing home admission than *nonspecific*, with *dementia & motility digestive*, *chronic ulcers & peripheral vascular*, and *mental & neurodegenerative* standing out. These latter patterns, together with *heart & circulatory*, also had a higher risk of home care need.

Table 4. Data for the unadjusted survival models using the multimorbidity patterns as the time-varying covariate.

Pattern[a,b]	Mortality		Nursing home admission		Home care need	
	Multimorbidity & age, HR[c] (95% CI)	Multimorbidity & frailty, HR (95% CI)	Multimorbidity & age, HR (95% CI)	Multimorbidity & frailty, HR (95% CI)	Multimorbidity & age, HR (95% CI)	Multimorbidity & frailty, HR (95% CI)
Allergy & migraine	0.39 (0.39-0.70)	0.43 (0.42-0.44)	1.68 (1.61-1.76)	1.53 (1.45-1.61)	2.50 (2.41-2.59)	2.98 (2.87-3.10)
Autoimmune & metabolic	0.97 (0.96-0.99)	N/A[d]	2.51 (2.40-2.62)	N/A	4.66 (4.52-4.81)	N/A
Chronic ulcers & peripheral vascular	6.94 (6.85-7.03)	6.87 (6.78-6.96)	22.28 (21.47-23.12)	23.43 (22.58-24.31)	26.25 (25.44-27.08)	34.36 (33.24-35.52)
Dementia & motility digestive	N/A	2.88 (2.86-2.92)	N/A	22.95 (22.25-23.67)	N/A	25.38 (24.67-26.12)
Diabetes & obesity	0.7 (0.69-0.71)	0.62 (0.61-0.63)	1.62 (1.56-1.69)	1.42 (1.37-1.48)	0.68 (0.67-0.69)	3.23 (3.13-3.34)
Genitourinary & respiratory	0.8 (0.79-0.81)	1.05 (1.04-1.07)	1.22 (1.16-1.28)	1.48 (1.42-1.55)	1.69 (1.63-1.75)	2.97 (2.86-3.07)
Heart & circulatory	3.14 (3.10-3.18)	3.07 (3.03-3.10)	7.71 (7.45-7.98)	8.13 (7.86-8.42)	15.64 (15.23-16.07)	21.59 (20.97-22.22)
Mental & neurodegenerative	3.56 (3.52-3.60)	1.96 (1.93-1.99)	24.41 (23.66-25.18)	10.91 (10.53-11.30)	20.46 (19.93-21.00)	16.98 (16.46-17.51)
Neuromusculoskeletal	0.32 (0.31-0.33)	0.31 (0.31-0.32)	1.61 (1.54-1.68)	1.43 (1.36-1.49)	3.36 (3.25-3.46)	3.83 (3.70-3.96)
Peripheral vascular & respiratory	2.04 (2.01-2.07)	1.98 (1.95-2.01)	3.99 (3.82-4.17)	4.02 (3.85-4.20)	7.53 (7.30-7.77)	9.54 (9.23-9.87)
Respiratory	0.9 (0.89-0.92)	0.87 (0.86-0.89)	3.11 (2.98-3.26)	2.98 (2.84-3.12)	5.78 (5.60-5.97)	7.10 (6.86-7.35)
AIC[e]	9,605,078[f]	9,637,872	10,261,814	9,996,581[f]	10,945,802	10,924,083[f]
R^2	0.37[f]	0.32	0.21	0.38[f]	0.36	0.39[f]
AUC[g]	0.713[f]	0.705	0.718[f]	0.717	0.719	0.721[f]

[a] Nonspecific was used as reference.

[b] All *P*<.001.

[c] HR: hazard ratio.

[d] N/A: not applicable.

[e] AIC: Akaike Information Criterion.

[f] The set of patterns that achieved better performance for each metric and outcome.

[g] AUC: area under the receiver operating characteristic curve.

Study of the Trajectories

Over the follow-up, individuals changed patterns an average of 1.75 (*multimorbidity & age*) and 1.85 (*multimorbidity & frailty*) times, while 45.1% (656,778/1,456,052) of the individuals remained in the same pattern. Multimedia Appendix 6 shows that these values depended on the length of the trajectory. For example, people with 5 years of follow-up had an average of 1.55 different patterns in their trajectory, while around 54.5% (52,634/96,578, 54.5% in *multimorbidity & age*; 52,586/96,578, 54.4% in *multimorbidity & frailty*) remained in the same pattern. In those with 10 years of follow-up, these values changed to 2.15 and 26.6%, respectively. The prevalence of most patterns varied with age, as shown in Figure 3. The prevalence of *nonspecific* showed the largest reduction, as most of its members developed diseases as they aged and shifted to more disease-specific patterns. On the other hand, *heart & circulatory* and *mental & neurodegenerative* for *multimorbidity & age*, and *dementia & motility digestive* for *multimorbidity & frailty* showed the highest increases in prevalence with age.

Figure 4 shows the transition matrices for both *multimorbidity & age* and *multimorbidity & frailty*. *Chronic ulcers & peripheral vascular*, *heart & circulatory*, *mental & neurodegenerative*, and *dementia & motility digestive* were the multimorbidity patterns most closely associated with mortality, as more than 55% of the individuals who started in them died during follow-up (Figure 4). On the other hand, the *allergy & migraine*, *neuromusculoskeletal*, and *nonspecific* clusters showed the lowest mortality. When considering frailty, the number of individuals transitioning to *mental & neurodegenerative* from any pattern was reduced. Nevertheless, considering the general stability of the trajectories, a relatively high percentage of individuals transitioned to *dementia & motility digestive*. Multimedia Appendix 7 (*multimorbidity & age*) and Multimedia Appendix 8 (*multimorbidity & frailty*) show the evolution of 3 subsets of 50 random individuals aged 65, 75, and 85 years, respectively, in 2010 to illustrate how multimorbidity trajectories vary with age. These figures show that the percentage of individuals starting in *nonspecific* decreased inversely with age at cohort entry, as did the percentage of individuals remaining in this pattern throughout follow-up.

Figure 3. Prevalence of each multimorbidity pattern for each age. (A) Multimorbidity & age; (B) multimorbidity & frailty. For each age, the information considered is from the individuals of that age in any time, regardless of the year of the study.

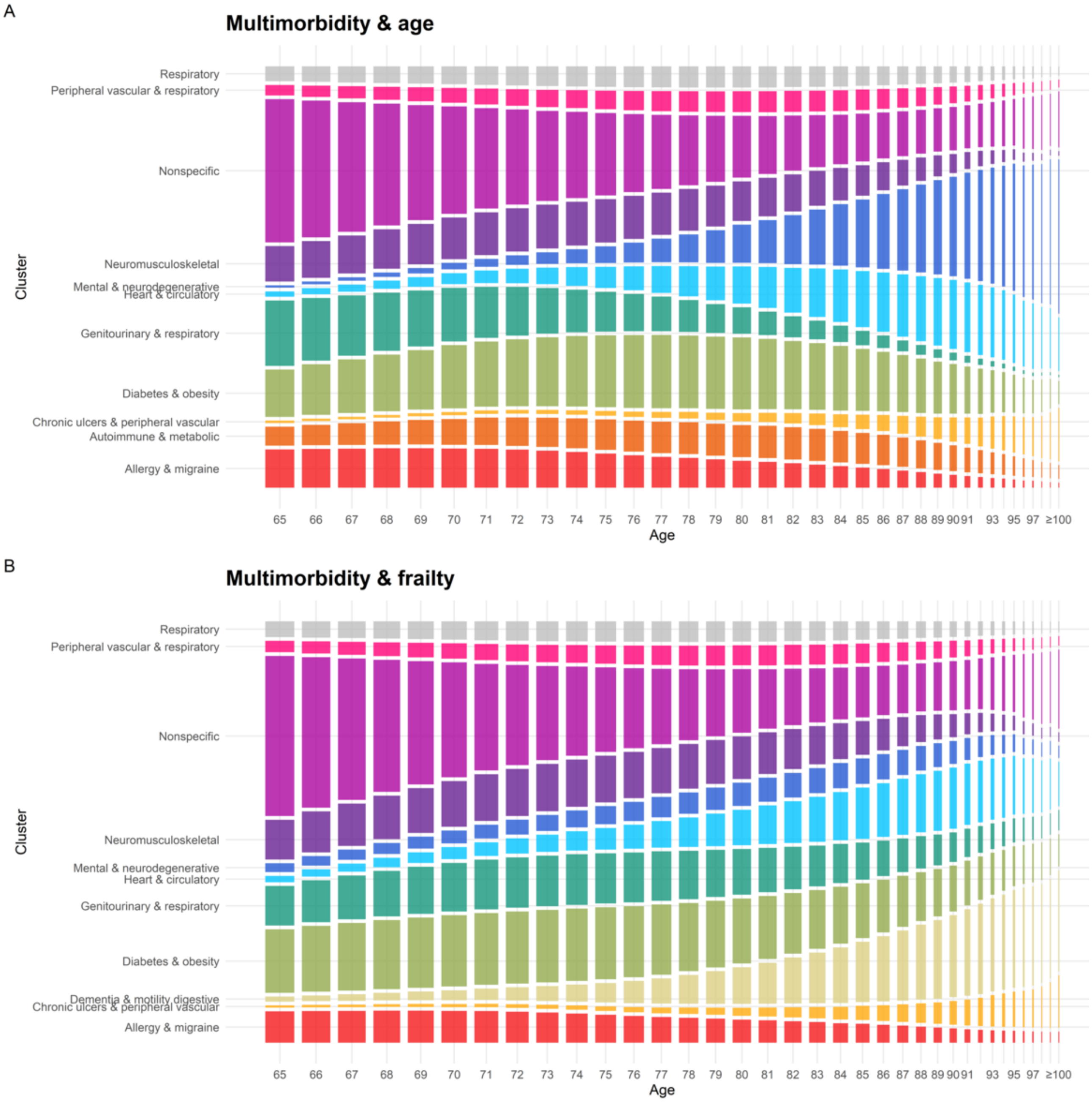
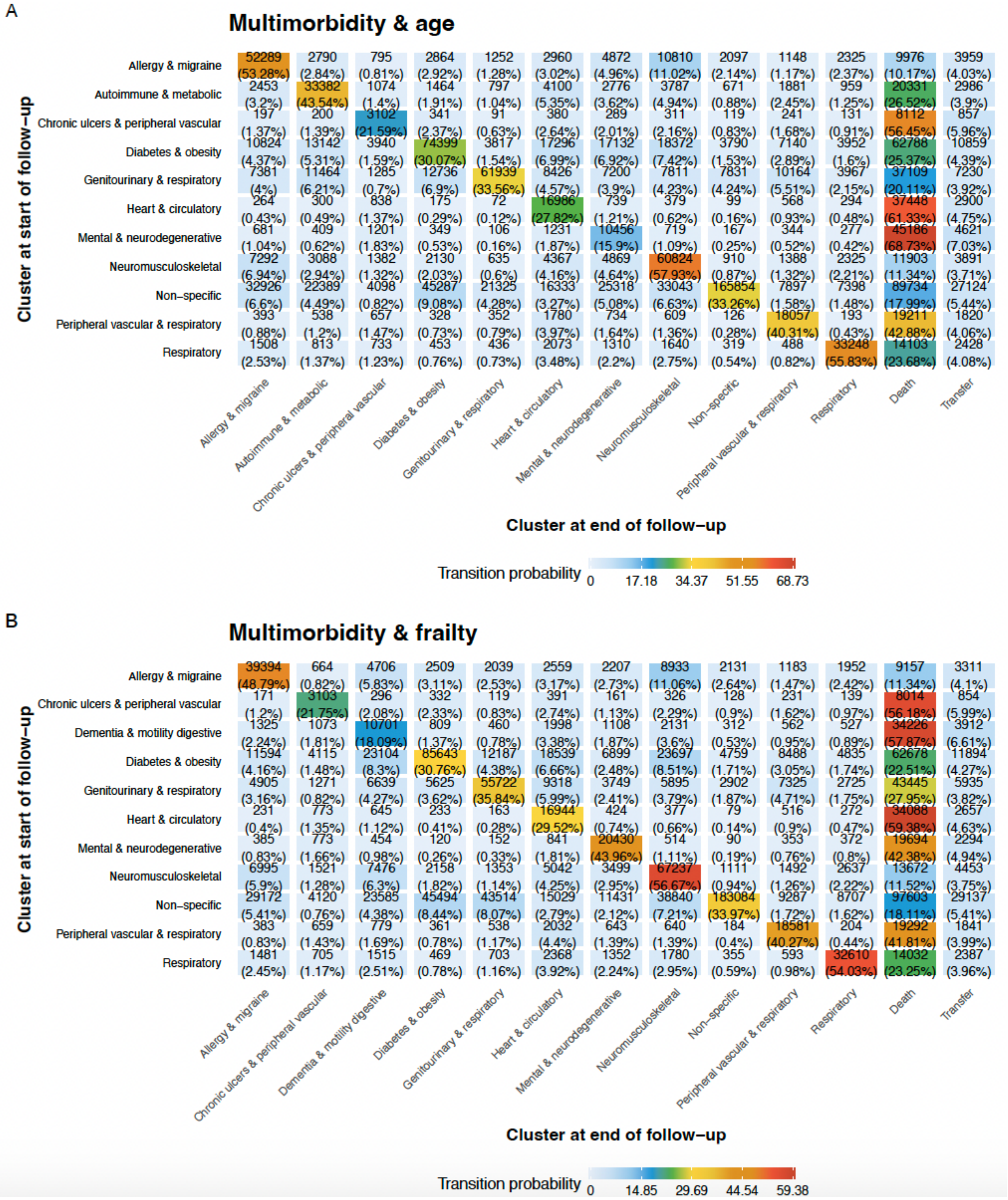


Figure 4. Transition matrices for multimorbidity & age (A) and multimorbidity & frailty (B) with k=11. Each cell shows the proportion of individuals transitioning from their initial pattern (y-axis) to the last pattern observed (x-axis).



	Allergy & migraine	Autoimmune & metabolic	Chronic ulcers & peripheral vascular	Diabetes & obesity	Genitourinary & respiratory	Heart & circulatory	Mental & neurodegenerative	Neuromusculoskeletal	Non–specific	Peripheral vascular & respiratory	Respiratory	Death	Transfer
Allergy & migraine	52289 (53.28%)	2790 (2.84%)	795 (0.81%)	2864 (2.92%)	1252 (1.28%)	2960 (3.02%)	4872 (4.96%)	10810 (11.02%)	2097 (2.14%)	1148 (1.17%)	2325 (2.37%)	9976 (10.17%)	3959 (4.03%)
Autoimmune & metabolic	2453 (3.2%)	33382 (43.54%)	1074 (1.4%)	1464 (1.91%)	797 (1.04%)	4100 (5.35%)	2776 (3.62%)	3787 (4.94%)	671 (0.88%)	1881 (2.45%)	959 (1.25%)	20331 (26.52%)	2986 (3.9%)
Chronic ulcers & peripheral vascular	197 (1.37%)	200 (1.39%)	3102 (21.59%)	341 (2.37%)	91 (0.63%)	380 (2.64%)	289 (2.01%)	311 (2.16%)	119 (0.83%)	241 (1.68%)	131 (0.91%)	8112 (56.45%)	857 (5.96%)
Diabetes & obesity	10824 (4.37%)	13142 (5.31%)	3940 (1.59%)	74399 (30.07%)	3817 (1.54%)	17296 (6.99%)	17132 (6.92%)	18372 (7.42%)	3790 (1.53%)	7140 (2.89%)	3952 (1.6%)	62788 (25.37%)	10859 (4.39%)
Genitourinary & respiratory	7381 (4%)	11464 (6.21%)	1285 (0.7%)	12736 (6.9%)	61939 (33.56%)	8426 (4.57%)	7200 (3.9%)	7811 (4.23%)	7831 (4.24%)	10164 (5.51%)	3967 (2.15%)	37109 (20.11%)	7230 (3.92%)
Heart & circulatory	264 (0.43%)	300 (0.49%)	838 (1.37%)	175 (0.29%)	72 (0.12%)	16986 (27.82%)	739 (1.21%)	379 (0.62%)	99 (0.16%)	568 (0.93%)	294 (0.48%)	37448 (61.33%)	2900 (4.75%)
Mental & neurodegenerative	681 (1.04%)	409 (0.62%)	1201 (1.83%)	349 (0.53%)	106 (0.16%)	1231 (1.87%)	10456 (15.9%)	719 (1.09%)	167 (0.25%)	344 (0.52%)	277 (0.42%)	45186 (68.73%)	4621 (7.03%)
Neuromusculoskeletal	7292 (6.94%)	3088 (2.94%)	1382 (1.32%)	2130 (2.03%)	635 (0.6%)	4367 (4.16%)	4869 (4.64%)	60824 (57.93%)	910 (0.87%)	1388 (1.32%)	2325 (2.21%)	11903 (11.34%)	3891 (3.71%)
Non–specific	32926 (6.6%)	22389 (4.49%)	4098 (0.82%)	45287 (9.08%)	21325 (4.28%)	16333 (3.27%)	25318 (5.08%)	33043 (6.63%)	165854 (33.26%)	7897 (1.58%)	7398 (1.48%)	89734 (17.99%)	27124 (5.44%)
Peripheral vascular & respiratory	393 (0.88%)	538 (1.2%)	657 (1.47%)	328 (0.73%)	352 (0.79%)	1780 (3.97%)	734 (1.64%)	609 (1.36%)	126 (0.28%)	18057 (40.31%)	193 (0.43%)	19211 (42.88%)	1820 (4.06%)
Respiratory	1508 (2.53%)	813 (1.37%)	733 (1.23%)	453 (0.76%)	436 (0.73%)	2073 (3.48%)	1310 (2.2%)	1640 (2.75%)	319 (0.54%)	488 (0.82%)	33248 (55.83%)	14103 (23.68%)	2428 (4.08%)

	Allergy & migraine	Chronic ulcers & peripheral vascular	Dementia & motility digestive	Diabetes & obesity	Genitourinary & respiratory	Heart & circulatory	Mental & neurodegenerative	Neuromusculoskeletal	Non–specific	Peripheral vascular & respiratory	Respiratory	Death	Transfer
Allergy & migraine	39394 (48.79%)	664 (0.82%)	4706 (5.83%)	2509 (3.11%)	2039 (2.53%)	2559 (3.17%)	2207 (2.73%)	8933 (11.06%)	2131 (2.64%)	1183 (1.47%)	1952 (2.42%)	9157 (11.34%)	3311 (4.1%)
Chronic ulcers & peripheral vascular	171 (1.2%)	3103 (21.75%)	296 (2.08%)	332 (2.33%)	119 (0.83%)	391 (2.74%)	161 (1.13%)	326 (2.29%)	128 (0.9%)	231 (1.62%)	139 (0.97%)	8014 (56.18%)	854 (5.99%)
Dementia & motility digestive	1325 (2.24%)	1073 (1.81%)	10701 (18.09%)	809 (1.37%)	460 (0.78%)	1998 (3.38%)	1108 (1.87%)	2131 (3.6%)	312 (0.53%)	562 (0.95%)	527 (0.89%)	34226 (57.87%)	3912 (6.61%)
Diabetes & obesity	11594 (4.16%)	4115 (1.48%)	23104 (8.3%)	85643 (30.76%)	12187 (4.38%)	18539 (6.66%)	6899 (2.48%)	23697 (8.51%)	4759 (1.71%)	8488 (3.05%)	4835 (1.74%)	62678 (22.51%)	11894 (4.27%)
Genitourinary & respiratory	4905 (3.16%)	1271 (0.82%)	6639 (4.27%)	5625 (3.62%)	55722 (35.84%)	9318 (5.99%)	3749 (2.41%)	5895 (3.79%)	2902 (1.87%)	7325 (4.71%)	2725 (1.75%)	43445 (27.95%)	5935 (3.82%)
Heart & circulatory	231 (0.4%)	773 (1.35%)	645 (1.12%)	233 (0.41%)	163 (0.28%)	16944 (29.52%)	424 (0.74%)	377 (0.66%)	79 (0.14%)	516 (0.9%)	272 (0.47%)	34088 (59.38%)	2657 (4.63%)
Mental & neurodegenerative	385 (0.83%)	773 (1.66%)	454 (0.98%)	120 (0.26%)	152 (0.33%)	841 (1.81%)	20430 (43.96%)	514 (1.11%)	90 (0.19%)	353 (0.76%)	372 (0.8%)	19694 (42.38%)	2294 (4.94%)
Neuromusculoskeletal	6995 (5.9%)	1521 (1.28%)	7476 (6.3%)	2158 (1.82%)	1353 (1.14%)	5042 (4.25%)	3499 (2.95%)	67237 (56.67%)	1111 (0.94%)	1492 (1.26%)	2637 (2.22%)	13672 (11.52%)	4453 (3.75%)
Non–specific	29172 (5.41%)	4120 (0.76%)	23585 (4.38%)	45494 (8.44%)	43514 (8.07%)	15029 (2.79%)	11431 (2.12%)	38840 (7.21%)	183084 (33.97%)	9287 (1.72%)	8707 (1.62%)	97603 (18.11%)	29137 (5.41%)
Peripheral vascular & respiratory	383 (0.83%)	659 (1.43%)	779 (1.69%)	361 (0.78%)	538 (1.17%)	2032 (4.4%)	643 (1.39%)	640 (1.39%)	184 (0.4%)	18581 (40.27%)	204 (0.44%)	19292 (41.81%)	1841 (3.99%)
Respiratory	1481 (2.45%)	705 (1.17%)	1515 (2.51%)	469 (0.78%)	703 (1.16%)	2368 (3.92%)	1352 (2.24%)	1780 (2.95%)	355 (0.59%)	593 (0.98%)	32610 (54.03%)	14032 (23.25%)	2387 (3.96%)

Discussion

This study aimed to assess how frailty contributes to the characterization of multimorbidity patterns, as identified through clustering techniques. In a Mediterranean cohort of 1,456,052 people aged over 65 years, 2 sets of 11 multimorbidity patterns were identified based on the presence of chronic conditions. One considered age and the other considered the number of frailty deficits. The consideration of frailty modified multimorbidity patterns, revealing patterns with better goodness of fit to the outcomes related to frail aging, such as nursing home admission and home care need, and gathering more frail individuals in patters characterized by more limiting conditions. The better fit to aging-related outcomes when considering frailty has been previously reported [48]. Moreover, the trajectories of multimorbidity patterns were different when considering frailty. When considering only conditions and age, and not frailty, the patterns showed better goodness of fit with the outcome of death, and described an additional disease-related pattern.

When considering frailty, an additional pattern more specific to aging was identified (*dementia & motility digestive*). Distinguishing this pattern from *mental & neurodegenerative* can enable more personalized clinical treatment. In the former case, palliative care would be appropriate, while the latter would call for active treatments to delay disease progression. Patterns defined by less limiting conditions for daily life, such as *diabetes & obesity*, had a lower prevalence of frailty when frailty was considered, in contrast with more limiting patterns such as *heart & circulatory*. For this same reason, the *genitourinary & respiratory* pattern also had a higher proportion of men when frailty was considered, as it included more individuals with prostate diseases, which usually lead to more frailty [49,50], while the same pattern included more women when frailty was not considered, as it involved addictions (ie, drug, alcohol, or tobacco use), which usually involve less frailty than prostate diseases.

Most of the 11 multimorbidity patterns were similarly described whether frailty was considered or not and could be classified into concordant or discordant multimorbidity patterns depending on whether the conditions defining the pattern shared pathophysiology or approaches to clinical management [51,52]. For example, *heart & circulatory*, *chronic ulcers & peripheral vascular*, *respiratory*, and *diabetes & obesity* are patterns of concordant multimorbidity, while *genitourinary & respiratory* and *peripheral vascular & respiratory* are patterns of discordant multimorbidity. Discordant conditions might be grouped together because of shared risk factors, such as smoking or alcohol intake, rather than shared pathophysiology. This is the case for *genitourinary & respiratory* and *peripheral vascular & respiratory*, which mainly included men, with a higher prevalence of smokers and high-risk drinkers. Another risk factor these individuals may share could be their genetics, which can also influence the development of multimorbidity [53]. This discordance makes treatment more complex [52]; thus, identifying patients who follow a discordant multimorbidity pattern can signal a need for integrated care.

Regarding the trajectories of multimorbidity patterns, the *heart & circulatory* and *mental & neurodegenerative* patterns (when not considering frailty), and the *dementia & digestive motility* pattern (when considering frailty) showed the most changes over time. Their prevalence increased the most with aging, and the patients in these groups had the highest probability of transitioning to death at the end of their follow-up. This may be due to the high prevalence of frailty in these patterns, and the association of frailty with death [54]. The *nonspecific* pattern also had a high probability of transition, as it included healthier individuals who transitioned to other patterns as diseases appeared with age. *Peripheral vascular & respiratory* involved a burden of heart conditions, which could lead the trajectory toward *heart & circulatory*. However, many of these individuals died, possibly before the onset of the cardiac diseases could be recorded in the EHRs. On the other hand, *neuromusculoskeletal* and *respiratory* were the patterns with the fewest transitions during follow-up, with more than 50% of the individuals remaining in the same pattern throughout the study period. This may be because both patterns tended to evolve toward disability rather than death or the development of other comorbidities, and nowadays, treatments are more effective in maintaining the status of patients. In our study, trajectories included a mean of 1.8 different multimorbidity patterns. Even people who had 10 years of follow-up did not make 3 changes on average; thus, the trajectories can be considered quite stable. This result is similar to that found in previous research in this same population, albeit with slightly different methods, where 59% of individuals did not change their pattern [23] over 5 years of follow-up (54.5% in this study).

In terms of patient-based multimorbidity patterns reported in the literature, the substantial variability [55] could be attributed to differences in the populations or to the lack of consensus on how many and which diseases should be considered in determining multimorbidity [9]. However, the most commonly reported patterns from multimorbidity data include cardiac, cognitive, psychiatric, musculoskeletal, respiratory, and genitourinary system diseases [27,28], and these patterns have also been found in this work. Few studies have described multimorbidity trajectories [19], and none considered frailty in their definition; thus, our study is pioneering in this line of research. Only 1 scoping review on multimorbidity trajectories compiled evidence from 34 studies, finding significant associations between multimorbidity and adverse outcomes [56]. However, the heterogeneity of the described methods and the long-term conditions considered in each study preclude a robust comparison.

This study has strengths and limitations. We used a large high-quality database [57] along with standardized and validated tools to identify both multimorbidity [43] and frailty [20,21]. EHRs are a representation of real-world data and may, despite cleaning, contain mistakes inherent to daily clinical practice. This could represent an information bias, but SIDIAP has implemented several standardized quality protocols to avoid it [40]. In addition, we considered the absence of any condition or frailty information in the EHRs to indicate an absence of that condition or frailty status in the individual. Therefore, some chronic conditions or frailty deficits could be underreported among patients who visit primary care centers less frequently, constituting an information bias. Consequently, only data available in the EHRs were considered, and to avoid the creation of unreal records, no missing values were imputed. The SIDIAP database is representative of the population of Catalonia [40], so its use does not imply a selection bias. We excluded people who did not visit primary care during the entire study period in order to eliminate those with private health insurance; however, we cannot rule out that some were healthy individuals. After the first filtering, a criterion based on the predominance of diseases in the identified patterns had to be defined and used to make the final selection. The inclusion of all potential diagnoses would have entailed greater complexity, which would have hindered both the interpretation of findings and the comparison with other studies. Other studies have proceeded in the same way [22,23]. Clustering is an unsupervised exploratory technique whose results depend on the population. Therefore, different patterns could be identified from a database in another region. However, the variables used to generate the patterns were obtained using electronic tools based on routine EHR data, such as SNAC-K and eFRAGICAP, so their acquisition is reproducible in information systems in other countries. This increases the international applicability of this study, helping to establish multimorbidity management worldwide. In addition, clustering techniques can suffer from dependency on random initialization, and there is no guarantee of optimal clustering. To minimize this disadvantage, 100 repetitions with different seeds were performed when optimizing the choice of the number of clusters. Particularly in fuzzy clustering techniques, the membership probability in the heuristic global cost function depends on the number of clusters, and specifying a wrong number of clusters may affect the clustering solution [58]. However, we have validated the number of clusters both analytically and clinically.

Care for older people requires holistic patient-centered care plans that are effectively coordinated and minimally disruptive, considering the social and family context in which health care activities are managed, decisions are made, and care is experienced. As a future line of work, these multimorbidity patterns could be used as adjustment covariates in prediction models for outcomes, such as those reported here, or others, such as emergency admission. Similarly, artificial intelligence–based models that predict the timing and direction of transitions between patterns can be developed. These models could help to improve and anticipate decision-making regarding end-of-life management. More work can be done on the study of trajectories, such as the development of care and treatment guidelines that simultaneously consider the current individual’s multimorbidity pattern and the trajectory over time. Sequence analysis, which combines longitudinal analysis and clustering, could also be used to identify trajectories [59]. The relationship between genetics and concordant and discordant multimorbidity patterns could also be studied.

This study took a person-centered approach, offering relevant information about the multimorbidity patterns and trajectories in the aging population based on age, frailty, and other health determinants. Multimorbidity and frailty can define aging, so both characteristics are relevant considerations made when designing and developing tools, such as multimorbidity patterns, to characterize the aging population. When considering the frailty of individuals, the estimation of outcomes, such as nursing home admission and home care need, improved, as did the characterization of the patterns themselves in terms of how limiting their main chronic conditions can be. The consideration of both multimorbidity and frailty can help to improve treatment guidelines, social assistance, and decision-making in primary care. If most patients in a primary care center follow patterns associated with frailty, increased spending on home-based care services and integrated care programs may be warranted, while a higher burden of multimorbidity would imply higher spending on physician visits. Therefore, and echoing other studies, we recommend that future research involving older populations consider frailty [60].

Acknowledgments

The project received a research grant from the Carlos III Institute of Health, Ministry of Economy and Competitiveness (Spain), awarded in 2019 under the Health Strategy Action 2013-2016, within the National Research Programme oriented to Societal Challenges, within the Technical, Scientific and Research National Plan 2013-2016 (reference PI19/00535), and the PFIS Grant FI20/00040, co-funded with European Union ERDF (European Regional Development Fund) funds. The funder had no role in the study design, data collection, data analysis, data interpretation, or writing of this work. The authors thank Meggan Harris for her role in editing and proofreading the manuscript.

Data Availability

The data used in this study are only available to the participating researchers, in accordance with current European and national laws. Thus, distribution of the data is not allowed. However, researchers from public institutions can request data from SIDIAP. Further information is available online [61].

Authors' Contributions

CV, MCB, ARL, and LACR participated in the study conceptualization, investigation, funding acquisition, and methodology definition. LACR and ARL performed the data curation, formal analysis, and visualization. LACR, MDC, and CV drafted the manuscript, and all authors participated in the review and editing of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
RECORD (Reporting of Studies Conducted using Observational Routinely-collected Data) statement.
[PDF File (Adobe PDF File), 147 KB - publichealth_v9i1e45848_app1.pdf]

Multimedia Appendix 2

Supplementary methods.
[PDF File (Adobe PDF File), 787 KB - publichealth_v9i1e45848_app2.pdf]

Multimedia Appendix 3
Longitudinal flowchart of included persons during the study period (year 2010-2019).
[DOCX File , 99 KB - publichealth_v9i1e45848_app3.docx]

Multimedia Appendix 4
The top 10 conditions in terms of the observed/expected ratio and exclusivity for each pattern for both multimorbidity & age and multimorbidity & frailty.
[XLSX File (Microsoft Excel File), 22 KB - publichealth_v9i1e45848_app4.xlsx]

Multimedia Appendix 5
Description of the participants gathered in each pattern.
[XLSX File (Microsoft Excel File), 21 KB - publichealth_v9i1e45848_app5.xlsx]

Multimedia Appendix 6
Statistics of the trajectories according to follow-up time.
[DOCX File , 85 KB - publichealth_v9i1e45848_app6.docx]

Multimedia Appendix 7
Trajectories of 50 random participants aged 65, 75, or 85 years in 2010 (baseline year) according to multimorbidity & age patterns.
[DOCX File , 761 KB - publichealth_v9i1e45848_app7.docx]

Multimedia Appendix 8
Trajectories of 50 random participants aged 65, 75, or 85 years in 2010 (baseline year) according to multimorbidity & frailty patterns.
[DOCX File , 747 KB - publichealth_v9i1e45848_app8.docx]

References

1. Barnett K, Mercer S, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 2012 Jul;380(9836):37-43 [FREE Full text] [doi: 10.1016/s0140-6736(12)60240-2]
2. Clegg A, Young J, Iliffe S, Rikkert M, Rockwood K. Frailty in elderly people. The Lancet 2013 Mar;381(9868):752-762 [FREE Full text] [doi: 10.1016/s0140-6736(12)62167-9]
3. Rockwood K, Mitnitski A, Song X, Steen B, Skoog I. Long-term risks of death and institutionalization of elderly people in relation to deficit accumulation at age 70. J Am Geriatr Soc 2006 Jun;54(6):975-979 [FREE Full text] [doi: 10.1111/j.1532-5415.2006.00738.x] [Medline: 16776795]
4. Villacampa-Fernández P, Navarro-Pardo E, Tarín JJ, Cano A. Frailty and multimorbidity: Two related yet different concepts. Maturitas 2017 Jan;95:31-35 [FREE Full text] [doi: 10.1016/j.maturitas.2016.10.008] [Medline: 27889050]
5. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar 01;56(3):M146-M156 [FREE Full text] [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
6. Mitnitski A, Mogilner A, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal 2001 Aug 08;1:323-336 [FREE Full text] [doi: 10.1100/tsw.2001.58] [Medline: 12806071]
7. Rodríguez-Mañas L, Féart C, Mann G, Viña J, Chatterji S, Chodzko-Zajko W, FOD-CC group (Appendix 1). Searching for an operational definition of frailty: a Delphi method based consensus statement: the frailty operative definition-consensus conference project. J Gerontol A Biol Sci Med Sci 2013 Jan;68(1):62-67 [FREE Full text] [doi: 10.1093/gerona/gls119] [Medline: 22511289]
8. Multimorbidity: a priority for global health research. The Academy of Medical Sciences. 2018. URL: https://acmedsci.ac.uk/file-download/82222577 [accessed 2023-06-02]
9. Vetrano D, Palmer K, Marengoni A, Marzetti E, Lattanzio F, Roller-Wirnsberger R, Joint Action ADVANTAGE WP4 Group. Frailty and Multimorbidity: A Systematic Review and Meta-analysis. J Gerontol A Biol Sci Med Sci 2019 Apr 23;74(5):659-666 [FREE Full text] [doi: 10.1093/gerona/gly110] [Medline: 29726918]
10. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev 2011 Sep;10(4):430-439. [doi: 10.1016/j.arr.2011.03.003] [Medline: 21402176]

11. Librero J, Peiró S, Ordiñana R. Chronic comorbidity and outcomes of hospital care: length of stay, mortality, and readmission at 30 and 365 days. J Clin Epidemiol 1999 Mar;52(3):171-179. [doi: 10.1016/s0895-4356(98)00160-7] [Medline: 10210233]
12. Roe L, Normand C, Wren M, Browne J, O'Halloran AM. The impact of frailty on healthcare utilisation in Ireland: evidence from the Irish longitudinal study on ageing. BMC Geriatr 2017 Sep 05;17(1):203 [FREE Full text] [doi: 10.1186/s12877-017-0579-0] [Medline: 28874140]
13. Han L, Clegg A, Doran T, Fraser L. The impact of frailty on healthcare resource use: a longitudinal analysis using the Clinical Practice Research Datalink in England. Age Ageing 2019 Sep 01;48(5):665-671 [FREE Full text] [doi: 10.1093/ageing/afz088] [Medline: 31297511]
14. García-Nogueras I, Aranda-Reneo I, Peña-Longobardo LM, Oliva-Moreno J, Abizanda P. Use of Health Resources and Healthcare Costs associated with Frailty: The FRADEA Study. J Nutr Health Aging 2017;21(2):207-214. [doi: 10.1007/s12603-016-0727-9] [Medline: 28112778]
15. Zhang J, Symons J, Agapow P, Teo J, Paxton C, Abdi J, et al. Best practices in the real-world data life cycle. PLOS Digit Health 2022 Jan;1(1):e0000003 [FREE Full text] [doi: 10.1371/journal.pdig.0000003] [Medline: 36812509]
16. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and Impact of Real-World Clinical Data for the Practicing Clinician. Adv Ther 2018 Nov 24;35(11):1763-1774 [FREE Full text] [doi: 10.1007/s12325-018-0805-y] [Medline: 30357570]
17. Ho Y, Hu F, Lee P. The Advantages and Challenges of Using Real-World Data for Patient Care. Clin Transl Sci 2020 Jan 27;13(1):4-7 [FREE Full text] [doi: 10.1111/cts.12683] [Medline: 31456349]
18. Wang S, Schneeweiss S, RCT-DUPLICATE Initiative, Franklin JM, Desai RJ, Feldman W, et al. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA 2023 Apr 25;329(16):1376-1385 [FREE Full text] [doi: 10.1001/jama.2023.4221] [Medline: 37097356]
19. Ho I, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. The Lancet Public Health 2021 Aug;6(8):e587-e597 [FREE Full text] [doi: 10.1016/s2468-2667(21)00107-9]
20. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age Ageing 2016 May;45(3):353-360 [FREE Full text] [doi: 10.1093/ageing/afw039] [Medline: 26944937]
21. Orfila F, Carrasco-Ribelles L, Abellana R, Roso-Llorach A, Cegri F, Reyes C, et al. Validation of an electronic frailty index with electronic health records: eFRAGICAP index. BMC Geriatr 2022 May 07;22(1):404 [FREE Full text] [doi: 10.1186/s12877-022-03090-8] [Medline: 35525922]
22. Violán C, Foguet-Boreu Q, Fernández-Bertolín S, Guisado-Clavero M, Cabrera-Bean M, Formiga F, et al. Soft clustering using real-world data for the identification of multimorbidity patterns in an elderly population: cross-sectional study in a Mediterranean population. BMJ Open 2019 Aug 30;9(8):e029594 [FREE Full text] [doi: 10.1136/bmjopen-2019-029594] [Medline: 31471439]
23. Violán C, Fernández-Bertolín S, Guisado-Clavero M, Foguet-Boreu Q, Valderas J, Vidal Manzano J, et al. Five-year trajectories of multimorbidity patterns in an elderly Mediterranean population using Hidden Markov Models. Sci Rep 2020 Oct 09;10(1):16879-16020 [FREE Full text] [doi: 10.1038/s41598-020-73231-9] [Medline: 33037233]
24. Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, Pons-Vigués M, Foguet-Boreu Q, Muñoz MA, et al. Multimorbidity patterns in the elderly: a prospective cohort study with cluster analysis. BMC Geriatr 2018 Jan 16;18(1):16 [FREE Full text] [doi: 10.1186/s12877-018-0705-7] [Medline: 29338690]
25. Marengoni A, Akugizibwe R, Vetrano D, Roso-Llorach A, Onder G, Welmer A, et al. Patterns of multimorbidity and risk of disability in community-dwelling older persons. Aging Clin Exp Res 2021 Feb;33(2):457-462 [FREE Full text] [doi: 10.1007/s40520-020-01773-z] [Medline: 33580869]
26. Tazzeo C, Rizzuto D, Calderón-Larrañaga A, Roso-Llorach A, Marengoni A, Welmer A, et al. Multimorbidity patterns and risk of frailty in older community-dwelling adults: a population-based cohort study. Age Ageing 2021 Nov 10;50(6):2183-2191 [FREE Full text] [doi: 10.1093/ageing/afab138] [Medline: 34228784]
27. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. J Clin Epidemiol 2014 Mar;67(3):254-266 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.09.021] [Medline: 24472295]
28. Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. PLoS One 2014;9(7):e102149 [FREE Full text] [doi: 10.1371/journal.pone.0102149] [Medline: 25048354]
29. Ng S, Tawiah R, Sawyer M, Scuffham P. Patterns of multimorbid health conditions: a systematic review of analytical methods and comparison analysis. Int J Epidemiol 2018 Oct 01;47(5):1687-1704 [FREE Full text] [doi: 10.1093/ije/dyy134] [Medline: 30016472]
30. Halkidi M, Batistakis Y, Vazirgiannis M. On Clustering Validation Techniques. Journal of Intelligent Information Systems 2001;17:107-145 [FREE Full text] [doi: 10.1023/A:1012801612483]

31. Ward R, Orkaby A, Dumontier C, Charest B, Hawley C, Yaksic E, et al. Trajectories of Frailty in the 5 Years Prior to Death Among U.S. Veterans Born 1927-1934. J Gerontol A Biol Sci Med Sci 2021 Oct 13;76(11):e347-e353 [FREE Full text] [doi: 10.1093/gerona/glab196] [Medline: 34244759]
32. Zamora-Sánchez JJ, Zabaleta-Del-Olmo E, Fernández-Bertolín S, Gea-Caballero V, Julián-Rochina I, Pérez-Tortajada G, et al. Profiles of Frailty among Older People Users of a Home-Based Primary Care Service in an Urban Area of Barcelona (Spain): An Observational Study and Cluster Analysis. J Clin Med 2021 May 13;10(10):2106 [FREE Full text] [doi: 10.3390/jcm10102106] [Medline: 34068296]
33. Bekić S, Babič F, Filipčić I, Trtica Majnarić L. Clustering of Mental and Physical Comorbidity and the Risk of Frailty in Patients Aged 60 Years or More in Primary Care. Med Sci Monit 2019 Sep 11;25:6820-6835 [FREE Full text] [doi: 10.12659/msm.915063]
34. Foreman K, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. The Lancet 2018 Nov;392(10159):2052-2090 [FREE Full text] [doi: 10.1016/s0140-6736(18)31694-5]
35. Alaeddini A, Jaramillo C, Faruqui S, Pugh M. Mining Major Transitions of Chronic Conditions in Patients with Multiple Chronic Conditions. Methods Inf Med 2018 Jan 24;56(05):391-400 [FREE Full text] [doi: 10.3414/me16-01-0135]
36. Jackson C, Dobson A, Tooth L, Mishra G. Body mass index and socioeconomic position are associated with 9-year trajectories of multimorbidity: A population-based study. Prev Med 2015 Dec;81:92-98 [FREE Full text] [doi: 10.1016/j.ypmed.2015.08.013] [Medline: 26311587]
37. Xu X, Mishra G, Dobson A, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLoS Med 2018 Mar;15(3):e1002516-e1002520 [FREE Full text] [doi: 10.1371/journal.pmed.1002516] [Medline: 29534066]
38. Kastner M, Cardoso R, Lai Y, Treister V, Hamid JS, Hayden L, et al. Effectiveness of interventions for managing multiple high-burden chronic diseases in older adults: a systematic review and meta-analysis. CMAJ 2018 Aug 27;190(34):E1004-E1012 [FREE Full text] [doi: 10.1503/cmaj.171391] [Medline: 30150242]
39. SIDIAP. URL: https://www.sidiap.org/index.php/ca/ [accessed 2023-01-05]
40. Recalde M, Rodríguez C, Burn E, Far M, García D, Carrere-Molina J, et al. Data Resource Profile: The Information System for Research in Primary Care (SIDIAP). Int J Epidemiol 2022 Dec 13;51(6):e324-e336 [FREE Full text] [doi: 10.1093/ije/dyac068] [Medline: 35415748]
41. Duque I, Domínguez-Berjón MF, Cebrecos A, Prieto-Salceda M, Esnaola S, Calvo Sánchez M, en nombre del Grupo de Determinantes Sociales de la Salud, iniciativa contexto de la Sociedad Española de Epidemiología. [Deprivation index by enumeration district in Spain, 2011]. Gac Sanit 2021;35(2):113-122 [FREE Full text] [doi: 10.1016/j.gaceta.2019.10.008] [Medline: 32014314]
42. Benchimol E, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015 Oct;12(10):e1001885 [FREE Full text] [doi: 10.1371/journal.pmed.1001885] [Medline: 26440803]
43. Calderón-Larrañaga A, Vetrano D, Onder G, Gimeno-Feliu L, Coscollar-Santaliestra C, Carfí A, et al. Assessing and Measuring Chronic Multimorbidity in the Older Population: A Proposal for Its Operationalization. J Gerontol A Biol Sci Med Sci 2017 Oct 01;72(10):1417-1423 [FREE Full text] [doi: 10.1093/gerona/glw233] [Medline: 28003375]
44. Chavent M, Kuentz V, Labenne A, Liquet B, Saracco J. PCAmixdata: Multivariate Analysis of Mixed Data. R Project. URL: https://cran.r-project.org/web/packages/PCAmixdata/index.html [accessed 2023-06-02]
45. Karlis D, Saporta G, Spinakis A. A Simple Rule for the Selection of Principal Components. Communications in Statistics - Theory and Methods 2003 Jan 04;32(3):643-666 [FREE Full text] [doi: 10.1081/sta-120018556]
46. Calinski T, Harabasz J. A dendrite method for cluster analysis. Comm. in Stats. - Theory & Methods 1974;3(1):1-27 [FREE Full text] [doi: 10.1080/03610927408827101]
47. Bezdek JC. Cluster Validity with Fuzzy Sets. Journal of Cybernetics 1973 Jan;3(3):58-73. [doi: 10.1080/01969727308546047]
48. Carrasco-Ribelles L, Roso-Llorach A, Cabrera-Bean M, Costa-Garrido A, Zabaleta-Del-Olmo E, Toran-Monserrat P, et al. Dynamics of multimorbidity and frailty, and their contribution to mortality, nursing home and home care need: A primary care cohort of 1 456 052 ageing people. EClinicalMedicine 2022 Oct;52:101610 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101610] [Medline: 36034409]
49. Bauer S, Walter L, Ensrud K, Suskind A, Newman J, Ricke W, et al. Assessment of Frailty and Association With Progression of Benign Prostatic Hyperplasia Symptoms and Serious Adverse Events Among Men Using Drug Therapy. JAMA Netw Open 2021 Nov 01;4(11):e2134427 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.34427] [Medline: 34817584]
50. Richardson K, Fox C, Maidment I, Steel N, Loke Y, Arthur A, et al. Anticholinergic drugs and risk of dementia: case-control study. BMJ 2018 Apr 25;361:k1315 [FREE Full text] [doi: 10.1136/bmj.k1315] [Medline: 29695481]
51. Forman D, Maurer M, Boyd C, Brindis R, Salive M, Horne F, et al. Multimorbidity in Older Adults With Cardiovascular Disease. J Am Coll Cardiol 2018 May 15;71(19):2149-2161 [FREE Full text] [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]

52. Skou S, Mair F, Fortin M, Guthrie B, Nunes B, Miranda J, et al. Multimorbidity. Nat Rev Dis Primers 2022 Jul 14;8(1):48 [FREE Full text] [doi: 10.1038/s41572-022-00376-4] [Medline: 35835758]
53. Masoli J, Pilling L, Frayling T. Genomics and multimorbidity. Age Ageing 2022 Dec 05;51(12):afac285 [FREE Full text] [doi: 10.1093/ageing/afac285] [Medline: 36469092]
54. Hanlon P, Nicholl BI, Jani BD, Lee D, McQueenie R, Mair FS. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. The Lancet Public Health 2018 Jul;3(7):e323-e332. [doi: 10.1016/s2468-2667(18)30091-4]
55. Busija L, Lim K, Szoeke C, Sanders K, McCabe M. Do replicable profiles of multimorbidity exist? Systematic review and synthesis. Eur J Epidemiol 2019 Nov;34(11):1025-1053 [FREE Full text] [doi: 10.1007/s10654-019-00568-5] [Medline: 31624969]
56. Cezard G, McHale C, Sullivan F, Bowles J, Keenan K. Studying trajectories of multimorbidity: a systematic scoping review of longitudinal approaches and evidence. BMJ Open 2021 Nov 22;11(11):e048485 [FREE Full text] [doi: 10.1136/bmjopen-2020-048485] [Medline: 34810182]
57. García-Gil MDM, Hermosilla E, Prieto-Alhambra D, Fina F, Rosell M, Ramos R, et al. Construction and validation of a scoring system for the selection of high-quality data in a Spanish population primary care database (SIDIAP). Inform Prim Care 2011;19(3):135-145 [FREE Full text] [doi: 10.14236/jhi.v19i3.806] [Medline: 22688222]
58. Duda RO, Hart PE, Stork DG. Pattern Classification, 2nd Edition. Nashville, TN: John Wiley & Sons; 2000.
59. Cezard G, Sullivan F, Keenan K. Understanding multimorbidity trajectories in Scotland using sequence analysis. Sci Rep 2022 Oct 01;12(1):16485 [FREE Full text] [doi: 10.1038/s41598-022-20546-4] [Medline: 36182953]
60. Howlett S, Rutenberg A, Rockwood K. The degree of frailty as a translational measure of health in aging. Nat Aging 2021 Aug;1(8):651-665 [FREE Full text] [doi: 10.1038/s43587-021-00099-3] [Medline: 37117769]
61. Data requests. SIDIAP. URL: https://www.sidiap.org/index.php/en/solicituds-en [accessed 2023-01-05]

Abbreviations

AIC: Akaike Information Criterion
ATC: Anatomical Therapeutic Classification
AUC: area under the receiver operating characteristic curve
CHI: Catalan Health Institute
EHR: electronic health record
ICD-10: International Classification of Diseases, 10th revision
OE: observed/expected
SIDIAP: Sistema d’Informació pel Desenvolupament de la Investigació a l’Atenció Primària
SNAC-K: Swedish National Study of Aging and Care in Kungsholmen

Edited by A Mavragani, T Sanchez; submitted 19.01.23; peer-reviewed by W Zhang, A Gimeno-Miguel, L Roth, S Puzhko, F Prazeres; comments to author 12.04.23; revised version received 02.05.23; accepted 25.05.23; published 27.06.23.

Please cite as:
Carrasco-Ribelles LA, Cabrera-Bean M, Danés-Castells M, Zabaleta-del-Olmo E, Roso-Llorach A, Violán C
Contribution of Frailty to Multimorbidity Patterns and Trajectories: Longitudinal Dynamic Cohort Study of Aging People
JMIR Public Health Surveill 2023;9:e45848
URL: https://publichealth.jmir.org/2023/1/e45848
doi:10.2196/45848
PMID:37368462

Original Paper

Quantifying Benefit-Risk Trade-Offs Toward Prophylactic Treatment Among Adult Patients With Hemophilia A in China: Discrete Choice Experiment Study

Limin Wang[1,2,3*], MPH; Shimeng Liu[4,5*], PhD; Shan Jiang[6], PhD; Chaofan Li[1,2,3], PhD; Liyong Lu[1,2,3], PhD; Yunhai Fang[7], PhD; Shunping Li[1,2,3], PhD

[1]Centre for Health Management and Policy Research, School of Public Health，Cheeloo College of Medicine, Shandong University, Jinan, China
[2]NHC Key Laboratory of Health Economics and Policy Research, Shandong University, Jinan, China
[3]Centre for Health Preference Research, Shandong University, Jinan, China
[4]School of Public Health, Fudan University, Shanghai, China
[5]NHC Key Laboratory of Health Technology Assessment, Fudan University, Shanghai, China
[6]Macquarie University Centre for the Health Economy, Macquarie Business School and Australian Institute of Health Innovation, Macquarie University, Sydney, Australia
[7]Shandong Hemophilia Treatment Center, Shandong Blood Center, Jinan, China
[*]these authors contributed equally

Corresponding Author:
Shunping Li, PhD
Centre for Health Management and Policy Research
School of Public Health，Cheeloo College of Medicine
Shandong University
44 Wenhuaxi Road
Jinan, 250014
China
Phone: 86 0531 88380161
Fax: 86 13188934998
Email: lishunping@sdu.edu.cn

Abstract

Background: Hemophilia A is a chronic condition that requires meticulous treatment and management. Patient preferences for prophylactic treatment can substantially influence adherence, outcomes, and quality of life, yet these preferences remain underexplored, particularly in China.

Objective: This study aimed to investigate the preferences for prophylactic treatment among Chinese adult patients with hemophilia A without inhibitors, considering clinical effectiveness, side effects, dosing mode, and dosing frequency.

Methods: A discrete choice experiment was used to elicit patient preferences for prophylactic treatment of hemophilia. The study was conducted across 7 provinces in China with socioeconomic and geographical diversity. Subgroup analysis was performed according to education level, geographic location, and treatment type, alongside the exploration of benefit-risk trade-offs.

Results: A total of 113 patients completed the discrete choice experiment questionnaire, and we included 102 responses for analysis based on predetermined exclusion criteria. The study found that patients prioritized reducing annual bleeding times and avoiding the risk of developing inhibitors over treatment process attributes. Subgroup analysis revealed that lower-educated patients and those from rural areas attached more importance to the dosing mode, likely due to barriers to self-administration. Patients demonstrated a clear understanding of benefit-risk trade-offs, exhibiting a willingness to accept an increased risk of developing inhibitors for improved clinical outcomes.

Conclusions: This study provides valuable insights into the preferences of patients with hemophilia A for prophylactic treatment in China. Understanding these preferences can enhance shared decision-making between patients and clinicians, fostering personalized prophylactic treatment plans that may optimize adherence and improve clinical outcomes.

(JMIR Public Health Surveill 2023;9:e45747) doi:10.2196/45747

KEYWORDS

benefit-risk assessment; discrete choice experiment; hemophilia A; patient preference; prophylactic treatment

Introduction

Hemophilia, an infrequent hemorrhagic disorder, is primarily triggered by the absence or deficiency of specific coagulation factors in the blood [1]. According to the World Federation of Hemophilia's 2018 global survey, China was found to have the second highest incidence of confirmed hemophilia cases globally [2]. With a prevalence of 2.73 per 100,000 individuals, hemophilia A (ie, classical hemophilia) constituted more than 80% of these instances [3]. The ensuing joint deformities and disabilities resulting from bleeding into tissues, joints, and muscles critically impair the quality of life for these patients [1].

Factor replacement therapy, classified into on-demand treatment (administered during bleeding episodes) and prophylactic treatment (routine administration to prevent bleeding), is the prevailing treatment modality [4]. However, the development of inhibitors, or an immune response to the therapy, constitutes the most severe adverse effect of this approach [5]. Patients with hemophilia A who have developed inhibitors undergo hemostasis and the removal of inhibitors [4]. In contrast, for patients with hemophilia A devoid of inhibitors, prophylactic treatment remains the optimal care strategy [4].

China has actively pursued and advocated prophylactic treatment measures appropriate to its national context, fostering comprehensive health management of patients with hemophilia [6,7]. Nevertheless, a key impediment remains low adherence to prophylactic treatment. Alarmingly, the rates of prophylactic treatment and treatment compliance are substantially lower in China compared to developed nations (4.1% vs 36.8% and 6.2% vs 40.5%, respectively) [8]. Previous studies have indicated that various factors, such as age, health status, annual bleeding times, infusion methods, dosing regimens, and cost, significantly influence the uptake of prophylactic treatment among adult patients with hemophilia [9,10]. Patients with hemophilia A receiving prophylactic treatment experience a reduction in annual bleeding times, yet they face an associated risk of developing inhibitors. This makes establishing a benefit-risk balance challenging. Therefore, assessments of potential benefits and harms can aid stakeholders, regulators, health technology assessors, and health professionals in understanding and communicating treatment-risk balance, ultimately enhancing adherence to prophylactic treatment [11].

The discrete choice experiment (DCE) is an established stated preference method, extensively used to quantify patients' preferences concerning health care [12-16]. Several DCEs have been conducted to elicit the treatment preferences of patients with hemophilia [17-23]. However, evidence regarding the preference for prophylactic treatment among adult patients with hemophilia A without inhibitors, particularly in terms of benefit-risk assessment, remains sparse in mainland China. This study endeavors to address this research gap by using a DCE as a patient decision aid. It aims to identify the preferences of adult patients with hemophilia A without inhibitors in China concerning the risk and benefit attributes when selecting prophylactic treatment.

Methods

Overview

The DCE is a renowned tool used to simulate the influence of diverse attributes of a service or commodity on individual preference [15,16]. In the scope of a DCE, participants are presented with the task of deciding between 2 or more hypothetical incentive scenarios, differentiated by several relevant dimensions, termed *attributes*. These attributes present varied configurations, referred to as *levels*, across the proposed alternatives [15].

Identification of Attributes and Levels

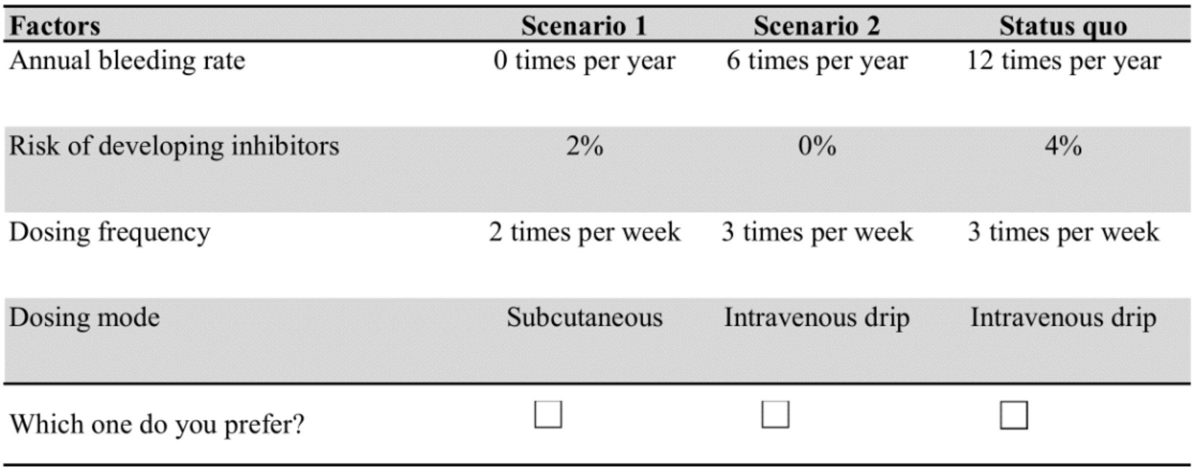
The onset of the DCE involved pinpointing relevant attributes. In accordance with research guidelines, a combination of qualitative and quantitative methods was applied to derive attributes and levels [24]. A total of 7 attributes were ascertained through literature reviews and expert consultations, including the annual bleeding times [21-23,25-28], the risk of developing inhibitors [17-21,27], dosing frequency [18,19,21,23,26-30], dosing mode [26,27,30], storage temperature [21,29], dosage form [18,20,22,26], and the risk of virus infection [17-20]. Following face-to-face patient interviews and expert focus group discussions, 3 attributes (storage temperature, dosage form, and virus infection risk) were dismissed. Due to its complexity, cost was also discarded as a function of individual factors such as weight, the type and specification of medication, and regional health insurance reimbursement rates (varying from 40% to 90%) [31]. Consequently, 4 attributes were selected for inclusion in this study: annual bleeding times, dosing frequency, dosing mode (these 3 represent benefits), and the risk of developing inhibitors (this represents risk). The levels of the 4 DCE attributes were sourced from drug instructions and clinical prophylactic treatment studies [5,32-35]. The chosen attributes and their corresponding levels are detailed in Table 1.

Table 1. Attributes and levels for discrete choice experiment questions.

Attributes	Levels
Annual bleeding times	• 0 times per year • 6 times per year • 12 times per year
Risk of developing inhibitors	• 0% • 2% • 4%
Dosing frequency	• 1 time per week • 2 times per week • 3 times per week
Dosing mode	• Intravenous drip • Intravenous push • Subcutaneous

Questionnaire and DCE Instrument Design

The D-optimal design approach was used to produce 18 choice tasks in SAS (version 9.2; SAS Institute). To alleviate cognitive load, these tasks were divided into 2 blocks. To reflect a more authentic clinical decision-making context, a status quo option was introduced in each choice task. An example of a DCE question is provided in Figure 1.



Factors	Scenario 1	Scenario 2	Status quo
Annual bleeding rate	0 times per year	6 times per year	12 times per year
Risk of developing inhibitors	2%	0%	4%
Dosing frequency	2 times per week	3 times per week	3 times per week
Dosing mode	Subcutaneous	Intravenous drip	Intravenous drip
Which one do you prefer?	☐	☐	☐

Figure 1. An example question in the discrete choice experiment (translated version).

The questionnaire encapsulated an introduction to the research background, attribute definitions, 9 DCE choice tasks, and sociodemographic inquiries. Preceding the official choice tasks, a practice task acclimated respondents to the forthcoming tasks. A question was repeated in each block to validate the internal consistency of the survey. The sociodemographic information gathered included sex, age, education level, employment status, and geographic location.

To ensure respondents comprehended the definition of the risk of developing inhibitors, they were posed a question regarding their understanding of risk percentages [36]. For instance, given a 4% risk of developing inhibitors, respondents were asked to determine how many patients out of 100 would potentially develop inhibitors. The successful answer to this question was a prerequisite for proceeding with the DCE choice tasks.

A pilot study was undertaken among adult patients with hemophilia A without inhibitors at the Shandong Province Hemophilia Treatment Center. The pilot sought to evaluate the understandability, acceptability, and validity of the questionnaire, resulting in revisions to the language and layout.

Sample Selection

This study's sample size was guided by a rule of thumb frequently used in DCE studies that the sample size should comprise no fewer than 75 respondents [37].

Participants were chiefly sourced from Shandong, Hebei, and Henan, which have the largest population of patients with hemophilia A registered with the Hemophilia Treatment Center Collaboration Network of China [38]. Additional patients were recruited from 4 other provinces: Jiangsu, Hubei, Hunan, and Chongqing. The inclusion criteria comprised a hemophilia A diagnosis, no inhibitors present at the time of recruitment, being aged 18 years or older, and receiving treatment for more than 50 exposure days. Exposure days refer to the cumulative number of days coagulation factor VIII was injected, which influences the likelihood of producing an inhibitor.

Data Collection

Formal data collection occurred between December 2021 and March 2022. To ensure data quality, we sent the electronic questionnaire to respondents first and then conducted the interview over the telephone in a one-on-one manner. The interviewers answered respondents’ questions. Through a warm-up section, the interviewers trained the respondents to familiarize themselves with the DCE choice tasks and guarantee they comprehended the whole questionnaire.

Model Specification

A mixed logit model was used to analyze the choice data in Stata (version 15.1; StataCorp), estimated through the simulated maximum likelihood approach [39,40]. The mixed logit model included an alternative-specific constant (ASC) indicative of the utility generated by the status quo option in comparison to non–status quo options. We hypothesized that the parameters of attribute levels would conform to a normal distribution. Parameter estimations were derived relative to the reference level within each attribute. The average preference value, termed *part-worth utility*, and the variability of the preference value among patients with hemophilia were represented by the mean and SD of a parameter, respectively. Model fit was determined by the Akaike information criterion (AIC), Bayesian information criterion (BIC), and log-likelihood ratio. To ensure the reliability of parameter estimates, we iteratively estimated the mixed logit model by incrementally increasing the number of random draws by 500, commencing with 50 draws. Estimation stability was achieved at 2500 draws, producing our final estimates [41].

We excluded patients who failed the consistency test or always chose the options on the left or right in all the choice tasks (ie, position bias) [15]. We also conducted a sensitivity check to examine whether the exclusion would significantly affect the results of the mixed logit model.

Dominant Preference Examination

Dominant preference refers to the phenomenon whereby a respondent’s choices are dictated by a single attribute, resulting in decisions that consistently favor the alternative with a superior level of 1 attribute in a choice task [42]. Such respondents avoid making trade-offs between attributes, impeding the analysis of relative importance between attributes [42]. We used nonparametric [42] and parametric [43] approaches to test for dominant preference. The nonparametric approach evaluated if respondents’ choices exhibited a pattern, that is, if the selected alternatives for the 7 tasks were invariably consistent with the alternatives with a higher level of a certain attribute in comparison to the other alternative in all 7 questions. The parametric approach compared the estimated coefficients of attribute levels between a full model encompassing all attributes and a reduced model containing only 1 attribute. Notable discrepancies between the coefficients indicated the presence of a dominant preference driven by that attribute. We iteratively inserted each attribute into the reduced model and compared the model estimates with the full model.

Attribute Relative Importance

The relative importance of each attribute was computed using the mixed logit model estimates through a widely used rescaling method [16,44]. The relative importance of each attribute was determined by dividing the range of coefficients within the attribute by the sum of all attribute ranges, which was subsequently rescaled to a 1-100 range. The highest value denoted the attribute perceived as most important by the respondents.

Interaction

A meticulous examination of all potential interactions between the characteristics of respondents and attribute levels was carried out using the multinomial logit model (MNL) [15]. Interaction terms were selected using a backward selection method based on the contribution of each term to model fit. The log-likelihood ratio test was used to compare the model specifications with a reduced model with 1 interaction term removed. If the removed term significantly influenced the model fit, the term was retained. Following the identification of interaction terms, we simulated a mixed logit model that incorporated the interactions to quantify the preference values assigned to the interaction terms by respondents.

Maximum Acceptable Risk

We quantified patients’ tolerance for the risk of developing inhibitors in exchange for improvements in other attributes. The outcome was termed the maximum acceptable risk (MAR) by patients. The MAR gauged the benefit-risk ratio that patients were willing to accept in terms of the trade-off between benefit and risk, given that the utility remained constant [45].

Subgroup Analysis

We conducted subgroup analyses by comparing the relative importance of attributes between different groups of respondents using individual-level preferences. We applied a mixed logit model to the variables used to categorize respondents, which encompassed education level, geographic location, and treatment type.

Ethics Approval

Ethical approval for this study was granted by the Center for Health Management and Policy Research, Shandong University (ECSHCMSDU20211102).

Results

Pilot Study

This pilot study incorporated a cohort of 15 patients, all of whom successfully completed the study. The derived model estimates were in alignment with our theoretical anticipations in terms of their coefficient signs and priority order, thereby affirming that the participants comprehended the choice tasks and that the quantity of questions was manageable. The mean time taken to conclude the study was approximately 15 minutes.

Patient Characteristics

We invited 140 patients who were eligible for inclusion, and 122 consented to participate (a response rate of 87%), of whom 113 completed the survey (a completion rate of 93%). On average, participants took 14.60 (SD 5.41) minutes to complete the questionnaire. Among those who completed the questionnaire, 10 participants did not pass the consistency test, while 1 participant exhibited position bias (ie, always choosing the left or right options); hence, we excluded them from further analysis. Thus, a total of 102 participants were included in the final analysis.

The demographic profile of the participants included in the analysis is presented in Table 2. The majority of patients were male (100/102, 98%), possessed an education level equivalent to high school or higher (54/102, 52.9%), were single (59/102, 57.8%), and resided in townships or rural areas (66/102, 64.7%). Most patients were diagnosed with severe hemophilia A (58/102, 56.9%) and received on-demand treatment (56/102, 54.9%).

Table 2. Sociodemographic characteristics of patients included in the analysis.

Characteristic	Patients included in the study (n=102)
Age (in years), mean (SD)	33.8 (7.5)
Sex, n (%)	
Male	100 (98)
Female	2 (1.9)
Education level, n (%)	
Middle school or below	48 (47.1)
High school or above	54 (52.9)
Marital status, n (%)	
Unmarried	59 (57.8)
Married	41 (40.2)
Divorced	2 (2)
Occupation status, n (%)	
Unemployed	42 (41.2)
Working full time	25 (24.5)
Working part time	11 (10.8)
Self-employed	9 (8.8)
Student	14 (13.7)
Retired	1 (1)
Geographic location, n (%)	
Urban	36 (35.3)
Rural	66 (64.7)
Severity of hemophilia A, n (%)	
Mild	3 (2.9)
Moderate	41 (40.2)
Severe	58 (56.9)
Treatment type, n (%)	
On-demand	56 (54.9)
Prophylaxis	41 (40.2)
Nonstandard prophylaxis	5 (4.9)

Patients' Preferences

Table 3 presents the mixed logit model results. The constant representing the status quo option was not statistically significant, indicating patients expressed no preference for the status quo compared with nonstatus quo alternatives.

Respondents expressed a higher likelihood of accepting either 0 or 6 instances of bleeding annually, as opposed to 12. Utility declined as the risk of developing inhibitors increased. In terms of dosing frequency, patients demonstrated a preference for 1 or 2 doses per week over 3. With respect to the mode of administration, intravenous pushes and subcutaneous injections were preferred over intravenous drips.

We did not identify the presence of dominant preferences through either parametric and nonparametric approaches. A sensitivity analysis was conducted by including the 11 patients who failed the internal validity test or exhibited position bias (Table S1 in Multimedia Appendix 1). We found no significant differences between the 2 models—one including and the other excluding these 11 patients—indicating that the exclusion did not affect the findings.

Table 3. Mixed logit model results.

	β	*P* value	SE	SD	*P* value	SE
Constant	0.72	.373	0.80	3.80	<.001	0.76
Annual bleeding times (reference level 12 per year)						
6 per year	1.80	<.001	0.25	0.00	.99	0.38
0 per year	3.64	<.001	0.39	1.49	<.001	0.31
Risk of developing inhibitors	−0.79	<.001	0.39	0.54	<.001	0.09
Dosing frequency (reference level 3 times per week)						
2 times per week	1.20	<.001	0.24	0.05	.88	0.32
1 time per week	1.58	<.001	0.25	0.01	.98	0.38
Dosing mode (reference level intravenous drip)						
Intravenous push	0.74	<.001	0.19	0.07	.84	0.33
Subcutaneous	1.35	<.001	0.29	1.95	<.001	0.34

Attribute Relative Importance

The frequency of annual bleeding events emerged as the most significant attribute from the perspective of the participants, with the attribute relative importance (ARI) score at 100. The risk of developing inhibitors ranked second in importance (ARI=87), while dosing frequency (ARI=43) and dosing mode (ARI=37) were considered comparatively less critical.

Interaction

The selection of interaction terms identified a significant interaction between the status-quo constant and the education level of participants (Table S2 in Multimedia Appendix 1). The positive preference value suggested that respondents with a higher level of education (ie, high school or above) were more likely to opt for the status quo alternative than those with a lower educational level (ie, middle school or below).

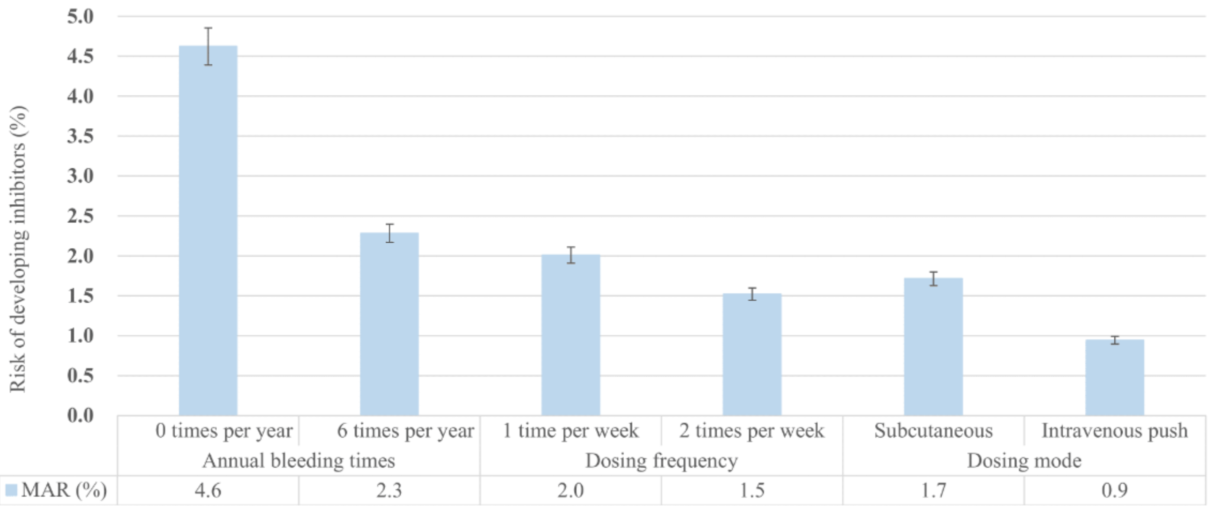
Subgroup Analysis

Subgroup analysis was performed based on variables such as education level, geographic location, and treatment type. Patients with lower education levels (middle school or below) considered the dosing mode more important than the dosing frequency. Conversely, patients with a higher education level (high school or above) placed greater importance on dosing frequency compared to dosing mode (Figure S1 in Multimedia Appendix 1). A similar contrast was observed between urban and rural respondents (Figure S2 in Multimedia Appendix 1). While urban respondents emphasized the dosing frequency over the dosing mode, rural respondents held the opposite view. The comparison of patients undergoing prophylaxis treatment versus those receiving on-demand treatment revealed a common preference pattern; both groups viewed dosing frequency as more important than dosing mode (Figure S3 in Multimedia Appendix 1).

Benefit-Risk Trade-Off

The results pertaining to MAR are depicted in Figure 2. For specified improvements in treatment attributes such as annual bleeding frequency, dosing frequency, and dosing mode, the MAR indicates the highest level of risk related to developing inhibitors that patients would tolerate, thereby hinting at patients’ willingness to accept a trade-off between potential benefits and risks.

Figure 2. Benefit-risk trade-off and maximum acceptable risk. MAR: maximum acceptable risk.



Patients were willing to accept an increase in the risk of developing inhibitors (MAR=4.6% for reducing annual bleeding instances from 12 to 0). They would also tolerate an increased risk of 2.3% for reducing the frequency of bleeding from 12 to 6 times per year. Furthermore, an increased risk of 2% would be acceptable to patients if the dosing frequency was decreased from 3 times a week to once a week, while an increased risk of 1.5% would be acceptable if the dosing frequency was reduced from 3 to 2 times a week. In terms of dosing mode, patients would tolerate a risk increase of 1.7% if the dosing model shifted from intravenous drip to subcutaneous injection and a risk increase of 0.9% if the dosing model was altered from intravenous drip to intravenous push.

Discussion

Overview

This is the first study investigating the preferences of patients with hemophilia A without inhibitors regarding prophylactic treatment by performing a DCE in China. We examined 4 attributes associated with prophylactic treatment, with results indicating patients placed higher value on clinical effectiveness (ie, reduction in annual bleeding times) and side effects (ie, risk of developing inhibitors) compared to aspects of the treatment process (ie, dosing frequency and mode). We also found that patients were willing to accept an increase of 4.6% in the risk of developing inhibitors for the reduction of bleeding from 12 times per year to 0 times per year.

This study showed that patients attach paramount importance to the reduction of annual bleeding times, an outcome consistent with previous investigations [21,25,27]. Our findings suggest that alterations in the mode and frequency of administration exert a relatively minimal influence on patients' preferences. In contrast, a previous study reported dissimilar results by using conjoint analysis to ascertain patient and parental preferences for prophylactic treatment [26]. Their study indicated that the frequency of prophylactic administration was deemed the most crucial attribute, followed by the annual bleeding times. This discrepancy could be attributed to differences in the standard of treatment across various countries. China primarily uses low-dose prophylaxis, which has comparatively less control over bleeding than standard-dose prophylaxis [46]. The aforementioned study, conducted in Australia, Britain, and the United States, demonstrated that reduced instances of breakthrough bleeds during prophylaxis (including 0, 1, 2, and 3 times per year) exerted a lesser influence on patient preferences. Given the current status of prophylactic treatment in China, it is plausible to propose that patients may favor a treatment regimen with enhanced bleeding control, which has implications for the development of personalized prophylactic treatment protocols [47].

Individual-Level Factors

Subgroup analysis revealed that patients with lower education levels prioritized the dosing mode. Patients residing in townships or rural areas also exhibited higher sensitivity toward the dosing mode compared to their urban counterparts. This could be attributed to the barriers they face in the self-administration of prophylactic infusions. Unlike typical oral or inhaled medications, coagulation factors require injection, necessitating specialized self-care skills or medical provider support. Self-injection therapy for patients with hemophilia remains challenging [48], particularly for patients with a lower level of education, who are less likely to master self-injection techniques. Patients in townships or rural areas often have a lower quality of life compared to those in urban areas [49]. Thus, it is essential to focus on self-management treatment services for patients with low education levels and those in rural areas. Such services could include intravenous injection training programs and consultation and guidance services [48].

Benefit-Risk Trade-Off

This study illustrated that patients demonstrated clear awareness of the benefit-risk assessment. The results suggested that patients were willing to accept a substantial risk of developing inhibitors for reducing the annual bleeding times, but were only willing to accept a minor risk alteration when changing drug administration from intravenous drip to intravenous push. The risk of developing inhibitors could be controlled by the treatment strategy [50,51]. Particularly, a heterogeneity analysis revealed that patients undergoing prophylactic treatment were more attuned to the risk of developing inhibitors. We propose that physicians must judiciously balance the trade-offs between bleeding and the risk of developing inhibitors when recommending prophylactic treatment for patients. This is of particular importance when optimal control of bleeding is achieved, as the risks of developing inhibitors may exceed the maximum acceptable range for the patient.

Limitations

This study fills a significant gap in our understanding of treatment preferences for people with hemophilia in mainland China. However, certain limitations persist. There could be potential sampling bias as we did not use random sampling. Nonetheless, to enhance the representativeness of our results, we examined the prophylactic preferences of adult patients with hemophilia A without inhibitors across 7 regions of China, accounting for geographical diversity and varied economic development status. Moreover, DCE, being a hypothetical choice scenario, cannot accurately track actual choice behavior. Therefore, to mitigate the impact of this potential discrepancy, we ensured the attribute levels were scientifically valid and realistic in the preliminary stage, thereby providing respondents with a credible profile for comparison and evaluation.

Conclusions

Our study's findings significantly contribute to understanding the preferences of adult patients with hemophilia A regarding prophylactic treatment. Acknowledging that patients will perceive trade-offs for the 4 attributes differently can enhance the dialogue between patients and clinicians on the risks and benefits of various prophylactic treatment modalities.

Acknowledgments

The authors would like to thank all participants for their contributions to this study. The authors would like to thank all the collaborators who cooperated to recruit patients with hemophilia A and the leaders of the organizations that supported this survey for patients with hemophilia. This work was supported by the National Natural Science Foundation of China (grant 72174110).

Authors' Contributions

SPL, LW, SML, CL, and YF designed the study. LW collected the data. LW, SML, SJ, and LL completed the data analyses and wrote the manuscript. SJ, SML, CL, LL, YF, and SPL edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary material for conditional logit models, results of mixed logit model with main effects and interactions, attribute relative importance perceived by high-education respondents vs low-education respondents, attribute relative importance perceived by urban respondents vs rural respondents, and attribute relative importance perceived by respondents receiving prophylactic treatment vs respondents receiving on-demand treatment.

[DOCX File , 132 KB - publichealth_v9i1e45747_app1.docx]

References

1. Peyvandi F, Garagiola I, Young G. The past and future of haemophilia: diagnosis, treatments, and its complications. Lancet 2016;388(10040):187-197. [doi: 10.1016/S0140-6736(15)01123-X] [Medline: 26897598]
2. Report on the Annual Global Survey 2018. World Federation of Hemophilia. 2019. URL: http://www1.wfh.org/publications/files/pdf-1731.pdf [accessed 2023-07-06]
3. Qu Y, Nie X, Yang Z, Yin H, Pang Y, Dong P, et al. The prevalence of hemophilia in mainland China: a systematic review and meta-analysis. Southeast Asian J Trop Med Public Health 2014 Mar;45(2):455-466. [Medline: 24968688]
4. Thrombosis and Hemostasis Group, Chinese Society of Hematology, Chinese Medical Association/Hemophilia Treatment Center Collaborative Network of China. [Chinese guidelines on the treatment of hemophilia (version 2020)]. Zhonghua Xue Ye Xue Za Zhi 2020;41(4):265-271 [FREE Full text] [doi: 10.3760/cma.j.issn.0253-2727.2020.04.001] [Medline: 32295333]
5. van den Berg HM, Fischer K, Carcao M, Chambost H, Kenet G, Kurnik K, PedNet Study Group. Timing of inhibitor development in more than 1000 previously untreated patients with severe hemophilia A. Blood 2019;134(3):317-320 [FREE Full text] [doi: 10.1182/blood.2019000658] [Medline: 31186271]
6. Yang R. [To implement prophylactic treatment and improve the quality of life in patients with hemophilia in China]. Zhonghua Xue Ye Xue Za Zhi 2015;36(7):537-538 [FREE Full text] [doi: 10.3760/cma.j.issn.0253-2727.2015.07.001] [Medline: 26304072]

7. Yang R, Poon MC, Luke KH, Zhao Y, Sun J, Wang X, et al. Building a network for hemophilia care in China: 15 years of achievement for the Hemophilia Treatment Center Collaborative Network of China. Blood Adv 2019;3(Suppl 1):34-37 [FREE Full text] [doi: 10.1182/bloodadvances.2019GS121524] [Medline: 31809551]
8. Sun J, Zhao Y, Yang R, Guan T, Iorio A, Chinese HERO study group. The demographics, treatment characteristics and quality of life of adult people with haemophilia in China - results from the HERO study. Haemophilia 2017;23(1):89-97. [doi: 10.1111/hae.13071] [Medline: 27599642]
9. Thornburg CD, Duncan NA. Treatment adherence in hemophilia. Patient Prefer Adherence 2017;11:1677-1686 [FREE Full text] [doi: 10.2147/PPA.S139851] [Medline: 29033555]
10. Lee Mortensen G, Strand AM, Almén L. Adherence to prophylactic haemophilic treatment in young patients transitioning to adult care: a qualitative review. Haemophilia 2018;24(6):862-872 [FREE Full text] [doi: 10.1111/hae.13621] [Medline: 30485633]
11. Tervonen T, Veldwijk J, Payne K, Ng X, Levitan B, Lackey LG, et al. Quantitative benefit-risk assessment in medical product decision making: a good practices report of an ISPOR task force. Value Health 2023;26(4):449-460 [FREE Full text] [doi: 10.1016/j.jval.2022.12.006] [Medline: 37005055]
12. Haghani M, Bliemer MCJ, Hensher DA. The landscape of econometric discrete choice modelling research. J Choice Model 2021;40:100303. [doi: 10.1016/j.jocm.2021.100303]
13. Asim S, Wang K, Nichini E, Yip FF, Zhu L, Fung HCE, et al. COVID-19 vaccination preferences among non-Chinese migrants in Hong Kong: discrete choice experiment. JMIR Public Health Surveill 2023 27;9:e40587 [FREE Full text] [doi: 10.2196/40587] [Medline: 36848242]
14. Tran BX, Do AL, Boyer L, Auquier P, Le HT, Le Vu MN, et al. Preference and willingness to pay for the regular COVID-19 booster shot in the Vietnamese population: theory-driven discrete choice experiment. JMIR Public Health Surveill 2023;9:e43055 [FREE Full text] [doi: 10.2196/43055] [Medline: 36599156]
15. Jiang S, Anis AH, Cromwell I, Mohammadi T, Schrader KA, Lucas J, et al. Health-care practitioners' preferences for the return of secondary findings from next-generation sequencing: a discrete choice experiment. Genet Med 2020;22(12):2011-2019 [FREE Full text] [doi: 10.1038/s41436-020-0927-x] [Medline: 32820245]
16. Jiang S, Ren R, Gu Y, Jeet V, Liu P, Li S. Patient preferences in targeted pharmacotherapy for cancers: a systematic review of discrete choice experiments. Pharmacoeconomics 2023;41(1):43-57 [FREE Full text] [doi: 10.1007/s40273-022-01198-8] [Medline: 36372823]
17. Brown TM, Pashos CL, Joshi AV, Lee WC. The perspective of patients with haemophilia with inhibitors and their care givers: preferences for treatment characteristics. Haemophilia 2011;17(3):476-482. [doi: 10.1111/j.1365-2516.2010.02401.x] [Medline: 21091851]
18. Mantovani LG, Monzini MS, Mannucci PM, Scalone L, Villa M, Gringeri A, Conan Study Group. Differences between patients', physicians' and pharmacists' preferences for treatment products in haemophilia: a discrete choice experiment. Haemophilia 2005;11(6):589-597. [doi: 10.1111/j.1365-2516.2005.01159.x] [Medline: 16236108]
19. Scalone L, Mantovani LG, Borghetti F, Von Mackensen S, Gringeri A. Patients', physicians', and pharmacists' preferences towards coagulation factor concentrates to treat haemophilia with inhibitors: results from the COHIBA study. Haemophilia 2009;15(2):473-486. [doi: 10.1111/j.1365-2516.2008.01926.x] [Medline: 19347988]
20. Mohamed AF, Epstein JD, Li-McLeod JM. Patient and parent preferences for haemophilia A treatments. Haemophilia 2011;17(2):209-214. [doi: 10.1111/j.1365-2516.2010.02411.x] [Medline: 21070493]
21. Fifer S, Kerr AM, Parken C, Hamrosi K, Eid S. Treatment preferences in people with haemophilia A or caregivers of people with haemophilia A: a discrete choice experiment. Haemophilia 2020;26(Suppl 5):30-40. [doi: 10.1111/hae.14037] [Medline: 32935396]
22. Chaugule SS, Hay JW, Young G. Understanding patient preferences and willingness to pay for hemophilia therapies. Patient Prefer Adherence 2015;9:1623-1630 [FREE Full text] [doi: 10.2147/PPA.S92985] [Medline: 26635471]
23. Lock J, de Bekker-Grob EW, Urhan G, Peters M, Meijer K, Brons P, 'OPTI-CLOT' study group. Facilitating the implementation of pharmacokinetic-guided dosing of prophylaxis in haemophilia care by discrete choice experiment. Haemophilia 2016;22(1):e1-e10. [doi: 10.1111/hae.12851] [Medline: 26612493]
24. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete choice experiments in health economics: past, present and future. Pharmacoeconomics 2019;37(2):201-226 [FREE Full text] [doi: 10.1007/s40273-018-0734-2] [Medline: 30392040]
25. Witkop M, Morgan G, O'Hara J, Recht M, Buckner TW, Nugent D, et al. Patient preferences and priorities for haemophilia gene therapy in the US: a discrete choice experiment. Haemophilia 2021;27(5):769-782 [FREE Full text] [doi: 10.1111/hae.14383] [Medline: 34310811]
26. Furlan R, Krishnan S, Vietri J. Patient and parent preferences for characteristics of prophylactic treatment in hemophilia. Patient Prefer Adherence 2015;9:1687-1694 [FREE Full text] [doi: 10.2147/PPA.S92520] [Medline: 26648701]
27. Mühlbacher AC, Sadler A, Lamprecht B, Juhnke C. Patient preferences in the treatment of hemophilia A: a best-worst scaling case 3 analysis. Value Health 2020;23(7):862-869 [FREE Full text] [doi: 10.1016/j.jval.2020.02.013] [Medline: 32762987]
28. Su J, Li N, Joshi N, Ng X, Botteman M, Shah R, et al. Patient and caregiver preferences for haemophilia A treatments: a discrete choice experiment. Haemophilia 2020;26(6):e291-e299. [doi: 10.1111/hae.14137] [Medline: 32937681]

29. Tischer B, Marino R, Napolitano M. Patient preferences in the treatment of hemophilia A: impact of storage conditions on product choice. Patient Prefer Adherence 2018;12:431-441 [FREE Full text] [doi: 10.2147/PPA.S151812] [Medline: 29618923]
30. Park YS, Hwang TJ, Cho GJ, Kim SK, Park SK, Kim JY, et al. Patients' and parents' satisfaction with, and preference for, haemophilia A treatments: a cross-sectional, multicentre, observational study. Haemophilia 2021;27(4):563-573 [FREE Full text] [doi: 10.1111/hae.14304] [Medline: 34128300]
31. Wang X. The impact of medical insurance reimbursement policy on the treatment level and disease burden of patients with rare diseases, taking hemophilia A as an example [D]. Tianjin University. 2020. URL: http://kns.cnki.net/kcms/detail/detail.aspx?doi=10.27356/d.cnki.gtjdu.2020.001068&dbcode=CMFD [accessed 2023-07-06]
32. Kempton CL. Inhibitors in previously treated patients: a review of the literature. Haemophilia 2010;16(102):61-65 [FREE Full text] [doi: 10.1111/j.1365-2516.2010.02235.x] [Medline: 20536987]
33. Mahlangu J, Oldenburg J, Paz-Priel I, Negrier C, Niggli M, Mancuso ME, et al. Emicizumab prophylaxis in patients who have hemophilia A without inhibitors. N Engl J Med 2018;379(9):811-822 [FREE Full text] [doi: 10.1056/NEJMoa1803550] [Medline: 30157389]
34. Blair HA. Emicizumab: a review in haemophilia A. Drugs 2019;79(15):1697-1707. [doi: 10.1007/s40265-019-01200-2] [Medline: 31542880]
35. Verma SP, Dutta TK, Mahadevan S, Nalini P, Basu D, Biswal N, et al. A randomized study of very low-dose factor VIII prophylaxis in severe haemophilia - a success story from a resource limited country. Haemophilia 2016;22(3):342-348. [doi: 10.1111/hae.12838] [Medline: 26987935]
36. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. J Natl Cancer Inst 2011;103(19):1436-1443 [FREE Full text] [doi: 10.1093/jnci/djr318] [Medline: 21931068]
37. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015;8(5):373-384 [FREE Full text] [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
38. Song X, Zhong J, Xue F, Chen L, Li H, Yuan D, et al. An overview of patients with haemophilia A in China: epidemiology, disease severity and treatment strategies. Haemophilia 2021;27(1):e51-e59. [doi: 10.1111/hae.14217] [Medline: 33245829]
39. Train KE. Discrete Choice Methods with Simulation, Second Edition. New York, NY: Cambridge University Press; 2009.
40. Hole AR. Fitting mixed logit models by using maximum simulated likelihood. Stata J 2007;7(3):388-401 [FREE Full text] [doi: 10.1177/1536867x0700700306]
41. Jiang S, Gu Y, Yang F, Wu T, Wang H, Cutler H, et al. Tertiary hospitals or community clinics? An enquiry into the factors affecting patients' choice for healthcare facilities in urban China. China Econ Rev 2020;63:101538. [doi: 10.1016/j.chieco.2020.101538]
42. Vaughan L, Bardsley M, Bell D, Davies M, Goddard A, Imison C, et al. Models of generalist and specialist care in smaller hospitals in England: A mixed-methods study. National Institute for Health Research. URL: https://www.ncbi.nlm.nih.gov/books/NBK568036/pdf/Bookshelf_NBK568036.pdf [accessed 2023-07-06]
43. Mohammadi T, Zhang W, Sou J, Langlois S, Munro S, Anis AH. A Hierarchical Bayes approach to modeling heterogeneity in discrete choice experiments: an application to public preferences for prenatal screening. Patient 2020;13(2):211-223. [doi: 10.1007/s40271-019-00402-w] [Medline: 31814082]
44. Krucien N, Watson V, Ryan M. Is best-worst scaling suitable for health state valuation? A comparison with discrete choice experiments. Health Econ 2017;26(12):e1-e16. [doi: 10.1002/hec.3459] [Medline: 27917560]
45. Johnson FR, Hauber AB, Poulos CM. A brief introduction to the use of stated-choice methods to measure preferences for treatment benefits and risks. ResearchGate. 2009. URL: https://www.researchgate.net/publication/237321750_A_Brief_Introduction_to_the_Use_of_Stated-Choice_Methods_to_Measure_Preferences_for_Treatment_Benefits_and_Risks [accessed 2023-06-27]
46. Wu R, Luke KH. The benefit of low dose prophylaxis in the treatment of hemophilia: a focus on China. Expert Rev Hematol 2017;10(11):995-1004. [doi: 10.1080/17474086.2017.1386096] [Medline: 28971697]
47. Petrini P, Valentino LA, Gringeri A, Re WM, Ewenstein B. Individualizing prophylaxis in hemophilia: a review. Expert Rev Hematol 2015;8(2):237-246. [doi: 10.1586/17474086.2015.1002465] [Medline: 25600578]
48. Teitel JM, Barnard D, Israels S, Lillicrap D, Poon MC, Sek J. Home management of haemophilia. Haemophilia 2004 Mar;10(2):118-133. [doi: 10.1046/j.1365-2516.2003.00853.x] [Medline: 14962201]
49. Carroll L, Benson G, Lambert J, Benmedjahed K, Zak M, Lee XY. Real-world utilities and health-related quality-of-life data in hemophilia patients in France and the United Kingdom. Patient Prefer Adherence 2019;13:941-957 [FREE Full text] [doi: 10.2147/PPA.S202773] [Medline: 31354248]
50. Brackmann HH, White GC, Berntorp E, Andersen T, Escuriola-Ettingshausen C. Immune tolerance induction: What have we learned over time? Haemophilia 2018 Apr;24 Suppl 3:3-14. [doi: 10.1111/hae.13445] [Medline: 29543371]
51. Witmer C, Young G. Factor VIII inhibitors in hemophilia A: rationale and latest evidence. Ther Adv Hematol 2013;4(1):59-72 [FREE Full text] [doi: 10.1177/2040620712464509] [Medline: 23610614]

Abbreviations

AIC: Akaike information criterion
ARI: attribute relative importance
ASC: alternative-specific constant
BIC: Bayesian information criterion
DCE: discrete choice experiment
MAR: maximum acceptable risk
MNL: multinomial logit model

Edited by M Focsa; submitted 17.01.23; peer-reviewed by WC Lee, X Ng, D Kukadiya; comments to author 16.03.23; revised version received 25.05.23; accepted 17.06.23; published 26.07.23.

Please cite as:
Wang L, Liu S, Jiang S, Li C, Lu L, Fang Y, Li S
Quantifying Benefit-Risk Trade-Offs Toward Prophylactic Treatment Among Adult Patients With Hemophilia A in China: Discrete Choice Experiment Study
JMIR Public Health Surveill 2023;9:e45747
URL: https://publichealth.jmir.org/2023/1/e45747
doi:10.2196/45747
PMID:37494098

Original Paper

Prediction Model for Postoperative Quality of Life Among Breast Cancer Survivors Along the Survivorship Trajectory From Pretreatment to 5 Years: Machine Learning–Based Analysis

Danbee Kang[1,2*], PhD; Hyunsoo Kim[1,2*], MS; Juhee Cho[1,2], MA, PhD; Zero Kim[3], PhD; Myungjin Chung[3], MD, PhD; Jeong Eon Lee[4], MD, PhD; Seok Jin Nam[4], MD, PhD; Seok Won Kim[4], MD, PhD; Jonghan Yu[4], MD, PhD; Byung Joo Chae[4], MD, PhD; Jai Min Ryu[4], MD, PhD; Se Kyung Lee[4], MD, PhD

[1]Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea
[2]Center for Clinical Epidemiology, Samsung Medical Center, Seoul, Republic of Korea
[3]Medical AI Research Center, Samsung Medical Center, Seoul, Republic of Korea
[4]Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Se Kyung Lee, MD, PhD
Department of Surgery
Samsung Medical Center
Sungkyunkwan University School of Medicine
81 Irwon-ro, Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 2 3410 3478
Fax: 82 2 3410 6982
Email: zzangdoc@gmail.com

Abstract

Background: Breast cancer is the most common cancer and the most common cause of cancer death in women. Although survival rates have improved, unmet psychosocial needs remain challenging because the quality of life (QoL) and QoL-related factors change over time. In addition, traditional statistical models have limitations in identifying factors associated with QoL over time, particularly concerning the physical, psychological, economic, spiritual, and social dimensions.

Objective: This study aimed to identify patient-centered factors associated with QoL among patients with breast cancer using a machine learning (ML) algorithm to analyze data collected along different survivorship trajectories.

Methods: The study used 2 data sets. The first data set was the cross-sectional survey data from the Breast Cancer Information Grand Round for Survivorship (BIG-S) study, which recruited consecutive breast cancer survivors who visited the outpatient breast cancer clinic at the Samsung Medical Center in Seoul, Korea, between 2018 and 2019. The second data set was the longitudinal cohort data from the Beauty Education for Distressed Breast Cancer (BEST) cohort study, which was conducted at 2 university-based cancer hospitals in Seoul, Korea, between 2011 and 2016. QoL was measured using European Organization for Research and Treatment of Cancer QoL Questionnaire Core 30 questionnaire. Feature importance was interpreted using Shapley Additive Explanations (SHAP). The final model was selected based on the highest mean area under the receiver operating characteristic curve (AUC). The analyses were performed using the Python 3.7 programming environment (Python Software Foundation).

Results: The study included 6265 breast cancer survivors in the training data set and 432 patients in the validation set. The mean age was 50.6 (SD 8.66) years and 46.8% (n=2004) had stage 1 cancer. In the training data set, 48.3% (n=3026) of survivors had poor QoL. The study developed ML models for QoL prediction based on 6 algorithms. Performance was good for all survival trajectories: overall (AUC 0.823), baseline (AUC 0.835), within 1 year (AUC 0.860), between 2 and 3 years (AUC 0.808), between 3 and 4 years (AUC 0.820), and between 4 and 5 years (AUC 0.826). Emotional and physical functions were the most important features before surgery and within 1 year after surgery, respectively. Fatigue was the most important feature between 1 and 4 years. Despite the survival period, hopefulness was the most influential feature on QoL. External validation of the models showed good performance with AUCs between 0.770 and 0.862.

Conclusions: The study identified important factors associated with QoL among breast cancer survivors across different survival trajectories. Understanding the changing trends of these factors could help to intervene more precisely and timely, and potentially prevent or alleviate QoL-related issues for patients. The good performance of our ML models in both training and external validation sets suggests the potential use of this approach in identifying patient-centered factors and improving survivorship care.

(JMIR Public Health Surveill 2023;9:e45212) doi:10.2196/45212

KEYWORDS

breast cancer survivor; quality of life; machine learning; trajectory; predict; develop; breast cancer; survivor; cancer; oncology; algorithm; model; QoL

Introduction

Breast cancer is the most common cancer and the most common cause of cancer death in women worldwide [1]. In the past years, breast cancer prognosis has significantly improved over time. Currently, the 5-year survival rates are in the range of 90%, and 10-year survival is about 80%. Given the increase in survival, a survivorship care plan is necessary over time, with particular attention to the quality of life (QoL) [2]. However, for many survivors, cancer survivorship is characterized by uncertainty regarding follow-up care and unmet psychosocial needs [3].

To develop tailored interventions and to provide appropriate survivorship care, it is necessary to find predictors for QoL during different phases of survivorship [4,5]. Although some predictors for QoL have been identified in several studies [6], almost all focused on 1 specific predictor. Fewer models have made individual predictions on QoL due to the complexity of clinical profiles and the inability to consider relevant interactions a priori. In addition, according to a recent cohort study, the QoL and the QoL-related factors change over time [7]. However, it is difficult to generate those models using traditional statistical methods.

To overcome the limitation of traditional models, a few machine learning (ML) models have been proposed in the literature that predict the QoL of breast cancer survivors. However, there were only a few ML models for QoL prediction with limitations [6,8-12]. First, most models did not fully include multidimensional factors. Although some studies included patient-centered factors such as functional impairment and psychological symptoms, they still missed key variables for the QoL of long-term survivors, such as spiritual well-being [13-16]. Second, only a few studies examined the predictors of QoL for long-term survivors. Third, QoL-related factors are known to change over time due to their multilayer and multidimensional characteristics [2], but previous models did not identify predictors as time varying. Fourth, the previous prediction models for QoL were difficult to interpret, and their overall prediction values were limited. Recently, it is possible to develop an ML algorithm that allows for interpretation [17]. Thus, this study aimed to identify patient-centered factors associated with QoL using an ML algorithm to analyze data from a cohort of Korean patients with breast cancer along different survivorship trajectories.

Methods

Study Population and Design

To produce a robust tool to identify factors associated with QoL during different survival phases, 2 different data sets were used. These included (1) the cross-sectional survey data from the Breast Cancer Information Grand Round for Survivorship (BIG-S) study to develop a model and (2) the longitudinal cohort data from the Beauty Education for Distressed Breast Cancer (BEST) cohort study to validate the model.

Development Set

The BIG-S study recruited consecutive breast cancer survivors (BCS) who visited the outpatient breast cancer clinic at the Samsung Medical Center in Seoul, Republic of Korea, between November 2018 and April 2019. The BIG-S study included survivors aged over 20 years and who did not have secondary cancer, metastasis, or recurrence. A total of 6265 survivors agreed to participate in the BIG-S study: before surgery (n=1980) and 1 year (n=653), 2 years (n=1265), 3 years (n=921), 4 years (n=682), and 5 years (n=764) after surgery.

External Validation Set

The BEST study (n=432) was conducted at 2 university-based cancer hospitals in Seoul, Republic of Korea, to evaluate the effect of cancer treatment-induced altered body image and QoL. Subjects were eligible to participate if they were between 18 and 65 years of age, had a diagnosis of breast cancer (ductal carcinoma in situ, stages I-III), had no sign of metastasis, were expected to have breast cancer surgery, and did not receive preoperative chemotherapy or radiation therapy [15]. There were 323 patients before surgery, and 297, 215, 214, and 232 patients who were followed prior to surgery and at 1, 2, 3, and 5 years following surgery, respectively.

Measures

In this study, the target variable was poor QoL, which was measured using a 7-point Likert scale with the European Organization for Research and Treatment of Cancer (EORTC) QoL Questionnaire Core 30 questionnaire. The single item has been validated to measure overall QoL [18].

To determine the factors associated with QoL, information about sociodemographics; diagnosis and treatment; and physical, psychological, social, and spiritual well-being was included based on a literature review (Table S1 in Multimedia Appendix

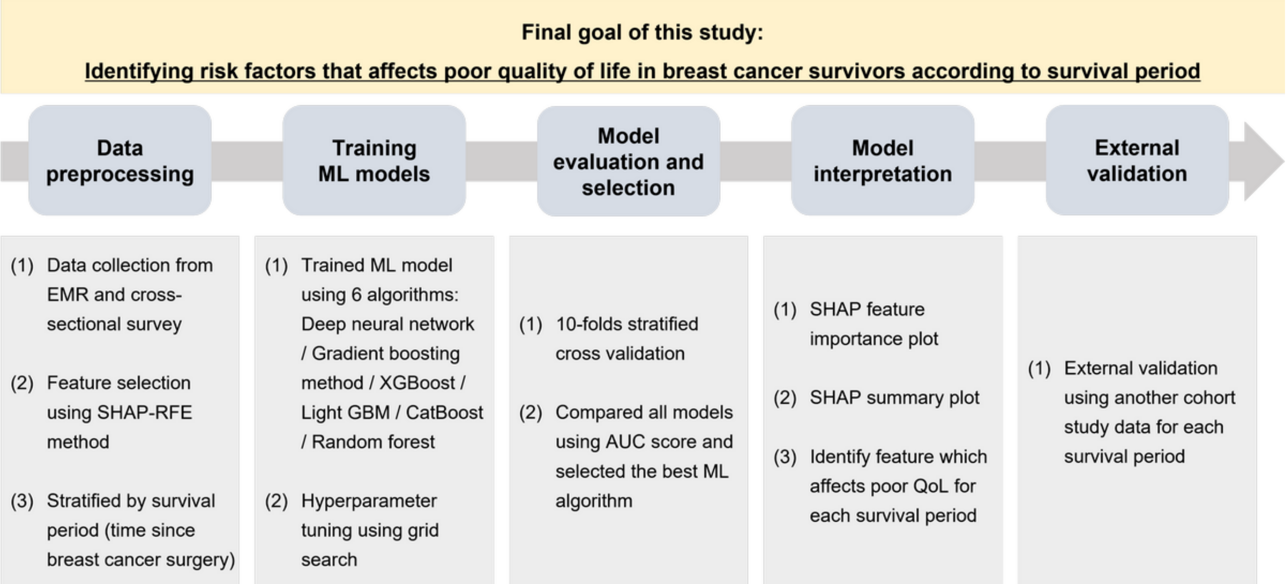
1) [5,15]. Sociodemographic factor data, including education level, marital status, monthly house income, working status during the survey, drinking status, and smoking status, were obtained using a standard questionnaire. Diagnosis and treatment data were obtained from electronic medical records. These data included types of operations, locations of tumors, comorbidities, laboratory test results, pathology stage, and type of treatment (chemotherapy, hormone therapy, target therapy, and radiotherapy).

In patient-reported outcomes, we followed the recommendation from International Consortium for Health Outcomes Measurement. To measure physical, psychological, and social well-being, the EORTC QoL Questionnaire Core 30 and Breast Cancer-Specific Module were used, and related symptoms and functions were evaluated. These included fatigue, pain, nausea and vomiting, emotional function, body image, social function, and role functioning. Spiritual well-being was evaluated using 3 questions from the Spiritual Well-being Domain of the Korean version of QoL of Cancer Survivors questionnaire [19]. In order to measure menopause symptoms, the Menopause Rating Scale (MRS) was used. The MRS included 11 items in 3 dimensions, including somatic-vegetative, psychological, and urogenital. The composite scores (score range 0-44) were based on adding the scores of the items from the respective dimensions.

Statistical Analysis

This study was conducted in five steps: (1) data preprocessing, (2) training ML models, (3) model evaluation and selection, (4) model interpretation, and (5) external validation (Figure 1). The target variable of "poor QoL" was defined as a score lower than 66 on the global health status scale (range 0-100). Factors associated with QoL were selected from the BIG-S data set. Since some of the treatment-related variables were not in the data collected prior to surgery, 37 and 45 features were selected for data preprocessing from variables collected before surgery and after surgery, respectively. In all algorithms, missing values were forward filled with the closest observation. If no past value was present, the training set mean was imputed by matching the participants' ages and pathology stages.

Figure 1. Workflow of machine learning. AUC: area under the curve; EMR: electronic media record; ML: machine learning; QoL: quality of life; SHAP: Shapley Additive Explanations; SHAP-RFE: shapley additive explanations-recursive feature elimination.



To train the ML models, the feature selection method and the recursive feature elimination method based on the Shapley Additive Explanation (SHAP) method were used to reduce the model complexity and to remove unnecessary features that generate noise in the prediction model. The SHAP method is one of the explainable artificial intelligence methods [20]. Through the Shapley values obtained using the SHAP method, how much a variable affects the outcome prediction and how the variable affects the outcome in each instance can be observed. For model evaluation and selection, we compared the performance of 6 different algorithms, including the deep neural network, gradient boosting machine, XGBoost, light gradient boosting machine, CatBoost, and random forest. For training the models, the grid search method was used for hyperparameter tuning. Hyperparameters are parameters that directly affect the learning process of the model and are determined by the user to improve model performance and avoid overfitting. After specifying the possible value range of hyperparameters for each model (Table S2 in Multimedia Appendix 1), models were trained using all possible combinations of hyperparameters, and then the optimal combinations were selected.

To validate and evaluate the model, 10-fold stratified cross-validation was used. The entire training data set was divided into 10-fold equal size subsamples by stratifying for the outcome variables. A single subsample was retained as the validation data for testing the model, and the remaining 9 subsamples were used as training data and the process was repeated 10 times. Using the 10 cross-validation results, the area under the receiver operating characteristic curve (AUC) scores were averaged for each model, and the final model with the highest mean AUC was selected. In this study, we also used SHAP values to interpret feature contributions and assess the clinical significance of predictive models. According to a previous study, the SHAP value is the measurement of the marginal contribution of each feature in different combinations.

The SHAP value of a feature can be interpreted as the difference between the model's predicted value when that feature is included versus when it is excluded, taking into account all possible combinations of other features. The base value, on the other hand, is the average predicted value of the model for all samples. When the SHAP value of a feature is positive, it means that including that feature has a positive effect on the predicted value, while a negative SHAP value indicates a negative effect. Overall, SHAP values help to explain how each feature contributes to the model's predictions, providing insight into the model's decision-making process [21,22].

Finally, external validation was confirmed using the BEST cohort data set, which was a completely different data set from that used for model training. The poor QoL group was predicted by inputting the external validation data set into the final model that trained the entire training data set using the ML algorithm selected by the survival period. It is notable that between the 3- and 4-year models, there was no validation data set because there was no participant follow-up within the BEST cohort for these time periods. The validation performances were also evaluated using AUC, accuracy, F1 score, sensitivity, and specificity, and were also compared with training performance.

All analyses were performed in the Python 3.7 programming environment (Python Software Foundation) and used the scikit-learn package and TensorFlow Keras framework.

Ethics Approval

The study was approved by the institutional review board of the Samsung Medical Center, Seoul, Republic of Korea, in the development set (SMC-2018-08-070) and external validation set (SMC-2011-07-019). Informed consent was obtained from all study participants.

Results

Characteristics of Participants

All 6265 participants were included in the full analysis data set. The mean age of the study participants was 50.6 (SD 8.66) years and 46.8% (n=2004) of participants were stage 1. In the training data set, 48.3% (n=3026) of the participants were classified into the poor QoL group (Table 1). The proportion of patients with breast cancer with poor QoL was 67.4% (n=1335) at diagnosis, and 41.8% (n=273), 39.3% (n=497), 40.1% (n=369), 36.1% (n=246), and 40.1% (n=306) patients had poor QoL at 1, 2, 3, 4, and 5 years after surgery, respectively (Table 1).

Table 1. Characteristics of participants.

Characteristics	Preoperation (n=1980)	Within 1 year (n=653)	Between 1 and 2 years	Between 2 and 3 years (n=921)	Between 3 and 4 years (n=682)	Between 4 and 5 years (n=764)	*P* value
Age (years), mean (SD)	50.54 (8.74)	49.12 (9.10)	49.78 (8.62)	50.96 (8.68)	50.29 (8.39)	52.37 (8)	<.001
BMI (kg/m^2), n (%)							.001
Underweight (<18.5)	110 (5.6)	34 (5)	56 (4)	45 (5)	25 (4)	36 (5)	
Normal (18.5-23)	985 (49.7)	352 (53.9)	675 (53.4)	506 (54.9)	392 (57.5)	416 (54.5)	
Overweight (23-25)	385 (19.4)	137 (21)	281 (22.2)	202 (21.9)	134 (19.6)	156 (20.4)	
Obese (≥25)	500 (25.3)	130 (19.9)	253 (20)	168 (18.2)	131 (19.2)	156 (20.4)	
Education, n (%)							<.001
Middle school or lower	103 (5.2)	58 (8.9)	88 (7)	68 (7.4)	65 (9.5)	82 (10.7)	
High school	324 (16.4)	217 (33.2)	445 (35.2)	315 (34.2)	223 (32.7)	271 (35.5)	
University graduates or higher	1553 (78.4)	378 (57.9)	732 (57.9)	538 (58.4)	394 (57.8)	411 (53.8)	
Working status at survey, n (%)							<.001
Working	868 (43.8)	371 (56.8)	713 (56.4)	524 (56.9)	366 (53.7)	417 (54.6)	
Not working	1112 (56.2)	282 (43.2)	552 (43.6)	397 (43.1)	316 (46.3)	347 (45.4)	
Marital status at survey, n (%)							.70
Single	185 (9.3)	61 (9.3)	101 (8)	68 (7.4)	67 (9.8)	55 (7.2)	
Married	1617 (81.7)	541 (82.8)	1043 (82.5)	764 (83)	554 (81.2)	636 (83.2)	
Divorced	118 (6)	31 (4.7)	75 (5.9)	52 (5.6)	37 (5.4)	42 (5.5)	
Bereavement	60 (3)	20 (3.1)	46 (3.6)	37 (4)	24 (3.5)	31 (4.1)	
Monthly family income (US $), n (%)							.001
≤$2000	479 (24.2)	142 (21.7)	284 (22.5)	219 (23.8)	154 (22.6)	187 (24.5)	
$2000-$4000	636 (32.1)	205 (31.4)	364 (28.8)	246 (26.7)	163 (23.9)	218 (28.5)	
>$4000	865 (43.7)	306 (46.9)	617 (48.8)	456 (49.5)	365 (53.5)	359 (47)	
Drinking status, n (%)							
Never	977 (49.3)	314 (48.1)	598 (47.3)	433 (47)	307 (45)	369 (48.3)	
Past	689 (34.8)	287 (44)	500 (39.5)	324 (35.2)	225 (33)	195 (25.5)	
Current	314 (15.9)	52 (8)	167 (13.2)	164 (17.8)	150 (22)	200 (26.2)	
Smoking status, n (%)							<.001
Never smoker	1789 (90.4)	602 (92.2)	1137 (89.9)	850 (92.3)	633 (92.8)	710 (92.9)	
Ever smoker	191 (9.6)	51 (8)	128 (10.1)	71 (7.7)	49 (7)	54 (7)	
Comorbidity (yes), n (%)	672 (33.9)	235 (36)	485 (38.3)	373 (40.5)	285 (41.8)	350 (45.8)	<.001
Physical activity (yes), n (%)	688 (34.7)	579 (88.7)	1138 (90)	802 (87.1)	601 (88.1)	665 (87)	<.001
Pathology stage, n (%)							<.001
0 or CR (NRT)	—[a]	119 (18.2)	212 (16.8)	141 (15.3)	78 (11)	68 (9)	
I	—	321 (49.2)	573 (45.3)	419 (45.5)	309 (45.3)	382 (50)	
II	—	181 (27.7)	391 (30.9)	292 (31.7)	223 (32.7)	248 (32.5)	
III or IV	—	32 (5)	89 (7)	69 (7.5)	72 (10.6)	66 (8.6)	
Type of surgery, n (%)							<.001
Mastectomy with reconstruction	—	112 (17.2)	213 (16.8)	172 (18.7)	119 (17.4)	95 (12.4)	

Characteristics	Preoperation (n=1980)	Within 1 year (n=653)	Between 1 and 2 years	Between 2 and 3 years (n=921)	Between 3 and 4 years (n=682)	Between 4 and 5 years (n=764)	*P* value
Mastectomy without reconstruction	—	112 (17.2)	198 (15.7)	136 (14.8)	150 (22)	165 (21.6)	
Breast conservation surgery	—	429 (65.7)	854 (67.5)	613 (66.6)	413 (60.6)	504 (66)	
Chemotherapy (yes), n (%)	—	160 (24.5)	419 (33.1)	365 (39.6)	305 (44.7)	389 (50.9)	<.001
Radiation therapy (yes), n (%)	—	482 (73.8)	968 (76.5)	675 (73.3)	496 (72.7)	580 (75.9)	.24
Hormone therapy (yes), n (%)	—	505 (77.3)	981 (77.5)	721 (78.3)	534 (78.3)	618 (80.9)	.44
Target therapy (yes), n (%)	—	84 (13)	186 (14.7)	140 (15.2)	98 (14.4)	116 (15.2)	.72

[a] Not available.

In the validation set, the mean age was 46.5 (SD 7.87) years, and 47.1% (n=428) of the participants were stage 1 (Table S3 in Multimedia Appendix 1). Compared to the training set, patients in the external validation set were relatively younger. Among these participants, 48.6% (n=573) were classified as having poor QoL. Patients with poor QoL prior to surgery and 1, 2, 3, and 5 years after surgery made up 70.4% (n=100), 53.2% (n=255), 49.4% (n=79), 48.5% (n=95), and 33.8% (n=72) of the groups, respectively (Table S4 in Multimedia Appendix 1).

Performances of Machine Learning Models for Each Survival Period

The available features in the training data set were used to build QoL prediction models based on 6 ML algorithms (Table S4 in Multimedia Appendix 1). From the whole data set, between 9 and 16 features were selected using the SHAP-RFE method. The AUC values of 6 ML algorithms were all over 0.75. Among 6 ML algorithms associated with the survival periods, all the final models were over 0.8 (Table S5 in Multimedia Appendix 1). The best predictive performances were observed using the CatBoost algorithm for all survival periods: overall (AUC 0.823), baseline (AUC 0.835), within 1 year (AUC 0.860), between 2 and 3 years (AUC 0.808), between 3 and 4 years (AUC 0.820), and between 4 and 5 years (AUC 0.826) (Table 2). All 5 model evaluation metric averages calculated through 10-fold stratified cross-validation for each survival period were higher than 0.7 and the AUC exceeded 0.8 (0.804-0.860), showing that the ML models performed well.

Table 2. Performance metrics by survival period.

Survival period	Overall[a]	Baseline[a]	Within 1 years[a]	Between 1 and 2 years[b]	Between 2 and 3 years[a]	Between 3 and 4 years[a]	Between 4 and 5 years[a]
AUC[c]	0.823	0.835	0.860	0.804	0.808	0.820	0.826
Accuracy	0.756	0.774	0.818	0.765	0.767	0.783	0.793
F1 score	0.707	0.817	0.782	0.705	0.709	0.723	0.752
Sensitivity	0.749	0.753	0.787	0.722	0.721	0.792	0.782
Specificity	0.761	0.815	0.839	0.793	0.797	0.777	0.801

[a] Observed using CatBoost algorithm.

[b] Observed using a gradient boosting algorithm.

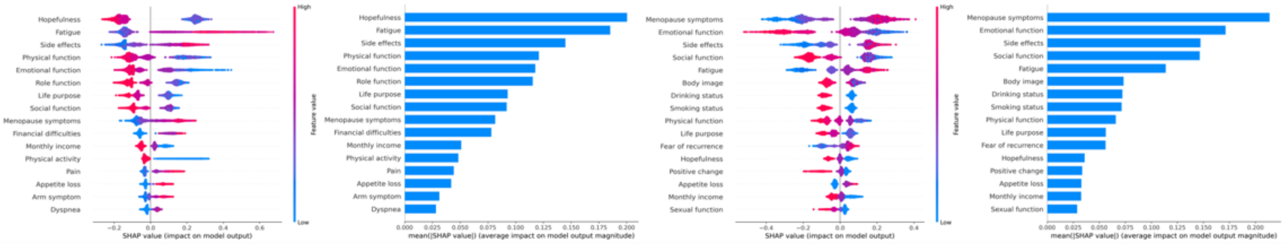
[c] AUC: area under the curve.

Important Features for Each Survival Period

The most important prognostic features for each survival period were identified using the feature importance from the SHAP method (Figure 2).

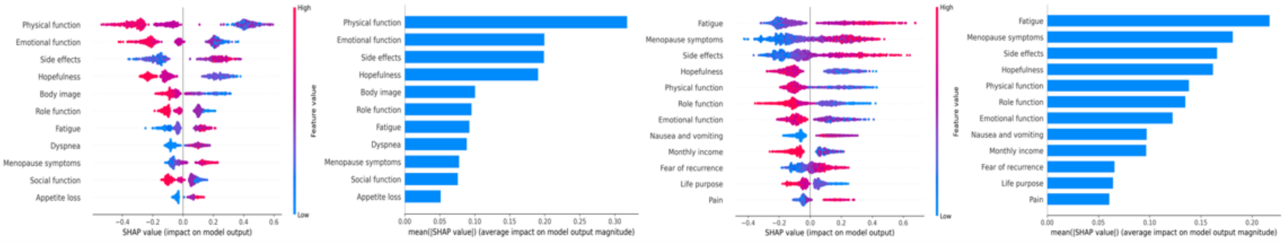
Regardless of survival period, hopefulness (SHAP value 0.2005) was the most important feature, and fatigue, side effects, physical function, emotional function, and role function were also important features. By the survival period, menopause symptoms (SHAP value 0.2137) and emotional function (SHAP value 0.1715) were the most important features prior to breast cancer surgery (Table S6 in Multimedia Appendix 1). For the within 1-year period, physical function (SHAP value 0.3177) was the most important feature, followed by emotional function, side effects, hopefulness, and body image. For the periods between 1-2, 2-3, and 3-4 years, fatigue (SHAP values 0.2172, 0.1819, and 0.1503, respectively) was the most important feature, followed by menopause symptoms, social function, and emotional function. For the period between 4 and 5 years, hopefulness (SHAP value 0.2370) was the most important feature, followed by physical function, dyspnea, financial difficulties, monthly income, menopause symptoms, side effects, and emotional function (Figure 3).

Figure 2. Summary plot of Shapley Additive Explanation.



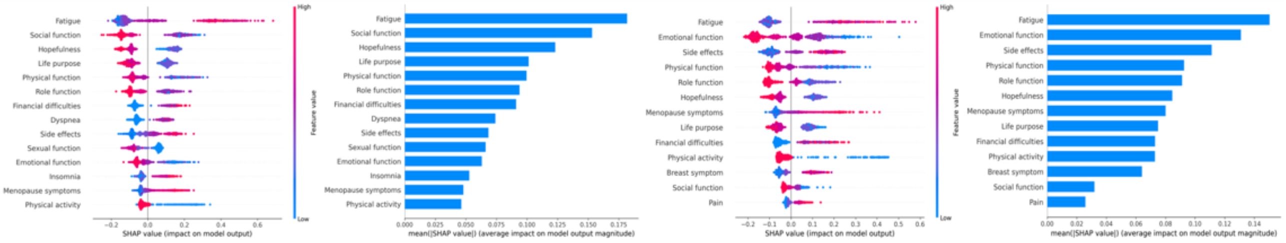
(A) Overall survival time (CatBoost)

(B) Baseline, before cancer surgery (CatBoost)



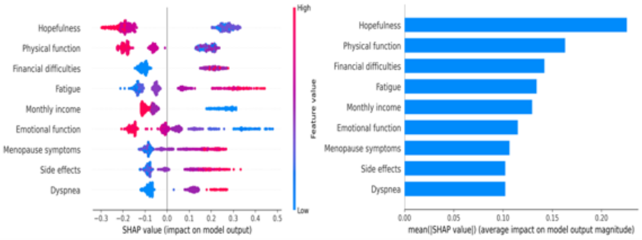
(C) Within 1 year after breast cancer surgery (CatBoost)

(D) Between 1 and 2 years after breast cancer surgery (Gradient Boosting Method)



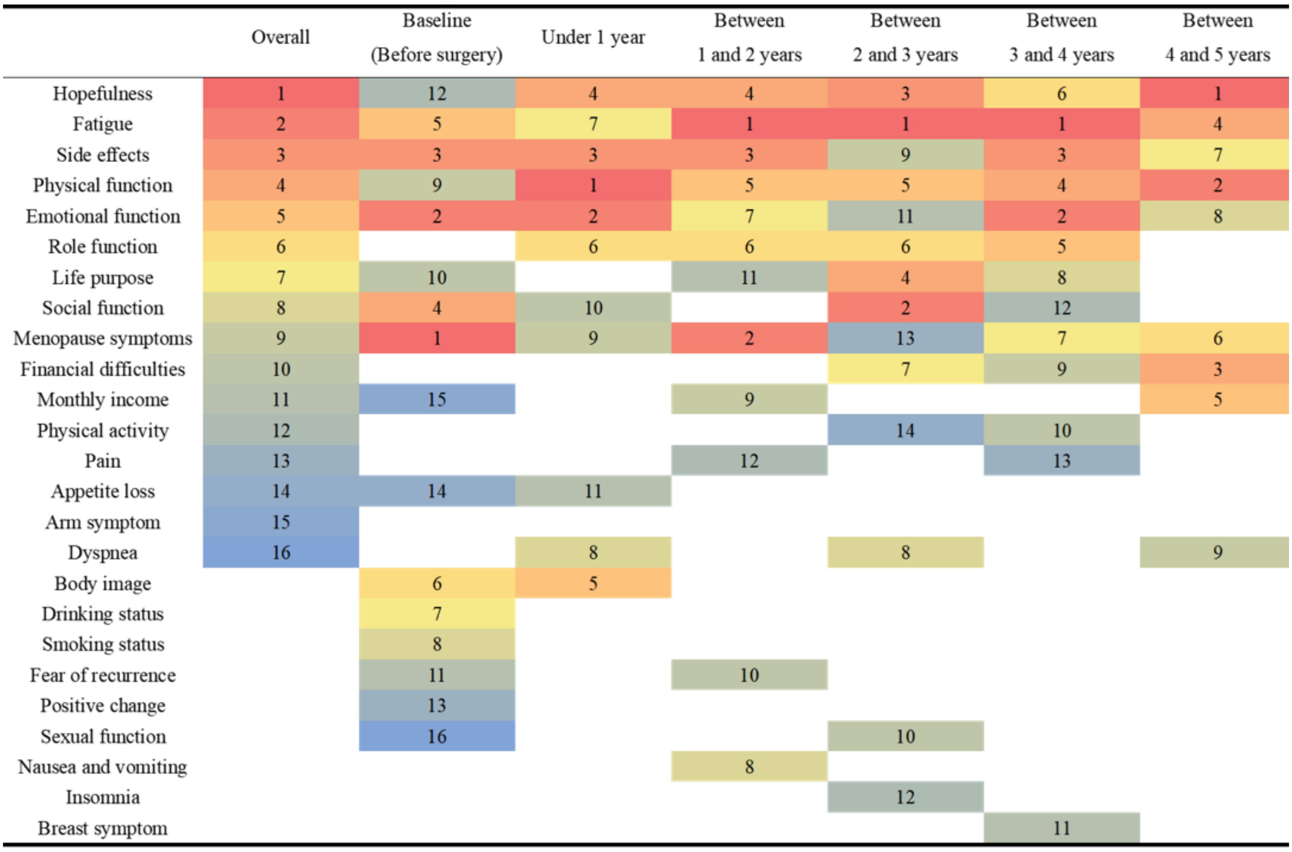
(E) Between 2 and 3 years after breast cancer surgery (CatBoost)

(F) Between 3 and 4 years after breast cancer surgery (CatBoost)



(G) Between 4 and 5 years after breast cancer surgery (CatBoost)

Figure 3. Rank of feature obtained by Shapley Additive Explanation value.



	Overall	Baseline (Before surgery)	Under 1 year	Between 1 and 2 years	Between 2 and 3 years	Between 3 and 4 years	Between 4 and 5 years
Hopefulness	1	12	4	4	3	6	1
Fatigue	2	5	7	1	1	1	4
Side effects	3	3	3	3	9	3	7
Physical function	4	9	1	5	5	4	2
Emotional function	5	2	2	7	11	2	8
Role function	6		6	6	6	5	
Life purpose	7	10		11	4	8	
Social function	8	4	10		2	12	
Menopause symptoms	9	1	9	2	13	7	6
Financial difficulties	10				7	9	3
Monthly income	11	15		9			5
Physical activity	12				14	10	
Pain	13			12		13	
Appetite loss	14	14	11				
Arm symptom	15						
Dyspnea	16		8		8		9
Body image		6	5				
Drinking status		7					
Smoking status		8					
Fear of recurrence		11		10			
Positive change		13					
Sexual function		16			10		
Nausea and vomiting				8			
Insomnia					12		
Breast symptom						11	

External Validation

All 5 model evaluation metric averages calculated in the external validation set were higher than 0.7 (Table 3).

For external validation, the receiver operating characteristic curves for each survival period were used to calculate the AUC.

When the final trained models for each survival period were externally validated using the BEST data set, the validation AUC was between 0.770 and 0.862, and the differences from the results of the 10-fold stratified cross-validation for training were from 0.009 to 0.056 (Figure 4).

Table 3. Performance metrics for external validation for survival period.

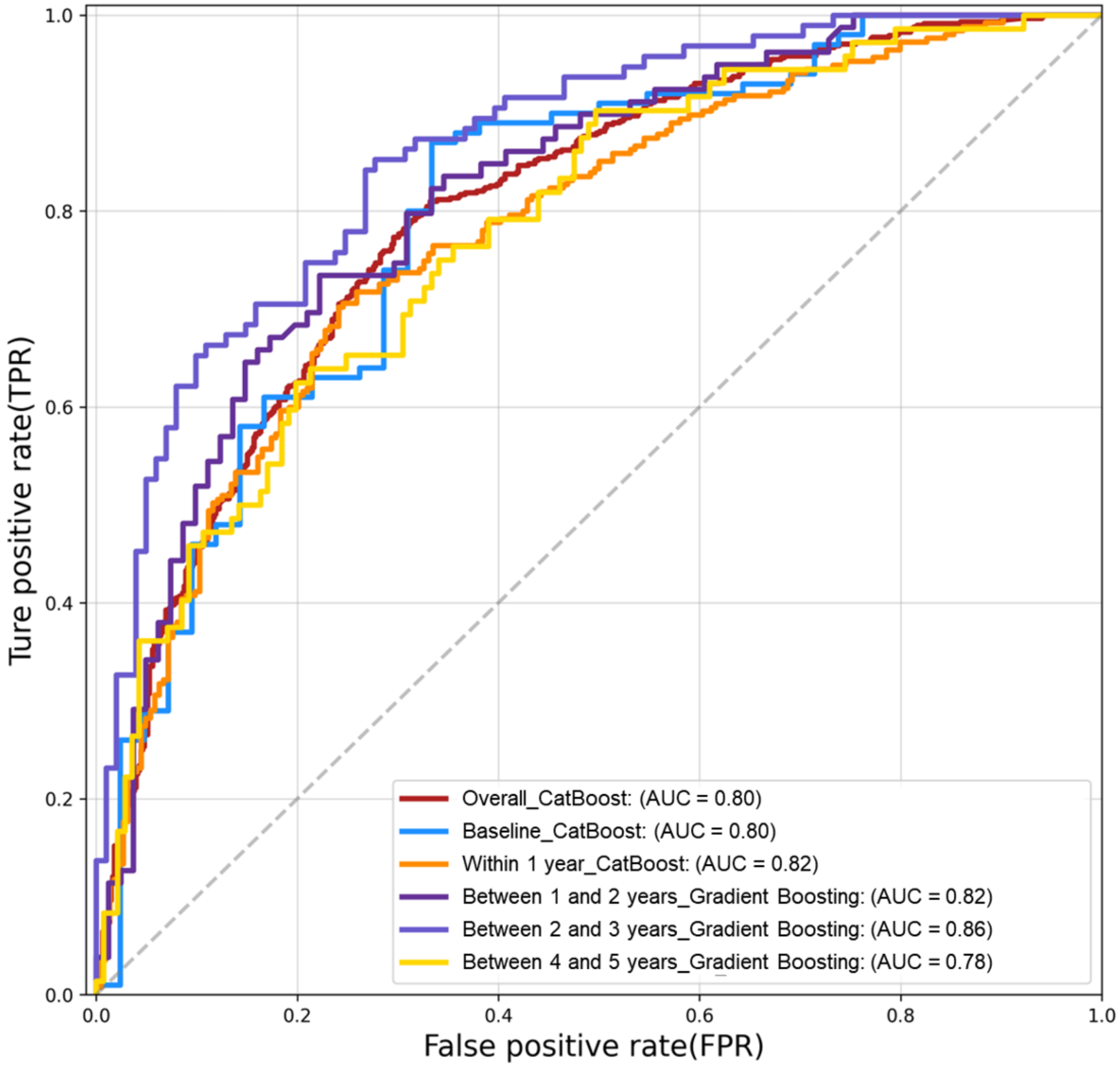
Survival period	Overall[a]	Baseline[a]	Within 1 years[a]	Between 1 and 2 years[b]	Between 2 and 3 years[a]	Between 4 and 5 years[a]
AUC[c]	0.800	0.799	0.778	0.816	0.863	0.779
Accuracy	0.810	0.737	0.729	0.756	0.786	0.742
F1 score	0.866	0.750	0.734	0.748	0.794	0.621
Sensitivity	0.870	0.810	0.702	0.734	0.853	0.625
Specificity	0.667	0.668	0.759	0.778	0.723	0.801

[a] Observed using CatBoost algorithm.

[b] Observed using a gradient boosting algorithm.

[c] AUC: area under the curve.

Figure 4. Receiver operating characteristics curve. AUC: area under the curve.



Discussion

Principal Findings

In this study, we developed and validated factors associated with QoL including physical, psychological, economic, spiritual, and social dimensions by survivorship trajectory using an ML algorithm. The developed model and external validation model performance were good for all survival trajectories. Before surgery, menopause symptoms and emotional function were important features. Within 1 year after the surgery period, physical function was the most important feature. Between 1 and 4 years, fatigue was the most important feature. Regardless of the survival period, hopefulness was the most influential feature of spiritual well-being.

In this study, the AUC for evaluating model performances surpassed 0.8 for all survival periods, and the results of external validation using data collected in other studies were also greater than 0.77. This performance is much better than that of previous studies that predicted QoL using ML modeling, which reported values ranging from 0.476 to 0.793 [6,9-12]. These ML-based breast cancer QoL prediction models were developed with not only clinical and sociodemographic factors but also with the integration of information from multiple factors, thus ensuring better model performance. Furthermore, this study stratified the model by time periods following surgery and found that there were different factors associated with QoL during each time period.

Prior to breast cancer surgery, menopause symptoms and emotional function were selected as important features that affect QoL in BCS. Among menopause symptoms, the most important factor was in the psychological domain, which included depressive mood, irritability, and anxiety. According to a previous study, depression and anxiety are the 2 most common psychiatric comorbidities encountered in patients with breast cancer [6,7]. Patients with breast cancer may experience depression or anxiety at any stage of their illness, from prediagnosis to the terminal phase of the illness. Studies in Western countries have shown that the prevalence of depression ranges from 1% to 56%, whereas the prevalence of depression found in Asian studies was between 12.5% and 31% [23]. Thus, timely psychosocial care should be needed for newly diagnosed distressed patients with cancer.

In this study, physical function was the most important feature that affected QoL in the group that was within 1 year after breast cancer surgery. This result was consistent with previous reports that BCS are susceptible to physical functioning–related problems and often experience treatment-related declines in their physical functioning capabilities within the 1-year period following their cancer diagnosis [24]. Treatment-related systemic side effects that occur after completion of treatment affect physical function, and poor physical function negatively affects the QoL of BCS [25]. Furthermore, physical functioning–related problems may persist even after treatment has been completed [26]. Persistent physical symptoms can increase fatigue and hinder patients' return to normal life, thereby reducing their QoL. Future research should focus on the development and testing of interventions for managing physical function in order to improve the QoL of patients with breast cancer.

Between 1 and 4 years after breast cancer surgery, fatigue is the most important feature that affects QoL in BCS. Cancer-related fatigue is one of the most distressing and common posttreatment sequelae among survivors of early-stage breast cancer [27]. More than 30% of patients with breast cancer experience persistent fatigue symptomatology up to 10 years after completion of treatment [28]. Cancer-related fatigue can result in substantial adverse physical, psychosocial, and socioeconomic consequences and has a negative impact on overall QoL. For BCS 1 year after diagnosis, reducing the burden of fatigue might be a preferable approach to improve their QoL and focusing on fatigue symptoms can help to enhance the long-term survivors' QoL [29]. Cancer-related fatigue is considered a complex symptom, with multidimensional and intricate aspects. The existence of physical, psychological, and emotional disturbance has been proven [30], and numerous evidence-based interventions for the management of fatigue have been recommended [31], most of them being complex nonpharmacological interventions. In order to address all dimensions of fatigue, nonpharmacological interventions should be tested and assessed.

Hopefulness was the most important feature in all survival period models, especially, between 4 and 5 years. Spiritual well-being was a predictor of improved QoL and is one of the important outcomes to measure in BCS [32]. Previous studies have indicated that survivors who had more hope in their lives were more likely to have better QoL [33]. Hope could help patients find a sense of health in the midst of disease to cope with various cancer symptoms and fear of recurrence and to find meaning and peace of mind [34]. These positive effects of hope might also improve QoL in BCS. Therefore, patient-centered interventions that help survivors find purpose in life by focusing on themes such as planning for life after cancer and value-based sources of meaning to have hope should be provided.

In this study, we performed an external validation to test the generalizability of our models, which is a strength compared to the previous study that did not perform external validation [6,9-12]. This aspect is important as it demonstrates the effectiveness of our models and their potential to be applied to other settings. Through external validation, we could assess our models' robustness and confirm their ability to provide accurate predictions in new and independent data sets. This enhances the reliability and use of our models, and highlights the potential of ML approaches in improving survivorship care for patients with breast cancer.

This study has several limitations. First, it is a cross-sectional study, and the directions of the associations between QoL; symptoms; and physical, psychosocial, and spiritual functions could be interchangeable. In fact, patients who had a poor QoL might report poorer function status. Second, QoL was measured using a single item from the EORTC-C30, and this might not be a reliable method to measure an individual's QoL. However, this single question has been validated to measure a person's overall QoL, and it has been widely used in different cultures and countries, including Korea. Lastly, the results of our study might not be generalizable to other cancer survivors in other settings. Further studies with various types of cancer survivors are necessary to confirm the study findings and its generalizability.

Despite these limitations, this study had several strengths. First, we included physical, psychological, economic, spiritual, and social dimensions and clinical factors. Second, we developed a prediction model to predict QoL from pretreatment to 5 years after surgery. Third, we developed different ML-based QoL surveillance models across survivorship. Fourth, we used SHAP methods, which allow for the interpretation of the model by the reader. Fifth, we performed external validation and the models showed good performance.

Conclusions

The results of this study may provide valuable information on developing a patient-centered survival care plan. Understanding the changing trends of influencing factors associated with QoL during different survival trajectories could help health care professionals intervene timely and appropriately in order to prevent or alleviate factors more precisely.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, Information and Communications Technology, and Future Planning (2017R1E1A1A0107764214), Future Medicine 20×30 Project of the Samsung Medical Center (SMX1210831), Amorepacific Corporation, and the Korea Breast Cancer Foundation.

Data Availability

The data supporting this study’s findings are available on request from the corresponding author (SKL). The data are not publicly available because they contain information that could compromise research participant consent.

Authors' Contributions

DK, JC, and SKL conceived and designed the study; JEL, SJN, SWK, JY, BJC, and SKL constructed the data; HSK, DK, JC, ZK, and MJC contributed toward analysis; DK, HSK, and JC wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 46 KB - publichealth_v9i1e45212_app1.docx]

References

1. Breast cancer. World Health Organization. 2018. URL: https://www.who.int/news-room/fact-sheets/detail/breast-cancer [accessed 2023-06-20]
2. Nardin S, Mora E, Varughese FM, D'Avanzo F, Vachanaram AR, Rossi V, et al. Breast cancer survivorship, quality of life, and late toxicities. Front Oncol 2020;10:864-864 [FREE Full text] [doi: 10.3389/fonc.2020.00864] [Medline: 32612947]
3. Hewitt M, Ganz PA, National Research Council, Institute of Medicine, American Society of Clinical Oncology. In: Hewitt M, Ganz PA, editors. From Cancer Patient to Cancer Survivor: Lost in Transition: An American Society of Clinical Oncology and Institute of Medicine Symposium. Washington, DC: National Academies Press; 2006.
4. Nekhlyudov L, Mollica MA, Jacobsen PB, Mayer DK, Shulman LN, Geiger AM. Developing a quality of cancer survivorship care framework: implications for clinical care, research, and policy. J Natl Cancer Inst 2019;111(11):1120-1130 [FREE Full text] [doi: 10.1093/jnci/djz089] [Medline: 31095326]
5. Ong WL, Schouwenburg MG, van Bommel ACM, Stowell C, Allison KH, Benn KE, et al. A standard set of value-based patient-centered outcomes for breast cancer: the international consortium for health outcomes measurement (ICHOM) initiative. JAMA Oncol 2017;3(5):677-685. [doi: 10.1001/jamaoncol.2016.4851] [Medline: 28033439]
6. Carmona-Bayonas A, Calderón C, Hernández R, Montes AF, Castelo B, Ciria-Suarez L, et al. Prediction of quality of life in early breast cancer upon completion of adjuvant chemotherapy. NPJ Breast Cancer 2021;7(1):92 [FREE Full text] [doi: 10.1038/s41523-021-00296-8] [Medline: 34257315]
7. Di Meglio A, Havas J, Gbenou AS, Martin E, El-Mouhebb M, Pistilli B, et al. Dynamics of long-term patient-reported quality of life and health behaviors after adjuvant breast cancer chemotherapy. J Clin Oncol 2022;40(27):3190-3204 [FREE Full text] [doi: 10.1200/JCO.21.00277] [Medline: 35446677]
8. Hueting TA, van Maaren MC, Hendriks MP, Koffijberg H, Siesling S. The majority of 922 prediction models supporting breast cancer decision-making are at high risk of bias. J Clin Epidemiol 2022;152:238-247 [FREE Full text] [doi: 10.1016/j.jclinepi.2022.10.016] [Medline: 36633901]
9. Shi HY, Tsai JT, Chen YM, Culbertson R, Chang HT, Hou MF. Predicting two-year quality of life after breast cancer surgery using artificial neural network and linear regression models. Breast Cancer Res Treat 2012;135(1):221-229. [doi: 10.1007/s10549-012-2174-6] [Medline: 22836876]
10. Tsai JT, Hou MF, Chen YM, Wan TTH, Kao HY, Shi HY. Predicting quality of life after breast cancer surgery using ANN-based models: performance comparison with MR. Support Care Cancer 2013;21(5):1341-1350. [doi: 10.1007/s00520-012-1672-8] [Medline: 23203653]
11. Savić M, Kurbalija V, Ilić M, Ivanović M, Jakovetić D, Valachis A, et al. Analysis of machine learning models predicting quality of life for cancer patients. In: Proceedings of the 13th International Conference on Management of Digital EcoSystems. 2021 Nov Presented at: MEDES '21; November 1-3, 2021; Virtual Event, Tunisia p. 35-42. [doi: 10.1145/3444757.3485103]
12. Nuutinen M, Korhonen S, Hiltunen AM, Haavisto I, Poikonen-Saksela P, Mattson J, et al. Impact of machine learning assistance on the quality of life prediction for breast cancer patients. In: Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5: BIOSTEC. 2022 Presented at: BIOSTEC 2022; February 9-11, 2022; Vienna, Austria p. 344-352. [doi: 10.5220/0010786900003123]

13. Shen A, Qiang W, Wang Y, Chen Y. Quality of life among breast cancer survivors with triple negative breast cancer: role of hope, self-efficacy and social support. Eur J Oncol Nurs 2020;46:101771. [doi: 10.1016/j.ejon.2020.101771] [Medline: 32506010]
14. Ho PJ, Gernaat SAM, Hartman M, Verkooijen HM. Health-related quality of life in Asian patients with breast cancer: a systematic review. BMJ Open 2018;8(4):e020512 [FREE Full text] [doi: 10.1136/bmjopen-2017-020512] [Medline: 29678980]
15. Kang D, Kim IR, Choi EK, Yoon JH, Lee SK, Lee JE, et al. Who are happy survivors? Physical, psychosocial, and spiritual factors associated with happiness of breast cancer survivors during the transition from cancer patient to survivor. Psychooncology 2017;26(11):1922-1928. [doi: 10.1002/pon.4408] [Medline: 28233366]
16. Kang D, Kim N, Han G, Kim S, Kim H, Lim J, et al. Divorce after breast cancer diagnosis and its impact on quality of life. Palliat Support Care 2022;20(6):807-812. [doi: 10.1017/S1478951521001711] [Medline: 36942579]
17. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. Entropy (Basel) 2020;23(1):18 [FREE Full text] [doi: 10.3390/e23010018] [Medline: 33375658]
18. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 03;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
19. Cho J, Kang D, Kim IR, Kim WS, Ferrell B, Kim SJ. Validation of the Korean version of the quality of life-cancer survivors (QOL-CS-K) questionnaire in lymphoma survivors. Cancer Res Treat 2018;50(1):204-211 [FREE Full text] [doi: 10.4143/crt.2017.091] [Medline: 28361524]
20. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30. 2017 Presented at: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017); 4-9 December 2017; Long Beach, California, USA p. 4766.
21. Loh HW, Ooi CP, Seoni S, Barua PD, Molinari F, Acharya UR. Application of explainable artificial intelligence for healthcare: a systematic review of the last decade (2011-2022). Comput Methods Programs Biomed 2022;226:107161. [doi: 10.1016/j.cmpb.2022.107161] [Medline: 36228495]
22. Wang X, Yang F, Zhu M, Cui H, Wei J, Li J, et al. Development and assessment of assisted diagnosis models using machine learning for identifying elderly patients with malnutrition: cohort study. J Med Internet Res 2023;25:e42435 [FREE Full text] [doi: 10.2196/42435] [Medline: 36917167]
23. Zainal NZ, Nik-Jaafar NR, Baharudin A, Sabki ZA, Ng CG. Prevalence of depression in breast cancer survivors: a systematic review of observational studies. Asian Pac J Cancer Prev 2013;14(4):2649-2656 [FREE Full text] [doi: 10.7314/apjcp.2013.14.4.2649] [Medline: 23725190]
24. Ganz PA, Kwan L, Stanton AL, Krupnick JL, Rowland JH, Meyerowitz BE, et al. Quality of life at the end of primary treatment of breast cancer: first results from the moving beyond cancer randomized trial. J Natl Cancer Inst 2004;96(5):376-387 [FREE Full text] [doi: 10.1093/jnci/djh060] [Medline: 14996859]
25. Park JH, Jun EY, Kang MY, Joung YS, Kim GS. Symptom experience and quality of life in breast cancer survivors. J Korean Acad Nurs 2009;39(5):613-621. [doi: 10.4040/jkan.2009.39.5.613] [Medline: 19901491]
26. Ganz PA, Kwan L, Stanton AL, Bower JE, Belin TR. Physical and psychosocial recovery in the year after primary treatment of breast cancer. J Clin Oncol 2011;29(9):1101-1109 [FREE Full text] [doi: 10.1200/JCO.2010.28.8043] [Medline: 21300931]
27. Runowicz CD, Leach CR, Henry NL, Henry KS, Mackey HT, Cowens-Alvarado RL, et al. American Cancer Society/American Society of Clinical Oncology breast cancer survivorship care guideline. CA Cancer J Clin 2016;66(1):43-73 [FREE Full text] [doi: 10.3322/caac.21319] [Medline: 26641959]
28. Abrahams HJG, Gielissen MFM, Schmits IC, Verhagen CAHHVM, Rovers MM, Knoop H. Risk factors, prevalence, and course of severe fatigue after breast cancer treatment: a meta-analysis involving 12 327 breast cancer survivors. Ann Oncol 2016;27(6):965-974 [FREE Full text] [doi: 10.1093/annonc/mdw099] [Medline: 26940687]
29. Xia J, Tang Z, Deng Q, Yang R, Wang J, Yu J. Predictors of the quality of life in Chinese breast cancer survivors. Breast Cancer Res Treat 2018;167(2):537-545. [doi: 10.1007/s10549-017-4512-1] [Medline: 28965272]
30. Curt GA, Breitbart W, Cella D, Groopman JE, Horning SJ, Itri LM, et al. Impact of cancer-related fatigue on the lives of patients: new findings from the fatigue coalition. Oncologist 2000;5(5):353-360 [FREE Full text] [doi: 10.1634/theoncologist.5-5-353] [Medline: 11040270]
31. Berger AM, Mooney K, Alvarez-Perez A, Breitbart WS, Carpenter KM, Cella D, National Comprehensive Cancer Network. Cancer-related fatigue, version 2.2015. J Natl Compr Canc Netw 2015;13(8):1012-1039 [FREE Full text] [doi: 10.6004/jnccn.2015.0122] [Medline: 26285247]
32. Pahlevan Sharif S, Lehto RH, Amiri M, Ahadzadeh AS, Sharif Nia H, Haghdoost AA, et al. Spirituality and quality of life in women with breast cancer: the role of hope and educational attainment. Palliat Support Care 2021;19(1):55-61. [doi: 10.1017/S1478951520000383] [Medline: 32580795]
33. McClain CS, Rosenfeld B, Breitbart W. Effect of spiritual well-being on end-of-life despair in terminally-ill cancer patients. Lancet 2003;361(9369):1603-1607. [doi: 10.1016/S0140-6736(03)13310-7] [Medline: 12747880]

34. Rustøen T, Cooper BA, Miaskowski C. The importance of hope as a mediator of psychological distress and life satisfaction in a community sample of cancer patients. Cancer Nurs 2010;33(4):258-267. [doi: 10.1097/NCC.0b013e3181d6fb61] [Medline: 20467303]

Abbreviations

AUC: area under the curve
BCS: breast cancer survivor
BEST: Beauty Education for Distressed Breast Cancer
BIG-S: Breast Cancer Information Grand Round for Survivorship
EORTC: European Organization for Research and Treatment of Cancer
ML: machine learning
MRS: Menopause Rating Scale
QoL: quality of life
SHAP: Shapley Additive Explanation

Edited by A Mavragani, T Sanchez; submitted 20.12.22; peer-reviewed by N Jiwani, O Ijezie; comments to author 20.04.23; revised version received 02.05.23; accepted 13.06.23; published 24.08.23.

Please cite as:
Kang D, Kim H, Cho J, Kim Z, Chung M, Lee JE, Nam SJ, Kim SW, Yu J, Chae BJ, Ryu JM, Lee SK
Prediction Model for Postoperative Quality of Life Among Breast Cancer Survivors Along the Survivorship Trajectory From Pretreatment to 5 Years: Machine Learning–Based Analysis
JMIR Public Health Surveill 2023;9:e45212
URL: https://publichealth.jmir.org/2023/1/e45212
doi:10.2196/45212
PMID:37309655

Original Paper

Trends in Opioid Medication Adherence During the COVID-19 Pandemic: Retrospective Cohort Study

Amir Marashi[1], PhD; David Warren[1], MRes; Gary Call[2], MD; Mark Dras[1], PhD

[1]School of Computing, Macquarie University, Macquarie Park, Australia
[2]Gainwell Technologies, Tysons, VA, United States

Corresponding Author:
Mark Dras, PhD
School of Computing
Macquarie University
4 Research Park Drive
Macquarie Park, 2109
Australia
Phone: 61 9850 9580
Email: mark.dras@mq.edu.au

Abstract

Background: The recent pandemic had the potential to worsen the opioid crisis through multiple effects on patients' lives, such as the disruption of care. In particular, good levels of adherence with respect to medication for opioid use disorder (MOUD), recognized as being important for positive outcomes, may be disrupted.

Objective: This study aimed to investigate whether patients on MOUD experienced a drop in medication adherence during the recent COVID-19 pandemic.

Methods: This retrospective cohort study used Medicaid claims data from 6 US states from 2018 until the start of 2021. We compared medication adherence for people on MOUD before and after the beginning of the COVID-19 pandemic in March 2020. Our main measure was the proportion of days covered (PDC), a score that measures patients' adherence to their MOUD. We carried out a breakpoint analysis on PDC, followed by a patient-level beta regression analysis with PDC as the dependent variable while controlling for a set of covariates.

Results: A total of 79,991 PDC scores were calculated for 37,604 patients (age: mean 37.6, SD 9.8 years; sex: n=17,825, 47.4% female) between 2018 and 2021. The coefficient for the effect of COVID-19 on PDC score was –0.076 and was statistically significant (odds ratio 0.925, 95% CI 0.90-0.94).

Conclusions: The COVID-19 pandemic was negatively associated with patients' adherence to their medication, which had declined since the beginning of the pandemic.

(JMIR Public Health Surveill 2023;9:e42495) doi:10.2196/42495

KEYWORDS

COVID-19; opioid crisis; opioids; medication for opioid use disorder; MOUD; pandemic; public health; opioid; medication; treatment; care; patient; opioid use disorder; beta regression analysis; breakpoint analysis

Introduction

The opioid crisis is an ongoing problem in the United States. Recent reports and studies have shown an increase in the number of overdoses in 2020, suggesting an association between the COVID-19 pandemic and the worsening of the opioid crisis [1-3]. Disruptions in medical services and care due to lockdowns, stay-at-home orders [4], and the scarcity of health care services might be some of the factors involved in this association. Patient adherence to medications is an important factor in the success of medication for opioid use disorder (MOUD), and higher adherence is associated with a reduction in mortality rate [5]. Pharmacists, health care experts, and public and private health technology companies have expressed their concerns about medication adherence amid the pandemic, both in the United States [6,7] and internationally [8-11]. Increases in emergency medical services responses, including the administration of naloxone [12], suggest that these concerns are well founded.

A recent study [13] examined the changes in the prescribing of opioid analgesics and buprenorphine for opioid use disorder (OUD) throughout the COVID-19 pandemic. By looking at aggregate prescription levels, it found that existing patients who receive opioid analgesics and buprenorphine for OUD generally maintained access to these medications during the COVID-19 pandemic, whereas the initiation of buprenorphine remained at a low rate through August 2020. Another study investigating patients' access to chronic medications during the COVID-19 pandemic used a dichotomous measurement for patients' adherence; it found that since the beginning of the pandemic, patients were less likely to discontinue OUD therapy (buprenorphine or naloxone) [14]. However, both aggregate analyses and studies similar to Clement et al [14] that use dichotomous variables can conceal more fine-grained, individual-level changes.

In this study, we examined the trend of medication adherence for US Medicaid patients with an OUD diagnosis from 6 US states who were on opioid medications from January 2018 to March 2021. We calculated the proportion of days covered (PDC) as a method of measuring medication adherence for each patient in the 3 months following the first visit for an opioid-related disorder or opioid dependence indication, according to International Classification of Diseases, 10th revision (ICD-10) codes, and calculated the average monthly PDC score in each month. Then, we examined the adherence trend across time and investigated the association between the pandemic and PDC score, as well as the association between the pandemic and opioid overdose in a 3-month time window.

Methods

Overview

We analyzed US Medicaid claims data for 6 US states. The data set is provided by HMS (a Gainwell Technologies company) through the Digital Health Cooperative Research Centre and comprises Medicaid claims (inpatient, outpatient, dental, and pharmaceutical), a few demographic characteristics of the patients (age, sex, zip code, and Medicaid enrollment periods), and their unique identifiers. The data set links their claims across time, different settings of care, and multiple episodes of enrollment. Diagnosis codes associated with the patient, procedures codes, prescription date, dosage, duration, and other prescription details are among the information available in the data set. The data set covers claims from 2015 to 2021 and comprises over 1.16 billion claims for over 14.2 million unique individuals.

We used a continuous measurement of adherence called PDC as a proxy for medication adherence. We defined PDC as the ratio of the number of days a patient could have their opioid treatment medication to the number of days in the period of interest.

We created a subset of patients with at least one relevant F11 diagnosis ICD-10 code (codes F11.1 opioid abuse; F11.2 opioid dependence; and F11.9 opioid use, unspecified), who have filled their opioid treatment prescriptions in the 3-month window following the first visit with an F11 diagnosis code. We excluded the patients whose first F11 visit happened before 2018. For each patient with at least three months of opioid treatment medication, we calculated and assigned 3 monthly PDC scores for each month following their medical visit. To investigate the association between the COVID-19 pandemic and medication adherence, we took 2 separate approaches.

First, in an exploratory analysis, we plotted the average PDC score for all patients in each month to visualize any changes in the trend of the average PDC score since the beginning of the pandemic. To detect whether such changes exist and to avoid subjective judgements, we used an algorithm described by Bai and Perron [15] for simultaneous estimation of multiple breakpoints implemented in the *strucchange* R package [16]. To define how many breakpoints should be considered, we calculated the Bayesian information criterion (BIC) estimates for different numbers of breakpoints and chose the number that minimizes the BIC. The algorithm detects the time(s) where possible breakpoints occurred and the CIs for the detected times.

In the second approach, we applied a patient-level beta regression model to examine the effect of the COVID-19 pandemic (modelled as a binary variable with a value equal to 1 for months after March 2020) on the PDC score. We controlled for a set of other covariates, including the age, sex, state, and comorbidities of each patient. To reduce the dimensionality of comorbidities, we categorized diagnosis codes into Elixhauser general comorbidity categories [17] according to the Agency for Health Research Quality rules using the *icd* R package by Wasey and Lang [18]. The Elixhauser categories condense many dozens of ICD-10 (and other) administrative codes into 30 broader groupings (as termed by Elixhauser et al [17]), such as "Alcohol abuse" (ICD-10 codes F10, E52, etc) or "Drug abuse" (ICD-10 codes F11.x-F16.x, etc). In a separate model, we repeated the analysis with data inclusive of the patients with only 1 or 2 months of PDC scores to see the sensitivity of the model.

Ethical Considerations

The data were provided in deidentified form by HMS. On this basis, an ethical review by Macquarie University deemed it as having negligible risk, in line with Australia's National Statement on Ethical Conduct in Human Research 2007 (as updated in 2018).

Results

We calculated a total of 79,991 PDC scores for 37,604 patients in our subset OUD sample (age: mean 37.6, SD 9.8 years; sex: n=17,825, 47.4% female) between 2018 and 2021 from Medicaid claims data of 6 US states. Overall, 7816 patients who did not have a PDC score in the month following their F11 diagnosis or had a month without a PDC score between their first and second scores were excluded from the analysis. From the remaining patients, the most common treatment approach involved buprenorphine prescription (24,731/29,788, 83.02%), followed by prescriptions for naltrexone alone (3707/29,788, 12.44%); only small proportions were prescribed buprenorphine and naltrexone together (778/,29,788, 2.61%) or methadone (572/29,788, 1.92%). In terms of prescription filling patterns, only 24.9% (7417/29,788) filled their MOUD prescriptions in the month following their F11 diagnosis (mean PDC score 79.6), another 24.9% (7417/29,788) filled 2 months of MOUD prescriptions (mean PDC score 91.7), and 50.2% (14,954/29,788) filled 3 or more months of MOUD prescriptions (mean PDC score 96.1). To understand whether any drop in adherence as measured by PDC was related to patients discontinuing versus continuing OUD therapy, we separated out patients based on how many months were covered. As Figure 1 shows, the PDC score was lower for patients with fewer covered months. To focus on the change during the COVID-19 period for patients that were more adherent to their medication, for the main part of the analysis, we excluded the patients with less than 3 months of PDC scores.

Figure 2 shows the result of applying the breakpoint detection algorithm on the time series of the PDC scores for the selected patients. The shaded area denotes the time after the pandemic was announced (March 2020). The dash line represents the detected breakpoint by the algorithm in July 2020, and the red line represents the CI from April to August.

The results of the beta regression analysis with a logit link for features with statistically significant coefficients are reported in Table 1. The dependent variable is the calculated PDC score between 0 and 1 for each patient in each month. The independent variable of interest, named "COVID," is a binary variable that is 0 for PDC scores at months before March 2020 and 1 for PDC scores from March 2020 onward. The control variables in the model included age group, sex, the month from which the PDC is calculated (PDC first month, PDC second month, and PDC third month), state identifiers, and Elixhauser comorbidity groups calculated based on the clinical history of each patient. Although the independent variables in the model did not explain much of the variance of the PDC score (R^2 the model is 0.055), the coefficient for the effect of "COVID" was –0.76 and was statistically significant (odds ratio 0.925, 95% CI 0.90-0.94). This result is in line with findings of the exploratory analysis and breakpoint detection. The model inclusive of the patients with only 1 or 2 PDC scores showed similar results. The results showed that male patients had lower PDC scores (OR 0.97, 95% CI 0.96-0.99), whereas patients in all age groups older than the reference category (aged 17-29 years) had higher PDC scores. A history of obesity and alcohol use disorder were positively associated with the PDC score.

Figure 1. Average proportion of days covered (PDC) over time, a comparison of patients with 1, 2, and 3 months of PDC scores.

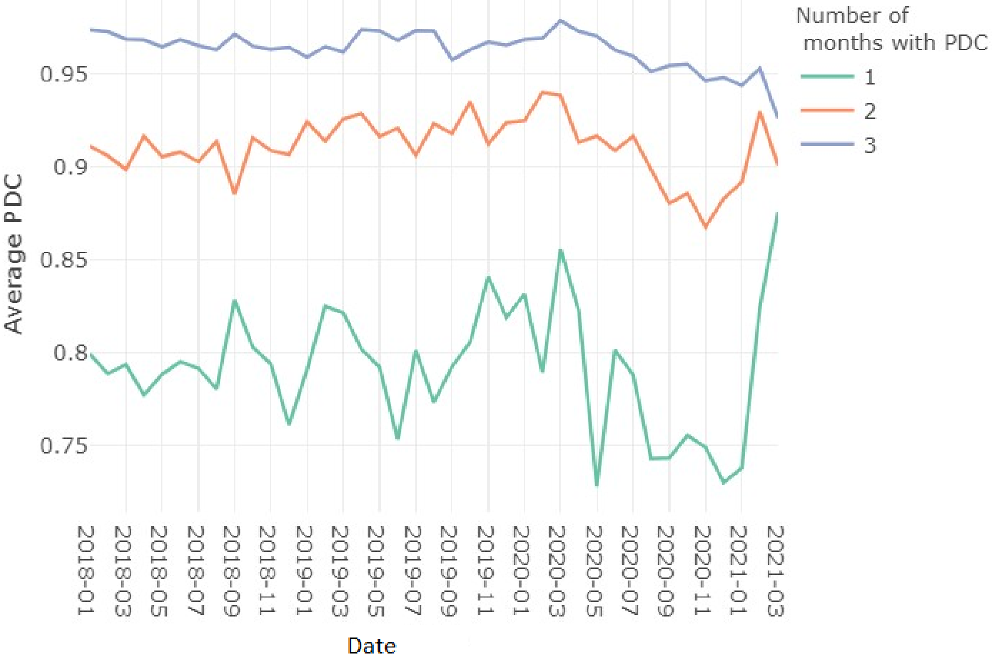


Figure 2. Detected breakpoints for patients with 3 months of proportion of days covered (PDC) scores.

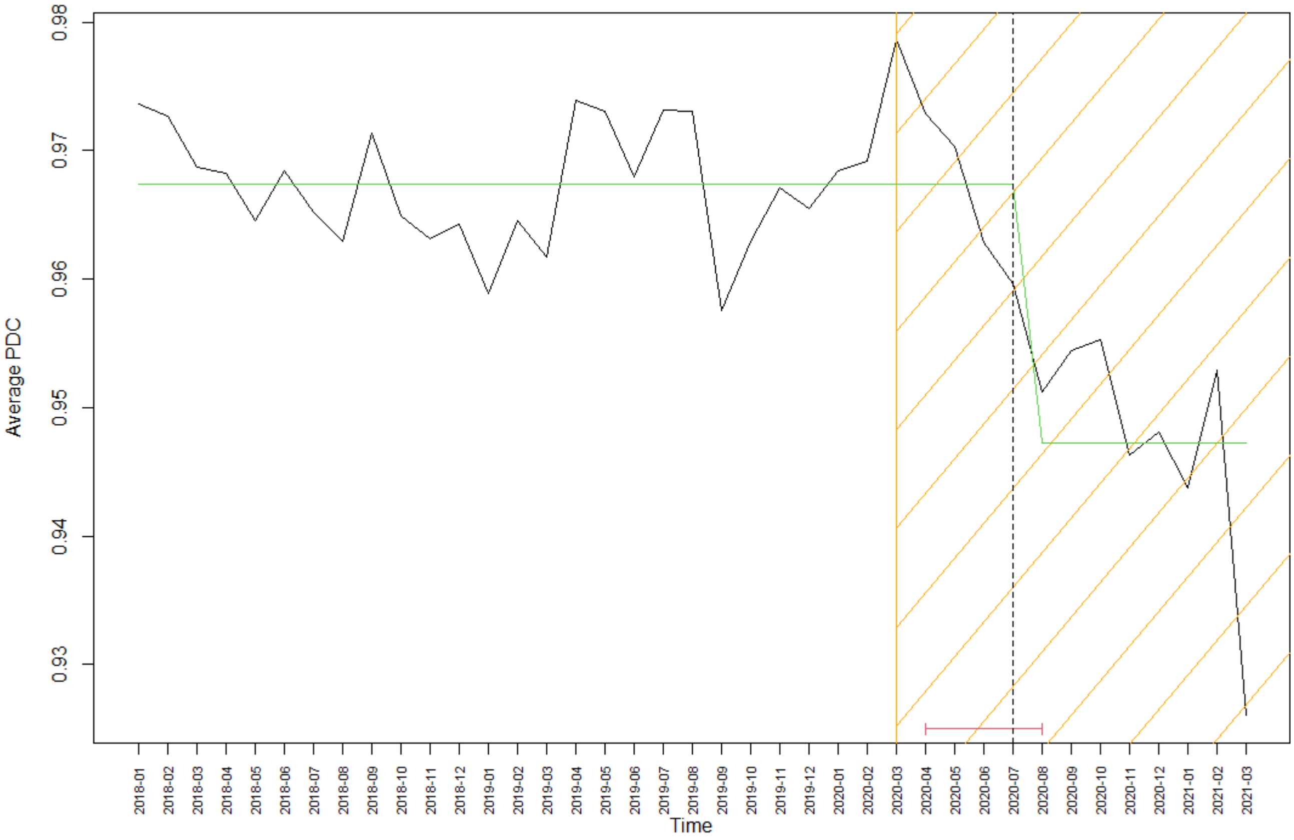


Table 1. The coefficients of the beta regressions models.

Feature name	Dependent variable: PDC[a] score			
	Patients with at least 3 months of PDC scores		All patients	
	Value	*P* value	Value	*P* value
COVID, beta (SE)	–0.076 (0.010)	<.001	–0.081 (0.009)	<.001
Age group (years), beta (SE)				
<16	–0.012 (0.262)	.96	0.173 (0.151)	.25
30-39	0.027 (0.012)	.02	0.032 (0.010)	.002
40-49	0.044 (0.014)	.002	0.041 (0.012)	.001
50-59	0.045 (0.018)	.01	0.055 (0.016)	.001
≥60	0.068 (0.031)	.03	0.084 (0.027)	.002
Sex, male, beta (SE)	–0.023 (0.009)	.01	–0.026 (0.008)	.002
PDC month, beta (SE)				
Second	0.017 (0.011)	.12	0.022 (0.009)	.02
Third	–0.018 (0.011)	.10	–0.025 (0.011)	.02
Number of months of PDC, beta (SE)	N/A[b]	N/A	0.254 (0.006)	<.001
Number of F11 codes, beta (SE)	–0.0001 (0.000)	.001	–0.0001 (0.000)	.001
State, beta (SE)				
2	–0.041 (0.045)	.36	–0.051 (0.039)	.19
3	–0.056 (0.022)	.01	–0.108 (0.019)	.001
4	–0.084 (0.020)	.001	–0.124 (0.017)	<.001
5	–0.036 (0.025)	.14	–0.052 (0.020)	.01
6	–0.210 (0.013)	<.001	–0.308 (0.011)	<.001
Elixhauser category names [17] for the relevant variables , beta (SE)				
Neuro other[c]	–0.038 (0.017)	.03	–0.015 (0.015)	.30
Obesity	0.019 (0.017)	.25	0.030 (0.015)	.04
FluidsLytes[d]	–0.034 (0.018)	.06	–0.022 (0.015)	.16
Alcohol[e]	0.030 (0.018)	.09	0.098 (0.015)	<.001
Drugs[f]	–0.011 (0.012)	.38	–0.020 (0.011)	.06
Constant	3.203 (0.027)	<.001	1.976 (0.027)	<.001
Observations, n	50,265	N/A	71,120	N/A
R^2	0.055	N/A	0.137	N/A
Log likelihood	176,084.60	N/A	225,324.60	N/A

[a] PDC: proportion of days covered.

[b] N/A: not applicable.

[c] Neuro other: other neurological disorders.

[d] FluidsLytes: fluid and electrolyte disorders.

[e] Alcohol: alcohol abuse, referring to alcohol use disorder.

[f] Drug: drug abuse, referring to substance use disorder.

Discussion

Our retrospective cohort study showed a change in adherence to MOUD during the COVID-19 pandemic; this change in adherence did not show up in previous aggregate and dichotomous-variable studies [13,14]. The results showed that patients' adherence to their medication dropped between April and August 2020 and remained low until the end of 2020. Our breakdown of patients by the number of covered months showed that this was true even for patients who had relatively good levels of adherence. (As a rule of thumb, a PDC score of 0.8

has been considered as being adherent, and Warren et al [19] supported this empirically.) This association of a drop in adherence and the COVID-19 period holds when taking into account other factors that also showed a strong association. In particular, state is a significant factor, which is likely to represent a range of factors (urban or rural distribution of population, education, etc). Although these data were not available at an individual level, a machine learning–based study of adherence to MOUD found that demographic proxy data did improve models [19]; that study nevertheless still found a state effect (across the 2 states studied) even after including these data. However, differences in response to COVID-19 could also be a factor. Other factors found to be significant here—sex, age group, etc—are also broadly in line with Warren et al [19].

These results that establish an association between a drop in adherence and the COVID-19 period do not tell us the cause of the drop in adherence, such as patients ceasing to go out to get or fill medication, patients tapering themselves, or some other cause. A small-scale study (429 patients) of opioid agonist treatment use during COVID-19 in Sydney, Australia [20], found that few patients disengaged from treatment, although the social restrictions there differed from the US cohort in this study, as did the distribution of medications (66% methadone in the Australian study). Given the importance of adherence to positive treatment outcomes for opioid addiction, interventions may be warranted. These could range from simple interventions, such as sending reminders and follow-up messages, to more complex ones, such as increased frequency of scheduled follow-up visits with treatment providers and, in some instances, case management. In the context of pandemic restrictions, telehealth offers some possibilities for interventions; the Sydney study [20] found it generally adequate for most clinical reviews, and a more recent study of telemedicine in a US OUD context [21] found that for a range of indicators, telemedicine constituted a comparable alternative to in-person OUD care.

Acknowledgments

We would like to thank the Digital Health Cooperative Research Centre (DHCRC) for access to the data set and Donna Price for support in making use of it. The content is solely the responsibility of the authors and does not necessarily represent the official views of the DHCRC. We would also like to thank the anonymous reviewers for very helpful suggestions.

Data Availability

Data cannot be shared publicly because the authors are not the owners of the data. Data are available from the Digital Health Cooperative Research Centre for researchers who meet the criteria for access to confidential data; access will be the same as that for the authors of this paper. Information about the process, along with a form to apply for data use and data usage agreements, are available on the web [22].

Conflicts of Interest

None declared.

References

1. Slat S, Thomas J, Lagisetty P. Coronavirus disease 2019 and opioid use-a pandemic within an epidemic. JAMA Health Forum 2020 May 01;1(5):e200628 [FREE Full text] [doi: 10.1001/jamahealthforum.2020.0628] [Medline: 36218504]
2. A record number of Americans died from drug overdoses in 2020. The Economist. 2021 Jul 15. URL: https://www.economist.com/graphic-detail/2021/07/15/a-record-number-of-americans-died-from-drug-overdoses-in-2020 [accessed 2023-08-10]
3. Currie JM, Schnell MK, Schwandt H, Zhang J. Trends in drug overdose mortality in Ohio during the first 7 months of the COVID-19 pandemic. JAMA Netw Open 2021 Apr 01;4(4):e217112 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.7112] [Medline: 33852004]
4. Mason M, Arukumar P, Feinglass J. The pandemic stay-at-home order and opioid-involved overdose fatalities. JAMA 2021 Jun 22;325(24):2495-2496. [doi: 10.1001/jama.2021.6700] [Medline: 33890982]
5. Ma J, Bao Y, Wang R, Su M, Liu M, Li J, et al. Effects of medication-assisted treatment on mortality among opioids users: a systematic review and meta-analysis. Mol Psychiatry 2019 Dec 22;24(12):1868-1883. [doi: 10.1038/s41380-018-0094-5] [Medline: 29934549]
6. Mollison C. Concerns about medication adherence grow amid COVID-19 pandemic. Pharmacy Times. 2020 Apr 3. URL: https://www.pharmacytimes.com/view/concerns-about-medication-adherence-grow-amid-covid-19-pandemic [accessed 2023-08-10]
7. COVID-19 widens disparities in medication adherence between patient populations. AllazoHealth. 2020 Mar 27. URL: https://allazohealth.com/resources/covid-19-adherence/ [accessed 2023-08-10]
8. Gomes T, Kitchen SA, Murray R. Measuring the burden of opioid-related mortality in Ontario, Canada, during the COVID-19 pandemic. JAMA Netw Open 2021 May 03;4(5):e2112865 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.12865] [Medline: 34037734]

9. Ágh T, van Boven JF, Wettermark B, Menditto E, Pinnock H, Tsiligianni I, et al. A cross-sectional survey on medication management practices for noncommunicable diseases in Europe during the second wave of the COVID-19 pandemic. Front Pharmacol 2021 Jun 7;12:685696 [FREE Full text] [doi: 10.3389/fphar.2021.685696] [Medline: 34163364]
10. Garfield S, Wheeler C, Boucher C, Etkind M, Lloyd J, Norton J, et al. Medicines management at home during the COVID-19 pandemic: a qualitative study exploring the UK patient/carer perspective. Int J Pharm Pract 2021 Oct 18;29(5):458-464 [FREE Full text] [doi: 10.1093/ijpp/riab050] [Medline: 34343311]
11. Kretchy IA, Asiedu-Danso M, Kretchy J. Medication management and adherence during the COVID-19 pandemic: perspectives and experiences from low-and middle-income countries. Res Social Adm Pharm 2021 Jan;17(1):2023-2026 [FREE Full text] [doi: 10.1016/j.sapharm.2020.04.007] [Medline: 32307319]
12. Khoury D, Preiss A, Geiger P, Anwar M, Conway KP. Increases in naloxone administrations by emergency medical services providers during the COVID-19 pandemic: retrospective time series study. JMIR Public Health Surveill 2021 May 27;7(5):e29298 [FREE Full text] [doi: 10.2196/29298] [Medline: 33999828]
13. Currie JM, Schnell MK, Schwandt H, Zhang J. Prescribing of opioid analgesics and buprenorphine for opioid use disorder during the COVID-19 pandemic. JAMA Netw Open 2021 Apr 01;4(4):e216147 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.6147] [Medline: 33856474]
14. Clement J, Jacobi M, Greenwood BN. Patient access to chronic medications during the COVID-19 pandemic: evidence from a comprehensive dataset of US insurance claims. PLoS One 2021 Apr 1;16(4):e0249453 [FREE Full text] [doi: 10.1371/journal.pone.0249453] [Medline: 33793663]
15. Bai J, Perron P. Estimating and testing linear models with multiple structural changes. Econometrica 1998 Jan;66(1):47-78. [doi: 10.2307/2998540]
16. Zeileis A, Leisch F, Hornik K, Kleiber C, Hansen B, Merkle EC. Package 'strucchange'. J Stat Softw 2015 [FREE Full text]
17. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998 Jan;36(1):8-27. [doi: 10.1097/00005650-199801000-00004] [Medline: 9431328]
18. Wasey JO, Lang M. icd: comorbidity calculations and tools for ICD-9 and ICD-10 codes. R Project. 2020. URL: https://CRAN.R-project.org/package=icd [accessed 2023-08-10]
19. Warren D, Marashi A, Siddiqui A, Eijaz A, Pradhan P, Lim D, et al. Using machine learning to study the effect of medication adherence in Opioid Use Disorder. PLoS One 2022 Dec 15;17(12):e0278988 [FREE Full text] [doi: 10.1371/journal.pone.0278988] [Medline: 36520864]
20. Lintzeris N, Deacon RM, Hayes V, Cowan T, Mills L, Parvaresh L, et al. Opioid agonist treatment and patient outcomes during the COVID-19 pandemic in south east Sydney, Australia. Drug Alcohol Rev 2022 Jul;41(5):1009-1019 [FREE Full text] [doi: 10.1111/dar.13382] [Medline: 34520592]
21. Hailu R, Mehrotra A, Huskamp HA, Busch AB, Barnett ML. Telemedicine use and quality of opioid use disorder treatment in the US during the COVID-19 pandemic. JAMA Netw Open 2023 Jan 03;6(1):e2252381 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.52381] [Medline: 36692880]
22. HMS and DHCRC. Digital Health CRC. URL: https://outgoing.digitalhealthcrc.com/hms-data-sets/ [accessed 2023-08-10]

Abbreviations

BIC: Bayesian information criterion
ICD-10: International Classification of Diseases, 10th revision
MOUD: medication for opioid use disorder
OUD: opioid use disorder
PDC: proportion of days covered

Edited by T Sanchez, A Mavragani; submitted 06.09.22; peer-reviewed by K Conway, MM Hasan; comments to author 19.03.23; revised version received 09.05.23; accepted 24.07.23; published 01.09.23.

Please cite as:
Marashi A, Warren D, Call G, Dras M
Trends in Opioid Medication Adherence During the COVID-19 Pandemic: Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e42495
URL: https://publichealth.jmir.org/2023/1/e42495
doi:10.2196/42495
PMID:37656492

Original Paper

Combinatorial Use of Machine Learning and Logistic Regression for Predicting Carotid Plaque Risk Among 5.4 Million Adults With Fatty Liver Disease Receiving Health Check-Ups: Population-Based Cross-Sectional Study

Yuhan Deng[1,2], MSc; Yuan Ma[3], PhD; Jingzhu Fu[4,5,6], PhD; Xiaona Wang[7], MSc; Canqing Yu[4,5,6,8], PhD; Jun Lv[4,5,6,8], PhD; Sailimai Man[2,4,5,6], PhD; Bo Wang[2,5,8], PhD; Liming Li[4,5,6,8], MD

[1]Chongqing Research Institute of Big Data, Peking University, Chongqing, China
[2]Meinian Institute of Health, Beijing, China
[3]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
[4]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[5]Peking University Health Science Center Meinian Public Health Institute, Beijing, China
[6]Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China
[7]MJ Health Screening Center, Beijing, China
[8]Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing, China

Corresponding Author:
Bo Wang, PhD
Meinian Institute of Health
13 Floor, Health Work, Huayuan Road, Haidian District
Beijing, 100083
China
Phone: 86 010 82097560
Email: paul@meinianresearch.com

Abstract

Background: Carotid plaque can progress into stroke, myocardial infarction, etc, which are major global causes of death. Evidence shows a significant increase in carotid plaque incidence among patients with fatty liver disease. However, unlike the high detection rate of fatty liver disease, screening for carotid plaque in the asymptomatic population is not yet prevalent due to cost-effectiveness reasons, resulting in a large number of patients with undetected carotid plaques, especially among those with fatty liver disease.

Objective: This study aimed to combine the advantages of machine learning (ML) and logistic regression to develop a straightforward prediction model among the population with fatty liver disease to identify individuals at risk of carotid plaque.

Methods: Our study included 5,420,640 participants with fatty liver from Meinian Health Care Center. We used random forest, elastic net (EN), and extreme gradient boosting ML algorithms to select important features from potential predictors. Features acknowledged by all 3 models were enrolled in logistic regression analysis to develop a carotid plaque prediction model. Model performance was evaluated based on the area under the receiver operating characteristic curve, calibration curve, Brier score, and decision curve analysis both in a randomly split internal validation data set, and an external validation data set comprising 32,682 participants from MJ Health Check-up Center. Risk cutoff points for carotid plaque were determined based on the Youden index, predicted probability distribution, and prevalence rate of the internal validation data set to classify participants into high-, intermediate-, and low-risk groups. This risk classification was further validated in the external validation data set.

Results: Among the participants, 26.23% (1,421,970/5,420,640) were diagnosed with carotid plaque in the development data set, and 21.64% (7074/32,682) were diagnosed in the external validation data set. A total of 6 features, including age, systolic blood pressure, low-density lipoprotein cholesterol (LDL-C), total cholesterol, fasting blood glucose, and hepatic steatosis index (HSI) were collectively selected by all 3 ML models out of 27 predictors. After eliminating the issue of collinearity between features, the logistic regression model established with the 5 independent predictors reached an area under the curve of 0.831 in the internal validation data set and 0.801 in the external validation data set, and showed good calibration capability graphically. Its predictive performance was comprehensively competitive compared with the single use of either logistic regression or ML algorithms. Optimal predicted probability cutoff points of 25% and 65% were determined for classifying individuals into low-, intermediate-, and high-risk categories for carotid plaque.

Conclusions: The combination of ML and logistic regression yielded a practical carotid plaque prediction model, and was of great public health implications in the early identification and risk assessment of carotid plaque among individuals with fatty liver.

(JMIR Public Health Surveill 2023;9:e47095) doi:10.2196/47095

KEYWORDS

machine learning; carotid plaque; health check-up; prediction; fatty liver; risk assessment; risk stratification; cardiovascular; logistic regression

Introduction

Carotid plaque is an independent risk factor for cerebral stroke [1], myocardial infarction [2], and atherosclerotic cardiovascular disease [2], which are all leading causes of death and disability worldwide [3,4], presenting severe economic burden in both developed and developing countries [5]. Nearly 20% of stroke cases were caused by carotid atherosclerotic plaque [6]. The rupture or shedding of carotid plaque can lead to thrombosis and has become the major cause of cerebrovascular accidents [3,7]. It has been proposed that nearly one-third of Chinese adults were experiencing from carotid plaque [8]. With the growing aging population and the acceleration in urbanization, the incidence rates of cardiovascular disease in China would increase steadily in the next few decades [9]. Thus, early detection of carotid plaque can bring great benefits in the timely and active prevention of stroke and other cerebrovascular and cardiovascular diseases. It is necessary to develop effective tools to identify carotid plaque in the asymptomatic population and curb its progression at an early stage.

Recently, several studies have demonstrated significant associations exist between fatty liver disease and coronary artery disease, including carotid plaque and carotid stenosis [10-12]. Individuals with fatty liver disease were proved to have an elevated risk of developing carotid plaque. However, although liver ultrasound has been incorporated into the routine check-up program, carotid artery ultrasound examination is not prevalent due to cost-effectiveness reasons for the asymptomatic population [13]. As the most widely used method for evaluating carotid plaque [14], the low prevalence of carotid ultrasound may result in missed detection of such plaque populations, especially in individuals with fatty liver. Thus, identifying carotid plaque patients in the population with fatty liver is more cost-effective and is of great public health implications for the prevention of cardiovascular disease.

Wu et al [13] developed a carotid plaque risk prediction tool among asymptomatic population based on machine learning (ML) algorithms, including extreme gradient boosting (XGBoost), gradient boosting decision tree, random forest (RF), and support vector machine, and achieved good performance, but the substantial complexity of the model may limit its practical use, while the commonly used risk prediction tools in the cardiovascular field, including Framingham risk score [15] and its modified model [16], were mostly based on traditional statistical models, including logistic regression and cox proportional-hazards regression. Although numerous studies have demonstrated that ML algorithms outperformed traditional statistical models in predictive performance throughout medical fields [17-19] due to their capability to analyze and learn the complex interactions and nonlinear associations among variables [17,20,21], the latter still own irreplaceable strengths, including their natural transparency, interpretability, and robustness, which boost their practicality in clinical research [22]. Therefore, using ML algorithms alone or traditional regression methods alone to train prediction models usually results in either accurate but complicated black boxes or practical but unsatisfactory-performed scoring systems.

In this study, we combined ML, including RF, XGBoost, and elastic net (EN) with logistic regression together to develop a straightforward and practical risk prediction model to help better identify individuals at risk of carotid plaque in the population with fatty liver disease. We also provided robust cutoff points for carotid plaque risk stratification and verified the results on an external data set.

Methods

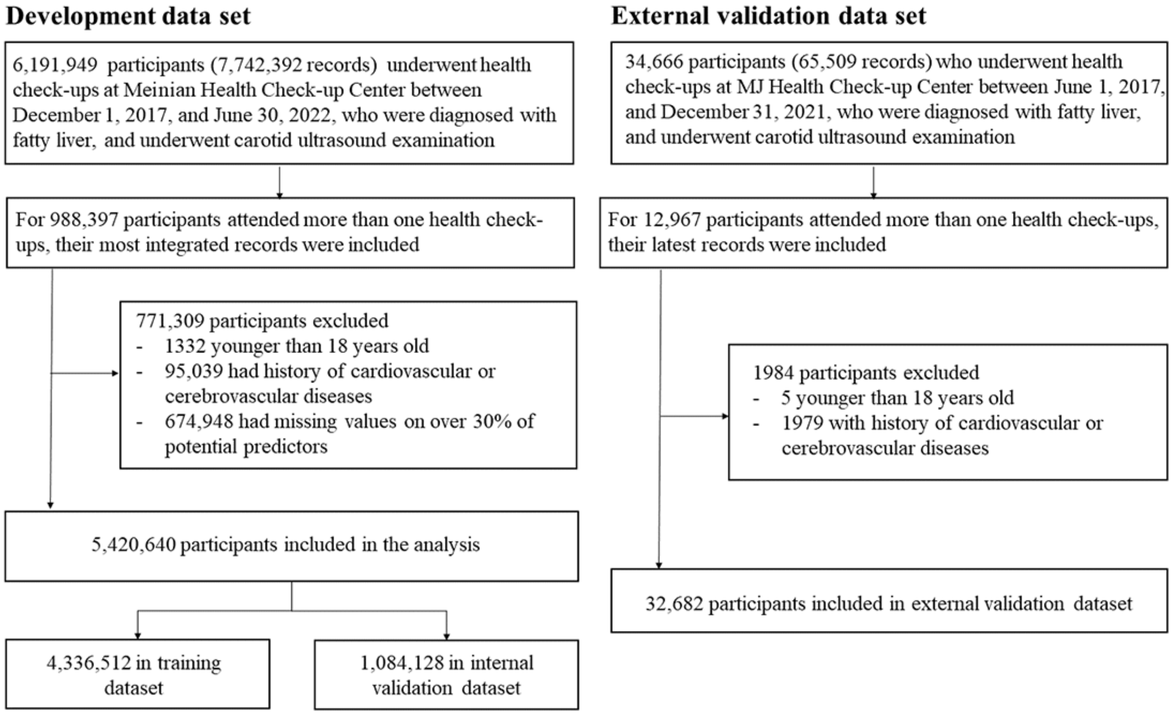
Data Source and Study Participants

Participants who attended health check-ups at Meinian Health Check-up Center and MJ Health Check-up Center were involved in this study for model development and external validation, respectively.

Development Data Set

Meinian Health Check-up Program is the largest check-up chain in China. It provides routine health check-up services for the whole population, with its check-up centers covering all 31 provinces in mainland China. Initially, participants diagnosed with fatty liver through hepatic ultrasonography and received carotid artery ultrasound examination between January 1, 2017, and June 30, 2022, were included. For those who attended 2 or more check-ups, the most integrated record was included in the analysis. Participants younger than 18 years, had missing values on over 30% of potential predictors, or had a history of cardiovascular or cerebrovascular diseases were excluded from the analysis. Finally, a total of 5,420,640 participants were included in the study for model development (Figure 1).

Figure 1. Flowchart of the study participants.



External Validation Data Set

MJ Health Check-up Center is a clinic in Beijing, China, it provides comprehensive health check-ups for the participants. Participants who underwent check-ups between June 1, 2017, and December 31, 2021, were diagnosed with fatty liver, and underwent carotid ultrasound examination were included. After excluding those younger than 18 years, or who had a history of cardiovascular or cerebrovascular diseases, 32,682 participants were included (Figure 1).

Ethical Considerations

The study was reviewed and approved by the institutional review board of Peking University Health Science Center (approval ID: IRB00001052-19077). The requirement for informed consent of participants was waived due to the use of deidentified data obtained as part of routine health check-ups.

Potential Predictors and Outcomes

All potential factors associated with carotid plaque reported by recent studies were considered. Considering the accessibility of the variables in the database, a total of 27 potential factors were extracted: (1) demographic characteristics: sex and age; (2) physical examination indicators: weight, height, BMI, systolic blood pressure (SBP), diastolic blood pressure, and heart rate; (3) laboratory examination indicators: total cholesterol (TC), triglyceride, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol (LDL-C), fasting blood glucose (FBG), alanine transaminase (ALT), aspartate aminotransferase (AST), ALT/AST, direct bilirubin, total bilirubin, alkaline phosphatase (ALP), uric acid, blood platelet count (PLT), white blood cell count, creatinine, and hepatic steatosis index (HSI, which was calculated as follows: HSI=8×(ALT/AST)+BMI+2 (if diabetes mellitus)+2 (if female) [12]); and (4) medical history: hypertension, diabetes, and hyperlipidemia.

The outcome was defined as whether the participant was diagnosed with carotid plaque by carotid artery ultrasound examination. Specifically, the common carotid arteries, the bifurcation, and the external and internal carotid arteries were examined on each side by experienced sonographers operating a Doppler ultrasound system (Sonoscape S50, China) with a linear 7.5 MHz probe under standardized protocols. The distance between the leading edge of the lumen-intima echo and the leading edge of the media-adventitia echo was defined as carotid intima-media thickness. Carotid plaque was accounted as a discrete, focal wall thickening ≥1.5 mm or focal thickening >50% greater than the surrounding carotid intima-media thickness in any of the arterial segments above [23].

Data Preprocessing and Statistical Analysis

The development data set was randomly divided into a training set (4,336,512/5,420,640, 80%) and an internal validation set (1,084,128/5,420,640, 20%). The training set was used for feature selection and model development, while the internal validation data set, together with the external validation data set, were used for model evaluation.

In the training set, missing data were imputed with the mean of each variable. And the imputed values derived from the training set were further used for missingness imputation in internal and external validation data sets. Outliers in the training set were defined as values distributed less than 1% or more than 99% quantile of the whole participants, and all the outliers were regarded as missing values.

Normally distributed continuous variables were presented as means with SDs and used the Student *t* test for statistical analysis. Non-normally distributed continuous variables were presented as a median and interquartile range and the Wilcoxon rank-sum test was used for comparison. Categorical variables were presented as counts and percentages and compared using the chi-square test.

Considering the limitations of *P* values in detecting group differences in large sample sizes, we used standardized mean difference (SMD) as an alternative method to compare the between-group differences. Unlike *P* values, SMD allows for standardized comparisons across groups despite differences in sample size, measurement scales, or variance [24]. An absolute value of SMD<0.20 can be considered as a small difference, and an absolute value of SMD<0.10 suggests a negligible difference.

Feature Selection and Model Development

Three ML algorithms were used for feature selection, including RF, EN, and XGBoost. To tune hyperparameters in the training set, 5-fold cross-validation was conducted. The top important features coselected by 3 algorithms were used for model development.

RF and XGBoost are 2 popular ensemble learning algorithms. Both of them use decision trees to construct their models. RF generates multiple decision trees in parallel by conducting random sampling and random feature selection, and the final prediction is made by aggregating the votes from all decision trees. Feature importance in RF can be measured by evaluating the mean Gini index of each feature across multiple trees.

In contrast, XGBoost builds decision trees sequentially, where each tree is trained to correct the errors of the previous ones, and eventually, the prediction is obtained by summing the results of all trees. Feature importance in XGBoost can be quantified by calculating the average gain that a feature brings when it is chosen as the splitting variable in any decision tree.

EN model extends logistic regression by adding L1 and L2 regularization terms to overcome multicollinearity and perform feature selection. The importance of each feature in EN model can be estimated by examining the magnitude of its coefficients.

Logistic regression was used to train the final prediction model by using the features selected. Collinearity was checked through clustering analysis and the most representative feature, that is owning the smallest value of $1\text{-}R^2$ in each cluster was further selected to train logistic regression.

Model performance was assessed using discrimination and calibration. Discrimination was evaluated by area under the receiving operating characteristic curve (AUROC), and calibration was investigated through the calibration curve and Brier score. The performance of the established model was compared with a model using backward selection logistic regression without ML based-feature selection, as well as the above 3 ML models. In addition, decision curve analysis was performed to see whether the net benefit would promote when using the prediction model.

Risk Stratification

Risk cutoff values were determined based on the Youden index, predicted risk probability distribution, and the prevalence rate of carotid plaque in the internal validation data set to divide participants into high-risk, intermediate-risk, and low-risk. Specifically, Youden index was used to identify the high-risk group, and then a cutoff value was selected for the remaining individuals based on the distribution of predicted probabilities determined by our prediction model and adjusted according to the prevalence rate of carotid plaque of the 2 groups below this cutoff point to achieve intermediate and low-risk stratification, and the effectiveness of these cutoff points was verified on an external validation data set.

All procedures were performed in SAS (version 9.4; SAS Institute) and Python (version 3.7; Python Software Foundation).

Results

Characteristics of Study Participants

Among the 5,420,640 participants in the development data set, 26.23% (1,421,970/5,420,640) were diagnosed with carotid plaque. The differences in all the potential predictors between participants with and without carotid plaque were statistically significant. Participants who developed carotid plaque were older, more likely to be female, and had higher SBP, diastolic blood pressure, TC, high-density lipoprotein cholesterol, LDL-C, FBG, direct bilirubin, total bilirubin, ALP, and lower height, weight, BMI, heart rate, triglyceride, ALT, AST, ALT/AST, uric acid, white blood cell count, creatinine, HSI when compared to their carotid plaque-free counterparts. The prevalence of hypertension, hyperlipidemia, and diabetes mellitus was also higher in the carotid plaque group compared to the carotid plaque-free group (Table 1). Among the 32,682 records in external validation, 21.64% (7074/32,682) were recorded for developing carotid plaque. The characteristics are presented in Table S1 of Multimedia Appendix 1. The differences between the development data set and external validation data set are presented in Table S2 of Multimedia Appendix 1.

Table 1. Characteristics of study participants in the development data set.

Characteristic	Total (N=5,420,640)	Carotid plaque			SMD[a]
		Yes (N=1,421,970)	No (N=3,998,670)	*P* value	
Sex, n (%)				<.001	–0.01
Male	3,667,424 (67.66)	955,704 (67.21)	2,711,720 (67.82)		
Female	1,753,216 (32.34)	466,266 (32.79)	1,286,950 (32.18)		
Age (years), mean (SD)	49.00 (39.00, 57.00)	57.00 (51.00, 64.00)	45.00 (36.00, 54.00)	<.001	1.13
HT[b] (cm), median (IQR)	167.00 (160.50-173.00)	166.00 (159.10-171.50)	168.00 (161.00-173.50)	<.001	–0.22
WT[c] (kg), mean (SD)	75.11 (12.34)	73.39 (11.40)	75.72 (12.60)	<.001	–0.19
BMI (kg/m^2), mean (SD)	26.90 (3.19)	26.75 (3.02)	26.96 (3.25)	<.001	–0.06
SBP[d] (mm Hg), mean (SD)	132.06 (18.53)	139.37 (19.67)	129.46 (17.38)	<.001	0.53
DBP[e] (mm Hg), mean (SD)	80.83 (12.15)	82.79 (12.22)	80.13 (12.05)	<.001	0.22
HR[f], (times/minute), mean (SD)	71.67 (8.35)	71.31 (8.46)	71.80 (8.31)	<.001	–0.06
TC[g] (mmol/L), mean (SD)	5.21 (1.03)	5.34 (1.09)	5.16 (1.00)	<.001	0.17
TG[h] (mmol/L), median (IQR)	1.76 (1.24-2.55)	1.76 (1.26-2.51)	1.76 (1.24-2.57)	.07	0.07
HDL-C[i] (mmol/L), median (IQR)	1.28 (1.10-1.44)	1.29 (1.12-1.47)	1.27 (1.09-1.43)	<.001	0.10
LDL-C[j] (mmol/L), mean (SD)	3.10 (0.83)	3.19 (0.87)	3.07 (0.81)	<.001	0.14
FBG[k] (mmol/L), median (IQR)	5.45 (4.99-6.02)	5.71 (5.19-6.52)	5.37 (4.94-5.87)	<.001	0.33
ALT[l] (U/L), median (IQR)	26.00 (18.30-38.50)	23.40 (17.30-33.09)	27.00 (19.00-40.32)	<.001	–0.22
AST[m] (U/L), median (IQR)	22.00 (18.00-27.00)	21.90 (18.00-26.30)	22.00 (18.00-27.40)	<.001	–0.05
ALT/AST, median (IQR)	1.20 (0.94-1.52)	1.10 (0.88-1.36)	1.24 (0.97-1.58)	<.001	–0.02
DBIL[n] (µmol/L), mean (SD)	3.71 (1.73)	3.74 (1.83)	3.70 (1.70)	<.001	0.02
TBIL[o] (µmol/L), median (IQR)	13.65 (10.86-14.66)	13.65 (11.06-14.90)	13.65 (10.80-14.60)	<.001	0.04
ALP[p] (U/L), mean (SD)	77.86 (19.04)	79.74 (19.82)	77.19 (18.71)	<.001	0.13
UA[q] (µmol/L), mean (SD)	369.13 (95.66)	358.66 (91.36)	372.85 (96.87)	<.001	–0.15
PLT[r] (10^9/L), mean (SD)	224.53 (57.17)	217.54 (56.48)	227.01 (57.21)	<.001	–0.17
WBC[s] (10^9/L), mean (SD)	6.36 (4.74)	6.35 (5.11)	6.36 (4.60)	.01	–0.01
Cr[t] (µmol/L), mean (SD)	69.36 (17.10)	69.02 (17.84)	69.49 (16.82)	<.001	–0.03
HSI[u], mean (SD)	37.83 (98.56)	36.96 (4.63)	38.14 (114.72)	<.001	–0.01
Hypertension, n (%)				<.001	0.49
Yes	2,190,601 (40.41)	823,098 (57.88)	1,367,503 (34.20)		
No	3,230,039 (59.59)	598,872 (42.12)	2,631,167 (65.80)		
Hyperlipidemia, n (%)				<.001	0.06
Yes	2,458,476 (45.35)	676,089 (47.55)	1,782,387 (44.57)		
No	2,962,164 (54.65)	745,881 (52.45)	2,216,283 (55.43)		
Diabetes mellitus, n (%)				<.001	0.34
Yes	717,876 (13.24)	318,281 (22.38)	399,595 (9.99)		

Characteristic	Total (N=5,420,640)	Carotid plaque			SMD[a]
		Yes (N=1,421,970)	No (N=3,998,670)	*P* value	
No	4,702,764 (86.76)	1,103,689 (77.62)	3,599,075 (90.01)		

[a] SMD: standardized mean difference.

[b] HT: height.

[c] WT: weight.

[d] SBP: systolic blood pressure.

[e] DBP: diastolic blood pressure.

[f] HR: heart rate.

[g] TC: total cholesterol.

[h] TG: triglyceride.

[i] HDL-C: high-density lipoprotein cholesterol.

[j] LDL-C: low-density lipoprotein cholesterol.

[k] FBG: fasting blood glucose.

[l] ALT: alanine transaminase.

[m] AST: aspartate aminotransferase.

[n] DBIL: direct bilirubin.

[o] TBIL: total bilirubin.

[p] ALP: alkaline phosphatase.

[q] UA: uric acid.

[r] PLT: blood platelet count.

[s] WBC: white blood cell count.

[t] Cr: creatinine.

[u] HSI: Hepatic Steatosis Index.

Feature Importance and Model Performance

Age, SBP, LDL-C, TC, FBG, and HSI were found to be the top important features through all 3 ML algorithms. These features ranked in the top 10 features of all 3 algorithms and were selected out of the 27 features to train the logistic regression model (Figure 2). Cluster analysis showed high collinearity existed between LDL-C and TC, thus the more informative one, LDL-C, was selected to develop the final model (Table 2). The formula for predicting the risk of carotid plaque, as determined by the final prediction model is given in the following equation:

Figure 2. Feature importance of the potential predictors on carotid plaque in population with fatty liver disease generated by (A) RF, (B) EN, and (C) XGBoost. The features highlighted in dark color represent those coselected by all 3 algorithms. ALP: alkaline phosphatase; ALT: alanine transaminase; AST: aspartate aminotransferase; Cr: creatinine; DB: diabetes; DBIL: direct bilirubin; DBP: diastolic blood pressure; EN: elastic net; FBG: fasting blood glucose; HDL-C: high-density lipoprotein cholesterol; HLP: Hyperlipidemia; HR: heart rate; HSI: hepatic steatosis index; HT: height; HTN: hypertension; LDL-C: low-density lipoprotein cholesterol; PLT: blood platelet count; RF: random forest; SBP: systolic blood pressure; TBIL: total bilirubin; TC: total cholesterol; TG: triglyceride; UA: uric acid; WBC: white blood cell count; WT: weight; XGBoost: extreme gradient boosting.

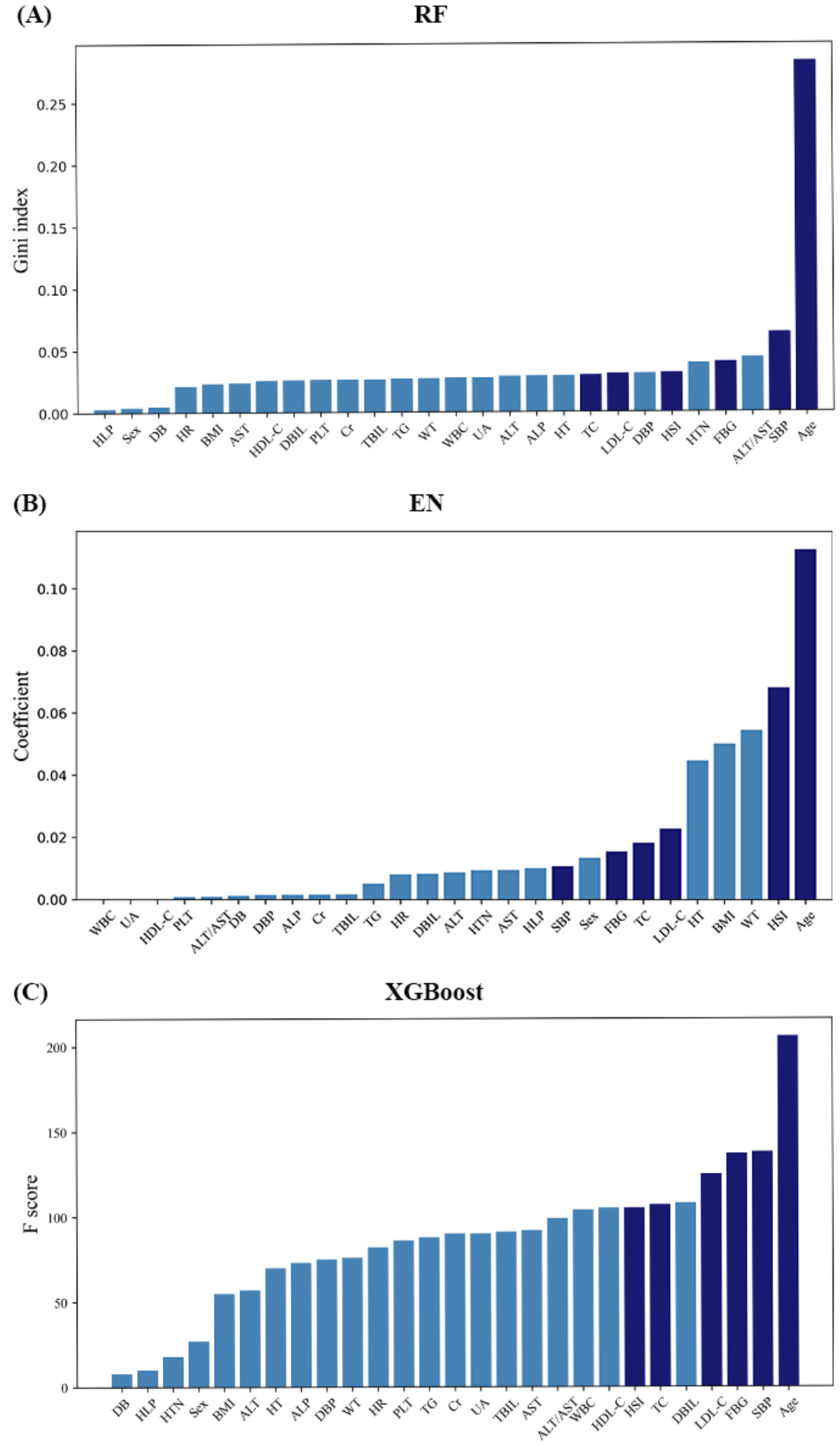


Table 2. Carotid plaque prediction model in population with fatty liver disease based on logistic regression.

Variable	β	SE	OR[a] (95% CI)
Intercept	–7.97380	0.015400	
Age	0.09230	0.000133	1.097 (1.096-1.097)
SBP[b]	0.01080	0.000070	1.011 (1.011-1.011)
FBG[c]	0.09840	0.000684	1.103 (1.102-1.105)
LDL-C[d]	0.16330	0.001480	1.177 (1.174-1.181)
HSI[e]	–0.00949	0.000289	0.991 (0.990-0.991)

[a] OR: odds ratio.

[b] SBP: systolic blood pressure.

[c] FBG: fasting blood glucose.

[d] LDL-C: low-density lipoprotein cholesterol.

[e] HSI: Hepatic Steatosis Index.

Model performance was evaluated in internal and external validation data sets, respectively, and the area under the curves achieved 0.831 and 0.801, respectively, both showing good discrimination capability (Figure 3A). The calibration curve in the internal validation data set lies tightly against the diagonal, while the external deviates a little, but still indicates good calibration capability (Figure 3B). When compared with the backward selection logistic regression model, which consisted of 15 features, or the 3 ML models involving 27 features, the prediction model we established with only 5 features was still competitive. The number of features and model performance in each model are shown in Table 3.

Figure 3. Model performance in discrimination and calibration for predicting the risk of carotid plaque in population with fatty liver disease evaluated by (A) ROC curves and (B) calibration curves. AUC: area under the curve; ROC: receiver operating characteristic.

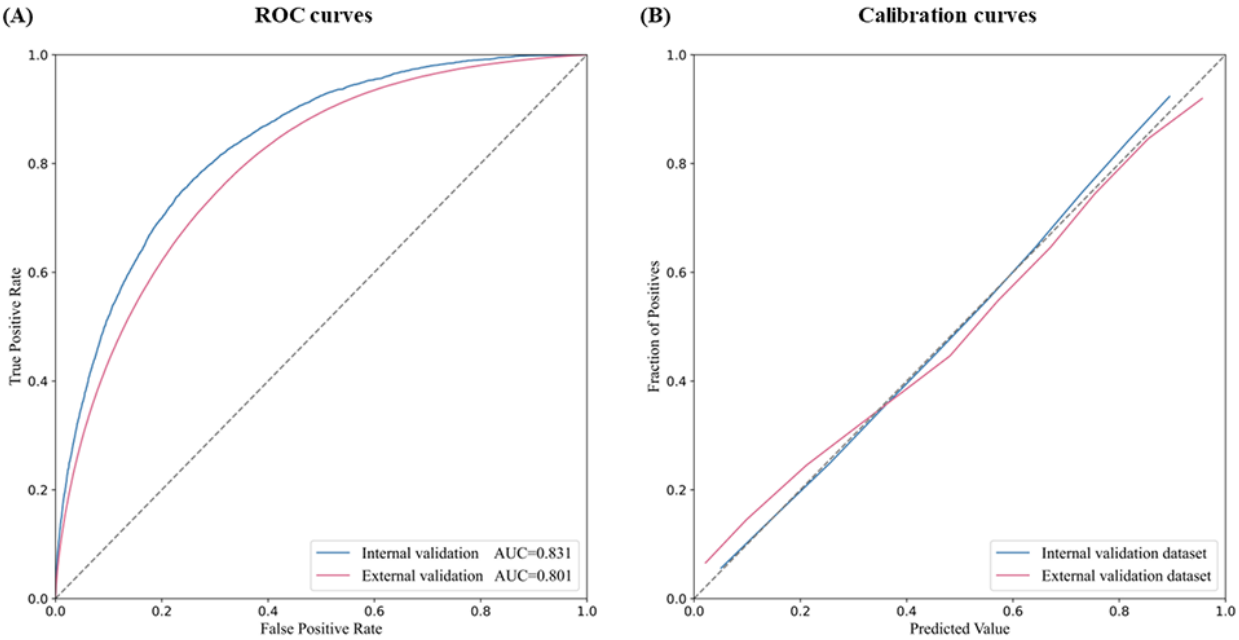


Table 3. The comparison of model performance for predicting the risk of carotid plaques in population with fatty liver disease in the internal validation data set.

Model	Features, n	AUC[a]	Brier score
LR-ML[b]	5	0.831	0.125
LR-BS[c]	13	0.822	0.139
RF[d]	27	0.832	0.151
EN[e]	27	0.834	0.178
XGBoost[f]	27	0.831	0.150

[a] AUC: area under the curve.

[b] LR-ML: ML-based feature selection logistic regression.

[c] LR-BC: backward selection logistic regression.

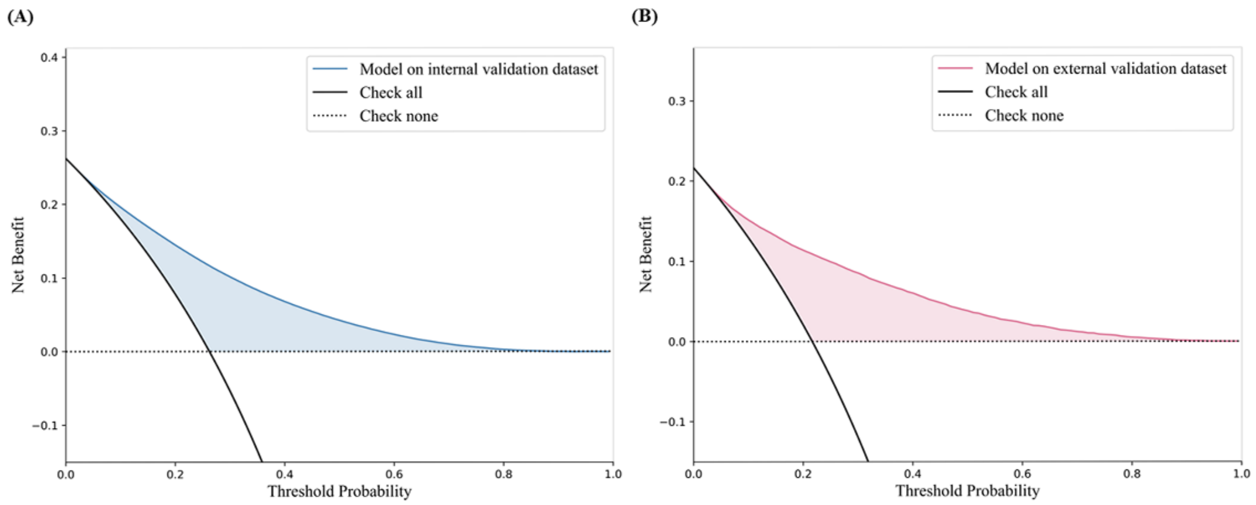
[d] RF: random forest.

[e] EN: elastic net.

[f] XGBoost: extreme gradient boosting.

The decision curve analysis showed that the application of the prediction model achieved promoted net benefits throughout all threshold probabilities both in internal and external validation data sets, indicating prospective utility in the real-world scenario (Figure 4).

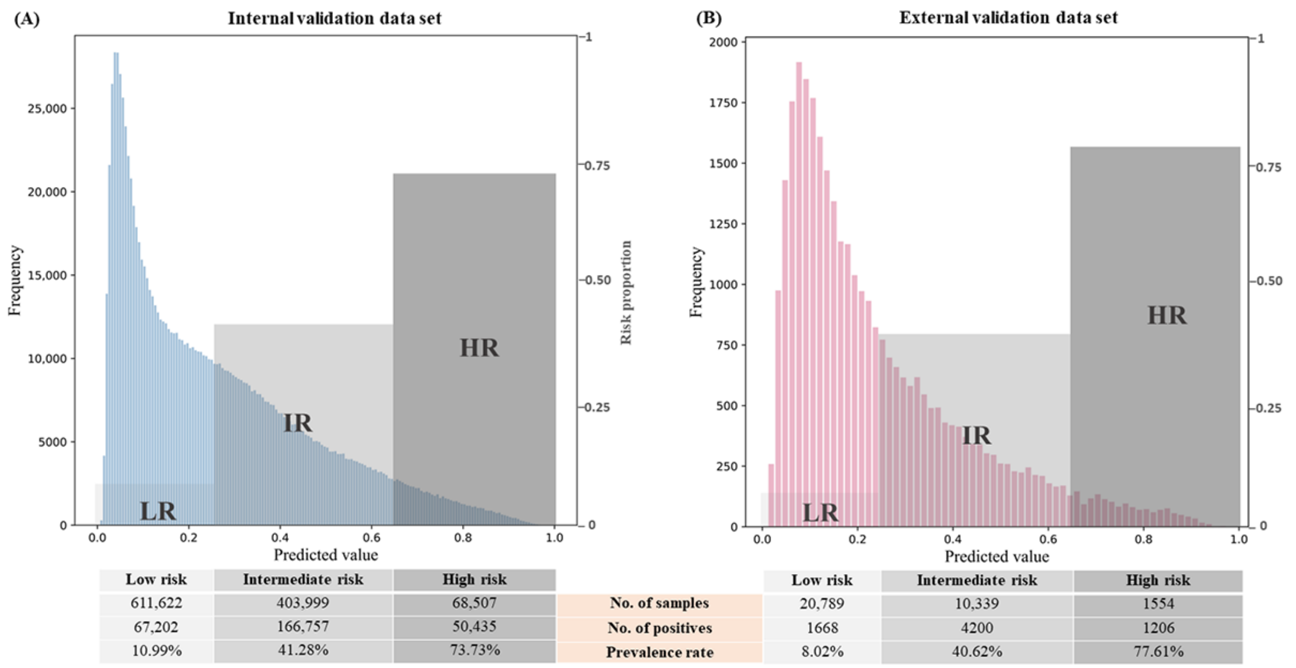
Figure 4. Decision curve analysis for predicting the risk of carotid plaque in population with fatty liver disease in (A) internal validation data set and (B) external validation data set.



Risk Stratification

The predicted risk probabilities of participants developing carotid plaque in 2 validation data sets were calculated and their histograms were presented in Figure 5. Using the Youden index, a threshold of 65% was used to categorize individuals as being at high risk, while a cutoff value of 25% was used to distinguish those at intermediate risk and low risk based on the distribution of predicted probabilities and the prevalence of carotid plaque in the remaining population. Our result revealed the prevalence of carotid plaque of 73.73%, 41.28%, and 10.99% for the high-risk, intermediate-risk, and low-risk groups, respectively. Upon application of these defined cutoff points to the external validation data set, we observed comparable prevalence rates of carotid plaque within each risk group when compared to the internal validation data set. Notably, the high-risk, intermediate-risk, and low-risk groups exhibited prevalence rates of carotid plaque amounting to 77.61%, 40.62%, and 8.02%, respectively. These findings indicate that the selected risk cutoff points can successfully stratify individuals with fatty liver disease into varying degrees of severity in terms of the risk of developing carotid plaque. The probability distribution, sample size, and prevalence rate in each level are also shown in Figure 5.

Figure 5. Probability distribution and risk classification plot generated by the carotid plaque prediction model in population with fatty liver disease in (A) internal validation data set and (B) external validation data set. The blue and pink colored columns represent the number of participants on different predicted probabilities, and the predicted probabilities are split into low risk, intermediate risk, and high risk by 0.25 and 0.65. Different levels of risks are presented by gray pillars of different opacities, the height of each pillar corresponds to risk proportion, which is calculated by the prevalence rate in each risk level. HR: high risk; IR: intermediate risk; LR: low risk.



Low risk	Intermediate risk	High risk		Low risk	Intermediate risk	High risk
611,622	403,999	68,507	**No. of samples**	20,789	10,339	1554
67,202	166,757	50,435	**No. of positives**	1668	4200	1206
10.99%	41.28%	73.73%	**Prevalence rate**	8.02%	40.62%	77.61%

Discussion

In this study, we established a practical and straightforward carotid plaque prediction model in population with fatty liver disease. By using only 5 features (Age, SBP, FBG, LDL-C, and HSI) coselected by 3 ML algorithms, the model achieved an AUROC of 0.831 and exhibited good calibration properties. Our study derived robust cutoff points of 25% and 65% for carotid plaque risk probability, enabling effective risk stratification and facilitating clinical decision-making regarding the need for carotid ultrasonography examination. These findings have practical implications for early detection and prevention of this condition, which can improve patient outcomes and reduce health care costs.

We identified specific features as strong predictors of the outcome. Age was selected as the top important feature by all 3 models, indicating its strong relationship with carotid plaque, prior studies have also drawn the same conclusion [8,25]. Previous research has demonstrated that increased SBP is a strong predictor of the development of carotid plaque, which was consistent with our findings [26-30]. Additionally, our study found evidence linking increased blood lipid levels, such as TC and LDL-C, and elevated glucose levels to a higher prevalence of carotid plaque. These findings are in line with prior research and support the notion that managing modifiable cardiovascular risk factors, such as dyslipidemia and hyperglycemia, is critical for reducing the likelihood of carotid plaque development [31-33]. HSI is a surrogate score for the noninvasive assessment of steatosis in patients with fatty liver [34] and is also a screening tool for nonalcoholic fatty liver disease [35]. In our study, we regarded this index as a continuous feature reflecting the severity of liver steatosis to predict carotid plaque. Although a cross-sectional study involving 768 patients with type 2 diabetes mellitus (T2DM) showed those with carotid plaque have significantly higher HSI (*P*<.001) compared with their healthy counterparts [36], our study came to the opposite conclusion. The following reason may explain the paradox. For all of the participants who have already been diagnosed with fatty liver disease, the severity may lead to behavior or lifestyle change and ulteriorly affect the development of carotid plaque. However, the lifestyle-related variables and diagnostic time of fatty liver were not included in our study, which may generate the opposite result with other studies.

Our findings revealed that the logistic regression model, comprising only 5 variables coselected by 3 ML algorithms, attained nearly equivalent area under the curve values as the ML models which included all 27 variables, but exhibited superior calibration capability. These results clearly indicated the superiority of adopting a combined approach. In regard to similar research, our predictive model remains competitive. For instance, in Wu's [13] investigations, the XGBoost model based on 34 variables acquired an AUROC of 0.8635, whereas our model, employing solely 5 variables, yielded a comparable AUROC value of 0.831 while preserving a more comprehensible and lucid modeling framework. In practical applications, our model accurately predicts outcomes using routine, easily measurable, and obtainable variables, indicating the potential for effective clinical implementation.

We aimed to identify high-risk individuals who may benefit from carotid ultrasonography screening for carotid plaque. Therefore, it is vital not only to estimate an individual's risk probability but also to determine optimal risk cutoff points for precise risk stratification and corresponding clinical guidance.

Although the Youden index is commonly used to determine the optimal cutoff predicted probability for risk stratification, it typically results in a binary classification of high-risk and low-risk groups [37-39]. However, our results have shown that the low-risk group often comprises a larger population with a wide range of risk probabilities from 0% to 65% when performing high-low risk stratification using the Youden index alone. Although this approach can effectively identify the high-risk group, providing identical guidance to individuals within the low-risk group with significantly different risk probabilities is not appropriate and unscientific. To achieve a more comprehensive risk stratification based on our large sample data, we further stratified the low-risk group generated by the Youden index into low and intermediate risk categories using risk probability distribution and prevalence rate of carotid plaque in each group. With our substantial sample size, selecting and adjusting cutoff points based on the distribution of risk probability and prevalence rate across different strata is achievable. This novel approach enables a more nuanced risk stratification beyond the binary classification of high and low-risk groups, potentially leading to the development of personalized health care plans.

We have developed customized health care recommendations for each risk group, providing precise guidance for carotid artery ultrasound examinations. Our results indicate that individuals in the high-risk group with a prevalence of carotid plaque of over 70% should strongly consider undergoing carotid artery ultrasound for definitive diagnosis. For those in the intermediate-risk group with a prevalence of over 40%, the examination is still advised, considering their individual financial circumstances. Furthermore, low-risk individuals with a prevalence of approximately 10% do not require a carotid artery ultrasound examination. By personalizing our approach based on an individual’s risk level and financial situation, we can effectively identify those who require further testing and optimize the cost-effectiveness of screening programs.

Currently, there are no established criteria for determining which populations require carotid ultrasound screening. Several guidelines and recommendations have been proposed to identify populations that may benefit from carotid ultrasound screening. For example, the guidelines for carotid artery ultrasound examination in Chinese health check-up populations specify that the evaluation standards for individuals undergoing carotid artery ultrasound include those at risk of hypertension, coronary atherosclerotic heart disease, stroke, and diabetes; high-risk populations such as smokers, overweight and obese individuals; individuals with moderate or higher cardiovascular risk assessment; and other suitable populations aged middle-aged or older. The American Heart Association recommends carotid ultrasound screening for asymptomatic patients who are over 65 years of age, men aged 55 to 75 years with a history of smoking or other risk factors, and women aged 55 to 75 years with a history of cardiovascular disease or other risk factors. However, the existing guidelines are primarily focused on risk stratification at the population level. Our objective, therefore, is to develop a prediction model for personalized risk stratification to enable better decision-making support in determining the need for carotid artery ultrasound monitoring on an individual basis. This approach would lead to more precise and personalized health care recommendations for the individuals under consideration.

Our prediction model and the cutoff points were verified on an independent external data set. The model was also able to accurately predict the risk of carotid plaque for each individual and the cutoff points remain robust in identifying different risk levels of groups, confirming the generalizability and applicability of our approach.

Several limitations need to be noted. First, due to the limitation of the database, some lifestyle variables, like smoking or drinking status [40,41], were not included in our model, which may affect the predictive performance to some extent. Second, because of the high calculation time cost caused by the huge sample size, bootstrap sampling was not used to generate a 95% CI of the performance metrics, a single measurement may not be forceful enough. Third, we opted to exclude individuals with cardiovascular and cerebrovascular diseases from our study sample, in recognition of their potential differences in baseline characteristics, health care–seeking behavior, and management strategies relative to the general population. These factors could introduce significant confounding effects and hinder the predictive performance of our model. Therefore, we excluded individuals with cardiovascular and cerebrovascular diseases at the beginning. However, this may have reduced the representativeness of our sample and introduced some selection bias. Fourth, although the check-up centers included in this study covered all provinces and all 3 economic zones (the eastern zone, central zone, and western zone) in mainland China, the check-up population may not be entirely representative of the general population, which may have biased our study towards a healthier group and thus limit the applicability and generalizability of our model to the broader population. Fifth, the cross-sectional design of our study means that the temporal relationship between the predictors and the outcome cannot be established, and there may be reverse causation or confounding effects that we have not accounted for.

In conclusion, we developed a prediction model that uses a set of routine and quantitative variables obtained from health checkup programs to estimate the risk of carotid plaque in individuals with fatty liver disease. The resulting model is cost-effective, easy to use, and demonstrated strong predictive performance. This approach provides a means for personalized risk assessment of carotid plaque and derives robust cutoff points for carotid plaque risk stratification, with potential implications for improving the cost-effectiveness of carotid ultrasound detection.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (grants 82192901, 82192904, and 82192900), Ministry of Science and Technology of the People's Republic of China (grant 2020YFC2003400), the National Natural Science Foundation of China (grant 91846303), and the Ministry of Science and Technology of the People's Republic of China (grant 2020YFC2004703). The funders had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Data Availability

The data are not publicly available as they are individual-level health check-up data**,** but are available from the corresponding author on reasonable request.

Authors' Contributions

YD contributed to the study design, data analysis, and manuscript drafting. YM and JF contributed to the study design and manuscript revision. XW and CY contributed to manuscript review and editing. JL contributed to funding acquisition. SM, BW, and LL contributed to manuscript review, editing, and funding acquisition. SM, BW, and LL contributed equally to the paper as co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of study participants in the external validation data set or characteristics of study participants in the development data set and external validation data set.
[DOCX File , 28 KB - publichealth_v9i1e47095_app1.docx]

References

1. Liu J, Zhou X, Lin H, Lu X, Zheng J, Xu E, et al. Deep learning based on carotid transverse B-mode scan videos for the diagnosis of carotid plaque: a prospective multicenter study. Eur Radiol 2023;33(5):3478-3487. [doi: 10.1007/s00330-022-09324-y] [Medline: 36512047]
2. Zhang Y, Wu Z, Li X, Wei J, Zhang Q, Wang J. Association between the triglyceride-glucose index and carotid plaque incidence: a longitudinal study. Cardiovasc Diabetol 2022;21(1):244 [FREE Full text] [doi: 10.1186/s12933-022-01683-6] [Medline: 36380351]
3. Xie J, Li Y, Xu X, Wei J, Li H, Wu S, et al. CPTV: classification by tracking of carotid plaque in ultrasound videos. Comput Med Imaging Graph 2023;104:102175 [FREE Full text] [doi: 10.1016/j.compmedimag.2022.102175] [Medline: 36630795]
4. Ding X, Wang X, Wu J, Zhang M, Cui M. Triglyceride-glucose index and the incidence of atherosclerotic cardiovascular diseases: a meta-analysis of cohort studies. Cardiovasc Diabetol 2021;20(1):76 [FREE Full text] [doi: 10.1186/s12933-021-01268-9] [Medline: 33812373]
5. Zhao W, Wu Y, Shi M, Bai L, Tu J, Guo Z, et al. Sex differences in prevalence of and risk factors for carotid plaque among adults: a population-based cross-sectional study in Rural China. Sci Rep 2016;6:38618 [FREE Full text] [doi: 10.1038/srep38618] [Medline: 27922121]
6. Saba L, Saam T, Jäger HR, Yuan C, Hatsukami TS, Saloner D, et al. Imaging biomarkers of vulnerable carotid plaques for stroke risk prediction and their potential clinical implications. Lancet Neurol 2019;18(6):559-572. [doi: 10.1016/S1474-4422(19)30035-3] [Medline: 30954372]
7. Jiao Y, Qin Y, Zhang Z, Zhang H, Liu H, Li C. Early identification of carotid vulnerable plaque in asymptomatic patients. BMC Cardiovasc Disord 2020;20(1):429 [FREE Full text] [doi: 10.1186/s12872-020-01709-5] [Medline: 33003997]
8. Clarke R, Du H, Kurmi O, Parish S, Yang M, Arnold M, China Kadoorie Biobank Collaborative Group. Burden of carotid artery atherosclerosis in Chinese adults: implications for future risk of cardiovascular diseases. Eur J Prev Cardiol 2017;24(6):647-656 [FREE Full text] [doi: 10.1177/2047487317689973] [Medline: 28128654]
9. Weiwei C, Runlin G, Lisheng L, Manlu Z, Wen W, Yongjun W, et al. Outline of the report on cardiovascular diseases in China, 2014. Eur Heart J Suppl 2016;18(Suppl F):F2-F11 [FREE Full text] [doi: 10.1093/eurheartj/suw030] [Medline: 28533724]
10. Guo YC, Zhou Y, Gao X, Yao Y, Geng B, Cui QH, et al. Association between nonalcoholic fatty liver disease and carotid artery disease in a community-based Chinese population: a cross-sectional study. Chin Med J (Engl) 2018;131(19):2269-2276 [FREE Full text] [doi: 10.4103/0366-6999.241797] [Medline: 30246712]
11. Kozakova M, Palombo C, Eng MP, Dekker J, Flyvbjerg A, Mitrakou A, RISC Investigators. Fatty liver index, gamma-glutamyltransferase, and early carotid plaques. Hepatology 2012;55(5):1406-1415 [FREE Full text] [doi: 10.1002/hep.25555] [Medline: 22334565]

12. Mandraffino G, Morace C, Franzè MS, Nassisi V, Sinicropi D, Cinquegrani M, et al. Fatty liver as potential biomarker of atherosclerotic damage in familial combined hyperlipidemia. Biomedicines 2022;10(8):1770 [FREE Full text] [doi: 10.3390/biomedicines10081770] [Medline: 35892670]
13. Wu D, Cui G, Huang X, Chen Y, Liu G, Ren L, et al. An accurate and explainable ensemble learning method for carotid plaque prediction in an asymptomatic population. Comput Methods Programs Biomed 2022;221:106842. [doi: 10.1016/j.cmpb.2022.106842] [Medline: 35569238]
14. Grant EG, Benson CB, Moneta GL, Alexandrov AV, Baker JD, Bluth EI, et al. Carotid artery stenosis: gray-scale and doppler US diagnosis--society of radiologists in ultrasound consensus conference. Radiology 2003;229(2):340-346. [doi: 10.1148/radiol.2292030516] [Medline: 14500855]
15. Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham study. Am J Cardiol 1976;38(1):46-51. [doi: 10.1016/0002-9149(76)90061-8] [Medline: 132862]
16. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham heart study. Circulation 2008;117(6):743-753 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.699579] [Medline: 18212285]
17. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. Eur Heart J 2017;38(23):1805-1814 [FREE Full text] [doi: 10.1093/eurheartj/ehw302] [Medline: 27436868]
18. Kakadiaris IA, Vrigkas M, Yen AA, Kuznetsova T, Budoff M, Naghavi M. Machine learning outperforms ACC / AHA CVD risk calculator in MESA. J Am Heart Assoc 2018;7(22):e009476 [FREE Full text] [doi: 10.1161/JAHA.118.009476] [Medline: 30571498]
19. Xi Y, Wang H, Sun N. Machine learning outperforms traditional logistic regression and offers new possibilities for cardiovascular risk prediction: a study involving 143,043 Chinese patients with hypertension. Front Cardiovasc Med 2022;9:1025705 [FREE Full text] [doi: 10.3389/fcvm.2022.1025705] [Medline: 36451926]
20. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018;319(13):1317-1318. [doi: 10.1001/jama.2017.18391] [Medline: 29532063]
21. Schauberger G, Tanaka LF, Berger M. A tree-based modeling approach for matched case-control studies. Stat Med 2023;42(5):676-692 [FREE Full text] [doi: 10.1002/sim.9637] [Medline: 36631256]
22. van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. BMC Med Res Methodol 2014;14:137 [FREE Full text] [doi: 10.1186/1471-2288-14-137] [Medline: 25532820]
23. Stein JH, Korcarz CE, Hurst RT, Lonn E, Kendall CB, Mohler ER, American Society of Echocardiography Carotid Intima-Media Thickness Task Force. Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: a consensus statement from the American society of echocardiography carotid intima-media thickness task force. endorsed by the society for vascular medicine. J Am Soc Echocardiogr 2008;21(2):93-111 [FREE Full text] [doi: 10.1016/j.echo.2007.11.011] [Medline: 18261694]
24. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition. Hillsdale NJ: Lawrence Erlbaum Associates; 1988:1-17.
25. Song P, Xia W, Zhu Y, Wang M, Chang X, Jin S, et al. Prevalence of carotid atherosclerosis and carotid plaque in Chinese adults: a systematic review and meta-regression analysis. Atherosclerosis 2018;276:67-73. [doi: 10.1016/j.atherosclerosis.2018.07.020] [Medline: 30036743]
26. Azcui Aparicio RE, Carrington MJ, Ball J, Abhayaratna W, Stewart S, Haluska B, et al. Association of traditional risk factors with carotid intima-media thickness and carotid plaque in asymptomatic individuals with a family history of premature cardiovascular disease. Int J Cardiovasc Imaging 2021;38:739-749. [doi: 10.1007/s10554-021-02459-x] [Medline: 34731395]
27. Dong S, Gao J, Wang C, Liu J, Gu H, Tu J, et al. Association between blood pressure components and the presence of carotid plaque among adults aged 45 years and older: a population-based cross-sectional study in rural China. Blood Press Monit 2019;24(5):234-240 [FREE Full text] [doi: 10.1097/MBP.0000000000000396] [Medline: 31469693]
28. Sturlaugsdottir R, Aspelund T, Bjornsdottir G, Sigurdsson S, Thorsson B, Eiriksdottir G, et al. Prevalence and determinants of carotid plaque in the cross-sectional REFINE-reykjavik study. BMJ Open 2016;6(11):e012457 [FREE Full text] [doi: 10.1136/bmjopen-2016-012457] [Medline: 27884845]
29. Selwaness M, Hameeteman R, Van 't Klooster R, Van den Bouwhuijsen Q, Hofman A, Franco OH, et al. Determinants of carotid atherosclerotic plaque burden in a stroke-free population. Atherosclerosis 2016;255:186-192 [FREE Full text] [doi: 10.1016/j.atherosclerosis.2016.10.030] [Medline: 27806835]
30. Jiang SM, Sun XF, Gu HX, Chen YS, Xi CS, Qiao X, et al. Effects of decline in renal function with age on the outcome of asymptomatic carotid plaque in healthy adults: a 5-year follow-up study. Chin Med J (Engl) 2012;125(15):2649-2657. [Medline: 22931970]
31. Zhou Z, Ong KL, Whelton SP, Allison MA, Curtis AJ, Blaha MJ, et al. Impact of blood lipids on 10-year cardiovascular risk in individuals without dyslipidemia and with low risk factor burden. Mayo Clin Proc 2022;97(10):1883-1893. [doi: 10.1016/j.mayocp.2022.03.025] [Medline: 35760597]

32. Lind L. Flow-mediated vasodilation was found to be an independent predictor of changes in the carotid plaque status during a 5-year follow-up. J Atheroscler Thromb 2014;21(2):161-168 [FREE Full text] [doi: 10.5551/jat.18572] [Medline: 24126123]
33. Wu J, Wang Y, Kang K, Wang A, Xu J, Zhao X. Association between cumulative exposure to different lipid parameters and risk of newly developed carotid plaque. Stroke Vasc Neurol 2021;6(3):359-365 [FREE Full text] [doi: 10.1136/svn-2020-000430] [Medline: 33468640]
34. Preveden T, Veres B, Ruzic M, Pete M, Bogic S, Kovacevic N, et al. Triglyceride-glucose index and hepatic steatosis index for the assessment of liver steatosis in HCV patients. Minerva Gastroenterol (Torino) 2023;69(2):254-260 [FREE Full text] [doi: 10.23736/S2724-5985.22.03168-0] [Medline: 35343668]
35. Lee JH, Kim D, Kim HJ, Lee CH, Yang JI, Kim W, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. Dig Liver Dis 2010;42(7):503-508. [doi: 10.1016/j.dld.2009.08.002] [Medline: 19766548]
36. Wang C, Cai Z, Deng X, Li H, Zhao Z, Guo C, et al. Association of hepatic steatosis index and fatty liver index with carotid atherosclerosis in type 2 diabetes. Int J Med Sci 2021;18(14):3280-3289 [FREE Full text] [doi: 10.7150/ijms.62010] [Medline: 34400897]
37. Zhang Y, Fu X, Xie X, Yan D, Wang Y, Huang W, et al. A novel model forecasting perioperative red blood cell transfusion. Sci Rep 2022;12(1):16127 [FREE Full text] [doi: 10.1038/s41598-022-20543-7] [Medline: 36167791]
38. Lange T, Backhaus SJ, Schulz A, Evertz R, Kowallick JT, Bigalke B, et al. Cardiovascular magnetic resonance-derived left atrioventricular coupling index and major adverse cardiac events in patients following acute myocardial infarction. J Cardiovasc Magn Reson 2023;25(1):24 [FREE Full text] [doi: 10.1186/s12968-023-00929-w] [Medline: 37046343]
39. Wang Z, Zhang L, Li S, Xu F, Han D, Wang H, et al. The relationship between hematocrit and serum albumin levels difference and mortality in elderly sepsis patients in intensive care units-a retrospective study based on two large database. BMC Infect Dis 2022;22(1):629 [FREE Full text] [doi: 10.1186/s12879-022-07609-7] [Medline: 35850582]
40. Yang D, Iyer S, Gardener H, Della-Morte D, Crisby M, Dong C, et al. Cigarette smoking and carotid plaque echodensity in the Northern Manhattan study. Cerebrovasc Dis 2015;40(3-4):136-143 [FREE Full text] [doi: 10.1159/000434761] [Medline: 26227885]
41. Hu J, Su F, Ren X, Cao L, Zhou Y, Fu Y, et al. Prediction of carotid plaque by blood biochemical indices and related factors based on fisher discriminant analysis. BMC Cardiovasc Disord 2022;22(1):371 [FREE Full text] [doi: 10.1186/s12872-022-02806-3] [Medline: 35965318]

Abbreviations

ALP: alkaline phosphatase
ALT: alanine transaminase
AST: aspartate aminotransferase
AUROC: area under the receiver operating characteristic curve
EN: elastic net
FBG: fasting blood glucose
HSI: Hepatic Steatosis Index
LDL-C: low-density lipoprotein cholesterol
ML: machine learning
PLT: blood platelet count
RF: random forest
SBP: systolic blood pressure
SMD: standardized mean difference
T2DM: type 2 diabetes mellitus
TC: total cholesterol
XGBoost: extreme gradient boosting

Edited by A Mavragani, T Sanchez; submitted 07.03.23; peer-reviewed by N Jiwani, T Yang, S Hulse; comments to author 11.04.23; revised version received 28.04.23; accepted 25.07.23; published 07.09.23.

Please cite as:
Deng Y, Ma Y, Fu J, Wang X, Yu C, Lv J, Man S, Wang B, Li L
Combinatorial Use of Machine Learning and Logistic Regression for Predicting Carotid Plaque Risk Among 5.4 Million Adults With Fatty Liver Disease Receiving Health Check-Ups: Population-Based Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e47095
URL: https://publichealth.jmir.org/2023/1/e47095
doi:10.2196/47095
PMID:37676713

Original Paper

L-Shaped Association of Serum Chloride Level With All-Cause and Cause-Specific Mortality in American Adults: Population-Based Prospective Cohort Study

Xinran Hou[1,2], MD, PhD; Wei Xu[3], MD; Chengliang Zhang[2,4], MD, PhD; Zongbin Song[1,2], MD, PhD; Maoen Zhu[1,2], MD; Qulian Guo[1,2], MD, PhD; Jian Wang[1,2], MD, PhD

[1]Department of Anesthesiology, Xiangya Hospital, Central South University, Changsha, China
[2]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China
[3]Department of Anesthesiology, Hunan Provincial Maternal and Child Health Care Hospital, University of South China, Changsha, China
[4]Department of Cardiovascular Surgery, Xiangya Hospital, Central South University, Changsha, China

Corresponding Author:
Jian Wang, MD, PhD
Department of Anesthesiology
Xiangya Hospital
Central South University
87 Xiangya road
Changsha, 410008
China
Phone: 86 13739077976
Email: jianwang@csu.edu.cn

Abstract

Background: Chloride is the most abundant anion in the human extracellular fluid and plays a crucial role in maintaining homeostasis. Previous studies have demonstrated that hypochloremia can act as an independent risk factor for adverse outcomes in various clinical settings. However, the association of variances of serum chloride with long-term mortality risk in general populations has been rarely investigated.

Objective: This study aims to assess the association of serum chloride with all-cause and cause-specific mortality in the general American adult population.

Methods: Data were collected from 10 survey cycles (1999-2018) of the National Health and Nutrition Examination Survey. All-cause mortality, cardiovascular disease (CVD) mortality, cancer mortality, and respiratory disease mortality data were obtained by linkage to the National Death Index through December 31, 2019. After adjusting for demographic factors and relevant lifestyle, laboratory items, and comorbid factors, weighted Cox proportional risk models were constructed to estimate hazard ratios and 95% CIs for all-cause and cause-specific mortality.

Results: A total of 51,060 adult participants were included, and during a median follow-up of 111 months, 7582 deaths were documented, 2388 of CVD, 1639 of cancer, and 567 of respiratory disease. The weighted Kaplan-Meier survival analyses showed consistent highest mortality risk in individuals with the lowest quartiles of serum chloride. The multivariate-adjusted hazard ratios from lowest to highest quartiles of serum chloride (≤101.2, 101.3-103.2, 103.2-105.0, and ≥105.1 mmol/L) were 1.00 (95% CI reference), 0.77 (95% CI 0.67-0.89), 0.72 (95% CI 0.63-0.82), and 0.77 (95% CI 0.65-0.90), respectively, for all-cause mortality (*P* for linear trend<.001); 1.00 (95% CI reference), 0.63 (95% CI 0.51-0.79), 0.56 (95% CI 0.43-0.73), and 0.67 (95% CI 0.50-0.89) for CVD mortality (*P* for linear trend=.004); 1.00 (95% CI reference), 0.67 (95% CI 0.54-0.84), 0.65 (95% CI 0.50-0.85), and 0.65 (95% CI 0.48-0.87) for cancer mortality (*P* for linear trend=.004); and 1.00 (95% CI reference), 0.68 (95% CI 0.41-1.13), 0.59 (95% CI 0.40-0.88), and 0.51 (95% CI 0.31-0.84) for respiratory disease mortality (*P* for linear trend=.004). The restricted cubic spline analyses revealed the nonlinear and L-shaped associations of serum chloride with all-cause and cause-specific mortality (all *P* for nonlinearity<.05), in which lower serum chloride was prominently associated with higher mortality risk. The associations of serum chloride with mortality risk were robust, and no significant additional interaction effect was detected for all-cause mortality and CVD mortality (*P* for interaction>.05).

Conclusions: In American adults, decreased serum chloride concentrations were independently associated with increased all-cause mortality, CVD mortality, cancer mortality, and respiratory disease mortality. Our findings suggested that serum chloride may serve as a promising cost-effective health indicator in the general adult population. Further studies are warranted to explore the potential pathophysiological mechanisms underlying the association between serum chloride and mortality.

(JMIR Public Health Surveill 2023;9:e49291) doi:10.2196/49291

KEYWORDS

serum chloride; all-cause mortality; cause-specific mortality; National Health and Nutrition Examination Survey; National Death Index

Introduction

Background

Chloride is the most abundant anion in human extracellular fluid, accounting for approximately one-third of plasma tonicity and two-thirds of all plasma negative charges [1], and is predominantly confined to the extracellular fluid compartment [2]. Serum chloride plays an important role in many physiological processes in the body, including the maintenance of osmotic pressure, electroneutrality of body fluids, muscular activity, and regulation of blood pressure [3,4]. Serum chloride levels are predominantly regulated by the gastrointestinal tract and kidney. Chloride is secreted in the gastric juice as hydrochloric acid and can be absorbed by the entire intestine during digestion. In the United States, the daily intake of chloride for adult men is 7.8 to 11.8 g, whereas that for adult women is 5.8 to 7.8 g [5]. The primary source of chloride in the body is table salt (sodium chloride) in the diet as well as salt-containing foods [6]. The kidney can filter 19,440 mmol of chloride every day, 99.1% of which is reabsorbed, and only 180 mmol of chloride is excreted into urine, maintaining serum chloride within a narrow range [7].

Many clinical situations are related to the dysregulation of serum chloride. Serum chloride can be reduced either by chloride depletion, such as vomiting, diarrhea, burns, diuretic use, renal diseases, adrenal insufficiency, and metabolic alkalosis, or by water retention, such as excessive water intake, congestive heart failure, liver cirrhosis, and nephrotic syndrome. In contrast, serum chloride can be elevated by the loss of hypotonic fluids (eg, fever), administration of hypertonic fluid, or metabolic acidosis [2]. Chloride is a crucial constituent in diagnostic tests across numerous clinical scenarios. However, unlike other serum electrolytes such as sodium, potassium, and calcium, the importance of serum chloride is often overlooked despite its extensive measurement in routine blood biochemical tests. In recent years, serum chloride has exhibited a robust independent prognostic value, which seems even better than serum sodium, in patients with heart failure [8-11], chronic kidney disease (CKD) [12], and pulmonary arterial hypertension [13] and in patients who are critically ill [14-16]. For instance, in patients diagnosed with congestive heart failure, both hypochloremia and hyperchloremia have been associated with higher rates of mortality and hospitalization [11,17]. Similarly, lower chloride levels have been associated with an increased risk of mortality and progression to end-stage renal disease in patients with CKD [12,14]. Lower chloride level at 6 months after pulmonary arterial hypertension could predict long-term mortality [13]. In addition, research conducted on patients who are critically ill has demonstrated that both low and high levels of chloride are associated with higher rates of mortality and dysfunction in organs [14-16]. However, most studies were conducted among specific patient populations (eg, patients who were hospitalized or participants with heart disease or kidney disease), and researchers mainly focused on patients with hypochloremia or hyperchloremia. Knowledge of the relationship between the generally distributed serum chloride and mortality from the community-based population remains limited and requires further clarification. Understanding the potential association between chloride levels and mortality in the general population is essential for identifying individuals at higher risk and providing preventive strategies.

Objective

To address this knowledge gap, we used a nationally representative cohort from the National Health and Nutrition Examination Survey (NHANES) 1999-2018 data set to investigate whether serum chloride was associated with all-cause and cause-specific mortality in the general American adult population.

Methods

Study Design and Participants Selection

The NHANES has been a continuous nationally representative survey investigating the health and nutritional status of the noninstitutionalized population of the United States since 1999 [18]. In this cohort study, we analyzed data in 10 cycles of NHANES from 1999-2018. Participants were restricted to those aged ≥18 years, not pregnant at the physical examination, and with eligible serum chloride measurement and mortality follow-up information.

Ethical Considerations

All participants in the study signed a written informed consent form, and the NHANES data collection was approved by the Research Ethics Review Board of the National Center for Health Statistics [19]. The NHANES has strict protocols and procedures in place to ensure confidentiality and protect against identification. As our study was a secondary data analysis, which lacked personal identifiers, and the data in NHANES are freely accessible to the public on the web [20], it did not require review by an institutional review board [19].

Measurement of Serum Chloride Concentration

Blood samples were collected and processed in the mobile examination center by certified laboratory professionals and then stored in biorepositories. Serum chloride was determined as part of the routine biochemistry profile with Beckman Synchron LX20 or Beckman Coulter UniCel DxC800 (Beckman

Coulter, Inc), both of which used indirect (or diluted) ion-selective electrode methodology for the determination of serum chloride concentrations. Details about the laboratory procedure have been described on the NHANES website [20].

Assessment of Covariates

The sociodemographic characteristics, namely, sex, age, race, education level, marital status, and family income-to-poverty ratio (PIR), were collected using standardized questionnaires during the household interview, as previously described [21].

Lifestyle factors were evaluated using a combination of data from physical examinations and questionnaires. BMI was calculated as weight (kg) divided by the square of height (m). Smoking status was classified as never (smoked <100 cigarettes/lifetime), former (≥100 cigarettes/lifetime but do not currently smoke), and current smokers. Alcohol users were classified as never (<12 drinks in a lifetime), former (≥12 drinks in a lifetime but did not drink in the last year), mild (≤1 drink/d for women or ≤2 drinks/d for men), moderate (≤2 drinks/d for women or ≤3 drinks/d for men), and heavy drinkers (>2 drinks/d for women or >3 drinks/d for men), where a drink referred to an alcoholic drink equivalent of 12 oz of beer, 4 oz of wine, or 1 oz of liquor (such as whiskey or gin) [22]. Diet quality was assessed by the Healthy Eating Index-2015 (HEI-2015), reflecting adherence to the 2015 to 2020 Dietary Guidelines for Americans [23]. Physical activity was quantified in metabolic equivalent tasks (METs) multiplied by exercise time (minutes) per week (METs min/wk) [24].

For laboratory variables, serum sodium, potassium, and bicarbonate levels were measured similarly to serum chloride levels, as described in the aforementioned method section. The estimated glomerular filtration rate (eGFR) was calculated using the CKD Epidemiology Collaboration equation (2009) [25].

Comorbid conditions were evaluated by combined information from NHANES data sets, as previously described [21]. Hypertension was determined by systolic blood pressure≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, having been diagnosed by a physician, or current use of antihypertensive drugs. Diabetes was determined by fasting glucose level ≥7.0 mmol/L, random blood glucose level ≥11.1 mmol/L, Hemoglobin A1c level ≥6.5%, 2-hour glucose level ≥11.1 mmol/L after oral glucose tolerance tests, having been diagnosed by a physician or current use of the antidiabetic drug. Chronic obstructive pulmonary disease (COPD) was determined by first second of expiratory volume/forced vital capacity <0.7 after the use of bronchodilators in spirometry, having been diagnosed by a physician, or current use of drugs for COPD. Coronary heart disease (CHD), stroke, and cancer were assessed based on whether the participants had been diagnosed with these conditions by a doctor.

Ascertainment of All-Cause and Cause-Specific Mortality

Mortality status for people who took part in the NHANES study was ascertained through probabilistic matching with the National Death Index Public-Use Linked Mortality Files up to December 31, 2019 [26]. The primary outcome was all-cause and cause-specific mortality, classified according to the international classification of diseases-10 codes recorded as the leading cause of death, in which cardiovascular disease (CVD) mortality corresponded to I00-I09, I11, I13, I20-I51, and I60-I69; cancer mortality corresponded to C00-C97; and respiratory disease mortality corresponded to J09-J18 and J40-J47.

In this study, the survival time has been calculated as months from the date of examination to the date of death or the end of follow-up (December 31, 2019), whichever came first.

Statistical Analysis

Sample weights, clustering, and stratification were used in the data analyses to account for the stratified, multistage probability design of NHANES, in which the combined sample weight was calculated as 2/5×4-year sample weight for examinations in mobile examination centers for the period from 1999 to 2002 and 1/5×2-year sample weight for examinations in mobile examination centers for other survey cycles according to the NHANES analytic guidelines [27,28].

Data are presented as the survey-weighted mean (SE) for continuous variables or frequency (survey-weighted percentage) for categorical variables. Differences in baseline information were detected using survey-weighted ANOVA or chi-square test.

Survey-weighted Kaplan-Meier survival curves were plotted, and overall log-rank tests and pairwise comparisons between group levels with Benjamini and Hochberg corrections [29] were used to compare the cumulative survival in different serum chloride quartiles. Survey-weighted Cox proportional hazards models were used to estimate the hazard ratios (HRs) and 95% CIs for the association of serum chloride quartiles with all-cause and cause-specific mortality, with the first quartile as the reference, and the linear trend was evaluated by assigning a median value to each category as a continuous variable. The crude and multivariable models with different covariate adjustments were constructed to estimate mortality risk according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline [30], in which sex, age, and race were adjusted in model 1; education, marital status, PIR, BMI, smoking, alcohol use, HEI-2015, and physical activity were additionally adjusted in model 2; serum sodium, serum potassium, serum bicarbonate, eGFR, and use of diuretics as well as comorbidity of hypertension, diabetes, CHD, stroke, COPD, and cancer were additionally adjusted in model 3. Moreover, serum chloride was calculated as a continuous variable to estimate the mortality risk with each millimole per liter increase of serum chloride in different multivariate models. The proportion of missing data in the covariates was summarized, and in the primary analyses, the listwise deletion for missing variables was adopted for multivariate models.

To evaluate the potential nonlinear association of mortality with serum chloride, the restricted cubic spline analysis was performed in multivariable-adjusted models, in which the number of knots was determined based on the Akaike information criterion.

The subgroup analyses were performed according to the dichotomic stratified variables, namely, sex; age at recruitment;

race; education level; marital status; PIR; BMI; smoker; drinker; HEI-2015; physical activity (physically active [≥600 MET-min/wk] vs physically inactive [<600 MET-min/wk]); and comorbidity of hypertension, diabetes, stroke, COPD, cancer, and CKD, determined by eGFR ≤60 mL/min/1.73 m^2 or albuminuria ≥30 mg/g based on the Kidney Diseases Improving Global Outcomes guideline [31], using the fully adjusted model except for the specific stratification variable. The likelihood ratio test also inspected interactions of serum chloride with the stratification variables.

Robustness of the results was examined using several sensitivity analyses. First, participants who died within the first 2 years of follow-up were excluded to reduce potential reverse causation. Second, participants with extremely low or high serum chlorides (<1.5×IQR below the lower quartile or >1.5×IQR above the upper quartile) were excluded to test the effect of potential outliers. Third, participants with hypoalbuminemia (serum albumin level <35 g/L) were excluded to evaluate the impact of malnutrition on the relation between serum chloride and mortality. Fourth, hypochloremia has been confirmed to be associated with higher mortality in patients with heart failure; therefore, we additionally adjusted the history of congestive heart failure. Fifth, potassium-sparing diuretics and nonpotassium-sparing diuretics may have different effects on serum chloride concentration and outcomes [17], so we separately adjusted these 2 types of diuretics. Sixth, we used the multiple imputation method to impute all missing covariates to test the influence of these missing variables. Finally, to explore the potential influence of unmeasured confounders on our risk estimates, we performed an E-value analysis, which quantifies the required magnitude of an unmeasured confounder that would negate the observed association between serum chloride and mortality [32].

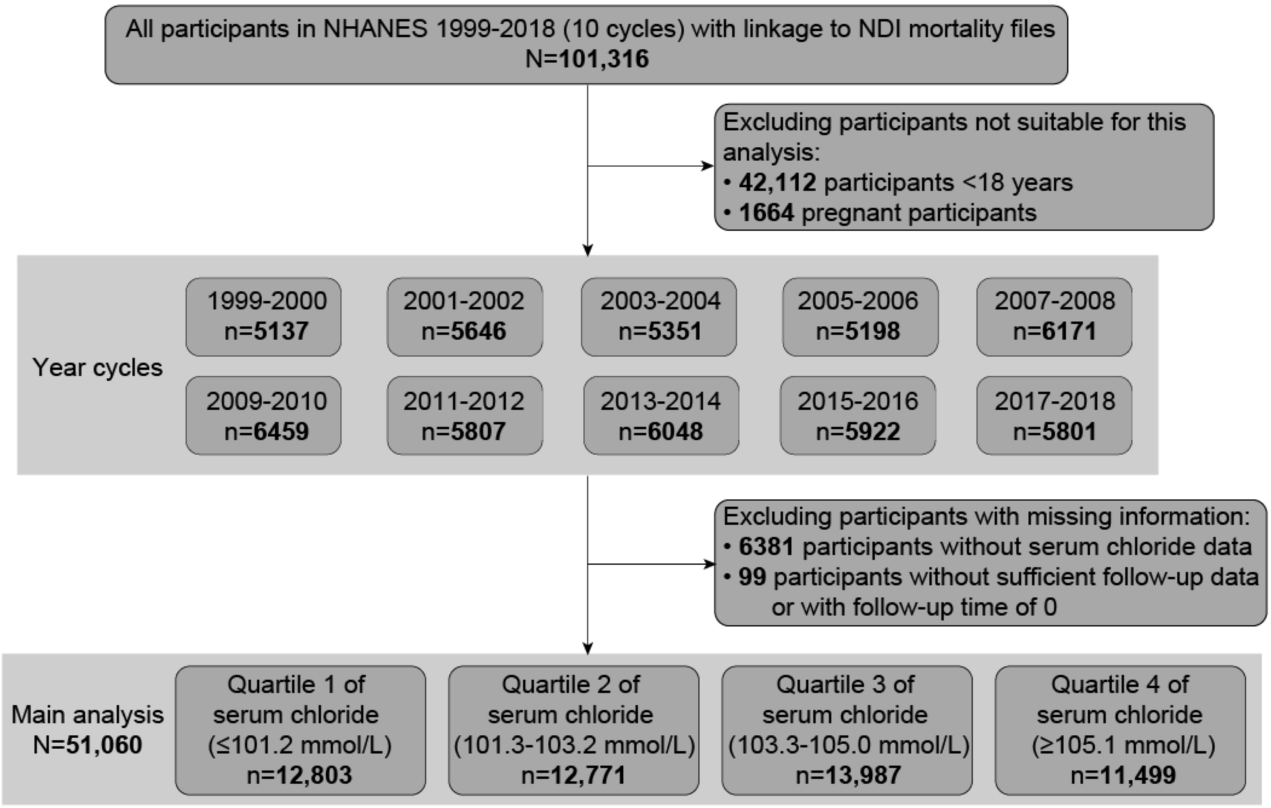
All statistical analyses were performed using the R program (version 4.2.1; R Foundation for Statistical Computing [33]) with package “survey” (version 4.1-1) for analysis of complex survey samples, package “survminer” (version 0.4.9) for survival curves plotting, package “rms” (version 6.6-0) for restricted cubic spline analysis, package “mice” (version 3.15.0) for multiple imputation, and package “Evalue” (version 4.1.3) for sensitivity analyses for unmeasured confounding and other biases. A 2-sided *P*<.05 was considered statistically significant.

Results

Selection of Participants

Of the total 101,316 participants in 10 cycles of NHANES (1999-2018), we excluded 42,112 participants aged <18 years, 1664 pregnant participants, 6381 participants without serum chloride data, and 99 participants without eligible follow-up data. The remaining 51,060 eligible participants were included in the final analysis (Figure 1).

Figure 1. Flowchart of participant selection from the US National Health and Nutrition Examination Survey (NHANES) 1999-2018. NDI: National Death Index.



Baseline Characteristics of Participants

Among the 51,060 participants, representing 209 million American adults, 25,632 (50.95%) were female; mean age was 46.4 (SD 16.6) years; and 22,199 (68.48%) were non-Hispanic White. The median serum chloride of the participants was 103.2 (IQR 101.2-105.0) mmol/L, and according to the quartiles, serum chloride values (mmol/L) were categorized into 4 groups: quartile 1 (Q1; ≤101.2), quartile 2 (Q2; 101.3-103.2), quartile 3 (Q3; 103.3-105.0), and quartile 4 (Q4; ≥105.1).

During a median follow-up time of 111 (IQR 59-169) months, overall 7582 deaths were documented, of which 1958 (24.76%) were attributed to diseases of the heart, 1639 (22.72%) were attributed to malignant neoplasms, 430 (5.02%) were attributed to cerebrovascular diseases, 409 (6.16%) were attributed to chronic lower respiratory diseases, and 158 (1.85%) were attributed to influenza and pneumonia, and the leading causes of death were reorganized into CVD, cancer, and respiratory disease in the following analyses (Multimedia Appendix 1).

The population baseline characteristic profiles were compared across the quartiles of serum chloride (Table 1). Briefly, individuals with lower serum chloride were more likely to be older, male, non-Hispanic White, and nonmarried; to have higher HEI-2015, lower serum sodium, higher serum bicarbonate, lower eGFR, more use of diuretics; and to have a higher prevalence of hypertension, diabetes, CHD, stroke, COPD, and cancer (Table 1).

Table 1. Baseline characteristics of participants in different serum chloride quartiles from the US National Health and Nutrition Examination Survey 1999-2018 (N=51,060).

	Total	Q1[a] (≤101.2)	Q2[b] (101.3-103.2)	Q3[c] (103.3-105.0)	Q4[d] (≥105.1)	*P* value
Sample size, n	51,060	12,803	12,771	13,987	11,499	N/A[e]
Sex, n (%)						<.001
Female	25,632 (50.95)	5722 (45.94)	6006 (48.16)	7337 (53.06)	6567 (57.48)	
Male	25,428 (49.05)	7081 (54.06)	6765 (51.84)	6650 (46.94)	4932 (42.52)	
Age (years), mean (SD)	46.4 (16.6)	49.1 (17.2)	45.7 (16.4)	45.4 (16.3)	45.5 (16.1)	<.001
Race, n (%)						<.001
Black	10,619 (10.7)	2463 (9.6)	2590 (10.11)	2906 (10.79)	2660 (12.59)	
Mexican	9388 (8.32)	2170 (7.11)	2318 (8.07)	2689 (8.87)	2211 (9.36)	
White	22,199 (68.48)	5776 (70.01)	5601 (68.86)	6041 (68.64)	4781 (66.04)	
Other	8854 (12.49)	2394 (13.28)	2262 (12.97)	2351 (11.7)	1847 (12.02)	
Education, n (%)						.003
Below high school	12,867 (16.71)	3335 (17.9)	3057 (16.26)	3441 (16.74)	3034 (18.53)	
High school or equivalent	10,994 (23.23)	2785 (24.9)	2706 (23.83)	2941 (23.51)	2562 (23.97)	
College or above	23,630 (56.67)	5878 (57.19)	6031 (59.92)	6578 (59.75)	5143 (57.5)	
Marital status, n (%)						<.001
Never married	9906 (17.95)	2433 (18.69)	2638 (19.42)	2716 (18.06)	2119 (17.79)	
Married	24,924 (53.92)	6124 (53.74)	6212 (55.81)	6989 (57.07)	5599 (55.66)	
Other	14,210 (25.11)	3825 (27.57)	3414 (24.77)	3736 (24.87)	3235 (26.55)	
Family income-to-poverty ratio, mean (SD)	2.98 (1.63)	2.97 (1.63)	3.04 (1.63)	3.01 (1.63)	2.89 (1.64)	<.001
BMI (kg/m^2), mean (SD)	28.7 (6.5)	28.6 (6.6)	28.3 (6.2)	28.5 (6.4)	29.3 (6.7)	<.001
Smoker, n (%)						<.001
Never	26,271 (52.57)	6446 (52.63)	6683 (54.72)	7291 (55.03)	5851 (52.76)	
Former	11,903 (24.14)	3265 (26.15)	2928 (25.14)	3154 (23.86)	2556 (23.63)	
Now	10,110 (20.93)	2501 (21.22)	2396 (20.15)	2723 (21.1)	2490 (23.62)	
Alcohol user, n (%)						.001
Never	6538 (10.47)	1598 (11.42)	1612 (11.27)	1751 (11.64)	1577 (12.59)	
Former	7600 (12.87)	1813 (13.99)	1750 (13.14)	2139 (14.62)	1898 (15.94)	
Mild	14,440 (31.93)	3690 (36.19)	3771 (36.99)	3925 (35.72)	3054 (33.35)	
Moderate	6464 (15.15)	1497 (16.13)	1633 (17.34)	1849 (17.29)	1485 (16.84)	
Heavy	8742 (19.14)	2230 (22.26)	2188 (21.26)	2351 (20.73)	1973 (21.28)	
Healthy Eating Index-2015, mean (SD)	50.2 (13.6)	51.2 (14.0)	50.4 (13.9)	50.0 (13.5)	49.0 (13.0)	<.001
Physical activity (metabolic equivalent tasks min/wk), mean (SD)	3423 (5638.7)	3359.1 (5659.6)	3466.5 (5812.3)	3315.3 (5476.3)	3581.7 (5604.0)	.07
Serum sodium (mmol/L), mean (SD)	139.3 (2.3)	137.8 (2.5)	139.1 (2.0)	139.7 (1.9)	140.6 (1.9)	<.001
Serum potassium (mmol/L), mean (SD)	4.01 (0.33)	4.00 (0.36)	4.01 (0.32)	4.01 (0.32)	4.03 (0.31)	.009
Serum bicarbonate (mmol/L), mean (SD)	24.8(2.3)	25.5 (2.3)	25.1 (2.2)	24.7 (2.1)	23.9 (2.1)	<.001
Estimated glomerular filtration rate (mL/min/1.73 m^2), mean (SD)	94.7 (20.8)	92.6 (21.2)	94.9 (20.3)	95.8 (20.4)	95.4 (21.4)	<.001
Diuretics use, n (%)						<.001
No	44,379 (88.83)	9941 (80.45)	11,188 (89.51)	12,692 (92.54)	10,558 (93.47)	
Yes	6632 (11.1)	2838 (19.55)	1573 (10.49)	1290 (7.46)	931 (6.53)	

	Total	Q1[a] (≤101.2)	Q2[b] (101.3-103.2)	Q3[c] (103.3-105.0)	Q4[d] (≥105.1)	*P* value
Hypertension, n (%)						<.001
No	30,501 (63.73)	6380 (54.28)	7791 (64.88)	9003 (68.28)	7327 (67.77)	
Yes	20,545 (36.25)	6416 (45.72)	4977 (35.12)	4983 (31.72)	4169 (32.23)	
Diabetes, n (%)						<.001
No	42,533 (87.34)	9591 (81.06)	10,773 (88.11)	12,118 (89.91)	10,051 (90.54)	
Yes	8523 (12.66)	3210 (18.94)	1998 (11.89)	1869 (10.09)	1446 (9.46)	
Coronary heart disease , n (%)						<.001
No	45,305 (93.01)	11,280 (95.32)	11,285 (96.65)	12,451 (97.06)	10,289 (96.77)	
Yes	2040 (3.43)	674 (4.68)	478 (3.35)	466 (2.94)	422 (3.23)	
Stroke, n (%)						<.001
No	45,663 (93.89)	11,440 (96.58)	11,381 (97.39)	12,506 (97.47)	10,336 (97.25)	
Yes	1837 (2.73)	557 (3.42)	420 (2.61)	461 (2.53)	399 (2.75)	
Chronic obstructive pulmonary disease , n (%)						<.001
No	45,581 (92.92)	11,347 (95.19)	11,383 (96.37)	12,494 (96.24)	10,357 (96.32)	
Yes	2009 (3.85)	672 (4.81)	436 (3.63)	491 (3.76)	410 (3.68)	
Cancer, n (%)						<.001
No	43,107 (87.48)	10,702 (89.05)	10,787 (91.35)	11,805 (90.72)	9813 (91.13)	
Yes	4404 (9.14)	1310 (11.05)	1016 (8.65)	1154 (9.28)	924 (8.87)	

[a] Q1: quartile 1.

[b] Q2: quartile 2.

[c] Q3: quartile 3.

[d] Q4: quartile 4.

[e] N/A: not applicable.

Associations of Serum Chloride Quartiles With All-Cause and Cause-Specific Mortality

The unadjusted survey-weighted Kaplan-Meier survival analyses, followed by the log-rank test and pairwise comparisons between groups with corrections, showed that individuals with the lowest quartiles of serum chloride consistently bear the highest mortality risk (Figure 2).

After multivariable adjustment, the all-cause mortality risk was lower in the Q2-Q4 group of serum chloride than in the Q1 group. Briefly, compared with the reference group (Q1), the HRs for all-cause mortality were estimated to be 0.77 (95% CI 0.67-0.89) for the Q2 group, 0.72 (95% CI 0.63-0.82) for the Q3 group, and 0.77 (95% CI 0.65-0.90) for the Q4 group (*P* for linear trend<.001; Table 2). As a continuous linear variable, every millimole per liter increment in serum chloride was associated with a 4% reduced risk of all-cause mortality (HR 0.96, 95% CI 0.94-0.98; Multimedia Appendix 2).

Figure 2. Survey-weighted cumulative hazard curves and log-rank tests comparing mortality risks owing to all-cause (A), cardiovascular disease (CVD; B), cancer (C), and respiratory disease (D) for individuals in different serum chloride quartiles from the US National Health and Nutrition Examination Survey 1999-2018. Q1: quartile 1; Q2: quartile 2; Q3: quartile 3; Q4: quartile 4.

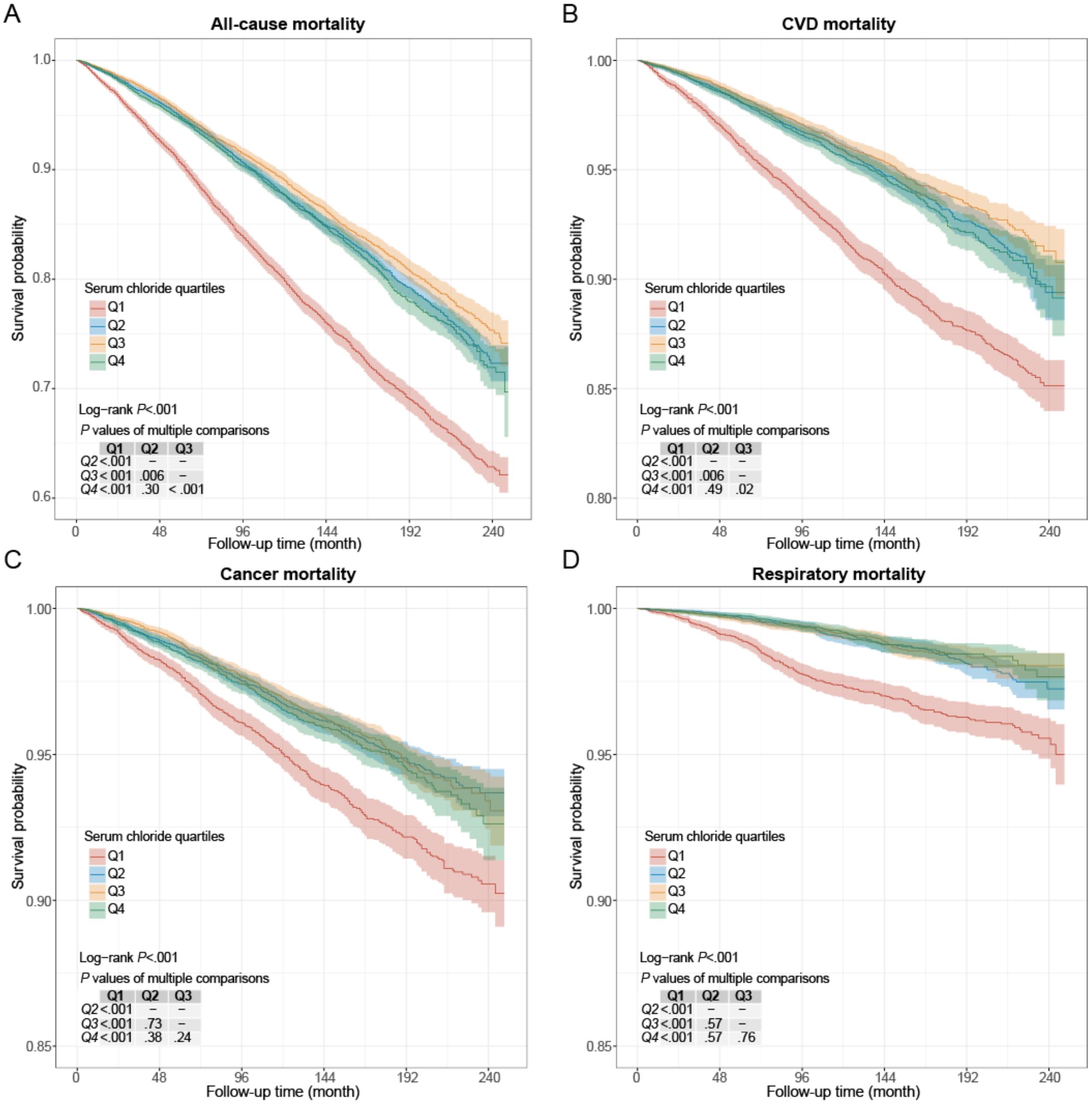


Table 2. Survey-weighted multivariate analyses of the associations of categorical serum chloride serum chloride with all-cause and cause-specific mortality for adults from the US National Health and Nutrition Examination Survey 1999-2018 (N=51,060)[a].

	Q1[b] (≤101.2)	Q2[c] (101.3-103.2)		Q3[d] (103.3-105.0)		Q4[e] (≥105.1)		*P* value for trend[f]
	HR[g] (95% CI)	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	
All-cause mortality								
Crude	1.00 (reference)	0.58 (0.53-0.63)	<.001	0.51 (0.47-0.56)	<.001	0.59 (0.53-0.65)	<.001	<.001
Model 1[h]	1.00 (reference)	0.71 (0.65-0.76)	<.001	0.65 (0.60-0.70)	<.001	0.73 (0.67-0.79)	<.001	<.001
Model 2[i]	1.00 (reference)	0.73 (0.65-0.82)	<.001	0.67 (0.59-0.75)	<.001	0.72 (0.63-0.81)	<.001	<.001
Model 3[j]	1.00 (reference)	0.77 (0.67-0.89)	<.001	0.72 (0.63-0.82)	<.001	0.77 (0.65-0.90)	.001	<.001
Cardiovascular disease mortality								
Crude	1.00 (reference)	0.54 (0.47-0.62)	<.001	0.45 (0.39-0.53)	<.001	0.55 (0.46-0.65)	<.001	<.001
Model 1	1.00 (reference)	0.65 (0.58-0.74)	<.001	0.57 (0.49-0.66)	<.001	0.68 (0.58-0.80)	<.001	<.001
Model 2	1.00 (reference)	0.60 (0.50-0.72)	<.001	0.53 (0.42-0.66)	<.001	0.65 (0.53-0.81)	<.001	<.001
Model 3	1.00 (reference)	0.63 (0.51-0.79)	<.001	0.56 (0.43-0.73)	<.001	0.67 (0.50-0.89)	.006	.004
Cancer mortality								
Crude	1.00 (reference)	0.55 (0.46-0.66)	<.001	0.58 (0.49-0.70)	<.001	0.63 (0.51-0.76)	<.001	<.001
Model 1	1.00 (reference)	0.64 (0.53-0.76)	<.001	0.67 (0.56-0.81)	<.001	0.75 (0.62-0.90)	.003	.002
Model 2	1.00 (reference)	0.69 (0.54-0.86)	.001	0.67 (0.52-0.86)	.002	0.72 (0.58-0.91)	.005	.005
Model 3	1.00 (reference)	0.67 (0.54-0.84)	<.001	0.65 (0.50-0.85)	.002	0.65 (0.48-0.87)	.004	.004
Respiratory mortality								
Crude	1.00 (reference)	0.44 (0.34-0.57)	<.001	0.40 (0.31-0.51)	<.001	0.38 (0.28-0.50)	<.001	<.001
Model 1	1.00 (reference)	0.54 (0.41-0.70)	<.001	0.49 (0.39-0.62)	<.001	0.47 (0.36-0.62)	<.001	<.001
Model 2	1.00 (reference)	0.63 (0.39-1.02)	.06	0.57 (0.39-0.82)	.003	0.50 (0.32-0.81)	.004	.001
Model 3	1.00 (reference)	0.68 (0.41-1.13)	.14	0.59 (0.40-0.88)	.009	0.51 (0.31-0.84)	.008	.004

[a] Data were fitted to a multivariate Cox proportional hazards model for data from a complex survey design.

[b] Q1: quartile 1.

[c] Q2: quartile 2.

[d] Q3: quartile 3.

[e] Q4: quartile 4.

[f] The test for trend was based on the variable containing the median value for each quartile.

[g] HR: hazard ratio.

[h] Model 1: adjusted for sex, age, and race.

[i] Model 2: adjusted for sex, age, race, education, marital status, family income-to-poverty ratio, BMI, smoking, alcohol use, Healthy Eating Index-2015, and physical activity.

[j] Model 3: adjusted for sex; age; race; education; marital status; family income-to-poverty ratio; BMI; smoking; alcohol use; Healthy Eating Index-2015; physical activity; serum sodium; serum potassium; serum bicarbonate; estimated glomerular filtration rate; use of diuretics; and comorbidity or history of hypertension, diabetes, coronary heart disease, stroke, chronic obstructive pulmonary disease, and cancer.

Similarly, CVD mortality was consistently lower in the Q2-Q4 group of serum chloride than in the Q1 group in all models. After being fully adjusted, the HRs for CVD mortality were estimated to be 0.63 (95% CI 0.51-0.79) for the Q2 group, 0.56 (95% CI 0.43-0.73) for the Q3 group, and 0.67 (95% CI 0.50-0.89) for the Q4 group compared with the Q1 group (*P* for linear trend=.004; Table 2). As a continuous linear variable, every millimole per liter increment in serum chloride was associated with a 7% reduced risk of CVD mortality (HR 0.93, 95% CI 0.89-0.97; Multimedia Appendix 2).

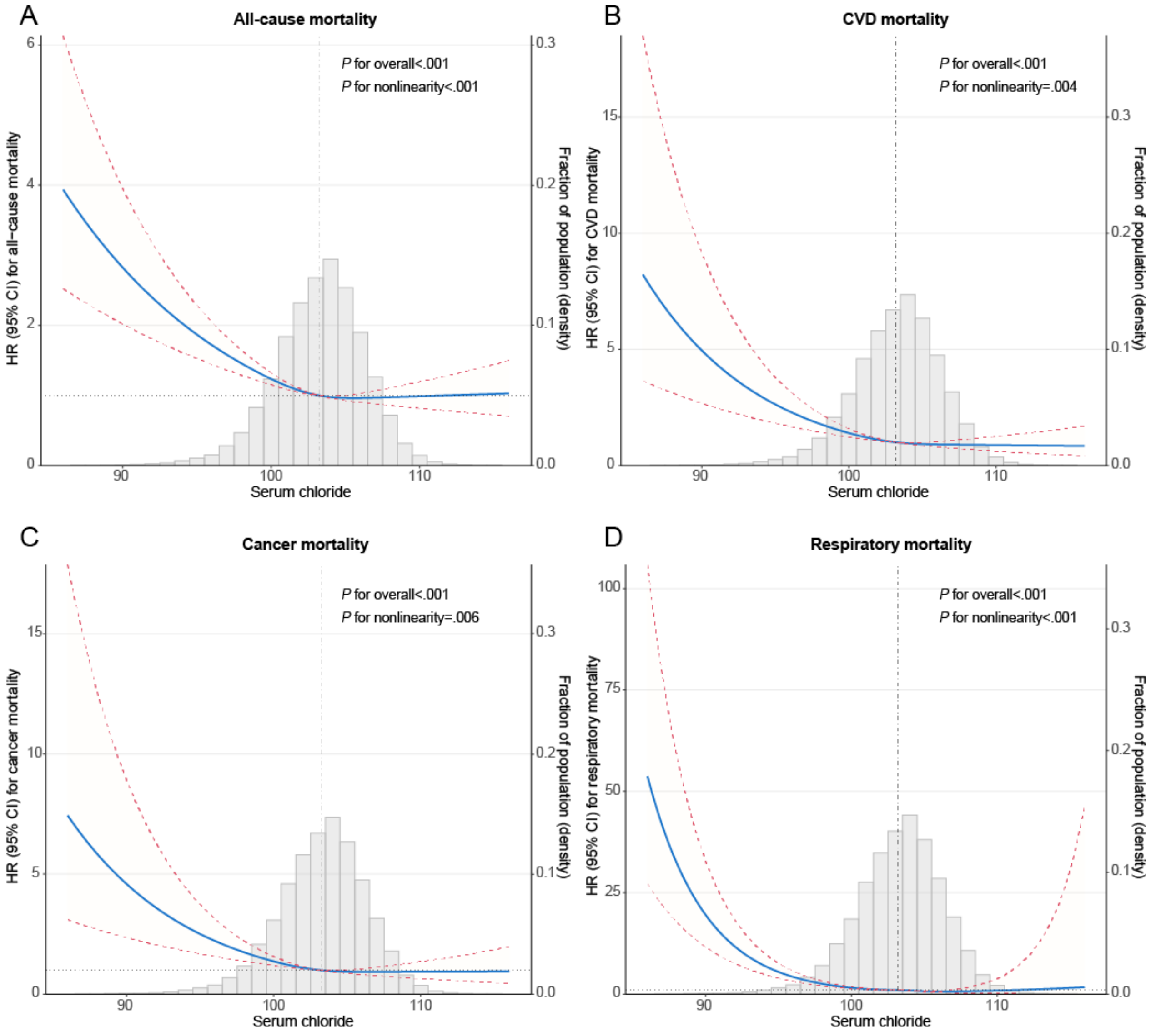
Moreover, cancer mortality was consistently lower in the Q2-Q4 group of serum chloride than in the Q1 group. With the total adjustment, the HRs for cancer mortality were estimated to be 0.67 (95% CI 0.54-0.84) for the Q2 group, 0.65 (95% CI 0.50-0.85) for the Q3 group, and 0.65 (95% CI 0.48-0.87) for the Q4 group compared with the Q1 group (*P* for linear trend=.004; Table 2). As a continuous linear variable, every millimole per liter increment in serum chloride was associated with a 6% reduced risk of cancer disease mortality (HR 0.94, 95% CI 0.89-0.98; Multimedia Appendix 2).

Similarly, consistently lower respiratory mortality was observed in Q3 and Q4 groups than in the Q1 group. With the full adjustment, the HRs for respiratory mortality were estimated to be 0.68 (95% CI 0.41-1.13) for the Q2 group, 0.59 (95% CI 0.40-0.88) for the Q3 group, and 0.51 (95% CI 0.31-0.84) for the Q4 group compared with the Q1 group (*P* for linear trend=.004; Table 2), and every millimole per liter increment in serum chloride was associated with a 15% reduced risk of respiratory mortality (HR 0.85, 95% CI 0.80-0.92; Multimedia Appendix 2).

The Nonlinear Analyses of the Association Between Continuous Serum Chloride and Mortality

Spline models with fully adjusted covariates were constructed to profile a more direct relationship between serum chloride and mortality. An L-shaped association was observed between serum chloride and all-cause mortality (nonlinear *P*<.001), in which all-cause mortality risk decreased steeply until around the median of serum chloride and then tended to be stable (Figure 3A). Similarly, L-shaped associations were observed between serum chloride and CVD mortality (nonlinear *P*=.004; Figure 3B), cancer mortality (nonlinear *P*=.006; Figure 3C), and respiratory mortality (nonlinear *P*<.001; Figure 3D).

Figure 3. Survey-weighted restricted cubic spline analyses of the associations of continuous serum chloride with all-cause mortality (A), cardiovascular disease (CVD) mortality (B), cancer mortality (C), and respiratory mortality (D), and the probability distribution histogram is represented in the background. All models were adjusted for sex; age; race; education; marital status; family income-to-poverty ratio; BMI; smoking status; alcohol use; Healthy Eating Index-2015; physical activity; serum sodium; serum potassium; serum bicarbonate; estimated glomerular filtration rate; use of diuretics; and comorbidity or history of hypertension, diabetes, coronary heart disease, stroke, chronic obstructive pulmonary disease, and cancer. Solid blue lines are multivariable-adjusted HR estimations, and the dashed red lines are the corresponding 95% CIs. The reference point was set at the median (103.2 mmol/L). HR: hazard ratio.



Subgroup Analyses of the Association of Serum Chloride With Mortality

Subgroup analyses were performed to depict the potential different associations of serum chloride quartiles and mortality in subpopulations. For the association with the primary outcome, all-cause mortality, no significant interactions were found between serum chloride and all these stratification variables (all *P* for interaction>.05), implying similar associations of serum chloride with all-cause mortality were detected in these subgroups (Table 3). Similarly, consistent associations of serum chloride with CVD mortality were detected across different subpopulations, and no significant interactions were detected (all *P* for interaction>.05; Multimedia Appendix 3). Regarding cancer mortality, more prominent associations of low chloride with mortality risk were detected in smokers and participants with hypertension, as shown by the interaction effect of serum chloride with smoking status (*P* for interaction=.04) and comorbid hypertension (*P* for interaction=.047; Multimedia Appendix 4). With regard to respiratory mortality, more prominent associations of low chloride with mortality risk were detected in adults with less physical activity, as indicated by its potential interaction effect with serum chloride (*P* for interaction=.02; Multimedia Appendix 5).

Table 3. Subgroup analyses of the associations of serum chloride quartiles with all-cause mortality for adults from the US National Health and Nutrition Examination Survey 1999-2018 (N=51,060)[a].

	Sample size, n	Q1[b] (≤101.2)	Q2[c] (101.3-103.2)	Q3[d] (103.3-105.0)	Q4[e] (≥105.1)	*P* value for trend	*P* value for interaction
		HR[f] (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)		
Sex							.66
Female	25,632	1.00 (reference)	0.77 (0.63-0.95)	0.80 (0.64-1.00)	0.83 (0.65-1.05)	.09	
Male	25,428	1.00 (reference)	0.76 (0.63-0.91)	0.66 (0.55-0.79)	0.74 (0.60-0.91)	.002	
Age (years)							.59
≤60	35,508	1.00 (reference)	0.81 (0.63-1.06)	0.78 (0.61-1.00)	0.77 (0.56-1.07)	.09	
>60	15,552	1.00 (reference)	0.71 (0.61-0.82)	0.69 (0.59-0.82)	0.78 (0.64-0.95)	.004	
Race							.25
White	22,199	1.00 (reference)	0.77 (0.66-0.89)	0.69 (0.58-0.81)	0.76 (0.63-0.92)	.002	
People of racial and ethnic minorities	28,861	1.00 (reference)	0.81 (0.64-1.03)	0.87 (0.70-1.08)	0.78 (0.59-1.04)	.09	
Education							.87
High school or below	23,861	1.00 (reference)	0.75 (0.63-0.89)	0.67 (0.55-0.80)	0.71 (0.59-0.86)	<.001	
College or above	23,630	1.00 (reference)	0.82 (0.67-1.00)	0.81 (0.66-1.00)	0.86 (0.66-1.13)	.22	
Marital status							.29
Married	24,924	1.00 (reference)	0.81 (0.69-0.96)	0.75 (0.64-0.88)	0.73 (0.60-0.89)	<.001	
Other	24,116	1.00 (reference)	0.73 (0.59-0.91)	0.69 (0.56-0.85)	0.85 (0.65-1.10)	.13	
Family income-to-poverty ratio							.19
≤1.3	14,912	1.00 (reference)	0.74 (0.58-0.94)	0.73 (0.59-0.90)	0.83 (0.63-1.09)	.18	
>1.3	31,676	1.00 (reference)	0.80 (0.68-0.94)	0.72 (0.60-0.85)	0.73 (0.60-0.89)	<.001	
BMI							.65
≤28	26,350	1.00 (reference)	0.75 (0.65-0.87)	0.68 (0.56-0.81)	0.75 (0.61-0.91)	.001	
>28	23,827	1.00 (reference)	0.80 (0.64-0.98)	0.77 (0.63-0.94)	0.80 (0.62-1.03)	.07	
Smoker							.11
Nonsmoker	26,271	1.00 (reference)	0.79 (0.65-0.96)	0.84 (0.68-1.04)	1.05 (0.81-1.35)	.90	
Smoker	22,013	1.00 (reference)	0.75 (0.63-0.89)	0.65 (0.53-0.79)	0.65 (0.53-0.79)	<.001	
Current drinker							.97
No	14,138	1.00 (reference)	0.78 (0.63-0.96)	0.71 (0.57-0.88)	0.78 (0.60-1.02)	.047	
Yes	29,646	1.00 (reference)	0.78 (0.66-0.93)	0.74 (0.63-0.88)	0.78 (0.64-0.96)	.007	
Healthy Eating Index -2015							.14
≤50	24,430	1.00 (reference)	0.69 (0.57-0.82)	0.66 (0.54-0.80)	0.75 (0.60-0.94)	.006	
>50	23,604	1.00 (reference)	0.88 (0.74-1.04)	0.79 (0.65-0.95)	0.80 (0.66-0.98)	.02	
Physical activity							.95
Inactive	13,114	1.00 (reference)	0.80 (0.67-0.95)	0.77 (0.63-0.93)	0.87 (0.69-1.09)	.10	
Active	23,689	1.00 (reference)	0.76 (0.62-0.92)	0.67 (0.56-0.81)	0.70 (0.55-0.90)	.003	
Diuretics use							.14
No	44,379	1.00 (reference)	0.81 (0.69-0.95)	0.72 (0.61-0.85)	0.75 (0.63-0.89)	<.001	
Yes	6632	1.00 (reference)	0.67 (0.52-0.85)	0.73 (0.58-0.93)	0.90 (0.66-1.24)	.09	
Hypertension							.46
No	30,501	1.00 (reference)	0.88 (0.69-1.12)	0.85 (0.66-1.10)	0.96 (0.71-1.29)	.74	
Yes	20,545	1.00 (reference)	0.72 (0.62-0.83)	0.65 (0.55-0.76)	0.68 (0.57-0.81)	<.001	

	Sample size, n	Q1[b] (≤101.2)	Q2[c] (101.3-103.2)	Q3[d] (103.3-105.0)	Q4[e] (≥105.1)	*P* value for trend	*P* value for interaction
		HR[f] (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)		
Diabetes							.32
No	42,533	1.00 (reference)	0.81 (0.69-0.94)	0.71 (0.61-0.83)	0.76 (0.63-0.91)	.002	
Yes	8523	1.00 (reference)	0.66 (0.52-0.84)	0.73 (0.57-0.93)	0.80 (0.59-1.09)	.04	
Coronary heart disease							.48
No	45,305	1.00 (reference)	0.76 (0.66-0.88)	0.71 (0.62-0.82)	0.78 (0.66-0.92)	.001	
Yes	2040	1.00 (reference)	0.89 (0.67-1.18)	0.73 (0.50-1.08)	0.72 (0.47-1.10)	.11	
Stroke							.76
No	45,663	1.00 (reference)	0.78 (0.67-0.90)	0.71 (0.62-0.82)	0.76 (0.64-0.90)	<.001	
Yes	1837	1.00 (reference)	0.65 (0.40-1.07)	0.87 (0.51-1.46)	0.91 (0.52-1.60)	.66	
Chronic obstructive pulmonary disease							.69
No	45,581	1.00 (reference)	0.79 (0.68-0.92)	0.73 (0.62-0.84)	0.77 (0.65-0.91)	.001	
Yes	2009	1.00 (reference)	0.69 (0.45-1.07)	0.72 (0.47-1.10)	0.84 (0.48-1.46)	.42	
Cancer							.54
No	43,107	1.00 (reference)	0.83 (0.72-0.96)	0.78 (0.67-0.91)	0.85 (0.71-1.03)	.053	
Yes	4404	1.00 (reference)	0.57 (0.42-0.76)	0.52 (0.39-0.68)	0.51 (0.36-0.73)	<.001	
Chronic kidney disease							.19
No	41,298	1.00 (reference)	0.84 (0.71-1.00)	0.77 (0.65-0.92)	0.79 (0.64-0.96)	.01	
Yes	9215	1.00 (reference)	0.68 (0.57-0.82)	0.63 (0.52-0.76)	0.72 (0.56-0.94)	.007	

[a] Data were calculated using svycoxph to fit a multivariate Cox proportional hazards model to data from a complex survey design, with adjustment of dichotomic sex; age; race; education; marital status; family income-to-poverty ratio; BMI; smoking; alcohol use; Healthy Eating Index-2015; physical activity; serum sodium; serum potassium; serum bicarbonate; use of diuretics; and comorbidity or history of hypertension, diabetes, coronary heart disease, stroke, chronic obstructive pulmonary disease, cancer, and chronic kidney disease, except for the specific stratification variable. The test for trend was based on the variable containing the median value for each quartile. The interaction effect was evaluated using the likelihood ratio test.

[b] Q1: quartile 1.

[c] Q2: quartile 2.

[d] Q3: quartile 3.

[e] Q4: quartile 4.

[f] HR: hazard ratio.

Sensitivity Analyses of the Association of Serum Chloride With Mortality

All sensitivity analyses showed a robust association of serum chloride with all-cause mortality and cause-specific mortality. Briefly, after excluding participants who died in the first 2 years (992/51,060, 1.94%), the association of categorical serum chloride with all-cause and cause-specific mortality remained stable (Multimedia Appendix 6). The association of categorical serum chloride with mortality was slightly attenuated in the fully adjusted model, especially with cancer mortality, when excluding participants with potential serum chloride outliers (936/51,060, 1.83%; Multimedia Appendix 7). In addition, excluding participants with possible hypoalbuminemia (714/51,060, 1.4%) did not cause substantial shifts in the results (Multimedia Appendix 8). Additional adjustment of history of congestive heart failure or separate adjustment of potassium-sparing diuretics and nonpotassium-sparing diuretics did not cause substantial shifts in the results (Multimedia Appendix 9). Furthermore, there were some missing data on covariates (Multimedia Appendix 10), and multiple imputation of them strengthened the associations of serum chloride with all-cause and cause-specific mortality, probably owing to the increased power (Multimedia Appendix 11).

Finally, we calculated E-values to assess the sensitivity to unmeasured confounding. Briefly, the all-cause mortality risk of low serum chloride (Q1) could be influenced only when the unmeasured covariates had risk ratios >1.92 for both serum chloride and all-cause mortality, and similarly, to account for the association we observed between serum chloride and cause-specific mortality and to address any unmeasured confounding, the strength of the relationship would need to be >2.34, 2.35, and 2.78 (Multimedia Appendix 12).

Discussion

Principal Findings

In this study, we used a prospective cohort with a relatively large sample size to explore the link between serum chloride level and all-cause and cause-specific mortality in American adults. After adjusting for demographic factors, relevant lifestyle, and comorbid factors, low serum chloride levels were significantly associated with higher risks of all-cause mortality, CVD mortality, cancer mortality, and respiratory disease mortality within a certain range. Further analysis revealed an L-shaped relationship between serum chloride levels and all-cause and specific-cause mortality in the general population. The associations of serum chloride and mortality risk were consistent across different subgroup populations and were rather robust in different settings. These findings suggested that serum chloride may serve as a promising cost-effective health indicator in the general adult population.

The widely acknowledged normal range of serum chloride for an adult is 96 to 106 mmol/L [2], and in our study population, the serum chloride was 103.2 (101.2-105.0) mmol/L, which subtly fluctuated according to sex, age, or race as well as lifestyle factors and clinical situations. We found that participants with lower serum chloride tended to be older and have a lower BMI. Considering that low chloride levels may be a symptom of malnutrition [34], lower BMI and older age might reflect a nutritional deficiency in the population. In addition, lower chloride levels are also observed in people with comorbidities such as hypertension, kidney disease, and diabetes mellitus.

The association of remarkable abnormality of serum chloride, hypochloremia, and hyperchloremia with mortality risk has been investigated in various hospital settings. Thus, serum chloride level (<100 mmol/L) in patients with hypertension was independently associated with a 20% higher mortality rate, as shown by a cohort of 12,968 patients with hypertension [35]. Similarly, serum chloride levels were found to be inversely associated with all-cause mortality in patients with CKD [14] and with intensive care unit mortality in patients with cirrhosis who were critically ill [36]. However, in the community-based population, the association between serum chloride distribution and mortality has been scarcely discussed and seems contradictory. Consistent with our study, a cohort study in the general population in Belgium that was followed up for 10 years showed that serum chloride levels <100 mmol/L were associated with increased all-cause, CVD, and non-CVD mortality [37]. Although we drew a conclusion similar to their study, we studied a different target population with a larger sample size and included different covariates. Another cohort study including 16,483 participants in NHANES III (1988-1994) followed up to 2011 found that low serum sodium level was associated with increased cardiovascular mortality, whereas lower serum chloride level was not, which was inconsistent with this study [38]. In our study, we included 51,060 participants (1999-2018) and followed them up to 2019. We included more covariates and conducted a more comprehensive analysis to confirm the results. Thus, the differences in the sample size, follow-up time, and statistical method might explain this discrepancy.

In our study, we found that low chloride level was not only associated with all-cause mortality but also with cause-specific mortality, including CVD mortality, cancer mortality, and respiratory mortality. Consistent with our findings, various studies have shown that lower chloride levels are associated with CVD mortality and play a prognostic role in CVD [11,37,39,40]. Several hypotheses have been proposed regarding the mechanism behind this. First, lower chloride levels or hypochloremia could increase the activity of a family of with no lysine kinases, which leads to diuretic resistance [41]. Second, as chloride, not sodium, is the principal regulator of renin release from the juxtaglomerular apparatus, hypochloremia can decrease chloride delivery to the macula densa and hence increase the renin release, which has long-term deleterious effects in patients with heart issues, such as venous congestion and pathologic myocardial remodeling [42]. Third, hypochloremia may impair the activity of the cardiac chloride channel, which may result in impaired cardiac contractility and arrhythmia and lead to sudden death [43,44]. Furthermore, electrolyte imbalance including hypochloremia can result in the dysregulation of myocyte intracellular pH, which also contributes to arrhythmia [45].

Studies on the association between serum chloride and cancer mortality are rare. A retrospective study in China demonstrated that low chloride levels were associated with increased mortality in patients with colorectal cancer after surgery [46]. Another prospective study in a hospital in Saudi Arabia reported that lower chloride level was a unique prognostic marker for high-risk polyps, which was associated with colorectal cancer [47]. Several hypotheses may elucidate the underlying mechanism. A cancer diagnosis may induce chronic stress in patients [48], which causes relative disruption of adrenal hormones and results in a decrease of chloride level [49]. Tumor-associated abnormal expression of some ion channel proteins can also lead to electrolyte disorders including hypochloremia and thus lead to adverse outcomes [50]. In addition, our subgroup analysis showed that smokers and adults with hypertension might be more susceptible to cancer mortality related to low serum chloride levels; however, further studies are needed to clarify the potential mechanisms.

Literature on the association between serum chloride and respiratory mortality is also limited. Consistent with our study, a retrospective cohort study in Glasgow showed that mortality in patients with COPD increased linearly as serum chloride levels <105 mmol/L [51]. Lower chloride level is a metabolic indicator of chronic respiratory acidosis [52]. Patients with COPD usually have an increase in partial pressure of carbon dioxide and bicarbonate; thus, chloride decrease may be a renal response to the increasing partial pressure of carbon dioxide to assure the electroneutrality. Moreover, our subgroup analysis indicated that participants with less physical activity may bear a higher respiratory mortality risk with low chloride compared with the counterparts with active physical activity, and further studies are needed to fully elucidate the mechanism.

Compared with Q1 and Q4 groups, participants in the Q2 and Q3 groups of serum chloride tend to have relative lower hazards, probably because their serum chloride levels (101.3-105.0 mmol/L) represents an ideal homeostasis which may, in some degree, avoid the risk of adverse events related to hypochloremia or hyperchloremia. Surprisingly, our study demonstrated that all-cause and specific-cause mortality did not increase when the serum chloride level was >105.1 mmol/L, which is in accordance with a recent community-based study conducted in

Israel [53] and studies involving people with high blood pressure [35] or those with chronic heart problems [9]. However, some inconsistent results are also noteworthy, in which a retrospective cohort in South Korea showed that the 90-day mortality was increased in patients after noncardiac surgery when serum chloride level was >110 mmol/L [54], and early hyperchloremia was associated with higher odds of death or disability at 90-day, in patients with acute ischemic stroke [55], and high serum chloride level at admission was associated with higher in-hospital mortality in patients in medical intensive care units [56]. The significant difference in the clinical characteristics, demographics, and laboratory variables in the different population settings may partially explain this discrepancy. Moreover, previous literature suggests that hyperchloremia is often hospital acquired, whereas hypochloremia is community acquired [57], in which people with unfavorably high serum chloride levels or hyperchloremia are more likely to be patients in the hospital but not individuals from the community, as in our study.

Fluctuation of serum chloride usually occurs with concomitant changes in serum sodium, potassium, and bicarbonate; therefore, we adjusted them in the final models and confirmed the chloride-specific risk independent of hyponatremia or acid-base disturbances. Serum chloride is largely influenced by renal function and the use of diuretics; therefore, we adjusted eGFR and diuretic use in the final model and performed stratified analysis according to the renal function and diuretic use, validating the consistent association of low serum chloride with mortality risk. In our study, the cutoff point for low serum chloride was set as the first quartile of chloride distribution in the population (101.2 mmol/L), whereas hypochloremia was commonly defined as ≤96 mmol/L [10,58], and some researchers selected 100 mmol/L as the cutoff point [13,35]; therefore, the differences in definition criteria should be taken into account when explaining the effect size of low serum chloride on mortality risk.

This study had several strengths, including the relatively large sample size (>50,000) and the nationally representative sample design, which facilitates the generalization of our findings. Moreover, we took into account a multitude of potential confounding factors, including sociodemographic factors, lifestyle factors, laboratory tests, drug use, and comorbidities with extensive statistical methods, which guaranteed the robustness of our findings.

Limitations

However, some limitations of this study must be acknowledged. First, causality cannot be determined owing to the observational study design, which requires randomized controlled studies for confirmation. Second, the serum chloride concentration was investigated only once at baseline, which may be influenced by temporary factors; therefore, repeated measurements are advisable in future studies. Third, the study participants were US adult civilians; thus, the associations and size effect need to be confirmed when trying to generalize the results to other populations. Finally, the possibility of residual and unknown confounding cannot be entirely excluded.

Conclusions

In conclusion, we found that decreased serum chloride concentrations were independently associated with increased all-cause mortality, CVD mortality, cancer mortality, and respiratory disease mortality in the American adult population, and these findings emphasized the importance of monitoring serum chloride in health evaluation. However, further studies are still warranted to verify the predictive role of serum chloride and to explore the potential pathophysiological mechanisms underlying the differences in mortality.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (82001195 to JW; 82001194 to XH); Natural Science Foundation of Hunan Province (2021JJ40990 to XH; 2020JJ5926 to JW); and Health Commission of Hunan province (202104110613 to WX).

Data Availability

The data sets were derived from National Health and Nutrition Examination Survey 1999-2018 [59], and only publicly available data were used in this study. The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

XH designed the study, collected and analyzed the data, and drafted the manuscript. WX, ZS, and CZ contributed critically to the data analysis. MZ and QG revised the manuscript. JW designed the study and drafted the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The distribution of leading causes of death in the enrolled participants with survey-weighted percentages from the US National Health and Nutrition Examination Survey 1999-2018.
[PNG File , 105 KB - publichealth_v9i1e49291_app1.png]

Multimedia Appendix 2
Survey-weighted multivariate analyses of the associations of continuous serum chloride with all-cause and cause-specific mortality for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 42 KB - publichealth_v9i1e49291_app2.doc]

Multimedia Appendix 3
Subgroup analyses of the associations of serum chloride quartiles with cardiovascular disease mortality for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 110 KB - publichealth_v9i1e49291_app3.doc]

Multimedia Appendix 4
Subgroup analyses of the associations of serum chloride quartiles with cancer mortality for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 110 KB - publichealth_v9i1e49291_app4.doc]

Multimedia Appendix 5
Subgroup analyses of the associations of serum chloride quartiles with respiratory disease mortality for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 101 KB - publichealth_v9i1e49291_app5.doc]

Multimedia Appendix 6
Survey-weighted multivariate analyses of the associations of categorical serum chloride with all-cause and cause-specific mortality after excluding participants who died in the first 2 years for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 62 KB - publichealth_v9i1e49291_app6.doc]

Multimedia Appendix 7
Survey-weighted multivariate analyses of the associations of categorical serum chloride with all-cause and cause-specific mortality after excluding participants with potential serum chloride outliers for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 62 KB - publichealth_v9i1e49291_app7.doc]

Multimedia Appendix 8
Survey-weighted multivariate analyses of the associations of categorical serum chloride with all-cause and cause-specific mortality after excluding participants with possible hypoalbuminemia for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 63 KB - publichealth_v9i1e49291_app8.doc]

Multimedia Appendix 9
Survey-weighted multivariate analyses of the associations of categorical serum chloride with all-cause and cause-specific mortality after additionally adjusting the history of congestive heart failure and separately adjusted use of potassium-sparing diuretics and potassium-wasting diuretics for adults from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 55 KB - publichealth_v9i1e49291_app9.doc]

Multimedia Appendix 10
Summary of missing data in covariates in analyses from the US National Health and Nutrition Examination Survey 1999-2018.
[PNG File , 98 KB - publichealth_v9i1e49291_app10.png]

Multimedia Appendix 11
Survey-weighted multivariate analyses of the associations of categorical serum chloride with all-cause and cause-specific mortality for adults after imputation all missing covariates from the US National Health and Nutrition Examination Survey 1999-2018.
[DOC File , 62 KB - publichealth_v9i1e49291_app11.doc]

Multimedia Appendix 12

E-value analysis for categorical serum chloride with all-cause and cause-specific mortality in adult participants from the US National Health and Nutrition Examination Survey 1999-2018.

[DOC File , 43 KB - publichealth_v9i1e49291_app12.doc]

References

1. Berend K, van Hulsteijn LH, Gans RO. Chloride: the queen of electrolytes? Eur J Intern Med 2012 Apr;23(3):203-211. [doi: 10.1016/j.ejim.2011.11.013] [Medline: 22385875]
2. Morrison G. Serum chloride. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. Third Edition. Boston, MA: Butterworth Publishers; 1990.
3. Pfortmueller CA, Uehlinger D, von Haehling S, Schefold JC. Serum chloride levels in critical illness-the hidden story. Intensive Care Med Exp 2018 Apr 13;6(1):10 [FREE Full text] [doi: 10.1186/s40635-018-0174-5] [Medline: 29654387]
4. Goto K, Kitazono T. Chloride ions, vascular function and hypertension. Biomedicines 2022 Sep 18;10(9):2316 [FREE Full text] [doi: 10.3390/biomedicines10092316] [Medline: 36140417]
5. Powers F. The role of chloride in acid-base balance. J Intraven Nurs 1999;22(5):286-291. [Medline: 10776193]
6. Campbell S. Dietary reference intakes: water, potassium, sodium, chloride, and sulfate. Nutr MD 2004 Jun;30(6):1-4 [FREE Full text]
7. Yunos NM, Bellomo R, Story D, Kellum J. Bench-to-bedside review: chloride in critical illness. Crit Care 2010;14(4):226 [FREE Full text] [doi: 10.1186/cc9052] [Medline: 20663180]
8. Kazory A, Ronco C. Emergence of chloride as an overlooked cardiorenal connector in heart failure. Blood Purif 2020;49(1-2):219-221 [FREE Full text] [doi: 10.1159/000503774] [Medline: 31851979]
9. Grodin JL, Verbrugge FH, Ellis SG, Mullens W, Testani JM, Tang WH. Importance of abnormal chloride homeostasis in stable chronic heart failure. Circ Heart Fail 2016 Jan;9(1):e002453 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.115.002453] [Medline: 26721916]
10. Testani JM, Hanberg JS, Arroyo JP, Brisco MA, Ter Maaten JM, Wilson FP, et al. Hypochloraemia is strongly and independently associated with mortality in patients with chronic heart failure. Eur J Heart Fail 2016 Jun;18(6):660-668 [FREE Full text] [doi: 10.1002/ejhf.477] [Medline: 26763893]
11. Cuthbert JJ, Pellicori P, Rigby A, Pan D, Kazmi S, Shah P, et al. Low serum chloride in patients with chronic heart failure: clinical associations and prognostic significance. Eur J Heart Fail 2018 Oct;20(10):1426-1435 [FREE Full text] [doi: 10.1002/ejhf.1247] [Medline: 29943886]
12. Zhou L, Wang X, Zhan X, Feng X, Wang N, Peng F, et al. Serum chloride and mortality in patients on continuous ambulatory peritoneal dialysis: a multi-center retrospective study. EClinicalMedicine 2021 Sep 16;41:101133 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101133] [Medline: 34585124]
13. Naal T, Abuhalimeh B, Khirfan G, Dweik RA, Tang WH, Tonelli AR. Serum chloride levels track with survival in patients with pulmonary arterial hypertension. Chest 2018 Sep;154(3):541-549 [FREE Full text] [doi: 10.1016/j.chest.2018.04.022] [Medline: 29698719]
14. Mandai S, Kanda E, Iimori S, Naito S, Noda Y, Kikuchi H, et al. Association of serum chloride level with mortality and cardiovascular events in chronic kidney disease: the CKD-ROUTE study. Clin Exp Nephrol 2017 Mar;21(1):104-111. [doi: 10.1007/s10157-016-1261-0] [Medline: 27039905]
15. Semmler G, Scheiner B, Balcar L, Paternostro R, Simbrunner B, Pinter M, et al. Disturbances in sodium and chloride homeostasis predict outcome in stable and critically ill patients with cirrhosis. Aliment Pharmacol Ther 2023 Jul;58(1):71-79. [doi: 10.1111/apt.17507] [Medline: 37016513]
16. Zhang Z, Xu X, Fan H, Li D, Deng H. Higher serum chloride concentrations are associated with acute kidney injury in unselected critically ill patients. BMC Nephrol 2013 Oct 28;14:235 [FREE Full text] [doi: 10.1186/1471-2369-14-235] [Medline: 24164963]
17. Grodin JL, Testani JM, Pandey A, Sambandam K, Drazner MH, Fang JC, et al. Perturbations in serum chloride homeostasis in heart failure with preserved ejection fraction: insights from TOPCAT. Eur J Heart Fail 2018 Oct;20(10):1436-1443 [FREE Full text] [doi: 10.1002/ejhf.1229] [Medline: 29893446]
18. About the national health and nutrition examination survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/about_nhanes.htm [accessed 2023-10-30]
19. NCHS Ethics Review Board (ERB) approval. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm [accessed 2023-10-30]
20. National health and nutrition examination survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2023-10-30]
21. Hou X, Hu J, Liu Z, Wang E, Guo Q, Zhang Z, et al. L-shaped association of serum calcium with all-cause and CVD mortality in the US adults: a population-based prospective cohort study. Front Nutr 2023 Jan 5;9:1097488 [FREE Full text] [doi: 10.3389/fnut.2022.1097488] [Medline: 36687714]

22. Chen CJ, Brown WM, Moomaw CJ, Langefeld CD, Osborne J, Worrall BB, et al. Alcohol use and risk of intracerebral hemorrhage. Neurology 2017 May 23;88(21):2043-2051 [FREE Full text] [doi: 10.1212/WNL.0000000000003952] [Medline: 28446657]
23. Reedy J, Lerman JL, Krebs-Smith SM, Kirkpatrick SI, Pannucci TE, Wilson MM, et al. Evaluation of the healthy eating index-2015. J Acad Nutr Diet 2018 Sep;118(9):1622-1633 [FREE Full text] [doi: 10.1016/j.jand.2018.05.019] [Medline: 30146073]
24. Kuo CP, Tsai MT, Lee KH, Lin YP, Huang SS, Huang CC, et al. Dose-response effects of physical activity on all-cause mortality and major cardiorenal outcomes in chronic kidney disease. Eur J Prev Cardiol 2022 Mar 25;29(3):452-461. [doi: 10.1093/eurjpc/zwaa162] [Medline: 33704426]
25. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF3, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009 May 05;150(9):604-612 [FREE Full text] [doi: 10.7326/0003-4819-150-9-200905050-00006] [Medline: 19414839]
26. 2019 public-use linked mortality files. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/data-linkage/mortality-public.htm [accessed 2023-10-30]
27. Johnson CL, Paulose-Ram R, Ogden CL, Carroll MD, Kruszon-Moran D, Dohrmann SM, et al. National health and nutrition examination survey: analytic guidelines, 1999-2010. Vital Health Stat 2 2013 Sep(161):1-24 [FREE Full text] [Medline: 25090154]
28. Akinbami LJ, Chen TC, Davy O, Ogden CL, Fink S, Clark J, et al. National health and nutrition examination survey, 2017-March 2020 prepandemic file: sample design, estimation, and analytic guidelines. Vital Health Stat 1 2022 May(190):1-36. [Medline: 35593699]
29. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 2018 Dec 05;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457 [FREE Full text] [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
31. Kidney Disease: Improving Global Outcomes (KDIGO) Glomerular Diseases Work Group. KDIGO 2021 Clinical practice guideline for the management of glomerular diseases. Kidney Int 2021 Oct;100(4S):S1-276 [FREE Full text] [doi: 10.1016/j.kint.2021.05.021] [Medline: 34556256]
32. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the e-value. Ann Intern Med 2017 Aug 15;167(4):268-274. [doi: 10.7326/M16-2607] [Medline: 28693043]
33. The R project for statistical computing. The Comprehensive R Archive Network. URL: https://www.r-project.org/ [accessed 2023-11-02]
34. Jacocks MA. Fluids. Electrolytes. Nutrition. In: General Surgery. New York, NY: Springer; 1996.
35. McCallum L, Jeemon P, Hastie CE, Patel RK, Williamson C, Redzuan AM, et al. Serum chloride is an independent predictor of mortality in hypertensive patients. Hypertension 2013 Nov;62(5):836-843. [doi: 10.1161/HYPERTENSIONAHA.113.01793] [Medline: 23980073]
36. Ji Y, Li L. Lower serum chloride concentrations are associated with increased risk of mortality in critically ill cirrhotic patients: an analysis of the MIMIC-III database. BMC Gastroenterol 2021 May 01;21(1):200 [FREE Full text] [doi: 10.1186/s12876-021-01797-3] [Medline: 33933032]
37. De Bacquer D, De Backer G, De Buyzere M, Kornitzer M. Is low serum chloride level a risk factor for cardiovascular mortality? J Cardiovasc Risk 1998 Jun;5(3):177-184. [Medline: 10201555]
38. He X, Liu C, Chen Y, He J, Dong Y. Risk of cardiovascular mortality associated with serum sodium and chloride in the general population. Can J Cardiol 2018 Aug;34(8):999-1003. [doi: 10.1016/j.cjca.2018.03.013] [Medline: 29935992]
39. Hanberg JS, Rao V, Ter Maaten JM, Laur O, Brisco MA, Perry Wilson F, et al. Hypochloremia and diuretic resistance in heart failure: mechanistic insights. Circ Heart Fail 2016 Aug;9(8):10.1161/CIRCHEARTFAILURE.116.003180 e003180 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.116.003180] [Medline: 27507113]
40. Grodin JL, Simon J, Hachamovitch R, Wu Y, Jackson G, Halkar M, et al. Prognostic role of serum chloride levels in acute decompensated heart failure. J Am Coll Cardiol 2015 Aug 11;66(6):659-666 [FREE Full text] [doi: 10.1016/j.jacc.2015.06.007] [Medline: 26248993]
41. Gupta R, Testani J, Collins S. Diuretic resistance in heart failure. Curr Heart Fail Rep 2019 Apr;16(2):57-66 [FREE Full text] [doi: 10.1007/s11897-019-0424-1] [Medline: 30762178]
42. Arora N. Serum chloride and heart failure. Kidney Med 2023 Apr;5(4):100614 [FREE Full text] [doi: 10.1016/j.xkme.2023.100614] [Medline: 36911181]
43. Duan DY, Liu LL, Bozeat N, Huang ZM, Xiang SY, Wang GL, et al. Functional role of anion channels in cardiac diseases. Acta Pharmacol Sin 2005 Mar;26(3):265-278. [doi: 10.1111/j.1745-7254.2005.00061.x] [Medline: 15715921]
44. Orchard CH, Cingolani HE. Acidosis and arrhythmias in cardiac muscle. Cardiovasc Res 1994 Sep;28(9):1312-1319. [doi: 10.1093/cvr/28.9.1312] [Medline: 7954638]
45. Orchard CH, Kentish JC. Effects of changes of pH on the contractile function of cardiac muscle. Am J Physiol 1990 Jun;258(6 Pt 1):C967-C981. [doi: 10.1152/ajpcell.1990.258.6.C967] [Medline: 2193525]

46. Li Q, Dai W, Jia H, Li Y, Xu Y, Li X, et al. Prognostic impact of hypochloremia in patients with stage I to III colorectal cancer after radical resection. Dis Colon Rectum 2018 Nov;61(11):1273-1280. [doi: 10.1097/DCR.0000000000001196] [Medline: 30239393]
47. Algassim A, Semaan T, Aldaher MA, Alluhaydan A, Jaddoh A, Al-Zubide S, et al. Serum chloride and admission status are potential prognostic markers of high-risk polyps: a prospective characterization of colorectal polyps in a tertiary hospital in Saudi Arabia. Cureus 2022 Jun;14(6):e26116 [FREE Full text] [doi: 10.7759/cureus.26116] [Medline: 35875269]
48. DiMatteo MR, Lepper HS, Croghan TW. Depression is a risk factor for noncompliance with medical treatment: meta-analysis of the effects of anxiety and depression on patient adherence. Arch Intern Med 2000 Jul 24;160(14):2101-2107. [doi: 10.1001/archinte.160.14.2101] [Medline: 10904452]
49. Harper JR, Ginn WMJ, Taylor WJ. Bilateral adrenal hemorrhage--a complication of anticoagulant therapy. Case report and review of the literature. Am J Med 1962 Jun;32:984-988. [doi: 10.1016/0002-9343(62)90045-1] [Medline: 13904779]
50. Ousingsawat J, Spitzner M, Puntheeranurak S, Terracciano L, Tornillo L, Bubendorf L, et al. Expression of voltage-gated potassium channels in human and mouse colonic carcinoma. Clin Cancer Res 2007 Mar 01;13(3):824-831. [doi: 10.1158/1078-0432.CCR-06-1940] [Medline: 17289873]
51. Pellicori P, McConnachie A, Carlin C, Wales A, Cleland JG. Predicting mortality after hospitalisation for COPD using electronic health records. Pharmacol Res 2022 May;179:106199 [FREE Full text] [doi: 10.1016/j.phrs.2022.106199] [Medline: 35381341]
52. Krapf R, Beeler I, Hertner D, Hulter HN. Chronic respiratory alkalosis. The effect of sustained hyperventilation on renal regulation of acid-base equilibrium. N Engl J Med 1991 May 16;324(20):1394-1401. [doi: 10.1056/NEJM199105163242003] [Medline: 1902283]
53. Shafat T, Novack V, Barski L, Haviv YS. Community-based serum chloride abnormalities predict mortality risk. PLoS One 2023;18(2):e0279837 [FREE Full text] [doi: 10.1371/journal.pone.0279837] [Medline: 36809243]
54. Oh TK, Do SH, Jeon YT, Kim J, Na HS, Hwang JW. Association of preoperative serum chloride levels with mortality and morbidity after noncardiac surgery: a retrospective cohort study. Anesth Analg 2019 Dec;129(6):1494-1501. [doi: 10.1213/ANE.0000000000003958] [Medline: 31743168]
55. Qureshi AI, Huang W, Gomez FE, Malhotra K, Arora N, Chandrasekaran PN, et al. Early hyperchloremia and outcomes after acute ischemic stroke. J Stroke Cerebrovasc Dis 2022 Aug;31(8):106523. [doi: 10.1016/j.jstrokecerebrovasdis.2022.106523] [Medline: 35633589]
56. Yılmaz H, Güney BC. Evaluation of the effect of hyperchloremia on the prognosis and mortality of medical intensive care patients: a single-center study. Turk J Med Sci 2022 Oct;52(5):1689-1696 [FREE Full text] [doi: 10.55730/1300-0144.5510] [Medline: 36422489]
57. Shao M, Li G, Sarvottam K, Wang S, Thongprayoon C, Dong Y, et al. Dyschloremia is a risk factor for the development of acute kidney injury in critically ill patients. PLoS One 2016 Aug 4;11(8):e0160322 [FREE Full text] [doi: 10.1371/journal.pone.0160322] [Medline: 27490461]
58. Zandijk AJ, van Norel MR, Julius FE, Sepehrvand N, Pannu N, McAlister FA, et al. Chloride in heart failure: the neglected electrolyte. JACC Heart Fail 2021 Dec;9(12):904-915 [FREE Full text] [doi: 10.1016/j.jchf.2021.07.006] [Medline: 34857174]
59. NHANES questionnaires, datasets, and related documentation. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx [accessed 2023-11-02]

Abbreviations

CHD: coronary heart disease
CKD: chronic kidney disease
COPD: chronic obstructive pulmonary disease
CVD: cardiovascular disease
eGFR: estimated glomerular filtration rate
HR: hazard ratio
MET: metabolic equivalent task
NHANES: National Health and Nutrition Examination Survey
PIR: family income-to-poverty ratio
Q1: quartile 1
Q2: quartile 2
Q3: quartile 3
Q4: quartile 4
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 29.05.23; peer-reviewed by WC Tseng, K Malhotra, C Zhang, B Shu, Q Chen; comments to author 21.08.23; revised version received 31.08.23; accepted 17.10.23; published 13.11.23.

Please cite as:
Hou X, Xu W, Zhang C, Song Z, Zhu M, Guo Q, Wang J
L-Shaped Association of Serum Chloride Level With All-Cause and Cause-Specific Mortality in American Adults: Population-Based Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e49291
URL: https://publichealth.jmir.org/2023/1/e49291
doi:10.2196/49291
PMID:37955964

Original Paper

Lung Cancer Risk Prediction Nomogram in Nonsmoking Chinese Women: Retrospective Cross-sectional Cohort Study

Lanwei Guo[1], MD, PhD; Qingcheng Meng[2], MD; Liyang Zheng[1], BSc; Qiong Chen[1], MD; Yin Liu[1], MD; Huifang Xu[1], MD; Ruihua Kang[1], MD; Luyao Zhang[1], MD; Shuzheng Liu[1], MD; Xibin Sun[1], MPH; Shaokai Zhang[1], MD

[1]Department of Cancer Epidemiology and Prevention, Henan Engineering Research Center of Cancer Prevention and Control, Henan International Joint Laboratory of Cancer Prevention, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China

[2]Department of Radiology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China

Corresponding Author:
Shaokai Zhang, MD
Department of Cancer Epidemiology and Prevention, Henan Engineering Research Center of Cancer Prevention and Control, Henan International Joint Laboratory of Cancer Prevention
The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital
No.127 Dongming Road
Zhengzhou, 450008
China
Phone: 86 37165587361
Email: shaokaizhang@126.com

Abstract

Background: It is believed that smoking is not the cause of approximately 53% of lung cancers diagnosed in women globally.

Objective: The study aimed to develop and validate a simple and noninvasive model that could assess and stratify lung cancer risk in nonsmoking Chinese women.

Methods: Based on the population-based Cancer Screening Program in Urban China, this retrospective, cross-sectional cohort study was carried out with a vast population base and an immense number of participants. The training set and the validation set were both constructed using a random distribution of the data. Following the identification of associated risk factors by multivariable Cox regression analysis, a predictive nomogram was developed. Discrimination (area under the curve) and calibration were further performed to assess the validation of risk prediction nomogram in the training set, which was then validated in the validation set.

Results: In sum, 151,834 individuals signed up to take part in the survey. Both the training set (n=75,917) and the validation set (n=75,917) were comprised of randomly selected participants. Potential predictors for lung cancer included age, history of chronic respiratory disease, first-degree family history of lung cancer, menopause, and history of benign breast disease. We displayed 1-year, 3-year, and 5-year lung cancer risk–predicting nomograms using these 5 factors. In the training set, the 1-year, 3-year, and 5-year lung cancer risk areas under the curve were 0.762, 0.718, and 0.703, respectively. In the validation set, the model showed a moderate predictive discrimination.

Conclusions: We designed and validated a simple and noninvasive lung cancer risk model for nonsmoking women. This model can be applied to identify and triage people at high risk for developing lung cancers among nonsmoking women.

(JMIR Public Health Surveill 2023;9:e41640) doi:10.2196/41640

KEYWORDS

lung cancer; risk model; forecasting; validation; female; nonsmokers

Introduction

China has the most lung cancer death cases around the world in 2020. In 2020, according to estimates provided by the International Agency for Research on Cancer, there were approximately 1.80 million cases of deadly lung cancer globally. China accounted for 39.8% of these cases [1]. In China, the continuous rise in lung cancer deaths during the past 2 decades was attributed to the rising prevalence of lung cancer in women [2]. Additionally, 50% or more of lung cancers in women in Southeast Asia were diagnosed in nonsmokers [3-5]. Most of the Chinese lung cancer cases were reported to be clinically progressed in 2012-2014, with 64.6% of them being stage III-IV lung cancers [6]. The lung cancer survival rate in China, which was defined as 5 years as standard, grew up to 20% between 2003 and 2015 [7]. The prognosis of lung cancer is strongly associated with the stage in which it was detected; the 5-year survival rate ranges from 0% in cases detected in patients with stage IV cancer to >80% in cases detected in stage I and whose patients underwent surgery [8].

Started in 2002, the National Lung Screening Trial indicated that low-dose computed tomography screening may decrease lung cancer deaths by 20% [9]. However, this project only screened people (41% women) at high risk for lung cancer based on age and smoking history (aged 55-74 years, smoked no fewer than 30 pack-years, and had no more than 15 years of having quit smoking). Women in China have their own characteristics of lung cancer risk factor exposure and incidence patterns, the most critical of which is that although the smoking rate among women is much lower than that of high-income countries such as the United States (2.4% in China and 23.6% in the United States), the lung cancer frequencies are relatively similar (22.8/100,000 in China and 30.8/100,000 in the United States, based on the standardized lung cancer incidence rate of the world population) [10,11]. This finding shows that the existing worldwide guidelines for lung cancer screening focused on smoking as the primary predictor for high-risk individuals, which would be inappropriate for Chinese women, particularly for nonsmoking women. Therefore, determining a way that accurately forecasts the risk of lung cancer in nonsmoking women and directing them toward the more cost-effective low-dose computed tomography screening is a feasible method for achieving efficient early diagnosis and treatment of lung cancer.

Earlier research has developed numerous lung cancer risk predictive models related to specific population demographics [12-41]; however, few of the predictive methods focused on nonsmoking women in mainland China [42]. Consequently, the development of lung cancer risk predicting tools for nonsmoking Chinese women according to consistently established risk factors in earlier studies has become a top goal [43]. Nevertheless, this goal is demanding and difficult. In contrast to the findings of lung cancer caused by tobacco, there are no identified risk variables for the progression of lung cancer in nonsmoking women. Although other risk factors were suggested, their relative importance varies greatly between geographical locations [3,4,44,45]. It was observed that the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) models, which included nearly 2000 Asian nonsmokers and only 7 cases of lung cancer, could be inapplicable to Asian nonsmokers [46]. Among the nonsmokers who participated in the PLCO study (n=65,711), none of them had a 6-year risk that was greater than 0.0151.

On the basis of the Cancer Screening Program in Urban China (CanSPUC), we created such a model [47]. In this paper, we aimed to create and internally validate a lung cancer risk predicting model for nonsmoking Chinese women, with the focus on established risk factors for lung cancer routinely available in general cancer-screening settings.

Methods

Data Source and Subjects

This retrospective, cross-sectional cohort study was carried out inside the scope of CanSPUC, a continuing statewide cancer-screening program for China’s urban population. CanSPUC is designed to detect the 5 most common malignancies, including lung cancer, colorectal cancer, upper gastrointestinal cancer, liver cancer, and female breast cancer. The CanSPUC approach was detailed in previous studies [47,48]. All of the qualified subjects were questioned by highly skilled staff to gather information about their exposure to risk variables and to assess their cancer risk using a specific cancer risk score system. The household registration system was used in local communities to identify eligible permanent residents who were aged 40-74 years and asymptomatic for lung cancer with no history of cancer diagnosis. Individuals who were unable to give informed consent, had a medical disability and were unlikely to complete curative lung cancer surgery, had a history of lung cancer, had received treatment for or had evidence of any cancer within the past 5 years (with the exception of nonmelanoma skin cancer and most in situ carcinomas), or had symptoms suggestive of lung cancer (including unexplained weight loss of >7.5 kg within the past 12 months or unexplained hemoptysis) were not eligible to participate. In October 2013, CanSPUC was implemented in Henan province, which encompassed 8 cities with complete cancer registration data (Zhengzhou, Zhumadian, Anyang, Luoyang, Nanyang, Jiaozuo, Puyang, and Xinxiang). We examined the data collected over the first 6 years (from October 2013 to October 2019) in the Henan province for our research. Only nonsmoking women were included in this investigation.

Ethics Approval

The Ethics Committee of the Affiliated Cancer Hospital of Zhengzhou University and Henan Cancer Hospital evaluated and authorized the research (no.2021-KY-0028-001). Our sample was drawn from retrospective encounters documented in the electronic health record; these data were deidentified for both sets of analyses and did not require informed consent.

Outcome, Variables, and Measurements

All new cases of lung cancer were identified by matching with the cancer registry database in Henan province, China (by unique ID number), and histologically confirmed between October 1, 2013, and March 10, 2020. In Henan province, records of lung cancer are first submitted to local cancer registries by the hospitals and medical institutions and then submitted to the Henan Provincial Central Cancer Registry of China by the local cancer registries. The International Classification of Diseases, Tenth Revision was used to classify newly diagnosed lung malignancies by site. Lung cancers were identified by the International Classification of Diseases, Tenth Revision code of C33-C34. To find possible lung cancer risk variables, self-reported information was collected (Textbox 1).

Textbox 1. Self-reported information collected.

1. **Demographic factors, such as age, ethnicity, educational status, marital status, height, and body weight**
    - A low-educational level was defined as elementary school or less, a medium educational level as junior or senior high school, and a high-educational level as college or above
    - According to the "Guidelines for the Prevention and Control of Overweight and Obesity in Chinese Adults," BMI was dependent on the individual's height as well as weight and segmented to "<18.5 kg/m^2," "18.5-23.9 kg/m^2," "24.0-27.9 kg/m^2," and "≥28.0 kg/m^2" categories [49]
2. **Dietary habit**
    - Dietary intake of the following food in the past 2 years: vegetables (green-leafy plants and fungi except potatoes, sweet potatoes, and starches) <2.5 kg/week or ≥2.5 kg/week; roughage (all other grains except for white flour and rice) <0.5 kg/week or ≥0.5 kg/week; and fruit <1.25 kg/week or ≥1.25 kg/week. The weight of the food was measured prior to cooking
3. **Living environment, behavior, and habits**
    - Cooking oil fume exposure: exposure is considered as "none or a little" if chimneys, fume extractors, or smokeless pots were used during cooking; otherwise, it was considered as "a lot"
    - Passive smoking: regular living or employment in an enclosed area where people routinely smoke was regarded as "yes"; otherwise, it was regarded as "no"
    - Alcohol consumption: "current" referred to those who had consumed alcohol at least once weekly on average for more than 6 months; "former" referred to those who had ceased drinking; "never" referred to those who had never consumed alcohol
    - Physical activity: swimming; taijiquan, qigong, or walking; long range running; aerobics; sporting events (such as basketball, table tennis, badminton, etc); Yangko dancing or fast walking; and other physical activities (such as mountains climbing, rope jumping, and shuttlecock kicking). Subjects who engage in at least three sessions of practice weekly for a total of ≥90 minutes weekly were classified as engaging in "heavy physical activity"; otherwise, they were classified as engaging in "moderate or no physical activity"
4. **Psychology and emotions, such as a history of serious trauma and more than 6 months of mental depression**
    - Serious trauma was described as a major illness or death of a family member, family conflict and separation, significant loss of property, unexpected job loss, severe unexpected physical injury, violent danger, etc
5. **Comorbidities, such as chronic respiratory disease, tuberculosis, chronic bronchitis, emphysema, asthma bronchiectasis, hypertension, hyperlipidemia, and diabetes**
    - Every self-reported case of comorbidity required an evaluation from a professional medical facility
6. **Family history of lung cancer**
    - Whether first-degree relatives, second-degree relatives, or third-degree relatives had lung cancer or not
7. **Physiology and fertility**
    - Including age of menarche (<12 years or ≥12 years), menopause (yes or no), fertility status (yes or no), lactation status (yes or no), history of benign breast illness (yes or no), and a history of reproductive system surgery (yes or no)

Statistical Analysis

To contrast the profiles of those who have lung cancer and those without cancer, descriptive statistics, presented as percentages for categorical data, were used. Chi-square tests were used to examine the univariate correlation between baseline characteristics and lung cancer progression. For continuous variables, mean (SD) or median (IQR) were used.

In this investigation, the integrated model was applied to generate a nomogram to measure the 1-year, 3-year, and 5-year estimations of the lung cancer risk in the training set, according to the independently prognostic variables using the stepwise multivariable Cox regression (P_{entry}=.15 and P_{stay}=.10). The calibration curve was used to determine the nomogram's validity. By applying 50% and 84% quantiles, the risk predictions were grouped into the low-risk group, medium-risk group, and high-risk group, as suggested previously [50]. As per the risk prediction model, Kaplan-Meier curves were displayed for the low-risk group, medium-risk group, and high-risk group for lung cancer. The log-rank analyses were performed to compare the 3 curves. Receiver operating characteristic curves and the area under the curve (AUC) were used to quantify the prediction performance of 1-year, 3-year, and 5-year lung cancer risk estimations in the training set and validation set. By comparing observed and predicted probabilities, the bootstrap sampling method was used to evaluate the calibration of the current model.

All statistical analysis was carried out via R (version 4.0.3; R Foundation for Statistical Computing) and SAS (version 9.4; SAS Institute) software. The nomogram was drawn using the *rms* package. The receiver operating characteristic curves were drawn by using the *survivalROC* package. Using the *ggplot2* package, a calibration curve was created. All of the tests were done using 2-tailed hypotheses, and *P*<.05 was determined to be statistically significant.

Results

Characteristics of the Study Population

This research consisted of a total of 151,834 qualified participants with an average age of 55.34 (SD 8.65) years. The subjects were randomly separated into a training set of 75,917 and a validation set of 75,917 (Figure 1). By March 2020, 204 lung cancer cases occurred within 151,834 subjects, resulting in an incident density of 42.24 per 100,000 person-years. Lung cancer cases were more frequent in those who were older (*P*<.001), had a history of respiratory illness (*P*=.001), had a first-degree family history of lung cancer (*P*=.02), and had menopause (*P*<.001). Extra features are shown in Table 1 and Table S1 in Multimedia Appendix 1.

Figure 1. Flow chart of participants included in this analysis.

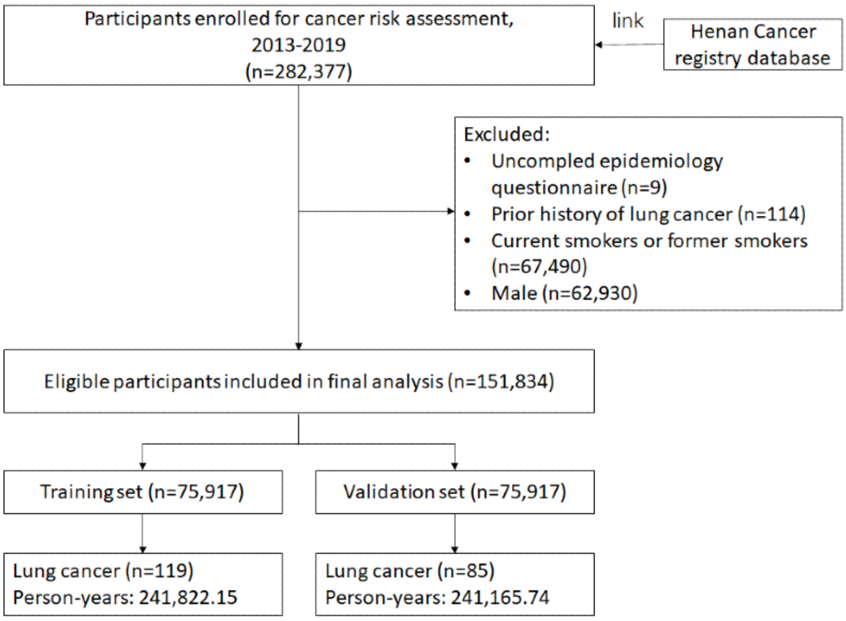


Table 1. Comparison of baseline characteristics between the non–lung cancer and lung cancer groups using chi-square test in the training set.

Variables	Total (N=75,917)[a]	Non–lung cancer[b]	Lung cancer[b]	χ^2 (*df*)	*P* value
All participants	75,917 (100)	75,798 (99.84)	119 (0.16)		
Person-years, median (IQR)	2.95 (1.73-4.83)	2.95 (1.73-4.83)	1.56 (0.83-2.38)		
Demographic characteristics					
Age (years), mean (SD)	55.37 (8.65)	55.36 (8.65)	60.37 (7.20)		
Age (years), n (%)				47.96 (6)	*<.001* [c]
40-44	9226 (12.15)	9221 (99.95)	5 (0.05)		
45-49	13,558 (17.86)	13,551 (99.95)	7 (0.05)		
50-54	14,389 (18.95)	14,376 (99.91)	13 (0.09)		
55-59	11,857 (15.62)	11,838 (99.84)	19 (0.16)		
60-64	12,927 (17.03)	12,889 (99.71)	38 (0.29)		
65-69	10,181 (13.41)	10,151 (99.71)	30 (0.29)		
70-74	3779 (4.98)	3772 (99.81)	7 (0.19)		
Race, n (%)				0.05 (1)	.83
Han nationality	74,431 (98.04)	74,314 (99.84)	117 (0.16)		
Others	1486 (1.96)	1484 (99.87)	2 (0.13)		
Education[d], n (%)				0.12 (2)	.94
Low	16,139 (21.26)	16,115 (99.85)	24 (0.15)		
Medium	49,922 (65.76)	49,842 (99.84)	80 (0.16)		
High	9856 (12.98)	9841 (99.85)	15 (0.15)		
Marriage, n (%)				1.89 (1)	.17
Unmarried, divorce, or widowed	3193 (4.21)	3191 (99.94)	2 (0.06)		
Married	72,724 (95.79)	72,607 (99.84)	117 (0.16)		
BMI (kg/m^2), n (%)				1.84 (3)	.61
<18.5	1133 (1.49)	1133 (100)	0 (0)		
18.5-24.0	35,445 (46.69)	35,388 (99.84)	57 (0.16)		
24.0-28.0	30,729 (40.48)	30,681 (99.84)	48 (0.16)		
≥28.0	8610 (11.34)	8596 (99.84)	14 (0.16)		
Dietary habit, n (%)					
Vegetables intake (kg/week)				0.01 (1)	.92
≥2.5	39,282 (51.74)	39,221 (99.84)	61 (0.16)		
<2.5	36,635 (48.26)	36,577 (99.84)	58 (0.16)		
Fruit intake (kg/week)				0.71 (1)	.40
≥1.25	43,683 (57.54)	43,610 (99.83)	73 (0.17)		
<1.25	32,234 (42.46)	32,188 (99.86)	46 (0.14)		
Roughage intake (kg/week)				0.64 (1)	.42
≥0.5	51,713 (68.12)	51,636 (99.85)	77 (0.15)		
<0.5	24,204 (31.88)	24,162 (99.83)	42 (0.17)		
Living environment, behavior, and habits, n (%)					
Cooking oil fume exposure				0.05 (1)	.82
None or a little	65,819 (86.70)	65,715 (99.84)	104 (0.16)		

Variables	Total (N=75,917)[a]	Non–lung cancer[b]	Lung cancer[b]	χ^2 (*df*)	*P* value
A lot	10,098 (13.3)	10,083 (99.85)	15 (0.15)		
Passive smoking				0.63 (1)	.43
No	49,045 (64.6)	48,964 (99.83)	81 (0.17)		
Yes	26,872 (35.4)	26,834 (99.86)	38 (0.14)		
Alcohol drinking				0.11 (2)	.95
Never	71,567 (94.27)	71,454 (99.84)	113 (0.16)		
Current	3647 (4.8)	3642 (99.86)	5 (0.14)		
Former	703 (0.93)	702 (99.86)	1 (0.14)		
Physical activity				3.19 (1)	.07
Moderate or no	40,014 (52.71)	39,961 (99.87)	53 (0.13)		
Heavy	35,903 (47.29)	35,837 (99.82)	66 (0.18)		
Psychology and emotions, n (%)					
History of a severe trauma				1.22 (1)	.27
No	65,199 (85.88)	65,101 (99.85)	98 (0.15)		
Yes	10,718 (14.12)	10,697 (99.8)	21 (0.2)		
Mental depression for over 6 months				0.00 (1)	.98
No	64,379 (84.8)	64,278 (99.84)	101 (0.16)		
Yes	11,538 (15.2)	11,520 (99.84)	18 (0.16)		
Comorbidities, n (%)					
History of chronic respiratory disease				11.53 (1)	*.001*
No	64,070 (84.39)	63,983 (99.86)	87 (0.14)		
Yes	11,847 (15.61)	11,815 (99.73)	32 (0.27)		
History of tuberculosis				1.24 (1)	.27
No	74,895 (98.65)	74,779 (99.85)	116 (0.15)		
Yes	1022 (1.35)	1019 (99.71)	3 (0.29)		
History of chronic bronchitis				3.44 (*1*)	.06
No	66,728 (87.9)	66,630 (99.85)	98 (0.15)		
Yes	9189 (12.1)	9168 (99.77)	21 (0.23)		
History of emphysema				3.21 (1)	.07
No	75,204 (99.06)	75,088 (99.85)	116 (0.15)		
Yes	713 (0.94)	710 (99.58)	3 (0.42)		
History of asthma bronchiectasis				1.27 (1)	.26
No	73,473 (96.78)	73,360 (99.85)	113 (0.15)		
Yes	2444 (3.22)	2438 (99.75)	6 (0.25)		
History of hypertension				1.66 (1)	.20
No	60,976 (80.32)	60,886 (99.85)	90 (0.15)		
Yes	14,941 (19.68)	14,912 (99.81)	29 (0.19)		
History of hyperlipidemia				1.67 (1)	.20
No	63,309 (83.39)	63,215 (99.85)	94 (0.15)		
Yes	12,608 (16.61)	12,583 (99.8)	25 (0.2)		
History of diabetes				0.00 (1)	.98
No	70,767 (93.22)	70,656 (99.84)	111 (0.16)		

Variables	Total (N=75,917)[a]	Non–lung cancer[b]	Lung cancer[b]	χ^2 (*df*)	*P* value
Yes	5150 (6.78)	5142 (99.84)	8 (0.16)		
First-degree family history of lung cancer, n (%)				5.15 (1)	*.02*
No	69,955 (92.15)	69,852 (99.85)	103 (0.15)		
Yes	5962 (7.85)	5946 (99.73)	16 (0.27)		
Physiology and fertility, n (%)					
Age of menarche (years)				0.34 (1)	.56
<12	1910 (2.52)	1908 (99.9)	2 (0.1)		
≥12	74,007 (97.48)	73,890 (99.84)	117 (0.16)		
Menopause				29.26 (1)	*<.001*
No	26,927 (35.47)	26,913 (99.95)	14 (0.05)		
Yes	48,990 (64.53)	48,885 (99.79)	105 (0.21)		
Fertility status				1.67 (1)	.20
No	1047 (1.38)	1047 (100)	0 (0)		
Yes	74,870 (98.62)	74,751 (99.84)	119 (0.16)		
Lactation status				0.06 (1)	.80
No	4233 (5.58)	4227 (99.86)	6 (0.14)		
Yes	71,684 (94.42)	71,571 (99.84)	113 (0.16)		
History of benign breast disease				3.61 (1)	.06
No	53,977 (71.1)	53,883 (99.83)	94 (0.17)		
Yes	21,940 (28.9)	21,915 (99.89)	25 (0.11)		
History of reproductive system surgery				1.75 (1)	.19
No	60,480 (79.67)	60,391 (99.85)	89 (0.15)		
Yes	15,437 (20.33)	15,407 (99.81)	30 (0.19)		

[a] Percentages in this column have denominators of N=75,917.

[b] Percentages in these columns have the n value in the “Total” column in the same row as the denominators.

[c] Italicized values indicate statistical significance.

[2] Low=primary school or below; medium=junior or senior high school; high=undergraduate degree or above.

Development of the Lung Cancer Risk Assessment Model

Table 2 displays the hazard ratios (HRs) with its 95% CI for every indicator. In the training set, age (≥55 years: HR 1.34, 95% CI 0.38-4.80; ≥60 years: HR 2.33, 95% CI 0.67-8.11; ≥65 years: HR 2.41, 95% CI 0.69-8.49; ≥70 years: HR 1.79, 95% CI 0.43-7.40), history of chronic respiratory disease (HR 1.94, 95% CI 1.24-3.04), first-degree family history of lung cancer (HR 1.60, 95% CI 0.91-2.83), menopause (HR 2.16, 95% CI 0.90-5.19), and history of benign breast disease (HR 0.58, 95% CI 0.36-0.94) were independent risk factors for lung cancer. Consequently, we applied these parameters to construct the model. We drew 1-year, 3-year, and 5-year risk–predicting nomograms for lung cancer (Figure 2A).

Table 2. Multivariable Cox regression prediction model of lung cancer risk in the training set.

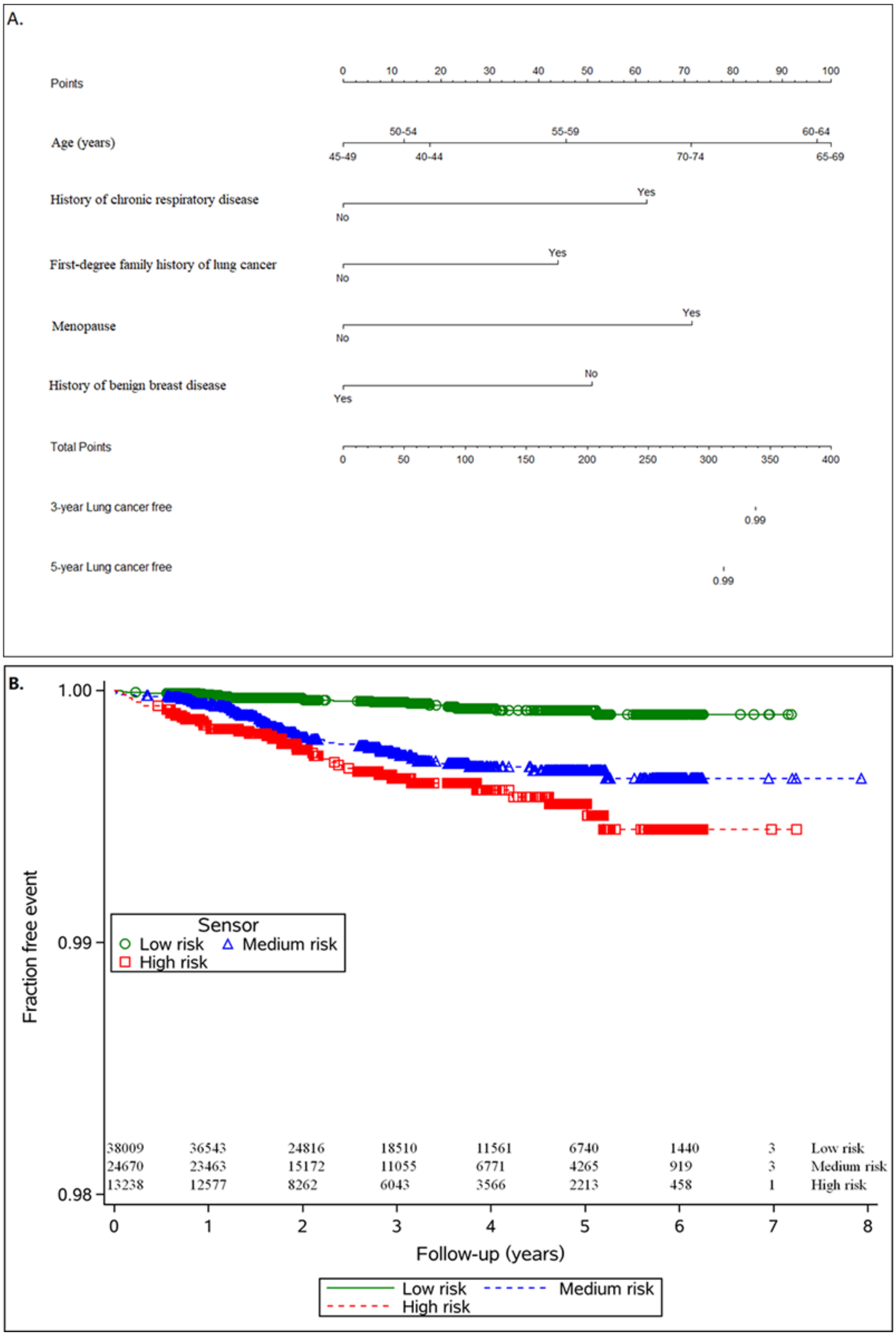
Variables	β coefficient	SE	HR[a] (95% CI)	χ^2 (*df*)	*P* value
Age (years)					
40-44	N/A[b]	N/A	1.00	N/A	N/A
45-49	–0.19	0.59	0.83 (0.26-2.64)	0.10 (1)	.75
50-54	–0.06	0.62	0.94 (0.28-3.19)	0.01 (1)	.93
55-59	0.30	0.65	1.34 (0.38-4.80)	0.21 (1)	.65
60-64	0.85	0.64	2.33 (0.67-8.11)	1.78 (1)	.18
65-69	0.88	0.64	2.41 (0.69-8.49)	1.89 (1)	.17
70-74	0.58	0.72	1.79 (0.43-7.40)	0.65 (1)	.42
History of chronic respiratory disease					
No	N/A	N/A	1.00	N/A	N/A
Yes	0.66	0.23	1.94 (1.24-3.04)	8.45 (1)	*.004* [c]
First-degree family history of lung cancer					
No	N/A	N/A	1.00	N/A	N/A
Yes	0.47	0.29	1.60 (0.91-2.83)	2.63 (1)	.11
Menopause					
No	N/A	N/A	1.00	N/A	N/A
Yes	0.77	0.45	2.16 (0.90-5.19)	2.95 (1)	.09
History of benign breast disease					
No	N/A	N/A	1.00	N/A	N/A
Yes	–0.55	0.25	0.58 (0.36-0.94)	4.97 (1)	*.03*

[a] HR: hazard ratio.

[b] N/A: not applicable.

[c] Italicized values indicate statistical significance.

Figure 2. (A) Nomogram to calculate the personal 1-year, 3-year, and 5-year risk of lung cancer, and (B) the lung cancer incidence across different cancer risk categories.



Predictive Performance of the Model

The risk predictions were categorized into low-risk, medium-risk, and high-risk categories, and a log-rank test revealed significant differences across the 3 groups (Figure 2B; *P*<.001).

By using this model, the AUC for 1-year, 3-year, and 5-year lung cancer risk in the training set was 0.762, 0.718, and 0.703, respectively. The model yielded a greater AUC for passive smokers (1-year: 0.787, 3-year: 0.715, and 5-year: 0.745) than for nonpassive smokers (1-year: 0.741, 3-year: 0.721, and 5-year: 0.689; Figure 3). Calibration was acceptable, with very similar observed and predicted hazards (Figure 4).

Figure 3. The receiver operating characteristic curves of prediction models in the training set. (A) Whole population; (B) Nonpassive smokers; (C) Passive smokers. AUC: area under the curve.

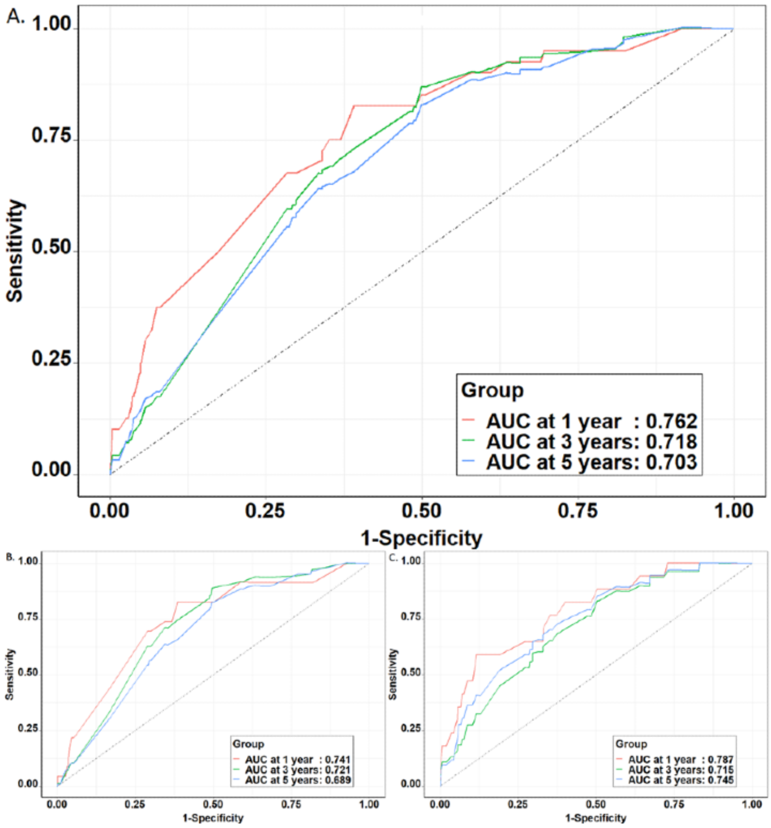
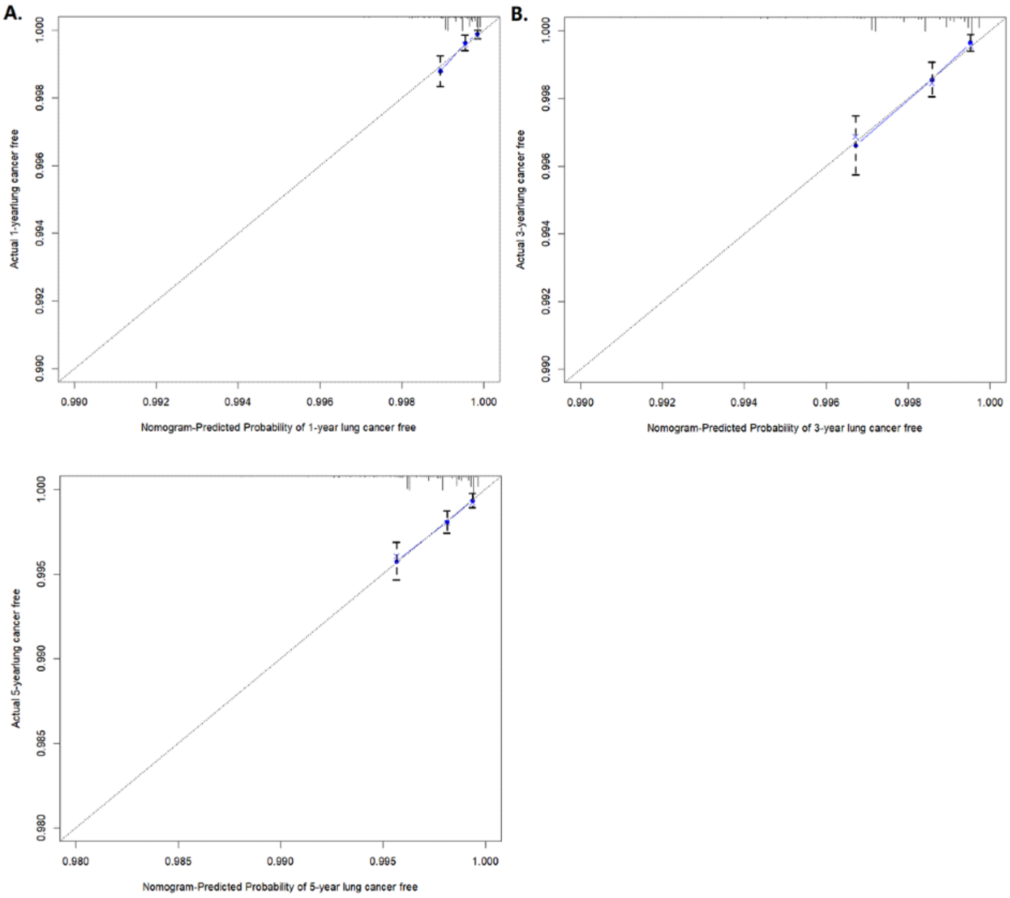


Figure 4. Calibration curves of the nomogram for (A) 1-year, (B) 3-year and (C) 5-year lung cancer–free rates in the training set.



Validation of the Lung Cancer Risk Model

The model demonstrated a moderate predictive discrimination in the validation set, with AUCs of 0.646, 0.658, and 0.650 for 1-year, 3-year, and 5-year lung cancer risks, respectively (Figure S1 in Multimedia Appendix 1), and satisfactory calibration of relative risk (Figure S2 in Multimedia Appendix 1).

Discussion

We constructed and validated a simple risk predictive model internally for lung cancer in nonsmoking women relying on 5 commonly accessible factors such as demographics (age), comorbidities (chronic respiratory disease), first-degree family history of lung cancer, and fertility (menopause and history of benign breast disease). Our results showed that the model has moderate discriminatory accuracy and goodness of fit for both nonpassive smokers and passive smokers.

Multiple lung cancer risk variables were discovered for nonsmoking women, such as passive smoking [51,52], prior lung diseases (tuberculosis, chronic bronchitis, emphysema, and prior lung disorders [chronic obstructive pulmonary disease]) [53], indoor radon [54], cooking oil fume [55], and a family history of lung cancer [56]. The established risk variables for lung cancer, such as age, a family history of lung cancer, and a history of chronic respiratory disease, are similar to the findings. Age is the most important risk variable for lung cancer in nonsmoking women according to our survey, which found that the risk was more than 2.4 times higher in the age group of 65-69 years than 40-44 years.

Menopause was associated with an increased risk of developing lung cancer, with an overall odds ratio of 1.33 (95% CI 0.90-1.96), according to a pooled analysis of nested case-control data [57], which is consistent with our findings. Interestingly, we found that women with a history of benign breast disease were less likely to develop lung cancer, possibly because these women may be more careful about their lifestyle and diet after developing breast disease than those who did not. This finding will need to be validated in future studies.

Besides the accurate indicators, risk predicting models should achieve performance standards for discrimination (the differentiation capacity to distinguish lung cancer cases from control ones) and calibration (defined as the consistency between observed and predicted risk for lung cancer). Since 2010, the substantial growth in the numbers of investigations on lung cancer risk predicting models shows the necessity of using predictive models to drive population triage. Initially, models, such as the Bach model [12], Spitz model [13], Liverpool Lung Project model [14], and PLCO$_{M2012}$ model [58], emphasized the importance of applying the classic epidemiological risk variables, including age, smoking history, personal history of disease, and family history of cancer. To the best of our knowledge, this study is one of the few studies to model the prediction of lung cancer risk among nonsmoking Chinese women. Due to the fact that each model was created in a distinct population with different baseline risks and lengths of follow-up, it is challenging to compare the discriminating performance of risk predictive models. The discriminating ability of every model was quite equal, with C-statistics ranging between 0.72 and 0.86. Compared to prior research, our models showed comparable predictive performance.

In understanding our findings, certain strengths and limitations should be carefully considered. Our research is conducted on a large population-based cancer-screening program in mainland China, which is a strong point. In addition, the variables included in this model could be easily collected and updated without any imaging, sophisticated testing, or calculation. Furthermore, the model will be used as a convenient method to triage high-risk people among nonsmoking women, and it will be involving in public health initiatives, such as recommendations regarding the control of lung cancer in nonsmokers. Nonetheless, the statistics based on self-report may be susceptible to social acceptability bias as well as recall bias. Since data collection and quality control were carried out to a high standard, the vast volume of information can be relied upon. Second, the performance of our risk-predicting model was not validated against an external data set before it was used. The findings of the internal calibration, on the other hand, suggest that this model will function satisfactorily when applied to a variety of populations.

In conclusion, a large-scale lung cancer–screening project in China served as the foundation for the creation and internal calibration of a straightforward risk predictive model for lung cancer in nonsmoking women. The model has moderate discrimination and could be used as a tool for triaging high-risk people to prevent lung cancer in nonsmoking women. To validate the concept in external populations, additional prospective studies are needed.

Acknowledgments

This study was supported by the Natural Science Foundation of Henan Province (No. 212300410261) and the training project for young and middle-aged excellent talents in health science and technology innovation of Henan province (YXKC2022045). We sincerely thank all the members of the Cancer Screening Program in Urban China from the National Cancer Center of China and Henan province. We are also grateful to the participants for be involved in this study.

Data Availability

The data sets for this manuscript are not publicly available because all our data are under regulation of both the National Cancer Center of China and The Affiliated Cancer Hospital of Zhengzhou University and Henan Cancer Hospital. Requests to access the data sets should be directed to SZ.

Authors' Contributions

LG and SZ contributed to the conception and design. LG and L Zheng contributed to statistical analyses. LG, QM, L Zheng, QC, YL, HX, RK, L Zhang, SL, XS, and SZ contributed to data acquisition and data interpretation. LG drafted the article. All authors revised the manuscript and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary table and figures.
[DOCX File , 169 KB - publichealth_v9i1e41640_app1.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Zhang SW, Zheng RS, Yang ZX, Zeng HM, Sun KX, Gu XY, et al. Trend analysis on incidence and age at diagnosis for lung cancer in cancer registration areas of China, 2000-2014. Article in Chinese. Zhonghua Yu Fang Yi Xue Za Zhi 2018 Jun 06;52(6):579-585. [doi: 10.3760/cma.j.issn.0253-9624.2018.06.005] [Medline: 29886678]
3. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. Nat Rev Cancer 2007 Oct;7(10):778-790. [doi: 10.1038/nrc2190] [Medline: 17882278]
4. Couraud S, Zalcman G, Milleron B, Morin F, Souquet PJ. Lung cancer in never smokers--a review. Eur J Cancer 2012 Jun;48(9):1299-1311. [doi: 10.1016/j.ejca.2012.03.007] [Medline: 22464348]
5. Scagliotti GV, Longo M, Novello S. Nonsmall cell lung cancer in never smokers. Curr Opin Oncol 2009 Mar;21(2):99-104. [doi: 10.1097/CCO.0b013e328321049e] [Medline: 19532009]
6. Shi JF, Wang L, Wu N, Li JL, Hui ZG, Liu SM, LuCCRES Group. Clinical characteristics and medical service utilization of lung cancer in China, 2005-2014: overall design and results from a multicenter retrospective epidemiologic survey. Lung Cancer 2019 Feb;128:91-100. [doi: 10.1016/j.lungcan.2018.11.031] [Medline: 30642458]
7. Zeng H, Chen W, Zheng R, Zhang S, Ji JS, Zou X, et al. Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. Lancet Glob Health 2018 May;6(5):e555-e567 [FREE Full text] [doi: 10.1016/S2214-109X(18)30127-X] [Medline: 29653628]
8. Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. Chest 2017 Jan;151(1):193-203. [doi: 10.1016/j.chest.2016.10.010] [Medline: 27780786]
9. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011 Aug 04;365(5):395-409 [FREE Full text] [doi: 10.1056/NEJMoa1102873] [Medline: 21714641]
10. Li Q, Hsia J, Yang G. Prevalence of smoking in China in 2010. N Engl J Med 2011 Jun 23;364(25):2469-2470. [doi: 10.1056/NEJMc1102459] [Medline: 21696322]
11. Cepeda-Benito A, Doogan NJ, Redner R, Roberts ME, Kurti AN, Villanti AC, et al. Trend differences in men and women in rural and urban U.S. settings. Prev Med 2018 Dec;117:69-75. [doi: 10.1016/j.ypmed.2018.04.008] [Medline: 29627511]
12. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, et al. Variations in lung cancer risk among smokers. J Natl Cancer Inst 2003 Mar 19;95(6):470-478. [doi: 10.1093/jnci/95.6.470] [Medline: 12644540]
13. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, et al. A risk model for prediction of lung cancer. J Natl Cancer Inst 2007 May 02;99(9):715-726. [doi: 10.1093/jnci/djk153] [Medline: 17470739]
14. Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, et al. The LLP risk model: an individual risk prediction model for lung cancer. Br J Cancer 2008 Jan 29;98(2):270-276 [FREE Full text] [doi: 10.1038/sj.bjc.6604158] [Medline: 18087271]
15. Etzel CJ, Kachroo S, Liu M, D'Amelio A, Dong Q, Cote ML, et al. Development and validation of a lung cancer risk prediction model for African-Americans. Cancer Prev Res (Phila) 2008 Sep;1(4):255-265 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-08-0082] [Medline: 19138969]
16. Spitz MR, Etzel CJ, Dong Q, Amos CI, Wei Q, Wu X, et al. An expanded risk prediction model for lung cancer. Cancer Prev Res (Phila) 2008 Sep;1(4):250-254 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-08-0060] [Medline: 19138968]
17. Young RP, Hopkins RJ, Hay BA, Epton MJ, Mills GD, Black PN, et al. Lung cancer susceptibility model based on age, family history and genetic variants. PLoS One 2009 Apr 23;4(4):e5302 [FREE Full text] [doi: 10.1371/journal.pone.0005302] [Medline: 19390575]
18. D'Amelio AMJ, Cassidy A, Asomaning K, Raji OY, Duffy SW, Field JK, et al. Comparison of discriminatory power and accuracy of three lung cancer risk models. Br J Cancer 2010 Jul 27;103(3):423-429 [FREE Full text] [doi: 10.1038/sj.bjc.6605759] [Medline: 20588271]

19. Raji OY, Agbaje OF, Duffy SW, Cassidy A, Field JK. Incorporation of a genetic factor into an epidemiologic model for prediction of individual risk of lung cancer: the Liverpool Lung Project. Cancer Prev Res (Phila) 2010 May;3(5):664-669. [doi: 10.1158/1940-6207.CAPR-09-0141] [Medline: 20424129]
20. Maisonneuve P, Bagnardi V, Bellomi M, Spaggiari L, Pelosi G, Rampinelli C, et al. Lung cancer risk prediction to select smokers for screening CT--a model based on the Italian COSMOS trial. Cancer Prev Res (Phila) 2011 Nov;4(11):1778-1789. [doi: 10.1158/1940-6207.CAPR-11-0026] [Medline: 21813406]
21. Tammemagi CM, Pinsky PF, Caporaso NE, Kvale PA, Hocking WG, Church TR, et al. Lung cancer risk prediction: Prostate, Lung, Colorectal And Ovarian Cancer Screening Trial models and validation. J Natl Cancer Inst 2011 Jul 06;103(13):1058-1068 [FREE Full text] [doi: 10.1093/jnci/djr173] [Medline: 21606442]
22. Tammemagi MC, Lam SC, McWilliams AM, Sin DD. Incremental value of pulmonary function and sputum DNA image cytometry in lung cancer risk prediction. Cancer Prev Res (Phila) 2011 Apr;4(4):552-561. [doi: 10.1158/1940-6207.CAPR-10-0183] [Medline: 21411501]
23. Hoggart C, Brennan P, Tjonneland A, Vogel U, Overvad K, Østergaard JN, et al. A risk model for lung cancer incidence. Cancer Prev Res (Phila) 2012 Jun;5(6):834-846 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-11-0237] [Medline: 22496387]
24. Li H, Yang L, Zhao X, Wang J, Qian J, Chen H, et al. Prediction of lung cancer risk in a Chinese population using a multifactorial genetic model. BMC Med Genet 2012 Dec 10;13(1):118 [FREE Full text] [doi: 10.1186/1471-2350-13-118] [Medline: 23228068]
25. Raji OY, Duffy SW, Agbaje OF, Baker SG, Christiani DC, Cassidy A, et al. Predictive accuracy of the Liverpool Lung Project risk model for stratifying patients for computed tomography screening for lung cancer: a case-control and cohort validation study. Ann Intern Med 2012 Aug 21;157(4):242-250 [FREE Full text] [doi: 10.7326/0003-4819-157-4-201208210-00004] [Medline: 22910935]
26. Park S, Nam BH, Yang HR, Lee JA, Lim H, Han JT, et al. Individualized risk prediction model for lung cancer in Korean men. PLoS One 2013 Feb 07;8(2):e54823 [FREE Full text] [doi: 10.1371/journal.pone.0054823] [Medline: 23408946]
27. Spitz MR, Amos CI, Land S, Wu X, Dong Q, Wenzlaff AS, et al. Role of selected genetic variants in lung cancer risk in African Americans. J Thorac Oncol 2013 Apr;8(4):391-397 [FREE Full text] [doi: 10.1097/JTO.0b013e318283da29] [Medline: 23454887]
28. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection criteria for lung-cancer screening. N Engl J Med 2013 Feb 21;368(8):728-736 [FREE Full text] [doi: 10.1056/NEJMoa1211776] [Medline: 23425165]
29. Veronesi G, Maisonneuve P, Rampinelli C, Bertolotti R, Petrella F, Spaggiari L, et al. Computed tomography screening for lung cancer: results of ten years of annual screening and validation of cosmos prediction model. Lung Cancer 2013 Dec;82(3):426-430. [doi: 10.1016/j.lungcan.2013.08.026] [Medline: 24099665]
30. El-Zein RA, Lopez MS, D'Amelio AM, Liu M, Munden RF, Christiani D, et al. The cytokinesis-blocked micronucleus assay as a strong predictor of lung cancer: extension of a lung cancer risk prediction model. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2462-2470 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0462] [Medline: 25172871]
31. Li K, Hüsing A, Sookthai D, Bergmann M, Boeing H, Becker N, et al. Selecting high-risk individuals for lung cancer screening: a prospective evaluation of existing risk models and eligibility criteria in the German EPIC cohort. Cancer Prev Res (Phila) 2015 Sep;8(9):777-785. [doi: 10.1158/1940-6207.CAPR-14-0424] [Medline: 26076698]
32. Marcus MW, Chen Y, Raji OY, Duffy SW, Field JK. LLPi: Liverpool Lung Project risk prediction model for lung cancer incidence. Cancer Prev Res (Phila) 2015 Jun;8(6):570-575. [doi: 10.1158/1940-6207.CAPR-14-0438] [Medline: 25873368]
33. Wang X, Ma K, Cui J, Chen X, Jin L, Li W. An individual risk prediction model for lung cancer based on a study in a Chinese population. Tumori 2015 Feb 14;101(1):16-23. [doi: 10.5301/tj.5000205] [Medline: 25702657]
34. Marcus MW, Raji OY, Duffy SW, Young RP, Hopkins RJ, Field JK. Incorporating epistasis interaction of genetic susceptibility single nucleotide polymorphisms in a lung cancer risk prediction model. Int J Oncol 2016 Jul;49(1):361-370 [FREE Full text] [doi: 10.3892/ijo.2016.3499] [Medline: 27121382]
35. Wang X, Ma K, Chi L, Cui J, Jin L, Hu J, et al. Combining telomerase reverse transcriptase genetic variant rs2736100 with epidemiologic factors in the prediction of lung cancer susceptibility. J Cancer 2016;7(7):846-853 [FREE Full text] [doi: 10.7150/jca.13437] [Medline: 27162544]
36. Wu X, Wen CP, Ye Y, Tsai M, Wen C, Roth JA, et al. Personalized risk assessment in never, light, and heavy smokers in a prospective cohort in Taiwan. Sci Rep 2016 Nov 02;6:36482 [FREE Full text] [doi: 10.1038/srep36482] [Medline: 27805040]
37. Muller DC, Johansson M, Brennan P. Lung cancer risk prediction model incorporating lung function: development and validation in the UK Biobank prospective cohort study. J Clin Oncol 2017 Mar 10;35(8):861-869 [FREE Full text] [doi: 10.1200/JCO.2016.69.2467] [Medline: 28095156]
38. Weber M, Yap S, Goldsbury D, Manners D, Tammemagi M, Marshall H, et al. Identifying high risk individuals for targeted lung cancer screening: independent validation of the PLCO risk prediction tool. Int J Cancer 2017 Jul 15;141(2):242-253 [FREE Full text] [doi: 10.1002/ijc.30673] [Medline: 28249359]

39. Charvat H, Sasazuki S, Shimazu T, Budhathoki S, Inoue M, Iwasaki M, JPHC Study Group. Development of a risk prediction model for lung cancer: The Japan Public Health Center-based Prospective Study. Cancer Sci 2018 Mar 21;109(3):854-862 [FREE Full text] [doi: 10.1111/cas.13509] [Medline: 29345859]
40. Katki HA, Kovalchik SA, Petito LC, Cheung LC, Jacobs E, Jemal A, et al. Implications of nine risk prediction models for selecting ever-smokers for computed tomography lung cancer screening. Ann Intern Med 2018 Jul 03;169(1):10-19 [FREE Full text] [doi: 10.7326/M17-2701] [Medline: 29800127]
41. Markaki M, Tsamardinos I, Langhammer A, Lagani V, Hveem K, Røe OD. A validated clinical risk prediction model for lung cancer in smokers of all ages and exposure types: a HUNT study. EBioMedicine 2018 May;31:36-46 [FREE Full text] [doi: 10.1016/j.ebiom.2018.03.027] [Medline: 29678673]
42. Lyu ZY, Li N, Chen SH, Wang G, Tan FW, Feng XS, et al. Exploratory research on developing lung cancer risk prediction model in female non-smokers. Article in Chinese. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Nov 06;54(11):1261-1267. [doi: 10.3760/cma.j.cn112150-20200805-01093] [Medline: 33147927]
43. Lam S. Lung cancer screening in never-smokers. J Thorac Oncol 2019 Mar;14(3):336-337 [FREE Full text] [doi: 10.1016/j.jtho.2018.12.019] [Medline: 30782380]
44. Toh C, Gao F, Lim W, Leong S, Fong K, Yap S, et al. Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. J Clin Oncol 2006 May 20;24(15):2245-2251. [doi: 10.1200/JCO.2005.04.8033] [Medline: 16710022]
45. Sisti J, Boffetta P. What proportion of lung cancer in never-smokers can be attributed to known risk factors? Int J Cancer 2012 Jul 15;131(2):265-275 [FREE Full text] [doi: 10.1002/ijc.27477] [Medline: 22322343]
46. Tammemägi MC, Church TR, Hocking WG, Silvestri GA, Kvale PA, Riley TL, et al. Evaluation of the lung cancer risks at which to screen ever- and never-smokers: screening rules applied to the PLCO and NLST cohorts. PLoS Med 2014 Dec;11(12):e1001764 [FREE Full text] [doi: 10.1371/journal.pmed.1001764] [Medline: 25460915]
47. Guo LW, Chen Q, Shen YC, Meng QC, Zheng LY, Wu Y, et al. Evaluation of a low-dose computed tomography lung cancer screening program in Henan, China. JAMA Netw Open 2020 Nov 02;3(11):e2019039 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19039] [Medline: 33141158]
48. Guo L, Zhang S, Liu S, Zheng L, Chen Q, Cao X, et al. Determinants of participation and detection rate of upper gastrointestinal cancer from population-based screening program in China. Cancer Med 2019 Nov;8(16):7098-7107 [FREE Full text] [doi: 10.1002/cam4.2578] [Medline: 31560836]
49. Chen C, Lu FC, Department of Disease Control Ministry of Health, PR China. The guidelines for prevention and control of overweight and obesity in Chinese adults. Biomed Environ Sci 2004;17 Suppl:1-36. [Medline: 15807475]
50. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 2013 Mar 06;13:33 [FREE Full text] [doi: 10.1186/1471-2288-13-33] [Medline: 23496923]
51. Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. BMJ 1997 Oct 18;315(7114):980-988 [FREE Full text] [doi: 10.1136/bmj.315.7114.980] [Medline: 9365295]
52. Stayner L, Bena J, Sasco AJ, Smith R, Steenland K, Kreuzer M, et al. Lung cancer risk and workplace exposure to environmental tobacco smoke. Am J Public Health 2007 Mar;97(3):545-551. [doi: 10.2105/AJPH.2004.061275] [Medline: 17267733]
53. Brenner DR, McLaughlin JR, Hung RJ. Previous lung diseases and lung cancer risk: a systematic review and meta-analysis. PLoS One 2011 Mar 31;6(3):e17479 [FREE Full text] [doi: 10.1371/journal.pone.0017479] [Medline: 21483846]
54. Darby S, Hill D, Auvinen A, Barros-Dios JM, Baysson H, Bochicchio F, et al. Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. BMJ 2005 Jan 29;330(7485):223 [FREE Full text] [doi: 10.1136/bmj.38308.477650.63] [Medline: 15613366]
55. Zhao Y, Wang S, Aunan K, Seip HM, Hao J. Air pollution and lung cancer risks in China--a meta-analysis. Sci Total Environ 2006 Aug 01;366(2-3):500-513. [doi: 10.1016/j.scitotenv.2005.10.010] [Medline: 16406110]
56. Lissowska J, Foretova L, Dabek J, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, et al. Family history and lung cancer risk: international multicentre case-control study in Eastern and Central Europe and meta-analyses. Cancer Causes Control 2010 Jul;21(7):1091-1104. [doi: 10.1007/s10552-010-9537-2] [Medline: 20306329]
57. Jin K, Hung RJ, Thomas S, Le Marchand L, Matsuo K, Seow A, et al. Hormonal factors in association with lung cancer among Asian women: a pooled analysis from the International Lung Cancer Consortium. Int J Cancer 2021 May 01;148(9):2241-2254 [FREE Full text] [doi: 10.1002/ijc.33405] [Medline: 33210298]
58. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection criteria for lung-cancer screening. N Engl J Med 2013 Feb 21;368(8):728-736 [FREE Full text] [doi: 10.1056/NEJMoa1211776] [Medline: 23425165]

Abbreviations

AUC: area under the curve
CanSPUC: Cancer Screening Program in Urban China
HR: hazard ratio
PLCO: Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial

Edited by A Mavragani; submitted 03.08.22; peer-reviewed by M Dai, L Wang; comments to author 25.10.22; revised version received 04.11.22; accepted 25.11.22; published 06.01.23.

Please cite as:
Guo L, Meng Q, Zheng L, Chen Q, Liu Y, Xu H, Kang R, Zhang L, Liu S, Sun X, Zhang S
Lung Cancer Risk Prediction Nomogram in Nonsmoking Chinese Women: Retrospective Cross-sectional Cohort Study
JMIR Public Health Surveill 2023;9:e41640
URL: https://publichealth.jmir.org/2023/1/e41640
doi:10.2196/41640
PMID:36607729

Original Paper

Association of Depression With Precautionary Behavior Compliance, COVID-19 Fear, and Health Behaviors in South Korea: National Cross-sectional Study

Hyerine Shin[1], MSN; Ji-Su Kim[1], PhD; HyunHae Lee[1], MSN

[1] Department of Nursing, Chung-ang University, Seoul, Republic of Korea

Corresponding Author:
Ji-Su Kim, PhD
Department of Nursing
Chung-ang University
84 Heukseok-ro, Dongjak-gu
Seoul, 06974
Republic of Korea
Phone: 82 10 7117 5088
Email: jisu80@cau.ac.kr

Abstract

Background: As of January 2022, the number of people infected with COVID-19 worldwide has exceeded 350 million. As the COVID-19 pandemic continues, people are affected in a wide range of areas of life, which in turn causes numerous psychological problems. Depression is a serious problem for people who have suffered from COVID-19. Depression can worsen COVID-19 precautionary behavior compliance or the health behavior itself. In addition, these depressive symptoms may have different characteristics depending on the individual's gender.

Objective: The aim of this study was to determine whether depression is a factor that may affect COVID-19 fear, precautionary behavior compliance, and health behavior, and how these characteristic trends differ by gender.

Methods: This was a secondary analysis of data from the 2020 Korea Community Health Survey (KCHS), a national cross-sectional survey conducted with complex sampling analysis. In 2020, the KCHS included COVID-19–related questions. For this study, we used the KCHS data from both the COVID-19–related questions and the Patient Health Questionnaire-9 scale. After weighting the data according to the KCHS guidelines, we calculated the distribution of men and women according to depression level. The data were collected using multiple-choice questions related to precautionary behavior compliance, COVID-19–related fears, and health behavior changes.

Results: Of the 204,787 participants, those who were clinically depressed had a greater tendency to not comply with precautionary behaviors. Regarding COVID-19, "fear" showed a decreasing trend in both men (adjusted odds ratio [AOR] 0.72, 95% CI 0.61-0.83) and women (AOR 0.74, 95% CI 0.63-0.86) with clinically relevant depression. Moreover, for both men and women, health behaviors deteriorated as depression intensified; the AOR for sleep duration changes was 2.28 (95% CI 2.00-2.59) in men and was 2.15 (95% CI 1.96-2.36) in women. Notably, the responses of clinically depressed women revealed a doubled increase in both their drinking (AOR 2.25, 95% CI 1.88-2.70) and smoking (AOR 2.71, 95% CI 1.95-3.77) habits compared with those of nondepressed women.

Conclusions: Both men and women with more severe depression were more likely to violate precautionary health behaviors as their depression worsened. Health behaviors also deteriorated for both genders, but women tended to show a greater change. Therefore, additional studies and interventions for vulnerable groups such as severely depressed people are needed. More research is also necessary to develop interventions based on statistical comparisons of men and women.

(JMIR Public Health Surveill 2023;9:e42677) doi:10.2196/42677

KEYWORDS

COVID-19; precautionary behaviors; COVID-19 fear; health behavior deterioration; gender differences

Introduction

As of January 25, 2022, 349,641,119 people worldwide had been infected with COVID-19 and the death toll had reached 5,592,266 [1]. Since the World Health Organization's March 11, 2020, declaration that the world was facing a pandemic, various restrictions have been implemented to prevent the global spread of the virus. These restrictions have included lockdowns for all or some educational facilities and nonessential businesses, travel limitations, quarantine measures, social isolation, and mandatory mask-wearing [2-4]. However, COVID-19 has had a profound impact on all countries, racial groups, and socioeconomic classes [5]. Secondary problems arise from these restrictions.

Among all the problems that have occurred, mental health issues related to COVID-19 are among the most important health concerns and have received considerable attention. Direct stress caused by illness and concerns about medical service interruption, employment, financial stability, and social contact restrictions may cause psychological problems [6-9]. Additionally, COVID-19 fear is associated with several mental health problems such as anxiety, depression, psychological distress, and sleep issues [10-12]. In particular, depression is one of the most representative mental health problems during COVID-19, which is associated with deterioration in health behaviors, including decreased physical activity and sleep and increased smoking [13-15].

Moreover, various empirical studies have reported gender differences in depression associated with the COVID-19 pandemic [16-18]. Depression is known to have some gender-specific characteristics, and depression rates are generally higher in women regardless of COVID-19 [19,20]. For example, among Australian adults, women were found to be more likely to have clinically significant depression and anxiety [17]. In the case of COVID-19, these differences still apply, as indicated by a study of adults aged 25-69 years showing that women reported more depression, anxiety, stress, loneliness, and a higher rate of increased cannabis use compared to pre-COVID-19 levels [16].

Mental health is not only a concern in and of itself, but may also impede health behaviors or COVID-19 precautionary actions [15,21,22]. Olapegba et al [23] reported that psychological distress caused by COVID-19 fear hinders precautionary behavior compliance. Moreover, depression can lead to even more significant health problems by reducing COVID-19 precautionary behavior compliance and general healthy behavior. In addition, given that severe degrees of fear can negatively impact an individual's ability to respond appropriately during the COVID-19 pandemic [24], the impact of depression may vary in severity, and, at the same time, gender differences may also exist. To date, studies have been conducted on the gender gap in depression related to COVID-19 or the following problems: medical service interruption, employment, financial stability, social contact restrictions [25-27], and the impact of depression on COVID-19 precautionary behavior compliance [22,28]. However, studies that have analyzed the severity of depression according to gender are scarce. Therefore, the purpose of this study was to determine whether depression is a factor that worsens COVID-19 fears, precautionary behavior compliance, and health behavior, and whether these characteristic trends differ by gender.

Methods

Study Design

This was a secondary analysis of data from a national cross-sectional survey: the 2020 Korea Community Health Survey (KCHS) conducted by the Korea Centers for Disease Control and Prevention [29]. The Community Health Survey, initiated in 2008, is an annual, anonymous, computer-assisted personal interview survey that evaluates the health status and behaviors of South Korean adults. A trained surveyor visits the households selected as samples and conducts a one-on-one interview using a laptop computer equipped with a survey program. The 2020 Community Health Survey covered 18 geographical areas and comprised 142 survey questions, including those pertaining to health behaviors such as exercise, smoking, and drinking. Given the global health context in 2020, the survey for that year was revised to include questions related to the COVID-19 pandemic.

The data were collected from August 16, 2020, to October 31, 2020 [29]. The target population of 2020 KCHS was adults aged 19 years or older, and a sample group was extracted through a composite sample design to minimize sample bias [30]. Next, in the first official extraction, probability proportional to size sampling was performed to ensure that the extraction probability was proportional to the number and size of households by housing type in Tong, Ban, and Ri (the smallest administrative district units). The second extraction was a systematic extraction based on the number of households in Tong, Ban, and Ri. In total, 229,269 people were surveyed; after excluding incomplete responses, the data from 204,787 respondents were available for analysis. To request the KCHS data for use in the present study, we signed a written pledge to confirm our intentions and submitted a usage plan describing the purpose for which we intended to use the data if, after the Korea Centers for Disease Control and Prevention reviewed the plan, the data request was approved. Only those who have received approval can download and use the raw data of KCHS.

Ethical Approval

This study was conducted after review from the Institutional Review Board (IRB) of Chung-ang University. Because this was a secondary analysis of data from a national cross-sectional survey, the study received an exemption by the IRB (approval number 1041078-202201-HR-002).

Measures

General Participant Characteristics

To determine the participants' general characteristics, we included data about their sociodemographic characteristics and individual health. Sociodemographic characteristics were assessed according to age, gender, educational level, monthly household income, profession, and region. Age was divided into four generational groups based on the criteria used in the

Korean Longitudinal Study of Aging and previous studies [31-33]: 20 to 44 years (youth), 45 to 64 years (middle age), and 65 years or older (older adults). According to a previous study, education levels were categorized as follows: (1) elementary or less, (2) middle school, (3) high school, and (4) college or higher. Monthly household income was classified into quartiles as Q1 (<25%), Q2 (25%-49%), Q3 (50%-74%), and Q4 (≥75%), and professions were organized according to the Korean version of the Standard Classification of Occupations and previous studies [34,35]. After referring to previous research, we reclassified occupations into four categories: white (white collar), pink (sales and service), blue (agricultural, forestry, fishing, and military), and unemployed. Region was divided into urban and rural areas.

Questions related to individual health conditions included drinking at least once a month, current smoking status, stress awareness, and subjective health status. Possible answers for drinking at least once per month, current smoking status, and stress awareness were either “yes” or “no.” Subjective health status was defined by the self-evaluation of one’s health as “very good” or “good,” and possible answers were “poor” and “fair.”

Depression Assessment

The Patient Health Questionnaire-9 (PHQ-9) was used to measure self-reported depression, with possible scores ranging from 0 to 27 [35]. PHQ-9 consists of nine items used to evaluate the diagnostic criteria for major depressive disorder in accordance with the Diagnostic and Statistical Manual of Mental Disorders, fifth edition. A high PHQ-9 score indicates a higher level of depressive symptoms, with 0-4 representing no depression, 5-9 indicating mild depression, 10-14 representing moderate depression, 15-19 indicating moderately severe depression, and 20-27 representing severe depression. In general, a PHQ-9 score ≥10 indicates major or clinically relevant depression [36]. To ensure greater clarity in the explanation of the results, we divided the PHQ-9 scores into three categories: 0-4 (none), 5-9 (mild), and ≥10 (clinically relevant).

Precautionary Behavior Compliance

Precautionary behavior compliance was defined as whether personal quarantine rules (eg, wearing a mask indoors, wearing a mask outdoors) and social distancing rules (eg, maintaining a healthy distance) were observed for the past week with the intention of helping to prevent the spread of COVID-19. According to a previous study [37], eight factors were used to evaluate precautionary behavior compliance as follows: in the last 7 days, had the individual (1) covered his or her mouth and nose while sneezing or coughing, (2) experienced daily ventilation at least twice a day, (3) had everyday spaces disinfected once per day, (4) wore a mask in facilities used by unspecified people, (5) worn a mask when it was impossible to keep an adequate distance, (6) maintained a 2-meter distance between themselves and others, (7) limited visiting sick people, and (8) limited going out or attending gatherings and events? Based on the KCHS’s indicator definition, for each of the eight items, only those who answered “very much” or “yes” were considered to have complied with each precautionary behavior. For example, only the respondents who answered “very much” or “yes” to the question “covered his or her mouth and nose while sneezing or coughing” were considered to have performed precautionary behavior.

COVID-19 Fear

COVID-19 fear is defined as a psychological concern due to the spread of COVID-19 and is assessed by five factors: fear of infection, dying from infection, public criticism, a family member getting infected, and economic loss due to infection. Based on the KCHS’s indicator definition, only those who answered “strongly agree” or “agree” for any of the above five factors were evaluated as having COVID-19 fear of each factor.

Health Behavior Deterioration

Criteria pertaining to physical activity, sleep duration, fast food or energy drink consumption, food consumption, alcohol consumption, and smoking were included to investigate health behavior deterioration during the COVID-19 pandemic. Possible responses were “increased,” “same,” “decreased,” or “not applicable.” With reference to a previous study [37], health behavior deterioration was measured as follows: decreased physical activity, changes in sleep duration (either increased or decreased), increased consumption of instant meals/soda, increased consumption of delivery food, increased alcohol consumption, and an increase in the number of cigarettes smoked per day.

Data Analysis

The distribution of men and women according to depression level was calculated after weighting based on the KCHS guidelines. The weighted values of the distribution of men and women according to depression level are shown as descriptive statistics such as frequency, percentage, and standard error. The participants’ general characteristics according to depression level and gender were also analyzed using descriptive statistics such as frequency, percentage, and standard error. We performed a χ^2 test to confirm the general distribution of the characteristics according to the depression level of the participants.

Depression levels, gender-specific precautionary behavior noncompliance, COVID-19 fear, and health behavior deterioration were analyzed using logistic regression. After controlling for general characteristics (age, educational level, monthly household income, profession, region, heavy drinking, current smoking status, stress awareness, and subjective health status), logistic regression analysis was used to identify differences in depression level and gender-specific precautionary behavior noncompliance, COVID-19 fear, and health behavior deterioration. The adjusted odds ratios (AORs) and 95% CIs for failure to comply with precautionary behaviors, COVID-19–related fear, and health behavior deterioration were calculated. All results are presented as weighted values. The results were statistically analyzed using SPSS version 26.0 (IBM Corp, Armonk, NY, USA).

Results

Table 1 shows the weighted and unweighted distributions of depression in the population according to gender. This study included 204,787 participants (92,739 men and 112,048 women).

Approximately 2% and 3.5% of men and women were classified as having clinically relevant depression, respectively.

Table 2 shows the participants' general characteristics, and Table 3 and Table 4 further summarize these data by depression levels for men and women, respectively. To better represent the general population of South Korea, the percentages and standard errors are presented as weighted values. The largest group of men with clinically relevant depression (48.8%) were aged 20-44 years, and 44.6% of men with clinically relevant depression had a college degree or higher. Approximately 41.6% belonged to the income level Q1, and the rate of unemployment reached 45.5%. Although there was no significant difference in the level of mental health by region, 45.6% of the men with clinically relevant depression reported heavy drinking and 44.5% were current smokers. Among the men with clinically relevant depression, only 21.3% answered that they had stress, but 79.3% evaluated their subjective health condition as poor. Among the women with clinically relevant depression, 60.3% were unemployed. Notably, self-reported heavy drinking increased to 19.4% and current smoking increased to 11.7%, which was five times higher than that of women without depression symptoms

Table 1. Overall and weighted distribution of depression in the sample population (N=204,787).

Depression level[a]	Unweighted, n (%)	Weighted[b]	
		Participants, n (%)	SE
Men (n=92,739)			
None (0-4)	83,026 (89.5)	16,832,027 (88.4)	0.1
Mild (5-9)	7889 (8.5)	1,799,468 (9.5)	0.1
Clinically relevant (≥10)	1824 (2.0)	399,490 (2.1)	0.1
Women (n=112,048)			
None (0-4)	93,305 (83.3)	15,943,821 (82.3)	0.2
Mild (5-9)	14,771 (13.2)	2,683,383 (13.9)	0.1
Clinically relevant (≥10)	3972 (3.5)	736,847 (3.8)	0.1

[a] Classified according to Patient Health Questionnaire-9 scores.

[b] Data were weighted to yield nationally representative estimates (total N=38,395,036).

Table 2. Participants' general characteristics (N=204,787).

Characteristics	Unweighted, n	Weighted[a]	
		Participants, n	% (SE)
Age (years)			
20-44	60,924	16,207,118	42.2 (0.1)
45-64	77,873	14,448,396	37.6 (0.1)
65 and over	65,990	7,739,521	20.2 (0.1)
Educational level			
Elementary or less	45,804	4,453,550	11.6 (0.1)
Middle school	23,051	3,089,900	8.1 (0.1)
High school	58,976	11,267,036	29.3 (0.1)
College or higher	76,956	19,584,549	51.0 (0.2)
Monthly household income quartile (Q)			
Q1 (lowest)	72,643	9,499,542	24.7 (0.1)
Q2	51,997	9,871,080	25.7 (0.1)
Q3	41,606	9,416,797	24.5 (0.1)
Q4 (highest)	38,541	9,607,616	25.0 (0.2)
Profession[b]			
White	38,617	9,999,414	26.0 (0.1)
Pink	25,798	5,195,447	13.5 (0.1)
Blue	59,749	8,538,958	22.2 (0.1)
None	80,623	14,661,216	38.2 (0.1)
Region			
Urban	58,299	16,465,931	42.9 (0.1)
Rural	146,488	21,929,104	57.1 (0.1)
Heavy drinking			
Never	86,787	13,072,593	34.0 (0.1)
No	67,789	13,778,747	35.9 (0.1)
Yes	50,211	11,543,695	30.1 (0.1)
Current smoking status			
Never	134,123	24,290,719	63.3 (0.1)
Former	37,613	7,183,170	18.7 (0.1)
Current	33,051	6,921,146	18.0 (0.1)
Stress awareness			
No	45,478	9,869,519	25.7 (0.1)
Yes	159,309	28,525,516	74.3 (0.1)
Subjective health status			
Poor	107,046	18,202,506	47.4 (0.1)
Fair	97,741	20,192,529	52.6 (0.1)

[a] Data were weighted to yield nationally representative estimates (total N=38,395,036).

[b] Classified into white (white collar), pink (sales and service), blue (agricultural, forestry, fishing, and military), and none (unemployed).

Table 3. Comparison of men's general characteristics by depression level (N=92,739).

Characteristics	No depression (PHQ-9[a] 0-4)		Mild depression (PHQ-9 5-9)		Clinically relevant depression (PHQ-9≥10)		*F* [b]	*df*	*P* value
	Unweighted[c], n	Weighted[d], n (%, SE)	Unweighted, n	Weighted, n (%, SE)	Unweighted, n	Weighted, n (%, SE)			
Age (years)							26.63	3.88, 61,102.47	<.001
20-44	25,820	7,320,210 (43.5, 0.2)	2762	875,763 (48.7, 0.7)	614	194,869 (48.8, 1.5)			
45-64	32,700	6,460,358 (38.4, 0.2)	2639	595,349 (33.1, 0.7)	561	116,955 (29.3, 1.3)			
65 and over	24,506	3,051,459 (18.1, 0.2)	2488	328,356 (18.2, 0.5)	649	87,666 (21.9, 1.1)			
Educational level							27.64	5.73, 90,089.21	<.001
Elementary or less	10,998	1,075,442 (6.4, 0.1)	1340	144,948 (8.1, 0.3)	418	51,493 (12.9, 0.8)			
Middle school	9269	1,194,121 (7.1, 0.1)	930	138,617 (7.7, 0.3)	250	40,617 (10.2, 0.8)			
High school	26,371	5,020,817 (29.8, 0.2)	2347	549,575 (30.5, 0.6)	555	129,132 (32.3, 1.4)			
College or higher	36,388	9,541,647 (56.7, 0.2)	3272	966,328 (53.7, 0.7)	601	178,248 (44.6, 1.5)			
Monthly household income quartile (Q)							62.22	5.93, 93,308.82	<.001
Q1 (lowest)	25,438	3,589,381 (21.3, 0.2)	3115	505,037 (28.1, 0.6)	949	166,331 (41.6, 1.4)			
Q2	21,809	4,321,547 (25.7, 0.2)	1864	459,319 (25.5, 0.6)	391	94,942 (23.8, 1.4)			
Q3	18,316	4,347,118 (25.8, 0.2)	1534	422,973 (23.5, 0.6)	257	72,562 (18.2, 1.2)			
Q4 (highest)	17,463	4,573,982 (27.2, 0.2)	1376	412,139 (22.9, 0.6)	227	65,655 (16.4, 1.1)			
Profession[e]							42.96	5.94, 93,506.52	<.001
White	17,865	4,938,141 (29.3, 0.2)	1533	470,702 (26.2, 0.7)	251	84,165 (21.1, 1.3)			
Pink	8371	2,013,564 (12.0, 0.2)	819	229,551 (12.8, 0.5)	165	45,913 (11.5, 1.0)			
Blue	34,082	5,513,592 (32.8, 0.2)	2737	558,664 (31.0, 0.7)	464	87,410 (21.9, 1.2)			
None	22,708	4,366,730 (25.9, 0.2)	2800	540,551 (30.0, 0.6)	944	182,002 (45.5, 1.5)			
Region							1.05	1.99, 31,436.14	.35
Urban	23,214	7,100,130 (42.2, 0.2)	2438	778,683 (43.3, 0.7)	559	169,067 (42.3, 1.5)			
Rural	59,812	9,731,897 (57.8, 0.2)	5451	1,020,785 (56.7, 0.7)	1265	230,423 (57.7, 1.5)			
Heavy drinking							16.41	3.98, 62,618.59	<.001
Never	24,042	3,914,812 (23.3, 0.2)	2416	418,786 (23.3, 0.6)	725	127,952 (32.0, 1.3)			

Characteristics	No depression (PHQ-9[a] 0-4)		Mild depression (PHQ-9 5-9)		Clinically relevant depression (PHQ-9≥10)		F [b]	*df*	*P* value
	Unweighted[c], n	Weighted[d], n (%, SE)	Unweighted, n	Weighted, n (%, SE)	Unweighted, n	Weighted, n (%, SE)			
No	24,751	5,072,726 (30.1, 0.2)	2202	514,892 (28.6, 0.7)	405	89,483 (22.4, 1.2)			
Yes	34,233	7,844,489 (46.6, 0.2)	3271	865,790 (48.1, 0.7)	694	182,055 (45.6, 1.5)			
Smoking status							53.93	3.98, 62,603.76	<.001
Never	25,752	5,621,708 (33.4, 0.2)	1914	474,901 (26.4, 0.6)	423	93,702 (23.4, 1.5)			
Former	30,840	5,750,630 (34.2, 0.2)	3004	613,040 (34.1, 0.7)	654	128,119 (32.1, 1.3)			
Current	26,434	5,459,689 (32.4, 0.2)	2971	711,527 (39.5, 0.7)	747	177,669 (44.5, 1.4)			
Stress awareness							2598.88	1.99, 31,378.61	<.001
No	14,374	3,431,190 (20.4, 0.2)	3997	1,021,365 (56.8, 0.7)	1342	314,591 (78.7, 1.2)			
Yes	68,652	13,400,837 (79.6, 0.2)	3892	778,103 (43.2, 0.7)	482	84,899 (21.3, 1.2)			
Subjective health status							867.4	2.00, 31,446.08	<.001
Poor	36,364	6,707,215 (39.8, 0.2)	5450	1,150,997 (64.0, 0.7)	1510	316,919 (79.3, 1.2)			
Fair	46,662	10,124,812 (60.2, 0.2)	2439	648,471 (36.0, 0.7)	314	82,571 (20.7, 1.2)			

[a] PHQ-9: Patient Health Questionnaire-9.

[b] Rao-Scott composite sample χ^2 test.

[c] Data were not weighted (total N= 204,787).

[d] Data were weighted to yield nationally representative estimates (total N=38,395,036).

[e] Classified into white (white collar), pink (sales and service), blue (agricultural, forestry, fishing, and military), and none (unemployed).

Table 4. Comparison of women's general characteristics by depression level (N=112,048).

Characteristics	No depression (PHQ-9[a] 0-4)		Mild depression (PHQ-9 5-9)		Clinically relevant depression (PHQ-9≥10)		*F* [b]	*df*	*P* value
	Unweighted[c], n	Weighted[d], n (%, SE)	Unweighted, n	Weighted, n (%, SE)	Unweighted, n	Weighted, n (%, SE)			
Age (years)							64.38	3.91, 62,134.61	<.001
20-44	26,185	6,354,538 (39.8, 0.2)	4323	1,137,258 (42.4, 0.5)	1220	324,480 (44.0, 1.0)			
45-64	36,214	6,187,323 (38.8, 0.2)	4674	880,246 (32.8, 0.5)	1085	208,165 (28.3, 0.9)			
65 and over	30,906	3,401,959 (21.3, 0.2)	5774	665,878 (24.8, 0.4)	1667	204,202 (27.7, 0.8)			
Educational level							41.03	5.81, 92,331.79	<.001
Elementary or less	26,408	2,480,195 (15.6, 0.1)	5111	528,808 (19.7, 0.4)	1529	172,665 (23.4, 0.7)			
Middle school	10,646	1,411,759 (8.9, 0.1)	1563	234,528 (8.7, 0.3)	393	70,257 (9.6, 0.6)			
High school	25,060	4,590,105 (28.8, 0.2)	3635	759,285 (28.3, 0.5)	1008	218,122 (29.6, 0.9)			
College or higher	31,191	7,461,761 (46.8, 0.2)	4462	1,160,761 (43.3, 0.5)	1042	275,803 (37.4, 1.0)			
Monthly household income quartile (Q)							95.29	5.94, 94,443.28	<.001
Q1 (lowest)	34,270	4,045,405 (25.4, 0.2)	6709	876,383 (32.7, 0.5)	2162	317,006 (43.0, 1.0)			
Q2	23,614	4,117,698 (25.8, 0.2)	3483	697,517 (26.0, 0.5)	836	180,058 (24.4, 0.9)			
Q3	18,471	3,866,590 (24.3, 0.2)	2481	579,397 (21.6, 0.5)	547	128,155 (17.4, 0.8)			
Q4 (highest)	16,950	3,914,127 (24.5, 0.2)	2098	530,085 (19.8, 0.5)	427	111,628 (15.1, 0.8)			
Profession[e]							36.44	5.90, 93,721.40	<.001
White	16,222	3,802,298 (23.8, 0.2)	2277	577,758 (21.5, 0.5)	469	126,349 (17.1, 0.8)			
Pink	14,037	2,420,483 (15.2, 0.2)	1926	380,193 (14.2, 0.4)	480	105,743 (14.3, 0.7)			
Blue	19,458	2,021,899 (12.7, 0.1)	2520	296,602 (11.1, 0.3)	488	60,792 (8.3, 0.5)			
None	43,588	7,699,140 (48.3, 0.2)	8048	1,428,829 (53.2, 0.5)	2535	443,963 (60.3, 1.0)			
Region							0.049	1.99, 31,747.28	.95
Urban	26,358	6,928,218 (43.5, 0.2)	4522	1,167,254 (43.5, 0.5)	1208	322,577 (43.8, 1.0)			
Rural	66,947	9,015,602 (56.5, 0.2)	10,249	1,516,128 (56.5, 0.5)	2764	414,270 (56.2, 1.0)			
Heavy drinking							56.69	3.98, 63,305.92	<.001
Never	49,462	7,109,633 (44.6, 0.2)	7909	117,411 (43.8, 0.5)	2233	326,999 (44.4, 1.0)			

Characteristics	No depression (PHQ-9[a] 0-4)		Mild depression (PHQ-9 5-9)		Clinically relevant depression (PHQ-9≥10)		*F* [b]	*df*	*P* value
	Unweighted[c], n	Weighted[d], n (%, SE)	Unweighted, n	Weighted, n (%, SE)	Unweighted, n	Weighted, n (%, SE)			
No	34,385	6,794.290 (42.6, 0.2)	4900	1,040,293 (38.8, 0.5)	1146	267,063 (36.2, 1.0)			
Yes	9458	2,039,897 (12.8, 0.2)	1962	468,678 (17.4, 0.4)	593	142,785 (19.4, 0.8)			
Smoking status							350.97	3.97, 63,128.55	<.001
Never	89,302	1,513,9982 (95.0, 0.1)	13,416	2,371,959 (88.4, 0.3)	3316	588,466 (79.9, 0.8)			
Former	2120	455,587 (2.8, 0.1)	714	173,855 (6.5, 0.3)	281	61,940 (8.4, 0.6)			
Current	1883	348,251 (2.2, 0.1)	641	137,568 (5.1, 0.2)	375	86,441 (11.7, 0.6)			
Stress awareness							3865.37	1.99, 31,722.31	<.001
No	15,760	3,131,642 (19.6, 0.2)	7157	1,413,307 (52.7, 0.5)	2848	557,423 (75.6, 0.8)			
Yes	77,545	1,281,2178 (80.4, 0.2)	7614	1,270,075 (47.3, 0.5)	1124	179,424 (24.4, 0.8)			
Subjective health status							1144.07	1.99, 31,746.02	<.001
Poor	49,326	7,556,843 (47.4, 0.2)	11,041	1,874,240 (69.8, 0.5)	3355	596,292 (80.9, 0.8)			
Fair	43,979	8,386,977 (52.6, 0.2)	3730	809,142 (30.2, 0.5)	617	140,555 (19.1, 0.8)			

[a] PHQ-9: Patient Health Questionnaire-9.

[b] Rao-Scott composite sample χ^2 test.

[c] Data were not weighted (total N= 204,787).

[d] Data were weighted to yield nationally representative estimates (total N=38,395,036).

[e] Classified into white (white collar), pink (sales and service), blue (agricultural, forestry, fishing, and military), and none (unemployed).

Table 5 illustrates how compliance with precautionary behaviors, COVID-19 fear, and health behavior deterioration differed according to gender and mental health determined using logistic regression analysis after adjusting for general characteristics. The crude odds ratios from the unadjusted model are shown in Multimedia Appendix 1. The precautionary behavior noncompliance rate was found to increase as the depression worsened; this was the case for most items compared to the group without depressive symptoms. The AOR of the noncompliance rate regarding wearing a mask indoors was lower in participants with mild depressive symptoms compared to that for participants with clinically relevant depression, which was the case for both men and women (Table 5). However, in the case of noncompliance with precautionary behaviors related to social distancing, the trend was different between men and women. Among men who had mild depressive symptoms, nonobservance of maintaining physical distance increased by more than that among men with clinically relevant depression. By contrast, the AOR of not maintaining physical distance in women was higher for those with clinically relevant depression than for those with mild depression, showing a tendency to worsen as mental health deteriorated.

The data also showed different characteristics of COVID-19 fear between men and women. The AOR of the fear of infection was similar for women and men with clinically relevant depression. However, the fear of death due to infection was significant only in women for those with mild depressive symptoms. In the case of public criticism, a significant association appeared only among women with clinically relevant depression.

As depression increased in severity in both men and women, health behaviors likewise deteriorated. Changes in sleep duration, an increase in drinking amount, and smoking frequency more than doubled for both men and women with clinically relevant depression. Regarding delivery food consumption, the AOR was the highest for men with mild depressive symptoms, which was similar to that for women with clinically relevant depression.

Table 5. Adjusted odds ratios (95% CIs) for failure to comply with precautionary behaviors, health behavior deterioration during the COVID-19 outbreak, and COVID-19–related fear according to depression levels.[a,b,c]

COVID-19–related questions	Men (n=92,739)			Women (n=112,048)		
	No depression[d] (0-4)	Mild depression (5-9)	Clinically relevant depression (≥10)	No depression[d] (0–4)	Mild depression (5–9)	Clinically relevant depression (≥10)
Failure to comply with precautionary behaviors						
Not covering mouth while coughing	1	*1.29 (1.13-1.46)* [e]	*1.56 (1.26-1.93)*	1	*1.38 (1.24-1.55)*	*1.65 (1.40-1.95)*
No proper ventilation	1	*1.54 (1.29-1.83)*	*1.73 (1.28-2.34)*	1	*1.59 (1.34-1.90)*	*2.22 (1.75-2.81)*
Not performing regular disinfection	1	*1.23 (1.16-1.31)*	*1.32 (1.16-1.50)*	1	*1.19 (1.13-1.25)*	*1.24 (1.13-1.36)*
Not wearing a mask indoors	1	*1.54 (1.04-2.27)*	*1.81 (1.07-3.06)*	1	*1.49 (1.02-2.18)*	*2.04 (1.21-3.44)*
Not wearing a mask when it was hard to maintain distance	1	*1.44 (1.12-1.85)*	1.42 (0.90-2.23)	1	*1.57 (1.21-2.02)*	*1.93 (1.35-2.77)*
Not keeping the minimum recommended physical distance	1	*1.41 (1.23-1.61)*	*1.34 (1.03-1.74)*	1	*1.49 (1.33-1.67)*	*1.57 (1.29-1.90)*
Not refraining from visiting hospitalized patients	1	*1.82 (1.41-2.35)*	*1.89 (1.14-3.13)*	1	*1.29 (1.02-1.62)*	1.35 (0.93-1.94)
Not refraining from going out	1	*1.55 (1.31-1.85)*	*1.51 (1.13-2.00)*	1	*1.56 (1.35-1.81)*	*1.36 (1.06-1.75)*
COVID-19–related fears						
Fear of infection	1	0.96 (0.87-1.05)	*0.72 (0.61-0.86)*	1	1.05 (0.95-1.15)	*0.74 (0.63-0.86)*
Fear of dying from infection	1	1.05 (0.98-1.12)	1.07 (0.94-1.22)	1	*1.09 (1.03-1.15)*	1.04 (0.94-1.14)
Fear of public criticism	1	0.95 (0.88-1.04)	0.85 (0.72-1.00)	1	1.07 (0.98-1.16)	*0.86 (0.75-0.99)*
Fear of a family member getting infected	1	1.06 (0.92-1.22)	0.83 (0.66-1.04)	1	1.12 (0.99-1.26)	0.94 (0.76-1.15)
Fear of economic loss due to infection	1	1.11 (1.00-1.23)	0.97 (0.79-1.19)	1	1.04 (0.95-1.13)	0.94 (0.80-1.10)
Health behavior deterioration						
Decreased physical activity	1	*1.17 (1.10-1.25)*	*1.4 (1.24-1.58)*	1	*1.28 (1.22-1.35)*	*1.31 (1.21-1.43)*
Changes in sleep duration	1	*1.62 (1.51-1.74)*	*2.28 (2.00-2.59)*	1	*1.71 (1.62-1.80)*	*2.15 (1.96-2.36)*
Increased consumption of instant meals/soda	1	*1.47 (1.35-1.60)*	*1.64 (1.40-1.93)*	1	*1.36 (1.27-1.45)*	*1.6 (1.43-1.80)*
Increased consumption of delivery food	1	*1.28 (1.19-1.39)*	*1.25 (1.07-1.45)*	1	*1.23 (1.15-1.31)*	*1.26 (1.12-1.41)*
Increased alcohol drinking	1	*1.51 (1.33-1.71)*	*1.95 (1.58-2.42)*	1	*1.53 (1.36-1.73)*	*2.25 (1.88-2.70)*
Increased smoking frequency	1	*1.64 (1.44-1.87)*	*2.55 (2.07-3.15)*	1	*1.53 (1.13-2.08)*	*2.71 (1.95-3.77)*

[a] Depression level was classified according to scores on the nine-item Patient Health Questionnaire.

[b] Logistic regression model adjusted for age, educational level, monthly household income, profession, region, heavy drinking, current smoking status, stress awareness, and subjective health status.

[c] Data were weighted to yield nationally representative estimates (total N=38,395,036).

[d] Reference category.

[e] Values in italics indicate a statistically significant association.

Discussion

Principal Findings

The purpose of this study was to determine the effect of an individual's depression level on noncompliance precautionary behavior, COVID-19 fear, and health behavior deterioration, and to compare the different trends according to gender. Overall, there was a significant difference in noncompliance with precautionary behavior and health behavior deterioration according to the degree of mental health. Additionally, some components pertaining to COVID-19 fear were different according to depression status. Men and women with clinically relevant depression were more likely to not comply with precautionary behaviors. Furthermore, as mental health deteriorated, health behaviors declined, which tended to be worse in women.

In the case of men with clinically relevant depression, not performing regular ventilation and not wearing a mask indoors significantly increased. Similarly, women with clinically relevant depression were more than twice as likely to not ventilate regularly or not wear a mask indoors than were women who were not depressed. These results are similar to those of several previous studies [23,38,39]; however, in some important ways, they conflict with the results of other studies. For example, one study indicated that poor mental health increased preventive activities, while another suggested that there was no association between precautionary behaviors and mental distress in those with severe mental illness [40,41]. This dichotomy may also exist in relation to COVID-19 fear. Generally, feelings of distress such as anxiety and fear are known to increase precautionary behavior [23,42,43]. However, in our study, the participants tended to have decreased levels of fear as their mental health deteriorated. This association between fear and precautionary behaviors suggests that reduced fear may increase precautionary behavior noncompliance.

Generally, health behaviors decrease as mental health worsens. Notably, women with clinically relevant depression showed a dramatic increase in heavy alcohol consumption and smoking frequency. Previous studies have also reported that women's drinking behaviors increased during the COVID-19 pandemic [44] and that depressive symptoms were also associated with an increase in heavy drinking [45].

However, an increase in social responsibility and a decrease in social support could worsen the mental health of women. In the context of the COVID-19 pandemic, women have been facing a variety of high-demand roles, including childcare and elder care, which most countries have traditionally imposed on them [17,46-48]. In addition, it is generally well known that social support has more of an effect on mental health for women than for men [49-51]. Restrictions on social gatherings, such as social distancing and guidelines that formed part of various quarantine measures, might have weakened the vital social support for women. In fact, various studies have provided empirical evidence that social distancing can affect mental health [52-54]. Moreover, research suggests that reduced physical activity due to social distancing may impair mental health [55]. Given that mental health is associated with smoking and drinking [56], changes in increased social responsibility and reduced social support can also worsen women's mental health and lead to dramatic increases in drinking and smoking.

Overall, the trend of noncompliance with precautionary behaviors and a decrease in health behaviors according to worsening depression showed a greater change in women than in men. Although precautionary behavior compliance is the most important factor in responding to pandemic-related situations [57], health behaviors constitute the basic foundation for maintaining one's physical and mental health, not only in the present but also in the future. Therefore, interventions that consider the degree of depression are necessary to alleviate depression, which could affect the implementation of precautionary behavior and health behavior. Additionally, further studies should examine different ways of implementing such interventions depending on gender.

Study Limitations

This study had several limitations. First, the study did not consider time changes; it was difficult to ascertain causal relationships between the effects of depression and precautionary and health behaviors. In particular, since the KCHS data did not confirm the antecedent relationship between depression and COVID-19, it could not be confirmed whether the depression was due to COVID-19 or other factors. Second, in this study, the trends of precautionary behavior compliance, COVID-19 fear, and health behaviors according to the degree of depression were identified by gender, but it was not shown how vulnerable women were compared to men and their statistical significance. Third, because the data used in this research were collected in the early stages of the COVID-19 pandemic, the precautionary behavior compliance rate created a fairly high temporary bias. In other words, if the data were collected at the end of the pandemic, the results of precautionary or health behaviors may have changed. Finally, in light of the fact that a self-reporting survey was conducted, there is a possibility that measurement errors may have occurred because of the inherent difficulty in checking whether COVID-19 precautionary behaviors were actually being followed and whether health behavior had actually changed. However, this study was cross-sectional and used a representative sample from the entire Republic of Korea. Additionally, this study is valuable in that it examined how depressive symptoms and gender affect precautionary behaviors, even when social aspirations and interest in precautionary behaviors are high.

Conclusion

In this study, we were able to identify how precautionary behavior compliance, COVID-19 fear, and health behaviors change according to the degree of depression and trends according to gender. We found that regardless of gender, people suffering from clinically relevant depression were highly likely to infringe on precautionary behaviors and deteriorate health behaviors. Additionally, this trend was more noticeable in women. These results can be useful for developing related interventions and for future studies that consider both mental status and gender.

Acknowledgments

This research was supported by the Chung-Ang University Graduate Research Scholarship in 2021.

Data Availability

The data set used in this study and more information about the data can be obtained from the official Korea Community Health Survey website [29].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Crude odds ratios (95% CIs) for failure to comply with precautionary behaviors, health behavior deterioration during the COVID-19 outbreak, and COVID-19–related fear according to depression levels.
[DOCX File , 17 KB - publichealth_v9i1e42677_app1.docx]

References

1. WHO Coronavirus (COVID-19) Dashboard. 2022 Jan 25. URL: https://covid19.who.int/ [accessed 2022-01-25]
2. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health 2020 May;5(5):e261-e270 [FREE Full text] [doi: 10.1016/S2468-2667(20)30073-6] [Medline: 32220655]
3. Thompson RN. Epidemiological models are important tools for guiding COVID-19 interventions. BMC Med 2020 May 25;18(1):152 [FREE Full text] [doi: 10.1186/s12916-020-01628-4] [Medline: 32448247]
4. Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet 2020 Apr 25;395(10233):1382-1393 [FREE Full text] [doi: 10.1016/S0140-6736(20)30746-7] [Medline: 32277878]
5. Shanafelt T, Ripp J, Trockel M. Understanding and addressing sources of anxiety among health care professionals during the COVID-19 pandemic. JAMA 2020 Jun 02;323(21):2133-2134. [doi: 10.1001/jama.2020.5893] [Medline: 32259193]
6. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Global Health 2020 Jul 06;16(1):57 [FREE Full text] [doi: 10.1186/s12992-020-00589-w] [Medline: 32631403]
7. Fancourt D, Steptoe A, Bu F. Trajectories of anxiety and depressive symptoms during enforced isolation due to COVID-19 in England: a longitudinal observational study. Lancet Psychiatry 2021 Feb;8(2):141-149 [FREE Full text] [doi: 10.1016/S2215-0366(20)30482-X] [Medline: 33308420]
8. O'Connor RC, Wetherall K, Cleare S, McClelland H, Melson AJ, Niedzwiedz CL, et al. Br J Psychiatry 2021 Jun;218(6):326-333 [FREE Full text] [doi: 10.1192/bjp.2020.212] [Medline: 33081860]
9. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021 Nov 06;398(10312):1700-1712 [FREE Full text] [doi: 10.1016/S0140-6736(21)02143-7] [Medline: 34634250]
10. Voitsidis P, Nikopoulou VA, Holeva V, Parlapani E, Sereslis K, Tsipropoulou V, et al. The mediating role of fear of COVID-19 in the relationship between intolerance of uncertainty and depression. Psychol Psychother 2021 Sep;94(3):884-893 [FREE Full text] [doi: 10.1111/papt.12315] [Medline: 33216444]
11. Saravanan C, Mahmoud I, Elshami W, Taha MH. Knowledge, anxiety, fear, and psychological distress about COVID-19 among university students in the United Arab Emirates. Front Psychiatry 2020;11:582189 [FREE Full text] [doi: 10.3389/fpsyt.2020.582189] [Medline: 33192728]
12. Lin C, Broström A, Griffiths MD, Pakpour AH. Investigating mediated effects of fear of COVID-19 and COVID-19 misunderstanding in the association between problematic social media use, psychological distress, and insomnia. Internet Interv 2020 Sep;21:100345 [FREE Full text] [doi: 10.1016/j.invent.2020.100345] [Medline: 32868992]
13. Taylor S, Asmundson GJG. Negative attitudes about facemasks during the COVID-19 pandemic: the dual importance of perceived ineffectiveness and psychological reactance. PLoS One 2021;16(2):e0246317 [FREE Full text] [doi: 10.1371/journal.pone.0246317] [Medline: 33596207]
14. Wang S, Quan L, Chavarro JE, Slopen N, Kubzansky LD, Koenen KC, et al. Associations of depression, anxiety, worry, perceived stress, and loneliness prior to infection with risk of post-COVID-19 conditions. JAMA Psychiatry 2022 Nov 01;79(11):1081-1091 [FREE Full text] [doi: 10.1001/jamapsychiatry.2022.2640] [Medline: 36069885]
15. Kim HR, Kim J. Stress, depression, and unhealthy behavior changes among patients with diabetes during COVID-19 in Korea. Healthcare 2022 Feb 04;10(2):303 [FREE Full text] [doi: 10.3390/healthcare10020303] [Medline: 35206917]
16. Brotto LA, Chankasingh K, Baaske A, Albert A, Booth A, Kaida A, et al. The influence of sex, gender, age, and ethnicity on psychosocial factors and substance use throughout phases of the COVID-19 pandemic. PLoS One 2021;16(11):e0259676 [FREE Full text] [doi: 10.1371/journal.pone.0259676] [Medline: 34807908]
17. Hammarberg K, Tran T, Kirkman M, Fisher J. Sex and age differences in clinically significant symptoms of depression and anxiety among people in Australia in the first month of COVID-19 restrictions: a national survey. BMJ Open 2020 Nov 24;10(11):e042696 [FREE Full text] [doi: 10.1136/bmjopen-2020-042696] [Medline: 33234659]
18. Zhang W, Walkover M, Wu YY. The challenge of COVID-19 for adult men and women in the United States: disparities of psychological distress by gender and age. Public Health 2021 Sep;198:218-222 [FREE Full text] [doi: 10.1016/j.puhe.2021.07.017] [Medline: 34481278]

19. Parker G, Brotchie H. Gender differences in depression. Int Rev Psychiatry 2010;22(5):429-436. [doi: 10.3109/09540261.2010.492391] [Medline: 21047157]
20. Kiely KM, Brady B, Byles J. Gender, mental health and ageing. Maturitas 2019 Nov;129:76-84. [doi: 10.1016/j.maturitas.2019.09.004] [Medline: 31547918]
21. Asmundson GJG, Taylor S. How health anxiety influences responses to viral outbreaks like COVID-19: What all decision-makers, health authorities, and health care professionals need to know. J Anxiety Disord 2020 Apr;71:102211 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102211] [Medline: 32179380]
22. Stickley A, Matsubayashi T, Sueki H, Ueda M. COVID-19 preventive behaviours among people with anxiety and depressive symptoms: findings from Japan. Public Health 2020 Dec;189:91-93 [FREE Full text] [doi: 10.1016/j.puhe.2020.09.017] [Medline: 33189941]
23. Olapegba PO, Chovwen CO, Ayandele O, Ramos-Vera C. Fear of COVID-19 and preventive health behavior: mediating role of post-traumatic stress symptomology and psychological distress. Int J Ment Health Addict 2022;20(5):2922-2933 [FREE Full text] [doi: 10.1007/s11469-021-00557-4] [Medline: 34121960]
24. Ahorsu DK, Imani V, Lin C, Timpka T, Broström A, Updegraff JA, et al. Associations between fear of COVID-19, mental health, and preventive behaviours across pregnant women and husbands: an actor-partner interdependence modelling. Int J Ment Health Addict 2022 Jun 11;20(1):68-82 [FREE Full text] [doi: 10.1007/s11469-020-00340-x] [Medline: 32837427]
25. Prowse R, Sherratt F, Abizaid A, Gabrys RL, Hellemans KGC, Patterson ZR, et al. Coping with the COVID-19 pandemic: examining gender differences in stress and mental health among university students. Front Psychiatry 2021;12:650759 [FREE Full text] [doi: 10.3389/fpsyt.2021.650759] [Medline: 33897499]
26. Simha A, Prasad R, Ahmed S, Rao NP. Effect of gender and clinical-financial vulnerability on mental distress due to COVID-19. Arch Womens Ment Health 2020 Dec;23(6):775-777 [FREE Full text] [doi: 10.1007/s00737-020-01097-x] [Medline: 33420600]
27. Xue B, McMunn A. Gender differences in unpaid care work and psychological distress in the UK Covid-19 lockdown. PLoS One 2021;16(3):e0247959 [FREE Full text] [doi: 10.1371/journal.pone.0247959] [Medline: 33662014]
28. Zhang W, Yang X, Zhao J, Yang F, Jia Y, Cui C, et al. Depression and psychological-behavioral responses among the general public in China during the early stages of the COVID-19 pandemic: survey study. J Med Internet Res 2020 Sep 04;22(9):e22227 [FREE Full text] [doi: 10.2196/22227] [Medline: 32886066]
29. Korea Community Health at a Glance 2020: Korea Community Health Survey (KCHS). Korea Disease Control and Prevention Agency. 2020. URL: https://chs.kdca.go.kr/chs/rawDta/rawDtaProvdMain.do#2 [accessed 2022-01-25]
30. Thomas SL, Heck RH. Analysis of large-scale secondary data in higher education research: potential perils associated with complex sampling designs. Res High Educ 2001;42:517-540. [doi: 10.1023/A:1011098109834]
31. About KLoSA. KEIS. URL: https://survey.keis.or.kr/klosa/klosa01.jsp [accessed 2023-01-03]
32. Han MA, Kim HR. Smoking behavior changes during COVID-19 among Korean adults. Am J Health Behav 2021 Nov 15;45(6):1031-1040. [doi: 10.5993/AJHB.45.6.7] [Medline: 34969414]
33. Chung S, Lee E. Patterns of time use across the life span in Korea: a latent class analysis and age and gender differences. Soc Indic Res 2016 Oct 27;134(3):1135-1155. [doi: 10.1007/s11205-016-1463-6]
34. Korean Standard classification of occupations. Korean Standard Statistical Classification Portal. 2017. URL: http://kssc.kostat.go.kr/ksscNew_web/ekssc/common/selectIntroduce.do?part=2&top_menu=101&bbsId=isco_s&categoryNameCode=801&categoryMenu=001 [accessed 2023-01-03]
35. Han S, Ko Y, Moon JE, Cho YS. Working hours are closely associated with depressive mood and suicidal ideation in Korean adults: a nationwide cross-sectional study. Sci Rep 2021 Nov 29;11(1):23102. [doi: 10.1038/s41598-021-02574-8] [Medline: 34845294]
36. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
37. Lee GB, Jung SJ, Yiyi Y, Yang JW, Thang HM, Kim HC. Socioeconomic inequality in compliance with precautions and health behavior changes during the COVID-19 outbreak: an analysis of the Korean Community Health Survey 2020. Epidemiol Health 2022;44:e2022013 [FREE Full text] [doi: 10.4178/epih.e2022013] [Medline: 35008144]
38. Liang W, Duan Y, Shang B, Hu C, Baker JS, Lin Z, et al. Precautionary behavior and depression in older adults during the COVID-19 pandemic: an online cross-sectional study in Hubei, China. Int J Environ Res Public Health 2021 Feb 14;18(4):1853 [FREE Full text] [doi: 10.3390/ijerph18041853] [Medline: 33672885]
39. Imtiaz A, Khan NM, Hossain MA. COVID-19 in Bangladesh: measuring differences in individual precautionary behaviors among young adults. Z Gesundh Wiss 2022;30(6):1473-1484 [FREE Full text] [doi: 10.1007/s10389-020-01453-2] [Medline: 33432285]
40. Lee ATC, Cheng GWH, Lin C, Wong BHC, Lam LCW. Do people with mental health problems have lower adherence to precautionary measures in COVID-19 pandemic? A cross-sectional observational study in Hong Kong. BMJ Open 2021 Aug 12;11(8):e046658 [FREE Full text] [doi: 10.1136/bmjopen-2020-046658] [Medline: 34385242]
41. Pinkham A, Ackerman RA, Depp CA, Harvey PD, Moore RC. COVID-19-related psychological distress and engagement in preventative behaviors among individuals with severe mental illnesses. NPJ Schizophr 2021 Feb 03;7(1):7. [doi: 10.1038/s41537-021-00136-5] [Medline: 33536446]

42. Harper CA, Satchell LP, Fido D, Latzman RD. Functional fear predicts public health compliance in the COVID-19 pandemic. Int J Ment Health Addict 2021;19(5):1875-1888 [FREE Full text] [doi: 10.1007/s11469-020-00281-5] [Medline: 32346359]
43. Kuper-Smith B, Doppelhofer LM, Oganian Y, Rosenblau G, Korn CW. Risk perception and optimism during the early stages of the COVID-19 pandemic. R Soc Open Sci 2021 Nov;8(11):210904 [FREE Full text] [doi: 10.1098/rsos.210904] [Medline: 34804569]
44. Pollard MS, Tucker JS, Green HD. Changes in adult alcohol use and consequences during the COVID-19 pandemic in the US. JAMA Netw Open 2020 Sep 01;3(9):e2022942 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.22942] [Medline: 32990735]
45. Nesoff ED, Gutkind S, Sirota S, McKowen AL, Veldhuis CB. Mental health and economic stressors associated with high-risk drinking and increased alcohol consumption early in the COVID-19 pandemic in the United States. Prev Med 2021 Dec;153:106854 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106854] [Medline: 34695505]
46. McGinty EE, Presskreischer R, Han H, Barry CL. Psychological distress and loneliness reported by US adults in 2018 and April 2020. JAMA 2020 Jul 07;324(1):93-94 [FREE Full text] [doi: 10.1001/jama.2020.9740] [Medline: 32492088]
47. Holingue C, Badillo-Goicoechea E, Riehm KE, Veldhuis CB, Thrul J, Johnson RM, et al. Mental distress during the COVID-19 pandemic among US adults without a pre-existing mental health condition: Findings from American trend panel survey. Prev Med 2020 Oct;139:106231 [FREE Full text] [doi: 10.1016/j.ypmed.2020.106231] [Medline: 32758507]
48. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. Int J Environ Res Public Health 2020 Jun 07;17(11):4065 [FREE Full text] [doi: 10.3390/ijerph17114065] [Medline: 32517294]
49. Wiklund M, Malmgren-Olsson E, Ohman A, Bergström E, Fjellman-Wiklund A. Subjective health complaints in older adolescents are related to perceived stress, anxiety and gender - a cross-sectional school study in Northern Sweden. BMC Public Health 2012 Nov 16;12(1):993 [FREE Full text] [doi: 10.1186/1471-2458-12-993] [Medline: 23158724]
50. Van Droogenbroeck F, Spruyt B, Keppens G. Gender differences in mental health problems among adolescents and the role of social support: results from the Belgian health interview surveys 2008 and 2013. BMC Psychiatry 2018 Jan 10;18(1):6 [FREE Full text] [doi: 10.1186/s12888-018-1591-4] [Medline: 29320999]
51. Milner A, Krnjacki L, LaMontagne AD. Age and gender differences in the influence of social support on mental health: a longitudinal fixed-effects analysis using 13 annual waves of the HILDA cohort. Public Health 2016 Nov;140:172-178. [doi: 10.1016/j.puhe.2016.06.029] [Medline: 27527844]
52. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb 29;395(10225):676 [FREE Full text] [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
53. Marroquín B, Vine V, Morgan R. Mental health during the COVID-19 pandemic: effects of stay-at-home policies, social distancing behavior, and social resources. Psychiatry Res 2020 Nov;293:113419 [FREE Full text] [doi: 10.1016/j.psychres.2020.113419] [Medline: 32861098]
54. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
55. Jacob ST, Crozier I, Fischer WA, Hewlett A, Kraft CS, Vega MDL, et al. Ebola virus disease. Nat Rev Dis Primers 2020 Feb 20;6(1):13 [FREE Full text] [doi: 10.1038/s41572-020-0147-3] [Medline: 32080199]
56. Ferreira VR, Jardim TV, Sousa ALL, Rosa BMC, Jardim PCV. Smoking, alcohol consumption and mental health: Data from the Brazilian study of Cardiovascular Risks in Adolescents (ERICA). Addict Behav Rep 2019 Jun;9:100147 [FREE Full text] [doi: 10.1016/j.abrep.2018.100147] [Medline: 31193769]
57. Advice for the public: Coronavirus disease (COVID-19). World Health Organization. 2022. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public [accessed 2023-01-03]

Abbreviations

AOR: adjusted odds ratio
IRB: institutional review board
KCHS: Korea Community Health Survey
PHQ-9: Patient Health Questionnaire-9

Edited by A Mavragani, T Sanchez; submitted 13.09.22; peer-reviewed by F Adorni, H Tehrani; comments to author 20.12.22; revised version received 09.01.23; accepted 28.01.23; published 22.02.23.

Please cite as:
Shin H, Kim JS, Lee H
Association of Depression With Precautionary Behavior Compliance, COVID-19 Fear, and Health Behaviors in South Korea: National Cross-sectional Study
JMIR Public Health Surveill 2023;9:e42677
URL: https://publichealth.jmir.org/2023/1/e42677
doi:10.2196/42677
PMID:36716130

Original Paper

Projection of Premature Cancer Mortality in Hunan, China, Through 2030: Modeling Study

Wenqiong Wu[1*], MA; Jing Wang[2,3*], MPH; Xian-zhen Liao[4], MA; Kekui Xu[4], BA; Yanhua Zou[4], BA; Zhaohui Shi[4], BA; Yingyun Hu[4], MA; Haifan Xiao[4], PhD; Can Li[4], PhD; Shiyu Cao[4], MA; Shiyu Wang[4], BA; Jia Guo[4], BA; Zhicheng Luo[3], MPH; Mengjiao Liu[4], BA; Mengyao Xu[4], BA; Donghui Jin[5], MA; Mengshi Chen[3], PhD; Zhongxi Fu[5*], MA; Shipeng Yan[4*], PhD

[1]Department of Radiotherapy, Hunan Cancer Hospital, Changsha, China
[2]Chengdu Center for Disease Control and Prevention, Chengdu, China
[3]Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, China
[4]Department of Cancer Prevention and Control, Hunan Cancer Hospital, Changsha, China
[5]Hunan Provincial Center for Disease Control and Prevention, Changsha, China
[*]these authors contributed equally

Corresponding Author:
Shipeng Yan, PhD
Department of Cancer Prevention and Control
Hunan Cancer Hospital
283 Tongzipo Road
Changsha, 410013
China
Phone: 86 88651233
Email: yanshipeng@hnca.org.cn

Abstract

Background: The United Nations Sustainable Development Goals for 2030 include reducing premature mortality from noncommunicable diseases by one-third. Although previous modeling studies have predicted premature mortality from noncommunicable diseases, the predictions for cancer and its subcategories are less well understood in China.

Objective: The aim of this study was to project premature cancer mortality of 10 leading cancers in Hunan Province, China, based on various scenarios of risk factor control so as to establish the priority for future interventions.

Methods: We used data collected between 2009 and 2017 from the Hunan cancer registry annual report as empirical data for projections. The population-attributable fraction was used to disaggregate cancer deaths into parts attributable and unattributable to 10 risk factors: smoking, alcohol use, high BMI, diabetes, physical inactivity, low vegetable and fruit intake, high red meat intake, high salt intake, and high ambient fine particulate matter (PM2.5) levels. The unattributable deaths and the risk factors in the baseline scenario were projected using the proportional change model, assuming constant annual change rates through 2030. The comparative risk assessment theory was used in simulated scenarios to reflect how premature mortality would be affected if the targets for risk factor control were achieved by 2030.

Results: The cancer burden in Hunan significantly increased during 2009-2017. If current trends for each risk factor continued to 2030, the total premature deaths from cancers in 2030 would increase to 97,787 in Hunan Province, and the premature mortality (9.74%) would be 44.47% higher than that in 2013 (6.74%). In the combined scenario where all risk factor control targets were achieved, 14.41% of premature cancer mortality among those aged 30-70 years would be avoided compared with the business-as-usual scenario in 2030. Reductions in the prevalence of diabetes, high BMI, ambient PM2.5 levels, and insufficient fruit intake played relatively important roles in decreasing cancer premature mortality. However, the one-third reduction goal would not be achieved for most cancers except gastric cancer.

Conclusions: Existing targets on cancer-related risk factors may have important roles in cancer prevention and control. However, they are not sufficient to achieve the one-third reduction goal in premature cancer mortality in Hunan Province. More aggressive risk control targets should be adopted based on local conditions.

(JMIR Public Health Surveill 2023;9:e43967) doi:10.2196/43967

KEYWORDS
cancer; forecasting; premature mortality; risk factors; Sustainable Development Goals

Introduction

Cancer is a leading cause of death worldwide, accounting for nearly 10 million deaths in 2020 or nearly one in six deaths [1]. In China, cancer affected more than 4.56 million people and caused 3 million deaths in 2020, accounting for 30.2% of all deaths in the world [2]. Furthermore, cancer is a leading cause of disability-adjusted life years (DALYs), accounting for 17.64% of all DALYs in China in 2019, which was nearly double the global value of 9.88% [3]. The high mortality and disability related to cancer pose a major disease burden worldwide. Cancer not only directly affects the lives of those diagnosed but also has its sequelae impact the family, society, and economy. The socioeconomic burden of cancer has been reported to be substantial, which includes both the direct health care costs and the lost productivity through premature mortality of the working population [4,5].

Cancer is a disease commonly believed to be preventable. In a nationwide study on the risk factors of cancer across 31 provinces of mainland China, Chen et al [6] found that nearly half (45.2%) of all cancer deaths could be attributed to 23 potentially controllable risk factors, which were classified into five categories: behavior, diet, metabolism, environment, and infection. This suggests that millions of deaths could be prevented or delayed each year if appropriate strategies were developed to control these modifiable risk factors, which would undoubtedly greatly reduce the cancer burden in China.

In 2016, the United Nations (UN) set a target to achieve a relative one-third reduction from the 2015 level in premature mortality from noncommunicable diseases (NCDs), including cancer, by 2030 in its Sustainable Development Goal (SDG) target 3.4 [7]. In response to UN SDG target 3.4, the Chinese government issued the Healthy China 2030 Plan with disease prevention and longevity improvement as two of its primary goals [8,9]. A series of policies targeting cancer prevention and treatment was introduced by the government to reduce the burden of premature cancer death in an effort to accomplish SDG target 3.4 of a one-third reduction in premature mortality of NCDs overall. However, considering the vast territory in China and substantial regional differences in socioeconomic and hygienic conditions, the ability to meet SDG target 3.4 will vary greatly across the provinces. Therefore, there is an urgent need for China to implement tailored, region-specific cancer control measures based on reliable data and valid evidence from studies focusing on the local populations.

Hunan Province, located in the central-south of China, is one of the most populous provinces with over 73 million residents. The gross domestic product of Hunan Province was 4.18 trillion yuan (approximately US $606.03 billion) in 2020, ranking 9th among 32 provincial administrative divisions in mainland China. Cancer incidence and mortality rates in Hunan Province in 2018 were 248.24/100,000 and 154.50/100,000, respectively, representing a medium level across the nation [10]. The burdens of years of life lost and DALYs of cancer in Hunan Province were reported to be significantly higher than the national averages [11]. Of note, due to the special local lifestyle habits, Hunan Province has a much higher burden of certain cancer types than other regions, such as oral cavity cancer and nasopharynx cancer, with obviously higher incidences and mortality than national average levels [12,13].

Since publication of the World Health Organization (WHO) Global Monitoring Framework in 2013 [14], several modeling analyses have been conducted to examine the effects of selected risk factor control on reducing NCD mortality under different scenarios worldwide. Kontis et al [15] set six risk factor scenarios based on the Global Monitoring Framework, and estimated their impact on global and regional NCD mortality between 2010 and 2025 using a time-based, population impact fraction formula. They found that achieving all of these targets could efficiently reduce premature mortality from the four main NCDs, including cancers, by nearly 25% globally and in some specific regions [15]. Su et al [16] used an age-period-cohort model for prediction, which suggested that targets of reducing tobacco use could be more ambitious in Taiwan to meet the goal of a 25% reduction in premature cardiovascular disease mortality. Other studies indicated that the risk factor targets recommended by the WHO would be sufficient to help achieve the goal of a one-third reduction in premature mortality for all NCDs combined, but not for certain major subdivisions such as cancers [17,18].

In general, most previous modeling studies focused on overall NCDs at a national level, and there is limited evidence on cancer and its subcategories at a local provincial level. Moreover, the WHO's voluntary global targets did not include dietary and environmental factors, which are known to be important risk factors for cancer, and it remains unknown whether and how control of these factors may help with cancer prevention.

In light of such research gaps, we performed this study to project premature mortality from cancer in Hunan Province through 2030 under different risk factor control scenarios. Specifically, we projected whether SDG target 3.4 can be met for cancer prevention in Hunan Province and how many deaths from cancers can be prevented if all selected risk factors were controlled. The risk factors were selected based on the Global Monitoring Framework [14] while adding dietary and environmental factors. Projecting future premature mortality under various risk factor control scenarios is crucial for informing decision-making on cancer control and allocating the limited clinical and public health resources to curb the increasing deaths caused by cancer. Our results will help policy makers better formulate priorities for interventions that focus on these risk factors for cancer prevention and control.

Methods

Ethics Approval

This analysis centered on publicly available data with no identifiable information on the subjects studied. Therefore, research ethics board approval was not required for this study.

Selection of Cancer Sites and Related Risk Factors

Cancer sites were selected based on the rank of cancer deaths in the Hunan cancer registry annual report series over the past 10 years, while also taking into account endemic cancer types associated with special local lifestyle habits in Hunan, such as oral cavity cancer and nasopharynx cancer. Finally, 10 leading subcategories were selected, including cancers of the lung, esophagus, liver, stomach, pancreas, prostate, breast, colorectum, oral cavity, and nasopharynx.

Correspondingly, risk factors were selected based on the following criteria: (1) causally associated with cancers as evidenced by the latest Global Burden of Disease (GBD) study [3] or the Continuous Update Project (CUP) Expert Report (low vegetable and fruit intake were included due to consistent evidence showing their causal associations with cancers [19]); (2) available data of exposure levels from officially representative surveys or epidemiological studies; (3) potentially modifiable by available interventions; and (4) recognized as one of the leading global or national causes of disease burden. Ultimately, 10 risk factors were selected, including smoking, alcohol use, physical inactivity, high BMI, low vegetable intake, low fruit intake, high red meat intake, high salt intake, diabetes, and fine particulate matter (PM2.5). Details of cancer sites and related risk factors are shown in Multimedia Appendix 1.

Data Sources

The National Program of Cancer Registries (NPCR), launched in 2008, is responsible for the collection, evaluation, and publication of cancer data in China. In Hunan Province, 70.4% of cancer patients have registered in the NPCR to date [20], and all cancer cases are coded according to the International Classification of Diseases, 10th Revision. The details of data collection, management, and analysis for cancer registration in China have been described elsewhere [21]. In this study, we extracted the cause-specific cancer death rates, along with annual population data, between 2009 and 2017 from Hunan Statistical Yearbook 2020 to estimate the number of cancer deaths throughout the province. Information on risk factors was mainly obtained from the Chinese Chronic Disease and Risk Factor Surveillance (CCDRFS) survey, which is an ongoing, nationally representative survey involving a set of standard questionnaire interviews, physical examinations, and biological sampling [22]. In this analysis, we used exposure data in the years 2010 and 2018 to represent 2009 and 2017, respectively, due to a lack of investigations. Regarding PM2.5, since systematic monitoring in China did not start until 2013, we extracted data for Hunan from China Regional Estimates (V4.CH.03) of the Saint Louis University Atmospheric Composition Analysis Group, which applied a geographically weighted regression to calibrate regional PM2.5 concentrations through ground-based observations. Overall, population exposures to these risk factors were measured using metrics related to their variable types. For instance, smoking status and alcohol use were measured as the prevalence of people exposed, while dietary factors were measured as continuous variables. The specific details of risk factor measurements are shown in Multimedia Appendix 1.

To ensure data quality, the relative risk (RR) estimates for risk-cancer pairs were preferentially derived from summary results published by the GBD series and the CUP Expert Report. If these were not available, priority was given to meta-analyses or systematic reviews conducted in China or Asia. Studies that provided RRs on our predefined metrics were preferred and estimates for both genders were assumed to be equal if no separate values were available.

Constructing Risk Factor Scenarios

Based on the Global Monitoring Framework and Healthy China 2030, we constructed 12 separate scenarios of risk factor exposure for the year 2030. Among them, the baseline scenario projected cancer mortality to 2030 assuming that all risk factors continue to follow current trends (see Multimedia Appendix 2 for details on the risk factor exposure estimation). The baseline scenario was simulated using a proportional change model based on the historical trends of risk factors. The other 11 scenarios projected cancer mortality assuming that each of the 10 risk factors achieved the target of domestic or foreign control standards, both separately and in combination. Specifically, the first 10 scenarios were modeled when each risk factor achieved its target, respectively, and the last scenario was modeled when all 10 risk factors achieved their targets. Targets for red meat, vegetable, and fruit intake were set according to the currently available literature. Given that the current annual average PM2.5 concentration in Hunan Province has reached grade II of China’s air quality standards, we used grade I of 15 μg/m^3 as the target for PM2.5, which was also consistent with phase III of the WHO’s air quality guidelines [23]. Targets for other risk factors were set according to the WHO’s voluntary global targets [14]. Among the 10 risk factors, categorical exposures were lowered directly to the target levels in 2030, whereas targets for continuous exposures were established by shifting the population distributions left or right, assuming a constant distribution for each age-sex stratum. Exceptionally, metrics for exposures such as BMI and diabetes were held constant. Table 1 provides the details of the risk factors for each of the 12 scenarios.

Table 1. Scenario specifications in risk factor exposure projection according to the World Health Organization Global Monitoring Framework.

Scenario	Scenario specification
Natural trend	Age- and sex-specific risk factor exposures were projected assuming the annual change rate remained similar to that between 2009 and 2017.
Harmful alcohol use	Age- and sex-specific prevalence of harmful alcohol use is reduced relatively by 10% from the 2013 level. All other risk factors follow the natural trends.
Smoking	Age- and sex-specific prevalence of smoking in 2030 is reduced relatively by 30% from the 2013 level. All other risk factors follow the natural trends.
Physical inactivity	Age- and sex-specific prevalence of physical inactivity in 2030 is 10% relatively lower than that in 2013. All other risk factors follow the natural trends.
Diabetes	Age- and sex-specific prevalence of diabetes in 2030 is the same as in 2013. All other risk factors follow the natural trends.
High BMI	Age- and sex-specific distributions of BMI in 2030 are the same as in 2013. All other risk factors follow the natural trends.
Low vegetable intake	Age- and sex-specific prevalence of low vegetable intake in 2030 is reduced relatively by 30% from the 2013 level. All other risk factors follow the natural trends.
Low fruit intake	Age- and sex-specific prevalence of low fruit intake in 2030 is reduced relatively by 30% from the 2013 level. All other risk factors follow the natural trends.
High red meat intake	Age- and sex-specific prevalence of high red meat intake in 2030 is reduced relatively by 30% from the 2013 level. All other risk factors follow the natural trends.
High salt intake	Age- and sex-specific mean population salt intake in 2030 is reduced relatively by 30% from the 2013 level. All other risk factors follow the natural trends.
PM2.5[a]	The annually averaged PM2.5 concentration in 2030 is reduced to 15 μg/m^3, according to grade I of air quality standard GB3095-2012. All other risk factors follow the natural trends.
All targets are achieved in 2030	All targets described above are achieved in 2030.

[a] PM2.5: fine particulate matter.

Projection of Cancer Mortality

Our analysis was focused on examining premature mortality under 12 different scenarios. Consistent with the global documents, we defined premature cancer mortality as the probability of dying from cancers between the ages of 30 and 70 years [24]. To predict mortality for 2030, we considered the annual total deaths to be a function of two separate projections: a trend of deaths potentially driven by the selected risk factors and a business-as-usual (BAU) trend unattributable to these risk factors. Based on these trends, several steps were involved. First, according to the comparative risk assessment [25], we used the population-attributable fraction (PAF) to divide annual deaths into two parts that were attributable or unattributable to the specified risk factors. The PAF estimates the fraction of health outcomes that would be eliminated if the exposures were altered to ideally counterfactual distributions, and different formulas were applied for discrete and continuous variables in the calculations [25]. Second, we projected the unattributable deaths using the proportional change model, assuming that changes in this part would continue to follow the trends observed between 2009 and 2017. Risk factors in the baseline scenario were projected in the same manner. For the other 11 scenarios, the targeted change of each factor was distributed evenly between 2017 and 2030 to obtain an annual PAF value. Third, the total deaths for different scenarios in 2030 could be estimated using the unattributable deaths and PAFs, and the premature mortality could be calculated using age-specific death rates (in 5-year age groups) with a life table method. The levels in 2013 were considered as the baseline for calculations of relative reductions in this study. Analyses for PAF were performed using MATLAB 7.0, while other data were analyzed in SAS 9.4. Details of the analysis methods are shown in Multimedia Appendix 3.

Results

Premature Mortality From Cancers 2009-2017

Table 2 shows the estimated number of cancer deaths and mortality by gender in Hunan Province, China, from 2009 to 2017. The number of cancer deaths increased each year for both genders. The age-standardized mortality rates showed an increasing trend, with an annual percentage change of 1.60% for the total population, 2.19% for men, and 1.65% for women.

The premature mortality rates from selected cancers are shown in Table 3 and the trend analysis results are shown in Table 4. The average annual percentage change (AAPC) showed a significant increasing trend for all cancers combined (2.61%), with the largest AAPC occurring in oral cavity cancer (17.26%), followed by esophageal cancer (7.60%), colorectal cancer (7.13%), pancreatic cancer (7.10%), and lung cancer (4.06%). For the other cancers, the premature mortality remained stable with nonsignificant AAPC values, although a decreasing trend was observed in stomach cancer (–1.16%) and liver cancer (–1.92%), while an increasing trend was observed in nasopharynx cancer (2.00%), prostate cancer (11.91%), and breast cancer (4.82%).

Table 2. Estimated cancer deaths and mortality rates in Hunan Province, China, 2009-2017.

Year	Total			Men			Women		
	Deaths, n	Mortality rate (1/100,000)	Standardized mortality rate (1/100,000)	Deaths, n	Mortality rate (1/100,000)	Standardized mortality rate (1/100,000)	Deaths, n	Mortality rate (1/100,000)	Standardized mortality rate (1/100,000)
2009	81,159	117.56	89.53	50,480	140.88	108.62	30,679	92.49	69.51
2010	83,332	127.03	87.13	53,938	159.69	110.53	29,394	92.07	62.85
2011	88,269	123.61	83.53	56,536	152.84	104.19	31,733	92.34	62.21
2012	95,044	132.08	91.14	62,883	168.78	117.78	32,161	93.11	63.74
2013	96,173	133.84	87.14	63,640	171.43	112.76	32,533	94.71	61.18
2014	102,821	142.47	92.58	67,338	180.00	119.36	35,483	102.51	65.26
2015	102,310	140.91	91.50	66,543	176.93	117.34	35,767	102.75	65.11
2016	108,097	147.48	97.57	70,590	185.88	125.61	37,507	106.52	68.81
2017	111,719	152.91	99.38	72,708	192.26	128.18	39,011	111.00	69.91
APC[a] (95%CI)	N/A[b]	3.13 (2.50-3.80)	1.60 (0.50-2.70)	N/A	3.43 (2.40-4.50)	2.19 (0.60-3.40)	N/A	2.54 (1.70-3.30)	1.65 (0.60-2.70)
P value	N/A	<.001	.011	N/A	<.001	.002	N/A	<.001	.007

[a] APC: annual percentage change.

[b] N/A: not applicable.

Table 3. Premature mortality (%) from selected cancers in Hunan Province, China, 2009-2017.[a]

Year	Lung cancer	Gastric cancer	Liver cancer	Colorectal cancer	Esophageal cancer	Pancreatic cancer	Nasopharynx cancer	Oral cavity cancer	Prostate cancer	Breast cancer	All cancers
2009	1.95	0.66	1.36	0.44	0.16	0.11	0.18	0.05	0.02	0.65	6.64
2010	2.02	0.60	1.64	0.34	0.16	0.10	0.19	0.04	0.07	0.32	6.53
2011	1.87	0.50	1.63	0.36	0.23	0.08	0.25	0.02	0.03	0.36	6.26
2012	2.25	0.55	1.46	0.54	0.29	0.12	0.28	0.05	0.03	0.49	7.06
2013	2.15	0.53	1.13	0.48	0.24	0.13	0.23	0.08	0.07	0.44	6.74
2014	2.42	0.57	1.27	0.55	0.25	0.13	0.23	0.09	0.06	0.55	7.31
2015	2.41	0.48	1.24	0.56	0.26	0.14	0.22	0.09	0.07	0.57	7.22
2016	2.55	0.56	1.35	0.62	0.29	0.16	0.23	0.11	0.07	0.62	7.83
2017	2.57	0.59	1.39	0.63	0.26	0.17	0.21	0.11	0.07	0.62	7.88

[a] Premature mortality is the probability of dying between the ages of 30 and 70 years from a specific cause that was calculated using the life table method.

Table 4. Annual percentage change (APC) in premature mortality from selected cancers in Hunan Province, China, 2009-2017.

Cancer type	APC (95% CI)	*P* value
Lung	4.06 (2.50 to 5.60)	<.001
Gastric	–1.16 (–4.00 to 1.80)	.38
Liver	–1.92 (–5.40 to 1.70)	.24
Colorectal	7.13 (3.00 to 11.50)	.005
Esophageal		
2009-2012	21.59 (2.20 to 44.60)	.04
2012-2017	0.03 (–7.50 to 8.00)	.99
Average	7.60 (1.60 to 13.90)	<.001
Pancreatic	7.10 (2.90 to 11.40)	.005
Nasopharynx		
2009-2012	13.91 (3.20 to 25.70)	.02
2012-2017	–4.59 (–8.70 to –0.30)	.04
Average	2.00 (–1.30 to 5.30)	.20
Oral cavity	17.26 (3.60 to 32.80)	.02
Prostate	11.91 (–1.30 to 26.80)	.07
Breast	4.82 (–2.40 to 12.60)	.17
All cancers	2.61 (1.40 to 3.80)	.001

Premature Mortality in the Baseline Scenario by 2030

Table 5 shows a comparison of premature deaths and mortality between 2013 and projected for 2030 in the baseline scenario where all risk factors continue their current trends. In 2030, an estimated 97,787 people would die prematurely from all cancers, with a premature mortality rate of 9.74%, which was 44.47% higher than that in 2013 (Table 5). For the number of premature deaths, lung cancer was expected to account for the largest proportion of total cancer deaths (29,337 deaths), followed by liver cancer (15,545 deaths), colorectal cancer (12,342 deaths), oral cavity cancer (7470 deaths), and esophageal cancer (5977 deaths). All cancers showed increases in the number of premature deaths except for gastric cancer, which was expected to decrease from 4104 in 2013 to 3053 in 2030. The mortality rate of all cancers was estimated to increase by 81.55% from 143.54 to 260.59 per 100,000 people. All cancers exhibited increasing trends in mortality rate except for gastric cancer, which was expected to decrease by 24.62% from 2013 to 2030.

The premature mortality rates for all cancers and each subcategory were consistently higher in men than in women, with differences of more than 5-fold (Table 5). For all cancers, relative increments of 61.47% and 13.46% were observed among men and women, respectively. For men, all cancers showed substantial increases in premature mortality, ranging from 38.25% for lung cancer to 1194.33% for oral cavity cancer, except for gastric cancer, which showed a decreasing trend (–32.51%). For women, most cancers showed increases in premature mortality, ranging from 14.04% for lung cancer to 158.84% for colorectal cancer, except for esophageal cancer (–64.61%), gastric cancer (–52.16%), pancreatic cancer (–45.08%), and nasopharynx cancer (–21.86%). Despite significant gender discrepancies, premature mortality for the whole population generally showed an increasing trend, with the greatest increase occurring in oral cavity cancer (971.53%).

Table 5. Deaths and premature mortality of main cancers for people aged 30-70 years in 2013 and projections for 2030 if risk factor trends continue in Hunan Province, China.

Disease	2013			2030			Percent change		
	Premature deaths, n	Mortality rate (1/100,000)	Premature mortality (%)	Premature deaths, n	Mortality rate (1/100,000)	Premature mortality (%)	Premature deaths	Mortality rate	Premature mortality
Men									
Total	36,886	188.18	8.82	71,920	382.42	14.24	94.98	103.22	61.47
Lung cancer	12,700	64.79	3.28	22,599	120.17	4.54	77.95	85.47	38.35
Gastric cancer	2785	14.21	0.69	2290	12.17	0.46	–17.78	–14.30	–32.51
Liver cancer	7373	37.61	1.71	12,360	65.72	2.61	67.64	74.72	52.34
Colorectal cancer	2405	12.27	0.60	7539	40.09	1.60	213.49	226.74	164.82
Esophageal cancer	1624	8.29	0.43	5916	31.46	1.35	264.19	279.58	210.91
Pancreatic cancer	591	3.01	0.15	1414	7.52	0.28	139.34	149.45	82.95
Nasopharynx cancer	1508	7.70	0.34	2852	15.16	0.63	89.07	97.06	82.95
Oral cavity cancer	591	3.01	0.13	7255	38.58	1.71	1128.22	1180.13	1194.33
Prostate cancer	253	1.29	0.07	2504	13.31	0.52	889.01	930.81	592.15
Other cancers	7046	23.68	1.72	7193	38.25	1.46	2.08	61.53	–14.85
Women									
Total	17,687	97.25	4.51	25,867	138.19	5.11	46.25	42.10	13.46
Lung cancer	3409	18.75	0.95	6738	36.00	1.38	97.63	92.02	44.93
Gastric cancer	1319	7.25	0.36	763	4.08	0.17	–42.13	–43.78	–52.16
Liver cancer	1928	10.60	0.51	3185	17.01	0.59	65.19	60.50	14.04
Colorectal cancer	1380	7.59	0.36	4803	25.66	0.92	248.04	238.15	158.84
Esophageal cancer	112	0.61	0.04	61	0.33	0.01	–45.30	–46.86	–64.61
Pancreatic cancer	365	2.01	0.10	271	1.45	0.06	–25.73	–27.84	–45.08
Nasopharynx cancer	497	2.73	0.11	400	2.14	0.09	–19.52	–21.81	–21.86
Oral cavity cancer	101	0.56	0.03	215	1.15	0.04	112.05	106.02	55.81
Breast cancer	1827	10.04	0.44	3185	17.01	0.69	74.37	69.42	57.56
Other cancers	6748	29.51	1.69	6245	33.36	1.27	–7.45	13.03	–24.88
Both genders									
Total	54,572	143.54	6.74	97,787	260.59	9.74	79.19	81.55	44.47
Lung cancer	16,109	42.18	2.15	29,337	78.18	2.96	82.12	85.33	37.57
Gastric cancer	4104	10.79	0.53	3053	8.14	0.32	–25.61	–24.62	–39.93
Liver cancer	9301	24.35	1.13	15,545	41.42	1.59	67.13	70.11	40.72
Colorectal cancer	3785	9.97	0.48	12,342	32.89	1.25	226.09	229.86	158.76
Esophageal cancer	1736	4.52	0.24	5977	15.93	0.68	244.29	252.39	182.06
Pancreatic cancer	956	2.52	0.13	1685	4.49	0.17	76.26	78.18	30.58
Nasopharynx cancer	2006	5.26	0.23	3252	8.67	0.35	62.15	64.77	53.36
Oral cavity cancer	692	1.81	0.08	7470	19.91	0.87	979.24	1001.08	971.53
Prostate cancer	253	1.29	0.07	2504	13.31	0.52	889.01	930.81	592.15
Breast cancer	1827	10.04	0.44	3185	17.01	0.69	74.37	69.42	57.56
Other cancers	13,794	26.54	1.70	13,438	35.81	1.36	–2.58	34.91	–19.87

Premature Deaths Avoided and Premature Mortality in Multiple Scenarios

Table 6 shows the projected premature deaths and mortality in 2030 with all risk factors under control, which was further compared with the rates for 2013 and the baseline scenario. Compared with the baseline scenario where all risk factors continue their current trends, 14.90% of premature deaths from cancers would be avoided if all risk factor targets were achieved. Lung cancer had the largest decrease in the number of cancer deaths compared to the baseline scenario (–27.30%), followed by colorectal cancer (–19.88%), esophageal cancer (–17.73%), and breast cancer (–14.08%). A further comparison of avoided cancer deaths by gender showed that men (–15.53%) benefited much more through these combined risk factor control targets than women (–13.16%).

The modeling scenarios seek to avert one-third of premature mortality by 2030. However, this goal is hard to accomplish. For all cancers combined, premature mortality among the total population was expected to increase by 23.65% compared to the baseline year of 2013, even if all risk factor control targets were reached by 2030 (Table 6, Figure 1).

For subcategories, all cancers showed increases in premature mortality compared with that in 2013 in the combined risk factor control target-achieved scenarios, except for gastric cancer with a decrease of 41.63% in the total population (34.45% for men and 53.40% for women). However, it should be noted that the premature deaths and mortality of gastric cancer would still decrease substantially even if all risk factors continue their current trends, as shown in the baseline scenario. A decrease in premature mortality was also found for women in esophageal cancer (72.63%), pancreatic cancer (48.71%), and nasopharynx cancer (29.44%) under combined target–achieved scenarios. Although the combined risk factor control targets failed to achieve the one-third reduction of the cancer mortality rate set by the UN, they could still lead to notable decreases in premature mortality compared with the baseline scenario in 2030.

Moreover, the impact on cancer premature mortality varied substantially across different risk factors. For all cancers combined, diabetes and low fruit intake were the top two leading risk factors of cancer premature mortality for both genders. For instance, halting the rise in the prevalence of diabetes may contribute to nearly half of the reductions in cancer premature mortality for both genders (Figure 1). The impact of other risk factors on cancer premature mortality depended on gender. A halt in the rise of BMI had the third-largest impact for men, whereas for women (Figure 2), it was the reduction in ambient PM2.5 levels (Figure 3). For risk factors such as smoking, high red meat intake, and physical inactivity, the predefined risk factor control targets seemed to be insufficient to have pronounced benefits on cancer premature mortality. For risk factors such as harmful alcohol use and high salt intake, the predefined control targets set by the WHO showed much smaller reductions in mortality than the baseline BAU scenario, indicating the need for setting more ambitious targets. In all simulated scenarios, men were projected to have greater gains than women for all cancers after risk factor controls. Scenario projections by gender are shown in Multimedia Appendix 4.

Table 6. Premature cancer mortality for people aged 30-70 years in 2030 if all risk factor targets are achieved in Hunan Province, China, and the comparison with baseline values.

Disease	2030 (if all risk factor targets are achieved)		Percent change compared with baseline in 2013, %		Percent change compared with baseline in 2030, %	
	Deaths, n	Premature mortality, %	Deaths	Premature mortality	Deaths	Premature mortality
Men						
Total	60,750	12.15	64.70	37.71	–15.53	–14.72
Lung cancer	16,207	3.26	27.62	–0.69	–28.28	–28.22
Gastric cancer	2220	0.45	–20.29	–34.45	–3.05	–2.88
Liver cancer	10,782	2.25	46.24	31.48	–12.76	–13.69
Colorectal cancer	5956	1.27	147.66	109.73	–21.00	–20.80
Esophageal cancer	4870	1.11	199.82	156.42	–17.67	–17.53
Pancreatic cancer	1267	0.25	114.53	62.42	–10.36	–11.22
Nasopharynx cancer	2794	0.61	85.23	79.30	–2.03	–2.00
Oral cavity cancer	6878	1.62	1064.35	1126.49	–5.20	–5.24
Prostate cancer	2429	0.50	859.36	572.08	–3.00	–2.90
Other cancers	7193	1.46	2.08	–14.85	0	0
Women						
Total	22,463	4.46	27.00	–1.09	–13.16	–12.83
Lung cancer	5120	1.05	50.17	10.71	–24.01	–23.61
Gastric cancer	743	0.17	–43.68	–53.40	–2.68	–2.59
Liver cancer	2811	0.51	45.80	–0.45	–11.74	–12.71
Colorectal cancer	3932	0.76	184.92	112.25	–18.14	–18.00
Esophageal cancer	47	0.01	–57.72	–72.63	–22.70	–22.66
Pancreatic cancer	253	0.05	–30.66	–48.71	–6.65	–6.60
Nasopharynx cancer	360	0.08	–27.61	–29.44	–10.05	–9.71
Oral cavity cancer	186	0.04	83.38	35.34	–13.52	–13.14
Breast cancer	2737	0.59	49.83	35.58	–14.08	–13.95
Other cancers	6245	1.27	–7.45	–24.88	0	0
Both genders						
Total	83,213	8.33	52.48	23.65	–14.90	–14.41
Lung cancer	21,327	2.15	32.39	0.11	–27.30	–27.23
Gastric cancer	2963	0.31	–27.81	–41.63	–2.96	–2.82
Liver cancer	13,593	1.37	46.15	21.61	–12.55	–13.58
Colorectal cancer	9888	1.00	161.25	107.52	–19.88	–19.80
Esophageal cancer	4917	0.56	183.26	132.30	–17.73	–17.64
Pancreatic cancer	1520	0.15	59.05	16.93	–9.77	–10.45
Nasopharynx cancer	3154	0.34	57.26	48.79	–3.01	–2.98
Oral cavity cancer	7064	0.82	920.53	912.83	–5.44	–5.48
Prostate cancer	2429	0.50	859.36	572.08	–3.00	–2.90
Breast cancer	2737	0.59	49.83	35.58	–14.08	–13.95
Other cancers	13,438	1.36	–2.58	–19.87	0	0

Figure 1. Probability of premature death due to cancers for people between ages 30 and 70 years in Hunan Province, China.

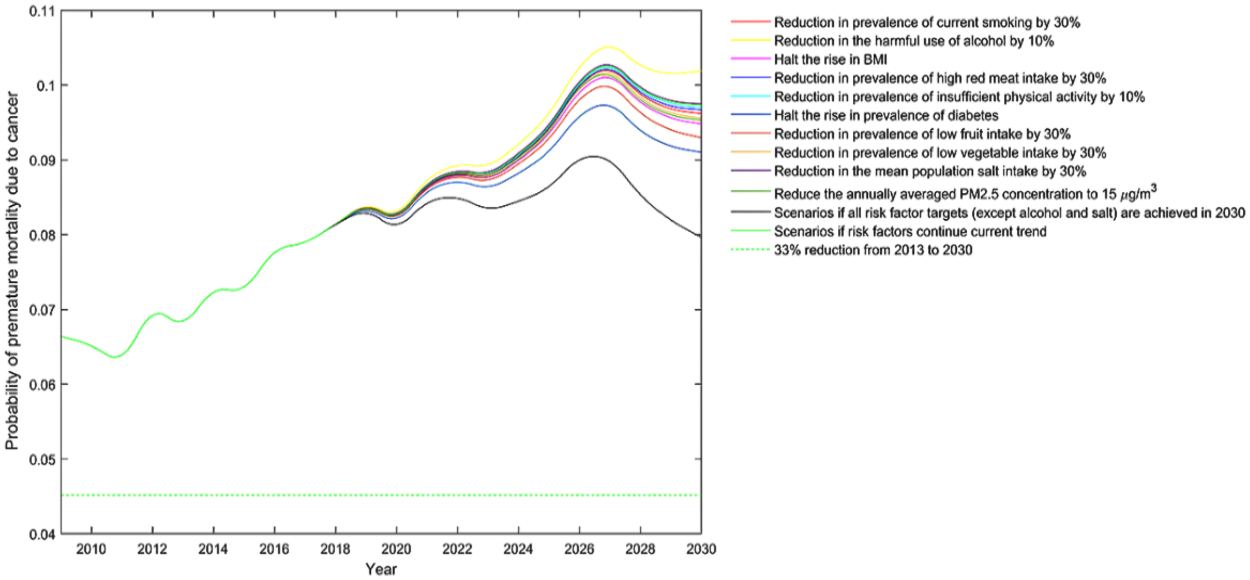


Figure 2. Probability of premature death due to cancers in men aged 30-70 years in Hunan Province, China.

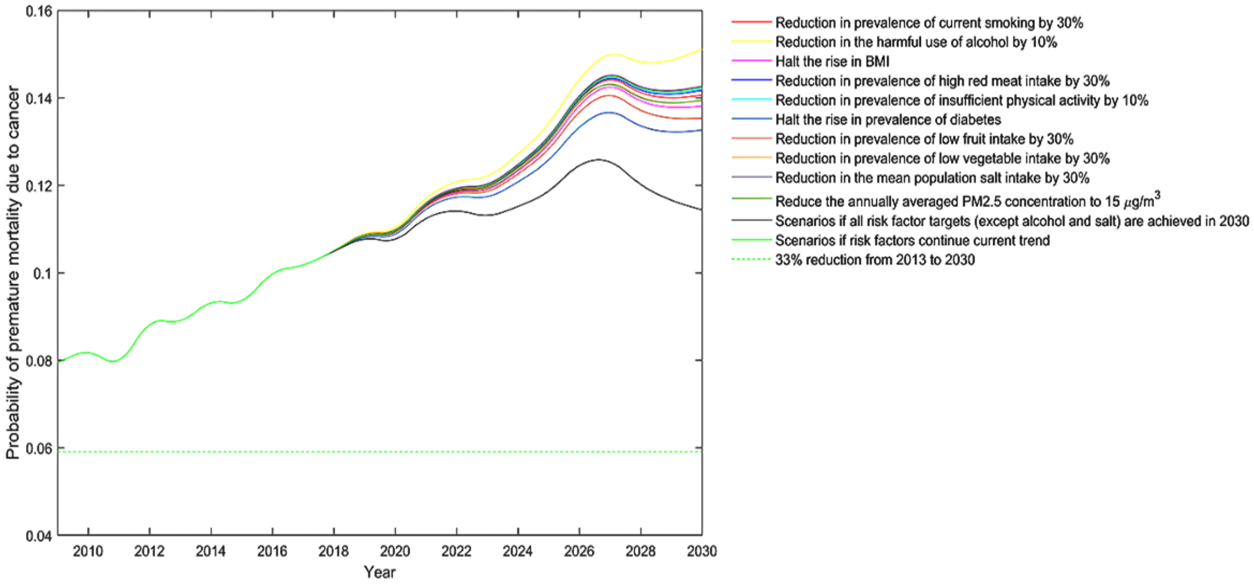
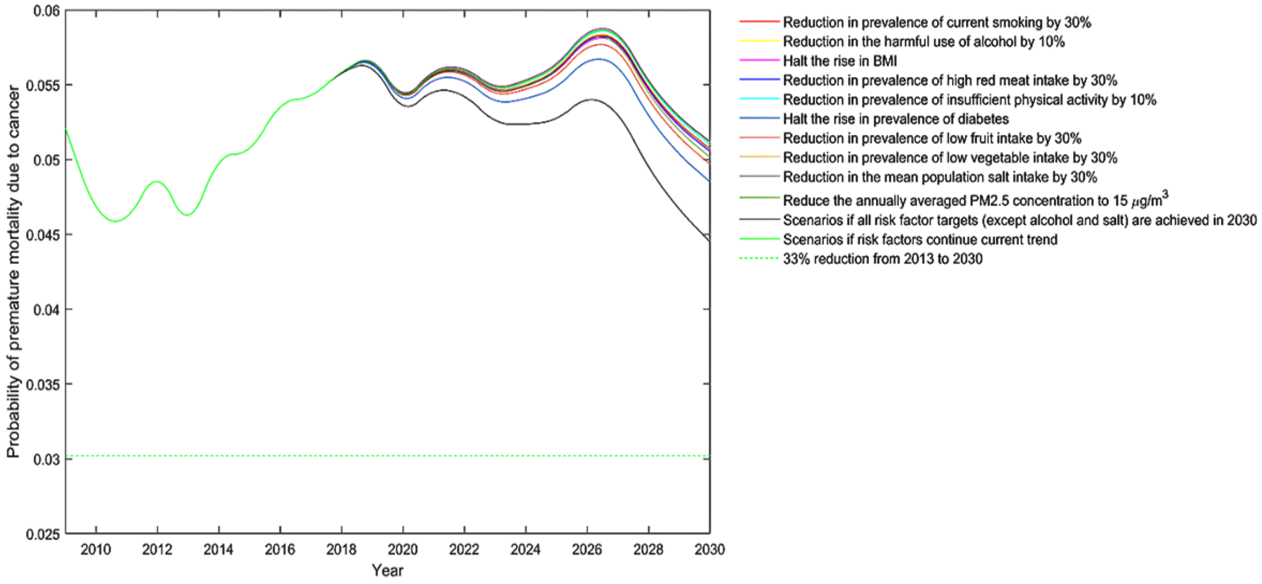


Figure 3. Probability of premature death due to cancers in women aged 30-70 years in Hunan Province, China.



Discussion

Principal Findings

In this study, we projected premature mortality from 10 leading cancers in Hunan Province under 12 different risk factor control scenarios in 2030 and evaluated whether SDG target 3.4 can be met for cancer prevention. The results suggest that the one-third reduction goal in premature cancer mortality would not be achieved in Hunan, even if all related risk factor control targets were reached by 2030. This finding is similar to that of a previous study conducted by Li et al [17], showing that achieving all related risk targets would lead to a one-third reduction for all NCDs combined, but not for cancers, in China. These findings suggest that meeting SDG target 3.4 for cancers would require extra efforts from both national and local governments. Notably, Li et al [17] only included six risk factors, in which merely three (ie, smoking, BMI, and physical inactivity) were modeled in the projection of cancer burdens. It seems difficult to achieve dramatic reductions in premature cancer mortality by only controlling for a limited number of risk factors; thus, further inclusion of more risk factors is needed for simulation in the whole country. Although our study contained a wider range of risk factors, the predicted premature cancer mortality in Hunan Province failed to reduce by one-third in 2030, which may be due to the combined effect of several factors such as the conservativeness of some risk factor targets and population aging. According to Chen et al [26], the proportion of people aged over 55 years is expected to increase to more than 30% in Hunan in 2030 [26]. In addition, epidemiological studies have shown that cancer mortality is significantly higher in people over 55 years old than in other age groups, and approximately 27% of cancer deaths were attributed to population aging in China in 2019 [10,21,26], which makes it more difficult to reverse the status quo. Hence, more stringent targets on key risk factors may be required in Hunan.

As for various cancer subcategories, it is interesting that risk factor control targets appeared to generate relatively minor additional benefits for cancers that had already experienced dramatic reductions under the BAU scenario. For other cancers that had not experienced reductions under the BAU scenario, the joint control of all related risk factors may generate larger additional benefits. For instance, the control of all five modifiable risk factors of lung cancer would reduce premature mortality by 28.2% for men and by 23.6% for women compared to the BAU scenario in 2030. These findings suggest that it is more cost-effective to control risk factors for cancers with tendencies toward worse conditions.

Through modeling, our estimates also illustrated significant discrepancies in premature mortality reduction across various risk factor control targets. In parallel with the previous study conducted by Li et al [17], improving physical inactivity resulted in minimal reduction in premature cancer mortality, which might be largely because physical inactivity was only causally related to two cancer sites (ie, colorectal cancer and breast cancer). Another explanation may be that the mild reduction of 10% in physical inactivity was not sufficient to produce a significant reduction in premature cancer mortality in 2030. A similar result was also seen for high red meat intake, which was only causally associated with breast and colorectal cancers. In comparison, smoking is causally associated with a wide range of cancers and showed the largest effect in reducing premature cancer mortality by smoking control. As the largest tobacco producer and consumer globally, China has enacted a set of antismoking policies and regulations after signing the WHO's Framework Convention on Tobacco Control in 2003 [27], which include public smoking bans, a rise in tobacco taxes, and warnings on the cigarette pack. The serial national surveys on smoking showed that the smoking prevalence in China had been declining steadily in recent years, although remaining at a high plateau among men [28]. Hunan Province has a consistently higher smoking prevalence than the national average level [29], possibly due to a later implementation of smoking restrictions. Recently, Hunan Province launched a series of feasible and effective smoking restriction policies, which have greatly reduced the smoking rate. Several rounds of CCDRFS surveys have also indicated a favorably moderate decrease of the smoking prevalence in Hunan, which may partly account for the relatively lower benefit in mortality reduction through the control of smoking than the control of high BMI and diabetes that have been rising steadily in Hunan. In addition, due to the especially low smoking prevalence among women in Hunan, the WHO target of a 30% reduction may produce much less benefit for women than that among men.

The past decade has also seen significant decreases in the mean population salt intake and prevalence of harmful alcohol use in Hunan, both of which led to more favorable trends than those of the WHO targets. Of note, the prevalence of current alcohol use was still maintained at a high level and was much higher in men than women, especially for hazardous and harmful alcohol use. The high alcohol use prevalence among men in China may be explained by the traditional Chinese culture that encourages drinking as a socially acceptable way to show their dominant positions in society. Disturbingly, there was a substantial increase in alcohol use following the rapid economic transition in China [30]; yet, the current legislation on alcohol use only focuses on problematic alcohol use such as drunk driving, rather than on alcohol use per se. Evidence has shown that alcohol use is the fourth-leading modifiable risk factor that contributes to cancer mortality among Chinese men [6], and abstinence from alcohol would increase life expectancy by 0.77 years [31]. These findings suggest that more effective measures are warranted to reduce alcohol use. The WHO has recommended some cost-effective alcohol prevention measures named "best buys," which included increasing taxes, enforcing bans on alcohol advertising, and restricting the availability of alcohol [32]. With respect to high salt intake, studies have shown that the daily salt intake in China had been decreasing slowly, which was largely attributed to the government's multicomponent strategies, including labeling, media campaigns, and voluntary reformulation of the salt industry, among others [33,34]. However, the mean salt intake in 2019 was approximately 9.3 grams per day [35], which was still much higher than the WHO's recommended level of 6.0 grams per day. Hence, there is still room for improvement in reducing the salt intake among Chinese residents.

Apart from the above risk factors, whether SDG target 3.4 can be realized in 2030 largely depends on the high-impact factors, including diabetes, high BMI, and insufficient intake of fruits and vegetables. China has the largest population of individuals with diabetes in the world, and the prevalence of diabetes has been sharply increasing in recent decades in China, including in Hunan Province [36]. Without proper measures to control blood glucose, China will continue to have the largest population of diabetes in 2030 with a predicted number of 140 million [37]. Previous studies discovered that high concentrations of glucose could provide steady energy for the growth of tumor cells [38,39], and thus diabetes was identified as an independent risk factor for several cancer types [40]. It was estimated that achieving the risk control target for diabetes in 2030 would avoid 57,400 deaths from NCDs in China [17], and our study also predicted notable benefits in premature cancer mortality reduction from the control of diabetes. In China, a range of programs such as primary diabetic health care have been promoted to curb the rapid rise in diabetes [41]. However, due to the imbalanced regional economic development, the capacity of primary medical services is relatively backward in the central and western regions of China, resulting in a lack of standardized diabetic management and low screening rates of diabetes-related complications [42]. Thus, there is still a long way to go to meet the commitments of the UN agenda.

Overweight and obesity prevalence among adolescents and adults has been increasing steadily in China, including Hunan Province, for the past two decades [35]. This is largely attributable to high-calorie dietary habits and reduced physical activity. Researchers have projected that the prevalence of overweight and obesity, if not controlled effectively, might reach 65.3% in adults and 31.8% in adolescents by 2030, with more than 800 million people reaching the overweight and obese categories in China [43]. The government has made great efforts to prevent overweight and obesity, including several school-based programs (such as the Healthy Children Action Plan and National Nutrition Campus Program) and community-based programs (such as the National Healthy Lifestyle Action) to promote healthy diet and exercise among the population [43]. However, existing policies on overweight and obesity are still fragmented in China; thus, coordinated and multisectoral strategies are needed to reach the target of halting the rise in overweight and obesity by 2030.

Fruit and vegetable intake is an indispensable part of a healthy diet. Studies have shown that the average daily intake of fruit in Hunan Province has been increasing steadily in recent years, while the intake of vegetables has been declining [44,45], which reflects a remarkable transition in dietary structure. Despite these trends, the proportions of adults who reached the recommended levels of Chinese dietary guidelines for both fruit and vegetable intake remained at low levels. Globally, inadequate consumption of fruits and vegetables has been validated to contribute to a large portion of the cancer burden in many countries such as the United Kingdom, Japan, and China [3,6,46]. In China, an unhealthy diet including low fruit and vegetable intake has accounted for more than 10% of DALYs [3]. Dietary guidelines with different recommendations have been developed worldwide to tackle the problem. The WHO recommends a total consumption of at least 400 grams per day of fruits and vegetables for adults, while in China the recommended level is 500 grams per day. Although the Global NCD Action Plan has included the prevalence of inadequate intake of fruits and vegetables as a monitoring indicator, no specific target has been set by the WHO to date. In our study, a 30% reduction for inadequate fruit and vegetable intake displayed relatively high benefits in cancer mortality reduction, further supporting the beneficial role of dietary risk control in achieving SDG target 3.4. Currently, health education with a focus on children is the major approach for promoting a healthy diet in China, which indeed has increased people's awareness [47]. Nevertheless, given that other factors such as availability and prices may also influence eating behavior, there is still much work to be done.

We also validated our projection with other methods or assumptions. In the baseline scenario, where all risk factors continue their current trends, the number of premature deaths from cancers increased from 65,443 in 2017 to 97,787 in 2030.

Given that approximately 4 million deaths are expected to occur in 2030 in China [48], an increase of only 32,344 premature deaths in such a populous province may be underestimated. Nevertheless, we further estimated the total number of cancer deaths for all ages and found that the value increased from 111,719 in 2017 to 235,578 in 2030 (see Multimedia Appendix 5), which indicated that people older than 70 years represented the major group of cancer deaths. In addition, assuming that cancer mortality trends continued to 2030 with constant change, we validated the predicted deaths in the BAU scenario using the proportional change model, which showed similar results with only a 3.6% difference in total cancer deaths and a 1.6% difference in overall premature mortality (see Multimedia Appendix 6). Therefore, we believe our projections to be credible.

Limitations

Some limitations of our work s*f*hould be considered. First, Hunan launched its cancer registry program in 2009, and the work at the early stages might be imperfect with low population coverage. Even by the end of 2020, the cancer registration in Hunan had only covered approximately 70% of the population, without achieving full coverage, which may to some extent cause a certain bias in the estimation of actual cancer deaths. Nevertheless, with the efforts of local governments, the coverage of cancer registration has rapidly expanded and the data quality has steadily improved in recent years. Some of the data were even cited in monographs of the International Agency for Research on Cancer. In addition, our data for risk factors were drawn from the CCDRFS survey; therefore, all limitations in estimates of levels in the CCDRFS study apply to this analysis.

Second, we used RR estimates primarily from the GBD series and the CUP Expert Report due to the lack of high-quality meta-analyses and prospective cohort studies in China, which may make our results statistically unstable. However, with more and more large cohort studies being carried out in China, more reliable RRs for China could be available in future studies. Third, due to a lack of dynamic monitoring data, some region-specific risk factors such as betel quid, which is classified as a class 1 carcinogen and has a high prevalence in Hunan, has not been included in this analysis. Considering that the Chinese government has made great efforts on sales restriction and increasing the public awareness of its harm in recent years, the prevalence of betel quid chewing in Hunan may decline steadily; thus, modeling without consideration of betel quid may lead to overestimation of future cancer mortality. Furthermore, there are potential interactions among the selected risk factors; however, due to the absence of information on most interactions, we simply calculated their combined effects on cancers based on the assumption of independence, which may lead to some uncertainties in our results. Hence, solid evidence–based joint RRs on cancers are warranted for future studies.

Fourth, we did not investigate the impact of population aging on premature cancer mortality due to insufficient technological conditions and time. In further studies, we will try to examine the fractions and trends attributable to population aging and its interactions with various risk factors on premature cancer mortality. Fifth, since the current health outcome reflects the cumulative effect of past exposures, risk factors such as smoking and alcohol were subdivided by duration and amount whenever possible. However, no lag effect was considered when data on specific information were unavailable. Nevertheless, calculations of PAFs in our study referred to the comparative risk assessment model from the latest GBD series, and the results were similar to those of previously published literature [6].

Conclusions

In summary, this modeling study illustrates that the absolute burden of premature deaths due to cancers will continue to increase over the next dozen years in the Hunan province of China. Notable health gains could be achieved by addressing unhealthy risk factors for cancers. However, existing targets on related risk factors are not sufficient, particularly in men, to achieve the one-third reduction goal in premature cancer mortality. More aggressive risk targets based on local conditions are urgently needed.

Acknowledgments

We thank all those who provided data for our analyses. This research was supported by the Science and Technology Innovation Program of Hunan Province (grant number 2022SK2050); Hunan Cancer Hospital Climb Plan (grant number YF2021007); and the Natural Science Foundation of Hunan Province, China (grant number 2021JJ40326).

Data Availability

The data will be available from the corresponding author on request.

Authors' Contributions

WW and SY secured funding for this work. JW and WW are joint first authors, and equally conceived, designed, and performed the work. XL, KX, YZ, ZS, YH, HX, CL, SC, SW, JG, ZL, ML, MX, DJ, ZF, MC, and SY contributed to the data collection and extraction. CL, ZL, and SY revised the manuscript. All authors gave final approval and agree to be accountable for all aspects of the work, ensuring integrity and accuracy.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details on risk factors and cancer sites.
[DOCX File , 21 KB - publichealth_v9i1e43967_app1.docx]

Multimedia Appendix 2
Risk factor exposure estimation.
[DOCX File , 25 KB - publichealth_v9i1e43967_app2.docx]

Multimedia Appendix 3
Analysis methods.
[DOCX File , 20 KB - publichealth_v9i1e43967_app3.docx]

Multimedia Appendix 4
Scenario projections for each cancer by gender.
[DOCX File , 416 KB - publichealth_v9i1e43967_app4.docx]

Multimedia Appendix 5
Death projection for people of all ages.
[DOCX File , 24 KB - publichealth_v9i1e43967_app5.docx]

Multimedia Appendix 6
Comparison between the baseline scenario and proportional change model.
[DOCX File , 25 KB - publichealth_v9i1e43967_app6.docx]

References

1. Cancer. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/cancer [accessed 2023-02-17]
2. China: Globocan 2020. Global Cancer Observatory. URL: https://gco.iarc.fr/today/data/factsheets/populations/160-china-fact-sheets.pdf [accessed 2023-02-17]
3. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
4. Pearce A, Sharp L, Hanly P, Barchuk A, Bray F, de Camargo Cancela M, et al. Productivity losses due to premature mortality from cancer in Brazil, Russia, India, China, and South Africa (BRICS): A population-based comparison. Cancer Epidemiol 2018 Apr;53:27-34. [doi: 10.1016/j.canep.2017.12.013] [Medline: 29353153]
5. Hofmarcher T, Lindgren P, Wilking N, Jönsson B. The cost of cancer in Europe 2018. Eur J Cancer 2020 Apr;129:41-49 [FREE Full text] [doi: 10.1016/j.ejca.2020.01.011] [Medline: 32120274]
6. Chen W, Xia C, Zheng R, Zhou M, Lin C, Zeng H, et al. Disparities by province, age, and sex in site-specific cancer burden attributable to 23 potentially modifiable risk factors in China: a comparative risk assessment. Lancet Glob Health 2019 Feb;7(2):e257-e269 [FREE Full text] [doi: 10.1016/S2214-109X(18)30488-1] [Medline: 30683243]
7. Transforming our world: the 2030 agenda for sustainable development. United Nations. URL: https://sdgs.un.org/2030agenda [accessed 2023-02-17]
8. Tan X, Liu X, Shao H. Healthy China 2030: a vision for health care. Value Health Reg Issues 2017 May;12:112-114. [doi: 10.1016/j.vhri.2017.04.001] [Medline: 28648308]
9. The plan for "Healthy China 2030". CPC Central Committee, State Council. URL: http://www.gov.cn/xinwen/2016-10/25/content_5124174.htm [accessed 2023-02-17]
10. Zou Y, Liao X, Xu K, Xiao H, Hu Y, Shi Z. Cancer incidence and mortality in Hunan Cancer Registration Areas in 2018. China Cancer 2022 Feb 15;31(4):241-248.
11. Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 Sep 28;394(10204):1145-1158 [FREE Full text] [doi: 10.1016/S0140-6736(19)30427-1] [Medline: 31248666]
12. Hu Y, Zhong R, Li H, Zou Y. Effects of betel quid, smoking and alcohol on oral cancer risk: a case-control study in Hunan Province, China. Subst Use Misuse 2020 Jun 22;55(9):1501-1508. [doi: 10.1080/10826084.2020.1750031] [Medline: 32569534]
13. Huang W, Zhu S, Zou Y, Shi Z, Xu K. Incidence and mortality of oral cancer in registered regions of Hunan in 2009–2015. China Cancer 2017 Jun 27;26(7):507-514.
14. Global action plan for the prevention and control of noncommunicable diseases 2013-2020. World Health Organization. URL: http://apps.who.int/iris/bitstream/10665/94384/1/9789241506236_eng.pdf [accessed 2023-02-17]

15. Kontis V, Mathers CD, Rehm J, Stevens GA, Shield KD, Bonita R, et al. Contribution of six risk factors to achieving the 25×25 non-communicable disease mortality reduction target: a modelling study. Lancet 2014 Aug 02;384(9941):427-437. [doi: 10.1016/S0140-6736(14)60616-4] [Medline: 24797573]
16. Su S, Lee W, Chen T, Wang H, Su T, Jeng J, et al. An evaluation of the 25 by 25 goal for premature cardiovascular disease mortality in Taiwan: an age-period-cohort analysis, population attributable fraction and national population-based study. Heart Asia 2017 Aug 07;9(2):e010905 [FREE Full text] [doi: 10.1136/heartasia-2017-010905] [Medline: 29467836]
17. Li Y, Zeng X, Liu J, Liu Y, Liu S, Yin P, et al. Can China achieve a one-third reduction in premature mortality from non-communicable diseases by 2030? BMC Med 2017 Jul 11;15(1):132 [FREE Full text] [doi: 10.1186/s12916-017-0894-5] [Medline: 28693510]
18. Roth GA, Nguyen G, Forouzanfar MH, Mokdad AH, Naghavi M, Murray CJ. Estimates of global and regional premature cardiovascular mortality in 2025. Circulation 2015 Sep 29;132(13):1270-1282. [doi: 10.1161/CIRCULATIONAHA.115.016021] [Medline: 26408271]
19. Diet, activity and cancer. World Cancer Research Fund. URL: http://www.wcrf.org/dietandcancerreport [accessed 2023-02-17]
20. Xiao Y. Hunan cancer registry annual report. Changsha: Central South University Press; 2018.
21. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016 Jan 25;66(2):115-132 [FREE Full text] [doi: 10.3322/caac.21338] [Medline: 26808342]
22. Li Y, Wang L, Jiang Y, Zhang M, Wang L. Risk factors for noncommunicable chronic diseases in women in China: surveillance efforts. Bull World Health Organ 2013 Sep 01;91(9):650-660. [doi: 10.2471/blt.13.117549]
23. WHO global air quality guidelines: particulate matter (PM2.5 and PM10) , ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. URL: https://www.who.int/publications/i/item/9789240034228?ua=1 [accessed 2023-02-17]
24. Global status report on noncommunicable diseases 2014. World Health Organization. URL: https://www.who.int/publications/i/item/9789241564854 [accessed 2023-02-17]
25. Ezzati M, Lopez AD, Rodgers AA, Murray CJL. Comparative quantification of health risks : global and regional burden of disease attributable to selected major risk factors. World Health Organization. URL: https://apps.who.int/iris/handle/10665/42770 [accessed 2023-02-17]
26. Chen Y, Guo F, Wang J, Cai W, Wang C, Wang K. Provincial and gridded population projection for China under shared socioeconomic pathways from 2010 to 2100. Sci Data 2020 Mar 09;7(1):83. [doi: 10.1038/s41597-020-0421-y] [Medline: 32152299]
27. Hu T, Lee AH, Mao Z. WHO Framework Convention on Tobacco Control in China: barriers, challenges and recommendations. Glob Health Promot 2013 Dec;20(4):13-22 [FREE Full text] [doi: 10.1177/1757975913501910] [Medline: 24297769]
28. China reported health hazards of smoking 2020. National Health Commission of the PRC. URL: http://www.gov.cn/xinwen/2021-05/30/content_5613994.htm [accessed 2023-02-17]
29. Wang M, Luo X, Xu S, Liu W, Ding F, Zhang X, et al. Trends in smoking prevalence and implication for chronic diseases in China: serial national cross-sectional surveys from 2003 to 2013. Lancet Respir Med 2019 Jan;7(1):35-45. [doi: 10.1016/S2213-2600(18)30432-6] [Medline: 30482646]
30. Manthey J, Shield K, Rylett M, Hasan O, Probst C, Rehm J. Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. Lancet 2019 Jun 22;393(10190):2493-2502. [doi: 10.1016/S0140-6736(18)32744-2] [Medline: 31076174]
31. Jiang Y, Liu S, Ji N, Zeng X, Liu Y, Zhang M, et al. Deaths attributable to alcohol use and its impact on life expectancy in China, 2013. Zhonghua Liu Xing Bing Xue Za Zhi 2018 Jan 10;39(1):27-31. [doi: 10.3760/cma.j.issn.0254-6450.2018.01.005] [Medline: 29374890]
32. Xu Q, Zhou M, Jin D, Zeng X, Qi J, Yin L, et al. Projection of premature mortality from noncommunicable diseases for 2025: a model based study from Hunan Province, China, 1990-2016. PeerJ 2020 Nov 03;8:e10298 [FREE Full text] [doi: 10.7717/peerj.10298] [Medline: 33194444]
33. Tan M, He FJ, Wang C, MacGregor GA. Twenty-four-hour urinary sodium and potassium excretion in China: a systematic review and meta-analysis. J Am Heart Assoc 2019 Jul 16;8(14):e012923 [FREE Full text] [doi: 10.1161/JAHA.119.012923] [Medline: 31295409]
34. Hyseni L, Elliot-Green A, Lloyd-Williams F, Kypridemos C, O'Flaherty M, McGill R, et al. Systematic review of dietary salt reduction policies: evidence for an effectiveness hierarchy? PLoS One 2017 May 18;12(5):e0177535 [FREE Full text] [doi: 10.1371/journal.pone.0177535] [Medline: 28542317]
35. National Health Commission of the People's Republic of China. Report on Chinese residents' chronic diseases and nutrition 2020. Acta Nutrimenta Sinica 2020 Dec 31;42(6):521.
36. Hu C, Jia W. Diabetes in China: epidemiology and genetic risk factors and their clinical utility in personalized medication. Diabetes 2018 Jan;67(1):3-11. [doi: 10.2337/dbi17-0013] [Medline: 29263166]
37. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation

Diabetes Atlas, 9 edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
38. Blanc-Lapierre A, Spence A, Karakiewicz PI, Aprikian A, Saad F, Parent M. Metabolic syndrome and prostate cancer risk in a population-based case-control study in Montreal, Canada. BMC Public Health 2015 Sep 18;15(1):913 [FREE Full text] [doi: 10.1186/s12889-015-2260-x] [Medline: 26385727]
39. Klement RJ, Kämmerer U. Is there a role for carbohydrate restriction in the treatment and prevention of cancer? Nutr Metab 2011 Oct 26;8(1):75 [FREE Full text] [doi: 10.1186/1743-7075-8-75] [Medline: 22029671]
40. Huang Y, Cai X, Qiu M, Chen P, Tang H, Hu Y, et al. Prediabetes and the risk of cancer: a meta-analysis. Diabetologia 2014 Nov 11;57(11):2261-2269. [doi: 10.1007/s00125-014-3361-2] [Medline: 25208757]
41. Shi R, Guo X, Zhang Q. Current status and outlook of diabetes self-management education and support in Chinese adults with type 2 diabetes. Chin J Diabetes Mellitus 2021 Feb 27;13(2):121-124. [doi: 10.3760/cma.j.cn115791-20201012-00610]
42. Cai C, Jia W. Community healthcare for diabetes in China. Sci Sin Vitae 2018 Aug 7;48(8):820-826. [doi: 10.1360/n052018-00048]
43. Wang Y, Zhao L, Gao L, Pan A, Xue H. Health policy and public health implications of obesity in China. Lancet Diabetes Endocrinol 2021 Jul;9(7):446-461. [doi: 10.1016/S2213-8587(21)00118-2] [Medline: 34097869]
44. Fu Z, Liu J, Liu H, Jin D. Dietary patterns of urban residents from 1982 to 2012 in Hunan. Zhong Nan Da Xue Xue Bao Yi Xue Ban 2014 Jul;39(7):713-717. [doi: 10.11817/j.issn.1672-7347.2014.07.011] [Medline: 25080917]
45. Xiao Y, Su C, Ouyang Y, Zhang B. Trends of vegetables and fruits consumption among Chinese adults aged 18 to 44 years old from 1991 to 2011. Zhonghua Liu Xing Bing Xue Za Zhi 2015 Mar;36(3):232-236. [Medline: 25975399]
46. Whiteman DC, Wilson LF. The fractions of cancer attributable to modifiable factors: a global review. Cancer Epidemiol 2016 Oct;44:203-221. [doi: 10.1016/j.canep.2016.06.013] [Medline: 27460784]
47. Di Sebastiano KM, Murthy G, Campbell KL, Desroches S, Murphy RA. Nutrition and cancer prevention: why is the evidence lost in translation? Adv Nutr 2019 May 01;10(3):410-418 [FREE Full text] [doi: 10.1093/advances/nmy089] [Medline: 30915435]
48. Cao W, Chen H, Yu Y, Li N, Cheng WQ. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. Chinese Med J 2021 Apr 05;134(7):783-791. [doi: 10.1097/cm9.0000000000001474]

Abbreviations

AAPC: average annual percentage change
BAU: business as usual
CCDRFS: Chinese Chronic Disease and Risk Factor Surveillance
CUP: Continuous Update Project
DALY: disability-adjusted life year
GBD: global burden of disease
NCD: noncommunicable disease
NPCR: National Program of Cancer Registries
PAF: population-attributable fraction
PM2.5: fine particulate matter
RR: relative risk
SDG: Sustainable Development Goal
UN: United Nations
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 01.11.22; peer-reviewed by L Du, W Chen; comments to author 21.12.22; revised version received 21.01.23; accepted 30.01.23; published 06.03.23.

Please cite as:
Wu W, Wang J, Liao XZ, Xu K, Zou Y, Shi Z, Hu Y, Xiao H, Li C, Cao S, Wang S, Guo J, Luo Z, Liu M, Xu M, Jin D, Chen M, Fu Z, Yan S
Projection of Premature Cancer Mortality in Hunan, China, Through 2030: Modeling Study
JMIR Public Health Surveill 2023;9:e43967
URL: https://publichealth.jmir.org/2023/1/e43967
doi:10.2196/43967
PMID:36877566

Original Paper

Disparity in Lung Cancer Screening Among Smokers and Nonsmokers in China: Prospective Cohort Study

Le Wang[1*], PhD; Youqing Wang[1,2*], MS; Fei Wang[3,4*], PhD; Yumeng Gao[5], BMed; Zhimei Fang[6]; Weiwei Gong[7], MS; Huizhang Li[1], MS; Chen Zhu[1], MS; Yaoyao Chen[1], BMed; Lei Shi[8], MD; Lingbin Du[1,2], MPH; Ni Li[3,4], PhD

[1]Department of Cancer Prevention, Zhejiang Cancer Hospital, Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China

[2]Zhejiang Key Laboratory of Diagnosis & Treatment Technology on Thoracic Oncology (Lung and Esophagus), Hangzhou, China

[3]Office of Cancer Screening, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[4]Chinese Academy of Medical Sciences Key Laboratory for National Cancer Big Data Analysis and Implement, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[5]Hwa Mei Hospital, University of Chinese Academy of Sciences, Ningbo, China

[6]Kecheng District People's Hospital of Quzhou, Quzhou, China

[7]Zhejiang Provincial Centre for Disease Control and Prevention, Hangzhou, China

[8]Department of Radiology, Zhejiang Cancer Hospital, Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Lingbin Du, MPH
Department of Cancer Prevention
Zhejiang Cancer Hospital, Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences
No.1 East Banshan Road, Gongshu District
Hangzhou, 310022
China
Phone: 86 571 88122219
Fax: 86 571 88122219
Email: dulb@zjcc.org.cn

Abstract

Background: Low-dose computed tomography (LDCT) screening is effective in reducing lung cancer mortality in smokers; however, the evidence in nonsmokers is scarce.

Objective: This study aimed to evaluate the participant rate and effectiveness of one-off LDCT screening for lung cancer among smokers and nonsmokers.

Methods: A population-based prospective cohort study was performed to enroll participants aged between 40 and 74 years from 2013 to 2019 from 4 cities in Zhejiang Province, China. Participants who were evaluated as having a high risk of lung cancer from an established risk score model were recommended to undergo LDCT screening. Follow-up outcomes were retrieved on June 30, 2020. The uptake rate of LDCT screening for evaluated high-risk participants and the detection rate of early-stage lung cancer (stage 0-I) were calculated. The lung cancer incidence, lung cancer mortality, and all-cause mortality were compared between the screened and nonscreened groups.

Results: At baseline, 62.56% (18,818/30,079) of smokers and 6% (5483/91,455) of nonsmokers were identified as high risk (*P*<.001), of whom 41.9% (7885/18,818) and 66.31% (3636/5483) underwent LDCT screening (*P*<.001), respectively. After a median follow-up of 5.1 years, 1100 lung cancer cases and 456 all-cause death cases (116 lung cancer death cases) were traced. The proportion of early-stage lung cancer among smokers was 60.3% (173/287), which was lower than the proportion of 80.3% (476/593) among nonsmokers (*P*<.001). Among smokers, a higher proportion was found in the screened group (72/106, 67.9%) than the nonscreened group (56/114, 49.1%; *P*=.005), whereas no significance was found (42/44, 96% vs 10/12, 83%; *P*=.20) among nonsmokers. Compared with participants who were not screened, LDCT screening in smokers significantly increased lung cancer incidence (hazard ratio [HR] 1.39, 95% CI 1.09-1.76; *P*=.007) but reduced lung cancer mortality (HR 0.52, 95% CI 0.28-0.96; *P*=.04) and all-cause mortality (HR 0.47, 95% CI 0.32-0.69; *P*<.001). Among nonsmokers, no significant results were found for lung cancer incidence (*P*=.06), all-cause mortality (*P*=.89), and lung cancer mortality (*P*=.17).

Conclusions: LDCT screening effectively reduces lung cancer and all-cause mortality among high-risk smokers. Further efforts to define high-risk populations and explore adequate lung cancer screening modalities for nonsmokers are needed.

(JMIR Public Health Surveill 2023;9:e43586) doi:10.2196/43586

KEYWORDS

lung cancer; screening; smoker; nonsmoker

Introduction

Background

Lung cancer is the second most commonly diagnosed cancer and the leading cause of cancer-related death worldwide; in 2020, there were an estimated 2.2 million new cases and 1.8 million deaths [1]. In 2019, more than one-third of all newly diagnosed lung cancers and approximately 40% of global cancer-related deaths occurred in China [1], where lung cancer was the leading cause of disability-adjusted life years and years of life lost [2], resulting in a high disease and socioeconomic burden. These trends underscore the need for continued efforts to improve lung cancer outcomes.

Screening with low-dose computed tomography (LDCT) is effective in reducing lung cancer mortality in rigorously randomized controlled trials such as the National Lung Screening Trial (NLST) and the Dutch-Belgian lung cancer screening trial (Nederlands Leuvens Screening Onderzoek [NELSON]); however, such trials have high LDCT uptake rates exceeding 90% [3,4]. These high uptake rates are not attainable in real-world settings; thus, evidence from real-world screening settings with imperfect uptake rates is crucial to inform appropriate and effective screening policies. Moreover, current findings are only applicable to smokers, who are considered as a high-risk lung cancer subgroup. The prevalence of nonsmoking-related lung cancer in East Asian countries is higher than that in Europe and the United States [5]. In China, tobacco smoking is responsible for 75% of lung cancer cases in men and 18% in women [6]. Therefore, the effectiveness of LDCT screening among nonsmokers at high risk of lung cancer still needs to be evaluated.

Objectives

The Cancer Screening Program in Urban China (a large public health service project) conducted in 2012 included both smokers and nonsmokers and targeted 5 types of cancer (lung cancer, female breast cancer, esophageal and gastric cancer, colorectal cancer, and liver cancer) [7]. In Zhejiang, the smoking rate of people aged ≥15 years was over 20%, which is relatively low among all provinces in China [8]. Nevertheless, Zhejiang Province has a higher lung cancer incidence than the national average [9,10]. These phenomena make Zhejiang an appropriate setting to study the effectiveness of lung cancer screening among smokers and nonsmokers. Using lung cancer screening data from the Cancer Screening Program in Urban China conducted in the first 7 years between October 2013 and September 2019 in Zhejiang, we evaluated the effectiveness of one-off LDCT screening in the early detection and reduction of lung cancer mortality and all-cause mortality among smokers and nonsmokers, respectively.

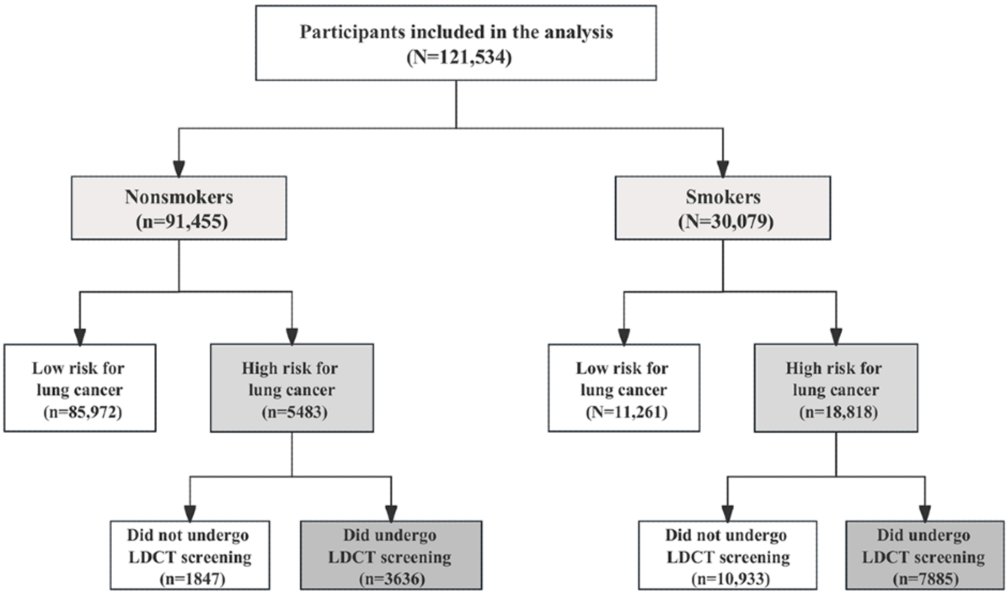
Methods

Study Design and Participants

We performed a population-based prospective study using the framework of the Cancer Screening Program in Urban China [11]. Trained staff called or visited residents aged 40-74 years living in selected communities of the participating cities. Patients with a history of cancer and those undergoing treatment for other serious medical and surgical diseases were excluded. For this screening program, only participants at high risk of lung cancer (refer to Risk Assessment Procedure) were recommended to undergo LDCT free of charge at a tertiary-level hospital designated by the program. All participants provided written informed consent.

We used data from lung cancer screening conducted between October 2013 and September 2019 in Zhejiang Province, which covered 4 cities (Hangzhou, Ningbo, Quzhou, and Jinhua). Overall, 121,534 eligible individuals participated in the lung cancer screening program (Figure 1). For this study, smokers were defined as those who had previously smoked or were currently smoking tobacco more than once per day for at least 6 months. The smokers and nonsmokers enrolled in this study were classified into 3 groups: the screened group (high-risk participants who underwent LDCT screening), the nonscreened group (high-risk participants who did not undergo LDCT screening), and the low-risk group.

Figure 1. Flow diagram of the study population. LDCT: low-dose computed tomography.



Ethics Approval

The study was approved by the ethics committee of the Chinese Academy of Medical Sciences's Cancer Hospital (approval number 15-070/997) and Zhejiang Cancer Hospital (approval number IRB-2022-271).

Risk Assessment Procedure

Eligible participants completed a cancer-related risk assessment questionnaire designed by the Cancer Screening Program in Urban China before LDCT that included questions regarding cigarette smoking history, occupational exposure to hazardous substances, frequent exercise, chronic respiratory diseases, family history of lung cancer, dietary intake of fresh vegetables in the previous year, and passive smoking. We adopted the sex-specific risk score systems derived from the Harvard Cancer Risk Index to evaluate the risk of lung cancer [12,13]. Each risk factor was assigned a score by an expert panel based on the magnitude of its association with lung cancer. Cumulative risk scores were calculated and divided by the average risk score in the general population, yielding final individual relative risks. Individuals with a relative risk >2.0 or aged ≥50 years with a smoking index of ≥400 (number of cigarettes smoked per day multiplied by years of smoking) were defined as being at high risk for lung cancer [12,13].

Individuals labeled as being at high risk for lung cancer were recommended to undergo a free LDCT scan with ≥16 slices at a tertiary-level hospital designated by the program. These participants then underwent a process of shared decision-making that included information about the potential benefits and harms of screening with LDCT to ensure that their decisions on LDCT scans were based on their free will.

Follow-up Data

To account for potential immortal time bias, where individuals in the screened group had to survive (be alive and event free) until the LDCT scan was conducted, the cohort entry date was defined as the date of screening in the screened group. For individuals in the nonscreened group, the cohort entry date was estimated based on the screening date of the individual in the screened group whose risk assessment date was closest to that of the nonscreened group. Time to lung cancer occurrence was calculated from the cohort entry date until the earliest occurrence of lung cancer, death, or administrative censoring (June 30, 2020). Accordingly, time to lung cancer death or all-cause death was calculated from the cohort entry date until death or administrative censoring, whichever occurred first.

Outcome Assessment and Quality Control

The primary outcomes of interest were incidence of lung cancer, lung cancer mortality, and all-cause mortality. The secondary outcomes were the proportion of early-stage lung cancer (stage 0-I) and the participation rate of LDCT. Lung cancer was defined according to the International Classification of Diseases (10th revision) and was coded as C34. Outcome data were retrieved from national linkages, including the cancer registry system and death surveillance system, every 6 months.

Paper-based standardized documentation forms (epidemiological questionnaire and LDCT report) were collected from trained staff and physicians. Form validity was checked and entered into the data management system by trained study staff. A consistency check was performed, and if inconsistencies were identified, errors were corrected by retrieving the original records. Each participant had a unique identification number that was used to track all individual-related documentation forms. All data were transmitted to the Central Data Management Team of the National Cancer Center of China, where databases were constructed and analyzed.

Definitions of Covariates

Covariates from the baseline survey included demographic characteristics (age: 40-54 years and 55-74 years; sex; education level: low [primary school or below], medium [primary school to high school], and high [high school or above]; and BMI), lifestyle factors (smoking status, passive smoking, occupational exposure to hazardous substances, and frequent exercise), family history of lung cancer, and baseline comorbidities (chronic respiratory diseases, digestive diseases, hepatobiliary diseases, hypertension, diabetes, and hyperlipidemia). Passive smoking referred to involuntary inhalation of tobacco smoke. Occupational exposure to hazardous substances included occupational exposure to asbestos, rubber, dust, pesticides, radiation, beryllium, uranium, and radon for at least 1 year. Frequent exercises were defined as exercises conducted at least 3 times per week for ≥30 minutes each. Respiratory diseases included pulmonary tuberculosis, chronic bronchitis, emphysema, asthmatic bronchiectasis, silicosis, and pneumoconiosis.

Statistical Analysis

Baseline study population characteristics were summarized using frequencies and percentages for categorical variables and arithmetic means and SDs for continuous variables. Baseline factors were compared between the nonsmoker and smoker groups using a 2-tailed Student *t* test or chi-square test. Group-specific participation rates of LDCT screening for high-risk populations by common factors were also calculated. Multivariable logistic regression was used to explore the potential determinants of the participant rate of LDCT screening. For lung cancer outcomes, cumulative incidence was estimated using the cumulative incidence function, accounting for the competing risk of mortality. Lung cancer mortality was estimated using the cumulative mortality function, accounting for the competing risk of death from other causes. Gray tests were used to assess differences between the groups for cumulative lung cancer incidence and lung cancer mortality. All-cause mortality was calculated using Kaplan-Meier analysis with a log-rank test to assess differences between groups. Owing to the small number of deaths among nonsmokers, we only used a Cox proportional hazards model among smokers to calculate the hazard ratios (HRs) and 95% CI of LDCT screening with lung cancer incidence, mortality, and all-cause mortality. All hypothesis tests were 2 sided. Statistical analyses were performed using R (version 3.5.1; R Foundation for Statistical Computing) and SAS (version 9.4; SAS Institute Inc), and *P*<.05 was considered statistically significant.

Results

Baseline Characteristics

In total, 121,534 participants aged 40-74 years who were enrolled in the program were included in this study; of them, 75.3% (n=91,453) were nonsmokers and 24.7% (n=30,079) were smokers. Compared with nonsmokers, smokers were older; were more likely to be male; were more likely to have medium and high educational levels; were more overweight and obese; and had more occupational exposure to hazardous substances, passive smoking, family history of lung cancer, chronic respiratory diseases, digestive diseases, hepatobiliary diseases, hypertension, hyperlipidemia, and diabetes but were less likely to exercise frequently (*P*<.001; Table 1). Comparisons of baseline characteristics by risk assessment and screening status for smokers and nonsmokers are, respectively, presented in Multimedia Appendices 1 and 2.

Table 1. Baseline characteristics of the study population (N=121,534).

Characteristics	Overall[a] (N=121,534)	Nonsmoker group[b] (n=91,455)	Smoker group[c] (n=30,079)	*P* value[d]
Demographic characteristics				
Age (years), mean (SD)	56.5 (8.3)	56.4 (8.4)	56.8 (8.0)	<.001
40-54, n (%)	49,948 (41.10)	38,166 (41.73)	11,782 (39.17)	
55-74, n (%)	71,586 (58.90)	53,289 (58.27)	18,297 (60.83)	
Sex, n (%)				<.001
Male	51,623 (42.48)	22,499 (24.60)	29,124 (96.83)	
Female	69,911 (57.52)	68,956 (75.40)	955 (3.17)	
Education, n (%)				<.001
Low	36,691 (30.19)	28,925 (31.63)	7766 (25.82)	
Medium	70,489 (58.00)	51,977 (56.83)	18,512 (61.54)	
High	14,354 (11.81)	10,553 (11.54)	3801 (12.64)	
BMI (kg/m^2), n (%)				<.001
<18.5	3294 (2.71)	2645 (2.90)	649 (2.16)	
18. 5-24	68,948 (56.82)	52,997 (58.05)	15,951 (53.12)	
24-28	40,842 (33.66)	29,630 (32.45)	11,212 (37.33)	
≥28	8250 (6.80)	6031 (6.61)	2219 (7.39)	
Lifestyle factors				
Occupational exposure to hazardous substances, n (%)				<.001
No	106,361 (87.52)	82,044 (89.71)	24,317 (80.84)	
Yes	15,173 (12.48)	9411 (10.29)	5762 (19.16)	
Passive smoking, n (%)				<.001
No	74,087 (61.11)	63,153 (69.15)	10,934 (36.55)	
Yes	47,158 (38.89)	28,173 (30.85)	18,985 (63.45)	
Frequent exercise, n (%)				<.001
No	64,019 (52.68)	45,355 (49.59)	18,664 (62.05)	
Yes	57,515 (47.32)	46,100 (50.41)	11,415 (37.95)	
Family history of lung cancer, n (%)				<.001
No	98,370 (88.69)	74,851 (90.17)	23,519 (84.30)	
Yes	12,539 (11.31)	8159 (9.83)	4380 (15.70)	
Baseline comorbidity				
Chronic respiratory diseases, n (%)				<.001
No	105,151 (86.52)	81,371 (88.97)	23,780 (79.06)	
Yes	16,383 (13.48)	10,084 (11.03)	6299 (20.94)	
Digestive diseases, n (%)				<.001
No	90,832 (74.74)	70,115 (76.71)	20,717 (68.88)	
Yes	30,702 (25.26)	21,340 (23.29)	9362 (31.12)	
Hepatobiliary diseases, n (%)				<.001
No	90,549 (74.51)	70,159 (76.71)	20,390 (67.79)	
Yes	30,985 (25.49)	21,296 (23.29)	9689 (32.21)	
Hypertension, n (%)				<.001
No	78,772 (71.20)	60,271 (72.72)	18,501 (66.66)	

Characteristics	Overall[a] (N=121,534)	Nonsmoker group[b] (n=91,455)	Smoker group[c] (n=30,079)	*P* value[d]
Yes	31,858 (28.80)	22,605 (27.28)	9253 (33.34)	
Hyperlipidemia, n (%)				<.001
No	91,603 (82.81)	69,381 (83.72)	22,222 (80.07)	
Yes	19,018 (17.19)	13,487 (16.28)	5531 (19.93)	
Diabetes, n (%)				<.001
No	101,246 (91.52)	76,417 (92.21)	24,829 (89.46)	
Yes	9379 (8.48)	6455 (7.79)	2924 (10.54)	

[a] Overall: 200 participants without information on BMI, 289 participants without information on passive smoking, 10,625 participants without information on family history of lung cancer, 10,904 participants without information on hypertension, 10,913 participants without information on hyperlipidemia, and 10,909 participants without information on diabetes.

[b] Nonsmokers: 152 participants without information on BMI, 129 participants without information on passive smoking, 8445 participants without information on family history of lung cancer, 8579 participants without information on hypertension, 8587 participants without information on hyperlipidemia, and 8583 participants without information on diabetes.

[c] Smokers: 48 participants without information on BMI, 160 participants without information on passive smoking, 2180 participants without information on family history of lung cancer, 2325 participants without information on hypertension, 2326 participants without information on hyperlipidemia, and 2326 participants without information on diabetes.

[d] *P* values were generated by using the chi-square or *t* test (2-tailed) by smoking status.

Participation Rates for LDCT Screening for High-risk Participants

Among smokers, 62.56% (18,818/30,079) were identified as being at high risk for lung cancer, which was significantly higher than 5.8% (5483/94,455) among nonsmokers (*P*<.001). Among the high-risk participants, 41.9% (7885/18,818) underwent LDCT screening among smokers, which was significantly lower than 66.31% (3636/5483) among nonsmokers (*P*<.001; Figure 1). The data on LDCT participation by subgroup are shown in Table 2. Results from multivariable logistic regression models showed that female participants (odds ratio [OR] 1.33, 95% CI 1.11-1.58; *P*<.001), older participants (OR 1.24, 95% CI 1.16-1.33; *P*<.001), participants with occupational exposure to hazardous substances (OR 1.50, 95% CI 1.38-1.62; *P*<.001), participants with a family history of lung cancer (OR 1.74, 95% CI 1.60-1.90; *P*<.001), participants with chronic respiratory diseases (OR 1.40, 95% CI 1.29-1.51; *P*<.001), participants with digestive diseases (OR 1.24, 95% CI 1.16-1.33; *P*<.001), participants with hepatobiliary diseases (OR 1.35, 95% CI 1.25-1.45; *P*<.001), participants with hyperlipidemia (OR 1.19, 95% CI 1.10-1.30; *P*<.001), and participants with diabetes (OR 1.16, 95% CI 1.04-1.29; *P*<.001) had higher LDCT participation rates among smokers. Among nonsmokers, participants with occupational exposure to hazardous substances (OR 1.37, 95% CI 1.19-1.58; *P*<.001), participants with a history of passive smoking (OR 1.22, 95% CI 1.04-1.44; *P*<.001), participants with a family history of lung cancer (OR 1.83, 95% CI 1.62-2.08; *P*<.001), and participants with hepatobiliary diseases (OR 1.37, 95% CI 1.21-1.56; *P*<.001) had higher LDCT participation rates.

Table 2. Participation rates of low-dose computed tomography (LDCT) screening for high-risk population by smoking status (n =24,301).

Characteristics	Nonsmoker group[a]				Smoker group[b]			
	Participants at high risk, n	Underwent LDCT, n (%)	*P* value[c]	OR[d] (95% CI)	Participants at high risk, n	Underwent LDCT, n (%)	*P* value[c]	OR[d] (95% CI)
Demographic characteristics								
Age (years)			.005				<.001	
40-54	2374	1623 (68.37)		—[e]	7217	2779 (38.51)		Reference
55-74	3109	2013 (64.75)		—	11,601	5106 (44.01)		1.24 (1.16-1.33)
Sex[f]			—				<.001	
Male	0	0 (0)		—	18,178	7524 (41.39)		Reference
Female	5483	3636 (66.31)		—	640	361 (56.41)		1.33 (1.11-1.58)
Education			.65				.48	
Low	1889	1261 (66.75)		—	5304	2247 (42.36)		—
Medium	3054	2026 (66.34)		—	11,540	4833 (41.88)		—
High	540	349 (64.63)		—	1974	805 (40.78)		—
BMI			.42				<.001	
<18.5	187	114 (60.96)		—	445	174 (39.10)		—
18.5-24	3156	2107 (66.76)		—	10,066	4089 (40.62)		—
24-28	1736	1153 (66.42)		—	6882	2978 (43.27)		—
≥28	393	257 (65.39)		—	1393	634 (45.51)		—
Lifestyle factors								
Occupational exposure to hazardous substances			<.001				<.001	
No	3716	2338 (62.92)		Reference	14,712	5668 (38.53)		Reference
Yes	1767	1298 (73.46)		1.37 (1.19-1.58)	4106	2217 (53.99)		1.50 (1.38-1.62)
Passive smoking			.001				<.001	
No	949	586 (61.75)		Reference	5735	2140 (37.31)		—
Yes	4463	2998 (67.17)		1.22 (1.04-1.44)	12,935	5649 (43.67)		—
Frequent exercise			.76				.41	
No	4205	2784 (66.21)		—	13,486	5676 (42.09)		—
Yes	1278	852 (66.67)		—	5332	2209 (41.43)		—
Family history of lung cancer			<.001				<.001	
No	2098	1238 (59.01)		Reference	13,583	5213 (38.38)		Reference
Yes	3200	2305 (72.03)		1.83 (1.62-2.08)	3764	2255 (59.91)		1.74 (1.60-1.90)
Baseline comorbidity								
Chronic respiratory diseases			.12				<.001	
No	333	234 (70.27)		—	13,440	4875 (36.27)		Reference
Yes	5150	3402 (66.06)		—	5378	3010 (55.97)		1.40 (1.29-1.51)
Digestive diseases			.03				<.001	
No	2158	1394 (64.60)		—	12,246	4519 (36.90)		Reference
Yes	3325	2242 (67.40)		—	6572	3366 (51.22)		1.24 (1.16-1.33)
Hepatobiliary diseases			<.001				<.001	
No	2233	1355 (60.7)		Reference	12,186	4361 (35.79)		Reference
Yes	3250	2281 (70.2)		1.37 (1.21-1.56)	6632	3524 (53.14)		1.35 (1.25-1.45)
Hypertension			.80				<.001	

Characteristics	Nonsmoker group[a]				Smoker group[b]			
	Participants at high risk, n	Underwent LDCT, n (%)	*P* value[c]	OR[d] (95% CI)	Participants at high risk, n	Underwent LDCT, n (%)	*P* value[c]	OR[d] (95% CI)
No	3254	2162 (66.44)		—	11,319	4630 (40.90)		—
Yes	1506	995 (66.07)		—	5736	2612 (45.54)		—
Hyperlipidemia			—				<.001	
No	3174	2068 (65.15)		—	13,458	5367 (39.88)		Reference
Yes	1583	1086 (68.60)		—	3596	1873 (52.09)		1.19 (1.10-1.30)
Diabetes			.21				<.001	
No	4301	2864 (66.59)		—	15,225	6356 (41.75)		Reference
Yes	457	291 (63.68)		—	1829	885 (48.39)		1.16 (1.04-1.29)

[a] Nonsmokers: 11 participants without information on BMI, 71 participants without information on passive smoking, 185 participants without information on family history of lung cancer, 723 participants without information on hypertension, 726 participants without information on hyperlipidemia, and 725 participants without information on diabetes.

[b] Smokers: 32 participants without information on BMI, 148 participants without information on passive smoking, 1471 participants without information on family history of lung cancer, 1763 participants without information on hypertension, 1764 participants without information on hyperlipidemia, and 1764 participants without information on diabetes.

[c] *P* values were generated by using the chi-square and Cochran-Armitage test statistics for differences where appropriate; missing values are not included.

[d] OR: odds ratio. ORs are presented for variables with significance in the multivariate logistic regression.

[e] The variable is excluded from the final model and OR is not available.

[f] All nonsmokers were female .

Overall Lung Cancer Incidence and Mortality

After a median follow-up time of 5.1 years (IQR 3.1-5.9 years), 377 lung cancer cases, 202 all-cause death cases, and 67 lung cancer death cases were observed among smokers, and 733 lung cancer cases, 254 all-cause death cases, and 49 lung cancer death cases were observed among nonsmokers (Table 3). The crude lung cancer incidence density was 277.57 (95% CI 250.92-307.05) per 100,000 person-years in smokers and 178.51 (95% CI 166.05-191.92) per 100,000 person-years in nonsmokers (Table 4), resulting in a crude rate ratio of 1.555 (95% CI 1.547-1.563). The crude all-cause mortality rate and lung cancer mortality rate in smokers was 147.89 (95% CI 128.83-169.76) per 100,000 person-years and 49.05 (95% CI 38.61-62.33) per 100,000 person-years, respectively, which were higher than 61.63 (95% CI 54.50-69.69) per 100,000 person-years and 11.89 (95% CI 8.99-15.73) per 100,000 person-years in nonsmokers, yielding a crude rate ratio of 2.4 (95% CI 2.382-2.417) and 4.126 (95% CI 4.057-4.195), respectively. Among lung cancer patients with available data on stage and histological information, the proportion of stage 0-I was 60.3% (173/287) among smokers and lower than the proportion of 80.3% (476/593) among nonsmokers (*P*<.001), and the proportion of adenocarcinoma was 61.2% (194/317) among smokers and lower than the proportion of 90.2% (607/673) among nonsmokers (*P*<.001; Table 3). Subgroup analyses showed that the proportion of stage 0-I was 67.9% (72/106) in the screened group, which was significantly higher than 49.1% (56/114) in the nonscreened group (*P*=.005) among smokers, and no significance was found of the proportion of stage 0-Ibetween the screened group (42/44, 96%) and the nonscreened group (10/12, 83%) among nonsmokers (*P*=.20).

Table 3. Lung cancer incidence cases and deaths among all participants (N=121,534).

	Nonsmoker[a]				Smoker[b]			
	All (n=91,455)	Low risk (n=85,972)	Nonscreened (n=1847)	Screened (n=3636)	All (n=30,379)	Low risk (n=11,261)	Nonscreened (n=10,933)	Screened (n=7885)
Incident cases, n								
All cases	733	662	17	54	377	86	148	143
Stage , n (%)								
0-I	476 (80.3)	424 (79)	10 (83.3)	42 (95.5)	173 (60.3)	45 (67.2)	56 (49.1)	72 (67.9)
II	24 (4.1)	23 (4.2)	1 (8.3)	0 (0)	19 (6.6)	3 (4.5)	8 (7)	8 (7)
III	32 (5.4)	31 (5.8)	1 (8.3)	0 (0)	47 (16.4)	8 (11.9)	22 (19.3)	17 (15)
IV	61 (10.3)	59 (11)	0 (0)	2 (4.5)	48 (16.7)	11 (16.4)	28 (24.6)	9 (8.5)
Histological type, n (%)								
Adenocarcinoma	607 (90.2)	546 (89.1)	16 (94.1)	45 (93.8)	194 (61.2)	54 (71.1)	59 (48.4)	81 (68.1)
Squamous cell carcinoma	34 (5.1)	33 (5.4)	0 (0)	1 (2.1)	79 (24.9)	16 (21.1)	39 (32)	24 (20.2)
Small-cell carcinoma	13 (1.9)	13 (2.1)	0 (0)	0 (0)	31 (9.8)	4 (5.3)	18 (14.8)	8 (7.6)
Others	19 (2.8)	21 (3.4)	1 (5.9)	2 (4.2)	13 (4.1)	2 (2.6)	6 (4.9)	5 (4.2)
Death cases, n								
All-cause death cases	254	246	3	5	202	64	102	36
Lung cancer death cases	49	48	1	0	67	13	40	14

[a] Nonsmokers: 140 lung cancer cases without information on stage and 60 lung cancer cases without information on histological type.

[b] Smokers: 90 lung cancer cases without information on stage and 60 lung cancer cases without information on the histological type.

Table 4. Lung cancer incidence density and mortality rate by subgroups.

	Nonsmoker				Smoker			
	Cases, n	Crude rate[a] (95% CI)	Adjusted rate[a,b] (95% CI)	*P* value[c]	Cases, n	Crude rate[a] (95% CI)	Adjusted rate[a,b] (95% CI)	*P* value
Lung cancer incidence density								
Overall	733	178.51 (166.05-191.92)	179.78 (166.14-194.54)	N/A[d]	377	277.57 (250.92-307.05)	219.86 (157.77-306.40)	N/A
Subgroup				<.001[c]				<.001[c]
Low-risk group	662	170.23 (157.74-183.70)	173.28 (159.64-188.08)	N/A	86	157.80 (127.74-194.94)	144.60 (67.76-308.57)	N/A
Nonscreened group	17	223.72 (139.08-359.88)	129.32 (80.39-208.03)	.06[e]	148	306.10 (260.55-359.61)	132.65 (112.91-155.84)	<.001
Screened group	54	382.33 (292.82-499.21)	223.37 (171.03-291.72)	N/A	143	433.71 (368.14-510.96)	389.49 (249.18-608.81)	N/A
All-cause mortality rate								
Overall	254	61.63 (54.50-69.69)	72.62 (63.54-83.01)	N/A	202	147.89 (128.84-169.76)	74.83 (54.98-103.73)	N/A
Subgroup				.29[c]				<.001[c]
Low-risk group	246	63.04 (55.63-71.43)	73.09 (63.92-83.58)	N/A	64	117.10 (91.65-149.61)	84.83 (37.31-192.89)	N/A
Nonscreened group	3	39.29 (12.67-121.81)	22.56 (7.28-69.95)	.89[e]	102	210.0 (172.9-255.0)	91.03 (74.97-110.53)	<.001[e]
Screened group	5	35.02 (14.57-84.13)	20.74 (8.63-49.86)	N/A	36	107.93 (77.86-149.63)	46.03 (33.16-63.88)	N/A
Lung cancer mortality rate								
Overall	49	11.89 (8.99-15.73)	15.98 (11.85-21.55)	N/A	67	49.05 (38.61-62.33)	32.4 (15.75-66.64)	N/A
Subgroup				.43[c]				<.001[c]
Low-risk group	48	12.3 (9.3-16.3)	23.2 (18.6-29.1)	N/A	13	25.7 (16.0-41.3)	23.8 (13.8-41.0)	N/A
Nonscreened group	1	13.1 (1.8-93.0)	7.5 (1.1-53.4)	.17[e]	40	82.3 (60.4-112.3)	35.7 (26.2-48.6)	<.001[e]
Screened group	0[f]	0[f]	0[f]	N/A	14	41.97 (24.86-70.87)	17.7 (10.47-29.94)	N/A

[a] Rate is the number of cases per 100,000 person-year.

[b] Rate is adjusted by age group and gender.

[c] *P* value is generated by comparing the 3 groups (low-risk group, nonscreened group, and screened group).

[d] N/A: not applicable.

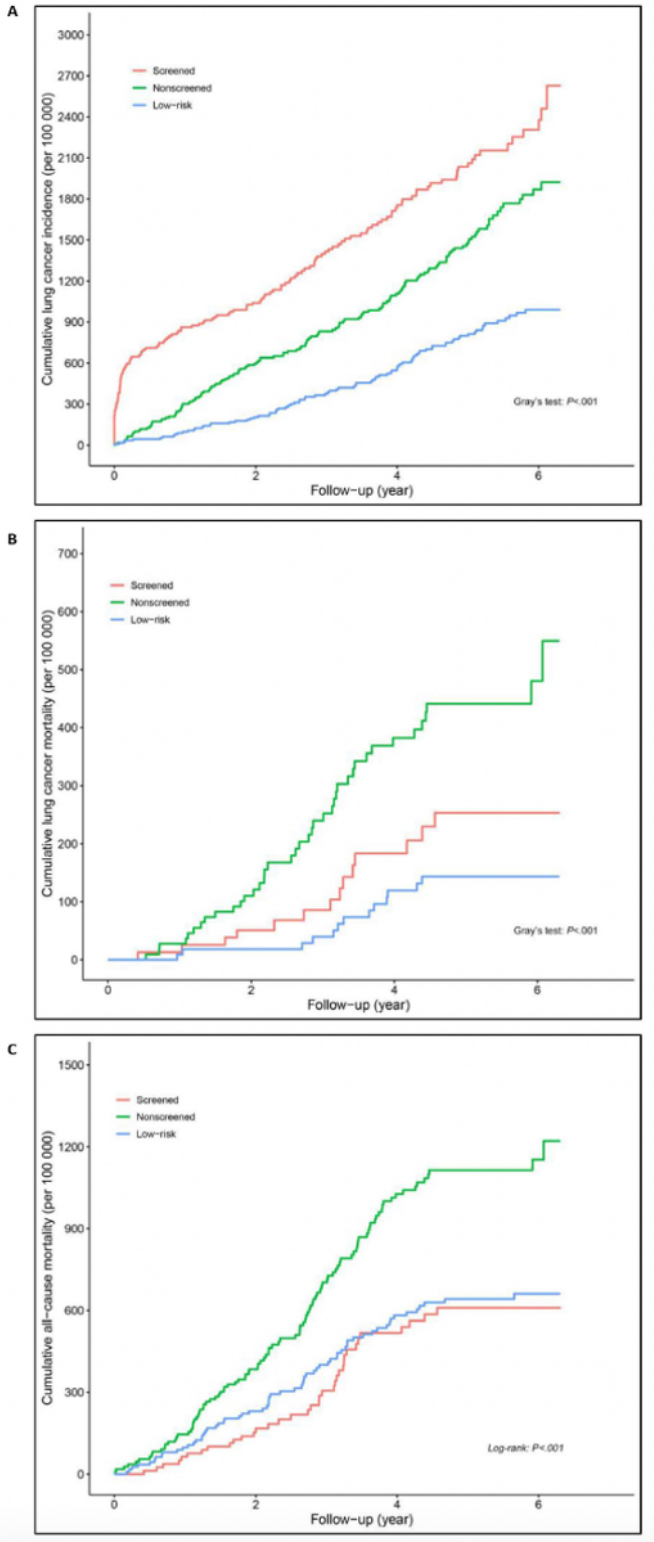
[e] *P* value is generated by comparing the 2 groups (nonscreened group and screened group).

[f] No lung cancer death cases were reported for the group.

Lung Cancer Incidence and Mortality in Smokers

For smokers, the crude lung cancer incidence densities in the low-risk, nonscreened, and screened groups were 157.80 (95% CI 127.74-194.94) per 100,000 person-year, 306.10 (95% CI 260.55-359.61) per 100,000 person-year, and 433.71 (95% CI 368.14-510.96) per 100,000 person-year, respectively (*P*<.001; Figure 2A-C; Table 4). After adjusting for potential confounders, participants at high risk had significantly higher lung cancer incidence intensity (HR 2.20, 95% CI 1.72-2.82; *P*<.001), lung cancer mortality (HR 2.93, 95% CI 1.60-5.37; *P*<.001), and all-cause mortality rate (HR 1.42, 95% CI 1.06-1.91; *P*=.02) than participants at low risk, and participants in the screened group had significantly higher lung cancer incidence (HR 1.39, 95% CI 1.09-1.76; *P*=.007) but lower lung cancer mortality (HR 0.52, 95% CI 0.28-0.96; *P*=.04) and all-cause mortality (HR 0.47, 95% CI 0.32-0.69; *P*<.001) than the nonscreened group (Multimedia Appendix 3).

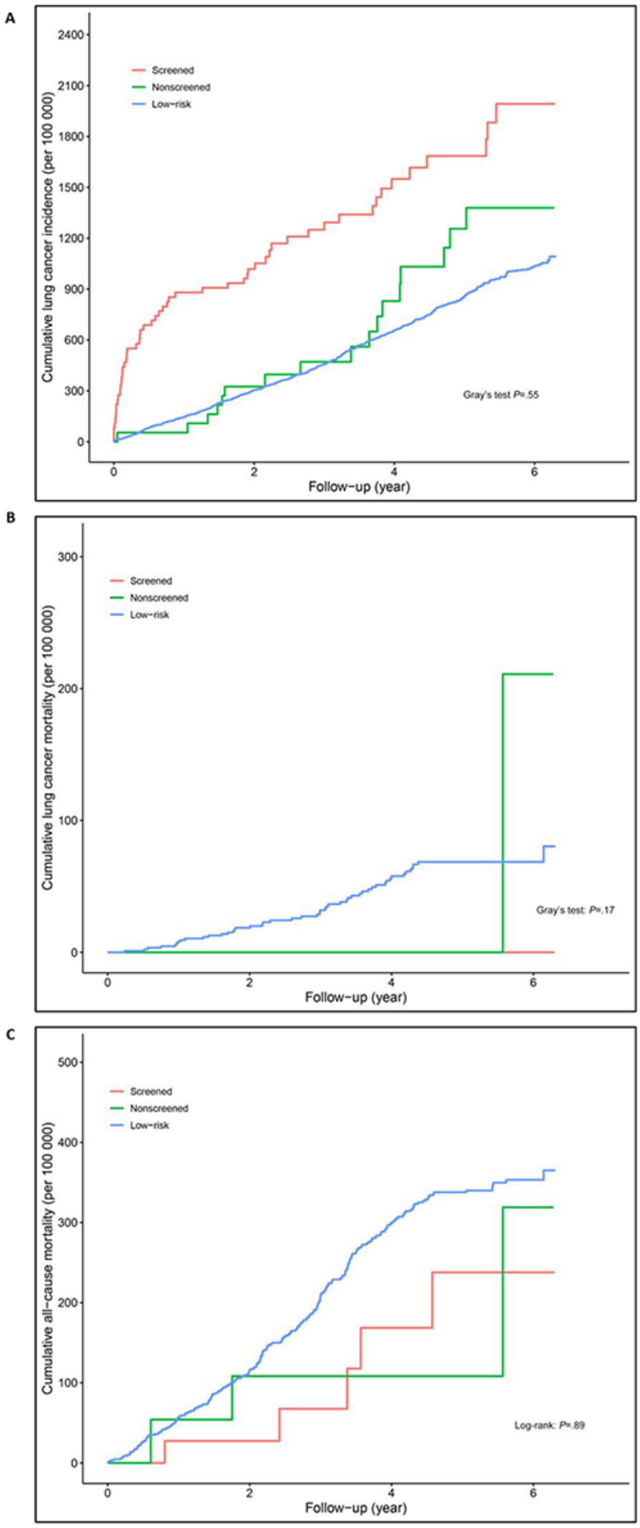
Figure 2. (A) Cumulative intensity of lung cancer incidence, (B) cumulative lung cancer mortality, and (C) cumulative all-cause mortality in smokers.



Lung Cancer Incidence and Mortality in Nonsmokers

For nonsmokers, lung cancer incidence densities in the low-risk, nonscreened, and screened groups were 170.23 (95% CI 157.74-183.70) per 100,000 person-year, 223.72 (95% CI 139.08-359.88) per 100,000 person-year, and 382.33 (95% CI 292.82-499.21) per 100,000 person-year, respectively. No significant differences were detected in lung cancer intensity (*P*=.06), all-cause mortality (*P*=.89), and lung cancer mortality (*P*=.17) between the screened and nonscreened groups (Table 4; Figure 3).

Figure 3. (A) Cumulative intensity of lung cancer incidence, (B) cumulative lung cancer mortality, and (C) cumulative all-cause mortality in nonsmokers.



Discussion

Principal Findings

In this multicenter-based prospective cohort study involving ≥120,000 participants, we found a higher participation rate of LDCT screening among nonsmokers than among smokers (3636/5483, 66.31% vs 7885/18,818, 41.9%). Statistically significant reduced lung cancer mortality and all-cause mortality were observed among smokers in the screened group compared with those in the nonscreened group, with a higher proportion of early-stage lung cancer (72/106, 67.9% in the screened group and 56/114, 49.1% in the nonscreened group). However, we failed to find a significant mortality reduction among screened nonsmokers, relative to nonscreened nonsmokers.

Comparison With Prior Work

The LDCT screening participation rate for nonsmokers was higher than that for smokers, as reported in other studies [13,14]. Studies have consistently shown that smokers have lower recognition of tobacco harm, lower health awareness, or preferred not to be notified of their status of lung cancer [14-18]. Women accounted for the majority of nonsmokers, and women had better awareness of cancer prevention than men [19]. This called for a multipronged approach to enhance engagement and extensively scale up lung cancer screening among smokers in China. As expected, nonsmokers and smokers with lung cancer risk factors (such as occupational exposure to hazardous substances, family history of lung cancer, and history of hepatobiliary diseases) had better LDCT screening compliance, as these disease-related factors may raise people's awareness of lung cancer. For individuals with high lung cancer risk but low screening compliance, promotion of and training on the benefits and need for lung cancer screening are urgently required. Therefore, health education should target these residents to develop good living habits and improve cancer awareness.

As expected, a higher lung cancer incidence density was detected in the high-risk and screened groups, and most lung cancers among the nonsmoker participants were detected in the early stages (476/593, 80.3% were in stage 0-I), in contrast to those among smokers (173/287, 60.3% were in stage 0-I). Moreover, our study showed that the proportion of early-stage lung cancer was also high in the nonscreened group (10/12, 83% vs 56/114, 49.1%). A multicenter hospital-based study in China reported that the proportion of patients with stage I lung cancer was only 19% [20]. This disparity implies that screening programs involving health education, even without the provision of free LDCT examination, might increase the detection rate of early-stage lung cancer when it is still curable.

In our study, the most frequent histological type was adenocarcinoma, with a proportion of 90.2% (607/673) among nonsmokers and 61.2% (194/317) among smokers. East Asian female never-smokers tend to develop adenocarcinoma, with the majority developing from oncogenic mutations [21]. Significantly higher proportions of adenocarcinoma among nonsmokers can be explained by the higher proportion of women among nonsmokers. In addition, in the screened group, the proportion of stage 0-I disease among nonsmokers was higher than that among smokers (42/44, 96% vs 72/106, 67.9%). Similar results were found in a population screened in Korea (92.7% vs 63.6%) [22]. These findings strongly suggest that LDCT screening may be an effective strategy to detect early-stage lung cancer in nonsmokers and smokers in Asia, even though the incidence density of lung cancer is lower in nonsmokers. Despite the poor prognosis of lung cancer, patients with stage I disease have a 5-year survival rate of >75% [23,24]. Therefore, our results establish the potential value of lung cancer screening programs for smokers and nonsmokers.

Over a median follow-up time of 5.1 years, among smokers who were evaluated as high risk, we found a remarkable 48% lung cancer mortality reduction by screening with LDCT. LDCT screening has high potential benefit in decreasing lung cancer mortality worldwide, as shown in the NLST, NELSON, and the German Lung Cancer Screening Intervention Trial studies, which demonstrated that LDCT screening reduced lung cancer mortality by 20%, 24%, and 26%, respectively [3,4,25]. The reduction in our study was greater because we focused on the smoker subgroup. In the Multicentric Italian Lung Detection trial, a similar reduction (39%) was observed in lung cancer–related mortality following LDCT screening among smoking high-risk participants with ≥20 pack-years [26]. As tobacco smoking is a major risk factor for lung cancer [27], we conclude that participants with a history of smoking are likely to derive the greatest benefit from screening. Unlike in trials, the participants were randomly allocated into the LDCT screening group and control group, whereas in our study, the nonscreened group was determined from high-risk participants who declined to participate in further LDCT examination. This indicates that individuals in the nonscreened group (who are at high risk of lung cancer but did not undergo an LDCT scan) were more likely to have poor health awareness and less likely to seek medical assistance. This could be explained by the lowest proportion of early-stage lung cancer, highest all-cause mortality rate, and highest lung cancer mortality rate among the nonscreened group, followed by the low-risk and screened groups. Therefore, we strongly recommend LDCT screening for lung cancer in regions with high smoking rates [28].

The promising effects of the LDCT screening program observed in this study may be because of sufficient medical resources and health care infrastructure in Zhejiang. Zhejiang is a coastal province with a relatively high economy level in China, of which the gross domestic product per capita was US $14,600 in 2020, which was 1.40 times that of the national average and 1.45 times that of the global average. Due to its rapidly growing economy and a well-developed medical system, cancer screening has been well received and undertaken since the 1970s [28,29]. With screening practices in place for decades, Zhejiang has successfully established relatively mature screening networks in local or primary care and has developed multidisciplinary teams of professionals, resulting in a successful and effective LDCT screening program.

We observed a significant 53% reduction in all-cause mortality following LDCT screening among smokers, which was higher than that reported by the NLST (6.7%); however, several European randomized controlled trials, such as the detection and screening of early lung cancer with novel imaging technology and NELSON trials [4,30], found no significant reduction. Much larger sample sizes are needed to detect a significant reduction in all-cause mortality [31]; the sample size in our study was adequate. The consistent reduction in lung cancer mortality and all-cause mortality that we observed in our study is reassuring and supports the accuracy of cause of death assignment and suggests no important harms of screening [32]. In addition to early detection of lung cancer, LDCT screening provides an opportunity to detect cardiovascular disease, pulmonary disease, and extrapulmonary neoplasms [33-35], as a beneficial reduction in all-cause mortality may result from clinically significant incidental findings.

We did not observe a significant reduction in either lung cancer mortality or all-cause mortality by screening for nonsmokers.

This may be because (1) a high proportion of early-stage lung cancers might lead to a favorable prognosis and a limited number of deaths; (2) considering the inadequate sample size as well as the limited follow-up time, no lung cancer deaths were traced in the follow-up period of this study among nonsmokers; or (3) we lacked accurate risk stratification strategies of nonsmokers for lung cancer screening. In East Asia, approximately one-third of lung cancers are unrelated to smoking [36,37]. Between 60% and 80% of women with lung cancer in Asia are nonsmokers compared with between 15% and 20% in the United States and Europe [5]. The current recommendation in the United States is that nonsmokers should not be screened; however, this recommendation is based on modeling in a predominantly White population, making the conclusion likely not applicable to Asian countries with a higher proportion of lung cancers in nonsmokers [38,39]. Moreover, in a population-based ecological cohort study among young women in Taiwan of China who rarely smoke, a significant upward trend in lung cancer incidence after LDCT screening but a stable trend in mortality was observed [40], which is suggestive of screening-related lung cancer overdiagnosis. Nevertheless, LDCT screening decreased the lung cancer mortality rate (HR 0.41) among nonsmokers in Hitachi City, Japan [41], suggesting that LDCT screening is effective in preventing lung cancer death among nonsmokers. In this study, the high adherence to LDCT screening and the high proportion of early-stage lung cancer may reflect a high level of conscientiousness among nonsmokers, indicating that the screening program involving health education has been advantageous for nonsmokers. However, it remains unclear whether LDCT screening can help nonsmokers in reducing mortality. Hence, longer follow-up data are urgently required to evaluate the effects of LDCT screening among nonsmokers. A risk prediction model that is applicable to Asian populations to identify nonsmokers with a high risk of lung cancer for LDCT screening must be established.

Limitations

Our study had some limitations. First, this study may not be representative of the entire general population of China, but it can provide a scientific basis for other regions with similar socioeconomic status. Second, this was a real-world study rather than a randomized controlled trial, which might have led to residual confounders. We admit that a health volunteer effect existed in our study. Our program increased the health awareness of all participants involved in the program, including those in the screened, nonscreened, and low-risk groups. This can explain why the early-stage distributions were higher among all groups than in the general population. These findings suggest that screening programs involving health education may also increase the early detection of lung cancer and potentially decrease disease-specific mortality. Third, the median follow-up of 5.1 years in this study might be insufficient to trace cases and achieve consistent findings. Therefore, additional studies with extended follow-up times should be conducted to further evaluate the screening effects. Fourth, the smoker group was not categorized into heavy and light smokers because of the limited number of cases, and male participants without available information on smoking history were excluded. Passive smoking was only considered for women, which potentially induced gender disparity in smokers and nonsmokers. Finally, the outcome data presented here were obtained from the cancer registry and death surveillance systems, which might have resulted in a misclassification bias. Nevertheless, the cancer registry data in Zhejiang definitively met the requirements of the International Agency for Research on Cancer and the International Association of Cancer Registries, and the data were included in *Cancer Incidence in Five Continents* (*Volume VIII-XI*) consecutively [42-44]. Thus, the data quality obtained from Zhejiang Province was reliable.

Conclusions

In conclusion, our findings suggest effective reduction of all-cause and lung cancer mortality among smokers following LDCT screening, whereas evidence of LDCT screening effectiveness for nonsmokers is still insufficient. It is important to not only target smokers but also identify nonsmokers at high risk of developing lung cancer to implement lung cancer screening programs and maximize screening benefits in China. Our study has significant health service implications, thus providing promising evidence to support the implementation of a national lung cancer screening program and to define optimal guidelines for lung cancer screening in China.

Acknowledgments

This study was funded by the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (grant KFJ-STS-QYZD-2021-08-001), the National Natural Science Foundation of China (grant 82273700), and the nonprofit Central Research Institute Fund of Chinese Academy of Medical Sciences (grant 2019PT320027).

Authors' Contributions

LD and NL contributed equally as the corresponding authors. LD, NL, LW, YW, and FW contributed to the conception and design of the study. LD, NL, LW, HL, CZ, and YW did the data curation; LW and FW performed formal data analysis; LD and NL did the funding acquisition and project administration; YG, ZF, WG, HL, CZ, YC, and LS performed the field investigation and supervision; LW, YW, and FW drafted the paper; and YG, ZF, WG, HL, CZ, YC, LS, LD, and NL interpreted the results. All authors contributed to data interpretation and rewriting of the paper. All authors reviewed and approved the final version. All authors had full access to all the data and were responsible for the decision to submit the manuscript. The corresponding author had full access to all the data in the study and had responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics of the study population of nonsmokers.
[DOCX File , 44 KB - publichealth_v9i1e43586_app1.docx]

Multimedia Appendix 2
Baseline characteristics of the study population of smokers.
[DOCX File , 44 KB - publichealth_v9i1e43586_app2.docx]

Multimedia Appendix 3
Effects of risk evaluation and low-dose computed tomography screening on lung cancer incidence, lung cancer mortality, and all-cause mortality among smokers.
[DOCX File , 36 KB - publichealth_v9i1e43586_app3.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Global Burden of Disease 2019 Cancer Collaboration. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: a systematic analysis for the global burden of disease study 2019. JAMA Oncol 2022 Mar 01;8(3):420-444 [FREE Full text] [doi: 10.1001/jamaoncol.2021.6987] [Medline: 34967848]
3. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011 Aug 04;365(5):395-409 [FREE Full text] [doi: 10.1056/NEJMoa1102873] [Medline: 21714641]
4. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020 Feb 06;382(6):503-513. [doi: 10.1056/NEJMoa1911793] [Medline: 31995683]
5. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. Nat Rev Cancer 2007 Oct;7(10):778-790. [doi: 10.1038/nrc2190] [Medline: 17882278]
6. Wang JB, Jiang Y, Wei WQ, Yang GH, Qiao YL, Boffetta P. Estimation of cancer incidence and mortality attributable to smoking in China. Cancer Causes Control 2010 Jun;21(6):959-965. [doi: 10.1007/s10552-010-9523-8] [Medline: 20217210]
7. Wang Y, Shi J, Du L, Huang H, Wang L, Zhu J, et al. Health-related quality of life in patients with esophageal cancer or precancerous lesions assessed by EQ-5D: a multicenter cross-sectional study. Thorac Cancer 2020 Apr;11(4):1076-1089 [FREE Full text] [doi: 10.1111/1759-7714.13368] [Medline: 32130756]
8. Wang D, Zhu Y, Li C, Wang Y, Wang S, Wu S, et al. Relationship between cigarette smoking and hearing loss in China: a cross-sectional study in Zhejiang province. Tob Induc Dis 2021 May 26;19:40 [FREE Full text] [doi: 10.18332/tid/135440] [Medline: 34079432]
9. Wang YQ, Li HZ, Gong WW, Chen YY, Zhu C, Wang L, et al. Cancer incidence and mortality in Zhejiang Province, Southeast China, 2016: a population-based study. Chin Med J (Engl) 2021 Jul 29;134(16):1959-1966 [FREE Full text] [doi: 10.1097/CM9.0000000000001666] [Medline: 34334631]
10. Zheng R, Zhang S, Zeng H, Wang S, Sun K, Chen R, et al. Cancer incidence and mortality in China, 2016. J Natl Cancer Cent 2022 Mar;2(1):1-9 [FREE Full text] [doi: 10.1016/j.jncc.2022.02.002]
11. Wang L, Shi JF, Zhu J, Huang HY, Bai YN, Liu GX, Health Economic Evaluation Working Group of the Cancer Screening Program in Urban China. Health-related quality of life and utility scores of patients with breast neoplasms in China: a multicenter cross-sectional survey. Breast 2018 Jun;39:53-62. [doi: 10.1016/j.breast.2018.03.004] [Medline: 29597131]
12. Colditz GA, Atwood KA, Emmons K, Monson RR, Willett WC, Trichopoulos D, et al. Harvard report on cancer prevention volume 4: Harvard Cancer Risk Index. Risk Index Working Group, Harvard Center for Cancer Prevention. Cancer Causes Control 2000 Jul;11(6):477-488. [doi: 10.1023/a:1008984432272] [Medline: 10880030]
13. Li N, Tan F, Chen W, Dai M, Wang F, Shen S, National Lung Cancer Screening programme group. One-off low-dose CT for lung cancer screening in China: a multicentre, population-based, prospective cohort study. Lancet Respir Med 2022 Apr;10(4):378-391. [doi: 10.1016/S2213-2600(21)00560-9] [Medline: 35276087]
14. Liang D, Shi J, Li D, Wu S, Jin J, He Y. Participation and yield of a lung cancer screening program in Hebei, China. Front Oncol 2021 Jan 10;11:795528 [FREE Full text] [doi: 10.3389/fonc.2021.795528] [Medline: 35083151]

15. Silvestri GA, Nietert PJ, Zoller J, Carter C, Bradford D. Attitudes towards screening for lung cancer among smokers and their non-smoking counterparts. Thorax 2007 Feb;62(2):126-130 [FREE Full text] [doi: 10.1136/thx.2005.056036] [Medline: 17101739]
16. Quaife SL, Vrinten C, Ruparel M, Janes SM, Beeken RJ, Waller J, et al. Smokers' interest in a lung cancer screening programme: a national survey in England. BMC Cancer 2018 May 02;18(1):497 [FREE Full text] [doi: 10.1186/s12885-018-4430-6] [Medline: 29716550]
17. Castro S, Sosa E, Lozano V, Akhtar A, Love K, Duffels J, et al. The impact of income and education on lung cancer screening utilization, eligibility, and outcomes: a narrative review of socioeconomic disparities in lung cancer screening. J Thorac Dis 2021 Jun;13(6):3745-3757 [FREE Full text] [doi: 10.21037/jtd-20-3281] [Medline: 34277066]
18. Wu FZ, Wu YJ, Chen CS, Yang SC. Impact of smoking status on lung cancer characteristics and mortality rates between screened and non-screened lung cancer cohorts: real-world knowledge translation and education. J Pers Med 2022 Jan 02;12(1):26 [FREE Full text] [doi: 10.3390/jpm12010026] [Medline: 35055341]
19. Wang L, Wang Y, Li H, Zhu C, Du L. Awareness of core knowledge about cancer prevention and its influencing factors among residents in Zhejiang Province. J Prev Med 2021;33(8):757-762 [FREE Full text]
20. Shi JF, Wang L, Wu N, Li JL, Hui ZG, Liu SM, LuCCRES Group. Clinical characteristics and medical service utilization of lung cancer in China, 2005-2014: overall design and results from a multicenter retrospective epidemiologic survey. Lung Cancer 2019 Feb;128:91-100. [doi: 10.1016/j.lungcan.2018.11.031] [Medline: 30642458]
21. Ha SY, Choi S, Cho JH, Choi HJ, Lee J, Jung K, et al. Lung cancer in never-smoker Asian females is driven by oncogenic mutations, most often involving EGFR. Oncotarget 2015 Mar 10;6(7):5465-5474 [FREE Full text] [doi: 10.18632/oncotarget.2925] [Medline: 25760072]
22. Kang HR, Cho JY, Lee SH, Lee YJ, Park JS, Cho YJ, et al. Role of low-dose computerized tomography in lung cancer screening among never-smokers. J Thorac Oncol 2019 Mar;14(3):436-444 [FREE Full text] [doi: 10.1016/j.jtho.2018.11.002] [Medline: 30445189]
23. Liang W, Shao W, Jiang G, Wang Q, Liu L, Liu D, et al. Chinese multi-institutional registry (CMIR) for resected non-small cell lung cancer: survival analysis of 5,853 cases. J Thorac Dis 2013 Dec;5(6):726-729 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2013.12.32] [Medline: 24409347]
24. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, Advisory Boards, and Participating Institutions. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. J Thorac Oncol 2016 Jan;11(1):39-51 [FREE Full text] [doi: 10.1016/j.jtho.2015.09.009] [Medline: 26762738]
25. Becker N, Motsch E, Trotter A, Heussel CP, Dienemann H, Schnabel PA, et al. Lung cancer mortality reduction by LDCT screening-results from the randomized German LUSI trial. Int J Cancer 2020 Mar 15;146(6):1503-1513 [FREE Full text] [doi: 10.1002/ijc.32486] [Medline: 31162856]
26. Pastorino U, Silva M, Sestini S, Sabia F, Boeri M, Cantarutti A, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol 2019 Oct 01;30(10):1672 [FREE Full text] [doi: 10.1093/annonc/mdz169] [Medline: 31168572]
27. Triphuridet N, Henschke C. Landscape on CT screening for lung cancer in Asia. Lung Cancer (Auckl) 2019 Sep 30;10:107-124 [FREE Full text] [doi: 10.2147/LCTT.S192643] [Medline: 31686936]
28. Wang Z, Han W, Zhang W, Xue F, Wang Y, Hu Y, et al. Mortality outcomes of low-dose computed tomography screening for lung cancer in urban China: a decision analysis and implications for practice. Chin J Cancer 2017 Jul 14;36(1):57 [FREE Full text] [doi: 10.1186/s40880-017-0221-8] [Medline: 28709441]
29. Du L, Mao W, Wang X, Yu C, Li H. The past and present of Zhejiang cancer prevention. China Cancer 2013;22(10):766-770 [FREE Full text] [doi: 10.11735/j.issn.1004-0242.2013.10.A001]
30. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, DANTE Study Group. Long-term follow-up results of the DANTE trial, a randomized study of lung cancer screening with spiral computed tomography. Am J Respir Crit Care Med 2015 May 15;191(10):1166-1175. [doi: 10.1164/rccm.201408-1475OC] [Medline: 25760561]
31. Heijnsdijk EA, Csanádi M, Gini A, Ten Haaf K, Bendes R, Anttila A, et al. All-cause mortality versus cancer-specific mortality as outcome in cancer screening trials: a review and modeling study. Cancer Med 2019 Oct;8(13):6127-6138 [FREE Full text] [doi: 10.1002/cam4.2476] [Medline: 31422585]
32. Dobbin KK, Ebell M. Should we expect all-cause mortality reductions in large screening studies? Br J Gen Pract 2018 Jun;68(671):290-291 [FREE Full text] [doi: 10.3399/bjgp18X696545] [Medline: 29853594]
33. Godoy MC, White CS, Erasmus JJ, Wu CC, Truong MT, Munden RF, et al. Extrapulmonary neoplasms in lung cancer screening. Transl Lung Cancer Res 2018 Jun;7(3):368-375 [FREE Full text] [doi: 10.21037/tlcr.2018.06.05] [Medline: 30050774]
34. Cai Q, Triphuridet N, Zhu Y, You N, Yip R, Yankelevitz DF, et al. Bronchiectasis in low-dose CT screening for lung cancer. Radiology 2022 Aug;304(2):437-447. [doi: 10.1148/radiol.212547] [Medline: 35438565]

35. Zhu Y, Yip R, Shemesh J, Jirapatnakul AC, Yankelevitz DF, Henschke CI. Combined aortic valve and coronary artery calcifications in lung cancer screening as predictors of death from cardiovascular disease. Eur Radiol 2020 Dec;30(12):6847-6857. [doi: 10.1007/s00330-020-07049-4] [Medline: 32725329]
36. Handy CE, Quispe R, Pinto X, Blaha MJ, Blumenthal RS, Michos ED, et al. Synergistic opportunities in the interplay between cancer screening and cardiovascular disease risk assessment: together we are stronger. Circulation 2018 Aug 14;138(7):727-734. [doi: 10.1161/CIRCULATIONAHA.118.035516] [Medline: 30359131]
37. Lam S. Lung cancer screening in never-smokers. J Thorac Oncol 2019 Mar;14(3):336-337 [FREE Full text] [doi: 10.1016/j.jtho.2018.12.019] [Medline: 30782380]
38. Yano T, Miura N, Takenaka T, Haro A, Okazaki H, Ohba T, et al. Never-smoking nonsmall cell lung cancer as a separate entity: clinicopathologic features and survival. Cancer 2008 Sep 01;113(5):1012-1018 [FREE Full text] [doi: 10.1002/cncr.23679] [Medline: 18618510]
39. Ten Haaf K, de Koning HJ. Should never-smokers at increased risk for lung cancer be screened? J Thorac Oncol 2015 Sep;10(9):1285-1291 [FREE Full text] [doi: 10.1097/JTO.0000000000000593] [Medline: 26287320]
40. Gao W, Wen CP, Wu A, Welch HG. Association of computed tomographic screening promotion with lung cancer overdiagnosis among Asian women. JAMA Intern Med 2022 Mar 01;182(3):283-290 [FREE Full text] [doi: 10.1001/jamainternmed.2021.7769] [Medline: 35040922]
41. Nawa T, Nakagawa T, Mizoue T, Kusano S, Chonan T, Hayashihara K, et al. A decrease in lung cancer mortality following the introduction of low-dose chest CT screening in Hitachi, Japan. Lung Cancer 2012 Dec;78(3):225-228. [doi: 10.1016/j.lungcan.2012.09.012] [Medline: 23069269]
42. Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB. Cancer Incidence in Five Continents. Volume VIII. Lyon, France: International Agency for Research on Cancer; 2002.
43. Shin HR, Masuyer E, Ferlay J, Curado MP, Asian Contributors to CI5 IX4. Cancer in Asia - incidence rates based on data in cancer incidence in five continents IX (1998-2002). Asian Pac J Cancer Prev 2010;11 Suppl 2:11-16 [FREE Full text] [Medline: 20553065]
44. Bray F, Ferlay J, Laversanne M, Brewster DH, Gombe Mbalawa C, Kohler B, et al. Cancer incidence in five continents: inclusion criteria, highlights from Volume X and the global status of cancer registration. Int J Cancer 2015 Nov 01;137(9):2060-2071 [FREE Full text] [doi: 10.1002/ijc.29670] [Medline: 26135522]

Abbreviations

HR: hazard ratio
LDCT: low-dose computed tomography
NELSON: Nederlands Leuvens Screening Onderzoek
NLST: National Lung Screening Trial
OR: odds ratio

Edited by Y Khader; submitted 17.10.22; peer-reviewed by L Guo, S Yan, M Dai; comments to author 02.12.22; revised version received 17.01.23; accepted 05.02.23; published 14.03.23.

Please cite as:
Wang L, Wang Y, Wang F, Gao Y, Fang Z, Gong W, Li H, Zhu C, Chen Y, Shi L, Du L, Li N
Disparity in Lung Cancer Screening Among Smokers and Nonsmokers in China: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e43586
URL: https://publichealth.jmir.org/2023/1/e43586
doi:10.2196/43586
PMID:36917151

Original Paper

Comparing the Effectiveness, Tolerability, and Acceptability of Heated Tobacco Products and Refillable Electronic Cigarettes for Cigarette Substitution (CEASEFIRE): Randomized Controlled Trial

Pasquale Caponnetto[1,2], PhD; Davide Campagna[1,3,4,5], MD; Marilena Maglia[3], DPhil; Francesca Benfatto[3], BA; Rosalia Emma[4], PhD; Massimo Caruso[1,6], PhD; Grazia Caci[7], MD; Barbara Busà[8], MD; Alfio Pennisi[9], MD; Maurizio Ceracchi[10], MD; Marcello Migliore[11,12], MD; Maria Signorelli[13], PhD

[1]Center of Excellence for the Acceleration of Harm Reduction, University of Catania, Catania, Italy
[2]Department of Educational Sciences, University of Catania, Catania, Italy
[3]Eclat Spin off srl, University of Catania, Catania, Italy
[4]Department of Clinical & Experimental Medicine, University of Catania, Catania, Italy
[5]Unità Operativa Complessa Medicina Accettazione Urgenza, University Teaching Hospital, University of Catania, Catania, Italy
[6]Department of Biomedical and Biotechnological Sciences, University of Catania, Catania, Italy
[7]Unit of Infectious Diseases, Department of Clinical and Experimental Medicine, University of Messina, Messina, Italy
[8]Dipartimento Emergenza-Urgenza, Farmacia presidio ospedaliero centro Azienda Ospedaliera di Rilievo Nazionale e di Alta Specializzazione Garibaldi, Catania, Italy
[9]Department of Respiratory Medicine, Hospital Clinics "Musumeci-Gecas", Catania, Italy
[10]Fullcro Srl, Roma, Italy
[11]Department of Surgery and Medical Specialties, University of Catania, Catania, Italy
[12]Minimally Invasive Thoracic Surgery and New Technology, University Hospital of Catania, Catania, Italy
[13]Department of Clinical and Experimental Medicine, Psychiatry Unit, University of Catania, Catania, Italy

Corresponding Author:
Pasquale Caponnetto, PhD
Center of Excellence for the Acceleration of Harm Reduction
University of Catania
Via Santa Sofia
Catania
Italy
Phone: 39 0954781124
Email: p.caponnetto@unict.it

Abstract

Background: People who smoke and who face challenges trying to quit or wish to continue to smoke may benefit by switching from traditional cigarettes to noncombustible nicotine delivery alternatives, such as heated tobacco products (HTPs) and electronic cigarettes (ECs). HTPs and ECs are being increasingly used to quit smoking, but there are limited data about their effectiveness.

Objective: We conducted the first randomized controlled trial comparing quit rates between HTPs and ECs among people who smoke and do not intend to quit.

Methods: We conducted a 12-week randomized noninferiority switching trial to compare effectiveness, tolerability, and product satisfaction between HTPs (IQOS 2.4 Plus) and refillable ECs (JustFog Q16) among people who do not intend to quit. The cessation intervention included motivational counseling. The primary endpoint of the study was the carbon monoxide–confirmed continuous abstinence rate from week 4 to week 12 (CAR weeks 4-12). The secondary endpoints included the continuous self-reported ≥50% reduction in cigarette consumption rate (continuous reduction rate) from week 4 to week 12 (CRR weeks 4-12) and 7-day point prevalence of smoking abstinence.

Results: A total of 211 participants completed the study. High quit rates (CAR weeks 4-12) of 39.1% (43/110) and 30.8% (33/107) were observed for IQOS-HTP and JustFog-EC, respectively. The between-group difference for the CAR weeks 4-12 was not significant (*P*=.20). The CRR weeks 4-12 values for IQOS-HTP and JustFog-EC were 46.4% (51/110) and 39.3% (42/107), respectively, and the between-group difference was not significant (*P*=.24). At week 12, the 7-day point prevalence of smoking abstinence values for IQOS-HTP and JustFog-EC were 54.5% (60/110) and 41.1% (44/107), respectively. The most frequent adverse events were cough and reduced physical fitness. Both study products elicited a moderately pleasant user experience, and the between-group difference was not significant. A clinically relevant improvement in exercise tolerance was observed after switching to the combustion-free products under investigation. Risk perception for conventional cigarettes was consistently higher than that for the combustion-free study products under investigation.

Conclusions: Switching to HTPs elicited a marked reduction in cigarette consumption among people who smoke and do not intend to quit, which was comparable to refillable ECs. User experience and risk perception were similar between the HTPs and ECs under investigation. HTPs may be a useful addition to the arsenal of reduced-risk alternatives for tobacco cigarettes and may contribute to smoking cessation. However, longer follow-up studies are required to confirm significant and prolonged abstinence from smoking and to determine whether our results can be generalized outside smoking cessation services offering high levels of support.

Trial Registration: ClinicalTrials.gov NCT03569748; https://clinicaltrials.gov/ct2/show/NCT03569748

(JMIR Public Health Surveill 2023;9:e42628) doi:10.2196/42628

KEYWORDS

harm reduction; heat not burn; electronic cigarettes; smoking cessation; smoking reduction; e-cigarette; public health; psychological well-being

Introduction

According to the World Health Organization, smoking is endemic, with more than 1.1 billion smokers worldwide, resulting in approximately 7 million premature deaths every year [1]. Deaths are primarily due to lung cancer and fatal complications of ischemic heart disease and chronic obstructive pulmonary disease (COPD) [1-3]. The risk of developing diseases has been shown to significantly reduce when stopping smoking [4,5].

Achieving cessation is challenging, because quit rates are low, relapse rates are high, and many smokers wish to continue to smoke [6,7]. The compulsion to smoke is difficult to break, and even for those who do quit smoking, relapse is the norm. For unsupported quit attempts, 80% of people relapse in the first month [7,8], and among people who smoke and use treatments, 75% fail within 6 months, with the large majority resuming smoking within 2 weeks [9]. Even among those who quit smoking during hospitalization and intended to stay quit, 25% relapsed on the first day after discharge [10].

Substitution of combustible tobacco cigarettes with less harmful combustion-free nicotine delivery alternatives (NDAs), such as electronic cigarettes (ECs) and heated tobacco products (HTPs), is now a relatively new option available to smokers [6,11-13].

ECs operate by heating an element that vaporizes a solution. HTPs consist of a holder that electronically transfers controlled heat to tobacco sticks that generate a nicotine-containing aerosol. Since the emission aerosols of combustion-free NDAs are produced at a much lower vaporizing/heating temperature compared to that of combustion (which generally starts above 400 °C), they contain less harmful and potentially harmful chemicals than tobacco smoke [14-19].

Although not completely risk free, EC and HTP use may help respiratory patients to achieve sustained abstinence from cigarette smoking, with clinically relevant health gains [20,21]. The most recent Cochrane review concluded that ECs with nicotine increased quit rates, and compared to nicotine replacement therapy, a risk ratio (RR) of 1.53 was reported, indicating, in absolute terms, an additional 3 quitters for every 100 using ECs [22]. However, formal demonstration of the efficacy of HTPs for smoking cessation is not yet available [23].

HTPs often mimic a hand-to-mouth experience that is very similar to that of conventional tobacco cigarettes. By mimicking the experience of tobacco smoking and its associated rituals, the use of HTPs can provide adequate compensatory physical and behavioral effects [24,25], likely serving as an effective method of relapse prevention [21].

It is not clear if HTPs provide a more gratifying smoking experience compared to ECs, and a direct comparison between the 2 types of products has never been investigated. With this in mind, we conducted a prospective randomized noninferiority switching trial to compare effectiveness, tolerability, and product satisfaction/adoption between HTPs and refillable ECs among people who smoke and do not intend to quit.

Methods

Study Participants

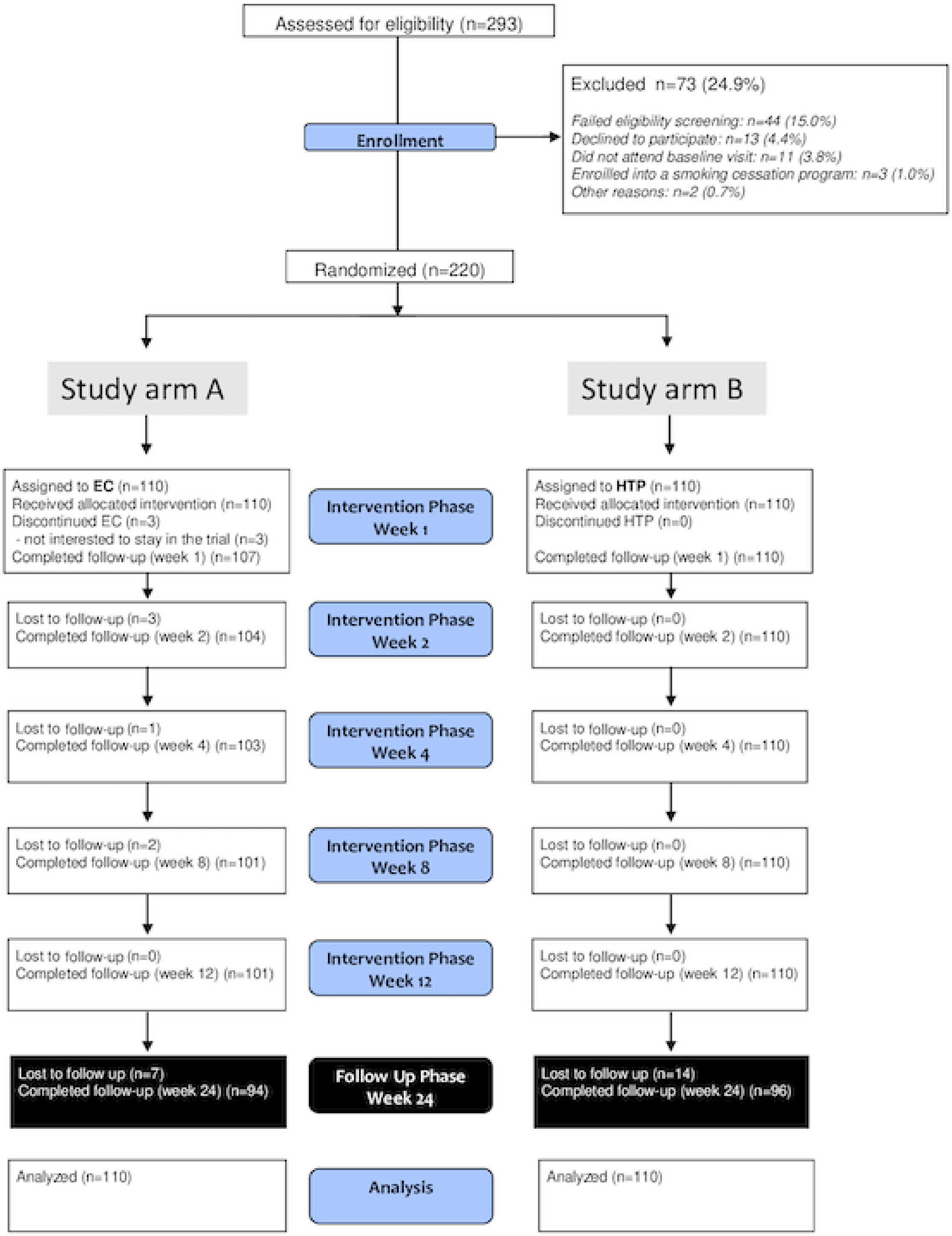
Eligibility criteria have been described previously in detail [26]. In brief, adult people smoking ≥10 cigarettes per day for the past year, having an exhaled carbon monoxide (eCO) level of ≥7 ppm, not intending to quit in the next 30 days, and interested in switching to combustion-free NDAs were recruited among hospital/university staff, via social media, or through word of mouth. Unwillingness to quit was confirmed by the answer "No" to the following questions: "Do you plan to quit smoking within the next 30 days?" and "Do you wish to participate in a smoking cessation program?" The exclusion criteria were as follows: (1) history of depression, panic disorder, psychosis, or bipolar disorder; (2) significant history of alcoholism or drug/chemical abuse within 12 months prior to screening; (3) known clinically significant diseases that, in the opinion of the investigator, would jeopardize the safety of the participant or impact the validity of the study results; (4) use of any tobacco/nicotine delivery device (except for own brands of cigarettes) within the last 3 months; and (5) use of nicotine replacement therapy or other smoking cessation therapies within the last 3 months.

Trial Design and Study Visits

Details of the study design and protocol have been previously published [26]. In brief, this was a 12-week, randomized, 2 parallel arm, open-label, noninferiority trial conducted to compare effectiveness, tolerability, adoption rates, and acceptability between HTPs and ECs in regular smokers (Multimedia Appendix 1). The trial consisted of a total of 8 study visits (see Multimedia Appendix 1). Participants could choose 1 out of 3 different flavors for each class of products and were provided with their preferred flavor for the whole duration of the study. Motivational counseling [27,28] was offered throughout the study to maximize study product adherence, to favor transition away from combustible tobacco cigarettes, and to prevent relapse back to smoking. Activities carried out during study visits are detailed in Multimedia Appendix 2.

The study was conducted in accordance with the Guideline for Good Clinical Practice and followed the Consolidated Standards of Reporting Trials (CONSORT) reporting guidelines for randomized studies (Figure 1).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. EC: electronic cigarette; HTP: heated tobacco product.



Study Products

HTPs and ECs were provided for the whole duration of the intervention phase (12 weeks).

HTPs

Participants randomized to the HTP arm of the study received IQOS 2.4 Plus consisting of a pen-like holder into which a tobacco stick is inserted and heated, and a battery case to recharge the holder after each use. IQOS 2.4 Plus was the only HTP available on the Italian market when this trial was designed. The device is to be used with tobacco sticks specifically processed and manufactured for IQOS (named *HEETS*). Participants could choose from 3 varieties of tobacco sticks (*HEETS* Amber, rich tobacco; *HEETS* Yellow, smooth tobacco; and *HEETS* Turquoise, menthol-flavored tobacco), which were available for sale on the Italian market at the time of the study.

ECs

Participants randomized to the EC arm of the study received JustFog Q16 Starter Kit consisting of a battery and a 1.9-mL refillable tank fitted with a 1.6-Ohm nichrome coil. Participants could choose from 3 varieties of e-liquid flavors (Puff Riserva Country 16 mg, sweet tobacco flavor; Puff Riserva Toscana 16 mg, full tobacco flavor; and Puff Artic 16 mg, menthol flavor;

all 3 are formulated in 50% propylene glycol/40% vegetable glycerin/10% H_2O), which were chosen for the trial by an expert panel to match the sensory experiences of the 3 IQOS tobacco sticks selected for the study.

Study Endpoints

The primary efficacy endpoint of the study was the continuous abstinence rate (CAR) from week 4 to week 12 (CAR weeks 4-12). Abstinence from smoking was defined as eCO-verified (<10 ppm) self-reported abstinence from cigarette smoking. CAR weeks 4-12 was used to compare quit rates between IQOS-HTP and JustFog-EC.

The secondary efficacy endpoints were the 7-day point prevalence of abstinence at week 12 and the continuous reduction rate (CRR) from week 4 to week 12 (CRR weeks 4-12). Smoking reduction was self-reported. A reduction in the number of cigarettes smoked per day of 50% from baseline was considered of importance and excluded those labeled as CAR weeks 4-12. The CRR weeks 4-12 was used to compare reduction rates between IQOS-HTP and JustFog-EC. Participants who could not be classified as CAR weeks 4-12 or CRR weeks 4-12 were considered to have failed. Safety reporting details have been previously published [26].

Product satisfaction was investigated by using the following questionnaires adapted and validated for EC and HTP use: modified Cigarette Evaluation Questionnaire (mCEQ) and modified Smoking Cue Appeal Survey (mSCAS). Risk perception was assessed by using the Perceived Risk Instrument for conventional cigarettes (PRI-P CC) and the Perceived Risk Instrument for reduced risk products (PRI-P RRP). The effect on quality of life was investigated by questionnaires (ie, EQ-5D-5L and EQ VAS) and by measuring changes in body weight and exercise tolerance (ie, $\dot{V}O_2$ max by the Chester step test). Self-reported EC or HTP use at each study visit was verified against the product use check and reported in the electronic case report form (e-CRF). The product use check was used to calculate daily consumption.

Study Assessments

The assessments carried out during study visits are listed in Multimedia Appendix 2 and included the following: (1) number of cigarettes smoked per day; (2) adverse events; (3) eCO levels evaluated with a calibrated handheld device (MicroCO); (4) resting blood pressure and heart rate taken with a semiautomated oscillometric sphygmomanometer (Smart Pressure, CA-MI); (5) body weight, height, body fat, visceral fat, fat-free mass, body muscle mass, bone body mass, metabolic age, and water content taken with a body composition analyzer (Tanita SC-240, Tanita); (6) BMI calculated by dividing weight by height square (kg/m^2); and (7) Chester step test to determine maximal aerobic capacity (ie, $\dot{V}O_2$ max).

Other measurements included the following questionnaires: (1) Fagerstrom Test for Cigarette Dependence [29]; (2) mCEQ [30]; (3) mSCAS [31]; (4) PRI-P CC and PRI-P RRP [32]; and (5) EQ-5D-5L and EQ VAS, a standardized measure of health-related quality of life [33].

Secondary analyses of blood pressure, heart rate, BMI, and Chester step test results by smoking phenotype classification will be reported in separate papers.

Ethical Considerations

The Ethical Review Board of Azienda Ospedaliero Universitaria "Policlinico-V. Emanuele," Universita` di Catania, Italy, reviewed and approved the study (approval reference number: 215/2017/PO). All participants provided written informed consent prior to participation in the study. The study has been registered at ClinicalTrial.gov (trial registration ID: NCT03569748). Study data are deidentified, and participants did not receive compensation.

Statistical Methods

A detailed description of the sample size calculation can be found in the published research protocol [26].

All the analyses were performed using SAS Version 9.4 (SAS Institute Inc). The primary efficacy endpoint of the study in the experimental study group was calculated with a noninferiority threshold of 15%. An α level of .05 was considered. Quit rates were evaluated on an intention-to-treat basis. CAR weeks 4-12 percentages and odds ratios (ORs) were calculated and used to compare quit rates between the IQOS-HTP and JustFog-EC study groups with the chi-square test. Moreover, CRR weeks 4-12 percentages and ORs were calculated and used to compare reduction rates between the IQOS-HTP and JustFog-EC study groups with the chi-square test. For the 7-day point prevalence of smoking abstinence and reduction, percentages were calculated at each study visit to illustrate trends.

Safety data were presented as descriptive statistics separately by study group. Any events documented in the period from the point of product randomization (V1) until the end of the intervention phase at 12 weeks when study products were withdrawn (V6) were considered as relevant for safety analysis.

Descriptive statistics of product acceptability measures (ie, mCEQ and mSCAS) were presented as summary tables by study group and study visit. Changes in mCEQ and mSCAS scores within and between study groups were analyzed using the Wilcoxon signed rank test and Wilcoxon rank sum test, respectively.

For risk perception, descriptive statistics of PRI-P CC and PRI-P RRP values were presented as summary tables by study group and study visit. Changes in PRI-P scores within and between study groups were analyzed using the Wilcoxon signed rank test and Wilcoxon rank sum test, respectively.

EQ-5D-5L, EQ VAS, body weight, and exercise tolerance were presented as descriptive statistics, and within- and between-group comparisons were carried out using the Wilcoxon signed rank test and Wilcoxon rank sum test, respectively.

A multiple logistic regression model was prepared to identify variables able to influence the primary outcome CAR weeks 4-12. We performed an a priori selection of variables able to act as determinants, effect modifiers, or confounders of quitting success. The continuous variables were categorized according to cutoffs based clinically. The univariate analysis was performed considering the subgroups CAR and no CAR as the outcome. The following factors, for which a statistical difference was detected, were included in the model: gender, daily cigarette consumption, and 4 psychological aspects included as domains in the questionnaires (ie, product satisfaction, psychological reward, enjoyment, and craving).

Results

Baseline Characteristics

The CONSORT flow diagram of the study subjects is shown in Figure 1. A total of 220 smokers were enrolled in the study, with 211 (95.9%) participants completing the intervention phase. Baseline characteristics of the participants are shown by study product assignment in Table 1 and were comparable between the study groups. On average, participants were Caucasian adults (approximately 41 years old), were mostly men (approximately 57.3%), had smoked about a pack daily for approximately 24 years, had a moderate Fagerstrom Test for Cigarette Dependence score of about 5, and had an average of 2 quit attempts in the past.

Table 1. Baseline characteristics of the participants.

Characteristic	HTP[a] group (N=110)	EC[b] group (N=110)	*P* value
Sex, n			.06
Female	56 (50.9)	70 (63.6)	
Male	54 (49.1)	40 (36.4)	
Age (years), mean (SD)	41.3 (16.1)	41.3 (16.9)	.97
Education level, n (%)			.84
Primary school	2 (1.8)	1 (0.9)	
Secondary school	20 (18.2)	24 (21.8)	
High school	64 (58.2)	61 (55.5)	
University	24 (21.8)	24 (21.8)	
Cigarettes per day, mean (SD)	22.6 (10.1)	22.8 (10.9)	.91
Exhaled carbon monoxide (ppm), mean (SD)	26.2 (13.6)	26.9 (15.4)	.71
Fagerstrom Test for Cigarette Dependence score	5.8 (2.1)	6.0 (2.2)	.47
Years of smoking, mean (SD)	24.5 (15.6)	22.9 (16.0)	.44
Number of quit attempts, mean (SD)	1.9 (2.7)	1.9 (2.5)	.96

[a] HTP: heated tobacco product.

[b] EC: electronic cigarette.

Smoking Abstinence and Reduction Rates

Smoking abstinence rates (CAR weeks 4-12), reduction rates (CRR weeks 4-12), and 7-day point prevalence of smoking abstinence and reduction are shown in Figure 2, and Figures 3A and 3B.

Figure 2. Smoking abstinence rates and smoking reduction rates. CAR: continuous abstinence rate; CRR: continuous reduction rate; EC: electronic cigarette; HTP: heated tobacco product.

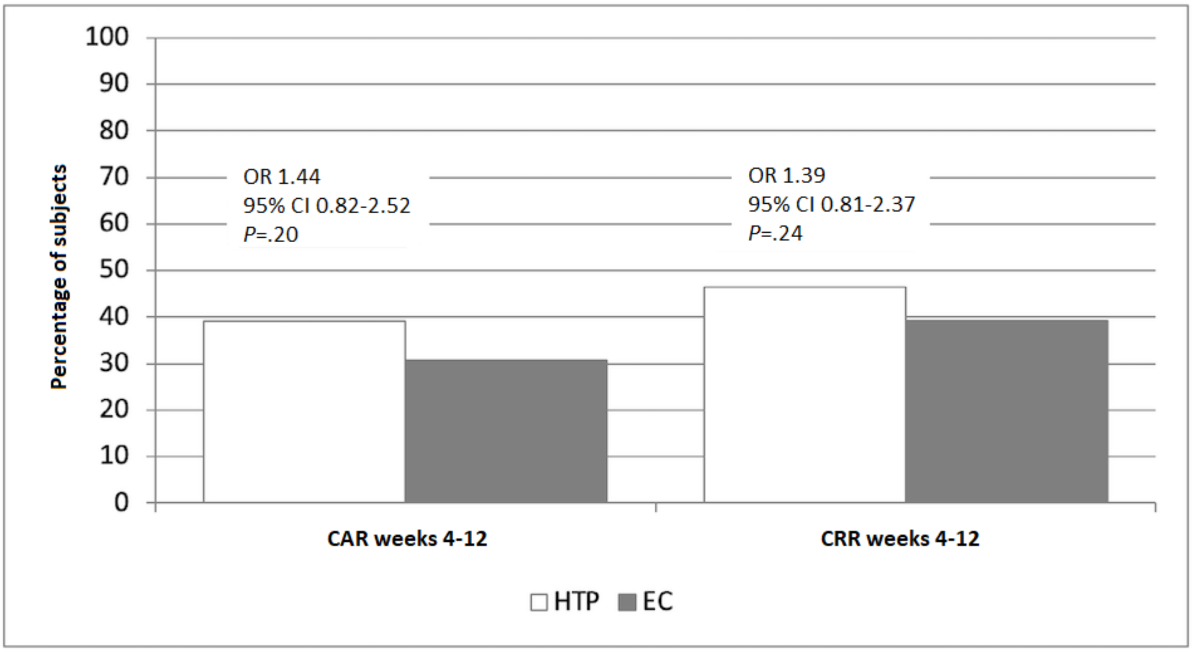
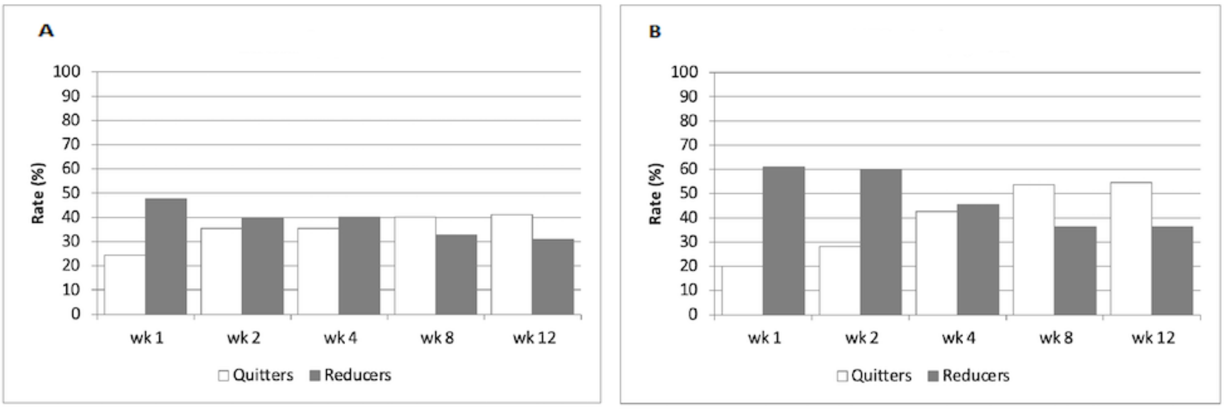


Figure 3. Seven-day prevalence of smoking abstinence and reduction in the electronic cigarette (A) and heated tobacco product (B) study groups.



High quit rates, evaluated on an intention-to-treat basis, were observed for both study groups; the CAR weeks 4-12 values for IQOS-HTP and JustFog-EC were 39.1% (43/110) and 30.8% (33/107), respectively (Figure 2). The between-group difference for the CAR weeks 4-12 was not significant (*P*=.20; chi-square test) and satisfied the noninferiority criteria of the study that differences in quit rates between products did not exceed 15% (OR 1.44, 95% CI 0.82-2.52).

High reduction rates, assessed among participants who were still smoking, were also reported, and the CRR weeks 4-12 values for IQOS-HTP and JustFog-EC were 46.4% (51/110) and 39.3% (42/107), respectively (Figure 2). The between-group difference for the CRR weeks 4-12 was not significant (*P*=.24; chi-square test; OR 1.39, 95% CI 0.81-2.37).

The 7-day point prevalence of smoking abstinence was >20% throughout the intervention phase, with values peaking at week 8 (V5) and week 12 (V6) (Figures 3A and 3B). The 7-day point prevalence of smoking abstinence values for IQOS-HTP and JustFog-EC were 53.6% (59/110) and 40.2% (43/107) at week 8, and 54.5% (60/110) and 41.1% (44/107) at week 12, respectively.

The 7-day point prevalence of smoking reduction (ie, dual use) was higher during the first 2 weeks of the intervention phase (Figures 3A and 3B). The 7-day point prevalence of smoking reduction values for IQOS-HTP and JustFog-EC were 60.9% (67/110) and 47.7% (51/107) at week 1, and 60.0% (66/110) and 39.3% (42/107) at week 2, respectively.

For both study products, Figures 3A and 3B also show progressive reduction in the proportion of dual use in the study, which was paralleled by rising prevalence of exclusive single use by the end of the intervention phase.

Product Preference, Acceptability, and Risk Perception

Among participants in the EC study arm, 50.9% (56/110) chose Puff Riserva Country, 30.9% (34/110) chose Puff Riserva Tuscan, and 18.2% (20/110) chose Puff Artic e-liquid. Among participants in the HTP study arm, 56.4% (62/110) chose *HEETS* Amber, 33.6% (37/110) chose *HEETS* Yellow, and 10.0% (11/110) chose *HEETS* Turquoise tobacco sticks. Technical issues (eg, device malfunctions) were relatively uncommon (Multimedia Appendix 3).

Appeal of the study products was analyzed using the mCEQ and mSCAS. No significant within-group changes in the mCEQ and mSCAS scores were observed (Wilcoxon signed rank test; Multimedia Appendix 4). Between-group changes were also not significantly different (Wilcoxon rank sum test; Multimedia Appendix 4). Moderate liking of the study products, mild psychological reward, moderate enjoyment of the respiratory tract sensation, and craving reduction with minimal aversion were noted (Multimedia Appendix 4). The mSCAS showed that the study products elicited a moderately pleasant user experience (Multimedia Appendix 4).

As expected, risk perception for conventional cigarettes was consistently higher than for the combustion-free study products (Multimedia Appendices 4 and 5). Within-group changes in PRI-P CC scores were small but significantly higher for both study groups (IQOS-HTP, *P*<.001; JustFog-EC, *P*=.003; Wilcoxon signed rank test). Between-group comparisons were not statistically significant (Wilcoxon rank sum test). No significant within- or between-group changes were observed for PRI-P RRP.

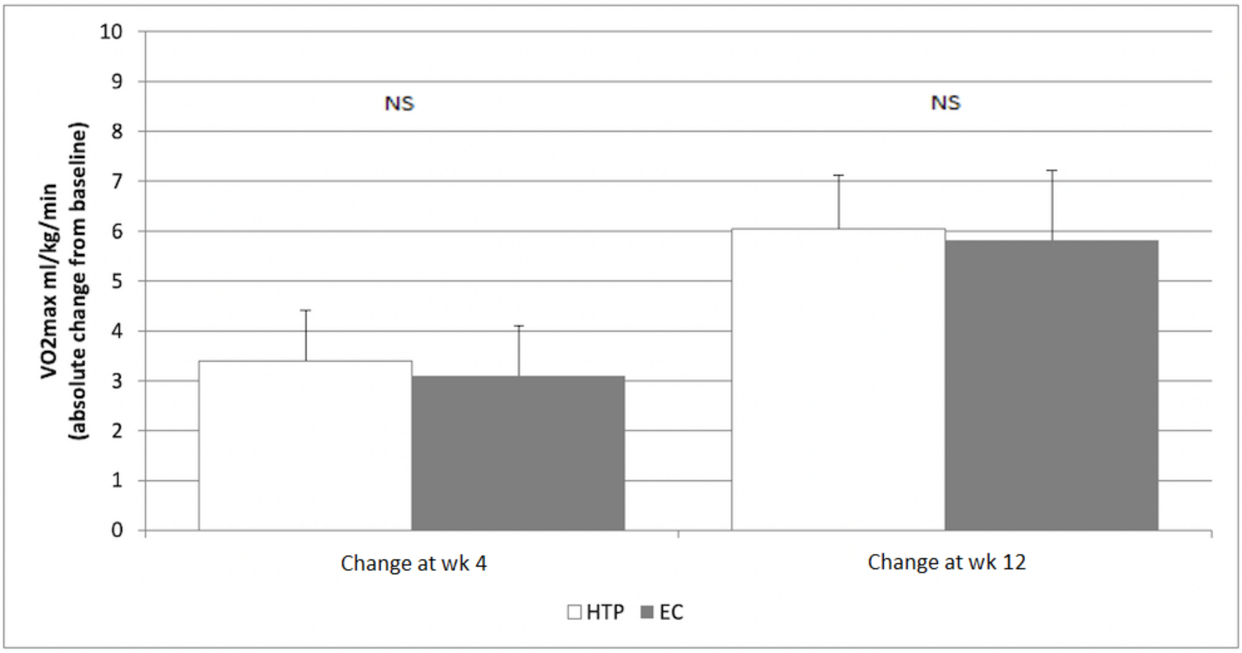
Consumption patterns of conventional tobacco cigarettes, vaping products, and HTPs throughout the study are illustrated in Multimedia Appendices 6 and 7.

Evaluation of Participant Well-being

EQ-5D-5L and EQ VAS results are summarized in Multimedia Appendix 8. No significant changes were observed between the 2 study groups. Within-group analyses for both IQOS-HTP and JustFog-EC showed small but significant changes in all EQ-5D-5L domains, with the exception of domain 2. Regarding EQ VAS, within-group analyses for both IQOS-HTP and JustFog-EC showed small but significant changes from baseline (IQOS-HTP and JustFog-EC; *P*<.001 for all comparisons).

Changes in exercise tolerance between study products were not significant (Figure 4; Multimedia Appendix 8). However, a significant improvement from baseline was observed after switching to the combustion-free products under investigation. For JustFog-EC, reported changes of 2.6 and 7.0 mL/kg/min were noted at week 4 and week 12, respectively (*P*<.001), while for IQOS-HTP, reported changes of 3.4 and 6.4 mL/kg/min were noted at week 4 and week 12, respectively (*P*=.007) (Figure 4; Multimedia Appendix 8). These $\dot{V}O_2$ max improvements were consistently greater than the minimum clinically important difference (MCID) defined as an improvement in the anaerobic threshold of at least 2 mL O_2/kg/min.

Figure 4. Chester step test results. EC: electronic cigarette; HTP: heated tobacco product; NS: not significant.



Predictors of Smoking Abstinence

The results of the abovementioned logistic regression model showed the following evidence: males were less likely to achieve CAR weeks 4-12 compared to females (OR 0.457, 95% CI 0.249-0.840), subjects who had a high consumption of the product were likely to achieve CAR weeks 4-12 compared to those who had a low consumption of the product (OR 0.450, 95% CI 0.212-0.956), and subjects who had a high craving reduction were likely to achieve CAR weeks 4-12 compared to those who had a low craving reduction (OR 0.391, 95% CI 0.186-0.825).

Adverse Events

The reported numbers of adverse events are listed in Multimedia Appendix 9. Most adverse events were rated as mild or moderate and did not led to discontinuation of product use in either study group. In general, the most commonly reported adverse events were cough and reduced physical fitness. Oropharyngeal irritation and dyspnea were more frequent in the EC group than in the HTP group. No serious adverse events were reported during the study. No significant changes in the mean resting heart rate, blood pressure, and BMI during product use were observed between and within study groups.

Discussion

HTP use elicited a marked reduction in cigarette consumption, resulting in almost 40% abstinence from smoking by the end of the study. When present, adverse events were mild and transitory. This is the first study to directly compare ECs to HTPs, showing comparable effectiveness and tolerability between the JustFog-EC and IQOS-HTP study groups.

RCT findings for IQOS-HTP are in agreement with the findings of an American Cancer Society analysis that found a substantial decline in cigarette sales after the introduction of IQOS in Japan [34] and an observational study of COPD patients that reported substantially attenuated or ceased cigarette consumption in the long term after switching to IQOS [21], but are different from the findings of a recent online survey of Korean adults showing a low probability of quitting among IQOS users [35].

Multiple factors contributed to the high quit rate observed among IQOS-HTP and JustFog-EC users: (1) participants were keen to switch to combustion-free NDAs; (2) personalized motivational counseling was administered by psychologists proficient in both smoking cessation and harm reduction at each study visit; (3) top selling products in their respective category (ie, JustFog for ECs and IQOS for HTPs) were given for free in the study and most participants found these products appealing; (4) study products had a pharmacokinetic profile of nicotine uptake mimicking that of conventional cigarettes [36,37]; (5) participants perceived the study products as less harmful than their own cigarettes; and (6) regular use of study products relieved cigarette-induced symptoms and improved exercise tolerance.

The same factors might have contributed to the comparable effectiveness between ECs and HTPs and to the progressive transition from dual use to solo use by the end of the intervention phase. Dual use is known for being a common transitory state, with transitions to solo use taking variable time to occur [38,39]. In our switching trial, IQOS users reported increased dual use in the first 2 weeks compared to JustFog users, but this quickly stabilized by week 4. This could indicate different learning curves for the 2 products.

For this study, we selected the top selling products in their respective category (ie, ECs and HTPs). Both IQOS-HTP and JustFog-EC performed well in the study, as technical issues (eg, malfunctions) were relatively uncommon. Participants enjoyed using both study products, with mCEQ scores showing mild to moderate positive responses in terms of product acceptability, cigarette craving reduction, and physical and psychological reward; mSCAS scores indicating a moderately pleasant user experience; and consumption data revealing regular constant product use throughout the study. This is consistent with the notion that a positive sensorial experience and product enjoyment can contribute to the effectiveness of combustion-free products in terms of cessation outcomes [40-42]. Analyses of PRI-P scoring showed that the study products were perceived to be much less harmful than combustible cigarettes, confirming findings from previous studies [43,44]. IQOS-HTP was perceived to be slightly riskier than JustFog-EC, and in the authors' opinion, this is probably because IQOS shows marked similarities with conventional cigarettes. Moreover, regular use of the study products provided adequate control of cravings (thereby serving as an effective method of relapse prevention), reduced symptoms, and had an overall positive impact on physical fitness, with similar improvements for IQOS-HTP and JustFog-EC. This may also explain why the trend in quit rates increased over time in this switching study; this is discordant to what is generally observed in standard smoking cessation studies in which success rates decline over time.

Adverse events were mild and did not led to discontinuation of product use in either study group. None of the participants abused the products under investigation in terms of excessive daily consumption. In some participants, HTP use was associated with mild cough and reduced physical fitness in line with previous observations [23]. However, the frequency of these symptoms was much lower by the end of the study compared to baseline. Previous smoking history is a key confounder when evaluating the health effects of combustion-free nicotine alternatives in switching studies, as shown by the progressive reduction in the frequency of symptoms by the end of this study. Oropharyngeal irritation was more frequent in the EC group than in the HTP group, probably because of the relatively high level of propylene glycol (a respiratory irritant) in the vaping products under investigation (formulated in 50% propylene glycol/40% vegetable glycerin/10% H_2O). This common irritative response has been shown to be transient and is of uncertain prognostic value [45].

A clinically relevant improvement in exercise tolerance was observed after switching to the combustion-free products under investigation as early as 4 weeks. Greater improvement was observed at 12 weeks as there was a much higher prevalence of quitters by the end of the intervention phase compared to 4 weeks. This is in agreement with the improvement in the level of exercise tolerance shown in prospective studies of COPD patients who switched to ECs [20] and HTPs [21]. The time-dependent improvement in exercise tolerance that occurs after switching may be explained by the marked decline in carbon monoxide exposure and in carboxyhemoglobin levels following cigarette substitution with combustion-free alternatives [46,47].

The trial had strengths and limitations. First, among the innovative features of this randomized controlled switching study, adherence to the study products was enhanced by offering a selection of different products to choose from according to preference/liking. Three aromas of tobacco sticks and three e-liquid flavors were provided to best match participants' sensorial experiences. Nonetheless, these choices remain limited (ie, only 3 different flavors for each class of products) and product specific, thus reducing the generalizability of the study findings. In addition, multiple flavor use is common among e-cigarette users, and switching between flavors is frequently reported even during daily use [48,49]. Du et al also reported that only 1.8% of regular e-cigarette users were using only 1 flavor on a regular basis [50]. Thus, the effectiveness of vaping products for smoking substitution may be further improved.

Second, after close scrutiny, we chose to offer the best vaping devices and HTPs available on the Italian market at the time of the study. More details about the selection process have been published previously [26]. Obviously, product assignment could not be blinded, and strong product preference (IQOS-HTP vs JustFog-EC) could have introduced an allocation bias. Only 3 subjects dropped out soon after randomization when they learned that their product allocation (ie, JustFog-EC) was not their preferred one (ie, IQOS-HTP). However, we cannot exclude that if JustFog-EC was seen as an inferior option, participants in this study group might have put less effort into their switching attempt than those allocated to IQOS-HTP. Nonetheless, the CARs in the JustFog-EC group were at least as high as previously reported [22].

Third, study products were provided in combination with personalized motivational counseling administered by psychologists proficient in both smoking cessation and harm reduction. Provision of expert guidance in the context of a switching trial conducted at specialized smoking cessation services may limit the generalizability of the study findings. Changes in tobacco/nicotine use behavior and product use will be investigated in a separate follow-up study under real-life conditions.

In conclusion, this study confirmed the effectiveness of ECs for cigarette substitution and smoking cessation [22,51], and revealed for the first time that HTP use can promote abstinence from cigarette smoking in combination with motivational counseling. HTPs provided a comparable experience to ECs. Moreover, these results were paralleled by a marked reduction in reported symptoms. Based on the findings of this study, HTPs may represent a valuable addition to the arsenal of reduced-risk products in terms of their smoking substitution potential, but longer follow-up studies are required to confirm significant and prolonged abstinence from smoking and to determine whether our results can be generalized outside smoking cessation services offering high levels of support.

Acknowledgments

The authors would like to thank the staff at Centro Prevenzione Cura Tabagismo (CPCT) for their invaluable assistance with subject recruitment. We are also grateful to Fabio Cibella for the fruitful scientific advice about the statistical analysis plan and Adrian Shalivari for the refinement of the electronic data capture system. The authors' thoughts and prayers go out to all the victims of the Russia-Ukraine conflict. This research was supported by an investigator-initiated study award by Philip Morris Products SA (PMI.IIS.2016.006). Philip Morris Products SA had no role in the design of the study and will not have any role during its execution, data analysis, or data interpretation, or during writing of the manuscript.

Data Availability

We may share deidentified individual participant-level data that underlie the results reported in this article. Data will be available on receipt of a request detailing the study hypothesis and statistical analysis plan. All requests should be sent to the corresponding author. Based on the scientific rigor of the proposal, the study authors will discuss all requests and decide whether data sharing is appropriate. All applicants will be asked to sign a data access agreement.

Conflicts of Interest

PC, DC, MS, M Migliore, and M Caruso are full-time employees of the University of Catania, Italy. RE is a fixed-term researcher at BIOMETEC, University of Catania, Italy. M Maglia and FB are fixed-term researchers at Centro per la Prevenzione e Cura del Tabagismo, University of Catania, Italy. BB is a full-time employee of ARNAS Garibaldi, Catania, Italy. AP is a full-time employee of Casa di Cura Musumeci-Gecas, Gravina di Catania, Italy. GC is a full-time employee of the Unit of Infectious Diseases, Department of Clinical and Experimental Medicine, University of Messina, Italy. M Ceracchi is an employee at FullCRO, Roma, Italy.

Multimedia Appendix 1
Study diagram.
[DOCX File , 83 KB - publichealth_v9i1e42628_app1.docx]

Multimedia Appendix 2
Study schedule and assessments.
[DOCX File , 20 KB - publichealth_v9i1e42628_app2.docx]

Multimedia Appendix 3
Detailed technical issues.
[DOCX File , 15 KB - publichealth_v9i1e42628_app3.docx]

Multimedia Appendix 4
Summary of the measures of product preference, acceptability, and risk perception.
[DOCX File , 18 KB - publichealth_v9i1e42628_app4.docx]

Multimedia Appendix 5
Risk perception.
[DOCX File , 263 KB - publichealth_v9i1e42628_app5.docx]

Multimedia Appendix 6
Consumption data for study participants (per protocol population).
[DOCX File , 20 KB - publichealth_v9i1e42628_app6.docx]

Multimedia Appendix 7
Average daily consumption.
[DOCX File , 86 KB - publichealth_v9i1e42628_app7.docx]

Multimedia Appendix 8
Summary of the measures of participant well-being.
[DOCX File , 20 KB - publichealth_v9i1e42628_app8.docx]

Multimedia Appendix 9
Adverse Events.
[DOCX File , 16 KB - publichealth_v9i1e42628_app9.docx]

Multimedia Appendix 10
CONSORT 2010 checklist.
[PDF File (Adobe PDF File), 113 KB - publichealth_v9i1e42628_app10.pdf]

References

1. Tobacco. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/tobacco [accessed 2023-03-12]
2. WHO report on the global tobacco epidemic, 2011: warning about the dangers of tobacco. World Health Organization. URL: https://apps.who.int/iris/handle/10665/44616 [accessed 2023-03-12]
3. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. The Health Consequences of Smoking—50 Years of Progress A Report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention (US); 2014.
4. US Dept of Health and Human Services. The Health Benefits of Smoking Cessation: A Report of the Surgeon General. Rockville, MD: US Department of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1990.
5. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004 Jun 22;328(7455):1519. [doi: 10.1136/bmj.38142.554479.ae]
6. O'Leary R, Polosa R. Tobacco harm reduction in the 21st century. DAT 2020 Jul 13;20(3):219-234. [doi: 10.1108/dat-02-2020-0007]
7. Benowitz NL. Nicotine Addiction. N Engl J Med 2010 Jun 17;362(24):2295-2303. [doi: 10.1056/nejmra0809890]
8. Hughes J, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. Addiction 2004 Jan;99(1):29-38. [doi: 10.1111/j.1360-0443.2004.00540.x] [Medline: 14678060]
9. Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), Office on Smoking and Health (US). How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention (US); 2010.
10. Mussulman LM, Scheuermann TS, Faseru B, Nazir N, Richter KP. Rapid relapse to smoking following hospital discharge. Prev Med Rep 2019 Sep;15:100891 [FREE Full text] [doi: 10.1016/j.pmedr.2019.100891] [Medline: 31193919]
11. Polosa R, Rodu B, Caponnetto P, Maglia M, Raciti C. A fresh look at tobacco harm reduction: the case for the electronic cigarette. Harm Reduct J 2013 Oct 04;10(1):19 [FREE Full text] [doi: 10.1186/1477-7517-10-19] [Medline: 24090432]

12. Polosa R, Farsalinos K, Prisco D. Health impact of electronic cigarettes and heated tobacco systems. Intern Emerg Med 2019 Sep 14;14(6):817-820. [doi: 10.1007/s11739-019-02167-4] [Medline: 31414334]
13. Nakama C, Tabuchi T. Use of heated tobacco products by people with chronic diseases: The 2019 JASTIS study. PLoS One 2021 Nov 18;16(11):e0260154 [FREE Full text] [doi: 10.1371/journal.pone.0260154] [Medline: 34793517]
14. Farsalinos KE, Polosa R. Safety evaluation and risk assessment of electronic cigarettes as tobacco cigarette substitutes: a systematic review. Ther Adv Drug Saf 2014 Apr 13;5(2):67-86 [FREE Full text] [doi: 10.1177/2042098614524430] [Medline: 25083263]
15. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on the Review of the Health Effects of Electronic Nicotine Delivery Systems. In: Eaton DL, Kwan LY, Stratton K, editors. Public Health Consequences of E-Cigarettes. Washington, DC: National Academies Press (US); 2018.
16. Statement on the potential toxicological risks from electronic nicotine (and non-nicotine) delivery systems (E(N)NDS – e-cigarettes). Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT). URL: https://cot.food.gov.uk/sites/default/files/2020-09/COT%20E%28N%29NDS%20statement%202020-04.pdf [accessed 2023-03-12]
17. Haziza C, de La Bourdonnaye G, Merlet S, Benzimra M, Ancerewicz J, Donelli A, et al. Assessment of the reduction in levels of exposure to harmful and potentially harmful constituents in Japanese subjects using a novel tobacco heating system compared with conventional cigarettes and smoking abstinence: A randomized controlled study in confinement. Regul Toxicol Pharmacol 2016 Nov;81:489-499 [FREE Full text] [doi: 10.1016/j.yrtph.2016.09.014] [Medline: 27693654]
18. Gale N, McEwan M, Camacho OM, Hardie G, Proctor CJ, Murphy J. Changes in biomarkers after 180 days of tobacco heating product use: a randomised trial. Intern Emerg Med 2021 Nov 01;16(8):2201-2212 [FREE Full text] [doi: 10.1007/s11739-021-02798-6] [Medline: 34196886]
19. Caruso M, Emma R, Distefano A, Rust S, Poulas K, Zadjali F, Replica Project Group. Electronic nicotine delivery systems exhibit reduced bronchial epithelial cells toxicity compared to cigarette: the Replica Project. Sci Rep 2021 Dec 17;11(1):24182 [FREE Full text] [doi: 10.1038/s41598-021-03310-y] [Medline: 34921164]
20. Polosa R, Morjaria JB, Prosperini U, Busà B, Pennisi A, Malerba M, et al. COPD smokers who switched to e-cigarettes: health outcomes at 5-year follow up. Ther Adv Chronic Dis 2020 Oct 10;11:2040622320961617 [FREE Full text] [doi: 10.1177/2040622320961617] [Medline: 33101622]
21. Polosa R, Morjaria JB, Prosperini U, Busà B, Pennisi A, Gussoni G, et al. Health outcomes in COPD smokers using heated tobacco products: a 3-year follow-up. Intern Emerg Med 2021 Apr 23;16(3):687-696 [FREE Full text] [doi: 10.1007/s11739-021-02674-3] [Medline: 33754228]
22. Hartmann-Boyce J, Lindson N, Butler AR, McRobbie H, Bullen C, Begh R, et al. Electronic cigarettes for smoking cessation. Cochrane Database Syst Rev 2022 Nov 17;11(11):CD010216-CD010217 [FREE Full text] [doi: 10.1002/14651858.CD010216.pub7] [Medline: 36384212]
23. Tattan-Birch H, Hartmann-Boyce J, Kock L, Simonavicius E, Brose L, Jackson S, et al. Heated tobacco products for smoking cessation and reducing smoking prevalence. Cochrane Database Syst Rev 2022 Jan 06;1(1):CD013790 [FREE Full text] [doi: 10.1002/14651858.CD013790.pub2] [Medline: 34988969]
24. Xu SS, Meng G, Yan M, Gravely S, Quah ACK, Ouimet J, et al. Reasons for Regularly Using Heated Tobacco Products among Adult Current and Former Smokers in Japan: Finding from 2018 ITC Japan Survey. Int J Environ Res Public Health 2020 Oct 31;17(21):8030 [FREE Full text] [doi: 10.3390/ijerph17218030] [Medline: 33142757]
25. Tompkins CNE, Burnley A, McNeill A, Hitchman SC. Factors that influence smokers' and ex-smokers' use of IQOS: a qualitative study of IQOS users and ex-users in the UK. Tob Control 2021 Jan 15;30(1):16-23 [FREE Full text] [doi: 10.1136/tobaccocontrol-2019-055306] [Medline: 31941822]
26. Caponnetto P, Caruso M, Maglia M, Emma R, Saitta D, Busà B, et al. Non-inferiority trial comparing cigarette consumption, adoption rates, acceptability, tolerability, and tobacco harm reduction potential in smokers switching to Heated Tobacco Products or electronic cigarettes: Study protocol for a randomized controlled trial. Contemp Clin Trials Commun 2020 Mar;17:100518 [FREE Full text] [doi: 10.1016/j.conctc.2020.100518] [Medline: 31956726]
27. Rosen MI, Ryan C, Rigsby M. Motivational Enhancement and MEMS Review to Improve Medication Adherence. Behav. change 2012 Feb 22;19(4):183-190. [doi: 10.1375/bech.19.4.183]
28. Miller WR, Rollnick S. Motivational Interviewing: Preparing People for Change, 2nd Edition. New York, NY: Guilford Press; 2002.
29. Fagerström K. Determinants of tobacco use and renaming the FTND to the Fagerstrom Test for Cigarette Dependence. Nicotine Tob Res 2012 Jan 24;14(1):75-78. [doi: 10.1093/ntr/ntr137] [Medline: 22025545]
30. Cappelleri JC, Bushmakin AG, Baker CL, Merikle E, Olufade AO, Gilbert DG. Confirmatory factor analyses and reliability of the modified cigarette evaluation questionnaire. Addict Behav 2007 May;32(5):912-923. [doi: 10.1016/j.addbeh.2006.06.028] [Medline: 16875787]
31. Murray H, McHugh R, Rowley T, Sirota A, Otto M. Development and validation of the smoking cue appeal survey. Subst Use Misuse 2010 Nov;45(13):2152-2161 [FREE Full text] [doi: 10.3109/10826084.2010.481771] [Medline: 20469971]

32. Cano S, Chrea C, Salzberger T, Alfieri T, Emilien G, Mainy N, et al. Development and validation of a new instrument to measure perceived risks associated with the use of tobacco and nicotine-containing products. Health Qual Life Outcomes 2018 Sep 21;16(1):192 [FREE Full text] [doi: 10.1186/s12955-018-0997-5] [Medline: 30241527]
33. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
34. Stoklosa M, Cahn Z, Liber A, Nargis N, Drope J. Effect of IQOS introduction on cigarette sales: evidence of decline and replacement. Tob Control 2020 Jul 17;29(4):381-387. [doi: 10.1136/tobaccocontrol-2019-054998] [Medline: 31209129]
35. Kim J, Lee S, Kimm H, Lee J, Lee C, Cho H. Heated tobacco product use and its relationship to quitting combustible cigarettes in Korean adults. PLoS One 2021 May 7;16(5):e0251243 [FREE Full text] [doi: 10.1371/journal.pone.0251243] [Medline: 33961641]
36. Phillips-Waller A, Przulj D, Pesola F, Smith K, Hajek P. Nicotine Delivery and User Ratings of IQOS Heated Tobacco System Compared With Cigarettes, Juul, and Refillable E-Cigarettes. Nicotine Tob Res 2021 Oct 07;23(11):1889-1894 [FREE Full text] [doi: 10.1093/ntr/ntab094] [Medline: 33983450]
37. Goldenson NI, Augustson EM, Chen J, Shiffman S. Pharmacokinetic and subjective assessment of prototype JUUL2 electronic nicotine delivery system in two nicotine concentrations, JUUL system, IQOS, and combustible cigarette. Psychopharmacology (Berl) 2022 Mar 20;239(3):977-988 [FREE Full text] [doi: 10.1007/s00213-022-06100-0] [Medline: 35184228]
38. Brouwer AF, Jeon J, Hirschtick JL, Jimenez-Mendoza E, Mistry R, Bondarenko IV, et al. Transitions between cigarette, ENDS and dual use in adults in the PATH study (waves 1-4): multistate transition modelling accounting for complex survey design. Tob Control 2020 Nov 16;31(3):424-431 [FREE Full text] [doi: 10.1136/tobaccocontrol-2020-055967] [Medline: 33199541]
39. Conner TS, Zeng J, Blank M, He V, Hoek J. A Descriptive Analysis of Transitions from Smoking to Electronic Nicotine Delivery System (ENDS) Use: A Daily Diary Investigation. Int J Environ Res Public Health 2021 Jun 10;18(12):6301 [FREE Full text] [doi: 10.3390/ijerph18126301] [Medline: 34200773]
40. Dawkins L, Turner J, Roberts A, Soar K. 'Vaping' profiles and preferences: an online survey of electronic cigarette users. Addiction 2013 Jun;108(6):1115-1125 [FREE Full text] [doi: 10.1111/add.12150] [Medline: 23551515]
41. Berg CJ, Barr DB, Stratton E, Escoffery C, Kegler M. Attitudes toward E-Cigarettes, Reasons for Initiating E-Cigarette Use, and Changes in Smoking Behavior after Initiation: A Pilot Longitudinal Study of Regular Cigarette Smokers. Open J Prev Med 2014 Oct;4(10):789-800 [FREE Full text] [doi: 10.4236/ojpm.2014.410089] [Medline: 25621193]
42. DiPiazza J, Caponnetto P, Askin G, Christos P, Maglia MLP, Gautam R, et al. Sensory experiences and cues among E-cigarette users. Harm Reduct J 2020 Oct 15;17(1):75 [FREE Full text] [doi: 10.1186/s12954-020-00420-0] [Medline: 33059666]
43. East KA, Tompkins CNE, McNeill A, Hitchman SC. 'I perceive it to be less harmful, I have no idea if it is or not:' a qualitative exploration of the harm perceptions of IQOS among adult users. Harm Reduct J 2021 Apr 13;18(1):42 [FREE Full text] [doi: 10.1186/s12954-021-00490-8] [Medline: 33849549]
44. Sutanto E, Miller C, Smith D, O'Connor R, Gravely S, Hammond D, et al. Perceived relative harm of heated tobacco products (IQOS), e-cigarettes, and cigarettes among adults in Canada: Findings from the ITC Project. Tob Induc Dis 2020 Sep 21;18(September):81-85 [FREE Full text] [doi: 10.18332/tid/127233] [Medline: 33013277]
45. Farsalinos K, Romagna G, Tsiapras D, Kyrzopoulos S, Voudris V. Characteristics, perceived side effects and benefits of electronic cigarette use: a worldwide survey of more than 19,000 consumers. Int J Environ Res Public Health 2014 Apr 22;11(4):4356-4373 [FREE Full text] [doi: 10.3390/ijerph110404356] [Medline: 24758891]
46. Caponnetto P, Maglia M, Prosperini G, Busà B, Polosa R. Carbon monoxide levels after inhalation from new generation heated tobacco products. Respir Res 2018 Aug 31;19(1):164 [FREE Full text] [doi: 10.1186/s12931-018-0867-z] [Medline: 30170593]
47. Beatrice F, Massaro G. Exhaled Carbon Monoxide Levels in Forty Resistant to Cessation Male Smokers after Six Months of Full Switch to Electronic Cigarettes (e-Cigs) or to A Tobacco Heating Systems (THS). Int J Environ Res Public Health 2019 Oct 15;16(20):3916 [FREE Full text] [doi: 10.3390/ijerph16203916] [Medline: 31618949]
48. Diamantopoulou E, Barbouni A, Merakou K, Lagiou A, Farsalinos K. Patterns of e-cigarette use, biochemically verified smoking status and self-reported changes in health status of a random sample of vapeshops customers in Greece. Intern Emerg Med 2019 Sep 11;14(6):843-851. [doi: 10.1007/s11739-018-02011-1] [Medline: 30635832]
49. Farsalinos KE, Romagna G, Tsiapras D, Kyrzopoulos S, Spyrou A, Voudris V. Impact of flavour variability on electronic cigarette use experience: an internet survey. Int J Environ Res Public Health 2013 Dec 17;10(12):7272-7282 [FREE Full text] [doi: 10.3390/ijerph10127272] [Medline: 24351746]
50. Du P, Bascom R, Fan T, Sinharoy A, Yingst J, Mondal P, et al. Changes in Flavor Preference in a Cohort of Long-Term Electronic Cigarette Users. Annals ATS 2020 May;17(5):573-581. [doi: 10.1513/annalsats.201906-472oc]
51. Tobacco: preventing uptake, promoting quitting and treating dependence. NICE. URL: https://www.nice.org.uk/guidance/ng209/resources/tobacco-preventing-uptake-promoting-quitting-and-treating-dependence-pdf-66143723132869 [accessed 2023-03-12]

Abbreviations

CAR: continuous abstinence rate
CONSORT: Consolidated Standards of Reporting Trials
COPD: chronic obstructive pulmonary disease
CRR: continuous reduction rate
EC: electronic cigarette
eCO: exhaled carbon monoxide
HTP: heated tobacco product
mCEQ: modified Cigarette Evaluation Questionnaire
mSCAS: modified Smoking Cue Appeal Survey
NDA: nicotine delivery alternative
OR: odds ratio
PRI-P CC: Perceived Risk Instrument for conventional cigarettes
PRI-P RRP: Perceived Risk Instrument for reduced risk products

---

Edited by A Mavragani, G Eysenbach; submitted 12.09.22; peer-reviewed by N Goldenson, K Farsalinos; comments to author 18.11.22; revised version received 07.12.22; accepted 04.02.23; published 04.04.23.

Please cite as:
Caponnetto P, Campagna D, Maglia M, Benfatto F, Emma R, Caruso M, Caci G, Busà B, Pennisi A, Ceracchi M, Migliore M, Signorelli M
Comparing the Effectiveness, Tolerability, and Acceptability of Heated Tobacco Products and Refillable Electronic Cigarettes for Cigarette Substitution (CEASEFIRE): Randomized Controlled Trial
JMIR Public Health Surveill 2023;9:e42628
URL: https://publichealth.jmir.org/2023/1/e42628
doi:10.2196/42628
PMID:37014673

Original Paper

Health Care–Seeking Behaviors, Disease Progression, Medications, Knowledge of, and Attitudes Toward Systemic Lupus Erythematosus in China: Cross-sectional Survey Study

Zonglin Dai[1], MPH, PhD; Xinxiang Huang[2], MMed; Fei Yuan[3], MMed; Tianwang Li[4], PhD; Baozhao Xie[5], MMed; He Lin[6], BMed; Pingting Yang[7], PhD; Xueyi Li[8], MMed; Shuiming Xu[9], BMed; Jinjun Zhao[10], PhD; Yukai Wang[11], PhD; Xiang Peng[12], PhD; Simin Wei[13], BMed; Wei Huang[14], MMed; Jingyang Li[15], MMed; Jing Liang[16], BMed; Xiuhua Liu[17], BMed; Yongliang Chu[18], MMed; Zhiming Zhang[19], BMed; Renpeng Zhang[20], BMed; Eric H Y Lau[1,21*], PhD; Zhiming Lin[22*], PhD

[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, Hong Kong

[2]Department of Rheumatology and Immunology, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[3]Department of Rheumatology, Dongguan People's Hospital, Dongguan, China

[4]Department of Rheumatology and Immunology, Guangdong Second Provincial General Hospital, Guangzhou, China

[5]Department of Rheumatology and Immunology, The Seventh Affiliated Hospital, Guangxi Medical University, Wuzhou, China

[6]Department of Rheumatology and Immunology, Fujian Provincial Hospital, Shengli Clinical Medical College of Fujian Medical University, Fujian, China

[7]Department of Rheumatology and Immunology, The First Affiliated Hospital of China Medical University, Shenyang, China

[8]Department of Rheumatology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

[9]Department of Rheumatology and Immunology, Ganzhou Municipal Hospital, Ganzhou, China

[10]Department of Rheumatology, Nanfang Hospital, Southern Medical University, Guangzhou, China

[11]Division of Rheumatology, Shantou Central Hospital, Shantou, China

[12]Department of Rheumatology and Immunology, The Sixth Affiliated Hospital of Guangzhou Medical University (Qingyuan People's Hospital), Qingyuan, China

[13]Division of Rheumatology, Guigang People's Hospital, Guigang, China

[14]Department of Rheumatology and Immunology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, China

[15]Division of Rheumatology, The Affiliated Zhuzhou Hospital of Xiangya Medical College of Central South University, Zhuzhou, China

[16]Department of Rheumatology, Huadu Affiliated Hospital of Southern Medical University, Guangzhou, China

[17]Division of Rheumatology, Liuzhou Worker's Hospital, The Fourth Affiliated Hospital of Guangxi Medical University, Liuzhou, China

[18]Division of Rheumatology, Guangdong Province Hospital of Traditional Chinese Medicine, Zhuhai Branch, Zhuhai, China

[19]Division of Rheumatology, The Second Affiliated Hospital of Fujian University of Traditional Chinese Medicine, Fuzhou, China

[20]Division of Rheumatology, The First People's Hospital of Zhaotong, Zhaotong, China

[21]Laboratory of Data Discovery for Health, Hong Kong Science Park, Hong Kong, Hong Kong

[22]Department of Rheumatology and Immunology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Zhiming Lin, PhD
Department of Rheumatology and Immunology
The Third Affiliated Hospital of Sun Yat-sen University
600 Tianhe Road
Guangzhou, 510630
China
Phone: 86 13632346950
Email: lzm-zj99@163.com

Abstract

Background: Systemic lupus erythematosus (SLE) is a systemic autoimmune disease involving multiple organs throughout the body. The health care–seeking behaviors, disease progression of SLE, and patients' knowledge of and attitudes toward SLE have not been characterized in China.

Objective: The aim of this study was to depict the health care–seeking behaviors, disease progression, and medications in patients with SLE and to examine the factors associated with their disease flares, knowledge, and attitudes toward SLE in China.

Methods: We conducted a cross-sectional survey in 27 provinces in China. Descriptive statistical methods were used to depict the demographic characteristics, health care–seeking behaviors, medications, and health status. Multivariable logistic regression models were used to identify the factors associated with disease flares, medication changes, and attitudes toward SLE. An ordinal regression model was used to examine the factors associated with the knowledge of the treatment guidelines.

Results: We recruited 1509 patients with SLE, and 715 had lupus nephritis (LN). Approximately 39.96% (603/1509) of the patients with SLE were primarily diagnosed with LN, and 12.4% (112/906) developed LN (mean time 5.2 years) from non-LN. Patients whose registered permanent residences or workplaces in other cities from the same province and adjacent provinces seeking health care accounted for 66.9% (569/850) and 48.8% (479/981) of the patients with SLE in the provincial capital cities, respectively. Mycophenolate mofetil was the most commonly used immunosuppressive drug in patients without LN (185/794, 23.3%) and patients with LN (307/715, 42.9%). Femoral head necrosis (71/228, 31.1%) and hypertension (99/229, 43.2%) were the most common adverse event (AE) and chronic disease during treatment, respectively. Change of hospitals for medical consultation (odds ratio [OR] 1.90, 95% CI 1.24-2.90) and development of 1 chronic disease (OR 3.60, 95% CI 2.04-6.24) and AE (OR 2.06, 95% CI 1.46-2.92) and more were associated with disease flares. A pregnancy plan (OR 1.58, 95% CI 1.18-2.13) was associated with changes in medication. Only 242 (16.03%) patients with SLE were familiar with the treatment guidelines, and patients with LN tended to be more familiar with the disease (OR 2.20, 95% CI 1.81-2.68). After receiving treatment, 891 (59.04%) patients changed their attitudes toward SLE from fear to acceptance, and patients with college education or higher (OR 2.09, 95% CI 1.10-4.04) were associated with a positive attitude toward SLE.

Conclusions: A large proportion of patients seeking health care in the provincial capital cities of China migrated from other cities. Persistent monitoring of potential AEs and chronic diseases during SLE treatment and managing patients who changed hospitals for medical consultation are essential for controlling disease flares. Patients had insufficient knowledge about SLE treatment guidelines and would benefit from health education to maintain a positive attitude toward SLE.

(JMIR Public Health Surveill 2023;9:e44541) doi:10.2196/44541

KEYWORDS

systemic lupus erythematosus; health care–seeking behaviors; disease progression; medications; knowledge; attitudes

Introduction

Systemic lupus erythematosus (SLE) is a systemic autoimmune disease involving multiple organs throughout the body [1]. With continuous improvements in diagnosis and treatment, the 5-year survival rate of patients with SLE has increased significantly to more than 90% in the past 2 decades [2]. In China, different regions have different levels of medical resources, and the provincial capital cities have better medical resources, which attract patients with SLE from other cities in the province or surrounding provinces [3,4]. Patients may also switch hospitals during treatment to seek better health care in China. However, the variation in medical insurance reimbursements and the pursuit of health care in different cities will result in a higher economic burden to patients, including the costs of transportation and time [5,6]. Therefore, it is crucial to study patients’ health care–seeking behaviors to provide a basis and guidance for the allocation of medical resources in China in the future.

Disease flare is a state of increased disease activity and has been reported to occur in more than 20% of the patients within 2 years after remission [7-10]. Disease progression, including disease flares, has become the most common cause of hospitalization for patients with SLE, and the risk of damage accrual has doubled [11,12]. In addition to 30%-60% of patients diagnosed with lupus nephritis (LN), a common SLE complication at the onset, some patients may develop LN during treatment [13-15]. However, only few studies have examined the disease progression of LN in patients with no LN in China. After the diagnosis of SLE, patients are given medications, including hydroxychloroquine, glucocorticoids, and immunosuppressive and biological agents. Later, the accompanying increase of adverse events (AEs) or chronic diseases may potentially impact the change in these medications [16]. Further, medications may become potentially changed if patients with SLE plan to become pregnant [17-19]. A retrospective study in Japan revealed that 9.8% of the patients with SLE had changed their medications, but the factors associated with the medication change were not assessed [20]. Understanding reasons for the change in the medications may improve the clinical management of patients with SLE in China [21].

The updated 2020 Chinese guidelines for the diagnosis and treatment of SLE and the 2019 Chinese guidelines for the diagnosis and treatment of LN (hereafter referred to as the treatment guidelines) are further divided to standardize the diagnosis and treatment standards, which not only provide a basis for the diagnosis and treatment of SLE but also provide a transparent guide for patients with SLE [21,22]. Knowledge about SLE can increase a patient’s adaptation to SLE and its treatment [23]. Insufficient general medical knowledge of the disease would result in doubts about medications, thereby increasing the risk of poor treatment adherence [24]. Nevertheless, there is a dearth of studies focusing on patients' familiarity with the treatment guidelines, which reflects the patient's knowledge of the disease to some extent. Socioeconomic and psychosocial factors impact the prognosis of SLE [25]. Negative attitudes, including fear and depression, reduce treatment compliance [26]. A study conducted in Germany showed that negative emotion was associated with 11% reduction in treatment adherence in patients with SLE [27]. A study in China reported that patients with SLE who had anxiety and depression were more likely to have a lower quality of life [28]. Hence, it is essential to explore patients' attitudes toward SLE from the time of their diagnosis to treatment in China. To sum up, the purposes of our study were to examine the disease progression and health care–seeking behaviors of patients with SLE, depict their medication use, analyze the factors associated with disease flares, and assess the predictors of familiarity with the treatment guidelines and their attitudes toward SLE.

Methods

Study Design

The data set in our study was based on an electronic-based questionnaire designed by rheumatologists, epidemiologists, statisticians, and clinicians from multiple centers. The survey was primarily answered by 20 patients and adjusted according to their feedback, and the final version consisted of 4 parts and 51 questions. The first part of the survey was the basic introduction to the study, which required the patient to read the relevant operations carefully and be familiar with the purpose of the study. The second part was about the baseline characteristics of the patients; the third part was about the time of diagnosis, health status, and medications; and the last part was about the knowledge of the treatment guidelines and attitudes toward SLE. The set of questions included fill-in-the-blank, multiple-choice questions with only 1 answer, and multiple-choice questions with several possible answers. To compare the regularity of disease occurrence, we designed the survey by using a large number of time indicators (eg, when were you first diagnosed with SLE and when were you first diagnosed with LN). The detailed contents of the survey are shown in Multimedia Appendix 1.

Ethics Approval

Ethics approval for this study was obtained from the ethics committee of the Third Affiliated Hospital of Sun Yat-sen University (reference [2021] 02-312-01) on October 19, 2021. All patients signed the informed consent before participating and answering the questions in this study. Information such as the objectives, methods, and expected outcomes from this study was presented in the consent form. The privacy and confidentiality protection description for the participants was provided in the consent form.

Tool and Data Collection

This survey was designed using a web-based questionnaire tool (WJX), and a quick response code was generated for distribution in hospitals in China. The survey was filled out by patients with SLE who scanned the quick response code under the guidance of the doctors at each visit. The study period was from June 2021 to December 2021, and data were collected from 105 hospitals and 27 provinces. The inclusion criteria were inpatients or outpatients who were diagnosed with SLE and who signed informed consent. Detailed information on the settings is shown in Table S1 of Multimedia Appendix 2.

To ensure the completeness and credibility of the survey, we made each question mandatory to answer. The key questions were asked in a way (eg, from the diagnosis of the disease to treatment, how many chronic diseases have you developed) such that they were verified by the questions in the subsequent sections (eg, the chronic disease(s) you have developed is/are). The survey had to be filled out in the presence of at least one research doctor. If a patient was not familiar with the way of using the tool or did not understand the way of answering in the relevant columns, the doctor would fill it out.

Definitions

Health care–seeking behavior refers to the seeking of health care in the provincial capital cities and changing of hospitals during treatment (ie, patients sought health care in a hospital different from where they were primarily diagnosed and treated). Disease progression refers to disease flares and the development of AEs and chronic diseases. Disease flares refer to developing LN from non-LN and relapse in all patients with SLE. The Systemic Lupus Erythematosus Disease Activity Index was used to measure relapse in patients without LN if the score increased by at least 4. The reoccurrence of proteinuria or rapid increase of serum creatinine levels was considered as relapse for patients diagnosed with LN [29,30]. Changes in medication meant switching therapy among immunosuppressive drugs and biological agents after the initial treatment. Development of AEs or chronic diseases during treatment referred to the number of AEs or chronic diseases that increased during treatment compared to that in the patients before they were diagnosed with SLE. Patients having a pregnancy plan refers to those with a desire and plan to have children.

Knowledge of the disease was measured by familiarity with the treatment guidelines, which consisted of 3 categories: unfamiliar, less familiar, and familiar. Patients who had not heard of the treatment guidelines were considered unfamiliar. Patients who did not read the treatment guidelines but obtained limited diagnostic and treatment information from other sources were considered less familiar. Patients who read the treatment guidelines in detail and judged their own conditions and guided their medication based on the treatment guidelines were regarded to be familiar with the treatment guidelines. Attitudes toward the disease varied from diagnosis to treatment and were categorized into 4 types: fear throughout treatment (fear →fear), from fear to acceptance (fear →acceptance), from acceptance to fear (acceptance→fear), and acceptance throughout treatment (acceptance→acceptance).

Statistical Analyses

Descriptive statistical methods were used to depict the demographic and clinical characteristics of patients with no LN and patients with LN. Proportions were calculated and presented in a graphical and tabular form. Progression of SLE, patients' health behaviors that included self-discontinuation of patients who stopped the treatment prematurely, irregular treatment of patients who had irregular follow-up and medications, knowledge of the treatment guidelines, and attitudes toward the disease were depicted in a Sankey diagram. Multivariable-adjusted logistic regression models were used to explore the factors associated with the development of LN from non-LN, relapse, and changes in medication. An ordinal regression model was adopted to examine the factors associated with the knowledge of the treatment guidelines. One of the key assumptions of the ordinal regression model is the parallel regression assumption, which requires the coefficients in the cumulative binary logistic regression models to be consistent. The Brant test compares the separate fits to binary logistic regression models and hence was suitably used for assessing the parallel regression assumption for ordinal regression [31]. A logistic regression model was used to investigate the factors associated with attitudes toward the disease.

Results

Characteristics of the Patients

The data of 1509 patients were collected: 794 (52.62%) patients without LN and 715 (47.38%) patients with LN, with a mean age of 34.9 years and 34.6 years, respectively (Table 1). The mean age of the patients with a primary diagnosis of SLE with non-LN and LN was 29 years and 28.4 years, respectively; 1412 (93.57%) patients with SLE were females, and 616 (40.82%) had a college or higher education. The monthly income of the patients with LN was lower than that of the patients without LN (*P*=.01). Relapse (*P*<.001), changes in medication (*P*<.001), and development of AEs (*P*=.003) and chronic diseases (*P*<.001) in patients with LN were higher than those in patients without LN. Only 242 (16.03%) patients with SLE were familiar with the treatment guidelines, and patients with LN were more familiar with the treatment guidelines than those without LN (*P*<.001); 891 (59.04%) patients with SLE changed their attitudes toward the disease from fear at primary diagnosis to acceptance during the treatment, whereas the proportion of patients' attitude changing from acceptance to fear was higher in patients with LN (47/715, 6.6%) than in patients without LN (27/794, 3.4%). Of the 1509 patients with SLE, 489 (32.41%) sought health care in a hospital different from the hospital where they were primarily diagnosed and treated, and 719 (47.64%) had a pregnancy plan. The detailed information of the patients is presented in Table 1.

Figure 1 shows the progression of SLE, health behaviors, and the proportion of patients regarding their familiarity with the disease and attitudes. The mean time for the treatment for patients was 6.4 years. Approximately 39.96% (603/1509) of the patients with SLE were primarily diagnosed with LN, and 12.4% (112/906) of the patients without LN developed LN in the later stage, with a mean time of 5.2 years. Self-discontinuation, irregular treatment, and drug resistance in patients who developed LN from non-LN versus those in patients who experienced relapse were 49.1% (55/112) versus 44.4% (242/545), 36.6% (41/112) versus 2.8% (15/545), and 9.8% (11/112) versus 5.3% (29/545), respectively. Approximately 90.58% (1367/1509) of the patients with SLE developed a positive attitude after treatment.

Table 1. Characteristics of the patients with systemic lupus erythematosus (N=1509).

	Patients with no LN[a] (n=794)	Patients with LN (n=715)	*P* value
Age (years), mean (SD)	34.9 (10.4)	34.6 (10)	.63
Age (years) at diagnosis, mean (SD)	29.0 (10.6)	28.4 (10)	.22
Gender, n (%)			.34
Male	46 (5.8)	51 (7.1)	
Female	748 (94.2)	664 (92.9)	
Education, n (%)			.19
Junior high school or lower	262 (33)	260 (36.4)	
Senior high school	191 (24.1)	280 (25.2)	
College or higher	341 (42.9)	275 (38.5)	
Monthly income (¥[b]), n (%)			.01
<3000	352 (44.3)	372 (52)	
3000-4999	294 (37)	231 (32.3)	
≥5000	148 (18.7)	112 (15.7)	
Basic treatment, n (%)			
Hydroxychloroquine	682 (85.9)	596 (83.4)	.20
Glucocorticoids	622 (78.3)	589 (82.4)	.06
Relapse[c], n (%)			<.001
No	568 (71.5)	396 (55.4)	
Yes	226 (28.5)	319 (44.6)	
Changes in medication[d], n (%)			<.001
None	615 (77.5)	464 (64.9)	
Once	121 (15.2)	152 (21.3)	
Twice	39 (4.9)	63 (8.8)	
Three times or more	19 (2.4)	36 (5)	
Change of hospitals[e], n (%)			.60
No	542 (68.3)	478 (66.9)	
Yes	252 (31.7)	237 (33.1)	
Development of adverse events during treatment[f], n (%)			.003
No	694 (87.4)	589 (82.4)	
One	88 (11.1)	98 (13.7)	
Two or more	12 (1.5)	28 (3.9)	
Development of chronic diseases during treatment[f], n (%)			<.001
No	704 (88.7)	576 (80.6)	
One	65 (8.2)	94 (13.1)	
Two or more	25 (3.1)	45 (6.3)	
Pregnancy plan, n (%)			.77
No	419 (52.8)	371 (51.9)	
Yes	375 (47.2)	344 (48.1)	
Knowledge of the treatment guidelines, n (%)			<.001
Unfamiliar	432 (54.4)	245 (34.3)	
Less familiar	263 (33.1)	327 (45.7)	

	Patients with no LN[a] (n=794)	Patients with LN (n=715)	*P* value
Familiar	99 (12.5)	143 (20)	
Attitudes toward the disease[g], n (%)			.03
Fear→fear	35 (4.4)	33 (4.6)	
Fear→acceptance	485 (61.1)	406 (56.8)	
Acceptance→fear	27 (3.4)	47 (6.6)	
Acceptance→acceptance	247 (31.1)	229 (32)	

[a] LN: lupus nephritis.

[b] CNY ¥1=US $0.15.

[c] Systemic Lupus Erythematosus Disease Activity Index ≥4 for patients without LN and reoccurrence of proteinuria in patients diagnosed with LN.

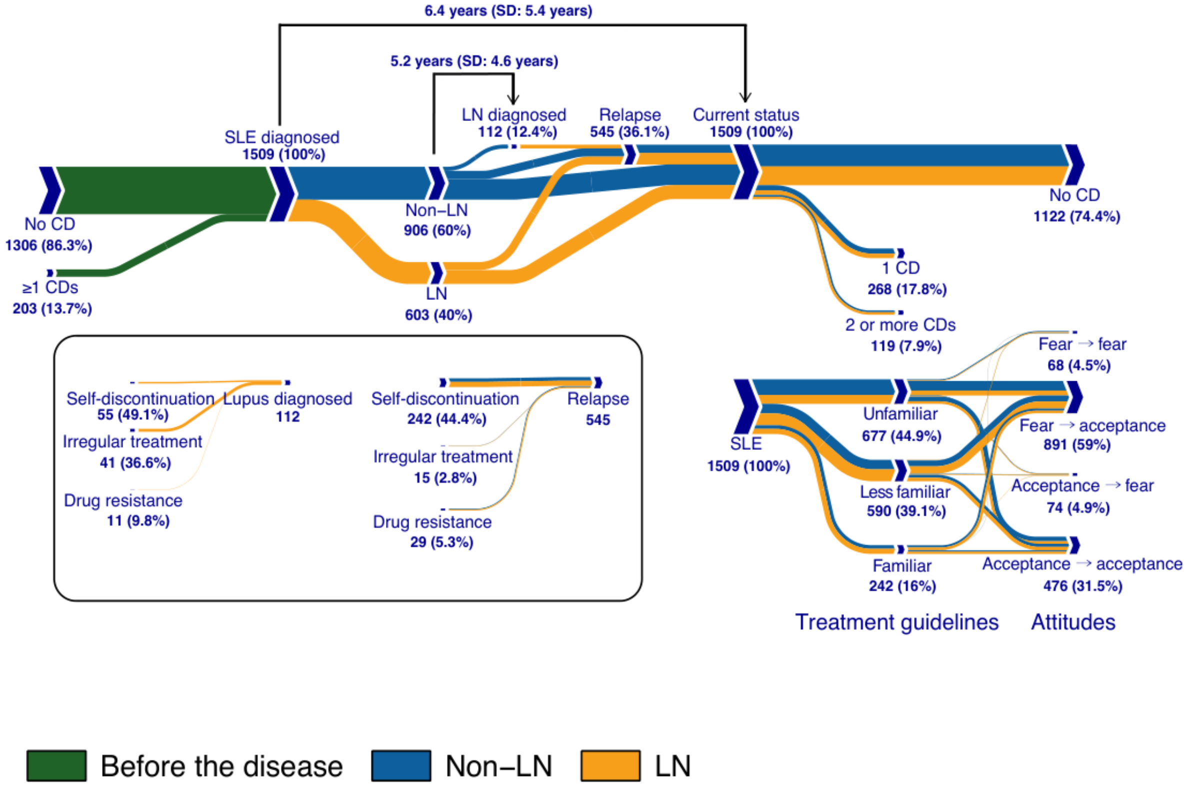
[d] Switching therapy among immunosuppressive drugs and biological agents after the initial treatment due to various reasons such as pregnancy plan, drug effectiveness, or serious adverse events.

[e] Patients sought health care in a hospital different from where they were primarily diagnosed and treated.

[f] The development of adverse events or chronic diseases during the treatment compared to that before patients were diagnosed with systemic lupus erythematosus.

[g] The attitude of patients from diagnosis to treatment: fear throughout treatment (fear→fear), from fear to acceptance (fear→acceptance), from acceptance to fear (acceptance→fear), and acceptance throughout treatment (acceptance→acceptance).

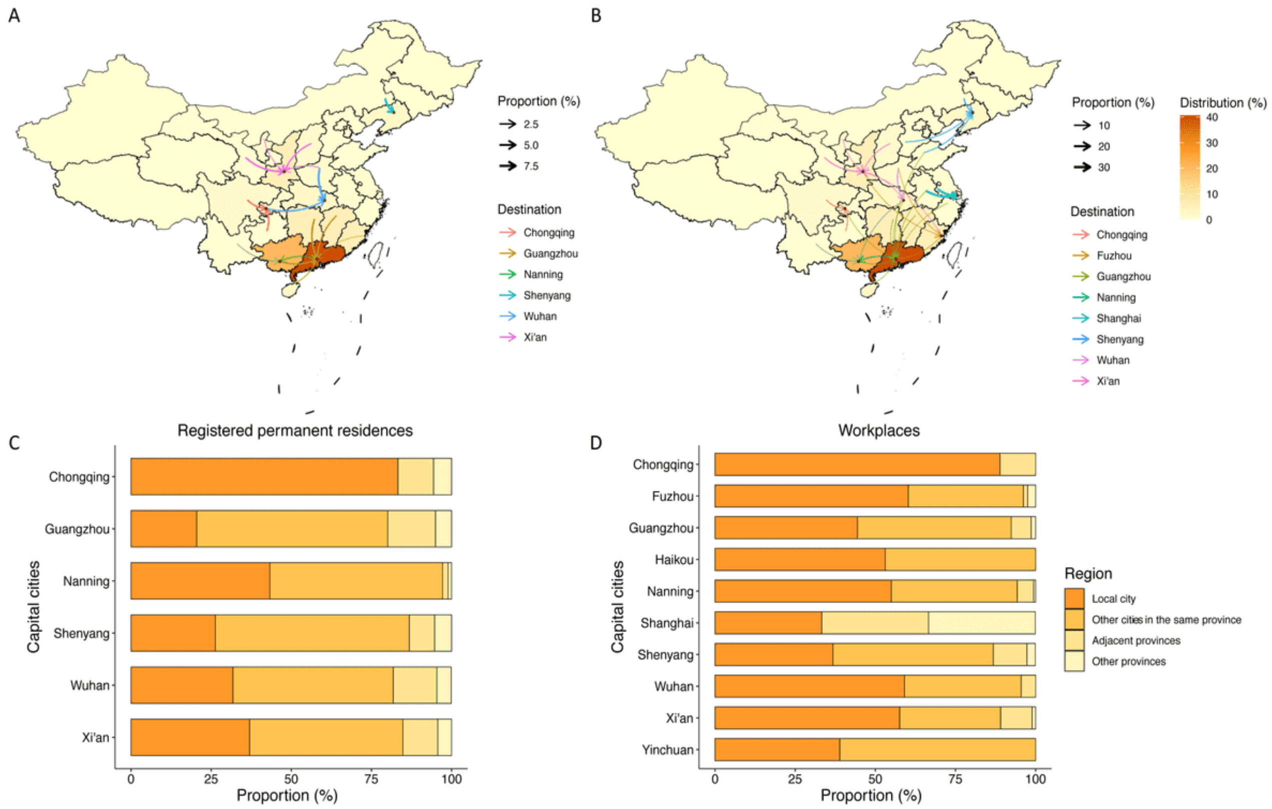
Figure 1. Progression of systemic lupus erythematosus, health behaviors, patients' knowledge of, and attitudes toward the disease. CD: chronic disease; LN: lupus nephritis; SLE: systemic lupus erythematosus.



Health Care–Seeking Behaviors

The distribution of the responses of patients with SLE and the proportions of patients seeking health care in hospitals in provincial capital cities determined by their registered permanent residence or workplaces are presented in Figure 2. Four provincial capital cities (Wuhan, Nanning, Guangzhou, and Shenyang) attracted over half of the total patients whose registered permanent residences were located in other cities from the same province and the adjacent provinces, ranging from 55.6% (99/178) in Nanning to 74.5% (374/502) in Guangzhou. The proportion reached 66.9% (569/850) in the 6 provincial capital cities. Ten provincial capital cities attracted 48.8% (479/981) of the patients whose workplaces were located in other cities in the same province and from the adjacent provinces, among which Guangzhou and Nanning attracted 54.2% (272/502) and 43.8% (78/178) of the patients, respectively.

Figure 2. Distribution of the responses of patients with systemic lupus erythematosus in provinces and the health-seeking behaviors in patients from adjacent provinces when considering the registered permanent residences (A) or the workplaces (B), and the proportion of patients from different cities in provincial capital cities when considering the registered permanent residences (C) or the workplaces (D).

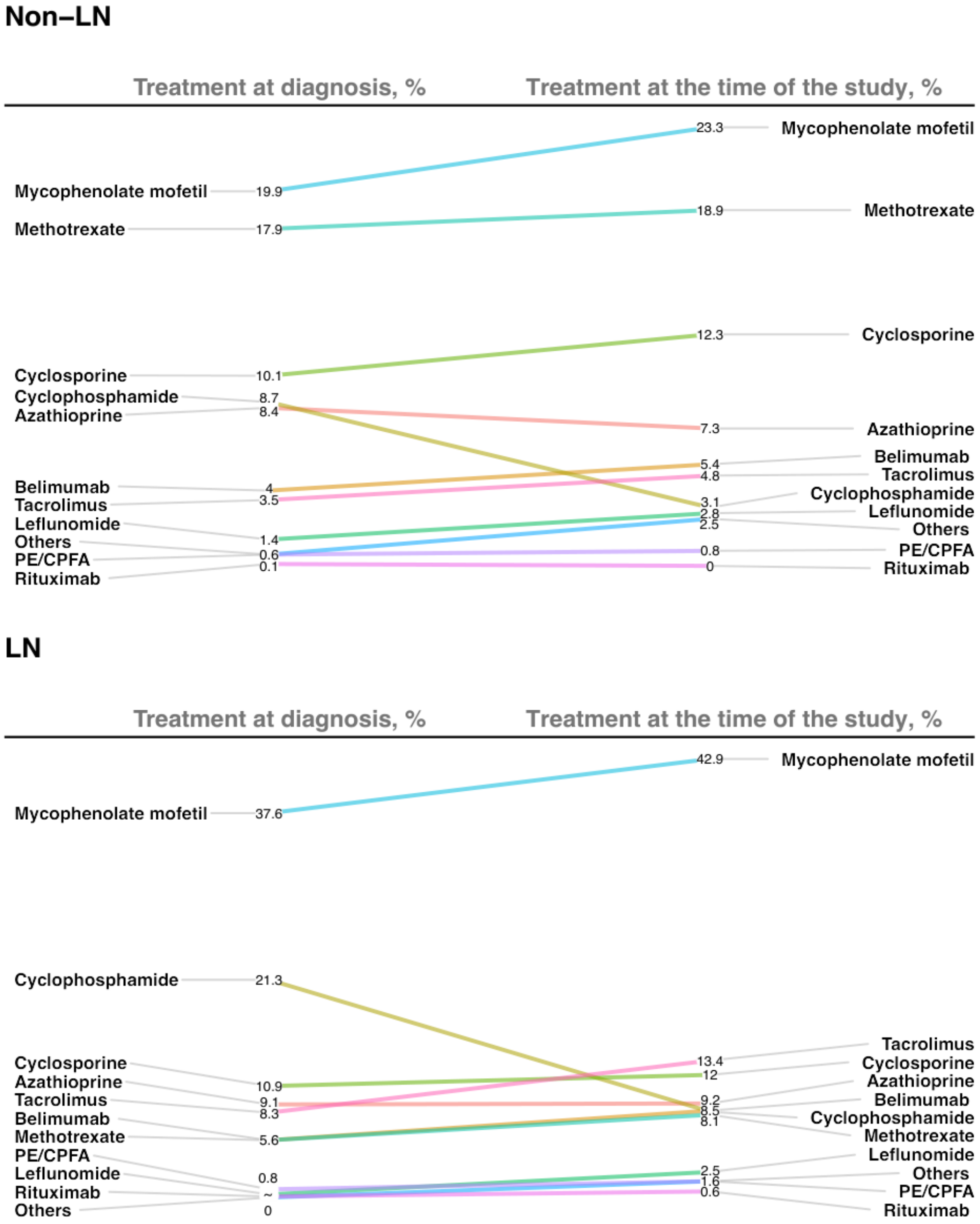


Medications and the Development of AEs and Chronic Diseases

The most common medication used by patients without LN at the time of the study was mycophenolate mofetil (185/794, 23.3%), followed by methotrexate (150/794, 18.9%) and cyclosporine (98/794, 12.3%). These 3 types of immunosuppressive drugs were also the most commonly used in patients without LN at the time of diagnosis, with 19.9% (158/794) on mycophenolate mofetil, 17.9% (142/794) on methotrexate, and 10.1% (80/794) on cyclosporine, respectively (Figure 3). For patients with LN, the most common medication used at the time of the study was mycophenolate mofetil (307/715, 42.9%), followed by tacrolimus (96/715, 13.4%) and cyclosporine (86/715, 12%). Cyclophosphamide (CYC) (152/715, 21.3%) was the second most commonly used immunosuppressive drug at the time of diagnosis, and only 8.4% (60/715) of the patients used it at the time of the study. Belimumab, one of the biological agents, was used less frequently at the time of the study in patients without LN (43/794, 5.4%) and in patients with LN (61/715, 8.5%).

Concerning the development of AEs and chronic diseases, femoral head necrosis (71/228, 31.1%), cataracts (63/228, 27.6%), and retinal macular degeneration (29/228, 12.7%) were the most common serious AEs that occurred during the treatment, while hypertension (99/229, 43.2%), dyslipidemia (45/229, 19.7%), and chronic gastritis (32/229, 14%) were the 3 common chronic diseases developed during the treatment (Table S2 and Table S3 of Multimedia Appendix 2). The highest educational level (college or higher), change of hospitals, development of AEs and chronic diseases, pregnancy plan, and LN were associated with changes in medication.

Figure 3. Medication use in patients with systemic lupus erythematosus at diagnosis and treatment during the time of this study. CPFA: continuous plasma filtration absorption; LN: lupus nephritis; PE: plasma exchange.



Factors Associated With Disease Flares, Knowledge of, and Attitudes Toward SLE

Change of hospitals (odds ratio [OR] 1.90, 95% CI 1.24-2.90) and development of 1 chronic disease (OR 3.60, 95% CI 2.04-6.24) or 2 or more chronic diseases (OR 4.67, 95% CI 2.13-9.95) were associated with LN development (Table 2). In addition, the development of 1 AE and 2 or more AEs during treatment compared with no AE development and LN status compared with non-LN status was associated with a higher risk of relapse. More factors, namely, age and college or higher education, were significantly associated with medication changes.

Only patients with LN (OR 2.20, 95% CI 1.81-2.68) were associated with familiarity with the knowledge of the treatment guidelines (Table 3). The parallel regression assumption was satisfied by the Brant test (Table S4 of Multimedia Appendix 2). Higher education, including senior high school and college or higher, was associated with a positive attitude to the disease, whereas patients with LN were associated with the risk of a negative attitude, together with the development of 1 AE (table 3).

Table 2. Multivariable-adjusted logistic regression analysis of factors associated with disease flares and changes in medication[a].

	LN development[b] (n=906), odds ratio (95% CI)	Relapse[c] (n=1448), odds ratio (95% CI)	Changes in medication[d] (N=1509), odds ratio (95% CI)
Age	0.98 (0.95-1.00)	1.00 (0.99-1.01)	1.02 (1.00-1.03)[e]
Gender			
Male	Reference	Reference	Reference
Female	1.06 (0.47-2.79)	1.00 (0.63-1.62)	1.58 (0.95-2.72)
Education			
Junior high school or lower	Reference	Reference	Reference
Senior high school	1.41 (0.80-2.50)	1.31 (0.97-1.81)	1.33 (0.95-1.86)
College or higher	1.28 (0.73-2.27)	1.16 (0.85-1.58)	1.72 (1.24-2.39)[e]
Monthly income (¥[f])			
<3000	Reference	Reference	Reference
3000-4999	0.71 (0.43-1.15)	0.92 (0.70-1.20)	1.18 (0.89-1.57)
≥5000	0.60 (0.30-1.14)	0.75 (0.52-1.08)	1.03 (0.71-1.50)
Change of hospitals[g]			
No	Reference	Reference	Reference
Yes	1.90 (1.24-2.90)[e]	2.23 (1.75-2.84)[e]	1.95 (1.52-2.50)[e]
Development of adverse events during treatment[h]			
No	Reference	Reference	Reference
One	1.73 (0.98-2.97)	2.06 (1.46-2.92)[e]	2.86 (2.04-4.02)[e]
Two or more	2.13 (0.73-5.87)	4.29 (1.77-12.06)[e]	4.50 (2.19-9.83)[e]
Development of chronic diseases during treatment[h]			
None	Reference	Reference	Reference
One	3.60 (2.04-6.24)[e]	3.43 (2.34-5.09)[e]	1.90 (1.31-2.75)[e]
Two or more	4.67 (2.13-9.95)[e]	2.70 (1.51-4.92)[e]	2.61 (1.50-4.59)[e]
Pregnancy plan			
No	Reference	Reference	Reference
Yes	0.87 (0.52-1.45)	1.01 (0.76-1.34)	1.58 (1.18-2.13)[e]
Health status			
Non–lupus nephritis	N/A[i]	Reference	Reference
Lupus nephritis	N/A	2.03 (1.61-2.56)[e]	1.78 (1.40-2.27)[e]

[a] All factors were mutually adjusted.

[b] LN: lupus nephritis. The development from non–lupus nephritis to lupus nephritis and the numbers in the analysis included patients with non–lupus nephritis and lupus nephritis in the later stage, excluding patients diagnosed with lupus nephritis.

[c] Systemic Lupus Erythematosus Disease Activity Index ≥4 for patients without lupus nephritis and reoccurrence of proteinuria in patients diagnosed with lupus nephritis. Patients who received treatment without achieving remission were excluded from the analysis.

[d] Switching therapy among immunosuppressive drugs and biological agents after the initial treatment due to various reasons such as pregnancy plan, drug effectiveness, or serious adverse events.

[e] Significant estimates (95% CI).

[f] CNY ¥1=US $0.15.

[g] Patients sought health care in a hospital different from where they were primarily diagnosed and treated.

[h] The development of adverse events or chronic diseases during treatment compared to before patients were diagnosed with systemic lupus erythematosus.

[i] Not applicable.

Table 3. Ordinal logistic regression analysis of factors associated with knowledge of the treatment guidelines and logistic regression of factors associated with the attitude toward the disease[a].

	Knowledge of the treatment guidelines	Attitudes toward the disease[b]	
	Familiar or less familiar versus unfamiliar (N=1509), odds ratio (95% CI)	Fear→acceptance versus fear→fear (n=959), odds ratio (95% CI)	Acceptance→fear versus acceptance→acceptance (n=550), odds ratio (95% CI)
Age	0.99 (0.98-1.00)	1.02 (0.99-1.05)	0.96 (0.93-0.99)[c]
Gender			
Male	Reference	Reference	Reference
Female	1.33 (0.89-1.99)	0.53 (0.08-1.82)	2.27 (0.76-9.80)
Education			
Junior high school or lower	Reference	Reference	Reference
Senior high school	1.08 (0.83-1.31)	2.06 (1.05-4.23)[c]	0.63 (0.31-1.22)
College or higher	1.04 (0.80-1.34)	2.09 (1.10-4.04)[c]	0.71 (0.35-1.40)
Monthly income (¥[d])			
<3000	Reference	Reference	Reference
3000-4999	1.19 (0.95-1.50)	1.43 (0.80-2.63)	0.75 (0.40-1.36)
≥5000	1.28 (0.94-1.73)	2.23 (0.90-6.37)	0.75 (0.28-1.80)
Change of hospitals[e]			
No	Reference	Reference	Reference
Yes	0.92 (0.74-1.13)	1.15 (0.68-2.00)	0.99 (0.53-1.76)
Development of adverse events during treatment[f]			
No	Reference	Reference	Reference
One	0.83 (0.61-1.12)	0.93 (0.45-2.13)	2.73 (1.36-5.35)[c]
Two or more	0.80 (0.42-1.52)	0.37 (0.13-1.26)	0.94 (0.05-5.93)
Development of chronic diseases during treatment[f]			
None	Reference	Reference	Reference
One	0.95 (0.69-1.32)	0.66 (0.32-1.44)	1.10 (0.45-2.44)
Two or more	1.25 (0.77-2.02)	0.77 (0.25-2.96)	1.92 (0.58-5.48)
Pregnancy plan			
No	Reference	Reference	Reference
Yes	0.79 (0.63-1.00)	1.21 (0.65-2.27)	1.76 (0.95-3.28)
Health status			
Non-LN[g]	Reference	Reference	Reference
LN	2.20 (1.81-2.68)[c]	1.07 (0.64-1.80)	1.86 (1.11-3.16)[c]

[a] All factors were mutually adjusted.

[b] Attitudes of patients from diagnosis to treatment: fear throughout treatment (fear→fear), from fear to acceptance (fear→acceptance), from acceptance to fear (acceptance→fear), and acceptance throughout treatment (acceptance→acceptance).

[c] Significant estimates (95% CI).

[d] CNY ¥1=US $0.15.

[e] Patients sought health care in a hospital different from where they were primarily diagnosed and treated.

[f] The development of adverse events or chronic diseases during the treatment compared to that before patients were diagnosed with systemic lupus erythematosus.

[g] LN: lupus nephritis.

Discussion

Principal Results

A noteworthy finding in our study is that 55.6% (473/850) of the patients with SLE who sought health care in the provincial capital cities migrated from other cities in the same province when considering their registered permanent residences. When considering the workplace, hospitals in provincial capital cities also lured 42.9% (421/981) of the patients with SLE from the same provinces to seek health care even if the workplaces of these patients were not located in the capital city. In addition, compared with general provincial capitals, high-income provincial capitals also attracted patients from adjacent provinces such as Guangzhou and Xi'an. This can be partly explained by the push-pull theory, as the provincial capital city has more resources in the province, including more hospital options, more medical experts, and better medical technology, which could be considered as the social factors for patients to achieve better quality of life [32]. However, out-of-town visits will increase patients' economic burdens such as transportation and accommodation costs. Therefore, targeted measures should be adopted by various departments for these patients. Hospitals in provincial capital cities can establish a unique medical channel that can provide personal support for people who migrate from other cities to seek medical care through effective identification. The medical insurance departments of each province should simplify the medical reimbursement process for patients from other cities in the same province and the adjacent provinces to reduce the economic burden of patients.

Factors associated with relapse in patients with SLE have been reported in many studies, including the increased anti-double-stranded deoxyribonucleic acid, B lymphocyte stimulator, renal organ involvement, deficiency of hydroxychloroquine, premature discontinuation, and poor compliance [22,33-35]. Consistent with the factors above, we found that patients with LN had a higher risk associated with relapse, among which 49.1% (55/112) of the patients self-discontinued the treatment. In addition, our results showed that the development of 1 or more AEs and chronic diseases during treatment increased the risk of disease flares, including the development of LN from non-LN and relapse in all patients with SLE. The mean time of disease progression in Chinese cohorts was 5.2 years, corresponding to an international inception cohort study with a mean time of 4.6 years [14]. The outcomes in our study may extend the previous finding [14].

A novel finding in our study is that switching from the first hospital where patients were primarily diagnosed and treated was associated with their risk of development of LN and relapse and was significantly associated with changes in their medications. This finding implies that the hospital that the patient first visited may have provided improper treatment. Requesting more meticulous care and treatment for patients with a record of switching hospitals from other cities, especially in the provincial capital cities, is imperative because the major public hospitals have a large number of patients from other cities, with outpatient visits and hospitalizations increasing. A study in 1 hospital in Kunming city, a provincial city in the Yunnan province, showed that outpatient visits and hospitalizations increased from 188,530 and 6620 in 2001 to 369,510 and 12,380 in 2010, respectively [36]. Another hospital in Shanghai showed that outpatient visits increased from 734,923 in 2013 to 881,376 in 2017 [37]. Although medical resources are more concentrated in provincial capital cities, individualized management of patients in those cities, including consideration of the patient's diagnosis and treatment history, at the initial hospital and the patient's home city is indispensable [3,4].

Apart from factors such as older age, higher education, development of at least 1 AE and chronic disease, change of hospitals, and LN development that were associated with changes in medication, nearly half of the patients with SLE had a pregnancy plan, which was also a significant factor (OR 1.58, 95% CI 1.18-2.13). It has been confirmed that using CYC leads to ovarian failure [38]. Our study shows that the use of CYC decreased by 60.5% (92/152) in the current treatment (at the time of the study) compared with that during the primary treatment, and the proportion of patients using CYC among those with a pregnancy plan was significantly lower than that among those without a pregnancy plan (Table S5 of Multimedia Appendix 2). This finding shows that in addition to considering clinical symptoms, the plans and the needs of the patients are also worth considering by doctors in the treatment process.

The familiarity with treatment guidelines reflects a patient's knowledge of the disease to a certain extent [39]. Our results indicated that only 12.5% (99/794) of the patients without LN and 20% (143/715) of the patients with LN were familiar with the treatment guidelines. Moreover, only patients with LN were more familiar with the treatment guidelines than patients without LN. The limited proportion of patients with knowledge of the treatment guidelines in China can be worrisome, as these results revealed 2 issues. First, the knowledge of the disease was inadequate; therefore, the treatment decision was more passive or physician-led. Furthermore, the lack of education and knowledge of SLE would increase the risk of disease flares [24]. Second, patients were more familiar with the disease in a severer state than when the disease was in a mild stage, indicating that patients focused less on their disease in the early stage. A study has shown that patients' knowledge of the disease could be improved by reading the treatment guide [39]. Therefore, the potential solution is to condense and simplify the relevant content of diagnosis and treatment in the treatment guidelines and to distribute the guidelines to doctors and patients in electronic or paper versions in a simple booklet for their study and reference. In addition, patients should be rendered a comprehensive understanding of the disease when the disease status is mild rather than after the disease has worsened.

Patients' attitudes toward SLE, on the whole, reflected acceptance, with approximately 90.58% (1367/1509) of the patients expressing a positive attitude during the treatment period, although nearly half of the patients felt fear at the time of diagnosis. An important factor was patients with LN who had a higher risk of developing a fearful attitude at diagnosis than patients without LN. One way to improve the patient's attitude is by providing health education. Studies have indicated that good education is associated with higher knowledge of diseases and increased self-awareness of one's health and assessment of health care [40-42]. Further, the interaction between doctors and patients can be strengthened, including using the evaluation function of the survey (WJX) to test patients' understanding of the latest research progress on the disease and monitoring patients' attitude changes to the disease at different stages, which can be achieved through a series of scoring and incentive measures [43].

Limitations In This Study

Our findings have some limitations and must be interpreted with caution. First, although we set the time points for collecting variables, the chronological order of the factors and the outcomes might not be explicit. For example, we dynamically described a change in the development of chronic diseases before and after the diagnosis of SLE. However, whether its increase occurs before or after the development of LN or relapse has not been elucidated. This limitation is because, on the one hand, the study design was cross-sectional. On the other hand, even in the existing medical database, the corresponding time point is relatively vague; therefore, reverse causality or simultaneity may appear. Second, a more detailed measurement of health emotions may improve health decision-making [44-46]. Our study lacked such measurement of attitudes toward SLE, and a further study evaluating patients' emotions on a numerical scale may improve the findings' accuracy. Third, our study might have responder bias. Since Guangdong is the most populous province in China, more samples were collected from Guangdong, while samples from other regions were more evenly distributed. Furthermore, recall bias might arise in the data collection process, especially for those who had received long-term treatment for more than 10 years. It was challenging for patients to recall their physical condition and medication regimen before the diagnosis. In order to reduce this limitation, we set the number of entries and specific names in the survey to achieve consistency and to ensure that recall bias was minimized.

Conclusions

In our study, a large proportion of patients seeking health care in provincial capital cities migrated from other cities in the same province and the adjacent provinces. Mycophenolate mofetil was the most commonly used immunosuppressive drug in the treatment of patients without LN and patients with LN at the time of this study. Femoral head necrosis was the most common AE, followed by cataract and retinal macular degeneration, while hypertension, dyslipidemia, and chronic gastritis were the 3 common chronic diseases developed during the treatment. Change of the hospital where patients were primarily diagnosed and treated and the development of AEs and chronic diseases were associated with disease flares and changes in medication. A pregnancy plan was also associated with changes in medication. Patients had insufficient knowledge about the treatment guidelines, and patients without LN tended to be less familiar with the treatment guidelines compared to patients with LN. Patients with SLE who had college education or higher were associated with a positive attitude rather than those who had only a junior high school education or lower. Our findings suggest that disease flares should be controlled by monitoring potential AEs and chronic diseases during treatment and by effectively managing patients who switched hospitals for their medical consultation. In addition, patients would benefit from health education to maintain a positive attitude toward SLE.

Acknowledgments

We thank all the doctors and patients who participated in our study. This work was supported by the Ten and Five Project of the Third Affiliated Hospital of Sun Yat-sen University (grant SW201901) and the Distinguished Young Scholar Candidates Program of the Third Affiliated Hospital of Sun Yat-sen University (grant A2675).

Data Availability

The data are not publicly available due to the potential violation that may compromise the privacy of the research participants. However, data may be available upon proper request from the corresponding author.

Authors' Contributions

ZD contributed to the study design, data analysis, and manuscript drafting. EHYL contributed to the study design, data interpretation, manuscript revision, revision, and approval. ZL contributed to study design, manuscript revision, and approval. The other authors were responsible for data collection and interpretation. All authors critically approved the final manuscript.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The surveys of 7 older patients (age range 67-86 years) among the 1509 participants were filled out by their doctors because these patients could not operate the digital survey or understand the questions. The surveys were answered by doctors in the presence of these patients who also confirmed the responses. There was no conflict of interest for the doctors who assisted the 7 older patients.

Multimedia Appendix 1
Survey used in this study.
[DOCX File , 26 KB - publichealth_v9i1e44541_app1.docx]

Multimedia Appendix 2
Supplementary Tables 1-5. Table S1 shows the settings of patients with systemic lupus erythematosus from registered permanent residences or workplaces in mainland China. Table S2 and Table S3 present the number of adverse events and chronic diseases developed in patients with systemic lupus erythematosus during the treatment, respectively. Brant test for each parameter regarding the parallel regression assumption in an ordinal logistic regression model is reflected in Table S4, and Table S5 is the use of medications in patients with systemic lupus erythematosus regarding the pregnancy plan at the time of the study.
[DOCX File , 26 KB - publichealth_v9i1e44541_app2.docx]

References

1. Cojocaru M, Cojocaru IM, Silosi I, Vrabie CD. Manifestations of systemic lupus erythematosus. Maedica (Bucur) 2011 Oct;6(4):330-336 [FREE Full text] [Medline: 22879850]
2. Tektonidou MG, Lewandowski LB, Hu J, Dasgupta A, Ward MM. Survival in adults and children with systemic lupus erythematosus: a systematic review and Bayesian meta-analysis of studies from 1950 to 2016. Ann Rheum Dis 2017 Dec;76(12):2009-2016. [doi: 10.1136/annrheumdis-2017-211663] [Medline: 28794077]
3. Guo Q, Luo K. Concentration of Healthcare Resources in China: The Spatial-Temporal Evolution and Its Spatial Drivers. Int J Environ Res Public Health 2019 Nov 20;16:23 [FREE Full text] [doi: 10.3390/ijerph16234606] [Medline: 31757034]
4. Nong S, Chen Z. Whither the roads lead to? Estimating association between urbanization and primary healthcare service use with chinese prefecture-level data in 2014. PLoS One 2020;15(6):e0234081 [FREE Full text] [doi: 10.1371/journal.pone.0234081] [Medline: 32492048]
5. Hilmola O, Henttu V. Transportation costs do matter: simulation study from hospital investment decision. JM2 2016 May 09;11(2):560-584. [doi: 10.1108/jm2-10-2013-0053]
6. Jo C. Cost-of-illness studies: concepts, scopes, and methods. Clin Mol Hepatol 2014 Dec;20(4):327-337 [FREE Full text] [doi: 10.3350/cmh.2014.20.4.327] [Medline: 25548737]
7. Carlavilla A, Gutiérrez E, Ortuño T, Morales E, González E, Praga M. Relapse of lupus nephritis more than 10 years after complete remission. Nephrol Dial Transplant 2005 Sep;20(9):1994-1998. [doi: 10.1093/ndt/gfh954] [Medline: 15985510]
8. El Hachmi M, Jadoul M, Lefèbvre C, Depresseux G, Houssiau FA. Relapses of lupus nephritis: incidence, risk factors, serology and impact on outcome. Lupus 2003;12(9):692-696. [doi: 10.1191/0961203303lu444oa] [Medline: 14514132]
9. Kakati S, Teronpi R, Barman B. Frequency, pattern and determinants of flare in systemic lupus erythematosus: A study from North East India. The Egyptian Rheumatologist 2015 Oct;37(4):S55-S59 [FREE Full text] [doi: 10.1016/j.ejr.2015.08.002]
10. Adamichou C, Bertsias G. Flares in systemic lupus erythematosus: diagnosis, risk factors and preventive strategies. Mediterr J Rheumatol 2017 Mar;28(1):4-12 [FREE Full text] [doi: 10.31138/mjr.28.1.4] [Medline: 32185248]
11. Edwards CJ, Lian TY, Badsha H, Teh CL, Arden N, Chng HH. Hospitalization of individuals with systemic lupus erythematosus: characteristics and predictors of outcome. Lupus 2003;12(9):672-676. [doi: 10.1191/0961203303lu452oa] [Medline: 14514129]
12. Ugarte-Gil MF, Acevedo-Vásquez E, Alarcón GS, Pastor-Asurza CA, Alfaro-Lozano JL, Cucho-Venegas JM, GLADEL. The number of flares patients experience impacts on damage accrual in systemic lupus erythematosus: data from a multiethnic Latin American cohort. Ann Rheum Dis 2015 Jun;74(6):1019-1023. [doi: 10.1136/annrheumdis-2013-204620] [Medline: 24525909]
13. Hanly JG, O'Keeffe AG, Su L, Urowitz MB, Romero-Diaz J, Gordon C, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. Rheumatology (Oxford) 2016 Feb;55(2):252-262 [FREE Full text] [doi: 10.1093/rheumatology/kev311] [Medline: 26342222]
14. Mahajan A, Amelio J, Gairy K, Kaur G, Levy RA, Roth D, et al. Systemic lupus erythematosus, lupus nephritis and end-stage renal disease: a pragmatic review mapping disease severity and progression. Lupus 2020 Aug;29(9):1011-1020 [FREE Full text] [doi: 10.1177/0961203320932219] [Medline: 32571142]
15. Saxena R, Mahajan T, Mohan C. Lupus nephritis: current update. Arthritis Res Ther 2011;13(5):240 [FREE Full text] [doi: 10.1186/ar3378] [Medline: 22078716]
16. Fanouriakis A, Bertsias G, Boumpas DT. Hydroxychloroquine dosing in systemic lupus erythematosus: response to 'Comment on the 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus by Fanouriakis by Costedoat-Chalumeau. Ann Rheum Dis 2020 Aug;79(8):e91. [doi: 10.1136/annrheumdis-2019-215615] [Medline: 31201171]
17. Knight CL, Nelson-Piercy C. Management of systemic lupus erythematosus during pregnancy: challenges and solutions. Open Access Rheumatol 2017;9:37-53 [FREE Full text] [doi: 10.2147/OARRR.S87828] [Medline: 28331377]
18. Mok CC, Wong RW. Pregnancy in systemic lupus erythematosus. Postgrad Med J 2001 Mar;77(905):157-165 [FREE Full text] [doi: 10.1136/pmj.77.905.157] [Medline: 11222822]
19. Palmsten K, Simard JF, Chambers CD, Arkema EV. Medication use among pregnant women with systemic lupus erythematosus and general population comparators. Rheumatology (Oxford) 2017 Apr 01;56(4):561-569 [FREE Full text] [doi: 10.1093/rheumatology/kew448] [Medline: 28013193]

20. Miyazaki C, Sruamsiri R, Mahlich J, Jung W. Treatment patterns and medical cost of systemic lupus erythematosus patients in Japan: a retrospective claims database study. J Med Econ 2020 Jul;23(7):786-799 [FREE Full text] [doi: 10.1080/13696998.2020.1740236] [Medline: 32149541]
21. Li M, Zhao Y, Zhang Z, Huang C, Liu Y, Gu J, Chinese Rheumatology Association, National Clinical Research Center for Dermatologic Immunologic Diseases, Chinese Systemic Lupus Erythematosus Treatment Research Group. 2020 Chinese Guidelines for the Diagnosis and Treatment of Systemic Lupus Erythematosus. Rheumatol Immunol Res 2020 Dec;1(1):5-23 [FREE Full text] [doi: 10.2478/rir-2020-0009] [Medline: 36465077]
22. Diagnosis and Treatment Guidelines Writing Group. Chinese guidelines for diagnostic and treatment of lupus nephritis. Natl Med J China 2019 Nov 26:3441-3455 [FREE Full text] [doi: 10.3760/cma.j.issn.0376-2491.2019.44.001]
23. Szymona-Pałkowska K, Janowski K, Pedrycz A, Mucha D, Ambroży T, Siermontowski P, et al. Knowledge of the Disease, Perceived Social Support, and Cognitive Appraisals in Women with Urinary Incontinence. Biomed Res Int 2016;2016:3694792 [FREE Full text] [doi: 10.1155/2016/3694792] [Medline: 28097132]
24. Zhang L, Luan W, Geng S, Ye S, Wang X, Qian L, et al. Lack of patient education is risk factor of disease flare in patients with systemic lupus erythematosus in China. BMC Health Serv Res 2019 Jun 13;19(1):378 [FREE Full text] [doi: 10.1186/s12913-019-4206-y] [Medline: 31196083]
25. Sutcliffe N, Clarke AE, Gordon C, Farewell V, Isenberg DA. The association of socio-economic status, race, psychosocial factors and outcome in patients with systemic lupus erythematosus. Rheumatology (Oxford) 1999 Nov;38(11):1130-1137. [doi: 10.1093/rheumatology/38.11.1130] [Medline: 10556268]
26. Jin J, Sklar GE, Min Sen Oh V, Chuen Li S. Factors affecting therapeutic compliance: A review from the patient's perspective. Ther Clin Risk Manag 2008 Feb;4(1):269-286 [FREE Full text] [doi: 10.2147/tcrm.s1458] [Medline: 18728716]
27. Chehab G, Sauer GM, Richter JG, Brinks R, Willers R, Fischer-Betz R, et al. Medical adherence in patients with systemic lupus erythematosus in Germany: predictors and reasons for non-adherence - a cross-sectional analysis of the LuLa-cohort. Lupus 2018 Sep;27(10):1652-1660. [doi: 10.1177/0961203318785245] [Medline: 30020020]
28. Shen B, Feng G, Tang W, Huang X, Yan H, He Y, et al. The quality of life in Chinese patients with systemic lupus erythematosus is associated with disease activity and psychiatric disorders: a path analysis. Clin Exp Rheumatol 2014;32(1):101-107. [Medline: 24029416]
29. Parikh SV, Nagaraja HN, Hebert L, Rovin BH. Renal flare as a predictor of incident and progressive CKD in patients with lupus nephritis. Clin J Am Soc Nephrol 2014 Feb;9(2):279-284 [FREE Full text] [doi: 10.2215/CJN.05040513] [Medline: 24262502]
30. Pope JE. Measuring flares in systemic lupus erythematosus. Rheumatology (Oxford) 2014 Dec;53(12):2134-2135. [doi: 10.1093/rheumatology/keu274] [Medline: 25065015]
31. Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 1990 Dec;46(4):1171-1178. [Medline: 2085632]
32. Urbański M. Comparing Push and Pull Factors Affecting Migration. Economies 2022 Jan 10;10(1):21. [doi: 10.3390/economies10010021]
33. Petri MA, van Vollenhoven RF, Buyon J, Levy RA, Navarra SV, Cervera R, BLISS-52BLISS-76 Study Groups. Baseline predictors of systemic lupus erythematosus flares: data from the combined placebo groups in the phase III belimumab trials. Arthritis Rheum 2013 Aug;65(8):2143-2153 [FREE Full text] [doi: 10.1002/art.37995] [Medline: 23754628]
34. Hajji M, Harzallah A, Kaaroud H, Barbouch S, Hamida FB, Abdallah TB. Factors associated with relapse of lupus nephritis: A single center study of 249 cases. Saudi J Kidney Dis Transpl 2017;28(6):1349-1355 [FREE Full text] [doi: 10.4103/1319-2442.220863] [Medline: 29265046]
35. Zeng X, Zheng L, Rui H, Kang R, Chen J, Chen H, et al. Risk factors for the flare of systemic lupus erythematosus and its influence on prognosis: a single-center retrospective analysis. Adv Rheumatol 2021 Jul 02;61(1):43. [doi: 10.1186/s42358-021-00202-7] [Medline: 34215349]
36. Ning S, Duan S, Wenning R. The Forecast of the Number of Outpatients and Discharged Patients in the Twelfth Five-year Plan Period. Chinese Medical Record 2011:48-50 [FREE Full text] [doi: 10.3969/j.issn.1672-2566.2011.10.029]
37. Zhou GJ, Jiang XB, Yuan D, Luo JF, Sun H. Seasonal variation trend analysis on outpatient visits in a dermatology hospital from 2013 to 2017. Shanghai Journal of Preventive Medicine 2019:70-73 [FREE Full text] [doi: 10.19428/j.cnki.sjpm.2019.18892]
38. McDermott EM, Powell RJ. Incidence of ovarian failure in systemic lupus erythematosus after treatment with pulse cyclophosphamide. Ann Rheum Dis 1996 Apr;55(4):224-229 [FREE Full text] [doi: 10.1136/ard.55.4.224] [Medline: 8733438]
39. Konttinen YT, Santavirta N, Honkanen V, Sandelin S, Schauman L, Grönblad M. Systemic lupus erythematosus patient guide: influence on knowledge of the disease. Ann Rheum Dis 1991 Dec;50(12):900-902 [FREE Full text] [doi: 10.1136/ard.50.12.900] [Medline: 1768155]
40. Tsou M. Association of Education, Health Behaviors, Concerns, and Knowledge with Metabolic Syndrome among Urban Elderly in One Medical Center in Taiwan. International Journal of Gerontology 2017 Sep;11(3):138-143 [FREE Full text] [doi: 10.1016/j.ijge.2016.09.006]

41. Andersen R, Newman JF. Societal and Individual Determinants of Medical Care Utilization in the United States. Milbank Q 2005;83:1-28 [FREE Full text] [doi: 10.1111/j.1468-0009.2005.00428.x]
42. Raghupathi V, Raghupathi W. The influence of education on health: an empirical assessment of OECD countries for the period 1995-2015. Arch Public Health 2020;78:20 [FREE Full text] [doi: 10.1186/s13690-020-00402-5] [Medline: 32280462]
43. Yu S, Alper HE, Nguyen A, Brackbill RM, Turner L, Walker DJ, et al. The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. BMC Med Res Methodol 2017 Apr 26;17(1):77 [FREE Full text] [doi: 10.1186/s12874-017-0353-1] [Medline: 28446131]
44. Ellis EM, Klein WM, Orehek E, Ferrer RA. Effects of Emotion on Medical Decisions Involving Tradeoffs. Med Decis Making 2018 Oct 20;38(8):1027-1039. [doi: 10.1177/0272989x18806493]
45. Gengler AM. Emotions and Medical Decision-Making. Soc Psychol Q 2020 Apr 16;83(2):174-194. [doi: 10.1177/0190272519876937]
46. Yukalov VI. Quantification of emotions in decision making. Soft Comput 2021 Nov 05;26(5):2419-2436. [doi: 10.1007/s00500-021-06442-5]

Abbreviations

AE: adverse event
CYC: cyclophosphamide
LN: lupus nephritis
OR: odds ratio
SLE: systemic lupus erythematosus

Edited by A Mavragani, T Sanchez; submitted 23.11.22; peer-reviewed by W Caldwell, A Ramaprasad, L Lafrado; comments to author 02.02.23; revised version received 13.02.23; accepted 28.02.23; published 07.04.23.

Please cite as:
Dai Z, Huang X, Yuan F, Li T, Xie B, Lin H, Yang P, Li X, Xu S, Zhao J, Wang Y, Peng X, Wei S, Huang W, Li J, Liang J, Liu X, Chu Y, Zhang Z, Zhang R, Lau EHY, Lin Z
Health Care–Seeking Behaviors, Disease Progression, Medications, Knowledge of, and Attitudes Toward Systemic Lupus Erythematosus in China: Cross-sectional Survey Study
JMIR Public Health Surveill 2023;9:e44541
URL: https://publichealth.jmir.org/2023/1/e44541
doi:10.2196/44541
PMID:37027203

Original Paper

Using the New "Life's Essential 8" Metrics to Evaluate Trends in Cardiovascular Health Among US Adults From 2005 to 2018: Analysis of Serial Cross-sectional Studies

Cheng Li[1*], BA; Yanzhi Li[2*], MSc; Min Zhao[3], MD; Cheng Zhang[4], MD; Pascal Bovet[5], MD; Bo Xi[1], MD

[1]Department of Epidemiology, School of Public Health, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

[2]Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, China

[3]Department of Nutrition and Food Hygiene, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

[4]Key Laboratory of Cardiovascular Remodeling and Function Research, Chinese Ministry of Education, Chinese National Health Commission and Chinese Academy of Medical Sciences, The State and Shandong Province Joint Key Laboratory of Translational Cardiovascu, Shandong University, Jinan, China

[5]Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland

[*]these authors contributed equally

Corresponding Author:
Bo Xi, MD
Department of Epidemiology, School of Public Health, Qilu Hospital
Cheeloo College of Medicine
Shandong University
44 Wenhuaxi Road
Jinan, 250012
China
Phone: 86 88382141
Email: xibo2007@126.com

Abstract

Background: The recently published "Life's Essential 8" (LE8) by the American Heart Association has overcome some limitations in evaluating cardiovascular health (CVH) in the previous "Life's Simple 7."

Objective: We aimed to examine the secular trends in CVH, as assessed by the LE8, in US adults from 2005 to 2018.

Methods: Using cross-sectional data from the National Health and Nutrition Examination Survey between 2005-2006 and 2017-2018, we calculated the age-standardized mean scores of overall CVH and each of the LE8 components, where a higher score (range 0-100 points) means a better health status. A total of 21,667 adults aged 20-79 years were included in this analysis.

Results: The overall CVH did not significantly change between 2005-2006 and 2017-2018 (65.5, 95% CI 63.9-67.1 to 65.0, 95% CI 62.8-67.1; *P*=.82). The individual metrics did not significantly change for diet (41.0, 95% CI 38.0-43.9 to 41.5, 95% CI 36.5-46.6; *P*=.94), physical activity (57.5, 95% CI 53.0-61.9 to 53.0, 95% CI 48.7-57.3; *P*=.26), and blood pressure (68.4, 95% CI 65.2-71.5 to 68.6, 95% CI 65.3-71.9, *P*=.35), improved for nicotine exposure (64.7, 95% CI 61.1-68.4 to 71.9, 95% CI 67.7-76.2; *P*<.001), sleep health (83.7, 95% CI 81.6-85.7 to 84.1, 95% CI 81.2-87.1; *P*=.006), and blood lipids (61.6, 95% CI 59.1-64.0 to 67.0, 95% CI 63.5-70.4; *P*<.001), and worsened for BMI (63.4, 95% CI 59.7-67.1 to 56.2, 95% CI 52.5-59.9; *P*<.001) and blood glucose (83.9, 95% CI 82.4-85.4 to 77.4, 95% CI 74.5-80.3; *P*<.001).

Conclusions: According to the LE8, the overall CVH did not change among US adults from 2005 to 2018, as well as 3 components (diet, physical activity, and blood pressure). Other metrics such as nicotine exposure, blood lipids, and sleep health improved, while BMI and blood glucose deteriorated over time.

(JMIR Public Health Surveill 2023;9:e45521) doi:10.2196/45521

KEYWORDS

trends; cardiovascular health; primordial prevention; adult; nutrition examination; survey; diet; physical activity; data collection; cross-sectional

Introduction

Cardiovascular disease (CVD) is a major public health issue worldwide [1]. In the United States, nearly one in 10 adults (aged ≥20 years) are suffering from CVD (mainly coronary heart disease and stroke) [2], and the annual direct and indirect costs due to CVD in the United States are estimated to amount to US $378 billion [2]. Although the age-standardized mortality rate attributable to CVD has largely declined over the past decades in the United States, the total health and economic burdens due to CVD remain huge, partly due to population growth and aging [3,4].

Given that a large part of the CVD burden is attributable to a limited number of health and behavioral factors, the American Heart Association (AHA) promoted 7 cardiovascular health (CVH) metrics (also known as the Life's Simple 7 [LS7]) in 2010 [5], including 4 behavioral factors (no smoking, sufficient physical activity [PA], healthy diet, and having a normal BMI), and 3 health factors (normal levels of blood lipids, blood pressure [BP], and blood glucose). Each of these 7 metrics was categorized by scores of 0, 1, and 2 points to represent a poor, intermediate, and ideal health status, respectively. The overall CVH score in the LS7 was calculated as the sum of the scores of these 7 metrics, which can range from 0 (worst health) to 14 points (optimal health). Unfortunately, the prevalence of people with ideal CVH (ie, with high LS7 score) has been consistently extremely low (eg, <1% have an optimal score of 14 points) [2].

The LS7 was widely used in the past decade and had played an important role in promoting CVH in the United States and around the world. However, there are some limitations to the LS7 [6]. First, the categorization (poor, intermediate, and ideal) of the LS7 components inherently reduces a precise assessment of the risk factors that all actually have a graded relation with CVD outcomes. Second, the definitions of poor, intermediate, and ideal categories for each component are arbitrary (eg, PA duration from 1 minute to 149 minutes per week are all categorized in the same intermediate category). Third, several social or behavioral CVD risk factors such as psychological factors and sleep are not included in the LS7. Fourth, some of the LS7 metrics are not assessed comprehensively. For example, the LS7 diet metric was assessed by the sole intake of 4 foods and nutrients (ie, fruits and vegetables, fish, whole grains, sweetened beverages, and sodium).

The newly released Life's Essential 8 (LE8) metrics proposed by the AHA [6] addressed several of the limitations of LS7 (Table S1 in Multimedia Appendix 1). For example, some score categories have been further defined—sleep health has been added as an eighth metric, the definition of a healthy diet has been expanded, the use of inhaled nicotine-delivery system and secondhand smoke exposure have been considered in addition to combustible cigarette use, hemoglobin A_{1c} level has been added in addition to fasting glucose, and non–high-density lipoprotein cholesterol is used rather than total cholesterol to assess blood lipids [6]. One recent study reported that 47.3% of young individuals who were evaluated as having ideal CVH by the LS7 were reclassified into the low CVH category by the LE8 [7], suggesting that the new LE8 can reduce the misclassification of CVH status. However, although the AHA acknowledged the importance of psychological health and well-being and strongly encouraged more routine assessment and intervention in clinical settings [6], these factors are also not considered in the LE8 as obligatory indicators.

The new CVH score defined by the LE8 has recently been shown to be inversely associated with the risk of all-cause and CVD mortality [8-10]. One study based on 23,110 adults in the United States indicated that every 10-score increase in overall CVH score could decrease the risk of all-cause mortality by 14% and CVD mortality by 19% [8]. Understanding the secular trends of overall CVH and its components is useful to inform and guide targeted health care and public health policies by the US government and other relevant organizations [11]. Therefore, based on the newly published CVH concept of the LE8, we examined the secular trends in overall CVH and its each component among the US adult population based on 7 serial nationally representative cross-sectional surveys conducted between 2005-2006 (which we refer to as "2005" hereafter) and 2017-2018 (which we refer to as "2018").

Methods

Study Population

Data were obtained from the National Health and Nutrition Examination Survey (NHANES), which consists of serial cross-sectional surveys conducted by the National Center for Health Statistic of the US Centers for Disease Control and Prevention to evaluate the health and nutritional status in US adults and children. Participants were selected using a multistage cluster probability sampling method, and all eligible participants were invited to complete a household interview and a physical examination. The questionnaire includes information on demographic, socioeconomic, lifestyle, and other health-related variables, and the physical examination consists of anthropometric and biological measurements. Data from the NHANES are publicly available [12]. The NHANES began in 1999-2000, and subsequent surveys were carried out every 2 years. In this study, we used the data from the 7 survey waves that were conducted from 2005 to 2018 since information on sleep health (1 component of the LE8) was initially collected since 2005. A total of 39,749 adults aged ≥20 years were initially included. We excluded individuals aged >79 years (n=2776), those with missing values on CVH metrics (total of 10,355, including 7996 due to missing data on diet [missing values on all two 24-hour dietary recalls: n=3876; missing on a 24-hour dietary recall: n=4120]) or demographic variables (n=2037), pregnant women (n=481), women who were breastfeeding (n=198), or individuals with self-reported CVD history (n=2235; Figure S1 in Multimedia Appendix 1). A total of 21,667 adults aged 20-79 years were included in this analysis.

We performed analyses in consideration of several demographic characteristics: age, sex, race or ethnicity, educational level, marital status, and household income. We classified participants into three age groups: 20-39, 40-64, and 65-79 years [13]. Self-reported race or ethnicity was classified as Hispanic, non-Hispanic White, non-Hispanic Black, and "Other" (including mainly Asian participants). To characterize an individual's socioeconomic status, we used the ratio of family income to poverty by dividing family income by the federal poverty threshold for the survey year, adjusting for household size, categorized as <1.30 (low income), 1.30-2.99 (middle income), and ≥3.00 (high income) [14]. Educational level was divided into 4 groups: <high school graduate, high school graduate, some college or associate degree, and college graduate or above. Marital status was categorized as married, divorced/separated/widowed, and unmarried/cohabitation status.

Ethics Approval

The NHANES was approved by the National Center for Health Statistics Research Ethics Review Board. Written informed consent was obtained from all participants. This study was exempted from ethical approval by the institutional review board of Shandong University given the use of deidentified and open access data.

Quantification of CVH

The LE8 score is based on 4 "behavioral factors" (diet, PA, smoking, and sleep health) and 4 "health factors" (BMI, blood lipids, blood glucose, and BP). Diet quality was assessed using the Healthy Eating Index 2015 (HEI-2015) based on data from 2 interviewer-administered 24-hour dietary recalls. Diet was assessed using the "Food Patterns Equivalents Database 14" from the US Department of Agriculture [15]. PA was assessed through self-reported frequency and duration of moderate (resulting in light sweating or a small increase in breathing or heart rate) and vigorous activity (resulting in heavy sweating or a large increase in breathing or heart rate) over the past 30 days. The PA duration was calculated by the frequency of PA in a week multiplied by the duration of PA each time. Nicotine exposure was assessed based on self-reported consumption of combustible cigarettes and use of e-cigarettes and other tobacco products in the previous 30 days, as well as self-reported secondhand smoke exposure in a participant's household. Sleep health was assessed from the question on usual sleep duration per day. Height and weight were objectively measured using a stadiometer and digital weight scale, respectively, and BMI was calculated as weight (kg)/height squared (m^2). BP was measured for three consecutive times after 5 minutes of seated rest, and the average of the second and third readings was used (or the average of the first and second measurements for participants who only had 2 readings or using the first measurement for participants who only had 1 reading). Information on the use of antihypertensive drugs was obtained from the interviewed questionnaire at home. Blood samples were collected at a mobile examination center and sent to a central laboratory for the assessment of blood lipids, fasting glucose, and hemoglobin A_{1c}. Data on the use of lipid-lowering medications, insulin, or oral hypoglycemic agents were obtained from the interviewed questionnaire at home. Changes for some biochemical indexes in measurement methods and used instruments over time were adjusted according to the official recommendation to make them comparable across different survey years [16]. The 8 behavioral and health factors and the thresholds used to allocate scores for each component in detail are shown in Table S2 in Multimedia Appendix 1. Scores for each component ranged from 0 to 100 points according to the AHA scoring algorithm [6], with a higher score meaning a better health status. The overall CVH score was calculated as the mean of the sum of all 8 metrics and similarly ranged from 0 (if the mean score of all components was 0) to 100 (optimal CVH). In this study, scores for the overall CVH and individual components were categorized into poor (0-50 points), intermediate (50-79 points), and high (80-100 points) status according to the AHA recommendation [6].

Statistical Analysis

Differences in percentages of demographic characteristics across the 7 survey waves were assessed with the chi-square test. We calculated the age-standardized mean scores and their 95% CI for the overall CVH score and for each of the 8 CVH components. We performed stratified analysis by sex, age group, educational level, marital status, race or ethnicity, and family income category. We used linear regression model to estimate the linear trends between 2005 and 2018 with adjustments for sex, age group, educational level, marital status, and family income category if appropriate. Data were standardized for age using a direct standardization method based on the age distribution of the 2018 US population (20-39 years: 38.6%; 40-64 years: 44.6%; 65-79 years: 16.8%) [17]. Since oversampling was done in some particular subgroups of the total population in the NHANES to increase the reliability and precision of health indicators in the specific population, the appropriate sample weights, as well as strata and primary sampling units provided by the NHANES were used to make the data national representative of the US population. All statistical analyses were performed using Stata software (version 16.0, Stata Corp), and 2-sided *P* values of <.05 were considered statistically significant.

Results

A total of 21,667 participants aged 20-79 years were included in this study. The distributions of sex, age group, race or ethnicity, marital status, and family income did not differ across the survey waves. Participants with an educational level lower than high school decreased from 13.9% in 2005 to 8.6% in 2018 (*P*=.01; Table 1).

The age-standardized mean overall CVH score changed from 65.5 (95% CI 63.9-67.1) in 2005 to 65.0 (95% CI 62.8-67.1) in 2018 (Figure 1), but the result was not statistically significant (*P* for trend=.82), as well as in almost all subgroups (Table 2). Trends in crude overall CVH score over time showed a similar pattern (Table S3 in Multimedia Appendix 1). Adults with age-standardized high (≥80 points), intermediate (50-79 points), and low (<50 points) overall CVH scores accounted for about 20%, 65%, and 15% in the US population, respectively, in each survey wave (Figure 2). The results were similar when based on the crude proportion (Figure S2 in Multimedia Appendix 1). The overall CVH score significantly worsened between 2005 and 2018 in older, Hispanic, and non-Hispanic White individuals, but no difference was found in other covariate categories.

The age-standardized mean scores of diets, PA, and BP did not significantly change from 2005 to 2018 (Figure 1), as well as in almost all subgroups (Table S4 in Multimedia Appendix 1). The age-standardized mean score of PA fluctuated over time. The age-standardized mean scores improved for nicotine exposure (64.7, 95% CI 61.1-68.4 to 71.9, 95% CI 67.7-76.2), sleep health (83.7, 95% CI 81.6-85.7 to 84.1, 95% CI 81.2-87.1), and blood lipids (61.6, 95% CI 59.1-64.0 to 67.0, 95% CI 63.5-70.4) from 2005 to 2018. However, the age-standardized mean scores significantly decreased for BMI (63.4, 95% CI 59.7-67.1 to 56.2, 95% CI 52.5-59.9) and blood glucose (83.9, 95% CI 82.4-85.4 to 77.4, 95% CI 74.5-80.3; Figure 1, Table S4 in Multimedia Appendix 1). The results were similar within subgroups of sex, age, race or ethnicity, marital status, and family income (Table S4 in Multimedia Appendix 1). Similar trends were found based on the crude scores of the 8 metrics (Table S5 in Multimedia Appendix 1). Among all 8 metrics, blood glucose, sleep health, and nicotine exposure had the largest proportion (>50%) of high score (≥80 points); PA and diets had the largest proportion (40%-50%) of the low score (<50 points) over time (Figure 2). The results were similar when based on crude proportions (Figure S2 in Multimedia Appendix 1). Distributions of scores categories for 8 individual CVH metrics according to the specific scoring algorithm in detail are shown in Figure S3 in Multimedia Appendix 1 (aged-standardized proportion) and Figure S4 in Multimedia Appendix 1 (crude proportion).

By 2018, Americans who were female (66.9, 95% CI 63.9-70.0), young adults aged 20-34 years (69.8, 95% CI 67.2-72.4), non-Hispanic White (65.9, 95% CI 63.1-68.6), or in the “Other” category (including mainly Asian individuals: 66.4, 95% CI 62.9-70.0) had (or tended to had) higher (better) age-standardized mean LE8 scores of overall CVH than those who were male (63.0, 95% CI 60.8-65.2, older adults aged 65-79 years (61.1, 95% CI 59.4-62.9) and non-Hispanic Black (60.1, 95% CI 57.8-62.5), and the age-standardized mean LE8 scores of overall CVH were also much higher in Americans with higher educational level (73.1, 95% CI 70.0-76.1) versus lower educational level (57.1, 95% CI 53.5-60.6), in those who were married (66.3, 95% CI 64.2-68.5) versus those who were unmarried (60.4, 95% CI 56.6-64.2), and in those with higher family income (68.3, 95% CI 65.7-71.0) versus those with lower family income (59.0, 95% CI 55.4-62.6; Table 2).

Table 1. Characteristics of the US adult National Health and Nutrition Examination surveys, 2005-2018.

Characteristics	2005-2006 (n=2791)	2007-2008 (n=3250)	2009-2010 (n=3514)	2011-2012 (n=3113)	2013-2014 (n=3292)	2015-2016 (n=2942)	2017-2018 (n=2765)	*P* value[a]
Sex, %								.49
Male	48.5	46.5	48.3	49.7	49.5	49.5	48.5	
Female	51.5	53.5	51.7	50.3	50.5	50.5	51.5	
Age group (years), %								.19
20-39	39.5	40.7	38.9	39.9	39.4	38.9	39.7	
40-64	48.2	48.7	49.6	49.2	48.3	46.5	45.0	
65-79	12.3	10.6	11.5	10.9	12.3	14.6	15.3	
Race or ethnicity, %								.36
Hispanic	10.5	13.5	13.5	14.5	15.0	15.0	15.6	
Non-Hispanic White	73.6	71.6	70.2	67.5	65.4	66.3	63.2	
Non-Hispanic Black	10.9	10.1	10.3	10.7	10.9	9.4	10.4	
Other	5.0	4.8	6.0	7.3	8.7	9.3	10.8	
Educational level, %								.01
<High school graduate	13.9	18.3	15.8	13.2	12.5	11.0	8.6	
High school graduate	24.3	24.0	22.4	19.6	20.3	21.2	27.0	
Some college or associate degree	32.2	29.5	31.5	32.8	34.2	33.5	31.7	
College graduate or above	29.6	28.2	30.3	34.4	33.0	34.3	32.7	
Marital status, %								.47
Married	59.0	56.0	57.6	53.2	56.6	55.7	53.6	
Divorced/separated/widowed	16.8	16.5	14.9	15.8	16.7	15.1	16.6	
Unmarried/cohabitation	24.2	27.5	27.5	31.0	26.7	29.2	29.8	
Ratio of family income to poverty,[b] %								.11
<1.30	14.6	19.3	21.0	22.7	23.3	18.5	19.4	
1.30-2.99	27.4	28.4	26.6	27.1	24.4	30.6	26.5	
≥3.00	58.0	52.3	52.4	50.2	52.3	50.9	54.1	

[a] Differences in percentages of demographic characteristics across the 7 survey waves were assessed with the chi-square test.

[b] The ratio of family income to the federal poverty threshold, which was adjusted for household size.

Figure 1. Trends in age-standardized mean score (95% CI) of overall cardiovascular health and its 8 components based on the Life's Essential 8 in US adults from 2005-2006 to 2017-2018. BP: blood pressure; CVH: cardiovascular health; non–HDL-C: non–high-density lipoprotein cholesterol; PA: physical activity.

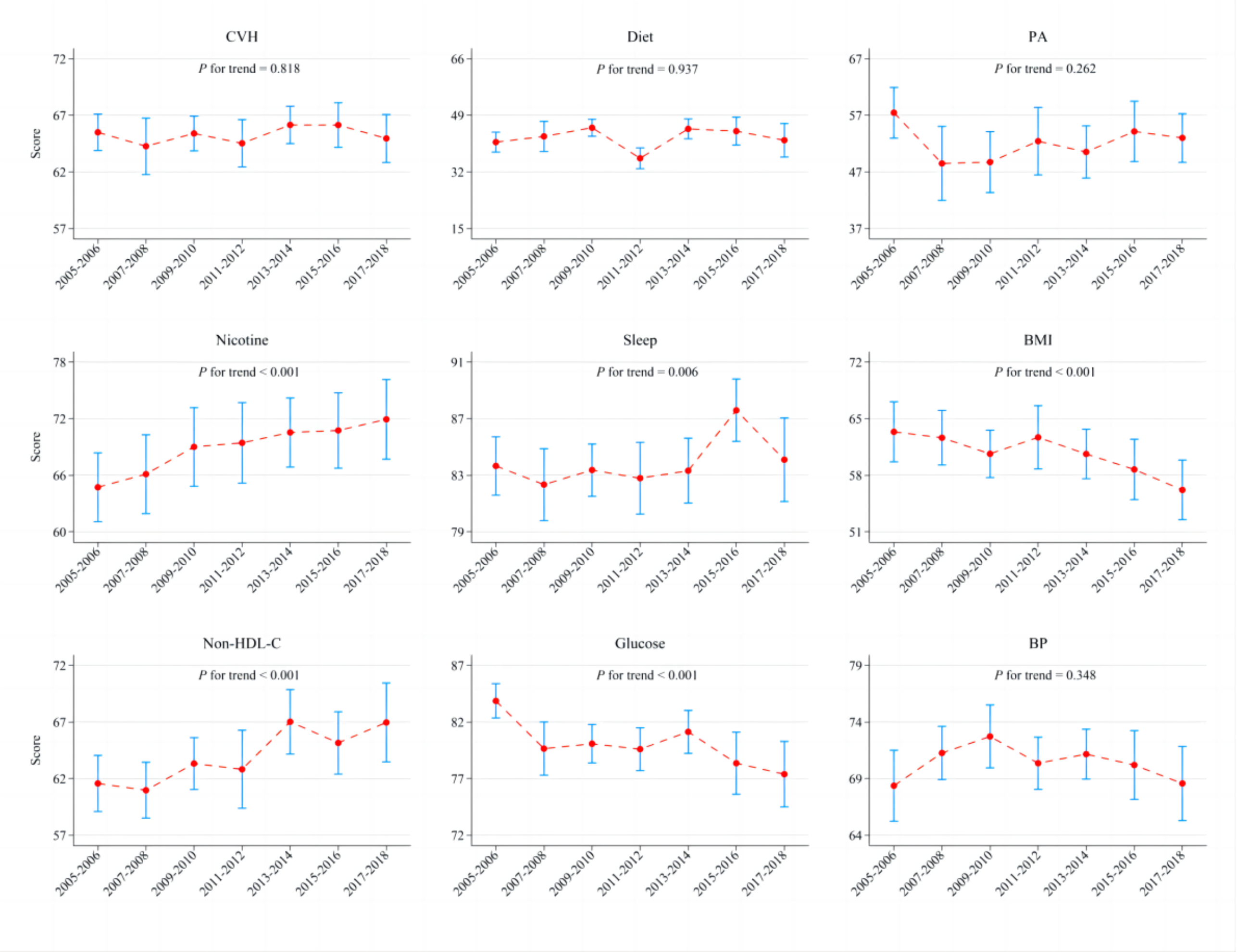


Table 2. Trends in the age-standardized mean score (95% CI) of total cardiovascular health in US adults, overall and by sex, age, race and ethnicity, educational level, marital status, family income to poverty ratio category, 2005-2018.[a]

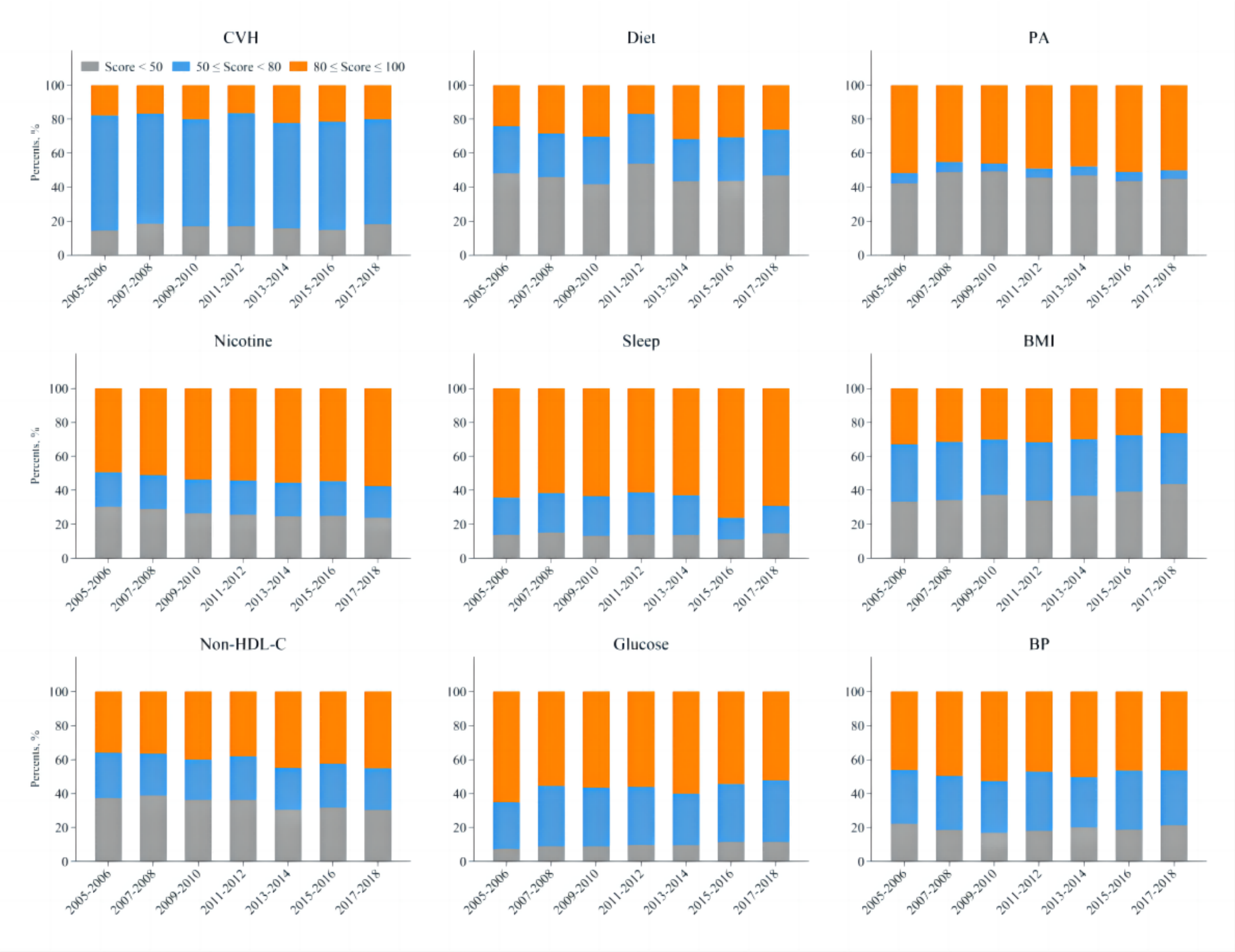
Characteristics	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014	2015-2016	2017-2018	*P* for trend[b]
Overall	65.5 (63.9-67.1)	64.3 (61.8-66.8)	65.4 (63.9-66.9)	64.5 (62.4-66.6)	66.2 (64.5-67.8)	66.1 (64.2-68.1)	65.0 (62.8-67.1)	.82
Sex								
Male	63.5 (61.5-65.5)	62.2 (59.7-64.7)	63.4 (61.2-65.5)	62.6 (60.1-65.0)	64.6 (62.4-66.8)	64.2 (61.8-66.6)	63.0 (60.8-65.2)	.77
Female	67.6 (65.6-69.5)	66.1 (63.1-69.1)	67.3 (65.7-69.0)	66.5 (64.0-68.9)	67.7 (65.7-69.7)	68.2 (65.9-70.5)	66.9 (63.9-70.0)	.46
Age group (years)								
20-39	68.9 (67.8-69.9)	67.8 (65.6-70.0)	69.4 (67.4-71.5)	69.7 (67.6-71.8)	69.7 (67.6-71.7)	70.0 (68.5-71.6)	69.8 (67.2-72.4)	.49
40-64	63.5 (61.7-65.3)	62.1 (59.4-64.8)	63.5 (62.4-64.5)	61.4 (59.3-63.5)	64.0 (62.7-65.4)	63.5 (61.3-65.6)	62.2 (60.4-64.1)	.94
65-79	63.2 (60.8-65.5)	62.0 (59.4-64.5)	61.2 (59.6-62.8)	61.1 (59.0-63.1)	63.8 (62.3-65.2)	64.3 (61.9-66.7)	61.1 (59.4-62.9)	.02
Race or ethnicity								
Hispanic	64.5 (62.2-66.7)	62.7 (60.9-64.4)	61.9 (59.3-64.6)	61.9 (59.6-64.2)	64.7 (62.1-67.3)	62.9 (61.1-64.6)	63.3 (60.7-66.0)	.02
Non-Hispanic White	66.4 (64.2-68.6)	64.9 (61.3-68.6)	66.7 (64.8-68.6)	65.3 (62.6-68.0)	66.8 (64.7-68.9)	67.4 (65.2-69.6)	65.9 (63.1-68.6)	.007
Non-Hispanic Black	59.7 (57.1-62.4)	59.4 (56.7-62.1)	58.1 (55.5-60.8)	60.0 (57.3-62.7)	60.9 (58.9-62.9)	60.0 (56.8-63.2)	60.1 (57.8-62.5)	.70
Other (including mainly Asian individuals)	65.7 (58.3-73.1)	67.6 (61.4-73.8)	68.7 (63.9-73.6)	68.8 (65.4-72.3)	69.5 (65.3-73.7)	68.2 (63.8-72.5)	66.4 (62.9-70.0)	.93
Educational level								
<High school graduate	59.1 (56.3-62.0)	57.1 (54.2-60.0)	58.0 (55.6-60.4)	56.3 (53.1-59.6)	59.4 (57.0-61.8)	57.7 (54.2-61.2)	57.1(53.5-60.6)	.36
High school graduate	60.8 (58.6-63.0)	60.6 (57.7-63.6)	59.8 (57.1-62.5)	60.0 (55.9-64.0)	59.9 (57.1-62.7)	59.9 (57.0-62.8)	61.4 (58.5-64.3)	.75
Some college or associate degree	65.1 (62.7-67.5)	64.2 (61.6-66.9)	64.5 (62.7-66.3)	63.2 (60.9-65.5)	64.7 (62.0-67.5)	65.4 (62.5-68.3)	62.0 (59.8-64.1)	.40
College graduate or above	72.7 (70.0-75.4)	72.2 (69.4-74.9)	74.3 (71.8-76.9)	71.4 (68.9-74.0)	73.9 (71.7-76.2)	73.5 (71.2-75.9)	73.1 (70.0-76.1)	.57
Marital status								
Married	66.0 (64.0-68.1)	65.0 (62.1-68.0)	66.2 (64.2-68.1)	66.0 (63.6-68.3)	67.5 (65.1-69.9)	67.2(64.9-69.5)	66.3 (64.2-68.5)	.94
Divorced/separated/widowed	63.5 (60.1-66.8)	59.2 (55.0-63.5)	62.6 (58.9-66.2)	60.3 (57.0-63.7)	62.5 (60.1-64.9)	60.9(56.3-65.4)	63.1 (60.1-66.2)	.74
Unmarried/cohabitation	64.8 (62.1-67.5)	63.6 (60.6-66.5)	63.0 (59.3-66.6)	62.4 (58.3-66.6)	65.6 (61.3-69.9)	64.3(61.1-67.5)	60.4 (56.6-64.2)	.31
Ratio of family income to poverty								
<1.30	59.7 (57.1-62.4)	58.4 (54.7-62.1)	58.7 (56.3-61.1)	58.6 (54.9-62.4)	60.2 (57.6-62.7)	60.0 (57.4-62.5)	59.0 (55.4-62.6)	.81
1.30-2.99	62.2 (59.9-64.6)	62.7 (60.3-65.1)	63.2 (61.0-65.3)	62.9 (60.5-65.3)	62.5 (60.2-64.9)	63.3 (60.6-66.1)	62.4 (59.6-65.2)	.66
≥3.00	68.4 (66.4-70.5)	67.3 (64.2-70.4)	69.0 (66.8-71.1)	68.0 (65.0-71.0)	70.2 (68.0-72.3)	69.6 (67.3-71.8)	68.3 (65.7-71.0)	.88

[a] Higher score denotes better cardiovascular health.

[b] Linear trends were examined using linear regression model, with adjustment for sex, age, race or ethnicity, educational level, marital status, and the

ratio of family income to poverty.

Figure 2. Trends in distributions of age-standardized scores categories of low (<50 points), median (50-79 points), and high (80-100 points) status of overall cardiovascular health and its 8 components based on the Life's Essential 8 in US adults from 2005-2006 to 2017-2018 (higher score denotes better CVH). BP: blood pressure; CVH: cardiovascular health; non–HDL-C: non–high-density lipoprotein cholesterol; PA: physical activity.



Discussion

Principal Findings

To our knowledge, this study is the first to examine the secular trends in CVH among US adults based on the new LE8 scoring algorithm proposed by the AHA. Using the 7 cross-sectional surveys between 2005 and 2018, the overall CVH score did not significantly change over time. However, the overall CVH status significantly worsened between 2005 and 2018 in older adults, Hispanic participants, and non-Hispanic White participants. There was no significant change over time for 3 CVH components (diet, PA, and BP) but 3 components (nicotine exposure, sleep health, and blood lipids) significantly improved, whereas 2 components (BMI and blood glucose) significantly, and markedly, worsened over time.

Although a number of public health and health care policies and programs have been implemented to prevent CVD since 2010 [18-20], the decline in age-standardized CVD mortality rate has slowed down or even stagnated considerably in recent years among US adults [2,21]. For example, the age-standardized mortality rate for heart disease declined between 1999 and 2010 by 8.3%, but only by 1.7% between 2010 and 2017 [21]. One important reason for this deceleration of the CVD decline may be that some CVD risk factors have not significantly improved or even deteriorated. This is consistent with other studies, which showed that the mean number of ideal CVH metrics based on the LS7 had not significantly changed in the past years [22,23]. For example, the Behavioral Risk Factor Surveillance System data showed a nonsignificant change in the mean number of ideal CVH metrics from 3.89 in 2011 to 3.96 in 2017 [22]. Another NHANES study among US adults showed that the mean number of ideal CVH components did not significantly change from 2007 to 2018 in both female participants (4.40 to 4.48) and male participants (3.97 to 3.93) [23]. In addition, previous studies among adults based on the LS7 also showed that there was no significant improvement in CVH among US adults [24,25]. Results of our study, which are based on the new LE8 scoring algorithm, are consistent with no change in CVH reported by the abovementioned studies based on the LS7, highlighting that there is still a long way to go to improve CVH among US adults.

The stable trends in the overall CVH score may mask disparities in trends in CVH components. We found that the mean scores of 4 CVH metrics including nicotine exposure (increase), BMI (decrease), non–high-density lipoprotein cholesterol (increase), and blood glucose (decrease) changed over time but not all in the same direction, and some (diet) did not change over time. Unhealthy diet is a major risk factor for CVD [26], but unfortunately, the dietary component gathered the lowest mean score [24]. Despite some improvement in some dietary components, the overall dietary quality did not significantly improve [27]. This emphasizes the need for interventions on those dietary components that tend to decline or do not improve over time. Actually, a number of interventions have already been implemented during the past years in the United States [28]. The smoking rate decreased from 24.8% in 1999-2000 to 18.1% in 2017-2018 [29], and this translated into improved mean LS7 scores of smoking from 1.53 in 2007-2010 to 1.60 in 2015-2018 [24], as well as in our study. Data from the NHANES showed that mean BMI increased from 28.0 kg/m^2 in 1999-2000 to 29.8 kg/m^2 in 2017-2018 [29], with the corresponding obesity prevalence dramatically increasing from 27.5% to 43.0% [30]. He et al [29] reported that the mean total cholesterol decreased from 203.3 mg/dL in 1999-2000 to 188.5 mg/dL in 2017-2018, which is consistent with our finding that the LE8 blood lipid component improved over time. Previous studies showed that the estimated prevalence of diabetes increased from 9.8% in 1999-2000 to 14.3% in 2017-2018 [31], consistent with the decreasing LS7 mean score of blood glucose from 1.42 in 2007-2010 to 1.29 in 2015-2018 [24], as well as the decreasing LE8 mean score observed in our study.

However, there were also several inconsistent findings in our study compared with previous studies regarding the secular trends in 2 CVH components (PA and BP) [24,32,33]. It is reported that the estimated prevalence of meeting the PA guideline target increased from 26.0% in 1998 to 37.4% in 2018 [32]. However, the overall mean score of PA in our study still remained unchanged and at a low level. As for hypertension, the prevalence as defined by the 2017 American College of Cardiology/AHA decreased (suggesting an improvement) from 48.1% in 1999-2000 to 44.1% in 2015-2016 [33]. However, another study reported that the mean BP score based on the LS7 slightly decreased (suggesting a deterioration) from 1.28 in 2007-2010 to 1.24 in 2015-2018 [24]. Inconsistent with the above 2 studies, we did not observe a significant change in the mean LE8 component score of BP between 2005 and 2018. In addition, the national data showed that more than half of the US population did not reach the sufficient sleep duration (20-64 years: 7-9 hours per day, 65-79 years: 7-8 hours per day) recommended by the National Sleep Foundation [34]. Inappropriate sleep duration (<6 hours or ≥9 hours) has been associated with an increased risk of CVD and related mortality [35]. Although previous studies reported a persistent deterioration of sleep quality in the US population from 2005 to 2018 [36,37], we observed an opposite trend with a significant increase (improvement) in the mean LE8 component score of sleep health from 2005 to 2018. These disparities might be partly due to methodological issues in assessing sleep quality across studies.

Consistent with previous studies [13,14], the overall LE8 score of CVH and its components differed markedly across the subgroups of sex, age, race or ethnicity, and socioeconomic status, with individuals who were female, younger, non-Hispanic White, and those with better socioeconomic status being more likely to have a better CVH. Although trends in the CVH metrics over time in these subgroups were mostly consistent with those in the overall population, there are some interesting findings. First, non-Hispanic White participants and the "Other" group (including mainly Asian participants) had a higher overall CVH score than non-Hispanic Black participants, but the difference decreased over time, partly due to deterioration in CVH over time among non-Hispanic White participants. The deterioration of 2 CVH components (diet and PA) among non-Hispanic White participants may be the reason for the worsening trend in overall CVH score in this subgroup. This highlights the need to consider LE8 trends according to racial or ethnic groups when designing CVD prevention and control policy and programs. Second, none of the CVH metrics significantly improved among older participants (65-79 years) over time. CVH usually markedly decreases with age [38], largely driven by the strong relation among hypertension, dyslipidemia, and type 2 diabetes with age. The CVH in older adults is a main challenge, as most hard CVD outcomes develop at an older age. In general, the less ideal CVH at all ages found in this study stresses the need for life course approaches to CVH and the need for high-risk individual-level approaches (clinical care) at middle and older ages. Third, the mean LE8 component of BMI (obesity) markedly deteriorated in all subgroups. This trend is particularly worrying, given the major role of adiposity on glucose metabolism, including insulin resistance, hypertension, and more generally poor CVH [39,40]. This highlights the need to strengthen preventive and medical approaches to weight gain prevention and control, including a broader use of novel effective clinical approaches [41]. Fourth, socioeconomic status such as educational level and family income are important social determinants of CVD [42]. Consistent with abundant literature on this topic, our results show that participants with higher socioeconomic status had much better CVH and the CVH gap (as assessed with the LE8 metrics) between individuals with low versus high socioeconomic status did not decrease over time, highlighting the need to further address social disparities in CVD prevention and control.

Strengths and Limitations

Our study is the first to examine the secular trends in CVH from 2005 to 2018 among US adults based on the nationally representative NHANES data using the new AHA LE8 metrics. However, several limitations should be mentioned. First, our study was based on data from 7 surveys over a rather short time period (<15 years). Second, data on diet, PA, nicotine exposure, and sleep health were obtained by self-report, which may be inaccurate due to recall and social desirability biases. Third, the 8 behavioral and health components are given equal weight (implying the same predictive performance) in calculating the LE8 overall CVH score. Fourth, although the 8 components constitute major or important risk/preventive factors for CVH, a number of other factors (eg, familial history/genetics, other conditions, or comorbidities) can largely alter CVD risk, particularly at the individual level. Fifth, missing values on relevant variables, especially for dietary variables, might have affected the representativeness of the study population.

However, the inclusion of data on one 24-hour dietary recall showed similar results (age-standardized score did not significantly change from 40.46 in 2005-2006 to 40.47 in 2017-2018, *P* for trend=.77) with the primary ones.

Conclusions

In summary, based on the LE8 metrics, the overall CVH score did not significantly change from 2005 to 2018 among US adults. Very few individuals had CVH scores in good or optimal ranges, implying a large scope for improvement for most adults. Three CVH components remained unchanged over time (diet, PA, and BP), 3 improved (nicotine exposure, sleep health, and blood lipids), and 2 worsened (BMI and blood glucose), suggesting avenues for improving CVH status among US adults.

Acknowledgments

The authors thank the National Center for Health Statistics of the Centers for Disease Control and Prevention for sharing the NHANES data. This work was supported by the Innovation Team of “Climbing” Program of Shandong University, and Youth Team of Humanistic and Social Science of Shandong University (20820IFYT1902) and National Key Research and Development Plan: Real-Time Intelligent Active Intervention on Integration of Ten Important Chronic Diseases (2020YFC2003504-2).

Data Availability

The data sets generated and analyzed during this study are publicly available on the US Centers for Disease Control and Prevention website [12].

Authors' Contributions

BX conceptualized the study. CL drafted the first draft of the manuscript. YL analyzed the data. BX, MZ, CZ, and PB critically revised the manuscript for key intellectual content. All authors approved the final version of the manuscript. BX is the guarantor. The corresponding author attests that all the listed authors meet the authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOC File , 7845 KB - publichealth_v9i1e45521_app1.doc]

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
2. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation 2022 Feb 22;145(8):e153-e639 [FREE Full text] [doi: 10.1161/CIR.0000000000001052] [Medline: 35078371]
3. Birger M, Kaldjian AS, Roth GA, Moran AE, Dieleman JL, Bellows BK. Spending on cardiovascular disease and cardiovascular risk factors in the United States: 1996 to 2016. Circulation 2021;144(4):271-282 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.120.053216] [Medline: 33926203]
4. Dunbar SB, Khavjou OA, Bakas T, Hunt G, Kirch RA, Leib AR, et al. Projected costs of informal caregiving for cardiovascular disease: 2015 to 2035: a policy statement from the American heart association. Circulation 2018;137(19):E558-E577. [doi: 10.1161/cir.0000000000000570]
5. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, American Heart Association Strategic Planning Task ForceStatistics Committee. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic impact goal through 2020 and beyond. Circulation 2010;121(4):586-613. [doi: 10.1161/CIRCULATIONAHA.109.192703] [Medline: 20089546]
6. Lloyd-Jones DM, Allen NB, Anderson CAM, Black T, Brewer LC, Foraker RE, American Heart Association. Life’s essential 8: updating and enhancing the American Heart Association’s construct of cardiovascular health: a presidential advisory from the American heart association). Circulation 2022;146(5):e18-e43 [FREE Full text] [doi: 10.1161/CIR.0000000000001078] [Medline: 35766027]
7. Shetty NS, Parcha V, Patel N, Yadav I, Basetty C, Li CY, et al. AHA life's essential 8 and ideal cardiovascular health among young adults. Am J Prev Cardiol 2023;13:100452 [FREE Full text] [doi: 10.1016/j.ajpc.2022.100452] [Medline: 36636126]

8. Yi J, Wang L, Guo X, Ren X. Association of Life's Essential 8 with all-cause and cardiovascular mortality among US adults: a prospective cohort study from the NHANES 2005-2014. Nutr Metab Cardiovasc Dis 2023 Jan 28 (forthcoming). [doi: 10.1016/j.numecd.2023.01.021] [Medline: 36775706]
9. Isiozor NM, Kunutsor SK, Voutilainen A, Laukkanen JA. Life's essential 8 and the risk of cardiovascular disease death and all-cause mortality in Finnish men. Eur J Prev Cardiol 2023:zwad040. [doi: 10.1093/eurjpc/zwad040] [Medline: 36753230]
10. Jin C, Li J, Liu F, Li X, Hui Y, Chen S, et al. Life's Essential 8 and 10-Year and lifetime risk of atherosclerotic cardiovascular disease in China. Am J Prev Med. Preprint posted online on February 20, 2023 Feb 20 (forthcoming). [doi: 10.1016/j.amepre.2023.01.009] [Medline: 36813641]
11. Cyriac J, Jenkins S, Patten CA, Hayes SN, Jones C, Cooper LA, et al. Improvements in diet and physical activity-related psychosocial factors among African Americans using a mobile health lifestyle intervention to promote cardiovascular health: the FAITH! (Fostering African American Improvement in Total Health) app pilot study. JMIR Mhealth Uhealth 2021;9(11):e28024 [FREE Full text] [doi: 10.2196/28024] [Medline: 34766917]
12. National Center for Health Statistics. Centre for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2023-02-26]
13. Lloyd-Jones DM, Ning HY, Labarthe D, Brewer L, Sharma G, Rosamond W, et al. Status of. Circulation cardiovascular health in US adults and children using the American Heart Association’s new “life’s essential 8” metrics: prevalence estimates from the national health and nutrition examination survey (NHANES), 2013 through 2018 2022;146(11):822-835. [doi: 10.1161/circulationaha.122.060911]
14. O'Hearn M, Lauren BN, Wong JB, Kim DD, Mozaffarian D. Trends and disparities in cardiometabolic health among U.S. adults, 1999-2018. J Am Coll Cardiol 2022;80(2):138-151. [doi: 10.1016/j.jacc.2022.04.046] [Medline: 35798448]
15. Food patterns equivalents database: methodology and user guides. Agricultural Research Service, US Department of Agriculture. URL: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fped-methodology/ [accessed 2022-08-07]
16. National Health and Nutrition Examination Survey 2015-2016 Data Documentation, Codebook, and Frequencies Plasma Fasting Glucose (GLU_I). URL: https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/GLU_I.htm [accessed 2023-02-20]
17. Population Pyramids of the World from 1950 to 2100. URL: https://www.populationpyramid.net/united-states-of-america/2018/ [accessed 2023-02-26]
18. Frieden TR, Berwick DM. The “million hearts” initiative: preventing heart attacks and strokes. N Engl J Med 2011;365(13):e27. [doi: 10.1056/nejmp1110421]
19. Tomaselli GF, Harty M, Horton K, Schoeberl M. The American Heart Association and the million hearts initiative. Circulation 2011;124(16):1795-1799. [doi: 10.1161/cir.0b013e3182327084]
20. Goldstein LB, Whitsel LP, Meltzer N, Schoeberl M, Birnbaum J, Nelson S, et al. American Heart Association and nonprofit advocacy: past, present, and future. Circulation 2011;123(7):816-832. [doi: 10.1161/cir.0b013e31820a5528]
21. Shah NS, Lloyd-Jones DM, O'Flaherty M, Capewell S, Kershaw K, Carnethon M, et al. Trends in cardiometabolic mortality in the United States, 1999-2017. JAMA 2019;322(8):780-782 [FREE Full text] [doi: 10.1001/jama.2019.9161] [Medline: 31454032]
22. Parcha V, Kalra R, Suri SS, Malla G, Wang TJ, Arora G, et al. Geographic variation in cardiovascular health among American adults. Mayo Clin Proc 2021;96(7):1770-1781 [FREE Full text] [doi: 10.1016/j.mayocp.2020.12.034] [Medline: 33775420]
23. Cho SMJ, Haidermota S, Honigberg MC, Natarajan P. Sex differences in temporal trends of cardiovascular health in young US adults. J Am Heart Assoc 2022;11(11):e024790. [doi: 10.1161/jaha.121.024790]
24. Zhu ZB, Bundy JD, Mills KT, Bazzano LA, Kelly TN, Zhang YH, et al. Secular trends in cardiovascular health in US adults (from NHANES 2007 to 2018). Am J Cardiol 2021;159:121-128. [doi: 10.1016/j.amjcard.2021.08.025] [Medline: 34656312]
25. He J, Zhu ZB, Bundy JD, Mills KT, Bazzano LA, Kelly TN, et al. Abstract P101: cardiovascular health in US adults before and after publication of the American heart association 2020 impact goals: results from NHANES 2005-2016. Circulation 2020;141(Suppl_1):AP101. [doi: 10.1161/circ.141.suppl_1.p101]
26. Sofi F, Cesari F, Abbate R, Gensini GF, Casini A. Adherence to mediterranean diet and health status: meta-analysis. BMJ 2008;337:a1344 [FREE Full text] [doi: 10.1136/bmj.a1344] [Medline: 18786971]
27. Zhang FF, Liu J, Rehm CD, Wilde P, Mande JR, Mozaffarian D. Trends and Disparities in diet quality among US adults by supplemental nutrition assistance program participation status. JAMA Netw Open 2018;1(2):e180237. [doi: 10.1001/jamanetworkopen.2018.0237]
28. Patnode CD, Henderson JT, Coppola EL, Melnikow J, Durbin S, Thomas RG. Interventions for tobacco cessation in adults, including pregnant persons: updated evidence report and systematic review for the US preventive services task force. JAMA 2021;325(3):280-298. [doi: 10.1001/jama.2020.23541] [Medline: 33464342]
29. He J, Zhu ZB, Bundy JD, Dorans KS, Chen J, Hamm LL. Trends in cardiovascular risk factors in US adults by race and ethnicity and socioeconomic status, 1999-2018. JAMA 2021;326(13):1286-1298 [FREE Full text] [doi: 10.1001/jama.2021.15187] [Medline: 34609450]

30. Ogden CL, Fryar CD, Martin CB, Freedman DS, Carroll MD, Gu Q, et al. Trends in obesity prevalence by race and hispanic origin- 1999-2000 to 2017-2018. JAMA 2020;324(12):1208-1210 [FREE Full text] [doi: 10.1001/jama.2020.14590] [Medline: 32857101]
31. Wang L, Li XG, Wang ZX, Bancks MP, Carnethon MR, Greenland P, et al. Trends in prevalence of diabetes and control of risk factors in diabetes among US adults, 1999-2018. JAMA 2021;326(8):704-716 [FREE Full text] [doi: 10.1001/jama.2021.9883] [Medline: 34170288]
32. Whitfield GP, Hyde ET, Carlson SA. Participation in leisure-time aerobic physical activity among adults, national health interview survey, 1998-2018. J Phys Act Health 2021;18(S1):S25-S36 [FREE Full text] [doi: 10.1123/jpah.2021-0014] [Medline: 34465654]
33. Liu BY, Sun YB, Xu GF, Rong S, Zhang J, Du Y, et al. Long-term trends in hypertension and elevated blood pressure among U.S. adults, 1999 to 2016. J Am Coll Cardiol 2018;72(17):2089-2091 [FREE Full text] [doi: 10.1016/j.jacc.2018.07.086] [Medline: 30336835]
34. Mireku MO, Rodriguez A. Sleep duration and waking activities in relation to the national sleep foundation's recommendations: an analysis of US population sleep patterns from 2015 to 2017. Int J Environ Res Public Health 2021;18(11):6154 [FREE Full text] [doi: 10.3390/ijerph18116154] [Medline: 34200277]
35. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011;32(12):1484-1492. [doi: 10.1093/eurheartj/ehr007] [Medline: 21300732]
36. Wang X, Ma H, Gupta S, Heianza Y, Fonseca V, Qi L. The joint secular trends of sleep quality and diabetes among US adults, 2005-2018. J Clin Endocrinol Metab 2022;107(11):3152-3161. [doi: 10.1210/clinem/dgac401] [Medline: 35776011]
37. Caraballo C, Mahajan S, Valero-Elizondo J, Massey D, Lu Y, Roy B, et al. Evaluation of temporal trends in racial and ethnic disparities in sleep duration among US adults, 2004-2018. JAMA Netw Open 2022;5(4):e226385 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.6385] [Medline: 35389500]
38. Rodgers JL, Jones J, Bolleddu SI, Vanthenapalli S, Rodgers LE, Shah K, et al. Cardiovascular risks associated with gender and aging. J Cardiovasc Dev Dis 2019;6(2):19 [FREE Full text] [doi: 10.3390/jcdd6020019] [Medline: 31035613]
39. Mozaffarian D. Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: a comprehensive review. Circulation 2016;133(2):187-225 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.018585] [Medline: 26746178]
40. Singh GM, Danaei G, Farzadfar F, Stevens GA, Woodward M, Wormser D, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group; Asia-Pacific Cohort Studies Collaboration (APCSC), Diabetes Epidemiology: Collaborative analysis of Diagnostic criteria in Europe (DECODE), Emerging Risk Factor Collaboration (ERFC), Prospective Studies Collaboration (PSC). The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS One 2013;8(7):e65174 [FREE Full text] [doi: 10.1371/journal.pone.0065174] [Medline: 23935815]
41. Lingvay I, Sumithran P, Cohen RV, le Roux CW. Obesity management as a primary treatment goal for type 2 diabetes: time to reframe the conversation. Lancet 2022 Jan;399(10322):394-405. [doi: 10.1016/s0140-6736(21)01919-x]
42. Schultz WM, Kelli HM, Lisko JC, Varghese T, Shen J, Sandesara P, et al. Socioeconomic status and cardiovascular outcomes: challenges and interventions. Circulation 2018 May 15;137(20):2166-2178 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.117.029652] [Medline: 29760227]

Abbreviations

AHA: American Heart Association
BP: blood pressure
CVD: cardiovascular disease
CVH: cardiovascular health
LE8: Life’s Essential 8
LS7: Life’s Simple 7
NHANES: National Health and Nutrition Examination Survey
PA: physical activity

Edited by A Mavragani, T Sanchez; submitted 05.01.23; peer-reviewed by C Paquet, M Anderson, K Gupta, Y Chen; comments to author 16.02.23; revised version received 06.03.23; accepted 27.03.23; published 08.05.23.

Please cite as:
Li C, Li Y, Zhao M, Zhang C, Bovet P, Xi B
Using the New "Life's Essential 8" Metrics to Evaluate Trends in Cardiovascular Health Among US Adults From 2005 to 2018: Analysis of Serial Cross-sectional Studies
JMIR Public Health Surveill 2023;9:e45521
URL: https://publichealth.jmir.org/2023/1/e45521
doi:10.2196/45521
PMID:37155232

Original Paper

Associations Between Personality Traits and Energy Balance Behaviors in Emerging Adulthood: Cross-Sectional Study

Katrina E Champion[1*], PhD; Cath Chapman[1*], PhD; Matthew Sunderland[1], PhD; Tim Slade[1], PhD; Emma Barrett[1], PhD; Erin Kelly[1], PhD; Lexine Stapinski[1], PhD; Lauren A Gardner[1], PhD; Maree Teesson[1], PhD; Nicola C Newton[1], PhD

The Matilda Centre for Research in Mental Health and Substance Use, The University of Sydney, Sydney, Australia

[*] these authors contributed equally

Corresponding Author:
Cath Chapman, PhD
The Matilda Centre for Research in Mental Health and Substance Use
The University of Sydney
Level 6, Jane Foss Russell Building
Sydney, 2006
Australia
Phone: 61 2 8627 9007
Email: cath.chapman@sydney.edu.au

Abstract

Background: Internalizing and externalizing personality traits are robust risk factors for substance use and mental health, and personality-targeted interventions are effective in preventing substance use and mental health problems in youth. However, there is limited evidence for how personality relates to other lifestyle risk factors, such as energy balance–related behaviors, and how this might inform prevention efforts.

Objective: This study aimed to examine concurrent cross-sectional associations between personality traits (ie, hopelessness, anxiety sensitivity, impulsivity, and sensation seeking) and sleep, diet, physical activity (PA), and sedentary behaviors (SB), 4 of the leading risk factors for chronic disease, among emerging adults.

Methods: Data were drawn from a cohort of young Australians who completed a web-based, self-report survey in 2019 during early adulthood. A series of Poisson and logistic regressions were conducted to examine the concurrent associations between the risk behaviors (sleep, diet, PA, and sitting and screen time) and personality traits (hopelessness, anxiety sensitivity, impulsivity, and sensation seeking) among emerging adults in Australia.

Results: A total of 978 participants (mean age 20.4, SD 0.5 years) completed the web-based survey. The results indicated that higher scores on hopelessness were associated with a greater daily screen (risk ratio [RR] 1.12, 95% CI 1.10-1.15) and sitting time (RR 1.05, 95% CI 1.0-1.08). Similarly, higher scores on anxiety sensitivity were associated with a greater screen (RR 1.04, 95% CI 1.02-1.07) and sitting time (RR 1.04, 95% CI 1.02-1.07). Higher impulsivity was associated with greater PA (RR 1.14, 95% CI 1.08-1.21) and screen time (RR 1.06, 95% CI 1.03-1.08). Finally, higher scores on sensation seeking were associated with greater PA (RR 1.08, 95% CI 1.02-1.14) and lower screen time (RR 0.96, 95% CI 0.94-0.99).

Conclusions: The results suggest that personality should be considered when designing preventive interventions for lifestyle risk behaviors, particularly in relation to SB, such as sitting and screen time.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12612000026820; https://tinyurl.com/ykwcxspr

(JMIR Public Health Surveill 2023;9:e42244) doi:10.2196/42244

KEYWORDS

personality; emerging adulthood; screen time; sedentary; sleep; diet; physical activity; prevention; mental health; risk factor; sedentary behavior; chronic disease

Introduction

Chronic diseases, such as cancer, cardiovascular disease, and mental disorders, are the leading causes of disability and death across the globe [1]. It is estimated that almost one-third of chronic diseases could be prevented by reducing exposure to modifiable lifestyle risk behaviors [2], such as physical inactivity, poor diet, alcohol use and smoking, poor sleep, and sedentary behaviors (SB; sitting and screen time) [3,4]. These risk behaviors are highly prevalent among adolescents in Australia [5] and worldwide [6] and are associated with significant health consequences, including increased risk of obesity [7] and emotional and behavioral problems [8,9]. They typically have their onset early in life, and once established, they track into adulthood, increasing the risk for chronic disease over the life course [4,10-12]. Identifying and implementing innovative strategies to prevent or modify these lifestyle risk behaviors during adolescence is critical for promoting lifelong health.

One plausible target for preventing lifestyle risk behaviors and associated harms is personality traits. Numerous studies have demonstrated robust associations between personality traits and substance use [13], particularly with respect to the internalizing personality profiles of *anxiety sensitivity* (fear of anxiety-related physical sensations) and *hopelessness* (tendency toward low mood; worthlessness; and negative beliefs about oneself, the world, and the future), and the externalizing profiles of *sensation seeking* (elevated need for stimulation and intolerance to boredom) and *impulsivity* (rapid decision-making and poor response inhibition) [14]. Each personality trait has been shown to be associated with specific patterns of substance misuse, motivations for use, and comorbid psychopathology. For example, hopelessness has been associated with the early onset of alcohol use and depressive symptoms, while impulsivity has been associated with drinking problems, tobacco use, other drug use, and conduct problems [15]. These personality profiles have been shown to be reliable predictors of future substance use and mental health problems and are therefore important targets for screening and prevention [13,15-17].

Improved understanding of the way in which personality traits are associated with other highly prevalent, yet modifiable risk behaviors among youth has considerable implications for tailoring prevention programs to account for particular personality styles. To date, much of the research examining associations between personality and energy balance–related behaviors (ie, health behaviors related to energy intake and energy expenditure) such as dietary intake, SBs (sitting and screen time), physical activity (PA), and sleep has focused on children and early adolescents [18,19]. For example, a 2019 cross-sectional study among 8-11 year olds in the United States found that adequate sleep duration and reduced screen time were associated with less impulsive behavior among children [18]. On the contrary, there is little research on the associations between personality and lifestyle risk behaviors during emerging adulthood (18-25 years), a critical, yet poorly understood period characterized by numerous personal and social role changes, increased autonomy over lifestyle choices, declines in health behaviors [20,21], and increased risk of anxiety, depression, and substance use disorders [22]. The limited existing literature on this population has found relationships between sleepiness and impulsivity [23] and extraversion and SB [24]. However, further research among larger samples of emerging adults and addressing a broader range of energy balance–related risk behaviors is required. Therefore, this study aimed to examine cross-sectional associations between 4 personality traits (hopelessness, anxiety sensitivity, impulsivity, and sensation seeking) and 5 energy balance–related behaviors (screen time, sitting time, fruit and vegetable intake, sleep duration, and moderate-to-vigorous physical activity [MVPA]) among a sample of emerging Australian adults.

Methods

Study Design

The sample was derived from the long-term follow-up of an existing cohort of young Australians who participated in a cluster randomized controlled trial (RCT). A total of 2190 grade 8 students (54% male; mean age 13.3, range 11.9-15.4 years) from 26 Australian secondary schools in New South Wales and Victoria (17 private or independent and 9 public) were originally recruited to the RCT in 2012. This 4-arm cluster RCT evaluated the impact of universal and selective interventions (separately and together) for the prevention of alcohol and other drug use delivered in grade 8. Intervention content did not focus on the behaviors measured for this study. All participants had received health education as usual (either as part of the control arm of the study in grade 8) or over the period of follow-up in grades 9 and 10 at school. This trial is registered with the Australian New Zealand Clinical Trials Registry (ACTRN12612000026820), and full details of the RCT, including detailed descriptions of the interventions, adherence, and attrition are published elsewhere [25-27]. Despite the longitudinal nature of the cohort, given data on healthy lifestyle behaviors were only obtained during the final follow-up period, this study uses cross-sectional data collected from 978 participants (mean age 20.4, SD 0.5 years) who completed the 7-year follow-up assessment in 2019.

Procedure

Using multiple sources of locator information provided during previous assessments (including email, phone number, postal address, Facebook handle, and parents' email address), participants were contacted and invited to complete a web-based, self-report survey at a time and location of their choice.

Ethics Approval

All procedures were approved by the Human Research Ethics Committees of the University of New South Wales (HC16881) and The University of Sydney (HREC 2018/845). All participants were required to provide digital informed consent prior to participation, which allows secondary analyses, such as this study, without additional consent. To ensure privacy and confidentiality, all data were deidentified. Respondents were compensated AUD $30 (USD $19.80) (or an equivalent voucher) for their time.

Measures

Overview

All measures were self-report and were completed digitally.

Demographics

The following demographic variables were assessed: sex (male, female, and other), employment status (full-time, part-time, unemployed, and other), and tertiary education in progress or completed (none, trade/technical, and university/college).

Personality

Personality was measured using the Substance Use Risk Profile Scale (SURPS), a 23-item tool developed to assess 4 personality traits relevant to substance use: impulsivity, hopelessness, anxiety sensitivity, and sensation seeking [13,16]. The tool measures 4 dimensions of personality: impulsivity (eg, I often don't think things through before I speak), sensation seeking (eg, I enjoy new and exciting experiences even if they are unconventional), hopelessness (eg, I feel that I am a failure), and anxiety sensitivity (eg, I get scared when I experience unusual bodily sensations). The SURPS is scored on a 4-point Likert scale (strongly agree—strongly disagree) and has demonstrated acceptable to good internal consistency and reliability (α=.62-.86) as well as good concurrent and predictive validity among a comparable sample of Australian youth [15]. The total scores for each of the 4 subscales were calculated and converted into *z*-scores based on the total sample mean and SD.

Fruit and Vegetable Consumption

Fruit and vegetable intake were assessed using 2 validated items commonly used in health research [28,29]: "About how many serves of fruit/vegetables do you usually have each day?" Possible response options were "don't eat fruit/vegetables," "1 serve or less," "2-3 serves," "4-5 serves," and "6 serves or more." Participants were provided with written information about what constitutes one serving of fruit and vegetables. In line with the Australian dietary guidelines [30], responses were dichotomized so that poor fruit intake was defined as less than 2 servings per day, and insufficient vegetable intake was classified as less than 4-5 servings per day. A new variable was constructed to represent both fruit and vegetable intake with two categories: (1) adequate fruit or vegetable intake and (2) inadequate fruit and vegetable intake.

Physical Activity

Self-reported MVPA was assessed using items from the International Physical Activity Questionnaire-Short Form (IPAQ-SF), which has demonstrated good psychometric properties in a diverse range of samples [31]. Respondents were asked to indicate how many days during the past 7 days and for how long each day (in hours and minutes) they had performed vigorous PAs (eg, heavy lifting, digging, aerobics, or fast bicycling). Participants were also asked to report the number of days they did moderate PAs (eg, carrying light loads, bicycling at a regular pace, or doubles tennis, excluding walking) in the past 7 days and how much time they usually spent doing these activities on one of those days (in hours and minutes). The number of hours spent performing both vigorous and moderate PA in the past 7 days was summed and divided by 7 to calculate the average number of hours per day engaged in MVPA.

Sedentary Behaviors

Self-reported daily sitting time (in hours) was assessed using a single item from the IPAQ-SF [31]:

> *How many hours do you spend sitting in a typical 24-hour day (eg, travelling to/from school, university or work; at school, university or work; watching television, using a computer at home and leisure time).*

To assess screen time, respondents were asked to report the amount of time (hours and minutes) in the past week they typically spent (1) watching television or videos during their free time and (2) using a computer during their free time (including computers, laptops, Xbox, PlayStation, iPads or other tablets, smartphones, YouTube, Facebook or other social media, and the internet). Separate items were used to assess average screen time on weekdays and weekend days. The total number of hours per week spent on both TV and computer during their free time was calculated by multiplying the sum of weekday TV and computer time by 5, multiplying the sum of weekend day TV and computer time by 2, and dividing by 7 to generate an average screen time per day over a typical week.

Sleep Duration

To assess sleep duration (in hours), respondents were asked "How many hours in each 24-hour day do you usually spend sleeping (including at night and naps)?" The total number of hours was used as a continuous indicator of sleep.

Statistical Analyses

Given prior work has demonstrated sex differences in the prevalence of lifestyle risk behaviors [32,33], descriptive statistics (means and frequencies) were determined for each of the risk behaviors by sex. Inspection of the outcomes in terms of the underlying distribution were used to inform the selection of appropriate regression models. A series of Poisson (for MVPA, sleep, screen time, sitting time) with a log link, and logistic regressions (for fruit or vegetable intake) with a logit link were estimated to examine the associations between risk behaviors and personality type. Rate ratios (RRs) derived by exponentiating the coefficients from Poisson regressions were presented for semicontinuous count data (MVPA, sleep, screen time, and sitting time), whereas odds ratios (ORs) derived by exponentiating the coefficients from logistic regressions were presented for binary data (fruit or vegetable intake). Initial data inspection indicated that the semicontinuous count variables (MVPA, sleep, screen time, and sitting time) demonstrated a significant negative skew, indicative of count data and suitable for Poisson models in comparison to linear models. Separate models were estimated to examine the association of each personality type with the multiple health behavior outcomes with sex, education status, and employment included in each model as covariates. Additional sensitivity analyses were conducted by reestimating the Poisson and logistic regression models for each outcome but included all personality scores in a single model as well as the additional covariates of sex, education status, and employment to determine independent effects of personality scores on each risk behavior. All analyses were conducted in SAS (version 9.4; SAS Institute).

Results

Sample Characteristics

A total of 978 participants (n=478, 49% male; mean age 20.4, SD 0.4 years) completed the web-based survey. Table 1 summarizes the sample characteristics, and Table 2 reports descriptive statistics for the lifestyle risk behaviors by sex among those who provided adequate data for the risk behaviors. Overall, the means and SDs of the raw personality scores are 13.6 (4.1) for negative thinking, 12.0 (2.9) for anxiety sensitivity, 10.0 (2.7) for impulsivity, and 15.7 (3.8) for sensation seeking.

Table 1. Sociodemographic characteristics of the sample at the 2019 assessment wave.

Characteristic	Values
Age (years)	
Mean (SD)	20.4 (0.4)
Range	18.2-22.0
Sex, n (%)	
Male	478 (48.9)
Female	499 (51.1)
Tertiary education (completed or in progress), n (%)	
None	36 (4)
Trade or technical	88 (9)
University or college	854 (87.3)
Employment status, n (%)	
Full-time	100 (10.2)
Part-time	707 (72.3)
Unemployed	157 (16.1)
Other	14 (1)

Table 2. Prevalence of lifestyle risk behaviors by sex at the 2019 assessment wave.

	MVPA[a] (n=829; h/day), mean (SD)	Sleep duration, (n=977; h/day), mean (SD)	Screen time (n=880; h/day), mean (SD)	Sitting time (n=974; h/day), mean (SD)	Fruit or vegetable (n=977), % inadequate intake/day[b]
Total sample	1.6 (2.7)	7.9 (2.4)	7.7 (6.1)	7.4 (3.4)	84.4
Male	2.0 (2.6)	7.8 (2.2)	8.1 (6.7)	7.3 (3.4)	89.3
Female	1.2 (2.7)	7.9 (2.5)	7.3 (5.4)	7.5 (3.3)	79.8

[a] MVPA: moderate-to-vigorous physical activity.

[b] Combines participants with inadequate fruit or vegetable intake.

Associations Between Personality and Lifestyle Risk Behaviors

Table 3 presents the RR, OR, and 95% CI from the regression models. Results indicated that a 1 SD higher score on hopelessness was associated with a 12% higher rate of screen time (RR 1.12, 95% CI 1.10-1.15) and a 5% higher rate of sitting time (RR 1.05, 95% CI 1.03-1.08). Similarly, a 1 SD higher score on anxiety sensitivity was associated with a 4% higher rate of both screen time and sitting time (RR 1.04, 95% CI 1.02-1.07; RR 1.04, 95% CI 1.02-1.07). A 1 SD higher score on impulsivity was associated with a 14% higher rate of MVPA (RR 1.14, 95% CI 1.08-1.21) and a 6% higher rate of screen time (RR 1.06, 95% CI 1.03-1.08). Finally, a 1 SD higher score on sensation seeking was associated with an 8% higher rate of MVPA (RR 1.08, 95% CI 1.02-1.14) and a 4% lower rate of screen time (RR 0.96, 95% CI 0.94-0.99). Personality was not associated with inadequate fruit and vegetable intake in any of the models (RRs 0.88-1.18; all CIs included 1.0). Sensitivity analyses adjusted for all other personality scores in the models are provided in Multimedia Appendix 1 and demonstrate the robustness of the key findings.

Table 3. Rate ratios, odds ratios, and 95% CI from the regression models examining associations between personality type and lifestyle risk behaviors at the 2019 assessment wave.

Personality type	MVPA[a] (n=829), RR[b] (95% CI)	Sleep duration (n=977), RR (95% CI)	RR (95% CI)	Sitting time (n=974), RR (95% CI)	Inadequate fruit or vegetable[c] (n=977), OR[d] (95% CI)
Hopelessness	0.95 (0.90-1.01)	1.00 (0.98-1.02)	*1.12 (1.10-1.15)*[e]	*1.05 (1.03-1.08)*	1.18 (0.99-1.41)
Anxiety sensitivity	1.02 (0.97-1.08)	1.01 (0.99-1.03)	*1.04 (1.02-1.07)*	*1.04 (1.02-1.07)*	0.91 (0.76-1.09)
Impulsivity	*1.14 (1.08-1.21)*	0.99 (0.97-1.01)	*1.06 (1.03-1.08)*	0.99 (0.96-1.01)	1.05 (0.87-1.25)
Sensation seeking	*1.08 (1.02-1.14)*	0.99 (0.97-1.02)	*0.96 (0.94-0.99)*	0.98 (0.96-1.00)	0.88 (0.74-1.06)

[a] MVPA: moderate-to-vigorous physical activity.

[b] RR: risk ratio.

[c] Odds ratios derived from logistic regression with reference category as adequate fruit and vegetable intake. All models controlled for sex, employment status, and education.

[d] OR: odds ratio.

[e] Italics indicates significance at the *P*<.05 level.

Discussion

Principal Results

This study examined the cross-sectional associations between personality traits (hopelessness, anxiety sensitivity, sensation seeking, and impulsivity) and sleep duration, diet, PA, and SB, 4 of the leading risk factors for chronic disease among a sample of emerging Australian adults. We found that hopelessness was associated with greater screen and sitting time, and anxiety sensitivity was associated with higher screen and sitting time. In terms of the externalizing personality traits, impulsivity was associated with greater screen time and MVPA, while sensation seeking was associated with greater MVPA but less screen time. There were no significant associations between any of the personality traits and sleep or diet.

Comparison With Prior Work

This study adds to the limited literature on personality traits and energy balance–related behaviors during emerging adulthood. The finding that participants high on hopelessness engaged in greater daily screen and sitting time among our sample of Australian 20 year olds is consistent with prior research among young people. For example, a longitudinal study conducted among a community-based sample of Canadian youth (spanning ages 10-21 years) similarly found that higher levels of depression were associated with greater screen time [34]. Although the personality trait of hopelessness measured in our study is not a direct measure of depressive symptomology, previous research has shown that the 2 constructs are moderately to strongly correlated [15,35]. The finding is also consistent with coping styles associated with hopelessness (and depressive symptomatology) such as withdrawal from others, avoidance of in-person social interaction, and distraction [13]. Relatedly, the present findings support previous work that has found the symptoms of depression to be associated with reduced PA among both adolescents and young adults [34,36], likely reflective of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) Diagnostic Criteria for depression of reduced physical movement and loss of energy. Our findings that impulsivity was associated with greater screen time, and sensation seeking with less screen time, is largely consistent with prior work [18]. For example, a small study among young Australian adults found that extraversion, which is correlated with the SURPS sensation-seeking scale [16], was negatively associated with both sitting time and leisure screen time [24], and a study among Spanish adolescents showed that extraversion was linked to a reduced risk of excessive use of social network sites [19]. Taken together, it seems that although young people exhibiting high levels of sensation seeking may have an increased risk of substance use and mental health problems [15], they are less likely to engage in risky levels of screen time and physical inactivity. Those exhibiting hopelessness and anxiety sensitivity, on the other hand, are more likely to engage in risky levels of screen time and sitting time as well as being at risk of substance use and mental health problems [15].

Interestingly, we did not find evidence of an association between the 4 personality traits and inadequate diet or less sleep. The lack of an association between personality and dietary intake is likely due to the fact that only fruit and vegetable consumption was measured, and we did not assess other dietary behaviors such as intake of sugar-sweetened beverages or junk food. Indeed, prior research suggests that individuals high on impulsivity are more likely to find it difficult to resist impulses related to unhealthy food choices [37], and impulsivity is associated with excessive overeating of high-calorie food or food addiction [38]. Similarly, our results in terms of sleep and personality somewhat contradict prior work that has found associations between sleep problems and impulsivity among young adults in the United States; however, the US-based study assessed daytime sleepiness, rather than total sleep duration [23]. It is also worth highlighting that self-reported sleep duration in the present sample was relatively high, with a mean of 7.9 (SD 2.4) hours of sleep reported per day. Although this is comparable to a recent Australian cohort where young adults (aged 18-24 years) reported sleeping 7.2 hours on average per night [39], it falls within the Australian health guidelines for sleep (7-9 hours per night) among young adults, so this may have limited our ability to examine associations between personality and sleep duration among those at the lower end of the sleep spectrum. Finally, our study did not assess sleep quality, which recent studies have suggested could be a better measure than sleep quantity as an index for assessing sleep, and further that in people with an average sleep duration of 7 hours, average sleep *quality* may be better related to health and other factors [40]. Future studies that include more nuanced sleep measures may shed more light on the relationship between personality traits and sleep. Similarly, levels of MVPA were also high among our sample. While a recent Australian study using self-reported data from the National Health Survey, estimated over half (55% or 1.2 million) of all 18-24 year olds were sufficiently active for their age [41], future studies with samples reporting MVPA levels at the lower end of the spectrum would be useful.

Implications for Prevention

The findings from this study have important implications for the prevention of chronic disease risk factors, especially excessive screen time, a highly prevalent behavior that has increased among children and adults in the wake of the COVID-19 pandemic [42,43]. Indeed, in 2019, before the global pandemic, this sample of young adults reported very high rates of daily recreational screen time, with an average of 7.7 (SD 6.1) hours of screen time per day (including watching TV or videos, using a computer, or playing videogames). Excessive sitting time was also prevalent, with participants reporting sitting a mean of 7.4 (SD 3.4) hours per day. High levels of SB have been associated with a range of negative health consequences [44]. Among adolescents, for example, excessive screen time is linked to markers of adiposity and cardiometabolic disease risk [45], poor mental health [46] and quality of life [47], and increased risk of all-cause, cardiovascular disease, and cancer-related mortality, and incidence of these diseases, in adults [48,49]. Therefore, public health approaches to modify SB are critical. Prior research has suggested that understanding the ways in which personality and risk behaviors are related may be important for designing more effective and appealing public health campaigns [36]. For example, a campaign that links reductions in SB to changes in mood may be more appealing to young adults than a campaign that simply focuses on moving more.

In addition, our results suggest that identifying young people with high levels of hopelessness, anxiety sensitivity, and impulsivity could be a means of intervening to reduce SB, such as sitting and screen time, and increasing MVPA. For example, a program targeting young people with high hopelessness could also teach young people strategies to improve both their mood and physical health, such as reducing screen time, spending time with others outdoors, and engaging in PA. Designed to target young people based on personality risk factors for substance misuse and other emotional and behavioral problems, the *Preventure* intervention is a selective, personality-targeted program that has been shown to reduce alcohol use, alcohol-related harms, illicit substance use, depressive symptoms, delinquency, truancy, bullying, and conduct problems in RCTs in North America and Australia up to 7 years post intervention [26,50-53]. It is a brief, manual-based intervention that includes character profiles and case studies to teach young people about their target personality style, help them to explore ways of coping with unhelpful thoughts and emotions related to their personality, and to identify and challenge personality-specific cognitions that lead to problematic behaviors [13,54]. Adaptations to programs such as Preventure could include specific personality-tailored content aimed at lifestyle risk behaviors such as screen time and sleep, as well as concurrent targeting of several related lifestyle risk behaviors, potentially offering synergistic effects on multiple risk behaviors [55].

Limitations

There are several limitations that should be considered. First, our measures of lifestyle risk behaviors were self-reported and were limited to simple indicators of each behavior. Future research should aim to assess a broader range of outcomes, including sugary foods and drinks, sleep quality, and specific types of screen time, including device and content viewed, and different types of PA (eg, active transport). Future research could also consider the inclusion of objective measures of lifestyle risk behaviors, noting that the benefits need to be weighed against the higher costs and practicalities of assessment as well as the potential barriers to engagement among young adults. Second, the SURPS differs from other widely used measures of personality, such as the Big 5 model of personality (neuroticism, extraversion, openness, agreeableness, and conscientiousness). Numerous studies have demonstrated the SURPS to be a reliable tool for identifying personality traits associated with substance misuse and other emotional and behavioral problems among adolescents [15-17,56]. Noting this, a key advantage of the SURPS is that its items measure personality traits and not substance use. This enables the identification of adolescents at greater risk for substance use and related risk behaviors prior to their entrenchment, which is particularly beneficial for the implementation of selective or targeted preventive interventions.

Third, the original sample was designed as a cluster RCT for a school-based prevention program on harmful alcohol use rather than a representative sample of Australian young adults, and covariates included in the analyses were limited to those measured in the original study. There were a large number of schools in the trial, and they represented both independent and public schools in 2 states in Australia. However, the sample was largely Australian-born, English-speaking adolescents, and although this is in line with the general population [57], they were primarily well educated and employed, limiting the generalizability of our findings. This is particularly important, given the prevalence of risk factors is not uniform across populations. For example, low socioeconomic status youth in Australia are 22 times more likely to use alcohol at risky levels and 5 times more likely to smoke [58], and 12% less likely to be sufficiently physically active [59] than higher socioeconomic status youth. Co-designed interventions for disadvantaged populations are clearly needed, including interventions delivered in the home and school settings as well as public health initiatives to address the underlying structural and social determinants of health. Future research examining the associations between personality and chronic disease risk factors among young adults from diverse and disadvantaged backgrounds is required to inform such interventions, along with the inclusion of a broader range of potential covariates related to socioeconomic status and disadvantage.

Fourth, we note that there is considerable attrition at the time point when lifestyle risk behaviors were obtained compared to the baseline sample (55%), and therefore, some degree of response bias might limit our results. Comprehensive attrition analyses are reported elsewhere [26]; however, those who were missing at time points 6 and 7 compared to those who were retained were more likely to be male, less likely to have used alcohol or have hazardous levels of drinking at baseline, and demonstrated lower mean impulsivity and sensation-seeking scores at baseline. Replication of these findings in other samples would support the robustness of the findings reported here. Finally, this analysis was limited to cross-sectional examinations, as the lifestyle risk behaviors (ie, diet, PA, and SB) were only assessed at the early adulthood assessments of the RCT, limiting our ability to examine causal and bidirectional relationships between personality and lifestyle risk behaviors. Nonetheless, this study focuses on the critical and understudied emerging adulthood period and suggests that young adults exhibiting high levels of specific personality traits may benefit from targeted interventions to improve both physical and mental health. Further longitudinal research examining personality traits and energy balance–related behaviors across adolescence and into early adulthood will be important for understanding when to intervene.

Conclusions

This study provides new knowledge about the associations between personality and chronic disease risk factors among young adults. Given the high prevalence of these risk behaviors, their known relationship to important chronic disease outcomes, and the demonstrated effects of personality-targeted programs on other domains of health, the study suggests that personality should be considered when designing preventive interventions for lifestyle risk behaviors, especially high levels of SB, such as screen and sitting time, and low levels of PA. These findings have the potential to inform future development of programs tailored to particular personality styles to improve physical and mental health among young Australians.

Acknowledgments

We acknowledge the schools, students, and teachers who participated in this research. The research team also acknowledges Ms Anna Smout, Dr Siobhan Lawler, and Dr Jennifer Debenham for their assistance with data collection for this study. This study was funded through a project grant from the National Health and Medical Research Council (NHMRC; APP1124958), a Centre of Research Excellence Grant (PREMISE APP1134909), an NHMRC Career Development Fellowship to NCN (APP1166377), an NHMRC Early Career Fellowship to KEC (APP 1120641), and an NHMRC Leadership Fellowship (APP1195284) to MT. The funding body played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Data Availability

The data that support the findings of this study are available on reasonable request from the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary table.
[DOCX File , 14 KB - publichealth_v9i1e42244_app1.docx]

References

1. WHO's Global Health Estimates. World Health Organization. 2020. URL: https://www.who.int/data/global-health-estimates [accessed 2023-12-01]
2. Australian Institute of Health and Welfare. Australia's Health 2016. Canberra: Australian Institute of Health and Welfare; 2016.
3. Ezzati M, Riboli E. Behavioral and dietary risk factors for noncommunicable diseases. N Engl J Med 2013;369(10):954-964 [FREE Full text] [doi: 10.1056/NEJMra1203528] [Medline: 24004122]
4. Ding D, Rogers K, van der Ploeg H, Stamatakis E, Bauman AE. Traditional and emerging lifestyle risk behaviors and all-cause mortality in middle-aged and older adults: evidence from a large population-based Australian cohort. PLoS Med 2015;12(12):e1001917 [FREE Full text] [doi: 10.1371/journal.pmed.1001917] [Medline: 26645683]
5. Champion KE, Chapman C, Gardner LA, Sunderland M, Newton NC, Smout S, Heath4Life team. Lifestyle risks for chronic disease among Australian adolescents: a cross-sectional survey. Med J Aust 2022;216(3):156-157 [FREE Full text] [doi: 10.5694/mja2.51333] [Medline: 34747039]
6. Inchley J, Currie D, Budisavljevic S, Torsheim T, Jåstad A, Cosma A, et al. Spotlight on Adolescent Health and Well-being: Findings From the 2017/2018 Health Behaviour in School-aged Children (HBSC) Survey in Europe and Canada: International Report: Volume 2: Key Findings. Copenhagen: WHO Regional Office for Europe; 2020.

7. Cappuccio FP, Taggart FM, Kandala NB, Currie A, Peile E, Stranges S, et al. Meta-analysis of short sleep duration and obesity in children and adults. Sleep 2008;31(5):619-626 [FREE Full text] [doi: 10.1093/sleep/31.5.619] [Medline: 18517032]
8. Boers E, Afzali MH, Newton N, Conrod P. Association of screen time and depression in adolescence. JAMA Pediatr 2019;173(9):853-859 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.1759] [Medline: 31305878]
9. Loewen OK, Maximova K, Ekwaru JP, Faught EL, Asbridge M, Ohinmaa A, et al. Lifestyle behavior and mental health in early adolescence. Pediatrics 2019;143(5):e20183307 [FREE Full text] [doi: 10.1542/peds.2018-3307] [Medline: 31004047]
10. Juonala M, Viikari JSA, Raitakari OT. Main findings from the prospective cardiovascular risk in young finns study. Curr Opin Lipidol 2013;24(1):57-64 [FREE Full text] [doi: 10.1097/MOL.0b013e32835a7ed4] [Medline: 23069987]
11. Kaikkonen JE, Mikkilä V, Magnussen CG, Juonala M, Viikari JSA, Raitakari OT. Does childhood nutrition influence adult cardiovascular disease risk?--insights from the young finns study. Ann Med 2013;45(2):120-128 [FREE Full text] [doi: 10.3109/07853890.2012.671537] [Medline: 22494087]
12. Scandiffio JA, Janssen I. Do adolescent sedentary behavior levels predict type 2 diabetes risk in adulthood? BMC Public Health 2021;21(1):969 [FREE Full text] [doi: 10.1186/s12889-021-10948-w] [Medline: 34022833]
13. Conrod PJ. Personality-targeted interventions for substance use and misuse. Curr Addict Rep 2016;3(4):426-436 [FREE Full text] [doi: 10.1007/s40429-016-0127-6] [Medline: 27909645]
14. Kelly EV, Grummitt L, Teesson M, Newton NC. Associations between personality and uptake of tobacco smoking: do they differ across adolescence? Drug Alcohol Rev 2019;38(7):818-822. [doi: 10.1111/dar.12975] [Medline: 31418960]
15. Newton NC, Barrett EL, Castellanos-Ryan N, Kelly E, Champion KE, Stapinski L, et al. The validity of the substance use risk profile scale (SURPS) among Australian adolescents. Addict Behav 2016;53:23-30. [doi: 10.1016/j.addbeh.2015.09.015] [Medline: 26441044]
16. Woicik PA, Stewart SH, Pihl RO, Conrod PJ. The substance use risk profile scale: a scale measuring traits linked to reinforcement-specific substance use profiles. Addict Behav 2009;34(12):1042-1055. [doi: 10.1016/j.addbeh.2009.07.001] [Medline: 19683400]
17. Castellanos-Ryan N, O'Leary-Barrett M, Sully L, Conrod P. Sensitivity and specificity of a brief personality screening instrument in predicting future substance use, emotional, and behavioral problems: 18-month predictive validity of the substance use risk profile scale. Alcohol Clin Exp Res 2013;37(Suppl 1):E281-E290. [doi: 10.1111/j.1530-0277.2012.01931.x] [Medline: 22974180]
18. Guerrero MD, Barnes JD, Walsh JJ, Chaput JP, Tremblay MS, Goldfield GS. 24-hour movement behaviors and impulsivity. Pediatrics 2019;144(3):e20190187 [FREE Full text] [doi: 10.1542/peds.2019-0187] [Medline: 31413180]
19. Yañez AM, Bennasar-Veny M, Leiva A, García-Toro M. Implications of personality and parental education on healthy lifestyles among adolescents. Sci Rep 2020;10(1):7911 [FREE Full text] [doi: 10.1038/s41598-020-64850-3] [Medline: 32404935]
20. Dunn EC, Masyn KE, Jones SM, Subramanian SV, Koenen KC. Measuring psychosocial environments using individual responses: an application of multilevel factor analysis to examining students in schools. Prev Sci 2015;16(5):718-733 [FREE Full text] [doi: 10.1007/s11121-014-0523-x] [Medline: 25421872]
21. Corder K, Winpenny E, Love R, Brown HE, White M, Sluijs EV. Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. Br J Sports Med 2019;53(8):496-503 [FREE Full text] [doi: 10.1136/bjsports-2016-097330] [Medline: 28739834]
22. Koenen KC, Rudenstine S, Susser E, Galea S. A Life Course Approach to Mental Disorders. Oxford, United Kingdom: Oxford University Press; 2013.
23. Grant JE, Chamberlain SR. Sleepiness and impulsivity: findings in non-treatment seeking young adults. J Behav Addict 2018;7(3):737-742 [FREE Full text] [doi: 10.1556/2006.7.2018.71] [Medline: 30203693]
24. Burnett CM, Allen MS, Vella SA. Personality and sedentary behaviour in Australian adults. Int J Sport Exerc Psychol 2016;16(3):215-220. [doi: 10.1080/1612197x.2016.1212083]
25. Newton NC, Stapinski L, Slade T, Champion KE, Barrett EL, Chapman C, et al. Pathways to prevention: protocol for the CAP (Climate and Preventure) study to evaluate the long-term effectiveness of school-based universal, selective and combined alcohol misuse prevention into early adulthood. BMC Public Health 2018;18(1):643 [FREE Full text] [doi: 10.1186/s12889-018-5554-y] [Medline: 29783974]
26. Newton NC, Stapinski LA, Slade T, Sunderland M, Barrett EL, Champion KE, et al. The 7-year effectiveness of school-based alcohol use prevention from adolescence to early adulthood: a randomized controlled trial of universal, selective, and combined interventions. J Am Acad Child Adolesc Psychiatry 2022;61(4):520-532 [FREE Full text] [doi: 10.1016/j.jaac.2021.10.023] [Medline: 34823025]
27. Teesson M, Newton NC, Slade T, Chapman C, Birrell L, Mewton L, et al. Combined prevention for substance use, depression, and anxiety in adolescence: a cluster-randomised controlled trial of a digital online intervention. Lancet Digit Health 2020;2(2):e74-e84 [FREE Full text] [doi: 10.1016/S2589-7500(19)30213-4] [Medline: 33334564]
28. Nguyen B, Ding D, Mihrshahi S. Fruit and vegetable consumption and psychological distress: cross-sectional and longitudinal analyses based on a large Australian sample. BMJ Open 2017;7(3):e014201 [FREE Full text] [doi: 10.1136/bmjopen-2016-014201] [Medline: 28298322]

29. Rustihauser IHE, Webb K, Abraham B, Allsopp R. Evaluation of Short Dietary Questions From the 1995 National Nutrition Survey. Herston, Australia: Australian Food and Nutrition Monitoring Unit; 2001.
30. National Health and Medical Research Council. The Australian dietary guidelines. Australian Government, Department of Health and Aged Care. 2013. URL: https://www.health.gov.au/resources/publications/the-australian-dietary-guidelines [accessed 2023-06-01]
31. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-1395 [FREE Full text] [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
32. MacArthur GJ, Smith MC, Melotti R, Heron J, Macleod J, Hickman M, et al. Patterns of alcohol use and multiple risk behaviour by gender during early and late adolescence: the ALSPAC cohort. J Public Health (Oxf) 2012;34(Suppl 1):i20-i30 [FREE Full text] [doi: 10.1093/pubmed/fds006] [Medline: 22363027]
33. Gardner LA, Champion KE, Parmenter B, Grummitt L, Chapman C, Sunderland M, The Health Life Team, et al. Clustering of six key risk behaviors for chronic disease among adolescent females. Int J Environ Res Public Health 2020;17(19):7211 [FREE Full text] [doi: 10.3390/ijerph17197211] [Medline: 33023089]
34. Gunnell KE, Flament MF, Buchholz A, Henderson KA, Obeid N, Schubert N, et al. Examining the bidirectional relationship between physical activity, screen time, and symptoms of anxiety and depression over time during adolescence. Prev Med 2016;88:147-152. [doi: 10.1016/j.ypmed.2016.04.002] [Medline: 27090920]
35. Caspi A, Begg D, Dickson N, Harrington H, Langley J, Moffitt TE, et al. Personality differences predict health-risk behaviors in young adulthood: evidence from a longitudinal study. J Pers Soc Psychol 1997;73(5):1052-1063. [doi: 10.1037//0022-3514.73.5.1052] [Medline: 9364760]
36. Pinto Pereira SM, Geoffroy MC, Power C. Depressive symptoms and physical activity during 3 decades in adult life: bidirectional associations in a prospective cohort study. JAMA Psychiatry 2014;71(12):1373-1380 [FREE Full text] [doi: 10.1001/jamapsychiatry.2014.1240] [Medline: 25321867]
37. Marques ICF, Ting M, Cedillo-Martínez D, Pérez-Cueto FJA. Effect of impulsivity traits on food choice within a nudging intervention. Nutrients 2020;12(5):1402 [FREE Full text] [doi: 10.3390/nu12051402] [Medline: 32423143]
38. Maxwell AL, Gardiner E, Loxton NJ. Investigating the relationship between reward sensitivity, impulsivity, and food addiction: a systematic review. Eur Eat Disord Rev 2020;28(4):368-384. [doi: 10.1002/erv.2732] [Medline: 32142199]
39. Adams RJ, Appleton SL, Taylor AW, Gill TK, Lang C, McEvoy RD, et al. Sleep health of Australian adults in 2016: results of the 2016 sleep health foundation national survey. Sleep Health 2017;3(1):35-42. [doi: 10.1016/j.sleh.2016.11.005] [Medline: 28346149]
40. Kohyama J. Which is more important for health: sleep quantity or sleep quality? Children (Basel) 2021;8(7):542 [FREE Full text] [doi: 10.3390/children8070542] [Medline: 34202755]
41. Physical activity. Australian Government, Australian Institute of Health and Welfare. 2021. URL: https://www.aihw.gov.au/reports-data/behaviours-risk-factors/physical-activity/overview [accessed 2023-06-01]
42. Trott M, Driscoll R, Irlado E, Pardhan S. Changes and correlates of screen time in adults and children during the COVID-19 pandemic: a systematic review and meta-analysis. EClinicalMedicine 2022;48:101452 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101452] [Medline: 35615691]
43. Gardner LA, Debenham J, Newton NC, Chapman C, Wylie FE, Osman B, et al. Lifestyle risk behaviours among adolescents: a two-year longitudinal study of the impact of the COVID-19 pandemic. BMJ Open 2022 01;12(6):e060309 [FREE Full text] [doi: 10.1136/bmjopen-2021-060309] [Medline: 35649588]
44. Park JH, Moon JH, Kim HJ, Kong MH, Oh YH. Sedentary lifestyle: overview of updated evidence of potential health risks. Korean J Fam Med 2020;41(6):365-373 [FREE Full text] [doi: 10.4082/kjfm.20.0165] [Medline: 33242381]
45. Saunders TJ, Chaput JP, Tremblay MS. Sedentary behaviour as an emerging risk factor for cardiometabolic diseases in children and youth. Can J Diabetes 2014;38(1):53-61 [FREE Full text] [doi: 10.1016/j.jcjd.2013.08.266] [Medline: 24485214]
46. Hoare E, Milton K, Foster C, Allender S. The associations between sedentary behaviour and mental health among adolescents: a systematic review. Int J Behav Nutr Phys Act 2016;13(1):108 [FREE Full text] [doi: 10.1186/s12966-016-0432-4] [Medline: 27717387]
47. Gopinath B, Hardy LL, Baur LA, Burlutsky G, Mitchell P. Physical activity and sedentary behaviors and health-related quality of life in adolescents. Pediatrics 2012;130(1):e167-e174. [doi: 10.1542/peds.2011-3637] [Medline: 22689863]
48. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. Ann Intern Med 2015;162(2):123-132. [doi: 10.7326/M14-1651] [Medline: 25599350]
49. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia 2012;55(11):2895-2905 [FREE Full text] [doi: 10.1007/s00125-012-2677-z] [Medline: 22890825]
50. Castellanos N, Conrod P. Brief interventions targeting personality risk factors for adolescent substance misuse reduce depression, panic and risk-taking behaviours. J Ment Health 2009;15(6):645-658. [doi: 10.1080/09638230600998912]
51. Conrod PJ, Castellanos N, Mackie C. Personality-targeted interventions delay the growth of adolescent drinking and binge drinking. J Child Psychol Psychiatry 2008;49(2):181-190. [doi: 10.1111/j.1469-7610.2007.01826.x] [Medline: 18211277]

52. Conrod PJ, Castellanos-Ryan N, Strang J. Brief, personality-targeted coping skills interventions and survival as a non-drug user over a 2-year period during adolescence. Arch Gen Psychiatry 2010;67(1):85-93 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2009.173] [Medline: 20048226]
53. Newton NC, Conrod PJ, Slade T, Carragher N, Champion KE, Barrett EL, et al. The long-term effectiveness of a selective, personality-targeted prevention program in reducing alcohol use and related harms: a cluster randomized controlled trial. J Child Psychol Psychiatry 2016;57(9):1056-1065. [doi: 10.1111/jcpp.12558] [Medline: 27090500]
54. Edalati H, Conrod PJ. A review of personality-targeted interventions for prevention of substance misuse and related harm in community samples of adolescents. Front Psychiatry 2018;9:770 [FREE Full text] [doi: 10.3389/fpsyt.2018.00770] [Medline: 30723431]
55. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
56. Krank M, Stewart SH, O'Connor R, Woicik PB, Wall AM, Conrod PJ. Structural, concurrent, and predictive validity of the substance use risk profile scale in early adolescence. Addict Behav 2011;36(1-2):37-46. [doi: 10.1016/j.addbeh.2010.08.010] [Medline: 20826056]
57. National health survey: first results, 2017-18. Australian Bureau of Statistics. 2018. URL: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-first-results/latest-release [accessed 2023-06-01]
58. Australia's children. Australian Institute of Health and Welfare. 2020. URL: https://www.aihw.gov.au/reports/children-youth/australias-children/contents/health/the-health-of-australias-children [accessed 2023-06-01]
59. Australia's youth. Australian Institute of Health and Welfare. 2021. URL: https://www.aihw.gov.au/reports/children-youth/australias-youth/contents/about [accessed 2023-06-01]

Abbreviations

IPAQ-SF: International Physical Activity Questionnaire-Short Form
MVPA: moderate-to-vigorous physical activity
OR: odds ratio
PA: physical activity
RCT: randomized controlled trial
RR: risk ratio
SB: sedentary behavior
SURPS: Substance Use Risk Profile Scale

Edited by A Mavragani; submitted 06.10.22; peer-reviewed by T Baranowski, J de Brito, A Martinko; comments to author 22.11.22; revised version received 26.02.23; accepted 24.05.23; published 15.06.23.

Please cite as:
Champion KE, Chapman C, Sunderland M, Slade T, Barrett E, Kelly E, Stapinski L, Gardner LA, Teesson M, Newton NC
Associations Between Personality Traits and Energy Balance Behaviors in Emerging Adulthood: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e42244
URL: https://publichealth.jmir.org/2023/1/e42244
doi:10.2196/42244
PMID:37318870

Original Paper

The Association Between Family Health and Frailty With the Mediation Role of Health Literacy and Health Behavior Among Older Adults in China: Nationwide Cross-Sectional Study

Haomiao Li[1*], PhD; Yibo Wu[2*], MPharm; Zhongliang Bai[3,4*], PhD; Xiwu Xu[5], MD; Dai Su[6], PhD; Jiangyun Chen[7], PhD; Ruibo He[1,8], PhD; Ju Sun[1], PhD

[1]School of Political Science and Public Administration, Wuhan University, Wuhan, China
[2]School of Public Health, Peking University, Beijing, China
[3]School of Health Services Management, Anhui Medical University, Hefei, China
[4]Faculty of Education, Health and Wellbeing, University of Wolverhampton, Wolverhampton, United Kingdom
[5]Beijing Hospital, Beijing, China
[6]Department of Health Management and Policy, School of Public Health, Capital Medical University, Beijing, China
[7]Institute of Health Management, Southern Medical University, Guangzhou, China
[8]College of Finance and Public Administration, Hubei University of Economics, Wuhan, China
[*]these authors contributed equally

Corresponding Author:
Ju Sun, PhD
School of Political Science and Public Administration
Wuhan University
Bayi road No.299
Wuhan, 430000
China
Phone: 86 18971431636
Email: sunphh@126.com

Abstract

Background: Family health develops from the intersection of the health of each family member and their interactions and capacities as well as the family's internal and external resources. Frailty is the most prominent and typical clinical manifestation during population aging. Family health may be effective in addressing frailty, and this association may be mediated by health literacy and health behaviors. Until now, it is unclear whether and how family health affects frailty in older adults.

Objective: This study aimed to examine the associations between family health and frailty and the mediation roles of health literacy and health behaviors.

Methods: A total of 3758 participants aged ≥60 years were recruited from a national survey conducted in 2022 in China for this cross-sectional study. Family health was measured using the Short Form of the Family Health Scale. Frailty was measured using the Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight (FRAIL) scale. Potential mediators included health literacy and health behaviors (not smoking, not having alcohol intake, physical exercise for ≥150 minutes per week, longer sleep duration, and having breakfast every day). Ordered logistic regression was applied to explore the association between family health and frailty status. Mediation analysis based on Sobel tests was used to analyze the indirect effects mediated by health literacy and behaviors, and the Karlson-Holm-Breen method was used to composite the indirect effects.

Results: Ordered logistic regression showed that family health is negatively associated with frailty (odds ratio 0.94, 95% CI 0.93-0.96) with covariates and potential mediators controlled. This association was mediated by health literacy (8.04%), not smoking (1.96%), longer sleep duration (5.74%), and having breakfast every day (10.98%) through the Karlson-Holm-Breen composition.

Conclusions: Family health can be an important intervention target that appears to be negatively linked to frailty in Chinese older adults. Improving family health can be effective in promoting healthier lifestyles; improving health literacy; and delaying, managing, and reversing frailty.

(JMIR Public Health Surveill 2023;9:e44486) doi:10.2196/44486

KEYWORDS

family health; frailty; health literacy; health behavior; healthy aging

Introduction

Background

The rapid aging of populations will lead to a greater chronic disease burden on the whole society [1]. Diseases such as hypertension, diabetes, cancer, stroke, arthritis, dementia, and related disease burdens will continue to increase [2]. From the perspective of disease occurrence and development, the immune functions of older adults gradually decrease with age, their susceptibility to adverse stimuli increases, and disease incidence rates will inevitably increase. However, there are substantially individual differences in the occurrence, development, prognosis, and progression of diseases and the consequent quality of life. These differences are mainly due to frailty [3,4]. *Frailty* refers to a homeostatic imbalance in the body, between unhealthy and serious damaging status due to a decline in physical, psychological, or cognitive function. It is the most prominent and typical clinical manifestation during population aging [5]. In addition, "prefrailty," which is defined as a complex multifactorial and multidimensional state associated with physiological and deleterious processes that develop over time, is an intermediate stage in the progression from robust to frailty [6]. Compared with frailty, prefrailty has a higher prevalence but a lower level of damage to older adults. In addition, prefrailty is more clinically reversible [7,8], which highlights the importance of early identification and intervention to reverse it or prevent the progression to frailty. Previous studies have shown that both prefrailty and frailty are associated with an increased risk of hospitalization, functional decline, progression to long-term care, and death [9]. Delaying the occurrence of prefrailty or frailty and promoting their reversal can play a substantial role in the promotion of healthy aging.

Family-oriented health promotions are promising and effective strategies because the family unit is both a resource and a priority group that needs preventive and curative services across its members' life course [10]. In particular, "family health" as a relatively new concept is receiving more attention from scholars. Weiss-Laxer et al [11] defined family health as a resource at the family unit level that develops from the intersection of the health of each family member and their interactions and capacities with the family's physical, social, emotional, economic, and medical resources. Family health integrates the key elements of the previous concepts of family structure, family function, and family social network; seeks to strengthen the ability of families to obtain external resources and sociality; and emphasizes health-related elements, linking individual health with social health. Therefore, family health may be an important target for health intervention. Regarding strategies for preventing and controlling frailty, researchers have yet to determine whether interventions through family health are effective.

Health literacy and health behaviors may be effective pathways to address the association between family health and frailty. Health literacy is defined as a cognitive and social skill that determines an individual's motivation and ability to access, understand, and use information in a way that promotes and maintains health. Low health literacy is associated with a poor understanding of one's medical condition, poor self-care, delayed care seeking, and lower use of preventive services; it can also affect disease management and outcomes in patients with chronic conditions [12,13]. Adequate health literacy can prevent frailty in older adults and plays a positive role in intervention and the management of frail, community-dwelling older adults [14,15]. Residents' health literacy may be affected by family structure, income level, information delivery, and other family members' education levels. Behaviors are the most concrete, visible aspect of family functioning. Family structure, processes, and cognitions are expressed through family behaviors [16]. There are a range of mechanisms underlying relationships between family and health behaviors, including promoting health-seeking or health treatment behaviors providing access, opportunities, and resources for a range of health behaviors [17]. Previous studies have also demonstrated the relationship between healthy behaviors, such as diet and exercise, and frailty [5,18]. Nevertheless, the associations between family health, health literacy, and health behaviors and frailty have not been explored.

Objective and Hypotheses

China is one of the countries with the fastest aging population and faces unprecedented challenges in the face of such a rapid aging process. Therefore, frailty prevention and management is essential and urgent in China, and family health may be an effective subject for intervention. This study explored whether and how family health could decrease the frailty risk in older populations. We proposed two hypotheses: (1) family health is positively associated with the prevalence of prefrailty and frailty and (2) the association between family health and frailty status is mediated by health literacy and health behaviors.

Methods

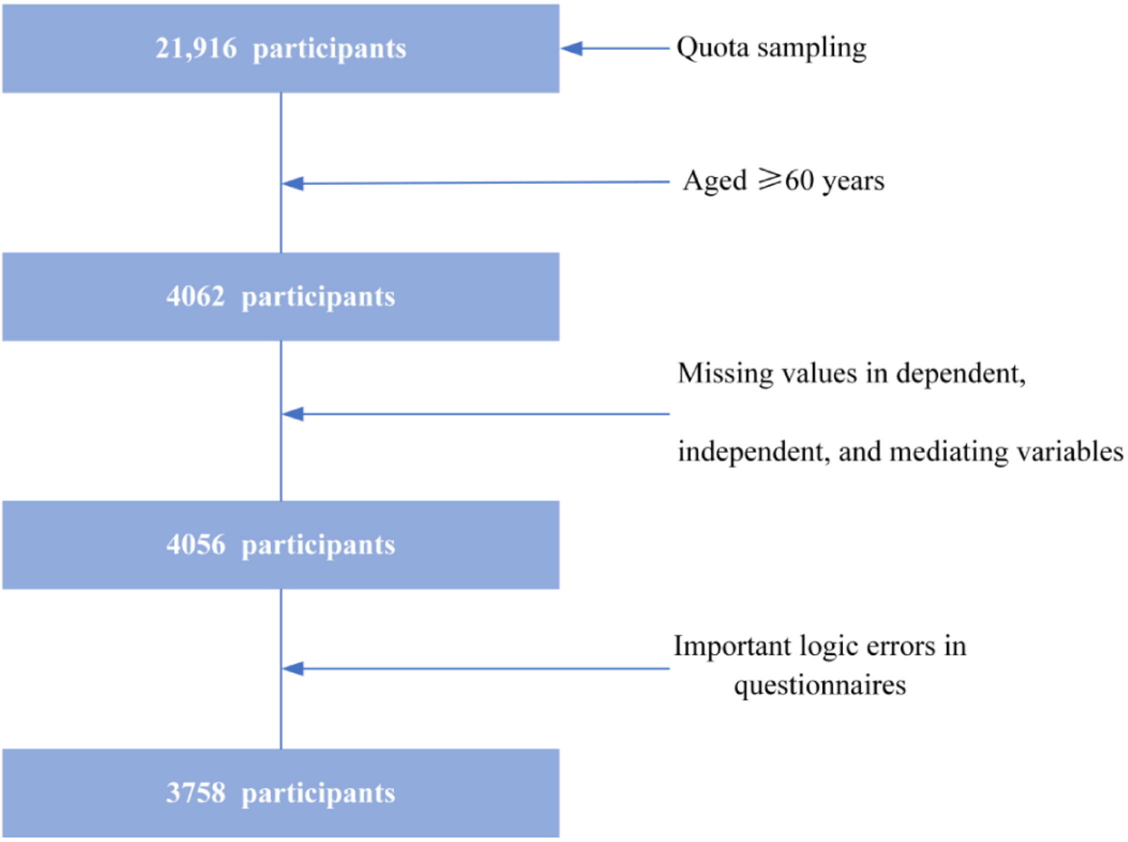
Sampling and Participants

The data are from a national survey, conducted from June 20, 2022, to August 31, 2022, in 148 cities; 202 districts and counties; 390 townships, towns, or streets; 780 communities or villages (excluding Hong Kong, Macao, and Taiwan) from 23 provinces; 5 autonomous regions; and 4 municipalities directly under the central government in China, initiated by the Peking University School of Public Health. The sampling ratio was determined based on the population proportion provided by the seventh national census data. At least 500, 1000, 1500, 2000, or 2500 individuals were sampled from each province, autonomous region, or municipality directly under the central government. The sample size was estimated to be 20,000. Finally, in the municipal, district, county, township or town, street, and community level, individual were sampled according to quota attribute, including sex and age.

To conduct the survey, the investigator set up a questionnaire survey site in the health service center or a relevant health service station of the sampling community in charge and then posted a poster and issued a paper or electronic invitation to recruit respondents. The investigator checked the identities of all respondents, solicited their informed consent, and determined whether the respondents met the inclusion criteria for the study participants. The targeted participants had to be aged >12 years, able to understand each item of the questionnaire, and able to complete the questionnaire on their own or with the help of an investigator. People who were confused, were experiencing mental health difficulties or cognitive impairment, or were unwilling to participate in the survey were excluded. All the participants voluntarily participated in the study and signed an informed consent form. In total, 21,916 questionnaires were collected after quota sampling. The survey protocol has been published [19].

In this study, the data of participants aged <60 years; those whose questionnaires did not include details on the independent, dependent, and potential mediating variables; and those whose questionnaires had logical errors (mainly contradictions and discrepancies between important variables such as age, family type, and marital status) were also excluded (details are shown in Figure 1). In total, 3758 participants were included in the final analysis.

Figure 1. Sample selection process for this cross-sectional study.



Variables

Frailty

Frailty was measured using the Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight (FRAIL) scale [20], which has been validated in older Chinese populations (Multimedia Appendix 1) [21-23]. We assessed the presence of fatigue and loss of body weight by participants responding "yes" to the following items in the self-reported questionnaire: "Were you tired most of last week?" and "Have you experienced an unexplained loss of more than 5% of your body weight in the last year?" The presence of resistance and ambulation problems were assessed by a "yes" answer to the following questions: "Can you go up a staircase?" and "Can you walk a block (500 meters) away?" Illness was assessed based on the total number of chronic diseases that participants had (at least 5 diseases had to be present), and the number was then dichotomized into a binary variable. All the abovementioned 5 variables were coded as 0 ("no") or 1 ("yes"), with 1 indicating the presence of deficits. The total deficits were summed to calculate a frailty score that ranged from 0 to 5. On the basis of previous studies, participants who scored 0 were defined as robust, those who scored 1 or 2 were defined as prefrail, and those who scored ≥3 were defined as frail [24].

Family Health

Family health was measured using the Short Form of the Family Health Scale, which was translated into Chinese with the consent of the original author. The Chinese version of the Short Form of the Family Health Scale has good reliability and validity and can be used to assess the level of family health of Chinese residents [25]. It contains four dimensions: (1) family, social, or emotional health processes; (2) family healthy lifestyle; (3) family health resources; and (4) family external social support. The scores of the items and calculations for family health ranged from 10 to 50, with a higher score indicating better family health (Table 1). Cronbach α of the sample in this study was .634, which was deemed as indicating acceptable reliability, and the Kaiser-Meyer-Olkin value was 0.702, which was deemed as indicating acceptable validity.

Table 1. Short Form of the Family Health Scale.

Item number[a]	Dimensions	Items
1	Family, social, or emotional health processes	We support each other.
2	Family, social, or emotional health processes	I feel safe in my family relationships.
3	Family, social, or emotional health processes	We stay hopeful even in difficult times.
4	Family healthy lifestyle	We help each other in seeking health care services when needed (such as making physician's appointments).
5	Family healthy lifestyle	We help each other make healthy changes.
6	Family health resources	We do not trust doctors and other health professionals.
7	Family health resources	My family did not have enough money at the end of the month after bills were paid.
8	Family health resources	My family did not have adequate housing.
9	Family external social supports	We have people outside of our family we can turn to when we have problems at school or work.
10	Family external social supports	If we needed financial help, we have people outside of our family we could turn to for a loan (eg, for RMB 1000 [US $140]).

[a] Items 1 to 5 and 9 to 10 were positively scored (strongly disagree=1, somewhat disagree=2, neither agree nor disagree=3, somewhat agree=4, and strongly agree=5), and items 6 to 8 were negatively scored (strongly disagree=5, somewhat disagree=4, neither agree nor disagree=3, somewhat agree=2, and strongly agree=1). The total score is the sum of each item, with a score ranging from 10 to 50, with a higher score indicating better family health.

Mediators

The mediators measured in this study were health literacy and health behaviors. Health literacy was measured using the new short-form health literacy instrument, which included 9 items (Multimedia Appendix 2). Scores ranged from 9 to 36 points, with a high score indicating higher levels of health literacy. Health behaviors included smoking, alcohol intake, physical exercise, sleep, and diet. Smoking status and alcohol intake were set as binary variables (1=yes and 0=no). Physical exercise was assessed by weekly exercise time, including aggravating activities (power sports, fast running, ball games, aerobics, and fast cycling), moderate-intensity physical activity (handling goods, medium-speed cycling, jogging, and table tennis, excluding walking), and light aerobic exercise (walking for at least 10 min). Exercise time was calculated by summing up all activity time (minute), and ≥150 minutes was deemed as a healthy style (0: <150 minutes and 1: ≥150 minutes) [26]. Sleep behavior was assessed by sleep duration per night (1: <5 hours, 2: 5-6 hours, 3: 6-7 hours, and 4: >7 hours). Diet behavior was measured based on whether the participant had breakfast every day (1=yes and 0=no).

Covariates

The covariates were initially identified based on previous studies and general knowledge. Then, the enrolled covariates were selected based on their association with the independent variable and their impact on the change in the association between the independent and dependent variables. Age, sex, and potential mediators were included as fixed covariates to be controlled. Other covariates were included as potential confounders in the final models if they changed the estimates of the effect of family health on frailty status by >10% or were significantly associated with frailty status based on generalized linear regression [27]. The final covariates included age, sex (0=male and 1=female), family type (1=core family, 1=backbone family, 2=joint family, 3=conjugal family, 4=single-parent family, and 5=other), residence (0=urban and 1=rural), marital status (0=married and 1=divorced, widowed, or unmarried), number of children (grouped by 0, 1, 2, and ≥3), public insurance coverage (0=not covered and 1=covered), and BMI and self-rated health (scoring from 0 to 100).

Statistical Analysis

Sociodemographic characteristics, health literacy, health behaviors, and family health of participants were summarized using frequencies (percentages) or means and SDs, grouped by frailty status. Statistical differences were tested using Kruskal-Wallis one-way analysis for the continuity variables and chi-square tests for the categorical variables.

Ordered logistic regressions were applied to assess the association between family health, including its 4 dimensions; health behaviors; and frailty, with covariates controlled. Odds ratios (ORs) and 95% Cis for the risk of frailty were estimated. A generalized additive model and smoothing curve fitting were further used to address the potential nonlinear relationship between family health and the risk of frailty, with potential mediators and covariates controlled. The generalized additive model allows us to fit the model using a nonlinear smoothing term without prior knowledge of the relationship between the dependent and independent variables, which allows the association to be visualized more intuitively [28].

Sobel tests were then applied to measure the significance of the mediating effects of health literacy and health behaviors. To assess the effect sizes of the mediators, Karlson-Holm-Breen (KHB) methods were used to composite the indirect effects with the ordered logistic model. For health behaviors, the healthier groups were used as references.

The significance level was set at a *P* value of <.05 for all the hypothesis tests. Data were analyzed using Stata (version 17.0; StataCorp) [29] and R (version 3.6.3; R Foundation for Statistical Computing).

Sensitivity and Subgroup Analysis

Several sensitivity analyses were conducted to validate the results. First, as the prevalence of frailty was relatively low in this study, we set the absolute value of the frailty score (0-5 points) as the dependent variable and repeated the KHB decomposition methods. As the frailty score followed a Poisson distribution, we applied negative binomial regression as the model for analysis during the KHB decomposition. Second, in the questionnaire, physical exercise included some high-intensity exercises, which may not be universal or normal among older populations. Therefore, we used the number of days walking (at least 10 minutes) per week instead of physical exercise time. Third, the optimal sleep duration for older adults was not consistent with previous studies. Therefore, we used self-rated sleep quality (0=*very bad*, 1=*bad*, 2=*good*, and 3=*very good*) instead of sleep duration.

Furthermore, the definition of “older adults” was varied in recent studies. The participants in this study were adults aged ≥60 years, which was consistent with some previous studies and with the retirement policy in China [30,31]. However, in geriatrics, the definition for older adults is age ≥65 years [32]. In cardiovascular medicine, “older” was defined as age ≥75 years [33]. Therefore, to better clarify the association between family health and frailty in “true” older adults, we performed subgroup analysis in different age groups, including “≥65 years old” and “≥75 years old.”

Ethics Approval and Informed Consent

This quantitative study was performed in accordance with the guideline “involves people of biomedical research ethics review method (try out)” of the China Ministry of Health national drug supervision and administration of the quality control standard for clinical trials (2003), medical instrument clinical trial regulations (2004), and the Declaration of Helsinki. The investigators obtained ethics approval from the Shaanxi Institute of International Trade and Commerce (JKWH-2022-02). All applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed over the course of this study. All interviewees provided written informed consent to participate in this study upon recruitment, and they participated voluntarily without compensation. All the study data were anonymous.

Results

Demographic Characteristics

Among the 3758 participants, 2327 (61.92%) were robust, 1342 (35.71%) were prefrail, and 89 (2.37%) were frail. The mean score of family health was 37.86 (SD 6.44), with 39.45 (SD 6.05), 35.18 (SD 6.25), and 36.63 (SD 5.53) in the robust, prefrail, and frail groups, respectively. We further compared the mean scores of family health between the prefrail and frail groups, and the difference was not significant (*P*=.237). Participants in the frail group were older than those in the other groups. The ratios for illiteracy; divorced, widowed, or unmarried; and smoking were also higher in the frail group than in the other groups, whereas self-rated health, health literacy, and the ratio of usually exercising (≥150 min/wk) were lower than those of the other groups (Table 2).

Table 2. Demographic characteristics and family health levels for participants aged ≥60 years by frailty status.

	All (N=3758)	Robust (n=2327)	Prefrail (n=1342)	Frail (n=89)	*P* value
Age (years), mean (SD)	68.81 (6.26)	68.28 (6.03)	69.37 (6.40)	74.33 (6.92)	<.001
Sex, n (%)					.51
Male	1866 (49.65)	1173 (50.41)	650 (48.44)	43 (48.31)	
Female	1892 (50.35)	1154 (49.59)	692 (51.56)	46 (51.69)	
Family type[a], n (%)					<.001
Core family	336 (8.94)	190 (8.17)	139 (10.36)	7 (7.86)	
Backbone family	1876 (49.92)	1228 (52.77)	612 (45.6)	36 (40.45)	
Joint family	237 (6.31)	138 (5.93)	92 (6.86)	7 (7.86)	
Conjugal family	897 (23.87)	587 (25.23)	286 (21.31)	24 (26.97)	
Single-parent family	143 (3.81)	67 (2.88)	70 (5.22)	6 (6.74)	
Other	269 (7.16)	117 (5.03)	143 (10.66)	9 (10.11)	
Residence, n (%)					<.001
Urban	2110 (56.15)	1408 (60.51)	664 (49.48)	38 (42.69)	
Rural	1648 (43.85)	919 (39.49)	678 (50.52)	51 (57.3)	
Marital status, n (%)					<.001
Married	3200 (85.15)	2034 (87.41)	1102 (82.12)	64 (71.91)	
Divorced, widowed, or unmarried	558 (14.85)	293 (12.59)	240 (17.88)	25 (28.09)	
Number of children, n (%)					<.001
0	277 (7.37)	96 (4.13)	175 (13.04)	6 (6.74)	
1	1212 (32.25)	780 (33.52)	420 (31.3)	12 (13.48)	
2	1232 (32.78)	804 (34.55)	402 (29.96)	26 (29.21)	
≥3	1037 (27.59)	647 (27.8)	345 (25.71)	45 (50.56)	
Public insurance coverage, n (%)					<.001
No	297 (7.9)	113 (4.86)	180 (13.41)	4 (4.49)	
Yes	3461 (92.1)	2214 (95.14)	1162 (86.59)	85 (95.51)	
BMI (kg/m^2), mean (SD)	21.91 (3.61)	22.10 (3.57)	21.62 (3.67)	21.38 (3.52)	<.001
Self-rated health, mean (SD)	70.08 (18.35)	73.92 (16.99)	64.41 (18.42)	55.10 (21.60)	<.001
Health literacy, mean (SD)	25.04 (5.29)	26.05 (4.96)	23.47 (5.37)	22.09 (5.41)	<.001
Physical exercise (minutes per week), n (%)					<.001
<150	557 (14.82)	275 (11.82)	246 (18.33)	36 (40.45)	
≥150	3201 (85.18)	2052 (88.18)	1096 (81.67)	53 (59.55)	
Alcohol intake, n (%)					.18
No	3237 (86.14)	2019 (86.76)	1144 (85.25)	74 (83.15)	
Yes	521 (13.86)	308 (13.24)	198 (14.75)	15 (16.85)	
Smoking status, n (%)					<.001
No	3125 (83.16)	2019 (86.76)	1040 (77.5)	66 (74.16)	
Yes	633 (16.84)	308 (13.24)	302 (22.5)	23 (25.84)	
Sleep duration per night (hours), n (%)					<.001
<5	255 (6.79)	88 (3.78)	152 (11.33)	15 (16.85)	
5-6	801 (21.31)	389 (16.72)	391 (29.14)	21 (23.6)	
6-7	1271 (33.82)	794 (34.12)	453 (33.76)	24 (26.97)	

	All (N=3758)	Robust (n=2327)	Prefrail (n=1342)	Frail (n=89)	*P* value
>7	1431 (38.08)	1056 (45.38)	346 (25.78)	29 (32.58)	
Have breakfast every day, n (%)					<.001
No	939 (24.99)	381 (16.37)	529 (39.42)	29 (32.58)	
Yes	2819 (75.01)	1946 (83.63)	813 (60.58)	60 (67.41)	
Family health, mean (SD)	37.86 (6.44)	39.45 (6.05)	35.18 (6.25)	36.63 (5.53)	<.001

[a] "Core family" refers to a family consisting of parents and unmarried children; "Conjugal family" refers to a family consisting of parents and married children; "Backbone family" refers to a family consisting of only husband and wife; "Joint family" means a family of parents or more married children or siblings after marriage; "Single-parent family" means a family consisting of divorced, widowed, or unmarried single fathers or mothers and their children or adopted children; and "Other" consists of the following: "Intergenerational family," referring to a family with only 2 generations, and the parents left the family for some reasons; "Dink family," referring to a voluntary infertile family consisting of 2 couples; "Single family," referring to not being married at the age of marriage or not married after divorce but living alone; and reformed families, cohabitation families, and gay families.

The Association Between Family Health and Frailty

The ordered logistic regression (Table 3) showed that better family health was associated with a lower risk of frailty (OR 0.93, 95% CI 0.91-0.94), with covariates adjusted (model 1). After potential mediators were entered into the model (model 2), the association between family health and frailty was still significant (OR 0.94, 95% CI 0.93-0.96; $P<.001$). Model 2 had a greater pseudo R^2 and lower Akaike information criterion and Bayesian information criterion, which indicated better model fitting performance. Meanwhile, except for no alcohol intake ($P=.74$), high health literacy ($P=.01$), not smoking ($P=.009$), physical exercise for ≥150 minutes per week ($P<.001$), more than 5 hours of sleep (all $P<.05$), and having breakfast every day ($P<.001$) were significantly associated with a decreased risk of frailty.

After adjusting for covariates and potential mediators, the smooth curve based on the generalized additive model suggested that family health tended to be linearly associated with frailty risk, and frailty risk decreased with increased family health (Figure 2).

We further analyzed the associations between the 4 dimensions of family health and frailty; the increases in all dimensions were associated with decreased frailty risk (Multimedia Appendix 3).

Table 3. The association between family health and frailty for participants aged ≥60 years explored by ordered logistic regression.

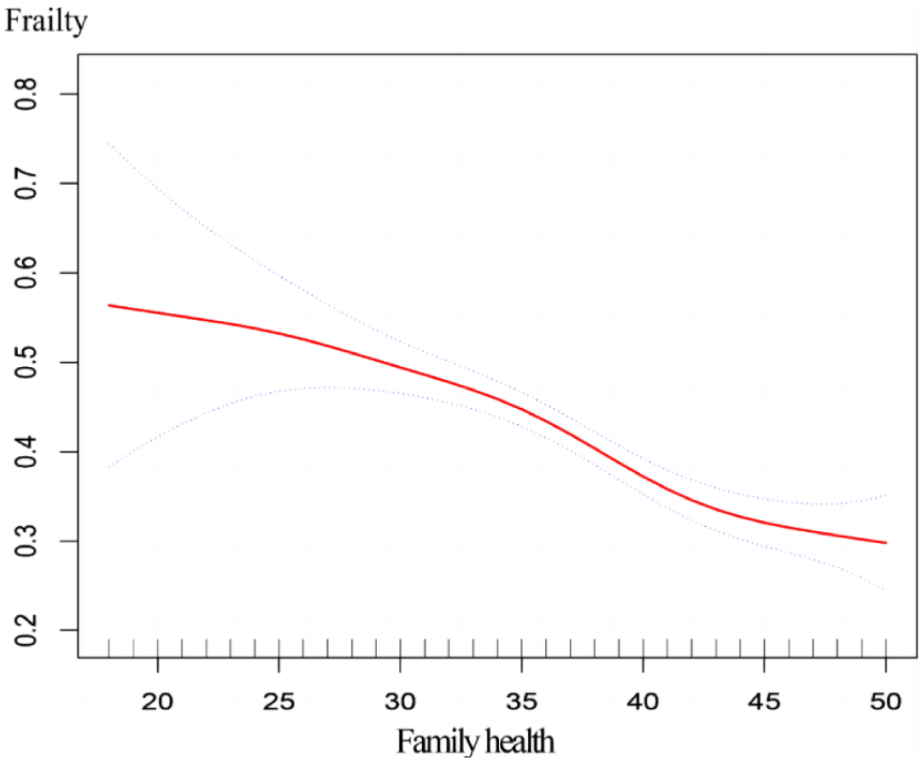
	Model 1		Model 2	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value
Family health	0.93 (0.91-0.94)	<.001	0.94 (0.93-0.96)	<.001
Age	1.05 (1.04-1.07)	<.001	1.05 (1.04-1.06)	<.001
Sex (ref[b]: male)	1.13 (0.98-1.31)	.08	1.15 (0.99-1.35)	.07
BMI	0.99 (0.97-1.01)	.22	0.99 (0.97-1.01)	.45
Family type (ref: core family)				
Backbone family	0.70 (0.54-0.90)	.005	0.65 (0.5-0.84)	.001
Joint family	1.08 (0.75-1.56)	.67	0.95 (0.66-1.37)	.79
Conjugal family	0.61 (0.46-0.81)	.001	0.59 (0.44-0.79)	<.001
Single-parent family	1.01 (0.63-1.60)	.98	0.88 (0.55-1.42)	.60
Other	0.94 (0.65-1.36)	.74	0.84 (0.57-1.22)	.36
Residence (ref: urban)	1.31 (1.13-1.52)	<.001	1.24 (1.06-1.45)	.006
Marital status (ref: married)	1.06 (0.83-1.34)	.65	1.05 (0.82-1.34)	.70
Number of children (ref: 0)				
1	0.70 (0.52-0.96)	.03	0.95 (0.69-1.32)	.78
2	0.66 (0.48-0.90)	.009	0.98 (0.7-1.37)	.92
≥3	0.59 (0.43-0.82)	.001	0.86 (0.61-1.22)	.41
Self-rated health	0.98 (0.97-0.98)	<.001	0.98 (0.98-0.98)	<.001
Public insurance coverage (ref: not covered)	0.53 (0.41-0.69)	<.001	0.6 (0.46-0.78)	<.001
Health literacy	N/A[c]	N/A	0.98 (0.96-0.99)	.01
Smoking status (ref: smoke)	N/A	N/A	0.76 (0.61-0.93)	.009
Alcohol intake (ref: drink)	N/A	N/A	1.04 (0.83-1.30)	.74
Physical exercise (ref: <150 min/wk)	N/A	N/A	0.57 (0.47-0.70)	<.001
Sleep duration per night (ref: <5 hours)				
5-6 hours	N/A	N/A	0.64 (0.48-0.87)	.004
6-7 hours	N/A	N/A	0.42 (0.31-0.56)	<.001
>7 hours	N/A	N/A	0.31 (0.23-0.41)	<.001
Have breakfast every day (ref: no)	N/A	N/A	0.63 (0.52-0.75)	<.001
Pseudo R^2	0.121	N/A	0.16	N/A
Akaike information criterion	5013.02	N/A	4833.66	N/A
Bayesian information criterion	5125.19	N/A	4995.69	N/A

[a] OR: odds ratio.

[b] ref: reference.

[c] N/A: not applicable.

Figure 2. Smooth curve fitting for the association between family health and frailty, with all covariates and potential mediators controlled.



Mediation Analysis

Mediation analysis revealed that the total effect of family health on frailty was −0.016, and the association between family health and frailty was mediated by health literacy, not smoking, sleep duration, and having breakfast every day. The Sobel test results were significant (*P*<.001). The KHB composition revealed that the 4 mediators reduced the total effect of family health on frailty by 26.72%. In addition, the mediating effects of health literacy and having breakfast constituted 8.04% and 10.98%, respectively, which were much higher than those of the other 3 mediators. Further details are presented in Figure 3 and Table 4.

We explored the mediating effects of health literacy, not smoking, sleep duration, and having breakfast on the 4 dimensions of family health and frailty (Table 5). The results were similar, although the mediating effect of sleep duration on the association between family health resources and frailty was not significant (*P*=.354).

Figure 3. Mediation analysis of the association between family health and frailty. ***Indicates that the statistically significant association at α=.001 level.

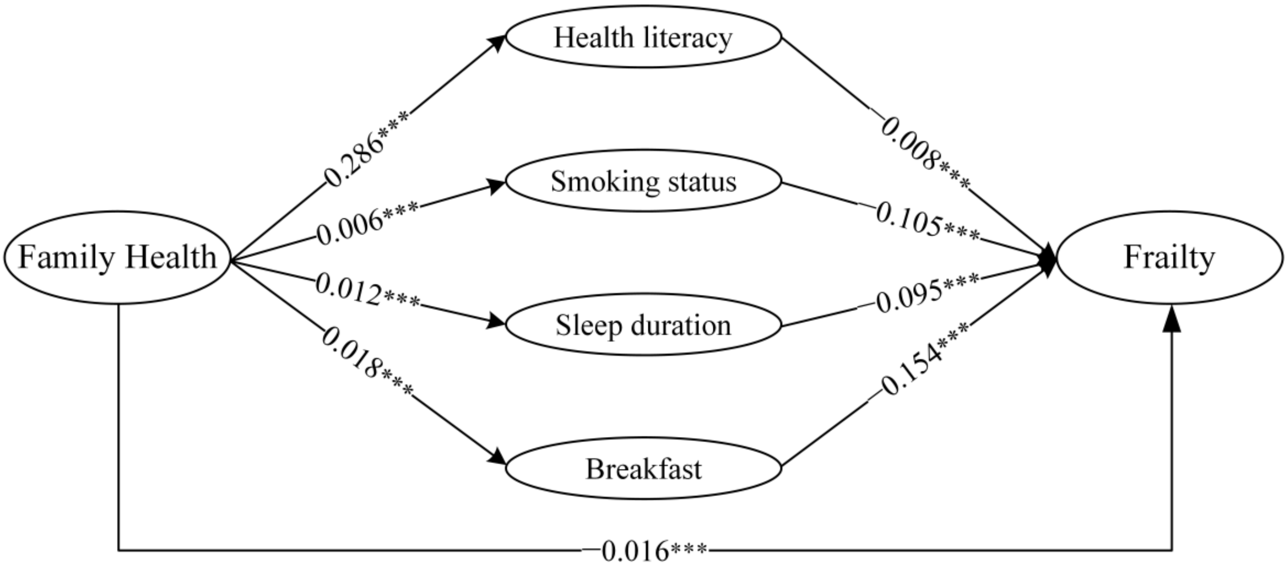


Table 4. The mediating effect of health literacy and health behaviors on family health and frailty explored by Sobel tests and Karlson-Holm-Breen (KHB) decomposition methods.

	Health literacy	Smoking status	Alcohol intake	Physical exercise	Sleep duration	Breakfast
Family health→mediator, β	.286[a]	.006[a]	.001	−.002	.012[a]	.018[a]
Mediator→frailty, β	−.008[a]	−.105[a]	−.023	−.138[a]	−.095[a]	−.154[a]
Indirect effect, β	−.002[a]	−.001[a]	−.000	.000	−.001[a]	−.003[a]
Direct effect, β	−.014[a]	−.016[a]	−.016[a]	−.017[a]	−.015[a]	−.014[a]
Total effect, β	−.016[a]	−.016[a]	−.016[a]	−.016[a]	−.016[a]	−.016[a]
Proportion of total effect that is mediated	0.147	0.042	0.001	−0.016	0.070	0.166
Sobel test	−0.002[a]	−0.001[a]	−0.000	−0.000	−0.001[a]	−0.003[a]
KHB decomposition						
Proportion of mediation effect (%)	8.04	1.96	__[b]	—	5.74	10.98

[a] *P*<.001.

[b] Alcohol intake and physical exercise were not taken into KHB analysis as the mediation effects were not significant.

Table 5. Mediation analysis and Karlson-Holm-Breen (KHB) decomposition of health literacy and health behaviors mediating the associations between dimensions of family health and frailty.

	Health literacy	Smoking status	Sleep duration	Breakfast
Family, social, or emotional health processes				
Indirect effect, β	−.007[a]	−.002[a]	−.003[a]	−.007[a]
Direct effect, β	−.029[a]	−.034[a]	−.034[a]	−.030[a]
Total effect, β	−.037[a]	−.037[a]	−.037[a]	−.037[a]
KHB decomposition				
Proportion of mediation effect (%)	10.78	2.89	6.20	13.23
Family healthy lifestyle				
Indirect effect, β	−.012[a]	−.004[a]	−.004[a]	−.010[a]
Direct effect, β	−.032[a]	−.041[a]	−.040[a]	−.034[a]
Total effect, β	−.044[a]	−.044[a]	−.044[a]	−.044[a]
KHB decomposition				
Proportion of mediation effect (%)	16.00	3.67	7.86	16.31
Family health resources				
Indirect effect, β	.003[a]	.001[b]	−.000	−.002[a]
Direct effect, β	−.019[a]	−.017[a]	−.015[a]	−.014[a]
Total effect, β	−.016[a]	−.016[a]	−.016[a]	−.016[a]
KHB decomposition				
Proportion of mediation effect (%)	−13.33	−3.99	2.47	7.57
Family external social supports				
Indirect effect, β	−.012[a]	−.003[a]	−.003[a]	−.008[a]
Direct effect, β	−.026[a]	−.035[a]	−.035[a]	−.030[a]
Total effect, β	−.038[a]	−.038[a]	−.038[a]	−.038[a]
KHB decomposition				
Proportion of mediation effect (%)	17.23	3.67	6.63	14.01

[a] *P*<.001.

[b] *P*<.01.

Sensitivity Analysis

In sensitivity analysis (Multimedia Appendix 4), with deficits of frailty (scoring from 0 to 5) set as the dependent variable, KHB methods were used, and a negative abnormal regression model was used. The results were consistent with those of the main analysis. When using walking (at least 10 minutes) days per week to replace physical exercise time, the mediating effect was significant (*P*<.001), and the path to family health was positively associated with physical exercise and further decreased frailty risk. When using self-rated sleep quality to replace sleep duration, the results revealed that better family health was associated with better sleep quality and further decreased frailty risk, with a significant mediating effect (*P*<.001).

Subgroup Analysis

To better clarify the association between family health and frailty in “true” older adults, we performed subgroup analyses in different age groups, including “≥65 years old” and “≥75 years old.” First, through ordered logistic regressions, the negative associations between family health and frailty were significant in the 2 subgroups (for those aged ≥65 years: OR 0.94, 95% CI 0.93-0.96, *P*<.001; for those aged ≥75 years: OR 0.96, 95% CI 0.93-0.99, *P*=.007). For those aged ≥65 years, the mediating effects of health literacy, smoking status, sleep duration, and breakfast remained significant. However, for those aged ≥75 years, only the mediating effect of breakfast was significant (for health literacy: *P*=.554; for smoking status: *P*=.266; for sleep duration: *P*=.081). The details are presented in Table 6.

Table 6. Table6. Subgroup mediation analysis stratified by age based on Sobel tests and Karlson-Holm-Breen (KHB) decomposition methods.

	Health literacy	Smoking status	Sleep duration	Breakfast
≥65 years				
Indirect effect, β	−.002[a]	−.001[a]	−.001[b]	−.002[b]
Direct effect, β	−.015[b]	−.016[b]	−.016[b]	−.014[b]
Total effect, β	−.017[b]	−.017[b]	−.017[b]	−.017[b]
KHB decomposition				
Proportion of mediation effect (%)	4.87	1.33	5.92	9.30
≥75 years				
Indirect effect, β	−.001	−.000	−.001	−.003[a]
Direct effect, β	−.014[b]	−.014[b]	−.013[b]	−.012[a]
Total effect, β	−.014[b]	−.014[b]	−.014[b]	−.014[b]
KHB decomposition				
Proportion of mediation effect (%)	—[c]	—	—	18.48

[a] *P*<.01.

[b] *P*<.001.

[c] Variables did not taken into KHB decomposition analysis as their mediating effects were not significant.

Discussion

Principal Findings

To our knowledge, this is the first study to assess and divide the effects of family health on frailty. The results demonstrate that family health is negatively associated with the prevalence of prefrailty and frailty, and this association can be mediated by health literacy and certain health behaviors (eg, not smoking, longer sleep duration, and having breakfast every day).

The prevalence of prefrailty and frailty in this study was 35.71% and 2.37%, respectively, which was lower than that reported in many previous studies in China. Furthermore, family health did not vary significantly between the prefrail and frail groups. This may be because the main survey method was conducted based on face-to-face reporting through an electronic questionnaire. If a respondent was mentally alert but not active enough to answer the questionnaire, the investigator conducted a one-on-one inquiry to complete the questionnaire on their behalf. Most participants were relatively healthy and less vulnerable than frail older adults. Therefore, the representativeness of the study sample might have been limited. In addition, frailty was measured using the FRAIL scale, which could not be validated by other frailty measures in this study, such as the frailty index proposed by Rockwood et al [34]. However, the findings of this study are important. Prefrail older adults are at a higher risk than robust older adults in experiencing frailty, adverse outcomes, and mortality [8]. In addition, older people with prefrailty are more likely to transition back to a state of robust health than those who are frail [35]. Therefore, health promotion for prefrail populations represents an important opportunity to prevent decline and dependence, enhance health, and reduce disability and the need for care. Therefore, the findings of this study provide important intervention strategies to manage frailty and have substantial implications for health policy and public health. Moreover, the results of the smooth curve and generalized additive model indicated a decreased risk of frailty and increased family health. Previous studies have indicated that both prefrail and frail older adults are highly susceptible to adverse health effects such as falls, disability, institutionalization, and hospitalization, all of which increase the social burden caused by population aging [36]. Therefore, improvement in family health is effective in preventing or delaying the onset of prefrailty or frailty.

Most studies associated with frailty management concentrated more on individuals, specifically exercise, nutritional intervention, multicomponent interventions, and individually tailored geriatric care models [37]. However, most individuals cannot exist away from their family, and family members are dependent on each other. Therefore, older adults may be far more affected by their family than by the outside world. As described by the family system theory, the family is a cohesive social unit that operates like a system with its own rules and responsibilities. Each family member has a profound impact on the choices of other members of the family, with the results passed on from generation to generation [38]. Therefore, frailty intervention in family units has great potential and may achieve twice the results with half the effort. This also indicates that promoting healthy aging through family health is effective and promising.

Previous studies have shown that the role of family is essential in predicting better health-related quality of life among older adults [39]. The presence of the family acts as the main source of social support in the acceptance of the aging process as well as helping to motivate participation in daily activities and improving self-esteem [40]. Psychosocial factors, such as individual preferences and values concerning food and sports or physical activities, affect health-related family interactions. Sociocultural factors indirectly affect family interactions via individuals who transfer these influences into their family lives [41]. In addition, the mediators in this study, including health literacy, smoking, sleep, and diet, have been shown to be effective in frailty interventions [42-44]. Family health, which covers a wider dimension of health-related social and internal support, interactions, and resources than family function or family climate [45], contributes to higher levels of health literacy and healthier behaviors and thus can decrease the risk of prefrailty and frailty. It is worth noting that in this study, the KHB composition revealed that the 4 mediators (health literacy, not smoking, sleep duration, and having breakfast every day) reduced the total effect of family health on frailty by only 26.72%, indicating that the proportion of mediated effects is not high. Family health is a comprehensive concept associated with individual's health in many aspects. The frailty status is also complex and is affected by many risk factors. Therefore, this study indicated potential paths for addressing the association between family health and frailty. Other mechanisms, especially the clinical, psychological, or physiological paths, should be explored in future research.

It is intriguing that the mediating effect of physical exercise on the association between family health and frailty was not significant when measured by weekly time ≥150 minutes but significant when measured by days of walking for at least 10 minutes per week. First, the negative association between physical exercise and frailty risk was significant in this study, which has also been supported by many previous studies [46-48]. Second, physical exercise, especially intensive and initiative exercises, was more popular among younger adults than older adults in China, and these exercise behaviors may be not easy to form among older adults, even when encouraged or urged by their family members. However, this study still indicated that some easy exercises, such as walking, which are also associated with decreased frailty risk [49], could be promoted by improving family health.

Family health resources, one of the dimensions of frailty, are negatively associated with health literacy and not smoking and also negatively associated with frailty. This may be because health sources mainly contribute to other behaviors, such as health-seeking behavior. Doctors' attempts to improve patients' health literacy and motivate them to quit smoking may be inadequate.

In addition, this study indicated that family health was associated with decreased frailty risk in older adults, which was validated in older adults aged ≥60 years, ≥65 years, and ≥75 years. However, the mediating roles of health literacy and health behaviors were not consistent. Most mediating effects in this study were not significant, particularly for those aged ≥75 years. A possible explanation may be found in the individual heterogeneity, complexity of disease conditions, and greater vulnerability to physical and psychological status for the relatively older populations [50]. Therefore, intervention in frailty status is more difficult in relatively older populations, and the associations between family health and frailty may be more significantly mediated by factors related to disease control and treatment. This also indicates that the mechanisms of the impact of family health on frailty are complex and require further exploration through prospective analyses within different age groups.

To improve family health, we suggest that it is necessary to broaden the previous perspective of family structure and family function, enhance the role of family health, make full use of and strengthen the internal relationships of the family, enhance the external support of the family, improve the family social network, and carry out health management of the older population with the family as the unit. Especially in China, with the acceleration of urbanization and the deepening of population aging, increasingly complex family structures and a broader range of social determinants of health have raised many challenges for health strategies. A family-centered healthy aging promotion strategy is feasible, but there is a need to consider the family's internal and social characteristics and to develop more scientific health promotion and management strategies suited to local conditions with the development of medical technology. Future research should further explore intervention strategies for family health and the causal effects of family health on individual health to promote healthy aging.

Limitations

This study has a few limitations. First, some participants were excluded from the data analysis because of missing data or logical errors. We compared our data with the population sample survey data from the National Bureau of Statistics (NBS) in China and found that the constituting ratio of the population aged ≥65 years among those aged ≥60 years (62.77%) was lower than that of NBS data (75.01%) and that the sex ratio (female or male) in this study was 0.99, which was slightly higher than that of NBS data (0.93). In addition, people who were confused, experienced mental health difficulties, or had cognitive impairments were not enrolled in the survey. However, mental or cognitive deficits are also risk factors for frailty. Therefore, the representativeness of the study findings should be interpreted with caution. Second, as mentioned earlier, the sample size of the frail population was small, making it difficult to conduct further analyses of the mechanism by which family health affects frailty. Third, all variables were self-reported, and some health-related variables may not be accurate because of recall bias. Fourth, this study had the inherent limitations of a cross-sectional study. Fifth, health behaviors that may be associated with frailty were not comprehensively or precisely determined in this survey. For example, the frequency of different physical exercises, nutritional intake, and self-adjustment of mood should be further explored. However, the association between family health and a decreased risk of prefrailty and frailty was consistently found across multiple models and subgroups, which offers important hints and inspiration for future studies. Longitudinal and prospective analyses of the causal effect of family health on frailty and its underlying mechanisms should be conducted in future studies.

Conclusions

In conclusion, family health can be an important intervention target that appears to be negatively linked to frailty in Chinese older adults. The mediation roles played by health literacy and health behaviors suggest that they can be effective in improving family health to promote healthier lifestyles as well as improving health literacy to delay, manage, and reverse frailty. The analysis related to the dimensions of family health may explain the mechanisms of the associations between family health, health literacy or behaviors, and frailty and guide future interventions. Strategies to intervene in frailty through family health in healthy aging and national public health strategies deserve more attention in the future.

Acknowledgments

The authors gratefully acknowledge the survey team for conducting the survey and for cleaning the data. The authors would also like to thank all the participants who participated in the study for their time and involvement. This work was supported by the China Scholarship Council (grant 202209095002) and the National Social Science Foundation (grant 16BGL150). The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the manuscript. The corresponding author had full access to all data in the study and had the final responsibility for the decision to submit for publication.

Data Availability

The data sets generated and analyzed during this study are not publicly available because the data still need to be used for other research but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Fried frailty phenotype and measurement.
[DOCX File , 12 KB - publichealth_v9i1e44486_app1.docx]

Multimedia Appendix 2
Health literacy and Assignment Criteria.
[DOCX File , 12 KB - publichealth_v9i1e44486_app2.docx]

Multimedia Appendix 3
The associations between dimensions of family health and frailty based on ordered logistic regression.
[DOCX File , 11 KB - publichealth_v9i1e44486_app3.docx]

Multimedia Appendix 4
Sensitivity analysis.
[DOCX File , 13 KB - publichealth_v9i1e44486_app4.docx]

References

1. Porter ME. A strategy for health care reform-toward a value-based system. N Engl J Med 2009 Jul 09;361(2):109-112. [doi: 10.1056/NEJMp0904131] [Medline: 19494209]
2. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, et al. The burden of disease in older people and implications for health policy and practice. Lancet 2015 Feb 07;385(9967):549-562. [doi: 10.1016/S0140-6736(14)61347-7] [Medline: 25468153]
3. Clark D, Kotronia E, Ramsay SE. Frailty, aging, and periodontal disease: basic biologic considerations. Periodontol 2000 2021 Oct;87(1):143-156 [FREE Full text] [doi: 10.1111/prd.12380] [Medline: 34463998]
4. Ribeiro AR, Howlett SE, Fernandes A. Frailty-a promising concept to evaluate disease vulnerability. Mech Ageing Dev 2020 Apr;187:111217. [doi: 10.1016/j.mad.2020.111217] [Medline: 32088282]
5. Brivio P, Paladini MS, Racagni G, Riva MA, Calabrese F, Molteni R. From healthy aging to frailty: in search of the underlying mechanisms. Curr Med Chem 2019;26(20):3685-3701 [FREE Full text] [doi: 10.2174/0929867326666190717152739] [Medline: 31333079]
6. Sezgin D, Liew A, O'Donovan MR, O'Caoimh R. Pre-frailty as a multi-dimensional construct: a systematic review of definitions in the scientific literature. Geriatr Nurs 2020 Mar;41(2):139-146. [doi: 10.1016/j.gerinurse.2019.08.004] [Medline: 31466806]
7. Ofori-Asenso R, Lee Chin K, Mazidi M, Zomer E, Ilomaki J, Ademi Z, et al. Natural regression of frailty among community-dwelling older adults: a systematic review and meta-analysis. Gerontologist 2020 May 15;60(4):e286-e298. [doi: 10.1093/geront/gnz064] [Medline: 31115434]

8. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
9. Frost R, Belk C, Jovicic A, Ricciardi F, Kharicha K, Gardner B, et al. Health promotion interventions for community-dwelling older people with mild or pre-frailty: a systematic review and meta-analysis. BMC Geriatr 2017 Jul 20;17(1):157 [FREE Full text] [doi: 10.1186/s12877-017-0547-8] [Medline: 28728570]
10. Barnes MD, Hanson CL, Novilla LB, Magnusson BM, Crandall AC, Bradford G. Family-centered health promotion: perspectives for engaging families and achieving better health outcomes. Inquiry 2020 Jan;57:46958020923537 [FREE Full text] [doi: 10.1177/0046958020923537] [Medline: 32500768]
11. Weiss-Laxer NS, Crandall A, Okano L, Riley AW. Building a foundation for family health measurement in national surveys: a modified Delphi expert process. Matern Child Health J 2020 Mar;24(3):259-266. [doi: 10.1007/s10995-019-02870-w] [Medline: 31912378]
12. Sudore RL, Yaffe K, Satterfield S, Harris TB, Mehta KM, Simonsick EM, et al. Limited literacy and mortality in the elderly: the health, aging, and body composition study. J Gen Intern Med 2006 Aug;21(8):806-812 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00539.x] [Medline: 16881938]
13. White S, Chen J, Atchison R. Relationship of preventive health practices and health literacy: a national study. Am J Health Behav 2008 May;32(3):227-242. [doi: 10.5555/ajhb.2008.32.3.227] [Medline: 18067463]
14. Huang CH, Lai YC, Lee YC, Teong XT, Kuzuya M, Kuo KM. Impact of health literacy on frailty among community-dwelling seniors. J Clin Med 2018 Nov 26;7(12):481 [FREE Full text] [doi: 10.3390/jcm7120481] [Medline: 30486284]
15. Shin HR, Choi EY, Kim SK, Lee HY, Kim YS. Health literacy and frailty in community-dwelling older adults: evidence from a nationwide cohort study in South Korea. Int J Environ Res Public Health 2021 Jul 27;18(15):7918 [FREE Full text] [doi: 10.3390/ijerph18157918] [Medline: 34360213]
16. Feinberg M, Hotez E, Roy K, Ledford CJ, Lewin AB, Perez-Brena N, et al. Family health development: a theoretical framework. Pediatrics 2022 May 01;149(Suppl 5):e2021053509I. [doi: 10.1542/peds.2021-053509I] [Medline: 35503316]
17. Lau RR, Quadrel MJ, Hartman KA. Development and change of young adults' preventive health beliefs and behavior: influence from parents and peers. J Health Soc Behav 1990 Sep;31(3):240-259. [Medline: 2133479]
18. da Silva VD, Tribess S, Meneguci J, Sasaki JE, Garcia-Meneguci CA, Carneiro JA, et al. Association between frailty and the combination of physical activity level and sedentary behavior in older adults. BMC Public Health 2019 Jun 07;19(1):709 [FREE Full text] [doi: 10.1186/s12889-019-7062-0] [Medline: 31174515]
19. Wang YJ, Kaierdebieke A, Fan SY, Zhang RF, Huang MJ, Li H, et al. Study protocol: a cross-sectional study on psychology and behavior investigation of Chinese residents, PBICR. Psychosom Med Res 2022;4(3):19 [FREE Full text] [doi: 10.53388/202219]
20. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. J Nutr Health Aging 2012 Jul;16(7):601-608 [FREE Full text] [doi: 10.1007/s12603-012-0084-2] [Medline: 22836700]
21. Si H, Jin Y, Qiao X, Tian X, Liu X, Wang C. Comparison of 6 frailty screening tools in diagnostic properties among Chinese community-dwelling older people. Geriatr Nurs 2021 Jan;42(1):276-282. [doi: 10.1016/j.gerinurse.2020.08.017] [Medline: 32948340]
22. Li Y, Zou Y, Wang S, Li J, Jing X, Yang M, et al. A pilot study of the FRAIL scale on predicting outcomes in Chinese elderly people with type 2 diabetes. J Am Med Dir Assoc 2015 Aug 01;16(8):714.e7-714.12. [doi: 10.1016/j.jamda.2015.05.019] [Medline: 26139596]
23. Ma Y, Hou L, Yang X, Huang Z, Yang X, Zhao N, et al. The association between frailty and severe disease among COVID-19 patients aged over 60 years in China: a prospective cohort study. BMC Med 2020 Sep 07;18(1):274 [FREE Full text] [doi: 10.1186/s12916-020-01761-0] [Medline: 32892742]
24. Chou YC, Tsou HH, Chan DD, Wen CJ, Lu FP, Lin KP, et al. Validation of clinical frailty scale in Chinese translation. BMC Geriatr 2022 Jul 20;22(1):604 [FREE Full text] [doi: 10.1186/s12877-022-03287-x] [Medline: 35858829]
25. Wang F, Wu Y, Sun X, Wang D, Ming WK, Sun X, et al. Reliability and validity of the Chinese version of a short form of the family health scale. BMC Prim Care 2022 May 06;23(1):108 [FREE Full text] [doi: 10.1186/s12875-022-01702-1] [Medline: 35524178]
26. Mok A, Khaw KT, Luben R, Wareham N, Brage S. Physical activity trajectories and mortality: population based cohort study. BMJ 2019 Jun 26;365:l2323 [FREE Full text] [doi: 10.1136/bmj.l2323] [Medline: 31243014]
27. Jaddoe VW, de Jonge LL, Hofman A, Franco OH, Steegers EA, Gaillard R. First trimester fetal growth restriction and cardiovascular risk factors in school age children: population based cohort study. BMJ 2014 Jan 23;348:g14 [FREE Full text] [doi: 10.1136/bmj.g14] [Medline: 24458585]
28. Cai X, Li N, Hu J, Wen W, Yao X, Zhu Q, et al. Nonlinear relationship between Chinese visceral adiposity index and new-onset myocardial infarction in patients with hypertension and obstructive sleep apnoea: insights from a cohort study. J Inflamm Res 2022 Feb 02;15:687-700 [FREE Full text] [doi: 10.2147/JIR.S351238] [Medline: 35140499]
29. What's new in Stata. Stata. URL: https://www.stata.com/ [accessed 2023-06-12]

30. Ling Y, Song Z, Yu Y, Jiang T. Dealing with an aging China-delaying retirement or the second-child policy? PLoS One 2021 Jan 07;16(1):e0242252 [FREE Full text] [doi: 10.1371/journal.pone.0242252] [Medline: 33411757]
31. Jia L, Du Y, Chu L, Zhang Z, Li F, Lyu D, COAST Group. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671 [FREE Full text] [doi: 10.1016/S2468-2667(20)30185-7] [Medline: 33271079]
32. Shwe PS, Ward SA, Thein PM, Junckerstorff R. Frailty, oral health and nutrition in geriatrics inpatients: a cross-sectional study. Gerodontology 2019 Sep;36(3):223-228. [doi: 10.1111/ger.12397] [Medline: 30861197]
33. Reddy VY, Möbius-Winkler S, Miller MA, Neuzil P, Schuler G, Wiebe J, et al. Left atrial appendage closure with the watchman device in patients with a contraindication for oral anticoagulation: the ASAP study (ASA Plavix feasibility study with watchman left atrial appendage closure technology). J Am Coll Cardiol 2013 Jun 25;61(25):2551-2556 [FREE Full text] [doi: 10.1016/j.jacc.2013.03.035] [Medline: 23583249]
34. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495 [FREE Full text] [doi: 10.1503/cmaj.050051] [Medline: 16129869]
35. Gill TM, Gahbauer EA, Allore HG, Han L. Transitions between frailty states among community-living older persons. Arch Intern Med 2006 Feb 27;166(4):418-423. [doi: 10.1001/archinte.166.4.418] [Medline: 16505261]
36. Hanlon P, Nicholl BI, Jani BD, Lee D, McQueenie R, Mair FS. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. Lancet Public Health 2018 Jul;3(7):e323-e332 [FREE Full text] [doi: 10.1016/S2468-2667(18)30091-4] [Medline: 29908859]
37. Walston J, Buta B, Xue QL. Frailty screening and interventions: considerations for clinical practice. Clin Geriatr Med 2018 Feb;34(1):25-38 [FREE Full text] [doi: 10.1016/j.cger.2017.09.004] [Medline: 29129215]
38. Rothbaum F, Rosen K, Ujiie T, Uchida N. Family systems theory, attachment theory, and culture. Fam Process 2002;41(3):328-350. [doi: 10.1111/j.1545-5300.2002.41305.x] [Medline: 12395563]
39. Abdala GA, Kimura M, Duarte YA, Lebrão ML, dos Santos B. [Religiousness and health-related quality of life of older adults]. Rev Saude Publica 2015;49:55 [FREE Full text] [doi: 10.1590/S0034-8910.2015049005416] [Medline: 26274870]
40. Andrade JM, Drumond Andrade FC, de Oliveira Duarte YA, Bof de Andrade F. Association between frailty and family functionality on health-related quality of life in older adults. Qual Life Res 2020 Jun;29(6):1665-1674. [doi: 10.1007/s11136-020-02433-5] [Medline: 32020563]
41. Wäsche H, Niermann C, Bezold J, Woll A. Family health climate: a qualitative exploration of everyday family life and health. BMC Public Health 2021 Jun 29;21(1):1261 [FREE Full text] [doi: 10.1186/s12889-021-11297-4] [Medline: 34187447]
42. Gimeno-Mallench L, Sanchez-Morate E, Parejo-Pedrajas S, Mas-Bargues C, Inglés M, Sanz-Ros J, et al. The relationship between diet and frailty in aging. Endocr Metab Immune Disord Drug Targets 2020;20(9):1373-1382. [doi: 10.2174/1871530320666200513083212] [Medline: 32400339]
43. Liu X, Wang C, Qiao X, Si H, Jin Y. Sleep quality, depression and frailty among Chinese community-dwelling older adults. Geriatr Nurs 2021 May;42(3):714-720. [doi: 10.1016/j.gerinurse.2021.02.020] [Medline: 33836251]
44. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
45. Güney E, Karataş Okyay E, Uçar T. Families' health behavior: validity and reliability of the Turkish version of the family health climate scale. Soc Work Public Health 2021 Aug 18;36(6):707-722. [doi: 10.1080/19371918.2021.1948484] [Medline: 34340641]
46. Angulo J, El Assar M, Álvarez-Bustos A, Rodríguez-Mañas L. Physical activity and exercise: strategies to manage frailty. Redox Biol 2020 Aug;35:101513 [FREE Full text] [doi: 10.1016/j.redox.2020.101513] [Medline: 32234291]
47. Ricci NA, Cunha AI. Physical exercise for frailty and cardiovascular diseases. Adv Exp Med Biol 2020;1216:115-129. [doi: 10.1007/978-3-030-33330-0_12] [Medline: 31894552]
48. Nagai K, Miyamato T, Okamae A, Tamaki A, Fujioka H, Wada Y, et al. Physical activity combined with resistance training reduces symptoms of frailty in older adults: a randomized controlled trial. Arch Gerontol Geriatr 2018 May;76:41-47. [doi: 10.1016/j.archger.2018.02.005] [Medline: 29455058]
49. Yu R, Wu WC, Leung J, Hu SC, Woo J. Frailty and its contributory factors in older adults: a comparison of two Asian regions (Hong Kong and Taiwan). Int J Environ Res Public Health 2017 Sep 21;14(10):1096 [FREE Full text] [doi: 10.3390/ijerph14101096] [Medline: 28934150]
50. Shen F, Chen J, Yang R, Yang J, Li H. Association between control status of blood pressure and frailty among middle-aged and older adults with hypertension in China: a longitudinal study. BMJ Open 2022 Mar 14;12(3):e056395 [FREE Full text] [doi: 10.1136/bmjopen-2021-056395] [Medline: 35288389]

Abbreviations

FRAIL: Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight
KHB: Karlson-Holm-Breen

NBS: National Bureau of Statistics
OR: odds ratio

Edited by A Mavragani; submitted 21.11.22; peer-reviewed by Y Zhou, V Chongthammakun; comments to author 24.01.23; revised version received 07.02.23; accepted 17.05.23; published 27.06.23.

Please cite as:
Li H, Wu Y, Bai Z, Xu X, Su D, Chen J, He R, Sun J
The Association Between Family Health and Frailty With the Mediation Role of Health Literacy and Health Behavior Among Older Adults in China: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e44486
URL: https://publichealth.jmir.org/2023/1/e44486
doi:10.2196/44486
PMID:37368463

Original Paper

Addressing Behavioral Barriers to COVID-19 Testing With Health Literacy–Sensitive eHealth Interventions: Results From 2 National Surveys and 2 Randomized Experiments

Carissa Bonner[1], MPH, PhD; Carys Batcup[1], MSc; Erin Cvejic[1], PhD; Julie Ayre[1], PhD; Kristen Pickles[1], PhD; Tessa Copp[1], PhD; Samuel Cornell[1], MSc; Brooke Nickel[1], PhD; Mustafa Dhahir[1], BPharm; Kirsten McCaffery[1], PhD

School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

Corresponding Author:
Carissa Bonner, MPH, PhD
School of Public Health
Faculty of Medicine and Health
University of Sydney
Edward Ford Building A27
Sydney, 2006
Australia
Phone: 61 2 9351 7125
Email: carissa.bonner@sydney.edu.au

Abstract

Background: Polymerase chain reaction (PCR) testing for COVID-19 was crucial in Australia's prevention strategy in the first 2 years of the pandemic, including required testing for symptoms, contact with cases, travel, and certain professions. However, several months into the pandemic, half of Australians were still not getting tested for respiratory symptoms, and little was known about the drivers of and barriers to COVID-19 PCR testing as a novel behavior at that time.

Objective: We aimed to identify and address COVID-19 testing barriers, and test the effectiveness of multiple eHealth interventions on knowledge for people with varying health literacy levels.

Methods: The intervention was developed in 4 phases. Phase 1 was a national survey conducted in June 2020 (n=1369), in which testing barriers were coded using the capability-opportunity-motivation-behavior framework. Phase 2 was a national survey conducted in November 2020 (n=2034) to estimate the prevalence of testing barriers and health literacy disparities. Phase 3 was a randomized experiment testing health literacy–sensitive written information for a wide range of barriers between February and March 2021 (n=1314), in which participants chose their top 3 barriers to testing to view a tailored intervention. Phase 4 was a randomized experiment testing 2 audio-visual interventions addressing common testing barriers for people with lower health literacy in November 2021, targeting young adults as a key group endorsing misinformation (n=1527).

Results: In phase 1, barriers were identified in all 3 categories: capability (eg, understanding which symptoms to test for), opportunity (eg, not being able to access a PCR test), and motivation (eg, not believing the symptoms are those of COVID-19). Phase 2 identified knowledge gaps for people with lower versus higher health literacy. Phase 3 found no differences between the intervention (health literacy–sensitive text for top 3 barriers) and control groups. Phase 4 showed that a fact-based animation or a TikTok-style video presenting the same facts in a humorous style increased knowledge about COVID-19 testing compared with government information. However, no differences were found for COVID-19 testing intentions.

Conclusions: This study identified a wide range of barriers to a novel testing behavior, PCR testing for COVID-19. These barriers were prevalent even in a health system where COVID-19 testing was free and widely available. We showed that key capability barriers, such as knowledge gaps, can be improved with simple videos targeting people with lower health literacy. Additional behavior change strategies are required to address motivational issues to support testing uptake. Future research will explore health literacy strategies in the current context of self-administered rapid antigen tests. The findings may inform planning for future COVID-19 variant outbreaks and new public health emergencies where novel testing behaviors are required.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12621000876897, https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=382318 ; Australian New Zealand Clinical Trials Registry ACTRN12620001355965, https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=380916&isReview=true

(JMIR Public Health Surveill 2023;9:e40441) doi:10.2196/40441

KEYWORDS

behavior change; health literacy; COVID-19; testing; infectious disease; public health

Introduction

The Role of Polymerase Chain Reaction Testing in COVID-19

The behavior of individuals has been crucial to the control of COVID-19, from self-isolating and testing to vaccination uptake [1]. A key preventive behavior in the early stages of the pandemic was polymerase chain reaction (PCR) testing for COVID-19 [2]. From 2020 to 2021, COVID-19 prevention strategies were often reliant on people getting a PCR test. This could be required when community members had been in contact with a positive case, had COVID-19 symptoms (eg, fever, cough, or sore throat), needed to travel from an outbreak area to another region, or worked in certain professions (eg, health workers). In Australia, community members were required to self-isolate at home until they returned a negative PCR test result, and this test-trace-isolate strategy was used to determine the need for short-term localized restrictions until linked clusters of cases were brought under control [3].

Testing Barriers

In early 2020, there was little research on COVID-19 testing behaviors, given the very new nature of this issue, so little was known about the barriers to testing or how to address this. Media reports at this time suggested different barriers existed across countries, which was confirmed in subsequent research. For example, countries such as Tanzania had major issues with opportunity barriers in terms of limited access to COVID-19 tests and fake testing kits [4]. Cost was a barrier in other countries, such as the United States, where the government and health insurers did not cover the testing [5], disproportionately affecting certain groups such as immigrant and noncitizen communities, who may also fear financial and legal repercussions from testing positive [6]. Testing was sometimes limited to certain criteria (eg, only if you have symptoms or regardless of exposure to COVID-19 cases) because of the lack of supply or staff resource issues [7]. There were also issues with delivering tests and transporting samples to remote areas [8]. Inadequate communication and low community knowledge about which symptoms require testing and the process to follow also impacted uptake [7].

The Australian Context

Australia was fortunate to have efficient and free testing widely available from the start of the pandemic, although this varied by location. PCR testing clinics were established nationally, including drive-through options to minimize contact with others and results sent by SMS text messages within a short period [9,10]. However, despite the high accessibility of PCR testing, flu tracking data suggested that many more people had respiratory symptoms than were getting tested [11]. At the time of the study, it was unclear why the uptake was so low, but we hypothesized that COVID-19 testing communication did not address the needs of varying health literacy levels in the community. Similar to many other countries [12], Australian national surveys showed that people with lower health literacy were less likely to know about COVID-19 symptoms and prevention measures [13] and were more likely to agree with misinformation about COVID-19 [14].

Theoretical Framework

According to the capability-opportunity-motivation-behavior (COM-B) model [15], health prevention behaviors can be conceptualized in terms of 3 main drivers: physical and psychological *capability* (eg, having the physical ability to drive to or walk up the stairs to access a testing center and knowing what to do if you have symptoms), physical and social *opportunity* (eg, the availability of testing centers in your area and social norms that make testing and self-isolation acceptable), and automatic and reflective *motivation* (eg, fear of a painful test and an explicit belief that it is important to get tested for symptoms) [2]. In early 2020, we used this framework as the basis for a new research program on the novel behavior of COVID-19 PCR testing.

Objective

This program aimed to develop and test eHealth interventions to overcome COVID-19 PCR testing barriers and address the varying health literacy needs of the community. The interventions were developed and evaluated in 4 phases from June 2020 to November 2021.

- Phase 1, in June 2020, aimed to identify the *range* of barriers to COVID-19 testing.
- Phase 2, in November 2020, aimed to estimate the *prevalence* of barriers to COVID-19 testing and to target interventions for the most important issues.
- Phase 3, from February to March 2021, aimed to test the efficacy of providing health literacy–sensitive written information (ie, adapted for people with lower health literacy) for all capability and motivation barriers identified in phase 2, where individuals could view information to make a plan for their top 3 barriers to testing.
- Phase 4, in November 2021, aimed to address design issues in phase 3 and test the efficacy of providing health literacy–sensitive audio-visual interventions (simple animation or TikTok-style video) for a smaller selection of common barriers for people with lower health literacy, identified in phase 2.

Methods

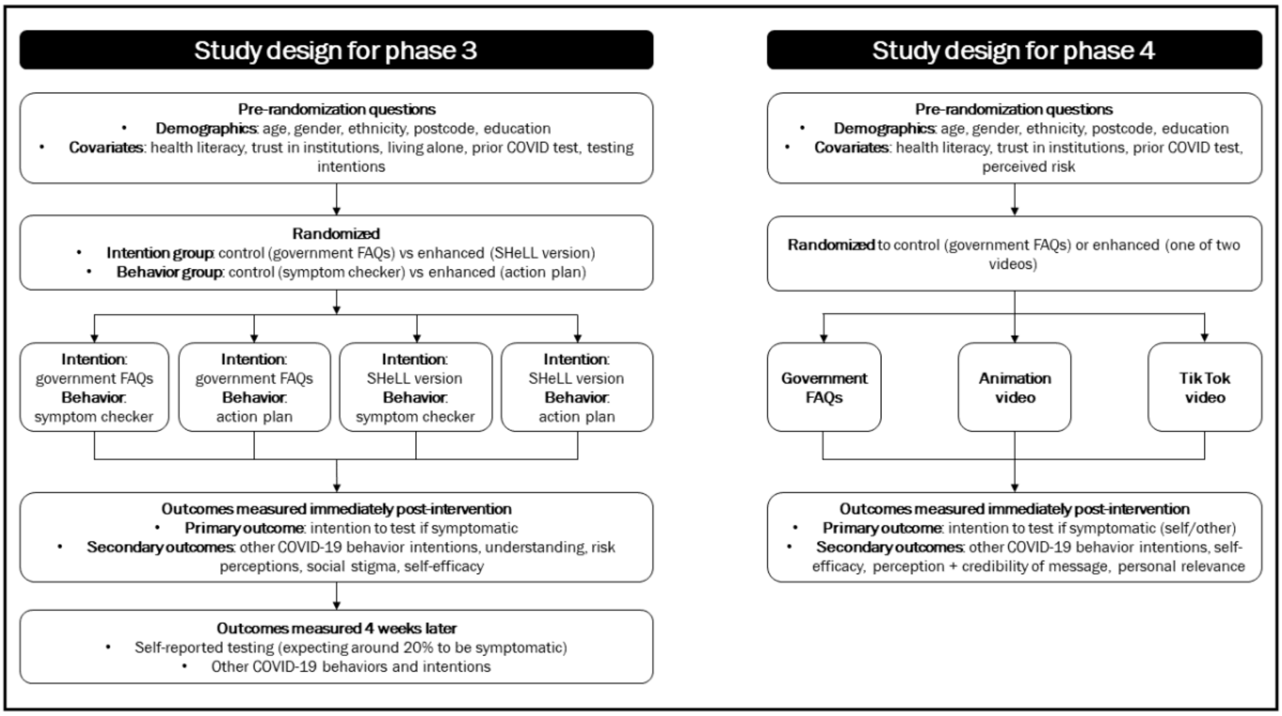
Ethics Approval

Ethics approval was obtained from the University of Sydney Human Research Ethics Committee (project number 2020/781), and the experiments were preregistered on the Australia New Zealand Trial Registry (ACTRN12621000876897 [16]; ACTRN12620001355965 [17]). All data were collected and stored anonymously, but participants could provide contact details to receive compensation via points for panel members and prize draws for gift vouchers if recruited via social media.

Study Design

The eHealth interventions were developed and tested in 4 phases. Phase 1 was a national survey conducted in June 2020, in which testing barriers were elicited and coded using the COM-B framework. Phase 2 was a national survey conducted in November 2020 to estimate the prevalence of testing barriers and health literacy disparities. Phase 3 was a randomized experiment testing health literacy–sensitive written information for a wide range of capability and motivation barriers from February to March 2021, in which participants chose their top 3 barriers to testing to view a tailored intervention. Phase 4 was a randomized experiment testing 2 audio-visual interventions addressing common testing barriers for people with lower health literacy in November 2021, targeting young adults as a key group endorsing misinformation (Figure 1).

Figure 1. Study design for phases 3 and 4. FAQ: frequently asked question; SHeLL: Sydney Health Literacy Lab.



Study Population

Phase 1 was open to any Australian adult who responded to the advertisements on social media. Phases 2 and 3 recruited a nationally representative sample of Australian adults based on age (equal groups above and below the age of 40 years), gender (equal groups for male and female; no quota for other categories), and education (equal groups for university degree and no degree). Phase 4 targeted younger Australian adults (aged <40 years) using the same market research panel.

Recruitment

Phase 1 used advertisements to target social media users, in which participants entered a prize draw to win a gift voucher. Phases 2 and 3 recruited a nationally representative sample via a market research panel company, where participants received points across multiple studies that they can redeem for various incentives including gift vouchers. Phase 4 involved both social media advertisements and panel recruitment. Different states in Australia were targeted at different times of recruitment so that participants were only recruited when there were very few or no cases in their state (when a test-trace-isolate strategy can be effective for containing spread).

Sample Size

For each phase, we recruited the following number of people: 1369 in phase 1; 20,349 in phase 2; 1314 in phase 3; and 1527 in phase 4.

Data Collection

All 4 phases involved piloting with a convenience sample from the target participant group before recruitment to identify and correct any grammar or navigation issues. The interventions were tested by a consumer representative who provided feedback to refine the study materials before formal data collection.

In phase 1, social media users were asked to participate in a series of 10-minute surveys, with COVID-19 testing questions included in June 2020. The New South Wales State Health Department provided a short list of testing barriers to be included along with an open response option for other perceived or experienced barriers. Phase 2 recruited a nationally representative sample, in which eligible panel members were invited to participate through the company's usual channels. After providing informed consent, participants completed a 10-minute survey, which included selecting COVID-19 testing barriers from a list and ranking them in order of importance.

A similar procedure was used for phases 3 and 4, with recruitment through a market research panel company, but participants were randomized to view different versions of COVID-19 testing information. In phase 3, they were randomized to view written government information or the intervention and completed the outcome questions. Those randomized to the intervention selected 3 relevant barriers, viewed health literacy–sensitive information about those issues, and created a plan for 1 chosen barrier. They were asked to create an action plan to help them overcome their barriers and received weekly reminders with a screenshot of their action plan by email. After 4 weeks, all participants received a 5-minute survey. In phase 4, social media users aged 18-39 years on Facebook and Instagram were targeted with advertisements, and further participants were recruited via the panel company. The following advertisement text was used for social media: “We want to hear from you! Complete a short survey about COVID-19 and be in with the chance to win a $20 gift card.” All participants answered a 10-minute survey. Those randomized to the intervention groups viewed a short audio-visual intervention, whereas those in the control group viewed standard written government information. Access to the outcome questions to complete the survey was enabled after 73 seconds for the animation and 65 seconds for the TikTok (the lengths of the audio-visual intervention) to increase the chance that participants viewed the intervention.

Outcomes

Survey questions for phases 1 and 2 are provided in Multimedia Appendix 1. In each phase, we measured variables shown to be associated with differences in the understanding of COVID-19 symptoms and prevention measures in our previous research [13,18]: age, gender, language, health literacy, trust, living alone, and prior COVID test. For the trials, our primary outcome was the intention to undergo testing for COVID-19 if symptomatic (measured in a broad way for phase 3 and a more specific way for phase 4 to increase sensitivity). The secondary outcomes included intentions about other prevention behaviors (self-isolation if symptomatic, social distancing of 1.5 m, washing hands regularly, and wearing masks in crowded indoor areas), understanding of messaging, risk perceptions, social stigma, and self-efficacy (ie, confidence in overcoming perceived barriers to testing). In phase 3 only, self-reported prevention behavior and intentions were assessed after 1 month, with our prior survey data suggesting that >20% of participants would experience symptoms over that time (Table 1).

Table 1. Primary outcome measures for phases 3 and 4.

Outcome and items	Response options	Phase 3	Phase 4
Intention to get tested for COVID-19			
Over the next 4 weeks I plan to get tested if I have COVID-19 symptoms (cough, sore throat, fever)	1=strongly disagree to 7=strongly agree	✓	✓
Imagine you woke up with a sore throat tomorrow. Would you get tested straight away?	1=extremely unlikely to 5=extremely likely		✓
Most people my age would get tested after seeing this information, if they develop symptoms	1=strongly disagree to 7=strongly agree		✓
Intention to engage in other preventive behaviors			
Over the next 4 weeks I plan to...Stay home if I have COVID-19 symptoms (cough, sore throat, fever)/Stay 1.5m away from others that I don't live with where I can/Wash my hands or use sanitiser to protect me and others from COVID-19/Wear a mask in crowded indoor areas	1=strongly disagree to 7=strongly agree	✓	✓
Knowledge			
Can you name 3 signs or symptoms that are associated with COVID-19?	Written answer	✓	
When someone has signs that they might have COVID-19 (e.g. a cough or a sore throat), they should...	Wait until their symptoms are bad enough, then get tested/get tested straight away and self-isolate at home until they get their test results/get tested straight away and carry on as normal until they get their test results/self-isolate at home until their symptoms go away	✓	
If someone has a sore throat and wants to get tested for COVID-19, they should go... (choose all that apply)	To the emergency department/to their GP[a]/to a COVID-19 testing center	✓	
If someone gets tested for COVID-19, and they need groceries while waiting for their result, they should ... (choose all that apply)	Ask someone to get the groceries for them/wear a mask at the shops/go to the shops quickly to get essentials/order online	✓	
What are the 6 main COVID-19 symptoms you should get tested for? Select from the list	Sore throat/loss of taste or smell/digestive issues/cough/muscle aches/vomiting/fever/conjunctivitis/runny nose/shortness of breath or difficulty breathing/diarrhea		✓
When do you need to get tested?	Any mild/slight symptoms/moderate/uncomfortable symptoms/symptoms are severe/disrupt your plans		✓
When do you need to get tested?	Any length of time with symptoms/symptoms lasting 2 days/symptoms lasting 3 days		✓
When do you need to get tested?	One or more symptoms in a single day/2 or more symptoms in a single day/3 or more symptoms in a single day		✓
If you have symptoms of COVID-19, you should get tested when... Select all that apply	There are hotel quarantine cases in your state/there are local cases in your community/there are no local cases in your community		✓
Should you get tested if you have unusual or new cold-like symptoms that you think are due to: Cold weather/a cold/flu/hayfever/allergies	1=yes definitely to 5=no definitely not		✓

[a] GP: general practitioner.

Analyses

For phases 1 and 2, participant characteristics and survey question responses were reported descriptively using a content analysis approach for open responses and chi-square tests to compare responses across health literacy levels. In total, 2 researchers mapped the text from open responses to the components of the COM-B model, with discussion to resolve discrepancies. For phases 3 and 4, analyses were conducted using planned contrasts between the intervention arms and control arm, implemented in the regression models. Continuous outcomes were analyzed using linear regression to estimate marginal mean differences, dichotomous outcomes were analyzed using generalized linear models with a modified Poisson approach (log link and robust SEs) to estimate relative risks, and count variables were analyzed using Poisson regression to estimate relative risks. In phase 3, analyses were controlled for age, gender, language, health literacy, trust, living alone, and previous COVID-19 testing. In phase 4, positive baseline intention, age, gender, language, health literacy, trust, and perceived COVID-19 risk in Australia were controlled for. Interactions between health literacy and randomized conditions were also explored.

Materials

Images and text for the intervention in phase 3 are provided in Multimedia Appendix 2. The health literacy–sensitive version of the text was developed by applying health literacy guidelines [18], using a web-based tool that provides objective feedback on the complexity of health information (eg, grade reading score, passive voice, and medical jargon) [19], and incorporating consumer feedback. We used the Sydney Health Literacy Lab Editor developed by our team [19] to meet the recommended grade 8 school level by simplifying complex words, sentences, and grammar. Participants randomized to the intervention were asked to choose their top 3 barriers to testing and then selected 1 barrier for the health literacy–sensitive action plan. This was adapted from our previous studies on health-related lifestyle changes to address intention-behavior gaps [20,21]. The web-based action plan used an *if-then* format (eg, “If I don’t want to get tested because there aren’t many cases in my area, then I will remind myself that every new outbreak of COVID-19 starts with one new case”), a format that has shown to improve various behavioral outcomes including smoking cessation, physical activity, and healthy eating [22-28]. Health literacy principles were applied to the if-then plan (eg, simple language, images, and breaking down tasks into smaller steps), as previous research has shown that this can improve the effectiveness of if-then plans for people with low health literacy [20,29]. The *if* options reflect the top 10 barriers for people with lower health literacy identified in phase 2:

- I would prefer to isolate instead.
- I’m not sure this symptom is one that needs testing.
- I have symptoms of COVID-19 but I don’t think they are bad enough.
- I have symptoms of COVID-19 but I think it’s a cold or hay fever.
- I’m worried the test is painful.
- I’m worried about spreading my illness on the way to the testing centre.
- There aren’t many cases in my area.
- I’m worried I will catch COVID-19 when I get tested or on the way to the testing centre.
- I’m not sure what to do.
- I’d like my doctor’s advice first.

Participants could then select a solution (*then* option). Solutions were generated in collaboration with our consumer representative.

Images and text for the phase 4 intervention are provided in Multimedia Appendix 3. There were 2 audio-visual interventions: one was an animation and the other was a TikTok-style video, which was more humorous. Both covered the following barriers to COVID-19 testing, which were identified as the most prevalent knowledge issues for people with lower health literacy in phase 2:

- I know what symptoms I have and don’t believe they are COVID-19 ones e.g. hay fever/normal cold.
- I’m not sure my symptoms are bad enough.
- It is unlikely I have COVID-19 because there aren’t many cases in my area.
- I’m not sure this symptom needs testing.

Results

Participant characteristics for each phase are provided in Multimedia Appendix 4.

Phase 1: Identification of COVID-19 Testing Barriers

Most people (1151/1369, 84.07%) agreed that they would get tested if they had COVID-19 symptoms (cough, fever, and sore throat), with 49.23% (674/1369) of people strongly agreeing. For self-isolation, 95.98% (1314/1369) of participants agreed to some extent that they would stay home if they had symptoms and 69.47% (951/1369) of participants strongly agreed. Most participants (982/1369, 71.73%) said they would get tested “no matter what.” The most common barriers selected from the list provided (Multimedia Appendix 1) were that testing is painful (153/1369, 11.18%), not knowing how to get tested (98/1369, 7.16%), and worry about getting infected at the testing center (81/1369, 5.92%). All other barriers were <3% (forgetting, worried what others think, too hard or expensive, doesn’t work or don’t trust results, and no one else getting tested). Many participants (136/1369, 9.93%) indicated other reasons, with 136 open responses that included many additional barriers to testing than those provided in the survey question. Table 2 maps all the barriers identified in open survey responses to the COM-B drivers of behavior (capability, opportunity, and motivation).

Table 2. Phase 1 COVID-19 testing behavior barriers mapped to the capability-opportunity-motivation-behavior (COM-B) model.

COM-B barriers	Example quotes from “other” barriers (10% of the sample)
Physical capability	
My disability means I can’t get a test	“Fear of injury from getting tested—I have a deviated septum and narrow sinuses.”
I need a ramp/disability provisions for the testing centre	“Mobility and suppressed immune system make travel difficult.” “Testing facilities can’t accommodate my disability so it’s better for me just to stay home.”
I physically can’t access testing centre	“I’m homebound due to severe disability so I know arranging testing will be super hard.”
Psychological capability	
I have been getting conflicting information or being told not to get tested even with symptoms	“When I first got sore throat, headache, aches, cough, and partner very lethargic I got tested, and had the impression from testers that it wasn’t necessary with only those symptoms.”
I’m not sure how to get tested	“I dont know how to drive so I dont know how to get to places that test.”
I’m not sure my symptoms are bad enough	“Believing that I don’t meet the threshold eg when is a cough a cough or a runny nose with spicy food a runny nose? Otherwise I would get tested.”
I’m not sure this symptom is one that needs testing	“I am not sure whether I would be eligible to be tested.”
Physical opportunity	
The testing centres are hard to access—too far away from me	“Hard to get to where I can get tested.”
The opening hours of the testing centres don’t suit me	“Being able to get to the testing clinic within opening hours when caring for children.”
I don’t want to take public transport	“Public transport required to get to testing centre is inappropriate.”
I will need to take time off work	“hard to find the time with increased workload.”
I don’t have enough time to get tested	“Just tooooo busy—work during COVID.”
I don’t have childcare	“Depends on if I can get childcare and can find the time.”
If I get tested, it will impact me financially	“My husbands test took 10 days to get results. Which meant he could not work. He is self employed which means he also did not get paid and lost 8 days of income.”
Social opportunity	
I am worried what others will think of me having a test/being positive	“I will certainly feel worried about what others think of me, especially people who lives in the same household.”
Reflective motivation	
I had a bad experience when I got tested previously	“we had a bad experience getting my son tested...I can’t afford to be off work for 5 days just to chase down a negative result.”
I think the process of testing or pre-testing requirements is too much hassle/pointless	“There is a very tiny chance it will actually be COVID-19 so waste of time.”
I think the test is painful	“It hurt last time I got tested.”
I am worried about spreading my illness	“Have to ask someone to take me to a clinic and don’t want to get them sick.”
I’m worried I will catch covid myself whilst getting tested or en route	“I’m immunocompromised so don’t want to be around others in case I get sick.”
I don’t want to hear that I’m positive	“Will only get tested if I have a fever. Also, ‘too scared’ to know the results!”
I don’t think the testing works or results are reliable enough	“the accuracy of tests does concern me a bit.”
I don’t want to self-isolate after the test	“requirement to quarantine between test & getting results.”
I would prefer to self-isolate instead (or do another type of test)	“Don’t think it’s necessary to test, just isolate.”
Symptoms due to something else	“The symptoms are the same as my normal June sniffles I wouldn’t bother, unless I get other unusual symptoms.”

COM-B barriers	Example quotes from “other” barriers (10% of the sample)
I have already got tested for these symptoms and was negative	“I was tested last week. If this residual cough remains I am unlikely to get retested. If I get new symptoms I would get retested.”
I have the symptoms but will wait for them to get worse first/threshold	“I’ll get tested if the symptoms last for more than 24 hours” “if its just something like runny nose but no other symptoms show up—No but if I have more than 1 then I’ll get tested.”
I don’t believe it can be covid as it’s so rare	“No community transmission, highly unlikely to be COVID-19.”
I will only get a test if advised to by my GP[a]	“Get tested if my dr refers me to be tested.”
I think getting tested may result in problems with my visa or job	“Might be excluded from starting a new job.”
Misinformation or myths/different views on managing COVID	“I believe in herd immunity.”
Automatic motivation	
I am scared of the test	“Afraid of the uncomfortable test.”

[a] GP: general practitioner.

Phase 2: Prevalence of Barriers to Testing for COVID-19

The aim of phase 2 was to estimate the prevalence of COVID-19 testing barriers in a nationally representative sample and explore health literacy disparities to identify priority issues for intervention.

Table 3 presents the prevalence and importance of barriers among the participants who selected any barrier (941/2034, 46.26%). The top barriers were related to motivation: “I know what symptoms I have and don’t believe they are COVID-19 ones e.g. hayfever/normal cold” (selected by 562/2034, 27.63%, with 200/2034, 9.83% ranking it most important) and “It is unlikely I have COVID-19 because there aren’t many cases in my area” (366/2034, 17.99% selected, with 104/2034, 5.11% ranking it most important). Capability issues were also common: “I’m not sure my symptoms are bad enough” (387/2034, 19.03% selected, with 109/2034, 5.36% ranking it most important) and “I’m not sure this symptom is one that needs testing” (306/2034, 15.04% selected, with 66/2034, 3.24% ranking it most important). Social opportunity issues were uncommon: 5.75% (117/2034) of participants were worried about what others might think if they got a positive COVID-19 test result, and 3.69% (75/2034) of participants worried about what others would think if they got tested at all. Physical opportunity issues included disabilities, access (especially the distance to travel to a testing center), and time restrictions.

When we compared participants with low health literacy versus high health literacy, we found similar results for the top 10 barriers, covering reflective motivation and psychological capability. However, there were significant differences between the 2 groups. People with low health literacy were more likely to select certain capability issues (I’m not sure how to get tested, $P<.001$; I’m not sure if this symptom needs testing, $P=.01$) and motivation issues (I would prefer to self-isolate instead, $P<.001$; I think the test is painful, $P=.002$; I’m worried about spreading my illness, $P<.001$, I’m worried I will catch COVID-19, $P=.001$; I don’t want to hear that I’m positive, $P<.001$; and I don’t trust people who are asking me to take a test, $P=.02$), compared with the people with higher health literacy (Table 3).

Table 3. Phase 2 prevalence of barriers in nationally representative sample and health literacy disparities (November 2020; n=2034).

Barrier[a]	Selections (n=2034), n (%)	People to rank it first (n=2034), n (%)	Low health literacy (n=334), n (%)	High health literacy (n=1681), n (%)	*P* value for health literacy
I would prefer to self-isolate instead	341 (16.8)	73 (3.6)	71 (21.3)	194 (11.5)	<.001
I know what symptoms I have and don’t believe they are COVID-19 ones e.g. hay fever/normal cold	562 (27.6)	200 (9.8)	61 (18.3)	241 (14.3)	.07
I think the test is painful	276 (13.6)	67 (3.3)	49 (14.7)	154 (9.2)	.002
I’m not sure my symptoms are bad enough	387 (19)	109 (5.4)	45 (13.5)	182 (10.8)	.16
I’m not sure this symptom is one that needs testing	306 (15)	66 (3.2)	41 (12.3)	134 (8)	.01
I am worried about spreading my illness on the way to the testing centre	207 (10.2)	39 (1.9)	39 (11.7)	105 (6.2)	<.001
It is unlikely I have COVID-19 because there aren’t many cases in my area	366 (18)	104 (5.1)	37 (11.1)	159 (9.5)	.36
I’m worried I will catch COVID-19 myself whilst getting tested or on the way to the testing centre	219 (10.8)	34 (1.7)	37 (11.1)	104 (6.2)	.001
I’m not sure how to get tested	123 (6)	27 (1.3)	36 (10.8)	59 (3.5)	<.001
I will only get tested if my GP[b] tells me I should	187 (9.2)	63 (3.1)	30 (9.0)	126 (7.5)	.35
I don’t want to take public transport	143 (7)	31 (1.5)	24 (7.2)	80 (4.8)	.07
I have already got tested for these symptoms and was negative	150 (7.4)	34 (1.7)	24 (7.2)	71 (4.2)	.02
I think the process of testing is too much effort	148 (7.3)	15 (0.7)	24 (7.2)	78 (4.6)	.05
I don’t want to hear that I’m positive for COVID-19	85 (4.2)	5 (0.2)	21 (6.3)	42 (2.5)	<.001
I have been told not to get tested even if I have symptoms	113 (5.6)	33 (1.6)	18 (5.4)	65 (3.9)	.20
I don’t trust people who are asking me to take a test	72 (3.5)	16 (0.8)	17 (5.1)	45 (2.7)	.02
I am worried what others will think of me being positive for COVID-19	117 (5.8)	29 (1.4)	16 (4.8)	66 (3.9)	.47
I don’t want to take up resources for testing so that others can’t get tested	88 (4.3)	8 (0.4)	15 (4.5)	41 (2.4)	.04
I am worried what others will think of me having a test	75 (3.7)	6 (0.3)	14 (4.2)	43 (2.6)	.10
I don’t think the testing works or results are reliable enough	66 (3.2)	15 (0.7)	13 (3.9)	44 (2.6)	.20
I had a bad experience when I got tested before	89 (4.4)	10 (0.5)	12 (3.6)	49 (2.9)	.51
I don’t understand why I need to get tested	48 (2.4)	12 (0.6)	9 (2.7)	34 (2)	.44
I don’t want to self-isolate after the test	118 (5.8)	16 (0.8)	9 (2.7)	46 (2.7)	.97
I have the symptoms but will wait for them to get worse first	88 (4.3)	9 (0.4)	6 (1.8)	45 (2.7)	.35
I think getting tested may result in problems with my visa or with official bodies	33 (1.6)	5 (0.2)	6 (1.8)	21 (1.2)	.43

[a] Ordered from highest to lowest frequency in the low health literacy group (top 10 used in phase 3); participants could select more than 1 barrier.

[b] GP: general practitioner.

Phase 3: Randomized Experiment to Test the Effect of Health Literacy–Sensitive Written Information About COVID-19 Testing Barriers

The aim of phase 3 was to address the top 10 testing barriers for people with lower health literacy. Opportunity barriers, such as physical inaccessibility, could not be addressed by a communication intervention, so this phase focused on capability and motivation issues. For the intervention, participants chose 3 relevant barriers from the 10 provided and viewed health literacy–sensitive versions of written government information about COVID-19 testing (eg, reading level reduced to grade 8 by replacing complex words and sentences with simpler options). Intervention participants then selected 1 barrier to make an "if-then" action plan for how they would get around this issue if they needed to get a COVID-19 test for symptoms. Immediately after the intervention, no differences were found for intervention versus control (written government frequently asked questions about COVID-19 testing) on intentions, knowledge, or any other psychological outcomes. After 4 weeks, 57.71% (790/1369) of respondents completed follow-up measures, but there were no differences between intervention and control (government tool that tailored information to the local context; eg, state-specific COVID-19 testing requirements). No significant differences were found in health literacy levels when included as an interaction term (Table 4).

Table 4. Phase 3 experiment results comparing intervention and control groups (February-March 2021; n=1314).

Outcomes[a]	Control (n=668), n (%)	Intervention (n=645), n (%)	Effect estimate (95% CI)	*P* value
Testing intentions (self—if symptomatic next 4 weeks; 7=strongest intentions)	5.8 (1.6)	5.8 (1.6)	0.07 (−0.09 to 0.22)	.40
Self-isolation intentions (7=strongest intentions)	6.2 (1.2)	6.1 (1.3)	−0.03 (−0.16 to 0.10)	.63
Knowledge total score (count/6)[b]	5.5 (0.9)	5.4 (1.0)	0.98 (0.94 to 1.03)	.42
Perceived risk (5=highest perceived risk)	2.8 (0.8)	2.8 (0.8)	−0.01 (−0.10 to 0.08)	.79
Threat to Australia (10=very serious threat)	6.7 (2.4)	6.7 (2.4)	0.04 (−0.21 to 0.29)	.77
Self-efficacy (7=highest self-efficacy)	5.8 (1.0)	5.7 (1.1)	0.00 (−0.11 to 0.11)	.96
COVID-19 stigma (7=highest stigma)	3.4 (1.2)	3.6 (1.3)	0.08 (−0.05 to 0.21)	.25
Distancing intentions (7=strongest intentions)	6.3 (1.1)	6.2 (1.2)	−0.03 (−0.15 to 0.08)	.56
Hand washing intentions (7=strongest intentions)	6.4 (1.0)	6.4 (1.1)	−0.04 (−0.14 to 0.07)	.49
Mask wearing intentions (7=strongest intentions)	6.0 (1.4)	6.0 (1.4)	0.02 (−0.13 to 0.16)	.81
Follow-up: Self-reported testing behavior[c]	39 (9.4)	39 (10.5)	1.00 (0.67 to 1.51)	.99
Follow-up: Testing intentions (7=strongest intentions)	5.8 (1.5)	5.7 (1.6)	−0.10 (−0.30 to 0.11)	.36
Follow-up: Testing intentions (controlling for baseline intentions) (7=strongest intentions)	—[d]	—	−0.11 (−0.30 to 0.08)	.26
Follow-up: Self-isolation intentions (7=strongest intentions)	6.2 (1.2)	6.1 (1.2)	−0.10 (−0.25 to 0.06)	.24
Follow-up: Self-isolation intentions (controlling for baseline intentions; 7=strongest intentions)	—	—	−0.05 (−0.20 to 0.09)	.47

[a] Continuous outcomes were analyzed using linear regression to estimate marginal mean differences.

[b] Count variables were analyzed using Poisson regression to estimate relative risks.

[c] Dichotomous outcomes were analyzed using generalized linear models with a modified Poisson approach (log link and robust SEs) to estimate relative risks.

[d] Not available.

Phase 4: Randomized Experiment to Test the Effect of Health Literacy–Sensitive Audio-Visual Interventions About COVID-19 Testing Barriers in Adults With Lower Health Literacy

The aim of phase 4 was to develop a more targeted communication intervention with further refined testing of outcome measures. We selected 4 capability (knowledge) barriers from the most prevalent issues for people with lower health literacy in phase 2 and developed 2 audio-visual intervention scripts to address these: a simple animation in the style of Australian government advertisements and a TikTok-style video developed from the same information by a pharmacist with a large social media following for COVID-19 information videos. Immediately postintervention, we found that the animation and TikTok versions were more effective than the written government information for increasing knowledge about COVID-19 testing but not testing intentions or other psychological outcomes (based on a multiple-comparison adjusted α level of .025). No significant differences were found in health literacy levels when included as an interaction term (Table 5).

Table 5. Phase 4 experiment results comparing intervention and health literacy groups (November 2021; n=1527).

Outcome[a]	Government text (n=509)	Animation (n=514)	TikTok (n=504), n (%)	Animation vs government text		TikTok vs government text		Condition × health literacy interaction (*P* value)
				Effect estimate (95% CI)	*P* value	Effect estimate (95% CI)	*P* value	
Positive testing intentions (self—if symptoms tomorrow)[b], n (%)	246 (48.3)	294 (57.2)	277 (55)	1.08 (0.99 to 1.17)	.08	1.10 (1.00 to 1.20)	.04	.70
Positive testing intentions (other—if symptoms tomorrow)[b], n (%)	270 (53)	264 (51.4)	260 (51.6)	0.93 (0.83 to 1.03)	.17	0.95 (0.85 to 1.07)	.40	.31
Positive testing intentions (self—if symptoms next 4 weeks)[b], n (%)	358 (70.3)	373 (72.6)	367 (72.8)	1.00 (0.93 to 1.07)	.97	1.02 (0.95 to 1.09)	.56	.24
Knowledge total score (count/6), median (IQR)[c]	3.0 (2.0 to 4.0)	4.0 (3.0 to 5.0)	4.0 (3.0 to 5.0)	1.33 (1.24 to 1.42)	<.001	1.25 (1.17 to 1.34)	<.001	.84
Self-efficacy, mean (SD)	5.6 (1.2)	5.7 (1.1)	5.7 (1.1)	0.07 (−0.05 to 0.20)	.26	0.13 (0.01 to 0.26)	.04	.13
Perceived effectiveness, mean (SD)	3.7 (0.8)	3.6 (1.0)	3.6 (0.9)	−0.09 (−0.17 to −0.01)	.04	−0.09 (−0.17 to 0.00)	.04	.84
Message credibility, mean (SD)	5.3 (1.4)	5.3 (1.5)	5.3 (1.4)	−0.04 (−0.16 to 0.08)	.53	−0.06 (−0.18 to 0.06)	.30	.20
Personal relevance, mean (SD)	4.2 (1.4)	4.2 (1.5)	4.1 (1.5)	−0.09 (−0.24 to 0.06)	.22	−0.13 (−0.28 to 0.02)	.09	.10
Other behavioral intentions (averaged: self-isolation, distancing, hand washing, and mask wearing), mean (SD)	5.7 (1.3)	5.8 (1.3)	5.8 (1.3)	0.08 (−0.05, 0.21)	.23	0.10 (−0.02, 0.23)	.12	.17

[a] Continuous outcomes were analyzed using linear regression to estimate marginal mean differences.

[b] Dichotomous outcomes were analyzed using generalized linear models with a modified Poisson approach (log link and robust SEs) to estimate relative risks.

[c] Count variables were analyzed using Poisson regression to estimate relative risks.

Discussion

This study aimed to identify and address barriers to COVID-19 testing and test the effectiveness of multiple eHealth interventions on knowledge for people with varying health literacy levels.

Principal Findings

Phase 1 identified a wide range of barriers to COVID-19 testing that had not been previously described in the COVID-19 literature. These covered all 3 behavioral drivers in the COM-B model. Phase 2 found that motivation and capability barriers were far more prevalent than opportunity barriers in Australia at the time of this study. Many barriers were reported as more prevalent among people with lower health literacy. Phases 3 and 4 tested different ways to address capability and motivation barriers. Phase 3 found no differences between standard government text about COVID-19 testing and a tailored text intervention that addressed many different barriers using health literacy design principles. Phase 4 found that audio-visual interventions to address key knowledge barriers are more effective than written government information for improving knowledge, but this was not enough to shift COVID-19 testing intentions in adjusted analyses.

Comparison With Prior Work

Since our Australian research was conducted, new papers have been published on the issue of COVID-19 testing, including the move from PCR to rapid antigen tests (RATs). Survey, interview, and media analysis studies have identified similar capability, opportunity, and motivation barriers in other countries, including the United States, the United Kingdom, the United Arab Emirates, and Jordan [30-34], as well as in

Australia (eg, in a recent qualitative study [35]). Sample populations have included the general public, parents, university students and staff, people experiencing homelessness, and specific cultural groups [36-44]. This study is unique in terms of mapping barriers to the COM-B theoretical framework, estimating the prevalence of key barriers to testing in people with diverse health literacy, and testing communication interventions to address capability and motivation barriers. We have shown that it is possible to increase COVID-19 testing knowledge when key barriers are explicitly addressed using simple audio-visual intervention formats that are cheap and quick to produce. However, it is important to acknowledge that a mass communication strategy will not address all barriers.

Interpretation and Implications

The results of this study provide new insights into identifying and addressing behavioral barriers to COVID-19 testing, which is central to understanding and controlling COVID-19 and future pandemics. The large range of barriers identified in this study reflects the fact that COVID-19 testing was a complex new behavior at the time of the research, requiring a multifaceted approach to improve uptake depending on key barriers in different communities. Although the Australian PCR testing system was free and widely accessible for the first 2 years of the pandemic, the shift to using RATs introduced new opportunity barriers that were not an issue at the time of these studies. For example, it was very difficult to locate RATs during a period of low supply and high demand in January 2022, and there were issues with price gouging that made this unaffordable for many Australians [45]. Government regulation and funded RATs were subsequently introduced to limited groups, such as schoolchildren and pensioners [46,47].

Future Research

The capability and motivation issues identified in this study apply to PCR testing, but there are likely to be additional barriers to RATs, which were not approved in Australia at the time of the study. Different barriers encountered for RATs are being investigated in subsequent research, including individuals' ability to understand instructions, perform self-testing, and interpret the results correctly (eg, see trial ACTRN12622001517763). Concerns have been raised about the misinterpretation of negative results from RATs, which have a high error rate if the test is not used within the recommended period after exposure to a COVID-19 case [48]. Another avenue for further work is to partner with the media to avoid the identification and stigmatization of individuals with positive test results in future disease outbreaks [49]. Media reports and anecdotal data from frontline health professionals may be a useful way to quickly identify emerging local issues that could inform the measurement of testing barriers to make them more relevant to local communities.

Strengths and Limitations

This research began as an unfunded and rapidly developing response to the COVID-19 pandemic. By addressing methodological issues and building on the findings in each phase, we were able to better target the final intervention and show the value of audio-visual formats in addressing common knowledge barriers among people with varying health literacy needs.

Phase 1 identified the range of barriers to COVID-19 testing in Australia for the first time, but the prevalence of the most important barriers could not be ascertained from these findings because of the reliance on open responses and a nonrepresentative sample. The next phase addressed these methodological issues. Phase 2 identified the prevalence of barriers to COVID-19 testing in a nationally representative sample and highlighted important health literacy disparities. However, even the second phase was not representative of all community groups, particularly those from culturally and linguistically diverse backgrounds, which has been identified as a key area of need in Australia and worldwide. We conducted a separate survey with these groups using interpreters to conduct the survey in phase 1 in preferred languages as a partnership with the Western Sydney Local Health District [50-52].

In phase 3, the intervention's highly tailored design meant that there was considerable heterogeneity in the intervention elements that participants received. This may have contributed to the lack of observed effect. It is possible that there were ceiling effects for testing intentions when participants assumed that they would get tested when they were not currently or recently thinking about the logistics of getting tested. It is also possible that participants did not engage with the text-based intervention content. We attempted to address these methodological issues in the final phase 4 by focusing on a consistent set of key knowledge barriers in a more targeted group with lower baseline testing intentions (younger and lower education), using a more engaging intervention format (animation with text and audio and a social media–style video), and including more sensitive measures of testing intention to avoid potential biases.

The audio-visual interventions produced for phase 4 have information that is specific to the Australian context and may not be useful in other countries but can be used as a starting point for new knowledge interventions. We found that a simple and relatively cheap animation focused on key messages or a TikTok-style video that incorporated humor was both effective for increasing knowledge but not testing intentions. However, the findings may not be generalizable to other contexts, particularly where opportunity issues such as cost or physical access to testing are a problem. We expect that there will be additional barriers to RATs. Nevertheless, this study provides a comprehensive list of testing barriers that may help us better prepare for future variants or the next pandemic.

Conclusions

To prepare for future pandemics, capability, opportunity, and motivation issues need to be addressed to increase testing behaviors for novel viruses, particularly in groups with lower health literacy. Audio-visual interventions can be used to address key knowledge issues in target populations. Our findings support broader advice from international experts on behavior change, highlighting the importance of diagnosing behavioral barriers to increase adherence to COVID-19 prevention behaviors.

Acknowledgments

The authors thank the participants of the longitudinal COVID-19 survey for their ongoing participation in this research, Wendy Liang for providing consumer input for the written intervention, and MD for developing the TikTok version of the intervention. The surveys were not specifically funded, but in-kind support was provided by the authors through research fellowships. The experimental trials were funded by the Marie Bashir Institute (now known as the Sydney Infectious Diseases Institute) at the University of Sydney. CB is supported by a National Health and Medical Research Council and Heart Foundation Early Career Fellowship (number 1122788).

KM is supported by a National Health and Medical Research Council Principal Research Fellowship (number 1121110).

Data Availability

Anonymous data are available from the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Phase 1 and 2 questionnaires.
[DOCX File , 60 KB - publichealth_v9i1e40441_app1.docx]

Multimedia Appendix 2
Screenshots from phase 3.
[PDF File (Adobe PDF File), 1681 KB - publichealth_v9i1e40441_app2.pdf]

Multimedia Appendix 3
Phase 4 materials.
[DOCX File , 19719 KB - publichealth_v9i1e40441_app3.docx]

Multimedia Appendix 4
Participant characteristics.
[DOCX File , 20 KB - publichealth_v9i1e40441_app4.docx]

Multimedia Appendix 5
CONSORT-eHEALTH checklist (V. 1.6.1).
[PDF File (Adobe PDF File), 640 KB - publichealth_v9i1e40441_app5.pdf]

References

1. Michie S, West R. Behavioural, environmental, social, and systems interventions against COVID-19. BMJ 2020 Jul 28;370:m2982. [doi: 10.1136/bmj.m2982] [Medline: 32723732]
2. West R, Michie S, Rubin GJ, Amlôt R. Applying principles of behaviour change to reduce SARS-CoV-2 transmission. Nat Hum Behav 2020 May;4(5):451-459. [doi: 10.1038/s41562-020-0887-9] [Medline: 32377018]
3. COVID-19 (Coronavirus) statistics. The New South Wales Ministry of Health. 2020. URL: https://www.health.nsw.gov.au/news/Pages/20201231_00.aspx [accessed 2021-04-24]
4. Nakkazi E. Obstacles to COVID-19 control in east Africa. Lancet Infect Dis 2020 Jun;20(6):660 [FREE Full text] [doi: 10.1016/S1473-3099(20)30382-0] [Medline: 32473138]
5. McElfish PA, Purvis R, James LP, Willis DE, Andersen JA. Perceived barriers to COVID-19 testing. Int J Environ Res Public Health 2021 Feb 25;18(5):2278 [FREE Full text] [doi: 10.3390/ijerph18052278] [Medline: 33668958]
6. Capps R, Gelatt J. Barriers to COVID-19 testing and treatment: immigrants without health coverage in the United States. Migration Policy Institute. 2020 May. URL: https://www.migrationpolicy.org/research/covid-19-testing-treatment-immigrants-health-insurance [accessed 2022-05-01]
7. Frieden J. Lack of health literacy a barrier to grasping COVID-19. MedPage Today. 2020 Jun 10. URL: https://www.medpagetoday.com/infectiousdisease/covid19/87002 [accessed 2020-09-08]
8. Bruton B, Edwards N. Barriers to mass testing for COVID-19 in Africa. Atlantic Council. 2020 Apr 17. URL: https://www.atlanticcouncil.org/blogs/africasource/barriers-to-mass-testing-for-covid-19-in-africa/ [accessed 2020-09-08]

9. What you need to know about coronavirus (COVID-19). Australian Government Department of Health and Aged Care. 2020 Sep 10. URL: https://www.health.gov.au/news/health-alerts/novel-coronavirus-2019-ncov-health-alert/what-you-need-to-know-about-coronavirus-covid-19 [accessed 2020-09-23]
10. Maasdorp J. How long should it take to get a COVID-19 test result? Here's the breakdown for each state. ABC News. 2020 Aug 11. URL: https://www.abc.net.au/news/2020-08-11/how-long-should-it-take-to-get-a-coronavirus-covid19-test-result/12545150 [accessed 2020-09-23]
11. Flu tracking reports. FluTracking. URL: https://info.flutracking.net/reports/australia-reports/ [accessed 2023-03-23]
12. Wolf MS, Serper M, Opsasnick L, O'Conor RM, Curtis L, Benavente JY, et al. Awareness, attitudes, and actions related to COVID-19 among adults with chronic conditions at the onset of the U.S. outbreak: a cross-sectional survey. Ann Intern Med 2020 Jul 21;173(2):100-109 [FREE Full text] [doi: 10.7326/M20-1239] [Medline: 32271861]
13. McCaffery KJ, Dodd RH, Cvejic E, Ayrek J, Batcup C, Isautier JM, et al. Health literacy and disparities in COVID-19-related knowledge, attitudes, beliefs and behaviours in Australia. Public Health Res Pract 2020 Dec 09;30(4):30342012 [FREE Full text] [doi: 10.17061/phrp30342012] [Medline: 33294907]
14. Pickles K, Cvejic E, Nickel B, Copp T, Bonner C, Leask J, et al. COVID-19 misinformation trends in Australia: prospective longitudinal national survey. J Med Internet Res 2021 Jan 07;23(1):e23805 [FREE Full text] [doi: 10.2196/23805] [Medline: 33302250]
15. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
16. McCaffery K. The effect of enhanced messaging on COVID-19 testing intentions and behaviour. Australian New Zealand Clinical Trials Registry. 2021 Jul 07. URL: https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=382318&isReview=true [accessed 2022-03-29]
17. McCaffery K. Optimising COVID-19 testing intentions and behaviour with enhanced messaging. Australian New Zealand Clinical Trials Registry. 2020 Dec 16. URL: https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=380916&isReview=true [accessed 2022-03-29]
18. Agency for healthcare research and quality, Rockville, health literacy universal precautions toolkit. 2nd edition. Agency for Healthcare Research and Quality. 2015. URL: https://www.ahrq.gov/health-literacy/improve/precautions/index.html [accessed 2023-06-08]
19. Ayre J, Isautier J, Copp T, Lawrence F, Mac O, Perks L, et al. Health literacy editor. Sydney Health Literacy Lab. 2021 Mar. URL: https://ses.library.usyd.edu.au/handle/2123/24642 [accessed 2023-03-01]
20. Ayre J, Cvejic E, Bonner C, Turner RM, Walter SD, McCaffery KJ. Effects of health literacy, screening, and participant choice on action plans for reducing unhealthy snacking in Australia: a randomised controlled trial. PLoS Med 2020 Nov 03;17(11):e1003409 [FREE Full text] [doi: 10.1371/journal.pmed.1003409] [Medline: 33141834]
21. Bonner C, Batcup C, Ayre J, Cvejic E, Trevena L, McCaffery K, et al. The impact of health literacy-sensitive design and heart age in a cardiovascular disease prevention decision aid: randomized controlled trial and end-user testing. JMIR Cardio 2022 Apr 15;6(1):e34142 [FREE Full text] [doi: 10.2196/34142] [Medline: 35436208]
22. Bélanger-Gravel A, Godin G, Amireault S. A meta-analytic review of the effect of implementation intentions on physical activity. Health Psychol Rev 2013;7(1):23-54 [FREE Full text] [doi: 10.1080/17437199.2011.560095]
23. Carrero I, Vilà I, Redondo R. What makes implementation intention interventions effective for promoting healthy eating behaviours? A meta-regression. Appetite 2019 Sep 01;140:239-247. [doi: 10.1016/j.appet.2019.05.024] [Medline: 31125588]
24. Kompf J. Implementation intentions for exercise and physical activity: who do they work for? A systematic review. J Phys Act Health 2020 Mar 01;17(3):349-359. [doi: 10.1123/jpah.2018-0720] [Medline: 31923898]
25. Adriaanse MA, Vinkers CD, De Ridder DT, Hox JJ, De Wit JB. Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. Appetite 2011 Feb;56(1):183-193 [FREE Full text] [doi: 10.1016/j.appet.2010.10.012] [Medline: 21056605]
26. McWilliams L, Bellhouse S, Yorke J, Lloyd K, Armitage CJ. Beyond "planning": a meta-analysis of implementation intentions to support smoking cessation. Health Psychol 2019 Dec;38(12):1059-1068 [FREE Full text] [doi: 10.1037/hea0000768] [Medline: 31414843]
27. Armitage CJ. A volitional help sheet to encourage smoking cessation: a randomized exploratory trial. Health Psychol 2008 Sep;27(5):557-566. [doi: 10.1037/0278-6133.27.5.557] [Medline: 18823182]
28. Armitage CJ, Arden MA. A volitional help sheet to increase physical activity in people with low socioeconomic status: a randomised exploratory trial. Psychol Health 2010 Dec;25(10):1129-1145. [doi: 10.1080/08870440903121638] [Medline: 20309777]
29. Ayre J, Bonner C, Cvejic E, McCaffery K. Randomized trial of planning tools to reduce unhealthy snacking: implications for health literacy. PLoS One 2019 Jan 17;14(1):e0209863 [FREE Full text] [doi: 10.1371/journal.pone.0209863] [Medline: 30653531]
30. Sim M. Barriers to COVID-19 testing and testing hesitancy: a rapid scoping review. The SPOR Evidence Alliance. 2021 Jan 08. URL: https://sporevidencealliance.ca/wp-content/uploads/2021/05/Barriers-to-COVID-19-testing-and-Testing-Hesitancy_Full-Report_2021.02.03.pdf [accessed 2022-05-01]

31. Shahrour G, Jaradat D, Dardas LA. Barriers related to COVID-19 testing intention. Public Health Nurs 2021 Nov;38(6):978-983 [FREE Full text] [doi: 10.1111/phn.12950] [Medline: 34313354]
32. Graham MS, May A, Varsavsky T, Sudre CH, Murray B, Kläser K, et al. Knowledge barriers in a national symptomatic-COVID-19 testing programme. PLOS Glob Public Health 2022 Jan 19;2(1):e0000028 [FREE Full text] [doi: 10.1371/journal.pgph.0000028] [Medline: 36962066]
33. Clipman SJ, Wesolowski A, Mehta SH, Agarwal S, Cobey SE, Cummings DA, et al. SARS-CoV-2 testing in Florida, Illinois, and Maryland: access and barriers. medRxiv. Preprint posted online December 24, 2020 2023 [FREE Full text] [doi: 10.1101/2020.12.23.20248789] [Medline: 33398298]
34. AlKetbi LM, Nagelkerke N, Abdelbaqi H, Al Blooshi F, Al Saedi M, Al mansoori S, et al. SARS-Cov-2 testing in the United Arab Emirates: population attitudes and beliefs. medRxiv. Preprint posted online February 23, 2021 2023 [FREE Full text] [doi: 10.1101/2021.02.19.21251841]
35. Robinson PA, Levy DC, Hooker C, Shaban RZ, Nahidi S, Leask J, et al. COVID-19 testing decisions and behaviours in two Australian cities. Health Promot J Austr 2023 Apr;34(2):587-594 [FREE Full text] [doi: 10.1002/hpja.599] [Medline: 35332631]
36. Islam MI, Chadwick V, Martiniuk A. Predictors of COVID testing among Australian youth: insights from the longitudinal study of Australian children. medRxiv. Preprint posted online Nov 01, 2021 2021 Nov 01 [FREE Full text] [doi: 10.1101/2021.10.31.21265627]
37. Unger JB, Soto DW, Lee R, Deva S, Shanker K, Sood N, et al. COVID-19 testing in schools: perspectives of school administrators, teachers, parents, and students in southern California. medRxiv. Preprint posted online Oct 14, 2021 2023 [FREE Full text] [doi: 10.1101/2021.10.14.21265000]
38. French CE, Denford S, Brooks-Pollock E, Wehling H, Hickman M. Low uptake of COVID-19 lateral flow testing among university students: a mixed methods evaluation. Public Health 2022 Mar;204:54-62 [FREE Full text] [doi: 10.1016/j.puhe.2022.01.002] [Medline: 35176622]
39. Mowbray F, Woodland L, Smith LE, Amlôt R, Rubin GJ. Is my cough a cold or COVID? A qualitative study of COVID-19 symptom recognition and attitudes toward testing in the UK. Front Public Health 2021 Aug 18;9:1194 [FREE Full text] [doi: 10.3389/fpubh.2021.716421]
40. Teasdale CA, Borrell LN, Shen Y, Kimball S, Rinke ML, Rane MS, et al. COVID-19 testing among US children, parental preferences for testing venues, and acceptability of school-based testing. Public Health Rep 2022 Mar;137(2):362-369 [FREE Full text] [doi: 10.1177/00333549211065518] [Medline: 35023416]
41. Vandrevala T, Alidu L, Hendy J, Shafi S, Ala A. 'It's possibly made us feel a little more alienated': how people from ethnic minority communities conceptualise COVID-19 and its influence on engagement with testing. J Health Serv Res Policy 2022 Apr;27(2):141-150 [FREE Full text] [doi: 10.1177/13558196211054961] [Medline: 34978500]
42. Wanat M, Logan M, Hirst JA, Vicary C, Lee JJ, Perera R, et al. Perceptions on undertaking regular asymptomatic self-testing for COVID-19 using lateral flow tests: a qualitative study of university students and staff. BMJ Open 2021 Sep 02;11(9):e053850 [FREE Full text] [doi: 10.1136/bmjopen-2021-053850] [Medline: 34475190]
43. Knight KR, Duke MR, Carey CA, Pruss G, Garcia CM, Lightfoot M, et al. COVID-19 testing and vaccine acceptability among homeless-experienced adults: qualitative data from two samples. J Gen Intern Med 2021 Oct 26;37(4):823-829 [FREE Full text] [doi: 10.1007/s11606-021-07161-1]
44. Watson D, Baralle NL, Alagil J, Anil K, Ciccognani S, Dewar-Haggart R, et al. How do we engage people in testing for COVID-19? A rapid qualitative evaluation of a testing programme in schools, GP surgeries and a university. BMC Public Health 2022 Feb 14;22(1):305 [FREE Full text] [doi: 10.1186/s12889-022-12657-4] [Medline: 35164728]
45. Concerning pricing of rapid antigen tests. Australian Competition and Consumer Commission. 2022 Jan 17. URL: https://www.accc.gov.au/media-release/concerning-pricing-of-rapid-antigen-tests [accessed 2022-03-29]
46. National cabinet statement. Prime Minister of Australia. URL: https://www.pm.gov.au/media/national-cabinet [accessed 2022-03-29]
47. Free family and staff RAT packs. Government of New South Wales. 2022 Feb 21. URL: https://education.nsw.gov.au/news/latest-news/free-family-and-staff-rat-packs- [accessed 2022-03-29]
48. Cook H. Australia COVID: experts urge caution on RAT accuracy as watchdog probes test kit complaints. Sydney Morning Herald. URL: https://www.smh.com.au/national/experts-urge-caution-on-rat-accuracy-as-watchdog-probes-test-kit-complaints-20220203-p59tgj.html [accessed 2022-03-30]
49. Southerton C. Queensland's coronavirus controversy: past pandemics show us public shaming could harm public health. The Conversation. 2020 Jul 31. URL: https://theconversation.com/queenslands-coronavirus-controversy-past-pandemics-show-us-public-shaming-could-harm-public-health-143699 [accessed 2021-04-24]
50. Ayre J, Muscat DM, Mac O, Batcup C, Cvejic E, Pickles K, et al. Main COVID-19 information sources in a culturally and linguistically diverse community in Sydney, Australia: a cross-sectional survey. Patient Educ Couns 2022 Aug;105(8):2793-2800 [FREE Full text] [doi: 10.1016/j.pec.2022.03.028] [Medline: 35369997]

51. Muscat DM, Ayre J, Mac O, Batcup C, Cvejic E, Pickles K, et al. Psychological, social and financial impacts of COVID-19 on culturally and linguistically diverse communities in Sydney, Australia. BMJ Open 2022 May 10;12(5):e058323 [FREE Full text] [doi: 10.1136/bmjopen-2021-058323] [Medline: 35537788]
52. Ayre J, Muscat DM, Mac O, Batcup C, Cvejic E, Pickles K, et al. COVID-19 testing and vaccine willingness: cross-sectional survey in a culturally diverse community in Sydney, Australia. Health Equity 2022 Dec 28;6(1):965-974 [FREE Full text] [doi: 10.1089/heq.2021.0171] [Medline: 36636119]

Abbreviations

COM-B: capability-opportunity-motivation-behavior
PCR: polymerase chain reaction
RAT: rapid antigen test

Edited by T Sanchez, A Mavragani; submitted 21.06.22; peer-reviewed by D Ivanković, H Pilabré; comments to author 24.01.23; revised version received 31.03.23; accepted 12.05.23; published 29.06.23.

Please cite as:
Bonner C, Batcup C, Cvejic E, Ayre J, Pickles K, Copp T, Cornell S, Nickel B, Dhahir M, McCaffery K
Addressing Behavioral Barriers to COVID-19 Testing With Health Literacy–Sensitive eHealth Interventions: Results From 2 National Surveys and 2 Randomized Experiments
JMIR Public Health Surveill 2023;9:e40441
URL: https://publichealth.jmir.org/2023/1/e40441
doi:10.2196/40441
PMID:37172319

Original Paper

Evaluation of Social Isolation Trajectories and Incident Cardiovascular Disease Among Middle-Aged and Older Adults in China: National Cohort Study

Lan Guo[1*], PhD; Wanxin Wang[1*], PhD; Jingman Shi[1], PhD; Xinyu Zheng[1], MSc; Yilin Hua[1], MSc; Ciyong Lu[1], PhD

School of Public Health, Sun Yat-sen University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Ciyong Lu, PhD
School of Public Health
Sun Yat-sen University
No 74 Zhongshan 2nd
Guangzhou, 510080
China
Phone: 86 87332477
Email: luciyong@mail.sysu.edu.cn

Abstract

Background: Although the association between social isolation and the risk of subsequent cardiovascular disease (CVD) is well documented, most studies have only assessed social isolation at a single time point, and few studies have considered the association using repeatedly measured social isolation.

Objective: This study aimed to examine the association between social isolation trajectories and incident CVD in a large cohort of middle-aged and older adults.

Methods: This study used data from 4 waves (wave 1, wave 2, wave 3, and wave 4) of the China Health and Retirement Longitudinal Study. We defined the exposure period as from June 2011 to September 2015 (from wave 1 to wave 3) and the follow-up period as from September 2015 to March 2019 (wave 4). On the basis of the inclusion and exclusion criteria, our final analytic sample included 8422 individuals who had no CVD in the China Health and Retirement Longitudinal Study in waves 1 to 3 and were fully followed up in wave 4. Social isolation was ascertained using an extensively used questionnaire at 3 consecutive, biennial time points from waves 1 to 3, and individuals were assigned to 3 predefined social isolation trajectories based on their scores at each wave (consistently low, fluctuating, and consistently high). Incident CVD included self-reported physician-diagnosed heart disease and stroke combined. Cox proportional hazard models estimated the associations of social isolation trajectories with risks of incident CVD, adjusting for demographics, health behaviors, and health conditions.

Results: Of the 8422 participants (mean age 59.76, SD 10.33 years at baseline), 4219 (50.09%) were male. Most of the participants (5267/8422, 62.54%) had consistently low social isolation over time and 16.62% (1400/8422) of the participants had consistently high social isolation over the exposure period. During the 4-year follow-up, 746 incident CVDs occurred (heart disease: 450 cases and stroke: 336 cases). Compared with individuals with consistently low social isolation, those with fluctuating social isolation (adjusted hazard ratio 1.27, 95% CI 1.01-1.59) and consistently high social isolation (adjusted hazard ratio 1.45, 95% CI 1.13-1.85) had higher risks for incident CVD after adjusting for demographics (ie, age, sex, residence, and educational level), health behaviors (ie, smoking status and drinking status), and health conditions (ie, BMI; history of diabetes, hypertension, dyslipidemia, chronic kidney disease; use of diabetes medications, hypertension medications, and lipid-lowering therapy; and depressive symptoms scores).

Conclusions: In this cohort study, middle-aged and older adults with fluctuating and consistently high social isolation exposure had higher risks of the onset of CVD than those without the exposure. The findings suggest that routine social isolation screenings and efforts to improve social connectedness merit increased attention for preventing CVD among middle-aged and older adults.

(JMIR Public Health Surveill 2023;9:e45677) doi:10.2196/45677

KEYWORDS

social isolation trajectories; cardiovascular disease; CVD; cohort study; middle-aged and older adults; China; social isolation

Introduction

Background

Social isolation, defined as an objective reduction in social network size and frequency of social contact, has been shown to be a critical factor contributing to physical and mental health risks [1]. Previous evidence has suggested that it is particularly common in older adults, who may experience decreasing economic resources, physical impairment, or the loss of contemporaries, all of which can limit social contact [2,3]. As the global population ages, there has been an increasing interest in assessing the health effects of social isolation in older adults. The impact of COVID-19 has amplified concerns regarding social isolation among older adults, leading to greater public health awareness [4,5]. In China, it is projected that the proportion of empty-nest families (referring to older individuals with no children or whose children have already left home) will reach 90% [6], highlighting the need for increased attention to social isolation among middle-aged and older adults.

Cardiovascular disease (CVD) is one of the leading causes of global mortality and a major contributor to disability in older adults [7]. Several epidemiological studies have examined the association between social isolation and CVD; however, the results have been inconsistent. Some prospective studies have found that social isolation was associated with a higher risk of developing CVDs, such as coronary heart disease [8], heart failure [9], and stroke [10]. A meta-analysis with 16 cohorts from Europe, North America, Australia, Japan, and Asian Russia showed that social isolation and loneliness are associated with approximately 30% higher risk of coronary heart disease and stroke, even after controlling for age, sex, and socioeconomic status [3]. However, a study that included individuals from the UK Biobank found that the association between social isolation and incident acute myocardial infarction or stroke was attenuated after considering other risk factors, including biological factors (BMI, diastolic and systolic blood pressure, and grip strength), health behaviors (alcohol consumption, physical activity, and smoking), depressive symptoms, socioeconomic factors (education, household income, and Townsend Deprivation Index), and history of chronic illness [11]. A study using data from the English Longitudinal Study of Ageing (ELSA) did not observe a significant association between social isolation and incident CVD over a mean follow-up period of 5.4 years [12]. Similarly, another meta-analysis, including 5 longitudinal studies from Australian adults, reported no association between social isolation and the incidence of CVD [13]. Moreover, there is increasing evidence that social isolation is associated with CVD risk factors such as increased hypertension, smoking, psychological factors (eg, depression), and biological mechanisms (eg, inflammation and stress reactivity) [14-16]. It has been well established that lifestyle [17] and health conditions [18] are associated with an increased risk of CVD outcomes, and therefore, the confounding or covariate influences of these factors on the association between social isolation and incident CVD should be considered when estimating the aforementioned association between social isolation and CVD. Moreover, the level of exposure to social isolation may be dynamic and change over time, with some individuals experiencing long-term social isolation, whereas others experiencing a temporary increase in social isolation [19]. However, prior studies evaluating the association between social isolation and CVD have used social isolation at only a single time point [8-11], which may not provide a complete picture of the role of social isolation in CVD risk. An in-depth understanding of social isolation trajectories may provide clues for identifying individuals with different time-varying levels (eg, consistently high, fluctuating, or consistently low) over their life course. To reduce the risk of CVD in middle-aged and older people, it is crucial to identify its association with social isolation to facilitate research into the development of targeted prevention or early intervention programs.

Objectives

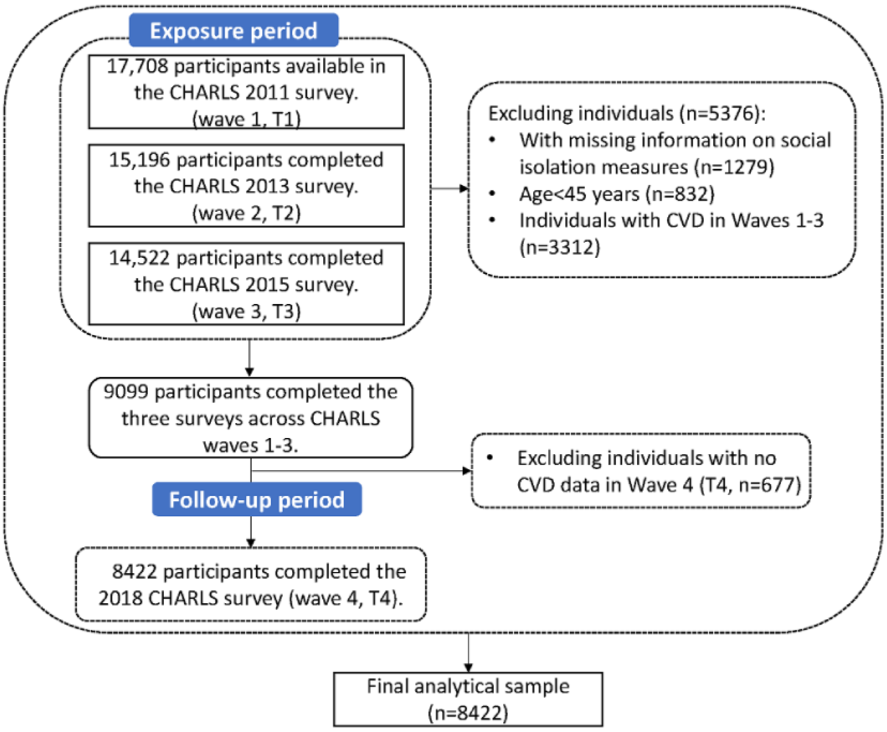
Therefore, to address this research gap, using the nationally representative data from the China Health and Retirement Longitudinal Study (CHARLS), we conducted longitudinal analyses to assess the associations of social isolation trajectories over a period of 4 years with incident CVD among the middle-aged and older Chinese population.

Methods

Study Population

This cohort study used data from CHARLS, an ongoing nationally representative cohort study of Chinese adults aged ≥45 years and their families, which has been previously described in detail [20]. The baseline survey was conducted between June 2011 and March 2012 using a multistage stratified probability-proportional-to-size (PPS) sampling method. In the first stage, 150 county-level units were randomly selected using a PPS sampling technique from a sampling frame containing all county-level units except Tibet. The sample was stratified by region and within each region by urban districts or rural counties and per capita gross domestic product statistics. The final sample of 150 counties was distributed across 28 provinces. The study sample used the lowest level of government organization consisting of administrative villages (*cun* in Chinese) in rural areas and neighborhoods (*shequ* or *juweihui* in Chinese) in urban areas as primary sampling units. A total of 3 primary sampling units within each county-level unit were selected using PPS sampling. At baseline, face-to-face computer-assisted personal interview data were collected from 17,708 participants in 10,257 households, who were recruited from 450 villages and resident communities across China. The response rate of the baseline survey was 80.5%. All participants were followed up every 2 years after the baseline survey using a face-to-face computer-assisted personal interview [20]. The CHARLS data include 4 waves of data (June 2011 to March 2019). We defined from June 2011 to September 2015 (from T1 to T3) as our exposure period (where T1 refers to wave 1, T2 to wave 2, and T3 to wave 3) and from September 2015 to March 2019 (T4, wave 4) as our follow-up period. On the basis of our inclusion and exclusion criteria, our final analytic sample included 8422 individuals who did not have CVD in CHARLS in waves 1 to 3 and were fully followed up in wave 4. Detailed process of participant selection is presented in Figure 1.

Figure 1. Flow diagram of participants included in the study. A total of 14,522 individuals had data available for waves 1-3. Individuals could be excluded due to multiple reasons. CHARLS: China Health and Retirement Longitudinal Study; CVD: cardiovascular disease.



Primary Exposure—Social Isolation Trajectories

In each biennial questionnaire, a social isolation index was generated based on social network ties and social activity or engagement [21]. Consistent with previous literature [22,23], the index consists of 4 indicators: currently unmarried (never married, separated, divorced, and widowed), living alone, having less than weekly contact (face-to-face, telephone, or email) with their children, and participating in social activities less than once a month. Detailed definitions of each indicator are provided in Methods in Multimedia Appendix 1 [22-27]. By summing the indicators, the total score ranged from 0 to 4, with higher scores indicating a greater level of social isolation. We performed Bartley sphericity test to analyze the validity structure of the social isolation index, and the *P* values across the first 3 waves were <.001. The rejection of the null hypothesis of Bartley sphericity test suggests that the validity structure of the social isolation index was good [28]. Participants were further categorized as nonisolated (score <2) and isolated (score ≥2) [22,23,29]. We created social isolation trajectories using the dichotomized social isolation variable for each wave (isolated vs nonisolated) [30]. Guided by theory and prior empirical work, we constructed 3 social isolation trajectories (consistently low, fluctuating, and consistently high) a priori based on combinations of social isolation status across the first 3 waves (wave 1 to 3) [31]. Consistently high was defined as isolated at 3 time points across the assessment period, consistently low was defined as nonisolated at 3 time points, and fluctuating encompassed the other social isolation trajectories that did not fit the above classifications. A visual depiction of the trajectories is provided in Table S1 in Multimedia Appendix 1.

CVD Outcomes

Incident CVD events were defined as the first occurrence of CVD during the follow-up period. In line with previous literature using CHARLS data [32,33], incident CVD events, including heart disease and stroke, were measured using the following standardized questions: “Have you been told by a doctor that you have been diagnosed with a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems?” or “Have you been told by a doctor that you have been diagnosed with a stroke?” Participants who reported heart disease or stroke during the follow-up period were defined as having an incident CVD.

Covariates

At baseline, information on the sociodemographic status and health-related factors was collected by trained interviews using a structured questionnaire. Demographic variables included age, sex, area of residence (rural or urban), and educational level (no formal education, primary school or below, middle or high school, and college or above). Health-related factors included smoking and drinking status (never, former, and current), BMI, depressive symptoms, hypertension, diabetes, dyslipidemia, chronic kidney disease, and the use of medications or therapies for hypertension, diabetes, and dyslipidemia. Detailed definitions of the covariates are provided in Multimedia Appendix 1.

Depressive symptoms were assessed using the 10-item Center for Epidemiology Scale for Depression (CESD-10), which has been validated and widely used among Chinese adults [34]. The CESD-10 consists of 10 items: (1) bothered by little things, (2) had trouble concentrating, (3) felt depressed, (4) everything was an effort, (5) felt hopeful, (6) felt fearful, (7) sleep was restless,

(8) felt happy, (9) felt lonely, and (10) could not get going. Each depressive symptom item in the past week was measured from 0 (rarely or none of the time [<1 day]) to 3 (most or all of the time [5-7 days]). The sum of CESD-10 scores ranges from 0 to 30, with higher scores indicating a higher level of depressive symptoms severity. A total score of ≥12 was used as the cutoff for having depressive symptoms [35]. Of note, subjective loneliness was measured using the CESD-10.

A subgroup of 5196 CHARLS participants underwent measurements of metabolic biomarkers, including fasting plasma glucose, total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, high-sensitivity C-reactive protein (hs-CRP), and serum creatinine. The estimated glomerular filtration rate was calculated using the Chronic Kidney Disease Epidemiology Collaboration’s 2009 creatinine equation [36].

Statistical Analysis

Data were described as mean (SD) or median (IQR) for continuous variables, and frequency with percentage was used to describe categorical variables. First, baseline characteristics were summarized based on the social isolation trajectory group and compared between participants using the chi-square test, ANOVA, or Kruskal-Wallis test, as appropriate. Second, we computed the incidence rates of CVD per 1000 person-years in CHARLS. We also calculated the follow-up time as the time elapsed from the date of the last interview, either the date of CVD diagnosis or the date of the latest interview (March 2019) in which the individual participated. Cox proportional hazard models estimated the hazard ratios (HRs) and 95% CIs for the associations between the social isolation trajectory group and incident CVD using the consistently low social isolation trajectory as the reference. Four models were estimated: model 1, an unadjusted model; in model 2, age and sex were adjusted; in model 3, age, sex, residence, educational level, smoking status, and drinking status were adjusted; and in model 4, the variables in model 3 plus BMI; history of diabetes, hypertension, dyslipidemia, and chronic kidney disease; use of diabetes medications, hypertension medications, and lipid-lowering therapy; and depressive symptoms scores were adjusted. We used interaction items and subgroup analyses to assess whether the potential association between the social isolation trajectory group and incident CVD was moderated by the following characteristics: age, sex, residence, educational level, smoking status, drinking status, diabetes, hypertension, dyslipidemia, chronic kidney disease, BMI, and depressive symptom scores. Four sensitivity analyses were conducted to assess the robustness of our study findings: (1) we further adjusted for metabolic biomarkers in the subgroup of 5196 participants who underwent metabolic examinations, (2) we additionally adjusted for health conditions and behaviors at wave 3 (T3) to estimate changes in associations due to potentially mediating health behavioral pathways, (3) we repeated analyses using the complete data set (7472 participants), and (4) we used inverse-probability weighting to examine the impact of potential selection bias by excluding individuals based on our exposure period [37]. All analyses were conducted using Stata (version 17.0; StataCorp) and R (version 4.2.1; R Foundation for Statistical Computing). Two-sided *P*<.05 was considered statistically significant.

Ethics Approval

The CHARLS study was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015), and ethics approval for the use of the CHARLS data was obtained from the University of Newcastle Human Research Ethics Committee. This study was conducted in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology reporting guidelines. All procedures were performed in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all the participants.

Results

Characteristics of Participants

Of the 8422 participants included in the primary analyses, the mean age of the study population at baseline was 59.76 (SD 10.33) years, and the proportion of male participants was 50.09% (4219/8422). Although the differences between included and excluded participants were relatively small, participants excluded from analyses were older and were more likely to be female; live in rural areas; and have a lower educational level, higher depressive symptoms scores, and lower BMI (Table S2 in Multimedia Appendix 1). Table 1 describes the characteristics of the participants according to their social isolation trajectory group. Among these middle-aged and older adults, the majority (5267/8422, 62.54%) had consistently low social isolation over time and 16.62% (1400/8422) had consistently high social isolation over the exposure period. Individuals with consistently high versus consistently low social isolation were more likely to be older (mean age 66.08, SD 10.55 years vs mean age 58.41, SD 9.88 years; *P*<.001), be female (811/1400, 57.93% vs 2592/5267, 49.21%; *P*<.001), live in rural area (1241/1400, 88.64% vs 3868/5267, 73.44%; *P*<.001), have no formal education (908/1400, 64.86% vs 2029/5267, 38.52%; *P*<.001), be a nonsmoker (855/1400, 61.07% vs 3130/5267, 59.42%; *P*=.04), be a nondrinker (897/1400, 64.07% vs 3148/5267, 59.77%; *P*=.002), have higher CESD-10 scores (mean 10.41, SD 6.70 vs mean 7.67, SD 5.98), have a history of hypertension (602/1400, 43% vs 1866/5267, 35.43%; *P*<.001), less use of diabetes mediations (29/1400, 2.07% vs 184/5267, 3.49%; *P*=.02), have higher systolic pressure (mean 135.65, SD 27.09 mm Hg vs mean 130.36, SD 24.88 mm Hg), have lower triglyceride (mean 118.71, SD 96.77 mg/dL vs mean 131.32, SD 101.87 mg/dL; *P*=.002) but higher high-density lipoprotein cholesterol (mean 54.11, SD 15.94 mg/dL vs mean 50.93, SD 15.28 mg/dL; *P*<.001), and have a lower estimated glomerular filtration rate (mean 88.42, SD 28.99 mL/min/1.73 m^2 vs mean 91.98, SD 36.09 mL/min/1.73 m^2; *P*=.02).

Table 1. Characteristics of the participants according to social isolation trajectory group.

Characteristics	Overall (N=8422)	Social isolation trajectory group			*P* value[a]
		Consistently low (n=5267)	Fluctuating (n=1755)	Consistently high (n=1400)	
Age (years), mean (SD)	59.76 (10.33)	58.41 (9.88)	58.78 (9.57)	66.08 (10.55)	<.001
Sex, male, n (%)	4219 (50.09)	2675 (50.79)	955 (54.41)	589 (42.07)	<.001
Area of residence[b], n (%)					<.001
Rural	6617 (78.57)	3868 (73.43)	1508 (85.93)	1241 (88.64)	
Urban	1803 (21.41)	1398 (26.54)	247 (14.07)	158 (11.28)	
Educational level[b], n (%)					<.001
No formal education	3698 (43.91)	2029 (38.52)	761 (43.36)	908 (64.86)	
Primary school or below	1936 (22.99)	1213 (23.03)	442 (25.18)	281 (20.07)	
Middle or high school	2573 (30.55)	1843 (34.99)	529 (30.14)	201 (14.36)	
College or above	208 (2.47)	177 (3.36)	21 (1.2)	10 (0.71)	
Smoking status[b], n (%)					.04
Nonsmoker	5064 (60.13)	3130 (59.43)	1079 (61.48)	855 (61.07)	
Former smoker	762 (9.05)	454 (8.62)	172 (9.8)	136 (9.71)	
Current smoker	2503 (29.72)	1623 (30.81)	486 (27.69)	394 (28.14)	
Drinking status[b], n (%)					.002
Nondrinker	5046 (59.91)	3148 (59.77)	1001 (57.04)	897 (64.07)	
Former drinker	1172 (13.91)	740 (14.05)	261 (14.87)	171 (12.21)	
Current drinker	2188 (25.98)	1368 (25.97)	491 (27.98)	329 (23.5)	
Depressive symptoms scores[b,c], mean (SD)	8.27 (6.25)	7.67 (5.98)	8.45 (6.34)	10.41 (6.7)	<.001
Health conditions (yes), n (%)					
Diabetes	878 (10.42)	542 (10.29)	191 (10.88)	145 (10.36)	.78
Hypertension	3066 (36.4)	1866 (35.43)	598 (34.07)	602 (43)	<.001
Dyslipidemia	2459 (29.2)	1556 (29.54)	516 (29.4)	387 (27.64)	.37
Chronic kidney disease	480 (5.7)	294 (5.58)	110 (6.27)	76 (5.43)	.50
History of medication use (yes)[b], n (%)					
Diabetes medications	267 (3.17)	184 (3.49)	54 (3.08)	29 (2.1)	.02
Hypertension medications	1385 (16.44)	906 (17.2)	241 (13.73)	238 (17.2)	.003
Lipid-lowering therapy	330 (3.92)	215 (4.08)	70 (3.99)	45 (3.3)	.33
BMI (kg/m^2), mean (SD)	23.21 (2.44)	23.19 (2.47)	23.32 (2.25)	23.13 (2.54)	.18
Blood pressure (mm Hg), mean (SD)					
Systolic pressure	131.11 (25.14)	130.36 (24.88)	129.57 (23.84)	135.65 (27.09)	<.001
Diastolic pressure	75.85 (12.13)	76.18 (12.08)	75.27 (12.02)	75.39 (12.43)	.02
Metabolic biomarkers[d]					
Fasting plasma glucose (mg/dL), mean (SD)	109.11 (33.48)	109.83 (36.14)	107.80 (29.48)	108.29 (28.25)	.16
HbA1c (%), mean (SD)	5.26 (0.78)	5.28 (0.83)	5.24 (0.70)	5.23 (0.67)	.24
Total cholesterol (mg/dL), mean (SD)	194.34 (37.71)	194.19 (37.93)	193.71 (36.69)	195.63 (38.23)	.50
Triglyceride (mg/dL), mean (SD)	127.75 (97.66)	131.32 (101.87)	125.14 (85.54)	118.71 (96.77)	.002

Characteristics	Overall (N=8422)	Social isolation trajectory group			*P* value[a]
		Consistently low (n=5267)	Fluctuating (n=1755)	Consistently high (n=1400)	
High-density lipoprotein (mg/dL), mean (SD)	51.78 (15.42)	50.93 (15.28)	52.25 (15.22)	54.11 (15.94)	<.001
Low-density lipoprotein (mg/dL), mean (SD)	117.92 (34.61)	117.91 (34.58)	117.22 (34.31)	118.90 (35.11)	.56
hs-CRP[e] (mg/L), median (IQR)	1.00 (1.63)	1.04 (1.68)	0.92 (1.36)	1.01 (1.78)	.15
eGFR[f] (mL/min/1.73 m^2), mean (SD)	90.82 (34.11)	91.98 (36.09)	89.36 (31.72)	88.42 (28.99)	.02

[a] *P* value was based on the chi-square test, ANOVA, or the Kruskal-Wallis test where appropriate.

[b] Missing data: 1 for the area of residence, 7 for educational level, 93 for smoking, 16 for drinking, 629 for depressive symptoms scores, 67 for diabetes medications, 37 for hypertension medications, and 178 for lipid-lowering therapy.

[c] Depressive symptom scores were measured using the 10-item Center for Epidemiology Scale for Depression, ranging from 0 to 30, with higher scores indicating a higher level of depressive symptom severity.

[d] Measured in the subpopulation of 5196 participants.

[e] hs-CRP: high-sensitivity C-reactive protein.

[f] eGFR: estimated glomerular filtration rate.

Associations of Social Isolation With Incident CVD

During the follow-up period between 2015 and 2019, a total of 746 incident CVDs occurred (heart disease: 450 cases and stroke: 336 cases). The incidence rate of CVD was 28.14 per 1000 person-years among individuals with consistently low social isolation, 34.81 per 1000 person-years among the fluctuating social isolation group, and 35.26 per 1000 person-years among those with consistently high social isolation. Table 2 shows that after adjusting for covariates in models 2 to 4, compared with individuals with consistently low social isolation (reference), those with fluctuating social isolation (adjusted HR 1.27, 95% CI 1.01-1.59; model 4) and those with consistently high social isolation (adjusted HR 1.45, 95% CI 1.13-1.85; model 4) had a higher risk for incident CVD. For CVD components, compared with the reference group, individuals with fluctuating social isolation (adjusted HR 1.59, 95% CI 1.16-2.17; model 4) and those with consistently high social isolation (adjusted HR 1.75, 95% CI 1.25-2.47; model 4) had a higher risk of stroke, but not heart disease (Table 2).

As shown in Figure 2, the association between the social isolation trajectory group and incident CVD events was not moderated by age, sex, residence, smoking status, drinking status, diabetes, hypertension, dyslipidemia, chronic kidney disease, BMI, and depressive symptoms scores. The association between consistently high social isolation and incident CVD was significant among participants with middle or high school or above educational level (adjusted HR 2.02, 95% CI 1.25-3.27; *P*=.03 for interaction). Moreover, we assessed the risk of incident CVD between fluctuating and consistently high social isolation subgroups. As shown in Table 3, we found that compared with individuals with fluctuating social isolation, those with consistently high social isolation were not significantly associated with an increased risk of incident CVD (*P*>.05).

Intercorrelations of the metabolic biomarkers used in this study are presented in Table S3 in Multimedia Appendix 1. Sensitivity analyses showed that after further adjusting for metabolic biomarkers, the associations of fluctuating social isolation and consistently high social isolation with incident stroke remained statistically significant (Table S4 in Multimedia Appendix 1). Our findings were not largely attenuated by health conditions and behaviors at wave 3 (Table S5 in Multimedia Appendix 1), and similar results were found when complete data analyses were performed (Table S6 in Multimedia Appendix 1) or using inverse-probability weighting (Table S7 in Multimedia Appendix 1).

Table 2. Cox proportional hazard ratios (HRs) for the association of social isolation trajectories with incident cardiovascular disease (CVD).

Outcome	Number of cases, n	Incidence rate (per 1000 person-years)	HR (95% CI)			
			Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
CVD						
Social isolation trajectory group						
Consistently low	428	28.14	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Fluctuating	175	34.81	1.26 (1.02-1.56)	1.26 (1.02-1.56)	1.25 (1.00-1.55)	1.27 (1.01-1.59)
Consistently high	143	35.26	1.45 (1.16-1.80)	1.40 (1.11-1.76)	1.38 (1.09-1.75)	1.45 (1.13-1.85)
Heart disease						
Social isolation trajectory group						
Consistently low	273	17.77	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Fluctuating	103	20.24	1.06 (0.83-1.47)	1.07 (0.77-1.49)	1.04 (0.74-1.43)	1.03 (0.75-1.40)
Consistently high	74	18.08	1.10 (0.83-1.47)	1.11 (0.83-1.48)	1.06 (0.79-1.43)	1.11 (0.78-1.58)
Stroke						
Social isolation trajectory group						
Consistently low	181	11.60	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Fluctuating	80	15.41	1.43 (1.06-1.94)	1.43 (1.05-1.93)	1.44 (1.06-1.96)	1.59 (1.16-2.17)
Consistently high	75	18.08	1.83 (1.36-2.47)	1.66 (1.21-2.27)	1.68 (1.21-2.32)	1.75 (1.25-2.47)

[a] Model 1 was an unadjusted model.

[b] Model 2 was adjusted for age and sex.

[c] Model 3 was adjusted for age, sex, residence, educational level, smoking status, and drinking status.

[d] Model 4 was adjusted as model 3 plus BMI; history of diabetes, hypertension, dyslipidemia, chronic kidney disease; use of diabetes medications, hypertension medications, lipid-lowering therapy; and depressive symptoms scores.

Figure 2. Longitudinal association of social isolation trajectories with incident cardiovascular disease (CVD) during the follow-up period. Graphs show hazard ratios (HRs) and 95% CIs for (A) CVD, (B) heart disease, and (C) stroke, adjusted for age; sex; residence; educational level; smoking status; drinking status; BMI; history of diabetes, hypertension, dyslipidemia, and chronic kidney disease; use of diabetes medications, hypertension medications, and lipid-lowering therapy; and depressive symptoms scores. *Consistently low social isolation (reference).

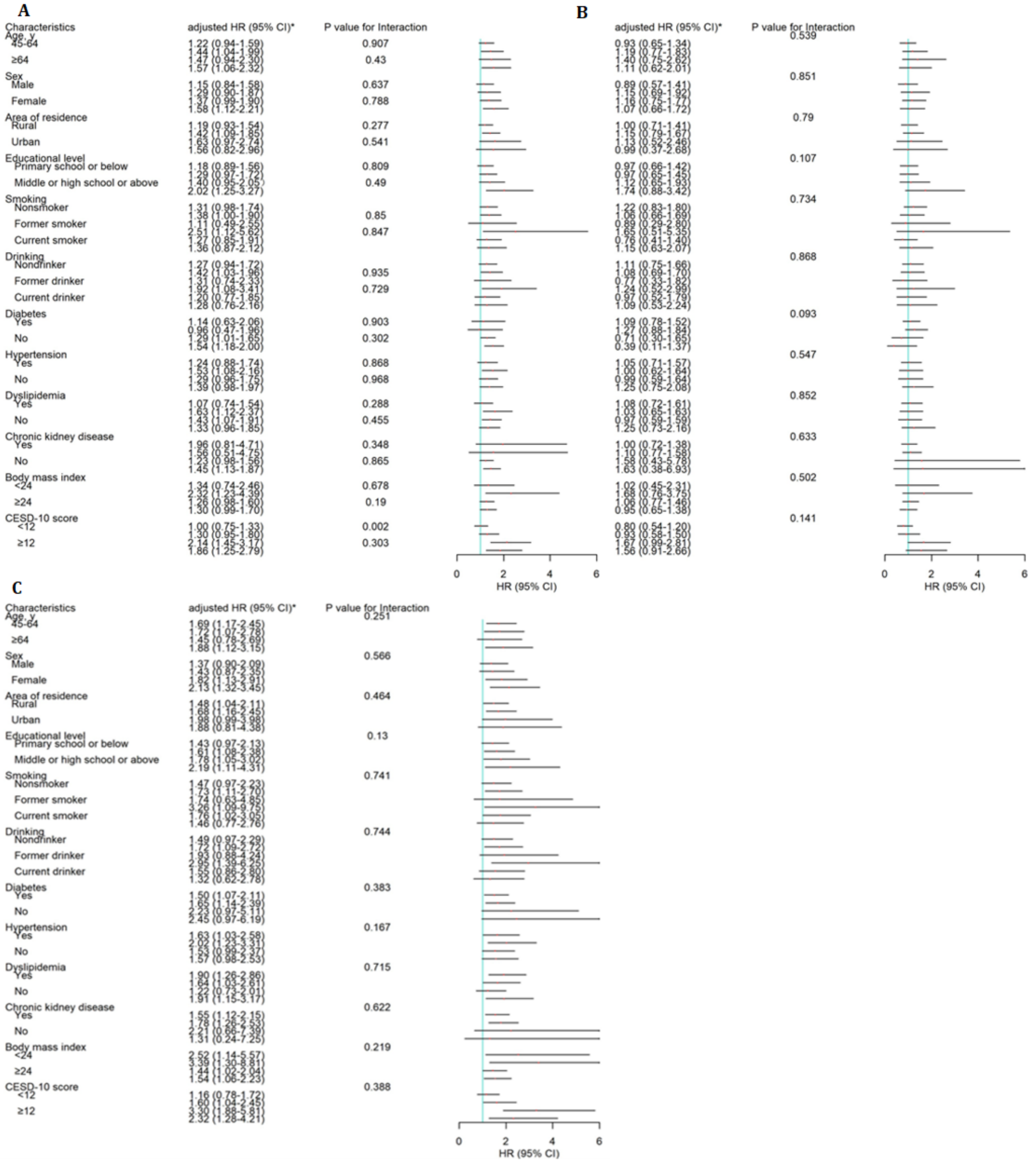


Table 3. Cox proportional hazard ratios (HRs) for the association of social isolation trajectories with incident cardiovascular disease (CVD).

Outcome	Number of cases, n	Incidence rate (per 1000 person-years)	HR (95% CI)			
			Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
CVD						
Social isolation trajectory group						
Fluctuating	175	34.81	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Consistently high	143	35.26	1.15 (0.88-1.48)	1.10 (0.83-1.45)	1.11 (0.84-1.47)	1.12 (0.83-1.50)
Heart disease						
Social isolation trajectory group						
Fluctuating	103	20.24	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Consistently high	74	18.08	0.96 (0.66-1.40)	0.96 (0.64-1.43)	0.97 (0.65-1.46)	1.08 (0.70-1.65)
Stroke						
Social isolation trajectory group						
Fluctuating	80	15.41	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Consistently high	75	18.08	1.28 (0.90-1.81)	1.19 (0.82-1.73)	1.20 (0.82-1.75)	1.09 (0.74-1.62)

[a] Model 1 was an unadjusted model.

[b] Model 2 was adjusted for age and sex.

[c] Model 3 was adjusted for age, sex, residence, educational level, smoking status, and drinking status.

[d] Model 4 was adjusted as model 3 plus BMI; history of diabetes, hypertension, dyslipidemia, chronic kidney disease; use of diabetes medications, hypertension medications, and lipid-lowering therapy; and depressive symptoms scores.

Discussion

Principal Findings

In this cohort study of middle-aged and older Chinese adults, individuals in the fluctuating and consistently high social isolation trajectory groups had a 27% and 45% higher risk for developing incident CVD than those with consistently low social isolation. The association between consistently high social isolation and incident CVD remained statistically significant after demographics, baseline health behaviors and conditions, and baseline depressive symptoms (including loneliness measurement) had been taken into account in multivariable models. However, after further adjusting for metabolic biomarkers, only the associations of fluctuating social isolation and consistently high social isolation with incident stroke remained statistically significant.

Several cross-sectional and longitudinal studies have suggested that the presence of social isolation is associated with a higher risk of coronary heart disease, heart failure [9], stroke [10], CVD diagnoses [38-40], and CVD risks [41]. Nagayoshi et al [10] demonstrated that community-dwelling men and women in the United States who reported having a small social network had an approximately 40% greater risk of incident stroke compared with their counterparts who reported a large social network, even after adjusting for demographics, socioeconomic variables, behavioral risk factors, and major stroke risk factors. Freak-Poli et al [40] reported a positive association between social isolation and incident CVDs, particularly stroke. However, these studies only assessed social isolation at a single time point, which may not provide a complete picture of the role of social isolation over the life span. In reality, the level of exposure to social isolation may be dynamic and vary over time [19]. Our study observed that 62.54% (5267/8422) of the study population had consistently low social isolation, 16.62% (1400/8422) had consistently high social isolation, and 20.83% (1755/8422) had fluctuating social isolation over the exposure period. Research on the longitudinal association between social isolation trajectories and CVD events is limited. Our study used repeated measures of social isolation and found that repeated occurrences of social isolation (ie, consistently high social isolation) may increase the risk of incident CVD, particularly stroke. The potential mechanisms underlying the association include elevated adrenaline levels in the blood, increased sympathetic nervous system activity, dysregulated heart rate and blood pressure, and modulated cardiovascular reactivity, as well as the activation of the hypothalamic-pituitary-adrenocortical axis and the sympathetic nervous system, or dysregulated inflammation [14,15,42]. In addition, social isolation is associated with unhealthy behavioral factors (eg, smoking, drinking, obesity, and lack of medication compliance) and psychological status (eg, depression), which may increase vascular risk and are considered CVD risk factors [43-47]. Consistent with previous literature [42,48-50], our study also observed that individuals in the consistently high social isolation group were more likely to be older, be female, live in more deprived areas, have higher systolic pressure, have higher depressive symptoms scores, and have poor medication compliance relative to those with consistently low social isolation. Our adjusted Cox regression models indicated that these factors might modestly explain the observed association between social isolation and CVD events. Furthermore, our sensitivity analysis, which was additionally adjusted for health conditions and behaviors at wave 3, resulted in attenuated risk estimates. Similarly, a large-scale prospective study using data from the UK Biobank reported that the association of social isolation with acute myocardial infarction and stroke was attenuated after adjusting for demographics, biological factors, health behaviors, and depressive symptoms [11]. Another large-scale prospective study using data from the UK Biobank and Million Women's study found that after controlling for established cardiovascular risk factors (eg, age, sex, and self-rated health) and health behaviors, no significant overall association between social isolation and nonfatal coronary heart disease and stroke was observed. However, in contrast, social isolation had clear associations with fatal coronary heart disease and stroke [49]. Furthermore, the association between social isolation and increased levels of inflammatory biomarkers, such as hs-CRP, are well documented [51,52]. In this study, after adjusting for hs-CRP and other metabolic biomarkers, the association of fluctuating social isolation and consistently high social isolation with incident stroke remained significant, suggesting that these findings are robust. Apart from the abovementioned mechanisms linking social isolation and CVD events, there may also be trajectory-specific mechanisms that require further investigation.

In addition, our study found that the association between the social isolation trajectory group and incident CVD events was not moderated by age, sex, residence, smoking status, drinking status, diabetes, hypertension, dyslipidemia, chronic kidney disease, BMI, and depressive symptoms scores. Only educational level modestly moderated the association between social isolation and incident CVD, with the highest HR observed in individuals with middle or high school education or above educational level. Similarly, Cené et al [9] did not observe the moderation effects of age, race, and ethnicity on the association between social isolation and incident heart failure among older women in the United States. Bu et al [39] reported no evidence of a moderation effect of age, sex, socioeconomic status, and CVD risk (established by obesity, high cholesterol, hypertension, diabetes, smoking, diet, physical activity, abnormal sleep, and depression) on the association between social isolation and CVD events among participants from the ELSA. Although we have no clear explanation for this finding, our study indicates that the social isolation trajectory and incident CVD may not differ in individuals with different ages, sexes, residences, BMI, or other abovementioned characteristics. However, for individuals with higher educational levels, the influences of social isolation trajectory on incident CVD may be greater than that for those with lower education levels. Nevertheless, the observed moderation effects of educational level or other unobserved findings still require further investigation.

Our study also found that individuals in the consistently high social isolation group did not have a higher risk of incident CVD than those with fluctuating social isolation. This result may be because individuals with fluctuating social isolation may also have increased risks of other health conditions, such as depressive symptoms, hypertension, and obesity, which may, in turn, increase their risk of developing CVDs [53,54]. Moreover, the fluctuating social isolation trajectory group in our study included a subgroup of individuals who were older, lived in rural areas, and had higher depressive symptoms scores than those with consistently low social isolation. On the basis of our findings, we recommend early identification and lifestyle interventions for socially isolated individuals as a means of combating CVDs and promoting healthy aging, with particular attention paid to the persistent occurrence of social isolation. Future research is needed to examine how frequently social isolation should be assessed and how many assessments would be sufficient to present the course of social isolation to determine high-risk groups for CVDs.

Strengths and Limitations

The strengths of our study include its prospective design, repeated measures of social isolation, a long follow-up period, and a wide range of covariates that include health conditions and behaviors and metabolic biomarkers. Repeated social isolation measurements over a long exposure period can minimize the likelihood of reverse causation or the potential bidirectionality of the estimated association between social isolation and CVD event. However, this study has several limitations. First, the diagnosis of CVD relied on self-reported data, which may be subject to reporting bias. Although medical records were not available in the CHARLS [33], previous large-scale studies using data from the Health and Retirement Study [55] or the ELSA [56] found that self-reported CVDs correspond well with CVDs coded according to the International Classification of Diseases in medical records. Second, individuals with missing social isolation measures and those who experienced CVD during the exposure assessment period were excluded from the analyses, which may limit the generalizability of our findings to middle-aged and older adults. Third, the predefined trajectory group definition of social isolation in our study may be only one possible representation of the accumulative progression of social isolation over time. Fourth, our study sample only included Chinese adults, and therefore our study findings may not be fully generalizable to other ethnicities. Fifth, although our social isolation index was based on previous research [22,23], it was not a standardized scale or instrument. Sixth, our analyses did not consider physical activity as a covariate because it was not assessed using a standardized scale and may be an intermediate variable between social isolation and CVD [57-61]. Finally, in one of our sensitivity analyses presented in Table S4 in Multimedia Appendix 1, all metabolic biomarkers were adjusted to test the robustness of our main findings. However, some of these markers appeared to be moderately correlated, which may have led to collinearity issues.

Conclusions

In this cohort study, both fluctuating and consistently high social isolation exposure were associated with a higher risk of incident CVD among middle-aged and older Chinese adults, especially for stroke. Our findings provide new evidence supporting the longitudinal link between social isolation trajectories and CVD, highlighting the importance of routine screening for social isolation, and providing preventive or intervention strategies to both fluctuate isolated or consistently isolated individuals to reduce the incidence of CVD and promote healthy aging. In addition, our results suggest that efforts to improve social connectedness and reduce social isolation among middle-aged and older adults are recommended.

Acknowledgments

The authors would like to thank the China Health and Retirement Longitudinal Study (CHARLS) research team for providing the data. LG and CL had full access to all the data in the study and took responsibility for the integrity of the data and accuracy of the data analysis.

Data Availability

The original data sets are publicly available on the website of the China Health and Retirement Longitudinal Study (CHARLS) [62]. The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

LG and WW contributed equally to this work. All authors were equally involved in the concept and design of this study. All authors did data acquisition, analysis, or interpretation. LG drafted the manuscript. LG, WW, JS, XZ, and YH critically revised the manuscript for important intellectual content. LG performed the statistical analysis. CL and LG provided the administrative, technical, or material support. CL and LG supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Details of the measured variables, trajectory definitions, baseline survey characteristics, intercorrelations of metabolic markers, and sensitivity analyses.

[DOCX File , 45 KB - publichealth_v9i1e45677_app1.docx]

References

1. Holt-Lunstad J. Loneliness and social isolation as risk factors: the power of social connection in prevention. Am J Lifestyle Med 2021 May 06;15(5):567-573 [FREE Full text] [doi: 10.1177/15598276211009454] [Medline: 34646109]
2. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. Proc Natl Acad Sci U S A 2013 Apr 09;110(15):5797-5801 [FREE Full text] [doi: 10.1073/pnas.1219686110] [Medline: 23530191]
3. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul 01;102(13):1009-1016 [FREE Full text] [doi: 10.1136/heartjnl-2015-308790] [Medline: 27091846]
4. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 Oct;32(10):1217-1220 [FREE Full text] [doi: 10.1017/S1041610220000988] [Medline: 32450943]
5. Cosco TD, Fortuna K, Wister A, Riadi I, Wagner K, Sixsmith A. COVID-19, social isolation, and mental health among older adults: a digital catch-22. J Med Internet Res 2021 May 06;23(5):e21864 [FREE Full text] [doi: 10.2196/21864] [Medline: 33891557]
6. Zhang HH, Jiang YY, Rao WW, Zhang QE, Qin MZ, Ng CH, et al. Prevalence of depression among empty-nest elderly in China: a meta-analysis of observational studies. Front Psychiatry 2020 Jul 07;11:608 [FREE Full text] [doi: 10.3389/fpsyt.2020.00608] [Medline: 32733289]
7. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
8. Golaszewski NM, LaCroix AZ, Godino JG, Allison MA, Manson JE, King JJ, et al. Evaluation of social isolation, loneliness, and cardiovascular disease among older women in the US. JAMA Netw Open 2022 Feb 01;5(2):e2146461 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.46461] [Medline: 35107574]
9. Cené CW, Leng XI, Faraz K, Allison M, Breathett K, Bird C, et al. Social isolation and incident heart failure hospitalization in older women: women's health initiative study findings. J Am Heart Assoc 2022 Mar;11(5):e022907 [FREE Full text] [doi: 10.1161/JAHA.120.022907] [Medline: 35189692]
10. Nagayoshi M, Everson-Rose SA, Iso H, Mosley Jr TH, Rose KM, Lutsey PL. Social network, social support, and risk of incident stroke: atherosclerosis risk in communities study. Stroke 2014 Oct;45(10):2868-2873 [FREE Full text] [doi: 10.1161/STROKEAHA.114.005815] [Medline: 25139878]

11. Hakulinen C, Pulkki-Råback L, Virtanen M, Jokela M, Kivimäki M, Elovainio M. Social isolation and loneliness as risk factors for myocardial infarction, stroke and mortality: UK Biobank cohort study of 479 054 men and women. Heart 2018 Sep;104(18):1536-1542. [doi: 10.1136/heartjnl-2017-312663] [Medline: 29588329]
12. Valtorta NK, Kanaan M, Gilbody S, Hanratty B. Loneliness, social isolation and risk of cardiovascular disease in the English Longitudinal Study of Ageing. Eur J Prev Cardiol 2018 Sep;25(13):1387-1396. [doi: 10.1177/2047487318792696] [Medline: 30068233]
13. Freak-Poli R, Phyo AZ, Hu J, Barker SF. Are social isolation, lack of social support or loneliness risk factors for cardiovascular disease in Australia and New Zealand? A systematic review and meta-analysis. Health Promot J Austr 2022 Oct;33 Suppl 1(Suppl 1):278-315 [FREE Full text] [doi: 10.1002/hpja.592] [Medline: 35263481]
14. Paul E, Bu F, Fancourt D. Loneliness and risk for cardiovascular disease: mechanisms and future directions. Curr Cardiol Rep 2021 May 07;23(6):68 [FREE Full text] [doi: 10.1007/s11886-021-01495-2] [Medline: 33961131]
15. Xia N, Li H. Loneliness, social isolation, and cardiovascular health. Antioxid Redox Signal 2018 Mar 20;28(9):837-851 [FREE Full text] [doi: 10.1089/ars.2017.7312] [Medline: 28903579]
16. Sugaya N, Yamamoto T, Suzuki N, Uchiumi C. The transition of social isolation and related psychological factors in 2 mild lockdown periods during the COVID-19 pandemic in Japan: longitudinal survey study. JMIR Public Health Surveill 2022 Mar 08;8(3):e32694 [FREE Full text] [doi: 10.2196/32694] [Medline: 35107428]
17. Cruz-Cobo C, Bernal-Jiménez M, Vázquez-García R, Santi-Cano MJ. Effectiveness of mHealth interventions in the control of lifestyle and cardiovascular risk factors in patients after a coronary event: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022 Dec 02;10(12):e39593 [FREE Full text] [doi: 10.2196/39593] [Medline: 36459396]
18. Oh SH, Kim D, Hwang J, Kang JH, Kwon Y, Kwon JW. Association of uncontrolled hypertension or diabetes mellitus with major adverse cardiovascular events and mortality in South Korea: population-based cohort study. JMIR Public Health Surveill 2023 Feb 03;9:e42190 [FREE Full text] [doi: 10.2196/42190] [Medline: 36735297]
19. Xiang X, Lai PH, Bao L, Sun Y, Chen J, Dunkle RE, et al. Dual trajectories of social isolation and dementia in older adults: a population-based longitudinal study. J Aging Health 2021 Jan;33(1-2):63-74 [FREE Full text] [doi: 10.1177/0898264320953693] [Medline: 32865104]
20. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China health and retirement longitudinal study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
21. Nicholson Jr NR. Social isolation in older adults: an evolutionary concept analysis. J Adv Nurs 2009 Jun;65(6):1342-1352. [doi: 10.1111/j.1365-2648.2008.04959.x] [Medline: 19291185]
22. Luo F, Guo L, Thapa A, Yu B. Social isolation and depression onset among middle-aged and older adults in China: moderating effects of education and gender differences. J Affect Disord 2021 Mar 15;283:71-76. [doi: 10.1016/j.jad.2021.01.022] [Medline: 33524661]
23. Lin L, Cao B, Chen W, Li J, Zhang Y, Guo VY. Association of adverse childhood experiences and social isolation with later-life cognitive function among adults in China. JAMA Netw Open 2022 Nov 01;5(11):e2241714 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.41714] [Medline: 36367722]
24. Wang JG, Chia YC, Chen CH, Park S, Hoshide S, Tomitani N, et al. What is new in the 2018 Chinese hypertension guideline and the implication for the management of hypertension in Asia? J Clin Hypertens (Greenwich) 2020 Mar;22(3):363-368 [FREE Full text] [doi: 10.1111/jch.13803] [Medline: 31955513]
25. Joint committee for guideline revision. 2016 Chinese guidelines for the management of dyslipidemia in adults. J Geriatr Cardiol 2018 Jan;15(1):1-29. [doi: 10.11909/j.issn.1671-5411.2018.01.011] [Medline: 29434622]
26. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2011 Jan;34 Suppl 1(Suppl 1):S62-S69 [FREE Full text] [doi: 10.2337/dc11-S062] [Medline: 21193628]
27. Chen D, Sun H, Lu C, Chen W, Guo VY. The association between hypertriglyceridemic-waist phenotype and chronic kidney disease: a cohort study and meta-analysis. Sci Rep 2022 Feb 04;12(1):1935 [FREE Full text] [doi: 10.1038/s41598-022-05806-7] [Medline: 35121773]
28. da Silva AM, Padula RS. Factor structure and short version of the modified Fresno test to assess the use of the evidence-based practice in physiotherapists. BMC Med Educ 2021 Feb 27;21(1):135 [FREE Full text] [doi: 10.1186/s12909-021-02535-9] [Medline: 33639895]
29. Gale CR, Westbury L, Cooper C. Social isolation and loneliness as risk factors for the progression of frailty: the English Longitudinal Study of Ageing. Age Ageing 2018 May 01;47(3):392-397 [FREE Full text] [doi: 10.1093/ageing/afx188] [Medline: 29309502]
30. Soh Y, Tiemeier H, Kawachi I, Berkman LF, Kubzansky LD. Eight-year depressive symptom trajectories and incident stroke: a 10-year follow-up of the HRS (Health and Retirement Study). Stroke 2022 Aug;53(8):2569-2576 [FREE Full text] [doi: 10.1161/STROKEAHA.121.037768] [Medline: 35603598]
31. Howard AF, Tan de Bibiana J, Smillie K, Goddard K, Pritchard S, Olson R, et al. Trajectories of social isolation in adult survivors of childhood cancer. J Cancer Surviv 2014 Mar;8(1):80-93 [FREE Full text] [doi: 10.1007/s11764-013-0321-7] [Medline: 24202698]

32. Li H, Zheng D, Li Z, Wu Z, Feng W, Cao X, et al. Association of depressive symptoms with incident cardiovascular diseases in middle-aged and older Chinese adults. JAMA Netw Open 2019 Dec 02;2(12):e1916591 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.16591] [Medline: 31800066]
33. Gao K, Cao LF, Ma WZ, Gao YJ, Luo MS, Zhu J, et al. Association between sarcopenia and cardiovascular disease among middle-aged and older adults: findings from the China health and retirement longitudinal study. EClinicalMedicine 2022 Jan 10;44:101264. [doi: 10.1016/j.eclinm.2021.101264] [Medline: 35059617]
34. Boey KW. Cross-validation of a short form of the CES-D in Chinese elderly. Int J Geriatr Psychiatry 1999 Aug;14(8):608-617. [doi: 10.1002/(sici)1099-1166(199908)14:8<608::aid-gps991>3.0.co;2-z] [Medline: 10489651]
35. Chen H, Mui AC. Factorial validity of the center for epidemiologic studies depression scale short form in older population in China. Int Psychogeriatr 2014 Jan;26(1):49-57. [doi: 10.1017/S1041610213001701] [Medline: 24125553]
36. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro 3rd AF, Feldman HI, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009 May 05;150(9):604-612 [FREE Full text] [doi: 10.7326/0003-4819-150-9-200905050-00006] [Medline: 19414839]
37. Narduzzi S, Golini MN, Porta D, Stafoggia M, Forastiere F. [Inverse probability weighting (IPW) for evaluating and "correcting" selection bias]. Epidemiol Prev 2014 Sep;38(5):335-341. [Medline: 25387748]
38. Naito R, Leong DP, Bangdiwala SI, McKee M, Subramanian SV, Rangarajan S, et al. Impact of social isolation on mortality and morbidity in 20 high-income, middle-income and low-income countries in five continents. BMJ Glob Health 2021 Mar;6(3):e004124 [FREE Full text] [doi: 10.1136/bmjgh-2020-004124] [Medline: 33753400]
39. Bu F, Zaninotto P, Fancourt D. Longitudinal associations between loneliness, social isolation and cardiovascular events. Heart 2020 Sep;106(18):1394-1399 [FREE Full text] [doi: 10.1136/heartjnl-2020-316614] [Medline: 32461329]
40. Freak-Poli R, Ryan J, Neumann JT, Tonkin A, Reid CM, Woods RL, et al. Social isolation, social support and loneliness as predictors of cardiovascular disease incidence and mortality. BMC Geriatr 2021 Dec 13;21(1):711 [FREE Full text] [doi: 10.1186/s12877-021-02602-2] [Medline: 34922471]
41. Hu J, Fitzgerald SM, Owen AJ, Ryan J, Joyce J, Chowdhury E, et al. Social isolation, social support, loneliness and cardiovascular disease risk factors: a cross-sectional study among older adults. Int J Geriatr Psychiatry 2021 Nov;36(11):1795-1809. [doi: 10.1002/gps.5601] [Medline: 34231940]
42. Gronewold J, Engels M, van de Velde S, Cudjoe TK, Duman EE, Jokisch M, et al. Effects of life events and social isolation on stroke and coronary heart disease. Stroke 2021 Jan;52(2):735-747. [doi: 10.1161/STROKEAHA.120.032070] [Medline: 33445957]
43. National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Health and Medicine Division, Board on Behavioral, Cognitive, and Sensory Sciences, Board on Health Sciences Policy, Committee on the Health and Medical Dimensions of Social Isolation and Loneliness in Older Adults. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington, DC, USA: National Academies Press; Feb 27, 2020.
44. Shankar A, McMunn A, Banks J, Steptoe A. Loneliness, social isolation, and behavioral and biological health indicators in older adults. Health Psychol 2011 Jul;30(4):377-385. [doi: 10.1037/a0022826] [Medline: 21534675]
45. Canham SL, Mauro PM, Kaufmann CN, Sixsmith A. Association of alcohol use and loneliness frequency among middle-aged and older adult drinkers. J Aging Health 2016 Mar;28(2):267-284 [FREE Full text] [doi: 10.1177/0898264315589579] [Medline: 26082130]
46. Hajek A, Kretzler B, König HH. The association between obesity and social isolation as well as loneliness in the adult population: a systematic review. Diabetes Metab Syndr Obes 2021 Jun 17;14:2765-2773 [FREE Full text] [doi: 10.2147/DMSO.S313873] [Medline: 34168476]
47. Cho JH, Olmstead R, Choi H, Carrillo C, Seeman TE, Irwin MR. Associations of objective versus subjective social isolation with sleep disturbance, depression, and fatigue in community-dwelling older adults. Aging Ment Health 2019 Sep;23(9):1130-1138 [FREE Full text] [doi: 10.1080/13607863.2018.1481928] [Medline: 30284454]
48. Gronewold J, Kropp R, Lehmann N, Schmidt B, Weyers S, Siegrist J, Heinz Nixdorf Recall Study Investigative Group. Association of social relationships with incident cardiovascular events and all-cause mortality. Heart 2020 Sep;106(17):1317-1323 [FREE Full text] [doi: 10.1136/heartjnl-2019-316250] [Medline: 32165451]
49. Smith RW, Barnes I, Green J, Reeves GK, Beral V, Floud S. Social isolation and risk of heart disease and stroke: analysis of two large UK prospective studies. Lancet Public Health 2021 Apr;6(4):e232-e239 [FREE Full text] [doi: 10.1016/S2468-2667(20)30291-7] [Medline: 33662329]
50. Bowling A. Social networks and social support among older people and implications for emotional well-being and psychiatric morbidity. Int Rev Psychiatry 1994 Jan;6(1):41-58 [FREE Full text] [doi: 10.3109/09540269409025242]
51. Smith KJ, Gavey S, RIddell NE, Kontari P, Victor C. The association between loneliness, social isolation and inflammation: a systematic review and meta-analysis. Neurosci Biobehav Rev 2020 May;112:519-541. [doi: 10.1016/j.neubiorev.2020.02.002] [Medline: 32092313]
52. Heffner KL, Waring ME, Roberts MB, Eaton CB, Gramling R. Social isolation, C-reactive protein, and coronary heart disease mortality among community-dwelling adults. Soc Sci Med 2011 May;72(9):1482-1488 [FREE Full text] [doi: 10.1016/j.socscimed.2011.03.016] [Medline: 21492978]

53. Luo M, Li L. Social isolation trajectories in midlife and later-life: patterns and associations with health. Int J Geriatr Psychiatry 2022 May;37(5). [doi: 10.1002/gps.5715] [Medline: 35470478]
54. Lay-Yee R, Matthews T, Moffitt T, Poulton R, Caspi A, Milne B. Are trajectories of social isolation from childhood to mid-adulthood associated with adult depression or suicide outcomes. Soc Psychiatry Psychiatr Epidemiol 2023 Mar;58(3):373-382 [FREE Full text] [doi: 10.1007/s00127-022-02389-6] [Medline: 36456781]
55. Gilsanz P, Walter S, Tchetgen Tchetgen EJ, Patton KK, Moon JR, Capistrant BD, et al. Changes in depressive symptoms and incidence of first stroke among middle-aged and older US adults. J Am Heart Assoc 2015 May 13;4(5):e001923 [FREE Full text] [doi: 10.1161/JAHA.115.001923] [Medline: 25971438]
56. Xie W, Zheng F, Yan L, Zhong B. Cognitive decline before and after incident coronary events. J Am Coll Cardiol 2019 Jun 25;73(24):3041-3050 [FREE Full text] [doi: 10.1016/j.jacc.2019.04.019] [Medline: 31221251]
57. Siegmund LA, Distelhorst KS, Bena JF, Morrison SL. Relationships between physical activity, social isolation, and depression among older adults during COVID-19: a path analysis. Geriatr Nurs 2021 Sep;42(5):1240-1244 [FREE Full text] [doi: 10.1016/j.gerinurse.2021.08.012] [Medline: 34507833]
58. Mazo GZ, Fank F, Franco PS, Capanema BD, Pereira FD. Impact of social isolation on physical activity and factors associated with sedentary behavior in older adults during the COVID-19 pandemic. J Aging Phys Act 2022 Feb 01;30(1):148-152. [doi: 10.1123/japa.2020-0456] [Medline: 34388700]
59. Schrempft S, Jackowska M, Hamer M, Steptoe A. Associations between social isolation, loneliness, and objective physical activity in older men and women. BMC Public Health 2019 Jan 16;19(1):74 [FREE Full text] [doi: 10.1186/s12889-019-6424-y] [Medline: 30651092]
60. Dempsey PC, Rowlands AV, Strain T, Zaccardi F, Dawkins N, Razieh C, et al. Physical activity volume, intensity, and incident cardiovascular disease. Eur Heart J 2022 Dec 07;43(46):4789-4800. [doi: 10.1093/eurheartj/ehac613] [Medline: 36302445]
61. Ahmadi MN, Clare PJ, Katzmarzyk PT, Del Pozo Cruz B, Lee IM, Stamatakis E. Vigorous physical activity, incident heart disease, and cancer: how little is enough? Eur Heart J 2022 Dec 07;43(46):4801-4814 [FREE Full text] [doi: 10.1093/eurheartj/ehac572] [Medline: 36302460]
62. China Health and Retirement Longitudinal Study (CHARLS). URL: https://charls.pku.edu.cn/ [accessed 2023-06-26]

Abbreviations

CESD-10: 10-item Center for Epidemiology Scale for Depression
CHARLS: China Health and Retirement Longitudinal Study
CVD: cardiovascular disease
ELSA: English Longitudinal Study of Ageing
HR: hazard ratio
hs-CRP: high-sensitivity C-reactive protein
PPS: probability-proportional-to-size

Edited by A Mavragani, T Sanchez; submitted 12.01.23; peer-reviewed by A Hayixibayi, C Yang; comments to author 12.04.23; revised version received 24.04.23; accepted 08.05.23; published 30.06.23.

Please cite as:
Guo L, Wang W, Shi J, Zheng X, Hua Y, Lu C
Evaluation of Social Isolation Trajectories and Incident Cardiovascular Disease Among Middle-Aged and Older Adults in China: National Cohort Study
JMIR Public Health Surveill 2023;9:e45677
URL: https://publichealth.jmir.org/2023/1/e45677
doi:10.2196/45677
PMID:37389914

Original Paper

Public Preference Heterogeneity and Predicted Uptake Rate of Upper Gastrointestinal Cancer Screening Programs in Rural China: Discrete Choice Experiments and Latent Class Analysis

Ruyue Liu[1,2,3], MPH; Qiuxia Li[3], BM; Yifan Li[4], BS; Wenjian Wei[4], BM; Siqi Ma[4], BS; Jialin Wang[4], BM; Nan Zhang[4], PhD

[1]Center for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[2]National Health Commission Key Lab of Health Economics and Policy Research, Shandong University, Jinan, China
[3]School of Public Health, Weifang Medical University, Weifang, China
[4]Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China

Corresponding Author:
Nan Zhang, PhD
Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences
Shandong First Medical University and Shandong Academy of Medical Sciences
440 Jiyan Road, Huaiyin District
Jinan, 250117
China
Phone: 86 15562637630
Email: nanzhang@vip.126.com

Abstract

Background: Rapid increases in the morbidity and mortality of patients with upper gastrointestinal cancer (UGC) in high-incidence countries in Asia have raised public health concerns. Screening can effectively reduce the incidence and mortality of patients with UGC, but the low population uptake rate seriously affects the screening effect.

Objective: We aimed to determine the characteristics that influence residents' preference heterogeneity for a UGC-screening program and the extent to which these characteristics predict residents' uptake rates.

Methods: A discrete choice experiment was conducted in 1000 residents aged 40-69 years who were randomly selected from 3 counties (Feicheng, Linqu, and Dongchangfu) in Shandong Province, China. Each respondent was repeatedly asked to choose from 9 discrete choice questions of 2 hypothetical screening programs comprising 5 attributes: screening interval, screening technique, regular follow-up for precancerous lesions, mortality reduction, and out-of-pocket costs. The latent class logit model was used to estimate residents' preference heterogeneity for each attribute level, their willingness to pay, and the expected uptake rates.

Results: Of the 1000 residents invited, 926 (92.6%) were included in the final analyses. The mean age was 57.32 (SD 7.22) years. The best model contained 4 classes of respondents (Akaike information criterion=7140.989, Bayesian information criterion=7485.373) defined by different preferences for the 5 attributes. In the 4-class model, out of 926 residents, 88 (9.5%) were assigned to class 1, named as the negative latent type; 216 (3.3%) were assigned to class 2, named as the positive integrated type; 434 (46.9%) were assigned to class 3, named as the positive comfortable type; and 188 (20.3%) were assigned to class 4, named as the neutral quality type. For these 4 latent classes, "out-of-pocket cost" is the most preferred attribute in negative latent type and positive integrated type residents (45.04% vs 66.04% importance weights), whereas "screening technique" is the most preferred factor in positive comfortable type residents (62.56% importance weight) and "screening interval" is the most valued attribute in neutral quality type residents (47.05% importance weight). Besides, residents in different classes had common preference for painless endoscopy, and their willingness to pay were CNY ¥385.369 (US $59.747), CNY ¥93.44 (US $14.486), CNY ¥1946.48 (US $301.810), and CNY ¥3566.60 (US $552.961), respectively. Residents' participation rate could increase by more than 89% (except for the 60.98% in class 2) if the optimal UGC screening option with free, follow-up for precancerous lesions, 45% mortality reduction, screening every year, and painless endoscopy was implemented.

Conclusions: Public preference heterogeneity for UGC screening does exist. Most residents have a positive attitude toward UGC screening, but their preferences vary in selected attributes and levels, except for painless endoscopy. Policy makers should consider these heterogeneities to formulate UGC-screening programs that incorporate the public's needs and preferences to improve participation rates.

(JMIR Public Health Surveill 2023;9:e42898) doi:10.2196/42898

KEYWORDS

upper gastrointestinal cancer; screening programs; discrete choice experiment; latent class logit model; public preference heterogeneity; uptake rate

Introduction

Upper gastrointestinal cancer (UGC; including gastric cancer and esophageal cancer) is one of the most common malignant neoplasms worldwide, and 1.60 million new UGC cases were reported and 1.31 million UGC deaths occurred in 2020, with China alone accounting for more than half of the cases and deaths, respectively [1]. Among patients diagnosed with UGC, the 5-year survival rate is approximately 30% [2]. The main reason for the poor prognosis is that most patients are diagnosed at an advanced stage when tumors are not resectable and nonsurgical therapeutic modalities are ineffective [3]. Nevertheless, UGC can be detected and treated with endoscopic screening at an earlier stage, with a survival rate of >90% [4]. Endoscopic screening, the gold standard for the early diagnosis of UGC, has been widely adopted in many countries and proven to be highly effective in reducing the morbidity and mortality rates of UGC [5,6]. Furthermore, recent studies have confirmed that endoscopic screening is a more cost-effective screening method for UGC than no screening [7-9]. However, the effectiveness of endoscopic screening is reduced by the low population uptake rate.

In some high-income countries, such as Japan and South Korea, the participation rate is lower than 50% despite the launch of nationwide endoscopic screening programs for gastric cancer for a long time [10,11]. Similarly, as a low-income country with high UGC incidence and mortality, China has performed endoscopic screening for patients with gastric cancer and esophageal cancer in more than 110 high-risk areas throughout the country since 2005, but residents' compliance to date was still only 48.62% [6]. This low uptake rate for UGC screening has become a huge public health challenge and needs to be addressed to maximize the benefits that can be achieved by endoscopic screening. Recent studies have shown that the characteristics (attributes and levels) of the UGC-screening program may impact whether individuals participate in the screening [12-16]. Nevertheless, if and how these characteristics affect residents' screening preferences remains poorly understood.

The discrete choice experiment (DCE) is a stated preference method that has been used from 1990s to obtain individuals' preferences for health care and for a wide range of health care topics [17]. In a DCE, participants are presented with alternative options that are systematically described according to several attributes and are asked to make a choice among these options [18]. Their choice can be analyzed using a discrete choice model. The latent class logit (LCL) model is the most commonly used choice model for exploring individuals' preference heterogeneity, which permits the potential classification of the target population and estimation of their preference heterogeneity in different classes [19]. However, to date, no published studies have used LCL to explore individuals' UGC-screening preference heterogeneity.

Therefore, in this study, we constructed a DCE and used the LCL model to elicit public preference heterogeneity regarding UGC-screening programs and explore the influence of personal characteristics on the variation, willingness to pay (WTP), and uptake rate in their choices. These results will be helpful for policy makers in understanding the heterogeneity of public preferences, resulting in a sustainable and effective UGC-screening policy.

Methods

Sampling and Participants

This is an extension of our previously published study [12], which also used contemporaneous data but where the new methods are used to obtain more essential and meaningful findings and conclusions. In our study, a stratified (cluster) random sampling method was used to select participants. First, located in the east, central, and west of Shandong Province, 3 cities (Weifang, Taian, and Liaocheng) were chosen as sample areas, which represent high, medium, and low economic levels according to the gross domestic product (GDP) per capita (2020), respectively. Then, 3 counties (Linqu, Feicheng, and Dongchangfu) were selected from each sample area as the survey points, respectively. Finally, rural residents aged 40 to 69 years in 2 to 4 villages were randomly chosen from each county for a face-to-face DCE questionnaire survey.

DCE Design

Overview

In this study, a DCE was conducted to determine the factors (attributes and attribute levels) that influence residents' preference heterogeneity to participate in a UGC-screening program and the weightage that residents give to the selected attributes and attribute levels. The DCE was developed according to the methodological standards issued by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) [20]. The design process is as follows. More details on the DCE design process, sample size calculation, and data collection are available in our previously published study [12].

Selection of Attributes and Attribute Levels

The attributes and attribute levels of the DCE were derived from a literature review [5,12-14,21-29], expert interviews, and focus group discussions with residents of the target population.

Finally, the survey results described 5 attributes with 2 to 4 levels: out-of-pocket costs (CNY ¥0 [US $0], CNY ¥100 [US $15.503], CNY ¥300 [US $46.511], and CNY ¥500 [US $77.519]), screening interval (every year, every 2 years, every 5 years, and once in a lifetime), screening technique (endoscopy and painless endoscopy), mortality reduction (15%, 30%, 45%, and 60%), and regular follow-up for precancerous lesions (yes or no). Notably, the "mortality reduction" attribute in this study refers to the respondents' expected reduction extent of mortality because of UGC if they chose to participate in UGC screening.

Questionnaire Design

SAS (version 9.4) software was used to design the questionnaire. The combination of 5 attributes with 2 to 4 levels yielded 256 (ie, $4^3 \times 2^2$) possible scenarios using full factor design [30]. To minimize the respondents' burden, a smaller fractional factorial design was selected according to the D-efficiency criteria. The final design randomly generated 16 choice sets into 2 blocks with 8 choice tasks. For each choice task, 2 screening programs (options A and B) and an opt-out option (whether to choose to be screened in real life) were included. In addition, a rationality test (choice set 1=choice set 5) was included in the DCE to investigate residents' understanding of the questionnaire [31]. Table 1 presents an example of a discrete choice task.

The questionnaire also contained sociodemographic questions, including age, sex (male and female), educational level, and medical insurance. As a result, the questionnaire had 2 sections: the first section comprised sociodemographic questions and the second comprised information about the 9 DCE choice sets (Multimedia Appendices 1-4).

Table 1. Example of a discrete choice task[a].

Attributes	Option A	Option B
Out-of-pocket costs (CNY ¥ or US $)[b]	100 or 15.503	300 or 46.511
Screening interval	Every 2 years	Every year
Regular follow-up for precancerous lesions	Yes	No
Mortality reduction (%)	15	45
Screening technique	Endoscopy	Painless endoscopy
Which of these options would you prefer?	Option A	Option B
Would you screen as you choose in real life?	Yes	No

[a] See the study by Liu et al [12] for more details.

[b] CNY ¥6.45=US $1.00.

Pilot Study

Before the main study, we conducted a pilot test with residents of the target population to ascertain whether respondents could manage the length of the questionnaire and examine the validity, acceptability, and intelligibility of the questionnaire.

In total, 30 residents were recruited for the pilot test and interviewed by 8 trained investigators. Respondents were provided with descriptions of the attributes and levels and were given the opportunity to comment on the questionnaire design and layout. They found the questionnaire acceptable and easy to complete. No major changes were made to the questionnaires.

Sample Size and Data Collection

Calculating the optimal DCE sample size is complicated by the fact that it depends on the question format, complexity of the choice tasks, degree of heterogeneity in the target population, and need to conduct subgroup analyses [21,32]. However, there is a generally accepted rule of thumb proposed by Orem and Johnson [33,34] for calculating sample size:

The required sample size depends on the largest number of levels for any 1 attribute (*c*), number of alternatives per choice task (*a*), and number of choice tasks (*t*). Therefore, this questionnaire required at least 125 respondents [(500 × 4) / (8 × 2)] to estimate the main effect alone. Because 2 blocks were included in the design, there was a need for at least 300 respondents (125 × 2). Other studies found that the sample size of foreign DCE studies in the field of health care ranged from 150 to 1200, and 77% of the studies had a sample size 600, whereas most of the studies in China had a sample size below 600 [30,35,36]. Therefore, we combined the abovementioned literature reviews, the need for LCL model analysis, and an expected response rate of 20% to determine a total sample size of 1000 residents.

The final survey was conducted from April 1, 2021, to May 31, 2021. A total of 1000 permanent residents of the sample area were invited to participate in the survey, and they were assisted in completing the questionnaire by 8 trained investigators. A 2-step survey was conducted to interview the residents. In the first step, respondents were asked to describe their basic information, such as age, sex, and health insurance. In the second step, they were repeatedly asked to choose between 2 alternative options from the per-choice task (a total of 9 choice tasks).

Statistical Analysis

Overview

Statistical analyses were performed using Stata 16.0 software, with the commands *lclogit2* and *lclogitml2* for the analysis of the preference data. All attributes other than cost were treated as classification variables coded by dummy variables, whereas the cost parameter was modeled as a continuous variable to allow for WTP estimation. Sociodemographic data were summarized using descriptive statistics.

LCL Model and Model Fitting

The LCL model is the most commonly used preference heterogeneity model, assuming that preferences differ across respondents and that respondents can be grouped probabilistically according to distinct preference classes, each corresponding to a unique program preference [37,38]. It is based on random utility theory [39]:

$$U = ASC + \beta_1 \text{ (screening interval: every year)} + \beta_2 \text{ (screening interval: every 2 years)} + \beta_3 \text{ (screening interval: every 5 years)} + \beta_4 \text{ (follow-up: yes)} + \beta_5 \text{ (follow-up: yes)} + \beta_6 \text{ (mortality reduction: 30\%)} + \beta_7 \text{ (mortality reduction: 45\%)} + \beta_8 \text{ (mortality reduction: 60\%)} + \beta_9 \text{ (screening technique: endoscopy screening)} + \beta_{10} \text{ (screening technique: painless endoscopy screening)} + \beta_{11} \text{ (out-of-pocket costs)} + \varepsilon$$

where U refers to the utility of the respondents’ (in particular, latent class) choice of option across the different choice sets in the formal investigation stage; ASC (alternative specific constant) represents a fixed constant term (reflecting the screening attitudes of respondents in different classes), which was used to capture unobservable influences beyond attributes present in the choice sets; β provides quantitative information on the strength of preference for each attribute level; and ε is an unobservable random component. A positive β coefficient indicates that the attribute level combination is preferred relative to the reference level, whereas a negative β coefficient indicates the opposite. The higher the absolute value of β coefficient, the stronger is the preference for that level relative to the reference level. Then, using the expectation-maximization algorithm, we constructed an LCL model with different numbers of classes. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to compare the model fit, with lower AIC and BIC values indicating better models [40].

Parameter Estimation

Importance weights are a measure of the importance of an attribute relative to other attributes in the model for an individual’s preferences [41]. It is calculated by dividing the maximum utility of an attribute by the total utility of all the attributes.

WTP refers to the fee that residents of different classes were willing to spend or want to be compensated for the change in a certain attribute level in a specific program. To calculate the respondents’ WTP, the estimation of cost attributes was used as a measure of the marginal utility of money. The ratio of the value of the coefficient of other attribute levels to the negative of the cost attribute was calculated to elicit respondents’ WTP for contribution to a UGC-screening program.

From a policy perspective, the uptake rate is a useful output for estimating the effect of policies yet to be implemented, such as the change in the participation rate of rural residents with increasing screening costs [39]. The logit probability of choosing alternative *i* rather than alternative *j* is given by the following equation where *x* is a vector of the attribute coefficients:

With the level of attribute improving from *k* (baseline reference level) to *g*, the change in the uptake rate when choosing the baseline program is given by the following equation:

Ethics Approval

The Institutional Ethical Review Board of Shandong Cancer Hospital and Institute approved this study (reference no. SDTHEC201909001).

Results

Study Participation

A total of 959 of 1000 invited residents completed the questionnaire (for a response rate of 95.9%), of which 33 residents failed the consistency test. The sensitivity analysis indicated that there were no substantial differences in preference between the full sample and those who passed the consistency test. However, from the model fitting results, the likelihood, AIC, and BIC values of model 2 (removing samples of residents who did not pass the consistency test) were smaller than those of model 1, indicating that model 2 has a better fit (Multimedia Appendix 5). Considering the accuracy of the results, the 926 respondents who passed the consistency test were finally included in the preference estimation using the LCL model. Table 2 presents the characteristics of the 926 respondents. In all, 44.7% (414/926) of respondents were aged 50 to 59 years, and 95.9% (888/926) had a partner. Most respondents (743/926, 80.2%) had no family history of cancer, and 56.2% (520/926) had experience in cancer screening.

Table 2. Characteristics of study participants[a].

Characteristics	Respondents (n=926)
Age (years), mean (SD)	57.32 (7.22)
Sex, n (%)	
Male	315 (34)
Female	611 (66)
Age group (years), n (%)	
40-49	139 (15)
50-59	414 (44.7)
60-70	373 (40.3)
Marital status[b], n (%)	
With a partner	888 (95.9)
Without a partner	38 (4.1)
Annual family income (CNY ¥>)[c], n (%)	
<10,000	434 (46.9)
10,000-29,999	294 (31.7)
≥30,000	198 (21.4)
Location, n (%)	
Linqu	322 (34.8)
Feicheng	310 (33.5)
Dongchangfu	294 (31.7)
Family history of cancer[d], n (%)	
Yes	183 (19.8)
No	743 (80.2)
Screening for cancer, n (%)	
Ever	520 (56.2)
Never	406 (43.8)

[a] For more details, see the study by Liu et al [12].

[b] Marital status: with a partner, reflecting that the individual is married and the spouse is alive; and without a partner, including single, divorced, and widowed.

[c] The per capita gross domestic product in 2020 in Linqu, Feicheng, and Dongchangfu was CNY ¥39,910, CNY ¥80,696, and CNY ¥50,726, respectively. The average exchange rate between US $ and CNY ¥ in 2021 was US $1=CNY ¥6.45.

[d] History of cancer in blood relatives, including parents, grandparents, siblings, uncles, aunts, and cousins.

Model Fitting Results

To select the appropriate number of classes, different numbers of classes were tested, from 2 to 5. The AIC and BIC values were used to select the final number of clusters that best fit the data. As presented in Table 3, there is a gradual decline in the values of AIC and BIC with an increase in the latent class number. When the number of latent classes is 4, the values reach a minimum (AIC=7140.989; BIC=7485.373) and subsequently increase. Accordingly, the 4-class model has the best fit for the data. The respondents’ preference heterogeneity was analyzed by constructing an LCL model with 4 classes.

Table 3. Results of model fitting under different classes.

Class	Observation, n	*df*	LL[a]	AIC[b]	BIC[c]
2	22,224	21	−4139.361	8320.722	8488.910
3	22,224	32	−3728.246	7520.492	7776.777
4	22,224	43	−3527.495	7140.989	7485.373
5	22,224	54	−3533.606	7175.213	7607.695

[a] LL: likelihood.

[b] AIC: Akaike information criterion.

[c] BIC: Bayesian information criterion.

Preferences Estimation

Table 4 shows that respondents' preferences varied at each attribute level, except for painless endoscopy. The detailed preference differences and the naming of different classes are as follows.

The residents (88/926, 9.5%) in class 1 responded negatively for UGC screening (β=3.372; *P*<.001), but they preferred lower screening costs. Theoretically, the respondents' negative screening attitudes can be changed once the cost of screening programs is reduced, so they were defined as the negative latent type (NLT).

Table 4. Parameters estimation results of latent class logit model.

Attributes and levels	Class 1 (n=88; class share=0.095)		Class 2 (n=216; class share=0.233)		Class 3 (n=434; class share=0.469)		Class 4 (n=188; class share=0.203)	
	β coefficient[a]	*P* value[b]	β coefficient	*P* value	β coefficient	*P* value	β coefficient	*P* value
ASC[c] (opt-out[d])	3.372	*<.001*	−3.843	*<.001*	−2.165	*<.001*	.076	.85
Screening interval (once in a lifetime[e])								
Every year	.514	.19	.432	*.003*	−.280	.33	2.294	*<.001*
Every 2 years	.298	.37	.015	.92	.141	.72	2.336	*<.001*
Every 5 years	.764	*.03*	.359	*.02*	.286	.37	1.277	*<.001*
Follow-up (yes[e])								
No	.396	*.01*	−.274	*<.001*	−.535	*.01*	−.247	*.02*
Morality reduction (15%[e])								
30%	−.161	.60	−.039	.82	.605	.13	.503	.007
45%	−.477	.13	.366	*.03*	.147	.67	.867	*<.001*
60%	−.285	.37	−.208	*.02*	.124	.69	1.142	*<.001*
Screening technique (endoscopy[e])								
Painless endoscopy	2.804	*<.001*	.612	*<.001*	4.173	*<.001*	1.088	*<.001*
Out-of-pocket cost	−.007	*<.001*	−.007	*<.001*	−.002	.004	−.0003	.32

[a] β coefficient: it reflects the values of each attribute level and the horizontal regression coefficient.

[b] Italicized *P* values denote a significance level at *P*<.05.

[c] ASC: alternative specific constant.

[d] Specific constant item for opt-out.

[e] Reference, reflecting the reference level for each attribute.

The residents (216/926, 23.3%) in class 2 responded most positively for UGC screening (β=−3.843; *P*<.001), and they preferred the program with every year, follow-up, 45% mortality reduction, and painless endoscopy, indicating that they had equal consideration in each attribute, so they were defined as the positive integrated type (PIT).

The residents (434/926, 46.8%) in class 3 responded positively for UGC screening (β=−2.165; *P*<.001), and they preferred follow-up, painless endoscopy, and lower out-of-pocket cost attribute levels over the screening interval and mortality reduction attributes. This group of people paid more attention to the comfort and experience of the screening process; therefore, they were defined as a positive comfortable type (PCT).

The residents (188/926, 20.3%) in class 4 had a neutral attitude toward UGC screening (β=.076; *P*>.05), and they preferred screening programs with shorter screening intervals, painlessness, follow-up, and a much lower risk of death regardless of the costs, indicating that they place more value on the quality and effect of the screening program, so they were defined as the neutral quality type (NQT).

Participants' WTP

Table 5 shows that NLT residents were willing to spend CNY ¥385.369 (US $59.747) to improve screening technique, that is, to upgrade endoscopy to painless endoscopy; PIT residents had more WTP for screening every year (CNY ¥65.966 [US $10.227]) compared with once in a lifetime, and they should be compensated CNY ¥41.776 (US $6.477) if there is no regular follow-up for precancerous lesions; PCT residents had the highest WTP for painless endoscopy (CNY ¥1946.482 [US $301.780]), and the participants should be compensated CNY ¥249.753 (US $ 38.721) if there is no follow-up after screening in theory. For the NQT population, there was no statistically substantial difference in costs attribute.

Table 5. Results of participants' willingness to pay (WTP).

Attributes and levels	Class 1 (n=88)		Class 2 (n=216)		Class 3 (n=434)		Class 4 (n=188)	
	WTP	*P* value[a]	WTP	*P* value	WTP	*P* value	WTP	*P* value
Screening interval (once in a lifetime[b])								
Every year	70.622	.19	65.97	*<.001* [b]	−130.71	.35	7524.14	.33
Every 2 years	40.950	.39	2.35	.91	65.56	.74	7660.38	.34
Every 5 years	104.981	*.051*	54.73	*.02*	133.22	.33	4187.49	.33
Follow-up (yes[b])								
No	54.381	.09	−41.78	*<.001*	−249.75	*.02*	−810.56	.38
Morality reduction (15%[b])								
30%	−22.088	.06	−5.98	.81	281.99	.18	1649.85	.37
45%	−65.495	.14	55.93	.02	68.79	.67	2841.79	.36
60%	−39.177	.38	−31.78	.13	57.63	.71	3744.69	.35
Screening technique (endoscopy[b])								
Painless endoscopy	385.369	*<.001*	93.44	*<.001*	1946.48	*<.001*	3566.60	.34

[a] Italicized *P* values denote a significance level at *P*<.05.

[b] Reference, reflecting the reference level in each attribute; WTP reflects residents' willingness to pay for a certain screening program.

Personal Characteristics Influencing Factors

Sociodemographic characteristics influencing factors are presented in Table 6. Class 4 was automatically identified as the reference class in the LCL model. Accordingly, the residents living in Linqu were included in class 1 easily, whereas male respondents and those with an experience of UGC screening were more likely to be excluded from class 1. Residents who had a family history of cancer were more likely to be assigned to class 2 than those without a family history. Compared with female respondents, male respondents were less likely to be included in class 3, and residents aged 60 to 69 years were more likely to be included in class 3.

Table 6. Personal characteristics influencing factors.

Attributes and levels	Class 1		Class 2		Class 3	
	β (SE)	*P* value[a]	β (SE)	*P* value	β (SE)	*P* value
Location (Feicheng[b])						
Linqu	2.083 (0.386)	*<.001*	−.008 (0.325)	.98	.883 (0.269)	*<.001*
Dongchangfu	−.575 (0.518)	.27	.181 (0.319)	.57	.530 (0.279)	.06
Annual family income[c] (¥<10,000[b])						
10,000-29,999	−.143 (0.427)	.74	−.392 (0.342)	.25	0.067 (0.291)	.81
≥30,000	.371 (0.383)	.33	.409 (0.318)	.19	0.428 (0.283)	.13
Sex (female[b])						
Male	−1.095 (0.356)	*.002*	−.094 (0.271)	.73	−.608 (0.230)	.008
Age group (years; 40-49[b])						
60-69	1.244 (0.492)	*.01*	.951 (0.406)	*.01*	.844 (0.315)	.007
50-59	0.592 (0.445)	.18	.640 (0.361)	.07	.338 (0.270)	.21
Education level (university and above[b])						
Below junior high school	−2.341 (1.003)	*.02*	.327 (1.266)	.79	−.496 (0.791)	.53
Junior high and high school	−2.737 (0.957)	.004	.124 (1.237)	.92	−.802 (0.762)	.29
Family history of cancer[d](no[b])						
Yes	−.647 (0.363)	.08	−.769 (0.303)	.01	−.396 (0.238)	.09
Screening for cancer[e] (never[b])						
Ever	−.657 (0.310)	.03	−.217 (0.272)	.42	−.144 (0.223)	.52
Constant	.801 (1.114)	.47	−.535 (1.348)	.69	.722 (0.869)	.41

[a] Italicized *P* values denote a significance level at *P*<.05.

[b] Reference, reflecting the reference level.

[c] The average exchange rate between US $ and CNY ¥ in 2021 was US $1=CNY ¥6.45.

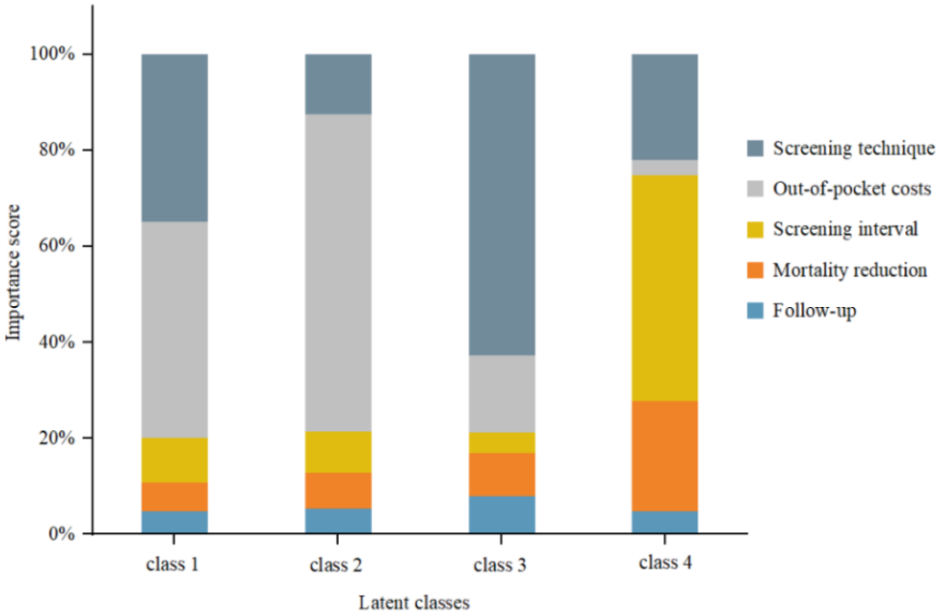
[d] Family history of cancer, including parents, grandparents, siblings, uncles, aunts, and cousins having a history of cancer.

[e] Screening for cancer means attending endoscopy screening at any time in the past.

Relative Importance of Attributes

As presented in Figure 1, we calculated the importance weights for the 5 attributes in different classes. The out-of-pocket cost attribute was the factor that most influenced class 1 and class 2 residents' willingness to participate in UGC screening (45.04% and 66.04% importance weights, respectively). The screening technique was the most important for class 3 respondents (62.56% importance weight). The screening interval was the most important for class 4 residents (47.05% importance weight) and the out-of-pocket cost was not important (3.07% importance weight).

Figure 1. Relative importance of attributes in different classes. Relative scores expressed in percent as weighted out of 100.

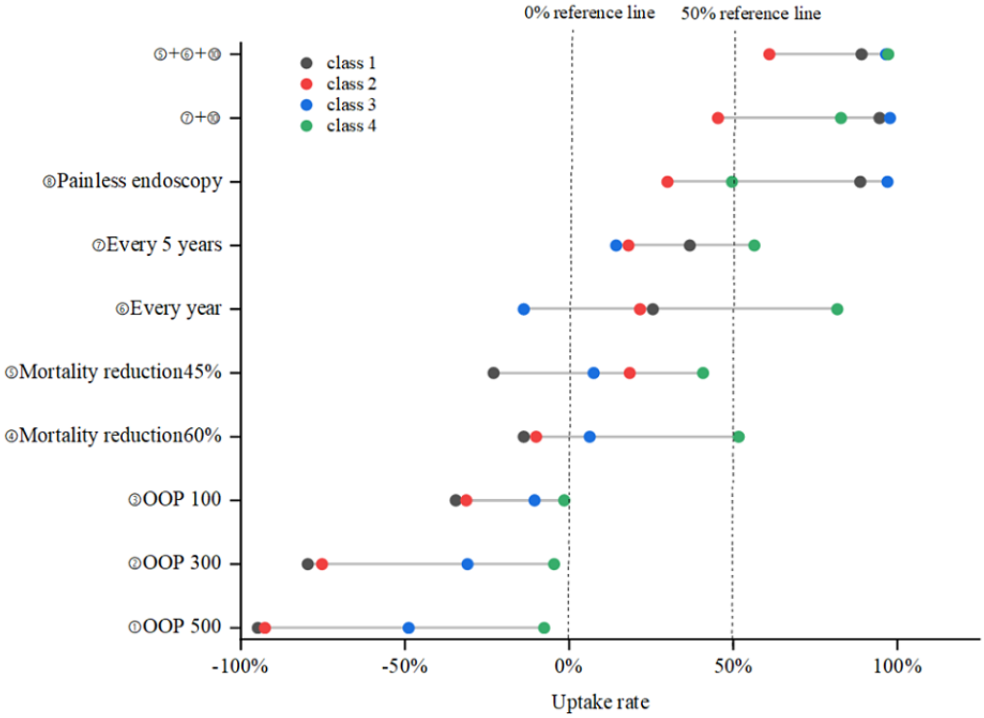


Uptake Rate

Figure 2 depicts the change in choice probability for different classes of respondents, as attributes and levels were changed at the baseline level (baseline: once in a lifetime, follow-up, 15% mortality reduction, endoscopy, and CNY ¥0). The screening uptake rate of residents in classes 1, 2, and 3 dramatically decreased by ≥48% (94.84%, 92.67%, and 48.91%, respectively) when the out-of-pocket costs increased to CNY ¥500 (US $71.903), whereas the uptake rate in class 4 remained constant.

For class 1 and class 3 residents, the participation rate was strongly driven by painless endoscopy, with an increase of 88.66% and 96.97%, respectively. Among residents in class 4, the participation rate would increase by 81.69% if the screening interval was improved from once in a lifetime to every year. The expected uptake rate for the best UGC-screening program (CNY ¥0, 45% mortality reduction, every year screening, and painless endoscopy) would increase by 89% and above in classes 1, 3, and 4, and class 2 residents' participation rate would remain ≥60%.

Figure 2. Uptake rate of residents in different classes. Null percent and 50% reference line represents the reference in uptake rate as 0% and 50%, respectively. The horizontal comparison between points can see the change of the participation rate of residents in different classes under the same attribute level, and the longitudinal comparison between points can show the change of participation rate of individuals in the same class under the change of different attribute levels. OOP: out-of-pocket costs.



Discussion

Principal Findings

To the best of our knowledge, this is the first study to explore public preference heterogeneity for a UGC-screening program by classifying the population. This study is a continuation and extension of our previous study [12]. We used novel analytic methods to explore population preference heterogeneity and found that preference heterogeneity exists when residents choose to participate in UGC-screening programs. Preference heterogeneity was explorable: the target population was divided into 4 categories, including NLT, PIT, PCT, and NQT. Residents in different classes had different preferences for the selected attributes and attribute levels, except for the screening technique attribute (preferably painless endoscopy). On the basis of this, we found that the optimal screening strategy with free, 45% mortality reduction, screening every year, and painless endoscopy could ensure that the screening participation rate of class 1, class 3, and class 4 residents reached ≥89%, and the screening uptake rate of residents in class 2 could reach 60.98%.

Nearly half (434/926, 46.8%) of the respondents were divided into the PCT group, indicating that they paid more attention to their experience and comfort levels regarding the screening procedure rather than the outcome. This is consistent with previous findings, in which the author found that residents' screening participation rate was significantly associated with severe pain and endoscopy experience [13]. Similarly, a DCE study of immediate family members with UGC indicated that participants preferred a pain-none program, increasing the participation rate by 58.85% compared with the reference level (pain-mild) [42]. Further analysis revealed that the uptake rate of residents of different classes was mostly driven by painless endoscopy. Accordingly, if painless endoscopic screening technology can be widely promoted, it will be able to satisfy the screening needs and preferences of residents in different categories, greatly improving population screening enthusiasm and participation rate. Areas with a good economic level and sufficient screening resources should take painless endoscopy as the essential element, paying more attention to the population screening experience.

Compared with other classes, NQT residents valued the screening quality and outcome most, such as screening interval and mortality reduction attributes (47.05% and 23% importance weights, respectively), and their screening compliance increased once the accuracy of endoscopy was improved. An unlabeled DCE assessing the attributes of an optimal esophageal adenocarcinoma screening test found that test accuracy generally outweighs the importance of potential pain and discomfort [13]. Consistent with our results, various DCEs evaluating preferences for breast, cervical, and colorectal cancer screening reported that attributes related to test accuracy were more important than attributes related to the screening procedure [43-45]. Although these studies targeted different cancer types and populations, their results provide face validity for the results of this study. This study provides an important reference for the clinical development of scientific and effective screening. Improving the accuracy of screening techniques and health care services is the main way to satisfy residents' growing demand for health care services and to increase their participation rate in a UGC screening program.

All residents, except those in class 4, strongly preferred the UGC-screening program with lower costs, particularly for the residents in class 1, and their negative attitude will be theoretically improved if the screening fee is reduced. These results are consistent with a recent South Australian study in which the uptake rate of participants decreased by 48% after the screening fee was changed from A$0 to A$500 (US $0 to US $328.62) [24]. Li et al [42] also pointed out that the probability of screening program selection increased by 11.30% when the cost changed from CNY ¥600 (US $86.28) to CNY ¥200 (US $28.76). Other studies regarding lung cancer screening and colorectal cancer screening found that cost-related attributes had an important impact on the residents' participation rate, and they all preferred the screening program with a lower cost [18,46]. However, it is unrealistic to provide a nationwide free UGC-screening strategy because of the large population size and limited medical resources. These results provide support and guidance for policy makers to explore the multifinancing and cost sharing of different interventions.

Only 9.5% of the respondents had a negative response for UGC screening, which is consistent with the screening status in rural China. These residents did not proactively seek medical attention or predisease screening influenced by their family income and the traditional concept of medical treatment. Other studies have also pointed out that some residents do not think it is necessary to screen before the onset of noticeable symptoms [47]. Therefore, policy makers should strengthen primary prevention and emphasize the importance of screening for cancer to promote more high-risk groups to take the initiative to screen, reduce their morbidity and mortality, and reduce the socioeconomic burden.

Differences in health care delivery systems across countries may affect public-related medical needs and preferences. A label DCE study conducted in the Netherlands indicated that respondents preferred video capsule over saliva swab and least preferred endoscopy, but a study performed in China showed that respondents have a strong preference for endoscopy [12,22]. In addition, we found that the population in different countries has a common preference for out-of-pocket cost attributes, preferring free or less-expensive screening programs [13,42,48,49]. Notably, in this study, we did not investigate the preference differences between health care systems in China and other countries. Accordingly, caution should be exercised when applying our results to other countries with different health care systems.

Limitations

This study has some limitations that should be considered. First, DCE as a stated preference method differs from revealed preferences (ie, a difference between what people say they will choose and what they actually choose) [50]. The revealed preferences should be examined after implementing the UGC-screening programs. Second, the development of UGC-screening programs depends not only on the population uptake rate but also on the cost-effectiveness of those programs.

In this study, we mainly analyzed the residents' screening uptake rate. However, other studies have reported that UGC screening is always more cost-effective than not screening [7,8]. Third, the respondents were selected from Shandong Province and interviewed during a specific period; thus, the results need to be carefully applied to all rural residents in China and other resource-limited countries.

Conclusions

Most rural residents have a positive attitude toward UGC screening, but they expressed preference heterogeneity in selected attributes and levels, except for painless endoscopy. The optimal UGC-screening program that is free with 45% mortality reduction, screening every year, and painless endoscopy should be implemented to maximize participation rates if resources permit. In areas with limited health resources, the population uptake rate can be controlled by adjusting steerable screening attributes such as screening intervals and out-of-pocket costs. In addition, differentiated screening programs suitable for residents belonging to different categories should be developed within certain regions. These findings will provide input for the design of a UGC-screening strategy that incorporates the public's preference heterogeneity to improve their participation rate and provide practical evidence for other resource-limited countries with high UGC incidence and mortality.

Acknowledgments

The authors wish to thank all the residents, collaborators, and graduate students who participated in the field survey and discrete choice experiment study, as well as the experts and work staff in the sample regions who assisted with participant recruitment and data collection.

RL was affiliated with the School of Public Health, Weifang Medical University at the time of the study and is currently affiliated with the Center for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University and the National Health Commission Key Lab of Health Economics and Policy Research, Shandong University.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Discrete choice experiment (DCE) questionnaires (block A) in Chinese.
[DOCX File , 34 KB - publichealth_v9i1e42898_app1.docx]

Multimedia Appendix 2
Discrete choice experiment (DCE) questionnaires (block A) in English.
[DOCX File , 35 KB - publichealth_v9i1e42898_app2.docx]

Multimedia Appendix 3
Discrete choice experiment (DCE) questionnaires (block B) in Chinese.
[DOCX File , 34 KB - publichealth_v9i1e42898_app3.docx]

Multimedia Appendix 4
Discrete choice experiment (DCE) questionnaires (block B) in English.
[DOCX File , 36 KB - publichealth_v9i1e42898_app4.docx]

Multimedia Appendix 5
Sensitivity analysis.
[DOCX File , 19 KB - publichealth_v9i1e42898_app5.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]

2. Zeng H, Chen W, Zheng R, Zhang S, Ji JS, Zou X, et al. Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. Lancet Glob Health 2018 May;6(5):e555-e567 [FREE Full text] [doi: 10.1016/S2214-109X(18)30127-X] [Medline: 29653628]
3. Liu M, He Z, Guo C, Xu R, Li F, Ning T, et al. Effectiveness of intensive endoscopic screening for esophageal cancer in China: a community-based study. Am J Epidemiol 2019 Apr 01;188(4):776-784. [doi: 10.1093/aje/kwy291] [Medline: 30608546]
4. Wang GQ, Abnet CC, Shen Q, Lewin KJ, Sun XD, Roth MJ, et al. Histological precursors of oesophageal squamous cell carcinoma: results from a 13 year prospective follow up study in a high risk population. Gut 2005 Feb;54(2):187-192 [FREE Full text] [doi: 10.1136/gut.2004.046631] [Medline: 15647178]
5. Chen R, Liu Y, Song G, Li B, Zhao D, Hua Z, et al. Effectiveness of one-time endoscopic screening programme in prevention of upper gastrointestinal cancer in China: a multicentre population-based cohort study. Gut 2021 Feb;70(2):251-260 [FREE Full text] [doi: 10.1136/gutjnl-2019-320200] [Medline: 32241902]
6. Wei WQ, Chen ZF, He YT, Feng H, Hou J, Lin DM, et al. Long-term follow-up of a community assignment, one-time endoscopic screening study of esophageal cancer in China. J Clin Oncol 2015 Jun 10;33(17):1951-1957 [FREE Full text] [doi: 10.1200/JCO.2014.58.0423] [Medline: 25940715]
7. Xia R, Li H, Shi J, Liu W, Cao M, Sun D, et al. Cost-effectiveness of risk-stratified endoscopic screening for esophageal cancer in high-risk areas of China: a modeling study. Gastrointest Endosc 2022 Feb;95(2):225-35.e20. [doi: 10.1016/j.gie.2021.08.008] [Medline: 34418461]
8. Xia R, Zeng H, Liu W, Xie L, Shen M, Li P, et al. Estimated cost-effectiveness of endoscopic screening for upper gastrointestinal tract cancer in high-risk areas in China. JAMA Netw Open 2021 Aug 02;4(8):e2121403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.21403] [Medline: 34402889]
9. Yang J, Wei WQ, Niu J, Liu ZC, Yang CX, Qiao YL. Cost-benefit analysis of esophageal cancer endoscopic screening in high-risk areas of China. World J Gastroenterol 2012 May 28;18(20):2493-2501 [FREE Full text] [doi: 10.3748/wjg.v18.i20.2493] [Medline: 22654446]
10. Kim Y, Jun JK, Choi KS, Lee HY, Park EC. Overview of the national cancer screening programme and the cancer screening status in Korea. Asian Pac J Cancer Prev 2011;12(3):725-730 [FREE Full text] [Medline: 21627372]
11. Hamashima C. Current issues and future perspectives of gastric cancer screening. World J Gastroenterol 2014 Oct 14;20(38):13767-13774 [FREE Full text] [doi: 10.3748/wjg.v20.i38.13767] [Medline: 25320514]
12. Liu R, Lu Y, Li Y, Wei W, Sun C, Zhang Q, et al. Preference for endoscopic screening of upper gastrointestinal cancer among Chinese rural residents: a discrete choice experiment. Front Oncol 2022 Jul 27;12:917622 [FREE Full text] [doi: 10.3389/fonc.2022.917622] [Medline: 35965546]
13. Peters Y, van Grinsven E, van de Haterd M, van Lankveld D, Verbakel J, Siersema PD. Individuals' preferences for esophageal cancer screening: a discrete choice experiment. Value Health 2020 Aug;23(8):1087-1095 [FREE Full text] [doi: 10.1016/j.jval.2020.03.013] [Medline: 32828222]
14. Gupta M, Beebe TJ, Dunagan KT, Schleck CD, Zinsmeister AR, Talley NJ, et al. Screening for Barrett's esophagus: results from a population-based survey. Dig Dis Sci 2014 Aug;59(8):1831-1850 [FREE Full text] [doi: 10.1007/s10620-014-3092-8] [Medline: 24652109]
15. Chang JY, Talley NJ, Locke 3rd GR, Katzka DA, Schleck CD, Zinsmeister AR, et al. Population screening for Barrett esophagus: a prospective randomized pilot study. Mayo Clin Proc 2011 Dec;86(12):1174-1180 [FREE Full text] [doi: 10.4065/mcp.2011.0396] [Medline: 22134936]
16. Sijben J, Peters Y, van der Velden K, Rainey L, Siersema PD, Broeders MJ. Public acceptance and uptake of oesophageal adenocarcinoma screening strategies: a mixed-methods systematic review. EClinicalMedicine 2022 Apr 04;46:101367 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101367] [Medline: 35399814]
17. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl Health Econ Health Policy 2003;2(1):55-64. [Medline: 14619274]
18. Cheng J, Pullenayegum E, Marshall DA, Marshall JK, Thabane L. An empirical comparison of methods for analyzing correlated data from a discrete choice survey to elicit patient preference for colorectal cancer screening. BMC Med Res Methodol 2012 Feb 20;12:15 [FREE Full text] [doi: 10.1186/1471-2288-12-15] [Medline: 22348526]
19. Kongsted A, Nielsen AM. Latent class analysis in health research. J Physiother 2017 Jan;63(1):55-58 [FREE Full text] [doi: 10.1016/j.jphys.2016.05.018] [Medline: 27914733]
20. Bridges JF, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR good research practices for conjoint analysis task force. Value Health 2011 Jun;14(4):403-413 [FREE Full text] [doi: 10.1016/j.jval.2010.11.013] [Medline: 21669364]
21. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. Value Health 2013 Jan;16(1):3-13 [FREE Full text] [doi: 10.1016/j.jval.2012.08.2223] [Medline: 23337210]
22. Peters Y, Siersema PD. Public preferences and predicted uptake for esophageal cancer screening strategies: a labeled discrete choice experiment. Clin Transl Gastroenterol 2020 Nov;11(11):e00260 [FREE Full text] [doi: 10.14309/ctg.0000000000000260] [Medline: 33105164]

23. Arnold M, Laversanne M, Brown LM, Devesa SS, Bray F. Predicting the future burden of esophageal cancer by histological subtype: international trends in incidence up to 2030. Am J Gastroenterol 2017 Aug;112(8):1247-1255. [doi: 10.1038/ajg.2017.155] [Medline: 28585555]
24. Bulamu NB, Chen G, Bright T, Ratcliffe J, Chung A, Fraser RJ, et al. Preferences for surveillance of Barrett's oesophagus: a discrete choice experiment. J Gastrointest Surg 2019 Jul;23(7):1309-1317. [doi: 10.1007/s11605-018-4049-6] [Medline: 30478530]
25. Verbeek RE, Leenders M, Ten Kate FJ, van Hillegersberg R, Vleggaar FP, van Baal JW, et al. Surveillance of Barrett's esophagus and mortality from esophageal adenocarcinoma: a population-based cohort study. Am J Gastroenterol 2014 Aug;109(8):1215-1222. [doi: 10.1038/ajg.2014.156] [Medline: 24980881]
26. Holmberg D, Santoni G, von Euler-Chelpin MC, Färkkilä M, Kauppila JH, Maret-Ouda J, et al. Incidence and mortality in upper gastrointestinal cancer after negative endoscopy for gastroesophageal reflux disease. Gastroenterology 2022 Feb;162(2):431-8.e4 [FREE Full text] [doi: 10.1053/j.gastro.2021.10.003] [Medline: 34627859]
27. van Dam L, Hol L, de Bekker-Grob EW, Steyerberg EW, Kuipers EJ, Habbema JD, et al. What determines individuals' preferences for colorectal cancer screening programmes? A discrete choice experiment. Eur J Cancer 2010 Jan;46(1):150-159 [FREE Full text] [doi: 10.1016/j.ejca.2009.07.014] [Medline: 19683432]
28. Brinkmann M, Diedrich L, Hemmerling M, Krauth C, Robra BP, Stahmeyer JT, et al. Heterogeneous preferences for colorectal cancer screening in Germany: results of a discrete choice experiment. Value Health 2023 Jan;26(1):104-114 [FREE Full text] [doi: 10.1016/j.jval.2022.07.012] [Medline: 36031478]
29. Zhou M, Thayer WM, Bridges JF. Using Latent class analysis to model preference heterogeneity in health: a systematic review. Pharmacoeconomics 2018 Feb;36(2):175-187. [doi: 10.1007/s40273-017-0575-4] [Medline: 28975582]
30. Cheung GW, Rensvold RB. Testing factorial invariance across groups: a reconceptualization and proposed new method. J Manage 1999 Feb;25(1):1-27 [FREE Full text] [doi: 10.1177/014920639902500101]
31. Johnson FR, Yang JC, Reed SD. The internal validity of discrete choice experiment data: a testing tool for quantitative assessments. Value Health 2019 Feb;22(2):157-160 [FREE Full text] [doi: 10.1016/j.jval.2018.07.876] [Medline: 30711059]
32. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015 Oct;8(5):373-384 [FREE Full text] [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
33. Orme BK. Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research. 2nd edition. Madison, WI, USA: Research Publishers LLC; 2010.
34. Johnson R, Orme B. Getting the most from CBC. Sawtooth Software Inc. 2003. URL: https://sawtoothsoftware.com/resources/technical-papers/getting-the-most-from-cbc [accessed 2022-06-12]
35. Phisalprapa P, Ngorsuraches S, Wanishayakorn T, Kositamongkol C, Supakankunti S, Chaiyakunapruk N. Estimating the preferences and willingness-to-pay for colorectal cancer screening: an opportunity to incorporate the perspective of population at risk into policy development in Thailand. J Med Econ 2021 Jan;24(1):226-233 [FREE Full text] [doi: 10.1080/13696998.2021.1877145] [Medline: 33467947]
36. Liu S. A study on the employment preference of health management undergraduates in the northern region of China: A Discrete Choice Experiment. Shandong University. Jinan, China: Shandong University; 2019 May 27. URL: https://tinyurl.com/2kkjtnj5 [accessed 2022-07-30]
37. Yoo HI. lclogit2: an enhanced command to fit latent class conditional logit models. Stata J 2020 Jun;20(2):405-425 [FREE Full text] [doi: 10.1177/1536867x20931003]
38. Pacifico D, Yoo HI. lclogit: a stata command for fitting latent-class conditional logit models via the expectation-maximization algorithm. Stata J 2013 Sep 01;13(3):625-639 [FREE Full text] [doi: 10.1177/1536867x1301300312]
39. Ryan M, Kolstad JR, Rockers PC, Dolea C. How to conduct a discrete choice experiment for health workforce recruitment and retention in remote and rural areas: a user guide with case studies. The World Bank. 2012 Dec 20. URL: https://tinyurl.com/28xa42aw [accessed 2022-07-12]
40. Evans NJ. Assessing the practical differences between model selection methods in inferences about choice response time tasks. Psychon Bull Rev 2019 Aug;26(4):1070-1098 [FREE Full text] [doi: 10.3758/s13423-018-01563-9] [Medline: 30783896]
41. Grande D, Mitra N, Shah A, Wan F, Asch DA. Public preferences about secondary uses of electronic health information. JAMA Intern Med 2013 Oct 28;173(19):1798-1806 [FREE Full text] [doi: 10.1001/jamainternmed.2013.9166] [Medline: 23958803]
42. Li HQ, Xue H, Yuan H, Wan GY, Zhang XY. Preferences of first-degree relatives of gastric cancer patients for gastric cancer screening: a discrete choice experiment. BMC Cancer 2021 Aug 26;21(1):959 [FREE Full text] [doi: 10.1186/s12885-021-08677-9] [Medline: 34445987]
43. de Bekker-Grob EW, Donkers B, Veldwijk J, Jonker MF, Buis S, Huisman J, et al. What factors influence non-participation most in colorectal cancer screening? A discrete choice experiment. Patient 2021 Mar;14(2):269-281 [FREE Full text] [doi: 10.1007/s40271-020-00477-w] [Medline: 33150461]

44. Li S, Liu S, Ratcliffe J, Gray A, Chen G. Preferences for cervical cancer screening service attributes in rural China: a discrete choice experiment. Patient Prefer Adherence 2019 May 30;13:881-889 [FREE Full text] [doi: 10.2147/PPA.S201913] [Medline: 31239646]
45. Chavez-MacGregor M, Clarke CA, Lichtensztajn DY, Giordano SH. Delayed initiation of adjuvant chemotherapy among patients with breast cancer. JAMA Oncol 2016 Mar;2(3):322-329 [FREE Full text] [doi: 10.1001/jamaoncol.2015.3856] [Medline: 26659132]
46. Zhao Z, Du L, Wang L, Wang Y, Yang Y, Dong H. Preferred lung cancer screening modalities in China: a discrete choice experiment. Cancers (Basel) 2021 Dec 03;13(23):6110 [FREE Full text] [doi: 10.3390/cancers13236110] [Medline: 34885217]
47. Teng Y, Cao M, Chen W. Development, status, and challenges of cancer screening in China. China Cancer 2022;31(07):481-487.
48. Pignone MP, Crutchfield TM, Brown PM, Hawley ST, Laping JL, Lewis CL, et al. Using a discrete choice experiment to inform the design of programs to promote colon cancer screening for vulnerable populations in North Carolina. BMC Health Serv Res 2014 Nov 30;14:611 [FREE Full text] [doi: 10.1186/s12913-014-0611-4] [Medline: 25433801]
49. Wong SF, Norman R, Dunning TL, Ashley DM, Khasraw M, Hayes TM, et al. A discrete choice experiment to examine the preferences of patients with cancer and their willingness to pay for different types of health care appointments. J Natl Compr Canc Netw 2016 Mar;14(3):311-319. [doi: 10.6004/jnccn.2016.0036] [Medline: 26957617]
50. Hauber AB, González JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR conjoint analysis good research practices task force. Value Health 2016 Jun;19(4):300-315 [FREE Full text] [doi: 10.1016/j.jval.2016.04.004] [Medline: 27325321]

Abbreviations

AIC: Akaike information criterion
ASC: alternative specific constant
BIC: Bayesian information criterion
DCE: discrete choice experiment
GDP: gross domestic product
ISPOR: International Society for Pharmacoeconomics and Outcomes Research
LCL: latent class logit
NLT: negative latent type
NQT: neutral quality type
PCT: positive comfortable type
PIT: positive integrated type
UGC: upper gastrointestinal cancer
WTP: willingness to pay

Edited by A Mavragani; submitted 23.09.22; peer-reviewed by W Chen, L Guo; comments to author 28.02.23; revised version received 20.03.23; accepted 25.04.23; published 10.07.23.

Please cite as:
Liu R, Li Q, Li Y, Wei W, Ma S, Wang J, Zhang N
Public Preference Heterogeneity and Predicted Uptake Rate of Upper Gastrointestinal Cancer Screening Programs in Rural China: Discrete Choice Experiments and Latent Class Analysis
JMIR Public Health Surveill 2023;9:e42898
URL: https://publichealth.jmir.org/2023/1/e42898
doi:10.2196/42898
PMID:37428530

Original Paper

The Impact of Metabolic Syndrome Risk Factors on Lung Function Impairment: Cross-Sectional Study

Rafael Molina-Luque[1,2], RN, MSc, PhD; Guillermo Molina-Recio[1,2], RN, MSc, PhD; Domingo de-Pedro–Jiménez[3], RN, MSc, PhD; Carlos Álvarez Fernández[4], MSc, MD; María García–Rodríguez[5], MSc, DN, PhD; Manuel Romero-Saldaña[1,2], RN, MSc, PhD

[1]Estilos de Vida, Innovación y Salud, Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), Córdoba, Spain
[2]Departamento de Enfermería, Famarcología y Fisioterapia, Universidad de Córdoba, Córdoba, Spain
[3]Indorama Ventures Química, Sociedad Limitado Unipersonal, Polígono Industrial Guadarranque, San Roque, Cádiz, Spain
[4]Departamento de Salud y Seguridad del Trabajo, Ayuntamiento de Córdoba, Córdoba, Spain
[5]Departamento de Enfermería y Nutrición, Facultad de Ciencias Biomédicas y de la Salud, Villaviciosa de Odón, Madrid, Spain

Corresponding Author:
Guillermo Molina-Recio, RN, MSc, PhD
Estilos de Vida, Innovación y Salud
Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC)
Avd Menéndez Pidal N/N
Córdoba, 14004
Spain
Phone: 34 957 21 37 00
Email: en1moreg@uco.es

Abstract

Background: Metabolic syndrome (MetS) is a constellation of risk factors increasingly present in the world's population. People with this syndrome are at an increased risk of cardiovascular disease and type 2 diabetes mellitus. Moreover, evidence has shown that it affects different organs. MetS and its risk factors are independently associated with impaired lung function, which can be quantified through spirometric variables.

Objective: This study aims to determine whether a high number of MetS criteria is associated with increased lung function decline.

Methods: We conducted a descriptive cross-sectional study with a random sample of 1980 workers. Workers with acute respiratory pathology (eg, influenza), chronic respiratory pathology (eg, chronic bronchitis), or exposure to substances harmful to the lungs (eg, organic and inorganic dust) were not included. MetS was established based on harmonized criteria, and lung function was assessed according to spirometric variables. On the basis of these, classification into restrictive lung disease (RLD), obstructive lung disease, and mixed lung disease (MLD) was performed. In addition, the association between MetS and lung function was established based on analysis of covariance, linear trend analysis, and multiple linear regression.

Results: MetS was associated with worse lung function according to all the spirometric parameters analyzed (percentage of predicted forced expiratory volume in 1 second: mean 83, SD 13.8 vs mean 89.2, SD 12.8; $P<.001$ and percentage of predicted forced vital capacity: mean 85.9, SD 11.6 vs mean 92, SD 11.3; $P<.001$). Moreover, those diagnosed with MetS had a higher prevalence of lung dysfunction (41% vs 21.9%; $P<.001$), RLD (23.4% vs 11.2%; $P<.001$), and MLD (7.3% vs 2.2%; $P<.001$). Furthermore, an increasing number of MetS criteria was associated with a greater impairment of pulmonary mechanics ($P<.001$). Similarly, with an increasing number of MetS criteria, there was a significant linear trend ($P<.001$) in the growth of the prevalence ratio of RLD (0 criteria: 1, 1: 1.46, 2: 1.52, 3: 2.53, 4: 2.97, and 5: 5.34) and MLD (0 criteria: 1, 1: 2.68, 2: 6.18, 3: 9.69, and 4: 11.37). Regression analysis showed that the alteration of all MetS risk factors, adjusted for various explanatory variables, was significantly associated with a worsening of spirometric parameters, except for forced expiratory volume in 1 second/forced vital capacity.

Conclusions: The findings have shown that an increase in cardiometabolic risk factors is associated with a more significant worsening of spirometric variables and a higher prevalence of RLD and MLD. As spirometry could be a crucial tool for monitoring patients at risk of developing chronic pathologies, we conclude that this inexpensive and easily accessible test could help detect changes in lung function in patients with cardiometabolic disorders. This highlights the need to consider the importance of cardiometabolic health in lung function when formulating public health policies.

(JMIR Public Health Surveill 2023;9:e43737) doi:10.2196/43737

KEYWORDS

cardiometabolic risk factor; lung function; metabolic syndrome; restrictive lung disease; spirometry

Introduction

Background

The impact of noncommunicable chronic diseases worldwide is increasingly important. Their prevalence and social and economic impact have increased in the recent years, placing them at the center of public health interest [1-4].

Among the most worrying pathological conditions, metabolic syndrome (MetS) stands out because it is a risk factor for cardiovascular disease and type 2 diabetes mellitus (DM), which are, in turn, the noncommunicable chronic diseases with the most significant impact worldwide [5,6]. MetS, according to the harmonized definition, is characterized by the comorbid presence of ≥3 of the following risk factors: dysglycemia, increased blood pressure, hypertriglyceridemia, abdominal obesity, and altered high-density lipoprotein (HDL) cholesterol levels [6]. The prevalence of MetS exceeds 30% in several countries and is expected to increase in the coming years [7,8]. All these make MetS one of the public health challenges of the 21st century; therefore, early detection is essential to avoid associated complications, even at an early age [9,10].

Lung Dysfunction and MetS

In addition to cardiovascular pathologies and type 2 DM, MetS has been associated with alterations in other systemic organs, most notably alterations in lung function [11,12]. For example, researchers have observed that MetS is associated with the worsening of parameters measuring pulmonary mechanics, leading to respiratory diseases (asthma, chronic obstructive pulmonary disease, pulmonary hypertension, etc) [12-14]. Several MetS-related lung pathologies are currently the most prevalent, have significant morbidity and mortality, and are a major public health concern [15,16]. This link could be explained by insulin resistance, one of the main pathophysiological mechanisms of MetS, although its involvement in the modification of pneumocyte function is not yet clear [17-19].

Parameters for pulmonary function status are obtained through spirometry, which is the test of choice to assess lung function [20]. It is simple to execute for those who are well instructed and is routinely performed by health care professionals, for example, in primary care or during health examinations of workers [21,22]. A comparison of the values obtained in the test with the theoretical values, estimated from validated formulas (according to age, sex, height, etc), provides information on the lung condition [23,24]. It is, therefore, an essential public health tool for primary, secondary, and tertiary prevention of pathologies with a high impact on the population [25].

In this context, health care professionals, who play an essential role in disease prevention and health promotion, have a tool that, when used efficiently, would help to detect problems beyond pulmonary deterioration [26]. Regarding the latter, given that the onsets of MetS is reflected in spirometric variables, and its risk factors also do so independently, it is relevant to know whether a progressive deterioration of lung function is caused by the individual presence of the risk factors characterizing MetS. Thus, through public health programs (hospitals, communities, workplaces, etc) that include the study of these parameters from the perspective of cardiometabolic alterations, it would be possible to detect the risk factors and treat them early to prevent the development of MetS.

Therefore, the researchers hypothesized that the increase in the number of MetS components is associated with a greater decline in spirometric parameters and, therefore, an alteration in lung function. On this basis, the objective of this study was to determine how spirometric variables are affected because of the increase in the number of the MetS risk factors present.

Methods

Design, Population, and Sample

A descriptive cross-sectional study was conducted in the working population of the Cordoba City Council (Spain). The minimum sample size estimated was 1685 workers based on an expected prevalence of MetS of 14.9%, a power of 80%, a precision of 1.5%, and a confidence of 95%. The sample was randomly selected from workers who underwent a health examination between 2015 and 2019 at the occupational health unit of the City Council. All employees of the Cordoba City Council (blue-collar workers and white-collar workers) undergo an annual occupational health test. Each of the measurements carried out is stored in a computer program of the occupational health unit, which allows for the subsequent extraction of anonymized data. For this study, the number of workers selected was increased by 20% over the minimum sample size calculated to compensate for losses owing to noncompliance with the eligibility criteria.

Workers who were diagnosed with any acute respiratory pathology (influenza, common cold, etc) or chronic pulmonary disease (chronic bronchitis, emphysema, pneumonia, chronic obstructive pulmonary disease, etc) and those who could not perform the spirometry test according to the protocol were excluded from the study. In addition, the occupational risk of the positions held by the workers was assessed, and it was confirmed that there was no exposure to organic and inorganic dust or substances of high molecular weight that could lead to occupational lung pathologies or affect, transiently, the spirometry test results.

Variables and Measures

Lung function was studied through spirometric parameters: forced expiratory volume in 1 second (FEV1; in L), forced vital capacity (FVC; in L), and the FEV1/FVC ratio. To determine the status of pulmonary function, the expected values of FEV1 and FVC were estimated and related to those obtained in spirometry by calculating the percentage of predicted FEV1 (FEV1%) and percentage of predicted FVC (FVC%) [24].

In addition, the percentage reductions of FEV1 and FVC with respect to the theoretical or expected values of FEV1 and FVC were calculated:

FEV1 reduction (%) = ([FEV1 – expected FEV1]/FEVI) × 100 **(1)**

FVC reduction (%) = ([FVC – expected FVC]/FVC) × 100 **(2)**

Workers were also categorized according to their dysfunctional pattern [27]: normal: FEV1/FVC>0.7 and FVC%>0.8; restrictive lung disease (RLD): FEV1/FVC>0.7 and FVC%<0.8; obstructive lung disease (OLD): FEV1/FVC<0.7 and FVC%>0.8; and mixed lung disease (MLD): FEV1/FVC<0.7 and FVC%<0.8. In addition, the aforementioned categories were grouped into a new dichotomous variable with the following values: normal and lung dysfunction (encompassing RLD, OLD, and MLD).

The spirometry test was performed using the DATOSPIR 120 C spirometer (Silbemed) and following the recommendations of the Spanish Society of Pneumology and Thoracic Surgery [27]. Workers were instructed not to use bronchodilators during the previous 6 to 48 hours (depending on the drug), to avoid caffeine during the previous 8 hours, not to smoke during the previous 24 hours, avoid alcohol and sedatives during the last 4 hours, and not to exercise during the previous 30 minutes. During the measurement, the worker wore a nose clip and was seated upright with his back against the backrest. Once the posture was correct, the worker was asked to inhale as much air as possible and then exhale as fast and hard as possible until they were told to stop. Workers had to perform at least 3 valid tests, with the highest spirometric values being chosen. The researchers considered a test valid when the test had a rapid start (back-extrapolated volume <5% of the FVC) and was performed without hesitation, involved continuous expiration for a duration not <6 seconds, had no abrupt end (last changes in volume lower than 0.025 L for ≥1 s), and had no anomalies in the technique (cough, new inhalation, among others). Overall, 2 or 3 acceptable maneuvers were necessary for spirometry test's interpretation in which the difference between the 2 best FVC and FEV1 measures was not >0.2 L.

Moreover, risk factors included in the harmonized criteria for the diagnosis of MetS were considered: waist circumference (WC) ≥102 cm in men and ≥88 cm in women, triglycerides ≥150 mg/dL, HDL cholesterol <50 mg/dL in women and <40 mg/dL in men, blood pressure ≥130/85 mm Hg or being on treatment for high blood pressure and fasting glucose ≥100 mg/dL, being on antidiabetic treatment, or diagnosed with DM. Thus, the presence of ≥3 of these risk factors implies a diagnosis of MetS [6].

The independent variables collected were age (years), sex (female and male), smoking status (nonsmoker, ex-smoker, and smoker), physical activity level (light, moderate, and heavy), weight (kg), height (cm), BMI (kg/m^2), WC (cm), hip circumference (cm), body fat percentage (Equation Córdoba for Estimation of Body Fat [28]), waist to hip ratio, waist to height ratio, systolic blood pressure (mm Hg), diastolic blood pressure (mm Hg), fasting plasma glucose (mg/dL), HDL cholesterol (mg/dL), and triglycerides (mg/dL). In addition, the diagnoses of DM (yes or no) and hypertension (yes or no) were collected.

Anthropometric variables were measured according to international recommendations [29]. Weight and height were collected using an Atlántida S11 stadiometer and scale (Básculas y Balanzas Añó-Sayol), with an accuracy of 0.1 kg and 0.1 cm, respectively. BMI was calculated based on these variables, and workers were categorized into normal weight, overweight, and obese groups, as proposed by the World Health Organization. WC was measured at end-expiration at the midpoint between the last rib and the iliac crest. Hip circumference was assessed at the most prominent point of the buttocks. Both variables were measured using a flexible tape, with the worker standing and their feet placed together. Blood pressure was measured according to the recommendations of the manual of arterial hypertension in primary care clinical practice [30], with the patients seated, using a calibrated digital sphygmomanometer (Omron M3, Omron Healthcare). All measurements were performed by specialized personnel to minimize the coefficient of variation. Each measurement was repeated 3 times, and the mean was calculated. Finally, physical activity was estimated in metabolic equivalents using the International Physical Activity Questionnaire.

Ethics Approval

The study protocol complied with the Declaration of Helsinki for medical research involving human participants and was approved by the Andalusian Biomedical Research Ethics Committee (4427/Acta number 295).

Informed Consent

All participants were informed, verbally and in writing, about the objectives of the health study. The researchers obtained informed consent following the current regulations.

Statistical Analysis

Quantitative variables were presented as mean and SD, and qualitative variables were presented as absolute frequencies, percentages, and prevalence ratios. The goodness of fit of the quantitative variables to a normal distribution was studied using a Kolmogorov-Smirnov test with Lilliefors correction.

Hypothesis testing was performed with the 2-tailed Student *t* test for 2 means; *z* tests for independent proportions; and chi-square and Fisher exact tests, when necessary, for qualitative variables. In addition, linear trend tests, analysis of covariance, and multiple linear regression were used to determine the effect of other factors and covariates on the variation in spirometric parameters.

For all statistical analyses, an α error probability of <5% was accepted, and CIs were calculated at 95%. SPSS Statistics (version 22.0; IBM Corp) and EPIDAT 4.2 (Department of Sanidade, Xunta de Galicia) were used for statistical analysis.

Results

Description of the Sample

Of the 2069 randomly selected workers, 168 (8.12%) were excluded because they were diagnosed with an acute respiratory pathology or a chronic pulmonary pathology or were exposed to compounds that could alter their lung function, and in 41 (1.98%) other workers, it was impossible to assess MetS. Finally, 44 (2.13%) other workers were excluded because their clinical records did not include some spirometric variables. Thus, the final sample size was 1816.

The age of the sample ranged from 18.8 to 67 years. The prevalence of MetS was 18.3% (95% CI 16.5%-20.1%), being significantly higher in men (20.6%, 95% CI 18.5%-23.4%) than in women (16%, 95% CI 13.8%-18.5%; *P*<.001). BMI was higher among those with MetS (mean difference [MD] 7.1 kg/m^2; *P*<.001). This group showed a higher prevalence of overweight and obesity (95.9% vs 49.3%; *P*<.001). In addition, in the MetS group, higher abdominal adiposity was observed according to the following parameters: (1) WC (MD 19.5 cm; *P*<.001), (2) waist to hip ratio (MD 0.1; *P*<.001), and (3) waist to height ratio (MD 0.12; *P*<.001).

All the other variables analyzed showed differences between those with and those without MetS, except for height (MD 0.2 cm; *P*=.73). Table 1 summarizes the main variables according to the presence or absence of MetS.

Table 1. Characteristics of the sample according to the presence of metabolic syndrome (MetS).

Variables	Total (n=1816)	With MetS (n=329, 18.1%)	Without MetS (n=1487, 81.9%)	*P* value[a]
Age (years), mean (SD)	43.8 (10.7)	48.6 (9.5)	42.8 (10.7)	<.001
Sex, n (%)				.01
Female	919 (50.6)	146 (44.4)	773 (52)	
Male	897 (49.4)	183 (55.6)	714 (48)	
Height (cm), mean (SD)	167.9 (9.3)	167.9 (9.3)	167.9 (9.4)	.98
Weight (kg), mean (SD)	75.6 (17.3)	91.7 (16.7)	72.1 (15.3)	<.001
BMI (kg/m^2), mean (SD)	26.7 (5.4)	32.5 (5.3)	25.5 (4.5)	<.001
Underweight, n (%)	36 (2)	0 (0)	36 (2.4)	
Normal weight, n (%)	731 (40.3)	13 (4)	718 (48.3)	
Overweight, n (%)	635 (35)	95 (28.9)	540 (36.3)	
Obesity, n (%)	414 (22.8)	221 (67.2)	193 (13)	
Waist (cm), mean (SD)	89.6 (14.1)	105.5 (11.2)	86 (12.1)	<.001
Hip (cm), mean (SD)	102.7 (10.1)	111 (10.9)	100.9 (9)	<.001
WHR[b], mean (SD)	0.87 (0.097)	0.95 (0.09)	0.85 (0.09)	<.001
WHtR[c], mean (SD)	0.53 (0.081)	0.63 (0.07)	0.51 (0.07)	<.001
Body fat (%), mean (SD)	31.4 (8.4)	38.9 (8.2)	29.8 (7.6)	<.001
Physical activity, n (%)				.02
Light	121 (6.7)	31 (9.4)	90 (6.1)	
Moderate	457 (25.2)	93 (28.3)	364 (24.5)	
Heavy	1238 (68.2)	205 (62.3)	1033 (69.5)	
Smoking habit, n (%)				<.001
Nonsmoker	882 (48.6)	130 (39.5)	752 (50.6)	
Ex-smoker	312 (17.2)	74 (22.5)	238 (16)	
Smoker	620 (34.1)	124 (37.7)	496 (33.4)	
Glucose (mg/dL), mean (SD)	99.5 (28.1)	119.8 (46.8)	95.1 (19.3)	<.001
HDL[d] cholesterol (mg/dL), mean (SD)	63.2 (14.5)	54.8 (13.2)	65.1 (14.1)	<.001
SBP[e] (mm Hg), mean (SD)	123.7 (16.9)	137.4 (16.7)	120.6 (15.4)	<.001
DBP[f] (mm Hg), mean (SD)	76.9 (10.5)	86 (9)	74.9 (9.7)	<.001
Triglycerides (mg/dL), mean (SD)	108.3 (68.8)	174.2 (101.3)	93.7 (48.4)	<.001
HBP[g], n (%)	704 (38.8)	277 (84.2)	427 (28.7)	<.001
Type 2 DM[h], n (%)	110 (6.1)	73 (22.2)	37 (2.5)	<.001
FEV1[i] (L), mean (SD)	3.1 (0.8)	2.8 (0.7)	3.1 (0.8)	<.001
FEV1%[j], mean (SD)	88.2 (13.2)	83 (13.8)	89.2 (12.8)	<.001
FEV1 reduction[k], mean (SD)	−16.3 (20.7)	−24.4 (24.4)	−14.5 (19.4)	<.001
FVC[l] (L), mean (SD)	4 (0.9)	3.7 (0.8)	4 (0.9)	<.001
FVC%[m], mean (SD)	90.9 (11.6)	85.9 (11.6)	92 (11.3)	<.001
FVC reduction[n], mean (SD)	−11.9 (15.2)	−18.7 (17.5)	−10.4 (14.3)	<.001
FEV1/FVC, mean (SD)	0.77 (0.07)	0.76 (0.07)	0.78 (0.07)	<.001
FEV1/FVC<0.7, n (%)	218 (12)	58 (17.6)	160 (10.8)	<.001

Variables	Total (n=1816)	With MetS (n=329, 18.1%)	Without MetS (n=1487, 81.9%)	*P* value[a]
FVC%<0.8, n (%)	300 (16.5)	101 (30.7)	199 (13.4)	<.001
Lung dysfunction, n (%)	461 (25.4)	135 (41)	326 (21.9)	<.001
Restrictive lung disease	243 (13.4)	77 (23.4)	166 (11.2)	<.001
Obstructive lung disease	161 (8.9)	34 (10.3)	127 (8.5)	.30
Mixed lung disease	57 (3.1)	24 (7.3)	33 (2.2)	<.001

[a] With MetS versus without MetS.

[b] WHR: waist to hip ratio.

[c] WHtR: waist to height ratio.

[d] HDL: high-density lipoprotein.

[e] SBP: systolic blood pressure.

[f] DBP: diastolic blood pressure.

[g] HBP: high blood pressure.

[h] DM: diabetes mellitus.

[i] FEV1: forced expiratory volume in 1 second.

[j] FEV1%: percentage of predicted forced expiratory volume in 1 second.

[k] FEV1 reduction (%) = ([FEV1 – expected FEV1]/FEV1) × 100.

[l] FVC: forced vital capacity.

[m] FVC%: percentage of predicted forced vital capacity.

[n] FVC reduction (%) = ([FVC – expected FVC]/FVC) × 100.

MetS and Lung Function (Spirometric Parameters)

In relation to lung function, workers with MetS showed lower values of FEV1 (MD 0.3 L; *P*<.001), FVC (MD 0.32 L; *P*<.001), and FEV1/FVC ratio (MD 0.015; *P*<.001). Similarly, higher proportions of FEV1/FVC<0.7 (17.6% vs 10.8%; *P*<.001) and FVC%<0.8 (30.7% vs 13.4%; *P*<.001) were found among participants with MetS. This translated into the presence of MetS being associated with a higher prevalence of lung dysfunction (41% vs 21.9%; *P*<.001). Table 1 provides more detailed results.

Individual MetS Criteria and Lung Function (Spirometric Parameters)

Several adjusted multiple linear regression models are presented in Table 2, where the dependent variables, spirometric variables, are represented in the rows, and the independent variables are represented in the columns. The adjusted multiple linear regression study confirms that a quantitative modification of the variables related to MetS criteria (increased WC, elevated plasma triglycerides, raised blood pressure, increased blood glucose, and decreased HDL) causes an alteration of the spirometric variables (Table 2). Specifically, it is observed that an increase in abdominal adiposity and insulin resistance–related variables lead to a significant decrease in FVC and FEV1. Finally, MetS was associated with reductions of 0.220 L and 0.277 L in FEV1 and FVC, respectively.

Table 2. Multiple linear regression for spirometric parameters (n=1816)[a].

Variable	Waist	Glucose	HDL[b]	Triglycerides	High blood pressure	Metabolic syndrome
FEV1[c]						
β	−.004	−.003	.003	−.001	−.078	−.220
Standardized β	−.065	−.116	.066	−.130	−.050	−.110
t test (*df*)	−3.728 (1809)	−7.410 (1809)	4.026 (1809)	−8.090 (1809)	−2.858 (1809)	−6.929 (1809)
R^2 adjusted	0.554	0.564	0.554	0.566	0.552	0.562
P value	<.001	<.001	<.001	<.001	<.001	<.001
FEV1%[d]						
β	−.146	−.071	.096	−.034	−1.607	−6.046
Standardized β	−.156	−.152	.106	−.180	−.059	−.177
t test (*df*)	−6.186 (1809)	−6.597 (1809)	4.390 (1809)	−7.656 (1809)	−2.345 (1809)	−7.596 (1809)
R^2 adjusted	0.059	0.061	0.049	0.069	0.042	0.069
P value	<.001	<.001	<.001	<.001	.02	<.001
FEV1 reduction[e]						
β	−.237	−.105	.127	−.051	−3.029	−9.147
Standardized β	−.162	−.142	.089	−.171	−.071	−.170
t test (*df*)	−6.397 (1809)	−6.201 (1809)	3.678 (1809)	−7.239 (1809)	−2.803 (1809)	−7.278 (1809)
R^2 adjusted	0.055	0.054	0.041	0.061	0.038	0.061
P value	<.001	<.001	<.001	<.001	.005	<.001
FVC[f]						
β	−.004	−.004	.004	−.002	−.096	−.277
Standardized β	−.064	−.115	.067	−.121	−.050	−.115
t test (*df*)	−3.790 (1809)	−7.532 (1809)	4.184 (1809)	−7.714 (1809)	−2.977 (1809)	−7.390 (1809)
R^2 adjusted	0.576	0.586	0.577	0.586	0.575	0.585
P value	<.001	<.001	<.001	<.001	<.001	<.001
FVC%[g]						
β	−.147	−.064	.096	−.029	−1.531	−5.941
Standardized β	−.178	−.154	.121	−.172	−.064	−.197
t test (*df*)	−7.016 (1809)	−6.640 (1809)	4.969 (1809)	−7.256 (1809)	−2.511 (1809)	−8.422 (1809)
R^2 adjusted	0.045	0.042	0.032	0.047	0.022	0.056
P value	<.001	<.001	<.001	<.001	.01	<.001
FVC reduction[h]						
β	−.202	−.088	.124	−.038	−2.126	−7.966
Standardized β	−.187	−.163	.118	−.173	−.068	−.202
t test (*df*)	−7.365 (1809)	−7.041 (1809)	4.870 (1809)	−7.270 (1809)	−2.655 (1809)	−8.638 (1809)
R^2 adjusted	0.047	0.045	0.031	0.046	0.022	0.057
P value	<.001	<.001	<.001	<.001	.008	<.001
FEV1/FVC						
β	−.004	−.006	−.003	−.005	−.055	−.099
Standardized β	−.009	−.023	−.007	−.052	−.004	−.006

Variable	Waist	Glucose	HDL[b]	Triglycerides	High blood pressure	Metabolic syndrome
t test (*df*)	−0.376 (1809)	−1.043 (1809)	−0.316 (1809)	−2.284 (1809)	−0.161 (1809)	−0.247 (1809)
R^2 adjusted	0.126	0.127	0.126	0.129	0.126	0.126
P value	.71	.30	.75	.02	.87	.81

[a] The variables in the rows are the dependent variables, and the variables in the columns are the independent variables. Models were adjusted by age, height, smoking habit, sex, and physical activity (dichotomized: light and active [moderate and heavy]).

[b] HDL: high-density lipoprotein.

[c] FEV1: forced expiratory volume in 1 second (L).

[d] FEV1%: percentage of predicted forced expiratory volume in 1 second.

[e] FEV1 reduction (%) = ([FEV1 − expected FEV1]/FEV1) × 100.

[f] FVC: forced vital capacity (L).

[g] FVC%: percentage of predicted forced vital capacity.

[h] FVC reduction (%) = ([FVC – expected FVC]/FVC) × 100.

Number of MetS Criteria and Lung Function (Spirometric Parameters)

Table 3 shows the effect of the spirometric variables as a function of the number of MetS components.

FVC and FVC% showed a decrease in values with an increasing number of MetS criteria. However, the differences found between having 1 or 2 criteria (FVC: MD 0.046 L, *P*=.16; FVC%: MD 1.330, *P*=.08; FVC reduction: MD 1.481%, *P*=.13), 3 or 4 criteria (FVC: MD 0.107 L, *P*=.06; FVC%: MD 2.066, *P*=.13; FVC reduction: MD 3.257%, *P*=.07), 3 or 5 criteria (FVC: MD 0.165 L, *P*=.15; FVC%: MD 4.322, *P*=.21; FVC reduction: MD 7.228%, *P*=.11), and 4 or 5 criteria (FVC: MD 0.058 L, *P*=.16; FVC%: MD 2.256, *P*=.53; FVC reduction: MD 3.971%, *P*=.40) were not significant in any of the variables studied. Nevertheless, it is worth noting that the percentage reduction of FVC from what was expected showed a clear linear downward trend, with an MD of 16 (SD 20) percentage points between those with 0 criteria and those with 5 criteria (*P*<.001; Figure 1). These results were also observed for FEV1, FEV1%, and FEV1 reduction. These findings indicate that an increase in the number of MetS criteria is associated with a significant reduction in lung function.

Table 3. Lung function according to the number of metabolic syndrome criteria.

Variables	Number of components, mean (SD)						*P* value[a]
	0 (n=540)	1 (n=554)	2 (n=393)	3 (n=223)	4 (n=95)	5 (n=11)	
FEV1[b]	3.16 (0.44)	3.09 (0.42)[c]	3.04 (0.44)[c]	2.93 (0.43)[d]	2.85 (0.44)[d]	2.77 (0.45)[d]	<.001
FEV1%[e]	91 (13)	89.1 (12.7)[c]	87.5 (12.7)[c]	84 (12.8)[d]	81.6 (12.8)[d]	80.8 (12.7)[d]	<.001
FEV1 reduction[f]	−12 (20)	−15.2 (20)[c]	−17.2 (20)[c]	−22.6 (20.1)[d]	−25.8 (20.1)[d]	−28.1 (20.1)[d]	<.001
FVC[g]	4.07 (0.51)	3.99 (0.49)[c]	3.95 (0.5)[c]	3.78 (0.49)[d]	3.68 (0.5)[d]	3.62 (0.49)[d]	<.001
FVC%[h]	93.5 (11.5)	91.8 (11.3)[c]	90.4 (11.3)[c]	86.7 (11.4)[d]	84.6 (11.4)[d]	82.3 (11.3)[d]	<.001
FVC reduction[i]	−8.7 (14.8)	−10.8 (14.8)[c]	−12.3 (14.8)[c]	−17.2 (14.8)[d]	−20.4 (14.8)[d]	−24.4 (14.8)[d]	<.001
FEV1/FVC	77.7 (6.6)	77.4 (6.5)	77 (6.5)	77.3 (6.5)	77.1 (6.5)	77.8 (6.5)	.73

[a] Analysis of covariance was adjusted for sex, age, height, smoking habits, and physical activity (dichotomized: light and active [moderate and heavy]).

[b] FEV1: forced expiratory volume in 1 second (L).

[c] The presence of the same symbol in the same row indicates that there were no significant differences between pulmonary function and the number of components.

[d] The presence of the same symbol in the same row indicates that there were no significant differences between pulmonary function and the number of components.

[e] FEV1%: percentage of predicted forced expiratory volume in 1 second.

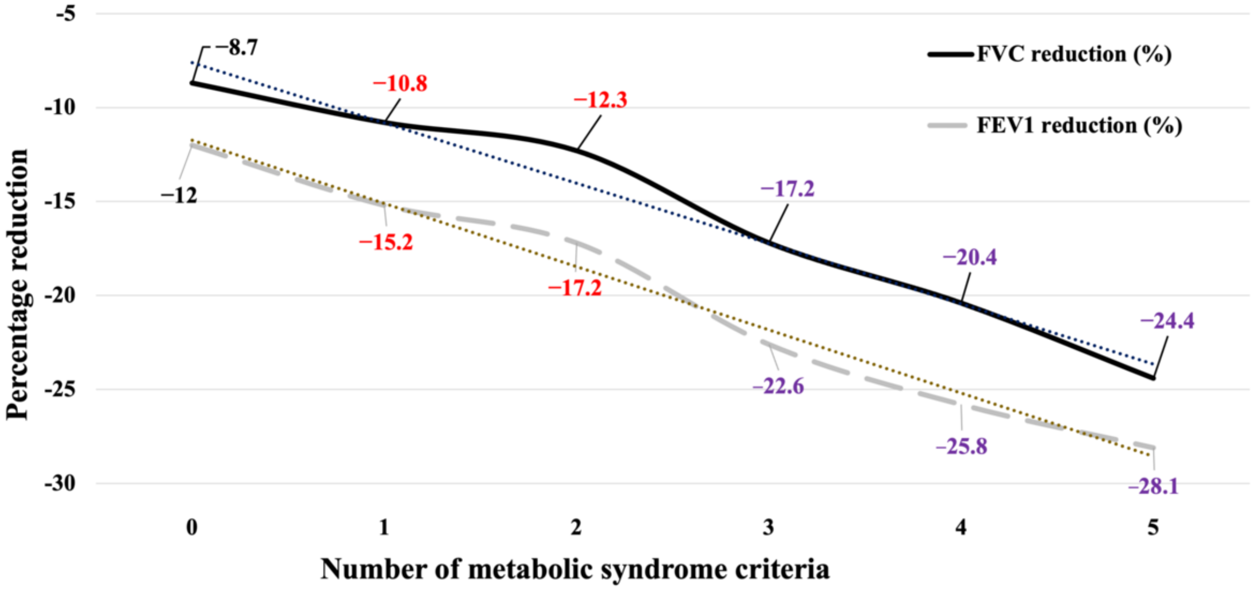
[f] FEV1 reduction (%) = ([FEV1 − expected FEV1]/FEV1) × 100.

[g] FVC: forced vital capacity (L).

[h] FVC%: percentage of predicted forced vital capacity.

[i] FVC reduction (%) = ([FVC − expected FVC]/FVC) × 100.

Figure 1. Percentage reductions of forced expiratory volume in 1 second (FEV1) and forced vital capacity (FVC) according to number of metabolic syndrome criteria. Same color means that there were no significant differences between values within a variable.



Number of MetS Criteria and Lung Dysfunction

This effect on pulmonary function variables was reflected in the distribution of the prevalence of respiratory disorders. In other words, the alteration of the spirometric values detailed in the *MetS and Lung Function (Spirometric Parameters), Individual MetS Criteria and Lung Function (Spirometric Parameters)* and *Number of MetS Criteria and Lung Function (Spirometric Parameters)* sections has clinical repercussions in the appearance of pathological respiratory patterns. In this respect, the proportion of participants who presented FEV1/FVC<0.7 (0: 8.9%, 1: 11.2%, 2: 12.7%, 3: 16.6%, 4: 21.1%, 5: 9.1%; *P*=.006) and FVC%<0.8 (0: 9.3%, 1: 14.4%, 2: 17.6%, 3: 28.7%, 4: 33.7%, 5: 45.5%; *P*<.001) was significantly higher among those with the most risk factors. This trend was also observed in the rates of lung dysfunction (0: 17.4%, 1: 23.6%, 2: 25.7%, 3: 38.1%, 4: 46.3%, 5: 45.5%; *P*<.001), in RLD (0: 8.5%, 1: 12.5%, 2: 13%, 3: 21.5%, 4: 25.3%, 5: 45.5%; *P*<.001), and in MLD (0: 0.7%, 1: 2%, 2: 4.6%, 3: 7.2%, 4: 8.4%, 5: 0%; *P*<.001) but not in OLD (0: 8.1%, 1: 9.2%, 2: 8.1%, 3: 9.4%, 4: 12.6%, 5: 9.1%; *P*=.79). Table 4 shows the prevalence ratios as a function of the number of components. The results show a significant linear increase in the prevalence ratio of all the pulmonary disorders studied with an increasing number of components (*P*<.001), except for OLD.

Table 4. Prevalence ratios for lung disorders according to the number of metabolic syndrome criteria.

Variables	Number of components, prevalence ratios (95% CI)						*P* value[a]	*P* value[b]
	0 (n=540)	1 (n=554)	2 (n=393)	3 (n=223)	4 (n=95)	5 (n=11)		
FEV1[c]/FVC[d]<0.7	1 (reference)[e]	1.26 (0.88-1.80)	1.43 (0.99-2.08)	1.87 (1.25-2.78)	2.37 (1.48-3.8)	1.02 (0.16-6.78)	.006	<.001
FVC%[f]<0.8	1 (reference)	1.56 (1.12-2.18)	1.9 (1.35-2.66)	3.1 (2.22-4.34)	3.64 (2.47-5.35)	4.91 (2.43-9.88)	<.001	<.001
Lung dysfunction	1 (reference)	1.36 (1.07-1.72)	1.48 (1.15-1.89)	2.19 (1.71-2.81)	2.66 (2-3.53)	3.13 (1.77-5.54)	<.001	<.001
RLD[g]	1 (reference)	1.46 (1.03-2.08)	1.52 (1.05-2.22)	2.53 (1.74-3.67)	2.97 (1.91-4.62)	5.34 (2.64-10.8)	<.001	<.001
OLD[h]	1 (reference)	1.13 (0.77-1.66)	1 (0.65-1.55)	1.16 (0.70-1.90)	1.55 (0.85-2.82)	1.12 (0.17-7.39)	.79	.33
MLD[i]	1 (reference)	2.68 (0.86-8.37)	6.18 (2.11-18.13)	9.69 (3.27-28.65)	11.37 (3.49-37.01)	0	<.001	<.001

[a] Chi-square test.

[b] Linear trend.

[c] FEV1: forced expiratory volume in 1 second (L).

[d] FVC: forced vital capacity (L).

[e] Reference category in statistical analysis.

[f] FVC%: percentage of predicted forced vital capacity.

[g] RLD: restrictive lung disease.

[h] OLD: obstructive lung disease.

[i] MLD: mixed lung disease.

Discussion

Principal Findings

This study aimed to identify the association between MetS and changes in spirometric parameters and precisely determine whether a higher number of MetS components is related to a worse state of lung function.

The results show that the presence of MetS is associated with worse lung function. In this study, participants with MetS showed lower mean FEV1, FEV1%, FVC, and FVC%. These findings were confirmed when adjusting for different explanatory variables, showing a clear effect of MetS on different spirometric parameters. In the case of FEV1/FVC, although its mean was lower in the MetS group, this association was lost when adjusting for other independent variables. This trend has been evidenced in populations of different ethnicities across cross-sectional and longitudinal designs, although discrepancies in FEV1/FVC have been observed [31-36]. Kim et al [32] showed that after 6 years of follow-up, participants who had MetS at the beginning of the study or developed MetS during the study had a greater decline in FVC and FEV1 than those who were healthy. However, the difference was not significant in the case of FEV1. Ford et al [33] found that participants with MetS had lower FEV1, FEV1%, FVC, and FVC% and higher FEV1/FVC.

These changes in spirometric parameters have clinical relevance, as they translate into the development of lung dysfunction. However, there are inconsistencies in the type of alteration most present in participants with MetS. Some authors state that OLD is more prevalent [12,34,37], whereas others show that RLD is predominant [12,33,36,38-40]. The results of our study show a higher proportion of participants with FEV1/FVC<0.7 and FVC%<0.8 among those diagnosed with MetS. This was reflected in a higher prevalence of lung dysfunction, RLD, and MLD, with no difference observed in OLD. In contrast to our results, Buchman et al [34] found that among men and women aged >60 years with MetS, there was no higher prevalence of FEV1/FVC<0.7 compared with those without MetS. However, we did not analyze by age group, so we do not know what happens specifically in those aged >60 years. Scarlata et al [39] observed that participants with RLD had 3-fold higher odds of MetS (95% CI 1.16-7.89) than those with normal spirometry, similar to findings by other authors [33]. Inconsistencies in the obstructive pattern may be (1) because some studies do not consider the mixed pattern, which may increase the percentage of patients with an obstructive pattern [37] or (2) because of the noninclusion of the restrictive pattern in the study [34]. However, it seems clear that there is a greater tendency for the presence of RLD in people with MetS [12,33,36,40].

Despite the clear association of MetS with lung dysfunction, the pathophysiological process remains unclear. However, several mechanisms that could be crucial in the process have been described, such as insulin resistance and low-grade inflammation, both of which are present in MetS [17,18,41], which is reflected in the lungs through different markers [42]. Therefore, it seems reasonable to state that damage in other body systems, identified according to MetS risk factors, is associated with lung damage [41]. In this regard, our results have shown that each MetS component (WC, triglycerides, HDL, glucose, and high blood pressure) can significantly modify each spirometric parameter once adjusted for explanatory variables. This fact has also been evidenced by other authors [37,38,43,44].

The accumulation of systemic damage, measured by the number of MetS criteria, has been associated with more significant organ damage [45-48]. In this regard, we have shown that having a higher number of MetS criteria was related to a significantly worse lung function (spirometric parameters, except for FEV1/FVC), although no statistically significant differences were found between some analysis groups defined by the number of MetS criteria. In addition, linear percentage reductions of FVC and FEV1 were observed among the different groups. This inverse relationship has been observed in other populations, highlighting the association of central obesity with lung function decline [31,33,35-38,49,50].

These findings are essential for public health in any country. Therefore, it could be recommended that when health care professionals observe a progressive loss of lung function in a patient, they should be alerted and explore for possible cardiometabolic impairment. Concerning the latter, it is noteworthy that a significant linear increase in the prevalence ratios of pulmonary disorders was observed when the number of MetS components increased. Specifically, the decreases in FVC% below 80% and in the restrictive pattern stand out. In this regard, Lee et al [38] found a linear growth trend in the prevalence of RLD in men but not in women. Chen et al [50] reported a significant linear trend between the accumulation of MetS components and the decline in FVC% and FEV1% in both men and women. Similar results were found by Yoon et al [36]. They also reported that a significant increase in the odds ratio for RLD was observed with an increasing number of risk factors, with the magnitude varying according to the adjustment made. This trend was not observed for OLD [36].

Regardless of the type of lung disorder associated with MetS, several researchers have highlighted the need to intervene in these patients to ensure better respiratory health [12,32,33,50]. From our perspective and based on the results obtained, we believe that governments should promote public health programs that include the detection of cardiometabolic disorders and their impacts on lung health. Although some organizations consider the magnitude of the most critical public health problems as risk factors for lung disease [15,51,52], only a few consider metabolic disorders such as obesity [53]. The accumulation of visceral fat, the primary pathophysiological mechanism in the development of insulin resistance and thus of major metabolic disturbances, requires public health attention [54]. Therefore, metabolic disorders, such as MetS, should be included in the programs more concretely, as the evidence is increasingly strong regarding the relationship between metabolic disorders and lung disorders [55,56], even pointing to lung cancer [57].

With this inclusion, health professionals could be encouraged to perform routine spirometry in health examinations to detect early lung disorders, both those of both respiratory origin and those of cardiometabolic origin, not only in the working population (higher risk) but also in the general population [25]. Health professionals are on the frontline, and public health must provide strategies to diagnose and prevent lung disorders and protect and promote health [58], which has been advocated for years [59,60]. This idea gains relevance in a context where spirometry may be underused or, when used, misused [61,62].

Limitations and Strengths

Owing to the epidemiological design, it is impossible to establish a cause-effect relationship between the increase in the number of cardiometabolic risk factors and the progressive alteration of spirometric variables. It would be interesting to conduct a retrospective cohort study to determine whether exposure to different cardiometabolic alterations modifies lung function and increases pulmonary disease incidence. Another possible limitation is the small number of participants in the group with 5 MetS criteria, which makes it difficult to analyze their lung function status. However, given the sample's representativeness, we consider that its small size is linked to what can be found in real settings, where the presence of these individuals is scarce. Moreover, given the particular characteristics of the sample (workers, ie, the working population), extrapolation to the general population (more sedentary) is challenging. However, workers represent a significant proportion of the people, meaning that the participants analyzed cover an essential demographic stratum. Although workers' smoking habits have been controlled, including the number of annual packs of cigarettes in data collection could improve the accuracy of the analysis. Despite the limitations, the robustness of the statistical tests, the control of confounding variables, and the large sample size make the results consistent. Furthermore, the methodology used has facilitated comparison with similar studies, making it possible to reinforce knowledge on the topic.

Conclusions

The presence of a greater number of MetS criteria (cardiometabolic risk factors) has been associated with increased lung dysfunction and a higher prevalence of pulmonary impairments, especially those of the restrictive and mixed types. In addition, the findings show that the comorbid occurrence of risk factors leads to a deterioration of FVC and FEV1. The findings highlight the need for governments to consider the importance of cardiometabolic health in lung function when formulating public health policies that are to be implemented in hospitals, health centers, companies, etc. In this context, spirometry could be crucial for health professionals to monitor patients at a risk of developing chronic pathologies. In addition, routine spirometry helps detect changes in lung function related to cardiometabolic disorders. In short, spirometry is an accessible method that should be used to prevent and provide early care for diseases, including those not of pulmonary origin.

Data Availability

The data sets generated and analyzed during this study are not publicly available because they belong to a public institution but are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Noncommunicable diseases. World Health Organization. 2022 Sep 16. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2023-04-29]
2. NCD Countdown 2030 collaborators. NCD countdown 2030: worldwide trends in non-communicable disease mortality and progress towards sustainable development goal target 3.4. Lancet 2018 Sep 22;392(10152):1072-1088 [FREE Full text] [doi: 10.1016/S0140-6736(18)31992-5] [Medline: 30264707]
3. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, Biryukov S. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Br Dent J 2015 Oct 09;219(7):329 [FREE Full text] [doi: 10.1038/sj.bdj.2015.751]
4. Global Burden of Cardiovascular Diseases Collaboration, Roth GA, Johnson CO, Abate KH, Abd-Allah F, Ahmed M, et al. The burden of cardiovascular diseases among US states, 1990-2016. JAMA Cardiol 2018 May 01;3(5):375-389 [FREE Full text] [doi: 10.1001/jamacardio.2018.0385] [Medline: 29641820]
5. Yusuf S, Joseph P, Rangarajan S, Islam S, Mente A, Hystad P, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet 2020 Mar 07;395(10226):795-808 [FREE Full text] [doi: 10.1016/S0140-6736(19)32008-2] [Medline: 31492503]
6. Alberti K, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009 Oct 20;120(16):1640-1645. [doi: 10.1161/circulationaha.109.192644]
7. Fernández-Bergés D, Cabrera de León A, Sanz H, Elosua R, Guembe MJ, Alzamora M, et al. Síndrome metabólico en España: prevalencia y riesgo coronario asociado a la definición armonizada y a la propuesta por la OMS. Estudio DARIOS. Rev Esp Cardiol (Engl Ed) 2012 Mar;65(3):241-248. [doi: 10.1016/j.recesp.2011.10.015] [Medline: 22305818]
8. Shin D, Kongpakpaisarn K, Bohra C. Trends in the prevalence of metabolic syndrome and its components in the United States 2007-2014. Int J Cardiol 2018 May 15;259:216-219. [doi: 10.1016/j.ijcard.2018.01.139] [Medline: 29472026]
9. Christian Flemming GM, Bussler S, Körner A, Kiess W. Definition and early diagnosis of metabolic syndrome in children. J Pediatr Endocrinol Metab 2020 Jul 28;33(7):821-833. [doi: 10.1515/jpem-2019-0552] [Medline: 32568734]
10. Gesteiro E, Megía A, Guadalupe-Grau A, Fernandez-Veledo S, Vendrell J, González-Gross M. Early identification of metabolic syndrome risk: a review of reviews and proposal for defining pre-metabolic syndrome status. Nutr Metab Cardiovasc Dis 2021 Aug 26;31(9):2557-2574 [FREE Full text] [doi: 10.1016/j.numecd.2021.05.022] [Medline: 34244048]
11. Mendrick DL, Diehl AM, Topor LS, Dietert RR, Will Y, La Merrill MA, et al. Metabolic syndrome and associated diseases: from the bench to the clinic. Toxicol Sci 2018 Mar 01;162(1):36-42 [FREE Full text] [doi: 10.1093/toxsci/kfx233] [Medline: 29106690]
12. Baffi CW, Wood L, Winnica D, Strollo PJJ, Gladwin MT, Que LG, et al. Metabolic syndrome and the lung. Chest 2016 Jun;149(6):1525-1534 [FREE Full text] [doi: 10.1016/j.chest.2015.12.034] [Medline: 26836925]
13. Lee K. Lung function and impaired kidney function in relation to metabolic syndrome. Int Urol Nephrol 2017 Jul;49(7):1217-1223. [doi: 10.1007/s11255-017-1567-1] [Medline: 28299524]
14. Kwon S, Lee M, Crowley G, Schwartz T, Zeig-Owens R, Prezant DJ, et al. Dynamic metabolic risk profiling of world trade center lung disease: a longitudinal cohort study. Am J Respir Crit Care Med 2021 Nov 01;204(9):1035-1047. [doi: 10.1164/rccm.202006-2617oc]
15. Chronic respiratory diseases. World Health Organization. URL: https://www.who.int/health-topics/chronic-respiratory-diseases#tab=tab_1 [accessed 2023-06-26]
16. Safiri S, Carson-Chahhoud K, Noori M, Nejadghaderi SA, Sullman MJ, Ahmadian Heris J, et al. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019. BMJ 2022 Jul 27;378:e069679 [FREE Full text] [doi: 10.1136/bmj-2021-069679] [Medline: 35896191]
17. Lee YB, Kim YS, Lee DH, Kim HY, Lee JI, Ahn HS, et al. Association between HOMA-IR and lung function in Korean young adults based on the Korea National Health and Nutrition Examination Survey. Sci Rep 2017 Sep 15;7(1):11726 [FREE Full text] [doi: 10.1038/s41598-017-11337-3] [Medline: 28916790]

18. McNeill JN, Lau ES, Zern EK, Nayor M, Malhotra R, Liu EE, et al. Association of obesity-related inflammatory pathways with lung function and exercise capacity. Respir Med 2021 Jul;183:106434 [FREE Full text] [doi: 10.1016/j.rmed.2021.106434] [Medline: 33964816]
19. Yang G, Han YY, Forno E, Yan Q, Rosser F, Chen W, et al. Glycated hemoglobin A1c, lung function, and hospitalizations among adults with asthma. J Allergy Clin Immunol Pract 2020 Nov;8(10):3409-15.e1 [FREE Full text] [doi: 10.1016/j.jaip.2020.06.017] [Medline: 32569755]
20. Global initiative for chronic obstructive lung disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. 2020. URL: https://goldcopd.org/wp-content/uploads/2019/12/GOLD-2020-FINAL-ver1.2-03Dec19_WMV.pdf [accessed 2022-04-29]
21. Represas-Represas C, Botana-Rial M, Leiro-Fernández V, González-Silva AI, García-Martínez A, Fernández-Villar A. Short- and long-term effectiveness of a supervised training program in spirometry use for primary care professionals. Arch Bronconeumol 2013 Sep;49(9):378-382. [doi: 10.1016/j.arbres.2013.01.001] [Medline: 23481409]
22. Ranzieri S, Corradi M. Conducting spirometry in occupational health at COVID-19 times: international standards. Med Lav 2021 Apr 20;112(2):95-106 [FREE Full text] [doi: 10.23749/mdl.v112i2.11420] [Medline: 33881003]
23. de Vicente CM, de Mir Messa I, Amigo SR, Vernetta AT, Gartner S, Serrano II, et al. Validación de las ecuaciones de referencia para la espirometría forzada en niños sanos preescolares españoles. Arch Bronconeumol 2018 Jan;54(1):24-30. [doi: 10.1016/j.arbres.2017.07.019] [Medline: 28935165]
24. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J 2012 Dec;40(6):1324-1343 [FREE Full text] [doi: 10.1183/09031936.00080312] [Medline: 22743675]
25. Spirometry monitoring technology. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/niosh/topics/spirometry/spirola.html [accessed 2023-06-26]
26. Vaz Fragoso CA, Van Ness PH, McAvay GJ. FEV1 as a standalone spirometric predictor and the attributable fraction for death in older persons. Respir Care 2020 Feb;65(2):217-226 [FREE Full text] [doi: 10.4187/respcare.07012] [Medline: 31662447]
27. García-Rio F, Calle M. Normativa sobre la espirometría (revisión 2013). Barcelona, Spain: Separ; 2013.
28. Molina-Luque R, Romero-Saldaña M, Álvarez-Fernández C, Bennasar-Veny M, Álvarez-López Á, Molina-Recio G. Equation Córdoba: a simplified method for estimation of body fat (ECORE-BF). Int J Environ Res Public Health 2019 Nov 15;16(22):4529 [FREE Full text] [doi: 10.3390/ijerph16224529] [Medline: 31731813]
29. Callaway CW, Chumlea WC, Bouchard C, Himes JH, Lohman TG, Martin AD, et al. Circumferences. In: Lohman TG, Roche AF, Martorell R, editors. Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics Books; 1998:44-45.
30. Sociedad Española de Medicina Interna (SEMI) y Sociedad Española de Medicina Familiar y Comunitaria (SEMFYC). [Improving the care of people with chronic conditions in Spain. Common position statement of the Spanish society of internal medicine and the Spanish society of family and community medicine]. Aten Primaria 2020 Jan;52(1):1-2 [FREE Full text] [doi: 10.1016/j.aprim.2019.11.005] [Medline: 31874698]
31. Qin L, Yang Z, Zhang W, Gu H, Lu S, Shi Q, et al. Association between metabolic syndrome and lung function in middle-aged and elderly Chinese individuals. Diabetes Res Clin Pract 2015 Jun;108(3):e46-e48. [doi: 10.1016/j.diabres.2015.03.007] [Medline: 25818886]
32. Kim SK, Bae JC, Baek JH, Jee JH, Hur KY, Lee MK, et al. Decline in lung function rather than baseline lung function is associated with the development of metabolic syndrome: a six-year longitudinal study. PLoS One 2017 Mar 27;12(3):e0174228 [FREE Full text] [doi: 10.1371/journal.pone.0174228] [Medline: 28346522]
33. Ford ES, Cunningham TJ, Mercado CI. Lung function and metabolic syndrome: findings of national health and nutrition examination survey 2007–2010. J Diabetes 2014 Nov 27;6(6):603-613 [FREE Full text] [doi: 10.1111/1753-0407.12136] [Medline: 26677470]
34. Buchmann N, Norman K, Steinhagen-Thiessen E, Demuth I, Eckardt R. Lung function in elderly subjects with metabolic syndrome and type II diabetes : data from the berlin aging study II. Z Gerontol Geriatr (Forthcoming) 2015 Oct 27. [doi: 10.1007/s00391-015-0958-0] [Medline: 26508108]
35. Inomoto A, Fukuda R, Phn JD, Kato G, Rpt RK, Rpt KH, Rpt, et al. The association between the body composition and lifestyle affecting pulmonary function in Japanese workers. J Phys Ther Sci 2016 Oct;28(10):2883-2889 [FREE Full text] [doi: 10.1589/jpts.28.2883] [Medline: 27821955]
36. Yoon H, Gi MY, Cha JA, Yoo CU, Park SM. The association between the metabolic syndrome and metabolic syndrome score and pulmonary function in non-smoking adults. Diab Vasc Dis Res 2018 Mar;15(2):131-138 [FREE Full text] [doi: 10.1177/1479164117746022] [Medline: 29233016]
37. Wang F, Tian D, Zhao Y, Li J, Chen X, Zhang Y. High-density lipoprotein cholesterol: a component of the metabolic syndrome with a new role in lung function. Evid Based Complement Alternat Med 2021 Jun 2;2021:6615595-6615599 [FREE Full text] [doi: 10.1155/2021/6615595] [Medline: 34188689]

38. Lee YY, Tsao YC, Yang CK, Chuang CH, Yu W, Chen JC, et al. Association between risk factors of metabolic syndrome with lung function. Eur J Clin Nutr 2020 May;74(5):811-817 [FREE Full text] [doi: 10.1038/s41430-018-0369-6] [Medline: 31427761]
39. Scarlata S, Fimognari FL, Cesari M, Giua R, Franco A, Pasqualetti P, et al. Lung function changes in older people with metabolic syndrome and diabetes. Geriatr Gerontol Int 2013 Oct;13(4):894-900. [doi: 10.1111/ggi.12026] [Medline: 23311827]
40. Lim SY, Rhee EJ, Sung KC. Metabolic syndrome, insulin resistance and systemic inflammation as risk factors for reduced lung function in Korean nonsmoking males. J Korean Med Sci 2010 Oct;25(10):1480-1486 [FREE Full text] [doi: 10.3346/jkms.2010.25.10.1480] [Medline: 20890430]
41. Wielscher M, Amaral AF, van der Plaat D, Wain LV, Sebert S, Mosen-Ansorena D, et al. Genetic correlation and causal relationships between cardio-metabolic traits and lung function impairment. Genome Med 2021 Jun 21;13(1):104 [FREE Full text] [doi: 10.1186/s13073-021-00914-x] [Medline: 34154662]
42. Brandao-Rangel MA, Moraes-Ferreira R, Oliveira-Junior MC, Santos-Dias AL, Bachi A, Gabriela-Pereira G, et al. Pulmonary function changes in older adults with and without metabolic syndrome. Sci Rep 2021 Aug 30;17337:11 [FREE Full text] [doi: 10.21203/rs.3.rs-271946/v1]
43. Yamamoto Y, Oya J, Nakagami T, Uchigata Y. Association between lung function and metabolic syndrome independent of insulin in Japanese men and women. Jpn Clin Med 2014 Apr 24;5:1-8 [FREE Full text] [doi: 10.4137/JCM.S13564] [Medline: 24812534]
44. Koo HK, Kim DK, Chung HS, Lee CH. Association between metabolic syndrome and rate of lung function decline: a longitudinal analysis. Int J Tuberc Lung Dis 2013 Nov;17(11):1507-1514. [doi: 10.5588/ijtld.12.0906] [Medline: 24125459]
45. Aboonabi A, Meyer RR, Singh I. The association between metabolic syndrome components and the development of atherosclerosis. J Hum Hypertens 2019 Dec;33(12):844-855. [doi: 10.1038/s41371-019-0273-0] [Medline: 31636352]
46. Okada R, Yasuda Y, Tsushita K, Wakai K, Hamajima N, Matsuo S. The number of metabolic syndrome components is a good risk indicator for both early- and late-stage kidney damage. Nutr Metab Cardiovasc Dis 2014 Mar;24(3):277-285. [doi: 10.1016/j.numecd.2013.08.004] [Medline: 24418372]
47. Eguchi K, Schwartz JE, Roman MJ, Devereux RB, Gerin W, Pickering TG. Metabolic syndrome less strongly associated with target organ damage than syndrome components in a healthy, working population. J Clin Hypertens (Greenwich) 2007 May;9(5):337-344 [FREE Full text] [doi: 10.1111/j.1524-6175.2007.06474.x] [Medline: 17485969]
48. Gui MH, Ling Y, Liu L, Jiang JJ, Li XY, Gao X. Effect of metabolic syndrome score, metabolic syndrome, and its individual components on the prevalence and severity of angiographic coronary artery disease. Chin Med J (Engl) 2017 Mar 20;130(6):669-677 [FREE Full text] [doi: 10.4103/0366-6999.201611] [Medline: 28303849]
49. Yu D, Simmons D. Association between lung capacity measurements and abnormal glucose metabolism: findings from the Crossroads study. Diabet Med 2014 May;31(5):595-599. [doi: 10.1111/dme.12346] [Medline: 24151940]
50. Chen WL, Wang CC, Wu LW, Kao TW, Chan JY, Chen YJ, et al. Relationship between lung function and metabolic syndrome. PLoS One 2014 Oct 9;9(10):e108989 [FREE Full text] [doi: 10.1371/journal.pone.0108989] [Medline: 25299452]
51. Risks and causes of lung cancer. Cancer Research UK. URL: https://www.cancerresearchuk.org/about-cancer/lung-cancer/risks-causes [accessed 2023-06-26]
52. COPD causes and risk factors. American Lung Association. URL: https://www.lung.org/lung-health-diseases/lung-disease-lookup/copd/what-causes-copd#:~:text=COPD%20Risk%20Factors,coal%20or%20wood%20burning%20stove [accessed 2023-06-26]
53. How public health issues affect lung health. American Lung Association. URL: https://www.lung.org/lung-health-diseases/wellness/public-health-issues-lung-health [accessed 2023-06-26]
54. Neeland IJ, Ross R, Després JP, Matsuzawa Y, Yamashita S, Shai I, International Atherosclerosis Society, International Chair on Cardiometabolic Risk Working Group on Visceral Obesity. Visceral and ectopic fat, atherosclerosis, and cardiometabolic disease: a position statement. Lancet Diabetes Endocrinol 2019 Sep;7(9):715-725 [FREE Full text] [doi: 10.1016/S2213-8587(19)30084-1] [Medline: 31301983]
55. Fekete M, Szollosi G, Tarantini S, Lehoczki A, Nemeth AN, Bodola C, et al. Metabolic syndrome in patients with COPD: causes and pathophysiological consequences. Physiol Int 2022 Mar 03;109(1):90-105. [doi: 10.1556/2060.2022.00164] [Medline: 35238797]
56. Chan SM, Selemidis S, Bozinovski S, Vlahos R. Pathobiological mechanisms underlying metabolic syndrome (MetS) in chronic obstructive pulmonary disease (COPD): clinical significance and therapeutic strategies. Pharmacol Ther 2019 Jun;198:160-188 [FREE Full text] [doi: 10.1016/j.pharmthera.2019.02.013] [Medline: 30822464]
57. Sin S, Lee CH, Choi SM, Han KD, Lee J. Metabolic syndrome and risk of lung cancer: an analysis of Korean National Health Insurance Corporation database. J Clin Endocrinol Metab 2020 Nov 01;105(11):dgaa596. [doi: 10.1210/clinem/dgaa596] [Medline: 32860708]
58. Respiratory disease: applying All Our Health. Public Health England. URL: https://www.gov.uk/government/publications/respiratory-disease-applying-all-our-health/respiratory-disease-applying-all-our-health#:~:text=Respiratory%20disease%20covers%20a%20wide,interstitial%20lung%20disease%20and%20mesothelioma [accessed 2023-06-26]

59. Schluger NW, Koppaka R. Lung disease in a global context. A call for public health action. Ann Am Thorac Soc 2014 Mar;11(3):407-416. [doi: 10.1513/AnnalsATS.201312-420PS] [Medline: 24673697]
60. GBD 2019 Chronic Respiratory Diseases Collaborators. Global burden of chronic respiratory diseases and risk factors, 1990-2019: an update from the Global Burden of Disease Study 2019. EClinicalMedicine 2023 May 13;59(3):101936 [FREE Full text] [doi: 10.1016/j.eclinm.2023.101936] [Medline: 37229504]
61. Ramalho SH, Shah AM. Lung function and cardiovascular disease: a link. Trends Cardiovasc Med 2021 Feb;31(2):93-98 [FREE Full text] [doi: 10.1016/j.tcm.2019.12.009] [Medline: 31932098]
62. Perret J, Yip SW, Idrose NS, Hancock K, Abramson MJ, Dharmage SC, et al. Undiagnosed and 'overdiagnosed' COPD using postbronchodilator spirometry in primary healthcare settings: a systematic review and meta-analysis. BMJ Open Respir Res 2023 Apr;10(1):e001478 [FREE Full text] [doi: 10.1136/bmjresp-2022-001478] [Medline: 37130651]

Abbreviations

DM: diabetes mellitus
FEV1%: percentage of predicted forced expiratory volume in 1 second
FEV1: forced expiratory volume in 1 second
FVC%: percentage of predicted forced vital capacity
FVC: forced vital capacity
HDL: high-density lipoprotein
MD: mean difference
MetS: metabolic syndrome
MLD: mixed lung disease
OLD: obstructive lung disease
RLD: restrictive lung disease
WC: waist circumference

Edited by A Mavragani, T Sanchez; submitted 22.10.22; peer-reviewed by J Edwards, UK Chalwadi, GP Arutyunov; comments to author 15.03.23; revised version received 17.05.23; accepted 18.07.23; published 05.09.23.

Please cite as:
Molina-Luque R, Molina-Recio G, de-Pedro–Jiménez D, Álvarez Fernández C, García–Rodríguez M, Romero-Saldaña M
The Impact of Metabolic Syndrome Risk Factors on Lung Function Impairment: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e43737
URL: https://publichealth.jmir.org/2023/1/e43737
doi:10.2196/43737
PMID:37669095

Original Paper

Prevalence, Regional Variations, and Predictors of Overweight, Obesity, and Hypertension Among Healthy Reproductive-Age Indian Women: Nationwide Cross-Sectional Polycystic Ovary Syndrome Task Force Study

Mohd Ashraf Ganie[1*], MBBS, MD, DM[‡]; Subhankar Chowdhury[2*], MBBS, MD, DM; Vanita Suri[3*], MBBS, MD; Beena Joshi[4*], MSc, PhD; Prasanta Kumar Bhattacharya[5*], MD, PhD; Sarita Agrawal[6*], MBBS, MD; Neena Malhotra[7*], MBBS, MD; Rakesh Sahay[8*], MD, DM; Puthiyaveettil Khadar Jabbar[9*], MD, DM; Roya Rozati[10*], MBBS, MD; Imtiyaz Ahmad Wani[1*], MSc, PhD; Amlin Shukla[11*], MD; Taruna Arora[11*], PhD; Haroon Rashid[1*], PhD

[1]Department of Endocrinology, Sher-i-Kashmir Institute of Medical Sciences, Srinagar, India

[2]Department of Endocrinology, Institute of Postgraduate Medical Education & Research, Kolkata, India

[3]Department of Obstetrics & Gynaecology, Postgraduate Institute of Medical Education and Research, Chandigarh, India

[4]Department of Operational Research, National Institute for Research in Reproductive Health, Indian Council of Medical Research (ICMR), Mumbai, India

[5]Department of General Medicine, North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, Meghalaya, India

[6]Department of Obstetrics & Gynaecology, All India Institute of Medical Sciences, Chatisgarh, India

[7]Department of Obstetrics & Gynaecology, All India Institute of Medical Sciences, New Delhi, India

[8]Department of Endocrinology, Osmania Medical College, Hyderabad, India

[9]Department of Endocrinology, Government Medical College, Kerala, India

[10]Department of Obstetrics & Gynaecology, Maternal Health & Research Trust, Hyderabad, India

[11]Reproductive Biology and Maternal Health, Indian Council of Medical Research (ICMR), New Delhi, India

[‡]ICMR PCOS Task Force Study Group

[*]all authors contributed equally

Corresponding Author:
Mohd Ashraf Ganie, MBBS, MD, DM
Department of Endocrinology, Sher-i-Kashmir Institute of Medical Sciences
8-JKB Soura Link
Ward 1P, Room No. 118
Srinagar, 190011
India
Phone: 91 9419041546
Fax: 91 2405501
Email: ashraf.endo@gmail.com

Abstract

Background: A clear understanding of the anthropometric and sociodemographic risk factors related to BMI and hypertension categories is essential for more effective disease prevention, particularly in India. There is a paucity of nationally representative data on the dynamics of these risk factors, which have not been assessed among healthy reproductive-age Indian women.

Objective: This cross-sectional polycystic ovary syndrome (PCOS) task force study aimed to assess the anthropometric and sociodemographic characteristics of healthy reproductive-age Indian women and explore the association of these characteristics with various noncommunicable diseases.

Methods: We conducted a nationwide cross-sectional survey from 2018 to 2022 as part of the Indian Council of Medical Research–PCOS National Task Force study, with the primary aim of estimating the national prevalence of PCOS and regional phenotypic variations among women with PCOS. A multistage random sampling technique was adopted, and 7107 healthy women (aged 18-40 years) from 6 representative geographical zones of India were included in the study. The anthropometric indices and sociodemographic characteristics of these women were analyzed. Statistical analysis was performed to assess the association between exposure and outcome variables.

Results: Of the 7107 study participants, 3585 (50.44%) were from rural areas and 3522 (49.56%) were from urban areas. The prevalence of obesity increased from 8.1% using World Health Organization criteria to 40% using the revised consensus guidelines for Asian Indian populations. Women from urban areas showed higher proportions of overweight (524/1908, 27.46%), obesity (775/1908, 40.62%), and prehypertension (1008/1908, 52.83%) categories. A rising trend of obesity was observed with an increase in age. Women aged 18 to 23 years were healthy (314/724, 43.4%) and overweight (140/724, 19.3%) compared with women aged 36 to 40 years with obesity (448/911, 49.2%) and overweight (216/911, 23.7%). The proportion of obesity was high among South Indian women, with 49.53% (531/1072) and 66.14% (709/1072), using both World Health Organization criteria and the revised Indian guidelines for BMI, respectively. BMI with waist circumference and waist-to-height ratio had a statistically significant linear relationship (r=0.417; P<.001 and r=0.422; P<.001, respectively). However, the magnitude, or strength, of the association was relatively weak ($0.3<|r|<0.5$). Statistical analysis showed that the strongest predictors of being overweight or obese were older age, level of education, wealth quintile, and area of residence.

Conclusions: Anthropometric and sociodemographic characteristics are useful predictors of overweight- and obesity-related syndromes, including prehypertension, among healthy Indian women. Increased attention to the health of Indian women from public health experts and policy makers is warranted. The findings of this study can be leveraged to offer valuable insights, informing health decision-making and targeted interventions that mitigate risk factors of overweight, obesity, and hypertension.

International Registered Report Identifier (IRRID): RR2-10.2196/23437

(JMIR Public Health Surveill 2023;9:e43199) doi:10.2196/43199

KEYWORDS

Indian Council of Medical Research–polycystic ovary syndrome study; ICMR-PCOS study; noncommunicable diseases; disease surveillance; obesity; hypertension; anthropometry; healthy reproductive-age women; socioeconomic status

Introduction

Background

Over the past few decades, a disturbing epidemiological transition has revealed that chronic noncommunicable diseases (NCDs) are on the rise both at the global and regional levels. The etiology of NCDs has remained vague and has been attributed to many factors, ranging from genetic predispositions to human lifestyles. The lack of physical exercise, high BMI, increased blood pressure (BP), and obesity are a few factors that affect people's health and predispose them to the risk of NCDs [1,2]. Obesity has now emerged as a major worldwide health concern and is probably the main cause of the deranged metabolism leading to metabolic syndrome, including insulin resistance (IR), type 2 diabetes mellitus, hypertension, and dyslipidemia [3,4]. Scientific evidence reveals that women with obesity are more vulnerable to infertility, miscarriage, and other reproductive complications [5]. Researchers have brought to light the fact that a fast-paced lifestyle has taken a toll on human health and is a contributor to the rise of various diseases and disorders. The issue of obesity and overweight has increased to epidemic proportions; one-third of the global population is overweight or obese, and according to the global burden of disease, >4 million people die each year as a result of these issues. According to the Indian Council of Medical Research (ICMR)–India Diabetes study 2015, the prevalence rates of obesity and central obesity vary from 11.8% to 31.3% and from 16.9% to 36.3%, respectively [6]. Another cross-sectional study reported an age-adjusted prevalence of hypertension of 10.9% in women, 12.5% in urban areas, and 10.6% in rural areas [7]. Recent reports from the World Health Organization (WHO) revealed that, among adults aged ≥18 years, 13% are obese and 39% are overweight [8]. The 2019 to 2021 National Family Health Survey (NFHS)–5 anthropometric data revealed that 19% of women aged between 15 and 49 years are underweight, 24% are overweight or obese, and 57% have a BMI within the normal range [9].

A global strategy for improving health, early detection, and prevention of disease epidemics can be developed by adopting appropriate, accessible, affordable, and scalable health solutions. Anthropometry is the single most universally applicable, inexpensive, noninvasive, and reliable method available for the assessment of body composition, size, and proportions among other methods such as biochemical and radiological assessments. Anthropometric indices are being used to predict the increased risk of several diseases at both the individual and population levels [10-12]. These indices are combinations of anthropometric measurements (BMI, waist circumference [WC], and waist-to-hip ratio [WHR]) that are vital for the interpretation of measurements and assessment of NCD risk factors in the population [13,14]. The most frequently used anthropometric indices are BMI and WC as predictors of overweight- and obesity-related syndromes, including hypertension. The existing literature includes waist-to-height ratio (WHtR) as an additional anthropometric indicator for the assessment of overweight- and obesity-related syndromes, including hypertension [9,15,16]. Several epidemiological studies have consistently reported an association between hypertension and BMI, WC, WHR, and WHtR [17,18].

The correlation between obesity and hypertension is a major concern worldwide. The prevalence of obesity in India is not uniform and varies according to age, gender, geographical environment, and socioeconomic status (SES). Furthermore, many sociodemographic indicators such as education, SES, and occupation and ethnic and racial differences are linked to health status through multiple pathways; however, their clear relationships have not yet been fully elucidated and hitherto remain obscure. The NFHS conducted in India has depicted substantial differences using WHO criteria in the prevalence of overweight and obesity among adults across regions, implying the role of geographic and regional factors. However, there is a paucity of similar literature among reproductive age–matched healthy women using revised consensus guidelines for Indian populations. Myriad previous studies have revealed that the prevalence of prehypertension is an important risk and predisposing factor for developing hypertension in the future [19]. Making use of Western anthropometric standards for low- and middle-income countries (LMICs) has remained controversial [20]; hence, country-specific anthropometric data become indispensable in this regard. It has been very well acknowledged that no standard is universally acceptable as the definition of indices such as optimal growth differs with ethnicity and genetic makeup. In addition, these anthropometric parameters are not uniform and vary considerably from population to population. However, these anthropometric indicators have high predictive value for assessing risk factors of NCDs, thus offering a robust mechanism for the identification of risk factors in relatively easy and cost-effective ways in a local setting.

Objectives

The pattern of overall lifestyle varies from region to region, which necessitates a comprehensive national-level research study, especially among young, healthy reproductive-age women across diverse regions of the country. India has called for the highest priority for the health needs of women, who constitute nearly 50% of the country's population. Taking a cue, the ICMR has taken an initiative under the ICMR–polycystic ovary syndrome (PCOS) Task Force study to address women-centric issues, wherein attention has also been paid to the assessment of various aberrations among women from representative regions of India using anthropometric indices. As the prevalence of overweight, obesity, and hypertension is rising in LMICs such as India, it becomes imperative to understand the dynamics of the various sociodemographic factors and their association with the incidence of overweight, obesity, and BP profile among age-matched healthy women from representative regions of India. With this endeavor, this study attempted to address the gaps in NFHS anthropometric data using nationally representative data through the adoption of revised country-specific guidelines.

Methods

Setting

We used data from the ICMR-PCOS Task Force study conducted across representative regions of India [21]. The data were collected from both rural and urban areas in 6 geographical zones of the country.

Ethics Approval

The study was conducted in accordance with the Helsinki Declaration of 1975 and has been approved by the institutional ethics committees at all 10 study sites as: 131/IEC-SKIMS/2017-101- Sher-i-Kashmir Institute of Medical Sciences, Srinagar; PGI/IEC/2017/47-Postgraduate Institute of Medical Education and Research, Chandigarh;IEC-34/09.02.2017- All India Institute of Medical Sciences, Delhi; IEC/2017/057- Institute of Post Graduate Medical Education and Research, Kolkata; NEIGR/IEC/2018/02- North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences; ECR/300/Inst/AP/2017-Osmania Medical College, Hyderabad; IEC/MHRT/302- Maternal Health and Research Trust, Hyderabad;173/IEC-AIIMSRPR/2017- All India Institute of Medical Sciences, Raipur; D/ICEC/Sci-33/37/2017- National Institute for Research in Reproductive Health, Mumbai and IEC.No.07/09/2017/MCT- Government Medical College, Trivandrum.

Informed Consent and Safety Measures

Written informed consent was obtained from the participants, ensuring their understanding of the purpose of the study and the obligations and possible consequences of their participation. During the data collection process, face masks, hand gloves, and hand sanitizer were provided to the data collection team and study participants to protect the data collectors and participants from COVID-19 infection.

Study Sites

This study involved the participation of the following sites: All India Institute of Medical Sciences, Delhi; All India Institute of Medical Sciences, Raipur; Government Medical College, Trivandrum; Institute of Post Graduate Medical Education and Research, Kolkata; Maternal Health and Research Trust, Hyderabad; North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, Shillong; National Institute for Research in Reproductive Health, Mumbai; Osmania Medical College, Hyderabad; Postgraduate Institute of Medical Education and Research, Chandigarh; and Sher-i-Kashmir Institute of Medical Sciences, Srinagar.

Sampling Design

This multicenter study involved the recruitment of apparently healthy women aged 18 to 40 years using a multistage sampling technique from randomly selected polling booths across urban and rural areas from the 6 zones (ie, North, South, East, West, Northeast, and Central India) of the country. The states and union territories were grouped into 6 geographical regions based on their sociocultural similarities, with a small change from the grouping (Table 1 shows the composition of each zone). Details of the sample size and sampling plan are available in our published protocol in *JMIR Research Protocols* [21].

Table 1. States and union territories in each region.

Region or zone	State and UT[a]
East	West Bengal
West	Maharashtra
North	Jammu and Kashmir (UT), Chandigarh (UT), and Delhi (UT)
Northeast	Meghalaya
Central	Chhattisgarh
South	Telangana

[a] UT: Union Territory.

Study Population

We approached a total of 13,166 women as per the voter ID list, of whom 3137 (23.83%) were either not eligible or refused to give consent and were excluded. Of the total 10,029 women, 2922 (29.14%) were excluded using the inclusion and exclusion criteria (Textbox 1). A total of 7107 women were enrolled as apparently healthy controls with complete sociodemographic details, and after considering the missing clinical information, our analysis included 3877 (54.55%) apparently healthy women from representative regions of the country (Figure 1).

Textbox 1. Inclusion and exclusion criteria.

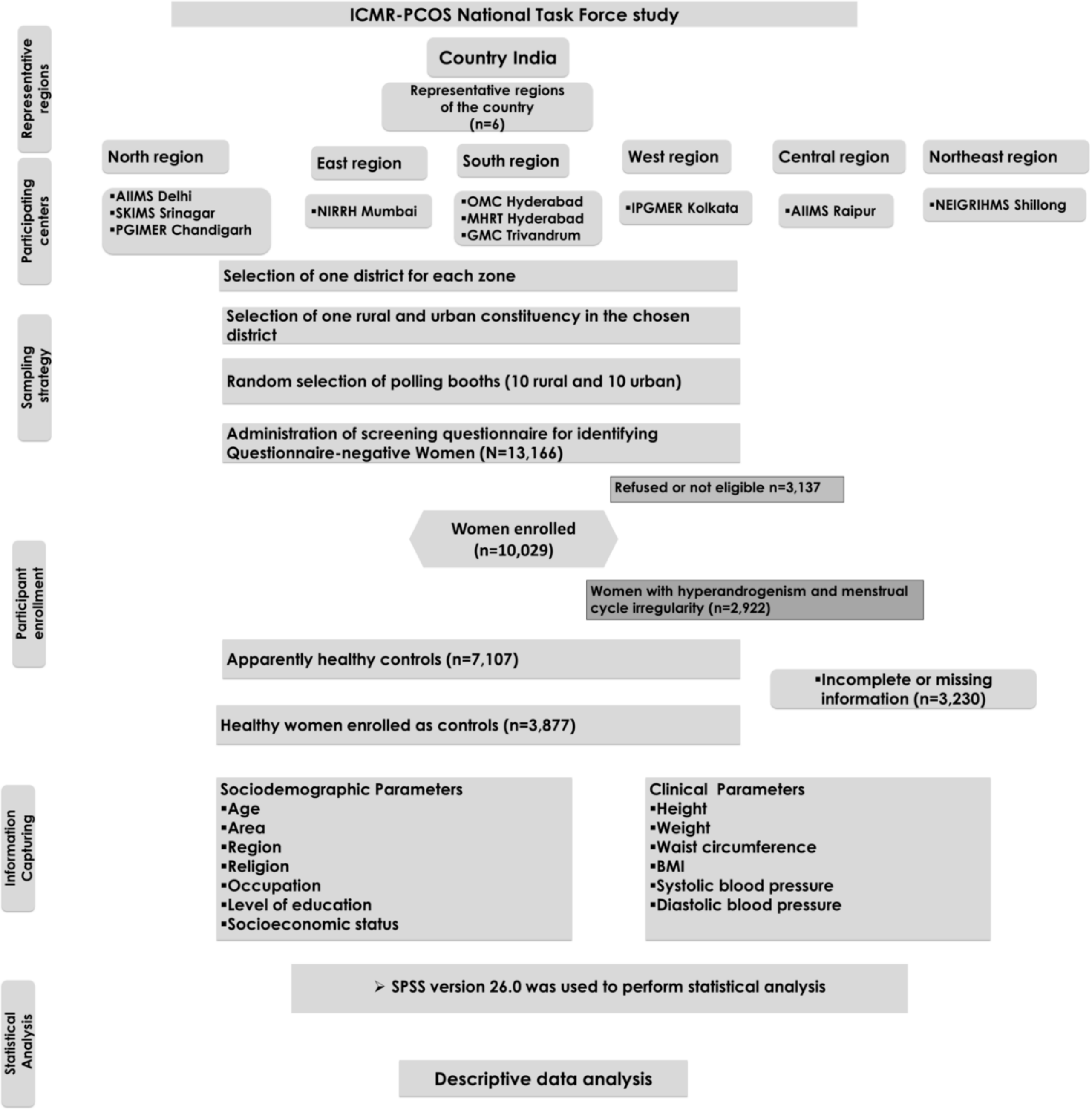
Inclusion criteria

- Healthy women aged between 18-40 years
- Women who were permanent residents of that area (>1 y)
- Willingness to participate in the study and sign an informed consent form

Exclusion criteria

- Participants with hypothyroidism, subclinical hypothyroidism, hyperthyroidism, hyperprolactinemia, type 2 diabetes, exogenous Cushing syndrome, premature ovarian failure, and hypopituitarism
- Pregnant or lactating women and those with cognitive or physical limitations that prevented them from answering the questionnaire

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram showing selection of participants in the cross-sectional polycystic ovary syndrome (PCOS) Task Force study. AIIMS: All India Institute of Medical Sciences; GMC: Government Medical College; ICMR: Indian Council of Medical Research; IPGMER: Institute of Post Graduate Medical Education and Research; MHRT: Maternal Health and Research Trust; NEIGRIHMS: North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences; NIRRH: National Institute for Research in Reproductive Health; OMC: Osmania Medical College; PGIMER: Postgraduate Institute of Medical Education and Research; SKIMS: Sher-i-Kashmir Institute of Medical Sciences.



Data Collection

Well-trained research staff collected the requisite data as appropriate at each study site. After obtaining written consent, an in-depth questionnaire was presented to each woman enrolled from the community (rural and urban) to obtain demographic information about her age, religion, marital status, level of education, SES, occupation, and type of family. Anthropometric data collected included height, weight, WHR, and WC. BP readings were taken at 5-minute intervals. The junior medical officers were responsible for the overall management of the field team at all sites. The investigators at all participating centers monitored the data collection process and checked the data during entry.

Tools and Techniques

Sociodemographic Information

Sociodemographic information pertaining to age, sex, address, occupation, education, SES, marital status, and religion was obtained using a paper-based questionnaire. Information related to anthropometric and related parameters in study participants was obtained through in-person physical evaluation during community visits.

Anthropometric Measurements

Participants included in the study underwent a range of anthropometric measurements in accordance with the instruction manual designed before the initiation of the study. Trained research staff were engaged to take the measurements and record the clinical data from all participants, including age, body weight (kg), height (cm), BMI (kg/m^2), WHR, and BP. The anthropometric measurements were performed using anthropometric tools (SECA 213, 813, and 203 scales) with similar specifications at all participating sites. The equipment was maintained and calibrated as per the manufacturer's instructions. BMI was calculated using the following formula: $BMI = weight\ (kg) / height^2\ (m^2)$. BP was measured 3 times with an automated sphygmomanometer in a sitting position after at least 5 minutes of rest.

Height was measured using a portable stadiometer (SECA 213) with precision nearest to 0.1 cm. Before measurement, the stadiometer was checked using calibration rods. During measurement, the participants were asked to take off their shoes. To ensure correct measurements, the shoulder, buttocks, calf, heel, and back of the head of each participant were to touch the vertical stand of the stadiometer.

Body weight was measured using a digital scale (SECA 813) with a precision of 0.1 kg. The accuracy of the scale was checked by weighing an object of known weight. The participants were asked to stand straight on the electronic weighing scale barefoot and with light clothing, and the weight displayed on the screen was recorded.

WC was measured using nonstretchable measuring tape (SECA), and the participants were asked to raise their clothing above the waist to avoid any interference with the measurements. WC was measured on bare skin with the participant in a standing position and at the narrowest indentation at the midpoint between the costal margin and iliac crest.

BP was measured using an automatic BP monitor (OMRON HEM-7124). To ensure adherence to WHO criteria for the measurement of BP, the participants were asked to sit on a chair and allowed to rest as required with their feet on the ground and their palm facing upward. All BP measurements were performed preferably on the left arm, ensuring that the cuff was wrapped over the bare upper arm, 2 to 3 cm above the antecubital fossa. The average of 3 consecutive measurements was taken as the BP of the participant.

Categorization of Variables

The primary outcome variables of our study included prevalence of overweight, obesity, and hypertension. The variables used were age; place of residence; geographical region; marital status; occupation; SES; and biological risk factors such as BMI, WC, and WHtR of the study participants. BMI was calculated as the ratio of weight (in kg) to the square of height (in m; $BMI = weight / height^2$). For the WHR, WC was divided by hip circumference. BMI was categorized as healthy, underweight, overweight, and obese. According to the WHO classification of obesity, BMI is categorized as healthy (<24.9 kg/m^2), overweight (25-29.9 kg/m^2), and obese (>30 kg/m^2) [22]. However, in this study, BMI was calculated based on the revised consensus guidelines for India [23,24]. Using this scale, the participants were categorized as underweight (<18.5 kg/m^2), healthy (18.5-22.9 kg/m^2), overweight (23.0-24.9 kg/m^2), and obese (≥25 kg/m^2).

The explanatory variables were selected based on published reports and the data set's structure. The study collected information pertaining to the participants' age (18-23, 24-29, 30-35, and 36-40 y), place of residence (urban and rural), geographical location (East, West, North, South, Northeast, and Central India), marital status, educational level (ie, illiterate, up to primary school, up to secondary school, up to high school, and above), SES using the BG Prasad scale (eg, lower class, middle class, lower middle class, and upper class), and religion (ie, Hindu, Muslim, Christian, Sikh, and others).

Definition and Categorization of Hypertension

Hypertension was defined according to the Eighth Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 8) criteria. All the participants with a systolic BP (SBP) of ≥140 mm Hg or a diastolic BP (DBP) of ≥90 mm Hg or who reported taking medication at present for the treatment of high BP were classified as having hypertension. As per the JNC 8, (1) stage-1 hypertension was defined as an SBP of 140 to 159 mm Hg or DBP of 90 to 99 mm Hg, and (2) prehypertension was defined as an SBP of >120 mm Hg but less than the cutoff for hypertension or DBP between 80 and 89 mm Hg in the absence of a diagnosis of hypertension or treatment with medication for high BP [23].

Statistical Analysis

The data from all study participants were entered into Excel (Microsoft Corp) sheets, and we used SPSS (version 26.0; IBM Corp) for data analysis. Descriptive statistical analysis was performed, and categorical variables were stratified into groups and then presented as numbers and percentages. Pearson correlation analysis was performed between BMI and WC and WHtR. A logistic regression model was used to determine the influence of demographic characteristics on general and central obesity. Variables included in the multivariate logistic regression analysis were age, region, SES, education, religion, marital status, and occupation. The data on the distribution of BMI categories using WHO criteria and the revised consensus guidelines for India were presented using bar charts. A 2-tailed P value of <.05 was considered statistically significant.

Results

Description

Apparently healthy women who met the inclusion criteria were enrolled as controls in the study. The sociodemographic information (age, sex, address, occupation, education, SES, marital status, and religion) of the study participants is shown in Table 2. Of the 7107 study participants, 3585 (50.44%) were from rural areas and 3522 (49.56%) were from urban areas. Most of the women (2323/7107, 32.69%) were aged 30 to 35 years, followed by 24 to 29 years (1869/7107, 26.3%), 36 to 40 years (1609/7107, 22.64%), and 18 to 23 years (1306/7107, 18.38%). Most women (4851/7107, 68.26%) were Hindu, followed by Muslim (1300/7107, 18.29%), Christian (837/7107, 11.78%), Sikh (88/7107, 1.24%), and other religions (31/7107, 0.44%). Approximately 73.8% (5245/7107) of women were married, and 24.22% (1721/7107) were unmarried, with approximately 1.98% (141/7107) of women being either widowed or un-remarried. Among the enrolled women, 22.37% (1590/7107) had a high school education, 18.62% (1323/7107) had a middle school education, 20.22% (1437/7107) had a postgraduate education, only 2.41% (171/7107) had received a professional education or honors, and 14.65% (1041/7107) were uneducated. Furthermore, 82.43% (5858/7107) of women were unemployed or unskilled, and 2.41% (171/7107) were in some profession. Most of them (2176/7107, 30.62%) belonged to the middle class, followed by the lower-middle class (1721/7107, 24.22%). Of all the enrolled women, 13.02% (925/7107) belonged to the upper class, 22.39% (1591/7107) belonged to the upper-middle class, and 9.62% (684/7107) belonged to the lower class.

Table 2. Sociodemographic characteristics of the study population from representative regions of the country (n=7107).

Variable	North (n=2216), n (%)	Northeast (n=829), n (%)	East (n=744), n (%)	Central (n=681), n (%)	South (n=1690), n (%)	West (n=947), n (%)	Total, n (%)
Age group (years)							
18-23	441 (19.9)	218 (26.3)	120 (16.1)	218 (32)	252 (14.9)	57 (6)	1306 (18.4)
24-29	496 (22.4)	267 (32.2)	199 (26.7)	203 (29.8)	481 (28.5)	223 (23.5)	1869 (26.3)
30-35	765 (34.5)	221 (26.7)	260 (34.9)	198 (29.1)	530 (31.4)	349 (36.9)	2323 (32.7)
36-40	514 (23.2)	123 (14.8)	165 (22.2)	62 (9.1)	427 (25.3)	318 (33.6)	1609 (22.6)
Area of residence							
Rural	1152 (52)	431 (52)	391 (52.6)	294 (43.2)	822 (48.6)	495 (52.3)	3585 (50.4)
Urban	1064 (48)	398 (48)	353 (47.4)	387 (56.8)	868 (51.4)	452 (47.7)	3522 (49.6)
Religion							
Hindu	1314 (59.3)	123 (14.8)	481 (64.7)	638 (93.7)	1390 (82.2)	905 (95.6)	4851 (68.3)
Muslim	817 (36.9)	9 (1.1)	263 (35.3)	25 (3.7)	164 (9.7)	22 (2.3)	1300 (18.3)
Christian	3 (0.1)	679 (81.9)	0 (0)	12 (1.8)	136 (8)	7 (0.7)	837 (11.8)
Sikh	82 (3.7)	0 (0)	0 (0)	4 (0.6)	0 (0)	2 (0.2)	88 (1.2)
Other	0 (0)	18 (2.2)	0 (0)	2 (0.3)	0 (0)	11 (1.2)	31 (0.4)
Marital status							
Unmarried	609 (27.5)	317 (38.2)	113 (15.2)	226 (33.2)	236 (14)	220 (23.2)	1721 (24.2)
Married	1585 (71.5)	473 (57.1)	630 (84.7)	447 (65.6)	1406 (83.2)	704 (74.3)	5245 (73.8)
Un-remarried	12 (0.5)	27 (3.3)	0 (0)	2 (0.3)	12 (0.7)	13 (1.4)	66 (0.9)
Widowed	10 (0.5)	12 (1.4)	1 (0.1)	6 (0.9)	36 (2.1)	10 (1.1)	75 (1.1)
Level of education							
Illiterate	348 (15.7)	25 (3)	274 (36.8)	28 (4.1)	221 (13.1)	145 (15.3)	1041 (14.6)
Primary school certificate	188 (8.5)	97 (11.7)	192 (25.8)	50 (7.3)	106 (6.3)	78 (8.2)	711 (10)
Middle school certificate	368 (16.6)	187 (22.6)	243 (32.7)	137 (20.1)	233 (13.8)	155 (16.4)	1323 (18.6)
High school certificate	546 (24.6)	208 (25.1)	24 (3.2)	158 (23.2)	481 (28.5)	173 (18.3)	1590 (22.4)
Intermediate or post–high school diploma	221 (10)	132 (15.9)	8 (1.1)	127 (18.6)	231 (13.7)	112 (11.8)	831 (11.7)
Graduate or postgraduate	531 (24)	174 (21)	3 (0.4)	178 (26.1)	296 (17.5)	255 (26.9)	1437 (20.2)
Profession or honors	14 (0.6)	3 (0.4)	0 (0)	3 (0.4)	122 (7.2)	29 (3.1)	171 (2.4)
Occupation							
Unemployed	1851 (83.5)	517 (62.4)	133 (17.9)	537 (78.9)	1177 (69.6)	563 (59.5)	4778 (67.2)
Unskilled worker	71 (3.2)	211 (25.5)	388 (52.2)	67 (9.8)	208 (12.3)	135 (14.3)	1080 (15.2)
Semiskilled worker	74 (3.3)	43 (5.2)	212 (28.5)	51 (7.5)	110 (6.5)	81 (8.6)	571 (8)
Skilled worker	104 (4.7)	15 (1.8)	11 (1.5)	21 (3.1)	86 (5.1)	72 (7.6)	309 (4.3)
Arithmetic skill job	8 (0.4)	0 (0)	0 (0)	2 (0.3)	27 (1.6)	25 (2.6)	62 (0.9)
Semiprofessional	24 (1.1)	13 (1.6)	0 (0)	3 (0.4)	42 (2.5)	27 (2.9)	109 (1.5)
Professional	64 (2.9)	23 (2.8)	0 (0)	0 (0)	40 (2.4)	44 (4.6)	171 (2.4)

Variable	North (n=2216), n (%)	Northeast (n=829), n (%)	East (n=744), n (%)	Central (n=681), n (%)	South (n=1690), n (%)	West (n=947), n (%)	Total, n (%)
Other	20 (0.9)	7 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	27 (0.4)
Socioeconomic status							
Upper class	180 (8.1)	5 (0.6)	17 (2.3)	163 (23.9)	333 (19.7)	227 (24)	925 (13)
Upper-middle class	431 (19.4)	15 (1.8)	466 (62.6)	204 (30)	265 (15.7)	210 (22.2)	1591 (22.4)
Middle class	541 (24.4)	320 (38.6)	234 (31.5)	180 (26.4)	714 (42.2)	187 (19.7)	2176 (30.6)
Lower-middle class	601 (27.1)	488 (58.9)	27 (3.6)	105 (15.4)	289 (17.1)	211 (22.3)	1721 (24.2)
Lower class	463 (20.9)	1 (0.1)	0 (0)	29 (4.3)	89 (5.3)	112 (11.8)	694 (9.8)
Type of family							
Living alone	70 (3.2)	3 (0.4)	0 (0)	10 (1.5)	57 (3.4)	7 (0.7)	147 (2.1)
Nuclear	1346 (60.7)	723 (87.2)	569 (76.5)	380 (55.8)	997 (59)	638 (67.4)	4653 (65.5)
Joint	734 (33.1)	103 (12.4)	172 (23.1)	287 (42.1)	596 (35.3)	297 (31.4)	2189 (30.8)
Extended	64 (2.9)	0 (0)	3 (0.4)	4 (0.6)	40 (2.4)	5 (0.5)	116 (1.6)
Other	2 (0.1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (0)

In addition, we observed that the prevalence of overweight, obesity, and hypertension among women showed an association with sociodemographic and other characteristics. We analyzed how overweight, obesity, and prehypertension varied when nonmodifiable factors such as age and modifiable factors such as educational level were considered in accordance with the revised consensus guidelines for India. A rising trend of obesity was observed with an increase in age. Women aged between 36 and 40 years (448/911, 49.2%) were more obese compared with those in other age groups. As expected, we observed the lowest percentage of obesity among women aged 18 to 23 years. The incidence of overweight was highest (341/1289, 26.45%) among women aged between 30 and 35 years and lowest among women aged between 18 and 23 years. In addition, we found a lower percentage of prehypertension (324/1946, 16.65%) among younger women aged between 18 and 23 years than in those aged between 30 and 35 years.

Using the revised consensus guidelines for India, we found varying proportions of BMI categories and prehypertension across rural and urban areas. Urban areas showed the highest proportions of women with obesity (775/1908, 40.62%), overweight (524/1908, 27.46%), and prehypertension (1008/1908, 52.83%) compared with women from rural areas. The proportion of underweight showed an almost decreasing trend with an increase in SES and vice versa for obesity rates. Upper-class and upper–middle-class women presented the lowest rates of underweight (17/308, 5.5% and 25/740, 3.4%, respectively). However, middle-class women presented a high incidence of underweight, obesity, and prehypertension (80/1331, 6.01%; 552/1331, 41.47%; and 738/1946, 37.92%, respectively). Women with a higher level of education (professional or honors) showed the lowest rates of overweight (6/34, 18%) and prehypertension (14/1946, 0.72%), and none were underweight. Women with a middle school education and illiterate women had higher obesity rates (299/665, 45% and 269/614, 43.8%, respectively). Women who had received a high school, graduate, or postgraduate education showed high rates of prehypertension (482/1946, 24.77% and 359/1946, 18.45%, respectively). Predictably, we observed high percentages of prehypertension (1409/1946, 72.4%) and overweight (706/2836, 24.89%) among unemployed women and the lowest rates of prehypertension (31/1946, 1.59%) among women who were either semiprofessional or professional (Table 3).

Table 4 shows the pattern of BMI and hypertension categories in each zone stratified by BMI, central obesity (WHtR and WC), and BP profiles. The comparative distribution of BMI categories of the 3877 women using the WHO and revised consensus guidelines for India are shown in Tables 4 and 5, in which regional variations are quite evident. Using WHO criteria and the revised consensus guidelines for India, the highest proportion of healthy women was observed in the Northeast (126/164, 76.8% and 99/164, 60.4%, respectively), and the proportion of overweight women was 42.1% (59/140) and 27.9% (39/140) in the East region, respectively. Accordingly, we found the highest rates of obesity in southern India (151/1072, 14.09% using the WHO criteria and 531/1072, 49.53% using the revised consensus BMI guidelines for India).

Table 3. Association between sociodemographic variables and BMI categories stratified by the revised consensus guidelines for India and stages of hypertension.

Variable	Underweight (n=204), n (%)	Healthy (n=1180), n (%)	Overweight (n=941), n (%)	Obese (n=1552), n (%)	Prehypertension (n=1946), n (%)
Age group (years)					
18-23	72 (9.9)	314 (43.4)	140 (19.3)	198 (27.3)	324 (16.6)
24-29	63 (6.6)	303 (31.8)	244 (25.6)	343 (36)	484 (24.9)
30-35	45 (3.5)	340 (26.4)	341 (26.5)	563 (43.7)	673 (34.6)
36-40	24 (2.6)	223 (24.5)	216 (23.7)	448 (49.2)	465 (23.9)
Area of residence					
Rural	128 (6.5)	647 (32.9)	417 (21.2)	777 (39.5)	938 (48.2)
Urban	76 (4)	533 (27.9)	524 (27.5)	775 (40.6)	1008 (52.8)
Religion					
Hindu	109 (4.2)	747 (28.5)	677 (25.8)	1086 (41.5)	1323 (68)
Muslim	73 (7.6)	308 (32)	199 (20.7)	383 (39.8)	525 (27)
Christian	20 (9.7)	102 (49.3)	33 (15.9)	52 (25.1)	53 (2.7)
Sikh	2 (2.4)	19 (22.9)	32 (38.6)	30 (36.1)	43 (2.2)
Other	0 (0)	4 (80)	0 (0)	1 (20)	2 (0.1)
Level of education					
Illiterate	38 (6.2)	161 (26.2)	146 (23.8)	269 (43.8)	379 (19.5)
Primary school certificate	22 (6)	95 (25.8)	93 (25.3)	158 (42.9)	198 (10.2)
Middle school certificate	31 (4.7)	180 (27.1)	155 (23.3)	299 (45)	321 (16.5)
High school certificate	55 (5.8)	312 (32.8)	223 (23.4)	361 (38)	482 (24.8)
Intermediate or post–high school diploma	32 (7.1)	155 (34.4)	104 (23.1)	160 (35.5)	193 (9.9)
Graduate or postgraduate	26 (3.3)	263 (33.1)	214 (27)	291 (36.6)	359 (18.4)
Profession or honors	0 (0)	14 (41.2)	6 (17.6)	14 (41.2)	14 (0.7)
Socioeconomic status					
Upper class	17 (5.5)	87 (28.2)	79 (25.6)	125 (40.6)	133 (6.8)
Upper-middle class	25 (3.4)	226 (30.5)	187 (25.3)	302 (40.8)	353 (18.1)
Middle class	80 (6)	366 (27.5)	333 (25)	552 (41.5)	738 (37.9)
Lower-middle class	48 (4.9)	331 (34)	222 (22.8)	372 (38.2)	461 (23.7)
Lower class	34 (6.5)	170 (32.4)	120 (22.9)	201 (38.3)	261 (13.4)
Occupation					
Unemployed	152 (5.4)	878 (31)	706 (24.9)	1100 (38.8)	1409 (72.4)
Unskilled worker	19 (4.7)	102 (25.2)	98 (24.3)	185 (45.8)	218 (11.2)
Semiskilled worker	19 (7.5)	84 (32.9)	53 (20.8)	99 (38.8)	137 (7)
Skilled worker	8 (4.3)	57 (30.6)	37 (19.9)	84 (45.2)	93 (4.8)
Arithmetic skill job	0 (0)	9 (33.3)	5 (18.5)	13 (48.1)	16 (0.8)
Semiprofessional	4 (6.9)	19 (32.8)	13 (22.4)	22 (37.9)	31 (1.6)
Professional	2 (2.2)	27 (30)	25 (27.8)	36 (40)	31 (1.6)

Table 4. Distribution of BMI stratified by the revised consensus guidelines for India and the World Health Organization (WHO) criteria along with stratified analysis of hypertension and various anthropometric indices of the study population.

Category and variable	North (n=2214), n (%)	Northeast (n=164), n (%)	East (n=140), n (%)	Central (n=126), n (%)	South (n=1072), n (%)	West (n=161), n (%)
BMI category (revised Indian guidelines)						
Underweight (<18.5 kg/m^2)	96 (4.3)	13 (7.9)	3 (2.1)	15 (11.9)	56 (5.2)	21 (13)
Healthy (18.5-22.9 kg/m^2)	695 (31.4)	99 (60.4)	36 (25.7)	55 (43.7)	240 (22.4)	55 (34.2)
Overweight (23-24.9 kg/m^2)	582 (26.3)	27 (16.5)	39 (27.9)	19 (15.1)	245 (22.9)	29 (18)
Obese (≥25 kg/m^2)	841 (38)	25 (15.2)	62 (44.3)	37 (29.4)	531 (49.5)	56 (34.8)
BMI category (WHO criteria)						
Underweight (<18.5 kg/m$^{2)}$	96 (4.3)	13 (7.9)	3 (2.1)	15 (11.9)	56 (5.2)	21 (13)
Healthy (18.5-24.9 kg/m^2)	1277 (57.7)	126 (76.8)	75 (53.6)	74 (58.7)	485 (45.2)	84 (52.2)
Overweight (25-29.9 kg/m^2)	699 (31.6)	25 (15.2)	59 (42.1)	33 (26.2)	380 (35.4)	44 (27.3)
Obese (≥30 kg/m^2)	142 (6.4)	0 (0)[a]	3 (2.1)	4 (3.2)	151 (14.1)	12 (7.5)
Waist category						
Central_obesity_WHtr[b]	1192 (53.8)	86 (52.4)	27 (19.3)	80 (63.5)	481 (44.9)	97 (60.2)
Central_obesity_WC[c]	498 (22.5)	13 (7.9)	7 (5)	38 (30.2)	164 (15.3)	48 (29.8)
Hypertension category (JNC 8[d] criteria)						
Normotensive	1173 (53)	134 (81.7)	70 (50)	73 (57.9)	363 (33.9)	118 (73.3)
Prehypertension	1041 (47)	30 (18.3)	70 (50)	53 (42.1)	709 (66.1)	43 (26.7)

[a] No participant identified in this category.

[b] WHtR: waist-to-height ratio.

[c] WC: waist circumference.

[d] JNC 8: Eighth Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure.

Table 5. Comparative distribution of BMI categories (obese, overweight, healthy, and underweight) of the women using the World Health Organization (WHO) criteria and the revised consensus guidelines for India (n=3877).

	Participants, n (%)
BMI (WHO)	
Obese	312 (8.05)
Overweight	1240 (31.98)
Healthy	2121 (54.71)
Underweight	204 (5.26)
BMI (Indian)	
Obese	1552 (40.03)
Overweight	941 (24.27)
Healthy	1180 (30.44)
Underweight	204 (5.26)

A clear variation in the frequency of central obesity, WHtR, and WC was noted among women across the regions. The central obesity WHtR and WC were highest among Central Indian women (80/126, 63.5% and 38/126, 30.2%, respectively).

We also analyzed how BP profile varied across regions using JNC 8 criteria. After stratifying by region, our analysis showed that the women who were enrolled as healthy controls in the study either were normotensive (healthy BP profile) or fell under the prehypertension category. The Northeast region (Shillong) showed the highest incidence of normotensive women (134/164, 81.7%) and the lowest rates of women with prehypertension (30/164, 18.3%). However, prehypertension was found to be highest among South Indian women (709/1072, 66.14%), followed by women from East (70/140, 50%), North (1041/2214, 47.02%), and Central (53/126, 42.1%) India, as well as 26.7% (43/161) from the Western region of the country. Table 6 shows the multivariate logistic regression analysis on the association between demographic characteristics (age, region, residence, occupation, SES, education, marital status, and religion of participants) and general and central obesity. We observed that age, region, religion, and occupation were significantly associated with obesity. Obesity was significantly higher in older age groups (36-40 y; odds ratio [OR] 2.57, 95% CI 1.906-3.471; *P*<.001), central regions (OR 1.76, 95% CI 1.144-2.719; *P*=.01), and the Muslim population (OR 2.73, 95% CI 2.245-3.313; *P*<.001).

Table 6. Multivariate logistic regression analysis on the association between demographic characteristics (age, region, residence, occupation, socioeconomic status, education, marital status, and religion of the participants) and general and central obesity.

Characteristic	Central obesity (based on WC[a])		General obesity (based on WHO[b] guidelines)	
	OR[c] (95% CI)	*P* value	OR (95% CI)	*P* value
Age group (years)				
18-23[d]	N/A[e]	N/A	N/A	N/A
24-29	1.199 (0.909-1.582)	.20	1.082 (0.725-1.1615)	.70
30-35	1.574 (1.184-2.091)	.002	1.132 (0.755)-1.699)	.55
36-40	2.572 (1.906-3.471)	<.001	1.334 (0.871-2.044)	.19
Region				
North[d]	N/A	N/A	N/A	N/A
Northeast	0.462 (0.196-1.086)	.08	0 (0-0)	>.99
East	0.180 (0.081-0.400)	<.001	0.274 (0.083-0.899)	.03
Central	1.764 (1.144-2.719)	.01	0.406 (0.144-1.146)	.09
South	0.686 (0.549-0.858)	.001	2.116 (1.587-2.823)	<.001
West	1.284 (0.877-1.879)	.20	0.861 (0.453-1.635)	.65
Residence				
Rural[d]	N/A	N/A	N/A	N/A
Urban	0.757 (0.643-0.890)	.001	0.642 (0.507-0.811)	<.001
Occupation				
Unemployed[d]	N/A	N/A	N/A	N/A
Unskilled worker	0.829 (0.597-1.149)	.26	1.240 (0.840-1.831)	.28
Semiskilled worker	1.101 (0.787-1.540)	.57	0.787 (0.454-1.364)	.39
Skilled worker	1.088 (0.757-1.564)	.65	0.968 (0.547-1.713)	.91
Arithmetic skill job	0.974 (0.370-2.563)	.96	2.058 (0.673-6.288)	.21
Semiprofessional	1.718 (0.938-3.145)	.08	1.424 (0.606-3.348)	.42
Professional	1.028 (0.597-1.771)	.92	2.275 (1.103-4.695)	.03
Other	4.029 (1.614-10.062)	.003	2.168 (0.484-9.716)	.31
Socioeconomic status				
Upper class[d]	N/A	N/A	N/A	N/A
Upper-middle class	0.518 (0.388-0.691)	<.001	0.384 (0.252-0.584)	<.001
Middle class	0.388 (0.295-0.509)	<.001	0.352 (0.242-0.512)	<.001
Lower-middle class	0.431 (0.326-0.568)	<.001	0.343 (0.234-0.504)	<.001
Lower class	0.477 (0.352-0.647)	<.001	0.266 (0.167-0.424)	<.001
Educational level				
Illiterate[d]	N/A	N/A	N/A	N/A
Primary school certificate	0.815 (0.582-1.142)	.24	0.666 (0.425-1.043)	.08
Middle school certificate	1.083 (0.828-1.417)	.56	0.722 (0.504-1.036)	.08
High school certificate	1.076 (0.843-1.372)	.56	0.697 (0.501-0.968)	.03
Intermediate or post–high school diploma	0.823 (0.596-1.137)	.24	0.555 (0.352-0.875)	.01
Graduate or postgraduate	1.161 (0.885-1.521)	.28	0.497 (0.333-0.742)	.001
Profession or honors	1.714 (0.758-3.873)	.20	0.473 (0.145-1.544)	.22
Marital status				

Characteristic	Central obesity (based on WC[a])		General obesity (based on WHO[b] guidelines)	
	OR[c] (95% CI)	*P* value	OR (95% CI)	*P* value
Unmarried[d]	N/A	N/A	N/A	N/A
Married	1.136 (0.887-1.455)	.31	1.136 (0.783-1.648)	.50
Un-remarried	1.081 (0.441-2.654)	.86	0.167 (0.019-1.485)	.11
Widowed	1.085 (0.442-2.664)	.86	0.625 (0.174-2.241)	.47
Religion				
Hindu[d]	N/A	N/A	N/A	N/A
Muslim	2.728 (2.245-3.313)	<.001	1.133 (0.844-1.521)	.41
Christian	1.027 (0.534-1.974)	.94	0.312 (0.094-1.037)	.06
Sikh	0.897 (0.515-1.563)	.70	0.806 (0.313-2.074)	.65
Other	1.687 (0.145-19.598)	.68	0.000 (0.000-)	>.99

[a] WC: waist circumference.

[b] WHO: World Health Organization.

[c] OR: odds ratio.

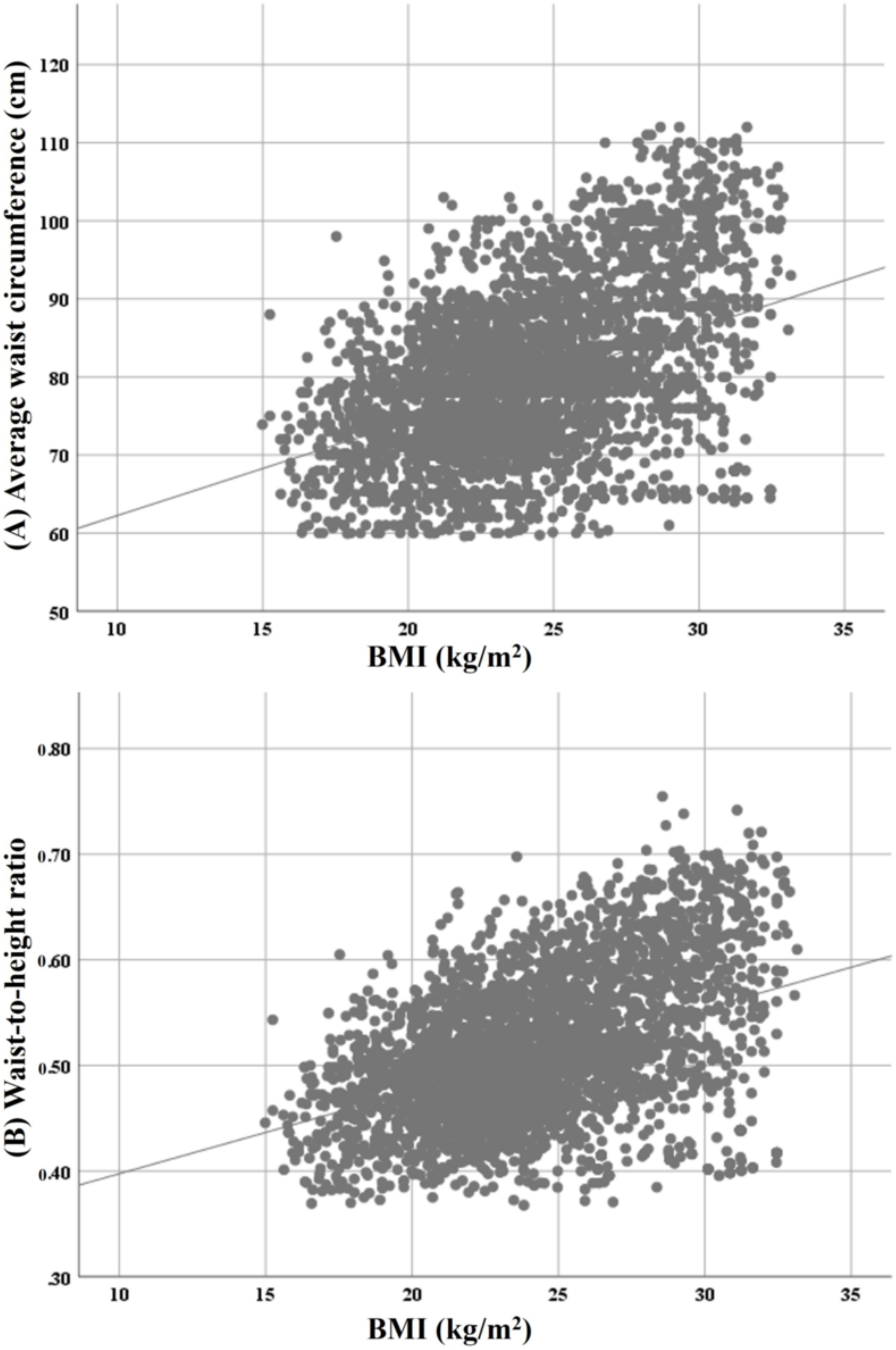
[d] Reference group.

[e] N/A: not applicable.

Correlation Between BMI and Overweight Indicators

Pearson correlation analysis was conducted, and we observed a positive but weak association between these variables. Figure 2 illustrates the correlation between BMI and (1) WC and (2) WHtR of the participants.

Figure 2. Correlation between BMI and (A) waist circumference and (B) waist-to-height ratio.



BMI and WC

BMI and WC showed a statistically significant linear relationship (*r*=0.417; *P*<.001). The direction of the relationship was positive, meaning that these variables tended to increase together (ie, greater WC was associated with greater BMI). However, the magnitude, or strength, of the association was relatively weak (0.3<|*r*|<0.5).

BMI and WHtR

Similarly, BMI and WHtR showed a statistically significant linear relationship (*r*=0.422; *P*<.001). The direction of the relationship was positive, implying that these variables tended to increase together (ie, greater WHtR was associated with greater BMI). However, the magnitude, or strength, of the association was relatively weak (0.3<|*r*|<0.5).

Discussion

Principal Findings

The ongoing ICMR-PCOS Task Force study is the first pan-Indian study conducted among urban and rural women across representative geographical regions. To the best of our knowledge, this is the first report on the pattern and prevalence of overweight, obesity, and hypertension among apparently healthy women with region-wise categorization into North, East, West, South, Northeast, and Central India. This part of the study uses a specific scenario of the risks to women's health, such as overweight and obesity, influenced by multiple determinants of health. Although the prevalence of overweight and obesity affects all segments of the population, there are notable trends among certain sections wherein women are more likely to be obese or gain weight. There are a number of studies available in the literature that use the WHO criteria for BMI cutoffs for defining overweight and obesity prevalence in India. However, these are scientifically challenged as these guidelines are based on studies of White individuals (defined as BMI ≥30 for obesity). Our study provides the comparative estimates of overweight and obesity based on the revised consensus guidelines for India, that is, it defines BMI as underweight (<18.5 kg/m^2), healthy (18.5-22.9 kg/m^2), overweight (23.0-24.9 kg/m^2), and obese (≥25 kg/m^2).

Our results showed that overweight, obesity, and hypertension are also prevalent among apparently healthy women across the country. Health concerns in women start to creep in when entering their second and third decades of life. Our results also showed a rising trend of obesity with an increase in age. While going through age categories, we found that women aged between 36 and 40 years were more obese (448/911, 49.2%). Overweight was observed more among women aged between 30 and 35 years. A significant proportion of women from other age groups was found to be either overweight or obese. The rising incidence of overweight and obesity is a challenge across all age groups [24].

Our findings showed a high proportion of healthy women across rural and urban areas using the revised Indian criteria (647/1969, 32.86% and 533/1908, 27.94%, respectively). These findings are in agreement with the statement under "Patterns by background characteristics of NFHS-5," which states that the proportion of healthy women was higher in rural areas than in urban areas. Our results show higher rates of obesity among urban women than in women from rural areas. These results are in agreement with the data analysis of the NFHS for the epidemiology of obesity in Indian adults and similar to the findings of other studies [9,25-28]. Mishra et al [29] also revealed that the urban population was at a higher risk of obesity than the rural population, in consonance with our findings. We believe that lifestyle and behavioral factors are the main reasons for the differences between urban and rural populations.

Our results showed that both obesity and overweight are significantly prevalent, along with cross-regional variations. The revised consensus guidelines for BMI in Indians showed an increased proportion of obese women in South India (531/1072, 49.53%), followed by East (62/140, 44.3%) and North (841/2214, 37.99%) India, whereas Northeast India showed the lowest rates of obesity (25/164, 15.2%). These findings are partially in agreement with the findings of the NFHS-5, which reported the highest proportion of obese women in Puducherry (south), followed by Chandigarh and Delhi (north). These cross-regional variations highlight the role of geographical factors in these disorders. Levels of urbanization may be the reason for the lower prevalence in the northeastern and eastern regions than in other regions. The northeastern, eastern, and central regions are the 3 least urbanized regions in India [30]. A systematic review reported that variations in the prevalence of obesity in India are due to diverse geographical conditions, lifestyles, and dietary patterns. As expected, we found that only a small proportion of women belonging to the upper class were underweight, and a good proportion (125/308, 40.6%) of women from the same class presented obesity. These findings are similar to those of a cross-sectional study that reported high rates of overweight and obesity (71.8%) among the socioeconomically well-off community of Marwaris, West Bengal, India. This can be attributed to their sedentary lifestyle and high-calorie food intake compared with women of lower SES. Obesity is a triggering factor for various diseases; we used Homeostatic Model Assessment of IR as an indicator of IR and found that 49.77% (661/1328) of participants had IR, with Homeostatic Model Assessment of IR of >2.

Our results showed that women with the highest level of education (with a profession or honors) were not underweight and showed the lowest rates of overweight (6/34, 18%) and prehypertension (14/1946, 0.72%) incidence compared with women who were illiterate or had other levels of education. These trends resemble the findings in other regions of the world, wherein a better education appears to be associated with a lower likelihood of obesity, especially among women. Data analyses of individual-level national health surveys in countries such as Australia, Canada, England, and Korea also reported a linear relationship between the number of years spent in full-time education and the probability of obesity, with most educated individuals displaying lower rates of the condition [31]. Several other studies have reported a strong education gradient in BMI or obesity, with better educated women being less likely to be overweight or obese [32-34]. Our results suggest that general education campaigns and the design and implementation of effective policies can play a significant role in reducing obesity to curb the epidemic.

Hypertension is of particular importance to public health because of its association with a number of diseases contributing to premature deaths. The incidence of hypertension has increased drastically in many LMICs, including India. Our analysis showed patterns of higher incidence of prehypertension among urban women compared with rural women, and a rising trend was observed with an increase in age. These findings refute the age-old concept that hypertension is not a major problem in rural India. The increasing trend of prehypertension with age is in agreement with the NFHS-5, which reported an increase in all categories of hypertension, including prehypertension, with age for both women and men. We also found patterns of greater prevalence of prehypertension among women belonging to the middle class (738/1946, 37.92%) and the lower middle class (461/1946, 23.69%) than among women from the upper class, with an incidence rate of 6.89% (133/1931). This may be linked to the tendencies toward sedentary living and eating unhealthy foods, among other things. These findings are in contradiction to the findings of studies from Nigeria and Kenya that reported a greater prevalence of hypertension among wealthier groups [35,36]. Prehypertension was found to be more prevalent in women who had a lower educational level (high school) and women who were illiterate (379/1946, 19.48%) compared with women with higher levels of education (professional or honors). These findings support the belief that a higher educational level increases awareness and information regarding hypertension and, subsequently, encourages adaptation to healthy lifestyles, and that could be the possible reason behind it. Reddy et al [37] reported based on a cross-sectional study that hypertension was significantly more prevalent in the group with a lower educational level compared with that with a higher educational level. Our results showed that age, level of education, occupation, geographical background, and SES were significant predictors of hypertension. The possible explanations for the lower incidence rates of overweight, obesity, and hypertension in women with a higher education and from the upper socioeconomic class in this study may be attributed to good awareness of preventive measures, better adherence to medical advice, and lifestyle modifications. We believe that there is an urgent need to implement strategies to increase awareness of hypertension and promote healthier lifestyles.

Region-wise categorization into North, East, West, South, Central, and Northeast India allowed us to capture and present the regional estimates of prehypertension and BMI categories among healthy subsets of women, which showed the highest rates of prehypertension among women from southern parts of India (709/1072, 66.14%) followed by East Indian women (70/140, 50%) and the lowest rates among Northeast Indian women. Overall, the identification of women who might face greater challenges with their disease management suggests a role for relevant health authorities in executing appropriately targeted public health measures and policies to ensure good hypertension management in this population.

Strengths

The major strength of our study is its large sample size with apparently healthy reproductive-age women recruited from the 6 regions of India. Hence, it provides region-specific incidences of overweight, obesity, and prehypertension. The findings of our study may be used to plan interventions or policies for the prevention and control of obesity and hypertension among different populations. The results of this study should be interpreted in the context of the potential limitations. Furthermore, we adopted a common instruction manual and standard equipment of similar specifications at all study centers to ensure the accuracy and reliable quality of all the measurements throughout the entire study period so as to generate quality data.

Limitations

There are some limitations to our study. First, to overcome population heterogeneity, we attempted to divide the study areas geographically into North, East, West, South, Northeast, and Central India and arrive at region-specific rates. However, the cultural and dietary practices vary drastically state-wise, and this may have a bearing on the results. Second, we did not report on the awareness and control of obesity and BP and the association of sociodemographic factors such as SES and nutrition with height. We used the BG Prasad social classification, which is applicable to rural and urban areas, but it only considers income and excludes the other SES criteria. In addition, this study provides estimates from 6 regions that include only a few states and union territories and did not include several major Indian states such as Rajasthan and Madhya Pradesh. Above all, the findings are limited to only feminine (female individuals) aged between 18 and 40 years.

Conclusions

To conclude, overweight, obesity, and hypertension epidemics are highly prevalent among healthy subsets of reproductive-age Indian women across rural and urban strata with cross-regional variations. Age, level of education, geographical background, and SES were found to be predictors of BMI and hypertension categories. Genetic and environmental factors such as diet and activity level are likely contributors to these aberrations; however, role of the type of family and marital status need to be explored further. Our results highlight the importance of both nonmodifiable (such as age) and modifiable (such as educational level) factors in determining patterns of overweight, obesity, and hypertension. Education plays an important role in tackling overweight and obesity. This study may serve as an indicator providing early clues and recommending further research using more rigorous scientific methods. In light of our findings, there is a need for targeted attention to both prevent future epidemics in a population that has not developed obesity or hypertension and help maintain healthy weight and BP in individuals with hypertension. Public health awareness regarding overweight, obesity, and BP needs to be tailored differently for women considering the regional, rural, and urban dimensions. We advocate for multidisciplinary efforts to address these issues, and policy makers need to consider what level of evidence should be deemed sufficient to prompt action and tackle these issues.

Acknowledgments

The authors are grateful to the Indian Council of Medical Research, New Delhi, India, for financial assistance in carrying out the study. They acknowledge the expert committee for their valuable suggestions and appreciate the assistance of the clinical and nonclinical staff of all the participating sites in collecting the data. The authors also thank the Multidisciplinary Research Unit, Sher-i-Kashmir Institute of Medical Sciences, funded by the Department of Health Research, Government of India, for providing the necessary support during the study. All authors declare that they had insufficient funding to support the open-access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this paper. The Indian Council of Medical Research–polycystic ovary syndrome study group members were Mohd Ashraf Ganie, Bharati Kulkarni, Amlin Shukla, Taruna Arora, Subhankar Chowdhury, Vanita Suri, Beena Joshi, Prasanta Kumar Bhattacharya, Sarita Agrawal, Neena Malhotra, Rakesh Sahay, Puthiyaveettil Khadar Jabbar, Roya Rozati, Imtiyaz Ahmad Wani, Gaivee Vinam Meshram, Shouvik Choudhury, V Sreenivas, Mukesh Srivastava, Abilash Nair, Parvaiz Koul, Zafar Amin Shah, Shariq Masoodi, Tabassum Parvez, and Ghulam Nabi Yatoo, Naseer Ahmed, Tariq Gojwari, Muzafar Maqsood Wani, Seema Qayoom, Wahid Khan, Rama Walia, Dipankar De, Aashima Arora, Tulika Singh, Neena Khanna, Nitish Naik, Shyam Prakash, Nandita Gupta, Rajesh Sagar, Jai Bhagwan Sharma, Devasenathipathy Kandasamy, Narendra Kuber Bodhey, Eli Mohapatra, Sabah Siddiqui, Gourisankar Kamilya, Archana Singh, Pradip Mukhopadhyay, Mousumi Mukhopadhyay, Nehar Ranjan Sarkar, Alice Abraham Ruram, Jessy Abraham, Saroj Kumar Pati, Dibakar Sahu, Lokesh Kumar Singh, Donboklang Lynser, Manika Agrawal, Ashok Anand, Shahina Begum, Anushree Patil, Shrabani Mukherjee, Hemangini Thakkar, Aruna Ramaiah, Lakshman Rao, Sudha Bindu, Malathi Ponnuru, Nadeem Ahmad, Sujatha Rani, Aleem Ahmed Khan, Sultan Rizwan Ahmed, Prabhakar Rao, Nirmala C, Anuja Elizabeth George, Jayasree Leelamma, Jayakumari, Gaurav thakur, Saba Noor, Shaikh Sharmeen, Aafia Rashid, Wasia Showkat, Rohina Bashir, Pieu Adhikary, Subhasish Pramanik, Rahul Harish, Mudasir Makhdoomi, Mudasir Fayaz, Nafeez Rehman, Neha Ravi, Ajay Kumar, Siffali Chandrakar, Nithlesh Kumar, Sudipta Banerjee, Gazala Hasan, Nitin Deshpande, Himali Vadhan, Rita Parab, Humaira Minhaj, Balaji Bhasker, Shaik Iqubal Ahmed, Ram Babu, and Arya Suresh.

Data Availability

The data used and analyzed during this study are available from the chief coordinator of the study and can be obtained upon reasonable request.

Conflicts of Interest

None declared.

References

1. Reddy KS, Shah B, Varghese C, Ramadoss A. Responding to the threat of chronic diseases in India. Lancet 2005 Nov;366(9498):1744-1749. [doi: 10.1016/s0140-6736(05)67343-6]
2. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001 May 03;344(18):1343-1350. [doi: 10.1056/nejm200105033441801]
3. Nearly one-third of the world's population is obese or overweight, new data show. The Institute for Health Metrics and Evaluation. 2014 May 28. URL: https://www.healthdata.org/news-events/newsroom/news-releases/nearly-one-third-worlds-population-obese-or-overweight-new-data [accessed 2022-07-05]
4. WHO Consultation on Obesity (1997: Geneva, Switzerland) , World Health Organization, Programme of Nutrition, Family and Reproductive Health. Obesity : preventing and managing the global epidemic : report of a WHO Consultation on Obesity, Geneva, 3-5 June 1997. World Health Organization. 1998. URL: https://apps.who.int/iris/handle/10665/63854 [accessed 2022-07-05]
5. Meldrum DR, Morris MA, Gambone JC. Obesity pandemic: causes, consequences, and solutions-but do we have the will? Fertil Steril 2017 Apr;107(4):833-839 [FREE Full text] [doi: 10.1016/j.fertnstert.2017.02.104] [Medline: 28292617]
6. Ahirwar R, Mondal PR. Prevalence of obesity in India: a systematic review. Diabetes Metab Syndr 2019 Jan;13(1):318-321. [doi: 10.1016/j.dsx.2018.08.032] [Medline: 30641719]
7. Ghosh S, Kumar M. Prevalence and associated risk factors of hypertension among persons aged 15-49 in India: a cross-sectional study. BMJ Open 2019 Dec 16;9(12):e029714 [FREE Full text] [doi: 10.1136/bmjopen-2019-029714] [Medline: 31848161]
8. Obesity. World Health Organization. URL: https://www.who.int/health-topics/obesity#tab=tab_1 [accessed 2022-07-05]
9. National Family Health Survey (NFHS-5) INDIA Volume I Report. International Institute for Population Sciences. URL: https://rchiips.org/nfhs/ [accessed 2022-07-18]
10. Gibson RS. Principles of Nutritional Assessment. Oxford, UK: Oxford University Press; 2005.
11. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan;363(9403):157-163. [doi: 10.1016/s0140-6736(03)15268-3]
12. Kuciene R, Dulskiene V. Associations between body mass index, waist circumference, waist-to-height ratio, and high blood pressure among adolescents: a cross-sectional study. Sci Rep 2019 Jul 01;9(1):9493 [FREE Full text] [doi: 10.1038/s41598-019-45956-9] [Medline: 31263167]
13. Seidell JC, Kahn HS, Williamson DF, Lissner L, Valdez R. Report from a Centers for Disease Control and Prevention Workshop on use of adult anthropometry for public health and primary health care. Am J Clin Nutr 2001 Jan;73(1):123-126. [doi: 10.1093/ajcn/73.1.123] [Medline: 11124761]

14. Gateva A, Kamenov Z. Anthropometric indices of visceral obesity and cardiovascular risk factors in patients with polycystic ovarian syndrome. In: Proceedings of the 15th International & 14th European Congress of Endocrinology. 2012 Presented at: 15th International & 14th European Congress of Endocrinology; May 5-9, 2012; Florence, Italy.
15. Goh LG, Dhaliwal SS, Welborn TA, Lee AH, Della PR. Anthropometric measurements of general and central obesity and the prediction of cardiovascular disease risk in women: a cross-sectional study. BMJ Open 2014 Feb 06;4(2):e004138 [FREE Full text] [doi: 10.1136/bmjopen-2013-004138] [Medline: 24503301]
16. Lee CM, Huxley RR, Wildman RP, Woodward M. Indices of abdominal obesity are better discriminators of cardiovascular risk factors than BMI: a meta-analysis. J Clin Epidemiol 2008 Jul;61(7):646-653. [doi: 10.1016/j.jclinepi.2007.08.012] [Medline: 18359190]
17. Duncan MJ, Mota J, Vale S, Santos MP, Ribeiro JC. Associations between body mass index, waist circumference and body shape index with resting blood pressure in Portuguese adolescents. Ann Hum Biol 2013 Mar 18;40(2):163-167. [doi: 10.3109/03014460.2012.752861] [Medline: 23327095]
18. Nasarudin SH, Ahmad N. Correlation between prehypertension and obesity indices among young adults. Malays J Public Health Med 2016;16(3):235-240 [FREE Full text]
19. Prabhakaran D, Jeemon P, Ghosh S, Shivashankar R, Ajay VS, Kondal D, et al. Prevalence and incidence of hypertension: results from a representative cohort of over 16,000 adults in three cities of South Asia. Indian Heart J 2017 Jul;69(4):434-441 [FREE Full text] [doi: 10.1016/j.ihj.2017.05.021] [Medline: 28822507]
20. Deurenberg P. Universal cut-off BMI points for obesity are not appropriate. Br J Nutr 2001 Feb 09;85(2):135-136. [doi: 10.1079/bjn2000273] [Medline: 11280336]
21. Ganie MA, Chowdhury S, Suri V, Joshi B, Bhattacharya PK, Agrawal S, et al. Evaluation of the prevalence, regional phenotypic variation, comorbidities, risk factors, and variations in response to different therapeutic modalities among Indian women: proposal for the Indian council of medical research-polycystic ovary syndrome (ICMR-PCOS) study. JMIR Res Protoc 2021 Aug 27;10(8):e23437 [FREE Full text] [doi: 10.2196/23437] [Medline: 34448720]
22. Obesity : preventing and managing the global epidemic : report of a WHO consultation. World Health Organization. 1999. URL: https://apps.who.int/iris/handle/10665/42330 [accessed 2022-08-02]
23. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). JAMA 2014 Feb 05;311(5):507-520. [doi: 10.1001/jama.2013.284427] [Medline: 24352797]
24. Sathish T, Kapoor N. Normal weight obesity and COVID-19 severity: a poorly recognized link. Diabetes Res Clin Pract 2020 Nov;169:108521 [FREE Full text] [doi: 10.1016/j.diabres.2020.108521] [Medline: 33096190]
25. International Institute for Population Sciences (IIPS) and ICF. National family health survey (NFHS-4) 2015-16. International Institute for Population Sciences. 2017 Dec. URL: https://dhsprogram.com/pubs/pdf/fr339/fr339.pdf [accessed 2022-08-04]
26. Verma M, Das M, Sharma P, Kapoor N, Kalra S. Epidemiology of overweight and obesity in Indian adults - a secondary data analysis of the National Family Health Surveys. Diabetes Metab Syndr 2021 Jul;15(4):102166. [doi: 10.1016/j.dsx.2021.06.003] [Medline: 34186375]
27. Babu Rao B, Junapudi SS. A comparative study of prevalence of overweight and obesity among urban, and rural population of South India. Int J Community Med Public Health 2019 Feb 22;6(3):1091-1095. [doi: 10.18203/2394-6040.ijcmph20190591]
28. Mehata S, Shrestha N, Ghimire S, Atkins E, Karki DK, Mishra SR. Association of altitude and urbanization with hypertension and obesity: analysis of the Nepal Demographic and Health Survey 2016. Int Health 2021 Feb 24;13(2):151-160 [FREE Full text] [doi: 10.1093/inthealth/ihaa034] [Medline: 32623453]
29. Mishra D, Naorem K, Saraswathy KN. Angiotensin-converting enzyme gene insertion/deletion polymorphism and cardiometabolic risk factors: a study among bhil tribal population from two environmental settings. Biochem Genet 2018 Aug 13;56(4):295-314. [doi: 10.1007/s10528-018-9845-x] [Medline: 29435690]
30. Chatterjee M. Regional variation in urbanisation in India and the emergence of new towns. Centre for Urban Economic Studies, Department of Economics, University of Calcutta. 2014. URL: https://www-sre.wu.ac.at/ersa/ersaconfs/ersa14/e140826aFinal01459.pdf [accessed 2022-08-21]
31. Devaux M, Sassi F, Church J, Cecchini M, Borgonovi F. Exploring the relationship between education and obesity. OECD J Econ Stud 2011;2011(1):121-150 [FREE Full text] [doi: 10.1787/eco_studies-2011-5kg5825v1k23]
32. Lleras-Muney A. The relationship between education and adult mortality in the United States. Rev Econ Stud 2005 Jan;72(1):189-221. [doi: 10.1111/0034-6527.00329]
33. Sánchez-Vaznaugh EV, Kawachi I, Subramanian SV, Sánchez BN, Acevedo-Garcia D. Do socioeconomic gradients in body mass index vary by race/ethnicity, gender, and birthplace? Am J Epidemiol 2009 May 01;169(9):1102-1112. [doi: 10.1093/aje/kwp027] [Medline: 19299405]
34. Kenkel D, Lillard D, Mathios A. The roles of high school completion and GED receipt in smoking and obesity. J Labor Econ 2006 Jul;24(3):635-660. [doi: 10.1086/504277]
35. Addo J, Smeeth L, Leon DA. Hypertension in sub-Saharan Africa. Hypertension 2007 Dec;50(6):1012-1018. [doi: 10.1161/hypertensionaha.107.093336]

36. Okoduwa SI, Umar IA, Ibrahim S, Bello F, Ndidi US. Socio-economic status of patients with type 2 diabetes and hypertension attending the Ahmadu Bello University Teaching Hospital, Zaria, North-West Nigeria. Glob J Health Sci 2014 Oct 09;7(1):280-287 [FREE Full text] [doi: 10.5539/gjhs.v7n1p280] [Medline: 25560354]
37. Reddy KS, Prabhakaran D, Jeemon P, Thankappan KR, Joshi P, Chaturvedi V, et al. Educational status and cardiovascular risk profile in Indians. Proc Natl Acad Sci U S A 2007 Oct 09;104(41):16263-16268 [FREE Full text] [doi: 10.1073/pnas.0700933104] [Medline: 17923677]

Abbreviations

BP: blood pressure
DBP: diastolic blood pressure
ICMR: Indian Council of Medical Research
IR: insulin resistance
JNC 8: Eighth Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure
LMICs: low- and middle-income countries
NCD: noncommunicable disease
NFHS: National Family Health Survey
OR: odds ratio
PCOS: polycystic ovary syndrome
SBP: systolic blood pressure
SES: socioeconomic status
WC: waist circumference
WHO: World Health Organization
WHR: waist-to-hip ratio
WHtR: waist-to-height ratio

Edited by A Mavragani, T Sanchez; submitted 22.11.22; peer-reviewed by L Vizza, S Jameela; comments to author 15.03.23; revised version received 05.04.23; accepted 16.05.23; published 06.09.23.

Please cite as:
Ganie MA, Chowdhury S, Suri V, Joshi B, Bhattacharya PK, Agrawal S, Malhotra N, Sahay R, Jabbar PK, Rozati R, Wani IA, Shukla A, Arora T, Rashid H
Prevalence, Regional Variations, and Predictors of Overweight, Obesity, and Hypertension Among Healthy Reproductive-Age Indian Women: Nationwide Cross-Sectional Polycystic Ovary Syndrome Task Force Study
JMIR Public Health Surveill 2023;9:e43199
URL: https://publichealth.jmir.org/2023/1/e43199
doi:10.2196/43199
PMID:37672315

Original Paper

The Impact of Physical Activity Intensity on the Dynamic Progression of Cardiometabolic Multimorbidity: Prospective Cohort Study Using UK Biobank Data

Bao-Peng Liu[1], MPH, PhD; Jia-Hui Zhu[1], BMed; Li-Peng Wan[1], MM; Zhen-Yu Zhao[1], BMed; Xinting Wang[1], PhD; Cun-Xian Jia[1], PhD

Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

Corresponding Author:
Cun-Xian Jia, PhD
Department of Epidemiology, School of Public Health
Cheeloo College of Medicine
Shandong University
Wenhua Xi Road, No 44
Jinan, 250012
China
Phone: 86 88382141 ext 8809
Email: jiacunxian@sdu.edu.cn

Abstract

Background: Although many studies have reported on the associations between the amount of physical activity (PA) and the transitions of cardiometabolic multimorbidity (CMM), the evidence for PA intensity has not been fully evaluated.

Objective: This study aimed to explore the impact of PA intensity on the dynamic progression of CMM.

Methods: The prospective cohort of this study using data from the UK Biobank included 359,773 participants aged 37-73 years who were recruited from 22 centers between 2006 and 2010. The diagnoses of CMM, which included the copresence of type 2 diabetes (T2D), ischemic heart disease, and stroke, were obtained from *first occurrence fields* provided by the UK Biobank, which included data from primary care, hospital inpatient record, self-reported medical condition, and death registers. The PA intensity was assessed by the proportion of vigorous PA (VPA) to moderate to vigorous PA (MVPA). Multistate models were used to evaluate the effect of PA intensity on the dynamic progression of CMM. The first model (model A) included 5 transitions, namely free of cardiometabolic disease (CMD) to first occurrence of CMD (FCMD), free of CMD to death, FCMD to CMM, FCMD to mortality, and CMM to mortality. The other model (model B) used specific CMD, namely T2D, ischemic heart disease, and stroke, instead of FCMD and included 11 transitions in this study.

Results: The mean age of the included participants (N=359,773) was 55.82 (SD 8.12) years at baseline, and 54.55% (196,271/359,773) of the participants were female. Compared with the participants with no VPA, participants with intensity levels of >0.75 to <1 for VPA to MVPA had a 13% and 27% lower risk of transition from free of CMD to FCMD (hazard ratio [HR] 0.87, 95% CI 0.83-0.91) and mortality (HR 0.73, 95% CI 0.66-0.79) in model A, respectively. The HR for the participants with no moderate PA was 0.82 (95% CI 0.73-0.92) compared with no VPA. There was a substantially protective effect of higher PA intensity on the transitions from free of CMD to T2D and from T2D to mortality, which reveals the importance of PA intensity for the transitions of T2D. More PA and greater intensity had a synergistic effect on decreasing the risk of the transitions from free of CMD to FCMD and mortality. Male participants, younger adults, adults with a higher BMI, current or previous smokers, and excessive alcohol drinkers could obtain more benefits from higher PA intensity for the lower risk of at least 1 transition from free of CMD, then to CMM, and finally to mortality.

Conclusions: This study suggests that higher PA intensity is an effective measure for preventing CMM and mortality in the early period of CMM development. Relevant interventions related to higher PA intensity should be conducted.

(JMIR Public Health Surveill 2023;9:e46991) doi:10.2196/46991

KEYWORDS

physical activity intensity; PA; dynamic progression; cardiometabolic multimorbidity; cohort study; CMM

Introduction

Background

Multimorbidity, defined as the coexistence of ≥2 chronic conditions from noncommunicable disease, mental disease, or infectious disease of long duration, is an unignorable public health issue that brings great physical and mental burden to families [1]. Cardiometabolic multimorbidity (CMM), namely the copresence of type 2 diabetes (T2D), ischemic heart disease (IHD), and stroke, is associated with significantly multiplicative mortality risk and lower life expectancy [2,3]. The prevalence of CMM for adults in the United States is 14.4% and is higher among the male population and older adults [4]. Previous studies characterized a troubling rise in the prevalence of CMM from 1999 to 2018 [4]. Early prevention is needed in consideration of the many adverse health outcomes related to CMM, such as cognitive decline, dementia, depression, and even a worse COVID-19 prognosis [5-8]. Previous studies have mentioned that many variables such as handgrip strength, sleep pattern, beverage consumption, shift work, and ambient air pollution are independent factors for predicting CMM [9-13]. Exploring more factors affecting the early stage of CMM, which means free of cardiometabolic disease (CMD) or first occurrence of CMD (FCMD), is necessary for preventing adverse and severe outcomes in the future.

A high level of physical activity (PA), especially for a longer duration of moderate to vigorous PA (MVPA), has been shown to be the common protective factor for CMD, including T2D [14-16], IHD [17-19], stroke [20-22], CMM [23-26], and death [27-30]. Previous studies also reported that a higher intensity of PA, namely a higher proportion of vigorous PA (VPA), could bring additional benefits for reducing the risk of mortality [31-34]. However, there is limited evidence for the associations of PA intensity with CMD and CMM. In other words, it is unclear whether a higher intensity of PA has a positive effect on decreasing the risk of mortality among the participants with CMD or CMM. Furthermore, previous studies also reported that the amount of PA, namely the duration or metabolic equivalent task (MET) for PA, could affect all the stages from healthy status to CMD, then to CMM, and finally to mortality [23,24]. To the best of our knowledge, no previous studies have evaluated the association of PA intensity with all the transitions simultaneously. Considering the studies focusing on a single transition might overlook some competing risks of mortality, future studies should investigate how PA intensity affects the dynamic progression from the free of CMD status to CMM and from other status to mortality.

Objectives

This study aimed to explore the associations of PA intensity with the dynamic progression of CMM, namely from free of CMD to FCMD, then to CMM, and finally to mortality. In addition, the effect of PA intensity on the transitions from specific CMD, namely T2D, IHD, and stroke, to CMM and mortality was further assessed. Furthermore, the effect of modification by sociodemographic factors and the amount of PA was also assessed. The findings of this study would provide some evidence of the benefits and help relevant departments to develop some interventions related to the PA intensity for CMM and mortality.

Methods

Study Design and Participants

The UK Biobank (UKB) is an ongoing population-based and prospective cohort, which includes >500,000 participants aged 37 to 73 years. The details of recruitment for the UKB can be found elsewhere [35]. The main aims of the UKB are to explore genetic and nongenetic determinants of the diseases among middle-aged adults and older adults. Briefly, UKB participants were recruited from 22 centers in the United Kingdom (England, Wales, and Scotland) between 2006 and 2010 at baseline. After obtaining electronic consent for the use of deidentified data, every participant finished a self-completed touchscreen questionnaire, including socioeconomic factors and lifestyles. In addition, computer-assisted interviews, physical measurements, and sample collection were also conducted. The records from national health-related hospitals, primary care, and death registers were also matched with the included participants, which could help us comprehensively understand their health status.

In this study, we excluded participants whose records were withdrawn or could not be followed up from the cohort. In addition, the participants with missing PA data at baseline were also excluded. Finally, we excluded the participants with CMD at baseline to avoid inverse causality. Finally, 359,773 participants were included in this study (Figure S1 in Multimedia Appendix 1).

Measures

Exposures

In this study, PA was assessed by using the short-form International Physical Activity Questionnaire, which has good validity and reliability [36]. The MET of moderate PA (MPA) and VPA calculated by the UKB was used in this study (data field 22038 and 22039). MVPA, which is the sum of MET associated with MPA and VPA, was used to assess the amount of PA according to the recommendations of the World Health Organization PA guidelines [37]. First, the amount of PA was categorized into <600, 600 to 1200, and >1200 MET minutes per week associated with MVPA. Second, PA intensity, which is the ratio of VPA to MVPA, was measured by using the MET of VPA to MVPA as follows: MET (VPA) / MET (MVPA). The PA intensity ranged from 0 to 1, and a higher ratio indicated higher PA intensity. Different categories (0, >0 to 0.25, >0.25 to 0.5, >0.5 to 0.75, >0.75 to <1, and 1) of PA intensity were used to assess the associations with all the transitions from the free of CMD status to FCMD, then to CMM, and finally to mortality. In addition, the ordinal scale, which coded the PA intensity from 0 to 5, was also used to measure the associations.

Outcomes

The incidence, progression, and prognosis of CMM, namely the occurrence of at least 2 diseases of T2D, IHD, and stroke, were the main outcomes in this study. The diagnoses and time of occurrence were obtained from *first occurrence fields* provided by UKB (data category: 2409) and were coded using the International Classification of Disease, 10th version, which included data from primary care, hospital inpatient record, self-reported medical condition, and death registers. T2D was coded by E11 and E14, as recommended in previous studies related to CMM diagnoses [9,24]. IHD was coded by I20-I25, and stroke was coded by I60-I69.

The records related to all-cause mortality were identified by linking to death registries of the National Health Service Information Centre for participants from England and Wales and the National Health Service Central Register for participants from Scotland [35]. The participants entered the cohort from the date of being recruited at the centers and exited at the date of death, occurrence of outcomes at respective progression, or censorship.

Covariates

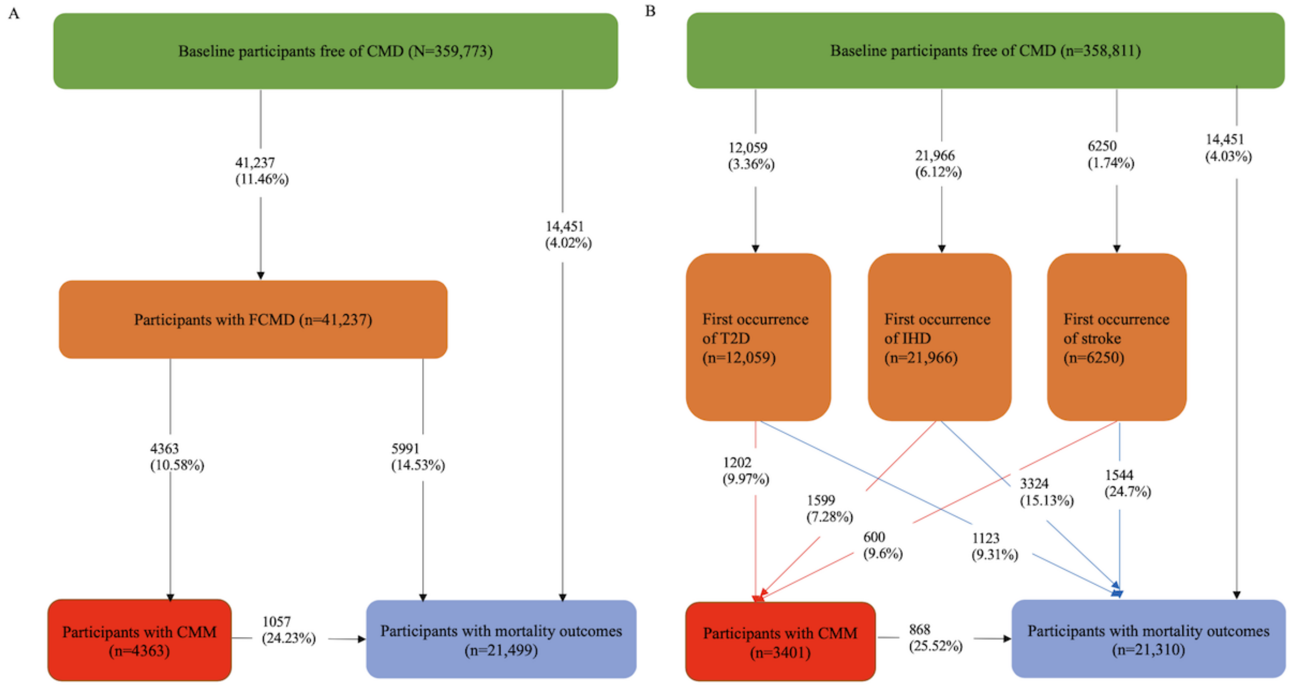
The covariates in this study, which were chosen according to previous studies on CMM [9,23,24], included age, sex, ethnicity, socioeconomic status (SES), BMI, household income before tax per year, study center region, education group, smoking status, and alcohol drinking status. According to previous studies [38,39], an a priori–defined directed acyclic graph [40] was drawn to show the relationship between the covariates and outcomes (Figure S2 in Multimedia Appendix 1). Ethnicity was classified as either White or Others. The SES was measured using the Townsend area deprivation index [41]. Higher scores indicated greater socioeconomic deprivation and lower SES, and quartiles of the score for SES were included in the analyses. The data field and definition of other included covariates are shown in Table S1 in Multimedia Appendix 1. Missing values for the covariates were regarded as a classification when performing the analyses.

Statistical Analysis

We used R (version 4.1.2; R Foundation for Statistical Computing) software to perform statistical analyses. Initially, traditional Cox proportional hazards models were used to explore the association of the amount of PA and intensity with FCMD, CMM, and all-cause mortality. No obvious violations to the proportional hazards assumption were noted for interested exposures. All the models were adjusted for age, sex, ethnicity, SES, BMI, household income before tax per year, study center region, education group, smoking status, and alcohol drinking status.

Furthermore, multistate models [42,43] were used to assess the effect of PA amount and intensity on the transitions from free of CMD to FCMD, then to CMM, and finally to mortality. The multistate model, which was an extension of the traditional Cox proportional hazards model and competing risk model, was widely used to explore the effect of interested risk factors on multiple transitions related to multistate outcomes. In this study, we adopted 2 series of multistate models, as recommended by previous studies [9,23,24]. The first model (model A) included 5 transitions, namely free of CMD to FCMD, free of CMD to death, FCMD to CMM, FCMD to mortality, and CMM to mortality (Figure 1). The other model (model B) used specific CMD, namely T2D, IHD, and stroke, instead of FCMD and included 11 transitions (Figure 1). When the participants had the same date for different transitions, we adopted a theoretically time of pushing entering date forward by 0.5 day, which was based on previous studies related to a similar topic [9,24]. When performing model B, we excluded 962 participants with the same date of diagnosis for T2D, IHD, or stroke to explore disease-specific associations, as we were unsure which disease occurred first.

Figure 1. Transitions from free of cardiometabolic disease (CMD) to mortality outcomes: (A) 5-transition model and (B) 11-transition model. CMM: cardiometabolic multimorbidity; FCMD: first cardiometabolic disease; IHD: ischemic heart disease; T2D: type 2 diabetes.



A series of subgroup analyses related to age (≤60 vs >60 y), sex (female vs male), BMI (underweight or normal weight vs overweight), smoking status (never vs previous or current), and alcohol drinking status (daily or almost daily vs ≤1-3 times/wk) were performed to identify the susceptible populations associated with PA intensity. *Z* test was used to compare the estimates of different subgroups, as recommended by Altman and Bland [44].

We performed several sensitivity analyses to verify the robustness of the estimates of model A. First, we used different time intervals (0.5 y, 1 y, 2 y, 3 y, and 5 y) to theoretically change the entering time to test the stability of the estimates. Second, we excluded the participants who entered different transitions on the same day. Third, we omitted the participants with FCMD for the first 2 years of follow-up to account for reverse causality. Fourth, considering the interactive effect of cancer and CMD, we reanalyzed the associations by omitting the participants with cancer at baseline. In addition, the participants with missing covariate values were excluded from the analyses. Finally, considering the correlation of air pollution with PA and CMM, we separately added particulate matter 2.5, particulate matter 10, nitric oxide, and nitrogen dioxide at baseline in the multistate models. The data of air pollution were estimated using the Land Use Regression model developed as part of the European Study of Cohorts for Air Pollution Effects and can be found in the UKB database (field id: 24003, 24004, 24005, and 24006).

Ethical Considerations

UKB received ethics approval from the North West Multi-centre Research Ethics Committee (reference 21/NW/0157). All participants provided written informed consent before enrollment in the study, which was conducted in accordance with the principles of the Declaration of Helsinki. The data used in this study were anonymized and deidentified for privacy and confidentiality protection. More details about ethics approval can be found on the UKB website.

Results

Participant Characteristics at Baseline and Transitions From Baseline to Mortality

Of the included participants, 54.55% (196,271/359,773) were female and 94.91% (341,474/359,773) were White. The mean age of the participants was 55.82 (SD 8.12) years at baseline. Moreover, 64.12% (230,676/359,773) of the participants were overweight and obese, and 88.68% (319,043/359,773) of the participants were recruited from England; 56.06% (201,671/359,773) of the participants never smoked, and 70.21% (252,589/359,773) of the participants drank daily, almost daily, or more than once per week. Furthermore, 38.9% (139,963/359,773), 19.55% (70,329/359,773), and 41.55% (149,481/359,773) of the included participants had MVPA of <600, 600 to 1200, and >1200 MET minutes per week, respectively. The participants with lower PA intensity were more likely to be female; be of younger age, belong to a different ethnicity, have a lower SES, have a higher BMI, have lower income, have a lower education level, engage in more frequent drinking, and have less PA; and be current smokers. All the differences for different groups of PA intensity were statistically significant. More characteristics for included participants are shown in Table 1.

There were 2 patterns for the transitions from free of CMD to FCMD, then to CMM, and finally to mortality. The 5-transition model and 11-transition model are shown in Figure 1. For pattern A, 11.46% (41,237/359,773) of the participants developed FCMD and 4.02% (14,451/359,773) of the participants among baseline participants free of CMD died eventually. Moreover, 10.58% (4363/41,237) of the participants developed CMM, and 14.53% (5991/41,237) of the participants died after FCMD. Finally, 24.23% (1057/4363) of the participants died after a CMM diagnosis. For pattern B, 3.36% (12,059/358,811), 6.12% (21,966/358,811), and 1.74% (6250/358,811) of the participants, among the baseline participants free of CMD, first developed T2D, IHD, and stroke, respectively. Moreover, 8.44% (3401/40275) of the participants developed CMM, and 14.88% (5991/40275) of the participants died after a diagnosis of T2D, IHD, and stroke. Eventually, 25.52% (868/3401) of the participants died after having a diagnosis of CMM. More details are presented in Figure 1.

Table 1. Descriptive statistics by physical activity (PA) intensity (N=359,773).

Characteristics[a]	Overall	No MVPA[b] (n=47,514)	PA intensity: the proportion of VPA[c] to MVPA					
			0 (n=94,989)	>0 to 0.25 (n=42,934)	>0.25 to 0.5 (n=65,347)	>0.5 to 0.75 (n=75,480)	>0.75 to <1 (n=24,606)	1 (n=8903)
Sex, n (%)								
Female	196,271 (54.55)	26,608 (56)	57,040 (60.05)	24,332 (56.67)	35,951 (55.02)	36,792 (48.74)	11,154 (45.33)	4394 (49.35)
Male	163,502 (45.44)	20,906 (44)	37,949 (39.95)	18,602 (43.33)	29,396 (44.98)	38,688 (51.25)	13,452 (54.67)	4509 (50.64)
Age (years), mean (SD)	55.82 (8.12)	55.60 (7.91)	57.05 (7.92)	57.54 (7.87)	56.00 (8.11)	54.45 (8.15)	53.11 (8.01)	53.47 (8.1)
Ethnicity, n (%)								
White	341,474 (94.91)	44,652 (93.98)	90,312 (95.08)	41,432 (96.5)	62,396 (95.48)	71,274 (94.43)	23,188 (94.24)	8220 (92.33)
Others	17,256 (4.8)	2677 (5.63)	4429 (4.66)	1386 (3.23)	2768 (4.23)	4002 (5.3)	1349 (5.48)	645 (7.24)
Unknown	1043 (0.29)	185 (0.39)	248 (0.26)	116 (0.27)	183 (0.28)	204 (0.27)	69 (0.28)	38 (0.04)
Socioeconomic status, n (%)[d]								
First quartile	89,899 (24.99)	10,834 (22.8)	22,203 (23.37)	11,253 (26.21)	16,946 (25.93)	19,662 (26.05)	6722 (27.31)	2279 (25.6)
Second quartile	89,761 (24.95)	11,309 (23.8)	23,192 (24.41)	11,122 (25.9)	16,653 (25.48)	18,996 (25.17)	6341 (25.77)	2148 (24.13)
Third quartile	89,832 (24.97)	11,689 (24.6)	23,831 (25.09)	10,849 (25.27)	16,404 (25.1)	18,852 (24.98)	6009 (24.42)	2198 (24.69)
Fourth quartile	89,827 (24.97)	13,613 (28.65)	25,640 (26.99)	9666 (22.51)	15,264 (23.36)	17,874 (23.68)	5501 (22.36)	2269 (25.48)
Unknown	454 (0.13)	69 (0.14)	123 (0.13)	44 (0.1)	80 (0.12)	96 (0.13)	33 (0.13)	9 (0.1)
BMI (kg/m^2), n (%)								
Underweight (<18.5)	1939 (0.54)	296 (0.62)	638 (0.67)	244 (0.57)	325 (0.5)	313 (0.41)	95 (0.39)	28 (0.31)
Normal weight (18.5-24.9)	125,609 (34.91)	12,981 (27.32)	30,952 (32.58)	16,229 (37.8)	24,794 (37.94)	28,015 (37.11)	9536 (38.75)	3102 (34.84)
Overweight (25.0-29.9)	154,404 (42.92)	19,466 (40.97)	40,039 (42.15)	18,632 (43.4)	28,244 (43.22)	33,311 (44.13)	10,843 (44.07)	3869 (43.46)
Obese (≥30)	76,272 (21.2)	14,359 (30.22)	22,929 (24.14)	7691 (17.91)	11,774 (18.02)	13,574 (17.98)	4076 (16.56)	1869 (20.99)
Unknown	1549 (0.43)	412 (0.87)	431 (0.45)	138 (0.32)	210 (0.32)	267 (0.35)	56 (0.23)	35 (0.39)
Household income before tax/y (£[e]), n (%)								
<18,000	62,366 (17.33)	9501 (20.15)	19,972 (21.02)	8107 (18.88)	10,569 (16.17)	10,553 (13.98)	2556 (10.39)	1108 (12.44)
18,000-30,999	78,380 (21.78)	9601 (20.21)	21,950 (23.11)	11,395 (26.54)	14,942 (22.86)	14,902 (19.74)	4174 (16.96)	1416 (15.9)
31,000-51,999	86,831 (24.13)	11,069 (23.3)	21,890 (23.04)	10,544 (24.56)	16,194 (24.78)	18,796 (24.9)	6184 (25.13)	2154 (24.19)
52,000-100,000	71,816 (19.96)	9043 (19.03)	15,859 (16.7)	6560 (15.28)	12,909 (19.75)	17,951 (23.78)	7112 (28.9)	2382 (26.75)
>100,000	19,873 (5.52)	2129 (4.48)	3726 (3.92)	1551 (3.61)	3532 (5.4)	5634 (7.46)	2557 (10.39)	744 (8.36)
Unknown	40,507 (11.26)	6171 (12.99)	11,592 (12.2)	4777 (11.13)	7201 (11.02)	7644 (10.13)	2023 (8.22)	1099 (12.34)

Characteristics[a]	Overall	No MVPA[b] (n=47,514)	PA intensity: the proportion of VPA[c] to MVPA					
			0 (n=94,989)	>0 to 0.25 (n=42,934)	>0.25 to 0.5 (n=65,347)	>0.5 to 0.75 (n=75,480)	>0.75 to <1 (n=24,606)	1 (n=8903)
Study center region, n (%)								
England	319,043 (88.68)	41,593 (87.54)	83,956 (88.38)	38,619 (89.95)	58,252 (89.15)	66,963 (88.72)	21,773 (88.49)	7887 (88.59)
Wales	14,829 (4.12)	2309 (4.86)	3965 (4.17)	1639 (3.82)	2496 (3.82)	3057 (4.05)	996 (4.05)	367 (4.12)
Scotland	25,901 (7.2)	3612 (7.6)	7068 (7.44)	2676 (6.23)	4599 (7.04)	5460 (7.23)	1837 (7.46)	649 (7.29)
Education group, n (%)								
College or university degree	129,911 (36.11)	14,438 (30.39)	30,719 (32.34)	14,510 (33.8)	24,849 (38.03)	30,087 (39.86)	11,581 (47.06)	3727 (41.86)
Any school degree (A-level[f], AS-level[g], O-level[h], GCSE[i], and CSE[j])	138,483 (38.49)	19,240 (40.49)	37,065 (39.02)	17,251 (40.18)	24,817 (37.98)	28,106 (37.24)	8707 (35.38)	3297 (37.03)
Vocational qualification (NVQ[k], HND[l], or HNC[m]) or other professional qualifications	40,545 (11.27)	5054 (10.64)	10,947 (11.52)	5524 (12.87)	7747 (11.85)	8154 (10.8)	2329 (9.46)	790 (8.87)
None of the above	48,412 (13.46)	8351 (17.57)	15,565 (16.39)	5414 (12.61)	7548 (11.55)	8644 (11.45)	1878 (7.63)	1012 (11.37)
Unknown	2422 (0.67)	431 (0.91)	693 (0.73)	235 (0.55)	386 (0.59)	489 (0.65)	111 (0.45)	77 (0.86)
Smoking status, n (%)								
Never	201,671 (56.06)	25,457 (53.58)	51,901 (54.64)	23,844 (55.54)	37,116 (56.8)	43,488 (57.61)	14,705 (59.76)	5160 (57.96)
Previous	121,014 (33.64)	15,493 (32.61)	31,669 (33.34)	15,045 (35.04)	22,407 (34.29)	25,374 (33.61)	8061 (32.76)	2965 (33.3)
Current	36,181 (10.06)	6385 (13.44)	11,144 (11.73)	3951 (9.2)	5706 (8.73)	6454 (8.55)	1787 (7.26)	754 (8.47)
Unknown	907 (0.25)	179 (0.38)	275 (0.29)	94 (0.22)	118 (0.18)	164 (0.22)	53 (0.21)	24 (0.27)
Alcohol drinking status, n (%)								
Daily or almost daily	77,268 (21.48)	9664 (20.34)	20,644 (21.73)	10,029 (23.36)	14,441 (22.1)	15,753 (20.87)	5028 (20.43)	1709 (19.19)
Above 1 time/wk	181,021 (50.31)	21,422 (45.08)	44,118 (46.44)	21,470 (50.01)	34,130 (52.23)	41,242 (54.64)	13,925 (56.59)	4714 (52.95)
1-3 times/mo	39,390 (10.95)	5740 (12.08)	10,895 (11.47)	4604 (10.72)	6943 (10.62)	7664 (10.15)	2563 (10.42)	981 (11.02)
Special occasions only	37,073 (10.3)	6042 (12.72)	11,479 (12.08)	4323 (10.07)	5969 (9.13)	6511 (8.63)	1850 (7.52)	899 (10.1)
Never	24,830 (6.9)	4601 (9.68)	7791 (8.2)	2496 (5.81)	3840 (5.88)	4277 (5.67)	1231 (5)	594 (6.67)
Unknown	191 (0.05)	45 (0.09)	62 (0.06)	12 (0.03)	24 (0.04)	33 (0.04)	9 (0.04)	6 (0.07)
PA amount (MET[n] min/wk), n (%)								
<600	139,963 (38.9)	47,514 (100)	58,656 (61.75)	3780 (8.8)	12,729 (19.48)	10,695 (14.17)	2726 (11.08)	3863 (43.39)
600-1200	70,329 (19.55)	0 (0)	17,960 (18.91)	9121 (21.24)	16,019 (24.51)	17,574 (23.25)	7118 (28.93)	2537 (28.5)

Characteristics[a]	Overall	No MVPA[b] (n=47,514)	PA intensity: the proportion of VPA[c] to MVPA					
			0 (n=94,989)	>0 to 0.25 (n=42,934)	>0.25 to 0.5 (n=65,347)	>0.5 to 0.75 (n=75,480)	>0.75 to <1 (n=24,606)	1 (n=8903)
>1200	149,481 (41.55)	0 (0)	18,373 (19.34)	30,033 (69.95)	36,599 (56.01)	47,211 (62.5)	14,762 (60)	2503 (28.11)

[a] Unknown included prefer not to answer, do not know, and missing values in the UK Biobank database.

[b] MVPA: moderate to vigorous PA.

[c] VPA: vigorous PA.

[d] Socioeconomic status was measured by using the Townsend area deprivation index.

[e] A currency exchange rate of £1=US $1.26 is applicable.

[f] A-level: Advanced level.

[g] AS-level: Advanced Subsidiary level.

[h] O-level: Ordinary level.

[i] GCSE: General Certificate of Secondary Education.

[j] CSE: General Certificate of Education.

[k] NVQ: National Vocational Qualification.

[l] HND: Higher National Diploma.

[m] HNC: Higher National Certificate.

[n] MET: metabolic equivalent task.

Traditional Cox Proportional Hazards Model

The estimates of the associations of PA amount and intensity with FCMD, CMM, and mortality are presented in Table S2 in Multimedia Appendix 1. More MVPA with 600-1200 and >1200 MET minutes per week lowered the risk of FCMD, CMM, and mortality than MVPA with <600 MET minutes per week. A decreased risk of FCMD, CMM, and mortality was also found in the participants with more VPA to MVPA (>0 to <1) than in the participants with no VPA. However, the associations of no MPA, namely total VPA with FCMD (*P*=.073), CMM (*P*=.294), and mortality (*P*=.123), were not statistically significant compared with no VPA. Similar results were found when the ordinal scale of PA intensity was used in the models.

Multistate Models

Multistate models for the associations of PA amount (Table S3 in Multimedia Appendix 1) and intensity (Table 2) with all the transitions were performed. First, meeting the recommendations of the World Health Organization, namely no less than 600 MET minutes per week for MVPA were found to be negatively associated with transitions from free of CMD to FCMD and mortality. In addition, >1200 MET minutes per week for MVPA was significantly associated with a decreased risk of transition from FCMD to CMM. When focusing on specific FCMD, more PA was found to be associated with a decreased risk of transitions from free of CMD to T2D, IHD, stroke, and mortality as well as transitions from stroke to CMM and from T2D to mortality.

A greater proportion of VPA to MPA, namely >0.5 for the proportion of VPA to MVPA, was associated with a lower risk of different transitions. Our analyses showed that compared with the participants with no VPA, participants with >0.75 to <1 of VPA to MVPA had a 13% and 27% lower risk of transition from free of CMD to FCMD (hazard ratio [HR] 0.87, 95% CI 0.83-0.91) and mortality (HR 0.73, 95% CI 0.66-0.79), respectively. Furthermore, the HR for the participants with no MPA was 0.82 (95% CI 0.73-0.92) compared with those with no VPA. The participants with >0.75 to <1 VPA to MVPA had the lowest risk of transitions from free of CMD to FCMD and mortality. The decreased risk of transitions from free of CMD to FCMD and mortality was also related to higher PA intensity in consideration of the ordinal scale. However, we did not find significant associations of PA intensity with other transitions in model A.

Model B showed different associations of PA intensity with the transitions related to specific FCMD. Specifically, a greater proportion of VPA to MPA was associated with a lower risk of transitions from free of CMD to T2D and stroke. In addition, a decreased risk of transitions from stroke to CMM and from T2D to mortality was found to be associated with >0.5 to 0.75 of VPA to MVPA. It was noted that lower PA intensity, that is, >0 to 0.5 of VPA to MVPA, was also found to be associated with a lower risk of transition from T2D to mortality. The ordinal scale of higher PA intensity was significantly associated with a lower risk of transition from free of CMD to T2D, stroke, and mortality.

Table 2. Associations of physical activity (PA) intensity with transitions from free of cardiometabolic disease (CMD) to first occurrence of CMD (FCMD), cardiometabolic multimorbidity (CMM), and mortality.

Transitions	PA intensity: the proportion of VPA[a] to MVPA[b], HR[c] (95% CI)						
	0	>0 to 0.25	>0.25 to 0.5	>0.5 to 0.75	>0.75 to <1	1	Ordinal scale
Model A[d]							
From free of CMD to FCMD	Reference	0.90 (0.87-0.93)[e]	0.90 (0.87-0.92)[e]	0.89 (0.87-0.92)[e]	0.87 (0.83-0.91)[e]	0.97 (0.90-1.03)	0.97 (0.97-0.98)[e]
From free of CMD to mortality	Reference	0.90 (0.85-0.95)[e]	0.85 (0.81-0.89)[e]	0.80 (0.76-0.84)[e]	0.73 (0.67-0.79)[e]	0.82 (0.73-0.93)[f]	0.94 (0.92-0.95)[e]
From FCMD to CMM	Reference	0.91 (0.82-1.01)	0.98 (0.89-1.08)	0.91 (0.83-1.00)	1.04 (0.90-1.21)	1.05 (0.85-1.30)	1.00 (0.97-1.02)
From FCMD to mortality	Reference	0.93 (0.85-1.02)	0.92 (0.85-1.00)	0.93 (0.86-1.01)	0.94 (0.83-1.08)	0.98 (0.82-1.18)	0.99 (0.97-1.01)
From CMM to mortality	Reference	1.07 (0.86-1.32)	0.97 (0.80-1.18)	1.04 (0.85-1.26)	0.93 (0.66-1.32)	1.16 (0.76-1.77)	1.01 (0.96-1.06)
Model B [d]							
From free of CMD to T2D[g]	Reference	0.74 (0.69-0.79)[e]	0.77 (0.72-0.81)[e]	0.77 (0.73-0.82)[e]	0.73 (0.66-0.79)[e]	0.99 (0.88-1.12)	0.94 (0.93-0.95)[e]
From free of CMD to IHD[h]	Reference	0.96 (0.91-1.00)	0.95 (0.91-0.99)[i]	0.97 (0.93-1.01)	0.95 (0.89-1.01)	1.02 (0.93-1.11)	1.00 (0.99-1.01)
From free of CMD to stroke	Reference	1.06 (0.97-1.15)	0.97 (0.90-1.05)	0.91 (0.85-0.99)[i]	0.91 (0.81-1.03)	0.75 (0.62-0.92)[f]	0.97 (0.95-0.99)[f]
From free of CMD to mortality	Reference	0.90 (0.85-0.95)[e]	0.85 (0.81-0.89)[e]	0.80 (0.76-0.84)[e]	0.73 (0.67-0.79)[e]	0.82 (0.73-0.93)[f]	0.94 (0.92-0.95)[e]
From T2D to CMM	Reference	0.87 (0.70-1.08)	1.01 (0.84-1.21)	0.98 (0.82-1.18)	0.91 (0.66-1.25)	1.12 (0.76-1.64)	1.00 (0.96-1.05)
From IHD to CMM	Reference	0.96 (0.78-1.18)	0.84 (0.69-1.02)	0.95 (0.80-1.14)	0.74 (0.52-1.05)	1.04 (0.70-1.55)	0.97 (0.93-1.02)
From stroke to CMM	Reference	0.94 (0.79-1.10)	0.89 (0.76-1.04)	0.84 (0.72-0.98)[i]	0.96 (0.75-1.22)	1.09 (0.78-1.53)	0.98 (0.94-1.01)
From T2D to mortality	Reference	0.80 (0.71-0.90)[e]	0.86 (0.77-0.95)[f]	0.87 (0.79-0.97)[f]	1.03 (0.87-1.21)	1.08 (0.86-1.36)	0.99 (0.96-1.01)
From IHD to mortality	Reference	0.82 (0.62-1.08)	0.98 (0.77-1.25)	0.90 (0.70-1.16)	1.30 (0.91-1.86)	0.84 (0.41-1.70)	1.01 (0.95-1.08)
From stroke to mortality	Reference	1.09 (0.93-1.28)	1.05 (0.90-1.22)	1.01 (0.86-1.17)	0.78 (0.59-1.04)	0.77 (0.48-1.23)	0.98 (0.94-1.02)
From CMM to mortality	Reference	0.95 (0.75-1.21)	0.92 (0.75-1.15)	1.00 (0.81-1.24)	0.92 (0.63-1.34)	1.08 (0.66-1.76)	1.00 (0.95-1.05)

[a] VPA: vigorous PA.

[b] MVPA: moderate to vigorous PA.

[c] HR: hazard ratio.

[d] The models were adjusted for age, sex, ethnicity, socioeconomic status, BMI, household income, study center region, education group, smoking status, alcohol drinking status, and PA amount. Model A included 5 transitions from free of CMD to FCMD, then to CMM, and finally to mortality (Figure 1A), and model B included 11 transitions by including specific FCMD (Figure 1B).

[e] *P*<.001.

[f] *P*<.01.

[g] T2D: type 2 diabetes.

[h] IHD: ischemic heart disease.

[i] *P*<.05.

Subgroup Analyses

Several important findings were involved in the subgroup analyses of PA amount, and details are presented in Table 3. First, lower PA (<600 MET min/wk) accompanied by lower PA intensity, that is, >0 to 0.25 of VPA to MVPA, was not significantly associated with a decreased risk of any transition compared with no VPA. Second, the participants with a greater proportion of VPA to MPA were highly associated with a reduced risk of transitions from free of CMD to FCMD and mortality regardless of the amount of PA. Third, higher PA intensity was associated with a decreased risk of transitions from free of CMD to T2D and stroke for the participants with the recommended amount of PA (no less than 600 MET min/wk) and IHD for the participants under the threshold of the recommended amount of PA (<600 MET min/wk). These findings were similar when using the ordinal scale of PA intensity.

A significant effect of modification related to age, sex, PA amount, BMI, smoking, and alcohol drinking status was found among at least 1 transition (Table 4; Tables S4 and S5 in Multimedia Appendix 1). The protective effect of higher PA intensity on the transition from free of CMD to FCMD was larger among younger participants, participants with overweight, and previous or current smokers. Specifically, the decreased risk of the transitions from free of CMD to IHD and stroke associated with higher PA intensity was only found in previous or current smokers. Male participants, current or previous smokers, excessive alcohol drinkers, and participants with more PA could have more benefits of higher PA intensity for reducing mortality from free of CMD. In addition, the protective effect of higher PA intensity on the transitions from FCMD to mortality was more pronounced among the current or previous smokers. Younger populations aged ≤60 years could benefit from higher PA intensity when having CMM.

Table 3. Associations of physical activity (PA) intensity with transitions from free of cardiometabolic disease (CMD) to first occurrence of CMD (FCMD), cardiometabolic multimorbidity (CMM), and mortality by PA amount.

Transitions	PA intensity: the proportion of VPA[a] to MVPA[b], HR[c] (95% CI)						
	0	>0 to 0.25	>0.25 to 0.5	>0.5 to 0.75	>0.75 to <1	1	Ordinal scale
<600 MET[d] min/wk							
Model A[e]							
From free of CMD to FCMD	Reference	0.96 (0.87-1.06)	0.88 (0.83-0.94)[f]	0.89 (0.83-0.95)[f]	0.89 (0.78-1.02)	1.04 (0.94-1.14)	0.98 (0.96-0.99)[g]
From free of CMD to mortality	Reference	0.91 (0.76-1.07)	0.85 (0.76-0.94)[g]	0.76 (0.67-0.85)[f]	0.69 (0.53-0.88)[g]	0.87 (0.73-1.04)	0.94 (0.91-0.96)[f]
From FCMD to CMM	Reference	1.28 (0.97-1.69)	0.82 (0.66-1.01)	0.75 (0.59-0.96)[h]	1.16 (0.75-1.80)	0.93 (0.68-1.27)	0.96 (0.92-1.00)
From FCMD to mortality	Reference	1.07 (0.83-1.37)	0.97 (0.82-1.14)	0.90 (0.74-1.08)	1.08 (0.74-1.57)	0.91 (0.70-1.19)	0.98 (0.95-1.02)
From CMM to mortality	Reference	0.77 (0.40-1.51)	0.97 (0.61-1.54)	1.15 (0.70-1.89)	0.82 (0.26-2.56)	1.33 (0.74-2.39)	1.03 (0.94-1.13)
Model B[e]							
From free of CMD to T2D[i]	Reference	0.94 (0.79-1.12)	0.81 (0.73-0.91)[f]	0.90 (0.81-1.02)	0.87 (0.69-1.10)	1.14 (0.97-1.34)	0.98 (0.96-1.01)
From free of CMD to IHD[j]	Reference	0.93 (0.81-1.07)	0.90 (0.83-0.98)[h]	0.87 (0.80-0.96)[g]	0.90 (0.75-1.08)	1.03 (0.90-1.18)	0.98 (0.96-1.00)[h]
From free of CMD to stroke	Reference	1.09 (0.85-1.39)	0.99 (0.85-1.16)	0.94 (0.79-1.12)	0.90 (0.64-1.28)	0.85 (0.64-1.12)	0.98 (0.94-1.01)
From free of CMD to mortality	Reference	0.91 (0.76-1.07)	0.85 (0.76-0.94)[g]	0.76 (0.67-0.85)[f]	0.69 (0.54-0.88)[g]	0.87 (0.73-1.04)	0.94 (0.91-0.96)[f]
From T2D to CMM	Reference	1.07 (0.62-1.83)	0.89 (0.61-1.31)	0.75 (0.49-1.17)	1.33 (0.63-2.83)	1.01 (0.60-1.70)	0.98 (0.90-1.06)
From IHD to CMM	Reference	1.33 (0.82-2.14)	0.81 (0.54-1.20)	0.82 (0.54-1.24)	0.57 (0.18-1.80)	0.88 (0.51-1.55)	0.95 (0.87-1.03)
From stroke to CMM	Reference	0.94 (0.54-1.63)	0.76 (0.53-1.08)	0.72 (0.48-1.08)	0.99 (0.44-2.22)	0.80 (0.46-1.39)	0.93 (0.86-1.00)
From T2D to mortality	Reference	0.85 (0.57-1.25)	0.81 (0.64-1.02)	0.81 (0.62-1.05)	1.41 (0.90-2.19)	1.00 (0.71-1.40)	0.98 (0.93-1.03)
From IHD to mortality	Reference	2.60 (1.46-4.64)[g]	0.61 (0.32-1.16)	0.62 (0.29-1.34)	0.64 (0.16-2.61)	1.10 (0.45-2.71)	0.93 (0.81-1.06)
From stroke to mortality	Reference	1.43 (0.90-2.29)	1.41 (1.06-1.88)[h]	1.24 (0.87-1.75)	0.70 (0.29-1.71)	0.93 (0.49-1.75)	1.04 (0.97-1.12)
From CMM to mortality	Reference	0.81 (0.38-1.73)	1.15 (0.71-1.88)	1.30 (0.76-2.21)	1.18 (0.38-3.72)	1.53 (0.78-3.00)	1.08 (0.98-1.20)
600-1200 MET min/wk							
Model A[e]							
From free of CMD to FCMD	Reference	0.88 (0.81-0.95)[f]	0.88 (0.83-0.94)[f]	0.91 (0.85-0.97) [g]	0.95 (0.87-1.04)	0.96 (0.84-1.09)	0.98 (0.97-1.00)[h]
From free of CMD to mortality	Reference	0.95 (0.84-1.07)	0.89 (0.80-0.99)[h]	0.85 (0.77-0.95)[g]	0.80 (0.68-0.94)[g]	0.93 (0.75-1.17)	0.96 (0.93-0.98)[g]
From FCMD to CMM	Reference	0.86 (0.67-1.09)	0.98 (0.81-1.20)	0.99 (0.81-1.20)	0.86 (0.63-1.17)	1.39 (0.95-2.03)	1.01 (0.96-1.06)
From FCMD to mortality	Reference	0.89 (0.72-1.09)	0.92 (0.77-1.09)	0.92 (0.77-1.09)	0.97 (0.76-1.25)	0.98 (0.68-1.42)	0.99 (0.95-1.03)

Transitions	PA intensity: the proportion of VPA[a] to MVPA[b], HR[c] (95% CI)						
	0	>0 to 0.25	>0.25 to 0.5	>0.5 to 0.75	>0.75 to <1	1	Ordinal scale
From CMM to mortality	Reference	0.84 (0.49-1.44)	0.97 (0.64-1.46)	1.02 (0.68-1.54)	1.20 (0.64-2.27)	0.97 (0.45-2.12)	1.02 (0.92-1.12)
Model B[e]							
From free of CMD to T2D	Reference	0.77 (0.66-0.89)[f]	0.72 (0.63-0.81)[f]	0.84 (0.75-0.95)[g]	0.87 (0.74-1.04)	0.92 (0.72-1.18)	0.96 (0.93-0.99)[h]
From free of CMD to IHD	Reference	0.90 (0.81-1.00)[h]	0.94 (0.86-1.02)	0.97 (0.89-1.06)	1.03 (0.91-1.16)	1.05 (0.87-1.25)	1.00 (0.98-1.03)
From free of CMD to stroke	Reference	1.04 (0.87-1.25)	0.97 (0.83-1.14)	0.84 (0.71-1.00)[h]	0.89 (0.70-1.13)	0.71 (0.48-1.06)	0.95 (0.91-0.99)[h]
From free of CMD to mortality	Reference	0.95 (0.84-1.07)	0.89 (0.80-0.99)[h]	0.86 (0.77-0.95)[g]	0.80 (0.68-0.94)[g]	0.94 (0.75-1.18)	0.96 (0.93-0.98)[g]
From T2D to CMM	Reference	1.09 (0.70-1.72)	0.90 (0.59-1.36)	1.05 (0.71-1.53)	0.56 (0.26-1.21)	1.43 (0.66-3.10)	0.98 (0.89-1.08)
From IHD to CMM	Reference	0.95 (0.57-1.59)	1.08 (0.71-1.65)	1.17 (0.78-1.75)	1.02 (0.54-1.93)	1.63 (0.75-3.55)	1.05 (0.96-1.16)
From stroke to CMM	Reference	0.77 (0.52-1.14)	0.70 (0.50-0.99)[h]	0.92 (0.67-1.26)	0.83 (0.52-1.34)	1.50 (0.86-2.62)	1.00 (0.92-1.08)
From T2D to mortality	Reference	0.75 (0.57-1.00)	0.81 (0.64-1.02)	0.83 (0.65-1.04)	1.02 (0.75-1.40)	1.19 (0.77-1.85)	0.99 (0.93-1.05)
From IHD to mortality	Reference	0.45 (0.20-1.01)	0.96 (0.57-1.63)	0.86 (0.48-1.52)	1.09 (0.51-2.34)	0.76 (0.18-3.14)	1.00 (0.87-1.15)
From stroke to mortality	Reference	1.01 (0.70-1.44)	0.94 (0.69-1.30)	0.98 (0.70-1.36)	0.74 (0.42-1.29)	0.14 (0.02-1.03)	0.94 (0.86-1.02)
From CMM to mortality	Reference	0.62 (0.32-1.21)	0.72 (0.43-1.22)	1.02 (0.65-1.61)	1.29 (0.64-2.60)	0.90 (0.36-2.25)	1.01 (0.90-1.13)
>1200 MET min/wk							
Model A[e]							
From free of CMD to FCMD	Reference	0.91 (0.87-0.96)[f]	0.92 (0.88-0.97)[g]	0.91 (0.86-0.95)[f]	0.85 (0.79-0.91)[f]	0.86 (0.75-0.99)[h]	0.97 (0.96-0.98)[f]
From free of CMD to mortality	Reference	0.83 (0.76-0.90)[f]	0.79 (0.73-0.86)[f]	0.75 (0.69-0.81)[f]	0.67 (0.59-0.75)[f]	0.62 (0.48-0.80)[f]	0.92 (0.90-0.93)[f]
From FCMD to CMM	Reference	0.88 (0.76-1.03)	1.03 (0.89-1.20)	0.93 (0.80-1.07)	1.11 (0.90-1.36)	0.95 (0.60-1.52)	1.01 (0.97-1.05)
From FCMD to mortality	Reference	0.93 (0.82-1.06)	0.92 (0.81-1.04)	0.95 (0.84-1.07)	0.91 (0.76-1.09)	1.14 (0.80-1.62)	0.99 (0.96-1.02)
From CMM to mortality	Reference	1.07 (0.80-1.44)	0.89 (0.66-1.19)	0.93 (0.70-1.25)	0.77 (0.48-1.23)	1.02 (0.37-2.78)	0.95 (0.88-1.03)
Model B[e]							
From free of CMD to T2D	Reference	0.79 (0.71-0.87)[f]	0.86 (0.78-0.94)[g]	0.80 (0.73-0.88)[f]	0.71 (0.62-0.82)[f]	0.84 (0.64-1.09)	0.95 (0.93-0.97)[f]
From free of CMD to IHD	Reference	0.98 (0.91-1.05)	0.98 (0.92-1.05)	0.99 (0.93-1.06)	0.93 (0.85-1.02)	0.97 (0.81-1.16)	0.99 (0.98-1.01)
From free of CMD to stroke	Reference	0.93 (0.82-1.04)	0.86 (0.76-0.96)[h]	0.82 (0.73-0.93)[g]	0.82 (0.69-0.96)[h]	0.62 (0.42-0.91)[h]	0.94 (0.91-0.97)[f]
From free of CMD to mortality	Reference	0.83 (0.76-0.90)[f]	0.79 (0.73-0.86)[f]	0.75 (0.69-0.81)[f]	0.67 (0.59-0.75)[f]	0.62 (0.48-0.80)[f]	0.92 (0.90-0.93)[f]

Transitions	PA intensity: the proportion of VPA[a] to MVPA[b], HR[c] (95% CI)						
	0	>0 to 0.25	>0.25 to 0.5	>0.5 to 0.75	>0.75 to <1	1	Ordinal scale
From T2D to CMM	Reference	0.75 (0.54-1.03)	1.04 (0.78-1.38)	0.97 (0.73-1.29)	0.94 (0.61-1.44)	1.13 (0.49-2.60)	1.02 (0.95-1.10)
From IHD to CMM	Reference	0.83 (0.62-1.12)	0.73 (0.54-0.98)[h]	0.86 (0.65-1.13)	0.62 (0.38-1.01)	1.05 (0.46-2.39)	0.94 (0.87-1.02)
From stroke to CMM	Reference	1.04 (0.81-1.33)	1.05 (0.83-1.34)	0.89 (0.70-1.13)	1.05 (0.75-1.48)	1.18 (0.60-2.34)	0.98 (0.92-1.05)
From T2D to mortality	Reference	0.86 (0.72-1.02)	0.94 (0.79-1.11)	0.96 (0.81-1.12)	1.03 (0.82-1.30)	1.15 (0.73-1.82)	1.02 (0.97-1.06)
From IHD to mortality	Reference	0.74 (0.49-1.10)	1.06 (0.74-1.53)	0.95 (0.66-1.36)	1.47 (0.91-2.37)	0.44 (0.06-3.18)	1.06 (0.96-1.17)
From stroke to mortality	Reference	1.11 (0.87-1.41)	0.99 (0.78-1.27)	1.00 (0.79-1.26)	0.82 (0.56-1.21)	1.27 (0.59-2.73)	0.98 (0.92-1.04)
From CMM to mortality	Reference	0.87 (0.62-1.20)	0.78 (0.57-1.07)	0.79 (0.58-1.09)	0.64 (0.38-1.07)	0.66 (0.21-2.10)	0.92 (0.84-1.00)

[a] VPA: vigorous PA.

[b] MVPA: moderate to vigorous PA.

[c] HR: hazard ratio.

[d] MET: metabolic equivalent task.

[e] The models were adjusted for age, sex, ethnicity, socioeconomic status, BMI, household income, study center region, education group, smoking status, alcohol drinking status, and PA amount. Model A included 5 transitions from free of CMD to FCMD, then to CMM, and finally to mortality (Figure 1A), and model B included 11 transitions by including specific FCMD (Figure 1B).

[f] *P*<.001.

[g] *P*<.01.

[h] *P*<.05.

[i] T2D: type 2 diabetes.

[j] IHD: ischemic heart disease.

Table 4. Subgroup analyses for the associations of physical activity (PA) intensity with transitions from free of cardiometabolic disease (CMD) to first occurrence of CMD (FCMD), cardiometabolic multimorbidity (CMM), and mortality.

Factors for subgroup analyses	Model A[a], HR[b] (95% CI)				
	From free of CMD to FCMD	From free of CMD to mortality	From FCMD to CMM	From FCMD to mortality	From CMM to mortality
Age (years)					
≤60	0.95 (0.94-0.96)[c]	0.92 (0.90-0.94)[c]	0.97 (0.93-1.00)	0.96 (0.93-1.00)[d]	0.89 (0.81-0.98)[d]
>60	0.97 (0.96-0.98)[c]	0.94 (0.93-0.96)[c]	0.98 (0.95-1.01)	0.98 (0.95-1.00)	1.05 (0.99-1.11)
Sex					
Male	0.97 (0.96-0.98)[c]	0.92 (0.91-0.94)[c]	1.01 (0.98-1.04)	0.98 (0.96-1.01)	1.03 (0.98-1.10)
Female	0.97 (0.95-0.98)[c]	0.96 (0.94-0.98)[c]	0.99 (0.95-1.03)	1.01 (0.98-1.04)	0.97 (0.89-1.06)
BMI (kg/m^2)					
<25	0.99 (0.97-1.00)	0.93 (0.91-0.95)[c]	0.98 (0.93-1.04)	0.97 (0.94-1.01)	0.93 (0.83-1.04)
≥25	0.97 (0.96-0.98)[c]	0.94 (0.92-0.95)[c]	0.99 (0.97-1.02)	0.99 (0.96-1.01)	1.02 (0.97-1.08)
Smoking status					
Never	0.99 (0.97-1.00)[e]	0.95 (0.93-0.96)[c]	1.00 (0.97-1.04)	1.01 (0.98-1.04)	0.99 (0.92-1.07)
Previous or current	0.96 (0.95-0.97)[c]	0.92 (0.91-0.94)[c]	0.98 (0.95-1.02)	0.96 (0.94-0.99)[e]	1.02 (0.96-1.08)
Alcohol drinking status					
≤1-3 times/wk	0.97 (0.96-0.98)[c]	0.94 (0.93-0.96)[c]	0.99 (0.96-1.02)	0.99 (0.97-1.01)	0.98 (0.93-1.04)
Daily or almost daily	0.98 (0.96-0.99)[e]	0.91 (0.89-0.93)[c]	1.01 (0.96-1.07)	0.97 (0.93-1.01)	1.08 (0.97-1.19)

[a] The models were adjusted for age, sex, ethnicity, socioeconomic status, BMI, household income, study center region, education group, smoking status, alcohol drinking status, and PA amount. Age, sex, BMI, smoking status, and alcohol drinking status were not included in the models for the subgroup analyses. Model A included 5 transitions from free of CMD to FCMD, then to CMM, and finally to mortality (Figure 1A). PA intensity refers to the proportion of vigorous PA to moderate to vigorous PA, and an ordinal scale of PA intensity was used in the subgroup analyses.

[b] HR: hazard ratio.

[c] *P*<.001.

[d] *P*<.05.

[e] *P*<.01.

Sensitivity Analysis

We did not find substantial changes for the associations of PA intensity with the transitions from free of CMD to FCMD, then to CMM, and finally to mortality by performing additional analyses (Table S6 in Multimedia Appendix 1) using different time intervals (0.5 y, 1 y, 2 y, 3 y, and 5 y) for the participants entering different transitions on the same day, excluding the participants who entered different transitions on the same day, excluding FCMD in the first 2-year follow-up, excluding the participants with cancer, and excluding the participants with missing values at baseline for the covariates, which revealed robust estimates in this study.

Discussion

Principal Findings

This is the first prospective cohort study using large-scale population–based data to explore the associations of PA intensity with different transitions from free of CMD, then to CMM, and finally to mortality. This study reported the effect of the proportion of VPA to MVPA on the incidence, progression, and prognosis of CMM using multistate models, and the principal findings are discussed next. Compared with the participants with no VPA, the participants with >0.75 to <1 of VPA to MVPA had the lowest risk of transitions from free of CMD to FCMD and mortality regardless of model A or B. The protective effect of higher PA intensity on the transitions from free of CMD to T2D (*P*<.001) and from T2D to mortality (*P*<.001) was statistically significant, which revealed the importance of PA intensity for the transitions of T2D. More PA and greater intensity had a synergistic effect on decreasing the risk of transitions from free of CMD to FCMD and mortality. Male participants, younger adults, participants with a higher BMI, current or previous smokers, and excessive alcohol drinkers could obtain more benefits from higher PA intensity for the lower risk of at least 1 transition from free of CMD, then to CMM, and finally to mortality.

Interpretations, Implications, and Comparisons With Existing Literature

In line with previous studies focusing on PA amount, this study had similar findings that meeting the recommendations of PA amount could significantly decrease the risk of FCMD, CMM, and mortality, regardless of whether the traditional Cox proportional hazards regression model or multistate model was used [23-30]. This study also provided some evidence for the additional benefits of more PA (>1200 MET min/wk) for FCMD, CMM, and mortality. It is worth noting that more PA did not provide a lower risk of mortality when including CMM in the consideration of estimates from the multistate model. The aforementioned findings might be explained by the fact that the participants with CMM might have cardiovascular organic lesions and the benefits from more PA might be offset.

In line with previous studies [31-34], this study also found that higher PA intensity was significantly associated with all-cause mortality. In addition, this study added new evidence to recent knowledge that a greater proportion of VPA to MVPA (>0.75 to <1) could provide more benefits, including a decreased risk of 27% and 13% for the transitions from free of CMD to mortality and FCMD, respectively. Simultaneously, it was noted that no strong evidence was observed that a higher proportion of VPA to MVPA was associated with a lower risk of transition from free of CMD to IHD, which indicated that a suitable intensity was more effective for preventing IHD than no VPA. Another important finding in this study was that median PA intensity (the proportion of VPA to MVPA >0 to 0.75) was associated with a lower risk of transition from T2D to mortality, which suggests that stronger interventions of higher PA intensity for the participants with T2D are necessary to decrease all-cause mortality. In addition, we did not find that higher PA intensity was associated with other transitions related to CMM, which emphasizes that higher PA intensity at baseline plays an important role in the primary and secondary prevention for CMM and has a weaker role in the prognosis of CMM.

A national cohort study of >400,000 US adults reported a decreased risk of all-cause mortality associated with the synergistic effect of more PA and greater intensity [32]. This study reported similar results for the associations of an interactive effect of PA amount and intensity on transitions from free of CMD to mortality. The aforementioned associations were also significant in the transition from free of CMD to FCMD, which not only emphasized the importance of PA amount but also highlighted the importance of higher PA intensity for healthy outcomes and the development of CMM.

Inconsistent with a previous study reporting no interactions for the associations of sex with all-cause mortality [31,32], this study found that male participants could benefit more from higher PA intensity for reducing mortality from the free of CMD status. In line with the study by Wang et al [32], we also found significant interactions of current or previous smoking status and PA intensity with all-cause mortality. These differences might be caused by the different models used in this study. Instead of traditional Cox proportional hazard models used in previous studies, this study used the multistate models, which introduced a competing risk for all-cause mortality and could more validly show the transitions from free of CMD to mortality. In addition, this study provided additional evidence for identifying sensitive populations for other transitions in the incidence, progression, and prognosis of CMM, which could help develop preventive, targeted strategies for sensitive populations in the transitions.

Strengths and Limitations

This study has several strengths. First, this study used multistate models considering the competing risk for all-cause mortality to assess the associations of PA intensity, namely the proportion of VPA to MVPA with the transitions from free of CMD to FCMD, then to CMM, and finally to mortality. Second, this study used a large-scale population–based cohort, which allowed us to explore the associations among different transitions using the prospective design. Third, long-term follow-ups of UKB in this study could provide a sufficient duration to observe the incidence, progression, and prognosis of CMM. Certain limitations should be approached with caution when interpreting the results of this study. First, the PA amount and intensity were self-reported, and more studies using portable equipment are needed to verify these associations. In addition, we only captured the baseline characteristics of PA in this study. Further studies should focus on the effects of the trajectory of PA on the transitions. In addition, although UKB had a low recruitment rate [35] and a limited age group from 37 to 73 years, a previous study proved that the associations of PA with mortality were not materially changed [45].

Clinical Implications and Conclusions

This study suggests that higher PA intensity is an effective measure for preventing CMM and mortality in the early period of CMM development. Higher PA intensity accompanied by higher PA levels could provide the lowest risk of all-cause mortality and FCMD. Targeted populations including male participants, younger adults, people with overweight, current or previous smokers, and excessive alcohol drinkers could obtain more benefits from higher PA intensity for FCMD and all-cause mortality. Relevant interventions related to higher PA intensity should be performed.

Acknowledgments

The authors would like to thank the participants of the UK Biobank. This research was conducted using the UK Biobank Resource under application number 91536. B-PL was supported by the National Natural Science Foundation of China (NSFC; #82103954) and Shandong Provincial Natural Science Foundation (#ZR2021QH310). C-XJ was supported by the NSFC (#81761128033). The funding sources were not involved in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Data Availability

UK Biobank data can be accessed by submitting an application through the UK Biobank website.

Authors' Contributions

B-PL and C-XJ conceptualized the research question. B-PL conducted data analysis and managed the data. B-PL, J-HZ, L-PW, Z-YZ, and XW drafted the paper. All authors provided critical feedback on the paper at various stages and approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study flow, design, covariates, and additional analyses.
[DOCX File , 459 KB - publichealth_v9i1e46991_app1.docx]

References

1. Multimorbidity: a priority for global health research. The Academy of Medical Sciences. 2018 Apr. URL: https://acmedsci.ac.uk/file-download/82222577 [accessed 2023-08-30]
2. Emerging Risk Factors Collaboration, Di Angelantonio E, Kaptoge S, Wormser D, Willeit P, Butterworth AS, et al. Association of cardiometabolic multimorbidity with mortality. JAMA 2015 Jul 07;314(1):52-60 [FREE Full text] [doi: 10.1001/jama.2015.7008] [Medline: 26151266]
3. Canoy D, Tran J, Zottoli M, Ramakrishnan R, Hassaine A, Rao S, et al. Association between cardiometabolic disease multimorbidity and all-cause mortality in 2 million women and men registered in UK general practices. BMC Med 2021 Oct 28;19(1):258 [FREE Full text] [doi: 10.1186/s12916-021-02126-x] [Medline: 34706724]
4. Cheng X, Ma T, Ouyang F, Zhang G, Bai Y. Trends in the prevalence of cardiometabolic multimorbidity in the United States, 1999-2018. Int J Environ Res Public Health 2022 Apr 14;19(8):4726 [FREE Full text] [doi: 10.3390/ijerph19084726] [Medline: 35457593]
5. Dove A, Marseglia A, Shang Y, Grande G, Vetrano DL, Laukka EJ, et al. Cardiometabolic multimorbidity accelerates cognitive decline and dementia progression. Alzheimers Dement 2022 Jun 16 (forthcoming). [doi: 10.1002/alz.12708] [Medline: 35708183]
6. Maddaloni E, D'Onofrio L, Alessandri F, Mignogna C, Leto G, Pascarella G, CoViDiab Study Group. Cardiometabolic multimorbidity is associated with a worse Covid-19 prognosis than individual cardiometabolic risk factors: a multicentre retrospective study (CoViDiab II). Cardiovasc Diabetol 2020 Oct 01;19(1):164 [FREE Full text] [doi: 10.1186/s12933-020-01140-2] [Medline: 33004045]
7. Tai XY, Veldsman M, Lyall DM, Littlejohns TJ, Langa KM, Husain M, et al. Cardiometabolic multimorbidity, genetic risk, and dementia: a prospective cohort study. Lancet Healthy Longev 2022 Jun;3(6):e428-e436 [FREE Full text] [doi: 10.1016/S2666-7568(22)00117-9] [Medline: 35711612]
8. Huang ZT, Luo Y, Han L, Wang K, Yao SS, Su HX, et al. Patterns of cardiometabolic multimorbidity and the risk of depressive symptoms in a longitudinal cohort of middle-aged and older Chinese. J Affect Disord 2022 Mar 15;301:1-7. [doi: 10.1016/j.jad.2022.01.030] [Medline: 34999125]
9. Luo H, Zhang Q, Yu K, Meng X, Kan H, Chen R. Long-term exposure to ambient air pollution is a risk factor for trajectory of cardiometabolic multimorbidity: a prospective study in the UK Biobank. EBioMedicine 2022 Oct;84:104282 [FREE Full text] [doi: 10.1016/j.ebiom.2022.104282] [Medline: 36174399]
10. Lu Y, Li G, Ferrari P, Freisling H, Qiao Y, Wu L, et al. Associations of handgrip strength with morbidity and all-cause mortality of cardiometabolic multimorbidity. BMC Med 2022 Jun 03;20(1):191 [FREE Full text] [doi: 10.1186/s12916-022-02389-y] [Medline: 35655218]
11. He L, Ma T, Li J, Luo Y, Zhang G, Cheng X, et al. Adherence to a healthy sleep pattern and incidence of cardiometabolic multimorbidity among hypertensive patients: a prospective study of UK Biobank. Sleep 2022 Oct 10;45(10):zsac141 [FREE Full text] [doi: 10.1093/sleep/zsac141] [Medline: 35738866]
12. Luo Y, He L, Ma T, Li J, Bai Y, Cheng X, et al. Associations between consumption of three types of beverages and risk of cardiometabolic multimorbidity in UK Biobank participants: a prospective cohort study. BMC Med 2022 Aug 18;20(1):273 [FREE Full text] [doi: 10.1186/s12916-022-02456-4] [Medline: 35978398]
13. Yang L, Luo Y, He L, Yin J, Li T, Liu S, et al. Shift work and the risk of cardiometabolic multimorbidity among patients with hypertension: a prospective cohort study of UK Biobank. J Am Heart Assoc 2022 Sep 06;11(17):e025936 [FREE Full text] [doi: 10.1161/JAHA.122.025936] [Medline: 36036170]

14. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ. Physical activity and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis. Eur J Epidemiol 2015 Jul;30(7):529-542. [doi: 10.1007/s10654-015-0056-z] [Medline: 26092138]
15. Kolb H, Martin S. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. BMC Med 2017 Jul 19;15(1):131 [FREE Full text] [doi: 10.1186/s12916-017-0901-x] [Medline: 28720102]
16. Smith AD, Crippa A, Woodcock J, Brage S. Physical activity and incident type 2 diabetes mellitus: a systematic review and dose-response meta-analysis of prospective cohort studies. Diabetologia 2016 Dec;59(12):2527-2545 [FREE Full text] [doi: 10.1007/s00125-016-4079-0] [Medline: 27747395]
17. Bahls M, Leitzmann MF, Karch A, Teumer A, Dörr M, Felix SB, et al. Physical activity, sedentary behavior and risk of coronary artery disease, myocardial infarction and ischemic stroke: a two-sample Mendelian randomization study. Clin Res Cardiol 2021 Oct 27;110(10):1564-1573 [FREE Full text] [doi: 10.1007/s00392-021-01846-7] [Medline: 33774696]
18. Koolhaas CM, Dhana K, Golubic R, Schoufour JD, Hofman A, van Rooij FJ, et al. Physical activity types and coronary heart disease risk in middle-aged and elderly persons: the Rotterdam study. Am J Epidemiol 2016 Apr 15;183(8):729-738. [doi: 10.1093/aje/kwv244] [Medline: 27022033]
19. Winzer EB, Woitek F, Linke A. Physical activity in the prevention and treatment of coronary artery disease. J Am Heart Assoc 2018 Mar 08;7(4):e007725 [FREE Full text] [doi: 10.1161/JAHA.117.007725] [Medline: 29437600]
20. Hung SH, Ebaid D, Kramer S, Werden E, Baxter H, Campbell BC, et al. Pre-stroke physical activity and admission stroke severity: a systematic review. Int J Stroke 2021 Dec 23;16(9):1009-1018. [doi: 10.1177/1747493021995271] [Medline: 33527883]
21. Kramer SF, Hung SH, Brodtmann A. The impact of physical activity before and after stroke on stroke risk and recovery: a narrative review. Curr Neurol Neurosci Rep 2019 Apr 22;19(6):28. [doi: 10.1007/s11910-019-0949-4] [Medline: 31011851]
22. MacDonald CJ, Madika AL, Gomes R, Severi G, Sibon I, Debette S, et al. Physical activity and stroke among women - a non-linear relationship. Prev Med 2021 Sep;150:106485. [doi: 10.1016/j.ypmed.2021.106485] [Medline: 33647351]
23. Freisling H, Viallon V, Lennon H, Bagnardi V, Ricci C, Butterworth AS, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. BMC Med 2020 Jan 10;18(1):5 [FREE Full text] [doi: 10.1186/s12916-019-1474-7] [Medline: 31918762]
24. Han Y, Hu Y, Yu C, Guo Y, Pei P, Yang L, China Kadoorie Biobank Collaborative Group. Lifestyle, cardiometabolic disease, and multimorbidity in a prospective Chinese study. Eur Heart J 2021 Sep 07;42(34):3374-3384 [FREE Full text] [doi: 10.1093/eurheartj/ehab413] [Medline: 34333624]
25. Xie H, Li J, Zhu X, Li J, Yin J, Ma T, et al. Association between healthy lifestyle and the occurrence of cardiometabolic multimorbidity in hypertensive patients: a prospective cohort study of UK Biobank. Cardiovasc Diabetol 2022 Oct 01;21(1):199 [FREE Full text] [doi: 10.1186/s12933-022-01632-3] [Medline: 36183084]
26. He L, Biddle SJ, Lee JT, Duolikun N, Zhang L, Wang Z, et al. The prevalence of multimorbidity and its association with physical activity and sleep duration in middle aged and elderly adults: a longitudinal analysis from China. Int J Behav Nutr Phys Act 2021 Jun 10;18:77. [doi: 10.1186/s12966-021-01150-7]
27. Lee DH, Rezende LF, Joh HK, Keum N, Ferrari G, Rey-Lopez JP, et al. Long-term leisure-time physical activity intensity and all-cause and cause-specific mortality: a prospective cohort of US adults. Circulation 2022 Aug 16;146(7):523-534 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.121.058162] [Medline: 35876019]
28. Samitz G, Egger M, Zwahlen M. Domains of physical activity and all-cause mortality: systematic review and dose-response meta-analysis of cohort studies. Int J Epidemiol 2011 Oct;40(5):1382-1400 [FREE Full text] [doi: 10.1093/ije/dyr112] [Medline: 22039197]
29. von Rosen P, Dohrn IM, Hagströmer M. Association between physical activity and all-cause mortality: a 15-year follow-up using a compositional data analysis. Scand J Med Sci Sports 2020 Jan;30(1):100-107. [doi: 10.1111/sms.13561] [Medline: 31581345]
30. Blond K, Brinkløv CF, Ried-Larsen M, Crippa A, Grøntved A. Association of high amounts of physical activity with mortality risk: a systematic review and meta-analysis. Br J Sports Med 2020 Oct;54(20):1195-1201. [doi: 10.1136/bjsports-2018-100393] [Medline: 31406017]
31. Gebel K, Ding D, Chey T, Stamatakis E, Brown WJ, Bauman AE. Effect of moderate to vigorous physical activity on all-cause mortality in middle-aged and older Australians. JAMA Intern Med 2015 Jun 01;175(6):970-977 [FREE Full text] [doi: 10.1001/jamainternmed.2015.0541] [Medline: 25844882]
32. Wang Y, Nie J, Ferrari G, Rey-Lopez JP, Rezende LF. Association of physical activity intensity with mortality: a national cohort study of 403 681 US adults. JAMA Intern Med 2021 Mar 01;181(2):203-211 [FREE Full text] [doi: 10.1001/jamainternmed.2020.6331] [Medline: 33226432]
33. Zhao M, Veeranki SP, Li S, Steffen LM, Xi B. Beneficial associations of low and large doses of leisure time physical activity with all-cause, cardiovascular disease and cancer mortality: a national cohort study of 88,140 US adults. Br J Sports Med 2019 Nov 19;53(22):1405-1411. [doi: 10.1136/bjsports-2018-099254] [Medline: 30890520]
34. Rey Lopez JP, Gebel K, Chia D, Stamatakis E. Associations of vigorous physical activity with all-cause, cardiovascular and cancer mortality among 64 913 adults. BMJ Open Sport Exerc Med 2019 Aug 12;5(1):e000596 [FREE Full text] [doi: 10.1136/bmjsem-2019-000596] [Medline: 31548909]

35. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015 Mar;12(3):e1001779 [FREE Full text] [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
36. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
37. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/336656 [accessed 2023-08-30]
38. Del Pozo Cruz B, Ahmadi MN, Lee IM, Stamatakis E. Prospective associations of daily step counts and intensity with cancer and cardiovascular disease incidence and mortality and all-cause mortality. JAMA Intern Med 2022 Nov 01;182(11):1139-1148 [FREE Full text] [doi: 10.1001/jamainternmed.2022.4000] [Medline: 36094529]
39. Ding D, Van Buskirk J, Nguyen B, Stamatakis E, Elbarbary M, Veronese N, et al. Physical activity, diet quality and all-cause cardiovascular disease and cancer mortality: a prospective study of 346 627 UK Biobank participants. Br J Sports Med 2022 Jul 10:bjsports-2021-105195. [doi: 10.1136/bjsports-2021-105195] [Medline: 35811091]
40. Textor J, van der Zander B, Gilthorpe MS, Liskiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int J Epidemiol 2016 Dec 01;45(6):1887-1894 [FREE Full text] [doi: 10.1093/ije/dyw341] [Medline: 28089956]
41. Tyrrell J, Jones SE, Beaumont R, Astley CM, Lovell R, Yaghootkar H, et al. Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. BMJ 2016 Mar 08;352:i582 [FREE Full text] [doi: 10.1136/bmj.i582] [Medline: 26956984]
42. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007 May 20;26(11):2389-2430. [doi: 10.1002/sim.2712] [Medline: 17031868]
43. Wreede LC, Fiocco M, Putter H. mstate: an R Package for the analysis of competing risks and multi-state models. J Stat Softw 2011 Jan 4;38(7):1-30. [doi: 10.18637/jss.v038.i07]
44. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003 Jan 25;326(7382):219 [FREE Full text] [doi: 10.1136/bmj.326.7382.219] [Medline: 12543843]
45. Stamatakis E, Owen KB, Shepherd L, Drayton B, Hamer M, Bauman AE. Is cohort representativeness Passé? Poststratified associations of lifestyle risk factors with mortality in the UK Biobank. Epidemiology 2021 Mar 01;32(2):179-188 [FREE Full text] [doi: 10.1097/EDE.0000000000001316] [Medline: 33492009]

Abbreviations

CMD: cardiometabolic disease
CMM: cardiometabolic multimorbidity
FCMD: first occurrence of cardiometabolic disease
HR: hazard ratio
IHD: ischemic heart disease
MET: metabolic equivalent task
MPA: moderate physical activity
MVPA: moderate to vigorous physical activity
PA: physical activity
SES: socioeconomic status
T2D: type 2 diabetes
UKB: UK Biobank
VPA: vigorous physical activity

Edited by A Mavragani, T Sanchez; submitted 04.03.23; peer-reviewed by S Robinson, T Liu; comments to author 12.05.23; revised version received 02.06.23; accepted 10.08.23; published 25.09.23.

Please cite as:
Liu BP, Zhu JH, Wan LP, Zhao ZY, Wang X, Jia CX
The Impact of Physical Activity Intensity on the Dynamic Progression of Cardiometabolic Multimorbidity: Prospective Cohort Study Using UK Biobank Data
JMIR Public Health Surveill 2023;9:e46991
URL: https://publichealth.jmir.org/2023/1/e46991
doi:10.2196/46991
PMID:37747776

Original Paper

Spatiotemporal Patterns of Esophageal Cancer Burden Attributable to Behavioral, Metabolic, and Dietary Risk Factors From 1990 to 2019: Longitudinal Observational Study

Peng Li[1,2], PhD; Jing Jing[3], PhD; Wenjun Liu[4], PhD; Jizhao Wang[1], MD, PhD; Xin Qi[2], PhD; Guangjian Zhang[1], MD, PhD

[1]Department of Thoracic Surgery, The First Affiliated Hospital of Xian Jiaotong University, Xi'an, China

[2]Department of Epidemiology and Health Statistics, School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, China

[3]College of Geography and Environment, Baoji University of Arts and Sciences, Baoji, China

[4]Department of Cardiovascular Medicine, The First Affiliated Hospital of Xian Jiaotong University, Xi'an, China

Corresponding Author:
Guangjian Zhang, MD, PhD
Department of Thoracic Surgery
The First Affiliated Hospital of Xian Jiaotong University
Yanta West Road
Xi'an, 710061
China
Phone: 86 85323860
Email: michael8039@xjtu.edu.cn

Abstract

Background: Esophageal cancer (EC) is the sixth leading cause of cancer-related burden with distinct regional variations globally. Although the burden of EC has decreased, the specific reasons for this decline are still unclear.

Objective: This study aims to uncover the spatiotemporal patterns of EC risk–attributable burden in 204 countries and territories from 1990 to 2019 so that prevention and control strategies of EC can be prioritized worldwide.

Methods: We extracted EC risk–attributable deaths, disability-adjusted life years (DALYs), age-standardized mortality rates (ASMRs), and age-standardized DALY rates (ASDRs) from the global burden of disease (GBD) study from 1990 to 2019, in terms of behavioral, metabolic, and dietary factors by age, sex, and geographical location. Average annual percentage change (AAPC) was used to assess the long-term trends in the ASMRs and ASDRs of EC due to specific risk factors.

Results: Between 1990 and 2019, the greatest decrease in EC burden was attributed to low intake of fruits and vegetables. An AAPC of –2.96 (95% CI –3.28 to –2.63) and –3.12 (95% CI –3.44 to –2.79) in ASMR and ASDR was attributable to a low-fruit diet, while an AAPC of –3.60 (95% CI –3.84 to –3.36) and –3.64 (95% CI –3.92 to –3.35) in ASMR and ASDR was attributed to a low-vegetable diet. However, the trends in ASMRs and ASDRs due to high BMI showed significant increases with an AAPC of 0.52 (95% CI 0.29-0.75) in ASMR and 0.42 (95% CI 0.18-0.66) in ASDR from 1990 to 2019 compared to significant decreases in other attributable risks with AAPC<0 ($P<.05$). East Asia had the largest decrease in EC burden due to low-vegetable diets, with an AAPC of –11.00 (95% CI –11.32 to –10.67) in ASMR and –11.81 (95% CI –12.21 to –11.41) in ASDR, followed by Central Asia, whereas Western Sub-Saharan Africa had the largest increase in ASMR and ASDR due to high BMI, with an AAPC of 3.28 (95% CI 3.14-3.42) and 3.09 (95% CI 2.96-3.22), respectively. China had the highest EC burden attributed to smoking, alcohol use, high BMI, and low-fruit diets. Between 1990 and 2019, there was a significant decrease in EC burden attributable to smoking, alcohol use, chewing tobacco, low-fruit diets, and low-vegetable diets in most countries, wherein a significant increase in the EC burden was due to high BMI.

Conclusions: Our study shows that smoking and alcohol consumption are still the leading risk factors of EC burden and that EC burden attributable to low intake of fruits and vegetables has shown the largest decline recently. The risks of ASMRs and ASDRs of EC showed distinct spatiotemporal patterns, and future studies should focus on the upward trend in the EC burden attributed to high BMI.

(JMIR Public Health Surveill 2023;9:e46051) doi:10.2196/46051

KEYWORDS

esophageal cancer; age-standardized mortality rate; disability-adjusted life years; attributable risk; average annual percentage change

Introduction

In 2019, esophageal cancer (EC) was the sixth leading cause of cancer deaths and disability-adjusted life years (DALYs), with over 0.54 million deaths and 1.17 million DALYs, representing 50% of all cancers worldwide [1]. Early diagnosis and treatment have been implemented to reduce EC incidence and mortality rates and to improve survival rates over years. However, patients with EC still have low 5-year survival rates (about 30.3%) due to unclear and nonspecific early clinical symptoms, and most patients are diagnosed at an advanced stage [2]. Therefore, EC is still an important public health and clinical challenge for early diagnosis and prevention, especially in low-income countries. Besides, EC burden greatly varies with geographical locations, and the "EC belt" has substantially high EC incidence and mortality rates worldwide [3]. The disparities in the burden of cancer and the poor treatment outcomes across regions worldwide are reflected in the Sustainable Development Goal target 3.4 that focuses on reducing premature mortality from noncommunicable diseases (eg, cancers) by one-third by promoting mental health and well-being by 2030 through prevention and treatment [4].

EC has 2 main histological subtypes, that is, esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC), which have analogous and distinct etiologies with modifiable risk factors (eg, smoking, alcohol use, obesity, diet) [5,6]. EC cases with over 85% ESCC histological subtype cluster in low-income countries and has recently shown a significantly downward trend worldwide [7,8]. Previous studies have reported the epidemiological characteristics of EC, such as incidence, mortality, and DALY rates in several countries or regions over several years [9,10]. Age-standardized incidence rates of EC in China have declined significantly from 1990 to 2019, and the decrease is more pronounced among females than among males [11]. Moreover, there are significant differences in the risk factors between sexes [11]. There was a significant male predominance in EC cases with a male-to-female incidence ratio of 3.3:1 for ESCC and 6.7:1 for EAC [8]. Further, the interaction effect of sex and smoking in ESCC has been reported in a cohort study with 13 years of median follow-up [12]. However, comprehensive understanding of EC risk–attributable burden does not exist in many countries, leading to a void in the up-to-date prevention strategies and adjustment in the EC treatment research focus. At this point, the latest changes in the specified risk-attributable burden of EC should be scheduled at global, regional, and national scales to elucidate the trend in EC cases that are attributable to primary factors (eg, smoking, alcohol use, metabolism, diet). Addressing the primary factors contributing to EC burden can lead to advances in the prognosis and treatment of EC in countries worldwide that require EC prevention and control measures.

To date, the Global Burden of Disease (GBD) study has assessed cancer burden by broadly considering the major modifiable risk factors across countries, ages, and sexes worldwide over time. To our knowledge, GBD is the only database that quantifies and provides a comprehensive picture of the cancer burden attributable to a set of risk factors through stringent quality assurance procedures to model at different scales over time. Further, GBD 2019 provides the latest information on each risk-outcome pair for EC in 204 countries and territories and its aggregation regions from 1990 to 2019, which enables the evaluation of the risk-attributable burden for low-income countries. Herein, we analyzed the spatiotemporal patterns in the burden of EC attributable to specific behavioral, metabolic, and dietary risk factors at global, regional, and national levels over 3 decades, which can help policy makers implement prevention measures and tackle the growing EC burden driven by specific risk factors.

Methods

Data Source

We collected the annual number of EC (International Classification of Diseases 10th revision code C15) deaths, DALYs, age-standardized mortality rates (ASMRs), and age-standardized DALY rates (ASDRs) related to risk-attributable factors by sex and age at global, regional, and national levels from GBD 2019 estimations from 1990 to 2019. The data included EC risk–attributable deaths, DALYs, ASMRs, and ASDRs in 204 countries and territories and were clustered in 21 GBD regions according to geographical proximity and economic similarities [13]. Subsequently, according to the sociodemographic index (SDI) ranging from 0 to 1, 204 countries and territories were classified under 5 SDI levels: low SDI (<0.45), low-middle SDI (≥0.45 to <0.61), middle SDI (≥0.61 to <0.69), high-middle SDI (≥0.69 to <0.80), and high SDI (≥0.80). The SDI is a composite indicator of a country's lag-distributed income per capita, average years of schooling (aged >15 years), and fertility rate in females younger than 25 years [14].

Exposure Selection and Definition

By reviewing previous literature and studies in terms of the leading risk factors of EC burden, we targeted the EC burden attributable to the most prominent risk factors, including behavioral risk (ie, smoking, alcohol use, and chewing tobacco), metabolic risk (ie, high BMI), and diet risk (ie, diet with less vegetables and fruits) [15,16]. We focused on these leading factors (established EC risk factors with available data in GBD) to analyze the spatiotemporal patterns of EC burden in different regions in order to implement EC prevention and intervention measures. In this study, smoking is defined as the prevalence of current use of any smoked tobacco product and prevalence of the former use of any smoked tobacco product. Alcohol use is defined as the average daily alcohol consumption of pure alcohol (measured in grams per day) by current drinkers who had consumed alcohol during the past 12 months. Chewing tobacco is defined as the current use of any chewing tobacco product. High BMI is defined as BMI above the normal range (18.5-24.9 kg/m^2). Diet low in vegetables and fruits is defined as the average daily consumption of vegetables and fruits (measured in grams per day). These variables are well-established risk factors for EC with available data in GBD and whose detailed definitions and risk gradients for EC are elaborated elsewhere [17]. Because EC incidence and mortality have been hardly reported among individuals younger than 40 years [18], we considered 0-39 years old as one age group, individuals aged 40-94 years as smaller subgroups each with an age range of 5 years, and >95 years old as the last age group.

Statistical Analysis

Estimation of the EC Burden

Data on deaths, DALYs, ASMRs, and ASDRs were reported with numbers and their 95% uncertainty intervals (UIs), generating values at the 2.5th and 97.5th percentile from 1000 draws. The methods for the estimation of explaining uncertainty in the risk-attributable burden have been described in detail elsewhere [15]. Briefly, the GBD comparative risk assessment was divided into 6 key processes for each risk-outcome estimation. First, we identified convincing or probable evidence for causal association with EC based on systematic reviews and meta-regression. Second, we estimated the relative risks for each risk-outcome pair as risk-attributable function of EC by using GBD's meta-regression-Bayesian, regularized, and trimmed method. Third, we analyzed the model risk exposure level and distribution by age, sex, location, and year by using Bayesian meta-regression modelling (DisMod-MR 2.1). Fourth, we determined the theoretical minimum risk exposure level based on a counterfactual scenario (a given population receives the optimal level of risk exposure) and the 85th percentile of exposure in cohorts and trial studies. Fifth, we calculated the population attributable fractions for each risk-outcome pair by age, sex, location, and year. Sixth, we estimated the EC burden attributable to specified risk factors, wherein deaths and DALYs were multiplied by the corresponding risk factor population attributable fraction [19,20].

Analysis of the Time Trends of EC Burden

An average annual percentage change (AAPC) was used to estimate the long-term trends in ASMR and ASDR in EC cases attributable to each risk factor from 1990 to 2019. In this study, AAPC was widely used to analyze a summary measure of time trend over an entire period, which was fitted in a regression by the natural logarithm of the ASMR and ASDR from 1990 to 2019, that is, $y = \alpha+\beta x+$, where $y = \ln$ (ASMR or ASDR) and x = calendar year. The AAPC was calculated as $100 \times (\exp(\beta)-1)$, and its 95% CI can also be obtained from the fitted regression model. The time trend of ASMR or ASDR was defined as a significant increase if AAPC>0 and its lower boundary of 95% CI>0 (*P*<.05). Inversely, the time trend of ASMR or ASDR was defined as a significant decrease if the AAPC<0 and its upper boundary of 95% CI<0 (*P*<.05). Otherwise, it was not a significant trend over time. In addition, we assessed the time trends of ASMR and ASDR attributable to each specific risk factor for EC by age group and sex at regional, SDI, and national levels from 1990 to 2019. These statistical analyses were conducted using Joinpoint trend analysis software (version 4.9.1.0; Statistical Methodology and Applications Branch), and corresponding geographical patterns were presented by ArcGIS 10.6 (ESRI). A *P* value less than .05 was considered statistically significant.

Ethical Considerations

This study did not involve human participants and animals. Ethics approval was not applicable for this study, as this study used existing good quality modeled data that were aggregated at the population level.

Results

Global Trends in EC Risk–Attributable Burden

As Table 1 shows, although the total number of EC risk–attributable deaths and DALYs globally increased in both sexes combined between 1990 and 2019, the corresponding ASMRs and ASDRs decreased. The deaths attributable to smoking increased from 134,682 (95% UI 105,311-152,927) in 1990 to 203,328 (95% UI 170,461-236,522) in 2019, while the corresponding ASMR decreased from 3.42 (95% UI 2.69-3.86) per 100,000 people in 1990 to 2.48 (95% UI 2.08-2.89) per 100,000 people in 2019. The DALYs and ASDRs attributable to smoking followed a similar pattern during the same period. Besides, in 2019, the higher global EC burden was attributable to specifically behavioral, metabolic, and dietary factors among males (Table 1).

Table 1. Global burden of esophageal cancer attributable to behavioral, metabolic, and diet factors in 1990 and 2019.

Risk factor	1990				2019			
	Death (persons, 95% UI[a])	ASMR[b] (95% UI)	DALY[c] (95% UI)	ASDR[d] (95% UI)	Death (persons, 95% UI)	ASMR (95% UI)	DALY (95% UI)	ASDR (95% UI)
Smoking								
B[e]	134,682 (105,311-152,927)	3.42 (2.69-3.86)	3,475,304 (2,674,422-3,967,053)	84.38 (65.34-96.30)	203,328 (170,462-236,512)	2.48 (2.08-2.89)	4,746,524 (3,983,523-5,544,229)	56.71 (47.57-66.14)
M[f]	120,286 (92,296-137,575)	6.70 (5.22-7.62)	3,160,591 (2,404,539-3,635,360)	162.14 (123.89-186.04)	187,234 (156,028-219,347)	4.95 (4.12-5.78)	4,425,432 (3,688,962-5,188,276)	111.18 (92.75-130.29)
F[g]	14,396 (11,141-17,351)	0.69 (0.53-0.83)	314,713 (242,655-382,305)	14.73 (11.41-17.87)	16,094 (12,861-19,289)	0.37 (0.29-0.44)	321,092 (260,159-383,914)	7.33 (5.94-8.76)
Alcohol use								
B	65,669 (47,318-84,125)	1.64 (1.19-2.11)	1,792,113 (1,277,343-2,281,636)	42.97 (30.73-54.66)	113,600 (84,063-144,686)	1.38 (1.02-1.76)	2,818,188 (2,109,628-3,573,631)	33.55 (25.09-42.54)
M	58,665 (41,912-74,762)	3.17 (2.30-4.04)	1,628,732 (1,160,091-2,068,043)	81.57 (58.12-103.62)	103,883 (77,092-132,860)	2.71 (2.01-3.46)	2,609,021 (1,943,902-3,312,769)	64.84 (48.32-82.39)
F	7004 (4724-9732)	0.34 (0.23-0.47)	163,381 (110,272-227,901)	7.62 (5.15-10.61)	9718 (6701-13,143)	0.22 (0.15-0.30)	209,167 (145,394-281,635)	4.80 (3.33-6.46)
Chewing tobacco								
B	9398 (6604-12,589)	0.24 (0.17-0.31)	261,073 (182,523-351,920)	6.25 (4.38-8.40)	18,277 (12,696-24,652)	0.22 (0.15-0.30)	475,862 (327,634-643,826)	5.66 (3.91-7.66)
M	6463 (4059-9238)	0.35 (0.22-0.50)	184,101 (115,166-263,107)	9.16 (5.75-13.08)	12,277 (7441-17,718)	0.32 (0.19-0.46)	330,479 (200,616-476,315)	8.14 (4.94-11.74)
F	2935 (1797-4434)	0.14 (0.08-0.21)	76,972 (46,998-116,401)	3.57 (2.18-5.38)	6000 (3649-8965)	0.14 (0.08-0.21)	145,383 (88,251-216,523)	3.34 (2.03-4.99)
High BMI								
B	35,283 (10,182-75,927)	0.90 (0.26-1.97)	917,929 (256,226-1,976,144)	22.26 (6.29-47.89)	89,904 (27,879-171,255)	1.09 (0.34-2.10)	2,202,314 (681,901-4,173,080)	26.27 (8.12-49.89)
M	23,273 (3778-54,281)	1.28 (0.21-2.99)	634,186 (103,387-1,476,891)	32.05 (5.21-74.68)	65,511 (12,587-135,689)	1.71 (0.33-3.54)	1,653,597 (320,884-3,422,784)	41.06 (7.95-85.11)
F	12,010 (640-30,139)	0.58 (0.03-1.44)	283,743 (15,239-714,960)	13.20 (0.71-33.16)	24,393 (1296-53,852)	0.56 (0.03-1.23)	548,718 (29,096-1,198,556)	12.59 (0.67-27.51)
Diet low in fruits								
B	51,867 (17,816-92,689)	1.32 (0.45-2.37)	1,358,518 (473,229-2,387,195)	32.83 (11.40-57.97)	51,210 (15,227-108,734)	0.63 (0.19-1.33)	1,249,775 (384,470-2,595,057)	14.96 (4.60-31.05)
M	33,861 (11,023-61,135)	1.86 (0.60-3.39)	924,139 (307,556-1,665,445)	46.68 (15.36-83.83)	35,701 (10,018-78,761)	0.94 (0.26-2.09)	892,474 (262,854-1,927,948)	22.31 (6.49-48.36)
F	18,006 (6627-32,143)	0.86 (0.31-1.53)	434,380 (160,567-761,632)	20.12 (7.43-35.37)	15,509 (5185-30,419)	0.35 (0.12-0.70)	357,301 (128,558-668,342)	8.23 (2.98-15.36)
Diet low in vegetables								
B	21,591 (2291-44,769)	0.55 (0.06-1.15)	550,996 (57,581-1,145,982)	13.41 (1.41-27.88)	17,176 (2549-33,958)	0.21 (0.03-0.42)	420,309 (64,154-827,694)	5.03 (0.77-9.91)
M	14,339 (1421-29,787)	0.80 (0.08-1.67)	379,595 (37,236-789,676)	19.45 (1.92-40.34)	11,573 (1709-23,211)	0.31 (0.04-0.62)	288,897 (43,912-575,697)	7.23 (1.09-14.42)
F	7252 (797-14,894)	0.35 (0.04-0.71)	171,401 (19,849-350,757)	7.96 (0.92-16.31)	5603 (868-10,917)	0.13 (0.02-0.25)	131,412 (20,890-258,144)	3.03 (0.48-5.95)

[a] UI: uncertainty interval.

[b] ASMR: age-standardized mortality rate per 100,000 people.

[c] DALY: disability-adjusted life year of persons.

[d] ASDR: age-standardized disability-adjusted life year rate per 100,000 people.

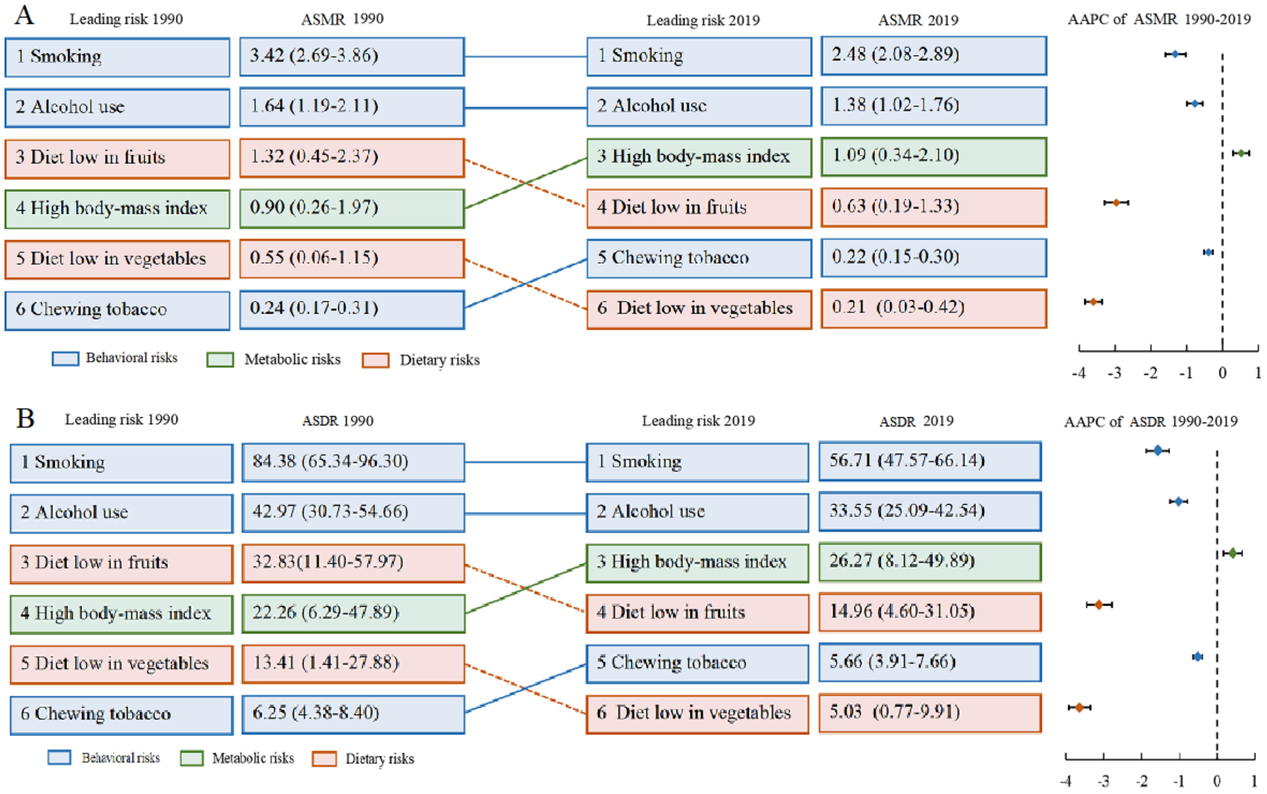
[e] B: both male and female.

[f] M: male.

[g] F: female.

Behavioral risks include smoking, alcohol use, and chewing tobacco; metabolic risks include high BMI; and diet risks include diets low in fruits and vegetables. Figure 1 shows the changes in EC ASMR and ASDR by specific risk factors for both sexes combined from 1990 to 2019. Globally, smoking and alcohol use were the leading risk factors for EC burden, which declined significantly in the past 3 decades. The ASMRs attributable to diets low in fruits and vegetables had the most significant decrease, falling from the third and fifth positions in 1990 to the fourth and sixth positions in 2019, respectively. However, EC burden related to high BMI ranked third during this period. Notably, although the ranks of ASMR and ASDR attributable to chewing tobacco increased, the burden of EC attributed to chewing tobacco declined significantly (*P*<.05).

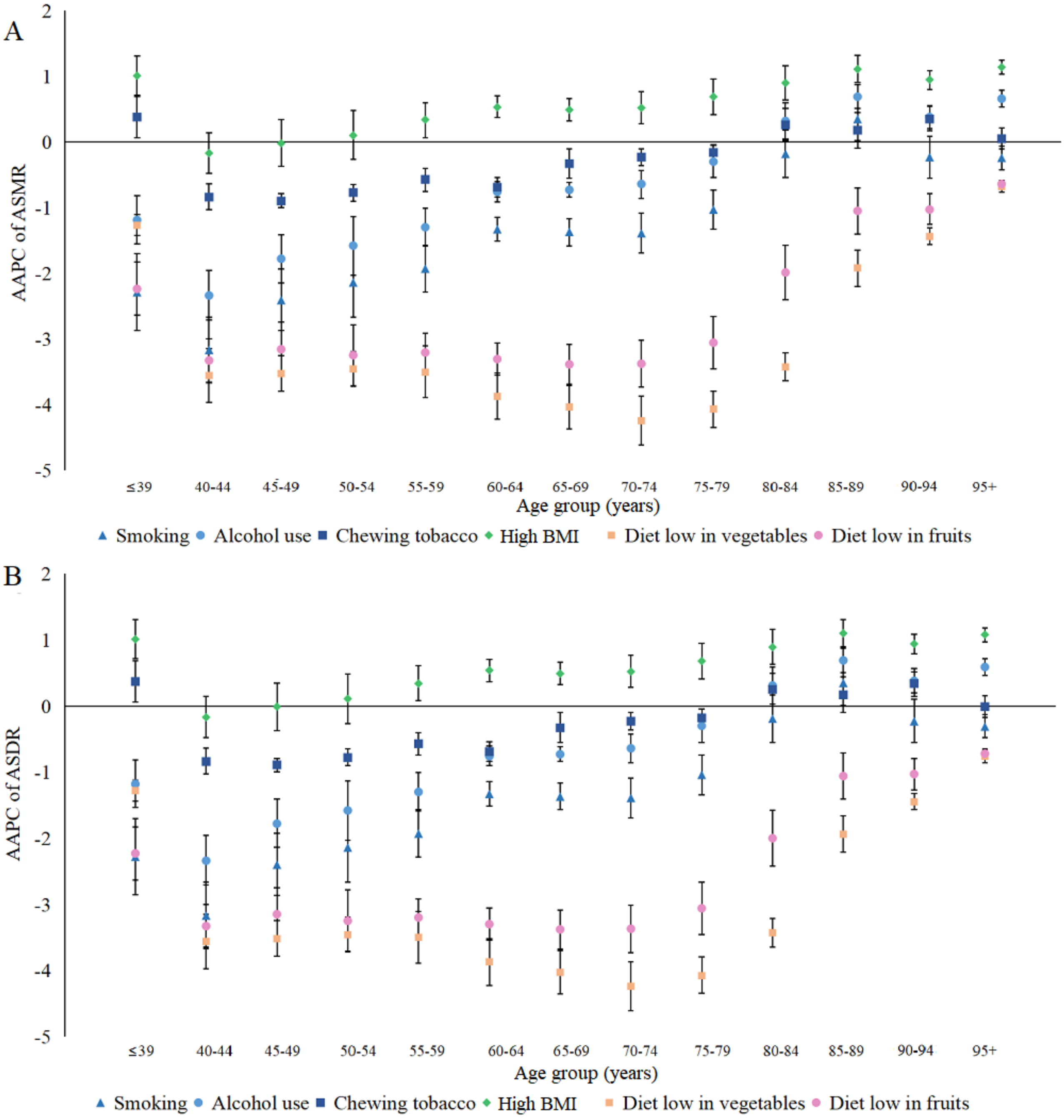
Figure 1. The changes in esophageal cancer (A) age-standardized mortality rate and (B) age-standardized disability-adjusted life year rate attributable to specific risk factors for both sexes combined in 1990-2019 globally. Dashed lines indicate decrease in rank. Solid lines indicate increase or no change in rank. Data in parentheses are 95% CIs. AAPC: average annual percentage change; ASDR: age-standardized disability-adjusted life year rate; ASMR: age-standardized mortality rate.



Global Trends in EC Risk–Attributable Burden by Age Groups

Figure 2 shows that smoking-related ASMRs and ASDRs declined in each age group except for the 85-89 years age group, with the highest decrease noted in the 40-44 years age group. Similarly, alcohol-related ASMR and ASDR declined in the 40-44 years age group, while it increased in the >80 years age group and peaked in the 85-89 years age group. The ASMRs and ASDRs attributable to diets low in fruits and vegetables declined in all age groups, but that attributable to high BMI increased in most age groups outside the subgroups within the age range of 40-54 years significantly (*P*<.05). The AAPCs of ASMR and ASDR for males and females are presented in Figures S1-S2 of Multimedia Appendix 1.

Figure 2. The average annual percentage change of (A) age-standardized mortality rate and (B) age-standardized disability-adjusted life year rate for both sexes combined by age groups in 1990-2019 globally. AAPC: average annual percentage change; ASDR: age-standardized disability-adjusted life year rate; ASMR: age-standardized mortality rate.

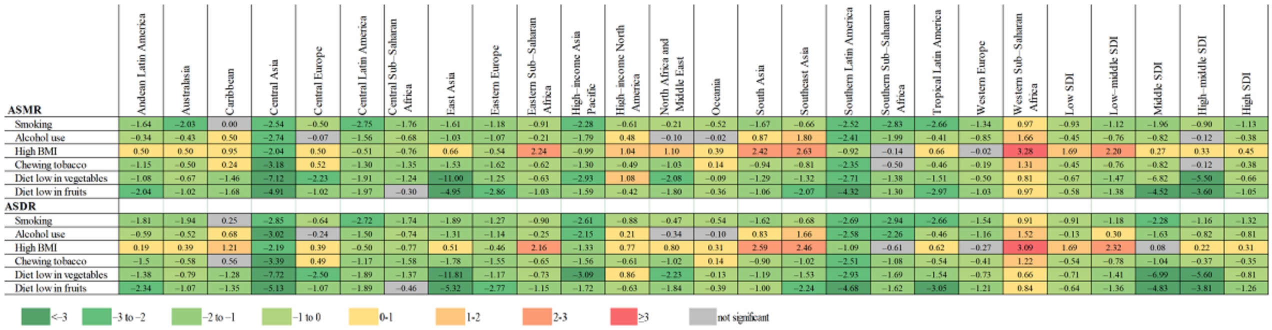


Regional and SDI Trends in EC Risk–Attributable Burden

As Figure 3 shows, there were significant downward trends in risk-attributable ASMRs and ASDRs in most regions with AAPC<0 compared to significant increases in high BMI–related ASMRs and ASDRs (*P*<.05). The EC burden attributable to each risk factor increased in Western Sub-Saharan Africa, while it declined significantly in Central Asia, Central Latin America, Eastern Europe, high-income Asia–Pacific, and Southern Latin America (*P*<.05). Among the risk factors, diets low in vegetables caused the most decline in the EC burden in East Asia and Central Asia, while high BMI caused the most increase in the ASMRs and ASDRs in Western Sub-Saharan Africa. Southern Sub-Saharan Africa and Central Asia showed the highest decline in smoking-related and alcohol use–related ASMRs and ASDRs, respectively. Figure 3 shows that the most dramatic decrease in EC burden was ascribed to low-vegetable diets in countries with middle SDI, followed by countries with high-middle SDI. In contrast, countries with low-middle SDI showed a sharp increase in EC burden, which is attributable to high BMI. In addition, Figures S1-S2 of Multimedia Appendix 1 show the changes in ASMRs and ASDRs attributable to risk factors among males and females in 21 GBD regions and 5 SDI levels. Table S1 of Multimedia Appendix 1 shows that the preponderance of EC burden for both sexes combined, in terms of the number of deaths and DALYs, were attributable to smoking, alcohol use, high BMI, and low-fruit diets in East

Asia, while the highest number of deaths and DALYs attributable to chewing tobacco and low-vegetable diets occurred in South Asia in 2019 (Multimedia Appendix 1). Table S2 in Multimedia Appendix 1 shows distinct differences and risk gradients of ASMRs and ASDRs in 21 GBD regions and 5 SDI levels.

Figure 3. The average annual percentage change for esophageal cancer risk–attributable age-standardized mortality rate and age-standardized disability-adjusted life year rate in 21 global burden of disease regions classified by 5 sociodemographic index levels for both sexes in 1990-2019. Dark green indicates the highest decline in the average annual percentage change, dark red indicates the highest increase in the average annual percentage change, and grey indicates no significant average annual percentage change. AAPC: average annual percentage change; ASDR: age-standardized disability-adjusted life year rate; ASMR: age-standardized mortality rate; SDI: sociodemographic index.



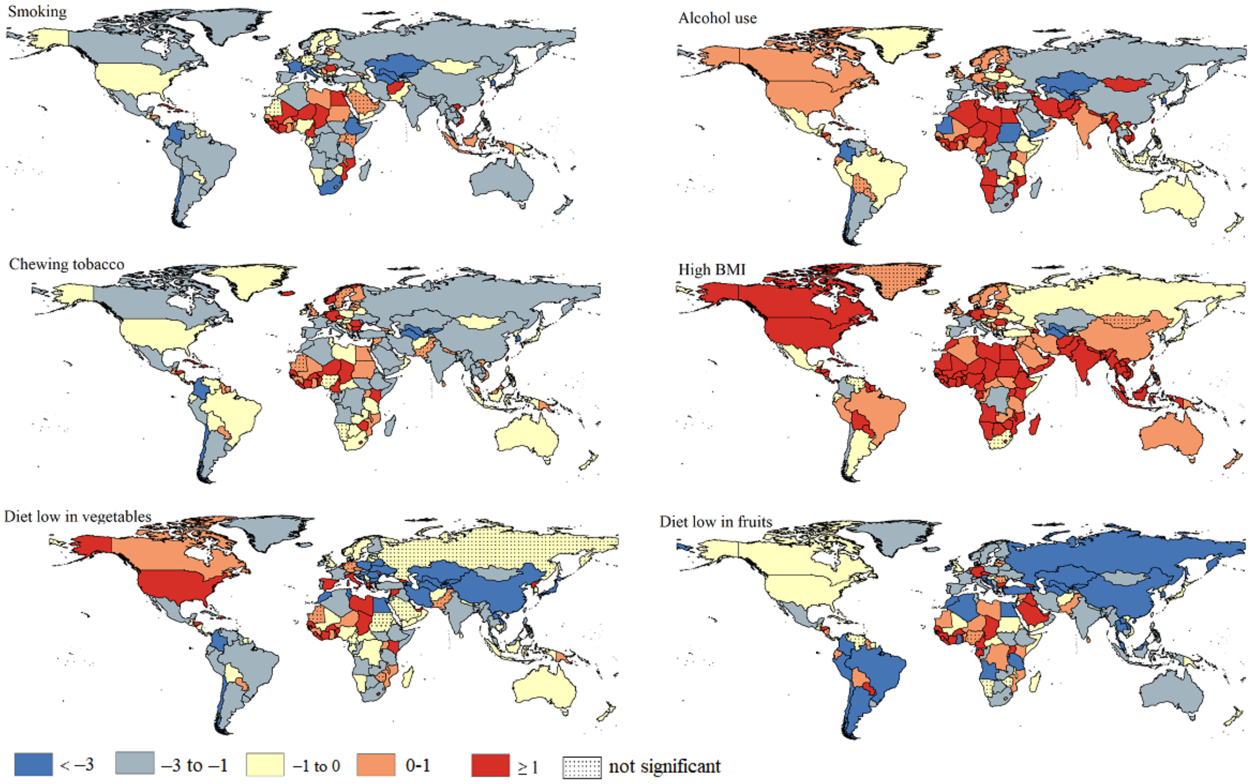
	Andean Latin America	Australasia	Caribbean	Central Asia	Central Europe	Central Latin America	Central Sub–Saharan Africa	East Asia	Eastern Europe	Eastern Sub–Saharan Africa	High–income Asia Pacific	High–income North America	North Africa and Middle East
ASMR													
Smoking	–1.64	–2.03	0.00	–2.54	–0.50	–2.75	–1.76	–1.61	–1.18	–0.91	–2.28	–0.61	–0.21
Alcohol use	–0.34	–0.43	0.50	–2.74	–0.07	–1.56	–0.68	–1.03	–1.07	–0.21	–1.79	0.48	–0.10
High BMI	0.50	0.50	0.95	–2.04	0.50	–0.51	–0.76	0.66	–0.54	2.24	–0.99	1.04	1.10
Chewing tobacco	–1.15	–0.50	0.24	–3.18	0.52	–1.30	–1.35	–1.53	–1.62	–0.62	–1.30	–0.49	–1.03
Diet low in vegetables	–1.08	–0.67	–1.46	–7.12	–2.23	–1.91	–1.24	–11.00	–1.25	–0.63	–2.93	1.08	–2.08
Diet low in fruits	–2.04	–1.02	–1.68	–4.91	–1.02	–1.97	–0.30	–4.95	–2.86	–1.03	–1.59	–0.42	–1.80
ASDR													
Smoking	–1.81	–1.94	0.25	–2.85	–0.64	–2.72	–1.74	–1.89	–1.27	–0.90	–2.61	–0.88	–0.47
Alcohol use	–0.59	–0.52	0.68	–3.02	–0.24	–1.50	–0.74	–1.31	–1.14	–0.25	–2.15	0.21	–0.34
High BMI	0.19	0.39	1.21	–2.19	0.39	–0.50	–0.77	0.51	–0.46	2.16	–1.33	0.77	0.80
Chewing tobacco	–1.5	–0.58	0.56	–3.39	0.49	–1.17	–1.58	–1.78	–1.55	–0.65	–1.56	–0.61	–1.02
Diet low in vegetables	–1.38	–0.79	–1.28	–7.72	–2.50	–1.89	–1.37	–11.81	–1.17	–0.73	–3.09	0.86	–2.23
Diet low in fruits	–2.34	–1.07	–1.35	–5.13	–1.07	–1.89	–0.46	–5.32	–2.77	–1.15	–1.72	–0.63	–1.84

	Oceania	South Asia	Southeast Asia	Southern Latin America	Southern Sub–Saharan Africa	Tropical Latin America	Western Europe	Western Sub–Saharan Africa	Low SDI	Low–middle SDI	Middle SDI	High–middle SDI	High SDI
ASMR													
Smoking	–0.52	–1.67	–0.66	–2.52	–2.83	–2.66	–1.34	0.97	–0.93	–1.12	–1.96	–0.90	–1.13
Alcohol use	–0.02	0.87	1.80	–2.41	–1.99	–0.41	–0.85	1.66	–0.45	–0.76	–0.82	–0.12	–0.38
High BMI	0.39	2.42	2.63	–0.92	–0.14	0.66	–0.02	3.28	1.69	2.20	0.27	0.33	0.45
Chewing tobacco	0.14	–0.94	–0.81	–2.35	–0.50	–0.46	–0.19	1.31	–0.45	–0.76	–0.82	–0.12	–0.38
Diet low in vegetables	–0.09	–1.29	–1.32	–2.71	–1.38	–1.51	–0.50	0.81	–0.67	–1.47	–6.82	–5.50	–0.66
Diet low in fruits	–0.36	–1.06	–2.07	–4.32	–1.30	–2.97	–1.03	0.97	–0.58	–1.38	–4.52	–3.60	–1.05
ASDR													
Smoking	–0.54	–1.62	–0.68	–2.69	–2.94	–2.66	–1.54	0.91	–0.91	–1.18	–2.28	–1.16	–1.32
Alcohol use	–0.10	0.83	1.66	–2.58	–2.26	–0.46	–1.16	1.52	–0.13	0.30	–1.63	–0.82	–0.81
High BMI	0.31	2.59	2.46	–1.09	–0.61	0.62	–0.27	3.09	1.69	2.32	0.08	0.22	0.31
Chewing tobacco	0.14	–0.90	–1.02	–2.51	–1.08	–0.54	–0.41	1.22	–0.54	–0.78	–1.04	–0.37	–0.35
Diet low in vegetables	–0.13	–1.19	–1.53	–2.93	–1.69	–1.54	–0.73	0.66	–0.71	–1.41	–6.99	–5.60	–0.81
Diet low in fruits	–0.39	–1.00	–2.24	–4.68	–1.62	–3.05	–1.21	0.84	–0.64	–1.36	–4.83	–3.81	–1.26

National Trends in EC Risk–Attributable Burden

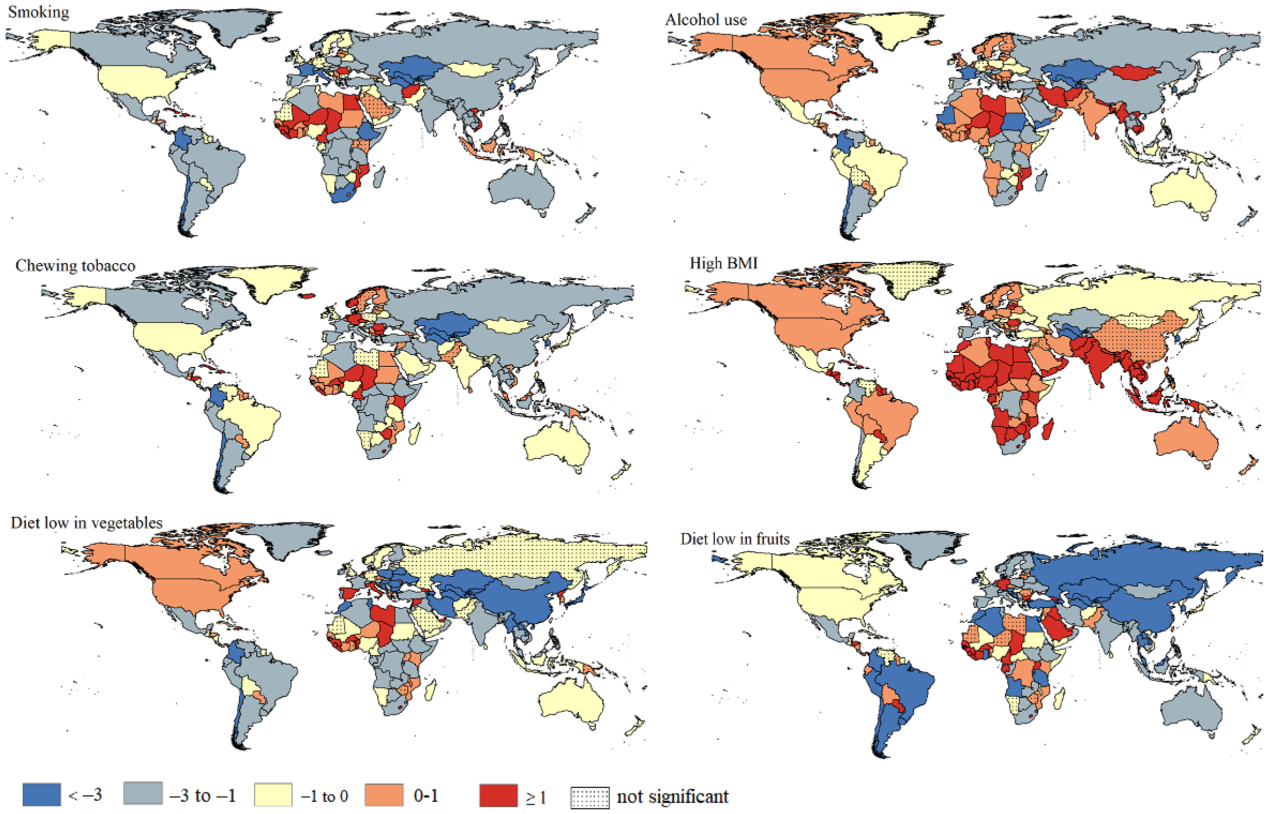
Figure 4 shows the trends in risk-attributable ASMR in both sexes combined in 204 countries and territories from 1990 to 2019. There was a significant downward trend in ASMRs attributable to smoking, alcohol use, chewing tobacco, low-vegetable diets, and low-fruit diets in 136 (66.7%), 93 (45.6%), 118 (57.8%), 141 (69.1%), and 137 (67.2%) countries of all the 204 countries and territories, respectively. In contrast, a significantly upward trend in ASMR attributable to high BMI was observed in 139 (68.1%) countries and territories. Turkmenistan showed the highest decline in ASMRs attributable to smoking (AAPC=–5.21, 95% CI –5.99 to –4.42), high BMI (AAPC=–3.21, 95% CI –4.03 to –2.39), chewing tobacco (AAPC=–4.77, 95% CI –5.51 to –4.03), and low-vegetable diets (AAPC=–13.67, 95% CI –14.69 to –12.64), while the highest decline in ASMR attributable to alcohol use (AAPC=–8.02, 95% CI –10.94 to –5.00) and low-fruit diets (AAPC=–7.99, 95% CI –8.63 to –7.34) occurred in Sudan and Albania. São Tomé, Príncipe, and Northern Mariana had the highest increase in ASMR attributable to smoking (AAPC=3.40, 95% CI 3.17-3.62) and chewing tobacco (AAPC=3.58, 95% CI 3.08-4.08). Besides, Vietnam had the highest increase in ASMR attributable to alcohol use (AAPC=8.61, 95% CI 7.77-9.45) and high BMI (AAPC=5.94, 95% CI 5.55-6.34), but United Arab Emirates had the highest increase in ASMR attributable to low-vegetable diets (AAPC=11.01, 95% CI 9.47-12.56) and low-fruit diets (AAPC=5.10, 95% CI 4.50-5.72).

Figure 4. The average annual percentage change of esophageal cancer age-standardized mortality rate attributable to specific risk factors for both sexes combined in 204 countries and territories from 1990 to 2019.



As Figure 5 shows, from 1990 to 2019, significant decreases in ASDR attributable to smoking, alcohol use, chewing tobacco, low-vegetable diets, and low-fruit diets were found in 143 (70.1%), 96 (47.1%), 116 (56.9%), 145 (71.1%), and 143 (70.1%) countries of the 204 countries and territories, while 124 (60.7%) countries and territories experienced significant increases in ASDR attributable to high BMI. During this period, Turkmenistan had the highest decrease in ASDR attributable to smoking (AAPC=–5.31, 95% CI –6.09 to –4.53) and low-vegetable diets (AAPC=–14.49, 95% CI –15.66 to –13.31), while Uzbekistan had the highest decrease in ASDR attributable to chewing tobacco (AAPC=–4.97, 95% CI –5.48 to –4.26) and high BMI (AAPC=–3.37, 95% CI –3.89 to –2.84). The changes in the risk-attributable burden in other countries and territories are presented in Figure 5.

Figure 5. The average annual percentage change of esophageal cancer age-standardized disability-adjusted life year rate attributable to specific risk factors for both sexes combined in 204 countries and territories from 1990 to 2019.



As Table S3 of Multimedia Appendix 1 shows, globally in 2019, China was the worst affected country with the highest number of EC deaths and DALYs attributable to smoking, alcohol use, high BMI, and low-fruit diets due to its large population and aging, accounting for about 50% of the total number of EC deaths and DALYs. China was followed by India, which had the highest number of EC-related deaths and DALYs attributable to chewing tobacco and low-vegetable diets.

Discussion

Principal Findings

This study uses the latest GBD 2019 data to analyze the spatiotemporal patterns in the burden of EC attributable to specific risk factors from 1990 to 2019, and our results show that the burden of EC varied with different risk factors at regional and national levels, among which EC disparities had different gradients of risk factors during this period. EC risk–attributable burden declined in most regions, countries, and territories, except for the increasing burden of EC attributed to high BMI. Our findings may be useful for the prognosis and treatment of EC by identifying the burden of modifiable risk factors worldwide for making clinical decisions about the risk attribution, stratification, and prevention.

Although the past 3 decades have witnessed a significant decrease in smoking-related EC burden, smoking still dominates among the risk-attributable factors. Moreover, the risk rank of EC burden attributable to chewing tobacco increased by 1 gradient from 1990 to 2019. The DALY of EC is estimated to be driven by tobacco smoking and chewing tobacco worldwide [10]. Compared with those in 1990, EC deaths and DALYs reduced by 55.9% and 42.1% in 2019, respectively, along with the corresponding ASMRs and ASDRs that annually declined by –1.18% and –1.41%, respectively, of which smoking resulted in 40.6% deaths [21]. Between 1990 and 2019, the trends in EC burden attributable to smoking largely varied across geographical locations, and there was a significant decrease mostly in regions outside the Caribbean and Western Sub-Saharan Africa. In 2003, the World Health Organization issued the “Framework Convention on Tobacco Control,” in which 168 countries joined this action and further worked out their own tobacco control strategies. EC deaths attributable to smoking continue to decrease with the reduction of smoking prevalence in high-income countries, paralleling with a decrease of 32.2% for males and 28.8% for females in age-standardized prevalence of smoking since 1990 [22,23]. Therefore, reducing smoking prevalence is an effective way to mitigate the EC burden.

Alcohol consumption still remains the second leading risk factor for EC burden from 1990 to 2019. The clear differences in EC burden over time attributable to alcohol use in GBD regions, as shown in our study, provides the most up-to-date information on the trends in EC risk–attributable burden. Approximately 22.6% of the deaths and 33.8% DALYs in EC are attributed to alcohol use worldwide [10,22]. Although alcohol-related EC deaths and DALYs increased by 13.1% and 8.3%, the change in the percentage of alcohol-related EC deaths and DALYs reduced by 14.4% and 17.1% from 1990 to 2017 [10,24]. The EC burden attributable to alcohol use declined in most regions (eg, Eastern Europe) concomitantly with the decrease in alcohol per capita consumption (from 11.2 L per capita to 9.8 L per capita) [25,26]. In contrast, a significant increase in EC burden related to alcohol use was observed in Southeast Asia, South Asia, Western Sub-Saharan Africa, and North America, largely because alcohol per capita consumption was high from 1990 to 2017 and is expected to remain high till 2030 [27-29]. For example, alcohol per capita consumption increased by 34% in Southeast Asia, ranging from 3.5 L per capita in 2010 to 4.7 L per capita in 2017 [27]. Thus, alcohol-related EC burden largely varies across geographical locations, and this variation should be carefully analyzed in the future.

The EC burden attributable to high BMI, in comparison with other risk factors, increased in most regions. This is consistent with the rising incidence of high BMI reported across the world, which has increased from 1975 to 2016 by nearly 12% for males, 24% for females, and 40% for both sexes combined [30]. Geographical differences in EC burden related to high BMI followed a geographical pattern of increasing obesity prevalence, with the largest absolute increase in high-income western countries, Central Asia, Middle East, and North Africa, and with the most relative increase in East and South-Eastern Asia [31]. However, previous studies have indicated that high BMI (or obesity) was a driver for EAC and not for ESCC [23]. This study also found that the geographical pattern of EC burden related to high BMI with AAPC>3 was generally consistent with that of EAC [32,33]. The slightly increased burden of EC attributable to high BMI, especially in typical ESCC high-risk regions such as China, suggest that EAC may show an upward trend due to increased prevalence of high BMI.

Consistent with that reported in previous studies [21,34], the burden of EC attributable to low-fruit and low-vegetable diets declined significantly, even if it remained stable or increased slightly in some regions. This aligns with the globally low consumption of fruits by 16.58% and vegetables by 25.60% from 1990 to 2017 [34,35]. The high dietary fibers of vegetables and fruits show a protective effect by generating inositol hexaphosphate to curb the growth rate of EC cells by reducing cellular proliferation and stimulating apoptosis [36]. Another plausible explanation is that high dietary fibers in vegetables and fruits decrease the levels of systemic inflammation factors such as tumor necrosis factor-α receptor-2 and interleukin-6, which may cause carcinogenesis [37]. The largest decrease in EC burden attributable to low intake of fruits and vegetables occurred in East Asia because of higher incomes and better health care. However, an upward trend in EC burden associated with low intake of vegetables occurred in North America and that associated with low intake of fruits occurred in Western Sub-Saharan Africa. In light of these variations, policy makers in each country should tailor related measures to promote well-balanced diets according to the local risk factors.

Risk Exposure of EC burden Among Sexes and Age Groups

In terms of the spatiotemporal pattern in the risk exposure of EC burden, important differences in the EC burden between sexes and age groups in different locations should be mentioned. The time trend in EC burden attributed to smoking and chewing tobacco among sexes and age groups followed a similar pattern with a significant decrease, but the gradient showed a distinct difference among locations. This is possibly caused by the reduction of smoking prevalence in different regions, wherein males in 135 countries and females in 68 countries showed a prominent decrease in smoking prevalence [22]. Similarly, the greater decreases in EC burden due to alcohol use among females and in the 40-44 years age group indicate that interventions should target males and >45-year-old populations. Besides, the time trends of the EC burden attributable to high BMI for males and females did not follow the same pattern—with an increasing AAPC for males and a decreasing AAPC for females within the 40-64 years old subgroups. A previous study documented that high BMI resulted in ASDR increase of 12.7% for males and 26.8% for females, accounting for nearly 20% of the EC burden [38]. In the past decades, the prevalence of obesity increased by quadruple in men and more than double in women, and the time trend of ASMR attributable to high BMI showed a significant upward trend in 170 countries and territories, a downward trend only in 8 countries and territories, and a stable trend in 26 countries and territories [31,39]. The reason for this phenomenon is not completely understood, but the global prevalence of high BMI and obesity as well as deaths and DALY-related overweight are observed to be parallel with our findings across age groups, implying that sex differences in the prevalence of high BMI and obesity in the age groups are a potential risk factor [38,40]. Accordingly, the sex differences in EC burden related to high BMI should be addressed by ad hoc strategies, which are imperative for implementing regulations and policies targeting specific disorders.

Public Health Implications

Our findings provide important evidence for tailored, country-specific policy development and inventions to address the geographical disparities in the EC burden, which can be used by governments at the national level to meet the challenges of the substantial burden of EC and develop strategies based on the priority of specific risk-attributable EC burden. Strategies such as behavioral and metabolic modifications (including smoking control, alcohol control policies and regulations, weight management tips) and access to abundant fruits and vegetables are imperative to face the challenges of EC. In light of the increase in EC burden related to high BMI in most countries, we recommend that public health policy makers and decision makers should prioritize their agenda by designing and implementing ad hoc effective policies to restrain the upward momentum of EAC. Moreover, encouraging balanced diets and reducing alcohol consumption and tobacco should constitute the essential component of EC prevention strategies. Educating the general public regarding risk factor modification is urgently needed in countries with heavy EC burden. Although it is not clear how diet influences EC development and progression, our study shows that EC burden attributable to low intake of fruits and vegetables has decreased significantly.

Strengths and Limitations

This study provides a comprehensive spatiotemporal analysis of the EC burden attributable to specific risk factors, thereby becoming a potential reference for developing interventions that address similar health problems by identifying the most prominent risk factor and monitoring the effectiveness of the change of risk-attributable burden of EC over time. Further studies are required to conduct early interventions targeting behavioral and metabolic risk factors of individuals based on our findings. Moreover, the accelerating growth of EC burden related to high BMI alerts us to prevent and curb this issue as early as possible. We hope these limitations can be solved in the future. The 2 distinct histological subtypes of EC, that is, ESCC and EAC, have completely different risk factors, trends, and geographical patterns, but data on ESCC and EAC are unavailable in GBD. Thus, we did not analyze the changes of EC burden attributable to risk factors by subtype. Second, some countries had no EC data because of the lack of registries, or even if they had data in several registries that did not cover the whole country and territory, the estimates were obtained by modelling. The EC data in countries without cancer registries were modeled, but the model did not solve the geographical approximates of EC, because the distribution of EC showed clear clusters. Third, we did not analyze the spatiotemporal changes of EC burden attributable to other risk factors such as consumption of hot meals, eating red or processed meat, physical activity, and infectious agents, because GBD has no data on these variables. This may overestimate the burden of EC attributable to the included risk factors in this study. Fourth, this study did not capture the potential synergy effects between risk factors in which some combinations might be multiplicative. For some risk factors such as behavior and diet, the joint effect of smoking and alcohol use on the burden of EC is very important for framing public policy. Further, more detailed work is needed to strengthen the evidence base for understanding the mediation effects of sex, age, and geographical locations on the risk-attributable burden of EC.

Conclusion

This study analyzes the spatiotemporal patterns in the risk-attributable burden of EC over the past 3 decades, suggesting that smoking and alcohol use remained the dominant risk factors and drive the burden of EC compared with other risk-outcome pairs. The EC burden caused by smoking, alcohol use, chewing tobacco, low-fruit diets, and low-vegetable diets has decreased over the last 3 decades, but the EC burden attributable to high BMI has increased significantly. In addition, changes in the burden of EC due to modifiable risk factors varied widely across regions, with the largest decrease in EC burden in East Asia, especially in China. Although the exact reasons for the marked geographical differences are still unclear, a large number of epidemiological studies are ongoing to yield valuable findings. Our findings targeted specific risk factors of EC burden at different geographical scales to elucidate the key modifiable risk factors for the prevention and control of EC in line with achieving the target 3.4 of the Sustainable Development Goals. Further, understanding the epidemiological trend and risk factor stratification of EC are the key pathways for public health and clinical decisions regarding risk stratification, screening, and prevention. These results might be conducive for identifying the leading modifiable risk factors for countries that might not have previous local research on EC burden and risk factor exposure.

Acknowledgments

We greatly appreciate the works of the Global Burden of Disease Study 2019 collaborators and thank all the researchers and collaborators who were involved in this study for their efforts and contribution. This work was funded by the Shanxi Provincial Administration of Traditional Chinese Medicine, Esophageal Cancer Comprehensive Treatment and Accelerated Recovery Innovation Team (grant 2022-SLRH-LJ-009), and the National Youth Science Foundation (grant 82102976).

Data Availability

Publicly available data sets were analyzed in this study. The processed data sets and statistical methods in the analysis are available from the corresponding author upon reasonable request.

Authors' Contributions

PL and GZ designed the study. PL drafted and revised the manuscript. JJ and WL collected and cleaned the data. JW checked the accuracy of data and prepared the tables. XQ performed statistical analysis. All authors reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary data.
[PDF File (Adobe PDF File), 3206 KB - publichealth_v9i1e46051_app1.pdf]

References

1. Li Y, Xu J, Gu Y, Sun X, Dong H, Chen C. The disease and economic burdens of esophageal cancer in China from 2013 to 2030: dynamic cohort modeling study. JMIR Public Health Surveill 2022 Mar 02;8(3):e33191-e33202 [FREE Full text] [doi: 10.2196/33191] [Medline: 34963658]

2. An L, Zheng R, Zeng H, Zhang S, Chen R, Wang S, et al. The survival of esophageal cancer by subtype in China with comparison to the United States. Int J Cancer 2023 Jan 15;152(2):151-161. [doi: 10.1002/ijc.34232] [Medline: 35913753]
3. Grille VJ, Campbell S, Gibbs JF, Bauer TL. Esophageal cancer: the rise of adenocarcinoma over squamous cell carcinoma in the Asian belt. J Gastrointest Oncol 2021 Jul;12(Suppl 2):S339-S349 [FREE Full text] [doi: 10.21037/jgo-2019-gi-08] [Medline: 34422398]
4. GBD 2019 Colorectal Cancer Collaborators. Global, regional, and national burden of colorectal cancer and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Gastroenterol Hepatol 2022 Jul;7(7):627-647 [FREE Full text] [doi: 10.1016/S2468-1253(22)00044-9] [Medline: 35397795]
5. Ho ALK, Smyth EC. A global perspective on oesophageal cancer: two diseases in one. The Lancet Gastroenterology & Hepatology 2020 Jun;5(6):521-522. [doi: 10.1016/s2468-1253(20)30047-9] [Medline: 32246940]
6. Han Y, Zhu X, Hu Y, Yu C, Guo Y, Hang D, et al. Electronic health record-based absolute risk prediction model for esophageal cancer in the Chinese population: model development and external validation. JMIR Public Health Surveill 2023 Mar 15;9:e43725 [FREE Full text] [doi: 10.2196/43725] [Medline: 36781293]
7. Malhotra GK, Yanala U, Ravipati A, Follet M, Vijayakumar M, Are C. Global trends in esophageal cancer. J Surg Oncol 2017 Apr;115(5):564-579. [doi: 10.1002/jso.24592] [Medline: 28320055]
8. Wang S, Zheng R, Arnold M, Abnet C, Zeng H, Zhang S, et al. Global and national trends in the age-specific sex ratio of esophageal cancer and gastric cancer by subtype. Int J Cancer 2022 Nov 01;151(9):1447-1461 [FREE Full text] [doi: 10.1002/ijc.34158] [Medline: 35678331]
9. Huang J, Koulaouzidis A, Marlicz W, Lok V, Chu C, Ngai CH, et al. Global burden, risk factors, and trends of esophageal cancer: an analysis of cancer registries from 48 countries. Cancers (Basel) 2021 Jan 05;13(1):141-156 [FREE Full text] [doi: 10.3390/cancers13010141] [Medline: 33466239]
10. GBD 2017 Oesophageal Cancer Collaborators. The global, regional, and national burden of oesophageal cancer and its attributable risk factors in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol 2020 Jun;5(6):582-597 [FREE Full text] [doi: 10.1016/S2468-1253(20)30007-8] [Medline: 32246941]
11. Li S, Chen H, Man J, Zhang T, Yin X, He Q, et al. Changing trends in the disease burden of esophageal cancer in China from 1990 to 2017 and its predicted level in 25 years. Cancer Med 2021 Mar;10(5):1889-1899 [FREE Full text] [doi: 10.1002/cam4.3775] [Medline: 33586344]
12. Soroush A, Malekzadeh R, Roshandel G, Khoshnia M, Poustchi H, Kamangar F, et al. Sex and smoking differences in the association between gastroesophageal reflux and risk of esophageal squamous cell carcinoma in a high-incidence area: Golestan Cohort Study. Int J Cancer 2023 Mar 15;152(6):1137-1149. [doi: 10.1002/ijc.34313] [Medline: 36214797]
13. Results of global burden of disease, injuries, and risk factors study. Institute for Health Metrics and Evaluation. URL: https://vizhub.healthdata.org/gbd-results [accessed 2022-10-22]
14. GBD 2019 Diseases Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
15. GBD 2019 Cancer Risk Factors Collaborators. The global burden of cancer attributable to risk factors, 2010-19: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2022 Aug 20;400(10352):563-591 [FREE Full text] [doi: 10.1016/S0140-6736(22)01438-6] [Medline: 35988567]
16. Domper Arnal MJ, Ferrández Arenas Á, Lanas Arbeloa Á. Esophageal cancer: risk factors, screening and endoscopic treatment in Western and Eastern countries. World J Gastroenterol 2015 Jul 14;21(26):7933-7943 [FREE Full text] [doi: 10.3748/wjg.v21.i26.7933] [Medline: 26185366]
17. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1923-1994 [FREE Full text] [doi: 10.1016/S0140-6736(18)32225-6] [Medline: 30496105]
18. Li P, Jing J, Li R, Ge M, Jia P, Hu W, et al. Upper gastrointestinal cancer in China: spatial epidemiologic evidence from screening areas. Cancer Prev Res (Phila) 2020 Nov;13(11):935-946. [doi: 10.1158/1940-6207.CAPR-20-0139] [Medline: 32655009]
19. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
20. Foreman KJ, Lozano R, Lopez AD, Murray CJ. Modeling causes of death: an integrated approach using CODEm. Popul Health Metr 2012 Jan 06;10:1 [FREE Full text] [doi: 10.1186/1478-7954-10-1] [Medline: 22226226]
21. Cai Y, Lin J, Wei W, Chen P, Yao K. Burden of esophageal cancer and its attributable risk factors in 204 countries and territories from 1990 to 2019. Front Public Health 2022;10:952087 [FREE Full text] [doi: 10.3389/fpubh.2022.952087] [Medline: 36148334]
22. GBD 2019 Tobacco Collaborators. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990-2019: a systematic analysis from the Global Burden of

Disease Study 2019. Lancet 2021 Jun 19;397(10292):2337-2360 [FREE Full text] [doi: 10.1016/S0140-6736(21)01169-7] [Medline: 34051883]
23. Xie W, Yang T, Zuo J, Ma Z, Yu W, Hu Z, et al. Chinese and global burdens of gastrointestinal cancers from 1990 to 2019. Front Public Health 2022;10:941284 [FREE Full text] [doi: 10.3389/fpubh.2022.941284] [Medline: 35910886]
24. Lin L, Li Z, Yan L, Liu Y, Yang H, Li H. Global, regional, and national cancer incidence and death for 29 cancer groups in 2019 and trends analysis of the global cancer burden, 1990-2019. J Hematol Oncol 2021 Nov 22;14(1):197 [FREE Full text] [doi: 10.1186/s13045-021-01213-z] [Medline: 34809683]
25. GBD 2016 Alcohol Drug Use Collaborators. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Psychiatry 2018 Dec;5(12):987-1012 [FREE Full text] [doi: 10.1016/S2215-0366(18)30337-7] [Medline: 30392731]
26. Karriker-Jaffe KJ, Room R, Giesbrecht N, Greenfield TK. Alcohol's harm to others: opportunities and challenges in a public health framework. J Stud Alcohol Drugs 2018 Mar;79(2):239-243 [FREE Full text] [doi: 10.15288/jsad.2018.79.239] [Medline: 29553351]
27. Manthey J, Shield K, Rylett M, Hasan OSM, Probst C, Rehm J. Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. The Lancet 2019 Jun;393(10190):2493-2502. [doi: 10.1016/s0140-6736(18)32744-2]
28. Freedman ND, Murray LJ, Kamangar F, Abnet CC, Cook MB, Nyrén O, et al. Alcohol intake and risk of oesophageal adenocarcinoma: a pooled analysis from the BEACON Consortium. Gut 2011 Aug;60(8):1029-1037 [FREE Full text] [doi: 10.1136/gut.2010.233866] [Medline: 21406386]
29. Sewram V, Sitas F, O'Connell D, Myers J. Tobacco and alcohol as risk factors for oesophageal cancer in a high incidence area in South Africa. Cancer Epidemiol 2016 Apr;41:113-121. [doi: 10.1016/j.canep.2016.02.001] [Medline: 26900781]
30. Sung H, Siegel RL, Torre LA, Pearson-Stuttard J, Islami F, Fedewa SA, et al. Global patterns in excess body weight and the associated cancer burden. CA Cancer J Clin 2019 Mar;69(2):88-112 [FREE Full text] [doi: 10.3322/caac.21499] [Medline: 30548482]
31. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet 2017 Dec 16;390(10113):2627-2642 [FREE Full text] [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]
32. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. Gut 2015 Mar;64(3):381-387. [doi: 10.1136/gutjnl-2014-308124] [Medline: 25320104]
33. Salem M, Puccini A, Xiu J, Raghavan D, Lenz HJ, Korn WM, et al. Comparative molecular analyses of esophageal squamous cell carcinoma, esophageal adenocarcinoma, and gastric adenocarcinoma. Oncologist 2018 Nov;23(11):1319-1327 [FREE Full text] [doi: 10.1634/theoncologist.2018-0143] [Medline: 29866946]
34. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: 10.1016/S0140-6736(19)30041-8] [Medline: 30954305]
35. Abnet CC, Corley DA, Freedman ND, Kamangar F. Diet and upper gastrointestinal malignancies. Gastroenterology 2015 May;148(6):1234-1243.e4 [FREE Full text] [doi: 10.1053/j.gastro.2015.02.007] [Medline: 25680671]
36. Qin X, Jia G, Zhou X, Yang Z. Diet and esophageal cancer risk: an umbrella review of systematic reviews and meta-analyses of observational studies. Adv Nutr 2022 Dec 22;13(6):2207-2216 [FREE Full text] [doi: 10.1093/advances/nmac087] [Medline: 36041184]
37. Zang Z, Liu Y, Wang J, Liu Y, Zhang S, Zhang Y, et al. Dietary patterns and severity of symptom with the risk of esophageal squamous cell carcinoma and its histological precursor lesions in China: a multicenter cross-sectional latent class analysis. BMC Cancer 2022 Jan 21;22(1):95 [FREE Full text] [doi: 10.1186/s12885-022-09206-y] [Medline: 35062901]
38. Dai H, Alsalhe TA, Chalghaf N, Riccò M, Bragazzi NL, Wu J. The global burden of disease attributable to high body mass index in 195 countries and territories, 1990-2017: An analysis of the Global Burden of Disease Study. PLoS Med 2020 Jul;17(7):e1003198 [FREE Full text] [doi: 10.1371/journal.pmed.1003198] [Medline: 32722671]
39. Zhi X, Kuang X, Liu K, Li J. The global burden and temporal trend of cancer attributable to high body mass index: Estimates from the Global Burden of Disease Study 2019. Front Nutr 2022;9:918330 [FREE Full text] [doi: 10.3389/fnut.2022.918330] [Medline: 35958256]
40. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug 30;384(9945):766-781 [FREE Full text] [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]

Abbreviations

AAPC: average annual percentage change
ASDR: age-standardized disability-adjusted life year rate
ASMR: age-standardized mortality rate

DALY: disability-adjusted life year
EAC: esophageal adenocarcinoma
EC: esophageal cancer
ESCC: esophageal squamous cell carcinoma
GBD: global burden of disease
SDI: sociodemographic index
UI: uncertainty interval

Edited by A Mavragani, T Sanchez; submitted 28.01.23; peer-reviewed by WF Khaw, G Jin; comments to author 02.03.23; revised version received 16.03.23; accepted 05.09.23; published 06.10.23.

Please cite as:
Li P, Jing J, Liu W, Wang J, Qi X, Zhang G
Spatiotemporal Patterns of Esophageal Cancer Burden Attributable to Behavioral, Metabolic, and Dietary Risk Factors From 1990 to 2019: Longitudinal Observational Study
JMIR Public Health Surveill 2023;9:e46051
URL: https://publichealth.jmir.org/2023/1/e46051
doi:10.2196/46051
PMID:37801354

Original Paper

The Effect of Sleep on Metabolism, Musculoskeletal Disease, and Mortality in the General US Population: Analysis of Results From the National Health and Nutrition Examination Survey

Ting Lei[1,2,3*], MD, PhD; Mingqing Li[1,2,3*], MD; Hu Qian[4], MD; Junxiao Yang[1,2,3], MD; Yihe Hu[1,2,3,5,6], MD; Long Hua[7,8], MD

[1]Department of Orthopedic Surgery, Xiangya Hospital, Central South University, Changsha, China
[2]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China
[3]Hunan Engineering Research Center of Biomedical Metal and Ceramic Implants, Xiangya Hospital, Central South University, Changsha, China
[4]Department of Orthopaedic Surgery, Affiliated Hospital of Zunyi Medical University, Zunyi, China
[5]Department of Orthopedic Surgery, The First Affiliated Hospital, Zhejiang University, Hangzhou, China
[6]College of Medicine, Zhejiang University, Hangzhou, China
[7]Department of Orthopedic Surgery, The First Affiliated Hospital, Xinjiang Medical University, Urumqi, China
[8]Key Laboratory of High Incidence Disease Research in Xinjiang, Ministry of Education, Xinjiang Medical University, Urumqi, China
[*]these authors contributed equally

Corresponding Author:
Long Hua, MD
Department of Orthopedic Surgery
The First Affiliated Hospital
Xinjiang Medical University
Xinyi Road
Urumqi, 830000
China
Phone: 86 15084715437
Email: hualong_xmu@163.com

Abstract

Background: Sleep is an important physiological behavior in humans that is associated with the occurrence and development of various diseases. However, the association of sleep duration with health-related outcomes, including obesity-related factors, musculoskeletal diseases, and mortality because of different causes, has not been systematically reported.

Objective: This study aims to systematically investigate the effect of sleep duration on health-related outcomes.

Methods: Overall, 54,664 participants with sleep information from 8 survey cycles of the National Health and Nutrition Examination Survey (2005-2020) were included in the analysis. Health-related outcomes comprised obesity-related outcomes (ie, BMI, obesity, waist circumference, and abdominal obesity), metabolism-related outcomes (ie, uric acid, hyperuricemia, and bone mineral density [BMD]), musculoskeletal diseases (ie, osteoarthritis [OA] and rheumatoid arthritis [RA]), and mortality because of different causes. The baseline information of participants including age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes was also collected as covariates. Information about the metabolism index, disease status, and covariates was acquired from the laboratory, examination, and questionnaire data. Survival information, including survival status, duration, and cause of death, was obtained from the National Death Index records. Quantile regression models and Cox regression models were used for association analysis between sleep duration and health-related outcomes. In addition, the threshold effect analysis, along with smooth curve fitting method, was applied for the nonlinear association analysis.

Results: Participants were divided into 4 groups with different sleep durations. The 4 groups showed significant differences in terms of baseline data (*P*<.001). The quantile regression analysis indicated that participants with increased sleep duration showed decreased BMI (β=−.176, 95% CI −.220 to −.133; *P*<.001), obesity (odds ratio [OR] 0.964, 95% CI 0.950-0.977; *P*<.001), waist circumference (β=−.219, 95% CI −.320 to −.117; *P*<.001), abdominal obesity (OR 0.975, 95% CI 0.960-0.990; *P*<.001), OA (OR 0.965, 95% CI 0.942-0.990; *P*=.005), and RA (OR 0.940, 95% CI 0.912-0.968; *P*<.001). Participants with increased sleep duration also showed increased BMD (β=.002, 95% CI .001-.003; *P*=.005), as compared with participants who slept <5.5 hours. A significant saturation effect of sleep duration on obesity, abdominal obesity, and hyperuricemia was detected through smooth curve fitting and threshold effect analysis (sleep duration>inflection point). In addition, a significant threshold effect of sleep duration on BMD (*P*<.001); OA (*P*<.001); RA (*P*<.001); and all-cause (*P*<.001), cardiovascular disease−cause (*P*<.001), cancer-cause (*P*=.005), and diabetes-cause mortality (*P*<.001) was found. The inflection point was between 6.5 hours and 9 hours.

Conclusions: The double-edged sword effect of sleep duration on obesity-related outcomes, embolism-related diseases, musculoskeletal diseases, and mortality because of different causes was detected in this study. These findings provided epidemiological evidence that proper sleep duration may be an important factor in the prevention of multisystem diseases.

(JMIR Public Health Surveill 2023;9:e46385) doi:10.2196/46385

KEYWORDS

sleep duration; mortality; clinical outcomes; threshold effect; National Health and Nutrition Examination Survey

Introduction

Background

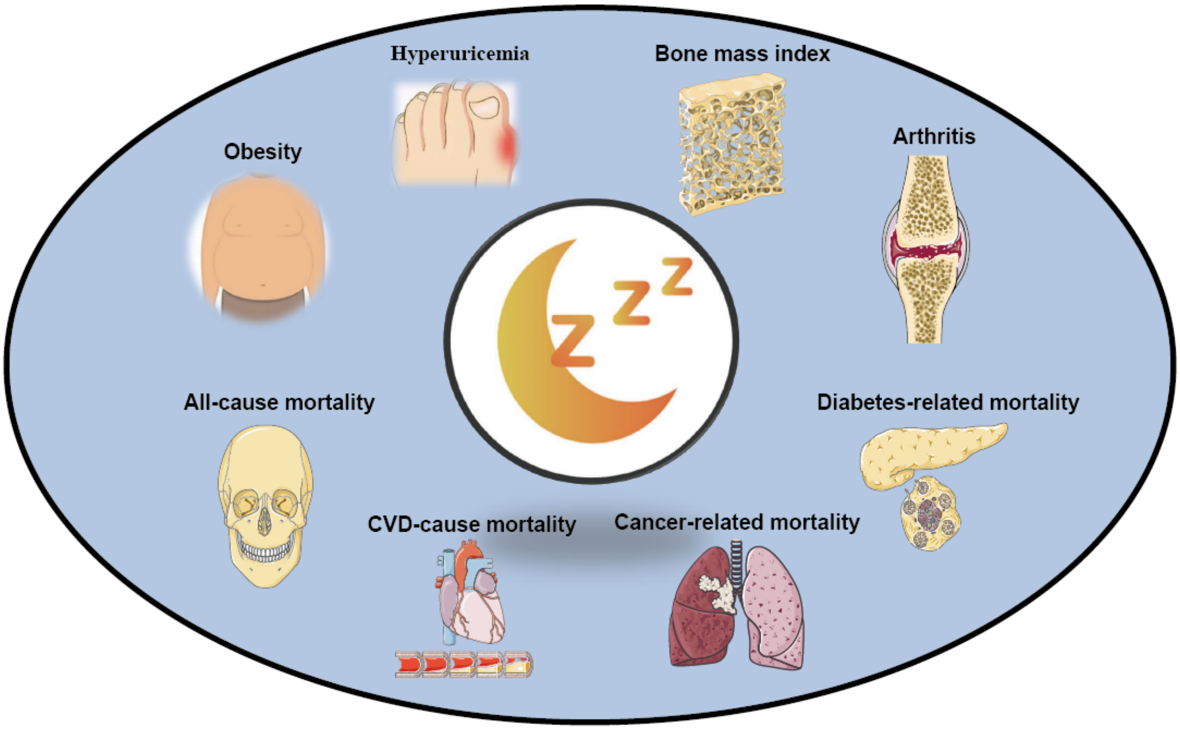
A third of our lives is spent sleeping, a crucial activity for maintaining our physiological functions. It is increasingly believed to influence the development of many diseases such as diabetes and death. Some studies have found that sleep deprivation can reduce energy expenditure by slowing metabolism and may also increase the risk of cardiovascular disease (CVD) [1]. Many meta-analyses have shown that short and long sleep durations are associated with all-cause mortality [2-4]. Moreover, sleep may improve musculoskeletal disease at the psychological level [5-8]. It has also been reported that sleep duration is associated with obesity and diabetes [9-11]. However, the problems that cannot be ignored are that there are opposite results reported, differences between different types of studies, diverse target populations and reference groups, insufficient sample size, and other factors; therefore, the benefits of sleep duration on human disease and mortality are still in an uncertain stage.

In addition, different ages, sex, race, education, marital status, body size, lifestyle habits, and the presence of chronic diseases all influenced the outcome variables [12,13]. They are similar to spiderwebs, which have complex connections and restrict or promote each other. At present, few studies include multiple influential factors to establish a prediction model of multiple-disease and multiple-factor mortality and a large sample size to evaluate the correlation between sleep duration and obesity, metabolism-related factors, musculoskeletal disease, and different types of mortalities.

Objectives

The US National Health and Nutrition Examination Survey (NHANES) database collected sleep data from participants in the survey program between 2005 and 2020. We used these data to evaluate the complex relationship between sleep duration and obesity metabolism, musculoskeletal disease, and mortality and to elaborate on their essential roles systematically (Figure 1).

Figure 1. The schematic graph of the relationship between sleep duration and metabolism, musculoskeletal disease, and mortality in the general US population. CVD: cardiovascular diseases.



Methods

Data Source and Population

All data used in this study were acquired from 8 survey cycles (2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, 2015-2016, 2017-2018, and 2019-2020) of the NHANES program. The NHANES database is designed to survey the nutritional health of the US population every 2 years. It uses a stratified sampling strategy that includes certain groups, such as people aged >60 years and other ethnic groups in the United States. The project includes demographic, inspection, laboratory, and questionnaire data. All data used in this study are available on the NHANES website [14].

Ethical Considerations

The investigation protocol for this national health program in the United States was reviewed and approved by the National Center for Health Statistics, as one of the departments of the Centers for Disease Control and Prevention. All participants were given an informed consent form outlining the program details, and they would sign it before being included [15]. Furthermore, this study was reviewed and approved by the ethics committee of Xiangya Hospital (202309188).

Measure of Sleep Duration

In the NHANES program, 8 survey cycles (2005-2020) monitored sleep disorders in people aged >16 years. A detailed measurement method for testing sleep disorders can be found on the website [14]. Sleep information survey collected the enrolled population's sleep duration and wake duration, including sleep and wake duration on weekdays and weekends. Statistical sleep time was rounded to the nearest half hour. The questions were asked at home by trained interviewers using a computer-assisted personal interview system to ensure data integrity, consistency, and analysis usefulness.

Obesity Evaluation

BMI assesses obesity, which is obtained as weight/square of height (kg/m^2), and participants' obesity status was then determined by BMI, with a BMI of ≥30 being considered obese. For the assessment of abdominal obesity, each participant's waist circumference (WC; in cm) was obtained for each survey period. Abdominal obesity was defined as WC of >102 cm for men and >88 cm for women. Bone mineral density (BMD) was measured using the dual-energy x-ray absorptiometry method, which has the advantages of high speed, ease of use, and low radiation [16]. Blood uric acid (UA) detection is based on UA oxidation by uricase, under the action of 4-aminophenazone, to produce color products, which the instrument can detect.

Assessment of Musculoskeletal Diseases

We evaluated musculoskeletal diseases including hyperuricemia (HUA), osteoarthritis (OA), and rheumatoid arthritis (RA). The diagnosis of HUA is determined using the enrolled patient's serum UA concentration. HUA is diagnosed with serum UA >7mg/mL in men or 6 mg/mL in women. The OA or RA status of participants was obtained through a questionnaire survey. The questions included the following: Have you ever been diagnosed with arthritis? Once the answer was confirmed, they were asked what type of arthritis was diagnosed. Participants were divided into OA, RA, and other groups based on their answers to these 2 questions. A previous study showed an 81% agreement between self-reported OA and clinically diagnosed OA [17].

Determination of Mortality Outcomes

Mortality information for NHANES program participants is available for download on the National Death Index website, the database used by the National Center for Health Statistics to register death certificate records. The mortality data used in this study were followed up to 2019. National Death Index mortality data included the number of NHANES participants, mortality status, causes of death, and survival time since enlisting in the NHANES program. The cause of death was determined according to the International Statistical Classification of Diseases, 10th edition. The National Institute of Health divided the causes of death into heart disease (054-068) and cancer (019-043). In addition, whether the participants died because of high blood pressure and diabetes was included in the data. In this study, all mortality rates were identified as all-cause mortality, and those related to malignancy were identified as cancer-cause mortality. In addition, deaths related to heart disease or hypertension were identified as CVD-cause mortality. Finally, diabetes-related deaths were identified as diabetes-cause mortality [18].

Assessment of Covariates

Covariate data were collected and included demographic data such as age, sex, race, education, and marital status. The race consisted of Mexican American, non-Hispanic Black, non-Hispanic White, other Hispanic, and other race. Education level included <9th grade, 9th to 11th grade, high school graduate, some college or Associate of Arts degree, college graduate or above, and others. Marital status included married, widowed, divorced, separated, never married, living with a partner, and others. Data also included BMI, total energy intake, physical activity, alcohol consumption, smoking, high blood pressure, and diabetes. There are subdivisions for each type of data. For example, hypertension is defined as a systolic blood pressure >140 mm Hg and a diastolic blood pressure >90 mm Hg. Diabetes is diagnosed by fasting blood glucose >7.0 mmol/L or the use of hypoglycemic drugs.

Statistical Analysis

This study aimed to investigate the association between sleep duration and various health-related outcomes using cross-sectional and follow-up data. The analysis between sleep duration and mortality information owing to different causes belongs to the prospective part, as the mortality information was collected after some period of follow-up, whereas the analysis between sleep duration and the other outcomes belongs to the cross-sectional part, as the information of these outcomes was collected at the same time with that of sleep duration and the covariates information. For the prospective part, the Cox quantile regression risk model was used for association analysis between sleep duration and mortality, and the hazard ratio (HR) was used for assessing the effect of sleep duration on mortality-related outcomes. For the cross-sectional part, the quantile regression model was used to analyze the correlation between sleep duration and other outcomes, and the odds ratio (OR) was used for assessing the effect of sleep duration on other health-related outcomes. For continuous variables, β values were used to measure the effect of risk factors on outcome measures. An OR or HR value >1 indicates that the factor is a risk factor. An OR or HR value <1 indicates that the element is protective. A β value of >0 indicates that the factor positively correlates with the outcome measure and vice versa. First, all the participants were equally grouped into 4 quantiles from short to long sleep duration, and the incidence or effective size of clinical outcomes in the first quantile (Q1) was set as the reference. Second, 3 regression models (Cox quantile or quantile regression model) with different covariables were constructed. A nonadjusted model without covariates is first used for the analysis. Model 1 was adjusted for age, sex, race, education, and marital status, and model 2 was adjusted for age, sex, race, education, marital status, BMI, total energy intake, physical activity, smoking, alcohol consumption, hypertension, and diabetes. In addition, baseline data between the 4 groups of participants were compared using the chi-square test (classified data) and the Kruskal-Wallis test (quantitative data). The smooth curve fitting method was used to analyze the nonlinear relationship between sleep duration and these outcomes. According to the results of the smooth curve fitting, the threshold effect analysis was used to investigate whether there were significant differences between the linear model and the 2-piecewise model. A *P* value of <.05 for logarithmic likelihood ratio test was considered to detect the threshold effect of sleep duration on outcomes. A *P* value <.05 was considered statistically significant in all statistical analyses. R software (R Foundation for Statistical Computing) was used for statistical analysis [19].

Results

Baseline Data of Participants in Different Sleep Duration

A total of 54,896 participants with sleep questionnaire information in the NHANES (2005-2020) were included. After excluding 232 participants with insufficient sleep questionnaire information, 54,664 participants were finally included (Figure 2). Significant differences between the 4 groups were detected in terms of various baseline data, including race composition, educational background, marital status, BMI, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes (Table 1).

In addition, significant differences between the 4 groups were detected in terms of analyzed outcomes, including obesity-related outcomes, musculoskeletal index, and diseases, as well as mortality because of different causes (Table 2).

Figure 2. The flow diagram of screening eligible participants from 8 survey cycles (2005-2020) of the National Health and Nutrition Examination Survey (NHANES) program and research structure of association analysis between sleep duration and various clinical outcomes. BMD: bone mineral density; CVD: cardiovascular disease; HUA: hyperuricemia; OA: osteoarthritis; RA: rheumatoid arthritis; UA: uric acid.

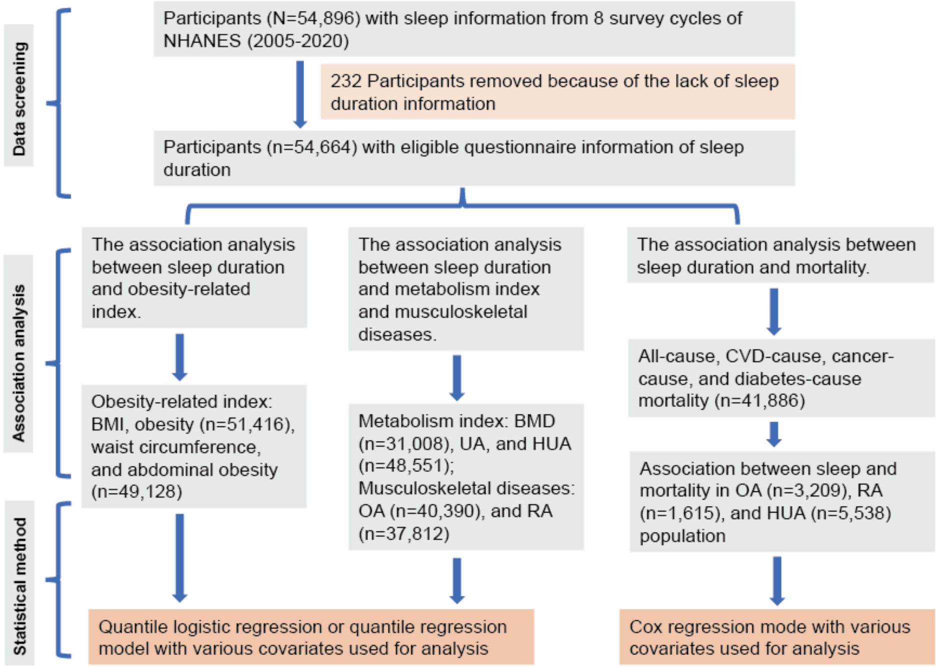


Table 1. Baseline data comparison of 54,664 participants in different groups divided by sleep duration on the basis of the National Health and Nutrition Examination Survey program (2005-2020; N=54,664; Q1, Q2, Q3, and Q4 represented the intervals of different sleep durations).

	Q1 (n=6862)	Q2 (n=10,524)	Q3 (n=14,198)	Q4 (n=23,080)	*P* value
Sleep duration (hours)	1-5.5	6-6.5	7-7.5	8-14.5	N/A[a]
Age (years), mean (SD)	47.76 (17.97)	46.03 (18.21)	45.76 (18.81)	46.65 (21.37)	<.001
Sex, n (%)					<.001
Male	3452 (50.31)	5345 (50.79)	7111 (50.08)	10,695 (46.34)	
Female	3410 (49.69)	5179 (49.21)	7087 (49.92)	12,385 (53.66)	
Race (Hispanic origin), n (%)					<.001
Mexican American	827 (12.05)	1618 (15.37)	2223 (15.66)	3940 (17.07)	
Non-Hispanic Black	2445 (35.63)	2792 (26.53)	2645 (18.63)	4667 (20.22)	
Non-Hispanic White	2180 (31.77)	3791 (36.02)	5990 (42.19)	9393 (40.70)	
Other Hispanic	704 (10.26)	1050 (9.98)	1331 (9.37)	2176 (9.43)	
Other race	706 (10.29)	1273 (12.10)	2009 (14.15)	2904 (12.58)	
Education level, n (%)					<.001
<9th grade	684 (10.60)	883 (9.17)	1154 (9.03)	2344 (11.78)	
9th-11th grade	1070 (16.58)	1316 (13.67)	1483 (11.60)	2845 (14.30)	
High school graduate	1669 (25.87)	2235 (23.22)	2764 (21.62)	4602 (23.13)	
Some college or AA[b] degree	2095 (32.47)	3026 (31.44)	3670 (28.71)	5685 (28.57)	
College graduate or above	929 (14.40)	2156 (22.40)	3706 (28.99)	4378 (22.00)	
Refused	4 (0.06)	4 (0.04)	4 (0.03)	7 (0.04)	
Do not know	1 (0.02)	6 (0.06)	3 (0.02)	39 (0.20)	
Marital status, n (%)					<.001
Married	2938 (44.83)	5040 (51.28)	7305 (55.84)	10,202 (49.86)	
Widowed	715 (10.91)	914 (9.30)	1201 (9.18)	2644 (12.92)	
Divorced	987 (15.06)	1232 (12.54)	1433 (10.95)	2394 (11.70)	
Separated	283 (4.32)	306 (3.11)	276 (2.11)	473 (2.31)	
Never married	1138 (17.37)	1677 (17.06)	2057 (15.73)	3396 (16.60)	
Living with partner	487 (7.43)	649 (6.60)	802 (6.13)	1332 (6.51)	
Refused	5 (0.08)	10 (0.10)	6 (0.05)	15 (0.07)	
Do not know	0 (0)	0 (0)	1 (0.01)	4 (0.02)	
BMI (kg/m^2), mean (SD)	30.09 (7.91)	29.33 (7.19)	28.68 (6.82)	28.49 (7.04)	<.001
Total energy intake (kcal), mean (SD)	2136.61 (1147.03)	2166.06 (1035.51)	2159.49 (992.91)	2064.69 (978.45)	<.001
Physical activity, n (%)					<.001
No	3938 (58.14)	6033 (58.04)	8368 (59.77)	14,298 (63.53)	
Yes	2835 (41.86)	4361 (41.96)	5633 (40.23)	8209 (36.47)	
Alcohol consumption, n (%)					N/A
No	2035 (36.44)	2679 (31.83)	3763 (33.68)	6590 (38.29)	
Yes	3550 (63.56)	5737 (68.17)	7409 (66.32)	10,620 (61.71)	
Smoking, n (%)					<.001
No	3230 (49.43)	5471 (55.89)	7778 (59.59)	12,107 (58.44)	
Yes	3304 (50.57)	4318 (44.11)	5274 (40.41)	8609 (41.56)	
Hypertension, n (%)					<.001
No	4095 (59.81)	6980 (66.39)	10,079 (71.05)	15,542 (67.45)	

	Q1 (n=6862)	Q2 (n=10,524)	Q3 (n=14,198)	Q4 (n=23,080)	*P* value
Yes	2752 (40.19)	3534 (33.61)	4107 (28.95)	7500 (32.55)	
Diabetes, n (%)					<.001
No	5684 (85.15)	9040 (87.86)	12,434 (89.67)	19,728 (87.29)	
Yes	991 (14.85)	1249 (12.14)	1433 (10.33)	2873 (12.71)	

[a] N/A: not applicable.

[b] AA: Associate of Arts.

Table 2. The comparison of health-related metabolism index, disease information, and mortality information in different groups divided by sleep duration among 54,664 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020; N=54,446; Q1, Q2, Q3, and Q4 represented the intervals of different sleep durations).

	Q1 (n=6862)	Q2 (n=10,524)	Q3 (n=14,198)	Q4 (n=23,080)	*P* value
Sleep duration (hours)	1-5.5	6-6.5	7-7.5	8-14.5	N/A[a]
Obesity, n (%)					<.001
No	3713 (57.16)	6142 (61.34)	8647 (64.45)	13,999 (65.14)	
Yes	2783 (42.84)	3871 (38.66)	4770 (35.55)	7491 (34.86)	
Waist circumference, mean (SD)	100.46 (17.81)	98.87 (17.09)	97.44 (16.38)	97.18 (17.02)	<.001
Abdominal obesity, n (%)					<.001
No	2600 (42.08)	4368 (45.40)	6173 (47.83)	9641 (47.21)	
Yes	3579 (57.92)	5254 (54.60)	6734 (52.17)	10,779 (52.79)	
Femur BMD[b], mean (SD)	1.02 (0.19)	1.02 (0.18)	1.01 (0.18)	1.00 (0.18)	<.001
Uric acid, mean (SD)	5.50 (1.47)	5.46 (1.45)	5.38 (1.40)	5.37 (1.46)	<.001
Hyperuricemia, n (%)					<.001
No	4935 (80.43)	7741 (81.78)	10,550 (83.28)	16,747 (82.57)	
Yes	1201 (19.57)	1725 (18.22)	2118 (16.72)	3534 (17.43)	
Osteoarthritis, n (%)					<.001
No	4267 (86.25)	7101 (88.41)	9680 (88.53)	14,169 (85.99)	
Yes	680 (13.75)	931 (11.59)	1254 (11.47)	2308 (14.01)	
Rheumatoid arthritis, n (%)					<.001
No	4267 (89.31)	7101 (93.47)	9680 (94.91)	14,169 (92.98)	
Yes	511 (10.69)	496 (6.53)	519 (5.09)	1069 (7.02)	
All-cause mortality, n (%)					<.001
No	4959 (87.32)	7830 (90.33)	10,064 (92.09)	14,529 (87.47)	
Yes	720 (12.68)	838 (9.67)	864 (7.91)	2082 (12.53)	
CVD[c]-cause mortality, n (%)					<.001
No	5430 (95.62)	8367 (96.53)	10,645 (97.41)	15,849 (95.41)	
Yes	249 (4.38)	301 (3.47)	283 (2.59)	762 (4.59)	
Cancer-cause mortality, n (%)					<.001
No	5514 (97.09)	8475 (97.77)	10,703 (97.94)	16,169 (97.34)	
Yes	165 (2.91)	193 (2.23)	225 (2.06)	442 (2.66)	
Diabetes-cause mortality, n (%)					<.001
No	5581 (98.27)	8579 (98.97)	10,842 (99.21)	16,389 (98.66)	
Yes	98 (1.73)	89 (1.03)	86 (0.79)	222 (1.34)	

[a] N/A: not applicable.

[b] BMD: bone mineral density.

[c] CVD: cardiovascular disease.

Correlation Between Sleep Duration and Musculoskeletal Diseases, Obesity, and Metabolism-Related Outcomes

As shown in Tables 3-6, a negative association between sleep duration and negative association was detected in the whole quantile of all the 3 models. In model 2, the β values of sleep duration on BMI in Q2, Q3, and Q4 were −.308, −.687, and −.931, respectively, and the ORs for obesity in Q2, Q3, and Q4 were 0.924, 0.874, and 0.810, respectively. The β values of sleep duration on WC in Q2, Q3, and Q4 of model 2 were −.593, −1.449, and −1.456, respectively. A negative correlation with abdominal obesity in Q3 (OR 0.898, 95% CI 0.831-0.971) and Q4 (OR 0.871, 95% CI 0.809-0.937) in model 2 was also detected. In contrast, a positive correlation with BMD in Q3 (OR 0.012, 95% CI 0.006-0.008) and Q4 (OR 0.008, 95% CI

0.002-0.013) in model 2 was found, whereas no significant association was detected between sleep duration and UA as well as HUA. In addition, a negative correlation between sleep duration and OA was detected in Q2 (OR 0.839, 95% CI 0.733-0.959), Q3 (OR 0.780, 95% CI 0.686-0.887), and Q4 (OR 0.771, 95% CI 0.683-0.870) in model 2. Similarly, a negative association between sleep duration and RA was detected in the whole quantile of all the 3 models. Subsequently, the nonlinear association using the smooth curve fitting method indicated that sleep duration showed a negative association with BMI, obesity, WC, and abdominal obesity (Figures 3A and 4D). In addition, a n-shaped association between sleep duration and BMD and HUA was detected (Figures 3E and 4F). However, an L-shaped association was observed between sleep duration and OA and RA (Figures 3G and 4H). The threshold effect analysis detected a significant threshold effect of sleep duration on BMI, obesity, and WC (Table 7) as well as on BMD, HUA, OA, and RA (Table 8). The relative inflection points in BMI, obesity, and WC were 8 hours. The relative inflection points in BMD, HUA, OA, and RA were 7.5 hours, 5 hours, 8 hours, and 6.5 hours, respectively. Below the inflection point, negative correlations were detected with BMI, obesity, WC, OA, and RA, whereas a positive correlation was observed between BMD and HUA.

Table 3. The association between sleep duration and obesity-related outcomes using the quantile regression method among 51,416 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020). Model 1 was adjusted for age, sex, race, educational level, and marital status; model 2 was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

Body index	β (95% CI)	*P* value
BMI		
Nonadjusted (N=51,416)		
1-5.5	0 (0)	N/A[a]
6-6.5	−.756 (−.979 to .534)	<.001
7-7.5	−1.407 (−1.618 to 1.196)	<.001
8-14.5	−1.595 (−1.793 to 1.397)	<.001
Model 1 (N=45,818)		
1-5.5	0 (0)	N/A
6-6.5	−.433 (−.658 to .208)	<.001
7-7.5	−.843 (−1.059 to .627)	<.001
8-14.5	−1.034 (−1.236 to .831)	<.001
Model 2 (N=37,979)		
1-5.5	0 (0)	N/A
6-6.5	−.308 (−.545 to .070)	.01
7-7.5	−.687 (−.916 to .458)	<.001
8-14.5	−.931 (−1.146 to .716)	<.001
WC[b]		
Nonadjusted (N=49,128)		
1-5.5	0 (0)	N/A
6-6.5	−1.587 (−2.129 to 1.044)	<.001
7-7.5	−3.015 (−3.529 to 2.500)	<.001
8-14.5	−3.276 (-3.759 to 2.793)	<.001
Model 1 (N=43,676)		
1-5.5	0 (0)	N/A
6-6.5	−.860 (−1.386 to .334)	.001
7-7.5	−1.875 (−2.380 to 1.370)	<.001
8-14.5	−1.791 (−2.266 to 1.316)	<.001
Model 2 (N=37,195)		
1-5.5	0 (0)	N/A
6-6.5	−.593 (−1.143 to .042)	.03
7-7.5	−1.449 (−1.979 to 0.919)	<.001
8-14.5	−1.456 (−1.955 to 0.957)	<.001

[a] N/A: not applicable.

[b] WC: waist circumference.

Table 4. The association between sleep duration and obesity-related outcomes using the quantile regression method among 51,416 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020).

Obesity outcomes	OR[a] (95% CI)	*P* value
Obesity		
Nonadjusted (N=51,416)		
1-5.5	1 (0)	N/A[b]
6-6.5	0.841 (0.789-0.896)	<.001
7-7.5	0.736 (0.693-0.782)	<.001
8-14.5	0.714 (0.675-0.755)	<.001
Model 1 (N=45,818)		
1-5.5	1 (0)	N/A
6-6.5	0.903 (0.845-0.966)	.003
7-7.5	0.844 (0.791-0.900)	<.001
8-14.5	0.799 (0.752-0.849)	<.001
Model 2 (N=37,979)		
1-5.5	1 (0)	N/A
6-6.5	0.924 (0.857-0.996)	.04
7-7.5	0.874 (0.813-0.940)	<.001
8-14.5	0.810 (0.756-0.867)	<.001
Abdominal obesity		
Nonadjusted (N=49,128)		
1-5.5	1 (0)	N/A
6-6.5	0.874 (0.819-0.932)	<.001
7-7.5	0.792 (0.745-0.842)	<.001
8-14.5	0.812 (0.767-0.860)	<.001
Model 1 (N=43,676)		
1-5.5	1 (0)	N/A
6-6.5	0.940 (0.875-1.011)	.10
7-7.5	0.871 (0.812-0.933)	<.001
8-14.5	0.856 (0.802-0.914)	<.001
Model 2 (N=37,195)		
1-5.5	1 (0)	N/A
6-6.5	0.961 (0.886-1.041)	.33
7-7.5	0.898 (0.831-0.971)	.007
8-14.5	0.871 (0.809-0.937)	<.001

[a] OR: odds ratio.

[b] N/A: not applicable.

Table 5. The association between sleep duration and bone mineral density, uric acid level, using the quantile regression method among participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020). Model 1 was adjusted for age, sex, race, educational level, and marital status; model 2 was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

Metabolism index	β (95% CI)	*P* value
BMD[a] (g/cm^2)		
Nonadjusted (N=31,008)		
1-5.5	0 (0)	N/A[b]
6-6.5	.002 (−.005 to .009)	.61
7-7.5	−.001 (−.008 to .006)	.74
8-14.5	−.020 (−.026 to −.013)	<.001
Model 1 (N=27,966)		
1-5.5	0 (0)	N/A
6-6.5	.004 (−.002 to .010)	.16
7-7.5	.006 (.001 to .012)	.03
8-14.5	.001 (−.005 to .006)	.84
Model 2 (N=24,485)		
1-5.5	0 (0)	N/A
6-6.5	.006 (−.000 to .012)	.06
7-7.5	.012 (.006 to .017)	<.001
8-14.5	.008 (.002 to .013)	.006
UA[c](mg/dL)		
Nonadjusted (N=48,551)		
1-5.5	0 (0)	N/A
6-6.5	−.043 (−.089 to .004)	.07
7-7.5	−.117 (−.161 to −.073)	<.001
8-14.5	−.134 (−.176 to −.093)	<.001
Model 1 (N=43,437)		
1-5.5	0 (0)	N/A
6-6.5	−.014 (−.058 to .030)	.53
7-7.5	−.060 (−.102 to −.018)	.005
8-14.5	−.044 (−.084 to .005)	.03
Model 2 (N=36,060)		
1-5.5	0 (0)	N/A
6-6.5	.008 (−.037 to .053)	.73
7-7.5	.008 (−.035 to .052)	.70
8-14.5	.017 (−.024 to .057)	.43

[a] BMD: bone mineral density.

[b] N/A: not applicable.

[c] UA: uric acid.

Table 6. The association between sleep duration and hyperuricemia, osteoarthritis, and rheumatoid arthritis using the quantile regression method among participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020).

Diseases	OR[a] (95% CI)	*P* value
HUA[b]		
Nonadjusted (N=48,511)		
1-5.5	1 (0)	N/A[c]
6-6.5	0.916 (0.844-0.994)	.03
7-7.5	0.825 (0.763-0.892)	<.001
8-14.5	0.867 (0.806-0.933)	<.001
Model 1 (N=43,437)		
1-5.5	1 (0)	N/A
6-6.5	0.987 (0.906-1.075)	.77
7-7.5	0.921 (0.848-1.001)	.05
8-14.5	0.949 (0.879-1.025)	.18
Model 2 (N=36,060)		
1-5.5	1 (0)	
6-6.5	1.041 (0.944-1.149)	.42
7-7.5	1.059 (0.963-1.164)	.24
8-14.5	1.057 (0.967-1.155)	.22
OA[d]		
Nonadjusted (N=40,390)		
1-5.5	1 (0)	N/A
6-6.5	0.823 (0.740-0.915)	<.001
7-7.5	0.813 (0.736-0.898)	<.001
8-14.5	1.022 (0.932-1.121)	.64
Model 1 (N=40,390)		
1-5.5	1 (0)	N/A
6-6.5	0.796 (0.709-0.895)	<.001
7-7.5	0.705 (0.630-0.788)	<.001
8-14.5	0.728 (0.657-0.808)	<.001
Model 2 (N=31,502)		
1-5.5	1 (0)	N/A
6-6.5	0.839 (0.733-0.959)	.01
7-7.5	0.780 (0.686-0.887)	<.001
8-14.5	0.771 (0.683-0.870)	<.001
RA[e]		
Nonadjusted (N=37,812)		
1-5.5	1 (0)	N/A
6-6.5	0.583 (0.513-0.664)	<.001
7-7.5	0.448 (0.394-0.509)	<.001
8-14.5	0.630 (0.564-0.704)	<.001
Model 1 (N=37,812)		
1-5.5	1 (0)	N/A

Diseases	OR[a] (95% CI)	*P* value
6-6.5	0.647 (0.564-0.741)	<.001
7-7.5	0.515 (0.450-0.589)	<.001
8-14.5	0.587 (0.521-0.660)	<.001
Model 2 (N=29,476)		
1-5.5	1 (0)	N/A
6-6.5	0.667 (0.572-0.779)	<.001
7-7.5	0.561 (0.481-0.654)	<.001
8-14.5	0.631 (0.551-0.724)	<.001

[a] OR: odds ratio.

[b] HUA: hyperuricemia.

[c] N/A: not applicable.

[d] OA: osteoarthritis.

[e] RA: rheumatoid arthritis.

Figure 3. The nonlinear association analysis between sleep duration and (A) BMI; (B) obesity; (C) waist circumference; (D) abdominal obesity; (E) bone mineral density; (F) hyperuricemia; (G) osteoarthritis; (H) rheumatoid arthritis; and (I) Aal-cause, (J) cardiovascular disease (CVD)–cause, (K) cancer-cause, and (L) diabetes-cause mortality among participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020) through smooth curve fitting. Note: Except for diabetes-cause mortality, the analysis was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes; the analysis of diabetes-cause mortality was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, and hypertension.

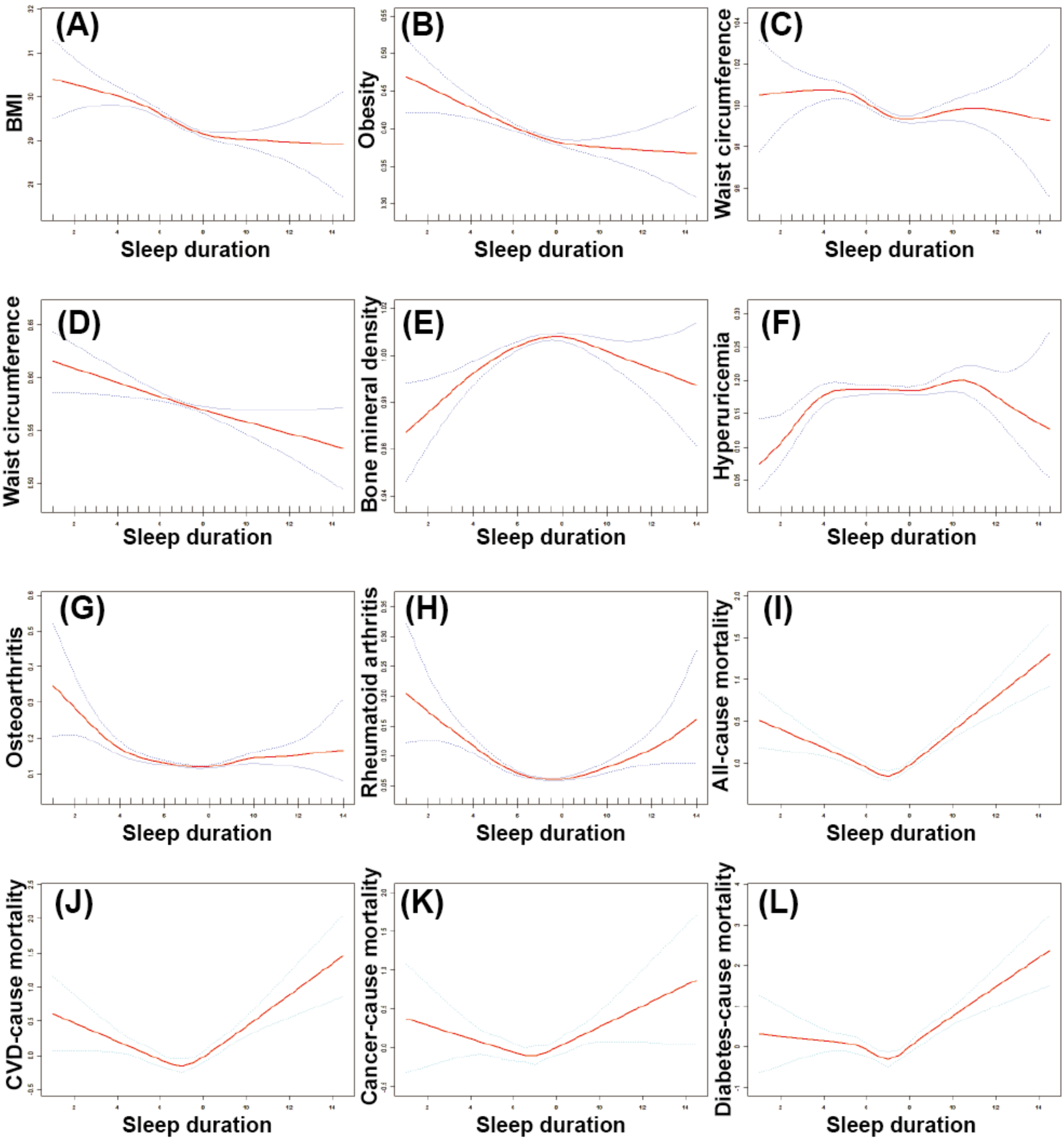


Figure 4. The nonlinear association analysis between sleep duration and mortality of participants experiencing diseases or metabolism disorders through smooth curve fitting: the nonlinear association between sleep and all-cause mortality in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020; (A) all-cause mortality, (B) cardiovascular disease (CVD)–cause mortality, (C) cancer-cause mortality, and (D) diabetes-cause mortality of participants in the US hyperuricemia population; the nonlinear association between sleep and (E) all-cause mortality, (F) CVD-cause mortality, (G) cancer-cause mortality, and (H) diabetes-cause mortality of participants in the US osteoarthritis population; the nonlinear association between sleep and (I) all-cause mortality, (J) CVD-cause mortality, (K) cancer-cause mortality, and (L) diabetes-cause mortality of participants in the US rheumatoid arthritis population. Except for diabetes-cause mortality, the analysis was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes; the analysis of diabetes-cause mortality was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, and hypertension.

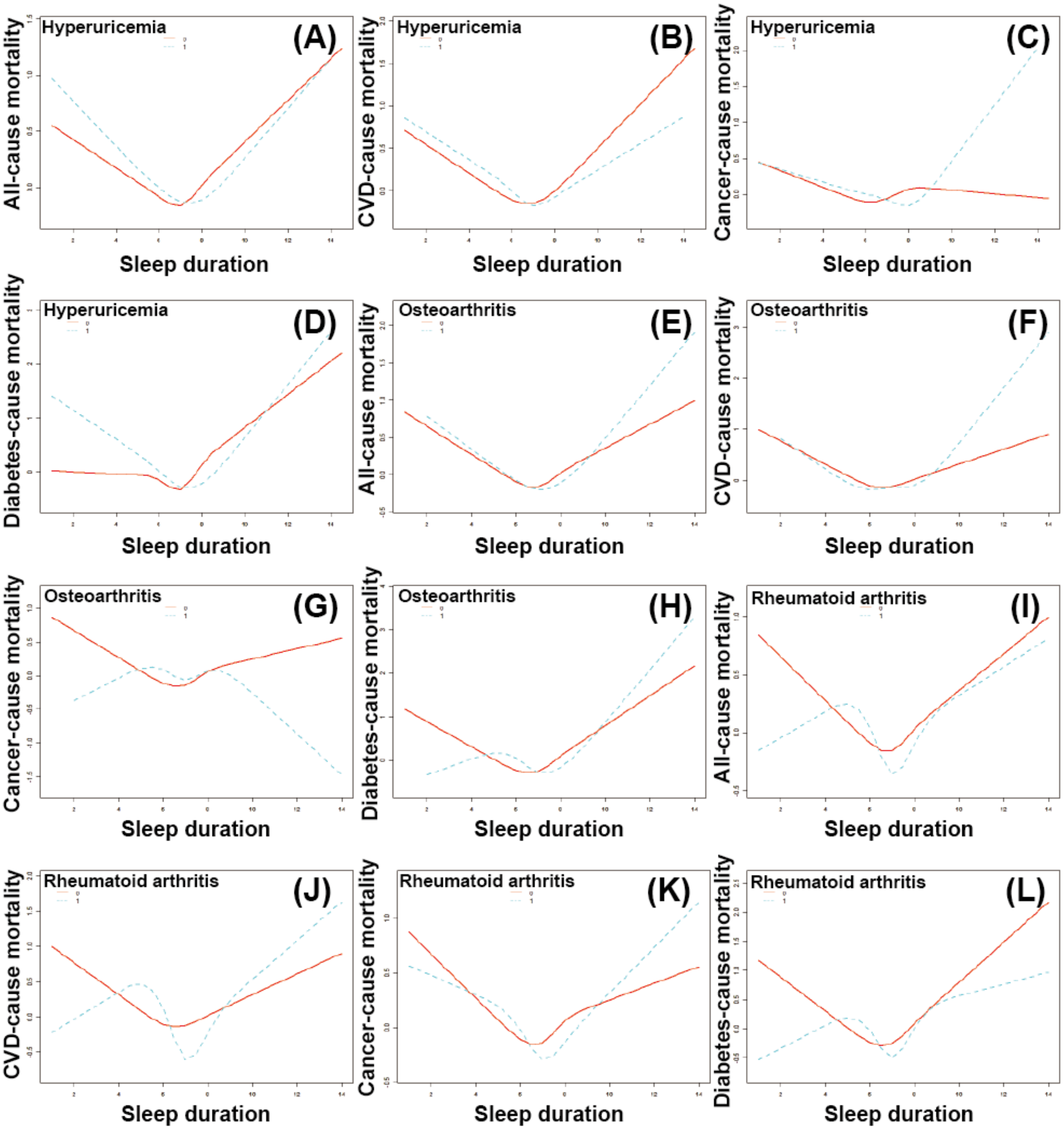


Table 7. Threshold effect analysis[a] of sleep duration on obesity-related outcomes among 37,979 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020).

Obesity-related outcomes	BMI (N=37,979), β (95% CI)	Obesity (N=37,979), OR[b] (95% CI)	WC[c] (n=37,195), β (95% CI)	AB obesity[d] (n=37,195), OR (95% CI)
Effect analysis by one linear model	−0.176[e] (−0.220 to −0.133)	0.964[e] (0.950 to 0.977)	−0.219[e] (−0.320 to −0.117)	0.975[e] (0.960 to 0.990)
Effect analysis by 2-fold piecewise model				
Effect difference	−0.249[e] (−0.309 to −0.190)	0.943[e] (0.925 to 0.961)	−0.407[e] (−0.546 to −0.269)	0.963[e] (0.943 to 0.982)
Sleep duration<inflection point	0.006 (−0.104 to 0.116)	1.018 (0.983 to 1.054)	0.258[f] (0.000 to 0.516)	1.008 (0.970 to 1.047)
Sleep duration≥inflection point	0.255[e] (0.114 to 0.397)	1.079[g] (1.032 to 1.129)	0.666[e] (0.335 to 0.997)	1.047 (0.997 to 1.099)

[a] The model used for the threshold effect analysis was adjusted for age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

[b] OR: odds ratio.

[c] WC: waist circumference.

[d] AB obesity: abdominal obesity.

[e] *P*<.001.

[f] *P*<.05.

[g] *P*<.01.

Table 8. Threshold effect analysis[a] of sleep duration on bone mineral density (BMD), hyperuricemia, osteoarthritis, and rheumatoid arthritis (RA) among 36,060 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020; N=36,060).

Obesity-related outcomes	BMD (n=24,485), β (95% CI)	Hyperuricemia (N=36,060), OR[b] (95% CI)	Osteoarthritis (n=31,502), OR (95% CI)	RA (n=29,476), OR (95% CI)
Effect analysis by one linear model	0.002[c] (0.001 to 0.003)	1.015 (0.997 to 1.033)	0.965[c] (0.942 to 0.990)	0.940[d] (0.912 to 0.968)
Effect analysis by 2-fold piecewise model				
Sleep duration<inflection point	0.006[d] (0.004 to 0.007)	1.139[c] (1.035 to 1.253)	0.901[d] (0.871 to 0.932)	0.726[d] (0.681 to 0.773)
Sleep duration≥inflection point	−0.005[d] (−0.007 to −0.002)	1.002 (0.981 to 1.023)	1.135[d] (1.070 to 1.203)	1.103[d] (1.055 to 1.154)
Effect difference	−0.010[d] (−0.014 to −0.006)	0.880[e] (0.793 to 0.976)	1.259[d] (1.166 to 1.360)	1.520[d] (1.385 to 1.668)

[a] The model used for threshold effect analysis was adjusted for age, sex, race, educational level, marital status, BMI, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

[b] OR: odds ratio.

[c] *P*<.01.

[d] *P*<.001.

[e] *P*<.05.

Association Between Sleep Duration and Mortality in General Population

As shown in Table 9, a significantly negative correlation between sleep duration and all-cause mortality was found in Q2 (HR 0.844) and Q3 (HR 0.749) of model 2, respectively. Consistently, a significantly negative correlation with CVD-cause mortality in Q2 (HR 0.773) in model 2 was discovered. For cancer-cause mortality, a significant HR in Q2 (HR 0.801) and Q3 (HR 0.798) of model 1 was detected, which was absent in model 2. For diabetes-cause mortality, a significantly negative HR was detected in Q2 (HR 0.654) and Q3 (HR 0.0.562) of model 1 as well as Q3 (HR 0.633) of model 2. The analysis results of smooth curve fitting indicated an L-shaped association between sleep duration and all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality (Figures 3I and 4L). The threshold effect analysis detected a significant threshold effect of sleep duration on all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality, and an inflection point of 6.5 hours was detected (Table 10). When below the inflection point, a significantly protective HR of sleep duration was detected in terms of all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality. However, when above the inflection point, a significant risk HR of sleep duration was detected in terms of all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality.

Table 9. The association between sleep duration and hazard ratios for mortality in population using the quantile regression method among 51,416 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020).

	Nonadjusted	*P* value	Model 1	*P* value	Model 2	*P* value
All-cause mortality, n (%)	41,886 (100)	N/A[a]	39,501 (100)	N/A	31,419 (100)	N/A
1-5.5, HR[b] (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.752 (0.681-0.831)	<.001	0.800 (0.724-0.884)	<.001	0.844 (0.751-0.949)	.005
7-7.5, HR (95% CI)	0.651 (0.589-0.718)	<.001	0.708 (0.641-0.783)	<.001	0.749 (0.666-0.842)	<.001
8-14.5, HR (95% CI)	1.171 (1.076-1.275)	<.001	0.969 (0.888-1.056)	.47	0.975 (0.879-1.081)	.63
CVD[c]-cause mortality, n (%)	41,886 (100)	N/A	39,501 (100)	N/A	31,419 (100)	N/A
1-5.5, HR (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.781 (0.661-0.924)	.004	0.843 (0.712-0.998)	.047	0.909 (0.745-1.107)	.37
7-7.5, HR (95% CI)	0.617 (0.520-0.731)	<.001	0.689 (0.579-0.818)	<.001	0.773 (0.632-0.945)	.01
8-14.5, HR (95% CI)	1.241 (1.076-1.432)	.003	1.010 (0.873-1.169)	.89	1.020 (0.857-1.215)	.82
Cancer-cause mortality, n (%)	41,886 (100)	N/A	39,501 (100)	N/A	31,419 (100)	N/A
1-5.5, HR (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.756 (0.614-0.931)	.008	0.801 (0.650-0.987)	.04	0.805 (0.632-1.025)	.08
7-7.5, HR (95% CI)	0.738 (0.604-0.903)	.003	0.798 (0.650-0.978)	.03	0.813 (0.642-1.030)	.09
8-14.5, HR (95% CI)	1.078 (0.902-1.290)	.41	0.925 (0.771-1.110)	.40	0.964 (0.780-1.192)	.74
Diabetes-cause mortality, n (%)	41,886 (100)	N/A	39,501 (100)	N/A	31,419 (100)	N/A
1-5.5, HR (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.587 (0.441-0.782)	<.001	0.654 (0.490-0.873)	.004	0.738 (0.525, 1.038)	.08
7-7.5, HR (95% CI)	0.476 (0.357-0.636)	<.001	0.562 (0.419-0.754)	<.001	0.633 (0.446-0.899)	.01
8-14.5, HR (95% CI)	0.922 (0.726-1.169)	.50	0.842 (0.661-1.072)	.16	1.072 (0.805-1.429)	.63

[a] N/A: not applicable.

[b] HR: heart rate.

[c] CVD: cardiovascular disease.

Table 10. Threshold effect analysis of sleep duration on mortality with different causes among 31,419 participants in 8 survey cycles in the National Health and Nutrition Examination Survey program (2005-2020; N=31,419).

Mortality	All-cause mortality (N=31,419)[a], HR[b] (95% CI)	Cardiovascular disease–cause mortality (N=31,419)[a], HR (95% CI)	Cancer-cause mortality (N=31,419)[a], HR (95% CI)	Diabetes-cause mortality (n=32,173)[c], HR (95% CI)
Effect analysis by one linear model	1.036[d] (1.012-1.060)	1.043[e] (1.003-1.084)	1.028 (0.980-1.079)	1.018 (0.955-1.084)
Effect analysis by 2-fold piecewise model				
Sleep duration<inflection point	0.847[f] (0.806-0.889)	0.835[f] (0.770-0.906)	0.892[e] (0.804-0.990)	0.813[d] (0.709-0.933)
Sleep duration≥inflection point	1.169[f] (1.129-1.210)	1.188[f] (1.122-1.258)	1.120[d] (1.041-1.206)	1.150[d] (1.050-1.260)
Effect difference	1.381[f] (1.285-1.484)	1.422[f] (1.262-1.603)	1.255[f] (1.077-1.462)	1.414[f] (1.161-1.722)

[a] The model used for the threshold effect analysis was adjusted for age, sex, race, educational level, marital status, total energy intake, BMI, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

[b] HR: hazard ratio.

[c] The model used for threshold effect analysis was adjusted for age, sex, race, educational level, marital status, total energy intake, BMI, physical activity, alcohol consumption, smoking, and hypertension.

[d] *P*<.01.

[e] *P*<.05.

[f] *P*<.001.

Association Between Sleep Duration and Mortality in RA, OA, and HUA Populations

As shown in Table 11, in the HUA population, the HR of all-cause mortality in Q2, Q3, and Q4 was 0.784 (95% CI 0.633-0.972; *P*=.03), 0.669 (95% CI 0.538-0.832; *P*<.001), and 0.786 (95% CI 0.649-0.953; *P*=.01), respectively. The HR of CVD-cause mortality in Q3 was 0.667 (95% CI 0.471-0.944; *P*=.02). The HR of diabetes-cause mortality in Q2 and Q3 was 0.587 (95% CI 0.345-0.996; *P*=.048) and 0.464 (95% CI 0.266-0.810; *P*=.007), respectively. In the OA population, sleep duration was identified as a protective factor for OA in Q3 of all-cause mortality. The HR for all-cause mortality in Q3 was 0.647 (95% CI 0.487-0.861; *P*=.003). In the RA population, sleep duration was identified as a protective factor for RA in Q3 (HR 0.500, 95% CI 0.338-0.742; *P*<.001) of all-cause mortality and Q3 (HR 0.302, 95% CI 0.143-0.638; *P*=.002) of CVD-cause mortality. The analysis results of the smooth curve fitting showed an L-shaped association between sleep duration and mortality in participants with HUA (Figure 4) and a n-shaped association between sleep duration and cancer-cause mortality in the OA population. The threshold effect analysis found a significant threshold effect of sleep duration on all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality in the HUA population, which was also detected for all-cause and CVD-cause mortality in the OA and RA populations as well as diabetes-cause mortality in the OA population. In the OA population, the detected inflection points of sleep duration for all-cause, CVD-cause, and diabetes-cause mortality were 6.5 hours, 8.5 hours, and 7.5 hours, respectively (Table 12).

Table 11. The effect of sleep duration on mortality of population experiencing musculoskeletal disorders using the quantile regression method among 5538 participants in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020). The model was adjusted by age, sex, race, educational level, marital status, total energy intake, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

	HUA[a] population	*P* value	OA[b] population	*P* value	RA[c] population	*P* value
All-cause mortality, n (%)	5538 (100)	N/A[d]	3209 (100)	N/A	1615 (100)	N/A
1-5.5, HR[e] (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.784 (0.633-0.972)	.03	0.782 (0.591-1.034)	.08	0.956 (0.683-1.337)	.79
7-7.5, HR (95% CI)	0.669 (0.538-0.832)	<.001	0.647 (0.487-0.861)	.003	0.500 (0.338-0.742)	<.001
8-14.5, HR (95% CI)	0.786 (0.649-0.953)	.01	0.863 (0.670-1.110)	.25	0.913 (0.675-1.236)	.56
CVD-cause mortality, n (%)	5538 (100)	N/A	3209 (100)	N/A	1615 (100)	N/A
1-5.5, HR (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.818 (0.583-1.147)	.24	0.772 (0.471-1.266)	.30	0.985 (0.575-1.688)	.96
7-7.5, HR (95% CI)	0.667 (0.471-0.944)	.02	0.795 (0.492-1.285)	.35	0.302 (0.143-0.638)	.002
8-14.5, HR (95% CI)	0.804 (0.594-1.088)	.16	1.097 (0.720-1.672)	.67	0.791 (0.477-1.311)	.36
Cancer-cause mortality, n (%)	5538 (100)	N/A	3209 (100)	N/A	1615 (100)	N/A
1-5.5, HR (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.919 (0.565-1.494)	.73	1.489 (0.778-2.848)	.23	0.817 (0.421-1.584)	.55
7-7.5, HR (95% CI)	0.796 (0.486-1.303)	.36	1.080 (0.552-2.114)	.82	0.522 (0.246-1.107)	.09
8-14.5, HR (95% CI)	0.947 (0.610-1.469)	.81	1.211 (0.646-2.272)	.55	0.833 (0.463-1.497)	.54
Diabetes-cause mortality, n (%)	5725 (100)	N/A	3345 (100)	N/A	1669 (100)	N/A
1-5.5, HR (95% CI)	1 (0)	N/A	1 (0)	N/A	1 (0)	N/A
6-6.5, HR (95% CI)	0.587 (0.345-0.996)	.048	0.763 (0.338-1.723)	.52	0.970 (0.342-2.746)	.95
7-7.5, HR (95% CI)	0.464 (0.266-0.810)	.007	0.737 (0.323-1.685)	.47	0.396 (0.101-1.552)	.18
8-14.5, HR (95% CI)	0.661 (0.418-1.046)	.08	0.975 (0.481-1.975)	.94	1.220 (0.490-3.036)	.67

[a] HUA: hyperuricemia.

[b] OA: osteoarthritis.

[c] RA: rheumatoid arthritis.

[d] N/A: not applicable.

[e] HR: heart rate.

Table 12. Threshold effect analysis[a] of sleep duration on mortality with different causes among 3345 participants with osteoarthritis (OA), 1669 participants with rheumatoid arthritis (RA), and 5725 participants with hyperuricemia (HUA) in 8 survey cycles of the National Health and Nutrition Examination Survey program (2005-2020).

Specific population	All-cause mortality	CVD[b]-cause mortality	Cancer-cause mortality	Diabetes-cause mortality
OA population, n (%)	3209 (100)	3209 (100)	3209 (100)	3345 (100)
Effect analysis by one linear model, HR[c] (95% CI)	1.037 (0.980-1.098)	1.153[d] (1.052-1.264)	0.978 (0.863-1.109)	1.164 (0.988-1.370)
Effect analysis by 2-fold piecewise model, HR (95% CI)				
Sleep duration<inflection point	0.755[e] (0.671-0.851)	0.961 (0.855-1.079)	1.003 (0.874-1.152)	0.851 (0.665-1.087)
Sleep duration≥inflection point	1.236[e] (1.140-1.339)	1.778[e] (1.470-2.150)	0.558 (0.138-2.259)	1.663[e] (1.288-2.148)
Effect difference	1.636[e] (1.377-1.942)	1.850[e] (1.424-2.404)	0.556 (0.132-2.339)	1.956[d] (1.274-3.003)
RA population, n (%)	1615 (100)	1615 (100)	1615 (100)	1669 (100)
Effect analysis by one linear model, HR (95% CI)	1.000 (0.934-1.070)	0.984 (0.875-1.106)	0.980 (0.859-1.119)	1.073 (0.876-1.314)
Effect analysis by 2-fold piecewise model, HR (95% CI)				
Sleep duration<inflection point	0.845[f] (0.739-0.965)	0.753[d] (0.608-0.933)	0.794 (0.619-1.019)	0.985 (0.782-1.241)
Sleep duration≥inflection point	1.133[f] (1.017-1.262)	1.215[f] (1.014-1.457)	1.161 (0.940-1.434)	1.893 (0.897-3.995)
Effect difference	1.341[d] (1.091-1.649)	1.614[d] (1.151-2.264)	1.462 (0.984-2.173)	1.921 (0.807-4.575)
HUA population, n (%)	5538 (100)	5538 (100)	5538 (100)	5725 (100)
Effect analysis by one linear model, HR (95% CI)	0.979 (0.937-1.023)	0.978 (0.913-1.049)	1.042 (0.946-1.149)	0.988 (0.881-1.108)
Effect analysis by 2-fold piecewise model, HR (95% CI)				
Sleep duration<inflection point	0.852[f] (0.797-0.911)	0.852[d] (0.766-0.947)	0.942 (0.843-1.053)	0.743[f] (0.633-0.871)
Sleep duration≥inflection point	1.216[f] (1.113-1.328)	1.212[d] (1.055-1.393)	1.731[f] (1.300-2.305)	1.546[f] (1.264-1.891)
Effect difference	1.427[f] (1.250-1.629)	1.423[d] (1.154-1.755)	1.838[f] (1.300-2.597)	2.083[f] (1.527-2.840)

[a] The model used for the threshold effect analysis was adjusted for age, sex, race, educational level, marital status, total energy intake, BMI, physical activity, alcohol consumption, smoking, hypertension, and diabetes.

[b] CVD: cardiovascular disease.

[c] HR: hazard ratio.

[d] *P*<.01.

[e] *P*<.001.

[f] *P*<.05.

Below the inflection point, a significantly decreased HR for all-cause mortality was detected, whereas a significantly increased HR for all-cause, CVD-cause, and diabetes-cause mortality was detected above the inflection point. In the RA population, the detected inflection point of sleep duration for all-cause and CVD-cause mortality was 6.5 hours. Below the inflection point, a significantly decreased HR for all-cause and CVD-cause mortality was detected. Above the inflection point, a significantly increased HR for all-cause and CVD-cause mortality was detected. In the HUA population, the inflection point of sleep duration for all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality in model 2 was 7.5 hours, 7.5 hours, 9 hours, and 7.5 hours, respectively. Below the inflection point, a significantly decreased HR for all-cause, CVD-cause, and diabetes-cause mortality was detected. Above the inflection point, a significantly increased HR for all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality was detected.

Discussion

Principal Findings

In this study, we found that increased sleep duration was associated with decreased BMI, obesity, WC, abdominal obesity, OA, and RA but increased BMD. The nonlinear analysis indicated that sleep duration caused a saturation effect on obesity, abdominal obesity, and HUA. In addition, a significant threshold effect of sleep duration on BMD; OA; RA; and all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality was found.

Sleep is an indispensable physiological activity for human beings that is closely associated with physical and mental health and the incidence and development of various diseases. Despite the critical significance of healthy sleep, sleep disorders have gradually increased over the past 2 decades, with more than a third of adults and 57% of teenagers in the United States experiencing insomnia [20]. With aggravating the worrying sleep dilemma, accumulating interests have been aroused in exploring sleep's physiological and pathological role in human beings. According to the International Classification of Sleep Disorders published by the American Academy of Sleep Medicine, the diagnosis of sleep disorders can be divided into 7 types, including insomnia, sleep-related breathing disorders, central disorders of hypersomnolence, circadian rhythm sleep-wake disorders, parasomnias, sleep-related movement disorders, and other sleep disorders [21]. These sleep disorders would significantly impair sleep duration, which is essential to maintain mental health [22,23].

Many studies have investigated the association between obesity and sleep duration, with controversial conclusions achieved. A negative association between obesity and sleep duration has been detected in some studies [24-27], whereas none of these studies have a sample size of>10,000 participants. In addition, Kripke et al [28] reported a U-shaped association between BMI and sleep duration in 636,095 female participants in the United States, with an inflection point of sleep duration (7 to 8 hours). In comparison, no significant association between BMI and sleep duration was detected in 480,841 male participants. To the best of our knowledge, we are the first to find the saturation effect of sleep duration on obesity-related outcomes, indicating a negative association between obesity and sleep duration when sleep duration is <8 hours. Some epidemiological studies have pointed out that sleep duration may cause obesity and endocrine changes such as decreased glucose tolerance and insulin sensitivity, increased evening concentrations of cortisol, increased levels of ghrelin, decreased levels of leptin, and increased hunger and appetite [29,30]. Although only a few experimental studies have explored the underlying mechanism behind this phenomenon, it could be hypothesized that the underlying mechanism is complex and works in multiple networks.

In this study, we have found a n-shaped association between BMD and sleep duration, with the inflection point of 7.5 hours sleep duration detected. On the basis of the NHANES program 2017-2018, Lee et al [6] found that the osteoporosis incidence of female participants aged >50 years with a sleep duration of <5 hours was significantly higher than that of normal participants (OR 7.35) in the NHANES program 2005-2010, which was almost consistent with the findings of Shiao et al [31]. In addition, Tang et al [32] found that participants with shorter or longer sleep durations showed lower BMD than those with normal sleep duration (7 to 9 hours), which was consistent with our findings. It should not be ignored that some studies reported no significant association between sleep duration and BMD [33] and osteoporosis [34]. These epidemiological studies revealed that sleep duration seems to be a double-edged sword in affecting the change of BMD. However, more experimental evidence should be provided to demonstrate this phenomenon. Swanson et al [35] found that the level of P1NP, a biomarker of bone formation, decreased when the sleep circadian rhythm was disrupted, which was more pronounced among younger people and indicated a detrimental effect of sleep disorders on normal bone turnover. In addition, it was reported that growth hormone might be an important effective agent involving sleep duration affecting bone mass [36]. In general, many epidemiological studies have demonstrated the detrimental effect of sleep disorders on BMD, although exploring specific molecular mechanisms needs further work.

UA is the terminal metabolite of purine compounds, of which approximately 80% are endogenous and 20% are derived from exogenous food rich in purine or nucleic acid proteins [37]. Under normal physiological conditions, the production and catabolism of UA are in dynamic homeostasis, mainly regulated by liver synthesis and excretion from or absorption in the kidney and intestine [38]. Excessive serum UA, HUA, was reported to be closely associated with various diseases, including gout, chronic kidney disease, and CVD [39,40]. A few epidemiological studies have investigated the association between sleep duration and UA or HUA. Lee et al [41] found a U-shaped association between sleep duration and UA in Korean women, indicating that female individuals with a sleep duration of 7 to 8 hours showed the lowest UA level. In addition, many related studies have reported the negative association that increased sleep duration was significantly associated with a lower risk of HUA [42-44]. However, in this study, we found a threshold effect between sleep duration and HUA, indicating a positive association between HUA and sleep duration when sleep duration was <5 hours. This finding was different from that of previous studies. Owing to the lack of experimental evidence and controversy regarding epidemiological studies, it could be challenging to draw an affirmative conclusion related to the association between sleep duration and UA level up till now.

OA is one of the most common musculoskeletal diseases, seriously impairing the health and social functioning of the US population. Recent epidemiological studies have investigated the association between sleep duration and OA. On the basis of the Korea NHANES program, Park et al [45] found a U-shaped association between sleep duration and OA in the older woman population (aged >50 years), with a lower incidence of OA occurring in the group reporting a 7 to 8 hours sleep duration. Ni et al [46] found that the increased OA incidence was positively correlated with a shorter sleep duration. These studies were conducted within a specific population in Korea. At the same time, there were few studies to investigate the association between sleep duration and OA in the general US population. What we performed in this study tried to answer this question, and we found a threshold effect of sleep duration on OA incidence in the US population. The OA incidence increased with the increase in sleep duration; however, this synergistically positive association disappeared when sleep duration increased to 8 hours. It is worth noting that Ni et al [46] found that patients with OA aged >60 years in the NHANES program (2011 to 2018) experienced a higher incidence of sleep trouble [46]. These studies and our findings have demonstrated a positive association between sleep duration and OA incidence, although the causal association between them is yet to be clarified. It is well known that progressing OA pathology into the advanced stage causes severe pain and impairs the patient's function and sleep quality [47,48]. Therefore, the more profound association between sleep duration and OA, such as a causal relationship, should be verified by more prospective clinical studies.

As an autoimmune disease, RA is caused by multiple factors, including genetic susceptibility and environmental factors [49]. Kim et al [50] found that the RA incidence in the Korean population decreased with an increase in sleep duration when sleep duration was <6 hours. They also found that severe RA pain was closely associated with sleep duration when it was <6 hours [50]. Meanwhile, Wu et al [51] found that shorter sleep duration, <6 hours, and poor sleep quality were positively associated with RA incidence. This study found that increased sleep duration was negatively associated with RA. A threshold effect (inflection point: 6.5 hours) was detected in the association analysis between sleep duration and RA. The association we detected between sleep duration (<6.5 hours) and RA was almost consistent with the aforementioned studies. It is worth noting that we were the first to find that excessive sleep duration (>6.5 hours) was a risk factor for RA. However, the causal relationship between sleep duration and RA was challenging to be verified because the pain symptom was one of the main characteristics of RA, which could also cause detrimental influences on sleep quality. Notably, Gao et al [52] performed a Mendelian randomization study to investigate the causal relationship between sleep traits and RA. They found no association between sleep duration and frequent insomnia, any insomnia, sleep duration, or snoring, although the casual effect of shorter sleep duration (<6 hours) on RA was detected through inverse variance weighted and weighted median methods, which provided further evidence to support our findings. Despite this accumulating epidemiological evidence and continuous insight into the pathogenesis of RA, research on how sleep duration or traits affect the occurrence and development of RA still needs to be completed. Further experimental studies focusing on this perspective may yield a new understanding of the pathogenesis of RA.

Many epidemiological studies have widely investigated the association between sleep duration and mortality [2,53,54]. The U-shaped association (inflection point: 6.5 hours) detected in this study between sleep duration and all-cause, CVD-cause, and cancer-cause mortality was first detected in the general US population, which was consistent with a previous study involving a specific population, such as older or woman populations [55-57]. With the change in dietary structure, diabetes has gradually become one of the most common aging diseases, causing serious detrimental effects on health and longevity [58,59]. However, mortality, mainly or partly because of diabetes, has been less investigated, especially for investigating the effect of life behaviors on diabetes-cause mortality in the general or specific population. Hou et al [60] investigated the association between diabetes-cause mortality and carbohydrate intake behaviors and found that low-quality carbohydrates at dinner were positively associated with diabetes-cause mortality. We first explored the association between sleep duration and diabetes-cause mortality, with a U-shaped association detected in the general US population. We also detected a similar U-shaped association between sleep duration and diabetes-cause mortality in the OA and HUA populations. These findings provide a strong epidemiological evidence that lifestyle behaviors, such as dietary and sleep behaviors, may have played an essential role in affecting poor diabetes-related outcomes. A large amount of epidemiological evidence have been provided to demonstrate that shorter or excessive sleep duration causes a detrimental effect on mortality outcomes. Unlike previous studies, we are the first to identify a threshold effect of sleep duration on mortality in the OA, RA, and HUA populations. This discovery offers new clinical evidence about the impact of sleep duration on mortality outcomes within these specific populations, highlighting the potential benefits of maintaining an appropriate sleep duration for individuals in these groups. There are some potential explanations from the perspective of mechanisms to clarify the negative association between shorter sleep duration and mortality, including increased cortisol secretion, altered growth hormone metabolism, inflammation activation, and changes in circulating leptin and ghrelin [2,61,62]. Although no such hypothesis has been provided to clarify the detrimental effect of excessive sleep duration on mortality in the general or specific population, further experimental studies must be conducted.

Study Limitations

Some limitations should be addressed when generalizing the conclusions. First, as a cross-sectional study, this study highlighted the association between sleep duration and various outcomes; however, it was not possible to infer the causal association between sleep duration and these outcomes. Mendelian randomization analysis could be an alternative method to investigate their causal association, which may be conducted to further determine their causal association in the US population. Second, the diagnosis of OA and RA was defined through a questionnaire survey method, which may lead to misdiagnosing or missed diagnosis of some participants and impair the evidence level of our conclusion. Third, although we included various covariates for adjustment in the analysis, there must be missing covariates associated with the outcomes analyzed in this study. For example, genetic factors, family history, and environmental exposures have been considered as the risk factors for some diseases, such as RA and obesity. Owing to the limitation of data availability, it was difficult to include these individual information as covariates. Furthermore, the sleep duration information of children was absent in the NHANES program. Therefore, we should be careful when generalizing the conclusions to specific populations. In addition, the quality of sleep is an important factor influencing the health status; however, the lack of this individual information in the NHANES program makes it difficult to analyze the role of sleep quality on individual health, which should be investigated in future studies. In general, more prospective and related experimental studies or Mendelian randomization studies are needed to further validate the conclusions of this study.

Conclusions

In our study, which included a total of 54,664 participants across 8 survey cycles of the NHANES program, we observed that longer sleep duration was associated with decreased BMI, obesity, WC, abdominal obesity, OA, and RA but increased BMD. Furthermore, we found that sleep duration caused a saturation effect on obesity, abdominal obesity, and HUA. In addition, a significant threshold effect of sleep duration on BMD; OA; RA; and all-cause, CVD-cause, cancer-cause, and diabetes-cause mortality was found. In conclusion, this study provided further evidence about the double-edged sword effect of sleep duration on health, and proper sleep duration plays a critical role in maintaining good health status. These findings suggested that ensuring adequate sleep duration should be an integral part of public health programs.

Acknowledgments

This study was supported by the Natural Science Foundation of China (grant 82360437). The authors would like to thank the National Health and Nutrition Examination Survey (NHANES) team and its participants.

Data Availability

All data used in this study are available on the National Health and Nutrition Examination Survey (NHANES) website [63].

Authors' Contributions

TL, ML, HQ, and JY collected and analyzed the data. TL was a major contributor to this manuscript. ML contributed to the ethics approval progress. LH takes full responsibility for the work, including the study design, access to data, and the decision to submit and publish the manuscript.

Conflicts of Interest

None declared.

References

1. Knutson KL, Van Cauter E. Associations between sleep loss and increased risk of obesity and diabetes. Ann N Y Acad Sci 2008 Jun 28;1129(1):287-304 [FREE Full text] [doi: 10.1196/annals.1417.033] [Medline: 18591489]
2. Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 2010 May;33(5):585-592 [FREE Full text] [doi: 10.1093/sleep/33.5.585] [Medline: 20469800]
3. Svensson T, Saito E, Svensson AK, Melander O, Orho-Melander M, Mimura M, et al. Association of sleep duration with all- and major-cause mortality among adults in Japan, China, Singapore, and Korea. JAMA Netw Open 2021 Sep 01;4(9):e2122837 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.22837] [Medline: 34477853]
4. Wang C, Bangdiwala SI, Rangarajan S, Lear SA, AlHabib KF, Mohan V, et al. Association of estimated sleep duration and naps with mortality and cardiovascular events: a study of 116 632 people from 21 countries. Eur Heart J 2019 May 21;40(20):1620-1629 [FREE Full text] [doi: 10.1093/eurheartj/ehy695] [Medline: 30517670]
5. Hirsch JK, Altier HR, Offenbächer M, Toussaint L, Kohls N, Sirois FM. Positive psychological factors and impairment in rheumatic and musculoskeletal disease: do psychopathology and sleep quality explain the linkage? Arthritis Care Res (Hoboken) 2021 Jan 04;73(1):55-64. [doi: 10.1002/acr.24440] [Medline: 32937011]
6. Lee CL, Tzeng HE, Liu WJ, Tsai CH. A cross-sectional analysis of the association between sleep duration and osteoporosis risk in adults using 2005-2010 NHANES. Sci Rep 2021 Apr 27;11(1):9090 [FREE Full text] [doi: 10.1038/s41598-021-88739-x] [Medline: 33907283]
7. Martínez-Rodríguez A, Rubio-Arias J, Ramos-Campo DJ, Reche-García C, Leyva-Vela B, Nadal-Nicolás Y. Psychological and sleep effects of tryptophan and magnesium-enriched Mediterranean diet in women with fibromyalgia. Int J Environ Res Public Health 2020 Mar 26;17(7):2227 [FREE Full text] [doi: 10.3390/ijerph17072227] [Medline: 32224987]
8. Zou L, Yeung A, Quan X, Boyden SD, Wang H. A systematic review and meta-analysis of mindfulness-based (Baduanjin) exercise for alleviating musculoskeletal pain and improving sleep quality in people with chronic diseases. Int J Environ Res Public Health 2018 Jan 25;15(2):206 [FREE Full text] [doi: 10.3390/ijerph15020206] [Medline: 29370149]
9. Dashti HS, Ordovás JM. Genetics of sleep and insights into its relationship with obesity. Annu Rev Nutr 2021 Oct 11;41(1):223-252. [doi: 10.1146/annurev-nutr-082018-124258] [Medline: 34102077]
10. St-Onge MP. Sleep-obesity relation: underlying mechanisms and consequences for treatment. Obes Rev 2017 Feb 06;18 Suppl 1(S1):34-39. [doi: 10.1111/obr.12499] [Medline: 28164452]
11. Reutrakul S, Van Cauter E. Sleep influences on obesity, insulin resistance, and risk of type 2 diabetes. Metabolism 2018 Jul;84:56-66. [doi: 10.1016/j.metabol.2018.02.010] [Medline: 29510179]
12. Petrov ME, Long DL, Grandner MA, MacDonald LA, Cribbet MR, Robbins R, et al. Racial differences in sleep duration intersect with sex, socioeconomic status, and U.S. geographic region: the REGARDS study. Sleep Health 2020 Aug;6(4):442-450 [FREE Full text] [doi: 10.1016/j.sleh.2020.05.004] [Medline: 32601040]
13. Troxel WM, Lee L, Hall M, Matthews KA. Single-parent family structure and sleep problems in Black and White adolescents. Sleep Med 2014 Feb;15(2):255-261 [FREE Full text] [doi: 10.1016/j.sleep.2013.10.012] [Medline: 24424100]

14. Di H, Guo Y, Daghlas I, Wang L, Liu G, Pan A, et al. Evaluation of sleep habits and disturbances among US adults, 2017-2020. JAMA Netw Open 2022 Nov 01;5(11):e2240788 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.40788] [Medline: 36346632]
15. NCHS Ethics Review Board (ERB) approval. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm#print [accessed 2018-04-01]
16. Njeh CF, Fuerst T, Hans D, Blake GM, Genant HK. Radiation exposure in bone mineral density assessment. Appl Radiat Isot 1999 Jan;50(1):215-236. [doi: 10.1016/s0969-8043(98)00026-8] [Medline: 10028639]
17. March LM, Schwarz JM, Carfrae BH, Bagge E. Clinical validation of self-reported osteoarthritis. Osteoarthritis Cartilage 1998 Mar;6(2):87-93 [FREE Full text] [doi: 10.1053/joca.1997.0098] [Medline: 9692063]
18. Gregg EW, Cheng YJ, Srinivasan M, Lin J, Geiss LS, Albright AL, et al. Trends in cause-specific mortality among adults with and without diagnosed diabetes in the USA: an epidemiological analysis of linked national survey and vital statistics data. Lancet 2018 Jun;391(10138):2430-2440. [doi: 10.1016/s0140-6736(18)30314-3]
19. Tippmann S. Programming tools: adventures with R. Nature 2015 Jan 01;517(7532):109-110. [doi: 10.1038/517109a] [Medline: 25557714]
20. Mattingly SM. Sleep, health, productivity, and the double-edged sword of technology. XRDS Crossroads ACM Mag Stud 2021 Sep 21;28(1):24-27. [doi: 10.1145/3481827]
21. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. Chest 2014 Nov;146(5):1387-1394. [doi: 10.1378/chest.14-0970] [Medline: 25367475]
22. Matricciani LA, Olds TS, Blunden S, Rigney G, Williams MT. Never enough sleep: a brief history of sleep recommendations for children. Pediatrics 2012 Mar;129(3):548-556. [doi: 10.1542/peds.2011-2039] [Medline: 22331340]
23. Matricciani L, Bin YS, Lallukka T, Kronholm E, Dumuid D, Paquet C, et al. Past, present, and future: trends in sleep duration and implications for public health. Sleep Health 2017 Oct;3(5):317-323. [doi: 10.1016/j.sleh.2017.07.006] [Medline: 28923186]
24. Gangwisch JE, Malaspina D, Boden-Albala B, Heymsfield SB. Inadequate sleep as a risk factor for obesity: analyses of the NHANES I. Sleep 2005 Oct;28(10):1289-1296. [doi: 10.1093/sleep/28.10.1289] [Medline: 16295214]
25. Eisenmann JC, Ekkekakis P, Holmes M. Sleep duration and overweight among Australian children and adolescents. Acta Paediatr 2006 Aug 1;95(8):956-963. [doi: 10.1080/08035250600731965] [Medline: 16882569]
26. Ko GT, Chan JC, Chan AW, Wong PT, Hui SS, Tong SD, et al. Association between sleeping hours, working hours and obesity in Hong Kong Chinese: the 'better health for better Hong Kong' health promotion campaign. Int J Obes (Lond) 2007 Feb 23;31(2):254-260. [doi: 10.1038/sj.ijo.0803389] [Medline: 16718283]
27. Cappuccio FP, Taggart FM, Kandala NB, Currie A, Peile E, Stranges S, et al. Meta-analysis of short sleep duration and obesity in children and adults. Sleep 2008 May;31(5):619-626 [FREE Full text] [doi: 10.1093/sleep/31.5.619] [Medline: 18517032]
28. Kripke DF, Garfinkel L, Wingard DL, Klauber MR, Marler MR. Mortality associated with sleep duration and insomnia. Arch Gen Psychiatry 2002 Feb;59(2):131-136. [doi: 10.1001/archpsyc.59.2.131] [Medline: 11825133]
29. Zimberg IZ, Dâmaso A, Del Re M, Carneiro AM, de Sá Souza H, de Lira FS, et al. Short sleep duration and obesity: mechanisms and future perspectives. Cell Biochem Funct 2012 Aug;30(6):524-529. [doi: 10.1002/cbf.2832] [Medline: 22473743]
30. Felső R, Lohner S, Hollódy K, Erhardt É, Molnár D. Relationship between sleep duration and childhood obesity: systematic review including the potential underlying mechanisms. Nutr Metab Cardiovasc Dis 2017 Sep;27(9):751-761. [doi: 10.1016/j.numecd.2017.07.008] [Medline: 28818457]
31. Shiao YC, Chen WT, Chen WL. Association of short sleep duration and trabecular bone score. Sci Rep 2021 Oct 06;11(1):19821 [FREE Full text] [doi: 10.1038/s41598-021-99410-w] [Medline: 34615958]
32. Tang Y, Wang S, Yi Q, Xia Y, Geng B. Sleep pattern and bone mineral density: a cross-sectional study of National Health and Nutrition Examination Survey (NHANES) 2017-2018. Arch Osteoporos 2021 Oct 23;16(1):157. [doi: 10.1007/s11657-021-01025-1] [Medline: 34689259]
33. Marques EA, Figueiredo P, Gudnason V, Lang T, Sigurdsson G, Sigurdsson S, et al. Associations of 24-hour sleep duration and CT-derived measurements of muscle and bone: the AGES-Reykjavik study. Exp Gerontol 2017 Jul;93:1-6 [FREE Full text] [doi: 10.1016/j.exger.2017.04.002] [Medline: 28404506]
34. Moradi S, Shab-Bidar S, Alizadeh S, Djafarian K. Association between sleep duration and osteoporosis risk in middle-aged and elderly women: a systematic review and meta-analysis of observational studies. Metabolism 2017 Apr;69:199-206. [doi: 10.1016/j.metabol.2017.01.027] [Medline: 28162775]
35. Swanson CM, Shea SA, Wolfe P, Cain SW, Munch M, Vujovic N, et al. Bone turnover markers after sleep restriction and circadian disruption: a mechanism for sleep-related bone loss in humans. J Clin Endocrinol Metab 2017 Oct 01;102(10):3722-3730 [FREE Full text] [doi: 10.1210/jc.2017-01147] [Medline: 28973223]
36. Stich FM, Huwiler S, D'Hulst G, Lustenberger C. The potential role of sleep in promoting a healthy body composition: underlying mechanisms determining muscle, fat, and bone mass and their association with sleep. Neuroendocrinology 2022;112(7):673-701 [FREE Full text] [doi: 10.1159/000518691] [Medline: 34348331]

37. Maiuolo J, Oppedisano F, Gratteri S, Muscoli C, Mollace V. Regulation of uric acid metabolism and excretion. Int J Cardiol 2016 Jun 15;213:8-14 [FREE Full text] [doi: 10.1016/j.ijcard.2015.08.109] [Medline: 26316329]
38. Lima WG, Martins-Santos ME, Chaves VE. Uric acid as a modulator of glucose and lipid metabolism. Biochimie 2015 Sep;116:17-23. [doi: 10.1016/j.biochi.2015.06.025] [Medline: 26133655]
39. Prasad Sah OS, Qing YX. Associations between hyperuricemia and chronic kidney disease: a review. Nephrourol Mon 2015 May 23;7(3):e27233 [FREE Full text] [doi: 10.5812/numonthly.7(3)2015.27233] [Medline: 26290849]
40. Zhang S, Wang Y, Cheng J, Huangfu N, Zhao R, Xu Z, et al. Hyperuricemia and cardiovascular disease. Curr Pharm Des 2019;25(6):700-709. [doi: 10.2174/1381612825666190408122557] [Medline: 30961478]
41. Lee YC, Son DH, Kwon YJ. U-shaped association between sleep duration, C-reactive protein, and uric acid in Korean women. Int J Environ Res Public Health 2020 Apr 13;17(8):2657 [FREE Full text] [doi: 10.3390/ijerph17082657] [Medline: 32294936]
42. Yu X, Gong S, Chen J, Zhang H, Shen Z, Gu Y, et al. Short sleep duration increases the risk of hyperuricemia among Chinese adults: findings from the China Health and Nutrition Survey. Sleep Med 2021 Aug;84:40-45. [doi: 10.1016/j.sleep.2021.05.014] [Medline: 34091320]
43. An Y, Li X, Ouyang F, Xiao S. Association between nocturnal sleep duration and the risk of hyperuricemia among Chinese government employees: a cross-sectional study. Front Public Health 2022 Nov 23;10:1055778 [FREE Full text] [doi: 10.3389/fpubh.2022.1055778] [Medline: 36504942]
44. Papandreou C, Babio N, Díaz-López A, Martínez-González MA, Becerra-Tomas N, Corella D, et al. Sleep duration is inversely associated with serum uric acid concentrations and uric acid to creatinine ratio in an elderly mediterranean population at high cardiovascular risk. Nutrients 2019 Apr 01;11(4):761 [FREE Full text] [doi: 10.3390/nu11040761] [Medline: 30939748]
45. Park HM, Kwon YJ, Kim HS, Lee YJ. Relationship between sleep duration and osteoarthritis in middle-aged and older women: a nationwide population-based study. J Clin Med 2019 Mar 13;8(3):356 [FREE Full text] [doi: 10.3390/jcm8030356] [Medline: 30871193]
46. Ni J, Zhou W, Cen H, Chen G, Huang J, Yin K, et al. Evidence for causal effects of sleep disturbances on risk for osteoarthritis: a univariable and multivariable Mendelian randomization study. Osteoarthritis Cartilage 2022 Mar;30(3):443-450 [FREE Full text] [doi: 10.1016/j.joca.2021.11.021] [Medline: 34890811]
47. Turk DC, Cohen MJ. Sleep as a marker in the effective management of chronic osteoarthritis pain with opioid analgesics. Semin Arthritis Rheum 2010 Jun;39(6):477-490. [doi: 10.1016/j.semarthrit.2008.10.006] [Medline: 19136144]
48. Sasaki E, Tsuda E, Yamamoto Y, Maeda S, Inoue R, Chiba D, et al. Nocturnal knee pain increases with the severity of knee osteoarthritis, disturbing patient sleep quality. Arthritis Care Res (Hoboken) 2014 Jul;66(7):1027-1032 [FREE Full text] [doi: 10.1002/acr.22258] [Medline: 24470323]
49. Costenbader KH, Chang SC, De Vivo I, Plenge R, Karlson EW. Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking. Arthritis Res Ther 2008;10(3):R52 [FREE Full text] [doi: 10.1186/ar2421] [Medline: 18462498]
50. Kim JH, Park EC, Lee KS, Lee Y, Shim S, Kim J, et al. Association of sleep duration with rheumatoid arthritis in Korean adults: analysis of seven years of aggregated data from the Korea National Health and Nutrition Examination Survey (KNHANES). BMJ Open 2016 Dec 13;6(12):e011420 [FREE Full text] [doi: 10.1136/bmjopen-2016-011420] [Medline: 27965246]
51. Wu W, Yang J, Gu Y, Chen X, Tan X. Dose-response relationship between sleep and rheumatoid arthritis. Am J Health Behav 2020 Jan 01;44(1):40-53. [doi: 10.5993/AJHB.44.1.5] [Medline: 31783931]
52. Gao RC, Sang N, Jia CZ, Zhang MY, Li BH, Wei M, et al. Association between sleep traits and rheumatoid arthritis: a mendelian randomization study. Front Public Health 2022 Jun 30;10:940161 [FREE Full text] [doi: 10.3389/fpubh.2022.940161] [Medline: 35844889]
53. Irwin MR, Olmstead R, Carroll JE. Sleep disturbance, sleep duration, and inflammation: a systematic review and meta-analysis of cohort studies and experimental sleep deprivation. Biol Psychiatry 2016 Jul 01;80(1):40-52 [FREE Full text] [doi: 10.1016/j.biopsych.2015.05.014] [Medline: 26140821]
54. Yin J, Jin X, Shan Z, Li S, Huang H, Li P, et al. Relationship of sleep duration with all-cause mortality and cardiovascular events: a systematic review and dose-response meta-analysis of prospective cohort studies. J Am Heart Assoc 2017 Sep 09;6(9):e005947 [FREE Full text] [doi: 10.1161/JAHA.117.005947] [Medline: 28889101]
55. Xiao Q, Blot WJ, Matthews CE. Weekday and weekend sleep duration and mortality among middle-to-older aged White and Black adults in a low-income southern US cohort. Sleep Health 2019 Oct;5(5):521-527 [FREE Full text] [doi: 10.1016/j.sleh.2019.04.008] [Medline: 31204307]
56. Gupta K, Nagalli S, Kalra R, Gupta R, Mahmood S, Jain V, et al. Sleep duration, baseline cardiovascular risk, inflammation and incident cardiovascular mortality in ambulatory U.S. Adults: national health and nutrition examination survey. Am J Prev Cardiol 2021 Aug 17;8:100246 [FREE Full text] [doi: 10.1016/j.ajpc.2021.100246] [Medline: 34485966]
57. Kabat GC, Xue X, Kamensky V, Zaslavsky O, Stone KL, Johnson KC, et al. The association of sleep duration and quality with all-cause and cause-specific mortality in the Women's Health Initiative. Sleep Med 2018 Oct;50:48-54. [doi: 10.1016/j.sleep.2018.05.015] [Medline: 29982090]

58. Furman D, Campisi J, Verdin E, Carrera-Bastos P, Targ S, Franceschi C, et al. Chronic inflammation in the etiology of disease across the life span. Nat Med 2019 Dec;25(12):1822-1832 [FREE Full text] [doi: 10.1038/s41591-019-0675-0] [Medline: 31806905]
59. Lin D, Xiao M, Zhao J, Li Z, Xing B, Li X, et al. An overview of plant phenolic compounds and their importance in human nutrition and management of type 2 diabetes. Molecules 2016 Oct 15;21(10):1374 [FREE Full text] [doi: 10.3390/molecules21101374] [Medline: 27754463]
60. Hou W, Han T, Sun X, Chen Y, Xu J, Wang Y, et al. Relationship between carbohydrate intake (quantity, quality, and time eaten) and mortality (total, cardiovascular, and diabetes): assessment of 2003-2014 national health and nutrition examination survey participants. Diabetes Care 2022 Dec 01;45(12):3024-3031. [doi: 10.2337/dc22-0462] [Medline: 36174119]
61. Spiegel K, Tasali E, Penev P, Van Cauter E. Brief communication: sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. Ann Intern Med 2004 Dec 07;141(11):846-850. [doi: 10.7326/0003-4819-141-11-200412070-00008] [Medline: 15583226]
62. Taheri S, Lin L, Austin D, Young T, Mignot E. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. PLoS Med 2004 Dec;1(3):e62 [FREE Full text] [doi: 10.1371/journal.pmed.0010062] [Medline: 15602591]
63. National Health and Nutrition Examination Survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2023-10-24]

Abbreviations

BMD: bone mineral density
CVD: cardiovascular disease
HR: hazard ratio
HUA: hyperuricemia
NHANES: National Health and Nutrition Examination Survey
OA: osteoarthritis
OR: odds ratio
RA: rheumatoid arthritis
UA: uric acid
WC: waist circumference

Edited by A Mavragani, T Sanchez; submitted 10.02.23; peer-reviewed by A Qin, H Liao, DK Yon; comments to author 17.05.23; revised version received 22.05.23; accepted 26.09.23; published 07.11.23.

Please cite as:
Lei T, Li M, Qian H, Yang J, Hu Y, Hua L
The Effect of Sleep on Metabolism, Musculoskeletal Disease, and Mortality in the General US Population: Analysis of Results From the National Health and Nutrition Examination Survey
JMIR Public Health Surveill 2023;9:e46385
URL: https://publichealth.jmir.org/2023/1/e46385
doi:10.2196/46385
PMID:37934562

Original Paper

Changing Trends in the Global Burden of Cataract Over the Past 30 Years: Retrospective Data Analysis of the Global Burden of Disease Study 2019

Bo Jiang[1*], MD; Tianhong Wu[1*], MM; Weiming Liu[1], MD; Gaoqin Liu[1], MD; Peirong Lu[1], MD, PhD

Department of Ophthalmology, The First Affiliated Hospital of Soochow University, Suzhou, China

[*]these authors contributed equally

Corresponding Author:
Peirong Lu, MD, PhD
Department of Ophthalmology
The First Affiliated Hospital of Soochow University
188 Shizi Street
Suzhou, 215006
China
Phone: 86 51265215191
Email: lupeirong@suda.edu.cn

Abstract

Background: Cataracts now account for the largest proportion of the global burden of blindness and vision loss. Understanding the changing trends in the global burden of cataracts over the past 30 years and the next 15 years is of clear significance for the prevention and control of cataracts in key populations. As far as we know, research on the future burden of cataracts is lacking.

Objective: This study aims to assess the global burden of cataracts over the past 30 years by using age-period-cohort modeling and to estimate trends in the next 15 years.

Methods: Data were obtained from the Global Burden of Disease Study 2019, the United Nations Development Programme, and the WHO (World Health Organization) Global Health Observatory data repository. The assessment of trends and disparities in the number and rate of disability-adjusted life years (DALYs) for cataracts from 1990 to 2019 was conducted. The association between the age-standardized DALY rate (ASDR) and the socio-demographic index (SDI), human development index (HDI), national levels of particulate matter <2.5 µm in diameter ($PM_{2.5}$), and ambient ultraviolet radiation (UVR) was determined using linear regression analysis. Additionally, we used the Nordpred (Harald Fekjær and Bjørn Møller) age-period-cohort model to predict the cataract burden from 2020 to 2034.

Results: Globally, the number of DALYs due to cataract increased from 3,492,604 (95% uncertainty interval [UI] 2,481,846-4,719,629) in 1990 to 6,676,281 (95% UI 4,761,210-9,006,193) in 2019. The ASDRs due to cataract decreased from 93.17 (95% UI 66.14-125.32) in 1990 to 82.94 (95% UI 59.06-111.75) in 2019, with an average annual percentage change of –0.37 (95% CI –0.44 to –0.3; $P<.001$). Age, female sex, air pollution, smoking, high fasting plasma glucose levels, and a high body mass index were risk factors for the burden of cataracts. SDI and HDI were negatively correlated with ASDRs of cataracts, while $PM_{2.5}$ and UVR were positively associated with them. Higher DALY rates were also associated with lower SDI ($R^2=0.1939$; $P<.001$), lower HDI ($R^2=0.2828$; $P<.001$), national $PM_{2.5}$ concentration ($R^2=0.1874$; $P<.001$), and ambient UVR levels ($R^2=0.2354$; $P<.001$). The prediction model suggested that the number of DALYs due to cataract will continue to rise globally, while the cataract DALY rate will continue to decrease.

Conclusions: While the ASDR of cataracts has decreased, there has been a notable increase in the number of DALYs over the past 30 years. Projections suggest that the global burden of cataracts will continue to rise over the next 15 years. To address this challenge, appropriate prevention and treatment policies must be implemented.

(JMIR Public Health Surveill 2023;9:e47349) doi:10.2196/47349

KEYWORDS

burden; cataract; disability-adjusted life-years; human development index; prediction; risk factor

Introduction

With the progressive aging of the world's population, health issues, such as ocular health, have drawn more attention [1]. Cataracts are among the main causes of blindness globally, accounting for nearly half of all cases of blindness in low-income countries and 5% in high-income countries [2,3]. Recently, researchers have focused on the disability-adjusted life years (DALYs), which are defined as the sum of years lived with disability and years of life lost due to premature death and reflect the difference between the real state of health and the standard condition [4,5]. Moreover, DALYs attributed to cataracts exceed 90% in low-income countries and generally indicate the sum of years lived with disability, underscoring the significant impairments cataracts deal to the patients' quality of life [4]. As the global burden of cataracts continues to grow due to the rise in global population and aging, this phenomenon will inevitably increase the economic burden, decrease the quality of life, and lead to cognitive function impairments [3,5]. Therefore, investigating the trends in the global cataract burden and associated risk factors will allow more targeted prevention strategies to be initiated, thereby promoting targeted prevention and treatment of cataracts.

A trend analysis of the Global Burden of Disease (GBD) Study 2019 data showed that the global burden of visual impairment due to cataracts exhibited a substantial increase, with the prevalence rate rising by 58.45% and the DALY rate increasing by 32.18% over the past 30 years [6]. Cataract is a multifactorial disease manifested by transparency loss of the eye lens, owing to tissue degradation and protein clumping, as well as pathogenic environmental and genetic factors [7]. Previous publications have reported risk factors such as aging, female sex, diabetes, excess weight, diet and nutritional status, smoking, drinking, exposure to ultraviolet light, socioeconomic status, and air pollution [8-23]. In addition, cataracts can also be innate, owing to a variety of genetic and molecular mechanisms [7]. However, the extent to which these risk factors affect cataracts remains unclear. Additionally, the long-term global trend of DALYs due to cataracts has not been visually demonstrated and analyzed, necessitating further research.

To address these gaps, we conducted an evaluation of the global trend in the DALYs of cataract by year, age, and sex from 1990 to 2019 and estimated its prediction over the next 15 years. Additionally, we conducted correlation analyses between age-standardized DALY rates (ASDRs) and the socio-demographic index (SDI) and human development index (HDI) separately. We then estimated the proportion of DALYs that could be attributed to several important risk factors for cataracts, including air pollution, smoking, high fasting plasma glucose, and a high BMI. Finally, the association between HDI, particulate matter <2.5 µm in diameter ($PM_{2.5}$), and ultraviolet radiation (UVR) and ASDRs of cataracts was analyzed. Our results represented a significant extension of previous reports and offered valuable insights for future initiatives aimed at preventing cataracts.

Methods

Data Sources

This study was based on the GBD Study 2019, the United Nations Development Programme, and the WHO (World Health Organization) Global Health Observatory data repository. We collected related data on cataracts and SDI from the GBD database. The HDI data were obtained from the United Nations Development Programme, and concentrations of $PM_{2.5}$ and average daily UVR data were collected from the WHO Global Health Observatory data repository. We estimated the global population using data from the GBD database.

Study Variable Description

Study results are presented by number of DALYs, ASDRs, crude DALY rates, and average annual percentage changes (AAPCs) between 1990 and 2019 using 95% uncertainty intervals (UIs) and 95% CIs, age, sex, SDI, HDI, $PM_{2.5}$ concentration, and UVR. The SDI gauges a nation's economic development using per capita income distribution, average schooling years, and the fertility rate of women under 25 years of age [24]. The SDI ranges from 0 to 1, with higher values indicating greater economic advancement. The HDI was used as a measure of health, education, and living conditions [25]; it is calculated from life expectancy, education years, schooling expectancy, and per capita gross national income. HDI values range from 0 to 1, with higher values indicating greater socioeconomic development. The UNDP classifies countries into subgroups based on HDI: low (<0.550), moderate (0.550-0.699), high (0.700-0.799), and very high (≥0.800). The $PM_{2.5}$ concentration is a common measurement for air pollution in a country and is defined as the annual average concentration of suspended fine particulate matter smaller than 2.5 µm in diameter. The UVR was defined as the average daily level of ambient solar UVR (in J/m^2).

Statistical Analysis

Temporal trends in DALYs due to cataract were assessed using the Joinpoint regression software (version 4.9.0.0; National Cancer Institute). This model uses segmented regression to capture disease distribution patterns over time. For each juncture point, the annual percentage change (APC) was computed, accompanied by a 95% CI. Additionally, the AAPC was determined as the weighted mean of individual APCs, offering a consolidated overview of the study period's overall trend. In summary, a rising cataract rate was noted if both the APC and the lower 95% CI bound exceeded 0, while a declining rate was indicated if both values fell below 0. Linear regression analysis was used to determine the relationship between ASDR and SDI, HDI, national $PM_{2.5}$, and ambient UVR. To predict DALYs due to cataract from 2020 to 2035, we used the Nordpred age-period-cohort model implemented in the Nordpred package in R (version 4.2.2; R Foundation for Statistical Computing). To validate the stability of the prediction results, the Bayesian age-period-cohort model was further applied to perform a sensitivity analysis using the R package "BAPC." All statistical analyses and data visualization were conducted using R, the Joinpoint Regression Program, and GraphPad Prism 9.0

(GraphPad Software, Inc). A significance level of α=.05 was used, with *P*<.05 indicating statistical significance.

Ethical Considerations

The authors received an exemption from ethical review from the institutional review board of the First Affiliated Hospital of Soochow University, and the study was also exempt from ethical review by the National Health Research Ethics Council of China. The research adhered to the principles of the Declaration of Helsinki. Due to the retrospective nature of the study, the requirement for informed consent regarding data processing and disease models was waived.

Results

Temporal Trends of Disease Burden Due to Cataract From 1990 to 2019

Globally, the number of DALYs due to cataracts increased from 3,492,604 (95% UI 2,481,84-4,719,629) to 6,676,281.11 (95% UI 4,761,210-9,006,193) in 2019 (Table 1). The ASDR exhibited a dramatic decrease from 1990 to 1995 (APC=–0.63; *P*<.05), a dramatic increase from 1995 to 2000 (APC=1.05; *P*<.05), and another dramatic decrease from 2000 to 2010 (APC=–0.56; *P*<.05), followed by a slight increase from 2010 to 2017 (APC=0.25; *P*>.05). However, a remarkable decrease in ASDR was observed between 2017 and 2019 (APC=–3.21; *P*<.05) (Figure 1).

Table 1. Disability-adjusted life years (DALYs) of cataract and its average annual percentage changes from 1990 to 2019 at the global and regional levels.

Characteristics	1990		2019		1990-2019	*P* value
	Value (95% UI[a])	ASDR[b] (95% UI)	Number (95% UI)	ASDR (95% UI)	AAPC[c] (95% CI)	
Global	3,492,605 (2,481,846-4,719,629)	93.17 (66.14-125.32)	6,676,281 (4,761,211-9,006,194)	82.94 (59.06-111.75)	–0.37 (–0.44 to –0.3)	<.001
Sex						
Female	2,025,405 (1,442,842-2,729,748)	97.44 (69.34-130.79)	3,928,327 (2,798,531-5,275,330)	89.82 (63.97-120.6)	–0.25 (–0.33 to –0.16)	<.001
Male	1,467,199 (1,038,553-1,989,939)	88.46 (62.56-118.7)	2,747,954 (1,954,020-3,725,738)	74.91 (53.35-101.18)	–0.54 (–0.6 to –0.49)	<.001
SDI[d] region						
High SDI	226,706 (161,075-309,553)	21.86 (15.5-29.71)	407,175 (289,682-556,589)	20.57 (14.49-28.1)	–0.2 (–0.23 to –0.17)	<.001
High-middle SDI	527,854 (374,258-709,159)	54.05 (38.26-71.81)	1,079,895 (770,486-1,455,305)	53.72 (38.45-72.57)	0.03 (–0.1 to 0.15)	.687
Middle SDI	1,238,750 (878,931-1,687,529)	139.91 (99.06-188.58)	2,469,946 (1,747,303-3,330,111)	107.61 (75.96-144.28)	–0.87 (–0.95 to –0.79)	<.001
Low-middle SDI	1,126,634 (800,056-1,538,168)	216.76 (153.53-292.32)	2,009,356 (1,432,490-2,712,777)	160.43 (114.24-215.23)	–1 (–1.04 to –0.97)	<.001
Low SDI	370,745 (264,016-506,463)	180.68 (128.63-244.99)	706,519 (502,216-953,371)	153.38 (108.77-206.59)	–0.55 (–0.57 to –0.54)	<.001
South–East Asia, East Asia, and Oceania						
East Asia	453,089 (319,903-623,402)	63.9 (45-86.66)	1,097,095 (767,203-1,501,496)	57.37 (40.44-77.8)	–0.26 (–0.56 to –0.03)	.083
Southeast Asia	632,885 (447,942-866,995)	279.06 (198.51-378.15)	1,079,918 (764,186-1,467,494)	195.67 (138.4-262.38)	–1.22 (–1.24 to –1.2)	<.001
Oceania	4572 (3228-6213)	181.65 (128.83-242.25)	9854 (7084-13,440)	163.72 (116.76-222.47)	–0.38 (–0.45 to –0.31)	<.001
Sub-Saharan Africa						
Eastern Sub-Saharan Africa	110,011 (78,056-152,266)	156.81 (112.06-215.18)	195,926 (138,353-268,562)	129.15 (91.78-176.19)	–0.67 (–0.7 to –0.63)	<.001
Central Sub-Saharan Africa	5896 (4127-8269)	32.86 (22.99-45.35)	11,403 (7827-15,764)	26.13 (18.05-35.82)	–0.79 (–0.81 to –0.76)	<.001
Southern Sub-Saharan Africa	30,885 (21,678-42,767)	119.28 (84.16-164.15)	41,255 (29,230-56,630)	80.48 (57.65-110.15)	–1.35 (–1.38 to –1.33)	<.001
Western Sub-Saharan Africa	129,494 (91,237-176,192)	165.62 (117.94-224.66)	258,586 (182,989-352,960)	152.14 (107.71-205.16)	–0.3 (–0.34 to –0.26)	<.001
South Asia	1,303,763 (924,771-1,774,799)	273.27 (194.6-366.72)	2,544,708 (1,812,812-3,422,950)	198.39 (142.54-264.15)	–1.07 (–1.09 to –1.04)	<.001
Latin America and Caribbean						
Caribbean	13,961 (9858-19,244)	56.3 (39.63-76.84)	22,146 (15,531-30,169)	42.9 (30.02-58.37)	–0.93 (–0.94 to –0.93)	<.001
Central Latin America	67,300 (47,685-90,784)	90.64 (64.33-121.65)	146,616 (104,295-197,476)	64.7 (46-87.16)	–1.17 (–1.2 to –1.14)	<.001
Tropical Latin America	76,104 (54,205-104,197)	96.82 (68.48-130.47)	164,285 (117,096-218,938)	70.86 (50.45-94.07)	–1.06 (–1.13 to –0.99)	<.001

Characteristics	1990		2019		1990-2019	*P* value
	Value (95% UI[a])	ASDR[b] (95% UI)	Number (95% UI)	ASDR (95% UI)	AAPC[c] (95% CI)	
Andean Latin America	26,053 (18,322-35,146)	140.25 (99-188.37)	52,092 (36,890-69,969)	96.22 (68.28-129.43)	–1.31 (–1.45 to –1.17)	<.001
North Africa and Middle East	211,075 (149,404-288,143)	143.15 (101.05-195.59)	376,380 (264,296-513,801)	98.21 (69.44-134.06)	–1.3 (–1.32 to –1.27)	<.001
Central Europe, Eastern Europe, and Central Asia						
Central Asia	32,930 (23,072-44,833)	77.49 (54.76-104.16)	41,305 (28,608-56,756)	66.22 (45.77-89.99)	–0.54 (–0.57 to –0.51)	<.001
Eastern Europe	86,076 (60,611-115,280)	33.49 (23.53-44.81)	100,478 (70,300-136,392)	28.94 (20.29-39.35)	–0.5 (–0.52 to –0.48)	<.001
Central Europe	28,961 (19,939-39,952)	21.26 (14.81-28.96)	42,190 (29,051-58,232)	19.36 (13.34-26.63)	–0.32 (–0.33 to –0.31)	<.001
High-income regions						
Southern Latin America	16,633 (11,565-22,343)	38.56 (26.81-51.96)	27,402 (19,214-37,237)	32.5 (22.74-44.1)	–0.58 (–0.61 to –0.55)	<.001
Western Europe	153,209 (107,373-210,703)	26.68 (18.64-36.59)	246,102 (172,914-338,143)	25.31 (17.75-34.81)	–0.18 (–0.2 to –0.16)	<.001
North America	65,177 (46,278-88,463)	18.06 (12.84-24.71)	112,884 (80,414-152,299)	17.43 (12.34-23.78)	–0.12 (–0.15 to –0.08)	<.001
Australasia	4791 (3315-6506)	21.32 (14.83-29.03)	10,416 (7336-14,137)	20.3 (14.25-27.82)	–0.17 (–0.21 to –0.13)	<.001
Asia Pacific	39,741 (28,017-54,026)	21.09 (14.83-28.65)	95,238 (67,107-130,641)	19.5 (13.59-26.89)	–0.27 (–0.29 to –0.25)	<.001

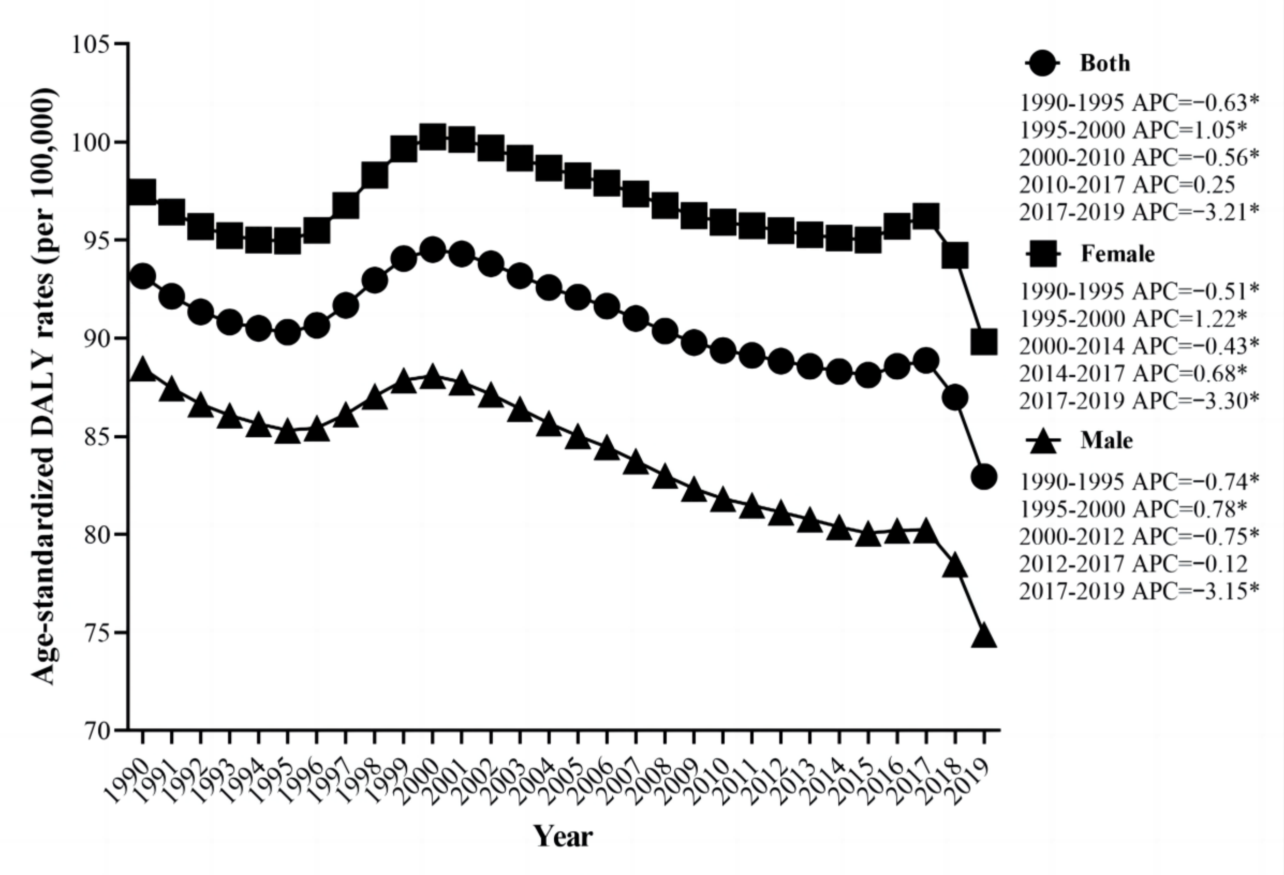
[a] UI: uncertainty interval.

[b] ASDR: age-standardized DALY rate.

[c] AAPC: average annual percent change.

[d] SDI: socio-demographic index.

Figure 1. Joinpoint regression analysis of global age-standardized disability-adjusted life-years (DALYs) rates due to cataract for both female and male from 1990 to 2019. **P*<.05; APC: annual percentage change.



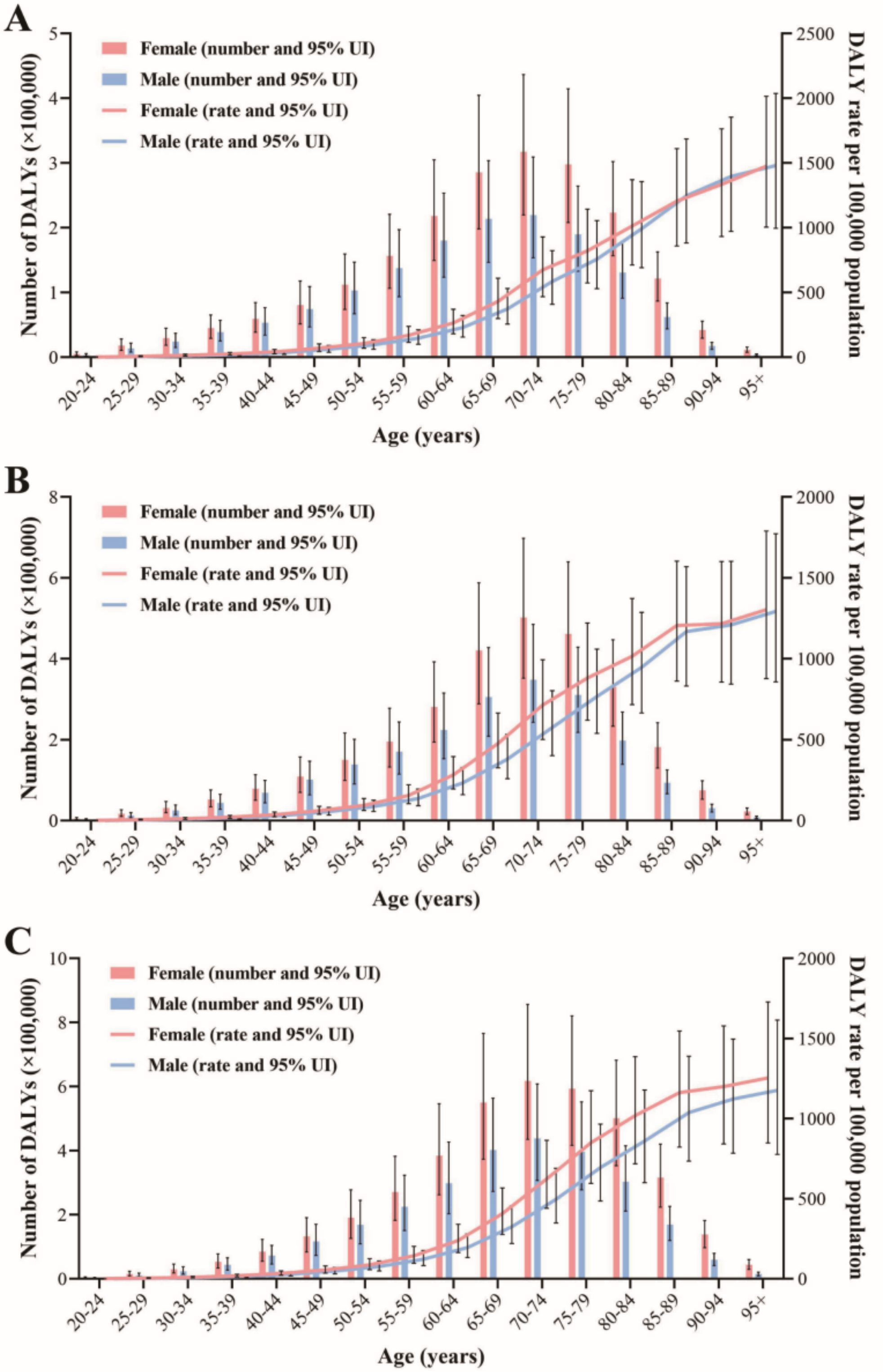
Regional Differences in Disease Burden Due to Cataract

At the regional level in 1990 and 2019, the number of DALYs due to cataracts was highest in South Asia and lowest in Oceania (Table 1). The ASDR was highest in Southeast Asia in 1990 and highest in South Asia in 2019. From 1990 to 2019, the AAPC of ASDR in Southern Sub-Saharan Africa was lowest at –1.35 (95% CI –1.38 to –1.33; *P*<.001), followed by Andean Latin America at –1.31 (95% CI –1.45 to –1.17; *P*<.001). In contrast, the AAPC of ASDR in North America was highest at –0.12 (95% CI –0.15 to –0.08; *P*<.001).

Age-Specific and Gender-Specific Disease Burden Due to Cataract

With increasing age, the global number of DALYs due to cataracts gradually increased (up to the age of 75 years) in 1990, reaching a peak at the ages of 70-74 years, after which it declined with increasing age. Similar trends were observed in 2005 and 2019. The DALYs due to cataracts were relatively low in individuals aged less than 30 years or more than 95 years. The crude DALY rates of cataracts rose with increasing age, with the fastest increment observed between the ages of 60 and 85 years in 1990. In 2005 and 2019, the general trend of the crude DALY rates was similar to 1990; however, the rate of increase after the age of 85 years slowed down more significantly (Figure 2). The global number of DALYs due to cataracts was lower in males than in females in 1990 and 2019, while ASDR was also smaller in males (Table 1). Moreover, the ASDR of male cataracts decreased more than that of females between 1990 and 2019, with AAPCs of –0.54 (95% CI –0.60 to –0.49; *P*<.001) and –0.25 (95% CI –0.33 to –0.16; *P*<.001), respectively (Table 1 and Figure 1).

Figure 2. Global number of disability-adjusted life years (DALYs) and crude DALY rates due to cataract by age and sex in (A) 1990, (B) 2005, and (C) 2019. The black bar indicates the upper and lower limits of the 95% uncertainty intervals (UIs) for females and males, respectively.



Characteristics of the Cataract Burden Attributable to Risk Factors

Areas with high SDI have lower ASDRs for cataracts than those with low SDI, with high-income North America having the lowest ASDRs and South Asia having the highest between 1990 and 2019 (Table 1 and Figure 3A). In addition, the ASDRs of cataracts were lower in countries with a high SDI than in those with a low SDI. Linear regression analysis revealed that ASDRs were negatively correlated with SDI (R^2=0.1939; P<.001; Figure 3B). ASDRs of cataracts due to air pollution, smoking, high fasting plasma glucose, and high BMI in 21 GBD regions in 2019 were obtained and presented in Figure 4. Apart from high SDI regions, air pollution accounted for the largest proportion of the 4 listed factors leading to cataracts, which increased as the SDI declined. The scatter plots between ASDRs of cataract and other country-level indicators were presented in Multimedia Appendix 1, and a linear trend was observed between ASDRs of cataract and HDI level (R^2=0.2828; P<.001), national $PM_{2.5}$ concentration (R^2=0.1874; P<.001), and ambient UVR (R^2=0.2354; P<.001). Based on the aforementioned results, HDI, national $PM_{2.5}$ concentration, and ambient UVR were included as independent variables to construct the multivariate linear regression model for ASDRs. A significant relationship was identified (P<.001), and this model accounted for 37.6% of the variance in the national differences in cataract burden.

Figure 3. Age-standardized disability-adjusted life year (DALY) rate (per 100,000 persons) of cataract by socio-demographic index (SDI) for (A) 21 Global Burden of Disease (GBD) Study regions from 1990 to 2019, and (B) 204 countries and territories in 2019. The black line represents the expected values based on age-standardized DALY rates and SDI in all locations.

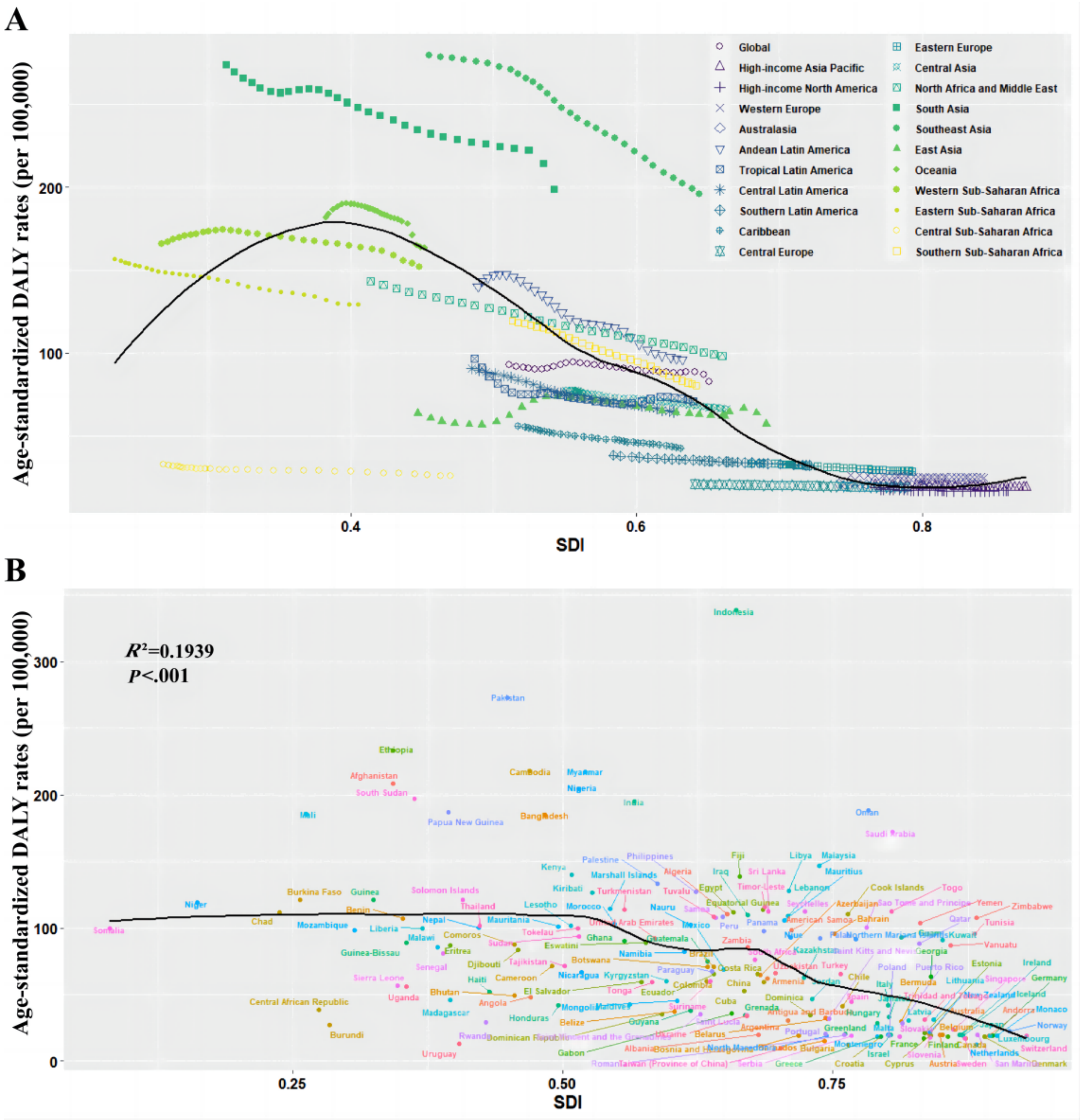
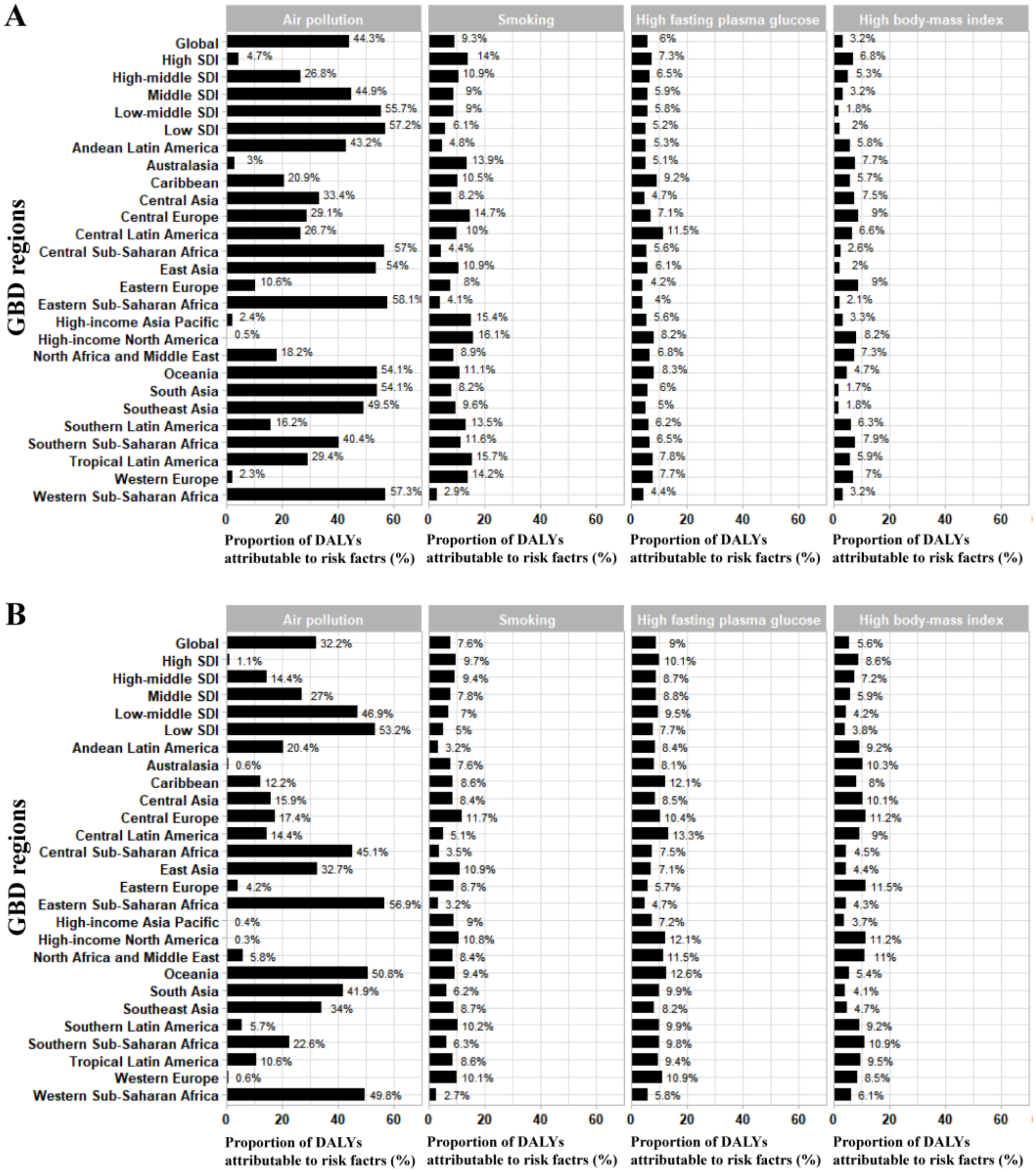


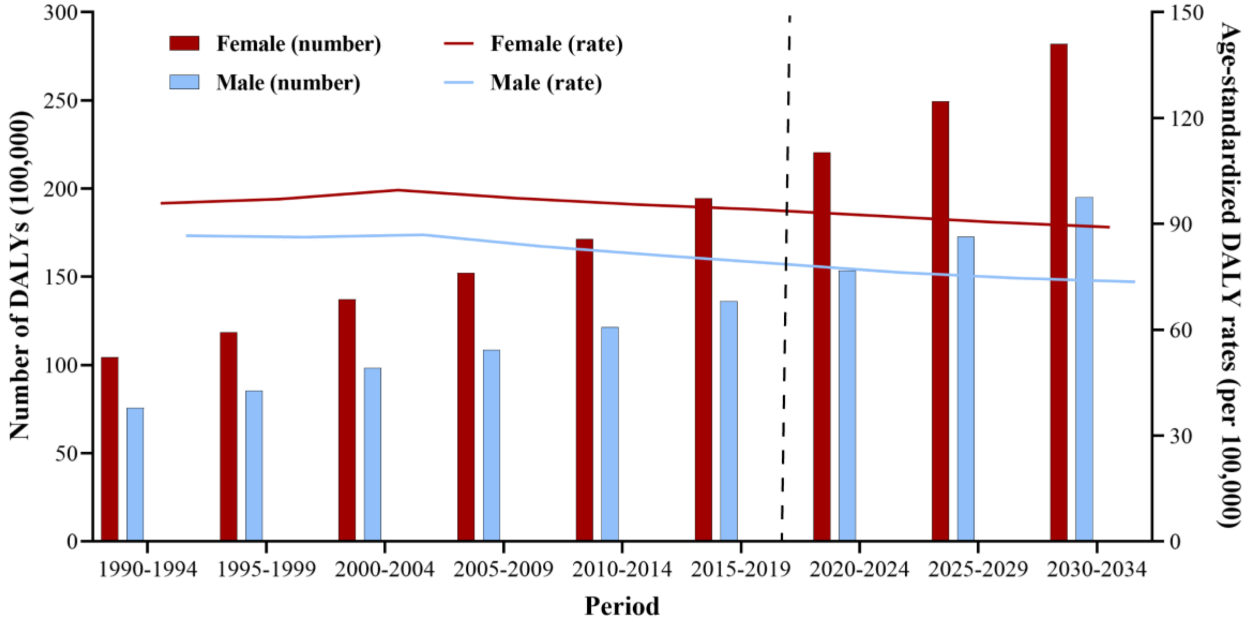
Figure 4. The proportion of cataract disability-adjusted life year (DALY) attributable to air pollution, smoking, high fasting plasma glucose, and high body-mass index for (A) 21 Global Burden of Disease (GBD) Study regions in 1990 and (B) 2019. SDI: socio-demographic index.



Predictions of DALYs Due to Cataract

We predicted that, due to population growth and aging, the DALYs due to cataracts for both sexes are expected to continue to increase over the next 15 years. Further, over the next 15 years, a decrease in the ASDRs of cataracts is expected, particularly among women (Figure 5 and Multimedia Appendix 2). While the patterns of changes in DALY rates in men were similar to those in women, the DALY rates of cataracts were notably lower among men. To assess the robustness of our predictions, we further performed a sensitivity analysis using the BAPC package to predict the future global burden of cataracts; the results showed consistent trends with the abovementioned results (Multimedia Appendix 3).

Figure 5. Trends in global number of disability-adjusted life years (DALYs) and age-standardized DALY rates (per 100,000 persons) due to cataract by sex: observed (1990-2019) and predicted (2020-2034).



Discussion

Principal Findings

This study reported on the number of DALYs and ASDRs due to cataracts from 1990 to 2019, revealing disparities by region, nation, age, and sex. Although the ASDRs decreased overall, the overall DALYs due to cataracts still increased. Our findings showed that over the last 3 decades, higher DALY rates were also associated with lower SDI, HDI, national $PM_{2.5}$, and ambient UVR. Furthermore, we predicted that the global burden of cataracts will continue to increase over the next 15 years due to the growth and aging of the global population. To the best of our knowledge, this report acts as an essential complement to previous studies, and the results of our research can provide crucial insights into evidence-based health care planning and reduce the global burden of cataracts by allocating resources to cataract prevention.

Global policy making and economic restructuring can guide the prevention and treatment of cataract. Initially, the WHO and International Agency for the Prevention of Blindness started the VISION 2020 global initiative in 1999 to prevent, control, and eliminate avoidable blindness at the national level by 2020 [26]. In line with earlier studies, the number of DALYs in cataracts has increased globally over the last 30 years, although the ASDRs fell between 1990 and 2019 [6,10,20]. According to our findings, low-income regions exhibited higher ASDRs compared to high-income regions, indicating a greater burden of cataract in the former. Several factors could potentially explain this pattern, including elevated UVR levels in southern countries and the prevalence of conditions such as arterial hypertension and diabetes, which might contribute to the proliferation of this ailment. Notably, in response to VISION 2020, high-income regions allocated funding for diverse programs aimed at reducing rates of low vision and blindness. The adoption of more proactive diagnostics, early detection, and treatment strategies could potentially bolster the performance of high-income regions [27]. Addressing this phenomenon warrants increased attention and further investigations into the mechanisms underlying these disparities.

Although previous studies have implicated many risk factors for cataracts, age is uniquely important. Cataracts can be divided by age-related, pediatric, and secondary causes. Age-related cataract is the most common type in adults, starting between the ages of 45 and 50 years [28]. Aging is the most meaningful risk factor for cataract because structural and functional changes to the lens can promote cataract development. The lens consists of soluble proteins, and the fiber cell organization in the lens is tight, which enables intercellular space reduction. These mechanisms prevent light content and support lens transparency. Increasing age leads to slower protein metabolism, reducing their solubilization, which in turn impairs lens transparency and may cause other pathological risk factors to have cumulatively detrimental effects [29,30].

According to our findings, both the number of DALYs and ASDRs caused by cataracts in females was greater than that in males from 1990 to 2019. The following is consistent with other studies that found similar situations of sex disparity in ocular diseases, such as cataracts, age-related macular degeneration, and diabetic retinopathy [5,21,30,31]. Possible explanations for these results include differences in follow-up time points and longer life expectancies in female patients [10]. Another explanation could be related to the inequality between men and women, as women can be disadvantaged in terms of access to education, employment opportunities, income distribution, and health care services [32]. Other research has also focused on the function of estrogen in cataract formation, which has been shown to play a role due to antioxidant effects on reactive oxygen species, which are important factors for cataracts. Menopause is hypothesized to cause an elevated risk of cataract in older females, caused by the withdrawal effect [9]. More attention and further studies regarding female eye care are required to explore the possible mechanisms between sex and cataract.

Furthermore, health inequalities are largely caused by socioeconomic disparities. According to our findings, the ASDRs of cataracts varied significantly between regions and nations and were lower in high-income regions and countries, while SDI and HDI were negatively correlated with cataract ASDR. A higher quality of eye care and increasing opportunities for cataract surgery may explain this association in countries with high SDI and HDI scores. Although surgery is the definitive therapy for cataract, many major barriers to cataract treatment still exist [33]. One aspect is that the outcome of cataract surgery is related to the quality of the surgery. Access to quality cataract surgery remains difficult in many countries, while the coverage of cataract surgery is influenced by GDP and health expenditure [18,20]. The lack of education and medical knowledge, as well as the cost of surgery, may discourage patients in low-income areas from choosing treatment [19]. As a result, the cataract surgical rate is over 10,000 in developed countries and below 500 in low-income countries [34]. Thus, in low-income countries, the emphasis on reducing the burden of cataracts in the formation of health policies should be as important as in high-income countries.

We also found that air pollution, smoking, high fasting plasma glucose, and high BMI were all significant cataract risk factors that contributed to most of the burden. In this study, air pollution referred mainly to household air pollution from solid fuels, which was associated with a higher risk of cataracts in agreement with previous studies [35,36]. Smoking is a risk factor for many diseases, including cataracts, glaucoma, and age-related macular degeneration, with studies reporting an elevated risk of nuclear sclerotic cataract in patients of both sexes [37-39]. Tobacco contains large quantities of toxic heavy metals, including cadmium, lead, and copper, and the accumulation of these metals on the lens can eventually contribute to cataract development [15,40,41]. More specifically, the ratio of the indirect impact of tobacco to the total impact of cadmium exceeded 50% [42]. This study also indicated that high fasting blood glucose levels were correlated with cataracts. The Blue Mountains Eye Study recently demonstrated that fasting blood glucose levels were associated with long-term cortical cataract formation and long-term progression of cortical, nuclear, and posterior subcapsular cataracts, which is consistent with our findings [43]. Similarly, the Beaver Dam Eye Study evaluated diabetic participants as having an increased risk of developing cataracts [44]. The study proposed strategies to address the increased risk of cataracts in diabetic patients, which include targeting the polyol pathway, oxidative stress, and nonenzymatic glycation of lens proteins [45,46]. Moreover, a study indicated that diabetic retinopathy is an independent risk factor for cataract formation in patients with type 2 diabetes, alongside BMI, HbA_{1c}, and insulin usage [12]. The Blue Mountain Eye Study also showed an increased risk of cataract in relation to metabolic syndrome [47]. Therefore, obesity is correlated with increased DALY rates in cataract. Obesity, especially central obesity, is associated with insulin resistance, which indirectly leads to type 2 diabetes [48]. In addition, the lipid composition of the lens changes during cataract formation, possibly because high BMI and excess body weight alter lipid metabolism in the lens [14]. The results of this study may have implications for public health, as cataracts are susceptible to air pollution, smoking, high fasting plasma glucose, and high BMI, and increased health awareness may be a cost-effective intervention to prevent blindness in patients with cataract.

This study illustrated that exposure to $PM_{2.5}$ and solar UVR accelerated cataract development. Current evidence suggests that air pollution may affect ocular diseases, including cataracts, age-related macular degeneration, and glaucoma [22]. After exposure to air pollutants such as $PM_{2.5}$, NO_2 (nitrogen dioxide), and NO_x (nitrogen oxide), the risk of cataract surgery increased by 5%, but the effects were relatively small [49]. The most common assessment of air pollution is $PM_{2.5}$. The Canadian Longitudinal Study on Aging illustrated that $PM_{2.5}$ had a borderline-adjusted correlation with cataracts [23]. Evidence suggests that exposure to high-dose UVR can cause acute photokeratitis and photoconjunctivitis while chronic exposure to low-dose UVR can cause cataract, pterygium, and corneal and conjunctival squamous cell carcinoma [50]. Interestingly, cataract DALY rates were lower in urbanized areas because the work of most individuals could be completed indoors, and their UVR exposure could be at a lower level [51]. Evidence suggests that oxidation accumulation, antioxidant potential, and repair function of the lens decrease with age, which are all related to age-related cataracts. The most significant cause of cataract may be the formation of reactive oxygen species caused by UVR. This decreased the activity of a series of enzymes that catalyze the reduction of hydrogen peroxide and peroxide radicals [50]. Our results indicate that improving the living environment and increasing UV radiation protection may be effective interventions to delay the occurrence of cataract blindness, while the pathogenic roles of $PM_{2.5}$ and UV light in cataracts require further study.

While the DALY rates for cataract decreased overall in most age and sex groups between 1990 and 2019, there was a significant increase in the total number of DALYs due to cataract. This upward trend in cataract DALYs, coupled with an aging population and expanding economy, is likely to lead to a greater economic burden on health care. We anticipate that the number of DALYs due to cataract will continue to rise steadily from 2020 to 2034. In light of the anticipated increase in the global burden of cataract, efforts should be made to strengthen prevention and management strategies for both established and emerging risk factors such as air pollution, smoking, and UV radiation. Effective global policies aimed at reducing the prevalence of these risk factors could play a critical role in mitigating the impact of cataract and reduce its burden.

Study Limitations

Although this study enabled us to examine the global cataract burden and its relationship to risk factors with high confidence, some limitations should be noted. First, all data come from the GBD 2019, UNDP, and GHO. As such, the authenticity and reliability of the results depend on the data source. Second, cataract is a multifactorial disease, and some key factors, such as congenital, traumatic, and drug-induced factors, have not been analyzed in this study; therefore, the factors included in our model may also be affected. Third, the prevalence and incidence of cataracts were not investigated in this study. Moreover, this study did not conduct a decomposition analysis of the disease burden of cataract over time. This omission indicates a potential gap in understanding the complex dynamics underlying the changing disease burden of cataract. Therefore, it is essential to explore and gather more information on cataract to improve our understanding of this important public health issue.

Conclusions

Over the past 3 decades, the global burden of cataract has increased, as evidenced by the increasing number of DALYs. Furthermore, given the changing demographics of the world, this trend is expected to persist over the coming 15 years. Age, female sex, air pollution, smoking, high fasting plasma glucose, high BMI, UVR, and low socioeconomic status were risk factors for the burden of cataracts. Our findings showed that the older adults and female populations in low-income countries have a greater burden caused by cataract; therefore, health services for cataract should be strengthened for these populations. Effective resource allocation and health-service planning will be essential in managing the anticipated increase in the global burden of cataract in the next 15 years.

Acknowledgments

We highly appreciate the work done by the Global Burden of Disease, the United Nations Development Programme, and the WHO (World Health Organization) Global Health Observatory. This work was supported by the Jiangsu Provincial Medical Innovation Team (CXTDA2017039) and the National Natural Science Foundation in China (81671641, 82271113, and 81970830).

Data Availability

The data analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

PRL and BJ had full access to all the data in the study and took responsibility for the integrity of the data, the accuracy of the data analysis, and the concept and design of the study. All authors were responsible for the acquisition, analysis, or interpretation of data. PRL, BJ, and THW drafted the manuscript. PRL, GQL, and BJ were responsible for the critical revision of the manuscript for important intellectual content. BJ did the statistical analysis. PRL and GQL obtained the funding. PRL was responsible for administrative, technical, or material support and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Associations between age-standardized disability-adjusted life-year (DALY) rates (per 100,000 persons) of cataract and HDI (A), national PM2.5 (B), and ambient ultraviolet radiation (C) in 2019. HDI: human development index; PM2.5: particle matter <2.5 μm in diameter.

[PNG File , 124 KB - publichealth_v9i1e47349_app1.png]

Multimedia Appendix 2

Trends in global number of disability-adjusted life-years (DALYs) and age-standardized DALY rates (per 100,000 persons) due to cataract by sex: observed (1990-2019) and predicted (2020-2034).

[DOC File , 32 KB - publichealth_v9i1e47349_app2.doc]

Multimedia Appendix 3

Trends in global number of disability-adjusted life-years (DALYs) and age-standardized DALY rates (per 100,000 persons) due to cataract by sex: observed (1990-2019) and predicted (2020-2034).

[PNG File , 120 KB - publichealth_v9i1e47349_app3.png]

References

1. Singh RR, Maurya P. Visual impairment and falls among older adults and elderly: evidence from longitudinal study of ageing in India. BMC Public Health 2022;22(1):2324 [FREE Full text] [doi: 10.1186/s12889-022-14697-2] [Medline: 36510173]

2. Samuel M, Abdulkadir H, Girma M, Glagn M. Assessment of knowledge and attitude of cataract and their associated factors among adults in Arba Minch Zuria Woreda, Southern Ethiopia. Clin Ophthalmol 2021;15:2913-2920 [FREE Full text] [doi: 10.2147/OPTH.S320873] [Medline: 34262251]
3. Ang MJ, Afshari NA. Cataract and systemic disease: a review. Clin Exp Ophthalmol 2021;49(2):118-127. [doi: 10.1111/ceo.13892] [Medline: 33426783]
4. Ono K, Hiratsuka Y, Murakami A. Global inequality in eye health: country-level analysis from the global burden of disease study. Am J Public Health 2010;100(9):1784-1788 [FREE Full text] [doi: 10.2105/AJPH.2009.187930] [Medline: 20634443]
5. Xu X, Wu J, Yu X, Tang Y, Tang X, Shentu X. Regional differences in the global burden of age-related macular degeneration. BMC Public Health 2020;20(1):410 [FREE Full text] [doi: 10.1186/s12889-020-8445-y] [Medline: 32228540]
6. Han X, Zou M, Liu Z, Sun Y, Young CA, Zheng D, et al. Time trends and heterogeneity in the disease burden of visual impairment due to cataract, 1990-2019: a global analysis. Front Public Health 2023;11:1140533 [FREE Full text] [doi: 10.3389/fpubh.2023.1140533] [Medline: 37077196]
7. Shiels A, Hejtmancik JF. Mutations and mechanisms in congenital and age-related cataracts. Exp Eye Res 2017;156:95-102 [FREE Full text] [doi: 10.1016/j.exer.2016.06.011] [Medline: 27334249]
8. Agrawal S, Fledderjohann J, Ghosh S. Risk factors for self-reported cataract symptoms, diagnosis, and surgery uptake among older adults in India: findings from the WHO SAGE data. Glob Public Health 2021;16(11):1771-1785. [doi: 10.1080/17441692.2020.1836246] [Medline: 33091324]
9. Zetterberg M, Celojevic D. Gender and cataract: the role of estrogen. Curr Eye Res 2015;40(2):176-190. [doi: 10.3109/02713683.2014.898774] [Medline: 24987869]
10. Lou L, Ye X, Xu P, Wang J, Xu Y, Jin K, et al. Association of sex with the global burden of cataract. JAMA Ophthalmol 2018;136(2):116-121 [FREE Full text] [doi: 10.1001/jamaophthalmol.2017.5668] [Medline: 29242928]
11. Fang R, Yu YF, Li EJ, Lv NX, Liu ZC, Zhou HG, et al. Global, regional, national burden and gender disparity of cataract: findings from the global burden of disease study 2019. BMC Public Health 2022;22(1):2068 [FREE Full text] [doi: 10.1186/s12889-022-14491-0] [Medline: 36369026]
12. Alabdulwahhab KM. Senile cataract in patients with diabetes with and without diabetic retinopathy: a community-based comparative study. J Epidemiol Glob Health 2022;12(1):56-63 [FREE Full text] [doi: 10.1007/s44197-021-00020-6] [Medline: 34874549]
13. Jacques PF, Moeller SM, Hankinson SE, Chylack LT, Rogers G, Tung W, et al. Weight status, abdominal adiposity, diabetes, and early age-related lens opacities. Am J Clin Nutr 2003;78(3):400-405 [FREE Full text] [doi: 10.1093/ajcn/78.3.400] [Medline: 12936921]
14. Sella R, Afshari NA. Nutritional effect on age-related cataract formation and progression. Curr Opin Ophthalmol 2019;30(1):63-69. [doi: 10.1097/ICU.0000000000000537] [Medline: 30320615]
15. Lehrer S, Rheinstein PH. Marijuana smoking and cataract. J Fr Ophtalmol 2022;45(3):267-271 [FREE Full text] [doi: 10.1016/j.jfo.2021.12.008] [Medline: 35093261]
16. Lindblad BE, Håkansson N, Philipson B, Wolk A. Alcohol consumption and risk of cataract extraction: a prospective cohort study of women. Ophthalmology 2007;114(4):680-685. [doi: 10.1016/j.ophtha.2006.07.046] [Medline: 17173974]
17. Löfgren S. Solar ultraviolet radiation cataract. Exp Eye Res 2017;156:112-116. [doi: 10.1016/j.exer.2016.05.026] [Medline: 27260484]
18. Deng Y, Yang D, Yu JM, Xu JX, Hua H, Chen RT, et al. The association of socioeconomic status with the burden of cataract-related blindness and the effect of ultraviolet radiation exposure: an ecological study. Biomed Environ Sci 2021;34(2):101-109 [FREE Full text] [doi: 10.3967/bes2021.015] [Medline: 33685568]
19. Sonron EA, Tripathi V, Hariharan S. The impact of sociodemographic and socioeconomic factors on the burden of cataract in Small Island Developing States (SIDS) in the Caribbean from 1990 to 2016. Ophthalmic Epidemiol 2020;27(2):132-140. [doi: 10.1080/09286586.2019.1700534] [Medline: 31818167]
20. Lou L, Wang J, Xu P, Ye X, Ye J. Socioeconomic disparity in global burden of cataract: an analysis for 2013 with time trends since 1990. Am J Ophthalmol 2017;180:91-96. [doi: 10.1016/j.ajo.2017.04.008] [Medline: 28428050]
21. He M, Wang W, Huang W. Variations and trends in health burden of visual impairment due to cataract: a global analysis. Invest Ophthalmol Vis Sci 2017;58(10):4299-4306 [FREE Full text] [doi: 10.1167/iovs.17-21459] [Medline: 28846778]
22. Grant A, Leung G, Freeman EE. Ambient air pollution and age-related eye disease: a systematic review and meta-analysis. Invest Ophthalmol Vis Sci 2022;63(9):17 [FREE Full text] [doi: 10.1167/iovs.63.9.17] [Medline: 35960515]
23. Grant A, Leung G, Aubin MJ, Kergoat MJ, Li G, Freeman EE. Fine particulate matter and age-related eye disease: the Canadian longitudinal study on aging. Invest Ophthalmol Vis Sci 2021;62(10):7 [FREE Full text] [doi: 10.1167/iovs.62.10.7] [Medline: 34369984]
24. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019. Lancet 2020;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
25. Wang W, He M, Li Z, Huang W. Epidemiological variations and trends in health burden of glaucoma worldwide. Acta Ophthalmol 2019;97(3):e349-e355 [FREE Full text] [doi: 10.1111/aos.14044] [Medline: 30790460]

26. Bourne RRA. Vision 2020: where are we? Curr Opin Ophthalmol 2020;31(2):81-84. [doi: 10.1097/ICU.0000000000000647] [Medline: 31904597]
27. Jiang B, Jiang C, Li J, Lu P. Trends and disparities in disease burden of age-related macular degeneration from 1990 to 2019: results from the global burden of disease study 2019. Front Public Health 2023;11:1138428 [FREE Full text] [doi: 10.3389/fpubh.2023.1138428] [Medline: 37265519]
28. Liu YC, Wilkins M, Kim T, Malyugin B, Mehta JS. Cataracts. Lancet 2017;390(10094):600-612. [doi: 10.1016/S0140-6736(17)30544-5] [Medline: 28242111]
29. Budnar P, Tangirala R, Bakthisaran R, Rao CM. Protein aggregation and cataract: role of age-related modifications and mutations in α-crystallins. Biochemistry (Mosc) 2022;87(3):225-241 [FREE Full text] [doi: 10.1134/S000629792203004X] [Medline: 35526854]
30. Hejtmancik JF, Kantorow M. Molecular genetics of age-related cataract. Exp Eye Res 2004;79(1):3-9 [FREE Full text] [doi: 10.1016/j.exer.2004.03.014] [Medline: 15183095]
31. Xu Y, Wang A, Lin X, Xu J, Shan Y, Pan X, et al. Global burden and gender disparity of vision loss associated with diabetes retinopathy. Acta Ophthalmol 2021;99(4):431-440 [FREE Full text] [doi: 10.1111/aos.14644] [Medline: 33124190]
32. Courtright P, Lewallen S. Why are we addressing gender issues in vision loss? Community Eye Health 2009;22(70):17-19 [FREE Full text] [Medline: 19888362]
33. Upaphong P, Thonusin C, Choovuthayakorn J, Chattipakorn N, Chattipakorn SC. The possible positive mechanisms of pirenoxine in cataract formation. Int J Mol Sci 2022;23(16):9431 [FREE Full text] [doi: 10.3390/ijms23169431] [Medline: 36012695]
34. Burton MJ, Ramke J, Marques AP, Bourne RRA, Congdon N, Jones I, et al. The Lancet global health commission on global eye health: vision beyond 2020. Lancet Glob Health 2021;9(4):e489-e551 [FREE Full text] [doi: 10.1016/S2214-109X(20)30488-5] [Medline: 33607016]
35. Chan KH, Yan M, Bennett DA, Guo Y, Chen Y, Yang L, et al. Long-term solid fuel use and risks of major eye diseases in China: a population-based cohort study of 486,532 adults. PLoS Med 2021;18(7):e1003716 [FREE Full text] [doi: 10.1371/journal.pmed.1003716] [Medline: 34324491]
36. Patel M, Shrestha MK, Manandhar A, Gurung R, Sadhra S, Cusack R, et al. Effect of exposure to biomass smoke from cooking fuel types and eye disorders in women from hilly and plain regions of Nepal. Br J Ophthalmol 2022;106(1):141-148. [doi: 10.1136/bjophthalmol-2020-316766] [Medline: 32920527]
37. Kelly SP, Thornton J, Edwards R, Sahu A, Harrison R. Smoking and cataract: review of causal association. J Cataract Refract Surg 2005;31(12):2395-2404. [doi: 10.1016/j.jcrs.2005.06.039] [Medline: 16473237]
38. Perez CI, Singh K, Lin S. Relationship of lifestyle, exercise, and nutrition with glaucoma. Curr Opin Ophthalmol 2019;30(2):82-88. [doi: 10.1097/ICU.0000000000000553] [Medline: 30562241]
39. Vyas A, Raman S, Surya J, Sen S, Raman R. The need for artificial intelligence based risk factor analysis for age-related macular degeneration: a review. Diagnostics (Basel) 2022;13(1):130 [FREE Full text] [doi: 10.3390/diagnostics13010130] [Medline: 36611422]
40. Nam GE, Han K, Ha SG, Han BD, Kim DH, Kim YH, et al. Relationship between socioeconomic and lifestyle factors and cataracts in Koreans: the Korea national health and nutrition examination survey 2008-2011. Eye (Lond) 2015;29(7):913-920 [FREE Full text] [doi: 10.1038/eye.2015.66] [Medline: 25976646]
41. Prokofyeva E, Wegener A, Zrenner E. Cataract prevalence and prevention in Europe: a literature review. Acta Ophthalmol 2013;91(5):395-405 [FREE Full text] [doi: 10.1111/j.1755-3768.2012.02444.x] [Medline: 22715900]
42. Wang W, Schaumberg DA, Park SK. Cadmium and lead exposure and risk of cataract surgery in U.S. adults. Int J Hyg Environ Health 2016;219(8):850-856 [FREE Full text] [doi: 10.1016/j.ijheh.2016.07.012] [Medline: 27460785]
43. Kanthan GL, Mitchell P, Burlutsky G, Wang JJ. Fasting blood glucose levels and the long-term incidence and progression of cataract: the Blue Mountains eye study. Acta Ophthalmol 2011;89(5):e434-e438 [FREE Full text] [doi: 10.1111/j.1755-3768.2011.02149.x] [Medline: 21457484]
44. Klein BE, Klein R, Lee KE. Diabetes, cardiovascular disease, selected cardiovascular disease risk factors, and the 5-year incidence of age-related cataract and progression of lens opacities: the Beaver Dam eye study. Am J Ophthalmol 1998;126(6):782-790. [doi: 10.1016/s0002-9394(98)00280-3] [Medline: 9860001]
45. Javadi MA, Zarei-Ghanavati S. Cataracts in diabetic patients: a review article. J Ophthalmic Vis Res 2008;3(1):52-65 [FREE Full text] [Medline: 23479523]
46. Kyselova Z, Stefek M, Bauer V. Pharmacological prevention of diabetic cataract. J Diabetes Complications 2004;18(2):129-140. [doi: 10.1016/S1056-8727(03)00009-6] [Medline: 15120709]
47. Tan JSL, Wang JJ, Mitchell P. Influence of diabetes and cardiovascular disease on the long-term incidence of cataract: the Blue Mountains eye study. Ophthalmic Epidemiol 2008;15(5):317-327. [doi: 10.1080/09286580802105806] [Medline: 18850468]
48. Yazıcı D, Sezer H. Insulin resistance, obesity and lipotoxicity. Adv Exp Med Biol 2017;960:277-304. [doi: 10.1007/978-3-319-48382-5_12] [Medline: 28585204]

49. Chua SYL, Khawaja AP, Desai P, Rahi JS, Day AC, Hammond CJ, et al. The association of ambient air pollution with cataract surgery in UK Biobank participants: prospective cohort study. Invest Ophthalmol Vis Sci 2021;62(15):7 [FREE Full text] [doi: 10.1167/iovs.62.15.7] [Medline: 34874411]
50. Ivanov IV, Mappes T, Schaupp P, Lappe C, Wahl S. Ultraviolet radiation oxidative stress affects eye health. J Biophotonics 2018;11(7):e201700377 [FREE Full text] [doi: 10.1002/jbio.201700377] [Medline: 29603665]
51. Zhu M, Yu J, Gao Q, Wang Y, Hu L, Zheng Y, et al. The relationship between disability-adjusted life years of cataracts and ambient erythemal ultraviolet radiation in China. J Epidemiol 2015;25(1):57-65 [FREE Full text] [doi: 10.2188/jea.JE20140017] [Medline: 25263793]

Abbreviations

AAPC: average annual percentage change
APC: annual percentage change
ASDR: age-standardized DALY rate
DALY: disability-adjusted life year
GBD: Global Burden of Disease
HDI: human development index
NO2: nitrogen dioxide
NOx: nitrogen oxide
PM2.5: particulate matter <2.5 µm in diameter
SDI: socio-demographic index
UVR: ultraviolet radiation
UI: uncertainty interval
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 19.03.23; peer-reviewed by F Liu, H Guo; comments to author 07.08.23; revised version received 21.08.23; accepted 24.10.23; published 05.12.23.

Please cite as:
Jiang B, Wu T, Liu W, Liu G, Lu P
Changing Trends in the Global Burden of Cataract Over the Past 30 Years: Retrospective Data Analysis of the Global Burden of Disease Study 2019
JMIR Public Health Surveill 2023;9:e47349
URL: https://publichealth.jmir.org/2023/1/e47349
doi:10.2196/47349
PMID:38051579

Original Paper

Association Between Purchase of Over-the-Counter Medications and Ovarian Cancer Diagnosis in the Cancer Loyalty Card Study (CLOCS): Observational Case-Control Study

Hannah R Brewer[1], BSc, MSc, PhD; Yasemin Hirst[2], BA, MSc, PhD; Marc Chadeau-Hyam[3], BA, PhD; Eric Johnson[3], BA; Sudha Sundar[4], MPhil, MBBS; James M Flanagan[1], BSc, PhD

[1]Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom
[2]Department of Behavioural Science and Health, University College London, London, United Kingdom
[3]School of Public Health, Faculty of Medicine, Imperial College London, London, United Kingdom
[4]Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, United Kingdom

Corresponding Author:
James M Flanagan, BSc, PhD
Department of Surgery and Cancer
Faculty of Medicine
Imperial College London
4th Floor IRDB, Hammersmith Campus, Du Cane Rd
London, W12 0NN
United Kingdom
Phone: 44 2075942127
Email: j.flanagan@imperial.ac.uk

Abstract

Background: Over-the-counter (OTC) medications are frequently used to self-care for nonspecific ovarian cancer symptoms prior to diagnosis. Monitoring such purchases may provide an opportunity for earlier diagnosis.

Objective: The aim of the Cancer Loyalty Card Study (CLOCS) was to investigate purchases of OTC pain and indigestion medications prior to ovarian cancer diagnosis in women with and without ovarian cancer in the United Kingdom using loyalty card data.

Methods: An observational case-control study was performed comparing purchases of OTC pain and indigestion medications prior to diagnosis in women with (n=153) and without (n=120) ovarian cancer using loyalty card data from two UK-based high street retailers. Monthly purchases of pain and indigestion medications for cases and controls were compared using the Fisher exact test, conditional logistic regression, and receiver operating characteristic (ROC) curve analysis.

Results: Pain and indigestion medication purchases were increased among cases 8 months before diagnosis, with maximum discrimination between cases and controls 8 months before diagnosis (Fisher exact odds ratio [OR] 2.9, 95% CI 2.1-4.1). An increase in indigestion medication purchases was detected up to 9 months before diagnosis (adjusted conditional logistic regression OR 1.38, 95% CI 1.04-1.83). The ROC analysis for indigestion medication purchases showed a maximum area under the curve (AUC) at 13 months before diagnosis (AUC=0.65, 95% CI 0.57-0.73), which further improved when stratified to late-stage ovarian cancer (AUC=0.68, 95% CI 0.59-0.78).

Conclusions: There is a difference in purchases of pain and indigestion medications among women with and without ovarian cancer up to 8 months before diagnosis. Facilitating earlier presentation among those who self-care for symptoms using this novel data source could improve ovarian cancer patients' options for treatment and improve survival.

Trial Registration: ClinicalTrials.gov NCT03994653; https://clinicaltrials.gov/ct2/show/NCT03994653

(JMIR Public Health Surveill 2023;9:e41762) doi:10.2196/41762

KEYWORDS

ovarian cancer; early diagnosis; transactional data; health informatics; cancer risk; medication; self-medication; self-care; over-the-counter medication; nonspecific symptoms; pain medication; indigestion medication

Introduction

Ovarian cancer is the eighth most common cancer in females worldwide and is the sixth most common cause of cancer-related death in females in the United Kingdom [1,2]. There is an urgent need for earlier diagnosis to improve survival outcomes since 5-year survival rates decrease from 93% for women diagnosed at the earliest stage to 20% for women diagnosed at the latest stage[2]; however, a population-based ovarian cancer screening program is not currently recommended in the United Kingdom [3]. One possible approach for improvement is to encourage earlier help-seeking behavior among women experiencing symptoms such as abdominal/pelvic pain, bloating, loss of appetite, and urinary symptoms, which can often be associated with other more common illnesses [4]. Awareness of these nonspecific symptoms among women in the United Kingdom is low [5,6]; consequently, the average time between experiencing a symptom and presenting to primary care with these symptoms is 39 days [7].

A proof-of-concept study conducted in 2019 concluded that there may be an increase in purchases of pain and indigestion medications 10-12 months before ovarian cancer diagnosis to treat symptoms [8], although larger studies were needed to validate this finding. The Cancer Loyalty Card Study (CLOCS) was therefore proposed to investigate this further through an observational case-control study design, which hypothesized that a significant change in purchasing behavior, specifically an increase in self-medication prior to presentation in primary care, could be an indication for early signs of ovarian cancer [9]. The primary objective of the CLOCS is to investigate the purchasing behaviors of women with and without ovarian cancer using commercial data collected through loyalty card use at two participating high street retailers in the United Kingdom with the aim to reduce delays in ovarian cancer diagnosis.

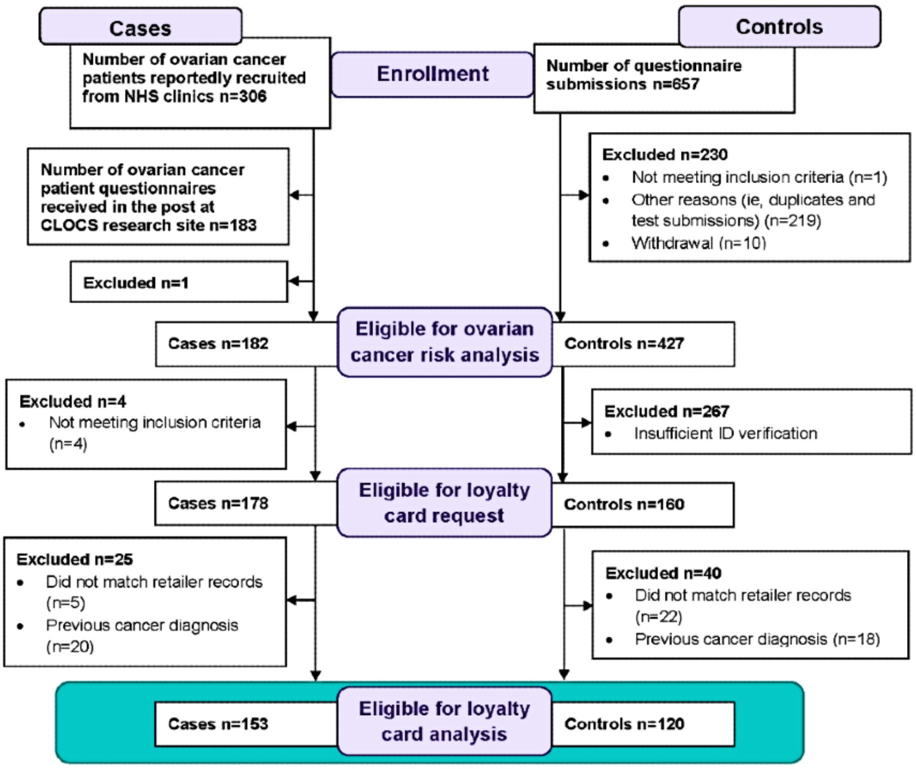
Methods

Study Design, Setting, and Participants

The CLOCS case-control study design and protocol have been described previously [9]. Eligible participants were women in the United Kingdom, 18 years or older, who owned at least one of the participating retailers' loyalty cards. Patients diagnosed with ovarian cancer within 2 years of recruitment to the study until January 31, 2022, were considered cases, whereas controls did not have a previous diagnosis of ovarian cancer. When face-to-face interactions between patients with ovarian cancer and clinics were minimized due to the COVID-19 pandemic, the protocol was amended to allow for patients to be recruited from National Health Service (NHS) clinics over the phone.

Patients were recruited from 12 NHS clinic sites in England, Wales, and Scotland from November 1, 2019, to January 31, 2022. Participants without ovarian cancer were recruited indirectly through an online experiment that investigated the effects of an animated decision aid on willingness to take part in the CLOCS in July 2020 (unpublished), and directly through paid social media advertisements using Meta (formerly known as Facebook), a test recruitment email from one of the participating retailers, word of mouth, and the advocate network VOICE [10] from September 1, 2020, to May 31, 2021. A CONSORT (Consolidated Standards of Reporting Trials) diagram of CLOCS participant recruitment and eligibility for analysis is shown in Figure 1.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram for Cancer Loyalty Card Study (CLOCS) case-control recruitment and loyalty card data collection. NHS: National Health Service.



After obtaining explicit consent, up to 6 years of loyalty card purchase history, starting from the date of recruitment, was requested from two high street retailers, hereafter referred to as high street retailer 1 (HSR1) and high street retailer 2 (HSR2), for eligible participants using the name and number on their loyalty card(s). One of the participating retailers offers a supply of health and beauty items, while the other stocks a wider range of health, beauty, and grocery items, thereby providing CLOCS with a variety of purchases related to the participants' intentions to shop at either retailer. If these details matched the retailers' records, then the participant purchase history was transferred to the CLOCS team. If the details did not match, then the purchase history was not shared, and the participant was recontacted up to two times to clarify their details. Once clarified, the participant's card history was requested once again as described previously [9].

Patient and Public Involvement

The CLOCS team included two ovarian cancer patient advocates who were involved in the project from the inception of the idea to study design, ethical approval applications, and development of recruitment strategies throughout the CLOCS. We presented two public-facing seminars to engage with the wider audience about the CLOCS in December 2020 and March 2021 during Ovarian Cancer Awareness month. Furthermore, we held three annual scientific meetings inviting academics, patient representatives, and participating high street retailers.

Variables

Participants completed a short questionnaire about ovarian cancer risk factors, including ethnicity, marital status, BMI, age at menarche, menopausal status, age at menopause, parity, breastfeeding, hysterectomy, tubal ligation, cancer history, endometriosis, aspirin use, oral contraceptive (OC) use, hormone replacement therapy (HRT) use, family history of ovarian and breast cancers, vaping, and cigarette smoking. The questionnaire also requested information about the symptoms experienced (if any) and number of visits to the general practitioner (GP) in the year before participating in the study for controls or leading up to cancer referral or diagnosis for cases.

In addition, participants were asked about which food and pharmacy stores they shop at regularly, at which stores they own a loyalty card, frequency of use of those loyalty cards, and where they heard about the study (for controls only). The questionnaire was amended to include a question about the number of people in their household. To gain a better understanding of the impact of COVID-19 on purchase history patterns and ovarian cancer diagnosis delays, the questionnaire was amended to include a question about whether participants had been diagnosed with COVID-19 (responses limited to the following: diagnosed and recovered, diagnosed and still ill, suspected but not formally diagnosed, did not have COVID-19). The date of diagnosis, type, stage, grade, *BRCA1* and *BRCA2* mutation status, and any surgical outcomes for cases were provided by the clinical team at their respective recruitment sites.

Statistical Methods

Ovarian Cancer Risk

Unconditional logistic regression was used to calculate the risk of ovarian cancer among CLOCS participants, first using an unadjusted model and then with adjustment for age, OC use (ever/never), menopausal status (premenopausal/postmenopausal), and HRT use (ever/never).

An ovarian cancer risk score was calculated for all participants stratified by age (<50 years and ≥50 years) using an additive model described previously [11]. The risk score was the sum of the product of the median log odds ratio (OR) and the values of the following risk factors: low- and high-dose aspirin use (regular use ever/never), BMI (continuous), breastfeeding (ever/never), number of months breastfeeding (continuous), endometriosis (ever diagnosed/never), family history of breast cancer in first-degree relative (ever/never), family history of ovarian cancer in first-degree relative (ever/never), hysterectomy and no HRT use (ever/never), age at the end of the last pregnancy (continuous), age at menarche (continuous), menopausal status (pre-/postmenopausal), OC use (ever/never), duration of OC use (continuous), number of nonfull-term pregnancies (continuous), number of full-term pregnancies (continuous), and tubal ligation status (ever/never). Instead of limiting the analysis to participants with complete data or eliminate risk factors from the risk score, missing values were imputed using the k-nearest neighbors imputation method in the Visualization and Imputation of Missing Values (*vim*) package in R. Participants were grouped into low-, medium-, or high-risk score groups using cut-off points based on tertiles of the risk score among controls.

Purchase History Analysis

For purchase history analyses, two controls were matched with each case according to age (continuous) and which HSR loyalty card they had (HSR1 only, HSR2 only, HSR1 and HSR2), with replacement. The analysis was focused on the period of 24 months prior to diagnosis for cases to maximize the amount of purchase history data available from participants and to detect a difference in purchases between cases and controls up to 10 months before diagnosis, according to the prior power calculations [9]. Owing to the differing periods of recruitment of cases and control, where possible, the same 24 months of purchases from controls were aligned with each matched case leading up to diagnosis; otherwise, the latest 24 months of purchases were used for controls. Purchases of items commonly used to treat symptoms of ovarian cancer, such as stomach pain and indigestion, were averaged and cumulatively summed for each participant per month over 24 months. Pain and indigestion medication purchases were selected using the categories predetermined by the retailers.

For statistical analysis, purchases for each individual were summed on a sliding 3-month window to smooth the signal and to represent a pattern of individual purchase behavior for cases and controls comparing all purchases to the purchase of both pain and indigestion medication bought together. The Fisher exact test was used to calculate the ORs, 95% CIs, and significance of enrichment at each month prior to diagnosis.

A conditional logistic regression model was used to investigate the association between purchase of medications that treat pain and indigestion over the 24-month period as well as pain and indigestion medication purchases separately, where the medication purchase was the exposure and ovarian cancer diagnosis was the outcome.

A receiver operating characteristic (ROC) analysis was performed for purchases of pain and indigestion medications separately at each month prior to diagnosis for cases and the latest 24 months available for controls. For cases, the ROC analysis was repeated and stratified by stage at diagnosis, where stages 1 and 2 were grouped together and considered the early stage and stages 3 and 4 were considered the late stage.

Sensitivity analyses were performed to assess the impact of several factors on purchase behaviors among CLOCS participants. First, conditional logistic regression was used to assess the association between ovarian cancer diagnosis and purchases of pain and indigestion medications before and after the COVID-19 pandemic (before and after March 2020). Second, to account for seasonal purchase behaviors (eg, painkillers linked to flu and cold season in autumn and winter), the conditional logistic regression model was repeated with adjustment for seasons. Finally, to investigate whether the number of members in a household affects the association between ovarian cancer and purchases, the model was run with adjustment for the number of members in the household and OC use.

Further sensitivity analyses related to ovarian cancer symptom behaviors were performed using the same conditional logistic regression model with adjustment for whether participants reportedly attended the GP in the 12 months prior to diagnosis for cases and prior to participation for controls, as well as with adjustment for whether participants reported experiencing nonspecific ovarian cancer symptoms.

All statistical analyses were conducted in the ISO27001-certified Secure Enclave environment at Imperial College London using R (version 4.1.2).

Ethics Considerations

The CLOCS was reviewed and approved by the North West-Greater Manchester South Research Ethics Committee (19/NW/0427). Explicit consent for the research team to request loyalty card data from participating retailers and to take part in the study was given from all participants.

Results

Participant Characteristics

As of January 31, 2022, 182 cases and 441 controls consented to take part in the CLOCS. As shown in the CONSORT diagram in Figure 1, one case was excluded for providing insufficient consent and 14 control participants withdrew or were excluded. Of the remaining 427 controls, only 160 (37.5%) provided sufficient ID verification and were therefore eligible for loyalty card data requests. There were 4 cases who completed the risk factor questionnaire but did not provide loyalty card details and therefore were excluded from purchase history requests.

The characteristics of CLOCS participants are shown in Table 1. On average, cases were about 13 years older than controls. In terms of loyalty card ownership, 47% of cases and 33% of controls owned cards from both retailers, while the remaining participants had one or the other. The majority of cases who took part in the CLOCS were diagnosed with serous ovarian cancer (75.8%) and of those, 94% were diagnosed at stage 3 or 4 (see Table S1 in Multimedia Appendix 1 for details).

Ovarian cancer risk and risk scores among participants are shown in Table S2 and Table S3, respectively, of Multimedia Appendix 1. The majority of cases (46.7%) fell into the high ovarian cancer risk score tertile, while 24.7% had a medium risk score and 28.6% of cases were in the low risk score tertile.

Purchase histories were requested for 178 cases and 160 controls from HSR1 and HSR2, as described in Figure 1. After consent, ID verification, matching loyalty card details with retailers' records, and recontacting participants who did not match to clarify loyalty card details, purchase histories from 173 (97.2%) cases and 138 (86.3%) controls were transferred to the CLOCS team for analysis. There were 5 (2.8%) cases and 22 (13.8%) controls whose loyalty card details did not match the records at either HSR1, HSR2, or both. There were 20 cases and 18 controls with a previous cancer diagnosis before participating in the study and were excluded from all purchase history analyses (participants with a previous nonmelanoma skin cancer were included). After exclusion of participants with a previous cancer diagnosis and matching cases and controls on age and cardholder status, purchase histories of 153 cases and 306 controls (77 unique individuals) were included in purchase history analyses. The median age difference for matched pairs was 0.3 years (IQR –0.4 to 1.7 years). Among this reduced set of participants, only OC use was associated with ovarian cancer risk and was therefore the only risk factor adjusted for in subsequent purchase history analyses.

Table 1. Characteristics of participants.

Characteristic	Cases (n=182)	Controls (n=427)
Age (years), mean (SD)[a]	64.7 (10.9)	51.6 (13.7)
Age group (years), n (%)		
18-39	6 (3.6)	61 (16.2)
40-49	7 (4.7)	96 (23.8)
50-59	32 (21.3)	125 (31.8)
60-69	64 (36.7)	107 (21.2)
≥70	73 (33.7)	33 (5.9)
Ethnicity, n (%)		
White	173 (95.1)	400 (93.7)
Nonwhite	8 (4.4)	11 (2.6)
Prefer not to say	0 (0)	0 (0)
Missing	1 (0.5)	16 (3.7)
Loyalty card, n (%)		
HSR1[b] card only	48 (26.4)	169 (39.6)
HSR2[c] card only	44 (24.2)	116 (27.2)
Both HSR1 and HSR2	86 (47.3)	142 (33.3)
Neither	4 (2.2)	0 (0)
Loyalty card use, n (%)		
Not at all	1 (0.5)	2 (0.5)
Not very often	9 (4.9)	10 (2.3)
Sometimes	18 (9.9)	42 (9.8)
Often	42 (23.1)	117 (27.4)
All the time	109 (59.9)	256 (60.0)
Missing	3 (1.6)	0 (0)
Number of members in the household (including participant)[d], n (%)		
1	32 (22.2)	43 (14.5)
2	75 (52.1)	120 (40.5)
≥3	29 (20.1)	131 (44.3)
Missing	8 (5.6)	2 (0.7)

[a] Missing data for n=5 (0.9%) controls.

[b] HSR1: high street retailer 1.

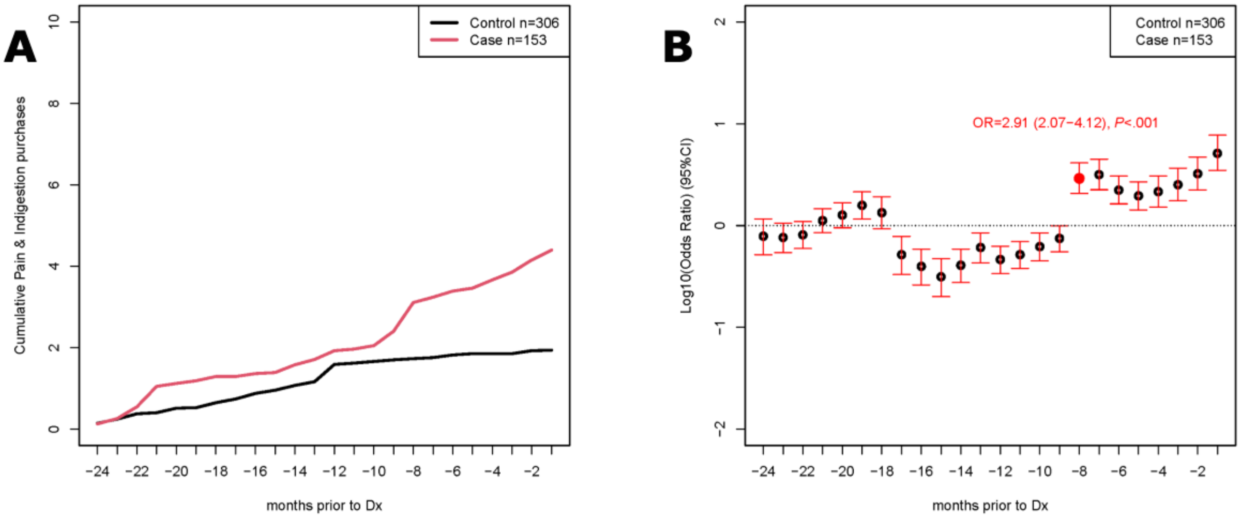
[c] HSR2: high street retailer 2.

[d] Number of members in the household question was added to the questionnaire after amendment and during active recruitment. Missingness describes those who participated before the question was available.

Cumulative Purchases

The average cumulative number of purchases of pain and indigestion medications together are shown in Figure 2A. Throughout the 24 months, there was an increased number of these purchases among cases before diagnosis compared with that of controls, with an increased rate of purchases 8 months before diagnosis. The mean numbers of purchases of pain and indigestion medications were greater for cases in the 6 months, 12 months, and 24 months before diagnosis, with a significant association between these purchases and ovarian cancer diagnosis in the 6 months before diagnosis (OR 1.05, 95% CI 1.01-1.10; Table S4 in Multimedia Appendix 1). The Fisher exact test (Figure 2B) showed maximum discrimination between cases and controls 8 months before diagnosis (OR 2.91, 95% CI 2.07-4.12).

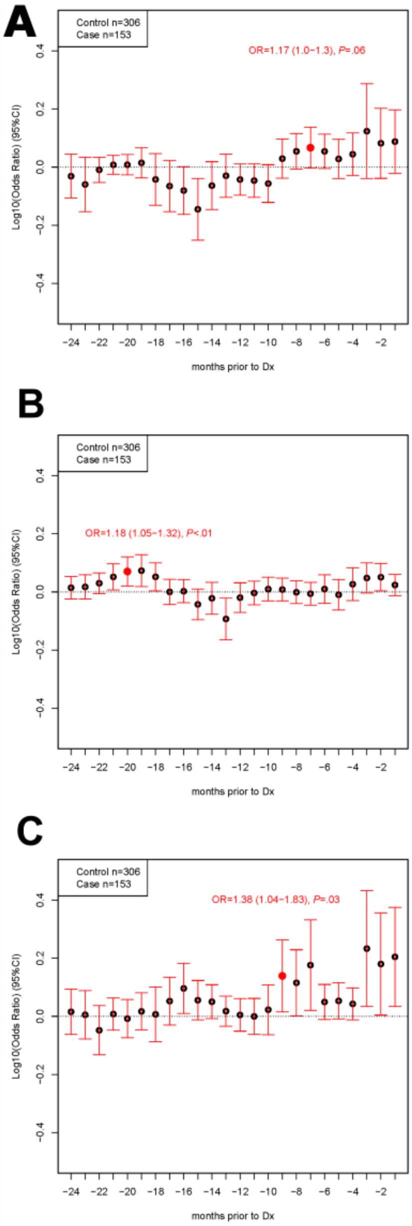
Figure 2. Combined pain and indigestion medications in the Cancer Loyalty Card Study (CLOCS). Cases (n=153) were matched with controls with respect to age and store card use (HSR1 only, HSR2 only, or HSR1+HSR2) at a 2:1 ratio with replacement (n=306 controls in matched analysis, 77 unique individuals). Participants with a previous diagnosis of cancer were excluded. Where possible, the same 24 months of purchases from controls were aligned to each matched case leading to their diagnosis (–24 months to –1 month on the X-axis); otherwise, the last 24 months of purchases were used for controls. (A) Average cumulative purchases of combined pain and indigestion medications in cases (red) and controls (black). (B) Purchases summed on a sliding 3-month window for cases and controls comparing all purchases to purchase of both pain and indigestion medication bought together. The Fisher exact test was used to calculate the odds ratios and significance of enrichment at each month prior to diagnosis. Dx: diagnosis; HSR: high street retailer.



Adjusted Models

As shown in Figure 3A, the likelihood of being diagnosed with ovarian cancer was significantly lower at 15 months before diagnosis in combined purchases of pain and indigestion medications, and there was no significant increase in risk of ovarian cancer diagnosis. However, when pain medication purchases were considered alone, there was a significantly increased likelihood of ovarian cancer 19 months before diagnosis (OR 1.18, 95% CI 1.05-1.32), as shown in Figure 3B. There was a significantly increased association between ovarian cancer diagnosis and indigestion medication purchases alone 9 months before diagnosis (OR 1.38, 95% CI 1.04-1.83), as shown in Figure 3C and Table S5 of Multimedia Appendix 1.

Figure 3. Conditional logistic regression for purchases prior to diagnosis (Dx). Cases (n=153) were matched with controls according to age and store card use (HSR1 only, HSR2 only, or HSR1+HSR2) at a 2:1 ratio with replacement (n=306 controls in matched analysis, 77 unique individuals). Participants with a previous diagnosis of cancer were excluded. Where possible, the same 24 months of purchases from controls were aligned to each matched case leading to their diagnosis (–24 months to –1 month on the X-axis); otherwise, the last 24 months of purchases were used for controls. Conditional logistic regression of matched case-control sets adjusting for oral contraceptive pill use and household number for combined sum of pain and indigestion medication (A), pain medication only (B), or indigestion medication only (C). HSR: high street retailer; OR: odds ratio.

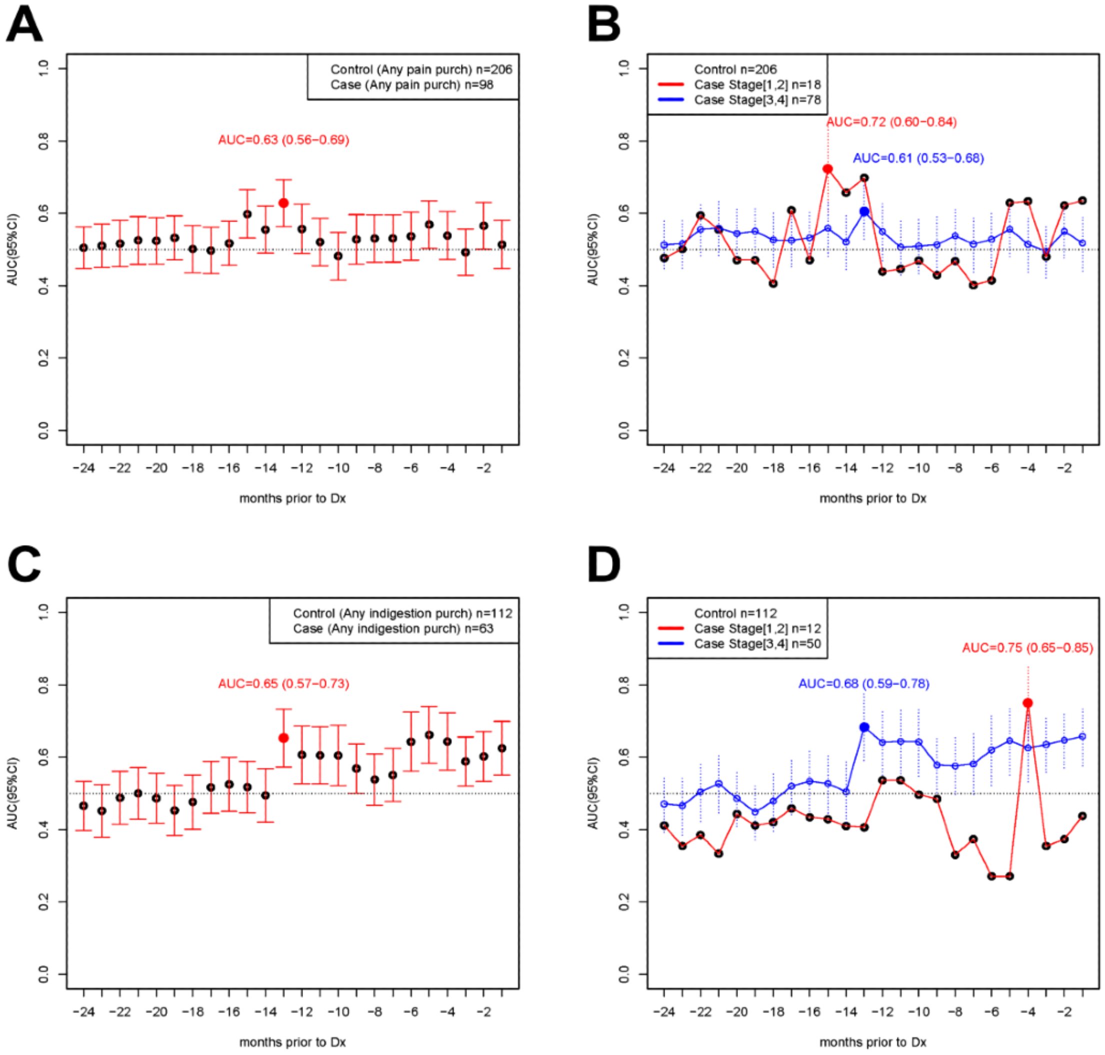


ROC Analyses

Restricting the analysis to only participants who purchased any pain medication over the 24-month period, 98 cases and 206 controls remained. The ROC analysis for pain medication purchases alone (Figure 4A and Table S6 in Multimedia Appendix 1), excluding participants who did not purchase any pain medication, showed a maximum area under the curve (AUC) at 13 months before diagnosis (AUC=0.63, 95% CI 0.56-0.69). When further stratified on early stage at diagnosis, the AUC reached a maximum of 0.72 (95% CI 0.60-0.84) 15 months before diagnosis (Figure 4B).

Among those who only purchased any indigestion medication over the 24-month period, 63 cases and 112 controls remained. The ROC analysis for indigestion medication purchases alone (Figure 4C and Table S6 of Multimedia Appendix 1) also showed a maximum AUC at 13 months before diagnosis (AUC=0.65, 95% CI 0.57-0.73), and this further improved when restricted to late-stage ovarian cancer diagnosis (AUC=0.68, 95% CI 0.59-0.78), as shown in Figure 4D. Overall, the AUC was greater for late-stage diagnosis than early-stage diagnosis in all 24 months before diagnosis.

Figure 4. Receiver operating characteristic (ROC) analysis stratifying by stage of diagnosis (Dx). Purchases in cases and controls were aligned as per Figure 3. (A) ROC analysis of pain medications at each month prior to diagnosis in cases (n=98) and controls (n=206) that purchased any pain medication in the 24-month period. Area under the curve (AUC) and 95% CIs are presented. (B) Stratification into cases diagnosed at a low stage (1,2) (n=18, red line) or high stage (3,4) (n=78, blue line). (C) ROC analysis of indigestion medications at each month prior to diagnosis in cases (n=63) and controls (n=112) that purchased any indigestion medication in the 24-month period. AUC and 95% CIs are presented. (D) Stratification into cases diagnosed at a low stage (1,2) (n=12, red line) or high stage (3,4) (n=50, blue line).



Sensitivity Analyses

Adjusting for purchase behaviors before and after the COIVD-19 pandemic, different seasons, number of household members, and OC use showed little to no effect on the association of purchases with ovarian cancer over 24 months before diagnosis (Tables S7-S9 in Multimedia Appendix 1).

After adjustment for the ovarian cancer risk score (Tables S10-S12 in Multimedia Appendix 1), there was no significant change in the OR for detecting cancer symptoms among CLOCS participants who purchased pain and indigestion medications compared with only adjusting for OC use and number of household members.

Discussion

Principal Results

The primary aim of the CLOCS was to investigate purchasing behaviors of women with and without ovarian cancer before diagnosis to determine whether there is a difference in purchases of OTC medications among ovarian cancer patients before diagnosis. Purchases among CLOCS participants indicated an increase of pain and indigestion medication purchases prior to diagnosis of ovarian cancer. Specifically, we detected an increase in purchases of pain and indigestion medications bought together at 8 months before diagnosis. When considered separately in conditional logistic regression models adjusting for confounders, the signal for increased purchases of indigestion medications occurred from 9 months before diagnosis. For pain medication, we observed a period of increased purchases at 19 months prior to diagnosis, but no significant differences in the 12 months leading up to diagnosis. This suggests that there may be a difference in the timeline along which the patients with ovarian cancer included in this study may have experienced and cared for symptoms that can be treated with pain or indigestion medication.

Comparison With Prior Work and Clinical Implications

The guidelines for investigation and initial assessment of ovarian cancer in symptomatic women vary greatly internationally partially due to the nonspecific nature of symptoms [12], and almost 1 in 4 cancer patients present with a broad range of abdominal symptoms before diagnosis [13]. Specifically, for an effective treatment of ovarian cancer, while symptom presentation may not always correlate with early-stage diagnosis, an earlier diagnosis even at stage IV could mean that patients may receive treatments that could improve their chances of surviving 1 year or longer. In the context of this study, when stratified by stage of diagnosis, for those diagnosed at an early stage, we identified that the purchase of pain medications occurred at 15 months before diagnosis, whereas indigestion medications were much more closely associated to diagnosis, at 4 months. For those diagnosed at a late stage, we identified indigestion medication purchases as early as 13 months prior to diagnosis and then in every month from 6 months up to diagnosis, but no significant change in pain medication purchases was found in the 12 months prior to diagnosis. From a clinical point of view, this highlights a unique interval in the tumor development where symptoms are presenting at different times. However, our sample sizes were too small to further stratify according to patient characteristics based on their OTC use associated with reported symptoms and stage at diagnosis.

While these differences could reflect the nonspecific nature of symptoms that occur during the development of ovarian cancer, they could also indicate the complex symptom appraisal processes for nonspecific cancer symptoms and presentation in health care [14]. For example, patients with pain 1 year before diagnosis might have already seen a health care professional following the initial presentation of their symptoms. However, in the absence of other symptoms not warranting an urgent cancer referral (ie, diagnostic pathway in the United Kingdom where patients with suspected cancer are seen within 2 weeks of presentation), patients' symptoms may have been treated for other illnesses, resulting in a longer diagnostic interval from the first onset of symptoms [15]. In fact, we identified no differences in the number of GP visits in the 12 months leading up to diagnosis or participation in the study between cases and controls, indicating that transactional data could be useful to further explore self-care behaviors even after patients may have presented in health care for their symptoms. A previous study reported that 36% of ovarian cancer patients presented to their GP at least three times before diagnosis [16], while approximately 25.8% of cases in the CLOCS reported attending three or more visits to the GP before diagnosis (see Table S3 in Multimedia Appendix 1). Much lower rates of help-seeking behaviors could be attributed to the fact that approximately 64% of cases were diagnosed with ovarian cancer throughout the COVID-19 pandemic, during which many people avoided contacting or seeking help about their symptoms from their local GP [17]. Our results further emphasize the importance of safety nets to be in place between community pharmacies and primary care to ensure that patients who continue to care for their ongoing symptoms are identified and referred promptly.

Alternatively, the duration and the effectiveness of self-care using OTC indigestion and pain medication may be related to the type of abdominal symptom experienced [13]. However, there is little qualitative evidence on patients' accounts of when they experienced pain and indigestion, or how they coped with these symptoms and cared for them before they received a diagnosis [18]. Nevertheless, the number of cases in the CLOCS with early-stage diagnosis was small; therefore, further studies with a larger number of patients with early-stage diagnosis would be needed to verify if self-care behaviors and earlier presentation are moderated by symptom development and individuals' symptom appraisal processes.

Strengths and Limitations

The main strength of this study is that this is the first study to investigate transactional data for understanding individuals' self-care behaviors prior to a cancer diagnosis using a case-control study design. The use of prospectively collected data prior to cancer diagnosis avoids potential recall bias that would have occurred from querying as to when patients purchased OTC products for nonspecific symptoms using questionnaires. In the CLOCS, exposure to pain and indigestion medication purchases was measured using loyalty card data that are recorded by retailers, thereby adjusting for the potential recall bias as well as respondents' bias to provide desirable outcomes. For example, although there was a significant difference between purchases of indigestion medications among cases and controls, the self-reported indigestion symptoms were not significantly different between cases and controls in this study (see Table S13 in Multimedia Appendix 1). Recall bias is a common issue reported for case-control studies; thus, to overcome recall bias associated with symptoms experienced before diagnosis, we aimed to recruit patients soon after their cancer diagnosis and/or exclude patients diagnosed before 2018.

One of the main limitations of this study is that purchase of an item is not the same as consumption of the item. The majority of participants reported that they lived with others in their household, and therefore could have been purchasing pain and indigestion medications for others instead of directly consuming these themselves. This limitation is applied equally to cases and controls who might also be buying products for others. There was no difference in the effect of purchases and ovarian cancer diagnosis when adjusting for other household members. Similarly, the analysis was restricted to purchases recorded only when participants used their loyalty cards, and any items purchased without the loyalty card were not recorded. However, 85% of participants reported that they used their loyalty card often or all the time. It is also likely that participants could be shopping for these medications at other retailers not included in this study, resulting in an underrepresentation of purchases with the limited participating retailers. Therefore, it is possible that the significant association observed here could be stronger with the inclusion of purchases of these items from other retailers, although further studies including more retailers are needed to verify this possibility.

Furthermore, most of the timeline for recruitment of cases and controls to this case-control study occurred during the COVID-19 pandemic. As a result of the paused recruitment and necessary amendments, the study did not meet the original goal to recruit 500 cases and 500 controls, and the sample size for analysis was smaller than originally intended. Due to the nature of recruitment of cases and controls, there was on average a 13-year age difference between cases and controls, reflecting an age bias in acceptability of sharing data, participating in research, and online recruitment of controls (HRB, YH, and JMF preliminary data, August 8, 2022). However, this age difference was accounted for in all analyses either through age adjustment or matching on age.

Future Research

This study highlights the significant timeframe (9 to 19 months prior to diagnosis) in the patient interval in cancer diagnosis for when individuals who have yet to receive an ovarian cancer diagnosis and may have been reacting to the bodily changes and caring for their symptoms using OTC medications to maintain their health status. In the landscape of cancer epidemiology and symptom presentation research [5,6], this result contradicts with the timelines reported for the first symptoms recorded on primary care–based records to calculate the patient interval, which potentially highlights the need to have better data collection methods on patient well-being.

Likewise, this study only reported the temporal associations between OTC purchases and ovarian cancer diagnosis, and future data analyses will include understanding the characteristics of changing patterns of purchases, as well as combining other purchase categories that are not linked to symptom management using unbiased machine-learning approaches. Future analysis may also consider the association between OTC purchases and other diseases with similar symptoms, such as gastrointestinal diseases. With growing literature on utilizing other data sources in health research, transactional data have already been used to evaluate dietary patterns, and such analysis is considered to be a tool that can supplement our understanding of people's dietary behaviors [19,20]. Similarly, this type of surveillance could provide an opportunity to utilize loyalty card programs, where available around the world, to assess purchases related to OTC treatment of many cancer types to improve earlier detection.

Conclusions

To date, the CLOCS has identified an increase in purchases of OTC pain and indigestion medications among ovarian cancer patients before diagnosis compared with those of women without ovarian cancer. Further studies with larger numbers of ovarian cancer patients, diagnosed at different stages, and more participating retailers are needed to verify these findings, which can lead to the future development of an alert system for individuals to seek medical attention for the symptoms they are experiencing sooner than they might otherwise.

Acknowledgments

The authors would like to acknowledge all of the participants in the Cancer Loyalty Card Study (CLOCS), the participating retailers, and the National Health Service (NHS) sites that recruited ovarian cancer patients to this study. We would like to acknowledge Deborah Tanner and Fiona Murphy who have acted as ovarian cancer patient representatives since the inception of the CLOCS project. This study was funded by Cancer Research UK Early Detection and Diagnosis Project (grant C38463/A26726) with additional support from the Cancer Research UK Imperial Centre, National Institute for Health Research Imperial Biomedical Research Centre, and Ovarian Cancer Action Research Centre.

Data Availability

Access to all Cancer Loyalty Card Study (CLOCS) purchase history data collected from the retailers is restricted due to the data sharing agreements in place and the sensitive nature of the data. However, we welcome other researchers to contact us for collaborations. The principal investigator and research team will review each proposal and, if approved, the analysis will be conducted by a member of the CLOCS team.

Authors' Contributions

JMF, YH, and HRB conceived the study. HRB and JMF conducted all statistical analyses. HRB, JMF, and YH reviewed the results. EJ provided expertise in information governance and data management. HRB wrote the first draft of the manuscript. YH, JMF, SS, and MC-H commented on and reviewed the final draft of the manuscript. All authors agree on the final version of the manuscript.

Conflicts of Interest

SS declares receiving honoraria from Astra Zeneca, Glaxo Smith Kline and MSD (Merck & Co) for speakers fees. The author has no conflicts in relation to this manuscript. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 97 KB - publichealth_v9i1e41762_app1.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Cancer Statistics for the UK. Cancer Research UK. 2021. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk [accessed 2021-08-01]
3. Menon U, Gentry-Maharaj A, Burnell M, Singh N, Ryan A, Karpinskyj C, et al. Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. Lancet 2021 Jun 05;397(10290):2182-2193 [FREE Full text] [doi: 10.1016/S0140-6736(21)00731-5] [Medline: 33991479]
4. Huepenbecker SP, Sun CC, Fu S, Zhao H, Primm K, Giordano SH, et al. Factors impacting the time to ovarian cancer diagnosis based on classic symptom presentation in the United States. Cancer 2021 Nov 15;127(22):4151-4160 [FREE Full text] [doi: 10.1002/cncr.33829] [Medline: 34347287]
5. Brain KE, Smits S, Simon AE, Forbes LJ, Roberts C, Robbé IJ, ICBP Module 2 Working Group. Ovarian cancer symptom awareness and anticipated delayed presentation in a population sample. BMC Cancer 2014 Mar 10;14:171 [FREE Full text] [doi: 10.1186/1471-2407-14-171] [Medline: 24612526]
6. Low EL, Waller J, Menon U, Jones A, Reid F, Simon AE. Ovarian cancer symptom awareness and anticipated time to help-seeking for symptoms among UK women. J Fam Plann Reprod Health Care 2013 Jul;39(3):163-171 [FREE Full text] [doi: 10.1136/jfprhc-2012-100473] [Medline: 23709609]
7. Sundar S, Neal RD, Kehoe S. Diagnosis of ovarian cancer. BMJ 2015 Sep 01;351:h4443. [doi: 10.1136/bmj.h4443] [Medline: 26328593]
8. Flanagan JM, Skrobanski H, Shi X, Hirst Y. Self-care behaviors of ovarian cancer patients before their diagnosis: proof-of-concept study. JMIR Cancer 2019 Jan 17;5(1):e10447 [FREE Full text] [doi: 10.2196/10447] [Medline: 30664464]
9. Brewer HR, Hirst Y, Sundar S, Chadeau-Hyam M, Flanagan JM. Cancer Loyalty Card Study (CLOCS): protocol for an observational case-control study focusing on the patient interval in ovarian cancer diagnosis. BMJ Open 2020 Sep 08;10(9):e037459 [FREE Full text] [doi: 10.1136/bmjopen-2020-037459] [Medline: 32900761]
10. Improving detection of ovarian cancer using store loyalty cards (Cancer Loyalty Card Study). Voice Global. URL: https://tinyurl.com/5a6b5v93 [accessed 2023-01-11]
11. Clyde MA, Palmieri Weber R, Iversen ES, Poole EM, Doherty JA, Goodman MT, on behalf of the Ovarian Cancer Association Consortium. Risk prediction for epithelial ovarian cancer in 11 United States-based case-control studies: incorporation of epidemiologic risk factors and 17 confirmed genetic loci. Am J Epidemiol 2016 Oct 15;184(8):579-589 [FREE Full text] [doi: 10.1093/aje/kww091] [Medline: 27698005]
12. Funston G, Van Melle M, Baun ML, Jensen H, Helsper C, Emery J, et al. Variation in the initial assessment and investigation for ovarian cancer in symptomatic women: a systematic review of international guidelines. BMC Cancer 2019 Nov 01;19(1):1028 [FREE Full text] [doi: 10.1186/s12885-019-6211-2] [Medline: 31676000]
13. Koo MM, von Wagner C, Abel GA, McPhail S, Hamilton W, Rubin GP, et al. The nature and frequency of abdominal symptoms in cancer patients and their associations with time to help-seeking: evidence from a national audit of cancer diagnosis. J Public Health 2018 Sep 01;40(3):e388-e395 [FREE Full text] [doi: 10.1093/pubmed/fdx188] [Medline: 29385513]
14. Koo MM, Swann R, McPhail S, Abel GA, Elliss-Brookes L, Rubin GP, et al. Presenting symptoms of cancer and stage at diagnosis: evidence from a cross-sectional, population-based study. Lancet Oncol 2020 Jan;21(1):73-79 [FREE Full text] [doi: 10.1016/S1470-2045(19)30595-9] [Medline: 31704137]
15. Renzi C, Whitaker KL, Wardle J. Over-reassurance and undersupport after a 'false alarm': a systematic review of the impact on subsequent cancer symptom attribution and help seeking. BMJ Open 2015 Feb 04;5(2):e007002 [FREE Full text] [doi: 10.1136/bmjopen-2014-007002] [Medline: 25652803]
16. Lyratzopoulos G, Neal RD, Barbiere JM, Rubin GP, Abel GA. Variation in number of general practitioner consultations before hospital referral for cancer: findings from the 2010 National Cancer Patient Experience Survey in England. Lancet Oncol 2012 Apr;13(4):353-365 [FREE Full text] [doi: 10.1016/S1470-2045(12)70041-4] [Medline: 22365494]
17. Quinn-Scoggins HD, Cannings-John R, Moriarty Y, Whitelock V, Whitaker KL, Grozeva D, et al. Cancer symptom experience and help-seeking behaviour during the COVID-19 pandemic in the UK: a cross-sectional population survey. BMJ Open 2021 Sep 16;11(9):e053095 [FREE Full text] [doi: 10.1136/bmjopen-2021-053095] [Medline: 34531224]
18. Evans J, Ziebland S, McPherson A. Minimizing delays in ovarian cancer diagnosis: an expansion of Andersen's model of 'total patient delay'. Fam Pract 2007 Feb;24(1):48-55. [doi: 10.1093/fampra/cml063] [Medline: 17158183]
19. Clark SD, Shute B, Jenneson V, Rains T, Birkin M, Morris MA. Dietary patterns derived from UK supermarket transaction data with nutrient and socioeconomic profiles. Nutrients 2021 Apr 27;13(5):1481 [FREE Full text] [doi: 10.3390/nu13051481] [Medline: 33925712]
20. Green MA, Watson AW, Brunstrom JM, Corfe BM, Johnstone AM, Williams EA, et al. Comparing supermarket loyalty card data with traditional diet survey data for understanding how protein is purchased and consumed in older adults for the UK, 2014-16. Nutr J 2020 Aug 13;19(1):83 [FREE Full text] [doi: 10.1186/s12937-020-00602-3] [Medline: 32791968]

Abbreviations

AUC: area under the curve
CLOCS: Cancer Loyalty Card Study
CONSORT: Consolidated Standards of Reporting Trials
GP: general practitioner
HRT: hormone replacement therapy
HSR1: high street retailer 1
HSR2: high street retailer 2
NHS: National Health Service
OC: oral contraceptive
OR: odds ratio
OTC: over the counter
ROC: receiver operating characteristic

Edited by G Eysenbach; submitted 08.08.22; peer-reviewed by S Zhang; comments to author 13.10.22; revised version received 18.10.22; accepted 30.11.22; published 26.01.23.

Please cite as:
Brewer HR, Hirst Y, Chadeau-Hyam M, Johnson E, Sundar S, Flanagan JM
Association Between Purchase of Over-the-Counter Medications and Ovarian Cancer Diagnosis in the Cancer Loyalty Card Study (CLOCS): Observational Case-Control Study
JMIR Public Health Surveill 2023;9:e41762
URL: https://publichealth.jmir.org/2023/1/e41762
doi:10.2196/41762
PMID:36701184

Original Paper

Pediatric Injury Surveillance From Uncoded Emergency Department Admission Records in Italy: Machine Learning–Based Text-Mining Approach

Danila Azzolina[1*], PhD; Silvia Bressan[2*], PhD; Giulia Lorenzoni[3], PhD; Giulia Andrea Baldan[3], MA; Patrizia Bartolotta[3], MD; Federico Scognamiglio[3], MA; Andrea Francavilla[3], MD; Corrado Lanera[3], PhD; Liviana Da Dalt[2*], MD; Dario Gregori[3*], PhD

[1]Department of Environmental and Preventive Sciences, University of Ferrara, Ferrara, Italy

[2]Department of Women's and Children's Health, University of Padova, Padua, Italy

[3]Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padova, Padua, Italy

[*]these authors contributed equally

Corresponding Author:
Dario Gregori, PhD
Unit of Biostatistics, Epidemiology and Public Health
Department of Cardiac, Thoracic, Vascular Sciences, and Public Health
University of Padova
Via Leonardo Loredan 18
Padua, 35128
Italy
Phone: 39 049 8275384
Email: dario.gregori@unipd.it

Abstract

Background: Unintentional injury is the leading cause of death in young children. Emergency department (ED) diagnoses are a useful source of information for injury epidemiological surveillance purposes. However, ED data collection systems often use free-text fields to report patient diagnoses. Machine learning techniques (MLTs) are powerful tools for automatic text classification. The MLT system is useful to improve injury surveillance by speeding up the manual free-text coding tasks of ED diagnoses.

Objective: This research aims to develop a tool for automatic free-text classification of ED diagnoses to automatically identify injury cases. The automatic classification system also serves for epidemiological purposes to identify the burden of pediatric injuries in Padua, a large province in the Veneto region in the Northeast Italy.

Methods: The study includes 283,468 pediatric admissions between 2007 and 2018 to the Padova University Hospital ED, a large referral center in Northern Italy. Each record reports a diagnosis by free text. The records are standard tools for reporting patient diagnoses. An expert pediatrician manually classified a randomly extracted sample of approximately 40,000 diagnoses. This study sample served as the gold standard to train an MLT classifier. After preprocessing, a document-term matrix was created. The machine learning classifiers, including decision tree, random forest, gradient boosting method (GBM), and support vector machine (SVM), were tuned by 4-fold cross-validation. The injury diagnoses were classified into 3 hierarchical classification tasks, as follows: injury versus noninjury (task A), intentional versus unintentional injury (task B), and type of unintentional injury (task C), according to the World Health Organization classification of injuries.

Results: The SVM classifier achieved the highest performance accuracy (94.14%) in classifying injury versus noninjury cases (task A). The GBM method produced the best results (92% accuracy) for the unintentional and intentional injury classification task (task B). The highest accuracy for the unintentional injury subclassification (task C) was achieved by the SVM classifier. The SVM, random forest, and GBM algorithms performed similarly against the gold standard across different tasks.

Conclusions: This study shows that MLTs are promising techniques for improving epidemiological surveillance, allowing for the automatic classification of pediatric ED free-text diagnoses. The MLTs revealed a suitable classification performance, especially for general injuries and intentional injury classification. This automatic classification could facilitate the epidemiological surveillance of pediatric injuries by also reducing the health professionals’ efforts in manually classifying diagnoses for research purposes.

(JMIR Public Health Surveill 2023;9:e44467) doi:10.2196/44467

KEYWORDS

machine learning; pediatrics; child and adolescent health; text mining; injury; death; surveillance; pediatric admission; hospitalization; patient record; unintentional injury; emergency department; emergency; epidemiological surveillance

Introduction

Unintentional injury is the leading cause of morbidity and mortality in children worldwide [1]. An injury is defined as tissue damage that occurs secondary to acute exposure to physical agents (eg, thermal, kinetic, chemical, electrical, or electrical energy or water) or chemicals (eg, poisoning). An injury can be fatal or nonfatal and can occur unintentionally or as a result of purposeful acts of harm (intentional) [2]. Unintentional injuries can be prevented or controlled because they are potentially understandable and predictable [3].

In 2013, 15.4% of 2.6 million unintentional injuries worldwide involved a fatal outcome for children between 1 and 14 years of age [1]. In Europe, 42,000 children and adolescents aged 0-19 years died of unintentional injuries in 2004 [4]. Moreover, a considerable number of children may incur some form of disability as a result of injury, often with lifelong consequences [5]. Decreasing the injury burden is the main challenge for child and adolescent public health policies over the next century [6]. For this reason, public health departments must pay more attention to the problem to implement prevention policies [6].

Injury surveillance is made difficult by a series of logistic and structural challenges, the most important of which is the accurate coding of injury mechanisms, products involved, types of injury, and body parts involved, given that emergency department (ED) admission and discharge records are largely based on narrative free-text notes [7]. Injury surveillance integrated with timely data dissemination is crucial for planning and evaluating prevention policies [8] and quantifying injury burden and related risk factors [9,10].

In Italy and other European and newly developed countries [11], narratives and free-text records are standard tools for reporting patient diagnoses. Automatic classification of such free-text information using machine learning techniques (MLTs) would be a powerful tool to improve injury surveillance [12].

This is true, especially for the ED, where physicians and medical personnel often face stressful situations from a clinical and management perspective [13]. Within this general framework, it could be promising to provide an automated MLT-based system aimed at facilitating free-text diagnosis encoding, by also limiting an additional burden for the overwhelmed medical staff. This MLT-based system could be tailored for research and epidemiological surveillance purposes. Furthermore, this surveillance system could be promising for pediatric injury surveillance purposes because most of the incidents that occur on the ground are referred to such departments, especially large pediatric EDs [13].

The literature over the past 10 years indicates an increasing interest in the automated categorization of free-text diagnoses due to the increased availability of documents in digital form [14,15]. Automatic MLT classifiers can learn (during the training phase) from a set of manually classified documents with complex free-text lexical patterns. A properly trained MLT tool can categorize a free-text record into its corresponding class. The advantages of this approach over manual methods are efficiency and saving time (in terms of expert labor) for free-text classification [15].

Statistical text mining methods can also be useful tools to classify electronic ED admission records and properly identify unintentional injury events [15].

This study represents, to our knowledge, the first effort in the literature in proposing an automatic injury classification system based on the free-text data of pediatric ED diagnoses. We propose this algorithm to facilitate injury epidemiological surveillance. The system is aimed at limiting the burden of health care professionals, who are overburdened by patient care and management tasks, in manually classifying diagnoses for epidemiological research.

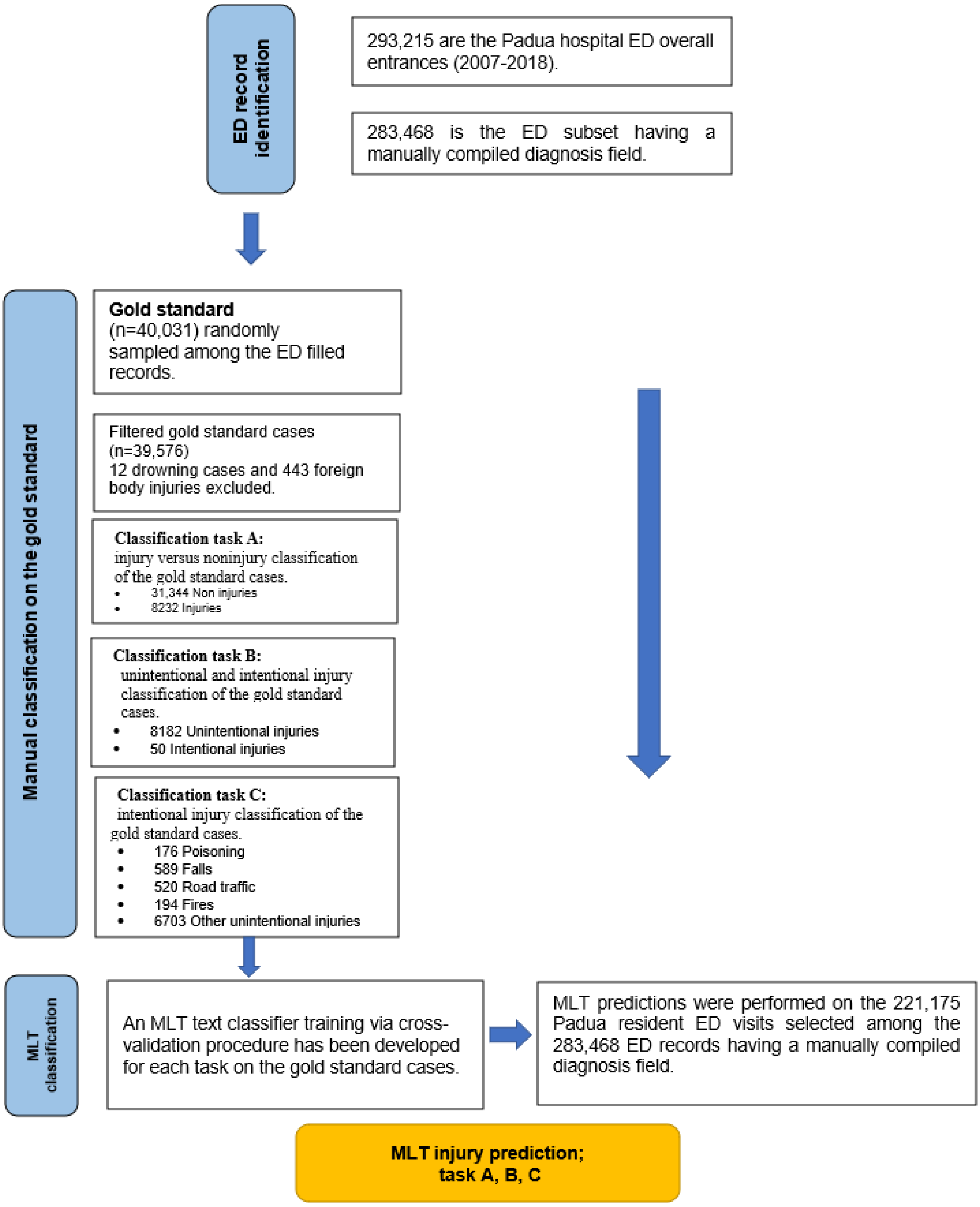
Methods

Data Selection

The study included 283,468 pediatric ED records with a filled discharge diagnosis field among 293,215 records [16] from the local electronic medical record system of Padova University Hospital in Northeast Italy between 2007 and 2018 (Figure 1).

The average ED annual workload is approximately 25,000 visits. The upper age limit to access the pediatric ED is 15 years. A higher and more variable age limit applies to children followed by the Department of Pediatrics for chronic illnesses. The Padova Hospital Pediatrics ED is characterized by high patient turnover with an average hospitalization time of 4-5 days. The number of admissions after ED access is approximately 850 per year.

Figure 1. Study flowchart—emergency department (ED) selection and gold standard identification together with manual injury classification procedure. Machine learning technique (MLT) cross-validation and prediction procedures for tasks A, B, and C are represented by the dark grey box.



Ethics Approval

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Azienda Ospedaliera of Padova (Hospital Ethics Committee; 0022925) on April 8, 2021. The analysis is carried out on observational data of a secondary nature; however, patients signed a consent form to allow the data to be used for scientific purposes at the time of collection. The records of each patient are kept anonymous with an appropriate identification key excluding personal information. No Compensation has been provided for subjects involved in the research.

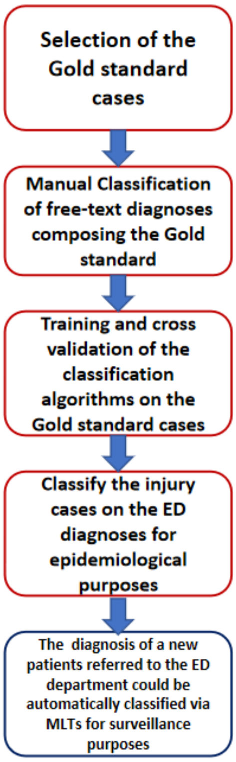
Learning Algorithm for Epidemiological Surveillance

The definition of the free-text classification algorithm and its use for epidemiological surveillance purposes consisted of several phases, as reported in Figure 2. These phases are the following:

1. A training set was defined as a gold standard and was composed of a random sample of ED diagnoses.

2. The gold standard diagnoses were manually classified by an expert physician into injury versus noninjury (task A), intentional versus unintentional injury (task B), and type of unintentional injury (task C), according to the World Health Organization (WHO) classification of injuries.
3. The preprocessed and manually classified gold standard cases were used to train the MLTs algorithms in classifying the diagnoses according to tasks A, B, and C. Several algorithms were considered to define the MLT tools; the most performing algorithms would be considered for the predictive tool definition.
4. The trained tool served to automatically classify the remaining ED diagnoses by providing a proof of concept of the injury epidemiology in the geographical area referring to the Padua ED center.
5. Once optimal algorithms were defined, they could be used to classify diagnoses on a new ED referral pediatric patient by defining an automated epidemiological surveillance system.

Figure 2. Flowchart of learning algorithm development for injury epidemiological surveillance. ED: emergency department; MLT: machine learning technique.



Gold Standard Definition

A randomly extracted subset of 40,031 ED records was manually classified (Figure 1) by a clinician for the following 3 classification tasks:

1. Injury or noninjury events, characterized as classification task A.
2. Intentional or unintentional injury events as classification task B (performed on diagnoses classified as injury in classification task A).
3. Unintentional injury category as classification task C (performed on records classified as unintentional injury in classification task B).

In this study, unintentional injury ED records were classified based on the WHO [17] classification of 5 types of unintentional injuries: road traffic injuries, poisoning, falls, fires, and drowning. We added a sixth category of unintentional injury exclusive of the WHO categories called “other unintentional injury.” In addition, the WHO classes represented by fewer than 15 records were not considered for cross-validation of the automatic MLT classifier.

Foreign body or choking injury events were also excluded because they constitute a separate epidemiological category conducive to purpose-specific studies [18].

Data Preprocessing

Free-text diagnoses were preprocessed by removing punctuation, as reported in the literature [19], stop words, white spaces, and numbers, leaving only word stems. All words were converted to the lowercase font.

After cleaning the text corpus, the free-text data were represented via a document-term matrix (DTM) that represented the diagnostic text data in the form of a matrix. The rows of the matrix reported the sentence for the single diagnosis, and the columns of the matrix represented the single word that composed the sentence. The DTM was filled by weighting each word term with the inverse of its frequency.

The most frequent words in the ED free-text records were also reported considering the unprocessed DTM in manually classified gold standard cases.

Different MLT algorithms were trained and tuned to classify injury diagnoses, as follows:

- The decision tree (DT), random forest (RF), and gradient boosting method (GBM) tree-based models.
- The support vector machine (SVM) method served as a comparison tool against the tree-based models.

MLT Classifier Cross-Validation

The filtered manual classification of 39,576 cases (Figure 1) served as the gold standard for training and cross-validating (4-fold) the MLT tool. In the literature, it is documented that 4- or 5-fold cross-validation is appropriate to minimize the SD accuracy estimate of a tuned model [20].

Tree-Based Methods and the Comparator

Decision Trees (DTs)

DTs are classification or regression models based on a top-down methodology in which, starting from the root node, binary splits of data are generated until a certain criterion is satisfied. The classification error rate has been considered as the fraction of training observations in a particular tree partition that doesn’t belong to the most widely occurring class [21].

The DT is a classification method that in several cases suffers from overfitting; for this reason, ensemble methods are provided in the literature. For example, bagging or bootstrap aggregation is a technique for reducing the variance of an estimated prediction function. Bagging seems to work especially well with high variance and low-bias procedures, such as trees [22].

Random Forest (RF)

Rf is a modification of the bagging method that constructs a large collection of poorly correlated trees and then calculates the average. In many problems, the performance of RF is high; RFs are also easy to train and regularize. As a result, they have become quite popular [23]. The RF tree-based algorithm involves the computation of hundreds to thousands of DTs and merges them to increase the generalizability of the model. The DT combination essentially takes the form of an ensemble method. Weak learner (or single DT) pooling is used as a strategy to obtain more powerful learners [22].

Gradient Boosting Method (GBM)

The GBM is based on sequential boosting improvements of weak classifiers (high bias and low variance). The GBM sequentially adds one classifier at a time so that the next classifier is trained to improve the previously trained DT. In contrast, the RF algorithm trains each classifier independently of the others [22].

Support Vector Machine (SVM) as the Comparator

The main objective of the SVM algorithm is to find an optimal hyperplane of a feature space of N dimensions (where N is the number of features) that distinctly classifies the data points into a binary partition [24]. Several hyperplanes may separate the resulting classes of data points. The SVM algorithm considers the hyperplanes that maximize the margin (the distance between the data points of the classes). The SVM algorithm was selected as a comparator for tree-based algorithms. The kernel hyperplane approach has been considered for the computation.

Classification Tasks

Three classification tasks were considered for the analysis, as follows:

- Classification of injury versus noninjury events (task A)
- Classification of intentional versus unintentional injury events (task B)
- Classification of unintentional injuries (task C) based on the WHO categorization (poisoning, road traffic, falls, fires and burns, drowning, and other unintentional injuries) [5].

Performance Evaluation

The performance of the MLT classifiers was evaluated using cross-validated accuracy and Kappa agreement values compared to the gold standard. In particular, the training set represented by the gold standard is one of the largest such sets produced to classify injuries.

For classification scenarios that involved severe class imbalance (where the minority class is represented by less than 15% of the cases), balanced accuracy was reported.

The mean and maximum accuracies computed for all classes that included unintentional injuries in the gold standard records were also calculated. Other performance measures were reported concerning positive and negative predictive values, sensitivity, and specificity.

MLT Predictions

MLT predictions for classification tasks A, B, and C were calculated for admissions to the ED of children residing in Padua (221,175 records; Figure 1). Subsequently, Poisson 95% CIs for injury incidence rates over the Padova province resident child (aged 0-18 years) population were computed to compare the predictions of the different MLT methods. The person-time was identified in the period of 2007-2018 by considering the official Italian statistic data source ISTAT [25]. The number of cases in the period was estimated using the RF, GBM, DT, and SVM algorithms.

Synthesis of Data

Summary statistics of the gold-standard case data were reported as follows: continuous data were summarized as first quartile, median, and third quartile; categorical data were reported as percentages and absolute frequencies. Wilcoxon-type tests were performed for continuous variables, and Pearson chi-square test or Fisher exact test, as appropriate, were performed for categorical variables.

The computations were performed using R 3.4.2 (R Foundation for Statistical Computing) [26] with the caret package [27] as a machine learning R interface and the rms package [28] for descriptive and standard statistical analyses.

Results

Gold Standard Description

The gold standard set used to train the MLT classifiers (39,576 cases) was composed of 19,659 female and 19,917 male individual admissions (Table 1). The sample was mainly composed of Italians, and 33,474 (85%) gold standard cases were aged between 1 month and 15 years (Table 1).

Injury events were mainly represented by Italian male children between 6 and 15 years of age (Table 1). Among the 8232 injuries in Table 1, only 50 (0.6%) cases were intentional injuries.

Manually classified WHO unintentional injury drowning cases were not considered in the analyses because there were only 12 such cases (Figure 1). Falls and road traffic injuries were the main types of unintentional cases in the gold standard set (Table 1).

Table 1. Characteristics of the gold standard cases. Continuous data are reported as medians (first and third quartiles); categorical data are reported as percentages and absolute frequencies. Wilcoxon-type tests were performed for continuous variables; Pearson chi-square test or Fisher exact test, as appropriate, were performed for categorical variables.

Characteristics	Noninjury (N=31,344), n (%)	Injury (N=8232), n (%)	Overall (N=39,576), n (%)	*P* value
Gender				<.001
Female	15,762 (50)	3897 (47)	19,659 (50)	
Male	15,582 (50)	4335 (53)	19,917 (50)	
Age				<.001
1-28 days	4056 (13)	103 (1)	4159 (11)	
29 days-1 year	6035 (19)	702 (9)	6737 (17)	
1-3 years	5293 (17)	1388 (17)	6681 (17)	
4-5 years	5270 (17)	1387 (17)	6657 (17)	
6-10 years	4720 (15)	1945 (24)	6665 (17)	
11-15 years	4170 (13)	2564 (31)	6734 (17)	
≥16 years	1800 (6)	143 (2)	1943 (5)	
Nationality				<.001
Other countries	14,875 (47)	3395 (41)	18,270 (46)	
Italian	516,469 (53)	4837 (59)	21,306 (54)	
Season				<.001
Spring	7590 (24)	2114 (26)	9704 (25)	
Summer	7540 (24)	2381 (29)	9921 (25)	
Autumn	7783 (25)	2175 (26)	9958 (25)	
Winter	8431 (27)	1562 (19)	9993 (25)	
Manual classification of events				<.001
Noninjury	31,344 (100)	—[a]	31,344 (79)	
Injury (intentional)	—	50 (1)	50 (0)	
Injury (unintentional: poisoning)	—	176 (2)	176 (0)	
Injury (unintentional: falls)	—	589 (7)	589 (1)	
Injury (unintentional: road traffic)	—	520 (6)	520 (1)	
Injury (unintentional: fires and burns)	—	194 (2)	194 (0)	
Injury (unintentional: other)	—	6703 (81)	6703 (17)	

[a] Not applicable.

Free-Text Diagnosis Description

The free-text diagnosis field of the gold standard cases was preprocessed, and the DTM was synthesized based on word occurrence. A manual frequency evaluation of noninjury diagnoses found the most frequent words were "high," "respiratory," "tract," "inflammatory," and "fever," whereas the most frequent words in injury diagnoses were "skull," "trauma," "wound," "fracture," and "hand" (Figure S1 in Multimedia Appendix 1).

Regarding the unintentional injury classes (Figure S2 in Multimedia Appendix 1), "trauma" was the most frequent word found in falls, road traffic, and other unintentional injury diagnoses.

Burn events were mainly described with hand-related attributes such as "right" and "left." No drowning events were reported.

The words “ingestion” and “suspected” were mainly associated with poisoning diagnoses.

MLT Classifier Performance

Classification Task A: Injury Versus Noninjury

The average cross-validated accuracy reported as percentages for the different MLT classifiers was greater than 85% in every case and was very similar for the SVM, GBM, and RF models (Table 2).

A high level of agreement among the methods can be found using the kappa measure, which was greater than 79% in every case with the exception of DT (Table 2). In addition, the kappa performance of the RF and SVM methods was very similar.

Most disagreement cases stemmed from a mismatch with the gold standard, and the methods exhibited a high level of consistency in this regard. For example, the percentage of correctly evaluated injuries was very similar for the SVM and RF disagreement cases (Table S1 in Multimedia Appendix 1).

Table 2. Injury versus noninjury classification task comparative cross-validated accuracies and kappa scores of the machine learning technique (MLT) classifiers.

Feature	Random forest	Decision tree	Gradient boosting method	Support vector machine
Overall accuracy (%)	94.09	88.38	94.1	94.14
Overall kappa (%)	81.76	58.66	81.63	81.77
Sensitivity	0.980475	0.993396	0.985065	0.983218
Specificity	0.795633	0.381413	0.7518	0.790492
Positive predictive value	0.945387	0.852819	0.934727	0.944237
Negative predictive value	0.918657	0.941199	0.933117	0.928852

Classification Task B: Intentional Versus Unintentional Injury

The balanced accuracy performance was greater than 70% for RF, GBM, DT, and SVM (Table 3).

Considering the other metrics (ie, sensitivity, specificity, negative predictive values, and positive predictive values), the algorithms were able to classify unintentional cases, as the negative predicted values were greater than 99% in every unintentional case. Intentional injuries, in contrast, were misclassified in several cases, and the positive predictive values were less than 2% in every case (Table 3).

Table 3. Intentional versus unintentional injury classification task comparative cross-validated performance measures of the machine learning technique (MLT) classifiers.

Feature	Random forest	Decision tree	Gradient boosting method	Support vector machine
Balanced accuracy (%)	71.33	75.63	76.65	70.86
Sensitivity	0.6867	0.7	0.7644	0.64
Specificity	0.7401	0.7787	0.7563	0.7631
Positive predictive value	0.0151	0.018	0.0178	0.0154
Negative predictive value	0.9976	0.9978	0.9982	0.9973

Classification Task C: Unintentional Injury Category

The algorithms were trained and tuned on different subclasses of unintentional injuries, and the scores of a balanced accuracy measure were found to be relatively greater and similar for RF, SVM, and GBM, and smaller for DT (Table 4).

For the other metrics, all the algorithms correctly identified poisoning and other unspecified injuries. For these classes, the classification positive predictive values were greater than 60% in every case for the different MLT methods (Table S2 in Multimedia Appendix 1). The metrics reported in Table S2 (Multimedia Appendix 1) also revealed decreased performance for the identification of trauma-related injuries, such as “burns,” “falls,” and “road traffic,” where the sensitivity was less than 15%.

Table 4. Unintentional injury classification task (in every case, maximum accuracy is achieved for the poisoning class).

Feature	Random forest	Decision tree	Gradient boosting method	Support vector machine
Maximum balanced accuracy (%)	91.42	81.09	91.63	91.62
Mean balanced accuracy (%)	64.59	59.26	65.07	64.95

Predicted MLT Injury Incidence Rates

For the 221,175 ED visits studied (Table S3 in Multimedia Appendix 1), the median age was 4 years, and the majority of the children were of Italian nationality (172,577, 78%). The estimated number of injury cases was similar for RF, GBM, and SVM but relatively lower for DT (Table S4 in Multimedia Appendix 1).

The estimated incidence rates of ED entrance for Padova residents (2007-2018) were very similar across the GBM, RF, and SVM methods and slightly lower for DT (Figure S3 in Multimedia Appendix 1).

Discussion

Principal Findings

In Italy, unintentional injuries in children have been scarcely investigated. Padova Hospital is an important health center in Northeast Italy, characterized by a high number of daily ED visits due to unintentional injuries. Analysis of the Padova Hospital ED database represents a suitable starting point for the development of a reliable and generalizable epidemiological injury surveillance system.

The free-text classification may improve the epidemiologic surveillance of pediatric ED injuries. However, the manual classification of free-text diagnoses is often time-consuming and requires highly trained clinicians [29]. On the contrary, automated text classification approaches require relatively fixed data sources and can improve the efficiency and timeliness of ED surveillance systems [29].

Several MLT methods are currently used to perform automatic text classification. The methodological comparison of different MLTs is useful to achieve a valid and accurate text classification [30]. This study demonstrates that ensemble tree-based resampling methods (RF and GBM) and SVMs are consistent with each other [31], reporting good classification accuracy [11] over different classification tasks, as corroborated in the literature [12]. DT is known to have a high variance when using training or test sets different from the same data set because it is prone to overfitting. Moreover, the optimal choice of an MLT classifier should be integrated and tailored to gold-standard data characteristics, such as the number of classes, class imbalance, and the correlation structure of predictors. In the literature, ensemble methods (ie, RF and GBM) have been shown to be more robust in relation to these previously mentioned issues [32].

Cross-validation is a useful method that limits overfitting and allows tuning of DT parameters to optimize model accuracy [33]. The best classifier performance in this study was achieved on the task of identifying injury versus noninjury cases in ED visits (task A).

Regarding surveillance, the implications of our results are clear; current injury surveillance systems are largely based on mortality or hospital discharge data [13]. However, thousands of pediatric patients are treated in the EDs and subsequently discharged [13]. In Italy and other European and newly developed countries, ED data often contain narratives and free text to describe patient diagnoses [11]. Thus, an automated ED surveillance system, not requiring additional physician work, would be a suitable tool for comprehensive surveillance of childhood injuries. A negative predictive value of at least 99% was found to identify unintentional injuries; this indicates that there is a high probability that the cases identified as unintentional by this algorithm were unintentional. MLTs are capable of correctly classifying unintentional cases, which are highly prevalent injury events. In the literature [34], unintentional injury ED visits were found nearly 20 times more than intentional injury ED visits in the United States, and the pattern is similar in European countries [35].

The algorithms in this study performed poorly on the identification of intentional events. The reason for this poor performance was due to a lack of intentional injury cases (n=50, 0.6%) in the data [19]. Other methods are needed to develop a more accurate free-text classifier for intentional injury events. Poor performance was also evident in the distinction of trauma-related injuries (eg, falls and road traffic injuries).

MLTs (especially ensemble algorithms) have shown good classification performance in poisoning events. From an epidemiological perspective, poisoning events remain the third most common cause of unintentional injuries in Italy and Europe [4].

Large EDs are important sources of surveillance for pediatric diseases, especially for trauma and injury-related issues, given that most of such events refer to these departments [13]. However, the staff employed in such facilities often work in stressful situations, and the time and human resources to devote to data collection and accurate diagnosis coding may be very limited. In this general framework, our proposed MLT-based tool could facilitate the automatic classification of events for surveillance purposes. Once implemented, this algorithm could be easily improved by accumulating more data on less prevalent injury categories. It is hereby possible to obtain a general overview of the phenomenon on the territory by monitoring its epidemiological evolution over time. This system could facilitate the timely activation of intervention policies, regardless of the alarming concentrations of injury events.

Moreover, it is also important to improve surveillance systems using classified ED data integrated with hospital discharge or mortality records to design effective injury prevention programs and interventions. In this general context, the proposed ML-based injury classification tool could be a first step toward addressing the burden of pediatric injuries from a new holistic perspective [36].

Limitations

One first limitation of this study is that the data used for injury classification provide little information on what happens between ED admission and the final diagnosis. Moreover, as the triage service is extremely operator dependent, human factors represent an important confounding aspect of injury classification.

Another possible limitation is the small prevalence of certain types of injury, such as intentional injuries and drowning (among unintentional ones); this issue makes the algorithm's performance on these types of events lacking. The injury and unintentional injury classifications constitute the leading classification task for this research; however, further research developments are needed to enrich the diagnosis data on these types of injuries and train the classification machine for a more refined surveillance tool. Moreover, the poorly represented classes of unintentional injuries constitute an issue to be deepened from a technical standpoint. Within this framework, further research is needed to develop algorithms tailored to handle severe class imbalance.

Another point to explore is the generalization of the algorithm; the MLT performance may be influenced by a training process performed on diagnoses data retrieved from the same center, where the referring physicians could maintain the same writing style across the data set. For this reason, data from other centers would be needed to generalize the validity of the epidemiological tool. Despite this limitation, this tool constitutes a proof of concept of an epidemiological surveillance attempt performed using a machine trained on data from a large pediatric ED referral center in Northeast Italy.

Conclusions

This research paper reports an MLT-based free-text classification application conducted for the epidemiological surveillance of pediatric injuries. The algorithms have been trained considering the free-text diagnoses data of the Padova University Hospital ED unit, a large referral center in Northeast Italy.

The results of this study, for the injury classification task, showed that MLTs are a promising tool for improving epidemiological surveillance, allowing for the characterization of pediatric injuries in the ED by considering the free-text diagnoses as data sources.

The reported classification performance is satisfactory, especially for general injuries and intentional injury classification. These research results could facilitate the surveillance of a phenomenon that is often not easy to identify. Moreover, the approach could save time for health professionals working in the ED in manually classifying diagnoses for research purposes.

Data Availability

The database is provided from an administrative data source of the University Hospital of Padua. They could be made available upon motivated request to the authors.

Authors' Contributions

DA, DG, and SB prepared the original draft. DA, GAB, GL, CL, FS, AF, and PB wrote, reviewed, and edited the manuscript. Formal analysis was conducted by DA. The methodology was designed by DA and DG. The study was supervised by DG and LDD.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[DOCX File , 127 KB - publichealth_v9i1e44467_app1.docx]

References

1. Alonge O, Khan UR, Hyder AA. Our shrinking globe: implications for child unintentional injuries. Pediatr Clin North Am 2016 Feb;63(1):167-181. [doi: 10.1016/j.pcl.2015.08.009] [Medline: 26613695]
2. Deal LW, Gomby DS, Zippiroli L, Behrman RE. Unintentional injuries in childhood: analysis and recommendations. Future Child 2000;10(1):4. [doi: 10.2307/1602823]
3. Judy K. Unintentional injuries in pediatrics. Pediatr Rev 2011 Oct;32(10):431-8; quiz 439. [doi: 10.1542/pir.32-10-431] [Medline: 21965710]
4. Sethi D, Towner E, Vincenten J, et al.. European report on child injury prevention. Copenhagen: World Health Organization. Regional Office for Europe; 2008.
5. WHO. World report on child injury prevention. URL: https://www.who.int/publications/i/item/9789241563574 [accessed 2019-02-04]
6. Krug EG, Sharma GK, Lozano R. The global burden of injuries. Am J Public Health 2000 Apr;90(4):523-526. [doi: 10.2105/ajph.90.4.523] [Medline: 10754963]
7. Nijman S, Blatter B. SM 04-1421 Can autotext regcognition software for coding injuries replace manual coding? findings from IDB/DISS in the netherlands. Inj Prev BMJ Publishing Group Ltd 2018;24(Suppl 2):A266. [doi: 10.1136/injuryprevention-2018-safety.736]
8. Mitchell RJ, McClure RJ, Williamson AM, McKenzie K. Implementing the national priorities for injury surveillance. Med J Aust 2008 Apr 07;188(7):405-408. [doi: 10.5694/j.1326-5377.2008.tb01685.x] [Medline: 18393744]

9. Hyder AA, Sugerman DE, Puvanachandra P, Razzak J, El-Sayed H, Isaza A, et al. Global childhood unintentional injury surveillance in four cities in developing countries: a pilot study. Bull World Health Organ 2009 May;87(5):345-352 [FREE Full text] [doi: 10.2471/blt.08.055798] [Medline: 19551252]
10. Hochdorn A, Oliveira A, Lorenzoni G, Francavilla A, Baldas S, Berchialla P, et al. Monitoring public perception of health risks in Brazil and Italy: cross-cultural research on the risk perception of choking in children. Children (Basel) 2021 Jun 24;8(7):541 [FREE Full text] [doi: 10.3390/children8070541] [Medline: 34202693]
11. Lorenzoni G, Bressan S, Lanera C, Azzolina D, Da Dalt L, Gregori D. Analysis of unstructured text-based data using machine learning techniques: the case of pediatric emergency department records in Nicaragua. Med Care Res Rev 2021 Apr 29;78(2):138-145. [doi: 10.1177/1077558719844123] [Medline: 31030615]
12. Vallmuur K, Nanda G, Lehto M. 347 Using machine learning to categorise Emergency Department data for product safety surveillance. Inj Prev 2016 Sep 01;22(Suppl 2):A127.2-A1A127. [doi: 10.1136/injuryprev-2016-042156.347]
13. Adirim TA, Wright JL, Lee E, Lomax TA, Chamberlain JM. Injury surveillance in a pediatric emergency department. Am J Emerg Med 1999 Oct;17(6):499-503. [doi: 10.1016/s0735-6757(99)90184-5] [Medline: 10530522]
14. Sebastiani F. Machine learning in automated text categorization. ACM Comput Surv 2002 Mar;34(1):1-47. [doi: 10.1145/505282.505283]
15. Morrison A, Stone DH, Doraiswamy N, Ramsay L. Injury surveillance in an accident and emergency department: a year in the life of CHIRPP. Arch Dis Child 1999 Jun;80(6):533-536 [FREE Full text] [doi: 10.1136/adc.80.6.533] [Medline: 10332002]
16. Ocagli H, Azzolina D, Bressan S, Bottigliengo D, Settin E, Lorenzoni G, et al. Epidemiology and trends over time of foreign body injuries in the pediatric emergency department. Children (Basel) 2021 Oct 19;8(10):938 [FREE Full text] [doi: 10.3390/children8100938] [Medline: 34682203]
17. Sminkey L. World report on child injury prevention. Inj Prev 2008 Feb;14(1):69-64. [doi: 10.1136/ip.2007.018143] [Medline: 18245322]
18. Passali D, Gregori D, Lorenzoni G, Cocca S, Loglisci M, Passali FM, et al. Foreign body injuries in children: a review. Acta Otorhinolaryngol Ital 2015 Oct;35(4):265-271. [Medline: 26824213]
19. Lanera C, Minto C, Sharma A, Gregori D, Berchialla P, Baldi I. Extending PubMed searches to ClinicalTrials.gov through a machine learning approach for systematic reviews. J Clin Epidemiol 2018 Nov;103:22-30. [doi: 10.1016/j.jclinepi.2018.06.015] [Medline: 29981872]
20. Moss H, Leslie D, Rayson P. Using JK fold cross validation to reduce variance when tuning NLP models. ArXiv Preprint posted online Jun 19, 2018. [doi: 10.48550/arXiv.1806.07139]
21. Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software; 1984.
22. Flaxman AD, Vahdatpour A, Green S, James SL, Murray CJ, Population HMRC. Random forests for verbal autopsy analysis: multisite validation study using clinical diagnostic gold standards. Popul Health Metr 2011 Aug 04;9:29 [FREE Full text] [doi: 10.1186/1478-7954-9-29] [Medline: 21816105]
23. Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2(3):18-22.
24. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: 10.1007/BF00994018]
25. ISTAT Demography. URL: http://demo.istat.it/ [accessed 2023-05-24]
26. The R Project for Statistical Computing. URL: https://www.R-project.org [accessed 2023-05-24]
27. Kuhn M. Building predictive models in R using the caret package. J Stat Soft 2008;28(5):1-26. [doi: 10.18637/jss.v028.i05]
28. Harrell F. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer-Verlag; 2001.
29. Maenner MJ, Yeargin-Allsopp M, Van Naarden Braun K, Christensen DL, Schieve LA. Development of a machine learning algorithm for the surveillance of autism spectrum disorder. PLoS One 2016;11(12):e0168224 [FREE Full text] [doi: 10.1371/journal.pone.0168224] [Medline: 28002438]
30. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med 2001 Aug;23(1):89-109. [doi: 10.1016/s0933-3657(01)00077-x] [Medline: 11470218]
31. AhmedK A, Aljahdali S, Naimatullah Hussain S. Comparative prediction performance with support vector machine and random forest classification techniques. IJCA 2013 May 17;69(11):12-16. [doi: 10.5120/11885-7922]
32. Zhou L. Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. Knowl Based Syst 2013 Mar;41:16-25. [doi: 10.1016/j.knosys.2012.12.007]
33. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. New York: Springer; 2009:745.
34. Ballesteros M, Williams D, Mack K, Simon T, Sleet D. The epidemiology of unintentional and violence-related injury morbidity and mortality among children and adolescents in the United States. Int J Environ Res Public Health 2018 Mar 28;15(4):616 [FREE Full text] [doi: 10.3390/ijerph15040616] [Medline: 29597289]
35. Rockett IR, Regier MD, Kapusta ND, Coben JH, Miller TR, Hanzlick RL, et al. Leading causes of unintentional and intentional injury mortality: United States, 2000–2009. Am J Public Health 2012 Nov;102(11):e84-e92. [doi: 10.2105/ajph.2012.300960]

36. Elboray S, Elawdy MY, Dewedar S, Elezz NA, El-Setouhy M, Smith GS, et al. Knowledge, attitudes, and practices of family physicians and nurses regarding unintentional injuries among children under 15 years in Cairo, Egypt. Int J Inj Contr Saf Promot 2017 Mar;24(1):24-31 [FREE Full text] [doi: 10.1080/17457300.2015.1056808] [Medline: 26176681]

Abbreviations

DT: decision tree
ED: emergency department
GBM: gradient boosting method
MLT: machine learning technique
RF: random forest
SVM: support vector machine
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 20.11.22; peer-reviewed by N Jiwani, S Raiden; comments to author 14.02.23; revised version received 07.03.23; accepted 23.03.23; published 12.07.23.

Please cite as:
Azzolina D, Bressan S, Lorenzoni G, Baldan GA, Bartolotta P, Scognamiglio F, Francavilla A, Lanera C, Da Dalt L, Gregori D
Pediatric Injury Surveillance From Uncoded Emergency Department Admission Records in Italy: Machine Learning–Based Text-Mining Approach
JMIR Public Health Surveill 2023;9:e44467
URL: https://publichealth.jmir.org/2023/1/e44467
doi:10.2196/44467
PMID:37436799

Original Paper

Enhancing the Predictive Power of Google Trends Data Through Network Analysis: Infodemiology Study of COVID-19

Amanda MY Chu[1], PhD; Andy C Y Chong[2], MSc; Nick H T Lai[3], BSc; Agnes Tiwari[4,5], PhD; Mike K P So[3], PhD

[1]Department of Social Sciences and Policy Studies, The Education University of Hong Kong, Hong Kong, Hong Kong

[2]School of Nursing, Tung Wah College, Hong Kong, Hong Kong

[3]Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Hong Kong, Hong Kong

[4]School of Nursing, Hong Kong Sanatorium & Hospital, Hong Kong, Hong Kong

[5]Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, Hong Kong

Corresponding Author:
Mike K P So, PhD
Department of Information Systems, Business Statistics and Operations Management
The Hong Kong University of Science and Technology
Clear Water Bay
Hong Kong
Hong Kong
Phone: 852 23587726
Email: immkpso@ust.hk

Abstract

Background: The COVID-19 outbreak has revealed a high demand for timely surveillance of pandemic developments. Google Trends (GT), which provides freely available search volume data, has been proven to be a reliable forecast and nowcast measure for public health issues. Previous studies have tended to use relative search volumes from GT directly to analyze associations and predict the progression of pandemic. However, GT's normalization of the search volumes data and data retrieval restrictions affect the data resolution in reflecting the actual search behaviors, thus limiting the potential for using GT data to predict disease outbreaks.

Objective: This study aimed to introduce a merged algorithm that helps recover the resolution and accuracy of the search volume data extracted from GT over long observation periods. In addition, this study also aimed to demonstrate the extended application of merged search volumes (MSVs) in combination of network analysis, via tracking the COVID-19 pandemic risk.

Methods: We collected relative search volumes from GT and transformed them into MSVs using our proposed merged algorithm. The MSVs of the selected coronavirus-related keywords were compiled using the rolling window method. The correlations between the MSVs were calculated to form a dynamic network. The network statistics, including network density and the global clustering coefficients between the MSVs, were also calculated.

Results: Our research findings suggested that although GT restricts the search data retrieval into weekly data points over a long period, our proposed approach could recover the daily search volume over the same investigation period to facilitate subsequent research analyses. In addition, the dynamic time warping diagrams show that the dynamic networks were capable of predicting the COVID-19 pandemic trends, in terms of the number of COVID-19 confirmed cases and severity risk scores.

Conclusions: The innovative method for handling GT search data and the application of MSVs and network analysis to broaden the potential for GT data are useful for predicting the pandemic risk. Further investigation of the GT dynamic network can focus on noncommunicable diseases, health-related behaviors, and misinformation on the internet.

(JMIR Public Health Surveill 2023;9:e42446) doi:10.2196/42446

KEYWORDS

internet search volumes; network analysis; pandemic risk; health care analytics; network connectedness; infodemiology; infoveillance; mobile phone; COVID-19

Introduction

Background

Since the discovery of the first case of COVID-19 in December 2019, the pandemic has continued to develop and spread globally for >3 years. With the new wave of COVID-19 that began in early 2022 owing to the Omicron variant [1], confirmed cases rose to 540 million and deaths rose to 6.3 million by the end of June 2022 [2]. To deal with such serious disease outbreaks, real-time disease surveillance plays a crucial role for policy makers in their efforts to effectively implement timely health measures and allocate resources. However, disease surveillance generally relies on laboratory testing results [3], which are limited by testing time and capacity [4]. Relevant clinical data, such as data for mortality, morbidity, and utility rate of the health system, are regarded as another layer of reliable population-scale indicators for tracking the spread of a pandemic [5]. Unfortunately, these official statistics usually lag behind the infection situation [4,6]. As a result, there is a need to explore other possible population-scale data as complementary sources for tracking pandemic risk in real time without the limitations of the traditional surveillance methods.

Internet use is a common human behavior worldwide, as reflected in the rapidly growing number of internet users and social media users globally [7]. In particular, the internet has become an important source of health care information [8]. Therefore, internet search data are now a valuable population-level resource for tracking disease development, complementing the traditional surveillance methods [9]. Consequently, an increasing number of infodemiological studies have sought to address the need for public health enhancement by applying internet user–contributed, health-related content. Investigations of infoveillance, an emerging type of public health syndromic surveillance that is based on information from the web [10], have been proposed for tracking various health issues [11-13].

Google Trends (GT) is a Google website that analyzes the popularity of search queries and thus provides a platform for the public's use in retrieving the real-time search patterns of internet users worldwide [14]. As an official Google service, the aggregate GT data range from multiple sources of Google search engines, including web searches, image searches, news searches, Google shopping, and YouTube searches. There has been an upward trend in public health and epidemiological research on the use of GT in the last decade [15], and GT data have been proven to be reliable forecast and nowcast measures for official statistics and social topics, such as unemployment [16] and waves of pathogenic infections [17,18]. As the data in the repository are freely available and updated in real time, GT is considered to be a timely infoveillance tool that can complement the traditional surveillance methods well [9,19]. In particular, because the impact of an infectious disease will likely attract the public's attention, many previous studies have attempted to investigate the associations between search volumes of internet queries and disease outbreak trends, such as with the outbreaks of Ebola [20] and Middle East respiratory syndrome coronavirus [21]. A methodology framework was also developed for working with GT data in infodemiology and infoveillance [22]. However, although web-based data are a good indicator for predicting changes in human behaviors [23], the application of GT data to health care research is relatively novel and has been evolving quickly only recently [24,25].

Although access to GT data is free, there are several limitations in retrieving the data. GT does not provide the actual number of searches for a particular topic of interest at a particular time and in a particular geographic region. Instead, it only shows "interest over time" by providing relative search volume (RSV) time series data, normalized to a time and geographic region, using a scale from 0 to 100. A value of 100 represents the peak popularity of a particular search interest at a particular time and location, and a value of 30 indicates that a particular search interest was 30% as popular as that of the peak search activity. Therefore, GT data over long time horizons will result in heavily normalized data with reduced resolution [18], thus rendering the gathered data noisy and limiting the predictive power of the data. The normalization of the GT data does offer convenience in comparison, but it also poses limitations in tracking search behaviors longitudinally with a high frequency, such as daily updating of search volumes. Early research on GT data also indicated that it is a nontrivial endeavor to aggregate information on GT data from multiple keywords. Therefore, there is a need to search for an effective method of retrieving GT data that increases the resolution of RSV to reflect a trend of interest and that integrates search data from multiple keywords.

Previous studies have tended to use the retrieved GT data of specified search queries for correlation analyses and modeling [9,11,19,23,26]. Many previous studies have been conducted on COVID-19–related keywords [18,27-30] and have demonstrated the importance of using multiple keywords. Although the inclusion of multiple search queries can help provide valuable information for tracking human behavior and public health issues [11], the linkage among the search queries remains a challenge for empirical studies seeking to analyze the data systematically [31]. Furthermore, we cannot ignore the potential of connectedness among search interests in reflecting the complexity of information searching behaviors among internet users. Understanding the complex search behavior will be crucial when using search data to develop an effective infoveillance system for infection outbreaks.

Objectives

This study sought to extend the application of GT data in predicting disease outbreaks by demonstrating the application of merged internet search volumes and to propose a dynamic network approach to tracking COVID-19 pandemic risk. Specifically, we introduced the construction of GT networks over time, in which the network nodes were represented by different keywords. Our proposed GT network method can help incorporate internet search information from multiple keywords into a network. Using GT network structures and related network statistics, we were able to study the dependence among keywords—for example, COVID-19–related words—and learn topological features of search volumes of different keywords. We also examined the predictive power of the proposed application of our GT network method through trend analysis and dynamic time warping (DTW). In the case of COVID-19, which we studied in this research, we investigated the implications of topological features on the COVID-19 pandemic risk.

Methods

This section describes the methodology used in this study. A set of custom Python scripts was used for data acquisition, cleansing, analysis, and most of the data visualization. *Gephi* 0.9.7 from Gephi Consortium [32] was used to create the base network graphs.

Data Acquisition

In this study, the aggregated data from GT were retrieved for analysis. We used the time series data of “interest over time” as the RSV. The numerical data, ranging between 0 and 100, indicated the “search interest relative to the highest point” for the defined filtering region and the period. However, according to Google’s own definition, the value of “50” means the search term is “half as popular,” whereas “0” means low search volume was recorded for the term (ie, <1% of the searches attributed to the most searched term in the search list) [33]. This reflects that the “interest over time” can be regarded as a latent variable of the actual volume of searching. However, noise may be introduced through rounding and normalization. Although noise is unavoidable, we attempted to alleviate the effects of noise through a shingling technique, as proposed in the *Transforming Multiple RSVs Into a Single MSV* section, and to focus on the analysis of the constructed temporal network.

Discovery of Related Keywords

Many previous studies used “related queries” to widen their scope of keyword coverage. Although this approach increases the number of search terms drastically, we opted to use “related topics” (RTs) because of their unambiguous characteristics and the inclusion of translations of the same concept into other languages.

For example, if a researcher is interested in understanding the search volume for Apple Inc, the computer and smartphone manufacturing company, when one enters only the word “apple,” GT will return the results for that exact word, without much interpretation. Therefore, the resulting data can hypothetically include a news report for the computer company’s latest product and the nutritional information for the fruit. Instead, researchers can search for a “topic” that filters out searches irrelevant to the computer manufacturer by picking “Apple (Technology Company).”

However, to the best of our knowledge, GT does not provide a public repository of the available topics. Yet, when a user enters certain keywords, GT will recommend a list of related queries and topics that can be used as our starting seed topics. On the basis of this method, we entered several queries, namely, “COVID-19” and “coronavirus,” which yields 3 RTs: “Coronavirus disease 2019 (Disease),” “Severe acute respiratory syndrome coronavirus 2 (Virus),” and “Coronavirus (Virus).” We continuously downloaded the RTs of the 3 seed topics we yielded in a 7-day sliding window from January 29, 2020, to March 4, 2022. All the keywords prompted by GT were included in the pool of keywords.

To remove the regional effects from our results, keywords related to specific regions were excluded. We determined that a term was region related if it was categorized by GT as one of the following: “country,” “state,” “province,” “prefecture,” “municipality,” “island,” “region,” “county,” “capital,” “city,” “autonomous community,” and “town.” For example, when we receive an RT “Stay-at-home order (Topic)” (machine tag: “/g/11hf9srvdz”) from GT, we will include it in our network, as it was determined as a “Topic” and is not in our exclusion category list. In contrast, if we receive an RT “Germany (Country in Europe)” (machine tag: “/m/0345h”) from GT, we will discard the topic, as the topic was categorized as “country” by Google, which is included in our exclusion category list.

After filtering, 228 topics were selected for this research, referred to as the topic set (see Textbox 1 for the search terms used).

Textbox 1. Selected keywords related to the COVID-19 pandemic, used in the construction of our temporal network.

Coronavirus, Vaccine, COVID-19 vaccine, COVID-19 testing, Test, Signs and symptoms, Symptom, Symptoms of COVID-19, Strain, Severe acute respiratory syndrome coronavirus 2, Worldometers, Statistics, Variant, AZD1222, AstraZeneca, Nasal congestion, Mutation, Emmanuel Macron, Traffic light, GOV.UK, Bill, Preventive healthcare, Northern Beaches Council, Nueva Cepa, Certificate, Virus, Pfizer, BNT162b2, Side effect, Rapid antigen test, Bell's palsy, Transmission, Isolation, Ivermectin, Food and Drug Administration, mRNA-1273, Centers for Disease Control and Prevention, COVID-19 pandemic in New York City, Polymerase chain reaction, Guideline, British Columbia, CVS Pharmacy, Drive-through, CVS Health, Mink, Regions of Italy, Galicia, TousAntiCovid, contagium, Tier 1 network, Blood type, Tier 2 network, Nasopharyngeal swab, 2020 coronavirus pandemic in Scotland, Incubation period, Case, Donald Trump, 2019–20 coronavirus pandemic related application, NHS COVID-19, National Health Service, Dashboard, Common cold, Silvio Berlusconi, Pandemic, Medical test, Departments of France, Serology, Infection prevention and control, Swine influenza, Herman Cain, Hydroxychloroquine, Provinces of Vietnam, COVID-19 pandemic in Victoria, Sore throat, University of Oxford, COVID-19 party, Jair Bolsonaro, World Health Organization, Airborne transmission, 2020 coronavirus pandemic in Canada, Worldometer COVID-19 Dashboard, County, COVID-19 pandemic in Texas, Asymptomatic, Disease outbreak, Rapid diagnostic test, Johns Hopkins University, Mortality rate, Antibody, Maharashtra Police, DG Police Office Mumbai, COVID-19 pandemic in West Bengal expected, Oise, Lysol, Infection, 2019–20 coronavirus pandemic, Death, Chinese language, Influenza, Pneumonia, China virus, Cruise ship, Cure, 2020 coronavirus pandemic in Singapore, Pangolins, Coronavirus disease 2019, SARS, Antigen, Immunoglobulin G, Variety, Prefectures of Japan, Coronavirus Alpha variant, Immunity, RNA, RNA virus, Peplomer, Roche Holding AG, Severe acute respiratory syndrome coronavirus, Roche, Disease cluster, Genome, Yahoo Japan Corporation, Robert Koch Institute, Medical laboratory, Ebola, Avian influenza, Breaking news, Nucleic acid test, NHK, Municipalities of Japan, Federal Office of Public Health of the Swiss Confederation, Test method, Immune system, Immunoglobulin M, Falling Number, Südwestrundfunk, interactive map, Subsidy, Remdesivir, Microscope, Molecular biology, Study group, Oct-04, Bulletin board system, Typhoon, Host, Lockdown, DNA, Hoax, Risk, Dwayne Johnson, 爆サイ, Assay, Shinzo Abe, South African Revenue Service, PTT Bulletin Board System, Reverse transcription polymerase chain reaction, Severance package, Health facility, Dabie bandavirus, Sequela, Inpatient care, Health professional, Go To Campaign, COVID-19 Contact-Confirming Application, Jul-02, U.S. state, Middle East respiratory syndrome–related coronavirus, Chemical structure, Basic reproduction number, Micrometre, HIV/AIDS, HIV, Accuracy and precision, Potency, Icelandic language, Interactivity, Wind wave, SARS outbreak, Genetics, Infectious disease, Bats, Patent, MERS, Spanish flu, Tasuku Honjo, Robert Koch, Kumiko Okae, Ministry of Health, Labour and Welfare of Japan, Website about COVID-19 pandemic in Vietnam, temporarily closed school, Case fatality rate, Mask, event, biological weapon, Diamond Princess, Vaccination rule, Point-of-care testing, Stay-at-home order ordinance, Incidence, Tagesschau, Districts of Germany, hot spot, Press conference, Podkarpackie Voivodeship, Dashboard, National Institute for Public Health and the Environment, Upper Austria, Child benefit, Demonstration, Travel warning, COVID-19 pandemic in Aichi prefecture, Amitabh Bachchan, Tönnies Holding, Patanjali Ayurved, Mayu Watanabe, Logistics, Public transport, Hygiene, Public health, Emergency medical services, Statistic, Advice, and News ticker

Downloading Search Volume Data From GT

To systematically retrieve data from GT, we used the library *pytrends* [34]. We executed the custom script regularly to collect the RSV data. Each request that was sent gathered a set of 30-day “interest over time” data :

For example, the RSV data gathered for the “Coronavirus (virus)” from January 1, 2020, to January 30, 2020 were:

We then repeated the requests to collect data consecutively. We list all the terms in Textbox 1 (topic set). Our data collection period for this research was from January 1, 2020 (t= 2020-01-30) to January 31, 2022 (t= 2022-01-31).

Merging of Multiple Short Time Series Trends Into a Glandular Time Series Over a Longer Period

Inspired by Park et al [35], we devised a range-based merging algorithm.

Let be an n-day actual search volume time series for a topic on the days lying in the interval , and let $X_{i,t}$ be the actual search volume for a topic i on date t:

Assuming that GT applied the normalization with a common correction factor C_t to the entire actual search volume time series, the RSV returned can be considered as :

For any 2 consecutive GT RSV data sets, and , where $v = u + 1$ and have the same data collection duration of n days, based on our previous assumption, and are normalized with the correction factors C_u and C_v within the interval $[u - (n - 1), u]$ and $[v - (n - 1), v]$, respectively.

Given $X_{(i,v)} = X_{(i,u+1)}$:

where

Through the moment estimator of , we can subsequently obtain the estimate of with respect to the RSV , :

For example, considering the correction factor of the data point for the term “Coronavirus (virus)” on January 31, 2020, from 2 RSV series of n=30 is:

Thus, the adjusted data point for “Coronavirus (virus)” on January 31, 2020, is as follows:

We can thus obtain a 31-day merged search volume (MSV) for “Coronavirus (virus),” as shown in Table 1.

Through multiple iterations of estimations using a total of *p* RSV time series, as shown in Tables 2 and 3, we can finally obtain an MSV time series for the search term from *t* − (*n* − 1) to *t* + (*p* − 1):

Continuing from our previous example, we can eventually obtain the MSV time series for “Coronavirus (virus)” as follows:

In this study, we opted to use RSV time series data with a length of 30 days (n=30), and the RSV time series for January 30, 2020, was used as the baseline against which other subsequent RSV series were calibrated.

Table 1. A sample calculation of the correction factor and adjusted data point of January 31, 2020, using data sets from January 1, 2020, to January 30, 2020, and from January 2, 2020, to January 31, 2020. Data set keyword: “Coronavirus.”

Date	Data downloaded on January 30, 2020	Data downloaded on January 31, 2020	Correction factor for new data point on January 31, 2020	Merged Google Trends time series
January 1, 2020	0	Not available in the sequence	Out of scope	0
January 2, 2020	0	0	0 / 0 = 1.000	0
January 3, 2020	0	0	0 / 0 = 1.000	0
...	...	...	...	...
January 27, 2020	75	78	75 / 78 = 0.962	75
January 28, 2020	90	91	90 / 91 = 0.989	90
January 29, 2020	89	88	89 / 88 = 1.011	89
January 30, 2020	100	100	100 / 100 = 1.000	100
January 31, 2020	Not available in the sequence	100		100 × 1.003 = *100.3*

Table 2. A sample calculation of the correction factor and adjusted data point of February 1, 2020, using data sets from January 2, 2020, to January 31, 2020, and from January 3, 2020, to February 1, 2020. Data set keyword: "Coronavirus."

Date	Data downloaded on January 31, 2020	Data downloaded on February 1, 2020	Correction factor for new data point on February 1, 2020	Merged Google Trends time series
January 2, 2020	0	Not available in the sequence	Out of scope	0
January 3, 2020	0	0	0 / 0 = 1.000	0
...	...	...	...	...
January 27, 2020	78	77	78 / 77 = 1.013	75
January 28, 2020	91	91	91 / 91 = 1.000	90
January 29, 2020	88	89	88 / 89 = 0.989	89
January 30, 2020	100	95	100 / 95 = 1.053	100
January 31, 2020	100	100	100 / 100 = 1.000	100 × 1.003 = 100.3
February 1, 2020	Not available in the sequence	72		72 × 1.008 = *72.60*

Table 3. A sample calculation of the correction factor and adjusted data point of February 1, 2022, using data sets from January 1, 2022, to January 30, 2022, and from January 2, 2022, to January 21, 2022. Data set keyword: "Coronavirus."

Date	Data downloaded on January 30, 2022	Data downloaded on January 31, 2022	Correction factor for new data point on January 31, 2022	Merged Google Trends time series
January 1, 2022	55	Not available in the sequence	Out of scope	76 × 0.3007 = 22.85
January 2, 2022	64	60	24.88 / 60 = 0.4146	90 × 0.2764 = 24.88
...	...	...	...	...
January 26, 2022	94	90	38.17 / 90 = 0.4241	95 × 0.4018 = 38.17
January 27, 2022	89	87	36.14 / 87 = 0.4154	88 × 0.4107 = 36.14
January 28, 2022	86	82	33.87 / 82 = 0.4130	82 × 0.4107 = 36.14
January 29, 2022	83	77	32.08 / 77 = 0.4167	80 × 0.4010 = 33.87
January 30, 2022	79	76	31.77 / 76 = 0.4181	79 × 0.4022 = 31.77
January 31, 2022	Not available in the sequence	78		78 × 0.4149 = *32.36*

Construction of Daily Search Volume Networks Using Merged GT Time Series Data

Let N be the size of the search term set. We considered the temporal network construction in an R-day rolling window. To better focus on the changes of MSVs and to ease the magnitude of differences between MSVs, we followed the studies by Chu et al [36] and So et al [37] to obtain square-rooted-differenced data. Thus, the square-rooted-differenced MSV matrix ($\mathbf{Y}_t$) at time t is as follows:

A covariance matrix (ρ_t) of each search term in the square-rooted-differenced MSV matrix ($\mathbf{Y}_t$) can be determined, where the correlation between the search terms i and j at time t is .

Using the correlation matrix ρ_t, we can then construct an undirected network of search terms with an adjacency matrix $\mathbf{A}_t$ formation at time t as:

In this study, we used a rolling window length (R) of 30 days to construct the network.

Calculation of Network Statistics

Network Density

On the basis of the method proposed by Chu et al [36], we defined the network density (D_t) of a network at time t as follows:

where E_t is the number of edges in the network at time t and C_t is the number of possible connections between all nodes (ie, search terms) at time t. D_t assesses the connectedness of the nodes present at time t. Higher values represent higher connectedness at a given moment.

Global Clustering Coefficient

As defined by So et al [37], the global clustering coefficient (GC_t) was calculated based on the local clustering coefficients $c_{i,t}$ of all nodes in the network at the time (ie,). We first defined the calculation of the local clustering coefficients for the search term i as follows:

where $e_{i,t}$ is the “number of connected pairs among the neighbors of vertex i at time t” and $k_{i,t}$ is the “number of neighbors for each vertex.”

Thus, the global clustering coefficient (GC_t) can be calculated as follows:

GC_t represents the ratio of the number of established connections (edges) to the number of possible connections. The higher the value of the coefficient, the more interconnected the network nodes are.

Calculation of DTW

DTW is an algorithm for measuring the similarity between 2 time series by determining the shortest total distances between the 2 series [38]. It is useful to detect the leader-follower interaction. We used the Python library *dtw-python* (version 1.1.12) to create the DTW-aligned plots [39]. The algorithm in the library associates a node from the query time series (in our case, network statistics, network density, and global clustering coefficients) to a node from the template time series (confirmed COVID-19 cases in and severity risk score [SRS]) with the lowest distances.

Calculation of DTW Metric

To quantify the DTW results, we adopted the DTW-based metric proposed by Laperre et al [40]. The proposed algorithm augments the existing DTW algorithm by including a “forward-looking only” constraint, that is, a latent variable (such as network density and global clustering coefficient) observed at time $t + p$ can only be mapped to a historical or the current observation (such as number of COVID-19 cases or SRSs) from t to $t + p$, where the time shift value . A time shift value of 0 indicates that the latent variable moves in sync with the observation, whereas a value >0 indicates a leading effect of the latent variable on the observation.

Calculation of the SRS

Adopted from the study by So et al [37], the SRS represents the presymptomatic transmission owing to possible interaction between the susceptible population in one country and currently infected cases in another country, before the confirmed cases are identified and force isolated or quarantined. Higher scores indicate stronger signals that the pandemic is uncontrolled, which could be regarded as a systematic risk assessment.

Ethical Considerations

Ethics approval is not required as no individually identifiable information was gathered, processed, or analyzed in this research.

Results

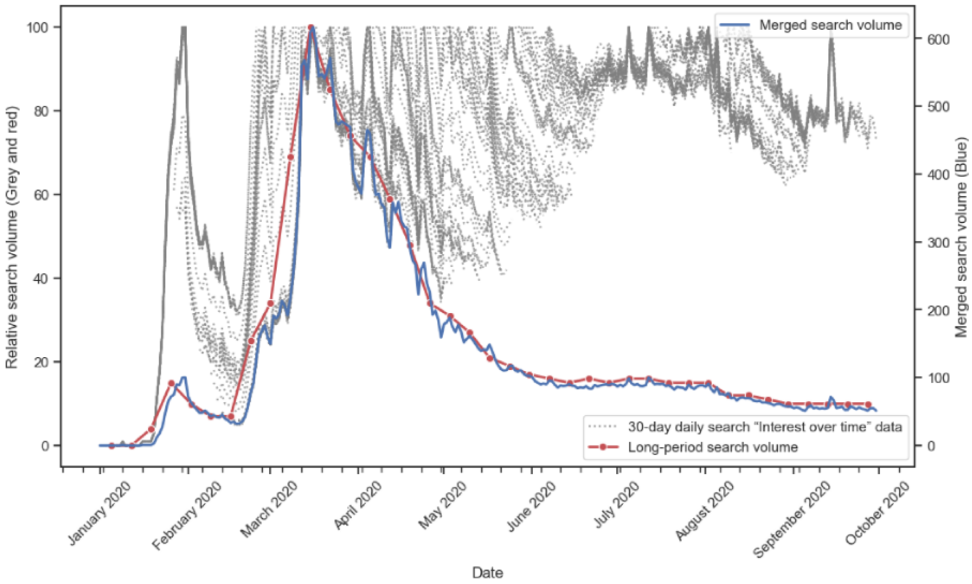
Transforming Multiple RSVs Into a Single MSV

To extract a search volume time series to reflect the trend of a specific search interest, we adopted a data transformation algorithm to construct MSVs using the rolling window approach. The primary idea was to produce a long time series of MSVs that could serve as a proxy for the search volume of keywords. Thus, the MSV series can be used to draw statistical inferences and predictions. MSV was produced by aggregating the RSV data while also exhibiting the trend indicated by the RSVs. With MSV, we were able to extend the investigation horizon beyond 9 months for daily observations and beyond 270 weeks for weekly observations, which was of interest because 9 months and 270 weeks are the longest periods for retrieving daily and weekly RSV data, respectively, from GT. Therefore, MSVs could be very helpful for tracking search interests over a long period in public health studies, such as in following the status of the COVID-19 pandemic, which has now lasted for >2.5 years.

We focused on “interest over time” as the proxy for the actual search volumes in Google. In this study, we did not specify any geographic regions but instead adopted the global web search data to obtain a more generic and globally inclusive searching phenomenon. Using these filter settings, the daily RSV data were extracted from GT for further data transformation, following the calculation procedure shown in Tables 1-3. Using the ratio of the “interest over time” data points for the same date in the 2 consecutive sequences, we calculated the correction factors for each day that existed in both sequences (as tabulated in the fourth column in the tables) to obtain the averages of the factors (as shown in the last cell in the fourth columns). The newly discovered data points (in this example, the data points on January 31, 2020; February 1, 2020; and January 31, 2022, respectively) were multiplied by the appropriate average correction factors (in this example, 1.003, 1.008, and 0.4149, respectively) to obtain the adjusted data point in the merged series. This process was repeated consecutively, as shown in Tables 1-3.

We further visualized and herein presented the process of creating the MSV series in Figure 1. Using the GT RSV data of the COVID-19 pandemic during the first 9 months as a demonstration, the 30-day “interest over time” data set (shown as 245 dotted lines) for the dates between January 1, 2020, and September 30, 2020, was downloaded from GT. We then calculated the next-day interest based on the data from the previous 29-day data and replicated the steps to compile the final MSV series (shown as the blue line in Figure 1) with 274 data points of daily MSV. The daily RSVs shown in the 245 dotted lines produced the relative search interest over a fixed period of 30 days. Extending the investigation period to longer than 30 days could reduce the sensitivity of the daily RSV in reflecting the daily time trend of the search interest. In contrast, the data directly retrieved from GT between January 1, 2020, and September 30, 2020, contained only 39 data points of weekly RSVs (shown by the red line in Figure 1). This reduction of data points implied that we needed to lower the time interval of investigation to 1 week, meaning that we could not study a daily dynamic in the search interest. Owing to the constraints of GT, we could only retrieve weekly RSVs of an interval of at most 270 weeks. Using our proposed MSV approach, we merged the 733 RSV series into one singular, extended MSV time series, aiming to recover the daily trend of the search interest. In fact, the patterns of the MSV (the blue line) and the weekly RSV (the red line) were similar, indicating that the transformation algorithm for constructing the MSV was reasonable in the sense that the trend of the search interest revealed by the MSV was consistent with the trend of the search interest revealed by the weekly RSV provided by GT. This example shows that although GT restricts the search data retrieval into weekly data points over a long period, our proposed approach could recover the daily search volume over the same investigation period to facilitate subsequent research analyses.

Figure 1. An example of the process of creating a merged search volume series (data set keyword: “Coronavirus”; dates: from January 31, 2020, to September 30, 2020). Dotted lines: the 30-day “interest over time” relative search volume (RSV) data set from Google Trends (GT). Red line: the weekly RSV series retrieved from GT. Blue line: the daily merged search volume series using our transformation algorithm to merge RSVs through a rolling window approach.



Selection of Coronavirus-Related Search Keywords

To understand the associations between the search interest and the pandemic spread, we explored a list of coronavirus-related keywords based on the metrics of “related topic” provided by GT. We searched the related keywords for 3 seed topics: “Coronavirus disease 2019,” “Severe acute respiratory syndrome coronavirus 2,” and “Coronavirus,” with a 7-day sliding window.

By combining the shorter lists of keywords returned by GT when searching for the 3 seed terms in each 7-day period over 108 weeks (from January 1, 2020, to January 31, 2022), we obtained a raw list of keywords. Then, we filtered out the duplicates and irrelevant keywords and eventually settled on a keyword set of 228 elements (search topics), which are listed in Textbox 1. Using the transformation algorithm proposed in the *Transforming Multiple RSVs Into a Single MSV* section, the MSV of each keyword was calculated for dynamic network analysis, which involved compiling network statistics using related keywords as network nodes for a pandemic risk assessment.

MSVs and COVID-19 Waves

To understand the relationship between internet search interest and pandemic development, the MSV trends of several selected keywords were compared in parallel with the newly confirmed COVID-19 cases (Figures 2 and 3). Owing to the severity of the COVID-19 pandemic in the United States compared with other countries, we evaluated the MSV data against the new COVID-19 cases in the United States in Figure 2, in addition to the global cases, as shown in Figure 3. We noticed a similar pattern between MSV data and the number of COVID-19 cases in the United States and globally. Therefore, we used the situation with COVID-19 in the United States as a basis for further discussion, and the peaks of the 5 waves of COVID-19 in the United States were highlighted in different colors as a reference to illustrate the different stages in the development of the pandemic. Diverse associations with the waves can be observed in the trend comparisons. For example, drastic reactions in MSVs can be observed for the keywords "Coronavirus," "Symptom," "Airborne transmission," "Travel warning," "Hygiene," and "Emergency medical service" in Figures 2A-2D, 2F, and 2H before the number of COVID-19 cases reached the first wave peak (highlighted in light blue) in early April 2020. These abrupt increases in the search interest of the general pandemic-related keywords provide evidence of the capabilities of MSVs in the early detection of outbreak risks.

In contrast, the MSV trends for "Vaccine" and "Rapid diagnostic test" in Figures 2G and 2I show different variations during the latter waves of the COVID-19 pandemic. Although vaccination and efficient diagnostic tests are typically developed in the later phase of an emerging infectious disease, the corresponding MSV trends still aligned with symbolic milestones in COVID-19 control efforts. This occurred when effective rapid diagnostic tests and vaccines for COVID-19 became available in the United States, during and after the third pandemic waves, respectively. These diverse associations between the search interest of keywords and pandemic developments also motivated and supported our approach to investigating the dynamic connectedness among the coronavirus-related search keywords through network analysis.

We also compared the MSV trends of the selected keywords with the number of confirmed new COVID-19 cases worldwide in Figure 3. In Figure 3, we can draw similar conclusions to Figure 2. We also noted an increase in confirmed cases of COVID-19 globally from April to May 2021, which correlated with the sharp rise in the MSVs of "Airborne transmission" and "Vaccine."

Figure 2. Comparison of US-confirmed new cases (blue line) sourced from the World Health Organization and the merged search volume of selected keywords, including (A) Coronavirus, (B) Symptom, (C) Airborne transmission, (D) Travel warning, (E) Public transport, (F) Hygiene, (G) Vaccine, (H) Emergency medical services, and (I) Rapid diagnostic test (red line), constructed by transforming the relative search volume of Google Trends from January 1, 2020, to February 3, 2022. Highlighted areas: the 5 wave peaks observed in confirmed new cases in the United States.

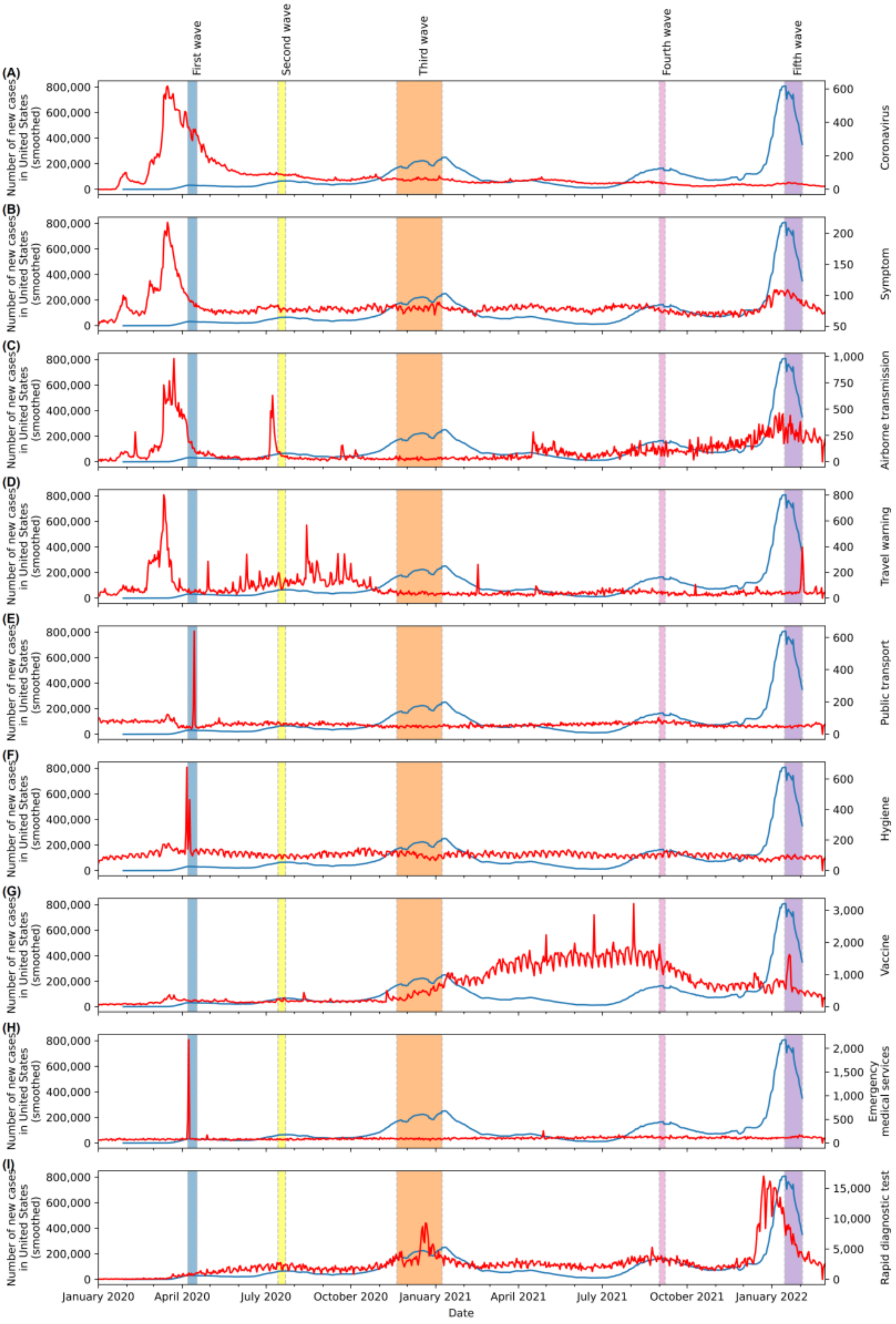
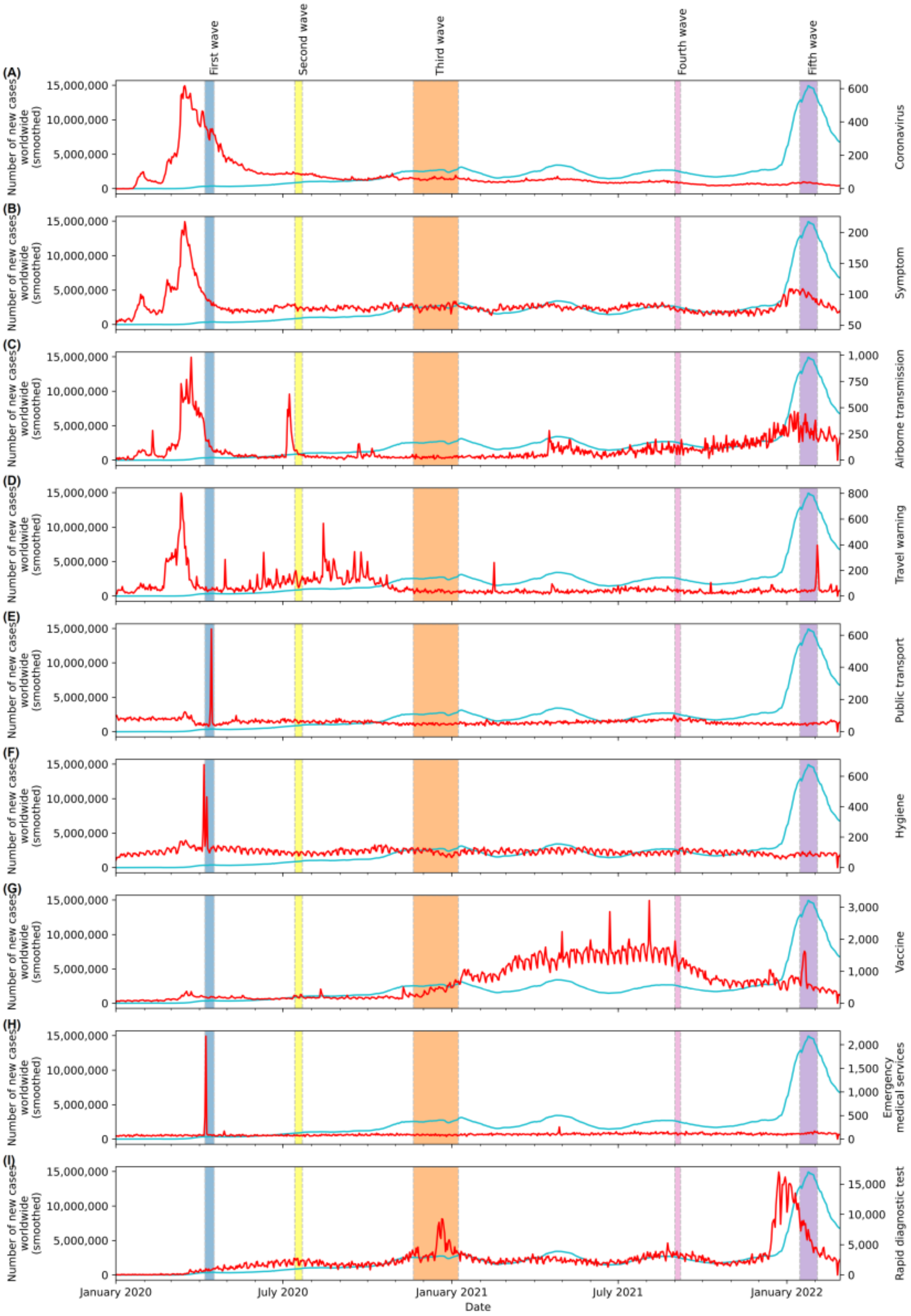


Figure 3. Comparison of worldwide confirmed new cases (cyan line) sourced from the World Health Organization and the merged search volume of selected keywords, including (A) Coronavirus, (B) Symptom, (C) Airborne transmission, (D) Travel warning, (E) Public transport, (F) Hygiene, (G) Vaccine, (H) Emergency medical services, and (I) Rapid diagnostic test (red line), constructed by transforming the relative search volume of Google Trends from January 1, 2020, to February 3, 2022. Highlighted areas: the 5 wave peaks observed in confirmed new cases in the United States.



Dynamic GT Network Construction

After computing the MSVs of all the selected keywords, we constructed a dynamic GT network based on the correlations between the search volume data of pairs of selected keywords. In the network analysis terminology, we established an edge (a connection) between 2 keywords (also known as nodes in a network) if the corresponding correlation coefficients were significantly high (when the correlation coefficient is >0.5; the threshold we adopted) [36,37]. This approach allowed us to quantify the connectedness of the relevant search queries using temporal dynamic network statistics.

As the search volume of keywords from a common topic (COVID-19, in our case) represents people's concerns about the topic, we anticipated that any common trend or comovement of the COVID-19–related keywords could indicate people's interest in the pandemic as they looked for relevant helpful information through internet searches. An increase in the MSVs of different keywords simultaneously might indicate an increasing concern about the common topic associated with the keywords.

To incorporate common trends in the search interest of multiple keywords, network analysis is a viable approach [41-45]. A network is a natural configuration that integrates common trends of information about multiple keywords together, from which we can perform a network analysis; for example, based on network statistics, we can understand the general interest in a collection of keywords simultaneously. Through the connectedness of a GT network [43,46-48], we quantified the coherence of changes in search interest for different keywords over time and studied the implications of the pandemic risk in this research. The set of network nodes is represented by the set of keywords, which is constant over time. However, the set of edges changes over time to reflect time-varying patterns of search interests. By construction, the MSV represents a proxy time series of the search volume, and the correlation of changes in the MSVs of 2 keywords exceeding a certain threshold can be interpreted as a strong comovement whose occurrence can be used to define an edge of the 2 keywords. For example, if the correlation between changes in the MSV of "Coronavirus" and "Symptom" at time *t* is >0.5 (the correlation threshold we took in this research), we created an edge between the nodes "Coronavirus" and "Symptom" in the GT network at time *t*. This construction mechanism used the same rationale that was used in the pandemic network constructions in other studies [43,46,49]; however, instead of using the number of confirmed COVID-19 cases in regions, we used the MSVs of keywords.

In addition to using 0.5 as the correlation threshold, we used 2 alternating correlation thresholds—0.4 and 0.6 for sensitivity analysis. A comparison of the network density and the global clustering coefficient constructed by the 3 different correlation thresholds at 0.4, 0.5 (the threshold used in this study), and 0.6 is shown in Figure 4. The pattern of the network density plots with thresholds of 0.4 and 0.6 is very similar to that of the one with a threshold of 0.5. We also observed a similar phenomenon in the global clustering coefficient, with an exception in March 2020, where the coefficient with a threshold of 0.6 showed a marginally narrower peak than those with thresholds of 0.4 and 0.5.

Following the studies by Chu et al [36], So et al [37], and So et al [46], the correlations of changes in MSV at time *t* were calculated using a rolling window approach that was based on the past *R*-day (including time *t*) MSV changes of keywords; therefore, the topology of the edges in the GT network would reflect the search interest characteristics of multiple keywords at time *t*. Ultimately, we had a constellation of dynamic GT networks constructed by correlations of changes in MSVs among keywords. Experience from the early research papers on COVID-19 revealed that the dynamic GT network, which summarized the infodemiology of relevant keywords, could provide an effective visualization of a pandemic situation [46], early warning signals of the pandemic [36], and pandemic risk assessments [37,50].

Figure 5 depicts a snapshot of the dynamic GT networks for 4 significant events: "Public Health Emergency of International Concern" on January 31, 2020, "COVID-19 was declared as a global pandemic by WHO" on March 11, 2020, "the Delta variant was named" on May 21, 2021, and "the Omicron variant was named" on November 26, 2021. Figure 5 also highlights the nodes with the top 5 most frequently appearing topic types: diseases (blue nodes), viruses (orange), vaccines (green), government agencies (red), and pharmaceutical companies (purple). It is obvious that the nodes related to diseases (blue nodes) occurred in the central cluster in all 4 snapshots. Several virus-related nodes (orange nodes), such as another official name for COVID-19 ("severe acute respiratory syndrome coronavirus 2" or "SARS-CoV-2"), were also included in the central clusters in most of the 4 time points we studied. These 2 observations matched our intuition that the general public seeks unknown information about COVID-19 through the search engine.

It was found that the vaccine-related nodes (green nodes) had a high degree of connectedness on all sampling days, even before any COVID-19 vaccine was approved for administration. In Figures 5C and 5D, disease-related and virus-related nodes drifted away from the cluster center, whereas nodes related to vaccines and pharmaceutical industry companies moved toward the center of the cluster.

To study the network macroscopically, we computed 2 network statistics over time—the network density and the global clustering coefficient—by using a 30-day rolling window. The network density is a measure of the node connectedness of a network, whereas the global clustering coefficient represents the ratio of the number of established connections (edges) to the number of possible connections. The higher their values, the more interconnected the network nodes are. Figure 6 presents a time series of the GT network density (blue line) and a time series of the GT network clustering coefficient (green line). In Figure 6, we also mark the 4 time points from Figure 5 to assess the relationships among the 2 network statistics, network density and global clustering coefficient, and the propagation of COVID-19. In the very early stage of COVID-19, on January 31, 2020, both the network density and global clustering coefficient were in the middle level of approximately 0.2 for the network density and 0.12 for the global clustering coefficient (the maximum value is 1.0). Probably because of the warnings of an emerging public health concern and because some keywords are not directly related to the pandemic, the search interest intensity was already quite high in January 2020. Both network statistics climbed to a peak of approximately 0.35, which alerted people to the severity of the disease's spread, thus inducing more search interest for the keywords at the same time. Subsequently, the network density remained at a relatively low level of approximately 0.15, even when the Delta variant was named on May 31, 2021. This reflects a relatively lower attention to the pandemic among the general public owing to people's adaptation to and fatigue from the long-term antipandemic living environment. What is particularly alarming is the occurrence of another peak in the 2 network statistics during October to early November 2021, before the Omicron variant was named. The GT network density and global clustering coefficients show a substantial increase in the public's interest in searching for keywords, probably indicating growing public awareness about and panic toward the next phase of the pandemic. Not long after that, in early 2022, the Omicron variant had a great impact on many countries.

Figure 4. A comparison of network density and clustering coefficient constructed by 3 different correlation thresholds at 0.4, 0.5 (the threshold used in the study) and 0.6.

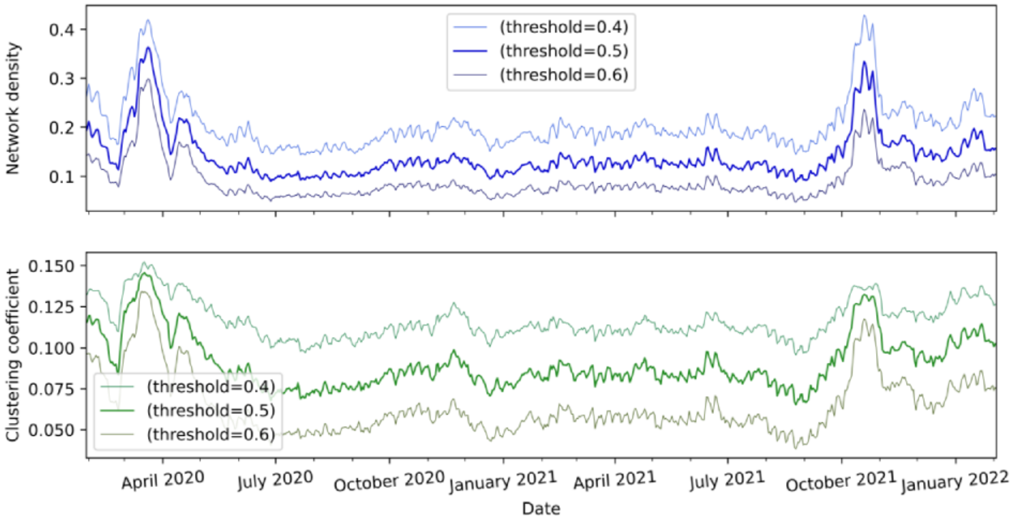
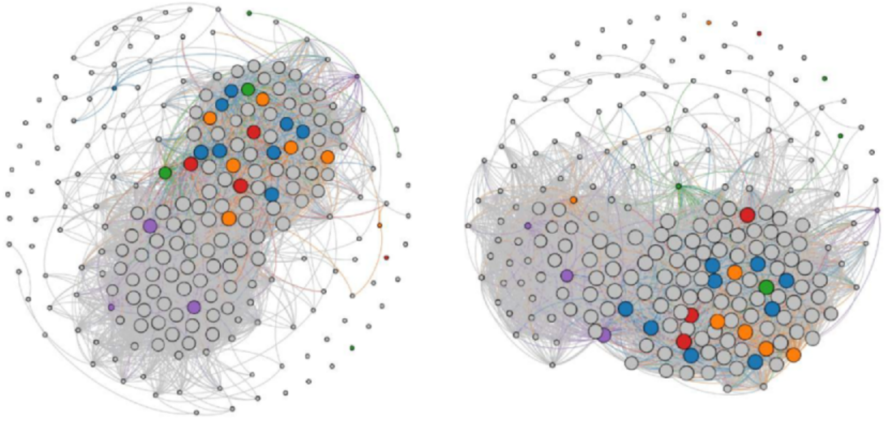
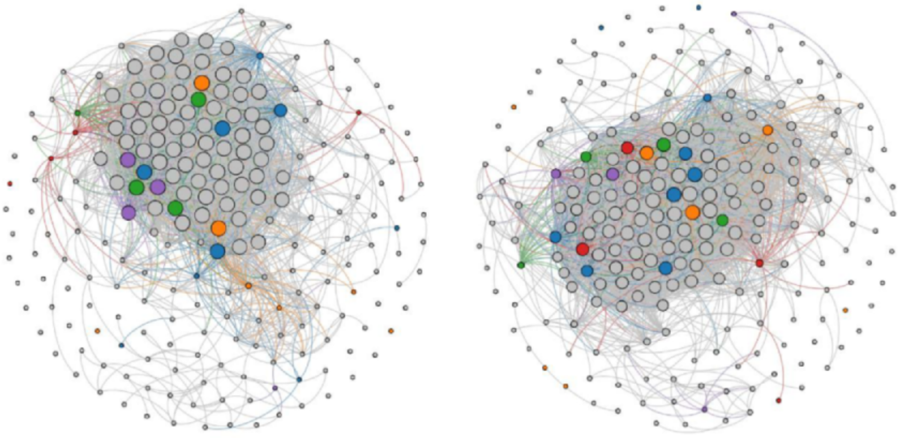


Figure 5. Snapshots of the dynamic Google Trends network. Snapshots were taken of the network on (A) January 31, 2020, (B) March 11, 2020, (C) April 31, 2021, and (D) November 26, 2021. Nodes are highlighted with different colors to represent the top 5 categories: diseases (blue), viruses (orange), vaccines (green), government agencies (red), and pharmaceutical companies (purple).

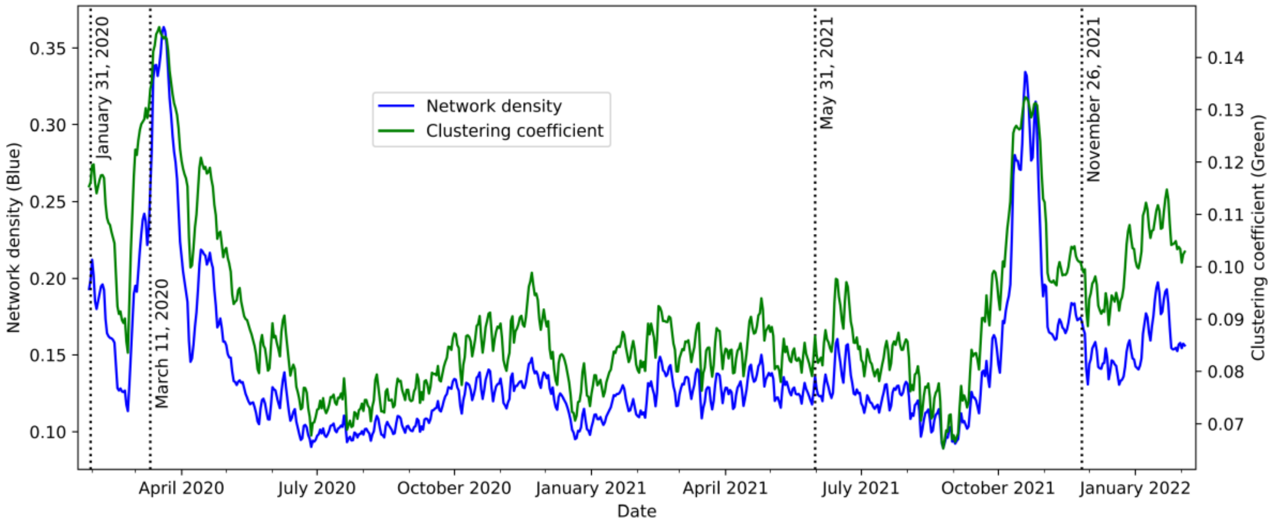


(A) January 31, 2020 **(B)** March 31, 2020



(C) May 31, 2021 **(D)** November 26, 2021

Figure 6. Network density and global clustering coefficients of the merged Google Trends network over time. The occurrences of significant events are marked with dotted lines. The 4 significant events marked are as follows: (1) COVID-19 was declared a “Public Health Emergency of International Concern” (on January 31, 2020), (2) COVID-19 was declared a pandemic (on March 11, 2020), (3) the Delta variant was named (on April 31, 2021), and (4) the Omicron variant was named (on November 26, 2021).



Comparing Network Statistics and SRSs Using DTW

Figures 7 and 8 illustrate the association between network statistics and the COVID-19 pandemic propagation in multiple directions. As mentioned in the *MSVs and COVID-19 Waves* section, the waves of COVID-19 cases represent the progression of the pandemic. As Figure 7 shows, there was an upward trend and peak in the dynamic network before every peak of the COVID-19 pandemic wave (highlighted in yellow). This shows that the dynamic networks of related keywords were capable of providing an early signal of not only the first outbreak but all 5 outbreaks throughout the COVID-19 pandemic (as shown in Figures 7A and 7B).

We further examined the predictive power of the dynamic network for the coronavirus-related keywords using the DTW algorithm [38,39]. By evaluating the similarity between the network and the newly confirmed cases in the United States, the resulting paths (the dotted lines) connect the 2 series of data and show an asymmetric step pattern. The consistently backslashed direction of the dotted lines suggests the lag-lead relationship between the dynamic network and the variation in the daily confirmed cases.

To evaluate the pandemic predictive capability of the dynamic network among the selected keywords in multiple layers, we also computed the DTW scores between the network and SRSs). The SRS, adopted from the studies by So et al [37,50], is regarded as a systematic measure that assesses the risk of the COVID-19 outbreak worldwide. A higher SRS represents a stronger signal, indicating that the global pandemic is uncontrolled. The DTW diagram in Figure 8 shows the lead-lag relationship of the SRSs throughout the entire observation period of the COVID-19 pandemic. These findings indicate that the dynamic network of related search keywords can predict the severity of the global risks of a pandemic. If regional search data are used, the predictability of the regional pandemic risk is also possible.

Figure 7. The 2-way dynamic time warping diagrams of the confirmed new COVID-19 cases in the United States (blue line) and the network statistics (black line) of the selected keywords in Google Trends: (A) network density trend versus confirmed new cases in the United States and (B) trend of the global clustering coefficient versus confirmed new cases in the United States. Yellow highlighted areas mark the 5 wave peaks observed in confirmed new cases in the United States.

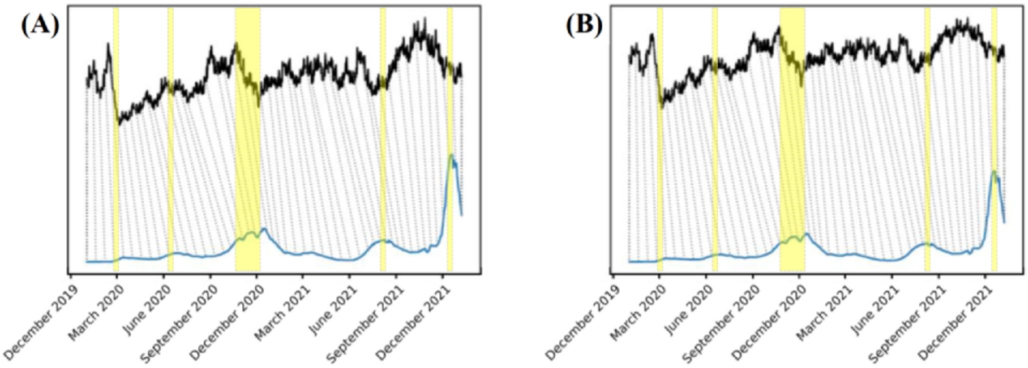
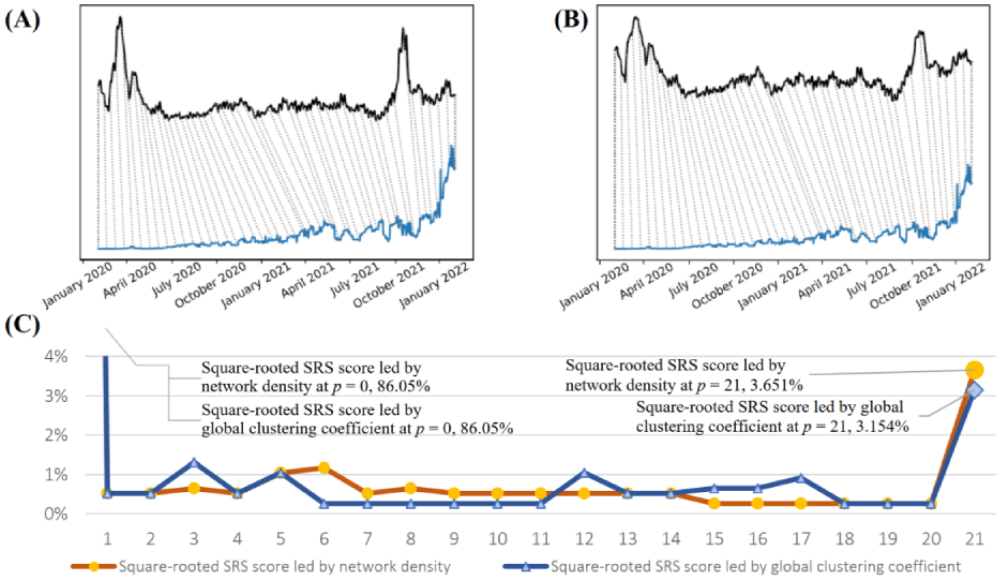


Figure 8. An illustrative evaluation of dynamic time warping (DTW) connections of the severity risk scores (SRSs) of the COVID-19 pandemic in the United States and the network statistics. Top: the 2-way DTW diagrams of the SRSs of the COVID-19 pandemic in the United States (blue line) and the network statistics (black line) of the selected keywords in Google Trends: (A) network density trend versus SRSs and (B) trend of global clustering coefficients versus SRSs. Bottom: (C) The summary plot of the time shifts (p) in each connection in (A), drawn with orange line, and (B), drawn with blue line. The data points at time shift $p = 0$ are muted in the plot.



Quantify the Asynchronous Movement of Statistics

We further quantify our observations in the 2-way DTW plots in Figures 8A and 8B by tabulating the portion of different time shift values, which shows the equivalently compelling result. As shown in Figure 8C, most of the connections between the SRS series and the network statistics series were on days 0 and 21.

We are not particularly interested in non–time-shifted DTW connections. Owing to the design of the DTW algorithm, the start and end of the pair of latent variables and observations will often warp to each other one-on-one, without any time shift, resulting in a nontrivial portion of the connections between any 2 time series defaults to $P= 0$ (ie, no time shift). Thus, we can safely assume that there is an insignificant amount of information in these connections, and we also masked the corresponding points in Figure 8C.

In contrast, the substantial portion of 21-day-forward-shifted connections offers an interesting insight. As many of the DTW connections detected have some form of forward time shift, the connections indicate that the movement of summary network statistics has a certain leading effect on the COVID-19 observed statistics. On the basis of our visual inspection and the DTW metric, the potential of summary network statistics in predicting real-life COVID-19 events is quite evident.

Dynamic Property of the MSV Network in the Pandemic Progression

In addition to the magnitude of the dynamic network, we explored the pattern alterations during the progression of the pandemic. In Figure 9, the construction of the networks on March 11, 2020, and October 23, 2021, are depicted, where the MSV network peaks are observable before the first and the fifth COVID-19 outbreaks, respectively. In addition to network density, the global clustering coefficients for these 2 dates also form peaks, indicating high clustering coefficients in the network of the selected coronavirus-related keywords during those periods.

On March 11, 2020, the dense network was represented by noticeable clusters in Figure 9A, although a small separation was found between the 2 clusters. Although the nodes of search queries “Travel warning,” “Airborne transmission,” and “Vaccine” are connected within the cluster, the nodes for “rapid diagnostic test” are located farther away from it.

Comparatively, we observe a single large cluster in Figure 9B, which shows a generally high connectedness among the coronavirus-related keywords. Moreover, the indicated node “rapid diagnostic test” has shifted into the main cluster with another node, “vaccine.” This may reflect the increasing importance and availability of effective rapid diagnostic tests for COVID-19 worldwide in the latter phase of the pandemic’s progression. Conversely, the nodes “Travel warning” and “Airborne transmission” are located outside the main cluster, as shown in Figure 9B. This change in their position may imply a reduced focus in the community owing to a relatively sufficient understanding of COVID-19 transmission and control. This finding suggests that the proposed approach to a dynamic MSV network is a possible infoveillance application that would allow nonstationary patterns of search interests during the progression of a pandemic.

We also included the square root of the SRS [37,50] in Figure 10 for comparison. The pandemic network statistics showed an upward trend that aligned with the variations in SRSs. The GT network statistics even provided a sensible signal in late 2021, alerting to the latter peak of SRS, which probably reflected the notable impact of the Omicron variant worldwide.

Figure 9. Visualization of the network construction using selected dates: (A) snapshot on March 11, 2020, and (B) snapshot on October 23, 2021. Highlighted nodes: search queries “rapid diagnostic test,” “vaccine,” “airborne transmission,” and “travel warning.” Large red and blue circles: observable clusters.

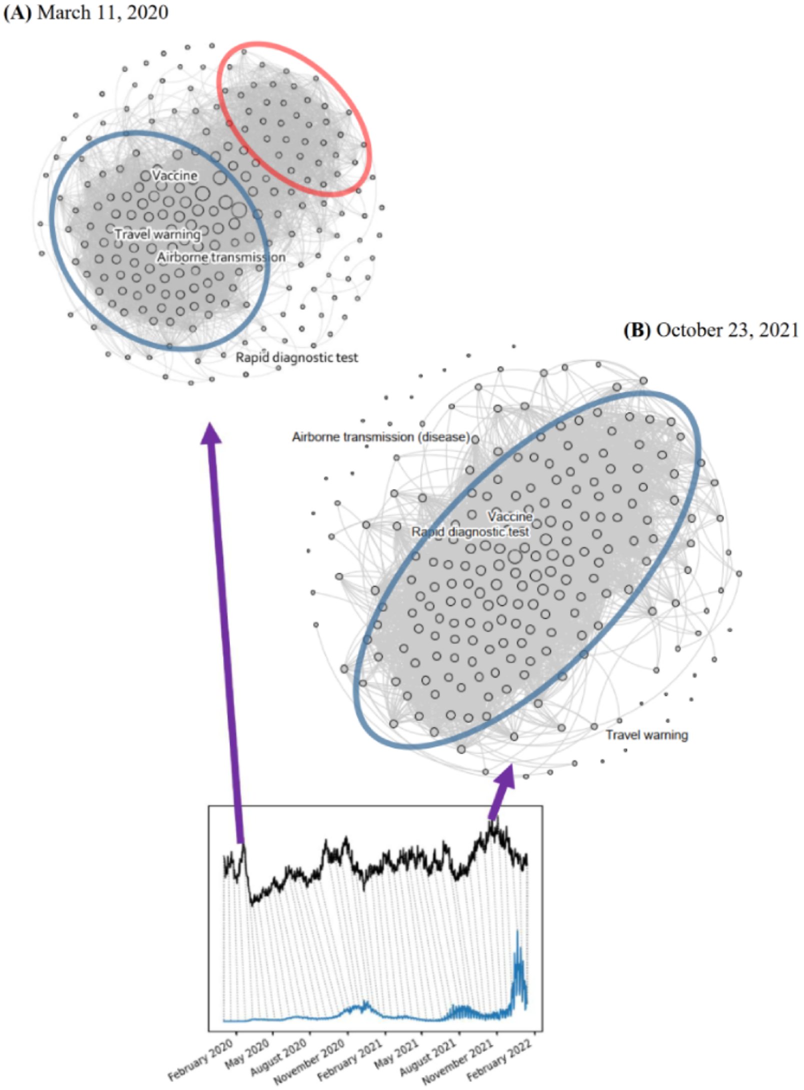
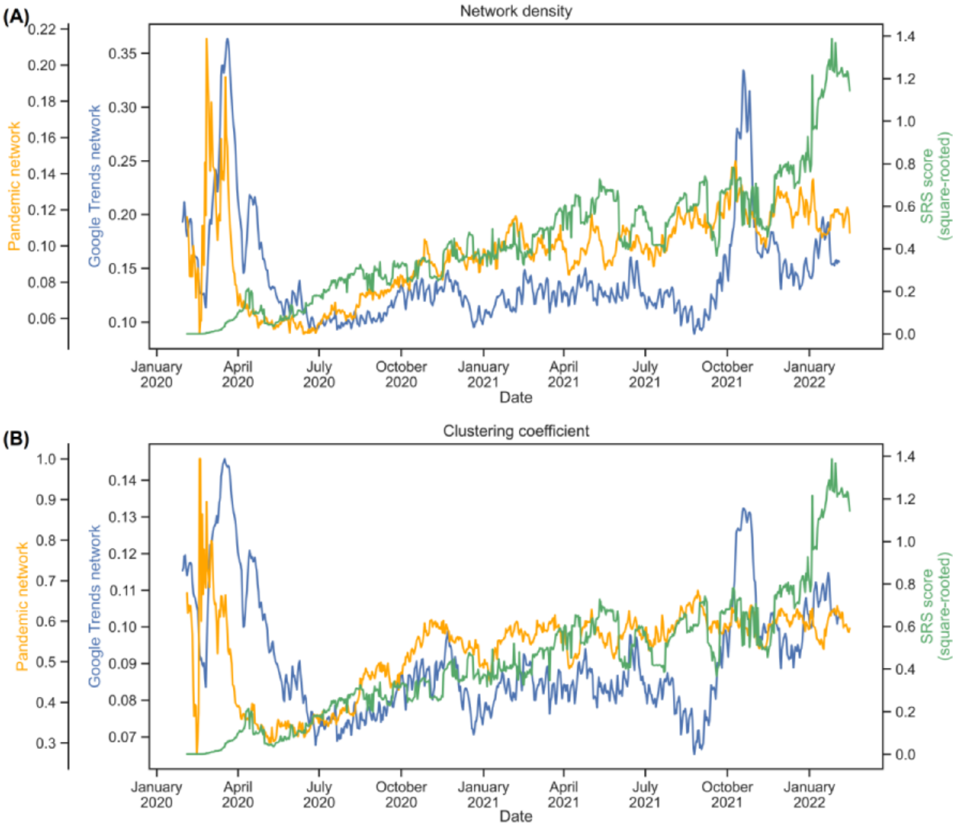


Figure 10. Two network statistics of the merged Google Trends (GT) network and the pandemic network, compared against the square roots of the severity risk scores (SRSs): (A) comparison with the GT network density and (B) comparison with the GT global clustering coefficient.



Discussion

Principal Findings

This study introduced an innovative method for handling GT search data in research analysis, including an algorithm for constructing an MSV series and a dynamic GT network.

The original data extracted from GT have several limitations that affected the analysis in this study. First, the search volume data available in GT are normalized to the specified time and region [33]. The time series data of "interest over time" of a particular keyword only indicate the search interest relative to the highest point for a defined region and period. The resulting search data were also rescaled as integers ranging from 0 to 100 [33]. The normalization and rounding processes may lead to data inaccuracy in representing the search interest. In addition, GT restricts the availability of data over a long period. When one requests search data over a long duration, only weekly or monthly data will be returned, rather than daily data [25]. This restriction greatly reduces the resolution of the search data and affects further investigation [31,51]. Therefore, this study introduced a merged algorithm that helps recover the resolution and accuracy of the search volume data extracted from GT over long observation periods. The suggested MSVs can broaden the potential value of GT data in long-term surveillance of health phenomena.

Strengths and Limitations

Using our novel merged algorithm, we observed similar phenomena in search volumes, which have been suggested in previous work, such as certain search terms correlated with COVID-19–confirmed cases [9], and there were increasing or decreasing searches for some keywords before major COVID-19 events [51]. However, unlike other previous infoveillance studies using GT [9,15-21,23-31,52-54], which often have to either sacrifice data resolution or data time horizon, we are able to use higher-resolution data with shorter delays for prolonged periods.

In addition to the merged algorithm, we demonstrated the potential of network analysis as a useful tool for infodemiology and infoveillance. Network analysis offers not only an integrated perspective of multiple time series but also cross-sectional visualization. In this study, we illustrated the cross-sectional network constructions at different timestamps in Figure 5 to investigate the clustering effect and associations of various keywords at different moments. We further presented an overall view of a temporal network via 2 summary statistics, namely, network density and global clustering coefficient, and revealed the lead-lag relationships between the official COVID-19 statistics and the summary statistics. The promising findings suggest that network analysis could be an important component for developing new predictive models.

Interestingly, the vaccine-related nodes (green nodes) in Figure 5 had a high degree of connectedness on all sampling days, even before any COVID-19 vaccine was approved for administration. It may be inferred that before the start of the pandemic, the public, health experts, and researchers searched for information about the coronavirus and its vaccines, resulting in a high number of correlations at the early stage, as shown in Figure 5A. Later, as the pandemic hit harder and the public began to understand more about COVID-19, people sought information more targeted at recovering from the COVID-19 pandemic, such as information on vaccinations and medications, as demonstrated in Figures 5C and 5D. In Figures 5C and 5D, disease-related and virus-related nodes drifted away from the cluster center, whereas nodes related to vaccines and pharmaceutical industry companies moved toward the center of the cluster.

Organizations such as government agencies and pharmaceutical companies were also on the radar of public awareness. At the beginning of the COVID-19 pandemic, members of the public understandably sought information on preventive measures and guidance from their government. As the pandemic continued, the public became familiar with social distancing and testing measures, and their attention shifted to vaccinations. As a result, the vaccines and the manufacturers of COVID-19 vaccines thus became part of the focal points, as illustrated in Figures 5C and 5D, where vaccine-related nodes (green) and nodes of pharmaceutical companies (purple) move toward the center of the cluster.

Our findings also revealed the capability of MSVs for the early detection of outbreak risk of emerging infectious diseases. Early signals are critical in controlling an emerging infection outbreak, as they allow policy makers and governments to impose rapid measures to curb the propagation of the disease. On the basis of our study findings, the MSVs of the generic keywords, such as the name of the infectious disease, symptoms, and transmission method, can provide an early signal of the risk of a potentially serious outbreak. This early detection may be related to self-diagnosis or information-seeking behaviors conducted via the internet, similar to the findings by Mangono et al [51]. Previous studies have raised the issue that a growing number of people will search the internet for health information about emerging diseases and their bodily changes and will make self-diagnoses [53] and seek preventive measures [23]. This causes the popularity of the related keywords to increase rapidly before an actual outbreak. In contrast, public transportation is considered as a factor in the spread of infectious diseases during the early stage of a pandemic. Therefore, the related keywords (eg, “travel warning” and “public transport” in Figures 2D and 2E) could be an indicator of citizens’ mobility intentions and provide useful information for predicting the spread of a pandemic.

In addition to discovering the early detection capability, we also revealed that different related keywords could have increased in popularity in search engines based on their roles at different stages of pandemic development. In the 5 observable waves of the COVID-19 pandemic in the United States since January 2020, each wave consisted of unique epidemiologic characteristics and caused the US government to take diverse actions to respond to the outbreaks [55,56]. Unlike other chronic diseases, the continual evolution of a pandemic and policy making may drastically vary the association between the public’s search interests and the progression of a disease of concern. Lu and Reis [19] also elaborated that the association between some search keywords and a disease’s progression may not be stable and persistent in the later waves of an outbreak because of increased public education about the particular disease.

By aligning the trends of MSVs with the COVID-19 pandemic waves, we discovered the dynamic property of multiple search queries when associated with the pandemic’s progression. This dynamic search pattern may be related to a shift in the public’s search interests regarding the same aspect of the pandemic. For example, personal hygiene was a crucial precautionary measure promoted during the first wave of the COVID-19 outbreak [54]. With appropriate health education and promotion, the public then developed a fundamental understanding of the measures, such as handwashing and mask wearing. With the 3 waves of COVID-19, people raised the demand for vaccines as an effective precaution against the disease [52]. As can be seen in our findings (Figures 2F and 2G), there was a shift of public search interests in precautionary measures from “hygiene” to “vaccine” after the third wave. Similarly, people with COVID-19 symptoms apparently considered “emergency medical service” during the early phases of the pandemic [57], but with the available “rapid diagnostic test” in 2021, those with symptoms tended to conduct the rapid tests before seeking medical services [58]. This yielded an observable shift in public search interests in diagnostic procedures from “emergency medical service” to “rapid diagnostic test” during the third wave. The shifting search interest may have been owing to changes in information-seeking behaviors before public policy implementation [51].

As the dynamic relationship between the internet data and the pandemic trends in the long run limited the potential of the conventional predictive models using search volumes as predictors [31], we mitigated this problem by considering the connectedness of the keywords, which could reflect the actual linkages of public concerns and human information-seeking behaviors. In this study, we provided an innovative application of network statistics that quantified the connectedness of MSV data in processing the prediction. Although multiple COVID-19 variants occurred that varied the pandemic’s development and the corresponding implementation of health measures [54], the resulting network trends could still be longitudinally associated with the 5 COVID-19 waves in the United States. By using DTW, the dynamic network of the MSVs was further justified as a leader of the trend of US-confirmed cases and of the severity of the global pandemic risk. Although previous studies revealed that linkages and patterns existed among the search queries specific to the pandemic [19,51], our proposed method took the nonstationary pattern among the relevant search queries into account during pandemic predictions. The method was able to demonstrate that the network of the merged GT volumes can be considered as a new form of official web-based data, not only associated with the pandemic but also capable of predicting its development.

This study’s approach using network analysis is innovative in the field of health and epidemiology research. Our early ground research provided a comprehensive empirical foundation for applying network statistics to investigate diverse pandemic-related issues, including infection topology [37], travel restrictions [43,47,49], and financial impacts [36,45,48]. In particular, pandemic networks using COVID-19 confirmed cases have proved their capability in tracking and providing early warnings of infection outbreaks [44,46]. We adapted the construction of a pandemic network to an infodemiological setting by investigating the dynamic GT network. In Figure 10, we compare our GT network statistics with the corresponding pandemic network statistics proposed in multiple relevant studies [36,37,44,46,47,49]. The pandemic network density and global clustering coefficients were shown to provide early warning signals of the COVID-19 pandemic [36,37,43]. Interestingly, both the GT network density and global clustering coefficients behaved similarly with the pandemic network statistics at the early stage of COVID-19 progression. On the basis of this comparison, we anticipate that both sets of network statistics can convey and deliver useful measures for pandemic risk and progression. As search data are considered to be complementary infoveillance measures, our proposed methodology of GT network analysis provides an important scientific discovery of the potential usefulness of search volume data in infodemiological research.

In addition to data analysis, the importance of keyword selection in infodemiology research has also been emphasized in the literature [59]. The procedure of excluding noisy queries is known to be important in ensuring the validity of research findings. However, there is a lack of a suitable framework for research on internet searching patterns during a pandemic. To prevent a stringent keyword selection by excluding potential keywords, we adopted a novel keyword selection approach based on relatively explicit data for “related topics” in GT. Although this approach helped us discover the keywords relevant to the pandemic and facilitated our interpretation of the findings, future studies in infodemiology could also adopt this approach to strengthen their selection of search queries.

We realize that our network consisted of regional topics associated with countries that have significantly higher than average internet coverage, such as Switzerland (internet penetration rate [IPR]: 98%), the United Kingdom (IPR: 98%), Japan (IPR: 94%), and the United States (IPR: 92%) [7]. Furthermore, IPR is skewed by the level of development of a country. However, because the proportion of those topics is as large as other nonregional terms, it is safe to assume that the network highlighted the interests of internet users in those regions, but the impact of the regional topics was nonsignificant and marginal.

Furthermore, we intentionally included most of the keywords we obtained from GT, as our research focuses on the worldwide public behavioral pattern during the COVID-19 epidemic. Despite the seeming irrelevance of some keywords, all the keywords collectively indicate the information the public received, with various levels of trustworthiness. Although only keywords for GT categories associated with regional effects were excluded in our study, future studies can examine the determined GT categories and exclude the keywords that are pathologically, psychosocially, or epidemiologically unrelated to their specific health concerns. In addition to the time and labor involved, researchers should consider the objectivity and validity of manually including or excluding certain keywords to develop the most rigorous data sets in their research.

Conclusions

As the growth of the epidemic in the United States generally relates to the transmission across multiple countries, our research investigates the linkage between worldwide public behavioral patterns and the COVID-19 pandemic in the United States. Future studies should further investigate the association between regional GT data and noncommunicable diseases.

Although we did not focus on investigating the spread of misinformation during COVID-19, research has shown that receiving accurate and inaccurate information is inversely correlated with individuals’ compliance with COVID-19 health guidelines [60-63]. Our data also show that the public may actively seek information on misinformation, as evidenced by the noninsignificant search interest in the term “hydroxychloroquine” (Textbox 1). Owing to the nature of GT, we cannot confirm whether the users were searching for misinformation out of their beliefs in certain misinformation, the inclusion of users’ interest in the information of different trustworthiness [60-62] and users’ interest in those who propagate said information [40] is critical to a comprehensive infoveillance instrument.

To the best of our knowledge, this study is the first of its kind to use a rolling window ensemble method [45,64] to recover data resolution in GT. This is also the first study to investigate the network statistics of GT data for predicting the development of a pandemic. The results also provide insights for researchers in further investigating the patterns of web-based searching behaviors, which is an important focus in infodemiological research. With the ongoing rapid growth of internet use, people will rely increasingly more on web-based information for decision-making. As the search volume data are freely available on the GT platform, it is expected that GT will become more common as a complementary, free, and timely data resource for traditional surveillance methods of health-related events. Future studies can focus on expanding the dynamic networks of search keywords to investigate noncommunicable diseases, health-related behaviors, and misinformation on the internet.

Acknowledgments

The work was supported by the research grants from the Department of Social Sciences and Policy Studies, The Education University of Hong Kong and, and The Hong Kong University of Science and Technology research grant “Risk Analytics and Applications” (grant SBMDF21BM07). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The source code and data sets generated during this study are available in the Zenodo repository [65].

Authors' Contributions

AMYC and MKPS conceptualized the study. MKPS and NHTL collected and analyzed the data. AMYC, ACYC, NHTL, AT, and MKPS interpreted the results. AMYC, ACYC, and NHTL drafted the manuscript. MKPS and AT finalized the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Weekly epidemiological update on COVID-19 - 1 February 2022. World Health Organization. 2022 Feb 01. URL: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---1-february-2022 [accessed 2022-06-27]
2. COVID-19 Weekly epidemiological update - 6 July 2022. World Health Organization. 2022 Jul 06. URL: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---6-july-2022 [accessed 2022-07-28]
3. Peto J. COVID-19 mass testing facilities could end the epidemic rapidly. BMJ 2020 Mar 22;368:m1163. [doi: 10.1136/bmj.m1163] [Medline: 32201376]
4. Omori R, Mizumoto K, Chowell G. Changes in testing rates could mask the novel coronavirus disease (COVID-19) growth rate. Int J Infect Dis 2020 May;94:116-118 [FREE Full text] [doi: 10.1016/j.ijid.2020.04.021] [Medline: 32320809]
5. Considerations for implementing and adjusting public health and social measures in the context of COVID-19: interim guidance, 14 June 2021. World Health Organization. 2021 Jun 14. URL: https://www.who.int/publications/i/item/considerations-in-adjusting-public-health-and-social-measures-in-the-context-of-covid-19-interim-guidance [accessed 2022-05-23]
6. Fox SJ, Lachmann M, Tec M, Pasco R, Woody S, Du Z, et al. Real-time pandemic surveillance using hospital admissions and mobility data. Proc Natl Acad Sci U S A 2022 Feb 15;119(7):e2111870119 [FREE Full text] [doi: 10.1073/pnas.2111870119] [Medline: 35105729]
7. Kemp S. Digital 2022: global overview report. DataReportal. 2022 Jan 26. URL: https://datareportal.com/reports/digital-2022-global-overview-report [accessed 2022-06-27]
8. Fernández-Luque L, Bau T. Health and social media: perfect storm of information. Healthc Inform Res 2015 Apr;21(2):67-73 [FREE Full text] [doi: 10.4258/hir.2015.21.2.67] [Medline: 25995958]
9. Li C, Chen LJ, Chen X, Zhang M, Pang CP, Chen H. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from internet searches and social media data, China, 2020. Euro Surveill 2020 Mar;25(10):2000199 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.10.2000199] [Medline: 32183935]
10. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
11. Brownstein JS, Nahari AD, Reis BY. Internet search patterns reveal firearm sales, policies, and deaths. NPJ Digit Med 2020 Nov 20;3(1):152 [FREE Full text] [doi: 10.1038/s41746-020-00356-6] [Medline: 33299072]
12. Allem J, Dormanesh A, Majmundar A, Unger JB, Kirkpatrick MG, Choube A, et al. Topics of nicotine-related discussions on Twitter: infoveillance study. J Med Internet Res 2021 Jun 07;23(6):e25579 [FREE Full text] [doi: 10.2196/25579] [Medline: 34096875]
13. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaeen M, Nurmambetova E, Lucini F, et al. Digital public health surveillance: a systematic scoping review. NPJ Digit Med 2021 Mar 03;4(1):41 [FREE Full text] [doi: 10.1038/s41746-021-00407-6] [Medline: 33658681]
14. Google trends. Google. URL: https://trends.google.com/ [accessed 2022-04-06]
15. Jun SP, Yoo HS, Choi S. Ten years of research change using google trends: from the perspective of big data utilizations and applications. Technol Forecast Soc Change 2018 May;130:69-87 [FREE Full text] [doi: 10.1016/j.techfore.2017.11.009]
16. Borup D, Schütte EC. In search of a job: forecasting employment growth using Google trends. J Bus Econ Stat 2020 Aug 03;40(1):186-200 [FREE Full text] [doi: 10.1080/07350015.2020.1791133]
17. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
18. Mavragani A, Gkillas K. COVID-19 predictability in the United States using Google Trends time series. Sci Rep 2020 Nov 26;10(1):20693 [FREE Full text] [doi: 10.1038/s41598-020-77275-9] [Medline: 33244028]
19. Lu T, Reis BY. Internet search patterns reveal clinical course of COVID-19 disease progression and pandemic spread across 32 countries. NPJ Digit Med 2021 Feb 11;4(1):22 [FREE Full text] [doi: 10.1038/s41746-021-00396-6] [Medline: 33574582]
20. Alicino C, Bragazzi NL, Faccio V, Amicizia D, Panatto D, Gasparini R, et al. Assessing Ebola-related web search behaviour: insights and implications from an analytical study of Google Trends-based query volumes. Infect Dis Poverty 2015 Dec 10;4:54 [FREE Full text] [doi: 10.1186/s40249-015-0090-9] [Medline: 26654247]
21. Shin SY, Seo DW, An J, Kwak H, Kim SH, Gwack J, et al. High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. Sci Rep 2016 Sep 06;6:32920 [FREE Full text] [doi: 10.1038/srep32920] [Medline: 27595921]

22. Mavragani A, Ochoa G. Google trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill 2019 May 29;5(2):e13439 [FREE Full text] [doi: 10.2196/13439] [Medline: 31144671]
23. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, R Niakan Kalhori S. Predicting COVID-19 incidence through analysis of google trends data in Iran: data mining and deep learning pilot study. JMIR Public Health Surveill 2020 Apr 14;6(2):e18828 [FREE Full text] [doi: 10.2196/18828] [Medline: 32234709]
24. Arora VS, McKee M, Stuckler D. Google trends: opportunities and limitations in health and health policy research. Health Policy 2019 Mar;123(3):338-341. [doi: 10.1016/j.healthpol.2019.01.001] [Medline: 30660346]
25. Mavragani A, Ochoa G, Tsagarakis KP. Assessing the methods, tools, and statistical approaches in Google trends research: systematic review. J Med Internet Res 2018 Nov 06;20(11):e270 [FREE Full text] [doi: 10.2196/jmir.9366] [Medline: 30401664]
26. Ortiz-Martínez Y, Garcia-Robledo JE, Vásquez-Castañeda DL, Bonilla-Aldana DK, Rodriguez-Morales AJ. Can Google® trends predict COVID-19 incidence and help preparedness? The situation in Colombia. Travel Med Infect Dis 2020 Sep;37:101703 [FREE Full text] [doi: 10.1016/j.tmaid.2020.101703] [Medline: 32360323]
27. Chang YW, Chiang WL, Wang WH, Lin CY, Hung LC, Tsai YC, et al. Google Trends-based non-English language query data and epidemic diseases: a cross-sectional study of the popular search behaviour in Taiwan. BMJ Open 2020 Jul 05;10(7):e034156 [FREE Full text] [doi: 10.1136/bmjopen-2019-034156] [Medline: 32624467]
28. Lampos V, Majumder MS, Yom-Tov E, Edelstein M, Moura S, Hamada Y, et al. Tracking COVID-19 using online search. NPJ Digit Med 2021 Feb 08;4(1):17 [FREE Full text] [doi: 10.1038/s41746-021-00384-w] [Medline: 33558607]
29. Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with internet search volumes: a Google TrendsTM analysis. Int J Infect Dis 2020 Jun;95:192-197 [FREE Full text] [doi: 10.1016/j.ijid.2020.04.033] [Medline: 32305520]
30. Costola M, Iacopini M, Santagiustina CR. Google search volumes and the financial markets during the COVID-19 outbreak. Financ Res Lett 2021 Oct;42:101884 [FREE Full text] [doi: 10.1016/j.frl.2020.101884] [Medline: 34903954]
31. Turk PJ, Tran TP, Rose GA, McWilliams A. A predictive internet-based model for COVID-19 hospitalization census. Sci Rep 2021 Mar 03;11(1):5106 [FREE Full text] [doi: 10.1038/s41598-021-84091-2] [Medline: 33658529]
32. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. Proc Int AAAI Conf Web Soc Media 2009;3(1):361-362 [FREE Full text] [doi: 10.1609/icwsm.v3i1.13937]
33. FAQ about Google Trends data. Google. URL: https://support.google.com/trends/answer/4365533?hl=en [accessed 2022-05-06]
34. Hogue J, DeWilde B. pytrends 4.7.3. pypi. 2020 May 22. URL: https://pypi.org/project/pytrends/ [accessed 2022-05-07]
35. Park T, Kwak H, An J. Building Longitudinal Google Trends to Measure Dynamic Local-Level Issue Attention. GitHub. 2020. URL: https://github.com/typark99/GoogleTrends_Covid19 [accessed 2021-11-03]
36. Chu AM, Tiwari A, So MK. Detecting early signals of COVID-19 global pandemic from network density. J Travel Med 2020 Aug 20;27(5):taaa084 [FREE Full text] [doi: 10.1093/jtm/taaa084] [Medline: 32463088]
37. So MK, Chu AM, Tiwari A, Chan JN. On topological properties of COVID-19: predicting and assessing pandemic risk with network statistics. Sci Rep 2021 Mar 04;11(1):5112 [FREE Full text] [doi: 10.1038/s41598-021-84094-z] [Medline: 33664280]
38. Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans Acoust Speech Signal Process 1978 Feb;26(1):43-49 [FREE Full text] [doi: 10.1109/TASSP.1978.1163055]
39. Giorgino T. Computing and visualizing dynamic time warping alignments in R: the dtw package. J Stat Softw 2009 Aug 14;31(7):1-24 [FREE Full text] [doi: 10.18637/jss.v031.i07]
40. Laperre B, Amaya J, Lapenta G. Dynamic time warping as a new evaluation for Dst forecast with machine learning. Front Astron Space Sci 2020 Jul 22;7:39 [FREE Full text] [doi: 10.3389/fspas.2020.00039]
41. Newman MEJ. Coauthorship networks and patterns of scientific collaboration. Proc Natl Acad Sci U S A 2004 Apr 06;101 Suppl 1(Suppl 1):5200-5205 [FREE Full text] [doi: 10.1073/pnas.0307545100] [Medline: 14745042]
42. Newman M. Networks. Oxford, UK: Oxford University Press; 2018.
43. Chu AM, Tsang JT, Chan JN, Tiwari A, So MK. Analysis of travel restrictions for COVID-19 control in Latin America through network connectedness. J Travel Med 2020 Dec 23;27(8):taaa176 [FREE Full text] [doi: 10.1093/jtm/taaa176] [Medline: 32970124]
44. Chu AM, Chan JN, Tsang JT, Tiwari A, So MK. Analyzing cross-country pandemic connectedness during COVID-19 using a spatial-temporal database: network analysis. JMIR Public Health Surveill 2021 Mar 29;7(3):e27317 [FREE Full text] [doi: 10.2196/27317] [Medline: 33711799]
45. So MK, Chan LS, Chu AM. Financial network connectedness and systemic risk during the COVID-19 pandemic. Asia-Pac Financial Mark 2021 Jun 14;28(4):649-665 [FREE Full text] [doi: 10.1007/s10690-021-09340-w]
46. So MK, Tiwari A, Chu AM, Tsang JT, Chan JN. Visualizing COVID-19 pandemic risk through network connectedness. Int J Infect Dis 2020 Jul;96:558-561 [FREE Full text] [doi: 10.1016/j.ijid.2020.05.011] [Medline: 32437929]
47. Tiwari A, So MK, Chong AC, Chan JN, Chu AM. Pandemic risk of COVID-19 outbreak in the United States: an analysis of network connectedness with air travel data. Int J Infect Dis 2021 Feb;103:97-101 [FREE Full text] [doi: 10.1016/j.ijid.2020.11.143] [Medline: 33212255]

48. So MK, Chu AM, Chan TW. Impacts of the COVID-19 pandemic on financial market connectedness. Financ Res Lett 2021 Jan;38:101864 [FREE Full text] [doi: 10.1016/j.frl.2020.101864]
49. So MK, Mak AS, Chu AM. Assessing systemic risk in financial markets using dynamic topic networks. Sci Rep 2022 Feb 17;12(1):2668 [FREE Full text] [doi: 10.1038/s41598-022-06399-x] [Medline: 35177679]
50. COVID-19 analytics. GitHub. URL: https://covid-19-dev.github.io/ [accessed 2022-02-17]
51. Mangono T, Smittenaar P, Caplan Y, Huang VS, Sutermaster S, Kemp H, et al. Information-seeking patterns during the COVID-19 pandemic across the United States: longitudinal analysis of Google Trends data. J Med Internet Res 2021 May 03;23(5):e22933 [FREE Full text] [doi: 10.2196/22933] [Medline: 33878015]
52. An L, Russell DM, Mihalcea R, Bacon E, Huffman S, Resnicow K. Online search behavior related to COVID-19 vaccines: infodemiology study. JMIR Infodemiology 2021 Nov 12;1(1):e32127 [FREE Full text] [doi: 10.2196/32127] [Medline: 34841200]
53. Tkachenko N, Chotvijit S, Gupta N, Bradley E, Gilks C, Guo W, et al. Google Trends can improve surveillance of Type 2 diabetes. Sci Rep 2017 Jul 10;7(1):4993 [FREE Full text] [doi: 10.1038/s41598-017-05091-9] [Medline: 28694479]
54. Husnayain A, Fuad A, Su EC. Applications of Google Search trends for risk communication in infectious disease management: a case study of the COVID-19 outbreak in Taiwan. Int J Infect Dis 2020 Jun;95:221-223 [FREE Full text] [doi: 10.1016/j.ijid.2020.03.021] [Medline: 32173572]
55. National COVID-19 preparedness plan. The White House. URL: https://www.whitehouse.gov/covidplan/ [accessed 2022-07-07]
56. Public health determination and order regarding suspending the right to introduce certain persons from countries where a quarantinable communicable disease exists. Federal Register. 2022 May 23. URL: https://www.federalregister.gov/documents/2022/04/06/2022-07306/public-health-determination-and-order-regarding-suspending-the-right-to-introduce-certain-persons [accessed 2022-05-23]
57. Margus C, Sondheim SE, Peck NM, Storch B, Ngai KM, Ho HE, et al. Discharge in pandemic: suspected COVID-19 patients returning to the emergency department within 72 hours for admission. Am J Emerg Med 2021 Jul;45:185-191 [FREE Full text] [doi: 10.1016/j.ajem.2020.08.034] [Medline: 33046303]
58. Public health surveillance for COVID-19: interim guidance. World Health Organization. URL: https://apps.who.int/iris/handle/10665/351761 [accessed 2022-08-14]
59. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of google trends in health care research: a systematic review. PLoS One 2014 Oct 22;9(10):e109583 [FREE Full text] [doi: 10.1371/journal.pone.0109583] [Medline: 25337815]
60. Barua Z, Barua S, Aktar S, Kabir N, Li M. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. Prog Disaster Sci 2020 Dec;8:100119 [FREE Full text] [doi: 10.1016/j.pdisas.2020.100119] [Medline: 34173443]
61. Vaal S, Schofield MB, Baker IS, Roberts BL. Narcissism, national narcissism, COVID-19 conspiracy belief, and social media use as predictors of compliance with COVID-19 public health guidelines. Curr Psychol (Forthcoming) 2022 Oct 05:1-8 [FREE Full text] [doi: 10.1007/s12144-022-03715-6] [Medline: 36213568]
62. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021 Mar 05;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
63. Dutta B, Peng MH, Chen CC, Sun SL. Role of infodemics on social media in the development of people's readiness to follow COVID-19 preventive measures. Int J Environ Res Public Health 2022 Jan 25;19(3):1347 [FREE Full text] [doi: 10.3390/ijerph19031347] [Medline: 35162369]
64. Chu AM, Chan TW, So MK, Wong WK. Dynamic network analysis of COVID-19 with a latent pandemic space model. Int J Environ Res Public Health 2021 Mar 19;18(6):3195 [FREE Full text] [doi: 10.3390/ijerph18063195] [Medline: 33808764]
65. Chu AM, Chong AC, Lai HT, Tiwari A, So MK. Source code for "enhancing the predictive power of Google Trends data through network analysis: an infodemiology study of COVID-19". Zenodo. 2023 Jun 27. URL: https://zenodo.org/record/8084831 [accessed 2023-06-27]

Abbreviations

DTW: dynamic time warping
GT: Google Trends
IPR: internet penetration rate
MSV: merged search volume
RSV: relative search volume
RT: related topic
SRS: severity risk score

Edited by A Mavragani; submitted 05.09.22; peer-reviewed by WK Wong, KF Lam, CY Sin, B Lam, S Yan, J Cuff; comments to author 11.05.23; revised version received 01.06.23; accepted 29.06.23; published 07.09.23.

Please cite as:
Chu AMY, Chong ACY, Lai NHT, Tiwari A, So MKP
Enhancing the Predictive Power of Google Trends Data Through Network Analysis: Infodemiology Study of COVID-19
JMIR Public Health Surveill 2023;9:e42446
URL: https://publichealth.jmir.org/2023/1/e42446
doi:10.2196/42446
PMID:37676701

Original Paper

Comparison of the Users’ Attitudes Toward Cannabidiol on Social Media Platforms: Topic Modeling Study

Yongjie Li[1], PhD; Xiangyu Yan[1], PhD; Zekun Wang[1], PhD; Mingchang Ma[2], MSc; Bo Zhang[1], PhD; Zhongwei Jia[1,3,4], PhD

[1]School of Public Health, Peking University, Beijing, China
[2]School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing, China
[3]Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University, Beijing, China
[4]Center for Drug Abuse Control and Prevention, National Institute of Health Data Science, Beijing, China

Corresponding Author:
Zhongwei Jia, PhD
School of Public Health, Peking University
38 Xueyuan Rd, Haidian District
Beijing
China
Phone: 86 18813052163
Email: urchinjj@163.com

Abstract

Background: As one of the major constituents of the cannabis sativa plant, cannabidiol (CBD) is approved for use in medical treatment and cosmetics because of its potential health benefits. With the rapid growth of the CBD market, customers purchase these products, and relevant discussions are becoming more active on social media.

Objective: In this study, we aimed to understand the users’ attitudes toward CBD products in various countries by conducting text mining on social media in countries with different substance management policies.

Methods: We collected posts from Reddit and Xiaohongshu, conducted topic mining using the latent Dirichlet allocation model, and analyzed the characteristics of topics on different social media. Subsequently, a co-occurrence network of high-frequency keywords was constructed to explore potential relationships among topics. Moreover, we conducted sentiment analysis on the posts’ comments and compared users’ attitudes toward CBD products on Reddit and Xiaohongshu using chi-square test.

Results: CBD-related posts on social media have been rapidly increasing, especially on Xiaohongshu since 2019. A total of 1790 posts from Reddit and 1951 posts from Xiaohongshu were included in the final analysis. The posts on the 2 social media platforms, Reddit and Xiaohongshu, were categorized into 7 and 8 topics, respectively, by the latent Dirichlet allocation model, and these topics on the 2 social media were grouped into 5 themes. Our study showed that the themes on Reddit were mainly related to the therapeutic effects of CBD, whereas the themes on Xiaohongshu concentrated on cosmetics, such as facial masks. Theme 2 (CBD market information) and theme 3 (attitudes toward CBD) on Reddit had more connections with other themes in the co-occurrence network, and theme 3 and theme 1 (CBD therapeutic effects) had a high co-occurrence frequency (22,803/73,865, 30.87%). Meanwhile, theme 1 (CBD cosmetics) on Xiaohongshu had various connections with others (169,961/384,575, 44.19%), and the co-occurrence frequency of theme 4 (CBD ingredients) and theme 1 was relatively prominent (27,128/49,312, 55.01%). Overall, users’ comments tended to be positive for CBD-related information on both Reddit and Xiaohongshu, but the percentage was higher on Xiaohongshu (82.25% vs 86.18%; $P<.001$), especially in cosmetics and medical health care products.

Conclusions: The CBD market has grown rapidly, and the topics related to CBD on social media have become active. There are apparent differences in users’ attitudes toward CBD in countries with different substance management policies. Targeted CBD management measures should be formulated to suit the prevalence of CBD use of each country.

(JMIR Public Health Surveill 2023;9:e34132) doi:10.2196/34132

KEYWORDS

cannabidiol; drug policy; latent Dirichlet allocation; social media; sentiment analysis

Introduction

Background

Cannabidiol (CBD) is a nonintoxicating principal constituent of the cannabis sativa plant discovered in 1940 [1]. In recent years, CBD has been widely used in cosmetics, food supplements, beverages, electronic cigarette liquids, and prescription pharmaceuticals. There are dozens of different forms of CBD products on the market, such as oils, capsules, edibles, and so on. Unlike tetrahydrocannabinol (THC), which is the main psychoactive component of cannabis, relevant studies have shown that CBD is a relatively safe and nonaddictive substance [2]. Moreover, there is evidence that CBD has variety of pharmacological effects on the receptor system, producing analgesic, anti-inflammatory, antianxiety, and antipsychotic effects [3]. Relevant clinical studies have also shown that CBD has the potential to treat anxiety disorders, panic disorders, and pain [4-7]. In addition, Tóth et al [8] suggested that CBD may be helpful in treating some skin symptoms, such as dryness and itching. Nevertheless, owing to the lack of large-scale randomized controlled trials, the clinical effectiveness and safety of CBD are uncertain [9,10] and its potential side effects have not yet been confirmed [11].

In the context of the global legalization of cannabis, government policies regarding CBD are inconsistent. In the United States, most states allow CBD for medical use and residents can purchase it from pharmacies or on the web [12]. However, products containing CBD with therapeutic claims are tightly regulated by the Food and Drug Administration, and except Epidiolex, no products have been approved by the Food and Drug Administration [13]. As the second largest CBD market, most European countries have legalized CBD products (THC not exceeding 0.2%), and those products with medical claims must be authorized as medical products [14,15]. Most customers in Europe are still free to purchase CBD products on the web. The acceleration of global cannabis legalization will lead to a further boom in the CBD market in the United States and European countries. In contrast, in countries such as China and Singapore, illicit drugs are strictly regulated by the government, including heroin, cannabis, and its derivatives. The legality of CBD has not been approved, and its use in food, medicine, and cosmetics is prohibited, as well as web-based or offline trade [16,17]. However, because of the fast-growing interest in lifestyle upgrades, customers in these countries have also begun to pay attention to CBD products [18].

With the explosive growth in the popularity of CBD products, related topics on social media are steadily increasing. People, especially adults aged between 18 and 35 years, are keen to share their feelings on social media. Moreover, users can comment on topics that they are interested in to express their thoughts and interact with others. Social media is an essential channel for CBD product marketing. Users often read related promotions when surfing on social media. Approximately 38% of customers discovered CBD through internet searches [19]. Keeping abreast of topics on social media will assist us in understanding the trends in CBD products, which have a substantial impact on people's behavior [20]. Previous studies have demonstrated the potential of social media in perceiving product trends in CBD [21,22]. However, these studies focus on social media data in the United States and European countries, where CBD has been approved for use and sale.

Objectives

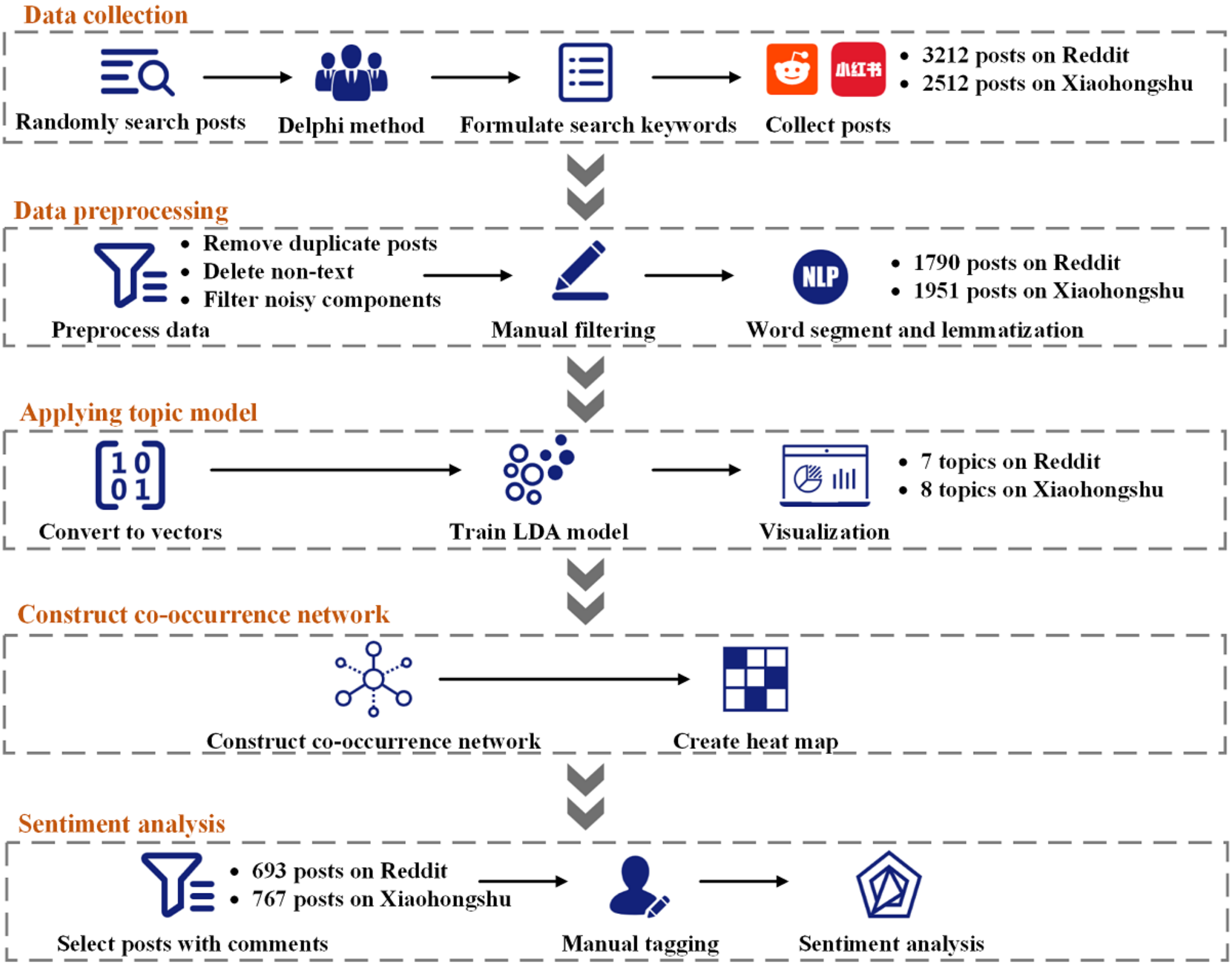
Propelled by the spread of information and promotion by celebrities, CBD products are landing on social media in China. As China implements strict drug regulatory policies, the popularity of CBD products on social media in China will bring new challenges to current policies, and there is no relevant research to date. This study aimed to extract and analyze topics on social media in China and other countries to explore users' attitudes toward CBD products and compare the differences with social media users. We hope to provide suggestions to policy makers to cope with the rapid development of the CBD market.

Methods

Data Collection and Preprocessing

First, we chose 2 popular social media platforms, Reddit and Xiaohongshu. Reddit is an English language–centric social community that includes a variety of topics, and its user base is primarily Western countries [23]. A report on the demographics of Reddit users showed that their locations were mainly from the United States, Australia, Brazil, Canada, Denmark, Norway, and other countries, which were also representative countries where CBD products were used. High school and college education accounted for the highest proportions, and the users' age ranged from 18 to 49 years [24,25]. As one of China's fastest growing social media apps, Xiaohongshu has approximately 100 million active users, most of whom are from China [26]. Xiaohongshu encourages its users to post different contents of their lives, and users can interact based on interests, such as fashion, food, and many other aspects. A study emphasized that posts on Xiaohongshu had a substantial influence on the purchasing behavior of Chinese consumers [27]. Considering that marijuana is strictly regulated in China, related posts are not allowed to be published. As Xiaohongshu is a platform connecting domestic and overseas users, CBD-related posts highlighted the current attitude of domestic users. Therefore, these posts on Xiaohongshu, as a representative, were used to analyze attitudes of users in China. We formulated a keyword list to retrieve CBD-related posts from the 2 social media platforms (Figure 1). In the beginning, we randomly searched 100 posts containing *CBD* from the 2 social media platforms and then 2 experts independently reviewed these posts and selected the appropriate keywords. Finally, the Delphi method was used to determine the completed search keyword list (Textbox 1) [28]. We searched all posts and the related comments from social media based on the keyword list we formulated until May 31, 2021. Each post included the serial number that was used to filter duplicates, titles, detailed content, and published dates. The comments on the post were composed of text content and comment date (Multimedia Appendix 1, Figure S1). As unstructured data contained considerable noise, all the original data were further processed to facilitate subsequent analysis. Duplicate posts resulting from different search keywords were removed with their comments based on the serial number. Then, nontext data (eg, video and picture) or blank content in the posts or comments were deleted. Subsequently, unnecessary sentence components (eg, emoji, hyperlink, and punctuations) and meaningless words (common stop words in English and Chinese) in the text were removed. Moreover, regular expressions were used to eliminate nonalphanumeric characters. In addition, we noticed that there were still a few irrelevant posts in the data set. Two researchers independently filtered unrelated posts. If there was a discrepancy, a third researcher determined the final outcome. Finally, for the Chinese posts on Xiaohongshu, we executed a word segment on the sentence using Jieba 0.42. In addition, we performed lemmatization on Reddit posts to extract more textual information using NLTK 3.5. After the above preprocessing, all the remaining data constituted the corpora of Reddit and Xiaohongshu.

Figure 1. Flowchart of the study. LDA: Latent Dirichlet Allocation.



Textbox 1. Search keywords on Reddit and Xiaohongshu.

Search keywords on Reddit

- cbd oil
- cbd cream
- hemp oil
- cbd hemp oil
- cbd oil anxiety
- cbd for pain
- cbd vape
- cbd benefits
- cbd liquid
- cannabidiol
- cbd cannabidiol
- cannabidiol oil
- cannabidiol cbd oil
- aceite cannabidiol
- cannabidiol capsules
- cannabidiol gummies
- cannabidiol products

Search keywords on Xiaohongshu

- cbd 油
- cbd 护肤
- cbd 面霜
- cbd 成分
- cbd 产品
- cbd 食品
- cbd 软糖
- cbd 化妆品
- cbd 胶囊
- cbd 焦虑
- cbd 失眠
- cbd 水乳
- cbd 疼痛
- cbd 抑郁
- cbd 大麻油
- cbd 烟
- cbd 电子烟

Data Analysis

Applying Topic Model

The topic model is a statistical algorithm for discovering the latent semantic structures of large corpora that can offer insights into ways we can better understand the main idea of the text. Latent Dirichlet allocation (LDA) is an unsupervised generative probabilistic method for modeling extensive text, resulting in different topics and document clusters [29]. Before applying LDA to the topic model, the text in the corpora was converted to a vector of term counts using Scikit-learn, a Python library (Figure 1). Perplexity was used as the criterion for evaluating the effectiveness of the LDA model; the lower the perplexity, the better the model predicts [30]. To determine the appropriate number of topics in the corpus of documents, the 5-fold cross-validation was applied in LDA model training. We randomly shuffled the corpus and divided it into 5 groups, 4 of which were training data sets, and the remaining one was a validation data set [31,32]. The candidate number of topics ranged from 2 to 15. We also calculated coherence measures to evaluate the candidate number of topics. Topic coherence measures scored a single topic by measuring the degree of semantic similarity between high-scoring words in the topic. Coherence value (CV) was selected as the metric in this study, and the higher value indicated optimal topic coherence [33,34]. Then, during each training iteration, we calculated and compared the perplexity and coherence scores of different candidate topic numbers in the validation data sets and presented these results in a scatter plot. Therefore, the final number of topics in the corpus was determined using a combination of perplexity and coherence. After the LDA model training was completed, we visualized the generated topics and related keywords. In the visualization layout, topics were represented in the form of bubbles, and each bubble represented one topic. The distance between the bubbles was an approximation of the semantic relationship between the topics. The overlap of bubbles indicated that there were similar parts between the topics. On the basis of the outcomes in the layout and combined with a better expression of the semantics of these topics, we merged similar topics (overlapping and closer in the distance) extracted by the LDA model into new themes.

Construct Co-occurrence Network of High-Frequency Keywords

A co-occurrence network was used to provide an intuitive expression of potential relationships between topics within corpora, and it was the collective interconnection of various high-frequency keywords based on their paired presence within the posts in the corpora. The top 10 high-frequency keywords in each theme were regarded as representative high-frequency keywords. We then formed pairs of these high-frequency keywords (Figure 1). In this study, we defined "co-occurrence" as a pair of keywords appearing in the same post. For example, "cannabidiol" (keyword in theme 1) and "experience" (keyword in theme 3) appeared in the post, so this pair (cannabidiol-experience) was said to "co-occur." Moreover, we counted the frequency of co-occurrence of high-frequency keywords in the corresponding corpora. We then constructed a co-occurrence network diagram of high-frequency keywords, where the nodes with colors represented the keywords of themes and the thickness of the edges indicated the frequency of co-occurrence in the corresponding corpora. In addition, the corresponding heatmaps were created to provide the relationships between various themes intuitively. In the co-occurrence heatmap, the number represented the frequency of co-occurrence between different themes, and the pie chart indicated the proportion of 1 theme in all themes in each row. This work was performed using Gephi 0.9.2 (Gephi Consortium) and Seaborn modules in Python 3.6 (Python Software Foundation).

Sentiment Analysis

Social media is one of the most important ways for people to obtain information, and sentiment analysis can help understand a person's opinion about a particular subject or topic based on the comments on posts [35]. Therefore, we selected posts with comments to form subcorpora and performed sentiment analysis on them (Figure 1). As the comments were feedback on the content of posts, we first paid attention to the content discussed in these posts, and we found that most of them were divided into 3 subjects: cosmetics, food, and medical health care products. We assigned 2 researchers with medical backgrounds to annotate the subject of posts with labels. The 2 researchers were asked to independently classify the posts and reach a consensus on the final label results. If there was an inconsistency, an expert would participate and ensure the final agreement. Next, we arranged for 2 other researchers to independently annotate the comments based on the detailed text to determine sentiment polarity (positive or negative). Positive comments referred to those messages that users liked or accepted in the content of posts, and negative comments were just the opposite. The process of determining positive or negative comments was consistent with the above process. According to the sentiment polarity annotation results on the comments, we set a sentiment classification label for each post. For example, the posts that had more positive comments than negative ones would be classified as positive posts and vice versa. Finally, we drew a scatter plot of sentiment polarity. The horizontal coordinates indicated the subject classification of the post, and the vertical coordinates indicated the number of comments (positive posts used the number of positive comments and negative posts used the number of negative comments). The differences in sentiment distribution between the 2 social media platforms were verified using the chi-square test executed in SPSS (version 22.0; IBM Corp).

Ethics Approval

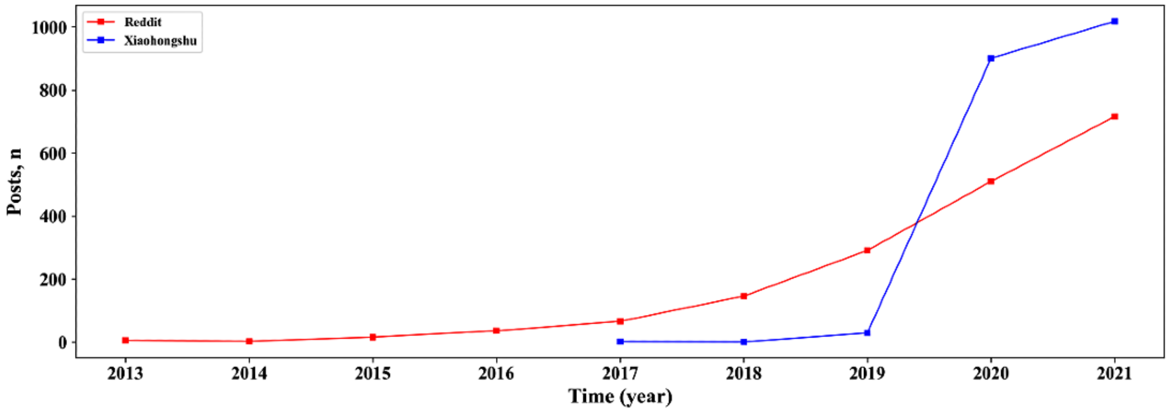
This study was approved by the Peking University Institutional Review Board (IRB00001052-16016).

Results

Basic Statistics Information

By the end of May 31, 2021, 3212 Reddit posts and 2512 Xiaohongshu posts were collected. After preprocessing, 1790 and 1951 posts were included in the statistical analyses. According to the statistics of the number of posts by years, it was apparent that there was a rising trend of CBD-related posts, both on Reddit and Xiaohongshu (Figure 2). In particular, compared with posts on Reddit, the attention paid to CBD on Xiaohongshu was hysteretic, and there had been explosive growth since 2019.

Figure 2. The number of posts on Reddit and Xiaohongshu per year.



Topic Modeling Result

After 5-fold cross-validation, perplexity was the lowest when the topic number was 7 on the Reddit corpus (199.34-201.91). Likewise, the highest CV using CV metric ranged from 0.63 to 0.67 with a topic number of 7. Therefore, 7 topics were identified from the posts on Reddit (Multimedia Appendix 1, Figures S2 and S3). The 7 topics were grouped into 5 themes: CBD product treatment and effect (theme 1, including 2 topics), market information related to CBD products (theme 2, including 2 topics), user attitudes toward CBD products (theme 3, including 1 topic), products and forms related to CBD (theme 4, including 1 topic), and consumer-oriented CBD product format (theme 5, including 1 topic). The keywords for each theme with the percentage of the given documents are summarized in Table 1. It showed that theme 1 was the most popular, accounting for 47.64% (43,197/90,676) of the keywords. Moreover, in theme 3 (17,565/90,676, 19.37%), many users expressed their desire to try CBD products, as advertisements on the CBD were readily available on similar websites. In addition, it was clear that there were various forms of CBD products in the consumer market, which were distinctive from the themes of Xiaohongshu. Perplexity and coherence varied for different values of the number of topics. On the basis of the perplexity and coherence of the Xiaohongshu corpus, we determined 8 as the optimal number of the topics with the minimum perplexity (137.15-140.70) and maximum coherence (0.64-0.68; Multimedia Appendix 1, Figures S4 and S5). We grouped these 8 topics into 5 themes (Table 2). The most obvious characteristic was that Xiaohongshu's themes were concentrated on cosmetics field. Theme 1 (including 3 topics) had the highest percentage (25,750/55,097, 46.74%) and was CBD-related cosmetics categories and functions, such as facial masks, creams, and essences. In addition, the themes on Reddit and Xiaohongshu both involved brands of CBD. However, there were not only domestic brands on Xiaohongshu but also international brands from the United States and Europe (theme 2, including 2 topics; 10,001/55,097, 18.15%). Moreover, theme 3 (including 1 topic) and theme 4 (including 1 topic) introduced the effect evaluation of CBD products and the raw materials of CBD cosmetics, respectively. Theme 5 (including 1 topic) described medical health care products with therapeutic benefits claims, and a similar theme also appeared in Reddit.

Table 1. Topic classification and keywords on Reddit corpora.

Classification and topics	Keywords	Values, %
Theme 1: CBD[a] production treatment and effect		
Topic 1: CBD production treatment	pain, anxiety, body, treatment, patients, use, reduce, sleep, disorders, cancer	25.56
Topic 2: Multiformat CBD products effect	hemp, make, gummies, benefit, plant, oil, cannabidiol, extract, high, health	22.08
Theme 2: Market information related to CBD products		
Topic 3: Company information of CBD	company, products, base, business, market, release, sell, service, live, develop	11.43
Topic 4: Brand and report of CBD	hemp, brand, share, report, sales, industry, grow, research, offer, hold, price	8.61
Theme 3: Attitudes toward for CBD products		
Topic 5: Attitude to use CBD products	like, know, want, feel, work, try, think, start, experience, really, need, good	19.37
Theme 4: Products and forms related to CBD		
Topic 6: Raw materials and other products	cannabis, marijuana, medical, extract, weed, capsule, flower, smoke, food	6.82
Theme 5: Consumer-oriented CBD product format		
Topic 7: Product formats and distribution channels	vape, bottle, liquid, spectrum, cream, free, online, shop, pure, website, order	6.13

[a] CBD: cannabidiol.

Table 2. Topic classification and keywords on Xiaohongshu corpora.

Classification and topics	Keywords	Values, %
Theme 1: CBD[a]-related cosmetics categories and functions		
Topic 1: Skin care cosmetics	skin (皮肤), sensitive (敏感), ingredient (成分), cream (面霜), skin care products (护肤品), lotion (水乳), absorb (吸收), oil skin (油皮), protect (护理)	19.33
Topic 2: Benefits of CBD mask	facial mask (面膜), repair (修护), stay up late (熬夜), hydrating (补水), emollient (舒缓), serum (精华液), fight against acne (消炎), antioxidant (抗氧化), emergency (急救)	15.67
Topic 3: Benefits of other CBD cosmetics	acne prone (痘痘), oil-control (控油), moisturize (保湿), purify (淡化), resurfacing (清爽), nutritious (舒缓), whitening (美白), antiwrinkle (抗皱), active (活化)	11.74
Theme 2: CBD-related cosmetics brands and research		
Topic 4: Study on the effectiveness of CBD	function (作用), efficacy (功效), research (研究), cell (细胞), health (健康), control (抑制), discovery (发现), stimulate (刺激), balance (平衡)	9.32
Topic 5: CBD-related cosmetics brands	products (产品), brand (品牌), industry (工业), extract (提取物), western (欧美), formulation (配方), America (美国), price (价格), domestic (国产)	8.83
Theme 3: Effect evaluation of CBD products		
Topic 6: Cosmetics effect evaluation	feel (感觉), really (真的), particular (特别), smell (味道), comfy (舒服), recommend (推荐), good (不错), try (试试), share (分享)	12.11
Theme 4: Main ingredients of CBD cosmetics		
Topic 7: Ingredient introduction	Cannabis (大麻), ingredient (成分), essence (精华), vitamin (维生素), essential oil (精油), extract (提取), add (添加), plant (植物), natural (天然)	11.98
Theme 5: Declared disease treatment effect		
Topic 8: CBD treatment of diseases	sleep (睡眠), anxiety (焦虑), relieve (缓解), insomnia (失眠), stress (压力), emotion (情绪), pain (疼痛), adjustment (调节), improve (改善)	11.02

[a] CBD: cannabidiol.

Co-occurrence Network of High-Frequency Keywords

For the themes on Reddit, there were wide connections between the different themes. Obviously, theme 2 had the most connections with other themes, which was 24.48% (83,733/342,066) followed by theme 3 that accounted for 21.59% (73,865/342,066; Figure 3A). The heatmap of high-frequency keywords showed that theme 3 and theme 1 had a high co-occurrence frequency (22,803/73,865, 30.87%); that is, users were interested in the therapeutic effects of CBD and expressed their thoughts on trying and some even used this type of product for the first time (Figure 4A). This study also showed that there were various forms of CBD products for users to try and choose, which was why the co-occurrence frequency of theme 5 and theme 3 was relatively high (6419/28,364, 22.63%). In contrast, as users in Xiaohongshu shared all kinds of posts about cosmetics, the themes of CBD-related cosmetics (theme 1) categories had various connections with others (169,961/384,575, 44.19%; Figure 3B). Furthermore, unlike the themes of Reddit mentioned above, the themes in Xiaohongshu that claimed the therapeutic effects (theme 5) were simply associated with cosmetics and derivatives of CBD ingredients (theme 4), which was only 6.89% (1758/25,503). Xiaohongshu users were more willing to spend time introducing ingredients in cosmetics, especially those related to CBD. Therefore, the co-occurrence frequency of theme 4 and theme 1 was relatively prominent (27,128/49,312, 55.01%; Figure 4B).

Figure 3. Co-occurrence network of high-frequency keywords on Reddit and Xiaohongshu. (A) The nodes represent the keywords of the themes on Reddit. The edge between two nodes (source node and target node) indicates that two keywords appear in the same post, and the thickness of the edges indicates the frequency of co-occurrence in the posts. The color of the edge is consistent with that of the source node. Theme 1: CBD production treatment and effects. Theme 2: market information related to CBD products. Theme 3: attitudes toward CBD products. Theme 4: products and forms related to CBD. Theme 5: consumer-oriented CBD product format. (B) The nodes represent keywords of the themes on Xiaohongshu. The edge between two nodes (source node and target node) indicates that two keywords appear in the same post, and the thickness of the edges indicates the frequency of co-occurrence in the posts. The color of the edge is consistent with that of the source node. Theme 1: CBD-related cosmetic categories and functions. Theme 2: CBD-related cosmetic brands and research. Theme 3: effect evaluation of CBD products. Theme 4: main ingredients of CBD cosmetics. Theme 5: declared disease treatment effect. CBD: cannabidiol.

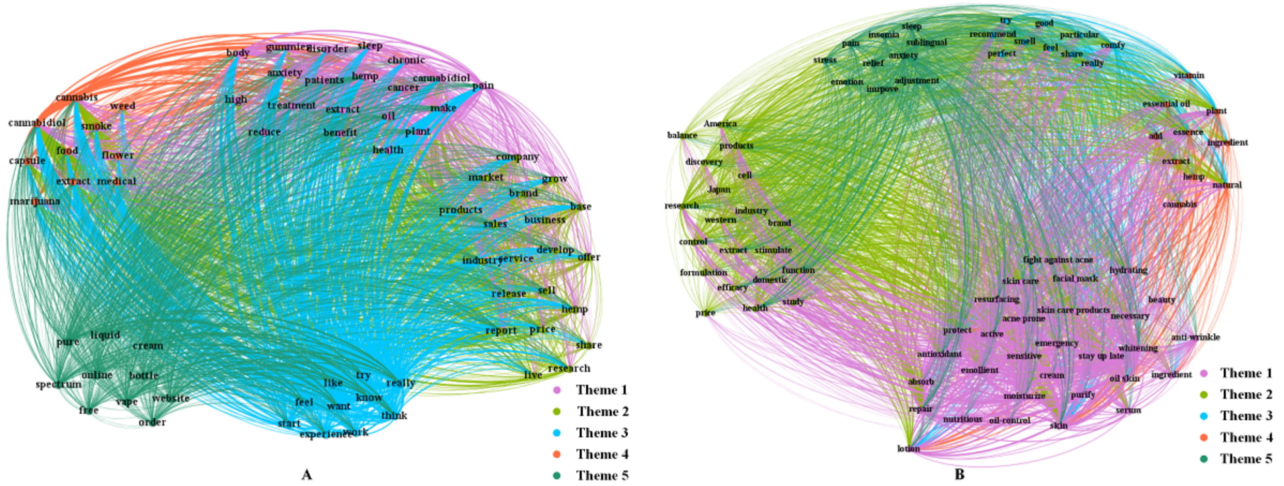
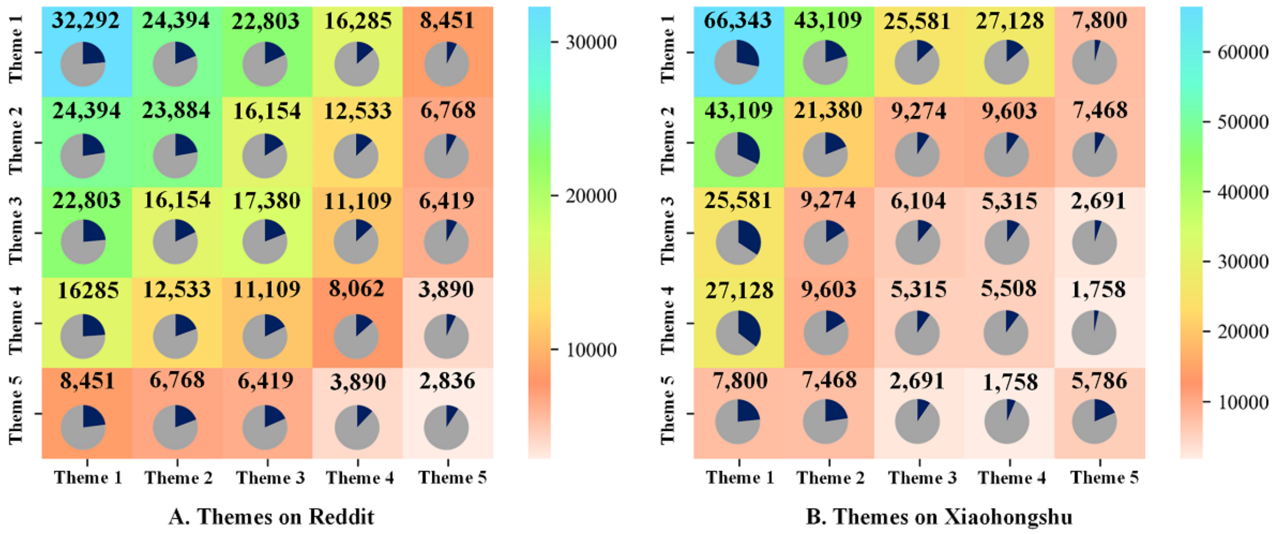


Figure 4. Co-occurrence heatmap of high-frequency keywords on Reddit and Xiaohongshu. The number is the frequency of co-occurrence between different themes, and the pie chart (blue part) indicates the proportion of one theme in all themes in each row. (A) The themes on Reddit—theme 1: CBD production treatment and effect, theme 2: market information related to CBD products, theme 3: attitudes toward CBD products, theme 4: products and forms related to CBD, and theme 5: consumer-oriented CBD product format. (B) The themes on Xiaohongshu—theme 1: CBD-related cosmetics categories and functions, theme 2: CBD-related cosmetic brands and research, theme 3: effect evaluation of CBD products, theme 4: main ingredients of CBD cosmetics, and theme 5: declared disease treatment effect. CBD: cannabidiol.



Distribution of Sentiment Polarity

After processing, there were 693 and 767 posts with comments on Reddit and Xiaohongshu, respectively (Figure 1). From the perspective of the classification subject of the posts, it was evident that the posts on Reddit were more focused on medical health care products (571/693, 82.4%), whereas the posts on Xiaohongshu had the highest proportion of posts on cosmetics (605/767, 78.88%; Figure 5). On Xiaohongshu, information about food containing CBD ingredients did not attract much attention (22/767, 2.87%), and discussions about cosmetics in Reddit were also less common (47/693, 6.78%). However, we noted that approximately 2.6% (18/693) of posts on Reddit reported using CBD products as alternative medicines, such as recreational cannabis and painkillers.

Overall, regardless of Reddit or Xiaohongshu, users' comments tended to be positive for CBD-related information, but the percentage was higher on Xiaohongshu (Reddit vs Xiaohongshu: 82.25% vs 86.18%; *P*<.001). Compared with Reddit, Xiaohongshu also had a higher percentage of positive comments on the subject of cosmetic and medical health care products (74.47% vs 91.57% in cosmetics and 82.84% vs 86.43% in medical health care products; *P*=.02). There was no significant difference in the proportion of posts with positive or negative emotions on the food subject between the 2 social media platforms (Table 3).

Figure 5. Sentiment distribution of posts on Reddit and Xiaohongshu. The horizontal coordinates indicate the subject classification of the post, and the vertical coordinates indicate the number of comments (positive posts used the number of positive comments, and negative posts used the number of negative comments).

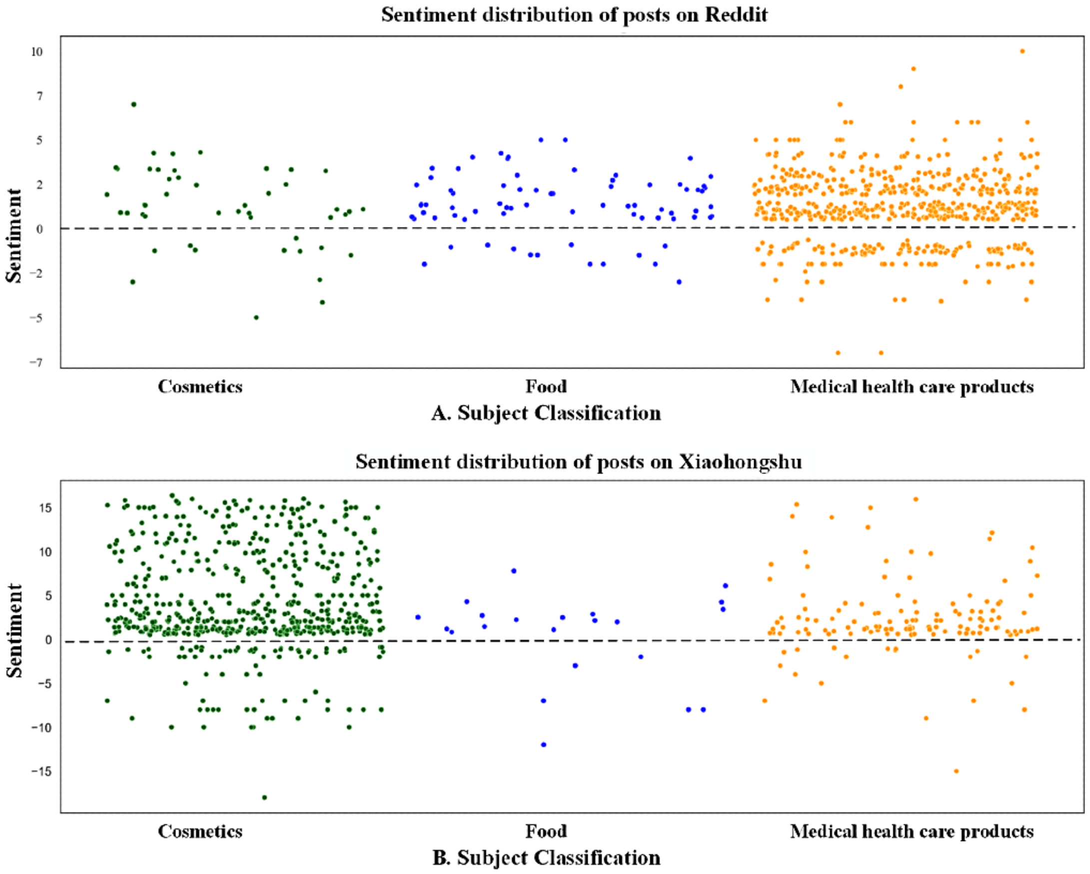


Table 3. The distribution of sentiment polarity on Reddit and Xiaohongshu.

	Overall			Cosmetics			Food			Medical health care products		
	Value, n (%)	Value, N	*P* value	Value, n (%)	Value, N	*P* value	Value, n (%)	Value, N	*P* value	Value, n (%)	Value, N	*P* value
Comments			<.001			.02			.30			.02
Reddit												
Positive	540 (82.25)	693		35 (74.47)	47		62 (82.67)	75		443 (82.84)	571	
Negative	153 (17.75)	693		12 (25.53)	47		13 (17.33)	75		128 (17.16)	571	
Xiaohongshu												
Positive	661 (86.18)	767		524 (91.57)	605		16 (72.73)	22		121 (86.43)	140	
Negative	106 (13.82)	767		81 (8.43)	605		6 (27.27)	22		19 (13.57)	140	

Discussion

Principal Findings

This novel study analyzed the posts related to CBD products on social media and compared users' attitudes toward CBD in countries with different substance management policies. The number of cannabis-related topics on social media was massive, and CBD was one of the emerging topics that attracted more and more attention. Benefiting from the formulated search keywords, the corpus in this study filtered out many data with high-noise levels, which also made the corpus relatively small. Therefore, we could build a more effective topic model to mine users' attitudes toward CBD products. The results showed that there were differences in the themes that users discussed on social media and their attitudes toward CBD had their own characteristics. Obviously, compared with some Western countries, there was a "vacuum period" when the relevant theme was first available on Xiaohongshu in China. Although China adheres to the strict drug control policy (including cannabis and its derivatives), with the popularity of legalization of medical and recreational cannabis, related themes have become increasingly active on social media in China [36,37]. The text-based topic model analysis in this study provided a reference for cross-cultural research. In traditional cross-cultural studies, qualitative methodologies were prominent and mostly conducted by cultural anthropologists in the form of ethnographies [38]. This required researchers to have a variety of cultural backgrounds, which may otherwise cause research bias. Language independence is commonly presented as one of the advantages of modern machine learning approaches to natural language processing (NLP), and it is an important type of scalability [39]. NLP technology can help us achieve cross-cultural quantitative analysis. In this study, we established an automated topic analysis process. A sequence of NLP functions was used to mine and extract topics from social media texts, and these topics were compared in different cultural contexts.

Social media serves as the vital source of information, as people are willing to share information on social media, and we can mine topic information from user-generated content on social media [40]. Although China is the largest producer of industrial hemp, its strict policy supervision has restricted the application of CBD. Cosmetics with CBD ingredients have attracted consumers, and it has been booming on social media. Even the international brands from the United States and Japan are also sold in the domestic market through e-commerce platforms (eg, Taobao and JD). Not only cosmetics but also some medical health care products that claimed to relieve anxiety or body pain have also appeared on social media in China. As more countries approved CBD for medical purposes, previous studies reported similar trends on other social media in Western countries [22]. Currently, there is a lack of scientific evidence for its effectiveness in treatment, and policy makers should pay attention to this phenomenon. Although the sale of cannabis and its derivatives is banned in China, CBD products for medical use have bypassed the legal supervision under the cover of cosmetics. In addition, many users claimed that taking CBD was good for health and even advocated anticancer effects, which may cause patients to eschew proven treatments and increase the disease burden [15]. Although CBD is touted for its therapeutic properties by consumers, public health professionals should be cautious when recommending CBD products and should ensure that regional and local laws are followed. Health care providers should clearly inform patients of the potential risks, especially in vulnerable groups such as children, older adults, and patients who are chronically ill. Moreover, our study found that users' attention to the 2 social media platforms was significantly different. Regarding CBD products, Xiaohongshu users posted more frequently about CBD in skin care, whereas Reddit users posted more frequently about the potential medicinal value of CBD products. Therefore, the government needs to strengthen the surveillance of the market and guide consumers to correctly understand CBD to avoid unnecessary risks owing to false propaganda.

There are various novel CBD product forms (eg, vaping and gummies) on Reddit, which are popular among teenagers. We noticed a close connection between this theme (theme 5) and the theme of therapeutic effects (theme 1) on Reddit. When teenagers are exposed to such information on social media, curiosity may drive them to try these CBD products, such as e-cigarettes, to relieve anxiety and further use CBD with high THC levels or even cannabis [41]. Although there were few related themes on Xiaohongshu, China would face the same situation with the rapid growth of the e-cigarette market. Therefore, it is urgent to conduct education and intervention in schools or entertainment venues. In addition, it is necessary to improve the detection of THC in CBD products. A previous study reported several CBD products with excessive THC in the European market [15]. The customs department should strengthen its supervision of the imported e-liquid. On Xiaohongshu, internet celebrities often introduced the main ingredient CBD and its relationship with cannabis when they shared the experience of using cosmetics. However, some of the contents lacked scientific evidence or were fake news. These kinds of fake news on social media mislead people and put public health at risk [42]. Therefore, social media should exercise professional responsibility to ensure that accurate information is published and disseminated. Policy makers need to penalize false or unsubstantiated CBD advertisement claims, especially for CBD foods with unclear efficacy. Any products making claims for therapeutic or medical use must be approved by regulatory authorities upon submission of data. In addition, regulators should strengthen the training of health care providers to ensure the safety and effectiveness of the prescriptions.

Finally, the sentiment polarity of comments on the 2 social media platforms reflected that users' attitudes toward CBD products differed. The almost overwhelmingly positive comments on cosmetics showed the popularity of CBD cosmetics among Chinese consumers. Most of the comments on Reddit focused on CBD health care products, but the proportion of negative comments was higher than that on Xiaohongshu. In contrast to China's drug policy, people in the West (ie, the United States and European countries) are free to buy and use medical CBD products. However, the therapeutic efficacy with actual experience may be lower than the advertised therapeutic efficacy, leading to more negative comments on social media. For example, 1 user said that the use of CBD oil did not alleviate body pain at all, and that the illness was finally cured by routine clinical treatment. In addition, the discussion on the legal status of CBD revolves around the question of whether it is a drug or food supplement. Many products were labeled as food to avoid regulation, but current data show that consumers, both in China and the West, are not particularly concerned about it. Overall, users were more positive about CBD products on social media in China, even though CBD is still illegal. As an emerging trend, domestic consumers know little about CBD and do not know enough about its management requirements and use risks, so the government should focus on publicizing CBD-related scientific evidence and management measures. Public health communities need to keep abreast of the latest research and implement evidence-based practice standards and guidelines to reduce the risk of potential misuse of CBD products.

Limitations

This study has several limitations. Our data were collected from social media, which are mostly used by internet-savvy young people. Therefore, there may be a lack of input from older adults who rarely use social media. However, considering that the primary consumers of CBD products are young people, our study represented most of the target population. In addition, the user demographics of Reddit and Xiaohongshu are not representative of all Western and Chinese populations. It is necessary to investigate more users beyond social media platforms to evaluate actual consumer attitudes. The NLP analysis is correlated with the amount of data, and the relatively precise search strategy used in this study may limit the size of the corpus. We will expand our corpus from other platforms to train a more comprehensive topic model. Also, as a formative study, more traditional measures of consumers' behavior are needed in addition to study on other social media platforms. Therefore, future research should consider analyzing more data on other social media platforms to explore users' attitudes toward CBD more comprehensively.

Conclusions

We collected posts related to CBD products from Reddit and Xiaohongshu and conducted theme mining using the LDA model. We analyzed the characteristics of the themes on social media and performed sentiment polarity on posts' comments to compare users' attitudes toward CBD products from different countries. The results showed that consumers in countries with different substance management policies had significantly different attitudes toward CBD. Therefore, more targeted CBD management measures should be formulated to suit each country's national conditions. In addition, we hope that this study can provide a reference for other countries where CBD has not been approved for use in medicines, cosmetics, and food.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant 72174004, 91546203, and 91846302).

Authors' Contributions

ZJ, BZ, and YL designed the study. YL and BZ collected the data. XY and BZ cleaned the data. YL, XY, and ZW analyzed the data. YL, ZB, MM, and ZW explained the results. YL and XY wrote the initial draft of the manuscript. YL, BZ, and ZW revised the report from preliminary draft to submission. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The visual layout of Latent Dirichlet Allocation (LDA) and perplexity and coherence with 5-fold cross-validation on documents of social media platforms.

[DOCX File , 1282 KB - publichealth_v9i1e34132_app1.docx]

References

1. The WHO Expert Committee on Drug Dependence (ECDD). World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/260546/9789241210188-eng.pdf [accessed 2021-09-07]
2. Expert Committee on Drug Dependence Fortieth Meeting. Cannabidiol (CBD) critical review report. World Health Organization. 2018 Jun 4. URL: https://www.who.int/medicines/access/controlled-substances/CannabidiolCriticalReview.pdf [accessed 2021-07-01]
3. MacCallum CA, Russo EB. Practical considerations in medical cannabis administration and dosing. Eur J Intern Med 2018 Mar;49:12-19. [doi: 10.1016/j.ejim.2018.01.004] [Medline: 29307505]
4. Chagas MH, Eckeli AL, Zuardi AW, Pena-Pereira MA, Sobreira-Neto MA, Sobreira ET, et al. Cannabidiol can improve complex sleep-related behaviours associated with rapid eye movement sleep behaviour disorder in Parkinson's disease patients: a case series. J Clin Pharm Ther 2014 Oct;39(5):564-566. [doi: 10.1111/jcpt.12179] [Medline: 24845114]
5. Shannon S, Lewis N, Lee H, Hughes S. Cannabidiol in anxiety and sleep: a large case series. Perm J 2019;23:18-041 [FREE Full text] [doi: 10.7812/TPP/18-041] [Medline: 30624194]
6. Lossignol D. Cannabinoids: a new approach for pain control? Curr Opin Oncol 2019 Jul;31(4):275-279. [doi: 10.1097/CCO.0000000000000523] [Medline: 30789867]
7. Urits I, Gress K, Charipova K, Habib K, Lee D, Lee C, et al. Use of cannabidiol (CBD) for the treatment of chronic pain. Best Pract Res Clin Anaesthesiol 2020 Sep;34(3):463-477. [doi: 10.1016/j.bpa.2020.06.004] [Medline: 33004159]
8. Tóth KF, Ádám D, Bíró T, Oláh A. Cannabinoid signaling in the skin: Therapeutic potential of the "C(ut)annabinoid" system. Molecules 2019 Mar 06;24(5):918 [FREE Full text] [doi: 10.3390/molecules24050918] [Medline: 30845666]
9. White CM. A review of human studies assessing Cannabidiol's (CBD) therapeutic actions and potential. J Clin Pharmacol 2019 Jul;59(7):923-934. [doi: 10.1002/jcph.1387] [Medline: 30730563]

10. Boyaji S, Merkow J, Elman RN, Kaye AD, Yong RJ, Urman RD. The role of Cannabidiol (CBD) in chronic pain management: an assessment of current evidence. Curr Pain Headache Rep 2020 Jan 24;24(2):4. [doi: 10.1007/s11916-020-0835-4] [Medline: 31980957]
11. Brown JD, Winterstein AG. Potential adverse drug events and drug-drug interactions with medical and consumer Cannabidiol (CBD) use. J Clin Med 2019 Jul 08;8(7):989 [FREE Full text] [doi: 10.3390/jcm8070989] [Medline: 31288397]
12. Mead A. The legal status of cannabis (marijuana) and cannabidiol (CBD) under U.S. law. Epilepsy Behav 2017 May;70(Pt B):288-291 [FREE Full text] [doi: 10.1016/j.yebeh.2016.11.021] [Medline: 28169144]
13. What you need to know (and what we're working to find out) about products containing cannabis or cannabis-derived compounds, including CBD. U.S. Food & Drug Administration. URL: https://www.fda.gov/consumers/consumer-updates/what-you-need-know-and-what-were-working-find-out-about-products-containing-cannabis-or-cannabis [accessed 2021-09-07]
14. Manthey J. Cannabis use in Europe: current trends and public health concerns. Int J Drug Policy 2019 Jun;68:93-96. [doi: 10.1016/j.drugpo.2019.03.006] [Medline: 31030057]
15. McGregor IS, Cairns EA, Abelev S, Cohen R, Henderson M, Couch D, et al. Access to cannabidiol without a prescription: a cross-country comparison and analysis. Int J Drug Policy 2020 Nov;85:102935 [FREE Full text] [doi: 10.1016/j.drugpo.2020.102935] [Medline: 32919298]
16. Shao G. Appendix: constitution of the people's Republic of China. In: Internet Law in China. Cambridge, UK: Chandos Publishing; 2012.
17. Misuse of drugs act 1973. Singapore Statutes Online. 2008. URL: https://sso.agc.gov.sg/Act/MDA1973 [accessed 2021-09-07]
18. China: a rising CBD superpower. Arcview Consultancy. URL: https://arcviewgroup.com/wp-content/uploads/2020/11/China-A-Rising-CBD-Superpower-August-2020-Whitepaper-ZL.pdf [accessed 2021-09-07]
19. Corroon J, Phillips JA. A cross-sectional study of cannabidiol users. Cannabis Cannabinoid Res 2018;3(1):152-161 [FREE Full text] [doi: 10.1089/can.2018.0006] [Medline: 30014038]
20. Whitehill JM, Trangenstein PJ, Jenkins MC, Jernigan DH, Moreno MA. Exposure to cannabis marketing in social and traditional media and past-year use among adolescents in states with legal retail cannabis. J Adolesc Health 2020 Feb;66(2):247-254 [FREE Full text] [doi: 10.1016/j.jadohealth.2019.08.024] [Medline: 31708374]
21. Tran T, Kavuluru R. Social media surveillance for perceived therapeutic effects of cannabidiol (CBD) products. Int J Drug Policy 2020 Mar;77:102688 [FREE Full text] [doi: 10.1016/j.drugpo.2020.102688] [Medline: 32092666]
22. Merten JW, Gordon BT, King JL, Pappas C. Cannabidiol (CBD): perspectives from Pinterest. Subst Use Misuse 2020;55(13):2213-2220. [doi: 10.1080/10826084.2020.1797808] [Medline: 32715862]
23. About of Reddit. Reddit. URL: https://www.redditinc.com/ [accessed 2021-09-07]
24. Regional distribution of desktop traffic to Reddit.com as of May 2022 by country. Statista. URL: https://www.statista.com/statistics/325144/reddit-global-active-user-distribution/ [accessed 2021-09-07]
25. Sattelberg W. The demographics of Reddit: who uses the site? Alphr. 2021 Apr 6. URL: https://www.alphr.com/demographics-reddit/ [accessed 2021-09-07]
26. About Xiaohongshu. Xiaohongshu. URL: https://www.xiaohongshu.com/protocols/about?language=en-US [accessed 2021-09-07]
27. Pemarathna R. Impact of Xiaohongshu on its user based and society: a review. Proc IRE 2019 Jun;2(11).
28. Taylor E. We agree, don't we? The Delphi method for health environments research. HERD 2020 Jan;13(1):11-23. [doi: 10.1177/1937586719887709] [Medline: 31887097]
29. Blei D, Ng A, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022.
30. Klakow D, Peters J. Testing the correlation of word error rate and perplexity. Speech Commun 2002 Sep;38(1-2):19-28. [doi: 10.1016/S0167-6393(01)00041-3]
31. Cross-validation of topic modelling. Free Range Statistics. 2017 Jan 5. URL: http://freerangestats.info/blog/2017/01/05/topic-model-cv [accessed 2021-09-07]
32. Zhao W, Chen JJ, Perkins R, Liu Z, Ge W, Ding Y, et al. A heuristic approach to determine an appropriate number of topics in topic modeling. BMC Bioinformatics 2015;16 Suppl 13(Suppl 13):S8 [FREE Full text] [doi: 10.1186/1471-2105-16-S13-S8] [Medline: 26424364]
33. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. 2015 Presented at: WSDM 2015: Eighth ACM International Conference on Web Search and Data Mining; Feb 2 - 6, 2015; Shanghai China. [doi: 10.1145/2684822.2685324]
34. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011 Presented at: EMNLP '11: Proceedings of the Conference on Empirical Methods in Natural Language Processing; Jul 27 - 31, 2011; Edinburgh United Kingdom.
35. Chang Y, Yeh W, Hsing Y, Wang C. Refined distributed emotion vector representation for social media sentiment analysis. PLoS One 2019;14(10):e0223317 [FREE Full text] [doi: 10.1371/journal.pone.0223317] [Medline: 31647844]
36. Lu L, Wang X. Drug addiction in China. Ann N Y Acad Sci 2008 Oct;1141(1):304-317. [doi: 10.1196/annals.1441.025] [Medline: 18991965]

37. Wilkinson ST, Yarnell S, Radhakrishnan R, Ball SA, D'Souza DC. Marijuana legalization: impact on physicians and public health. Annu Rev Med 2016;67:453-466 [FREE Full text] [doi: 10.1146/annurev-med-050214-013454] [Medline: 26515984]
38. Papayiannis S, Anastassiou-Hadjicharalambous X. Cross-cultural studies. In: Encyclopedia of Child Behavior and Development. Boston, MA: Springer; 2011.
39. Bender EM. On achieving and evaluating language-independence in NLP. Linguistic Issues Language Technol 2011 Oct 01;6. [doi: 10.33011/lilt.v6i.1239]
40. Westerman D, Spence PR, Van Der Heide B. Social media as information source: recency of updates and credibility of information. J Comput Mediat Comm 2013 Nov 08;19(2):171-183. [doi: 10.1111/jcc4.12041]
41. Wheeler M, Merten JW, Gordon BT, Hamadi H. CBD (Cannabidiol) product attitudes, knowledge, and use among young adults. Subst Use Misuse 2020;55(7):1138-1145. [doi: 10.1080/10826084.2020.1729201] [Medline: 32093530]
42. Naeem SB, Bhatti R, Khan A. An exploration of how fake news is taking over social media and putting public health at risk. Health Info Libr J 2021 Jun;38(2):143-149 [FREE Full text] [doi: 10.1111/hir.12320] [Medline: 32657000]

Abbreviations

CBD: cannabidiol
LDA: Latent Dirichlet Allocation
NLP: natural language processing
THC: tetrahydrocannabinol

Edited by A Mavragani, H Bradley; submitted 07.10.21; peer-reviewed by R Gore, AM Auvinen; comments to author 21.02.22; revised version received 14.05.22; accepted 25.11.22; published 11.01.23.

Please cite as:
Li Y, Yan X, Wang Z, Ma M, Zhang B, Jia Z
Comparison of the Users' Attitudes Toward Cannabidiol on Social Media Platforms: Topic Modeling Study
JMIR Public Health Surveill 2023;9:e34132
URL: https://publichealth.jmir.org/2023/1/e34132
doi:10.2196/34132
PMID:36630175

Original Paper

Vulnerability to HIV Infection Among International Immigrants in China: Cross-sectional Web-Based Survey

Yuyin Zhou[1,2], MD; Feng Cheng[3,4], PhD; Junfang Xu[1,2], PhD

[1]Center for Health Policy Studies, School of Public Health, Zhejiang University School of Medicine, Hangzhou, China
[2]Department of Pharmacy, Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[3]Vanke School of Public Health, Tsinghua University, Beijing, China
[4]Institute for Healthy China, Tsinghua University, Beijing, China

Corresponding Author:
Junfang Xu, PhD
Center for Health Policy Studies
School of Public Health
Zhejiang University School of Medicine
866 Yuhangtang Road
Hangzhou, 310058
China
Phone: 86 18801230482
Email: xujf2019@zju.edu.cn

Abstract

Background: The rising number of migrants worldwide, including in China given its recent rapid economic development, poses a challenge for the public health system to prevent infectious diseases, including sexually transmitted infections (STIs) caused by risky sexual behaviors.

Objective: The aim of this study was to explore the risky sexual behaviors of international immigrants living in China to provide evidence for establishment of a localized public health service system.

Methods: Risky sexual behaviors were divided into multiple sexual partners and unprotected sexual behaviors. Basic characteristics, sexual knowledge, and behaviors of international immigrants were summarized with descriptive statistics. Multivariate logistic regression analyses were used to identify factors associated with risky sexual behaviors, and the associations of demographic characteristics and risk behaviors with HIV testing and intention to test for HIV.

Results: In total, 1433 international immigrants were included in the study, 61.76% (n=885) of whom had never heard of STIs, and the mean HIV knowledge score was 5.42 (SD 2.138). Overall, 8.23% (118/1433) of the participants had been diagnosed with an STI. Among the 1433 international immigrants, 292 indicated that they never use a condom for homosexual sex, followed by sex with a stable partner (n=252), commercial sex (n=236), group sex (n=175), and casual sex (n=137). In addition, 119 of the international immigrants had more than three sex partners. Individuals aged 31-40 years were more likely to have multiple sexual partners (adjusted odds ratio [AOR] 2.364, 95% CI 1.149-4.862). Married participants were more likely to have unprotected sexual behaviors (AOR 3.096, 95% CI –1.705 to 5.620), whereas Asians were less likely to have multiple sexual partners (AOR 0.446, 95% CI 0.328-0.607) and unprotected sexual behaviors (AOR 0.328, 95% CI 0.219-0.492). Women were more likely to have taken an HIV test than men (AOR 1.413, 95% CI 1.085-1.841). Those who were married (AOR 0.577, 95% CI 0.372-0.894), with an annual disposable income >150,000 yuan (~US $22,000; AOR 0.661, 95% CI 0.439-0.995), considered it impossible to become infected with HIV (AOR 0.564, 95% CI 0.327-0.972), and of Asian ethnicity (AOR 0.330, 95% CI 0.261-0.417) were less likely to have an HIV test. People who had multiple sexual partners were more likely to have taken an HIV test (AOR 2.041, 95% CI 1.442-2.890) and had greater intention to test for HIV (AOR 1.651, 95% CI 1.208-2.258).

Conclusions: International immigrants in China exhibit risky sexual behaviors, especially those aged over 30 years. In addition, the level of HIV-related knowledge is generally low. Therefore, health interventions such as targeted, tailored programming including education and testing are urgently needed to prevent new HIV infections and transmission among international immigrants and the local population.

(JMIR Public Health Surveill 2023;9:e35713) doi:10.2196/35713

KEYWORDS

international immigrants; HIV; risky sexual behavior; China

Introduction

Under globalization, the number of international travelers has increased significantly worldwide during the past decade, from 2433 million in 2008 to 3188 million in 2019 [1,2]. Owing to its rapid economic development, China has become a popular country and attracted numerous international immigrants. For example, the National Bureau of Statistics showed that approximately 845,697 international immigrants lived in China in 2020, increasing by 251,865 compared to that recorded in 2010 [3,4]. Thus, an increasing number of international immigrants settled in China with different ethnic backgrounds and cultures, and this significant number of immigrants can inevitably cause some public health problems [5]. The health challenges posed by migration, not only for the immigrants themselves but also for residents, have become an important global health issue. For example, migration has proven to be a risk factor for sexually transmitted infections (STIs), especially with respect to the spread of HIV to residents of the destination country during unprotected sexual encounters [6]. Risky sexual behaviors might increase upon immigration given the removal of factors that might inhibit sexual freedom in the home country, including physical separation of partners or social networks, escaping from the pressure of home community or social culture, and especially an increase in casual sex [7].

The total number of people living with HIV/AIDS worldwide increased from 31.1 million in 2010 to 37.7 million in 2020 [8]. However, new HIV infections have declined by 31%, from 2.1 million in 2010 to 1.5 million in 2020 [8]. Migration has become one of the key issues related to HIV prevention and control in recent years considering the increasing number of HIV infections among migrants [9]. Indeed, since the beginning of the HIV epidemic, governments have been worried that migrants may be largely responsible for spreading HIV [10,11]. However, specific public health responses have not yet been established to monitor HIV among migrants, especially international immigrants, in most countries. Considering that the majority of international immigrants are sexually active young people and their interaction with local Chinese residents is inevitable, active responses for public health and related research are particularly important.

In China, the number of newly reported HIV-infected international immigrants reached up to 15,319 between 2004 and 2017 [12], although there is currently no mandatory screening for HIV or other STIs for international immigrants in China. Most studies regarding the sexual behaviors of international immigrants have been conducted in high-income countries such as the United States, Canada, and the United Kingdom [13-15]. The status of sexual behaviors among international immigrants living in China is currently unclear. Moreover, considering the differences in cultures and health services, interventions that are employed in high-income countries may not be suitable for international immigrants in China. Therefore, we aimed to explore the risky sexual behaviors of international immigrants living in China to provide evidence for establishment of a localized public health service system to address the significant increase in the number of international immigrants under globalization.

Methods

Participants

We used a cross-sectional web-based study design to collect data from international immigrants living in China by a snowball sampling method. Three to six international immigrants with different occupations, ages, home countries, and genders from Yiwu, Guangzhou, Beijing, and Hangzhou were invited to start the investigation. These immigrants were then asked to invite other international immigrants living in China to participate in the survey to expand the sample. We selected these locations in particular because Yiwu, as one of the largest commodity trading cities, attracts nearly half a million international immigrants every year [16]; Guangzhou is located in the core of the Pearl River Delta [17], and is also an important trading port representing one of the cities receiving a significant number of international immigrants; Beijing is the capital of China, which is bound to attract a substantial number of foreign people as the political center; and Hangzhou is the provincial capital of Zhejiang and located in the eastern coast of China, which has developed rapidly in recent years because of the high-tech service industry, resulting in a large foreign population.

The inclusion criteria of participants were: (1) international immigrants whose homeland was not China but were living in China, (2) above 18 years old, (3) could read English, and (4) willing to participate in the study. The exclusion criteria were: (1) under 18 years old; (2) Chinese nationality; and 3) did not pass the “attention check” in the questionnaire, which was used to identify careless respondents and improve the data quality.

The sample size was calculated based on the 31.14% HIV diagnosis risk among international immigrants according to previous studies [11,18], which required at least 330 participants. We also estimated the sample size based on 5-10 times the number of questions in the questionnaire, which required 650 respondents. To select the maximum representative sample from the population, 1460 participants were recruited and the responses of 1433 international immigrants were incorporated in the final analysis.

Data Collection

Considering the sensitivity of sexual issues and the COVID-19 pandemic, we used an online questionnaire rather than a face-to-face survey, which was conducted between January and September 2021, including sociodemographic information (eg, gender, age, marital status, education level, employment, and annual disposable income), HIV-related knowledge, STIs history (eg, diagnosis of genital herpes, syphilis, and condyloma in the past), HIV testing history (ie, have taken an HIV test in the past), and intention to have an HIV test in the future.

In addition, all participants were invited to recall their related sexual behaviors while living in China in the past year. Sexual behavior–related data included the number of sexual partners; condom use; stable, casual, commercial, homosexual, and group sexual behaviors; as well as illicit drug use. Risky sexual behaviors were classified as having multiple sexual partners and unprotected sexual behaviors. HIV-related knowledge was measured using the 8-item HIV Knowledge Questionnaire (HIV-KQ-8) [19]. If the answer was correct, a score of 1 was assigned, whereas a score of 0 was given if the answer was wrong or “do not know.” HIV/STIs knowledge was finally evaluated by the total scores, with a maximum score of 8 if all answers were correct; thus, a higher score represented a higher level of HIV/STIs knowledge among international immigrants. The questions regarding sexual behaviors were designed based on the guidelines of intervention work for the prevention of HIV/AIDS issued by the China Center for Disease Control and our previous study [20,21]. The reliability and validity of the questionnaires were confirmed using Cronbach α (.996) and the Kaiser-Meyer-Olkin test (0.961).

Ethics Considerations

The study protocol and consent procedure were approved by the Ethics Review Committee, School of Public Health, Zhejiang University (#2019-064). Informed consent information was provided before the questions; participants had the option to exit the survey after reading the informed consent information or to provide consent to continue. The confidentiality of individuals was properly protected in the management of the investigation and the processing of data.

Statistical Analysis

Descriptive statistics were used to analyze the basic characteristics, sexual-related knowledge, and sexual behaviors of international immigrants using frequency and percentage or mean (SD) as appropriate. Univariate and multivariate logistic regression analyses were used to identify the factors associated with risky sexual behaviors among international immigrants and the association with HIV test history and intention to receive an HIV test in the future. SPSS 24.0 statistical software (IBM, Armonk, NY, USA) was used to analyze all data. *P*<.05 was considered to indicate a statistically significant association.

Results

Basic Characteristics of International Immigrants in China

In total, 1433 international immigrants were included in the study, with a predominance of men (Table 1). The average age was approximately 25 years and most of the participants were well-educated with 68.95% (988/1433) reporting more than 11 years of education. The annual disposable income of the majority of the international immigrants was less than 50,000 yuan (~US $7200). The majority of the respondents were unmarried, and the largest proportion of immigrants were from Africa, followed by America, Europe, and Asia (Table 1).

Table 1. Basic characteristics of international immigrants in China (N=1433).

Characteristics	Value
Gender, n (%)	
Male	973 (67.90)
Female	460 (32.10)
Age (years), mean (SD)	24.97 (4.57)
Marital status, n (%)	
Unmarried	1263 (88.14)
Married	110 (7.68)
Windowed	4 (0.28)
Divorced	9 (0.63)
Other	47 (3.28)
Education level, n (%)	
Illiterate	49 (3.42)
1-5 years	299 (20.87)
6-10 years	97 (6.77)
11-12 years	157 (10.96)
>12 years	831 (58.00)
Employment, n (%)	
Employed	231 (16.12)
Unemployed	1202 (83.88)
Annual disposable income (yuan)[a], n (%)	
≤50,000	1040 (72.58)
50,001-100,000	212 (14.79)
100,001-150,000	79 (5.51)
>150,000	102 (7.12)
Home continent[b], n (%)	
Africa	679 (47.38)
America	67 (4.68)
Europe	40 (2.79)
Asia	610 (42.57)
Other	6 (0.42)

[a] Exchange rate on December 7, 2021: 6.3708 yuan=US $1.

[b] There were 31 responses missing.

Knowledge on HIV Among International Immigrants in China

Among the 1433 international immigrants, 61.76% (885) had never heard of STIs or HIV. Among the immigrants who had heard of HIV/STIs, the mean knowledge score on HIV was 5.42 (SD 2.138). The percentage of respondents with an HIV knowledge score <4, 4-6, and >6 was 6.98% (100/1433), 17.31% (248/1433), and 13.89% (199/1433), respectively.

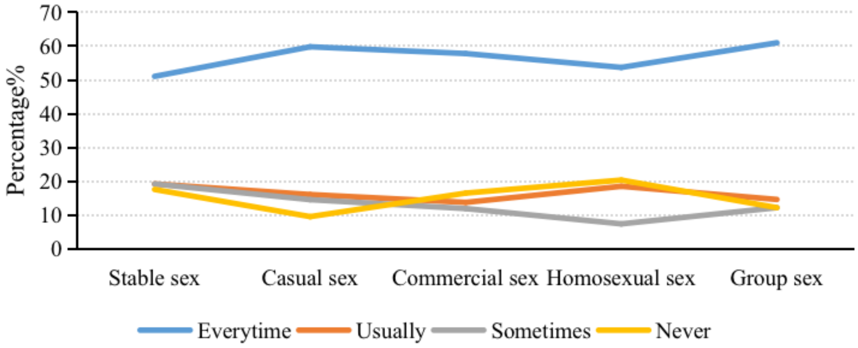
Self-Report of STIs Among International Immigrants

Overall, 91.77% (1315/1433) of international immigrants had not been diagnosed with any STIs, whereas 8.23% (118/1433) of the international immigrants had been diagnosed with at least one STI, including gonorrhea (16/1433, 1.12%), syphilis (12/1433, 0.84%), condyloma acuminatum (1/1433, 0.07%), herpes progenitalis (3/1433, 0.21%), hepatitis B (9/1433, 0.63%), and hepatitis C (2/1433, 0.14%). In addition, 4.82% (69/1433) of the participants had been diagnosed with HIV.

Risky Sexual Behaviors

As shown in Figure 1, more than 50% of international immigrants reported using a condom during sexual behaviors. However, 20.38% (292/1433) of international immigrants engaging in homosexual sex reported never using condoms, followed by 17.59% (252/1433) in stable sexual relationships, 16.47% (236/1433) engaging in commercial sex, 12.21% (175/1433) for group sex, and 9.56% (137/1433) for casual sex. In addition, 8.30% (119/1433) of international immigrants had more than three sexual partners.

Figure 1. Frequency of condom use according to sexual behavior among international immigrants.



Factors Associated With Risky Sexual Behaviors Among International Immigrants

The binary logistic regression model showed the factors affecting whether the respondents had multiple partners or unprotected sex as a risky sexual behavior (see Table S1 in Multimedia Appendix 1). We found that being aged 31-40 years was significantly associated with having multiple partners (adjusted odds ratio [AOR] 2.364, 95% CI 1.149-4.862). Married respondents indicated more unprotected risk behaviors (AOR 3.096, 95% CI –1.705 to 5.620) compared with unmarried respondents. Asians tended to have fewer multiple partners (AOR 0.446, 95% CI 0.328-0.607) and unprotected risk behaviors (AOR 0.328, 95% CI 0.219-0.492) compared with Africans.

Demographic Characteristics, Risk Behaviors, and Their Association With HIV Testing and Intention to Test for HIV

Table S2 in Multimedia Appendix 1 shows the associations of demographic characteristics and risk behaviors with HIV testing. Women had a higher probability of having taken an HIV test in the past than men (AOR 1.413, 95% CI 1.085-1.841). People who were married had less intention to take an HIV test (AOR 0.577, 95% CI 0.372-0.894) compared to unmarried respondents. Those considering it to be impossible to become infected with HIV were less likely to have ever taken an HIV test (AOR 0.523, 95% CI 0.281-0.975) or to intend to take an HIV test (AOR 0.564, 95% CI 0.327-0.972) compared with those who thought it could be possible to become infected. Compared to respondents who did not have multiple sex partners, those who reported having multiple sexual partners were more likely to intend to test for HIV (AOR 1.797, 95% CI 1.324-2.438). Furthermore, Asians were less likely to indicate an intention to be tested for HIV (AOR 0.330, 95% CI 0.261-0.417).

Discussion

Principal Findings

International immigrants in China exhibited risky sexual behaviors, with 22.40% and 13.19% of the respondents reporting having multiple sexual partners and unprotected sexual behaviors, respectively. These trends may be partly related to the differences in the new living environment faced by international immigrants from those of the homeland. For example, the absence of family and peer monitoring, language barriers, cultural barriers, and a sense of anonymity may offer immigrants more sexual freedom [22,23]. In addition, 4.82% of the respondents reported that they have been diagnosed with HIV. Therefore, the risk of HIV transmission and infection among international immigrants exists, and interventions are urgently needed to control the HIV epidemic caused by international migration in China.

HIV knowledge is a significant factor contributing to HIV prevention [14]. Previous studies also showed that enhanced knowledge moderated sexual risk behaviors, and HIV knowledge increased safer-sex intentions, condom use, abstinence, and HIV testing and treatment [24-27]. However, the majority of the international immigrants surveyed in our study had never heard of STIs, and the knowledge scores of the participants who were aware of HIV were also quite low. A study in Canada suggested that those born in sub-Saharan African also scored relatively low on the HIV knowledge questionnaire, which was consistent with our results [14]. Based on our analysis between Chinese nationals and international immigrants, the mean score of HIV-related knowledge of international immigrants who had heard of HIV (5.42, SD 2.14) was lower than that of Chinese nationals (6.68, SD 2.08), and there was a much higher number of international immigrants who had never heard of HIV (61.76%) compared to Chinese nationals (5.48%). These results were consistent with previous research [28]. The reasons may include that immigrants represent a marginalized group, who may not receive any HIV education in either their homeland or destination country. However, there are also discrepancies among studies related to the extent of HIV knowledge among immigrants. For example, a study in Thailand indicated that immigrants exhibited a fairly high standard of knowledge about the risk factors of HIV [5], and this knowledge increased among immigrants who had heard of HIV after they came to Thailand, which also showed that they obtained adequate education from the local government. Another study performed in South Korea also found that immigrant workers exhibited better knowledge of HIV compared with Korean respondents [29]. These differences may be caused by the fact that the participants in our study mostly migrated from low- and middle-income countries, where HIV-related resources are limited.

The age group of 30-60 years was significantly associated with having more risky sexual behaviors among the international immigrants surveyed in our study, which was consistent with the majority of previous related studies. This may be related to the fact that this age corresponds to a more physiologically sexually active group, who may have sex with multiple partners without protection. A previous study published in 2018 also reached a similar conclusion, showing that adults of this age group were more likely to have sex without a condom, new sexual partners, and multiple sexual partners [30]. In addition, consistent with a prior study showing that married people tended to have more high-risk sexual behaviors [31], we found that married individuals tended to have more unprotected sexual behaviors than unmarried individuals. There are several possible reasons for this difference. For example, a traditional culture that does not permit engaging in sexual behaviors before marriage could cause more risky sexual behaviors among married than unmarried individuals. In addition, married people may plan to have a baby, which would be accompanied by unprotected sex [32]. Moreover, some people find it shameful to use condoms once married or are pursuing the physical and psychological satisfaction from unprotected sex [13]. However, in our study, 16.13% of the married respondents indicated engaging in unprotected sexual behaviors not with their stable partners. For instance, a prior study also showed that married immigrant men tended to have affairs [33], and 28.48% of the migrant workers who lived apart from their spouses had extramarital sexual behaviors [34]. People engaging in casual sex, commercial sex, homosexual sex, and group sex are considered to be a high-risk population for STIs, and the frequency of condom use was rather low among international immigrants; therefore, health education is necessary and important for these international immigrants. Moreover, the environment of field testing for HIV typically has poor privacy protection, which may lead those at high risk of worrying about disclosure of personal information, especially for international immigrants who may be more insecure about whether they will be able to settle in the destination country if they are diagnosed with HIV.

With respect to the influence of the origin region of the immigrants, we found that Asians exhibited less risky sexual behaviors but also had reduced intention to be tested for HIV compared with Africans. The possible reasons for these differences may include that polygamy still exists in some areas of Africa, especially in some sub-Saharan African countries [35], and international immigrants from Africa tend to have little knowledge about HIV [36,37]. These factors may influence their sexual behaviors, especially increasing the risk of having multiple sexual partners and unprotected sexual behaviors. In addition, the fact that Africa had the highest number of people living with HIV in 2020, suggesting that Africans are most heavily affected by HIV [38], may cause their higher intention to receive an HIV test in the future compared with Asians. Among international immigrants, homosexuals were less likely to have ever tested for HIV than heterosexuals. Generally, homosexuals are considered to engage in risky sexual behaviors. Without an HIV test, the infection status would be unknown, which would place their partners at greater risk for the transmission of HIV and other STIs. Moreover, the proportion of HIV infections transmitted through homosexual behaviors has increased significantly in recent years [31]; thus, there is an urgent need to promote HIV testing among this high-risk group.

Limitations

This study had some limitations. First, it is possible that the participants were influenced by social and cultural norms to hide their risky sexual behavior, including a history of commercial or casual sexual behaviors, which may affect the research results, although an online survey could avoid the embarrassment caused by a face-to-face survey. Second, although the seed participants were selected to have different occupations, ages, homelands, and genders from different cities, a bias caused by snowball sampling may still exist, which could lead to misinterpretation of the related statistical results. In addition, considering the significant number of international immigrants living in China, our participants were mainly from lower socioeconomic strata with a lower annual disposable income, which represents a narrow band of the socioeconomic system. Therefore, the results may not be generalizable to the whole population. Third, the questionnaire regarding risky sexual behaviors was designed based on previous work and was amended according to our research participants and objectives. Thus, the validity of the questionnaire in the survey should be further tested.

Implications for Health Policy

Globalization leads to an inevitable trend of increased population mobility. Under this situation, it is necessary to take proactive public health approaches to manage the health of international immigrants, especially with respect to infectious diseases. Our study showed that the overall HIV-related knowledge of the international immigrants in China was low along with a low rate of intention to test for HIV. Therefore, HIV-related education is essential, such as by community advocacy and online propaganda delivered in their native languages, to address the gaps in HIV knowledge among international immigrants. More importantly, it is difficult for international immigrants to access HIV-related services because of the lack of specified health care and prevention upon arrival in the new country, which may cause immigrants with an HIV-positive status to remain undiscovered until some typical symptoms appear. To prevent the transmission of HIV, the basic public health care services provided by the Chinese government should be extended to cover international immigrants. Active promotion for antiviral treatment is less available for international immigrants with HIV infection than for local residents in China. Thus, an integrated health system for HIV prevention, monitoring, and treatment for immigrants and local residents will be important to reduce the transmission rate of HIV. Moreover, this is also necessary to reduce the inequities in access to health care and health outcomes for immigrants. It is also important to consider the wider health care and socioeconomic barriers immigrants face to be successful in improving HIV outcomes. In addition, as a convenient and efficient method, HIV self-testing could be promoted among international immigrants, such as setting up rapid detection kit vending machine in immigrant gathering communities [39]. Moreover, online measures, including HIV testing, education, and interventions, could be developed using the internet and big data to solve the privacy issues that are a major concern of international immigrants.

Conclusion

Risky sexual behaviors exist among international immigrants in China, including multiple partners and unprotected sexual behaviors, and the level of cognition of HIV knowledge was quite low, especially for married immigrants and those aged above 30 years. Therefore, there is an urgent need to promote health education and HIV testing specified for international immigrants. In addition, it is also important to develop and improve health care access for international immigrants under the current situation of increasing globalization.

Acknowledgments

We are grateful to all study participants for their valuable participation. This work was supported in part by the National Natural Science Foundation of China (Project No. 71904171); China Medical Board (Project No. 20-391 and 202033); Zhejiang Provincial Soft Science Research Project (Project No. 2021C35015); Research Fund, the Vanke School of Public Health, Tsinghua University (Project No. 2021ZZ004); Sanming Project of Medicine in Shenzhen (Project No. SZSM202111001); and National Key R&D Program of China during the 14th Five-Year Plan Period (Project No. 2021ZD0114102).

Availability of Data and Material

The views expressed in this publication do not represent the views of the Veterans Health Administration or the United States Government.

Authors' Contributions

All authors were responsible for the structure of this paper. YZ conducted the literature review, data analysis, and drafted the paper. JX contributed to the study’s conception and design, interpretation of the data, and critical revisions of the paper. FC provided interpretation and revisions of the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Factors associated with risky sexual behaviors (Table S1), and demographic characteristics, risk behaviors, and their association with HIV testing and intention to test for HIV (Table S2) among international immigrants.

[DOCX File , 45 KB - publichealth_v9i1e35713_app1.docx]

References

1. The number of inbound tourists reached 130 million last year. The Central People's Government of the People's Republic of China. 2009. URL: http://www.gov.cn/govweb/jrzg/2009-02/26/content_1243877.htm [accessed 2021-12-12]
2. China's inbound tourism development report 2020. Ministry of Culture and Tourism of the People's Republic of China. URL: https://www.mct.gov.cn/whzx/zsdw/zglyyjy/202102/t20210201_921194.htm [accessed 2021-12-12]
3. Tabulation on the 2010 Population Census of the People's Republic of China. National Bureau of Statistics. 2010. URL: http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm [accessed 2021-12-12]
4. Bulletin of the Seventh National Census (No. 8). National Bureau of Statistics. 2021. URL: http://www.stats.gov.cn/ztjc/zdtjgz/zgrkpc/dqcrkpc/ggl/202105/t20210519_1817701.html [accessed 2021-12-12]
5. Fuller TD, Chamratrithirong A. Knowledge of HIV risk factors among immigrants in Thailand. J Immigr Minor Health 2009 Apr 5;11(2):83-91. [doi: 10.1007/s10903-008-9163-1] [Medline: 18604584]
6. Matteelli A, Schlagenhauf P, Carvalho AC, Weld L, Davis XM, Wilder-Smith A, GeoSentinel Surveillance Network. Travel-associated sexually transmitted infections: an observational cross-sectional study of the GeoSentinel surveillance database. Lancet Infect Dis 2013 Mar;13(3):205-213. [doi: 10.1016/S1473-3099(12)70291-8] [Medline: 23182931]
7. Jirattikorn A, Tangmunkongvorakul A, Musumari PM, Ayuttacorn A, Srithanaviboonchai K, Banwell C, et al. Sexual risk behaviours and HIV knowledge and beliefs of Shan migrants from Myanmar living with HIV in Chiang Mai, Thailand. Int J Migr Health Soc Care 2020 Nov 02;16(4):543-556. [doi: 10.1108/ijmhsc-09-2019-0080]
8. 2021 UNAIDS Global AIDS Update — Confronting inequalities — Lessons for pandemic responses from 40 years of AIDS. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2021 Jul 14. URL: https://www.unaids.org/en/resources/documents/2021/2021-global-aids-update [accessed 2021-12-12]

9. Hargreaves S, Nellums L, Friedland JS. Time to rethink approaches to migrant health screening. Lancet 2016 Nov 19;388(10059):2456-2457 [FREE Full text] [doi: 10.1016/S0140-6736(16)31703-2] [Medline: 27742166]
10. Carballo M, Hargreaves S, Gudumac I, Maclean EC. Evolving migrant crisis in Europe: implications for health systems. Lancet Glob Health 2017 Mar;5(3):e252-e253 [FREE Full text] [doi: 10.1016/S2214-109X(17)30040-2] [Medline: 28193385]
11. Fakoya I, Álvarez-del Arco D, Woode-Owusu M, Monge S, Rivero-Montesdeoca Y, Delpech V, et al. A systematic review of post-migration acquisition of HIV among migrants from countries with generalised HIV epidemics living in Europe: mplications for effectively managing HIV prevention programmes and policy. BMC Public Health 2015 Jun 19;15(1):561 [FREE Full text] [doi: 10.1186/s12889-015-1852-9] [Medline: 26085030]
12. Wu Y. Epidemiological characteristics of infectious diseases and health service utilization for foreign population in China. Doctoral dissertation. China National Knowledge Infrastructure. 2020. URL: https://tinyurl.com/yckdpymv [accessed 2023-01-03]
13. Persichino J, Ibarra L. HIV and Latino migrant workers in the USA. Ethn Rac Stud 2012 Jan;35(1):120-134. [doi: 10.1080/01419870.2011.594167]
14. Tulloch HE, Balfour L, Kowal J, Tasca GA, Angel JB, Garber G, et al. HIV knowledge among Canadian-born and sub-Saharan African-born patients living with HIV. J Immigr Minor Health 2012 Mar 4;14(1):132-139. [doi: 10.1007/s10903-011-9480-7] [Medline: 21643728]
15. Barrett HR, Mulugeta B. Human immunodeficiency virus (HIV) and migrant "risk environments": the case of the Ethiopian and Eritrean immigrant community in the West Midlands of the UK. Psychol Health Med 2010 May 17;15(3):357-369. [doi: 10.1080/13548501003653192] [Medline: 20480439]
16. Yiwu's experience of expatriate management: inclusiveness to achieve "Little United Nations". ChinaNews. 2017. URL: http://www.chinanews.com/sh/2017/06-19/8255098.shtml [accessed 2021-12-12]
17. Cai X, Liu M, Wen Y. Spatial-temporal evolvement characteristics and mechanism of foreign immigrants in Guangzhou. J South China Normal Univ 2019;3:92-31 [FREE Full text]
18. Transcript of the regular press conference of the National Health Commission on November 23, 2018. National Health Commission of China. URL: http://www.nhc.gov.cn/wjw/xwdt/201811/b0ca3817224e44899a3248a0f6e48948.shtml [accessed 2022-06-12]
19. Han M, Hu Z. China HIV/AIDS Monitoring and Evaluation Framework (Trial). Beijing, China: People's Medical Publishing House; 2007.
20. Xu J, Luo Y, Dong H, Zhao G. The effects of internet exposure on sexual risk behavior among sexually experienced male college students in China: cross-sectional study. JMIR Public Health Surveill 2022 May 02;8(5):e31847 [FREE Full text] [doi: 10.2196/31847] [Medline: 35499864]
21. Xu J, Wang P, Cheng F. Health related behaviors among HIV-infected people who are successfully linked to care: an institutional-based cross-sectional study. Infect Dis Poverty 2020 Mar 10;9(1):28 [FREE Full text] [doi: 10.1186/s40249-020-00642-1] [Medline: 32169118]
22. Qiu P, Yang Y, Wu F, Cao X, Zhao S, Ma X. Progress and enlightenment of home and abroad researches on mental health of mobile population. Chin Mental Health J 2010;1:64-68 [FREE Full text]
23. UNAIDS and IOM. Migration and AIDS. Int Migr 2002 Dec 16;36(4):445-468. [doi: 10.1111/1468-2435.00056]
24. Scott-Sheldon LA, Carey MP, Carey KB, Cain D, Simbayi LC, Mehlomakhulu V, et al. HIV testing is associated with increased knowledge and reductions in sexual risk behaviours among men in Cape Town, South Africa. Afr J AIDS Res 2013 Dec;12(4):195-201 [FREE Full text] [doi: 10.2989/16085906.2013.863219] [Medline: 25871481]
25. Bazargan M, Kelly EM, Stein JA, Husaini BA, Bazargan SH. Correlates of HIV risk-taking behaviors among African-American college students: the effect of HIV knowledge, motivation, and behavioral skills. J Natl Med Assoc 2000 Aug;92(8):391-404. [Medline: 10992684]
26. Ford K, Wirawan DN, Muliawan P. Social influence, AIDS/STD knowledge, and condom use among male clients of female sex workers in Bali. AIDS Educ Prev 2002 Dec;14(6):496-504. [doi: 10.1521/aeap.14.8.496.24113] [Medline: 12512850]
27. Schroder KEE, Carey MP. Anger as a moderator of safer sex motivation among low-income urban women. J Behav Med 2005 Oct 25;28(5):493-506 [FREE Full text] [doi: 10.1007/s10865-005-9014-7] [Medline: 16247592]
28. Goldenberg SM, Liu V, Nguyen P, Chettiar J, Shannon K. International migration from non-endemic settings as a protective factor for HIV/STI risk among female sex workers in Vancouver, Canada. J Immigr Minor Health 2015 Mar 4;17(1):21-28 [FREE Full text] [doi: 10.1007/s10903-014-0011-1] [Medline: 24700025]
29. Jung M. Immigrant workers' knowledge of HIV/AIDS and their sexual risk behaviors: a respondent-driven sampling survey in South Korea. Sex Disabil 2011 Dec 17;30(2):199-208. [doi: 10.1007/s11195-011-9245-1]
30. Santa Maria D, Hernandez DC, Arlinghaus KR, Gallardo KR, Maness SB, Kendzor DE, et al. Current Age, age at first sex, age at first homelessness, and HIV risk perceptions predict sexual risk behaviors among sexually active homeless adults. Int J Environ Res Public Health 2018 Jan 26;15(2):218 [FREE Full text] [doi: 10.3390/ijerph15020218] [Medline: 29373568]
31. Jiang H, Hong H, Dong H, Jiang J, He L. HIV testing and risks of sexual behavior among HIV-negative men who have sex with men in Ningbo, China. Int J Environ Res Public Health 2020 Mar 19;17(4):1322 [FREE Full text] [doi: 10.3390/ijerph17041322] [Medline: 32092851]

32. Khan ME, Townsend JW, D'Costa S. Behind closed doors: a qualitative study of sexual behaviour of married women in Bangladesh. Cult Health Sex 2002 Jan;4(2):237-256. [doi: 10.1080/13691050110102253]
33. Viadro CI, Earp JAL. The sexual behavior of married Mexican immigrant men in North Carolina. Soc Sci Med 2000 Mar;50(5):723-735. [doi: 10.1016/s0277-9536(99)00305-6] [Medline: 10658852]
34. Li S, Zhong Z, Lei X, Zhang M, Xu X, Yan W. A study on sexual behavior status and social strategies of married migrant workers in Chongqing. Chinese Journal of Evidence-Based Medicine 2011;11(3):257-260 [FREE Full text]
35. Fenske J. African polygamy: past and present. J Dev Econ 2015 Nov;117:58-73. [doi: 10.1016/j.jdeveco.2015.06.005]
36. World Health Organization. Baseline report on global sexually transmitted infection surveillance 2012. Geneva: World Health Organization; 2013.
37. General situation of legal infectious diseases in China in 2020-2021. Bureau of Disease Control and Prevention. URL: http://www.nhc.gov.cn/jkj/s3578/202103/f1a448b7df7d4760976fea6d55834966.shtml [accessed 2021-12-12]
38. Key facts HIV. World Health Organization. 2020. URL: https://cdn.who.int/media/docs/default-source/hq-hiv-hepatitis-and-stis-library/key-facts-hiv-2020.pdf?sfvrsn=582c3f6e_8 [accessed 2021-12-12]
39. Cheng S, Han Z, Liu F, Cao H, Xu H, Li Y. Analysis on influencing factors on HIV testing behaviors in some foreigners in Guangzhou. Chinese Journal of Epidemiology 2020;41(07):1092-1096. [doi: 10.3760/cma.j.cn112338-20200219-00123]

Abbreviations

AOR: adjusted odds ratio
HIV-KQ-8: 8-item HIV Knowledge Questionnaire
STI: sexually transmitted infection

Edited by A Mavragani, H Bradley; submitted 14.12.21; peer-reviewed by H Akram, L Lafrado; comments to author 10.05.22; revised version received 04.07.22; accepted 25.11.22; published 10.01.23.

Please cite as:
Zhou Y, Cheng F, Xu J
Vulnerability to HIV Infection Among International Immigrants in China: Cross-sectional Web-Based Survey
JMIR Public Health Surveill 2023;9:e35713
URL: https://publichealth.jmir.org/2023/1/e35713
doi:10.2196/35713
PMID:36626224

Original Paper

Risk and Protective Profile of Men Who Have Sex With Men Using Mobile Voluntary HIV Counseling and Testing: Latent Class Analysis

Piao-Yi Chiou[1,2,3], PhD; Wei-Wen Tsao[1], BA; Kuan-Chia Lin[4,5], PhD; Yuan-Yuan Fang[6], PhD; Kuan-Yin Lin[7], PhD; Chia-Lin Li[3,8], PhD

[1]School of Nursing, College of Medicine, National Taiwan University, Taipei City, Taiwan
[2]Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan
[3]Research and Development Committee, Taiwan AIDS Nurse Association, Taipei, Taiwan
[4]Community Medicine Research Center, National Yang Ming Chiao Tung University, Taipei, Taiwan
[5]Medical Affairs, Cheng Shin General Hospital, Taipei, Taiwan
[6]Department of Post Baccalaureate Nursing, College of Medicine, I-Shou University, Kaohsiung, Taiwan
[7]Department of Internal Medicine, National Taiwan University Hospital, National Taiwan University College of Medicine, Taipei, Taiwan
[8]Center for Neuropsychiatric Research, National Health Research Institutes, Taipei, Taiwan

Corresponding Author:
Piao-Yi Chiou, PhD
School of Nursing
College of Medicine
National Taiwan University
No 2-1 Xuzhou Road, Zhongzheng District
Taipei City, 100
Taiwan
Phone: 886 2 23123123 ext 288427
Fax: 886 2 23219913
Email: purechiou2@gmail.com

Abstract

Background: Mobile voluntary counseling and testing (VCT) for HIV has been carried out to improve the targeting of at-risk populations and HIV case detection for men who have sex with men (MSM). However, the HIV-positive detection rate using this screening strategy has declined in recent years. This may imply unknown changes in risk-taking and protective features jointly influencing the testing results. These changing patterns in this key population remain unexplored.

Objective: The aim of this study was to identify the nuanced group classification of MSM who underwent mobile VCT using latent class analysis (LCA), and to compare the difference in characteristics and testing results between subgroups.

Methods: A cross-sectional research design and purposive sampling were applied between May 21, 2019, and December 31, 2019. Participants were recruited by a well-trained research assistant through social networking platforms, including the most popular instant messenger app Line, geosocial network apps dedicated to MSM, and online communities. Mobile VCT was provided to participants at an assigned time and place. Demographic characteristics and risk-taking and protective features of the MSM were collected via online questionnaires. LCA was used to identify discrete subgroups based on four risk-taking indicators—multiple sexual partners (MSP), unprotected anal intercourse (UAI), recreational drug use within the past 3 months, and history of sexually transmitted diseases—and three protective indicators—experience of postexposure prophylaxis, preexposure prophylaxis use, and regular HIV testing.

Results: Overall, 1018 participants (mean age 30.17, SD 7.29 years) were included. A three-class model provided the best fit. Classes 1, 2, and 3 corresponded to the highest risk (n=175, 17.19%), highest protection (n=121, 11.89%), and low risk and low protection (n=722, 70.92%), respectively. Compared to those of class 3, class 1 participants were more likely to have MSP and UAI within the past 3 months, to be ≥40 years of age (odds ratio [OR] 2.197, 95% CI 1.357-3.558; *P*=.001), to have HIV-positive results (OR 6.47, 95% CI 2.272-18.482; *P*<.001), and a CD4 count ≤349/μL (OR 17.50, 95% CI 1.223-250.357; *P*=.04). Class

2 participants were more likely to adopt biomedical preventions and have marital experience (OR 2.55, 95% CI 1.033-6.277; *P*=.04).

Conclusions: LCA helped derive a classification of risk-taking and protection subgroups among MSM who underwent mobile VCT. These results may inform policies for simplifying the prescreening assessment and more precisely recognizing those who have higher probabilities of risk-taking features but remain undiagnosed targets, including MSM engaging in MSP and UAI within the past 3 months and those ≥40 years old. These results could be applied to tailor HIV prevention and testing programs.

(JMIR Public Health Surveill 2023;9:e43394) doi:10.2196/43394

KEYWORDS

HIV testing; latent class analysis; men who have sex with men; mobile health; postexposure prophylaxis (PEP); preexposure prophylaxis (PrEP); risk-taking; anal sex; sexual partners; social networking

Introduction

Background

Owing to the success of antiretroviral therapy for people infected with HIV, and the promotion of HIV screening and prophylactic medications for at-risk populations [1-3], the annual incidence rate of HIV infections among adults has declined globally by 31% between 2010 and 2021 [4]. In Taiwan, an estimated 90% of people living with HIV (PLWH) know their HIV status and 93% of them have initiated antiretroviral treatment, 95% of whom have achieved viral suppression [5]. To reach the World Health Organization’s goal of eliminating HIV infection by 2030, improving the access of the remaining undiagnosed target groups to screening is a critical strategy in HIV case detection and epidemic control [6].

The remaining undiagnosed PLWH require aggressive testing measures to facilitate an early diagnosis. Mobile voluntary counseling and testing (VCT) for HIV is a novel screening method in Taiwan, carried out by disseminating recruitment messages through social networking platforms to deliver testing services to specific groups at a specific time and location [7]. Compared to an outreach model of screening based at a community station, mobile VCT is more flexible with respect to the time and location of testing; therefore, the targeting of at-risk populations and case discovery are improved in the mobile VCT model [8,9]. However, the number of newly diagnosed HIV infections in Taiwan has been declining annually since 2017 [10] and the number of people that are positive for HIV associated with mobile VCT is also decreasing [7], which implies unknown changes in HIV risk-taking patterns and the influence of protective features in people who are uninfected. A possible inference could be that people who were previously at risk of HIV may have reduced their risk-taking behaviors, including avoiding having multiple sexual partners (MSP) and having unprotected anal intercourse (UAI), and abstaining from recreational drug use; concurrently, they may have adopted higher biomedical protection measures, including regular HIV testing, and when necessary, taking postexposure prophylaxis (PEP) and preexposure prophylaxis (PrEP) [11-14]. To provide HIV testing resources to those with relatively high priority needs in a mobile VCT service, it is necessary to be able to precisely recognize those who encompass a higher probability of risk-taking features and a lower probability of protective features. However, the feature classification of MSM who participate in mobile VCT remains unknown.

Latent class analysis (LCA) is a mixture model that is used to identify latent heterogeneity driven by class membership according to responses to a set of observed variables [15]. The theoretical assumption underlying LCA is that an individual shares both latent and observed categorical variables, and membership in unobserved classes can cause or explain patterns of scores across assessment indicators [16,17]. More specifically, LCA helps to identify similar characteristics among individuals to derive subgroups. LCA has been used to discern patterns of HIV risk based on demographic characteristics [18,19] and behaviors [20,21], helping define and target distinct groups. According to the theoretical basis and hypotheses of LCA, as well as previous research findings, LCA is appropriate for analysis of unidentified patterns in MSM who receive mobile VCT.

Objective

The aim of this study was to identify the nuanced group classification based on the risk and protective profile of MSM who accepted mobile VCT using LCA, and then to compare the characteristics and HIV testing results between subgroups to further determine groups of individuals with different patterns.

Methods

Study Design

A cross-sectional research design and purposive sampling were applied. Participants were recruited via social networking platforms. After completing an online questionnaire, we offered mobile VCT at a designated time and location, according to the participants’ preferences, where they could obtain a rapid HIV test.

Participants

The inclusion criteria were age ≥20 years, being literate, self-identification as an MSM, and self-reported as participating in HIV risk-associated behavior. The exclusion criteria included no self-reported risk factors for HIV infection or being HIV-positive.

Ethical Considerations

This study was approved by the Research Ethics Committee of National Taiwan University (201903ES024), and adhered to the regulations of the Human Subjects Research Act issued by the Ministry of Health and Welfare of Taiwan and the Personal Data Protection Act issued by the National Development

Council in Taiwan. After explaining the study purpose and protocol, the participants were asked to complete an online consent form and questionnaires anonymously through a provided URL; all data were processed anonymously. Participants could withdraw at any time without providing a reason. A mask could be worn during the mobile VCT processes to protect privacy. Each participant was eligible to receive an electronic gift certificate worth US $3.27 after all procedures were completed.

Procedure and Data Collection

The study covered the period from May 21, 2019, to December 31, 2019. Participants were recruited in the same manner as described in a previous study [7]. Social networking platforms were used, including the most popular instant messenger app Line and geosocial network apps dedicated to MSM, such as Grindr, Hornet, Jack'd, Scruff, and Blued. Online communities (such as Facebook fan pages) were used to disseminate the study information and recruit participants. Testing appointments were discussed and booked through one-on-one messaging on each recruiting platform. Mobile VCT was provided by a well-trained research assistant at a designated time and location, such as a convenience store, coffee shop, park, or fast-food restaurant, for participants who live in Taipei City and New Taipei City. Study information was provided to the participants, whose identity was verified through previous messaging records. Prior to the test, participants were asked to fill out an online consent form and questionnaire, which took about 10-15 minutes to complete. Following submission of the questionnaire, the researcher and participant would review and confirm whether there were any missing or questionable answers to ensure completeness of the questionnaire. Pretest counseling was provided and a blood drop was collected via a rapid HIV testing kit. Testing results were available immediately and recorded by the research assistant, and posttest counseling and a referral to the hospital HIV case manager to confirm testing and obtain PEP and PrEP were provided. The whole screening process took approximately 30-40 minutes. The CD4+ T-cell counts of each participant were followed and recorded by the research assistant.

Instruments

Questionnaire

The online questionnaire included 18 closed-ended questions; six were focused on demographic characteristics and 12 were focused on risk-taking and protective features. Demographic characteristics of interest included age, education level, marital experience, monthly income, sexual orientation, and if participants were open about their MSM status. HIV risk-taking features included preferred sexual position; number of sexual partners within the past 3 months; frequency of condom use during the previous 10 occasions of anal intercourse within the past 3 months; experience of recreational drug use within the past 3 months, including the number of types of drugs; and history of sexually transmitted diseases (STDs; including syphilis, gonorrhea, genital herpes, chlamydia, human papillomavirus, pubic lice, and hepatitis A, B, and C). HIV-preventing features included awareness of PEP and PrEP, experience of using PEP and PrEP, experience of HIV testing, and frequency of HIV testing.

Rapid HIV Self-Testing Kits

The Alere Determine HIV-1/2 rapid testing kit was used to test for HIV-1/HIV-2 antibodies in the human serum or plasma. This kit is licensed by the Taiwan Food and Drug Administration, and has specificity and sensitivity values of 99.87% and 99.75%, respectively [22]. Results were obtained within approximately 15 minutes and were managed and recorded by trained staff.

Indicator Selection for LCA

Risk-taking– and protection-related indicators were considered conditionally independent and were defined based on previous studies [11-14] and expert input; risk-taking–related indicators were defined as behaviors that directly lead to HIV infection or the features that significantly increase the risk of HIV infection, and the protection-related indicators were defined as proactive HIV screening or preventive drug administration. Three specialists, including an infectious disease physician, HIV case manager, and HIV public health scientist, provided expert input. Finally, seven indicators were selected. The four risk indicators included MSP, UAI, recreational drug use within the past 3 months, and history of STDs. The three protective indicators included experience of PEP use, experience of PrEP use, and regular HIV testing (at 3-6–month intervals when having unprotected sex). All variables were dichotomized (yes vs no).

Statistical Analyses

All data were confirmed to be complete prior to analysis. LCA was performed using the poLCA package within R software (GitHub, Inc) to identify distinct classes based on model fit. The multiple fit statistics, including log likelihood (LL), likelihood ratio test statistic (G^2), Akaike information criterion (AIC), Bayesian information criterion (BIC), and χ^2 goodness of fit, were applied and considered prior to the decision on the model being made [17]. BIC was identified as the best reliable fit statistic model, which is an intuitive calculation method that is commonly used for model selection in univariate and multivariate logistic regression due to its simplicity, and considers the effect of a large sample size to prevent the model from being too complex as a result of high model accuracy [17,23,24]. In addition, we examined the similarity of item-response probability patterns within classes and the interpretability of class separation. Subgroup membership was assigned based on the class of each participant and the highest probability of belonging. A probability of ≥50% per item indicated that group members were more likely to possess a characteristic of interest [19]. The univariable multinomial logistic regression model was applied to determine differences in class characteristics and testing outcomes.

Results

Study Population

In total, 1023 participants made an appointment for mobile VCT via social networking platforms during the study period. Five (0.49%) participants did not attend the testing appointment and lost contact without a reason. A total of 1018 participants completed the questionnaire and entered the screening process. The participants' characteristics are presented in Table 1. The participants' median age was 30.17 (IQR 25-34) years. Most participants were in the age range of 20-29 years, had college- or university-level education, had never been married, and had a stable monthly income of approximately US $714-$1783. Most participants were MSM, and had had a come-out experience, preferred both top and bottom sex positions, had both casual partners and MSP, and frequently used condoms (7-9 times per 10 sexual encounters). Some participants self-reported using recreational drugs (116/1018, 11.39%) within the past 3 months, with the majority indicating the use of one or two substances. Approximately one-fifth had a history of STDs, including, human papillomavirus (106/222, 47.7%), syphilis (71/222, 31.98%), gonorrhea (50/222, 22.5%), hepatitis B (33/222, 14.9%), pubic lice (31/222, 13.96%), genital herpes (18/222, 8.1%), chlamydia (11/222, 4.95%), hepatitis A (11/222, 4.95%), and hepatitis C (10/222, 4.5%). Most participants were aware of PEP, although few had used it. Most participants were aware of PrEP, although again few had used it. Furthermore, most participants had experience of HIV testing, including testing every 3-6 months. Eighteen participants were newly HIV-positive; 9 of whom had a CD4 count ≥350/μL and nine of whom had a CD4 count ≤349/μL.

Table 1. Descriptive characteristics of the men who have sex with men using mobile voluntary HIV counseling and testing (N=1018).

Characteristic	Participants, n (%)
Age group (years)	
20-29	545 (53.54)
30-39	361 (35.46)
≥40	112 (11.00)
Education	
High school or below high school	121 (11.89)
College or university	733 (72.00)
Postgraduate	164 (16.11)
Marital experience	
No	985 (96.76)
Yes	33 (3.24)
Monthly income (US $)	
≤713	214 (21.02)
714-1783	640 (62.87)
≥1784	164 (16.11)
Sexual orientation	
MSM[a]	844 (82.91)
Bisexual	174 (17.09)
Have come out	
No	165 (16.21)
Yes	853 (83.79)
Preferred sexual position	
Top	265 (26.03)
Bottom	287 (28.19)
Both bottom and top	466 (45.78)
Sexual partner[b]	
No sexual partner	143 (14.05)
A single main sexual partner	235 (23.08)
Casual and multiple sexual partners	640 (62.87)
Condom use during past 10 times of anal intercourse[b]	
No sexual behavior	83 (8.15)
Every time used (10 times)	287 (28.19)
Frequently used (7-9 times)	346 (33.99)
Occasionally used (4-6 times)	144 (14.15)
Rarely used (1-3 times)	99 (9.72)
Never used (0 times)	59 (5.80)
Recreational drug use[b]	
No	902 (88.61)
Yes	116 (11.39)
Number of recreational drugs used (n=116)[b]	
1–2	76 (65.5)

Characteristic	Participants, n (%)
3–4	24 (20.7)
≥5	16 (13.8)
Experience of STDs[c]	
No	796 (78.19)
Yes	222 (21.81)
Have ever heard of and know of PEP[d]	
No	162 (15.91)
Yes	856 (84.09)
Experience of using PEP	
No	887 (87.13)
Yes	131 (12.87)
Have ever heard of and know of PrEP[e]	
No	121 (11.89)
Yes	897 (88.11)
Experience of using PrEP	
No	891 (87.52)
Yes	127 (12.48)
Experience of HIV testing	
No	162 (15.91)
Yes	856 (84.09)
HIV testing frequency (n=856)	
First time	162 (15.91)
Every 3-6 months	630 (61.89)
6 months and longer	226 (22.20)
HIV test result	
Negative	1000 (98.23)
Newly positive	18 (1.77)
CD4 count (n=18)	
≥350/μL	9 (50.0)
≤349/μL	9 (50.0)

[a] MSM: men who have sex with men.

[b] Within the past 3 months.

[c] STD: sexually transmitted disease (including syphilis, gonorrhea, genital herpes, chlamydia, human papillomavirus, pubic lice, and hepatitis A, B, and C).

[d] PEP: postexposure prophylaxis.

[e] PrEP: preexposure prophylaxis.

LCA Results

After entering the seven indicators selected for LCA, the outcomes of 2-6 latent classes were produced and compared. The final model was determined according to the lowest value of BIC and class separation, although the lowest values of LL, AIC, G^2, and χ^2 were inconsistent with those of BIC (Table 2). A three-class model was selected, based on fit statistics and model interpretation. Table 3 presents the probabilities of HIV risk-taking and the protective profile of the three-class model.

Classes 1, 2, and 3 represented the highest risk (n=175, 17.2%), highest protection (n=121, 11.9%), and low risk and low protection (n=722, 70.9%), respectively. Participants in class 1 had the highest risk of having MSP and UAI; they were also more likely to get tested regularly. Participants in class 2 had the highest probabilities of protective features, including experience of PEP and PrEP and regular HIV testing, whereas they had a medium risk of having MSP and UAI. Compared to those of classes 1 and 2, participants in class 3 had a lower risk of having MSP and UAI, and they tended to get regularly tested, which was considered an HIV-protection feature (Table 3).

Table 2. Goodness of fit indices for model selection.

Model	LL[a]	G^{2b}	AIC[c]	BIC[d]	χ^2	*df*	Number of parameters
1-class	–3683.55	471.03	7381.11	7415.59	894.34	120	7
2-class	–3566.28	236.48	7162.56	7236.44	270.48	112	15
3-class[e]	–3524.05	152.02	7094.10	7207.39	187.85	104	23
4-class	–3509.56	123.04	7081.12	7233.81	151.71	96	31
5-class	–3498.77	101.46	7075.53	7267.63	109.18	88	39
6-class	–3488.17	80.26	7070.34	7301.84	80.26	80	47

[a] LL: log likelihood.

[b] G^2: likelihood ratio test statistic.

[c] AIC: Akaike information criterion.

[d] BIC: Bayesian information criterion.

[e] The selected class based on the lowest BIC value.

Table 3. Posterior probabilities (%) for the three-class model (N=1018).[a]

Indicators	Class 1, highest risk (n=175, 17.2%)	Class 2, highest protection (n=121, 11.9%)	Class 3, low risk and low protection (n=722, 70.9%)
Risk-taking indicators			
Multiple sexual partners[b]	96.5	66.3	51.1
Unprotected anal intercourse[b,c]	100.0	58.8	52.7
Experience of recreational drug use[b]	23.2	13.4	7.1
Experience of STDs[d]	43.8	24.6	14.0
Protective indicators			
Experience of using PEP[e]	8.4	76.3	2.8
Experience of using PrEP[f]	16.1	67.0	1.3
Regular HIV testing[g]	74.8	94.4	51.7

[a] Probabilities greater than 50% indicate items for which members of a given class were more likely to be classified within that class.

[b] Within the past 3 months.

[c] Not using a condom every time during past 10 encounters of anal intercourse.

[d] STD: sexually transmitted disease (including syphilis, gonorrhea, genital herpes, chlamydia, human papillomavirus, pubic lice, and hepatitis A, B, and C).

[e] PEP: postexposure prophylaxis.

[f] PrEP: preexposure prophylaxis.

[g] Testing frequency between 3 and 6 months in the case of regular unprotected sex.

Univariable Multinomial Logistic Regression Results

Class 3 was used as a reference due to the lower probability of HIV risk and protective features. Compared to those of class 3, class 1 participants were more likely to be ≥40 years old, to have an HIV-positive test result, and a CD4 count ≤349/μL. Class 2 participants were more likely to have been married (Table 4 and Table 5).

Table 4. Demographics and testing results of the three classes (N=1018).

Variable	Class 1[a] (n=175, 17.19%), n (%)	Class 2[b] (n=121, 11.89%), n (%)	Class 3[c] (n=722, 70.92%), n (%)
Age group (years)			
20-29	85 (48.6)	69 (57.1)	391 (54.2)
30-39	58 (33.1)	39 (32.2)	264 (36.3)
≥40	32 (18.3)	13 (10.7)	67 (9.3)
Education			
High school or below high school	17 (9.7)	15 (12.4)	89 (12.3)
College or university	127 (72.6)	84 (69.4)	522 (72.3)
Postgraduate	31 (17.1)	22 (18.2)	111 (15.4)
Marital experience			
No	166 (94.9)	114 (94.2)	705 (97.6)
Yes	9 (5.1)	7 (5.8)	17 (2.4)
Monthly income (US $)			
≤713	42 (24.0)	20 (16.5)	152 (21.1)
714-1783	103 (58.9)	82 (76.8)	544 (63.0)
≥1784	30 (17.1)	19 (15.7)	115 (15.9)
Sexual orientation			
MSM[d]	142 (81.1)	104 (86.0)	598 (82.8)
Bisexual	33 (18.9)	17 (14.0)	124 (17.2)
Have come out			
No	26 (14.9)	21 (17.4)	118 (16.3)
Yes	149 (85.1)	100 (82.6)	604 (83.7)
Preferred sexual position			
Top	37 (21.1)	43 (28.1)	194 (26.9)
Bottom	53 (30.3)	38 (31.4)	196 (27.1)
Both bottom and top	85 (48.6)	49 (40.5)	332 (46.0)
HIV testing result			
Negative	166 (94.9)	118 (97.5)	716 (99.2)
Positive	9 (5.1)	3 (2.5)	6 (0.8)
CD4 count (n=18)			
≥350/μL	2 (22.2)	2 (66.7)	5 (83.3)
≤349/μL	7 (77.8)	1 (33.3)	1 (16.7)

[a] Class 1: highest risk.

[b] Class 2: highest protection.

[c] Class 3: low risk and low protection.

[d] MSM: men who have sex with men.

Table 5. Comparison of classes according to demographics and testing results (N=1018).

Variable	Class 1 versus Class 3 (highest risk vs low risk and low protection)		Class 2 versus Class 3 (highest protection vs low risk and low protection)	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value
Age group (years)				
20-29 (reference)	—[b]	—	—	—
30-39	1.01 (0.699-1.461)	.96	0.837 (0.549-1.277)	.41
≥40	2.197 (1.357-3.558)	.001	1.100 (0.576-2.099)	.77
Education				
High school or below high school (reference)	—	—	—	—
College or university	1.274 (0.732-2.216)	.39	0.955 (0.527-1.729)	.88
Postgraduate	1.462 (0.760-2.812)	.26	1.176 (0.576-2.399)	.66
Marital experience				
No (reference)	—	—	—	—
Yes	2.248 (0.985-5.133)	.06	2.546 (1.033-6.277)	.04
Monthly income (US $)				
≤713 (reference)	—	—	—	—
714-1783	0.819 (0.547-1.226)	.33	1.370 (0.813-2.309)	.24
≥1784	0.944 (0.557-1.600)	.83	1.256 (0.641-22.461)	.51
Sexual orientation				
MSM[c] (reference)	—	—	—	—
Bisexual	1.121 (0.733-1.715)	.60	0.788 (0.465-1.364)	.40
Have come out				
No (reference)	—	—	—	—
Yes	1.120 (0.706-1.775)	.63	0.930 (0.559-1.549)	.72
Preferred sexual position				
Top (reference)	—	—	—	—
Bottom	1.418 (0.891-2.256)	.14	1.106 (0.669-1.830)	.69
Both bottom and top	1.324 (0.878-2.053)	.17	0.842 (0.525-1.350)	.48
HIV testing result				
Negative	—	—	—	—
Positive	6.470 (2.272-18.428)	<.001	3.034 (0.749-12.297)	.12
CD4 count (n=18)				
≥350/μL	—	—	—	—
≤349/μL	17.5 (1.223-250.357)	.04	2.5 (0.100-62.605)	.58

[a] OR: odds ratio.

[b] Not applicable.

[c] MSM: men who have sex with men.

Discussion

Principal Results

LCA was used to derive three subgroups, based on best fit statistics and model interpretation, of MSM who underwent mobile VCT via social networking platforms. Two risk-taking indicators (having MSP and UAI) and three protective indicators (experience of PEP, experience of PrEP, and regular HIV testing) had probabilities >50% and were used to determine discrete classes in the final model. The profile of the highest risk group was class 1. The three classes were heterogeneous in age and marital history. These findings may help to identify subgroups suitable for targeted interventions and HIV testing.

MSP and UAI were two risk indicators that exceeded 50% in all three groups. Previous studies have confirmed that having experience of MSP and UAI can be one of the predictive factors for HIV-positivity rates of MSM [25,26]. The results of this study provide additional information; that is, the participants who were >40 years old were more likely to be grouped in class 1 and were at an increased risk of HIV and late diagnosis. Higher probabilities of MSP and UAI and age >40 years concurrently could be the first risk assessment items for setting priorities of those who need to be tested, especially when mobile VCT represents a competitive resource and needs to be applied precisely for those in the highest risk group. This result echoes a previous discussion that, compared to younger-aged MSM, middle-aged (40-50 years) adults are still at high risk for HIV infection [27]. The emergence of dating apps has enabled older MSM to connect with a different generation and to participate in condomless sexual activity [28]. Furthermore, feelings of loneliness and fatigue associated with HIV prevention may increase the risk of unprotected sexual encounters and the number of sexual partners among older MSM, thereby increasing the infection risk [29,30]. Additionally, the signs and symptoms of early HIV infection, including persistent influenza-like symptoms, herpes zoster, a single episode of bacterial pneumonia, thrombocytopenia, and lymphocytopenia, may be attributable to diseases of aging, and these people and their caregivers may neglect their risk for HIV and delay HIV testing; therefore, older adults are at increased risk of late diagnosis [31].

The other two risk indicators, drug use within the past 3 months and lifetime STDs, did not exceed a probability of 50% in the three subgroups. There are several possible reasons to explain this finding. Class 1 participants had a drug use probability of 23.2%, which was higher than those in the other classes and was lower than the probability of 77.3% from a previous report based on an online survey in a single Chinese city [32]. This probability was also lower than that previously reported (30%) in a pen-and-paper investigation, without any intervention, from 13 European cities [33]. This difference may be because the previous two studies investigated the experience of recreational drug use in a lifetime, which is longer than the past 3 months considered in this study. In addition, all participants in this study underwent counseling and testing; thus, the self-reported rate of drug use was likely an underestimate. That is, participants may be less likely to report drug use in face-to-face interactions than in online channels and paper questionnaires for fear of judgment and legal consequences. The probability of lifetime STDs (43.8%) in class 1 was higher than those in the other two classes, was comparable to that (43.3%) previously reported in a cross-sectional survey of American urban MSM undergoing a physical examination and STD testing [34], and was higher than that (13.0%) quoted in a study based on self-reports in a Chinese city [35]. Self-reporting of STDs may be limited by the feelings of shame and recall bias [36,37]. However, in this study, compared to the low-risk group, MSM in the highest risk group were more likely to disclose a history of STDs. These participants may believe that STDs are curable and have expected a medical referral for diagnosis and treatment through this consultation. These results suggest that in addition to MSP and UAI, lifetime STDs, especially a history of human papillomavirus, syphilis, and gonorrhea, and recreational drug use within the past 3 months could be considered as the minor risk-taking assessment items for arranging the priority for HIV testing.

The probabilities of the three protective indicators exceeded 50% in class 2 and were higher than those in the other two groups. It is noteworthy that the highest risk group, class 1, engaged in regular testing as the only protective measure, and the probability of taking PrEP and PEP was lower than that in class 2. Although this study did not investigate the reasons that hinder the use of PEP and PrEP in the highest risk group, previous studies have revealed that barriers include the lack of awareness of HIV risk behaviors, side effects and high cost of PrEP, the compliance difficulties and stigma of using PrEP [38,39], and low access to or tolerability of PEP [40]. Furthermore, age and income may have hindered the uptake of PrEP and PEP in this study. The proportion of participants aged ≥40 years in class 1 was 7.6% higher than that of class 2 and the proportion having a monthly income of ≤US $713 in class 1 was 7.5% higher than that of class 2. Older MSM may have a poorer economic status and were less likely to be aware, or have an inappropriate understanding of, PrEP than MSM ≤30 years old, which made them less likely to use it [41,42]. Improving HIV prevention may involve providing counseling on PEP and PrEP use during HIV testing for high-risk targets, including basic knowledge, side effect management, using reminders to ensure compliance, and providing financial and social support that increases the likelihood of preventive medication use [43].

Compared to those of class 3, class 2 participants were more likely to have been married. Studies have shown that approximately half of Chinese MSM intend to marry women and most of them expect to engage in extramarital homosexual relationships [44]. MSM who had recent UAI with unknown partners declared PrEP acquisition intentions [45]. Our findings still require further analysis to understand whether having experience of marriage could facilitate the adoption of HIV-prevention behaviors by MSM to protect their spouses.

Class 3 was the largest group and had lower risk and lower protective features than the other two classes. Over half of the participants in class 3 perceived themselves at risk of HIV infection and regularly undertook convenient, free, and anonymous HIV screening as a protective measure. Thus, HIV testing seems to be the best way to reach those who meet the characteristics of class 3. Accordingly, screening staff should continually reinforce this group of HIV risk-reducing behaviors and provide more protective options to promote the best preventive effect.

Limitations

This study had some limitations. First, convenient sampling was used and participants were recruited mainly in northern Taiwan, thereby reducing the generalizability of the findings and representativeness of the sample. Second, this research was conducted before the COVID-19 outbreak, and thus our results may not be inferenced and applied to the COVID-19 pandemic, since the pattern of risky and protective behaviors could have changed during that period. Third, self-reported data are likely subject to bias, which may affect the validity of these findings. Fourth, this study was cross-sectional, precluding discussions about causality. Fifth, this research did not consider the factors that may affect the acceptance of mobile VCT of MSM with regard to social and environmental aspects, such as avoidance of screening due to discriminatory attitudes of health care providers. In follow-up research, indicators of more dimensions could be added to promote a more comprehensive analysis.

Conclusions

MSM who underwent mobile VCT via social networking platforms exhibited different risk and protection profiles. LCA helped to derive a classification based on risk-taking and protective indicators. These results may inform policies for simplifying the prescreening assessment and precisely recognizing those who have higher risk-taking features but remain undiagnosed targets, including MSM engaging in MSP and UAI within the past 3 months and those ≥40 years old. Longitudinal studies could be conducted to evaluate the efficacy of this approach.

Acknowledgments

The authors would like to thank the Taiwan Centers for Disease Control (program number JK107053) for providing research funding. The authors thank all participants for undergoing mobile VCT and providing valuable information.

Conflicts of Interest

None declared.

References

1. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, HPTN 052 Study Team. Antiretroviral therapy for the prevention of HIV-1 transmission. N Engl J Med 2016 Sep 01;375(9):830-839 [FREE Full text] [doi: 10.1056/NEJMoa1600693] [Medline: 27424812]
2. Resch SC, Foote JHA, Wirth KE, Lasry A, Scott JA, Moore J, et al. Health impact and cost-effectiveness of HIV testing, linkage, and early antiretroviral treatment in the Botswana Combination Prevention Project. J Acquir Immune Defic Syndr 2022 Aug 01;90(4):399-407 [FREE Full text] [doi: 10.1097/QAI.0000000000002996] [Medline: 35420554]
3. Molina JM, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, ANRS IPERGAY Study Group. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. N Engl J Med 2015 Dec 03;373(23):2237-2246. [doi: 10.1056/NEJMoa1506273] [Medline: 26624850]
4. UNAIDS data 2021. The Joint United Nations Programme on HIV/AIDS. 2021. URL: https://www.unaids.org/sites/default/files/media_asset/JC3032_AIDS_Data_book_2021_En.pdf [accessed 2022-03-09]
5. Defeat AIDS and share love. Taiwan Center for Disease Control. 2021. URL: https://www.cdc.gov.tw/Bulletin/Detail/VVem5ruoxZhiPBFveSPiLQ?typeid=9 [accessed 2022-03-06]
6. Fast-track commitments to end AIDS by 2030. The Joint United Nations Programme on HIV/AIDS. 2016. URL: https://www.unaids.org/en/resources/documents/2016/fast-track-commitments [accessed 2022-03-06]
7. Chiou PY, Ko NY, Chien CY. Mobile HIV testing through social networking platforms: comparative study. J Med Internet Res 2021 Mar 26;23(3):e25031 [FREE Full text] [doi: 10.2196/25031] [Medline: 33769298]
8. Nanthaprut P, Manojai N, Chanlearn P, Mattawanon N, Chiawkhun P, Homkham N, et al. Comparison of HIV-positive incidence among transgender women and men who have sex with men at stand-alone and mobile voluntary counseling and testing Facilities in Chiang Mai Province, Thailand. AIDS Patient Care STDS 2021 Apr;35(4):116-125. [doi: 10.1089/apc.2020.0258] [Medline: 33835852]
9. Chamie G, Clark TD, Kabami J, Kadede K, Ssemmondo E, Steinfeld R, et al. A hybrid mobile approach for population-wide HIV testing in rural east Africa: an observational study. Lancet HIV 2016 Mar;3(3):e111-e119 [FREE Full text] [doi: 10.1016/S2352-3018(15)00251-9] [Medline: 26939734]
10. HIV/AIDS. Taiwan Center for Disease Control. 2022. URL: https://www.cdc.gov.tw/En/Category/ListContent/bg0g_VU_Ysrgkes_KRUDgQ?uaid=CqNo313w78G1fWhz429xDA [accessed 2022-03-09]
11. Punyacharoensin N, Edmunds WJ, De Angelis D, Delpech V, Hart G, Elford J, et al. Effect of pre-exposure prophylaxis and combination HIV prevention for men who have sex with men in the UK: a mathematical modelling study. Lancet HIV 2016 Feb;3(2):e94-e104. [doi: 10.1016/S2352-3018(15)00056-9] [Medline: 26847231]
12. Morgan E, Skaathun B, Lancki N, Jimenez AD, Ramirez-Valles J, Bhatia R, et al. Trends in HIV risk, testing, and treatment among MSM in Chicago 2004-2014: implications for HIV elimination planning. J Urban Health 2017 Oct;94(5):699-709 [FREE Full text] [doi: 10.1007/s11524-017-0175-9] [Medline: 28631059]
13. Ossa-Giraldo AC, Correa JS, Moreno CL, Blanquiceth Y, Flórez-Álvarez L, Contreras-Ramírez K, et al. Sexual Behaviors and Factors Associated with Condomless Sexual Practice in Colombian Men Who Have Sex with Men at High Risk of HIV Transmission. Arch Sex Behav 2021 Oct;50(7):3175-3190. [doi: 10.1007/s10508-020-01856-y] [Medline: 33829332]

14. Gras J, Pillet M, Antoni G, Cua E, Charreau I, Raffi F, IPERGAY Study group. Risk factors for HIV infection among men who have sex with men in the ANRS IPERGAY PrEP trial. Sex Transm Infect 2022 Aug;98(5):383-386. [doi: 10.1136/sextrans-2021-055199] [Medline: 34663696]
15. Magidson J, Vermunt JK. Latent class models. In: Kaplan D, editor. The SAGE handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: SAGE Publications; 2004:175-198.
16. Rabe-Hesketh S, Skrondal A. Classical latent variable models for medical research. Stat Methods Med Res 2008 Feb;17(1):5-32. [doi: 10.1177/0962280207081236] [Medline: 17855748]
17. Weller BE, Bowen NK, Faubert SJ. Latent Class Analysis: A Guide to Best Practice. J Black Psychol 2020 Jun 05;46(4):287-311. [doi: 10.1177/0095798420930932]
18. Dangerfield DT2, Harawa NT, Smith LR, Jeffries WL4, Baezconde-Garbanati L, Bluthenthal R. Latent Classes of Sexual Risk Among Black Men Who Have Sex with Men and Women. Arch Sex Behav 2018 Oct;47(7):2071-2080 [FREE Full text] [doi: 10.1007/s10508-017-1142-y] [Medline: 29541914]
19. Smith MK, Wei C, Liu C, Pan SW, Ong JJ, Tucker JD. Gender identity and sexual orientation in Chinese men who have sex with men: a latent class analysis. Arch Sex Behav 2020 Feb;49(2):721-731 [FREE Full text] [doi: 10.1007/s10508-019-01481-4] [Medline: 31571020]
20. Card KG, Lachowsky NJ, Cui Z, Carter A, Armstrong H, Shurgold S, et al. A latent class analysis of seroadaptation among gay and bisexual men. Arch Sex Behav 2018 Jan;47(1):95-106 [FREE Full text] [doi: 10.1007/s10508-016-0879-z] [Medline: 27987086]
21. Doyle CM, Maheu-Giroux M, Lambert G, Mishra S, Apelian H, Messier-Peet M, Engage Study Team. Combination HIV prevention strategies among Montreal gay, bisexual, and other men who have sex with men in the PrEP era: a latent class analysis. AIDS Behav 2021 Jan;25(1):269-283 [FREE Full text] [doi: 10.1007/s10461-020-02965-4] [Medline: 32648063]
22. "Abbott" DETERMINE HIV-1/2. Food and Drug Administration, Ministry of Health and Welfare. URL: https://info.fda.gov.tw/MLMS/H0001D3.aspx?LicId=06010855 [accessed 2022-06-30]
23. Selig K, Shaw P, Ankerst D. Bayesian information criterion approximations to Bayes factors for univariate and multivariate logistic regression models. Int J Biostat 2020 Oct 29;17(2):241-266. [doi: 10.1515/ijb-2020-0045] [Medline: 33119543]
24. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007 Dec 05;14(4):535-569. [doi: 10.1080/10705510701575396]
25. Dong Y, Liu S, Xia D, Xu C, Yu X, Chen H, et al. Prediction model for the risk of HIV infection among MSM in China: validation and stability. Int J Environ Res Public Health 2022 Jan 17;19(2):1010 [FREE Full text] [doi: 10.3390/ijerph19021010] [Medline: 35055826]
26. Burt RD, Thiede H. Variations in patterns of sexual risk behavior among Seattle-area MSM based on their HIV status, the HIV status of their partner and partner type. AIDS Behav 2012 Apr;16(3):599-607 [FREE Full text] [doi: 10.1007/s10461-011-9979-x] [Medline: 21691761]
27. Houle B, Mojola SA, Angotti N, Schatz E, Gómez-Olivé FX, Clark SJ, et al. Sexual behavior and HIV risk across the life course in rural South Africa: trends and comparisons. AIDS Care 2018 Nov;30(11):1435-1443 [FREE Full text] [doi: 10.1080/09540121.2018.1468008] [Medline: 29701073]
28. Queiroz AAFLN, Sousa AFL, Brignol S, Araújo TME, Reis RK. Vulnerability to HIV among older men who have sex with men users of dating apps in Brazil. Braz J Infect Dis 2019 Sep;23(5):298-306 [FREE Full text] [doi: 10.1016/j.bjid.2019.07.005] [Medline: 31472125]
29. Valdes B, Gattamorta K, Jones SG, Fenkl EA, De Santis JP. Social support, loneliness, depressive symptoms, and high-risk sexual behaviors of middle-aged Hispanic men who have sex with men. J Assoc Nurses AIDS Care 2019 Jan;30(1):98-110. [doi: 10.1097/JNC.0000000000000002] [Medline: 30586087]
30. Dangerfield DT2, Carmack CC, Gilreath TD, Duncan DT. Latent classes of sexual positioning practices and sexual risk among men who have sex with men in Paris, France. AIDS Behav 2018 Dec;22(12):4001-4008. [doi: 10.1007/s10461-018-2267-2] [Medline: 30173344]
31. Justice AC, Goetz MB, Stewart CN, Hogan BC, Humes E, Luz PM, et al. Delayed presentation of HIV among older individuals: a growing problem. Lancet HIV 2022 Apr;9(4):e269-e280. [doi: 10.1016/S2352-3018(22)00003-0] [Medline: 35218732]
32. Zhao P, Tang S, Wang C, Zhang Y, Best J, Tangthanasup TM, et al. Recreational drug use among Chinese MSM and transgender individuals: results from a national online cross-sectional study. PLoS One 2017 Jan 20;12(1):e0170024 [FREE Full text] [doi: 10.1371/journal.pone.0170024] [Medline: 28107391]
33. Rosińska M, Gios L, Nöstlinger C, Vanden Berghe W, Marcus U, Schink S, Sialon II Network. Prevalence of drug use during sex amongst MSM in Europe: results from a multi-site bio-behavioural survey. Int J Drug Policy 2018 May;55:231-241. [doi: 10.1016/j.drugpo.2018.01.002] [Medline: 29402683]
34. Solomon MM, Nureña CR, Tanur JM, Montoya O, Grant RM, McConnell JJ. Transactional sex and prevalence of STIs: a cross-sectional study of MSM and transwomen screened for an HIV prevention trial. Int J STD AIDS 2015 Oct;26(12):879-886 [FREE Full text] [doi: 10.1177/0956462414562049] [Medline: 25520017]

35. Song Y, Li X, Zhang L, Liu Y, Jiang S, Stanton B. Demographic and behavioral determinants of self-reported history of sexually-transmitted diseases (STDs) among young migrant men who have sex with men (MSM) in Beijing, China. Health Educ J 2012 Nov 01;71(6):699-708 [FREE Full text] [doi: 10.1177/0017896911422776] [Medline: 34421123]
36. Harte D, Mercey D, Jarman J, Benn P. Is the recall of men who have sex with men (MSM) diagnosed as having bacterial sexually transmitted infections (STIs) for re-screening a feasible and effective strategy? Sex Transm Infect 2011 Dec;87(7):577-582. [doi: 10.1136/sextrans-2011-050144] [Medline: 21965470]
37. Hood JE, Friedman AL. Unveiling the hidden epidemic: a review of stigma associated with sexually transmissible infections. Sex Health 2011 Jun;8(2):159-170. [doi: 10.1071/SH10070] [Medline: 21592429]
38. Franks J, Hirsch-Moverman Y, Loquere ASJ, Amico KR, Grant RM, Dye BJ, et al. Sex, PrEP, and stigma: experiences with HIV pre-exposure prophylaxis among New York City MSM participating in the HPTN 067/ADAPT Study. AIDS Behav 2018 Apr;22(4):1139-1149 [FREE Full text] [doi: 10.1007/s10461-017-1964-6] [Medline: 29143163]
39. Matacotta JJ, Rosales-Perez FJ, Carrillo CM. HIV preexposure prophylaxis and treatment as prevention - Beliefs and access barriers in men who have sex with men (MSM) and transgender women: a systematic review. J Patient Cent Res Rev 2020 Jul 27;7(3):265-274 [FREE Full text] [doi: 10.17294/2330-0698.1737] [Medline: 32760758]
40. Palich R, Martin-Blondel G, Cuzin L, Le Talec JY, Boyer P, Massip P, et al. Experiences of HIV postexposure prophylaxis (PEP) among highly exposed men who have sex with men (MSM). Sex Transm Infect 2017 Nov;93(7):493-498. [doi: 10.1136/sextrans-2016-052901] [Medline: 28739808]
41. Buffel V, Reyniers T, Masquillier C, Thunissen E, Nöstlinger C, Laga M, et al. Awareness of, willingness to take PrEP and its actual use among Belgian MSM at high risk of HIV infection: secondary analysis of the Belgian European MSM Internet Survey. AIDS Behav 2022 Jun;26(6):1793-1807. [doi: 10.1007/s10461-021-03526-z] [Medline: 34817766]
42. Hulstein SH, Zimmermann HML, de la Court F, Matser AA, Schim van der Loeff MF, Hoornenborg E, et al. Factors associated with the intention to use HIV preexposure prophylaxis for young and older men who have sex with men. Sex Transm Dis 2022 May 01;49(5):343-352 [FREE Full text] [doi: 10.1097/OLQ.0000000000001599] [Medline: 35001016]
43. Golub SA, Gamarel KE, Rendina HJ, Surace A, Lelutiu-Weinberger CL. From efficacy to effectiveness: facilitators and barriers to PrEP acceptability and motivations for adherence among MSM and transgender women in New York City. AIDS Patient Care STDS 2013 Apr;27(4):248-254 [FREE Full text] [doi: 10.1089/apc.2012.0419] [Medline: 23565928]
44. Wu W, Yan X, Zhang X, Goldsamt L, Chi Y, Huang D, et al. Potential HIV transmission risk among spouses: marriage intention and expected extramarital male-to-male sex among single men who have sex with men in Hunan, China. Sex Transm Infect 2020 Mar;96(2):151-156 [FREE Full text] [doi: 10.1136/sextrans-2018-053906] [Medline: 31171593]
45. Gamarel KE, Golub SA. Intimacy motivations and pre-exposure prophylaxis (PrEP) adoption intentions among HIV-negative men who have sex with men (MSM) in romantic relationships. Ann Behav Med 2015 Apr;49(2):177-186 [FREE Full text] [doi: 10.1007/s12160-014-9646-3] [Medline: 25124457]

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion
G^2: likelihood ratio test statistic
LCA: latent class analysis
LL: log likelihood
MSM: men who have sex with men
MSP: multiple sexual partners
OR: odds ratio
PEP: postexposure prophylaxis
PLWH: people living with HIV
PrEP: pre-exposure prophylaxis
STD: sexually transmitted disease
UAI: unprotected anal intercourse
VCT: voluntary counseling and testing

Edited by Y Khader; submitted 23.10.22; peer-reviewed by PH Liao, C Copen; comments to author 18.11.22; revised version received 06.01.23; accepted 11.01.23; published 16.02.23.

Please cite as:
Chiou PY, Tsao WW, Lin KC, Fang YY, Lin KY, Li CL
Risk and Protective Profile of Men Who Have Sex With Men Using Mobile Voluntary HIV Counseling and Testing: Latent Class Analysis
JMIR Public Health Surveill 2023;9:e43394
URL: https://publichealth.jmir.org/2023/1/e43394
doi:10.2196/43394
PMID:36795477

Original Paper

HIV Epidemiology, Care, and Treatment Outcomes Among Student and Nonstudent Youths Living With HIV in Southwest China Between 1996 and 2019: Historical Cohort Study

Na Wang[1,2*], PhD; Guanghua Lan[1*], PhD; Qiuying Zhu[1], MS; Huanhuan Chen[1], MS; Jinghua Huang[1], BA; Qin Meng[1], MS; Zhiyong Shen[1], MS; Shujia Liang[1], MS; Xiuling Wu[1], AD; Liuhong Luo[1], MS; Rongyi Ye[2], MPH; Jinli Chen[2], MPH; Shengkui Tan[2], PhD; Hui Xing[3], MS; Yiming Shao[3], PhD; Yuhua Ruan[3], PhD; Mei Lin[1], PhD

[1]Guangxi Key Laboratory of Major Infectious Disease Prevention Control and Biosafety Emergency Response, Guangxi Center for Disease Control and Prevention, Nanning, China

[2]Guangxi Key Laboratory of Environmental Exposomics and Entire Lifecycle Heath, School of Public Health, Guilin Medical University, Guilin, China

[3]State Key Laboratory of Infectious Disease Prevention and Control, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Mei Lin, PhD
Guangxi Key Laboratory of Major Infectious Disease Prevention Control and Biosafety Emergency Response
Guangxi Center for Disease Control and Prevention
18 Jinzhou Road
Nanning, 530027
China
Phone: 86 07712518689
Email: gxlinmei@126.com

Abstract

Background: Nearly one-third of new HIV infections occurred among youth in 2019 worldwide. Previous studies suggested that student youths living with HIV and nonstudent youths living with HIV might differ in some risk factors, transmission routes, HIV care, and disease outcomes.

Objective: This study aimed to compare the HIV epidemic, disease outcomes, and access to care among student and nonstudent youths living with HIV aged 16 to 25 years in Guangxi, China.

Methods: We performed a historical cohort study by extracting data on all HIV or AIDS cases aged 16 to 25 years in Guangxi, China, during 1996-2019 from the Chinese Comprehensive Response Information Management System of HIV or AIDS. We conducted analyses to assess possible differences in demographic and behavioral characteristics, HIV care, and disease outcomes between student and nonstudent youths living with HIV. Multivariate Cox regression was used to assess differences in mortality and virologic failure between student and nonstudent cases.

Results: A total of 13,839 youths aged 16 to 25 years were infected with HIV during 1996-2019. Among them, 10,202 cases were infected through sexual contact, most of whom were men (n=5507, 54%); 868 (8.5%) were students, and 9334 (91.5%) were not students. The number of student youths living with HIV was lower before 2006 but gradually increased from 2007 to 2019. In contrast, the nonstudent cases increased rapidly in 2005, then gradually declined after 2012. Student cases were mainly infected through homosexual contact (n=614, 70.7% vs n=1447, 15.5%; *P*<.001), while nonstudent cases were more likely to be infected through heterosexual contact (n=7887, 84.5% vs n=254, 29.3%; *P*<.001). Moreover, nonstudent cases had a significantly lower CD4 count than student cases at the time of HIV diagnosis (332 vs 362 cells/µL; *P*<.001). Nonstudents also had a delayed antiretroviral therapy (ART) initiation compared to students (93 days vs 22 days; *P*<.001). Furthermore, the mortality rate of 0.4 and 1.0 deaths per 100 person-years were recorded for student and nonstudent youths with HIV, respectively. Overall, the mortality risk in nonstudent cases was 2.3 times that of student cases (adjusted hazard ratio [AHR] 2.3, 95% CI 1.2-4.2; *P*=.008). The virologic failure rate was 2.3 and 2.6 per 100 person-years among student and nonstudent youths living with HIV, respectively. Nonstudent cases had double the risk of virologic failure compared to student cases (AHR 1.9, 95% CI 1.3-2.6; *P*<.001).

Conclusions: Nonstudent youths living with HIV might face a low CD4 count at the time of HIV diagnosis, delayed ART initiation, and increased risk of death and virologic failure. Thus, HIV prevention and interventions should target youths who dropped out of school early to encourage safe sex and HIV screening, remove barriers to HIV care, and promote early ART initiation to curb the HIV epidemic among youths.

(JMIR Public Health Surveill 2023;9:e38881) doi:10.2196/38881

KEYWORDS

HIV; student; nonstudent; antiretroviral therapy; mortality; China

Introduction

According to the Joint United Nations Programme on HIV/AIDS, 31% of new HIV infections globally occurred among youths aged 15 to 24 years in 2019 [1]. Sexual transmission, including heterosexual and homosexual transmission, was the primary route of HIV/AIDS transmission among young people [1-4]. In China, the incidence rate of HIV/AIDS increased fivefold from 0.27 cases per 100,000 in 2008 to 1.49 cases per 100,000 in 2017. HIV/AIDS is a major cause of death from infectious diseases in adolescents [5]. The number of youths living with HIV showed this increase in some cities [6]. Furthermore, nonstudents aged 15 to 24 years accounted for 81% of HIV cases compared to 19% of students in 2017 [7]. Students and nonstudents might differ in HIV risk and risk factors [8].

Antiretroviral therapy (ART), introduced globally since 1996, improves immune function by suppressing viral replication and dramatically declines the morbidity and mortality of people living with HIV [9]. ART early initiation—initiating ART as early as the day of HIV diagnosis—has numerous benefits, including reducing the HIV epidemic, improving the health status and quality of life, as well as extending the life expectancy of people living with HIV [10,11]. However, adolescents living with HIV in southern Africa are likely to have less access to HIV care and a hard time adhering to ART compared to adults, which in turn has resulted in negative consequences, such as poorer viral suppression and HIV-related deaths [12-14]. Student youths living with HIV may face more barriers to ART, such as stigma, fear of unintended disclosure of HIV status (due to the information on drug packaging and lack of privacy while taking pills), challenges in drug storage in school, inability to coordinate their studies and clinic-related activities, and lack of structured supporting systems in schools. These barriers may lead to sadness, anger, frustration, and stress among student youths living with HIV [15]. At the same time, the educational level, some demographic and behavioral characteristics, and low perceived risk for HIV are associated with late ART initiation in nonstudent youths living with HIV [6]. Thus, student and nonstudent youths with HIV aged between 16 and 25 years might also show differences in the ART initiation and treatment outcomes.

Currently, there is limited information on the demographic and behavioral characteristics, HIV infection routes, access to HIV care, long-term ART outcome, and mortality rate among student and nonstudent cases in a large population. Therefore, we conducted a historical cohort study from 1990 to 2019 to investigate the HIV epidemic, access to care, and HIV/AIDS disease outcomes among youths aged 16 to 25 years in Guangxi, China, to better understand the disease characteristics among student and nonstudent youths with HIV. The findings of this study will inform the design of AIDS programs, considering the differences in risk factors, transmission routes, the ART initiation rate, and treatment outcomes between students and nonstudents. In addition, the findings will guide more tailored HIV prevention and care strategies to curb the HIV epidemic among adolescents further.

Methods

Study Sites and Population

The study was conducted in Guangxi Zhuang Autonomous Region, southwest China. Guangxi was among the top 3 provinces and autonomous regions with the most HIV/AIDS cases in China from 2004 to 2007. Besides, it had the second highest number of HIV/AIDS cases between 2011 and 2013 [16]. Most HIV cases were infected through sexual contact [17], and youths were very vulnerable to HIV. Therefore, we designed this study to investigate HIV-positive youths aged 16 to 25 years (age at diagnosis) and living in Guangxi (at the time of data exporting, January 2021) between January 1, 1996, and December 31, 2020.

Data Sources

We used data extracted from the Chinese Comprehensive Response Information Management System of HIV/AIDS (CRIMS), which includes the National Case Reporting Database (NCRD) and the National Free Antiretroviral Therapy Database (NFATD). The NCRD contains data on all HIV/AIDS cases from local hospitals, Centers for Disease Control and Prevention (CDC) clinics, and blood banks, including demographic information, the HIV diagnosis date, reporting organization, as well as the first CD4 test reporting date and result. The first CD4 test was conducted after an HIV diagnosis. The NFATD contains information on the ART, such as viral load and time, ART regimens, interregional transfer, comorbid diseases, loss to follow-up, death, and time of death [18]. In our study, time for the initiation of the ART was used as the baseline time point, and then patients were invited to visit the clinic for follow-up at 15 days, 1 and 3 months, and every 3 months; their information about the ART was then updated in the NFATD accordingly. Patients were also recommended to test the viral load and CD4 count once every 12 months. Individual cases in the NFATD were matched to the NCRD database using the case's card numbers [18]. Anonymous and deidentification data were extracted from the NCRD and NFATD by designated staff at CDC offices in Guangxi.

Variables collected in this study are about the patient status at the time of HIV diagnosis and follow-up information, including whether they are a student or nonstudent, age, gender, ethnicity, marital status, education, time and sites of HIV diagnosis, infection route, CD4 count, the ART initiation, HIV viral load, and death. We collected the HIV-positive cases reported from January 1, 1996, to December 31, 2019. However, the viral load and death data were collected until December 31, 2020. This study included all newly diagnosed HIV-positive youths aged 16 to 25 years (age at diagnosis) each year living in Guangxi.

Ethics Approval

This study was approved by the institutional review board of Guangxi (GXIRB2016-0047-3).

Data Analyses

We compared the demographic and behavioral characteristics, HIV care, virological loads, and deaths between students and nonstudents living with HIV. Chi-square tests were used for categorical variables (eg, gender, ethnicity, occupation, marital status, education, infection route, calendar year of HIV diagnosis, health facilities providing HIV diagnosis, and the ART initiation), and *t* tests or nonparametric tests were used for continuous variables (eg, CD4 count at HIV diagnosis and time from HIV diagnosis until the ART initiation). Multivariable Cox regression was conducted to assess if students or nonstudents' status were associated with all-cause death and virologic failure, adjusting for sex, ethnicity, marital status, education, infection route, CD4 count before the ART initiation, and the World Health Organization (WHO) clinical stage before the ART.

Results

Demographic and Behavioral Characteristics Among Youths Living With HIV

A total of 13,839 youths aged 16 to 25 years were diagnosed with HIV or AIDS in Guangxi between 1996 and 2019. Among them, 883 (6.4%) were students, and 12,956 (93.6%) were nonstudents. The number of students living with HIV was low before 2006 but gradually increased from 2007 to 2019. In contrast, the nonstudent cases living with HV increased rapidly in 2005 and gradually declined after 2012. The number of student youths was one-third of the nonstudents youths living with HIV in 2019 (Figure 1). The student cases were predominantly male; the majority of nonstudent youths living with HIV were male except during 2008-2011 (Figure 2). The homosexual transmission was the main infection route among student youths living with HIV in Guangxi. However, among the nonstudent cases, injecting drugs was the main route of infection before 2005, followed by heterosexual transmission during 2006-2019. At the same time, homosexual transmission showed an increasing trend in recent years (Figure 3). The mean CD4 count at the time of diagnosis was slightly higher in student youths living with HIV and showed an increasing trend during 1996-2019 in both student and nonstudent youths living with HIV (Figure 4).

Of the 13,839 youths, 61.8% (n=8551) were male and 38.2% (n=5288) were female. The HIV-positive youths were mostly of Han ethnicity (n=8647, 62.5%), nonfarmer (n=8687, 62.8%), single (n=8768, 63.4%), and with middle school education level or below (n=9211, 66.6%). Homosexual contact was the primary route of transmission for students (n=614, 69.5%), while heterosexual contact was the main transmission route for nonstudents (n=7887, 60.9%; *P*<.001; Table 1).

Figure 1. Annual cases of student and nonstudent youths living with HIV in Guangxi, China, from 1996 to 2019.

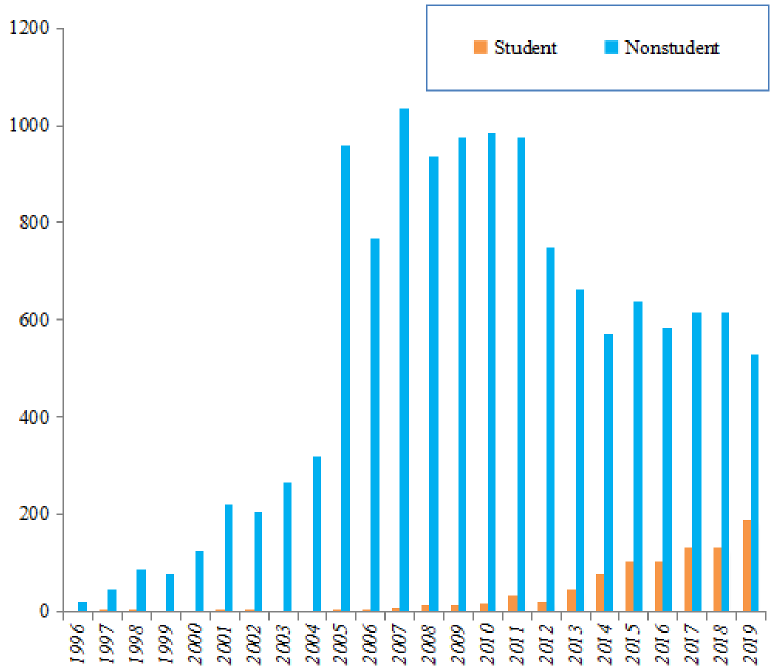


Figure 2. The gender of student (A) and nonstudent (B) youths living with HIV in Guangxi, China from 1996 to 2019.

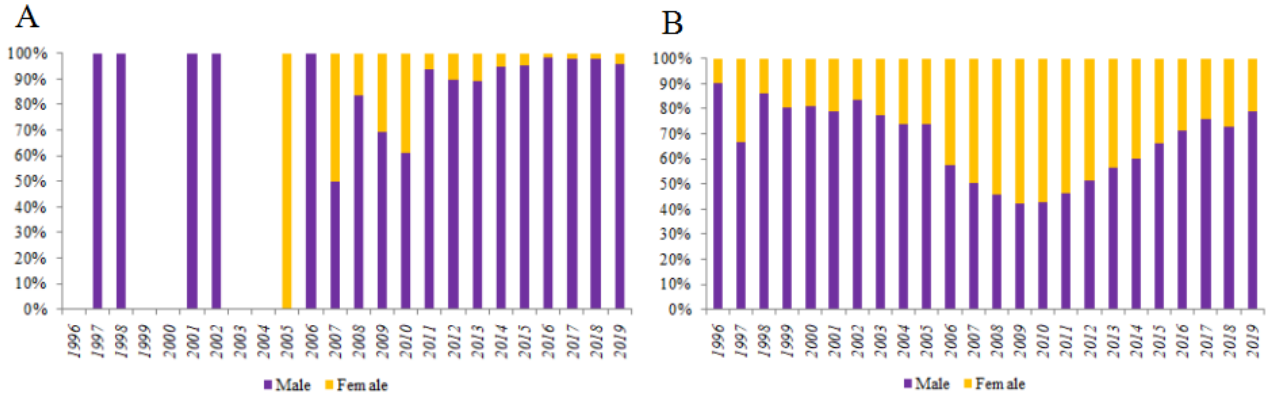


Figure 3. The infection routs of student(A) and nonstudent(B) youths living with HIV in Guangxi, China during 1996-2019.

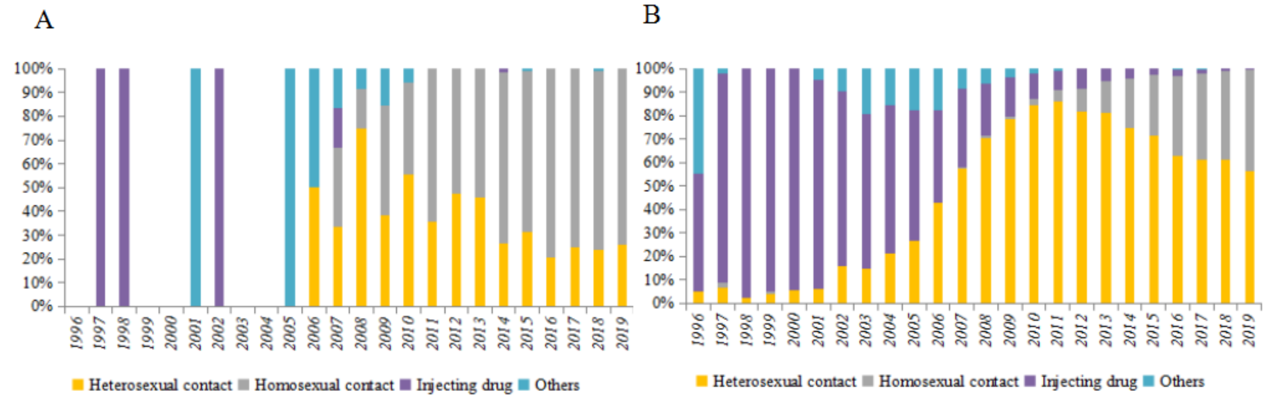


Figure 4. The mean CD4 count at diagnosis of student and nonstudent youths living with HIV in Guangxi, China from 1996 to 2019.

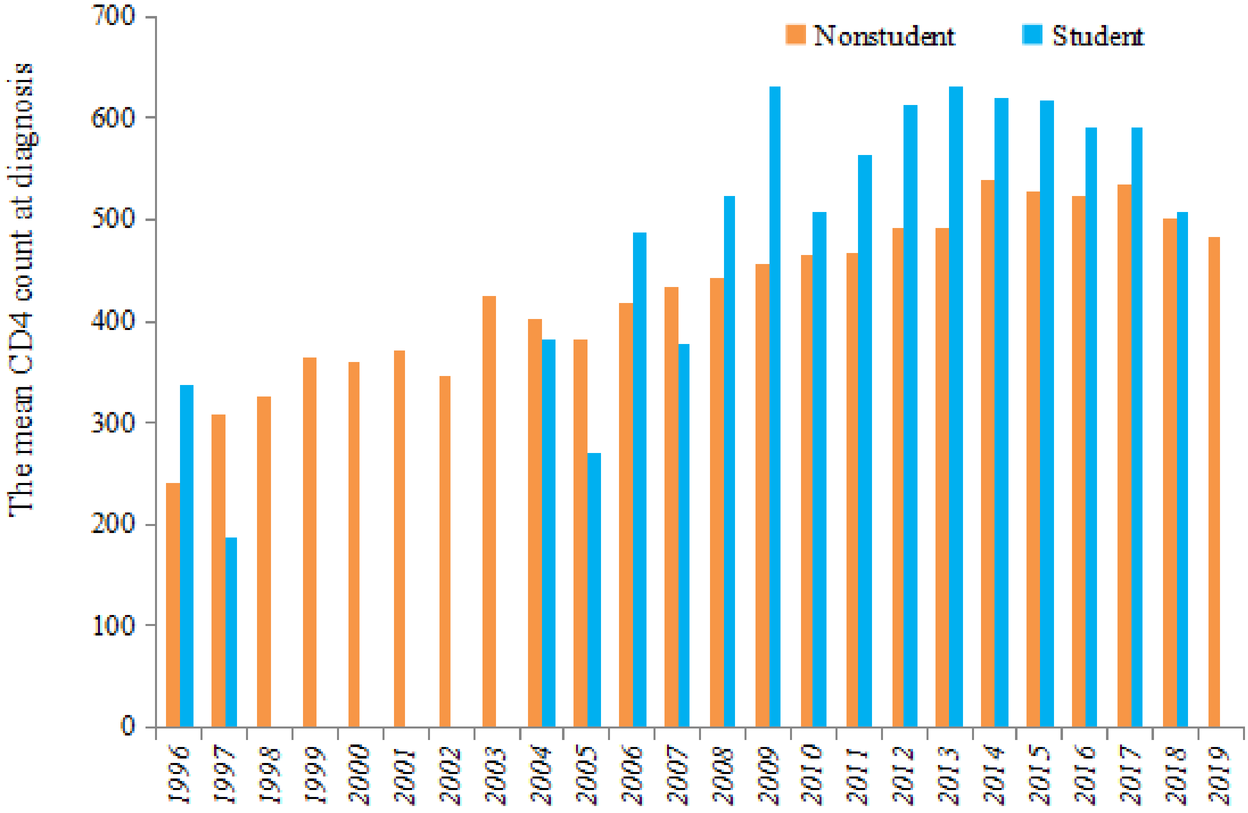


Table 1. Demographic and behavioral characteristics of youths living with HIV aged 16-25 years by student status in Guangxi, China, from 1996 to 2019.

Characteristics	Total, n (%)	Students, n (%)	Nonstudents, n (%)	*P* value
Total	13,839 (100)	883 (6.4)	12,956 (93.6)	
Gender				<.001
Male	8551 (61.8)	832 (94.2)	7719 (59.6)	
Female	5288 (38.2)	51 (5.8)	5237 (40.4)	
Ethnicity				<.001
Han	8647 (62.5)	601 (68.1)	8046 (62.1)	
Other	5192 (37.5)	282 (31.9)	4910 (37.9)	
Occupation				<.001
Farmer	5152 (37.2)	0 (0)	5152 (39.8)	
Other	8687 (62.8)	883 (100)	7804 (60.2)	
Marital status				<.001
Single	8768 (63.4)	875 (99.1)	7893 (60.9)	
Currently married	3962 (28.6)	5 (0.6)	3957 (30.5)	
Divorced, separated, or widowed	526 (3.8)	0 (0)	526 (4.1)	
Unknown	583 (4.2)	3 (0.3)	580 (4.5)	
Education				<.001
Middle school and below	9211 (66.6)	22 (2.5)	9189 (70.9)	
High school	1954 (14.1)	195 (22.1)	1759 (13.6)	
College and above	1739 (12.6)	661 (74.9)	1078 (8.3)	
Unknown	935 (6.8)	5 (0.6)	930 (7.2)	
Infection route				<.001
Heterosexual contact	8141 (58.8)	254 (28.8)	7887 (60.9)	
Homosexual contact	2061 (14.9)	614 (69.5)	1447 (11.2)	
Injecting drug	2960 (21.4)	5 (0.6)	2955 (22.8)	
Other	677 (4.9)	10 (1.1)	667 (5.1)	
Calendar year of HIV diagnosis				<.001
1996-2003 (pre-ART[a])	958 (6.9)	4 (0.5)	954 (7.4)	
2004-2007 (no standardized ART)	3050 (22)	9 (1)	3041 (23.5)	
2008-2012 (standardized ART)	4787 (34.6)	93 (10.5)	4694 (36.2)	
2013-2019 (TDF[b] added)	5044 (36.4)	777 (88)	4267 (32.9)	

[a] ART: antiretroviral therapy.

[b] TDF: tenofovir disoproxil fumarate.

Demographic and Behavioral Characteristics Among Youths Living With HIV Through Sexual Contact

A total of 10,202 youths aged 16-25 years were infected with HIV through sexual contact between 1996 and 2019 in Guangxi. Among them, 8.5% (868/10,202) were students, and 91.5% (9334/10,202) were nonstudents. Compared with nonstudent cases, student cases were more likely to be male (n=820, 94.5% vs n=4687, 50.2%; *P*<.001), nonfarmer (n=868, 100% vs n=5396, 57.8%; *P*<.001), single (n=862, 99.3% vs n=5555, 59.5%; *P*<.001), with an educational level of college or above (n=656, 75.6% vs n=1066, 11.4%; *P*<.001). Furthermore, student youths living with HIV were more likely to be infected through homosexual contact (n=614, 70.7% vs n=1447, 15.5%; *P*<.001) and diagnosed with HIV between 2013 and 2019 (n=774, 89.2% vs n=4134, 44.3%; *P*<.001; Table 2).

Table 2. Demographic and behavioral characteristics of youths aged 16-25 years living with HIV through sexual contact by student status in Guangxi, China, from 1996 to 2019.

Characteristics	Total, n (%)	Students, n (%)	Nonstudents, n (%)	*P* value
Total	10,202 (100)	868 (8.5)	9334 (91.5)	
Gender				<.001
Male	5507 (54)	820 (94.5)	4687 (50.2)	
Female	4695 (46)	48 (5.5)	4647 (49.8)	
Ethnicity				<.001
Han	6319 (61.9)	594 (68.4)	5725 (61.3)	
Other	3883 (38.1)	274 (31.6)	3609 (38.7)	
Occupation				<.001
Farmer	3938 (38.6)	0 (0)	3938 (42.2)	
Other	6264 (61.4)	868 (100.0)	5396 (57.8)	
Marital status				<.001
Single	6417 (62.9)	862 (99.3)	5555 (59.5)	
Currently married	3320 (32.5)	5 (0.6)	3315 (35.5)	
Divorced, separated, or widowed	404 (4)	0 (0)	404 (4.3)	
Unknown	61 (0.6)	1 (0.1)	60 (0.6)	
Education				<.001
Middle school and below	6577 (64.5)	19 (2.2)	6558 (70.3)	
High school	1834 (18)	193 (22.2)	1641 (17.6)	
College and above	1722 (16.9)	656 (75.6)	1066 (11.4)	
Unknown	69 (0.7)	0 (0)	69 (0.7)	
Infection route				<.001
Heterosexual contact	8141 (79.8)	254 (29.3)	7887 (84.5)	
Homosexual contact	2061 (20.2)	614 (70.7)	1447 (15.5)	
Calendar year of HIV diagnosis				<.001
1996-2003 (pre-ART[a])	92 (0.9)	0 (0)	92 (1)	
2004-2007 (no standardized ART)	1227 (12)	5 (0.6)	1222 (13.1)	
2008-2012 (standardized ART)	3975 (39)	89 (10.3)	3886 (41.6)	
2013-2019 (TDF[b] added)	4908 (48.1)	774 (89.2)	4134 (44.3)	

[a] ART: antiretroviral therapy.

[b] TDF: tenofovir disoproxil fumarate.

HIV Care Between Student and Nonstudent Youths Living With HIV

Among the HIV/AIDS cases infected through sexual contact reported from 2008 to 2019 in Guangxi, student youths living with HIV were more likely to receive an HIV diagnosis in CDC compared to nonstudent cases (n=458, 59.6% vs n=3099, 44.1%; *P*<.001). In addition, student cases had a higher average CD4 count (median 362 [IQR 262-481] cells/µL) compared to nonstudents (median 332 [IQR 204-459] cells/µL; *P*<.001) at diagnosis. Although most youths were on the ART in both groups, student youths with HIV who were on the ART were significantly higher in number compared to nonstudent youths (n=740, 96.2% vs n=6247, 88.8%; *P*<.001). The median time from HIV diagnosis until the ART initiation was 22 days for student cases and 93 days for nonstudent cases (*P*<.001; Table 3).

Table 3. HIV care of youths living with HIV who were infected through sexual contact (aged 16-25 years) by student status in Guangxi, China, from 2008 to 2019.

HIV care	Student (n=769)	Nonstudent (n=7034)	*P* value
Health facilities providing HIV diagnosis, n (%)			<.001
Hospital	270 (35.1)	3710 (52.7)	
CDC[a]	458 (59.6)	3099 (44.1)	
Other	41 (5.3)	225 (3.2)	
CD4 count at HIV diagnosis (cells/µL), median (IQR)	362 (262-481)	332 (204-459)	<.001
On ART, n (%)			<.001
Yes	740 (96.2)	6247 (88.8)	
No	29 (3.8)	787 (11.2)	
Time from HIV diagnosis until the ART[b] initiation by CD4 count (day), median (IQR)			
Overall (n=7803)	22 (10-85)	93 (16-979)	<.001
0-199 (n=2768)	26 (11-164)	151 (21-1349)	<.001
200-349 (n=2597)	20 (9-88)	80 (15-789)	<.001
≥350 (n=2291)	24 (11-74)	67 (14-739)	<.001
Never tested (n=147)	24 (18-57)	267 (18-1065)	.04

[a] CDC: Centers for Disease Control and Prevention.

[b] ART: antiretroviral therapy.

Mortality Rate Between Student and Nonstudent Youths Living With HIV

Among the 7803 youths infected through sexual contact and who started the ART from 2008 to 2019, a total of 45,610.4 person-years were observed, and 446 youths died before December 31, 2020. The overall death rate among nonstudent youths living with HIV was 1.0 deaths per 100 person-years, which was higher than that among student youths living with HIV (0.4 deaths per 100 person-years). After adjusting for sex, ethnicity, marital status, education, infection route, CD4 count before the ART initiation, and the WHO clinical stage before the ART, nonstudent youths living with HIV had mortality risk doubled compared to student youths with HIV (adjusted hazard ratio [AHR] 2.3, 95% CI 1.2-4.2; *P*=.008). Youths who had initiated a zidovudine-based regimen had a 30% decreased mortality rate compared to those who took a stavudine-based regimen (AHR 0.7, 95% CI 0.5-0.9; *P*=.002). Youths who initiated a tenofovir disoproxil fumarate–based regimen also had a decreased mortality rate compared to the stavudine-based regimen group (AHR 0.7, 95% CI 0.5-1.0; *P*=.03; Table 4).

Table 4. Mortality rate of youths living with HIV aged 16-25 years by student status in Guangxi, China, from 2008 to 2019.

Characteristics	Patients, n	Deaths, n	Person-years, n	Deaths per 100 person-years (95% CI)	HR[a] (95% CI)	*P* value	AHR[b] (95% CI)	*P* value
Overall	7803	446	45,610.4	1.0 (0.9-1.1)				
Age group (16-25 years)						<.001		.008
Student	769	11	2976.3	0.4 (0.2-0.7)	1.0		1.0	
Nonstudent	7034	435	42,634.2	1.0 (0.9-1.1)	3.1 (1.7-5.7)		2.3 (1.2-4.2)	
Calendar year of initiating the ART[c]						.003		.44
2008-2012	2417	229	22,653.7	1.0 (0.9-1.1)	1.0		1.0	
2013-2019	5386	217	22,956.8	0.9 (0.8-1.1)	0.7 (0.6-0.9)		1.1 (0.9-1.5)	
Initial regimen								
Stavudine-based	637	111	5606.3	2.0 (1.6-2.4)	1.0	—[d]	1.0	—
Zidovudine-based	3274	174	23,438.3	0.7 (0.6-0.9)	0.4 (0.3-0.5)	<.001	0.7 (0.5-0.9)	.002
TDF[e]-based	3724	154	15,170.1	1.0 (0.9-1.2)	0.4 (0.3-0.5)	<.001	0.7 (0.5-1.0)	.03
Other	168	7	1395.7	0.5 (0.2-1.1)	0.3 (0.1-0.5)	<.001	0.3 (0.2-0.7)	.005

[a] HR: hazard ratio.

[b] AHR: adjusted hazard ratio; adjusted for sex, ethnicity, marital status, education, infection route, CD4 count before the antiretroviral therapy (ART) initiation, and the World Health Organization (WHO) clinical stage before the ART.

[c] ART: antiretroviral therapy.

[d] Not applicable.

[e] TDF: tenofovir disoproxil fumarate.

Virologic Failure Between Student and Nonstudent Youths Living With HIV

We assessed the virologic failure among youths who started ART in 2008 since free and standardized ART programs started from 2008 in China. Of the 4119 youths who started ART during 2008-2019, a total of 24,388.6 person-years were observed, and 632 youths met the clinical criteria of virologic failure during the follow-up period. The virologic failure rate among nonstudent youths with HIV was 2.6 deaths per 100 person-years, which was higher than that of student youths with HIV (2.3 per 100 person-years). After being adjusted for sex, ethnicity, marital status, education, infection route, CD4 count before the ART initiation, and the WHO clinical stage before the ART, the risk of virologic failure among nonstudent youths with HIV had doubled compared to student youths (AHR 1.9, 95% CI 1.3-2.6; *P*<.001). Youths who had initiated the ART during 2013-2019 had a 30.9 times increased risk of virologic failure compared with those who initiated the ART during 2008-2012 (AHR 31.9, 95% CI 18.2-55.9; *P*<.001). Youths who initiated the ART with a zidovudine-based (AHR 1.5, 95% CI 1.2-2.0; *P*=.003) and TDF-based regimen (AHR 1.9, 95% CI 1.4-2.6; *P*<.001) had an increased risk of virologic failure compared to those who took a stavudine-based regimen (Table 5).

Table 5. Virologic failure (≥400 copies/mL) of youths living with HIV aged 16-25 years by student status in Guangxi, China, from 2008 to 2019.

Characteristics	Patients, n	Virologic failure, n	Person-years, n	Virologic failure per 100 person-years (95% CI)	HR[a] (95% CI)	*P* value	AHR[b] (95% CI)	*P* value
Overall	4199	632	24,388.6	2.6 (2.4-2.8)				
Age group (16-25 years)						.01		<.001
Student	428	38	1664.2	2.3 (1.7-3.1)	1.0		1.0	
Nonstudent	3691	594	22,724.4	2.6 (2.4-2.8)	0.7 (0.5-0.9)		1.9 (1.3-2.6)	
Calendar year of initiating the ART[c]						<.001		<.001
2008-2012	1247	270	12,004.0	2.3 (2.0-2.5)	1.0		1.0	
2013-2019	2872	362	12,384.6	2.9 (2.6-3.2)	46.7 (28.1-77.4)		31.9 (18.2-55.9)	
Initial regimen								
Stavudine-based	325	75	3034.0	2.5 (2.0-3.1)	1.0	—[d]	1.0	—
Zidovudine-based	1658	304	12,168.0	2.5 (2.2-2.8)	1.6 (1.2-2.0)	<.001	1.5 (1.2-2.0)	.003
TDF[e]-based	2052	241	8543.4	2.8 (2.5-3.2)	5.0 (3.7-6.7)	<.001	1.9 (1.4-2.6)	<.001
Other	84	12	643.1	1.9 (1.1-3.3)	0.6 (0.3-1.1)	.08	0.5 (0.3-1.0)	.05

[a] HR: hazard ratio.

[b] AHR: adjusted hazard ratio; adjusted for sex, ethnicity, marital status, education, infection route, CD4 count before the ART initiation, and the WHO clinical stage before the ART.

[c] ART: antiretroviral therapy.

[d] Not applicable.

[e] TDF: tenofovir disoproxil fumarate.

Discussion

To our knowledge, this is the first study investigating the characteristics of youths living with HIV in southwest China. Nonstudent youths living with HIV accounted for 93.6% (12956/13839) of the total HIV cases in youths during 1996-2019. This percentage was much higher than that in youths aged 15-24 years in Hangzhou, China, during 2012-2016 (70.4%) [6]. Given the low numbers before 2006, an obvious increasing trend was observed among student youths living with HIV, approximately one-third of nonstudent cases in 2019. Homosexual transmission was the main infection route among student youths living with HIV in Guangxi (n=614, 69.5%), while heterosexual contact was the primary transmission route among nonstudent cases (n=7887, 60.9%), which was different from the primary transmission route in Hangzhou (ie, homosexual transmission) [6]. Heterosexual contact is the primary mode of HIV transmission in China [7,19]. Although the mean CD4 count at the time of diagnosis was slightly higher in student cases and consistently increased from 1996 to 2019 among student and nonstudent youths living with HIV, there is still a need to promote frequent HIV testing for early HIV diagnosis in youths in the future.

Among youths who were infected with HIV through sexual contact, the 2 groups showed differences: student youths with HIV were mainly male (n=820, 94.5% vs n=4687, 50.2%) and unmarried (n=862, 99.3%), and homosexual contact was the primary transmission route (n=614, 70.7%). In general, youths were considered sexually active and less likely to use a condom [20]. Unprotected sexual activity raises the risk of HIV transmission among youths [21]. Moreover, our study demonstrated that most nonstudent youths living with HIV had an educational level of middle school or below (n=6558, 70.3%). Limited access to sexual education, leading to the lack of HIV awareness, could be a possible reason of high-risk sexual behaviors. Thus, these findings could be applied in tailored HIV prevention programs with the goal of reducing the HIV epidemic among youths. In addition, more attention should be given to homosexual student and nonstudent youths who dropped out before high school.

Furthermore, student youths living with HIV had declined CD4 levels at the time of HIV diagnosis, and the CD4 level for nonstudent cases was even lower. This might be a result of delayed HIV testing. Besides, more than half of the nonstudents with HIV (n=3710, 52.7%) received an HIV diagnosis in hospitals instead of CDCs, implying low awareness of HIV testing, as CDCs have voluntary counseling and testing, but clinics and hospitals usually conduct HIV tests as part of physical examination. Moreover, student youths with HIV (n=740, 96.2%) had a slightly higher rate of the ART compared to nonstudent cases. At the same time, nonstudents had a longer median time for the ART initiation (93 days) compared to students (22 days). Stigma and lack of knowledge are major factors leading to delayed ART [6]. Therefore, future programs should promote early diagnosis and treatment to enhance the treatment efficacy and improve outcomes [11], and consequently, reduce morbidity and mortality [22].

Nonstudnet youths living with HIV had a higher risk of experiencing virologic failure and death compared with student cases after the ART initiation. More attention should be given to nonstudents for preventing virologic failure and death. It could be explained by the findings of this study that nonstudent youths with HIV in Guangxi were less likely to adopt the ART, had a lower CD4 count at the time of HIV diagnosis, and a delayed ART initiation. Early ART initiation could lead to improved clinical outcomes [23]. Therefore, future programs should work on eliminating barriers to HIV care and promoting early ART initiation among youths, especially among nonstudents. In this study, 70.9% (n=9189) of nonstudents youths with HIV in Guangxi had an educational level of middle school or below, and only 8.3% (n=1078) obtained university or higher education. Previous studies found that low level of education is associated with delayed HIV care [24], and school-based interventions could increase HIV knowledge and awareness among students [25,26], which in turn could probably reduce their risk behaviors and promote benefit behaviors. In addition to the school-based interventions, various interventions targeting out-of-school youths should be promoted to increase ART awareness and reduce risky sexual behaviors among the youths.

Based on the findings of this study, tailored HIV programs could be more effective in HIV prevention among the youths. For nonstudents aged 16-25 years, who accounted for the majority of HIV cases, social media–oriented HIV education using platforms like Douyin and WeChat has a great potential to deliver future HIV programs due to their wide coverage and accessibility. For students, more attention could be given to the high-risk groups, such as homosexual men. At the same time, sex education and safe sex should be strengthened in middle school, high school, and college. Given the high mortality rate as well as high risk of virologic failure and delayed ART among nonstudent youths with HIV, increased health protection awareness is crucial to encourage early diagnosis and early ART initiation [27]. Moreover, peer counseling and group therapy might be good ways to address psychological barriers to HIV testing and care, including stigma and lack of support [28,29]. Furthermore, HIV prevention campaigns could also create more awareness among the general public and clinic doctors, reduce stigma, and create a safe environment for youths to discuss their behaviors [30].

However, this study had some limitations. First, the study used historical cohort study data extracted from existing public health databases, and some information that might be useful were not available, such as the details on ART adherence and its side effects. Second, we investigated all-cause mortality instead of AIDS-related mortality, which might give different results. Third, students and nonstudents were generalized, disregarding their level of education; for example, middle and high school students might have different behaviors in HIV care and treatment outcomes from college students. For nonstudents, different levels of education may result in different perceptions and concerns about health. Thus, we adjusted for the educational background in the multivariate analysis to correct these discrepancies. Despite these limitations, this study provides valuable information on HIV risk behaviors, HIV care, and treatment outcomes for both student and nonstudent youths living with HIV. Overall, these findings will guide future HIV prevention and intervention programming in youths.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (82160636, 12071366, 82160642, 81760616, and 82260670), the Guangxi Natural Science Foundation Project grants (2020GXNSFAA159020), Guangxi Key Laboratory of AIDS Prevention Control and Translation (ZZH2020010), technical application and promotion plan of Guilin City (2020011203-6), Ministry of Science and Technology of China (2022YFC2305200), Guangxi Bagui Honor Scholarship, Guangxi Medical and health key discipline construction project, and Chinese State Key Laboratory of Infectious Disease Prevention and Control.

Conflicts of Interest

None declared.

References

1. Global AIDS report 2020. Joint United Nations Programme on HIV/AIDS. URL: https://www.unaids.org/en/resources/documents/2020/unaids-data [accessed 2021-07-06]
2. van den Berg JJ, Gamarel KE, Westfall AO, Fortenberry JD, Hosek SG, Wilson CM, et al. Transmission risk among youth living with HIV in the US. J Adolesc Health 2020 Jul;67(1):61-68 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.01.008] [Medline: 32169529]
3. Ye R, Liu C, Tan S, Li J, Simoni JM, Turner D, et al. Factors associated with past HIV testing among men who have sex with men attending university in China: a cross-sectional study. Sex Health 2021;18(1):58. [doi: 10.1071/sh20088]
4. Gibbs A, Reddy T, Dunkle K, Jewkes R. HIV-prevalence in South Africa by settlement type: a repeat population-based cross-sectional analysis of men and women. PLoS One 2020 Mar 17;15(3):e0230105 [FREE Full text] [doi: 10.1371/journal.pone.0230105] [Medline: 32182255]
5. Dong Y, Wang L, Burgner DP, Miller JE, Song Y, Ren X, et al. Infectious diseases in children and adolescents in China: analysis of national surveillance data from 2008 to 2017. BMJ 2020 Apr 02;369:m1043 [FREE Full text] [doi: 10.1136/bmj.m1043] [Medline: 32241761]

6. Luo Y, Wu H, Zhang X, Li X, Scott S, Chen J, et al. HIV care continuum among newly diagnosed student and non-student youths between 2012 and 2016 in Hangzhou, China. Chin Med J (Engl) 2019 Jun 20;132(12):1420-1428 [FREE Full text] [doi: 10.1097/CM9.0000000000000264] [Medline: 31205099]
7. Zhang X, Wang N, Vermund SH, Zou H, Li X, Zhang F, et al. Interventions to improve the HIV continuum of care in China. Curr HIV/AIDS Rep 2019 Dec;16(6):448-457. [doi: 10.1007/s11904-019-00469-y] [Medline: 31776975]
8. Luo QQ, Luo YC, Ren XL. [Comparison of HIV infection risk between 15 to 24 year-old student men who have sex with men and non-student men who have sex with men: a cross-sectional study]. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Dec 06;54(12):1435-1440. [doi: 10.3760/cma.j.cn112150-20191217-00938] [Medline: 33333663]
9. Yoshimura K. Current status of HIV/AIDS in the ART era. J Infect Chemother 2017 Jan;23(1):12-16. [doi: 10.1016/j.jiac.2016.10.002] [Medline: 27825722]
10. Boyd M, Boffito M, Castagna A, Estrada V. Rapid initiation of antiretroviral therapy at HIV diagnosis: definition, process, knowledge gaps. HIV Med 2019 Mar 06;20 Suppl 1(S1):3-11 [FREE Full text] [doi: 10.1111/hiv.12708] [Medline: 30724450]
11. Teeraananchai S, Kerr S, Amin J, Ruxrungtham K, Law M. Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis. HIV Med 2017 Apr 31;18(4):256-266 [FREE Full text] [doi: 10.1111/hiv.12421] [Medline: 27578404]
12. Nachega J, Hislop M, Nguyen H, Dowdy D, Chaisson R, Regensberg L, et al. Antiretroviral therapy adherence, virologic and immunologic outcomes in adolescents compared with adults in southern Africa. J Acquir Immune Defic Syndr 2009 May 01;51(1):65-71 [FREE Full text] [doi: 10.1097/QAI.0b013e318199072e] [Medline: 19282780]
13. Zanoni BC, Archary M, Buchan S, Katz IT, Haberer JE. Systematic review and meta-analysis of the adolescent HIV continuum of care in South Africa: the Cresting Wave. BMJ Glob Health 2016 Oct 24;1(3):e000004 [FREE Full text] [doi: 10.1136/bmjgh-2015-000004] [Medline: 28588949]
14. Jobanputra K, Parker LA, Azih C, Okello V, Maphalala G, Kershberger B, et al. Factors associated with virological failure and suppression after enhanced adherence counselling, in children, adolescents and adults on antiretroviral therapy for HIV in Swaziland. PLoS One 2015 Feb 19;10(2):e0116144 [FREE Full text] [doi: 10.1371/journal.pone.0116144] [Medline: 25695494]
15. Apondi E, Wachira J, Ayikukwei R, Kafu C, Onyango J, Omollo M, et al. Barriers to ART adherence among school students living with HIV in Kenya. Afr J AIDS Res 2021 Oct 11;20(3):232-237. [doi: 10.2989/16085906.2021.1979606] [Medline: 34635018]
16. Sun X, Yang W, Tang S, Shen M, Wang T, Zhu Q, et al. Declining trend in HIV new infections in Guangxi, China: insights from linking reported HIV/AIDS cases with CD4-at-diagnosis data. BMC Public Health 2020 Jun 12;20(1):919 [FREE Full text] [doi: 10.1186/s12889-020-09021-9] [Medline: 32532238]
17. Chen H, Luo L, Pan SW, Lan G, Zhu Q, Li J, et al. HIV epidemiology and prevention in Southwestern China: trends from 1996-2017. CHR 2019 Sep 02;17(2):85-93. [doi: 10.2174/1570162x17666190703163838]
18. Mao Y, Wu Z, Poundstone K, Wang C, Qin Q, Ma Y, et al. Development of a unified web-based national HIV/AIDS information system in China. Int J Epidemiol 2010 Dec 26;39 Suppl 2(Suppl 2):ii79-ii89 [FREE Full text] [doi: 10.1093/ije/dyq213] [Medline: 21113041]
19. Dong Z, Ma L, Cai C, Gao G, Lyu F. Demographic features of identified PLWHA infected through commercial and nonmarital noncommercial heterosexual contact in China from 2015 to 2018: a retrospective cross-sectional study. BMC Infect Dis 2021 Jan 13;21(1):71 [FREE Full text] [doi: 10.1186/s12879-020-05757-2] [Medline: 33441089]
20. Habel MA, Leichliter JS, Dittus PJ, Spicknall IH, Aral SO. Heterosexual anal and oral sex in adolescents and adults in the United States, 2011–2015. Sexual Trans Dis 2018;45(12):775-782. [doi: 10.1097/olq.0000000000000889]
21. Setse R, Siberry G, Gravitt P, Moss W, Agwu A, Wheeling J, LEGACY Consortium. Correlates of sexual activity and sexually transmitted infections among human immunodeficiency virus-infected youth in the LEGACY cohort, United States, 2006. Pediatr Infect Dis J 2011 Nov;30(11):967-973 [FREE Full text] [doi: 10.1097/INF.0b013e3182326779] [Medline: 22001904]
22. Robertson M, Braunstein S, Hoover D, Li S, Nash D. Timeliness of human immunodeficiency virus diagnosis and antiretroviral treatment initiation in the era of universal testing and treatment. J Infect Dis 2019 Jul 19;220(4):648-656. [doi: 10.1093/infdis/jiz148] [Medline: 30997508]
23. Ford N, Migone C, Calmy A, Kerschberger B, Kanters S, Nsanzimana S, et al. Benefits and risks of rapid initiation of antiretroviral therapy. AIDS 2018 Jan 02;32(1):17-23 [FREE Full text] [doi: 10.1097/QAD.0000000000001671] [Medline: 29112073]
24. Minniear TD, Gaur AH, Thridandapani A, Sinnock C, Tolley EA, Flynn PM. Delayed entry into and failure to remain in HIV care among HIV-infected adolescents. AIDS Res Hum Retrovir 2013 Jan;29(1):99-104. [doi: 10.1089/aid.2012.0267]
25. Zheng Y, Zhang X, Sun X, Shi Y, Chang C. Evaluation of the college-based HIV/AIDS education policy in Beijing, China: a mixed method approach. Environ Health Prev Med 2020 Sep 10;25(1):50 [FREE Full text] [doi: 10.1186/s12199-020-00890-5] [Medline: 32912181]
26. Gao X, Wu Y, Zhang Y, Zhang N, Tang J, Qiu J, et al. Effectiveness of school-based education on HIV/AIDS knowledge, attitude, and behavior among secondary school students in Wuhan, China. PLoS One 2012 Sep 7;7(9):e44881 [FREE Full text] [doi: 10.1371/journal.pone.0044881] [Medline: 22970322]

27. Okal J, Lango D, Matheka J, Obare F, Ngunu-Gituathi C, Mugambi M, et al. "It is always better for a man to know his HIV status" - a qualitative study exploring the context, barriers and facilitators of HIV testing among men in Nairobi, Kenya. PLoS One 2020 Apr 15;15(4):e0231645 [FREE Full text] [doi: 10.1371/journal.pone.0231645] [Medline: 32294124]
28. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database Syst Rev 2017 Dec 24;4:CD011279 [FREE Full text] [doi: 10.1002/14651858.CD011279.pub3] [Medline: 28436583]
29. Liu Y, Vermund SH, Ruan Y, Liu H, Rivet Amico K, Simoni JM, et al. Peer counselling versus standard-of-care on reducing high-risk behaviours among newly diagnosed HIV-positive men who have sex with men in Beijing, China: a randomized intervention study. J Int AIDS Soc 2018 Feb 12;21(2):e25079 [FREE Full text] [doi: 10.1002/jia2.25079] [Medline: 29430845]
30. Njau B, Lisasi E, Damian DJ, Mushi DL, Boulle A, Mathews C. Feasibility of an HIV self-testing intervention: a formative qualitative study among individuals, community leaders, and HIV testing experts in northern Tanzania. BMC Public Health 2020 Apr 15;20(1):490 [FREE Full text] [doi: 10.1186/s12889-020-08651-3] [Medline: 32293370]

Abbreviations

AHR: adjusted hazard ratio
ART: antiretroviral therapy
CDC: Centers for Disease Control and Prevention
CRIMS: Chinese Comprehensive Response Information Management System
NCRD: National Case Reporting Database
NFATD: National Free Antiretroviral Therapy Database
WHO: World Health Organization

Edited by A Mavragani; submitted 22.04.22; peer-reviewed by X Zhang, X Zhang; comments to author 24.07.22; revised version received 13.08.22; accepted 29.11.22; published 24.02.23.

Please cite as:
Wang N, Lan G, Zhu Q, Chen H, Huang J, Meng Q, Shen Z, Liang S, Wu X, Luo L, Ye R, Chen J, Tan S, Xing H, Shao Y, Ruan Y, Lin M
HIV Epidemiology, Care, and Treatment Outcomes Among Student and Nonstudent Youths Living With HIV in Southwest China Between 1996 and 2019: Historical Cohort Study
JMIR Public Health Surveill 2023;9:e38881
URL: https://publichealth.jmir.org/2023/1/e38881
doi:10.2196/38881
PMID:36826980

Original Paper

Digital-Assisted Self-interview of HIV or Sexually Transmitted Infection Risk Behaviors in Transmasculine Adults: Development and Field Testing of the Transmasculine Sexual Health Assessment

Sari L Reisner[1,2,3,4], ScD; David R Pletta[3], MPH; Dana J Pardee[4], MBA; Madeline B Deutsch[5], MPH, MD; Sarah M Peitzmeier[6], MSPH, PhD; Jaclyn MW Hughto[7], MPH, PhD; Meg Quint[1], BS; Jennifer Potter[2,4,8], MD

[1]Brigham and Women's Hospital, Boston, MA, United States
[2]Harvard Medical School, Boston, MA, United States
[3]Harvard T.H. Chan School of Public Health, Boston, MA, United States
[4]Fenway Health, Boston, MA, United States
[5]University of California San Francisco, San Francisco, CA, United States
[6]University of Michigan, Ann Arbor, MI, United States
[7]Brown University School of Public Health, Providence, RI, United States
[8]Beth Israel Deaconess Medical Center, Boston, MA, United States

Corresponding Author:
Sari L Reisner, ScD
Brigham and Women's Hospital
221 Longwood Ave
5th Floor
Boston, MA, 02115
United States
Phone: 1 15712437532
Email: sreisner@bwh.harvard.edu

Abstract

Background: The sexual health of transmasculine (TM) people—those who identify as male, men, or nonbinary and were assigned a female sex at birth—is understudied. One barrier to conducting HIV- and sexually transmitted infection (STI)–related research with this population is how to best capture sexual risk data in an acceptable, gender-affirming, and accurate manner.

Objective: This study aimed to report on the community-based process of developing, piloting, and refining a digitally deployed measure to assess self-reported sexual behaviors associated with HIV and STI transmission for research with TM adults.

Methods: A multicomponent process was used to develop a digital-assisted self-interview to assess HIV and STI risk in TM people: gathering input from a Community Task Force; working with an interdisciplinary team of content experts in transgender medicine, epidemiology, and infectious diseases; conducting web-based focus groups; and iteratively refining the measure. We field-tested the measure with 141 TM people in the greater Boston, Massachusetts area to assess HIV and STI risk. Descriptive statistics characterized the distribution of sexual behaviors and HIV and STI transmission risk by the gender identity of sexual partners.

Results: The Transmasculine Sexual Health Assessment (TM-SHA) measures the broad range of potential sexual behaviors TM people may engage in, including those which may confer risk for STIs and not just for HIV infection (ie, oral-genital contact); incorporates gender-affirming language (ie, *genital* or *frontal* vs *vaginal*); and asks sexual partnership characteristics (ie, partner gender). Among 141 individual participants (mean age 27, SD 5 years; range 21-29 years; n=21, 14.9% multiracial), 259 sexual partnerships and 15 sexual risk behaviors were reported. Participants engaged in a wide range of sexual behaviors, including fingering or fisting (receiving: n=170, 65.6%; performing: n=173, 66.8%), oral-genital sex (receiving: n=182, 70.3%; performing: n=216, 83.4%), anal-genital sex (receptive: n=31, 11.9%; insertive: n=9, 3.5%), frontal-genital sex (receptive: n=105, 40.5%; insertive: n=46, 17.8%), and sharing toys or prosthetics during insertive sex (n=62, 23.9%). Overall barrier use for each sexual behavior ranged from 10.9% (20/182) to 81% (25/31). Frontal receptive sex with genitals and no protective barrier was the highest

(21/42, 50%) with cisgender male partners. In total, 14.9% (21/141) of participants reported a lifetime diagnosis of STI. The sexual history tool was highly acceptable to TM participants.

Conclusions: The TM-SHA is one of the first digital sexual health risk measures developed specifically with and exclusively for TM people. TM-SHA successfully integrates gender-affirming language and branching logic to capture a wide array of sexual behaviors. The measure elicits sexual behavior information needed to assess HIV and STI transmission risk behaviors. A strength of the tool is that detailed partner-by-partner data can be used to model partnership-level characteristics, not just individual-level participant data, to inform HIV and STI interventions.

(JMIR Public Health Surveill 2023;9:e40503) doi:10.2196/40503

KEYWORDS

transgender; sexual health; HIV; sexually transmitted infection; STI; epidemiology; mobile phone

Introduction

Background

Transmasculine (TM) people—those who identify as men, male, transgender men, or a nonbinary gender identity and were assigned female sex at birth [1]—are at risk of HIV and other sexually transmitted infections (STIs), such as gonorrhea, chlamydia, and human papillomavirus (HPV). Historically, research has largely focused on the risk of HIV and STIs in transfeminine populations, that is, those who identify as women, female, transgender women, or a nonbinary gender identity and were assigned male sex at birth [2,3]. Studies on HIV and STI burden in TM people remain scarce [2,4]. Moreover, there is a lack of consistency in the reported rates of STIs in transgender men, with 1 review finding STI ranges of 0% to 4.2% for syphilis, 0% to 10.5% for gonorrhea, and 1.2% to 11.1% for chlamydia [2]. It is important to capture the diverse sexual practices of TM people that may be relevant not only for HIV but also for STIs such as genital or anal HPV, gonorrhea, and chlamydia, as well as for counseling on family planning and contraception needs. Failure to collect complete detail regarding sexual behaviors and the timing and recency of those behaviors may result in misclassification of sexual risk exposures, especially given the heterogeneity of sexual behaviors, gender identities, and the bodies of TM people and their sexual partners.

TM people hold a wide range of sexual identities and engage in a variety of sexual behaviors based on their own and their partner's anatomy. Analyzing data from the US National Transgender Discrimination Survey, a sample of 2578 TM respondents identified their sexuality as gay (19%), bisexual (13%), and queer (51%) [5], supporting the need for measures that are able to capture the diversity of sexual partners and behaviors. Statistics show that TM people vary in terms of access to, and desire for, genital surgery. In the 2015 US. Transgender Survey, only 5% of TM respondents had a metoidioplasty (detachment of clitoral ligaments to lengthen an enlarged clitoris to create a neophallus) or phalloplasty (creation of a neophallus using extragenital tissue, typically from the forearm or thigh), but >25% and 19% indicated that they "someday" want a metoidioplasty and phalloplasty, respectively [6,7]. Regardless of anatomy, TM people may or may not engage in sexual behaviors that use their own or their partner's genitals. The wide range of anatomies and sexual identities confers varying sexual behaviors among TM people, which may include higher risk behaviors such as penetrative receptive sex with cisgender gay men that places TM people at risk for the acquisition of HIV and other STIs. Moreover, TM people who experience barriers to surgical access or who do not desire gender-affirming surgery with cervix removal are at risk of cervical cancer caused by high-risk HPV, which can be transmitted or acquired through oral sex, anal sex, and genital-genital contact. Although estimates of HPV are comparable between cisgender women and TM people [8-12], TM people are less likely to receive cervical cancer screenings according to recommended guidelines, underscoring the need for a sexual health measure that allows providers to better understand whether TM patients are at risk of HPV exposure [13]. Other STIs, including gonorrhea, chlamydia, and herpes simplex virus, also warrant attention.

Previous HIV or STI risk surveillance measures, such as the AIDS Risk Behavior Assessment and Risk Assessment Battery, are commonly used and adapted to assess HIV risk [14,15]. However, these measures have limitations. First, these measures use gendered, cisnormative language (eg, assumptions that those who have penises are men and use he or him pronouns), which can be harmful to TM people who are frequently misgendered and systematically misrepresented in data collection [16]. Lack of acceptability of measures in TM populations may lead to nonresponse bias and unwillingness to participate in research or share their sexual history honestly with providers. Furthermore, cisnormative language generates validity concerns when TM people are not able to answer questions accurately according to their or their partners' anatomies. Second, the measures often do not consider sexual practices that may be more common among TM people, such as the use of prosthetics or toys during sex. These sexual practices are common in TM communities and carry the risk of HPV, gonorrhea, chlamydia, and herpes simplex virus transmission or acquisition but are not assessed by traditional measures. Transgender and gender-diverse activists continue to call for inclusive HIV or STI research measures, notably stating, "A lack of research does not mean we are not at risk" [17,18]. Traditional measures may, therefore, underestimate HIV or STI risk in this population by failing to assess common sexual practices. Third, many measures fail to distinguish between insertive and receptive sexual practices, which have different implications for HIV or STI risk and the probability of transmission or acquisition [19]. To our knowledge, only 1 inclusive sexual health and reproductive health inventory has been published for transgender and nonbinary people, which captures sexual behavior, gender affirmation, pregnancy, abortion, and contraception and is of great use for research but may be too lengthy for providers in clinics to accurately and quickly assess HIV or STI risk in only TM people [20]. A study by Bauer et al [21] exploring the sexual health of TM people who have sex with men indicated that they developed and used measures that include a wide range of sexual behaviors, but they did not provide structure or enough details to accurately recreate those measures.

There is limited knowledge and lack of consensus about how best to capture sexual risk data in the TM population, both in terms of accurately characterizing sexual behaviors that confer specific risks for HIV and STI transmission and ensuring that the measure is gender affirming and acceptable for TM respondents. Digital-assisted self-interviewing uses computer- or technology-deployed methods (eg, laptops, electronic tablets, and smartphones) to collect survey data, wherein the respondent completes the assessment via self-reporting without an interviewer administering it. The capacity for programmed branch logic and the highly sensitive nature of sexual risk behavior questions makes this a method well suited for confidential data collection of self-reported HIV- and STI-related risks for TM adults [22,23].

Objectives

This study aimed to develop and test a robust and inclusive sexual health and risk assessment designed *with* and specifically *for* TM people. The aims were twofold: (1) to report on the community-based process of developing and refining a digital-assisted self-interview measure designed to assess self-reported sexual behaviors associated with HIV and STI transmission for research in TM adults, known as the Transmasculine Sexual Health Assessment (TM-SHA), and (2) to field and digitally implement the TM-SHA and characterize the sexual histories, sexual partnership, and sexual practices in a sample of TM adults.

Methods

Overview

This research was undertaken from March 2015 to September 2016 as part of a biobehavioral study to evaluate screening methods for the detection of high-risk HPV in TM adults (ClinicalTrials.gov NCT02401867) [24]. The main outcomes and broader study details have been reported elsewhere [8]. This secondary analysis focuses on the development, implementation, and refinement of the HIV and STI sexual risk measure, including reporting on HIV and STI transmission risk by the gender of sexual partners.

Ethics Approval

All study activities were approved by the Fenway Health Institutional Review Board (FWA00000145).

Consent to Participate

Informed consent procedures ensured that the people understood what was involved in participating in the study, what the risks and benefits were of participation, and voluntarily agreed to participate. The original informed consent also applied to this secondary analysis, for which all study data were anonymized and deidentified. Participants were compensated US $100 for study participation in the form of a prepaid American Express gift card.

Community-Scientific Partnership

A multicomponent community-based process was used to develop and refine a questionnaire to assess HIV and STI risk in TM people using digital-assisted self-interviewing technology. The process used community-based participatory research principles to work *with* not *on* TM communities and to cocreate and refine the sexual behavior assessment. The development process was guided by a 15-member research team that included multiple TM adults who worked closely with a 5-member Community Task Force of TM people and a 5-member multidisciplinary Scientific Advisory Board that comprised content experts in transgender medicine, clinical epidemiology, and infectious diseases research. This was a dynamic process involving community and scientific engagement. At the core of the enterprise was the following tenet: to prioritize being gender affirming and respectful, while collecting accurate data in a patient-centered manner to understand and address sexual health needs of TM people.

Development, Modification, and Refinement of the Measure

The TM-SHA was developed to assess sexual behaviors, partnerships, and protective barrier use that confer transmission risk of HIV and other STIs. The primary purpose was to measure the broad range of sexual behaviors engaged in by TM adults, including those that may confer risk for HIV, HPV, and other STIs, in a gender-affirming, sensitive, and accurate manner.

The basic TM-SHA assessment structure was modeled on an HIV risk assessment originally adapted from the AIDS Risk Behavior Assessment [14] by the LifeSkills trial, a multisite study testing the efficacy of an HIV behavioral intervention to reduce HIV transmission risk for young transgender women aged 16-29 years [20]. The measure was subsequently adapted with young adult transgender men who have sex with men aged 18-29 years in LifeSkills for Men [25]. These HIV risk measures specifically focused on sexual and other behaviors that confer risk for HIV infection, such as genital and anal intercourse and needle sharing in the context of injection drug use or hormone therapy. This study expanded on this to include a broader range of behaviors.

The development of the TM-SHA involved active collaboration between the research team, Task Force, and Scientific Advisory Board to ensure the creation of a gender-affirming assessment that reflected the sexual lives and experiences of TM adults. Among these groups, we discussed the topics of language and terminology, usability, question clarity and specificity, response options, and acceptability. A preliminary version of the TM-SHA was then pilot-tested with members of the Task Force and the research team, who provided feedback regarding the instrument’s functionality and design. After addressing feedback from the pilot test, we field-tested the TM-SHA in a sample of 141 TM adults, all of whom had cervixes and reported at least 1 sexual partner in the past 12 months (see *Field Test of the Assessment* section).

When reviewing the preliminary version of the measure, the following ideas emerged about the components of the ideal assessment: asking about an array of sexual behaviors, including those that are not typically associated with HIV risk; adding an introduction about the complexity of terminology; asking participants' surgical status and programming skip patterns to assess sexual risk accordingly; asking participants what language they use to describe their body parts and then having that populate subsequent questions about sexual behaviors and risks; and assessing sexual partner and partnership characteristics.

On the basis of these ideas, the following changes and refinements to the instrument were made: adding questions to assess the broad range of potential sexual behaviors TM people may engage in, including those which may confer risk for HPV and other STIs, not only for HIV infection; incorporating gender-affirming language to make questions and response options less gendered and binary sex specific (ie, using more generic terms like *genital* vs *vaginal*); and asking sexual partner and partnership characteristics to understand interpersonal context.

TM-SHA Measure

The TM-SHA measure is summarized in Table 1, and Table 2 presents the grid of sexual behaviors and frequency of barrier use (the full TM-SHA measurement and survey logic guidelines can be obtained by contacting the corresponding author). The TM-SHA assessed individual- and partnership-level risks. At the individual level, participants were asked to provide information about their sexual orientation and the total number of sexual partners in the past 36 months.

Participants were then asked to provide demographic and sexual risk–related data for up to 3 sexual partnerships from the past 12 months. These partnership-level data included the gender identity of the sexual partner, configuration of the sexual partnership, behaviors engaged in during sex, frequency of protective barrier use by sexual behavior, and whether the sexual partner had ever been diagnosed with an STI.

Sexual behavior data collected through the TM-SHA consisted of the following 15 sexual behaviors: (1) performing frontal (ie, vaginal) and/or anal penetration with a finger or fist, (2) receiving frontal and/or anal penetration with a finger or fist, (3) performing oral-genital sex, (4) receiving oral-genital sex, (5) performing oral-anal sex, (6) receiving oral-anal sex, (7) frontal insertive sex with genitals, (8) frontal receptive sex with genitals, (9) frontal insertive sex with a toy or prosthetic, (10) frontal receptive sex with a toy or prosthetic, (11) anal insertive sex with genitals, (12) anal receptive sex with genitals, (13) anal insertive sex with a toy or prosthetic, (14) anal receptive sex with a toy or prosthetic, and (15) insertive sex in which a toy or prosthetic was shared.

The use of a protective barrier can be operationalized from the measure as using a condom, internal condom, dental dam, or gloves during sexual activity. The frequency of protective barrier use can be coded as either engaged in sexual behavior with no use of a protective barrier or engaged in sexual behavior and used a protective barrier on at least 1 occasion.

Table 1. Overview of the Transmasculine Sexual Health Assessment.

Section, section #, and item		Type
Introduction		
1	This part of the survey will ask you about sex. Sex is a personal issue that can sometimes be sensitive or difficult to talk about, especially for those who are transgender or gender nonconforming because bodies do not always reflect identity. We will be asking you about different forms of sexual activity including the following: penetrative sex with fingers (“fingering”) or fists (“fisting”), oral-genital sex (putting your mouth on someone’s genitals or someone putting their mouth on your genitals), oral-anal sex (“rimming”), receptive and insertive frontal sex, receptive and insertive anal sex, and use of prosthetics/toys. By “receptive sex” we mean when a person inserts their genitals or a toy into your frontal opening or anus. By “insertive sex” we mean when your genitals are inserted into a frontal opening or anus without the assistance of a prosthetic/dildo/toy/etc. We will also be asking you about your use of barriers during sexual activity. Barriers differ depending on the sexual activity. Barrier methods include condoms, internal condoms, dental dams, and gloves. Except where indicated, all questions about sex refer to consensual sex or sex you experienced because you wanted to participate and not because you were forced, coerced, or otherwise made to have sex. Remember your answers to these questions will be kept completely private. Please try your best to answer each question.	Prompt
Demographics		
1	Please answer the following questions about your sexual identity, attractions, and recent sexual activity.	Prompt
2	Which of the following best describes your sexual identity or orientation today? Choose one... (1) gay/homosexual/same-gender attraction; (2) straight/heterosexual; (3) bisexual; (4) queer; (5) pansexual; (6) questioning; (7) asexual; (8) unsure; (9) I do not label my sexual orientation; (97) Other, please specify; and (99) I prefer not to answer.	Multiple choice
Sexual contact (last 36 months)		
1	Please think about the people you have had sexual contact with in the last 36 months (3 years).	Prompt
2	How many individuals have you had any form of sexual contact with in the last 36 months (3 years)? Sexual contact includes penetrative sex using fingers or fists, oral-genital sex, oral-anal sex, receptive and/or insertive frontal sex, receptive and/or insertive anal sex, and use of prosthetics/toys. If you would prefer not to answer, please enter “99,999.”	Open response
3	Please specify the gender(s) of the partner(s) you have had sexual contact with in the last 36 months (3 years) and the number of partner(s) of each gender. The total must be equal to the number you specified in the previous question. If you did not have sexual contact with a person of a specified gender, please enter 0. If you would prefer not to answer, please enter “99,999”; in each field that you do not wish to answer ... (1) cisgender/nontransgender man, (2) cisgender/nontransgender woman, (3) transgender man (FTM[a]), (4) transgender woman (MTF[b]), (5) male-assigned gender nonconforming/nonbinary person, and (6) female-assigned gender nonconforming/nonbinary person.	Multiple choice
4	Of the partners that you had sexual contact with in the last 36 months (3 years), how many did you have unprotected frontal and/or anal sex with (ie, without a condom or other barrier)? This includes penetrative sex using fingers or fists, oral-genital sex, oral-anal sex, receptive and/or insertive frontal sex, receptive and/or insertive anal sex, and use of prosthetics/toys. If you would prefer not to answer, please enter “99,999.”	Open response
Sexual contact (last 12 months)		
1	Please think about the people you have had sexual contact with in the last 12 months (1 year).	Prompt
2	How many individuals have you had any form of sexual contact with in the last 12 months (1 year)? Sexual contact includes penetrative sex using fingers or fists, oral-genital sex, oral-anal sex, receptive and/or insertive frontal sex, receptive and/or insertive anal sex, and use of prosthetics/toys. If you would prefer not to answer, please enter “99,999.”	Open response
3	Please specify the gender(s) of the partner(s) you have had sexual contact with in the last 12 months (1 year) and the number of partner(s) of each gender. The total must be equal to the number you specified in the previous question. If you did not have sexual contact with a person of a specified gender, please enter 0. If you would prefer not to answer, please enter “99,999”; in each field that you do not wish to answer... (1) cisgender/nontransgender man, (2) cisgender/nontransgender woman, (3) transgender man (FTM), (4) transgender woman (MTF), (5) male-assigned gender nonconforming/nonbinary person, and (6) female-assigned gender nonconforming/nonbinary person.	Multiple choice
4	Of the partners that you had sexual contact with in the last 12 months (1 year), how many did you have unprotected frontal and/or anal sex with (ie, without a condom or other barrier)? This includes penetrative sex using fingers or fists, oral-genital sex, oral-anal sex, receptive and/or insertive frontal sex, receptive and/or insertive anal sex, and use of prosthetics/toys. If you would prefer not to answer, please enter “99,999.”	Open response
Identifying number of sexual partners (last 12 months)		
1	We will now ask you about your 3 most recent sexual partners in the last 12 months (1 year).	Prompt

Section, section #, and item		Type
2	Thinking about the people that you have had sexual contact with in the last 12 months, please enter the initials of your most recent sexual partner. If you do not know the person's initials or do not feel comfortable providing that information, please create your own way of identifying this partner. [c]Note: if you have had sex with >1 partner at the same time, please choose only 1 partner for this question, and use the next question(s) to identify the other, concurrent partner(s).	Open response
3	Thinking about the people that you have had sexual contact with in the last 12 months, did you have another partner in addition to (partner 1 initials)? ... (1) yes, (2) no, and (99) I prefer not to answer.	Multiple choice
4	(If yes) Please enter the initials of this sexual partner. If you do not know the person's initials or do not feel comfortable providing that information, please create your own way of identifying this partner.	Open response
5	(If sex partners last 12 months >2) Thinking about the people that you have had sexual contact with in the last 12 months, did you have another partner in addition to (partner 1 initials) and (partner 2 initials)? ... (1) yes, (2) no, and (99) I prefer not to answer.	Multiple choice
6	(If yes) Please enter the initials of this sexual partner. If you do not know the person's initials or do not feel comfortable providing that information, please create your own way of identifying this partner.	Open response
Sexual partner one		
1	Please answer the following questions about your interactions with (partner 1 initials).	Prompt
2	What was (partner 1 initials)'s gender? ... (1) cisgender/nontransgender man, (2) cisgender/nontransgender woman, (3) transgender man (FTM), (4) transgender woman (MTF), (5) male-assigned gender nonconforming/nonbinary person; (6) female-assigned gender nonconforming/nonbinary person, and (99) I prefer not to answer.	Multiple choice
3	How would you describe your relationship with (partner 1 initials)? If >1 description applies, please select "Other" and describe... (1) married or in a civil partnership; (2) serious relationship (boyfriend/girlfriend/partner), someone you dated for a while and feel very close to; (3) casually dating but not serious; (4) poly (polyamorous); (5) open relationship/nonmonogamous; (6) sleeping with this person ("fuck buddy" or "booty call") but not dating; (7) dom/sub (dominant/submissive); (8) fluid bonded; (9) one night stand; (10) stranger or anonymous person; (11) exchange partner/sex work client; (97) other, please specify; and (99) I prefer not to answer.	Multiple choice
4	For the following questions, please identify whether or not you engaged in any of the following sexual activities in the last 12 months with this partner (partner 1 initials). If you respond "Yes," please specify how often a barrier was used. If you respond "No," please select "I did not engage in this activity" under barrier use. [c]Note: as a reminder, by "barrier" we mean condoms, internal condoms, dental dams, and gloves.	Prompt
5	Sexual behaviors and frequency of protective barrier use grid[d]	Multiple choice
6	STI[e] history[d]: Has this partner ever been diagnosed (by a physician, nurse, or other medical provider) with any of the following STIs?: (1) HIV, (2) HPV[f], or (3) HSV[g] type 1 or 2 (response options: yes, no, I do not know, and I prefer not to answer).	Multiple choice

[a] FTM: female-to-male.

[b] MTF: male-to-female.

[c] The text following "Note" was presented as clarification for participants completing the survey.

[d] See Table 2 for the grid of sexual behaviors and protective barrier use.

[e] STI: sexually transmitted infection.

[f] HPV: human papillomavirus.

[g] HSV: herpes simplex virus.

Table 2. Transmasculine Sexual Health Assessment: grid of sexual behaviors and frequency of protective barrier use.

	Activity performed with (initials of sexual partner)[a] in the last 12 months?			While engaged in this activity, how often was a barrier used?						
	Yes	No	I prefer not to answer	I did not engage in this activity	Never	Less than half the time	About half the time	More than half the time	Always	I prefer not to answer
Frontal and/or anal penetration with a finger (“fingering”) or fist (“fisting”) - performed										
Frontal and/or anal penetration with a finger (“fingering”) or fist (“fisting”) -received										
Oral-genital – performed										
Oral-genital – received										
Oral-anal (“rimming”) - performed										
Oral-anal (“rimming”) - received										
Frontal receptive - with genitals										
Frontal receptive - with toy or prosthetic										
Frontal insertive - with genitals										
Frontal insertive - with toy or prosthetic										
Anal receptive - with genitals										
Anal receptive - with toy or prosthetic										
Anal insertive - with genitals										
Anal insertive - with toy or prosthetic										
Insertive sex toys or prosthetics shared										

[a] Participants were asked to list the initials for each of their most 3 recent sexual partners in the last 12 months (Table 2). The initials for each sexual partners were prepopulated into the assessment grid for sexual behaviors.

Field Test of the Measure

The TM-SHA was subsequently deployed via a digital tablet in a sample of 141 TM people aged 21 to 50 years in the greater Boston, Massachusetts, area. Participants were recruited via convenience sampling methods (eg, flyers and word-of-mouth referral) and through medical providers. Study visits were held at a community health center that specialized in gender-affirming transgender care in Boston, Massachusetts [26], and lasted for approximately 3 hours. During the study visit, the participants completed a 45-minute cross-sectional survey in which the TM-SHA was embedded. The details of the study protocol have been described elsewhere [24]. After the completion of the TM-SHA measure, the participants were asked to provide optional qualitative, open-ended feedback about their experience with the assessment. The TM-SHA required 20 minutes to complete.

Data Analysis

Descriptive analyses characterized the distribution of TM adults’ sexual behaviors and HIV and STI transmission risk by the gender identity of sexual partners. The gender identity of sexual partners was coded into the following 4 groups: TM (combining transgender man and assigned female gender nonconforming or nonbinary person), transfeminine (combining transgender woman and assigned male gender nonconforming or nonbinary person), cisgender man, and cisgender woman. Descriptive statistics for the sample were calculated using TM adults as individual participants (N=141). This participant-level data set was then transposed from *short* to *long* formatting to reorient the sexual partnership data (n=259 partnerships). As participants provided data for up to 3 sexual partnerships within the past 12 months, the long-formatted data set allowed us to descriptively analyze protective barrier use within the sexual partnership data. Protective barrier use was operationalized as using a condom, internal condom, dental dam, or gloves during sexual activity versus not using the aforementioned protective barriers. Statistical analyses were conducted using SAS (version 9.4; SAS Institute). Themes from the open-ended feedback from the participants on the TM-SHA were summarized qualitatively in Excel (Microsoft Corporation).

Data Exclusion

In total, 9 participants were excluded from the data analysis because they did not report any partnership-level data.

Results

Sample Characteristics

The characteristics of participants (N=141) are presented in Table 3. The majority were aged between 21 and 29 years (102/141, 72.3%), with a mean age of 27 years. Approximately 75.2% (106/141) and 14.9% (21/141) of the participants were White and multiracial, respectively. Among all the participants, approximately 14.9% (21/141) reported a personal history of STI diagnosis.

Table 3. Demographic characteristics of transmasculine adults in field-testing sample (N=141).[a]

Characteristics	Values[b]
Age (years), n (%)	
21-24	45 (31.9)
25-29	57 (40.4)
30-34	25 (17.7)
≥35	14 (9.9)
Age (years), mean (SD)	27.4 (5.7)
Race, n (%)	
White	106 (75.2)
Black	4 (2.8)
Asian	8 (5.7)
Native Hawaiian or other Pacific Islander	1 (0.1)
Multiracial	21 (14.9)
Missing	1 (0.71)
Gender identity, n (%)	
Man or male	38 (26.9)
Transgender man (Female-to-Male)	71 (50.4)
Genderqueer or nonbinary	27 (19.2)
Another gender[a]	5 (3.5)
Sexual orientation, n (%)	
Gay, homosexual, or same-gender attraction	14 (9.9)
Straight or heterosexual	16 (11.4)
Bisexual	18 (12.8)
Queer	65 (46.1)
Pansexual	13 (9.2)
Asexual	3 (2.2)
Questioning or unsure	3 (2.1)
I do not label my sexual orientation	7 (5.0)
Missing	2 (1.4)
Lifetime STI[c] diagnosis[d], n (%)	
No	120 (85.1)
Yes	21 (14.9)
Number of sexual partners within the past 12 months, median (IQR)	2 (3)

[a] Includes agender, bigender, and write-in gender identities distinguishable from the categories provided.

[b] Percentages may not sum to 100% owing to rounding.

[c] STI: sexually transmitted infection.

[d] Includes any lifetime diagnosis of HIV, chlamydia, trichomoniasis, syphilis, gonorrhea, genital herpes, hepatitis B, hepatitis C, or another STI.

Sexual Partnership Characteristics

In Table 4, descriptive statistics of sexual partnerships (n=259) and practices (15 sexual risk behaviors) are shown, along with the frequency of protective barrier use by sexual behaviors. The participants engaged in a wide range of sexual behaviors, including fingering or fisting (receiving: 170/259, 65.6% and performing: 173/259, 66.8%), oral-genital sex (receiving: 182/259, 70.3% and performing: 216/259, 83.4%), anal-genital sex (receptive: 31/259, 11.9% and insertive: 9/259, 3.5%), frontal-genital sex (receptive: 105/259, 40.5% and insertive: 46/259, 17.8%), and sharing toys or prosthetics during insertive sex (62/259, 23.9%). The overall barrier use for each sexual behavior varied greatly, ranging from 10.9% (20/182) to 81% (25/31).

The prevalence of protective barrier use was heterogeneous across different partnerships and sexual behaviors. For example, frontal receptive sex with genitals with no protective barrier ranged from 2% (1/42; with a transfeminine person) to 50% (21/42; with a cisgender man). In our field test of TM-SHA, we found that 40% (42/105) of participants engaged in frontal receptive sex with genitals without the use of a protective barrier. Of those, more than half (22/42, 52%) engaged in this behavior with a transfeminine person or a cisgender man, which may confer risk for HPV and other STIs. One-third (14/42, 33%) of TM participants reported engaging in frontal receptive sex with genitals and no protective barrier with a cisgender woman partner, which is notable given that HPV is also spread by people without a penis. Notably, TM people engaged in insertive frontal and insertive anal sex, in which an enlarged clitoris (often achieved through hormone use) is inserted into a partner's frontal or anal cavity.

Table 4. Characterizing the sexual partners, sexual behaviors, and protective barrier use of transmasculine adults (n=259 partnerships) by gender identity of sexual partners and protective barrier use across varied sexual behaviors.

Sexual behavior	No protective barrier used, n (%)					Protective barrier used, n (%)					Total[a], n (%)
	Values[b]	Gender of sexual partners				Values[b]	Gender of sexual partners				
		Transgender masculine[c,d]	Transgender feminine[d,e]	Cisgender man[d]	Cisgender woman[d]		Transgender masculine[c,d]	Transgender feminine[d,e]	Cisgender man[d]	Cisgender woman[d]	
Performing frontal and/or anal penetration with a finger or fist	132 (76.3)	21 (16)	6 (5)	20 (15)	85 (64)	41 (24)	16 (39)	4 (10)	3 (7)	18 (44)	173 (66.8)
Receiving frontal and/or anal penetration with a finger or fist	134 (78.8)	19 (14)	11 (8)	51 (38)	53 (40)	36 (21)	14 (39)	2 (6)	4 (11)	16 (44)	170 (65.6)
Performing oral-genital sex	184 (85.2)	30 (16)	11 (6)	60 (33)	83 (45)	32 (15)	1 (3)	8 (250)	13 (41)	10 (31)	216 (83.4)
Receiving oral-genital sex	162 (89)	28 (17)	11 (7)	50 (31)	73 (45)	20 (11)	4 (20)	4 (20)	1 (5)	11 (55)	182 (70.3)
Performing oral-anal sex	39 (80)	4 (10)	3 (8)	15 (38)	17 (44)	10 (20)	1 (10)	4 (40)	0	5 (50)	39 (15)
Receiving oral-anal sex	34 (77)	8 (24)	1 (3)	15 (44)	10 (29)	10 (23)	3 (30)	2 (20)	0	5 (50)	44 (17)
Frontal insertive sex with genitals	29 (63)	6 (21)	0	0	23 (79)	17 (37)	2 (12)	3 (18)	8 (47)	4 (24)	46 (18)
Frontal receptive sex with genitals	42 (40)	6 (14)	1 (2)	21 (50)	14 (33)	63 (60)	2 (3)	14 (22)	45 (71)	2 (3)	105 (40.5)
Frontal insertive sex with toy or prosthetic	44 (39)	8 (18)	1 (2)	3 (7)	32 (73)	69 (61)	23 (33)	3 (4)	1 (1)	42 (61)	113 (43.6)
Frontal receptive sex with toy or prosthetic	41 (47)	10 (24)	3 (7)	10 (24)	18 (44)	46 (53)	13 (28)	5 (11)	4 (9)	24 (52)	87 (34)
Anal insertive sex with genitals	2 (22)	0	0	1 (50)	1 (50)	7 (78)	2 (29)	1 (14)	2 (29)	2 (29)	9 (3)
Anal receptive sex with genitals	6 (19)	0	1 (17)	5 (83)	0	25 (81)	0	4 (16)	21 (84)	0	31 (12)

Sexual behavior	No protective barrier used, n (%)					Protective barrier used, n (%)					Total[a], n (%)
	Values[b]	Gender of sexual partners				Values[b]	Gender of sexual partners				
		Transgender masculine[c,d]	Transgender feminine[d,e]	Cisgender man[d]	Cisgender woman[d]		Transgender masculine[c,d]	Transgender feminine[d,e]	Cisgender man[d]	Cisgender woman[d]	
Anal insertive sex with toy or prosthetic	17 (33)	1 (6)	6 (35)	7 (41)	3 (18)	34 (67)	8 (24)	7 (21)	10 (29)	9 (26)	51 (20)
Anal receptive sex with toy or prosthetic	11 (31)	3 (27)	2 (18)	4 (36)	2 (18)	25 (69)	8 (32)	3 (12)	3 (12)	11 (44)	36 (14)
Insertive sex in which toys or prosthetics were shared	21 (34)	3 (14)	2 (10)	2 (10)	14 (67)	41 (66)	14 (34)	3 (7)	4 (10)	20 (49)	62 (24)

[a] The denominator for percentages in the column is the total number of sexual partnerships (N=259).

[b] The denominator for percentages in the column is the total N for that sexual behavior.

[c] Includes transgender men and gender nonbinary assigned-female-at-birth partners.

[d] The denominator for percentages in the column is the N for barrier use or no barrier use.

[e] Includes transgender women and gender nonbinary assigned-male-at-birth partners.

Lessons Learned From Participants' Open-ended Feedback

Through open-ended feedback in the development and field-testing of TM-SHA, the team learned many lessons for measuring TM people's sexual risk (Textbox 1). The TM-SHA had high levels of acceptability among TM participants. No length-related concerns or participant burden concerns were reported for the measure. Skip logic ensured that participants circumvented any questions that were not relevant to them. Themes that emerged in the open-ended feedback after the completion of the TM-SHA included appreciation expressed by multiple participants for the measure's gender-affirming language and for the comprehensiveness of the assessment and how it asked about diverse behaviors and partnerships. Two participants noted confusion surrounding the terms "receptive" and "insertive" for describing some of the sexual behaviors and suggested further clarification in the future.

Textbox 1. Lessons learned in measuring sexual risk in transmasculine (TM) adults.

- Partner and work with TM community members to ensure gender-affirming, cultural competency, and responsiveness of measures.
- Use introductory language at the beginning to acknowledge, validate, and affirm differences in identity and language use for TM people. Specify terms to be used throughout the survey for body parts, sexual acts, etc, for example, the use of *frontal* sex instead of *vaginal* sex.
- Use simple concrete language, which can be challenging with jargon and sexual health terminology. For TM people, carefully select words and phrases to ensure item clarity, similar interpretation of the item across respondents, response options, and choices (eg, check all, mutually exclusive, and open ended).
- Consider the potential emotional impact that terminology and language may evoke and acknowledge sensitivity and constant evolution of language in TM communities.
- Capture the heterogeneity of sexual identities in TM people, diversity of sexual partnership types and behaviors, and interpersonal contexts.
- Validate all sexual practices TM people engage in, regardless of whether they confer high probability of HIV or sexually transmitted infection transmission risk, by asking about high-, low-, and varying-risk activities.
- Do not assume that TM people are limited to certain types of sex based on their gender identity or based on their anatomy (eg, questions should ask about insertive and receptive sex). For example, given the diversity of bodies, it was important to specify whether a sexual act was performed with or without a prosthetic or toy. Gender of one's sexual partner may not always correspond with their genitals.
- Provide space for open-ended feedback about the assessment and experience of completing the survey. For example, ask participants: Is there anything else you feel to be important that we did not address?

Discussion

Principal Findings

The TM-SHA is one of the first digital sexual health risk measures developed specifically with and exclusively for TM people. It successfully integrated gender-affirming language and branching logic that captures the wide array of sexual behaviors and partnerships in TM people. The TM-SHA was highly acceptable to TM adults, with no participant concerns regarding burdensomeness or length of the measure, and it was found to elicit the sexual behavior information needed to assess HIV or STI transmission risk behaviors. For the parent study's outcome of high-risk HPV, the measure included questions to better assess potential sexual exposures for other STIs besides HIV infection, given that multiple behaviors confer risk for high-risk HPV.

In collaboration with a Task Force and community input, the TM-SHA was adapted with gender-affirming language to measure the sexual behavior and protective barrier use of participants. The use of gender-affirming language may have 2 advantages for measuring the sexual risk behavior of TM people. First, gender-affirming language may improve the validity of the TM-SHA by providing a description of sexual behaviors and genitals in a way that better reflects some of the perspectives of TM people (eg, using *frontal* receptive sex instead of *vaginal* receptive sex) [27]. Second, the gender identity options listed on the TM-SHA reflect those commonly held by TM people and their sexual partners; the instrument served to further acknowledge and validate these gender identities.

The sexual health of TM people remains understudied [2,4], and a barrier to conducting HIV- and STI-related research with this study population is the lack of consensus on how to best capture sexual risk data in a standardized manner that is gender affirming, culturally responsive, and accurate. In the context of widespread stigma facing TM people, there is a need for high-quality measures that are cocreated with communities to maximize acceptability, for example, by avoiding stigmatizing language that may trigger medical and research mistrust and by cueing participants that researchers are knowledgeable and transcompetent in asking sensitive questions. Thus, this study fills an important gap for both researchers and clinicians aiming to capture the wide variety of sexual behaviors of TM people.

The TM-SHA provides a high level of specificity about sexual practices, some of which may confer greater risk for HIV and STIs than others. Additional research is warranted on how to optimize use and interpret the measure. For example, it may be helpful to develop a *scoring* schema to categorize someone for research as *low, moderate, or high* risk based on their reported protection of sexual behaviors or for clinical settings to appropriately counsel a patient on their STI risk or need for Papanicolaou testing. The growing functionality of electronic health records (EHRs) and patient portals offers an opportunity to integrate the TM-SHA and similar tools into clinical care provision. Patients could be asked to complete the TM-SHA before their visit via the patient portal, and the information elicited could be linked to the EHR and integrated into clinical care delivery.

Limitations and Future Directions

With regard to limitations, the current TM-SHA does not consider HIV prevention methods such as pre-exposure prophylaxis (PrEP) or postexposure prophylaxis to prevent HIV acquisition used by participants or their partners. TM people engage in sexual behaviors conferring eligibility for PrEP but have a low uptake of PrEP [28]. Future iterations of the measure may consider adding questions and skip patterns to assess current PrEP and postexposure prophylaxis use for both TM people and their partners. In addition, the current TM-SHA assumes that the participant had a cervix per eligibility criteria for the parent study. Future versions of the measure may consider adding questions and logic to assess gender-affirming surgeries, which may result in different anatomy and, therefore, sexual risks in participants. According to participant feedback, future versions of the measure would benefit from clarifying insertive and receptive sex. For example, in reporting on frontal receptive sex with genitals with a cisgender woman partner, it will be important to learn how people understand or interpret the terms *receptive* and *with genitals*. Cognitive interviewing methods may assist with this process and strengthen the evidence of validity.

The length of the survey and advanced skipped patterns used in the survey, which may be difficult for beginners to code, are best suited for preprogrammed electronic surveys. This may pose a challenge for those with low literacy rates and for medical providers hoping to use this measure in clinic [22]. An audio, digital-assisted self-interview or interviewer-administered assessment may be used to address the low literacy levels of respondents. Researchers and clinical providers may also adapt and shorten the TM-SHA for their research, clinical practice, and patient needs. For example, if HIV risk is the parameter of interest, the assessment could be shortened to only those sexual behaviors that confer a high HIV risk. Future research is needed to reduce the scope of the tool, ensure its accuracy, and externally validate the measure for clinical robustness. Such research could examine clinical outcomes (eg, laboratory or clinically confirmed STIs) alongside patient-reported outcomes and/or gather and compare patient-reported and provider-documented sexual behavior data for accuracy.

Finally, despite efforts to capture all types of sexual behaviors and partnerships, it is possible that not all behaviors were captured by the measure, as currently designed. Thus, open-ended questions at the end of the measure were added. This allows participants the space to disclose any other relevant information, including the behaviors they may want to report, which may be used in data analysis or to improve the measure for future use. A strength of the TM-SHA is that it is inclusive of polyamorous relationships in that for each partnership they report on, participants are asked to describe the type of partnership (eg, casual or polyamorous). In addition, although participants were not asked directly about multipartner sex (eg, group sex), the assessment asked about sexual behaviors with multiple partners. Asking explicitly about multipartner sex is recommended in the future.

Conclusions

This study offers a digital measure designed with and specifically for TM people for future use in TM sexual health research and clinical care. The tool garnered high levels of acceptability from TM participants, was feasible to implement in field-testing, and reported no concerns about participant burden or length in completing the assessment. A strength of this measure is that detailed partner-by-partner data can be used to model partnership-level characteristics in addition to individual-level participant data. Additional psychometric evaluation of the measure is necessary, including longitudinal data, to assess the performance of the TM-SHA over time for research purposes. Clinically, the TM-SHA provides a standardized sexual history measure for TM people, which may help clinicians enhance trust with their patients and more effectively identify sexual behaviors that warrant HIV or STI screening in clinical care. Integration of the measure into clinical research and care via patient portals linked to EHR may facilitate the delivery of culturally responsive and gender-affirming sexual health care.

Acknowledgments

This study was funded by the Patient-Centered Outcomes Research Institute (PCORI; contract CER-1403-12625). All methods were reviewed by the PCORI Scientific Committee and found to be in accordance with their methodology standards. PCORI was not involved in the collection, analysis, or interpretation of study data. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors would like to thank the Community Task Force, Scientific Advisory Board, and the research team who came together and collaborated on this study to ensure transmasculine people's engagement and participation. The authors also thank the transmasculine people who field-tested the measure and provided input.

Data Availability

The data sets analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

SLR, MBD, SMP, JMWH, and JP conceptualized the study. SLR, DRP, JMWH, and SMP curated data. SLR and DRP conducted the formal analysis. SLR, MBD, SMP, and JP participated in acquiring funding. SLR, MBD, JMWH, SMP, and JP were study investigators. SLR, MBD, JMWH, SMP, and JP contributed to the study methodology. SLR and DJP conducted project administration. SLR and DJP participated in managing resources. SLR and DRP obtained software. SLR, MBD, DJP, and JP provided supervision for the study. SLR, JMWH, and JP validated study findings. SLR and DRP conducted data visualization. SLR, DRP, DJP, and MQ wrote the initial draft manuscript. SLR, DRP, DJP, MBD, JMWH, SMP, MQ, and JP reviewed, edited, and provided critical feedback on the final manuscript.

Conflicts of Interest

None declared.

References

1. LGBTQIA+ Glossary of Terms for Health Care Teams. Health Resources and Services Administration (HRSA). URL: https://www.lgbtqiahealtheducation.org/wp-content/uploads/2020/02/Glossary-2022.02.22-1.pdf [accessed 2023-02-13]
2. Herbst JH, Jacobs ED, Finlayson TJ, McKleroy VS, Neumann MS, Crepaz N, HIV/AIDS Prevention Research Synthesis Team. Estimating HIV prevalence and risk behaviors of transgender persons in the United States: a systematic review. AIDS Behav 2008 Jan;12(1):1-17. [doi: 10.1007/s10461-007-9299-3] [Medline: 17694429]
3. Injustice at every turn: a report of the national transgender discrimination survey. National LGBTQ Task Force Action Fund. URL: https://www.thetaskforce.org/injustice-every-turn-report-national-transgender-discrimination-survey/ [accessed 2023-03-13]
4. Van Gerwen OT, Jani A, Long DM, Austin EL, Musgrove K, Muzny CA. Prevalence of sexually transmitted infections and human immunodeficiency virus in transgender persons: a systematic review. Transgend Health 2020 Jun 01;5(2):90-103 [FREE Full text] [doi: 10.1089/trgh.2019.0053] [Medline: 32656353]
5. Reisner SL, Pardo ST, Gamarel KE, White Hughto JM, Pardee DJ, Keo-Meier CL. Substance use to cope with stigma in healthcare among U.S. Female-to-male trans masculine adults. LGBT Health 2015 Dec;2(4):324-332 [FREE Full text] [doi: 10.1089/lgbt.2015.0001] [Medline: 26788773]
6. James S, Herman J, Rankin S, Keisling M, Mottet L, Anafi M. The Report of the 2015 U.S. Transgender Survey. National Center for Transgender Equality. 2016. URL: https://transequality.org/sites/default/files/docs/usts/USTS-Full-Report-Dec17.pdf [accessed 2023-03-13]
7. Djordjevic ML, Stojanovic B, Bizic M. Metoidioplasty: techniques and outcomes. Transl Androl Urol 2019 Jun;8(3):248-253 [FREE Full text] [doi: 10.21037/tau.2019.06.12] [Medline: 31380231]

8. Reisner SL, Deutsch MB, Peitzmeier SM, White Hughto JM, Cavanaugh TP, Pardee DJ, et al. Test performance and acceptability of self- versus provider-collected swabs for high-risk HPV DNA testing in female-to-male trans masculine patients. PLoS One 2018;13(3):e0190172 [FREE Full text] [doi: 10.1371/journal.pone.0190172] [Medline: 29538411]
9. Sargent A, Bailey A, Almonte M, Turner A, Thomson C, Peto J, ARTISTIC Study Group. Prevalence of type-specific HPV infection by age and grade of cervical cytology: data from the ARTISTIC trial. Br J Cancer 2008 May 20;98(10):1704-1709 [FREE Full text] [doi: 10.1038/sj.bjc.6604324] [Medline: 18392052]
10. Dunne EF, Unger ER, Sternberg M, McQuillan G, Swan DC, Patel SS, et al. Prevalence of HPV infection among females in the United States. JAMA 2007 Feb 28;297(8):813-819. [doi: 10.1001/jama.297.8.813] [Medline: 17327523]
11. Chesson HW, Dunne EF, Hariri S, Markowitz LE. The estimated lifetime probability of acquiring human papillomavirus in the United States. Sex Transm Dis 2014 Nov;41(11):660-664 [FREE Full text] [doi: 10.1097/OLQ.0000000000000193] [Medline: 25299412]
12. Hariri S, Unger ER, Sternberg M, Dunne EF, Swan D, Patel S, et al. Prevalence of genital human papillomavirus among females in the United States, the National Health And Nutrition Examination Survey, 2003-2006. J Infect Dis 2011 Aug 15;204(4):566-573. [doi: 10.1093/infdis/jir341] [Medline: 21791659]
13. Dhillon N, Oliffe JL, Kelly MT, Krist J. Bridging barriers to cervical cancer screening in transgender men: a scoping review. Am J Mens Health 2020;14(3):1557988320925691 [FREE Full text] [doi: 10.1177/1557988320925691] [Medline: 32489142]
14. Donenberg GR, Emerson E, Bryant FB, Wilson H, Weber-Shifrin E. Understanding AIDS-risk behavior among adolescents in psychiatric care: links to psychopathology and peer relationships. J Am Acad Child Adolesc Psychiatry 2001 Jun;40(6):642-653 [FREE Full text] [doi: 10.1097/00004583-200106000-00008] [Medline: 11392341]
15. Navaline HA, Snider EC, Petro CJ, Tobin D, Metzger D, Alterman AI, et al. Preparations for AIDS vaccine trials. An automated version of the Risk Assessment Battery (RAB): enhancing the assessment of risk behaviors. AIDS Res Hum Retroviruses 1994;10 Suppl 2:S281-S283. [Medline: 7865319]
16. McLemore KA. A minority stress perspective on transgender individuals' experiences with misgendering. Stigma and Health 2018 Feb;3(1):53-64. [doi: 10.1037/sah0000070]
17. No Data No More. AIDS Vaccine Advocacy Coalition (AVAC). URL: https://www.avac.org/no-data-no-more [accessed 2023-03-13]
18. No Data No More Press Conference. YouTube. 2021. URL: https://www.youtube.com/watch?v=k4PEzndtrBM [accessed 2023-03-13]
19. Maheux AJ, Zhou Y, Thoma BC, Salk RH, Choukas-Bradley S. Examining sexual behavior among U.S. Transgender adolescents. J Sex Res 2021 Oct;58(8):1050-1060 [FREE Full text] [doi: 10.1080/00224499.2020.1855408] [Medline: 33356600]
20. Moseson H, Lunn MR, Katz A, Fix L, Durden M, Stoeffler A, et al. Development of an affirming and customizable electronic survey of sexual and reproductive health experiences for transgender and gender nonbinary people. PLoS One 2020 May 4;15(5):e0232154 [FREE Full text] [doi: 10.1371/journal.pone.0232154] [Medline: 32365110]
21. Bauer GR, Redman N, Bradley K, Scheim AI. Sexual health of trans men who are gay, bisexual, or who have sex with men: results from Ontario, Canada. Int J Transgend 2013 Apr;14(2):66-74 [FREE Full text] [doi: 10.1080/15532739.2013.791650] [Medline: 24971043]
22. Schroder KE, Carey MP, Vanable PA. Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. Ann Behav Med 2003 Oct;26(2):104-123 [FREE Full text] [doi: 10.1207/S15324796ABM2602_03] [Medline: 14534028]
23. Johnson AM, Copas AJ, Erens B, Mandalia S, Fenton K, Korovessis C, et al. Effect of computer-assisted self-interviews on reporting of sexual HIV risk behaviours in a general population sample: a methodological experiment. AIDS 2001 Jan 05;15(1):111-115. [doi: 10.1097/00002030-200101050-00016] [Medline: 11192852]
24. Reisner SL, Deutsch MB, Peitzmeier SM, White Hughto JM, Cavanaugh T, Pardee DJ, et al. Comparing self- and provider-collected swabbing for HPV DNA testing in female-to-male transgender adult patients: a mixed-methods biobehavioral study protocol. BMC Infect Dis 2017 Jun 23;17(1):444 [FREE Full text] [doi: 10.1186/s12879-017-2539-x] [Medline: 28645254]
25. Reisner SL, Hughto JM, Pardee DJ, Kuhns L, Garofalo R, Mimiaga MJ. LifeSkills for men (LS4M): pilot evaluation of a gender-affirmative HIV and STI prevention intervention for young adult transgender men who have sex with men. J Urban Health 2016 Feb;93(1):189-205 [FREE Full text] [doi: 10.1007/s11524-015-0011-z] [Medline: 26753882]
26. Reisner SL, Bradford J, Hopwood R, Gonzalez A, Makadon H, Todisco D, et al. Comprehensive transgender healthcare: the gender affirming clinical and public health model of Fenway Health. J Urban Health 2015 Jun;92(3):584-592 [FREE Full text] [doi: 10.1007/s11524-015-9947-2] [Medline: 25779756]
27. Ragosta S, Obedin-Maliver J, Fix L, Stoeffler A, Hastings J, Capriotti MR, et al. From 'shark-week' to 'Mangina': an analysis of words used by people of marginalized sexual orientations and/or gender identities to replace common sexual and reproductive health terms. Health Equity 2021 Sep 01;5(1):707-717 [FREE Full text] [doi: 10.1089/heq.2021.0022] [Medline: 34909540]

28. Golub SA, Fikslin RA, Starbuck L, Klein A. High rates of PrEP eligibility but low rates of PrEP access among a national sample of transmasculine individuals. J Acquir Immune Defic Syndr 2019 Sep 01;82(1):e1-e7 [FREE Full text] [doi: 10.1097/QAI.0000000000002116] [Medline: 31232834]

Abbreviations

EHR: electronic health record
HPV: human papillomavirus
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection
TM: transmasculine
TM-SHA: Transmasculine Sexual Health Assessment

Edited by A Mavragani; submitted 23.06.22; peer-reviewed by S Sarejloo, P Serrano, E Davidovich, U Belkind; comments to author 09.12.22; revised version received 29.12.22; accepted 12.01.23; published 17.03.23.

Please cite as:
Reisner SL, Pletta DR, Pardee DJ, Deutsch MB, Peitzmeier SM, Hughto JMW, Quint M, Potter J
Digital-Assisted Self-interview of HIV or Sexually Transmitted Infection Risk Behaviors in Transmasculine Adults: Development and Field Testing of the Transmasculine Sexual Health Assessment
JMIR Public Health Surveill 2023;9:e40503
URL: https://publichealth.jmir.org/2023/1/e40503
doi:10.2196/40503
PMID:36930204

Original Paper

Peer-Led Community-Based Support Services and HIV Treatment Outcomes Among People Living With HIV in Wuxi, China: Propensity Score–Matched Analysis of Surveillance Data From 2006 to 2021

Xiaojun Meng[1*], MSc; Hanlu Yin[1*], MSc; Wenjuan Ma[1*], MSc; Jing Gu[1*], MSc; Zhen Lu[2], MSc; Thomas Fitzpatrick[3], MD; Huachun Zou[2,4], PhD

[1]Wuxi Municipal Center for Disease Control and Prevention, Wuxi, China
[2]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China
[3]Division of Allergy and Infectious Diseases, University of Washington, Seattle, WA, United States
[4]Kirby Institute, University of New South Wales, Sydney, Australia
[*]these authors contributed equally

Corresponding Author:
Huachun Zou, PhD
School of Public Health (Shenzhen), Sun Yat-sen University
No. 66, Gongchang Road, Guangming District
Shenzhen, 518107
China
Phone: 86 20 87335651
Email: zouhuachun@mail.sysu.edu.cn

Abstract

Background: Community-based organizations deliver peer-led support services to people living with HIV. Systematic reviews have found that peer-led community-based support services can improve HIV treatment outcomes; however, few studies have been implemented to evaluate its impact on mortality using long-term follow-up data.

Objective: We aimed to evaluate the associations between the receipt of peer-led community-based support services and HIV treatment outcomes and survival among people living with HIV in Wuxi, China.

Methods: We performed a propensity score–matched retrospective cohort study using data collected from the Chinese National HIV/AIDS Comprehensive Information Management System for people living with HIV in Wuxi, China, between 2006 and 2021. People living with HIV who received adjunctive peer-led community-based support for at least 6 months from a local community-based organization (exposure group) were matched to people living with HIV who only received routine clinic-based HIV care (control group). We compared the differences in HIV treatment outcomes and survival between these 2 groups using Kaplan-Meier curves. We used competing risk and Cox proportional hazards models to assess correlates of AIDS-related mortality (ARM) and all-cause mortality. We reported adjusted subdistribution hazard ratio and adjusted hazard ratio with 95% CIs.

Results: A total of 860 people living with HIV were included (430 in the exposure group and 430 in the control group). The exposure group was more likely to adhere to antiretroviral therapy (ART; 396/430, 92.1% vs 360/430, 83.7%; *P*<.001), remain retained in care 12 months after ART initiation (402/430, 93.5% vs 327/430, 76.1%; *P*<.001), and achieve viral suppression 9 to 24 months after ART initiation (357/381, 93.7% vs 217/243, 89.3%; *P*=.048) than the control group. The exposure group had significantly lower ARM (1.8 vs 7.0 per 1000 person-years; *P*=.01) and all-cause mortality (2.3 vs 9.3 per 1000 person-years; *P*=.002) and significantly higher cumulative survival rates (*P*=.003). The exposure group had a 72% reduction in ARM (adjusted subdistribution hazard ratio 0.28, 95% CI 0.09-0.95) and a 70% reduction in all-cause mortality (adjusted hazard ratio 0.30, 95% CI 0.11-0.82). The nonrandomized retrospective nature of our analysis prevents us from determining whether peer-led community-based support caused the observed differences in HIV treatment outcomes and survival between the exposure and control groups.

Conclusions: The receipt of peer-led community-based support services correlated with significantly improved HIV treatment outcomes and survival among people living with HIV in a middle-income country in Asia. The 15-year follow-up period in this study allowed us to identify associations with survival not previously reported in the literature. Future interventional trials are needed to confirm these findings.

(JMIR Public Health Surveill 2023;9:e43635) doi:10.2196/43635

KEYWORDS

people living with HIV; community health workers; community-based organizations; peer-led services; HIV treatment outcomes

Introduction

Survival Improved Among People Living With HIV

According to the 2020 global AIDS report, 38 million people are living with HIV globally, of whom 25.4 million are receiving antiretroviral therapy (ART) [1]. If the *95-95-95* targets updated by the Joint United Nations Program on HIV/AIDS in December 2020 are to be achieved by 2025 [2], nearly 86% of all people living with HIV worldwide will need to be initiated on ART and achieve viral suppression. China has achieved considerable success in optimizing the HIV care continuum in recent decades. Currently, an estimated 83.4% of all people living with HIV in China are receiving ART, of whom 94.2% have achieved viral suppression [3].

The overall life expectancy of people living with HIV with good adherence to ART is approaching that of the general population in many high-income countries [4,5]; however, significant differences in survival persist in resource-limited settings [6]. The use of ART to improve survival among people living with HIV depends on early initiation of treatment, adherence to ART, and retention in care [7-9]. AIDS-related mortality (ARM) remains high in certain contexts, particularly because of advanced immunosuppression at the time of ART initiation [10,11]. In addition, an increasing proportion of deaths among people living with HIV are attributed to non-ARM (NARM) [12]. A recent global systematic review showed many people living with HIV died of cardiovascular disease, non-AIDS malignancies, and liver disease [13]. To improve health outcomes among people living with HIV, effective interventions are needed to reduce both ARM and NARM.

Community Health Workers Promote the Quality of HIV Care Among People Living With HIV

Task shifting involves the rational redistribution of tasks from the professional health workforce to community-based organizations (CBOs) and community health workers (CHWs) and is recognized by the World Health Organization (WHO) as an important strategy to optimize HIV care [14]. CHWs can play a valuable role in engaging and retaining people living with HIV in care [15,16]. CHWs strengthen the quality of HIV services as well as promote dignity and improve the quality of life for people living with HIV enrolled in those services [14]. Trials in several low-resource settings found CHWs can reduce the rates of virologic failure on ART and loss to follow-up [16-18]. A cluster randomized controlled trial in Zimbabwe found using CHWs to implement a community-based differentiated service delivery model that adapted HIV services to the preferences and expectations of adolescents significantly improved HIV viral suppression [19]. The WHO recommends community-based interventions to support ART adherence and retention in care while acknowledging that the evidence to support this recommendation is weak. Most previous evaluations of peer-led community-based interventions for people living with HIV have been conducted in high-income countries or sub-Saharan Africa, and few have reported survival as an outcome.

CBOs and CHWs have played a critical role in optimizing the HIV care continuum in China, especially in efforts to increase access to ART and improve retention in care [20]. The Chinese government has implemented a nationwide task-shifting strategy for HIV services, supporting the scale-up of CHWs through the China AIDS Fund for CBOs. Few previous studies have evaluated the effect of peer support delivered by CHWs on HIV treatment outcomes in China. We performed a propensity score–matched analysis of longitudinal data extracted from a database of HIV treatment outcomes in China to evaluate the impact of a multicomponent peer support model on ART adherence, viral suppression, and survival among people living with HIV in Wuxi, China.

Methods

Study Design

Using routine surveillance records on people living with HIV, we conducted a propensity score–matched cohort study of individuals who were served by local CBOs on HIV care. Convenient sampling method was used to enroll people living with HIV who had received peer-led community-based support services from local CBOs in the exposure group. In the meantime, we used propensity score matching (PSM) to enroll a comparable control group from other people living with HIV who had never received such a service.

Study Setting

Wuxi is a city of 8.5 million people in China’s Jiangsu Province. As of December 2021, a total of 4111 people living with HIV were registered as living in Wuxi. Rainbow Family is currently the only CBO registered with the local government to provide services for people living with HIV in Wuxi. This CBO has been providing support services to people living with HIV since June 2006.

Standard of Care for People Living With HIV in China

In 2003, China implemented free ART for people living with HIV [21]. Under this policy, all people living with HIV in China can obtain free HIV care and ART from health care providers at infectious diseases hospitals or clinics affiliated with local centers for disease control and prevention (CDCs). National guidelines suggest routine follow-up services should be delivered at 0.5, 1, 2, 3, 6, 9, and 12 months after ART initiation to monitor for adverse drug effects and adjust ART as necessary.

At 13 months after ART initiation, follow-up visits are recommended every 3 months, with CD4 and viral load testing performed at least every 12 months [22].

Peer-Led Community-Based Support Services Provided by Rainbow Family

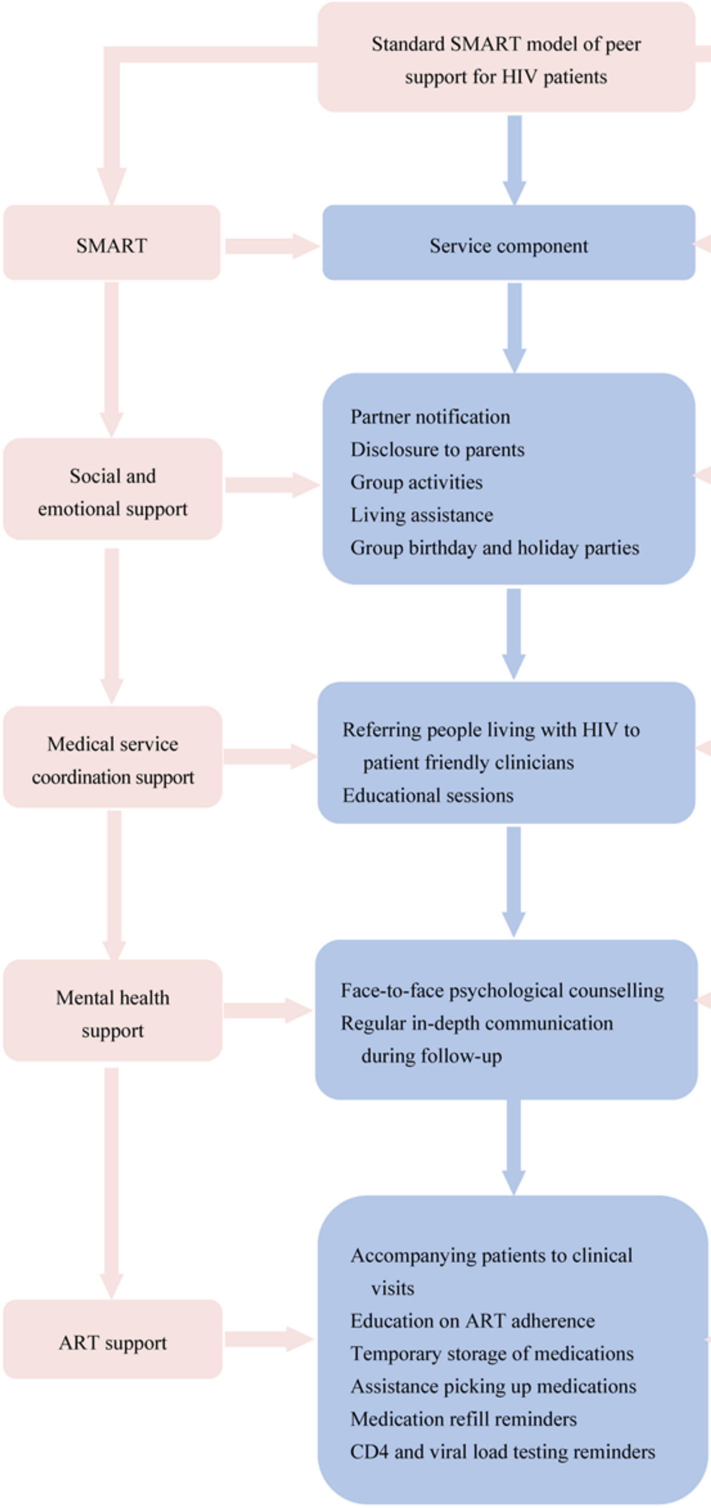
In Wuxi, all newly diagnosed people living with HIV are informed of their diagnosis and receive their initial HIV care visit at a local CDC clinic or infectious diseases hospital. At this initial visit, health care providers will introduce people living with HIV to the peer-led community-based support services offered at Rainbow Family and provide a referral if desired. A CHW from Rainbow Family will be assigned to any referred person living with HIV in Wuxi. Not all newly diagnosed people living with HIV opt to use the peer support services provided by CBOs. Factors influencing decision-making include, but are not limited to, self-acceptance of HIV status, willingness to expose their sexual orientation to others in the CBOs, and willingness to be followed up by peer CHWs in daily life.

CHWs at Rainbow Family are people living with HIV themselves and recruited as volunteers working for CBOs in their spare time who have documented adherence to ART for at least 12 months and demonstrated strong interpersonal and communication skills. Before providing services to clients, CHWs at Rainbow Family receive education on HIV pathogenesis, prevention, and treatment, as well as training on how to provide psychological support and referrals to other medical services. CHW trainings are led by professional health care workers, including public health practitioners from the Wuxi municipal CDC, HIV doctors and nurses, mental health counselors, and experienced peer educators. After passing the professional training, the CHWs will be assigned clients who are people living with HIV and initiate the peer-led community-based support according to guidelines.

The peer-led community-based support model provided by Rainbow Family was designed as a patient-centered peer support model. CHWs build trust and rapport with people living with HIV through regular face-to-face communication, contact through WeChat (an instant messaging software widely used in China), telephone calls, home visits, and group activities. These regular interactions allow CHWs to provide health care services and support to people living with HIV that are tailored to the individual needs of clients.

A social, medical, and mental ART support (SMART) model is used as a conceptual framework to guide this service (Figure 1). This model emphasizes four forms of support: (1) social and emotional support to navigate relationships with sex partners, spouses, family members, and friends, particularly assisting people living with HIV in coping with the emotional stress associated with partner notification and disclosure of HIV status; (2) medical service coordination support to overcome HIV-related stigma and discrimination from health care providers by referring people living with HIV to clinicians with experience providing nondiscriminatory care; (3) mental health support to connect people living with HIV with in-person psychological counseling to address concurrent anxiety, depression, or other mental health disorders; and (4) ART support to provide people living with HIV with additional education emphasizing the importance of ART adherence, accompanying patients to clinic visits to help them navigate ART initiation and maintenance, and providing reminders regarding medication refills and laboratory testing to overcome treatment fatigue.

Figure 1. The social, medical, and mental antiretroviral therapy (SMART) model of community-based peer-led support for people living with HIV in Wuxi, China. ART: antiretroviral therapy.



Data Collection

In accordance with national HIV treatment guidelines, HIV diagnoses are reported via the Chinese National HIV/AIDS Comprehensive Information Management System (CNHCIMS) within 24 hours. Information on sociodemographic characteristics, risk behaviors, CD4 count, HIV viral load, and mortality is recorded in the CNHCIMS. We extracted case report data from the CNHCIMS on sociodemographic characteristics and risk factors for HIV transmission for all people living with HIV in Wuxi between January 1, 2006, and December 31, 2019. We also extracted follow-up data detailing care retention, ART adherence, CD4 and viral load monitoring, mortality, and the cause of death for the same people living with HIV between January 1, 2006, and December 31, 2021.

Exposure Group

Rainbow Family collects and records the name and government-issued ID card number of all people living with HIV who have been assigned a CHW and received peer-led community-based support services. We obtained a list of people living with HIV who ever received peer-led community-based support from Rainbow Family, and persons on this list were linked to the CNHCIMS data by their unique government-issued ID card number. People living with HIV were assigned to the exposure group if they met the following inclusion criteria: (1) age ≥18 years at HIV diagnosis; (2) initiated ART on and before December 31, 2019; and (3) received peer-led community-based support from Rainbow Family for at least 6 months.

Control Group

People living with HIV in Wuxi identified in the CNHCIMS who had not received peer-led community-based support from Rainbow Family were eligible to be assigned to the control group in our analysis. To adjust for potential confounders, PSM was used to match people living with HIV in the exposure group with controls. Sex, age at diagnosis, ethnicity, education, marital status, occupation, the route of HIV infection, and baseline CD4 count are factors found to be significantly associated with survival among people living with HIV in previous studies and therefore were included in our PSM model [12,23,24]. In addition, before June 2016, people living with HIV in China were only eligible to initiate ART if their CD4 count was <500 cells/μL; therefore, the year of HIV diagnosis was also included in our PSM model.

When implementing PSM, propensity scores were generated using logistic regression, with the receipt versus nonreceipt of peer-led community-based support as the dependent variable. Propensity scores were generated for each of the imputed data sets, and an average propensity score for each observation was calculated across data sets. Each patient who received peer-led community-based support services (exposure group) was matched 1:1 with a patient who received routine support (control group) according to their propensity scores, using the nearest neighbor matching method without replacement.

Study Variables

Two sets of variables were extracted from the CNHCIMS. The first set was variables in HIV case report data set evaluated at diagnosis, which involved sex (male and female), age at diagnosis (continuous), ethnicity (Han and others), education (primary or less, secondary, and tertiary), marital status (unmarried, married, and divorced or widowed), occupation (student and others), the route of HIV infection (men who have sex with men and others including heterosexual behavior, drug use, and blood donation), the date of HIV diagnosis, and baseline CD4 count (cells/μL). All these variables were categorized into 2 or 3 subgroups where necessary and applicable. The other set was variables in the follow-up data set recorded after HIV diagnosis, in which drug collection, ART interruption, and scheduled follow-up visit were used to measure adherence to ART, and CD4 and viral load were used to measure retention in care.

Outcomes

All data on study outcomes were extracted from the CNHCIMS, and all outcomes were assessed using these extracted data. Adherence to ART was defined as never having a documented period of ART interruption and never having a period when a scheduled follow-up clinic visit was missed >3 months. Retention in care was defined as having received either CD4 or viral load monitoring on at least 1 occasion ≥12 months after ART initiation. People living with HIV who were on ART for at least 6 months and had a documented HIV viral load ≤1000 copies per mL were considered to be virally suppressed [25].

Information on the cause of death was obtained from diagnosis codes in the CNHCIMS. Causes of death were classified into categories of ARM, NARM, and all-cause mortality according to the coding causes of death in HIV Project protocol [26]. A death was classified as AIDS-related if the cause of death was an AIDS-defining condition. All other deaths were classified as non–AIDS-related. Non–AIDS-related deaths were grouped into several subcategories based on diagnosis codes: non–AIDS-related infections, malignancy, liver disease, cardiovascular disease, respiratory disease, accident and suicide and overdose, and other. ARM, NARM, and all-cause mortality were calculated using the number of AIDS-related, non–AIDS-related, and all deaths and the corresponding duration of follow-up, respectively.

Statistical Analysis

Descriptive statistics were used to summarize baseline demographic characteristics and HIV treatment outcomes for both the exposure and control groups. Differences between groups were compared using the Pearson chi-square test. Person-years (PY) was used to estimate ARM, NARM, and all-cause mortality. The Kaplan-Meier method was used to generate survival curves, and the log rank test was conducted to assess differences in survival time between the exposure and control groups. Competing risk models were used to identify factors potentially associated with ARM. Subdistribution hazard ratios were reported with corresponding 95% CIs. Cox proportional hazards models were used to assess correlates of all-cause mortality with reporting hazard ratios and its 95% CIs. Sex, age at diagnosis, ethnicity, marital status, education, occupation, the mode of transmission, the date of HIV diagnosis, baseline CD4 count, the receipt of peer-led community-based support, adherence to ART, and retention in care were included in the adjusted analysis for correlates of mortality and mortality, in which variables with a *P* value <.05 were retained in the final model. Statistical analyses were carried out using Stata (version 15.0; Stata Corp), and survival curves were completed using R software (version 3.6.0; R Core Team).

Ethics Approval

This study was approved by the Ethics Review Committee of Wuxi Municipal Centre for Disease Control and Prevention (WXCDC2022014). A unique ID number was used to protect each participant's privacy.

Results

Participant Characteristics

Data describing a total of 2794 people living with HIV were retrieved from the CNHCIMS between 2006 and 2019, among whom 430 met the criteria to be assigned to the exposure group. There were significant differences in baseline characteristics, including sex (*P*=.005), age at diagnosis (*P*<.001), education (*P*<.001), marital status (*P*<.001), occupation (*P*<.001), and baseline CD4 count (*P*=.001), comparing the 430 and 2364 people living with HIV who did and did not meet the inclusion criteria to be assigned to the exposure group, respectively. Using PSM, 430 people living with HIV were assigned to the control group. After PSM, people living with HIV in the exposure and control groups were comparable in all recorded baseline characteristics (Table 1).

Table 1. Baseline characteristics of people living with HIV assigned to the exposure and control groups before and after propensity score matching (PSM).

Variables	Before PSM			After PSM		
	Exposure group (n=430), n (%)	Other people living with HIV (n=2364), n (%)	*P* value	Exposure group (n=430), n (%)	Control group (n=430), n (%)	*P* value
Sex			.005			.77
Male	365 (84.9)	2117 (89.6)		365 (84.9)	368 (85.6)	
Female	65 (15.1)	247 (10.5)		65 (15.1)	62 (14.4)	
Age (years) at diagnosis			<.001			.89
<30	196 (45.6)	741 (31.4)		196 (45.6)	198 (46.1)	
≥30	234 (54.4)	1623 (68.6)		234 (54.4)	232 (54)	
Ethnicity			.61			.73
Han ethnicity	425 (98.8)	2329 (98.5)		425 (98.8)	426 (99.1)	
Others	5 (1.2)	35 (1.5)		5 (1.2)	4 (0.9)	
Education			<.001			.27
Primary or less	16 (3.7)	280 (11.8)		16 (3.7)	23 (5.4)	
Secondary	234 (54.4)	1404 (59.4)		234 (54.4)	214 (49.7)	
Tertiary	180 (41.9)	680 (28.8)		180 (41.9)	193 (44.9)	
Marital status			<.001			.09
Unmarried	210 (48.8)	859 (36.3)		210 (48.8)	228 (53)	
Married	147 (34.2)	790 (33.4)		147 (34.2)	118 (27.5)	
Divorced and widowed	73 (17)	715 (30.3)		73 (17)	84 (19.5)	
Occupation			<.001			.90
Student	41 (9.5)	69 (2.9)		41 (9.5)	42 (9.8)	
Others	389 (90.5)	2295 (97.1)		389 (90.5)	388 (90.2)	
Route of HIV infection			.23			.94
MSM[a]	262 (60.9)	1511 (63.9)		262 (60.9)	261 (60.7)	
Others	168 (39.1)	853 (36.1)		168 (39.1)	169 (39.3)	
Date of HIV diagnosis[b]			.50			.62
Before June 2016	275 (64)	1472 (62.3)		275 (64)	268 (62.3)	
After June 2016	155 (36.1)	892 (37.7)		155 (36.1)	162 (37.7)	
Baseline CD4 count (cells/µL)			.001			.76
<200	126 (29.3)	899 (38)		126 (29.3)	122 (28.4)	
≥200	304 (70.7)	1465 (62)		304 (70.7)	308 (71.6)	

[a] MSM: men who have sex with men.

[b] Since June 2016, all people living with HIV in China have been encouraged to initiate antiretroviral therapy regardless of CD4 cell count.

Among all 860 people living with HIV included in our analysis, 733 (85.2%) were men, 394 (45.8%) were diagnosed with HIV aged <30 years, 851 (99%) were of Han ethnicity, 373 (43.4%) had completed high school or had a higher level of educational achievement, and 83 (9.7%) were current students. More than half were unmarried (438/860, 50.9%) and were men who have sex with men (523/860, 60.8%). More than one-third (317/860, 36.9%) were diagnosed after June 2016, and just under one-third (248/860, 28.8%) had a CD4 count <200 cells/µL at the time of HIV diagnosis.

ART Adherence and Retention in Care

Compared with the control group, people living with HIV in the exposure group were more likely to adhere to ART (396/430, 92.1% vs 360/430, 83.7%; χ_1^2=14.2; *P*<.001) and remain retained in care (402/430, 93.5% vs 327/430, 76.1%; χ_1^2=50.7; *P*<.001) 12 months after ART initiation (Table 2). The proportion of people living with HIV who received viral load monitoring increased over time for both the exposure group (from 88/430, 20.5% at 9-12 months after HIV diagnosis to 381/430, 88.6% at 9-24 months after HIV diagnosis; χ_4^2=501.5; trend: *P*<.001) and control group (from 59/430, 13.7% to 243/430, 56.5%; χ_4^2=211.4; trend: *P*<.001; Table 3). Rates of viral load monitoring in the exposure group were significantly higher than the rates in the control group from 9 to 24 months after HIV diagnosis (all *P* values <.05). Rates of viral load suppression were significantly higher in the exposure group than the control group from 9 to 24 months after HIV diagnosis (357/381, 93.7% vs 217/243, 89.3%; χ_1^2=3.9; *P*=.048).

Table 2. Antiretroviral therapy (ART) adherence and retention in HIV care between the exposure and control groups (n=860).

Variables	Total (n=860), n (%)	Exposure group (n=430), n (%)	Control group (n=430), n (%)	Chi-square (*df*)	*P* value
Adhere to ART[a,b]				14.2 (1)	<.001
Yes	756 (87.9)	396 (92.1)	360 (83.7)		
No	104 (12.1)	34 (7.9)	70 (16.3)		
Retention in care[c]				50.7 (1)	<.001
Yes	729 (84.8)	402 (93.5)	327 (76.1)		
No	131 (15.2)	28 (6.5)	103 (24)		

[a] ART: antiretroviral therapy.

[b] Adherence to ART was defined as never having a documented period of ART interruption and never having a period when a scheduled follow-up clinic visit was missed by more than 3 months.

[c] Retention in care was defined as receiving either CD4 or viral load monitoring within 12 months after ART initiation.

Table 3. Viral suppression between the exposure and control groups (n=860).

Variables	Exposure group			Control group		
	Values, n/N (%)	Chi-square (*df*)	Trend (*P* value)	Values, n (%)	Chi-square (*df*)	Trend (*P* value)
Viral load testing after enrollment (months)		501.5 (4)	<.001		211.4 (4)	<.001
9-12	88/430 (20.5)			59/430 (13.7)		
9-15	211/430 (49.1)			115/430 (26.7)		
9-18	292/430 (67.9)			169/430 (39.3)		
9-21	341/430 (79.3)			207/430 (48.1)		
9-24	381/430 (88.6)			243/430 (56.5)		
Viral load suppression after enrollment (months)		0.02 (4)	.89		0.1 (4)	.79
9-12	83/88 (94.3)			52/59 (88.1)		
9-15	196/211 (92.9)			105/115 (91.3)		
9-18	271/292 (92.8)			154/169 (91.1)		
9-21	317/341 (93)			187/207 (90.3)		
9-24	357/381 (93.7)			217/243 (89.3)		

ARM and All-Cause Mortality

During the study period, 25 individuals died within 4316.8 PY of follow-up. Among the 25 people living with HIV who died, 19 (76%) and 6 (24%) were attributed to ARM and NARM, respectively. Reasons for ARM were AIDS-related opportunistic infections (9/19, 47%), AIDS-related malignancies (6/19, 31%), and HIV wasting syndrome (4/19, 21%). Compared with the control group, the exposure group had lower rates of ARM (1.8 vs 7.0 per 1000 PY; rate difference −5.2; *P*=.01) and all-cause mortality (2.3 vs 9.3 per 1000 PY; rate difference −7.0; *P*=.002). There was no statistically significant difference in NARM (0.5

vs 2.3 per 1000 PY; rate difference −1.9; *P*=.12) comparing the 2 groups.

Survival analysis showed that people living with HIV in the exposure group had higher cumulative survival rates (χ_1^2=9.1; *P*=.003) compared with controls (Figure 2). After adjusting for other factors, multivariate analysis showed that the receipt of peer-led community-based support was associated with a 72% (adjusted subdistribution hazard ratio 0.28, 95% CI 0.09-0.95) reduction in ARM and a 70% (adjusted hazard ratio 0.30, 95% CI 0.11-0.82) reduction in all-cause mortality. Other factors associated with ARM and all-cause mortality are presented in Table 4.

Figure 2. All-cause mortality between the exposure and control groups.

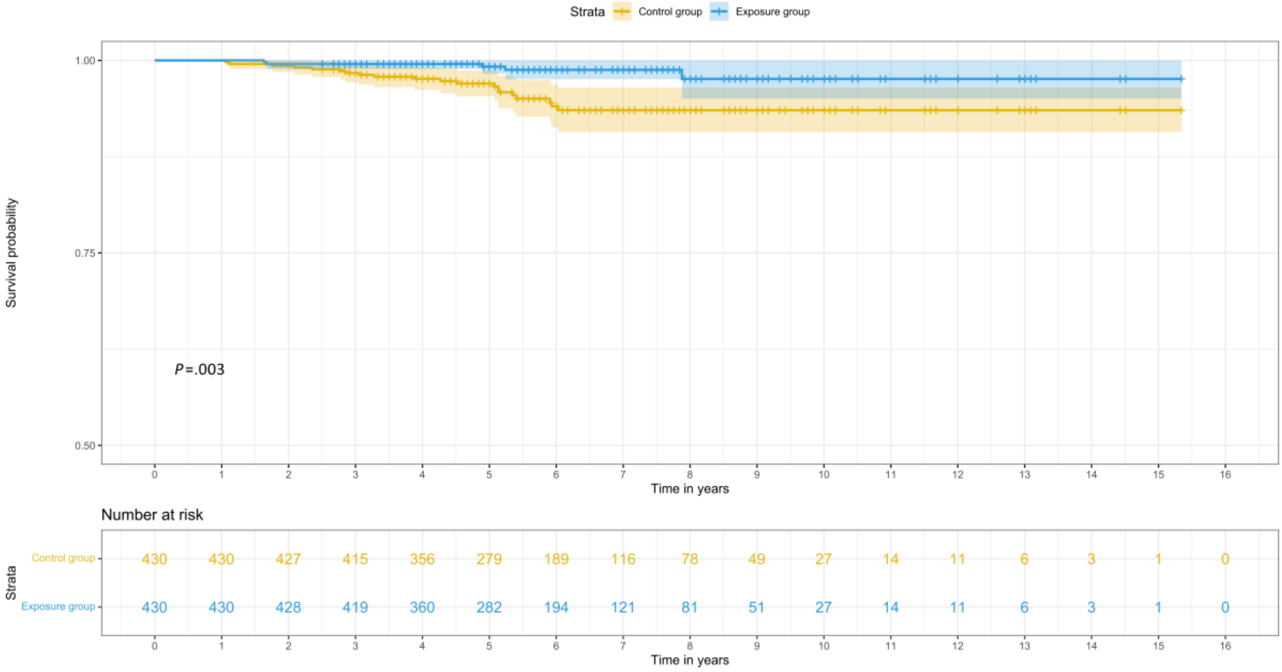


Table 4. Correlates of AIDS-related mortality (ARM) and all-cause mortality among people living with HIV.

Characteristics	PY[a] of follow-up	ARM			All-cause mortality		
		Deaths, n	Mortality rate (per 1000 PY; %)	Adjusted SHR[b] (95% CI)	Deaths, n	Mortality rate (per 1000 PY; %)	Adjusted HR[c] (95% CI)
Sex							
Male	3650.43	17	4.7	Reference	22	6.0	Reference
Female	666.47	2	3.0	0.32 (0.05-1.87)[d]	3	4.5	0.30 (0.08-1.12)[d]
Age (years) at diagnosis							
<30	1934.85	4	2.1	Reference	4	2.1	Reference
≥30	2382.05	15	6.3	1.99 (0.52-7.61)	21	8.8	2.44 (0.56-10.77)
Ethnicity							
Han ethnicity	4206.67	18	4.3	Reference	24	5.7	Reference
Others	110.23	1	9.1	1.46 (0.24-8.79)	1	9.1	1.25 (0.15-10.15)[d]
Marital status							
Unmarried	2143.83	6	2.8	Reference	6	2.8	Reference
Married	1424.30	6	4.2	0.67 (0.20-2.28)	10	7.0	1.08 (0.28-4.16)
Divorced and widowed	748.77	7	9.3	1.10 (0.29-4.15)	9	12.0	1.43 (0.35-5.85)
Education							
Primary or less	196.50	1	5.1	Reference	3	15.3	Reference
Secondary	2320.90	14	6.0	1.85 (0.16-21.30)[d]	18	7.8	0.67 (0.18-2.47)
Higher	1799.50	4	2.2	1.34 (0.07-25.00)[d]	4	2.2	0.36 (0.06-1.98)
Occupation							
Student	397.19	1	2.5	Reference	1	2.5	Reference
Others	3919.71	18	4.6	0.61 (0.07-5.76)[d]	24	6.1	0.76 (0.07-7.71)
Mode of HIV transmission							
MSM[e]	2554.63	6	2.3	Reference	9	3.5	Reference
Others	1762.27	13	7.4	3.17 (1.19-8.41)	16	9.1	1.82 (0.67-4.94)
Date of HIV diagnosis							
Before June 2016	2997.84	15	5.0	Reference	20	6.7	Reference
After June 2016	1319.06	4	3.0	0.69 (0.16-2.97)[d]	5	3.8	0.43 (0.14-1.30)[d]
Baseline CD4 count (cells/µL)							
<200	1290.49	13	10.1	Reference	15	11.6	Reference
≥200	3026.40	6	2.0	0.24 (0.09-0.66)	10	3.3	0.39 (0.17-0.90)
Receipt of peer-led community-based support							
No	2144.74	15	7.0	Reference	20	9.3	Reference
Yes	2172.12	4	1.8	0.28 (0.09-0.95)	5	2.3	0.30 (0.11-0.82)
Adherence to ART[f]							

Characteristics	PY[a] of follow-up	ARM			All-cause mortality		
		Deaths, n	Mortality rate (per 1000 PY; %)	Adjusted SHR[b] (95% CI)	Deaths, n	Mortality rate (per 1000 PY; %)	Adjusted HR[c] (95% CI)
Yes	3760.96	10	2.7	Reference	15	4.0	Reference
No	555.94	9	16.2	4.55 (1.74-12.50)	10	18.0	3.70 (1.56-8.33)
Retention in care							
No	855.34	6	7.0	Reference	10	11.7	Reference
Yes	3461.56	13	3.8	0.46 (0.14-1.54)	15	4.3	0.27 (0.11-0.64)

[a] PY: person-years.

[b] SHR: subdistribution hazard ratio.

[c] HR: hazard ratio.

[d] These variables were tested associated with mortality at P>.20, which were not included in the initial adjusted model.

[e] MSM: men who have sex with men.

[f] ART: antiretroviral therapy.

Discussion

Principal Findings

In this propensity score–matched analysis of HIV treatment outcomes and mortality among people living with HIV in Wuxi between 2006 and 2021, we found that people living with HIV who received peer-led community-based support had better ART adherence, higher rates of retention in care, and improved survival rates than people living with HIV who had not received these services. Rates of viral suppression and compliance to viral load monitoring guidelines were high in the exposure group, with >90% of people living with HIV who had participated in the multicomponent support services program achieving viral suppression 9 to 24 months after HIV diagnosis. Most previous studies evaluating the relationship between peer-led support and HIV treatment outcomes were implemented in low- to middle-income countries (LMICs) in Africa or high-income countries, with few previous studies having been conducted in Asia [27,28]. Previous evaluations of similar interventions also reported improvements in ART adherence, viral suppression, and retention in care, although few have evaluated associations with ARM or all-cause mortality, given the limited follow-up periods [15,29]. Our study adds to the existing literature through demonstrating associations between peer-led community-based support services and improved HIV treatment outcomes, including lower ARM and all-cause mortality, over 15 years of follow-up in a middle-income country in Asia.

The Impact of Peer Support on Improving ART Adherence, Retention in Care, and Viral Suppression

People living with HIV in China who received at least 6 months of adjunctive peer-led community-based support had better ART adherence, viral suppression, and retention in HIV care than those in matched controls who only received standard clinic-based HIV care. Differences in retention in care were particularly pronounced, with >93% and only 76% of people living with HIV in the exposure and control groups retained in HIV care for ≥12 months after ART initiation, respectively. Our findings are similar to previous observational and experimental evaluations of support programs delivered by CHWs to people living with HIV in LMICs [15]. A prospective observational cohort study in Rwanda found that 85% of people living with HIV who received supplementary community-based support services remained retained in care and virally suppressed 1 year after ART initiation, which was a significantly higher proportion than people living with HIV who only received standard clinic-based HIV care (79%) [17]. Several meta-analyses have shown that community-based HIV initiatives and peer support are superior to standard clinic-based care to improve retention in care, ART adherence, and viral suppression [30,31]. There are many possible explanations as to why the receipt of peer-led community-based support was associated with improved HIV treatment outcomes in our analysis. The fear of unintended disclosure of HIV status, stigma, and discrimination have been identified as negatively influencing ART adherence and retention in care [32,33]. The SMART support framework used by CHWs at Rainbow Family explicitly addresses these potential barriers. It includes social and emotional support to help people living with HIV cope with stresses associated with partner notification and the disclosure of HIV status, as well as assistance navigating stigma and discrimination from health care providers. Previous studies have also found anxiety, depression, and treatment fatigue as negatively impacting ART adherence and retention in care [34,35]. The psychological support provided by CHWs in this study included specific strategies to address concurrent anxiety and depression among people living with HIV. CHWs at Rainbow Family also provided reminders regarding medication refills and laboratory testing in efforts to mitigate treatment fatigue, which may have improved rates of ART adherence and viral load monitoring among people living with HIV in the exposure group.

The Impact of Peer Support on Reductions of ARM and All-Cause Mortality

Compared with matched controls, the receipt of peer-led community-based support services was associated with 72% and 70% reductions in ARM and all-cause mortality, respectively, among people living with HIV in China. These findings are consistent with previous studies conducted in Rwanda [17] and South Africa [36] where support services delivered by CHWs in community settings were associated with lower mortality among people living with HIV after ART initiation. The improved HIV treatment outcomes observed among people living with HIV who received supportive services from Rainbow Family were likely a major contributing factor to the higher cumulative survival in the exposure group. Decades of research have shown that retention in care and ART adherence are essential to achieving and maintaining viral suppression, improving immune function, and thereby reducing ARM among people living with HIV [37-42]. Our results suggest that the observed reductions in all-cause mortality between the exposure and control groups were primarily driven by reductions in ARM as NARM did not significantly differ between the 2 groups. Although CHWs at Rainbow Family did provide general education on healthy lifestyle choices beyond HIV care, including the importance of physical activity, balanced diet, and smoking cessation, these were secondary considerations and were not explicitly emphasized by any particular component of the SMART support framework. As the proportion of deaths attributable to NARM among people living with HIV continues to grow worldwide, the scope of CHW support for people living with HIV may need to be expanded to address causes of mortality beyond opportunistic infections and HIV- or AIDS-associated malignancies, such as cardiovascular disease, substance use disorder, renal disease, and liver disease [13].

Differences in Various Peer-Led Community-Based Support Models

The peer-led community-based support services provided by Rainbow Family adopted a multicomponent SMART framework that attempted to provide broad support for HIV care across multiple domains, including HIV education and counseling, ART adherence and laboratory monitoring reminders, social and emotional support for interpersonal relationships and disclosure of HIV status, and mental health counseling. Although the WHO guidelines encourage task shifting of HIV services to CHWs in resource-limited settings, little guidance is available as to which types of support services or delivery models should be prioritized in task-shifting efforts to optimize health outcomes among people living with HIV [14]. Systemic reviews of CHWs interventions for people living with HIV in LMICs have found that support services vary considerably. Many CHW-based or peer-led interventions described in the literature include regular phone calls, SMS text messages, or home visits to answer HIV care questions and reminders to obtain ART refills and laboratory monitoring [31]. Other components that have been inconsistently incorporated into CHW-based or peer-led interventions for people living with HIV include directly observed therapy, referrals to other medical or social work services, harm reduction and substance use management, and various forms of psychosocial support [31]. The wide variety of intervention models as well as substantial differences in study design, setting, and measured outcomes complicate efforts to perform comparative analyses to determine which forms of peer-led community-based support are the most important to improve HIV treatment outcomes among people living with HIV [27]. Future qualitative and implementation science work may be particularly helpful in identifying core components of peer-led community-based support services for people living with HIV and optimizing service delivery across different settings in LMICs.

Engaging CBOs in HIV Care in China

Although the Chinese government has established the China AIDS Fund to support the development of grassroots organizations including CBOs providing service to people living with HIV, there are only a small number of people living with HIV who are served by peer CHWs. As CBOs in China are required to register as an official nongovernmental organization in a local Civil Affairs Bureau and most of them do not have sufficient personnel and funding to support themselves to achieve this, they have to apply for governmental funding on HIV as grassroots organizations through cooperation with local CDCs or specialized infectious disease hospital, which limits their development and growth owing to lacking endorsement from the government [43,44]. Our study showed that providing peer-led community-based differentiated support through peer CHWs in CBOs were likely to significantly improve the survival of people living with HIV. Accordingly, we strongly suggest that the central government and local governments at all levels should formulate more supportive policies and invest more to engage CBOs in HIV care. Because people living with HIV in China are generally experiencing a high degree of stigma owing to the conservative social environment [45], many infected persons are not willing to join CBOs to work as a peer CHW. Consequently, multilateral and multilevel efforts are urgently needed to overcome stigma and discrimination against people living with HIV to encourage them to engage in providing peer-led community-based differentiated support services.

Comparison With Prior Work

To the best of our knowledge, this is the first study to evaluate associations between the receipt of a multicomponent peer-led support service and survival among people living with HIV over a 15-year follow-up period and in a country in Asia based on a propensity score–matched analysis.

Limitations

Our study has several important limitations. The nonrandomized retrospective nature of our analysis prevents us from determining whether peer-led community-based support caused the observed differences in HIV treatment outcomes and survival between the exposure and control groups. Confounding factors may be responsible for some or all of our observed associations. For example, it is possible people living with HIV who were experiencing a mental health crisis or socioeconomic instability were both systematically more likely to have poor HIV treatment outcomes and less likely to remain connected to support services at Rainbow Family for at least 6 months. We used PSM to match individuals in the exposure and control groups to control for many common confounders that are known to influence HIV treatment outcomes; however, the possibility of unmeasured confounders influencing our results remains. Of note, by limiting our analytical sample to the 430 people living with HIV who received peer-led community-based support and 430 of matched controls, the representativeness of our sample may have been impacted, and people living with HIV included in this analysis may not be representative of people living with HIV in China more broadly. Therefore, the generalizability of our results must be interpreted with caution. In addition, although Rainbow Family was the only CBO registered with the local government to provide community-based support services to people living with HIV in Wuxi between 2006 and 2021, it is possible that other local CBOs or nongovernmental organizations were providing support services to some of the persons included in our analysis that were unknown to the local government or these authors, resulting in misclassification bias.

Conclusions

This study adds additional evidence supporting the use of peer-led community-based HIV initiatives and peer support interventions in LMICs, demonstrating that these programs are likely impactful in resource-limited settings beyond sub-Saharan Africa. The SMART framework used to guide the multicomponent HIV support services provided by Rainbow Family may be a model for similar services in other LMICs, particularly in other parts of China and East Asia where comprehensive peer-led support for people living with HIV remains less common. Further studies conducted across diverse settings are needed to confirm the effect and cost-effectiveness of such interventions.

Acknowledgments

This study was supported by grants of the Programme of Science and Technology Development Plan of Wuxi Municipal Science and Technology Bureau (grant Y20222006), Top Talent Support Program for Young and Middle-aged People of Wuxi Municipal Health Commission (grant BJ2020099), the Natural Science Foundation of China Excellent Young Scientists Fund (grant 82022064), Natural Science Foundation of China International and Regional Research Collaboration Project (grant 72061137001), the Sanming Project of Medicine in Shenzhen (grant SZSM201811071), the High Level Project of Medicine in Longhua, Shenzhen (grant HLPM201907020105), the Shenzhen Science and Technology Innovation Commission Basic Research Program (grant JCYJ20190807155409373), Special Support Plan for High-Level Talents of Guangdong Province (grant 2019TQ05Y230), and the Fundamental Research Funds for the Central Universities (grant 58000-31620005). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

XM and HZ conceived the study. XM, HY, WM, and HZ designed the study protocol and analysis plan. XM, HY, TF, and HZ wrote the manuscript. JG provided consultation on data analysis. WM, JG, and TF edited the manuscript and assisted with the interpretation of results. ZL critically reviewed the manuscript. XM is the cocorresponding author. All authors critically reviewed and substantively revised the manuscript. All authors have approved the final version of the manuscript for publication.

Conflicts of Interest

None declared.

References

1. Global AIDS update 2020. Joint United Nations Programme on HIV/AIDS. Geneva, Switzerland: UNAIDS; 2020. URL: https://www.unaids.org/sites/default/files/media_asset/2020_global-aids-report_en.pdf [accessed 2022-04-23]
2. Frescura L, Godfrey-Faussett P, Feizzadeh A A, El-Sadr W, Syarif O, Ghys PD, on and behalf of the 2025 testing treatment target Working Group. Achieving the 95 95 95 targets for all: a pathway to ending AIDS. PLoS One 2022 Aug 4;17(8):e0272405 [FREE Full text] [doi: 10.1371/journal.pone.0272405] [Medline: 35925943]
3. Zhao Y, Han M, Ma Y, Li D. Progress towards the 90-90-90 targets for controlling HIV - China, 2018. China CDC Wkly 2019 Nov;1(1):5-7 [FREE Full text] [Medline: 34594590]
4. Nakagawa F, Lodwick RK, Smith CJ, Smith R, Cambiano V, Lundgren JD, et al. Projected life expectancy of people with HIV according to timing of diagnosis. AIDS 2012 Jan 28;26(3):335-343. [doi: 10.1097/QAD.0b013e32834dcec9] [Medline: 22089374]
5. May MT, Ingle SM. Life expectancy of HIV-positive adults: a review. Sex Health 2011 Dec;8(4):526-533. [doi: 10.1071/SH11046] [Medline: 22127039]
6. van Sighem AI, Gras LA, Reiss P, Brinkman K, de Wolf F, ATHENA national observational cohort study. Life expectancy of recently diagnosed asymptomatic HIV-infected patients approaches that of uninfected individuals. AIDS 2010 Jun 19;24(10):1527-1535. [doi: 10.1097/QAD.0b013e32833a3946] [Medline: 20467289]
7. Adeoti AO, Dada M, Elebiyo T, Fadare J, Ojo O. Survey of antiretroviral therapy adherence and predictors of poor adherence among HIV patients in a tertiary institution in Nigeria. Pan Afr Med J 2019 Jul 31;33:277 [FREE Full text] [doi: 10.11604/pamj.2019.33.277.18711] [Medline: 31692880]
8. Mann M, Diero L, Kemboi E, Mambo F, Rono M, Injera W, et al. Antiretroviral treatment interruptions induced by the Kenyan postelection crisis are associated with virological failure. J Acquir Immune Defic Syndr 2013 Oct 01;64(2):220-224 [FREE Full text] [doi: 10.1097/QAI.0b013e31829ec485] [Medline: 24047971]

9. Zhao Y, Wu Z, McGoogan JM, Sha Y, Zhao D, Ma Y, et al. Nationwide cohort study of antiretroviral therapy timing: treatment dropout and virological failure in China, 2011-2015. Clin Infect Dis 2019 Jan 01;68(1):43-50 [FREE Full text] [doi: 10.1093/cid/ciy400] [Medline: 29771296]
10. Kouanda S, Meda IB, Nikiema L, Tiendrebeogo S, Doulougou B, Kaboré I, et al. Determinants and causes of mortality in HIV-infected patients receiving antiretroviral therapy in Burkina Faso: a five-year retrospective cohort study. AIDS Care 2012;24(4):478-490. [doi: 10.1080/09540121.2011.630353] [Medline: 22148973]
11. Otwombe KN, Laher F, Tutu-Gxashe T, Gray G, Mohapi L. The effect of a maturing antiretroviral program on early mortality for patients with advanced immune-suppression in Soweto, South Africa. PLoS One 2013 Nov 28;8(11):e81538 [FREE Full text] [doi: 10.1371/journal.pone.0081538] [Medline: 24312317]
12. Gheibi Z, Shayan Z, Joulaei H, Fararouei M, Beheshti S, Shokoohi M. Determinants of AIDS and non-AIDS related mortality among people living with HIV in Shiraz, southern Iran: a 20-year retrospective follow-up study. BMC Infect Dis 2019 Dec 30;19(1):1094 [FREE Full text] [doi: 10.1186/s12879-019-4676-x] [Medline: 31888541]
13. Farahani M, Mulinder H, Farahani A, Marlink R. Prevalence and distribution of non-AIDS causes of death among HIV-infected individuals receiving antiretroviral therapy: a systematic review and meta-analysis. Int J STD AIDS 2017 Jun;28(7):636-650. [doi: 10.1177/0956462416632428] [Medline: 26868158]
14. Task shifting: global recommendations and guidelines. World Health Organization. Geneva, Switzerland: World Health Organization; 2007. URL: https://apps.who.int/iris/bitstream/handle/10665/43821/9789241596312_eng.pdf [accessed 2022-05-11]
15. Mwai GW, Mburu G, Torpey K, Frost P, Ford N, Seeley J. Role and outcomes of community health workers in HIV care in sub-Saharan Africa: a systematic review. J Int AIDS Soc 2013 Sep 10;16(1):18586 [FREE Full text] [doi: 10.7448/IAS.16.1.18586] [Medline: 24029015]
16. Lubega M, Tumwesigye NM, Kadobera D, Marrone G, Wabwire-Mangen F, Peterson S, et al. Effect of community support agents on retention of people living with HIV in pre-antiretroviral care: a randomized controlled trial in eastern Uganda. J Acquir Immune Defic Syndr 2015 Oct 01;70(2):e36-e43. [doi: 10.1097/QAI.0000000000000723] [Medline: 26079842]
17. Franke MF, Kaigamba F, Socci AR, Hakizamungu M, Patel A, Bagiruwigize E, et al. Improved retention associated with community-based accompaniment for antiretroviral therapy delivery in rural Rwanda. Clin Infect Dis 2013 May;56(9):1319-1326. [doi: 10.1093/cid/cis1193] [Medline: 23249611]
18. Lifson AR, Workneh S, Hailemichael A, Demisse W, Slater L, Shenie T. Implementation of a peer HIV community support worker program in rural Ethiopia to promote retention in care. J Int Assoc Provid AIDS Care 2017;16(1):75-80 [FREE Full text] [doi: 10.1177/2325957415614648] [Medline: 26518590]
19. Mavhu W, Willis N, Mufuka J, Bernays S, Tshuma M, Mangenah C, et al. Effect of a differentiated service delivery model on virological failure in adolescents with HIV in Zimbabwe (Zvandiri): a cluster-randomised controlled trial. Lancet Glob Health 2020 Feb;8(2):e264-e275 [FREE Full text] [doi: 10.1016/S2214-109X(19)30526-1] [Medline: 31924539]
20. Zhang X, Chen L, Ning Y, Bai J, Dong Z, Wang R. An exploration of involvemen tof non-governmental organizations (NGOs) in comprehensive management of PLWHA. Chin J AIDS STD 2011;17(1):61-64 [FREE Full text]
21. Zhang FJ, Pan J, Yu L, Wen Y, Zhao Y. Current progress of China's free ART program. Cell Res 2005 Nov;15(11-12):877-882. [doi: 10.1038/sj.cr.7290362] [Medline: 16354563]
22. Ma Y, Ma Y, Lu H, Wang L, Mao Y, Shen L, et al. CD4 count and viral load monitoring for people living with HIV. In: National Guidlines on Follow-up for People Living With HIV. Beijing, China: National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention; Sep 9, 2016.
23. Mangal TD, Meireles MV, Pascom AR, de Almeida Coelho R, Benzaken AS, Hallett TB. Determinants of survival of people living with HIV/AIDS on antiretroviral therapy in Brazil 2006-2015. BMC Infect Dis 2019 Feb 28;19(1):206 [FREE Full text] [doi: 10.1186/s12879-019-3844-3] [Medline: 30819120]
24. Li L, Yuan T, Wang J, Fitzpatrick T, Li Q, Li P, et al. Sex differences in HIV treatment outcomes and adherence by exposure groups among adults in Guangdong, China: a retrospective observational cohort study. EClinicalMedicine 2020 May;22:100351 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100351] [Medline: 32510049]
25. Limbada M, Macleod D, Situmbeko V, Muhau E, Shibwela O, Chiti B, HPTN 071 (PopART) study team. Rates of viral suppression in a cohort of people with stable HIV from two community models of ART delivery versus facility-based HIV care in Lusaka, Zambia: a cluster-randomised, non-inferiority trial nested in the HPTN 071 (PopART) trial. Lancet HIV 2022 Jan;9(1):e13-e23 [FREE Full text] [doi: 10.1016/S2352-3018(21)00242-3] [Medline: 34843674]
26. Ingle SM, May MT, Gill MJ, Mugavero MJ, Lewden C, Abgrall S, Antiretroviral Therapy Cohort Collaboration. Impact of risk factors for specific causes of death in the first and subsequent years of antiretroviral therapy among HIV-infected patients. Clin Infect Dis 2014 Jul 15;59(2):287-297 [FREE Full text] [doi: 10.1093/cid/ciu261] [Medline: 24771333]
27. Øgård-Repål A, Berg RC, Fossum M. Peer support for people living with HIV: a scoping review. Health Promot Pract 2023 Jan;24(1):172-190 [FREE Full text] [doi: 10.1177/15248399211049824] [Medline: 34693785]
28. Genberg BL, Shangani S, Sabatino K, Rachlis B, Wachira J, Braitstein P, et al. Improving engagement in the HIV care cascade: a systematic review of interventions involving people living with HIV/AIDS as peers. AIDS Behav 2016 Oct;20(10):2452-2463 [FREE Full text] [doi: 10.1007/s10461-016-1307-z] [Medline: 26837630]

29. Hacking D, Mgengwana-Mbakaza Z, Cassidy T, Runeyi P, Duran LT, Mathys RH, et al. Peer mentorship via mobile phones for newly diagnosed HIV-positive youths in clinic care in Khayelitsha, South Africa: mixed methods study. J Med Internet Res 2019 Dec 10;21(12):e14012 [FREE Full text] [doi: 10.2196/14012] [Medline: 31821150]
30. Berg RC, Page S, Øgård-Repål A. The effectiveness of peer-support for people living with HIV: a systematic review and meta-analysis. PLoS One 2021 Jun 17;16(6):e0252623 [FREE Full text] [doi: 10.1371/journal.pone.0252623] [Medline: 34138897]
31. Dave S, Peter T, Fogarty C, Karatzas N, Belinsky N, Pant Pai N. Which community-based HIV initiatives are effective in achieving UNAIDS 90-90-90 targets? A systematic review and meta-analysis of evidence (2007-2018). PLoS One 2019 Jul 17;14(7):e0219826 [FREE Full text] [doi: 10.1371/journal.pone.0219826] [Medline: 31314764]
32. Izudi J, Okoboi S, Lwevola P, Kadengye D, Bajunirwe F. Effect of disclosure of HIV status on patient representation and adherence to clinic visits in eastern Uganda: a propensity-score matched analysis. PLoS One 2021 Oct 19;16(10):e0258745 [FREE Full text] [doi: 10.1371/journal.pone.0258745] [Medline: 34665842]
33. Mhode M, Nyamhanga T. Experiences and impact of stigma and discrimination among people on antiretroviral therapy in Dar es salaam: a qualitative perspective. AIDS Res Treat 2016;2016:7925052 [FREE Full text] [doi: 10.1155/2016/7925052] [Medline: 27110395]
34. van Wyk BE, Davids LA. Challenges to HIV treatment adherence amongst adolescents in a low socio-economic setting in Cape Town. South Afr J HIV Med 2019 Oct 28;20(1):1002 [FREE Full text] [doi: 10.4102/sajhivmed.v20i1.1002] [Medline: 31745433]
35. Truong M, Rane MS, Govere S, Galagan SR, Moosa M, Stoep AV, et al. Depression and anxiety as barriers to art initiation, retention in care, and treatment outcomes in KwaZulu-Natal, South Africa. EClinicalMedicine 2021 Jan;31:100621 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100621] [Medline: 33490927]
36. Fatti G, Meintjes G, Shea J, Eley B, Grimwood A. Improved survival and antiretroviral treatment outcomes in adults receiving community-based adherence support: 5-year results from a multicentre cohort study in South Africa. J Acquir Immune Defic Syndr 2012 Dec 01;61(4):e50-e58. [doi: 10.1097/QAI.0b013e31826a6aee] [Medline: 22842842]
37. Chesney MA. Factors affecting adherence to antiretroviral therapy. Clin Infect Dis 2000 Jun;30 Suppl 2:S171-S176. [doi: 10.1086/313849] [Medline: 10860902]
38. Bangsberg DR, Moss AR, Deeks SG. Paradoxes of adherence and drug resistance to HIV antiretroviral therapy. J Antimicrob Chemother 2004 May;53(5):696-699. [doi: 10.1093/jac/dkh162] [Medline: 15044425]
39. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Joint United Nations Programme on HIV/AIDS. Geneva, Switzerland: UNAIDS; 2014. URL: https://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf [accessed 2022-06-30]
40. Chaiyachati KH, Ogbuoji O, Price M, Suthar AB, Negussie EK, Bärnighausen T. Interventions to improve adherence to antiretroviral therapy: a rapid systematic review. AIDS 2014 Mar;28 Suppl 2:S187-S204. [doi: 10.1097/QAD.0000000000000252] [Medline: 24849479]
41. Wakooko P, Gavamukulya Y, Wandabwa JN. Viral load suppression and associated factors among HIV patients on antiretroviral treatment in Bulambuli district, eastern Uganda: a retrospective cohort study. Infect Dis (Auckl) 2020 Nov 5;13:1178633720970632 [FREE Full text] [doi: 10.1177/1178633720970632] [Medline: 33223836]
42. Buchbinder M, Blue C, Rennie S, Juengst E, Brinkley-Rubinstein L, Rosen DL. Practical and ethical concerns in implementing enhanced surveillance methods to improve continuity of HIV care: qualitative expert stakeholder study. JMIR Public Health Surveill 2020 Sep 04;6(3):e19891 [FREE Full text] [doi: 10.2196/19891] [Medline: 32886069]
43. Wu Z, Wang Y, Detels R, Rotheram-Borus MJ. China AIDS policy implementation: reversing the HIV/AIDS epidemic by 2015. Int J Epidemiol 2010 Dec;39 Suppl 2(Suppl 2):ii1-ii3 [FREE Full text] [doi: 10.1093/ije/dyq220] [Medline: 21113031]
44. Cheng W, Cai Y, Tang W, Zhong F, Meng G, Gu J, et al. Providing HIV-related services in China for men who have sex with men. Bull World Health Organ 2016 Mar 01;94(3):222-227 [FREE Full text] [doi: 10.2471/BLT.15.156406] [Medline: 26966334]
45. Zhang C, Li X, Liu Y, Qiao S, Zhang L, Zhou Y, et al. Emotional, physical and financial burdens of stigma against people living with HIV/AIDS in China. AIDS Care 2016;28 Suppl 1(sup1):124-131 [FREE Full text] [doi: 10.1080/09540121.2016.1146206] [Medline: 26886415]

Abbreviations

ARM: AIDS-related mortality
ART: antiretroviral therapy
CBO: community-based organization
CDC: center for disease control and prevention
CHW: community health worker
CNHCIMS: Chinese National HIV/AIDS Comprehensive Information Management System
LMIC: low- to middle-income country
NARM: non–AIDS-related mortality

PSM: propensity score matching
PY: person-years
SMART: social, medical, and mental antiretroviral therapy support
WHO: World Health Organization

Edited by Y Khader; submitted 19.10.22; peer-reviewed by H Joulaei, M Fararouei; comments to author 20.12.22; revised version received 02.01.23; accepted 11.01.23; published 24.03.23.

Please cite as:
Meng X, Yin H, Ma W, Gu J, Lu Z, Fitzpatrick T, Zou H
Peer-Led Community-Based Support Services and HIV Treatment Outcomes Among People Living With HIV in Wuxi, China: Propensity Score–Matched Analysis of Surveillance Data From 2006 to 2021
JMIR Public Health Surveill 2023;9:e43635
URL: https://publichealth.jmir.org/2023/1/e43635
doi:10.2196/43635
PMID:36961492

Original Paper

The Impact of COVID-19 Lockdown on Cases of and Deaths From AIDS, Gonorrhea, Syphilis, Hepatitis B, and Hepatitis C: Interrupted Time Series Analysis

Xinsheng Wu[1], MD; Xinyi Zhou[1], MD; Yuanyi Chen[1], MD; Ke Zhai[1], MD; Ruoyao Sun[1], MD; Ganfeng Luo[1], PhD; Yi-Fan Lin[1], PhD; Yuwei Li[1], PhD; Chongguang Yang[1*], PhD; Huachun Zou[1*], PhD

School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, Guangdong, China

*these authors contributed equally

Corresponding Author:
Huachun Zou, PhD
School of Public Health (Shenzhen)
Sun Yat-sen University
66 Gongchang Road
Guangming District
Shenzhen, Guangdong, 518107
China
Phone: 86 20 8733 5651
Email: zouhuachun@mail.sysu.edu.cn

Abstract

Background: China implemented a nationwide lockdown to contain COVID-19 from an early stage. Previous studies of the impact of COVID-19 on sexually transmitted diseases (STDs) and diseases caused by blood-borne viruses (BBVs) in China have yielded widely disparate results, and studies on deaths attributable to STDs and BBVs are scarce.

Objective: We aimed to elucidate the impact of COVID-19 lockdown on cases, deaths, and case-fatality ratios of STDs and BBVs.

Methods: We extracted monthly data on cases and deaths for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C between January 2015 and December 2021 from the notifiable disease reporting database on the official website of the National Health Commission of China. We used descriptive statistics to summarize the number of cases and deaths and calculated incidence and case-fatality ratios before and after the implementation of a nationwide lockdown (in January 2020). We used negative binominal segmented regression models to estimate the immediate and long-term impacts of lockdown on cases, deaths, and case-fatality ratios in January 2020 and December 2021, respectively.

Results: A total of 14,800,330 cases of and 127,030 deaths from AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C were reported from January 2015 to December 2021, with an incidence of 149.11/100,000 before lockdown and 151.41/100,000 after lockdown and a case-fatality ratio of 8.21/1000 before lockdown and 9.50/1000 after lockdown. The negative binominal model showed significant decreases in January 2020 in AIDS cases (–23.4%; incidence rate ratio [IRR] 0.766, 95% CI 0.626-0.939) and deaths (–23.9%; IRR 0.761, 95% CI 0.647-0.896), gonorrhea cases (–34.3%; IRR 0.657, 95% CI 0.524-0.823), syphilis cases (–15.4%; IRR 0.846, 95% CI 0.763-0.937), hepatitis B cases (–17.5%; IRR 0.825, 95% CI 0.726-0.937), and hepatitis C cases (–19.6%; IRR 0.804, 95% CI 0.693-0.933). Gonorrhea, syphilis, and hepatitis C showed small increases in the number of deaths and case-fatality ratios in January 2020. By December 2021, the cases, deaths, and case-fatality ratios for each disease had either reached or remained below expected levels.

Conclusions: COVID-19 lockdown may have contributed to fewer reported cases of AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C and more reported deaths and case-fatality ratios of gonorrhea, syphilis, and hepatitis C in China.

(JMIR Public Health Surveill 2023;9:e40591) doi:10.2196/40591

KEYWORDS

COVID-19; AIDS; gonorrhea; syphilis; hepatitis B; hepatitis C; China; nonpharmaceutical; intervention; STD; disease; virus; cases; deaths; fatality; data

Introduction

China implemented stringent nonpharmaceutical interventions (NPIs), including city lockdowns and traffic restrictions, social distancing, and school closures in the early stages of the outbreak of COVID-19 to contain its spread [1-7]. These NPIs contributed to low morbidity and mortality of COVID-19 in the country [8-10]. However, it is necessary to keep a constant focus on other important infectious diseases. The numbers of documented influenza and pneumonia deaths in the United States in 2020 were reported to increase by 7.5% compared to 2019 [11], and modeling predicted that HIV deaths in African countries would increase [12,13].

As humans migrate and communicate, sexually transmitted diseases (STDs) and diseases caused by blood-borne viruses (BBVs) are a serious global burden [14-16]. Although China has established a national STD surveillance sentinel system, increased publicity and education, and improved treatment guidelines, STDs and BBVs are still on the rise. HIV/AIDS morbidity, mortality, and new HIV infections (defined as newly acquired blood HIV antibodies, as shown by positive results in ELISA and Western blotting and a large viral load in PCR testing) continually increased each year in China from 2004 to 2016 (from 0.235/100,000, 0.057/100, and 1.020/100,000 in 2004, respectively, to 3.990/100,000, 1.034/100, and 6.442/100,000, respectively, in 2016) [17]. Surveillance data from 2004 to 2013 in China showed that the incidence of hepatitis C, HIV infection, and syphilis increased by 19.2% (95% CI 15.9%-22.6%) annually [18].

Understanding changes in AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C over time is vital to contain the disease burdens of STDs and BBVs. Many studies have mainly focused on changes in respiratory infectious diseases during COVID-19 [19-21]. There is limited and conflicting evidence on changes in STDs and BBVs during the pandemic. A large tertiary care hospital in Chicago [22] reported the rate of presumed active infection with syphilis increased from 1.2% (prepandemic; June 2019-March 2020) to 1.8% (during the pandemic; April 2020-June 2020). Two referral centers in Greece [23] recorded 10% and 36.9% decreases in syphilis and gonorrhea cases, respectively, in 2020. A study of the STD and BBV service of an Italian clinic [24] did not find statistically significant differences compared with the previous 4 years. In addition, previous studies of the impact of COVID-19 on STDs and BBVs in China have yielded widely disparate results [25,26], and data on deaths attributable to STDs and BBVs are scarce. We aimed to elucidate the impact of COVID-19 lockdown on cases, deaths, and case-fatality ratios for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China using interrupted time series analysis. This may provide evidence of the impact of COVID-19 lockdown on STDs and BBVs at the national level and inform planning for efficient control strategies in the postpandemic era.

Methods

Study Design

China launched a national online routine reporting system for selected infectious diseases in 2003 that covers all 31 provinces in mainland China. Health facilities at various levels are required to report cases through standard case report forms within 2 to 24 hours of the detection of a notifiable infectious disease.

We extracted national data on monthly reported cases and deaths for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C, which are mainly or can be sexually transmitted, from January 2015 to December 2021 from the official website of the China Health and Wellness Commission [27]. The lockdown period was considered to start in January 2020 with the Wuhan city lockdown (on January 23, 2020) and massive restrictions implemented nationwide. Although some regions had eased restrictions before April 8 (when Wuhan lifted its restrictions), they tended to take stronger NPI measures to deal with the possible importation of COVID-19 cases from Wuhan. Based on these considerations, we defined the lockdown period as ending in April 2020. AIDS cases are defined as a diagnosis of AIDS. Syphilis cases include both primary and secondary syphilis. Cases that were reported as cases of notifiable infectious disease and died as a result of that infectious disease are reported as deaths, excluding deaths due to accidents or noninfectious diseases.

Statistical Analysis

The primary outcome was the number of cases and deaths reported by month for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C before and after lockdown implementation. The secondary outcomes were the incidence and case-fatality ratios. We defined the incidence (per 100,000 population) as the mean value of annual new reported cases divided by the population size and the case-fatality ratio (per 1000 population) as the mean value of annual reported deaths divided by annual new reported cases.

We first summarized monthly reported data on cases and deaths for the 5 STDs and BBVs as the median (IQR) before, during, and after lockdown and calculated the incidence and case-fatality ratios before and after lockdown. The Kruskal-Wallis test was used to compare secondary outcomes before and after lockdown. Since overdispersion occurred in almost all outcomes (Multimedia Appendix 1, Table S1), we conducted an interrupted time series analysis by fitting a negative binominal segmented regression model to estimate the immediate effects of lockdown on cases, deaths, and case-fatality ratios, as well as trends after lockdown [28]. The negative binominal model included (1) a time variable, (2) a dummy variable representing the pre- and postlockdown periods, and (3) an interaction term between time and the dummy variables (Multimedia Appendix 1).

We adjusted for seasonality and long-term trends to control their effects and allow the direct study of changes in public health issues (ie, testing and transmission), as follows: (1) the reporting of STD and BBV data is influenced by individual and clinical activity, so we adjusted for seasonal variation by including a Fourier term consisting of 2 sine-cosine pairs in the model [29]; (2) the reporting of STD and BBV data has long-term trends, so we adjusted for long-term trends by using the prelockdown model to predict the expected outcomes after lockdown (ie, a counterfactual scenario). We analyzed the annual trends of the models for the 5 diseases to determine their seasonality.

The incidence rate ratio (IRR) was calculated by comparing the fitted numbers from the model with the expected numbers from the contemporaneous counterfactual. By subtracting a period from the time variable, we were able to center time in January 2020 and December 2021 to estimate the impact on different timepoints [30]. To calculate the trend after lockdown, we added the coefficients associated with time and the time-dummy interaction. We adjusted the standard errors of the model parameters using the Newey-West method to calculate a 95% CI for the IRR in both January 2020 and December 2021, with lag taking the optimal value calculated [31].

We conducted a sensitivity analysis of the number of pairs of Fourier terms (1, 3, and 5 pairs). All statistical tests were 2-sided, and *P*<.05 was considered statistically significant. We performed all analyses in R (version 4.0.4; R Foundation for Statistical Computing; Multimedia Appendix 1).

Ethical Considerations

This study was a secondary analysis of online reported data with no identifying information available to the researchers. Therefore, ethical review was not applicable.

Results

Overview of Study

From January 2015 to December 2021, 14,800,330 cases of AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C were reported in China, with a notified incidence of 149.11/100,000 before lockdown and 151.41/100,000 after lockdown. The incidence (mean value of annual incidence) change in these 5 infectious diseases before and after lockdown is shown in Table 1. There were slight but not significant (*P*>.05) changes in the incidence of the 5 diseases after lockdown. The median number of monthly reported cases was 177,556 (IQR 164,546-187,722) during the 60 months before lockdown, 155,491 (IQR 121,585-157,244) during the first 3 months of nationwide lockdown (January 2020 to March 2020), and 184,401 (IQR 180,396-190,466) during the 21 months after lockdown eased (April 2020 to December 2021). Figure 1 shows the annual seasonal pattern in the case counts of the 5 diseases. AIDS peaked in May and October to November, gonorrhea in July and November, syphilis in May to June, and hepatitis B and C in May.

In total, 127,030 deaths were reported during the study period, with the case-fatality ratio changing from 8.15/1000 before lockdown to 9.50/1000 after lockdown. The change in case-fatality ratios (ie, the mean value of annual case-fatality ratios) for the 5 infectious diseases before and after lockdown is shown in Table 2. There was a significant (*P*=.03) increase in the gonorrhea case-fatality ratio after lockdown (from 0.01/1000 to 0.04/1000). The median number of monthly deaths due to the five STDs and BBVs was 1360 (IQR 1208-1692) before lockdown, 1078 (IQR 1054-1270) during lockdown, and 1761 (IQR 1602-1942) after lockdown eased.

Table 1. Case numbers and incidence rates for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China from January 2015 to December 2021. Incidence was calculated as the mean number of annual new cases divided by the population size (per 100,000).

	Cases[a], n	Incidence before lockdown[b,c]	Incidence after lockdown[b,d]	*P* value	Cases before lockdown[c,e], median (IQR)	Cases during lockdown[e,f], median (IQR)	Cases after lockdown eased[e,g], median (IQR)
All 5 diseases	14,800,330	149.11	151.41	.70	177,556 (164,546-187,722)	155,491 (121,585-157,244)	184,401 (180,396-190,466)
AIDS	428,715	4.31	4.40	.70	5124 (4291-6098)	2759 (2446-3784)	5484 (5039-6124)
Gonorrhea	846,935	8.65	8.38	.70	10,436 (8954-11,343)	4661 (4092-6458)	10,874 (10,551-11,264)
Syphilis	3,612,684	36.15	37.56	.44	42,702 (38,898-46,394)	39,671 (30,560-40,412)	46,538 (44,438-47,999)
Hepatitis B	8,219,292	82.72	84.29	.70	97,461 (90,780-103,851)	88,150 (69,828-89,588)	101,701 (99,319-105,393)
Hepatitis C	1,692,704	17.27	16.78	.44	20,400 (19,224-21,502)	16,718 (12,893-17,002)	20,438 (20,001-21,254)

[a] Total number of cases from January 2015 to December 2021.

[b] Per year.

[c] Before lockdown: January 2015 to December 2019.

[d] After lockdown: From January 2020 to December 2021.

[e] Per month.

[f] During lockdown: From January 2020 to March 2020.

[g] After lockdown eased: From April 2020 to December 2021.

Figure 1. Negative binominal models for monthly cases of AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C by year in China from January 2015 to December 2021. Total cases (A), AIDS (B), gonorrhea (C), syphilis (D), hepatitis B (E) and hepatitis C (F).

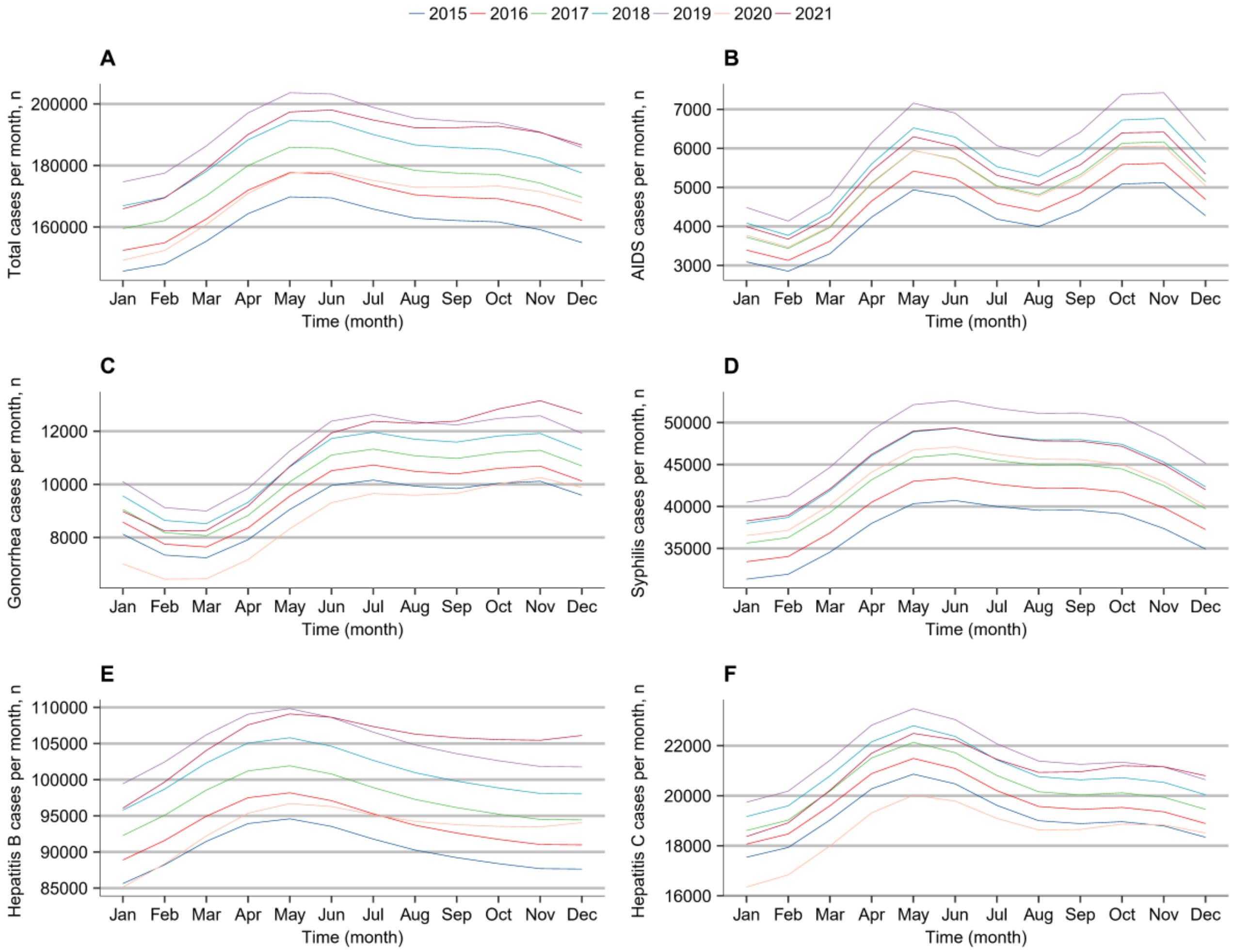


Table 2. Deaths and case-fatality ratios for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China from January 2015 to December 2021. The case-fatality ratio represents the mean number of annual deaths divided by annual new cases (per 1000).

	Deaths[a], n	Case-fatality ratio before lockdown[b,c]	Case-fatality ratio after lockdown[c,d]	*P* value	Deaths before lockdown[b,e], median (IQR)	Deaths during lockdown[e,f], median (IQR)	Deaths after lockdown eased[e,g], median (IQR)
All 5 diseases	127,030	8.15	9.50	.44	1360 (1208-1692)	1078 (1054-1270)	1761 (1602-1942)
AIDS	122,508	271.38	316.23	.44	1304 (1156-1635)	971 (962-1194)	1700 (1554-1874)
Gonorrhea	15	0.01	0.04	.03	0 (0-0)	0 (0-2)	0 (0-0)
Syphilis	516	0.14	0.16	>.99	6 (4-8)	9 (8-17)	6 (4-7)
Hepatitis B	3168	0.38	0.41	.24	38 (31-42)	37 (34-58)	37 (33-44)
Hepatitis C	823	0.49	0.48	.70	9 (7-12)	12 (9-14)	10 (8-11)

[a] Total number from January 2015 to December 2021.

[b] Before lockdown: From January 2015 to December 2019.

[c] Per year.

[d] After lockdown: From January 2020 to December 2021.

[e] Per month.

[f] During lockdown: From January 2020 to March 2020.

[g] After lockdown eased: From April 2020 to December 2021.

Cases and Deaths

AIDS

In the first month of lockdown, there was a 23.4% (IRR 0.766, 95% CI 0.626-0.939; Figure 2A and Table 3) decrease in the number of reported AIDS cases; after lockdown, the monthly number remained unchanged (IRR 1.005, 95% CI 0.990-1.020); by December 2021, AIDS cases were still below the expected level (IRR 0.715, 95% CI 0.586-0.874).

In the first month of lockdown, the number of reported AIDS deaths showed a 23.9% decline (IRR 0.761, 95% CI 0.647-0.896); after lockdown, the number showed an increasing trend of 1.1% (IRR 1.011, 95% CI 1.000-1.021) per month; by December 2021, the number of AIDS deaths remained below the expected level (IRR 0.754, 95% CI 0.654-0.869). In the first month of lockdown, there was no significant change in the AIDS case-fatality ratio (IRR 1.004, 95% CI 0.895-1.125); after lockdown, the monthly number remained unchanged (IRR 1.006, 95% CI 0.997-1.016); by December 2021, the AIDS case-fatality ratio remained in line with the expected level (IRR 1.075, 95% CI 0.924-1.252).

Figure 2. Monthly numbers, trends, and fitted negative binominal segmented regression models for 5 sexually transmitted diseases and diseases caused by blood-borne viruses in China from January 2015 to December 2021: AIDS (A), gonorrhea (B), syphilis (C), hepatitis B (D), and hepatitis C (E). The light blue shaded areas indicate the period during the lockdown, from January 2020 to March 2020, and the dark blue shaded areas indicate the period after lockdown eased, from April 2020 to December 2021. Case-fatality ratios represent the number of deaths divided by the number of new cases (per 1000). NPI: nonpharmaceutical intervention.

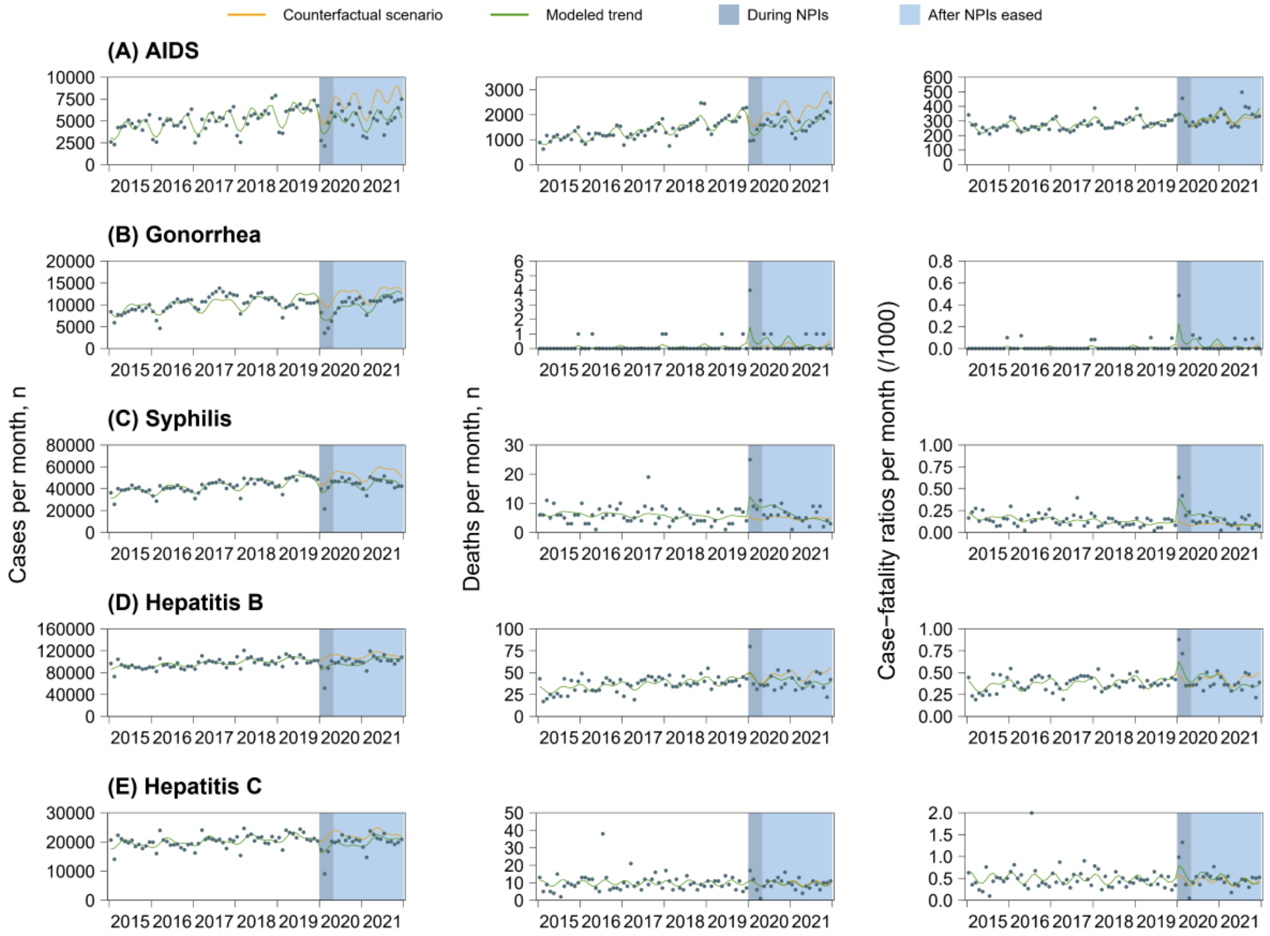


Table 3. Negative binominal segmented regression models of the impact of COVID-19 lockdown on AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China from January 2015 to December 2021. Trend represents the slope change per month. Case-fatality ratio represents the number of deaths divided by the number of new cases. Autocorrelation was addressed for all diseases using Newey-West standard errors to calculate CIs, with lag taking the optimal value calculated.

Disease	IRR[a] (95% CI) at lockdown[b]	IRR (95% CI) at study end[c]	Trend (95% CI) before lockdown[d]	Trend (95% CI) after lockdown[e]
AIDS				
Cases	0.766 (0.626-0.939)	0.715 (0.586-0.874)	1.008 (1.006-1.009)	1.005 (0.990-1.020)
Deaths	0.761 (0.647-0.896)	0.754 (0.654-0.869)	1.011 (1.009-1.013)	1.011 (1.000-1.021)
Case-fatality ratio	1.004 (0.895-1.125)	1.075 (0.924-1.252)	1.003 (1.002-1.004)	1.006 (0.997-1.016)
Gonorrhea				
Cases	0.657 (0.524-0.823)	0.952 (0.722-1.257)	1.005 (1.001-1.008)	1.021 (1.007-1.035)
Deaths	5.902 (1.090-31.959)	0.558 (0.038-8.266)	1.015 (0.982-1.049)	0.916 (0.804-1.043)
Case-fatality ratio	9.676 (1.764-53.089)	0.398 (0.018-8.890)	1.012 (0.977-1.048)	0.880 (0.766-1.012)
Syphilis				
Cases	0.846 (0.763-0.937)	0.818 (0.751-0.892)	1.005 (1.004-1.006)	1.004 (0.998-1.010)
Deaths	2.380 (1.455-3.893)	0.735 (0.421-1.284)	0.995 (0.990-1.001)	0.946 (0.913-0.980)
Case-fatality ratio	3.173 (1.877-5.365)	0.810 (0.442-1.484)	0.990 (0.984-0.996)	0.933 (0.897-0.971)
Hepatitis B				
Cases	0.825 (0.726-0.937)	0.968 (0.896-1.045)	1.003 (1.002-1.004)	1.010 (1.002-1.018)
Deaths	1.038 (0.854-1.261)	0.704 (0.590-0.842)	1.006 (1.003-1.009)	0.989 (0.979-0.999)
Case-fatality ratio	1.312 (0.997-1.726)	0.709 (0.556-0.902)	1.003 (1.000-1.006)	0.976 (0.958-0.995)
Hepatitis C				
Cases	0.804 (0.693-0.933)	0.950 (0.854-1.057)	1.002 (1.001-1.004)	1.010 (1.001-1.019)
Deaths	1.006 (0.781-1.296)	0.926 (0.672-1.275)	1.000 (0.994-1.006)	0.996 (0.982-1.011)
Case-fatality ratio	1.394 (1.080-1.800)	0.906 (0.653-1.258)	0.997 (0.991-1.004)	0.979 (0.966-0.992)

[a] IRR: incidence rate ratio.

[b] January 2020.

[c] December 2021.

[d] From January 2015 to December 2019.

[e] From January 2020 to December 2021.

Gonorrhea

In the first month of lockdown, there was a 34.3% decline (IRR 0.657, 95% CI 0.524-0.823; Figure 2B and Table 3) in the number of reported gonorrhea cases; after lockdown, there was a 2.1% (IRR 1.021, 95% CI 1.007-1.035) monthly increasing trend; by December 2021, the number of gonorrhea cases had returned to the expected level (IRR 0.952, 95% CI 0.722-1.257).

In the first month of lockdown, the number of reported gonorrhea deaths showed an increase of 490.2% (IRR 5.902, 95% CI 1.090-31.959); after lockdown, the number remained largely unchanged (IRR 0.916, 95% CI 0.804-1.043); by December 2021, the number of gonorrhea deaths had returned to below the expected level (IRR 0.558, 95% CI 0.038-8.266). In the first month of lockdown, there was an 867.6% increase (IRR 9.676, 95% CI 1.764-53.089) in the gonorrhea case-fatality ratio; after lockdown, the number remained unchanged (IRR 0.880, 95% CI 0.766-1.012); by August 2021, the gonorrhea case-fatality ratio had returned to below the expected level (IRR 0.398, 95% CI 0.018-8.890).

Syphilis

In the first month of lockdown, the number of reported syphilis cases showed a 15.4% decrease (IRR 0.846, 95% CI 0.763-0.937; Figure 2C and Table 3); after lockdown, the monthly number per month was unchanged (IRR 1.004, 95% CI 0.998-1.010); by December 2021, the number of syphilis cases was still below the expected level (IRR 0.818, 95% CI 0.751-0.892).

In the first month of lockdown, the number of reported syphilis deaths showed an increase of 138% (IRR 2.380, 95% CI 1.455-3.893); after lockdown, the number per month showed a decreasing trend of 5.4% (IRR 0.946, 95% CI 0.913-0.980); by December 2021, the number of syphilis deaths had returned to below the expected level (IRR 0.735, 95% CI 0.421-1.284). In the first month of lockdown, the syphilis case-fatality ratio showed an increase of 217.3% (IRR 3.173, 95% CI

1.877-5.365); after lockdown, the number per month showed a decreasing trend of 6.7% (IRR 0.933, 95% CI 0.897-0.971); by December 2021, the syphilis case-fatality ratio had returned to the expected level (IRR 0.810, 95% CI 0.442-1.484).

Hepatitis B

In the first month of lockdown, there was a 17.5% (IRR 0.825, 95% CI 0.726-0.937; Figure 2D and Table 3) decrease in the number of reported hepatitis B cases; after lockdown, there was an increasing trend of 1% (IRR 1.010, 95% CI 1.002-1.018) per month; by December 2021, the number of hepatitis B cases had returned to the expected level (IRR 0.968, 95% CI 0.896-1.045).

In the first month of lockdown, there was no significant change in the number of reported hepatitis B deaths (IRR 1.038, 95% CI 0.854-1.261); after lockdown, the number showed a decreasing trend of 1.1% (IRR 0.989, 95% CI 0.979-0.999) per month; by December 2021, the number of hepatitis B deaths was significantly lower than the expected level (IRR 0.704, 95% CI 0.590-0.842). In the first month of lockdown, there was no significant change in the hepatitis B case-fatality ratio (IRR 1.312, 95% CI 0.997-1.726); after lockdown, it showed a decreasing trend of 2.4% (IRR 0.976, 95% CI 0.958-0.995); by December 2021, the hepatitis B case-fatality ratio was significantly lower than the expected level (IRR 0.709, 95% CI 0.556-0.902).

Hepatitis C

In the first month of lockdown, there was a 19.6% (IRR 0.804, 95% CI 0.693-0.933; Figure 2E and Table 3) decrease in the number of reported hepatitis C cases; after lockdown, there was a 1% increasing trend (IRR 1.010, 95% CI 1.001-1.019) per month; by December 2021, the number of hepatitis C cases had returned to the expected level (IRR 0.950, 95% CI 0.854-1.057).

In the first month of lockdown, there was no significant change in the number of reported hepatitis C deaths (IRR 1.006, 95% CI 0.781-1.296); after lockdown, the monthly number remained unchanged (IRR 0.996, 95% CI 0.982-1.011); by December 2021, the number of hepatitis C deaths was not significantly different from the expected level (IRR 0.926, 95% CI 0.672-1.275). In the first month of lockdown, there was a 39.4% increase (IRR 1.394, 95% CI 1.080-1.800) in the hepatitis C case-fatality ratio; after lockdown, it showed a decreasing trend of 2.1% (IRR 0.979, 95% CI 0.966-0.992) per month; by December 2021, the hepatitis C case-fatality ratio was not significantly different from the expected level (IRR 0.906, 95% CI 0.653-1.258).

Sensitivity Analysis

Sensitivity analysis showed that the number of pairs of the Fourier term had little effect on the results (Multimedia Appendix 1, Tables S2, S3, and S4).

Discussion

Principal Findings

This study found that COVID-19 lockdown had a significant impact on the number of reported cases and deaths and case-fatality ratios for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China. In the first month of lockdown, significant decreases were seen in reported AIDS cases (–23.4%) and deaths (–23.9%), reported gonorrhea cases (–34.3%), reported syphilis cases (–15.4%), reported hepatitis B cases (–17.5%), and reported hepatitis C cases (–19.6%). Gonorrhea, syphilis, and hepatitis C showed small increases in the number of reported deaths and case-fatality ratios in January 2020. By December 2021, nearly two years after the lockdown, the reported cases, deaths, and case-fatality ratios for each disease either reached or remained below expected levels.

Our study found that monthly reported cases of each disease showed an alarming upward trend before the COVID-19 pandemic, which was consistent with the overall trend worldwide. Since the 1980s, the number of new reported cases of STDs and BBVs has declined significantly with comprehensive prevention efforts around the world [32-34]. However, over the past decade, several countries and regions have reported an increasing trend in reported cases of STDs and BBVs. From 2009 to 2019, reported syphilis cases in Sri Lanka and reported gonorrhea and syphilis cases in Thailand showed an increasing trend [35]. A similar trend has been observed in several European countries [36]. With the widespread global epidemic of HIV, STDs and BBVs have become an increasingly serious public health problem.

In addition, the seasonality of reported cases of STDs and BBVs shown in this study is consistent with existing research. A study from Melbourne, Australia [37] found that urethral gonorrhea diagnoses among men who have sex with men (odds ratio [OR] 1.23, 95% CI 1.04-1.46) and nongonococcal urethritis diagnoses among men who have sex with women (OR 1.11, 95% CI 1.03-1.20) were higher in summer compared with winter. This may be due to increased chances of infection due to increased biological sexual desire, number of sexual partners, and sexual activity during the hot summer months [38,39], as well as increased willingness to test for STDs and BBVs in the population due to World AIDS Day (December 1). In addition, the annual seasonal pattern of monthly reported cases of the 5 infectious diseases was largely due to the annual activities of clinics and health facilities. Every year around April, health facilities in different regions in China receive their annual assignments, so most peaks in disease reporting are in the last 3 quarters. In the first quarter, there is a major holiday in China (Chinese New Year), leading to a low disease reporting level. Similar trends were observed in our study, suggesting that our data source and analytic tools were robust.

Decreases in reported AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C cases during the COVID-19 lockdown have also been observed in other countries. A national study in the United States [40] found that the numbers of weekly reported cases of chlamydia (–49.8%), gonorrhea (–71.2%), and syphilis (–63.7%) were much lower in 2020 than in week 15 of 2019. A study in Catalonia, Spain [41] found that there were 51% fewer reported cases of STDs and BBVs than expected since the beginning of the COVID-19 pandemic, reaching an average of 56% during the lockdown, with the greatest decrease of 72% for chlamydia and the least, of 22%, for syphilis. Difficulties and unwillingness to seek medical care and possible underreporting were considered as reasons for the decrease in the 5 diseases. According to a study on health service use during COVID-19 in China, all-cause visits showed a decrease of 47.6% (ranging from 24.8% in township health centers to 70% in primary care clinics) and inpatient volume showed a decrease of 47.7% (ranging from 26.8% in other health facilities to 57.8% in first-level hospitals) in February 2020 [42]. Another study, in Hong Kong, found that emergency department visits decreased by 27.4%, from 1,426,259 in 2019 to 1,035,562 in 2020 [43]. Several studies have reported that patients have delayed or avoided seeking care for fear of being infected with COVID-19 [44,45].

The decrease in the number of reported cases may also be due to NPIs such as mobility restrictions and social distancing measures affecting people's sexual activity. A study of 967 Chinese youth [46] found that 22% of participants reported decreased sexual desire, 41% decreased frequency of sexual intercourse, 30% increased frequency of masturbation, 20% decreased alcohol consumption before or during sexual activity, and 31% deterioration in partner relationships during the pandemic and related containment measures. An online survey conducted in China [47] showed that 44% of people reported a decrease in sexual partners during COVID-19. In addition, frequent long-term lockdowns reduce interpersonal contact and public gatherings, which may also reduce the risk of transmission of STDs and BBVs in the population.

The immediate decrease in reported AIDS deaths in the first month of lockdown may be due to COVID-19 deaths among people living with AIDS being attributed to COVID-19 rather than AIDS. According to the National Health Commission of China, the cumulative number of deaths nationwide reached 4636 by the end of 2021 [48]. However, we are concerned that incorrect death certificates might obscure the truth about the deaths of HIV patients, especially during the COVID-19 pandemic, when a high work burden might have resulted in incorrect findings and documentation within hospitals and clinics. In the future, it would be valuable to elucidate the reasons behind changes in the number of deaths if more data become available. In addition, AIDS deaths were lower than the expected level 24 months after lockdown, which may be explained by the effectiveness of measures taken in China to prevent antiretroviral therapy (ART) interruption. NPIs adopted by countries or regions around the world during the COVID-19 pandemic were thought to affect the HIV care continuum for people living with HIV/AIDS [12,30,49,50]. The Chinese Center for AIDS/STD Control and Prevention released a nationwide directive on January 26, 2020, under which people living with HIV/AIDS could obtain 1 month of ART from any local HIV care clinic or hospital [51]. Designated hospitals and clinics across the country also contributed to the maintenance of ART. Hospitals in Shenzhen increased the supply of medications for a single ART collection and adopted courier delivery to ensure uninterrupted treatment for people living with HIV/AIDS during the outbreak [52]. Together, these efforts mitigated the disruption of ART to the greatest extent possible.

The increase in reported deaths or case-fatality ratios for gonorrhea, syphilis, and hepatitis C may be associated with shifts in medical resources and services to COVID-19 patients. China heavily focused the use of medical supplies, human resources, and health resources during the pandemic to respond to the explosive increase in confirmed COVID-19 cases, with some hospitals shutting down due to the risk of COVID-19 infection and some admitting only severe and critically ill COVID-19 cases [53]. Reduced willingness to seek medical care and access to care may have contributed to the disruption of treatment and the increase in deaths. However, it is important to note that the absolute numbers of deaths from gonorrhea, syphilis, and hepatitis C were very small and should not be overinterpreted.

The implementation of massive lockdowns did not change the increasing trend of reported cases of gonorrhea, hepatitis B, and hepatitis C. These short-term trends may not necessarily be stable due to the immediate impact of lockdown on STDs and BBVs in January 2020 and will need to be observed over a longer period. As China rapidly brought the domestic COVID-19 pandemic under control, the health system was able to shift resources and human resources to address other infectious diseases. Concerns about seeking health care were alleviated, with health facility visits and inpatient volume showing continued statistically significant increases after March 2020, reaching 89% and 91% of expected levels, respectively, in June 2020 [42]. Meanwhile, frequent and persistent NPIs in China such as social distancing continued to affect sexual activity and limit the spread of STDs and BBVs after the Wuhan lockdown was lifted (on April 7, 2020). By December 2021, reported cases, deaths, and case-fatality ratios for AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C either reached or remained below expected levels.

Our study used 7 years of nationwide data to estimate the impact of COVID-19 on AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China using 4 indicators adjusted for long-term trends and seasonality. To our knowledge, this is the first study to use interrupted time series analyses to elucidate the number of reported STD and BBV cases and deaths in China during the COVID-19 pandemic. However, our study has several limitations. First, it is an ecological study and cannot demonstrate a causal relationship between COVID-19 lockdowns and changes in trends related to AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China. Second, our study was based on publicly available data, which limited what kinds of data we could obtain. Due to a lack of data reported on a weekly basis, we could only define the time point for lockdown as January 2020, which may be inaccurate. Due to a lack of provincial data, we were unable to provide results for different provinces or regions in this study. Potentially, deaths due to COVID-19 and incorrect death certificates in hospitals and clinics during the COVID-19 pandemic could have affected the accuracy of this study's results. The lack of prevalence data is also a limitation, given that the prevalence of many infectious diseases is unknown. Although the case-fatality ratio cannot be interpreted as the overall risk of death from an infection, it is the most commonly used metric because most countries collect this information [18,54,55]. As prevalence is unknown, the case-fatality ratio not only provides a crude estimate of the risk of death (such as in the COVID-19 outbreak [56]), but also allows for cross-sectional comparisons. Finally, this study was a secondary analysis of online reported data with no identifying information available to the researchers, which limits our consideration of disease intersections in the analysis. The subjects of our study should be interpreted as reported cases of a single disease rather than as individuals.

Conclusion

In summary, these findings suggest that AIDS, gonorrhea, syphilis, hepatitis B, and hepatitis C in China were transiently affected by COVID-19 lockdown. Our findings may promote the development of STD and BBV policies and control measures in the context of the COVID-19 pandemic and the persistence of lockdowns. More long-term observations are needed to investigate trends in STDs and BBVs over time.

Acknowledgments

This study was supported by the Natural Science Foundation of China Excellent Young Scientists Fund (82022064), the Special Support Plan for High-Level Talents of Guangdong Province (2019TQ05Y230), the High Level Project of Medicine in Longhua, Shenzhen (HLPM201907020105), the Sanming Project of Medicine in Shenzhen (SZSM201811071), the Shenzhen Nanshan District San-Ming Project (SZSM202103008), and the "Pearl River Talent Plan" Innovation and Entrepreneurship Team Project of Guangdong Province (2021ZT09Y544). No funding parties had any role in the design of the study or in the explanation of the data. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit it for publication.

Data Availability

The data sets generated during and/or analyzed during the current study are available in the public, open-access China Health and Wellness Commission repository [27].

Authors' Contributions

HZ, CY, XW, XZ, and YC conceived and designed the study in consultation with the other authors. YC contributed to data collection. XZ reviewed scientific literature. XW contributed to data analysis and presentation. All authors contributed to the interpretation of the study findings. XW drafted the report with all authors critically reviewing the paper. All authors saw and approved the final report. HZ and CY contributed equally to this work. XW, XZ and YC contributed equally to this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Formula for negative binominal regression models and supplementary tables.

[DOCX File , 33 KB - publichealth_v9i1e40591_app1.docx]

References

1. The State Council's announcement on the arrangement of public holidays in 2020. The State Council of the People's Republic of China. URL: http://www.gov.cn/zhengce/content/2019-11/21/content_5454164.htm [accessed 2023-04-13]
2. The State Council's announcement on extending the Lunar New Year Holiday in 2020. The State Council of the People's Republic of China. URL: http://www.gov.cn/zhengce/content/2020-01/27/content_5472352.htm [accessed 2023-01-13]
3. Press conference on Jan 26. National Health Commission of the People's Republic of China. URL: http://www.nhc.gov.cn/xcs/fkdt/202001/12ec9062d5d041f38e210e8b69b6d7ef.shtml [accessed 2023-04-13]
4. The announcement of cancellations of major events including temple fairs in Beijing. The People's Government of Beijing Municipality. URL: http://www.beijing.gov.cn/ywdt/gzdt/t1614497.htm [accessed 2023-04-13]
5. The announcement from Wuhan's headquarters on the novel coronavirus prevention and control. The State Council of the People's Republic of China. URL: http://www.gov.cn/xinwen/2020-01/23/content_5471751.htm [accessed 2023-04-13]
6. Shen J, Duan H, Zhang B, Wang J, Ji JS, Wang J, et al. Prevention and control of COVID-19 in public transportation: Experience from China. Environ Pollut 2020 Nov;266(Pt 2):115291 [FREE Full text] [doi: 10.1016/j.envpol.2020.115291] [Medline: 32829124]
7. Liu N, Tan JC, Li J, Li S, Cai Y, Wang H. COVID-19 pandemic: Experiences in China and implications for its prevention and treatment worldwide. Curr Cancer Drug Targets 2020;20(6):410-416. [doi: 10.2174/1568009620666200414151419] [Medline: 32286947]
8. Tian H, Liu Y, Li Y, Wu C, Chen B, Kraemer MUG, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 2020 May 08;368(6491):638-642 [FREE Full text] [doi: 10.1126/science.abb6105] [Medline: 32234804]

9. Oraby T, Tyshenko MG, Maldonado JC, Vatcheva K, Elsaadany S, Alali WQ, et al. Modeling the effect of lockdown timing as a COVID-19 control measure in countries with differing social contacts. Sci Rep 2021 Feb 08;11(1):3354 [FREE Full text] [doi: 10.1038/s41598-021-82873-2] [Medline: 33558571]
10. SeyedAlinaghi S, Afsahi AM, MohsseniPour M, Behnezhad F, Salehi MA, Barzegary A, et al. Late complications of COVID-19; a systematic review of current evidence. Arch Acad Emerg Med 2021;9(1):e14 [FREE Full text] [doi: 10.22037/aaem.v9i1.1058] [Medline: 33681819]
11. Ahmad FB, Anderson RN. The leading causes of death in the US for 2020. JAMA 2021 May 11;325(18):1829-1830 [FREE Full text] [doi: 10.1001/jama.2021.5469] [Medline: 33787821]
12. Jewell BL, Mudimu E, Stover J, ten Brink D, Phillips AN, Smith JA, et al. Potential effects of disruption to HIV programmes in sub-Saharan Africa caused by COVID-19: results from multiple mathematical models. The Lancet HIV 2020 Sep;7(9):e629-e640. [doi: 10.1016/s2352-3018(20)30211-3] [Medline: 32771089]
13. Jewell BL, Smith JA, Hallett TB. Understanding the impact of interruptions to HIV services during the COVID-19 pandemic: A modelling study. EClinicalMedicine 2020 Sep;26:100483 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100483] [Medline: 33089116]
14. Kirkcaldy RD, Weston E, Segurado AC, Hughes G. Epidemiology of gonorrhoea: a global perspective. Sex Health 2019 Sep;16(5):401-411 [FREE Full text] [doi: 10.1071/SH19061] [Medline: 31505159]
15. Kreisel KM, Spicknall IH, Gargano JW, Lewis FM, Lewis RM, Markowitz LE, et al. Sexually transmitted infections among US women and men: Prevalence and incidence estimates, 2018. Sex Transm Dis 2021 Apr 01;48(4):208-214. [doi: 10.1097/OLQ.0000000000001355] [Medline: 33492089]
16. Pham-Kanter GB, Steinberg MH, Ballard RC. Sexually transmitted diseases in South Africa. Genitourin Med 1996 Jun;72(3):160-171 [FREE Full text] [doi: 10.1136/sti.72.3.160] [Medline: 8707316]
17. Qiao Y, Xu Y, Jiang D, Wang X, Wang F, Yang J, et al. Epidemiological analyses of regional and age differences of HIV/AIDS prevalence in China, 2004-2016. Int J Infect Dis 2019 Apr;81:215-220 [FREE Full text] [doi: 10.1016/j.ijid.2019.02.016] [Medline: 30797071]
18. Yang S, Wu J, Ding C, Cui Y, Zhou Y, Li Y, et al. Epidemiological features of and changes in incidence of infectious diseases in China in the first decade after the SARS outbreak: an observational trend study. Lancet Infect Dis 2017 Jul;17(7):716-725 [FREE Full text] [doi: 10.1016/S1473-3099(17)30227-X] [Medline: 28412150]
19. Huang Q, Wood T, Jelley L, Jennings T, Jefferies S, Daniells K, NPIsImpactOnFlu Consortium, et al. Impact of the COVID-19 nonpharmaceutical interventions on influenza and other respiratory viral infections in New Zealand. Nat Commun 2021 Feb 12;12(1):1001 [FREE Full text] [doi: 10.1038/s41467-021-21157-9] [Medline: 33579926]
20. Sakamoto H, Ishikane M, Ueda P. Seasonal influenza activity during the SARS-CoV-2 Outbreak in Japan. JAMA 2020 May 19;323(19):1969-1971 [FREE Full text] [doi: 10.1001/jama.2020.6173] [Medline: 32275293]
21. Wu D, Lu J, Liu Y, Zhang Z, Luo L. Positive effects of COVID-19 control measures on influenza prevention. Int J Infect Dis 2020 Jun;95:345-346 [FREE Full text] [doi: 10.1016/j.ijid.2020.04.009] [Medline: 32283283]
22. Stanford K, Almirol E, Schneider J, Hazra A. Rising syphilis rates during the COVID-19 pandemic. Sex Transm Dis 2021 Jun 01;48(6):e81-e83. [doi: 10.1097/OLQ.0000000000001431] [Medline: 33783406]
23. Apalla Z, Lallas A, Mastraftsi S, Giannoukos A, Noukari D, Goula M, et al. Impact of COVID-19 pandemic on STIs in Greece. Sex Transm Infect 2022 Feb;98(1):70 [FREE Full text] [doi: 10.1136/sextrans-2021-054965] [Medline: 33653880]
24. Bonato F, Ferreli C, Satta R, Rongioletti F, Atzori L. Syphilis and the COVID-19 pandemic: Did the lockdown stop risky sexual behavior? Clin Dermatol 2021;39(4):710-713 [FREE Full text] [doi: 10.1016/j.clindermatol.2020.11.006] [Medline: 34809778]
25. Chen S, Zhang X, Zhou Y, Yang K, Lu X. COVID-19 protective measures prevent the spread of respiratory and intestinal infectious diseases but not sexually transmitted and bloodborne diseases. J Infect 2021 Jul;83(1):e37-e39 [FREE Full text] [doi: 10.1016/j.jinf.2021.04.018] [Medline: 33895225]
26. Yan X, Wang X, Zhang X, Wang L, Zhang B, Jia Z. The epidemic of sexually transmitted diseases under the influence of COVID-19 in China. Front Public Health 2021;9:737817 [FREE Full text] [doi: 10.3389/fpubh.2021.737817] [Medline: 34976912]
27. Infectious Disease Prevention and Control. The China Health and Wellness Commission. URL: http://www.nhc.gov.cn/jkj/s2907/new_list.shtml?tdsourcetag=s_pcqq_aiomsg [accessed 2023-04-13]
28. Xiao H, Augusto O, Wagenaar BH. Reflection on modern methods: a common error in the segmented regression parameterization of interrupted time-series analyses. Int J Epidemiol 2021 Jul 09;50(3):1011-1015. [doi: 10.1093/ije/dyaa148] [Medline: 33097937]
29. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017 Feb 01;46(1):348-355 [FREE Full text] [doi: 10.1093/ije/dyw098] [Medline: 27283160]
30. Dorward J, Khubone T, Gate K, Ngobese H, Sookrajh Y, Mkhize S, et al. The impact of the COVID-19 lockdown on HIV care in 65 South African primary care clinics: an interrupted time series analysis. Lancet HIV 2021 Mar;8(3):e158-e165 [FREE Full text] [doi: 10.1016/S2352-3018(20)30359-3] [Medline: 33549166]

31. Wu X, Wu G, Ma Y, Huang X, Yang Y, Cai Y, et al. The impact of COVID-19 non-pharmaceutical interventions on HIV care continuum in China: An interrupted time series analysis. Lancet Reg Health West Pac 2022 Dec;29:100569 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100569] [Medline: 35966023]
32. Fenton KA, Lowndes CM. Recent trends in the epidemiology of sexually transmitted infections in the European Union. Sex Transm Infect 2004 Aug;80(4):255-263 [FREE Full text] [doi: 10.1136/sti.2004.009415] [Medline: 15295121]
33. Shao C, Xu W, Ye G. Sexually transmitted disease control in China (1949-1994). Chin Med Sci J 1996 Dec;11(4):252-257. [Medline: 9387394]
34. Steen R, Wi TE, Kamali A, Ndowa F. Control of sexually transmitted infections and prevention of HIV transmission: mending a fractured paradigm. Bull World Health Organ 2009 Nov;87(11):858-865 [FREE Full text] [doi: 10.2471/blt.08.059212] [Medline: 20072772]
35. Sharma M, Rewari BB, Aditama TY, Turlapati P, Dallabetta G, Steen R. Control of sexually transmitted infections and global elimination targets, South-East Asia Region. Bull World Health Organ 2021 Apr 01;99(4):304-311 [FREE Full text] [doi: 10.2471/BLT.20.254003] [Medline: 33953448]
36. Syphilis and congenital syphilis in Europe - A review of epidemiological trends (2007–2018) and options for response. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/publications-data/syphilis-and-congenital-syphilis-europe-review-epidemiological-trends-2007-2018 [accessed 2023-04-13]
37. Cornelisse VJ, Chow EPF, Chen MY, Bradshaw CS, Fairley CK. Summer heat: a cross-sectional analysis of seasonal differences in sexual behaviour and sexually transmissible diseases in Melbourne, Australia. Sex Transm Infect 2016 Jun;92(4):286-291. [doi: 10.1136/sextrans-2015-052225] [Medline: 26546343]
38. Chow EPF, Vodstrcil LA, Fairley CK. Seasonal variations in kissing and sexual activities among men who have sex with men in Melbourne, Australia: implications for seasonal sexually transmissible infection preventions and interventions. Sex Health 2020 Apr;17(2):149-154. [doi: 10.1071/SH19046] [Medline: 32135076]
39. Wellings K, Macdowall W, Catchpole M, Goodrich J. Seasonal variations in sexual activity and their implications for sexual health promotion. J R Soc Med 1999 Feb;92(2):60-64 [FREE Full text] [doi: 10.1177/014107689909200204] [Medline: 10450213]
40. Pagaoa M, Grey J, Torrone E, Kreisel K, Stenger M, Weinstock H. Trends in nationally notifiable sexually transmitted disease case reports during the US COVID-19 pandemic, January to December 2020. Sex Transm Dis 2021 Oct 01;48(10):798-804 [FREE Full text] [doi: 10.1097/OLQ.0000000000001506] [Medline: 34224523]
41. Sentís A, Prats-Uribe A, López-Corbeto E, Montoro-Fernandez M, Nomah DK, de Olalla PG, Catalan HIVSTI Surveillance Group. The impact of the COVID-19 pandemic on Sexually Transmitted Infections surveillance data: incidence drop or artefact? BMC Public Health 2021 Sep 07;21(1):1637 [FREE Full text] [doi: 10.1186/s12889-021-11630-x] [Medline: 34493244]
42. Xiao H, Dai X, Wagenaar BH, Liu F, Augusto O, Guo Y, et al. The impact of the COVID-19 pandemic on health services utilization in China: Time-series analyses for 2016-2020. Lancet Reg Health West Pac 2021 Apr;9:100122 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100122] [Medline: 34327438]
43. Wai AK, Wong CK, Wong JY, Xiong X, Chu OC, Wong MS, et al. Changes in Emergency Department Visits, Diagnostic Groups, and 28-Day Mortality Associated With the COVID-19 Pandemic: A Territory-Wide, Retrospective, Cohort Study. Ann Emerg Med 2022 Feb;79(2):148-157 [FREE Full text] [doi: 10.1016/j.annemergmed.2021.09.424] [Medline: 34742591]
44. Czeisler, Marynak K, Clarke KE, Salah Z, Shakya I, Thierry JM, et al. Delay or avoidance of medical care because of COVID-19-related concerns - United States, June 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 11;69(36):1250-1257 [FREE Full text] [doi: 10.15585/mmwr.mm6936a4] [Medline: 32915166]
45. Mantica G, Riccardi N, Terrone C, Gratarola A. Non-COVID-19 visits to emergency departments during the pandemic: the impact of fear. Public Health 2020 Jun;183:40-41 [FREE Full text] [doi: 10.1016/j.puhe.2020.04.046] [Medline: 32417567]
46. Li G, Tang D, Song B, Wang C, Qunshan S, Xu C, et al. Impact of the COVID-19 pandemic on partner relationships and sexual and reproductive health: cross-sectional, online survey study. J Med Internet Res 2020 Aug 06;22(8):e20961 [FREE Full text] [doi: 10.2196/20961] [Medline: 32716895]
47. Li W, Li G, Xin C, Wang Y, Yang S. Challenges in the practice of sexual medicine in the time of COVID-19 in China. J Sex Med 2020 Jul;17(7):1225-1228 [FREE Full text] [doi: 10.1016/j.jsxm.2020.04.380] [Medline: 32418751]
48. Update on the COVID-19 outbreak as at 24:00 on 31 December, 2021. National Health Commission of China. URL: http://www.nhc.gov.cn/xcs/yqtb/202201/64fae96558284e548f31f67bb541b8f2.shtml [accessed 2023-04-13]
49. Yang X, Zeng C, Tam CC, Qiao S, Li X, Shen Z, et al. HIV service interruptions during the COVID-19 pandemic in China: the role of COVID-19 challenges and institutional response from healthcare professional's perspective. AIDS Behav 2022 Apr;26(4):1270-1278 [FREE Full text] [doi: 10.1007/s10461-021-03484-6] [Medline: 34613522]
50. Medina N, Alastruey-Izquierdo A, Bonilla O, Ortíz B, Gamboa O, Salazar LR, Fungired. Impact of the COVID-19 pandemic on HIV care in Guatemala. Int J Infect Dis 2021 Jul;108:422-427 [FREE Full text] [doi: 10.1016/j.ijid.2021.06.011] [Medline: 34119676]

51. UNAIDS and China working together during the COVID-19 outbreak to ensure that people living with HIV continue to get treatment. Joint United Nations Programme on HIV/AIDS. URL: https://www.unaids.org/en/resources/presscentre/pressreleaseandstatementarchive/2020/february/20200218_china_covid19 [accessed 2023-04-13]
52. Wang H. HIV care during the coronavirus disease-2019 pandemic in Shenzhen, China. Curr Opin HIV AIDS 2020 Nov;15(6):341-344 [FREE Full text] [doi: 10.1097/COH.0000000000000653] [Medline: 32969977]
53. Ma Xiaowei, director of the National Health Commission: Wuhan Critical Care Designated Hospital has a concentration of more than 90 national and provincial medical teams. National Health Commission of the People's Republic of China. URL: http://www.nhc.gov.cn/wjw/xwdt/202002/ccf7b9e3d53d4f6d866aaa305c0e6d76.shtml [accessed 2020-02-29]
54. Randremanana R, Andrianaivoarimanana V, Nikolay B, Ramasindrazana B, Paireau J, Ten Bosch QA, et al. Epidemiological characteristics of an urban plague epidemic in Madagascar, August-November, 2017: an outbreak report. Lancet Infect Dis 2019 May;19(5):537-545 [FREE Full text] [doi: 10.1016/S1473-3099(18)30730-8] [Medline: 30930106]
55. Campèse C, Jarraud S, Che D. [Legionnaire's disease: surveillance in France in 2005]. Med Mal Infect 2007 Nov;37(11):716-721. [doi: 10.1016/j.medmal.2007.05.007] [Medline: 17988811]
56. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet 2020 Feb 15;395(10223):470-473 [FREE Full text] [doi: 10.1016/S0140-6736(20)30185-9] [Medline: 31986257]

Abbreviations

ART: antiretroviral therapy
BBV: blood-borne virus
IRR: incidence rate ratio
NPI: nonpharmaceutical intervention
OR: odds ratio
STD: sexually transmitted disease

Edited by T Sanchez, A Mavragani; submitted 28.06.22; peer-reviewed by K Tang, H Xiao; comments to author 25.10.22; revised version received 07.11.22; accepted 12.01.23; published 03.05.23.

Please cite as:
Wu X, Zhou X, Chen Y, Zhai K, Sun R, Luo G, Lin YF, Li Y, Yang C, Zou H
The Impact of COVID-19 Lockdown on Cases of and Deaths From AIDS, Gonorrhea, Syphilis, Hepatitis B, and Hepatitis C: Interrupted Time Series Analysis
JMIR Public Health Surveill 2023;9:e40591
URL: https://publichealth.jmir.org/2023/1/e40591
doi:10.2196/40591
PMID:36634257

Original Paper

Predictive Modeling of Lapses in Care for People Living with HIV in Chicago: Algorithm Development and Interpretation

Joseph A Mason[1], BS; Eleanor E Friedman[1], MS, PhD; Samantha A Devlin[1], MS; John A Schneider[1], MD, MPH; Jessica P Ridgway[1], MD, MS

[1] The Chicago Center for HIV Elimination, Department of Medicine, University of Chicago, Chicago, IL, United States

Corresponding Author:
Joseph A Mason, BS
The Chicago Center for HIV Elimination
Department of Medicine
University of Chicago
5837 S. Maryland Ave.
Chicago, IL, 60637
United States
Phone: 1 773 702 9586
Email: joemason@bsd.uchicago.edu

Abstract

Background: Reducing care lapses for people living with HIV is critical to ending the HIV epidemic and beneficial for their health. Predictive modeling can identify clinical factors associated with HIV care lapses. Previous studies have identified these factors within a single clinic or using a national network of clinics, but public health strategies to improve retention in care in the United States often occur within a regional jurisdiction (eg, a city or county).

Objective: We sought to build predictive models of HIV care lapses using a large, multisite, noncurated database of electronic health records (EHRs) in Chicago, Illinois.

Methods: We used 2011-2019 data from the Chicago Area Patient-Centered Outcomes Research Network (CAPriCORN), a database including multiple health systems, covering the majority of 23,580 people with an HIV diagnosis living in Chicago. CAPriCORN uses a hash-based data deduplication method to follow people across multiple Chicago health care systems with different EHRs, providing a unique citywide view of retention in HIV care. From the database, we used diagnosis codes, medications, laboratory tests, demographics, and encounter information to build predictive models. Our primary outcome was lapses in HIV care, defined as having more than 12 months between subsequent HIV care encounters. We built logistic regression, random forest, elastic net logistic regression, and XGBoost models using all variables and compared their performance to a baseline logistic regression model containing only demographics and retention history.

Results: We included people living with HIV with at least 2 HIV care encounters in the database, yielding 16,930 people living with HIV with 191,492 encounters. All models outperformed the baseline logistic regression model, with the most improvement from the XGBoost model (area under the receiver operating characteristic curve 0.776, 95% CI 0.768-0.784 vs 0.674, 95% CI 0.664-0.683; $P<.001$). Top predictors included the history of care lapses, being seen by an infectious disease provider (vs a primary care provider), site of care, Hispanic ethnicity, and previous HIV laboratory testing. The random forest model (area under the receiver operating characteristic curve 0.751, 95% CI 0.742-0.759) revealed age, insurance type, and chronic comorbidities (eg, hypertension), as important variables in predicting a care lapse.

Conclusions: We used a real-world approach to leverage the full scope of data available in modern EHRs to predict HIV care lapses. Our findings reinforce previously known factors, such as the history of prior care lapses, while also showing the importance of laboratory testing, chronic comorbidities, sociodemographic characteristics, and clinic-specific factors for predicting care lapses for people living with HIV in Chicago. We provide a framework for others to use data from multiple different health care systems within a single city to examine lapses in care using EHR data, which will aid in jurisdictional efforts to improve retention in HIV care.

(JMIR Public Health Surveill 2023;9:e43017) doi:10.2196/43017

KEYWORDS

HIV; predictive model; lapse in care; retention in care; people living with HIV; Chicago; HIV care continuum; electronic health record; EHR

Introduction

Background

Retention in HIV care is an important step in the HIV care continuum for both public health efforts to end the HIV epidemic and for the individual health of people living with HIV [1]. The Centers for Disease Control and Prevention aims to reduce new HIV infections by at least 90% over the next 10 years [2]. Out of every 10 cases of HIV transmission, 8 are associated with people living with HIV who are not engaged in regular HIV medical care [2]. As part of the goal to reduce new HIV infections, it is important to ensure people living with HIV are receiving regular medical care. Currently, only 50.1% of people living with HIV in the United States are receiving continuous HIV care [3]. Therefore, increasing the percentage of people living with HIV retained in care is critical to ending the HIV epidemic in the United States.

Many methods to improve retention in care are being piloted across the United States [4]. In HIV clinical settings, the use of case managers to assist with appointment scheduling and health care navigation has been shown to improve retention [5]. Another initiative, Data to Care, is a strategy by public health departments to use HIV surveillance data to identify people living with HIV that have fallen out of care and attempt to relink them [6]. However, a study on the Data to Care program in Seattle and King County, Washington, found that only 50% of people living with HIV were successfully reengaged in care, suggesting a need for additional strategies to increase the number of people living with HIV in care [7]. One upstream intervention would be to identify people living with HIV who are more likely to fall out of care prior to lapsing in care, allowing them to be prioritized for intense case management and other retention interventions.

Prior Work

Previous studies have created predictive models to identify people who are likely to be at risk for lapses in HIV care. Ramachandran et al [8] used electronic health record (EHR) data from a single urban HIV clinic to predict retention in care. Similarly, Ridgway et al [9] used EHR data from multiple HIV clinics across the United States to predict retention in care. These studies identified factors associated with lapses in HIV care at the clinic or national level, but public health strategies to improve retention in care in the United States often occur at the city or county level. Local socio-structural factors may impact retention in care among people living with HIV within different cities in unique ways. To our knowledge, no prior studies have predicted which patients will lapse from HIV care within a single urban city (ie, using EHR data from multiple clinics within a single city or geographic region). Our study focuses on the city of Chicago, which is located in Cook County, Illinois, a priority jurisdiction for the Centers for Disease Control and Prevention's Ending the HIV Epidemic initiative [10]. The Chicago Department of Public Health (CDPH) is responsible for identifying and implementing HIV public health strategies, including the Ending the HIV Epidemic initiative, for both HIV prevention and HIV treatment within the city.

Goal of This Study

This study aimed to develop predictive models for determining which patients are at risk for lapses in HIV care in Chicago using the Chicago Area Patient-Centered Outcomes Research Network (CAPriCORN). CAPriCORN contains EHR data from a diverse group of 11 different health care systems in Chicago, including academic medical centers, community hospitals, and community health clinic networks [11]. Importantly, CAPriCORN contains data for 12.8 million people, including the majority of 23,580 diagnosed people living with HIV in the city of Chicago [12,13]. Data from CAPriCORN allowed us to follow people living with HIV who obtain HIV care across the different CAPriCORN health care facilities in Chicago. We used these data to identify those engaged in HIV care and to build predictive models to identify people living with HIV at risk for lapses in HIV care within the city of Chicago.

Methods

Study Population

We used deidentified EHR data obtained from 7 sites in CAPriCORN, selecting records from adults (aged 18 years and older) with encounters between January 1, 2011, and September 5, 2019. The hospitals included in the CAPriCORN research database use several different EHR systems, including Epic, Cerner, GE Centricity, eClinicalWorks, Sunquest, and Veterans Health Information Systems and Technology Architecture [11]. CAPriCORN was responsible for all harmonization of data across different EHR systems. Data from CAPriCORN are linked by using a hash-based data deduplication method, allowing researchers to identify and follow patients across the different Chicago health care systems [11]. All EHR data were provided at the encounter level, allowing us to incorporate changes to sociodemographic, medical diagnoses, and laboratory information over time. Data contained sociodemographic information, encounter information, diagnosis codes, laboratory data, and antiretroviral therapy (ART) medication or prescription drug orders. We used the following criteria to determine if an individual was HIV-positive: (1) a positive HIV laboratory test (confirmatory HIV antibody, p24 antigen, or HIV viral load > 20 copies/mL), (2) an HIV viral load test performed (regardless of results) and prescription for ART, excluding pre-exposure prophylaxis prescriptions, (3) an HIV diagnosis code (International Classification of Diseases [ICD] version 9 [ICD-9] codes 42, 079.53, 795.71, and V08; ICD version 10 [ICD-10] codes B20, R75, and Z21) and a prescription for ART, excluding pre-exposure prophylaxis, or (4) an HIV diagnosis code and 2 HIV viral load tests performed [14]. Individuals were only included in the study if they met one (or more) of the above criteria and had at least two HIV care encounters. An HIV care encounter was defined as an ambulatory encounter that was either (1) with an infectious disease (ID) provider (taxonomy code “207RI0200X” or “2080P0208X”), or (2) associated with an HIV viral load test or CD4 T cell test. This definition allowed for the inclusion of both people living with HIV who receive their HIV care from their primary care physician as well as those who receive HIV care from their ID physician. To avoid double counting of care visits, we considered visits and laboratory tests within 7 days of each other to be part of the same encounter.

Ethics Approval

This study was approved by the Chicago Area Institutional Review Board (IRB #00009693).

Outcome

The outcome of interest in all models was a lapse in HIV care, defined as not having a second HIV care encounter within 1 year after the original HIV care encounter. Because people living with HIV with stable viral loads may have longer time periods between their appointments (eg, from 6 to 9 months), a 1-year gap was chosen to distinguish between these patients and people living with HIV who had truly lapsed in care [9,15,16]. People living with HIV were censored from the study when <12 months of study time remained to ensure that all participants had the opportunity to be identified as having a lapse in care.

Predictor Variables

We initially created a list of 134 variables for inclusion in the models; however, 20 variables were excluded in the final models due to multicollinearity between these variables and those already in the model. Final models were built using 114 variables including patient demographics, history of lapse in care, site, provider specialty, insurance information, laboratory tests and results, ART medications, and patient diagnoses. Medical diagnoses were based on the presence of the ICD-9 and ICD-10 diagnosis codes. Specifically, we examined diagnoses of cancer, cardiovascular disease, diabetes, mental health disorders (including mood and anxiety disorders), obesity, sexually transmitted infections (STIs; ie, chlamydia, gonorrhea, hepatitis B, hepatitis C, and syphilis), as well as the use of alcohol, cannabis, opioids, sedatives, stimulants, and tobacco. CAPriCORN provided racial demographic information, with patients able to select more than one race. Laboratory test variables included HIV viral load, CD4 count, HIV genotype, STI tests, and other ID tests within the preceding 12 months. The complete list of variables, including those excluded due to multicollinearity, as well as their ICD-9 or ICD-10 codes, is available in Multimedia Appendices 1 and 2.

Model Development

Encounter data were randomly split into a training set (70% of total data) and a validation set (30% of total data). The split was performed while holding a constant ratio of the outcome variable. We created a baseline logistic regression model using only patient demographics and previous history of lapses in care allowing us to quantify how much more predictive capability was provided by including additional data related to laboratory tests and diagnoses. We then created 4 different predictive models: logistic regression, elastic net logistic regression, random forest, and XGBoost. The hyperparameters for the elastic net, random forest, and XGBoost models were tuned using a grid search and can be found in Multimedia Appendix 3. The hyperparameters were optimized using 5-fold cross-validation on the metric of the area under the receiver operating characteristic curve (AUC).

Model Performance

The performance of these 4 models was compared with each other and with the baseline logistic regression model. Model performance was evaluated using AUC via DeLong’s [17] method. Variable influence for each model was evaluated using the absolute value of the *t* statistic for the baseline and full logistic regression models, the value of coefficients for the elastic net logistic regression, the Gini index for the random forest, and the relative influence for the XGBoost model [18,19]. Each model can provide the association of variables; however, unlike the baseline regression, full logistic regression, and elastic net logistic regression models, the random forest and XGBoost models we created cannot provide the directionality of association. All analyses were performed using R (version 4.0.3; R Core Team).

Results

Study Population

The study population contained 16,930 people living with HIV with a total of 191,492 HIV care encounters during the study period. When examining the demographic information for participants based on information provided at their last HIV care encounter, people living with HIV in this study were predominantly male (75.2%, 12,733/16,930), Black (61.4%, 10,395/16,930), non-Hispanic or -Latino (84.2%, 14,249/16,930), and had all HIV care encounters at a single site within CAPriCORN (95.3%, 16,135/16,930). Based on each participant’s last 2 HIV care encounters, 25.1% (4257/16,930) of participants had a lapse in care (Table 1). Over the study period, 42.1% (7135/16,930) of participants had a change in their lapse in HIV care status (either from engaged to lapsed in care or vice versa). The different care sites served a wide range of number of patients; the smallest site served 224 unique people living with HIV, and the largest site served 10,035 unique people living with HIV. The proportion of lapses in care varied among sites. The worst-performing site was Site F, with 21.4% (61/292) of encounters resulting in a lapse in care. The best-performing site was Site G, with 3.7% (5283/141,864) of encounters resulting in a lapse in care. Over the study period, 94.4% (180,777/191,492) of HIV care encounters were followed by a second HIV care encounter within 1 year.

Table 1. Characteristics of people living with HIV in the Chicago Area Patient-Centered Outcomes Research Network in Chicago, Illinois, 2011-2019 (N=16,930).

Variable of interest and values	Frequency, n (%)
Lapses in HIV care	
No lapse in care	12,673 (74.9)
Lapse in care	4257 (25.1)
Sex	
Male	12,733 (75.2)
Female	4147 (24.5)
Unknown	50 (0.3)
Race	
African American or Black	10,395 (61.4)
White	5168 (30.5)
Multiracial	136 (0.8)
Other	276 (1.6)
Unknown	955 (5.6)
Ethnicity	
Hispanic or Latino	2681 (15.8)
Non-Hispanic or -Latino	14,249 (84.2)
Age (years)	
Less than 25	868 (5.1)
25 to less than 35	3080 (18.2)
35 to less than 45	3520 (20.8)
45 to less than 55	4587 (27.1)
55 or older	4875 (28.8)
Primary payer type	
Private	3030 (17.9)
Medicare	2861 (16.9)
Medicaid	5162 (30.5)
Other	1031 (6.1)
Unknown	4846 (28.6)
Site	
Multisite	795 (4.7)
Single site	16,135 (95.3)
Comorbidities	
Hypertension	5128 (30.3)
Mental health	5757 (34.0)
Tobacco	4068 (24.0)
Mood disorder	3914 (23.1)
Alcohol	1347 (8.0)
Opioids	1114 (6.6)
Gonorrhea laboratory test in the past year	
Yes	2157 (12.7)
No	14,773 (87.3)

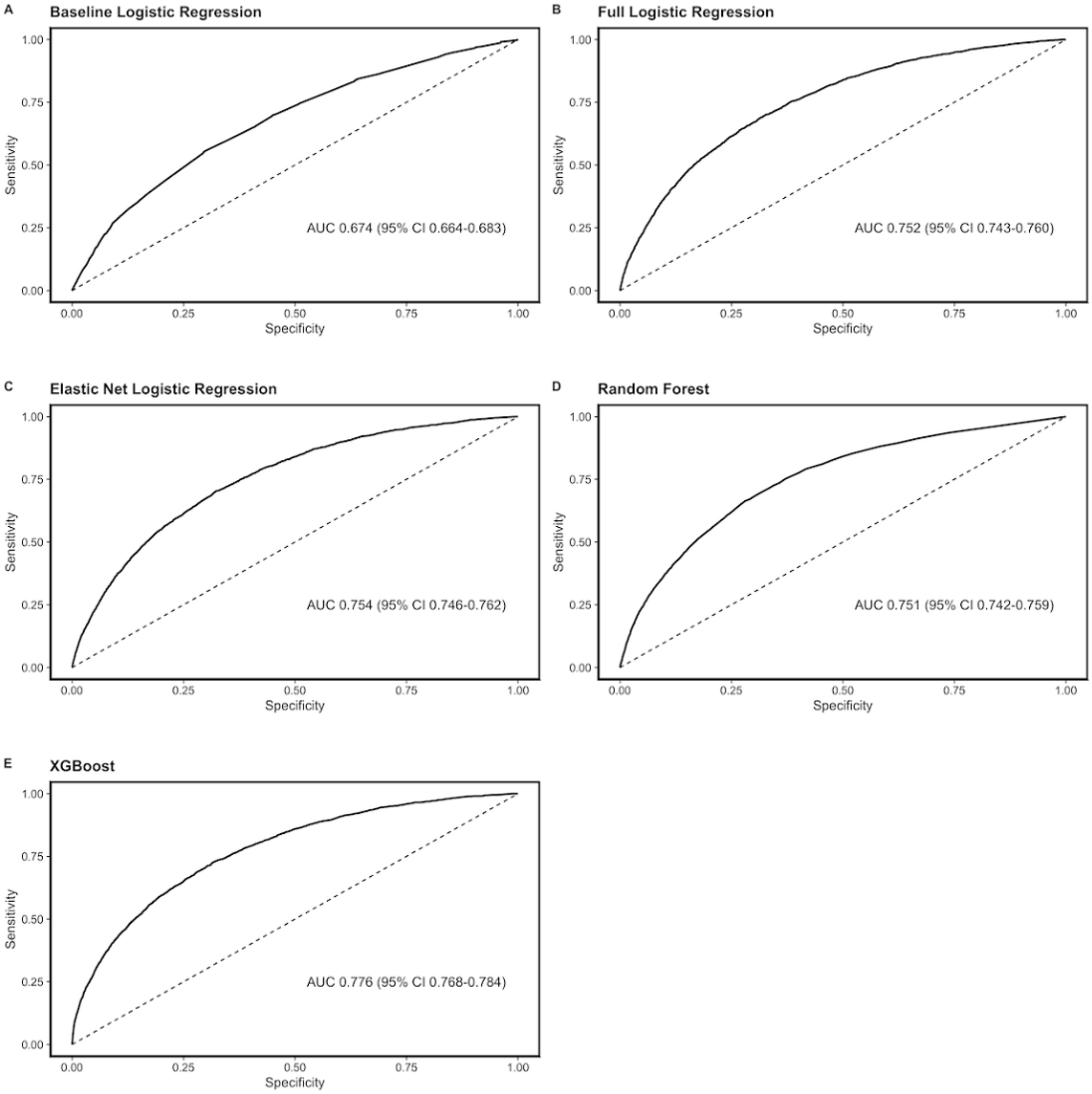
Variable of interest and values	Frequency, n (%)
Chlamydia laboratory test in the past year	
Yes	2419 (14.3)
No	14,511 (85.7)

Model Performance

Using our validation data set containing 57,447 encounters, all models created using 114 variables outperformed the baseline logistic regression model. The greatest improvement in AUC was seen with the XGBoost model (AUC 0.776, 95% CI 0.768-0.784 vs 0.674, 95% CI 0.664-0.683 for the baseline model; *P*<.001; Figure 1).

The XGBoost model had a significant difference in performance between the full logistic regression (AUC 0.776, 95% CI 0.768-0.784 vs 0.752, 95% CI 0.743-0.760; *P*<.001), elastic net logistic regression (AUC 0.776, 95% CI 0.768-0.784 vs 0.754, 95% CI 0.746-0.762; *P*<.001), and random forest (AUC 0.776, 95% CI 0.768-0.784 vs 0.751, 95% CI 0.742-0.759; *P*<.001) models. There was a significant difference in performance between the elastic net logistic regression model and the full logistic regression model (AUC 0.754, 95% CI 0.746-0.762 vs 0.752, 95% CI 0.743-0.760; *P*<.001); however, the magnitude of the difference between the 2 models was slight and may not be clinically meaningful. There was no significant difference in performance between the elastic net logistic regression model and the random forest model (AUC 0.754, 95% CI 0.746-0.762 vs 0.751, 95% CI 0.742-0.759; *P*=.43; Multimedia Appendix 4).

Figure 1. Prediction model receiver operator curves of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019. AUC: area under the receiver operating characteristic curve.



Predictor Variable Importance

The top 10 most influential predictors in each model are listed in Tables 2 and 3. Although each model had slightly different results, common predictors included no lapse in care in the prior year, being seen by an ID provider (vs a primary care provider), and Hispanic or Latino ethnicity. Influential predictors for the full logistic regression model were no lapse in care in the prior year (adjusted odds ratio [aOR] 0.42, 95% CI 0.39-0.46; *P*<.001), being seen by an ID provider (vs a primary care provider; aOR 0.54, 95% CI 0.49-0.58; *P*<.001), age of 55 years or older (vs 25 years to less than 35 years; aOR 0.56, 95% CI 0.51-0.61; *P*<.001), Hispanic ethnicity (aOR 0.57, 95% CI 0.52-0.63; *P*<.001), age of 45 years to <55 years (vs 25 years to less than 35 years; aOR 0.68, 95% CI 0.63-0.73; *P*<.001), and site F (aOR 6.39, 95% CI 3.89-10.47; *P*<.001).

Table 2. Most influential variables for baseline logistic regression and full logistic regression models of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019.

Model and variable	aOR[a] (95% CI)	*P* value
Baseline logistic regression		
No lapse in care in the prior 12 months	0.34 (0.32-0.36)	<.001
Hispanic ethnicity	0.32 (0.29-0.34)	<.001
Age of 55 years or older	0.53 (0.49-0.57)	<.001
Black race	0.41 (0.37-0.46)	<.001
Age of 45 years to less than 55 years	0.68 (0.63-0.73)	<.001
Unknown race	1.72 (1.52-1.94)	<.001
White race	0.63 (0.56-0.70)	<.001
Asian race	0.40 (0.31-0.51)	<.001
Age of 35 years to less than 45 years	0.86 (0.80-0.93)	<.001
Multiracial race	0.64 (0.41-0.99)	.047
Full logistic regression		
No lapse in care in the prior 12 months	0.42 (0.39-0.46)	<.001
Infectious disease provider	0.54 (0.49-0.58)	<.001
Age of 55 years or older	0.56 (0.51-0.61)	<.001
Hispanic ethnicity	0.57 (0.52-0.63)	<.001
Age of 45 years to less than 55 years	0.68 (0.63-0.73)	<.001
Site F	6.39 (3.89-10.47)	<.001
CD4 T cell percentage result greater than or equal to 28 in the prior 12 months	1.36 (1.25-1.48)	<.001
Gonorrhea laboratory test in the prior 12 months	0.35 (0.26-0.47)	<.001
Primary insurance unknown	0.56 (0.47-0.66)	<.001
Cancer diagnosis ever	0.78 (0.72-0.83)	<.001

[a] aOR: adjusted odds ratio.

Table 3. Top performing variables for elastic net logistic regression, random forest, and XGBoost models of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019.

Model and variable	Variable importance score
Elastic net logistic regression	
Site F	1.575
No lapse in care in the prior 12 months	–0.835
Chlamydia laboratory test in the prior 12 months	0.815
Site A	0.814
Gonorrhea laboratory test in the prior 12 months	–0.769
Site C	0.649
Infectious disease provider	–0.583
Hispanic ethnicity	–0.512
Primary insurance unknown	–0.506
Emergency medicine provider	0.504
Random forest	
Primary insurance Medicaid	216.8
No lapse in care in the prior 12 months	205.9
Hypertension diagnosis ever	200.0
Age 35 years to less than 45 years	195.0
Tobacco diagnosis ever	194.5
Age of 45 years to less than 55 years	193.0
Mental health diagnosis ever	190.5
CD4 T cell percentage result less than 28 in the prior 12 months	184.1
Mood disorder diagnosis ever	172.5
Syphilis diagnosis ever	171.4
XGBoost	
Site G	0.089
No lapse in care in the prior 12 months	0.065
Infectious disease provider	0.045
Site A	0.038
Previous viral load result suppressed	0.031
Age of 55 years or older	0.019
Hispanic ethnicity	0.019
Primary insurance Medicaid	0.019
CD4 T cell percentage result greater than or equal to 28 in the prior 12 months	0.018
HIV viral load laboratory result unsuppressed in the last 12 months	0.017

Discussion

Principal Results

Our study found that 25.1% (4257/16,930) of patients experienced a lapse in care based on their last eligible HIV care encounter, which is lower than the CDPH estimate of 33.8% of people living with HIV in Chicago not accessing HIV care [13]. Our lower percentage is likely due to two factors: (1) a slightly different patient care population and (2) different definitions being used to determine who was not in care. We defined a lapse in HIV care as not having a second HIV care encounter within 1 year after the original HIV care encounter; however, CDPH defines a lapse in HIV care as not having at least one HIV laboratory test (CD4 or viral load) within a year [13]. Additionally, we only calculated lapses in care among patients with at least two HIV care encounters, rather than among all people living with HIV within the jurisdiction of Chicago identified in the enhanced HIV/AIDS Reporting System used by CDPH [13]. By limiting our study population to people with at least two HIV care encounters, we may have selected patients with a history of retention in care, resulting in a patient population less likely to experience a lapse in care.

In this study, we developed predictive models for lapses in care among people living with HIV using linked data that incorporated information from multiple health care systems across Chicago. Our XGBoost model outperformed our baseline model, which included only information about demographics and past lapses in care. Across the different models, we assessed various HIV health care and STI-related factors in predicting the likelihood of lapses in HIV care. In all models, prior lapses in care were among the most influential predictors of future lapses in care. This observation aligns with other studies that have found previous retention in care to be important in predicting future retention [8,9,20]. Our models also showed that higher CD4 T cell counts are associated with an increased risk of lapsing in care, supporting previous findings [21-23]. A previous suppressed HIV viral load laboratory result and an unsuppressed HIV viral load laboratory result in the prior 12 months were found to be associated with lapsing in care; however, the directionality could not be determined from the XGBoost model. A gonorrhea laboratory test in the last year was predictive of a decreased risk of lapsing in care in the logistic regression model and the elastic net logistic regression model. A chlamydia laboratory test in the last year was predictive of an increased risk of lapsing in care in the elastic net logistic regression model.

In addition to HIV-related factors, we also examined the ability of sociodemographic characteristics to predict lapses in HIV care. Since CAPriCORN covers the majority of people living with HIV in Chicago, this is a unique opportunity to view how these factors impact lapses in care in Chicago. We found that younger age is a predictor for lapses in care, which has been seen in other studies [9,21,22,24-26]. Hispanic or Latino ethnicity was associated with a decreased risk of lapsing in care in this study, a finding supported by some previous literature but contradicted by others [9,21,22,25-28]. Another predictor for not lapsing in care was the female sex, consistent with previous studies [22,28].

In terms of general medical and health care factors, we found that people living with HIV with chronic comorbidities, particularly cancer, are less likely to lapse in care, as others have also found [23]. This lower risk of lapsing in care may result from people with other comorbidities being more likely to follow up with both their chronic disease health care provider as well as their ID provider [9,29]. The comorbidity of hypertension (having hypertension or being hypertension-free) was also associated with lapsing in care [9,21]. Unfortunately, this result comes from the random forest model, so we cannot assess the directionality of the association. Finally, despite examining data from a single city, we found that the HIV care site was associated with a higher risk of lapsing in care. The HIV care sites included in our study ranged from a small percentage to almost a fifth of encounters resulting in a lapse of care, which could be due to a myriad of clinic-level factors, such as different patient demographics, acceptance of multiple insurance types, provider familiarity and trust, different neighborhood locations, and varying levels of resources available for identifying patients at risk for lapses in care and reengaging them. Additionally, structural racism and disinvestment in certain neighborhoods in Chicago have disproportionately impacted marginalized communities and the clinics that serve these communities, which may also be reflected in clinic site differences [30]. Our findings regarding the impact of chronic comorbidity and clinic-level factors on risk for lapsing in care suggest that public health departments may benefit from using data from city-wide research networks like CAPriCORN to better understand which people living with HIV have access to HIV care and are engaged in care.

Consequently, our model could allow public health agencies to identify which groups of people living with HIV within Chicago are at the highest risk of lapsing in care, allowing resources to be targeted toward these patients *before* they lapse in care. The integration of research data with city- or county-based HIV surveillance data can provide a more complete picture of access to and engagement with HIV care than either data source alone. For instance, CDPH, like most city public health departments, can follow the entire population of people living with HIV in Chicago due to mandatory laboratory reporting of HIV tests. CDPH also has precise knowledge of who enters or leaves the jurisdiction. This information was unavailable to us in the CAPriCORN health network. However, CAPriCORN contains a wide range of information derived from the EHRs of patients, encompassing their entire range of medical illnesses and test results, instead of the public health departments' relatively limited HIV-related medical information. City health departments typically use viral load and CD4 count laboratory tests as proxies for HIV care visits, whereas we gathered information from both laboratory and clinic visits, allowing us to capture the full range of HIV care encounters and providing a unique way to look at the city-wide continuum of care. For instance, using the CAPriCORN data, we were able to determine that most patients receive care from a single site; only 4.7% (795/16,930) of patients received care from more than one site within the research network. This is less than a study that found that 8% of people living with HIV in Philadelphia received HIV care from multiple sites [31]. Having the ability to determine this important information about the care continuum demonstrates the benefits of incorporating data from EHRs through research networks like CAPriCORN into public health surveillance for HIV care.

Limitations

There are several limitations to this study. In order to deidentify the data set and protect patient privacy, CAPriCORN removed specific dates of encounters. Therefore, we used patient age in days to determine the relative time between encounters. The lack of dates also means that we cannot account for study-wide temporal trends. Some of our sites also had incomplete data. For example, 1 site was missing HIV laboratory data, resulting in no eligible HIV care encounters, meaning that we were unable to incorporate this site into our models. This incompleteness of data shows a larger limitation of city-wide research networks. It is critical for cities using our framework to focus on incorporating high-quality data, ensuring the most accurate and complete data are available for tracking engagement in HIV care. Additionally, we determined that an HIV care encounter included primary care visits in which an HIV viral load test or

CD4 test were performed. However, it is possible that some primary care encounters were for HIV care but did not include performing an HIV viral load test or CD4 test leading to misclassifications. Similarly, we focused on 1 definition of a lapse in HIV care as our outcome, but there are multiple standards for defining a lapse in HIV care [32]. Future studies should compare different standards, such as the Health Resources and Services Administration HIV/AIDS Bureau measure [33]. As we did not have access to individual patient records for review of medical records, we used ICD-9/10 codes to identify diagnoses, which might lack sensitivity and specificity in identifying patients with medical comorbidities. However, ICD-9/10 codes provide a ubiquitous way for researchers to build predictive models using easily accessible data even if they lack the resources to carry out a review of medical records for every patient. Previous studies have shown that using ICD-9/10 codes can be effective at scaling up the rapid identification of people living with HIV [14,34]. Future studies should include social history variables and unstructured fields or free-text notes to see if they improve overall model performance. This model is possible because of the existence of the CAPriCORN research network. For this work to be extended in other settings, a deidentified, centralized data source from multiple hospital systems would need to be created in other jurisdictions. Our study is retrospective, so the performance of the model in a prospective setting should be tested in future studies prior to implementation.

Conclusions

We built predictive models for lapses in HIV care for people living with HIV from 7 different health care systems across Chicago and identified health care factors found in structured EHR data that are important for predicting lapses in care. Our models also serve as a proof of concept to show that differences in site-based practices (ie, laboratory tests ordered, diagnosis codes, different EHR systems, etc) do not prevent the ability to develop predictive models for lapses in care within a city. To our knowledge, this is the first time predictive models of lapses in HIV care have been built using EHR data from a large research network of health care systems within a major urban city with a high HIV prevalence rate. By focusing on data from patients served by a single public health department jurisdiction, findings from this study may inform HIV public health strategies for the city of Chicago. This model framework can also be adapted and used in clinics to identify patients most at risk for lapses in care who would benefit from tailored interventions or resources, thereby mitigating the number of people living with HIV who fall out of care and supporting efforts to end the HIV epidemic.

Acknowledgments

The authors wish to thank CAPriCORN for providing the data used in this publication. This work was supported by the National Institutes of Health (K23 MH121190).

Conflicts of Interest

JPR reports receiving fees from Gilead Sciences for consulting. EEF receives funding from the Gilead Sciences Focus grant as well as the Ryan White grant.

Multimedia Appendix 1
Clinical variables included in prediction models of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019.
[XLSX File (Microsoft Excel File), 54 KB - publichealth_v9i1e43017_app1.xlsx]

Multimedia Appendix 2
ART codes included in prediction models of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019.
[XLSX File (Microsoft Excel File), 14 KB - publichealth_v9i1e43017_app2.xlsx]

Multimedia Appendix 3
Hyperparameter values used in prediction models of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019.
[DOCX File , 15 KB - publichealth_v9i1e43017_app3.docx]

Multimedia Appendix 4
Sensitivity and specificity values at different cut-off points for prediction models of lapses in care for people living with HIV in Chicago, Illinois, 2011-2019.
[DOCX File , 16 KB - publichealth_v9i1e43017_app4.docx]

References

1. Li Z, Purcell DW, Sansom SL, Hayes D, Hall HI. Vital Signs: HIV Transmission Along the Continuum of Care - United States, 2016. MMWR Morb Mortal Wkly Rep 2019 Mar 22;68(11):267-272 [FREE Full text] [doi: 10.15585/mmwr.mm6811e1] [Medline: 30897075]

2. Ending the HIV epidemic: HIV treatment is prevention. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/vitalsigns/end-hiv/index.html [accessed 2022-03-24]
3. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data United States and 6 dependent areas, 2019. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance/vol-26-no-2/index.html [accessed 2022-03-24]
4. Department of Health and Human Services. Ending the HIV epidemic. HIV.gov. 2021. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview [accessed 2022-03-29]
5. Davila JA, Miertschin N, Sansgiry S, Schwarzwald H, Henley C, Giordano TP. Centralization of HIV services in HIV-positive African-American and Hispanic youth improves retention in care. AIDS Care 2013;25(2):202-206. [doi: 10.1080/09540121.2012.689811] [Medline: 22708510]
6. Sweeney P, DiNenno EA, Flores SA, Dooley S, Shouse RL, Muckleroy S, et al. HIV data to care-using public health data to improve HIV care and prevention. J Acquir Immune Defic Syndr 2019;82(suppl 1):S1-S5 [FREE Full text] [doi: 10.1097/QAI.0000000000002059] [Medline: 31425388]
7. Chang EJ, Fleming M, Nunez A, Dombrowski JC. Predictors of successful HIV care re-engagement among persons poorly engaged in HIV care. AIDS Behav 2019;23(9):2490-2497 [FREE Full text] [doi: 10.1007/s10461-019-02491-y] [Medline: 30980279]
8. Ramachandran A, Kumar A, Koenig H, De Unanue A, Sung C, Walsh J, et al. Predictive analytics for retention in care in an urban HIV clinic. Sci Rep 2020;10(1):6421 [FREE Full text] [doi: 10.1038/s41598-020-62729-x] [Medline: 32286333]
9. Ridgway JP, Ajith A, Friedman EE, Mugavero MJ, Kitahata MM, Crane HM, et al. Multicenter development and validation of a model for predicting retention in care among people with HIV. AIDS Behav 2022;26(10):3279-3288 [FREE Full text] [doi: 10.1007/s10461-022-03672-y] [Medline: 35394586]
10. Jurisdictions: ending the HIV Epidemic. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/endhiv/jurisdictions.html [accessed 2022-05-18]
11. Kho AN, Hynes DM, Goel S, Solomonides AE, Price R, Hota B, CAPriCORN Team. CAPriCORN: Chicago area patient-centered outcomes research network. J Am Med Inform Assoc 2014;21(4):607-611 [FREE Full text] [doi: 10.1136/amiajnl-2014-002827] [Medline: 24821736]
12. CAPriCORN executive summary February 2022. CAPriCORN. 2022. URL: https://www.capricorncdrn.org/wp-content/uploads/2022/02/CAPriCORN-executive-summary-february2022.pdf [accessed 2022-07-27]
13. HIV/STI surveillance report. Chicago Department of Public Health. 2020. URL: https://www.chicago.gov/content/dam/city/depts/cdph/statistics_and_reports/2019_HIV_STI_REPORT_Final.pdf [accessed 2023-04-26]
14. Ridgway JP, Mason JA, Friedman EE, Devlin S, Zhou J, Meltzer D, et al. Comparison of algorithms for identifying people with HIV from electronic medical records in a large, multi-site database. JAMIA Open 2022;5(2):ooac033 [FREE Full text] [doi: 10.1093/jamiaopen/ooac033] [Medline: 35651521]
15. Edwards JK, Cole SR, Westreich D, Moore R, Mathews C, Geng E, CNICS Research Network. Loss to clinic and five-year mortality among HIV-infected antiretroviral therapy initiators. PLoS One 2014;9(7):e102305 [FREE Full text] [doi: 10.1371/journal.pone.0102305] [Medline: 25010739]
16. Laboratory testing schedule for monitoring people with HIV before and after initiation of antiretroviral therapy. National Institutes of Health. URL: https://clinicalinfo.hiv.gov/en/table/table-3-laboratory-testing-schedule-monitoring-people-hiv-and-after-initiation-antiretroviral [accessed 2022-06-16]
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837-845 [FREE Full text] [doi: 10.2307/2531595] [Medline: 3203132]
18. Breiman L. Random forests. Mach Learn 2011;45(1):5-32 [FREE Full text] [doi: 10.1023/A:1010933404324]
19. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001;29(5):1189-1232 [FREE Full text] [doi: 10.1214/aos/1013203451]
20. Pence BW, Bengtson AM, Boswell S, Christopoulos KA, Crane HM, Geng E, et al. Who will show? Predicting missed visits among patients in routine HIV primary care in the United States. AIDS Behav 2019;23(2):418-426 [FREE Full text] [doi: 10.1007/s10461-018-2215-1] [Medline: 30006790]
21. Giordano TP, Hartman C, Gifford AL, Backus LI, Morgan RO. Predictors of retention in HIV care among a national cohort of US veterans. HIV Clin Trials 2009;10(5):299-305 [FREE Full text] [doi: 10.1310/hct1005-299] [Medline: 19906622]
22. Fleishman JA, Yehia BR, Moore RD, Korthuis PT, Gebo KA, HIV Research Network. Establishment, retention, and loss to follow-up in outpatient HIV care. J Acquir Immune Defic Syndr 2012;60(3):249-259 [FREE Full text] [doi: 10.1097/QAI.0b013e318258c696] [Medline: 22531758]
23. Cunningham CO, Buck J, Shaw FM, Spiegel LS, Heo M, Agins BD. Factors associated with returning to HIV care after a gap in care in New York state. J Acquir Immune Defic Syndr 2014;66(4):419-427 [FREE Full text] [doi: 10.1097/QAI.0000000000000171] [Medline: 24751434]
24. Catz SL, McClure JB, Jones GN, Brantley PJ. Predictors of outpatient medical appointment attendance among persons with HIV. AIDS Care 1999;11(3):361-373 [FREE Full text] [doi: 10.1080/09540129947983] [Medline: 10474634]

25. Rebeiro P, Althoff KN, Buchacz K, Gill J, Horberg M, Krentz H, North American AIDS Cohort Collaboration on Research and Design. Retention among North American HIV-infected persons in clinical care, 2000-2008. J Acquir Immune Defic Syndr 2013;62(3):356-362 [FREE Full text] [doi: 10.1097/QAI.0b013e31827f578a] [Medline: 23242158]
26. Israelski D, Gore-Felton C, Power R, Wood MJ, Koopman C. Sociodemographic characteristics associated with medical appointment adherence among HIV-seropositive patients seeking treatment in a county outpatient facility. Prev Med 2001;33(5):470-475 [FREE Full text] [doi: 10.1006/pmed.2001.0917] [Medline: 11676589]
27. Sheehan DM, Fennie KP, Mauck DE, Maddox LM, Lieb S, Trepka MJ. Retention in HIV care and viral suppression: individual- and neighborhood-level predictors of racial/ethnic differences, Florida, 2015. AIDS Patient Care STDS 2017;31(4):167-175 [FREE Full text] [doi: 10.1089/apc.2016.0197] [Medline: 28414260]
28. Myers K, Li T, Baum M, Ibanez G, Fennie K. The individual, interactive, and syndemic effect of substance use, depression, education, and ethnicity on retention in HIV care. Int J STD AIDS 2021;32(2):184-193 [FREE Full text] [doi: 10.1177/0956462419890727] [Medline: 33323072]
29. Pecoraro A, Royer-Malvestuto C, Rosenwasser B, Moore K, Howell A, Ma M, et al. Factors contributing to dropping out from and returning to HIV treatment in an inner city primary care HIV clinic in the United States. AIDS Care 2013;25(11):1399-1406 [FREE Full text] [doi: 10.1080/09540121.2013.772273] [Medline: 23428205]
30. Smith JD, Davis P, Kho AN. Community-driven health solutions on Chicago's South Side. Stanf Soc Innov Rev 2021;19(3):A27-A29 https://ssir.org/articles/entry/community_driven_health_solutions_on_chicagos_south_side [FREE Full text] [doi: 10.48558/85p7-3113] [Medline: 34744527]
31. Yehia BR, Schranz AJ, Momplaisir F, Keller SC, Gross R, Frank I, et al. Outcomes of HIV-infected patients receiving care at multiple clinics. AIDS Behav 2014;18(8):1511-1522 [FREE Full text] [doi: 10.1007/s10461-013-0625-7] [Medline: 24077931]
32. Mugavero MJ, Westfall AO, Zinski A, Davila J, Drainoni ML, Gardner LI, Retention in Care (RIC) Study Group. Measuring retention in HIV care: the elusive gold standard. J Acquir Immune Defic Syndr 2012;61(5):574-580 [FREE Full text] [doi: 10.1097/QAI.0b013e318273762f] [Medline: 23011397]
33. Mugavero MJ, Davila JA, Nevin CR, Giordano TP. From access to engagement: measuring retention in outpatient HIV clinical care. AIDS Patient Care STDS 2010;24(10):607-613 https://www.liebertpub.com/doi/10.1089/apc.2010.0086 [FREE Full text] [doi: 10.1089/apc.2010.0086] [Medline: 20858055]
34. Paul DW, Neely NB, Clement M, Riley I, Al-Hegelan M, Phelan M, et al. Development and validation of an electronic medical record (EMR)-based computed phenotype of HIV-1 infection. J Am Med Inform Assoc 2018;25(2):150-157 [FREE Full text] [doi: 10.1093/jamia/ocx061] [Medline: 28645207]

Abbreviations

aOR: adjusted odds ratio
ART: antiretroviral therapy
AUC: area under the receiver operating characteristic curve
CAPriCORN: Chicago Area Patient-Centered Outcomes Research Network
CDPH: Chicago Department of Public Health
EHR: electronic health record
ICD: International Classification of Diseases
ID: infectious disease
STI: sexually transmitted infection

Edited by Y Khader; submitted 27.09.22; peer-reviewed by P Han, D Gopukumar; comments to author 23.01.23; revised version received 01.03.23; accepted 11.04.23; published 17.05.23.

Please cite as:
Mason JA, Friedman EE, Devlin SA, Schneider JA, Ridgway JP
Predictive Modeling of Lapses in Care for People Living with HIV in Chicago: Algorithm Development and Interpretation
JMIR Public Health Surveill 2023;9:e43017
URL: https://publichealth.jmir.org/2023/1/e43017
doi:10.2196/43017
PMID:37195750

Original Paper

Prevalence and Trends of Sexual Behaviors Among Young Adolescents Aged 12 Years to 15 Years in Low and Middle-Income Countries: Population-Based Study

Zhengyue Jing[1,2], PhD; Jie Li[2], PhD; Yi Wang[2], PhD; Chengchao Zhou[2,3,4], PhD

[1]Department of Social Security, School of Health Policy and Management, Nanjing Medical University, Nanjing, China
[2]Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[3]National Health Commission of China Key Lab of Health Economics and Policy Research, Shandong University, Jinan, China
[4]Institute of Health and Elderly Care, Shandong University, Jinan, China

Corresponding Author:
Chengchao Zhou, PhD
Centre for Health Management and Policy Research
School of Public Health, Cheeloo College of Medicine
Shandong University
44 Wen-hua-xi Road
Jinan, 250012
China
Phone: 86 053188381567
Fax: 86 053188382553
Email: zhouchengchao@sdu.edu.cn

Abstract

Background: Risky sexual behaviors remain significant public health challenges among adolescents. Nearly 90% of adolescents live in low and middle-income countries (LMICs), but few studies have used standardized methodology to monitor the prevalence and trends of sexual behaviors among adolescents in LMICs.

Objective: This study aimed to assess the prevalence of sexual behaviors (ever had sexual intercourse, multiple sexual partners, and condom use) among adolescents aged 12 years to 15 years as well as the trends in prevalence between 2003 and 2017.

Methods: For this population-based study, we used recent data from the Global School-based Student Health Survey conducted in 69 LMICs from 2003 to 2017 to assess the recent prevalence of sexual behaviors by using complex analysis and a random effects meta-analyses method. Using the chi-square trend test, we also assessed the trends in the prevalence of sexual behaviors in 17 countries that had conducted ≥1 round of surveys from 2003 to 2017.

Results: We included 145,277 adolescents aged 12 years to 15 years (64,719/145,277, 44.5% boys) from the 69 LMICs that had conducted ≥1 survey and 80,646 adolescents aged 12 years to 15 years (34,725/80,646, 43.1% boys) from the 17 LMICs that had conducted ≥1 round of surveys. The recent global prevalence of ever had sexual intercourse was 6.9% (95% CI 6.2%-7.6%) and was higher among boys (10.0%, 95% CI 9.1%-11.1%) than girls (4.2%, 95% CI 3.7%-4.7%) and among those aged 14 years to 15 years (8.5%, 95% CI 7.7%-9.3%) than those aged 12 years to 13 years (4%, 95% CI 3.4%-4.7%). Among adolescents who had ever had sex, the recent global prevalence of having multiple sexual partners was 52% (95% CI 50.4%-53.6%) and was higher among boys (58%, 95% CI 56.1%-59.9%) than girls (41.4%, 95% CI 38.9%-43.9%) and among those aged 14 years to 15 years (53.5%, 95% CI 51.6%-55.4%) than those aged 12 years to 13 years (49.7%, 95% CI 45.9%-53.5%). Among adolescents who had ever had sex, the recent global prevalence of condom use was 58.1% (95% CI 56.2%-59.9%) and was higher among girls (59.2%, 95% CI 56.4%-61.9%) than boys (57.7%, 95% CI 55.7%-59.7%) and among those aged 14 years to 15 years (59.9%, 95% CI 58.0%-61.8%) than those aged 12 years to 13 years (51.6%, 95% CI 47.5%-55.7%). Between the earliest and latest surveys, the overall prevalence of ever had sexual intercourse (3.1% decrease) and condom use (2.0% decrease) showed downward trends. The overall prevalence of having multiple sexual partners increased by 2.6%.

Conclusions: We provide evidence and important implication for policymakers to develop targeted policy support systems to prevent and reduce risky sexual behaviors among young adolescents in LMICs with a high prevalence of risky sexual behaviors.

(JMIR Public Health Surveill 2023;9:e45236) doi:10.2196/45236

KEYWORDS

risky sexual behaviors; early sexual intercourse; multiple sexual partners; condom use; young adolescents; low and middle-income countries

Introduction

Adolescence is a period of rapid change in physical, psychological, emotional, and cognitive development, during which adolescents become more interested in sexual behaviors. Although sexual behaviors are widely considered a normative and physiological component of adolescent development, risky sexual behaviors including early initiation of sexual intercourse, having multiple sexual partners, and condom nonuse remain significant public health challenges among adolescents due to their potentially deleterious effects on later sexual and reproductive health [1]. Sexual debut is not deterministic of later sexual activities and risks, but sexual intercourse initiated at an earlier than normative age (typically defined as 15 years or younger) exposes adolescents, particularly adolescent girls, to a variety of risks such as HIV infection and other sexually transmitted infections (STIs) [2,3]. Girls who initiate sexual intercourse early were found to be at higher risk of reporting unintended pregnancies and STIs than those who delay intercourse until late adolescence [4,5]. Early sexual behavior is a public health concern among adolescents both in low and middle-income countries (LMICs) and high-income countries (HICs), but the adverse consequences of early sexual debut for adolescents in LMICs are more severe than for those in HICs. For example, previous studies have found that adolescent birth rates in LMICs are more than double that of HICs and most STIs occur in LMICs [6,7]. In addition, early sexual intercourse has been associated with reporting negative social and psychological outcomes (such as suicidal behaviors) [8] and subsequent higher-risk behaviors including alcohol or drug use during sex, lower levels of condom use, and forced sex [9,10].

Having multiple sexual partners is a common practice among young people. A review of data from the Global School-based Student Health Survey (GSHS) in 21 countries between 2010 and 2016 found that 53.1% of adolescents aged 12 years to 15 years who had sexual intercourse reported having multiple sexual partners [11]. Having multiple sexual partners is a risky sexual behavior that increases the risk of HIV and other STI transmission among adolescents [12]. A previous study indicated that the high prevalence of HIV/AIDS in sub-Saharan Africa is driven by high levels of multiple sexual partnerships [13].

Condom use has been acknowledged as an effective way to reduce the risk of gonorrhea, herpes simplex virus type 2, syphilis, and other STIs [14]. A review indicated that consistent use of condoms could effectively reduce the incidence of HIV/AIDS by 80% [15]. Meanwhile, condom use was associated with a lower risk of unintended pregnancy among adolescents, and the failure rate of male condoms for unintended pregnancy was about 2% [16]. Despite this, the prevalence of condom use among adolescents remains low. A study reported that nearly two-thirds of adolescents aged 15 years to 21 years did not use condoms during their last sexual intercourse [17].

Healthy sexuality is a key component of adolescent development. Nearly 90% of adolescents live in LMICs, and monitoring the prevalence and trends of sexual behaviors among adolescents in LMICs can help public health and education sectors design appropriate prevention and intervention strategies to promote sexual and reproductive health in adolescents. Although previous studies have reported the prevalence of sexual behaviors among adolescents aged 13 years to 15 years in some countries [18], as more countries have been released in the GSHS data set, there is an urgent need to analyze more up-to-date data and examine the difference in the prevalence of sexual behaviors among young adolescents. Differences in the prevalence of adolescent sexual behavior across countries may be influenced by race, religion, the society's culture, and economic status. Age and gender are also important factors to consider. However, few studies have used standardized questionnaires to examine and compare the age and gender differences in the prevalences of risky sexual behaviors among young adolescents in LMICs. More importantly, the most effective policies to prevent risky sexual behaviors may change over time; therefore, identifying trends in the prevalence of sexual behaviors is important for these countries to understand the effectiveness of previously developed policies in combating risky sexual behaviors among young adolescents and thus adjust intervention policies to fit the changes in policy needs and impacts. Therefore, it is imperative to analyze trends in the prevalences of risky sexual behaviors among young adolescents to provide evidence for the development of intervention programs. However, to our knowledge, few previous studies have assessed recent trends in the prevalences of risky sexual behaviors among young adolescents in LMICs.

Therefore, this study aimed to use recent data collected in the GSHS from 2003 to 2017 to assess the prevalences of sexual behaviors including ever had sexual intercourse, multiple sexual partners, and condom use as well as their age and gender differences among adolescents aged 12 years to 15 years in 69 LMICs. We also aimed to evaluate trends in the prevalences of sexual behaviors among adolescents in 17 LMICs between 2003 and 2017.

Methods

Study Design and Participants

This study used the latest GSHS data (2003-2017) publicly available on the websites of the US Centers for Disease Control and Prevention (CDC) and World Health Organization (WHO). Developed by the WHO and US CDC, the GSHS is a school-based self-administered survey for young adolescents aged 12 years to 15 years. The goal of the GSHS is to assess the health behavior risks and protective factors of middle-school students across countries and ultimately help countries develop and provide health care programs, resources, and policies that promote the health of adolescents [19]. To ensure the comparability of data across countries, the GSHS survey used the same standardized procedure (including sampling strategy, study methodology, and questionnaire) in each country; the GSHS questionnaire was translated into the local language, and each country was free to select question modules.

The selection of respondents in each country participating in the GSHS was based on a 2-stage random cluster sampling procedure. In the first stage, schools were randomly selected from all middle schools in each country by using the probability proportionate to size sampling method. In the second stage, classes were randomly selected from each selected school. All the students in each selected class were eligible to participate in this survey. Finally, based on the available data on sexual behaviors in the GSHS database, for the analysis of the prevalence of sexual behaviors, this study used the most recent GSHS data from 69 countries in 6 WHO regions (including 17 from Africa, 30 from the Americas, 1 from the Eastern Mediterranean, 2 from Europe, 5 from Southeast Asia, and 14 from the Western Pacific) that had conducted at least one survey between 2003 and 2017. To analyze the trends in the prevalences of sexual behaviors, we used the GSHS data from 17 countries in 4 WHO regions (5 from Africa, 7 from the Americas, 2 from Southeast Asia, and 3 from the Western Pacific) that had conducted more than one round of surveys between the earliest and latest surveys.

Outcomes and Definitions

Sexual intercourse was measured in this study using the question: “Have you ever had sexual intercourse?” The answer was dichotomized as “Yes” or “No.” Sexual partners were measured in this study using the question: “During your life, with how many people have you had sexual intercourse?” Multiple sexual partners were defined as having 2 or more sexual partners during their lifetime. Condom use was measured in this study using the question: “The last time you had sexual intercourse, did you or your partner use a condom?” The answer was also dichotomized as “Yes” or “No.” The prevalences of multiple sexual partners and condom use were calculated for those having sexual intercourse.

Statistical Analysis

The complex sampling command in SPSS Version 22.0 (IBM Corp) was used to conduct the statistical analysis. To adjust the sampling survey method and the differences between sampled students and national students, all complex sample analyses in this study were performed using 3 weighted variables included in each GSHS data set: strata, primary sampling unit, and weights. Weighted prevalences and corresponding 95% CIs of ever had sexual intercourse, having had multiple sexual partners, and condom use were calculated by region and country and by sex and age. ArcGIS software was used to map the prevalence of sexual behaviors among adolescents in each country. Chi-square analysis was used to test for differences in the prevalences between sexes and age groups, and *P* values <.05 indicated that the difference was statistically significant. The chi-square trend test was used to test the secular trend in the prevalences between the earliest survey and latest survey, and *P* values <.05 represented significant downward or upward trends in prevalences over time. Meanwhile, due to the significant heterogeneity between countries, we used the random effects model to calculate the pooled overall and regional prevalences of ever had sexual intercourse through the meta-analysis module in Stata Version 11.0.

Ethical Considerations

The ethics committee (usually the Ministry of Health or Education) of each country included in the GSHS survey reviewed and approved the study protocol. This survey was anonymous, and the answers were protected by privacy law. All participants and their parents gave their informed written or verbal consent for participation prior to the survey.

Results

Participants

As shown in Table 1 145,277 adolescents aged 12 years to 15 years were included from 69 countries in 6 WHO regions that had conducted at least one survey between 2003 and 2017. The sample sizes ranged from 218 in Nauru to 18,031 in Argentina. A total of 80,646 adolescents aged 12 years to 15 years were included from 17 countries in 4 WHO regions that had conducted more than one round of surveys between 2003 and 2017.

Table 1. Survey characteristics of the Global School-based Student Health Surveys of adolescents aged 12 years to 15 years by country, 2003-2017.

Country	Survey year	Total survey sample, N	Response rate, n (%)
Africa			
Benin	2016	712	625 (87.8)
Botswana	2005	1391	1016 (73.0)
Eswatini			
	2003	6664	4558 (68.4)
	2013	1314	1089 (82.9)
Ghana			
	2007	4235	2520 (59.5)
	2012	1330	861 (64.7)
Liberia	2017	529	298 (56.3)
Malawi	2009	2185	1649 (75.5)
Mauritania	2010	1272	865 (68.0)
Mauritius			
	2011	2071	2032 (98.1)
	2017	1946	1515 (77.9)
Mozambique	2015	652	445 (68.3)
Namibia			
	2004	4492	2226 (49.6)
	2013	1918	1385 (72.2)
Senegal	2005	2633	2096 (79.6)
Seychelles			
	2007	1146	861 (75.1)
	2015	2060	1535 (74.5)
Sierra Leone	2017	1835	1101 (60)
Tanzania	2014	2590	1979 (76.4)
Uganda	2003	1868	1238 (66.3)
Zambia	2004	1315	382 (29)
Zimbabwe	2003	3877	2569 (66.3)
Americas			
Anguilla			
	2009	696	546 (78.4)
	2016	571	448 (78.4)
Antigua and Barbuda	2009	1198	1028 (85.8)
Argentina			
	2007	1523	1292 (84.8)
	2012	21,620	18,031 (83.4)
Bahamas	2013	1304	1034 (79.3)
Barbados	2011	1502	1240 (82.6)
Belize	2011	1597	1345 (84.2)
Bolivia	2012	2761	2318 (83.9)
British Virgin Islands	2009	1191	1008 (84.6)

Country	Survey year	Total survey sample, N	Response rate, n (%)
Cayman	2007	1265	915 (72.3)
Chile	2013	1342	1182 (88.1)
Colombia	2007	7963	7187 (90.3)
Costa Rica	2009	2259	2023 (89.6)
Curaçao	2015	1491	1261 (84.6)
Dominica	2009	1308	1047 (80)
Ecuador	2007	4508	3711 (82.3)
El Salvador	2013	1600	1438 (89.9)
Grenada	2008	1296	1018 (78.5)
Guatemala			
	2009	4461	4002 (89.7)
	2015	3591	2880 (80.2)
Guyana			
	2004	1060	855 (80.7)
	2010	1958	1519 (77.6)
Honduras	2012	1474	1253 (85)
Jamaica	2017	1057	885 (83.7)
Paraguay	2017	1972	1720 (87.2)
Peru	2010	2350	2260 (96.2)
Saint Kitts and Nevis	2011	1463	1206 (82.4)
Saint Lucia	2007	1070	926 (86.5)
Saint Vincent and the Grenadines	2007	1184	956 (80.7)
Suriname			
	2009	1043	934 (89.5)
	2016	1448	1300 (89.8)
Trinidad and Tobago			
	2011	2352	1961 (83.4)
	2017	2749	2205 (80.2)
Uruguay			
	2006	2876	2542 (88.4)
	2012	2855	2587 (90.6)
Venezuela	2003	3901	3052 (78.2)
Eastern Mediterranean			
Djibouti	2007	961	655 (68.2)
Europe			
Macedonia	2007	1538	1364 (88.7)
Tajikistan	2006	7457	5789 (77.6)
Southeast Asia			
Bhutan	2016	3268	2873 (87.9)
Indonesia			
	2007	3013	2784 (92.4)

Country	Survey year	Total survey sample, N	Response rate, n (%)
	2015	11,063	6992 (63.2)
Nepal	2015	4565	3392 (74.3)
Thailand			
	2008	2671	2309 (86.4)
	2015	4120	3399 (82.5)
Timor-Leste	2015	2053	961 (46.8)
Western Pacific			
Brunei Darussalam	2014	1822	1611 (88.4)
Cambodia	2013	1809	1471 (81.3)
Fiji			
	2010	1491	1229 (82.4)
	2016	1512	1184 (78.3)
French Polynesia	2015	1898	1678 (88.4)
Kiribati	2011	1337	1155 (86.4)
Lao People's Democratic Republic	2015	1636	1453 (88.8)
Malaysia	2011	16,248	13,880 (85.4)
Mongolia	2013	3695	3303 (89.4)
Nauru	2011	361	218 (60.4)
Samoa			
	2011	2153	676 (31.4)
	2017	1054	707 (67.1)
Vanuatu			
	2011	844	782 (92.7)
	2016	1278	900 (70.4)
Tuvalu	2013	675	455 (67.4)
Viet Nam	2013	1743	1605 (92.1)
Wallis and Futuna	2015	713	601 (84.3)
All countries combined			
More than 1 survey	—[a]	107,528	80,646 (75)
Most recent survey	—	182,051	145,277 (79.8)

[a] Not applicable.

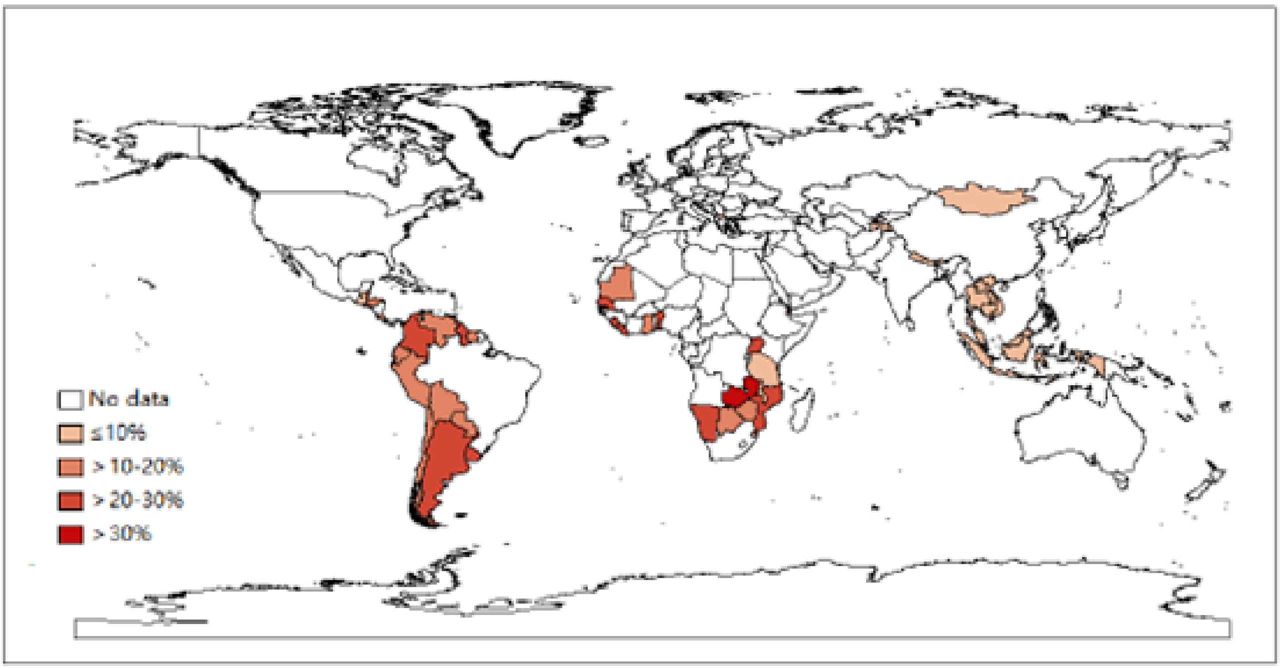
Prevalences of Sexual Behaviors in Young Adolescents

The overall prevalence of ever had sexual intercourse was 6.9% (95% CI 6.2%-7.6%) and was lowest in the Western Pacific and highest in the Americas (Table 2). Both gender and age differences were statistically significant in Africa, the Americas, and Southeast Asia. In Table S1 in Multimedia Appendix 1, nearly two-thirds (47 countries, 68%) of the 69 countries had a prevalence of ever had sexual intercourse >10%, while those in Europe and South Asia had a lower prevalence of <10% (Figure 1). Almost all countries had a higher prevalence for boys than girls, as well as for those aged 14 years to 15 years than those aged 12 years to 13 years.

Table 2. The prevalence of ever had sexual intercourse among young adolescents aged 12 years to 15 years by World Health Organization region, sex, and age in 2003-2017.

Region	Total sample, % (95% CI)	Boys, % (95% CI)	Girls, % (95% CI)	*P* value	12-13 years old, % (95% CI)	14-15 years old, % (95% CI)	*P* value
Total	6.9 (6.2-7.6)	10.0 (9.1-11.1)	4.2 (3.7-4.7)	<.001	4.0 (3.4-4.7)	8.5 (7.7-9.3)	<.001
Africa	12.9 (11.4-14.6)	20.1 (17.7-22.9)	6.2 (5.2-7.3)	<.001	8.7 (7.0-10.6)	15.2 (13.5-17.2)	<.001
Americas	17.8 (16.9-18.9)	24.4 (23.0-25.8)	12.1 (11.1-13.1)	<.001	9.5 (8.6-10.4)	22.2 (21.0-23.5)	<.001
Eastern Mediterranean	15.0 (12.0-18.6)	22.2 (17.9-27.1)	4.8 (3.1-7.5)	<.001	10.3 (4.5-21.9)	16.0 (13.3-19.0)	.24
Europe	3.0 (2.3-4.0)	5.1 (3.8-6.8)	0.9 (0.5-1.5)	<.001	2.4 (1.4-4.1)	3.2 (2.4-4.4)	.30
Southeast Asia	1.8 (1.4-2.3)	2.6 (2.0-3.4)	1.1 (0.8-1.6)	<.001	1.1 (0.7-1.6)	2.5 (2.0-3.1)	<.001
Western Pacific	1.2 (0.9-1.6)	1.5 (1.1-2.1)	0.9 (0.6-1.3)	.02	0.9 (0.6-1.2)	1.2 (0.9-1.7)	.11

Figure 1. Prevalence of ever had sexual intercourse among young adolescents aged 12 years to 15 years based on the most recent Global School-based Student Health Survey from 69 countries, 2003-2017.

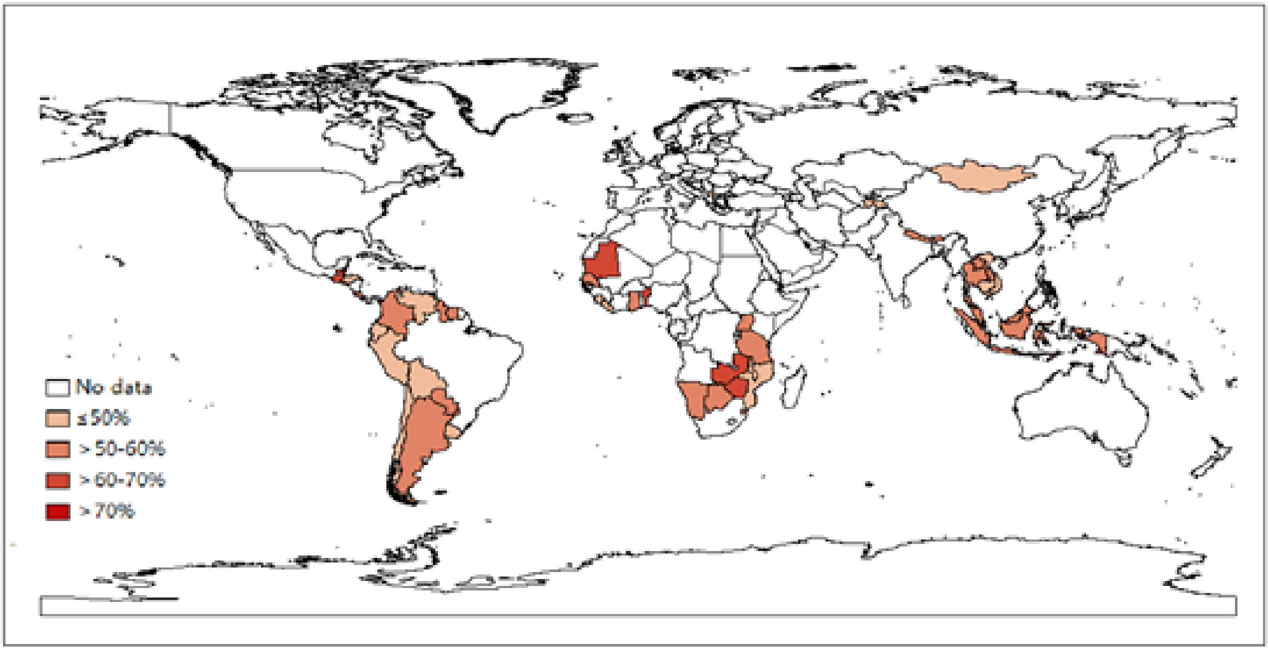


The prevalences of having multiple sexual partners and condom use among adolescents were based on those who ever had sexual intercourse. A total of 20,257 adolescents were included in the analysis of the prevalences of having multiple partners and condom use. The overall prevalence of having multiple sexual partners was 52.0% (95% CI 50.4%-53.6%) and was lowest in Western Pacific and highest in the Eastern Mediterranean (Table 3). The gender difference was statistically significant in all regions except for the Eastern Mediterranean, and the age difference was significant only in the Americas. In Table S2 in Multimedia Appendix 1, most (50 countries, 72%) of the 69 included countries had a prevalence of having multiple sexual partners >50% (Figure 2). The prevalence ranged from 2 times to 9 times higher among boys than among girls across countries, and the prevalence in adolescents aged 14 years to 15 years was higher than in those aged 12 years to 13 years in 47 of the 61 countries (data were unavailable for 8 countries in the Western Pacific).

Table 3. The prevalence of having multiple sexual partners among young adolescents aged 12 years to 15 years by World Health Organization region, sex, and age in 2003-2017.

Region	Total sample, % (95% CI)	Boys, % (95% CI)	Girls, % (95% CI)	*P* value	12-13 years old, % (95% CI)	14-15 years old, % (95% CI)	*P* value
Total	52.0 (50.4-53.6)	58.0 (56.1-59.9)	41.4 (38.9-43.9)	<.001	49.7 (45.9-53.5)	53.5 (51.6-55.4)	.12
Africa	52.3 (49.3-55.3)	54.7 (51.2-58.2)	45.2 (39.5-51.0)	.007	54.4 (46.7-61.9)	51.7 (47.9-55.5)	.59
Americas	52.3 (50.6-54.0)	59.9 (57.9-61.9)	38.7 (35.9-41.6)	<.001	46.9 (42.8-51.0)	53.5 (51.5-55.5)	.008
Eastern Mediterranean	71.2 (61.1-79.5)	72.3 (62.2-80.5)	63.9 (37.4-83.9)	.48	63.2 (39.1-82.2)	72.2 (62.1-80.4)	.40
Europe	40.8 (34.7-47.3)	44.8 (37.6-52.1)	16.4 (7.9-30.9)	.001	41.1 (28.1-55.4)	40.8 (34.0-47.9)	.97
Southeast Asia	56.0 (49.1-62.7)	60.4 (51.8-68.5)	47.5 (37.4-57.7)	.04	47.3 (34.0-61.0)	59.4 (51.6-66.6)	.13
Western Pacific	32.3 (22.1-44.4)	57.4 (51.9-62.9)	40.5 (34.0-47.4)	.001	53.0 (42.8-63.0)	51.7 (47.1-56.2)	.82

Figure 2. Prevalence of having multiple sexual partners among young adolescents aged 12 years to 15 years who ever had sexual intercourse based on the most recent Global School-based Student Health Survey from 69 countries, 2003-2017.

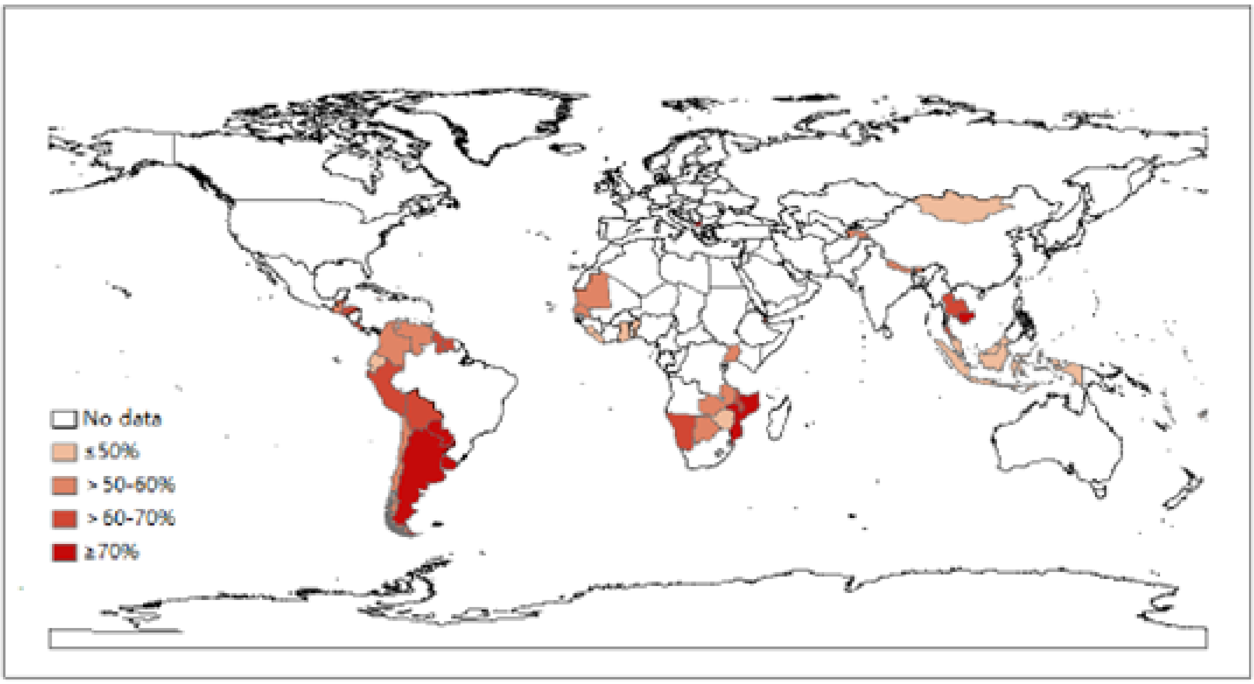


The overall prevalence of condom use at last sex was 58.1% (95% CI 56.2%-59.9%) and was lowest in the Western Pacific and highest in Europe (Table 4). The overall prevalence of condom use at last sex exceeded 50% in all regions, regardless of gender and age, except for among boys in Africa and the Western Pacific. In Table S3 in Multimedia Appendix 1, more than three-quarters (52 countries, 75%) of the included 69 countries had a prevalence >50% (Figure 3). In 38 of the 62 countries (data were unavailable for 7 countries in the Western Pacific), the prevalence in boys was higher than in girls, particularly in all countries in the Western Pacific. The prevalence in adolescents aged 14 years to 15 years was higher than in those aged 12 years to 13 years in 40 of the 62 countries, while all countries in Europe showed the opposite result.

Table 4. The prevalence of condom use at last sex among young adolescents aged 12 years to 15 years by World Health Organization region, sex, and age in 2003-2017.

Region	Total sample, % (95% CI)	Boys, % (95% CI)	Girls, % (95% CI)	*P* value	12-13 years old, % (95% CI)	14-15 years old, % (95% CI)	*P* value
Total	58.1 (56.2-59.9)	57.7 (55.7-59.7)	59.2 (56.4-61.9)	.35	51.6 (47.5-55.7)	59.9 (58.0-61.8)	<.001
Africa	47.8 (43.4-52.2)	45.9 (41.5-50.4)	53.4 (46.3-60.3)	.03	37.9 (30.7-45.7)	50.9 (46.4-55.3)	.002
Americas	65.2 (63.2-67.2)	66.1 (63.7-68.3)	63.8 (60.6-66.8)	.19	60.8 (56.7-64.7)	66.2 (63.9-68.5)	.02
Eastern Mediterranean	64.1 (51.6-75.0)	64.9 (51.4-76.3)	59.4 (31.5-82.3)	.72	64.6 (40.9-82.7)	64.1 (51.0-75.4)	.97
Europe	71.7 (65.2-77.4)	71.3 (64.5-77.3)	73.9 (62.2-83.0)	.62	72.9 (52.7-86.7)	71.4 (64.1-77.7)	.87
Southeast Asia	54.8 (48.6-60.9)	56.7 (48.5-64.5)	51.3 (41.5-61.0)	.36	55.7 (43.2-67.5)	54.5 (46.8-62.0)	.86
Western Pacific	47.6 (33.2-62.3)	49.0 (44.3-53.8)	31.0 (22.6-40.8)	<.001	44.4 (34.2-55.2)	42.8 (37.5-48.3)	.77

Figure 3. Prevalence of condom use at last sex among young adolescents aged 12 years to 15 years who ever had sexual intercourse based on the most recent Global School-based Student Health Survey from 69 countries, 2003-2017.



Trends in the Prevalences of Sexual Behaviors in Young Adolescents

As shown in Table 5, the overall prevalence of ever had sexual intercourse in young adolescents showed a downward trend over time between the earliest and latest surveys (3.1% decrease); a similar trend was observed by sex (boys, 3.9% decrease; girls, 2.5% decrease) and age group (12-13 years, 1.1% decrease; 14-15 years, 3.7% decrease). Specifically, in Table S4 in Multimedia Appendix 1, the prevalence of ever had sexual intercourse decreased in 9 countries (with the largest decrease of 24.6% in Samoa) and increased in 8 countries (with the largest increase of 7.9% in Seychelles).

Table 5. The trends in the prevalence of ever had sexual intercourse among young adolescents aged 12 years and 15 years between 2003 and 2017 or the earliest and latest surveys by World Health Organization region, sex, and age.

Region	Total sample, % (95% CI)	*P* value	Boys, % (95% CI)	*P* value	Girls, % (95% CI)	*P* value	12-13 years old, % (95% CI)	*P* value	14-15 years old, % (95% CI)	*P* value
Total	–3.1 (–3.5 to –2.7)	<.001	–3.9 (–4.5 to –3.3)	<.001	–2.5 (–2.9 to –2.1)	<.001	–1.1 (–2.0 to –0.2)	.02	–3.7 (–4.2 to –3.2)	<.001
Africa	–3.4 (–4.5 to –2.3)	.008	–2.8 (–4.7 to –0.9)	<.001	–3.7 (–4.9 to –2.5)	<.001	–3.9 (–7.8 to 0.1)	.31	–3.1 (–4.3 to –1.9)	.001
Americas	–2.6 (–3.5 to –1.7)	<.001	–6.6 (–8.1 to –5.1)	<.001	0 (–1.1 to 1.1)	.99	1.3 (–0.8 to 3.4)	.43	–2.6 (–3.8 to –1.4)	<.001
Southeast Asia	–1.0 (–1.5 to –0.5)	.008	–0.8 (–1.6 to 0.1)	.99	–1.2 (–1.8 to –0.6)	.002	–0.5 (–3.5 to 2.5)	.92	–1.2 (–1.9 to –0.5)	.02
Western Pacific	–3.2 (–4.7 to –1.7)	<.001	–4.6 (–7.7 to –1.5)	<.001	–1.9 (–3.2 to –0.6)	<.001	–3.1 (–10.6 to 4.4)	.32	–5.3 (–7.2 to –3.4)	<.001

A total of 13,051 adolescents were included in the trend analyses of the prevalences of having multiple partners and condom use. Table 6 shows that the overall prevalence of having multiple sexual partners increased by 2.6% over time; the upward trend was similar among girls and adolescents aged 14 years to 15 years (girls, 5.8% increase; age 14-15 years, 2.7% increase). All regions experienced an increase in the prevalence of having multiple sexual partners except for Africa. Specifically, in Table S5 in Multimedia Appendix 1, the prevalence of having multiple sexual partners increased in 9 countries (with the largest increase of 14.2% in Namibia), was unchanged in 2 countries, and decreased in 5 countries (with the largest decrease of 3.9% in Guyana).

Table 6. The trends in the prevalence of having multiple sexual partners among young adolescents aged 12 years to 15 years between 2003 and 2017 or the earliest and latest surveys by World Health Organization region, sex, and age.

Region	Total sample, % (95% CI)	*P* value	Boys, % (95% CI)	*P* value	Girls, % (95% CI)	*P* value	12-13 years old, % (95% CI)	*P* value	14-15 years old, % (95% CI)	*P* value
Total	2.6 (0.8 to 4.4)	<.001	0.4 (–1.8 to 2.6)	.56	5.8 (2.9 to 8.7)	.047	2.5 (–1.9 to 6.9)	.81	2.7 (0.7 to 4.7)	<.001
Africa	–0.2 (–3.9 to 3.5)	.95	–0.6 (–5.3 to 4.1)	.89	–0.5 (–6.2 to 5.2)	.93	27.5 (20.7 to 34.3)	<.001	–9.6 (–13.8 to –5.4)	.002
Americas	1.6 (–0.9 to 4.1)	.59	–0.5 (–3.5 to 2.5)	.06	5.9 (1.8 to 10.0)	<.001	1.0 (–5.1 to 7.1)	.26	2.0 (–0.7 to 4.7)	.07
Southeast Asia	7.4 (–3.9 to 18.7)	.25	4.4 (–10.1 to 18.9)	.42	8.2 (–8.1 to 24.5)	.38	–22.9 (–45.8 to 0.1)	.63	14.9 (2.2 to 27.6)	.046
Western Pacific	4.7 (–3.8 to 13.2)	.58	3.7 (–6.6 to 14.0)	.86	6.3 (–8.5 to 21.1)	.09	–8.9 (–37.0 to 19.2)	.15	4.9 (–4.1 to 13.9)	.80

In Table 7, the overall prevalence of condom use at last sex decreased by 2.0% over time; the downward trend was similar in boys (4.5% decrease) but increased for girls (1.8% increase). The decrease was not significant in adolescents aged 12 years to 13 years or 14 years to 15 years. Specifically, in Table S6 in Multimedia Appendix 1, the prevalence of condom use at last sex decreased in 10 countries (with the largest decrease of 9.7% in Ghana) and increased in 6 countries (with the largest increase of 14.6% in Namibia).

Table 7. The trends in the prevalence of condom use among young adolescents aged 12 years to 15 years between 2003 and 2017 or the earliest and latest surveys by World Health Organization region, sex, and age.

Region	Total sample, % (95% CI)	*P* value	Boys, % (95% CI)	*P* value	Girls, % (95% CI)	*P* value	12-13 years old, % (95% CI)	*P* value	14-15 years old, % (95% CI)	*P* value
Total	–2.0 (–3.7 to –0.3)	<.001	–4.5 (–6.6 to –2.4)	<.001	1.8 (–1.0 to 4.6)	.67	–2.4 (–6.5 to 1.7)	.64	–1.8 (–3.6 to 0.1)	.55
Africa	–3.6 (–7.2 to 0.1)	.57	–7.3 (–12.1 to –2.5)	.004	3.5 (–2.2 to 9.2)	.74	–12.6 (–20.4 to –4.8)	<.001	–0.3 (–4.5 to 3.9)	.97
Americas	–4.2 (–6.3 to –2.1)	<.001	–5.6 (–8.1 to –3.1)	.001	–1.8 (–5.4 to 1.8)	.75	–11.9 (–17.0 to –6.8)	<.001	–2.9 (–5.2 to –0.6)	.43
Southeast Asia	6.8 (–4.1 to 17.7)	.57	5.7 (–8.2 to 19.6)	.88	6.6 (–10.3 to 23.5)	.50	18.8 (–3.1 to 40.7)	.38	4.0 (–8.5 to 16.5)	.85
Western Pacific	2.2 (–6.4 to 10.8)	.72	4.0 (–6.5 to 14.5)	.91	–4.0 (–18.9 to 10.9)	.44	22.3 (–4.5 to 49.1)	.21	1.1 (–8.0 to 10.2)	.37

Discussion

Principal Findings

This study used the most recent GSHS data from 69 countries that had conducted at least one survey between 2003 and 2017 to assess the prevalences of sexual behaviors among adolescents aged 12 years to 15 years. Specifically, the overall prevalence of ever had sexual intercourse was 6.9%, and among adolescents who had ever had sex, the overall prevalences of having multiple sexual partners and using condom at last sex were 52.0% and 58.1%, respectively.

The prevalences of sexual behaviors among young adolescents varied widely across countries, and these differences may be due to several factors, including the social cultural environment, economic status, policy, race, and attitudes toward sex. For example, we found that the lowest prevalences of ever had sexual intercourse were in Indonesia, Malaysia, and Tajikistan, which can be explained by the religious culture embedded in these countries. Most people in these countries embrace Islam, which prohibits premarital sex and regards sexuality as a taboo and sensitive subject [20]. Especially in Indonesia, as the largest Islamic country with legal restrictions on pornography in the world [21], adolescents were generally considered to be more sexually conservative and bound due to the strict legal regulations and Islamic values.

The prevalences of sexual behaviors in boys were higher than in girls in most LMICs, which was consistent with previous studies in various countries, including HICs [22]. This finding may be due to the difference in sexual development between boys and girls. Previous studies have suggested that boys are more likely to report risky sexual behaviors and to initiate sexual intercourse earlier than girls [23,24]. Meanwhile, boys have more permissive attitudes about premarital sexual activity, while sexual intercourse in girls is often accompanied by trust, love, and a romantic relationship [25]. On the other hand, there are double standards in society for early sexual intercourse behaviors between boys and girls. Boys experience more sexual freedom, and sexual initiation in boys is regarded as a symbol of masculinity or rite of passage in some countries [26]. However, premarital sexual intercourse is stigmatized or labeled as indiscreet in girls, especially in Asian countries [23], and girls not only are bearing multiple social pressures but also have a higher risk of reporting unintended pregnancy and poor reproductive health. It is worth noting that the prevalence of condom use at last sex among boys in LMICs (57.7%) is lower than that from data obtained from HICs such as the United States (61% in those aged 13-19 years) and Australia (65.1% in those aged 10-12 years) and that future measures are needed to increase the use of condoms among not only girls but also boys in LMICs [27,28].

This study also provided evidence that the prevalences of sexual behaviors among adolescents aged 14 years to 15 years were higher than among those aged 12 years to 13 years in most countries, which has been shown in many previous studies [29]. This can be interpreted using the normal sexual development process. The expression of secondary sex characteristics is influenced by sexual-related hormones in puberty. A study showed that, during the ages of 12 years to 15 years, changes in hormones mature the reproductive system with age [30]. We speculated that sexual development may arouse sexual desire; thus, older adolescents are more likely to engage in sexual intercourse than younger adolescents.

Based on the GSHS data from 17 countries that had conducted more than one round of survey between the earliest and latest surveys, the overall prevalences of ever had sexual intercourse and condom use at last sex decreased over time, but the prevalence of having multiple sexual partners increased over time. To our knowledge, few studies have investigated the trends in sexual behaviors among young adolescents in LMICs, and this is the first study to explore the trends of sexual behaviors in young adolescents aged 12 years to 15 years in LMICs. An earlier study based on both the Youth Risk Behavior Survey (YRBS) and National Survey of Adolescent Males data found that the proportion of male adolescents aged 15 years to 17 years reporting sexual intercourse decreased by 9% from 1991 to 1997 and by 8% from 1988 to 1995 [31]. Another YRBS report based on recent surveillance data from the United States found that the percentages of high school students who ever had sex, had multiple sexual partners, or used a condom at last sex declined from 2009 to 2019 [32], with the downward trend in sexual intercourse and condom use at last sex consistent with our findings.

From a public health perspective, the decline in the prevalence of sexual intercourse among young adolescents aged 12 years to 15 years was a salutary trend, as sexual intercourse at an early age may put young adolescents at higher risk for poorer sexual and reproductive health outcomes. Most previous studies have identified factors influencing early sexual initiation from a static perspective, and only a few studies have explained the possible reasons for the decline in the prevalence of sexual intercourse. One study indicated that access to formal sexual education may reduce the number of adolescents younger than 15 years who have sexual intercourse [33]. Another study conducted in British Columbia suggested that the presence of protective factors (such as supportive schools and families, opportunities for community and school involvement) in adolescents' lives may contribute to declines in the proportion of adolescents in grades 7 to 12 who reported ever having sexual intercourse from 1992 to 2003 [34]. There is currently a lack of evidence explaining the change in adolescent sexual intercourse; future research should consider investigating factors that may affect changes in the prevalences of sexual behaviors among young adolescents.

The decreasing prevalence of condom use at last sex over time among young adolescents aged 12 years to 15 years is worrying. Previous studies have suggested that limited access to condoms (eg, restrictive laws or policies that provide contraceptives based on age or marital status) and the sociocultural environment (eg, the stigmatization of condoms) may be the main barriers to condom use among young people [35]. In addition, the decline in condom use at last sex among young adolescents might be explained by the lack of contraceptive knowledge, financial constraints, and the pursuit of sexual pleasure [36]. We found that the prevalence of having multiple sexual partners among adolescents who had ever had sexual intercourse increased over time, which was inconsistent with findings in HICs. The United States (2009-2019) and Canada (1992-2003) reported the prevalence of having multiple sexual partners decreased over time, by 5.2%, among high school students and, by 3.6%, among adolescents in grades 7 to 12 [32,34]. These previous studies conducted in Canada and the United States defined multiple sexual partners as having 3, 4, or more sexual partners during their lifetime, whereas the GSHS defined multiple sexual partners as having 2 or more partners in their lifetime, which may be one reason why the results of this study are inconsistent with previous results.

Implications and Contribution

Healthy sexual behaviors are an important component of adolescents' physical and psychological development. This study reported the prevalence and trends of sexual behaviors of young adolescents aged 12 years to 15 years in LMICs, which can provide evidence and important implications for policymakers to implement targeted programs that promote and improve the sexual and reproductive health of adolescents. The prevalences of adolescent sexual behaviors vary widely across countries, and there is an urgent need for the public health sector to develop targeted policy support systems to prevent and reduce risky sexual behaviors among young adolescents in LMICs with high prevalences of risky sexual behaviors. For example, the comprehensive sexuality education (CSE) intervention should be implemented, and its integration into the education system should be encouraged in countries where the prevalence of early sexual intercourse initiation is high or rising. Previous studies have indicated that CSE can increase sexual health knowledge and reduce risky sexual behaviors among young adolescents such as delaying the age of first sexual intercourse or ensuring that the age of sexual initiation is developmentally appropriate for the individual, increasing contraception use, and reducing the number of sexual partners [37,38]. The decline in condom use at last sex reported by some countries was a reminder that policymakers should continue to prioritize potential barriers to condom use among adolescents and explore solutions to these barriers. It is also important to consider gender and age differences among young adolescents when developing strategies to promote adolescent sexual health.

Limitations

This study also has several limitations. First, the prevalence of "having multiple sexual partners" or "condom use at last sex" among adolescents was based on those who have had sexual intercourse; the sample size of young adolescents who ever had sexual intercourse in some countries was extremely low, which may have caused selection bias and resulted in the lowest or highest prevalences of having multiple sexual partners and condom use in these countries. Therefore, the interpretation of the prevalences of having multiple sexual partners and condom use should be considered within the specific circumstances of the sample area. Second, the sexual behavior information was self-reported, which may have resulted in recall bias or intentional omission of key information. Although the interview questionnaire was anonymous and every effort was made to create a secure environment and maintain confidentiality of the respondents' information, the prevalences of sexual behaviors may still be underestimated or overestimated. Third, the GSHS was based on surveys of teenagers in school, which did not include those who did not attend school. Previous studies have shown that out-of-school students are more likely to initiate sexual behaviors than in-school students; thus, there may exist differences in sexual-related behaviors between these 2 groups, and the results of this study may not be suitable for generalization to all adolescents. Fourth, due to the data availability for each country, only 17 countries have conducted 2 rounds of surveys on sexual behaviors, and there are differences in the timing and length of surveys between countries. Although this can provide evidence for policy development in each country, our estimates for regional or global trends therefore need to be interpreted with caution. Finally, for some countries where sexual behavior surveys were conducted a long time ago (in the past decade or so), it is necessary to obtain data on recent adolescent sexual behaviors to identify recent trends.

Conclusions

In conclusion, this study found that the prevalences of sexual behaviors in young adolescents varied by country, region, sex, and age group. Over time, the overall prevalences of ever had sexual intercourse and condom use at last sex decreased, but the prevalence of having multiple sexual partners increased. This study can provide evidence and important implications for policymakers to develop a targeted policy support system to prevent and reduce risky sexual behaviors among young adolescents in LMICs with high prevalences of risky sexual behaviors.

Acknowledgments

We thank the World Health Organization and the US Centers for Disease Control for making the Global School-based Student Health Surveys (GSHS) data accessible for analysis and the country survey coordinators and participants involved in conducting the GSHS.

Data Availability

Data are available in a public, open access repository. It can be freely accessed via the World Health Organization Noncommunicable Diseases Microdata Repository [19].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary file.
[DOCX File , 102 KB - publichealth_v9i1e45236_app1.docx]

References

1. Pinyopornpanish K, Thanamee S, Jiraporncharoen W, Thaikla K, McDonald J, Aramrattana A, et al. Sexual health, risky sexual behavior and condom use among adolescents young adults and older adults in Chiang Mai, Thailand: findings from a population based survey. BMC Res Notes 2017 Dec 04;10(1):682 [FREE Full text] [doi: 10.1186/s13104-017-3055-1] [Medline: 29202883]
2. Layland EK, Ram N, Caldwell LL, Smith EA, Wegner L. Leisure boredom, timing of sexual debut, and co-occurring behaviors among South African adolescents. Arch Sex Behav 2021 Aug;50(6):2383-2394 [FREE Full text] [doi: 10.1007/s10508-021-02014-8] [Medline: 34401994]
3. Ahanhanzo YG, Sossa-Jérôme C, Sopoh G, Tchandana M, Azandjèmè C, Tchamdja T. Factors associated with early sexual intercourse among teenagers and young adults in rural south of Benin. J Public Health Afr 2018 Oct 01;9(2):681 [FREE Full text] [doi: 10.4081/jphia.2018.681] [Medline: 30687472]
4. Ghebremichael M, Larsen U, Paintsil E. Association of age at first sex with HIV-1, HSV-2, and other sexual transmitted infections among women in northern Tanzania. Sex Transm Dis 2009 Sep;36(9):570-576 [FREE Full text] [doi: 10.1097/OLQ.0b013e3181a866b8] [Medline: 19707109]
5. Schofield HT, Bierman KL, Heinrichs B, Nix RL, Conduct Problems Prevention Research Group. Predicting early sexual activity with behavior problems exhibited at school entry and in early adolescence. J Abnorm Child Psychol 2008 Nov 8;36(8):1175-1188 [FREE Full text] [doi: 10.1007/s10802-008-9252-6] [Medline: 18607716]
6. Salam RA, Faqqah A, Sajjad N, Lassi ZS, Das JK, Kaufman M, et al. Improving adolescent sexual and reproductive health: a systematic review of potential interventions. J Adolesc Health 2016 Oct;59(4S):S11-S28 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.05.022] [Medline: 27664592]
7. Kalamar AM, Bayer AM, Hindin MJ. Interventions to prevent sexually transmitted infections, including HIV, among young people in low- and middle-income countries: a systematic review of the published and gray literature. J Adolesc Health 2016 Sep;59(3 Suppl):S22-S31. [doi: 10.1016/j.jadohealth.2016.05.020] [Medline: 27562450]
8. Kim D, Kim H. Early initiation of alcohol drinking, cigarette smoking, and sexual intercourse linked to suicidal ideation and attempts: findings from the 2006 Korean Youth Risk Behavior Survey. Yonsei Med J 2010 Jan;51(1):18-26 [FREE Full text] [doi: 10.3349/ymj.2010.51.1.18] [Medline: 20046509]
9. Pettifor A, van der Straten A, Dunbar M, Shiboski S, Padian N. Early age of first sex: a risk factor for HIV infection among women in Zimbabwe. AIDS 2004 Jul 02;18(10):1435-1442. [doi: 10.1097/01.aids.0000131338.61042.b8] [Medline: 15199320]
10. O'Donnell L, O'Donnell CR, Stueve A. Early sexual initiation and subsequent sex-related risks among urban minority youth: the reach for health study. Fam Plann Perspect 2001;33(6):268-275 [FREE Full text] [Medline: 11804436]

11. Smith L, Jackson SE, Jacob L, Grabovac I, Nisbet LA, López-Sánchez GF, et al. Association between cannabis use and sexual behavior among adolescents aged 12-15 years in 21 low- and middle-income countries. Addict Behav 2019 Dec;99:106073. [doi: 10.1016/j.addbeh.2019.106073] [Medline: 31421584]
12. Alawode OA, Ogunwemimo H, Bolorunduro M, Awoleye AF. Age at sexual debut and multiple sexual partnerships among adolescents in Nigeria: an assessment of the mediating role of the knowledge of sexually transmitted infections. Adolescents 2021 Oct 12;1(4):421-432. [doi: 10.3390/adolescents1040032]
13. Wilson CN, Sathiyasusuman A. Associated risk factors of STIs and multiple sexual relationships among youths in Malawi. PLoS ONE 2015 Aug 6;10(8):e0134286. [doi: 10.1371/journal.pone.0134286]
14. Holmes K, Levine R, Weaver M. Effectiveness of condoms in preventing sexually transmitted infections. Bull World Health Organ 2004 Jun;82(6):454-461 [FREE Full text] [Medline: 15356939]
15. Weller S, Davis BK. Condom effectiveness in reducing heterosexual HIV transmission. Cochrane Database Syst Rev 2002;3(1):CD003255. [doi: 10.1002/14651858.CD003255] [Medline: 11869658]
16. Kaplan D, Feinstein R, Fisher M, Klein J, Olmedo L, Rome E, Committee on Adolescence. Condom use by adolescents. Pediatrics 2001 Jun;107(6):1463-1469. [doi: 10.1542/peds.107.6.1463] [Medline: 11389278]
17. Brown LK, DiClemente R, Crosby R, Fernandez MI, Pugatch D, Cohn S, Project Shield Study Group. Condom use among high-risk adolescents: anticipation of partner disapproval and less pleasure associated with not using condoms. Public Health Rep 2008 Sep 01;123(5):601-607 [FREE Full text] [doi: 10.1177/003335490812300510] [Medline: 18828415]
18. Woog V, Kågesten A. The sexual and reproductive health needs of very young adolescents aged 10–14 in developing countries: what does the evidence show? Guttmacher Institute. 2017 May. URL: https://www.guttmacher.org/report/srh-needs-very-young-adolescents-in-developing-countries [accessed 2023-05-15]
19. Global School-based Student Health Survey. World Health Organization NCD Microdata Repository. URL: https://extranet.who.int/ncdsmicrodata/index.php/catalog/GSHS [accessed 2023-05-15]
20. Lee LK, Chen PCY, Lee KK, Kaur J. Premarital sexual intercourse among adolescents in Malaysia: a cross-sectional Malaysian school survey. Singapore Medical Journal 2006;47(6):476-481 [FREE Full text]
21. Hald GM, Mulya TW. Pornography consumption and non-marital sexual behaviour in a sample of young Indonesian university students. Cult Health Sex 2013 Jun 20;15(8):981-996. [doi: 10.1080/13691058.2013.802013] [Medline: 23782270]
22. Nogueira Avelar E Silva R, Wijtzes A, van de Bongardt D, van de Looij-Jansen P, Bannink R, Raat H. Early sexual intercourse: prospective associations with adolescents physical activity and screen time. PLoS One 2016 Aug 11;11(8):e0158648 [FREE Full text] [doi: 10.1371/journal.pone.0158648] [Medline: 27513323]
23. Zuo X, Lou C, Gao E, Cheng Y, Niu H, Zabin LS. Gender differences in adolescent premarital sexual permissiveness in three Asian cities: effects of gender-role attitudes. J Adolesc Health 2012 Mar;50(3 Suppl):S18-S25 [FREE Full text] [doi: 10.1016/j.jadohealth.2011.12.001] [Medline: 22340852]
24. Mladenovic B, Donev D, Spasovski M. Gender analysis of sexual behaviour of senior high-school students in Skopje, R. Macedonia - cross-sectional study. Prilozi 2009 Jul;30(1):179-190. [Medline: 19736540]
25. Petersen JL, Hyde JS. Gender differences in sexual attitudes and behaviors: a review of meta-analytic results and large datasets. J Sex Res 2011 Mar 15;48(2-3):149-165. [doi: 10.1080/00224499.2011.551851] [Medline: 21409712]
26. Barker G, Ricardo C. Young men and the construction of masculinity in Sub-Saharan Africa : implications for HIV/AIDS, conflict, and violence. The World Bank. 2005. URL: https://documents1.worldbank.org/curated/en/481401468101357773/pdf/327120rev0PAPER0AFR0young0men0WP26.pdf [accessed 2023-05-15]
27. Koumans EH, Welch R, Warner DL. 70. Differences in adolescent condom use trends by global region. Journal of Adolescent Health 2020 Feb;66(2):S36-S37. [doi: 10.1016/j.jadohealth.2019.11.073]
28. Fisher C, Kauer S, Mikolajczak G, Ezer P, Kerr L, Bellamy R, et al. Prevalence rates of sexual behaviors, condom use, and contraception among Australian heterosexual adolescents. J Sex Med 2020 Dec;17(12):2313-2321. [doi: 10.1016/j.jsxm.2020.08.009] [Medline: 32948487]
29. Jee Y, Lee G. Prevalence of sexual experience among Korean adolescent: age-period-cohort analysis. Epidemiol Health 2020 Mar 03;42:e2020008 [FREE Full text] [doi: 10.4178/epih.e2020008] [Medline: 32124584]
30. Fortenberry JD. Puberty and adolescent sexuality. Horm Behav 2013 Jul;64(2):280-287 [FREE Full text] [doi: 10.1016/j.yhbeh.2013.03.007] [Medline: 23998672]
31. Santelli JS, Lindberg LD, Abma J, McNeely CS, Resnick M. Adolescent sexual behavior: estimates and trends from four nationally representative surveys. Family Planning Perspectives 2000 Jul;32(4):156-165 [FREE Full text] [doi: 10.2307/2648232]
32. Youth Risk Behavior Surveillance Data Summary & Trends Report. Centers for Disease Control and Prevention. 2009. URL: https://www.cdc.gov/nchhstp/dear_colleague/2020/dcl-102320-YRBS-2009-2019-report.html
33. Mueller TE, Gavin LE, Kulkarni A. The association between sex education and youth's engagement in sexual intercourse, age at first intercourse, and birth control use at first sex. J Adolesc Health 2008 Jan;42(1):89-96. [doi: 10.1016/j.jadohealth.2007.08.002] [Medline: 18155035]
34. Saewyc E, Taylor D, Homma Y, Ogilvie G. Trends in sexual health and risk behaviours among adolescent students in British Columbia. Can J Hum Sex 2008 Jan 01;17(1-2):1-13 [FREE Full text] [Medline: 23115485]

35. Harper CR, Steiner RJ, Lowry R, Hufstetler S, Dittus PJ. Variability in condom use trends by sexual risk behaviors: findings from the 2003–2015 National Youth Risk Behavior Surveys. Sexual Trans Dis 2018;45(6):400-405. [doi: 10.1097/olq.0000000000000763]
36. Society for Adolescent Health and Medicine. Condom availability in schools: a practical approach to the prevention of sexually transmitted infection/HIV and unintended pregnancy. J Adolesc Health 2017 Jun;60(6):754-757. [doi: 10.1016/j.jadohealth.2017.03.019] [Medline: 28532649]
37. Kemigisha E, Bruce K, Ivanova O, Leye E, Coene G, Ruzaaza GN, et al. Evaluation of a school based comprehensive sexuality education program among very young adolescents in rural Uganda. BMC Public Health 2019 Oct 28;19(1):1393 [FREE Full text] [doi: 10.1186/s12889-019-7805-y] [Medline: 31660918]
38. Fonner VA, Armstrong KS, Kennedy CE, O'Reilly KR, Sweat MD. School based sex education and HIV prevention in low- and middle-income countries: a systematic review and meta-analysis. PLoS One 2014;9(3):e89692. [doi: 10.1371/journal.pone.0089692] [Medline: 24594648]

Abbreviations

CDC: Center for Disease Control and Prevention
CSE: comprehensive sexuality education
GSHS: Global School–based Student Health Survey
HIC: high-income country
LMICs: low and middle-income countries
STI: sexually transmitted infection
WHO: World Health Organization
YRBS: Youth Risk Behavior Survey

Edited by A Mavragani; submitted 21.12.22; peer-reviewed by Z Hou, X Qu, X Yin; comments to author 27.01.23; revised version received 06.03.23; accepted 26.04.23; published 07.06.23.

Please cite as:
Jing Z, Li J, Wang Y, Zhou C
Prevalence and Trends of Sexual Behaviors Among Young Adolescents Aged 12 Years to 15 Years in Low and Middle-Income Countries: Population-Based Study
JMIR Public Health Surveill 2023;9:e45236
URL: https://publichealth.jmir.org/2023/1/e45236
doi:10.2196/45236
PMID:37285196

Original Paper

Health Information on Pre-Exposure Prophylaxis From Search Engines and Twitter: Readability Analysis

Albert Park[1], PhD; Fatima Sayed[1], MS; Patrick Robinson[2], MD, MPH; Latesha Elopre[3], MSPH, MD; Yaorong Ge[1], PhD; Shaoyu Li[4], PhD; Christian Grov[5], MPH, PhD; Patrick Sean Sullivan[6], DVM, PhD

[1]Department of Software and Information Systems, University of North Carolina Charlotte, Charlotte, NC, United States
[2]Department of Public Health Sciences, University of North Carolina Charlotte, Charlotte, NC, United States
[3]Division of Infectious Disease, University of Alabama at Birmingham, Birmingham, AL, United States
[4]Department of Mathematics and Statistics, University of North Carolina Charlotte, Charlotte, NC, United States
[5]Department of Community Health and Social Sciences, City University of New York, New York City, NY, United States
[6]Department of Epidemiology, Emory University, Atlanta, GA, United States

Corresponding Author:
Albert Park, PhD
Department of Software and Information Systems
University of North Carolina Charlotte
9201 University City Blvd
Woodward 310H
Charlotte, NC, 28223-0001
United States
Phone: 1 704 687 8668
Fax: 1 704 687 1652
Email: al.park@charlotte.edu

Abstract

Background: Pre-exposure prophylaxis (PrEP) is proven to prevent HIV infection. However, PrEP uptake to date has been limited and inequitable. Analyzing the readability of existing PrEP-related information is important to understand the potential impact of available PrEP information on PrEP uptake and identify opportunities to improve PrEP-related education and communication.

Objective: We examined the readability of web-based PrEP information identified using search engines and on Twitter. We investigated the readability of web-based PrEP documents, stratified by how the PrEP document was obtained on the web, information source, document format and communication method, PrEP modality, and intended audience.

Methods: Web-based PrEP information in English was systematically identified using search engines and the Twitter API. We manually verified and categorized results and described the method used to obtain information, information source, document format and communication method, PrEP modality, and intended audience. Documents were converted to plain text for the analysis and readability of the collected documents was assessed using 4 readability indices. We conducted pairwise comparisons of readability based on how the PrEP document was obtained on the web, information source, document format, communication method, PrEP modality, and intended audience, then adjusted for multiple comparisons.

Results: A total of 463 documents were identified. Overall, the readability of web-based PrEP information was at a higher level (10.2-grade reading level) than what is recommended for health information provided to the general public (ninth-grade reading level, as suggested by the Department of Health and Human Services). Brochures (n=33, 7% of all identified resources) were the only type of PrEP materials that achieved the target of ninth-grade reading level.

Conclusions: Web-based PrEP information is often written at a complex level for potential and current PrEP users to understand. This may hinder PrEP uptake for some people who would benefit from it. The readability of PrEP-related information found on the web should be improved to align more closely with health communication guidelines for reading level to improve access to this important health information, facilitate informed decisions by those with a need for PrEP, and realize national prevention goals for PrEP uptake and reducing new HIV infections in the United States.

(JMIR Public Health Surveill 2023;9:e48630) doi:10.2196/48630

KEYWORDS

pre-exposure prophylaxis; PrEP; health literacy; health education materials; readability; prophylaxis; health information; electronic health education; HIV; infection; Twitter

Introduction

Pre-exposure prophylaxis (PrEP) is proven to prevent HIV infections [1-3], and using PrEP to prevent new HIV infections among vulnerable populations has been identified as a critical strategy in the Ending the HIV Epidemic initiative [4]. PrEP use increased more than 10-fold in the United States in just 5 years after its approval in 2012 [5]. However, inequitable PrEP uptake persists in the United States [6-8], especially in the southern United States and among Black and Hispanic residents in the United States. When those who need PrEP the most do not receive it, we lose opportunities to maximize reductions in new HIV infections. One potential explanation for lower uptake among vulnerable populations for HIV infections is the lower level of health literacy in these groups (ie, people in the southern United States, people living in poverty, younger people, and people of Black race or Hispanic ethnicity) [9,10]. Thus, analyzing the readability (ie, how difficult a text is to understand) of web-based information about PrEP would help understand the effectiveness of available web-based PrEP information and identify opportunities to improve the understandability of information for people who might benefit from PrEP.

Health literacy is an individual’s ability to obtain and understand health information to make informed health decisions [11]. The US Department of Health and Human Services (HHS) [12,13], the National Institutes of Health (NIH) [13,14], the American Medical Association (AMA) [15], and the US Food and Drug Administration (FDA) [16] recommend health information for the public be written at the ninth-grade reading level or lower to ensure that most people in the US population (ie, the general public) can make informed health decisions. However, researchers have consistently found evidence that text-based health information resources, including web-based resources, are too complex for the general public [17-25]. For example, patient education materials for electronic cigarettes [22], mental health [17], pediatrics [23], medical specialties [24], diabetes mellitus [18], clinical orthopedics [19], human papillomavirus immunizations [20], and cardiovascular diseases [21] were found to require higher literacy levels than that recommended by the NIH, AMA, FDA, and HHS. Moreover, health information available from commercially funded sources was significantly more difficult to read than the information available from government-funded sources for some health conditions, such as diabetes, hypertension, depression, high cholesterol, arthritis, asthma, heartburn, obesity, influenza, and erectile dysfunction [26], but not for electronic cigarette information [22]. This complexity often led to comprehension errors [27,28].

The readability of web-based PrEP information has also been reported [25], but the previous assessment was limited to 100 unique websites and did not include information circulated on social media. Additionally, the previous assessment did not stratify the assessment, which limits our understanding of how to improve the readability of educational materials. For information related to PrEP use, the increasing complexity of educational materials has been associated with lower uptake of PrEP [29,30]. There are other factors that can impact the understandability of PrEP information and accessibility to certain populations. For example, the format of the information (eg, brochure and website), the information source (eg, US government entities and for-profit organizations), and the intended audience could impact the readability of and trust in materials. Thus, different forms and sources of information should be examined for readability; these data should be interpreted in light of intended audiences who might require different levels of complexity. Analyzing updated content is also important because new PrEP delivery methods continue to gain approval and become available to consumers. For example, with the recent approval of injectable PrEP [31], readability studies of PrEP information need to be updated to include this new delivery method.

Assessments of readability should also include information disseminated through social media [32,33], which is an important source of passive information for many young people. However, many recent studies on the readability of health-related content on the web [17-25] have either focused solely on information obtained actively through search engines [17,18,20-22,24,25] or within popular websites for consumers [19,23]. It is important to examine health information that is obtained both actively via search engines and passively via social media channels. The latter is critical because >80% of Americans use at least one social media platform [34] and major health agencies, such as World Health Organization (WHO) and Centers for Disease Control and Prevention (CDC), regularly communicate via Twitter feeds. Social media content can include news releases as well as discussions on various health topics (eg, sexually transmitted infections, emergency situations [35,36], and preventive measures [33]).

We aimed to improve our understanding of the readability of publicly available web-based and social media–based PrEP information by analyzing it based on its information source, document format, communication method, PrEP modality, and intended audience. Our findings can be used to develop communication strategies that are tailored to the needs of different objectives and improve health literacy and the understandability of PrEP information for the general public.

Methods

Data Collection

We collected examples of both passively (eg, social media posts) and actively (information sought through user-initiated web searches) obtained information.

Collecting Passively Obtained Information

For passively obtained information, we collected PrEP-related information exchanged through Twitter. Twitter was chosen because many health organizations currently use Twitter to communicate health messages, including information about PrEP and HIV prevention, to the public [37,38]. About 22% of US adults use Twitter (45% of 18- to 24-year-olds); many (42%) check the platform more than once a day [39]. Tweets are short, being limited to 280 characters, but often refer readers to longer text information through embedded URLs. Thus, we investigated patient-level information sources (ie, brochures and websites) that were disseminated by URLs in tweets. We assessed and compared the readability of PrEP information based on the method used to obtain information (ie, passively or actively), information source (ie, organization type), document format (eg, brochures and websites), PrEP modality (ie, oral and injectable), and the intended audience (ie, patient, provider, and general).

To assess passively obtained information on PrEP, we used PrEP-related tweets with URLs as a starting point. Using Twitter API [40], we first collected PrEP-related tweets from July 2012, when the FDA first approved the PrEP indication for a pharmaceutical medication [41], to January 31, 2022, based on PrEP-related keywords that were used in a previous study [42] and a new set of keywords related to injectable PrEP [31] (Textbox 1). To emulate how users might search for information about PrEP on the web, we started from the Twitter URLs and further assessed the contents of external URLs that were embedded in the original URL (ie, manual web crawling). For each URL visited, PrEP information was manually and systematically collected (eg, text about PrEP) from the URLs. We repeated this process until all the relevant information linked to the original URLs and their referred external URLs were collected.

Textbox 1. Pre-exposure prophylaxis (PrEP)–related keywords for tweet collection.

PrEP keywords: [(truvada) OR (hiv AND PrEP) OR (preexposure AND prophylaxis AND hiv) OR (pill AND prevent AND hiv) OR (pill AND protect AND hiv) OR (pill AND protect AND AIDS)] [42]

Injectable PrEP keywords: [(Apretude) OR (cabotegravir) OR (inject AND prevent AND hiv) OR (inject AND protect AND hiv)]

Collecting Actively Obtained Information

To assess actively obtained information, a systematic search of PrEP was conducted using 5 popular search engines (ie, Google, Yahoo, Bing, DuckDuckGo, and Ask) from April to June 2022. We simulated the behavior of general consumers using the following various search terms and phrases: "PrEP," "PrEP information," "What is PrEP," "How to use PrEP," "Pre-exposure Prophylaxis frequently asked questions," "how does PrEP work in HIV prevention," "PrEP brochure," "PrEP facts," "HIV and PrEP," "Truvada PrEP," "PrEP side effects," "how much does PrEP cost," or "how to take prep" until a search result saturation was reached (ie, new information was not obtained). Then for comparison purposes, we specifically searched for "medication guide," "patient information," "info for people," "information for the patient," "information for the user," and "patient version" of known PrEP medications, (eg, Truvada, Descovy, Apretude, Vocabria, and cabotegravir).

For each of the 5 search engines, search results were collected from the first page for each search engine because most users rarely read past the first page of search results [43] and because our focus was to analyze the most frequently accessed web-based PrEP information. We did not include sponsored listings, as we wanted to focus on the most relevant and unbiased information. We manually verified search results to ensure that we only included web pages with PrEP information. We then systematically and manually web crawled these pages, using links from the search results as a starting point to emulate user behavior. Several false-positive links were excluded from the search results that were related to different acronyms or abbreviations (eg, Professional Research Experience Program). We also excluded documents without any PrEP-related information (eg, content about drugs used only for HIV treatment). We also excluded papers published in peer-reviewed journals that are not intended for the general public and would skew the readability results. Several documents were part of a much bigger document (eg, product monograph and prescribing information). Despite their relevance, if the overall PrEP information document exceeded 32,767 characters, we only included sections on PrEP that were written for the prescription drug prescribers in our analyses. We felt that this would most closely emulate how lay people would read this text. Duplicate sources obtained from more than one search path, educational videos, and descriptive figures and images were excluded and removed, and the web pages were converted to plain text for analysis.

Ethical Considerations

We only analyzed publicly available documents in this study and did not analyze identifiable private information or involve any direct or indirect interactions with individuals. Per UNC Charlotte's policy (citation 45 CFR 46 Definitions), the study is exempt from institutional review board requirements because it does not meet the regulatory definitions of human subject research.

Data Categorization

Information Source

We categorized information sources according to the originator of the document. The originator type was first determined with manual Google searches about organizations. This information was supplemented with information from the "about us" section of websites. We also used information in the web-based sources about affiliations, funding sources, and the Crunchbase [44] to finalize their organization types. Several websites still had no explicit indication of their affiliations or funding sources; we have denoted them as "N/A" and categorized them separately in our comparison analyses.

The information sources for this study included national and state governmental agencies (ie, US-located government entities; eg, CDC and California Department of Public Health), non-US governmental agencies (eg, National Health Service of the

United Kingdom), other public health organizations (eg, WHO and Joint United Nations Programme on HIV/AIDS), nongovernment organizations (eg, Mayo Clinic), and for-profit organizations (eg, Gilead Sciences).

Document Format and Communication Method

Communication methods were categorized as brochures, information sheets, or websites by manually assessing each document. To identify brochures, we first searched for the keyword "brochure" in the URLs, and then, we manually verified other documents with similar presentations of information. Brochures were distinguished from information sheets by their inclusion of images, figures, or pictures. To identify the information sheets, we searched for the keywords "factsheet" and "sheet" from the remaining URLs. Then, we manually labeled other information sheets by searching for similar presentations of information, which contained only text (ie, no images, figures, and pictorial descriptions). Other resources were categorized as websites because they were hosted on a web page and did not meet the criteria for either brochure or information sheet.

PrEP Modality

All documents were manually categorized based on PrEP modality (ie, delivery methods) to allow comparison by modality. We classified documents with one or more keywords—"oral," "tablet," "pill," "Truvada," "Descovy," "daily PrEP," "PrEP 2-1-1," and "PrEP on demand"—in the title as "oral PrEP." We classified documents with the keywords "Cabotegravir," "Apretude," "long-acting," "injection PrEP," and "injectable" as "injectable PrEP." If a document had both types of keywords in the title, we categorized it as "oral and injectable PrEP." For documents without these keywords in the title, the modality was categorized based on the content of the document. Documents without PrEP modality descriptions were categorized as general PrEP.

Intended Audience

The intended audience was manually determined for each document, that is, patient, provider, or general public. We used document titles and the first paragraph to determine the target audience. To identify documents written for patients, we looked for the following titles: "Patient Information," "Medication Guide," "Patient Medication Information," "Information for the user," "Information for the patient," and "Overview for Patients." Then, we read the first paragraph to ensure its intended audience was patient and checked if it had statements similar to the following. "Read this carefully before you start taking [Drug name] and each time you get a refill. […]." Similarly, documents written for providers were identified if they had these or similar titles: "Clinical Guide," "Provider Frequently Asked Questions," "Provider FAQs," "Clinician Guidance," and "Guide for Medical Providers." We also read the first paragraph to check phrases similar to "Review this guide when prescribing [Drug name]." The intended audience for documents not meeting the criteria for patient or provider documents was determined as intended for the general public.

Measuring Readability

We used the Flesch-Kincaid grade level [45], Simple Measure of Gobbledygook Index [46], Coleman and Liau Index [47], and Automated Readability Index [48] to calculate the text readability, which is defined here as the grade level required to comprehend the text in the US education system. These indices are widely used in previous readability studies [23,24,26,49-53], and we applied them to the entire data set. To computationally perform readability analysis, the readability scores were first calculated based on the 4 readability indices using the open-source Python *textstat* package [54]. Given that different readability indices can generate a range of results, we used the mean of the 4 readability indices in our analyses to increase the reliability of our results but also report results generated by all 4 indices (Multimedia Appendix 1). We then conducted 5 sets of pairwise independent sample *t* tests to compare readability scores based on the method of obtaining the information, information source, document format and communication method, PrEP modality, and intended audience. When necessary, *P* values were adjusted using the prespecified Hommel procedure [55] to account for multiple comparisons within each set.

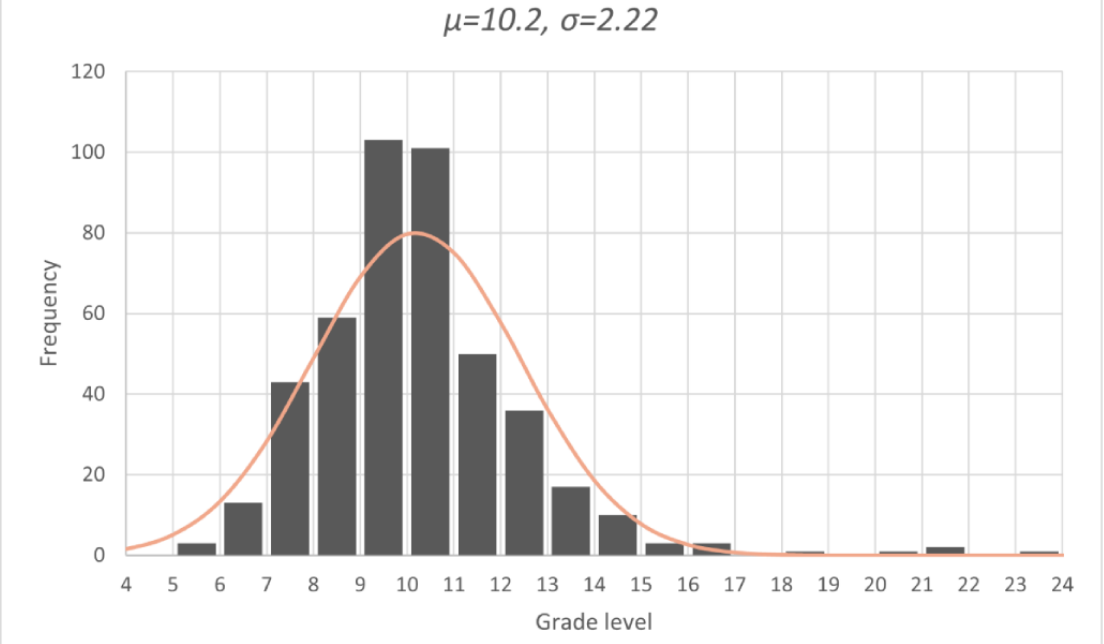
Results

Information Obtained Method

We collected a total of 463 documents, 194 from Twitter (ie, passively obtained information) and 269 from search engines (ie, actively obtained information). A total of 17 documents were duplicates. For analyses, we removed 1 set of the 17 duplicate documents. The documents were excluded in both sets, a total of 34 documents, when comparing the readability of documents based on methods for obtaining information. Using the average of the 4 reading level indices, the overall average reading level for web-based PrEP materials was a 10.20 grade level with an SE of 0.11. This is significantly above the recommended grade level of lower than ninth grade for the comprehension of health materials. The readability scores of materials retrieved from Twitter (9.8 grade level, SE 0.16) and search engines (10.5 grade level, SE 0.15) were significantly different (*P*=.002 by independent sample *t* tests).

We found that the distribution of readability scores for all documents was approximately normally distributed (ie, unimodal) with a slight skew to the right and a mode between the ninth and eleventh reading levels as shown in Figure 1. About 74% (n=328) of web-based sources required reading levels higher than ninth-grade.

Figure 1. The empirical distribution of the average readability level required for understanding web-based pre-exposure prophylaxis information.



Information Source

We found 172 documents from 45 for-profit entities (eg, Gilead Sciences and Nurx), 157 documents from 49 nongovernment organizations (eg, Mayo Clinic and Black AIDS Institute), 88 documents from 28 US governmental entities (eg, CDC and NIH), and 16 documents from 7 non-US governmental entities (eg, National Health Service of the United Kingdom and Health Service Executive-Ireland) or other public health organizations (eg, WHO and Joint United Nations Programme on HIV/AIDS). A total of 13 documents from organizations had no explicit indication of their affiliations or funding sources; they were excluded from our information source comparison analyses.

Some organizations and entities developed multiple PrEP-related documents, but they differed in content and topic. For example, some documents provided details about a particular PrEP drug (eg, Truvada and Descovy), while others provided information about accessing PrEP or PrEP cost. Still, others were different document types (eg, brochures). In our analyses, we treated each document as a separate data point.

Complete readability scores for each document are presented in Multimedia Appendix 1. The overall average grade reading level required to understand web-based PrEP information (10.2 grade level, SE 0.1) is considerably higher than the target grade level recommended by health agencies for health information. The mean grade reading levels and SEs (ie, of the 4 indices) from these organizations are displayed below and compared in Table 1. There was no significant difference in readability among organization types.

- Non-US government entities or other public health organizations: mean 10.6 (SE 0.7)
- For-profit entities: mean 10.4 (SE 0.2)
- The US-located government entities: mean 10.3 (SE 0.2)
- Nongovernment organizations: mean 9.9 (SE 0.2)

Table 1. Pairwise 2 sample *t* test of the average readability by source type.

Source type	*P* value	Adjusted *P* value
Non-US government entities or other public health organizations[a]	.22	.66
For-profit entities[a]	.04	.12
The US-located government entities[a]	.15	.46
Non-US government entities or other public health organizations[b]	.65	.74
For-profit entities[b]	.75	.75
Non-US government entities or other public health organizations[c]	.74	.74

[a] Versus nongovernment organizations.

[b] Versus the US-located government entities.

[c] Versus for-profit entities.

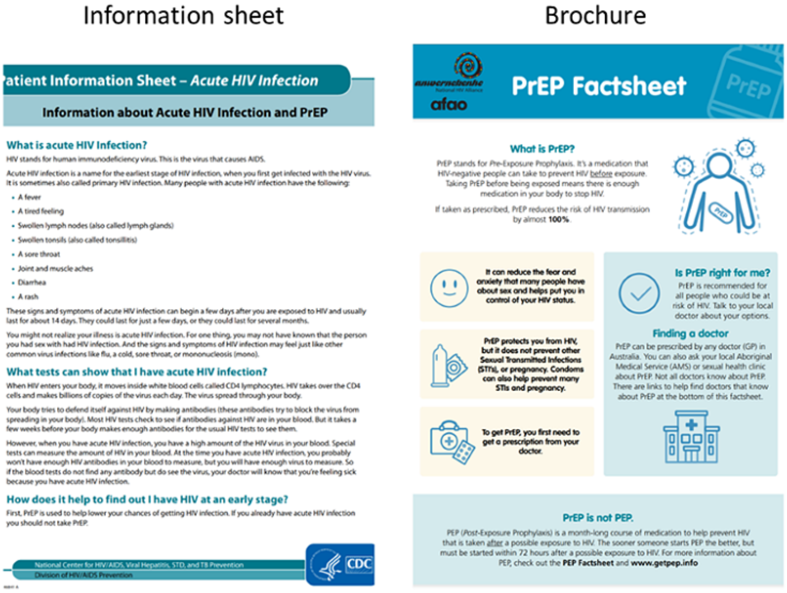
Document Format and Communication Method

The most common type of document was websites (n=384), including content from health organizations and health care professionals. Websites typically provide a comprehensive range of information on multiple related topics, such as HIV/AIDS and PrEP access, at a reading level of 10.2 grade (SE 0.46).

In our data set, information sheets (n=29) were in PDF file format and had titles containing phrases such as “Patient Medication,” “Medication Guide,” “Information for the user,” “Information sheet,” “Fact sheet,” “Facts,” and “Information booklet.” Compared to other document formats, they were also lengthy and detailed, including medication information and instructions for patient use. Most medication guides were written by the drug manufacturers (eg, Gilead Sciences and ViiV Healthcare) and FDA. Five information sheets were provided by nonprofit and non-US government health organizations. Information sheets were written at a reading level of 11.2 grade (SE 0.46).

Examples of each type of document are provided in Figure 2. Brochures (n=33) included user guides and leaflets containing commonly asked questions such as “What is PrEP” and “How to get PrEP.” Brochures were prepared by federal and state governmental organizations (n=15), non-US governmental organizations (n=2), nongovernmental organizations (n=10), and for-profit organizations (n=6). The 3 document types had significantly different readability levels: the readability of brochures was easier than both information sheets (adjusted *P*<.001) and websites (adjusted *P*=.003), with an average readability level at the recommended ninth grade level (SE 0.31).

Figure 2. Examples of document types used to disseminate pre-exposure prophylaxis (PrEP) information on the web: information sheet (top left), brochure (top right), website (bottom left), and website in a blog format (bottom right). More information about each document, including the source URL, can be found in Multimedia Appendix 1, rows 12, 272, 114, and 147. For a higher-resolution version of this figure, see Multimedia Appendix 2.



Information sheet

atient Information Sheet – *Acute HIV Infection*

Information about Acute HIV Infection and PrEP

What is acute HIV Infection?

HIV stands for human immunodeficiency virus. This is the virus that causes AIDS.

Acute HIV infection is a name for the earliest stage of HIV infection, when you first get infected with the HIV virus. It is sometimes also called primary HIV infection. Many people with acute HIV infection have the following:

- A fever
- A tired feeling
- Swollen lymph nodes (also called lymph glands)
- Swollen tonsils (also called tonsillitis)
- A sore throat
- Joint and muscle aches
- Diarrhea
- A rash

These signs and symptoms of acute HIV infection can begin a few days after you are exposed to HIV and usually last for about 14 days. They could last for just a few days, or they could last for several months.

You might not realize your illness is acute HIV infection. For one thing, you may not have known that the person you had sex with had HIV infection. And the signs and symptoms of HIV infection may feel just like other common virus infections like flu, a cold, sore throat, or mononucleosis (mono).

What tests can show that I have acute HIV infection?

When HIV enters your body, it moves inside white blood cells called CD4 lymphocytes. HIV takes over the CD4 cells and makes billions of copies of the virus each day. The virus spread through your body.

Your body tries to defend itself against HIV by making antibodies (these antibodies try to block the virus from spreading in your body). Most HIV tests check to see if antibodies against HIV are in your blood. But it takes a few weeks before your body makes enough antibodies for the usual HIV tests to see them.

However, when you have acute HIV infection, you have a high amount of the HIV virus in your blood. Special tests can measure the amount of HIV in your blood. At the time you have acute HIV infection, you probably won't have enough HIV antibodies in your blood to measure, but you will have enough virus to measure. So if the blood tests do not find any antibody but do see the virus, your doctor will know that you're feeling sick because you have acute HIV infection.

How does it help to find out I have HIV at an early stage?

First, PrEP is used to help lower your chances of getting HIV infection. If you already have acute HIV infection you should not take PrEP.

National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention
Division of HIV/AIDS Prevention

Brochure

PrEP Factsheet

What is PrEP?

PrEP stands for *Pre-Exposure Prophylaxis*. It's a medication that HIV-negative people can take to prevent HIV before exposure. Taking PrEP before being exposed means there is enough medication in your body to stop HIV.

If taken as prescribed, PrEP reduces the risk of HIV transmission by almost **100%**.

It can reduce the fear and anxiety that many people have about sex and helps put you in control of your HIV status.

PrEP protects you from HIV, but it does not prevent other Sexual Transmitted Infections (STIs), or pregnancy. Condoms can also help prevent many STIs and pregnancy.

To get PrEP, you first need to get a prescription from your doctor.

Is PrEP right for me?

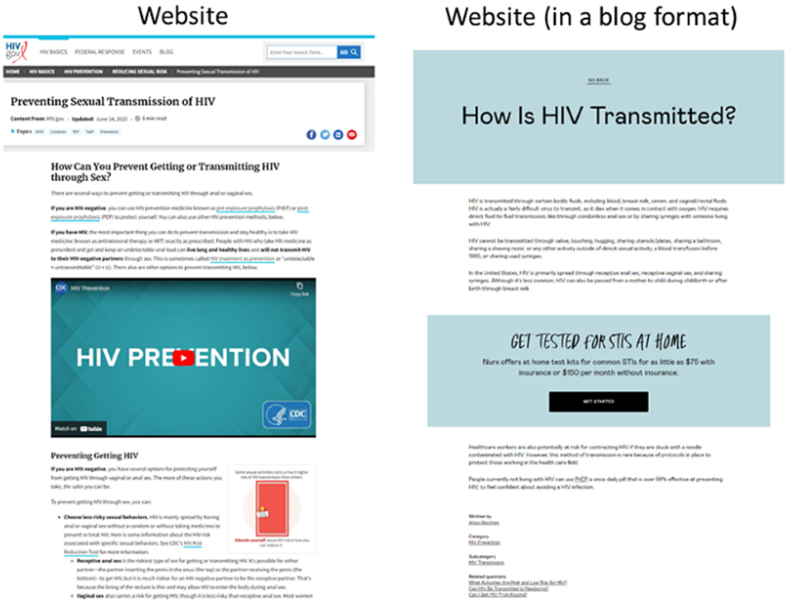
PrEP is recommended for all people who could be at risk of HIV. Talk to your local doctor about your options.

Finding a doctor

PrEP can be prescribed by any doctor (GP) in Australia. You can also ask your local Aboriginal Medical Service (AMS) or sexual health clinic about PrEP. Not all doctors know about PrEP. There are links to help find doctors that know about PrEP at the bottom of this factsheet.

PrEP is not PEP.

PEP (*Post-Exposure Prophylaxis*) is a month-long course of medication to help prevent HIV that is taken after a possible exposure to HIV. The sooner someone starts PEP the better, but must be started within 72 hours after a possible exposure to HIV. For more information about PEP, check out the **PEP Factsheet** and **www.getpep.info**



Website

Preventing Sexual Transmission of HIV

How Can You Prevent Getting or Transmitting HIV through Sex?

Preventing Getting HIV

Website (in a blog format)

How Is HIV Transmitted?

GET TESTED FOR STIS AT HOME

Nurx offers at home test kits for common STIs for as little as $75 with insurance or $150 per month without insurance.

PrEP Modality

We identified 220 documents relating to oral PrEP only, 14 documents relating to injectable PrEP only, and 14 documents mentioning both oral and injectable PrEP. The rest of the documents (n=198, 44%) did not mention a specific PrEP delivery method (ie, "general PrEP"); these latter documents were excluded from analyses. On average, the content on injectable PrEP (12.7 grade level, SE 0.9) was more difficult to comprehend than the content on oral PrEP (10.2 grade level, SE 0.2). Information on both forms of PrEP was intermediate in readability (10.8 grade level, SE 0.6). The comparison analysis results are shown in Table 2.

Table 2. Pairwise 2 sample *t* test of the average readability by PrEP[a] modality.

PrEP modality	*P* value	Adjusted *P* value
Injectable PrEP[b]	<.001	<.001
Oral and injectable PrEP[b]	.30	.30
Injectable PrEP[c]	.09	.09

[a] PrEP: pre-exposure prophylaxis.

[b] Versus oral PrEP.

[c] Versus oral and injectable PrEP.

Intended Audience

Most documents (n=426, 95.5%) were categorized as "general public" because they did not specify their intended audience. We found 14 documents specifically written for PrEP users and 6 specifically written for providers.

The readability grade level indices of content intended for PrEP providers (15.6 grade level, SE 1.59) were significantly higher when compared to the content intended for patients (10.2 grade level, SE 0.3; adjusted $P<.001$) and the general public (10.1 grade level, SE 0.1; adjusted $P<.001$). The reading level of content written for PrEP users was similar to that of materials for the general public (adjusted $P=.84$).

Discussion

Principal Findings

Our study highlights the challenging task of effectively educating diverse audiences about PrEP. Our results are consistent with previous work [25], which found that web-based PrEP information is often difficult to understand. We add to this previous work by including data disseminated in social media and stratifying web-based PrEP documents by how the PrEP document was obtained on the web, information source, document format and communication method, PrEP modality, and intended audience. We found PrEP materials from Twitter were significantly easier to comprehend than those identified by web searches, despite the fact Twitter was merely posting URLs of web-based information. It is possible that organizations selectively identify more comprehendible materials to promote through Twitter. The readability of web-based PrEP information from different organizational groups was all higher than recommended, and there was no significant difference between them (Table 1). This contrasts with a previous study that focused on different health conditions [26]. Informational documents intended for providers had significantly higher reading levels than those intended for patients or the general public. However, the materials intended for the patients and the public were still higher than the recommended ninth-grade reading level.

Brochures were the only type of PrEP materials that achieved the HHS and NIH target of ninth-grade reading level [13]. Brochures have been used in successful public health campaigns for facilitating behavior change, knowledge increase, and self-efficacy for other health conditions [56]. Brochures may be effective means of PrEP communication for most people who are not medical providers. However, brochures were not a common mode of conveying information: there were only 33 brochures in 446 unique documents that we identified.

As PrEP regimens become more complex (eg, injectable PrEP and 2-1-1 PrEP), educational materials might become more difficult to comprehend. For example, we identified that a higher reading level was required to understand materials about injectable PrEP. Yet, there is great interest in injectable PrEP: in a National HIV Behavioral survey of 314 people, injectable PrEP was preferred by 3 times more respondents compared to oral PrEP [31]. This calls for attention to improving the readability of materials describing injectable PrEP because difficulty in understanding injectable PrEP information could discourage people who would otherwise be interested in the new PrEP route of administration.

Our findings suggest that there is a need for health communication strategies and policies that support the development and dissemination of clear and concise PrEP information. Our finding suggests that health communication strategies can be enhanced by using brochures and plain language, which may include avoiding jargon. Policies can also be improved to support this effort. For instance, governments could increase funding for research into PrEP literacy and the development of PrEP communication guidelines. Additionally, governments could mandate that health care providers receive training on how to communicate about PrEP in a clear and easy-to-understand way for a variety of audiences. Together, these communication strategies and policies can help individuals to have access to clear and easy-to-understand information about PrEP, which can help to increase PrEP uptake and reduce HIV transmission.

Limitations

Our findings should be understood in the light of their limitations. First, individuals accessing PrEP information on the web may be a select subgroup of those interested in PrEP and might have higher or lower health literacy than those who seek information from other sources. However, given that the internet has become an increasingly popular resource for health information [32,57] and groups potentially eligible for PrEP [58] tend to have high internet use [59], it is likely that a substantial proportion of potential and current PrEP users would seek PrEP information on the web. Second, current general-purpose readability indices alone may not be a perfect measurement of comprehension and reading level [60,61], especially for assessing documents with nontext content, such as charts, graphs, and videos. For instance, we noted that all brochures in our data set contained images. Pictorial and graphic representations of information have been shown to be more effective than text-only messages [62]. Similarly, we found that some websites also contain PrEP-related informational videos; videos have been shown to be effective in modifying health behaviors, including promoting HIV testing [63]. In this study, we focused on textual information as text remains the primary medium for health communication on the web [64] and relied on the average scores of 4 readability indices to minimize the lack of sensitivity or breadth of measurements from using a single metric. Third, our analysis, although systematic, was not exhaustive. The web-based PrEP information included in our study was collected from a variety of sources, yet it is possible that some important information was missed. We limited our search to English language materials, and the first page of the search result imitated users' web-based behavior [43]. Lastly, we did not use generic names of PrEP medications, thus potentially not including PrEP information that only might use the generic names.

Future Directions

Despite public guidelines for writing health education materials, readability levels of web-based health materials, even those written by public health organizations, have either become more difficult [65] or reflected no improvement [19] with few exceptions [66]. AMA guidelines suggest addressing health literacy concerns by avoiding unnecessary details and lengthy background information [15]. This recommendation is consistent with our finding that documents in the brochure format, which were assessed as most readable, were concise, focused more on patient needs, and provided less background information. Readability is only one measure of the impact of health information, for example, measures of health literacy are evolving to include measures of the extent to which materials increase decision-making ability [67]. Thus, future research on health literacy should extend to whether educational materials are helpful to clients in making informed decisions.

Social media platforms have been widely used for public health communication and education, but research focused on measuring the effectiveness of social media as a channel for public health information is limited (with a few notable exceptions [68]). A better understanding of social media platforms' effectiveness should improve public health communication and the development of an impactful campaign for facilitating behavior change, knowledge increase, and self-efficacy related to public health needs—including PrEP uptake. Inequitable PrEP uptake persists among racial, ethnic, and gender minority communities [6,7]. Follow-up studies should (1) optimize the effectiveness of web-based educational materials in making informed decisions among health equity populations and (2) tailor messaging to address the information needs of specific populations to address inequitable PrEP uptake.

Future studies of the readability of web-based PrEP information should consider the following points. PrEP use and HIV prevention are global concerns, and it is important to understand how people in other languages are accessing and understanding information about PrEP. Future studies should examine materials in other languages. To provide a more comprehensive picture of the range of available web-based PrEP information, future work can expand the data collection process by expanding the number of result pages, search keywords, search engines, and search locations and by applying various filters to narrow and rank the search results. In addition, future work should consider exploring other measures of comprehension, especially for materials that include images, videos, and other nontext elements. We believe that these measures will strengthen the findings and provide a more comprehensive picture of how people are understanding information about PrEP.

Conclusions

The overall reading level of PrEP information found on the web was above that recommended for most potential users; if unaddressed, other challenges to supporting PrEP use by those most likely to benefit might be made more difficult because of a gap in comprehension of web-based information. The readability of web-based PrEP information needs to be improved to comply with health communication recommendations, reduce the barrier of a health literacy gap, facilitate informed decisions by those with a need for PrEP, and counter the HIV epidemic.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The readability scores for each document collected from search engine and Twitter.
[DOCX File , 247 KB - publichealth_v9i1e48630_app1.docx]

Multimedia Appendix 2
Examples of document types used to disseminate pre-exposure prophylaxis (PrEP) information on the web.
[DOCX File , 1340 KB - publichealth_v9i1e48630_app2.docx]

References

1. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, Bangkok Tenofovir Study Group. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet 2013;381(9883):2083-2090. [doi: 10.1016/S0140-6736(13)61127-7] [Medline: 23769234]
2. Haberer JE, Baeten JM, Campbell J, Wangisi J, Katabira E, Ronald A, et al. Adherence to antiretroviral prophylaxis for HIV prevention: a substudy cohort within a clinical trial of serodiscordant couples in East Africa. PLoS Med 2013;10(9):e1001511 [FREE Full text] [doi: 10.1371/journal.pmed.1001511] [Medline: 24058300]
3. Smith DK, Herbst JH, Rose CE. Estimating HIV protective effects of method adherence with combinations of preexposure prophylaxis and condom use among African American men who have sex with men. Sex Transm Dis 2015;42(2):88-92 [FREE Full text] [doi: 10.1097/OLQ.0000000000000238] [Medline: 25585067]
4. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019;321(9):844-845 [FREE Full text] [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
5. Sullivan PS, Giler RM, Mouhanna F, Pembleton ES, Guest JL, Jones J, et al. Trends in the use of oral emtricitabine/tenofovir disoproxil fumarate for pre-exposure prophylaxis against HIV infection, United States, 2012-2017. Ann Epidemiol 2018;28(12):833-840 [FREE Full text] [doi: 10.1016/j.annepidem.2018.06.009] [Medline: 30037634]
6. Siegler AJ, Mouhanna F, Giler RM, Weiss K, Pembleton E, Guest J, et al. The prevalence of pre-exposure prophylaxis use and the pre-exposure prophylaxis-to-need ratio in the fourth quarter of 2017, United States. Ann Epidemiol 2018;28(12):841-849 [FREE Full text] [doi: 10.1016/j.annepidem.2018.06.005] [Medline: 29983236]
7. Frieden TR, Foti KE, Mermin J. Applying public health principles to the HIV epidemic—how are we doing? N Engl J Med 2015;373(23):2281-2287 [FREE Full text] [doi: 10.1056/NEJMms1513641] [Medline: 26624243]
8. Sullivan PS, Siegler AJ. What will it take to meet UNAIDS targets for preexposure prophylaxis users? Curr Opin Infect Dis 2022;35(1):1-8. [doi: 10.1097/QCO.0000000000000809] [Medline: 34879049]
9. Sullivan PS, Mena L, Elopre L, Siegler AJ. Implementation strategies to increase PrEP uptake in the south. Curr HIV/AIDS Rep 2019;16(4):259-269 [FREE Full text] [doi: 10.1007/s11904-019-00447-4] [Medline: 31177363]
10. Lopez C, Kim B, Sacks K. Health literacy in the United States: enhancing assessments and reducing disparities. Milken Institute. 2022. URL: https://milkeninstitute.org/sites/default/files/2022-05/Health_Literacy_United_States_Final_Report.pdf [accessed 2023-08-04]
11. Healthy people 2020. US Department of Health and Human Services. Office of Disease Prevention and Health Promotion. 2020. URL: https://wayback.archive-it.org/5774/20220414160933/https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-health/interventions-resources/health-literacy [accessed 2023-08-07]
12. Walsh TM, Volsko TA. Readability assessment of internet-based consumer health information. Respir Care 2008;53(10):1310-1315 [FREE Full text] [Medline: 18811992]
13. Arkin EB. Making health communication programs work. U.S. Department of Health & Human Services, National Institutes of Health, National Cancer Institute. 2009. URL: https://www.cancer.gov/publications/health-communication/pink-book.pdf [accessed 2023-08-04]
14. How to write easy-to-read health materials. National Institutes of Health. 2015. URL: http://www.webcitation.org/6etJOZbZM [accessed 2023-08-04]
15. Weiss B. Health literacy. American Medical Association Foundation and American Medical Association. 2003. URL: http://lib.ncfh.org/pdfs/6617.pdf [accessed 2023-08-04]
16. Fischhoff B, Brewer NT, Downs JS. Communicating risks and benefits: an evidence-based user's guide. US Department of Health and Human Services, Food and Drug Administration (FDA). Silver Spring, MD; 2011. URL: https://www.fda.gov/about-fda/reports/communicating-risks-and-benefits-evidence-based-users-guide [accessed 2023-08-04]
17. Skierkowski DD, Florin P, Harlow LL, Machan J, Ye Y. A readability analysis of online mental health resources. Am Psychol 2019;74(4):474-483 [FREE Full text] [doi: 10.1037/amp0000324] [Medline: 30299121]
18. Lipari M, Berlie H, Saleh Y, Hang P, Moser L. Understandability, actionability, and readability of online patient education materials about diabetes mellitus. Am J Health Syst Pharm 2019;76(3):182-186. [doi: 10.1093/ajhp/zxy021] [Medline: 31408087]
19. Sabharwal S, Badarudeen S, Unes Kunju S. Readability of online patient education materials from the AAOS web site. Clin Orthop Relat Res 2008;466(5):1245-1250 [FREE Full text] [doi: 10.1007/s11999-008-0193-8] [Medline: 18324452]

20. MacLean SA, Basch CH, Ethan D, Garcia P. Readability of online information about HPV immunization. Hum Vaccin Immunother 2019;15(7-8):1505-1507 [FREE Full text] [doi: 10.1080/21645515.2018.1502518] [Medline: 30024828]
21. Ayyaswami V, Padmanabhan D, Patel M, Prabhu AV, Hansberry DR, Agarwal N, et al. A readability analysis of online cardiovascular disease-related health education materials. Health Lit Res Pract 2019;3(2):e74-e80 [FREE Full text] [doi: 10.3928/24748307-20190306-03] [Medline: 31049489]
22. Park A, Zhu SH, Conway M. The readability of electronic cigarette health information and advice: a quantitative analysis of web-based information. JMIR Public Health Surveill 2017;3(1):e1 [FREE Full text] [doi: 10.2196/publichealth.6687] [Medline: 28062390]
23. D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. Arch Pediatr Adolesc Med 2001;155(7):807-812 [FREE Full text] [doi: 10.1001/archpedi.155.7.807] [Medline: 11434848]
24. Agarwal N, Hansberry DR, Sabourin V, Tomei KL, Prestigiacomo CJ. A comparative analysis of the quality of patient education materials from medical specialties. JAMA Intern Med 2013;173(13):1257-1259 [FREE Full text] [doi: 10.1001/jamainternmed.2013.6060] [Medline: 23689468]
25. Kecojevic A, Basch CH, Garcia P. Readability analysis of online health information on preexposure prophylaxis (PrEP). Public Health 2020;182:53-55. [doi: 10.1016/j.puhe.2020.02.002] [Medline: 32171091]
26. Cochrane ZR, Gregory P, Wilson A. Readability of consumer health information on the internet: a comparison of U.S. government-funded and commercially funded websites. J Health Commun 2012;17(9):1003-1010. [doi: 10.1080/10810730.2011.650823] [Medline: 22512714]
27. Keselman A, Smith CA. A classification of errors in lay comprehension of medical documents. J Biomed Inform 2012;45(6):1151-1163 [FREE Full text] [doi: 10.1016/j.jbi.2012.07.012] [Medline: 22925723]
28. Smith CA, Hetzel S, Dalrymple P, Keselman A. Beyond readability: investigating coherence of clinical text for consumers. J Med Internet Res 2011;13(4):e104 [FREE Full text] [doi: 10.2196/jmir.1842] [Medline: 22138127]
29. Gombe MM, Cakouros BE, Ncube G, Zwangobani N, Mareke P, Mkwamba A, et al. Key barriers and enablers associated with uptake and continuation of oral pre-exposure prophylaxis (PrEP) in the public sector in Zimbabwe: qualitative perspectives of general population clients at high risk for HIV. PLoS One 2020;15(1):e0227632 [FREE Full text] [doi: 10.1371/journal.pone.0227632] [Medline: 31931514]
30. Muhumuza R, Ssemata AS, Kakande A, Ahmed N, Atujuna M, Nomvuyo M, et al. Exploring perceived barriers and facilitators of PrEP uptake among young people in Uganda, Zimbabwe, and South Africa. Arch Sex Behav 2021;50(4):1729-1742 [FREE Full text] [doi: 10.1007/s10508-020-01880-y] [Medline: 33954824]
31. Levy ME, Patrick R, Gamble J, Rawls A, Opoku J, Magnus M, et al. Willingness of community-recruited men who have sex with men in Washington, DC to use long-acting injectable HIV pre-exposure prophylaxis. PLoS One 2017;12(8):e0183521 [FREE Full text] [doi: 10.1371/journal.pone.0183521] [Medline: 28827821]
32. Fox S. Health topics: 80% of internet users look for health information online. Pew Research Center’s Internet & American Life Project. 2011. URL: http://www.webcitation.org/6gbUGjFS5 [accessed 2023-08-04]
33. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. BMJ Glob Health 2020;5(10):e004206 [FREE Full text] [doi: 10.1136/bmjgh-2020-004206] [Medline: 33097547]
34. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2023-08-04]
35. Lober WB, Flowers JL. Consumer empowerment in health care amid the internet and social media. Semin Oncol Nurs 2011;27(3):169-182. [doi: 10.1016/j.soncn.2011.04.002] [Medline: 21783008]
36. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. BMC Public Health 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]
37. Gatewood J, Monks SL, Singletary CR, Vidrascu E, Moore JB. Social media in public health: strategies to distill, package, and disseminate public health research. J Public Health Manag Pract 2020;26(5):489-492 [FREE Full text] [doi: 10.1097/PHH.0000000000001096] [Medline: 32732723]
38. Harris JK, Hawkins JB, Nguyen L, Nsoesie EO, Tuli G, Mansour R, et al. Using Twitter to identify and respond to food poisoning: the food safety STL project. J Public Health Manag Pract 2017;23(6):577-580 [FREE Full text] [doi: 10.1097/PHH.0000000000000516] [Medline: 28166175]
39. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. Pew Research Center. 2019. URL: https://www.pewresearch.org/short-reads/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [accessed 2023-08-04]
40. Roesslein J. Tweepy. 2009. URL: http://www.tweepy.org [accessed 2023-08-04]
41. Gilead Sciences, Inc.. U.S. Food and Drug Administration approves Gilead’s Truvada® for reducing the risk of acquiring HIV: first agent indicated for uninfected adults at high risk of acquiring HIV through Sex. Business Wire. 2012. URL: https://tinyurl.com/y858ew4t [accessed 2023-08-04]
42. Kakalou C, Lazarus JV, Koutkias V. Mining social media for perceptions and trends on HIV pre-exposure prophylaxis. Stud Health Technol Inform 2019;264:959-963. [doi: 10.3233/SHTI190366] [Medline: 31438066]

43. van Deursen AJAM, van Dijk JAGM. Using the internet: skill related problems in users' online behavior. Interact Comput 2009;21(5-6):393-402. [doi: 10.1016/j.intcom.2009.06.005]
44. Crunchbase. URL: https://web.archive.org/web/20230617200359/https://www.crunchbase.com/ [accessed 2023-08-04]
45. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas (automated readability index, fog count, and Flesch reading ease formula) for navy enlisted personnel. Institute for Simulation and Training. 1975. URL: https://stars.library.ucf.edu/istlibrary/56/ [accessed 2023-08-04]
46. McLaughlin GH. SMOG grading—a new readability formula. J Read 1969;12(8):639-646.
47. Coleman M, Liau TL. A computer readability formula designed for machine scoring. J Appl Psychol 1975;60(2):283-284. [doi: 10.1037/h0076540]
48. Smith EA, Senter RJ. Automated readability index. AMRL TR 1967:1-14. [Medline: 5302480]
49. Meade CD, Byrd JC. Patient literacy and the readability of smoking education literature. Am J Public Health 1989;79(2):204-206. [doi: 10.2105/ajph.79.2.204] [Medline: 2913844]
50. Malouff J, Gabrilowitz D, Schutte N. Readability of health warnings on alcohol and tobacco products. Am J Public Health 1992;82(3):464. [doi: 10.2105/ajph.82.3.464-a] [Medline: 1536374]
51. Tian C, Champlin S, Mackert M, Lazard A, Agrawal D. Readability, suitability, and health content assessment of web-based patient education materials on colorectal cancer screening. Gastrointest Endosc 2014;80(2):284-290. [doi: 10.1016/j.gie.2014.01.034] [Medline: 24674352]
52. Yin HS, Gupta RS, Tomopoulos S, Wolf MS, Mendelsohn AL, Antler L, et al. Readability, suitability, and characteristics of asthma action plans: examination of factors that may impair understanding. Pediatrics 2013;131(1):e116-e126. [doi: 10.1542/peds.2012-0612] [Medline: 23209106]
53. Terblanche M, Burgess L. Examining the readability of patient-informed consent forms. Open Access J Clin Trials 2010;2:157-162 [FREE Full text] [doi: 10.2147/oajct.s13608]
54. Bansal S, Aggarwal C. textstat. MIT. 2015. URL: https://pypi.python.org/pypi/textstat [accessed 2023-08-04]
55. Wright SP. Adjusted P-values for simultaneous inference. Biometrics 1992;48(4):1005-1013. [doi: 10.2307/2532694]
56. Anker AE, Feeley TH, McCracken B, Lagoe CA. Measuring the effectiveness of mass-mediated health campaigns through meta-analysis. J Health Commun 2016;21(4):439-456. [doi: 10.1080/10810730.2015.1095820] [Medline: 26953782]
57. Fox S, Duggan M. Health online 2013: one in three American adults have gone online to figure out a medical condition. Pew Research Center. 2013. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2023-08-04]
58. U.S. Centers for Disease Control and Prevention. Core indicators for monitoring the ending the HIV epidemic initiative (early release). HIV Surveillance Data Tables. 2021. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance-data-tables/vol-2-no-2/cdc-hiv-surveillance-tables-vol-2-no-2.pdf [accessed 2023-08-04]
59. Perrin A, ATSKE S. About three-in-ten U.S. adults say they are ‘almost constantly’ online. Pew Research Center. 2021. URL: https://www.pewresearch.org/short-reads/2021/03/26/about-three-in-ten-u-s-adults-say-they-are-almost-constantly-online/ [accessed 2023-08-04]
60. Meade CD, Smith CF. Readability formulas: cautions and criteria. Patient Educ Couns 1991;17(2):153-158. [doi: 10.1016/0738-3991(91)90017-y]
61. Davison A, Kantor RN. On the failure of readability formulas to define readable texts: a case study from adaptations. Read Res Q 1982;17(2):187-209. [doi: 10.2307/747483]
62. Hammond D. Health warning messages on tobacco products: a review. Tob Control 2011;20(5):327-337 [FREE Full text] [doi: 10.1136/tc.2010.037630] [Medline: 21606180]
63. Tuong W, Larsen ER, Armstrong AW. Videos to influence: a systematic review of effectiveness of video-based education in modifying health behaviors. J Behav Med 2014;37(2):218-233. [doi: 10.1007/s10865-012-9480-7] [Medline: 23188480]
64. Himelboim I, McCreery S. New technology, old practices: examining news websites from a professional perspective. Convergence 2012;18(4):427-444. [doi: 10.1177/1354856511429648]
65. Zhang D, Earp BE, Kilgallen EE, Blazar P. Readability of online hand surgery patient educational materials: evaluating the trend since 2008. J Hand Surg Am 2022;47(2):186.e1-186.e8 [FREE Full text] [doi: 10.1016/j.jhsa.2021.03.025] [Medline: 34023192]
66. Wong K, Levi JR. Readability trends of online information by the American Academy of Otolaryngology-Head and Neck Surgery Foundation. Otolaryngol Head Neck Surg 2017;156(1):96-102. [doi: 10.1177/0194599816674711] [Medline: 28045633]
67. Healthy people 2010 final review. U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/healthy_people/hp2010/hp2010_final_review.htm [accessed 2023-08-04]
68. Al-Dmour H, Masa'deh R, Salman A, Abuhashesh M, Al-Dmour R. Influence of social media platforms on public health protection against the COVID-19 pandemic via the mediating effects of public health awareness and behavioral changes: integrated model. J Med Internet Res 2020;22(8):e19996 [FREE Full text] [doi: 10.2196/19996] [Medline: 32750004]

Abbreviations

AMA: American Medical Association

CDC: Centers for Disease Control and Prevention
FAQ: frequently asked question
FDA: US Food and Drug Administration
HHS: US Department of Health and Human Services
HSE: Health and Safety Executive
NIH: National Institutes of Health
PrEP: pre-exposure prophylaxis
WHO: World Health Organization

Edited by A Mavragani; submitted 01.05.23; peer-reviewed by V Cantos, S Ha; comments to author 31.05.23; revised version received 21.06.23; accepted 26.06.23; published 04.09.23.

Please cite as:
Park A, Sayed F, Robinson P, Elopre L, Ge Y, Li S, Grov C, Sullivan PS
Health Information on Pre-Exposure Prophylaxis From Search Engines and Twitter: Readability Analysis
JMIR Public Health Surveill 2023;9:e48630
URL: https://publichealth.jmir.org/2023/1/e48630
doi:10.2196/48630
PMID:37665621

Original Paper

Moderating Effect of Pre-Exposure Prophylaxis Use on the Association Between Sexual Risk Behavior and Perceived Risk of HIV Among Brazilian Gay, Bisexual, and Other Men Who Have Sex With Men: Cross-Sectional Study

Kevin James Blair[1,2], MSc, MD; Thiago S Torres[3], PhD; Brenda Hoagland[3], MD, PhD; Daniel R B Bezerra[3], MSc; Valdilea G Veloso[3], MD, PhD; Beatriz Grinsztejn[3], MD, PhD; Jesse Clark[1], MD; Paula M Luz[3], MD, PhD

[1]South American Program in HIV Prevention Research, Division of Infectious Diseases, Department of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[2]Department of Surgery, University of California Los Angeles, Los Angeles, CA, United States

[3]Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Corresponding Author:
Kevin James Blair, MSc, MD
South American Program in HIV Prevention Research
Division of Infectious Diseases, Department of Medicine
University of California Los Angeles
10833 Le Conte Ave
72-227 CHS
Los Angeles, CA, 90095
United States
Phone: 1 6362368515
Email: KBlair@mednet.ucla.edu

Abstract

Background: Gay, bisexual, and other men who have sex with men (MSM) with a higher perceived risk of HIV are more aware of and willing to use pre-exposure prophylaxis (PrEP). PrEP is an effective HIV prevention strategy, but there is a lack of data on how PrEP use might moderate the relationship between sexual risk behavior and perceived risk of HIV. Moreover, most studies measure perceived risk of HIV via a single question.

Objective: We estimated the moderating effect of PrEP use on the association between sexual risk behavior and perceived risk of HIV, measured with the 8-item Perceived Risk of HIV Scale (PRHS), among Brazilian MSM.

Methods: A cross-sectional, web-based survey was completed by Brazilian Hornet app users aged ≥18 years between February and March 2020. We included data from cisgender men who reported sex with men in the previous 6 months. We evaluated the moderating effect of current PrEP use on the association between sexual risk behavior, measured via the HIV Incidence Risk Index for MSM (HIRI-MSM), and perceived risk of HIV, measured by the PRHS. Higher HIRI-MSM (range 0-45) and PRHS (range 10-40) scores indicate greater sexual behavioral risk and perceived risk of HIV, respectively. Both were standardized to *z* scores for use in multivariable linear regression models.

Results: Among 4344 cisgender MSM, 448 (10.3%) were currently taking PrEP. Current PrEP users had a higher mean HIRI-MSM score (mean 21.0, SD 9.4 vs mean 13.2, SD 8.1; $P<.001$) and a lower mean PRHS score (mean 24.6, SD 5.1 vs mean 25.9, SD 4.9; $P<.001$) compared to those not currently taking PrEP. In the multivariable model, greater HIRI-MSM scores significantly predicted increased PRHS scores ($\beta=.26$, 95% CI 0.22-0.29; $P<.001$). PrEP use moderated the association between HIRI-MSM and PRHS score (interaction term $\beta=–.30$, 95% CI –0.39 to –0.21; $P<.001$), such that higher HIRI-MSM score did not predict higher PRHS score among current PrEP users.

Conclusions: Our results suggest current PrEP users have confidence in PrEP's effectiveness as an HIV prevention strategy. PrEP's effectiveness, positive psychological impact, and the frequent HIV testing and interaction with health services required of PrEP users may jointly influence the relationship between sexual risk behavior and perceived risk of HIV among PrEP users.

(JMIR Public Health Surveill 2023;9:e45134) doi:10.2196/45134

KEYWORDS

HIV prevention; men who have sex with men; pre-exposure prophylaxis; Latin America; risk factors; risk perception; HIV; gay; prevention; health service

Introduction

Tenofovir-based oral pre-exposure prophylaxis (PrEP) is highly effective in preventing HIV infection in gay, bisexual, and other men who have sex with men (MSM) [1,2]. Higher perceived risk of HIV has been shown to be associated with awareness and acceptability of [3,4], interest in [5], and willingness to use PrEP [6-8] and is influential in decisions to initiate [9-13], continue or discontinue [14-17], and adhere to PrEP [18]. Taking PrEP has in turn also been shown to be associated with lower perceived risk and perceived severity of HIV [19,20] and lower sexual anxiety [21-23].

While several studies have explored the impact of PrEP use on sexual risk behaviors, such as condomless anal intercourse (CAI) [17,22,24,25], few have considered how PrEP use may moderate the relationship between sexual risk behavior and perceived risk of HIV. Potential PrEP users in the PrEP Brasil study with a higher perceived risk of HIV were more likely to say that they would discontinue condom use if they took PrEP [26]. Another study found that a higher perceived risk of HIV among PrEP users correlated with lower condom use but higher PrEP adherence [18]. A study from Amsterdam found PrEP use was associated with a lower perceived risk of HIV while sexual risk behaviors, such as CAI, were associated with a higher perceived risk, but the influence of PrEP use on the relationship between sexual risk behavior and perceived risk of HIV was not considered [19].

Most research on the perceived risk of HIV and PrEP use has measured perceived risk using a single question [5,6,18-20,27-30]. However, single question measures may be inadequate, given that perceived risk of HIV is a complex, multidimensional concept [20,27,29,31]. To address this, the 8-item Perceived Risk of HIV Scale (PRHS) was developed and subsequently validated for use in European and Brazilian Portuguese [32-34]. The PRHS covers multiple dimensions of the perceived risk of HIV, including cognitive likelihood assessment (ie, the chance of infection), intuitive assessment (ie, worry about infection), and salience of risk (ie, having thought about infection) [32].

Among Latin American countries, Brazil is a middle-income country of 203 million people and was the first country in the region to offer free national access to PrEP to eligible populations under its public health system [35,36]. As of June 2023, there were more than 64,000 PrEP users in the country, 82% of whom were MSM [13,37]. One recent study found that more than half of PrEP users in Brazil thought they had no risk of acquiring HIV, but, as with most other studies, a single question was used to measure perceived risk of HIV and the association between sexual behavior and perceived risk was not considered [38]. In this study, we used the PRHS to evaluate the moderating effect of current PrEP use on the association between sexual risk behavior and perceived risk of HIV among Brazilian MSM in the year 2020.

Methods

Study Design and Population

We administered a cross-sectional, web-based survey to a convenience sample of users of Hornet, a geosocial gay social network app, between February and March 2020. The survey was administered via Alchemer [39] and contained 118 questions written in Brazilian Portuguese. Requests for survey completion were sent twice and responses were collected over a 35-day period, rather than defining a sample size a priori. Additional study design details were described previously [40].

The overall study population included Hornet users ≥18 years old living in Brazil with completed surveys. We excluded those who incorrectly answered any of the 5 attention questions, which asked respondents to select a specific answer: “This question is merely a check. Please select option A from the responses below” [41]. People living with HIV were excluded. Additionally, we focused on cisgender men who reported having sex with other men in the previous 6 months, excluding those who self-identified with other genders (eg, cisgender woman, transgender man, transgender woman, or nonbinary).

Variables

Sociodemographics

Age, race, sexual orientation, education level, monthly income, Brazilian state of residence, and residence in the state’s capital city metropolitan area were collected. Responses for race included White, Black, Pardo (mixed race), Indigenous, Asian (Japanese, Chinese, Korean, among others), or prefer not to respond. Income was asked in reference to the 2020 monthly minimum wage (MW) in Brazil (BRL1039, ~US $190), which was grouped into low-income (no salary, 1-2× MW), middle-income (2-6× MW), and high-income (>6× MW).

Sexual Health and Behavior

We asked about the timing of the most recent HIV test, which was categorized as within the last 6 months, more than 6 months ago, or never tested. Knowledge about HIV was measured via the HIV/AIDS Knowledge Assessment (HIV-KA) tool, which is scored 0 to 12, with higher scores indicating greater HIV knowledge [40,42,43]. Respondents indicated whether they had a steady partner and, if so, their steady partner’s HIV status. We also asked about each of the following over the previous 6 months: number of male sex partners, any male sex partners known to be living with HIV, receptive CAI, sexually transmitted infection (STI) diagnoses (syphilis, urethral or rectal gonorrhea, urethral or rectal chlamydia), transactional sex (sex for money or other good), or chemsex (see Table 2 footnote).

Level of engagement in sexual risk behaviors was estimated via the HIV Incidence Risk Index for MSM (HIRI-MSM) score, calculated based on respondent age and the following behaviors with men in the previous 6 months: number of sex partners, number of sex partners living with HIV, receptive CAI with any HIV-status partner, insertive CAI with a partner living with HIV, and use of stimulants (cocaine; crack, pasta básica, or oxy; ecstasy; methamphetamines; gamma hydroxybutyrate or gamma butyrolactone; poppers; or other inhalants) [44,45]. HIRI-MSM is scored from 0 to 45, and a score of >10 has been considered an indication that the respondents may benefit from PrEP.

PrEP Use

We asked respondents if they had ever heard of PrEP (awareness), and those who responded affirmatively were asked about PrEP use via the question “Are you taking or have you taken PrEP?” Options included “No, I have never taken PrEP,” “Yes, I took PrEP but stopped,” or “Yes, I am currently taking PrEP.” Respondents who were not aware of PrEP were grouped with never PrEP users. For this analysis, never and past PrEP use were grouped together as “no current PrEP use.” To evaluate the robustness of the results from the regression models, a sensitivity analysis was performed in which we excluded past PrEP users from the regression models. Current PrEP users were asked about adherence based on the number of days they had taken PrEP over the previous week, categorized as every day, 4 to 6 days, or 3 or fewer days.

Outcome

Perceived risk of HIV was measured via the PRHS [32,34]. The PRHS consists of 8 Likert-scale questions covering different dimensions of perceived risk, as described in the introduction. Prior work showed the scale to be valid for use among MSM in Brazil [34]. Total scores range from 10 to 40, with higher scores indicating greater perceived risk of HIV.

Statistical Analyses

The outcome of this analysis was the PRHS score, and our primary aim was to assess the effect of the HIRI-MSM score and current PrEP use as explanatory variables as well as their potential interaction. All analyses were performed using R Software for Statistical Computing (version 4.0.3; R Core Team). Categorical variables were described using proportions and continuous variables using mean with SD and median with IQR. Cronbach α was used to assess the internal reliability of the PRHS and the HIV-KA tool [46]. We present and compare mean PRHS scores across various demographic, sexual health, and sexual risk behavior variables using the Student *t* test (2 variable categories) or ANOVA *F* test (3 or more categories). The comparison of PRHS scores across sexual risk behaviors was stratified according to current versus no current PrEP use.

We used multivariable linear regression models to estimate the association between HIRI-MSM score and PRHS score and the moderating effect of current PrEP use on that association. Due to the small sample size for individual variable categories, respondents with missing data and those who identified their race as Asian or Indigenous were excluded from the regression model. For inclusion in regression models, PRHS, HIRI-MSM, and HIV-KA scores were standardized to *z* scores by subtracting the mean overall score from each participant’s score and dividing the difference by the overall SD. Multivariable models were created by first including HIRI-MSM as the explanatory variable, followed by current PrEP use, and then an interaction term between the two. Finally, a full multivariable model included variables for which prior work had shown an association with perceived risk of HIV, including race, education, Brazilian state, sexual orientation, HIV knowledge, steady partner, transactional sex, and timing of last HIV test [4,19,22,27,30,34,47-52]. Age was not included as a covariate in the final model since it is accounted for in the HIRI-MSM score. We verified the fitted model’s assumptions using standard diagnostic plots and observed no violations. To aid in the visualization of the moderating effect of PrEP, we present a graphical representation of the interaction between the HIRI-MSM score and PrEP use created using the “effects” package in R [53].

Ethics Approval

This study received approval from the human subjects ethics committee at Instituto Nacional de Infectologia Evandro Chagas of Fundação Oswaldo Cruz (#CAAE 01777918.0.0000.5262) and was exempt from review by the University of California, Los Angeles institutional review board. All participants provided electronic informed consent before survey initiation.

Results

A total of 4344 cisgender MSM who reported sex with other men in the previous 6 months were included (Figure 1). The overall sample had a mean age of 34.4 (SD 10.2) years (Table 1). Most reported White race (n=2624, 60.4%), identified as gay or homosexual (n=3643, 83.9%), and had a university or higher-level education (n=2992, 68.9%). Most respondents lived in the Southeast region of Brazil, with the Brazilian states of São Paulo (n=2284, 52.6%) and Rio de Janeiro (n=919, 21.2%) being the most common. More than two-thirds lived in their state’s capital metropolitan area (n=2963, 68.2%). The majority (n=2321, 53.4%) of respondents had tested for HIV within the previous 6 months, and knowledge about HIV prevention was high (median HIV-KA score 11, IQR 10-12; α=.61). Nearly all (n=3996, 92%) were aware of PrEP, and 448 (10.3%) were currently taking PrEP.

The PRHS had high internal reliability (α=.75). Mean overall PRHS score was 25.8 (SD 4.9). The mean PRHS score was higher among gay-identifying MSM compared to other sexual orientations (26 vs 24.9; *P*<.001) and among university-educated respondents compared to those with secondary- or lower-level education (26 vs 25.4; *P*<.001). There was no significant difference in PRHS score across age, race, income, or Brazilian state. PRHS scores were higher among those who were aware of PrEP, as compared to those not aware (25.9 vs 24.3; *P*<.001), but were lower among MSM currently taking, compared to not currently taking PrEP (24.6 vs 25.9; *P*<.001).

Current PrEP users had significantly higher mean HIRI-MSM scores compared to MSM not currently taking PrEP (21.0 vs 13.2; *P*<.001) (Table 2). Among MSM not currently taking PrEP, higher PRHS scores significantly correlated with reporting a greater number of sex partners (*P*<.001), a steady partner living with HIV or with unknown HIV status (*P*<.001), any sex partner living with HIV (*P*<.001), receptive CAI (*P*<.001), STI diagnoses (*P*<.001), chemsex (*P*<.001), and transactional sex (*P*<.001). Conversely, among MSM currently taking PrEP, the only sexual risk behavior significantly associated with higher PRHS score was having a steady partner living with HIV or with unknown HIV status (*P*=.03). Among MSM currently taking PrEP, having taken PrEP every day in the previous week, as compared to 3 or fewer days, was associated with a lower mean PRHS score (24.4 vs 27.2; *P*=.05).

Higher HIRI-MSM scores significantly predicted higher PRHS scores in bivariate linear regression (model 1: β=.19, 95% CI 0.16-0.22; *P*<.001), meaning a 0.19 higher standardized PRHS score per 1 SD increase in standardized HIRI-MSM score (Table 3). The magnitude of the β coefficient for the HIRI-MSM score increased when PrEP use was included in the model (model 2: β=.23, 95% CI 0.20-0.26; *P*<.001), and when the HIRI:PrEP interaction term was added as an explanatory variable (model 3: β=.27, 95% CI 0.24-0.30; *P*<.001). In the fully adjusted model, standardized HIRI-MSM score had a positive association (β=.26, 95% CI 0.22-0.29; *P*<.001) and current PrEP use had a significant negative association (β=–.36, 95% CI –0.48 to –0.23; *P*<.001) with standardized PRHS scores. The estimated β coefficient for the HIRI-MSM:PrEP use interaction term indicated a significant moderating effect on the association between HIRI-MSM and PRHS score by current PrEP use (β=–0.30, 95% CI –0.39 to –0.21; *P*<.001). Results were unchanged when past PrEP users were excluded from regression models (Table S1 in Multimedia Appendix 1). The graphical representation of HIRI-MSM:PrEP use interaction shows how among those not using PrEP, higher PRHS scores were associated with higher HIRI-MSM scores. In contrast, among those taking PrEP, PRHS scores were unchanged by higher HIRI-MSM scores (Figure 2).

Among covariates in the full adjusted regression model (Table 3), having a steady partner who was HIV-negative compared to no steady partner was associated with a lower PRHS score (β=–.24, 95% CI –0.31 to –0.18; *P*<.001). Conversely, having a university or higher-level education compared to secondary- or lower-level (β=.12, 95% CI 0.06 to 0.19; *P*<.001), increasing HIV-KA score (β=.12, 95% CI 0.09 to 0.15; *P*<.001), identifying as gay compared to other sexual orientations (β=.11, 95% CI 0.03 to 0.19; *P*=.007), having a partner living with HIV or with unknown HIV status (β=.16, 95% CI 0.05 to 0.28; *P*=.004), and testing for HIV more than 6 months ago (β=.08, 95% CI 0.01 to 0.15; *P*=.02) were all associated with higher PRHS scores.

Figure 1. Flow diagram for study inclusion.

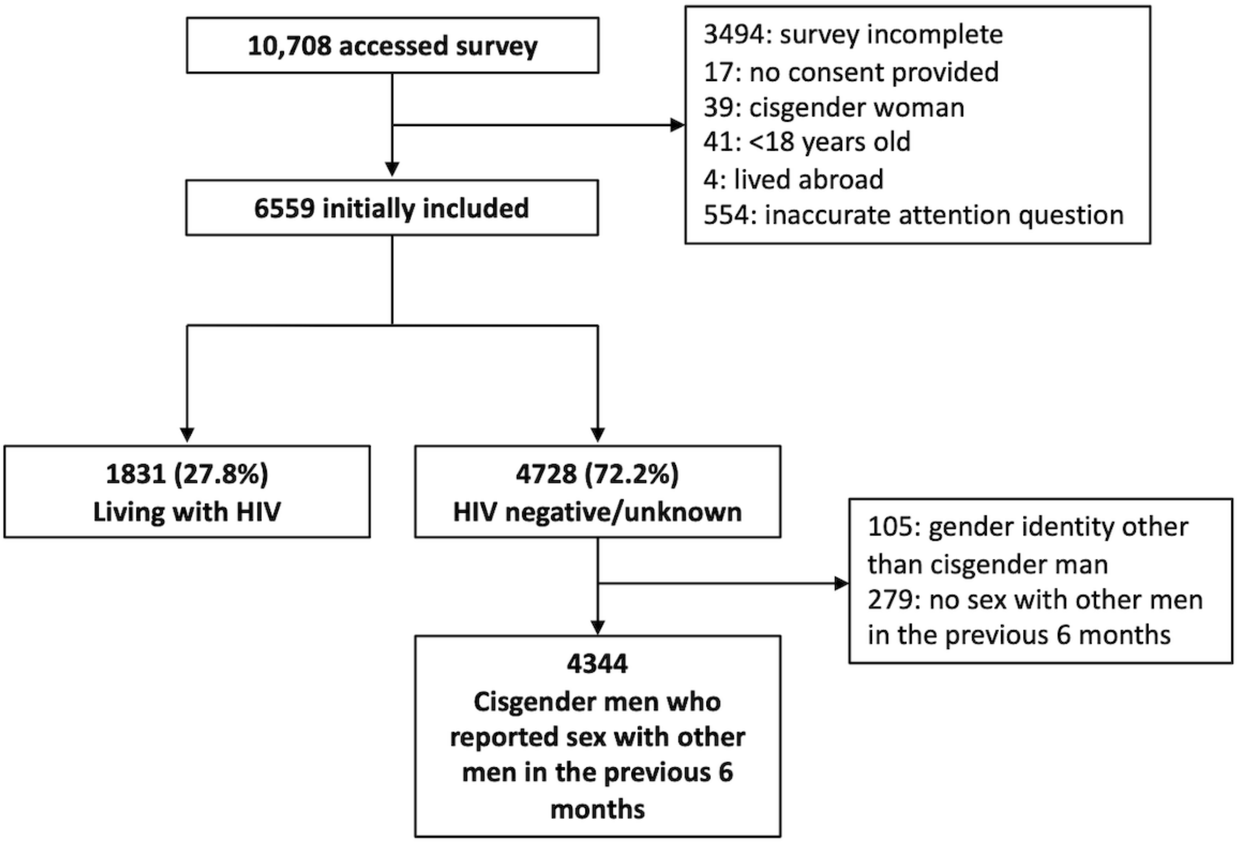


Table 1. Mean Perceived Risk of HIV Scale scores compared across demographic and select sexual health characteristics from a cross-sectional sample of Brazilian men who have sex with men in 2020.

Variables	Total, N=4344; n (%)	PRHS[a] score; mean (SD)	*P* value
Age group (years)			.58
18-24	644 (14.8)	25.6 (4.9)	
25-29	955 (22)	25.9 (4.8)	
30-39	1595 (36.7)	25.9 (4.9)	
≥40	1150 (26.5)	25.7 (4.9)	
Mean (SD)	34.4 (10.2)	N/A[b]	N/A
Race			.48
White	2,624 (60.4)	25.8 (4.8)	
Pardo	1112 (25.6)	25.8 (4.9)	
Black	406 (9.3)	26 (5)	
Asian	70 (1.6)	25.3 (4.6)	
Indigenous	30 (0.7)	24.3 (5.9)	
Unanswered	102 (2.3)	26.1 (5.3)	
Sexual orientation			<.001
Gay or homosexual	3643 (83.9)	26 (4.9)	
Other[c]	701 (16.1)	24.9 (4.8)	
Education level[d]			<.001
University or higher	2992 (68.9)	26 (4.9)	
Secondary or less	1276 (29.4)	25.4 (4.9)	
Unanswered	76 (1.7)	26.3 (5.2)	
Income level[e]			.39
High	1189 (27.4)	25.9 (4.9)	
Middle	1977 (45.5)	25.8 (4.8)	
Low	1178 (27.1)	25.6 (5.1)	
Brazilian state			.44
São Paulo	2284 (52.6)	25.8 (4.8)	
Rio de Janeiro	919 (21.2)	26 (5.1)	
Other	1141 (26.3)	25.7 (4.9)	
Live in state’s capital metropolitan area			.70
Yes	2963 (68.2)	25.8 (4.9)	
No	1381 (31.8)	25.7 (4.9)	
Timing of most recent HIV test			.06
Within the last 6 months	2321 (53.4)	25.7 (4.9)	
More than 6 months ago	1522 (35)	26 (4.8)	
Never tested	431 (9.9)	25.4 (4.8)	
Unanswered	70 (1.6)	25.9 (6.2)	
HIV-KA[f] score (range 0-12)			N/A
Mean (SD)	10.7 (1.6)	N/A	N/A
Median (IQR)	11 (10-12)	N/A	N/A
PrEP[g] Awareness			<.001

Variables	Total, N=4344; n (%)	PRHS[a] score; mean (SD)	*P* value
Yes	3996 (92)	25.9 (4.8)	
No	348 (8)	24.3 (5.3)	
PrEP use			<.001
Current	448 (10.3)	24.6 (5.1)	
No current	3896 (89.7)	25.9 (4.9)	
Past[h]	156 (3.6)	26.6 (4.9)	
Never[h]	3740 (86.1)	25.9 (4.9)	

[a] PRHS: Perceived Risk of HIV Scale.

[b] N/A: not applicable.

[c] "Other" includes responses of heterosexual, bisexual, pansexual, asexual, and other.

[d] Education levels of college and postgraduate were combined into "university or higher"; secondary, primary, or less than primary were combined into "secondary or less."

[e] We grouped no salary, 1×, and 2× the minimum wage as "low-income," 2-6× as "middle-income," and >6× as "high-income." Monthly minimum monthly wage in 2020 was BRL1039 (~US $190).

[f] HIV-KA: HIV/AIDS Knowledge Assessment; high scores indicate higher level of HIV knowledge.

[g] PrEP: pre-exposure prophylaxis.

[h] Past and never PrEP use were combined as no current PrEP use for further analyses.

Table 2. Mean Perceived Risk of HIV Scale scores compared across sexual risk behaviors and stratified by current versus no current pre-exposure prophylaxis use.

Variable	Current PrEP[a] use, n=448			No current PrEP use, n=3896		
	Respondents, n (%)	PRHS[b] score, mean (SD)	*P* value	Respondents, n (%)	PRHS score, mean (SD)	*P* value
HIRI-MSM[c] score			.82			<.001
Mean (SD)	21 (9.4)	N/A[d]		13.2 (8.1)	N/A	
>10	387 (86.4)	24.5 (5.1)		2309 (59.3)	26.8 (4.7)	
<10	61 (13.6)	24.7 (5.1)		1587 (40.7)	24.7 (4.8)	
Number of sex partners[e]			.57			<.001
1-5	111 (24.8)	25 (4.9)		2196 (56.4)	25.1 (4.8)	
6-10	105 (23.4)	24.4 (4.9)		818 (21)	26.5 (4.6)	
11-30	138 (30.8)	24.7 (5.4)		637 (16.4)	27.2 (4.5)	
30+	94 (21)	24 (5)		245 (6.3)	28.3 (5.4)	
Steady partner[e]			.03			<.001
No	305 (68.1)	24.5 (4.9)		2603 (66.8)	26.2 (4.8)	
Yes, HIV negative	96 (21.4)	23.8 (5.3)		1019 (26.2)	25 (4.9)	
Yes, living with HIV	33 (7.4)	26 (4.8)		120 (3.1)	27.4 (4.9)	
Yes, I don't know their HIV status	14 (3.1)	27.2 (6.8)		154 (4)	27 (4.6)	
Sex partner living with HIV[e]			.40			<.001
Yes	155 (34.6)	24.8 (5.3)		346 (8.9)	27.8 (4.8)	
No	293 (65.4)	24.4 (5)		3550 (91.1)	25.7 (4.8)	
Receptive CAI[e,f]			.67			<.001
Yes	294 (65.6)	24.6 (5.1)		1558 (40)	27 (4.7)	
No	154 (34.4)	24.4 (5.1)		2338 (60)	25.2 (4.8)	
STI[g] diagnosis[e]			.88			<.001
Yes	128 (28.6)	24.5 (5.3)		452 (11.6)	27.8 (4.8)	
No or unknown	320 (71.4)	24.6 (5)		3444 (88.4)	25.7 (4.8)	
Chemsex[e,h]			.91			<.001
Yes	168 (37.5)	24.5 (5.1)		837 (21.5)	27.2 (4.8)	
No	280 (62.5)	24.6 (5.1)		3059 (78.5)	25.6 (4.8)	
Transactional sex[i]			.09			.001
Yes	33 (7.4)	23.1 (5)		242 (6.2)	26.9 (5.7)	
No	415 (92.6)	24.7 (5.1)		3654 (93.8)	25.9 (4.8)	
PrEP adherence in past week[j]			.05			N/A
Every day	393 (87.7)	24.4 (5.1)		N/A	N/A	
4 to 6 days	39 (8.7)	25.4 (4.4)		N/A	N/A	
3 or fewer days	16 (3.6)	27.2 (6.3)		N/A	N/A	

[a] PrEP: pre-exposure prophylaxis.

[b] PRHS: Perceived Risk of HIV Scale.

[c] HIRI-MSM: HIV Incidence Risk Index for Men Who Have Sex With Men. The score is calculated based on respondent age and the following behaviors in the previous 6 months: number of male partners, number of male partners living with HIV, receptive condomless anal sex with any HIV-status partner, insertive condomless anal sex with a partner living with HIV, use of stimulants (cocaine; crack, basic paste, or oxy; ecstasy; methamphetamines [crystal or speed]; gamma hydroxybutyrate or gamma butyrolactone; poppers; or other inhalants). A score of >10 suggests the respondent may benefit from

pre-exposure prophylaxis.

[d] N/A: not applicable.

[e] Asked in reference to the previous 6 months.

[f] CAI: condomless anal intercourse.

[g] STI: sexually transmitted infection. The survey specified syphilis, urethral or rectal gonorrhea, and urethral or rectal chlamydia.

[h] Use of any of the following substances before or during sex: cocaine; crack, basic paste, or oxy; marijuana, hashish, or skank; ecstasy; methamphetamines (crystal or speed); gamma hydroxybutyrate or gamma butyrolactone; poppers; other inhalants; mephedrone; hallucinogens (LSD, mushroom tea, others); or others.

[i] Sex for money or some other good (eg, gifts, housing).

[j] Answered only by those currently taking PrEP (n=448).

Table 3. Linear regression models demonstrating the moderating effect of current PrEP use on the association between standardized PRHS and HIRI-MSM scores (n=4092).[a]

Variable	Model 1		Model 2		Model 3		Full model	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Intercept	0 (–0.03 to 0.03)	0.99	.05 (0.02 to 0.08)	.001	.06 (0.02 to 0.09)	<.001	–.11 (–0.21 to –0.01)	.04
HIRI-MSM[b] score, standardized	.19 (0.16 to 0.22)	<.001	.23 (0.20 to 0.26)	<.001	.27 (0.24 to 0.30)	<.001	.26 (0.22 to 0.29)	<.001
Current PrEP[c] use	N/A[d]	N/A	–.52 (–0.62 to –0.41)	<.001	–.31 (–0.43 to –0.19)	<.001	–.36 (–0.48 to –0.23)	<.001
HIRI:PrEP interaction	N/A	N/A	N/A	N/A	–.30 (–0.39 to –0.21)	<.001	–.30 (–0.39 to –0.21)	<.001
Race								
Black[e]	N/A	N/A	N/A	N/A	N/A	N/A	.04 (–0.07 to 0.14)	.49
Pardo[e]	N/A	N/A	N/A	N/A	N/A	N/A	.02 (–0.05 to 0.09)	.61
Education								
University or higher[e]	N/A	N/A	N/A	N/A	N/A	N/A	.12 (0.06 to 0.19)	<.001
State								
Rio de Janeiro[e]	N/A	N/A	N/A	N/A	N/A	N/A	.03 (–0.05 to 0.10)	.48
Other[e]	N/A	N/A	N/A	N/A	N/A	N/A	–.01 (–0.08 to 0.06)	.73
Sexual orientation								
Gay[e]	N/A	N/A	N/A	N/A	N/A	N/A	.11 (0.03 to 0.19)	.007
HIV-KA[f] score, standardized	N/A	N/A	N/A	N/A	N/A	N/A	.12 (0.09 to 0.15)	<.001
Steady partner								
HIV-negative[e]	N/A	N/A	N/A	N/A	N/A	N/A	–.24 (–0.31 to –0.18)	<.001
Living with HIV or HIV-unknown[e]	N/A	N/A	N/A	N/A	N/A	N/A	.16 (0.05 to 0.28)	.004
Transactional sex[e]	N/A	N/A	N/A	N/A	N/A	N/A	.11 (–0.01 to 0.23)	.073
Last HIV test								
6 or more months ago[e]	N/A	N/A	N/A	N/A	N/A	N/A	.08 (0.01 to 0.15)	.016
Never[e]	N/A	N/A	N/A	N/A	N/A	N/A	–.06 (–0.17 to 0.04)	.24
Adjusted R^2	0.034	N/A	0.057	N/A	0.066	N/A	0.103	N/A

[a] Given small sample sizes, Asian (n=70), Indigenous (n=30), and respondents with missing data (n=152) were excluded.

[b] HIRI-MSM: HIV Incidence Risk Index-men who have sex with men.

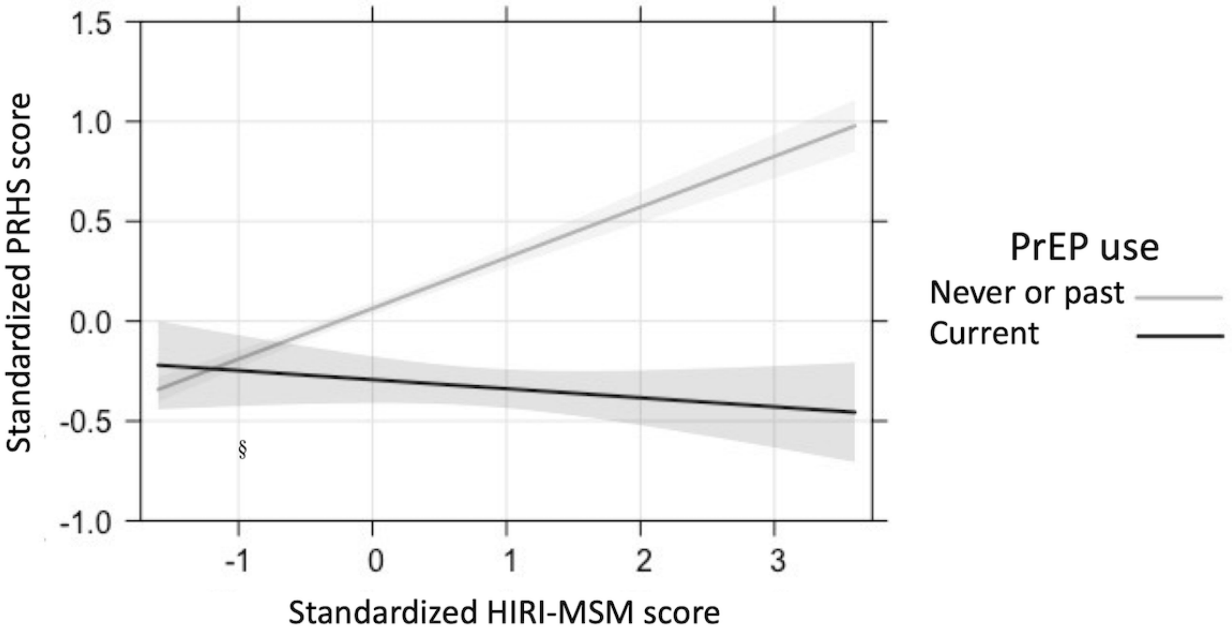
[c] PrEP: pre-exposure prophylaxis.

[d] N/A: not applicable.

[e] Reference categories are as follows: race: White; education: secondary or lower; Brazilian state: São Paulo; sexual orientation: other; steady partner: no steady partner; transactional sex: no; last HIV test: 6 months or less.

[f] HIV-KA: HIV/AIDS Knowledge Assessment.

Figure 2. Graphical representation of the joint effect of sexual risk behavior (measured via HIRI-MSM) and current PrEP use on perceived risk of HIV (measured via PRHS). The final model was adjusted for race, education, Brazilian state, sexual orientation, HIV knowledge, steady partner, transactional sex, and timing of last HIV test. HIRI-MSM: HIV Incidence Risk Index for men who have sex with men; PrEP: pre-exposure prophylaxis; PRHS: Perceived Risk of HIV Scale.



Discussion

Principal Findings

We found that current PrEP use among MSM in Brazil had a significant moderating effect on the association between sexual risk behavior and perceived risk of HIV. While there was an overall positive association between increasing HIRI-MSM and PRHS scores, the negative moderating effect of PrEP use resulted in no significant association between HIRI-MSM and PRHS scores among current PrEP users. Perceived risk of HIV was significantly positively associated with all sexual risk behavior variables among MSM not currently taking PrEP, but not among MSM currently taking PrEP.

One explanation for our findings is that Brazilian MSM who were taking PrEP were appropriately confident and optimistic about its effectiveness as an HIV prevention strategy. This understanding has been referred to as prevention optimism, which Holt and Murphy [54] define as “the belief that it is easier to avoid HIV infection or transmission because of PrEP and that it is more acceptable and safer to engage in condomless sex because the risk of HIV is reduced.” Prevention optimism could be a key mediator between sexual risk behavior and perceived risk of HIV and may be a useful adjunct to interpreting cross-sectional data [31,54]. While we did not specifically assess prevention optimism, future research could consider assessing agreement with statements such as “It is safe for me to have sex without condoms if I am using PrEP” [54]. There is a concern that the lower perceived risk of HIV among PrEP users may indirectly increase the risk of bacterial STIs via increased sexual risk behaviors, such as number of sexual partners and CAI [55], though data on the temporal relationship between PrEP use and changes in sexual behavior and STI acquisition are mixed [22,56,57]. Nevertheless, the lack of association between sexual risk behaviors and perceived risk of HIV among PrEP users in our study highlights the importance of including the prevention of bacterial STIs as a part of PrEP implementation policy. Other future considerations include the adoption of prophylaxis options for bacterial STIs to be taken by PrEP users [58].

Another important factor influencing the relationship between sexual risk behavior and perceived risk of HIV is the positive psychological impact of PrEP, which has been shown to be associated with reduced anxiety and fear related to sexual intercourse, and relatedly, increased sexual pleasure and intimacy [21-23,59-61]. PrEP use allows some MSM to engage in sexual behaviors that were previously the cause of anxiety and may facilitate connectedness within the MSM community [21,23]. Moreover, the ability of MSM taking PrEP to forgo condoms without worry about acquiring HIV can have a significant positive impact on sexual pleasure and satisfaction. In our study, we used a validated scale to capture PrEP’s impact on the multiple dimensions of the perceived risk of HIV, but future studies could more specifically consider the positive mental health impact of PrEP use among Brazilian MSM.

The frequent interaction with the health care system required of those taking PrEP may also influence an individual’s perceived risk of HIV and offer an opportunity for counseling interventions focused on reducing risk of bacterial STIs among PrEP users [4,22,56]. In Brazil, MSM receiving PrEP from the public health system must have appointments and HIV testing with their sexual health care provider every 3-6 months [62]. Regular HIV testing may contribute to lower perceived risk, since we found that those who had last tested for HIV within the previous 6 months had lower perceived risk of HIV. Additionally, the counseling provided at these visits could influence perceived risk by way of an improved understanding of how one's sexual behaviors relate to HIV risk and how PrEP may mitigate that risk [31]. Jaspal et al [4] proposed that perceived risk of HIV may mediate the pathway between HIV knowledge and sexual behavior. We previously found that PrEP use was associated with greater HIV knowledge among Brazilian MSM [40], and in this study we found greater HIV knowledge was associated with greater perceived risk of HIV among our overall sample. However, we did not explore the association among current PrEP users specifically, and future studies could compare the association between HIV knowledge and perceived risk of HIV stratified by PrEP use.

Like previous work from Brazil [47], we found higher education and reporting one's sexual orientation as gay to be associated with a higher perceived risk of HIV, whereas having a steady HIV-negative partner, compared to no steady partner, was associated with a lower perceived risk of HIV. Having tested for HIV more than 6 months ago, compared to within 6 months, was associated with a higher perceived risk of HIV among our sample, but findings from other studies have been mixed. MSM in the United States who had tested for HIV more than 1 year ago, compared to within the past year, were found to have lower perceived risk of HIV [27]. Another multicountry study found that as the years since the last HIV test increased, the perceived risk of HIV decreased among MSM in the United Kingdom but increased among MSM in Thailand [51].

Our use of the PRHS is unique among similar studies. Much of the prior research on perceived risk of HIV and PrEP use has measured perceived risk using agreement with a single Likert scale statement [6,18-20,27] or a single question with categorical responses of low, medium, and high risk [5,28-30]. The PRHS, in contrast, offers a more robust measure by assessing multiple dimensions of the perceived risk of HIV [32].

It is important to acknowledge that PrEP users are not a monolith, and the distribution of PRHS scores among PrEP users in our study demonstrated a range from low to high perceived risk of HIV. While we have demonstrated that PrEP use, on average, has a moderating effect on the relationship between sexual risk behavior and perceived risk of HIV, we did not study how this association may vary based on the type of PrEP user, and we did not have longitudinal data to assess how risk perception may change over time. Previous research has shown that PrEP users may fall into distinct groups depending on their perceived risk, sexual risk behaviors, and use of other prevention methods. A longitudinal study from France and Canada found that participants enrolled in the PrEP trial fell into distinct risk perception trajectories throughout the study follow-up, which they described as low-, medium-, and high perceived risk [18]. Similarly, a discrete choice experiment conducted among MSM in Singapore found that respondents fell into 3 different groups based on PrEP and condom preferences, and their perceived risk of HIV and STIs [63]. MSM PrEP-users from Australia fell into 4 distinct groupings based on a latent class analysis of risk behaviors and perceived risk of STIs, with highly variable views toward STI risk among the 4 groups [64]. Understanding the potential heterogeneity of PrEP users is critical in the implementation of future PrEP initiatives in Latin America and when considering interventions to reduce the risk of bacterial STIs among PrEP users.

Our study has several limitations. The cross-sectional nature of these data limited our ability to assert temporality between exposures of interest and perceived risk of HIV. These data were self-reported and may have been subject to recall or social desirability bias. We focused on cisgender MSM because of the small sample size of respondents reporting other gender identities; additional research is needed to characterize how PrEP use relates to sexual behavior and perceived risk of HIV among transgender, nonbinary, and gender-diverse individuals in Brazil. The majority of MSM in our sample were from either São Paulo or Rio de Janeiro, which may limit the generalizability of our findings, since previous research showed that perceived risk of HIV risk varies by region of Brazil [52]. Our findings are also specific to MSM who use Hornet and may not be generalizable to those who use other social networking app or those who do not have access. Additionally, the 10.3% of our sample who were taking PrEP at the time of this study may represent early adopters of PrEP and may not necessarily be representative of the population of MSM who will use PrEP as availability expands. The HIRI-MSM tool is a proxy measure of behavior risk of HIV, and other measures of sexual behavior may have different associations with perceived risk. Finally, the HIRI-MSM tool was developed for use in the United States and has not been specifically validated for use among Brazilian MSM, but we chose it because we preferred a continuous rather than dichotomous explanatory variable and because it has widespread familiarity among the HIV prevention research community and previous use among Brazilian MSM [52].

Conclusions

PrEP is highly effective at preventing HIV acquisition, and its use among MSM is associated with lower perceived risk of HIV compared to those not taking PrEP. We found that while higher HIRI-MSM scores were predictive of higher PRHS scores among our overall sample, the association was moderated by PrEP use, resulting in no significant association between sexual risk behavior and perceived risk of HIV among current PrEP users in Brazil. PrEP's objective efficacy, positive psychological impact, and the frequent HIV testing and interaction with the health care system required of PrEP users may jointly influence the relationship between sexual risk behavior and perceived risk of HIV, though additional research is needed to measure how each of these factors influences the moderating effect of PrEP. Future studies should explore the concept of prevention optimism and consider the temporal associations between PrEP use, sexual risk behaviors, and perceived HIV risk. Finally, the expansion of PrEP access in the Latin American region should consider how the lower perceived risk of HIV among PrEP users may necessitate targeted counseling on the risk of bacterial STIs as well as adjunct STI prevention modalities.

Acknowledgments

The abstract for this study was presented as a web-based poster presentation at the 24th International AIDS Conference, July 29, 2022, to August 2, 2022. This project was funded by the University of California Los Angeles' South American Program in HIV Prevention Research (SAPHIR) National Institute of Mental Health (NIMH) grant R25MH087222. This work was also supported by the UCLA Center for HIV Identification, Prevention, and Treatment Services (CHIPTS) NIMH grant P30MH58107. PML was supported by Programa Inova FIOCRUZ, Edital Geração do Conhecimento or VPPCB, National Council of Technological and Scientific Development (CNPq; #316401/2021-8), and Carlos Chagas Filho Foundation for Research Support in the State of Rio de Janeiro (FAPERJ; #E-26/201.133/2021). TST was supported by CNPq (#402916/2021-2 and #311871/2021-6) and FAPERJ (#E-26/211.577/2021 and #E-26/201.270/2022). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding institutions.

Data Availability

The data set analyzed during this study is available from the corresponding author on reasonable request.

Authors' Contributions

TST, BH, DRBB, VGV, BG, and PML were involved in study planning, survey design, survey administration, and data collection. KJB, TST, JC, and PML organized, analyzed, and interpreted the data. All authors reviewed and revised the draft paper and provided approval for submission of the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Linear regression models demonstrating the moderating effect of current PrEP use on the association between standardized PRHS and HIRI-MSM scores, excluding past PrEP users N=3946.

[DOCX File , 31 KB - publichealth_v9i1e45134_app1.docx]

References

1. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
2. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, iPrEx Study Team. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. Sci Transl Med 2012;4(151):151ra125 [FREE Full text] [doi: 10.1126/scitranslmed.3004006] [Medline: 22972843]
3. Torres TS, Luz PM, De Boni RB, de Vasconcellos MTL, Hoagland B, Garner A, et al. Factors associated with PrEP awareness according to age and willingness to use HIV prevention technologies: the 2017 online survey among MSM in Brazil. AIDS Care 2019;31(10):1193-1202. [doi: 10.1080/09540121.2019.1619665] [Medline: 31122033]
4. Jaspal R, Lopes B, Bayley J, Papaloukas P. A structural equation model to predict pre-exposure prophylaxis acceptability in men who have sex with men in Leicester, UK. HIV Med 2019;20(1):11-18 [FREE Full text] [doi: 10.1111/hiv.12667] [Medline: 30160367]
5. Herder T, Agardh A, Björkman P, Månsson F. Interest in taking HIV pre-exposure prophylaxis is associated with behavioral risk indicators and self-perceived HIV risk among men who have sex with men attending HIV testing venues in Sweden. Arch Sex Behav 2020;49(6):2165-2177 [FREE Full text] [doi: 10.1007/s10508-020-01740-9] [Medline: 32495243]
6. Chakrapani V, Newman PA, Shunmugam M, Rawat S, Baruah D, Nelson R, et al. PrEP eligibility, HIV risk perception, and willingness to use PrEP among high-risk men who have sex with men in India: a cross-sectional survey. AIDS Care 2022;34(3):301-309 [FREE Full text] [doi: 10.1080/09540121.2021.1887801] [Medline: 33615903]
7. Sun Z, Gu Q, Dai Y, Zou H, Agins B, Chen Q, et al. Increasing awareness of HIV pre-exposure prophylaxis (PrEP) and willingness to use HIV PrEP among men who have sex with men: a systematic review and meta-analysis of global data. J Int AIDS Soc 2022;25(3):e25883 [FREE Full text] [doi: 10.1002/jia2.25883] [Medline: 35255193]
8. Torres TS, Konda KA, Vega-Ramirez EH, Elorreaga OA, Diaz-Sosa D, Hoagland B, ImPrEP Study Group. Factors associated with willingness to use pre-exposure prophylaxis in Brazil, Mexico, and Peru: web-based survey among men who have sex with men. JMIR Public Health Surveill 2019;5(2):e13771 [FREE Full text] [doi: 10.2196/13771] [Medline: 31210143]
9. Blumenthal J, Jain S, Mulvihill E, Sun S, Hanashiro M, Ellorin E, et al. Perceived versus calculated HIV risk: implications for pre-exposure prophylaxis uptake in a randomized trial of men who have sex with men. J Acquir Immune Defic Syndr 2019;80(2):e23-e29 [FREE Full text] [doi: 10.1097/QAI.0000000000001888] [Medline: 30422909]
10. Cox J, Apelian H, Moodie EEM, Messier-Peet M, Hart TA, Grace D, Engage Study Team. Use of HIV pre-exposure prophylaxis among urban Canadian gay, bisexual and other men who have sex with men: a cross-sectional analysis of the

Engage cohort study. CMAJ Open 2021;9(2):E529-E538 [FREE Full text] [doi: 10.9778/cmajo.20200198] [Medline: 34021010]

11. Ramautarsing RA, Meksena R, Sungsing T, Chinbunchorn T, Sangprasert T, Fungfoosri O, et al. Evaluation of a pre-exposure prophylaxis programme for men who have sex with men and transgender women in Thailand: learning through the HIV prevention cascade lens. J Int AIDS Soc 2020;23(Suppl 3):e25540 [FREE Full text] [doi: 10.1002/jia2.25540] [Medline: 32602660]
12. Hoagland B, Moreira RI, De Boni RB, Kallas EG, Madruga JV, Vasconcelos R, PrEP Brasil Study Team. High pre-exposure prophylaxis uptake and early adherence among men who have sex with men and transgender women at risk for HIV infection: the PrEP Brasil demonstration project. J Int AIDS Soc 2017;20(1):21472 [FREE Full text] [doi: 10.7448/IAS.20.1.21472] [Medline: 28418232]
13. Bavinton BR, Grulich AE. HIV pre-exposure prophylaxis: scaling up for impact now and in the future. Lancet Public Health 2021;6(7):e528-e533 [FREE Full text] [doi: 10.1016/S2468-2667(21)00112-2] [Medline: 34087117]
14. Gombe MM, Cakouros BE, Ncube G, Zwangobani N, Mareke P, Mkwamba A, et al. Key barriers and enablers associated with uptake and continuation of oral pre-exposure prophylaxis (PrEP) in the public sector in Zimbabwe: qualitative perspectives of general population clients at high risk for HIV. PLoS One 2020;15(1):e0227632 [FREE Full text] [doi: 10.1371/journal.pone.0227632] [Medline: 31931514]
15. Sullivan PS, Siegler AJ. What will it take to meet UNAIDS targets for preexposure prophylaxis users? Curr Opin Infect Dis 2022;35(1):1-8. [doi: 10.1097/QCO.0000000000000809] [Medline: 34879049]
16. Whitfield THF, John SA, Rendina HJ, Grov C, Parsons JT. Why I quit pre-exposure prophylaxis (PrEP)? A mixed-method study exploring reasons for PrEP discontinuation and potential re-initiation among gay and bisexual men. AIDS Behav 2018;22(11):3566-3575 [FREE Full text] [doi: 10.1007/s10461-018-2045-1] [Medline: 29404756]
17. Powell VE, Gibas KM, DuBow J, Krakower DS. Update on HIV preexposure prophylaxis: effectiveness, drug resistance, and risk compensation. Curr Infect Dis Rep 2019;21(8):28 [FREE Full text] [doi: 10.1007/s11908-019-0685-6] [Medline: 31227999]
18. Di Ciaccio M, Sagaon-Teyssier L, Protière C, Mimi M, Suzan-Monti M, Meyer L, et al. Impact of HIV risk perception on both pre-exposure prophylaxis and condom use. J Health Psychol 2021;26(10):1575-1586. [doi: 10.1177/1359105319883927] [Medline: 31647330]
19. Basten M, den Daas C, Heijne JCM, Boyd A, Davidovich U, Rozhnova G, et al. The rhythm of risk: sexual behaviour, PrEP use and HIV risk perception between 1999 and 2018 among men who have sex with men in Amsterdam, The Netherlands. AIDS Behav 2021;25(6):1800-1809 [FREE Full text] [doi: 10.1007/s10461-020-03109-4] [Medline: 33269426]
20. Zimmermann HML, van Bilsen WPH, Boyd A, Prins M, van Harreveld F, Davidovich U, HIV Transmission Elimination Team Amsterdam. Prevention challenges with current perceptions of HIV burden among HIV-negative and never-tested men who have sex with men in the Netherlands: a mixed-methods study. J Int AIDS Soc 2021;24(8):e25715 [FREE Full text] [doi: 10.1002/jia2.25715] [Medline: 34449130]
21. Whitfield THF, Jones SS, Wachman M, Grov C, Parsons JT, Rendina HJ. The impact of pre-exposure prophylaxis (PrEP) use on sexual anxiety, satisfaction, and esteem among gay and bisexual men. J Sex Res 2019;56(9):1128-1135 [FREE Full text] [doi: 10.1080/00224499.2019.1572064] [Medline: 30777781]
22. Grov C, Westmoreland DA, D'Angelo AB, Pantalone DW. How has HIV pre-exposure prophylaxis (PrEP) changed sex? A review of research in a new era of bio-behavioral HIV prevention. J Sex Res 2021;58(7):891-913 [FREE Full text] [doi: 10.1080/00224499.2021.1936440] [Medline: 34180743]
23. Keen P, Hammoud MA, Bourne A, Bavinton BR, Holt M, Vaccher S, et al. Use of HIV pre-exposure prophylaxis (PrEP) associated with lower HIV anxiety among gay and bisexual men in Australia who are at high risk of HIV infection: results from the Flux study. J Acquir Immune Defic Syndr 2020;83(2):119-125 [FREE Full text] [doi: 10.1097/QAI.0000000000002232] [Medline: 31935203]
24. Traeger MW, Schroeder SE, Wright EJ, Hellard ME, Cornelisse VJ, Doyle JS, et al. Effects of pre-exposure prophylaxis for the prevention of human immunodeficiency virus infection on sexual risk behavior in men who have sex with men: a systematic review and meta-analysis. Clin Infect Dis 2018;67(5):676-686 [FREE Full text] [doi: 10.1093/cid/ciy182] [Medline: 29509889]
25. Storholm ED, Volk JE, Marcus JL, Silverberg MJ, Satre DD. Risk perception, sexual behaviors, and PrEP adherence among substance-using men who have sex with men: a qualitative study. Prev Sci 2017;18(6):737-747 [FREE Full text] [doi: 10.1007/s11121-017-0799-8] [Medline: 28578516]
26. Villela LM, Veloso VG, Hoagland B, Fernandes NM, Kallas EG, Madruga JV, et al. Possible risk compensation, attitudes, and beliefs among Brazilian individuals potentially eligible for pre-exposure prophylaxis. Int J STD AIDS 2021;32(9):806-815. [doi: 10.1177/0956462421992157] [Medline: 33615909]
27. Kahle EM, Sharma A, Sullivan SP, Stephenson R. HIV prioritization and risk perception among an online sample of men who have sex with men in the United States. Am J Mens Health 2018;12(4):676-687 [FREE Full text] [doi: 10.1177/1557988318776581] [Medline: 29781331]

28. Khawcharoenporn T, Mongkolkaewsub S, Naijitra C, Khonphiern W, Apisarnthanarak A, Phanuphak N. HIV risk, risk perception and uptake of HIV testing and counseling among youth men who have sex with men attending a gay sauna. AIDS Res Ther 2019;16(1):13 [FREE Full text] [doi: 10.1186/s12981-019-0229-z] [Medline: 31189481]
29. Carroll JJ, Heffron R, Mugo N, Ngure K, Ndase P, Asiimwe S, et al. Perceived risk among human immunodeficiency virus serodiscordant couples in East Africa taking oral pre-exposure prophylaxis. Sex Transm Dis 2016;43(8):471-475 [FREE Full text] [doi: 10.1097/OLQ.0000000000000472] [Medline: 27419812]
30. Seekaew P, Pengnonyang S, Jantarapakde J, Meksena R, Sungsing T, Lujintanon S, et al. Discordance between self-perceived and actual risk of HIV infection among men who have sex with men and transgender women in Thailand: a cross-sectional assessment. J Int AIDS Soc 2019;22(12):e25430 [FREE Full text] [doi: 10.1002/jia2.25430] [Medline: 31855324]
31. Evangeli M, Baker LLE, Pady K, Jones B, Wroe AL. What leads some people to think they are HIV-positive before knowing their diagnosis? A systematic review of psychological and behavioural correlates of HIV-risk perception. AIDS Care 2016;28(8):943-953. [doi: 10.1080/09540121.2016.1139040] [Medline: 26852667]
32. Napper LE, Fisher DG, Reynolds GL. Development of the perceived risk of HIV scale. AIDS Behav 2012;16(4):1075-1083 [FREE Full text] [doi: 10.1007/s10461-011-0003-2] [Medline: 21785873]
33. Martins A, Chaves C, Canavarro MC, Pereira M. Psychometric properties of the European Portuguese version of the perceived risk of HIV scale in the general population and HIV-uninfected partners from sero-different couples. BMC Public Health 2019;19(1):1336 [FREE Full text] [doi: 10.1186/s12889-019-7696-y] [Medline: 31640612]
34. Torres TS, Luz PM, Marins LMS, Bezerra DRB, Almeida-Brasil CC, Veloso VG, et al. Cross-cultural adaptation of the perceived risk of HIV scale in Brazilian Portuguese. Health Qual Life Outcomes 2021;19(1):117 [FREE Full text] [doi: 10.1186/s12955-021-01760-6] [Medline: 33836775]
35. The world factbook: Brazil. Central Intelligence Agency. 2020. URL: https://www.cia.gov/the-world-factbook/countries/brazil/ [accessed 2022-04-06]
36. HIV policy lab. Georgetown University. URL: https://oneill.law.georgetown.edu/projects/hiv-policy-lab/ [accessed 2022-04-06]
37. Painel PrEP. Departamento de Doenças de Condições Crônicas e Infecções Sexualmente Transmissíveis: Ministério da Saúde. URL: http://antigo.aids.gov.br/pt-br/painel-prep [accessed 2023-07-11]
38. Torres TS, Hoagland B, Bezerra DRB, Garner A, Jalil EM, Coelho LE, et al. Impact of COVID-19 pandemic on sexual minority populations in Brazil: an analysis of social/racial disparities in maintaining social distancing and a description of sexual behavior. AIDS Behav 2021;25(1):73-84 [FREE Full text] [doi: 10.1007/s10461-020-02984-1] [Medline: 32737817]
39. Give every customer a voice. Make every voice matter. Alchemer. URL: https://www.alchemer.com/ [accessed 2019-12-01]
40. Blair KJ, Torres TS, Hoagland B, Bezerra DRB, Veloso VG, Grinsztejn B, et al. Pre-exposure prophylaxis use, HIV knowledge, and internalized homonegativity among men who have sex with men in Brazil: a cross-sectional study. Lancet Reg Health Am 2022;6:100152 [FREE Full text] [doi: 10.1016/j.lana.2021.100152] [Medline: 35291206]
41. Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods 2012;17(3):437-455. [doi: 10.1037/a0028085] [Medline: 22506584]
42. Ferreira RC, Torres TS, Ceccato MDGB, Bezerra DRB, Thombs BD, Luz PM, et al. Development and evaluation of short-form measures of the HIV/AIDS knowledge assessment tool among sexual and gender minorities in Brazil: cross-sectional study. JMIR Public Health Surveill 2022;8(3):e30676 [FREE Full text] [doi: 10.2196/30676] [Medline: 35348470]
43. Ferreira RC, Torres TS, Marins LMS, Ceccato MDGB, Bezerra DRB, Luz PM. HIV knowledge and its correlation with the Undetectable = Untransmittable slogan in Brazil. Rev Saude Publica 2022;56:87 [FREE Full text] [doi: 10.11606/s1518-8787.2022056004168] [Medline: 36228232]
44. Smith DK, Pals SL, Herbst JH, Shinde S, Carey JW. Development of a clinical screening index predictive of incident HIV infection among men who have sex with men in the United States. J Acquir Immune Defic Syndr 2012;60(4):421-427 [FREE Full text] [doi: 10.1097/QAI.0b013e318256b2f6] [Medline: 22487585]
45. Preexposure prophylaxis for the prevention of HIV infection in the United States – 2017 update: a clinical practice guideline. US Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2017.pdf [accessed 2023-09-07]
46. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2018;48(6):1273-1296 [FREE Full text] [doi: 10.1007/s11165-016-9602-2]
47. Luz PM, Torres TS, Almeida-Brasil CC, Marins LMS, Veloso VG, Grinsztejn B, et al. High-risk sexual behavior, binge drinking and use of stimulants are key experiences on the pathway to high perceived HIV risk among men who have sex with men in Brazil. AIDS Behav 2021;25(3):748-757. [doi: 10.1007/s10461-020-03035-5] [Medline: 32940826]
48. Kesler MA, Kaul R, Liu J, Loutfy M, Gesink D, Myers T, et al. Actual sexual risk and perceived risk of HIV acquisition among HIV-negative men who have sex with men in Toronto, Canada. BMC Public Health 2016;16:254 [FREE Full text] [doi: 10.1186/s12889-016-2859-6] [Medline: 26969463]
49. Sharma A, Kahle EM, Sullivan SP, Stephenson R. Birth cohort variations across functional knowledge of HIV prevention strategies, perceived risk, and HIV-associated behaviors among gay, bisexual, and other men who have sex with men in

the United States. Am J Mens Health 2018;12(6):1824-1834 [FREE Full text] [doi: 10.1177/1557988318790875] [Medline: 30051744]
50. Gallagher T, Link L, Ramos M, Bottger E, Aberg J, Daskalakis D. Self-perception of HIV risk and candidacy for pre-exposure prophylaxis among men who have sex with men testing for HIV at commercial sex venues in New York City. LGBT Health 2014;1(3):218-224. [doi: 10.1089/lgbt.2013.0046] [Medline: 26789715]
51. Chard AN, Metheny N, Stephenson R. Perceptions of HIV seriousness, risk, and threat among online samples of HIV-negative men who have sex with men in seven countries. JMIR Public Health Surveill 2017;3(2):e37 [FREE Full text] [doi: 10.2196/publichealth.7546] [Medline: 28634155]
52. Torres TS, Marins LMS, Veloso VG, Grinsztejn B, Luz PM. How heterogeneous are MSM from Brazilian cities? An analysis of sexual behavior and perceived risk and a description of trends in awareness and willingness to use pre-exposure prophylaxis. BMC Infect Dis 2019;19(1):1067 [FREE Full text] [doi: 10.1186/s12879-019-4704-x] [Medline: 31856746]
53. Fox J, Weisberg S. An R Companion to Applied Regression, 3rd Edition. Thousand Oaks, CA: SAGE Publications, Inc; 2018.
54. Holt M, Murphy DA. Individual versus community-level risk compensation following preexposure prophylaxis of HIV. Am J Public Health 2017;107(10):1568-1571 [FREE Full text] [doi: 10.2105/AJPH.2017.303930] [Medline: 28817332]
55. Hart TA, Noor SW, Berlin GW, Skakoon-Sparling S, Tavangar F, Tan D, Engage Study. Pre-exposure prophylaxis and bacterial sexually transmitted infections (STIs) among gay and bisexual men. Sex Transm Infect 2023;99(3):167-172 [FREE Full text] [doi: 10.1136/sextrans-2021-055381] [Medline: 35701145]
56. Rojas Castro D, Delabre RM, Molina JM. Give PrEP a chance: moving on from the "risk compensation" concept. J Int AIDS Soc 2019;22(Suppl 6):e25351 [FREE Full text] [doi: 10.1002/jia2.25351] [Medline: 31468693]
57. Veloso VG, Cáceres CF, Hoagland B, Moreira RI, Vega-Ramírez H, Konda KA, ImPrEP Study Group. Same-day initiation of oral pre-exposure prophylaxis among gay, bisexual, and other cisgender men who have sex with men and transgender women in Brazil, Mexico, and Peru (ImPrEP): a prospective, single-arm, open-label, multicentre implementation study. Lancet HIV 2023;10(2):e84-e96 [FREE Full text] [doi: 10.1016/S2352-3018(22)00331-9] [Medline: 36565708]
58. Luetkemeyer AF, Donnell D, Dombrowski JC, Cohen S, Grabow C, Brown CE, DoxyPEP Study Team. Postexposure Doxycycline to Prevent Bacterial Sexually Transmitted Infections. N Engl J Med 2023 Apr 06;388(14):1296-1306. [doi: 10.1056/NEJMoa2211934] [Medline: 37018493]
59. Philpot S, Prestage G, Holt M, Haire B, Maher L, Hammoud M, et al. Gay and bisexual men's perceptions of pre-exposure prophylaxis (PrEP) in a context of high accessibility: an Australian qualitative study. AIDS Behav 2020;24(8):2369-2380. [doi: 10.1007/s10461-020-02796-3] [Medline: 31970582]
60. Mabire X, Puppo C, Morel S, Mora M, Rojas Castro D, Chas J, et al. Pleasure and PrEP: pleasure-seeking plays a role in prevention choices and could lead to PrEP initiation. Am J Mens Health 2019;13(1):1557988319827396 [FREE Full text] [doi: 10.1177/1557988319827396] [Medline: 30819060]
61. Gagliano N, Black S, Mniszak C, Morgan J, Lachowsky N, Knight R. Exchanging a "gnawing thought for "taking a pill once a day": examining sexual minority men's experiences with PrEP and mental health in a universal PrEP access context. AIDS Educ Prev 2021;33(5):411-423. [doi: 10.1521/aeap.2021.33.5.411] [Medline: 34596431]
62. Hoagland B, Torres TS, Bezerra DRB, Geraldo K, Pimenta C, Veloso VG, et al. Telemedicine as a tool for PrEP delivery during the COVID-19 pandemic in a large HIV prevention service in Rio de Janeiro-Brazil. Braz J Infect Dis 2020;24(4):360-364 [FREE Full text] [doi: 10.1016/j.bjid.2020.05.004] [Medline: 32504552]
63. Tan RKJ, Wang Y, Prem K, Harrison-Quintana J, Teo AKJ, Kaur N, et al. HIV pre-exposure prophylaxis, condoms, or both? Insights on risk compensation through a discrete choice experiment and latent class analysis among men who have sex with men. Value Health 2021;24(5):714-723 [FREE Full text] [doi: 10.1016/j.jval.2020.11.023] [Medline: 33933241]
64. Traeger MW, Murphy D, Ryan KE, Asselin J, Cornelisse VJ, Wilkinson AL, et al. Latent class analysis of sexual behaviours and attitudes to sexually transmitted infections among gay and bisexual men using PrEP. AIDS Behav 2022;26(6):1808-1820. [doi: 10.1007/s10461-021-03529-w] [Medline: 34782934]

Abbreviations

CAI: condomless anal intercourse
HIRI-MSM: HIV Incidence Risk Index for men who have sex with men
HIV-KA: HIV/AIDS Knowledge Assessment
MSM: men who have sex with men
MW: minimum wage
NIMH: National Institute of Mental Health
PrEP: pre-exposure prophylaxis
PRHS: Perceived Risk of HIV Scale
STI: sexually transmitted infection

Edited by A Mavragani, T Sanchez; submitted 16.12.22; peer-reviewed by A Algarin, S Chen; comments to author 06.06.23; revised version received 14.07.23; accepted 31.07.23; published 05.10.23.

Please cite as:
Blair KJ, Torres TS, Hoagland B, Bezerra DRB, Veloso VG, Grinsztejn B, Clark J, Luz PM
Moderating Effect of Pre-Exposure Prophylaxis Use on the Association Between Sexual Risk Behavior and Perceived Risk of HIV Among Brazilian Gay, Bisexual, and Other Men Who Have Sex With Men: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e45134
URL: https://publichealth.jmir.org/2023/1/e45134
doi:10.2196/45134
PMID:37796573

Original Paper

Pretreatment HIV Drug Resistance and the Molecular Transmission Network Among HIV-Positive Individuals in China in 2022: Multicenter Observational Study

Hongli Chen[1,2*], MPH; Jingjing Hao[1*], MS; Jing Hu[1], BSc; Chang Song[1], BSc; Yesheng Zhou[1], MPH; Miaomiao Li[1], MS; Jin Chen[1], MS; Xiu Liu[1], BSc; Dong Wang[1], BSc; Xiaoshan Xu[1], PhD; Peixin Xin[1], BSc; Jiaxin Zhang[1], MS; Lingjie Liao[1], PhD; Yi Feng[1], BSc; Dan Li[1], PhD; Stephen W Pan[3], PhD; Yiming Shao[1], PhD; Yuhua Ruan[1], PhD; Hui Xing[1], MS

[1] National Center for AIDS/STD Control and Prevention (NCAIDS), Chinese Center for Disease Control and Prevention (China CDC), Beijing, China
[2] Sichuan Nursing Vocational College, Chengdu, China
[3] Department of Public Health, The University of Texas at San Antonio, San Antonio, TX, United States
[*] these authors contributed equally

Corresponding Author:
Yuhua Ruan, PhD
National Center for AIDS/STD Control and Prevention (NCAIDS)
Chinese Center for Disease Control and Prevention (China CDC)
155 Changbai Road, Changping District
Beijing, 102206
China
Phone: 86 86 10 58900976
Email: ruanyuhua92@163.com

Abstract

Background: Emerging HIV drug resistance caused by increased usage of antiretroviral drugs (ARV) could jeopardize the success of standardized HIV management protocols in resource-limited settings.

Objective: We aimed to characterize pretreatment HIV drug resistance (PDR) among HIV-positive individuals and risk factors in China in 2022.

Methods: This cross-sectional study was conducted using 2-stage systematic sampling according to the World Health Organization’s surveillance guidelines in 8 provincial-level administrative divisions in 2022. Demographic information and plasma samples were obtained from study participants. PDR was analyzed using the Stanford HIV drug resistance database, and the Tamura-Nei 93 model in HIV-TRACE was used to calculate pairwise matches with a genetic distance of 0.01 substitutions per site. Logistic regression was used to identify and estimate factors associated with PDR.

Results: PDR testing was conducted on 2568 participants in 2022. Of the participants, 34.8% (n=893) were aged 30-49 years, 81.4% (n=2091) were male, and 3.2% (n=81) had prior ARV exposure. The prevalence of PDR to protease and reverse transcriptase regions, nonnucleoside reverse transcriptase inhibitors, nucleoside reverse transcriptase inhibitors, and protease inhibitors were 7.4% (n=190), 6.3% (n=163), 1.2% (n=32), and 0.2% (n=5), respectively. Yunnan, Jilin, and Zhejiang had much higher PDR incidence than did Sichuan. The prevalence of nonnucleoside reverse transcriptase inhibitor–related drug resistance was 6.1% (n=157) for efavirenz and 6.3% (n=163) for nevirapine. Multivariable logistic regression models indicated that participants who had prior ARV exposure (odds ratio [OR] 7.45, 95% CI 4.50-12.34) and the CRF55_01B HIV subtype (OR 2.61, 95% CI 1.41-4.83) were significantly associated with PDR. Among 618 (24.2%) sequences (nodes) associated with 253 molecular transmission clusters (size range 2-13), drug resistance mutation sites included K103, E138, V179, P225, V106, V108, L210, T215, P225, K238, and A98.

Conclusions: The overall prevalence of PDR in China in 2022 was modest. Targeted genotypic PDR testing and medication compliance interventions must be urgently expanded to address PDR among newly diagnosed people living with HIV in China.

(JMIR Public Health Surveill 2023;9:e50894) doi:10.2196/50894

KEYWORDS

HIV; human immunodeficiency virus; mutation; pretreatment drug resistance; risk factors; molecular transmission network

Introduction

With the implementation of the “Treat all” program in 2016 in China, all people living with HIV are eligible to be treated, with their consent, regardless of their CD4 cell count and viral load level. Early treatment has been shown to effectively prolong the life span of people living with HIV and prevent secondary transmission. However, treatment of all individuals diagnosed with HIV has the potential to increase the prevalence of HIV drug resistance (DR) [1-4]. Pretreatment HIV drug resistance (PDR) refers to resistance that is detected among antiretroviral drug (ARV)–naive people initiating antiretroviral therapy (ART) or people with prior ARV drug exposure [5]. PDR may compromise the effectiveness of ART, increase mortality, and lead to secondary transmission of HIV [6]. Thus, assessing PDR levels is important for informing national policies on ARV and for improving ART programs and service delivery.

In China, individual ART plans are determined primarily by national standards for first- and second-line regimens, and potential DR is not a major consideration among individuals preparing to initiate ART. Due partly to resource limitations, individual HIV genotype resistance testing is not routinely conducted prior to the initiation of ART in China. This approach to developing individual ART regimens may be associated with DR [7]. A national meta-analysis indicated that PDR rapidly rose from 4.05% in 2011 to 5% in 2017. Specifically, resistance to nonnucleoside reverse transcriptase inhibitors (NNRTIs) increased from 2.15% to 3.81% from 2012 to 2017 [8]. However, it is important to note that PDR testing and prevalence varies widely by geographic region. Other surveys conducted on PDR in certain provinces of China showed that the prevalence of PDR was 10.5% in Chongqing, 9.1% in Shenyang, 4.1% in Beijing, 6% in Shenzhen, and 10.8% in Pu’er County of Yunnan [1,9-12]. The rates of PDR among HIV-positive individuals in China are comparatively lower than those in other countries, such as Cameroon (10.4%), Namibia (12.7%), Israel (12.1%), Mozambique (14.4%), Mexico (14.8%), the United States (22.5%), and Japan (12.5%) [13-19]. Yet China still lacks PDR surveillance methods that meet World Health Organization (WHO) standards. Assessing PDR prevalence with WHO standardized protocols will facilitate more accurate estimates of PDR prevalence in China, which will in turn better inform evidence-based ART policies in China.

Molecular transmission network analysis can be used to identify the maximum number of clusters and relatively active clusters based on genetic distance thresholds. Such methods have been used to evaluate the efficacy of ART on preventing secondary HIV transmission [20]. Moreover, our understanding of DR has been greatly enhanced by studies integrating analysis of transmission characteristics, DR strains, and molecular transmission networks [10,21].

The WHO has identified DR strategic surveillance of PDR as a very high priority in countries using efavirenz (EFV)/nevirapine (NVP) for first-line ART and recommends implementation or updates every 3 years [22]. Based on WHO PDR surveillance guidelines, the aim of this novel multicenter observational study is to investigate the prevalence of PDR and drug resistance mutations (DRMs) to partial pol regions, including protease (PR) and reverse transcriptase regions (RT), in diverse provinces in China.

Methods

Study Design and Study Population

We selected 8 provincial-level administrative divisions (PLADs) in China as the study area to conduct a large multicenter observational cross-sectional study. Two-stage systematic sampling in each PLAD was conducted according to the WHO PDR surveillance guidelines [23]. Provinces were stratified into the following 3 categories based on their HIV incidence rates: high (>2.00 cases per 10,000 individuals; included Sichuan Province, Chongqing Municipality, and Yunnan Province), moderate (0.52-1.99 cases per 10,000 individuals; included Jiangsu Province and Zhejiang Province), and low (<0.52 cases per 10,000 individuals; included Jilin Province, Hebei Province, and Hubei Province).

Guided by WHO PDR surveillance guidelines, we calculated the minimum necessary sample size by assuming the prevalence of PDR was 10% and the CI range was ±5%. We sampled 11-20 clinics in each PLAD according to the number of HIV-positive individuals starting or restarting ART in the past 12 months. The minimum sample size in each PLAD was 158-179 samples. To ensure a minimum necessary sample size, we assumed that 40% of collected specimens would be lost to factors such as laboratory failure and a design effect. Expected and actual sample sizes for each province are shown in Multimedia Appendix 1. Recruitment and specimen collection of HIV-positive individuals occurred in 2022 as individuals initiated ART.

The study eligibility criteria for individuals included the following: (1) tested HIV-positive, (2) ≥18 years old, (3) started or restarted first-line ART between January and June 2022 (includes women exposed to ARV drugs for preventing mother-to-child transmission of HIV, people who have received pre-exposure prophylaxis, and individuals reinitiating first-line ART), and (4) provided informed consent.

Data Collection

Each study participant was assigned a confidential and unique patient identifier used for linking basic epidemiological information and specimens. Basic sociodemographic data were collected, including age, sex, ethnicity, education level, marital status, occupation, route of infection, prior ARV exposure, time elapsed since HIV diagnosis, and PLADs. Clinical information included patients’ CD4 cell count before ART and HIV subtype.

Laboratory Tests

Viral RNA was extracted from 200 µL plasma samples using the QIAamp viral RNA mini kit (Qiagen, Hilton, Germany) according to the operational instructions. Extracted RNA was used for first-round polymerase chain reaction, and the products were used for second-round polymerase chain reaction amplification; both rounds were performed in 25 μL volume reactions to amplify the HIV pol gene region (HXB2: 2253-3553 nt). This covered the PR region (4-99 amino acids) and the partial RT region (1-251 amino acids) [24]. The amplified products were sequenced by Sanger sequencing.

Subtype and Drug Resistance Analysis

Sequencher (version 4.10.1, Gene Codes Corp) software was used to recombine and edit the original sequence fragments and nucleosides. BioEdit (version 7.0.9, Informer Technologies) software was used to realign the HXB2 reference sequence. The sequences were realigned, and some common reference sequences (including A to K) were used to identify the subtype by constructing neighbor joining phylogenetic trees in MEGA (version 7.0.26, MEGA Software). Subtypes with a bootstrap value higher than 70% were determined using HXB2 as the reference sequence. Sequences were realigned using the HIV database website [25]. Realigned sequences were uploaded to the Stanford Drug Resistance Database [26] to identify the degree of resistance of the 12 listed drugs that the WHO recommends for monitoring, namely, EFV, NVP, abacavir, zidovudine, lamivudine, tenofovir, emtricitabine, stavudine, didanosine, lopinavir/r, atazanavir/r, and darunavir/r. Corresponding DR mutant genes with levels of 1, 2, 3, 4, and 5 were classified as sensitive, potential resistance, low resistance, moderate resistance, and high resistance, respectively. Only mutations characterized as low, moderate, or high resistance were considered resistant. All possible sample mutual contamination, laboratory contamination, and other sequence quality controls were monitored by constructing neighbor joining phylogenetic trees and using the WHO HIVDR QC tool [27].

Construction of the Molecular Network

Realigned sequences were conducted to calculate pairwise genetic distances though the Tamura-Nei 93 model in the transmission cluster engine, HIV-TRACE [28]. Sequences no longer than 1000 bp or containing ≥5% ambiguities were excluded. A threshold genetic distance of 0.01 substitutions per site was used to observe transmission events that may occur within 5-6 years [29]. Each individual in the molecular network was represented by a node, and each node in the molecular transmission network represented HIV-positive individuals who were matched with epidemiological information. We then linked nodes to each other if their pairwise genetic distance was up to 0.01 substitutions per site. Web pages were used to visualize molecular networks, which were based on technical guides for HIV transmission network monitoring and intervention [30]. The results and subsequent molecular network diagram are accessible on the web [31]. Our previous studies used other software, such as HYPHY to calculate the pairwise gene distance and Cytoscape for visualization [20,32,33]. Clusters were defined as containing 2 or more nodes with HIV-positive individuals in the same cluster having potential for transmission.

Statistical Analysis

All quantitative data analyses were performed using the Statistical Analysis System version 9.4 (SAS Institute). Two-sided *P* values <.05 were considered statistically significant. All variables were assessed using univariate binary logistic regression. Variables significantly associated (*P*<.05) with the outcome were entered into a stepwise regression model to remove collinear variables. The remaining variables were entered into a multivariate binary logistic regression model to identify correlates of PDR. The *P* value of the final multivariate logistic regression model was <.05.

Ethical Considerations

This study was approved by the ethics committee of the National Center for AIDS/STD Control and Prevention, China Centers for Disease Control and Prevention (approval number X140617334).

Results

Characteristics of the Study Participants

A total of 2869 HIV-positive individuals were initially recruited into the study. After removal of 16 HIV-positive individuals who were younger than 18 years and 9 individuals who had not started ART, the penultimate sample size was 2844. PDR test results were available for 90.3% (n=2568) of participants in RT/PR regions, thus yielding a final sample size of 2568. Of the 2568 study participants with valid PDR results, 34.8% (n=893) were aged 30-49 years, 81.4% (n=2091) were male, 86.8% (n=2229) were of Han ethnicity, 30.5% (n=783) had a junior high school education, 37% (n=950) were single, 36% (n=924) were farmers, and 57.6% (n=1480) and 37% (n=951) were infected through heterosexual and homosexual intercourse, respectively. In addition, 3.2% (n=81) of participants had prior ARV drug exposure, 84.5% (n=2170) were diagnosed in 2022, and 37.5% (n=962) had CD4 cell counts of 0-199 cells/mm^3 prior to ART initiation (Table 1).

Table 1. Characteristics of HIV-positive individuals in 2022 in China.

Characteristic	Participants (n=2568), n (%)
Age (years)	
18-29	606 (23.6)
30-49	893 (34.8)
50-69	894 (34.8)
≥70	175 (6.8)
Sex	
Male	2091 (81.4)
Female	477 (18.6)
Ethnicity	
Han	2229 (86.8)
Other	339 (13.2)
Education	
Primary and below	797 (31)
Junior high school	783 (30.5)
Senior high school	360 (14)
College	607 (23.6)
Missing	21 (0.9)
Marital status	
Single	950 (37)
Married or cohabiting	1144 (44.6)
Divorced or widowed	457 (17.8)
Missing	17 (0.6)
Occupation	
Farmer	924 (36)
Other	1644 (64)
Route of infection	
Heterosexual intercourse	1480 (57.6)
Homosexual intercourse	951 (37)
Intravenous drug use	21 (0.8)
Other	116 (4.5)
Prior ARV[a] exposure	
Yes	81 (3.2)
No	2487 (96.8)
CD4 cell count before ART[b] (cell/mm^3)	
0-199	962 (37.5)
≥200	1554 (60.5)
Missing	52 (2)

[a] ARV: antiretroviral drugs.

[b] ART: antiretroviral therapy.

PDR and Mutations

Among the 2568 included study participants, PDR was detected in 7.4% (n=190). Table 2 shows the prevalence of overall PDR in Hebei (13/175; 7.4%), Jilin (19/206; 9.2%), Jiangsu (24/335; 7.2%), Zhejiang (34/385; 8.8%), Hubei (20/362; 5.5%), Chongqing (20/378; 5.3%), Sichuan (17/350; 4.9%) and Yunnan (43/377; 11.4%) in 2022. The prevalence of PDR in Yunnan (*P*=.002), Jilin (*P*=.047), and Zhejiang (*P*=.04) was much higher than that in Sichuan. Additionally, the prevalence of PDR in Yunnan was significantly higher than those in Chongqing (*P*=.002) and Hubei (*P*=.004).

Of the 2568 cases studied, the prevalence of PDR to NNRTIs, nucleoside reverse transcriptase inhibitors (NRTIs), and protease inhibitors (PIs) were 6.3% (n=163), 1.2% (n=32), and 0.2% (n=5), respectively. NNRTI-related DR to EFV and NVP was 6.1% (n=157) and 6.3% (n=163), respectively. There were 14 NNRTI-related resistance mutations; the most common mutations were K103N/S (n=81, 3.2%), followed by V179D/E/IL (n=39, 1.5%), and E138A/G/K (n=31, 1.2%). Among NRTI-related DR, the most frequent PDR drug was abacavir (n=19, 0.7%), followed by stavudine (n=17, 0.7%), emtricitabine (n=15, 0.6%), lamivudine (n=15, 0.6%), didanosine (n=13, 0.5%), zidovudine (n=10, 0.4%), and tenofovir (n=8, 0.3%). We detected 10 NRTI-related resistance mutations, of which M184V/I (n=13, 0.5%) was the most common. Even in the PR region, 0.2% (n=5) PI-related PDR to atazanavir/r and lopinavir/r was detected, including the PI-related DRMs L90M/IM (n=3, 0.1%), N88T/D (n=1, 0.1%), and V82IM (n=1, 0.1%) (Table 3).

Table 2. Prevalence of pretreatment HIV drug resistance by provincial-level administrative divisions in 2022 in China.

Province	Prevalence, n (%)				Participants (n=2568), n
	PR/RT[a] (n=190)	NNRTIs[b] (n=163)	NRTIs[c] (n=32)	PIs[d] (n=5)	
Hebei	13 (7.4)	9 (5.1)	4 (2.3)	0 (0)	175
Jilin	19 (9.2)	17 (8.2)	4 (1.9)	0 (0)	206
Jiangsu	24 (7.2)	22 (6.5)	1 (0.2)	1 (0.2)	335
Zhejiang	34 (8.8)	27 (7)	7 (1.8)	0 (0)	385
Hubei	20 (5.5)	19 (5.2)	1 (0.2)	0 (0)	362
Chongqing	20 (5.3)	14 (3.7)	7 (1.8)	1 (0.2)	378
Sichuan	17 (4.9)	16 (4.5)	0 (0)	1 (0.2)	350
Yunnan	43 (11.4)	39 (10.3)	8 (2.1)	2 (0.5)	377
All provinces	190 (7.4)	163 (6.3)	32 (1.2)	5 (0.2)	2568

[a] PR/RT: protease and reverse transcriptase.

[b] NNRTIs: nonnucleoside reverse transcriptase inhibitors.

[c] NRTIs: nucleoside reverse transcriptase inhibitors.

[d] PIs: protease inhibitors.

Table 3. Pretreatment HIV drug resistance and mutations among HIV-positive individuals in 2022 in China.

Antiretroviral drug	Prevalence, n	% (95% CI)	HIV drug resistance mutations and combination of mutations, n (%)
PR/RT[a] region (NNRTIs[b], NRTIs[c], PIs[d])	190	7.4 (6.4-8.4)	
NNRTIs	163	6.3 (5.4-7.3)	
Efavirenz	157	6.1 (5.2-7)	K103N/S: 81 (3.2)
Nevirapine	163	6.3 (5.4-7.3)	V179D/E/IL: 39 (1.5) • V179D+E138A:14 (0.5) • V179E+E138G: 11 (0.1) • V179D+V106I: 2 (0.1) • V179E+K238N: 2 (0.1) • V179D+G190E: 1 (0.1) • V179D+Y181C+G190A+H221Y: 1 (0.1) • V179D+Y188F+K238T: 1 (0.1) • V179E+A98G: 1 (0.1) • V179E+K103N+E138A: 1 (0.1) • V179E+K103N+E138G: 1 (0.1) • V179E+K103N+V106M: 1 (0.1) • V179E+V108I: 1 (0.1) • V179E+H221Y: 1 (0.1) • V179IL+Y188L: 1 (0.1) E138A/G/K: 31 (1.2) V106A/M/I: 15 (0.6) • V106I+K103N: 5 (0.2) • V106I+V179D: 2 (0.1) • V106I+Y188L: 1 (0.1) • V106I+G190A: 1 (0.1) G190C/A/E/S: 13 (0.5) K101E/P: 9 (0.4) Y181C: 9 (0.4) A98G: 8 (0.5) P225H: 8 (0.3) Y188F/L/C: 6 (0.2) V108I: 5 (0.2) H221Y: 5 (0.2) K238T: 3 (0.1) • K238N+V179E: 2 (0.1) L100I/V: 2 (0.1)
NRTIs	32	1.2 (0.8-1.7)	
Abacavir	19	0.7 (0.4-1.1)	M184V/I: 13 (0.5)
Emtricitabine	15	0.6 (0.3-0.9)	T215A/S/I: 5 (0.2)
Lamivudine	15	0.6 (0.3-0.9)	T69D/DN/ADN/del: 5 (0.2)
Didanosine	13	0.5 (0.2-0.8)	D67N: 5 (0.2)
Stavudine	17	0.7 (0.3-1)	K70E/T/R: 5 (0.2)
Tenofovir	8	0.3 (0.1-0.5)	K65R: 3 (0.1)
Zidovudine	10	0.4 (0.1-0.6)	L210W/MRW: 2 (0.1) L74I: 2 (0.1) K219Q: 1 (0.1) Y115F: 1 (0.1)
PIs	5	0.2 (0-0.3)	
Atazanavir/r	4	0.1 (0-0.3)	L90M/IM: 3 (0.1)
Darunavir/r	0		N88T: 1 (0.1)
Lopinavir/r	4	0.1 (0-0.3)	V82IM: 1 (0.1)

Antiretroviral drug	Prevalence, n	% (95% CI)	HIV drug resistance mutations and combination of mutations, n (%)
Mutual drug resistance to NNRTIs and NRTIs	10	0.4 (0.1-0.6)	

[a] PR/RT: protease and reverse transcriptase.

[b] NNRTIs: nonnucleoside reverse transcriptase inhibitors.

[c] NRTIs: nucleoside reverse transcriptase inhibitors.

[d] PIs: protease inhibitors.

Factors Associated With PDR

Table 4 lists the factors associated with PDR in PR/RT regions. Compared with participants without prior ARV drug exposure, those who had prior ARV drug exposure had 7.45 times greater odds of PDR (odds ratio [OR] 7.45, 95% CI 4.50-12.34). Compared to participants with the CRF07_BC subtype, those with the CRF55_01B subtype had 2.61 times greater odds of PDR (OR 2.61, 95% CI 1.41-4.83).

Table 4. Factors associated with pretreatment HIV drug resistance (PDR) among HIV-positive individuals in 2022 in China.

Variable	Number (n=2568)	PDR (n=190), n (%)	Odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Age (years)						
18-29	606	49 (8.1)	1.00	N/A[a]	—[b]	—
30-49	893	68 (7.6)	0.74 (0.64-1.37)	.74	—	—
50-69	894	60 (6.7)	0.82 (0.55-1.21)	.32	—	—
≥70	175	13 (7.4)	0.91 (0.48-1.72)	.78	—	—
Sex						
Male	2091	149 (7.1)	1.00	N/A	—	—
Female	477	41 (8.6)	1.23 (0.85-1.76)	.27	—	—
Ethnicity						
Han	2229	158 (7.1)	1.00	N/A	—	—
Other	339	32 (9.4)	1.37 (0.92-2.04)	.12	—	—
Education						
Primary and below	797	63 (7.9)	1.00	N/A	—	—
Junior high school	783	57 (7.3)	0.92 (0.631.33)	.64	—	—
Senior high school	360	21 (5.8)	0.72 (0.43-1.20)	.21	—	—
College	607	49 (8.1)	1.02 (0.69-1.51)	.91	—	—
Missing	21	0 (0)	—	—	—	—
Marital status[c]						
Single	950	84 (8.8)	1.00	N/A	—	—
Married or cohabiting	1144	72 (6.3)	0.69 (0.50-0.96)	.03	—	—
Divorced or widowed	457	34 (7.4)	0.83 (0.55-1.26)	.38	—	—
Missing	17	0 (0)	—	.98	—	—
Occupation						
Farmer	924	58 (6.3)	1.00	N/A	—	—
Other	1644	132 (8)	0.27 (0.16-2.64)	.10	—	—
Route of infection[c]						
Heterosexual intercourse	1480	102 (6.9)	1.00	N/A	—	—
Homosexual intercourse	951	72 (7.6)	1.11 (0.81-1.51)	.53	—	—
Intravenous drug use	21	6 (28.6)	5.41 (2.05-14.23)	<.001	—	—
Other	116	10 (8.6)	1.28 (0.65-2.51)	.48	—	—
Prior ARV[d] exposure[c,e]						
No	2487	163 (6.6)	1.00	N/A	1.00	N/A
Yes	81	27 (33.3)	7.13 (4.37-11.62)	<.001	7.45 (4.50-12.34)	.001
CD4 cell count before ART (cell/mm^3)						
0-199	962	65 (6.8)	1.00	N/A	—	—
≥200	1554	122 (7.9)	1.18 (0.86-1.61)	.31	—	—
Missing	52	3 (5.8)	0.85 (0.26-2.79)	.78	—	—
HIV subtype[c,e]						
CRF07_BC	1137	69 (6.1)	1.00	N/A	—	N/A
CRF08_BC	376	31 (8.2)	1.39 (0.89-2.16)	.14	1.09 (0.68-1.73)	.73

Variable	Number (n=2568)	PDR (n=190), n (%)	Odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
CRF01_AE	786	60 (7.6)	1.28 (0.89-1.83)	.18	1.27 (0.89-1.83)	.19
CRF55_01B	105	14 (13.3)	2.38 (1.29-4.39)	.006	2.61 (1.41-4.83)	.002
B	77	8 (10.4)	1.79 (0.83-3.88)	.14	1.61 (0.73-3.57)	.24
Other	87	8 (9.2)	1.57 (0.73-3.38)	.25	1.43 (0.65-3.14)	.37
Provincial-level administrative divisions[c]						
Sichuan	350	17 (4.9)	1.00	N/A	—	—
Hebei	175	13 (7.4)	1.57 (0.75-3.32)	.24	—	—
Jilin	206	19 (9.2)	1.99 (1.01-3.92)	.047	—	—
Jiangsu	335	24 (7.2)	1.51 (0.80-2.87)	.21	—	—
Zhejiang	385	34 (8.8)	1.90 (1.04-3.46)	.04	—	—
Hubei	362	20 (5.5)	1.15 (0.59-2.23)	.69	—	—
Chongqing	378	20 (5.3)	1.09 (0.56-2.13)	.79	—	—
Yunnan	377	43 (11.4)	2.52 (1.41-4.51)	.002	—	—

[a] N/A: not applicable.

[b] Not available.

[c] Variables screened by stepwise regression.

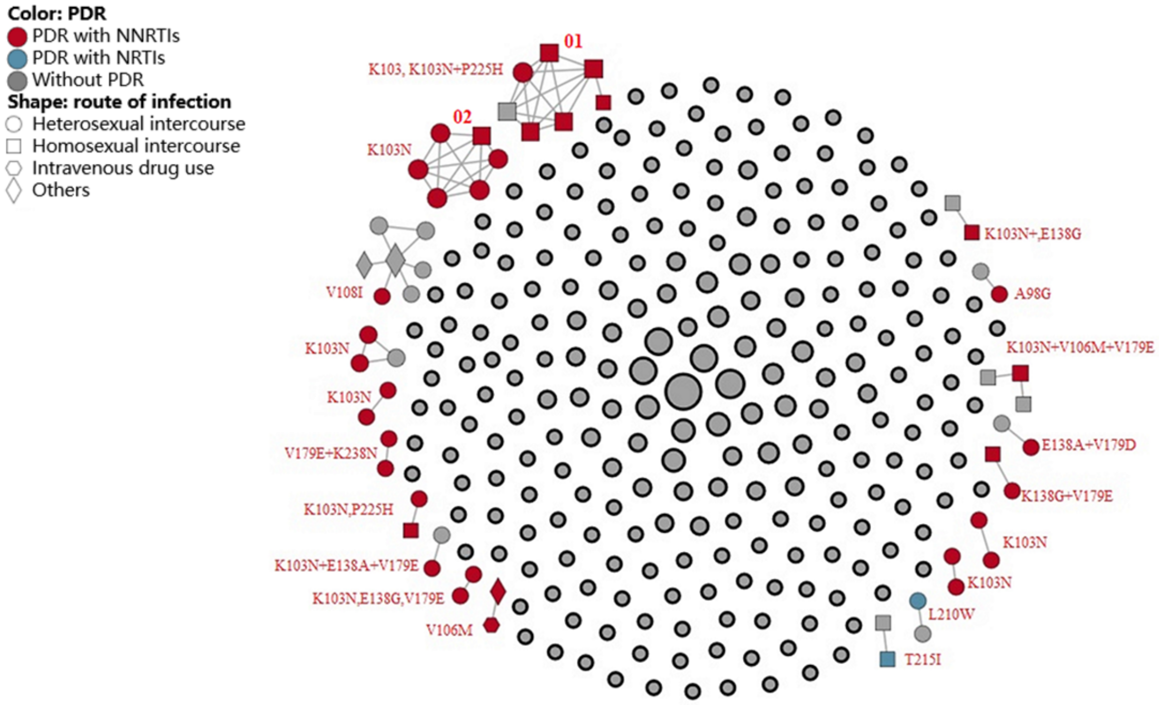
[d] ARV: antiretroviral drugs.

[e] Variables selected into multivariable logistic regression models (without prior ARV exposure serving as the reference group and with HIV subtype represented as 5 dummy variables with CRF07_BC serving as the reference group).

Molecular Transmission Network

Of the 2568 sequences in the PR/RT regions, 13 sequences were removed because the sequences in the PR/RT regions were shorter than 1000 bp or the number of mixed bases was >5%. Hence, a total of 2555 sequences in the PR/RT regions were used to construct a molecular transmission network. Under the optimal genetic distance threshold of 0.01 substitutions per site, 618 (24.2%) sequences (nodes) were linked to a total of 253 clusters (size range 2-13). We identified 38 sequences with PDR strains among 19 clusters. Most PDR strains linked to clusters had resistance to NNRTIs (36 of 38 sequences); DRM sites included K103, E138, V179, P225, V106, V108, L210, T215, P225, K238, and A98. Individual PDR status was not statistically significantly associated with belonging to a cluster. Cluster 01 with PDR consisted of 6 nodes in Jiangsu. Cluster 02 included HIV-positive individuals from 5 nodes in the same clinic sites in Zhejiang and 1 node in Hubei (Figure 1).

Figure 1. Molecular transmission network among HIV-positive individuals in 2022 in China.



Discussion

In 2022, a cross-sectional survey was conducted to examine the prevalence of PDR among HIV-positive individuals in 8 PLADs in China. The results showed that the overall prevalence of PDR was 7.4%. According to the WHO's qualitative classification of PDR prevalence levels, a prevalence below 5% is considered low, a prevalence between 5% and 15% is moderate, and a prevalence ≥15% is high [5]. Therefore, the overall prevalence of PDR in 2022 in China and those of 7 PLADs (all but Sichuan) were considered moderate. The prevalence of PDR in 13 PLADs and 31 PLADs in China was 6.8% and 4.4% in 2017 and 2018, respectively [34,35]. Monitoring and surveillance are needed to assess the extent to which resistant strains have spread, and new technologies and strategies should be applied to effectively curb the spread of PDR.

Our study revealed that the PDR rate of NNRTIs was 6.3%. However, the PDR NNRTI prevalence was 10.3% for Yunnan. We observed higher PDR rates in Zhejiang and Yunnan in 2022 compared to surveys conducted in 2018 (3.7% in Zhejiang and 3.8% in Yunnan) and this increase in PDR prevalence was driven mainly by NNRTI resistance. According to WHO ART guidelines developed in 2016, it is recommended to consider implementing supplementary first-line antiviral treatment plans when the prevalence of PDR is equal to or greater than 10% for NNRTIs in HIV-positive individuals who are starting first-line ART [36]. Further research is needed to identify the underlying reasons for the high NNRTI PDR rates in Yunnan. From 2014 to 2020, 21 of 30 countries that reported investigations to the WHO had an NVP or EFV PDR rate that exceeded 10% among HIV-positive individuals starting first-line ART [37]. In our study, we found that the prevalence of EFV and NVP DR was 6.1% and 6.3%, respectively. The main DRM sites identified were K103, E138, V179, and V106. These results are consistent with earlier surveys conducted in China in 2015 and 2018 [32,34]. Studies from Zimbabwe, Kenya, and Argentina have identified K103 as the most common DRM site for NNRTIs used as first-line antiviral drugs [38-40]. National HIV treatment strategies and regimens should be reconsidered given the high prevalence of NNRTI DR. In 2023, the *National Free HIV Antiviral Treatment Drug Treatment Manual* was released in China, and rilpivirine was added as the first-line treatment regimen [41]. The RPV resistance rate in our study was 5.3% (137/2568). The PDR after rilpivirine is applied to first-line treatment alternatives should be continuously monitored.

Our 2022 study indicated that PDR rates have increased in China over the past 7 years for NRTI DR (from 1.1% in 2015 and 0.8% in 2018 to 1.2% in 2022) and PI DR (from 0.02% in 2015 and 0.07% in 2018 to 0.2% in 2022) [34]. Nonetheless, PDR rates for NRTIs and PIs in China generally appear to be lower than the WHO's estimated regional PDR rates in Africa (NRTI PDR: 6.1%; PI PDR: 0.3%), the Americas (NRTI PDR: 6.4%; PI PDR: 0.8%), Southeast Asia (NRTI PDR: 3.1%; PI PDR: 0.4%), and the Western Pacific (NRTI PDR: 4%; PI PDR: 0%) [37].

Results of the multivariable analysis indicated that individuals who had prior ARV drug exposure had 7.45 times greater odds of PDR compared to those without prior ARV drug exposure. Such findings are consistent with research from Mexico, African countries, and Northeast China [42-44]. There was an increasing trend of the proportion of CRF55_01B samples, rising from 2.3% in 2015 to 3.9% in 2018 and 4.1% in 2022 in China [34]. Additionally, participants with the CRF55_01B subtype had 2.61 times greater odds of PDR (OR 2.61, 95% CI 1.41-4.83) compared to participants whose subtype was CRF07_BC. This may be due to association with V179 as a characteristic DR site of CRF55_01B, and V179 being frequently associated with E138G and causing low resistance to EFV and NVP [45].

Controlling the spread of CRF55_01B may be helpful for mitigating the proliferation of drug resistance. Furthermore, it is imperative to enhance the quality and effectiveness of intervention services for HIV-positive individuals.

In this study, we did not detect any statistically significant associations between having a PDR strain and belonging to clusters. Such results are similar with findings from Shijiazhuang city, Hebei Province and Liangshan region, Sichuan Province [46,47]. One possible explanation for these findings is that the transmission potential of DR strains is weaker than that of nonresistant wild strains [47,48]. The predominant subtype among the sequences in the molecular transmission network was CRF07_BC (representing 46.6% of strains), which is consistent with findings from a 2018 national molecular transmission network survey in China [32]. Clustered transmission of K103N DRMs in our study was observed in the molecular transmission network and is similar to findings from molecular transmission network studies in coastal Kenya [49]. It is worth noting that 6 DR strains in cluster 01 were from HIV-positive individuals at identical sites in Jiangsu Province. Epidemiological data indicated that all 6 were adult men aged younger than 30 years, of whom 5 reported homosexual transmission and 1 reported heterosexual transmission. Cluster 02 contained 6 samples with DR strains from HIV-positive individuals in the same clinic sites in Zhejiang Province, of whom 5 reported heterosexual transmission and 1 reported homosexual transmission, suggesting a possible clustered transmission of DR strains in that area.

Our study has several limitations. Due to sampling limitations, clinic sites were established in different counties and districts, and only 20-80 patients were enrolled in each site during a certain period of time. This may have resulted in smaller observable clusters (containing 2-13 nodes) in the molecular network and a lower network access rate. In addition, when detecting drug-resistant strains using genotyping, Sanger sequencing may not be able to detect up to 20% of less prevalent PDR strains [24]. This may have resulted in an underestimation of the PDR rate in HIV-positive individuals.

The findings from our 2022 PDR survey in China indicate a need to review and reconsider the standard ART protocol in some PLADs. Additional resources and strategies are needed in regions within China experiencing a higher prevalence of PDR. Possible interventions include extending pilot genotypic PDR testing to rapidly adjust medication regimens for HIV-positive individuals identified with PDR and targeted measures to improve medication compliance and prevention.

Acknowledgments

This work was supported by the Ministry of Science and Technology of China (grants 2022YFC2305205 and 2022YFC2305201), National Natural Science Foundation of China (grant 82260670), China Capital's Funds for Health Improvement and Research (grant 2022-1G-3011), Beijing Municipal Science & Technology Commission (grant D161100000416002), Guangxi Bagui Honor Scholarship, and Chinese State Key Laboratory of Infectious Disease Prevention and Control. The authors thank the study participants for their support and the staff at the clinical sites and provincial Chinese Center for Disease Control and Prevention for their help in collecting blood samples and data.

Authors' Contributions

HC, YR, HX, LL, YF, and YS were responsible for the study design and planning. HC, CS, YZ, XX, and ML were responsible for the statistics and figures. HC, J Hao, DL, XL, PX, DW, CS, J Hu, and JZ conducted the experiments and collected the data. HC, YR, and SWP drafted the manuscript. YR and HX guided the whole study and revised the article. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample size of HIV-positive individuals in 8 provincial-level administrative divisions in 2022 in China.
[DOCX File , 24 KB - publichealth_v9i1e50894_app1.docx]

References

1. Zhang T, Liao L, Shao Y, Feng Y, Ruan Y, Xing H. Relationship between drug resistance and death in HIV-infected patients receiving antiretroviral therapy - 7 PLADs, China, 2010-2019. China CDC Wkly 2021 May 02;3(14):291-297 [FREE Full text] [doi: 10.46234/ccdcw2021.068] [Medline: 34594868]
2. Volk JE, Nguyen DP, Hare CB, Marcus JL. HIV infection and drug resistance with unsupervised use of HIV pre-exposure prophylaxis. AIDS Res Hum Retroviruses 2018 May;34(4):329-330 [FREE Full text] [doi: 10.1089/AID.2017.0285] [Medline: 29262689]
3. Zhao Y, Zhang M, Shi CX, Huang J, Zhang F, Rou K, et al. Mortality and virological failure among HIV-infected people who inject drugs on antiretroviral treatment in China: an observational cohort study. Drug Alcohol Depend 2017 Jan 01;170:189-197 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.11.011] [Medline: 27987476]

4. Yu F, Li Q, Wang L, Zhao H, Wu H, Yang S, et al. Drug resistance to HIV-1 integrase inhibitors among treatment-naive patients in Beijing, China. Pharmgenomics Pers Med 2022;15:195-203 [FREE Full text] [doi: 10.2147/PGPM.S345797] [Medline: 35300056]
5. HIV drug resistance report 2019. World Health Organization. 2019. URL: https://apps.who.int/iris/handle/10665/325891 [accessed 2023-10-31]
6. Beck IA, Levine M, McGrath CJ, Bii S, Milne RS, Kingoo JM, et al. Pre-treatment HIV-drug resistance associated with virologic outcome of first-line NNRTI-antiretroviral therapy: a cohort study in Kenya. EClinicalMedicine 2020 Jan;18:100239 [FREE Full text] [doi: 10.1016/j.eclinm.2019.100239] [Medline: 31956856]
7. National Center for AIDS/STD Control and Prevention. National Handbook of Free AIDS Antiviral Drug Treatment (4th Edition). Beijing: People's Medical Publishing House; 2016.
8. Zuo L, Liu K, Liu H, Hu Y, Zhang Z, Qin J, et al. Trend of HIV-1 drug resistance in China: a systematic review and meta-analysis of data accumulated over 17 years (2001-2017). EClinicalMedicine 2020 Jan;18:100238 [FREE Full text] [doi: 10.1016/j.eclinm.2019.100238] [Medline: 31922125]
9. Lu D, Wu H, Yarla NS, Xu B, Ding J, Lu T. HAART in HIV/AIDS treatments: future trends. Infect Disord Drug Targets 2018;18(1):15-22. [doi: 10.2174/1871526517666170505122800] [Medline: 28474549]
10. Zhao B, Song W, Kang M, Dong X, Li X, Wang L, et al. Molecular network analysis reveals transmission of HIV-1 drug-resistant strains among newly diagnosed HIV-1 infections in a moderately HIV endemic city in China. Front Microbiol 2021;12:797771 [FREE Full text] [doi: 10.3389/fmicb.2021.797771] [Medline: 35069498]
11. Ye J, Hao M, Xing H, Zhang F, Wu H, Lv W, et al. Transmitted HIV drug resistance among individuals with newly diagnosed HIV infection: a multicenter observational study. AIDS 2020 Mar 15;34(4):609-619. [doi: 10.1097/QAD.0000000000002468] [Medline: 31895143]
12. Liu M, He X, Deng R, Tang S, Harypursat V, Lu Y, et al. Pretreatment drug resistance in people living with HIV: a large retrospective cohort study in Chongqing, China. HIV Med 2022 Mar;23 Suppl 1(Suppl 1):95-105 [FREE Full text] [doi: 10.1111/hiv.13253] [Medline: 35293098]
13. Hattori J, Shiino T, Gatanaga H, Mori H, Minami R, Uchida K, et al. Characteristics of transmitted drug-resistant HIV-1 in recently infected treatment-naive patients in Japan. J Acquir Immune Defic Syndr 2016 May 01;71(4):367-373. [doi: 10.1097/QAI.0000000000000861] [Medline: 26428230]
14. Taffa N, Roscoe C, Sawadogo S, De Klerk M, Baughman AL, Wolkon A, et al. Pretreatment HIV drug resistance among adults initiating ART in Namibia. J Antimicrob Chemother 2018 Dec 01;73(11):3137-3142 [FREE Full text] [doi: 10.1093/jac/dky278] [Medline: 30137412]
15. Kassaye SG, Grossman Z, Balamane M, Johnston-White B, Liu C, Kumar P, et al. Transmitted HIV drug resistance is high and longstanding in metropolitan Washington, DC. Clin Infect Dis 2016 Oct 15;63(6):836-843 [FREE Full text] [doi: 10.1093/cid/ciw382] [Medline: 27307507]
16. Tchouwa GF, Eymard-Duvernay S, Cournil A, Lamare N, Serrano L, Butel C, et al. Prevalence of pretreatment HIV drug resistance in Cameroon following a nationally representative WHO survey. J Antimicrob Chemother 2018 Oct 01;73(9):2468-2474. [doi: 10.1093/jac/dky221] [Medline: 29931063]
17. Wagner T, Zuckerman NS, Halperin T, Chemtob D, Levy I, Elbirt D, et al. Epidemiology and transmitted HIV-1 drug resistance among treatment-naïve individuals in Israel, 2010-2018. Viruses 2021 Dec 31;14(1):71 [FREE Full text] [doi: 10.3390/v14010071] [Medline: 35062274]
18. Ismael N, Wilkinson E, Mahumane I, Gemusse H, Giandhari J, Bauhofer A, et al. Molecular epidemiology and trends in HIV-1 transmitted drug resistance in Mozambique 1999-2018. Viruses 2022 Oct 09;14(9):1992 [FREE Full text] [doi: 10.3390/v14091992] [Medline: 36146798]
19. Matías-Florentino M, Chaillon A, Ávila-Ríos S, Mehta SR, Paz-Juárez HE, Becerril-Rodríguez MA, et al. Pretreatment HIV drug resistance spread within transmission clusters in Mexico City. J Antimicrob Chemother 2020 Mar 01;75(3):656-667 [FREE Full text] [doi: 10.1093/jac/dkz502] [Medline: 31819984]
20. Chen J, Chen H, Li J, Luo L, Kang R, Liang S, et al. Genetic network analysis of human immunodeficiency virus sexual transmission in rural Southwest China after the expansion of antiretroviral therapy: a population-based study. Front Microbiol 2022;13:962477 [FREE Full text] [doi: 10.3389/fmicb.2022.962477] [Medline: 36060743]
21. Yuan D, Du Z, Zhou J, Ye L, Su L, Yang H, et al. HIV-1 subtype diversity, drug resistance, and genetic transmission networks in men who have sex with men with virologic failure in antiretroviral therapy in Sichuan, China, 2011 to 2017. Medicine (Baltimore) 2019 Oct;98(43):e17585 [FREE Full text] [doi: 10.1097/MD.0000000000017585] [Medline: 31651864]
22. HIV drug resistance strategy, 2021 update. World Health Organization. 2021. URL: https://apps.who.int/iris/handle/10665/343175 [accessed 2023-10-31]
23. Surveillance of HIV drug resistance in adults initiating antiretroviral therapy (pre-treatment HIV drug resistance). World Health Organization. 2014. URL: http://apps.who.int/iris/bitstream/handle/10665/112802/9789241507196_eng.pdf?sequence=1 [accessed 2023-10-31]
24. Li M, Liang S, Zhou C, Chen M, Liang S, Liu C, et al. HIV drug resistance mutations detection by next-generation sequencing during antiretroviral therapy interruption in China. Pathogens 2021 Mar 25;10(3):264 [FREE Full text] [doi: 10.3390/pathogens10030264] [Medline: 33668946]

25. HIVAlign. HIV Sequence Database. URL: https://www.hiv.lanl.gov/content/sequence/VIRALIGN/viralign.html [accessed 2023-10-31]
26. HIV drug resistance database. Stanford University. URL: https://hivdb.stanford.edu/ [accessed 2023-10-31]
27. WHO HIVDR QC Tool. URL: https://sequenceqc-dev.bccfe.ca/who_qc [accessed 2023-10-31]
28. Kosakovsky Pond SL, Weaver S, Leigh Brown AJ, Wertheim JO. HIV-TRACE (TRAnsmission Cluster Engine): a tool for large scale molecular epidemiology of HIV-1 and other rapidly evolving pathogens. Mol Biol Evol 2018 Jul 01;35(7):1812-1819 [FREE Full text] [doi: 10.1093/molbev/msy016] [Medline: 29401317]
29. PS18-1802 list of attachments. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hiv/funding/announcements/ps18-1802/attachments.html [accessed 2023-10-31]
30. WHO, UNODC, UNAIDS technical guide for countries to set targets for universal access to HIV prevention, treatment and care for injecting drug users. World Health Organization. URL: https://www.unodc.org/documents/hiv-aids/idu_target_setting_guide.pdf [accessed 2023-11-06]
31. HIV-TRACE. URL: https://veg.github.io/hivtrace-viz/ [accessed 2023-10-31]
32. Chen H, Hu J, Song C, Li M, Zhou Y, Dong A, et al. Molecular transmission network of pretreatment drug resistance among human immunodeficiency virus-positive individuals and the impact of virological failure on those who received antiretroviral therapy in China. Front Med (Lausanne) 2022;9:965836 [FREE Full text] [doi: 10.3389/fmed.2022.965836] [Medline: 36106325]
33. Kang R, Li J, Chen H, Tang Z, Pan SW, Luo L, et al. Using longitudinal genetic-network study to understand HIV treatment-as-prevention. AIDS 2021 May 01;35(6):947-955. [doi: 10.1097/QAD.0000000000002812] [Medline: 33443369]
34. Hao J, Zheng S, Gan M, Dong A, Kang R, Li M, et al. Changing proportions of HIV-1 subtypes and transmitted drug resistance among newly diagnosed HIV/AIDS individuals - China, 2015 and 2018. China CDC Wkly 2021 Dec 31;3(53):1133-1138 [FREE Full text] [doi: 10.46234/ccdcw2021.251] [Medline: 35036036]
35. Kang R, Liang S, Ma Y, Liang S, Xiao L, Zhang X, et al. Pretreatment HIV drug resistance in adults initiating antiretroviral therapy in China, 2017. Infect Dis Poverty 2020 May 24;9(1):54 [FREE Full text] [doi: 10.1186/s40249-020-00668-5] [Medline: 32448388]
36. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach, 2nd ed. World Health Organization. 2016. URL: https://www.who.int/publications/i/item/9789241549684 [accessed 2023-10-31]
37. HIV drug resistance report 2021. World Health Organization. 2021. URL: https://apps.who.int/iris/handle/10665/349340 [accessed 2023-10-31]
38. Mungati M, Mhangara M, Gonese E, Mugurungi O, Dzangare J, Ngwende S, et al. Pre-treatment drug resistance among patients initiating antiretroviral therapy (ART) in Zimbabwe: 2008-2010. BMC Res Notes 2016 Jul 10;9:302 [FREE Full text] [doi: 10.1186/s13104-016-2101-8] [Medline: 27287672]
39. Bissio E, Barbás MG, Bouzas MB, Cudolá A, Salomón H, Espínola L, et al. Pretreatment HIV-1 drug resistance in Argentina: results from a surveillance study performed according to WHO-proposed new methodology in 2014-15. J Antimicrob Chemother 2017 Mar;72(2):504-510. [doi: 10.1093/jac/dkw445] [Medline: 27789684]
40. Silverman RA, Beck IA, Kiptinness C, Levine M, Milne R, McGrath CJ, et al. Prevalence of pre-antiretroviral-treatment drug resistance by gender, age, and other factors in HIV-infected individuals initiating therapy in Kenya, 2013-2014. J Infect Dis 2017 Dec 19;216(12):1569-1578 [FREE Full text] [doi: 10.1093/infdis/jix544] [Medline: 29040633]
41. "National Free HIV Antiviral Drug Treatment Manual (2023 Edition)" officially released. National Center for AIDS/STD Control and Prevention, China CDC. URL: https://ncaids.chinacdc.cn/zxzx/zxdteff/202306/t20230616_266839.htm [accessed 2023-10-31]
42. García-Morales C, Tapia-Trejo D, Matías-Florentino M, Quiroz-Morales VS, Dávila-Conn V, Beristain-Barreda Á, et al. HIV pretreatment drug resistance trends in Mexico City, 2017-2020. Pathogens 2021 Dec 08;10(12):1587 [FREE Full text] [doi: 10.3390/pathogens10121587] [Medline: 34959542]
43. Crowell TA, Danboise B, Parikh A, Esber A, Dear N, Coakley P, et al. Pretreatment and acquired antiretroviral drug resistance among persons living with HIV in four African countries. Clin Infect Dis 2021 Oct 05;73(7):e2311-e2322 [FREE Full text] [doi: 10.1093/cid/ciaa1161] [Medline: 32785695]
44. Gao L, Xia H, Zeng R, Wu Y, Zaongo SD, Hu Y, et al. Pre-treatment and acquired antiretroviral drug resistance among people living with HIV in Tianjin, China. HIV Med 2022 Mar;23 Suppl 1:84-94. [doi: 10.1111/hiv.13252] [Medline: 35293099]
45. Liu Y, Li H, Wang X, Han J, Jia L, Li T, et al. Natural presence of V179E and rising prevalence of E138G in HIV-1 reverse transcriptase in CRF55_01B viruses. Infect Genet Evol 2020 Jan;77:104098. [doi: 10.1016/j.meegid.2019.104098] [Medline: 31678241]
46. Wang X, Liu X, Li F, Zhou H, Li J, Wang Y, et al. Epidemiological surveillance of HIV-1 transmitted drug resistance among newly diagnosed individuals in Shijiazhuang, northern China, 2014-2015. PLoS One 2018;13(6):e0198005 [FREE Full text] [doi: 10.1371/journal.pone.0198005] [Medline: 29870534]

47. Liu L, Dong A, Liao L, Feng Y, Shao Y, Liang S, et al. Survey of pretreatment HIV drug resistance and genetic transmission network analysis among HIV patients in a high drug-use area of southwest China. Curr HIV Res 2019;17(6):441-451 [FREE Full text] [doi: 10.2174/1570162X17666191128101426] [Medline: 31778107]
48. Hofstra LM, Nijhuis M, Pingen M, Mudrikova T, Riezebos-Brilman A, Simoons-Smit AM, et al. Evolution and viral characteristics of a long-term circulating resistant HIV-1 strain in a cluster of treatment-naive patients. J Antimicrob Chemother 2013 Jul;68(6):1246-1250. [doi: 10.1093/jac/dkt038] [Medline: 23467175]
49. Hassan AS, Esbjörnsson J, Wahome E, Thiong'o A, Makau GN, Price MA, et al. HIV-1 subtype diversity, transmission networks and transmitted drug resistance amongst acute and early infected MSM populations from coastal Kenya. PLoS One 2018;13(12):e0206177 [FREE Full text] [doi: 10.1371/journal.pone.0206177] [Medline: 30562356]

Abbreviations

ART: antiretroviral therapy
ARV: antiretroviral drugs
DR: drug resistance
DRM: drug resistance mutation
EFV: efavirenz
NNRTIs: nonnucleoside reverse transcriptase inhibitors
NRTIs: nucleoside reverse transcriptase inhibitors
NVP: nevirapine
OR: odds ratio
PDR: pretreatment HIV drug resistance
PIs: protease inhibitors
PLADs: provincial-level administrative divisions
PR: protease
RT: reverse transcriptase
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 17.07.23; peer-reviewed by Y Tao, X Han; comments to author 23.08.23; revised version received 10.09.23; accepted 06.10.23; published 17.11.23.

Please cite as:
Chen H, Hao J, Hu J, Song C, Zhou Y, Li M, Chen J, Liu X, Wang D, Xu X, Xin P, Zhang J, Liao L, Feng Y, Li D, Pan SW, Shao Y, Ruan Y, Xing H
Pretreatment HIV Drug Resistance and the Molecular Transmission Network Among HIV-Positive Individuals in China in 2022: Multicenter Observational Study
JMIR Public Health Surveill 2023;9:e50894
URL: https://publichealth.jmir.org/2023/1/e50894
doi:10.2196/50894
PMID:37976080

Original Paper

Predicting Risky Sexual Behavior Among College Students Through Machine Learning Approaches: Cross-sectional Analysis of Individual Data From 1264 Universities in 31 Provinces in China

Xuan Li[1], BSc; Hanxiyue Zhang[1], PhD; Shuangyu Zhao[1], BSc; Kun Tang[1], DPhil

[1] Vanke School of Public Health, Tsinghua University, Beijing, China

Corresponding Author:
Kun Tang, DPhil
Vanke School of Public Health
Tsinghua University
No 30 Shuangqing Road
Haidian District
Beijing, 100084
China
Phone: 86 13671129425
Email: tangk@mail.tsinghua.edu.cn

Abstract

Background: Risky sexual behavior (RSB), the most direct risk factor for sexually transmitted infections (STIs), is common among college students. Thus, identifying relevant risk factors and predicting RSB are important to intervene and prevent RSB among college students.

Objective: We aim to establish a predictive model for RSB among college students to facilitate timely intervention and the prevention of RSB to help limit STI contraction.

Methods: We included a total of 8794 heterosexual Chinese students who self-reported engaging in sexual intercourse from November 2019 to February 2020. We identified RSB among those students and attributed it to 4 dimensions: whether contraception was used, whether the contraceptive method was safe, whether students engaged in casual sex or sex with multiple partners, and integrated RSB (which combined the first 3 dimensions). Overall, 126 predictors were included in this study, including demographic characteristics, daily habits, physical and mental health, relationship status, sexual knowledge, sexual education, sexual attitude, and previous sexual experience. For each type of RSB, we compared 8 machine learning (ML) models: multiple logistic regression (MLR), naive Bayes (BYS), linear discriminant analysis (LDA), random forest (RF), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), deep learning (DL), and the ensemble model. The optimal model for both RSB prediction and risk factor identification was selected based on a set of validation indicators. An MLR model was applied to investigate the association between RSB and identified risk factors through ML methods.

Results: In total, 5328 (60.59%) students were found to have previously engaged in RSB. Among them, 3682 (41.87%) did not use contraception every time they had sexual intercourse, 3602 (40.96%) had previously used an ineffective or unsafe contraceptive method, and 1157 (13.16%) had engaged in casual sex or sex with multiple partners. XGBoost achieved the optimal predictive performance on all 4 types of RSB, with the area under the receiver operator characteristic curve (AUROC) reaching 0.78, 0.72, 0.94, and 0.80 for contraceptive use, safe contraceptive method use, engagement in casual sex or with multiple partners, and integrated RSB, respectively. By ensuring the stability of various validation indicators, the 12 most predictive variables were then selected using XGBoost, including the participants' relationship status, sexual knowledge, sexual attitude, and previous sexual experience. Through MLR, RSB was found to be significantly associated with less sexual knowledge, more liberal sexual attitudes, single relationship status, and increased sexual experience.

Conclusions: RSB is prevalent among college students. The XGBoost model is an effective approach to predict RSB and identify corresponding risk factors. This study presented an opportunity to promote sexual and reproductive health through ML models, which can help targeted interventions aimed at different subgroups and the precise surveillance and prevention of RSB among college students through risk probability prediction.

(JMIR Public Health Surveill 2023;9:e41162) doi:10.2196/41162

KEYWORDS

risky sexual behavior; sexually transmitted infections; college students; machine learning; prediction; students; risk factor; STI; intervention; China; sex

Introduction

Risky sexual behavior (RSB) is defined as sexual activities that are more likely to lead to the risk of sexually transmitted infections (STIs) and unwanted pregnancies [1], including sex without contraceptive use and sex with an ineffective or unsafe contraceptive method. It is a serious issue among college students worldwide, especially in low- and middle-income countries [2,3]. Among 5 types of contraceptive methods under the classification criterion from the World Health Organization (WHO) [4], condoms and hormonal contraceptive methods are regarded as safe and highly effective for adolescents [5]. Studies have shown that consistent condom use is low among college students [6]. In previous research, it was estimated that approximately 40% of students did not use a condom during their last sexual encounter [7]. Casual sex and sex with multiple partners have also been regarded as RSBs [8-10]. Nearly half of the college students had casual sex experience [11,12], and approximately 50.7% of students were sexually active, with 42.3% of students having multiple sexual partners [13]. Thus, to help reduce the influence of RSB, it is important to intervene and prevent it among college students by identifying relevant risk factors and making RSB predictions.

A number of association studies have focused on the risk factors for RSB. A systematic review of 30 papers summarized 11 aspects of risk factors of RSB including sociodemographics, gender roles, substance use, and partner characteristics [14]. It was widely validated that being male [15,16], drinking alcohol [17], experiencing poverty, and experiencing peer pressure [18] are significant risk factors for RSB [19,20]. Mental health has also been linked to RSB, with higher depression resulting in more RSB and STIs [20,21]. In addition, numerous public health and sociological studies have found that romantic relationship status [22,23], sexual knowledge [1], and sexual attitude have a significant influence on RSB [24,25]. However, few studies have fully used these identified factors to make RSB predictions.

Previous predictive studies on RSB have mostly been based on conventional regression models, which have high limitations of assumption and less ideal effects on RSB prediction. To fill such gaps, machine learning (ML) offers a possible alternative for factor identification and outcome prediction. In the past few years, a large number of studies have emerged using ML to predict the occurrence of STIs, which have achieved ideal performance [24,25]. However, RSB, as the main transmission route for STIs, has gained little attention from ML for prediction.

The objective of this study was to develop an ML-based model to precisely predict RSB in college students. Through a cross-province survey in China, this study aimed to develop a series of ML models to predict different types of RSB among Chinese college students. By comparison, we adopted the optimal model to identify key risk factors to help recognize and predict college students' engagement in RSB, thus facilitating more precise intervention and prevention.

Methods

Participants and Research Procedures

We conducted a large-scale and internet-based survey, the National College Student Survey on Sexual and Reproductive Health in 2020 (NCSS-SRH 2020), sponsored by the China Family Planning Association (CFPA). Through multistage sampling from November 2019 to February 2020, a total of 55,757 Chinese college students from 241 universities completed the questionnaire survey. Voluntary participants were recruited using snowball sampling, and informed consent was obtained from each participant before completing the survey.

Among all participants, 1177 (2.11%) were excluded for either failing the attention check questions, ignoring the informed consent, or being outside the age range for college students (15-24 years old according to a standard definition of late adolescence and young adulthood by WHO) [26,27]. Samples with duplicated answers and variables with missing values over 5% were deleted. For the remaining variables, missing values were imputed with the use of multiple imputation. Due to the constraint of sexual orientation and past sexual experience, 8794 (15.78%) self-identified heterosexual students with sexual intercourse experience were finally included in the analyses.

RSB Outcomes

We classified RSB from 4 perspectives: (1) whether contraception was used, (2) whether the contraceptive method was effective and safe, (3) whether participants engaged in casual sex or sex with multiple partners, and (4) the integration of the former 3 perspectives.

Contraception use was evaluated through 2 dimensions. Regarding frequency, contraception use was measured by the question "Do you use contraception while having sex every time you have sexual intercourse?" Regarding practices, contraception use was measured by the question "Did you/your partner use contraception the last time you had sex?" For both questions, a "no" response was considered to indicate RSB. According to WHO guidance, 4 contraceptive types with over 10 specific methods were investigated as contraception use: hormonal contraceptive methods, intrauterine devices, emergency contraception, and condoms [4].

The effectiveness and safety of contraception was mainly determined by the method of contraception. If someone reported using "emergency contraception," "external ejaculation," or a "safe period" for contraception, they were considered a member of the RSB group. The effectiveness and safety of contraception were also evaluated through the frequency and practice dimensions, which were, respectively, measured by the question of whether such approaches were taken every time or the last time of sexual intercourse.

Casual or multiple sex was assessed with the following questions: "Have you ever had sex through 'booty calls', ‘one-night stands’, ‘buying sex’, or ‘sex with multiple partners’?” As before, a “yes” response was considered RSB.

Finally, integrated RSB dimension was the combination of contraception use, the safety of contraception, and casual sex or sex with multiple partners. As long as 1 of these 3 types of behaviors occurred, we considered the student to have met the criteria for integrated RSB.

Predictors

We included 126 potential predictors, including baseline characteristics (sex, age, ethnicity, religion, income, parental information, etc), daily habits (exercise, appearance and popularity, mobile phone addiction, alcohol and tobacco consumption, etc), physical and mental health status, relationship status, sexual knowledge, sexual education, sexual attitude, previous sexual experience, experience of sexual harassment and assault, etc. Among them, the age of the participant, the age of the participant’s partner, the income and expenditure of the participant, and the frequency of sports were treated as continuous variables. Other variables, including the degree of agreement, the frequency of participation, and the order of evaluation, were treated as either binary variables or ordered categorical variables.

Continuous variables were standardized, and categorical variables were split into multidimensional Boolean values before applying ML models. For each type of RSB, we used ML models to select and identify key variables to predict RSB among college students.

Statistical Analyses

For descriptive statistics of baseline characteristics, continuous data were presented as the mean (SD) or the median (IQR), and the Student *t*-test or the Kruskal-Wallis test was applied depending on the normality distribution. Categorical data were presented in the form of counts with percentages, and the chi-square test or the Fisher exact test was applied. $P<.05$ in a 2-tailed test was considered statistically significant in these tests. In addition, we developed multivariable mixed models with a logit link function to learn the specific linear relationship between RSB and the key variables screened by ML. The results were presented in the form of point estimates of coefficients and corresponding 95% CIs, and statistical significance was accepted when $P<.05$.

Model Development and Validation

We used 8 ML approaches: multiple logistic regression (MLR), naive Bayes (BYS), linear discriminant analysis (LDA), random forest (RF), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), deep learning (DL), and the ensemble model. The ensemble model used the average values of all other models’ predictive values to perform classification. The data set was split into a training set and a test set randomly in a ratio of 8:2, with 7035 (80%) samples in the training set and 1759 (20%) samples in the test set. Our models were built on the training set and then applied to the test set for RSB prediction. Model discrimination was assessed through the receiver operator characteristic (ROC) curve, and model performance was assessed through accuracy, precision, recall, F_1-score, the area under the receiver operator characteristic curve (AUROC), and the root-mean-square error (RMSE) calculated on the test set, which were presented in the form of the mean (SD).

Through comparison, we chose the optimal model for RSB prediction and included the most predictable variables in the model. The appropriate number of predictable variables was determined by the turning points of the model performance indicators. If all indicators did not change significantly through statistical testing when a new variable was added, we considered the turning point to have been reached.

To obtain the optimal performance for each model, we adopted the minimum distance (MD) method to select cut-off points to discriminate predictive values into 0 or 1. The MD method regarded the point closest to (0,1) on the ROC curve as the optimal cut-off point. To ensure the reliability and minimize the sensitivity of the results, we used a 10-fold cross-validation method to select the tuning hyperparameters as well as cut-off points. In addition, we repeated this process 10 times to minimize the influence of accidental circumstances.

All models were generated using R version 4.0.3 (R Core Team and the R Foundation for Statistical Computing). We used the *glmnet* package for MLR, the *e1071* package for BYS, the *MASS* package for LDA, the *randomForest* package for the RF, the *gbm* package for GBM, the *xgboost* package for XGBoost, and the *h20* package for DL.

Ethical Considerations

Ethical approval was obtained from the Institutional Review Board of Tsinghua University (IRB no. 20190083). All participants provided informed consent online, which was set before answering the questionnaire and emphasized the autonomy of participating and the ability to withdraw at any time. The privacy of personal information was protected throughout the study via anonymous data collection, and confidentiality was maintained by asking participants to provide honest answers. Eligible participation in this survey was voluntary and was not compensated.

Results

Baseline Characteristics

The baseline characteristics of the participants are shown in Table 1. Our sample covered all provincial-level administrative regions in China and achieved a relatively good balance in the sample division of eastern, central, western, and northeastern regions (n=4758, 54.11%, n=1484, 16.88%, n=2140, 24.33%, and n=412, 4.69%, respectively), as well as the sex ratio (males: n=3918, 44.55%; females: n=4876, 55.45%). A total of 8794 students were included, among which 3682 (41.87%) did not use contraception every time they engaged in sexual intercourse, 3602 (40.96%) did not use safe or effective contraceptive methods every time they engaged in sexual intercourse, 1157 (13.16%) had casual sex or sex with multiple partners, and 5328 (60.59%) had experience with at least 1 of those 3 former behaviors before. In addition, 307 (3.49%) participants did not use contraception and 2140 (24.33%) used ineffective or unsafe contraceptive methods during their last sexual intercourse, the details of which are shown in Multimedia Appendix 1. The distribution of basic characteristics differed greatly in terms of RSB. Between the 2 groups with and without integrated RSB, the region of residence, sex, age, ethnicity, religious beliefs, urbanization of hometown, left-behind experience, migration experience, and self-assessment of family finances were significantly different (*P*<.05) among college students.

Table 1. Baseline characteristics of participants (N=8794) grouped by different types of RSB[a].

Characteristics	Overall	RSB type 1 (nonuse of contraception)[b]	RSB type 2 (ineffective or unsafe contraceptive method)[c]	RSB type 3 (casual sex or sex with multiple partners)[d]	RSB type 4 (integrated RSB)[e]
Region of residence, n (%); P_1<.001, P_2<.001, P_3=.007, P_4<.001[f]					
Eastern	4758 (54.11)	1770 (48.07)	1779 (49.39)	676 (58.43)	2697 (50.62)
Central	1484 (16.88)	660 (17.93)	627 (17.41)	163 (14.09)	910 (17.08)
Western	2140 (24.33)	1094 (29.71)	1026 (28.48)	263 (22.73)	1476 (27.70)
Northeastern	412 (4.69)	158 (4.29)	170 (4.72)	55 (4.75)	245 (4.60)
Sex, n (%); P_1<.001, P_2=.15, P_3<.001, P_4<.001					
Male	3918 (44.55)	1826 (49.59)	1638 (45.47)	646 (55.83)	2573 (48.29)
Female	4876 (55.45)	1856 (50.41)	1964 (54.53)	511 (44.17)	2755 (51.71)
Self-assessed gender-role conformity[g], n (%); P_1=.29, P_2=.70, P_3<.001, P_4=.52					
Low	277 (3.15)	104 (2.82)	109 (3.03)	60 (5.19)	177 (3.32)
Middle	4158 (47.28)	1759 (47.77)	1720 (47.75)	500 (43.22)	2515 (47.20)
High	4359 (49.57)	1819 (49.40)	1773 (49.22)	597 (51.60)	2636 (49.47)
Age (years), median (IQR); P_1<.001, P_2<.001, P_3<.001, P_4<.001	20.00 (19.00-21.00)	20.00 (19.00-21.00)	20.00 (19.00-21.00)	20.00 (19.00-22.00)	20.00 (19.00-21.00)
Ethnicity, n (%); P_1<.001, P_2<.001, P_3=.77, P_4<.001					
Han	7902 (89.86)	3240 (88.00)	3164 (87.84)	1043 (90.15)	4698 (88.18)
Minority	892 (10.14)	442 (12.00)	438 (12.16)	114 (9.85)	630 (11.82)
Religious beliefs, n (%); P_1=.005, P_2=.051, P_3=.17, P_4=.006					
No	8039 (91.41)	3329 (90.41)	3267 (90.70)	1045 (90.32)	4835 (90.75)
Yes	755 (8.59)	353 (9.59)	335 (9.30)	112 (9.68)	493 (9.25)
Average monthly expenditure (CNY)/US \$[h], median (IQR); P_1=.29, P_2=.64, P_3<.001, P_4=.47	1800.00 (1200.00-2500.00)/265.48 (176.99-368.73)	1800.00 (1200.00-2500.00)/265.48 (176.99-368.73)	1800.00 (1200.00-2500.00)/265.48 (176.99-368.73)	2000.00 (1500.00-3000.00)/294.98 (221.24-442.47)	1800.00 (1200.00-2500.00)/265.48 (176.99-368.73)
Urbanization of hometown, n (%); P_1<.001, P_2=.16, P_3<.001, P_4=.001					
Urban	4827 (54.89)	1893 (51.41)	1940 (53.86)	745 (64.39)	2846 (53.42)
Suburban	2669 (30.35)	1166 (31.67)	1103 (30.62)	281 (24.29)	1644 (30.86)
Rural	1298 (14.76)	623 (16.92)	559 (15.52)	131 (11.32)	838 (15.73)
Left-behind experience, n (%); P_1<.001, P_2<.001, P_3=.006, P_4<.001					
No	6184 (70.32)	2426 (65.89)	2450 (68.02)	854 (73.81)	3631 (68.15)
Yes	2610 (29.68)	1256 (34.11)	1152 (31.98)	303 (26.19)	1697 (31.85)
Migration experience, n (%); P_1=.04, P_2<.001, P_3=.14, P_4=.003					
No	6864 (78.05)	2830 (76.86)	2738 (76.01)	923 (79.78)	4101 (76.97)
Yes	1930 (21.95)	852 (23.14)	864 (23.99)	234 (20.22)	1227 (23.03)
Self-assessment of family finances[i], n (%); P_1<.001, P_2=.08, P_3=.14, P_4>=.005					
Low	652 (7.41)	316 (8.58)	293 (8.13)	89 (7.69)	434 (8.15)
Middle	7167 (81.50)	2985 (81.07)	2904 (80.62)	921 (79.60)	4303 (80.76)
High	975 (11.09)	381 (10.35)	405 (11.24)	147 (12.71)	591 (11.09)

[a] RSB: risky sexual behavior.

[b] Nonuse of contraception indicated that someone did not use contraception while having sex every time.

[c] Ineffective or unsafe contraception indicated that someone often used unsafe contraceptive methods (eg, emergency contraception, external ejaculation, and safe period).

[d] Casual sex or sex with multiple partners indicated that someone had engaged in casual sex or sex with multiple partners before.

[e] Integrated RSB was the combination of the former 3 types of RSB.

[f] P_1, P value of RSB type 1; P_2, P value of RSB type 2; P_3, P value of RSB type 3; P_4, P value of RSB type 4.

[g] Self-assessed gender role conformity is a 1-7–ordered categorical-scale question. We classified the responses into 3 groups: low conformity (1-2), middle conformity (3-5), and high conformity (6-7).

[h] CNY 1=US $0.145749.

[i] Self-assessment of family finances is a 1-7–ordered categorical-scale question. We classified the responses into 3 groups: low income (1-2), middle income (3-5), and high income (6-7).

Model Performance and Validation

To precisely identify RSB among Chinese college students, we used various ML models to execute RSB prediction on the test data set to choose the optimal model. The model performance for the 4 types of RSB is presented in Table 2, and the model performance for the other 2 types of RSB based on the last sexual intercourse is presented in Multimedia Appendix 1. Through the comparison from multiple rounds of experiments, it was obvious that some ML models had better efficiency than the traditional multilinear logistic regression model. Compared to the performance of MLR (for which the average AUROCs were 0.76, 0.71, 0.91, and 0.79, respectively), XGBoost, GBM, and the RF presented better performance in terms of accuracy, the F_1-score, and the AUROC.

To better understand model discrimination, we plotted ROC curves of all models on 4 types of RSB in Figure 1. Similarly, ROC curves of the other 2 types of RSB based on the last time of sexual intercourse are presented in Multimedia Appendix 1. It could be inferred that the curves of XGBoost and GBM were above the other curves, which suggests that these 2 models outperformed the others. Their average results were similar (average AUROCs were 0.77, 0.72, 0.94, and 0.80, respectively, in GBM and 0.78, 0.72, 0.94, and 0.80, respectively, in XGBoost). The ensemble model also played an effective role in predicting different types of RSB (average AUROCs were 0.77, 0.72, 0.93, and 0.80, respectively).

Table 2. Model performance among different types of RSB[a].

RSB	Model	Accuracy, mean (SD)	Precision, mean (SD)	Recall, mean (SD)	F_1-score, mean (SD)	AUROC[b], mean (SD)	RMSE[c], mean (SD)
Nonuse of contraception[d]							
	MLR[e]	0.70 (0.01)	0.63 (0.02)	0.70 (0.02)	0.66 (0.01)	0.76 (0.01)	0.44 (0.01)
	BYS[f]	0.66 (0.01)	0.58 (0.02)	0.68 (0.02)	0.62 (0.01)	0.71 (0.01)	0.53 (0.01)
	LDA[g]	0.70 (0.01)	0.63 (0.02)	0.71 (0.02)	0.66 (0.01)	0.76 (0.01)	0.44 (0.01)
	RF[h]	0.71 (0.01)	0.63 (0.02)	0.72 (0.02)	0.67 (0.01)	0.77 (0.01)	0.44 (0.00)
	GBM[i]	0.71 (0.01)	0.64 (0.02)	0.72 (0.02)	0.67 (0.01)	0.77 (0.01)	1.07 (0.01)
	XGBoost[j]	0.71 (0.01)	0.63 (0.02)	0.72 (0.03)	0.67 (0.01)	0.78 (0.01)	0.44 (0.01)
	DL[k]	0.65 (0.01)	0.58 (0.02)	0.65 (0.03)	0.61 (0.01)	0.70 (0.01)	0.52 (0.01)
	Ensemble	0.71 (0.01)	0.63 (0.02)	0.73 (0.02)	0.67 (0.01)	0.77 (0.01)	0.44 (0.01)
Ineffective or unsafe contraceptive method[l]							
	MLR	0.66 (0.01)	0.57 (0.02)	0.67 (0.02)	0.62 (0.01)	0.71 (0.01)	0.46 (0.01)
	BYS	0.63 (0.01)	0.54 (0.02)	0.65 (0.04)	0.59 (0.02)	0.68 (0.01)	0.53 (0.01)
	LDA	0.66 (0.01)	0.58 (0.02)	0.66 (0.03)	0.62 (0.01)	0.71 (0.01)	0.46 (0.01)
	RF	0.67 (0.01)	0.58 (0.02)	0.67 (0.03)	0.62 (0.01)	0.72 (0.01)	0.46 (0.00)
	GBM	0.67 (0.01)	0.59 (0.02)	0.66 (0.03)	0.62 (0.01)	0.72 (0.01)	1.08 (0.01)
	XGBoost	0.67 (0.01)	0.59 (0.02)	0.67 (0.03)	0.62 (0.01)	0.72 (0.01)	0.46 (0.00)
	DL	0.61 (0.01)	0.52 (0.02)	0.6 (0.04)	0.56 (0.02)	0.65 (0.02)	0.54 (0.01)
	Ensemble	0.67 (0.01)	0.58 (0.02)	0.67 (0.02)	0.62 (0.01)	0.72 (0.01)	0.46 (0.01)
Casual sex or sex with multiple partners[m]							
	MLR	0.83 (0.02)	0.44 (0.03)	0.84 (0.02)	0.57 (0.03)	0.91 (0.01)	0.26 (0.01)
	BYS	0.79 (0.02)	0.37 (0.03)	0.77 (0.03)	0.5 (0.02)	0.85 (0.01)	0.36 (0.01)
	LDA	0.83 (0.01)	0.43 (0.03)	0.85 (0.02)	0.57 (0.02)	0.90 (0.01)	0.27 (0.01)
	RF	0.87 (0.01)	0.5 (0.03)	0.87 (0.02)	0.64 (0.03)	0.94 (0.01)	0.24 (0.01)
	GBM	0.87 (0.01)	0.52 (0.03)	0.88 (0.02)	0.65 (0.02)	0.94 (0.01)	1.03 (0.00)
	XGBoost	0.88 (0.01)	0.53 (0.03)	0.88 (0.02)	0.66 (0.02)	0.94 (0.01)	0.23 (0.01)
	DL	0.83 (0.02)	0.43 (0.03)	0.83 (0.02)	0.57 (0.03)	0.90 (0.01)	0.29 (0.01)
	Ensemble	0.86 (0.02)	0.48 (0.03)	0.86 (0.02)	0.62 (0.03)	0.93 (0.01)	0.24 (0.01)
Integrated RSB[n]							
	MLR	0.72 (0.01)	0.80 (0.01)	0.71 (0.02)	0.75 (0.01)	0.79 (0.01)	0.43 (0.01)
	BYS	0.68 (0.01)	0.76 (0.01)	0.68 (0.03)	0.72 (0.02)	0.74 (0.01)	0.56 (0.01)
	LDA	0.71 (0.01)	0.79 (0.01)	0.70 (0.02)	0.74 (0.01)	0.77 (0.01)	0.44 (0.01)
	RF	0.73 (0.01)	0.80 (0.01)	0.73 (0.02)	0.77 (0.01)	0.79 (0.01)	0.43 (0.00)
	GBM	0.73 (0.01)	0.81 (0.01)	0.73 (0.02)	0.77 (0.01)	0.80 (0.01)	1.05 (0.01)
	XGBoost	0.73 (0.01)	0.81 (0.01)	0.73 (0.02)	0.77 (0.01)	0.80 (0.01)	0.43 (0.01)
	DL	0.67 (0.01)	0.77 (0.01)	0.66 (0.02)	0.71 (0.01)	0.73 (0.01)	0.5 (0.01)
	Ensemble	0.73 (0.01)	0.80 (0.01)	0.74 (0.02)	0.77 (0.01)	0.79 (0.01)	0.43 (0.01)

[a] RSB: risky sexual behavior.

[b] AUROC: area under the receiver operator characteristic curve.

[c] RMSE: root-mean-square error.

[d] Nonuse of contraception indicated that someone did not use contraception while having sex every time.

[e] MLR: multiple logistic regression.

[f] BYS: naive Bayes.

[g] LDA: linear discriminant analysis.

[h] RF: random forest.

[i] GBM: gradient boosting machine.

[j] XGBoost: extreme gradient boosting.

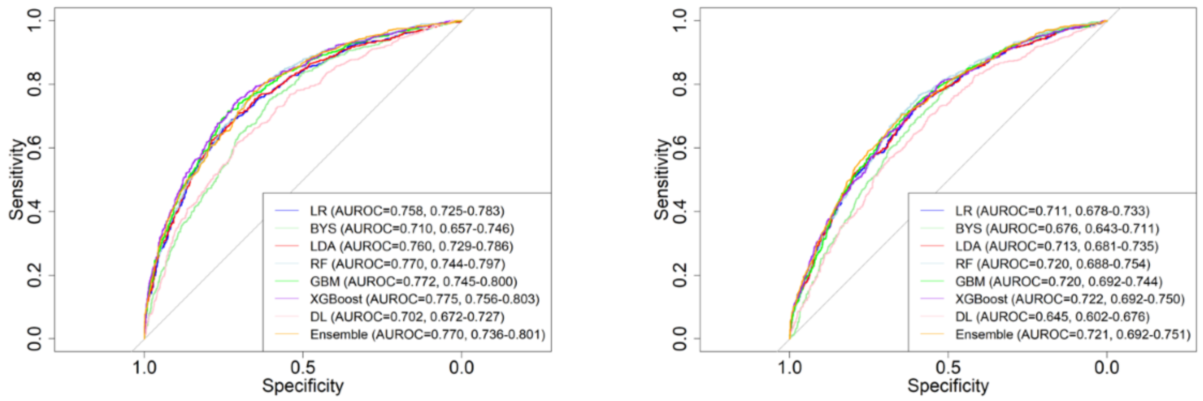
[k] DL: deep learning.

[l] Ineffective or unsafe contraception indicated that someone often used unsafe contraceptive methods (eg, emergency contraception, external ejaculation, and safe period).

[m] Casual sex or sex with multiple partners indicated that someone had engaged in casual sex or sex with multiple partners before.

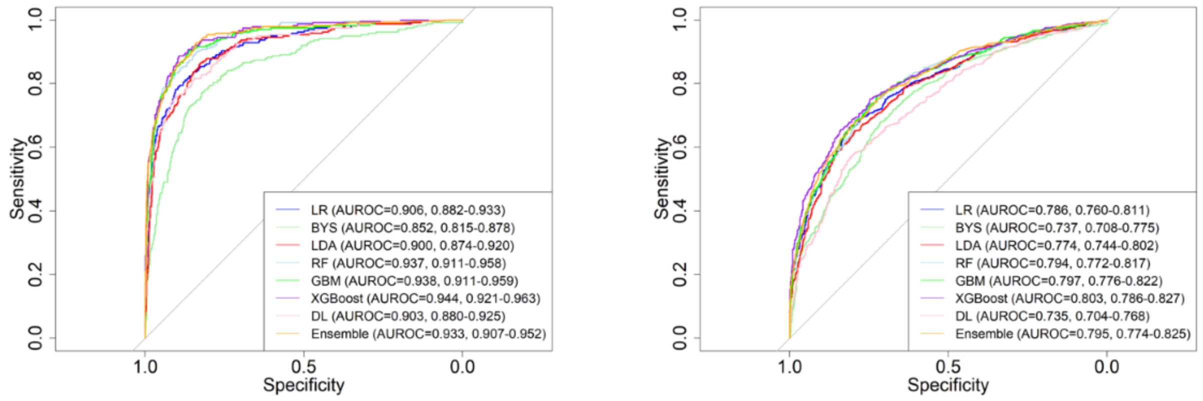
[n] Integrated RSB was the combination of the former 3 types of RSB.

Figure 1. AUROC curves among the different types of RSB. AUROC: area under the receiver operator characteristic curve; BYS: naive Bayes; DL: deep learning; GBM: gradient boosting machine; LDA: linear discriminant analysis; LR: logistic regression; RF: random forest; RSB: risky sexual behavior; XGBoost: extreme gradient boosting.



(a) Nonuse of contraception

(b) Ineffective or unsafe contraceptive method



(c) Casual sex or sex with multiple partners

(d) Integrated RSB

Variable Selection and Prediction for RSB

Through the comprehensive comparison of the evaluation indicators considering both efficiency and robustness, XGBoost was chosen to form the predicting model. To comprehensively predict RSB, we took integrated RSB as the outcome to select important variables, which were ranked in the order of importance according to the XGBoost model. Multimedia Appendix 1 shows the trend of 6 indicators as the number of variables increases, where points represent the average performance and lines represent the range. The turning points of the 6 indicators were 5th, 8th, 7th, 7th, 6th, and 12th, respectively. Thus, we finally chose 12 key variables for our prediction model, as presented in Figure 2. It could be inferred that RSB has multiple types of influencing factors, including relationship status, sexual knowledge, sexual attitudes, and previous sexual experience. The final prediction model we developed could estimate not only the probability of a student engaging in RSB but also the kind of RSB they were more likely to engage in.

To explore the specific association between the outcome variables and the predictive factors identified through XGBoost, we finally performed MLR, and the results of integrated RSB are presented in Table 3. It could be inferred that the lack of sexual knowledge and a liberal sexual attitude significantly increased the risk of RSB. In addition, regarding intimate relationship status, compared to the non-single group, the single group had a higher risk of RSB. Previous sexual experience also had a great influence on RSB. The greater the number of people participants had sexual intercourse with, the higher their risk of RSB. The more convenient the availability of contraceptives was, the less risk students would show RSB. The results of the other three types of RSB are shown in Multimedia Appendix 1.

Figure 2. The 12 most predictive variables selected by the XGBoost model. XGBoost: extreme gradient boosting.

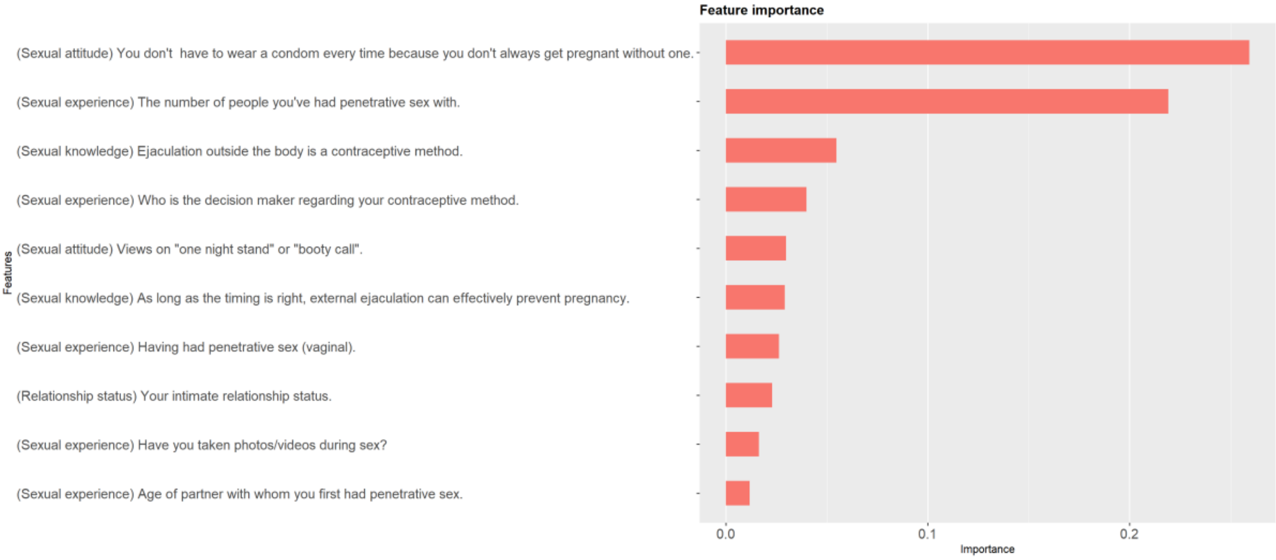


Table 3. Association between integrated RSB[a] and its key variables through MLR[b].

Variable and option	Coefficient estimate	SE	*Z* value	*P* value (>|*Z*|)
Sexual knowledge: ejaculation outside the body is a contraceptive method.				
Wrong	N/A[c]	N/A	N/A	N/A
Right	0.46	0.07	6.99	<.001
Sexual knowledge: having sex during a safe period is a contraceptive method.				
Wrong	N/A	N/A	N/A	N/A
Right	0.23	0.06	3.55	<.001
Sexual knowledge: as long as the timing is right, external ejaculation can effectively prevent pregnancy.				
Wrong	N/A	N/A	N/A	N/A
Right	0.42	0.06	–6.69	<.001
Sexual attitude: views on one-night stands or “booty calls.”				
I can accept it.	N/A	N/A	N/A	N/A
I can understand my friends doing this, but I cannot.	–0.44	0.07	–6.38	<.001
Totally unacceptable.	–0.28	0.07	–3.73	<.001
Sexual attitude: you don't have to wear a condom every time, because you don't always get pregnant without one.				
Strongly disagree	N/A	N/A	N/A	N/A
Relatively disagree	1.05	0.07	15.09	<.001
Not sure	1.28	0.11	11.66	<.001
Relatively agree	1.68	0.17	9.76	<.001
Strongly agree	1.25	0.25	5.07	<.001
Relationship status: your intimate relationship status.				
Single	N/A	N/A	N/A	N/A
Nonsingle	–0.39	0.06	–6.63	<.001
Sexual experience: having had penetrative sex (vaginal).				
Never has been and never will be acceptable.	N/A	N/A	N/A	N/A
Never has been but I can accept it in the future.	0.97	1.49	0.65	.52
I started before junior high school.	1.42	1.42	1.00	.32
I started since senior high school.	1.15	1.41	0.82	.42
I started since college.	0.81	1.41	0.57	.57
Sexual experience: have you taken photos/videos during sex?				
Never has been and never will be acceptable.	N/A	N/A	N/A	N/A
Never has been but I can accept it in the future.	0.01	0.07	0.16	.87
I started before junior high school.	0.17	0.35	0.49	.62
I started since senior high school.	0.45	0.13	3.48	.001
I started since college.	0.46	0.07	6.32	<.001
Sexual experience: age of partner with whom you first had penetrative sex.	0.01	0.01	1.38	.17
Sexual experience: the number of people you've had penetrative sex with.	0.33	0.02	13.93	<.001
Sexual experience: who is the decision maker regarding your contraceptive method?				
Myself.	N/A	N/A	N/A	N/A
My partner.	0.17	0.09	2.01	.04
By mutual negotiation.	0.11	0.06	1.82	.07
It depends.	1.30	0.13	10.06	<.001

Variable and option	Coefficient estimate	SE	*Z* value	*P* value (>|*Z*|)
Just use what we can find.	0.84	0.23	3.67	<.001
Others.	–0.31	0.33	–0.93	.36
Sexual experience: availability of contraceptives.				
Very convenient	N/A	N/A	N/A	N/A
Relatively convenient	0.30	0.06	5.41	<.001
Relatively inconvenient	0.41	0.08	4.91	<.001
Very inconvenient	0.28	0.14	2.02	.04

[a]RSB: risky sexual behavior.

[b]MLR: multiple logistic regression.

[c]N/A: not applicable as the baseline group.

Discussion

Principal Results

This study validated the effectiveness of ML models in predicting RSB among college students through comparisons of multiple models. Among various ML models, XGBoost performed the best in this task, with a higher accuracy, precision, F_1-score, and AUROC performance than others. Thus, we eventually used XGBoost to identify the 12 most predictive factors for total RSB, including relationship status, sexual knowledge, sexual attitudes, and previous sexual experience. This systematic process of data modeling as well as the accuracy of the final results indicated that ML approaches could have considerable value in RSB prediction and intervention among college students.

ML models were substantially superior to conventional regressions and should be recommended for more practical applications. Compared to the AUROC values of the MLR model (0.76, 0.71, 0.91, and 0.79 on the 4 types of RSB, respectively), XGBoost had a much higher effect in terms of AUROC values (0.78, 0.72, 0.94, and 0.80 on the 4 types of RSB, respectively). However, the capacity of explanation of MLR was nonnegligible. Thus, we finally used an MLR model again to investigate the linear association between outcomes and those important risk factors identified through XGBoost.

In this study, demographic characteristics and socioeconomic status were found to be significant factors of RSBs. Adolescents who were from an ethnic minority background, held religious beliefs, or had a lower family financial status tended to engage in more RSBs. This may be due to the lack of sexual education resources or specific religious customs. In addition, it was worth noting that students who had migration experience or were from rural hometowns were associated with increased RSBs. This finding was consistent with previous research and reflects the persisting gap in adolescent sexual and reproductive health between urban and rural areas [28,29]. Correlated with a lower level of education and socioeconomic status, rural-to-urban adolescents had less exposure to sexual knowledge and sex education [28,30].

In accordance with previous studies, romantic relationships were highly associated with RSB. It was found that college students with a romantic relationship had a significantly higher probability of using condoms during vaginal sex, oral sex, and anal sex [23]. On the one hand, students not in romantic relationships usually had fewer condom-carrying practices and a higher occurrence of unplanned sex. On the other hand, according to Rosenthal et al [31], rather than being concerned with the risk of STIs, students not in romantic relationships paid more attention to building intimacy through RSB, especially during casual sexual encounters. This is a worrisome mechanism since these students are more inclined to be unaware of each other’s health status, which could lead to a considerable risk of STIs.

This study also validated the role of sexual attitude in RSB. The more tolerant the students were toward condom nonuse and one-night stand, the higher their probability of engaging in RSB. In fact, it was widely validated that sexual attitude plays the most predictive influence in predicting RSB [32,33], which could be illustrated through the theory of planned behavior (TPB) that attitude is an activator of behavior [34]. Thus, sexual education should be facilitated to emphasize the importance and necessity of safe sex and dispel misunderstandings about it.

In addition, students’ previous sexual experience also had a large influence on the possibility of RSB, such as the number of sexual partners, the decision maker regarding the contraceptive method, and whether to take photos/videos during sex. This finding is also rational since well-practiced behavior would more likely recur due to the natural automation of initiates and controls [35].

Limitations

There are some limitations in this study. First, all outcomes and predictors were self-reported by participants, which may have caused recall bias and nonresponse bias. Since some questions were sensitive, participants may not have been willing to provide correct answers, such as on previous sexual experience, experience of sexual harassment, and assault. Second, the measurement of contraception use did not differentiate between the types of contraceptive methods used by the study population. Different contraceptive methods serve different functions. Condom use can prevent both STIs and unwanted pregnancies, while hormonal contraception is only effective for pregnancy prevention. Although the vast majority of participants used condoms for contraception, the predictive accuracy of contraceptive effectiveness may still have been obscured in this study because of the lack of differentiation across methods. Third, the definition of casual sex and sex with multiple partners was not well specified. Though these 2 types of behavior were undoubtedly validated as RSBs, the risks of contracting STIs and unwanted pregnancy can be kept relatively low with correct condom use. Fourth, although ML models have better predictive performance than traditional regression models, their explanatory performance is much weaker. We performed MLR to compensate for this drawback, but multicollinearity may exist among those selected risk factors, which could lead to inaccurate estimation. In particular, we had a large scale of many questions, with much similarity among them. Thus, questions could be divided into clusters to decrease the variable dimensions as well as strengthen the explanatory power. Fifth, since our model relied on a cross-sectional questionnaire, the outcomes and predictors were questioned concurrently. Under the same time window, it is difficult to identify the sequence of events. There may be causal inversion problems between predictors and outcomes, and thus, the prospective predictive efficiency of the model is hard to validate.

Comparison With Prior Work

There are considerable methodological, theoretical, and practical implications of this study. From the methodological aspect, we adopted a scientific and rigorous process to generate an RSB-predictive model using ML methods, which constitutes a research gap and urgent work to be done. The selection of the model, the adjustment of parameters, the comparison of indicators, and the finalization of variable numbers are of high reference value in the methodology. From a theoretical perspective, we identified a series of risk factors for RSB. We provided additional evidence for the association of demographic characteristics and socioeconomic status with RSB. Critical factors influencing RSB were also explored, including sexual attitude, sexual knowledge, relationship status, and sexual experience. Through the results presented, a comprehensive and evidence-based guideline was formed to facilitate more precise interventions and prevent RSB among adolescents and young adults. From a practical perspective, we developed a predictive model to help identify RSB among college students. Due to privacy concerns and the stigmatization of sexual behavior, it is often difficult to investigate the real prevalence of RSB among adolescents. Using the 12 predictors identified here, the model can predict not only the probability of a student engaging in RSB but also what kind of RSB they are more likely to engage in. With such a model, our study allows for more targeted intervention and prevention of RSB in students before they contract STIs, and thus, these students will be better able to avoid the various negative consequences of RSB, including STIs and unwanted pregnancy.

Conclusion

In summary, our study confirmed that ML approaches, especially XGBoost, have greater predictive effects for RSB than traditional regression models. Such ML-based assessment tools could generate new applications with considerable practical value, which would promote health at both the individual and the public level in the future.

Acknowledgments

The authors would like to express gratitude to the China Family Planning Association and the China Youth Network for their support in data collection and to all the students who participated in this survey.

Data Availability

The data that support the findings of this study are available upon reasonable request from the corresponding author (KT). The data are not publicly available due to research ethics board restrictions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[PDF File (Adobe PDF File), 472 KB - publichealth_v9i1e41162_app1.pdf]

References

1. Keto T, Tilahun A, Mamo A. Knowledge, attitude and practice towards risky sexual behaviors among secondary and preparatory students of Metu town, south western Ethiopia. BMC Public Health 2020 Sep 14;20(1):1394 [FREE Full text] [doi: 10.1186/s12889-020-09371-4] [Medline: 32928174]
2. Bellizzi S, Pichierri G, Menchini L, Barry J, Sotgiu G, Bassat Q. The impact of underuse of modern methods of contraception among adolescents with unintended pregnancies in 12 low- and middle-income countries. J Glob Health 2019 Dec;9(2):020429 [FREE Full text] [doi: 10.7189/jogh.09.020429] [Medline: 31673342]
3. World Health Organization. Adolescent Pregnancy. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-pregnancy [accessed 2022-10-01]

4. World Health Organization. Contraception. URL: https://www.who.int/health-topics/contraception#tab=tab_3 [accessed 2022-10-01]
5. Apter D. Contraception options: aspects unique to adolescent and young adult. Best Pract Res Clin Obstet Gynaecol 2018 Apr;48:115-127. [doi: 10.1016/j.bpobgyn.2017.09.010] [Medline: 29032945]
6. Pan X, Cong L, Ma Q, Xu G, Yu F, Zou Y. [Perception on AIDS infection risk and condom use among 2785 college students having had sexual experience in Zhejiang Province]. Zhonghua Liu Xing Bing Xue Za Zhi 2006 Jun;27(6):499-502. [Medline: 17152510]
7. Florence C. Reported Condom Use in Students Enrolled in a Personal Health and Wellness Course. 2012. URL: https://scholars.indstate.edu/bitstream/handle/10484/8210/Candace%20Florence.pdf?sequence=2&isAllowed=y [accessed 2023-01-09]
8. Garga S, Thomas M, Bhatia A, Sullivan A, John-Leader F, Pit S. Geosocial networking dating app usage and risky sexual behavior in young adults attending a music festival: cross-sectional questionnaire study. J Med Internet Res 2021 Apr 15;23(4):e21082 [FREE Full text] [doi: 10.2196/21082] [Medline: 33856354]
9. Thornton LC, Frick PJ, Ray JV, Wall Myers TD, Steinberg L, Cauffman E. Risky sex, drugs, sensation seeking, and callous unemotional traits in justice-involved male adolescents. J Clin Child Adolesc Psychol 2019 Dec 13;48(1):68-79. [doi: 10.1080/15374416.2017.1399398] [Medline: 29236522]
10. Homma Y, Wang N, Saewyc E, Kishor N. The relationship between sexual abuse and risky sexual behavior among adolescent boys: a meta-analysis. J Adolesc Health 2012 Jul;51(1):18-24 [FREE Full text] [doi: 10.1016/j.jadohealth.2011.12.032] [Medline: 22727072]
11. Longo LM, Ertl MM, Pazienza R, Agiliga AU, Dillon FR, Martin JL. Associations among negative urgency, sensation seeking, alcohol use, self-esteem, and casual sexual behavior for college students. Subst Use Misuse 2020 Dec 26;55(5):796-805. [doi: 10.1080/10826084.2019.1703748] [Medline: 31876218]
12. Owen J, Fincham FD, Moore J. Short-term prospective study of hooking up among college students. Arch Sex Behav 2011 Apr 4;40(2):331-341. [doi: 10.1007/s10508-010-9697-x] [Medline: 21203816]
13. Anteneh ZA. Prevalence and correlates of multiple sexual partnerships among private college students in Bahir Dar City, Northwest Ethiopia. Sci J Public Health 2013;1(1):9-17. [doi: 10.11648/j.sjph.20130101.12]
14. Medina-Perucha L, Family H, Scott J, Chapman S, Dack C. Factors associated with sexual risks and risk of STIs, HIV and other blood-borne viruses among women using heroin and other drugs: a systematic literature review. AIDS Behav 2019 Jan 2;23(1):222-251 [FREE Full text] [doi: 10.1007/s10461-018-2238-7] [Medline: 30073636]
15. Menon J, Mwaba S, Thankian K, Lwatula C. Risky sexual behaviour among university students. Int STD Res Rev 2016 Jan 10;4(1):1-7. [doi: 10.9734/isrr/2016/25462]
16. Somba MJ, Mbonile M, Obure J, Mahande MJ. Sexual behaviour, contraceptive knowledge and use among female undergraduates' students of Muhimbili and Dar es Salaam Universities, Tanzania: a cross-sectional study. BMC Womens Health 2014 Aug 07;14(1):94 [FREE Full text] [doi: 10.1186/1472-6874-14-94] [Medline: 25099502]
17. Yi S, Tuot S, Yung K, Kim S, Chhea C, Saphonn V. Factors associated with risky sexual behavior among unmarried most-at-risk young people in Cambodia. AJPHR 2014 Jan 7;2(5):211-220. [doi: 10.12691/ajphr-2-5-5]
18. Adeoti YF. Predisposing Factors Influencing Risky Sexual Behaviours as Expressed by Undergraduates in Osun State Nigeria. 2016. URL: https://tinyurl.com/2dnnzuch [accessed 2023-01-09]
19. Amare T, Yeneabat T, Amare Y. A systematic review and meta-analysis of epidemiology of risky sexual behaviors in college and university students in Ethiopia. J Environ Public Health 2019 Mar 20;2019:4852130-4852138 [FREE Full text] [doi: 10.1155/2019/4852130] [Medline: 31015844]
20. Jackson JM, Seth P, DiClemente RJ, Lin A. Association of depressive symptoms and substance use with risky sexual behavior and sexually transmitted infections among African American female adolescents seeking sexual health care. Am J Public Health 2015 Oct;105(10):2137-2142. [doi: 10.2105/ajph.2014.302493]
21. Vasilenko SA, Kugler KC, Butera NM, Lanza ST. Patterns of adolescent sexual behavior predicting young adult sexually transmitted infections: a latent class analysis approach. Arch Sex Behav 2015 Apr 22;44(3):705-715 [FREE Full text] [doi: 10.1007/s10508-014-0258-6] [Medline: 24449152]
22. Long L, Han Y, Tong L, Chen Z. Association between condom use and perspectives on contraceptive responsibility in different sexual relationships among sexually active college students in China: a cross-sectional study. Medicine (Baltimore) 2019 Jan;98(1):e13879 [FREE Full text] [doi: 10.1097/MD.0000000000013879] [Medline: 30608410]
23. Ssewanyana D, Sebena R, Petkeviciene J, Lukács A, Miovsky M, Stock C. Condom use in the context of romantic relationships: a study among university students from 12 universities in four Central and Eastern European countries. Eur J Contracept Reprod Health Care 2015 Feb 09;20(5):350-360. [doi: 10.3109/13625187.2014.1001024] [Medline: 25664484]
24. Bao Y, Medland NA, Fairley CK, Wu J, Shang X, Chow EP, et al. Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. J Infect 2021 Jan;82(1):48-59. [doi: 10.1016/j.jinf.2020.11.007] [Medline: 33189772]
25. Feller D, Zucker J, Yin MT, Gordon P, Elhadad N. Using clinical notes and natural language processing for automated HIV risk assessment. J Acquir Immune Defic Syndr 2018 Feb 01;77(2):160-166 [FREE Full text] [doi: 10.1097/QAI.0000000000001580] [Medline: 29084046]

26. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. Lancet 2011 Jun;377(9783):2093-2102. [doi: 10.1016/s0140-6736(11)60512-6]
27. World Health Organization. Older Adolescent (15 to 19 years) and Young Adult (20 to 24 years) Mortality. URL: https://tinyurl.com/mt8hvbpy [accessed 2022-10-01]
28. Sudhinaraset M, Astone N, Blum RW. Migration and unprotected sex in Shanghai, China: correlates of condom use and contraceptive consistency across migrant and nonmigrant youth. J Adolesc Health 2012 Mar;50(3 Suppl):S68-S74 [FREE Full text] [doi: 10.1016/j.jadohealth.2011.12.007] [Medline: 22340859]
29. Li S, Huang H, Cai Y, Xu G, Huang F, Shen X. Characteristics and determinants of sexual behavior among adolescents of migrant workers in Shangai (China). BMC Public Health 2009 Jun 19;9(1):195 [FREE Full text] [doi: 10.1186/1471-2458-9-195] [Medline: 19538756]
30. Zhao S, Gao E, Zabin LS. Unmet needs for reproductive health knowledge among unmarried migrant youth. J Reprod Contracept 2008 Dec;19(4):227-238. [doi: 10.1016/s1001-7844(08)60025-5]
31. Rosenthal D, Gifford S, Moore S. Safe sex or safe love: competing discourses? AIDS Care 1998 Feb 27;10(1):35-47. [doi: 10.1080/09540129850124569] [Medline: 9536200]
32. McEachan RRC, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the theory of planned behaviour: a meta-analysis. Health Psychol Rev 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]
33. Albarracín D, Johnson BT, Fishbein M, Muellerleile PA. Theories of reasoned action and planned behavior as models of condom use: a meta-analysis. Psychol Bull 2001 Jan;127(1):142-161 [FREE Full text] [doi: 10.1037/0033-2909.127.1.142] [Medline: 11271752]
34. Bargh JA, Chen M, Burrows L. Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. J Pers Soc Psychol 1996;71(2):230-244. [doi: 10.1037/0022-3514.71.2.230]
35. Ouellette JA, Wood W. Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. Psychol Bull 1998 Jul;124(1):54-74. [doi: 10.1037/0033-2909.124.1.54]

Abbreviations

AUROC: area under the receiver operator characteristic curve
BYS: naive Bayes
DL: deep learning
GBM: gradient boosting machine
LDA: linear discriminant analysis
MD: minimum distance
ML: machine learning
MLR: multiple logistic regression
RF: random forest
ROC: receiver operator characteristic
RMSE: root-mean-square error
RSB: risky sexual behavior
STI: sexually transmitted infection
WHO: World Health Organization
XGBoost: extreme gradient boosting

Edited by G Eysenbach; submitted 17.07.22; peer-reviewed by A Eleftheriou, PY Chiou; comments to author 21.10.22; revised version received 10.11.22; accepted 30.11.22; published 25.01.23.

Please cite as:
Li X, Zhang H, Zhao S, Tang K
Predicting Risky Sexual Behavior Among College Students Through Machine Learning Approaches: Cross-sectional Analysis of Individual Data From 1264 Universities in 31 Provinces in China
JMIR Public Health Surveill 2023;9:e41162
URL: https://publichealth.jmir.org/2023/1/e41162
doi:10.2196/41162
PMID:36696166

Original Paper

Spatiotemporal Trends in Self-Reported Mask-Wearing Behavior in the United States: Analysis of a Large Cross-sectional Survey

Juliana C Taube[1], AB; Zachary Susswein[1], BSc; Shweta Bansal[1], PhD

[1]Department of Biology, Georgetown University, Washington, DC, United States

Corresponding Author:
Shweta Bansal, PhD
Department of Biology
Georgetown University
Reiss Science Building, Room 406
37th and O Streets, NW
Washington, DC, 20057
United States
Phone: 1 202 687 9256
Email: shweta.bansal@georgetown.edu

Abstract

Background: Face mask wearing has been identified as an effective strategy to prevent the transmission of SARS-CoV-2, yet mask mandates were never imposed nationally in the United States. This decision resulted in a patchwork of local policies and varying compliance, potentially generating heterogeneities in the local trajectories of COVID-19 in the United States. Although numerous studies have investigated the patterns and predictors of masking behavior nationally, most suffer from survey biases and none have been able to characterize mask wearing at fine spatial scales across the United States through different phases of the pandemic.

Objective: Urgently needed is a debiased spatiotemporal characterization of mask-wearing behavior in the United States. This information is critical to further assess the effectiveness of masking, evaluate the drivers of transmission at different time points during the pandemic, and guide future public health decisions through, for example, forecasting disease surges.

Methods: We analyzed spatiotemporal masking patterns in over 8 million behavioral survey responses from across the United States, starting in September 2020 through May 2021. We adjusted for sample size and representation using binomial regression models and survey raking, respectively, to produce county-level monthly estimates of masking behavior. We additionally debiased self-reported masking estimates using bias measures derived by comparing vaccination data from the same survey to official records at the county level. Lastly, we evaluated whether individuals' perceptions of their social environment can serve as a less biased form of behavioral surveillance than self-reported data.

Results: We found that county-level masking behavior was spatially heterogeneous along an urban-rural gradient, with mask wearing peaking in winter 2021 and declining sharply through May 2021. Our results identified regions where targeted public health efforts could have been most effective and suggest that individuals' frequency of mask wearing may be influenced by national guidance and disease prevalence. We validated our bias correction approach by comparing debiased self-reported mask-wearing estimates with community-reported estimates, after addressing issues of a small sample size and representation. Self-reported behavior estimates were especially prone to social desirability and nonresponse biases, and our findings demonstrated that these biases can be reduced if individuals are asked to report on community rather than self behaviors.

Conclusions: Our work highlights the importance of characterizing public health behaviors at fine spatiotemporal scales to capture heterogeneities that may drive outbreak trajectories. Our findings also emphasize the need for a standardized approach to incorporating behavioral big data into public health response efforts. Even large surveys are prone to bias; thus, we advocate for a social sensing approach to behavioral surveillance to enable more accurate estimates of health behaviors. Finally, we invite the public health and behavioral research communities to use our publicly available estimates to consider how bias-corrected behavioral estimates may improve our understanding of protective behaviors during crises and their impact on disease dynamics.

(JMIR Public Health Surveill 2023;9:e42128) doi:10.2196/42128

KEYWORDS

COVID-19; face mask; nonpharmaceutical interventions; spatiotemporal; United States; survey bias; survey; US; behavior; effectiveness; disease; decision-making; community; surveillance

Introduction

Human behavior plays a key role in infectious disease transmission [1,2]. Individuals' decisions to get vaccinated, reduce their contacts, or wear a face mask, for example, can have a tremendous impact on disease dynamics [3-5]. The COVID-19 pandemic has highlighted that we are grossly limited in our ability to accurately measure and predict human behavior in the face of a novel pathogen. Yet, knowledge of how human behaviors vary over time and space is critical to assess the effectiveness of mitigation strategies, to forecast disease surges, and to parameterize coupled disease-behavior models [6]. In particular, there is a paucity of data on how the frequency of face mask wearing varies across the United States over different phases of the pandemic. This lack of fine-scale spatiotemporal data has forced public health organizations to adopt an inefficient one-size-fits-all approach to encourage masking nationwide, rather than directing resources and messaging to areas with the lowest uptake. Here, we identify the spatiotemporal trends in self-reported data on mask-wearing behavior across the United States from a large survey distributed from September 2020 to May 2021.

Mask wearing has been identified as an effective strategy to reduce the transmission of SARS-CoV-2. At the individual level, masks decrease both the amount of viral particles dispersed by an infectious wearer and the amount of those inhaled by an uninfected wearer [7]. Modeling studies at the population level (eg, [4,8,9]) have suggested that mask wearing can limit SARS-CoV-2 transmission and COVID-19 deaths, including under scenarios where masks are not worn universally or are not completely effective at blocking transmission. Randomized controlled trials (eg, [10]) have also demonstrated that mask wearing is an effective community-level intervention against COVID-19. Despite limited information at the time, the Centers for Disease Control and Prevention (CDC) initially recommended mask wearing on April 3, 2020 [11]. Lack of a national mandate, though, resulted in a heterogeneous landscape of mask policies across states, counties, towns, and even individual businesses in the United States [12,13]. Compounded with this spatial heterogeneity in mandates is additional heterogeneity in compliance, documented by localized observational studies (eg, [14]). A collection of systematic, accurate data on mask-wearing levels across the United States is therefore essential to informing our understanding of the role of mask wearing in the US COVID-19 pandemic trajectory.

To address this gap, researchers and organizations have implemented extensive surveys on human behavior, including mask wearing (eg, [15-17]). These surveys hold exciting promise, yet they have contributed relatively little to our understanding of human behavior, due to significant sampling limitations. Larger surveys with sufficient power to detect trends at local geographic scales are often not designed to capture a representative sample of the population. Demographic biases arising from a nonrepresentative sample can be addressed with standard statistical tools, such as survey weights, but other forms of bias, particularly nonresponse and social desirability biases, are more challenging to correct. Surveys about salient public health issues are especially likely to suffer from response bias; COVID-19–cautious individuals may be overrepresented in a survey about COVID-19 behavior. However, without estimates of the proportion of individuals in a given region who are COVID-19 cautious, there is no way to use survey weights on this demographic. Likewise, respondents may be influenced by social desirability bias when self-reporting COVID-19–preventive behaviors, such as vaccination, social distancing, or mask wearing, so that they respond in a manner deemed favorable by society despite being inaccurate [18]. Without observational or ground-truth data to validate survey responses, quantifying this social desirability bias is difficult. Furthermore, it is critical that ground-truth data to correct biases in health behavior be used at a fine spatial and temporal scale to avoid further exacerbation of biases (eg, [19]).

The value of surveys on public health behaviors can be further restricted when data collection is at the national or state level. Coarse-grained spatiotemporal information about human behavior is of limited utility, providing only sparse insight into local trends. Collecting responses at the national or state level ignores spatial heterogeneity at these finer scales, preventing the identification of these local effects that can drive disease dynamics. Spatial heterogeneity in not only drivers of disease transmission, such as human behavior, but also disease prevalence has been well documented across pathogens (eg, [20-22]). For example, differences in connectivity between counties or states can affect the timing and geographic scale of disease spread, while national scale mobility data elide these key patterns [23,24]. Likewise, aggregation of vaccination data to the state level can hide spatial clustering of unvaccinated individuals, which undermines herd immunity and could drive sustained measles outbreaks in the United States [25,26]. Despite the importance of detailed local estimates on the drivers of disease incidence, few studies have analyzed human behavior during the COVID-19 pandemic nationally at these fine spatial scales. Furthermore, most surveys are not conducted for long enough to capture human behavior changes over time, leaving scant opportunity to assess the effects of changing public health messaging/guidance or disease prevalence on human behavior.

Here, we systematically characterize mask wearing across the United States at a fine spatiotemporal scale for 9 months using a national survey and account for the bias in this survey. By comparing survey demographics and vaccination statuses with accurate ground-truth data, we estimate and account for survey and response biases in our analysis of masking behavior. With these bias-corrected estimates, we characterize the spatiotemporal heterogeneity in masking behavior at the county-month level across the United States. Finally, we examine the differences between self-reported and community-reported estimates of masking using an additional survey question, seeking to understand whether these 2 measures are good predictors of one another. Our results are the most precise estimates of masking in the United States during the COVID-19 pandemic, providing insight into the local variation in behavior in response to public health messaging and changes in COVID-19 incidence.

Methods

Study Design

In this study, we sought to characterize the spatiotemporal heterogeneity in self-reported masking behavior in the United States from the fall of 2020 to the spring of 2021. Due to the small sample size in some counties, we used Bayesian binomial regression models to estimate mask-wearing proportions each month. Recognizing that surveys are subject to several types of bias, we used raking and resampling of responses to correct for unrepresentative samples and self-reported vaccination status compared to ground-truth vaccination data to quantify nonresponse and social desirability biases. With these estimates, we were able to identify spatiotemporal trends in bias-corrected masking behavior and compare these values to reported community levels of masking in a different survey question.

Survey Data and Processing

We analyzed self-reported mask-wearing survey responses for all 50 US states and the District of Columbia using data from the US COVID-19 Trends and Impact Survey (CTIS) [27]. The CTIS was created by the Delphi Research Group at Carnegie Mellon University and distributed through a partnership with Facebook. Beginning in September 2020, a random state-stratified sample of active Facebook users was invited daily to take the survey about COVID-19 and report how often they wore a mask in the past 5-7 days (the number of days changed from 5 to 7 on February 8, 2021). The answer options were (1) “All of the time,” (2) “Most of the time,” (3) “Some of the time,” (4) “A little of the time,” (5) “None of the time,” and (6) “I have not been in public in the last 5-7 days” (Multimedia Appendix 1, Figure S13). To dichotomize these responses for an analysis of the proportion of respondents wearing masks, we dropped respondents who had not been in public recently or did not respond to the masking question, and considered responses of “all of the time” and “most of the time” as masking, and all other responses as not masking. This cut-off is reasonable, considering the raw proportions of responses in each category for September through May (Multimedia Appendix 1, Figure S14). Due to sample size constraints, we aggregated these responses to the county-month scale. We ignored potential heterogeneity at smaller temporal (weekly) and spatial (zip code) scales due to the limited sample size.

By dichotomizing masking responses, we also lost information about the frequency with which people mask, though we expect the effect of this choice to be minimal (see Multimedia Appendix 1 for details).

Bayesian Binomial Regression Model

Due to the small sample sizes in some US counties, we used Bayesian binomial regression models to develop reliable estimates of the proportion of individuals masking in a given county-month. Population density was used as a fixed effect; masking behavior has previously been linked to population density, and this variable was easily available at the county scale [14,28]. We fit separate models for each month, allowing for a temporal trend without explicitly modeling it by specifying a parametric form. We defined M_i as the number of respondents masking in county *i* (eg, respondents who masked most or all of the time in the past 5-7 days), N_i as the total number of respondents in county *i* ($M_i \leq N_i$), and p_i as the county-level probability of a response consistent with masking. We used the following model to estimate and :

where $D_i = \log_{10}(\text{population density}_i)$ for county *i*. We ran the model using *brms* [29] with the *cmdstanR* [30] backend. We ran the sampler with 4 chains for 3000 iterations per chain. Sampler diagnostics indicated efficient exploration and that the model had converged, $n_{eff} < 0.25$, n_{eff} per iteration ≥ 0.25, ≤ 1.01, E-BFMI > 0.25, and no transitions hit the maximum tree depth. All Pareto-smoothed importance sampling leave-one-out (PSIS-LOO) k statistic values were below 0.71, indicating that the model was robust to the influence of individual observations, and the distribution of Pareto k statistics was uniform, indicating that the model captured essential features of the data [31]. Posterior predictive checks indicated a good model fit (Multimedia Appendix 1, Figure S15) as did the plots of observed versus predicted and residual values (Multimedia Appendix 1, Figures S16 and S17). We note that our binomial regression approach compensated for the small sample sizes in some county-months, but the resulting estimates depend on the validity of our model structure.

We also explored more complex model specifications that included state- or county-level random effects. However, both models suffered from a lack of convergence or overfitting and produced functionally similar results. Thus, we opted for the more parsimonious model presented earlier for our main findings; details of these additional models can be found in Multimedia Appendix 1 (Figures S20 and S21).

Survey Raking and Resampling

We were unable to use the provided weights for responses to the CTIS, due to spatial and temporal mismatch with the scales of our data analysis. Thus, we calculated county-month weights for each observation using the anesrake package [32] and the US Census American Community Survey’s 2018 5-year data on county age, sex, and education distributions. Age, sex, and education distributions were based on each county’s population over the age of 18 years to match the survey sample. We did not use race or ethnicity data in the raking scheme, as their inclusion substantially reduced algorithm convergence, but note that race/ethnicity was moderately correlated with education (Cramer's V>0.10). We then resampled from these responses using the calculated weights to estimate a raked masking proportion, which was fed into the binomial models, as described before. We excluded observations with missing age, sex, or education responses from the raking process and assigned equal weights to observations from county-months that did not converge (additional details in Multimedia Appendix 1).

Estimation of CTIS Masking Bias

Given the likelihood of sampling, nonresponse, and social desirability biases, we generated bias-corrected estimates of masking in the United States. In the absence of ground-truth masking data with which to calibrate these CTIS responses, we turned to a different survey question for which ground-truth data were available.

Beginning in late December 2020, the CTIS asked respondents whether they had received a COVID-19 vaccine. The response options were (1) "Yes, "(2) "No," and (3) "I don't know." Meanwhile, ground-truth vaccination data were collected by combining state-reported and CDC data to estimate the percentage of people vaccinated in each county in the United States [33,34]. A comparison of CTIS responses and ground-truth vaccination data revealed that the estimates of COVID-19 vaccination based on CTIS responses were much higher than true vaccination rates at the US county scale (Multimedia Appendix 1, Figures S18 and S19, [19,35]). Assuming that masking survey responses suffer from the same bias issues (in magnitude and direction) as vaccination responses, this result would suggest that CTIS responses also substantially overestimate masking behavior. Thus, we approximated survey bias by comparing the CTIS vaccination responses to the ground-truth vaccination data at the county level and incorporating this bias into the model of CTIS masking behavior.

Like the masking data, the CTIS vaccination response data suffers from small and unrepresentative samples in some counties. Thus, we resampled the responses from April and May 2021 according to the survey weights we generated before and then used a (frequentist) binomial generalized linear mixed-effects model to estimate p_i, the proportion of respondents who were vaccinated (assumed to be partial vaccination, with 1 of a 1-dose or 2-dose vaccine) at the county level each week (details in Multimedia Appendix 1).

Given these modeled CTIS county-level vaccination proportions, we compared them with the true vaccination data to calculate the expected bias in reported survey responses relative to ground-truth data in county *i*:

$$\text{bias}_i = \text{logit}(\text{CTIS estimated vaccination proportion}_i) - \text{logit}(\text{true vaccination proportion}_i)$$

To increase the stability of our bias estimates, we used a linear mixed-effects model. This mixed-effects model used random intercepts, which penalizes extreme coefficient estimates to the overall mean, and assumed that the residual error in the estimates was normally distributed. This model generated a penalized estimate of survey bias for each county from the difference in modeled reported vaccination and ground-truth vaccination:

This model was implemented using *lmer* in the *lme4* package [36]. If there were no responses in county *i* or a bias estimate could not be calculated, bias estimates for this county were imputed by taking the mean of surrounding county estimates.

We then incorporated these estimates into a Bayesian binomial regression model with an offset for bias to estimate the bias-corrected probability of reporting masking in county *i*. We defined M_i as the number of respondents masking in county *i* out of N_i total respondents and p_i as the county-level probability of a response consistent with masking. We used the following model to estimate and :

where D_i = $\log_{10}$(population density$_i$) for county *i*. The bias-corrected proportion of individuals masking, c_i, was calculated as

We ran the model using *brms* [29] with the *cmdstanR* [30] backend. We ran the sampler with 4 chains for 3000 iterations per chain. Sampler diagnostics indicated efficient exploration and that the model had converged: $n_{eff}>1000$, n_{eff} per iteration≥0.3, and ≤ 1.01.

Community-Reported Masking

Beginning November 24, 2020, the CTIS asked a question about masking in one's community: "In the past 7 days, when out in public places where social distancing is not possible, about how many people would you estimate wore masks?" The answer options were (1) "All of the people," (2) "Most of the people," (3) "Some of the people," (4) "A few of the people," (5) "None of the people," and (6) "I have not been out in public places in the past 7 days." We dichotomized these responses and aggregated them to the county-month the same way as the self-reported CTIS masking responses for December 2020 through May 2021. We then modeled these community masking estimates the same way we modeled the CTIS masking data using Bayesian binomial regression and resampling weighted by survey weights but without a bias offset.

Spatiotemporal Analysis

All analyses were completed in R version 4.1.3 (R Core Team and the R Foundation for Statistical Computing), and maps were produced using *choroplethr* [37]. Urban-rural classes were from the National Center for Health Statistics (NCHS) Urban-Rural Classification Scheme [38].

Ethical Considerations

This study was reviewed by the Institutional Review Board of Georgetown University and was determined not to be human subject research.

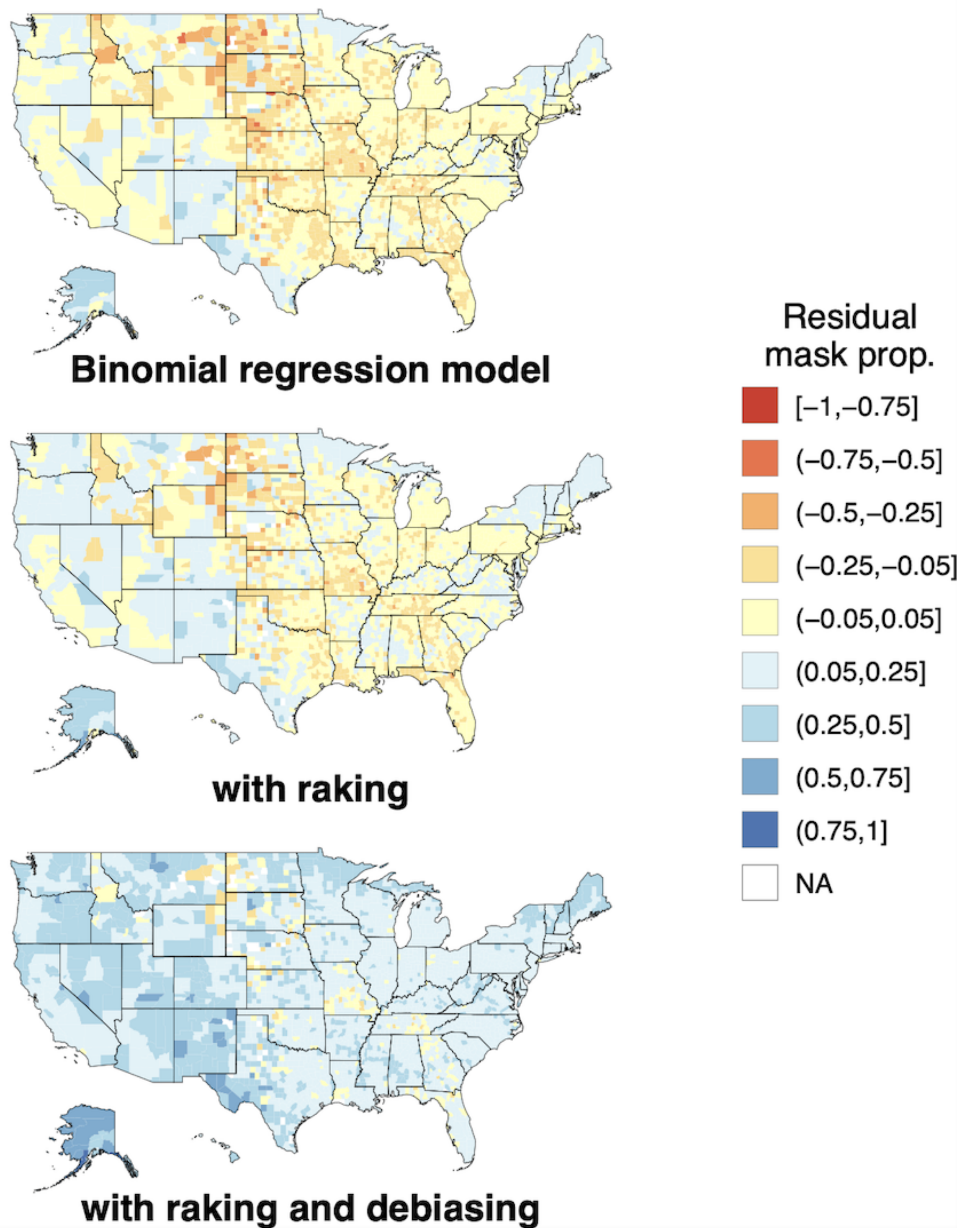
Results

To characterize the trends in the masking behavior in the United States during the COVID-19 pandemic, we used data from the CTIS conducted via Facebook from September 2020 through May 2021. Respondents self-reported how often they had worn a mask while in public in the past week (8,338,877 valid responses). We transformed these responses into a binary variable of masking or not masking and aggregated the responses to the county-month level to analyze spatiotemporal trends. To validate this data source, we analyzed a separate data set from Outbreaks Near Me and found consistent spatiotemporal patterns (Multimedia Appendix 1, Figures S1-S3), though both data sources suffer from issues of bias and small sample size. We addressed these issues in the CTIS data using binomial regression models to inform estimates of masking in counties with a small sample size, and raking and sample rebalancing on age, sex, and education to adjust for unrepresentative samples. Recognizing that CTIS responses to a question about vaccination overestimated the true vaccination rates, we quantified this bias for each county and used it to correct the estimates of masking behavior, assuming that vaccination and masking behavior responses were equally as biased. (Vaccination and mask wearing are both prosocial public health behaviors, which are socially desirable to report and are likely correlated [39-43].) We analyzed the overall spatial and temporal trends as well as fine-scale heterogeneity in the bias-corrected masking behavior estimates. Finally, we validated the bias-corrected CTIS values by comparing them to the respondents' estimates of the proportion of people masking in their community.

Spatially Heterogeneous Effects of the Binomial Regression Model, Survey Raking, and Debiasing

To demonstrate the spatially heterogeneous effects of our data-processing scheme, Figure 1 shows the difference between estimates from 3 separate models and the raw CTIS masking data. We refer to this difference as the residual, though it is only an indicator of model fit in Figure 1A. In Figures 1B and 1C, the residual values indicate where data corrections caused the largest changes in estimates compared to the original data. After modeling the data with binomial regression, estimates of masking proportions were higher than the observed values (Multimedia Appendix 1, Figure S4) in the central United States and slightly lower than the observed values in the Northeast, Northwest, and Southwest (Figure 1A). Adjusting for unrepresentative samples with raking and resampling and rerunning the binomial regression model had a minor effect on mask-wearing estimates, only exhibiting a slight decrease compared to the model without raking (Figure 1B). Correcting for survey biases using vaccination data in the binomial regression model run on raked survey responses systematically decreased masking proportions, as expected and denoted by increased residuals (Figure 1C). (A map showing the spatial distribution of these biases is given in Multimedia Appendix 1, Figure S19.) We refer to estimates from the model in Figure 1C as debiased or bias-corrected for the remainder of the paper. Our results reinforce that behavioral surveillance should be conducted carefully to limit bias initially.

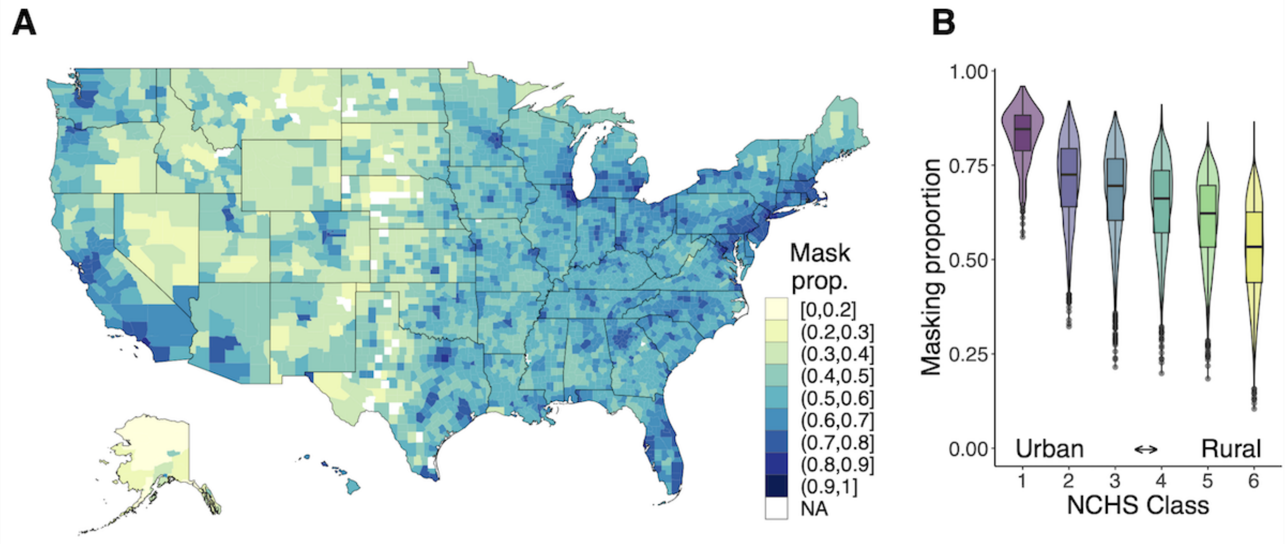
Figure 1. Visualization of spatially heterogeneous data-processing effects. (A) Residuals following the binomial regression model. (B) Residuals following the binomial regression model with raking/sample rebalancing. (C) Residuals following the binomial regression model with raking/sample rebalancing and an offset for bias. Residuals are defined as the difference between the modeled and the observed masking estimates at each analysis stage, where negative values indicate model estimates were higher than observed values and positive residuals indicate model estimates were lower than observed values. All maps are shown for February 2021. N/A: not applicable. See Multimedia Appendix 2 for a high-resolution image.



Masking Behavior Exhibits Spatial and Temporal Heterogeneity and Is Positively Associated With Population Density

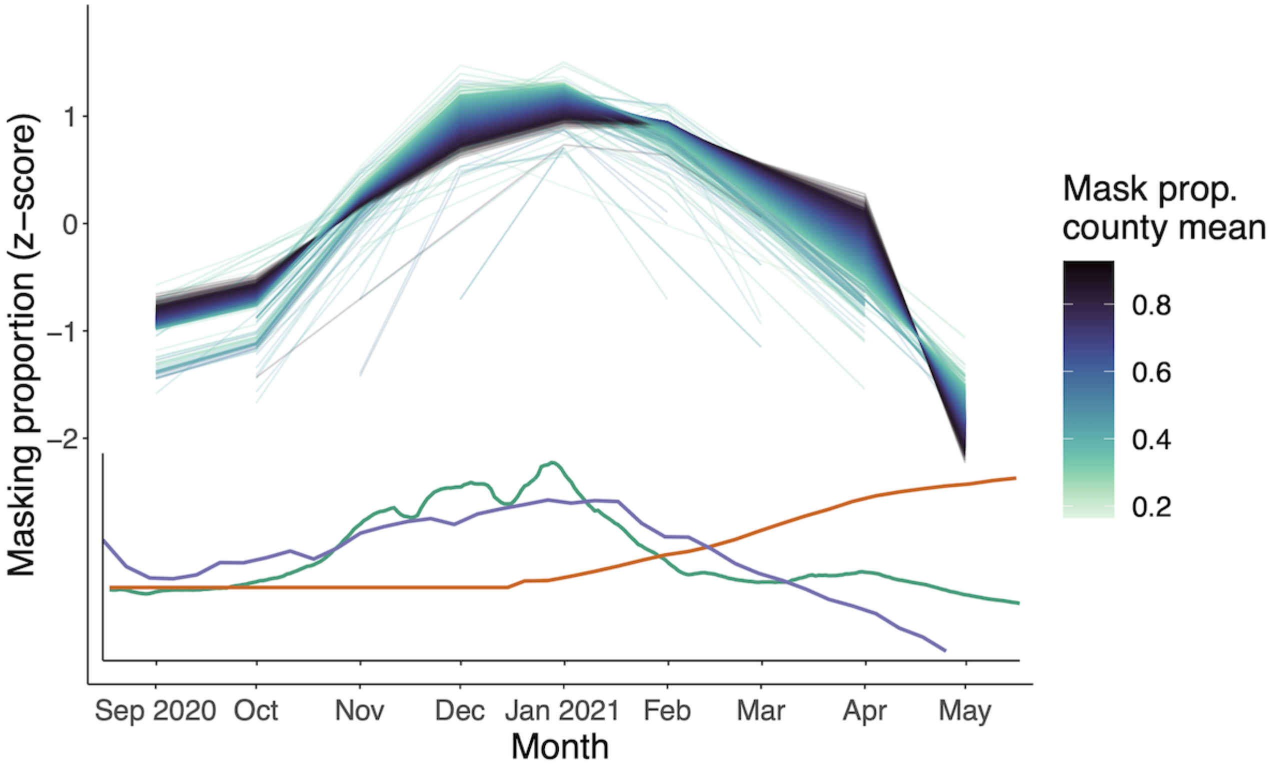
Using bias-corrected masking proportions from the CTIS, we found that masking behavior was spatially heterogeneous over all months (Moran's I between 0.68 and 0.70 for all months, Figures 2A and S5 in Multimedia Appendix 1). Bias-corrected masking proportions ranged from 0.11 to 0.96 and varied substantially within states, emphasizing the importance of analyzing masking behavior at finer scales than the state or Health and Human Services (HHS) region level. Masking proportions were closely linked to population density over the survey period: urban counties tended to have higher masking proportions than rural counties (Figure 2B). Although masking proportions ranged quite a bit within NCHS urban-rural classifications, all differences between NCHS classes were significant (Kruskal-Wallis and pairwise Wilcox test, n=27,842, all *P*<.001). Over all counties and survey months, the median fitted masking proportion in urban counties exceeded 0.8, while the median fitted masking proportion in the most rural counties was below 0.6.

Figure 2. Bias-corrected masking behavior is spatially heterogeneous and higher in urban areas. (A) Map of bias-corrected masking behavior in October 2020 reveals high spatial heterogeneity. Masking proportions vary substantially even within a single state. Spatial heterogeneity does not notably vary over time (Multimedia Appendix 1, Figure S5). A selection of other months in the study period are shown in Multimedia Appendix 1 (Figures S6-S8). (B) Breakdown of county masking proportions over all survey months by the NCHS urban-rural classification. A direct relationship between the median masking proportion and population density is observed. N/A: not applicable; NCHS: National Center for Health Statistics. See Multimedia Appendix 3 for a high-resolution image.



Masking behavior not only varied geographically but also temporally. Peak masking behavior was observed in January 2021, while the lowest masking proportions were observed in May 2021 (Figure 3). Counties with higher mean masking proportions fluctuated less than counties with lower mean masking proportions from September to April but experienced the largest differences from their mean values in May 2021. For context, we highlight that this decrease coincides with increasing proportions of vaccinated individuals in the United States (Figure 3, [33]), declining new infections [44], and decreasing reported worry about severe illness due to COVID-19 from the CTIS [27]. The policy context during this time was also shifting: On April 27, 2021, the CDC announced that fully vaccinated individuals no longer needed to wear masks outdoors [45], and on May 13, 2021, it announced that fully vaccinated individuals no longer had to wear masks indoors either [46]. Meanwhile, 49% of counties that ever had a mask mandate lifted it before May 1, 2021 (Multimedia Appendix 1, Figure S9). These announcements coincided with the observed decrease in masking in these months. Together, these analyses underscore the importance of tracking and analyzing mask wearing at fine spatial and across long temporal scales: further spatial and temporal aggregation of these data would have missed key heterogeneity previously not quantified and prevented future work from investigating the connection between policy and behavior change at appropriate granularity.

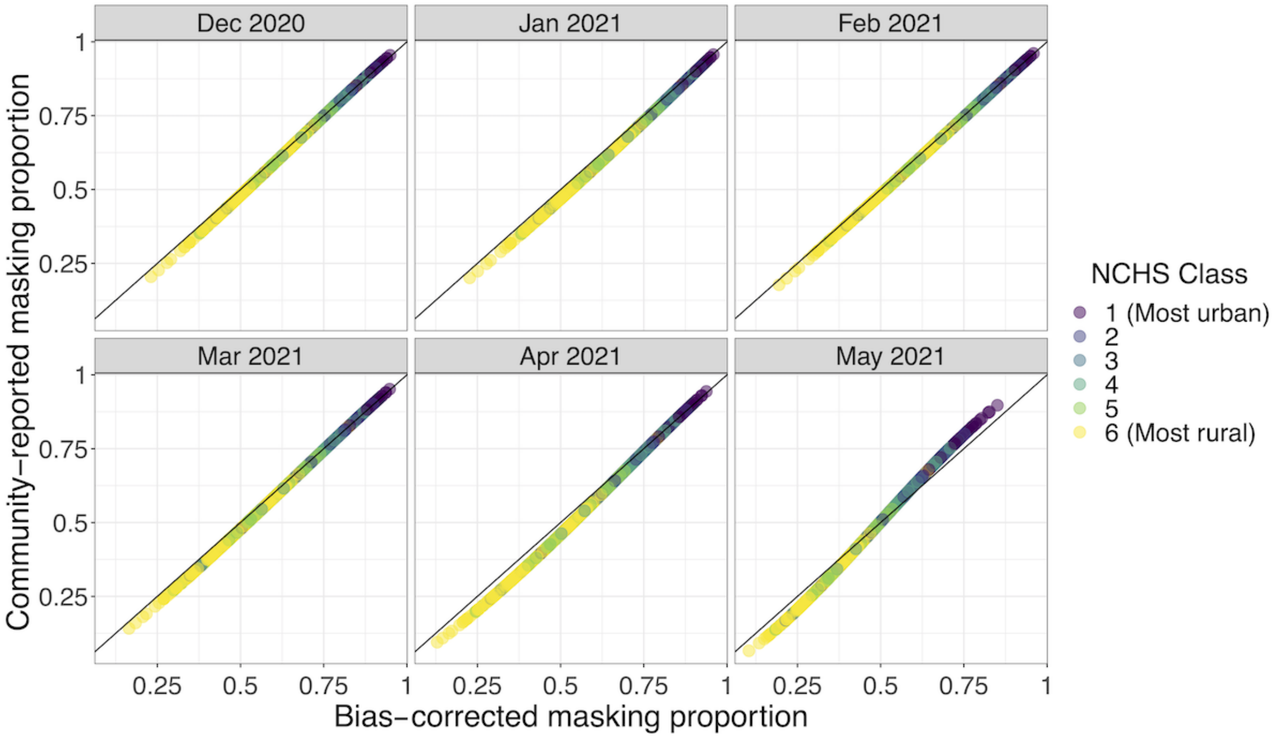
Figure 3. Bias-corrected masking behavior peaked in the winter of 2020-2021 and fell in the spring of 2021, mirroring new cases and increasing vaccinations. Top curves show the time series of the z-score of bias-corrected masking proportions for each county colored by the average masking proportion across the survey period. The inset plot shows z-scores of the 7-day rolling average of new cases (green), the proportion of individuals vaccinated nationally (orange), and the reported worry about severe illness from COVID-19 in CTIS respondents (purple). Z-scores are based on the mean and SD of each county's masking estimates over the survey period. CTIS: COVID-19 Trends and Impact Survey. See Multimedia Appendix 4 for a high-resolution image.



Community-Reported Masking Levels Are a Good Predictor of Bias-Corrected Self-Reported Estimates

Bias-corrected masking proportions were well approximated by modeled estimates of community-reported masking (Figure 4). The difference between the 2 mask-wearing proportion estimates ranged from –6% to 5% and became more apparent in April and May 2021, particularly in rural areas. In May 2021, though, community estimates in urban areas tended to overestimate bias-corrected individual masking estimates. This result is quantitatively affected by influential observations but is qualitatively robust (Multimedia Appendix 1, Figure S10). These results suggest that surveying participants about community masking may give less biased responses than asking individuals to report their own masking behavior, potentially reducing social desirability bias and capturing parts of the population that may be otherwise less likely to respond to the survey.

Figure 4. Community-reported masking gives a good estimate of bias-corrected self-reported masking. Community-reported masking refers to the CTIS question where individuals report how many people in their community are masking, which may decrease nonresponse and social desirability biases, compared to asking individuals to self-report their masking behavior. Point color denotes urban-rural classes. Comparisons of individual- and community-reported estimates at different analysis stages are shown in Multimedia Appendix 1 (Figures S11 and S12). CTIS: COVID-19 Trends and Impact Survey. See Multimedia Appendix 5 for a high-resolution image.



Discussion

Principal Findings

Despite the widespread adoption of face mask-wearing at points during the COVID-19 pandemic in the United States, the true prevalence of this behavior across temporal and spatial scales is largely unknown. Data on mask wearing have been collected through surveys, at varying spatiotemporal resolutions and with potentially varying survey biases (eg, [15-17]). Here, we characterized mask-wearing behavior across the United States using self-reported masking data from a large national online survey. We used Bayesian binomial regression models to remediate issues of small sample size, performed raking/sample balancing to address unrepresentative survey samples, and corrected for additional response biases using measurable bias in vaccination data. We observed substantial spatial heterogeneity in the masking behavior across urban versus rural counties, with some temporal changes in mean masking estimates at the county-month level, most notably a steep decline in masking in May 2021. We found that community-reported masking responses well approximated our bias-corrected masking estimates. Other work adds to this validation: Similar spatial heterogeneity was found in 2 other surveys, with some overlapping periods (Multimedia Appendix 1, Figures S1-S3, [17,47]), and our debiased estimates generally agree with those recorded in observational studies, including higher levels of masking in urban areas (Multimedia Appendix 1, Table S1, [14,28]). Our results reveal the landscape of the masking behavior across 3 distinct phases of the pandemic (presurge, surge during winter 2020-2021, and postsurge during the initial COVID-19 vaccination rollout). Our work also highlights the critical role that behavioral big data can play in the pandemic response, if such data are used with caution.

Implications

Our findings have several implications for the fields of infectious disease epidemiology and public health policy. We identified high spatial and moderate temporal heterogeneity in masking behavior at the county-level—patterns that are obscured if data are aggregated to the state or HHS region level. Contrary to our expectations, this level of spatial variability around the mean is consistent over time. Consequently, disease models should account for spatial variability in masking behavior but may only need to consider changes in masking dynamics over longer temporal scales. The high spatial heterogeneity we found in masking behavior also highlights the need for diverse and targeted public health approaches across the country rather than a single national program. Guidance set at the state level without regard for differences in local conditions may miss early opportunities to control disease spread, prematurely enforce public health restrictions, and contribute to fatigue with public health restrictions. Thus, we advocate for local behavioral data collection and geographically targeted public health policy for optimized resource use and efficient disease suppression.

Although county-level mask-wearing behavior varied across months, we observed little heterogeneity across counties in these temporal trends. The observed changes in masking behavior roughly correspond to national trends in new cases in the United States and self-reported worry about severe disease, as reported in the CTIS, though we did not determine causality or examine this relationship at the individual or county level. Because we modeled county-level averages, this observed correlation could be driven by a specific demographic group or subset of individuals modifying their masking behavior, rather than a uniform change in average mask uptake in a county's population. The sharp decrease in masking in May 2021 is contemporaneous with many states lifting mask mandates (Multimedia Appendix 1, Figure S9) and an announcement from the CDC that vaccinated individuals no longer have to wear masks outdoors (April 27, 2021 [45]) or indoors (May 13, 2021 [46]). It is plausible that these policy changes could have impacted masking behavior, both in vaccinated and in unvaccinated individuals, even though the change in CDC guidance did not apply to unvaccinated individuals [42,48]. More work is needed to explore the potential differences in mask wearing between vaccinated and unvaccinated individuals. Additional research could also focus on quantifying the impact of social norms on individuals' masking behavior at fine spatiotemporal resolution in the United States.

Recent work has highlighted the potential for big data sources to provide a measurement of spatially disaggregated social phenomena (eg, [49]), while other research points to the challenges of inferring high-quality estimates of behavior from such high-volume data sources [19]. In our work, we sought to steer away from "big data hubris" [50] by applying rigorous statistical methods to manage concerns about representativeness and bias and by conducting an internal validation of our model-based estimates [51]. In particular, we addressed representativeness by age, sex, and education to capture sociodemographic response bias. Motivated by an association between COVID-19–preventative behaviors [39], we debiased our masking estimates based on vaccination data to address additional nonresponse bias and social desirability bias. Finally, we internally validated our population-scale masking estimates of self-reported behaviors with responses of community behaviors, which may be subject to less nonresponse and social desirability bias compared to self-reporting questions [52]. We found that community-reported masking estimates agreed closely with bias-corrected self-reported masking behavior, highlighting that surveying participants about community behavior may be an avenue to reduce survey bias. We note, however, that this finding may not apply in all settings; self-reported masking behavior on a university campus closely matched observed masking levels, and questions about community masking were less accurate [53]. Although these implications for analysis of surveys on human behavior may not apply universally, similar results have been found in other infectious disease applications, including disease surveillance (using the CTIS data [27,54]) and early outbreak detection in social networks [55,56]. Our results further emphasize the promise of human social sensing going forward to make big data sources more meaningful [57].

Limitations

Nevertheless, our approach has some limitations that are important to consider. We were unable to deal with all representation or response biases, including the exclusion of individuals under 18 years of age; a lack of representativeness due to factors other than age, sex, and education; recall bias; dishonest responses; and other characteristics that may be predictive of nonresponse or social desirability bias, such as political leanings or belief in COVID-19 conspiracies [27]. Likewise, we could not account for how individuals with Facebook accounts may engage differently in COVID-19–preventive behaviors than non-Facebook users. It is unclear whether these biases that are unaccounted for would have a systematic or random effect on our results. However, we do expect their impact to be relatively small, particularly as our ground-truth–based debiasing approach adjusts masking estimates regardless of the source of nonresponse bias, and as supported by our community-reported masking analysis (since the community question captures some of the non-Facebook user population). We assumed that self-reported mask wearing is biased in the same magnitude and direction as the self-reported COVID-19 vaccination status—an assumption that should be tested in future research. Our approach also does not resolve issues of a small sample size; for example, the association we found between bias-corrected self-reported masking and community-reported masking is stronger between modeled estimates than between raw means. Although we attempted to correct for bias in our mask-wearing estimates, the point estimates for these values should be interpreted with caution. The goal of our work is not to produce point estimates of county-level mask-wearing behavior but instead to take advantage of the CTIS survey design and characterize the relative trends in mask-wearing behavior between and across counties. We advocate for additional observational studies with experimental designs that would allow for direct estimation of these quantities to improve behavioral surveillance estimates.

Conclusion

In summary, we produced the first accurate high-resolution spatiotemporal estimates of face mask wearing in the United States for the period from September 2020 through May 2021. Our work reveals that masking behavior is highly variable across the United States, suggesting that a one-size-fits-all approach to increasing mask-wearing behavior is likely to be ineffective. Instead, we identified regions of the country with higher and lower masking levels. These differences should be investigated going forward as public health organizations consider how to more effectively target these low-masking regions. For example, these communities may be more susceptible to mis- and disinformation regarding mitigation behaviors, which must be strategically confronted. Furthermore, this variability in behavior demonstrates the need to develop infectious disease dynamics models to analyze and predict how spatiotemporal trends in disease are affected by changes in human behavior, such as vaccination, contact patterns, and face mask wearing. Our analyses also address issues of survey bias, with the takeaway that, in the future, we should invest in a robust survey infrastructure that can recruit large representative samples with minimal bias, including using certain representative respondents as human social sensors to report on their communities.

Acknowledgments

The authors thank the Carnegie Mellon University (CMU) Delphi team for sharing the US COVID-19 Trends and Impact Survey (CTIS) data openly and freely and Alex Reinhart for his feedback on this work. We also thank the Outbreaks Near Me team at Boston Children’s Hospital and Momentive for data sharing and Benjamin Rader and John Brownstein for their feedback.

The research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM123007. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

Data aggregated to the county-month scale and all code to analyze these data will be made available on GitHub [58]. Individual survey responses cannot be shared by the authors, but researchers can refer to Ref. [59] if they would like to enter an agreement for data usage with the Carnegie Mellon University (CMU) Delphi team.

Authors' Contributions

JCT performed analyses, interpreted results, and drafted and edited the manuscript. ZS designed and guided analyses and edited the manuscript. SB designed the study, guided the analysis, interpreted results, and edited the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials.

[PDF File (Adobe PDF File), 29021 KB - publichealth_v9i1e42128_app1.pdf]

Multimedia Appendix 2

Visualization of spatially heterogeneous data-processing effects. (A) Residuals following the binomial regression model. (B) Residuals following the binomial regression model with raking/sample rebalancing. (C) Residuals following the binomial regression model with raking/sample rebalancing and an offset for bias. Residuals are defined as the difference between the modeled and the observed masking estimates at each analysis stage, where negative values indicate model estimates were higher than observed values and positive residuals indicate model estimates were lower than observed values. All maps are shown for February 2021. N/A: not applicable.

[PNG File , 4292 KB - publichealth_v9i1e42128_app2.png]

Multimedia Appendix 3

Bias-corrected masking behavior is spatially heterogeneous and higher in urban areas. (A) Map of bias-corrected masking behavior in October 2020 reveals high spatial heterogeneity. Masking proportions vary substantially even within a single state. Spatial heterogeneity does not notably vary over time (Multimedia Appendix 1, Figure S5). A selection of other months in the study period are shown in Multimedia Appendix 1 (Figures S6-S8). (B) Breakdown of county masking proportions over all survey months by the NCHS urban-rural classification. A direct relationship between the median masking proportion and population density is observed. N/A: not applicable; NCHS: National Center for Health Statistics.

[PNG File , 3213 KB - publichealth_v9i1e42128_app3.png]

Multimedia Appendix 4

Bias-corrected masking behavior peaked in the winter of 2020-2021 and fell in the spring of 2021, mirroring new cases and increasing vaccinations. Top curves show the time series of the z-score of bias-corrected masking proportions for each county colored by the average masking proportion across the survey period. The inset plot shows z-scores of the 7-day rolling average of new cases (green), the proportion of individuals vaccinated nationally (orange), and the reported worry about severe illness from COVID-19 in CTIS respondents (purple). Z-scores are based on the mean and SD of each county’s masking estimates over the survey period. CTIS: COVID-19 Trends and Impact Survey.

[PNG File , 3253 KB - publichealth_v9i1e42128_app4.png]

Multimedia Appendix 5

Community-reported masking gives a good estimate of bias-corrected self-reported masking. Community-reported masking refers to the CTIS question where individuals report how many people in their community are masking, which may decrease nonresponse

and social desirability biases, compared to asking individuals to self-report their masking behavior. Point color denotes urban-rural classes. Comparisons of individual- and community-reported estimates at different analysis stages are shown in Multimedia Appendix 1 (Figures S11 and S12). CTIS: COVID-19 Trends and Impact Survey.
[PNG File , 3105 KB - publichealth_v9i1e42128_app5.png]

References

1. Funk S, Salathé M, Jansen VAA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J R Soc Interface 2010 Sep 06;7(50):1247-1256 [FREE Full text] [doi: 10.1098/rsif.2010.0142] [Medline: 20504800]
2. Ferguson N. Capturing human behaviour. Nature 2007 Apr 12;446(7137):733. [doi: 10.1038/446733a] [Medline: 17429381]
3. Del Valle S, Hethcote H, Hyman J, Castillo-Chavez C. Effects of behavioral changes in a smallpox attack model. Math Biosci 2005 Jun;195(2):228-251. [doi: 10.1016/j.mbs.2005.03.006] [Medline: 15913667]
4. Worby CJ, Chang H. Face mask use in the general population and optimal resource allocation during the COVID-19 pandemic. Nat Commun 2020 Aug 13;11(1):4049 [FREE Full text] [doi: 10.1038/s41467-020-17922-x] [Medline: 32792562]
5. Bayham J, Kuminoff NV, Gunn Q, Fenichel EP. Measured voluntary avoidance behaviour during the 2009 A/H1N1 epidemic. Proc Biol Sci 2015 Nov 07;282(1818):20150814 [FREE Full text] [doi: 10.1098/rspb.2015.0814] [Medline: 26511046]
6. Verelst F, Willem L, Beutels P. Behavioural change models for infectious disease transmission: a systematic review (2010-2015). J R Soc Interface 2016 Dec;13(125):20160820 [FREE Full text] [doi: 10.1098/rsif.2016.0820] [Medline: 28003528]
7. Ueki H, Furusawa Y, Iwatsuki-Horimoto K, Imai M, Kabata H, Nishimura H, et al. Effectiveness of face masks in preventing airborne transmission of SARS-CoV-2. mSphere 2020 Oct 28;5(5):e00637-20. [doi: 10.1128/msphere.00637-20]
8. Eikenberry SE, Mancuso M, Iboi E, Phan T, Eikenberry K, Kuang Y, et al. To mask or not to mask: modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infect Dis Model 2020;5:293-308 [FREE Full text] [doi: 10.1016/j.idm.2020.04.001] [Medline: 32355904]
9. Stutt ROJH, Retkute R, Bradley M, Gilligan CA, Colvin J. A modelling framework to assess the likely effectiveness of facemasks in combination with 'lock-down' in managing the COVID-19 pandemic. Proc Math Phys Eng Sci 2020 Jun 10;476(2238):20200376 [FREE Full text] [doi: 10.1098/rspa.2020.0376] [Medline: 32821237]
10. Abaluck J, Kwong LH, Styczynski A, Haque A, Kabir MA, Bates-Jefferys E, et al. Impact of community masking on COVID-19: a cluster-randomized trial in Bangladesh. Science 2022 Jan 14;375(6577):eabi9069 [FREE Full text] [doi: 10.1126/science.abi9069] [Medline: 34855513]
11. Fisher KA, Barile JP, Guerin RJ, Vanden Esschert KL, Jeffers A, Tian LH, et al. Factors associated with cloth face covering use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):933-937 [FREE Full text] [doi: 10.15585/mmwr.mm6928e3] [Medline: 32673303]
12. Joo H, Miller GF, Sunshine G, Gakh M, Pike J, Havers FP, et al. Decline in COVID-19 hospitalization growth rates associated with statewide mask mandates - 10 states, March-October 2020. MMWR Morb Mortal Wkly Rep 2021 Feb 12;70(6):212-216 [FREE Full text] [doi: 10.15585/mmwr.mm7006e2] [Medline: 33571176]
13. Guy GP, Lee FC, Sunshine G, McCord R, Howard-Williams M, Kompaniyets L, CDC COVID-19 Response Team, Mitigation Policy Analysis Unit, CDC Public Health Law Program. Association of state-issued mask mandates and allowing on-premises restaurant dining with county-level COVID-19 case and death growth rates - United States, March 1-December 31, 2020. MMWR Morb Mortal Wkly Rep 2021 Mar 12;70(10):350-354 [FREE Full text] [doi: 10.15585/mmwr.mm7010e3] [Medline: 33705364]
14. Haischer MH, Beilfuss R, Hart MR, Opielinski L, Wrucke D, Zirgaitis G, et al. Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic. PLoS One 2020 Oct 15;15(10):e0240785 [FREE Full text] [doi: 10.1371/journal.pone.0240785] [Medline: 33057375]
15. Kapteyn A, Angrisani M, Bennett D, de BW, Darling J, Gutsche T. Tracking the effect of the COVID-19 pandemic on the lives of American households. Survey Res Methods 2020 Jun;14(2):186. [doi: 10.18148/srm/2020.v14i2.7737]
16. US Census Bureau. Measuring Household Experiences during the Coronavirus Pandemic. 2022. URL: https://www.census.gov/householdpulsedata [accessed 2023-01-06]
17. Katz J, Sanger-Katz M, Quealy K. A Detailed Map of Who Is Wearing Masks in the U.S. 2020 Jul 17. URL: https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html [accessed 2023-01-06]
18. Jakubowski A, Egger D, Nekesa C, Lowe L, Walker M, Miguel E. Self-reported vs directly observed face mask use in Kenya. JAMA Netw Open 2021 Jul 01;4(7):e2118830 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.18830] [Medline: 34328505]
19. Bradley VC, Kuriwaki S, Isakov M, Sejdinovic D, Meng X, Flaxman S. Unrepresentative big surveys significantly overestimated US vaccine uptake. Nature 2021 Dec;600(7890):695-700 [FREE Full text] [doi: 10.1038/s41586-021-04198-4] [Medline: 34880504]
20. Favier C, Schmit D, Müller-Graf CDM, Cazelles B, Degallier N, Mondet B, et al. Influence of spatial heterogeneity on an emerging infectious disease: the case of dengue epidemics. Proc R Soc Biol Sci 2005 Jun 07;272(1568):1171-1177 [FREE Full text] [doi: 10.1098/rspb.2004.3020] [Medline: 16024379]

21. Merler S, Ajelli M. The role of population heterogeneity and human mobility in the spread of pandemic influenza. Proc R Soc Biol Sci 2010 Feb 22;277(1681):557-565 [FREE Full text] [doi: 10.1098/rspb.2009.1605] [Medline: 19864279]
22. Susswein Z, Valdano E, Brett T, Rohani P, Colizza V, Bansal S. Ignoring spatial heterogeneity in drivers of SARS-CoV-2 transmission in the US will impede sustained elimination. medRxiv Aug 10 [FREE Full text] [doi: 10.1101/2021.08.09.21261807] [Medline: 34401885]
23. Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 2006 Apr 21;312(5772):447-451. [doi: 10.1126/science.1125237] [Medline: 16574822]
24. Cuadros DF, Xiao Y, Mukandavire Z, Correa-Agudelo E, Hernández A, Kim H, et al. Spatiotemporal transmission dynamics of the COVID-19 pandemic and its impact on critical healthcare capacity. Health Place 2020 Jul;64:102404 [FREE Full text] [doi: 10.1016/j.healthplace.2020.102404] [Medline: 32736312]
25. Truelove SA, Graham M, Moss WJ, Metcalf CJE, Ferrari MJ, Lessler J. Characterizing the impact of spatial clustering of susceptibility for measles elimination. Vaccine 2019 Jan 29;37(5):732-741 [FREE Full text] [doi: 10.1016/j.vaccine.2018.12.012] [Medline: 30579756]
26. Masters NB, Eisenberg MC, Delamater PL, Kay M, Boulton ML, Zelner J. Fine-scale spatial clustering of measles nonvaccination that increases outbreak potential is obscured by aggregated reporting data. Proc Natl Acad Sci U S A 2020 Nov 10;117(45):28506-28514. [doi: 10.1073/pnas.2011529117] [Medline: 33106403]
27. Salomon JA, Reinhart A, Bilinski A, Chua EJ, La Motte-Kerr W, Rönn MM, et al. The US COVID-19 Trends and Impact Survey: continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination. Proc Natl Acad Sci U S A 2021 Dec 21;118(51):e2111454118 [FREE Full text] [doi: 10.1073/pnas.2111454118] [Medline: 34903656]
28. Callaghan T, Lueck JA, Trujillo KL, Ferdinand AO. Rural and urban differences in COVID-19 prevention behaviors. J Rural Health 2021 Mar;37(2):287-295 [FREE Full text] [doi: 10.1111/jrh.12556] [Medline: 33619836]
29. Bürkner P. brms: an R package for Bayesian multilevel models using Stan. J Stat Softw 2017;80(1):1-28. [doi: 10.18637/jss.v080.i01]
30. Gabry J, Češnovar R. CmdStanR. URL: https://mc-stan.org/cmdstanr/index.html [accessed 2023-01-06]
31. Vehtari A, Simpson D, Gelman A, Yao Y, Gabry J. Pareto smoothed importance sampling. arXivstat Feb.
32. Pasek J. Anesrake: ANES Raking Implementation. 2018. URL: https://CRAN.R-project.org/package=anesrake [accessed 2023-01-06]
33. Tiu A, Susswein Z, Merritt A, Bansal S. Characterizing the spatiotemporal heterogeneity of the COVID-19 vaccination landscape. Am J Epidemiol 2022 Sep 28;191(10):1792-1802 [FREE Full text] [doi: 10.1093/aje/kwac080] [Medline: 35475891]
34. Tiu A, Bansal S. COVID-19 Vaccination Tracking. 2021. URL: https://github.com/bansallab/vaccinetracking [accessed 2023-01-06]
35. Reinhart A, Tibshirani R. Big data, big problems: responding to "Are we there yet?". arXivstat Sep.
36. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using Lme4. J Stat Softw 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
37. Lamstein A, Johnson B, Trulia, Inc.. choroplethr: Simplify the Creation of Choropleth Maps in R. 2022. URL: https://CRAN.R-project.org/package=choroplethr [accessed 2023-01-06]
38. Ingram DD, Franco SJ. 2013 NCHS Urban-Rural Classification Scheme for counties. Vital Health Stat 2 2014 Apr(166):1-73 [FREE Full text] [Medline: 24776070]
39. Latkin CA, Dayton L, Yi G, Colon B, Kong X. Mask usage, social distancing, racial, and gender correlates of COVID-19 vaccine intentions among adults in the US. PLoS One 2021 Feb 16;16(2):e0246970 [FREE Full text] [doi: 10.1371/journal.pone.0246970] [Medline: 33592035]
40. Lam CN, Kaplan C, Saluja S. Relationship between mask wearing, testing, and vaccine willingness among Los Angeles County adults during the peak of the COVID-19 pandemic. Transl Behav Med 2022 Mar 17;12(3):480-485 [FREE Full text] [doi: 10.1093/tbm/ibab150] [Medline: 34865166]
41. Moore JS, Virk H, Summers JA. Correlation of population factors, compliance with masking and social distance, vaccination, and COVID-19 infection in Central Appalachia. South Med J 2022 Jul;115(7):420-421 [FREE Full text] [doi: 10.14423/SMJ.0000000000001410] [Medline: 35777747]
42. Calamari LE, Tjaden AH, Edelstein SL, Weintraub WS, Santos R, Gibbs M, COVID-19 Community Research Partnership Study Group. Self-reported mask use among persons with or without SARS CoV-2 vaccination - United States, December 2020-August 2021. Prev Med Rep 2022 Aug;28:101857 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101857] [Medline: 35706687]
43. Floyd CJ, Joachim GE, Boulton ML, Zelner J, Wagner AL. COVID-19 vaccination and mask wearing behaviors in the United States, August 2020 - June 2021. Expert Rev Vaccines 2022 Oct;21(10):1487-1493. [doi: 10.1080/14760584.2022.2104251] [Medline: 35856246]
44. The New York Times. Coronavirus (Covid-19) Data in the United States. 2021. URL: https://github.com/nytimes/covid-19-data [accessed 2023-01-06]

45. Sun LH. CDC Says Fully Vaccinated Americans Can Go without Masks Outdoors, Except in Crowded Settings. 2021. URL: https://www.washingtonpost.com/health/2021/04/27/cdc-guidance-masks-outdoors/ [accessed 2023-01-06]
46. Abutaleb Y, McGinley L. CDC Says Fully Vaccinated Americans No Longer Need Masks Indoors or Outdoors in Many Cases. 2021. URL: https://www.washingtonpost.com/health/2021/05/13/cdc-says-fully-vaccinated-americans-no-longer-need-masks-indoors-or-outdoors-most-cases/ [accessed 2023-01-06]
47. Rader B, White LF, Burns MR, Chen J, Brilliant J, Cohen J, et al. Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study. Lancet Digital Health 2021 Mar;3(3):e148-e157 [FREE Full text] [doi: 10.1016/S2589-7500(20)30293-4] [Medline: 33483277]
48. Blake A. How Many Unvaccinated People Will Stop Wearing Masks Now?. 2021. URL: https://www.washingtonpost.com/politics/2021/05/14/how-many-unvaccinated-people-will-quietly-stop-wearing-masks-now/ [accessed 2023-01-06]
49. Chetty R, Jackson MO, Kuchler T, Stroebel J, Hendren N, Fluegge RB, et al. Social capital I: measurement and associations with economic mobility. Nature 2022 Aug 01;608(7921):108-121 [FREE Full text] [doi: 10.1038/s41586-022-04996-4] [Medline: 35915342]
50. Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
51. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big data for infectious disease surveillance and modeling. J Infect Dis 2016 Dec 01;214(suppl_4):S375-S379 [FREE Full text] [doi: 10.1093/infdis/jiw400] [Medline: 28830113]
52. Bradburn NM, Sudman S, Wansink B. Asking Questions: The Definitive Guide to Questionnaire Design -- For Market Research, Political Polls, and Social and Health Questionnaires, 2nd, Revised Edition. San Francisco, CA: Jossey-Bass; 2004.
53. Adaryukov J, Grunevski S, Reed DD, Pleskac TJ. I'm wearing a mask, but are they?: perceptions of self-other differences in COVID-19 health behaviors. PLoS One 2022;17(6):e0269625 [FREE Full text] [doi: 10.1371/journal.pone.0269625] [Medline: 35666754]
54. Astley CM, Tuli G, Mc Cord KA, Cohn EL, Rader B, Varrelman TJ, et al. Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. Proc Natl Acad Sci U S A 2021 Dec 21;118(51):e2111455118 [FREE Full text] [doi: 10.1073/pnas.2111455118] [Medline: 34903657]
55. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. PLoS One 2010 Sep 15;5(9):e12948 [FREE Full text] [doi: 10.1371/journal.pone.0012948] [Medline: 20856792]
56. Garcia-Herranz M, Moro E, Cebrian M, Christakis NA, Fowler JH. Using friends as sensors to detect global-scale contagious outbreaks. PLoS One 2014 Apr 9;9(4):e92413 [FREE Full text] [doi: 10.1371/journal.pone.0092413] [Medline: 24718030]
57. Galesic M, Bruine de Bruin W, Dalege J, Feld SL, Kreuter F, Olsson H, et al. Human social sensing is an untapped resource for computational social science. Nature 2021 Jul 30;595(7866):214-222. [doi: 10.1038/s41586-021-03649-2] [Medline: 34194037]
58. Taube J, Bansal S. Characterizing Spatiotemporal Trends in Self-Reported Mask-Wearing Behavior in the United States. URL: https://github.com/bansallab/spatial_masking [accessed 2023-01-06]
59. Delphi Research Group at Carnegie Mellon University. Delphi Epidata API. URL: https://cmu-delphi.github.io/delphi-epidata/symptom-survey/data-access.html [accessed 2023-01-06]

Abbreviations

CDC: Centers for Disease Control and Prevention
CTIS: COVID-19 Trends and Impact Survey
HHS: Health and Human Services
NCHS: National Center for Health Statistics

Edited by T Sanchez, A Mavragani; submitted 23.08.22; peer-reviewed by S Fox, N Masters; comments to author 03.11.22; revised version received 22.11.22; accepted 16.12.22; published 06.03.23.

Please cite as:
Taube JC, Susswein Z, Bansal S
Spatiotemporal Trends in Self-Reported Mask-Wearing Behavior in the United States: Analysis of a Large Cross-sectional Survey
JMIR Public Health Surveill 2023;9:e42128
URL: https://publichealth.jmir.org/2023/1/e42128
doi:10.2196/42128
PMID:36877548

Original Paper

Lifestyle Behavior Changes and Associated Risk Factors During the COVID-19 Pandemic: Results from the Canadian COVIDiet Online Cohort Study

Anne-Julie Tessier[1,2], RD, PhD; Audrey Moyen[1], BSc, RD; Claire Lawson[1], MSc; Aviva Ilysse Rappaport[1], RD, MSc; Hiba Yousif[1], RD, MSc; Chloé Fleurent-Grégoire[1], BSc, RD; Sophie Lalonde-Bester[1], BSc, RD; Anne-Sophie Brazeau[1], RD, PhD; Stéphanie Chevalier[1,2], RD, PhD

[1]School of Human Nutrition, Faculty of Agriculture and Environmental Sciences, McGill University, Sainte-Anne-de-Bellevue, QC, Canada
[2]Research Institute of the McGill University Health Centre, Montréal, QC, Canada

Corresponding Author:
Anne-Julie Tessier, RD, PhD
School of Human Nutrition
Faculty of Agriculture and Environmental Sciences
McGill University
21111 Lakeshore
Sainte-Anne-de-Bellevue, QC, H9X 3V9
Canada
Phone: 1 514 398 7925
Email: anne-julie.tessier@mail.mcgill.ca

Abstract

Background: The COVID-19 pandemic and related lockdowns have impacted lifestyle behaviors, including eating habits and physical activity; yet, few studies have identified the emerging patterns of such changes and associated risk factors.

Objective: This study aims to identify the patterns of weight and lifestyle behavior changes, and the potential risk factors, resulting from the pandemic in Canadian adults.

Methods: Analyses were conducted on 1609 adults (18-89 years old; n=1450, 90.1%, women; n=1316, 81.8%, White) of the Canadian COVIDiet study baseline data (May-December 2020). Self-reported current and prepandemic weight, physical activity, smoking status, perceived eating habits, alcohol intake, and sleep quality were collected through online questionnaires. Based on these 6 indicator variables, latent class analysis (LCA) was used to identify lifestyle behavior change patterns. Associations with potential risk factors, including age, gender, ethnicity, education, income, chronic diseases, body image perception, and changes in the stress level, living situation, and work arrangement, were examined with logistic regressions.

Results: Participants' mean BMI was 26.1 (SD 6.3) kg/m^2. Of the 1609 participants, 980 (60.9%) had a bachelor's degree or above. Since the pandemic, 563 (35%) had decreased income and 788 (49%) changed their work arrangement. Most participants reported unchanged weight, sleep quality, physical activity level, and smoking and alcohol consumption, yet 708 (44%) reported a perceived decrease in eating habit quality. From LCA, 2 classes of lifestyle behavior change emerged: healthy and less healthy (probability: 0.605 and 0.395, respectively; Bayesian information criterion [BIC]=15574, entropy=4.8). The healthy lifestyle behavior change group more frequently reported unchanged weight, sleep quality, smoking and alcohol intake, unchanged/improved eating habits, and increased physical activity. The less healthy lifestyle behavior change group reported significant weight gain, deteriorated eating habits and sleep quality, unchanged/increased alcohol intake and smoking, and decreased physical activity. Among risk factors, body image dissatisfaction (odds ratio [OR] 8.8, 95% CI 5.3-14.7), depression (OR 1.8, 95% CI 1.3-2.5), increased stress level (OR 3.4, 95% CI 2.0-5.8), and gender minority identity (OR 5.5, 95% CI 1.3-22.3) were associated with adopting less healthy behaviors in adjusted models.

Conclusions: The COVID-19 pandemic has appeared to have influenced lifestyle behaviors unfavorably in some but favorably in others. Body image perception, change in stress level, and gender identity are factors associated with behavior change patterns; whether these will sustain over time remains to be studied. Findings provide insights into developing strategies for supporting adults with poorer mental well-being in the postpandemic context and promoting healthful behaviors during future disease outbreaks.

Trial Registration: ClinicalTrials.gov NCT04407533; https://clinicaltrials.gov/ct2/show/NCT04407533

(JMIR Public Health Surveill 2023;9:e43786) doi:10.2196/43786

KEYWORDS

healthy lifestyle; behaviors; pandemic; COVID-19; public health; body image; gender; stress; risk factor; physical activity; public health; mental well-being; depression

Introduction

On March 11, 2020, the World Health Organization (WHO) declared COVID-19 (SARS-CoV-2) a global pandemic. Between March 12 and 22, all Canadian provinces had declared a public health emergency, and individuals were asked to stay at home and nonessential businesses were closed, including restaurants and gyms. Indoor and outdoor gatherings were also prohibited in large provinces to avoid spreading of the disease. Indeed, the pandemic imposed a sudden disruption in people’s usual routine. Although many adopted telework practice, others, such as frontline workers, were required to work on-site with additional protective equipment. In May 2020, measures became less stringent and nonessential businesses across Canada reopened gradually [1]. COVID-19 vaccines in adults were under clinical trial investigation, when a second wave hit in September 2020, leading to the enforcement of a second lockdown across provinces [1,2].

Confinement inevitably disturbs lifestyle behaviors, such as eating habits, physical activity level, and sleep, to varying degrees among individuals [3]. Although some observational studies and surveys have reported more home cooking, better diet quality, and increased physical activity levels [4,5], others have observed decreased physical activity, increased energy and alcohol intake, binge eating, reliance on take-out foods, and overall decreased diet quality [6-8]. The combination of poor diet quality, low physical activity, and greater alcohol intake, BMI, and smoking represents an important risk for severe COVID-19 [9], COVID-19 mortality [10], and chronic diseases during the postpandemic period [11]. Accumulating evidence shows that particular groups, such as women and younger individuals, were more susceptible to higher levels of stress and depression during the pandemic, possibly due to changes in home life among others [12-15]. These factors, combined with limited access to grocery shops [8], restaurants, and food banks, restricted budgets due to sudden unemployment, and changes in living arrangements may have affected how people modified their lifestyle during the pandemic [16,17]. Yet, it remains unclear which factors are major determinants of overall lifestyle pattern adoption in times of the pandemic, and no study has examined the lifestyle pattern change of adults in Canada.

Therefore, the objectives of this study were to characterize the potential impact of the COVID-19 pandemic on the lifestyle habits of Canadians and to identify patterns of weight and lifestyle behavior change and associated risk factors. We postulated that individuals experiencing greater levels of stress, isolation, changes in their living situation and work arrangement, and decreased income adopted less healthy lifestyle behaviors. Findings from this study may inform strategies to cope with the effects of current and potential future outbreaks in promoting and maintaining a healthy lifestyle.

Methods

Study Population

The web- and mobile-based Canadian COVIDiet cohort study was initiated on May 29, 2020, toward the end of the COVID-19 pandemic’s first wave of infections. The primary objective was to provide longitudinal data on the impact of the pandemic on eating behaviors, diet quality, and related lifestyle behaviors of Canadian adults. Based on the Canadian population estimated at 31 million adults and considering a 5% margin of error, 95% CI, and 50% response distribution [18], 400 participants were required per age strata (18-45, 45-65, ≥65 years). Following power allocation relative to population size [19], the target sample size was 1920 participants. Until December 5, 2020, 2959 adults from across Canada were recruited using social media advertisement, professional societies, and institutional email lists and from community centers, including food banks. Efforts were made to recruit across age groups, regions, and gender (men and women). Regions were delimited as Ontario, Québec, British Columbia, Prairies (Alberta, Manitoba, Saskatchewan), Atlantic provinces (Nova Scotia, New Brunswick, Newfoundland and Labrador, Prince Edward Island), and territories (Northwest Territories, Yukon, and Nunavut). Interested individuals were invited to contact the study coordinator by email or visit the study website for more information and for eligibility screening.

Eligible participants were 18 years of age or older, were living in Canada at the time of study entry, were able to read English or French, owned a smartphone or tablet, and had access to the internet. Individuals living in care homes or other institutional environments; hospitalized, pregnant, or breastfeeding; or with an active and uncontrolled acute disease that interfered with usual food intake were excluded. The COVIDiet study participants were followed every 3 months for 1 year to capture the impact of the pandemic across waves of infections and for 1 year thereafter to evaluate the longer-term effect. The current analysis was conducted on baseline (May-December 2020) and retrospective data from the baseline questionnaire. Individuals who did not fill out the baseline questionnaire (n=904), provided implausible responses (n=8), and had missing data on changes in weight and lifestyle behavior (eating habits, alcohol consumption, physical activity level, sleep quality and smoking; n=442) were excluded. The final sample included 1609 participants (Multimedia Appendix 1, Figure S1).

Ethical Considerations

Data were collected with online encrypted surveys using AirTable [20] and deidentified using generated ID codes.

Participants were eligible to win a CA $250 (US $182) gift card when they enrolled in the study; 1 draw occurred at every 500 participants enrolled. All participants provided online informed consent. This study was approved by the McGill University Research Ethics Board (20-04-064) and was registered on ClinicalTrials.gov (NCT04407533).

Questionnaires

Baseline Sociodemographics and Medical History

Participants were asked to provide demographic characteristics and changes in lifestyle habits since the pandemic. They were asked to report on age, self-identified gender (women, men, others), ethnicity (White, South Asian, Chinese, Black, Filipino, Latin American, Arabic, Southeast Asian, West Asian, Korean, Japanese, Indigenous, other), usual occupation (full-time, part-time, occasional/seasonal, self-employed, student, unemployed, caregiver, volunteer, retiree), highest or current level of education (grade 8 or lower, grade 9 or 10, high school diploma, vocational school, college diploma , university certificate, bachelor’s degree, graduate studies), usual annual household income (<CA $20,000/<US $14,560, CA $20,000-$50,000/US $14,560-$36,400, >CA $50,000 -$100,000/>US $36,400-$72,800, >CA $100,000-$150,000/>US $72,800-$109,199, >CA $150,000/>US $109,199), usual household/living situation (living with parents, grandparents, brothers and sisters, host family, roommates, partner, children, alone), province, area (urban, suburban, rural), height (m or ft), usual weight (kg or lb), perceived body image (satisfied, somewhat satisfied, not satisfied), living with a chronic disease (hypertension, osteoarthritis, mood and anxiety disorder, osteoporosis, diabetes, asthma, chronic obstructive pulmonary disease, ischemic heart disease, cancer, dementia, other), and time since diagnosis. Questions were retrieved from the Canadian Community Health Survey (CCHS) 2020 [21]. As a measure of restrictive public health measures burden, the pandemic wave at the time of study enrollment was used. Participants who completed the Rapid Response - Healthy Living questionnaire before the second wave was declared on September 23, 2020 (May 26-September 22, 2020) [22] were classified within the “lenient restrictions” category, while those who started the study after this cut-off (September 23-December 14, 2020) were classified within the “strict restrictions” category.

Other Lifestyle Behaviors

Participants were also asked to report on physical activity (≥ or <150 minutes/week [23]) and its intensity, based on the CCHS 2020 questionnaire. We queried on alcohol consumption (≤1 drink/week, 1-3 drinks/week, 4-8 drinks/week, 9-14 drinks/week, >15 drinks/week) using an adapted CCHS 2020 questionnaire to better capture short-term changes in alcohol intake since the start of the pandemic. Smoking (daily, occasionally, no) and sleeping time (hours) and quality (very good, fairly good, fairly bad, very bad) were also retrieved from the CCHS [24]. Participants answered according to the current situation and from before the pandemic. The current anxiety level was assessed using the General Anxiety Disorder-7 (GAD-7) [25], and depression was assessed using the 10-item Center for Epidemiologic Studies Depression Scale (CES-D-10; ≥10) [26]. We also queried on the self-reported change in the stress level compared to before the pandemic. If participants reported having 1 or more chronic diseases, questions were asked on medication and medication changes during the pandemic.

For each question, participants were given the option to respond with “I don’t know” or “I prefer not to answer.”

Weight and Lifestyle Behavior Changes

Baseline (during) versus prepandemic changes in weight and lifestyle factor status were categorized into 3 levels. For weight, a difference greater than +/–3% was considered an increase or a decrease in weight, respectively; participants with 0%-2.9% or insignificant weight change were categorized as having unchanged weight. For alcohol consumption, an upward change in 1 or more categories (<1 drink/week, 1-3 drinks/week, 4-8 drinks/week, 9-14 drinks/week, ≥15 drinks/week) was categorized as increased consumption, a downward change as decreased consumption, or no change as unchanged consumption. For physical activity level, a change in adherence to the WHO recommendations (yes or no) was categorized as increased (no to yes), decreased (yes to no), or unchanged (yes to yes or no to no). For eating habits, a perceived change was reported as improved, unchanged, or worsened. For sleep quality, an increase or decrease in 1 or more categories (very bad, fairly bad, fairly good, very good) was categorized as increased or decreased sleep quality, respectively, or otherwise categorized as unchanged. For smoking status (daily or occasionally/never), a change was categorized as increased (no to yes), decreased (yes to no), or unchanged (yes to yes, or no to no).

Statistical Analysis

Baseline characteristics were reported as means and SDs, and counts and percentages, as applicable. Latent class analysis (LCA) was used to identify lifestyle behavior change patterns based on 6 indicator variables: change in weight, alcohol consumption, physical activity, perceived eating habits, sleep quality, and smoking status. Models with 1, 2, and 3 classes were tested, and improvement in test statistics chi-square goodness of fit, Bayesian information criterion (BIC), Akaike information criterion (AIC), and entropy were used for model selection. Differences between participants’ characteristics and sociodemographics by class membership were examined using the chi-square for proportions, and the *t* test and Mann-Whitney U test for normally and nonnormally distributed data, respectively. Potential risk determinants of class membership were examined using logistic regressions. Model 1 was adjusted for age, gender, ethnicity, province, and burden of restrictive measure, and model 2 was adjusted for model 1 covariates and for body image perception, anxiety level, depression, chronic diseases, and changes in income, stress level, living situation, and work arrangement. Given the oversampling of women, gender-weighted sensitivity analysis was performed applying the Canadian 2019 Census gender proportions (women: 50.6%, men: 49.3%) [27] to improve the generalizability of results. Unweighted stratified analyses by age (< and ≥55 years), baseline BMI (< and ≥25 kg/m^2), and restrictive public health measure burden subgroups were conducted. P<.05 was considered statistically significant in all analyses. LCA and logistic regressions were performed using R version 4.1.2 (2021-11) in RStudio (R Foundation for Statistical Computing). This study's program is available on GitHub [28].

Results

Participants Characteristics

In total, 942 (46%) of 1609 participants were recruited between the end of the first wave of infections and the beginning of the second wave, that is, when lenient restrictive measures were in place. The other participants (1105/1609, 54%,) were recruited during the second wave of the pandemic, during strict restrictive measures. Table 1 shows prepandemic sociodemographic and lifestyle characteristics of participants. Participants mostly self-identified as women, were aged 18-89 years (mean 40.2, SD 15.5 years), and had a mean BMI of 26.1 (SD 6.3) kg/m^2. The majority were White, had at least a bachelor's degree, and reported having a household income > CA $50,000 (US $36,400) per year.

Table 1. Characteristics of participants, overall and by identified pattern of behavior change.

Characteristics	Overall (N=1609)	Healthy (n=974, 60.5%)	Less healthy (n=635, 39.5%)
Gender, n (%); P<.001[a]			
Men	115 (7.1)	86 (8.8)	29 (4.6)
Women	1461 (90.8)	873 (89.6)	588 (92.6)
Minority	22 (1.4)	6 (0.6)	16 (2.5)
Missing	11 (0.7)	9 (0.9)	2 (0.3)
Age (years) at baseline, mean (SD); range=18-89 years, P=.26[b]	40.2 (15.5)	40.5 (16.1)	39.6 (14.3)
Missing, n (%)[c]	20 (1.2)	N/A[d]	N/A
Ethnicity, n (%); P<.001[a]			
White	1307 (81.2)	796 (81.7)	511 (80.5)
Black	12 (0.7)	7 (0.7)	5 (0.8)
Hispanic	31 (1.9)	18 (1.8)	13 (2.0)
Asian	115 (7.1)	71 (7.3)	44 (6.9)
Indigenous	11 (0.7)	3 (0.3)	8 (1.3)
Multiethnic	106 (6.6)	63 (6.5)	43 (6.8)
Missing	27 (1.7)	16 (1.6)	11 (1.7)
Province or territory of residence, n (%); P<.001[a]			
Pacific region	186 (11.6)	106 (10.9)	80 (12.6)
Prairie provinces	310 (19.3)	173 (17.8)	137 (21.6)
Ontario	603 (37.5)	342 (35.1)	261 (41.1)
Québec	372 (23.1)	273 (28.0)	99 (15.6)
Atlantic region	120 (7.5)	67 (6.9)	53 (8.3)
Northern Territories	7 (0.4)	6 (0.6)	1 (0.2)
Missing	11 (0.7)	7 (0.7)	4 (0.6)
Residential area, n (%); P=.12[a]			
Urban	912 (56.7)	549 (56.4)	363 (57.2)
Suburban	445 (27.7)	284 (29.2)	161 (25.4)
Rural	247 (15.4)	138 (14.2)	109 (17.2)
Missing	5 (0.3)	3 (0.3)	2 (0.3)
BMI (kg/m^2), mean (SD); P<.001[b]	26.1 (6.3)	25.7 (6.2)	26.8 (6.5)
Missing, n (%)	77 (4.8)	N/A	N/A
Education, n (%); P=.96[a]			
Grade 9 or 10	5 (0.3)	4 (0.4)	1 (0.2)
High school diploma	197 (12.2)	116 (11.9)	81 (12.8)
Vocational school	14 (0.9)	8 (0.8)	6 (0.9)
Diploma from Collège d'enseignement general et professionnel (CEGEP) or community college	279 (17.3)	167 (17.1)	112 (17.6)
University certificate	91 (5.7)	54 (5.5)	37 (5.8)
Bachelor’s degree	599 (37.2)	370 (38.0)	229 (36.1)
Graduate studies (master’s degree or PhD)	381 (23.7)	231 (23.7)	150 (23.6)
Missing	43 (2.7)	24 (2.6)	19 (3.0)
Household income (CA $/US $ per year)[e], n (%); P=.08[a]			

Characteristics	Overall (N=1609)	Healthy (n=974, 60.5%)	Less healthy (n=635, 39.5%)
<20,000/<14,560	110 (6.8)	59 (6.1)	51 (8.0)
20,000-50,000/14,560-36,400	315 (19.6)	181 (18.6)	134 (21.1)
>50,000-100,000/>36,400-72,800	432 (26.8)	261 (26.8)	171 (26.9)
>100,000-150,000/>72,800-109,199	335 (20.8)	199 (20.4)	136 (21.4)
>150,000/>109,199	269 (16.7)	183 (18.8)	86 (13.5)
Missing	148 (9.2)	91 (9.3)	57 (9.0)
Employment status, n (%); P=.32[a]			
Full-time	734 (45.6)	436 (44.8)	298 (46.9)
Part-time	176 (10.9)	109 (11.2)	67 (10.6)
Student	103 (6.4)	63 (6.5)	40 (6.3)
Retired	145 (9.0)	99 (10.2)	46 (7.2)
Other (caregiver, volunteer, etc)	440 (27.3)	259 (26.6)	181 (28.5)
Missing	11 (0.7)	8 (0.8)	3 (0.5)
Chronic diseases, n (%); P=.08[a]			
0	1041 (64.7)	633 (65.0)	408 (64.3)
1	142 (8.8)	87 (8.9)	55 (8.7)
≥2	241 (15.0)	131 (13.4)	110 (17.3)
Missing	185 (11.5)	123 (12.6)	62 (9.8)
Smoker, n (%)	117 (7.3)	71 (7.3)	46 (7.2)
Alcohol consumption, n (%); P<.001[a]			
≥4 drinks/week	366 (22.7)	258 (26.5)	108 (17.0)
Restrictive measures at enrollment, n (%); P<.001[a]			
Strict	1033 (64.2)	584 (60.0)	449 (70.7)
Lenient	576 (35.8)	390 (40.0)	186 (29.3)
Use of dietary supplements, n (%); P=.02[a]			
Yes	596 (37.0)	340 (34.9)	256 (40.3)
No	999 (62.1)	627 (64.4)	372 (58.6)
Missing	14 (0.9)	7 (0.7)	7 (1.1)
Physical activity, n (%); P<.001[a]			
<150 minutes/week	592 (36.8)	581 (59.7)	436 (68.7)
≥150 minutes/week	1017 (63.2)	393 (40.3)	199 (31.3)
Since the pandemic, mean (SD)			
GAD-7[f], anxiety symptoms (/21); P<.001[b]	7.1 (5.3)	6.1 (4.8)	8.8 (5.5)
CES-D-10[g], depression symptoms (/30); P<.001[b]	11.1 (6.7)	9.3 (6.1)	13.7 (6.6)
Body image perception, n (%); P<.001[a]			
Satisfied	235 (14.6)	205 (21.0)	30 (4.7)
Somewhat satisfied	655 (40.7)	453 (46.5)	202 (31.8)
Not satisfied	707 (43.9)	306 (31.4)	401 (63.1)

Characteristics	Overall (N=1609)	Healthy (n=974, 60.5%)	Less healthy (n=635, 39.5%)
Missing	12 (0.7)	10 (1.0)	2 (0.3)

[a] P values are from chi-square tests unless otherwise indicated.

[b] P values are from 2-sided independent *t* tests.

[c] All percentages are unweighted.

[d] N/A: not applicable.

[e] CA $1.00=US $0.73.

[f] GAD-7: General Anxiety Disorder-7

[g] CES-D-10: 10-item Center for Epidemiologic Studies Depression Scale.

Self-Reported Changes in Weight, Lifestyle, and Living Factors

Since the pandemic, 35.1% (565/1609) participants had a decreased income, 49.0% (788/1609) changed their work arrangement, and 94.3% (1518/1609) kept the same living situation (1295/1609, 80.5%, were living with others, and 223/1609, 13.9%, were living alone). Regarding mental well-being, 52.2% (840/1609) participants reported a worsened mood and 63.8% (1026/1609) reported increased stress levels compared to before the pandemic. Among lifestyle factors, most of the population reported either no change or insignificant weight gain or loss and no change in alcohol intake, physical activity, smoking, and sleep quality, yet 44.1% (709/1609) reported a decreased perceived quality of eating habits (Table 2).

Table 2. Self-reported changes in sociodemographic status, weight, and lifestyle factors. All reported values are unweighted.

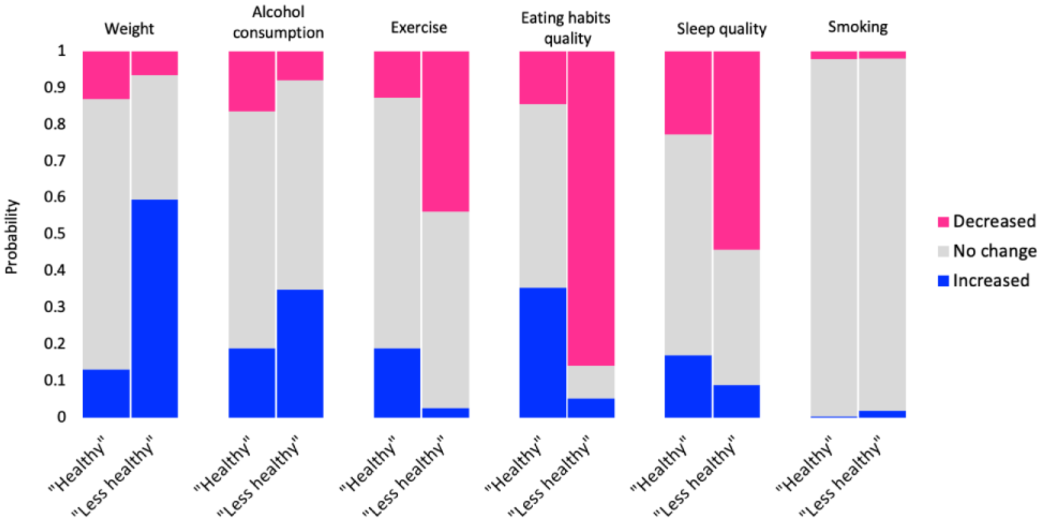
Factor	Increased, n (%)	Unchanged, n (%)	Decreased, n (%)
Weight	523 (32.5)	920 (57.2)	166 (10.3)
Perceived eating habit quality	369 (22.9)	531 (33.0)	709 (44.1)
Sleep quality	219 (13.6)	815 (50.7)	575 (35.7)
Physical activity level	196 (12.2)	1001 (62.2)	412 (25.6)
Smoking	17 (1.1)	1559 (96.9)	33 (2.1)
Alcohol consumption	412 (25.6)	991 (61.6)	206 (12.8)

Emergent Patterns of Change in Weight and Lifestyle Behaviors

Among the tested LCA models, the model with 2 emerging lifestyle behavior change classes had the best fit despite a slightly lower AIC compared to the model with 3 classes (Multimedia Appendix 2, Table S1) and was therefore selected for use. Participants with the class 1 membership (974/1609, 60.5%, population share probability) were more likely to report unchanged weight, sleep quality, smoking and alcohol intake, unchanged/improved eating habits, and increased physical activity. Participants with the class 2 membership (635/1609, 39.5%, population share probability) were more likely to report significant weight gain, deteriorated eating habits and sleep quality, unchanged/increased alcohol intake and smoking, and decreased physical activity. Class 1 was referred to as the healthy and class 2 as the less healthy pattern. Results from the LCA are presented in Figure 1.

Participants with the less healthy pattern were more likely to self-identify as women or a gender minority, to have a lower income prepandemic, and to live with 2 or more chronic diseases. They were also more likely to report a decreased income since the pandemic. Age, ethnicity, residential area, and change in the living situation did not differ between classes.

Figure 1. Two identified classes of change in weight and lifestyle patterns as a result of the pandemic. Classes were identified using LCA based on self-reported changes (prepandemic vs baseline) in weight, physical activity level, smoking, alcohol use, and sleep quality. LCA: latent class analysis.



Associations of Risk Factors With Identified Patterns of Lifestyle Behavior Change

Among risk factors, body image dissatisfaction (odds ratio [OR] 8.8, 95% CI 5.3-14.9), increased stress level (OR 3.4, 95% CI 2.0-5.8), depression (OR 1.7, 95% CI 1.2-2.4), and self-identifying as a gender minority (OR 5.5, 95% CI 1.3-22.3) were positively associated with adopting less healthy lifestyle behaviors in adjusted models (Figure 2 and Multimedia Appendix 2, Table S2). In addition, the province of Québec was inversely associated with less healthy lifestyle behavior changes (OR 0.51, 95% CI 0.33-0.80).

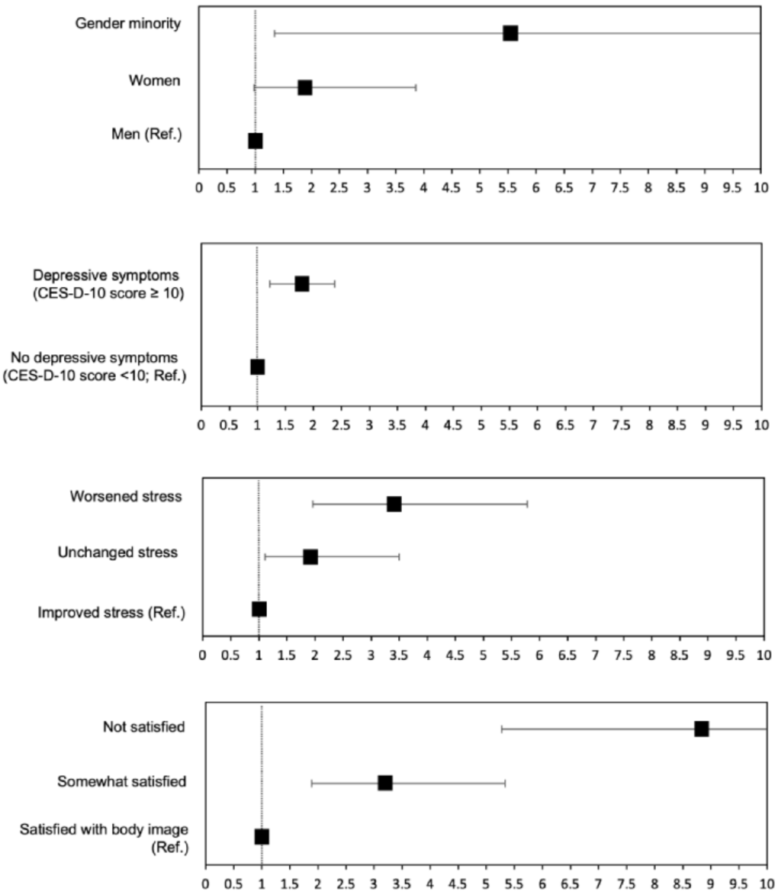
Gender-weighted sensitivity analysis showed a stronger positive association between women and the less healthy lifestyle behavior changes, and the association for gender minorities disappeared (Multimedia Appendix 2, Table S3). The results were otherwise consistent with the nonweighted analysis.

Stratified analysis by age group showed positive associations of body image dissatisfaction, symptoms of depression, increased stress, and women and minority genders with less healthy lifestyle behavior changes in younger adults (<55 years old) and positive associations of symptoms of anxiety and body image dissatisfaction with less healthy lifestyle behavior changes in older adults (Multimedia Appendix 2, Table S4). However, depressive symptoms, gender, and stress level were not determinants of lifestyle behavior change patterns in this subgroup. Living with a chronic disease was predictive of less healthy lifestyle behavior changes in older adults.

Stratified analysis by BMI group showed positive associations between body image dissatisfaction, symptoms of depression, increased stress levels, lenient restrictive measures at the time of enrollment, living with ≥2 chronic diseases, living situation change (lived with people and moved alone), and less healthy lifestyle behavior changes in those with a usual BMI of <25 kg/m^2 (Multimedia Appendix 2, Table S5). In participants with a usual BMI of ≥25 kg/m^2, increased stress levels, maintenance of the same work arrangement, and body image dissatisfaction were positively associated with less healthy lifestyle behavior changes. Living with ≥2 chronic diseases was predictive of less healthy lifestyle behavior changes in this population subgroup. Gender was not a determinant of the outcome within either BMI group.

Stratified analysis by the burden of restrictive measures at the time of enrollment showed body image dissatisfaction, symptoms of depression, and increased levels of stress to be positively associated with less heathy lifestyle behavior changes in both subgroups (Multimedia Appendix 2, Table S6). Living with ≥2 chronic diseases during lenient restrictive measures was predictive of less healthy lifestyle behavior changes. None of the participants recruited during lenient restrictive measures self-identified as a gender minority; gender was only significant within the subgroup of participants recruited during strict restrictive measures.

Figure 2. Unweighted adjusted ORs of the association between psychosocial factors and lifestyle pattern adoption (“less healthy” vs "healthy”). CES-D-10, 10-item Center for Epidemiologic Studies Depression Scale; OR, odds ratio; Ref., reference group. Tables S2 and S3 provide unweighted and gender-weighted ORs.



Discussion

Principal Findings

During the COVID-19 pandemic, more than half of the Canadian COVIDiet study adults reported no change in body weight, smoking status, exercise level, alcohol intake, and sleep quality; however, a dominant proportion (44%) reported worsened eating habits. Considering these lifestyle factors altogether, we identified 2 distinct lifestyle behavior change patterns: healthy and less healthy. The significant and independent determinants of a less healthy lifestyle behavior change pattern were body dissatisfaction, depression, increased stress levels since the pandemic, self-identifying as a gender minority, and the province of residence.

Several surveys have reported on changes in individual lifestyle factors during the pandemic in adults worldwide [10,29-40], but few observational studies have examined the overall lifestyle behavior change patterns and associated risk factors [41,42]. Similar to our findings and using comparable data-driven statistical approaches, one study found “favorable,” “unfavorable,” and “no change” lifestyle change groups in a French cohort [42], while another study identified “prohealthy,” “unhealthy,” and “constant” groups in a Polish population [41].

Adopting Healthy Lifestyle Behaviors

In this study, the healthy lifestyle behavior change pattern (974/1609, 60.5%) was mainly characterized by an increased physical activity level, unchanged or improved perceived eating habits, and stable weight, sleep quality, and smoking status. An increased physical activity level has also been reported in other countries [29,42]. It may have resulted from increased free time due to telework or from a desire to improve personal health in facing the disease outbreak. In addition, participants allocated to the healthy class were more active at baseline, which may have led them to coping with the pandemic-related stress with exercise [43], thus increasing their physical activity level, while decreasing stress levels. With regard to diet, stable and improved eating habits were observed in our study and elsewhere [4,29,42]. Staying at home may have favored the avoidance of less healthy food, such as fast food, and more frequent cooking for some individuals [4,29,42,44]. One study in the Canadian province of Québec (NutriQuebec study; N=853, with 87.2% women, 52.5% aged 50-69 years) showed a slight overall increase in diet quality (+1.1/100 points in the Healthy Eating Index-2015; 95% CI 0.6-1.5) during early lockdown (April-May 2020), especially in younger adults, in those with lower education, or in those with obesity [4]. Although not directly comparable, only 22.9% (369/1609) of the overall population reported improved eating habits and a larger proportion reported worsened eating habits (709/1609, 44.1%) in our study. In addition, only 35.5% (346/974) reported improved eating habits and the majority reported unchanged habits in the healthy lifestyle behavior change group. This suggests that the diet quality of Canadians may not have improved at a population level; in-depth dietary assessment will permit more robust conclusions. Lastly, the COVID-19 pandemic appears to have had disproportionate impacts on populations living with social inequalities [45]. Participants in the healthy pattern group were more likely to self-identify as men, to report higher and unchanged household incomes, to be satisfied with their body image, and to be living with less than 2 chronic diseases. The lifestyle behavior changes observed in this group may be reflective of social privilege.

Adopting Less Healthy Lifestyle Behaviors

Contrastingly, the less healthy lifestyle behavior change pattern (635/1609, 39.5%) was characterized by perceived worsened eating habits, decreased physical activity, decreased sleep quality, and weight gain. Similar patterns of lifestyle habit deterioration were observed in Canada, predominantly in British Columbia [46], and in other countries [47]. Another Canadian study (N=2338, with 90.2% women, February-April 2020) showed that although moderate-to-vigorous activity tracked with wearable devices returned to the prepandemic level 6 weeks after an initial decline in March 2020, the step count and light physical activity remained lower [48]. The lockdown imposed important modifications on daily physical activities, including work-related commuting interruption due to telework or unemployment, the prohibited practice of team sports, and the closure of fitness centers. Those at higher risk of complications and particularly worried of contracting COVID-19 possibly took extra isolation precautions and reduced their physical activity level. In addition, our data show that 1 independent determinant of the less healthy classification was increased levels of stress. Both objective and subjective measures of stress have been shown to hinder physical activity levels, especially in less active individuals [43]. Worsened eating habits may be explained by modified food supply and food choices due to more limited access to grocery stores, fresh food, and working from home. Increased snacking and sweet consumption have been reported in numerous global studies (reviewed in [49]) and may have contributed to participants' perception of worsened eating habits. In addition, the less healthy lifestyle behavior change class more likely reported increased alcohol intake. Altogether, these factors probably explain the observed body weight increase in this group. Comparably, in the NutriNet-Santé cohort study (France; N=37,252; weighted proportions: 52.3% women, 42.5% aged 25-50 years), among the 3 clusters identified, 1 reflected unfavorable changes that included decreased physical activity, increased snacking and sweet consumption, decreased vegetable and fruit consumption, and weight gain between March and May 2020 [50]. These changes were associated with being a woman, working from home, the presence of children at home, a lower income, and more depressive symptoms. In our study, individuals in the less healthy lifestyle behavior change group also more likely reported increased stress levels and anxiety and depressive symptoms and a greater proportion had more than 2 chronic diseases and reported a decreased income. A healthy lifestyle, including regular physical activity and high diet quality, were associated with reduced risk of chronic diseases [51] and mortality [52]. As a strategy to promote health during self-quarantine, the WHO regional office for Europe released physical activity and dietary guidelines [52]. However, none were made available in Canada in the early months of the pandemic. The impact of such guidelines on public health during the pandemic remains to be assessed.

Factors Associated With Lifestyle Behavior Changes

Among potential risk factors, body image satisfaction was the strongest determinant of lifestyle behavior change patterns during the pandemic, independent of several other factors, including gender, age, and income. Close to half of the studied population (707/1609, 43.9%) reported being not satisfied with their body image during the pandemic, a proportion higher than that observed in a population survey in Canada before the pandemic (15%) [53]. Body image dissatisfaction was also more frequent in those self-identifying as women and gender minorities, compared to men, which is typically observed in Western societies [54,55]. Evidence supports that individuals dissatisfied with their body image may have poorer mental health [56] and may engage in less healthy lifestyle behaviors, such as decreased physical activity [57]. A study showed an association of worsened body image dissatisfaction with increased psychological distress during the pandemic [58]. Yet, in our study, body image dissatisfaction was associated with less healthy lifestyle behavior changes independent of stress, anxiety, and depression. The reverse association is also possible, that is, those who adopted less healthy lifestyle behavior changes, including weight gain, may have experienced a greater level of dissatisfaction toward their body image. In our study, the majority of participants also reported increased stress levels since the pandemic (1026/1609, 63.8%), which was associated with less healthy lifestyle behavior changes. Compelling prepandemic evidence has supported an association of stress, anxiety, and depression with poorer lifestyle behaviors, including weight gain [59-63]. However, the pandemic appears to have exacerbated negative mental health in some individuals, putting them at risk of adopting poorer lifestyle behaviors [17] or vice versa [64].

Limited data of the impact of the COVID-19 pandemic on gender minorities have been collected and are available. In our study, self-identifying as a gender minority was associated with adopting less healthy lifestyle habits; although the sample size was small for this group (24/1609, 1.2%), the proportion was comparable to that reported in Canada (0.33% from Census data) [65]. Current evidence supports high rates of worsening physical and mental health outcomes during the pandemic among individuals belonging to gender minorities [66,67]. From stratified analyses, both self-identifying as a woman and self-identifying as a gender minority were associated with less healthy lifestyle behavior changes among younger participants (<55 years) only. Before the pandemic, women (compared to men) and gender-diverse individuals (compared to men and women) already reported poorer mental health in Canada. Since the beginning of social distancing, negative mental health has been reported in women and gender minorities in Canada [68] and elsewhere [58,69-71]. In people belonging to gender minorities, the lack or loss of access to community groups, stigmatization, and negative psychological outcomes of the pandemic may have contributed to the adoption of less healthy lifestyle behavior changes [66]. In younger women and gender minorities, the telework and at-home childcare arrangement may have further adversely impacted lifestyle behaviors [42,71].

Living in the province of Québec was significantly associated with adopting healthier lifestyle behaviors during the pandemic compared to Ontario and other large provinces. Yet, restrictive measures have been comparable across these provinces throughout the pandemic [1]. Some evidence suggests that the Québec population may have experienced healthier changes [4]. However, the association with healthier lifestyle behavior changes during the pandemic is most likely attributable to greater recruitment in this province during less stringent restrictive measures in our study (Québec: 301/576 during lenient and 71/1033 during strict restrictions; Ontario: 139/576 during lenient and 464/1033 during strict restrictions).

Strengths and Limitations

The strengths of this study include the overall large sample size and adequate representation across provinces. Another strength is the use of LCA, as opposed to using measures of central tendency, that allowed us to identify 2 classes and informed on distinct patterns within the population.

The study also has limitations. First, variations in restriction measures were observed across provinces between the start and the end of recruitment, which may have contributed to confounding residuals, although our models were adjusted for the pandemic phases during which participants responded to the questionnaire. Second, the use of self-reported lifestyle questionnaires may have introduced subjectivity to measurements. Third, the use of advertisement strategies combined with the nature of our study may have led to selection bias of more educated and healthier individuals. In addition, the over-/undersampling of specific groups of individuals, namely gender and age, may have limited the generalizability of our results to the Canadian population; however, our sensitivity analysis using gender-based calibration weights showed consistent results with our main analysis (Multimedia Appendix 2, Table S3). Fourth, although subgroup analysis by age, baseline BMI, and pandemic wave group was successfully conducted, it was not possible to examine the associations between risk factors and lifestyle behavior change patterns across ethnic and gender groups; further studies with a larger sample size within each, especially ethnic and gender minorities, are warranted.

Conclusion

Findings from this observational study showed that 39.5% (635/1609) of participants adopted less healthy lifestyle behaviors during the COVID-19 pandemic but that the majority (974/1609, 60.5%) adopted healthier behaviors. Factors such as body image perception, depression status, change in the stress level, and gender identity may have motivated these lifestyle behavior changes. It is yet to be determined whether the latter will sustain in the long term. In the postpandemic period, multidisciplinary strategies involving dietitians, psychologists, and kinesiologists may be considered to assist adults in (re)gaining mental and physical health. Lockdown-adapted public health recommendations and resources related to nutrition, physical activity, sleep, and smoking should be promoted during future outbreaks.

Acknowledgments

AJT, SC, and ASB elaborated the study design. SC, ASB, AIR, AM, CFG, CL, HY, SLB, and AJT contributed to participant recruitment. AM, CL, and AIR conducted the study and data collection. HY contributed to data preparation. AM and AJT contributed to table and figure preparation. AJT conducted statistical analyses and drafted the manuscript. SC obtained ethics approval and funding and supervised all activities. All authors have read, edited, and approved the final manuscript. This research was funded by McGill University.

Data Availability

The data sets generated and analyzed during the study are not publicly available due to ethical restrictions but are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

STROBE diagram showing the flow of participants in the Canadian COVIDiet cohort study. Inclusion criteria included reading and speaking English or French, owning a smartphone or tablet, and access to the internet. Exclusion criteria were living with an active and uncontrolled acute disease that interfered with usual food intake, being hospitalized, being pregnant, or living in care homes or other institutional environments.

[DOCX File , 196 KB - publichealth_v9i1e43786_app1.docx]

Multimedia Appendix 2

Supplemental Tables 1-6.
[DOCX File , 44 KB - publichealth_v9i1e43786_app2.docx]

References

1. Lawson T, Nathans L, Goldenberg A, Fimiani M, Boire-Schwab D, Waschuk G. COVID-19: emergency measures tracker. McCarthy Tétrault. 2022. URL: https://www.mccarthy.ca/en/insights/articles/covid-19-emergency-measures-tracker [accessed 2023-03-14]
2. Government of Canada announces major steps in treating and preventing COVID-19 through vaccines and therapies. Government of Canada. 2020. URL: https://www.canada.ca/en/innovation-science-economic-development/news/2020/08/government-of-canada-announces-major-steps-in-treating-and-preventing-covid-19-through-vaccines-and-therapies.html [accessed 2023-03-14]
3. Mattioli AV, Sciomer S, Cocchi C, Maffei S, Gallina S. Quarantine during COVID-19 outbreak: changes in diet and physical activity increase the risk of cardiovascular disease. Nutr Metab Cardiovasc Dis 2020 Aug 28;30(9):1409-1417 [FREE Full text] [doi: 10.1016/j.numecd.2020.05.020] [Medline: 32571612]
4. Lamarche B, Brassard D, Lapointe A, Laramée C, Kearney M, Côté M, et al. Changes in diet quality and food security among adults during the COVID-19-related early lockdown: results from NutriQuébec. Am J Clin Nutr 2021 Apr 06;113(4):984-992 [FREE Full text] [doi: 10.1093/ajcn/nqaa363] [Medline: 33398347]
5. Lesser IA, Nienhuis CP. The impact of COVID-19 on physical activity behavior and well-being of Canadians. Int J Environ Res Public Health 2020 May 31;17(11):3899 [FREE Full text] [doi: 10.3390/ijerph17113899] [Medline: 32486380]
6. Voyer CJ. Un sondage confirme que la pandémie modifie les habitudes de vie et la consommation d?alcoolet de cannabis des québécois. Coalition Poids. 2020. URL: https://cqpp.qc.ca/fr/salle-de-presse/un-sondage-confirme-que-la-pandemie-modifie-les-habitudes-de-vie-et-la-consommation-dalcool-et-de-cannabis-des-quebecois/ [accessed 2023-03-14]
7. Marty L, de Lauzon-Guillain B, Labesse M, Nicklaus S. Food choice motives and the nutritional quality of diet during the COVID-19 lockdown in France. Appetite 2021 Feb 01;157:105005 [FREE Full text] [doi: 10.1016/j.appet.2020.105005] [Medline: 33068666]
8. Nielsen DE, Labonté K, Karamanoglu I, Han HY, Tavanaei M, Duhamel P, et al. Longitudinal patterns of food procurement over the course of the COVID-19 pandemic: findings from a Canadian online household survey. Front Public Health 2021;9:752204 [FREE Full text] [doi: 10.3389/fpubh.2021.752204] [Medline: 35127611]
9. Sallis R, Young DR, Tartof SY, Sallis JF, Sall J, Li Q, et al. Physical inactivity is associated with a higher risk for severe COVID-19 outcomes: a study in 48 440 adult patients. Br J Sports Med 2021 Oct 13;55(19):1099-1105 [FREE Full text] [doi: 10.1136/bjsports-2021-104080] [Medline: 33849909]
10. Li X, Li J, Qing P, Hu W. COVID-19 and the change in lifestyle: bodyweight, time allocation, and food choices. Int J Environ Res Public Health 2021 Oct 08;18(19):10552 [FREE Full text] [doi: 10.3390/ijerph181910552] [Medline: 34639852]
11. Li Y, Schoufour J, Wang DD, Dhana K, Pan A, Liu X, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020 Jan 08;368:l6669 [FREE Full text] [doi: 10.1136/bmj.l6669] [Medline: 31915124]
12. Kowal M, Coll-Martín T, Ikizer G, Rasmussen J, Eichel K, Studzińska A, et al. Who is the most stressed during the COVID-19 pandemic? Data from 26 countries and areas. Appl Psychol Health Well Being 2020 Dec;12(4):946-966 [FREE Full text] [doi: 10.1111/aphw.12234] [Medline: 32996217]
13. Shin H, Kim J, Lee H. Association of between depression, precautionary behavior compliance, COVID-19 fear, and health behaviors: analysis of data from a national cross-sectional study. JMIR Public Health Surveill 2023 Feb 22;9:e42677 [FREE Full text] [doi: 10.2196/42677] [Medline: 36716130]
14. Son C, Hegde S, Smith A, Wang X, Sasangohar F. Effects of COVID-19 on college students' mental health in the United States: interview survey study. J Med Internet Res 2020 Sep 03;22(9):e21279 [FREE Full text] [doi: 10.2196/21279] [Medline: 32805704]
15. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
16. Hadar-Shoval D, Alon-Tirosh M, Asraf K, Tannous-Haddad L, Tzischinsky O. Lifestyle changes, emotional eating, gender, and stress during COVID-19 lockdown. Nutrients 2022 Sep 18;14(18):3868 [FREE Full text] [doi: 10.3390/nu14183868] [Medline: 36145243]
17. Caroppo E, Mazza M, Sannella A, Marano G, Avallone C, Claro AE, et al. Will nothing be the same again?: Changes in lifestyle during COVID-19 pandemic and consequences on mental health. Int J Environ Res Public Health 2021 Aug 10;18(16):8433 [FREE Full text] [doi: 10.3390/ijerph18168433] [Medline: 34444180]
18. Krejcie RV, Morgan DW. Determining sample size for research activities. Educ Psychol Meas 2016 Jul 02;30(3):607-610. [doi: 10.1177/001316447003000308]
19. Population estimates on July 1st, by age and sex. Statistics Canada. 2021. URL: https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=1710000501 [accessed 2023-03-14]
20. AirTable: your platform for building connected apps. AirTable. URL: https://www.airtable.com/product [accessed 2023-03-14]

21. Statistics C. Canadian Community Health Survey (CCHS) - annual component. Statistics Canada. 2020. URL: https://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=assembleInstr&a=1&&lang=en&Item_Id=1262397 [accessed 2023-03-14]
22. Speech from the throne: opening of the second session of the forty-third parliament. Government of Canada. URL: https://www.canada.ca/en/privy-council/campaigns/speech-throne/2020/speech-from-the-throne.html [accessed 2023-03-14]
23. Physical activity. World Health Organization. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2023-03-14]
24. Canadian Community Health Survey (CCHS) - rapid response - healthy living. Statistics Canada. 2019. URL: https://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=assembleInstr&lang=en&Item_Id=1262718 [accessed 2023-03-14]
25. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
26. Andresen EM. Screening for depression in well older adults: evaluation of a short form of the CES-D. Am J Prev Med 1994;10(2):77.
27. Sex at birth and gender: technical report on changes for the 2021 census. Statistics Canada. 2020. URL: https://www12.statcan.gc.ca/census-recensement/2021/ref/98-20-0002/982000022020002-eng.cfm [accessed 2023-03-14]
28. DrTessier / COVIDiet_study. GitHub. 2022 Oct 1. URL: https://github.com/DrTessier/COVIDiet_study [accessed 2023-03-14]
29. Di Renzo L, Gualtieri P, Pivari F, Soldati L, Attinà A, Cinelli G, et al. Eating habits and lifestyle changes during COVID-19 lockdown: an Italian survey. J Transl Med 2020 Jun 08;18(1):229 [FREE Full text] [doi: 10.1186/s12967-020-02399-5] [Medline: 32513197]
30. Balanzá-Martínez V, Kapczinski F, de Azevedo Cardoso T, Atienza-Carbonell B, Rosa AR, Mota JC, et al. The assessment of lifestyle changes during the COVID-19 pandemic using a multidimensional scale. Rev Psiquiatr Salud Ment (Engl Ed) 2021 Aug;14(1):16-26 [FREE Full text] [doi: 10.1016/j.rpsm.2020.07.003] [Medline: 32962948]
31. López-Moreno M, López MTI, Miguel M, Garcés-Rimón M. Physical and psychological effects related to food habits and lifestyle changes derived from Covid-19 home confinement in the Spanish population. Nutrients 2020 Nov 10;12(11):3445 [FREE Full text] [doi: 10.3390/nu12113445] [Medline: 33182816]
32. Cancello R, Soranna D, Zambra G, Zambon A, Invitti C. Determinants of the lifestyle changes during COVID-19 pandemic in the residents of Northern Italy. Int J Environ Res Public Health 2020 Aug 28;17(17):6287 [FREE Full text] [doi: 10.3390/ijerph17176287] [Medline: 32872336]
33. Hu Z, Lin X, Chiwanda Kaminga A, Xu H. Impact of the COVID-19 epidemic on lifestyle behaviors and their association with subjective well-being among the general population in mainland China: cross-sectional study. J Med Internet Res 2020 Aug 25;22(8):e21176 [FREE Full text] [doi: 10.2196/21176] [Medline: 32759103]
34. Souza TC, Oliveira LA, Daniel MM, Ferreira LG, Della Lucia CM, Liboredo JC, et al. Lifestyle and eating habits before and during COVID-19 quarantine in Brazil. Public Health Nutr 2022 Jan;25(1):65-75 [FREE Full text] [doi: 10.1017/S136898002100255X] [Medline: 34108067]
35. van der Werf ET, Busch M, Jong MC, Hoenders HJR. Lifestyle changes during the first wave of the COVID-19 pandemic: a cross-sectional survey in the Netherlands. BMC Public Health 2021 Jun 25;21(1):1226 [FREE Full text] [doi: 10.1186/s12889-021-11264-z] [Medline: 34172042]
36. Abouzid M, El-Sherif DM, Eltewacy NK, Dahman NBH, Okasha SA, Ghozy S, EARG Collaborators. Influence of COVID-19 on lifestyle behaviors in the Middle East and North Africa Region: a survey of 5896 individuals. J Transl Med 2021 Mar 30;19(1):129 [FREE Full text] [doi: 10.1186/s12967-021-02767-9] [Medline: 33785043]
37. Cervera-Martínez J, Atienza-Carbonell B, Mota JC, Bobes-Bascarán T, Crespo-Facorro B, Esteban C, et al. Lifestyle changes and mental health during the COVID-19 pandemic: a repeated, cross-sectional web survey. J Affect Disord 2021 Dec 01;295:173-182 [FREE Full text] [doi: 10.1016/j.jad.2021.08.020] [Medline: 34469856]
38. Alon-Tirosh M, Hadar-Shoval D, Asraf K, Tannous-Haddad L, Tzischinsky O. The association between lifestyle changes and psychological distress during COVID-19 lockdown: the moderating role of COVID-related stressors. Int J Environ Res Public Health 2021 Sep 15;18(18):9695 [FREE Full text] [doi: 10.3390/ijerph18189695] [Medline: 34574620]
39. Ferrante G, Camussi E, Piccinelli C, Senore C, Armaroli P, Ortale A, et al. Did social isolation during the SARS-CoV-2 epidemic have an impact on the lifestyles of citizens? Epidemiol Prev 2020;44(5-6 Suppl 2):353-362 [FREE Full text] [doi: 10.19191/EP20.5-6.S2.137] [Medline: 33412829]
40. Enriquez-Martinez OG, Martins MCT, Pereira TSS, Pacheco SOS, Pacheco FJ, Lopez KV, et al. Diet and lifestyle changes during the COVID-19 pandemic in Ibero-American countries: Argentina, Brazil, Mexico, Peru, and Spain. Front Nutr 2021;8:671004 [FREE Full text] [doi: 10.3389/fnut.2021.671004] [Medline: 34150828]
41. Górnicka M, Drywień ME, Zielinska MA, Hamułka J. Dietary and lifestyle changes during COVID-19 and the subsequent lockdowns among Polish adults: a cross-sectional online survey PLifeCOVID-19 study. Nutrients 2020 Aug 03;12(8):2324 [FREE Full text] [doi: 10.3390/nu12082324] [Medline: 32756458]
42. Deschasaux-Tanguy M, Druesne-Pecollo N, Esseddik Y, de Edelenyi FS, Allès B, Andreeva VA, et al. Diet and physical activity during the coronavirus disease 2019 (COVID-19) lockdown (March-May 2020): results from the French NutriNet-Santé cohort study. Am J Clin Nutr 2021 Apr 06;113(4):924-938 [FREE Full text] [doi: 10.1093/ajcn/nqaa336] [Medline: 33675635]

43. Stults-Kolehmainen MA, Sinha R. The effects of stress on physical activity and exercise. Sports Med 2014 Jan;44(1):81-121 [FREE Full text] [doi: 10.1007/s40279-013-0090-5] [Medline: 24030837]
44. Dou Z, Stefanovski D, Galligan D, Lindem M, Rozin P, Chen T. The COVID-19 pandemic impacting household food dynamics: a cross-national comparison of China and the U.S. SocArXiv Papers Preprint posted online May 23, 2020 [FREE Full text] [doi: 10.31235/osf.io/64jwy]
45. UN Secretary-General's policy brief: the impact of COVID-19 on women. United Nations. 2020. URL: https://www.unwomen.org/en/digital-library/publications/2020/04/policy-brief-the-impact-of-covid-19-on-women [accessed 2023-03-14]
46. He M, Xian Y, Lv X, He J, Ren Y. Changes in the COVID-19 body weight, physical activity, and lifestyle during the semi-lockdown period after the outbreak of COVID-19 in China: an online survey. Disaster Med Public Health Prep 2021 Apr 14;15(2):e23-e28 [FREE Full text] [doi: 10.1017/dmp.2020.237] [Medline: 32660669]
47. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. Int J Environ Res Public Health 2020 Jun 07;17(11):4065 [FREE Full text] [doi: 10.3390/ijerph17114065] [Medline: 32517294]
48. Di Sebastiano KM, Chulak-Bozzer T, Vanderloo LM, Faulkner G. Don't walk so close to me: physical distancing and adult physical activity in Canada. Front Psychol 2020;11:1895 [FREE Full text] [doi: 10.3389/fpsyg.2020.01895] [Medline: 32849110]
49. Bennett G, Young E, Butler I, Coe S. The impact of lockdown during the COVID-19 outbreak on dietary habits in various population groups: a scoping review. Front Nutr 2021 Mar 4;8:626432 [FREE Full text] [doi: 10.3389/fnut.2021.626432] [Medline: 33748175]
50. Han H, Cao Y, Feng C, Zheng Y, Dhana K, Zhu S, et al. Association of a healthy lifestyle with all-cause and cause-specific mortality among individuals with type 2 diabetes: a prospective study in UK Biobank. Diabetes Care 2022 Feb 01;45(2):319-329. [doi: 10.2337/dc21-1512] [Medline: 34857534]
51. Li Y, Pan A, Wang DD, Liu X, Dhana K, Franco OH, et al. Impact of healthy lifestyle factors on life expectancies in the US population. Circulation 2018 Jul 24;138(4):345-355 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.117.032047] [Medline: 29712712]
52. Stay physically active during self-quarantine. World Health Organization. URL: https://tinyurl.com/3e264bsv [accessed 2023-03-14]
53. 11% of Canadians “completely satisfied” with their looks: GfK study. Growth from Knowledge. 2015. URL: https://www.gfk.com/en-us/press/11-of-canadians-completely-satisfied-with-their-looks-gfk-study [accessed 2022-09-29]
54. Carter A, Forrest JI, Kaida A. Association between internet use and body dissatisfaction among young females: cross-sectional analysis of the Canadian Community Health Survey. J Med Internet Res 2017 Feb 09;19(2):e39 [FREE Full text] [doi: 10.2196/jmir.5636] [Medline: 28183688]
55. Voges MM, Giabbiconi C, Schöne B, Waldorf M, Hartmann AS, Vocks S. Gender differences in body evaluation: do men show more self-serving double standards than women? Front Psychol 2019;10:544 [FREE Full text] [doi: 10.3389/fpsyg.2019.00544] [Medline: 30930819]
56. Barnes M, Abhyankar P, Dimova E, Best C. Associations between body dissatisfaction and self-reported anxiety and depression in otherwise healthy men: a systematic review and meta-analysis. PLoS One 2020;15(2):e0229268 [FREE Full text] [doi: 10.1371/journal.pone.0229268] [Medline: 32097427]
57. Sabiston C, Pila E, Vani M, Thogersen-Ntoumani C. Body image, physical activity, and sport: a scoping review. Psychol Sport Exerc 2019 May;42:48-57 [FREE Full text] [doi: 10.1016/j.psychsport.2018.12.010]
58. Robinson E, Boyland E, Chisholm A, Harrold J, Maloney NG, Marty L, et al. Obesity, eating behavior and physical activity during COVID-19 lockdown: a study of UK adults. Appetite 2021 Jan 01;156:104853 [FREE Full text] [doi: 10.1016/j.appet.2020.104853] [Medline: 33038479]
59. Sarris J, Thomson R, Hargraves F, Eaton M, de Manincor M, Veronese N, et al. Multiple lifestyle factors and depressed mood: a cross-sectional and longitudinal analysis of the UK Biobank (N = 84,860). BMC Med 2020 Nov 12;18(1):354 [FREE Full text] [doi: 10.1186/s12916-020-01813-5] [Medline: 33176802]
60. Choi KW, Stein MB, Nishimi KM, Ge T, Coleman JR, Chen C, 23andMe Research Team, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, et al. An exposure-wide and mendelian randomization approach to identifying modifiable factors for the prevention of depression. Am J Psychiatry 2020 Oct 01;177(10):944-954 [FREE Full text] [doi: 10.1176/appi.ajp.2020.19111158] [Medline: 32791893]
61. Atzendorf J, Apfelbacher C, Gomes de Matos E, Kraus L, Piontek D. Patterns of multiple lifestyle risk factors and their link to mental health in the German adult population: a cross-sectional study. BMJ Open 2018 Dec 19;8(12):e022184 [FREE Full text] [doi: 10.1136/bmjopen-2018-022184] [Medline: 30573479]
62. Loprinzi PD, Mahoney S. Concurrent occurrence of multiple positive lifestyle behaviors and depression among adults in the United States. J Affect Disord 2014 Aug;165:126-130. [doi: 10.1016/j.jad.2014.04.073] [Medline: 24882189]
63. Buttery AK, Mensink GBM, Busch MA. Healthy behaviours and mental health: findings from the German Health Update (GEDA). Eur J Public Health 2015 Apr;25(2):219-225. [doi: 10.1093/eurpub/cku094] [Medline: 25061231]

64. Parameshwar A, Maiya GA, Kamath SU, Shastry BA, Ravishankar. Lifestyle modification with physical activity promotion on leptin resistance and quality of life in metabolic syndrome - a systematic review with meta-analysis. Curr Diabetes Rev 2021;17(3):345-355. [doi: 10.2174/1573399816666200211102917] [Medline: 32072915]
65. Canada is the first country to provide census data on transgender and non-binary people. Statistics Canada. 2022. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/220427/dq220427b-eng.htm [accessed 2023-03-14]
66. Nowaskie DZ, Roesler AC. The impact of COVID-19 on the LGBTQ+ community: comparisons between cisgender, heterosexual people, cisgender sexual minority people, and gender minority people. Psychiatry Res 2022 Mar;309:114391 [FREE Full text] [doi: 10.1016/j.psychres.2022.114391] [Medline: 35032764]
67. Flentje A, Obedin-Maliver J, Lubensky ME, Dastur Z, Neilands T, Lunn MR. Depression and anxiety changes among sexual and gender minority people coinciding with onset of COVID-19 pandemic. J Gen Intern Med 2020 Sep;35(9):2788-2790 [FREE Full text] [doi: 10.1007/s11606-020-05970-4] [Medline: 32556877]
68. Moyser M. Gender differences in mental health during the COVID-19 pandemic. Statistics Canada. 2020. URL: https://www150.statcan.gc.ca/n1/pub/45-28-0001/2020001/article/00047-eng.htm [accessed 2023-03-14]
69. Banks J, Xu X. The mental health effects of the first two months of lockdown during the COVID‐19 pandemic in the UK. Fiscal Stud 2020 Nov 30;41(3):685-708. [doi: 10.1111/1475-5890.12239]
70. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892 [FREE Full text] [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]
71. Nienhuis CP, Lesser IA. The impact of COVID-19 on women's physical activity behavior and mental well-being. Int J Environ Res Public Health 2020 Dec 04;17(23):9036 [FREE Full text] [doi: 10.3390/ijerph17239036] [Medline: 33291530]

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion
CCHS: Canadian Community Health Survey
CES-D-10: 10-item Center for Epidemiologic Studies Depression Scale
GAD-7: General Anxiety Disorder-7
LCA: latent class analysis
OR: odds ratio
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 14.11.22; peer-reviewed by M Stults-Kolehmainen, J Atzendorf; comments to author 29.01.23; revised version received 13.02.23; accepted 23.02.23; published 30.03.23.

Please cite as:
Tessier AJ, Moyen A, Lawson C, Rappaport AI, Yousif H, Fleurent-Grégoire C, Lalonde-Bester S, Brazeau AS, Chevalier S
Lifestyle Behavior Changes and Associated Risk Factors During the COVID-19 Pandemic: Results from the Canadian COVIDiet Online Cohort Study
JMIR Public Health Surveill 2023;9:e43786
URL: https://publichealth.jmir.org/2023/1/e43786
doi:10.2196/43786
PMID:36848226

Original Paper

US Adults Practicing Healthy Lifestyles Before and During COVID-19: Comparative Analysis of National Surveys

Hon Lon Tam[1], PhD; Sek Ying Chair[1], PhD; Isaac Sze Him Leung[2], PhD; Leona Yuen Ling Leung[3,4], DN; Alex Siu Wing Chan[5], PhD

[1]The Nethersole School of Nursing, The Chinese University of Hong Kong, Shatin, NT, China (Hong Kong)

[2]Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, China (Hong Kong)

[3]The Ronin Institute, Montclair, NJ, United States

[4]Canadian Academy of Independent Scholars, Vancouver, BC, Canada

[5]Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hung Hom, Kowloon, China (Hong Kong)

Corresponding Author:
Hon Lon Tam, PhD
The Nethersole School of Nursing
The Chinese University of Hong Kong
Rm 704B
7/F, Esther Lee Building
Shatin, NT
China (Hong Kong)
Phone: 852 39439306
Email: hltam@cuhk.edu.hk

Abstract

Background: Practicing healthy lifestyles can reduce the risk to develop noncommunicable diseases and the related mortality. Studies showed that practicing healthy lifestyles could enhance disease-free life expectancy and preserve bodily functions. However, engagement in healthy lifestyle behavior was suboptimal.

Objective: This study aimed to define individuals' lifestyle characteristics before and during COVID-19 and determine the factors associated with practicing a healthy lifestyle. This cross-sectional study was conducted using data from the 2019 and 2021 Behavioral Risk Factor Surveillance System surveys.

Methods: US individuals aged ≥18 years were interviewed via phone call. Healthy lifestyles were assessed through corresponding questions regarding the maintenance of optimal body weight, physical activity, daily consumption of at least five portions of fruits and vegetables, current smoking status, and alcohol consumption. Missing data were imputed using a package in the R statistical software. The effects of practicing a healthy lifestyle on cases without missing data and those with imputation were reported.

Results: There were 550,607 respondents (272,543 and 278,064 from 2019 and 2021, respectively) included in this analysis. The rates of practicing a healthy lifestyle were 4% (10,955/272,543) and 3.6% (10,139/278,064) in 2019 and 2021, respectively. Although 36.6% (160,629/438,693) of all 2021 respondents had missing data, the results of the logistic regression analysis for cases without missing data and those with imputation were similar. Of the cases with imputation, women (odds ratio [OR] 1.87) residing in urban areas (OR 1.24) with high education levels (OR 1.73) and good or better health status (OR 1.59) were more likely to practice healthier lifestyles than young individuals (OR 0.51-0.67) with a low household income (OR 0.74-0.78) and chronic health conditions (OR 0.48-0.74).

Conclusions: A healthy lifestyle should be strongly promoted at the community level. In particular, factors associated with a low rate of practice of healthy lifestyles should be targeted.

(JMIR Public Health Surveill 2023;9:e45697) doi:10.2196/45697

KEYWORDS

healthy lifestyle; health risk behaviors; habits; noncommunicable diseases; population surveillance; Behavioural Risk Factor Surveillance System; BRFSS

Introduction

Noncommunicable diseases (NCDs) are chronic health conditions including cardiovascular diseases and diabetes, which were responsible for approximately 74% deaths worldwide in 2019 [1]. Practicing a healthy lifestyle reduces the risk of NCDs and associated all-cause mortality [2-5]; it has been shown to reduce 78% and 42% of incident diabetes and hypertension, respectively [2,3]. The following 5 components determine a healthy lifestyle: optimal body weight, regular physical activity, healthy diet, smoking cessation, and limited alcohol consumption. Loef and Walach [4] conducted a meta-analysis of 15 studies and reported that strict adherence to most of these components could significantly reduce all-cause mortality (ie, from 26% to 68%). These findings were supported by the results of a cohort study involving 487,198 healthy adults; after a median follow-up of 10.2 years, Zhu et al [5] reported that practicing a healthy lifestyle could reduce up to 68% of the all-cause mortality and mortality rate associated with ischemic heart disease (58%), stroke (79%), and respiratory diseases (81%). A large-scale cohort study involving 111,562 participants with 28-34 years of follow-up showed that a healthy lifestyle could enhance disease-free life expectancy at 50 years of age by up to 31.1 years [6]. Another 3-year follow-up cohort study involving 3107 older adults found that practicing a healthy lifestyle could alleviate the deterioration of gait speed and cognitive function [7]. In brief, practicing a healthy lifestyle can reduce all-cause mortality, slow down the progression of NCDs, and preserve bodily functions. Accordingly, the World Health Organization (WHO) has recommended practicing healthy lifestyles to prevent the risk of NCDs across all age groups [8].

Despite this, engagement in healthy lifestyle behaviors has remained suboptimal, with some factors hindering individuals from pursing such lifestyles. Using a random sampling technique, Azizan et al [9] investigated adherence to a healthy diet among low-income individuals in Malaysia; only 34.8% of the 1450 participants followed a healthy diet. A low level of education and diabetes prevented individuals from following a healthy diet [9]. However, a cross-sectional study conducted in Iran comprising 500 patients with diabetes revealed that over 80% of the participants followed healthy diets and were physically active; the use of convenience sampling technique might be the reason of the high adherence rate of healthy diets and physical activity [10]. Another cross-sectional study conducted in Jordan, using random sampling, involving 1000 patients with hypertension reported that only 23% of the included participants practiced healthy lifestyles; being men and having a low level of hypertension knowledge prevented the participants from practicing healthy lifestyles [11]. On the other hand, fewer rural residents practice healthy lifestyles [12,13]. This might be why a higher mortality rate was noted in the rural areas [14]. Different sociodemographic conditions, comorbidities, and residential areas may influence individuals' decisions to practice healthy lifestyles.

A healthy lifestyle should be promoted at all ages. Studies have shown that practicing a healthy lifestyle may not always result in satisfactory outcomes and that many factors can influence individuals' decisions to pursue such lifestyles. However, COVID-19 has markedly altered our living environments and behaviors [15-17]. For example, a 3-month longitudinal study found that people spent less time visiting public places and became physical inactivity during COVID-19 [18]. Previous studies assessing healthy lifestyles were conducted at the community level and before the pandemic was announced. Our main objectives were to (1) define the individuals' characteristics and their practice of healthy lifestyles in 2019 and 2021 and (2) determine the factors associated with practicing a healthy lifestyle.

Methods

This cross-sectional study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guideline [19]. The study was performed using publicly available data from the Centers for Disease Control and Prevention (CDC).

Data Sources

Secondary analysis of data from the 2019 and 2021 Behavioral Risk Factor Surveillance System (BRFSS) surveys, which are national annual surveys that investigate health-related behaviors and health conditions of noninstitutionalized US adults aged ≥18 years, was performed [20]. The surveyor used the random-digit–dialing technique to generate random phone numbers for telephone and mobile phone interviews across 50 states in the United States, the District of Columbia, Guam, Puerto Rico, and the US Virgin Islands. The latest BRFSS survey was conducted in 2021. The median survey response rate of the 2019 BRFSS survey and 2021 BRFSS survey was 49.4% and 44%, respectively [20,21].

Study Population

Cases for which data regarding sociodemographic variables or chronic health conditions were missing and individuals who refused to answer questions were excluded from the study. We included 272,543/418,268 (65.2%) and 278,064/438,693 (63.4%) individuals from the 2019 and 2021 BRFSS 2021 surveys, respectively.

Variables

Sociodemographic Characteristics and Perceived Health

Following previous studies [9-12], several sociodemographic variables and the perceived health were extracted from the BRFSS surveys. Marital status, employment status, and housing were recategorized in this study, whereas the rest followed the categorizations from the BRFSS survey. In addition, some variables were recategorized in the logistic regression analysis, where perceived health was categorized as "fair or poor" (including the rating of fair and poor) and "good or better" (including the rating of good, very good, and excellent) and education level was recategorized as "high school or less" (including attended and graduated from high school) and "college and above" (including attended and graduated from college).

Healthy Lifestyle

A healthy lifestyle was defined by adherence to healthy behaviors and the avoidance of risky behaviors. Based on the items in the BRFSS and WHO recommendations [8], health behaviors were determined by three components: (1) optimal body weight, defined by a no response to the item “adults who have a body mass index greater than 25”; (2) physical activity, defined by the item “adults who reported doing physical activity or exercise during the past 30 days other than their regular job”; and (3) healthy diet, defined by the reports of the daily consumption of at least five portions of fruits and vegetables.

Following WHO recommendations [8], two risky behaviors were investigated: (1) current smoker status, which was directly surveyed using the BRFSS; and (2) excessive alcohol users, defined as heavy or binge drinkers.

Chronic Health Conditions

Several chronic health conditions, namely hypertension, high cholesterol, diabetes, asthma, chronic obstructive pulmonary disease, and depression, were assessed using the BRFSS survey data.

Statistical Analysis

Data were analyzed using the R statistical software (R Foundation for Statistical Computing). Logistic regression analysis was used to investigate which independent variables were associated with practicing healthy lifestyles (dependent variable). These independent variables included the sociodemographic characteristics, health status, and health-related behaviors. To inform future research, the latest 2021 data set was analyzed. Because the percentage of the excluded subjects was as high as 36.6% and the missing rate of outcome variables were associated with various covariates (missing at random instead of missing at completely random), multiple imputations were conducted to estimate the missing variables among the excluded subjects. Five imputed data sets were created by the *multiple imputation via chained equations (MICE)* package in R.

MICE imputation starts by a simple imputation such as mean imputations on all variables with missing value. Then, these imputed missing values are updated by running a series of regressions, whereby each variable with missing value is regressed upon other variables repeatedly until convergence. For cases with imputations, logistic regression analyses were performed for each of the imputed data sets. The regression coefficients were combined using the *pool()* function within the *MICE* package. In addition, we used the *detectseparation* package in R to check if there was any multicollinearity in our logistic regression analysis. The results of the 2021 data set logistic regression analyses were presented both for cases without missing data and for those with imputations. A *P* value <.05 was considered statistically significant.

Ethical Considerations

The surveys were conducted by US federal agencies where verbal consent was obtained from each participant prior to the data collection [20,21]. The BRFSS data set are allowed for secondary data analysis without permission required. No institutional review board approval or patient informed consent was required for secondary data analysis. The participants’ personal information had been deidentified in the data set prior to being publicly available [20,21].

Results

Individual Characteristics

We analyzed 550,607 respondents (272,543 and 278,064 from the 2019 and 2021 BRFSS surveys, respectively; Table 1). The majority of the respondents were female (n=284,876, 51.7%) and resided in urban areas (n=468,547, 85.1%); 35.2% (n=193,850) of the respondents were aged 65 years and older. Regarding health and related behaviors, 83% (n=456,846) of individuals were perceived to have a good or better health status. Hypertension (n=276,148, 50.2%) and high cholesterol (n=248,440, 45.1%) were the most commonly reported chronic diseases. The percentage of respondents who performed physical activity was greater in 2021 than in 2019 (214,943/278,064, 77.3% vs 204,733/272,543, 75.1%); in contrast, the percentage of smokers was lower in 2021 than in 2019 (36,983/278,064, 13.3% vs 38,662/272,543, 14.2%). The percentage of practicing a healthy lifestyle nonsignificantly decreased from 4% in 2019 to 3.6% in 2021.

Table 1. Individual characteristics among complete cases without missing data in 2019 and 2021.

Variables	2019 (n=272,543), n (%)	2021 (n=278,064), n (%)
Sociodemographic		
Sex		
Male	129,986 (47.7)	135,745 (48.8)
Female	142,557 (52.3)	142,319 (51.2)
Age range (years)		
18-24	13,910 (5.1)	11,020 (4)
25-34	28,659 (10.5)	28,729 (10.3)
35-44	34,174 (12.5)	39,392 (14.2)
45-54	41,268 (15.1)	44,794 (16.1)
55-64	58,183(21.4)	56,628 (20.3)
65 and older	96,349 (35.4)	97,501 (35.1)
Residential area		
Urban	230,874 (84.7)	237,673 (85.5)
Rural	41,669 (15.3)	40,391 (14.5)
Education level		
Attended high school or less	14,805 (5.4)	13,065 (4.7)
Graduated from high school	67,247 (24.7)	65,609 (23.6)
Attended college or above	77,622 (28.5)	77,630 (27.9)
Graduated from college or above	112,869 (41.4)	121,760 (43.8)
Marital status		
Single	52,796 (19.4)	54,306 (19.5)
Married	146,251 (53.7)	152,435 (54.8)
Divorced, widowed, or others	73,496 (27)	71,323 (25.6)
Employment		
Employed	145,075 (53.3)	152,050 (54.7)
Unemployed	27,083 (9.9)	26,550 (9.5)
Retired or unable to work	100,385 (36.8)	99,464 (35.8)
Household income (US $)		
Less than 15,000	22,051 (8.1)	15,772 (5.6)
15,000 to <25,000	39,395 (14.5)	26,772 (9.6)
25,000 to <35,000	26,803 (9.8)	33,851 (12.2)
35,000 to <50,000	37,991 (13.9)	38,514 (13.9)
50,000 or more	146,303 (53.7)	163,155 (58.7)
Housing		
Own	199,188 (73.1)	206,029 (74.1)
Rent	62,790 (23)	62,854 (22.6)
Other arrangement	10,565 (3.9)	9181 (3.3)
Health insurance		
Yes	252,202 (92.5)	264,984 (95.3)
No	20,341 (7.5)	13,080 (4.7)
Health and related behaviors		
Perceived health		

Variables	2019 (n=272,543), n (%)	2021 (n=278,064), n (%)
Excellent	44,293 (16.3)	49,072 (17.6)
Very good	94,103 (34.5)	98,468 (35.4)
Good	84,798 (31.1)	86,112 (31)
Fair	36,038 (13.2)	33,517 (12.1)
Poor	13,311 (4.9)	10,895 (3.9)
Chronic health condition		
Hypertension	110,699 (40.6)	165,449 (59.5)
High cholesterol	95,634 (35.1)	152,806 (55)
Diabetes	37,048 (13.6)	37,910 (13.6)
Asthma	25,755 (9.4)	27,300 (9.8)
COPD[a]	22,545 (8.3)	22,261 (8)
Depression	53,900 (19.8)	57,284 (20.6)
Report of comorbidities		
None	91,235 (33.5)	29,563 (10.6)
2 and more	101,630 (37.3)	159,570 (57.4)
Health behaviors		
Maintaining optimal body weight	83,875 (30.8)	80,754 (29)
Performing physical activity	204,733 (75.1)	214,943 (77.3)
Healthy diet	44,028 (16.2)	40,444 (14.5)
Risk behaviors		
Current smoker	38,662 (14.2)	36,983 (13.3)
Excessive alcohol user	44,171 (16.2)	44,798 (16.1)
Practicing healthy lifestyles		
Yes	10,955 (4)	10,139 (3.6)
No	261,590 (96)	267,925 (96.4)

[a] COPD: chronic obstructive pulmonary disease.

Predictors of Practicing Healthy Lifestyles

Table 2 shows the results of the logistic regression analysis for cases without missing data in 2019 and 2021 who practiced healthy lifestyles; similar results were obtained for 2019 and 2021. In 2021, female individuals (odds ratio [OR] 1.95; *P*<.001) who had a high level of education (OR 1.99; *P*<.001), were single (OR 1.10; *P*=.02), unemployed (OR 1.10; *P*=.001), and had perceived good health (OR 1.62; *P*<.001) were found to have a higher likelihood of practicing healthy lifestyles. Compared with adults aged 65 years and above, young adults aged 18-64 years tended not to practice healthy lifestyles (OR 0.53-0.66; *P*<.001). Individuals with a low household income have a low likelihood of practicing healthy lifestyles (OR 0.72-0.79; *P*<.001).

Individuals with chronic health conditions had a low likelihood of practicing healthy lifestyles. The data in Table 2 show that individuals with NCDs such as hypertension and diabetes had a low likelihood of practicing healthy lifestyles (OR 0.48-0.83; *P*<.001).

For the cases with imputations, the weighted estimates of the logistic regression analysis results from 5 imputed data sets are presented in Table 3 (n=438,693). The rates of missing outcomes of healthy lifestyles in each variable are reported in Table S1 in Multimedia Appendix 1; briefly, the rate varied from 14.1% (61,858/438,693) to 27.3% (119,823/438,693). The weighted estimates were in line with those for cases without missing data (Table 2). Multicollinearity was not detected in all regression analyses.

Table 2. Logistic regression analysis of practicing a healthy lifestyle among complete cases without missing data in 2019 and 2021.

Variables	2019 (n=272,543)		2021 (n=278,064)	
	Odds ratio	*P* value	Odds ratio	*P* value
Sex: female	2.03	<.001	1.95	<.001
Residence in urban areas	1.22	<.001	1.23	<.001
Education: college and above	1.90	<.001	1.99	<.001
Age range (years)				
18-24	0.80	<.001	0.59	<.001
25-34	0.55	<.001	0.53	<.001
35-44	0.54	<.001	0.58	<.001
45-54	0.56	<.001	0.54	<.001
55-64	0.69	<.001	0.66	<.001
65 and above	Reference		Reference	
Marital status				
Single	0.92	.09	1.10	.02
Married	1.06	.03	1.15	<.001
Divorced, widowed, or others	Reference		Reference	
Employment				
Employed	0.87	<.001	0.86	<.001
Unemployed	1.17	<.001	1.11	.001
Retired or unable to work	Reference		Reference	
Household income (US $)				
Less than 15,000	0.66	<.001	0.78	<.001
15,000 to <25,000	0.72	<.001	0.72	<.001
25,000 to <35,000	0.71	<.001	0.77	<.001
35,000 to <50,000	0.72	<.001	0.79	<.001
50,000 or more	Reference		Reference	<.001
Perceived good or better health	1.51	<.001	1.62	<.001
Hypertension	0.51	<.001	0.49	<.001
High cholesterol	0.74	<.001	0.71	<.001
Diabetes	0.49	<.001	0.48	<.001
Asthma	0.84	<.001	0.83	<.001
COPD[a]	0.68	<.001	0.58	<.001
Depression	0.73	<.001	0.71	<.001

[a] COPD: chronic obstructive pulmonary disease.

Table 3. Logistic regression analysis of practicing a healthy lifestyle among all cases with multiple imputation in 2021 (n=438,693).

Variables	B (SE)	Wald chi-square (*df*)	Odds ratio (95% CI)
Sex: female	0.628 (0.019)	1064.00 (1)	1.87 (1.80-1.95)[a]
Residence in urban areas	0.212 (0.026)	66.32 (1)	1.24 (1.17-1.30)[a]
Education: college and above	0.551 (0.025)	478.90 (1)	1.73 (1.65-1.82)[a]
Age range (years)			
18-24	–0.401 (0.047)	72.59 (1)	0.67 (0.61-0.73)[a]
25-34	–0.667 (0.037)	321.8 (1)	0.51 (0.47-0.55)[a]
35-44	–0.578 (0.036)	258.80 (1)	0.56 (0.52-0.60)[a]
45-54	–0.640 (0.037)	295.50 (1)	0.53 (0.49-0.46)[a]
55-64	–0.453 (0.029)	246.20 (1)	0.64 (0.60-0.67)[a]
65 and above	Reference		
Marital status			
Single	0.029 (0.030)	0.91 (1)	1.03 (0.97-1.09)
Married	0.098 (0.025)	15.55 (1)	1.10 (1.05-1.16)[a]
Divorced, widowed, or others	Reference		
Employment			
Employed	–0.148 (0.027)	30.58 (1)	0.86 (0.81-0.90)[a]
Unemployed	0.114 (0.033)	11.64 (1)	1.12 (1.05-1.20)[a]
Retired or unable to work	Reference		
Household income (US $)			
Less than 15,000	–0.294 (0.082)	12.70 (1)	0.75 (0.63-0.87)[b]
15,000 to <25,000	–0.302 (0.099)	9.35 (1)	0.74 (0.61-0.89)[c]
25,000 to <35,000	–0.267 (0.042)	39.45 (1)	0.77 (0.70-0.83)[a]
35,000 to <50,000	–0.245 (0.034)	52.58 (1)	0.78 (0.73-0.83)[a]
50,000 or more	Reference		
Perceived good or better health	0.465 (0.034)	191.60 (1)	1.59 (1.49-1.70)[a]
Hypertension	–0.674 (0.022)	917.20 (1)	0.51 (0.48-0.53)[a]
High cholesterol	–0.307 (0.019)	254.90 (1)	0.74 (0.70-0.76)[a]
Diabetes	–0.729 (0.04)	339.60 (1)	0.48 (0.44-0.52)[a]
Asthma	–0.168 (0.034)	23.92 (1)	0.85 (0.79-0.90)[a]
COPD[d]	–0.500 (0.048)	106.90 (1)	0.61 (0.55-0.66)[a]
Depression	–0.347 (0.025)	197.60 (1)	0.71 (0.67-0.74)[a]

[a] Significant result: *P* value <.001

[b] Significant result: .001≤ *P* value <.01

[c] Significant result: .01≤ *P* value <.05

[d] COPD: chronic obstructive pulmonary disease.

Discussion

Principal Findings

This study assessed the determinants of practicing healthy lifestyles at the national level using 2019 and 2021 BRFSS survey data. The distribution of sociodemographic variables and health and related behaviors were similar for both the surveys. However, the tendency to practice healthy lifestyles slightly decreased in 2021. The logistic regression analysis results indicated that female individuals residing in urban areas with a higher level of education were more likely to practice healthy lifestyles, whereas younger individuals with a low household income were less likely to do so. Although chronic health conditions hindered individuals from practicing healthy lifestyles, those with perceived good or better health were more likely to practice healthy lifestyles.

However, the overall rate of practicing a healthy lifestyle was still very low. Fang et al [22] analyzed 2013 BRFSS data and reported that only 1.7% of the assessed individuals with hypertension practiced healthy lifestyles. Other studies have shown that 2.1% of healthy adults and 6% of health care providers practiced healthy lifestyles [5,6]. Moreover, although more individuals performed physical activity, the percentage of individuals with optimal body weights and following healthy diets decreased. High BMI and unhealthy diet were major risk factors for disability and premature death worldwide [23,24]. In this study, we found that the individuals' residences considerably affected their decisions to practice healthy lifestyles. Rural residents reported that they received limited access to health-related information from health care providers [25]. Although the overall number of health care providers, such as physicians, nurse practitioners, and physician assistants, in the United States significantly increased between 2009 and 2017, the urban-rural disparities remained significant; the number and density of health care providers increased to a greater extent in urban areas than in rural areas [26]. Therefore, a healthy lifestyle should be strongly promoted, and urban-rural health disparities should be tackled urgently.

In concordance with previous studies [9,11,27], our results showed that individuals with a low household income, who were male, and had high school or lower levels of education were less likely to practice healthy lifestyles. Although married people were more likely to practice healthy lifestyles, results from another national survey revealed that such individuals were more likely to be overweight or obese [28]. In contrast, we found that young people (aged <65 years) were less likely to practice healthy lifestyles. Thus, early health promotion is required because the risk of developing NCDs exponentially increases after 50 years of age [29].

Practicing healthy lifestyles among individuals with hypertension were focused because hypertension has been the leading risk factor for global disease burden worldwide for a decade [23,30]. Although certain guidelines recommend healthy lifestyles and medical treatments to effectively control blood pressure [31-33], our analysis showed that individuals with hypertension were less likely to practice healthy lifestyles. A Finnish cohort study involving 41,225 participants followed up over 4 years reported that the initiation of antihypertensive treatment significantly decreased the tendency of individuals with hypertension to practice healthy lifestyles (OR 0.55-0.92) [34]. As blood pressure can indeed be controlled with medications, less attention has been paid to individuals with hypertension practicing healthy lifestyles. Approximately 82.3% of individuals with hypertension in the United States use antihypertensive medications [35]; they might not practice healthy lifestyles as their blood pressure is medically controlled.

Concern about around those with diabetes practicing healthy lifestyles have been raised, because these individuals showed the least likelihood of doing so. A healthy lifestyle has been indicated in various diabetes management guidelines [36,37]. However, our findings revealed that individuals with diabetes were the least likely to practice healthy lifestyle, showing the smallest OR among all variables. Individuals with diabetes have emphasized that education and support from health care providers can enhance their self-care tendencies [38]. Hence, new methods for diabetes care should be used.

Implications

Our study findings have strong implications for clinical practice, research, and policy making. The low likelihood of practicing healthy lifestyles highlights the need for strengthening current practices. Health education supported by printed materials, phone calls, and SMS text messaging may improve individuals' adherence to healthy lifestyles and medical treatments [39-43]. Given the rapid increase in smartphone ownership and advancing technologies [44], digital interventions such as website and smartphone apps have been shown to slightly improve adherence to a healthy lifestyle [45-47]. Supportive methods and digital interventions can be integrated into current health promotion activities to facilitate healthy lifestyles and widen the possibility of receiving health-related information and support from health care providers, especially for rural residents. Moreover, digital interventions can help overcome the barriers to public health measures caused by COVID-19 [18]. However, we noted that several sociodemographic characteristics can hinder individuals from practicing healthy lifestyles. Additionally, there may be associations between depression, ethnic background, and socioeconomic status [48]. Future studies can explore the associations and focus on developing specific interventions that promote healthy lifestyles among such individuals.

Rural residents showed a low likelihood of practicing healthy lifestyles; this highlights the need of support from policy makers. A cohort study of 25,014 middle-age and older adults with a 20-year follow-up revealed that performing any type of physical activity could significantly reduce the likelihood of hospitalization to a mean of 0.42 days per year; this reduced likelihood of hospitalization could reduce health care expenditures by approximately 7% [49]. Moreover, some amount of the saved health care expenditure could be granted as tax deductions to individuals who participate in healthy lifestyle programs and provide evidence of practicing such lifestyles (ie, reduced body weight). Such tax deductions can attract more people toward healthy lifestyles, further decreasing the overall health care expenditure. The shortage of health care providers may be managed through the use of digital interventions for providing education and support to residents in rural areas. However, more interventions should be developed to attract health care providers to work and remain in rural areas. For example, increased training in rural areas, the payment of locum relief, and loan repayment may help do so [50].

Limitations

This study has 4 key limitations. First, the variables in the secondary analysis were limited to the data set used. For example, only 1 question in the data set investigated the individuals' physical activity habits. Moreover, the question was rather general and inquired whether the individuals performed any type of physical activity or exercise during the past 30 days. Two essential components—frequency and intensity—were not explored. Thus, our study findings might not adequately represent the level of physical activity as recommended by the WHO [51]. Second, all variables were reported by individuals via phone call. Therefore, we might have encountered reporting biases arising from the participant's recall from memory, and the interviewers may have recorded the responses incorrectly. To address such issues, the CDC has created a program to identify errors or conditions of concern due to data entry by the interviewers [20,21]. Third, although the phone numbers used in each year were not repeated, it was unclear if some numbers would be used in the other survey years. Hence, it was possible that some people might have responded to both surveys in this study. Fourth, the overall rate of missing data was high at 36.6% in 2021, which could have impacted the generalizability of the results. However, we performed multiple imputations, and the pooled results of the imputed data were in line with those from the cases without missing outcomes. Hence, our imputed data provided an accurate estimation of the influence of sociodemographic variables and chronic health conditions on practicing healthy lifestyles and enhanced the generalizability of our findings.

Conclusions

Practicing a healthy lifestyle can have numerous health benefits across all age groups. We found that very few people practiced healthy lifestyles before and during COVID-19. Some sociodemographic variables and chronic health conditions were identified as significant factors affecting individuals' decisions to practice healthy lifestyles. Several interventions were suggested to strengthen the promotion of healthy lifestyles at the community level and to attract more health care providers to work and remain in rural areas.

Acknowledgments

The authors wish to thank Dr Erik Fung for his valuable comments.

Data Availability

The data used in this study (Behavioral Risk Factor Surveillance System [BRFSS] 2019 and 2021) are publicly available [20,21].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Missing outcome of healthy lifestyles in each variable.
[DOCX File , 20 KB - publichealth_v9i1e45697_app1.docx]

References

1. The top 10 causes of death. World Health Organization. 2022 Dec 9. URL: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [accessed 2022-12-18]
2. Schlesinger S, Neuenschwander M, Ballon A, Nöthlings U, Barbaresko J. Adherence to healthy lifestyles and incidence of diabetes and mortality among individuals with diabetes: a systematic review and meta-analysis of prospective studies. J Epidemiol Community Health 2020 May 19;74(5):481-487. [doi: 10.1136/jech-2019-213415] [Medline: 32075860]
3. Bai G, Zhang J, Zhao C, Wang Y, Qi Y, Zhang B. Adherence to a healthy lifestyle and a DASH-style diet and risk of hypertension in Chinese individuals. Hypertens Res 2017 Feb 08;40(2):196-202. [doi: 10.1038/hr.2016.119] [Medline: 27604345]
4. Loef M, Walach H. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. Prev Med 2012 Sep;55(3):163-170. [doi: 10.1016/j.ypmed.2012.06.017] [Medline: 22735042]
5. Zhu N, Yu C, Guo Y, Bian Z, Han Y, Yang L, China Kadoorie Biobank Collaborative Group. Adherence to a healthy lifestyle and all-cause and cause-specific mortality in Chinese adults: a 10-year prospective study of 0.5 million people. Int J Behav Nutr Phys Act 2019 Nov 04;16(1):98 [FREE Full text] [doi: 10.1186/s12966-019-0860-z] [Medline: 31685026]
6. Li Y, Schoufour J, Wang DD, Dhana K, Pan A, Liu X, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020 Jan 08;368:l6669 [FREE Full text] [doi: 10.1136/bmj.l6669] [Medline: 31915124]

7. Visser M, Wijnhoven HAH, Comijs HC, Thomése FGCF, Twisk JWR, Deeg DJH. A healthy lifestyle in old age and prospective change in four domains of functioning. J Aging Health 2019 Aug 29;31(7):1297-1314 [FREE Full text] [doi: 10.1177/0898264318774430] [Medline: 29809092]
8. Noncommunicable diseases. World Health Organization. 2022 Sep 16. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2022-12-22]
9. Azizan NA, Thangiah N, Su T, Majid HA. Does a low-income urban population practise healthy dietary habits? Int Health 2018 Mar 01;10(2):108-115. [doi: 10.1093/inthealth/ihy001] [Medline: 29462331]
10. Mirahmadizadeh A, Khorshidsavar H, Seif M, Sharifi MH. Adherence to medication, diet and physical activity and the associated factors amongst patients with type 2 diabetes. Diabetes Ther 2020 Feb 08;11(2):479-494 [FREE Full text] [doi: 10.1007/s13300-019-00750-8] [Medline: 31916213]
11. Alefan Q, Huwari D, Alshogran OY, Jarrah MI. Factors affecting hypertensive patients' compliance with healthy lifestyle. Patient Prefer Adherence 2019 Apr 23;13:577-585 [FREE Full text] [doi: 10.2147/PPA.S198446] [Medline: 31114171]
12. Tam HL, Chung SF, Wang Q. Urban-rural disparities in hypertension management among middle-aged and older patients: results of a 2018 Chinese national study. Chronic Illn 2022 May 22:17423953221102627. [doi: 10.1177/17423953221102627] [Medline: 35603631]
13. Muhammad T, Srivastava S, Hossain B, Paul R, Sekher TV. Decomposing rural-urban differences in successful aging among older Indian adults. Sci Rep 2022 Apr 19;12(1):6430 [FREE Full text] [doi: 10.1038/s41598-022-09958-4] [Medline: 35440788]
14. Cosby AG, McDoom-Echebiri MM, James W, Khandekar H, Brown W, Hanna HL. Growth and persistence of place-based mortality in the United States: the rural mortality penalty. Am J Public Health 2019 Jan;109(1):155-162. [doi: 10.2105/AJPH.2018.304787] [Medline: 30496008]
15. Aburto JM, Schöley J, Kashnitsky I, Zhang L, Rahal C, Missov TI, et al. Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries. Int J Epidemiol 2022 Feb 18;51(1):63-74 [FREE Full text] [doi: 10.1093/ije/dyab207] [Medline: 34564730]
16. Agrawal S, Makuch S, Dróżdż M, Strzelec B, Sobieszczańska M, Mazur G. The impact of the COVID-19 emergency on life activities and delivery of healthcare services in the elderly population. J Clin Med 2021 Sep 10;10(18):4089 [FREE Full text] [doi: 10.3390/jcm10184089] [Medline: 34575200]
17. Sum KK, Cai S, Law E, Cheon B, Tan G, Loo E, et al. COVID-19-related life experiences, outdoor play, and long-term adiposity changes among preschool- and school-aged children in Singapore 1 year after lockdown. JAMA Pediatr 2022 Mar 01;176(3):280-289 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.5585] [Medline: 35072692]
18. Young DR, Hong BD, Lo T, Inzhakova G, Cohen DA, Sidell MA. The longitudinal associations of physical activity, time spent outdoors in nature and symptoms of depression and anxiety during COVID-19 quarantine and social distancing in the United States. Prev Med 2022 Jan;154:106863 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106863] [Medline: 34774881]
19. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Prev Med 2007 Oct;45(4):247-251. [doi: 10.1016/j.ypmed.2007.08.012] [Medline: 17950122]
20. 2021 BRFSS survey data and documentation. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/brfss/annual_data/annual_2021.html [accessed 2022-09-22]
21. 2019 BRFSS survey data and documentation. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/brfss/annual_data/annual_2019.html [accessed 2022-09-22]
22. Fang J, Moore L, Loustalot F, Yang Q, Ayala C. Reporting of adherence to healthy lifestyle behaviors among hypertensive adults in the 50 states and the District of Columbia, 2013. J Am Soc Hypertens 2016 Mar;10(3):252-262.e3 [FREE Full text] [doi: 10.1016/j.jash.2016.01.008] [Medline: 26851000]
23. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
24. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: 10.1016/S0140-6736(19)30041-8] [Medline: 30954305]
25. Chen X, Orom H, Hay JL, Waters EA, Schofield E, Li Y, et al. Differences in rural and urban health information access and use. J Rural Health 2019 Jun 16;35(3):405-417 [FREE Full text] [doi: 10.1111/jrh.12335] [Medline: 30444935]
26. Zhang D, Son H, Shen Y, Chen Z, Rajbhandari-Thapa J, Li Y, et al. Assessment of changes in rural and urban primary care workforce in the United States from 2009 to 2017. JAMA Netw Open 2020 Oct 01;3(10):e2022914 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.22914] [Medline: 33112401]
27. Chang S, Chang Y, Wu L. Gender differences in lifestyle and risk factors of metabolic syndrome: do women have better health habits than men? J Clin Nurs 2019 Jun 06;28(11-12):2225-2234. [doi: 10.1111/jocn.14824] [Medline: 30786102]

28. Schoeppe S, Vandelanotte C, Rebar AL, Hayman M, Duncan MJ, Alley SJ. Do singles or couples live healthier lifestyles? trends in Queensland between 2005-2014. PLoS One 2018 Feb 28;13(2):e0192584 [FREE Full text] [doi: 10.1371/journal.pone.0192584] [Medline: 29489832]
29. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
30. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1923-1994 [FREE Full text] [doi: 10.1016/S0140-6736(18)32225-6] [Medline: 30496105]
31. Whelton PK, Carey RM, Aronow WS, Casey DEJ, Collins KJ, Himmelfarb CD, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. Hypertension 2018 Jun;71(6):e13-e115. [doi: 10.1161/hyp.0000000000000076] [Medline: 29133356]
32. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, ESC Scientific Document Group. 2018 ESC/ESH guidelines for the management of arterial hypertension. Eur Heart J 2018 Sep 01;39(33):3021-3104 [FREE Full text] [doi: 10.1093/eurheartj/ehy339] [Medline: 30165516]
33. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.120.15026] [Medline: 32370572]
34. Korhonen MJ, Pentti J, Hartikainen J, Ilomäki J, Setoguchi S, Liew D, et al. Lifestyle changes in relation to initiation of antihypertensive and lipid-lowering medication: a cohort study. J Am Heart Assoc 2020 Feb 18;9(4):e014168 [FREE Full text] [doi: 10.1161/JAHA.119.014168] [Medline: 32019405]
35. Samanic CM, Barbour KE, Liu Y, Wang Y, Fang J, Lu H, et al. Prevalence of self-reported hypertension and antihypertensive medication use by county and rural-urban classification - United States, 2017. MMWR Morb Mortal Wkly Rep 2020 May 08;69(18):533-539 [FREE Full text] [doi: 10.15585/mmwr.mm6918a1] [Medline: 32379728]
36. American Diabetes Association. Standards of medical care in diabetes-2022 abridged for primary care providers. Clin Diabetes 2022 Jan;40(1):10-38 [FREE Full text] [doi: 10.2337/cd22-as01] [Medline: 35221470]
37. IDF clinical practice recommendations for managing type 2 diabetes in primary care. International Diabetes Federation. 2017. URL: https://www.idf.org/component/attachments/attachments.html?id=1586&task=download [accessed 2022-12-22]
38. Ndjaboue R, Chipenda Dansokho S, Boudreault B, Tremblay M, Dogba MJ, Price R, et al. Patients' perspectives on how to improve diabetes care and self-management: qualitative study. BMJ Open 2020 Apr 29;10(4):e032762 [FREE Full text] [doi: 10.1136/bmjopen-2019-032762] [Medline: 32354775]
39. Tam HL, Wong EML, Cheung K. Effectiveness of educational interventions on adherence to lifestyle modifications among hypertensive patients: an integrative review. Int J Environ Res Public Health 2020 Apr 07;17(7):2513 [FREE Full text] [doi: 10.3390/ijerph17072513] [Medline: 32272601]
40. Yamaoka K, Tango T. Effects of lifestyle modification on metabolic syndrome: a systematic review and meta-analysis. BMC Med 2012 Nov 14;10(1):138 [FREE Full text] [doi: 10.1186/1741-7015-10-138] [Medline: 23151238]
41. Tam HL, Leung LYL, Wong EML, Cheung K, Chan ASW. Integration of text messaging interventions into hypertension management among older adults: A systematic review and meta-analysis. Worldviews Evid Based Nurs 2022 Feb 11;19(1):16-27. [doi: 10.1111/wvn.12549] [Medline: 35014147]
42. Head KJ, Noar SM, Iannarino NT, Grant Harrington N. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
43. Conn VS, Ruppar TM. Medication adherence outcomes of 771 intervention trials: systematic review and meta-analysis. Prev Med 2017 Jun;99:269-276 [FREE Full text] [doi: 10.1016/j.ypmed.2017.03.008] [Medline: 28315760]
44. Taylor K, Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. 2019 Feb 5. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2022-12-18]
45. Stavric V, Kayes NM, Rashid U, Saywell NL. The effectiveness of self-guided digital interventions to improve physical activity and exercise outcomes for people with chronic conditions: a systematic review and meta-analysis. Front Rehabil Sci 2022 Jun 24;3:925620 [FREE Full text] [doi: 10.3389/fresc.2022.925620] [Medline: 36188933]
46. Ma Y, Cheng HY, Cheng L, Sit JW. The effectiveness of electronic health interventions on blood pressure control, self-care behavioural outcomes and psychosocial well-being in patients with hypertension: a systematic review and meta-analysis. Int J Nurs Stud 2019 Apr;92:27-46. [doi: 10.1016/j.ijnurstu.2018.11.007] [Medline: 30690164]
47. Wong EM, Leung DYP, Chair S, Sit JWH. Effects of a web-based educational support intervention on total exercise and cardiovascular risk markers in adults with coronary heart disease. Worldviews Evid Based Nurs 2020 Aug 08;17(4):283-292. [doi: 10.1111/wvn.12456] [Medline: 32772509]

48. Budhwani H, Hearld KR, Chavez-Yenter D. Depression in racial and ethnic minorities: the impact of nativity and discrimination. J Racial Ethn Health Disparities 2015 Mar 11;2(1):34-42. [doi: 10.1007/s40615-014-0045-z] [Medline: 26863239]
49. Luben R, Hayat S, Wareham N, Pharoah P, Khaw K. Usual physical activity and subsequent hospital usage over 20 years in a general population: the EPIC-Norfolk cohort. BMC Geriatr 2020 May 06;20(1):165 [FREE Full text] [doi: 10.1186/s12877-020-01573-0] [Medline: 32375672]
50. Russell D, Mathew S, Fitts M, Liddle Z, Murakami-Gold L, Campbell N, et al. Interventions for health workforce retention in rural and remote areas: a systematic review. Hum Resour Health 2021 Aug 26;19(1):103 [FREE Full text] [doi: 10.1186/s12960-021-00643-7] [Medline: 34446042]
51. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020 Nov 25. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2022-12-20]

Abbreviations

BRFSS: Behavioral Risk Factor Surveillance System
CDC: Centers for Disease Control and Prevention
MICE: multiple imputation via chained equations
NCD: noncommunicable disease
OR: odds ratio
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 12.01.23; peer-reviewed by H Sakr, G Gonzalez-Escamilla; comments to author 15.02.23; revised version received 28.02.23; accepted 15.03.23; published 31.03.23.

Please cite as:
Tam HL, Chair SY, Leung ISH, Leung LYL, Chan ASW
US Adults Practicing Healthy Lifestyles Before and During COVID-19: Comparative Analysis of National Surveys
JMIR Public Health Surveill 2023;9:e45697
URL: https://publichealth.jmir.org/2023/1/e45697
doi:10.2196/45697
PMID:36940169

Original Paper

Joint Analysis of the Epidemic Evolution and Human Mobility During the First Wave of COVID-19 in Spain: Retrospective Study

Benjamin Steinegger[1*], PhD; Clara Granell[1*], PhD; Giacomo Rapisardi[1*], PhD; Sergio Gómez[1*], PhD; Joan Matamalas[2,3*], PhD; David Soriano-Paños[4*], PhD; Jesús Gómez-Gardeñes[4*], PhD; Alex Arenas[1*], PhD

[1]Universitat Rovira i Virgili, Tarragona, Spain
[2]Harvard Medical School, Boston, MA, United States
[3]Brigham and Women's Hospital, Boston, MA, United States
[4]Department of Condensed Matter Physics, University of Zaragoza, Zaragoza, Spain
[*]all authors contributed equally

Corresponding Author:
Alex Arenas, PhD
Universitat Rovira i Virgili
Avda Paisos Catalans 26
Tarragona, 43007
Spain
Phone: 34 977559687
Email: alexandre.arenas@urv.cat

Abstract

Background: The initial wave of the COVID-19 pandemic placed a tremendous strain on health care systems worldwide. To mitigate the spread of the virus, many countries implemented stringent nonpharmaceutical interventions (NPIs), which significantly altered human behavior both before and after their enactment. Despite these efforts, a precise assessment of the impact and efficacy of these NPIs, as well as the extent of human behavioral changes, remained elusive.

Objective: In this study, we conducted a retrospective analysis of the initial wave of COVID-19 in Spain to better comprehend the influence of NPIs and their interaction with human behavior. Such investigations are vital for devising future mitigation strategies to combat COVID-19 and enhance epidemic preparedness more broadly.

Methods: We used a combination of national and regional retrospective analyses of pandemic incidence alongside large-scale mobility data to assess the impact and timing of government-implemented NPIs in combating COVID-19. Additionally, we compared these findings with a model-based inference of hospitalizations and fatalities. This model-based approach enabled us to construct counterfactual scenarios that gauged the consequences of delayed initiation of epidemic response measures.

Results: Our analysis demonstrated that the pre–national lockdown epidemic response, encompassing regional measures and heightened individual awareness, significantly contributed to reducing the disease burden in Spain. The mobility data indicated that people adjusted their behavior in response to the regional epidemiological situation before the nationwide lockdown was implemented. Counterfactual scenarios suggested that without this early epidemic response, there would have been an estimated 45,400 (95% CI 37,400-58,000) fatalities and 182,600 (95% CI 150,400-233,800) hospitalizations compared to the reported figures of 27,800 fatalities and 107,600 hospitalizations, respectively.

Conclusions: Our findings underscore the significance of self-implemented prevention measures by the population and regional NPIs before the national lockdown in Spain. The study also emphasizes the necessity for prompt and precise data quantification prior to enacting enforced measures. This highlights the critical interplay between NPIs, epidemic progression, and human behavior. This interdependence presents a challenge in predicting the impact of NPIs before they are implemented.

(JMIR Public Health Surveill 2023;9:e40514) doi:10.2196/40514

KEYWORDS

epidemics; NPIs; nonpharmaceutical intervention; human behavior; Spain; COVID-19; mobility data; epidemic evolution; public health; surveillance; public health intervention; model-based inference

Introduction

Spain was among the strongest-hit countries worldwide during the first wave, with officially 28,000 fatalities attributed to COVID-19 [1]. In terms of excess deaths, there was an alarming increase of 80% [2], coupled with an attack rate of about 5% [3]. In response to the rapidly rising case numbers, the Spanish authorities implemented stringent containment measures to control the spread of SARS-CoV-2.

The first case in Spain was reported on January 31. The first local transmission was reported on February 26. On March 8, all autonomous communities (Comunidades Autonomas [CCAA]) had reported local transmission. The fast-rising case numbers led the authorities to impose a national lockdown on March 15. Bars, hotels, and restaurants had to close, and individuals were only allowed to leave their homes for work and essential shopping. However, the national lockdown was foregone by various regional measures. On March 10, Madrid and Álava (Basque Country) closed the entire educational system. Other CCAA followed shortly afterward. Madrid and Andalusia closed the gastronomy sector on March 13 and 14, respectively. Additional local measures consisted of the cancellation of festivities and football matches.

The Spanish government declared a state of emergency on March 14, a day before the lockdown took effect. Despite the lockdown being in place, case numbers were still rising toward the end of March [1]. As a result, the national authorities intensified the lockdown between March 28 and April 12, during which all nonessential economic activities came to a halt. The lockdown was then gradually lifted from May 2 onward under the competence of the local authorities (CCAA).

In short, the nonpharmaceutical interventions (NPIs) issued by the national and local authorities, which came along with a strong reduction in mobility, were eventually sufficient to mitigate the daily infections. A thorough analysis of the series of events that shaped the epidemic evolution during the first wave, as performed in various countries [4-10], is essential to design future mitigation strategies of COVID-19 or other emerging respiratory diseases [11,12].

The imposed NPIs and voluntarily practiced social distancing caused a strong reduction in contacts that was reflected by a decrease in mobility. Different studies show that mobility is a robust indicator for the evolution of the epidemic and, hence, the reproduction number [6,9,13-17]. The relationship was most evident during the early phase of the epidemic, when there was no efficient contact tracing in place and only minor use of face masks [16]. Accordingly, a joint analysis of mobility and epidemiological data can provide valuable insight into the evolution of the first wave in Spain.

Moreover, a combined examination provides an opportunity to investigate the factors influencing epidemic trends beyond the implementation of NPIs. Numerous earlier studies that assessed the epidemic's progression assumed sudden shifts in transmission rates as NPIs were implemented [4,18,19]. This method may potentially exaggerate the effects of NPIs, as it does not account for the voluntary behavior modifications of the population that could impact the epidemic's trajectory. By addressing this issue, the integrated analysis presented here offers a more comprehensive perspective on the forces driving the dynamics.

More specifically, in this study, we first evaluate large-scale mobile phone data to show how mobility evolved in the face of rising case numbers and during the lockdown. In particular, we analyze whether there was any change in mobility prior to the introduction of the lockdown. To determine the impact of mobility on the epidemic dynamics, we then contrast the evolution of mobility with epidemiological data, such as case numbers, hospitalizations, and fatalities. More specifically, we blend a direct analysis of the epidemiological data [5,18] and model-based inference [10,19-21]. Furthermore, leveraging the model-based inference, we evaluate the evolution of the underreporting of infections over time. Finally, although most previous studies have focused exclusively on the impact of NPIs, this model-based approach allows us also to build counterfactual scenarios and quantify how the epidemic response that anticipated the lockdown substantially reduced the impact of SARS-CoV-2 in Spain during the first wave.

Methods

Mobility Data

The data were provided by the Ministry of Transport, Mobility and Urban Agenda (MITMA) [22]. The raw data stemming from 1 mobile network provider consisted of the anonymized individual trajectories of about 13 million individuals. Beyond the intrinsic limitation of mobile phone data, we assumed that the 13 million individuals would provide a reasonable sample of the Spanish population. By using additional information, such as land usage, sociodemographic indicators, the transport network and the schedule of the public transport, the raw data were transformed into an origin-destination matrix by the MITMA. We directly used the origin-destination matrices and did not have access to the raw data. More details can be found in the methodological note provided by the MITMA [23]. Trips are recorded on the level of municipalities and aggregated on an hourly basis. Additionally, trips are separated into 6 different distance classes. To calculate the mobility reduction on a national or provincial level, we summed up the number of trips and compared it to the corresponding day during the reference period (February 14-20).

Reconstruction of Exposure Times

We reconstructed the exposure times for different autonomous communities and regions (CCAA) by using a deconvolution process with the symptom onset data. This method allowed us to estimate the time individuals were exposed to the virus before they started exhibiting symptoms, providing a more accurate understanding of the transmission dynamics within each region. By tracing back the exposure times, we can gain valuable insights into the infection patterns and better assess the effectiveness of various interventions and public health measures implemented across the CCAA. This information can then be used to improve and refine epidemic models and guide future decision-making to better control the spread of the virus. More specifically, we used the backprojNP function [24] from the *surveillance* package [25,26] in R. The method, initially proposed by Becker et al [27], infers the expected number of exposures, given the probability mass function of the incubation period, through a maximum likelihood deconvolution approach.

Credible intervals were calculated based on a bootstrap procedure [28]. We fixed the smoothing factor k=6, which corresponds to a centered rolling average of days. The bootstrap procedure made use of 1000 samples (B=1000). The incubation period distribution was fixed as a gamma distribution with mean 5.2 and SD 2.8 days [29]. The time series with symptom onsets was provided by the Centro Nacional de Epidemiología [1].

Estimation of Rt

From the median incidence, obtained by the reconstruction of the exposure times, we estimated the evolution of Rt using the *EpiEstim* package [30,31]. For the infectivity profile, we chose the generation time estimated by Ganyani et al [32]. To be more precise, we assumed a generation time following a gamma distribution with mean 5.2 (95% CI 3.78-6.78) days and SD 1.72 (95% CI 0.91-3.93) days. This generation time distribution corresponds to the estimation by Ganyani et al [32] for Singapore, while assuming the same incubation period distribution [29] as we did in the reconstruction of the exposure times. We assumed SDs of 1.0 and 1.2 days for the mean and the generation time, respectively. However, we bound the values for the mean and SD by the estimations of Ganyani et al [32]. We fixed a centered rolling average of 7 days. To bootstrap the credible intervals, we took 100 samples of the generation time distribution and considered 100 posteriors for each of these samples (n1=100 and n2=100, respectively).

Identifying the Linear Segments of Rt

To identify the linear segments, we use the R package *segmented* [33,34]. The method proposed by Muggeo [33,34] implements a maximum likelihood approach using linear predictors. The credible intervals are obtained through bootstrapping. As previously pointed out, we assumed 3 segments of Rt. An initial constant value R1 as the disease was spreading freely in Spain, as well as 2 linear evolving parts corresponding to the decrease in Rt toward the lockdown and the constant decrease observed during the lockdown.

Model

An alternative method for investigating the evolution of Rt is to perform model-based inference [10,19,20]. By fitting a minimal model to daily fatalities and hospitalizations, we can compare the results with those derived from the reported number of infections and assess whether the early decrease in Rt can be attributed to saturation in testing capacity. This approach is more reliable as fatalities and hospitalizations are less susceptible to fluctuating reporting rates. However, daily fatalities experience significant underreporting. Although the official number of COVID-19 fatalities during the first wave is approximately 28,000, excess deaths amount to around 50,000 [2]. Nonetheless, the reported fatalities and excess deaths follow similar trends, with excess deaths exhibiting a slower decline (Figure S13, Multimedia Appendix 1). Consequently, we considered the reported fatalities to be a sufficiently robust data stream.

We opted for a discrete-time model, informed by empirically derived distributions for the generation time [32], incubation time [29], and time from symptom onset to hospitalization and death [35-37]. The time between symptom onset and hospitalization or death varies significantly by age [35] and region [36]. Since our data were aggregated by age and location, we did not incorporate age stratification or geographical heterogeneity through metapopulations.

Moreover, we assumed an instantaneous reproduction number Rt^M with a functional form. The notation Rt^M distinguishes this from the reproduction number inferred from reported infections, Rt. The functional form of Rt^M was inspired by the Rt inferred from the reported infections.

We divided Rt^M into 3 linear segments: the “free” spreading phase before restrictions (constant value R1), the initiation of the epidemic response (linear decrease at time BP), and the lockdown (constant value R2) on March 15. We referred to the intersection between R1 and R2, or the moment Rt^M began to decrease, as the breakpoint BP. Along with R1, R2, and the initial number of infected individuals I0, the initiation of the epidemic response remained a free parameter. The assumption that Rt^M would reach a constant value R2 upon lockdown implementation was re-evaluated in the sensitivity analysis. This framework allowed us to assess the plausibility of an early decrease around March 5/6, as found in the reported infections.

We decided against using intensive care unit (ICU) admissions as a data stream for our inference. During the first wave, some CCAA reported current occupancy, while others reported new admissions, and many changed their reporting criteria over time. Additionally, only in Madrid and Catalonia, the hardest-hit CCAA, did we observe an earlier peak in the 70-79-year age group compared to younger groups, followed by a rapid decrease in ICU admissions (Figure S14, Multimedia Appendix 1). This strongly suggests that ICU admission criteria were adjusted due to health care system overload. Another indication of this is that hospital admissions and ICU admissions peak on the same day, even though health authorities report a 3-day delay from symptom onset to ICU admission compared to hospital admissions (Figure S15, Multimedia Appendix 1). These factors suggest that ICU admissions data are not a reliable data stream.

Given the form of Rt^M, the generation time distribution w(t), and the size of the population N, the daily incidence I_t on day t evolves as [38]:

We used the same generation time distribution as that for estimating Rt [32]. From the daily incidence, we propagated the symptom onset as well as the daily fatalities through convolution [14,19]. Given the incubation time distribution P(t), and the distribution of time from symptom onset to hospitalization H(t) and to death D(t), the daily number of individuals developing symptoms S_t, the daily hospitalizations G_t, and the daily fatalities F_t evolve as:

Model Fitting

We fixed the incubation time distribution as that for the estimation of exposure times [29]. We fixed the infection hospitalization ratio (IHR) and the infection fatality ratio (IFR) by dividing the total number of hospitalizations and fatalities, respectively, by the total number of infected individuals. The latter was fixed through a nationwide seroprevalence study that found an attack rate of 5% by the end of April 2020 [39]. This led to an IHR and an IFR of 4.54% and 1.18%, respectively.

We assumed the times from symptom onset to hospitalization and to death to follow a gamma distribution. Since case line data were not available, we fixed the shape factor as 2.5 and 2.2, respectively, according to Hawryluk et al [35]. However, the health authorities published the median and IQR of these distributions [37]. They reported a median time from symptom onset to hospitalization and death of 6 (IQR 3-9) and 11 (IQR 7-17) days, respectively. We fixed the scale values by performing a least-squares fit with respect to these data. This resulted in a shape and scale parameter for time between symptom onset to hospitalization and death of 2.68 and 5.85, respectively.

We adjusted the model to the daily hospitalizations and fatalities through a Markov chain Monte Carlo (MCMC) approach—to be more precise, through Hamiltonian Monte Carlo [40]. For the log-likelihood, we chose a negative binomial distribution with a dispersion parameter that was left as a free parameter. Motivated by previous findings [13,19], we fixed the prior of R1 as a normal distribution with mean 4.5 (SD 1.0). The prior of R2 was a uniform distribution between 0.4 and 1.0. Similarly, the prior of BP was flat between and days before lockdown. The initial number of infected individuals I0 was uniform between and 5000. We set the prior for the dispersion as a normal distribution with mean 10.0 and SD 5.0. The daily fatalities were not accurately adjusted using directly the distribution received from the least-squares fit. Therefore, we added the scale and shape parameter of the distribution for time between symptom onset and death as a free parameter. We fixed the prior as a normal distribution with a mean as that found from the least-squares fit and SD 0.1. The form of the inferred distribution from symptom onset to death is shown in Figure S20, Multimedia Appendix 1. The model was implemented in Stan [41,42]. We ran 6 chains with 4000 iterations, where 2000 iterations were used for warm-up. Gelman-Rubin convergence statistics [43], that is, potential scale reduction factors, were all smaller than 1.001. Posteriors and trace plots are shown in Figures S21 and S22, Multimedia Appendix 1, respectively.

Counterfactual Scenarios

A common question when analyzing the pandemic's evolution retrospectively is, What would have happened if we had acted earlier or later? Several studies have examined the impact of an earlier or later lockdown [4,7,20]. These detailed modeling efforts allow us to address this question by considering a shift in the entire epidemic response [7,20]. If the reproduction number is initially constant, as was the case in our analysis, the effect of shifting the epidemic response is entirely determined by R1, that is, the initial doubling time. In other words, if one shifts the epidemic response by the number of days equivalent to the doubling time, the attack rate, hospitalizations, and fatalities double.

Here, however, we focused on the epidemic response that occurred before the lockdown implementation. This encompassed regional measures, individual awareness, nationwide educational system closures, and the lockdown announcement. Essentially, we moved the breakpoint BP by a specific number of days. If the shift surpassed the lockdown date, the BP remained fixed on March 15.

Sensitivity Analysis

The first part of the sensitivity analysis consisted of relaxing the assumption that the reproduction number would reach a stable value (R2) on March 15, the day the lockdown took effect. To do so, we introduced a second breakpoint BP2 that defined when Rt^M would reach R2. For BP2, we chose a flat prior from March 15 onward. The detailed results of the fit are shown in Figures S16-S18, Multimedia Appendix 1. The other estimated parameters, as well as the counterfactual scenarios, were robust. The median of BP2 was found half a day later than March 15, with a credible interval between March 15 and 17. This added plausibility to our assumption that Rt^M would reach R2 on March 15. Furthermore, we considered a second generation time distribution taken from Ferretti et al [44]. The generation time corresponded to a Weibull distribution, compared to a gamma distribution in the main text, with mean 5.0 and SD 1.9 days. The results did not substantially change and are presented in Figures S23 and S24, Multimedia Appendix 1.

Ethical Considerations

The Research Ethics Committee of Rovira i Virgili University deemed this study as exempt from ethics review (URV.F01.04.00 ALTRES-2023-PRD-0001). The mobility data from the mobile phone provider were anonymized and aggregated. Individual traces were not available. No other ethical considerations apply to this work.

Results

The Evolution of Mobility

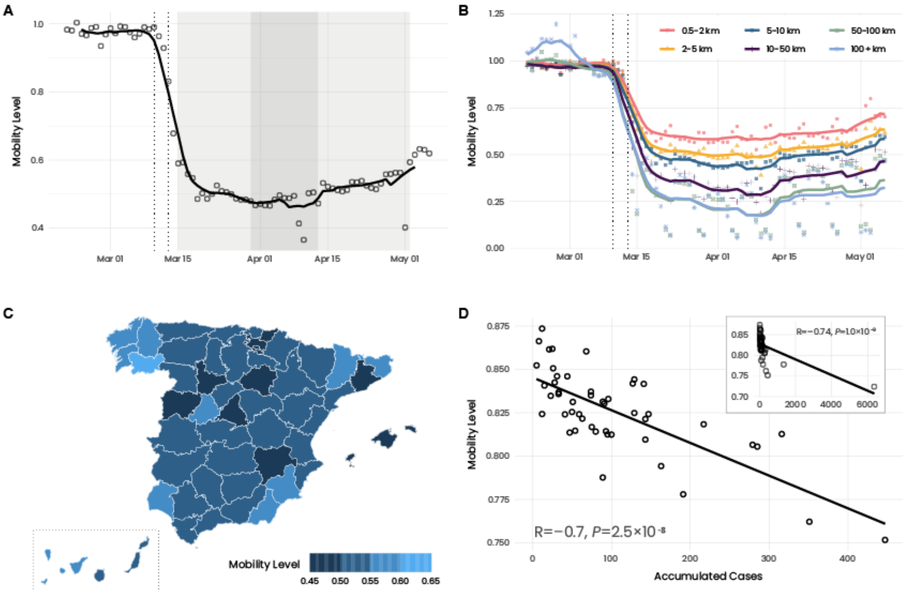
As highlighted previously, mobility is an indicator for the impact of NPIs. Here, we analyzed a data set provided by the MITMA that analyzed the evolution of mobility through anonymized mobile phone data from about 13 million users (see the Methods section for more details). The data indicated that already before the lockdown, there was a substantial reduction in mobility from March 10 onward (Figure 1A). The initiation of this decrease coincided with the first regional NPIs introduced in Madrid and in the Basque country. Furthermore, the early reduction was consistent with an increase in COVID-19–related searches on Google (Figure S1, Multimedia Appendix 1). During the lockdown, mobility was about 50% of the prepandemic level. Mobility increased from the end of the stronger lockdown toward the end of lockdown to about 60%.

Here, we referred to the aggregated, nationwide number of trips. However, the reduction in mobility was heterogeneous. For example, long-distance trips exhibited a much greater reduction than short trips, and long trips were only about 15%-20% of the prepandemic level (Figure 1B). A more detailed overview on the evolution of mobility during the lockdown is provided in Figures S2 and S3, Multimedia Appendix 1. Furthermore, there were regional heterogeneities. If we look at the provincial level (administrative subdivisions of CCAA), the strongest-hit province in terms of infections, Madrid, showed the highest reduction in mobility (Figure 1C). Similarly, highly affected provinces, such as Álava and Barcelona, were among the 5 provinces with the highest mobility reduction.

In fact, we found a pronounced correlation (R=–0.74) between the mobility level and the number of cases (Figure 1D) prior to the lockdown, which then substantially reduced during the lockdown (Figure S5, Multimedia Appendix 1). Interestingly, we observed a stronger correlation with the absolute number of cases than with the population size. The same held true for hospitalizations, ICU admissions, and fatalities (Figure S6, Multimedia Appendix 1).

We did not find regions with increased traffic on the days preceding the lockdown (Figure S7, Multimedia Appendix 1). In addition, the economic level or the fraction of individuals belonging to the working population did not seem to have been factors that determined the mobility level (Figures S8 and S9, Multimedia Appendix 1). However, we observed a general tendency that the reduction in mobility was greater in urban areas than in rural ones (Figures S8 and S9, Multimedia Appendix 1).

Figure 1. Evolution of mobility during the first wave. (A) Average, nationwide aggregated mobility before and during the lockdown. Dots indicate the data points, while the solid line shows a rolling centered 7-day average. Vertical dashed lines indicate the first NPIs introduced in Spain on March 10 (school closure in Madrid and Basque Country) and the announcement of the lockdown on March 13. The shaded area in light gray indicates lockdowns 1 and 3. In dark gray, we indicated lockdown 2, where, in addition, all nonessential economic activity was shut down. (B) Nationwide aggregated mobility but separating the distances of trips. Dots indicate data points, and the solid lines represent a rolling centered 7-day average. In general, long-distance trips showed a much higher reduction than shorter trips. (C) Mobility level during the lockdown (March 15-May 2) for the total number of trips in the provinces (administrative subdivisions of CCAA) of Spain. Ceuta and Melilla are not shown. Please note that the Canary Islands (islands at the bottom) were moved to be visible. (D) Correlation between the accumulated number of cases until the lockdown and the mobility level. Points represent the provinces of Spain. We excluded the provinces of Madrid and Barcelona since they represent statistical outliers due to their high number of cases. For completeness, the results with these provinces included are shown in Figure S4, Multimedia Appendix 1. The factor R and P denote the Pearson correlation coefficient and the associated *P* value. Similarly, the Spearman correlation coefficient was found to be –0.7 with *P*<.001. CCAA: Comunidades Autonomas; NPI: nonpharmaceutical intervention.



Estimation of Rt

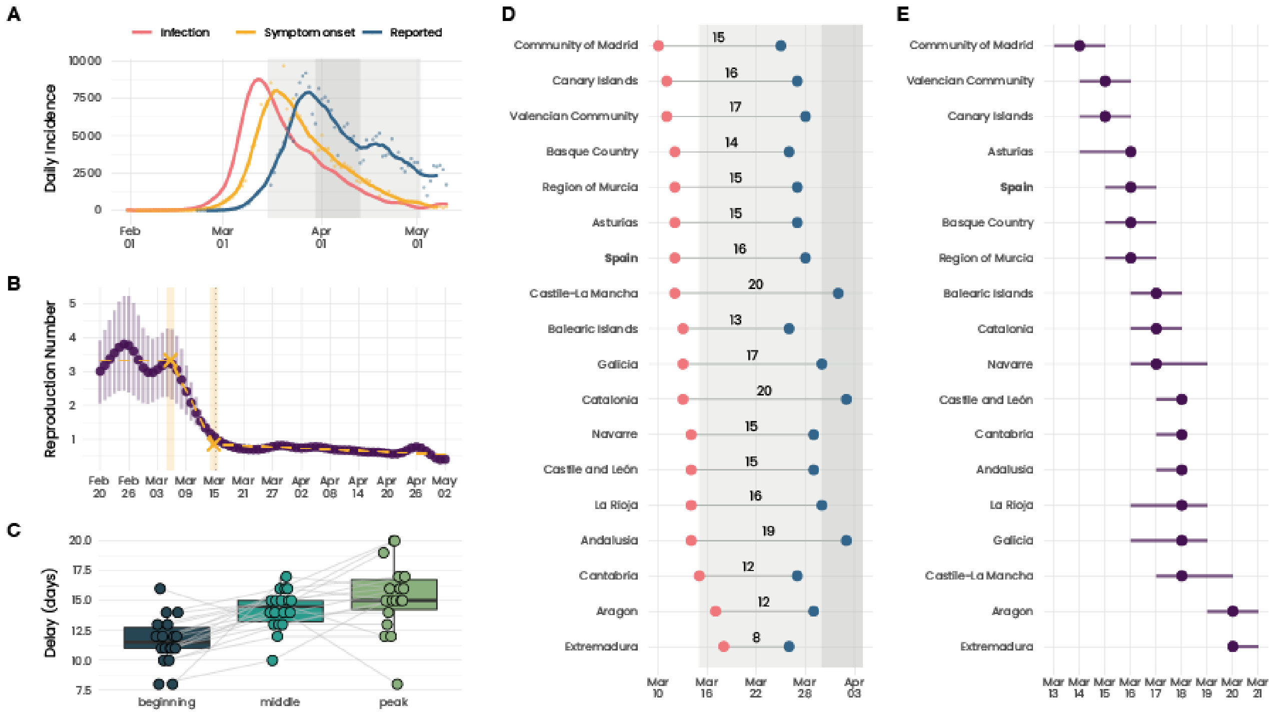
Figure 2A shows how the exposure times substantially anticipated the reported infections. The respective curves peaked with a delay of 16 days. In Figure 2B, we observe how the 3 linear segments capture the essential evolution of Rt. Additionally, we see a minimal variation in Rt during the lockdown, and the second breakpoint aligns with the lockdown implementation on March 15. Given that the average incubation time is about 5 days, this indicates a substantial delay from symptom onset until individuals were tested, results were received, and eventual positive cases were reported. We repeated the same analysis for each CCAA (Figure 2D and Figure S10, Multimedia Appendix 1). The delay ranged, among the CCAA, between 8 days in Extremadura and 20 days in Catalonia.

Given the exposure times, we inferred the reproduction number Rt shown in Figure 2B. We estimated that initially, the epidemic spread with a reproduction number of around 3. Rt started to decrease on March 5/6 according to the linear segments we identified. Furthermore, the first time Rt was below 1 was shortly after the beginning of the lockdown: between March 15 and 17. Repeating the same analysis for the regions, we observed that Rt dropped below 1 in all CCAA between March 13 and 21 (see Figure 2E). Please note that the nationwide data are strongly dominated by Madrid due to the high case numbers there. The evolution of the reproduction number in the CCAA is shown in Figure S11, Multimedia Appendix 1.

Although Rt dropping below 1 shortly after the lockdown was expected, the early decrease in Rt that was initiated on March 5/6 was rather surprising. At this point, no NPIs were in place and we did not observe any reduction in mobility. Furthermore, test-trace-isolate was implemented on such a small scale that it seems unlikely to have substantially contributed to the decrease. A possible explanation for this early reduction in Rt could be a saturation in testing capacity.

Looking at the temporal evolution of the delays between infection and reporting dates, we found an indication that the health care system was under increased strain to test and report infected individuals. The delay between infection and reporting continuously increased toward the peak in infections (Figure 2C). Although initially, we found a delay of 12 days, it increased steadily toward the peak of 16 days. Furthermore, this pattern was consistent across almost all CCAA. The only CCAA that showed a reduction in the delay toward the peak were Aragon and Extremadura (Figure S12, Multimedia Appendix 1). Interestingly, these were the CCAA with the lowest delay between the peak in infections and reported cases (Figure 2D) and were the last CCAA to have an Rt below 1 (Figure 2E). This further supports the hypothesis that there was a general overload on the testing facilities that impacted the evolution of Rt, which in turn may have led to the anticipated decrease in Rt.

Figure 2. Reconstruction of the exposure times and estimation of Rt. (A) Cases when they were reported (blue), onset of symptoms (yellow), and reconstructed exposure times. Dots indicate data points and the solid line a rolling centered 7-day average. The shaded area for the exposure times represents 95% credible intervals. (B) Estimation of Rt. Vertical bars indicate 95% credible intervals. The dashed yellow line represents the 3 linear segments we identified. The shaded yellow area shows the 95% credible interval for the 2 breakpoints. The 2 breakpoints are between March 5 (95% CI 4-6) and March 15 (95% CI 14-16), where the latter coincides with the implementation of the lockdown. (C) Delay between exposure time and reporting date. The position is defined with respect to the peak. To be more precise, we showed the time difference when both curves reached x% of their peak value. The 3 points (beginning, middle, and peak) correspond to 5%, 50%, and 100%, respectively, of daily infections compared to the peak incidence. Each point corresponds to a CCAA. Gray lines indicate how the delay evolves for each CCAA. We note that the delay steadily increased toward the peak in almost all CCAA. (D) Blue and red dots indicate the day the curves of the reported cases and exposure times reached their respective peaks. The number in between denotes the difference in days between these dates. (E) The day Rt was first below 1 in the different CCAA. Horizontal bars indicate 95% credible intervals. CCAA: Comunidades Autonomas.



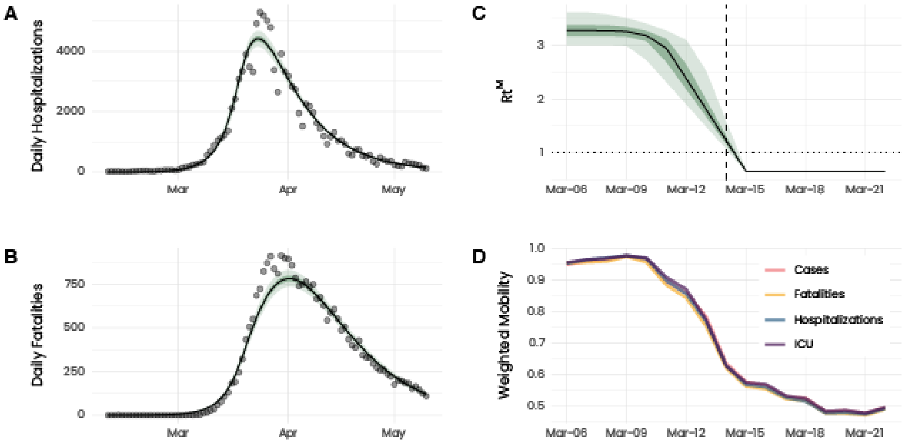
Model-Based Inference

Figure 3A,B shows the fit of the minimal epidemic model with respect to the fatalities and hospitalizations, respectively. The adjusted curve for the fatalities peaked slightly later compared to the real data. This could stem from the faster decrease in reported fatalities compared to the excess deaths. Looking at Rt^M, we found R1=3.27 (95% CI 3.01-3.61) and R2=0.66 (95% CI 0.64-0.67). Both R1 and R2 were consistent with the estimation from the reported cases, Rt. Given the generation time we used here, this resulted in a doubling time of 2.55 days (95% CI 2.32-2.78) during the free phase and a half lifetime of 9.11 (95% CI 8.72-9.56) days after the lockdown. The small doubling time suggests how the considerable delay between infection and reporting dates may have contributed to a substantial discrepancy between the actual epidemiological situation and the one that we could infer from the data. However, the low value of R2, together with the half-life time, shows the efficiency of the lockdown during the first wave and how it contributed to the low case numbers during the early summer months.

Let us turn now to the initiation of the epidemic response. The model adjustments yielded the initiation of the decrease in Rt^M on March 10 (95% CI 8-12), contrasting with the decrease in Rt on March 5/6. This suggests that the early decrease was due to an overload in the health care system, which eventually impeded testing infected individuals. The initiation of the decrease in Rt^M was consistent with the reduction in mobility. In Figure 3C,D, we contrast the reduction in mobility with Rt^M. Note that we weighted the mobility reduction by various daily epidemiological indicators—reported cases, fatalities, hospitalizations, and ICU admission—in the different provinces rather than by population. Hence, we accurately reflected the impact of mobility reduction on the evolution of the epidemic. Note that there was an additional, even though less rapid, decrease in mobility from March 15 onward. However, results from the sensitivity analysis (Figures S16-S18, Multimedia Appendix 1) suggest that the reproduction number reached a stable value timely after the lockdown. In this sense, the impact of the additional decrease in mobility on the epidemic seems not to have been substantial.

The eventual saturation of testing capacity also had consequences for ascertainment, that is, the fraction of cases that were detected compared to the total number of infections. We assumed our model output to be the total number of infections and compared it to the number of infections (exposure times) that we inferred previously. We noted that initially, ascertainment was only around 5% and subsequently increased. However, it then substantially dropped around May 5/6 before eventually starting to increase again. This substantial increase was consistent with the pronounced expansion in testing capacity in the middle of April (Figure S19, Multimedia Appendix 1) that enabled the detection of around 15% of the infections.

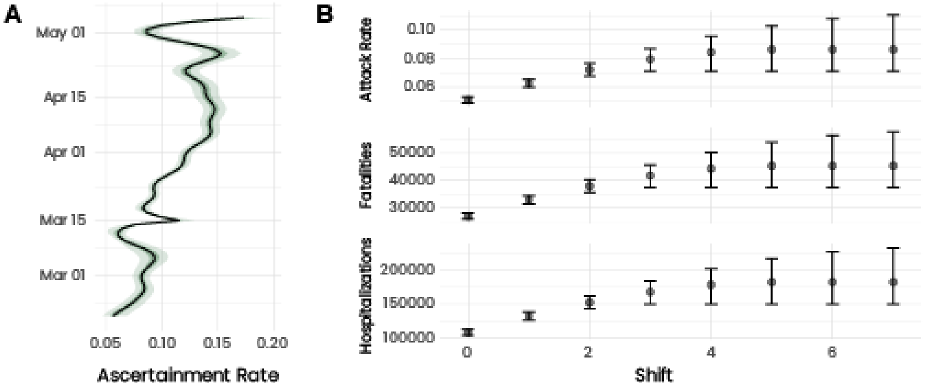
Figure 3. Model adjustment to epidemiological data. (A and B) Model adjustment to the daily hospitalizations and fatalities, respectively. The green and light-green shaded areas represent 50% and 95% credible intervals, respectively. The solid line represents the median. (C) Inferred evolution of Rt^M. The dashed vertical line indicates the implementation of the lockdown. (D) Aggregated mobility level of the Spanish provinces prior to the lockdown and shortly afterward. Instead of averaging by population size, we averaged by different daily epidemiological indicators to visualize the impact of mobility reduction on epidemic evolution. This is in analogy with the definition of Rt [31]. We observed that mobility and Rt^M started to decrease around the same time. ICU: intensive care unit.



Counterfactual Scenarios

The results of the counterfactual scenarios, shown in Figure 4B-D, indicate that the absence of an epidemic response before the lockdown would have resulted in an attack rate of 8.6% (95% CI 7.1-11.0), which is more than 50% higher than the actual value. The higher attack rate would then have resulted in 45,400 (95% CI 37,400-58,000) fatalities and 182,600 (95% CI 150,400-233,800) hospitalizations compared to 27,800 fatalities and 107,600 hospitalizations, respectively, that were reported. In other words, the results suggest that the pandemic response before the lockdown contributed substantially to limiting the impact of SARS-CoV-2 during the first wave in Spain.

Figure 4. Ascertainment and counterfactual scenarios. (A) Evolution of ascertainment in time. Ascertainment is defined as the ratio between the reported cases and the incidence from the adjusted model. We compared the date of infection (exposure time) instead of the reporting date. We noted a sudden decrease from March 5/6 toward the lockdown. Later, the testing capacity substantially increased, which increased ascertainment (Figure S19, Multimedia Appendix 1). (B) Counterfactual scenarios for the attack rate, death toll, and total number of hospitalizations. The counterfactual scenarios consist of shifting the breakpoint BP by x days. If the BP exceeded March 15 (ie, lockdown), no further shift was applied. We noted how the response before the implementation of the lockdown substantially contributed to limiting the impact of the epidemic.



Discussion

Principal Findings

Our analysis of large-scale mobility data revealed a significant reduction in mobility initiated before the lockdown. Additionally, we discovered a strong correlation between the decrease in mobility before the lockdown and the number of reported cases across Spanish provinces. Although the direct analysis of reported infections suggested a reduction in the reproduction number preceding the decrease in mobility, the model-based inference using hospital admissions and deaths primarily attributes this to a declining reporting rate as the epidemic progressed. The model-based inference indicates a simultaneous decrease in mobility and the reproduction number. According to our counterfactual scenarios, we estimate that the reduction in the reproduction number before the lockdown reduced the attack rate, hospital admissions, and deaths by over 30%.

The direct analysis of reported cases revealed a constant increase in the delay between exposure and notification as the epidemic peaked. An average delay of 16 days highlights the late testing of symptomatic individuals, a significant test turnaround time, and considerable notification delays. Studies have reported similar delays in other countries [17,20]. Such delays obstruct accurate evaluations of the epidemiological situation and hinder effective epidemic response management [45].

The increasing reporting delay also suggests a saturation in testing capacity. We hypothesized that the reduction in mobility before the decrease in Rt resulted from a worsening reporting rate. Our model-based approach supports this hypothesis, indicating an epidemic evolution consistent with the reduction in mobility, as found in other studies [15-17,46]. Moreover, ICU capacities seemed to have reached their limits. Specifically, ICU admissions among 70-79-year-olds peaked earlier than for younger age groups in the hardest-hit provinces, Madrid and Barcelona, indicating a change in admission criteria. These factors illustrate the immense pressure the Spanish health care system faced during the first wave.

The decrease in mobility began on March 9/10, 5-6 days before lockdown implementation. Linka et al [17] observed a similar decrease in mobility before lockdown introductions in other Western European countries. Identifying the exact cause of the early mobility decrease is challenging, as various factors are likely to have contributed to the decline. Worsening situations in many countries, reported in the media [47] and further disseminated through social networks [48], might have increased public awareness. In line with this increased awareness, the usage of SARS-CoV-2–related hashtags in Spain surged from March 9 onward [49]. We also found a significant correlation between mobility levels and the number of cases at the provincial level, suggesting the presence of risk-based individual awareness [50]. Other studies have identified risk-based awareness for SARS-CoV-2 [51,52], as well as for HIV [53] and measles [54].

The correlation between mobility and reported cases was especially high before lockdown implementation, as contact reduction mainly occurred voluntarily (Figure S5, Multimedia Appendix 1). However, the first autonomous communities announced educational system closure on March 9, with many others following suit. These factors are interconnected, as public awareness of COVID-19 influences voluntary social distancing, and individuals' attitudes toward the disease also impact authorities' decisions. Moreover, lockdown announcements can have similar effects as their implementations [55]. Our results should not be interpreted as lockdowns and NPIs being generally ineffective. Various studies have emphasized the impact of NPIs [18,56]. Our findings, however, highlight the complex interplay between risk-based and policy-induced behavioral changes and challenge a mechanistic understanding of NPIs, where any reduction in the reproduction number is solely attributed to policy [10,19]. In summary, studies analyzing the impact of public health policies should also account for voluntary behavioral changes beyond NPIs.

At the beginning of the model dynamics on February 10, we estimated around 1100 infectious individuals. In contrast, the first case of local transmission in mainland Spain was reported on February 26. This discrepancy underscores how SARS-CoV-2 initially spread silently through the population.

This observation aligns with the excess deaths attributed to influenza in February 2020 in Catalonia [57]. In line with this, the model consistently predicts more fatalities at the beginning of the epidemic than were reported.

Our counterfactual study demonstrates that the early decrease in mobility significantly contributed to mitigating the impact of the first wave in Spain. Our results suggest that if the epidemic response had started with the lockdown implementation on March 15, the attack rate would have been 8.6%, resulting in 180,000 hospitalizations and 45,000 fatalities, an increase of 60%-70% in these indicators. This finding indicates that in addition to the national lockdown, the combination of awareness and regional measures helped slow the spread of SARS-CoV-2 in Spain. Shifting the entire epidemic response, including the lockdown, would have led to twice/half as many infected individuals for every 2.5 days later/earlier. This emphasizes the importance of timely containment efforts in managing emerging epidemics [7,19,20,58].

Limitations

Several aspects of our analysis could be improved if more detailed data were available. For instance, the epidemiological model could be age-stratified if relevant data were available for Spain regarding symptom onset to hospitalization and death [36]. The same applies to geographical heterogeneity. If case line data were available at the regional level, autonomous communities could be treated separately [35]. In the absence of such data, age or location stratification is not feasible. Access to data that are not publicly available, unlike in other countries [4], would enable a more detailed and comprehensive analysis of the 2020 SARS-CoV-2 epidemic in Spain [11,12].

Conclusion

Summarizing, this study emphasizes that the behavioral response to an epidemic is multifaceted, driven by both voluntary decision-making and health authorities' policies. Although our results indicate that the reproduction number reached its lowest point during the lockdown in Spain, they also suggest a decrease before the implementation of any containment policies. Disentangling voluntary from policy-induced behavioral changes remains a future challenge, as this interplay involves individual psychology, societal dynamics, and pathogen spread [59]. However, understanding this complex interplay is crucial for designing better health policies and accurately assessing the need for public health interventions.

Acknowledgments

The authors acknowledge Prof Dan Larremore and Prof Miguel Hernán for their helpful comments on the manuscript. AA, BS, and SG acknowledge financial support from Spanish Ministerio de Economía, Industria y Competitividad (MINECO; grant nos. PGC2018-094754-B-C2 and PID2021-128005NB-C21), the Generalitat de Catalunya (grant nos. 2021SGR-0063 and 2020PANDE00098), and the Universitat Rovira i Virgili (grant no. 2019PFR-URV-B2-41). AA acknowledges the Joint Appointment Program at the Pacific Northwest National Laboratory (PNNL). The PNNL is a multiprogram national laboratory operated for the US Department of Energy by the Battelle Memorial Institute (contract no. DE-AC05-76RL01830) and the European Union’s Horizon Europe Programme under the Critical Action Planning Over Extreme-Scale Data (CREXDATA) project (grant agreement no. 101092749). AA also acknowledges support from the Generalitat de Catalunya (PDAD14/20/00001), Institución Catalana de Investigación y Estudios Avanzados (ICREA) Academia, and the James S McDonnell Foundation (grant no. 220020325). DS-P and JG-G acknowledge financial support from the Spanish Ministerio de Ciencia e Innovación (projects FIS2017-87519-P and PID2020-113582GB-I00), the Departamento de Industria e Innovación del Gobierno de Aragón y Fondo Social Europeo (Física Estadística y No-Lineal [FENOL] group E-19), and the Fundación Ibercaja and Universidad de Zaragoza (grant no. 224220). CG acknowledges financial support from Beatriz de Galindo (Ministerio de Ciencia, Innovación y Universidades). BS acknowledges financial support from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 713679 and from the Universitat Rovira i Virgili.

Data Availability

Supplementary information is available for this paper. Correspondence and requests for materials should be addressed to AA.

Authors' Contributions

AA, SG, and JG-G are joint senior authors. BS and AA designed the project. BS, JM, GR, and DS-P collected data. BS, JM, DS-P, GR, and CG analyzed the data. BS, JM, DS-P, GR, SG, CG, JG-G, and AA interpreted the results. BS, JM, DS-P, GR, SG, CG, JG-G, and AA wrote the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complementary figures.
[PDF File (Adobe PDF File), 4331 KB - publichealth_v9i1e40514_app1.pdf]

References

1. COVID-19 en Espana. Centro Nacional de Epidemilogía. 2020. URL: https://cnecovid.isciii.es/covid19/#documentaci%C3%B3n-y-datos [accessed 2023-05-09]
2. Estimacion del número de defunciones semanales (EDeS) durante el brote de COVID-19. Instituto Nacional de Estadística. 2020. URL: https://www.ine.es/experimental/defunciones/experimentaldefunciones.html [accessed 2023-05-09]
3. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, ENE-COVID Study Group. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020 Aug 22;396(10250):535-544 [FREE Full text] [doi: 10.1016/S0140-6736(20)31483-5] [Medline: 32645347]
4. Knock ES, Whittles LK, Lees JA, Perez-Guzman PN, Verity R, FitzJohn RG, et al. Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. Sci Transl Med 2021 Jul 14;13(602):eabg4262 [FREE Full text] [doi: 10.1126/scitranslmed.abg4262] [Medline: 34158411]
5. Li F, Li Y, Liu M, Fang L, Dean NE, Wong GWK, et al. Household transmission of SARS-CoV-2 and risk factors for susceptibility and infectivity in Wuhan: a retrospective observational study. Lancet Infect Dis 2021 May;21(5):617-628. [doi: 10.1016/s1473-3099(20)30981-6]
6. Marziano V, Guzzetta G, Rondinone BM, Boccuni F, Riccardo F, Bella A, et al. Retrospective analysis of the Italian exit strategy from COVID-19 lockdown. Proc Natl Acad Sci U S A 2021 Jan 26;118(4):e2019617118 [FREE Full text] [doi: 10.1073/pnas.2019617118] [Medline: 33414277]
7. Althaus CL, Probst D, Hauser A, Riou J. Time is of the essence: containment of the SARS-CoV-2 epidemic in Switzerland from February to May 2020. medRxiv Preprint posted online July 25, 2020 [FREE Full text] [doi: 10.1101/2020.07.21.20158014]
8. Lemaitre JC, Perez-Saez J, Azman AS, Rinaldo A, Fellay J. Assessing the impact of non-pharmaceutical interventions on SARS-CoV-2 transmission in Switzerland. Swiss Med Wkly 2020 May 18;150(2122):w20295 [FREE Full text] [doi: 10.4414/smw.2020.20295] [Medline: 32472939]
9. Di Domenico L, Pullano G, Sabbatini CE, Boëlle PY, Colizza V. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. BMC Med 2020 Jul 30;18(1):240 [FREE Full text] [doi: 10.1186/s12916-020-01698-4] [Medline: 32727547]
10. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science 2020 Jul 10;369(6500):eabb9789 [FREE Full text] [doi: 10.1126/science.abb9789] [Medline: 32414780]
11. García-Basteiro A, Alvarez-Dardet C, Arenas A, Bengoa R, Borrell C, Del Val M, et al. The need for an independent evaluation of the COVID-19 response in Spain. Lancet 2020 Aug;396(10250):529-530. [doi: 10.1016/s0140-6736(20)31713-x]
12. García-Basteiro AL, Legido-Quigley H, Álvarez-Dardet C, Arenas A, Bengoa R, Borrell C, et al. Evaluation of the COVID-19 response in Spain: principles and requirements. Lancet Public Health 2020 Nov;5(11):e575. [doi: 10.1016/s2468-2667(20)30208-5]
13. Arenas A, Cota W, Gómez-Gardeñes J, Gómez S, Granell C, Matamalas JT, et al. Modeling the spatiotemporal epidemic spreading of COVID-19 and the impact of mobility and social distancing interventions. Phys Rev X 2020 Dec 18;10(4):041055. [doi: 10.1103/physrevx.10.041055]
14. Wong G, Weiner Z, Tkachenko A, Elbanna A, Maslov S, Goldenfeld N. Modeling COVID-19 dynamics in Illinois under nonpharmaceutical interventions. Phys Rev X 2020 Nov 16;10(4):041033. [doi: 10.1103/PhysRevX.10.041033]
15. Deforche K, Vercauteren J, Müller V, Vandamme A. Joint Analysis of the Epidemic Evolution and Human Mobility During the First Wave of COVID-19 in Spain: Retrospective Study. BMC Public Health 2021 Apr 06;21(1):654 [FREE Full text] [doi: 10.1186/s12889-021-10676-1] [Medline: 33823820]
16. Nouvellet P, Bhatia S, Cori A, Ainslie KEC, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission. Nat Commun 2021 Feb 17;12(1):1090 [FREE Full text] [doi: 10.1038/s41467-021-21358-2] [Medline: 33597546]
17. Linka K, Goriely A, Kuhl E. Global and local mobility as a barometer for COVID-19 dynamics. Biomech Model Mechanobiol 2021 Apr 15;20(2):651-669 [FREE Full text] [doi: 10.1007/s10237-020-01408-2] [Medline: 33449276]
18. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nat Hum Behav 2020 Dec 16;4(12):1303-1312. [doi: 10.1038/s41562-020-01009-0] [Medline: 33199859]
19. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, Imperial College COVID-19 Response Team, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020 Aug 5;584(7820):257-261. [doi: 10.1038/s41586-020-2405-7] [Medline: 32512579]
20. Pei S, Kandula S, Shaman J. Differential effects of intervention timing on COVID-19 spread in the United States. Sci Adv 2020 Dec 04;6(49):eabd6370 [FREE Full text] [doi: 10.1126/sciadv.abd6370] [Medline: 33158911]
21. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. Proc Natl Acad Sci U S A 2020 May 12;117(19):10484-10491 [FREE Full text] [doi: 10.1073/pnas.2004978117] [Medline: 32327608]
22. Estudio de movilidad con big data. Ministerio de transportes, movilidad, y agenda urbana. 2020. URL: https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data [accessed 2023-05-09]

23. Informe metodologico. Ministerio de transportes, movilidad, y agenda urbana. 2021. URL: https://cdn.mitma.gob.es/portal-web-drupal/covid19/bigdata/mitma?estudio19informemetodologicov3.pdf [accessed 2023-05-09]
24. Hohle M. backprojNP: non-parametric back-projection of incidence cases to exposure. surveillance: Temporal and Spatio-Temporal Modeling and Monitoring of Epidemic Phenomena. 2019. URL: https://rdrr.io/cran/surveillance/man/backprojNP.html [accessed 2023-05-09]
25. Salmon M, Schumacher D, Höhle M. Monitoring count time series in R: aberration detection in public health surveillance. J Stat Soft 2016;70(10):1-35. [doi: 10.18637/jss.v070.i10]
26. Meyer S, Held L, Höhle M. Spatio-temporal analysis of epidemic phenomena using the R package surveillance. J Stat Soft 2017;77(11):1-55. [doi: 10.18637/jss.v077.i11]
27. Becker NG, Watson LF, Carlin JB. A method of non-parametric back-projection and its application to AIDS data. Stat Med 1991 Oct;10(10):1527-1542. [doi: 10.1002/sim.4780101005] [Medline: 1947509]
28. Yip PSF, Lam KF, Xu Y, Chau PH, Xu J, Chang W, et al. Reconstruction of the infection curve for SARS epidemic in Beijing, China using a back-projection method. Commun Stat Simul Comput 2008 Feb;37(2):425-433. [doi: 10.1080/03610910701792562]
29. Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei Province, China: a descriptive and modelling study. Lancet Infect Dis 2020 Jul;20(7):793-802. [doi: 10.1016/s1473-3099(20)30230-9]
30. Cori A. EpiEstim: estimate time-varying reproduction numbers from epidemic curves. CRAN R Project. 2021 Jan 9. URL: https://CRAN.R-project.org/package=EpiEstim [accessed 2023-05-09]
31. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol 2013 Nov 01;178(9):1505-1512 [FREE Full text] [doi: 10.1093/aje/kwt133] [Medline: 24043437]
32. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Euro Surveill 2020 Apr;25(17):2000257 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.17.2000257] [Medline: 32372755]
33. Muggeo VMR. segmented: an R package to fit regression models with broken-line relationships. R News 2008 May;8(1):20-25 [FREE Full text]
34. Muggeo VMR. Estimating regression models with unknown break-points. Stat Med 2003 Oct 15;22(19):3055-3071. [doi: 10.1002/sim.1545] [Medline: 12973787]
35. Hawryluk I, Mellan TA, Hoeltgebaum H, Mishra S, Schnekenberg RP, Whittaker C, et al. Inference of COVID-19 epidemiological distributions from Brazilian hospital data. J R Soc Interface 2020 Nov 25;17(172):20200596 [FREE Full text] [doi: 10.1098/rsif.2020.0596] [Medline: 33234065]
36. Faes C, Abrams S, Van Beckhoven D, Meyfroidt G, Vlieghe E, Hens N, Belgian Collaborative Group on COVID-19 Hospital Surveillance. Time between symptom onset, hospitalisation and recovery or death: atatistical analysis of Belgian COVID-19 patients. Int J Environ Res Public Health 2020 Oct 17;17(20):7560 [FREE Full text] [doi: 10.3390/ijerph17207560] [Medline: 33080869]
37. COVID-19. Informes previos. Instituto de Salud Carlos III. 2020. URL: https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx [accessed 2023-05-09]
38. Fraser C. Estimating individual and household reproduction numbers in an emerging epidemic. PLoS One 2007 Aug 22;2(8):e758 [FREE Full text] [doi: 10.1371/journal.pone.0000758] [Medline: 17712406]
39. Pastor-Barriuso R, Pérez-Gómez B, Hernán MA, Pérez-Olmeda M, Yotti R, Oteo-Iglesias J, ENE-COVID Study Group. Infection fatality risk for SARS-CoV-2 in community dwelling population of Spain: nationwide seroepidemiological study. BMJ 2020 Nov 27;371:m4509 [FREE Full text] [doi: 10.1136/bmj.m4509] [Medline: 33246972]
40. Neal R. MCMC using Hamiltonian dynamics. arxiv:1206.1901 Preprint posted online May 9, 2012. [doi: 10.48550/arXiv.1206.1901]
41. The Stan Core Library. Stan Development Team. 2021. URL: http://mc-stan.org/3 [accessed 2023-05-09]
42. RStan: the R interface to Stan. Stan Development Team. 2020. URL: http://mc-stan.org/ [accessed 2023-05-09]
43. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci 1992 Nov 1;7(4):457-472. [doi: 10.1214/ss/1177011136]
44. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
45. Marinović AB, Swaan C, van Steenbergen J, Kretzschmar M. Quantifying reporting timeliness to improve outbreak control. Emerg Infect Dis 2015 Feb;21(2):209-216 [FREE Full text] [doi: 10.3201/eid2102.130504] [Medline: 25625374]
46. Li Y, Wang X, Campbell H, Nair H. The association of community mobility with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 330 local UK authorities. Lancet Digital Health 2021 Oct;3(10):e676-e683. [doi: 10.1016/s2589-7500(21)00144-8]
47. Mach KJ, Salas Reyes R, Pentz B, Taylor J, Costa CA, Cruz SG, et al. News media coverage of COVID-19 public health and policy information. Humanit Soc Sci Commun 2021 Sep 28;8(1):220. [doi: 10.1057/s41599-021-00900-z]

48. Tsao S, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digital Health 2021 Mar;3(3):e175-e194. [doi: 10.1016/s2589-7500(20)30315-0]
49. Esquirol B, Prignano L, Díaz-Guilera A, Cozzo E. Characterizing Twitter users behaviour during the Spanish Covid-19 first wave. arXiv:2012.06550 Preprint posted online December 11, 2012. [doi: 10.48550/arXiv.2012.06550]
50. Funk S, Salathé M, Jansen VAA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J R Soc Interface 2010 Sep 06;7(50):1247-1256 [FREE Full text] [doi: 10.1098/rsif.2010.0142] [Medline: 20504800]
51. Weitz JS, Park SW, Eksin C, Dushoff J. Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. Proc Natl Acad Sci U S A 2020 Dec 22;117(51):32764-32771 [FREE Full text] [doi: 10.1073/pnas.2009911117] [Medline: 33262277]
52. Wambua J, Hermans L, Coletti P, Verelst F, Willem L, Jarvis CI, et al. The influence of risk perceptions on close contact frequency during the SARS-CoV-2 pandemic. Sci Rep 2022 Mar 25;12(1):5192 [FREE Full text] [doi: 10.1038/s41598-022-09037-8] [Medline: 35338202]
53. Prata N, Morris L, Mazive E, Vahidnia F, Stehr M. Relationship between HIV risk perception and condom use: evidence from a population-based survey in Mozambique. IFPP 2006 Dec;32(04):192-200. [doi: 10.1363/3219206]
54. Philipson T. Private vaccination and public health: an empirical examination for U.S. measles. J Hum Resour 1996;31(3):611-630. [doi: 10.2307/146268]
55. Hyland-Wood B, Gardner J, Leask J, Ecker UKH. Toward effective government communication strategies in the era of COVID-19. Humanit Soc Sci Commun 2021 Jan 27;8(1):1-11. [doi: 10.1057/s41599-020-00701-w]
56. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science 2021 Feb 19;371(6531):eabd9338 [FREE Full text] [doi: 10.1126/science.abd9338] [Medline: 33323424]
57. Coma Redon E, Mora N, Prats-Uribe A, Fina Avilés F, Prieto-Alhambra D, Medina M. Excess cases of influenza and the coronavirus epidemic in Catalonia: a time-series analysis of primary-care electronic medical records covering over 6 million people. BMJ Open 2020 Jul 29;10(7):e039369 [FREE Full text] [doi: 10.1136/bmjopen-2020-039369] [Medline: 32727740]
58. Huang X, Shao X, Xing L, Hu Y, Sin DD, Zhang X. The impact of lockdown timing on COVID-19 transmission across US counties. EClinicalMedicine 2021 Aug;38:101035 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101035] [Medline: 34308301]
59. Ferguson N. Capturing human behaviour. Nature 2007 Apr 12;446(7137):733. [doi: 10.1038/446733a] [Medline: 17429381]

Abbreviations

CCAA: Comunidades Autonomas
ICU: intensive care unit
IFR: infection fatality ratio
IHR: infection hospitalization ratio
MITMA: Ministry of Transport, Mobility and Urban Agenda
NPI: nonpharmaceutical intervention

Edited by A Mavragani, T Sanchez; submitted 24.06.22; peer-reviewed by R Guo, B Maier, JG Young; comments to author 14.11.22; revised version received 02.12.22; accepted 27.04.23; published 22.05.23.

Please cite as:
Steinegger B, Granell C, Rapisardi G, Gómez S, Matamalas J, Soriano-Paños D, Gómez-Gardeñes J, Arenas A
Joint Analysis of the Epidemic Evolution and Human Mobility During the First Wave of COVID-19 in Spain: Retrospective Study
JMIR Public Health Surveill 2023;9:e40514
URL: https://publichealth.jmir.org/2023/1/e40514
doi:10.2196/40514
PMID:37213190

Original Paper

Gender-Specific Association Between Perceived Stigma Toward Tuberculosis and Acceptance of Preventive Treatment Among College Students With Latent Tuberculosis Infection: Cross-Sectional Analysis

Yemin Yuan[1,2], PhD; Jin Jin[3], MSc; Xiuli Bi[3], MSc; Hong Geng[3], MSc; Shixue Li[1,4], PhD; Chengchao Zhou[1,4], PhD

[1]Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[2]Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[3]Department of Prevention and Control, Shandong Public Health Clinical Center, Jinan, China
[4]NHC Key Laboratory of Health Economics and Policy Research, Shandong University, Jinan, China

Corresponding Author:
Chengchao Zhou, PhD
Centre for Health Management and Policy Research, School of Public Health
Cheeloo College of Medicine
Shandong University
44 Wenhuaxi Road Jinan
Shandong
Jinan, 250012
China
Phone: 86 0531 8838 1567
Email: zhouchengchao@sdu.edu.cn

Abstract

Background: With the increasing enrollment scale of colleges, the number of students on campus has risen sharply in China. The number of students with tuberculosis (TB) and rifampicin-resistant TB in colleges has increased significantly. Preventive treatment of latent tuberculosis infection (LTBI) is an important means for TB prevention and control in colleges. At present, the acceptance of LTBI treatment among college students remains unclear. In addition, evidence shows stigma may be one of the key factors affecting acceptance of LTBI treatment. To date, there is little direct evidence on the gender-specific association between perceived stigma toward TB and acceptance of LTBI treatment among college students.

Objective: This study aimed to describe the acceptance of LTBI treatment among college students in an eastern province of China to explore the association between perceived stigma toward TB and acceptance of LTBI treatment and to examine the moderating effect of gender on the association.

Methods: Data were derived from the project on the evaluation of LTBI treatment and its effectiveness among college students in Shandong, China. In total, 1547 college students were included in the analysis. We considered covariates at the individual and family levels. Multilevel mixed-effects logistic regression was used to examine the moderating role of gender and also explore the association between perceived stigma toward TB and acceptance of LTBI treatment.

Results: The acceptance rate of LTBI treatment among the diagnosed college students was 46.7% (n=723). The proportion of female students (n=361, 51.5%) accepting LTBI treatment was higher than that of male students (n=362, 42.8%; *P*=.001). There was an interaction between perceived stigma toward TB and gender (OR 0.93, 95% CI 0.87-1.00; *P*=.06). Among college students with LTBI, perceived stigma toward TB was positively associated with acceptance of preventive treatment (OR 1.03, 95% CI 1.00-1.08, *P*=.05). Perceived stigma toward TB was positively associated with accepting LTBI treatment only among male students (OR 1.07, 95% CI 1.02-1.12; *P*=.005).

Conclusions: The acceptance rate of preventive treatment among college students with LTBI was low. Contrary to our expectations, perceived stigma toward TB was positively associated with acceptance of preventive treatment. Gender moderated this association; high perceived stigma toward TB was associated with acceptance of preventive treatment only in male gender. Gender-specific strategies are effective in improving the acceptability of LTBI treatment in colleges.

(JMIR Public Health Surveill 2023;9:e43972) doi:10.2196/43972

KEYWORDS

gender differences; perceived stigma; latent tuberculosis; treatment; acceptance; college students

Introduction

According to the Global Tuberculosis Report 2022, China has the third highest number of cases among 30 countries with a high burden of tuberculosis (TB). In China, the reported incidence of TB among students dropped from 20.6/100,000 enrolled students in 2010 to 13.4/100,000 in 2015 but increased to 17.5/100,000 in 2019 [1]. In 2019, TB cases in students accounted for 6.2% of all TB cases [2]. In recent years, the annual incidence of pulmonary TB among college students is higher than the national level for the same age group [3]. A systematic review and meta-analysis of TB outbreaks among students in mainland China showed that the incidence of TB was higher in colleges compared to high schools [4]. The number of students with TB and rifampicin-resistant TB increased significantly in China from 2015 to 2019, most of whom were college students [5]. With the expansion of national college enrollment and more schools being open, the prevention and control of TB in colleges have become challenging.

Latent tuberculosis infection (LTBI) is defined as a state of persistent immune response to stimulation by *Mycobacterium tuberculosis* antigens with no evidence of clinically manifest active TB [6]. Approximately 350 million people in China have an infection caused by *Mycobacterium tuberculosis* [7], which means that a considerable number of people are in the LTBI state. The population with LTBI is a large reservoir of TB and has a 5%-10% lifetime chance of developing active TB if LTBI is untreated [8]. LTBI treatment is one of the important interventions to finally meet the end TB targets. TB outbreaks still occur from time to time in Chinese colleges. LTBI treatment is an important means for TB prevention and control in colleges. The acceptance of LTBI treatment among college students remains unclear at present. Clarification of the LTBI treatment acceptance and identification of the key factors influencing it may be of great significance to implement this intervention among college students with LTBI.

Stigma is defined as “a social process, experienced or anticipated, characterized by exclusion, rejection, blame or devaluation that results from experience, perception or reasonable anticipation of an adverse social judgement about a person or group” [9]. TB-related stigma, including fear of infection, may be a barrier to initiation and successful completion of LTBI treatment [10,11]. A qualitative study in the Netherlands [12] found that the difficulty to differentiate LTBI from TB and the consequent fear of TB infection and disease had negatively impacted the screening or treatment of LTBI among asylum seekers and refugees. Another study in England [13] indicated that migrants in a college environment were reluctant to accept LTBI screening and treatment programs because they were concerned about reputational loss and stigma of being involved in a TB project. These findings indicate that stigma may be one of the key factors affecting acceptance of LTBI treatment. Based on existing findings, we hypothesize that perceived stigma toward TB is negatively associated with acceptance of preventive treatment among college students with LTBI.

Globally, many more adult men than women have been diagnosed with TB and have died from it [14]. A meta-analysis reported that men were disadvantaged in seeking and accessing TB care and suggested that men were a high-risk group requiring improved access to TB care [15]. In addition, studies found that substantial gender differences in the epidemiologic burden of TB were almost solely attributable to gender differences in disease incidence and treatment initiation, in both of which men were disadvantaged [16]. Studies in Nigeria and Australia showed that there were gender differences in TB diagnosis, treatment, and outcomes [17,18]. Gender differences in TB may be greater in China. The average male to female proportion in China was about 2.19:1, which was higher than the global average level (1.9:1) [19,20]. However, there are few studies on gender differences in LTBI-related aspects. We hypothesize that gender differences have an impact on the association between perceived stigma toward TB and the acceptance of LTBI treatment.

Given the importance of LTBI treatment, TB control in schools, and the critical role of stigma, the aim of this study was (1) to describe the acceptance of LTBI treatment among college students in an eastern province of China; (2) to explore the association between perceived stigma toward TB and the acceptance of LTBI treatment; and (3) to examine the moderating effect of gender on this association.

Methods

Study Setting and Participants

This study used data from the project on treatment and effectiveness evaluation of LTBI among college students in Shandong province, China. More information about procedures has been described in detail in our previous publication [21]. The study began in September 2020 and was conducted in 6 cities (Jinan, Qingdao, Yantai, Rizhao, Linyi, and Zibo) using a cluster random sampling method to select samples. We selected students aged ≥18 years with positive purified protein derivative test results (ie, induration diameter ≥10 mm or accompanied by blisters, papules, double circles, and lymphatitis) to participate in our questionnaire survey. We excluded patients previously diagnosed with active pulmonary TB or extrapulmonary TB as well as those with a history of mental illness or epilepsy. Respondents were instructed to fill out the electronic survey questionnaire by trained graduate students to ensure the quality of the questionnaire. In total, 1691 individuals were eligible, and 1631 completed the whole survey.

We used the following formula to calculate the sample size:

In this formula, *n*, δ, and *p* are estimations for sample size, allowable error, and population rate π, respectively. Because acceptance of LTBI treatment among college students with LTBI is not clear in China, we used the positive prevalence of the tuberculin skin test (TST) to calculate the sample size. The prevalence of positive TSTs in Chinese college students using 10mm cutoff value was 9%-49% [22,23]. We set $P=.09$, $\alpha=.05$ (2-tailed), and $\delta=0.009$. Considering the sampling method and the 20% lost follow-up rate, the TST screening sample was finally determined to include 20,330 people [21].

Ethical Considerations

This study protocol received ethics approval from the Ethical Committee of the School of Public Health in Shandong University (LL20200306). The investigation was conducted after the acquisition of written informed consent from all participants.

Measures

Acceptance of LTBI Treatment

Acceptance of LTBI treatment was measured by the question "Will you accept preventive treatment of latent tuberculosis infection?" The answer was divided into either accepting treatment or refusing treatment. We eliminated 84 samples whose answers were inconclusive. Finally, 1547 college students were included in the analysis.

Perceived Stigma Toward TB

Perceived stigma toward TB was assessed using 11 questions [24]. These problems originate in countries with a high burden of TB. More details about these questions are described in Multimedia Appendix 1. The answer to each question included "Yes" and "No." "Yes" indicated the presence of perceived stigma, and "No" indicated the absence of perceived stigma. The score ranged from 0 to 11. The higher the score, the higher the perceived stigma toward TB. The Cronbach coefficient was .880 in this study.

Covariates

In this study, gender was used as a moderating variable. Based on previous research [25,26], some other factors may also influence the acceptance of LTBI treatment. Therefore, we added covariates such as age (continuous), type of student (eg, medical students and nonmedical students), type of residence (eg, urban and rural), parents' levels of education (eg, primary school or below, junior high school, and high school or above), household income in 2019 (eg, quintile 1, quintile 2, quintile 3, and quintile 4), smoking (yes), current drinking (yes), physical activity level (eg, low, moderate, and high), exposure history of TB (eg, no, yes, or unknown), boarding experiences (yes), and knowledge of TB (continuous). Physical activity level was measured by the Chinese version of International Physical Activity Questionnaire Short Form [27]. The score of TB knowledge ranges from 0 to 45 (Multimedia Appendix 2). The higher the score, the higher the level of TB knowledge.

Statistical Analysis

First, a descriptive analysis was used to describe participants' characteristics with mean (SD) or median (IQR) values for continuous variables. Frequency and percentage values were used to describe categorical variables. Second, Pearson chi-square test (for categorical variables) and 2-tailed student *t* test or Wilcoxon rank-sum test (for continuous variables) were used to compare characteristics of male and female students. Then, since 1547 participants came from 6 cities and 16 colleges, multilevel mixed-effects logistic regression (Multimedia Appendix 3) was performed to examine the moderating role of gender. We added the multiplicative interaction term (perceived stigma toward TB × gender) in models. The association between perceived stigma toward TB and acceptance of LTBI treatment was also assessed by multilevel mixed-effects logistic regression. In model 1, no covariate was included. Model 2 was based on model 1, with additional adjusting for age, type of student, type of residence, father's level of education, mother's level of education, household income, smoking, current drinking, physical activity level, exposure history of TB, boarding experiences, and knowledge of TB. Odds ratios (ORs) and 95% CIs were presented as measures of effect. A 2-tailed $P<.1$ was considered as statistically significant. We conducted all analyses in Stata (version 14.2; Stata Corp).

Results

Table 1 shows the baseline characteristics of participants by gender. Of the 1547 participants, about 846 (54.7%) were male students. The average age of respondents was 18.5 (SD 0.8) years, and most of them were nonmedical students. The proportion of female medical specialty students was higher than that of male students ($P<.001$). In terms of lifestyle behaviors, male students had higher proportions of smoking and current drinking than female students ($P<.001$). The average TB knowledge score was 19 (SD 7.7). The level of TB knowledge in female students was higher than that in male students ($P=.05$). The median score of perceived stigma toward TB was 1 (IQR 0-4), with a statistical difference between male and female students ($P=.09$). A total of 723 (46.7%) students with LTBI were willing to accept preventive treatment. The proportion of female students accepting LTBI treatment was higher than that of male students ($P=.001$). Male students who accepted LTBI treatment had higher perceived stigma score compared to those who did not ($P=.02$; Multimedia Appendix 4).

The interaction between perceived stigma toward TB and gender and its effect on acceptance of LTBI treatment is shown in Table 2. After adjusting for covariates, there was an interaction between perceived stigma toward TB and gender (OR 0.93, 95% CI 0.87-1.00; $P=.06$).

Table 1. Participants' characteristics of latent tuberculosis infection (LTBI) college students by gender in Shandong, China, in 2020.

Variables	Total, n (%)	Male, n (%)	Female, n (%)	*P* value
Total	1547 (100)	846 (54.7)	701 (45.3)	
Age (years), mean (SD)	18.5 (0.8)	18.6 (0.8)	18.5 (0.8)	.25
Type of students, n (%)				<.001
Medical students	247 (16)	82 (9.7)	165 (23.5)	
Nonmedical students	1300 (84)	764 (90.3)	562 (76.5)	
Type of residence, n (%)				.28
Urban	740 (47.8)	394 (46.6)	346 (49.4)	
Rural	807 (52.2)	452 (53.4)	355 (50.6)	
Father's level of education, n (%)				.75
Primary school or below	268 (17.3)	141 (16.7)	127 (18.1)	
Junior high school	721 (46.6)	397 (46.9)	324 (46.2)	
High school or above	558 (36.1)	308 (36.4)	250 (35.7)	
Mother's level of education, n (%)				.24
Primary school or below	467 (30.2)	248 (29.3)	219 (31.2)	
Junior high school	650 (42)	348 (41.1)	302 (43.1)	
High school or above	430 (27.8)	250 (29.6)	180 (25.7)	
Household income[a], n (%)				.74
Quartile 1	426 (27.5)	231 (27.3)	195 (27.8)	
Quartile 2	394 (25.5)	211 (24.9)	183 (26.1)	
Quartile 3	376 (24.3)	203 (24)	173 (24.7)	
Quartile 4	351 (22.7)	201 (23.8)	150 (21.4)	
Smoking (yes)	159 (10.3)	151 (17.9)	8 (1.1)	<.001
Current drinking (yes)	382 (24.7)	324 (38.3)	58 (8.3)	<.001
Physical activity level, n (%)				.26
Low	558 (36.1)	306 (36.2)	252 (36)	
Moderate	509 (32.9)	265 (31.3)	244 (34.8)	
High	480 (31)	275 (32.5)	205 (29.2)	
Exposure history of TB[b], n (%)				.68
No	1,093 (70.6)	596 (70.5)	497 (70.9)	
Yes	58 (3.8)	35 (4.1)	23 (3.3)	
Unknown	396 (25.6)	215 (25.4)	181 (25.8)	
Boarding experiences (yes), n (%)	1264 (81.7)	700 (82.7)	564 (80.5)	.25
Knowledge of TB (score), mean (SD)	19.0 (7.7)	18.7 (7.9)	19.5 (7.4)	.05
Perceived stigma toward TB (score), median (IQR)	1 (4)	1 (4)	1 (4)	.09
Acceptance of LTBI treatment (yes), n (%)	723 (46.7)	362 (42.8)	361 (51.5)	.001

[a] Quartile 1: ≤US $5800; quartile 2: >US $5800 and ≤US $10,150; quartile 3: >US $ 10,150 and ≤14,500; quartile 4: >US $ 14,500. Average exchange rate in 2019: US $ 1=RMB 6.8967.

[b] TB: tuberculosis.

Table 2. The effect of gender on the cross-sectional interaction between perceived stigma toward tuberculosis (TB) and acceptance of latent tuberculosis infection (LTBI) treatment among college students in Shandong, China, in 2020.

Variables	Model 1[a]				Model 2[b]			
	OR[c]	SE	95% CI	*P* value	OR	SE	95% CI	*P* value
Perceived stigma toward TB	1.06	0.025	1.01-1.11	.005	1.07	0.025	1.02-1.12	.005
Gender (female)	1.80	0.261	1.36-2.39	<.001	1.94	0.300	1.43-2.62	<.001
Perceived stigma toward TB × gender	0.94	0.035	0.87-1.00	.07	0.93	0.035	0.87-1.00	.06

[a] Model 1 was unadjusted.

[b] Model 2 was adjusted for age, type of student, type of residence, father’s level of education, mother’s level of education, household income, smoking, current drinking, physical activity level, TB exposure history, boarding experiences, and knowledge of TB.

[c] OR: odds ratio.

Table 3 presents the relationship between perceived stigma toward TB and acceptance of LTBI treatment among male and female students. Perceived stigma toward TB was positively associated with acceptance of preventive treatment among college students with LTBI (OR 1.03, 95% CI 1.00-1.08; *P*=.05). Among male students, high perceived stigma toward TB was associated with accepting LTBI treatment when compared to low perceived stigma toward TB (OR 1.07, 95% CI 1.02-1.12; *P*=.005). However, there was no association between perceived stigma toward TB and acceptance of LTBI treatment among female students.

Table 3. Cross-sectional association between perceived stigma toward tuberculosis (TB) and acceptance of latent tuberculosis infection (LTBI) treatment among male and female students in Shandong, China, in 2020.

Variables	Model 1[a]				Model 2[b]			
	OR[c]	SE	95% CI	*P* value	OR	SE	95% CI	*P* value
Perceived stigma toward TB (n=1547)	1.04	0.019	1.00-1.07	.04	1.03	0.019	1.00-1.08	.05
Perceived stigma toward TB in male students (n=846)	1.06	0.025	1.02-1.11	.005	1.07	0.026	1.02-1.12	.005
Perceived stigma toward TB in female students (n=701)	1.00	0.029	0.94-1.06	.93	0.99	0.030	0.94-1.05	.87

[a] Model 1 was unadjusted.

[b] Model 2 was adjusted for age, type of students, type of residence, father’s level of education, mother’s level of education, household income, smoking, current drinking, physical activity level, exposure history of TB, boarding experiences, and knowledge of TB.

[c] OR: odds ratio.

Discussion

Principal Findings

In this study, only 723 (46.7%) college students diagnosed with LTBI were willing to accept preventive treatment. This figure is lower than what has been previously reported for other students’ acceptance rates of LTBI treatment, ranging between 63.9% and 87% in China [28,29]. The possible reason is that participants in this study were general college students, while previous research focused on students who had close contacts with TB and was related to the screening and management of students’ close contacts with TB [30]. For close contacts, schools actively organized screening and recommend LTBI treatments. Close contacts were more aware of TB than the student and more likely to accept LTBI treatment for fear of developing TB.

Our study observed a higher acceptance rate of LTBI treatment among female students than male students, contrary to previous studies in the United States and Canada [31]. One possible explanation for our findings is that female students have a higher level of TB knowledge than male students. Previous studies found that patients with LTBI who had low TB knowledge scores in student TB contacts or TB-designated hospitals were less likely to accept preventive treatment [28,32]. Students with higher knowledge of TB had some knowledge of the spread, harm, and prevention of TB. They wanted to reduce their risk of developing TB. When health workers recommended LTBI treatment, female students were more likely to accept it.

Our results showed that higher perceived stigma toward TB was positively related to acceptance of preventive treatment among college students with LTBI. This did run counter to our initial hypothesis, wherein we assumed that perceived stigma toward TB may be negatively associated with acceptance of LTBI treatment. No direct evidence has shown this association among college students, but prior studies found that TB-related stigma was negatively associated with LTBI treatment acceptance in other populations (eg, refugees and migrants) [12,13]. On the one hand, we speculate that the difference is due to people’s surroundings. College students live in a trustful and inclusive environment. College students did not score high on perceived stigma toward TB in our study, suggesting that they did not have embedded fears surrounding TB. Of course, it could also be that they did not know much about TB. After all, the TB knowledge score of college students in the study was 19.0 out of 45. However, TB stigma is relatively high in the community. Immigrants may feel singled out and stigmatized for other reasons [33,34]. One study in Eritrea [35] showed that restrained contacts due to fear of getting infected with TB, gossiping, and finger pointing in the community caused enacted and anticipated stigma among patients with TB. These findings indicate that the relationship between perceived stigma toward TB and acceptance of LTBI treatment may be quite opposite in different populations. On the other hand, LTBI is still different from TB infection. Although participants experienced enforced stigma, which could impede treatment initiation, their strong motivation to prevent becoming sick with TB resulted in acceptance of LTBI treatment.

It is worth noting that perceived stigma toward TB was positively associated with acceptance of LTBI treatment in male students, but there was no association between the stigma toward Tb and LTBI treatment acceptance in female students. Similar to previous studies on the impact of gender differences on TB epidemiological characteristics and access to services [15,16], we found gender differences played a part in the association between perceived stigma toward TB and acceptance of LTBI treatment. One possible factor that accounts for the disparity could be the different temperament and personality characteristics of male and female students. Women tend to be cautious, careful, worried, or pessimistic; men are more carefree, relaxed, and upbeat, even when most people are worried [36,37]. Therefore, male students may be more likely to accept LTBI treatment after feeling perceived stigma toward TB, whereas female students are not. Another plausible explanation is that the differences are related to men’s unhealthy behaviors, such as smoking and drinking. Studies found men were overrepresented in various pulmonary TB risk groups, notably alcoholic and tobacco smokers [16,38]. In this study, the proportion of male students who smoke and drink alcohol far exceeded that of female students. Despite higher perceived TB stigma, male students with LTBI were more likely to accept preventive treatment for fear of developing active TB. However, these pathways are speculative. Future research can explore the underlying mechanisms of the relationship between perceived stigma toward TB and LTBI treatment acceptance.

Perceived stigma toward TB was not associated with acceptance of LTBI treatment in female students. The reason may be that perceived stigma is not a key factor in female students accepting preventive treatment. In our study, the proportion of female medical specialty students was higher than that of male students. Moreover, female students have similar stigma scores but higher TB knowledge and LTBI treatment acceptance scores compared to male students. Of the added covariates, we found that exposure history of TB (Yes: OR 3.00, 95% CI 0.98-9.16; P=.05 and Unknown: OR 1.61, 95% CI 1.10-2.37; P=.02) and TB knowledge (OR 1.02, 95% CI 1.00-1.05; P=.08) were positively associated with acceptance of LTBI treatment among female students. In the future, research may focus on the factors influencing the acceptance of preventive treatment in women.

The current findings have some policy implications. Gender may be an important heterogeneous factor influencing the acceptance of LTBI treatment in colleges. It is recommended that colleges regularly and continuously carry out various forms of TB and LTBI knowledge publicity and education activities to raise students’ awareness of TB. The perceived stigma toward TB among male students needs to be addressed in improving their treatment acceptance. It is suggested that colleges provide suitable psychological counseling for male students. Overall, perhaps gender-specific management can improve the acceptability of LTBI treatment among college students.

This study has some limitations. First, the results were based on colleges in Shandong Province. Our findings were not necessarily generalizable to other settings within or outside China. Moreover, cross-sectional data used in this study were impossible to demonstrate the causality between perceived stigma toward TB and acceptance of LTBI treatment. Establishing and analyzing a multiyear longitudinal cohort of colleges covering different geographic regions in future studies could better elucidate the impact of gender on the relationship between perceived stigma toward TB and willingness for LTBI treatment.

Conclusions

In conclusion, we found that the acceptance rate of preventive treatment among college students with LTBI was low. The acceptance rate of female students was higher than that of male students. There were gender differences in the association between perceived stigma toward TB and acceptance of LTBI treatment. Only among male students, those with high perceived stigma toward TB were more likely to accept LTBI treatment. This study implies that gender-specific strategies in colleges may be effective in improving the acceptability of preventive treatment.

Acknowledgments

We are grateful to all participants for their effort. We thank the officials of local health agencies and staffs at the study sites for their cooperation.

This study was funded by the National Natural Science Foundation of China (grants 71974117 and 71774104), Cheeloo Youth Scholar Grant, and Shandong University (grants IFYT1810 and 2012DX006).

Data Availability

The data sets used and analyzed during this study can be obtained from the corresponding author upon request.

Authors' Contributions

YY and CZ designed the study. YY, JJ, XB, and HG collected the data. YY and CZ analyzed the data. YY drafted the manuscript. SL and CZ made critical revisions to the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Perceived stigma toward tuberculosis.
[DOC File , 16 KB - publichealth_v9i1e43972_app1.doc]

Multimedia Appendix 2
Knowledge of tuberculosis.
[DOC File , 34 KB - publichealth_v9i1e43972_app2.doc]

Multimedia Appendix 3
An introduction of multilevel mixed-effects logistic regression.
[DOC File , 43 KB - publichealth_v9i1e43972_app3.doc]

Multimedia Appendix 4
The prevalence of acceptance of LTBI treatment in Shandong, China, in 2020.
[DOC File , 19 KB - publichealth_v9i1e43972_app4.doc]

References

1. Chen H, Xia Y, Zhang C, Zhang H, Cheng J, Zhao Y. Tuberculosis in schools and requirements for prevention and control in China. China CDC Wkly 2021 Jan 15;3(3):58-60 [FREE Full text] [doi: 10.46234/ccdcw2021.005] [Medline: 34594957]
2. Cheng J, Chen H, Du X, Zhang H, Xia Y, Zhang C, et al. Applying an Automated-Alert System for Tuberculosis Control in Schools in China. China CDC Wkly 2020 Nov 13;2(46):886-887 [FREE Full text] [doi: 10.46234/ccdcw2020.237] [Medline: 34594792]
3. Yang X. Problems and countermeasures of tuberculosis prevention and control in colleges: a case study of Zibo Vocational College. Beijing: Chinese Antituberculosis Association 2021 Sep 28:118-121 [FREE Full text]
4. Bao H, Liu K, Wu Z, Wang X, Chai C, He T, et al. Tuberculosis outbreaks among students in mainland China: a systematic review and meta-analysis. BMC Infect Dis 2019 Nov 14;19(1):972 [FREE Full text] [doi: 10.1186/s12879-019-4573-3] [Medline: 31727001]
5. Su W, Ruan Y, Li T, Du X, Jiang J, He Y, et al. Epidemiological characteristics of rifampicin-resistant tuberculosis in students - China, 2015-2019. China CDC Wkly 2021 Jun 25;3(26):549-552 [FREE Full text] [doi: 10.46234/ccdcw2021.142] [Medline: 34594933]
6. Latent tuberculosis infection: updated and consolidated guidelines for programmatic management. World Health Organization. URL: https://www.who.int/publications/i/item/9789241550239 [accessed 2021-10-31]
7. Houben RMGJ, Dodd PJ. The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. PLoS Med 2016 Oct;13(10):e1002152 [FREE Full text] [doi: 10.1371/journal.pmed.1002152] [Medline: 27780211]
8. Mushtaq A. New guidelines for latent tuberculosis. Lancet Infect Dis 2020 Apr;20(4):414. [doi: 10.1016/S1473-3099(20)30181-X] [Medline: 32222207]
9. Scambler G. Health-related stigma. Sociol Health Illn 2009 Apr;31(3):441-455 [FREE Full text] [doi: 10.1111/j.1467-9566.2009.01161.x] [Medline: 19366430]
10. Rebeiro PF, Cohen MJ, Ewing HM, Figueiredo MC, Peetluk LS, Andrade KB, et al. Knowledge and stigma of latent tuberculosis infection in Brazil: implications for tuberculosis prevention strategies. BMC Public Health 2020 Jun 09;20(1):897 [FREE Full text] [doi: 10.1186/s12889-020-09053-1] [Medline: 32517671]
11. Alsdurf H, Hill PC, Matteelli A, Getahun H, Menzies D. The cascade of care in diagnosis and treatment of latent tuberculosis infection: a systematic review and meta-analysis. Lancet Infect Dis 2016 Nov;16(11):1269-1278. [doi: 10.1016/S1473-3099(16)30216-X] [Medline: 27522233]
12. Spruijt I, Haile DT, van den Hof S, Fiekert K, Jansen N, Jerene D, et al. Knowledge, attitudes, beliefs, and stigma related to latent tuberculosis infection: a qualitative study among Eritreans in the Netherlands. BMC Public Health 2020 Oct 23;20(1):1602 [FREE Full text] [doi: 10.1186/s12889-020-09697-z] [Medline: 33097021]
13. Walker C, Duffield K, Kaur H, Dedicoat M, Gajraj R. Acceptability of latent tuberculosis testing of migrants in a college environment in England. Public Health 2018 May;158:55-60. [doi: 10.1016/j.puhe.2018.02.004] [Medline: 29567507]

14. Global Tuberculosis Report 2021. World Health Organization. URL: https://www.who.int/publications/i/item/9789240037021 [accessed 2021-10-31]
15. Horton KC, MacPherson P, Houben RMGJ, White RG, Corbett EL. Sex differences in tuberculosis burden and notifications in low- and middle-income countries: a systematic review and meta-analysis. PLoS Med 2016 Sep;13(9):e1002119 [FREE Full text] [doi: 10.1371/journal.pmed.1002119] [Medline: 27598345]
16. Horton KC, Sumner T, Houben RMGJ, Corbett EL, White RG. A Bayesian approach to understanding sex differences in tuberculosis disease burden. Am J Epidemiol 2018 Nov 01;187(11):2431-2438 [FREE Full text] [doi: 10.1093/aje/kwy131] [Medline: 29955827]
17. Oshi SN, Alobu I, Ukwaja KN, Oshi DC. Investigating gender disparities in the profile and treatment outcomes of tuberculosis in Ebonyi state, Nigeria. Epidemiol Infect 2015 Apr;143(5):932-942. [doi: 10.1017/S095026881400291X] [Medline: 25355040]
18. Dale K, Tay E, Trauer JM, Trevan P, Denholm J. Gender differences in tuberculosis diagnosis, treatment and outcomes in Victoria, Australia, 2002-2015. Int J Tuberc Lung Dis 2017 Dec 01;21(12):1264-1271. [doi: 10.5588/ijtld.17.0338] [Medline: 29297447]
19. Liu K, Li T, Vongpradith A, Wang F, Peng Y, Wang W, et al. Identification and prediction of tuberculosis in Eastern China: analyses from 10-year population-based notification data in Zhejiang province, China. Sci Rep 2020 May 04;10(1):7425 [FREE Full text] [doi: 10.1038/s41598-020-64387-5] [Medline: 32367050]
20. Floyd K, Glaziou P, Zumla A, Raviglione M. The global tuberculosis epidemic and progress in care, prevention, and research: an overview in year 3 of the End TB era. Lancet Respir Med 2018 Apr;6(4):299-314. [doi: 10.1016/S2213-2600(18)30057-2] [Medline: 29595511]
21. Yuan Y, Jin J, Bi X, Geng H, Li S, Zhou C. Factors associated with refusal of preventive therapy after initial willingness to accept treatment among college students with latent tuberculosis infection in Shandong, China. BMC Infect Dis 2023 Jan 20;23(1):38 [FREE Full text] [doi: 10.1186/s12879-023-08005-5] [Medline: 36670356]
22. Huang W, Fang Z, Luo S, Lin S, Xu L, Yan B, et al. The effect of BCG vaccination and risk factors for latent tuberculosis infection among college freshmen in China. Int J Infect Dis 2022 Sep;122:321-326.2 [FREE Full text] [doi: 10.1016/j.ijid.2022.06.010] [Medline: 35700876]
23. Li B, Chen X, Yang Q, Zhang M, Liao M, Zhu X. Evaluation the impact of BCG on the mycobacterium tuberculosis specific IFN-y elispot assay. J Clin Pulm Med 2010 Jul 2010;15(7):955-957. [doi: 10.3969/j.issn.1009-6663.2010.07.027]
24. Sima BT, Belachew T, Abebe F. Knowledge, attitude and perceived stigma towards tuberculosis among pastoralists; Do they differ from sedentary communities? A comparative cross-sectional study. PLoS One 2017 Jul 17;12(7):e0181032 [FREE Full text] [doi: 10.1371/journal.pone.0181032] [Medline: 28715439]
25. Ren Z. Acceptability of preventive treatment among close contacts of pulmonary tuberculosis. Beijing: China CDC 2020 Jun 30:34-50. [doi: 10.27511/d.cnki.gzyyy.2020.000096]
26. Ma Z. The status and follow-up study of tuberculosis latent infection among freshmen in junior middle school, senior high school, university and junior college in Yining city, Xinjiang. Urumqi: Xinjiang Medical University 2022 Mar 1:32-34. [doi: 10.27433/d.cnki.gxyku.2022.000508]
27. Fan M, Lyu J, He P. [Chinese guidelines for data processing and analysis concerning the International Physical Activity Questionnaire]. Zhonghua Liu Xing Bing Xue Za Zhi 2014 Aug;35(8):961-964. [Medline: 25376692]
28. Li Y, Zheng YH, Lu LP, Yang MX, Zhou CM, Yuan ZA, et al. Acceptance of chemo-prophylaxis for latent tuberculosis infection among high school/college student contacts of tuberculosis patients in Shanghai, China. Biomed Environ Sci 2018 Apr;31(4):317-321 [FREE Full text] [doi: 10.3967/bes2018.041] [Medline: 29773096]
29. Zu X, Yao Y, Gong D, Wang Q, Ren Z, Cheng J. Acceptability of 6-month prophylactic isoniazid therapy in latently infected close contacts of tuberculosis patients. ChinJ Public Health 2020 Mar;36(3):369-374. [doi: 10.11847/zgggws1127769]
30. Code for prevention and control of tuberculosis in schools (2017 edition). National Health Commission of the People's Republic of China, Ministry of Education of the People's Republic of China. URL: http://www.moe.gov.cn/srcsite/A17/moe_943/s3285/201707/t20170727_310182.html [accessed 2021-10-31]
31. Colson PW, Hirsch-Moverman Y, Bethel J, Vempaty P, Salcedo K, Wall K, Tuberculosis Epidemiologic Studies Consortium. Acceptance of treatment for latent tuberculosis infection: prospective cohort study in the United States and Canada. Int J Tuberc Lung Dis 2013 Apr;17(4):473-479. [doi: 10.5588/ijtld.12.0697] [Medline: 23485381]
32. Li G, Xu H, Jing M, Wang P, Fan P, Xu L, et al. Reason analysis and countermeasure evaluation of latent tuberculosis infection refusing preventive treatment in a certain county. J Clin Pulm Med 2022 Sep 2022;27(9):1415-1420. [doi: 10.3969/j.issn.1009-6663.2022.09.025]
33. Parmer J, Macario E, Tatum K, Brackett A, Allen L, Picard R, et al. Latent tuberculosis infection: misperceptions among non-U.S.-born-populations from countries where tuberculosis is common. Glob Public Health 2022 Aug;17(8):1728-1742. [doi: 10.1080/17441692.2021.1947342] [Medline: 34228584]
34. Degeling C, Carter SM, Dale K, Singh K, Watts K, Hall J, et al. Perspectives of Vietnamese, Sudanese and South Sudanese immigrants on targeting migrant communities for latent tuberculosis screening and treatment in low-incidence settings: a report on two Victorian community panels. Health Expect 2020 Dec;23(6):1431-1440 [FREE Full text] [doi: 10.1111/hex.13121] [Medline: 32918523]

35. Gebreweld FH, Kifle MM, Gebremicheal FE, Simel LL, Gezae MM, Ghebreyesus SS, et al. Factors influencing adherence to tuberculosis treatment in Asmara, Eritrea: a qualitative study. J Health Popul Nutr 2018 Jan 05;37(1):1 [FREE Full text] [doi: 10.1186/s41043-017-0132-y] [Medline: 29304840]
36. Cai Q, Wang X, Liu H, Han J, He Q, Xia J, et al. Personality differences according to age and sex using the Chinese version of temperament and character inventory-revised. J China Med Univ 2014 Dec;43(12):1078-1083.
37. Chen Z, Lu X, Kitamura T. The factor structure of the Chinese version of the temperament and character inventory: factorial robustness and association with age and gender. Compr Psychiatry 2013 Apr;54(3):292-300. [doi: 10.1016/j.comppsych.2012.08.003] [Medline: 23273870]
38. Ben Jmaa M, Ben Ayed H, Koubaa M, Hammami F, Damak J, Ben Jemaa M. Is there gender inequality in the epidemiological profile of tuberculosis? Tunis Med 2020 Mar;98(3):232-240 [FREE Full text] [Medline: 32395817]

Abbreviations

LTBI: latent tuberculosis infection
OR: odds ratio
TB: tuberculosis
TST: tuberculin skin test

Edited by A Mavragani, T Sanchez; submitted 01.11.22; peer-reviewed by P Liu, T Li; comments to author 06.04.23; revised version received 20.04.23; accepted 24.05.23; published 14.06.23.

Please cite as:
Yuan Y, Jin J, Bi X, Geng H, Li S, Zhou C
Gender-Specific Association Between Perceived Stigma Toward Tuberculosis and Acceptance of Preventive Treatment Among College Students With Latent Tuberculosis Infection: Cross-Sectional Analysis
JMIR Public Health Surveill 2023;9:e43972
URL: https://publichealth.jmir.org/2023/1/e43972
doi:10.2196/43972
PMID:37314847

Original Paper

Association Between Comorbid Anxiety and Depression and Health Risk Behaviors Among Chinese Adolescents: Cross-Sectional Questionnaire Study

Meng Wang[1], BSc; Xingyue Mou[1], BSc; Tingting Li[1], MD, MPH; Yi Zhang[1], MD, MPH, PhD; Yang Xie[1], MD, MPH; Shuman Tao[2,3], MD, MPH, PhD; Yuhui Wan[1,2,3,4], MD, MPH, PhD; Fangbiao Tao[1,2,3,4], MD, MPH, PhD; Xiaoyan Wu[1,2,3,4], MD, MPH, PhD

[1]Department of Maternal, Child and Adolescent Health, School of Public Health, Anhui Medical University, Hefei, China

[2]Key Laboratory of Population Health Across Life Cycle, Ministry of Education of the People's Republic of China, Hefei, China

[3]Anhui Provincial Key Laboratory of Population Health and Aristogenics, Anhui Medical University, Hefei, China

[4]Key Laboratory of Study on Abnormal Gametes and Reproductive Tract, National Health Commission of the People's Republic of China, Hefei, China

Corresponding Author:
Xiaoyan Wu, MD, MPH, PhD
Department of Maternal, Child and Adolescent Health
School of Public Health
Anhui Medical University
No 81, Meishan Road
Shushan District
Hefei, 230032
China
Phone: 86 551 65161168
Email: xywu@ahmu.edu.cn

Abstract

Background: Comorbidity of psychiatric disorders such as depression and anxiety is very common among children and adolescents. Few studies have examined how comorbid anxiety and depression are associated with health risk behaviors (HRBs) in adolescents, which could inform preventative approaches for mental health.

Objective: We evaluated the association between HRBs and comorbid anxiety and depression in a large adolescent cohort.

Methods: We used data from 22,868 adolescents in the National Youth Cohort (China). Anxiety and depression symptoms were assessed using the 9-item Patient Health Questionnaire scale and the 7-item Generalized Anxiety Disorder scale, respectively. Comorbidity was determined by the coexistence of anxiety and depression. HRBs including poor diet, smoking, physical inactivity, and poor sleep, as well as the above HRB scores, were added to obtain the total HRB score (HRB risk index). Based on single and total HRB scores, we divided participants into low-, medium-, and high-risk groups. Potential confounders included gender, presence of siblings, regional economic level, educational status, self-rated health, parental education level, self-reported family income, number of friends, learning burden, and family history of psychosis. Correlation analysis was used to explore associations between single risk behaviors. Binary logistic regression estimated the association between HRBs and anxiety-depression comorbidity before and after adjusting for potential confounders.

Results: The comorbidity rate of anxiety and depression among Chinese adolescents was 31.6% (7236/22,868). There was a statistically significant association between each HRB ($P<.05$), and HRBs were positively associated with comorbid anxiety and depression in the above population. For single HRBs, adolescents with poor diet, smoking, and poor sleep (medium-risk) were more prone to anxiety-depression comorbidity after adjusting for confounders compared to low-risk adolescents. However, adolescents with all high-risk HRBs were more likely to have comorbid anxiety and depression after adjusting for confounders (poor diet odds ratio [OR] 1.50, 95% CI 1.39-1.62; smoking OR 2.17, 95% CI 1.67-2.81; physical inactivity OR 1.16, 95% CI 1.06-1.28; poor sleep OR 1.84, 95% CI 1.70-2.01). Moreover, in both unadjusted (medium risk OR 1.79, 95% CI 1.56-2.05; high risk OR 3.09, 95% CI 2.72-3.52) and adjusted (medium risk OR 1.57, 95% CI 1.37-1.80; high risk OR 2.33, 95% CI 2.03-2.68) models, HRB risk index, like clustered HRBs, was positively associated with anxiety-depression comorbidity, and the strength of the association was stronger than for any single HRB. In addition, we found that compared to girls, the association between clustered HRBs and anxiety-depression comorbidity was stronger in boys after adjustment.

Conclusions: We provide evidence that HRBs are related to comorbid anxiety and depression. Interventions that decrease HRBs may support mental health development in adolescence, with the potential to improve health and well-being through to adulthood.

(JMIR Public Health Surveill 2023;9:e46289) doi:10.2196/46289

KEYWORDS

health risk behaviors; anxiety; depression; comorbidity; adolescent; mental health; children; intervention; lifestyle behavior; mental disorder; public health; cross-sectional study

Introduction

Adolescence is a transitional period from childhood to adulthood and also a critical period for the formation of lifestyle behaviors that can significantly impact immediate and long-term physical and mental health. During this period, under the influence of various pressures, such as study burden, life, and social interaction, some health risk behaviors (HRBs) are prone to occur [1]. The World Health Organization (WHO) defines youth HRBs as behaviors that directly, indirectly, or potentially threaten the current or future health of adolescents, mainly including physical inactivity, unintentional injury, and premature and unsafe behaviors [2]. However, many adolescents do not conform to specific health behavior guidelines, and the prevalence of unhealthy behaviors among adolescents ranges from 2.2% to 42.6% [3]. Of note, most adolescents' HRBs do not exist alone, but tend to cluster together [4]. The harm caused by clustered HRBs is larger than single HRBs. Therefore, identifying clustered HRBs may be more practical than studying single HRBs.

Adolescents are prone to many HRBs, and HRBs can lead to psychosocial disorders, physical illness, and other problems, particularly anxiety and depression [5]. Anxiety and depression, as the most common mental illnesses in adolescents, usually present a comorbidity pattern. Comorbid anxiety and depression is a global public health concern, as about 25% to 50% of depressed children and adolescents experience anxiety, and 10% to 15% of anxious children and adolescents have depression worldwide [6]. Evidence suggests that anxiety-depression comorbidity during adolescence is not only associated with immense personal and family suffering but also causes other adverse outcomes, such as academic failure, impaired cognitive function, poor interpersonal relationships, and even suicide [7]. Possible reasons for these psychosocial disorders include genetic predisposing factors [8], environmental factors [9], and personal behavioral factors such as HRBs [10]. Health behaviors have been highlighted as a potentially important factor that may support better mental outcomes throughout the life course, especially in adolescence. A study using data from a nationally representative sample of Australian adolescents showed that early-onset HRBs were strongly associated with mental health outcomes and varied by sex [11]. Moreover, a study examining the relationship between clustered HRBs and mental health status among US students identified clusters of HRBs, and the results showed participants who reported clustered HRBs were more likely to report poorer mental health [12].

Both HRBs and anxiety-depression comorbidity are common in adolescents, and there might be an association between them [13]. Thus, it is essential to delve into associations between HRBs and comorbid anxiety and depression in adolescents, as well as to compare the association in each sex. However, little research has examined how clustered HRBs affect the development of anxiety-depression comorbidity or gender-specific associations in adolescents, which could inform targeted interventions aiming to reduce harmful behaviors and improve long-term well-being. Furthermore, previous research has pointed to exercise, smoking, diet, and sleep as playing important roles in the prevention and treatment of mental disorders [14]. Therefore, our study aims to understand the prevalence of comorbid anxiety and depressive symptoms among Chinese adolescents and determine the extent of associations between HRBs (poor diet, physical inactivity, smoking, and poor sleep) and comorbid anxiety and depression by controlling confounders and identifying the influence of gender on these associations.

Methods

Setting and Study Design

Participants were enrolled in the National Adolescent Health Survey Cohort, which is a nationally representative, cross-sectional study designed to track behaviors, physical health, and mental health in Chinese adolescents from October to December 2021. A multistage cluster sampling method was applied to recruit participants. First, we selected first-line cities (Shenzhen, Chongqing, Zhengzhou, and Shenyang) and second-line cities (Kunming, Xuzhou, Taiyuan, Nanchang) according to regional economic levels. Second, each district selected 2 rural junior and senior high schools and 2 urban junior and senior high schools. The study enrolled 27 schools in total (5 of these schools were combined junior and senior high schools). Each school selected all students from 4 to 6 classes (with no less than 200 students in each grade) to conduct a questionnaire survey. There were no less than 600 students in each school and no less than 2400 students in each district.

In this survey, 24,500 questionnaires were actually distributed, and 1632 invalid returned questionnaires were excluded. Among these, 426 (1.7%) respondents were unwilling to participate in the survey, 367 (1.5%) were not in school on the day of the survey, and 839 (3.4%) did not complete the questionnaire (>15% missing data or obvious logical errors). Thus, 22,868 valid questionnaires were finally completed for an effective response rate of 93.3%.

Measures

Demographic Factors

Sociodemographic data were collected by a self-administered questionnaire, including gender (boy or girl), age, presence of siblings (yes or no), regional economic level (first-line or second-line) [15], residential area (rural or urban), educational status (junior or senior high school), parental education level (primary school and below, junior high school, or senior high school and above), self-reported family income (low, medium, or high), number of friends (≤2 or ≥3), learning burden (high, medium, or low), and family history of psychosis (yes or no). It is worth noting that "first-line" cities are considered the most important cities in China, with the most developed economies, the densest populations, and the highest urban radiation capacity (this refers to the ability of a city to influence the surrounding area and other cities); examples are Shanghai and Shenzhen. "Second-line" cities are mostly provincial capitals, economically strong cities in eastern regions, or economically developed regional central cities, such as Taiyuan and Nanchang.

Health Risk Behaviors

Dietary habits were assessed using a questionnaire adapted from the American Adolescent Health Behavior Monitoring System [16]. Dietary risk was assessed primarily by asking the adolescents about their frequency of consumption of fruits, vegetables, fast food, and soft drinks in the last week [17]. For fruit/vegetable consumption, replies of ≥3, 2, and ≤1 represent risk scores of 0, 1, and 2, respectively. For fast food consumption, replies of 0, 1-4, and ≥5 represent risk scores of 0, 1, and 2, respectively. For soft drinks, replies of ≤1, 2, and 3 represent risk scores of 0, 1, and 2, respectively. We then added the risk scores for a total score of 0 to 8, categorized into 3 groups: 0-2 (low risk), 2-3 (medium risk) and 4-8 (high risk). It is worth noting that the composite risk score could equal 2 only if (1) any 2 of the 4 components were medium risk (0 + 0 + 1 + 1) or (2) if any 1 of the 4 components was high risk (0 + 0 + 0 + 2). The former case was classified as low risk, whereas the latter was classified as medium risk. The reason for this was that adolescents with low dietary risk should not have any dietary component assessed as high risk.

Smoking status during the last month was assessed by a question in the American Adolescent Health Behavior Surveillance System Adaptation Questionnaire [16]: "During the last month, how many days did you smoke?" According to other similar studies [18], responses were categorized into 3 groups: less than 2 days (low risk), 3 to 19 days (medium risk), and 20 days or more (high risk).

Sleep duration during the last month was assessed using the Munich Chronotype Questionnaire Short Form (MCTQ-SF) [19]: adolescents reported their bedtime, fall-asleep time, and morning wakeup time; we then computed sleep duration. According to sleep duration recommendations proposed by the WHO and others, adolescents aged 13 to 18 years should sleep 8 to 10 hours per 24 hours on a regular basis to promote optimal health [20]; thus, responses were categorized into 3 groups, including low risk (8 to 10 hours), medium risk (7 to 8 hours, not including 8 hours, or 10 to 11 hours, not including 10 hours), and high risk (less than 7 hours or more than 11 hours).

The Physical Activity Rating Scale (PAR-3) was used to measure physical activity levels in adolescents [21]. This scale examines the amount of physical activity from 3 aspects: intensity, frequency, and time of participating in physical activity. Physical activity is rated as intensity times frequency times time. Intensity, duration, and frequency of physical activity were assessed by the questions "When performing physical activity, which of the following intensities are most likely to be selected?" "How many minutes at a time did the above intensity activity last?" and "How many times was the above activity performed?" respectively. Intensity and frequency were graded from 1 to 5 with 1 to 5 points, and time was scored from 1 to 5 with 0 to 4 points. The score range of the amount of physical activity was from 0 to 100. Physical activity risk assessment criteria were as follows: score >43 (low risk), score 20-42 (medium risk) and score <19 (high risk).

To evaluate the clustering effect of HRBs on comorbid anxiety and depression, we constructed an overall HRB index. To indicate overall HRB risk, single HRB factors were given weighted risk scores: 0 (low risk), 0.5 (medium risk), and 1 (high risk). They were then summed into an overall index ranging from 0 to 4. To study different levels of HRB risk, the HRB risk index scores were divided into 3 categories: low risk (total score 0-0.5), medium risk (total score 1-1.5), and high risk (total score 2-4). The index was termed HRB risk index.

Comorbid Anxiety and Depression

Anxiety symptoms were assessed using the Generalized Anxiety Disorder–7 (GAD-7). The GAD-7 is a 7-item self-reported measure used for screening and evaluating generalized anxiety disorder symptoms [22]. Respondents indicate how much they have been bothered over the last two weeks by each GAD-7 symptom using a 4-point Likert scale ranging from 0 (not at all) to 3 (nearly every day). GAD-7 score ranges from 0 to 21 with scores of 0-4, 5-9, 10-14, and 15-21 representing normal, mild, moderate, and severe levels of anxiety, respectively. According to the wide use of a cutoff value of 5, we classified GAD-7 scale scores ≥5 as indicating the presence of anxiety symptoms, and scores <5 as indicating the absence of anxiety symptoms.

The prevalence of depressive symptoms was measured using the Patient Health Questionnaire–9 (PHQ-9), which is used for screening and evaluating depressive symptoms [23]. Respondents indicate how much they have been bothered over the last 2 weeks by each PHQ-9 symptom using a 4-point Likert scale ranging from 0 (not at all) to 3 (nearly every day). The PHQ-9 score ranges from 0 to 27, with scores of 0-4, 5-9, 10-14, and 15-27 representing normal, mild, moderate, and severe levels of depression, respectively. Based on the suggested cutoff, participants who scored 5 points or more were considered to be at risk for depressive symptoms in our study.

An adolescent would be considered to have comorbid anxiety and depression if their GAD-7 scale score was ≥5 and their PHQ-9 scale score was also ≥5. Adolescents who met the above criteria were divided into a comorbid anxiety and depression group and a noncomorbid group.

The Chinese versions of the above questionnaires on adolescent demography, behavior, and mental health have good reliability and validity and have been verified in the Chinese adolescent population.

Confounders

We adjusted for a series of potential confounders; these were defined as measures that were associated with both the exposure and the outcomes but were not on the causal pathway between the two. Confounders mainly included statistically significant variables, including gender, presence of any siblings, regional economic level, educational status, self-rated health, parental education level, self-reported family income, number of friends, learning burden, and family history of psychosis.

Ethical Considerations

The studies involving human participants were reviewed and approved by the Ethics Institutional Review Committee of Anhui Medical University (20200573), and informed consent was obtained from the schools, students, and parents.

Statistical Analysis

Statistical analyses were carried out with SPSS (version 23.0; IBM Corp). Descriptive analyses investigated associations between HRBs, comorbid anxiety and depression, and confounders. Continuous variables with normal distribution in our study were represented as means and SDs, while categorical variables were summarized as frequencies (n) and percentages (%). Chi-square tests were used to assess demographic characteristics stratified by the presence or absence of comorbid anxiety and depression. Spearman correlation was used to assess the relationship between each HRB. Binary logistic regression was used to investigate the relationships between single HRBs, clustered HRBs, and comorbid anxiety and depression in adolescents before and after adjusting for confounders, and a correlation heat map was produced with Origin (OriginLab). For a sensitivity analysis, we explored whether associations varied by gender. All statistical tests were 2-sided and had a significance level set at $P<.05$.

Results

Participant Characteristics

The characteristics of the 22,868 participants (11,578/22,868, 50.6% boys and 11,290/22,868, 49.4% girls) included in the analysis are shown in Table 1. The mean age of the participants was 14.6 (SD 1.8) years. Among the 22,868 participants, the occurrence of comorbid anxiety and depression was 31.6% (7236/22,868). There were gender differences among adolescents with anxiety-depression comorbidity, and girls (4133/11,290, 35.7%) were more likely to have such mental health problems than boys (3203/11,578; 27.7%; $P<.001$). Moreover, not having siblings, living in a first-line economic city, being a senior high school student, having parents with an education level of primary school or below, having a lower self-reported family income, having ≤2 friends, having a higher learning burden, and having a family history of psychosis were all associated with comorbid anxiety and depression ($P<.05$).

Multimedia Appendix 1, Table S1 presents the risk index scores of single HRBs and clustered HRBs and their population distribution among the adolescents as indicated in the Methods section on HRB risk index scores.

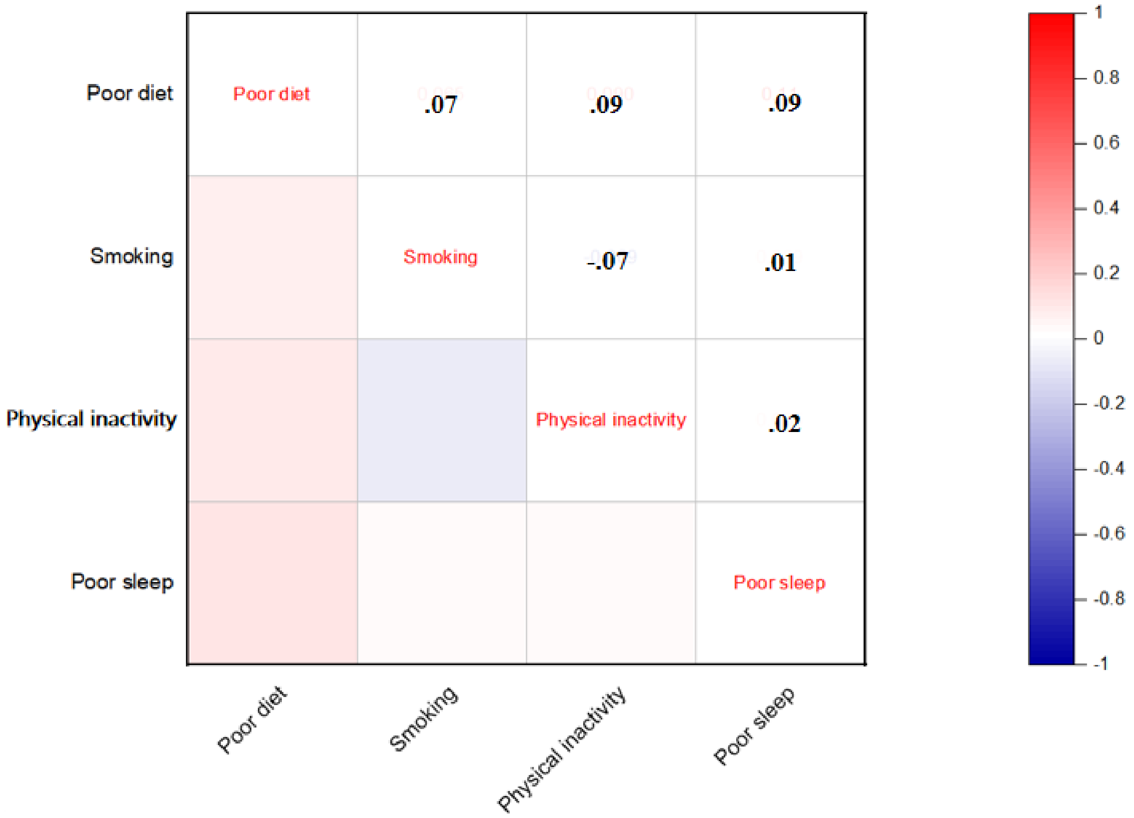
Table 1. Characteristics of the study sample.

Variable	Comorbid anxiety/depression (n=7236), n (%)	No comorbid anxiety/depression (n=15,632), n (%)	Chi-square (*df*)	*P* value
Gender			171.56 (1)	<.001
Boys (n=11,578)	3203 (27.7)	8375 (72.3)		
Girls (n=11,290)	4033 (35.7)	7257 (64.3)		
Any siblings			4.14 (1)	.04
Yes (n=6560)	2011 (30.7)	4549 (69.3)		
No (n=16,308)	5225 (32)	11,083 (68)		
Regional economic level			28.54 (1)	<.001
First-line (n=11,466)	3816 (33.3)	7650 (66.7)		
Second-line (n=11,402)	3420 (30)	7982 (70)		
Educational status			145.20 (1)	<.001
Junior high school (n=11,840)	3323 (28.1)	8517 (71.9)		
Senior high school (n=11,028)	3913 (35.5)	7115 (64.5)		
Father’s education level			17.15 (2)	<.001
Primary school and below (n=2890)	1009 (34.9)	1881 (65.1)		
Junior high school (n=8733)	2751 (31.5)	5982 (68.5)		
Senior high school and above (n=11,245)	3476 (30.9)	7769 (69.1)		
Mother’s education level			32.87 (2)	<.001
Primary school and below (n=4331)	1520 (35.1)	2811 (64.9)		
Junior high school (n=8350)	2633 (31.5)	5717 (68.5)		
Senior high school and above (n=10,187)	7236 (31.6)	15,632 (68.4)		
Self-reported family income			215.92 (2)	<.001
Low (n=2771)	1209 (43.6)	1562 (56.4)		
Medium (n=16,656)	5058 (30.4)	11,598 (69.6)		
High (n=786)	969 (28.2)	2472 (71.8)		
Number of friends			375.08 (1)	<.001
≤2 (n=16,039)	5512 (34.4)	10,527 (65.6)		
≥3 (n=6829)	1724 (25.2)	5105 (74.8)		
Learning burden			1013.24 (2)	<.001
Low (n=1481)	367 (24.8)	1114 (75.2)		
Medium (n=13,694)	3377 (24.7)	10,317 (75.3)		
High (n=7693)	3492 (45.4)	4201 (54.6)		
Family history of psychosis			67.95 (1)	<.001
Yes (n=1367)	570 (41.7)	797 (58.3)		
No (n=21,501)	6666 (31)	14,835 (69)		

Correlation of Various Health Risk Behaviors

Most studies have shown that HRBs often appear in clusters, which indicates there is usually a certain correlation between various HRBs. Thus, we also performed a correlation analysis of poor diet, smoking, physical inactivity, and poor sleep; the results are shown in Figure 1. There were statistically significant associations between each HRB (*P*<.05). Of note, there was evidence that other HRBs are positively related in addition to the negative correlation between physical inactivity and smoking. Furthermore, poor diet and poor sleep had the strongest correlation.

Figure 1. Correlation plot of health risk behaviors. Significant level: *P*<.05.



Association of Health Risk Behaviors With Comorbid Anxiety and Depression

Table 2 shows the chi-square associations between anxiety-depression comorbidity and HRBs in adolescents. There were statistically significant associations with anxiety-depression comorbidity for both single HRBs and clustered HRBs, all with a *P* value <.001. In terms of poor diet, physical inactivity, poor sleep, and clustered HRBs, we found that the higher the risk level of HRBs, the higher the likelihood of having comorbid anxiety and depression. However, the smoking medium-risk group had the highest rate of comorbid anxiety and depression compared to the smoking low-risk and high-risk groups.

Logistic regression analysis showed that HRBs were associated with comorbid anxiety and depression (Table 3). Regarding poor diet, adolescents in the medium-risk and high-risk groups were more likely to suffer from anxiety-depression comorbidity than those with low-risk HRBs. Adjusting for potential confounding factors partially attenuated these results, but comorbidity risk generally remained elevated with the increased risk of HRBs (medium risk OR 1.23, 95% CI 1.14-1.33; high risk OR 1.50, 95% CI 1.39-1.62). It is worth noting that the association between smoking and comorbid anxiety and depression was nonlinear, with a higher comorbidity rate in the medium-risk group (OR 2.74, 95% CI 2.14-3.51), whereas the comorbidity rate of anxiety and depression had a decreased tendency when risk increased (OR 2.17, 95% CI 1.67-2.81). The results were similar in the unadjusted model. Physical inactivity medium-risk and high-risk groups were associated with comorbid anxiety and depression without adjustment for any confounders, and the comorbidity rate increased with higher risk levels. However, results from the adjusted model confirmed evidence that there was a statistically significant association only in the high-risk group for physical inactivity (high risk OR 1.16, 95% CI 1.06-1.28). For poor sleep (medium risk OR 1.29, 95% CI 1.20-1.39; high risk OR 1.84, 95% CI 1.70-2.01) and clustered HRBs (medium risk OR 1.57, 95% CI 1.37-1.80; high risk OR 2.33, 95% CI 2.03-2.68), before and after adjusting the model, compared with the low risk group, other risk groups had a higher occurrence rate of anxiety-depression comorbidity, and the comorbidity rate increased as the risk level increased (*P*<.001).

On the basis of the above analysis, we performed a sensitivity analysis for gender stratification. After model adjustment, both OR values and 95% CIs were reduced (Multimedia Appendix 1, Table S2). For poor diet, physical inactivity, poor sleep, and clustered HRBs, boys at the same risk level had higher rates of anxiety-depression comorbidity than girls. Poor diet, poor sleep, and clustered HRB risk index were significant for both boys and girls before and after model adjustment, but no association of physical inactivity with anxiety-depression comorbidity was found for girls in the high-risk group before or after adjustment. It is worth noting that the comorbidity rate of anxiety and depression among girls based on smoking status (medium risk OR 3.94, 95% CI 2.24-6.90; high risk OR 5.00, 95% CI 1.77-14.12) was higher than boys (medium risk OR 2.49, 95% CI 1.88-3.28; high risk OR 2.07, 95% CI 1.58-2.72) at the same risk level, and the risk of smoking was nonlinearly associated with the comorbidity rate of anxiety and depression only in boys. Furthermore, logistic regression associations of clustered HRB risk index with comorbid anxiety and depression were stronger in boys than girls (Figure 2).

Table 2. Association of health risk behaviors with comorbid anxiety and depression.

HRBs[a]	Comorbid anxiety/depression, n (%)	No comorbid anxiety/depression, n (%)	Chi-square (*df*)	*P* value
Poor diet			234.88 (2)	<.001
Low risk (n=6082)	1519 (25)	4563 (75)		
Medium risk (n=7442)	2294 (30.8)	5148 (69.2)		
High risk (n=9344)	3423 (47.3)	5921 (63.4)		
Smoking			83.40 (2)	<.001
Low risk (n=22,336)	6972 (31.2)	15,364 (68.8)		
Medium risk (n=277)	145 (52.3)	132 (47.7)		
High risk (n=255)	119 (46.7)	136 (53.3)		
Physical inactivity			87.33 (2)	<.001
Low risk (n=3173)	837 (26.4)	2336 (73.6)		
Medium risk (n=5110)	1479 (28.9)	3631 (71.1)		
High risk (n=14,585)	4920 (33.7)	9665 (66.3)		
Poor sleep			474.57 (2)	<.001
Low risk (n=7472)	1793 (24)	5679 (76)		
Medium risk (n=8784)	2728 (31.1)	6056 (68.9)		
High risk (n=6612)	2715 (41.1)	3897 (31.6)		
HRB risk index			535.72 (2)	<.001
Low risk (n=1801)	294 (16.3)	1507 (83.7)		
Medium risk (n=8380)	2167 (25.9)	6213 (74.1)		
High risk (n=12,687)	4775 (37.6)	7912 (62.4)		

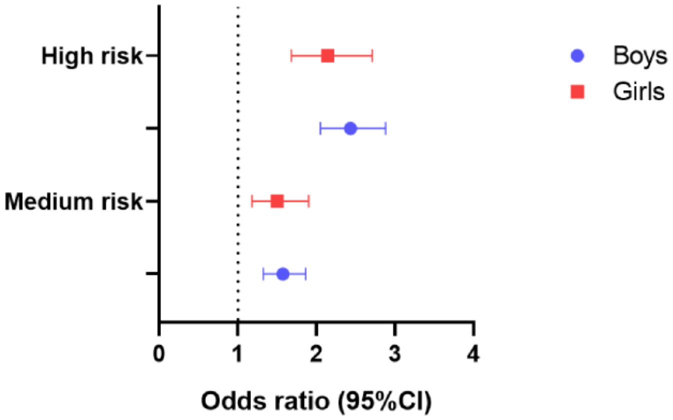
[a] HRB: health risk behavior.

Table 3. Association between health risk behaviors and comorbid anxiety and depression. All reference groups were the low risk groups. Model 1 did not adjust for any variable; model 2 adjusted for demographically meaningful variables such as gender and parental education.

HRBs[a]	Model 1 odds ratio (95% CI)	*P* value	Model 2 odds ratio (95% CI)	*P* value
Poor diet				
Medium risk	1.34 (1.24-1.44)	<.001	1.23 (1.14-1.33)	<.001
High risk	1.74 (1.62-1.87)	<.001	1.50 (1.39-1.62)	<.001
Smoking				
Medium risk	2.42 (1.91-3.07)	<.001	2.74 (2.14-3.51)	<.001
High risk	1.93 (1.51-2.47)	<.001	2.17 (1.67-2.81)	<.001
Physical inactivity				
Medium risk	1.14 (1.03-1.26)	.01	1.07 (0.96-1.18)	.22
High risk	1.42 (1.30-1.54)	<.001	1.16 (1.06-1.28)	<.001
Poor sleep				
Medium risk	1.43 (1.33-1.53)	<.001	1.29 (1.20-1.39)	<.001
High risk	2.21 (2.05-2.37)	<.001	1.84 (1.70-2.01)	<.001
HRB risk index				
Medium risk	1.79 (1.56-2.05)	<.001	1.57 (1.37-1.80)	<.001
High risk	3.09 (2.72-3.52)	<.001	2.33 (2.03-2.68)	<.001

[a] HRB: health risk behavior.

Figure 2. Logistic regression association between health risk behavior risk index and comorbid anxiety and depression in adjusted model. The adjusted model was adjusted for the presence of any siblings, regional economic level, educational status, self-rated health, parental education level, self-reported family income, number of friends, learning burden, and family history of psychosis.



Discussion

Principal Findings

This nationwide study showed that HRBs (poor diet, physical inactivity, smoking, and poor sleep), especially clustered HRBs, were positively associated with higher odds of comorbid anxiety and depression among Chinese adolescents. Furthermore, we confirmed boys had a stronger association between clustered HRBs and comorbid anxiety and depression than girls.

Comparison With Prior Work

The prevalence of comorbid anxiety and depression found in our study (7236/22,868, 31.6%) is higher than in another Chinese study (20.9%) [24] and a previous study carried out with college students in Brazil (27.8%) [25]. However, the prevalence in our study is lower than the estimate in an Italian study (the prevalence of comorbid anxiety and depression was 47%) [26]. The reason for the differences in comorbidity among these different studies may be differences in study areas, survey tools, and subjects, but the above studies all showed high comorbidity of anxiety and depression symptoms. As we all know, comorbid anxiety and depression is common in adolescents. The etiology of comorbid anxiety and depression is usually different from single anxiety or depression. In the common etiological model of comorbidity, anxiety and depression are caused by common environmental risk factors, and the common risk factors interact to promote comorbidity. This may become the core etiological model for explaining comorbid anxiety and depression in the future [27,28]. Furthermore, as shown, common modifiable risk factors for HRBs may lead to anxiety-depression comorbidity, and anxiety-depression comorbidity in turn can easily lead to risk behaviors in individuals, which may be a bidirectional-interaction process [29].

Those with a high HRB risk index, who comprised 55.5% (12,687/22,868) of the sample in our study, were more likely to have comorbid anxiety and depression. The results nonetheless indicate a nonlinear association between smoking and comorbid anxiety and depression, with a slightly higher likelihood of having comorbid anxiety and depression among those with medium smoking risk compared with high smoking risk. This may be related to the age of the adolescents, as a past study showed that the frequency with which adolescents engage in these risk behaviors across development is not always linear [30]. In addition, the reasons for the stronger association between smoking and comorbid anxiety and depression after controlling for confounders need to be explored. Particular behaviors may play different roles in the development of comorbid anxiety and depression in adolescence. Overall, although the results were attenuated, significant differences remained after adjusting for confounders.

Substantial evidence suggests that a range of psychiatric conditions are associated with individuals with adverse health behaviors compared to healthy control subjects [31]. However, few studies have examined the direct association between HRBs and comorbid anxiety and depression, and a majority of studies have examined associations between HRBs and single anxiety or depression. A study of health behaviors and mental health found that the group with the highest clustered health behaviors had the lowest risk of depression, anxiety, and stress [32]. Additionally, participants with multiple HRBs showed significantly increased risk of depression (OR 2.21, 95% CI 1.83-2.67) and anxiety (OR 2.32, 95% CI 1.85-2.92) [3]. Mortality risk and risk of poor health-related quality of life also increase in a dose-response manner as the number of HRBs increases [33]. These results indicate that mental health problems in adolescence are associated with a greater risk of engaging in a host of behaviors that increase individuals' risk for poor physical and mental health over the course of their lifespan.

Achieving good diet habits, not smoking, having adequate physical activity, and sleeping for an appropriate length of time have been individually associated with better mental health among adolescents [34-37]. However, the fact that the above HRBs have been considered separately is concerning, because research has shown that these behaviors are codependent and should be considered together. For example, in addition to being a risk factor for mental disease in and of itself, it is likely poor sleep also acts synergistically to increase the disease burden related to other risk behaviors, with evidence suggesting associations between poor sleep and risks of smoking, poor nutrition, harmful alcohol consumption, and physical inactivity (SNAP); this association may be stronger in people with mental health conditions than those without mental health conditions [38]. Combinations of HRBs have been shown to pose greater morbidity risk than the sum of their individual effects, suggesting a synergistic relationship between risk factors. Multiple-behavior change interventions may therefore have a greater potential for a positive impact on comorbid anxiety and depression than single-behavior change interventions.

There is a close association between HRBs and comorbid anxiety and depression, which may involve biological mechanisms. Multisystem dysfunction, inflammation, oxidative stress, and biological aging are considered to be key pathways in the evolution of mood disorders. Evidence suggests that engaging in HRBs also affects developing adolescents' brains. Adolescents' brains are sensitive to nicotine, especially in the still-developing prefrontal cortex, which is important for executive function and attention, and nicotine use affects cognition by modifying information processing. Thus, adolescent smokers are at increased risk for poor mental health in adulthood [39,40]. Furthermore, mental health in adolescence also worsens with poor nutrition and insufficient physical activity. Physical activity provides various biological effects via multiple mechanisms seen in animal models, including increased levels of brain-derived neurotrophic factor (BDNF) [41] and enhanced neurogenesis [42]. With increased expression of serotonin in animal models as a result of increased physical activity [43], this is theorized to explain much of the antidepressant effect of exercise. A reduction in oxidative stress and inflammatory cytokines has also been found to occur as a result of a healthy lifestyle [44], and this immunological effect may reduce bad mood via effects on the psychoneuroimmunological nexus [45].

Strengths and Limitations

Strengths of our study include the use of a large, nationally representative sample of adolescents, which allowed for adjustment for potential confounding variables that may have made our findings more compelling. Multistage cluster sampling and weighted statistical analyses adjusted for avoidable bias and allow us to generalize our data to the population level. Moreover, HRBs often appear in clusters, but few studies have investigated the 4 HRBs examined here (poor diet, physical inactivity, smoking, and poor sleep) at the same time, and there are relatively few investigations that have examined clustering of behaviors, including sleep. Our study is thus the first large-scale study to simultaneously investigate and combine specific HRBs with comorbid anxiety and depression. Last, sensitivity analyses, such as gender stratification of HRBs, were carried out so as to enable gender-specific measures.

Nonetheless, we recognize the following limitations in our study: First, it was cross-sectional, a study design that limits determination of the causal relationship of HRBs and comorbid anxiety and depression among adolescents. Our research is still ongoing, and further longitudinal analysis will be performed in the future. Second, data on sleep patterns were based on subjective responses of the adolescents, and we could not prevent recall bias. Further studies should be undertaken that include more objective measures of sleep patterns, such as polysomnography and actigraphy. Third, anxiety and depression symptoms were self-reported, which may have led to variation, but the GAD-7 and PHQ-9 are valid instruments with high internal consistency, good test-retest reliability, and good convergence with other types of instruments. Finally, confounding factors that may have contributed to anxiety-depression comorbidity may not have been measured and included in our analyses; future studies should take this into account and incorporate it into the analysis.

Conclusions

The prevalence of comorbid anxiety and depression was 31.6% (7236/22,868) in this national sample of Chinese adolescents. Single HRBs, including poor diet, smoking, physical inactivity, and poor sleep, were all associated with comorbid anxiety and depression in adolescents. Furthermore, clustered HRBs increase the risk of comorbid anxiety and depression compared to single HRBs. These findings should be replicated in longitudinal studies to determine the direction of this association. Our study provides evidence that maintaining a healthy lifestyle pattern, such as maintaining high levels of physical activity and not smoking, is advantageous to healthy development in Chinese adolescents.

Acknowledgments

We are extremely grateful to all participants and their parents who took part in the National Adolescents Cohort (China). This study was supported by the National Natural Science Foundation of China (grant 82173542) and the Grants for Scientific Research Level Improvement Plan from Anhui Medical University (grant 2022xkjT004).

Data Availability

The data sets generated during and/or analyzed during this study are not publicly available because they contain information that could compromise the privacy of research participants, but are available from the corresponding author on reasonable request.

Authors' Contributions

XW and FT conceived and designed the experiments; ST, YW, TL, XM, YX, YZ, and MW performed the study; MW, YX, and YZ analyzed the data; MW wrote the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table S1 (study population, distribution of main variables and risk scores) and Table S2 (association analysis of health risk behaviors with comorbidity of anxiety and depression by gender).

[DOCX File , 18 KB - publichealth_v9i1e46289_app1.docx]

References

1. McCloughen A, Foster K, Huws-Thomas M, Delgado C. Physical health and wellbeing of emerging and young adults with mental illness: an integrative review of international literature. Int J Ment Health Nurs 2012 Jun;21(3):274-288. [doi: 10.1111/j.1447-0349.2011.00796.x] [Medline: 22533335]
2. The health of young people: a challenge and a promise. World Health Organization. 1993. URL: https://apps.who.int/iris/handle/10665/37353 [accessed 2023-06-29]
3. Ye Y, Wang P, Qu G, Yuan S, Phongsavan P, He Q. Associations between multiple health risk behaviors and mental health among Chinese college students. Psychol Health Med 2016;21(3):377-385. [doi: 10.1080/13548506.2015.1070955] [Medline: 26222809]
4. Bennasar-Veny M, Yañez AM, Pericas J, Ballester L, Fernandez-Dominguez JC, Tauler P, et al. Cluster analysis of health-related lifestyles in university students. Int J Environ Res Public Health 2020 Mar 09;17(5):1776 [FREE Full text] [doi: 10.3390/ijerph17051776] [Medline: 32182922]
5. Yi X, Liu Z, Qiao W, Xie X, Yi N, Dong X, et al. Clustering effects of health risk behavior on mental health and physical activity in Chinese adolescents. Health Qual Life Outcomes 2020 Jul 03;18(1):211 [FREE Full text] [doi: 10.1186/s12955-020-01468-z] [Medline: 32620107]
6. Axelson DA, Birmaher B. Relation between anxiety and depressive disorders in childhood and adolescence. Depress Anxiety 2001;14(2):67-78. [doi: 10.1002/da.1048] [Medline: 11668659]
7. Choi K, Kim Y, Jeon H. Comorbid anxiety and depression: Clinical and conceptual consideration and transdiagnostic treatment. Adv Exp Med Biol 2020;1191:219-235. [doi: 10.1007/978-981-32-9705-0_14] [Medline: 32002932]
8. Wickersham A, Leightley D, Archer M, Fear NT. The association between paternal psychopathology and adolescent depression and anxiety: A systematic review. J Adolesc 2020 Feb;79:232-246. [doi: 10.1016/j.adolescence.2020.01.007] [Medline: 31986478]
9. Zhou S, Zhang L, Wang L, Guo Z, Wang J, Chen J, et al. Prevalence and socio-demographic correlates of psychological health problems in Chinese adolescents during the outbreak of COVID-19. Eur Child Adolesc Psychiatry 2020 Jun;29(6):749-758 [FREE Full text] [doi: 10.1007/s00787-020-01541-4] [Medline: 32363492]
10. Massetti GM, Thomas CC, King J, Ragan K, Buchanan Lunsford N. Mental health problems and cancer risk factors among young adults. Am J Prev Med 2017 Sep;53(3S1):S30-S39 [FREE Full text] [doi: 10.1016/j.amepre.2017.04.023] [Medline: 28818243]
11. Smout A, Newton NC, Slade T, O'Donoghue B, Chapman C. The relationship between early risk-taking behavior and mental health problems among a nationally representative sample of Australian youth. J Affect Disord 2020 Jul 01;272:239-248. [doi: 10.1016/j.jad.2020.04.040] [Medline: 32553364]
12. Jao NC, Robinson LD, Kelly PJ, Ciecierski CC, Hitsman B. Unhealthy behavior clustering and mental health status in United States college students. J Am Coll Health 2019;67(8):790-800 [FREE Full text] [doi: 10.1080/07448481.2018.1515744] [Medline: 30485154]
13. Yang Y, Qi Y, Cui Y, Li B, Zhang Z, Zhou Y, et al. Emotional and behavioral problems, social competence and risk factors in 6-16-year-old students in Beijing, China. PLoS One 2019;14(10):e0223970 [FREE Full text] [doi: 10.1371/journal.pone.0223970] [Medline: 31647827]
14. Firth J, Solmi M, Wootton RE, Vancampfort D, Schuch FB, Hoare E, et al. A meta-review of "lifestyle psychiatry": the role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. World Psychiatry 2020 Oct;19(3):360-380 [FREE Full text] [doi: 10.1002/wps.20773] [Medline: 32931092]
15. Xiao W, Xu H, Yu W, Li S, Li R, Jin Z, et al. Comparison of adverse childhood experience analytic approaches and associations with emotional and behavioral problems: A nationwide study among Chinese middle school students. J Affect Disord 2023 Mar 15;325:755-761. [doi: 10.1016/j.jad.2023.01.060] [Medline: 36682695]
16. Kann L, McManus T, Harris WA, Shanklin SL, Flint KH, Queen B, et al. Youth risk behavior surveillance - United States, 2017. MMWR Surveill Summ 2018 Jun 15;67(8):1-114 [FREE Full text] [doi: 10.15585/mmwr.ss6708a1] [Medline: 29902162]
17. Hutchesson MJ, Duncan MJ, Oftedal S, Ashton LM, Oldmeadow C, Kay-Lambkin F, et al. Latent class analysis of multiple health risk behaviors among Australian university students and associations with psychological distress. Nutrients 2021 Jan 28;13(2):425 [FREE Full text] [doi: 10.3390/nu13020425] [Medline: 33525585]

18. Oellingrath IM, De Bortoli MM, Svendsen MV, Fell AKM. Lifestyle and work ability in a general working population in Norway: a cross-sectional study. BMJ Open 2019 Apr 03;9(4):e026215 [FREE Full text] [doi: 10.1136/bmjopen-2018-026215] [Medline: 30948597]
19. Cheung FTW, Ho AWY, Chan JWY, Li X, Chan NY, Zhang J, et al. Validation of the Chinese version of the Munich Chronotype Questionnaire (MCTQ) in Hong Kong Chinese youths. Chronobiol Int 2022 May;39(5):678-689. [doi: 10.1080/07420528.2022.2025821] [Medline: 35086397]
20. Paruthi S, Brooks LJ, D'Ambrosio C, Hall WA, Kotagal S, Lloyd RM, et al. Recommended amount of sleep for pediatric populations: a consensus statement of the American Academy of Sleep Medicine. J Clin Sleep Med 2016 Jun 15;12(6):785-786 [FREE Full text] [doi: 10.5664/jcsm.5866] [Medline: 27250809]
21. Liang D. The stress level of college students and its relationship with physical exercise. Zhongguo Xin Li Wei Sheng Za Zhi 1994;8(1):5-6.
22. Mason JE, LeBouthillier DM, Asmundson GJ. Relationships between health behaviors, posttraumatic stress disorder, and comorbid general anxiety and depression. Cogn Behav Ther 2019 May;48(3):184-199. [doi: 10.1080/16506073.2018.1498119] [Medline: 30081716]
23. Levis B, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. BMJ 2019 Apr 09;365:l1476 [FREE Full text] [doi: 10.1136/bmj.l1476] [Medline: 30967483]
24. Zhou S, Wang L, Qi M, Yang X, Gao L, Zhang S, et al. Depression, anxiety, and suicidal ideation in Chinese university students during the COVID-19 pandemic. Front Psychol 2021;12:669833 [FREE Full text] [doi: 10.3389/fpsyg.2021.669833] [Medline: 34421725]
25. Jenkins PE, Ducker I, Gooding R, James M, Rutter-Eley E. Anxiety and depression in a sample of UK college students: a study of prevalence, comorbidity, and quality of life. J Am Coll Health 2021;69(8):813-819. [doi: 10.1080/07448481.2019.1709474] [Medline: 31995452]
26. Bertani D, Mattei G, Ferrari S, Pingani L, Galeazzi G. Anxiety, depression and personality traits in Italian medical students. Riv Psichiatr 2020;55(6):342-348. [doi: 10.1708/3503.34892] [Medline: 33349727]
27. Waszczuk MA, Zavos HMS, Gregory AM, Eley TC. The stability and change of etiological influences on depression, anxiety symptoms and their co-occurrence across adolescence and young adulthood. Psychol Med 2016 Jan;46(1):161-175 [FREE Full text] [doi: 10.1017/S0033291715001634] [Medline: 26310536]
28. Mathew AR, Pettit JW, Lewinsohn PM, Seeley JR, Roberts RE. Co-morbidity between major depressive disorder and anxiety disorders: shared etiology or direct causation? Psychol Med 2011 Oct;41(10):2023-2034 [FREE Full text] [doi: 10.1017/S0033291711000407] [Medline: 21439108]
29. Hoare E, Werneck AO, Stubbs B, Firth J, Collins S, Corder K, et al. Association of child and adolescent mental health with adolescent health behaviors in the UK Millennium Cohort. JAMA Netw Open 2020 Aug 03;3(8):e2011381 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.11381] [Medline: 32777059]
30. Chen P, Jacobson KC. Developmental trajectories of substance use from early adolescence to young adulthood: gender and racial/ethnic differences. J Adolesc Health 2012 Feb;50(2):154-163. [doi: 10.1016/j.jadohealth.2011.05.013] [Medline: 22265111]
31. Firth J, Siddiqi N, Koyanagi A, Siskind D, Rosenbaum S, Galletly C, et al. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. Lancet Psychiatry 2019 Aug;6(8):675-712. [doi: 10.1016/S2215-0366(19)30132-4] [Medline: 31324560]
32. Di Benedetto M, Towt CJ, Jackson ML. A cluster analysis of sleep quality, self-care behaviors, and mental health risk in Australian university students. Behav Sleep Med 2020;18(3):309-320. [doi: 10.1080/15402002.2019.1580194] [Medline: 30821507]
33. Oftedal S, Kolt GS, Holliday EG, Stamatakis E, Vandelanotte C, Brown WJ, et al. Associations of health-behavior patterns, mental health and self-rated health. Prev Med 2019 Jan;118:295-303. [doi: 10.1016/j.ypmed.2018.11.017] [Medline: 30476503]
34. Chaiton MO, Cohen JE, O'Loughlin J, Rehm J. A systematic review of longitudinal studies on the association between depression and smoking in adolescents. BMC Public Health 2009 Sep 22;9(1):356 [FREE Full text] [doi: 10.1186/1471-2458-9-356] [Medline: 19772635]
35. Poitras VJ, Gray CE, Borghese MM, Carson V, Chaput J, Janssen I, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. Appl Physiol Nutr Metab 2016 Jun;41(6 Suppl 3):S197-S239 [FREE Full text] [doi: 10.1139/apnm-2015-0663] [Medline: 27306431]
36. Chaput J, Gray CE, Poitras VJ, Carson V, Gruber R, Olds T, et al. Systematic review of the relationships between sleep duration and health indicators in school-aged children and youth. Appl Physiol Nutr Metab 2016 Jun;41(6 Suppl 3):S266-S282 [FREE Full text] [doi: 10.1139/apnm-2015-0627] [Medline: 27306433]
37. Lassale C, Batty GD, Baghdadli A, Jacka F, Sánchez-Villegas A, Kivimäki M, et al. Healthy dietary indices and risk of depressive outcomes: a systematic review and meta-analysis of observational studies. Mol Psychiatry 2019 Jul;24(7):965-986 [FREE Full text] [doi: 10.1038/s41380-018-0237-8] [Medline: 30254236]

38. Metse AP, Clinton-McHarg T, Skinner E, Yogaraj Y, Colyvas K, Bowman J. Associations between suboptimal sleep and smoking, poor nutrition, harmful alcohol consumption and inadequate physical activity ('SNAP risks'): A comparison of people with and without a mental health condition in an Australian community survey. Int J Environ Res Public Health 2021 Jun 01;18(11):5946 [FREE Full text] [doi: 10.3390/ijerph18115946] [Medline: 34206135]
39. Jin M, Yoon C, Ko H, Kim H, Kim A, Moon H, et al. The relationship of caffeine intake with depression, anxiety, stress, and sleep in Korean adolescents. Korean J Fam Med 2016 Mar;37(2):111-116 [FREE Full text] [doi: 10.4082/kjfm.2016.37.2.111] [Medline: 27073610]
40. Vermeulen JM, Wootton RE, Treur JL, Sallis HM, Jones HJ, Zammit S, et al. Smoking and the risk for bipolar disorder: evidence from a bidirectional Mendelian randomisation study. Br J Psychiatry 2021 Feb;218(2):88-94. [doi: 10.1192/bjp.2019.202] [Medline: 31526406]
41. Dauwan M, Begemann MJH, Slot MIE, Lee EHM, Scheltens P, Sommer IEC. Physical exercise improves quality of life, depressive symptoms, and cognition across chronic brain disorders: a transdiagnostic systematic review and meta-analysis of randomized controlled trials. J Neurol 2021 Apr;268(4):1222-1246 [FREE Full text] [doi: 10.1007/s00415-019-09493-9] [Medline: 31414194]
42. Ernst C, Olson A, Pinel J, Lam R, Christie B. Antidepressant effects of exercise: evidence for an adult-neurogenesis hypothesis? J Psychiatry Neurosci 2006 Mar;31(2):84-92 [FREE Full text] [Medline: 16575423]
43. Dey S, Singh R, Dey P. Exercise training: significance of regional alterations in serotonin metabolism of rat brain in relation to antidepressant effect of exercise. Physiol Behav 1992 Dec;52(6):1095-1099. [doi: 10.1016/0031-9384(92)90465-e] [Medline: 1283013]
44. Mastorakos G, Pavlatou M, Diamanti-Kandarakis E, Chrousos G. Exercise and the stress system. Hormones (Athens) 2005;4(2):73-89 [FREE Full text] [Medline: 16613809]
45. Heitmann H, Andlauer TFM, Korn T, Mühlau M, Henningsen P, Hemmer B, et al. Fatigue, depression, and pain in multiple sclerosis: How neuroinflammation translates into dysfunctional reward processing and anhedonic symptoms. Mult Scler 2022 Jun;28(7):1020-1027 [FREE Full text] [doi: 10.1177/1352458520972279] [Medline: 33179588]

Abbreviations

BDNF: brain-derived neurotrophic factor
GAD-7: Generalized Anxiety Disorder–7
HRB: health risk behavior
MCTQ-SF: Munich Chronotype Questionnaire Short Form
OR: odds ratio
PAR-3: Physical Activity Rating Scale
PHQ-9: Patient Health Questionnaire–9
SNAP: smoking, poor nutrition, harmful alcohol consumption, and physical inactivity
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 06.02.23; peer-reviewed by R Song, S Yoon; comments to author 26.04.23; revised version received 17.05.23; accepted 16.06.23; published 05.07.23.

Please cite as:
Wang M, Mou X, Li T, Zhang Y, Xie Y, Tao S, Wan Y, Tao F, Wu X
Association Between Comorbid Anxiety and Depression and Health Risk Behaviors Among Chinese Adolescents: Cross-Sectional Questionnaire Study
JMIR Public Health Surveill 2023;9:e46289
URL: https://publichealth.jmir.org/2023/1/e46289
doi:10.2196/46289
PMID:37405826

Review

Effectiveness of Multicomponent Interventions in Office-Based Workers to Mitigate Occupational Sedentary Behavior: Systematic Review and Meta-Analysis

Liying Zhou[1,2,3*], MD; Xinxin Deng[1,2,3*], MD; Kangle Guo[4], MD; Liangying Hou[1,2,3], MD, Dr med; Xu Hui[1,2,3], MD; Yanan Wu[1,2,3], MD; Meng Xu[1,2,3], MD; Yongsheng Wang[1,2,3], MD; Shanshan Liang[1,2,3], MD; Kehu Yang[1,2,3], Dr med, Prof Dr; Xiuxia Li[1,2,3], Dr med, Prof Dr

[1]Health Technology Assessment Center/Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China

[2]Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

[3]Key Laboratory of Evidence-Based Medicine and Knowledge Translation of Gansu Province, Lanzhou, China

[4]Infection Management Section, Gansu Provincial Hospital, Lanzhou, China

[*]these authors contributed equally

Corresponding Author:
Xiuxia Li, Dr med, Prof Dr
Health Technology Assessment Center/Evidence-Based Social Science Research Center
School of Public Health
Lanzhou University
199 Donggang West Road
Lanzhou, 730000
China
Phone: 86 13893624871
Email: lixiuxia@lzu.edu.cn

Abstract

Background: Sedentary time in workplaces has been linked to increased risks of chronic occupational diseases, obesity, and overall mortality. Currently, there is a burgeoning research interest in the implementation of multicomponent interventions aimed at decreasing sedentary time among office workers, which encompass a comprehensive amalgamation of individual, organizational, and environmental strategies.

Objective: This meta-analysis aims at evaluating the effectiveness of multicomponent interventions to mitigate occupational sedentary behavior at work compared with no intervention.

Methods: PubMed, Web of Science, EMBASE, and Cochrane Central Register of Controlled Trials (CENTRAL) databases were searched from database inception until March 2023 to obtain randomized controlled trials (RCTs) assessing the efficacy of multicomponent interventions on occupational sedentary behavior among office-based workers. Two reviewers independently extracted the data and assessed the risk of bias by using the Cochrane Collaboration’s risk of bias tool. The average intervention effect on sedentary time was calculated using Stata 15.1. Mean differences (MDs) with 95% CIs were used to calculate the continuous variables. Subgroup analyses were performed to determine whether sit-stand workstation, feedback, and prompt elements played an important role in multicomponent interventions. Further, the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) system was used to evaluate the certainty of evidence.

Results: A total of 11 RCTs involving 1894 patients were included in the analysis. Five studies were rated as low risk of bias, 2 as unclear risk of bias, and 4 as high risk. The meta-analysis results showed that compared with no intervention, multicomponent interventions significantly reduced occupational sitting time (MD=–52.25 min/8-h workday, 95% CI –73.06 to –31.44; *P*<.001) and occupational prolonged sitting time (MD=–32.63 min/8-h workday, 95% CI –51.93 to –13.33; *P*=.001) and increased occupational standing time (MD=44.30 min/8-h workday, 95% CI 23.11-65.48; *P*<.001), whereas no significant differences were found in occupational stepping time (*P*=.06). The results of subgroup analysis showed that compared with multicomponent interventions without installment of sit-stand workstations, multicomponent interventions with sit-stand workstation installment showed better effects for reducing occupational sitting time (MD=–71.95 min/8-h workday, 95% CI –92.94 to –51.15), increasing occupational standing time (MD=66.56 min/8-h workday, 95% CI 43.45-89.67), and reducing occupational prolonged sitting time (MD=–47.05 min/8-h workday, 95% CI –73.66 to –20.43). The GRADE evidence summary showed that all 4 outcomes were rated as moderate certainty.

Conclusions: Multicomponent interventions, particularly those incorporating sit-stand workstations for all participants, are effective at reducing workplace sedentary time. However, given their cost, further research is needed to understand the effectiveness of low-cost/no-cost multicomponent interventions.

(JMIR Public Health Surveill 2023;9:e44745) doi:10.2196/44745

KEYWORDS

sedentary behavior; workplace; multicomponent; systematic review

Introduction

For most working adults, the majority of their sedentary time is accrued during work time [1,2]. More sitting time was reported at work than in any other sitting activity such as watching television or using a computer at home on weekdays. Studies also revealed that the sitting time of work for full-time office workers accounted for approximately 60%-90% of the total daily sitting time in a workday [3,4]. In addition, there is evidence that working adults who spend long periods of time sitting at work do not necessarily compensate for their sitting by being more active outside work [5]. It is crucial to note that contemporary research indicates that excessive sedentary behavior is detrimentally linked to many health-related risks such as cardiovascular disease, unhealthy aging, musculoskeletal disorders, poor bone health, poor metabolic health, and all-cause mortality, especially when sedentary time accumulates in prolonged uninterrupted bouts [6,7]. The workplace has been highlighted by the World Health Organization as a vital setting for health promotion actions to reduce sedentary behavior [8].

According to 2 umbrella reviews, the utilization of electronic and mobile health tools such as mobile apps is associated with a reduction in sedentary behavior [9,10]. In addition, the recent umbrella reviews indicate that interventions targeting the physical environment, specifically the implementation of active workstations, represent the most efficacious category of interventions for mitigating sedentary behavior in workplaces [11,12]. Following the first randomized controlled trial (RCT) in 2013 [13], a surging multitude of subsequent RCTs has been undertaken to evaluate the efficacy of multicomponent interventions in mitigating sedentary behavior in workplaces among office-based employees [14,15]. The multicomponent interventions encompass a comprehensive amalgamation of diverse modalities, including individual strategies such as counselling, prompts, telephonic support, and motivational interviewing; environmental strategies such as active workstations, prompting posters, exercise sessions, and access to a gym; as well as organizational strategies such as workshops, site visits, consultations, and appointing team leaders, ambassadors, or management support. Currently, there is a burgeoning research interest for implementing multicomponent interventions aimed at mitigating sedentary behavior among office workers [16].

With the growing public health concerns surrounding sedentary behavior in the workplace among nonmanual employees, it is important to determine the effectiveness of multicomponent interventions for office-based workers. This is not only from a theoretical and best practice perspective but also due to the lack of a systematic summary on this topic. Multicomponent interventions are designed to address multiple levels of influence for sustainable behavior change, which aligns with the socioecological theory that emphasizes targeting multiple levels of influence [17]. Despite this, to date, evidence on the efficacy of multicomponent interventions has not been systematically summarized. Previous literature reviews on this topic have been constrained in their scope, as they have not concentrated specifically on the efficacy of multicomponent interventions in mitigating sedentary behavior. In addition, these reviews are outdated and have incorporated only a small number of RCTs in their analyses, which restricts the capacity to arrive at conclusive determinations regarding their impact [18,19].

Thus, this systematic review and meta-analysis includes newly published RCTs and aims to identify the effect of multicomponent interventions on sedentary workplace behavior based on more sedentary-related outcomes such as occupational sitting time, occupational standing time, occupational stepping time, and prolonged occupational sitting time. Furthermore, given that electronic and mobile health tools and physical environment interventions such as the implementation of an active workstation are independent and effective components affecting sedentary time, we performed subgroup analyses based on whether these components were included in the multicomponent interventions and attempted to identify the more effective components of multicomponent interventions.

Methods

Search Strategy

PubMed, Web of Science, EMBASE, and Cochrane Central Register of Controlled Trials (CENTRAL) databases were searched from database inception to March 1, 2023. The main search strategies were as follows: (occupation* or workplace* or employe* or office* or work-site or worker* or staff* or white-collar*) AND (sedentary or sitting or inactivity or "physical activity" or "physically active") AND (random* or blind* or singleblind* or doubleblind* or tripleblind* or RCT* or control*). The detailed search strategy is presented in Multimedia Appendix 1. In addition, the World Health Organization International Clinical Trials Registry Platform search portal, ClinicalTrials.gov, conference materials, and reference lists (backward and forward) of the studies identified using the above search strategy were searched manually for additional studies on March 1, 2023. We also searched relevant grey literature, including clinical guidelines, reports, and working papers through Google and grey literature databases [20].

Inclusion and Exclusion Criteria

Studies were included if they met all the eligibility criteria mentioned in Table 1.

Table 1. Inclusion and exclusion criteria for study selection.

	Inclusion criteria	Exclusion criteria
Participants	Workers aged ≥18 years who work in the office, do jobs that generally do not involve manual labor, or wear uniforms or work clothes were included	People working in the transportation industry (such as taxi drivers, truck drivers, bus drivers, and airline pilots) and those who operate heavy equipment (such as crane operators and bulldozer operators) were excluded
Intervention	Multicomponent interventions integrating individual strategies (eg, counselling, prompts, telephonic support, motivational interviews), environmental strategies (eg, active workstations, prompting posters, exercise sessions, access to a gym), and organizational strategies (eg, workshops, site visits, consultations, and appointment of team leaders, ambassadors, management support) aimed at changing sedentary behavior were eligible	Single-component interventions (such as implementation of a sit-stand workstation alone) or 2-level interventions consisting of only 2 types of strategies of the multicomponent intervention were excluded
Comparison	Usual work practices or waitlist	Other interventions such as single-component interventions, 2-level interventions, or multicomponent interventions were excluded
Outcome	Workplace sitting and activity outcomes during work hours: occupational sitting time, occupational standing time, occupational stepping time, and occupational prolonged sitting time as measured by self-report (eg, questionnaires) or using objective measures (eg, accelerometer) at the primary end point	Studies that reported daily sitting or activity outcomes during total waking hours rather than workplace sitting and activity outcomes specifically during work hours were excluded
Study design	Only studies with a parallel control (or treatment-comparison group) such as randomized controlled trials, controlled trials, cluster randomized controlled trials, or quasi-experimental studies were included in this review	Reviews, expert opinions, meta-analyses, or before-and-after studies were excluded
Language	English	Non-English

Study Selection and Data Extraction

Two independent reviewers performed title screening and data extraction of the retrieved studies, and any conflicts were resolved through a discussion. We used EndNote X9.1 software to omit duplicates. Subsequently, based on the inclusion and exclusion criteria, the 2 reviewers screened the titles and abstracts to discard irrelevant studies. Studies were removed from further review if both reviewers excluded them. Otherwise, full papers were obtained for a detailed review. We extracted the following data from the included studies by using a prespecified data form: general information (publication date, name of the first author, and study country/region), study population (age, gender, education, and employment status), multicomponent intervention (component, frequency, and duration of intervention; delivery mode; theoretical framework; status of sit-stand workstation installation; and inclusion of feedbacks and prompts), comparison intervention (waitlist or no intervention), outcomes (occupational sitting time, occupational standing time, occupational stepping time, and occupational prolonged sitting time [sitting time accumulated in bouts ≥30 minutes]), and follow-up time.

Risk of Bias Assessment

The Cochrane risk of bias tool was used to evaluate the quality of RCTs based on randomization and allocation concealment (selection bias), blinding of personnel and participants (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selection of reported results (reporting bias), and sources (other bias), and the RCTs were evaluated as having low, high, or unclear risk of bias [21]. Studies were rated as having a low risk of bias if all items were low risk. When 1 item was high risk, the study was rated as having a high risk of bias. The studies that did not fall into the abovementioned categories were rated as having an unclear risk of bias [22-24]. In all the included studies, blinding of the participants or personnel to the intervention and allocation concealment was not possible in accordance with the nature and aim of the interventions, wherein intervention participants typically underwent changes in environment (such as installation of sit-stand workstation). Therefore, the performance bias item and allocation concealment item were excluded from the bias assessment. However, for the allocation concealment item, trials were assessed as high risk if there was a contamination between intervention and control group participants, that is, if participants from the same office or company ended up in different groups. Participants in the control group are more likely to be less sedentary under the influence of individuals in the intervention group in the same office, regardless of group allocation [25-27]. Studies were considered to have a low risk of bias if measures were taken to minimize contamination, such as using cluster trials or assigning intervention and control participants to separate floors in the same building. Studies were classified as unclear risk of bias if there was insufficient information to determine the presence of either of the above conditions. For studies with multiple publications, we reviewed all relevant papers, including protocol papers, to ensure that the quality of the trial was judged on all available information.

Data Treatment and Statistical Analysis

Data synthesis was based on the general recommendations from the Cochrane Handbook for Systematic Reviews of Interventions [28]. Both the adjusted mean difference (AMD) and its SD were extracted for studies that reported AMD for the intervention and control groups after the intervention or change (change from baseline or change score) for time spent in sedentary behavior. For studies that did not report AMD or change, the unadjusted MD and SD were calculated from the means and SDs at baseline and after follow-up in each group. Therefore, the AMDs and unadjusted MDs were pooled together using the inverse variance method. For studies with multiple intervention arms, estimates were combined into a single group as a weighted average using Cochrane formula for combining subgroups if all the intervention arms are multicomponent. If, in a multiple arm trial, an arm was not a multicomponent intervention, then we excluded that arm.

We used minutes per 8 hours of work time at the workplace as the standard unit in the main group and subgroup analyses because it was the most commonly reported unit of measurement in the included trials. We converted all the measurement units for sitting at work into min/8-h workday, where needed and possible and assumed the data referred to a 5-day work week if this was not reported. Therefore, studies that reported min/8-h workday were combined with studies that reported min/d, min/workday, h/d, and min/week to measure the overall minutes in sedentary behavior per workday.

MD value with 95% CI based on the inverse variance method was used as the summary statistic for continuous variables, including sitting time, standing time, stepping time, and prolonged sitting time. Heterogeneity was evaluated using the Higgins I^2 value, and values ≤50% and >50% were considered to indicate low and high heterogeneity, respectively. MD values and the corresponding 95% CIs were calculated. $P<.05$ was considered to indicate statistical significance [29]. All statistical analyses were performed using Stata version 15.1 (StataCorp LLC). The random effects model was used. As mentioned in the Introduction, electronic and mobile health tools and physical environment interventions, especially installation of active workstations, are independent and effective components affecting sedentary time. Therefore, we performed subgroup analyses to determine whether sit-stand workstations, feedbacks (individual feedback on sedentary behavior at baseline and following the monitoring of activity during the trial), and prompts (individual prompts to regularly break up sitting through email, telephone, computer software, or prompting app) played an important role in multicomponent interventions. With regard to the sit-stand workstation element, we divided the included studies into 2 subgroups based on whether individual sit-stand workstations were installed in the trial: multicomponent intervention with sit-stand workstation installment and multicomponent intervention without sit-stand workstation installation. Based on prior research highlighting the significant influence of availability on the utilization of sit-stand workstation [30], we made the decision to mandate the complete availability of the sit-stand workstation within the subgroup of multicomponent interventions that involved the installation of such workstations. To maintain categorization consistency, the studies including workplaces that provide some sit-stand workstations available for shared use were classified into the subgroup of multicomponent intervention without sit-stand workstation installment. In the case of shared active workstations, office workers were required to temporarily vacate their original workstations to utilize shared facilities. In these instances, shared sit-stand workstation interventions were paired with other environmental strategy interventions such as access to a gym and the provision of company bikes but not with individual sit-stand workstations. Sensitivity analyses were performed on the outcome indicators of ≥10 studies to explore their potential sources and assess the robustness of these results. Egger test was used to assess publication bias [31].

Certainty Assessment

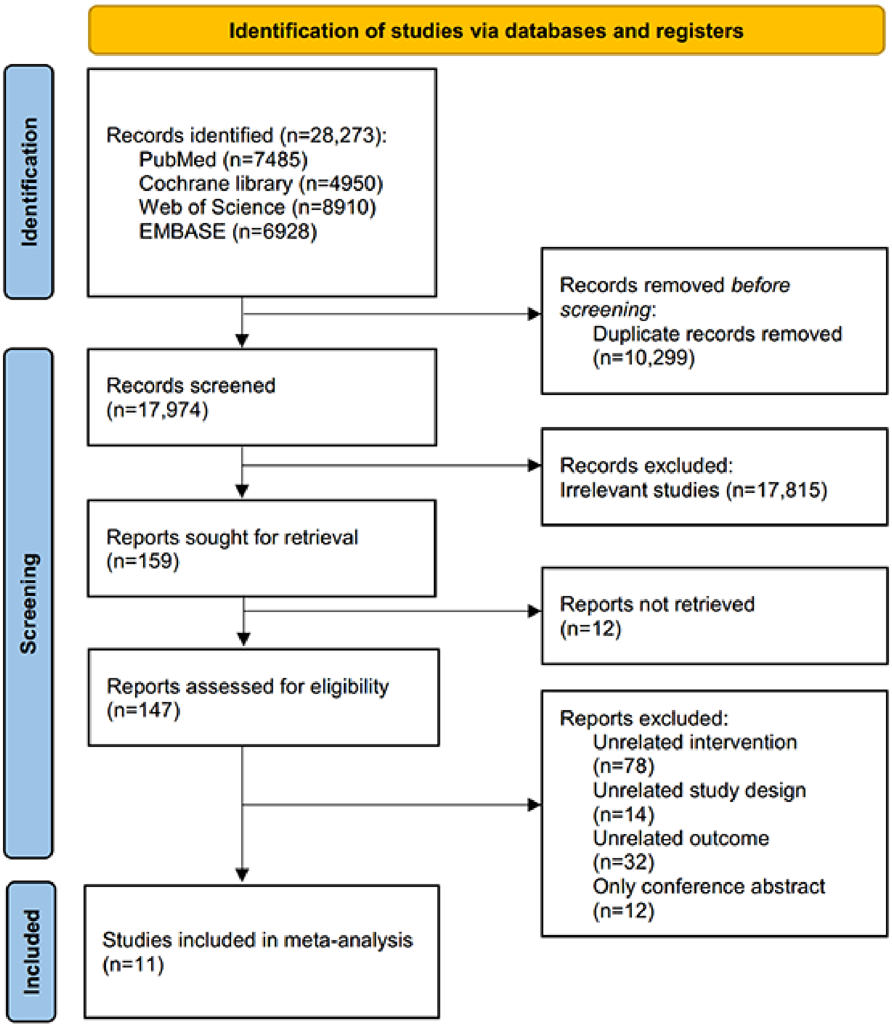
Two reviewers independently rated the certainty of evidence associated with specific outcomes and constructed a table summarizing the certainty of evidence findings using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) system [32,33]. The GRADE approach uses 5 domains, namely, risk of bias, inconsistency, indirectness, imprecision, and publication bias, which are assessed to determine the degree of confidence in the estimate of effect or association derived from the meta-analysis. Prespecified level of certainty was initially considered based on the study design (eg, high certainty for RCTs). Correspondingly, following the rating by reviewers, the certainty of evidence was changed and divided into no downgrade (not serious), downgrade by 1 level (serious), or downgrade by 2 levels (very serious). Finally, the certainty of evidence of each outcome was judged as different levels of evidence, that is, high, moderate, low, and very low.

Results

Study Selection

A flow diagram of the literature selection process is shown in Figure 1. A total of 28,273 relevant records were initially identified, of which 10,299 were excluded because of duplication. The titles and abstracts were screened, and 17,815 were deemed unsuitable. After reading the full text, another 159 studies were excluded because of inappropriate study design or topic of research. Finally, 11 studies were included in this meta-analysis [13-15,34-41].

Figure 1. Flowchart of the selection of randomized controlled trials that assessed the effects of multicomponent interventions on the occupational sedentary behavior of office-based workers.



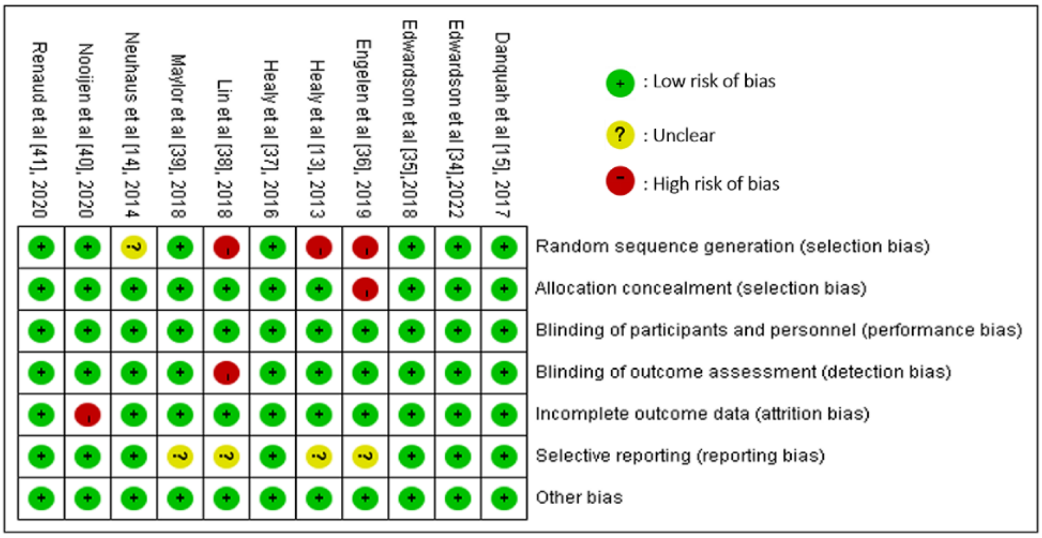
Main Characteristics of the RCTs

Of the 11 studies [13-15,34-41] included in this meta-analysis, 7 were cluster RCTs [15,34,35,37,39-41] and 4 were quasi-RCTs [13,14,36,38] (Multimedia Appendix 2). Regarding the location of the studies, 4 were conducted in Australia [13,14,36,37], 3 in England [34,35,39], and 1 each in Denmark [15], Taiwan [38], Sweden [40], and the Netherlands [41]. The studies were published between 2013 and 2022. All the trials compared the multicomponent intervention with no intervention among office-based workers. Multicomponent interventions aiming at reducing sedentary time in the workplace consisted of individual strategies, environmental strategies, and organizational strategies. In Multimedia Appendix 3 [13-15,34-41], a detailed account of the constituent elements and implementation particulars of multicomponent interventions implemented in each of the trials is presented. Seven of the 11 studies [13-15,34-37] had environmental components that implemented interventions to install sit-stand workstations, whereas the other 4 [38-41] did not. The studies included 1894 participants (1127 and 767 in the intervention and control groups, respectively) with a mean age ranging between 37.3 and 49.5 years. The sample size largely varied among the studies and ranged between 25 and 756. The intervention and follow-up periods varied from 1 month to 12 months. Seven studies [13, 14, 34, 35, 37, 39, 41] assessed sedentary behavior outcomes using the activPAL accelerometer, 3 studies [15, 36, 40] used the ActiGraph accelerometer, and only 1 [38] study used self-reported methods (ie, a self-efficacy scale). The dropout rates in the studies ranged from 1.98% to 35.52%.

Risk of Bias Assessment

As shown in Figure 2, the risk of bias was high in 4 studies [13,36,38,40], unclear in 2 [14,39], and low in 5 [15,34,35,37,41]. Regarding the random sequence generation assessment, 3 studies [13,36,38] did not adhere to random sequence generation, and thus we judged them to have a high risk of bias. Additionally, 1 trial [14] was assessed as unclear risk of bias because it gave no information about how randomization was done. For allocation concealment, 1 trial [36] was assessed as high risk due to contamination between the intervention and control group participants. Regarding outcome assessment, 1 trial [38] was rated as high risk of bias because its outcome measures were self-reported. Regarding incomplete outcome data, 1 study [40] was assessed as high risk of bias due to attrition rates of more than 25%. Regarding selection of the reported results, 4 trials [13,36,38,39] were assessed as unclear risk of bias due to there being no available trial registration or published protocol.

Figure 2. Risk of bias summary: the judgments of the reviewers about each risk of bias item for each of the included 11 randomized controlled trials using the Cochrane risk of bias tool [13-15,34-41].



	Renaud et al [41], 2020	Nooijen et al [40], 2020	Neuhaus et al [14], 2014	Maylor et al [39], 2018	Lin et al [38], 2018	Healy et al [37], 2016	Healy et al [13], 2013	Engelen et al [36], 2019	Edwardson et al [35],2018	Edwardson et al [34],2022	Danquah et al [15], 2017
Random sequence generation (selection bias)	+	+	?	+	−	+	−	−	+	+	+
Allocation concealment (selection bias)	+	+	+	+	+	+	+	−	+	+	+
Blinding of participants and personnel (performance bias)	+	+	+	+	+	+	+	+	+	+	+
Blinding of outcome assessment (detection bias)	+	+	+	+	−	+	+	+	+	+	+
Incomplete outcome data (attrition bias)	+	−	+	+	+	+	+	+	+	+	+
Selective reporting (reporting bias)	+	+	+	?	?	+	?	?	+	+	+
Other bias	+	+	+	+	+	+	+	+	+	+	+

+ : Low risk of bias

? : Unclear

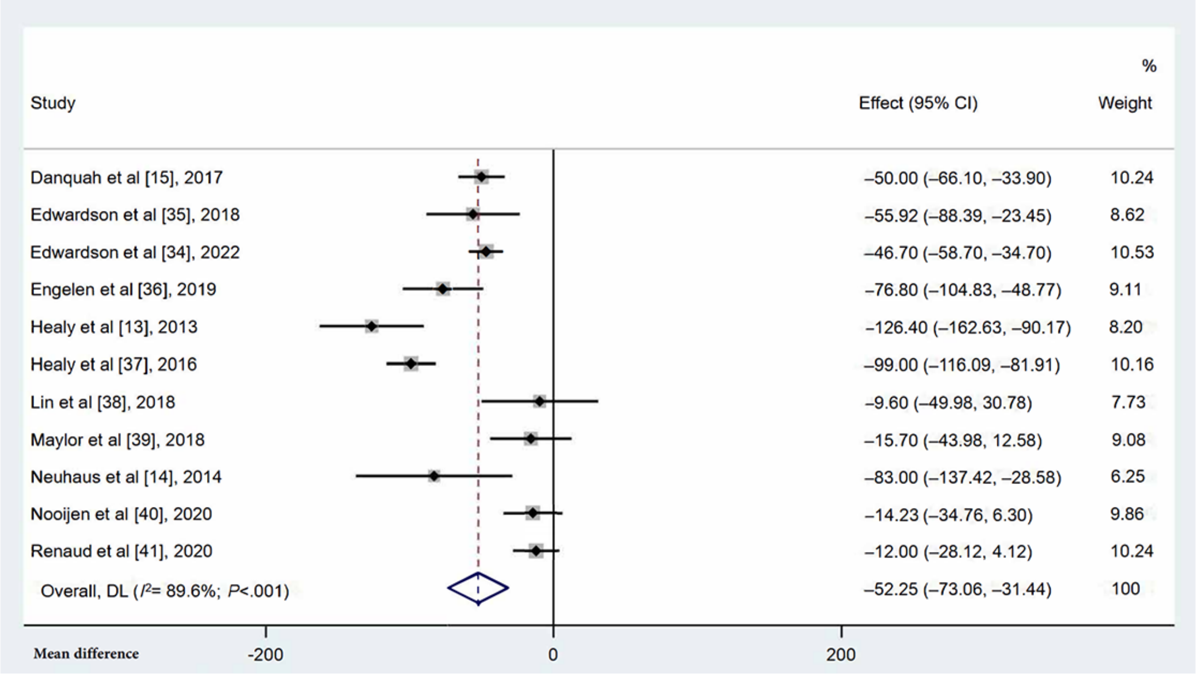
− : High risk of bias

Summary of the Intervention Effects

Occupational Sitting Time

All 11 included studies, with the multicomponent intervention and control groups including 1127 and 767 individuals, respectively, measured occupational sitting time (Figure 3). Participants in the multicomponent intervention groups had significantly shorter occupational sitting time than those in the control groups (MD=–52.25 min/8-h workday, 95% CI –73.06 to –31.44; I^2=89.6%; P<.001). Sensitivity analysis showed that no single study significantly affected the overall heterogeneity.

Figure 3. Forest plots for the effects of multicomponent interventions on occupational sitting time in office-based workers compared with usual practices [13-15,34-41]. Note: Weights are from random-effects model. DL: DerSimonian-Laird approach.

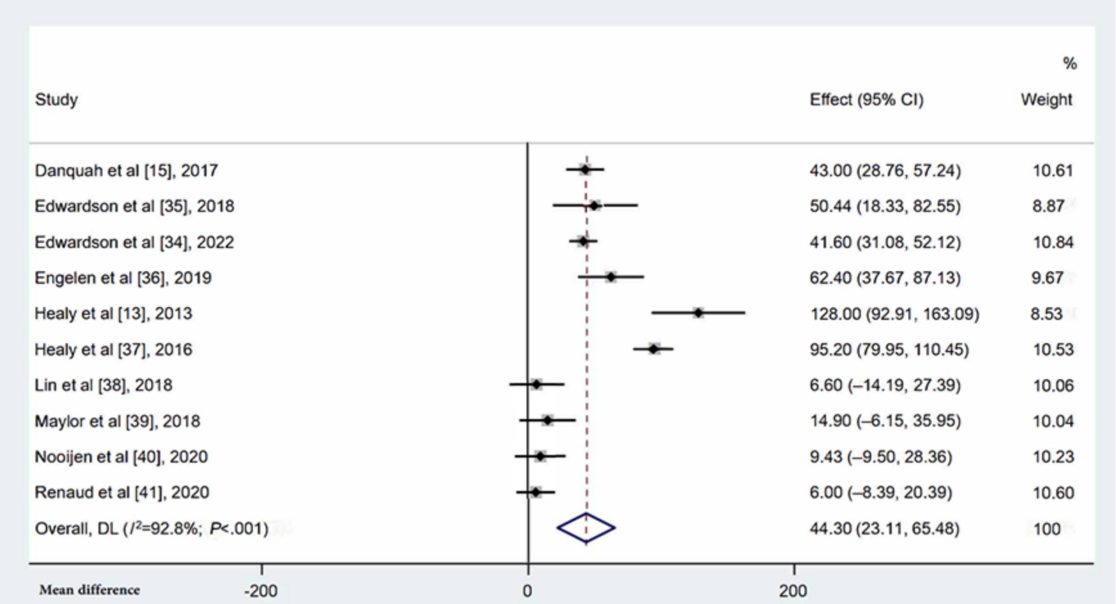


Study	Effect (95% CI)	% Weight
Danquah et al [15], 2017	–50.00 (–66.10, –33.90)	10.24
Edwardson et al [35], 2018	–55.92 (–88.39, –23.45)	8.62
Edwardson et al [34], 2022	–46.70 (–58.70, –34.70)	10.53
Engelen et al [36], 2019	–76.80 (–104.83, –48.77)	9.11
Healy et al [13], 2013	–126.40 (–162.63, –90.17)	8.20
Healy et al [37], 2016	–99.00 (–116.09, –81.91)	10.16
Lin et al [38], 2018	–9.60 (–49.98, 30.78)	7.73
Maylor et al [39], 2018	–15.70 (–43.98, 12.58)	9.08
Neuhaus et al [14], 2014	–83.00 (–137.42, –28.58)	6.25
Nooijen et al [40], 2020	–14.23 (–34.76, 6.30)	9.86
Renaud et al [41], 2020	–12.00 (–28.12, 4.12)	10.24
Overall, DL (I^2= 89.6%; P<.001)	–52.25 (–73.06, –31.44)	100

Occupational Standing Time

A total of 10 studies [13,15,34-41] with the multicomponent intervention and control groups including 1115 and 754 individuals, respectively, measured occupational standing time at the workplace (Figure 4). Participants who were in the multicomponent intervention group had significantly increased standing time compared with those in the control group (MD=44.30 min/8-h workday, 95% CI 23.11-65.48; I^2=92.8%; P<.001). Sensitivity analysis showed that no single study significantly affected the overall heterogeneity.

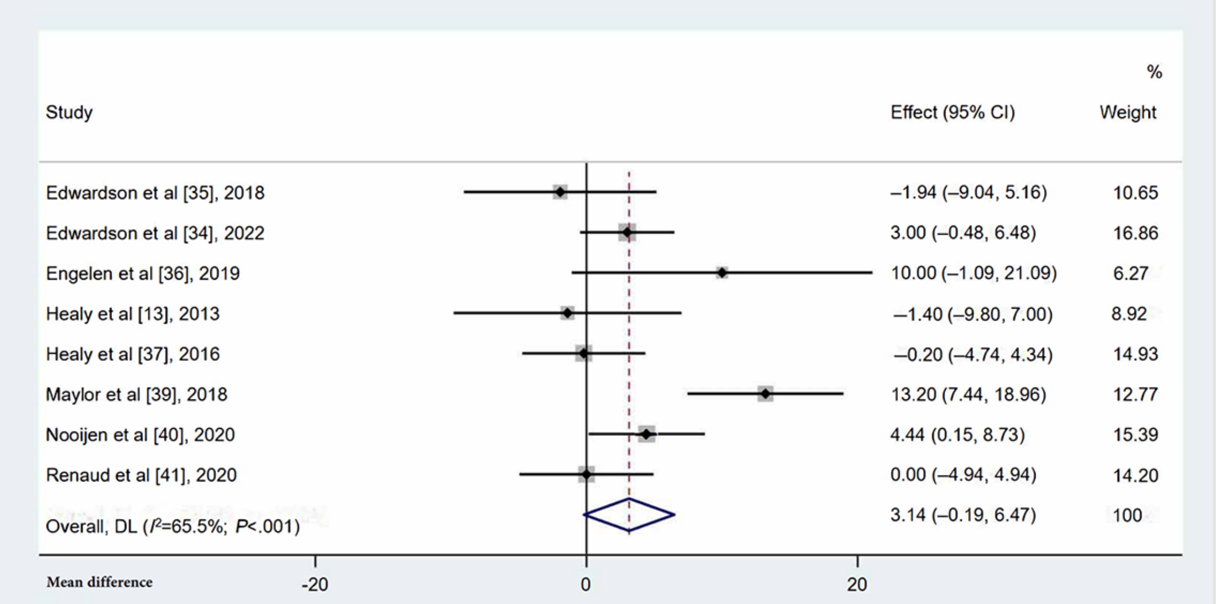
Figure 4. Forest plots for the effects of multicomponent interventions on occupational standing time in office-based workers compared with usual practices [13,15,34-41]. Note: Weights are from random-effects model. DL: DerSimonian-Laird approach.



Occupational Stepping Time

Of the 11 studies [13-15,34-41], 8 [13,34-37,39-41] reported the occupational stepping time of individuals (Figure 5). No statistically significant difference in occupational stepping time was observed between the 2 groups (MD=3.14 min/8-h workday, 95% CI –0.19 to 6.47; I^2=65.5%; P=.06).

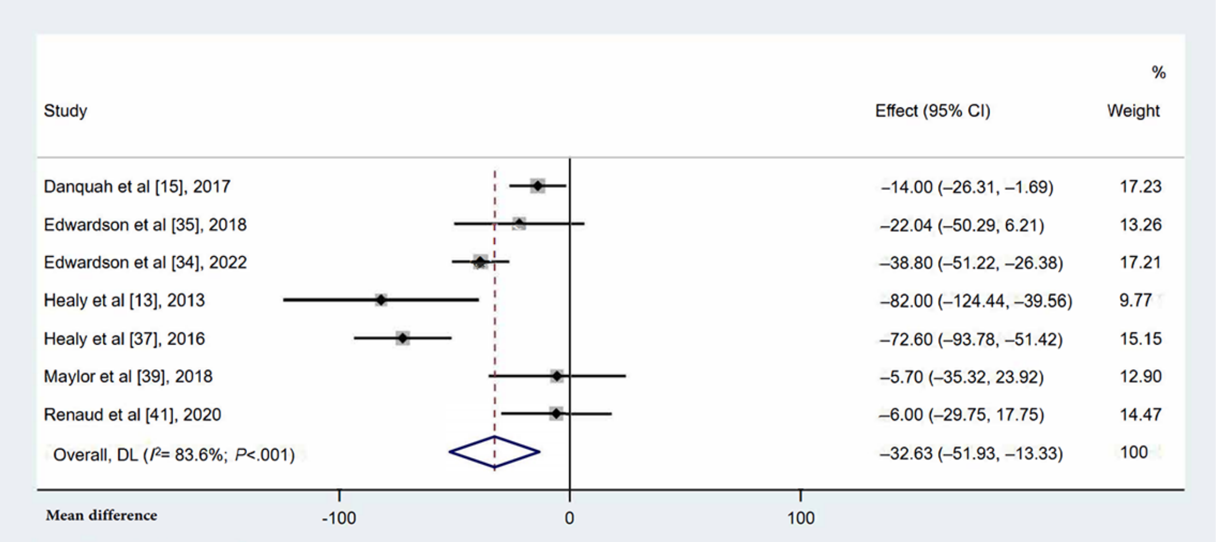
Figure 5. Forest plots for the effects of multicomponent interventions on occupational stepping time in office-based workers compared with usual practices [13,34-37,39-41]. Note: Weights are from random-effects model. DL: DerSimonian-Laird approach.



Occupational Prolonged Sitting Time

Seven studies [13,15,34,35,37,39,41] reported a reduction in occupational prolonged sitting time at work following multicomponent interventions (Figure 6). Multicomponent interventions significantly reduced prolonged sitting time (MD=–32.63 min/8-h workday, 95% CI –51.93 to –13.33; I^2=83.6%; P=.001).

Figure 6. Forest plots for the effects of multicomponent interventions on occupational prolonged sitting time in office-based workers compared with usual practices [13,15,34,35,37,39,41]. Note: Weights are from random-effects model. DL: DerSimonian-Laird approach.



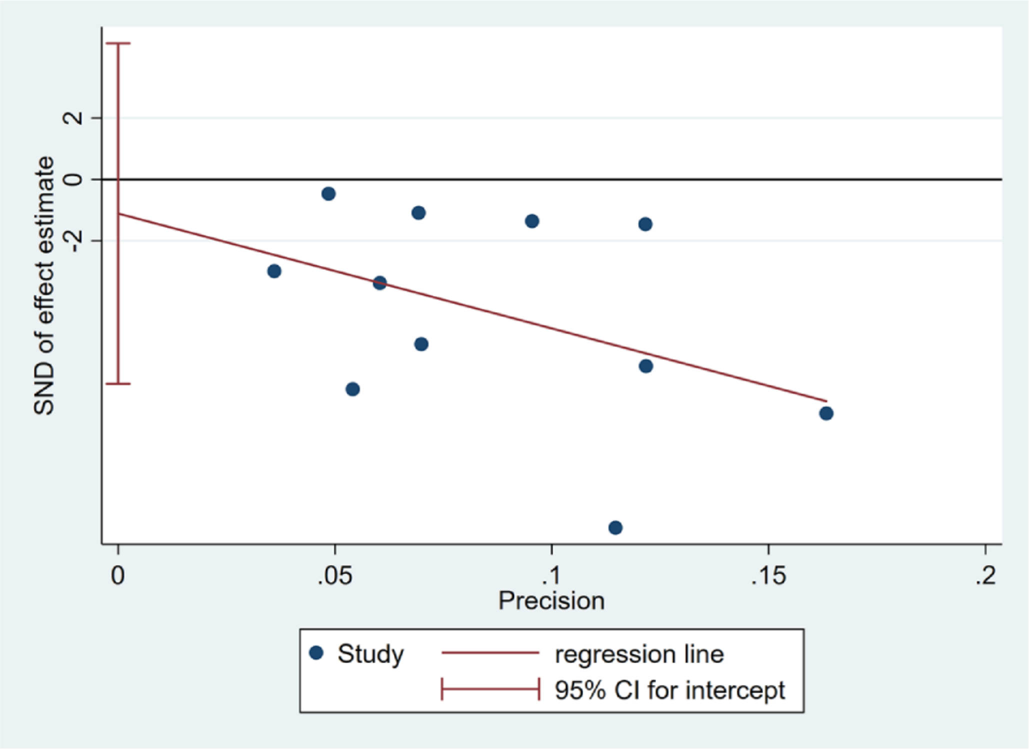
Subgroup Analyses

In the subgroup analyses for installment of sit-stand workstations during multicomponent interventions, the multicomponent intervention subgroup with sit-stand workstation installment showed considerable reduction in the occupational sitting time (MD=–71.95 min/8-h workday, 95% CI –92.94 to –51.15; I^2=84.2%; *P*<.001); however, the multicomponent intervention subgroup without sit-stand workstation installment showed a smaller reduction in the occupational sitting time (MD=–14.25 min/8-h workday, 95% CI –22.78 to –5.72; I^2=0%; *P*=.001). Furthermore, the results showed no heterogeneity in the subgroup without the installment of sit-stand workstations. The multicomponent intervention subgroup with the sit-stand workstation installment showed a great increase in the occupational standing time (MD=66.56 min/8-h workday, 95% CI 43.45-89.67; I^2=89.9%; *P*<.001), but the multicomponent intervention subgroup without the sit-stand workstation installment showed no such effect (MD=6.82 min/8-h workday, 95% CI –0.09 to 13.67; I^2=0%; *P*=.06). The results showed no heterogeneity in the subgroup analysis without the installment of sit-stand workstations. Further, the multicomponent intervention subgroup with the installment of sit-stand workstations showed a considerable reduction in occupational prolonged sitting time (MD=–47.05 min/8-h workday, 95% CI –73.66 to –20.43; I^2=88.7%; *P*=.001), but the multicomponent intervention subgroup without the installment of a sit-stand workstation showed a smaller reduction (MD=–11.49 min/8-h workday, 95% CI –22.78 to –0.19; I^2=0%; *P*=.04). On the contrary, the multicomponent intervention subgroup with the installment of sit-stand workstations showed no significant difference in stepping time (MD=1.53 min/8-h workday, 95% CI –0.90 to 3.96; I^2=5.8%; *P*=.22), but the subgroup without the installment of sit-stand workstations showed increased occupational stepping time (MD=5.22 min/8-h workday, 95% CI 0.31-10.13; I^2=74.9%; *P*=.04). On performing subgroup analysis based on inclusion of prompt and feedback elements in multicomponent interventions, no significant differences between subgroups were observed overall (Multimedia Appendix 4).

Publication Bias

A funnel plot based on Egger test was constructed, and no publication bias was found (*P*=.68) (Figure 7).

Figure 7. Egger publication bias plot for assessing publication bias among the 11 included randomized controlled trials. SND: standard normal deviation.



Certainty Assessment

The GRADE evidence summary of certainty showed that all 4 outcomes were rated as moderate certainty (occupational sitting time, occupational standing time, occupational stepping time, and occupational prolonged sitting time). The main reason we downgraded the certainty of evidence was limitations in study design, that is, high risk of bias in the studies, including nonrandomized allocation sequence, possible contamination between the intervention and control groups, self-reported outcome measures, a high rate of loss to follow-up, and unclear risk of bias in the studies, including no information about random allocation sequence generation, trial registration, or published protocol (Multimedia Appendix 5).

Discussion

Principal Findings

In this meta-analysis of RCTs, we included 11 studies [13-15,34-41] involving 1894 participants from 6 countries. The analyses showed, with moderate certainty of evidence, that multicomponent interventions were effective in reducing occupational sitting time, increasing occupational standing time, and reducing occupational prolonged sitting time. We did not find a statistically significant difference in the occupational stepping time outcome between the groups. Based on subgroup analysis results, multicomponent interventions with the installment of sit-stand workstations were significantly more effective in decreasing both occupational sitting time (*P*<.001) and occupational prolonged sitting time (*P*<.001) while increasing occupational standing time (*P*<.001). Conversely, the subgroup without the sit-stand workstation installment demonstrated only marginal reduction in occupational sitting time and occupational prolonged sitting time but demonstrated a minor increase in occupational stepping time. Minimal differences were observed between multicomponent intervention subgroups with the presence and absence of feedbacks and prompts in occupational sitting time, standing time, stepping time, and prolonged sitting time. More RCTs are still needed to strengthen the evidence body.

Potential Interpretations of the Findings

These findings may be explained by multicomponent interventions implemented in the included studies, which reduced sedentary behavior primarily by increasing sitting-to-standing transitions by using a sit-stand workstation rather than by increasing stepping [42]. Multicomponent interventions may primarily reduce occupational sitting and occupational prolonged sitting time and increase occupational standing time by promoting the use of sit-stand workstations among office workers in their environmental strategies through individual and organizational strategies. The provision of a sit-stand workstation permits continued work at a computer while standing as opposed to encouraging regular ambulation [43]. Therefore, sit-stand workstation installation may be required to significantly reduce occupational sitting time. However, the subgroup without the installment of sit-stand workstations exhibited only slight reductions in both occupational sitting time (MD=–14.25 min/8-h workday) and occupational prolonged sitting time (MD=–11.49 min/8-h workday) and a minor increase in occupational stepping time (MD=5.22 min/8-h workday). In addition, it should be noted that the increase in occupational stepping time associated with an 8-hour workday is minimal, and the effect size is negligible. Furthermore, the increased occupational sitting time only constitutes a small fraction of the overall reduction in occupational sitting time. The impact of increased occupational stepping time on reducing sedentary behavior in the workplace is limited. Consequently, we deduce that some multicomponent interventions implementing measures such as step/pedometer challenges, walking routes and resources, walking meetings, and lunch walks to reduce sedentary behaviors demonstrated limited effectiveness in increasing the stepping time of participants because most of these measures were of secondary and facilitative nature throughout the multicomponent intervention [35,44]. Moreover, the effectiveness of these measures may also be limited by space, feasibility, and convenience. Specifically, all interventions reduced the occupational total sitting time and prolonged sitting time, and most increased occupational standing time by using sit-stand workstations. These changes occurred across the workday among participants, although there was a wide individual variability in these changes [45,46].

This explanation can also be confirmed by Danquah et al [15], who suggested that the multicomponent intervention was effective in reducing sitting time and prolonged sitting time and in increasing the number of sit-to-stand transitions among office workers who already have sit-stand workstations (which is a standard equipment at most Danish workplaces, introduced for ergonomic considerations) [15]. In addition, Renaud et al [41] concluded that the multicomponent intervention had little to no effect in reducing occupational sitting time, which may be because of the relatively low intensity of the intervention, that is, it only involved the replacement of 25% of sitting workstations with sit-stand workstations. In addition, Michaud et al [47] conducted a cluster-RCT comparing multicomponent interventions with and without the incorporation of a sit-stand workstation and reported that the former intervention, with the installment of a sit-stand workstation, resulted in a reduction in occupational sitting time as well as total sitting time, thus strongly supporting our findings. Therefore, we believe that sit-stand workstations play a central role in multicomponent interventions and that the other components play a role in facilitating and increasing the effect.

Comparisons With Other Reviews

Our findings on the effects of multicomponent interventions are consistent with several other systematic reviews in this area, which have demonstrated the benefits of such interventions in reducing sedentary time among office workers [18,19]. However, in comparison with previous systematic reviews and meta-analyses that investigated the effectiveness of multicomponent interventions, we included 3 more outcomes (occupational standing time, occupational prolonged sitting time, and occupational stepping time). In addition, subgroup analyses, including the installation of sit-stand workstations, feedbacks, and prompts, were conducted to investigate the significant constituents of multicomponent interventions, predicated on the inclusion of sit-stand workstations, feedback, and prompt elements, which were found to be independent and efficacious factors influencing sedentary time in previous studies. In brief, we included 11 RCTs with 1894 participants, which was almost 3 times the number of RCTs and participants included in the previous meta-analyses. Nonetheless, we observed similar results suggesting that multicomponent interventions are effective in reducing sitting time. With a more robust analysis, we could conclude that multicomponent interventions are also effective in reducing prolonged sitting time and increasing standing time. Furthermore, we used the GRADE method to assess the certainty of the evidence for the effectiveness of multicomponent interventions in reducing occupational sitting time and occupational prolonged sitting time and increasing occupational standing time and occupational stepping time among office workers.

Assessment of Evidence Quality

Although the results of this meta-analysis are based on rigorously designed RCTs, they should be interpreted with caution because of the risk of bias. The 5 study designs [15,34,35,37,41] identified to have a low risk of bias met all standards. The remaining 4 studies [13,36,38,40] were deemed to have a high risk of bias, and 2 [14,39] were considered to have unclear risk of bias. This was mainly demonstrated in the fact that the allocation sequence was not randomized; there was possible contamination between the intervention and control groups; there was a high rate of loss to follow-up; and there was no available information about random allocation sequence generation, trial registration, or published protocol. This may have affected the authenticity of the reported results. Allocation concealment and blinding of participants and personnel was not possible once the study was underway because of the open plan nature of the environmental strategies in multicomponent interventions. However, 1 trial was assessed as high risk because of contamination between the intervention and control group participants; in this trial, participants from the same office or company were placed in different groups [36]. Participants in the control group were likely to be less sedentary because of the influence of the intervention group participants in the same office [25-27]. Therefore, while allocation concealment and blinding are not possible, the certainty of evidence was also downgraded 1 level to moderate due to the other limitations in the study design, such as the allocation sequence not being randomized, possible contamination between the intervention and control groups, a high rate of loss to follow-up, and no information about random allocation sequence generation, trial registration, or published protocol. These methodological limitations underscore the importance of future clinical trials adhering to robust study design principles and implementation guidelines. RCTs with rigorous randomization procedures should be prioritized to minimize bias and increase the validity of findings [48]. Additionally, adopting the CONSORT (Consolidated Standards of Reporting Trials) as the reporting standard can significantly improve study quality and transparency [49]. CONSORT guidelines provide a structured framework for reporting essential aspects of trial design, conduct, and analysis, thereby enabling readers to evaluate the study's validity and replicate the findings. To further enhance the scientific quality and reliability of RCTs, investigators should consider utilizing the Cochrane quality assessment tool [21]. This tool allows researchers to conduct a comprehensive self-examination of the study design, hypothesis formulation, data collection and analysis methods, and risk of bias assessment. By critically evaluating these aspects, researchers can identify and address potential limitations, thus strengthening the overall methodological rigor of the trial. Furthermore, the preregistration of studies and publishing research protocols can be beneficial for improving the transparency and trustworthiness of the evidence generated [50,51]. This practice helps reduce the risk of selective outcome reporting and ensures that the reported findings align with the initial study plan.

Strengths and Limitations

This review was based on RCTs with real-world data. We conducted a comprehensive systematic search of 4 primary databases. Any differences between reviewers during screening, data extraction, risk of bias assessment, and evidence certainty grading were resolved through a discussion. We used 4 outcome measures to assess the effects of multicomponent interventions on sedentary behavior in the workplace. Estimates from the fully adjusted models in each study were used in our analyses to reduce the potential for confounding. We used the GRADE approach to assess the certainty of evidence, which will allow the readers to clearly understand the uncertainty in results. Furthermore, we conducted sensitivity analyses to test the robustness of the results. No evidence of publication bias was found. However, when interpreting the results, the following limitations should be considered. First, the countries included in this review were mainly high-income countries, thereby limiting the generalizability of the results. Office workers in high-income countries pay more attention to physical health than those in low-income countries; accordingly, participants may have shown higher motivation during the trial. Second, differences among office workers may affect the reliability of the evidence obtained in this study, such as job type, length of work, levels of physical and cognitive loads across sectors and industries, and workload. Third, the limited sample size observed in certain studies, coupled with variations in measurement units, likely exerted an influential impact on the wide CIs, leading to imprecise estimations of effect sizes. Furthermore, some included studies have been evaluated to possess high and unclear risk of bias due to limitations in study designs. In the future, more high-quality cluster-RCTs with large sample sizes are needed to strengthen the evidence body. Fourth, a substantial degree of heterogeneity was evident in the findings. However, we found no heterogeneity in the multicomponent intervention without the installment of a sit-stand workstation, whereas the multicomponent intervention with the sit-stand workstation installment showed heterogeneity. A high degree of heterogeneity was still observed in the subgroup of multicomponent intervention with the sit-stand workstation installment. Therefore, we speculate that the installment of a sit-stand workstation as well as its specific type and function (eg, whether it was electronically regulated) constituted the primary sources of heterogeneity. Regrettably, due to inadequate description of these parameters in the included studies, we were unable to conduct more comprehensive subgroup analyses or meta-regression analyses to further explore the origins of heterogeneity.

Conclusions

Multicomponent interventions are indeed effective in reducing sedentary behavior in workplaces. In particular, multicomponent interventions including the component of sit-stand workstation installation are more effective in reducing sedentary behavior in workplaces. The reduction in sitting time was mainly replaced by standing time, which we presume primarily resulted from the increased use of the sit-stand workstation. Thus, the sit-stand workstation may be a key component of multicomponent interventions. Therefore, future research will benefit from assessing the relative contribution of different individual and organizational components for mitigating sedentary behavior in workplaces through factorial designs or multiple arms. Moreover, future studies should add detailed cost-effectiveness and time-of-day effect analyses while reporting multicomponent interventions. Further research to understand the effectiveness of low/no-cost multicomponent interventions is needed.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (72074103), the Fundamental Research Funds for the Central Universities (lzujbky-2021-ct06, lzujbky-2021-kb22), and the Gansu Special Project of Soft Science (20CX9ZA109). We would like to thank Shuya Ni, a master's student from Jinan University, for her help at various stages of this review.

Data Availability

The data set and any other materials of our study are available from the corresponding author on request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed search strategy for PubMed, Web of Science, EMBASE, and Cochrane Central Register of Controlled Trials databases.
[DOCX File , 29 KB - publichealth_v9i1e44745_app1.docx]

Multimedia Appendix 2
Basic characteristics of the included studies.
[DOCX File , 20 KB - publichealth_v9i1e44745_app2.docx]

Multimedia Appendix 3

Detailed account of the constituent elements and implementation particulars of the multicomponent interventions implemented in each of the included trials.

[DOCX File , 34 KB - publichealth_v9i1e44745_app3.docx]

Multimedia Appendix 4

Subgroup analyses of multicomponent intervention effects on the mitigation of occupational sedentary behavior based on the installation of sit-stand workstations, feedbacks, and prompts.

[DOCX File , 21 KB - publichealth_v9i1e44745_app4.docx]

Multimedia Appendix 5

Summary of findings for multicomponent interventions for mitigating occupational sedentary behavior among office-based workers.

[DOCX File , 18 KB - publichealth_v9i1e44745_app5.docx]

Multimedia Appendix 6

PRISMA checklist.

[DOCX File , 32 KB - publichealth_v9i1e44745_app6.docx]

References

1. Ishii K, Shibata A, Oka K. Work engagement, productivity, and self-reported work-related sedentary behavior among Japanese adults: a cross-sectional study. J Occup Environ Med 2018 Apr;60(4):e173-e177. [doi: 10.1097/JOM.0000000000001270] [Medline: 29280773]
2. Kazi A, Haslam C, Duncan M, Clemes S, Twumasi R. Sedentary behaviour and health at work: an investigation of industrial sector, job role, gender and geographical differences. Ergonomics 2019 Jan;62(1):21-30. [doi: 10.1080/00140139.2018.1489981] [Medline: 29925296]
3. Keown MK, Skeaff CM, Perry TL, Haszard JJ, Peddie MC. Device-measured sedentary behavior patterns in office-based university employees. J Occup Environ Med 2018 Dec;60(12):1150-1157. [doi: 10.1097/JOM.0000000000001467] [Medline: 30308622]
4. Rosenkranz SK, Mailey EL, Umansky E, Rosenkranz RR, Ablah E. Workplace sedentary behavior and productivity: a cross-sectional study. Int J Environ Res Public Health 2020 Sep 08;17(18):6535 [FREE Full text] [doi: 10.3390/ijerph17186535] [Medline: 32911740]
5. van Dommelen P, Coffeng JK, van der Ploeg HP, van der Beek AJ, Boot CRL, Hendriksen IJM. Objectively measured total and occupational sedentary time in three work settings. PLoS One 2016;11(3):e0149951 [FREE Full text] [doi: 10.1371/journal.pone.0149951] [Medline: 26937959]
6. Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, Tennant B, 2018 Physical Activity Guidelines Advisory Committee. Sedentary behavior and health: update from the 2018 physical activity guidelines advisory committee. Med Sci Sports Exerc 2019 Jun;51(6):1227-1241 [FREE Full text] [doi: 10.1249/MSS.0000000000001935] [Medline: 31095080]
7. Park JH, Moon JH, Kim HJ, Kong MH, Oh YH. Sedentary lifestyle: overview of updated evidence of potential health risks. Korean J Fam Med 2020 Nov;41(6):365-373 [FREE Full text] [doi: 10.4082/kjfm.20.0165] [Medline: 33242381]
8. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf [accessed 2023-06-30]
9. Fiedler J, Eckert T, Wunsch K, Woll A. Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. BMC Public Health 2020 Oct 23;20(1):1605 [FREE Full text] [doi: 10.1186/s12889-020-09700-7] [Medline: 33097013]
10. Rodríguez-González P, Iglesias D, Fernandez-Rio J, Gao Z. Effectiveness of interventions using apps to improve physical activity, sedentary behavior and diet: An umbrella review. Complement Ther Clin Pract 2023 Feb;50:101711. [doi: 10.1016/j.ctcp.2022.101711] [Medline: 36509032]
11. Lam K, Baurecht H, Pahmeier K, Niemann A, Romberg C, Biermann-Stallwitz J, et al. How effective and how expensive are interventions to reduce sedentary behavior? An umbrella review and meta-analysis. Obes Rev 2022 May;23(5):e13422. [doi: 10.1111/obr.13422] [Medline: 35068047]
12. Nguyen P, Le LK, Nguyen D, Gao L, Dunstan DW, Moodie M. The effectiveness of sedentary behaviour interventions on sitting time and screen time in children and adults: an umbrella review of systematic reviews. Int J Behav Nutr Phys Act 2020 Sep 21;17(1):117 [FREE Full text] [doi: 10.1186/s12966-020-01009-3] [Medline: 32958052]
13. Healy GN, Eakin EG, Lamontagne AD, Owen N, Winkler EAH, Wiesner G, et al. Reducing sitting time in office workers: short-term efficacy of a multicomponent intervention. Prev Med 2013 Jul;57(1):43-48 [FREE Full text] [doi: 10.1016/j.ypmed.2013.04.004] [Medline: 23597658]

14. Neuhaus M, Healy GN, Dunstan DW, Owen N, Eakin EG. Workplace sitting and height-adjustable workstations: a randomized controlled trial. Am J Prev Med 2014 Jan;46(1):30-40. [doi: 10.1016/j.amepre.2013.09.009] [Medline: 24355669]
15. Danquah IH, Kloster S, Holtermann A, Aadahl M, Bauman A, Ersbøll AK, et al. Take a Stand!-a multi-component intervention aimed at reducing sitting time among office workers-a cluster randomized trial. Int J Epidemiol 2017 Feb 01;46(1):128-140. [doi: 10.1093/ije/dyw009] [Medline: 27094749]
16. Nicolson GH, Hayes C, Darker C. A theory-based multicomponent intervention to reduce occupational sedentary behaviour in professional male workers: protocol for a cluster randomised crossover pilot feasibility study. Pilot Feasibility Stud 2020 Nov 10;6(1):175 [FREE Full text] [doi: 10.1186/s40814-020-00716-9] [Medline: 33292787]
17. Bohlen LC, Michie S, de Bruin M, Rothman AJ, Kelly MP, Groarke HNK, et al. Do combinations of behavior change techniques that occur frequently in interventions reflect underlying theory? Ann Behav Med 2020 Nov 01;54(11):827-842 [FREE Full text] [doi: 10.1093/abm/kaaa078] [Medline: 32959875]
18. Chu AHY, Ng SHX, Tan CS, Win AM, Koh D, Müller-Riemenschneider F. A systematic review and meta-analysis of workplace intervention strategies to reduce sedentary time in white-collar workers. Obes Rev 2016 May;17(5):467-481. [doi: 10.1111/obr.12388] [Medline: 26990220]
19. Shrestha N, Kukkonen-Harjula KT, Verbeek JH, Ijaz S, Hermans V, Pedisic Z. Workplace interventions for reducing sitting at work. Cochrane Database Syst Rev 2018 Dec 17;12(12):CD010912 [FREE Full text] [doi: 10.1002/14651858.CD010912.pub5] [Medline: 30556590]
20. OPENGREY.EU – Grey Literature Database. URL: http://www.opengrey.eu [accessed 2023-06-30]
21. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
22. Li X, Li Y, Guo K, Chen N, Cui X, Bai Z, et al. Evidence based social science in China paper 3: the quality of social science RCTs published from 2000-2020. J Clin Epidemiol 2022 Jan;141:64-73. [doi: 10.1016/j.jclinepi.2021.09.014] [Medline: 34520849]
23. Li Y, Cao L, Zhang Z, Hou L, Qin Y, Hui X, et al. Reporting and methodological quality of COVID-19 systematic reviews needs to be improved: an evidence mapping. J Clin Epidemiol 2021 Jul;135:17-28 [FREE Full text] [doi: 10.1016/j.jclinepi.2021.02.021] [Medline: 33657455]
24. Wang Y, Chen N, Guo K, Li Y, Yang C, Shang X, et al. Reporting and methodological quality of acupuncture network meta-analyses could be improved: an evidence mapping. J Clin Epidemiol 2023 Jan;153:1-12. [doi: 10.1016/j.jclinepi.2022.11.004] [Medline: 36351510]
25. Bergman F, Wahlström V, Stomby A, Otten J, Lanthén E, Renklint R, et al. Treadmill workstations in office workers who are overweight or obese: a randomised controlled trial. Lancet Public Health 2018 Nov;3(11):e523-e535 [FREE Full text] [doi: 10.1016/S2468-2667(18)30163-4] [Medline: 30322782]
26. Rhoads CH. The implications of "contamination" for experimental design in education. Journal of Educational and Behavioral Statistics 2011 Feb 01;36(1):76-104. [doi: 10.3102/1076998610379133]
27. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? BMJ 2001 Feb 10;322(7282):355-357 [FREE Full text] [doi: 10.1136/bmj.322.7282.355] [Medline: 11159665]
28. Higgins J, Thomas J, Chandler J. Cochrane Handbook for Systematic Reviews of Interventions. Oxford, UK: The Cochrane Collaboration; 2019.
29. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 06;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
30. Mantzari E, Galloway C, Wijndaele K, Brage S, Griffin SJ, Marteau TM. Impact of sit-stand desks at work on energy expenditure, sitting time and cardio-metabolic risk factors: Multiphase feasibility study with randomised controlled component. Prev Med Rep 2019 Mar;13:64-72 [FREE Full text] [doi: 10.1016/j.pmedr.2018.11.012] [Medline: 31304079]
31. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
32. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]
33. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]
34. Edwardson CL, Biddle SJH, Clemes SA, Davies MJ, Dunstan DW, Eborall H, et al. Effectiveness of an intervention for reducing sitting time and improving health in office workers: three arm cluster randomised controlled trial. BMJ 2022 Aug 17;378:e069288 [FREE Full text] [doi: 10.1136/bmj-2021-069288] [Medline: 35977732]
35. Edwardson CL, Yates T, Biddle SJH, Davies MJ, Dunstan DW, Esliger DW, et al. Effectiveness of the Stand More AT (SMArT) Work intervention: cluster randomised controlled trial. BMJ 2018 Oct 10;363:k3870 [FREE Full text] [doi: 10.1136/bmj.k3870] [Medline: 30305278]

36. Engelen L, Drayton BA, Young S, Daley M, Milton K, Bauman A, et al. Impact and process evaluation of a co-designed ‘Move More, Sit Less’ intervention in a public sector workplace. WOR 2019 Nov 21;64(3):587-599. [doi: 10.3233/wor-193020]
37. Healy GN, Eakin EG, Owen N, Lamontagne AD, Moodie M, Winkler EAH, et al. A cluster randomized controlled trial to reduce office workers' sitting time: effect on activity outcomes. Med Sci Sports Exerc 2016 Sep;48(9):1787-1797. [doi: 10.1249/MSS.0000000000000972] [Medline: 27526175]
38. Lin Y, Hong O, Lin C, Lu S, Chen M, Lee K. A "sit less, walk more" workplace intervention for office workers: long-term efficacy of a quasi-experimental study. J Occup Environ Med 2018 Jun;60(6):e290-e299. [doi: 10.1097/JOM.0000000000001299] [Medline: 29438155]
39. Maylor BD, Edwardson CL, Zakrzewski-Fruer JK, Champion RB, Bailey DP. Efficacy of a multicomponent intervention to reduce workplace sitting time in office workers: a cluster randomized controlled trial. J Occup Environ Med 2018 Sep;60(9):787-795. [doi: 10.1097/JOM.0000000000001366] [Medline: 29851741]
40. Nooijen CFJ, Blom V, Ekblom ?, Heiland EG, Larisch L, Bojsen-Møller E, et al. The effectiveness of multi-component interventions targeting physical activity or sedentary behaviour amongst office workers: a three-arm cluster randomised controlled trial. BMC Public Health 2020 Sep 01;20(1):1329 [FREE Full text] [doi: 10.1186/s12889-020-09433-7] [Medline: 32873260]
41. Renaud LR, Jelsma JG, Huysmans MA, van Nassau F, Lakerveld J, Speklé EM, et al. Effectiveness of the multi-component dynamic work intervention to reduce sitting time in office workers - Results from a pragmatic cluster randomised controlled trial. Appl Ergon 2020 Apr;84:103027. [doi: 10.1016/j.apergo.2019.103027] [Medline: 31987512]
42. Chau JY, der Ploeg HPV, van Uffelen JGZ, Wong J, Riphagen I, Healy GN, et al. Are workplace interventions to reduce sitting effective? A systematic review. Prev Med 2010 Nov;51(5):352-356. [doi: 10.1016/j.ypmed.2010.08.012] [Medline: 20801153]
43. MacEwen BT, MacDonald DJ, Burr JF. A systematic review of standing and treadmill desks in the workplace. Prev Med 2015 Jan;70:50-58. [doi: 10.1016/j.ypmed.2014.11.011] [Medline: 25448843]
44. Hadgraft NT, Brakenridge CL, LaMontagne AD, Fjeldsoe BS, Lynch BM, Dunstan DW, et al. Feasibility and acceptability of reducing workplace sitting time: a qualitative study with Australian office workers. BMC Public Health 2016 Sep 05;16(1):933 [FREE Full text] [doi: 10.1186/s12889-016-3611-y] [Medline: 27595754]
45. Neuhaus M, Eakin EG, Straker L, Owen N, Dunstan DW, Reid N, et al. Reducing occupational sedentary time: a systematic review and meta-analysis of evidence on activity-permissive workstations. Obes Rev 2014 Oct;15(10):822-838. [doi: 10.1111/obr.12201] [Medline: 25040784]
46. Torbeyns T, Bailey S, Bos I, Meeusen R. Active workstations to fight sedentary behaviour. Sports Med 2014 Sep;44(9):1261-1273. [doi: 10.1007/s40279-014-0202-x] [Medline: 24842828]
47. Michaud TL, You W, Estabrooks PA, Leonard K, Rydell SA, Mullane SL, et al. Cost and cost-effectiveness of the 'Stand and Move at Work' multicomponent intervention to reduce workplace sedentary time and cardiometabolic risk. Scand J Work Environ Health 2022 Jul 01;48(5):399-409 [FREE Full text] [doi: 10.5271/sjweh.4022] [Medline: 35333373]
48. Wang X, Chen Y, Akl EA, Tokalić R, Marušić A, Qaseem A, RIGHT working group. The reporting checklist for public versions of guidelines: RIGHT-PVG. Implement Sci 2021 Jan 11;16(1):10 [FREE Full text] [doi: 10.1186/s13012-020-01066-z] [Medline: 33430911]
49. Bennett JA. The Consolidated Standards of Reporting Trials (CONSORT): Guidelines for reporting randomized trials. Nurs Res 2005;54(2):128-132. [doi: 10.1097/00006199-200503000-00007] [Medline: 15778654]
50. Ioannidis JPA. How to make more published research true. PLoS Med 2014 Oct;11(10):e1001747 [FREE Full text] [doi: 10.1371/journal.pmed.1001747] [Medline: 25334033]
51. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. Proc Natl Acad Sci U S A 2018 Mar 13;115(11):2600-2606 [FREE Full text] [doi: 10.1073/pnas.1708274114] [Medline: 29531091]

Abbreviations

AMD: adjusted mean difference
CENTRAL: Cochrane Central Register of Controlled Trials
CONSORT: Consolidated Standards of Reporting Trials
GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
MD: mean difference
RCT: randomized controlled trial

Edited by A Mavragani, T Sanchez; submitted 01.12.22; peer-reviewed by A Martinko, G Healy; comments to author 23.02.23; revised version received 12.04.23; accepted 24.05.23; published 26.07.23.

Please cite as:
Zhou L, Deng X, Guo K, Hou L, Hui X, Wu Y, Xu M, Wang Y, Liang S, Yang K, Li X
Effectiveness of Multicomponent Interventions in Office-Based Workers to Mitigate Occupational Sedentary Behavior: Systematic Review and Meta-Analysis
JMIR Public Health Surveill 2023;9:e44745
URL: https://publichealth.jmir.org/2023/1/e44745
doi:10.2196/44745
PMID:37494100

Original Paper

Association Between Household Online Grocery Delivery Service Use and Food and Drink Purchase Behavior in England: Cross-Sectional Analysis

Amy Yau[1], PhD; Cherry Law[1,2], PhD; Laura Cornelsen[1], PhD; Jean Adams[3], PhD; Emma Boyland[4], PhD; Thomas Burgoine[3], PhD; Frank de Vocht[5,6], PhD; Martin White[3], MD; Steven Cummins[1], PhD

[1]Department of Public Health, Environments and Society, Faculty of Public Health and Policy, London School of Hygiene & Tropical Medicine, London, United Kingdom

[2]Department of Agri-Food Economics & Marketing, School of Agriculture, Policy and Development, University of Reading, Reading, United Kingdom

[3]Medical Research Council (MRC) Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom

[4]Department of Psychology, Institute of Population Health, University of Liverpool, Liverpool, United Kingdom

[5]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

[6]National Institute for Health Research Applied Research Collaboration West (NIHR ARC West), Bristol, United Kingdom

Corresponding Author:
Amy Yau, PhD
Department of Public Health, Environments and Society
Faculty of Public Health and Policy
London School of Hygiene & Tropical Medicine
15-17 Tavistock Place
London, WC1H 9SH
United Kingdom
Phone: 44 020 7636 8636
Email: amy.yau@lshtm.ac.uk

Abstract

Background: Online grocery delivery services (OGDSs) are a popular way of acquiring food. However, it is unclear whether OGDS use is associated with the healthiness of purchases and whether there are sociodemographic differences in OGDS use. If so, the increased prevalence of OGDS use may have implications for population diet, and differential OGDS use could contribute to diet inequalities.

Objective: This study aimed to examine whether OGDS use varies by sociodemographic characteristics and is associated with the amount and types of groceries purchased.

Methods: Item-level take-home food and drink purchase data (n=3,233,920 items) from households in London and the North of England were available from the 2019 UK Kantar fast-moving consumer goods panel (N=1911). Purchases were categorized as being bought online or in-store. We used logistic regression to estimate the likelihood of an above-median frequency of OGDS use by sociodemographic characteristics. We used Poisson regression to estimate the differences in energy and nutrients purchased by households that had above- and below-median OGDS use and the proportion of energy purchased from products high in fat, salt, and sugar (HFSS) online versus in-store among households that used both shopping methods (n=665).

Results: In total, 668 (35%) households used OGDSs at least once in 2019. Of the households that used OGDSs, the median use was 5 occasions in 2019. Households were more likely to have above-median use in London versus in the North of England (odds ratio 1.29, 95% CI 1.01-1.65) and if they had a higher annual household income (odds ratio 1.56, 95% CI 1.02-2.38 for ≥£50,000 [US $64,000] vs <£20,000 [$25,600]). Households with above-median OGDS use had a higher weekly mean purchase of energy by 1461 (95% CI 1448-1474) kcal per person compared with households with below-median OGDS use. For households that used a combination of in-store and online shopping, HFSS products made up a lower proportion (–10.1%, 95% CI –12% to –8.1%) of energy purchased online compared to in-store.

Conclusions: Differences in grocery purchases between households with above- and below-median OGDS use could have positive or negative consequences. The extra energy purchased among households with above-median OGDS use could lead to overconsumption or food waste, which has negative consequences for population and environmental health. Alternatively, this extra energy may be replacing out-of-home purchasing, which tends to be less healthy, and may be beneficial for the population diet. Households made fewer HFSS purchases when shopping online compared to in-store, which may be due to differences in the shopping environment or experience, such as fewer promotions and advertisements when shopping online or not having to transport and carry purchases home. As higher-income households used OGDS more frequently, the implications of this sociodemographic pattern on dietary inequalities must be explored.

(JMIR Public Health Surveill 2023;9:e41540) doi:10.2196/41540

KEYWORDS

food and beverages; food preferences; supermarkets; internet; consumer behavior; lifestyle; diet; inequality; food purchase; sociodemographic factors; grocery purchase; online grocery; online purchase; public health; online; delivery; grocery; diet

Introduction

There has been an increasing interest in whether the use of digital platforms to purchase foods impacts on the type of products purchased and, therefore, population diet and health. Much of this research has focused on the purchase of takeaway foods [1-3], but fewer studies have explored the use of online services for grocery purchases. The use of online grocery delivery services (OGDS) is an increasingly popular method of purchasing food and drink products for preparation and consumption at home. Although online delivery services have been available for over 20 years, they have grown rapidly in the past decade, with recent consumer adoption and growth in the use of these services further accelerated by the COVID-19 pandemic [4,5]. These services enable groceries to be ordered online for home delivery or to be picked up through "click and collect," where purchased products are selected and packed in-store ready for collection by the consumer. The industry predicts further growth, with around half of those who used OGDS for the first time during the pandemic intending to continue [6].

OGDS may plausibly have positive and negative impacts on dietary behaviors [7-10]. Positive impacts might arise through changes in the shopping experience compared to traditional in-store purchasing. For example, products are presented symbolically in online channels rather than physically in stores, which may decrease their vividness and reduce consumers' desire to purchase [11]. OGDS also increase the lag time between product choice and food acquisition [12]. It is theorized that people make more healthful food choices when the outcome (ie, acquisition or consumption of food) is further into the future [9], as there is a decreased focus on immediate gratification [11,12]. OGDS may also improve access to food, thereby potentially increasing access in communities with poor availability or quality of local food shopping opportunities or enabling those with poor access to transport to take advantage of home delivery [13-15]. Almost all households in Great Britain are covered by at least 1 supermarket delivery service, with 3 of the major supermarkets each providing coverage to over 98% of households [15].

There may also be negative effects on dietary behaviors. Sizing of products may be more difficult to determine online, leading to over- or underpurchasing. One UK study found that there were differences in the size of products, fewer price promotions, and less front-of-pack labeling for products available online when compared to similar products in-store [16]. Online delivery may also be more appealing for larger, heavier items and increase bulk buying (leading to overconsumption and waste), while consumers may be less inclined to purchase healthier perishable goods such as fresh fruit and vegetables [10].

Who accesses these services and what they purchase online may be influenced by sociodemographic characteristics, which could also plausibly influence dietary inequalities [8]. For example, many OGDSs have minimum spend requirements and delivery fees, meaning that more affluent households may have greater access to and use of these services compared to more disadvantaged households [10]. OGDS use may also be affected by household size, composition, age, and other sociodemographic characteristics. Similar differences have been observed for online delivery services for takeaway foods, where the use of these services was more likely among adults who were younger, male, from an ethnic minority background, highly educated, or had children [1]. Such differences may have implications for dietary inequalities if OGDS use influences purchasing behavior [7,8,11].

Whether OGDS affects the dietary quality of purchases has not been well studied. Some limited evidence from the United States suggests that the share of consumer expenditure on healthy products is higher when grocery purchases are made online compared to in-store [9,11,12,17]. From intervention studies that use online delivery services to improve diet and nutrition, including in low-income populations, there is some evidence that they can improve the healthiness of purchases [18,19].

Overall, the evidence on who uses OGDS and how the use of these services may have either positive or negative effects on dietary behaviors is limited. A deeper understanding of the ways that OGDSs are associated with purchasing behavior could help inform future policies and interventions aiming to improve population diet and reduce dietary inequalities. This study, using large-scale consumer panel data from England, aims to generate evidence in this area by (1) exploring the sociodemographic correlates of OGDS use and (2) exploring whether the use of OGDS is associated with overall grocery food and drink purchases and purchases of unhealthier products that are high in fat, salt, and sugar (HFSS).

Methods

Study Design

We used data on household take-home grocery food and drink purchasing from the 2019 Kantar fast-moving consumer goods (FMCG) panel. FMCG are products sold quickly and at a relatively low cost, including food and drink. Kantar, a commercial consumer data company, continually recruits around

30,000 households in the United Kingdom to a live panel via email and post using quota sampling. The households largely reflect the key sociodemographic characteristics of the geographical region they have been sampled from. The data available for this analysis were from a study examining the association between outdoor HFSS food and drink advertising restrictions in London and household food and drink purchases [20]. Households in our analysis were randomly selected from Kantar households in London or the North of England (North West, North East, and Yorkshire and the Humber regions). The included households (N=1911) recorded individual food and drink purchases (n=3,233,799 items) from December 31, 2018, to December 29, 2019. Our findings are reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Multimedia Appendix 1).

Ethical Considerations

Ethical approval was not required for the analysis of anonymized secondary data. Upon joining, participants agreed to the terms and conditions of the Kantar FMCG panel, which state that their data may be used for research purposes. Panelists receive around £100 (~US $128) worth of vouchers for their participation per year.

Purchases and Nutrient Data

Participating households recorded all food and drink items purchased and brought into the home using a handheld barcode scanner. Nonbarcoded products, such as loose fruits and vegetables, were recorded using bespoke barcodes. Kantar provided nutrient data based on direct assessment of product nutrient labels in outlets twice a year or using product nutrient label images provided by Brandbank, an FMCG product database. Where this nutritional information was not available, Kantar obtained nutritional values from similar products or an average value for the product type was used.

We used nutrient data provided by Kantar to assess the healthiness of products by categorizing purchases as HFSS or non-HFSS, based on the UK's nutrient profiling model (NPM), which was developed by the Food Standards Agency in 2004/2005 [21]. We chose this classification as it has been used in policies and regulations in the United Kingdom previously [22,23]. An NPM score for each product was calculated by adding points for energy, sugar, sodium, and saturated fat and subtracting points for protein, fiber, fruit, nut, and vegetable content. Information on the energy, sugar, sodium, saturated fat, protein, and fiber content of each purchase was provided by Kantar. Kantar categorizes product markets as high, mixed (medium) or low in fruit, nuts, and vegetables, which we then used to assign individual products with 5 (>80%), 1 (>40% and ≤80%), or 0 (≤40%) points for fruit, nut, and vegetable content. The higher the final NPM score, the less healthy the product. We used the recommended cut-offs of ≥4 for foods and ≥1 for drinks to classify a product as HFSS. We additionally split purchases into 35 food groups based on (1) product market and submarket classifications provided by Kantar and (2) whether products were healthier (non-HFSS) or less healthy (HFSS). These food groups have been used in other studies and are described in Table S1 in Multimedia Appendix 1 [24].

Sociodemographic Characteristics

Sociodemographic data are collected by Kantar annually. We had access to data on sex, age (years), and occupational social class of the main food shopper, number of adults and children in the household, and household income. Occupational social class was based on the National Readership Survey social grade [25]. We categorized occupational social class as high (AB—higher and intermediate managerial, administrative, and professional occupations), middle (C1C2—supervisory, clerical and junior managerial, administrative, and professional occupations), and low (DE—semiskilled and unskilled manual occupations and unemployed, including retired). Gross annual household income bands were available for 1608 (84.1%) households (<£20,000, £20,000-£29,999, £30,000-£39,999, £40,000-£49,999, and ≥£50,000 [A currency exchange rate of £1=US $1.28 is applicable.]). Height and weight were self-reported by 1544 (80.8%) of the main food shoppers and used to calculate BMI (kg/m^2).

Outcomes

Mode of Purchase

Panelists reported the type of store each purchase was from, including whether purchases were made through the internet. Purchases were determined to be online or in-store purchases according to this information. Click and collect services were included in OGDS, as store type information was based on where purchases were made rather than how they were delivered.

Purchases

Household purchases, stratified by online or in-store, were summed over the 1-year study period. We then divided this by household size and the 52 study weeks to obtain means per person per week for outcomes: energy (kcal), fat (g), saturated fat (g), sugar (g), salt (g), items (packs), and expenditure (£). To account for energy purchased, we also looked at mean weekly outcomes per 100 kcal purchased for fat (g/100 kcal), saturated fat (g/100 kcal), sugar (g/100 kcal), salt (g/100 kcal), and expenditure (£/100 kcal). We further explored the proportion of purchases that were HFSS and the proportion of energy purchased from HFSS products and from each of the 35 food groups.

Statistical Analyses

All items purchased on the same day using the same mode of purchase (online or in-store) were considered as having been purchased in the same shopping occasion. The use of OGDS was highly skewed (Figure S1 in Multimedia Appendix 1). Most households did not use OGDS in 2019, and many households that did use them did so infrequently. Among households that used OGDS, the median use was 5 occasions in 2019. We explored differences in purchases between households that used OGDS above and below the median as we theorized that households with very low OGDS use (eg, 1-2 occasions in a year) would not be representative of typical users.

Using a logistic regression model, we examined the association between sociodemographic characteristics (sex, age group, region, occupational social class, household income, number of adults in the household, and presence of children in the household) and OGDS use (above- or below-median). Using Poisson regression models, we explored the association between OGDS use and total purchases of energy, nutrients (overall and per 100 kcal), items, and expenditure (overall and per 100 kcal). Marginal effects were obtained, and pairwise comparison was used to estimate the difference in mean purchases per person per week. We also examined the association between OGDS use and the proportion of purchases that were HFSS using linear regression models. For households that used both online and in-store purchasing methods (n=665), we compared the mean proportion of energy purchased from HFSS products and the 35 food groups through the 2 shopping methods. To examine the associations between OGDS use and purchases independent of household sociodemographic characteristics, all models were adjusted for sex, age, and occupational social class of the main food shopper, region, household income, and proportion of household members that were children (number of children divided by household size). The logistic regression model used to investigate the association between sociodemographic characteristics and OGDS use was also mutually adjusted for other sociodemographic characteristics. All analyses were conducted in Stata SE 16 (Stata Corp).

Sensitivity Analyses

As the cut-offs for categorizing households by the frequency of OGDS could have been made in several ways, we also explore the differences between regular users (which we defined as households using OGDS ≥26 occasions in 2019) and nonregular users (<26 occasions) to see if our results were sensitive to changes in cut-off. As only 147 households were regular OGDS users, we did not use this as our main analysis due to reduced statistical power. We also ran the analyses comparing households that had any OGDS use (n=668) with those that had no OGDS use (n=1243). As mentioned above, we did not use this as our main analysis as we theorized that households with very low OGDS use would not be representative of typical users. Using a subsample of households where the main food shopper's BMI was known (n=1544), we also reran our analyses, additionally adjusting for BMI—a potential confounder of the relationship between OGDS use and purchasing behavior—to see if this changed our findings. BMI was categorized into 2 groups: not overweight (<25 kg/m^2) and overweight (≥25 kg/m^2).

Results

Study Population and OGDS Use

In total, 1911 households were included in our analyses (Table 1). Most households had a female main food shopper (n=1391, 72.8%), were in the middle social class group (n=1133, 59.3%), had 2 adults (n=1062, 55.6%), and had no children (n=1358, 71.1%). Households varied in their household income and age of the main shopper. Of the 1911 included households, 668 (35%) used OGDS at least once in 2019. Almost all households that used OGDS also shopped in-store (n=665, 99.6%). Of households that used OGDS, the median use was 5 (IQR 1-22) occasions in 2019. When stratified by above- (≥5 occasions) and below- (<5 occasions) median OGDS use, there were 353 (18.5%) households above and 1558 (81.5%) below the median including households that did not use OGDS at all. Without adjustment, female main shoppers, younger main shoppers, households in London, households with higher annual income, and households with children were more likely to have above-median OGDS use.

Table 1. Sociodemographic characteristics of the study sample.

	Total users (N=1911), n (%)	Above-median OGDS[a] users (n=353), n (%)	Below-median OGDS users (n=1558), n (%)	Chi-square (*df*)	*P* value
Sex of the main shopper					.05
Male	520 (27.2)	81 (23)	439 (28.2)	3.98 (1)	
Female	1391 (72.8)	272 (77.1)	1119 (71.8)		
Age of the main shopper (years)					<.001
18-34	244 (12.8)	49 (13.9)	195 (12.5)	31.84 (4)	
35-44	349 (18.3)	82 (23.2)	267 (17.1)		
45-54	501 (26.2)	117 (33.1)	384 (24.7)		
55-64	436 (22.8)	60 (17)	376 (24.1)		
≥65	381 (19.9)	45 (12.8)	336 (21.6)		
Region					.02
London	939 (49.1)	194 (55)	745 (47.8)	5.87 (1)	
North of England	972 (50.9)	159 (45)	813 (52.2)		
Occupational social class					.14
Low	334 (17.5)	64 (18.1)	270 (17.3)	4.00 (2)	
Middle	1133 (59.3)	194 (55)	939 (60.3)		
High	444 (23.2)	95 (26.9)	349 (22.4)		
Household income[b]					<.001
<£20,000	388 (20.3)	65 (18.4)	323 (20.7)	43.11 (5)	
£20,000-£29,999	323 (16.9)	49 (13.9)	274 (17.6)		
£30,000-£39,999	267 (14)	35 (9.9)	232 (14.9)		
£40,000-£49,999	217 (11.4)	57 (16.2)	160 (10.3)		
≥£50,000	413 (21.6)	110 (31.2)	303 (19.5)		
Unknown	303 (15.9)	37 (10.5)	266 (17.1)		
Adults in the household					.14
1	424 (22.2)	67 (19)	357 (22.9)	3.98 (2)	
2	1062 (55.6)	212 (60.1)	850 (54.6)		
≥3	423 (22.2)	74 (21)	351 (22.5)		
Children in the household					.001
Yes	553 (28.9)	128 (36.3)	425 (27.3)	11.29 (1)	
No	1358 (71.1)	225 (63.7)	1133 (72.7)		
BMI (kg/m^2)					.11
Not overweight	601 (31.5)	101 (28.6)	500 (32.1)	4.34 (2)	
Overweight	943 (49.4)	171 (48.4)	772 (49.6)		
Unknown	367 (19.2)	81 (23)	286 (18.4)		

[a] OGDS: online grocery delivery service.

[b] A currency exchange rate of £1=US $1.28 is applicable.

Food and Drink Purchases

Overall, households purchased a median of 9808 (IQR 6988-13,375) kcal and spent £21.40 (IQR £14.20-£31.40) per person per week (Table 2). Most food and drink items were purchased in-store, with 2,891,843 (89.4%) items purchased in-store and 341,956 (10.6%) items purchased online. HFSS purchases accounted for 1,231,882 (38.1%) of total items purchased and households purchased a median of 4.8 (IQR 3.2-6.9) HFSS items per person per week. Without adjustment, households with above-median OGDS use had higher purchases of energy and salt, purchased more items and spent more compared to households with below-median OGDS use. Households with above-median OGDS use also purchased marginally more salt per 100 kcal, spent more per 100 kcal purchased, and purchased marginally less sugar per 100 kcal compared to households with below-median OGDS use in our unadjusted analysis.

Table 2. Unadjusted sample purchases of energy, nutrients, items, and expenditure per person per week for households with above- and below-median online grocery delivery service (OGDS) use.

	Total users (N=1911), median (IQR)	Above-median OGDS users (n=353), median (IQR)	Below-median OGDS users (n=1558), median (IQR)	Chi-square (*df*=1)	*P* value
Energy (kcal)	9808 (6988-13,375)	10,220 (7302-13,801)	9728 (6836-13,332)	3.83	.05
Fat (g)	412.9 (290.9-574.8)	434.3 (306.3-598.0)	408.5 (288.5-571.9)	1.87	.17
Saturated fat (g)	155.0 (105.7-225.8)	161.1 (113.2-230.4)	153.2 (103.9-225.0)	2.20	.14
Sugar (g)	433.9 (294.8-623.6)	437.5 (307.9-623.2)	433.8 (290.5-623.6)	0.00	.94
Salt (g)	31.1 (21.4-44.1)	33.3 (23.8-45.9)	30.5 (21.1-43.3)	4.80	.03
Items (n)	14.7 (10.3-20.5)	16.0 (11.6-21.8)	14.5 (10.1-20.3)	7.21	.01
HFSS[a] items (n)	4.8 (3.2-6.9)	5.0 (3.5-7.1)	4.7 (3.1-6.8)	3.46	.06
Expenditure (£)[b]	21.4 (14.2-31.4)	24.2 (16.9-35.3)	20.6 (13.7-30.6)	18.61	<.001
Fat per 100 kcal	4.25 (3.90-4.60)	4.25 (3.91-4.59)	4.25 (3.90-4.61)	0.03	.87
Saturated fat per 100 kcal	1.62 (1.43-1.81)	1.64 (1.44-1.85)	1.62 (1.43-1.80)	0.18	.67
Sugar per 100 kcal	4.49 (3.83-5.18)	4.36 (3.77-5.11)	4.52 (3.85-5.20)	4.71	.03
Salt per 100 kcal	0.32 (0.27-0.37)	0.32 (0.28-0.37)	0.31 (0.27-0.37)	3.83	.05
Expenditure per 100 kcal	0.22 (0.18-0.27)	0.24 (0.20-0.29)	0.21 (0.17-0.26)	37.00	<.001

[a] HFSS: high in fat, salt, and sugar.

[b] A currency exchange rate of £1=US $1.28 is applicable.

Sociodemographic Correlates of OGDS Use

In our adjusted models, households where the main shopper was female were marginally more likely than those with male shoppers to have above-median OGDS use (odds ratio [OR] 1.38, 95% CI 1.04 to 1.83; Table 3). London households were more likely to have above-median OGDS use than households in the North of England (OR 1.29, 95% CI 1.01-1.65). Households in the highest 2 income groups (£40,000-£49,999 and ≥£50,000) were more likely to have above-median OGDS use than households with an income <£20,000 (OR 1.70, 95% CI 1.09-2.66 and OR 1.56, 95% CI 1.02-2.38, respectively). No difference was observed by age group of main shopper, occupational social class, number of adults, or the presence of children in the household.

Table 3. Sociodemographic differences in online grocery delivery service (OGDS) use (N=1911) using a logistic regression model adjusted for sex and age of the main food shopper, occupational social class, household income, the proportion of household members that were children, and region.

	Above-median OGDS use	
	Odds ratio (95% CI)	Percentage, mean (95% CI)
Sex of the main shopper		
Male	Reference	15.3 (12.2-18.3)
Female	*1.38 (1.04-1.83)*[a]	19.7 (17.6-21.8)
Age of the main shopper (years)		
18-34	Reference	18.9 (13.9-23.8)
35-44	1.15 (0.77-1.74)	21.1 (16.6-25.6)
45-54	1.28 (0.86-1.90)	22.8 (19.1-26.5)
55-64	0.75 (0.48-1.18)	14.9 (11.3-18.5)
≥65	0.65 (0.40-1.06)	13.3 (9.5-17.0)
Region		
North of England	Reference	16.7 (14.3-19.0)
London	*1.29 (1.01-1.65)*	20.3 (17.7-22.9)
Occupational social class		
Low	Reference	22.4 (17.5-27.3)
Middle	0.72 (0.51-1.01)	17.4 (15.2-19.6)
High	0.78 (0.52-1.18)	18.6 (15.0-22.2)
Household income[b]		
<£20,000	Reference	17.5 (13.3-21.7)
£20,000-£29,999	0.88 (0.58-1.34)	15.7 (11.7-19.8)
£30,000-£39,999	0.72 (0.45-1.15)	13.3 (9.2-17.4)
£40,000-£49,999	*1.70 (1.09-2.66)*	26.3 (20.4-32.1)
≥£50,000	*1.56 (1.02-2.38)*	24.7 (20.1-29.2)
Unknown	0.66 (0.42-1.05)	12.4 (8.7-16.2)
Adults in household		
1	Reference	18.1 (14.0-22.2)
2	1.12 (0.80-1.57)	19.8 (17.4-22.2)
≥3	0.84 (0.56-1.25)	15.7 (12.4-19.1)
Children in household		
No	Reference	18.1 (15.9-20.3)
Yes	1.08 (0.80-1.46)	19.2 (15.8-22.7)

[a] Italic formatting indicates significance at 95% CI.

[b] A currency exchange rate of £1=US $1.28 is applicable.

Association Between OGDS Use and Total Food and Drink Purchases

After adjustment for sociodemographic and household characteristics, households with above-median OGDS use had higher mean purchases of energy (1461 kcal, 95% CI 1448-1474 kcal), fat (64.8 g, 95% CI 62.1-67.6 g), saturated fat (28.5 g, 95% CI 26.8-30.2 g), sugar (63.9 g, 95% CI 61.1-66.7 g), salt (6.1 g, 95% CI 5.4-6.9 g), and items (3 items, 95% CI 2.5-3.5 items) per person per week (Table 4). These households also spent more on groceries per person per week (mean £6.30, 95% CI £5.70-£7.00) than those with below-median OGDS use. However, there was no difference in nutrients purchased per 100 kcal between the 2 groups. When food products and drink products were separated, food accounted for most of the differences observed, although the significant associations were the same for both food and drink (Table S2 in Multimedia Appendix 1).

Table 4. Incidence rate ratio (95% CI) and predicted difference in adjusted weekly mean (95% CI) purchases of energy, nutrients, items per person, and expenditure on groceries per person between households with above- and below-median online grocery delivery service (OGDS) use using a Poisson model adjusted for sex and age of the main food shopper, occupational social class, household income, the proportion of household members that were children, and region.

	Incidence rate ratio (95% CI)	Predicted outcome for above-median OGDS use (n=353), mean (95% CI)	Predicted outcome for below-median OGDS use (n=1558), mean (95% CI)	Predicted difference, mean (95% CI)
Energy (kcal)	1.14 (1.14 to 1.14)	11,927 (11,916 to 11,939)	10,466 (10,461 to 10,471)	*1461* (*1448* to *1474*)[a]
Fat (g)	1.14 (1.14 to 1.15)	512.3 (509.8 to 514.8)	447.4 (446.4 to 448.5)	*64.8* (*62.1* to *67.6*)
Saturated fat (g)	1.17 (1.16 to 1.18)	200.4 (198.8 to 201.9)	171.9 (171.2 to 172.5)	*28.5* (*26.8* to *30.2*)
Sugar (g)	1.13 (1.13 to 1.14)	544.6 (542.0 to 547.2)	480.7 (479.6 to 481.8)	*63.9* (*61.1* to *66.7*)
Salt (g)	1.18 (1.16 to 1.20)	40.3 (39.6 to 41.0)	34.2 (33.9 to 34.5)	*6.1* (*5.4* to *6.9*)
Items (n)	1.19 (1.16 to 1.22)	18.8 (18.4 to 19.3)	15.8 (15.6 to 16.0)	*3.0* (*2.5* to *3.5*)
Expenditure (£)[b]	1.27 (1.24 to 1.30)	29.9 (29.3 to 30.5)	23.6 (23.4 to 23.9)	*6.3* (*5.7* to *7.0*)
Fat per 100 kcal	1.00 (0.94 to 1.06)	4.3 (4.2 to 4.5)	4.3 (4.2 to 4.4)	–0.0 (–0.3 to 0.2)
Saturated fat per 100 kcal	1.02 (0.93 to 1.12)	1.7 (1.5 to 1.8)	1.6 (1.6 to 1.7)	0.0 (–0.1 to 0.2)
Sugar per 100 kcal	0.99 (0.93 to 1.04)	4.5 (4.3 to 4.8)	4.6 (4.5 to 4.7)	–0.1 (–0.3 to 0.2)
Salt per 100 kcal	1.03 (0.84 to 1.27)	0.3 (0.3 to 0.4)	0.3 (0.3 to 0.4)	0.0 (–0.1 to 0.1)
Expenditure per 100 kcal	1.13 (0.89 to 1.43)	0.3 (0.2 to 0.3)	0.2 (0.2 to 0.3)	0.0 (–0.0 to 0.1)

[a] Italic formatting indicates significance at 95% CI.

[b] A currency exchange rate of £1=US $1.28 is applicable.

Association Between OGDS Use and Purchases of HFSS Products

There were no differences in HFSS purchases between households that had above- and below-median OGDS use in our adjusted models (Table 5). At a more disaggregated level, households with above-median OGDS use purchased more or the same amount as households that had below-median use across healthier and less healthy food groups (Table S3 in Multimedia Appendix 1).

Table 5. Coefficient (95% CI) and predicted difference in adjusted weekly mean (95% CI) energy; nutrients; items from high in fat, salt, and sugar purchases as a proportion of all grocery food and drink purchases; and expenditure on products high in fat, salt, and sugar as a proportion of all food and drink grocery expenditure between households with above- and below-median online grocery delivery service (OGDS) use using a linear regression model adjusted for sex and age of the main food shopper, occupational social class, household income, proportion of household members that were children, and region.

	Coefficient (95% CI)	Predicted outcome for above-median OGDS use (n=353), mean (95% CI)	Predicted outcome for below-median OGDS use (n=1558), mean (95% CI)	Predicted difference, mean (95% CI)
Energy	0.00 (–0.01 to 0.01)	52.7 (51.8 to 53.7)	52.5 (52.0 to 52.9)	0.2 (–0.8 to 1.3)
Fat	0.01 (–0.00 to 0.02)	71.3 (70.3 to 72.2)	70.6 (70.2 to 71.1)	0.6 (–0.4 to 1.6)
Saturated fat	0.01 (–0.00 to 0.02)	77.0 (76.1 to 77.9)	76.0 (75.5 to 76.4)	1.0 (–0.0 to 2.0)
Sugar	–0.00 (–0.02 to 0.01)	56.5 (55.1 to 57.9)	56.7 (56.0 to 57.3)	–0.2 (–1.8 to 1.4)
Salt	0.01 (–0.01 to 0.02)	64.4 (63.3 to 65.5)	63.7 (63.2 to 64.2)	0.7 (–0.5 to 1.9)
Items	–0.00 (–0.01 to 0.01)	37.6 (36.7 to 38.6)	37.7 (37.2 to 38.1)	–0.0 (–1.1 to 1.0)
Expenditure	–0.00 (–0.01 to 0.01)	39.5 (38.5 to 40.5)	39.7 (39.2 to 40.2)	–0.2 (–1.4 to 0.9)

Association Between Mode of Shopping and Purchases Among OGDS Users

Among households that used both online and in-store shopping methods (n=665), the mean proportion of energy purchased from HFSS products was –10.1% (95% CI –12% to –8.1%) lower in online compared to in-store shopping occasions (43.3%, 95% CI 41.4%-45.2% vs 53.3%, 95% CI 52.5%-54.1%) in the adjusted model. When stratified by food group, online purchases had proportionately more energy from vegetables (1%, 95% CI 0.2%-1.8%), healthy nonmilk-based drinks (1.6%, 95% CI 0.7%-2.4%), and alcohol (1.2%, 95% CI 0.4%-2.1%) and proportionally less energy from some HFSS food groups, including puddings and biscuits (–3.3%, 95% CI –4.1% to –2.5%), chocolate and confectionery (–1.5%, 95% CI –2.2% to –0.7%), and savory snacks (–0.8, 95% CI –1.3% to –0.2%; Table S4 in Multimedia Appendix 1). Healthy nonmilk-based drinks were nonalcoholic and not milk-based (eg, milkshakes and coffee drinks). Examples of nonmilk-based drinks included carbonated drinks and juice drinks. Healthy refers to non-HFSS according to the UK nutrient profiling model. The proportion of energy was also lower from carbohydrate food groups (healthy bread: –3.1%, 95% CI –3.7% to –2.5%; less healthy bread: –0.7, 95% CI –0.8% to –0.5%; morning goods: –0.6%, 95% CI –0.8% to –0.4%), dairy products (reduced fat milk: –0.9, 95% CI –1.2% to –0.6%; high fat milk: –0.4, 95% CI –0.6% to –0.2%; less healthy cheese: –1.2, 95% CI –1.7% to –0.7%, healthy cheese: –0.01, 95% CI –0.02% to 0%), protein (–0.6%, 95% CI –1.1% to 0%), and ready meals (healthy: –1.2%, 95% CI –1.6% to –0.8%; less healthy: –0.7, 95% CI –1.2% to –0.2) for online shopping occasions.

Sensitivity Analyses

The difference in total grocery purchases between regular OGDS users and nonregular users was larger compared to the main analysis (2208 kcal, 95% CI 2188-2228 kcal vs 1461 kcal, 95% CI 1448-1474 kcal; Table S5 in Multimedia Appendix 1). There were no qualitative differences in the proportion of HFSS purchases between regular OGDS users versus nonregular users and above- versus below-median users (Table S6 in Multimedia Appendix 1). The association between households with any versus no OGDS use and total purchases was in the same direction as the main analysis, but the magnitude of difference was smaller (756 kcal, 95% CI 746-766 kcal vs 1461 kcal, 95% CI 1448-1474 kcal; Table S7 in Multimedia Appendix 1). Although the point estimates for HFSS purchases were similar for any versus no OGDS use and above- versus below-median use, a marginally higher proportion of energy, fat, and saturated fat from HFSS products was observed among households that had any OGDS use compared to households with none (Table S8 in Multimedia Appendix 1). Adjusting for BMI did not qualitatively change the association between OGDS use and total food and drink purchases (Table S9 in Multimedia Appendix 1) or OGDS use and HFSS purchases (Table S10 in Multimedia Appendix 1).

Discussion

Summary of Main Findings

Overall, in our sample, 35% (n=668) of London and the North of England households used OGDS at least once, with OGDS users having a median use of 5 occasions in 2019. Almost all households that used OGDS also purchased foods in-store. Higher-income households, households with a female main food shopper, and households in London versus the North of England had a higher likelihood of above-median OGDS use. Households with above-median OGDS use had higher mean purchases of groceries per person per week, in terms of energy (1461 kcal), number of food and drink items (3 items), and expenditure (£6.30). Nutrients purchased per 100 kcal, expenditure on purchases per 100 kcal, and the proportion of total purchases that were HFSS were similar in both groups. However, households that used both shopping modes purchased a 10.1% (95% CI 8.1%-12%) lower proportion of energy from HFSS products from their online compared with in-store purchases.

Interpretation of Findings

Almost no households in our sample exclusively used OGDS for their grocery shopping. OGDS use therefore appears to complement, rather than substitute, in-store grocery shopping. The extra energy purchased among households with above-median OGDS use, compared with below-median OGDS use, could lead to overconsumption or food waste, which has negative consequences for population health and environmental health. Alternatively, more energy purchased for in-home consumption could be because these households eat out less. If this is the case, additional grocery purchases may be beneficial for population diet as consuming out-of-home food is associated with weight gain and higher intakes of energy, fat, salt, and sugar [26,27].

No significant differences in the healthiness of purchasing between households according to their OGDS use were observed, with households with above-median OGDS use purchasing greater amounts of both HFSS and non-HFSS products. However, there was a difference in the healthiness of purchasing within households, with their online purchases having fewer HFSS products than their in-store purchases. This means that households made healthier choices online or preferentially purchased their healthier products online. As our analysis was a comparison within households, the differences observed here cannot be due to household characteristics and are likely due to factors that influence the shopping experience (eg, vividness of products, time lag between shopping and receiving items, exposure to promotions and front-of-pack labeling, preference for purchasing some items online or in-store, and ability to browse products online vs in-store) and, in turn, purchasing behavior [11,12,28]. Online shopping environments have been found to be more supportive of healthier diets through fewer advertisements and promotions and a lack of product placement to encourage HFSS purchases. A healthier food environment in an online setting, compared to a physical setting, is supported by studies from the United States, which found OGDS to be associated with purchasing healthier food [9,11,12,17]. The higher proportion of energy from fruit and vegetables in households' online shops suggests that households were not deterred from purchasing perishable goods without being able to choose products personally, which was a concern highlighted in previous research [10]. Differences in the online food environment, compared to the physical food environment, may have implications for population health, especially as OGDS use prevalence increases. Therefore, further research is needed to understand why households choose OGDS and how it relates to in-store purchasing and out-of-home purchasing to give us a better understanding of total food purchases. With the COVID-19 pandemic having encouraged greater use of OGDS [4,5], future studies should examine short- and long-term changes in food purchasing behaviors and their associations with population diet.

While we have some understanding of, and are trying to regulate, aspects of the in-store food environment [29], less is known about how online platforms influence purchases and whether and how these should (or should not) be regulated. Some studies have shown positive results for interventions that manipulate the online food environment to encourage healthier purchasing [28,30,31]. The online food environment could be used to increase access to fresh food, improve access to nutritional information, and reduce exposure to HFSS advertising [10,19]. While there are fewer promotions online compared to in-store in the United Kingdom (24% of products vs 32% of products), front-of-pack labeling was only found for 42% of products online compared to 74% in-store [16]. Regulations could be helpful in harnessing the elements of OGDSs that may improve population health while limiting the elements that may have negative impacts.

Strengths and Limitations

While online platforms for grocery purchasing have been available for many years, little is known about their association with food and drink purchasing behavior in the United Kingdom. Our findings contribute to the evidence base on OGDS use and food purchasing behavior. Purchase data are a good indicator of food and drink consumption [32,33]. However, they do not tell us about the intrahousehold distribution of purchases, which we assumed to be equal across all household members, including children. This, of course, will not have been the case and provides only an estimate, which likely underestimates purchased energy and nutrients per adult for households with children. Purchase data also do not tell us about food waste. Having data at the item level with the day on which purchases were made and whether this was online or in-store allowed us to establish OGDS use by number of occasions. However, it is unknown whether households may report online and in-store purchases differently. Households may also have forgotten to report purchases or chosen to not report less healthy purchases due to social desirability. The lower average weekly household grocery food expenditure in this analysis (£21.40) compared to the UK's Living Costs and Food Survey (£32.12) may be explained by the inclusion of only 2 UK regions in our study and the unweighted nature of our data, which limits its accuracy in reflecting population-level averages [34]. We were able to compare online and in-store purchases for households that used both modes of shopping, which removed potential confounding from household characteristics. However, we were only able to assess purchases brought into the home and could not consider out-of-home purchases, so we were unable to capture overall household food purchasing. We were also unable to equivalize household income as we did not have information on the age of the children in the household. However, all models were adjusted for the proportion of household members that were children. Errors in the nutrient data, imputed values where data were missing, and the crude estimation of fruit, nut, and vegetable content may have affected the accuracy of our categorization of products as HFSS or non-HFSS. We were unable to explore how online shopping interacts with other forms of food acquisition.

Conclusions

In this study, OGDS use was greater among higher-income and London households. Households that had above-median OGDS use purchased a mean of 1461 kcal more energy per person per week through their grocery purchases. The differences in grocery purchases between households with above- and below-median OGDS use could have positive or negative consequences. The extra energy purchased through groceries among households with above-median OGDS use could lead to overconsumption or food waste, a problem for population health and the environment. Alternatively, more energy purchased for in-home consumption could plausibly replace out-of-home consumption, which may be beneficial to population diet, as out-of-home food tends to be less healthy. Households purchased fewer HFSS purchases when shopping online compared to in-store, which may be due to differences in the shopping environment or experience, such as fewer promotions and advertisements when shopping online or not having to transport and carry purchases home. As higher-income households used OGDS more frequently, the implications of this sociodemographic patterning on dietary inequalities need to be explored. Further research is needed to investigate the relationship between OGDS use and the healthiness of purchasing in different population subgroups. With the online food environment becoming an increasingly important retail channel for household food purchasing, it should be an important consideration when designing food system policies and interventions that aim to improve population diet and reduce dietary inequalities.

Acknowledgments

This study was funded by the National Institute for Health and Care Research (NIHR) School for Public Health Research (grant PD-SPH-2015). The National Institute for Health and Care Research School for Public Health Research is a partnership between the Universities of Sheffield, Bristol, Cambridge, Imperial, University College London; The London School for Hygiene and Tropical Medicine; LiLaC—a collaboration between the Universities of Liverpool and Lancaster; and Fuse—The Centre for Translational Research in Public Health, a collaboration between Newcastle, Durham, Northumbria, Sunderland and Teesside Universities. AY, TB, MW, and SC are also funded by the Biotechnology and Biological Science Research Council through the Mandala project (grant BB/V004832/1). JA is a coinvestigator on the Mandala project and has funding from the NIHR Policy Research Programme and Public Health Research Programme, the Department of Health and Social Care via the Public Health-Policy Research Unit, and Economic and Social Research Council for SALIENT. SC is also funded by Health Data Research UK. Health Data Research UK is an initiative funded by the UK Research and Innovation, the Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities. JA, TB, and MW are supported by the Medical Research Council (grant MC_UU_00006/7). FdV is partly funded by the NIHR Applied Research Collaboration West (NIHR ARC West) at University Hospitals Bristol National Health Service Foundation Trust. LC is funded by a UK Medical Research Council fellowship (grant MR/P021999/1). The views expressed are those of the authors and do not necessarily represent those of any of the above-mentioned funders. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the terms of our data agreement with Kantar. Data are available to purchase by contacting Kantar directly.

Authors' Contributions

All authors developed the research questions. AY, CL, LC, and SC formally analyzed the data. AY and SC wrote the manuscript. All authors critically reviewed and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Files.
[DOCX File , 150 KB - publichealth_v9i1e41540_app1.docx]

References

1. Keeble M, Adams J, Sacks G, Vanderlee L, White CM, Hammond D, et al. Use of online food delivery services to order food prepared away-from-home and associated sociodemographic characteristics: a cross-sectional, multi-country analysis. Int J Environ Res Public Health 2020;17(14):5190 [FREE Full text] [doi: 10.3390/ijerph17145190] [Medline: 32709148]
2. Poelman MP, Thornton L, Zenk SN. A cross-sectional comparison of meal delivery options in three international cities. Eur J Clin Nutr 2020;74(10):1465-1473. [doi: 10.1038/s41430-020-0630-7] [Medline: 32332863]
3. Bates S, Reeve B, Trevena H. A narrative review of online food delivery in Australia: challenges and opportunities for public health nutrition policy. Public Health Nutr 2023;26(1):262-272 [FREE Full text] [doi: 10.1017/S1368980020000701] [Medline: 32515719]
4. Tighe D. Online grocery shopping in the United Kingdom (UK)—statistics and facts. Statista. URL: https://www.statista.com/topics/3144/online-grocery-shopping-in-the-united-kingdom/#dossierKeyfigures [accessed 2022-01-06]
5. Thompson C, Hamilton L, Dickinson A, Fallaize R, Mathie E, Rogers S, et al. Changes to household food shopping practices during the COVID-19 restrictions: evidence from the East of England. Health Place 2022;78:102906 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102906] [Medline: 36108358]
6. Wang XC, Kim W, Holguín-Veras J, Schmid J. Adoption of delivery services in light of the COVID pandemic: who and how long? Transp Res Part A Policy Pract 2021;154:270-286 [FREE Full text] [doi: 10.1016/j.tra.2021.10.012] [Medline: 34744329]
7. Pitts SBJ, Ng SW, Blitstein JL, Gustafson A, Niculescu M. Online grocery shopping: promise and pitfalls for healthier food and beverage purchases. Public Health Nutr 2018;21(18):3360-3376 [FREE Full text] [doi: 10.1017/S1368980018002409] [Medline: 30338751]
8. Khandpur N, Zatz LY, Bleich SN, Taillie LS, Orr JA, Rimm EB, et al. Supermarkets in cyberspace: a conceptual framework to capture the influence of online food retail environments on consumer behavior. Int J Environ Res Public Health 2020;17(22):8639 [FREE Full text] [doi: 10.3390/ijerph17228639] [Medline: 33233798]
9. Harris‐Lagoudakis K. Online shopping and the healthfulness of grocery purchases. Am J Agri Econ 2021;104(3):1050-1076 [FREE Full text] [doi: 10.1111/ajae.12262]
10. Fernandez MA, Raine KD. Digital food retail: public health opportunities. Nutrients 2021;13(11):3789 [FREE Full text] [doi: 10.3390/nu13113789] [Medline: 34836044]
11. Huyghe E, Verstraeten J, Geuens M, Van Kerckhove A. Clicks as a healthy alternative to bricks: how online grocery shopping reduces vice purchases. J Mark Res 2017;54(1):61-74. [doi: 10.1509/jmr.14.0490]
12. Milkman KL, Rogers T, Bazerman MH. I'll have the ice cream soon and the vegetables later: a study of online grocery purchases and order lead time. Mark Lett 2009;21(1):17-35. [doi: 10.1007/s11002-009-9087-0]
13. Brandt EJ, Silvestri DM, Mande JR, Holland ML, Ross JS. Availability of grocery delivery to food deserts in states participating in the online purchase pilot. JAMA Netw Open 2019;2(12):e1916444 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.16444] [Medline: 31790562]
14. Brar K, Minaker LM. Geographic reach and nutritional quality of foods available from mobile online food delivery service applications: novel opportunities for retail food environment surveillance. BMC Public Health 2021;21(1):458 [FREE Full text] [doi: 10.1186/s12889-021-10489-2] [Medline: 33676458]
15. Newing A, Hood N, Videira F, Lewis J. 'Sorry we do not deliver to your area': geographical inequalities in online groceries provision. Int Rev Retail Distrib Consum Res 2021;32(1):80-99 [FREE Full text] [doi: 10.1080/09593969.2021.2017321]
16. Bhatnagar P, Scarborough P, Kaur A, Dikmen D, Adhikari V, Harrington R. Are food and drink available in online and physical supermarkets the same? A comparison of product availability, price, price promotions and nutritional information. Public Health Nutr 2021;24(5):819-825 [FREE Full text] [doi: 10.1017/S1368980020004346] [Medline: 33109282]

17. Saphores JD, Xu L. E-shopping changes and the state of E-grocery shopping in the US—evidence from national travel and time use surveys. Res Transp Econ 2021;87:100864 [FREE Full text] [doi: 10.1016/j.retrec.2020.100864]
18. Locher I, Waselewski M, Sonneville K, Resnicow K, Chang T. Grocery delivery of healthy foods to pregnant young women with low incomes: feasibility and acceptability mixed methods study. JMIR Form Res 2020;4(12):e21602 [FREE Full text] [doi: 10.2196/21602] [Medline: 33361055]
19. Wyse R, Jackson JK, Delaney T, Grady A, Stacey F, Wolfenden L, et al. The effectiveness of interventions delivered using digital food environments to encourage healthy food choices: a systematic review and meta-analysis. Nutrients 2021;13(7):2255 [FREE Full text] [doi: 10.3390/nu13072255] [Medline: 34208869]
20. Yau A, Berger N, Law C, Cornelsen L, Greener R, Adams J, et al. Changes in household food and drink purchases following restrictions on the advertisement of high fat, salt, and sugar products across the transport for London network: a controlled interrupted time series analysis. PLoS Med 2022;19(2):e1003915 [FREE Full text] [doi: 10.1371/journal.pmed.1003915] [Medline: 35176022]
21. The nutrient profiling model. UK Department of Health and Social Care. 2011. URL: https://www.gov.uk/government/publications/the-nutrient-profiling-model [accessed 2022-01-05]
22. Regulation of advertising for less healthy food and drink: Implementation of new statutory restrictions. Ofcom. 2023 Jul. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0016/264013/Regulation-of-advertising-for-less-healthy-food-and-drink.pdf [accessed 2023-12-11]
23. TfL ad policy: Approval guidance food and non-alcoholic drink advertising. Transport for London. 2019. URL: http://content.tfl.gov.uk/policy-guidance-food-and-drink-advertising.pdf [accessed 2022-01-05]
24. Berger N, Cummins S, Smith RD, Cornelsen L. Recent trends in energy and nutrient content of take-home food and beverage purchases in Great Britain: an analysis of 225 million food and beverage purchases over 6 years. BMJ Nutr Prev Health 2019;2(2):63-71 [FREE Full text] [doi: 10.1136/bmjnph-2019-000036] [Medline: 33235959]
25. Social grade. National Readership Survey. 2018. URL: http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/ [accessed 2020-09-22]
26. Lachat C, Nago E, Verstraeten R, Roberfroid D, Van Camp J, Kolsteren P. Eating out of home and its association with dietary intake: a systematic review of the evidence. Obes Rev 2012;13(4):329-346. [doi: 10.1111/j.1467-789X.2011.00953.x] [Medline: 22106948]
27. Gesteiro E, García-Carro A, Aparicio-Ugarriza R, González-Gross M. Eating out of home: influence on nutrition, health, and policies: a scoping review. Nutrients 2022;14(6):1265 [FREE Full text] [doi: 10.3390/nu14061265] [Medline: 35334920]
28. Gorin AA, Raynor HA, Niemeier HM, Wing RR. Home grocery delivery improves the household food environments of behavioral weight loss participants: results of an 8-week pilot study. Int J Behav Nutr Phys Act 2007;4:58 [FREE Full text] [doi: 10.1186/1479-5868-4-58] [Medline: 18001469]
29. UK Statutory Instruments 2021 No. 1368. The food (promotion and placement) (England) regulations 2021. 2021. URL: https://www.legislation.gov.uk/uksi/2021/1368 [accessed 2023-08-30]
30. Koutoukidis DA, Jebb SA, Ordóñez-Mena JM, Noreik M, Tsiountsioura M, Kennedy S, et al. Prominent positioning and food swaps are effective interventions to reduce the saturated fat content of the shopping basket in an experimental online supermarket: a randomized controlled trial. Int J Behav Nutr Phys Act 2019;16(1):50 [FREE Full text] [doi: 10.1186/s12966-019-0810-9] [Medline: 31174547]
31. Shin S, van Dam RM, Finkelstein EA. The effect of dynamic food labels with real-time feedback on diet quality: results from a randomized controlled trial. Nutrients 2020;12(7):2158 [FREE Full text] [doi: 10.3390/nu12072158] [Medline: 32698540]
32. Griffith R, O'Connell M, Smith K. Measuring the quality of people's diets: a comparison of intake and purchase data. The Institute for Fiscal Studies. 2012. URL: http://www.ifs.org.uk/publications/6531 [accessed 2022-01-06]
33. Appelhans BM, French SA, Tangney CC, Powell LM, Wang Y. To what extent do food purchases reflect shoppers' diet quality and nutrient intake? Int J Behav Nutr Phys Act 2017;14(1):46 [FREE Full text] [doi: 10.1186/s12966-017-0502-2] [Medline: 28399887]
34. Family food 2018/19. UK Department for Environment Food and Rural Affairs. 2020. URL: https://www.gov.uk/government/statistics/family-food-201819/family-food-201819#household-spending-on-food [accessed 2022-01-06]

Abbreviations

FMCG: fast-moving consumer good
HFSS: high in fat, salt, and sugar
NPM: nutrient profiling model
OGDS: online grocery delivery service
OR: odds ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 05.08.22; peer-reviewed by S Shin, J Meinilä, R Abd Jalil; comments to author 07.12.22; revised version received 14.03.23; accepted 29.08.23; published 19.12.23.

Please cite as:
Yau A, Law C, Cornelsen L, Adams J, Boyland E, Burgoine T, de Vocht F, White M, Cummins S
Association Between Household Online Grocery Delivery Service Use and Food and Drink Purchase Behavior in England: Cross-Sectional Analysis
JMIR Public Health Surveill 2023;9:e41540
URL: https://publichealth.jmir.org/2023/1/e41540
doi:10.2196/41540
PMID:38113090

Protocol

Long-Acting Injectable Cabotegravir for HIV Preexposure Prophylaxis Among Sexual and Gender Minorities: Protocol for an Implementation Study

Beatriz Grinsztejn[1], MD, PhD; Thiago Silva Torres[1], PhD; Brenda Hoagland[1], MD, PhD; Emilia Moreira Jalil[1], MD, PhD; Ronaldo Ismerio Moreira[1], PhD; Gabrielle O'Malley[2], PhD; Starley B Shade[3], PhD; Marcos R Benedetti[1], MSc; Julio Moreira[4], BA; Keila Simpson[5], BA; Maria Cristina Pimenta[6], PhD; Valdiléa Gonçalves Veloso[1], MD, PhD; The ImPrEP CAB-Brasil Study Team[7]

[1]Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

[2]Department of Global Health, Schools of Medicine and Public Health, University of Washington, Seattle, WA, United States

[3]Institute for Global Health Sciences, Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, United States

[4]Grupo Arco-Íris de Cidadania LGBT, Rio de Janeiro, Brazil

[5]Associação Nacional de Travestis e Transexuais, Salvador, Brazil

[6]Brazilian Ministry of Health, Brasilia, Brazil

[7]See Acknowledgments

Corresponding Author:
Beatriz Grinsztejn, MD, PhD
Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz
Av. Brasil 4365 Manguinhos
Rio de Janeiro, 21040360
Brazil
Phone: 55 21 38659128
Email: gbeatriz@ini.fiocruz.br

Abstract

Background: Long-acting injectable cabotegravir (CAB-LA) for preexposure prophylaxis (PrEP) has proven efficacious in randomized controlled trials. Further research is critical to evaluate its effectiveness in real-world settings and identify effective implementation approaches, especially among young sexual and gender minorities (SGMs).

Objective: ImPrEP CAB Brasil is an implementation study aiming to generate critical evidence on the feasibility, acceptability, and effectiveness of incorporating CAB-LA into the existing public health oral PrEP services in 6 Brazilian cities. It will also evaluate a mobile health (mHealth) education and decision support tool, digital injection appointment reminders, and the facilitators of and barriers to integrating CAB-LA into the existing services.

Methods: This type-2 hybrid implementation-effectiveness study includes formative work, qualitative assessments, and clinical steps 1 to 4. For formative work, we will use participatory design methods to develop an initial CAB-LA implementation package and process mapping at each site to facilitate optimal client flow. SGMs aged 18 to 30 years arriving at a study clinic interested in PrEP (naive) will be invited for step 1. Individuals who tested HIV negative will receive mHealth intervention and standard of care (SOC) counseling or SOC for PrEP choice (oral or CAB-LA). Participants interested in CAB-LA will be invited for step 2, and those with undetectable HIV viral load will receive same-day CAB-LA injection and will be randomized to receive digital appointment reminders or SOC. Clinical appointments and CAB-LA injection are scheduled after 1 month and every 2 months thereafter (25-month follow-up). Participants will be invited to a 1-year follow-up to step 3 if they decide to change to oral PrEP or discontinue CAB-LA and to step 4 if diagnosed with HIV during the study. Outcomes of interest include PrEP acceptability, choice, effectiveness, implementation, and feasibility. HIV incidence in the CAB-LA cohort (n=1200) will be compared with that in a similar oral PrEP cohort from the public health system. The effectiveness of the mHealth and digital interventions will be assessed using interrupted time series analysis and logistic mixed models, respectively.

Results: During the third and fourth quarters of 2022, we obtained regulatory approvals; programmed data entry and management systems; trained sites; and performed community consultancy and formative work. Study enrollment is programmed for the second quarter of 2023.

Conclusions: ImPrEP CAB Brasil is the first study to evaluate CAB-LA PrEP implementation in Latin America, one of the regions where PrEP scale-up is most needed. This study will be fundamental to designing programmatic strategies for implementing and scaling up feasible, equitable, cost-effective, sustainable, and comprehensive alternatives for PrEP programs. It will also contribute to maximizing the impact of a public health approach to reducing HIV incidence among SGMs in Brazil and other countries in the Global South.

Trial Registration: Clinicaltrials.gov NCT05515770; https://clinicaltrials.gov/ct2/show/NCT05515770

International Registered Report Identifier (IRRID): PRR1-10.2196/44961

(JMIR Public Health Surveill 2023;9:e44961) doi:10.2196/44961

KEYWORDS

sexual and gender minorities; young; Brazil; HIV prevention; injectable preexposure prophylaxis; injectable PrEP; cabotegravir

Introduction

Background

In 2021, there were approximately 960,000 people living with HIV in Brazil [1]. In recent studies, the HIV prevalence rate was estimated to reach approximately 23% among Brazilian gay men, bisexual men, and other men who have sex with men (MSM) [2] and surpass 30% among *travesti* and transgender women (TGW) in some cities [3,4]. Young MSM aged ≤30 years are among the few populations with an increasing number of new HIV infections in Brazil, whereas HIV incidence has stabilized in the overall population [2]. Since December 2017, oral tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC) HIV preexposure prophylaxis (PrEP) have been freely available in Brazil through the national public health system (*Sistema Único de Saúde* [SUS]) [5]. As of October 2022, there were 566 health facilities across all Brazilian states providing oral PrEP to approximately 46,886 individuals [6]. Although 85% of current PrEP users in Brazil are MSM, PrEP use remains well below the initial projections [5].

Tenofovir-based oral PrEP has been effective in dramatically reducing population-level HIV incidence in multiple regions [7-11]. However, to realize the protective benefits of PrEP, individuals need to both initiate PrEP and remain on PrEP during periods of HIV vulnerability [12,13]. Among the individuals who initiated PrEP in the Brazilian SUS, discontinuation was common among MSM (approximately 40%) and TGW (>50%) [6]. PrEP Brasil, a daily oral PrEP demonstration study, found that although 74% of all participants had protective concentrations of tenofovir diphosphate at week 48, young participants and TGW participants had decreasing rates of protective drug concentrations throughout the study [14]. In addition, data from the ImPrEP study, which aimed to evaluate the feasibility of same-day PrEP initiation for MSM and TGW within the public health contexts of Brazil, Mexico, and Peru, identified lower PrEP adherence among TGW and individuals aged ≤30 years [15,16].

Owing to user adherence challenges, PrEP has not decreased HIV incidence among some population, particularly among populations with higher susceptibility (youth, racial and ethnic minorities, and TGW) [17-20]. Although HIV incidence in the ImPrEP study population was relatively low (0.85/100 person-years), it was higher among TGW, participants from Peru, those aged 18 to 30 years, Black or mixed-race participants, and those who were nonadherent to daily oral PrEP [15,16]. These disproportionately worse outcomes reinforce the potential of PrEP agents that do not require daily or planned oral dosing to increase their effective use among the most susceptible populations.

In multiple studies, reasons for discontinuing or never initiating daily oral PrEP include low HIV perceived risk, concerns about medication side effects, desire not to take a daily pill, stigma, competing life events, and challenges accessing the medication [13,21-23]. Long-acting PrEP agents may potentially address some of these concerns by providing consistent protection through infrequent, discrete administration, avoiding the need for daily adherence to an oral regimen and the related need for disclosure to sex partners and acquaintances [24].

Long-Acting Injectable Cabotegravir

Long-acting injectable cabotegravir (CAB-LA) is an integrase strand-transfer inhibitor, and its effectiveness for PrEP has been demonstrated in phase 2b/3 studies (HIV Prevention Trials Network [HPTN] 083 and HPTN 084) [25,26]. HPTN 083 was a double-blind double-dummy randomized controlled trial that assessed the effectiveness of CAB-LA for PrEP compared with oral tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC) PrEP among 4566 cisgender MSM and TGW under high susceptibility to HIV. The blinded phase of the study showed the superiority of CAB-LA, with a 66% reduction in new infections in the cabotegravir arm compared with the active TDF and FTC arm. Similarly, HPTN 084 observed impressive reductions in HIV infections in the CAB-LA arm compared with the TDF and FTC arm among cisgender women in Africa, with an 89% reduction in new infections in the cabotegravir arm when compared with the active TDF and FTC arm. Data from both studies led to the regulatory approval of CAB-LA for HIV prevention in the United States, Australia, South Africa, Malawi, and Zimbabwe. Approval is currently being sought in numerous countries, including Brazil. In July 2022, a recommendation for CAB-LA PrEP was issued by the World Health Organization (WHO) [27].

Although it is reassuring that CAB-LA has demonstrated comparable effectiveness among youth, Black MSM, and TGW in large clinical trials, the advantages of bypassing daily adherence behavior may not be realized if individuals cannot access injectable PrEP or do not return for repeat injections. Implementation studies are needed to provide crucial information about how interventions that are proven efficacious under randomized clinical trial conditions can be optimized for delivery within real-world service settings. The ImPrEP CAB Brasil implementation study aims to generate critical evidence to inform national policies and program implementers about how injectable CAB-LA PrEP can be provided to MSM, nonbinary persons, and transgender persons within public health PrEP services in Brazil. Its implementation objectives are to describe the facilitators of and barriers to integrating CAB-LA into routine public health clinics and to evaluate whether a mobile health (mHealth) intervention optimizes decision-making on the best PrEP regimen among PrEP users and WhatsApp text messaging improves the uptake of and adherence to CAB-LA injection appointments. The primary clinical objective is to evaluate the effectiveness of CAB-LA PrEP in reducing the risk of HIV acquisition in a study cohort of participants in a context where participants can exercise choice in their HIV prevention method (injectable CAB-LA PrEP or daily oral PrEP).

Methods

Study Design

The ImPrEP CAB Brasil (Figure 1) is an implementation study of CAB-LA PrEP among 1200 young adult (aged 18 to 30 years) sexual and gender minorities (SGMs) in 6 Brazilian cities: Rio de Janeiro, São Paulo, Manaus, Salvador, Florianópolis, and Campinas (Figure 2). Each participating site provides PrEP service as a part of the Brazilian SUS.

This is a type-2 hybrid implementation-effectiveness open-label cohort study with a convergent mixed methods approach (quantitative and qualitative). A type-2 hybrid design examines both effectiveness and implementation strategies within the same study [28]. The study will draw on selected implementation science constructs from the theoretical framework for acceptability (TFA) [29], the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework [30], and the framework of Proctor et al [31] to assess the implementation outcomes for the different study objectives outlined earlier. Relevant outcomes and their level of evaluation will include the following: acceptability at the individual health provider and participant levels, reach (choice and uptake) at the individual participant level, effectiveness at the individual participant level, implementation at the setting level (overall delivery), maintenance at the individual participant level, and feasibility at the setting level.

Individuals may participate in up to 4 different steps (Figure 3). In step 1, PrEP naive individuals who come to the study sites seeking PrEP will be exposed to education and counseling about HIV prevention and be informed of the availability of both oral and injectable PrEP. Half of the enrolled participants that tested negative for HIV will receive standard of care (SOC) prevention counseling, and the other half will be exposed to an mHealth education and decision support intervention in addition to SOC prevention counseling. Participants who choose to use injectable PrEP and with an undetectable HIV viral load will enroll in step 2 on the same day, whereas those who choose to use oral PrEP will be referred to the Brazilian Public Health PrEP program at the same clinic. Of the participants who enroll in step 2 and receive a CAB-LA injection, half will be randomized to receive a digital appointment reminder using WhatsApp (Meta Platforms, Inc; the most common messaging app in Brazil). Participants in step 2 who may decide to discontinue CAB-LA during the study follow-up will be invited to step 3 and switch to oral PrEP (TDF and FTC) or keep off PrEP with HIV and sexually transmitted infection (STI) monitoring and counseling for 1 year. Participants from either step 2 or step 3 who are diagnosed with an HIV infection during follow-up will be invited to be enrolled in step 4, where antiretroviral therapy (ART) effectiveness and ART resistance will be monitored for 1 year.

Figure 1. ImPrEP CAB Brasil logo.

Figure 2. Site distribution for the ImPrEP CAB Brasil study.

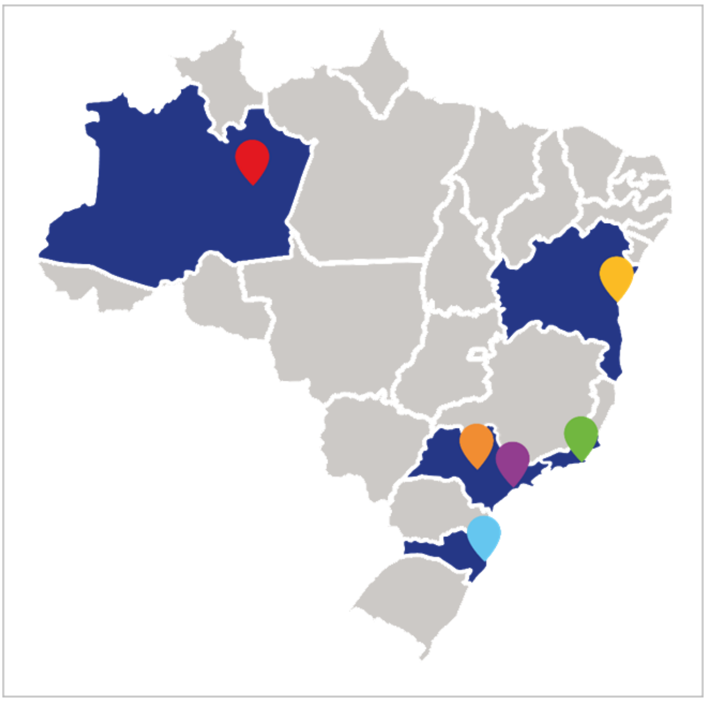
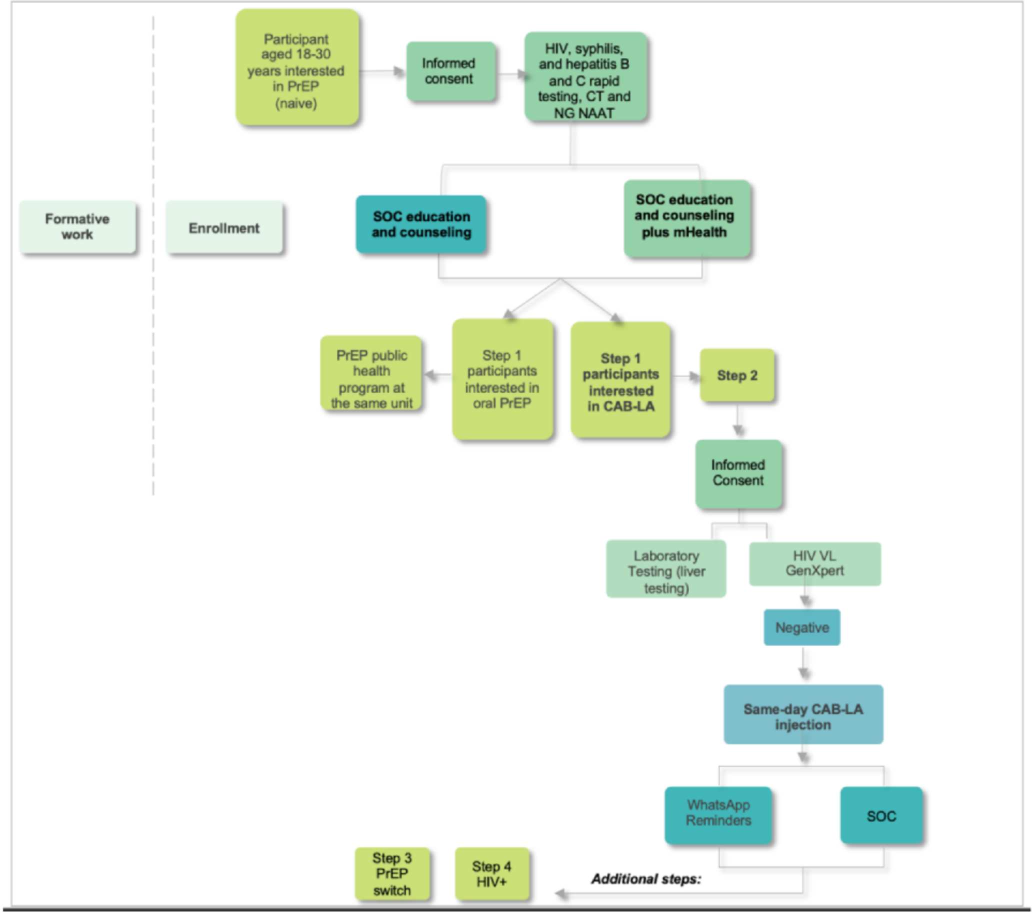


Figure 3. Overview of the study design and enrollment scheme. Only HIV-negative participants will perform standard of care (SOC) or mobile health (mHealth). Individuals living with HIV will be referred for antiretroviral therapy. CAB-LA: long-acting injectable cabotegravir; CT: Chlamydia trachomatis; NAAT: nucleic acid amplification testing; NG: Neisseria gonorrhea; PrEP: preexposure prophylaxis; VL: viral load.



Eligibility Criteria

Participants will be included if they (1) are MSM, nonbinary person (assigned as male at birth), TGW, or transgender men; (2) are aged 18 to 30 years; (3) attend one of the study clinics looking for PrEP; (4) are CAB-LA and TDF and FTC PrEP naive; (5) are willing and able to provide written informed consent and adhere to the study requirements; (6) have nonreactive or negative HIV test results, including both third-generation HIV rapid tests and an undetectable HIV RNA at enrollment for individuals choosing injectable CAB-LA; (7) have no report of hepatic dysfunction; and (8) are willing to undergo all required study procedures. The full list of exclusion criteria is described in Multimedia Appendix 1. The study team will provide checklists to aid the site teams in accessing the inclusion and exclusion criteria.

Study Procedures

Formative Work

We will conduct formative work using participatory design methods [32] to develop an initial implementation package for CAB-LA PrEP. Training modules will be developed in collaboration with health providers at the 2 sites that have had experience with CAB-LA PrEP as part of HPTN 083 (Rio de Janeiro and São Paulo) and will include slides, learning activities, and job aids (eg, education, and counseling scripts).

Process mapping [33,34] at each site will facilitate planning for optimal PrEP client flow, that is, for ensuring that PrEP users can move most efficiently and effectively through the various clinical steps needed to receive services while minimizing the work burden on health care providers and maintaining service quality. Engaging clinical staff in the process mapping exercise prompts the identification of anticipated facilitators and barriers associated with each step of service delivery as well as planning for barrier mitigation strategies. Thus, process maps can facilitate the development of "readiness" exercises once PrEP CAB-LA is approved and ready for scale-up in Brazil and elsewhere. Finally, process maps will also be used at the end of the study to guide health care providers' reflections on the facilitators of and barriers to different components of PrEP delivery. Guidelines for developing process maps and process map examples will be included in the implementation package. The implementation package developed in collaboration with providers from the CAB-LA–experienced sites will be used to prepare health care providers at the remaining 4 sites and will be revised based on their feedback.

Focus group discussions (FGDs) will be held with health care providers experienced in CAB-LA PrEP and members of the young SGM community to gather their input for the adaptation of an mHealth education and decision support tool. This tool, intended to be accessed via an electronic tablet with self-guided instructions, will be designed to convey information about both oral and CAB-LA PrEP in an engaging way, describing the benefits and drawbacks of each method, to help participants decide whether and which PrEP method may be appropriate for them. A preliminary storyboard will be developed and shared with FGD participants, and they will be asked to provide feedback on images, content, and sequence. Feedback will be incorporated into the programming of the mHealth education intervention. FGD participants will be also asked to review and refine the content of WhatsApp appointment reminders and provide input regarding the best timing for receiving the WhatsApp reminders. In addition, FGD participants will be asked about barriers and facilitators for CAB-LA PrEP implementation. Health care providers at the 2 study sites with experience in CAB-LA PrEP will be inquired about frequent questions asked by CAB-LA PrEP users about the medication as well as examples of education and counseling strategies they have used previously to facilitate CAB-LA user comprehension.

Step 1: mHealth Intervention

Overview

Step 1 focuses on evaluating the mHealth intervention and participants' choice of PrEP method (Table 1). Potential step 1 participants include individuals who come to a study clinic looking for PrEP, are PrEP naive, and have expressed an interest in taking PrEP (see eligibility criteria). All individuals accessing the service for PrEP will be assessed for study eligibility by a designated study staff member. Eligible participants will be asked whether they are interested to assess educational and counseling interventions for PrEP choice.

After signing step 1 informed consent, participants will undergo rapid tests for HIV, syphilis, and hepatitis B and C. Rectal, oral, and urine swab samples will be collected for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* nucleic acid amplification testing. If a reactive or positive result is obtained for HIV test, the respective person is deemed ineligible to receive PrEP. Additional testing to confirm suspected HIV infection will be performed in accordance with local guidelines. If an HIV infection is confirmed, the participants will receive counseling and be immediately referred for HIV treatment and care. All participants will be evaluated for signs or symptoms of a possible acute HIV infection. Those with suspected acute HIV infection will be referred for follow-up investigation within the public health system.

Participants at each site will receive counseling and education on the use of PrEP options per the local SOC from a health care provider until 100 participants indicate that they would like to start CAB-LA PrEP and enroll in step 2. Additional participants will receive an mHealth education and counseling intervention plus the SOC education and counseling until 100 participants exposed to both SOC and the mHealth intervention enroll in step 2 of the study at each site (200 participants enrolled per site).

Upon completing their exposure to SOC or SOC plus mHealth, participants will be asked to complete a brief survey administered by a study member that will capture the following information: (1) participants' knowledge and beliefs about the dosing, advantages, and disadvantages of the 2 PrEP options, (2) their method of choice for HIV prevention, and (3) the reasons for their preference for oral or CAB-LA PrEP. Participants who received the mHealth intervention will be asked additional questions about its acceptability and its value in supporting decision making on PrEP choice.

Table 1. Overview of the ImPrEP CAB Brasil study visits and schedule of procedures: steps 1 and 2 (year 1).

Procedures	Step 1 enrollment	Step 2							
		Enrollment	Month 1	Month 3	Month 5	Month 7	Month 9	Month 11	Month 13 or year 1
Administrative, behavioral, and clinical									
ICF[a] step 1	✓								
SOC[b] vs mHealth[c] and SOC	✓								
Health provider survey	✓								
PrEP[d] choice	✓								
ICF step 2		✓							
Randomization text message		✓							
Locator information		✓	✓	✓	✓	✓	✓	✓	✓
HIV counseling		✓	✓	✓	✓	✓	✓	✓	✓
Reproductive counseling[e]		✓	✓	✓	✓	✓	✓	✓	✓
Condom, lubricant, and HIVST[f]		✓	✓	✓	✓	✓	✓	✓	✓
Adherence counseling		✓	✓	✓	✓	✓	✓	✓	✓
HIV risk and mental health assessment		✓	✓	✓	✓	✓	✓	✓	✓
Medical history		✓	✓	✓	✓	✓	✓	✓	✓
Concomitant medication		✓	✓	✓	✓	✓	✓	✓	✓
Directed physical examination		✓	✓	✓	✓	✓	✓	✓	✓
Weight and vital signs		✓	✓	✓	✓	✓	✓	✓	✓
Height		✓							✓
HIVST report			✓	✓	✓	✓	✓	✓	✓
Laboratory									
Third-generation HIV rapid test	✓		✓	✓	✓	✓	✓	✓	✓
HIV RNA PCR[g] (GeneXpert)		✓							✓[h]
HBV[i] rapid test	✓			✓		✓		✓	✓
HCV[j] rapid test	✓			✓		✓		✓	✓
Syphilis test[k]	✓			✓		✓		✓	✓[h]
Liver function test (ALT[l])		✓		✓		✓		✓	✓
Oropharyngeal CT[m] and NG[n] test	✓			✓		✓		✓	✓[h]
Urine CT and NG test	✓			✓		✓		✓	✓[h]
Rectal swab CT and NG test	✓			✓		✓		✓	✓[h]
Urine pregnancy test		✓[e]	✓[e]	✓[e]	✓[e]	✓[e]	✓[e]	✓[e]	✓[e]

Procedures	Step 1 enrollment	Step 2							
		Enrollment	Month 1	Month 3	Month 5	Month 7	Month 9	Month 11	Month 13 or year 1
Plasma and Serum storage		✓	✓	✓	✓	✓	✓	✓	✓
Study medication									
Injection		✓	✓	✓	✓	✓	✓	✓	✓
ISR[o] evaluation		✓	✓	✓	✓	✓	✓	✓	✓

[a] ICF: informed consent form.

[b] SOC: standard of care.

[c] mHealth: mobile health.

[d] PrEP: preexposure prophylaxis.

[e] Only for transgender men.

[f] HIVST: HIV self-testing.

[g] PCR: polymerase chain reaction.

[h] Collect if it is the final study visit.

[i] HBV: hepatitis B.

[j] HCV: hepatitis C.

[k] Rapid test plus nontreponemal syphilis testing, as indicated.

[l] ALT: alanine aminotransferase.

[m] CT: *Chlamydia trachomatis*.

[n] NG: *Neisseria gonorrhea*.

[o] ISR: injection site reaction.

Qualitative Interview

A small subset of step 1 participants from each site (8 SOC participants per site and 12 SOC plus mHealth participants per site) will be asked to participate in a qualitative interview to evaluate their perceived advantages and disadvantages of different PrEP products and reasons for their choice. Those who received SOC plus mHealth will be asked to share their feedback on the acceptability of the mHealth intervention. Audio recordings will be transcribed *verbatim* by study staff.

PrEP Choice and Administration

Following the completion of step 1 procedures, participants will be asked by a health provider whether they choose oral PrEP or CAB-LA PrEP. Participants who indicate a preference for oral PrEP will be referred to receive it in the routine HIV prevention service and will not be navigated by the study team. The study team will not actively follow-up on these participants but will retain their participant ID in case they express interest in switching to CAB-LA PrEP later in the study period. Participants who choose CAB-LA PrEP will be invited on the same day to participate in step 2 of the study and will undergo additional consent, screening, HIV viral load testing, enrollment, and data collection procedures.

Step 2

Overview

Step 1 participants who chose CAB-LA PrEP after receiving either SOC counseling or SOC counseling plus mHealth will be eligible to participate on the same day in step 2 (Tables 1 and 2). Step 2 aims to understand the safety and effectiveness of CAB-LA PrEP in an environment of choice, assess an implementation strategy (WhatsApp appointment reminder) for improving adherence to CAB-LA injections, evaluate HIV testing strategies, evaluate the acceptability of STI sample self-collection and HIV self-testing, and understand the facilitators of and barriers to the integrated delivery of CAB-LA PrEP in the existing oral PrEP services.

HIV viral load testing will be performed in step 2 participants using the GeneXpert platform to confirm their eligibility. With a detectable HIV result, the person will be deemed ineligible to receive PrEP and will be referred to treatment and care through the public health system. Those who do not have detectable HIV results and do not meet any other exclusion criteria will receive their first CAB-LA injection.

The first follow-up visit will happen in 1 month, and the subsequent follow-up visits will happen once in every 2 months. At every follow-up visit, sites will perform HIV testing using the third-generation (antibody) HIV rapid tests before providing the next CAB-LA injection. Blood-based HIV self-testing will be performed by the study participant 24 hours before all follow-up visits, and the results will be reported by the participant during the study visit. If a participant does not bring the HIV self-testing report, the test will be performed at the clinic in a private space and before the study visit procedures. Regularly, additional HIV testing with stored samples will be conducted retrospectively, in batches, with fourth-generation HIV rapid test, Abbot ARCHITECT HIV Ag/AB combo assay, and the HIV-1 and HIV-2 RNA Viral Load COBAS 5800 platform. For positive or inconclusive results, the Geenius HIV1/2 confirmatory assay will be used. In all study visits, samples will be stored for additional retrospective HIV testing at the study central lab at INI-Fiocruz, as necessary. At any point of the study follow-up, participants with suspected HIV infection will be referred for follow-up investigation under study protocol. If HIV infection is confirmed at any point, CAB-LA injection will be stopped, and the participant will be immediately linked to care and ART initiation (according to the Brazilian guidelines) and will be invited to maintain a 1-year follow-up in the study with quarterly visits (step 4). Resistance testing will also be performed (if feasible, based on the number of viral copies).

Table 2. Overview of the ImPrEP CAB Brasil study visits and schedule of procedures: step 2 (year 2).

Procedures	Month 15	Month 17	Month 19	Month 21	Month 23	Month 25 or year 2
Administrative, behavioral, and clinical						
Locator information	✓	✓	✓	✓	✓	✓
HIV counseling	✓	✓	✓	✓	✓	✓
Reproductive counseling	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]
Condom, lubricant, and HIVST[b]	✓	✓	✓	✓	✓	✓
Adherence counseling	✓	✓	✓	✓	✓	✓
HIV Risk and Mental health assessment	✓	✓	✓	✓	✓	✓
Medical history	✓	✓	✓	✓	✓	✓
Concomitant medication	✓	✓	✓	✓	✓	✓
Directed physical examination	✓	✓	✓	✓	✓	✓
Weight and vital Signs	✓	✓	✓	✓	✓	✓
Height						✓
HIVST report	✓	✓	✓	✓	✓	✓
Laboratory						
Third-generation HIV rapid test	✓	✓	✓	✓	✓	✓
HIV RNA PCR[c] (GeneXpert)						✓
HBV[d] rapid test			✓			✓
HCV[e] rapid test			✓			✓
Syphilis test[f]	✓		✓		✓	✓
Liver function test (ALT[g])			✓			✓
Urine CT[h] and NG[i] test	✓		✓		✓	✓
Oropharyngeal CT and NG test	✓		✓		✓	✓
Rectal swab CT and NG test	✓		✓		✓	✓
Urine pregnancy test	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]
Plasma and Serum storage	✓	✓	✓	✓	✓	✓
Study medication						
Injection	✓	✓	✓	✓	✓	✓
ISR[j] evaluation	✓	✓	✓	✓	✓	✓

[a] Only for transgender men.

[b] HIVST: HIV self-testing.

[c] PCR: polymerase chain reaction.

[d] HBV: hepatitis B.

[e] HCV: hepatitis C.

[f] Rapid test plus nontreponemal syphilis testing, as indicated.

[g] ALT: alanine aminotransferase.

[h] CT: *Chlamydia trachomatis*.

[i] NG: *Neisseria gonorrhea*.

[j] ISR: injection site reaction.

Randomized Injection Reminders via WhatsApp

As part of the consent process for enrollment in step 2, participants will be informed that they may be selected to receive a WhatsApp appointment reminder 2 days before each scheduled injection visit. Participants will be randomly assigned (1:1) to either the SOC (verbal communication and paper appointment slip letting participants know when they should come back to the clinic) or SOC plus WhatsApp arm using simple or block randomization and will be informed of which arm they are assigned to during the step 2 enrollment visit. During the month 5 visit, participants who were randomized to receive WhatsApp appointment reminders will be asked to complete a brief participant survey to assess the acceptability of receiving WhatsApp appointment reminders. If a participant does not return for their month 5 visit, the survey will be administered at their next clinic visit.

Switching From Oral TDF and FTC PrEP to CAB-LA PrEP

Step 1 participants may choose to switch from oral TDF and FTC PrEP to CAB-LA PrEP only once during the study period. These participants will undergo all step 2 procedures prior to receiving the first CAB-LA injection. Individuals with hepatitis B surface antigen–positive results and indication of hepatitis B treatment will be counseled to remain on oral TDF and FTC for PrEP. If hepatitis B treatment is not indicated and the individual chooses to switch to CAB-LA, enrollment will be allowed unless exclusionary hepatic criteria are identified (alanine aminotransferase ≥5 × upper limit of normal).

Specific Visit Procedures

Participants included in step 2 to receive CAB-LA PrEP who decide to discontinue participation before completing the minimum 1 year and maximum 2 years of study follow-up will complete early discontinuation visit procedures (Table 3). These participants will be invited to participate in a 12-month extension follow-up to safely monitor the CAB-LA “tail” phase and will be offered oral PrEP (TDF and FTC). Some of them will be invited to participate in a qualitative interview to describe their experience with CAB-LA PrEP and the reason for decision to discontinue CAB-LA PrEP.

If a participant misses an appointment but wishes to continue with CAB-LA PrEP, reloading with CAB-LA may be necessary, and the participant will need repeat HIV rapid tests before receiving a new CAB-LA injection. Oral bridging with TDF and FTC will be offered to participants who anticipate not being able to attend their scheduled injections (planned missed injections) in the +7 or −7-day window, starting approximately 2 months after the last injection dose of CAB-LA.

Table 3. The ImPrEP CAB Brasil–specific visit procedures.

Procedures	Final visit	Early discontinuation[a]	Seroconversion	STI[b] interim visit	AE[c] interim visit
Administrative, behavioral, and clinical					
HIV counseling	✓	✓	✓	✓	
Reproductive counseling[d]	✓	✓	✓[e]	✓[e]	✓[e]
Condom, lubricant, and HIVST[f]	✓	✓	✓	✓	
Adherence counseling	✓	✓	✓		
HIV Risk and Mental health assessment	✓	✓	✓		✓
Medical history	✓	✓	✓	✓	✓
Concomitant medication	✓	✓	✓	✓	✓
Physical examination	✓	✓	✓	✓	✓
Weight and vital Signs	✓	✓	✓		
Height	✓	✓	✓		
HIVST report	✓	✓[g]	✓[g]	✓[g]	✓[g]
Laboratory					
Third-generation HIV rapid test	✓	✓	✓	✓	
HIV RNA PCR[h] (GeneXpert)	✓	✓		✓[i]	
HIV RNA PCR			✓		
HBV[j] rapid test	✓	✓	✓	✓	
HCV[k] rapid test	✓	✓	✓	✓	
Syphilis test	✓	✓	✓	✓	
Liver function test (ALT[l])	✓	✓	✓		✓[m]
Creatinine or ClCr[n]	✓[o]	✓[o]			
Oropharyngeal CT[p] and NG[q] test	✓	✓	✓	✓	
Urine CT and NG test	✓	✓	✓	✓	
Rectal swab CT and NG test	✓	✓	✓	✓	
Urine pregnancy test	✓[r]	✓[r]	✓[r]	✓[e]	✓[e]
CD4 cell count			✓		
HIV resistance testing			✓		
Plasma and serum storage	✓	✓	✓		
Study medication					
Injection	✓[d]				
ISR[s] evaluation	✓[d]				✓[t]
ART[u] initiation			✓		
TDF[v] and FTC[w] PrEP[x]	✓[y]	✓[y]			

[a] Any discontinuation occurring before the anticipated final study visit and not related to seroconversion.

[b] STI: sexually transmitted infection.

[c] AE: adverse event.

[d] Depending on the scenario that presents itself at the time of the final study visit.

[e] If pregnancy is suspected.

[f] HIVST: HIV self-test.

[g] If available.

[h] PCR: polymerase chain reaction.

[i] If HIV infection is suspected.

[j] HBV: hepatitis B.

[k] HCV: hepatitis C.

[l] ALT: alanine aminotransferase.

[m] Interim visits performed because of suspected hepatic adverse events should collect liver function tests.

[n] ClCr: creatinine clearance.

[o] If oral preexposure prophylaxis is initiated.

[p] CT: *Chlamydia trachomatis.*

[q] NG: *Neisseria gonorrhea.*

[r] Only for transgender men.

[s] ISR: injection site reaction.

[t] Interim visits performed because injection site reactions should follow specific protocol instructions.

[u] ART: antiretroviral therapy.

[v] TDF: tenofovir disoproxil fumarate.

[w] FTC: emtricitabine.

[x] PrEP: preexposure prophylaxis.

[y] In a final study visit or an early discontinuation visit, if the participant chooses to switch to oral preexposure prophylaxis, they will be followed in the 12-month follow-up study.

Study Step 3

Participants from step 2 will be invited to participate in step 3 if they meet the following eligibility requirements: (1) the participant decides to switch from CAB-LA to oral TDF and FTC during study follow-up or (2) the participant decides to early discontinue CAB-LA injections but does not want to start oral PrEP.

Participants who choose to switch from CAB-LA to oral TDF and FTC will need to complete the procedures related to the daily oral TDF and FTC enrollment visit (Table 3). The procedures are the same as those specified under both the early discontinuation visit and the final study visit, plus creatinine dosage and estimated creatinine clearance. Once enrolled in step 3, the follow-up consists of visits scheduled for the following intervals: month 1, month 4, month 7, month 10, and month 13 or year 1 (Table 4).

Participants who choose to keep off PrEP will be invited to be followed up in step 3. Participants will have quarterly visits with sexual risk assessment, HIV and STI testing and counseling, and the provision of condoms and lubricants.

Table 4. The ImPrEP CAB Brasil step 3 procedures: daily oral preexposure prophylaxis.

Procedures	Enrollment[a]	Month 1	Month 4	Month 7	Month 10	Month 13 or year 1
Administrative, behavioral, and clinical						
Locator information	✓	✓	✓	✓	✓	✓
HIV risk assessment and counseling	✓	✓	✓	✓	✓	✓
Condom, lubricant, and HIVST[b]	✓	✓	✓	✓	✓	✓
Adherence counseling	✓	✓	✓	✓	✓	✓
Medical history	✓	✓	✓	✓	✓	✓
Concomitant medication	✓	✓	✓	✓	✓	✓
Directed physical examination	✓	✓	✓	✓	✓	✓
Weight and vital Signs	✓	✓	✓	✓	✓	✓
Height	✓					✓
HIVST report		✓	✓	✓	✓	✓
Laboratory						
Third-generation HIV rapid test	✓	✓	✓	✓	✓	✓
HBV[c] rapid test	✓			✓		✓
HCV[d] rapid test	✓			✓		✓
Syphilis test	✓		✓	✓	✓	✓
Liver function test (ALT[e])[f]	✓			✓		✓
Creatinine or ClCr[g]	✓			✓		✓
Oropharyngeal CT[h] and NG[i] test	✓		✓	✓		✓
Urine CT and NG test	✓		✓	✓	✓	
Rectal swab CT and NG	✓		✓	✓	✓	
Plasma and serum storage	✓	✓	✓	✓	✓	✓
PrEP[j] medication						
TDF[k] and FTC[l] dispensation	✓	✓	✓	✓	✓	✓

[a] Enrollment procedures for oral preexposure prophylaxis are the same as those of the final study visit and early discontinuation visit.

[b] HIVST: HIV self-test.

[c] HBV: hepatitis B.

[d] HCV: hepatitis C.

[e] ALT: alanine aminotransferase.

[f] Only for those who switched from long-acting injectable cabotegravir.

[g] ClCr: creatinine clearance.

[h] CT: *Chlamydia trachomatis.*

[i] NG: *Neisseria gonorrhea*.

[j] PrEP: preexposure prophylaxis.

[k] TDF: tenofovir disoproxil fumarate.

[l] FTC: emtricitabine.

Step 4

Overview

Participants who are diagnosed with HIV while using CAB-LA during study follow-up will perform the procedures related to the *seroconversion visit* (Table 3) and will be invited to join step 4. HIV treatment will be initiated in accordance with the Brazilian guidelines. The participant will be requested to come to the study clinic for quarterly visits for 1 year. HIV RNA polymerase chain reaction will be evaluated in all visits. If a participant is identified with suspected virological failure, a new resistance test will be performed, and a new ART regimen will be selected based on the resistance test results.

Qualitative Assessments During the Study

At month 7, a select set of participants will be invited to participate in a semistructured qualitative interview guided by the TFA to understand their experiences with CAB-LA PrEP and with the clinic services, including WhatsApp appointment reminders and HIV self-testing. Up to 16 participants will be interviewed at each site. The aim is to sample approximately 8 individuals who continually adhered to their injection schedule, approximately 4 individuals who needed to reload, and 4 individuals who discontinued CAB-LA injections. Interviews will be conducted in Portuguese and audio recorded. Interviewers will complete a debriefing report summarizing the key points of the interview within 12 hours of completing the interview. Audio recordings will be transcribed *verbatim* by study staff.

Health Care Provider Survey

At the end of the implementation study, health providers will be asked to complete an in-depth interview on their experiences integrating CAB-LA in the existing PrEP services and their views on the mHealth tool and WhatsApp reminders. The interview questionnaire will seek to understand the facilitators of and barriers to CAB-LA service delivery and will be informed by implementation science constructs drawn from the TFA (affective attitude, burden, ethicality, intervention coherence, opportunity cost or relative advantage, perceived effectiveness, and self-efficacy) [29], Reach, Effectiveness, Adoption, Implementation, and Maintenance (implementation) [30], and Proctor et al [31] (feasibility). The definitions of the constructs are summarized in Multimedia Appendix 1. During these interviews, health care providers will also be asked to review their most recent process mapping, describe changes that happened with CAB-LA implementation, and describe ongoing “pain points." Short Likert-scale surveys will also be embedded in the questionnaires to gather quantitative data on acceptability.

Major Outcomes

The major clinical and implementation outcomes, levels of analysis, methods, and timing are described in Table 5.

Table 5. Summary of major clinical and implementation outcomes.

Major outcomes	Level	Methods and timing
Clinical		
HIV incidence (effectiveness)	• Individual	• HIV test completed at each visit and interim analysis after 12 months of enrollment and upon the completion of the study
Resistance to CAB-LA[a] or oral TDF[b] and FTC[c] PrEP[d] (effectiveness and safety)	• Individual	• HIV genotyping test at the time of seroconversion or suspected viral failure at any time during the 12-month follow-up
Implementation		
mHealth[e] intervention acceptability (from TFA[f])	• Individual participant (client or potential PrEP user)	• Survey of all participants receiving mHealth intervention immediately upon the completion of SOC[g] or SOC and mHealth education and counseling • Qualitative interview (12 participants per site who receive mHealth and SOC) during the enrollment period for step 1
mHealth intervention effectiveness (from RE-AIM[h])	• Individual participant (client or potential PrEP user)	• Knowledge survey of all participants receiving mHealth intervention immediately upon the completion of SOC or SOC and mHealth education and counseling • Qualitative interviews of 8 participants per site who receive SOC counseling and 12 participants who receive SOC and mHealth education and counseling about their knowledge and beliefs about oral and CAB-LA PrEP in relation to their choice • Adherence data at the end of the study
mHealth tool acceptability (TFA) and feasibility (from Proctor et al [31])	• Individual provider and clinic setting	• Interviews at the end of the study including questions that will generate both qualitative and quantitative data
WhatsApp appointment reminder acceptability (TFA)	• Individual client or PrEP user	• Brief survey given to each participant who received WhatsApp appointment reminders at their month 5 visit • Qualitative interviews (16 participants per site) conducted at month 7
WhatsApp appointment reminder effectiveness (from RE-AIM)	• Individual client or PrEP user	• Study adherence data at the middle and end of the study • Qualitative interviews (16 participants per site) conducted at month 7
WhatsApp appointment reminder feasibility (Proctor et al [31]) and implementation (RE-AIM)	• Individual provider and clinic setting	• Interviews at the end of the study including questions that will generate both qualitative and quantitative data • WhatsApp appointment reminder program data files
Acceptability (TFA) of STI[i] self-test	• Individual client or CAB-LA PrEP user	• Qualitative interviews (16 participants per site) conducted at month 7
Acceptability (TFA) and feasibility (Proctor) of STI self-test	• Individual provider and clinic setting	• Qualitative interviews at the end of the study
Acceptability (TFA) of HIV self-testing	• Individual client or CAB- LA PrEP user	• Qualitative interviews (16 participants per site) conducted at month 7
Implementation: barriers to and facilitators of the integrated delivery of CAB-LA PrEP in the existing oral PrEP services	• Individual participant (client or CAB-LA PrEP user) • Individual provider and clinic setting	• Qualitative interviews (16 participants per site) conducted at month 7 • Interviews at the end of the study including questions that will generate both qualitative and quantitative data
Feasibility of CAB-LA PrEP in public clinic PrEP services	• Individual provider and clinic setting	• Interviews at the end of the study including questions that will generate both qualitative and quantitative data • Review of routine study documentation

Major outcomes	Level	Methods and timing
Maintenance	Individual participant (client or CAB-LA PrEP user)	Adherence to injection visits at the middle and end of the study
Reach	Individual participant and cohort characteristics	Survey of all participants receiving mHealth intervention immediately upon the completion of SOC or SOC and mHealth education and counseling about the rationale for their PrEP choice Qualitative interview (12 participants per site who receive mHealth and SOC) during the enrollment period for step 1 that will include questions about the rational for their PrEP choice Study data on the number of participants enrolled and pace of enrollment and demographic data

[a] CAB-LA: long-acting injectable cabotegravir.

[b] TDF: tenofovir disoproxil fumarate.

[c] FTC: emtricitabine.

[d] PrEP: preexposure prophylaxis.

[e] mHealth: mobile health.

[f] TFA: theoretical framework for acceptability.

[g] SOC: standard of care.

[h] RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance.

[i] STI: sexually transmitted infection.

Statistical Analysis

Our major clinical outcome (primary analysis) will adopt an intent-to-treat approach and compare HIV incidence (and 95% CI) among individuals in the CAB-LA cohort with HIV incidence (and 95% CI) among individuals on daily PrEP within the Brazilian SUS. We will estimate the rate of HIV incidence with 95% CI in each group using Kaplan-Maier statistics. We will use a Cox proportional hazard model to assess the relative hazard (hazard ratio with 95% CI) between the 2 groups, overall and by site, age category, and gender.

It is anticipated that there will be an average follow-up of 1.5 years per person with at least 75% retention each year (N=1350 person-years of follow-up) and that incidence will be higher than that observed in HPTN 083 in Latin America (0.60 per 100 person-years). If the observed incidence is approximately 50% higher than that observed in HPTN 083, the study will be able to estimate an incidence of 0.85 (95% CI of 0.47-1.42) per 100 person-years (Table 6). If the observed incidence is approximately twice as high as that observed in HPTN 083, then the study will be able to estimate an incidence of 1.18 (95% CI of 0.72-1.82) per 100 person-years. If the comparison group is of similar size, then the incidence would need to be 2.2 to 2.5 times higher to observe statistically significant differences between CAB-LA and oral PrEP. However, if the comparison group is 3 times larger, then the incidence would need to be 2 to 2.2 times higher to observe statistically significant differences between CAB-LA and oral PrEP.

To reach the primary implementation objective, we will report descriptive statistics (numbers, proportions, and medians) of Likert-scale responses to the TFA constructs (affective attitude, burden, ethicality, intervention coherence, opportunity costs, and perceived effectiveness and self-efficacy) among health care providers.

Table 6. Observable differences given potential HIV incidence in intervention and comparison groups.

Group	Incidence (95% CI)
Intervention: CAB-LA[a]PrEP[b]	
1350 py[c] incidence/100 py	
8 cases	0.59 (0.28-1.13)
12 cases	0.89 (0.48-1.51)
18 cases	1.19 (0.71-1.88)
Comparison: oral PrEP in the Brazilian health system	
1350 py[c] incidence/100 py	
24 cases	1.78 (1.17-2.61)
30 cases	2.22 (1.53-3.13)
36 cases	2.67 (1.90-3.65)
2700 py incidence/100 py	
43 cases	1.59 (1.17-2.13)
54 cases	2.00 (1.52-2.59)
66 cases	2.44 (1.91-3.09)
4050 py incidence/100 py	
60 cases	1.48 (1.14-1.89)
78 cases	1.93 (1.53-2.39)
94 cases	2.32 (1.89-2.83)

[a] CAB-LA: long-acting cabotegravir.

[b] PrEP: preexposure prophylaxis.

[c] py: person-years.

Ethics Approval

The study protocol and informed consent forms were reviewed and approved by the Evandro Chagas National Institute of Infectious Diseases (INI), Fundação Oswaldo Cruz (Fiocruz) institutional review board (August 25, 2022; #CAAE 59166522.7.1001.5262) and the Research Ethics Review Committee at the WHO with respect to scientific content and compliance with applicable research and human participants' regulations. The protocol, site-specific informed consent forms, participant education and recruitment materials, and any requested documents and their subsequent modifications were also be reviewed and approved by the ethical review bodies responsible for the oversight of the research conducted at each of the study sites. Informed consent will be obtained in 2 study steps. In step 1, consent will be obtained from participants who visit the site looking for PrEP. In step 2, the participants will provide consent for the use of injectable CAB-LA PrEP. This study may be paused or stopped at any time by the local institutional review board or the study's principal investigator.

Results

During the third and fourth quarters of 2022, we obtained regulatory approvals; programmed data entry or management systems; trained sites; and performed community consultancy, process mapping, and formative work. Study enrollment is anticipated to happen in the first quarter of 2023. The planned activities are presented in Table 7.

Table 7. Planned activities for the ImPrEP CAB Brasil study (2022-2024).

Activities	Quarter 2 of 2022	Quarter 3 of 2022	Quarter 4 of 2022	Quarter 1 of 2023	Quarter 2 of 2023	Quarter 3 of 2023	Quarter 4 of 2023	Quarter 1 of 2024	Quarter 2 of 2024	Quarter 3 of 2024	Quarter 4 of 2024
Regulatory submission	✓										
Regulatory approvals		✓	✓								
Data entry or management systems programmed	✓	✓									
Drug importation and distribution				✓	✓						
Site equipping	✓	✓	✓								
SOP[a] and SSP[b] development	✓	✓	✓	✓							
Staff selection	✓	✓									
Staff training		✓	✓								
Process mapping			✓								
Formative work		✓	✓	✓							
Enrollment					✓	✓	✓				
Follow-up						✓	✓	✓	✓	✓	
Data consolidation							✓	✓	✓	✓	
Data analysis										✓	✓
Final report											✓
Manuscript submission											✓

[a] SOP: standard operational procedures.

[b] SSP: study-specific plan.

Discussion

Principal Findings

ImPrEP CAB Brasil is the first study to evaluate CAB-LA PrEP implementation in Latin America. Latin America is one of the regions where PrEP scale-up, as well as the reduction of inequalities in PrEP access, is most needed. We hereby describe the study protocol, providing details of the study's objectives, procedures, outcomes, and statistical analysis to aid researchers from Brazil, Latin America, and other regions during CAB-LA implementation within public health systems.

To date, no clinical or implementation results are available for injectable PrEP with CAB-LA in real-life settings. We expect the acceptability of CAB-LA in Brazil to be high. Data from a web-based survey and descriptive choice experiment have shown that injectable PrEP was preferred over other PrEP modalities in Brazil [35-37]. Our results regarding PrEP choice among individuals seeking PrEP in health units will increase this body of evidence. For instance, in the field of contraception, the increased availability of options and informed decision-making in the choice of contraception have improved the uptake and persistence of the methods used [38]. It is likely that expanding the method options for PrEP will similarly improve uptake and persistence.

Our study will also contribute to the expansion of the knowledge on the best HIV testing algorithm to be implemented in low- and middle-income countries when using CAB-LA for PrEP, which is a major need.

This study is coordinated by INI-Fiocruz, an internationally recognized research institution and the leading reference center in Latin America for infectious diseases. INI-Fiocruz has been working closely with the Brazilian Ministry of Health (BMOH), the Pan American Health Organization, Unitaid, and the WHO to address evidence gaps and identify key learnings to inform the introduction of injectable CAB-LA PrEP among the most susceptible populations. The INI-Fiocruz team has extensive experience in engaging SGMs in demand creation, providing HIV combination prevention services that foster PrEP uptake and retention and advocating for enhanced HIV prevention policies [39,40]. In addition to being a recognized academic institution, INI-Fiocruz is the largest HIV prevention and treatment center in Rio de Janeiro, Brazil, with over 1000 SGMs tested per year [41]. The INI-Fiocruz study team was responsible for planning, designing, and implementing the PrEP Brasil study [14], which provided critical evidence to support the BMOH in the incorporation of PrEP as a public health policy. The team also coordinated the ImPrEP study, which provided information for PrEP surveillance in Brazil, and PrEP implementation in Mexico and Peru [16].

This study is poised for informing policy makers, building public health staff capacity, and creating environments for successfully implementing the proposed prevention strategies in collaboration with the BMOH. Moreover, we expect the mHealth component to be highly acceptable and possibly implemented within the Brazilian public health system.

Conclusions

The ImPrEP CAB Brasil study will contribute to the generation of valuable information to inform the design of programmatic strategies for implementing CAB-LA PrEP in Brazil, maximizing the public health impact of PrEP in reducing HIV incidence among young SGMs in Brazil, Latin America, and other countries in the Global South.

Acknowledgments

This project was made possible thanks to Unitaid’s funding and support. Unitaid accelerates access to innovative health products and lays the foundation for their scale-up by countries and partners. Unitaid is a hosted partnership of the World Health Organization. ViiV Healthcare will provide the study medication. BG and TST are funded by the National Council of Technological and Scientific Development and Carlos Chagas Filho Foundation for Research Support in the State of Rio de Janeiro. The authors would like to thank the Ministry of Health of Brazil for their support.

The contributors associated with ImPrEP CAB Brasil Study Group are as follows: Anna Hellstrom, Alessandro Farias, Marcus V Lacerda, José Valdez Madruga, Josué N de Lima, Ronaldo Zonta, Roberta Trefiglio, Renato Lima, Antonio Carvalho, Sandra Wagner Cardoso, Débora Marra Faber Barreto, Egydio Sampaio Pessanha, Sandro Nazer, Carla de Barcellos Rocha, Claudio Gruber Mann, Nilo M Fernandes, Cristina Jalil, Lucilene Freitas, Eduardo Carvalheira Netto, Mônica Derrico, Flávia Lessa, Larissa Vilela, Laylla Monteiro, Josias Freitas, Toni Araújo, Daniel R B Bezerra, Luana M S Marins, Robson Pierre N. Silva.

Data Availability

This manuscript is a description of a study protocol; therefore, no data are available.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional information about ImPrEP CAB-Brasil protocol.
[DOCX File , 59 KB - publichealth_v9i1e44961_app1.docx]

References

1. UNAIDS Data 2021. Joint United Nations Programme on HIV/AIDS. 2021. URL: https://www.unaids.org/sites/default/files/media_asset/JC3032_AIDS_Data_book_2021_En.pdf [accessed 2023-03-15]
2. Coelho LE, Torres TS, Veloso VG, Grinsztejn B, Jalil EM, Wilson EC, et al. The prevalence of HIV among men who have sex with men (MSM) and young MSM in Latin America and the Caribbean: a systematic review. AIDS Behav 2021 Oct;25(10):3223-3237. [doi: 10.1007/s10461-021-03180-5] [Medline: 33587242]
3. Bastos FI, Bastos LS, Coutinho C, Toledo L, Mota JC, Velasco-de-Castro CA, “Divas Research Group”. HIV, HCV, HBV, and syphilis among transgender women from Brazil: assessing different methods to adjust infection rates of a hard-to-reach, sparse population. Medicine (Baltimore) 2018 May;97(1S Suppl 1):S16-S24 [FREE Full text] [doi: 10.1097/MD.0000000000009447] [Medline: 29794601]
4. Grinsztejn B, Jalil EM, Monteiro L, Velasque L, Moreira RI, Garcia AC, Transcender Study Team. Unveiling of HIV dynamics among transgender women: a respondent-driven sampling study in Rio de Janeiro, Brazil. Lancet HIV 2017 Apr;4(4):e169-e176 [FREE Full text] [doi: 10.1016/S2352-3018(17)30015-2] [Medline: 28188030]
5. Luz PM, Benzaken A, de Alencar TM, Pimenta C, Veloso VG, Grinsztejn B. PrEP adopted by the Brazilian national health system: what is the size of the demand? Medicine (Baltimore) 2018 May;97(1S Suppl 1):S75-S77 [FREE Full text] [doi: 10.1097/MD.0000000000010602] [Medline: 29794605]
6. Painel PrEP. Ministério da Saúde Brasil. 2022. URL: https://www.gov.br/aids/pt-br/assuntos/prevencao-combinada/prep-profilaxia-pre-exposicao/painel-prep [accessed 2023-03-15]

7. Sullivan A, Chiavenna C, Cartier A, Jaafar S, Mitchell H, Ogaz D, et al. BPD2/7 HIV and STI incidence among MSM users and non-users of HIV PrEP in England: results from the Impact trial. In: Proceedings of the 18th European AIDS Conference. 2021 Presented at: EACS '21; October 27-30, 2021; London, UK.
8. Scheer S, Hsu L, Schwarcz S, Pipkin S, Havlir D, Buchbinder S, et al. Trends in the San Francisco human immunodeficiency virus epidemic in the "getting to zero" era. Clin Infect Dis 2018 Mar 19;66(7):1027-1034 [FREE Full text] [doi: 10.1093/cid/cix940] [Medline: 29099913]
9. Grulich AE, Jin F, Bavinton BR, Yeung B, Hammoud MA, Amin J, Expanded PrEP Implementation in Communities New South Wales (EPIC-NSW) research group. Long-term protection from HIV infection with oral HIV pre-exposure prophylaxis in gay and bisexual men: findings from the expanded and extended EPIC-NSW prospective implementation study. Lancet HIV 2021 Aug;8(8):e486-e494. [doi: 10.1016/S2352-3018(21)00074-6] [Medline: 34217426]
10. Grulich AE, Guy R, Amin J, Jin F, Selvey C, Holden J, Expanded PrEP Implementation in Communities New South Wales (EPIC-NSW) research group. Population-level effectiveness of rapid, targeted, high-coverage roll-out of HIV pre-exposure prophylaxis in men who have sex with men: the EPIC-NSW prospective cohort study. Lancet HIV 2018 Nov;5(11):e629-e637. [doi: 10.1016/S2352-3018(18)30215-7] [Medline: 30343026]
11. Koss CA, Havlir DV, Ayieko J, Kwarisiima D, Kabami J, Chamie G, et al. HIV incidence after pre-exposure prophylaxis initiation among women and men at elevated HIV risk: a population-based study in rural Kenya and Uganda. PLoS Med 2021 Feb;18(2):e1003492 [FREE Full text] [doi: 10.1371/journal.pmed.1003492] [Medline: 33561143]
12. Haberer JE, Bangsberg DR, Baeten JM, Curran K, Koechlin F, Amico KR, et al. Defining success with HIV pre-exposure prophylaxis: a prevention-effective adherence paradigm. AIDS 2015 Jul 17;29(11):1277-1285 [FREE Full text] [doi: 10.1097/QAD.0000000000000647] [Medline: 26103095]
13. Laborde ND, Kinley PM, Spinelli M, Vittinghoff E, Whitacre R, Scott HM, et al. Understanding PrEP persistence: provider and patient perspectives. AIDS Behav 2020 Sep;24(9):2509-2519 [FREE Full text] [doi: 10.1007/s10461-020-02807-3] [Medline: 32048078]
14. Grinsztejn B, Hoagland B, Moreira RI, Kallas EG, Madruga JV, Goulart S, PrEP Brasil Study Team. Retention, engagement, and adherence to pre-exposure prophylaxis for men who have sex with men and transgender women in PrEP Brasil: 48 week results of a demonstration study. Lancet HIV 2018 Mar;5(3):e136-e145. [doi: 10.1016/S2352-3018(18)30008-0] [Medline: 29467098]
15. Konda KA, Torres TS, Mariño G, Ramos A, Moreira RI, Leite IC, ImPrEP Study Group. Factors associated with long-term HIV pre-exposure prophylaxis engagement and adherence among transgender women in Brazil, Mexico and Peru: results from the ImPrEP study. J Int AIDS Soc 2022 Oct;25 Suppl 5(Suppl 5):e25974 [FREE Full text] [doi: 10.1002/jia2.25974] [Medline: 36225148]
16. Veloso VG, Cáceres CF, Hoagland B, Moreira RI, Vega-Ramírez H, Konda KA, ImPrEP Study Group. Same-day initiation of oral pre-exposure prophylaxis among gay, bisexual, and other cisgender men who have sex with men and transgender women in Brazil, Mexico, and Peru (ImPrEP): a prospective, single-arm, open-label, multicentre implementation study. Lancet HIV 2023 Feb;10(2):e84-e96 [FREE Full text] [doi: 10.1016/S2352-3018(22)00331-9] [Medline: 36565708]
17. Serota DP, Rosenberg ES, Lockard AM, Rolle CP, Luisi N, Cutro S, et al. Beyond the biomedical: preexposure prophylaxis failures in a cohort of young black men who have sex with men in Atlanta, Georgia. Clin Infect Dis 2018 Aug 31;67(6):965-970 [FREE Full text] [doi: 10.1093/cid/ciy297] [Medline: 29635415]
18. Scott HM, Spinelli M, Vittinghoff E, Morehead-Gee A, Hirozawa A, James C, et al. Racial/ethnic and HIV risk category disparities in preexposure prophylaxis discontinuation among patients in publicly funded primary care clinics. AIDS 2019 Nov 15;33(14):2189-2195 [FREE Full text] [doi: 10.1097/QAD.0000000000002347] [Medline: 31436610]
19. Deutsch MB, Glidden DV, Sevelius J, Keatley J, McMahan V, Guanira J, iPrEx investigators. HIV pre-exposure prophylaxis in transgender women: a subgroup analysis of the iPrEx trial. Lancet HIV 2015 Dec;2(12):e512-e519 [FREE Full text] [doi: 10.1016/S2352-3018(15)00206-4] [Medline: 26614965]
20. Celum C, Hosek S, Tsholwana M, Kassim S, Mukaka S, Dye BJ, et al. PrEP uptake, persistence, adherence, and effect of retrospective drug level feedback on PrEP adherence among young women in southern Africa: results from HPTN 082, a randomized controlled trial. PLoS Med 2021 Jun;18(6):e1003670 [FREE Full text] [doi: 10.1371/journal.pmed.1003670] [Medline: 34143779]
21. Spinelli MA, Laborde N, Kinley P, Whitacre R, Scott HM, Walker N, et al. Missed opportunities to prevent HIV infections among pre-exposure prophylaxis users: a population-based mixed methods study, San Francisco, United States. J Int AIDS Soc 2020 Apr;23(4):e25472 [FREE Full text] [doi: 10.1002/jia2.25472] [Medline: 32294338]
22. Jin F, Amin J, Guy R, Vaccher S, Selvey C, Zablotska I, Expanded PrEP Implementation in Communities New South Wales (EPIC-NSW) research group. Adherence to daily HIV pre-exposure prophylaxis in a large-scale implementation study in New South Wales, Australia. AIDS 2021 Oct 01;35(12):1987-1996. [doi: 10.1097/QAD.0000000000002970] [Medline: 34101630]
23. Krakower D, Maloney KM, Powell VE, Levine K, Grasso C, Melbourne K, et al. Patterns and clinical consequences of discontinuing HIV preexposure prophylaxis during primary care. J Int AIDS Soc 2019 Feb;22(2):e25250 [FREE Full text] [doi: 10.1002/jia2.25250] [Medline: 30768762]

24. Coelho LE, Torres TS, Veloso VG, Landovitz RJ, Grinsztejn B. Pre-exposure prophylaxis 2.0: new drugs and technologies in the pipeline. Lancet HIV 2019 Nov;6(11):e788-e799. [doi: 10.1016/S2352-3018(19)30238-3] [Medline: 31558423]
25. Delany-Moretlwe S, Hughes JP, Bock P, Ouma SG, Hunidzarira P, Kalonji D, HPTN 084 study group. Cabotegravir for the prevention of HIV-1 in women: results from HPTN 084, a phase 3, randomised clinical trial. Lancet 2022 May 07;399(10337):1779-1789 [FREE Full text] [doi: 10.1016/S0140-6736(22)00538-4] [Medline: 35378077]
26. Landovitz RJ, Donnell D, Clement ME, Hanscom B, Cottle L, Coelho L, HPTN 083 Study Team. Cabotegravir for HIV prevention in cisgender men and transgender women. N Engl J Med 2021 Aug 12;385(7):595-608 [FREE Full text] [doi: 10.1056/NEJMoa2101016] [Medline: 34379922]
27. Guidelines on long-acting injectable cabotegravir for HIV prevention. World Health Organization. 2022 Jul 28. URL: https://www.who.int/publications/i/item/9789240054097 [accessed 2023-03-15]
28. Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2019 Oct;280:112513 [FREE Full text] [doi: 10.1016/j.psychres.2019.112513] [Medline: 31434011]
29. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
30. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019 Mar 29;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
31. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
32. Wooldridge AR, Carayon P, Hoonakker P, Hose BZ, Eithun B, Brazelton 3rd T, et al. Work system barriers and facilitators in inpatient care transitions of pediatric trauma patients. Appl Ergon 2020 May;85:103059 [FREE Full text] [doi: 10.1016/j.apergo.2020.103059] [Medline: 32174347]
33. Antonacci G, Reed JE, Lennox L, Barlow J. The use of process mapping in healthcare quality improvement projects. Health Serv Manage Res 2018 May;31(2):74-84 [FREE Full text] [doi: 10.1177/0951484818770411] [Medline: 29707978]
34. Trebble TM, Hansi N, Hydes T, Smith MA, Baker M. Process mapping the patient journey: an introduction. BMJ 2010 Aug 13;341:c4078. [doi: 10.1136/bmj.c4078] [Medline: 20709715]
35. de Aguiar Pereira CC, Torres TS, Luz PM, Hoagland B, Farias A, Brito JD, et al. Preferences for pre-exposure prophylaxis (PrEP) among men who have sex with men and transgender women at risk of HIV infection: a multicentre protocol for a discrete choice experiment in Brazil. BMJ Open 2021 Sep 27;11(9):e049011 [FREE Full text] [doi: 10.1136/bmjopen-2021-049011] [Medline: 34580095]
36. de Aguiar Pereira CC, Torres TS, Luz PM, Hoagland B, Farias A, Brito JD, et al. Preferences for pre-exposure prophylaxis (PrEP) among sexual and gender minorities: a discrete choice experiment in Brazil. Lancet Reg Health Am 2023 Mar;19:100432 [FREE Full text] [doi: 10.1016/j.lana.2023.100432]
37. Torres TS, Nascimento AR, Coelho LE, Konda KA, Vega-Ramirez EH, Elorreaga OA, et al. Preferences for PrEP modalities among gay, bisexual, and other men who have sex with men from Brazil, Mexico, and Peru: a cross-sectional study. Ther Adv Infect Dis 2023 Feb 18;10:20499361231153548 [FREE Full text] [doi: 10.1177/20499361231153548] [Medline: 36814515]
38. Gray AL, Smit JA, Manzini N, Beksinska M. Systematic review of contraceptive medicines “Does choice make a difference?”. Reproductive Health and HIV Research Unit. 2006 Oct. URL: https://toolkits.knowledgesuccess.org/sites/default/files/311684.pdf [accessed 2023-03-15]
39. Bezerra DR, Jalil CM, Jalil EM, Coelho LE, Carvalheira E, Freitas J, et al. Complementary recruitment strategies to reach men who have sex with men and transgender women: the experience of a large Brazilian HIV prevention service. AIDS Behav 2022 Aug;26(8):2643-2652. [doi: 10.1007/s10461-022-03609-5] [Medline: 35122578]
40. Bezerra DR, Jalil CM, Jalil EM, Coelho LE, Netto EC, Freitas J, et al. Comparing web-based venues to recruit gay, bisexual, and other cisgender men who have sex with men to a large HIV prevention service in Brazil: evaluation study. JMIR Form Res 2022 Aug 04;6(8):e33309 [FREE Full text] [doi: 10.2196/33309] [Medline: 35925658]
41. Teixeira SL, Jalil CM, Jalil EM, Nazer SC, da Costs Cruz Silva S, Veloso VG, et al. Evidence of an untamed HIV epidemic among MSM and TGW in Rio de Janeiro, Brazil: a 2018 to 2020 cross-sectional study using recent infection testing. J Int AIDS Soc 2021 Jun;24(6):e25743 [FREE Full text] [doi: 10.1002/jia2.25743] [Medline: 34132470]

Abbreviations

ART: antiretroviral therapy
BMOH: Brazilian Ministry of Health
CAB-LA: long-acting injectable cabotegravir
FGD: focus group discussion
Fiocruz: Fundação Oswaldo Cruz

FTC: emtricitabine
HPTN: HIV Prevention Trials Network
INI: Instituto Nacional de Infectologia Evandro Chagas
mHealth: mobile health
MSM: gay men, bisexual men, and other men who have sex with men
PrEP: preexposure prophylaxis
SGM: sexual and gender minority
SOC: standard of care
STI: sexually transmitted infection
SUS: Sistema Único de Saúde
TDF: tenofovir disoproxil fumarate
TFA: theoretical framework for acceptability
TGW: travesti and transgender women
WHO: World Health Organization

Edited by A Mavragani; submitted 10.12.22; peer-reviewed by M Spinelli, Z Ni; comments to author 03.01.23; revised version received 19.01.23; accepted 31.01.23; published 19.04.23.

Please cite as:
Grinsztejn B, Torres TS, Hoagland B, Jalil EM, Moreira RI, O'Malley G, Shade SB, Benedetti MR, Moreira J, Simpson K, Pimenta MC, Veloso VG, The ImPrEP CAB-Brasil Study Team
Long-Acting Injectable Cabotegravir for HIV Preexposure Prophylaxis Among Sexual and Gender Minorities: Protocol for an Implementation Study
JMIR Public Health Surveill 2023;9:e44961
URL: https://publichealth.jmir.org/2023/1/e44961
doi:10.2196/44961
PMID:37074775

Review

Reporting, Monitoring, and Handling of Adverse Drug Reactions in Australia: Scoping Review

Joel Fossouo Tagne[1,2,3], BBiomed, Mpharm, GC-RESIM; Reginald Amin Yakob[4], BBiomed, Mpharm, AACPA; Thu Ha Dang[5,6], MIPH, MD; Rachael Mcdonald[3,7,8], BAppSc, PGDip(Biomech), GCHE, PhD; Nilmini Wickramasinghe[1,8,9], BSc, Grad DipMgtSt, MBA, PhD

[1]Department of Health and Biostatistics, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia
[2]Centre for Health Analytics, Murdoch Children's Research Institute, Melbourne, Australia
[3]MedTechVic, Swinburne University of Technology, Melbourne, Australia
[4]Australian Association of Consultant Pharmacy, Sydney, Australia
[5]Department of Psychology, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia
[6]Behavioural Sciences Unit, Department of Health Services Research, Peter MacCallum Cancer Centre, Melbourne, Australia
[7]Department of Nursing and Allied Health, Occupational Therapy, Swinburne University of Technology, Melbourne, Australia
[8]Iverson Health Innovation Research Institute, Swinburne University of Technology, Melbourne, Australia
[9]Epworth Healthcare, Melbourne, Australia

Corresponding Author:
Joel Fossouo Tagne, BBiomed, Mpharm, GC-RESIM
Department of Health and Biostatistics, School of Health Sciences
Swinburne University of Technology
John Street
Hawthorn
Melbourne, 3122
Australia
Phone: 61 0412478610
Email: jfossouotagne@swin.edu.au

Abstract

Background: Adverse drug reactions (ADRs) are unintended consequences of medication use and may result in hospitalizations or deaths. Timely reporting of ADRs to regulators is essential for drug monitoring, research, and maintaining patient safety, but it has not been standardized in Australia.

Objective: We sought to explore the ways that ADRs are monitored or reported in Australia. We reviewed how consumers and health care professionals participate in ADR monitoring and reporting.

Methods: The Arksey and O'Malley framework provided a methodology to sort the data according to key themes and issues. Web of Science, Scopus, Embase, PubMed, CINAHL, and Computer & Applied Sciences Complete databases were used to extract articles published from 2010 to 2021. Two reviewers screened the papers for eligibility, extracted key data, and provided descriptive analysis of the data.

Results: Seven articles met the inclusion criteria. The Adverse Medicine Events Line (telephone reporting service) was introduced in 2003 to support consumer reporting of ADRs; however, only 10.4% of consumers were aware of ADR reporting schemes. Consumers who experience side effects were more likely to report ADRs to their doctors or pharmacists than to the drug manufacturer. The documentation of ADR reports in hospital electronic health records showed that nurses and pharmacists were significantly less likely than doctors to omit the description of the drug reaction, and pharmacists were significantly more likely to enter the correct classification of the drug reaction than doctors. Review and analysis of all ADR reports submitted to the Therapeutic Goods Administration highlighted a decline in physician contribution from 28% of ADR reporting in 2003 to 4% in 2016; however, within this same time period, hospital and community pharmacists were a major source of ADR reporting (ie, 16%). In 2014, there was an increase in ADR reporting by community pharmacists following the introduction of the GuildLink ADR web-based reporting system; however, a year later, the reporting levels dropped. In 2018, the Therapeutic Goods Administration introduced a black triangle scheme on the packaging of newly approved medicines, to remind and encourage ADR reporting on new medicines, but this was only marginally successful at increasing the quantity of ADR reports.

Conclusions: Despite the existence of national and international guidelines for ADR reporting and management, there is substantial interinstitutional variability in the standards of ADR reporting among individual health care facilities. There is room for increased ADR reporting rates among consumers and health care professionals. A thorough assessment of the barriers and enablers to ADR reporting at the primary health care institutional levels is essential. Interventions to increase ADR reporting, for example, the black triangle scheme (alert or awareness) or GuildLink (digital health), have only had marginal effects and may benefit from further improvement revisions and awareness programs.

(JMIR Public Health Surveill 2023;9:e40080) doi:10.2196/40080

KEYWORDS

pharmacovigilance; adverse drug reactions; primary care; digital health

Introduction

The World Health Organization defines pharmacovigilance (PV) as “the science and activities related to the detection, assessment, understanding and prevention of adverse effects of medications” [1]. Many harmful adverse drug reactions (ADRs) resulting from medication use go undetected or unreported to regulatory authorities [2]. The underreporting of ADRs remains a major threat to patient safety and is a substantial burden to established health care systems [2,3]. Most PV schemes are based on spontaneous and voluntary reporting [4]. In Australia, the Therapeutic Goods Administration (TGA) receive suspected ADRs reports from health care professionals (HCPs; physicians, pharmacists, nurses, etc) and the general public (consumers, patients, carers, members of the legal system, etc) [5,6]. The TGA regularly reviews available information originating from submitted ADR reports, and this informs drug safety decisions [4]. International studies have stated that less than 5% of ADRs are reported, including in jurisdictions where ADR reporting is mandatory [7]. Without a robust PV systems, ADRs may remain undetected for years, exposing patients to unanticipated health risks, and is a detriment to the health care system and taxpayers [8]. For example, the anti-inflammatory drug rofecoxib (Vioxx) was withdrawn from the market for high risk of myocardial infarctions after population exposures had reached millions, emphasizing the need for early detection of drug safety signals to ensure global health safety[7].

In Australia, approximately 400,000 consumers present to hospital emergency departments each year with medication-related problems [3,9]. Furthermore, about 7.2% to 11% of hospital admissions are specifically related to ADRs, of which approximately 50% are preventable [9,10]. Globally, it has been reported that 3.6% to 15.6% of hospital admissions are related to ADRs [11,12]. Complications related to ADRs may increase the mean length of hospital stay from 8 to 20 days [13]. According to the Pharmaceutical Society of Australia’s medication safety report (2019), health care expenditure for medication-related problems is estimated at AUD $1.4 billion (US $900,207) per annum [14]. Despite its limitations, spontaneous reporting remains the most common method for generating safety signals; however, it is estimated that only 6% of all ADRs that occur are reported [7,15,16]. The reasons for ADR underreporting have previously been described [10,17-19]; these may include limited knowledge and awareness of PV, reluctance to report by consumers or HCPs due to time constraints, and nonsupportive workplace structures. Although various interventions to improve ADR reporting have been implemented [20], ADR underreporting remains a limitation of the current PV system [10,21].

Previous studies have evaluated the effects of interventions for improving ADR reporting rates; however, the perspective of ADR reporting and handling within the Australian health care system remains largely unexplored [20-22]. There remains little integrative and collective knowledge on ADR reporting, monitoring, and handling in the Australian health care context [21]. This scoping review aimed to provide a comprehensive landscape of PV and ADR reporting, monitoring, and handling in Australia. This review may stimulate further research, policy makers, regulatory authorities, or software vendors to make decisions that may promote ADR reporting and improve patient safety within the Australian context.

Methods

Overview

The Arksey and O’Malley methodology framework was adopted for conducting this scoping review [23,24]. This framework uses a rigorous process of transparency that enables the replication of the search strategy, which increases the reliability of the study findings. The framework has six stages as follows.

Stage 1: Identifying the Research Question

The focus of this scoping review was to explore ADR reporting in the Australian health care system and the extent of participation by consumers or HCPs and to describe the different reporting systems. The core questions are (1) How are ADRs monitored, reported, and handled in the Australian health care system? and (2) What is the extent of participation by consumers or HCPs?

Stage 2: Identifying the Relevant Studies

The search strategy was developed by authors JFT and RAY. Six scientific databases were searched (Web of Science; Scopus; Embase; PubMed, including MEDLINE; CINAHL; and Computer & Applied Science Complete) for research articles published in English between 2010 to 2021. Furthermore, a reference list search (ie, backward reference search) and a cited reference search (ie, forward reference search) were carried out based on the full-text papers that met the study selection criteria. The reference search was repeated on newly identified papers until no additional relevant papers could be found. Multimedia Appendix 1 presents the full search strategy for each database.

Stage 3: Selecting Studies

The inclusion and exclusion criteria are presented in Textbox 1.

Five authors (JFT, RAY, THD, RM, and NW) independently screened the titles and abstracts of the studies to identify the ones for potential inclusion. Disagreements about exclusions were discussed until a consensus was reached. The full texts of the articles were then independently reviewed by 2 authors (JFT and RAY), who decided whether or not they met the selection criteria. In the case of disagreement, a third author (NW) made the final decision.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Adverse drug reactions (ADRs) in the Australian health care system (medication management context)
- Cross-sectional or survey-based studies

Exclusion criteria

- Not related to drugs
- Not related to ADR reporting, monitoring, or handling (eg, efficacy or effectiveness of a study's design)
- Not related to patients' reporting (eg, a safety study in animals)
- Qualitative studies, review manuscripts, editorials, letters, and news

Stage 4: Charting the Data

Two reviewers (JFT and RAY) identified characteristic data sets to describe the articles, which included the author(s), year of publication, setting (where the study was conducted, eg, hospital and community pharmacy), study population (eg, physicians), gender, age, intervention, and reporting rates. The data were represented in a logical and descriptive manner in line with the objectives of this review. As per guidelines, scoping reviews do not exclude research based on the research quality but rather identify areas that are lacking in research [23,24]. Therefore, methodological quality or risk of bias of the included articles were not needed [23,24].

Stage 5: Collating, Summarizing, and Reporting Results

The information extracted were summarized, organized, and discussed. A PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flow chart was developed to summarize the identified literature and study selection steps. The narrative and a numerical summary of the characteristics of the included studies were tabulated. In the conceptual analysis, the study aimed to describe 2 main themes: (1) how ADRs are monitored, handled, or reported in the Australian health care system and (2) the extent of participation by stakeholders. The results of the first theme are presented in Multimedia Appendix 2 as the type of reporting or intervention, PV system used, and the focus of the surveillance. The results of the second theme are presented by the study population, setting, and reporting rates.

Stage 6: Consultation Exercise

The last stage was to engage in a consultation exercise with stakeholders to discuss, inform, and validate the results of the study [23,24]. This scoping review was interdisciplinary [23,24]; the overall expertise covered medicine, clinical pharmacy, PV, epidemiology, public health, health informatics, and statistics. The participation of all the authors throughout the research process served the purpose of constant consultation and provided the rigor and strength to the study findings and conclusions.

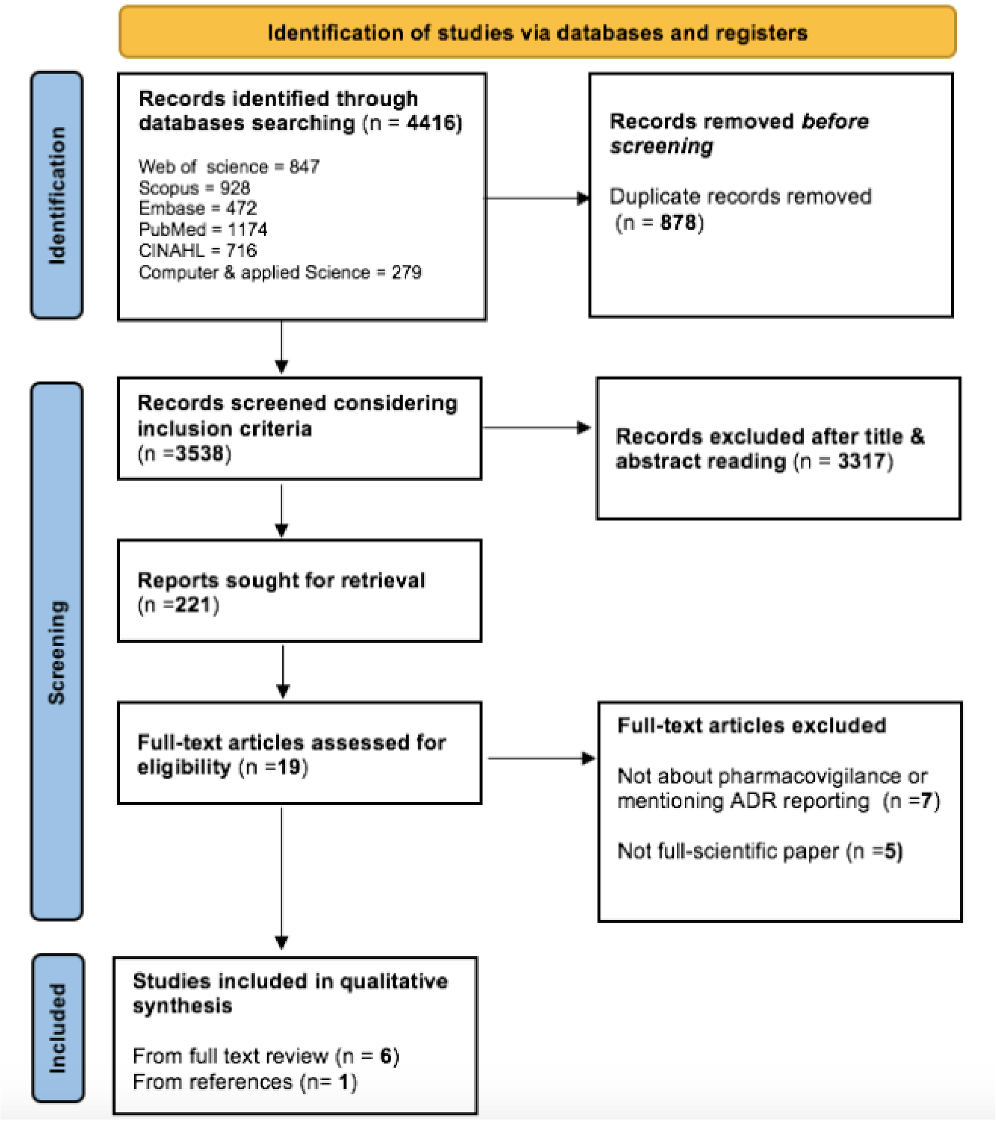
Results

Overview

Figure 1 (PRISMA flow chart) depicts the full review process and shows the number of citations excluded at each step. A total of 3538 citations were retrieved from the scientific electronic databases after duplicates (n=878) were removed. We screened the titles, abstracts, and full texts using the exclusion criteria and resulted with 6 citations to be included. The reference lists of the selected studies were also manually searched, and 1 article was identified; in total, 7 citations were relevant to the research question and included in the flow chart (stage 3). Multimedia Appendix 3 presents a list of the included studies.

The selected studies were published between 2013 and 2021. Among the studies, 2 focused on consumers and patients, 2 on health care providers, and 3 on information systems. In all, 5 of the 7 studies were in the hospital settings, and the rest were in the community. The findings of the selected studies are presented in 2 themes that relates back to the research questions, "ADR reporting, monitoring, and handling in the Australian health care context" and "participation." The characteristics of the studies' findings are summarized in Multimedia Appendix 2.

Figure 1. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flowchart of study selection and inclusion criteria. ADR: adverse drug reaction.



ADR Reporting, Monitoring, and Handling in Australia

Regulatory Monitoring of ADR

The Australian National Safety and Quality Health Service provides statements for accreditation and include standards for ADR management. Health services are required to have processes for documenting medication allergies and ADRs within the health care records and in the organization-wide incident-reporting system, as well as for submitting major or rare ADRs to the TGA. Providing written ADR information to both the patient and their general practitioner (GP) is a requirement in Australian hospitals [25]. However, despite national and international guidelines for ADR reporting and management, there is no established national protocol for reporting ADRs at the individual health care facility level [26]. As such, there exists substantial interinstitutional variability with respect to the timing and nature of ADR reports [26]. The responsibility for the timely collection and reporting of drug safety information mainly rests with the marketing authorization holder due to mandatory reporting requirements from regulatory agencies [27,28].

ADR Handling

An Adverse Drug Reaction Review Committee (ADRRC) is an internal hospital committee comprising of a multidisciplinary team that includes a senior pharmacist and a specialist clinician from at least one of dermatology, immunology, clinical pharmacology, infectious diseases, and general medicine [25,26]. The committee meets every 2 weeks to review all ADR reports; verify diagnoses, if required; organize allergic clinical referrals; and provide further risk mitigation measures through written recommendations to the clinicians involved as well as the patients and carers. Relevant ADR reports are forwarded to the national database at the TGA [26]. In the study of patient satisfaction with an ADR warning card at a tertiary hospital during the period between January 2013 and April 2016 (n=241), patients with suspected ADRs received a wallet-sized paper card titled “Temporary ADR warning card” with the name of the drug thought to be the causative agent. The ADR pharmacist collated the supporting evidence for the case including notes from the inpatient stay and information from the Australian

Database of Adverse Event Notifications. The collated evidence was sent to the ADRRC clinicians who reviewed the case and determined an outcome. If the ADR was confirmed, the patient was sent a letter containing the details of the reaction and a laminated card titled “Permanent ADR warning card,” and a copy of the letter was sent to the patient’s GP [25].

ADR Reporting Schemes

The Adverse Medicine Events Line is a telephone reporting service that was introduced in Australia in 2003 to allow consumers to report suspected ADRs to the TGA and receive advice about side effects [27]. In June 2014, a pharmacy software vendor, GuildLink, created a web-based Adverse Events Recording module that integrated directly into the TGA ADR web service. This integration allows community pharmacists to report ADRs directly to the TGA from their professional service program instead of having to manually complete a separate ADR reporting form. In 2018, the TGA introduced a black triangle scheme on the packaging of newly approved medicines to alert HCPs and consumers, serving as a prompt to report any ADRs associated with that medicine [28]

Participants and Extent of Involvement

Patient and Consumer

Direct consumer and patient reporting of ADRs to the TGA has been possible since 1964, and since 1990, there had been fewer than 7000 consumer ADR reports [27]. In 2009, there was a spike in consumer ADR reporting (1307/13,298, 9.8%) that was associated to influenza vaccines due to the H1N1 pandemic; however, consumer reporting fell to 3% in 2011 [27]. A study of consumer awareness to ADR reporting between September and October 2012 used 2 methods to survey respondents: a computer-assisted telephone interview (n=2484) and a Pureprofile web-based survey (n=2497). From the 2 groups combined, 46.3% of consumers indicated that side effects of prescription medicines were very common (88.4%) [27]. The awareness of consumer reporting schemes among the whole study group was low (10.4%) [27]. Among 217 respondents who had experienced a side effect and were aware of consumer reporting schemes, 46 (21.2%) reported the ADR, using one of the reporting schemes [27]. Consumers were more likely to report ADRs to their doctors or pharmacists than to the pharmaceutical industry [27,28]. Among the consumer who had experienced a side effect, 84.6% reported the event to an HCP, most often a GP [27]. In the study of patient satisfaction with an ADR warning card at a tertiary hospital during the period between January 2013 and April 2016 (n=241), 85% notified their doctor, 67% notified their family, and 40% notified their community pharmacist [25]. The likelihood of the causative agent being available in community pharmacies determined if participants were more likely to consider it necessary to inform their regular community pharmacist of the new ADR (P=.001) [25]. The majority of ADRs experienced by participants fell within 3 drug classes: antibiotics (32%), iron infusions (14%), and contrast media (13%) [25].

HCPs’ Participation

In the review and analysis of all ADR reports submitted to the TGA, 28% out of 10,981 reports received in 2003 and about 7% out of the 14,400 reports received in 2011 were from GPs [27]. In 2016, the TGA national PV data also highlighted that hospital and community pharmacists were a major source of ADR reporting, at a much higher rate than doctors (physicians; ie, 16% for pharmacists vs 4% for physicians) [26]. In fact, in their submission for the 2015 TGA review of Australian Medicines and Medical Devices regulations, the Consumers Health Forum argued for mandatory requirements for doctors and pharmacists to report ADRs [28]. However, the Royal Australasian College of Physicians expressed the need for Australian physicians to receive payment for completing ADR reports [27]. In the review of voluntary ADR reports by HCPs to an ADRRC from 2012-2016, of the 555 ADR reports, 471 (84.8%) were reported by hospital pharmacists, 52 (9.4%) by doctors, and 32 (5.7%) by other HCPs. The median time from the date of onset of an ADR to submission of an ADR report (ie, to the TGA) was 3 days [26]. In the study of knowledge and perspectives of ADR reporting by community pharmacists between January to February 2017 (n=263), 35.3% (n=82) of community pharmacists reported at least one ADR to the TGA in the previous 12 months, even though 88.4% (n=205) of the pharmacists encountered an ADR in a patient, and 65.9% (n=153) documented ADRs as part of a clinical intervention at least once a month [28]. The documentation of opioid and penicillin ADR reports in hospital electronic health records showed that nurses and pharmacists were significantly less likely than doctors to omit the reaction description and that pharmacists were significantly more likely to enter the correct classification than doctors (53%, 95% CI 50.52-55.47; P<.001) [29,30].

Discussion

To the best of our knowledge, this is the first scoping review to integrate and synthesize the available published and scientific literature on ADR reporting within the Australian health care context. Thus, this study provides a broad overview of the Australian PV landscape and may highlight potential areas for innovation or quality improvement.

Consumers Participation in ADR Reporting

The results of our review indicates that consumer reporting is low, as reported in other studies [27]; despite this, consumers who are aware of reporting schemes may be willing to report. The studies showed that consumers were generally more likely to report ADRs to doctors (physicians) or pharmacists [27,28]. The findings are consistent with a recent study reporting that community pharmacists are usually the first point of contact regarding medication-related issues as the most frequently visited HCPs in Australia [10]. Introducing measures to facilitate ADR reporting for both consumer and patients and HCPs can be beneficial. A 2020 systematic review of interventions to improve ADR reporting also concluded that there was scope to include community pharmacies to improve ADR reporting [22]. These findings were also reported in other reviews [20,21]. Robertson and Newby [27] published in 2013 that only 10.14% of consumers (respondents) were aware of ADR reporting schemes. In 2020, Li et al [31] reported the findings of a black triangle scheme (ADR awareness) introduced by the TGA in

2018 to increase awareness and promote the reporting of ADRs by consumers and HCPs [31]. The effect of the black triangle scheme was only marginally successful at increasing ADR reports, that is, there was an improvement in the overall quality of ADR reports submitted but no meaningful increase in the quantity of reports [31]. The study concluded that additional strategies were required to enhance the overall PV system in Australia [28,31]. Considering that there has only been a marginal change in ADR reporting since 2013 [27], it suggests that there is opportunity for regulators to provide more awareness, such as education and further research. The provision of feedback to consumers about their ADR reports and involving consumers in the ADR management process, for example, the provision of ADR Warning cards, resulted in a high level of patient satisfaction and may have positively contributed to patient awareness [25]. In addition, feedback to health care providers on their submitted ADR reports may also encourage further reporting [28].

Health Care Provider Participation in ADR Reporting

Consumers who experienced a side effect reported the event to an HCP, most often a GP (physician) [27]. However, the review and analysis of all ADR reports submitted to the TGA highlighted a decline in GPs' contribution, from 28% of ADR reporting in 2003 to 4% in 2016 [26,27]. Therefore, understanding the barriers to and facilitators of ADR reporting among GPs in Australia may inform future interventions. Furthermore, in 2016, the TGA data also highlighted that hospital pharmacists and community pharmacists were a major source of ADR reporting, at a much higher rate than doctors (16% from pharmacists vs 4% from GPs) [26]. A Canadian study exploring why clinicians do not report ADRs posited that previous studies had focused predominantly on the knowledge and attitudes of HCPs and framed underreporting as a failure of individual clinicians without investigating workplace structures or practices that may have influenced reporting [18]. The discrepancies in reporting between GPs and pharmacists are vital and may warrant further investigation into organization and workplace structures. This is particularly important given our results in the documentation of opioid and penicillin ADR reports in a hospital electronic health records that showed that nurses and pharmacists were significantly less likely than doctors to omit the reaction description and that pharmacists were significantly more likely to enter the correct classification than doctors [29,30]. It is important to note that the study did not mention the reasons for reporting differences, leaving room for further investigations. Conversely, despite nurses being less likely to omit the reaction description, their perspective and reporting rate were not discussed and would benefit from further exploration. Nevertheless, although new technologies, for example, "electronic health records," present as an opportunity to facilitate ADR reporting [21], it has been noted that in practice, these electronic systems can cause unexpected errors, desensitize clinicians to alerts, and increase the documentation burden [19]. This is evident in our findings, where in June 2014, a pharmacy software vendor, GuildLink, created a web-based Adverse Events Recording module that integrated directly into the TGA ADR web service [28]. This allowed community pharmacists to report ADRs directly to the TGA from their professional service program instead of having to manually complete a separate ADR reporting form [28]. In the follow-up from the GuildLink intervention, the quantity of reports received by the TGA was nearly as high as that of the previous year (2013); however, despite the positive start, the numbers declined again in 2015 [28]. The factors that may have influenced a sustained adoption of the GuildLink system were not mentioned or discussed. As new digital technologies emerge and continue to transform health care management [32], exploring existing technology and their shortcomings, for example, GuildLink, and evaluating their health impact present areas of opportunity for improving health quality and patient safety.

Repeated calls for mandated ADR reporting for HCPs have been made, and as discussed above, patients are more likely to discuss or report issues with their medications to the HCPs who initially dispensed or prescribed them (GPs or pharmacists). The Consumers Health Forum in their 2015 submission argued for mandatory requirements for doctors and pharmacists to report ADRs [23,28]. However, the Royal Australasian College of Physicians also expressed the need for Australian physicians to receive payment for completing ADR reports [27]. Furthermore, one of the studies in our results suggested financial incentives for ADR reporting was also highly regarded as a measure to improve reporting by pharmacist [28]. Considering these points, none of the studies in our results explored or noted the impact of mandatory reporting or incentives, for example, financial, on ADR reporting. Furthermore, the finding of the review also showed that despite national and international guidelines for ADR reporting and management, there is no established gold standard systems (policy or guideline) at the individual health care facility level in Australia [26]. Therefore, there exists substantial interinstitutional variability with respect to the timing and nature of reporting ADRs [26]. It was also not possible to pinpoint the net effect of interinstitutional variability in ADR reporting to the TGA, which requires further investigation.

Strength and Limitation of the Review

The strength of this scoping review is that it is the first review to attempt an integration and synthesis of the available literature on ADR reporting, monitoring, and handling in the Australian health care context; the extent of participation; and ADR reporting systems. Notwithstanding the value of this research, some limitations must be acknowledged. This review was limited to papers published in English; it is possible that other potentially relevant literature in other languages were omitted. Although this review was performed using multiple major databases, searching other databases may have yielded other relevant published papers. A quality assessment of the studies included in the review was not undertaken; however, it was not relevant for the aim and not always necessary for scoping reviews [23,24]. There was limited data from the included studies in regard to the variabilities in participants' age, gender, or educational status. The majority of the studies were undertaken in the hospital setting and predominately focused on pharmacists and physicians; therefore, applying these results to the entire Australian health care system may require additional considerations, for example, the impact of age, gender, task, or work setting. Different states, regions, and countries may differ in their infrastructures, capacities, general workplace culture, medical education programs, and economic status, which limits the generalizability of the findings of this review.

Conclusion

The exploratory nature of the scoping review helped to integrate and synthesize the literature to give us a broad overview of the PV and ADR-reporting landscape in Australia. This review is beneficial for identifying research gaps and serves as a vital step to identify key areas for digital interventions and quality improvement in ADR reporting. It is also hoped that the results of this scoping review send a message to regulators (eg, the TGA), software vendors, or policy makers concerning the work that is needed to develop a robust and standardized PV system in Australia. This review did not find a generally accepted "gold standard" protocol or framework for ADR reporting among individual health care facilities, and there is substantial interinstitutional variability with respect to the timing and nature of ADR reporting in Australia.

Despite the marginal effects from previous interventions aiming to enhance ADR reporting, for example, ADR awareness (black triangle scheme) or digital health interventions (GuildLink), there is opportunity to improve patient safety in relation to ADRs. Further evidence-based research is needed to guide ADR interventions design and implementation. To support these interventions, knowledge of the barriers and enablers to ADR reporting among consumers, pharmacists, physicians, and nurses at a primary health care institutional level is warranted.

Acknowledgments

This study was supported by the Department of Health and Biostatistics, School of Health Sciences, and the Iverson Health Innovation Research Institute, Swinburne University of Technology, Melbourne, Australia

Conflicts of Interest

None declared.

Multimedia Appendix 1

Full search strategy for each database.

[PNG File , 213 KB - publichealth_v9i1e40080_app1.png]

Multimedia Appendix 2

Characteristics of the study findings on adverse drug reactions (ADRs) reporting, monitoring, and handling.

[PNG File , 232 KB - publichealth_v9i1e40080_app2.png]

Multimedia Appendix 3

List of studies selected for data analysis and synthesis in this scoping review.

[PNG File , 167 KB - publichealth_v9i1e40080_app3.png]

References

1. Lardon J, Abdellaoui R, Bellet F, Asfari H, Souvignet J, Texier N, et al. Adverse drug reaction identification and extraction in social media: a scoping review. J Med Internet Res 2015 Jul 10;17(7):e171 [FREE Full text] [doi: 10.2196/jmir.4304] [Medline: 26163365]
2. Ribeiro-Vaz I, Silva A, Costa Santos C, Cruz-Correia R. How to promote adverse drug reaction reports using information systems - a systematic review and meta-analysis. BMC Med Inform Decis Mak 2016 Mar 01;16:27 [FREE Full text] [doi: 10.1186/s12911-016-0265-8] [Medline: 26926375]
3. Phillips AL, Nigro O, Macolino KA, Scarborough KC, Doecke CJ, Angley MT, et al. Hospital admissions caused by adverse drug events: an Australian prospective study. Aust Health Rev 2014 Feb;38(1):51-57. [doi: 10.1071/AH12027] [Medline: 24351707]
4. Fossouo J, McDonald R, Wickramasinghe N. Leveraging information technology in pharmacovigilance: benefits for pharmacists and pharmaceutical companies. In: Wickramasinghe N, Chalasani S, Sloane E, editors. Digital Disruption in Healthcare. Cham, Switzerland: Springer International Publishing; Jun 24, 2022:187-207.
5. Dong V, Sharan S. Pharmacovigilance - a regulator's perspective. Australian Government, Department of Health - Therapeutic Goods Administration. 2018 Aug 15. URL: https://www.tga.gov.au/sites/default/files/pharmacovigilance-regulator-perspective.pdf [accessed 2021-08-25]
6. Pedrós C, Vallano A, Cereza G, Mendoza-Aran G, Agustí A, Aguilera C, et al. An intervention to improve spontaneous adverse drug reaction reporting by hospital physicians: a time series analysis in Spain. Drug Saf 2009;32(1):77-83. [doi: 10.2165/00002018-200932010-00007] [Medline: 19132807]

7. Bailey C, Peddie D, Wickham ME, Badke K, Small SS, Doyle-Waters MM, et al. Adverse drug event reporting systems: a systematic review. Br J Clin Pharmacol 2016 Jul;82(1):17-29 [FREE Full text] [doi: 10.1111/bcp.12944] [Medline: 27016266]
8. Peddie D, Small SS, Badke K, Wickham ME, Bailey C, Chruscicki A, et al. Designing an adverse drug event reporting system to prevent unintentional reexposures to harmful drugs: study protocol for a multiple methods design. JMIR Res Protoc 2016 Aug 18;5(3):e169 [FREE Full text] [doi: 10.2196/resprot.5967] [Medline: 27538362]
9. Parameswaran Nair N, Chalmers L, Bereznicki BJ, Curtain C, Peterson GM, Connolly M, et al. Adverse drug reaction-related hospitalizations in elderly Australians: a prospective cross-sectional study in two Tasmanian hospitals. Drug Saf 2017 Jul 6;40(7):597-606. [doi: 10.1007/s40264-017-0528-z] [Medline: 28382494]
10. Fossouo Tagne J, Yakob RA, Mcdonald R, Wickramasinghe N. Barriers and facilitators influencing real-time and digital-based reporting of adverse drug reactions by community pharmacists: qualitative study using the Task-Technology Fit Framework. Interact J Med Res 2022 Oct 11;11(2):e40597 [FREE Full text] [doi: 10.2196/40597] [Medline: 36222800]
11. Bouvy JC, de Bruin ML, Koopmanschap MA. Epidemiology of adverse drug reactions in Europe: a review of recent observational studies. Drug Saf 2015 May 31;38(5):437-453 [FREE Full text] [doi: 10.1007/s40264-015-0281-0] [Medline: 25822400]
12. Ayalew MB, Tegegn HG, Abdela O. Drug related hospital admissions; a systematic review of the recent literatures. Bull Emerg Trauma 2019 Oct 05;7(4):339-346 [FREE Full text] [doi: 10.29252/beat-070401] [Medline: 31857995]
13. Davies EC, Green CF, Taylor S, Williamson PR, Mottram DR, Pirmohamed M. Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. PLoS One 2009;4(2):e4439 [FREE Full text] [doi: 10.1371/journal.pone.0004439] [Medline: 19209224]
14. Medicine safety: take care. Pharmaceutical Society of Australia. 2019 Jan. URL: https://www.psa.org.au/wp-content/uploads/2019/01/PSA-Medicine-Safety-Report.pdf [accessed 2023-01-06]
15. Parracha ER, Advinha AM, Lopes MJ, Oliveira-Martins S. Mobile apps for quick adverse drug reaction report: a scoping review. Pharmacoepidemiol Drug Saf 2023 Jan;32(1):19-27 [FREE Full text] [doi: 10.1002/pds.5542] [Medline: 36125022]
16. de Rosa M, Fenza G, Gallo A, Gallo M, Loia V. Pharmacovigilance in the era of social media: discovering adverse drug events cross-relating Twitter and PubMed. Future Gener Comput Syst 2021 Jan;114:394-402. [doi: 10.1016/j.future.2020.08.020]
17. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. Drug Saf 2009;32(1):19-31. [doi: 10.2165/00002018-200932010-00002] [Medline: 19132802]
18. Hohl CM, Small SS, Peddie D, Badke K, Bailey C, Balka E. Why clinicians don't report adverse drug events: qualitative study. JMIR Public Health Surveill 2018 Feb 27;4(1):e21 [FREE Full text] [doi: 10.2196/publichealth.9282] [Medline: 29487041]
19. Peddie D, Small SS, Badke K, Bailey C, Balka E, Hohl CM. Adverse drug event reporting from clinical care: mixed-methods analysis for a minimum required dataset. JMIR Med Inform 2018 Jun 28;6(2):e10248 [FREE Full text] [doi: 10.2196/10248] [Medline: 29954724]
20. Li R, Zaidi STR, Chen T, Castelino R. Effectiveness of interventions to improve adverse drug reaction reporting by healthcare professionals over the last decade: a systematic review. Pharmacoepidemiol Drug Saf 2020 Jan;29(1):1-8 [FREE Full text] [doi: 10.1002/pds.4906] [Medline: 31724270]
21. Li R, Curtis K, Zaidi ST, Van C, Castelino R. A new paradigm in adverse drug reaction reporting: consolidating the evidence for an intervention to improve reporting. Expert Opin Drug Saf 2022 Sep;21(9):1193-1204. [doi: 10.1080/14740338.2022.2118712] [Medline: 36031811]
22. Paudyal V, Al-Hamid A, Bowen M, Hadi MA, Hasan SS, Jalal Z, et al. Interventions to improve spontaneous adverse drug reaction reporting by healthcare professionals and patients: systematic review and meta-analysis. Expert Opin Drug Saf 2020 Sep 03;19(9):1173-1191. [doi: 10.1080/14740338.2020.1807003] [Medline: 32755492]
23. Daudt HML, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013 Mar 23;13:48 [FREE Full text] [doi: 10.1186/1471-2288-13-48] [Medline: 23522333]
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
25. Scott MW, McDonough M, Wong GO, Summers R, Spark MJ. Patient satisfaction with a hospital adverse drug reaction reporting system. J Pharm Pract Res 2020 Apr 13;50(2):152-159. [doi: 10.1002/jppr.1635]
26. Aung AK, Tang MJ, Adler NR, de Menezes SL, Goh MSY, Tee HW, et al. Adverse drug reactions reported by healthcare professionals: reaction characteristics and time to reporting. J Clin Pharmacol 2018 Oct 07;58(10):1332-1339. [doi: 10.1002/jcph.1148] [Medline: 29733431]
27. Robertson J, Newby DA. Low awareness of adverse drug reaction reporting systems: a consumer survey. Med J Aust 2013 Nov 18;199(10):684-686. [doi: 10.5694/mja13.10069] [Medline: 24237099]
28. Li R, Curtain C, Bereznicki L, Zaidi STR. Community pharmacists' knowledge and perspectives of reporting adverse drug reactions in Australia: a cross-sectional survey. Int J Clin Pharm 2018 Aug 10;40(4):878-889 [FREE Full text] [doi: 10.1007/s11096-018-0700-2] [Medline: 30097819]

29. Inglis JM, Caughey GE, Smith W, Shakib S. Documentation of adverse drug reactions to opioids in an electronic health record. Intern Med J 2021 Sep;51(9):1490-1496. [doi: 10.1111/imj.15209] [Medline: 33465262]
30. Inglis JM, Caughey GE, Smith W, Shakib S. Documentation of penicillin adverse drug reactions in electronic health records: inconsistent use of allergy and intolerance labels. Intern Med J 2017 Nov 05;47(11):1292-1297. [doi: 10.1111/imj.13558] [Medline: 28742226]
31. Li R, Curtis K, Zaidi STR, Van C, Castelino R. Effect of the black triangle scheme and its online educational campaign on the quantity and quality of adverse drug event reporting in Australia: a time series analysis. Expert Opin Drug Saf 2020 Jun 31;19(6):747-753. [doi: 10.1080/14740338.2020.1746762] [Medline: 32207361]
32. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. J Med Internet Res 2018 Feb 05;20(2):e41 [FREE Full text] [doi: 10.2196/jmir.8793] [Medline: 29402759]

Abbreviations

ADR: adverse drug reaction
ADRRC: Adverse Drug Reaction Review Committee
GP: general practitioner
HCP: health care professional
PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses
PV: pharmacovigilance
TGA: Therapeutic Goods Administration

Edited by A Mavragani; submitted 05.06.22; peer-reviewed by A Teixeira, C Sullivan, S Bhattacharjee, T Smith; comments to author 29.08.22; revised version received 01.12.22; accepted 22.12.22; published 16.01.23.

Please cite as:
Fossouo Tagne J, Yakob RA, Dang TH, Mcdonald R, Wickramasinghe N
Reporting, Monitoring, and Handling of Adverse Drug Reactions in Australia: Scoping Review
JMIR Public Health Surveill 2023;9:e40080
URL: https://publichealth.jmir.org/2023/1/e40080
doi:10.2196/40080
PMID:36645706

Original Paper

Comparing the Use of a Mobile App and a Web-Based Notification Platform for Surveillance of Adverse Events Following Influenza Immunization: Randomized Controlled Trial

A Brianne Bota[1], PhD; Julie A Bettinger[2], MPH, PhD; Shirley Sarfo-Mensah[1], MPH; Jimmy Lopez[2], BSc; David P Smith[1], MCogSc; Katherine M Atkinson[3], BSc; Cameron Bell[4], BEng; Kim Marty[2], BSc; Mohamed Serhan[1], BSc, HBsc; David T Zhu[1,5], BSc; Anne E McCarthy[6], MD, MSc, DTM&H; Kumanan Wilson[1,4,6,7], MD, MSc

[1]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[2]Vaccine Evaluation Center, BC Children's Hospital Research Institute, Vancouver, BC, Canada
[3]Karolinska Institutet, Solna, Sweden
[4]CanImmunize Inc, Ottawa, ON, Canada
[5]Department of Social and Behavioral Sciences, Yale School of Public Health, New Haven, CT, United States
[6]Department of Medicine, University of Ottawa, Ottawa, ON, Canada
[7]Bruyère Research Institute, Ottawa, ON, Canada

Corresponding Author:
Kumanan Wilson, MD, MSc
Clinical Epidemiology Program
Ottawa Hospital Research Institute
1053 Carling Avenue
Administrative Service Building, Box 684
Ottawa, ON, K1Y 4E9
Canada
Phone: 1 6137985555 ext 17921
Email: kwilson@toh.ca

Abstract

Background: Vaccine safety surveillance is a core component of vaccine pharmacovigilance. In Canada, active, participant-centered vaccine surveillance is available for influenza vaccines and has been used for COVID-19 vaccines.

Objective: The objective of this study is to evaluate the effectiveness and feasibility of using a mobile app for reporting participant-centered seasonal influenza adverse events following immunization (AEFIs) compared to a web-based notification system.

Methods: Participants were randomized to influenza vaccine safety reporting via a mobile app or a web-based notification platform. All participants were invited to complete a user experience survey.

Results: Among the 2408 randomized participants, 1319 (54%) completed their safety survey 1 week after vaccination, with a higher completion rate among the web-based notification platform users (767/1196, 64%) than among mobile app users (552/1212, 45%; *P*<.001). Ease-of-use ratings were high for the web-based notification platform users (99% strongly agree or agree) and 88.8% of them strongly agreed or agreed that the system made reporting AEFIs easier. Web-based notification platform users supported the statement that a web-based notification-only approach would make it easier for public health professionals to detect vaccine safety signals (91.4%, agreed or strongly agreed).

Conclusions: Participants in this study were significantly more likely to respond to a web-based safety survey rather than within a mobile app. These results suggest that mobile apps present an additional barrier for use compared to the web-based notification–only approach.

Trial Registration: ClinicalTrials.gov NCT05794113; https://clinicaltrials.gov/show/NCT05794113

(JMIR Public Health Surveill 2023;9:e39700) doi:10.2196/39700

KEYWORDS

active participant–centered reporting; health technology; adverse event reporting; mobile apps; immunization; vaccine; safety; influenza; campaign; apps; mobile; surveillance; pharmacovigilance

Introduction

Vaccine safety surveillance is a core component of vaccination programs to monitor the safety of vaccines for health care professionals, policy makers, and the public. The implementation of a vaccine safety surveillance program increases the public confidence in various vaccines that are introduced [1]. Serious adverse events following immunization (AEFIs) are rare but do occur. For this reason, good-quality pharmacovigilance is necessary to detect AEFIs. Digital technology offers the potential to enhance and improve AEFI surveillance.

Vaccine pharmacovigilance in Canada currently involves passive and active surveillance systems that are designed to detect even very rare events in the population of vaccine recipients. However, passive surveillance suffers from underreporting and reporting biases (based on age, severity, and type of vaccine), while active surveillance occurs only for adverse events (AEs) in hospitalized children for specific conditions [2,3]. The detection of safety signals can be slow for either system and does not allow for the calculation of population-based incidence rates.

Participant-centered active vaccine reporting offers a potentially more economical and sustainable mechanism to conduct large-scale pharmacovigilance as it allows for rapid identification of AEFIs with minimal human resource needs. The Canadian National Vaccine Safety (CANVAS) network was established in 2009 to provide enhanced monitoring for pandemic and seasonal influenza vaccines using rapid, web-based active surveillance. CANVAS provides timely influenza vaccine safety information annually, via a web-based survey, collected from over 50,000 adults and parents across Canada. CANVAS includes an unvaccinated control group, which provides a robust approach for conducting rapid evaluations of vaccine safety.

The use of mobile health (mHealth) has allowed researchers, policy makers, and health care practitioners to reach individuals who are often less accessible, which, in turn, encompasses a broader and more representative sample of the population [4]. mHealth has not only increased accessibility but also provided low-cost health care solutions for various populations [5]. The COVID-19 pandemic and influenza outbreaks disproportionally harm individuals from communities that may be less accessible [6]. The need to reach broader populations is important in vaccine safety research to better improve postmarket surveillance and could potentially be aided through low-cost solutions via mHealth. However, this needs to be evaluated.

mHealth is a rapidly growing field, and the near ubiquity of smartphones presents a unique opportunity to incorporate digital technologies to address public health issues, such as monitoring AEFIs, and facilitate communication between individuals and public health officials. With increased mobile device and app usage, the potential exists to capture, transmit, and monitor postimmunization experiences rapidly using self-reporting and personal mobile devices. For example, Australia has successfully implemented participant-centered digital AEFI reporting using SMS text messaging technology [7]. CANImmunize is a digital immunization tracking solution that could serve as an acceptable platform for digital AEFI reporting [8,9].

In this study, we will evaluate the effectiveness and feasibility of using a mobile app for reporting participant-centered seasonal influenza AEFIs compared to a web-based notification platform.

Methods

Study Procedures

A 2-centered randomized controlled trial (RCT) was conducted to evaluate the use of safety reporting via a mobile app compared to safety reporting via web-based CANVAS notifications among individuals receiving the influenza vaccine from October 6 to November 29, 2020, during the seasonal influenza vaccine campaign in Ottawa, Ontario, and Vancouver, British Columbia, Canada.

Individuals were recruited at the time of receiving their influenza vaccine. Eligibility criteria included the ability to speak English or French, having an active email address and telephone number, and being immunized with the seasonal influenza vaccine.

Randomization

After study enrollment, participants were randomized to receive the web-based safety survey either through the mobile app or were emailed a link to the web-based survey using a 4-block randomization design.

Web-Based Notification Arm

All participants randomized to the web-based notification arm received the following web-based CANVAS notifications [10,11]. Briefly, participants received an email notifying them of their registration in the study. Eight days following their influenza vaccine, participants received an email with the survey link asking them to complete their web-based influenza vaccine safety survey. Participants received a reminder email on day 11 if they did not complete their survey. Further details on CANVAS surveillance and a description of the questionnaire can be found in previous studies [10,12,13].

Mobile App Arm

Participants randomized to the mobile app arm received an email asking them to download the app and activate their account. Users who did not activate their account after 48 hours received a reminder email. Participants who activated their accounts could spontaneously report an AE through the app and were also notified of the day 8 survey through the app.

Eight days following their vaccination, mobile app participants who activated their accounts received a push notification on their phone to complete their survey. A reminder push notice was sent out on day 11 to participants who had not yet competed the day 8 survey. On November 16, 2020 (midway through the recruitment period), additional email reminders in the mobile app arm were implemented on days 2, 4, and 6 to remind participants to register for the app. All participants received a day 8 email directing them to use their CANImmunize account to complete their influenza vaccine survey. Access to the survey link also was available in the email reminder.

Usability Survey

Following completion of the safety survey, all participants were sent a separate link via email to complete a user experience survey. Participants were asked about their history of participating in the flu vaccine safety survey and whether they previously used the CANImmunize app. Using a Likert scale, participants were asked questions on (1) perceived ease of use, (2) perceived usefulness, (3) their attitudes and intention of use toward the platform, and (4) questions pertaining to vaccine confidence and safety.

Statistical Analysis

The data were analyzed within the intention-to-treat arm to which participants were randomized. The data were summarized descriptively. Differences in response rates were compared using a chi-square test. AEFI incidence was compared between groups using a chi-square test and presented with 95% CIs. Response time between groups was assessed using a Student *t* test. A Mann-Whitney test with adjustments for multiple comparisons was used to compare useability responses among new and previous web-based notification platform users. Significance was accepted as *P*<.05.

Ethical Considerations

All participants provided informed electronic consent upon completion of the web-based survey for primary data collection as well as secondary analyses of research data. The Ottawa Health Science Network Research Ethics Board (20200591-01H) and the BC Children’s Hospital Research Ethics Board approved this study. This was a substudy under the CANVAS protocol (OHSN REB: 20100715-01H, BC: H10-02274). All participants’ personal data were anonymized to protect their privacy and confidentiality. Participants did not receive any monetary compensation for participating in this study.

Results

Overview

Between October 6 and November 29, 2020, a total of 2408 individuals agreed to participate in the RCT (Vancouver, n=1409; Ottawa, n=999). In total, 1196 participants were randomized to the CANVAS arm and 1212 to the CANImmunize arm (Figure 1).

Overall, 1319 (54%) RCT participants completed the web-based safety survey. Participant demographics are presented in Table 1. The completion rate was higher among web-based notification platform users (64%, n=767) than among mobile app users (45%, n=552; *P*<.001).

Of the 552 mobile app respondents, 15% (N=87) created a CANImmunize account and 4.3% (n=24) already had an account, 35 (6.3%) accessed the survey using the app, and the remaining participants (n=517, 93.6%) accessed the survey in the reminder email.

Of the users who created a new account, 100 (90% of registered users) registered for their account after November 16, 2020, compared to 10 before. Initiation of the reminder emails improved survey completion in the CANImmunize arm, with 73% (n=404) of mobile app survey completions occurring after the implementation of additional reminder emails.

Figure 1. Overview of participant recruitment, randomization, and survey completion. RCT: randomized controlled trial.

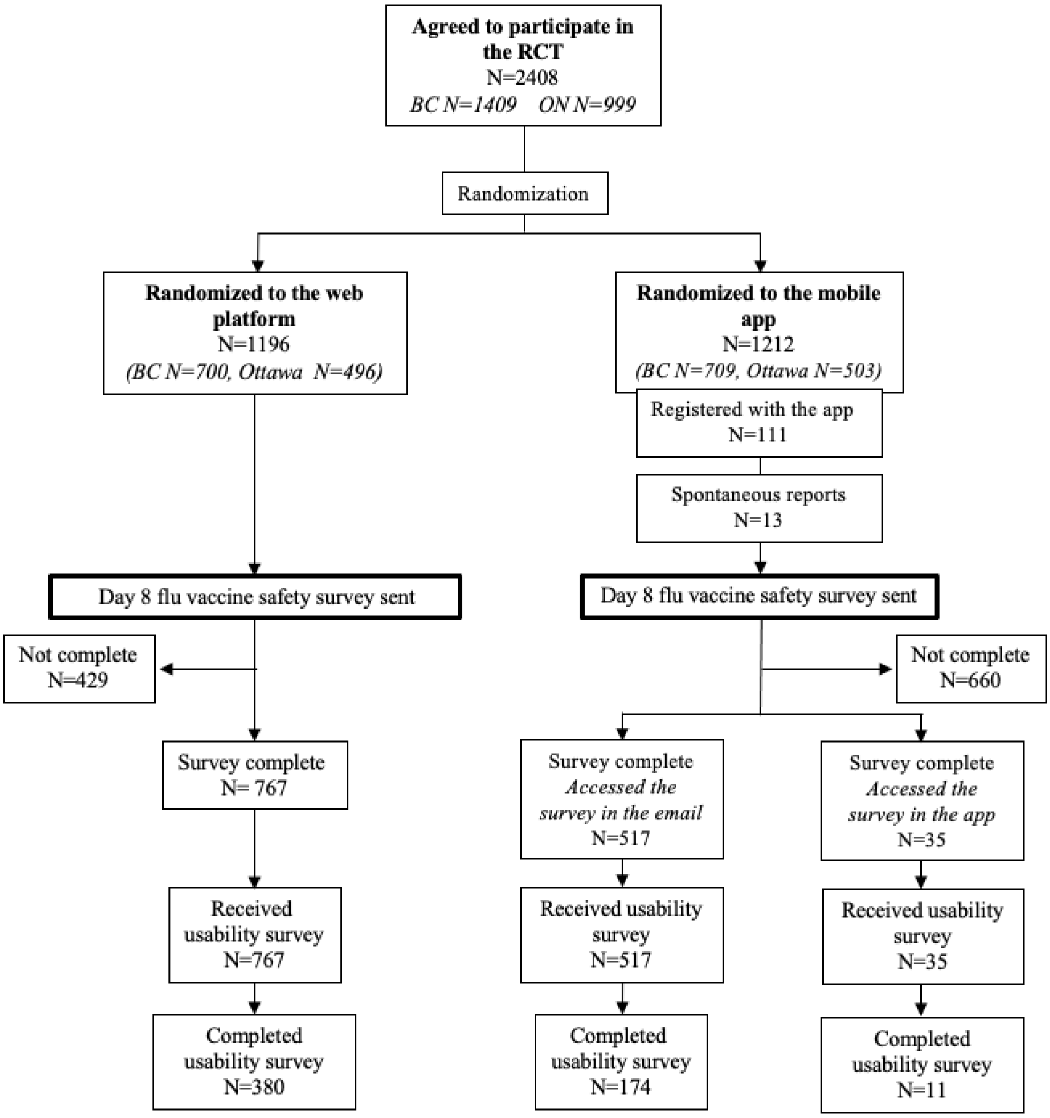


Table 1. Demographic details of participants (N=1319) who submitted a flu vaccine safety survey (intention-to-treat arm).

	Web-based notification (n=767), n (%)	Mobile app (n=552), n (%)	Total, n (%)
Sex			
Female	430 (56.1)	325 (58.9)	755 (57.2)
Male	333 (43.4)	226 (40.9)	559 (42.4)
Other	4 (0.5)	1 (0.2)	5 (0.4)
Age group			
6-23 months	36 (4.7)	28 (5.1)	64 (4.8)
2-4 years	87 (11.3)	81 (14.7)	168 (12.7)
5-9 years	138 (18.0)	74 (13.4)	212 (16.1)
10-14 years	96 (12.5)	66 (11.9)	162 (12.3)
15-19 years	23 (3.0)	14 (2.5)	37 (2.8)
20-29 years	43 (5.6)	32 (5.8)	75 (5.7)
30-39 years	120 (15.6)	114 (20.6)	234 (17.7)
40-49 years	142 (18.5)	83 (15.0)	225 (17.1)
50-64 years	76 (9.9)	55 (10.0)	131 (9.9)
65-79 years	5 (0.6)	4 (0.7)	9 (0.7)
80+ years	1 (0.1)	1 (0.2)	2 (0.2)
Number of Flu vaccines in the last 2 years			
0	78 (10.2)	53 (9.6)	131 (11.7)
1	141 (18.4)	100 (18.1)	241 (3.7)
2	548 (71.4)	399 (72.3)	947 (84.6)
Previous web-based notification platform usage			
Yes	171 (44.4)[a]	78 (40.2)[b]	249 (43.0)[c]
No	198 (51.4)[a]	98 (50.5)[b]	296 (51.1)[c]
Unknown	16 (4.2)[a]	18 (9.3)[b]	34 (5.9)[c]
Existing app users			
Yes	N/A[d]	26 (13.4)	N/A
No	N/A	158 (81.4)	N/A
Unknown	N/A	10 (5.2)	N/A

[a] Calculated on the basis of a total of 385 participants.

[b] Calculated on the basis of a total of 194 participants.

[c] Calculated on the basis of a total of 579 participants.

[d] N/A: not applicable.

Response Time

The mean response time for the day 8 survey was slightly longer for the mobile app group than for the web-based notification platform group (10.7, SD 3.9 days vs 10.1, SD 3.2 days, respectively; *P*=.001).

Spontaneous Reports

Spontaneous reports were available for mobile app users only. In total, 27 mobile app users accessed the spontaneous report survey, and 13 users submitted a spontaneous report, 1 of which was considered incomplete. The spontaneous reports were submitted on average 4.3 (SD 1.3, median 4) days after vacation. One of the 13 spontaneous reports was medically attended. Seven of the completed reports did not indicate any symptoms, suggesting that participants may have submitted their day 8 survey early or may have misunderstood the spontaneous report since they did not report any AEFIs despite the purpose of the survey being to gather AEFI information.

Event Reporting

Approximately 10% (n=134) of study participants reported experiencing a new or worsening health problem (78 web-based notification system users vs 56 mobile app users). In total, 43 of these participants (29 web-based notification system users vs 14 mobile app users) reported missing work or activities or consulting a medical professional (Table 2). No significant differences were observed in incidence rates between web-based and mobile app participants. The relative risk of developing or worsening of any health problem for the web-based notification arm relative to those in the mobile app arm was 1.002 (95% CI 0.7226-1.39). The relative risk of developing or worsening of any nonsevere health problem was 0.839 (95% CI 0.5702-1.2649). The relative risk of a health problem being severe enough to miss work or activities or a medical consultation was 1.491 (95% CI 0.7855-2.7619).

Table 2. Adverse event reporting rates among participants.

	Web-based notification system users, n (rate, 95% CI)	Mobile app users, n (rate, 95% CI)
Development or worsening of any health problem	78 (10.17, 8.12-12.53)	56 (10.14, 7.8-13.0)
Development or worsening of any nonsevere health problem	49 (6.39, 4.76-8.36)	42 (7.61, 5.5-10.1)
Adverse event severe enough to miss work or activities or a medical consultation	29 (3.78, 2.55-5.39)	14 (2.54, 1.4-4.2)

Usability and Perceived Usefulness

In total, 579 participants took the usability survey, with 194 mobile app and 385 web-based notification respondents, respectively. In total, 249 of those completing the usability survey had previously participated in the influenza vaccine safety survey (n=249, 43% of usability respondents) and 13% of usability respondents had used CANImmunize before (mobile app arm only, n=26).

Of the 35 mobile app users who completed the survey in the app, only 11 completed the useability survey; hence, we have only reported on the usability responses of the web-based notification platform users (Table 3). We conducted a sensitivity analysis among existing and new web platform users. There was no difference in responses between new and previous users of the web-based notification survey respondents (Multimedia Appendix 1).

Overall, 99% of web-based notification platform users agreed or strongly agreed that the platform was easy to use, and 88.8% of them thought that it made reporting AEFI easier.

When asked about perceived usefulness, 73.8% of web-based notification platform users agreed or strongly agreed that it will make vaccines safer, and 91.4% of them agreed or strongly agreed that it could make it easier for public health professionals to detect vaccine safety issues. The majority of participants (88.8%) agreed or strongly agreed that it made reporting a vaccine side effect easier, using this system was a good idea (93.2%), and that they would use it for additional vaccines (85.9%) or the COVID-19 vaccine (94.0%). Despite this, only 47.4% of web-based notification platform users who completed the usability survey reported that it increased their confidence in vaccine safety. Only 77.7% of users felt confident about data privacy and security.

Table 3. Perceived ease of use, usefulness, attitudes toward use, and intention to use among web-based notification survey users.

	Web-based notification survey (n=385)					
	Total respondents, n	Strongly disagree, n (%)	Disagree, n (%)	Neither agree nor disagree, n (%)	Agree, n (%)	Strongly agree, n (%)
Easy to use	380	1 (0.3)	0 (0)	2 (1.1)	61 (33.2)	316 (83.2)
Easy to open	382	1 (0.3)	0 (0)	2 (1.1)	75 (40.8)	304 (79.6)
Easy to access	377	0 (0)	0 (0)	2 (1.1)	72 (39.1)	303 (80.4)
This system will help make vaccines safer	382	2 (0.5)	3 (0.8)	95 (24.9)	132 (34.6)	150 (39.3)
This system increased my awareness of vaccine records	377	11 (2.9)	45 (11.9)	155 (41.1)	80 (21.2)	86 (22.8)
This system could make it easier for public health to detect safety issues with new vaccines	382	1 (0.3)	0 (0)	32 (8.4)	181 (47.4)	168 (44)
This system allows me to easily report a vaccine side effect (an adverse event) following immunization	376	5 (1.3)	3 (0.8)	34 (9.0)	150 (39.9)	184 (48.9)
Using this system to report vaccine side effects (adverse events following immunization) is a good idea	380	1 (0.3)	2 (0.5)	23 (6.1)	144 (37.9)	210 (55.3)
I feel confident about the privacy and security of my data in this system	381	1 (0.3)	3 (0.8)	81 (21.3)	152 (39.9)	144 (37.8)
If this was available for additional vaccines, I would use it	382	1 (0.3)-	1 (0.3)	52 (13.6)	142 (37.2)	186 (48.7)
I would use this system for a new COVID-19 vaccine	382	1 (0.3)	0 (0)	22 (5.8)	142 (37.2)	217 (56.8)
This system increased my confidence in the safety of vaccines	378	1 (0.3)	19 (5.0)	179 (47.4)	87 (23.0)	92 (24.3)
Getting vaccinated is a good way to protect myself and my family from disease	374	1 (0.3)	0 (0)	6 (1.6)	71 (19.0)	296 (79.1)
Vaccinating myself and my family is important for the health of others in my community	379	1 (0.3)	0 (0)	3 (0.8)	68 (17.9)	307 (81)
I am concerned about serious side effects of vaccines	381	32 (8.4)	94 (24.7)	80 (21.0)	106 (27.8)	69 (18.1)
New vaccines carry more risks than older vaccines	383	31 (8.1)	41 (10.7)	187 (48.8)	85 (22.2)	39 (10.2)

Discussion

Principal Findings

This study compared 2 digital reporting systems (a web-based notification platform vs a mobile app) for AEFI reporting during the 2020 seasonal influenza vaccination campaign. In total, 2408 individuals agreed to participate in the RCT. A total of 1196 participants were randomized to the CANVAS arm and 1212 to the CANImmunize arm (Figure 1). Overall, 1319 (54%) RCT participants completed the web-based safety survey. The completion rate was higher among web-based notification platform users (64%, n=767) than among mobile app users (45%, n=552). Ease of use was higher among web-based notification platform users. Our findings suggest that a web-based survey link is an acceptable approach for active, participant-centered AEFI reporting. We identified a number of concerns with mobile app reporting, which would need to be addressed to improve the acceptability and usability of app-based reporting.

To be successful, digital solutions must be easy to use, easy to access, and meet the needs of the users [14]. In this study, we saw lower rates of participant engagement in the mobile app arm and were unable to assess usability due to low participant responses. Participants were required to download the app and register their account. This added complexity and was likely a barrier to use.

Interestingly, the highest response rate in the mobile app group was observed among 30-39–year-old participants. This was possibly due to this demographic being having a higher level of technological literacy and comfort with mobile apps [15]. The lowest response rate among the mobile app users was observed among older adults aged >65 years. The low response rate among older adults may be due to less comfort with mobile apps or possibly barriers such as difficulty reading smaller font [16].

The lack of responses to the usability survey is also a potential concern as we were unable to assess exactly what needs to change to make mobile app reporting acceptable and user-friendly. The usability survey was provided as a separate email link after the safety survey was completed, and it may have been missed by users.

It is critical to make mHealth solutions accessible, and a barrier may be technological literacy or low perceived ease of use among some populations. It is important that mHealth apps are designed to be user-friendly for target demographics [17]. This is important in vaccine safety surveillance studies that need to capture data from a breadth of demographics.

Current literature suggests that the development of mHealth apps should follow 8 specific categories [17] with usability being one of the key categories. Usability refers to the app being adapted to the target population [17]. Making the app easy to use, with clear instructions, and the feedback of various community members who use the app is critical. Our finding—a lower response rate was observed among mobile app users than among the web-based notification platform users—implies that the mobile app may not have been optimally designed for usability for all of the target demographics. The app was possibly challenging to navigate or perceived to be so resulting in a lower response rate than that among users of the web-based notification system.

Enabling AEFI reporting from individuals who are already using a mobile app for another purpose would require less work for the user and may be a more successful approach. For example, 42% of Canadians report accessing websites, mobile apps, or other interactive web-based services to support or monitor their health [18]. Integrating use with these types of health monitoring apps would be an approach worth trialing. Indeed, since the completion of this study, the CANVAS-COVID safety survey has recruited control participants from among current CANImmunize users [19]. Other digital technologies have been used during the COVID-19 vaccine rollout. Social media was used to collect postmarket vaccine safety data [20-23], and the US Centers for Disease Control and Prevention created the V-safe After Vaccination Health Checker application, which allowed individuals to register and complete vaccine safety surveys after receiving their COVID-19 vaccine [23].

Based on the RCT conducted in this study, it was found that adding the extra step of using the mobile app introduced an additional barrier for participants who may have lower technological literacy. Future mHealth developments should include various participants from the target demographics in the development process to ensure usability.

Limitations

This study has several limitations. Mobile app users were required to register for a CANImmunize account for the study. Registration allowed for influenza vaccine data to be uploaded to the users' CANImmunize account; however, this additional step appeared to be a barrier to use as we saw low registration numbers. Study participants were randomized after study registration; hence, onsite personnel were not able to support participants randomized to the mobile app arm in activating their account, which may also have contributed to low account registration. However, this can be perceived as a strength as this provided more real-world effectiveness data of app usage as in a nonstudy setting, users would have to navigate this on their own without study staff assistance. The day 8 survey email served as a reminder to mobile app users to register their accounts if they had not yet done so, but it also provided a direct link to the web-based safety and usability surveys, which 94% of mobile app arm users who completed the survey chose to use, rather than registering for a mobile app account. As a result, we could not effectively evaluate user preference or experiences in the mobile app arm of the study. Another major limitation is that data about nonresponders were not collected, which could introduce risk for response bias, particularly since the response rate differed between the web-based notification arm (64%) and the mobile app arm (45%).

Conclusions

In summary, this study demonstrated high user acceptability with the web-based survey platform compared to that with a mobile app. Making AEFI reporting available to existing mobile app users may still be a viable initiative for participant-centered active reporting in Canada but would require further refinement.

Acknowledgments

We wish to thank the individuals who participated in this study. This project was supported by The Ottawa Hospital Academic Medical Organization (TOHAMO) and The Ottawa Hospital Foundation. The Canadian National Vaccine Safety Network as part of the Canadian Immunization Research Network is funded by a research grant from the Canadian Institutes of Health Research and the Public Health Agency of Canada (grant #151944; June 1, 2017, to December 31, 2022).

Data Availability

Data are available upon request.

Conflicts of Interest

KW, KA, and CB were cofounders and shareholders of CANImmunize Inc, a corporation that provides consumer facing solutions to immunization information systems. KW has served on the Medicago Independent Data Monitoring Committee and is a member of the Moderna Global Safety Core Consultancy Group. All other authors have no conflicts of interest to disclose.

Editorial notice

This randomized study was only registered retrospectively due to delays caused by the COVID-19 pandemic. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
Useability survey responses for new compared to previous users of the influenza vaccine safety survey web platform.
[DOCX File , 24 KB - publichealth_v9i1e39700_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 391 KB - publichealth_v9i1e39700_app2.pdf]

References

1. Cheng AC, Buttery JP. Vaccine safety: what systems are required to ensure public confidence in vaccines? Med J Aust 2022 Aug 15;217(4):189-190 [FREE Full text] [doi: 10.5694/mja2.51662] [Medline: 35843626]
2. IMPACT after 17 years: Lessons learned about successful networking. Paediatr Child Health 2009 Jan;14(1):33-39 [FREE Full text] [doi: 10.1093/pch/14.1.33] [Medline: 19436462]
3. Scheifele DW, Halperin SA, CPS/Health Canada, Immunization Monitoring Program, Active (IMPACT). Immunization Monitoring Program, Active: a model of active surveillance of vaccine safety. Semin Pediatr Infect Dis 2003 Jul;14(3):213-219. [doi: 10.1016/s1045-1870(03)00036-0] [Medline: 12913834]
4. Liu P, Astudillo K, Velez D, Kelley L, Cobbs-Lomax D, Spatz ES. Use of mobile health applications in low-income populations. Circ: Cardiovascular Quality and Outcomes 2020 Sep;13(9) [FREE Full text] [doi: 10.1161/circoutcomes.120.007031]
5. Western MJ, Armstrong MEG, Islam I, Morgan K, Jones UF, Kelson MJ. The effectiveness of digital interventions for increasing physical activity in individuals of low socioeconomic status: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2021 Nov 09;18(1):148 [FREE Full text] [doi: 10.1186/s12966-021-01218-4] [Medline: 34753490]
6. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 2020 Nov;74(11):964-968 [FREE Full text] [doi: 10.1136/jech-2020-214401] [Medline: 32535550]
7. Leeb A, Regan AK, Peters IJ, Leeb C, Leeb G, Effler PV. Using automated text messages to monitor adverse events following immunisation in general practice. Med J Aust 2014 Apr 21;200(7):416-418. [doi: 10.5694/mja13.11166] [Medline: 24794676]
8. Digitally transforming immunization practice for government, employers, health care and Canadians. CANImmunize. URL: https://www.canimmunize.ca/en/home [accessed 2023-01-13]
9. Houle S, Atkinson K, Paradis M, Wilson K. CANImmunize: A digital tool to help patients manage their immunizations. Can Pharm J (Ott) 2017;150(4):236-238 [FREE Full text] [doi: 10.1177/1715163517710959] [Medline: 29163723]
10. Bettinger J, De Serres G, Valiquette L, Vanderkooi O, Kellner J, Coleman B, Canadian Immunization Research Network. 2017/18 and 2018/19 seasonal influenza vaccine safety surveillance, Canadian National Vaccine Safety (CANVAS) Network. Euro Surveill 2020 Jun;25(22):2020 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.22.1900470] [Medline: 32524947]
11. Bettinger J, De Serres G, Valiquette L, Vanderkooi O, Kellner J, Coleman B, Canadian Immunization Research Network. 2017/18 and 2018/19 seasonal influenza vaccine safety surveillance, Canadian National Vaccine Safety (CANVAS) Network. Euro Surveill 2020 Jun;25(22):2017 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.22.1900470] [Medline: 32524947]
12. Bettinger J, Rouleau I, Gariepy MC, Bowie W, Valiquette L, Vanderkooi O, Public Health Agency of Canada/Canadian Institutes for Health Research Influenza Research Network. Successful methodology for large-scale surveillance of severe events following influenza vaccination in Canada, 2011 and 2012. Euro Surveill 2015 Jul 23;20(29):21189 [FREE Full text] [doi: 10.2807/1560-7917.es2015.20.29.21189] [Medline: 26227369]
13. Bettinger J, Vanderkooi O, MacDonald J, Kellner J. Rapid online identification of adverse events after influenza immunization in children by PCIRN's National Ambulatory Network. Pediatr Infect Dis J 2014 Oct;33(10):1060-1064. [doi: 10.1097/INF.0000000000000373] [Medline: 25361187]
14. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
15. van Laar E, van Deursen AJAM, van Dijk JAGM, de Haan J. Determinants of 21st-century skills and 21st-century digital skills for workers: a systematic literature review. SAGE Open 2020 Jan 24;10(1):215824401990017 [FREE Full text] [doi: 10.1177/2158244019900176]
16. Jung SO, Son YH, Choi E. E-health literacy in older adults: an evolutionary concept analysis. BMC Med Inform Decis Mak 2022 Jan 31;22(1):28 [FREE Full text] [doi: 10.1186/s12911-022-01761-5] [Medline: 35101005]
17. Llorens-Vernet P, Miró J. Standards for mobile health-related apps: systematic review and development of a guide. JMIR Mhealth Uhealth 2020 Mar 03;8(3):e13057 [FREE Full text] [doi: 10.2196/13057] [Medline: 32130169]

18. Digital Health Services Survey: What We Heard from Canadians. Government of Canada. 2021. URL: https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04573.html [accessed 2022-04-27]
19. Bettinger JA, Sadarangani M, De Serres G, Valiquette L, Vanderkooi OG, Kellner JD, Canadian Immunization Research Network. The Canadian National Vaccine Safety Network: surveillance of adverse events following immunisation among individuals immunised with the COVID-19 vaccine, a cohort study in Canada. BMJ Open 2022 Jan 20;12(1):e051254 [FREE Full text] [doi: 10.1136/bmjopen-2021-051254] [Medline: 35058258]
20. Alonso R, Chertcoff A, Leguizamón FDV, Galleguillos Goiry L, Eizaguirre MB, Rodríguez R, et al. Evaluation of short-term safety of COVID-19 vaccines in patients with multiple sclerosis from Latin America. Mult Scler J Exp Transl Clin 2021 Oct;7(4):20552173211061543 [FREE Full text] [doi: 10.1177/20552173211061543] [Medline: 34868630]
21. Djanas D, Yusirwan, Martini RD, Rahmadian, Putra H, Zanir A, Syahrial, et al. Survey data of COVID-19 vaccine side effects among hospital staff in a national referral hospital in Indonesia. Data Brief 2021 Jun;36:107098 [FREE Full text] [doi: 10.1016/j.dib.2021.107098] [Medline: 33969163]
22. Adam M, Gameraddin M, Alelyani M, Alshahrani MY, Gareeballah A, Ahmad I, et al. Evaluation of post-vaccination symptoms of two common COVID-19 vaccines used in Abha, Aseer Region, Kingdom of Saudi Arabia. PPA 2021 Sep;Volume 15:1963-1970. [doi: 10.2147/ppa.s330689]
23. Chapin-Bardales J, Myers T, Gee J, Shay DK, Marquez P, Baggs J, et al. Reactogenicity within 2 weeks after mRNA COVID-19 vaccines: findings from the CDC v-safe surveillance system. Vaccine 2021 Nov 26;39(48):7066-7073 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.019] [Medline: 34763946]

Abbreviations

AE: adverse event
AEFI: adverse event following immunization
CANVAS: Canadian National Vaccine Safety Network
mHealth: mobile health
RCT: randomized controlled trial

Edited by T Sanchez, A Mavragani; submitted 19.05.22; peer-reviewed by S Matsuda, R Poluru, N Jarour; comments to author 22.11.22; revised version received 13.01.23; accepted 30.01.23; published 08.05.23.

Please cite as:
Bota AB, Bettinger JA, Sarfo-Mensah S, Lopez J, Smith DP, Atkinson KM, Bell C, Marty K, Serhan M, Zhu DT, McCarthy AE, Wilson K
Comparing the Use of a Mobile App and a Web-Based Notification Platform for Surveillance of Adverse Events Following Influenza Immunization: Randomized Controlled Trial
JMIR Public Health Surveill 2023;9:e39700
URL: https://publichealth.jmir.org/2023/1/e39700
doi:10.2196/39700
PMID:37155240

Original Paper

Correlation Between Opioid Drug Prescription and Opioid-Related Mortality in Spain as a Surveillance Tool: Ecological Study

Alejandro Salazar[1,2,3], PhD; Soledad Moreno-Pulido[4], PhD; Pablo Prego-Meleiro[5,6], PhD; Jesús Henares-Montiel[7,8], PhD; José Pulido[5,7], PhD; Marta Donat[9], PhD; Gabriel Sotres-Fernandez[10,11], MD; Luis Sordo[5,7], PhD

[1]Observatory of Pain, University of Cádiz, Cádiz, Spain

[2]Biomedical Research and Innovation Institute of Cádiz (INiBICA), Research Unit, Puerta del Mar University Hospital, University of Cadiz, Cádiz, Spain

[3]Department of Statistics and Operational Research, University of Cádiz, Puerto Real, Spain

[4]Department of Mathematics, University of Cádiz, Puerto Real, Spain

[5]Department of Public Health and Maternal-Child Health, Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

[6]Department of Analytical Chemistry, Physical Chemistry and Chemical Engineering, Faculty of Pharmacy, University of Alcalá, Alcalá de Henares, Spain

[7]Center for Biomedical Research in Epidemiology and Public Health (CIBER), Madrid, Spain

[8]Andalusian School of Public Health, Granada, Spain

[9]National School of Health, Salud Carlos III Health Institute, Madrid, Spain

[10]Oncology Department, Hospital Quirón Salud, Madrid, Spain

[11]Department of Medicine, Faculty of Biomedical and Health Sciences, Universidad Europea de Madrid, Madrid, Spain

Corresponding Author:
Soledad Moreno-Pulido, PhD
Department of Mathematics
University of Cádiz
Avenida de la Universidad de Cádiz, s/n
Puerto Real, 11510
Spain
Phone: 34 956483474
Email: soledad.moreno@uca.es

Abstract

Background: Opioid drug prescription (ODP) and opioid-related mortality (ORM) have increased in Spain. However, their relationship is complex, as ORM is registered without considering the type of opioid (legal or illegal).

Objective: This ecological study aimed to examine the correlation between ODP and ORM in Spain and discuss their usefulness as a surveillance tool.

Methods: This was an ecological descriptive study using retrospective annual data (2000-2019) from the Spanish general population. Data were collected from people of all ages. Information on ODP was obtained from the Spanish Medicines Agency in daily doses per 1000 inhabitants per day (DHD) for total ODP, total ODP excluding those with better safety protocols (codeine and tramadol), and each opioid drug separately. Rates of ORM (per 1,000,000 inhabitants) were calculated based on deaths registered (International Classification of Diseases, 10th Revision codes) as opioid poisoning by the National Statistics Institute, derived from the drug data recorded by medical examiners in death certificates. Opioid-related deaths were considered to be those that indicated opioid consumption (accidental, infringed, or self-inflicted) as the main cause of death: death due to accidental poisoning (X40-X44), intentional self-inflicted poisoning (X60-X64), drug-induced aggression (X85), and poisoning of undetermined intention (Y10-Y14). A descriptive analysis was carried out, and correlations between the annual rates of ORM and DHD of the prescribed opioid drugs globally, excluding medications of the least potential risk of overdose and lowest treatment tier, were analyzed using Pearson linear correlation coefficient. Their temporal evolution was analyzed using cross-correlations with 24 lags and the cross-correlation function. The analyses were carried out using Stata and StatGraphics Centurion 19.

Results: The rate of ORM (2000-2019) ranged between 14 and 23 deaths per 1,000,000 inhabitants, with a minimum in 2006 and an increasing trend starting in 2010. The ODP ranged between 1.51 to 19.94 DHD. The rates of ORM were directly correlated with the DHD of total ODP (r=0.597; P=.006), total ODP without codeine and tramadol (r=0.934; P<.001), and every prescribed opioid except buprenorphine (*P*=.47). In the time analysis, correlations between DHD and ORM were observed in the same year, although not statistically significant (all *P*≥.05).

Conclusions: There is a correlation between greater availability of prescribed opioid drugs and an increase in opioid-related deaths. The correlation between ODP and ORM may be a useful tool in monitoring legal opiates and possible disturbances in the illegal market. The role of tramadol (an easily prescribed opioid) is important in this correlation, as is that of fentanyl (the strongest opioid). Measures stronger than recommendations need to be taken to reduce off-label prescribing. This study shows that not only is opioid use directly related to the prescribing of opioid drugs above what is desirable but also an increase in deaths.

(JMIR Public Health Surveill 2023;9:e43776) doi:10.2196/43776

KEYWORDS

opioid; overdose; drug overdose; opioid-related deaths; mortality; tramadol; fentanyl; substance use; substance misuse; substance abuse; ecological study; death

Introduction

The prescription of opioid drugs has grown enormously in Spain [1,2], with an increase of over 300% since 2001 [3,4]. This is not a local phenomenon. The same is occurring in many countries in Europe and around the world, resulting in a serious public health problem [5-7].

The increase in the opioid drug prescription (ODP) in Spain has been highlighted and registered in pharmaceutical surveillance systems [8], in addition to occasional investigations. This has resulted in recommendations to increase the precision of both prescription indications and duration of treatment [9,10], aligned with what is occurring in the international context [11]. The use of these medications has helped many people to manage their illnesses [9,11], especially those with chronic conditions and cancer [12]. However, the use of these medications is related to the potential appearance of adverse effects: the 2 most relevant are the development of an addiction to these drugs and death due to overdose [1,13].

The opioid crisis in North America is related to the medical prescription of these drugs. Both the number of people addicted to medically prescribed opioids and those addicted to other illegal opioids have increased dramatically in the last decade. The consequences for health, the loss of life years, and the economic costs are enormous [14-16].

The current data on the situation in Spain do not show a situation similar to that of North America [1,17]. The increase in ODP has not yet been accompanied by the phenomenon seen in other countries. However, concerns about the situation are important, and surveillance has been heightened in new profiles of addicted users in treatment as well as in the possible indirect indicators of an increase in consumption. This consumption has increased in an unequal way, based on different active ingredients and indications for use. There is an especially relevant increase in extrahospital use of tramadol and interhospital use of fentanyl [2,8,18], and at the same time, non-oncological use accounts for nearly 90% in certain cases [19].

Together with the registry of ODP, one of the most valuable indicators of the growth and type of use is the registry of deaths due to opioid overdose, also referred to as opioid-related mortality (ORM) [13,20]. This indicator reflects the number of deaths whose principal cause is overdose of any opioid and excludes the deaths due to other causes (such as terminal cancers) in which opioid treatment could exist [1,15,21]. The relationship between ODP and ORM is complex, as in Spain, they are recorded without differentiating between the type of opioid (legal or illegal). Therefore, determining the relationship between ODP and ORM can provide key information both about deaths due to prescribed opioid drugs and, in an indirect way, deaths due to illegal opioid consumption. The objective of this study was to examine the correlation between ODP and ORM in Spain from 2000 to 2019 and discuss their usefulness as a surveillance tool.

Methods

This was an ecological study using retrospective annual data (2000-2019) from the Spanish general population. Data were collected on people of all ages for ORM and ODP.

Data Source

To determine ORM, population data were collected from 2000 to 2019 through the Spanish National Statistics Institute [22], which produces a report of “statistics of deaths according to the cause of death derived from the drug data recorded by medical examiners in their death certificates” [23], following the criteria established by the World Health Organization (WHO) in the International Classification of Diseases, 10th Revision (ICD-10) [24]. This statistic provides information on mortality according to the cause of death. This is similar to methodology used in previous studies [1]. From these data, and according to the ICD-10 codes [24], we retrieved information on ORM. Opioid-related deaths were considered to be those that indicated opioid consumption (accidental, infringed, or self-inflicted) as the main cause of death: death due to accidental poisoning (X40-X44), intentional self-inflicted poisoning (X60-X64), drug-induced aggression (X85), and poisoning of undetermined intention (Y10-Y14). These codes excluded deaths due to opioid overdose in the contexts of compassionate sedation and euthanasia in terminal cases, given that in these cases, the codes were included in the disease codes. The complete list of codes is shown in Multimedia Appendix 1. These codes were used in ways that are similar to what has been carried out in prior publications [21], which also coincides with the codes for death related to the use of opioids as used in later publications [16,25,26]. These causes do not differentiate by the type of opioid, whether in the Spanish or international context. Some of the cited codes (X40, X60, and X85) do not fully correspond to the operational definitions that exist in the Spanish context. For each year (from 2000 to 2019), we report data on the number of ORM and ORM rates per 1,000,000 inhabitants (quotient between the number of opioid-related deaths and the total population).

The data on prescribed opioid drugs were obtained via the Spanish Agency of Medicines and Health Products (*Agencia Española de Medicamentos y Productos Sanitarios* [AEMPS]). All of the drugs require a medical prescription or special permissions to be sold in pharmacies or dispensed by health centers. The AEMPS receives state-level information on all of the prescriptions written for all drugs that contain opioids and separates them by active ingredient. In 2019, the complete list was made up of morphine, hydromorphone, oxycodone, oxycodone and naloxone, pethidine, fentanyl, dextropropoxyphene, pentazocine, buprenorphine, codeine and paracetamol, codeine and acetyl salicylic acid, codeine and ibuprofen, tramadol and paracetamol, tramadol and dexketoprofen, and tramadol and tapentadol. Prescribed opioid drugs were studied both globally and individually, which is explained later in this article. The measurement unit used was dose per 1000 inhabitants per day (DHD).

Statistical Analysis

Normal distribution was confirmed for the variables (ODP and ORM) using Shapiro-Wilk tests, and we observed that all of the variables followed a normal distribution (Multimedia Appendix 2). Subsequently, a description of the variables' evolution over time was carried out. We carried out an analysis of the correlation between the annual rates of ORM and the DHD of the prescribed opioid drugs globally, excluding medications of the least potential risk of overdose and lowest treatment tier [27] (codeine and tramadol), and separately, based on the frequency of their use or their growth over recent years from the initial descriptive analysis. The relationship between ORM and ODP was analyzed as follows: (1) total ODP, (2) total ODP excluding codeine, (3) total ODP excluding codeine and tramadol, (4) tramadol, (5) fentanyl, (6) tapentadol, (7) oxycodone (unifying presentations with and without naloxone), and (8) morphine. All correlations were calculated using Pearson linear correlation coefficient. Statistically significant differences were observed when $P<.05$.

Finally, an analysis of the evolution of each variable was carried out across time, using time series. The relationship between the total evolution of the mortality rate and the DHD of the rest of the drugs was analyzed via cross-correlations with 24 lags and the cross-correlation function. Using this method provides knowledge of the correlations of mortality over time t (in a determined year) and the DHD of each opioid drug in time $t-k$ (in annual spikes or intervals; in this case: 1 year prior, 2 years prior, etc). This could indicate a potential impact of the increase in a dose during year t on mortality in $t-k$ years (1 year later, 2 years later, etc), without necessarily demonstrating a causal relationship.

Time evolution was studied during the whole period of the data, for the periods 2000-2011 and 2012-2019, and for the periods 2000-2015 and 2016-2019. The fragmentation of these periods corresponded to changes in the trend in deaths observed in 2010 and the alert made by health authorities in 2016 about the increase in the use of these medications as well as a call for their rational use [28], respectively.

The analyses were carried out using the Stata statistical package (version 16.1; StataCorp LLC) and StatGraphics Centurion 19 (StatGraphics Technologies, Inc).

Ethics Approval

This study was a secondary analysis of data with no identifying information available to the researchers. It was approved by the institutional review board at the Hospital Clínico San Carlos (affiliated to Complutense University) on March 18, 2021 (approval 21/167-E).

Results

The temporal evolution of the rate of ORM from 2000 to 2019 ranged between 14 and 23 deaths per 1,000,000 inhabitants. Rates decreased progressively to a minimum in 2006, and after some fluctuations, there was a slight increase starting in 2010 that seemed to be stable over time (Figure 1).

In terms of the evolution of the DHD of the opioid drugs analyzed, we observed a continuous increase of more than 40% per year until 2006. Starting in 2006, this became more acute, especially in 2010 when the rate grew from 4.4 to 10 DHD, with a global increase from 2000 to 2019 of 1.51 to 19.83 DHD (12.6 times greater). The magnitude of tramadol stands out (prescribed alone or in conjunction with other nonopioid active ingredients); tramadol represents 63.7% of the total opioids prescribed, with an increase of 10% (of the total) since 2010. Codeine, however, declined from 17% of opioids in 2010 to 9% in 2019. Fentanyl accounted for 13% of prescribed opioids in 2019, followed by tapentadol (5%), buprenorphine (3.4%), and oxycodone (3%). Notable growth can be observed in all of these opioids. Fentanyl and tapentadol both grew by 1000% between 2000 and 2012, whereas oxycodone (first, without and, later, with naloxone) grew by 600% since 2011. The consumption of buprenorphine has been very stable over the past decade. Finally, morphine accounted for 1.3% of prescribed opioids, with an increase in consumption of 63% since 2010 (Figure 2).

Mortality rates for each year were directly correlated (the higher DHD of opioid drug, the higher the mortality rate) in all of the cases analyzed, except in the case of buprenorphine ($P=.47$). The strongest positive correlations were observed with global ODP without codeine ($r=0.902$; $P<.001$), global ODP ($r=0.597$; $P=.006$), global ODP without codeine or tramadol ($r=0.934$; $P<.001$), tapentadol ($r=0.873$; $P=.003$), and oxycodone ($r=0.835$; $P=.002$; Figure 3).

Figure 1. Evolution of opioid-related mortality per 1,000,000 inhabitants (2000-2019).

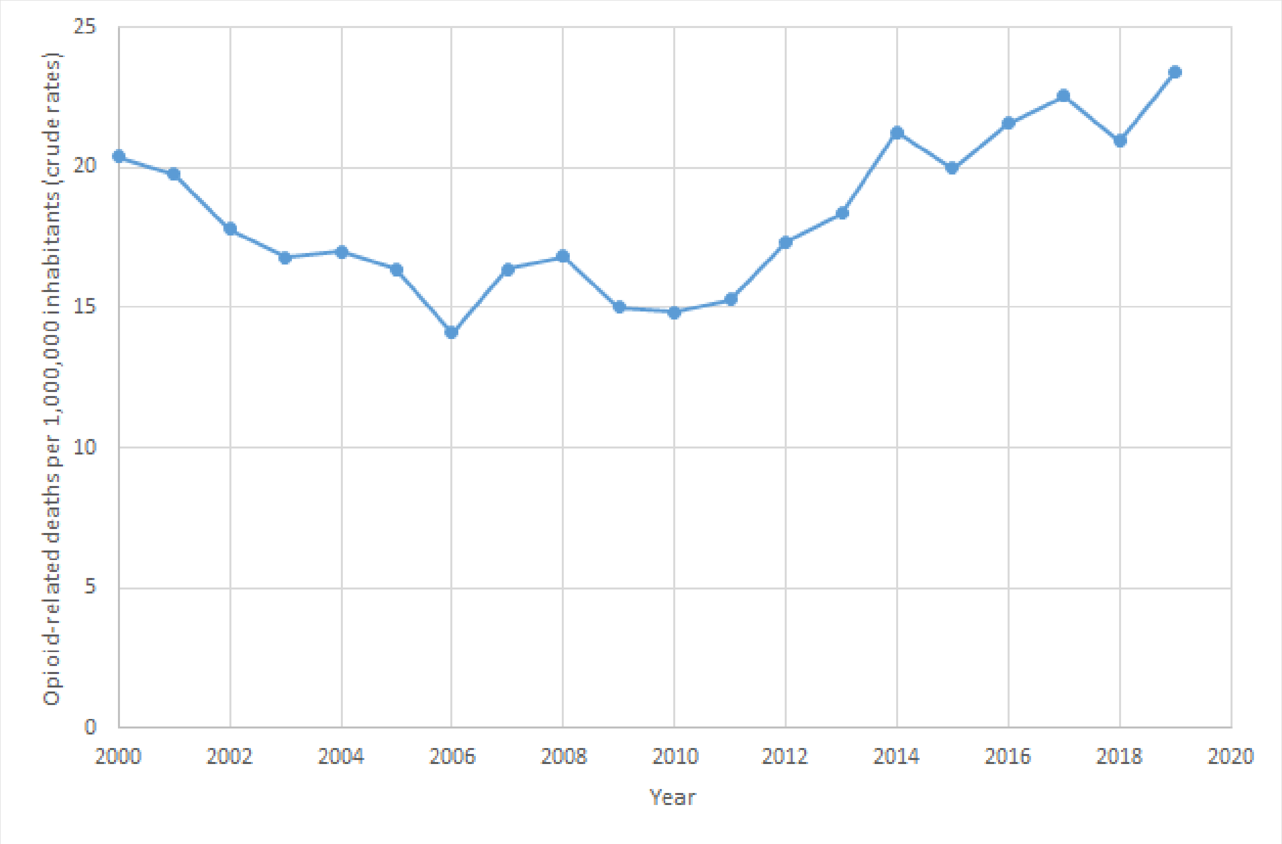


Figure 2. Opioid drug prescription evolution from 2000 to 2019 (unit: dose per 1000 inhabitants per day [DHD]).

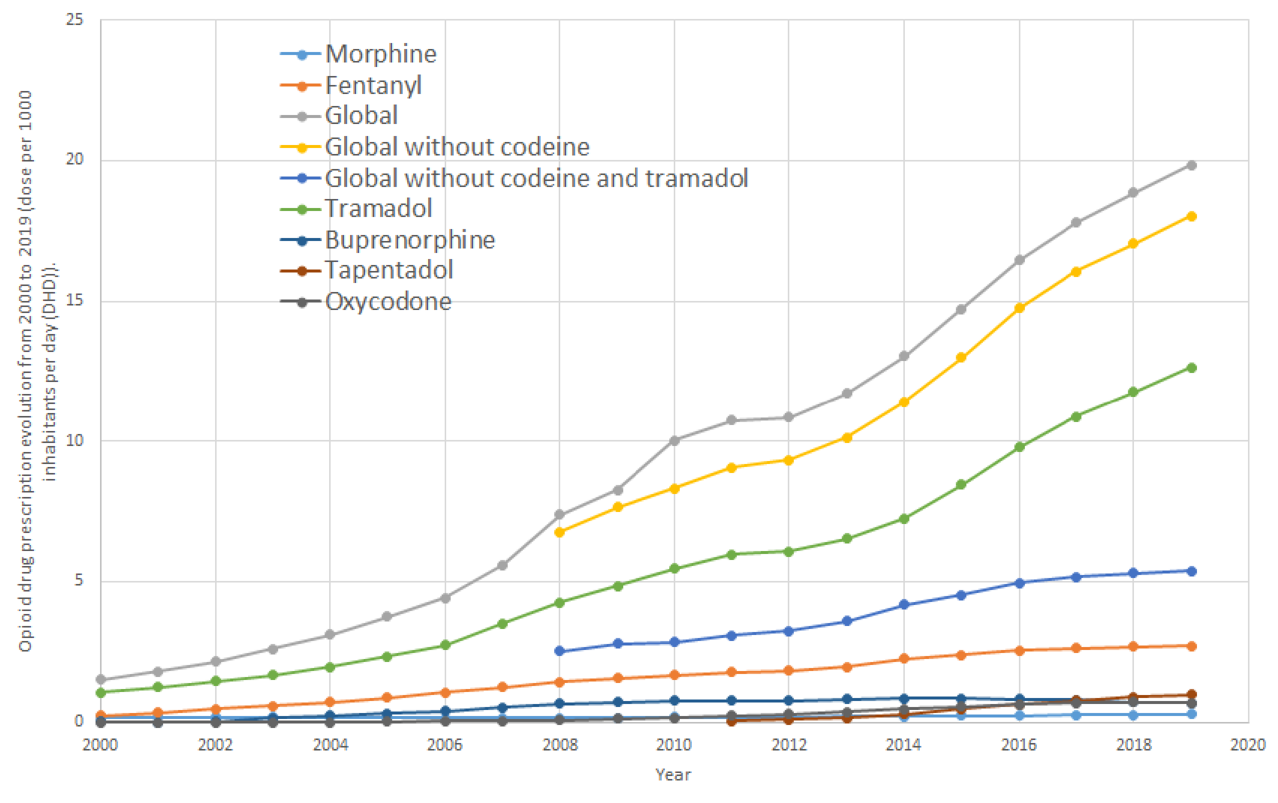
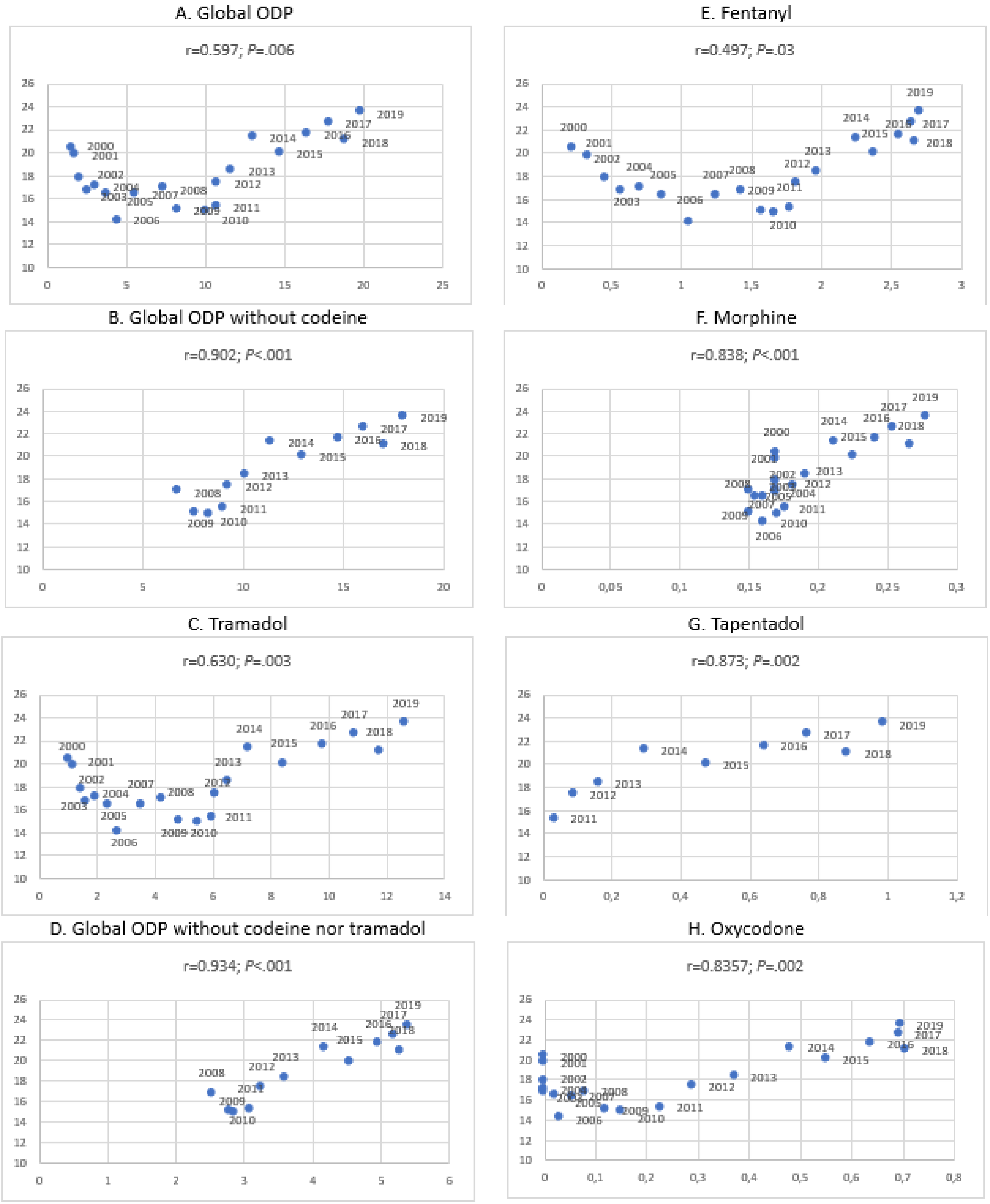
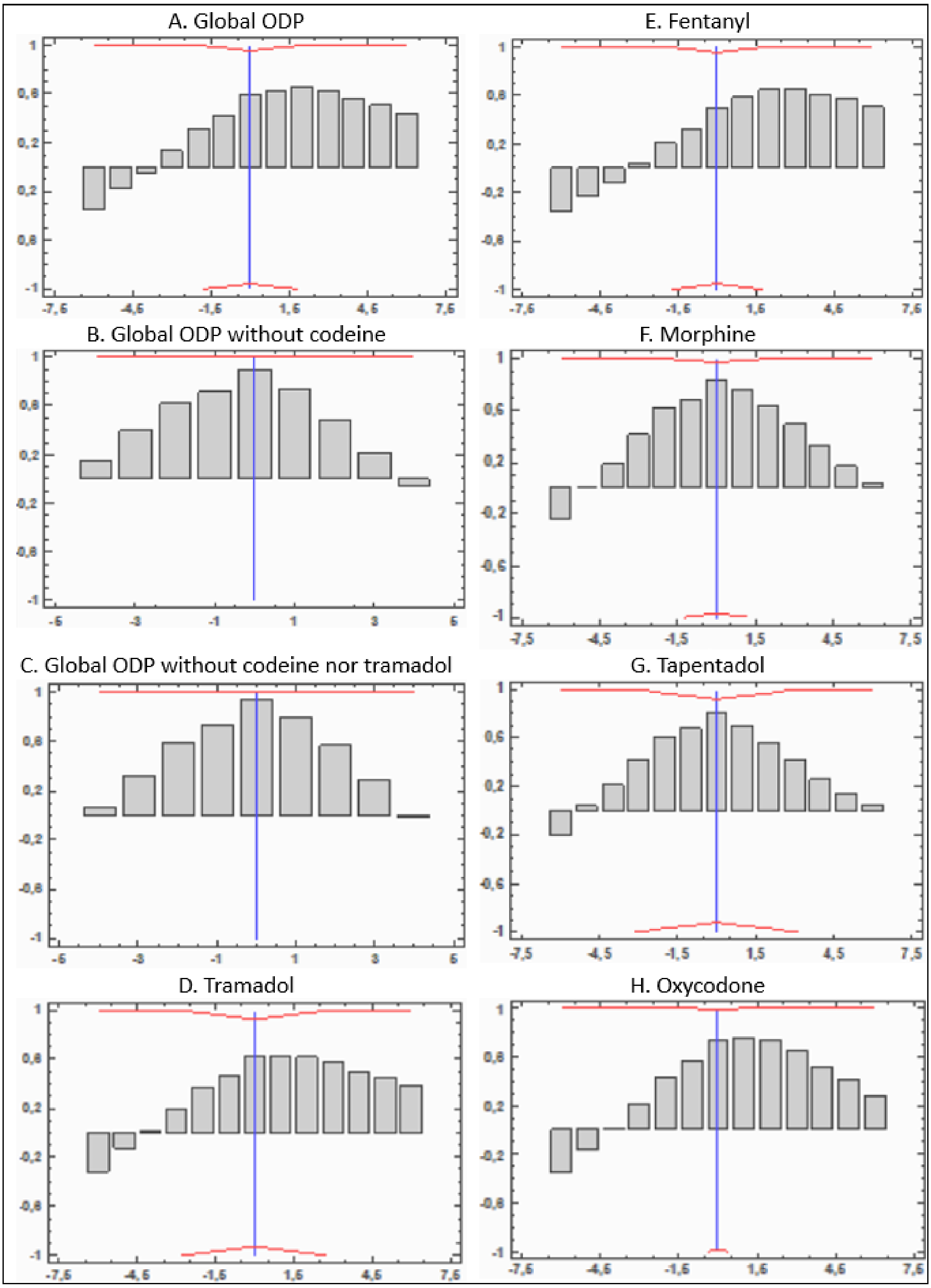


Figure 3. Correlation between opioid drug prescription (ODP; x-axis) and mortality rate per 1,000,000 inhabitants (y-axis).



[Figure 4](#) shows the cross-correlations between mortality rates and the DHD of the different drugs analyzed. In considering all of these drugs, the figure shows a greater correlation between prescribed DHD in a year and the mortality rate 3 years later, reaching a maximum in the second year. However, after eliminating prescribed drugs with codeine and those with codeine and tramadol from the totals, the correlations with the mortality rate can be observed for the same year and not so much for the following years. The same situation occurred when exclusively considering morphine and tapentadol, where the greatest correlations were observed in the same year and in the 2 following years, with a gradual dilution of the correlation in successive years. This same relationship can also be observed with tramadol alone or in conjunction with another medicine. Fentanyl seems to have an impact on mortality over the medium term, in the second or third years, whereas buprenorphine has a long-term impact of between 4 and 6 years. Despite this pattern, these differences were not significant (all $P \geq .05$). The segmented analysis of the periods of 2000-2010 and 2011-2019 as well as 2000-2016 and 2017-2019 did not result in findings different from what is described here.

Figure 4. Cross-correlation between opioid-related mortality rates and opioid drug prescription (ODP). y-axis: correlation; x-axis: delay in years.



Discussion

This study confirms the enormous increase in the prescription of opioid drugs in the last 2 decades in Spain. This increase was driven especially by tramadol, but fentanyl, due to its potency and hospital use, also plays a very relevant role. It also points out a correlation between the increases in ODP and ORM.

Considering the fact that illicit opioid use has decreased or remained the same in the period from 2001 to 2019 [17,20], this could suggest that deaths related to prescribed opioid drugs have increased.

The increase in ODP has been highlighted in the past by other studies carried out in Spain that use different data sources [2-4]. This study includes a larger geographic region, Spain's national territory, in which medical indications are homogenous [28]. This suggests that the progression of ODP is evolving in a parallel way in different regions in Spain [4,17,18]. This increase corresponds with the situation in most other European countries [5].

Many of the causes of this enormous increase have been pointed out. The use of these drugs [11,12,28], as well as their target population, has increased (in large part due to the fact that the aging population experiences more disease and related pain [29]). However, both in Spain and in other countries where this phenomenon is taking place, prescriptions are increasing at a faster rate than these other factors [5,30,31]. More opioids are being prescribed and for longer terms than that recommended by treatment guidelines. This issue is beyond the scope of this study, but it is one in which the role of prescriptions, compared to other factors such as socioeconomic level or residing in rural versus urban areas [18], seems to be especially important given its potential as an area for improvement [32,33].

In addition to describing this general phenomenon, this study shows the correlation with ORM, which coincides with what has been shown to occur in other contexts [33]. In Spain, it appears that in the years included in this study, there has not been a proliferation in illegal consumption of prescribed opioid drugs or other opioids [17,20]. Therefore, it can be inferred that it is the prescribed opioid drugs and not illegal opioids that are provoking this increase in deaths. This suggestion should be taken with great prudence, but it corresponds with what has occurred in environments with greater problems related to illegal opioids [14]. With a lower proportion of illegal opioids, there is a greater correlation between prescriptions and opioid-related mortality. This study found the potential mortality from unstudied opioids to be minimal and the correlation between opioid medications and overdose deaths to be consistent with these occurrences. An absence of correlation might have indicated that illegal drugs were operating in the market. The results tell us that the increase in overdoses is related to an increase in opioid prescribing.

This fact reflects the situation up till 2019, but we should not forget that in the future, prescribed opioid drugs could spread to the illegal market, as has happened in other places [16,21,34]. Above all, precautions should be taken to detect such a situation if it were to occur [20]. The ORM-ODP correlation could be a useful indicator both separately and when harmonized with other indicators such as hospital admissions. In fact, from 2016 to 2019, hospital admissions related to poisoning, other consequences of external causes (S00-T88), and external causes of morbidity (V00-Y99) increased in Spain by 20% [35].

The existence of the ORM-ODP correlation may be as informative as its possible nonexistence in the future. The lack of a correlation could suggest that illegal opioids have increased [14] or that there has been an improvement in prescription practices. Furthermore, opioid-related deaths are an indicator of opioid consumption. It is the least desired adverse effect, but there is a proportionality between its frequency and consumption [20]. Therefore, the surveillance of ORM could provide information about possible variation in other parameters, such as addiction or other related comorbidities (associated infections, dual pathology, etc). Therefore, the surveillance of trends in ORM and ODP separately, as well as the correlations between them, could play an important role in the global monitoring of opioid consumption and its consequences.

In establishing preventive measures, it is useful to focus on certain active ingredients, in addition to prescribed opioid drugs in general, especially in the case of tramadol and fentanyl, in which we found an ORM-ODP correlation. These are 2 medications with different indications and safety profiles.

Tramadol accounts for not only the majority of prescriptions but its increase has also been constant and parallel to total ODP. It belongs to the group of second-tier analgesic opioids as defined by the WHO [27]. The indications of tramadol for non-oncological pain can possibly explain this increase, as suggested in other Spanish studies [18] and coinciding with the European and North American contexts [36]. It is unknown whether this increase corresponds to the increase in its target population. The indications for tramadol and the number of potential patients have increased [11,12]; however, studies of its misuse and misprescription are scarce and contradict each other [36-38]. This study did find a correlation between ORM and tramadol prescriptions; therefore, it is important to reinforce the correct use of tramadol among those who prescribe it [28], as well as to intensify surveillance. It is still the most prescribed opioid drug, where a simple prescription is all that is required [8]. It is much easier to prescribe than others.

In the third analgesic tier [27], fentanyl stands out due to its growth and the ORM-ODP correlation. It is the strongest opioid, specifically recommended for the treatment of irruptive oncological pain. It is also the opioid with the greatest potential number of overdoses, although this also depends on the posology [11,12,39]. This increase is in the context of the notable increase in the consumption of the strongest opioids [2-4,10]. This trend goes beyond the scope of this study. The number of prescriptions of opioids in the third tier increased by 9.6% between 2019 and 2020 [19]. What is surprising is that 89% is related to non-oncological indications [19]. Pain management is an essential component of medical care, but these data are worrisome [3]. Both the increase, as well as the possible adverse effects, have been highlighted in Spain by health authorities, who issued an alert in 2016 to restrain prescriptions to specific indications and the doses to the length of the disease [28]. This study did not find a change over time in the correlations prior to or after that date.

Of the opioid drugs studied, only buprenorphine was not correlated with ORM. It is possible that this could be due to the fact that this opioid is usually used as a substitute treatment for other opioids [40], with a stable number of users [20]. Data on the dispensation of buprenorphine did not differentiate between its use as an analgesic and as treatment for addiction, and as such, it is impossible to interpret these data.

In terms of the temporal patterns described, after eliminating the safest active ingredients from the totals (codeine and codeine with tramadol), the correlations with the mortality rate shifted and could be observed in the same year and less so in successive years. Individually, for morphine, tramadol, and tapentadol, the years with the greatest number of deaths coincided with the years with the greatest number of prescriptions of these drugs. Although these results are not statistically significant (probably because the time series analyzed is short and had few data points [19 data-year/opioid]), they could point to the fact that the ORM-ODP relationship is not mediated by all opioids. They should be studied individually, as they do not share the same indications, safety profiles, or prescription processes. Future studies should study in further depth the prescription of these drugs.

This study includes a number of limitations. Although the data collected correspond to precise definitions that use ICD-10 codes, the reliability of the determination and coding of the cause of death depends on each professional for each death and not on the data source itself. The codes selected are exclusively for deaths due to overdose. There are also specific codes for deaths provoked by opioids in euthanasia or compassionate sedation in the context of palliative care that we did not include. The deaths due to these other causes are not included in what we refer to as ORM. Even so, in some cases, it is possible to omit using the overdose code to avoid legal or administrative problems, especially in the case of accidents. Thus, the number of deaths could be understated. On the other hand, the data did not permit us to disaggregate by sex nor age, which could limit these conclusions. However, these data are acceptable in the international context, as the data source is used by the European Monitoring Centre for Drugs and Drug Addiction [41].

It is important to point out that the Spanish system exclusively includes the main cause of death, and not multiple causes. This differs from other information systems, such as the US system, in which several codes are collected. These codes can be cross-examined to refine the cases in which (1) opioid use has been registered and (2) a death associated with these substances has been recorded. This is the reason why we talk about opioid-related deaths rather than directly attributable causes of mortality in our paper.

This study suggests that at the population level, there is a correlation between greater availability of prescribed opioid drugs and an increase in opioid-related deaths. The correlation between ODP and ORM may be a useful tool in monitoring legal opiates but also possible disturbances in the illegal market. The role of tramadol (an easily prescribed opioid) is important in this correlation, as is that of fentanyl (the strongest opioid). Measures stronger than recommendations need to be taken to reduce off-label prescribing. This study shows that opioid use is not only directly related to the prescribing of opioid drugs above what is desirable but also an increase in deaths.

Acknowledgments

This study has been funded by Instituto de Salud Carlos III (ISCIII) through the project “PI19/00982” and cofunded by the European Union. The authors are grateful to the National Plan on Drugs and Spanish Agency of Medicines and Health Products (*Agencia Española de Medicamentos y Productos Sanitarios* [AEMPS], *Ministerio de Sanidad, España*) for their full support and Emily Felt for translation and suggestions.

Data Availability

The Spanish population data [42] and the data on opioid-related mortality [43] are publicly available at the Spanish National Statistics Institute website.

The data prescribed opioid drugs were obtained via the Spanish Agency of Medicines and Health Products (*Agencia Española de Medicamentos y Productos Sanitarios* [AEMPS]) [44].

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of included codes.
[DOCX File , 12 KB - publichealth_v9i1e43776_app1.docx]

Multimedia Appendix 2
Shapiro-Wilk normality test in prescribed opioid drugs.
[DOCX File , 18 KB - publichealth_v9i1e43776_app2.docx]

References

1. Anderson K, Stowasser D, Freeman C, Scott I. Prescriber barriers and enablers to minimising potentially inappropriate medications in adults: a systematic review and thematic synthesis. BMJ Open 2014 Dec 08;4(12):e006544 [FREE Full text] [doi: 10.1136/bmjopen-2014-006544] [Medline: 25488097]
2. Bigal LM, Bibeau K, Dunbar S. Tramadol prescription over a 4-year period in the USA. Curr Pain Headache Rep 2019 Aug 06;23(10):76. [doi: 10.1007/s11916-019-0777-x] [Medline: 31388761]
3. Bosetti C, Santucci C, Radrezza S, Erthal J, Berterame S, Corli O. Trends in the consumption of opioids for the treatment of severe pain in Europe, 1990-2016. Eur J Pain 2019 Apr 04;23(4):697-707. [doi: 10.1002/ejp.1337] [Medline: 30407692]

4. Boslett AJ, Denham A, Hill EL. Using contributing causes of death improves prediction of opioid involvement in unclassified drug overdoses in US death records. Addiction 2020 Jul 27;115(7):1308-1317 [FREE Full text] [doi: 10.1111/add.14943] [Medline: 32106355]
5. Cahir C, Bennett K, Teljeur C, Fahey T. Potentially inappropriate prescribing and adverse health outcomes in community dwelling older patients. Br J Clin Pharmacol 2014 Jan 20;77(1):201-210 [FREE Full text] [doi: 10.1111/bcp.12161] [Medline: 23711082]
6. Campbell CI, Weisner C, Leresche L, Ray GT, Saunders K, Sullivan MD, et al. Age and gender trends in long-term opioid analgesic use for noncancer pain. Am J Public Health 2010 Dec;100(12):2541-2547 [FREE Full text] [doi: 10.2105/AJPH.2009.180646] [Medline: 20724688]
7. Cheema E, McGuinness K, Hadi MA, Paudyal V, Elnaem MH, Alhifany AA, et al. Causes, nature and toxicology of fentanyl-associated deaths: a systematic review of deaths reported in peer-reviewed literature. J Pain Res 2020 Dec 7;13:3281-3294 [FREE Full text] [doi: 10.2147/JPR.S280462] [Medline: 33324089]
8. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain--United States, 2016. JAMA 2016 Apr 19;315(15):1624-1645 [FREE Full text] [doi: 10.1001/jama.2016.1464] [Medline: 26977696]
9. Drug-related deaths and mortality in Europe: update from the EMCDDA expert network. European Monitoring Centre for Drugs and Drug Addiction. 2019. URL: https://tinyurl.com/3ukaeyh2 [accessed 2023-06-15]
10. Feingold D, Goldberger N, Haklai Z, Lev-Ran S. Fatal overdoses of opioids in Israel 2005-2014. Eur Addict Res 2017 Dec 21;23(6):276-283. [doi: 10.1159/000485033] [Medline: 29268270]
11. Garcia del Pozo J, Carvajal A, Viloria JM, Velasco A, Garcia del Pozo V. Trends in the consumption of opioid analgesics in Spain. higher increases as fentanyl replaces morphine. Eur J Clin Pharmacol 2008 Apr 20;64(4):411-415. [doi: 10.1007/s00228-007-0419-9] [Medline: 18157671]
12. Gomes T, Kitchen SA, Murray R. Measuring the burden of opioid-related mortality in Ontario, Canada, during the COVID-19 pandemic. JAMA Netw Open 2021 May 03;4(5):e2112865 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.12865] [Medline: 34037734]
13. Gomes T, Tadrous M, Mamdani MM, Paterson JM, Juurlink DN. The burden of opioid-related mortality in the United States. JAMA Netw Open 2018 Jun 01;1(2):e180217 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.0217] [Medline: 30646062]
14. González-Bermejo D, Rayón-Iglesias P, Rodríguez-Pascual A, Álvarez-Gutiérrez A, Fernández-Dueñas A, Montero-Corominas D, et al. Drug utilization study on immediate release fentanyl in Spain. prevalence, incidence, and indication. Pharmacoepidemiol Drug Saf 2021 Mar 14;30(3):371-378. [doi: 10.1002/pds.5118] [Medline: 32929809]
15. Grigoras CA, Karanika S, Velmahos E, Alevizakos M, Flokas M, Kaspiris-Rousellis C, et al. Correlation of opioid mortality with prescriptions and social determinants: a cross-sectional study of Medicare enrollees. Drugs 2018 Jan 20;78(1):111-121. [doi: 10.1007/s40265-017-0846-6] [Medline: 29159797]
16. Álvarez Mazariegos JA, Calvete Waldomar S, Fernández-Marcote Sánchez-Mayoral RM, Guardia Serecigni J, Henche Ruiz AI, Lligoña Garreta A, et al. Guía de consenso para el buen uso de analgésicos opioides: gestión de riesgos y beneficios. Socidrogalcohol. 2017. URL: https://tinyurl.com/yb492hvj [accessed 2023-06-15]
17. Häuser W, Morlion B, Vowles KE, Bannister K, Buchser E, Casale R, et al. European* clinical practice recommendations on opioids for chronic noncancer pain - part 1: role of opioids in the management of chronic noncancer pain. Eur J Pain 2021 May 02;25(5):949-968 [FREE Full text] [doi: 10.1002/ejp.1736] [Medline: 33655607]
18. Hurtado I, García-Sempere A, Peiró S, Sanfélix-Gimeno G. Increasing trends in opioid use from 2010 to 2018 in the region of Valencia, Spain: a real-world, population-based study. Front Pharmacol 2020 Dec 11;11:612556 [FREE Full text] [doi: 10.3389/fphar.2020.612556] [Medline: 33362564]
19. Jannetto PJ. The North American Opioid Epidemic. Ther Drug Monit 2021 Feb 01;43(1):1-5. [doi: 10.1097/FTD.0000000000000817] [Medline: 33337587]
20. Jiménez-Fernández B, Calomarde-Gómez C, López-Lazcano A, Lligoña A, Gual A, López-Pelayo H. Systematic review on the clinical management of chronic pain and comorbid opioid use disorder. Adicciones 2022 Apr 20:1680 [FREE Full text] [doi: 10.20882/adicciones.1680] [Medline: 35472158]
21. Jones W, Kurdyak P, Fischer B. Examining correlations between opioid dispensing and opioid-related hospitalizations in Canada, 2007-2016. BMC Health Serv Res 2020 Jul 22;20(1):677 [FREE Full text] [doi: 10.1186/s12913-020-05530-w] [Medline: 32698815]
22. Larney S, Peacock A, Tran L, Stockings E, Santomauro D, Santo T, et al. All-cause and overdose mortality risk among people prescribed opioids: a systematic review and meta-analysis. Pain Med 2020 Dec 25;21(12):3700-3711. [doi: 10.1093/pm/pnaa214] [Medline: 32951045]
23. Martínez Oró DP. Opioides en España. Ni repunte de heroína ni crisis de opioides a la americana. Opioids in Spain. Neither a heroin spike nor an American-style opioid crisis. Episteme. Investigación e Intervención social. 2019. URL: https://tinyurl.com/237jr84k [accessed 2023-06-15]
24. Portal Estadístico. Área de Inteligencia de Gestión. Ministerio de Sanidad. Gobierno de España. URL: https://tinyurl.com/2acyjwfm [accessed 2022-09-27]

25. Oliva A, Armas N, Dévora S, Abdala S. Opioid use trends in Spain: the case of the island of La Gomera (2016-2019). Naunyn Schmiedebergs Arch Pharmacol 2022 Feb 09;395(2):217-226 [FREE Full text] [doi: 10.1007/s00210-021-02193-0] [Medline: 34882264]
26. Pulido J, Sanchez-Niubo A, Llorens N, Hoyos J, Barrio G, Belza MJ, et al. Estimating the prevalence of recreational opioid use in Spain using a multiplier method. Int J Environ Res Public Health 2022 Apr 15;19(8):4815 [FREE Full text] [doi: 10.3390/ijerph19084815] [Medline: 35457681]
27. Qeadan F, Madden E, Mensah N, Tingey B, Herron J, Hernandez-Vallant A, et al. Epidemiological trends in opioid-only and opioid/polysubstance-related death rates among American Indian/Alaska Native populations from 1999 to 2019: a retrospective longitudinal ecological study. BMJ Open 2022 May 02;12(5):e053686 [FREE Full text] [doi: 10.1136/bmjopen-2021-053686] [Medline: 35501103]
28. Regueras E, Guzmán JL. Prescriptions of opioid medicines in Spain betwen on 2019 and 2020: what medical specialities are prescribing them and in what indications. Multidisciplinary Pain Journal 2021 Apr 5:5-12 [FREE Full text]
29. Reines SA, Goldmann B, Harnett M, Lu L. Misuse of tramadol in the United States: an analysis of the National Survey of Drug Use and Health 2002-2017. Subst Abuse 2020 Jun 03;14:1178221820930006 [FREE Full text] [doi: 10.1177/1178221820930006] [Medline: 32547049]
30. Richards GC, Aronson JK, Mahtani KR, Heneghan C. Global, regional, and national consumption of controlled opioids: a cross-sectional study of 214 countries and non-metropolitan territories. Br J Pain 2022 Feb 04;16(1):34-40 [FREE Full text] [doi: 10.1177/20494637211013052] [Medline: 35111312]
31. Rostam-Abadi Y, Gholami J, Amin-Esmaeili M, Safarcherati A, Mojtabai R, Ghadirzadeh MR, et al. Tramadol use and public health consequences in Iran: a systematic review and meta-analysis. Addiction 2020 Dec 25;115(12):2213-2242. [doi: 10.1111/add.15059] [Medline: 32196801]
32. Ruiz-López D, Alonso-Babarro A. Opioid consumption in the Community of Madrid (Spain) between 2004 and 2014. Rev Clin Esp (Barc) 2019 Oct;219(7):367-374. [doi: 10.1016/j.rce.2019.02.002] [Medline: 30914142]
33. Salazar A, Moreno S, de Sola H, Moral-Munoz J, Dueñas M, Failde I. The evolution of opioid-related mortality and potential years of life lost in Spain from 2008 to 2017: differences between Spain and the United States. Curr Med Res Opin 2020 Feb;36(2):285-291. [doi: 10.1080/03007995.2019.1684251] [Medline: 31635485]
34. Use of opioids in Spain (2010-2019). Observatory on medication use. Spanish Agency of Medicines and Medical Devices (AEMPS). 2020. URL: https://www.aemps.gob.es/medicamentos-de-uso-humano/observatorio-de-uso-de-medicamentos/utilizacion-de-medicamentos-opioides-en-espana/ [accessed 2020-09-15]
35. Official website of the Spanish National Statistics Institute. Spanish National Statistics Institute. 2018. URL: http://www.ine.es/ [accessed 2023-06-15]
36. Statistics of deaths according to the cause of death. methodology. Spanish National Statistics Institute. 2017. URL: https://www.ine.es/daco/daco42/sanitarias/metodologia_00.pdf [accessed 2023-06-15]
37. Spreen LA, Dittmar EN, Quirk KC, Smith MA. Buprenorphine initiation strategies for opioid use disorder and pain management: a systematic review. Pharmacotherapy 2022 May 25;42(5):411-427 [FREE Full text] [doi: 10.1002/phar.2676] [Medline: 35302671]
38. Storms H, Marquet K, Aertgeerts B, Claes N. Prevalence of inappropriate medication use in residential long-term care facilities for the elderly: a systematic review. Eur J Gen Pract 2017 Dec 08;23(1):69-77 [FREE Full text] [doi: 10.1080/13814788.2017.1288211] [Medline: 28271916]
39. World Drug Report 2021. United Nations Office on Drugs and Crime. 2021. URL: https://www.unodc.org/unodc/data-and-analysis/wdr2021.html [accessed 2022-06-15]
40. ATC/DDD Index 2022. World Health Organization Collaborating Centre for Drug Statistics Methodology. 2022. URL: https://www.whocc.no/atc_ddd_index/ [accessed 2022-09-27]
41. World Health Organization. ICD-10: International Statistical Classification of Diseases and Related Health Problems. Geneva, Switzerland: World Health Organization; 2004.
42. Population figures. Spanish National Statistics Institute. URL: https://www.ine.es/dyngs/INEbase/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=resultados&idp=1254735572981 [accessed 2023-06-15]
43. Death statistics according to cause of death. Spanish National Statistics Institute. URL: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176780&menu=resultados&idp=1254735573175#!tabs-1254736194710 [accessed 2023-06-15]
44. Use of opioid drugs in Spain. Spanish Agency of Medicines and Health Products (AEMPS). URL: https://www.aemps.gob.es/medicamentos-de-uso-humano/observatorio-de-uso-de-medicamentos/utilizacion-de-medicamentos-opioides-en-espana/ [accessed 2023-06-15]

Abbreviations

AEMPS: Agencia Española de Medicamentos y Productos Sanitarios (Spanish Agency of Medicines and Health Products)
DHD: dose per 1000 inhabitants per day

ICD-10: International Classification of Diseases, 10th Revision
ODP: opioid drug prescription
ORM: opioid-related mortality
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 24.10.22; peer-reviewed by H de Sola, JM Suelves, H Jin; comments to author 06.04.23; revised version received 08.05.23; accepted 26.05.23; published 28.06.23.

Please cite as:
Salazar A, Moreno-Pulido S, Prego-Meleiro P, Henares-Montiel J, Pulido J, Donat M, Sotres-Fernandez G, Sordo L
Correlation Between Opioid Drug Prescription and Opioid-Related Mortality in Spain as a Surveillance Tool: Ecological Study
JMIR Public Health Surveill 2023;9:e43776
URL: https://publichealth.jmir.org/2023/1/e43776
doi:10.2196/43776
PMID:37379061

Original Paper

Safety Signal Generation for Sudden Sensorineural Hearing Loss Following Messenger RNA COVID-19 Vaccination: Postmarketing Surveillance Using the French Pharmacovigilance Spontaneous Reporting Database

Hung Thai-Van[1,2,3*], MD, PhD; Marie-Blanche Valnet-Rabier[4*], PharmD, PhD; Maëva Anciaux[2], MD; Aude Lambert[5], PharmD; Anaïs Maurier[6], PharmD; Judith Cottin[7], PharmD; Tessa Pietri[8], PharmD; Alexandre Destère[9], PharmD; Marlène Damin-Pernik[10], PharmD; Fanny Perrouin[11], PharmD; Haleh Bagheri[12], PharmD, PhD

[1]Service d’Audiologie & d’Explorations Otoneurologiques, Hospices civils de Lyon, Lyon, France
[2]Département d'audiologie et d'Explorations otoneurologiques, Hôpital Universitaire de lyon 1, Villeurbanne, France
[3]Institut de l'Audition, Institut Pasteur, Paris, France
[4]Regional Pharmacovigilance Center of Franche-Comté, Centre Hospitalier Universitaire de Besançon, Besançon, France
[5]Regional Pharmacovigilance Center of Strasbourg, Hôpitaux Universitaires de Strasbourg, Strasbourg, France, Strasbourg, France
[6]Service de Pharmacosurveillance, Centre Régional de Pharmacovigilance du Centre Val de Loire, Centre Hospitalier Universitaire, Tours. France, Tours, France
[7]Regional Pharmacovigilance Center of Lyon, Service Hospitalo-Universitaire de Pharmaco-Toxicologie, Hospices Civils de Lyon, Lyon, France, Lyon, France
[8]Regional Pharmacovigilance Center of Marseille, Service de Pharmacologie Clinique et Pharmacovigilance, Aix Marseille Université, Assistance Publique Hôpitaux de Marseille, Marseille, France, Marseille, France
[9]Department of Pharmacology and Pharmacovigilance Center, Côte d'Azur University Medical Center, Nice, France, Nice, France
[10]Regional Pharmacovigilance Center of Saint-Etienne, Centre Hospitalier Universitaire de Saint-Étienne, Saint-Étienne, France, Saint Etienne, France
[11]Regional Pharmacovigilance Center of Nantes, Service de Pharmacologie Clinique, Nantes, France, Nantes, France
[12]Department of Medical and Clinical Pharmacology, Regional Pharmacovigilance Center of Toulouse, CIC1436, Toulouse University Hospital, Toulouse, France
*these authors contributed equally

Corresponding Author:
Haleh Bagheri, PharmD, PhD
Department of Medical and Clinical Pharmacology
Regional Pharmacovigilance Center of Toulouse, CIC1436
Toulouse University Hospital
37 Allées Jules Guesde
Toulouse, 31000
France
Phone: 33 561145902
Fax: 33 561255116
Email: haleh.bagheri@univ-tlse3.fr

Abstract

Background: The World Health Organization recently described sudden sensorineural hearing loss (SSNHL) as a possible adverse effect of COVID-19 vaccines. Recent discordant pharmacoepidemiologic studies invite robust clinical investigations of SSNHL after COVID-19 messenger RNA (mRNA) vaccines. This postmarketing surveillance study, overseen by French public health authorities, is the first to clinically document postvaccination SSNHL and examine the role of potential risk factors.

Objective: This nationwide study aimed to assess the relationship between SSNHL and exposure to mRNA COVID-19 vaccines and estimate the reporting rate (Rr) of SSNHL after mRNA vaccination per 1 million doses (primary outcome).

Methods: We performed a retrospective review of all suspected cases of SSNHL after mRNA COVID-19 vaccination spontaneously reported in France between January 2021 and February 2022 based on a comprehensive medical evaluation, including the evaluation of patient medical history, side and range of hearing loss, and hearing recovery outcomes after a minimum period of 3 months. The quantification of hearing loss and assessment of hearing recovery outcomes were performed according to a grading system modified from the Siegel criteria. A cutoff of 21 days was used for the delay onset of SSNHL. The primary outcome was estimated using the total number of doses of each vaccine administered during the study period in France as the denominator.

Results: From 400 extracted cases for tozinameran and elasomeran, 345 (86.3%) spontaneous reports were selected. After reviewing complementary data, 49.6% (171/345) of documented cases of SSNHL were identified. Of these, 83% (142/171) of SSNHL cases occurred after tozinameran vaccination: Rr=1.45/1,000,000 injections; no difference for the rank of injections; complete recovery in 22.5% (32/142) of cases; median delay onset before day 21=4 days (median age 51, IQR 13-83 years); and no effects of sex. A total of 16.9% (29/171) of SSNHL cases occurred after elasomeran vaccination: Rr=1.67/1,000,000 injections; rank effect in favor of the first injection (*P*=.03); complete recovery in 24% (7/29) of cases; median delay onset before day 21=8 days (median age 47, IQR 33-81 years); and no effects of sex. Autoimmune, cardiovascular, or audiovestibular risk factors were present in approximately 29.8% (51/171) of the cases. SSNHL was more often unilateral than bilateral for both mRNA vaccines (*P*<.001 for tozinameran; *P*<.003 for elasomeran). There were 13.5% (23/142) of cases of profound hearing loss, among which 74% (17/23) did not recover a serviceable ear. A positive rechallenge was documented for 8 cases.

Conclusions: SSNHL after COVID-19 mRNA vaccines are very rare adverse events that do not call into question the benefits of mRNA vaccines but deserve to be known given the potentially disabling impact of sudden deafness. Therefore, it is essential to properly characterize postinjection SSNHL, especially in the case of a positive rechallenge, to provide appropriate individualized recommendations.

(JMIR Public Health Surveill 2023;9:e45263) doi:10.2196/45263

KEYWORDS

mRNA COVID-19 vaccine; COVID-19; messenger RNA; tozinameran; elasomeran; sudden sensorineural hearing loss; audiogram; positive rechallenge; spontaneous reporting; postmarketing; surveillance; pharmacovigilance

Introduction

Background

SARS-CoV-2, the coronavirus responsible for COVID-19 disease, has spread rapidly, causing a global pandemic since early 2020. In late 2020, the first messenger RNA (mRNA)–based COVID-19 vaccine, Pfizer-BioNTech BNT162b2, Comirnaty (tozinameran), was authorized by the European Medicines Agency (EMA), followed by Moderna mRNA-1273, Spikevax (elasomeran). The monitoring of COVID-19 vaccines has been a major challenge in verifying their effectiveness in the general population and identifying any adverse drug reactions (ADRs) not yet observed in clinical trials. Postmarketing spontaneous pharmacovigilance systems provide a gold mine of data for drug safety research. Pharmacovigilance systems are designed to identify emerging signals using multiple data sources. Current pharmacovigilance is predominantly based on spontaneous reporting and mainly helpful in detecting type B effects (effects that are often allergic or idiosyncratic reactions; characteristically occur in only a minority of patients; are usually unrelated to dosage; and are serious, unexpected, and unpredictable) [1]. Signals may arise from one or multiple sources and may suggest a potentially new causal association or new aspect of a known risk of a drug. Indeed, because of the design of clinical trials, rare or delayed ADRs and those occurring in specific subpopulations could be identified only during postmarketing surveillance. Despite the use of many data sources to detect emerging signal, spontaneous reports continue to be the cornerstone of signal detection for most marketed drugs or vaccines, as they have been since the 1960s [2]. This challenge in COVID-19 vaccine surveillance seemed very important because of the large number of individuals (hundreds of thousands) exposed to these vaccines. In France, the National Agency for the Safety of Medicines and Health Products (*Agence Nationale de Sécurité du Médicament et des produits de santé* [ANSM]) is responsible for assessing the safety of vaccines through the continuous surveillance of all ADRs to quickly detect potential safety issues and take appropriate risk reduction measures. For the COVID-19 vaccination campaign, an enhanced surveillance system has been put in place [3]. If a safety signal was validated, the EMA Pharmacovigilance Risk Assessment Committee (PRAC) would make recommendations for any necessary changes to the product information or any further steps. For example, in May 2021, based on the cases reported in clinical trials and from vaccination campaigns, the EMA PRAC validated the potential signal about delayed injection site reaction for elasomeran and concluded that this information should be added to the product information of Spikevax [4]. Moreover, on October 28, 2022, the EMA PRAC recommended that heavy menstrual bleeding be added to the Summary of Product Characteristics of the mRNA vaccines (tozinameran and elasomeran) as a side effect of unknown frequency. This signal was supported by the ANSM's requested review of all spontaneous reports in European countries via EudraVigilance—the EMA's system for the management and analysis of information on suspected ADRs in the European Economic Area [5].

Sudden sensorineural hearing loss (SSNHL) is defined as a sensorineural hearing loss >30 dB over at least 3 consecutive audiometric frequencies occurring within <72 hours [6]. In cases where it occurs in the absence of reference audiometry, the diagnosis is made by comparing the hearing thresholds of the affected ear with those of the contralateral ear. SSNHL affects 5 to 20 individuals per 100,000 person-years [7,8], mainly those aged between 40 and 54 years [8,9]. Early detection and management of SSNHL are crucial to maximizing hearing recovery and the quality of life [10]. Although approximately 98% of SSNHL cases are unilateral [11], both their clinical presentation and prognosis vary greatly depending on the severity of the hearing loss [6], as well as the presence or absence of associated otoneurological manifestations such as tinnitus or balance disorders [6,11]. In most cases, there is no identifiable etiology, and SSNHL is classified as idiopathic [12]. The most likely pathophysiological mechanism for the development of idiopathic SSNHL is thought to be cochlear injury due to poor vascular perfusion or viral infection [13]. Disorders of the cochlear microcirculation due to altered plasma viscosity, platelet aggregability, red blood cell deformability, and endothelial dysfunction have been reported in patients with SSNHL [14,15].

According to the available literature, post–COVID-19 SSNHL is much less common than anosmia [16,17]. This may explain why post–COVID-19 SSNHL has been underdiagnosed by otolaryngologists compared with anosmia. However, several in-depth case studies have linked SSNHL cases to the COVID-19 outbreak [18-30]. The SARS-CoV-2 spike glycoprotein is known to have a particular tropism for human angiotensin-converting enzyme 2 receptors [31], which are expressed not only in the lung respiratory epithelium and the olfactory neuroepithelium but also in the vascular endothelium [32]. Accordingly, it was hypothesized that hearing loss in patients with COVID-19 might result from vascular endotheliitis with microthrombosis in the cochlea, auditory nerve, or hearing centers located in the temporal lobe [33]. Another mechanism advanced is the spread of viral meningitis from the subarachnoid space to the inner ear via the cochlear aqueduct, as supported by magnetic resonance imaging scans in clinical cases combining deafness and neurological symptoms [34]. A possible neurological manifestation of SARS-CoV-2 is the Guillain-Barré syndrome, which may induce auditory neuropathy [35]. Regardless of the pathophysiological mechanisms involved, it is interesting to note that increased high-frequency pure tone thresholds and reduced transient-evoked otoacoustic emissions, suggestive of cochlear outer hair cell damage, have been reported in asymptomatic patients who are COVID-19 positive [36].

SSNHL has recently emerged as a possible adverse effect of COVID-19 vaccines. Sporadic cases of unilateral SSNHL occurring shortly after the first or second administration of adenoviral vector– or mRNA-based vaccines have been recently reported, either isolated or in association with other cochleovestibular disorders (tinnitus, dizziness, and vertigo), in patients of both sexes with or without a history of audiovestibular, autoimmune, or cardiovascular diseases [37-40]. In February 2022, the World Health Organization (WHO) raised a signal detection about cases of SSNHL after COVID-19 vaccination, prompting the need for investigation. As of February 22, 2021, a total of 172 cases of hearing loss after COVID-19 vaccination (including 142 cases after tozinameran and 15 cases after elasomeran) were reported from 10 countries [41]. To date, 2 large-scale studies have investigated SSNHL after COVID-19 vaccine administration and reported conflicting results. One study, using data collected between December 2020 and July 2021 in the United States by the Centers for Disease Control and Prevention Vaccine Adverse Events Reporting System (a passive surveillance system), found no difference between the estimated incidence of SSNHL after COVID-19 vaccination with tozinameran, elasomeran, and ChAdOx1-S (0.6 to 28.0 cases per 100,000 person-years) and the historical incidence of SSNHL (11 to 77 cases per 100,000 person-years) [42,43]. By contrast, a retrospective cohort study conducted by the largest state-mandated health service organizations from Israel on data collected from December 2020 to May 2021 concluded that there is a proven, albeit small, risk (<1 per 100,000 vaccinated individuals) of SSNHL following immunization with tozinameran [44]. Recently, another retrospective cohort study performed on the Finish register–based countrywide data did not find a potential association between COVID-19 vaccinations and SSNHL. However, the results of these 3 studies were based on diagnosis codes in administrative databases, and the authors suggested the need for studies with more detailed clinical data with a more objective diagnosis of SSNHL [45]: none of these studies described the clinical characteristics of postvaccination SSNHL in terms of severity or duration or examined the facilitating role of potential risk factors, thereby prompting further research [46]. Moreover, the reliability of passive surveillance system databases when it comes to verifying SSNHL diagnostic criteria has been questioned [47]. To our knowledge, there is no large-scale, data-driven study investigating the cases of SSNHL after mRNA COVID-19 vaccine exposure identified using robust clinical measures of severity and recovery.

Objectives

After several spontaneous reports of postvaccination SSNHL to the pharmacovigilance system and requests from patients and physicians for documented data regarding the possibility of revaccination (dose 2 or 3) after a first event of post–mRNA vaccine SSNHL, a national population-based active surveillance study was conducted to perform a rigorous and objective analysis of spontaneously reported cases and assess this new potential signal.

Methods

Study Design and Settings

The French pharmacovigilance system was established at the end of the 1970s based on a decentralized network of Regional Pharmacovigilance Centers (RPVCs), with the mission of identifying, evaluating, and preventing adverse drug events after marketing authorization. The ANSM is a decisional authority and has mobilized the French network of 30 RPVCs to ensure the continuous monitoring and evaluation of ADRs of COVID-19 vaccines through the daily evaluation of spontaneous reports. Health care professionals or users can report potential ADRs following COVID-19 vaccines (similar to how they would following the administration of other drugs) directly to the RPVC or using a dedicated web portal [48]. Each report of a potential ADR is assessed through a careful clinical, chronological, semiological, and pharmacological analysis before being registered (anonymous registration) in a centralized French Pharmacovigilance Database (FPVDB). For this purpose, an intensive pharmacovigilance survey was set up, and 6 RPVC experts were appointed to prepare weekly pharmacovigilance reports for the first marketed vaccines (4 experts for tozinameran and 2 experts for elasomeran). In addition, a scientific monitoring committee was also set up, comprising ANSM staff members, pharmacovigilance experts in charge of the survey, and specialized medical experts (cardiologists; neurologists; ear, nose, and throat specialists; etc) for a collegial evaluation of the reports submitted daily [3]. In this study, we conducted a nationwide population-based surveillance study using this enhanced pharmacovigilance system to assess a potential signal for post–mRNA COVID-19 vaccine SSNHL.

Data Collection

We conducted a retrospective study to collect all ADRs recorded in the FPVDB occurring any time after mRNA COVID-19 vaccine administration from the beginning of the vaccination campaign (December 28, 2020) to February 2, 2022. Eligible SSNHL cases were selected using the Medical Dictionary for Regulatory Activity hierarchy, with preferred term selected from the narrow Standardized Medical Queries "Hearing impairment" [49]. The ANSM provided a line listing with all data related to ADR and the medical history generated by each RPVC. In addition, we requested medical records, audiometry examinations, magnetic resonance imaging, and all substantial examination results for each case. Two audiology experts analyzed all data according to the American Academy of Otolaryngology-Head and Neck Surgery Foundation guidelines [11]. The clinical analysis was performed on the basis of demographics (sex and age), risk factors (autoimmune, cardiovascular, or audiovestibular factors), the side of hearing loss (one or both ears affected), the presence of tinnitus or balance disorders, the degree of hearing loss (calculated by averaging hearing thresholds at 0.5, 1, 2, and 4 kHz) [50], the onset of SSNHL, and the rank of vaccination (first injection, second injection, or booster), with a particular focus on positive rechallenge cases. Following recent research on SSNHL [51,52], patients were categorized according to a grading system modified from the Siegel criteria [53]. This allowed both the quantification of hearing loss and assessment of hearing recovery outcomes using the same objective standards in all patients. The quantification of hearing loss included 5 grades covering all possible degrees of deafness: slight (grade 1), mild (grade 2), moderate to moderately severe (grade 3), severe (grade 4), and profound (grade 5). Recovery may be complete, partial, slight, or absent, or SSNHL may result in a nonserviceable ear (Tables 1 and 2).

Table 1. Modified Siegel criteria for hearing loss (HL) grades.

Level of HL	HL range (dB)
Grade 1	≤25
Grade 2	26-40
Grade 3	41-70
Grade 4	71-90
Grade 5	>90

Table 2. Modified Siegel criteria for hearing recovery outcomes.

Level of hearing recovery	Final hearing level (dB), gain (dB)
Complete recovery	≤25
Partial recovery	26-40, >15
Slight improvement	41-70, >15
No improvement	71-90, <15
Nonserviceable ear	>90

We classified the cases as medically or nonmedically documented. After a detailed review of complementary medical data, we excluded cases for which the attributable causes were different from vaccination or those that were nonevaluable because of missing data (nonmedically documented). Finally, the cases fulfilling the criteria of SSNHL were included and distinguished using a cutoff of 21 days. This cutoff value was recommended by the Safety Platform for Emergency Vaccines and the Brighton Collaboration guidelines for sensorineural hearing loss [54]. According to previous research on the auditory [43,44] or nonauditory [55] adverse events of mRNA vaccination, exposure could be defined as vaccination with an mRNA COVID-19 vaccine within the previous 21 days, and patients vaccinated beyond that date were considered unexposed. However, given the very limited number of large-scale studies available on the auditory adverse effects of mRNA COVID-19 vaccination, we also considered the spontaneous reporting of SSNHL after day 21 in the first instance.

A minimum of 3 months of follow-up was conducted (follow-up until August 1) for each SSNHL case to collect missing data and assess the potential hearing loss recovery.

Outcome Measure

Descriptive analysis was performed for the 2 mRNA COVID-19 vaccines (tozinameran and elasomeran) for all the items cited earlier. The primary outcome was the reporting rate (Rr) of SSNHL following mRNA COVID-19 vaccine administration per 1 million doses, which was calculated using the total number of mRNA COVID-19 vaccine doses administered over the study period as the denominator according to the month of administration. Since the beginning of the French vaccination campaign against COVID-19 to early February 2022, a total of 97,840,529 doses of tozinameran and 22,690,889 doses of elasomeran have been administered in France [56].

Statistical Analysis

Patient characteristics were expressed as count (percentage) for categorical variables and median (IQR) for quantitative variables according to the administered vaccine and for the whole cohort because of the few elasomeran cases. The Rr was expressed as cases per 1 million vaccine injections. The number of reported SSNHL cases and the number of vaccine injections were estimated by month during the study period. Finally, we performed a binomial test comparing a sample proportion with a hypothetical proportion to analyze the potential association of the Rr of SSNHL with sex, vaccination rank, and unilaterality or bilaterality of hearing loss (significance level: *P*<.05; R [version 4.2.1; R Foundation for Statistical Computing]).

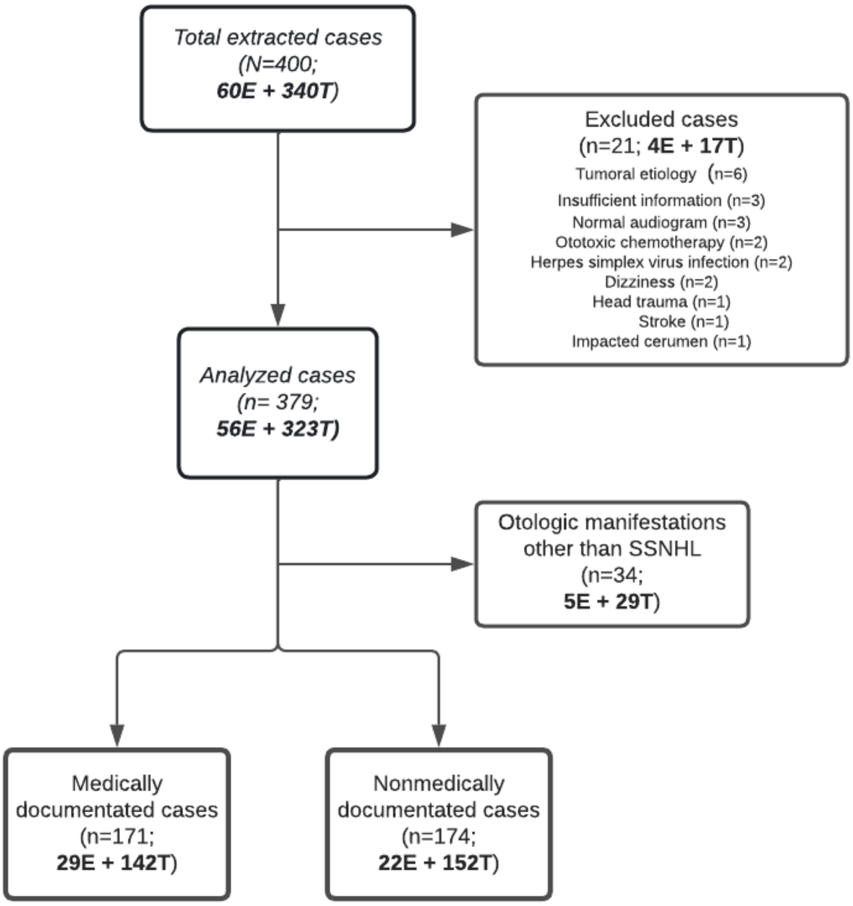
Ethics Approval

This nationwide postmarketing surveillance study was conducted with the authorization of the *Commission nationale de l'informatique et des libertés* (CNIL; National Commission on Informatics and Liberty; n 2014-302). The authorization was granted specifically for the French pharmacovigilance spontaneous reporting database placed under the responsibility of the ANSM.

Results

During the study period, 400 cases of SSNHL after COVID-19 vaccine administration (n=340, 85% for tozinameran and n=60, 15% for elasomeran) were registered in the FPVDB. After case review, a total of 42.8% (171/400) of cases fulfilled the criteria of SSNHL (Figure 1). Hearing assessment was available either through an initial audiogram test (118/171, 69% of cases) or through an expert medical evaluation performed at the follow-up (53/171, 30.9% of cases).

Figure 1. Flowchart for selection of cases of sudden sensorineural hearing loss (SSNHL) following messenger RNA (mRNA) vaccines. E: elasomeran; T: tozinameran.



Description of Post–Tozinameran Vaccine SSNHL Cases

A total of 323 cases of SSNHL after tozinameran vaccine administration were selected, of which 142 (43.9%) cases met the criteria for inclusion; 29 (8.9%) cases were consistent with otologic manifestations other than SSNHL, and 152 (47.1%) cases were not sufficiently documented to draw conclusions (Figure 1).

For these 142 cases, the delay onset was ≤21 days for 108 (76.1%) cases, with a median delay onset of 4 (range 2-9) days. Of these 142 cases, 84 (59.2%) were those of female patients. The median age was 51 (range 13-83) years, and 69% (98/142) of patients were aged 30 to 64 years. A total of 35.2% (50/142) of patients had a medical history, which was otoneurological in 11.9% (17/142) of cases. The vaccination rank was known for 88% (125/142) of cases, with the first injection involved in 42.3% (60/142) of cases. Steroids were administered orally in 47.2% (67/142) of cases. SSNHL was unilateral in 87.3% (124/142) of cases. The hearing loss grade ranged between 1 and 3 in 75.4% (107/142) of cases. Tinnitus was associated with SSNHL in 52.8% (75/142) of cases and vertigo in 28.9% (41/142) of cases. Complete recovery was observed in 22.5% (32/142) of cases, whereas no improvement was observed in 53.5% (76/142) of cases and nonserviceable ear in 8.5% (12/142) of cases (Tables 3 and 4).

Deafness was more often unilateral than bilateral (*P*<.001). No effect of sex or vaccination rank was observed. Case follow-up identified 3.5% (5/142) of cases of positive rechallenge (Table 5): four were those of female vaccine recipients and 1 was that of a male vaccine recipient (age range 32-72 years). SSNHL was unilateral in most cases, was of grades 2 (37/142, 26.1%) to 3 (46/142, 32.4%), occurred rapidly (<10 days) after the first and second doses in all cases, and completely resolved in 2.1% (3/142) of cases during the 3-month follow-up.

Table 3. Characteristics of sudden sensorineural hearing loss cases.

Characteristics	Tozinameran			Elasomeran		
	<21 days (n=108, 76.1%)	>21 days (n=34, 23.9%)	Total (n=142)	<21 days (n=26, 86.7%)	>21 days (n=3, 10.3%)	Total (n=29)
Sex, n (%)						
Male	43 (39.8)	15 (44.1)	58 (40.8)	13 (50)	1 (33.3)	14 (48.3)
Female	65 (60.2)	19 (55.9)	84 (59.2)	13 (50)	2 (66.7)	15 (51.7)
Age (years)						
Values, median (range)	50 (13-83)	51 (16-72)	51 (13-83)	51 (33-81)	43 (36-68)	47 (33-81)
0-18, n (%)	3 (2.8)	1 (2.9)	4 (2.8)	0 (0)	0 (0)	0 (0)
19-29, n (%)	7 (6.5)	0 (0)	7 (4.9)	0 (0)	0 (0)	0 (0)
30-49, n (%)	43 (39.8)	12 (35.3)	55 (38.7)	10 (38.5)	2 (66.7)	12 (41.4)
50-64, n (%)	26 (24.1)	17 (50)	43 (30.3)	9 (34.6)	0 (0)	9 (31)
65-74, n (%)	19 (17.6)	4 (11.8)	23 (16.2)	6 (23.1)	1 (33.3)	7 (24.1)
≥75, n (%)	10 (9.3)	0 (0)	10 (7)	1 (3.8)	0 (0)	1 (3.4)
Medical history, n (%)	38 (35.2)	12 (35.3)	50 (35.2)	7 (26.9)	2 (66.7)	9 (31)
CV[a]	9 (8.3)	5 (14.7)	14 (9.9)	2 (7.7)	1 (33.3)	3 (10.3)
ON[b]	12 (11.1)	5 (14.7)	17 (12)	3 (11.5)	1 (33.3)	4 (13.8)
AIM[c]	7 (6.5)	2 (5.9)	9 (6.3)	0 (0)	0 (0)	0 (0)
CV and ON	4 (3.7)	0 (0)	4 (2.8)	0 (0)	0 (0)	0 (0)
ON and AIM	3 (2.8)	0 (0)	3 (2.1)	1 (3.8)	0 (0)	1 (3.4)
Other etiologies	3 (2.8)	0	3 (2.1)	1 (3.8)	0 (0)	1 (3.4)
Time to onset[d] (days), median (IQR)	4 (2-9)	41 (25-67)	N/A[e]	8 (1-21)	50 (26-144)	N/A
Vaccination rank, n (%)						
First dose	47 (43.5)	13 (38.2)	60 (42.2)	12 (46.2)	1 (33.3)	13 (44.8)
Second dose	39 (36.1)	14 (41.2)	53 (37.3)	10 (38.5)	2 (66.7)	12 (41.4)
Booster	12 (11.1)	0 (0)	12 (8.5)	2 (7.7)	0 (0)	2 (6.9)
Unknown	10 (9.3)	7 (20.6)	17 (12)	2 (7.7)	0 (0)	2 (6.9)
Oral steroid administration, n (%)	48 (44.4)	19 (55.9)	67 (47.2)	14 (53.8)	2 (66.7)	16 (55.2)
Laterality, n (%)						
Unilateral	94 (87)	30 (88.2)	124 (87.3)	20 (76.9)	2 (66.7)	22 (75.9)
Bilateral	14 (13)	4 (11.8)	18 (12.7)	6 (23.1)	1 (33.3)	7 (24.1)
Hearing loss grades, n (%)						
Grade 1	22 (20.4)	2 (5.9)	24 (16.9)	4 (15.4)	3 (100)	7 (24.1)
Grade 2	27 (25)	10 (29.4)	37 (26.1)	5 (19.2)	0 (0)	5 (17.2)
Grade 3	35 (32.4)	11 (32.4)	46 (32.4)	7 (26.9)	0 (0)	7 (24.1)
Grade 4	13 (12)	6 (17.6)	19 (13.4)	3 (11.5)	0 (0)	3 (10.3)
Grade 5	11 (10.2)	5 (14.7)	16 (11.3)	7 (26.9)	0 (0)	7 (24.1)
Associated cochleovestibular disorders, n (%)						
Tinnitus	59 (54.6)	16 (47.1)	75 (52.8)	8 (30.8)	2 (66.7)	10 (34.5)
Vertigo and balance disorders	33 (30.6)	8 (23.5)	41 (28.9)	11 (42.3)	1 (33.3)	12 (41.4)
Time to recovery (days), median (range)	15 (5-67.5)	11 (8-22.5)	N/A	21 (2-90)	ND[f]	N/A
Positive rechallenge, n (%)	5 (4.6)	0 (0)	5 (3.5)	2 (7.7)	1 (33.3)	3 (10.3)

[a] CV: cardiovascular.

[b] ON: otoneurological.

[c] AIM: autoimmune disease.

[d] Interval between vaccine administration and symptom onset.

[e] N/A: not applicable.

[f] ND: not determined.

Table 4. Hearing recovery outcomes by grade of hearing loss.

Hearing loss	Hearing outcome									
	Tozinameran (n=142), n (%)					Elasomeran (n=29), n (%)				
	Complete recovery	Partial recovery	Slight improvement	No improvement	Nonserviceable ear	Complete recovery	Partial recovery	Slight improvement	No improvement	Nonserviceable ear
Grade 1	12 (8.5)	0 (0)	0 (0)	12 (8)	0 (0)	4 (13.8)	0 (0)	0 (0)	3 (10.3)	0 (0)
Grade 2	9 (6.3)	5 (3.5)	0 (0)	23 (16)	0 (0)	2 (6.9)	1 (3.4)	0 (0)	2 (6.9)	0 (0)
Grade 3	10 (7.1)	6 (4.2)	1 (0.7)	29 (20)	0 (0)	1 (3.4)	2 (6.9)	0 (0)	4 (13.8)	0 (0)
Grade 4	1 (0.7)	3 (2.1)	3 (2.1)	12 (8)	0 (0)	0 (0)	1 (3.4)	0 (0)	2 (6.9)	0 (0)
Grade 5	0 (0)	0 (0)	4 (2.8)	0	12 (8.5)	0 (0)	2 (6.9)	0 (0)	0 (0)	5 (17.2)

Table 5. Characteristics of the positive rechallenge cases.

Sex	Age (years)	Medical history	Vertigo	Tinnitus	Vaccination rank and delay onset (days)	Hearing loss side	Hearing loss grade[a]	Final hearing recovery
Tozinameran								
Female	61	None	Yes	Yes	• D1[b]—4 • D2[c]—7 • B[d]—7	Unilateral	3	No improvement
Male	47	None	No	Yes	• D1—9 • D2—9	Unilateral	3	No improvement
Female	74	Fluctuating hearing loss and vertigo punctually treated with steroids	No	Yes	• D1—8 • D2—9	Unilateral	3	Complete recovery
Female	32	Protein S deficiency	No	No	• D1—8 • D2—2	Unilateral	2	Complete recovery
Female	79	Diabetes mellitus stabilized since 1977, left Ménière' disease since 1977, hypothyroidism stabilized, high blood pressure stabilized since 1985, and hypercholesterolemia stabilized since 1985	Yes	Yes	• D1—8 • D2—8 • B—R-[e]	Unilateral (contralateral to the Ménière side)	2	Complete recovery
Elasomeran								
Female	45	None	No	Yes	• D1—5 • D2—<1 • B—R-[e]	Bilateral	2	Complete recovery
Male	71	None	No	Yes	• D1—1 • D2—1	Unilateral	2	Complete recovery
Female	41	High blood pressure	No	Yes	• D2—60 B—40	Bilateral	2	No improvement

[a] Measured after injection.

[b] D1: first vaccine injection.

[c] D2: second vaccine injection.

[d] B: booster.

[e] R-: Negative rechallenge (no recurrence of sudden sensorineural hearing loss; the booster of elasomeran was administered at half dose).

Description of Post–Elasomeran Vaccine SSNHL Cases

A total of 56 cases of SSNHL after elasomeran vaccine administration were selected, of which 29 (49%) cases met the inclusion criteria; 5 (8%) cases were consistent with otologic manifestations other than SSNHL, and 22 (37%) cases were not sufficiently documented to draw conclusions (Figure 1).

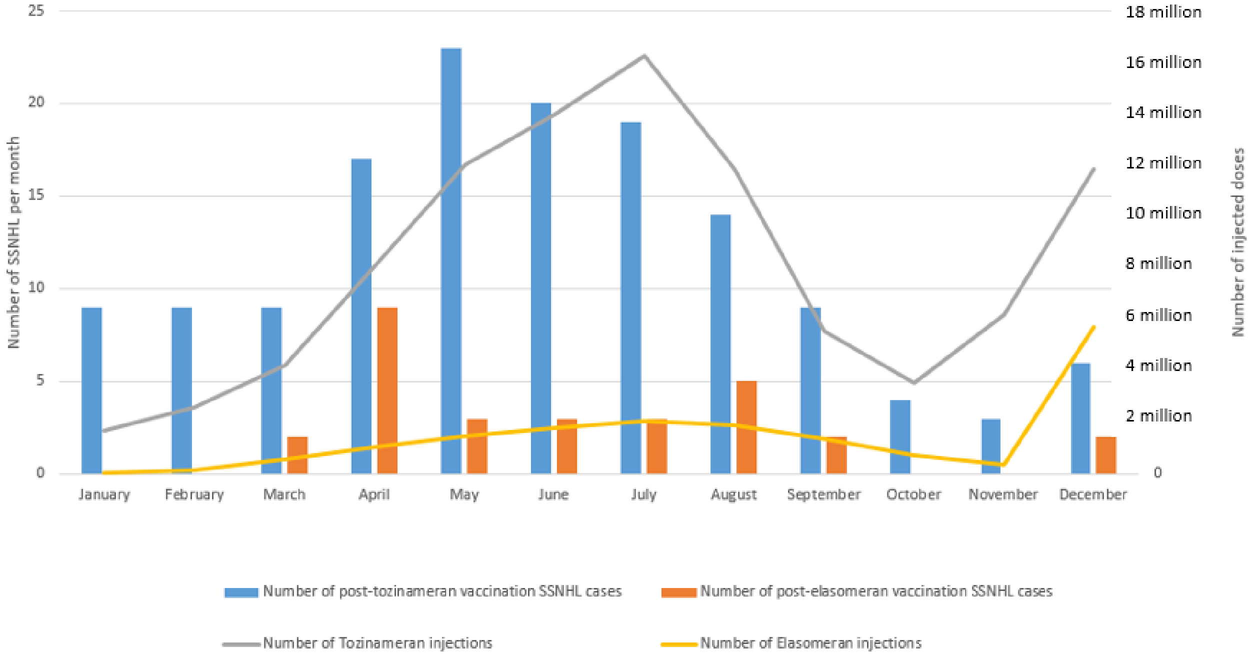
Among these 29 cases, the delay onset was ≤21 days for 26 (90%) cases, with a median delay onset of 8 (range 1-21) days, and 15 (52%) were those of female patients. The median age was 47 (range 33-81) years, and 72% (21/29) of patients were aged 30 to 64 years. A total of 31% (9/29) of patients had a medical history, which was otoneurological in 14% (4/29) of cases. The vaccination rank was known for 27 cases, with the first injection involved in 45% (13/29) of cases. Steroids were administered orally in 55% (16/29) of cases. SSNHL was unilateral in 76% (22/29) of cases. The hearing loss grade ranged from 1 to 3 in 66% (19/29) of cases. Tinnitus was associated with SSNHL in 34% (10/29) of cases and vertigo in 41% (12/29) of cases. Complete recovery was observed in 24% (7/29) of cases, whereas no improvement was observed in 38% (11/29) of cases and nonserviceable ear in 18% (5/29) of cases (Tables 3 and 4).

Deafness was more often unilateral than bilateral (*P*<.003). A rank effect was found in favor of the first dose (*P*<.03). No effects of sex or side of hearing loss were observed. Case follow-up identified 3 cases of positive rechallenge (Table 5): two (67%) were those of female vaccine recipients and 1 (3%) was that of a male vaccine recipient (age range 41-71 years). SSNHL was unilateral in most cases, was of grades 2 (5/29, 18%) to 3 (7/29, 24%), occurred rapidly (<10 days) after the first and second doses in 2 cases (in the third case, delay onset was >21 days, but with positive rechallenge with the booster), and completely resolved in 2 cases during the 3-month follow-up.

Rate of Spontaneous Reports

Figure 2 shows, by month, the number of spontaneous reports of SSNHL cases after tozinameran vaccine administration and after elasomeran vaccine administration, as well as the number of injections of each vaccine in 2021. The total Rr seemed similar for both vaccines and was estimated at 1.45/1,000,000 doses for tozinameran and 1.67/1,000,000 doses for elasomeran. During this period, a total of 86,993 and 20,265 spontaneous reports were registered for tozinameran and elasomeran, respectively, in the FPVDB, of which 22,619 (26%) and 3782 (18.66%) were serious cases, respectively. As a result, SSNHL cases accounted for 0.63% and 0.76% of the total serious adverse events recorded for tozinameran and elasomeran, respectively.

Figure 2. Reporting rate of sudden sensorineural hearing loss (SSNHL) cases for the 2 messenger RNA (mRNA) vaccines according to the month. Solid lines: the number of tozinameran (gray) and elasomeran (yellow) injections expressed in million doses (right y-axis) per month. Bar graph: the number of tozinameran-related (blue) and elasomeran-related (orange) SSNHL cases expressed in case number (left y-axis) per month.



Discussion

Principal Findings

The main objective of this nationwide postmarketing surveillance study was to conduct a continuous assessment of the safety of mRNA COVID-19 vaccines to generate potential safety signals related to serious and typical adverse effects and consider relevant measures that should be taken quickly [3]. Our results—based on all spontaneous reporting of mRNA vaccine–induced SSNHL registered in the FPVDB during >1 year, supported by intensive pharmacovigilance monitoring, and combined with administrative data on the doses of vaccines administered—suggested that the occurrence of SSNHL after mRNA COVID-19 vaccination could be defined as a very rare adverse effect, as it was estimated to be <2 per 1 million injections for both vaccines. The occurrence of SSNHL cases was found to parallel the number of injections of the most widely administered mRNA vaccine in France over the study period, with tozinameran accounting for 80% of the injections. This has not been verified for elasomeran injections, probably because of the smaller number of people exposed to this vaccine and the smaller number of SSNHL cases reported. The scarcity of this adverse effect supports the appropriateness of spontaneous reporting allowing a signal generation. Spontaneous reporting is particularly useful for identifying rare and idiosyncratic safety concerns. For example, 12 (57%) out of 21 drug withdrawals in France between 1998 and 2004 were solely based on spontaneous reports. Similarly, 8 (73%) out of 11 product withdrawals between 1999 and 2001 in the United Kingdom and the United States were based on early information from spontaneous reports. Some studies suggest that spontaneous reports are most useful for newly marketed products, specifically within 3 years after launch [2]. A larger investigation and disproportionality analysis of the spontaneous reporting of WHO database (Vigibase) could be useful to generate a signal. Moreover, a multicentric case-control study, methodologically appropriate to investigate rare events, might be of interest to conclude an association between mRNA COVID-19 vaccination and SSNHL, considering the different risk factors for SSNHL, as it has already been done for myocarditis or pericarditis [57,58]. That said, the Rr of mRNA COVID-19 vaccine–induced SSNHL cases remained below the estimated incidence of SSNHL cases in the general population (approximately 5 to 20 cases per 100,000 person-years in France). In addition, unlike cases of mRNA COVID-19 vaccine–induced myocarditis [58] and because of the lack of data, it should be difficult to compare the rate of SSNHL cases related to mRNA COVID-19 vaccine administration with that presumably related to COVID-19 infections.

Another contribution of this study is the first clinical description of SSNHL as an adverse effect of mRNA COVID-19 vaccination. Because strict diagnostic criteria were used, the risk of false-positive results when identifying SSNHL cases was reduced compared with previous research using passive surveillance system databases [47]. To name a few, conductive or posttraumatic hearing loss, vestibular neuritis, labyrinthitis, and fluctuating hearing loss related to a history of Ménière disease or acoustic neurinoma were excluded. Nevertheless, the clinical presentation of mRNA COVID-19 vaccine–induced SSNHL cases could include audiovestibular symptoms such as tinnitus, vertigo, or balance disorders associated with deafness. Although postvaccination hearing loss was most often found to be unilateral and of grades 1 to 3 (ie, slight to moderately severe hearing loss), there were also cases of bilateral and severe to profound hearing loss that may have required specific treatment in the absence of complete recovery for both vaccines studied. It is interesting to note that at the end of the follow-up period (minimal duration: 3 months), a complete recovery was achieved in less than one-quarter of the cases. Whereas SSNHL secondary to vaccination was found to be a very rare adverse event, it is of note that 10% of affected patients were diagnosed with a nonserviceable ear, which represents a potential indication for cochlear implantation. The rate of complete or partial recovery varied between 41% and 52% for hearing loss grades 1 to 3, whereas the rate of complete or partial recovery was 23% for grade 4 and only 9% for grade 5. This observation is in agreement with the evolution of SSNHL reported outside the vaccination context [51].

However, the rate of recovery could be greater because of missing feedback from patients who recovered completely. The most affected age group was intermediate (30-64 years), and the median age was approximately 50 years for both vaccines, which is consistent with the usual age of onset of idiopathic SSNHL [9,10]. Thus, there was no evidence for a more frequent occurrence of this adverse effect for the extreme age groups. Autoimmune, cardiovascular, audiological, or otoneurological risk factors were present in only one-third of the cases, definitively preventing the identification of a specific patient profile.

Furthermore, we should underline our findings related to the delay onset of SSNHL according to the vaccination rank. A rank effect was found only for elasomeran: the occurrence of SSNHL was higher after the first injection than after the second injection and the booster. In agreement with Safety Platform for Emergency Vaccines and Brighton Collaboration guidelines on postvaccinal SSNHL, as well as with previous data on audiological or cardiac adverse events with mRNA COVID-19 vaccines, we found that the median delay onset fell within the first 2 weeks after injection [42,43,54,57,58]. Wichova et al [39] also reported an average delay onset of 10.18 (SD 9) days, whereas the median (range) delay onset in the WHO report was 1 day (few hours to 19 days) [41].

Moreover, data supported a possible relationship between exposure to mRNA COVID-19 vaccines and the occurrence of SSNHL. Positive rechallenge cases were documented for 5 tozinameran and 3 elasomeran vaccinations, meaning that repeated exposure to a COVID-19 mRNA-based vaccine led to SSNHL for a minority of patients who had previously experienced the same adverse event. In a pharmacovigilance study, this result alone was sufficient to suggest a high probability of a causal relationship [59].

Since the beginning of the COVID-19 outbreak, the association between SARS-CoV-2 infection and the occurrence of SSNHL has been supported exclusively by cross-sectional studies conducted on limited case series [60]. Almost all reported cases were unilateral sudden deafness [26,28], which could be associated with disabling ipsilateral tinnitus [16,24] or vertigo with nystagmus beating toward the contralateral side [28]. Regardless of the clinical presentation, the estimation of the exact incidence of SSNHL attributable to COVID-19 infection may have been hampered by the difficulty in applying the same criteria for causal relationship across studies. These criteria could include, in addition to the certainty of viral infection by polymerase chain reaction test or the titration of specific antibodies, the temporal concordance between the infection and the onset of hearing loss; the presence of symptoms and signs consistent with concomitant labyrinthitis or neuritis; and, above all, the exclusion of any other reason that could be counted among the causes of SSNHL [61]. Recent updates have pointed out the need for the long-term documentation of audiovestibular symptoms in large cohorts of SARS-CoV-2–infected patients before conclusions can be drawn about their true incidence [62-64]. On the basis of the existing literature, SSNHL can be considered so far as a rare complication of COVID-19 infection [21,65,66].

Limitations

Our study has some limitations. An important step in assessing vaccine-related adverse events is to compare their rates in vaccinated and unvaccinated populations, as was done for myocarditis [58]. Because the studied adverse event is very rare, large population-based pharmacovigilance studies (Vigibase) or a case-control study appropriate for rare events is needed to further define the role of mRNA COVID-19 vaccination in the occurrence of SSNHL, as has been done for COVID-19 vaccines and menstrual cycle [67]. Herein, we conducted a "case series study," a well-known design in pharmacoepidemiology that differs from a cohort study or a case-control study. As such, our nationwide postmarketing surveillance did not include any control group, thus preventing the consideration of potentially important variables such as age or SSNHL time to onset. After drug marketing, case series are most useful for 2 related purposes: first, to quantify the incidence or rate of an adverse reaction and second, to detect any particular adverse effect not yet identified in clinical trials because of its scarcity [68]. Moreover, we cannot rule out the underreporting of ADRs, which may have biased the estimated rate of SSNHL. However, following the introduction of mRNA COVID-19 vaccines in France, a major effort has been made by the enhanced pharmacovigilance system to improve the reporting of any suspected ADR following the administration of these vaccines by health professionals or patients. By February 2022, approximately 80% of the French population had received the first 2 vaccine injections, and 52.6% had also received the booster [69]. The impressive number of reports of suspected ADRs (>80,000 between January 2021 and January 2022 in France) suggests that underreporting may have been very rare, particularly for potentially serious adverse events [3].

Other authors have retrospectively compared the incidence of SSNHL cases in the population exposed to mRNA COVID-19 vaccines with that in the general population using a population-based cohort and International Classification of Diseases, Ninth Revision, codes without any data on patients' medical history or risk factors [38,39]. In our clinical experience, this could lead to a selection of false-positive SSNHL cases. As previously done by Ouldali et al [70] for mRNA COVID-19 vaccine–induced pediatric inflammatory multisystem syndrome, we chose a restrictive method to select and describe medically confirmed SSNHL cases based on pharmacovigilance and complementary medical data, which is the added value of our survey. Nevertheless, the lack of available data for the general population precluded comparisons with historical incidence or incidence in the unvaccinated population. Another possible limitation is the underdiagnosis of SSNHL cases related to the COVID-19 outbreak due to the difficulties in accessing health care during the COVID-19 outbreak. However, disabling deafness remains a medical emergency and should lead to hospital admission.

Conclusions

This study identified very rare but severe cases of medically confirmed SSNHL after the administration of mRNA COVID-19 vaccines without calling into question the demonstrated benefit of vaccinating a large population in the context of rapid SARS-CoV-2 circulation. On the basis of our findings, particularly those related to the cases of positive rechallenge and grade-5 hearing loss, patients should undergo a comprehensive evaluation of any otologic manifestation occurring after COVID-19 vaccination. Early diagnosis of SSNHL is essential for positive outcomes. Each medical evaluation should be performed for 2 reasons: the prompt initiation of appropriate treatment to improve the prognosis of SSNHL and provision of expert counseling to patients regarding subsequent vaccine injections, given the evolution of SARS-CoV-2, its variants, and new adaptive vaccines. These data are new and of critical interest to patients, physicians, and regulatory policies to tailor public health messages and medical care.

Acknowledgments

The authors would like to acknowledge the support of French National Agency for the Safety of Medicines and Health Products (*Agence Nationale de Sécurité du Médicament et des produits de santé*).

The authors would like to thank all Regional Pharmacovigilance Centers (RPVCs) from the French Network of Regional Pharmacovigilance who generously provided their time for active surveillance and audiogram collection during this 1-year study: RPVCs Amiens, Angers, Bordeaux, Brest, Caen, Clermont-Ferrand, Dijon, Grenoble, Lille, Limoges, Nancy, Paris Hôpital Européen Georges-Pompidou, Paris Henri Mondor, Paris Fernand Widal, Paris Pitié Salpétrière, Paris Saint-Antoine, Paris Cochin-Port Royal, Poitiers, Reims, Rennes, and Rouen. The authors acknowledge the support of the Paris Hearing Institute from *Fondation pour l'Audition* (FPA IDA09) and are grateful to Hélène Boyer (*Direction de la Recherche en Santé*, *Hospices Civils de Lyon*, France) for carefully proofreading this manuscript. The authors acknowledge Samuel Tessier for statistical analysis.

The authors are grateful to all staff members of the French National Agency for the Safety of Medicines and Health Products (*Agence Nationale de Sécurité du Médicament et des produits de santé*) who interacted with them from the initiation to the completion of the study.

Data Availability

Data are available upon reasonable request to MBV-R and HB.

Authors' Contributions

HT-V, MBV-R, and HB contributed equally to this study. They had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. HT-V, MBV-R, and HB conceptualized and designed the study. All authors were involved in acquisition, analysis, and interpretation of the data. HT-V, MBV-R, MA, and HB drafted the manuscript. All authors critically revised the manuscript for important intellectual content. MBV-R and HB performed the statistical analysis. HT-V obtained the funding. AL, AM, JC, TP, AD, MD-P, and FP provided administrative, technical, or material support. HT-V, MBV-R, and HB supervised the study.

Conflicts of Interest

None declared.

References

1. Meyboom RH, Egberts AC, Edwards IR, Hekster YA, de Koning FH, Gribnau FW. Principles of signal detection in pharmacovigilance. Drug Saf 1997 Jun;16(6):355-365. [doi: 10.2165/00002018-199716060-00002] [Medline: 9241490]

2. Bate A, Hornbuckle K, Juhaeri J, Motsko SP, Reynolds RF. Hypothesis-free signal detection in healthcare databases: finding its value for pharmacovigilance. Ther Adv Drug Saf 2019;10:2042098619864744 [FREE Full text] [doi: 10.1177/2042098619864744] [Medline: 31428307]
3. Benkebil M, Gautier S, Gras-Champel V, Massy N, Micallef J, Valnet Rabier M. COVID-19 vaccines surveillance in France: a global response to a major national challenge. Anaesth Crit Care Pain Med 2021 Jun;40(3):100866 [FREE Full text] [doi: 10.1016/j.accpm.2021.100866] [Medline: 33895471]
4. COVID-19 vaccine safety update. European Medicines Agency. 2014 Jul 14. URL: https://tinyurl.com/yc7mbvj7 [accessed 2022-12-12]
5. Meeting highlights from the Pharmacovigilance Risk Assessment Committee (PRAC) 24 - 27 October 2022. European Medicines Agency. 2022 Oct 28. URL: https://www.ema.europa.eu/en/news/meeting-highlights-pharmacovigilance-risk-assessment-committee-prac-24-27-october-2022mittee-prac-24-27-october-2022 [accessed 2022-12-15]
6. Byl FM. Sudden hearing loss: eight years' experience and suggested prognostic table. Laryngoscope 1984;94(5):647-661. [doi: 10.1288/00005537-198405000-00014]
7. Mosnier I, Bouccara D, Atassi-Dumont M, Sterkers O. [Treatments of idiopathic sudden sensorineural hearing loss: retrospective study of 144 cases]. Rev Laryngol Otol Rhinol (Bord) 1998;119(2):119-128. [Medline: 9770055]
8. Charrier JB, Tran Ba Huy P. [Idiopathic sudden sensorineural hearing loss: a review]. Ann Otolaryngol Chir Cervicofac 2005 Feb;122(1):3-17. [doi: 10.1016/s0003-438x(05)82312-6] [Medline: 15851940]
9. Sciancalepore PI, de Robertis V, Sardone R, Quaranta N. Sudden sensorineural hearing loss: what factors influence the response to therapy? Audiol Res 2020 Jul 07;10(1):234 [FREE Full text] [doi: 10.4081/audiores.2020.234] [Medline: 32944207]
10. Yigider AP, Keskin M, Kufeciler L, Kocak HE. Topography of the lesion in idiopathic sudden sensorineural hearing loss. Acta Otolaryngol 2020 Nov 12;140(11):909-913. [doi: 10.1080/00016489.2020.1813328] [Medline: 32921217]
11. Chandrasekhar SS, Tsai Do BS, Schwartz SR, Bontempo LJ, Faucett EA, Finestone SA, et al. Clinical practice guideline: sudden hearing loss (update). Otolaryngol Head Neck Surg 2019 Aug;161(1_suppl):S1-45. [doi: 10.1177/0194599819859885] [Medline: 31369359]
12. Fetterman BL, Saunders JE, Luxford WM. Prognosis and treatment of sudden sensorineural hearing loss. Am J Otol 1996 Jul;17(4):529-536. [Medline: 8841697]
13. Kuhn M, Heman-Ackah SE, Shaikh JA, Roehm PC. Sudden sensorineural hearing loss: a review of diagnosis, treatment, and prognosis. Trends Amplif 2011 Sep 22;15(3):91-105 [FREE Full text] [doi: 10.1177/1084713811408349] [Medline: 21606048]
14. Schreiber BE, Agrup C, Haskard DO, Luxon LM. Sudden sensorineural hearing loss. Lancet 2010 Apr;375(9721):1203-1211. [doi: 10.1016/s0140-6736(09)62071-7]
15. Quaranta N, De Ceglie V, D'Elia A. Endothelial dysfunction in idiopathic sudden sensorineural hearing loss: a review. Audiol Res 2016 Apr 20;6(1):151 [FREE Full text] [doi: 10.4081/audiores.2016.151] [Medline: 27588164]
16. Cashman KA, Wilkinson ER, Zeng X, Cardile AP, Facemire PR, Bell TM, et al. Immune-mediated systemic vasculitis as the proposed cause of sudden-onset sensorineural hearing loss following lassa virus exposure in cynomolgus macaques. mBio 2018 Oct 30;9(5):01896-18 [FREE Full text] [doi: 10.1128/mBio.01896-18] [Medline: 30377282]
17. Passali GC, Bentivoglio AR. Comment to the article "Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study". Eur Arch Otorhinolaryngol 2020 Aug 7;277(8):2391-2392 [FREE Full text] [doi: 10.1007/s00405-020-06024-5] [Medline: 32383095]
18. de Melo GD, Lazarini F, Levallois S, Hautefort C, Michel V, Larrous F, et al. COVID-19-related anosmia is associated with viral persistence and inflammation in human olfactory epithelium and brain infection in hamsters. Sci Transl Med 2021 Jun 02;13(596):eabf8396 [FREE Full text] [doi: 10.1126/scitranslmed.abf8396] [Medline: 33941622]
19. Koumpa FS, Forde CT, Manjaly JG. Sudden irreversible hearing loss post COVID-19. BMJ Case Rep 2020 Oct 13;13(11):e238419 [FREE Full text] [doi: 10.1136/bcr-2020-238419] [Medline: 33051251]
20. Sriwijitalai W, Wiwanitkit V. Hearing loss and COVID-19: a note. Am J Otolaryngol 2020 May;41(3):102473 [FREE Full text] [doi: 10.1016/j.amjoto.2020.102473] [Medline: 32276732]
21. Cure E, Cumhur Cure M. Comment on "hearing loss and COVID-19: a note". Am J Otolaryngol 2020 Jul;41(4):102513 [FREE Full text] [doi: 10.1016/j.amjoto.2020.102513] [Medline: 32386897]
22. Karimi-Galougahi M, Naeini AS, Raad N, Mikaniki N, Ghorbani J. Vertigo and hearing loss during the COVID-19 pandemic – is there an association? Acta Otorhinolaryngol Ital 2020 Dec;40(6):463-465. [doi: 10.14639/0392-100x-n0820]
23. Lamounier P, Franco V, Ramos HV, Gobbo DA, Teixeira RP, dos Reis PC, et al. A 67-year-old woman with sudden hearing loss associated with SARS-CoV-2 infection. Am J Case Rep 2020 Oct 05;21:e927519. [doi: 10.12659/ajcr.927519]
24. Kilic O, Kalcioglu MT, Cag Y, Tuysuz O, Pektas E, Caskurlu H, et al. Could sudden sensorineural hearing loss be the sole manifestation of COVID-19? An investigation into SARS-COV-2 in the etiology of sudden sensorineural hearing loss. Int J Infect Dis 2020 Aug;97:208-211 [FREE Full text] [doi: 10.1016/j.ijid.2020.06.023] [Medline: 32535294]
25. Lang B, Hintze J, Conlon B. Coronavirus disease 2019 and sudden sensorineural hearing loss. J Laryngol Otol 2020 Oct 01;134(11):1026-1028. [doi: 10.1017/s0022215120002145]

26. Umashankar A, Prakash P, Prabhu P. Sudden sensorineural hearing loss post coronavirus disease: a systematic review of case reports. Indian J Otolaryngol Head Neck Surg 2022 Oct 10;74(Suppl 2):3028-3035 [FREE Full text] [doi: 10.1007/s12070-021-02756-w] [Medline: 34277383]
27. Beckers E, Chouvel P, Cassetto V, Mustin V. Sudden sensorineural hearing loss in COVID-19: a case report and literature review. Clin Case Rep 2021 Apr 11;9(4):2300-2304 [FREE Full text] [doi: 10.1002/ccr3.4019] [Medline: 33936683]
28. Ricciardiello F, Pisani D, Viola P, Cristiano E, Scarpa A, Giannone A, et al. Sudden sensorineural hearing loss in mild COVID-19: case series and analysis of the literature. Audiol Res 2021 Jul 01;11(3):313-326 [FREE Full text] [doi: 10.3390/audiolres11030029] [Medline: 34287226]
29. Saniasiaya J. Hearing loss in SARS-CoV-2: what do we know? Ear Nose Throat J 2021 Apr 05;100(2_suppl):152S-154S [FREE Full text] [doi: 10.1177/0145561320946902] [Medline: 32755405]
30. Lorenzo-Villalba N, Pierre L, Guerrero-Niño J, Jannot X, Andrès E. Hearing loss and sixth cranial nerve paresis after COVID-19. Eur J Case Rep Intern Med 2022;9(2):003221 [FREE Full text] [doi: 10.12890/2022_003221] [Medline: 35265560]
31. Chern A, Famuyide A, Moonis G, Lalwani A. Bilateral sudden sensorineural hearing loss and intralabyrinthine hemorrhage in a patient with COVID-19. Otol Neurotol 2021 Jan;42(1):e10-e14 [FREE Full text] [doi: 10.1097/MAO.0000000000002860] [Medline: 33301283]
32. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 2003 Nov 27;426(6965):450-454 [FREE Full text] [doi: 10.1038/nature02145] [Medline: 14647384]
33. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, et al. Endothelial cell infection and endotheliitis in COVID-19. Lancet 2020 May 02;395(10234):1417-1418 [FREE Full text] [doi: 10.1016/S0140-6736(20)30937-5] [Medline: 32325026]
34. Harenberg J, Jonas JB, Trecca EM. A liaison between sudden sensorineural hearing loss and SARS-CoV-2 infection. Thromb Haemost 2020 Sep 28;120(9):1237-1239. [doi: 10.1055/s-0040-1714370] [Medline: 32725610]
35. Degen C, Lenarz T, Willenborg K. Acute profound sensorineural hearing loss after COVID-19 pneumonia. Mayo Clin Proc 2020 Aug;95(8):1801-1803 [FREE Full text] [doi: 10.1016/j.mayocp.2020.05.034] [Medline: 32753155]
36. Ellul MA, Benjamin L, Singh B, Lant S, Michael BD, Easton A, et al. Neurological associations of COVID-19. Lancet Neurol 2020 Sep;19(9):767-783. [doi: 10.1016/s1474-4422(20)30221-0]
37. Mustafa M. Audiological profile of asymptomatic Covid-19 PCR-positive cases. Am J Otolaryngol 2020 May;41(3):102483 [FREE Full text] [doi: 10.1016/j.amjoto.2020.102483] [Medline: 32307189]
38. Jeong J, Choi HS. Sudden sensorineural hearing loss after COVID-19 vaccination. Int J Infect Dis 2021 Dec;113:341-343 [FREE Full text] [doi: 10.1016/j.ijid.2021.10.025] [Medline: 34670143]
39. Wichova H, Miller M, Derebery M. Otologic manifestations after COVID-19 vaccination: the house ear clinic experience. Otol Neurotol 2021 Oct 01;42(9):e1213-e1218 [FREE Full text] [doi: 10.1097/MAO.0000000000003275] [Medline: 34267103]
40. Ekobena P, Rothuizen LE, Bedussi F, Guilcher P, Meylan S, Ceschi A, et al. Four cases of audio-vestibular disorders related to immunisation with SARS-CoV-2 mRNA vaccines. Int J Audiol 2022 May 05:1-5. [doi: 10.1080/14992027.2022.2056721] [Medline: 35510657]
41. Rausch C, Qun-Ying Y. WHO Pharmaceuticals Newsletter No. 1, 2022. Geneva: World Health Organization; 2022.
42. Formeister EJ, Chien W, Agrawal Y, Carey JP, Stewart CM, Sun DQ. Preliminary analysis of association between COVID-19 vaccination and sudden hearing loss using US centers for disease control and prevention vaccine adverse events reporting system data. JAMA Otolaryngol Head Neck Surg 2021 Jul 01;147(7):674-676 [FREE Full text] [doi: 10.1001/jamaoto.2021.0869] [Medline: 34014263]
43. Formeister EJ, Wu MJ, Chari DA, Meek R, Rauch SD, Remenschneider AK, et al. Assessment of sudden sensorineural hearing loss after COVID-19 vaccination. JAMA Otolaryngol Head Neck Surg 2022 Apr 01;148(4):307-315 [FREE Full text] [doi: 10.1001/jamaoto.2021.4414] [Medline: 35201274]
44. Yanir Y, Doweck I, Shibli R, Najjar-Debbiny R, Saliba W. Association between the BNT162b2 Messenger RNA COVID-19 vaccine and the risk of sudden sensorineural hearing loss. JAMA Otolaryngol Head Neck Surg 2022 Apr 01;148(4):299-306 [FREE Full text] [doi: 10.1001/jamaoto.2021.4278] [Medline: 35201275]
45. Nieminen TA, Kivekäs I, Artama M, Nohynek H, Kujansivu J, Hovi P. Sudden hearing loss following vaccination against COVID-19. JAMA Otolaryngol Head Neck Surg 2023 Feb 01;149(2):133-140. [doi: 10.1001/jamaoto.2022.4154] [Medline: 36520464]
46. Ulrich AK, Sundaram ME, Osterholm MT. Rare sudden sensorineural hearing loss potentially associated with COVID-19 vaccination does not outweigh the benefit of COVID-19 vaccines. JAMA Otolaryngol Head Neck Surg 2022 Apr 01;148(4):315-316. [doi: 10.1001/jamaoto.2021.4279] [Medline: 35201285]
47. Briggs SE, Brenner MJ, Chandrasekhar SS. Sudden sensorineural hearing loss and COVID-19 vaccination. JAMA Otolaryngol Head Neck Surg 2022 Feb 01;148(2):196-197. [doi: 10.1001/jamaoto.2021.3384] [Medline: 34817571]
48. Portail de signalement de signalement des évènements sanitaires indésirables. Ministère de la santé. URL: https://signalement.social-sante.gouv.fr/ [accessed 2022-12-15]

49. Medical Dictionary for Regulatory Activities - MedDRA. Australian Government The Department of Health and Aged Care. URL: https://www.meddra.org/training/schedule/1154/Standardised%20MedDRA%20Queries% [accessed 2022-12-15]
50. What the numbers mean: an epidemiological perspective on hearing. NIH National Institute on Deafness and Other Communication Disorders. URL: https://www.nidcd.nih.gov/health/statistics/what-numbers-mean-epidemiological- [accessed 2023-02-05]
51. Cheng Y, Chu Y, Tu T, Shiao A, Wu S, Liao W. Modified Siegel's criteria for sudden sensorineural hearing loss: reporting recovery outcomes with matched pretreatment hearing grades. J Chin Med Assoc 2018 Nov;81(11):1008-1012 [FREE Full text] [doi: 10.1016/j.jcma.2018.03.012] [Medline: 29929833]
52. Kandakure VT, Kunjumon R, Dube Y, More MS, Garje S. A prospective study on post Covid sudden onset sensory neural hearing loss and its recovery. Indian J Otolaryngol Head Neck Surg 2023 Feb 14:1-7 [FREE Full text] [doi: 10.1007/s12070-023-03529-3] [Medline: 36817015]
53. Siegel LG. The treatment of idiopathic sudden sensorineural hearing loss. Otolaryngol Clin North Am 1975 Jun;8(2):467-473. [Medline: 1153209]
54. Sensorineural Hearing Loss: AESI Case Definition Companion Guide. Brighton Collaboration. 2023 Mar 12. URL: https://brightoncollaboration.us/sensorineural-hearing-loss-aesi-case-definition-companion-guide/ [accessed 2023-03-14]
55. Oster ME, Shay DK, Shimabukuro TT. Myocarditis cases after mRNA-based COVID-19 vaccination in the US-reply. JAMA 2022 May 24;327(20):2020-2021. [doi: 10.1001/jama.2022.5134] [Medline: 35608588]
56. Le tableau de bord de la vaccination. Ministère de la Santé et de la Prévention. 2021 May 20. URL: https://solidarites-sante.gouv.fr/grands-dossiers/vaccin-covid-19/article/le-tableau-de- [accessed 2022-12-15]
57. Wong H, Hu M, Zhou CK, Lloyd PC, Amend KL, Beachler DC, et al. Risk of myocarditis and pericarditis after the COVID-19 mRNA vaccination in the USA: a cohort study in claims databases. Lancet 2022 Jun;399(10342):2191-2199. [doi: 10.1016/s0140-6736(22)00791-7]
58. Le Vu S, Bertrand M, Jabagi M, Botton J, Drouin J, Baricault B, et al. Age and sex-specific risks of myocarditis and pericarditis following Covid-19 messenger RNA vaccines. Nat Commun 2022 Jun 25;13(1):3633 [FREE Full text] [doi: 10.1038/s41467-022-31401-5] [Medline: 35752614]
59. Aronson JK, Hauben M. Anecdotes that provide definitive evidence. BMJ 2006 Dec 16;333(7581):1267-1269 [FREE Full text] [doi: 10.1136/bmj.39036.666389.94] [Medline: 17170419]
60. Barda N, Dagan N, Ben-Shlomo Y, Kepten E, Waxman J, Ohana R, et al. Safety of the BNT162b2 mRNA Covid-19 vaccine in a nationwide setting. N Engl J Med 2021 Sep 16;385(12):1078-1090. [doi: 10.1056/nejmoa2110475]
61. Almufarrij I, Munro KJ. One year on: an updated systematic review of SARS-CoV-2, COVID-19 and audio-vestibular symptoms. Int J Audiol 2021 Dec 22;60(12):935-945. [doi: 10.1080/14992027.2021.1896793] [Medline: 33750252]
62. Satar B. Criteria for establishing an association between Covid-19 and hearing loss. Am J Otolaryngol 2020;41(6):102658 [FREE Full text] [doi: 10.1016/j.amjoto.2020.102658] [Medline: 32836039]
63. Jafari Z, Kolb BE, Mohajerani MH. Hearing loss, tinnitus, and dizziness in COVID-19: a systematic review and meta-analysis. Can J Neurol Sci 2022 Mar 12;49(2):184-195 [FREE Full text] [doi: 10.1017/cjn.2021.63] [Medline: 33843530]
64. Parrino D, Frosolini A, Toninato D, Matarazzo A, Marioni G, de Filippis C. Sudden hearing loss and vestibular disorders during and before COVID-19 pandemic: an audiology tertiary referral centre experience. Am J Otolaryngol 2022;43(1):103241 [FREE Full text] [doi: 10.1016/j.amjoto.2021.103241] [Medline: 34555789]
65. Fancello V, Fancello G, Hatzopoulos S, Bianchini C, Stomeo F, Pelucchi S, et al. Sensorineural hearing loss post-COVID-19 infection: an update. Audiol Res 2022 Jun 01;12(3):307-315 [FREE Full text] [doi: 10.3390/audiolres12030032] [Medline: 35735365]
66. Pool C, King TS, Pradhan S, Isildak H. Sudden sensorineural hearing loss and coronavirus disease 2019. J Laryngol Otol 2022 Apr 21;136(9):823-826. [doi: 10.1017/s0022215122000068]
67. Alvergne A, Kountourides G, Argentieri MA, Agyen L, Rogers N, Knight D, et al. A retrospective case-control study on menstrual cycle changes following COVID-19 vaccination and disease. iScience 2023 Apr 21;26(4):106401-106421 [FREE Full text] [doi: 10.1016/j.isci.2023.106401] [Medline: 36987520]
68. Strom B. Basic principles of clinical epidemiology relevant to pharmacoepidemiologic studies. In: Pharmacoepidemiology. Hoboken, NJ: Wiley-Blackwell; 2019.
69. Points de situation vaccination. Ministère de la Santé et de la Prévention. 2021. URL: https://tinyurl.com/2b5dryt6 [accessed 2022-12-10]
70. Ouldali N, Bagheri H, Salvo F, Antona D, Pariente A, Leblanc C, “French Covid-19 Paediatric Inflammation Consortium”£the “French Pharmacovigilance network”*. Hyper inflammatory syndrome following COVID-19 mRNA vaccine in children: a national post-authorization pharmacovigilance study. Lancet Reg Health Eur 2022 Jun;17:100393 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100393] [Medline: 35505833]

Abbreviations

ADR: adverse drug reaction

ANSM: Agence Nationale de Sécurité du Médicament et des Produits de santé (National Agency for the Safety of Medicines and Health Products)
CNIL: Commission nationale de l’informatique et des libertés (National Commission on Informatics and Liberty)
EMA: European Medicines Agency
FPVDB: French Pharmacovigilance Database
mRNA: messenger RNA
PRAC: Pharmacovigilance Risk Assessment Committee
RPVC: Regional Pharmacovigilance Center
Rr: reporting rate
SSNHL: sudden sensorineural hearing loss
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 23.12.22; peer-reviewed by S Blackburn, Y Chu; comments to author 27.01.23; revised version received 17.02.23; accepted 18.04.23; published 14.07.23.

Please cite as:
Thai-Van H, Valnet-Rabier MB, Anciaux M, Lambert A, Maurier A, Cottin J, Pietri T, Destère A, Damin-Pernik M, Perrouin F, Bagheri H
Safety Signal Generation for Sudden Sensorineural Hearing Loss Following Messenger RNA COVID-19 Vaccination: Postmarketing Surveillance Using the French Pharmacovigilance Spontaneous Reporting Database
JMIR Public Health Surveill 2023;9:e45263
URL: https://publichealth.jmir.org/2023/1/e45263
doi:10.2196/45263
PMID:37071555

Original Paper

Caregiver Perceptions of Children's and Adolescents' Psychosocial Functioning During the Stringent COVID-19 Lockdown Restrictions in Shanghai: Cross-sectional Study

Xu Liu[1*], PhD; Jing Wu[2*], PhD; Hongyang Yang[3,4*], PhD; Fangjie Zhao[1*], PhD; Yuchen Qin[1*], PhD; Jiali Wu[1], BSc; Hongli Yan[2], BSc; Yan Xu[1], BSc; Lulu Zhang[1], Prof Dr, PhD

[1]College of Health Service, Second Military Medical University, Shanghai, China
[2]School of Nursing, Second Military Medical University, Shanghai, China
[3]Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[4]Institute of Hospital Development Strategy, China Hospital Development Institute, Shanghai Jiao Tong University, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Lulu Zhang, Prof Dr, PhD
College of Health Service
Second Military Medical University
800# Xiangyin Road, Yangpu
Shanghai, 200433
China
Phone: 86 02181871421
Email: zllrmit@163.com

Abstract

Background: The COVID-19 pandemic represents a global health crisis. The Shanghai municipal government in China implemented strict and comprehensive pandemic control strategies in the first half of 2022 to eliminate a wave of COVID-19 infection. The pandemic and the resulting government responses have led to abrupt changes to families' daily lives, including the mental health of children and adolescents.

Objective: The aim of this paper is to examine the impact of COVID-19 exposure and the stringent lockdown measures on the daily life and mental health of children and adolescents and to provide suggestions on maintaining their mental health when similar public health emergencies occur in the future.

Methods: In this cross-sectional study, an anonymous survey was distributed online in May 1-15, 2022, in Shanghai. Individuals were eligible to participate if they were currently the caregiver of a child or adolescent (aged 4-17 years). Outcomes were psychosocial functioning of children and adolescents, as reported by parents, using the Pediatric Symptom Checklist-17. COVID-19 exposure and life changes were also reported. Multivariate logistic regression was used to analyze risk factors for poor psychosocial functioning.

Results: In total, 2493 valid questionnaires were analyzed. The rate of positive scores on the global Pediatric Symptom Checklist-17 scale was 16.5% (n=411). Internalizing, attention, and externalizing problem subscale positivity rates were 17.3% (n=431), 10.9% (n=272), and 8.9% (n=221), respectively. Caregivers reported that 64.2% (n=1601) and 20.7% (n=516) of the children's interactions with friends or peers and parents deteriorated, respectively. Compared with male caregivers, female caregivers were less likely to report psychosocial problems in children and adolescents (adjusted odds ratio [aOR] 0.68; 95% CI 0.53-0.88). Older children and those with lower COVID-19 Exposure and Family Impact Scales scores were less likely to have psychological problems (aOR 1.15; 95% CI 1.10-1.21). Compared with children with screen times <1 hour per day for recreation, those using screens for >3 hours had higher odds of psychological distress (aOR 2.09; 95% CI 1.47-1.97). Children who spent 1-2 hours exercising and had better interactions with friends or peers and parents showed a trend toward lower odds of psychological problems. Children and adolescents with worse sleep compared with preclosure were more likely to have psychological problems.

Conclusions: The prevalence of psychosocial problems among children and adolescents is relatively high. Being young, having more COVID-19 exposure, and having more screen times (>3 h/day), less exercise time (<30 min), worse sleep, and deteriorated interactions with friends or peers and parents were risk factors for poor psychosocial functioning. It is necessary for governments, communities, schools, and families to take appropriate countermeasures to reduce the negative impact of the stringent control measures on caregivers' parenting and psychosocial functioning of children and adolescents.

(JMIR Public Health Surveill 2023;9:e43689) doi:10.2196/43689

KEYWORDS

COVID-19 exposure; psychosocial function; parenting; children and adolescents; China

Introduction

The COVID-19 pandemic represents a global health crisis. According to the latest data, as of June 7, 2022, there have been over 529 million confirmed cases of COVID-19 globally and over 6 million deaths [1]. Governments across the world have implemented several public health and social measures, such as mask wearing, school closures, international travel restrictions, and stay-at-home orders to prevent COVID-19 infection [1]. The unpredictability and uncertainty of the COVID-19 pandemic, the associated lockdowns and containment strategies, and the resulting economic disruption have led to abrupt and unpredictable changes to families' daily lives and could increase the risk of mental health problems [2]. Previous studies showed that shielding behavior in both clinically extremely vulnerable and non–clinically extremely vulnerable populations during the first wave of the 2020 COVID-19 pandemic was associated with worse mental and physical well-being, suggesting that the adoption of such behavior may have resulted in avoidable detriments to physical and mental health [3].

In late February 2022, a wave of intense COVID-19 infection appeared in Shanghai, China. According to the Shanghai Municipal Health Commission, as of May 4, 2022, 593,336 cases have been identified, including 538,450 asymptomatic carriers [4]. To limit the spread of and subsequently eliminate COVID-19, the Shanghai municipal government implemented strict and comprehensive pandemic control strategies from March to May, including mask wearing, school and business closures, restrictions on domestic movement, public transport closures, and stay-at-home orders. According to the Public Health and Social Measures Severity Index issued by the World Health Organization, Shanghai's policy was level 4 (of 5) in terms of severity [1]. The strict and comprehensive pandemic control strategies in Shanghai are designed to reduce the number of people infected and to allow for early diagnosis and appropriate treatment for severe COVID-19 cases [4]. Shanghai's significant and stringent efforts against Omicron are essential for China as a whole to exit the pandemic situation.

Studies showed that the global pandemic and the resulting government responses have led to abrupt and unpredictable changes to damage to human health and creating burdens for families, health care systems, and societies [5-7]. As members of society, children and adolescents will not escape the repercussions of COVID-19, and inevitably, some children and families will experience these social costs differently. Efforts to contain the spread of COVID-19 have involved sudden, and often mandatory, physical distancing, thus removing many regular sources of social connection. School closures and quarantine orders may have contributed to a considerable proportion of the harms experienced by children [8] through a reduction in social contact with peers and teachers [9]. Many children have been unable to play or socialize outside the home. Social connections with classmates and peers were reduced, and adolescents especially suffer from a lack of social stimuli [10].

The mental health of children and adolescents during the COVID-19 crisis has attracted great attention [11]. In a systematic review of 36 studies from 11 countries, school closures and social lockdown during the first COVID-19 wave were associated with adverse mental health symptoms and negative health behaviors among children and adolescents [12]. Between 18% and 60% of children and adolescents scored above risk thresholds for distress, particularly anxiety and depressive symptoms [12]. Significant anxiety rates of 10%-19% and depressive symptom rates of 17%-39% have been estimated during the first COVID-19 wave [12]. In a longitudinal study, earlier internalization of symptoms increased the risk of clinically relevant depressive symptoms [13]. Another study suggested that 1 in every 4 young people globally are experiencing symptoms of clinical depression, while 1 in 5 are experiencing symptoms of clinical anxiety [8]. Assessments should be performed, and risk factors should be found as soon as possible to prevent serious mental problems from developing.

Longitudinal analysis indicated that more stringent policies and intense pandemics are associated with worse mental health [14]. This study aimed to examine the impact of COVID-19 exposure and the stringent lockdown measures on the daily life change and mental health of children and adolescents and provide suggestions on maintaining their mental health when similar public health emergencies occur in the future, leading to stringent lockdowns.

Methods

Recruitment

From May 1, 2022, to May 15, 2022, we conducted a municipal-wide web-based questionnaire survey in Shanghai, China. Individuals were eligible to participate if they were currently the caregivers of a child or adolescent (aged 4 to 17 years). Data were collected through the Wenjuanxing online questionnaire platform, on which only a fully completed questionnaire can be uploaded. An invitation letter, detailing the research aim and providing a QR code automatically generated by Wenjuanxing, was issued on widely used social platforms. Snowball sampling method and convenient sampling method were used [15]. Any individual interested in the study could use the QR code and fill in the questionnaire. In addition, we have taken a number of questionnaire quality-control measures, including the following: respondent as caregiver, who should be older than 20 years old; respondents with Shanghai internet protocol addresses; each social account can only fill in the questionnaire once; and respondents who took less than 2 minutes to complete the questionnaire were excluded.

Ethical Considerations

This study was approved by the Ethics Committee of Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine (Approval Number: XHEC-C-2022-043). Appropriate ethical considerations were applied during all the stages of this study. Participation in the study was entirely voluntary. The survey questionnaire was distributed among respondents who provided their informed consent on the first page and clicking on the "agree" button after reading about the purpose of the survey. The questionnaire was anonymous to ensure the confidentiality and reliability of data.

Measures

Demographic Characteristics of Caregivers and Children

Caregivers were asked to report the following information: age, gender, education level, living area, and number of children aged <18 years. In addition, the caregivers were asked to report the child's age (4-18 years), gender, daily exposure time to screens for learning during school closures, daily exposure time to screens for recreation during school closures, amount of exercise per day during school closures, child's sleep time compared to preclosure, child's interactions with friends or peers compared to preclosure, child's interactions with parents compared to preclosure, and child's psychosocial symptoms.

COVID-19 Exposure and Family Impact Scales

The COVID-19 Exposure and Family Impact Scales (CEFIS) is a caregiver-report measure used to examine the degree to which families are exposed to potentially traumatic aspects of the COVID-19 pandemic, and the perceived impact on child, caregiver, and family functioning. It has been used to examine the impact of COVID-19 on various chronically ill pediatric populations [16,17], as well as on healthy children [18]. The CEFIS consists of the following three primary scales: Exposure, Impact, and Distress. Our study used the Exposure scale. The Exposure scale contains 25 "Yes/No" items that assess whether families have been exposed to COVID-19–related events such as lockdowns, school closures, changes in employment status, or the virus itself. When completing the CEFIS, residents are asked to think about the period from March 2022 to the present. "Yes" responses are summed to yield a total score, with higher scores indicating greater exposure to COVID-19 and related events.

Pediatric Symptom Checklist-17 Items

The Pediatric Symptom Checklist-17 (PSC-17) is one of the most frequently used general pediatric psychosocial screening instruments [19]. It is a caregiver-report measure of children's psychosocial functioning. Parents rate the frequency of a variety of symptoms in their child (0=*never*, 1=*sometimes*, and 2=*often*). A screen can be considered "positive" based on total or domain scores. Total scores can be used to categorize a child as "at risk" or "not at risk" based on established cutoff scores. A score of ≥15 corresponds to at-risk status. On the three PSC-17 subscales, scores of ≥5 is for internalizing and ≥7 for externalizing; scores of attention ≥7 indicate risk. The PSC-17 has 3 subscales that specifically assess internalizing, externalizing, and attention-related problems.

Statistical Analysis

Descriptive statistics were generated for respondent characteristics and CEFIS Exposure subscale scores. Caregiver and child age, as continuous variables, were presented as median with interquartile range, and were analyzed by the Wilcoxon rank sum test. Categorical variables are presented as numbers and percentages and were compared between poor psychosocial health (PSC-17 score ≥15) and good psychosocial health (PSC-17 score <15) groups using Pearson χ^2 test or Fisher exact test. *P* values for trends were calculated using mental health as a binary categorical variable—event=PSC-17 score ≥15 (1), nonevent=PSC-17 score <15 (0)—and Pearson $\chi2$ trend test. To further examine potential risk and protective factors for mental health, univariate and multivariate logistic regression were performed. All respondent characteristic variables were included and adjusted in the multivariate logistic regression; unadjusted odds ratios, adjusted odds ratios (aORs), and their 95% CIs were calculated. All analyses were performed with SAS 9.4 (SAS Institute, Inc) and R statistical software (version 3.6.3; R Development Core Team). Statistical significance was set at *P*<.05, and all tests were 2-tailed.

Results

Descriptive Analysis of Participants

A total of 2540 valid questionnaires were returned. Questionnaires completed in <180 seconds were excluded, and 2493 valid questionnaires were finally analyzed. In total, 38.31% (955/2493) of the participants were male (Table 1), and 73.20% (n=1825) had 1 child (aged <18 years). In total, 52.55% (n=1310/2493) of the children were boys, and their average age was 10.99 (SD 4.18) years. Interactions with friends or peers and parents reportedly deteriorated in 64.22% (n=1601/2493) and 20.70% (n=516/2493) of the children, respectively.

The mean CEFIS score was 11.5 (SD 2.49). The scores for individual CEFIS item are summarized in Figure 1.

Table 1. Participant demographic characteristics (N=2493).

Characteristics	Total (N=2493)	PSC-17[a] score <15 (n=2082)	PSC-17 score ≥15 (n=411)	*P* value	*P* trend[b]
Caregiver’s age (years), median (IQR)	40 (36, 43)	40 (36, 43)	40 (36, 43)	.79	N/A[c]
Caregiver’s gender, n (%)				.008	N/A
Male	955 (38.31)	773 (80.94)	182 (19.06)		
Female	1538 (61.69)	1309 (85.11)	229 (14.89)		
Education level, n (%)				.49	.49
High school and below	444 (17.81)	379 (85.36)	65 (14.64)		
Junior college or BA	1532 (61.45)	1271 (82.96)	261 (17.04)		
Postgraduate and above	517 (20.74)	432 (83.56)	85 (16.44)		
Living area (square meters), n (%)				.31	.23
<30	890 (35.70)	730 (82.02)	160 (17.98)		
30-50	653 (26.19)	553 (84.69)	100 (15.31)		
>50	950 (38.11)	799 (84.11)	151 (15.89)		
Number of children (<18 years old), n (%)				.004	N/A
1	1825 (73.20)	1548 (84.82)	277 (15.18)		
≥2	668 (26.8)	534 (79.94)	134 (20.06)		
CEFIS[d] score, median (IQR)	11 (10, 13)	11 (9, 13)	13 (11, 14)	<.001	N/A
Child age (years), median (IQR)	11 (8, 14)	11 (8, 15)	11 (7, 14)	.08	N/A
Child age group, n (%)				.046	.51
4-6 years	473 (18.97)	402 (84.99)	71 (15.01)		
7-12 years	1111 (44.56)	905 (81.46)	206 (18.54)		
13-18 years	909 (36.47)	775 (85.26)	134 (14.74)		
Child’s gender, n (%)				.62	N/A
Boy	1310 (52.55)	1089 (83.13)	221 (16.87)		
Girl	1183 (47.45)	993 (83.94)	190 (16.06)		
Child’s length of daily exposure to screens for learning during closure, n (%)				.001	<.001
≤4 hours	686 (27.52)	593 (86.44)	93 (13.56)		
4-6 hours	637 (25.55)	543 (85.24)	94 (14.76)		
6-8 hours	613 (24.59)	508 (82.87)	105 (17.13)		
≥8 hours	557 (22.34)	438 (78.64)	119 (21.36)		
Child’s length of daily exposure to screens for recreation during closure, n (%)				<.001	<.001
≤1 hour	634 (25.43)	556 (87.70)	78 (12.30)		
1-2 hours	835 (33.49)	724 (86.71)	111 (13.29)		
2-3 hours	414 (16.61)	350 (84.54)	64 (15.46)		
≥3 hours	610 (24.47)	452 (74.10)	158 (25.90)		
Child’s length of exercise per day during closure, n (%)				<.001	<.001
≤30 minutes	1370 (54.96)	1087 (79.34)	283 (20.66)		
30 minutes to 1 hour	832 (33.37)	738 (88.70)	94 (11.30)		
1 hour to 2 hours	219 (8.78)	200 (91.32)	19 (8.68)		
≥2 hours	72 (2.89)	57 (79.17)	15 (20.83)		
Child’s sleep compared to preclosure, n (%)				<.001	<.001
Much better	92 (3.69)	85 (92.39)	7 (7.61)		

Characteristics	Total (N=2493)	PSC-17[a] score <15 (n=2082)	PSC-17 score ≥15 (n=411)	*P* value	*P* trend[b]
A little better	459 (18.41)	403 (87.80)	56 (12.20)		
No change	1186 (47.57)	1065 (89.8)	121 (10.20)		
A little worse	623 (24.99)	480 (77.05)	143 (22.95)		
Much worse	133 (5.34)	49 (36.84)	84 (63.16)		
Child’s interaction with friends or peers compared to preclosure, n (%)				<.001	<.001
Much better	45 (1.81)	40 (88.89)	5 (11.11)		
A little better	161 (6.46)	145 (90.06)	16 (9.94)		
No change	686 (27.52)	629 (91.69)	57 (8.31)		
A little worse	880 (35.30)	751 (85.34)	129 (14.66)		
Much worse	721 (28.91)	517 (71.71)	204 (28.29)		
Child’s interaction with parents compared to preclosure, n (%)				<.001	<.001
Much better	131 (5.25)	121 (92.37)	10 (7.63)		
A little better	564 (22.63)	495 (87.77)	69 (12.23)		
No change	1282 (51.43)	1162 (90.64)	120 (9.36)		
A little worse	398 (15.96)	267 (67.09)	131 (32.91)		
Much worse	118 (4.73)	37 (31.36)	81 (68.64)		

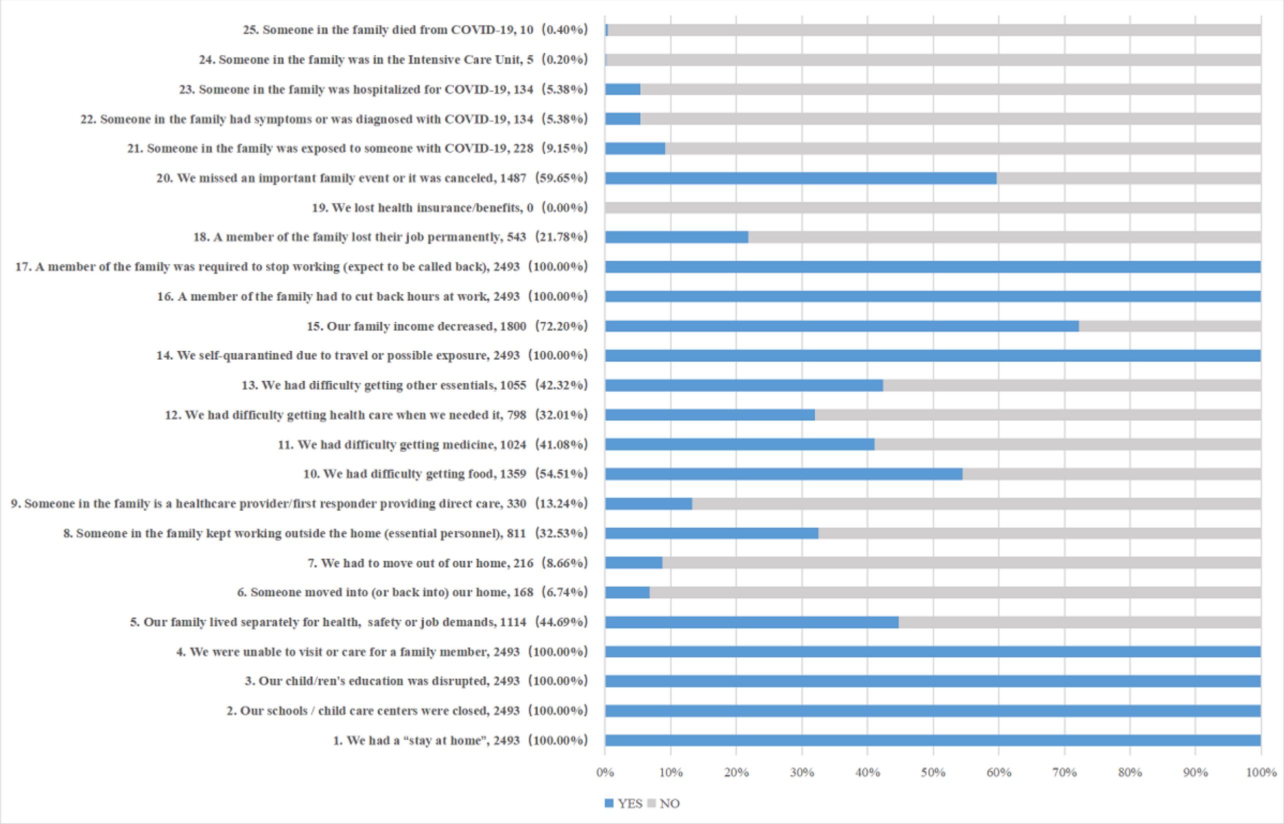
[a] PSC-17: Pediatric Symptom Checklist-17 Items.

[b] Pearson $\chi 2$ trend test.

[c] N/A: not applicable.

[d] CEFIS: COVID-19 Exposure and Family Impact Scales.

Figure 1. Individual COVID-19 Exposure and Family Impact Scales (CEFIS) item responses.



Bivariate Analysis of Psychosocial Functioning

As shown in Table 2, the proportion of positive screening scores on the global PSC-17 scale was 16.5% (411/2493). The internalizing, attention, and externalizing problem subscales had positive screening rates of 17.3% (431/2493), 10.9% (272/2493), and 8.9% (221/2493), respectively. Caregivers' gender, number of children (younger than 18 years old), CEFIS score, child age, child's length of daily exposure to screens for learning and recreation during closure, length of exercise per day, sleep, interaction with friends/peers and interaction with parents compared to pre-closure were significantly associated with psychosocial problems (Table 1).

Table 2. Caregiver reports of child mental health (N=2493).

Mental health	Positive screening scores, n (%)	Mean score (SD)
PSC-17[a] total problems	411 (16.5)	8.23 (6.72)
PSC-17 attention problems	272 (10.9)	3.43 (2.51)
PSC-17 internalizing problems	431 (17.3)	2.24 (2.29)
PSC-17 externalizing problems	221 (8.9)	2.55 (2.62)

[a] PSC-17: Pediatric Symptom Checklist-17 Items.

Logistic Regression Analysis

Multivariate logistic regression analysis showed that, compared with male caregivers, female caregivers were less likely to report psychosocial problems in children and adolescents (aOR 0.68; 95% CI 0.53-0.88). Older children and those with lower CEFIS scores were less likely to have psychological problems (aOR 1.15; 95% CI 1.10-1.21). Compared with children with screen times <1 hour per day for recreation, those using screens for >3 hours had higher odds of psychological distress (aOR 2.09; 95% CI 1.47-1.97). Compared with children who spent <0.5 hour exercising, those who spent 0.5-1 hour exercising had lower odds of psychological problems (aOR 0.70; 95% CI 0.52-0.94). Children who spent 1-2 hours exercising showed a trend toward lower odds of psychological problems. Children with worse sleep compared with preclosure and worse interactions with friends or peers and parents were more likely to have psychological problems (Table 3).

Table 3. Multivariate logistic regression analysis of characteristics associated with the mental health status of children and adolescents during the COVID-19–related school closures in Shanghai, China.

Characteristic	Unadjusted OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Caregiver's age	1.00 (0.98,1.01)	.65	1.01 (0.99,1.04)	.37
Caregiver's gender				
Male	Reference	N/A[b]	Reference	N/A
Female	0.74 (0.60,0.92)	.007	0.68 (0.53,0.88)	.004
Education level				
High school and below	Reference	N/A	Reference	N/A
Junior College or BA	1.20 (0.89,1.61)	.23	1.14 (0.78,1.66)	.50
Postgraduate and above	1.15 (0.81,1.63)	.44	1.09 (0.69,1.70)	.71
Living area (square meters)				
<30	Reference	N/A	Reference	N/A
30-50	0.83 (0.63,1.08)	.17	1.04 (0.76,1.42)	.82
>50	0.86 (0.68,1.10)	.23	0.90 (0.68,1.20)	.49
Number of children (<18 years old)				
1	Reference	N/A	Reference	N/A
≥2	1.40 (1.12,1.76)	.004	1.52 (1.16,1.99)	.002
CEFIS[c] score	1.24 (1.19,1.30)	<.001	1.15 (1.10,1.21)	<.001
Child age	0.98 (0.95,1.00)	.06	0.95 (0.91,0.99)	.01
Child's gender				
Boy	Reference	N/A	Reference	N/A
Girl	0.94 (0.76,1.17)	.59	0.85 (0.66,1.08)	.18
Child's length of daily exposure to screens for learning during closure				
≤4 hours	Reference	N/A	Reference	N/A
4-6 hours	1.10 (0.81,1.50)	.53	1.32 (0.92,1.90)	.13
6-8 hours	1.32 (0.97,1.78)	.07	1.38 (0.94,2.01)	.10
≥8 hours	1.73 (1.29,2.33)	<.001	1.20 (0.79,1.81)	.39
Child's length of daily exposure to screens for recreation during closure				
≤1 hour	Reference	N/A	Reference	N/A
1-2 hours	1.09 (0.80,1.49)	.57	1.26 (0.89,1.79)	.19
2-3 hours	1.30 (0.91,1.86)	.15	1.37 (0.91,2.07)	.13
≥3 hours	2.49 (1.85,3.36)	<.001	2.09 (1.47,2.97)	<.001
Child's length of exercise per day during closure				
≤30 minutes	Reference	N/A	Reference	N/A
30 minutes to 1 hour	0.49 (0.38,0.63)	<.001	0.70 (0.52,0.94)	.02
1 hour to 2 hours	0.36 (0.22,0.59)	<.001	0.61 (0.35,1.06)	.08
≥2 hours	1.01 (0.56,1.81)	.97	0.93 (0.45,1.92)	.84
Child's sleep compared to preclosure				
Much better	Reference	N/A	Reference	N/A
A little better	1.69 (0.74,3.83)	.21	1.42 (0.58,3.50)	.45
No change	1.38 (0.62,3.05)	.43	1.05 (0.44,2.54)	.91
A little worse	3.62 (1.64,7.99)	.001	1.85 (0.76,4.48)	.18
Much worse	20.82 (8.92,48.58)	<.001	5.32 (2.02,13.95)	<.001

Characteristic	Unadjusted OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Child's interaction with friends or partners compared to preclosure				
Much better, a little better, or no change	Reference	N/A	Reference	N/A
A little worse	1.79 (1.33,2.42)	<.001	1.19 (0.85,1.66)	.31
Much worse	4.12 (3.10,5.47)	<.001	1.70 (1.22,2.39)	.002
Child's interaction with parents compared to preclosure				
Much better	Reference	N/A	Reference	N/A
A little better	1.69 (0.84,3.37)	.14	1.53 (0.72,3.25)	.27
No change	1.25 (0.64,2.45)	.52	1.17 (0.56,2.44)	.68
A little worse	5.94 (3.01,11.69)	<.001	4.03 (1.90,8.54)	<.001
Much worse	26.48 (12.47,56.24)	<.001	10.51 (4.53,24.34)	<.001

[a] OR: odds ratio.

[b] N/A: not applicable.

[c] CEFIS: COVID-19 Exposure and Family Impact Scales.

Discussion

Principal Findings

This study examined the impact of the COVID-19 crisis on a sample of children and adolescents in Shanghai during the lockdown in the first half of 2022. This was the first study to assess the mental health of children and adolescents using the PSC-17 in China. Nearly one-sixth (16.5%, 411/2493) of the children and adolescents were reported to be with psychosocial function problems by their caregivers, which was similar to the results of a US study [20]. In another study, 18% of urban school-age children (5-11 years old) were at risk according to the total PSC-17 score, while 18% were at risk according to the PSC-17 internalizing symptoms subscale score [19]. Given the mental health impact on children and adolescents of COVID-19 exposure, measures should be taken to help families and parents optimize child health and development [21].

Psychological problems of children and adolescents varied cross different gender of reporting caregivers. Female caregivers reported less psychosocial problems in their children. The results were similar to those by Dave et al [22] but in contrast to the findings from Hagerman [23]. Dave et al [22] reported higher scores for externalizing behaviors, such as conduct problems and hyperactivity, among children (according to their fathers). Because of the stay-at-home policy, many fathers spent more time with their children, were involved in looking after them, and were more sensitive to externalizing behaviors than mothers, who were often the primary caregivers. The results showed that older children had fewer psychological problems. A possible reason for this might be that younger children were more susceptible to the crisis situation [24]. Those caregivers who had more than one child (younger than 18 years old) reported more psychological problems of their children. The reason might be that large families with more children living in cramped conditions might be under additional stress, including parenting-related stress, abuse, violence against children [25], and an increased risk of COVID-19 infection [26]. Government and community efforts are needed that target those at-risk households.

COVID-19 exposure was a risk factor for psychological problems of children and adolescents. The results of our study indicated severe COVID-19 exposure in our sample and increased risks of psychological problems. To eliminate COVID-19, stringent measures were implemented in Shanghai. Transport restrictions resulted in inconvenience in everyday life. Schools were closed, and children's education was disrupted. Everyone, except essential workers, had to stay at home. Some families had difficulty obtaining essentials such as food, medicine, health care, and so on. Moreover, the shutdown had economic repercussions for many families, such as decreased income or unemployment. The infected families and those in which someone was serving as an essential worker outside the home had more difficulties with parenting and meeting the growth and developmental demands of their children. Government and community efforts are needed that target at-risk households. Children who are separated from their caregivers also require particular attention, including those infected with or suspected of being infected with COVID-19, those who are quarantined in local hospitals or medical observation centers, and those whose caregivers are infected with COVID-19 [27].

In this study, the rate of psychosocial problems increased with recreational screen time, consistent with other studies [28]. School closures and strict containment measures have led to more families relying on technology and digital solutions for their children's learning, entertainment, and connection to the outside world [28,29]. Many children and adolescents are spending more time online, allowing for social interactions in some cases but also increasing the risk of cyberbullying, cybercrime, privacy infringements, and screen fatigue [10,30,31]. It has been proved that cyberchondria partially mediated the relationships between perceived severity of the COVID-19 pandemic and depression, anxiety, and stress in China [32]. As screen time increased, so too did sedentary behavior [33], which was associated with poor behavioral conduct or prosocial behavior and reduced sleep duration [34,35]. Worse sleep was a risk factor for the development of psychosocial problems in our study. Therefore, for children and adolescents, the time spent being sedentary should be limited, particularly the amount of recreational screen time [36]. Mutually agreed rules, limits, and parental monitoring of screen time were perceived as likely to be effective [37]. Interactive games and exercises can serve as a supplement for filling the time previously spent on screen.

Children and adolescents who spent 1-2 hours exercising showed a trend toward lower odds of psychological problems, and exercise for less than 30 minutes per day had a high probability of psychosocial problems. Psychosocial problems were linked to lower levels of physical activity [38-40]. Greater levels of physical activity are associated with improved physical fitness, bone health, cognitive outcomes, and mental health (reduced symptoms of depression) and less adiposity [36]. The World Health Organization recommends that children and adolescents perform at least 60 minutes of moderate-to-vigorous–intensity (mostly aerobic) physical activity per day. Vigorous-intensity aerobic activities, as well as those that strengthen muscle and bone, should be performed on at least 3 days a week [36]. High-quality physical education and supportive environments can promote physical and health literacy for long-lasting, healthy, active lifestyles [41]. Special attention should be paid to long-duration exercise. Previous studies had showed exercise durations of >3 hours were associated with worse mental health than exercising for 45 minutes [42]. In this study, exercise durations of 30 minutes to 1 hour were shown to be better for the psychosocial function of children and adolescents. Moreover, excessive exercise might cause fatigue and worsen mental health of children and adolescents [43].

In total, 27.9% of the caregivers reported that their interactions with children improved, which was significantly associated with lower positive screening scores. The stay-at-home policy might promote stronger relationships of children and adolescents with their parents [25] and has shown to improve their mental health [44]. Previous studies have reported that parent-child activities and exercises together can improve interactions between caregivers and children and amplifies the beneficial effects of sleep on mental health [45]. For example, mindfulness-promoting exercises (ie, yoga and tai chi) could be performed indoors by parents and children together [42]. In total, 64.2% (1601/2493) of the children and adolescents reported worse relationships with friends and classmates, and these children and adolescents were more likely to have poor mental health. There is an urgent need to increase the social interactions of children and adolescents during lockdown. Communities could redesign neighborhoods to provide children with spaces to play together and enjoy interactions [10]. Studies have shown that digitalized social contact can mitigate the potentially harmful effects of physical distancing on young people [9]. Specifically, active use of social media, for example, engaging in direct communication (ie, messaging) or posting directly on another person’s social media profile, can increase well-being and help maintain personal relationships. Time-limited and active use of social media for contact with peers might be beneficial for the mental health of children and adolescents.

Limitations

This study had several limitations. First, its cross-sectional nature limits the ability to establish a causal relationship between risk factors and mental health. Furthermore, there were no preclosure PSC-17 data for Chinese children and adolescents, so we could not determine whether the psychosocial problems detected in the participants were intensified during COVID-19 exposure. Second, the data were collected online because of the COVID-19 pandemic, and caregivers whose lives were severely disrupted and had limited time online might have been less likely to finish the questionnaire. Third, data for the main variable of interest, psychosocial functioning, were obtained from caregivers of the children and adolescents, which may cause bias. When compared to self-report or direct observation of youth behaviors, caregiver report may not be as accurate.

Conclusions

In this cross-sectional study conducted during a wave of COVID-19 infection in Shanghai, in which stringent responses were implemented, we found that the prevalence of psychosocial problems among children and adolescents was relatively high. Daily life of the children and adolescents changed a lot, including screen time, exercise, sleep, and interaction with parents and peers. Younger children, more COVID-19 exposure, more recreational screen time, and worse sleep were risk factors for poor psychosocial functioning. In particular, exercise durations of 1-2 hours and better interactions with friends or peers and parents benefit the psychosocial function of children and adolescents. The findings suggest that it is necessary for governments, communities, schools, and families to take appropriate countermeasures to reduce the impact of the lockdown measures on daily life and its subsequent influence on mental health of children and adolescents, which can be provided as suggestions on maintaining the mental health when similar public health emergencies occur in the future, leading to the stringent lockdown. In addition, it will be necessary to conduct longitudinal studies to clarify whether psychosocial problems are reduced, maintained, or intensified over time, as well as to identify subgroups of young people most at risk.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (grant 71774166 to XL,72174204 to LZ, and grant 71904200 to JW), the Chinese Defense Research Program of Science and Technology (grant 2019-JCJQ- JJ-065 to XL), Military Medical Science and Technology Youth Cultivation Program (grant 19QNP019 to JW), Shanghai Sailing Program (grant 20YF1457900 to YQ), Military Key Disciplines Construction Project (grant 03 to LZ), and the Sailing Talents Project of the Naval Military Medical University.

The authors would like to express their gratitude to EditSprings for the expert linguistic services provided.

Data Availability

The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Authors' Contributions

JW, HY, FZ, and YQ contributed equally as first co-authors. XL, HY, and LZ contributed to the conception and design. JW, HY, FZ, YX, JW, and XL drafted the manuscript. JW and YQ did the statistical analysis. XL and LZ supervised the study and the manuscript. All authors critically revised the manuscript for important intellectual content.

Conflicts of Interest

None declared.

References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2023-01-24]
2. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. The Lancet Psychiatry 2020 Sep;7(9):813-824 [FREE Full text] [doi: 10.1016/s2215-0366(20)30307-2]
3. Bachtiger P, Adamson A, Maclean WA, Kelshiker MA, Quint JK, Peters NS. Determinants of Shielding Behavior During the COVID-19 Pandemic and Associations With Well-being Among National Health Service Patients: Longitudinal Observational Study. JMIR Public Health Surveill 2021 Sep 20;7(9):e30460 [FREE Full text] [doi: 10.2196/30460] [Medline: 34298499]
4. Zhang X, Zhang W, Chen S. Shanghai's life-saving efforts against the current omicron wave of the COVID-19 pandemic. The Lancet 2022 May;399(10340):2011-2012. [doi: 10.1016/s0140-6736(22)00838-8]
5. Silk J, Scott L, Hutchinson E, Lu C, Sequeira SL, McKone KMP, et al. Storm Clouds and Silver Linings: Day-to-Day Life in COVID-19 Lockdown and Emotional Health in Adolescent Girls. J Pediatr Psychol 2022 Feb 03;47(1):37-48 [FREE Full text] [doi: 10.1093/jpepsy/jsab107] [Medline: 34664665]
6. Santomauro D, Mantilla HA, Shadid J. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. The Lancet 2021;398(10312):1700-1712 [FREE Full text] [doi: 10.1016/s0140-6736(21)02143-7]
7. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
8. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global Prevalence of Depressive and Anxiety Symptoms in Children and Adolescents During COVID-19: A Meta-analysis. JAMA Pediatr 2021 Nov 01;175(11):1142-1150 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.2482] [Medline: 34369987]
9. Orben A, Tomova L, Blakemore S. The effects of social deprivation on adolescent development and mental health. Lancet Child Adolesc Health 2020 Aug;4(8):634-640 [FREE Full text] [doi: 10.1016/S2352-4642(20)30186-3] [Medline: 32540024]
10. Clark H, Coll-Seck AM, Banerjee A, Peterson S, Dalglish SL, Ameratunga S, et al. After COVID-19, a future for the world's children? The Lancet 2020 Aug;396(10247):298-300. [doi: 10.1016/S0140-6736(20)31481-1]
11. Richardson D, Carraro A, Cebotari V, Gromada A, Rees G. Supporting families and children beyond COVID-19: social protection in high-income countries. Florence, Italy: UNICEF Office of Research-Innocenti. 2020. URL: https://tinyurl.com/w3ut5pym [accessed 2023-01-24]
12. Viner R, Russell S, Saulle R, Croker H, Stansfield C, Packer J, et al. School Closures During Social Lockdown and Mental Health, Health Behaviors, and Well-being Among Children and Adolescents During the First COVID-19 Wave: A Systematic Review. JAMA Pediatr 2022 Apr 01;176(4):400-409. [doi: 10.1001/jamapediatrics.2021.5840] [Medline: 35040870]
13. Vacaru SV, Beijers R, de Weerth C. Internalizing symptoms and family functioning predict adolescent depressive symptoms during COVID-19: A longitudinal study in a community sample. PLoS One 2022 Mar 18;17(3):e0264962 [FREE Full text] [doi: 10.1371/journal.pone.0264962] [Medline: 35303003]
14. Aknin LB, Andretti B, Goldszmidt R, Helliwell JF, Petherick A, De Neve J, et al. Policy stringency and mental health during the COVID-19 pandemic: a longitudinal analysis of data from 15 countries. Lancet Public Health 2022 May;7(5):e417-e426 [FREE Full text] [doi: 10.1016/S2468-2667(22)00060-3] [Medline: 35461592]
15. Midorikawa H, Tachikawa H, Taguchi T, Shiratori Y, Takahashi A, Takahashi S, et al. Demographics Associated With Stress, Severe Mental Distress, and Anxiety Symptoms During the COVID-19 Pandemic in Japan: Nationwide Cross-sectional Web-Based Survey. JMIR Public Health Surveill 2021 Nov 22;7(11):e29970 [FREE Full text] [doi: 10.2196/29970] [Medline: 34653018]

16. Enlow P, Phan T, Lewis A, Hildenbrand AK, Sood E, Canter KS, et al. Validation of the COVID-19 Exposure and Family Impact Scales. J Pediatr Psychol 2022 Mar 05;47(3):259-269 [FREE Full text] [doi: 10.1093/jpepsy/jsab136] [Medline: 34969064]
17. Kazak A, Alderfer M, Enlow P, Lewis AM, Vega G, Barakat L, et al. COVID-19 Exposure and Family Impact Scales: Factor Structure and Initial Psychometrics. J Pediatr Psychol 2021 Jun 03;46(5):504-513 [FREE Full text] [doi: 10.1093/jpepsy/jsab026] [Medline: 33749794]
18. Raviv T, Warren CM, Washburn JJ, Kanaley MK, Eihentale L, Goldenthal HJ, et al. Caregiver Perceptions of Children's Psychological Well-being During the COVID-19 Pandemic. JAMA Netw Open 2021 Apr 01;4(4):e2111103 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.11103] [Medline: 33914046]
19. Jellinek M, Bergmann P, Holcomb JM, Riobueno-Naylor A, Dutta A, Haile H, et al. Recognizing Adolescent Depression with Parent- and Youth-Report Screens in Pediatric Primary Care. J Pediatr 2021 Jun;233:220-226.e1. [doi: 10.1016/j.jpeds.2021.01.069] [Medline: 33548264]
20. Spencer AE, Oblath R, Dayal R, Loubeau JK, Lejeune J, Sikov J, et al. Changes in psychosocial functioning among urban, school-age children during the COVID-19 pandemic. Child Adolesc Psychiatry Ment Health 2021 Dec 02;15(1):73 [FREE Full text] [doi: 10.1186/s13034-021-00419-w] [Medline: 34857026]
21. Ghebreyesus TA, Russell C. Opportunities in crisis for optimising child health and development. The Lancet 2022 May;399(10337):1761-1763. [doi: 10.1016/s0140-6736(22)00608-0]
22. Davé S, Nazareth I, Senior R, Sherr L. A comparison of father and mother report of child behaviour on the Strengths and Difficulties Questionnaire. Child Psychiatry Hum Dev 2008 Dec 12;39(4):399-413. [doi: 10.1007/s10578-008-0097-6] [Medline: 18266104]
23. Hagerman TK, Carle AC, Houtrow AJ. Differences in Reporting Child Health Needs and Status and Developmental Concerns between Mothers and Fathers: Findings from the National Survey of Children's Health. J Pediatr 2020 May;220:254-257. [doi: 10.1016/j.jpeds.2019.12.059] [Medline: 32037155]
24. Calegaro VC, Ramos-Lima LF, Hoffmann MS, Zoratto G, Kerber N, Costa FCD, et al. Closed doors: Predictors of stress, anxiety, depression, and PTSD during the onset of COVID-19 pandemic in Brazil. J Affect Disord 2022 Aug 01;310:441-451 [FREE Full text] [doi: 10.1016/j.jad.2022.05.052] [Medline: 35569607]
25. Cluver L, Lachman JM, Sherr L, Wessels I, Krug E, Rakotomalala S, et al. Parenting in a time of COVID-19. Lancet 2020 Apr 11;395(10231):e64 [FREE Full text] [doi: 10.1016/S0140-6736(20)30736-4] [Medline: 32220657]
26. Hill DL, Christakis DA, Feudtner C. Illness in the Family and Shared Stress, Work, and Struggle. JAMA Pediatr 2022 Jun 01;176(6):549-550. [doi: 10.1001/jamapediatrics.2022.0491] [Medline: 35404432]
27. Liu JJ, Bao Y, Huang X, Shi J, Lu L. Mental health considerations for children quarantined because of COVID-19. The Lancet Child & Adolescent Health 2020 May;4(5):347-349. [doi: 10.1016/s2352-4642(20)30096-1]
28. Trott M, Driscoll R, Irlado E, Pardhan S. Changes and correlates of screen time in adults and children during the COVID-19 pandemic: A systematic review and meta-analysis. EClinicalMedicine 2022 Jun;48:101452 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101452] [Medline: 35615691]
29. Children at increased risk of harm online during global COVID-19 pandemic. Unicef. 2020. URL: https://www.unicef.org/press-releases/children-increased-risk-harm-online-during-global-covid-19-pandemic [accessed 2022-04-25]
30. Clark H, Coll-Seck AM, Banerjee A, Peterson S, Dalglish SL, Ameratunga S, et al. A future for the world's children? A WHO–UNICEF–Lancet Commission. The Lancet 2020 Feb;395(10224):605-658. [doi: 10.1016/s0140-6736(19)32540-1]
31. Rider EA, Ansari E, Varrin PH, Sparrow J. Mental health and wellbeing of children and adolescents during the covid-19 pandemic. BMJ 2021 Aug 24;374:n1730. [doi: 10.1136/bmj.n1730] [Medline: 34429302]
32. Han L, Zhan Y, Li W, Xu Y, Xu Y, Zhao J. Associations Between the Perceived Severity of the COVID-19 Pandemic, Cyberchondria, Depression, Anxiety, Stress, and Lockdown Experience: Cross-sectional Survey Study. JMIR Public Health Surveill 2021 Sep 16;7(9):e31052 [FREE Full text] [doi: 10.2196/31052] [Medline: 34478402]
33. Moore SA, Faulkner G, Rhodes RE, Brussoni M, Chulak-Bozzer T, Ferguson LJ, et al. Impact of the COVID-19 virus outbreak on movement and play behaviours of Canadian children and youth: a national survey. Int J Behav Nutr Phys Act 2020 Jul 06;17(1):85 [FREE Full text] [doi: 10.1186/s12966-020-00987-8] [Medline: 32631350]
34. Chauntry AJ, Bishop NC, Hamer M, Paine NJ. Sedentary behaviour, physical activity and psychobiological stress reactivity: A systematic review. Biol Psychol 2022 Jul;172:108374. [doi: 10.1016/j.biopsycho.2022.108374] [Medline: 35667480]
35. Huang S, Li Y, Zhang Y, Huang Y, Wu B, Zhang W, et al. Sleep, physical activity, sedentary behavior, and risk of incident dementia: a prospective cohort study of 431,924 UK Biobank participants. Mol Psychiatry 2022 Oct 14;27(10):4343-4354. [doi: 10.1038/s41380-022-01655-y] [Medline: 35701596]
36. WHO guidelines on physical activity and sedentary behaviour: at a glance. World Health Organization. URL: https://apps.who.int/iris/rest/bitstreams/1318299/retrieve [accessed 2023-01-24]
37. Minges KE, Owen N, Salmon J, Chao A, Dunstan DW, Whittemore R. Reducing youth screen time: qualitative metasynthesis of findings on barriers and facilitators. Health Psychol 2015 Apr;34(4):381-397 [FREE Full text] [doi: 10.1037/hea0000172] [Medline: 25822054]

38. Qin Z, Shi L, Xue Y, Lin H, Zhang J, Liang P, et al. Prevalence and Risk Factors Associated With Self-reported Psychological Distress Among Children and Adolescents During the COVID-19 Pandemic in China. JAMA Netw Open 2021 Jan 04;4(1):e2035487 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.35487] [Medline: 33496797]
39. Olive LS, Sciberras E, Berkowitz TS, Hoare E, Telford RM, O'Neil A, et al. Child and Parent Physical Activity, Sleep, and Screen Time During COVID-19 and Associations With Mental Health: Implications for Future Psycho-Cardiological Disease? Front Psychiatry 2021 Feb 10;12:774858 [FREE Full text] [doi: 10.3389/fpsyt.2021.774858] [Medline: 35242059]
40. Zhou J, Yuan X, Qi H, Liu R, Li Y, Huang H, et al. Prevalence of depression and its correlative factors among female adolescents in China during the coronavirus disease 2019 outbreak. Global Health 2020 Jul 28;16(1):69 [FREE Full text] [doi: 10.1186/s12992-020-00601-3] [Medline: 32723373]
41. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization. 2018. URL: https://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf [accessed 2023-01-24]
42. Chekroud S, Gueorguieva R, Zheutlin A, Paulus M, Krumholz H, Krystal J, et al. Association between physical exercise and mental health in 1·2 million individuals in the USA between 2011 and 2015: a cross-sectional study. The Lancet Psychiatry 2018 Sep;5(9):739-746 [FREE Full text] [doi: 10.1016/s2215-0366(18)30227-x]
43. Zhang F, Yin X, Bi C, Ji L, Wu H, Li Y, et al. Psychological symptoms are associated with screen and exercise time: a cross-sectional study of Chinese adolescents. BMC Public Health 2020 Nov 12;20(1):1695 [FREE Full text] [doi: 10.1186/s12889-020-09819-7] [Medline: 33183262]
44. Samji H, Wu J, Ladak A, Vossen C, Stewart E, Dove N, et al. Review: Mental health impacts of the COVID-19 pandemic on children and youth - a systematic review. Child Adolesc Ment Health 2022 May 28;27(2):173-189 [FREE Full text] [doi: 10.1111/camh.12501] [Medline: 34455683]
45. So M, Perry NB, Langenfeld AD, Barnes AJ. Adolescent Sleep and Mental Health Across Race/Ethnicity: Does Parent-Child Connectedness Matter? J Dev Behav Pediatr 2021 Apr 22;42(9):742-750. [doi: 10.1097/dbp.0000000000000958]

Abbreviations

aOR: adjusted odds ratio
CEFIS: COVID-19 Exposure and Family Impact Scales
PSC-17: Pediatric Symptom Checklist-17 Items

Edited by A Mavragani, T Sanchez; submitted 20.10.22; peer-reviewed by A AL-Asadi, L Shi; comments to author 01.12.22; revised version received 19.12.22; accepted 06.01.23; published 07.02.23.

Please cite as:
Liu X, Wu J, Yang H, Zhao F, Qin Y, Wu J, Yan H, Xu Y, Zhang L
Caregiver Perceptions of Children's and Adolescents' Psychosocial Functioning During the Stringent COVID-19 Lockdown Restrictions in Shanghai: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e43689
URL: https://publichealth.jmir.org/2023/1/e43689
doi:10.2196/43689
PMID:36749625

Original Paper

Safety of COVID-19 Vaccination in Patients With Breast Cancer: Cross-Sectional Study in China

Shaohua Zhang[1*], MD; Jianbin Li[1,2*], MD; Ruonan Xu[1], MD; Qianjun Chen[3], MD; Gang Sun[4], MD; Ying Lin[5], MD; Yali Cao[6], MD; Yiding Chen[7], MD; Cuizhi Geng[8], MD; Yuee Teng[9], MD; Jianyun Nie[10], MD; Xinzheng Li[11], MD; Guiying Xu[12], MD; Xinlan Liu[13], MD; Feng Jin[9], MD; Zhimin Fan[14], MD; Ting Luo[15], MD; Hong Liu[16], MD; Fu-sheng Wang[1], MD; Zefei Jiang[1], MD

[1]Fifth Medical Center of Chinese People's Liberation Army General Hospital, Beijing, China
[2]Beijing Institute of Biotechnology, Academy of Military Medical Sciences, Beijing, China
[3]Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangdong, China
[4]Cancer Hospital Affiliated to Xinjiang Medical University, Xinjiang, China
[5]The First Affiliated Hospital of Sun Yat-sen University, Guangdong, China
[6]Nanchang Third Hospital, Jiangxi, China
[7]The Second Affiliated Hospital of Medical College of Zhejiang University, Zhejiang, China
[8]Fourth Hospital of Hebei Medical University, Hebei, China
[9]The First Hospital of China Medical University, Liaoning, China
[10]Yunnan Cancer Hospital, Yunnan, China
[11]Shanxi Cancer Hospital, Shanxi, China
[12]Jilin Cancer Hospital, Jilin, China
[13]General Hospital of Ningxia Medical University, Ningxia, China
[14]The First Hospital of Jilin University, Jilin, China
[15]Sichuan Uniersity Huaxi Campus, Sichuan, China
[16]Tumor Hospital of Tianjin Medical University, Tianjin, China
[*]these authors contributed equally

Corresponding Author:
Zefei Jiang, MD
Fifth Medical Center of Chinese People's Liberation Army General Hospital
East street of Fengtai District
Beijing, 100071
China
Phone: 86 18518277411
Email: jiangzefei@csco.org.cn

Abstract

Background: The widespread use of vaccines against the novel coronavirus disease (COVID-19) has become one of the most effective means to establish a population immune barrier. Patients with cancer are vulnerable to COVID-19 infection, adverse events, and high mortality, and should be the focus of epidemic prevention and treatment. However, real-world data on the safety of vaccines for patients with breast cancer are still scarce.

Objective: This study aims to compare the safety of COVID-19 vaccines between patients vaccinated before or after being diagnosed with breast cancer.

Methods: Patients with breast cancer who sought medical advice from October 2021 to December 2021 were screened. Those who received COVID-19 vaccines were enrolled in this study to analyze the safety of the vaccines. The primary outcome was patient-reported adverse events (AEs). All events after vaccine injection were retrospectively documented from the patients.

Results: A total of 15,455 patients with breast cancer from 41 hospitals in 20 provinces in China were screened, and 5766 patients who received COVID-19 vaccines were enrolled. Of those enrolled, 45.1% (n=2599) of patients received vaccines before breast cancer diagnosis, 41.3% (n=2379) were vaccinated after diagnosis, and 13.6% (n=784) did not known the accurate date of vaccination or cancer diagnosis. Among the patients vaccinated after diagnosis, 85.4% (n=2032) were vaccinated 1 year after cancer diagnosis and 95.4% (n=2270) were vaccinated during early-stage cancer. Of all 5766 vaccinated patients, 93.9% (n=5415) received an inactivated vaccine, 3.7% (n=213) received a recombinant subunit vaccine, and 2.4% (n=138) received other vaccines, including adenovirus and mRNA vaccines. In the first injection of vaccines, 24.4% (n=10, 95% CI 11.2-37.5) of patients who received an adenovirus vaccine reported AEs, compared to only 12.5% (n=677, 95% CI 11.6-13.4) of those who received an inactivated vaccine. Patients with metastatic breast cancer reported the highest incidence of AEs (n=18, 16.5%, 95% CI 9.5-23.5). Following the second injection, patients who received an inactivated vaccine (n=464, 8.7%, 95% CI 8.0-9.5) and those who received a recombinant vaccine (n=25, 8.7%, 95% CI 5.5-12.0) reported the same incidence of AEs. No significant differences in patient-reported AEs were found between the healthy population and patients with breast cancer (16.4% vs 16.9%, respectively); the most common AEs were local pain (11.1% vs 9.1%, respectively), fatigue (5.5% vs 6.3%, respectively), and muscle soreness (2.3% vs 3.6%, respectively). The type of vaccine and time window of vaccination had little impact on patient-reported AEs.

Conclusions: Compared with patients vaccinated before breast cancer diagnosis, there were no significant differences in patient-reported AEs in the patients vaccinated after diagnosis. Thus, it is safe for patients with breast cancer, especially for those in the early stage, to receive COVID-19 vaccines.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2200055509; https://tinyurl.com/33zzj882

(JMIR Public Health Surveill 2023;9:e46009) doi:10.2196/46009

KEYWORDS

breast cancer; COVID-19 vaccines; patients reported adverse events; healthy population; vaccine safety

Introduction

The outbreak of COVID-19 has had a huge impact on the lives and work of people worldwide [1-3]. Although several drugs have been authorized for emergency use by many countries [4,5], inaccessibility and strong demand have limited their widespread use during the pandemic. The establishment of a population immune barrier through mass vaccination with COVID-19 vaccines may be the one of the most effective means to end the pandemic at this moment [6]. However, for certain populations, safety concerns of vaccines may have negative impacts on widespread vaccination, thus weakening the process of herd immunity. In this process, vaccination of specific patients is an important part of epidemic prevention and control [7,8].

Cancers are the second leading cause of death worldwide, with high morbidity and mortality rates in the general population. In the setting of the SARS-CoV-2 pandemic, patients with cancer are at high risk of infection, and the safety of these patients after COVID-19 vaccination is an important consideration. Aberrant immune responses in the context of underlying cancer and the use of anticancer therapies may contribute to impaired immune responses and altered reactogenicity following immunization against SARS-CoV-2 [9,10]. Therefore, patients with tumors have had a high risk of infection, high incidence of serious events, and high mortality in the COVID-19 pandemic [11]. As a special group, there is still confusion on whether the safety of COVID-19 vaccines in such patients is the same as in healthy individuals because of their immunosuppression [12,13].

For patients with breast cancer, there has been a history of receiving vaccines for cancer treatment [14]. However, the COVID-19 vaccination rate remains low in China due to inadequate health education, the lack of vaccine safety data, and even some other social status influences rather than vaccine shortages [15-17]. The fear of delaying cancer therapy has surpassed the fear of COVID-19 infection [18]. To better understand why some patients with cancer are unwilling to receive vaccines and the safety profiles of those who have received vaccines, we conducted this real-world study, called CSCO BC NCP-02, from cancer centers across China [19]. Notably, we found that half of the patients in the study received COVID-19 vaccines before cancer diagnosis. This allowed us to compare the differences in adverse events (AEs) between patients who were vaccinated before or after breast cancer diagnosis, thus providing more data to encourage patients with breast cancer to be vaccinated with a COVID-19 vaccine. We aimed to describe the clinical features and AEs of patients with breast cancer and identify risk factors associated with AEs in these patients.

Methods

Sample and Data

This was a cross-sectional study initiated by the Chinese Society of Clinical Oncology breast cancer (CSCO BC) committee. The inclusion criteria were as follows: (1) patients who received at least 1 dose of a COVID-19 vaccine, (2) patients diagnosed as having invasive breast cancer, (3) inpatients who visited the participating hospital from November 2021 to December 2021, and (4) patients with more than 3 months of overall survival. The exclusion criteria were as follows: (1) patients with microinvasion of ductal carcinoma in situ breast cancer and (2) patients with a second primary tumors.

Different types of vaccines were recorded. Dosing regimens were as follows: 1 dose for adenovirus vectored vaccines, 2 doses for inactivated viral vaccines and mRNA vaccines, and 3 doses for recombinant vaccines.

Procedures

Inpatients with breast cancer who were vaccinated were asked to provide retrospective answers through self-reported questionnaires. For highly professional medical issues, physicians were asked to assist the patients to complete questionnaires. Differences for those who received vaccines before cancer diagnosis versus after cancer diagnosis were compared according to the time of cancer diagnosis and vaccination. Patients were recruited in hospitals by the CSCO BC committee. By adopting a competitive enrollment method, only hospitals that contributed more than 20 valid questionnaires by the cutoff date were included.

Through telephone follow-ups or questionnaire surveys, data including epidemiology, demography, clinical information, previous and current therapies, type of vaccines, and patient-reported AEs and their severity after vaccination were collected nationwide. Medical workers of participating hospitals were responsible for data collection, and at least 1 professional doctor from each hospital was responsible for data identification. To ensure the quality of the data, we reviewed all the data, and we retrieved missing data by answering questions. All the data were checked by 2 physicians.

Ethical Considerations

This study was registered (ChiCTR2200055509) and was approved by the ethics board of the Fifth Medical Center of the Chinese People's Liberation Army General Hospital (KY-2022-7-45-1). Oral informed consent was obtained from participants, and secondary analysis was allowed without additional consent. The study data were deidentified. No compensation was involved in this study.

Measures of Variables

The outcome was the safety of the COVID-19 vaccines in patients with breast cancer, which was mainly calculated by dividing the number of AEs by the total population. Another observed outcome was the incidence of AEs, which was calculated by dividing the number of patients with AEs by the total population. Patients were divided into 3 groups according to their time of vaccination and cancer diagnosis: vaccinated before cancer diagnosis, vaccinated after diagnosis, or unknown. AEs were analyzed according to the vaccination time and different doses of the vaccines.

Data Analysis Procedure

SPSS 21.0 (IBM) was used for data statistics. Pearson χ^2 was used for data analysis and a 2-way ANOVA was used to detect data differences. Frequency tables were analyzed using the χ^2 test or Fisher exact test. A multivariable logistic regression model was fitted to examine possible differences in AEs. Odds ratios (ORs) and 95% CIs were calculated with Mantel-Haenszel models. A 2-sided $\alpha<.05$ was considered statistically significant. GraphPad Prism v6.0 was used as the drawing software.

Results

A total of 15,455 patients from 41 hospitals in China were screened (Figure 1), and 5766 patients were eligible in this analysis, of which 45.1% (n=2599) were vaccinated before cancer diagnosis, 41. 2% (n=2379) were vaccinated after diagnosis, and 13.7% (n=788) lacked an accurate diagnosis time or vaccination time. The baseline characteristics of the patients are detailed in Table 1.

Among the patients vaccinated before breast cancer diagnosis, 57.9% (n=1505) were diagnosed with breast cancer 1-3 months after vaccination and 34.4% (n=963) were diagnosed 4-6 months after vaccination. Among the patients vaccinated after breast cancer diagnosis, 85.4% (n=2032) received the vaccine 1 year after cancer diagnosis (Figure 2).

Of all the patients, 93.9% (n=5415) received an inactivated vaccine, 3.7% (n=213) received a recombinant vaccine, and 2.4% (n=138) received other vaccines, including adenovirus and mRNA vaccines. Statistical analysis showed that more patient-reported AEs occurred following the first dose than following the second dose (Figure 3A). Following the first injection, 24.4% (n=10, 95% CI 11.2-37.5) of patients who received an adenovirus vaccine reported AEs compared to 12.5% (n=677, 95% CI 11.6-13.4) for those who received an inactivated vaccine. Following the second dose, patients who received an inactivated vaccine (n=464, 8.7%, 95% CI 8.0-9.5) or a recombinant vaccine (n=25, 8.7%, 95% CI 5.5-12.0) reported the same incidence of AEs (Figure S1 in Multimedia Appendix 1).

Considering the timing of vaccination, of the patients vaccinated before diagnosis, 13.1% (n=340, 95% CI 11.8-14.4) and 8.6% (n=219, 95% CI 7.5-9.7) reported AEs following their first and second injections, respectively (Figure 3B). The cohort included patients in the early stage and patients with metastatic breast cancer. In the early stage, 10.1% (n=227, 95% CI 8.9-11.4) of patients reported AEs following the second dose. Of the patients with metastatic breast cancer, 16.5% (n=18, 95% CI 9.5-23.5) reported AEs after the first dose and 9.3% (n=10, 95% CI 3.8-14.7) after the second dose.

Of all the patients included in this study, a total of 16.4% (n=427) and 16.9% (n=402) of patients vaccinated before or after cancer diagnosis, respectively, reported AEs (Table 2). The most common AEs were local pain (11.1% vs 9.1%, respectively), fatigue (5.5% vs 6.3%, respectively), and muscle soreness (2.3% vs 3.6%, respectively). Other reported AEs included local swelling, local induration, headache, fever, joint pain, and nausea, among others. The incidence of grade 3/4 AEs was relatively low. The highest incidence was for fever, which was reported by 0.4% (n=11) of patients vaccinated before diagnosis and 0.2% (n=5) of those vaccinated after diagnosis.

To identify the factors associated with AEs, a multivariate analysis was conducted (Table S1 in Multimedia Appendix 1). The type of vaccine (odds ratio [OR]=0.937, 95% CI 0.689-1.273) and the time window of vaccination (OR=0.919, 95% CI 0.785-1.077) had little impact on AEs. Patients with low educational degrees reported a lower incidence of AEs (OR=0.794, 95% CI 0.680-0.927) (Table S2 in Multimedia Appendix 1).

Figure 1. Flow chart of the study.

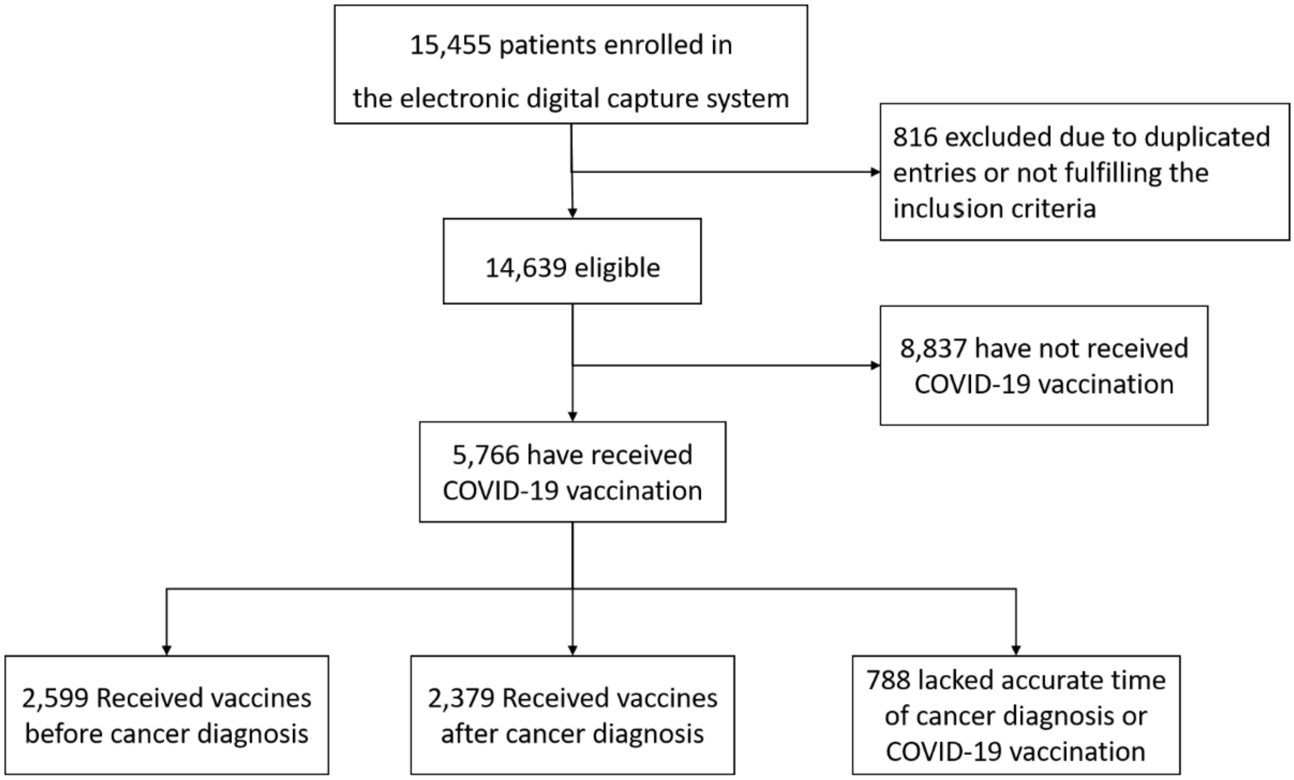


Table 1. Baseline characteristics of vaccinated patients.

Characteristic	Vaccinated before diagnosis (n=2599), n (%)	Vaccinated after diagnosis (n=2379), n (%)	Timing unknown (n=788), n (%)	Total (n=5766), n (%)
Age (years)				
≤45	904 (34.8)	935 (39.3)	211 (26.8)	2050 (35.6)
>45	1632 (62.8)	1345 (56.5)	431 (54.7)	3408 (59.1)
Unknown	63 (2.4)	99 (4.2)	146 (18.5)	308 (5.2)
Chronic illness				
No	2143 (82.5)	1841 (77.4)	657 (83.4)	1107 (19.2)
Yes	456 (17.5)	538 (22.6)	131 (16.6)	4659 (80.8)
Level of education				
Without bachelor degree	1575 (60.6)	1543 (64.8)	462 (58.6)	3580 (62.1)
With bachelor degree	1013 (39)	812 (34.1)	210 (26.6)	2035 (35.3)
Unknown	11 (0.4)	24 (1)	116 (14.7)	151 (2.5)
T stage				
T1-2	1927 (74.1)	1612 (67.8)	322 (40.8)	3861 (67)
T3-4	199 (7.7)	112 (4.7)	25 (3.2)	336 (5.8)
Unknown	473 (18.2)	655 (27.5)	441 (56)	1569 (27.2)
N stage				
N0	1115 (42.9)	1096 (46.1)	215 (27.3)	2426 (42.1)
N1-3	1051 (40.4)	727 (30.6)	139 (17.6)	1917 (33.2)
Unknown	433 (16.7)	556 (23.3)	434 (55.1)	1423 (24.6)
HR[a]				
Positive	1523 (58.6)	1755 (73.8)	246 (31.2)	3524 (61.1)
Negative	626 (24.1)	439 (18.5)	97 (12.3)	1162 (20.2)
Unknown	450 (17.3)	185 (7.7)	445 (56.5)	1080 (18.7)
HER2[b]				
Positive	825 (31.7)	618 (26)	118 (15)	1561 (27.1)
Negative	1324 (50.9)	1331 (55.9)	225 (28.5)	2880 (50)
Unknown	450 (17.3)	430 (18.1)	445 (56.5)	1325 (22.9)
Neoadjuvant				
Yes	709 (27.3)	373 (15.7)	95 (12.1)	1177 (20.4)
No	1890 (72.7)	2006 (84.3)	693 (87.9)	4589 (79.6)
Surgery				
Yes	2089 (80.4)	2295 (96.5)	403 (51.1)	4787 (83)
No	510 (19.6)	84 (3.5)	385 (48.9)	979 (17)
Adjuvant				
Yes	1704 (65.6)	2105 (88.5)	284 (36)	4093 (71)
No	895 (34.4)	274 (11.5)	504 (64)	1673 (29)
Metastatic				
Yes	128 (4.9)	203 (8.5)	11 (1.4)	342 (5.9)
No	2471 (95.1)	2176 (91.5)	777 (98.6)	5424 (94.1)

[a] HR: hormone receptor.

[b] HER2: human epidermal growth factor receptor 2.

Figure 2. Time interval between COVID-19 vaccination and cancer diagnosis. Those without accurate dates of breast cancer diagnosis or COVID-19 vaccination were excluded.

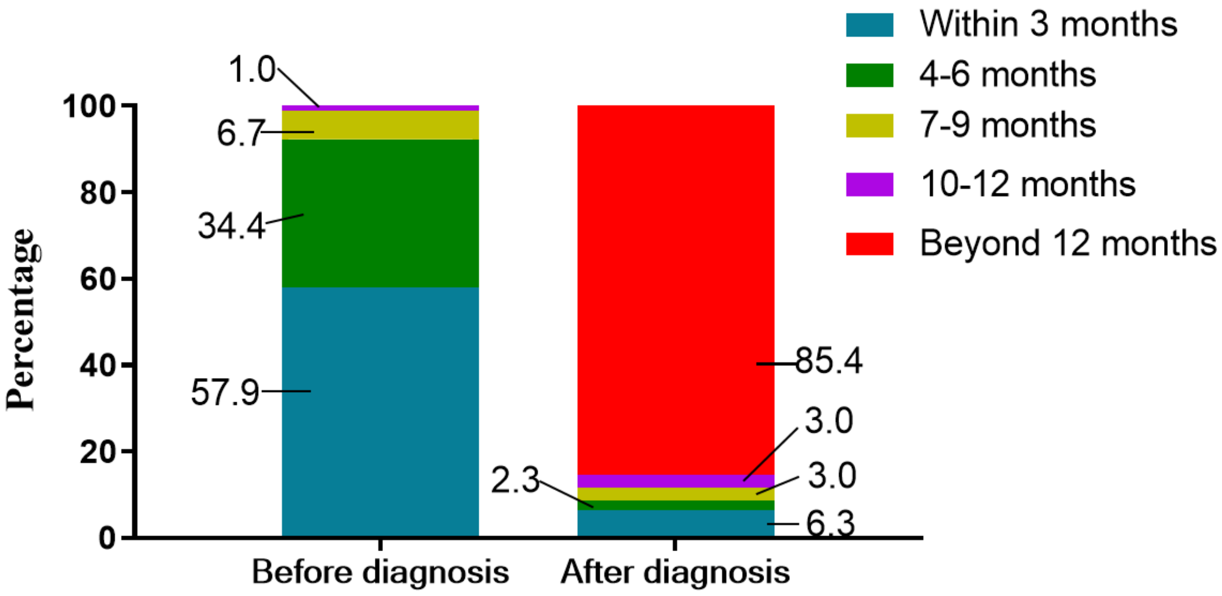


Figure 3. Distribution of adverse events. (A) Incidence of adverse events in patients vaccinated with different vaccines. (B) Incidence of adverse events in patients vaccinated at different stages of breast cancer.

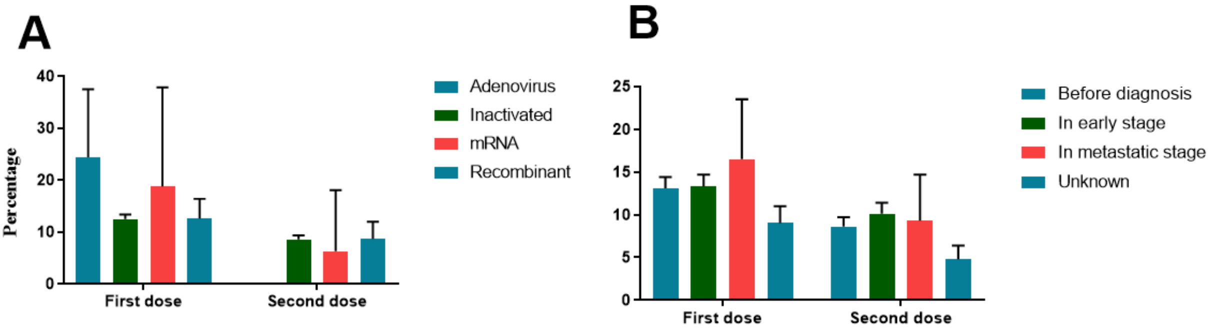


Table 2. Safety analysis of COVID-19 vaccines in the enrolled patients.

	Vaccinated before diagnosis (n=2599), n (%)	Vaccinated after diagnosis (n=2379), n (%)	Timing unknown (n=788), n (%)
Adverse events	427 (16.4)	402 (16.9)	106 (13.5)
Local reactions			
Local pain			
Any	288 (11.1)	218 (9.1)	38 (4.8)
Grade 3/4	2 (0.1)	3 (0.1)	0 (0)
Local swelling			
Any	43 (1.7)	33 (1.4)	7 (0.9)
Grade 3/4	2 (0.1)	2 (0.1)	0 (0)
Subcutaneous nodules			
Any	32 (1.2)	9 (0.4)	6 (0.8)
Grade 3/4	4 (0.1)	1 (<0.1)	0 (0)
Systemic reactions			
Fever			
Any	29 (1.1)	25 (1)	6 (0.6)
≥38 °C	11 (0.4)	5 (0.2)	3 (0.3)
Headache			
Any	48 (1.8)	41 (1.7)	8 (1)
Grade 3/4	2 (0.1)	1 (<0.1)	0 (0)
Fatigue			
Any	142 (5.5)	150 (6.3)	36 (4.6)
Grade 3/4	3 (0.1)	0 (0)	0 (0)
Muscle soreness			
Any	61 (2.3)	87 (3.6)	13 (1.7)
Grade 3/4	2 (0.1)	1 (<0.1)	0 (0)
Joint pain			
Any	18 (0.7)	24 (1)	4 (0.5)
Grade 3/4	2 (0.1)	3 (0.1)	0 (0)
Nausea			
Any	27 (1)	22 (0.9)	2 (0.3)
Grade 3/4	2 (0.1)	1 (<0.1)	0 (0)
Loss of appetite			
Any	12 (0.5)	10 (0.4)	4 (0.5)
Grade 3/4	2 (0.1)	1 (<0.1)	0 (0)
Anaphylaxis			
Any	6 (0.2)	15 (0.6)	0 (0)
Grade 3/4	0 (0)	0 (0)	0 (0)
Dizzy			
Any	4 (0.2)	12 (0.5)	0 (0)
Grade 3/4	0 (0)	0 (0)	0 (0)
Disturbance in respiration			
Any	0 (0)	5 (0.2)	0 (0)

	Vaccinated before diagnosis (n=2599), n (%)	Vaccinated after diagnosis (n=2379), n (%)	Timing unknown (n=788), n (%)
Grade 3/4	0 (0)	0 (0)	0 (0)
Breast pain			
Any	5 (0.2)	1 (<0.1)	0 (0)
Grade 3/4	0 (0)	0 (0)	0 (0)
Thirsty			
Any	6 (0.2)	0 (0)	0 (0)
Grade 3/4	0 (0)	0 (0)	0 (0)
Diarrhea			
Any	1 (<0.1)	2 (0.1)	0 (0)
Grade 3/4	0 (0)	1 (<0.1)	0 (0)
Angina pectoris			
Any	1 (<0.1)	4 (0.2)	0 (0)
Grade 3/4	0 (0)	1 (<0.1)	0 (0)
Others			
Any	23 (0.9)	20 (0.8)	0 (0)
Grade 3/4	0 (0)	0 (0)	0 (0)

Discussion

Principal Findings

This was a real-world study to compare patient-reported AEs among patients vaccinated before or after breast cancer diagnosis. The results have several implications. First, most patients vaccinated after breast cancer diagnosis received their vaccines 1 year after diagnosis, indicating that most patients with early breast cancer received vaccines after chemotherapy due to the short therapeutic period. Second, the incidence of AEs in patients with early-stage breast cancer was lower than in those with metastatic breast cancer. When compared with patients vaccinated before diagnosis, the reported AEs from patients vaccinated after diagnosis was not increased. Third, patients reported a higher incidence of AEs following their first dose compared to following their second dose. The most common AEs were local pain and fatigue. The type of vaccine and time window of vaccination had little impact on AEs.

In this study, of those vaccinated before diagnosis, 90% were diagnosed with breast cancer within 6 months of vaccination. Ipsilateral axillary swelling was one of most frequently reported local reactions to the COVID-19 vaccine, occurring in 11.6% and 16% of recipients following the first and second dose, respectively [20,21]. The lymphadenopathy caused by vaccination may encourage them to seek specialist treatment, which was an important reason for the higher rate of positive lymph nodes on diagnosis in those patients [22]. For patients vaccinated after their diagnosis with breast cancer, most of them received a COVID-19 vaccine 1 year after their diagnosis. Due to the fact that 91.5% of these patients were in the early stage, this contributed to a large proportion of patients being vaccinated without any chemotherapy, which echoed the recommendations in the Chinese expert consensus [23] and mitigated the negative impact on the immune system from chemotherapy [24,25]. The promotion of the consensus does have potential benefits for elevating the rate of vaccination for patients with cancer [26].

In terms of different kinds of vaccines, the majority of patients in this study chose inactivated vaccines (eg, Beijing Biotech, Wuhan Biotech, and Sinovac). Compared with other types of vaccines [27,28], inactivated vaccines presented relatively good safety data in patients with breast cancer [29], although no significant result was shown in the multivariate analysis. We also found that there were significant differences in the incidence of AEs in patients with breast cancer at different stages. Compared with patients in the early stage, the incidence of AEs was higher in late-stage patients after the first dose of vaccine. For one thing, the overlap of the toxicity from the vaccine and cancer treatment in advanced breast cancer may increase the incidence of AEs [30,31]. For another, considering the complexity of antigens, the immune response may have been more aggressive, which can also lead to a significant increase in AEs [32].

The common AEs in this study included local pain, fatigue, muscle soreness, local swelling, local induration, headache, and fever. The incidence of AEs in this study was much lower than in a published randomized controlled study [33]. In our multivariable analysis, we found that the time window of vaccination and type of vaccine had little influence on AEs. Notably, patients with different education levels exhibited a significant difference in the incidence of AEs; patients with a low degree of education had a lower incidence of AEs. The limitation of real-world studies attributed to the lack of effective guidance during the investigation and incomplete data from case report forms may be one of the reasons [34]. To avoid declines and delays in cancer treatment after COVID-19 quarantine restrictions [35], patients might conceal some discomfort from their physicians. During cancer therapy, some AEs might be regarded as treatment-related AEs and ignored by patients and clinicians. In addition, patients with severe AEs might be transferred to a specialist hospital rather than a cancer hospital, which can also result in an underestimated incidence of AEs in this study.

As a retrospective real-word study, inevitable selection bias is another limitation. Over 90% of patients received an inactivated vaccine, so the safety data of other vaccines in patients with breast cancer might be premature. We also paid less attention to the efficacy of vaccines, and the originally planned follow-up of infections has also been terminated due to indiscriminative infection of COVID-19 after the end of quarantine restrictions in China since December 2022. Fortunately, we have collected blood samples from some of the patients to explore changes in antibody titers, which might be a remaining effective alternative indicator to evaluate the efficacy of vaccines [36,37]. Additionally, owing to the high possibility of selection bias and measurement bias, although no data support the notion that COVID-19 vaccination would induce breast cancer, much importance should be attached to this kind of population.

Conclusions

Despite the limitations, our retrospective study highlights the safety of COVID-19 vaccination in patients with breast cancer, and these patients are recommended to receive COVID-19 vaccines, just like healthy individuals. Much attention should be paid to the first injection of vaccines. While limited by retrospective data, more evidence from high quality, double-blind, randomized control trials are expected.

Acknowledgments

This study was supported by the Emergency Key Program of Guangzhou Laboratory (grant EKPG21-30-4) and the Beijing Science and Technology Plan (grant Z1811000 01718215).

Data Availability

Access to the data set supporting the findings of this study can only legally be accessed on acceptance from the Chinese Society of Clinical Oncology breast cancer database collaborative group. Interested parties can contact the corresponding author for further information. We confirm that others would be able to access these data in the same manner as the authors. We also confirm that the authors do not have any special access privileges that others would not have.

Authors' Contributions

JL and ZJ formed the initial proposal, designed the study, and developed the protocol. SZ and JL wrote the first edition of this manuscript. JL analyzed the data and revised the manuscript. All authors were responsible for enrolment in this study, contributed to this research, and read and approved the final version of the manuscript before submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figure and tables.
[DOCX File , 166 KB - publichealth_v9i1e46009_app1.docx]

References

1. Woolf SH, Chapman DA, Lee JH. COVID-19 as the leading cause of death in the United States. JAMA 2021 Jan 12;325(2):123-124 [FREE Full text] [doi: 10.1001/jama.2020.24865] [Medline: 33331845]
2. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet 2022 Apr 16;399(10334):1513-1536 [FREE Full text] [doi: 10.1016/S0140-6736(21)02796-3] [Medline: 35279232]
3. Coccia M. Pandemic prevention: lessons from COVID-19. Encyclopedia 2021 May 31;1(2):433-444. [doi: 10.3390/encyclopedia1020036]
4. Cao Z, Gao W, Bao H, Feng H, Mei S, Chen P, et al. VV116 versus nirmatrelvir-ritonavir for oral treatment of Covid-19. N Engl J Med 2023 Mar 02;388(5):406-417 [FREE Full text] [doi: 10.1056/NEJMoa2208822] [Medline: 36577095]
5. Gottlieb RL, Vaca CE, Paredes R, Mera J, Webb BJ, Perez G, et al. Early remdesivir to prevent progression to severe Covid-19 in outpatients. N Engl J Med 2022 Jan 27;386(4):305-315. [doi: 10.1056/nejmoa2116846]
6. Andrews N, Tessier E, Stowe J, Gower C, Kirsebom F, Simmons R, et al. Duration of protection against mild and severe disease by Covid-19 vaccines. N Engl J Med 2022 Jan 27;386(4):340-350. [doi: 10.1056/nejmoa2115481]
7. Stock S, Carruthers J, Calvert C, Denny C, Donaghy JG, Goulding A, et al. SARS-CoV-2 infection and COVID-19 vaccination rates in pregnant women in Scotland. Nat Med 2022 Mar;28(3):504-512 [FREE Full text] [doi: 10.1038/s41591-021-01666-2] [Medline: 35027756]

8. Mayer MA, Vidal-Alaball J, Puigdellívol-Sánchez A, Marín Gomez FX, Leis A, Mendioroz Peña J. Clinical characterization of patients with COVID-19 in primary care in Catalonia: retrospective observational study. JMIR Public Health Surveill 2021 Mar 08;7(2):e25452 [FREE Full text] [doi: 10.2196/25452] [Medline: 33496668]
9. Tran S, Truong TH, Narendran A. Evaluation of COVID-19 vaccine response in patients with cancer: an interim analysis. Eur J Cancer 2021 Dec;159:259-274 [FREE Full text] [doi: 10.1016/j.ejca.2021.10.013] [Medline: 34798454]
10. Hosseini M, Seyedpour S, Khodaei B, Loghman A, Seyedpour N, Yazdi M, et al. Cancer vaccines for triple-negative breast cancer: a systematic review. Vaccines (Basel) 2023 Jan 09;11(1):146 [FREE Full text] [doi: 10.3390/vaccines11010146] [Medline: 36679991]
11. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. Lancet Oncol 2020 Mar;21(3):335-337. [doi: 10.1016/s1470-2045(20)30096-6]
12. Fendler A, de Vries EGE, GeurtsvanKessel CH, Haanen JB, Wörmann B, Turajlic S, et al. COVID-19 vaccines in patients with cancer: immunogenicity, efficacy and safety. Nat Rev Clin Oncol 2022 Jun;19(6):385-401 [FREE Full text] [doi: 10.1038/s41571-022-00610-8] [Medline: 35277694]
13. Ribas A, Sengupta R, Locke T, Zaidi S, Campbell K, Carethers JM, et al. Priority COVID-19 vaccination for patients with cancer while vaccine supply is limited. Cancer Discov 2021 Mar;11(2):233-236 [FREE Full text] [doi: 10.1158/2159-8290.CD-20-1817] [Medline: 33355178]
14. Zhou Y. HER2/neu-based vaccination with li-Key hybrid, GM-CSF immunoadjuvant and trastuzumab as a potent triple-negative breast cancer treatment. J Cancer Res Clin Oncol 2023 Aug;149(9):6711-6718 [FREE Full text] [doi: 10.1007/s00432-023-04574-9] [Medline: 36692548]
15. Laine C, Cotton D, Moyer DV. COVID-19 vaccine: promoting vaccine acceptance. Ann Intern Med 2021 Feb;174(2):252-253. [doi: 10.7326/m20-8008]
16. Villarreal-Garza C, Vaca-Cartagena BF, Becerril-Gaitan A, Ferrigno AS, Mesa-Chavez F, Platas A, et al. Attitudes and factors associated with COVID-19 vaccine hesitancy among patients with breast cancer. JAMA Oncol 2021 Aug 01;7(8):1242-1244 [FREE Full text] [doi: 10.1001/jamaoncol.2021.1962] [Medline: 34110371]
17. Chagpar AB, Coccia M. Factors associated with breast cancer mortality-per-incident case in low-to-middle income countries (LMICs). J Clin Oncol 2019 May 20;37(15_suppl):1566-1566. [doi: 10.1200/jco.2019.37.15_suppl.1566]
18. Song D, Shabani J, Jaiswal V, Cagliostro M, Rubinstein D, Alraies MC. Delayed doxorubicin induced cardiomyopathy in a breast cancer patient: a case report. Radiol Case Rep 2023 Mar;18(3):1256-1260 [FREE Full text] [doi: 10.1016/j.radcr.2022.12.048] [Medline: 36691410]
19. Coccia M. Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: a global analysis. Environ Res 2022 Mar;204:112314. [doi: 10.1016/j.envres.2021.112314]
20. Özütemiz C, Krystosek LA, Church AL, Chauhan A, Ellermann JM, Domingo-Musibay E, et al. Lymphadenopathy in COVID-19 vaccine recipients: diagnostic dilemma in oncologic patients. Radiology 2021 Jul;300(1):E296-E300 [FREE Full text] [doi: 10.1148/radiol.2021210275] [Medline: 33625300]
21. Bshesh K, Khan W, Vattoth AL, Janjua E, Nauman A, Almasri M, et al. Lymphadenopathy post-COVID-19 vaccination with increased FDG uptake may be falsely attributed to oncological disorders: a systematic review. J Med Virol 2022 May 07;94(5):1833-1845 [FREE Full text] [doi: 10.1002/jmv.27599] [Medline: 35060149]
22. Wolfson S, Kim E. Breast Cancer Screening and Axillary Adenopathy in the Era of COVID-19 Vaccination. Radiology 2023 Mar;306(2):e222040. [doi: 10.1148/radiol.222040] [Medline: 36219117]
23. Nie J, Wang T, Xu R, Li J, Yin Y, Wang X, et al. Chinese expert consensus on coronavirus disease 2019 vaccination for breast cancer patients. Transl Breast Cancer Res 2021 Oct;2:34-34. [doi: 10.21037/tbcr-21-31]
24. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature 2017 Jan 18;541(7637):321-330. [doi: 10.1038/nature21349] [Medline: 28102259]
25. Wack S, Patton T, Ferris LK. COVID-19 vaccine safety and efficacy in patients with immune-mediated inflammatory disease: review of available evidence. J Am Acad Dermatol 2021 Nov;85(5):1274-1284. [doi: 10.1016/j.jaad.2021.07.054] [Medline: 34363909]
26. Li S, Cui G, Kaminga AC, Cheng S, Xu H. Associations between health literacy, eHealth literacy, and COVID-19-related health behaviors among Chinese college students: cross-sectional online study. J Med Internet Res 2021 May 06;23(5):e25600. [doi: 10.2196/25600] [Medline: 33822734]
27. Zhu F, Guan X, Li Y, Huang J, Jiang T, Hou L, et al. Immunogenicity and safety of a recombinant adenovirus type-5-vectored COVID-19 vaccine in healthy adults aged 18 years or older: a randomised, double-blind, placebo-controlled, phase 2 trial. Lancet 2020 Aug 15;396(10249):479-488 [FREE Full text] [doi: 10.1016/S0140-6736(20)31605-6] [Medline: 32702299]
28. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416. [doi: 10.1056/nejmoa2035389]
29. Al Kaabi N, Zhang Y, Xia S, Yang Y, Al Qahtani MM, Abdulrazzaq N, et al. Effect of 2 inactivated SARS-CoV-2 vaccines on symptomatic COVID-19 infection in adults: a randomized clinical trial. JAMA 2021 Jul 06;326(1):35-45 [FREE Full text] [doi: 10.1001/jama.2021.8565] [Medline: 34037666]

30. Monin L, Laing AG, Muñoz-Ruiz M, McKenzie DR, del Molino del Barrio I, Alaguthurai T, et al. Safety and immunogenicity of one versus two doses of the COVID-19 vaccine BNT162b2 for patients with cancer: interim analysis of a prospective observational study. Lancet Oncol 2021 Jun;22(6):765-778. [doi: 10.1016/s1470-2045(21)00213-8]
31. De Placido P, Pietroluongo E, De Angelis C, Tafuro M, Barraco C, Giannatiempo R, et al. Safety and immunogenicity of the COVID-19 vaccine BNT162b2 for patients with breast and gynecological cancer on active anticancer therapy: results of a prospective observational study. Front Oncol 2022;12:951026 [FREE Full text] [doi: 10.3389/fonc.2022.951026] [Medline: 36059663]
32. Naranbhai V, Pernat CA, Gavralidis A, St Denis KJ, Lam EC, Spring LM, et al. Immunogenicity and reactogenicity of SARS-CoV-2 vaccines in patients with cancer: the CANVAX cohort study. J Clin Oncol 2022 Jan 01;40(1):12-23. [doi: 10.1200/jco.21.01891]
33. Li M, Wang H, Tian L, Pang Z, Yang Q, Huang T, et al. COVID-19 vaccine development: milestones, lessons and prospects. Signal Transduct Target Ther 2022 May 03;7(1):146 [FREE Full text] [doi: 10.1038/s41392-022-00996-y] [Medline: 35504917]
34. Carrozzino D, Patierno C, Guidi J, Berrocal Montiel C, Cao J, Charlson M, et al. Clinimetric criteria for patient-peported outcome measures. Psychother Psychosom 2021;90(4):222-232 [FREE Full text] [doi: 10.1159/000516599] [Medline: 34038901]
35. Li J, Wang H, Geng C, Liu Z, Lin Y, Nie J, et al. Suboptimal declines and delays in early breast cancer treatment after COVID-19 quarantine restrictions in China: a national survey of 8397 patients in the first quarter of 2020. EClinicalMedicine 2020 Sep;26:100503 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100503] [Medline: 32989430]
36. Luangdilok S, Wanchaijiraboon P, Pakvisal N, Susiriwatananont T, Zungsontiporn N, Sriuranpong V, et al. Immunogenicity after a third COVID-19 mRNA booster in solid cancer patients who previously received the primary heterologous CoronaVac/ChAdOx1 vaccine. Vaccines (Basel) 2022 Sep 26;10(10):1613 [FREE Full text] [doi: 10.3390/vaccines10101613] [Medline: 36298478]
37. Benati I, Coccia M. Global analysis of timely COVID-19 vaccinations: improving governance to reinforce response policies for pandemic crises. Int J Health Gov 2022 May 31;27(3):240-253. [doi: 10.1108/ijhg-07-2021-0072]

Abbreviations

AE: adverse events
CSCO BC: Chinese Society of Clinical Oncology breast cancer
OR: odds ratio

Edited by A Mavragani, T Sanchez; submitted 25.01.23; peer-reviewed by P Dabas, D Bondage, M Coccia, Y Hirst; comments to author 07.07.23; revised version received 26.07.23; accepted 31.10.23; published 07.12.23.

Please cite as:
Zhang S, Li J, Xu R, Chen Q, Sun G, Lin Y, Cao Y, Chen Y, Geng C, Teng Y, Nie J, Li X, Xu G, Liu X, Jin F, Fan Z, Luo T, Liu H, Wang FS, Jiang Z
Safety of COVID-19 Vaccination in Patients With Breast Cancer: Cross-Sectional Study in China
JMIR Public Health Surveill 2023;9:e46009
URL: https://publichealth.jmir.org/2023/1/e46009
doi:10.2196/46009
PMID:38060302

Original Paper

Understanding Child Wasting in Ethiopia: Cross-sectional Analysis of 2019 Ethiopian Demographic and Health Survey Data Using Generalized Linear Latent and Mixed Models

Girma Gilano[1], MPH; Samuel Hailegebreal[1], MPH; Sewunet Sako[1], MPH; Firehiwot Haile[2], MPH; Kasarto Gilano[3], BSc; Binyam Tariku Seboka[4], MPH; Kefita Kashala[5], MSc

[1]Department of Health Informatics, School of Public Health, College of Medicine and Health Sciences, Arba Minch University, Arba Minch, Southern Nation Nationalities Peoples' Region, Ethiopia

[2]Department of Public Health, School of Public Health, College of Medicine and Health Sciences, Arba Minch University, Arba Minch, Southern Nation Nationalities Peoples' Region, Ethiopia

[3]Department of Public Health, Arba Minch College of Health Sciences, Arba Minch, Southern Nation Nationalities Peoples' Region, Ethiopia

[4]Department of Health Informatics, School of Public Health, College of Medicine and Health Sciences, Dilla University, Dilla, Southern Nation Nationalities Peoples' Region, Ethiopia

[5]Department of Biomedical Science, School of Medicine, College of Medicine and Health Sciences, Arba Minch University, Arba Minch, Southern Nation Nationalities Peoples' Region, Ethiopia

Corresponding Author:
Girma Gilano, MPH
Department of Health Informatics, School of Public Health
College of Medicine and Health Sciences
Arba Minch University
Arba Minch, Southern Nation Nationalities Peoples' Region
Ethiopia
Phone: 251 913 930384
Fax: 251 46 8810820
Email: gilanog@yahoo.com

Abstract

Background: Wasting is an immediate, visible, and life-threatening form of undernutrition in children aged <5 years. Within a short time, wasting causes recurrent sickness, delayed physical and mental growth, impatience, poor feeding, and low body weight. The long-term consequences of wasting and undernutrition are stunting, inability to learn, poor health status, and poor work performance. Wasting remains a public health problem in Ethiopia. According to the World Health Organization, countries have to reduce undernutrition including child wasting to below 5% by 2025. Ethiopia is attempting to attain national and international targets of undernutrition while struggling with many problems.

Objective: This study aimed to identify the prevalence and associated factors of wasting to provide information for further renewing policy commitments.

Methods: We used community-based, cross-sectional data from the Ethiopian Mini Demographic and Health Survey. The survey was conducted in 9 regions and 2 city administrations. Two-stage cluster sampling was used to recruit study participants. In the first stage, enumerations areas were selected, and 28-35 households per enumeration area were selected in the second stage. Our analysis included 2016 women with children aged <5 years from the 2019 EMDHS data set. We dropped incomplete records and included all women who fulfilled the eligibility criteria. We used multilevel ordinal regression using Generalized Linear Latent and Mixed Models (GLLAMM) and predicted probability with log-likelihood ratio tests. Fulfilling the proportional odds model's assumption during the application of multilevel ordinary logistic regression was a cumbersome task. GLLAMM enabled us to perform the multilevel proportional odds model using an alternative method.

Results: In our analysis, wasting was 7.68% (95% CI 6.56%-8.93%). Around 26.82% of mothers never used antenatal care for their current child. Most mothers (52.2%) did not have formal education, and 86.8% did not have postnatal care for their children. Additionally, half (50.93%) of the mothers have ≥6 household members. Wasting was associated with feeding diverse foods (coefficient 4.90, 95% CI 4.90-4.98), female sex of the household head (–40.40, 95% CI –40.41 to –40.32), home delivery (–35.51, 95% CI –35.55 to –35.47), first (16.66, 95% CI, 16.60-16.72) and second (16.65, 95% CI 16.60-16.70) birth order, female child (–12.65, 95% CI –12.69 to –12.62), and household size of 1 to 3 (10.86, 95% CI 10.80-10.92).

Conclusions: According to the target set by World Health Organization for reducing undernutrition in children aged <5 years to below 5% by 2025, child wasting of 7.68% in Ethiopia should spark an immediate reaction from the government and stakeholders. Informed policy decisions, technology-based child-feeding education, and food self-sufficiency support could improve the current challenges. Additional effort is important to improve low maternal education, family planning, awareness of sex preferences, women empowerment, and maternal health services.

(JMIR Public Health Surveill 2023;9:e39744) doi:10.2196/39744

KEYWORDS

wasting; Generalized Linear Latent and Mixed Models; GLLAMM; under-five children; factors; Ethiopia

Introduction

Background

Child wasting is a key indicator used by the World Health Organization (WHO) to estimate the prevalence of childhood malnutrition. Child wasting refers to children who are too thin for their height because of recent rapid weight loss or failure to gain weight. A wasted child is one whose weight falls below 2 SDs of what is expected for the child's height. Moderately or severely wasted children are at a greater risk of death [1]. According to the United Nations Children's Fund (UNICEF), child wasting due to acute undernutrition piles additional life-costing health risks on children and usually arises as the result of maternal undernutrition, low birth weight, poor feeding and care practices, infection, food insecurity, and poverty [2].

The short-term consequences of undernutrition include frequent sickness, delayed physical and mental development, restlessness, poor feeding, and low body weight. In long run, it causes stunting, inability to learn, poor health status, and poor work performance [3-5]. Unfortunately, childhood undernutrition constraints are disproportionally large in low-income countries [6]. Despite this, evidence indicates that the international community is fragmented in regard to engaging against the problem [7]. Currently, the WHO developed a strategic plan outlining the urgent and accelerated action to end hunger and all forms of undernutrition by 2030 [8]. The status of the plan was not assessed at present. As part of this initiative, the government of Ethiopia is committed to ending wasting and stunting in children aged ≤2 years by 2030, through the effective coordination and collaboration of sectors, communities, and development partners [9]. According to the Ethiopian Growth and Transformation Plan II (GTP-II) evaluation, the rate of wasting was 9% in 2015 and 4.9% in 2020 [10]. However, the Ethiopian Health Sector Transformation Plan endorsed in the same year as GTP-II showed that wasting was 7% in 2020 and should be 5% by 2025 [11]. The discrepancy encourages further investigation.

Specific observations show that undernutrition largely occurs during the early time of childhood when higher energy is necessary for growth and development [6,12]. In other words, UNICEF, WHO, and World Bank all agree that a child who experiences wasting has a reduced immunity against various childhood infections and possibly result in death or poor mental and physical development later [13]. Nutritional inadequacy during the first 2 years is also prevalent in some Asian countries [14]; however, the severity of the problem is higher in African countries due to additional access and awareness constraints [15].

A study conducted in 35 low- and middle-income countries showed that the rate of wasting was still 12.9% in these countries [16]. Stunting, underweight, and wasting combined in sub-Saharan countries ranged from 12.14% in Benin to 0.58% in the Gambia [17]. Recently, children in many African countries still experience undernutrition [18]. Some of the prevalence are 18% in Niger, 15.5% in Burkina Faso, 12.7% in Mali, 11.1% in Comoros, 8.7% in Ethiopia, 6.2% in Namibia, 13% in Chad, and 10.5% in Sao Tomé and Príncipe [19]. It is 5.3% in Ghana [20] and 25.5% in Baka Pygmy [21]. Wasting is common in children aged <5 years and school-age children in Ethiopia. A new small-scale study in Ethiopia showed that 15.3 % of children aged 6-14 years are wasted [22], whereas another study also showed that 6.3% of children aged 6-12 years in the South Gondar Zone are wasted [23]. Another study showed that among children aged 6-59 months, the rate of wasting was 16.8% in Kersa [24], 18.2% in Dabat district [25], and 12.8 % North Wollo [26], showing that there is high wasting in various corners of the country.

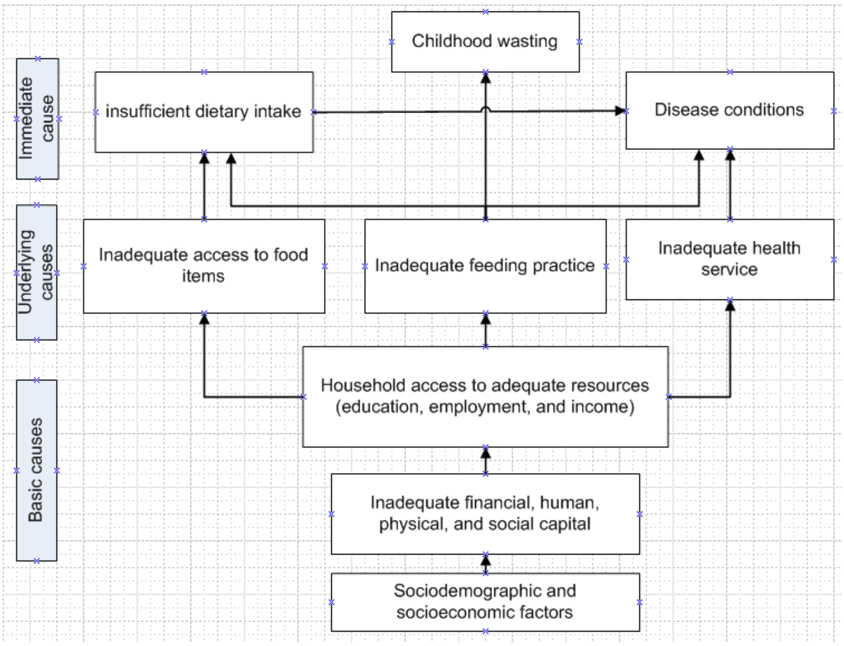
Recalling factors associated with wasting in low-income countries, one piece of evidence showed that child wasting was due to the quality of children's diet, low birth weight, maternal undernutrition, and poor complementary feeding [15]. In Ghana, wasting was associated with the wealth index of households, maternal educational status, and region of residence [27]. Another study in South Africa stated that wasting was associated with food insecurity and low awareness [28]. Household wealth, iron deficiency, recent diarrhea, and exclusive breastfeeding were the main associated factors of wasting in Somalia [29]. Generally, the age of the child, sex of the child, birth order, size of the child, education of the mother, wealth index of the family, working status of the mother, antenatal care, and postnatal care for that child are factors associated with wasting in the sub-Saharan region [17].

In Ethiopia, one study showed that wasting was significantly associated with the types of toilets and the sex of the household head [30]. In another study, wasting was associated with poor dietary diversity, late initiation of breastfeeding, absence of postnatal vitamin A supplementation, and maternal occupational status [25]. Child age, sex, dietary diversity score, husband or partner's educational level, and wealth index were associated with wasting in other studies [16,26,27,31-35]. According to the scaling-up management guideline, due to the climate liabilities and public health crises, Ethiopian children share a higher burden of wasting that needs undivided attention [36]. Despites this, relative to the WHO target for 2025, Ethiopia is struggling to reduce undernutrition of the targeted proportion [14]. Although wasting is an immediate, visible, and life-threatening form of undernutrition in children aged <5 years in peripheral and pastoralist areas in Ethiopia, studies usually do not focus on this undernutrition. Wasting is the first outcome of undernutrition, and early identification might help to control other undernutrition. It is challenging to comprehend wasting as a dichotomized variable because of its nature, which can be normal, moderate, or severe. Usually, studies dichotomize wasting as their main variable and this waste some information because of unnecessary merging. This has led to either over- or underestimation of wasting and not allowing appropriate policy decisions. Additionally, the inconsistent reports between GTP-II, Health Sector Transformation Plan, and some small scale studies mentioned above might show the existence of contextual gaps. The authors also believe that the unsuccessful previous efforts are because of the inability of identifying determinants in a country-representative sample that can show where to focus. Here, our main aim was to identify child wasting and its correlates using multilevel ordinary logistic regression to support policy direction.

Conceptual Framework

Figure 1 depicts a conceptual framework adapted from UNICEF's conceptual framework for determinants of child wasting. It summarizes the sequence in which the child wasting occurs [37].

Figure 1. A conceptual framework adapted from United Nations Children's Fund (UNICEF) determinants for childhood wasting.



Methods

Data Sources and Participants

We used the 2019 Ethiopian Mini Demographic and Health Survey (EMDHS) data for this analysis. This data were collected for the second EMDHS in 2019. In Ethiopia, there are 4 administrative levels (Ethiopia or federal, regions, zones, and woredas). The 9 regions are further divided into city administrations (Addis Ababa and Dire Dawa), agrarian regions (Tigray; Amhara; Oromia; Benishangul-Gumuz; Southern Nations, Nationalities, and People's Region; Gambela; and Harari), and pastoralists regions (Afar and Somali). We obtained data for 2016 eligible women with children aged <5 years from the 2019 EMDHS data set, downloaded from the MEASURE program web address, and extracted data elements necessary for this analysis. EMDHS 2019 used a sampling frame of all census enumeration areas (EAs) created for the 2019 Ethiopia Population and Housing Census (PHC) prepared by the Central Statistical Agency. There was a complete list of 149,093 EAs created for the 2019 PHC. An EA is a geographic area covering an average of 131 households. The sampling frame contained information about the EA location, type of residence (urban or rural), and an estimated number of residential houses focusing on key indicators for this survey. Each region was stratified into urban and rural areas, yielding 21 sampling strata. Samples of EAs were selected independently in each stratum in 2 stages. Finally, 305 EAs (93 in urban and 212 in rural areas) were selected with a probability proportional to EA size (based on the 2019 PHC frame) and with an independent selection in each sampling stratum. Either all women aged 15-49 years who were permanent residents of the selected households or visitors who slept the night before the survey were eligible for interview. The height and weight of children aged 0-59 months were collected, and women aged 15-49 years were interviewed using the Woman's Questionnaire [37].

Variables

The outcome variable of this study was wasting in children aged <5 years: if a child's weight for height is below 2 SDs from the expected weight-for-height median identified by the WHO for boys and girls [26]. In this study, we classified children as normal (when the weight-for-height *z* score is between –2 SDs to 2 SDs) [28]; moderate wasting (when the weight-for-height *z* score is between –3 SDs to –2 SDs); and severe wasting (when the weight-for-height *z* score is below –3 SDs), depending on the references from WHO 2006 guideline [38].

Independent Variables

We selected the following independent variables based on performance in previous evidence [17,23,25,26,30,32,33,39] and the availability of variables in the 2019 EMDHS data set. Age (mother and child), sex of the child, mother's educational status, head of household, wealth index, religion, residence, antenatal care, place of delivery, postnatal care, breastfeeding, anemia status of the mother, anemia status of the child, dietary diversity score, husband or partner's educational level, and birth order were independent variables.

Variable Definitions

Antenatal Care

Antenatal care visits were presented in groups of none, 1-3, or 4+. A mother may have no visits, 1-3 visits, or over 4 visits according to WHO.

Anemia Status of the Mother

The anemia state of the mother was defined as the percentage of women aged 15-49 years with mild, moderate, or severe anemia or with any anemia. It is the number of not pregnant women whose hemoglobin count is less than 12.0 grams per deciliter (g/dL) plus the number of pregnant women whose count is less than 11.0 g/dL.

Anemia Status of the Child

The anemia state of the child was defined as the percentage of children aged 6-59 months with mild, moderate, or severe anemia or with any anemia. This is when the hemoglobin count of a child is less than 11 grams per deciliter (g/dL).

Dietary Diversity Score

The dietary diversity score measured children aged <5 years who consumed a minimum of 5 of the 8 food groups (grains, roots and tubers, legumes and nuts, dairy products, meat [fish, poultry, and liver/organ meats], eggs, vitamin A–rich fruits and vegetables, other fruits and vegetables, and breast milk) in the past 24 hours.

Household Head

Households were the primary unit selected for interview in the Ethiopian Demographic and Health Survey (DHS). The definition of a household is a person or group of related or unrelated persons who live together in the same dwelling unit(s), who acknowledge one male or female adult as the head of the household, who share the same housekeeping arrangements, and who are considered a single unit.

Family Wealth Index

The questionnaire included queries concerning the household's ownership of several consumer items such as television and car; dwelling characteristics such as flooring material, type of drinking water source, toilet facilities; and other characteristics related to wealth status. Each household asset for which information was collected is assigned a weight or factor score generated through principal components analysis. The resulting asset scores were standardized about a standard normal distribution with a mean of 0 and an SD of 1. These standardized scores were then used to create the breakpoints that define wealth quintiles as lowest, second, middle, fourth, and highest [40].

Data Processing and Analysis

We used frequencies, weighted frequencies, means, SDs, and percentages or proportions to describe child wasting. Our data set contained many factors, so we checked multi-colinearity using the mean variance inflation factor (1.31), which was within an acceptable range. Before applying different models for analysis, we cleaned the data per the study criteria in Stata software (version 15.0; StataCorp) and weighted considering sampling weight, primary sampling unit, and strata before analyzing. Since wasting has normal, moderate, and severe categories, we applied multilevel ordinary logistic regression to fit the data. We conducted a bivariate analysis to identify candidate variables for multivariate analysis and selected variables with a *P* value <.20 [40,41] for the multivariable model. In current statistical analyses, studies conduct preanalysis filters to select variables for final models at a *P* value <.25 [42] and commonly at *P*<.20 [43]. We declared the final association at a *P* value of <.05. We presented the outputs of the model using coefficients and 95% CI.

Statistical Analysis—The Ordinal Logistic Regression Model

The data used in this analysis were hierarchical, which we could not analyze using binary logistic regression. Multilevel logistic regression was also not applicable since the response variable contained more than 2 categories. Therefore, we applied a special type of ordinal logistic regression (Generalized Linear Latent and Mixed Model [GLLAMM]) to account for the parallel planes and proportional odds assumptions. This model has been used to analyze clustered data [44]. We executed a mixed-effects ordinal logistic regression in a multilevel proportional odds model using GLLAMM. We used adaptive quadrature to estimate deviance and log-likelihood [45-47]. After fitting the full model, we also estimated posterior means and SDs of the latent variable. The marginal test gave us the expected response regarding the prior distribution of the latent variables so that we were able to look at the "marginal" or population-averaged effects of covariates [44,48].

Ethical Considerations and Consent to Participate

This study used secondary data from demographic and health survey data files. Initially, the authors formally requested access to the data sets from the MEASURE DHS team by completing the web-based request form [49]. Accordingly, permission to access the data and the letter of authorization were obtained from ICF International. Therefore, for this study, consent to participate is not applicable. We kept all data confidential, and no effort was made to identify households or individuals. The

Ethiopian Health Nutrition and Research Institute Review Board and the National Research Ethics Review Committee at the Ministry of Science and Technology of Ethiopia approved EMDHS 2019. The original informed consent allowed the free deidentified secondary analysis without additional consent. The authors also confirmed that all methods were carried out with relevant guidelines and regulations. The authors also ensured the study data were anonymous or deidentified for the confidentiality and privacy of the participants. According to the original consent, there was no compensation for this cross-sectional data acquisition.

Results

We analyzed the data of 2016 mothers with children aged <5 years and found an overall wasting of 7.68% (95% CI 6.56%-8.93%). Using weighted frequencies, we found that 1.86% of children were severely wasted and 5.79% were moderately wasted. Around half (48.68%) of the mothers of the children were in the age range of 25-34 years, where 26.82% of them did not have antenatal care for the current child. Over half (52.2%) of the mothers had no education, and 86.8% never had postnatal care for their current child. Half (50.93%) of the mothers had ≥6 household members. Additionally, 24.93% of mothers had a history of having 6 or more childbirths. More exceptionally, 36.6% of mothers gave birth to 2 children in the last 5 years. Many mothers (39.85%) did not use iron supplementation in the previous pregnancy. Additionally, 37.49% of mothers did not give their children solid or semisolid food yesterday, whereas 22.9% did not provide diverse food for their children (Table 1).

In the GLLAMM, variables such as giving diverse food to a child, sex of the child, total household members, sex of household head, wealth index, birth order, and maternal education were significantly associated with child wasting (all *P*<.001). Accordingly, a child fed diverse food had 4.90 times higher log odds of having normal body weight when other factors were kept constant (coefficient 4.90, 95% CI 4.90-4.98). If the sex of the household head was female, the log odds of having normal body weight was 40.40 times lower (–40.40, 95% CI –40.41 to –40.32), and when the child delivered at home, the log odds of having normal body weight was 35.51 times lower (–35.51, 95% CI –35.55 to –35.47). In other words, first and second birth-order children had higher log odds of falling into the normal body weight category (16.66, 95% CI 16.60-16.72; and 16.65, 95% CI 16.60-16.70, respectively). Families in the poor or middle-income category had 5.95 and 5.94 times lower log odds of having normal body weight (–5.95, 95% CI –6 to –5.90; and –5.94, 95% CI –6 to –5.88, respectively). The female sex of the child correlated with 12.65 times lower log odds of having normal body weight (–12.65, 95% CI –12.69 to –12.62) when all factors were constant. Children from smaller household sizes (1-3 and 4-5 members) had higher log odds of having the normal body weight (10.86, 95% CI 10.80-10.92; and 10.87, 95% CI10.91-10.92, respectively), and children whose mothers were educated (primary, secondary, and higher education) had higher log odds of having normal body weight (1.96, 95% CI 1.89-2.03; 1.99, 95% CI 1.92-2.04; and 1.99, 95% CI 1.90-2.10, respectively) (Table 2).

The equality between the nonproportional odds log-likelihood and postestimation log-likelihood shows that the assumptions hold for the model, and we applied the nonproportional model, although the proportional model almost performed similarly (Table 3).

Table 1. Sociodemographic characteristics of mothers in Ethiopia, from the 2019 Ethiopian Mini Demographic and Health Survey.

Variables	Weighted frequency, n (%)
Educational level	
No education	1052.89 (52.2)
Primary	745.1 (36.94)
Secondary	386.69 (7.85)
Higher	60.52 (3)
Number of children	
0-1	1113.56 (55.21)
2-4	738.13 (36.6)
≥5	165.26 (8.19)
Timing of the initial ANC[a]	
First trimester	541.28 (36.67)
Second trimester	778.08 (52.71)
Third trimester	156.68 (10.62)
Total number of children ever born	
<6	1514.12 (75.07)
≥6	502.81 (24.93)
Wasting	
Severe	37.41 (1.86)
Moderate	116.74 (5.79)
Normal	1862 (92.36)
Age group (years)	
15-24	548.74 (27.21)
25-34	981.94 (48.68)
35-49	486.25 (24.11)
Number of times child was fed solid food	
0 (not fed)	756.05 (37.49)
1-2	431.74 (21.41)
3-4	604.55 (29.97)
≥5	224.58 (11.13)
Wealth status	
Poor	896.88 (44.47)
Middle	348.89 (17.3)
Rich	771.15 (38.23)
Feeding diverse food	
No	461.87 (22.9)
Yes	1555.06 (77.1)
Took iron during ANC	
No	803.73 (39.85)
Yes	1213.19 (60.15)
Number of household members	
1-3	266.84 (13.23)
4-5	722.9 (35.84)

Variables	Weighted frequency, n (%)
≥6	1027.19 (50.93)
Number of ANC visits	
None	540.88 (26.82)
1-2	239 (11.86)
3-4	808.95 (40.11)
≥5	427.92 (21.22)

[a] ANC: antenatal care.

Table 2. The Generalized Linear Latent and Mixed Models of wasting in children aged <5 years in Ethiopia, from the 2019 Ethiopian Mini Demographic and Health Survey.

Variables	Coefficient (SE)	95% CI	*P* value
Diverse food fed to child			
No	—[a]	—	
Yes	4.9 (.02)	4.90 to 4.98	.001
Give milk to child			
No	–9 (.024)	–9.04 to –8.96	.001
Yes	—	—	
Sex of household head			
Male	—	—	
Female	–40.4 (.022)	–40.41 to –40.32	.001
Birth order			
First	16.66 (.03)	16.60 to 16.72	.001
Second	16.65 (.026)	16.60 to 16.70	.001
Third or above	—	—	
Place of delivery			
Home	–35.51 (.026)	–35.55 to –35.47	.001
Health facility	—	—	
Wealth index			
Poor	–5.95 (.02)	–6 to –5.90	.001
Middle	–5.94 (.30)	–6 to 5.88	.001
Rich	—	—	
Sex of the child			
Male	—	—	
Female	–12.65 (.017)	–12.69 to –12.62	.001
Number of household members			
1-3	10.86 (.03)	10.80 to 10.92	.001
4-5	10.87 (.02)	10.91 to 10.92	.001
≥6	—	—	
Maternal education			
No education	—	—	.
Primary	1.96 (.04)	1.89 to 2.03	.001
Secondary	1.99 (.03)	1.92 to 2.04	.001
Higher	1.99 (.48)	1.90 to 2.10	.001

[a] Not applicable.

Table 3. Parameter estimates and tests for wasting in children aged <5 years in Ethiopia, from the 2019 Ethiopian Mini Demographic and Health Survey.

Parameters	Proportional odds model	Nonproportional model
Variance	1.59	1.50
SE	.37	.36
Intracluster correlation coefficient	.32	.32
Log-likelihood of the model	–666	–671
Akaike information criteria of the model	570	558
Bayesian information criteria of the model	637	635
Severe stunting, σ (SE)	–4.44 (.74)	–3.84 (.51)
Moderate stunting, σ (SE)	–2.90 (.73)	–2.32 (.50)
Proportional model nested in nonproportional model	Likelihood ratio test: χ^2_9=11.04, *P*=.002	—[a]
Predicted probability or marginal test	Log-likelihood	–671

[a] Not applicable.

Discussion

Principal Findings

From our analysis, the overall wasting among children aged <5 years was 7.68% (95% CI 6.56%-8.93%). This finding is consistent with the 7.4% and 9.5% wasting rates in Afghanistan and India, respectively [27]. The finding is less than the average wasting rate (12.14%) in sub-Saharan countries [17]. The rate was 18% in Niger, 15.5% in Burkina Faso, 12.7% in Mali, and 11.1% in Comoros [19]. It is less than other study findings in Ethiopia that show wasting was 15.3% in Northwest Ethiopia [22], 16.8% in Kersa [24], 18.2% in Dabat district [25], and 12.8 % in North Wollo [26]. The clear discrepancy might show that Ethiopia needs to work very hard to reach the less than 5% wasting target set by WHO for 2025 [14]. Additionally, 52.2% of the mothers did not have an education. This is also evidenced by the 60.25% of those with no educational attainment from country-level studies in 2019 [39]. From other evidence, 59.7% of mothers have no education in the country [50], and other studies have confirmed the same [33,51,52]. Education is repeatedly the factor behind the most underachievement in the country and needs further efforts. Nearly forty percent (39.85%) of mothers did not use iron supplementation during pregnancy, which is also consistent with a finding of 34.6% in Somali [29] but sufficiently different from prenatal iron use of 45.3% in a community-based study in the country [22]. Half (50.93%) of mothers have 6 or more children, which is greater than the finding of 30.8% from another study [53]. A poor wealth index of 44.47% is consistent with the finding of 44.2% in another study [54] that shows birth control and economically empowering mothers could be the crucial target for future improvement.

There are different factors associated with all underachievements. Children who were fed diverse foods by their mothers had higher log odds of normal body weight. The finding is consistent with the results in other studies in the country [26,51]. Ethiopia has diverse cultures, which are the base for various feeding styles; however, as things change over time, natural products are no longer available in sufficient volumes and mothers might need further health education [55,56]. Female household heads were associated with the poor body weight of their children aged <5 years. In Ethiopia, men lead the household and assume all responsibilities [57]. However, when the family is separated or divorced, finding food and feeding children lies with women [37]. This might result in the potential for the undernutrition of children [39,58]. Women are poor and undereducated and have low awareness in the country, which needs attention when household responsibility falls into their hands [37,58]. Home delivery is associated with poor feeding practices (wasting) in children aged <5 years, as was also observed in other studies [31,33]. The reason might be the exposure of mothers to health professionals' counseling, which could be useful for the future. However, mothers who gave birth in health institutions might be mothers who already know what to do. Therefore, careful intervention selection is necessary. Additionally, children with first or second birth order might get enough attention during development, as opposed to those born later. This information is also available from studies conducted in various parts of the country [20,39]. A planned and economically sized family is always important. It seems that those born with many children do not give equal attention to every child, which needs attention [50]. Consequently, wasting was largely experienced by large-family children. This is consistent with another study conducted in the country [31]. The essence of family planning to reduce the risk of big families and maintain already large families might be vital. Moreover, female children had higher log odds for moderate or severe wasting, which is also consistent with another study [33]. Thus, the problem of family planning and management might have reduced most of these feeding problems. In most studies in the country, maternal education is the most prevalent problem, which was also seen in this study [33,49,51,59-61]. Generally, the current problems could have been reduced by maternal education, family planning, avoidance of sex preference, and access to pregnant women's services. Despite all the essential findings of this study, there are also some possible limitations. Disproportion of sampling, missing data, and secondary nature of the data are some limitations. The authors approached the problem through multilevel analysis, weighting, dropping records with missing data, and considering the time of data collection in the discussion.

Conclusions

Considering the target set by WHO to reduce undernutrition in children aged <5 years to below 5% by 2025, the current 7.68% wasting rate in children aged <5 years is very high and should spark immediate commitment from the government. Additionally, improving low maternal education, supporting women to use family planning, creating awareness of sex preferences, empowering women economically, and renewing commitment to improving maternal health services and follow-ups during pregnancy might be inclusive activities that need further encouragement. It was suggested that revisiting policy with current technology-based child-feeding education and advocating food self-sufficiency in the country could improve the current challenges

Acknowledgments

The authors are grateful to MEASURE Demographic and Health Survey, ICF International for allowing us to use the 2019 Ethiopian Mini Demographic and Health Survey data.

The authors received no specific funding for this work.

Data Availability

The survey data set used in this analysis is third-party data from the Demographic and Health Survey (DHS) website (www.dhsprogram.com) and permission to access the data is granted only to the registered DHS data users.

Authors' Contributions

GG wrote the method section and analyzed the Ethiopian Mini Demographic and Health Survey data, whereas GG, SH, SS, FH, KG, BTS, and KK were equally involved in the conception of the study, interpreted the results, and drafted and critically revised the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Maleta K. Undernutrition in Malawi. Malawi Med J 2006;18(4):190-207. [doi: 10.4314/mmj.v18i4.10922]
2. Early childhood development. UNICEF. 2020. URL: https://www.unicef.org/media/48886/file/UNICEF_Early_Moments_Matter_for_Every_Child-ENG.pdf [accessed 2022-04-21]
3. Martins VJB, Toledo Florêncio TMM, Grillo LP, do Carmo P Franco M, Martins PA, Clemente APG, et al. Long-lasting effects of undernutrition. Int J Environ Res Public Health 2011 Jun 26;8(6):1817-1846 [FREE Full text] [doi: 10.3390/ijerph8061817] [Medline: 21776204]
4. Consequences of undernutrition. HealthPhone. 2021. URL: http://iap.healthphone.org/consequences-of-undernutrition.html#:~:text=Home [accessed 2022-03-23]
5. Hannum E, Hu L. Chronic undernutrition, short-term hunger, and student functioning in rural northwest China. Int J Educ Dev 2017 May;54:26-38 [FREE Full text] [doi: 10.1016/j.ijedudev.2016.12.002] [Medline: 36017528]
6. Nutrition and care for children with wasting. UNICEF. 2021. URL: https://www.unicef.org/nutrition/child-wasting [accessed 2022-03-19]
7. Black MM, Walker SP, Wachs TD, Ulkuer N, Gardner JM, Grantham-McGregor S, et al. Policies to reduce undernutrition include child development. Lancet 2008 Feb 09;371(9611):454-455. [doi: 10.1016/S0140-6736(08)60215-9] [Medline: 18262026]
8. Strategic plan to reduce malnutrition in Africa adopted by WHO member states. World Health Organization. 2019 Aug 21. URL: https://www.afro.who.int/news/strategic-plan-reduce-malnutrition-africa-adopted-who-member-states [accessed 2022-04-20]
9. Seqota Declaration. Ministry of Health-Ethiopia. 2015. URL: https://www.moh.gov.et/site/Seqota_Declaration [accessed 2022-03-19]
10. Growth and Transformation Plan II (GTP II) (2015/16-2019/20). United Nations Environment Programme. 2016 Jan 01. URL: https://leap.unep.org/countries/et/national-legislation/growth-and-transformation-plan-ii-gtp-ii-201516-201920 [accessed 2022-03-19]
11. Health Sector Transformation Plan II (HSTP) (2021-2025). Ministry of Health-Ethiopia. 2021 Feb. URL: https://e-library.moh.gov.et/library/wp-content/uploads/2021/07/HSTP-II.pdf [accessed 2022-03-19]
12. Ministry of Finance-Ethiopia, UNICEF. National situation analysis of children and women in Ethiopia. UNICEF. 2019. URL: https://www.unicef.org/esa/media/5996/file/UNICEF-Ethiopia-National-Situation-Analysis-Children-Women-2019.pdf [accessed 2023-03-24]

13. Clifford C. Factsheet: child undernutrition in Africa. Africa Check. 2019 Sep 23. URL: https://africacheck.org/fact-checks/factsheets/factsheet-child-undernutrition-africa [accessed 2022-03-19]
14. Chaparro C, Oot L, Sethuraman K. Overview of the nutrition situation in four countries in South and Central Asia. FANTA III (Food and Nutrition Technical Assistance). 2014 Mar. URL: https://www.fantaproject.org/sites/default/files/download/South-and-Central-Asia-Nutrition-Overview-Mar2014.pdf [accessed 2023-01-23]
15. Torlesse H, Le MT. South Asia and child wasting – unravelling the conundrum. Field Exchange 2020 Oct(63):1-6 [FREE Full text]
16. Li Z, Kim R, Vollmer S, Subramanian SV. Factors associated with child stunting, wasting, and underweight in 35 low- and middle-income countries. JAMA Netw Open 2020 Apr 01;3(4):e203386 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3386] [Medline: 32320037]
17. Amadu I, Seidu A, Duku E, Boadu Frimpong J, Hagan Jnr JE, Aboagye RG, et al. Risk factors associated with the coexistence of stunting, underweight, and wasting in children under 5 from 31 sub-Saharan African countries. BMJ Open 2021 Dec 20;11(12):e052267 [FREE Full text] [doi: 10.1136/bmjopen-2021-052267] [Medline: 34930735]
18. Oyedele O. Undernutrition in Africa: economic implications. Int J Heal Sci Res 2017 Oct;7(10):252-255 [FREE Full text]
19. Serdula M. Diet, malnutrition in sub-Saharan Africa. Ann IFORD 1988 Dec;12(2):35-63. [Medline: 12178516]
20. Boah M, Azupogo F, Amporfro DA, Abada LA. The epidemiology of undernutrition and its determinants in children under five years in Ghana. PLoS One 2019 Jul 31;14(7):e0219665 [FREE Full text] [doi: 10.1371/journal.pone.0219665] [Medline: 31365528]
21. Funk SM, Palomo Guerra B, Bueno Zamora A, Ickowitz A, Poni NA, Abdou MA, Martín, et al. Understanding growth and malnutrition in Baka Pygmy children. Hum Ecol 2020 May 31;48(3):293-306. [doi: 10.1007/s10745-020-00161-5]
22. Mulu N, Mohammed B, Woldie H, Shitu K. Determinants of stunting and wasting in street children in Northwest Ethiopia: a community-based study. Nutrition 2022 Feb;94:111532 [FREE Full text] [doi: 10.1016/j.nut.2021.111532] [Medline: 34942423]
23. Yisak H, Tadege M, Ambaw B, Ewunetei A. Prevalence and determinants of stunting, wasting, and underweight among school-age children aged 6-12 years in South Gondar Zone, Ethiopia. Pediatric Health Med Ther 2021 Jan 18;12:23-33. [doi: 10.2147/PHMT.S287815] [Medline: 33500678]
24. Roba AA, Assefa N, Dessie Y, Tolera A, Teji K, Elena H, et al. Prevalence and determinants of concurrent wasting and stunting and other indicators of malnutrition among children 6-59 months old in Kersa, Ethiopia. Matern Child Nutr 2021 Jul 16;17(3):e13172 [FREE Full text] [doi: 10.1111/mcn.13172] [Medline: 33728748]
25. Tariku A, Bikis GA, Woldie H, Wassie MM, Worku AG. Child wasting is a severe public health problem in the predominantly rural population of Ethiopia: a community based cross-sectional study. Arch Public Health 2017 Jun 12;75(1):26 [FREE Full text] [doi: 10.1186/s13690-017-0194-8] [Medline: 28616226]
26. Anato A. Predictors of wasting among children under-five years in largely food insecure area of north Wollo, Ethiopia: a cross-sectional study. J Nutr Sci 2022 Feb 09;11:e8 [FREE Full text] [doi: 10.1017/jns.2022.8] [Medline: 35291271]
27. Harding KL, Aguayo V, Webb P. Factors associated with wasting among children under five years old in South Asia: implications for action. PLoS One 2018 Jul 03;13(7):e0198749 [FREE Full text] [doi: 10.1371/journal.pone.0198749] [Medline: 29969457]
28. Steenkamp L, Lategan R, Raubenheimer J. Moderate malnutrition in children aged five years and younger in South Africa: are wasting or stunting being treated? South African Journal of Clinical Nutrition 2016 Jul 28;29(1):27-31. [doi: 10.1080/16070658.2016.1215886]
29. Donkor WES, Mbai J, Sesay F, Ali SI, Woodruff BA, Hussein SM, et al. Risk factors of stunting and wasting in Somali pre-school age children: results from the 2019 Somalia micronutrient survey. BMC Public Health 2022 Feb 09;22(1):264 [FREE Full text] [doi: 10.1186/s12889-021-12439-4] [Medline: 35139826]
30. Kebede D, Aynalem A. Prevalence of undernutrition and potential risk factors among children below five years of age in Somali region, Ethiopia: evidence from 2016 Ethiopian demographic and health survey. BMC Nutr 2021 Oct 07;7(1):56 [FREE Full text] [doi: 10.1186/s40795-021-00460-0] [Medline: 34615555]
31. Afework E, Mengesha S, Wachamo D. Stunting and associated factors among under-five-age children in West Guji Zone, Oromia, Ethiopia. J Nutr Metab 2021 Feb 4;2021:8890725-8890728 [FREE Full text] [doi: 10.1155/2021/8890725] [Medline: 33614170]
32. Hirvonen K, Wolle A, Laillou A, Vinci V, Chitekwe S, Baye K. Child growth faltering dynamics in food insecure districts in rural Ethiopia. Matern Child Nutr 2021 Sep 15:e13262. [doi: 10.1111/mcn.13262] [Medline: 34523809]
33. Amare ZY, Ahmed ME, Mehari AB. Determinants of nutritional status among children under age 5 in Ethiopia: further analysis of the 2016 Ethiopia demographic and health survey. Global Health 2019 Nov 06;15(1):62 [FREE Full text] [doi: 10.1186/s12992-019-0505-7] [Medline: 31694661]
34. Seboka BT, Alene TD, Ngusie HS, Hailegebreal S, Yehualashet DE, Gilano G, et al. Spatial variations and determinants of acute malnutrition among under-five children in Ethiopia: evidence from 2019 Ethiopian Demographic Health Survey. Ann Glob Health 2021 Nov 22;87(1):114 [FREE Full text] [doi: 10.5334/aogh.3500] [Medline: 34900614]

35. Seboka BT, Hailegebreal S, Yehualashet DE, Gilano G, Kabthymer RH, Ewune HA, et al. Exploring spatial variations and determinants of dietary diversity among children in Ethiopia: spatial and multilevel analysis using EDHS (2011-2016). J Multidiscip Healthc 2021 Sep 21;14:2633-2650 [FREE Full text] [doi: 10.2147/JMDH.S327456] [Medline: 34584420]
36. Ministry of Health-Ethiopia. Scaling up management of wasting in Ethiopia: a case study. Concern Worldwide. 2021 Mar. URL: https://admin.concern.net/sites/default/files/documents/2021-06/05_Ethiopia_CMAM_EN.pdf [accessed 2022-04-12]
37. UNICEF, World Health Organization, Food and Agriculture Organization. Global action plan on child wasting: a framework for action to accelerate progress in preventing and managing child wasting and the achievement of the Sustainable Development Goals. World Health Organization. 2020. URL: https://cdn.who.int/media/docs/default-source/nutritionlibrary/publications/malnutrition/global-action-plan-child-wasting_58c36b82-d381-41c8-96f8-19332a8ea168.pdf?sfvrsn=79760143_3&download=true [accessed 2022-03-19]
38. Central Statistical Agency (CSA) [Ethiopia] and ICF. Ethiopia Demographic and Health Survey 2016. DHS Program. 2016. URL: https://dhsprogram.com/pubs/pdf/FR328/FR328.pdf [accessed 2022-03-19]
39. WHO Multicentre Growth Reference Study Group. WHO child growth standards based on length/height, weight and age. Acta Paediatr Suppl 2006 Apr;450:76-85. [doi: 10.1111/j.1651-2227.2006.tb02378.x] [Medline: 16817681]
40. Woldeamanuel BT. Socioeconomic, demographic, and environmental determinants of under-5 mortality in Ethiopia: evidence from Ethiopian Demographic and Health Survey, 2016. Child Development Research 2019 May 12;2019:1-15. [doi: 10.1155/2019/1073782]
41. Croft TN, Marshall AMJ, Allen CK. Guide to DHS statistics. DHS Program. 2020 May. URL: https://preview.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_DHS-7_v2.pdf [accessed 2022-03-26]
42. Bantie GM, Tamirat KS, Woya AA, Melese AA, Ayalew AF, Wubetu GA, et al. Cancers preventive practice and the determinants in Amhara regional state, Northwest Ethiopia. PLoS One 2022 May 19;17(5):e0267189 [FREE Full text] [doi: 10.1371/journal.pone.0267189] [Medline: 35587937]
43. Aynalem ZB, Yazew KG, Gebrie MH. Evidence-based practice utilization and associated factors among nurses working in Amhara Region Referral Hospitals, Ethiopia. PLoS One 2021 Mar 19;16(3):e0248834 [FREE Full text] [doi: 10.1371/journal.pone.0248834] [Medline: 33740000]
44. Heinze G, Wallisch C, Dunkler D. Variable selection - a review and recommendations for the practicing statistician. Biom J 2018 May 02;60(3):431-449 [FREE Full text] [doi: 10.1002/bimj.201700067] [Medline: 29292533]
45. Rabe-hesketh S, Skrondal A, Pickles A. GLLAMM Manual. U.C. Berkeley Division of Biostatistics Working Paper Series. 2004. URL: https://biostat.jhsph.edu/~fdominic/teaching/bio656/software/gllamm.manual.pdf [accessed 2022-04-20]
46. Rabe-Hesketh S, Skrondal A. Gllamm Companion. In: Multilevel Longitud Model Using Stata. 4th ed. Vol 1: Continuous Responses. College Station, TX: Stata Press; 2022.
47. Grilli L, Rampichini C. A review of random effects modelling using gllamm in Stata. University of Florence 2005 Oct 15:1-27 [FREE Full text]
48. Skrondal A, Rabe-Hesketh S. Generalized Linear Latent and Mixed Models with composite links and exploded likelihoods. GLLAMM. 2010. URL: http://www.gllamm.org/composite_conf.pdf [accessed 2022-04-24]
49. de Silva DBUS, Sooriyarachchi MR. Generalized Linear Multilevel Models for ordinal categorical responses: methods and application to medical data. Sri Lankan J App Stats 2012 Dec 02;12:83. [doi: 10.4038/sljastats.v12i0.4969]
50. The DHS Program. 2021. URL: https://www.dhsprogram.com [accessed 2022-02-10]
51. Takele K, Zewotir T, Ndanguza D. Understanding correlates of child stunting in Ethiopia using generalized linear mixed models. BMC Public Health 2019 May 22;19(1):626 [FREE Full text] [doi: 10.1186/s12889-019-6984-x] [Medline: 31118013]
52. Geda NR, Feng CX, Janzen B, Lepnurm R, Henry CJ, Whiting SJ. Infant and young child feeding practices in Ethiopia: analysis of socioeconomic disparities based on nationally representative data. Arch Public Health 2021 Mar 16;79(1):35 [FREE Full text] [doi: 10.1186/s13690-021-00555-x] [Medline: 33726847]
53. Getaneh Z, Melku M, Geta M, Melak T, Hunegnaw MT. Prevalence and determinants of stunting and wasting among public primary school children in Gondar town, northwest, Ethiopia. BMC Pediatr 2019 Jun 25;19(1):207 [FREE Full text] [doi: 10.1186/s12887-019-1572-x] [Medline: 31238889]
54. Comandini O, Cabras S, Marini E. Birth registration and child undernutrition in sub-Saharan Africa. Public Health Nutr 2016 Jul;19(10):1757-1767. [doi: 10.1017/S136898001500333X] [Medline: 26669828]
55. Mertens A, Benjamin-Chung J, Colford J, Hubbard A, van der Laan M, Coyle J. Child wasting and concurrent stunting in low- and middle-income countries. medRxiv. Preprint posted on online on December 31, 2022 . [doi: 10.1101/2020.06.09.20126979]
56. Bediye S, Nemi G, Makkar H. Ethiopian feed industry: current status, challenges and opportunities. Feedpedia. 2018 Feb. URL: https://www.feedipedia.org/content/ethiopian-feed-industry-current-status-challenges-and-opportunities [accessed 2022-03-19]
57. Makkar HPS, Addonizio E, Gizachew L. Charcterization of feeding systems in Ethiopia with a focus on dry areas. Feedpedia. 2018 Mar. URL: https://www.feedipedia.org/content/characterization-feeding-systems-ethiopia-focus-dry-areas [accessed 2022-04-20]

58. Households: Ethiopia. Euromonitor International. 2022 Jun. URL: https://www.euromonitor.com/households-ethiopia/report [accessed 2022-04-22]
59. The rise and fall of girls' education in Ethiopia. The Borgen Project. 2018 Jul 24. URL: https://borgenproject.org/the-rise-and-fall-of-girls-education-in-ethiopia/#:~:text=Top [accessed 2022-03-24]
60. Muche A, Melaku MS, Amsalu ET, Adane M. Using geographically weighted regression analysis to cluster under-nutrition and its predictors among under-five children in Ethiopia: evidence from demographic and health survey. PLoS One 2021 May 21;16(5):e0248156 [FREE Full text] [doi: 10.1371/journal.pone.0248156] [Medline: 34019545]
61. Akombi BJ, Agho KE, Hall JJ, Wali N, Renzaho AMN, Merom D. Stunting, wasting and underweight in sub-Saharan Africa: a systematic review. Int J Environ Res Public Health 2017 Aug 01;14(8):863 [FREE Full text] [doi: 10.3390/ijerph14080863] [Medline: 28788108]

Abbreviations

DHS: Ethiopian Demographic and Health Survey
EA: enumeration area
EMDHS: Ethiopian Mini Demographic and Health Survey
EPHC: Ethiopia Population and Housing Census
GLLAMM: Generalized Linear Latent and Mixed Models
GTP-II: Growth and Transformation Plan II
UNICEF: United Nations Children’s Fund
WHO: World Health Organization

Edited by A Mavragani; submitted 20.05.22; peer-reviewed by A AL-Asadi, S Ashraf, R Abd Jalil, K DiFilippo; comments to author 14.12.22; revised version received 17.12.22; accepted 27.12.22; published 08.02.23.

Please cite as:
Gilano G, Hailegebreal S, Sako S, Haile F, Gilano K, Seboka BT, Kashala K
Understanding Child Wasting in Ethiopia: Cross-sectional Analysis of 2019 Ethiopian Demographic and Health Survey Data Using Generalized Linear Latent and Mixed Models
JMIR Public Health Surveill 2023;9:e39744
URL: https://publichealth.jmir.org/2023/1/e39744
doi:10.2196/39744
PMID:36753309

Original Paper

Geographic, Demographic, and Socioeconomic Disparities and Factors Associated With Cancer Literacy in China: National Cross-sectional Study

Siyi He[1], MSc; He Li[1], MSc; Maomao Cao[1], MSc; Dianqin Sun[1], MSc; Fan Yang[1], MSc; Xinxin Yan[1], MSc; Shaoli Zhang[1], BSc; Changfa Xia[1], PhD; Yiwen Yu[1], MSc; Liang Zhao[1], BSc; Jufang Shi[1], PhD; Ni Li[1], PhD; Xue Qin Yu[2], PhD; Wanqing Chen[1*], PhD; Jie He[3*], MD

[1]Office of Cancer Screening, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]The Daffodil Centre - a joint venture with Cancer Council NSW, The University of Sydney, Sydney, Australia

[3]Thoracic Surgery Department, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Wanqing Chen, PhD
Office of Cancer Screening
National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College
No 17 Panjiayuan Nanli, Chaoyang
Beijing, 100021
China
Phone: 86 10 87787039
Fax: 86 10 67718227
Email: chenwq@cicams.ac.cn

Abstract

Background: Cancer literacy is associated with several health-related behaviors and outcomes. However, there is still a lack of nationwide surveys for cancer literacy in China.

Objective: This study aims to evaluate cancer literacy in China, explore disparities, and provide scientific evidence for policy makers.

Methods: A cross-sectional survey was conducted in mainland China in 2021 using the multistage probability proportional to the size sampling method. Both the reliability and validity of the questionnaire were evaluated. The awareness levels were adjusted by sampling weights and nonrepresentativeness weights to match the actual population distributions. The Rao-Scott adjusted chi-square test was applied to test geographic, demographic, and socioeconomic disparities. A generalized linear model was used to explore potential factors.

Results: A total of 80,281 participants aged 15-74 years were finally enrolled from 21 provinces, with an overall response rate of 89.32%. The national rate of cancer literacy was 70.05% (95% CI 69.52%-70.58%). The rates were highest regarding knowledge of cancer management (74.96%, 95% CI 74.36%-75.56%) but were lowest regarding basic knowledge of cancer (66.77%, 95% CI 66.22%-67.33%). Cancer literacy was highest in East China (72.65%, 95% CI 71.82%-73.49%), Central China (71.73%, 95% CI 70.65%-72.81%), and North China (70.73%, 95% CI 68.68%-72.78%), followed by Northeast (65.38%, 95% CI 64.54%-66.22%) and South China (63.21%, 95% CI 61.84%-64.58%), whereas Southwest (59.00%, 95% CI 58.11%-59.89%) and Northwest China (57.09%, 95% CI 55.79%-58.38%) showed a need for improvement. Demographic and socioeconomic disparities were also observed. Urban dwellers, the Han ethnic group, and population with higher education level or household income were associated with prior knowledge. The questionnaire showed generally good internal and external reliability and validity.

Conclusions: It remains important for China to regularly monitor levels of cancer literacy, narrow disparities, and strengthen health education for dimensions with poor performance and for individuals with limited knowledge to move closer to the goal of Healthy China 2030.

(JMIR Public Health Surveill 2023;9:e43541) doi:10.2196/43541

KEYWORDS

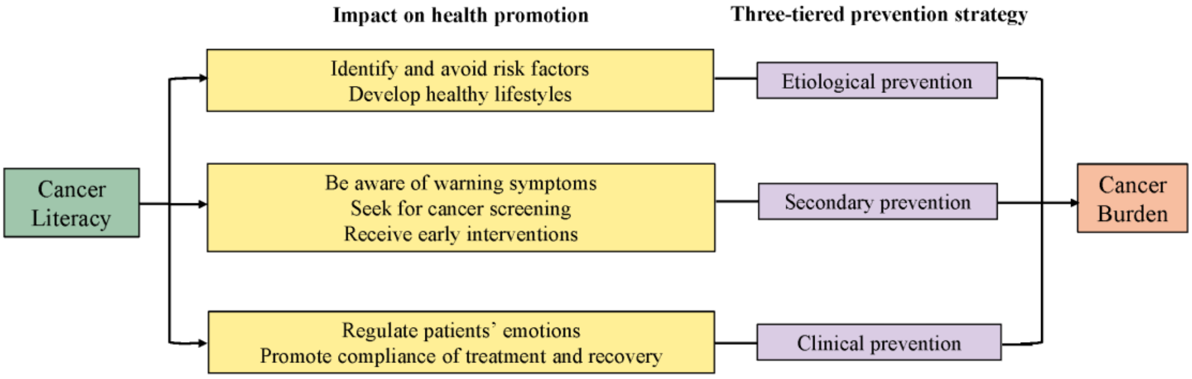
health literacy; cancer literacy; cross-sectional study; Healthy China Initiative; cancer control

Introduction

Cancer is a global issue that threatens public health, and 23.7% of new cancer cases and 30.2% deaths occurred in China in 2020 [1]. Meanwhile, regional and population disparities are evident in terms of the burden of cancer across China [2]. Cancer ranked the first and the third cause of death in urban and rural China in 2020, respectively, and more cancer-related resources, including qualified medical workers, high-level medical institutions, and health input, were allocated in East China [3]. Faced with this situation, the Chinese government has issued a series of documents [4-6] aiming to gradually promote comprehensive strategies to slow down the growth of cancer. Constant efforts have also been made to encourage health equity by increasing funds and enacting policies.

Policy makers in the United States have recognized the significance of health literacy, which may be a potential intervenable factor to reduce health disparities with decreased financial burdens on the health system [7,8]. Cancer literacy refers to all the knowledge a layperson needs to possess to understand the information and advice the health system has to offer with regard to preventing, diagnosing, and treating cancer [9]. Up to now, several studies have reported a consistent association between health literacy and some health-related behaviors or outcomes, including physical exercise, cancer screening, and mortality [8,10-12]. Therefore, the Chinese government has called for improvement in cancer literacy, with the aim of finally decreasing the burden of cancer on the basis of a 3-tiered prevention strategy (Figure 1) [13,14].

Figure 1. Cancer literacy might be associated with the burden of cancer through a 3-tiered prevention strategy.



Specifically, the Healthy China Initiative (2019-2030), the guiding principle for national health undertakings in China in the next decade, proposed phased targets for cancer literacy of 70% and 80% by 2022 and 2030, respectively [5]. The National Health Commission of China has formulated a series of the Three-year Action Plan for Cancer Prevention and Control, which repeatedly emphasized urgency and significance of improving cancer literacy [4,13]. As 1 of the 2 indicators for the assessment of government work in cancer control action in the Healthy China Initiative, the level of cancer literacy among the Chinese population needs to be evaluated to indicate the current status of progress in 2022 and to further promote the accomplishment of the next goal. However, there is still a lack of reliable measurement tools in China, and thus, national data on cancer literacy are not available. To fill this gap, the primary aim of this study was to evaluate the current level of cancer literacy in China using a well-developed questionnaire. The second aim was to assess disparities among regions and populations to provide references for health education.

Methods

Ethics Approval

This study protocol was approved by the ethics committee of the cancer hospital of the Chinese Academy of Medical Sciences (NCC-007739). We followed the Strengthening the Reporting of Observational Studies in Epidemiology reporting guideline [15].

Participants and Sample Size

A national cross-sectional survey was administered in 7 administrative divisions of mainland China using the multistage probability proportional to size sampling (PPS) method in 2021. Residents aged 15-74 years were selected, except for those registered in a collective residence. At least 2152 samples had to be included in each province to ensure their evaluations of cancer literacy (see details in Multimedia Appendix 1).

Sampling Method

The sampling process was divided into 6 stages in this national survey (Multimedia Appendix 1). Briefly, the primary sampling units of this study were provinces, and a total of 21 provinces were selected from 7 administrative units using the PPS method. The secondary sampling units (SSUs) were county-level administrative units, and the whole survey was conducted in urban and rural areas, which were classified according to the document issued by the Ministry of Civil Affairs. The PPS method was applied based on the *2020 China Statistical Yearbook* [14], and finally 3 streets or towns and 6 neighborhood or village committees were randomly selected in each SSU.

Simple random sampling was used to select households based on the electronic geographic listings of community households. One eligible family member was identified with the Kish grid method and selected to complete a questionnaire survey [16]. Participants were reached and reviewed by attempts tried per household.

Survey Tool

The tool applied in this study was designed by the National Cancer Center of China and contains a 3-level index framework constructed based on an extensive literature, public document, and expert opinions (Multimedia Appendix 2). The questionnaire consists of 13 true-or-false items, 13 single-choice items, and 11 multiple-choice items, and it covers 5 main dimensions: basic knowledge (9 scores, 24.3%), primary prevention (9 scores, 24.3%), early detection and treatment (10 scores, 27.0%), treatment (6 scores, 16.3%), and recovery of cancer (3 scores, 8.1%). Each item corresponds to the information recommended by the Chinese Health Commission after several rounds of deliberations, and all the items combined form a comprehensive assessment of individuals' cognition, attitude, skills, and behavior choices for cancer prevention and treatment [17]. In addition, basic information was also investigated, mainly age, sex, education level, ethnicity, family history of cancer, smoking status, annual income, and marital status.

Data Weighting and Calculation

For all items, 1 score was given for a complete correct answer, zero score was given for a wrong or partially correct answer, and the final score was defined as the sum of all the scores. Cancer literacy at the population level was calculated as the sum of the final scores of all respondents divided by the number of items to be answered. Meanwhile, the cancer literacy estimates were adjusted by the product of design weights and nonrepresentativeness weights, considering the inevitable inconsistencies of population distribution between the survey sample and the actual population (see details in Multimedia Appendix 1). SE values and the corresponding 95% CIs were also calculated.

Statistical Analysis

The completed questionnaires were included in the final analysis. Continuous variables are presented as the mean (SD) or median (IQR), and categorical variables are reported as counts and percentages. The internal and external reliability and validity of the questionnaire were evaluated. Item response theory (IRT) 2-parameter logistic analysis was used to evaluate the discrimination (*a* parameter) and difficulty (*b* parameter) of each item by dimension. The values of the Human Development Index (HDI), age-standardized incidence rates, and mortality rates for 7 administrative divisions were extracted to explore potential associations with cancer literacy [2,18]. The response rates were estimated based on both the unit nonresponse and the item nonresponse. R (version 4.1.0; R Foundation for Statistical Computing) and SAS (version 9.4; SAS Institute Inc) software were used to conduct data analysis. The Rao-Scott adjusted chi-square test was used to compare cancer literacy among subgroups, which is a modified chi-square test method for survey data by improving the performance in the goodness-of-fit problem and avoiding misleading results generated by clustering in the survey design [19,20]. When any assumptions of linear regression were violated (Linear relationship, Independence, Homoscedasticity, and Normality) [21], the levels of cancer literacy were grouped according to quartiles, and the multivariable logistic regression analysis was applied to explore the potential factors of cancer literacy, with adjustments of all the background variables except self-reported health status. A 2-sided $P<.05$ was considered statistically significant.

Results

Basic Characteristics

By December 31, 2021, a total of 80,281 questionnaires had been completed in mainland China, with an overall response rate of 89.32% (Table S1, Multimedia Appendix 3). The median age of all the participants was 46 (IQR 34-56) years, and the population with registered permanent residences in urban areas accounted for 51.24%, which was representative of Chinese citizens according to the Sixth Census in China. As shown in Table S2 (Multimedia Appendix 3), the survey sample overrepresented females (57.49%) and those with high levels of education (associate degree and above, 39.17%), whereas Western China (Southwest and Northwest, 14.58%) was underrepresented compared with the respective values of 21.19% from census data. Multistep adjustments of the above variables were carried out to ensure the representativeness of the main findings. Additionally, most participants were married (81.11%), never smokers (74.75%), and had no family history of cancer (68.50%). There were no logical errors in the collected questionnaires during random sampling checks, indicating good quality control of this survey.

Assessments of the Questionnaire

A panel composed of 26 experts in clinical medicine, preventive medicine, health economics, imaging diagnostics, and immunology assessed the difficulty and significance of the questionnaire framework, and the tool generally showed satisfactory content validity (Tables S3-S5, Multimedia Appendix 4). The cancer literacy rate of medical practitioners was different from that of nonmedical practitioners, that is, 80.15% (95% CI 76.08%-84.21%) and 69.83% (95% CI 69.29%-70.37%), respectively, suggesting reasonable external validity. The assessment of the internal consistency yielded an overall Cronbach coefficient of .92. The coefficient values of the 5 dimensions were 0.69, 0.79, 0.77, 0.76, and 0.59, respectively, and the former 4 dimensions were internally consistent on their own. Retest questionnaires were obtained 2 weeks later, and the correlation between the 2 rounds of testing was 0.75, indicating acceptable test-retest reliability. In addition, Table S6 (Multimedia Appendix 4) reports the IRT parameter estimates. Discrimination parameters for all the items ranged between 0.37 and 3.67, and the difficulty estimates were also in the reasonable range [22].

Overall Levels of Cancer Literacy

The national level of cancer literacy in China was 70.05% (95% CI 69.52%-70.58%). The awareness rates in terms of treatment and recovery of cancer were relatively high at 74.96% (95% CI 74.36%-75.56%) and 71.00% (95% CI 70.25%-71.76%), respectively, followed by the level of knowledge of primary prevention (70.81%, 95% CI 70.16%-71.46%). However, the corresponding levels of basic knowledge and early detection for cancer were below the target at 66.77% (95% CI 66.22%-67.33%) and 69.16% (95% CI 68.61%-69.71%), respectively. Detailed rates of cancer literacy in terms of more specific dimensions are shown in Table S7 (Multimedia Appendix 5). Levels of cancer literacy in terms of the early diagnosis of cancer, positive beliefs in cancer, and the physiological recovery were relatively low at 59.05% (95% CI 58.32%-59.79%), 63.35% (95% CI 62.72%-63.99%), and 66.13% (95% CI 65.16%-67.10%), respectively.

Disparities in Cancer Literacy Levels

There were distinct discrepancies among the geographical distributions of levels of cancer literacy (Figure 2). In general, the rates were highest in the population residing in East China (72.65%, 95% CI 71.82%-73.49%), Central China (71.73%, 95% CI 70.65%-72.81%), and North China (70.73%, 95% CI 68.68%-72.78%), followed by Northeast (65.38%, 95% CI 64.54%-66.22%) and South China (63.21%, 95% CI 61.84%-64.58%), and lowest in Southwest (59.00%, 95% CI 58.11%-59.89%) and Northwest China (57.09%, 95% CI 55.79%-58.38%). In general, the disease burden of cancer decreased with improvements in cancer literacy, and superior cancer literacy seems to be associated with a higher HDI value (Table S8, Multimedia Appendix 5).

Demographic and socioeconomic disparities in population distributions were also observed for cancer literacy within China (Table 1, Figure 3). Briefly, urban dwellers (vs rural dwellers, 71.25% vs 68.45%) and females (vs males, 70.73% vs 69.45%) had prior knowledge of cancer prevention and control. The rates also increased as the education level (*P*<.001) and annual household income (*P*<.001) increased. Although there was no significant trend in the rates of cancer literacy among age groups, individuals aged 35-54 years showed the highest overall and dimension-specific rates, and elderly individuals had better knowledge in terms of basic knowledge and early detection and treatment of cancer than those younger than 35 years. In addition, the levels of cancer literacy remained higher among people of the Han ethnic group (vs ethnic minorities, 70.43% vs 63.42%), married individuals (vs unmarried individuals, 70.38% vs 65.38%), never smokers (vs daily smokers, 70.58% vs 67.96%), and individuals with a family history of cancer (vs individuals without or with an unknown family history, 75.93% vs 69.45% or 64.34%). After multivariable adjustment, the above variables remained significantly associated with cancer literacy (Table 2).

Figure 2. Geographic disparities among provinces (A) and among administrative divisions (B) in terms of the rates of cancer literacy in China, 2021. The provinces included in each administrative division were determined following the government document of China, and all the provinces were allocated to 7 administrative divisions. AH, Anhui; BJ, Beijing; CQ, Chongqing; FJ, Fujian; GD, Guangdong; GS, Gansu; GX, Guangxi; GZ, Guizhou; HA, Henan; HB, Hubei; HE, Hebei; HI, Hainan; HK, Hong Kong; HL, Heilongjiang; HN, Hunan; JL, Jilin; JS, Jiangsu; JX, Jiangxi; LN, Liaoning; MO, Macao; NM, Inner Mongolia; NX, Ningxia; QH, Qinghai; SC, Sichuan; SD, Shandong; SH, Shanghai; SN, Shaanxi; SX, Shanxi; TJ, Tianjin; TW, Taiwan; XJ, Xinjiang; XZ, Tibet; YN, Yunnan; ZJ, Zhejiang.

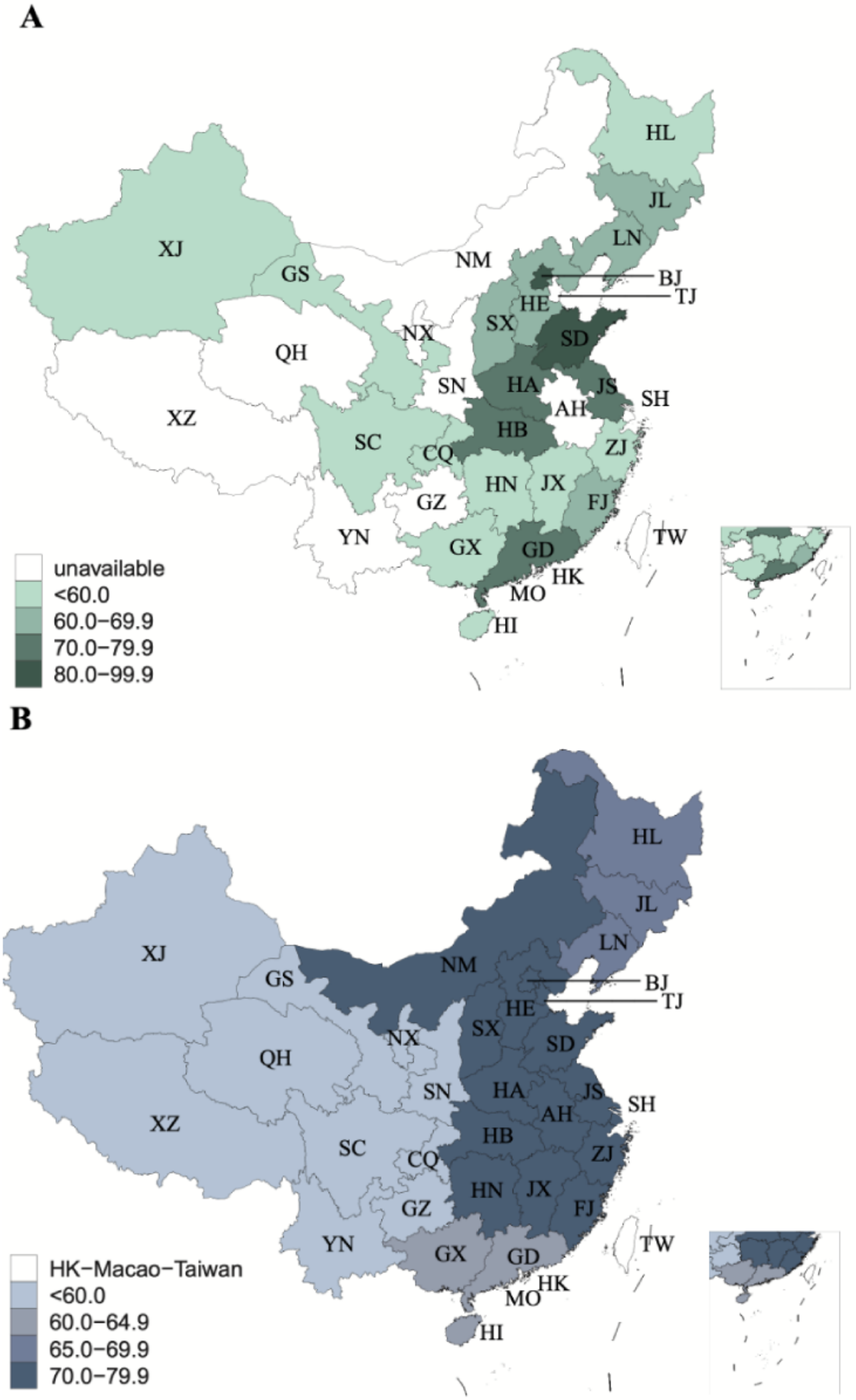


Table 1. The levels of cancer literacy by participant characteristics.

Characteristics	Sample size (n)	Cancer literacy, % (range)	*P* value
Sex			.02
Male	34,127	69.45 (68.72-70.19)	
Female	46,154	70.73 (69.96-71.51)	
Education level			<.001
Primary school and below	12,778	66.80 (65.79-67.82)	
Middle school	19,827	72.05 (71.56-72.54)	
High school	16,231	77.45 (77.02-77.87)	
Associate degree and above	31,445	82.25 (81.94-82.55)	
Type of registered permanent residence			<.001
Rural	41,138	68.45 (68.05-68.85)	
Urban	39,143	71.25 (70.36-72.14)	
Age (years)			<.001
15-24	5292	68.03 (64.52-71.54)	
25-34	15,183	67.77 (65.35-70.19)	
35-44	17,915	71.01 (69.64-72.38)	
45-54	19,179	71.41 (70.48-72.33)	
55-64	14,146	68.93 (67.94-69.91)	
65-74	8566	70.19 (68.95-71.42)	
Ethnicity			<.001
Han ethnic group	74,375	70.43 (69.88-70.98)	
Ethnic minority	5856	63.42 (60.90-65.94)	
Marital status			.20
Unmarried	10,855	65.38 (62.27-68.48)	
Married	65,112	70.38 (69.83-70.94)	
Separated	333	68.86 (63.28-74.43)	
Divorced	1954	69.12 (66.16-72.08)	
Widowed	1673	70.85 (68.40-73.30)	
Occupation			<.001
Primary industry	24,471	68.37 (67.65-69.09)	
Secondary industry	4919	70.47 (68.42-72.52)	
Tertiary industry	35,220	73.68 (72.43-74.92)	
Unemployed	14,828	69.65 (68.56-72.52)	
Smoking status			<.001
Daily smoker	12,187	67.96 (66.86-69.05)	
Light smoker	2563	70.97 (67.87-74.08)	
Former smoker	5282	71.42 (69.51-73.33)	
Never smoker	60,012	70.58 (69.95-71.21)	
Self-reported health status			.001
Good	24,216	68.37 (67.32-69.42)	
Relatively good	29,131	70.84 (69.93-71.75)	
Average	23,736	71.11 (70.22-72.00)	
Relatively poor	2631	67.08 (64.28-69.87)	

Characteristics	Sample size (n)	Cancer literacy, % (range)	*P* value
Poor	549	67.55 (61.89-73.21)	
Family history of cancer			<.001
Yes	18,129	75.93 (74.90-76.96)	
No	54,989	69.45 (68.84-70.05)	
Unknown	6801	64.34 (62.27-66.41)	
Annual household income per capita			<.001
<7500 CNY (<US $1117.00)	15,198	65.41 (63.92-66.89)	
7500-16,667 CNY (US $1117.00-2482.28)	21,313	70.51 (69.56-71.46)	
16,667-30,000 CNY (US $2482.28-4468.01)	12,817	71.88 (70.79-72.97)	
>30,000 CNY (>US $4468.01)	17,323	70.12 (68.75-71.49)	

Figure 3. Demographic and socioeconomic disparities in terms of the overall and dimension-specific rates of cancer literacy in China, 2021.

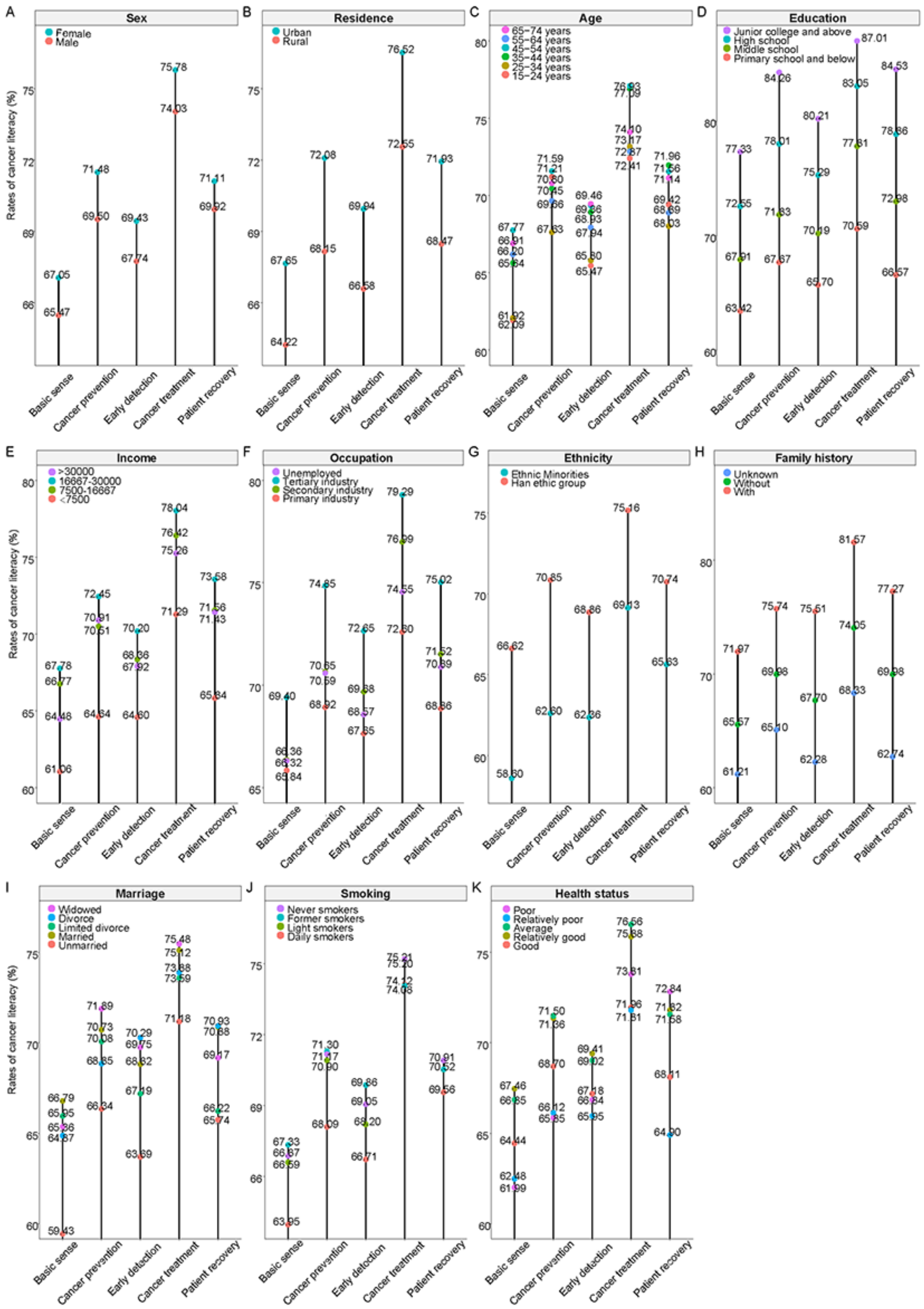


Table 2. Multivariate analysis of the associations between the included variables and the levels of cancer literacy.

Variables	OR[a] Level 2[b]	*P* value	OR[a] Level 3[b]	*P* value	OR[a] Level 4[b]	*P* value
Sex						
Male	1 (Reference)					
Female	1.06	.18	1.07	.12	1.20	<.001
Age (years)						
15-24	1 (Reference)					
25-34	1.12	.39	0.84	.19	0.79	.08
35-44	1.20	.14	0.89	.32	1.13	.35
45-54	1.27	.048	0.95	.65	1.13	.33
55-64	1.106	.62	0.76	.02	0.95	.70
65-74	1.00	.97	0.88	.27	1.25	.07
Smoking status						
Daily smoker	1 (Reference)					
Light smoker	1.08	.47	1.63	<.001	1.67	<.001
Former smoker	1.11	.15	1.88	<.001	1.85	<.001
Never smoker	1.11	.051	1.39	<.001	1.31	<.001
Type of registered permanent residence						
Urban	1 (Reference)					
Rural	1.03	.59	0.85	<.001	0.73	<.001
Education level						
Primary school and below	1 (Reference)					
Middle school	1.35	<.001	1.42	<.001	1.40	<.001
High school	1.74	<.001	2.16	<.001	2.34	<.001
Associate degree and above	1.83	.004	3.08	<.001	5.06	<.001
Ethnicity						
Han ethnic group	1 (Reference)					
Ethnic minority	0.93	.33	0.78	.004	0.80	.01
Family history of cancer						
Yes	1 (Reference)					
No	0.64	<.001	0.60	<.001	0.50	<.001
Unknown	0.47	<.001	0.48	<.001	0.38	<.001
Occupation						
Tertiary industry	1 (Reference)					
Primary industry	0.88	.03	0.84	.003	0.82	<.001
Secondary industry	0.86	.06	0.75	<.001	0.58	<.001
Unemployed	0.79	<.001	0.63	<.001	0.66	<.001
Annual household income per capita (RMB)[c]						
<7500 RMB (<US $1117.00)	1 (Reference)					
7500-16,667 RMB (US $1117.00-2482.28)	1.50	<.001	1.54	<.001	1.41	<.001
16,667-30,000 RMB (US $2482.28-4468.01)	1.58	<.001	1.48	<.001	1.36	<.001
>30,000 RMB (>US $4468.01)	1.43	<.001	1.18	.008	1.10	.12
Administrative divisions						

Variables	OR[a] Level 2[b]	*P* value	OR[a] Level 3[b]	*P* value	OR[a] Level 4[b]	*P* value
Northwest China	1 (Reference)					
Southwest China	0.73	.03	1.50	.02	2.18	<.001
South China	1.06	.67	1.76	.001	2.93	<.001
Northeast China	0.98	.89	2.12	<.001	2.62	<.001
North China	1.17	.23	3.46	<.001	5.72	<.001
Central China	2.25	<.001	3.79	<.001	5.28	<.001
East China	1.24	.07	3.27	<.001	6.98	<.001

[a] The OR represents the odds ratio that higher levels of cancer literacy (level 2/level 3/level 4, compared with the reference level [level 1]) will occur given a particular exposure and the odds of the outcome occurring in the absence of that exposure.

[b] The rates were classified into 4 groups according to quartiles before the multivariate logistic regression analysis, which were level 1 (<59.46%, as the reference), level 2 (59.46%-81.08%), level 3 (81.08%-91.89%), and level 4 (>91.89%).

[c] Classified according to quartiles.

Discussion

Principal Results

To our knowledge, this is the first national survey to measure the level of cancer literacy through completely objective items, aiming to provide references for the Chinese government as an evaluation of former health education practices and a review of current strategies. In this work, the standardized rate of cancer literacy for the Chinese population has reached the target value proposed in Healthy China 2030 for 2022 ahead of schedule. However, from the perspective of knowledge dimensions, the corresponding rate of basic knowledge of cancer remained relatively low, and the grasp of information regarding early diagnosis and treatment also needed improvement. Furthermore, there were significant disparities in the distributions among regions and populations with different characteristics. Gender, household registration type, education level, smoking status, annual household income per capita, ethnicity, occupation, and family history of cancer were associated with cancer literacy.

Comparison With Prior Work

Before the development of the questionnaire, the published evaluation tools for cancer literacy were systematically searched and summarized (see Table S9 in Multimedia Appendix 6 [9,10,23-34] for details). In brief, there are several established tools, such as the measure of awareness and beliefs about cancer [23] and the Cancer Awareness Measure [24]. These questionnaires contain single- or multidimensional assessments of cancer understanding, attitude, and skills and have been adopted in local survey practices. The Chinese government proposed that the current task was to raise awareness of systematic knowledge of cancer prevention and control. Accordingly, we collected key information on cancer that the general population was expected to know through a systematic review and the Delphi method [35] and formed an item pool. Existing questionnaires did not completely cover the abovementioned core knowledge, and the corresponding Chinese versions were not available or have not been validated in the Chinese population. Although there have been 3 related surveys administered during government-funded cancer screening programs in China [26-28], 2 questionnaires were not considered comprehensive, and the other multidimensional tool was developed to assess attitudes instead of knowledge, which were not suitable for a national survey on the evaluation of cancer literacy. On the other hand, the questionnaire in this study has a clear theoretical framework, comprehensive coverage of information, and acceptable reliability and validity performance, and is closely connected to the current situation in China.

Disparities among levels of cancer literacy in each dimension reflected the achievements and shortcomings of health education at this stage in China. Awareness of cancer management was relatively high, possibly because the Chinese government encouraged strengthening the professional advantages of health workers, especially in popularizing science during clinical activities [13]. Furthermore, low requirements were set for this dimension considering its strong professionalism, which might explain the better performance than in another survey [12]. The corresponding rate of cancer prevention was also acceptable [12,36], given that it was one of the main subjects in health education that was closest to daily life and would not raise concern. In general, the awareness level of primary prevention in this survey was significantly higher than a survey conducted in rural China during 2007-2014 (mean score 56.1, SD 19.3) and a Switzerland survey in 2009 (mean score 54.9, SD 23.9) [12,26], while still lower than a similar level reported by a UK study [24]. In contrast, the lower awareness of basic knowledge of cancer in China provided lessons for other developing countries. Very high proportions of the participants showed positive beliefs toward cancer outcomes in a survey conducted in 6 high-HDI countries, especially in Norway, and the main barrier to symptomatic presentation was the busy schedule [37]. However, in Chinese culture, people tend to delay diagnosis or treatment because of social stigma, fear, familial obligations, and embarrassment [38-40]. Thus, it is essential to emphasize the importance of timely interventions and address common misunderstandings through interactive communication that may affect individual choices [41,42]. Meanwhile, paying more attention to exerting the patient effect can help to eliminate misconceptions of cancer through experience sharing. In addition, there was room for improvement in skills, mainly the identification of warning symptoms [40,41,43]. A population-based survey in the United States reported awareness of an average of 8.43 (in 11) cancer symptoms, with the rates of recognition on each symptom mostly over 76% [44]. Another survey in developed countries also reported better awareness of warning symptoms, ranging from 70.09% (7.71 in 11, Sweden) to 79.09% (8.70 in 11, Canada) [37]. There are currently 4 cancer screening programs targeting different regions and populations in China [45]. During community mobilization and the screening process, health education carried out by community health workers and clinicians respectively might contribute to this. It is believed that cancer screening will be familiar to a wider range of individuals with the expansion of programs and synchronous health education. Moreover, appropriate examples of the application of knowledge in practice are expected to contribute to cultivating comprehensive cancer literacy.

There were evident variations in the levels of cancer literacy among regions within China. Several prospective studies have shown potential associations between cancer literacy and HDI values or burden of cancer, while related evidence remains limited and the majority was cross-sectional studies. Specifically, low health literacy has been proven to be related to poorer use of health care services and higher mortality [8,27], which might exacerbate the disparities in health outcomes, taking into account the existing differences in cancer burdens and health resources among regions. Urban residents, individuals with higher income, or those who finished compulsory education tended to have higher levels of awareness, possibly due to the more myriad ways that they obtain professional information and their stronger ability to absorb information [40,46]. Therefore, health education efforts should be focused on regions with lower life expectancy, years of schooling, and per capita gross national income, especially Western China. Significant disparities were also observed among populations. Populations with a family history of cancer or those older than 55 years showed acceptable levels of awareness because they might be more willing to acquire knowledge and receive synchronous education during health care [43]. This provided additional evidence to counter the stereotype that elderly individuals are not well adapted to information technologies [47]. Hence, increasing the enthusiasm of actively seeking information may improve the efficiency of health education. Participants occupied in tertiary industry showed higher cancer literacy because of higher coverage of health workers. In addition, higher awareness rates were found among nonsmokers, which was consistent with previous findings [48]. The gap in rates between genders was not distinct, consistent with the results of a national survey on health literacy [49].

Strength and Limitations

In comparison with previous surveys focusing on single types of cancer (mainly colorectal cancer [40] and cervical cancer [43,50]) or specific populations (such as students [51], females [43,46], and elderly people [40]), this study was the first large-scale cross-sectional survey that covered almost 70% of provinces in mainland China and provided more comprehensive information. With the aim of assessing government work, this survey informs current disparities in cancer literacy in China and essential references for improvements in health education and promotion. However, there were also several limitations in this work. First, because of the limitation of data accessibility, only the distributions of the most important variables were adjusted. As it was impossible to fully consider all possible factors and approach the census population, the reported cancer literacy might be overestimated or underestimated in this study. Second, this was a cross-sectional survey, which can indicate correlations but cannot infer causality between awareness and other variables. Third, it was not appropriate to compare the overall cancer literacy with that reported in other surveys directly, considering the differences in the applied questionnaires and the main purposes. However, as the questionnaire was developed based on published articles, the items included in each dimension were consistent with the commonly recognized knowledge, and comparisons on some specific aspects can be made with caution.

Conclusions

In summary, the weighted cancer literacy for Chinese residents was 70.05% in this cross-sectional survey. The awareness level was highest in the dimension of cancer management but was lowest for basic knowledge of cancer and the knowledge of early diagnosis and treatment. There were significant geographic, demographic, and socioeconomic disparities regarding cancer literacy, which calls for urgent efforts to carry out targeted health education for individuals with a heavier disease burden, lower income, or limited education level, especially in Southwest and Northwest China. Policy makers should conduct regular surveys to grasp the most current situations, address weaknesses in knowledge, correct misunderstandings, and move toward the goal of Healthy China 2030.

Acknowledgments

We want to thank the health workers of 21 provinces in Mainland China, who conducted the face-to-face questionnaire survey and provided detailed data. We appreciate the following collaborators for their contributions to the development and evaluation of the applied questionnaire: Min Yang, Xin Wang, Xiaofeng Bi, Nan Bi, Fengwei Tan, Wei Tang, Zhijie Wang, Zhen Huang, Yi Fang, Chunguang Guo, Sujun Han, Li Xiao, Xia Wan, Xibin Sun, Hua Dong, Shaokai Zhang, Yuqin Liu, Yubei Huang, Xiaobin Hu, Li Liu, Yunyong Liu, Yonggang Qian, Xiaojie Sun, Minjie Wang, Ning Wang, Jinyi Zhou, Qingfeng Wei, Yunfeng Xi, Zhihua Yin, Yashen Zhang, and Weifang Zheng. We especially appreciate Yinghua Li from the Chinese Health Education Center for her kind and professional guidance on this survey. This work was primarily supported with grants from Sanming Project of Medicine in Shenzhen (SZSM201911015), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Science (2019PT320027), and the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2021-I2M-1-033).

Data Availability

The study group welcomes potential collaboration to maximize the use of data. A data dictionary, a detailed study protocol, and the R programs can be reached by contacting the corresponding author of this paper. Due to Chinese legal restrictions and the current ethical approval for the study, data are not publicly available to share, but the research group can provide descriptive data in a table form. Requests can be made to WC or JH.

Authors' Contributions

WC and JH contributed equally as the corresponding authors. JH, WC, SH, JS, CX, NL, YY, and LZ contributed to the conception and design of the study. SH, HL, MC, and DS did the literature search and the construction of tables and figures. SH, FY, XY, and SZ verified the underlying data. SH, HL, MC, DS, CX, YY, LZ, JS, NL, XQY, WC, and JH contributed to data collection and quality control. SH, CX, LZ, JS, WC, and JH contributed to the administration and supervision. SH, WC, and JH drafted the paper. SH, HL, MC, DS, CX, JS, NL, XQY, WC, and JH interpreted the results. All authors contributed to data interpretation and rewriting the paper. All authors reviewed and approved the final version. All authors had full access to all the data and were responsible for the decision to submit the manuscript. The corresponding author had full access to all the data in the study and had responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
eMethods.
[DOCX File , 28 KB - publichealth_v9i1e43541_app1.docx]

Multimedia Appendix 2
Applied questionnaire.
[DOCX File , 25 KB - publichealth_v9i1e43541_app2.docx]

Multimedia Appendix 3
Response rate and baseline characteristics of enrolled participants.
[DOCX File , 20 KB - publichealth_v9i1e43541_app3.docx]

Multimedia Appendix 4
Performance of content validity evaluation and item response theory (IRT) analysis of the questionnaire.
[DOCX File , 35 KB - publichealth_v9i1e43541_app4.docx]

Multimedia Appendix 5
The levels of cancer literacy by administrative divisions.
[DOCX File , 24 KB - publichealth_v9i1e43541_app5.docx]

Multimedia Appendix 6
Summary of key information of the published evaluation tools for cancer literacy.
[DOCX File , 27 KB - publichealth_v9i1e43541_app6.docx]

References

1. Ferlay J, Ervik M, Lam F. Global Cancer Observatory: Cancer Today. International Agency for Research on Cancer. URL: https://gco.iarc.fr/today [accessed 2022-02-05]
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016;66(2):115-132 [FREE Full text] [doi: 10.3322/caac.21338] [Medline: 26808342]
3. National Health Commission of the People's Republic of China. 2021 China Health Statistics Yearbook. Beijing: Peking Union Medical College Press; 2021.
4. Notice on printing of three-year action plan for cancer prevention and control (2015-2017). National Health Commission of the People's Republic of China. 2015. URL: http://www.nhc.gov.cn/jkj/s5878/201509/656437bc5c7e4cd0afb581de85be998a.shtml [accessed 2020-09-23]
5. The healthy China action (2019-2030). National Health Commission of the People's Republic of China. 2019. URL: http://www.nhc.gov.cn/guihuaxxs/s3585u/201907/e9275fb95d5b4295be8308415d4cd1b2.shtml [accessed 2023-01-26]

6. Notice of the general office of the state council on printing China's mid- and long-term plan for the prevention and treatment of chronic diseases (2017-2025). State Council of the People's Republic of China. 2017. URL: http://www.nhc.gov.cn/bgt/gwywj2/201702/63b05a3bc7814a3686d5d37f0211f88c.shtml [accessed 2023-01-26]
7. Koh HK, Berwick DM, Clancy CM, Baur C, Brach C, Harris LM, et al. New federal policy initiatives to boost health literacy can help the nation move beyond the cycle of costly 'crisis care'. Health Aff (Millwood) 2012;31(2):434-443 [FREE Full text] [doi: 10.1377/hlthaff.2011.1169] [Medline: 22262723]
8. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
9. Diviani N, Schulz PJ. What should laypersons know about cancer? Towards an operational definition of cancer literacy. Patient Educ Couns 2011;85(3):487-492. [doi: 10.1016/j.pec.2010.08.017] [Medline: 20926225]
10. Diviani N, Schulz PJ. Association between cancer literacy and cancer-related behaviour: evidence from Ticino, Switzerland. J Public Health Res 2014;3(2):295 [FREE Full text] [doi: 10.4081/jphr.2014.295] [Medline: 25343138]
11. Koirala R, Gurung N, Dhakal S, Karki S. Role of cancer literacy in cancer screening behaviour among adults of Kaski district, Nepal. PLoS One 2021;16(7):e0254565 [FREE Full text] [doi: 10.1371/journal.pone.0254565] [Medline: 34255769]
12. Diviani N, Schulz PJ. First insights on the validity of the concept of cancer literacy: a test in a sample of Ticino (Switzerland) residents. Patient Educ Couns 2012;87(2):152-159. [doi: 10.1016/j.pec.2011.08.016] [Medline: 21945564]
13. Notice on printing health China action: cancer prevention and control implementation plan (2019-2022). National Health Commission of the People's Republic of China. 2019. URL: http://www.nhc.gov.cn/jkj/s5878/201909/2cb5dfb5d4f84f8881897e232b376b60.shtml [accessed 2023-01-26]
14. 2020 China statistical yearbook. National Bureau of Statistics of China. URL: http://www.stats.gov.cn/tjsj/ndsj/2020/indexch.htm [accessed 2023-01-26]
15. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies. STROBE. URL: https://www.equator-network.org/wp-content/uploads/2015/10/STROBE_checklist_v4_cross-sectional.pdf [accessed 2023-01-26]
16. Siddiqi K, Siddiqui F, Boeckmann M, Islam Z, Khan A, Dobbie F, et al. Attitudes of smokers towards tobacco control policies: findings from the studying tobacco users of Pakistan (STOP) survey. Tob Control 2020;31(1):112-116. [doi: 10.1136/tobaccocontrol-2020-055995] [Medline: 33318127]
17. Notice of the disease control bureau of the National Health Commission on carrying out the 2020 National Cancer Prevention and Control Publicity Week. National Health Commission of the People's Republic of China. URL: http://www.nhc.gov.cn/jkj/s5878/202003/8f30acddc9f84132a414233e8937d431.shtml [accessed 2023-01-26]
18. China national human development report special edition. United Nations Development Programme. 2019. URL: https://www.cn.undp.org/content/China/en/home/library/human_development/national-human-development-report-special-edition.html [accessed 2023-01-26]
19. Rao JNK, Scott AJ. On simple adjustments to chi-square tests with sample survey data. Ann Statist 1987;15(1):385-397. [doi: 10.1214/aos/1176350273]
20. Li F, Cui Y, Li Y, Guo L, Ke X, Liu J, et al. Prevalence of mental disorders in school children and adolescents in China: diagnostic data from detailed clinical assessments of 17,524 individuals. J Child Psychol Psychiatry 2022;63(1):34-46. [doi: 10.1111/jcpp.13445] [Medline: 34019305]
21. Yang K, Tu J, Chen T. Homoscedasticity: an overlooked critical assumption for linear regression. Gen Psychiatr 2019;32(5):e100148 [FREE Full text] [doi: 10.1136/gpsych-2019-100148] [Medline: 31673679]
22. Reise SP, Waller NG. Item response theory and clinical measurement. Annu Rev Clin Psychol 2009;5(1):27-48. [doi: 10.1146/annurev.clinpsy.032408.153553] [Medline: 18976138]
23. Simon AE, Forbes LJL, Boniface D, Warburton F, Brain KE, Dessaix A, ICBP Module 2 Working Group, ICBP Programme BoardAcademic Reference Group. An international measure of awareness and beliefs about cancer: development and testing of the ABC. BMJ Open 2012;2(6):e001758 [FREE Full text] [doi: 10.1136/bmjopen-2012-001758] [Medline: 23253874]
24. Stubbings S, Robb K, Waller J, Ramirez A, Austoker J, Macleod U, et al. Development of a measurement tool to assess public awareness of cancer. Br J Cancer 2009;101(Suppl 2):S13-S17 [FREE Full text] [doi: 10.1038/sj.bjc.6605385] [Medline: 19956157]
25. Su CC, Chen YM, Kuo BJ. Development and psychometric testing of the Cancer Knowledge Scale for Elders. J Clin Nurs 2009 Mar;18(5):700-707. [doi: 10.1111/j.1365-2702.2008.02489.x] [Medline: 19239537]
26. Li H, Zeng H, Zheng R, Zou X, Cao M, Sun D, et al. Association of cancer awareness levels with the risk of cancer in rural China: a population-based cohort study. Cancer 2020;126(20):4563-4571 [FREE Full text] [doi: 10.1002/cncr.33029] [Medline: 32780477]
27. Qiao L, Li B, Liu XX, Zhao YQ, Ma J, Li TY, et al. [Analysis of the level of the core knowledge and related factors of cancer prevention and treatment in the upper gastrointestinal cancer screening area of Sichuan Province in 2018]. Zhonghua Yu Fang Yi Xue Za Zhi 2019;53(11):1110-1114. [doi: 10.3760/cma.j.issn.0253-9624.2019.11.007] [Medline: 31683396]

28. Shi JF, Mao AY, Liu CC, Dong P, Ren JS, Wang K, et al. [Health literacy and awareness of cancer control in urban China, 2005-2017: overall design of a national multicenter survey]. Zhonghua Yu Fang Yi Xue Za Zhi 2020;54(1):108-112. [doi: 10.3760/cma.j.issn.0253-9624.2020.01.020] [Medline: 31914578]
29. Mazor KM, Roblin DW, Williams AE, Greene SM, Gaglio B, Field TS, et al. Health literacy and cancer prevention: two new instruments to assess comprehension. Patient Educ Couns 2012 Jul;88(1):54-60 [FREE Full text] [doi: 10.1016/j.pec.2011.12.009] [Medline: 22244323]
30. Morris NS, Field TS, Wagner JL, Cutrona SL, Roblin DW, Gaglio B, et al. The association between health literacy and cancer-related attitudes, behaviors, and knowledge. J Health Commun 2013;18 Suppl 1(Suppl 1):223-241 [FREE Full text] [doi: 10.1080/10810730.2013.825667] [Medline: 24093358]
31. Loo JL, Ang YK, Yim HS. Development and validation of a cancer awareness questionnaire for Malaysian undergraduate students of Chinese ethnicity. Asian Pac J Cancer Prev 2013;14(1):565-570 [FREE Full text] [doi: 10.7314/apjcp.2013.14.1.565] [Medline: 23534795]
32. Dumenci L, Matsuyama R, Riddle DL, Cartwright LA, Perera RA, Chung H, et al. Measurement of cancer health literacy and identification of patients with limited cancer health literacy. J Health Commun 2014;19 Suppl 2(0 2):205-224 [FREE Full text] [doi: 10.1080/10810730.2014.943377] [Medline: 25315594]
33. Barros A, Moreira L, Santos H, Ribeiro N, Santos-Silva F. Development of a Measurement Tool to Assess Students' Knowledge and Perceptions About Cancer (SKPaC). J Cancer Educ 2019 Jun;34(3):556-561. [doi: 10.1007/s13187-018-1341-9] [Medline: 29508229]
34. Echeverri M, Anderson D, Nápoles AM. Cancer Health Literacy Test-30-Spanish (CHLT-30-DKspa), a New Spanish-Language Version of the Cancer Health Literacy Test (CHLT-30) for Spanish-Speaking Latinos. J Health Commun 2016;21 Suppl 1(Suppl):69-78 [FREE Full text] [doi: 10.1080/10810730.2015.1131777] [Medline: 27043760]
35. de Villiers MR, de Villiers PJ, Kent AP. The Delphi technique in health sciences education research. Med Teach 2005;27(7):639-643. [doi: 10.1080/13611260500069947] [Medline: 16332558]
36. Munir R, Noureen N, Bashir M, Shoaib N, Ashraf A, Lisec J, et al. Cancer awareness measure (CAM) and cancer awareness measure mythical causes scale (CAM-MY) scores in Pakistani population. Sci Rep 2022;12(1):8887 [FREE Full text] [doi: 10.1038/s41598-022-13012-8] [Medline: 35614124]
37. Forbes LJL, Simon AE, Warburton F, Boniface D, Brain KE, Dessaix A, International Cancer Benchmarking Partnership Module 2 Working Group. Differences in cancer awareness and beliefs between Australia, Canada, Denmark, Norway, Sweden and the UK (the International Cancer Benchmarking Partnership): do they contribute to differences in cancer survival? Br J Cancer 2013;108(2):292-300 [FREE Full text] [doi: 10.1038/bjc.2012.542] [Medline: 23370208]
38. Taneja N, Chawla B, Awasthi AA, Shrivastav KD, Jaggi VK, Janardhanan R. Knowledge, attitude, and practice on cervical cancer and screening among women in India: a review. Cancer Control 2021;28:10732748211010799. [doi: 10.1177/10732748211010799] [Medline: 33926235]
39. Ng ES, Tan CH, Teo DC, Seah CY, Phua KH. Knowledge and perceptions regarding colorectal cancer screening among Chinese: a community-based survey in Singapore. Prev Med 2007;45(5):332-335. [doi: 10.1016/j.ypmed.2007.06.021] [Medline: 17707496]
40. Wong FMF. Factors associated with knowledge, attitudes, and practice towards colorectal cancer and its screening among people aged 50-75 years. Int J Environ Res Public Health 2021;18(8):4100 [FREE Full text] [doi: 10.3390/ijerph18084100] [Medline: 33924546]
41. McCutchan GM, Wood F, Edwards A, Richards R, Brain KE. Influences of cancer symptom knowledge, beliefs and barriers on cancer symptom presentation in relation to socioeconomic deprivation: a systematic review. BMC Cancer 2015;15:1000 [FREE Full text] [doi: 10.1186/s12885-015-1972-8] [Medline: 26698112]
42. Pedersen AF, Forbes L, Brain K, Hvidberg L, Wulff CN, Lagerlund M, et al. Negative cancer beliefs, recognition of cancer symptoms and anticipated time to help-seeking: an international cancer benchmarking partnership (ICBP) study. BMC Cancer 2018;18(1):363 [FREE Full text] [doi: 10.1186/s12885-018-4287-8] [Medline: 29609534]
43. Elshami M, Thalji M, Abukmail H, Al-Slaibi I, Alser M, Radaydeh A, et al. Knowledge of cervical cancer risk factors among Palestinian women: a national cross-sectional study. BMC Womens Health 2021;21(1):385 [FREE Full text] [doi: 10.1186/s12905-021-01510-2] [Medline: 34727914]
44. Sarma EA, Rendle KA, Kobrin SC. Cancer symptom awareness in the US: sociodemographic differences in a population-based survey of adults. Prev Med 2020;132:106005. [doi: 10.1016/j.ypmed.2020.106005] [Medline: 32004595]
45. Cao M, Li H, Sun D, He S, Yu Y, Li J, et al. Cancer screening in China: the current status, challenges, and suggestions. Cancer Lett 2021;506:120-127. [doi: 10.1016/j.canlet.2021.02.017] [Medline: 33684533]
46. Liu T, Li S, Ratcliffe J, Chen G. Assessing knowledge and attitudes towards cervical cancer screening among rural women in Eastern China. Int J Environ Res Public Health 2017;14(9):967 [FREE Full text] [doi: 10.3390/ijerph14090967] [Medline: 28846616]
47. Wagner LS, Wagner TH. The effect of age on the use of health and self-care information: confronting the stereotype. Gerontologist 2003;43(3):318-324. [doi: 10.1093/geront/43.3.318] [Medline: 12810895]

48. Schliemann D, Ismail R, Donnelly M, Cardwell CR, Su TT. Cancer symptom and risk factor awareness in Malaysia: findings from a nationwide cross-sectional study. BMC Public Health 2020;20(1):464 [FREE Full text] [doi: 10.1186/s12889-020-08581-0] [Medline: 32252721]
49. The national residents' health literacy level has increased to 23.15% in 2020. National Health Commission of the People's Republic of China. 2020. URL: http://www.nhc.gov.cn/xcs/s7847/202104/6cede3c9306a41eeb522f076c82b2d94.shtml [accessed 2023-01-26]
50. Ruddies F, Gizaw M, Teka B, Thies S, Wienke A, Kaufmann AM, et al. Cervical cancer screening in rural Ethiopia: a cross- sectional knowledge, attitude and practice study. BMC Cancer 2020;20(1):563 [FREE Full text] [doi: 10.1186/s12885-020-07060-4] [Medline: 32552740]
51. Osowiecka K, Yahuza S, Szwiec M, Gwara A, Kasprzycka K, Godawska M, et al. Students' knowledge about cervical cancer prevention in Poland. Medicina (Kaunas) 2021;57(10):1045 [FREE Full text] [doi: 10.3390/medicina57101045] [Medline: 34684082]

Abbreviations

HDI: Human Development Index
IRT: item response theory
PPS: probability proportional to size sampling
SSU: secondary sampling unit

Edited by Y Khader; submitted 15.10.22; peer-reviewed by E Neter, YK Lau; comments to author 02.12.22; revised version received 08.12.22; accepted 09.12.22; published 17.02.23.

Please cite as:
He S, Li H, Cao M, Sun D, Yang F, Yan X, Zhang S, Xia C, Yu Y, Zhao L, Shi J, Li N, Yu XQ, Chen W, He J
Geographic, Demographic, and Socioeconomic Disparities and Factors Associated With Cancer Literacy in China: National Cross-sectional Study
JMIR Public Health Surveill 2023;9:e43541
URL: https://publichealth.jmir.org/2023/1/e43541
doi:10.2196/43541
PMID:36800218

Original Paper

Peritoneal Dialysis Care in Mainland China: Nationwide Survey

Ping Li[1,2,3,4,5], MD; Xueying Cao[1,2,3,4,5], MD; Weicen Liu[1,2,3,4,5], PhD; Delong Zhao[1,2,3,4,5], MD; Sai Pan[1,2,3,4,5], MD; Xuefeng Sun[1,2,3,4,5], MD; Guangyan Cai[1,2,3,4,5], MD; Jianhui Zhou[1,2,3,4,5], MD; Xiangmei Chen[1,2,3,4,5], MD

[1]Department of Nephrology, First Medical Center of Chinese People's Liberation Army General Hospital, Beijing, China

[2]Nephrology Institute of the Chinese People's Liberation Army, First Medical Center of Chinese People's Liberation Army General Hospital, Beijing, China

[3]State Key Laboratory of Kidney Diseases, First Medical Center of Chinese People's Liberation Army General Hospital, Beijing, China

[4]National Clinical Research Center for Kidney Diseases, First Medical Center of Chinese People's Liberation Army General Hospital, Beijing, China

[5]Beijing Key Laboratory of Kidney Disease Research, First Medical Center of Chinese People's Liberation Army General Hospital, Beijing, China

Corresponding Author:
Xiangmei Chen, MD
Department of Nephrology
First Medical Center of Chinese People's Liberation Army General Hospital
28 Fuxing Road
Beijing, 100853
China
Phone: 86 66935462
Email: xmchen301@126.com

Abstract

Background: Peritoneal dialysis (PD) care in mainland China has been progressing in the past 10 years.

Objective: To complement information from the dialysis registry, a large-scale nationwide survey was conducted to investigate the current infrastructure and management of PD care at hospitals of different tiers.

Methods: A web-based multiple-choice questionnaire was distributed through the National Center for Nephrology Medical Quality Management and Control to PD centers of secondary and tertiary hospitals in October 2020. The 2-part survey collected the information of PD centers and the clinical management of patients on PD. A total of 788 effective surveys from 746 hospitals were voluntarily returned, and data were extracted and analyzed.

Results: The effective survey data covered 101,537 patients on PD, with 95% (96,460/101,537) in the tertiary hospitals. The median number of patients per PD center was 60 (IQR 21-152); this number was 32 (IQR 8-65) and 70 (IQR 27-192) for secondary and tertiary hospitals, respectively. There was a discrepancy in the availability of designated physical areas for different functions of PD care between the secondary and tertiary hospitals. The proportion of tertiary hospitals with PD training (P=.01), storage (P=.09), and procedure area (P<.001) was higher compared to secondary hospitals. PD catheter placement was performed in 96% (608/631) of the PD centers in tertiary hospitals, which was significantly higher compared to 86% (99/115) in secondary hospitals (P<.001). Automated PD was available in 55% (347/631) of the tertiary hospitals, which was significantly higher than that in secondary hospitals (37/115, 32%) according to the survey (P<.001). The most commonly performed PD module was continuous ambulatory peritoneal dialysis (772/788, 98%), followed by intermittent peritoneal dialysis (543/788, 69%). The overall reported nocturnal intermittent peritoneal dialysis was 31% (244/788); it was 28% (220/788) for continuous cycling peritoneal dialysis and 15% (118/788) for tidal peritoneal dialysis. Comparisons between the secondary and tertiary hospitals revealed no significant differences in prophylactic antibiotic use for PD catheter placement and therapeutic use for peritonitis. The first peritoneal equilibrium test was conducted in 58% (454/788) of patients at 4-6 weeks after initiation of PD, and 91% (718/788) reported at least one peritoneal equilibrium test per year. Overall, 79% (570/722) and 65% (469/722) of PD centers performed assessment for dialysis adequacy and residual kidney function, respectively; and 87% (685/788) of patients on PD were followed every 1 to 3 months for laboratory and auxiliary examinations.

Conclusions: This national survey reflects the current status and disparities of PD center management in mainland China. The study results suggest that the PD care needs to be more conveniently accessible in secondary hospitals, and quality management and staff training in secondary hospitals are still in high demand.

(JMIR Public Health Surveill 2023;9:e39568) doi:10.2196/39568

KEYWORDS

renal replacement therapy; peritoneal dialysis; medical quality; tertiary hospital; secondary hospital; China

Introduction

It was not until the late 1990s that peritoneal dialysis (PD) started to grow rapidly in academic hospitals in large cities of mainland China [1]. Although those top PD centers gained significantly improved clinical outcomes in the late 2000s, the accumulated experiences were shared through published materials of standard operating procedure (SOP) for PD management [2] and through organized training materials [3] and courses from 2011 to 2012.

In China, the pyramid of the health care system is constructed by 3 tiers of hospitals and community convenient clinics. There were about 3000 tertiary hospitals and more than 10,000 secondary hospitals in China according to the 2020 China Health Care Development Statistics [4]. The tertiary hospitals are the top-level hospitals usually located in regional capitals, serving as teaching hospitals as well as medical care providers to manage complex and incurable diseases. The primary hospitals and clinics are often the service providers for long-term disease follow-up, general prescription renewal, and preventive care. Health care workers at tertiary hospitals and in large cities tended to have better exposure to up-to-date medical knowledge, clinical management concept, and more training opportunities for advanced procedures. More patients in end-stage kidney disease (ESKD) choose PD as their renal replacement therapy due to continuing increase in the prevalence and incidence of ESKD [5,6], average educational level of the patients, availability and accessibility of PD resources, and better policy support. In the recent 10 years, PD service is available not only in leading tertiary hospitals but also in secondary hospitals, which has raised a concern about the sufficient governance of PD care quality.

In mainland China, there has been a nationwide registry for dialysis—Chinese National Renal Data System (CNRDS) [7-9]—along with some independent provincial-level databases [10-12] that mostly record clinical and socioeconomic outcomes. Detailed facility infrastructure data, especially for PD centers, are not available in these databases. Although some data related to the concept and behavior of PD care providers (such as the common follow-up parameters, check-up schedule, and the use of medicine for comorbidities and peritonitis) are included in the dialysis registries, these data are largely not spontaneously reported. In addition, some of the information about the reasons that hindered PD as the choice for renal replacement therapy are not collected by the registries. Moreover, previously reported results from large-scale, real-world studies or analysis of registry data did not include information from the recent 2 years [12-14]. Thus, there is a strong need to have another way to collect real-world PD care data and elicit relevant information for better PD care in the future.

To complement a comprehensive picture of current PD practice in mainland China, a national survey was designed focusing on hospital infrastructure for PD as well as conceptual modalities and goals for management of patients on PD. By October 2020, the time when the questionnaires were distributed, data from CNRDS showed that there were about 1000 hospitals offering PD service, with only few in primary hospitals.

Methods

Participants

From September to October 2020, according to the information in the CNRDS, renal divisions from tertiary and secondary hospitals known to have or likely to have PD programs were invited to the web-based survey. Through the National Center for Health Care Quality Management in Nephrology Diseases, the invitation QR code and questionnaire instruction pack was first sent to each provincial Center for Health Care Quality Management in Nephrology Diseases and then distributed to physicians whose practice field was PD. Survey results were voluntarily submitted by participating centers and practitioners.

Survey Design

A multiple-choice type of questionnaire was designed and tested in a pilot group before being distributed to the target respondents. Detailed questions included in the survey are listed in Table 1. The questionnaire consisted of 2 parts. The first part collected the information of the hospital or PD center, including the location and health care level of the facility (secondary or tertiary), availability of nephrology and dialysis care resources, regulatory rules on PD, number of current patients on PD, and the types of PD performed.

The second part of the survey addressed practice concepts or the current situation of clinical management for patients on PD. The questions covered PD catheter placement and care; use of peritoneal equilibrium test (PET); regular follow-up schedule and items; as well as management of common complications, including peritonitis, anemia, and hypertension.

Table 1. Survey items listed in the questionnaire for peritoneal dialysis (PD).

Category	Item
Facility-related survey items	• Name and location of the hospital • Administrative ranking of the hospital (secondary hospital or tertiary hospital) • Number of current patients on PD • Types of PD performed in your center • Diseases treated with PD in your center • Common situations hindering you from choosing PD • Functional areas designated for PD • Does your center register patients on PD in the CNRDS[a] [15]? • Is PD catheter placement performed in your center? • Do you perform PD catheter implantation surgery?
Concept-related survey items	• Antibiotic medicine routinely given right after PD catheter placement • Maximal interval for change of external connecting tube • Do you prescribe or perform peritoneal equilibrium test (PET)? • Time to administer the first PET • Frequency of repeat PETs • Common follow-up interval for patients on PD • Laboratory tests and examinations administered for PD follow-up • Antibiotics for treatment of peritonitis • Common hemoglobin level of patients on PD in your center • Use of antihypertensive drugs, number of drugs, and estimated percentage of patients

[a] CNRD: Chinese National Renal Data System.

Statistical Analysis

Descriptive statistical analysis was performed to evaluate PD items in the survey, and median (IQR) or proportion (frequency) were reported for continuous or categorical variables, as appropriate. Comparisons by health care level of the facility were conducted using Pearson chi-square and Wilcoxon tests for categorical and continuous variables, respectively, along with figures. R (version 4.03; R Core Team) and GraphPad Prism (version 8.0; GraphPad Software Inc) were used for data analysis and figures.

Ethical Considerations

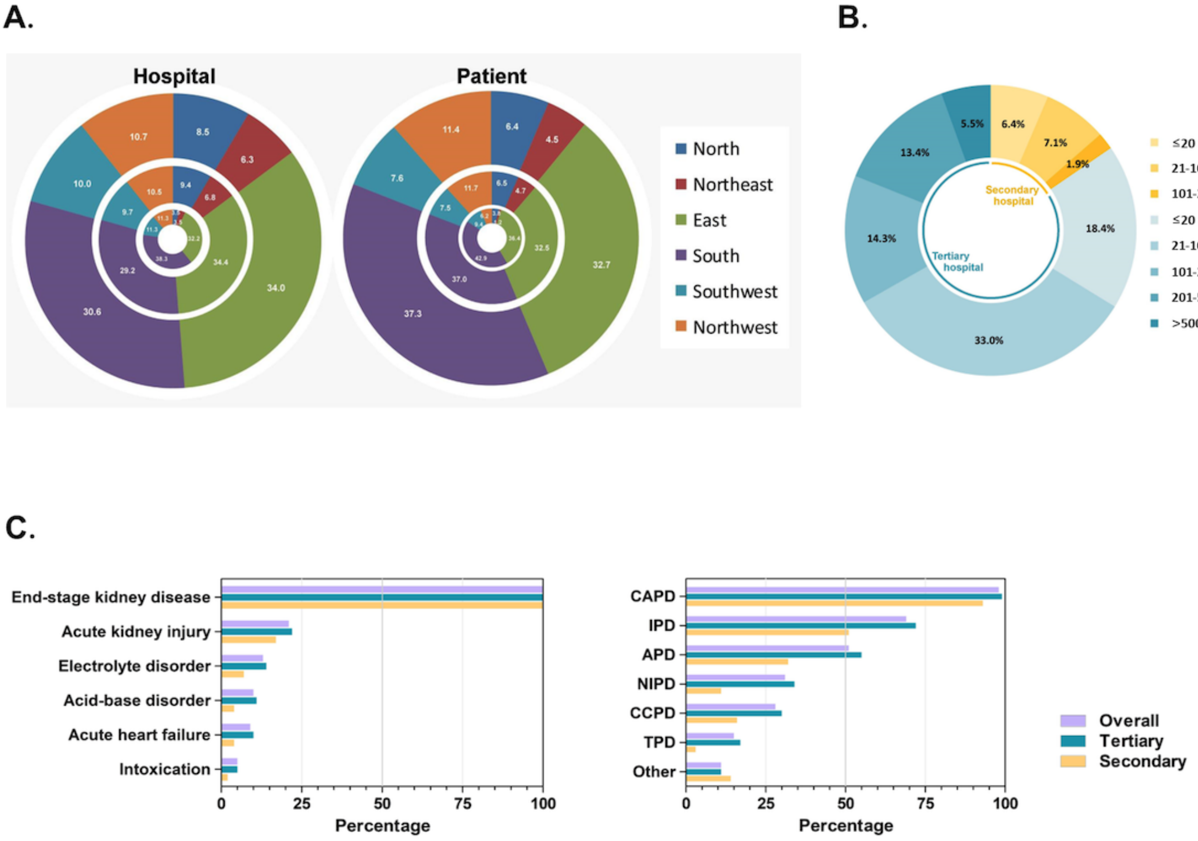
This work was carried out under the research program of National Health Care Quality Management in Nephrology Diseases of Chinese People's Liberation Army General Hospital. The research protocol was reviewed and exempted by the institutional review board of Chinese People's Liberation Army General Hospital. There was no collection of individual and identifiable patient data for the survey. The original informed consent from each participating hospital allows for secondary data analysis without additional consent. All questionnaires were voluntarily responded without compensation.

Results

Participating PD Centers and Patients

A total of 788 surveys from 746 hospitals located in all 31 provincial-level areas (ie, province, autonomous prefecture, or municipality directly under the central government) in mainland China were returned. These 746 centers covered 101,537 patients on PD, with 95% (96460/101,537) of them in the tertiary hospitals. Collected data were further grouped into 6 socioeconomic administrative areas (ie, North, Northeast, East, South, Southwest, and Northwest China); proportions of the hospitals and patients from each area are shown in Figure 1A. Overall, the median patient number per PD center was 60 (IQR 21-152), that is, 70 (IQR 26-192) for tertiary hospitals and 32 (IQR 8-65) for secondary hospitals (Figure 1B). There was no significant regional difference in the distribution of PD centers in terms of patient size. Among the 746 centers, 10 centers did not treat patients on PD, 6 (60%) of which were tertiary hospitals; however, among the 722 centers providing PD care, 611 (85%) were tertiary hospitals. All these indicate that tertiary hospitals are the primary provider of PD care in mainland China (*P*=.03).

Figure 1. Characteristics of hospitals and patients on peritoneal dialysis (PD) covered by the survey. (A) Proportions (%) of participated hospitals and patients in different regions of mainland China; inner ring: secondary hospitals; middle ring: tertiary hospitals; and outer ring: all hospitals. (B) Proportion of participated hospitals with different number of patients on peritoneal dialysis (PD). (C) Proportions of diseases treated with PD and the types of PD performed. APD: automated PD; CAPD: continuous ambulatory PD; CCPD: continuous cycling PD; IPD: intermittent PD; NIPD: nocturnal IPD; TPD: tidal PD.



Clinical Conditions Impacting PD Decision

Data from the 788 submitted surveys showed that besides chronic and acute renal failure, secondary hospitals tended to apply less PD compared to tertiary hospitals for acute congestive heart failure (*P*=.05), electrolyte (*P*=.03), acid-base disorders (*P*=.02), and drug intoxications (*P*=.10; Figure 1C). Between the secondary and tertiary hospitals, there was no significant difference in factors hindering PD decision when coming across clinical situations, such as extensive adhesion of the peritoneum (776/788, 98%), unrepaired hernia (721/788, 91%), and severe skin disease (671/788, 85%).

Treatment Module of PD

The most common PD type was continuous ambulatory peritoneal dialysis, which was reported by 98% (772/788) of the respondents. It was followed by intermittent peritoneal dialysis 69% (543/788). The overall reported nocturnal intermittent peritoneal dialysis, continuous cycling peritoneal dialysis, and tidal peritoneal dialysis were 31% (244/788), 28% (220/788), and 15% (118/788), respectively. Automated PD was available in 32% (37/115) of the secondary facilities and 55% (347/631) of the tertiary facilities by the end of October 2020 (Figure 1C). A combination of hemodialysis and PD was reported by 11 centers; daytime ambulatory PD was reported by 23 centers, and a combination of manual continuous ambulatory peritoneal dialysis and automated PD was reported by 1 center.

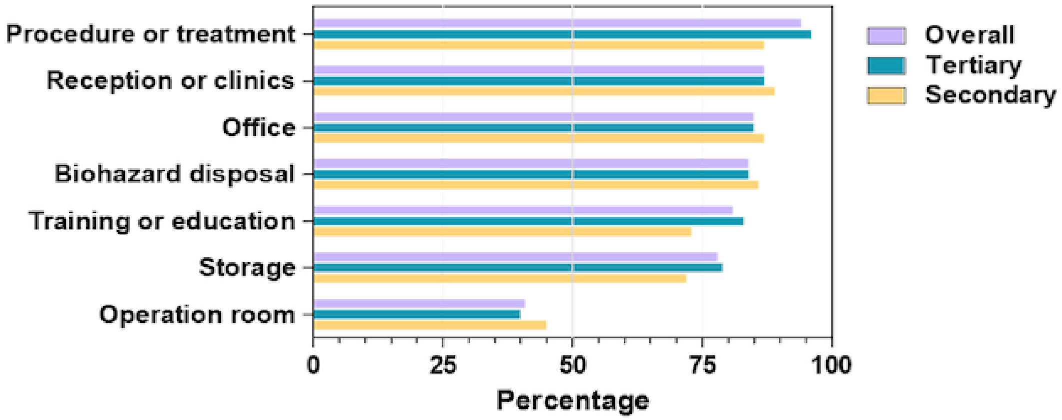
Management of PD Centers

Among the secondary centers, 17% (20/115) reported that there were no established management rules for the PD center, while this number was 7% (44/631) in tertiary hospitals (*P*=.002).

About 19% (22/115) of the secondary hospitals did not have specialized PD medical records for each patient, and this was 11% (69/631) in the tertiary hospitals (*P*=.01).

More than 80% of the centers reported separate functional areas designated for PD care, including areas for reception, procedure or treatment, biohazard disposal, and staff office (Figure 2), with no significant statistical difference between the secondary and tertiary hospitals. Compared to the secondary hospitals, tertiary centers were more likely to have designated areas for training or patient education (*P*=.01), storage (*P*=.09), and procedures (*P*<.001). Only 40% (289/722) of the PD centers had operation rooms exclusive for dialysis access, and there was no discernible difference between the secondary and tertiary centers.

Figure 2. Proportion of hospitals having functional areas designated for peritoneal dialysis use.



PD Catheter Implantation and Care

About 96% (608/631) of the participating tertiary hospitals performed implantation surgery for PD catheters, which involved 78% (515/662) of the survey participants; this was performed in 86% (99/115) of the secondary centers (vs tertiary centers; *P*<.001) and by 67% (85/126) of physicians in the secondary hospitals (vs tertiary hospitals; *P*=.01). Assessment and patient education were conducted before the implantation surgery in 100% of the 631 tertiary hospitals and 94% (108/115) in the secondary centers. Among the hospitals performing PD catheter implantation, 100% (746/746) provided preoperative assessment and patient education. Cephalosporin of the first or second generation was commonly used right after the surgery, accounting for 48.7% (384/788) of the antibiotics used. Other antibiotics included third or fourth-generation cephalosporin (72/788, 9.1%), quinolone (10/788, 1.3%), and aminoglycosides (4/788, 0.5%). About 40.4% (318/788) of the participants reported using antibiotics other than the ones mentioned above or not using them. There was no significant difference in postoperative antibiotic use between the tertiary and secondary hospitals.

There was a slight but significant difference in the frequency of replacing the external tube that connects to the PD catheter (*P*=.04). Among the secondary hospitals, 82% (94/115) changed the connecting tube every 6 months, and 15% (17/115) did so every 3 months; this was 90% (568/631) for every 6 months and 7% (44/631) for every 3 months in the tertiary hospitals. A few centers reported replacement frequency at every 9 months (about 2%) or every 12 months (<1%) from both secondary and tertiary hospitals.

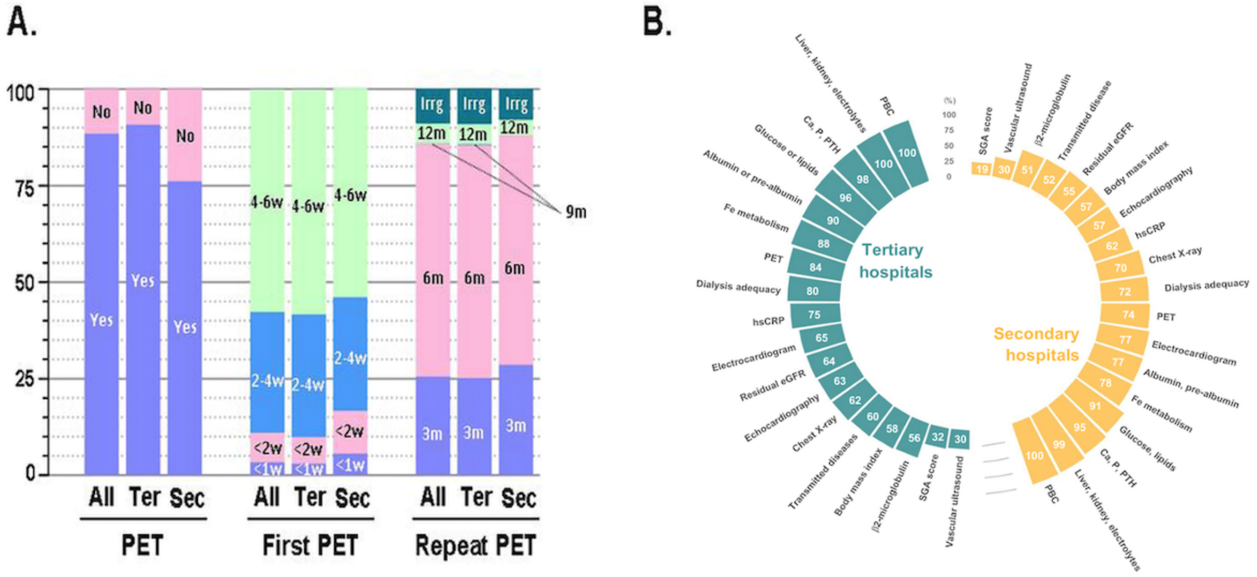
Management of PD Patient Follow-Up

The use of PET to assess the function of the peritoneum was reported in 91% (574/631) of the tertiary hospitals and 76% (87/115) of the secondary hospitals (*P*<.001; Figure 3A). Around 58% (454/788)of the first PETs were performed within 4-6 weeks after initiation of PD, and 31% (247/788) conducted the first PET between 2-4 weeks overall in tertiary and secondary hospitals. PET was repeated every 6 months in 60% (475/788) of the PD centers, and another 25%-29% (166/662 for tertiary hospitals and 36/126 for secondary hospitals) of the centers had repeat PET every 3 months (Figure 3A).

For PD long-term care, the most common follow-up interval was every month (385/788, 49%), followed by every 3 months (226/788, 29%), irregular (85/788, 11%), every 2 months (74/788, 9%), and every 6 months (18/788, 2%).

For the laboratory tests and imaging examinations, peripheral blood count was considered for each patient (788/788, 100%); blood biochemistry panel (ie, parameters for liver enzymes, kidney function, and electrolytes; 788/788, 100%); serum calcium, phosphate, intact parathyroid hormone (772/788, 98%); and glucose and lipid profiles (748/788, 95%) were routinely followed. Parameters for iron metabolism (*P*=.002), nutrition status (ie, albumin and pre-albumin; *P*<.001), and inflammation biomarker high-sensitive C-reactive protein (*P*=.003) were more likely to be followed up in the tertiary hospitals compared to the secondary hospitals (Figure 3B). Transmitted disease panel, BMI, serum β_2-macroglobulin, chest x-ray, electrocardiogram, and cardiovascular ultrasound were also checked in both tertiary and secondary centers, with checking rates ranging from 30% (236/788) to 67% (527/788). PET (*P*=.01), PD adequacy test (*P*=.06), and assessment for residual kidney function (*P*=.04) were administered more in tertiary hospitals than in secondary hospitals.

Figure 3. Follow-up Items for peritoneal dialysis. (A) Availability of peritoneal equilibrium test (PET) and the time it was performed (ter: tertiary hospital; sec: secondary hospital; w: week). (B) The follow-up items regularly applied in secondary and tertiary hospitals for patients on peritoneal dialysis. Ca: Calcium; eGFR: estimated glomerular filtration rate; Fe: Ferrum; hsCRP: high-sensitive C-reactive protein; P: Phosphorus; PBC: peripheral blood cell; PTH: parathyroid hormone; SGA: subjective global assessment.



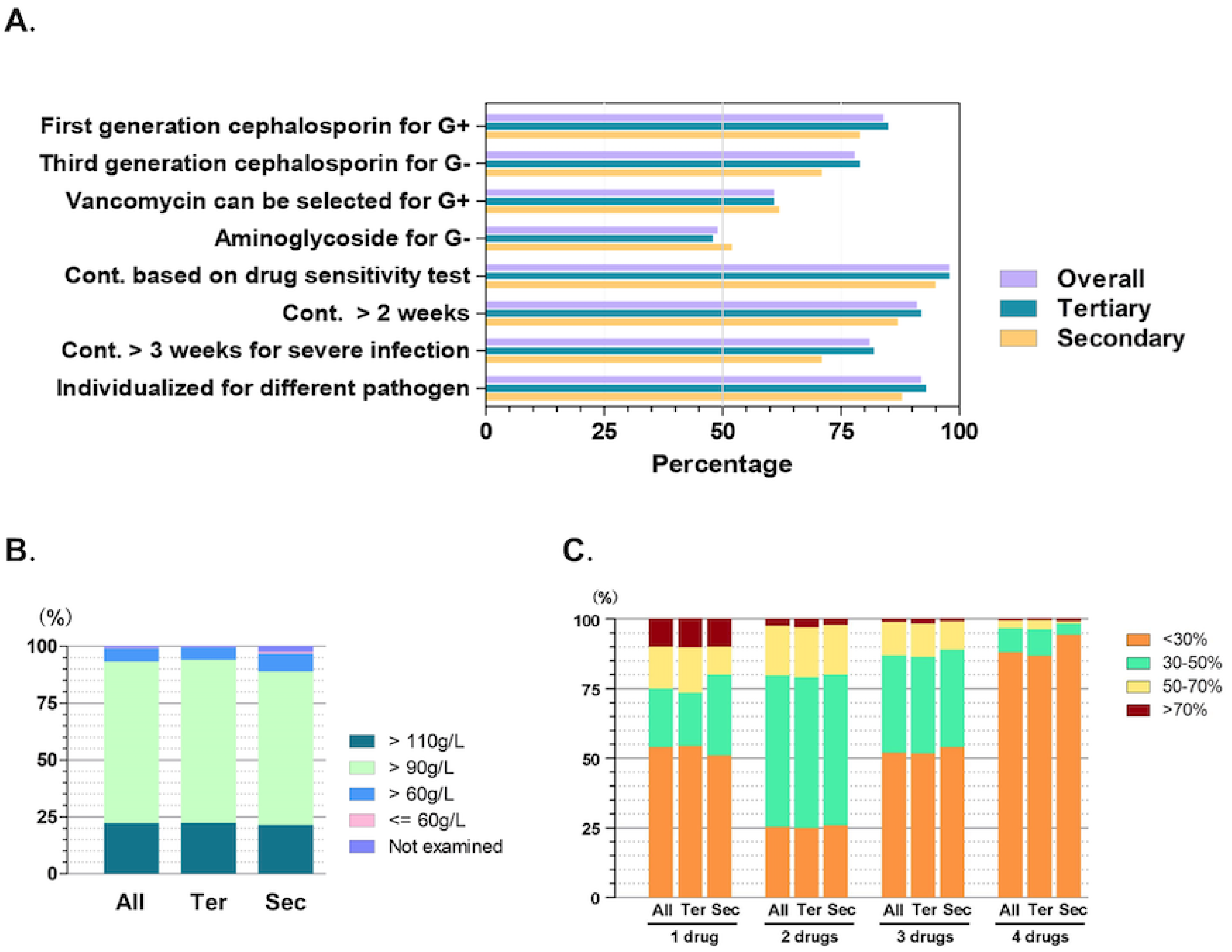
PD Complication Management

For infectious peritonitis, about 84% (659/788) of the participants chose first-generation cephalosporin for gram-positive bacteria, and 78% (612/788) chose third-generation cephalosporin for gram-negative pathogens (Figure 4A). Vancomycin was also commonly used for the treatment of gram-positive peritonitis, and aminoglycoside was used for gram-negative peritonitis. For severe infection, 29% (33/115) of the secondary hospitals did not continue antibiotic therapy longer than 3 weeks, compared to 18% (114/631) of the tertiary centers (*P*=.005). Awareness of applying drug sensitivity tests and individualized use of antibiotics was more than 92% (725/788) among the participants.

In the secondary hospitals, 89% (112/126) of the patients on PD had hemoglobin levels greater than 90 g/L, compared to 94% (623/662) of them in the tertiary centers (Figure 4B).

The majority of patients on PD took 1-3 antihypertensive drugs, with a total of 97% (764/788; 95% CI 95%-98%) of the respondents prescribing 1-3 drugs for more than 30% of the patients (Figure 4C). A combination of 2-3 antihypertensive medicines was more common than using a single drug (677/788, 86%, 95% CI 83%-88% vs 362/788, 46%, 95%CI 43%-50%; *P*<.001) or ≥4 drugs (vs 95/788, 12%, 95%CI 10%-14%; *P*<.001).

Figure 4. Complication management. (A) Antibiotic choice for peritonitis (G+: gram-positive; G–: gram-negative; Cont: continued). (B) Hemoglobin level in patients on peritoneal dialysis. (C) Number of antihypertensive drug used (ter: tertiary hospital; sec: secondary hospital).



Discussion

To our knowledge, this survey was the first nationwide study focusing on the current situation of PD management and PD practitioners' concepts and behavior in mainland China [1,12,14]. It covered about 80% of PD centers and nearly 90% of the estimated number of current patients on PD from all provincial-level areas, according to previously published registry data [9,13], as well as about 70% of PD centers and 80% of patients on PD, according to unpublished CNRDS 2021 data. The respondents who filled the survey were mostly physicians in charge of the PD centers or some chief nurses who were well knowledgeable about the PD center operations. The big number of the returned surveys and the personnel responding to the questionnaires suggested a faithful reflection of the real situations that may not be accurately captured or reported by the registries. The participating numbers of tertiary and secondary hospitals were about 20% and 1% of the total number of tertiary and secondary hospitals in mainland China, respectively, according to the 2020 China Health Care Development Statistics [4]. In this study, we reported that PD care was provided mostly by tertiary hospitals in mainland China, and about 5% of patients were managed by secondary hospitals. We found that tertiary hospitals had a more complete infrastructure of PD centers, performed more operations for PD catheter placement, treated more PD indication–related acute diseases, and conducted more assessments for peritoneal membrane characteristics and dialysis adequacy than the secondary hospitals. This study also showed that current concepts of the goals for long-term PD complication management were adequately adopted by health care providers (HCPs) in both tertiary and secondary hospitals, with a relatively similar compliance rate of the follow-up items.

Our study showed a couple of differences in PD care management between the secondary and tertiary hospitals. In the secondary hospitals, a lack of designated infrastructure for PD patient education was common, and the PD catheter placement procedure was significantly less performed. This discrepancy could be explained by the lack of clinical experience and human power as well as low financial incentive for PD catheter implantation. To overcome these barriers, it is necessary to implement a well-designed and feasible PD program for catheter implantation. Sufficient surgical training, a dedicated team and time, and protected reimbursement from payers would help to achieve this goal. Complex and more advanced PD treatment modules, such as cycler-facilitated nocturnal intermittent peritoneal dialysis or tidal peritoneal dialysis as well as a combination of hemodialysis and PD, were significantly less reported by secondary hospitals. We speculate that, along with the development of digital tools for PD patient management, more available web-based patient education resources, more use of PD cycler, and the shifting of reimbursement policy toward lower tiers of cities and health care facilities, more patients with incident dialysis will choose PD and more patients on PD will migrate to secondary and even primary health care centers for chronic PD management. In that regard, the proportion of patients on PD in rural and lower-tier facilities is expected to increase rapidly in the near future, which advocates strengthened implementation of standardized PD quality management and control at high priority.

Adequate patient follow-up scheme and good execution are necessities for quality PD care [16]. In China, a usual PD outpatient follow-up routine includes HCPs' tracking of the daily records of PD exchanges, measurement of blood pressure, prescriptions for examinations, and medication adjustment [17]. Our survey results showed that more than 85% of patients on PD were regularly followed every 1 to 3 months, and almost every patient was checked for anemia (low hemoglobin) and mineral bone markers (ie, calcium, phosphorus, and parathyroid hormone), which was similarly complied according to multiple international guidelines and recommendations [18-20]. These proportions were higher than that reported in China Kidney Disease Network 2016 Annual Data Report, which was around 90% [8]. The difference may be explained as follows: our survey covered wider regions and more patients on PD, whereas the China Kidney Disease Network Report used data sources from claim databases and included primary hospitals. Nonetheless, thanks to the continuous effort of improving health care accessibility and affordability from the government at all administrative levels, results from these 2 studies indicate that patients with ESKD are getting an adequate standard of care in mainland China.

In our survey, there were about 10% (10,150/101,537) of patients on PD whose hemoglobin level was below 90 g/L, the level at which rescue therapy for anemia should be considered [21]. These data were also similar to the data recently reported by the US Renal Data System [16]. Multiple guidelines recommend treating chronic kidney disease anemia to the hemoglobin level of 110-120 g/L [21,22] or around 110 g/L [18,19]. In our study, about 22% (175/788) of the respondents reported hemoglobin levels of their patients on PD as above 110 g/L, with no difference between the secondary and tertiary hospitals, indicating that further clinical effort may be needed to reduce the anemia burden in patients on PD. Multidrug therapy is usually required and recommended for the control of blood pressure in patients with chronic kidney disease worldwide [23-25]. In PD, antihypertensive treatment is dependent on not only the drug administration but also fluid volume control. Our study showed that 86% (87,321/101,537) of patients on PD took 2-3 antihypertensive medicines. Future studies on the relationships among blood pressure level; types of antihypertensive drugs; volume status; and other factors, such as retaining of residual kidney function, longevity of peritoneal membrane function, anemia treatment regimen, cardiovascular outcome, patient compliance, and health economics would provide guidance to improve PD outcomes as well as goal-directed value-based PD care [17,26].

Assessment for peritoneal membrane characteristics is helpful for appropriate PD prescription. International PD guidelines consistently recommended performing the first PET at 4-6 weeks after PD initiation and having it reevaluated every 6-12 months [27-30]. In our study, PET was not performed in 9% (57/631) of the tertiary hospitals and 24% (28/115) of the secondary hospitals. About 42% (334/788) of the first PETs was performed within 4 weeks after initiation of PD. Moreover, about 26% (202/788) of the respondents repeated PET every 3 months. All these indicate that Chinese HCPs use PET more frequently than those tests commonly recommended by different guidelines, suggesting that in-depth education about PET should be enhanced.

Infectious peritonitis or PD catheter–related infection is the major PD complication that affect long-term PD survival and require immediate empirical treatment with antibiotics and systemic long-term management [31,32]. In China, most participating hospitals implemented the same recommendations of the guideline of International Society for Peritoneal Dialysis [33] for the management of peritonitis in PD. Our survey revealed that 92% (726/788)of PD HCPs acknowledged the values of pathogen detections and drug sensitivity tests, continued antibiotic treatment for 2-3 weeks, and individualized antibiotic regimen for PD-related peritonitis [30,33]. For initial empirical therapy, 84% (659/788) of respondents chose first-generation cephalosporins for gram-negative pathogens, and 78% (612/788) chose third-generation cephalosporins for gram-negative pathogens. Besides the cephalosporins, vancomycin was also chosen for gram-positive pathogens by 61% (481/788) of the respondents and aminoglycoside for the gram-negative pathogens by 49% (386/788). These patterns of choice were in line with multiple guidelines or recommendations regarding PD-related peritonitis [29,30,33].

In our study, 95% (96,460/101,537) of the patients on PD were managed by tertiary hospitals, despite the fact that the number of secondary hospitals in China is about 5 times that of the tertiary hospitals. Although our survey covered a much larger PD patient population, it was likely that our data were skewed toward tertiary facilities. It is also possible that some transferring of patients on PD overlapped, and data were not completely aggregated for analysis in our study.

It is known that the quality of ESKD care is influenced by socioeconomic status and geographic regions [34]. Except for the number of responded hospitals and patients on PD, this study did not find significant regional differences in the answered items listed in the questionnaire. The dominance of patients in tertiary hospitals in this study might contribute to the diminished possible regional discrepancy because most of the tertiary hospitals similarly follow the SOP for PD, despite being from different regions. Due to the infrastructure inadequacy, there were still 13% (94/722) of the PD centers that did not have a dedicated PD reception area; 19% (137/722) had no PD training room, 22% (159/722) had no separated PD storage room, and only 41% (296/722) had their unique PD operation room. Furthermore, 9% (65/722) of PD centers had not established local SOP, and 12% (87/722) had no specific PD-related health records. These need to be improved in the future for sustainable growth of PD along with high-quality standards of care.

This study did not limit the qualification of participants from each hospital and used voluntary responses, and this could cause selection bias. In addition, disposing bias might be introduced during the process of aggregating data from the same facility. More studies on the relationship between PD clinical outcomes and hospital quality management can be more instructive for health care policy makers. Future studies addressing improved prognosis of patients on PD through strengthened measures will require longitudinal analyses using both registry and survey data.

In summary, this first national survey provided data reflecting the current status of PD center management in mainland China. It also collected HCPs' concepts for the routine practice of PD patient management. The discrepancy of PD care between the secondary and tertiary hospitals may guide future iteration of the SOP for PD and quality management, targeted HCP training for PD, as well as value-oriented patient education. To a broader extent, providing a better goal-oriented PD care would help to reduce the substantial health care costs and improve the quality of life of patients with ESKD.

Acknowledgments

The authors appreciated the efforts of all clinical practitioners who participated in this survey.

The authors would like to thank the contributions made by the study teams of the national center for health care quality management in kidney diseases: Wenhu Liu, Aili Jiang, Ying Li, Jinsheng Xu, Lihua Wang, Caili Wang, Hongli Lin, Li Yao, Lining Miao, Ping Luo, Jundong Jiao, Zhaohui Ni, Changying Xing, Jianghua Chen, Jianmei Zhou, Yi Yu, Jianxin Wan, Rong Wang, Aimin Zhong, Guohua Ding, Hong Liu, Wei Shi, Xiaomei Peng, Hong Li, Yani He, Li Wang, Yibin Yang, Zhu Zhou, Hongli Jiang, Jianqin Wang, Wenbo Hu, Menghua Chen, Jian Liu, Chen Lu, and Xiaoping Yang.

This study is supported by Beijing Municipal Science and Technology Commission (Z191100007619054) and the Key Research and Development Program of Science and Technology Ministry (2022YFC3602005).

Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of our research.

Conflicts of Interest

None declared.

References

1. Yao Q, Dong J, Feng T, Lindholm B. What PD research in China tells us. Perit Dial Int 2018 Dec 01;38(Suppl 2):S19-S24 [FREE Full text] [doi: 10.3747/pdi.2018.00027] [Medline: 30315043]
2. Chen X. editor. Peritoneal Dialysis Standard Operating Procedure (SOP) In Chinese. Beijing, China: People's Millitary Medical Press; 2010.
3. Chen X. Practical Course in Peritoneal Dialysis in Chinese. Beijing: People's Millitary Medical Press; 2013.
4. China Healthcare Development Statistics. URL: http://www.nhc.gov.cn/ [accessed 2022-01-09]
5. Liyanage T, Ninomiya T, Jha V, Neal B, Patrice HM, Okpechi I, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet 2015 May 16;385(9981):1975-1982. [doi: 10.1016/S0140-6736(14)61601-9] [Medline: 25777665]
6. Luyckx VA, Al-Aly Z, Bello AK, Bellorin-Font E, Carlini RG, Fabian J, et al. Publisher correction: sustainable development goals relevant to kidney health: an update on progress. Nat Rev Nephrol 2021 Oct 02;17(10):704 [FREE Full text] [doi: 10.1038/s41581-021-00473-9] [Medline: 34341564]
7. Cao X, Zhou J, Cai G, Tan N, Huang J, Xie X, et al. Predicting one-year mortality in peritoneal dialysis patients: an analysis of the China Peritoneal Dialysis Registry. Int J Med Sci 2015;12(4):354-361 [FREE Full text] [doi: 10.7150/ijms.11694] [Medline: 26019685]
8. Zhang L, Zhao MH, Zuo L, Wang Y, Yu F, Zhang H, CK-NET Work Group. China Kidney Disease Network (CK-NET) 2015 annual data report. Kidney Int Suppl (2011) 2019 Mar;9(1):e1-e81 [FREE Full text] [doi: 10.1016/j.kisu.2018.11.001] [Medline: 30828481]
9. Xie F, Zhang D, Wu J, Zhang Y, Yang Q, Sun X, et al. Design and implementation of the first nationwide, web-based Chinese Renal Data System (CNRDS). BMC Med Inform Decis Mak 2012 Feb 28;12(1):11 [FREE Full text] [doi: 10.1186/1472-6947-12-11] [Medline: 22369692]
10. Yao Q, Zhang W, Qian J. Peritoneal dialysis in shanghai. Perit Dial Int 2008 Jun;28 Suppl 3:S42-S45. [Medline: 18552262]
11. Zhang X, Chen Y, Cai Y, Tian X, Xiao J, Zhao Z, et al. Characteristics of patients initializing peritoneal dialysis treatment from 2007 to 2014: analysis from Henan Peritoneal Dialysis Registry data. Iran J Kidney Dis 2018 May;12(3):178-184 [FREE Full text] [Medline: 29891748]
12. Wu C, Chen X, Ying Wang A, Chen J, Gao H, Li G, et al. Peritoneal dialysis in Sichuan province of China – report from the Chinese National Renal Data System. Ren Fail 2018 Oct 22;40(1):577-582. [doi: 10.1080/0886022x.2018.1496933]

13. Yu X, Chen J, Ni Z, Chen N, Chen M, Dong J, et al. Number of daily peritoneal dialysis exchanges and mortality risk in a Chinese population. Perit Dial Int 2018 Dec 01;38(Suppl 2):S53-S63 [FREE Full text] [doi: 10.3747/pdi.2017.00283] [Medline: 30315040]
14. Yang C, Gao B, Zhao X, Su Z, Sun X, Wang H, et al. Executive summary for China Kidney Disease Network (CK-NET) 2016 annual data report. Kidney Int 2020 Dec;98(6):1419-1423. [doi: 10.1016/j.kint.2020.09.003] [Medline: 33276868]
15. Chinese National Renal Data System. URL: http://www.cnrds.net/TxLogin [accessed 2022-01-09]
16. USRDS. US renal data system 2019 annual data report: epidemiology of kidney disease in the United States. Am J Kidney Dis 2019 Oct 31;76(3):450. [doi: 10.1053/j.ajkd.2019.09.002] [Medline: 31704084]
17. Karkar A. Response to: International society for peritoneal dialysis practice recommendations: prescribing high-quality goal-directed peritoneal dialysis. Perit Dial Int 2020 Apr 13;40(4):425-426. [doi: 10.1177/0896860820917821]
18. Mikhail A, Brown C, Williams JA, Mathrani V, Shrivastava R, Evans J, et al. Renal association clinical practice guideline on anaemia of chronic kidney disease. BMC Nephrol 2017 Nov 30;18(1):345 [FREE Full text] [doi: 10.1186/s12882-017-0688-1] [Medline: 29191165]
19. Yap DYH, McMahon LP, Hao C, Hu N, Okada H, Suzuki Y, APSN HIF-PHI Recommendation Committee. Recommendations by the Asian Pacific society of nephrology (APSN) on the appropriate use of HIF-PH inhibitors. Nephrology (Carlton) 2021 Feb 09;26(2):105-118 [FREE Full text] [doi: 10.1111/nep.13835] [Medline: 33222343]
20. Kidney Disease: Improving Global Outcomes (KDIGO) CKD-MBD Update Work Group. KDIGO 2017 clinical practice guideline update for the diagnosis, evaluation, prevention, and treatment of chronic kidney disease-mineral and bone disorder (CKD-MBD). Kidney Int Suppl (2011) 2017 Jul;7(1):1-59 [FREE Full text] [doi: 10.1016/j.kisu.2017.04.001] [Medline: 30675420]
21. Locatelli F, Aljama P, Canaud B, Covic A, De Francisco A, Macdougall IC, Anaemia Working Group of European Renal Best Practice (ERBP). Target haemoglobin to aim for with erythropoiesis-stimulating agents: a position statement by ERBP following publication of the Trial to reduce cardiovascular events with Aranesp therapy (TREAT) study. Nephrol Dial Transplant 2010 Sep 29;25(9):2846-2850. [doi: 10.1093/ndt/gfq336] [Medline: 20591813]
22. Kliger AS, Foley RN, Goldfarb DS, Goldstein SL, Johansen K, Singh A, et al. KDOQI US commentary on the 2012 KDIGO clinical practice guideline for anemia in CKD. Am J Kidney Dis 2013 Nov;62(5):849-859. [doi: 10.1053/j.ajkd.2013.06.008] [Medline: 23891356]
23. KDIGO. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int Suppl 2013;3(1):1-150 [FREE Full text]
24. Dahnan M, Assabri AM, Khader YS. Risk factors for end-stage renal failure among patients on hemodialysis in Aljomhory hospital, Sa'adah governorate, Yemen: hospital-based case-control study. JMIR Public Health Surveill 2019 Sep 25;5(3):e14215 [FREE Full text] [doi: 10.2196/14215] [Medline: 31573930]
25. Zeng Q, Yang J, Wang Z, Liu H, Wang J, Yang T, et al. The epidemiological characteristics of noncommunicable diseases and malignant tumors in Guiyang, China: cross-sectional study. JMIR Public Health Surveill 2022 Oct 28;8(10):e36523 [FREE Full text] [doi: 10.2196/36523] [Medline: 36306160]
26. Vaios V, Georgianos PI, Liakopoulos V, Agarwal R. Assessment and management of hypertension among patients on peritoneal dialysis. CJASN 2018 Oct 19;14(2):297-305. [doi: 10.2215/cjn.07480618]
27. Blake PG, Bargman JM, Brimble KS, Davison SN, Hirsch D, McCormick BB, Canadian Society of Nephrology Work Group on Adequacy of Peritoneal Dialysis. Clinical practice guidelines and recommendations on peritoneal dialysis adequacy 2011. Perit Dial Int 2011 Mar 01;31(2):218-239. [doi: 10.3747/pdi.2011.00026] [Medline: 21427259]
28. van Biesen W, Heimburger O, Krediet R, Rippe B, La Milia V, Covic A, ERBP working group on peritoneal dialysis. Evaluation of peritoneal membrane characteristics: clinical advice for prescription management by the ERBP working group. Nephrol Dial Transplant 2010 Jul 04;25(7):2052-2062. [doi: 10.1093/ndt/gfq100] [Medline: 20203287]
29. Woodrow G, Fan SL, Reid C, Denning J, Pyrah AN. Renal Association Clinical Practice Guideline on peritoneal dialysis in adults and children. BMC Nephrol 2017 Nov 16;18(1):333 [FREE Full text] [doi: 10.1186/s12882-017-0687-2] [Medline: 29145808]
30. Ito Y, Tawada M, Yuasa H, Ryuzaki M. New Japanese society of dialysis therapy guidelines for peritoneal dialysis. Contrib Nephrol 2019;198:52-61. [doi: 10.1159/000496523] [Medline: 30991403]
31. Crabtree JH, Shrestha BM, Chow K, Figueiredo AE, Povlsen JV, Wilkie M, et al. Creating and maintaining optimal peritoneal dialysis access in the adult patient: 2019 update. Perit Dial Int 2019 Sep 01;39(5):414-436 [FREE Full text] [doi: 10.3747/pdi.2018.00232] [Medline: 31028108]
32. Szeto C, Li PK, Johnson DW, Bernardini J, Dong J, Figueiredo AE, et al. ISPD catheter-related infection recommendations: 2017 update. Perit Dial Int 2017 Mar 01;37(2):141-154. [doi: 10.3747/pdi.2016.00120] [Medline: 28360365]
33. Li PK, Szeto CC, Piraino B, de Arteaga J, Fan S, Figueiredo AE, et al. ISPD peritonitis recommendations: 2016 update on prevention and treatment. Perit Dial Int 2016 Sep 10;36(5):481-508 [FREE Full text] [doi: 10.3747/pdi.2016.00078] [Medline: 27282851]
34. Cho Y, Bello AK, Levin A, Lunney M, Osman MA, Ye F, et al. Peritoneal dialysis use and practice patterns: an international survey study. Am J Kidney Dis 2021 Mar;77(3):315-325 [FREE Full text] [doi: 10.1053/j.ajkd.2020.05.032] [Medline: 32800844]

Abbreviations

CNRDS: Chinese National Renal Data System
ESKD: end-stage kidney disease
HCP: health care provider
PD: peritoneal dialysis
PET: peritoneal equilibrium test
SOP: standard operating procedure

Edited by A Mavragani; submitted 15.05.22; peer-reviewed by J Chen, WC Lee; comments to author 20.10.22; revised version received 09.11.22; accepted 25.11.22; published 14.03.23.

Please cite as:
Li P, Cao X, Liu W, Zhao D, Pan S, Sun X, Cai G, Zhou J, Chen X
Peritoneal Dialysis Care in Mainland China: Nationwide Survey
JMIR Public Health Surveill 2023;9:e39568
URL: https://publichealth.jmir.org/2023/1/e39568
doi:10.2196/39568
PMID:36917165

Original Paper

Association Between Joint Physical Activity and Dietary Quality and Lower Risk of Depression Symptoms in US Adults: Cross-sectional NHANES Study

Jinghong Liang[1], MD; Shan Huang[1], MM; Nan Jiang[1], MM; Aerziguli Kakaer[1], MM; Yican Chen[1], MM; Meiling Liu[1], MM; Yingqi Pu[1], MM; Shaoyi Huang[1], MM; Xueya Pu[1], MM; Yu Zhao[1], MM; Yajun Chen[1], PhD

[1]Department of Maternal and Child Health, School of Public Health, Sun Yat-sen University, Guangzhou, China

Corresponding Author:
Yajun Chen, PhD
Department of Maternal and Child Health, School of Public Health, Sun Yat-sen University
No.74 Zhongshan 2nd Road, Yuexiu District
Guangzhou, 510080
China
Phone: 86 17372110875
Email: chenyj68@mail.sysu.edu.cn

Abstract

Background: Depression escalating public health concern and the modest efficacy of currently available treatments have prompted efforts to identify modifiable risk factors associated with depression symptoms. Physical inactivity, poor nutrition, or other lifestyle behaviors are among the potentially modifiable risk factors most consistently linked with depression. Past evidence regarding the single effect of physical activity (PA) or dietary quality (DQ) on reducing the risk of depression symptoms has been well-documented. However, the association of the joint effect of PA and DQ on depression symptoms has never been investigated in a representative sample of adults.

Objective: This study investigates the association between PA and depression symptoms and between DQ and depression symptoms, and their combined effects on US adults.

Methods: Data were obtained from the National Health and Nutrition Examination Survey (NHANES) 2007 to 2018 cycles. The primary exposures were DQ and PA, measured using the Healthy Eating Index (HEI)-2015 and the metabolic equivalent (MET) minutes per week reported in questionnaires, respectively. Depression symptoms were defined as a 9-item Patient Health Questionnaire (PHQ-9) score of ≥10. We created 4 lifestyle categories: healthy diet and active individuals, unhealthy diet but active individuals, healthy diet but inactive individuals, and unhealthy diet and inactive individuals. Participants were considered to have a healthy diet if they fell within the 60th percentile of the HEI-2015 or to be active if they met the current guidelines for PA. A survey-multivariable logistic regression approach was used to model adjust the variables relevant to the associations, and an age-adjusted prevalence for depression symptoms was calculated following the NHANES guidelines.

Results: In total, 19,295 participants represented a weighted number of 932.5 million adults aged 20 to 80 years in the noninstitutionalized US population. The total age-adjusted prevalence of depression symptoms among all respondents was 7.08% (1507/19,295). Of the respondents, 81.97% (15,816/19,295) met the PA recommendation and 26.79% (5170/19,295) scored at or above the 60th percentile on the HEI-2015. Depression symptoms were inversely associated with a higher level of PA (adjusted odds ratio [AOR] 0.819, 95% CI 0.716-0.938) and healthy DQ (AOR 0.809, 95% CI 0.701-0.931), respectively. A healthy diet combined with recommended PA was associated with a significantly lower risk of depression symptoms (AOR 0.658, 95% CI 0.538-0.803) than those who consumed an unhealthy diet but were physically active (AOR 0.890, 95% CI 0.765-1.038) or consumed a healthy diet but were physically inactive (AOR 1.077, 95% CI 0.817-1.406).

Conclusions: Our findings indicate that people with a healthy diet and recommended PA have a lower risk of depression symptoms than those with an unhealthy diet and a low level of PA. A healthy dietary habit and regular PA are potential precautions against depression.

(JMIR Public Health Surveill 2023;9:e45776) doi:10.2196/45776

KEYWORDS

physical activity; dietary quality; depression symptom; adults; NHANES

Introduction

Depression, a major worldwide public health concern commonly characterized by low emotion, diminished interest, and decreased energy and attention, has been affecting nearly 7.2% to 9.2% of the US population, particularly adults, from 2015 to 2020 [1]. Depression will continue to affect individuals, the health care system, and the whole society if left unprevented [2]. The growing prevalence of depression poses an increasing demand for highly effective, preventative approaches. Some studies have reported potentially protective effects of dietary quality (DQ) or physical activity (PA) on depression symptoms, as well as the biological actions of nutrients or behavior, both within and beyond the context of healthy behavior [3,4]. Understanding PA and DQ is vital to intervene against depression symptom risks during adulthood, as dietary pattern, PA, and other behaviors are the major triggers determining the well-being of mental health [5].

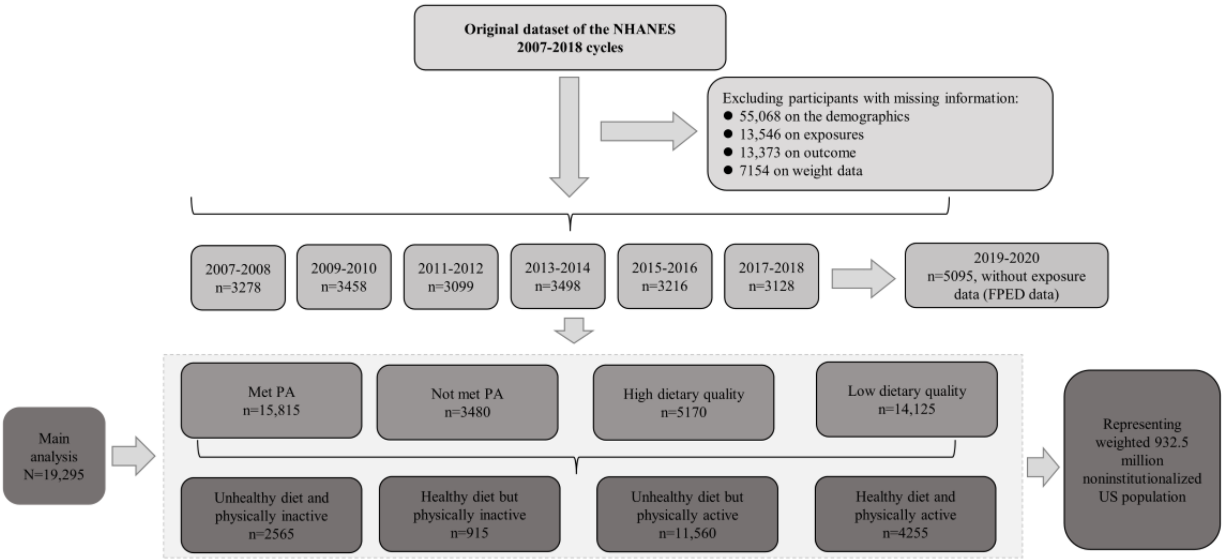
Nevertheless, robust evidence concerning the association between the single effect of DQ or PA and the prevalence of depression symptoms remains sparse and debated. For example, although PA was linked to a decreased risk of depression in several studies [6,7], these findings were not consistently replicated in another study [8]. Moreover, little is known about the combined influence of dietary behaviors and PA on depression symptoms in the US population. It can be argued that healthy diet and sufficient activity may be more effective than unhealthy diet and inactivity in promoting the mental health of adults. Nevertheless, empirical evidence for it is far from clear cut, and there exist limitations in prior related studies [9-14]. First, most previous studies simply examined isolated nutrients, such as energy intake and eating behavior, for depression symptoms, rather than the total DQ [9,10], which is a comprehensive representation of an individual's overall diet pattern. Second, the most important finding is that no studies so far have specifically targeted the association between DQ and PA and depression symptoms among the US population, except for some studies that investigated this exposure on other health indicators [11-13]. Furthermore, a few studies reported the trends for DQ with serial cross-sectional investigations using older data (updated to the year 2016) [14], but these findings need to be verified by more subsequent national data (National Health and Nutrition Examination Survey [NHANES] is updated to the 2020 cycle). Finally, as there might be no race-related information available in their studies, the impact of discrepancy on race or ethnicity on diets, which deserves more attention, was often neglected in previous studies. For example, Black people, commonly reported to have poor DQ, have the highest morbidity rates of diet-related diseases in comparison with other racial or ethnic groups in the United States [15,16]. Individual depression symptoms might be affected by lifestyle, depending on sex, age, and race, but no relevant study has examined the heterogeneity of populations. To compensate for these limitations, we sought to describe the prevalence of depression symptoms and its association with DQ and PA, considering age, race or ethnicity, sex, and sociodemographic status, among the US population using data from the continuous NHANES.

Methods

Study Design

The NHANES is an ongoing cross-sectional program conducted by the National Center for Health Statistics of the Centers for Disease Control and Prevention to investigate the public's general health-related behaviors, their socioeconomic and nutritional status, and the results of physical examinations. With a multistage, stratified probability sample selected from cities, blocks, households, and the number of people within households, the NHANES offers comprehensive data to represent the US civilian and noninstitutionalized population. The objectives, study design, and implementation of the NHANES have been introduced before [14]. The 24-hour dietary recall was administered in person by a trained interviewer using the United States Department of Agriculture automated multiple-pass method [17]. This study was conducted in accordance with the Declaration of Helsinki [18]. Flowchart of selection of NHANES participants can be found in Figure 1.

Figure 1. Flowchart of selection of National Health and Nutrition Examination Survey (NHANES) participants: unhealthy diet and physically inactive individuals (participants did not meet the US physical activity [PA] recommendation guideline and below the 60th percentile of the Healthy Eating Index [HEI]-2015 score), healthy diet but physically inactive individuals (participants did not meet the US PA recommendation guideline but at or above the 60th percentile of the HEI-2015 score), unhealthy diet but physically active individuals (participants met the US PA recommendation guideline but below the 60th percentile of the HEI-2015 score), healthy diet and physically active individuals (participants met the US PA recommendation guideline and at or above the 60th percentile of the HEI-2015 score). FPED: food patterns equivalents database.



Ethics Approval and Informed Consent

Written informed consent was obtained from all study participants, and the program was approved by the Ethics Review Board of the National Center for Health Statistics [19]. The NHANES data are free for public use and available on the web [20], so it was not necessary to obtain the agreement of the medical ethics committee board [21].

Assessment for DQ

With an interval of 3 to 10 days, two 24-hour recalls of all the food and drink consumed the day before the interview (midnight to midnight) were conducted. The Healthy Eating Index (HEI)-2015 is a comprehensive measurement of an individual's dietary pattern with overall and subcomponents scores, which adheres to the 2010 Dietary Guidelines for Americans [22]. The HEI-2015 contains 13 nutrient- and food-based components, which comprise 9 adequacy food components (total fruits, whole fruits, total vegetables, greens and beans, whole grains, dairy, total protein foods, seafood and plant proteins, and fatty acids) and 4 moderation components (refined grains, sodium, added sugars, and saturated fats). The total HEI-2015 score ranges from 0 to 100, with a higher score indicating greater overall DQ [22]. Participants whose average score of the 2 days was in or above the 60th percentile in the HEI were considered to have observed the dietary guidelines or consumed healthy foods [23].

Assessment for PA

Data were collected using the Global Physical Activity Questionnaire. Created by the World Health Organization (WHO), the Global Physical Activity Questionnaire was used to assess different domains of individuals' PA, such as leisure-time PA, occupation, and transportation PA [24,25]. In accordance with the WHO analysis guide, PA was converted to metabolic equivalent (MET) minutes of moderate to vigorous PA per week [24]. MET values vary with the type of exercise, and the NHANES offers the recommended MET values for each PA. PA was based on the MET values of type, frequency, and duration of activities per week, which was calculated using the following formula: PA (MET-min/wk) = MET × weekly frequency × duration of each PA [26]. PA=0 denotes participants who do not engage in any PA, else, it means that participants have constant or intermittent PA. The respondents were classified based on whether they met the American PA guideline (PA at moderate intensity should be done for 150 minutes [equivalent to 600 MET min/wk] a week or, at vigorous intensity, should be performed 75 min/wk for adults [27]).

Definitions of Lifestyle Groups for PA and DQ

According to previous studies [13,23], we created four lifestyle categories according to DQ and PA: (1) healthy diet and active individuals, (2) unhealthy diet but active individuals, (3) healthy diet but inactive individuals, and (4) unhealthy diet and inactive individuals. Participants were considered to have a healthy diet if they fell within the 60th percentile of the HEI or were active if they met the current guidelines for PA [13,27].

Outcome Measurements

The 9-item Patient Health Questionnaire (PHQ-9) is a brief self-report measure of depression symptoms in a primary care and research setting that has a well-established factor structure, reliability, and validity [28,29]. The range of scores for PHQ-9 is from 0 to 27, with higher scores reflecting greater severity, and a PHQ-9 score of ≥10 was recommended as the binary threshold to define the presence of depression symptoms, which has a sensitivity of 88% and a specificity of 88% for screening major depression symptoms [30].

Assessment of Covariates

According to prior research and clinical experts, potentially confounding and modifying variables were identified [31] as follow: age group (20-39, 40-59, or 60-80 years), sex (male or female), race or ethnicity (non-Hispanic White, non-Hispanic Black, Mexican American, Other races [including multiracial and other Hispanic]), the highest level of education achieved (less than 9th grade, 9th-11th grade [including 12th grade without diploma], high school graduate [general educational development or equivalent], college graduate or above, or some college or associate's degree), marital status (widowed or divorced or separated, never married, married or living with partner), poverty-to-income ratio (PIR; a family poverty index was used to estimate socioeconomic status, which is calculated by measuring the income thresholds of different types of households and updating them annually because of inflation, based on the Consumer Price Index. According to the commonly used cutoff levels for identifying eligibility for federal assistance, we classified these ratios as being either ≤100% or ≥100% of the poverty threshold, depending on whether they exceeded the poverty threshold), BMI (calculated based on the ratio of weight in kilograms to height in meters squared), smoking status (current, previous, or never), whether taking antidepressant or anxiolytic medications, sleep time (<7 h, 9 h, and >9 h), and sedentary behavior (SB; the amount of time spent sitting per day is considered SB, including sitting at a desk reading; playing cards; watching television or using a computer; sitting with friends; and traveling by car, bus, or train).

Statistical Analyses

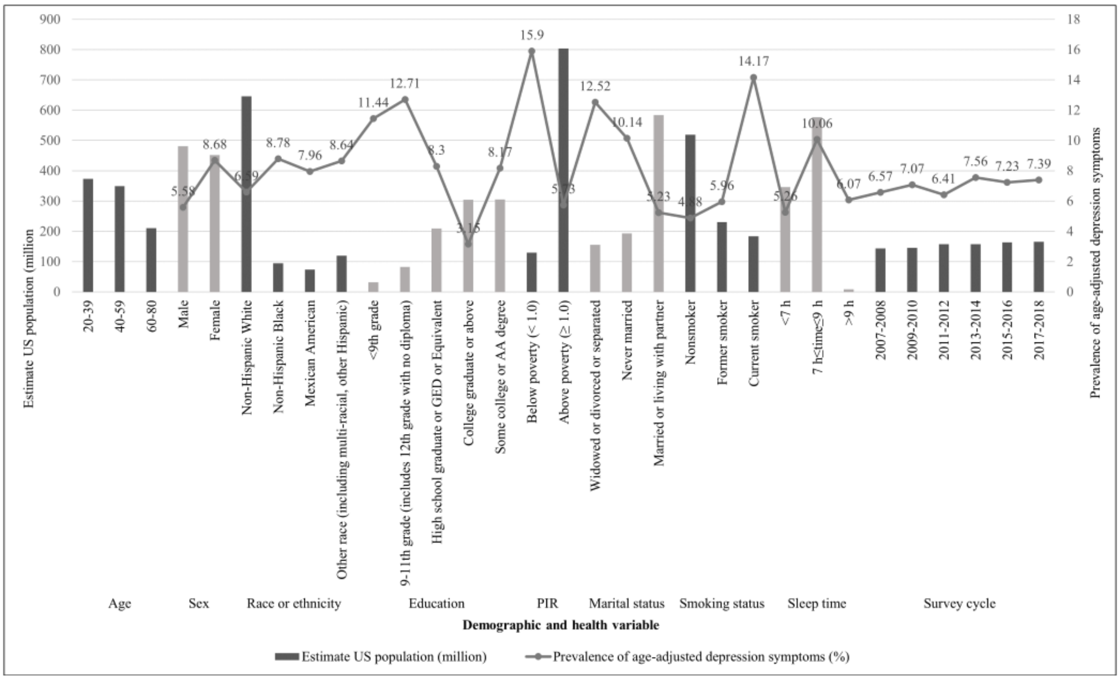
We used weighted samples and considered stratification and clustering in the design to generate nationally representative estimates that were applied to US residents [32]. The combined NHANES cycle and weights were constructed following the guidelines for continuous NHANES analyses [33]. Proportional outcome variables are presented as means (SD) with analysis of variance F and the proportional outcome variables as proportions (%) with analysis of the chi-square test to determine statistical significance by 4 groups. The estimates obtained from adjusted odds ratios (AORs) and corresponding 95% CIs were used to evaluate the associations between PA and DQ and depression symptoms. A survey-multivariable logistic regression approach was used to adjust the variables relevant to the associations. For each lifestyle group, the median values of health variables were treated as continuous variables in a multivariate linear regression model to estimate the trend between each lifestyle and depression symptom. We used 3 different models: a crude model; a model adjusted for age, sex, and race or ethnicity; and a model in which we additionally adjusted for age, sex, race or ethnicity, education, marital status, BMI, PIR, smoking status, sleep time, and SB. We also generated 3 indicators along with their 95% CI, namely, relative excess risk due to interaction (RERI), attributable proportion due to interaction (AP), and synergy index (SI), to estimate the interactive effect between PA and DQ [34,35]. The 95% CI of RERI and AP containing 0 indicates there is an additive interaction between the 2 exposures, and the greater the absolute value of RERI, the stronger the interaction between the 2 exposures (PA and DQ) [36]. For a 95% CI of SI, if it contains 1, it indicates that there exists an additive interaction between the 2 exposures (PA and DQ) simultaneously. An SI value greater than 1 indicates that the synergistic effect is enhanced when the factors were both exposures, while a value below 1 indicates that the effect of the 2 factors is weakened [36]. An age-adjusted prevalence for depression symptoms was calculated according to the NHANES guidelines [37]. To account for potential variability in the associations in terms of age, sex, and other demographic characteristics, exploratory subgroup analyses were performed with the interactions obtained, followed by sensitivity analyses after excluding individuals exposed to extreme PIR or individuals with extreme BMI values to identify the robustness of our findings. Moreover, alcohol use (former, never, and now), total energy, and whether taking antidepressant or anxiolytic medications were regarded as additional covariates to control their potential influence. The probability of the null hypothesis value (*P* value) below .05 was considered to reflect a statistically significant difference. All statistical analyses were conducted using R statistical programming language (X64 version 4.1.0; R Foundation for Statistical Computing).

Results

General Characteristics

Approximately, a weighted population of 932.5 million American residents was represented by 19,295 NHANES participants with valid data, and most of them were non-Hispanic White participants (645.6 million; Multimedia Appendix 1). Among the 19,295 respondents, 52.59% (n=10,147) were female, 44.76% (n=8637) were non-Hispanic White, 80.54% (n=15,541) had at least a high school education, and 20% (n=3859) lived in households with incomes below the federal poverty level. Overall, 81.97% (15,815/19,295) of the respondents met the PA recommendation, and 26.79% (5170/19,295) of participants scored at or above the 60th percentile on the HEI-2015. Moreover, participants with a healthy diet and regular exercise had an average HEI-2015 score of 69.32 (SD 0.19) and were active with 4476 (SD 105.15) MET minutes moderate to vigorous PA per week. In contrast, people who eat unhealthily and are inactive have an average score of 44.87 (SD 0.25) and only 316.90 (SD 4.24) minutes of moderate to vigorous PA per week. The age-adjusted prevalence of depression symptoms among respondents who did not meet the PA guidelines and had a low-quality diet was 9.51% (264/2565). In total, 4.19% (200/4255) of participants with depression symptoms in the group met the PA recommendations and had a higher DQ. The age-adjusted prevalence for another 2 groups was 7.54% (961/11,560; unhealthy diet but physically active) and 8% (82/915; healthy diet but physically inactive). We also found a relatively higher age-adjusted depression symptom prevalence in female (922/9148, 8.68%), low education participants (<9th grade: 152/1381, 11.44%; 9th-11th grade: 313/2373, 12.71%), and especially smokers (586/3982, 14.17%). Multimedia Appendix 1 and Figure 2 present the characteristics of the 19,295 US adults according to their lifestyles. The demographic characteristics and health variables of participants, both classified by depression symptom status and 6 survey cycles, are presented separately in Multimedia Appendices 2 and 3.

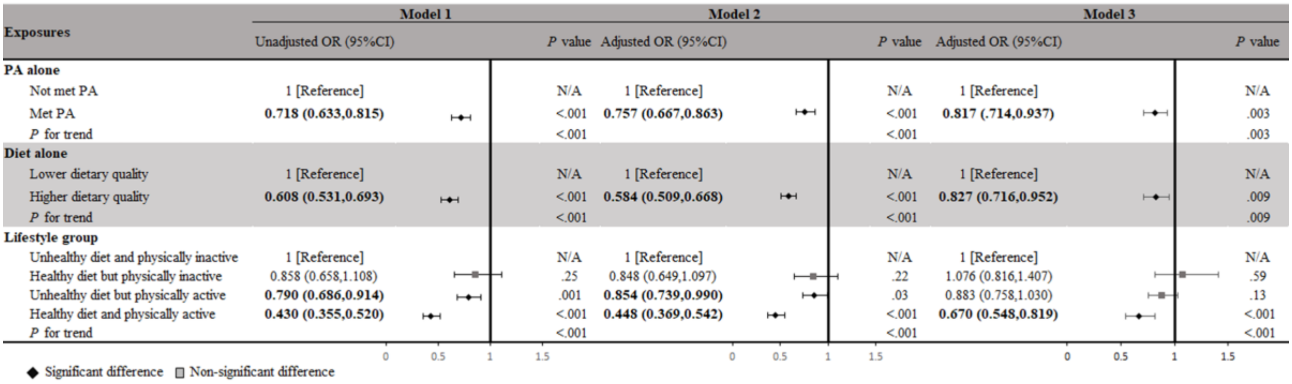
Figure 2. Prevalence of age-adjusted depression symptoms and estimate US population of different characteristics among US general adults, National Health and Nutrition Examination Survey 2007 to 2018. AA: Associate's Degree; GED: general educational development; PIR: poverty income ratio.



Association Between Single PA or DQ and Depression Symptoms

For the single effect of PA or DQ on depression symptoms, a statistically significant inverse relationship was found between the single PA effect and depression symptom prevalence in each model (*P*<.001; Figure 3). An inverse significant association was detected after controlling for multiple variables in the met PA guidelines group (AOR 0.817, 95% CI 0.714, 0.937; *P*=.003). Similarly, we noted a significant positive association between the single DQ effect and depression symptom prevalence: participants with a higher DQ had a lower risk of depression symptoms (AOR 0.827, 95% CI 0.716, 0.952; *P*=.009) compared with those with poor DQ (Figure 3).

Figure 3. Association between PA and dietary quality and their combined effect and prevalence of depression symptoms among 19,295 study participants. Significant results are in bold. Model 1: crude model, no covariates were adjusted; model 2: adjusted for age, sex, and race or ethnicity; model 3: adjusted for age, sex, race or ethnicity, education, marital status, BMI, poverty-to-income ratio (PIR), smoking status, sedentary behavior (SB), and sleep time. N/A: not applicable; OR: odds ratio; PA: physical activity.



Exposures	Model 1		Model 2		Model 3	
	Unadjusted OR (95%CI)	*P* value	Adjusted OR (95%CI)	*P* value	Adjusted OR (95%CI)	*P* value
PA alone						
Not met PA	1 [Reference]	N/A	1 [Reference]	N/A	1 [Reference]	N/A
Met PA	**0.718 (0.633,0.815)**	<.001	**0.757 (0.667,0.863)**	<.001	**0.817 (.714,0.937)**	.003
P for trend		<.001		<.001		.003
Diet alone						
Lower dietary quality	1 [Reference]	N/A	1 [Reference]	N/A	1 [Reference]	N/A
Higher dietary quality	**0.608 (0.531,0.693)**	<.001	**0.584 (0.509,0.668)**	<.001	**0.827 (0.716,0.952)**	.009
P for trend		<.001		<.001		.009
Lifestyle group						
Unhealthy diet and physically inactive	1 [Reference]	N/A	1 [Reference]	N/A	1 [Reference]	N/A
Healthy diet but physically inactive	0.858 (0.658,1.108)	.25	0.848 (0.649,1.097)	.22	1.076 (0.816,1.407)	.59
Unhealthy diet but physically active	**0.790 (0.686,0.914)**	.001	**0.854 (0.739,0.990)**	.03	0.883 (0.758,1.030)	.13
Healthy diet and physically active	**0.430 (0.355,0.520)**	<.001	**0.448 (0.369,0.542)**	<.001	**0.670 (0.548,0.819)**	<.001
P for trend		<.001		<.001		<.001

◆ Significant difference ▨ Non-significant difference

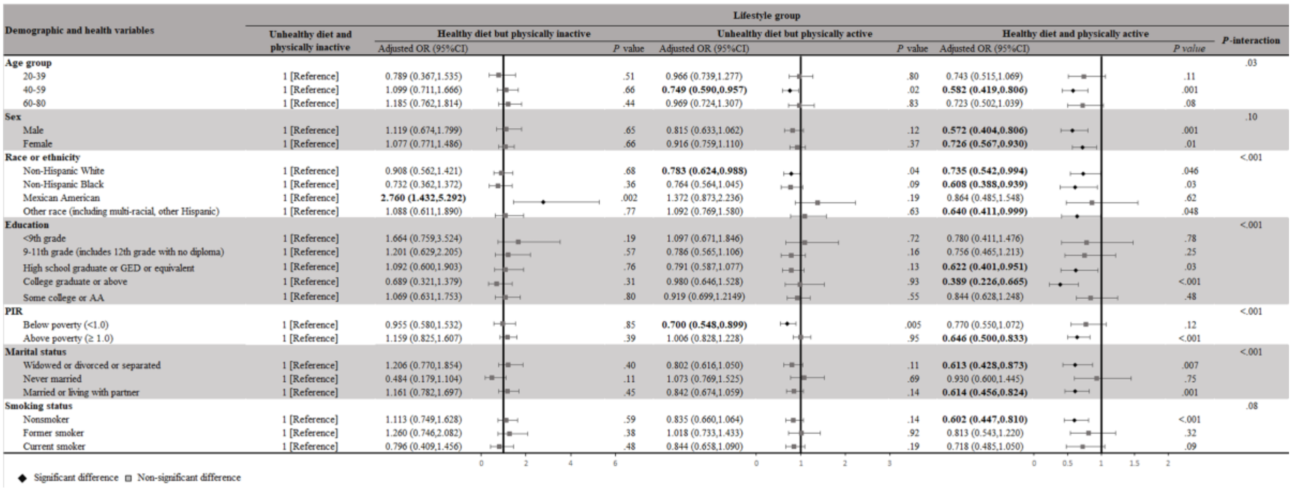
Joint Association Between PA and DQ and Depression Symptoms

Regarding the joint effect of PA and DQ on depression symptoms, when comparing depression symptom prevalence between groups with different lifestyles, participants consuming a healthier diet and being more physically active had a lower risk ratio of depression symptoms of 0.670 (95% CI 0.548-0.819; *P*<.001). Participants who consumed a healthy diet but were physically inactive did not have a significantly lower risk of depression symptoms (AOR 1.076, 95% CI 0.816-1.407; *P*=.59), as well as those who consumed an unhealthy diet and were physically active (AOR 0.883, 95% CI 0.758-1.030; *P*=.13). The results regarding the additive interaction between the 2 exposures (PA and DQ) showed that the RERI and AP between PA and DQ were statistically significant (RERI=−0.052, 95% CI −2.149 to 2.044; AP=−0.078, 95% CI −3.333 to 3.177), which meant there is an additive interaction between the 2 exposures. A similar significant interactive effect of SI (SI=1.189, 95% CI 0.001-1170.880) was observed, which indicates the synergistic effect of PA and DQ is enhanced in our model (Figure 3).

Subgroup Analyses and Sensitivity Analyses

The main analysis of the lifestyle-depression symptom association based on demographic subgroups showed similar AORs within each group (Figure 4). The AORs remained significant for participants having a healthy diet and PA, who were male (AOR 0.572, 95% CI 0.404-0.806; *P*<.001), who were non-Hispanic Black (AOR 0.608, 95% CI 0.388-0.939; *P*<.05), or had a PIR above the average level (AOR 0.646, 95% CI 0.500-0.833; *P*<.001). In the sensitivity analyses for depression symptoms attributed to lifestyle exposures, after further adjusting for several potential indicators (alcohol users [former, never, and now], total energy, whether taking antidepressant or anxiolytic medications, and sleep time), the results remained similar in that DQ and PA were positively correlated with mental health (Multimedia Appendix 4).

Figure 4. Associations between exposure to different lifestyle groups and incidence of depression symptoms in stratified analyses. Significant results are in bold. The multivariable model was adjusted for age, sex, race or ethnicity, education, marital status, BMI, PIR, smoking status, SB, and sleep time. AA: Associate's Degree; GED: general educational development; OR: odds ratio; PIR: poverty-to-income ratio; SB: sedentary behavior.



Demographic and health variables	Lifestyle group							P-interaction
	Unhealthy diet and physically inactive	Healthy diet but physically inactive		Unhealthy diet but physically active		Healthy diet and physically active		
		Adjusted OR (95%CI)	*P* value	Adjusted OR (95%CI)	*P* value	Adjusted OR (95%CI)	*P* value	
Age group								.03
20-39	1 [Reference]	0.789 (0.367,1.535)	.51	0.966 (0.739,1.277)	.80	0.743 (0.515,1.069)	.11	
40-59	1 [Reference]	1.099 (0.711,1.666)	.66	**0.749 (0.590,0.957)**	.02	**0.582 (0.419,0.806)**	.001	
60-80	1 [Reference]	1.185 (0.762,1.814)	.44	0.969 (0.724,1.307)	.83	0.723 (0.502,1.039)	.08	
Sex								.10
Male	1 [Reference]	1.119 (0.674,1.799)	.65	0.815 (0.633,1.062)	.12	**0.572 (0.404,0.806)**	.001	
Female	1 [Reference]	1.077 (0.771,1.486)	.66	0.916 (0.759,1.110)	.37	**0.726 (0.567,0.930)**	.01	
Race or ethnicity								<.001
Non-Hispanic White	1 [Reference]	0.908 (0.562,1.421)	.68	**0.783 (0.624,0.988)**	.04	**0.735 (0.542,0.994)**	.046	
Non-Hispanic Black	1 [Reference]	0.732 (0.362,1.372)	.36	0.764 (0.564,1.045)	.09	**0.608 (0.388,0.939)**	.03	
Mexican American	1 [Reference]	**2.760 (1.432,5.292)**	.002	1.372 (0.873,2.236)	.19	0.864 (0.485,1.548)	.62	
Other race (including multi-racial, other Hispanic)	1 [Reference]	1.088 (0.611,1.890)	.77	1.092 (0.769,1.580)	.63	**0.640 (0.411,0.999)**	.048	
Education								<.001
<9th grade	1 [Reference]	1.664 (0.759,3.524)	.19	1.097 (0.671,1.846)	.72	0.780 (0.411,1.476)	.78	
9-11th grade (includes 12th grade with no diploma)	1 [Reference]	1.201 (0.629,2.205)	.57	0.786 (0.565,1.106)	.16	0.756 (0.465,1.213)	.25	
High school graduate or GED or equivalent	1 [Reference]	1.092 (0.600,1.903)	.76	0.791 (0.587,1.077)	.13	**0.622 (0.401,0.951)**	.03	
College graduate or above	1 [Reference]	0.689 (0.321,1.379)	.31	0.980 (0.646,1.528)	.93	**0.389 (0.226,0.665)**	<.001	
Some college or AA	1 [Reference]	1.069 (0.631,1.753)	.80	0.919 (0.699,1.2149)	.55	0.844 (0.628,1.248)	.48	
PIR								<.001
Below poverty (<1.0)	1 [Reference]	0.955 (0.580,1.532)	.85	**0.700 (0.548,0.899)**	.005	0.770 (0.550,1.072)	.12	
Above poverty (≥ 1.0)	1 [Reference]	1.159 (0.825,1.607)	.39	1.006 (0.828,1.228)	.95	**0.646 (0.500,0.833)**	<.001	
Marital status								<.001
Widowed or divorced or separated	1 [Reference]	1.206 (0.770,1.854)	.40	0.802 (0.616,1.050)	.11	**0.613 (0.428,0.873)**	.007	
Never married	1 [Reference]	0.484 (0.179,1.104)	.11	1.073 (0.769,1.525)	.69	0.930 (0.600,1.445)	.75	
Married or living with partner	1 [Reference]	1.161 (0.782,1.697)	.45	0.842 (0.674,1.059)	.14	**0.614 (0.456,0.824)**	.001	
Smoking status								.08
Nonsmoker	1 [Reference]	1.113 (0.749,1.628)	.59	0.835 (0.660,1.064)	.14	**0.602 (0.447,0.810)**	<.001	
Former smoker	1 [Reference]	1.260 (0.746,2.082)	.38	1.018 (0.733,1.433)	.92	0.813 (0.543,1.220)	.32	
Current smoker	1 [Reference]	0.796 (0.409,1.456)	.48	0.844 (0.658,1.090)	.19	0.718 (0.485,1.050)	.09	

◆ Significant difference ▢ Non-significant difference

Discussion

Principal Findings

Comprising 19,295 participants with 6 waves (NHANES cycles: 2007-2018), the overall analysis indicated that lower depression symptom rates were seen in participants who either followed a healthier dietary pattern or were engaged in more PA and who both had better DQ and more PA. The association remained significant when other potential covariates were controlled for in sensitivity analyses.

The benefits of a healthy diet or PA alone for depression have been well-documented [3,38]. However, data from the latest Global Status Report on PA issued by the WHO in October 2022 showed that 27.5% of adults failed to meet the WHO's recommendation for PA that an adult should work out 150 to 300 minutes a week with moderate PA or 75 to 150 minutes a week at an aerobic intensity [27]. Consequently, among the 27.5% of adults with insufficient PA, 47% had health-related indicators of noncommunicable diseases caused by hypertension, and 43% had depression [39]. Similarly, our results also showed that 18.03% of US adult participants did not have adequate PA, as recommended by the WHO. The similarity of sample characteristics between our study and the global report makes our findings more reliable and increases their authenticity. In line with previous studies, our results showed that a single effect of PA or DQ was significantly associated with mental health [40,41]. This association is explained by the reaction of the endocannabinoid system to acute neuroendocrine stimulation and inflammation [42] and long-term brain adaptations, such as changes in neural architecture [43]. A Mendelian randomization study has investigated more than 100 potentially modifiable factors for their association with incident depression, but it only identified several modifiable risk factors for the prevention of depression, including PA-related domains and dietary [44]. PA showed a potential causal connection with depression in a large-scale genome-wide association study, which suggested more PA might help prevent depression [45].

Comparison With Prior Work

As noted earlier, there might exist a causal relationship between dietary habits and depression [44]. A comprehensive food-based diet pattern made up of multiple dietary components or food groupings, rather than single nutrients or foods, has become widespread and helps us understand health trajectories, as such dietary patterns can evaluate the interplay between nutrients and foods consumed simultaneously through a multidimensional approach with the impact of the overall diet considered [46,47]. Our findings revealed that a healthy diet was negatively associated with depression symptoms among adults by comparing their DQ. Recent studies using the same NHANES data to investigate the association between HEI and depression symptoms concluded that higher HEI scores significantly correlated with fewer depression symptoms [48]. For all age groups, a constantly optimized diet will probably offer substantial health gains, which are predicted to be larger with earlier optimization in diet [49]. The hypotheses for healthy dietary pattern are also applicable to the prevention of depression symptoms, probably because an inflammatory process can be inhibited by a high-quality diet that includes vegetables, fruits, whole grains, nuts, and legumes [50,51]. Several inflammatory markers, namely IL-1, IL-6, TNF-α, and hsCRP, have been found to be elevated in individuals with depression [52], and thus depression may be pathophysiologically underpinned by chronic inflammation. The researchers also found that people who closely followed the specific dietary patterns, such as the Dietary Approaches to Stop Hypertension (DASH) diet, Mediterranean diet, or Mediterranean-DASH Intervention for Neurodegenerative Delay diet—a newly developed one that incorporates elements from the Mediterranean and DASH diets, emphasizes plant-based food, and avoids animal-based food—tended to have lower depression risk [4,45,46]. Among the recommended diet for reducing depression risk are fruits, vegetables, legumes, wholegrain cereals, and nuts instead of unhealthy products such as fast foods, pastries, and soft drinks [53]. Despite the fact that existing evidence on DQ only discussed certain kinds of food, the healthy dietary pattern in our study is comprehensive enough to cover the aforementioned nutrients, whose potential positive impact on depression symptoms is worthy of prompt replication in larger populations [54]. As well-documented as either the single effect of diet or PA alone for preventing depression is, there are no relevant studies that have examined the effects of diet and PA as a whole on depression. PA and a reduction in certain "unhealthy" foods, such as fast foods and pastries, are widely recognized as independently protective against depression. However, these studies used self-reported PA measures and focused on a single nutrient or nutrient group rather than on the quality of the overall diet. Our study provides a new comprehensive diet pattern for depression prevention unlike previous findings focusing on single nutrients or food groups. A comprehensive DQ score captures not only individual intake levels but also higher-level interactions. Meanwhile, our results showed that PA exposure was also a recommended lifestyle. Specifically, this study emphasizes the importance of incorporating dietary behaviors into health behavior models when examining influence of PA on depression. On the basis of the most recent NHANES data (updated to the 2018 cycle) with 19,295 participants on behalf of approximately 932,548,888 US adults, the study measured health parameters for each of the 4 lifestyle groups to further demonstrate the joint benefits of adequate PA and healthy diet for depression symptoms among US adults. Furthermore, to enhance the robustness of our analyses, we have also considered a moderate number of potential confounders, such as SB, an important factor related to individual health that is independent of PA. On the basis of the findings of the integrated effect of PA and DQ on depression symptoms, we found that the risk of depression symptoms had reduced for participants who achieved the recommended PA level or consumed a healthy diet. However, when participants with different PA levels and DQ were classified into 4 lifestyle groups to examine the combined effect of PA and DQ, correlations of the 2 lifestyles (unhealthy diet but physically active and healthy diet but physically inactive) with depression symptoms (reported in Figure 3) were mild and moderate, respectively, without statistical significance. It showed that the combination of PA and DQ may be more effective than either factor alone in reducing depression symptoms. Hence, it may be hypothesized that regular PA and a healthy diet are both at work for the prevention of depression symptoms and neither can be omitted. In addition, our model has further demonstrated that there existed an additive interaction between PA and DQ, and the synergistic effect is enhanced on the improvement of depression symptoms when the recommended PA level and a healthy diet are simultaneously exposed. Such a synergistic effect may be linked to the Lifestyle Medicine (LM) preventive strategy, a multicomponent and multimodal therapeutic approach that combines nutrition, PA, and other behaviors management [55]. That the pathogenesis and progression of depression involve lifestyle determinants has been well established, although the effectiveness of these determinants in helping patients with depression is underappreciated in current first-line treatments [56]. The analysis of the integrated effect in our study agreed with the results of the meta-analysis that examined the effect of multicomponent LM interventions, which include PA, DQ, and smoking cessation, on relieving depression symptoms (nonclinical depression) among adults in comparison with the named control group [5]. This study deliberately excluded patients with clinical depression to examine the preventive effect of LM intervention on depression symptoms [5]. The results highlighted that patients with depression might benefit greatly from the LM intervention, whose potential effect in the prevention of depression symptoms may be valued by major clinical guidelines and mental health experts [57]. One meta-analysis analyzed 50 randomized controlled trials concerning LM intervention and revealed that the effect of multicomponent LM intervention on depression symptoms was greatly moderated by the number of lifestyle factors adopted [58], which was consistent with our main finding that the synergistic effect was enhanced when more lifestyle behaviors that are beneficial to depression symptoms have been involved in one's daily life. Considering that the multicomponent LM intervention could possibly be used in clinical practice of controlling depression, examining the preventative effect of this intervention on depression symptoms is therefore essential. LM may be a promising complementary treatment for depression symptoms. However, more research is needed to determine its effectiveness and applicability as an adjunct to existing first-line treatments and to understand the impact of LM interventions on different populations, such as adolescents and older adults [59]. On the basis of these findings, it is important to combine DQ and PA in national and community strategies for depression symptoms to obtain the optimal prevention effects.

Limitations

However, we acknowledge the possibility of limitations in our study that must be considered. First, PA was only measured by self-report, which is subject to recall and social desirability biases. Second, our cross-sectional study design made it difficult to establish causal links between diet and PA and depression. Nevertheless, evidence regarding the single effect of PA or a healthy diet on depression symptoms has been identified in previous Mendelian randomization studies [44,45], implying that reverse causation was less likely to occur in our study. Thus, high-quality prospective population-based studies should be conducted to verify this causal relationship. Third, as another common depression symptom population, children and adolescents are not included as they have not completed the relevant assessment concerning PA and depression symptoms. Thus, caution should be exercised when extrapolating this association in general adults to younger populations. Fourth, several potential risk factors, such as hypertension and diabetes mellitus, may also be the cause of the depression symptoms but are not considered as covariates in our study. Moreover, as our study focused on the PA pattern, we did not further explore the association between PA frequency, duration, and intensity and depression symptoms, and the dose-response relationship between them, which could be examined in a future study.

Conclusions

On the basis of the NHANES data from the 1999 to 2018 cycle, our study found that a higher level of PA incorporated with better DQ was inversely related to depression symptoms in a representative sample of the US population, which meant that participants experiencing the recommended PA level and consuming a healthy diet together enhance the synergistic effect on depression symptoms. Therefore, our findings help identify the elements of specific lifestyles that are effective for the prevention of depression symptoms and offer useful reference for decision makers to start relevant preventive programs and promote them to the public.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (number 82273650).

Data Availability

The data sets generated and analyzed during this study are available from the National Health and Nutrition Examination Survey (NHANES) databases 2007-2018 [60]. The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

JL conducted the database search, screened and extracted data for the meta-analysis, prepared extracted data for the procedures, and had primary responsibility in writing this paper. JL and Shan H performed the statistical analysis, interpretation of data, and drafted the initial manuscript. JL, NJ, AK, Yican C, ML, YP, Shaoyi H, XP, and YZ selected the articles, extracted the data, analyzed the data, and contributed to the discussion and editing. Yajun C supervised data collection and critically edited the final manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Adjusted weighted characteristics of 19,295 US adults ≥20 years, stratified by lifestyle groups, the National Health and Nutrition Examination Survey 2007 to 2018.

[DOCX File , 32 KB - publichealth_v9i1e45776_app1.docx]

Multimedia Appendix 2

Characteristics of 19,295 US adults ≥20 years, stratified by depression symptom status, the National Health and Nutrition Examination Survey 2007 to 2018.

[DOCX File , 30 KB - publichealth_v9i1e45776_app2.docx]

Multimedia Appendix 3

Characteristics of 19,295 US adults ≥20 years, stratified by survey cycle, the National Health and Nutrition Examination Survey 2007 to 2018.

[DOCX File , 32 KB - publichealth_v9i1e45776_app3.docx]

Multimedia Appendix 4

Association between lifestyle and risk of depression symptoms with the exclusion of extreme poverty-to-income ratio (PIR), BMI values, and adjustment for alcohol use, total energy intake, and use of antidepressant or anxiolytic medications.

[DOCX File , 20 KB - publichealth_v9i1e45776_app4.docx]

References

1. Goodwin R, Dierker L, Wu M, Galea S, Hoven C, Weinberger A. Trends in U.S. Depression prevalence from 2015 to 2020: the widening treatment gap. Am J Prev Med 2022 Nov;63(5):726-733 [FREE Full text] [doi: 10.1016/j.amepre.2022.05.014] [Medline: 36272761]

2. Daly M, Sutin A, Robinson E. Depression reported by US adults in 2017-2018 and March and April 2020. J Affect Disord 2021 Jan 01;278:131-135 [FREE Full text] [doi: 10.1016/j.jad.2020.09.065] [Medline: 32956962]
3. Carter T, Guo B, Turner D, Morres I, Khalil E, Brighton E, et al. Preferred intensity exercise for adolescents receiving treatment for depression: a pragmatic randomised controlled trial. BMC Psychiatry 2015 Oct 14;15:247 [FREE Full text] [doi: 10.1186/s12888-015-0638-z] [Medline: 26467764]
4. Perez-Cornago A, Sanchez-Villegas A, Bes-Rastrollo M, Gea A, Molero P, Lahortiga-Ramos F, et al. Relationship between adherence to Dietary Approaches to Stop Hypertension (DASH) diet indices and incidence of depression during up to 8 years of follow-up. Public Health Nutr 2017 Sep;20(13):2383-2392. [doi: 10.1017/S1368980016001531] [Medline: 27335121]
5. Gómez-Gómez I, Bellón J, Resurrección DM, Cuijpers P, Moreno-Peral P, Rigabert A, et al. Effectiveness of universal multiple-risk lifestyle interventions in reducing depressive symptoms: systematic review and meta-analysis. Prev Med 2020 May;134:106067. [doi: 10.1016/j.ypmed.2020.106067] [Medline: 32194097]
6. Bernaards CM, Jans MP, van den Heuvel SG, Hendriksen IJ, Houtman IL, Bongers PM. Can strenuous leisure time physical activity prevent psychological complaints in a working population? Occup Environ Med 2006 Jan;63(1):10-16 [FREE Full text] [doi: 10.1136/oem.2004.017541] [Medline: 16361400]
7. Hamer M, Molloy G, de Oliveira C, Demakakos P. Leisure time physical activity, risk of depressive symptoms, and inflammatory mediators: the English Longitudinal Study of Ageing. Psychoneuroendocrinology 2009 Aug;34(7):1050-1055. [doi: 10.1016/j.psyneuen.2009.02.004] [Medline: 19272716]
8. Ku P, Fox K, Chen LJ. Physical activity and depressive symptoms in Taiwanese older adults: a seven-year follow-up study. Prev Med 2009 Mar;48(3):250-255. [doi: 10.1016/j.ypmed.2009.01.006] [Medline: 19297687]
9. Bakian A, Huber R, Scholl L, Renshaw P, Kondo D. Dietary creatine intake and depression risk among U.S. adults. Transl Psychiatry 2020 Feb 03;10(1):52 [FREE Full text] [doi: 10.1038/s41398-020-0741-x] [Medline: 32066709]
10. Lai J, Hiles S, Bisquera A, Hure A, McEvoy M, Attia J. A systematic review and meta-analysis of dietary patterns and depression in community-dwelling adults. Am J Clin Nutr 2014 Jan;99(1):181-197 [FREE Full text] [doi: 10.3945/ajcn.113.069880] [Medline: 24196402]
11. Cuenca-García M, Ortega F, Ruiz J, González-Gross M, Labayen I, Jago R, HELENA Study Group. Combined influence of healthy diet and active lifestyle on cardiovascular disease risk factors in adolescents. Scand J Med Sci Sports 2014 Jun;24(3):553-562. [doi: 10.1111/sms.12022] [Medline: 23237548]
12. An R. Diet quality and physical activity in relation to childhood obesity. Int J Adolesc Med Health 2017 Apr 01;29(2):0045. [doi: 10.1515/ijamh-2015-0045] [Medline: 26351906]
13. Loprinzi PD, Smit E, Mahoney S. Physical activity and dietary behavior in US adults and their combined influence on health. Mayo Clin Proc 2014 Feb;89(2):190-198. [doi: 10.1016/j.mayocp.2013.09.018] [Medline: 24485132]
14. Liu J, Rehm C, Onopa J, Mozaffarian D. Trends in diet quality among youth in the United States, 1999-2016. JAMA 2020 Mar 24;323(12):1161-1174 [FREE Full text] [doi: 10.1001/jama.2020.0878] [Medline: 32207798]
15. Hiza H, Casavale K, Guenther P, Davis CA. Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. J Acad Nutr Diet 2013 Feb;113(2):297-306. [doi: 10.1016/j.jand.2012.08.011] [Medline: 23168270]
16. Kirkpatrick S, Dodd K, Reedy J, Krebs-Smith SM. Income and race/ethnicity are associated with adherence to food-based dietary guidance among US adults and children. J Acad Nutr Diet 2012 May;112(5):624-35.e6 [FREE Full text] [doi: 10.1016/j.jand.2011.11.012] [Medline: 22709767]
17. NCHS. National Health and Nutrition Examination Survey: MEC in-person dietary interviewers procedures manual. Atlanta: NCHS; Jul 1, 2009.
18. World Medical Association. Declaration of Helsinki ethical principles for medical research involving human subjects. J Int Biotechnol Law 2009;6(2):73-76. [doi: 10.1515/JIBL.2009.12]
19. NCHS Ethics Review Board (ERB) approval*. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm [accessed 2023-04-29]
20. National health and nutrition examination survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2023-04-29]
21. NCHS Ethics Review Board (ERB) approval. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm#print [accessed 2018-04-01]
22. Krebs-Smith S, Pannucci T, Subar A, Kirkpatrick S, Lerman J, Tooze J, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet 2018 Sep;118(9):1591-1602 [FREE Full text] [doi: 10.1016/j.jand.2018.05.021] [Medline: 30146071]
23. Ford E, Bergmann MM, Boeing H, Li C, Capewell S. Healthy lifestyle behaviors and all-cause mortality among adults in the United States. Prev Med 2012 Jul;55(1):23-27 [FREE Full text] [doi: 10.1016/j.ypmed.2012.04.016] [Medline: 22564893]
24. Organization. WH: Global Physical Activity Questionnaire (GPAQ) Analysis Guide. Global Physical Activity Questionnaire (GPAQ) Analysis Guide. URL: https://www whoint/ncds/surveillance/steps/resources/GPAQ_Analysis_Guidepdf [accessed 2021-10-11]
25. National Health and Nutrition Examination Survey. URL: https://wwwncdcgov/nchs/nhanes [accessed 2021-10-12]
26. Chen L, Cai M, Li H, Wang X, Tian F, Wu Y, et al. Risk/benefit tradeoff of habitual physical activity and air pollution on chronic pulmonary obstructive disease: findings from a large prospective cohort study. BMC Med 2022 Feb 28;20(1):70 [FREE Full text] [doi: 10.1186/s12916-022-02274-8] [Medline: 35220974]

27. Piercy K, Troiano R, Ballard R, Carlson S, Fulton J, Galuska D, et al. The physical activity guidelines for Americans. JAMA 2018 Nov 20;320(19):2020-2028 [FREE Full text] [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
28. Kroenke K, Spitzer R, Williams J, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010;32(4):345-359. [doi: 10.1016/j.genhosppsych.2010.03.006] [Medline: 20633738]
29. Kroenke K, Spitzer R, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
30. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals 2002 Sep;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
31. Kessler R, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003 Jun 18;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095] [Medline: 12813115]
32. National Health and Nutrition Examination Survey Data. National Center for Health Statistics. URL: http://wwwcdcgov/nchs/about/major/nhanes/nhanes2003-2004/nhanes03_04htm [accessed 2007-02-09]
33. National Health and Nutrition Examination Survey: Specifying Weight Parameters. 2013. URL: http://wwwcdcgov/nchs/tutorials/NHANES/SurveyDesign/Weighting/Task2htm [accessed 2013-05-10]
34. Knol M, VanderWeele T, Groenwold R, Klungel O, Rovers M, Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. Eur J Epidemiol 2011 Jun;26(6):433-438 [FREE Full text] [doi: 10.1007/s10654-011-9554-9] [Medline: 21344323]
35. Zhang Q, Zhang X, Li S, Liu H, Liu L, Huang Q, et al. Joint effect of urinary arsenic species and serum one-carbon metabolism nutrients on gestational diabetes mellitus: a cross-sectional study of Chinese pregnant women. Environ Int 2021 Nov;156:106741 [FREE Full text] [doi: 10.1016/j.envint.2021.106741] [Medline: 34217037]
36. Andersson T, Alfredsson L, Källberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. Eur J Epidemiol 2005;20(7):575-579. [doi: 10.1007/s10654-005-7835-x] [Medline: 16119429]
37. Module 8: Age Standardization and Population Estimates. Task 1a: How to Generate Age-Adjusted Prevalence Rates in SUDAAN. URL: https://wwwncdcgov/nchs/nhanes/tutorials/samplecodeaspx [accessed 2022-10-01]
38. Hughes CW, Barnes S, Barnes C, Defina LF, Nakonezny P, Emslie GJ. Depressed Adolescents Treated with Exercise (DATE): a pilot randomized controlled trial to test feasibility and establish preliminary effect sizes. Ment Health Phys Act 2013 Jun;6(2):62264 [FREE Full text] [doi: 10.1016/j.mhpa.2013.06.006] [Medline: 24244220]
39. Global status report on physical activity 2022. URL: https://wwwwhoint/publications/i/item/9789240059153 [accessed 2022-10-01]
40. Pearce M, Garcia L, Abbas A, Strain T, Schuch F, Golubic R, et al. Association between physical activity and risk of depression: a systematic review and meta-analysis. JAMA Psychiatry 2022 Jun 01;79(6):550-559 [FREE Full text] [doi: 10.1001/jamapsychiatry.2022.0609] [Medline: 35416941]
41. Schuch F, Vancampfort D, Firth J, Rosenbaum S, Ward P, Silva E, et al. Physical activity and incident depression: a meta-analysis of prospective cohort studies. Am J Psychiatry 2018 Jul 01;175(7):631-648. [doi: 10.1176/appi.ajp.2018.17111194] [Medline: 29690792]
42. Brellenthin A, Crombie K, Hillard C, Koltyn KF. Endocannabinoid and mood responses to exercise in adults with varying activity levels. Med Sci Sports Exerc 2017 Aug;49(8):1688-1696. [doi: 10.1249/MSS.0000000000001276] [Medline: 28319590]
43. Kandola A, Ashdown-Franks G, Hendrikse J, Sabiston C, Stubbs B. Physical activity and depression: towards understanding the antidepressant mechanisms of physical activity. Neurosci Biobehav Rev 2019 Dec;107:525-539. [doi: 10.1016/j.neubiorev.2019.09.040] [Medline: 31586447]
44. Choi K, Stein M, Nishimi K, Ge T, Coleman J, Chen C, 23andMe Research Team, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, et al. An exposure-wide and mendelian randomization approach to identifying modifiable factors for the prevention of depression. Am J Psychiatry 2020 Oct 01;177(10):944-954 [FREE Full text] [doi: 10.1176/appi.ajp.2020.19111158] [Medline: 32791893]
45. Choi K, Chen C, Stein M, Klimentidis Y, Wang M, Koenen K, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Assessment of bidirectional relationships between physical activity and depression among adults: a 2-sample mendelian randomization study. JAMA Psychiatry 2019 Apr 01;76(4):399-408 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.4175] [Medline: 30673066]
46. Arvaniti F, Panagiotakos DB. Healthy indexes in public health practice and research: a review. Crit Rev Food Sci Nutr 2008 Apr;48(4):317-327. [doi: 10.1080/10408390701326268] [Medline: 18409114]
47. Fransen H, Ocke M. Indices of diet quality. Curr Opin Clin Nutr Metab Care 2008 Jul 1;11(5):559-565. [doi: 10.1097/mco.0b013e32830a49db]
48. Wang K, Zhao Y, Nie J, Xu H, Yu C, Wang S. Higher HEI-2015 score is associated with reduced risk of depression: result from NHANES 2005-2016. Nutrients 2021 Jan 25;13(2):348 [FREE Full text] [doi: 10.3390/nu13020348] [Medline: 33503826]

49. Fadnes L, Økland JM, Haaland Ø, Johansson KA. Estimating impact of food choices on life expectancy: a modeling study. PLoS Med 2022 Feb;19(2):e1003889 [FREE Full text] [doi: 10.1371/journal.pmed.1003889] [Medline: 35134067]
50. Druart C, Alligier M, Salazar N, Neyrinck A, Delzenne NM. Modulation of the gut microbiota by nutrients with prebiotic and probiotic properties. Adv Nutr 2014 Sep;5(5):624S-633S [FREE Full text] [doi: 10.3945/an.114.005835] [Medline: 25225347]
51. Delzenne NM, Neyrinck AM, Cani PD. Modulation of the gut microbiota by nutrients with prebiotic properties: consequences for host health in the context of obesity and metabolic syndrome. Microb Cell Fact 2011 Aug 30;10 Suppl 1(Suppl 1):S10 [FREE Full text] [doi: 10.1186/1475-2859-10-S1-S10] [Medline: 21995448]
52. Berk M, Williams LJ, Jacka FN, O'Neil A, Pasco JA, Moylan S, et al. So depression is an inflammatory disease, but where does the inflammation come from? BMC Med 2013 Sep 12;11:200 [FREE Full text] [doi: 10.1186/1741-7015-11-200] [Medline: 24228900]
53. Opie R, Itsiopoulos C, Parletta N, Sanchez-Villegas A, Akbaraly T, Ruusunen A, et al. Dietary recommendations for the prevention of depression. Nutr Neurosci 2017 Apr;20(3):161-171. [doi: 10.1179/1476830515Y.0000000043] [Medline: 26317148]
54. Gianfredi V, Dinu M, Nucci D, Eussen S, Amerio A, Schram M, et al. Association between dietary patterns and depression: an umbrella review of meta-analyses of observational studies and intervention trials. Nutr Rev 2023 Feb 10;81(3):346-359. [doi: 10.1093/nutrit/nuac058] [Medline: 36240456]
55. Egger G, Binns A, Rossner SR. The emergence of "lifestyle medicine" as a structured approach for management of chronic disease. Med J Aust 2009 Feb 02;190(3):143-145. [doi: 10.5694/j.1326-5377.2009.tb02317.x] [Medline: 19203313]
56. Walsh R. Lifestyle and mental health. Am Psychol 2011 Oct;66(7):579-592 [FREE Full text] [doi: 10.1037/a0021769] [Medline: 21244124]
57. Young C, Trapani K, Dawson S, O'Neil A, Kay-Lambkin F, Berk M, et al. Efficacy of online lifestyle interventions targeting lifestyle behaviour change in depressed populations: a systematic review. Aust N Z J Psychiatry 2018 Sep;52(9):834-846. [doi: 10.1177/0004867418788659] [Medline: 30052063]
58. Wong V, Ho F, Shi N, Sarris J, Chung KF, Yeung WF. Lifestyle medicine for depression: a meta-analysis of randomized controlled trials. J Affect Disord 2021 Apr 01;284:203-216. [doi: 10.1016/j.jad.2021.02.012] [Medline: 33609955]
59. Sarris J, O’Neil A, Coulson CE, Schweitzer I, Berk M. Lifestyle medicine for depression. BMC Psychiatry 2014 Apr 10;14:107. [doi: 10.1186/1471-244x-14-107]
60. National Health and Nutrition Examination Survey (NHANES). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/participant.htm [accessed 2023-04-29]

Abbreviations

AOR: adjusted odds ratio
AP: Attributable Proportion due to interaction
DASH: Dietary Approaches to Stop Hypertension
DQ: dietary quality
HEI: Healthy Eating Index
LM: Lifestyle Medicine
MET: metabolic equivalent
NHANES: National Health and Nutrition Examination Survey
PA: physical activity
PHQ-9: 9-item Patient Health Questionnaire
PIR: poverty-to-income ratio
RERI: Relative Excess Risk due to Interaction
SB: sedentary behavior
SI: Synergy Index
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 16.01.23; peer-reviewed by B Xi, J Ran; comments to author 09.02.23; revised version received 19.02.23; accepted 01.03.23; published 10.05.23.

Please cite as:
Liang J, Huang S, Jiang N, Kakaer A, Chen Y, Liu M, Pu Y, Huang S, Pu X, Zhao Y, Chen Y
Association Between Joint Physical Activity and Dietary Quality and Lower Risk of Depression Symptoms in US Adults: Cross-sectional NHANES Study
JMIR Public Health Surveill 2023;9:e45776
URL: https://publichealth.jmir.org/2023/1/e45776
doi:10.2196/45776
PMID:37163324

Original Paper

Development of a Prediction Model for Healthy Life Years Without Activity Limitation: National Cross-sectional Study

Masahiro Nishi[1,2], MD, PhD; Reo Nagamitsu[2,3], MD; Satoaki Matoba[1], MD, PhD

[1]Department of Cardiovascular Medicine, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

[2]Department of Health and Welfare, Kyoto Prefectural Government, Kyoto, Japan

[3]Department of Epidemiology for Community Health and Medicine, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

Corresponding Author:
Masahiro Nishi, MD, PhD
Department of Cardiovascular Medicine
Graduate School of Medical Science
Kyoto Prefectural University of Medicine
465 Kajiicho, Hirokoji
Kawaramachi, Kamigyo
Kyoto, 602-8566
Japan
Phone: 81 75 251 5511
Email: nishim@koto.kpu-m.ac.jp

Abstract

Background: In some countries, including Japan—the leading country in terms of longevity, life expectancy has been increasing; meanwhile, healthy life years have not kept pace, necessitating an effective health policy to narrow the gap.

Objective: The aim of this study is to develop a prediction model for healthy life years without activity limitations and deploy the model in a health policy to prolong healthy life years.

Methods: The Comprehensive Survey of Living Conditions, a cross-sectional national survey of Japan, was conducted by the Japanese Ministry of Health, Labour and Welfare in 2013, 2016, and 2019. The data from 1,537,773 responders were used for modelling using machine learning. All participants were randomly split into training (n=1,383,995, 90%,) and test (n=153,778, 10%) subsets. Extreme gradient boosting classifier was implemented. Activity limitations were set as the target. Age, sex, and 40 types of diseases or injuries were included as features. Healthy life years without activity limitations were calculated by incorporating the predicted prevalence rate of activity limitations in a life table. For the wide utility of the model in individuals, we developed an application tool for the model.

Results: In the groups without (n=1,329,901) and with (n=207,872) activity limitations, the median age was 47 (IQR 30-64) and 69 (IQR 54-80) years, respectively (*P*<.001); female sex comprised 51.3% (n=681,794) in the group without activity limitations and 56.9% (n=118,339) in the group with activity limitations (*P*<.001). A total of 42 features were included in the feature set. Age had the highest impact on model accuracy, followed by depression or other mental diseases; back pain; bone fracture; other neurological disorders, pain, or paralysis; stroke, cerebral hemorrhage, or infarction; arthritis; Parkinson disease; dementia; and other injuries or burns. The model exhibited high performance with an area under the receiver operating characteristic curve of 0.846 (95% CI 0.842-0.849) with exact calibration for the average probability and fraction of positives. The prediction results were consistent with the observed values of healthy life years for both sexes in each year (range of difference between predictive and observed values: −0.89 to 0.16 in male and 0.61 to 1.23 in female respondents). We applied the prediction model to a regional health policy to prolong healthy life years by adjusting the representative predictors to a target prevalence rate. Additionally, we presented the health condition without activity limitations index, followed by the application development for individual health promotion.

Conclusions: The prediction model will enable national or regional governments to establish an effective health promotion policy for risk prevention at the population and individual levels to prolong healthy life years. Further investigation is needed to validate the model’s adaptability to various ethnicities and, in particular, to countries where the population exhibits a short life span.

(JMIR Public Health Surveill 2023;9:e46634) doi:10.2196/46634

KEYWORDS

healthy life years; machine learning; health policy; prediction model; health promotion; activity limitation; cross-sectional study; application tool; life expectancy

Introduction

Global public health, secure social systems, and advances in medical practice have contributed to the extension of life expectancy and healthy life years (referred to as the healthy life expectancy) of humans worldwide. With the growing recognition of the importance of taking into account the state of being alive or quality of life, "healthy life years" has come to be focused on as an integrated health indicator that combines not only mortality data but also data on the state of being alive. Healthy life years are not merely defined as life without disability or illness but include a holistic life of well-being. Although life expectancy has been increasing, healthy life years have not yet been kept pace, necessitating an effective health policy to narrow the gap [1,2].

There have been several measures to estimate healthy life years, which are used to evaluate national or regional health status. The World Health Organization has used the health-adjusted life expectancy, a measure of healthy life years based on a specialized health survey producing disability weight on various diseases, injuries, and sequelae [3-5]. In some countries, including Japan—the leading country in terms of longevity, a national survey is conducted to investigate healthy life years using a questionnaire for the presence of activity limitations.

To date, some determinants of healthy life years and the relevant activity limitations have been identified. Typical risk factors, such as obesity, hypertension, hyperglycemia, smoking, and excessive alcohol consumption, are negatively associated with a healthy life [6,7]. Physical activity and exercise [8-11] as well as a diverse healthy diet [12,13] are expected to prolong healthy life years. In addition to fatal diseases, several nonfatal conditions, such as mental health disorders, musculoskeletal problems, and ophthalmic diseases, are also crucial determinants of activity limitations [10,14,15].

Despite the increasing interest in a healthy life for public health campaign and individual health awareness, a prediction model of healthy life years with integrated features has not been reported. In this study, we sought to develop a prediction model for healthy life years without activity limitations using machine learning and to deploy the model to a health policy in prolonging healthy life years at the population and individual levels.

Methods

Data Description

The Comprehensive Survey of Living Conditions, a cross-sectional national survey, is conducted every 3 years by the Japanese Ministry of Health, Labour and Welfare to investigate the fundamental aspects of the nation's livelihood, such as health, medical care, welfare, pension, and income [16]. In health questionnaire of the Comprehensive Survey of Living Conditions, subjective symptoms, health problems in daily life, disease or injury under treatment, subjective health assessment, worries and stress, mental state, and receiving rate of health check-ups are surveyed. The response rate in 2019 was 72.5% according to the Japanese Ministry of Health, Labour and Welfare. Among the data in 2013, 2016, and 2019, data from 1,537,773 responders were used for the analysis. The activity limitations of responders were evaluated using responses to the questions, "Do you have any health problem which limits your daily activity?" Respondents who answered "yes" were categorized as the "activity limitations" group, and those who answered "no" were categorized into the "no activity limitation" group. Activity limitations, age, sex, and the 40 types of diseases or injuries under treatment were included in the analysis.

Model Description

The activity limitations, which were classified as binary, were set as model target; the "activity limitations" group was classified as 1, and the "no activity limitation" group as 0. Age, sex, and the 40 types of diseases or injuries under treatment were included as features. We implemented the extreme gradient boosting (XGB) classifier—a widely used supervised tree-based model, which uses labeled data sets to train a model [17-19]—for the binary classification of activity limitations using the scikit-learn 1.1.1 package [20] in Python 3.10.4. Using the "*train_test_split*" function, we randomly split the data set into training (n=1,383,995, 90%) and test (n=153,778, 10%) subsets [21,22]. We selected the best feature set using the recursive feature elimination function in the training data set, with the XGB classifier as an estimator. To determine the best hyperparameter values providing the highest model performance, we used *the GridSearchCV* function for training subset with the five-fold cross validation. A total of 480 XGB models were compared using different combinations of hyperparameters. Consequently, we selected a model with hyperparameters (n_estimators=200, max_depth=9, eta=0.1, min_child_weight=2, max_delta_step=5, and subsample=0.5) yielding the highest area under the receiver operating characteristic curve (AUROC). Finally, the model performance was evaluated for the test subsets.

The impact of the features on the model accuracy was estimated by permutation importance, which is defined as difference of error when a feature value is randomly shuffled, assigning 1.0 to the highest impact. The SHapley Additive exPlanations (SHAP) value, which explains a feature contribution on model output in each sample, was used to evaluate the effect of features on the model output [23]. We calculated the area under the curve and log loss, a measure of how close predictive probability is to observed value, as the model accuracy metrics for the XGB classifier, random forest, and logistic regression. Other metrics were calculated based on the confusion matrix. We used *calibration_curve* function for the model calibration between predictive probability and fraction of positives, dividing samples into 10 bins according to predictive probability. Cost-benefit was calculated to determine the optimal cut-off of prediction by the sum of true positive and true negative as 0, false positive as +1, and false negative as −1 in the confusion matrix. Healthy life years without activity limitations of female respondents in Kyoto Prefecture were predicted using the prediction model with the original and target prevalence rates of the representative diseases, which are much higher than the mean prevalence rates in the whole country. Random sampling was performed for populations with each disease to achieve target prevalence rate. The web-based application tool for the model was developed on a web application platform based on the programming code [24].

Statistical Analysis

General descriptive statistics were performed in R (version 4.2.0; R Core Team) [25]. Categorical values are represented as numbers (along with percentages), and numerical values are represented as medians (IQRs). The chi-square test was used for categorical values, and the Mann-Whitney *U* test was used for continuous values with a nonparametric distribution. *P*<.01 was considered statistically significant. The health condition without activity limitation (HCAL) index was calculated by subtracting the percentage of predictive probability for activity limitations from 100. Curve fitting was performed using third-order polynomial regression. Healthy life years without activity limitation were calculated using Sullivan's method, which is widely used to calculate life expectancy based on age-stratified mortality rate and life table, incorporating the prevalence rate of activity limitations to a life table in Japan [26,27].

Ethical Considerations

The study was approved by the ethics committee of Kyoto Prefectural University of Medicine (approval number ERB-C-2496). This study conformed to the principles outlined in the Declaration of Helsinki. Since this study used only existing national survey data, opt-out decline was adopted for participants on the university website instead of informed consent. The study data are anonymous. There was no compensation for participants.

Results

Participant Characteristics

The characteristics of participants (N=1,537,773) are described and stratified according to the presence of activity limitations (Table 1). In the groups without (n=1,329,901) and with activity limitations (n=207,872), the median age was 47 (IQR 30-64) and 69 (IQR 54-80) years, respectively (*P*<.001), and female sex comprised 51.3% (n=681,79) of the participants in the group without activity limitations (vs n=118,339, 56.9%; *P*<.001). Diseases under treatment, except for infertility, were also significantly different between the two groups (depression or other mental disease: n=13,727, 1% vs n=15,347, 7.4%; dementia: n=2420, 0.2% vs n=8667, 4.2%; stroke, cerebral hemorrhage, or infarction: n=8452, 0.6% vs n=10,818, 5.2%; angina or myocardial infarction: n=16,467, 1.2% vs n=13,043, 6.3%; rheumatoid arthritis: n=5153, 0.4% vs n=6239, 3%; arthritis: n=15,682, 1.2% vs n=19,753, 9.5%; back pain: n=42,856, 3.2% vs n=37,980, 18.3%; kidney disease: n=7415, 0.6% vs n=7866, 3.8%; malignant neoplasm or cancer: n=7594, 0.6% vs n=6249, 3%). All participants (N=1,537,773) were randomly split into training (n=1,383,995, 90%) and test (n=153,778, 10%) subsets with similar characteristics (Table S1 in Multimedia Appendix 1).

Table 1. Characteristics of participants stratified by presence of activity limitations.

Characteristics	All (N=1,537,773)	Without activity limitation (n=1,329,901)	With activity limitation (n=207,872)	*P* value
Age (years), median (IQR)	50 (32-67)	47 (30-64)	69 (54-80)	<.001
Sex (female), n (%)	800,133 (52)	681,794 (51.3)	118,339 (56.9)	<.001
Diabetes, n (%)	77,672 (5.1)	53,671 (4)	24,001 (11.5)	<.001
Thyroid disease, n (%)	19,811 (1.3)	14,360 (1.1)	5451 (2.6)	<.001
Depression or other mental disease, n (%)	29,074 (1.9)	13,727 (1)	15,347 (7.4)	<.001
Dementia, n (%)	11,087 (0.7)	2420 (0.2)	8667 (4.2)	<.001
Parkinson disease, n (%)	3194 (0.2)	533 (0)	2661 (1.3)	<.001
Other neurological disorders, pain, or paralysis, n (%)	11,028 (0.7)	4311 (0.3)	6717 (3.2)	<.001
Eye disease, n (%)	83,577 (5.4)	52,941 (4)	30,636 (14.7)	<.001
Ear disease, n (%)	16,411 (1.1)	9354 (0.7)	7057 (3.4)	<.001
Stroke, cerebral hemorrhage, or infarction, n (%)	19,270 (1.3)	8452 (0.6)	10,818 (5.2)	<.001
Angina and myocardial infarction, n (%)	29,510 (1.9)	16,467 (1.2)	13,043 (6.3)	<.001
Other cardiovascular disease, n (%)	28,703 (1.9)	15,653 (1.2)	13,050 (6.3)	<.001
Acute nasopharyngitis and common cold, n (%)	5125 (0.3)	3549 (0.3)	1576 (0.8)	<.001
Infertility, n (%)	1536 (0.1)	1359 (0.1)	177 (0.1)	.02
Dental disease, n (%)	80,560 (5.2)	63,668 (4.8)	16,892 (8.1)	<.001
Gout, n (%)	15,396 (1)	12,090 (0.9)	3306 (1.6)	<.001
Obesity, n (%)	8038 (0.5)	5013 (0.4)	3025 (1.5)	<.001
Dyslipidemia, n (%)	81,338 (5.3)	63,404 (4.8)	17,934 (8.6)	<.001
Hypertension, n (%)	206,103 (13.4)	153,500 (11.5)	52,603 (25.3)	<.001
Allergic rhinitis, n (%)	32,310 (2.1)	24,472 (1.8)	7838 (3.8)	<.001
Chronic obstructive pulmonary disease, n (%)	2250 (0.1)	813 (0.1)	1437 (0.7)	<.001
Asthma, n (%)	19,022 (1.2)	13,149 (1)	5873 (2.8)	<.001
Other respiratory disease, n (%)	15,134 (1)	8517 (0.6)	6617 (3.2)	<.001
Stomach or duodenum disease, n (%)	26,285 (1.7)	17,048 (1.3)	9237 (4.4)	<.001
Liver or gallbladder disease, n (%)	14,624 (1)	9283 (0.7)	5341 (2.6)	<.001
Other digestive disease, n (%)	18,656 (1.2)	11,028 (0.8)	7628 (3.7)	<.001
Atopic dermatitis, n (%)	14,353 (0.9)	11,553 (0.9)	2800 (1.3)	<.001
Other skin disease, n (%)	29,205 (1.9)	20,475 (1.5)	8730 (4.2)	<.001
Rheumatoid arthritis, n (%)	11,392 (0.7)	5153 (0.4)	6239 (3)	<.001
Arthritis, n (%)	35,435 (2.3)	15,682 (1.2)	19,753 (9.5)	<.001
Stiff shoulder, n (%)	43,474 (2.8)	28,093 (2.1)	15,381 (7.4)	<.001
Back pain, n (%)	80,836 (5.3)	42,856 (3.2)	37,980 (18.3)	<.001
Osteoporosis, n (%)	28,790 (1.9)	14,606 (1.1)	14,184 (6.8)	<.001
Kidney disease, n (%)	15,281 (1)	7415 (0.6)	7866 (3.8)	<.001
Prostatic hypertrophy, n (%)	19,932 (1.3)	12,293 (0.9)	7639 (3.7)	<.001
Menopausal or postmenopausal disorder, n (%)	3041 (0.2)	1980 (0.1)	1061 (0.5)	<.001
Bone fracture, n (%)	10,464 (0.7)	3345 (0.3)	7119 (3.4)	<.001
Other injury or burns, n (%)	10,230 (0.7)	5451 (0.4)	4779 (2.3)	<.001
Anemia or blood disease, n (%)	10,660 (0.7)	5980 (0.4)	4680 (2.3)	<.001
Malignant neoplasm or cancer, n (%)	13,843 (0.9)	7594 (0.6)	6249 (3)	<.001

Characteristics	All (N=1,537,773)	Without activity limitation (n=1,329,901)	With activity limitation (n=207,872)	*P* value
Pregnancy, puerperium, threatened abortion, or placenta previa, n (%)	2198 (0.1)	1565 (0.1)	633 (0.3)	<.001

Model Performance Evaluation

To create the model feature set, the AUROC was compared for each feature number. A total of 42 features were included in the feature set because they had the highest AUROC (Figure S1 in Multimedia Appendix 1). The feature impact estimated by permutation importance showed that age had the highest impact on model accuracy, followed by depression or other mental diseases; back pain; bone fracture; other neurological disorders, pain, or paralysis; stroke, cerebral hemorrhage, or infarction; arthritis; Parkinson disease; dementia; and other injuries or burns (Figure 1).

The accuracy metrics were compared for some learners. We selected the XGB classifier as a learner because it exhibited a high AUORC and low log loss compared with random forest and logistic regression (Table S2 in Multimedia Appendix 1). The model performance was evaluated by depicting the receiver operating characteristic curve, and the AUROC was 0.846 (95% CI 0.842-0.849; Figure 2A). The calibration plot exhibited exact calibration for the average probability and fraction of positives (Figure 2B). We set the cut-off of 0.31 according to the lowest absolute value of the mean cost-benefit (Figure 2C). The prediction results were consistent with the observed values of healthy life years for both sexes in each year (range of difference between predictive and observed value: −0.89 to 0.16 in male and 0.61 to 1.23 in female respondents; Figure 2D). Thus, the developed model exhibited a markedly high performance in predicting healthy life years without activity limitations.

Figure 1. Feature impact estimated by permutation importance. Permutation importance was calculated for features using test data.

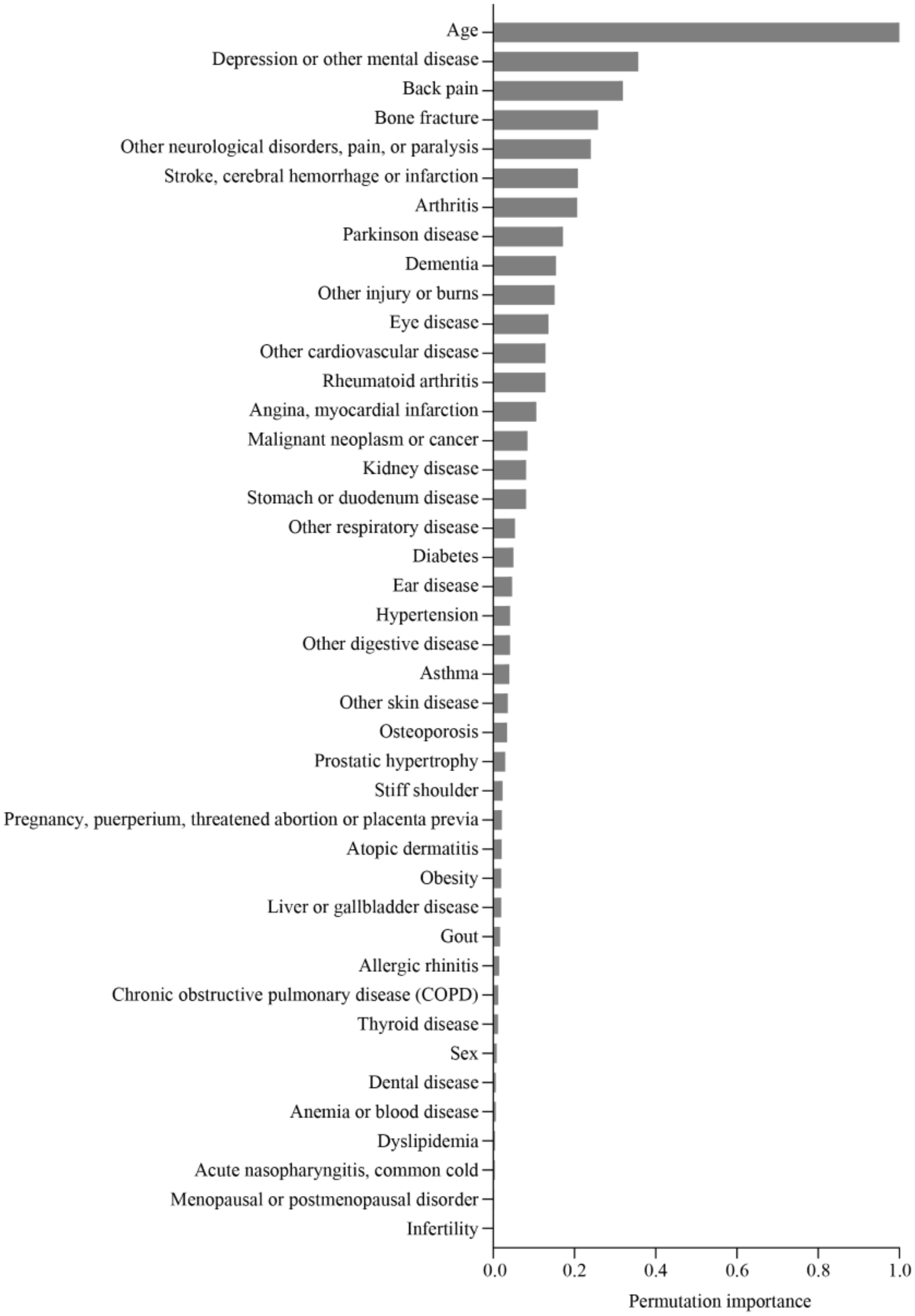
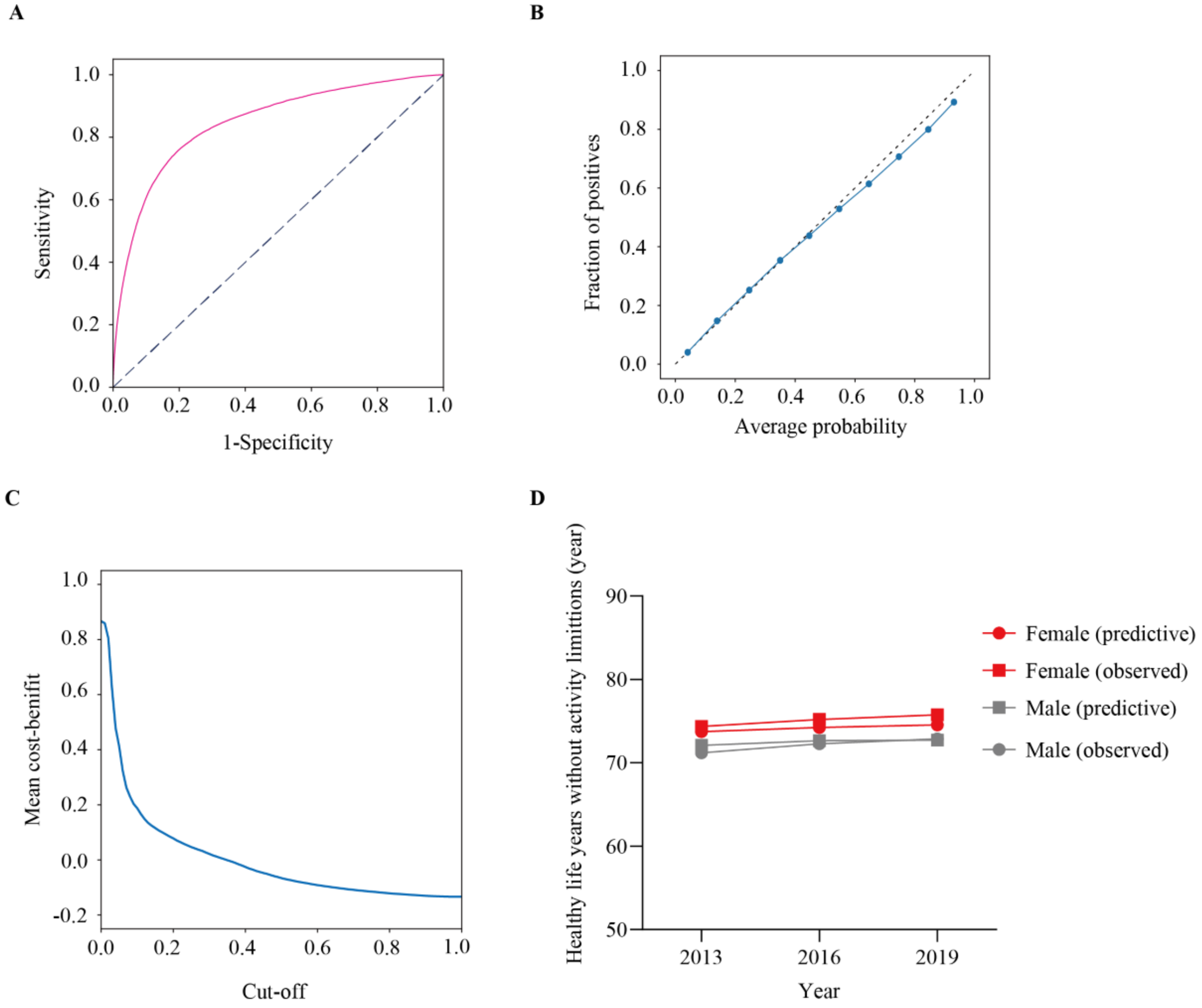


Figure 2. Evaluation of model performance. (A) Receiver operating characteristic curve for the model. Area under the receiver operating characteristic curve was 0.846 (95% CI 0.842-0.849). (B) Calibration plot for the model; samples were divided into 10 bins according to probability. (C) Mean cost-benefit curve. (D) Predictive and observed value of healthy life years for male and female respondents in each year.



Model Application for Population and Individual Health

For model application at the population level, we used the prediction model for a regional health policy regarding healthy life years (Table S3 in Multimedia Appendix 1). Healthy life years without activity limitations of females in Kyoto prefecture in Japan were predicted using the prediction model with the original and target prevalence rates of representative predictors, such as depression or other mental diseases, back pain, and stiff shoulder. As a result, healthy life years without activity limitations were simulated to be prolonged from 73.25 in the original to 73.81 in the target, a difference of 0.56. Herein, we demonstrate the use of a prediction model for a regional health policy to prolong healthy life years at the population level.

To enhance the interpretability of the feature effect on the model output, the SHAP value is displayed for each feature (Figure 3). This shows that age has the greatest effect on model output. HCAL index was decreased by aging (Figure 4). For the wide utility of the model in individuals, we developed a web-based application tool to display the HCAL index (Figure S2 in Multimedia Appendix 1).

Figure 3. Feature effect on model output. SHapley Additive exPlanations (SHAP) value was calculated for features using test data.

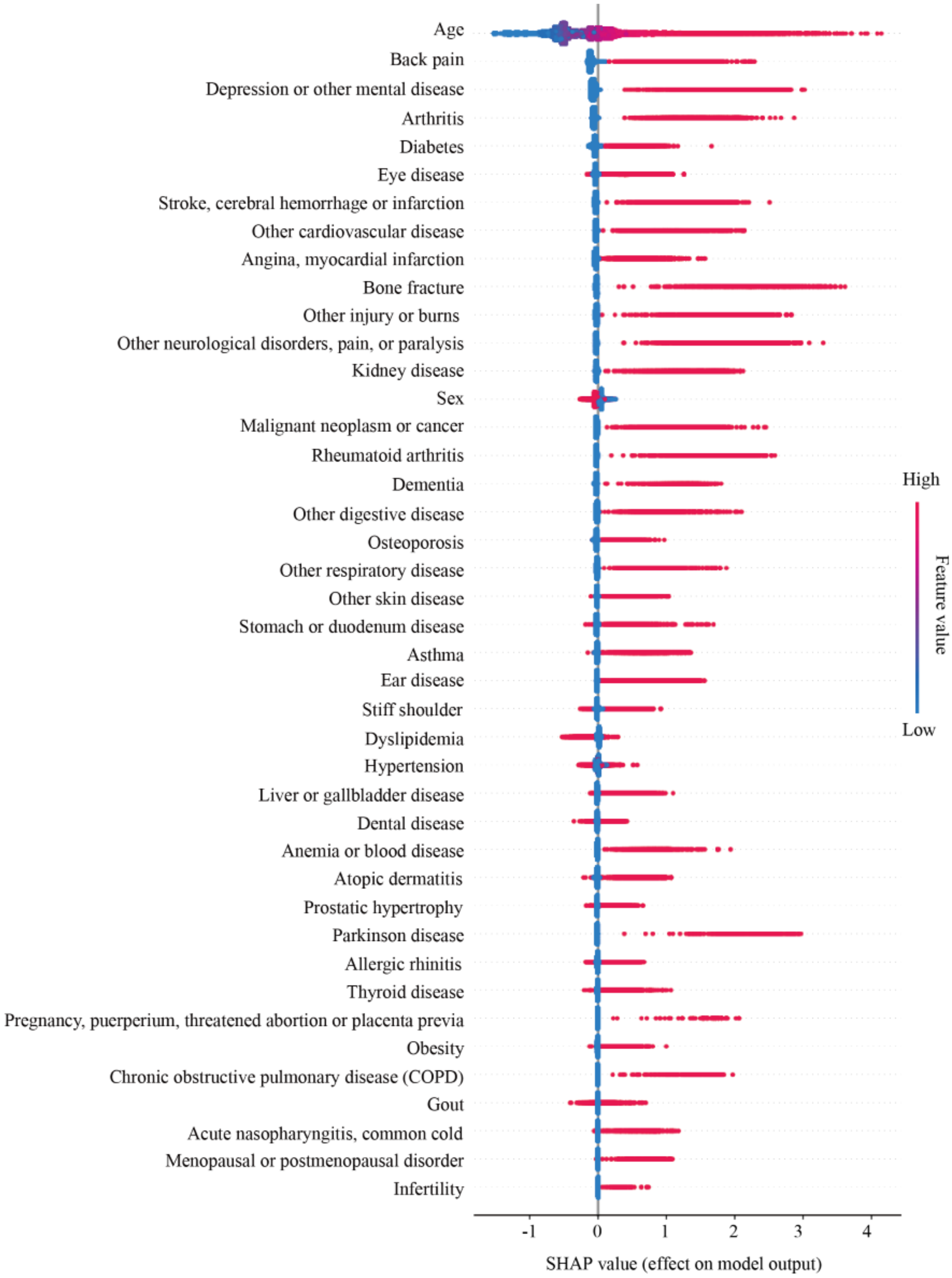
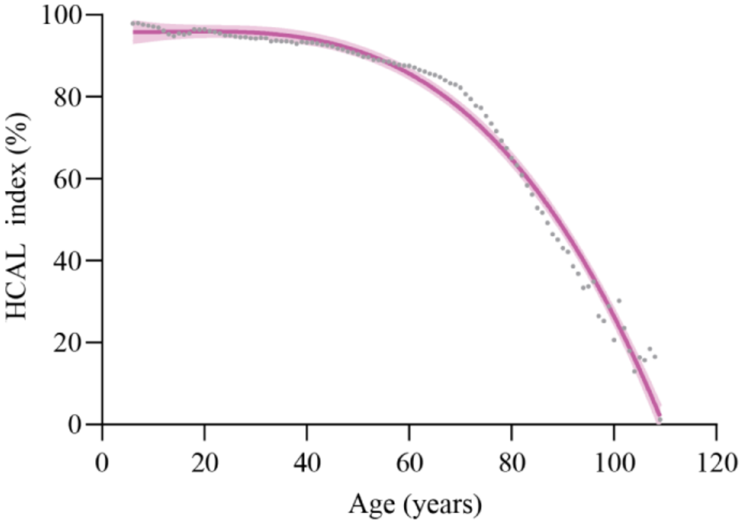


Figure 4. Health condition without activity limitation (HCAL) per age. The HCAL index indicates subtraction of the percentage of predictive probability from 100. Curve fitting was done using third-order polynomial regression. Error bar indicates 95% CI. Each dot represents mean HCAL index per age.



Discussion

Principal Findings

In this study, we developed a prediction model for healthy life years without activity limitations using machine learning by analyzing a cross-sectional national survey. The model exhibited markedly high performance with a high AUROC and subtle differences between the observed and predicted values of healthy life years without activity limitations. We applied the prediction model to a regional health policy to prolong healthy life years by adjusting the representative predictors to a target prevalence rate. Additionally, we presented the HCAL index, followed by the application development, for individual health promotion.

We estimated the feature impact on model accuracy by permutation importance and the effect on model output by the SHAP value. The impact of features on model accuracy showed that age had the highest impact, followed by depression or other mental disease; back pain; bone fracture; other neurological disorders, pain, or paralysis; stroke, cerebral hemorrhage, or infarction; arthritis; Parkinson disease; dementia; other injuries or burns. Interestingly, the high-impact features included several nonfatal conditions, such as mental disorders, musculoskeletal problems, and neurological diseases. Our findings were consistent with previous reports that suggest mental health disorders and musculoskeletal problems are crucial predictive factors for activity limitations [10,15]. Of mental disorders, schizophrenia and major depression have heavy disability weight according to the Global Burden of Disease Study [3]. A study using data from the Global Burden of Disease Study found no association between health output and common health system per prefecture in Japan [28]. These findings indicate that healthy life years without activity limitations largely rely on mental, musculoskeletal, or neurological causes rather than other typical lifestyle-related factors.

We leveraged machine learning to predict healthy life years without activity limitations. The presence of activity limitations assessed by a subjective questionnaire was used for the model target; nonetheless, healthy life years could be predicted accurately with the objective 42 features using machine learning. Machine learning facilitated model deployment by application development at the population and individual levels. Natural language processing has been applied to calculate the health-adjusted life expectancy using electronic medical records [29]. Machine learning combined with natural language processing for electronic medical records will provide a solution for global health issues regarding healthy life years.

We demonstrated the model application for population and individual health. Healthy life years without activity limitations of females in Kyoto prefecture were simulated using the prediction model with the original and target prevalence rates of representative predictors. Thus, the model could be used to present effective ways to prolong healthy life years for a regional health policy. Moreover, the application tool was developed using the model for wide utility in individual health promotion. The tool can be used in several situations, such as health check, patient education, and outpatient clinics. Our model was developed with machine learning and can be used for prediction of population-level healthy life years as well as individual health conditions, increasing its feasibility compared with other measures for healthy life years.

Limitations

This study had certain limitations, as it was based on a survey that included subjective data, and only data from Japan were used. Further investigation is needed to validate the model's adaptability to various ethnicities and, in particular, countries where the population has a short life span. For complexity of machine learning to explain and interpret, we used permutation importance and SHAP values for feature impact.

Conclusions

In conclusion, we developed a prediction model for healthy life years without activity limitations, using machine learning. The prediction model will enable the national or regional government to establish an effective health promotion policy for risk prevention at the population and individual levels to prolong healthy life years. It would be interesting to investigate the model's applicability to other countries and ethnicities.

Acknowledgments

We are grateful to Tomoyuki Yamamoto, Mika Yamashita, and Kumiko Katsuyama from the Department of Health and Welfare, Kyoto Prefectural Government, Japan.

This study was supported by Foundation for Total Health Promotion.

Data Availability

The data set generated and analyzed during this study is not publicly available due to restrictions imposed by Japanese Ministry of Health, Labour and Welfare (the data provider) but can be obtained by contacting the ministry based on reasonable request.

Authors' Contributions

MN was responsible for conception of the study. Formal analysis was performed by MN and RN. All authors participated in manuscript writing and approved the final manuscript. SM provided overall supervision. All the authors were responsible for the decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[PDF File (Adobe PDF File), 427 KB - publichealth_v9i1e46634_app1.pdf]

References

1. Salomon JA, Wang H, Freeman MK, Vos T, Flaxman AD, Lopez AD, et al. Healthy life expectancy for 187 countries, 1990–2010: a systematic analysis for the Global Burden Disease Study 2010. Lancet 2012 Dec;380(9859):2144-2162. [doi: 10.1016/s0140-6736(12)61690-0]
2. GBD 2017 DALYsHALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1859-1922 [FREE Full text] [doi: 10.1016/S0140-6736(18)32335-3] [Medline: 30415748]
3. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2129-2143. [doi: 10.1016/S0140-6736(12)61680-8] [Medline: 23245605]
4. Salomon JA, Haagsma JA, Davis A, de Noordhout CM, Polinder S, Havelaar AH, et al. Disability weights for the Global Burden of Disease 2013 study. Lancet Glob Health 2015 Nov;3(11):e712-e723 [FREE Full text] [doi: 10.1016/S2214-109X(15)00069-8] [Medline: 26475018]
5. Mortality and global health estimates. World Health Organization. URL: https://apps.who.int/gho/data/node.main.686?lang=en [accessed 2023-04-25]
6. Willcox BJ, He Q, Chen R, Yano K, Masaki KH, Grove JS, et al. Midlife risk factors and healthy survival in men. JAMA 2006 Nov 15;296(19):2343-2350. [doi: 10.1001/jama.296.19.2343] [Medline: 17105797]
7. Stenholm S, Head J, Kivimäki M, Kawachi I, Aalto V, Zins M, et al. Smoking, physical inactivity and obesity as predictors of healthy and disease-free life expectancy between ages 50 and 75: a multicohort study. Int J Epidemiol 2016 Aug;45(4):1260-1270 [FREE Full text] [doi: 10.1093/ije/dyw126] [Medline: 27488415]
8. Monma T, Takeda F, Noguchi H, Takahashi H, Tamiya N. The impact of leisure and social activities on activities of daily living of middle-aged adults: evidence from a national longitudinal survey in Japan. PLoS One 2016 Oct 27;11(10):e0165106 [FREE Full text] [doi: 10.1371/journal.pone.0165106] [Medline: 27788163]
9. Yamada M, Arai H. Self-management group exercise extends healthy life expectancy in frail community-dwelling older adults. Int J Environ Res Public Health 2017 May 15;14(5):531 [FREE Full text] [doi: 10.3390/ijerph14050531] [Medline: 28505140]
10. Myojin T, Ojima T, Kikuchi K, Okada E, Shibata Y, Nakamura M, et al. Orthopedic, ophthalmic, and psychiatric diseases primarily affect activity limitation for Japanese males and females: Based on the Comprehensive Survey of Living Conditions. J Epidemiol 2017 Feb;27(2):75-79 [FREE Full text] [doi: 10.1016/j.je.2016.09.007] [Medline: 28142015]
11. Monma T, Takeda F, Noguchi H, Takahashi H, Watanabe T, Tamiya N. Exercise or sports in midlife and healthy life expectancy: an ecological study in all prefectures in Japan. BMC Public Health 2019 Sep 09;19(1):1238 [FREE Full text] [doi: 10.1186/s12889-019-7570-y] [Medline: 31500600]

12. Imai T, Miyamoto K, Sezaki A, Kawase F, Shirai Y, Abe C, et al. Traditional Japanese diet score - association with obesity, incidence of ischemic heart disease, and healthy life expectancy in a global comparative study. J Nutr Health Aging 2019 Jun 24;23(8):717-724. [doi: 10.1007/s12603-019-1219-5] [Medline: 31560029]
13. Miyamoto K, Kawase F, Imai T, Sezaki A, Shimokata H. Dietary diversity and healthy life expectancy-an international comparative study. Eur J Clin Nutr 2019 Mar 13;73(3):395-400. [doi: 10.1038/s41430-018-0270-3] [Medline: 30104730]
14. Hashimoto S, Kawado M, Yamada H, Seko R, Murakami Y, Hayashi M, et al. Gains in disability-free life expectancy from elimination of diseases and injuries in Japan. J Epidemiol 2012;22(3):199-204 [FREE Full text] [doi: 10.2188/jea.je20110112] [Medline: 22343332]
15. Monma T, Takeda F, Noguchi H, Tamiya N. Age and sex differences of risk factors of activity limitations in Japanese older adults. Geriatr Gerontol Int 2016 Jun 04;16(6):670-678. [doi: 10.1111/ggi.12533] [Medline: 26044713]
16. Comprehensive Survey of Living Conditions 2019. Ministry of Health, Labour and Welfare, Japan. URL: https://www.mhlw.go.jp/toukei/list/20-21kekka.html [accessed 2023-04-25]
17. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: KDD '16; August 13-17; San Francisco, CA p. 13-17. [doi: 10.1145/2939672.2939785]
18. Yadaw AS, Li Y, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. The Lancet Digital Health 2020 Oct;2(10):e516-e525. [doi: 10.1016/s2589-7500(20)30217-x]
19. Wenzl FA, Kraler S, Ambler G, Weston C, Herzog SA, Räber L, et al. Sex-specific evaluation and redevelopment of the GRACE score in non-ST-segment elevation acute coronary syndromes in populations from the UK and Switzerland: a multinational analysis with external cohort validation. Lancet 2022 Sep;400(10354):744-756. [doi: 10.1016/s0140-6736(22)01483-0]
20. Fabian P, Gaël V, Alexandre G, Vincent M, Bertrand T, Olivier G. Machine learning in Python. JMLR 2011;12(85):2825-2830. [doi: 10.1201/9781420067194-20]
21. Nishi M, Uchino E, Okuno Y, Matoba S. Robust prognostic prediction model developed with integrated biological markers for acute myocardial infarction. PLoS One 2022;17(11):e0277260 [FREE Full text] [doi: 10.1371/journal.pone.0277260] [Medline: 36327332]
22. Kwong EW, Wu H, Pang GK. A prediction model of blood pressure for telemedicine. Health Informatics J 2018 Sep;24(3):227-244 [FREE Full text] [doi: 10.1177/1460458216663025] [Medline: 27496863]
23. Lundberg S, Lee S. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017(30):4768-4777.
24. AI Diagnostic Tool for Health Condition Without Activity Limitations. URL: https://self-check-health.herokuapp.com/ [accessed 2023-04-25]
25. The R Project for Statistical Computing. URL: https://www.R-project.org/ [accessed 2023-04-25]
26. Vital Statistics. Ministry of Health, Labour and Welfare JAPAN. URL: https://www.mhlw.go.jp/english/database/db-hw/index.html [accessed 2023-04-25]
27. Sullivan DF. A single index of mortality and morbidity. HSMHA Health Rep 1971 Apr;86(4):347-354 [FREE Full text] [Medline: 5554262]
28. Nomura S, Sakamoto H, Glenn S, Tsugawa Y, Abe SK, Rahman MM, et al. Population health and regional variations of disease burden in Japan, 1990-2015: a systematic subnational analysis for the Global Burden of Disease Study 2015. Lancet 2017 Sep 23;390(10101):1521-1538 [FREE Full text] [doi: 10.1016/S0140-6736(17)31544-1] [Medline: 28734670]
29. Ruan X, Li Y, Jin X, Deng P, Xu J, Li N, et al. Health-adjusted life expectancy (HALE) in Chongqing, China, 2017: an artificial intelligence and big data method estimating the burden of disease at city level. Lancet Reg Health West Pac 2021 Apr;9:100110 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100110] [Medline: 34379708]

Abbreviations

AUROC: area under the receiver operating characteristic curve
HCAL: health condition without activity limitation
SHAP: SHapley Additive exPlanations
XGB: extreme gradient boosting

Edited by A Mavragani, T Sanchez; submitted 20.02.23; peer-reviewed by J Zhu, Y Shao; comments to author 23.03.23; revised version received 27.03.23; accepted 18.04.23; published 17.05.23.

Please cite as:
Nishi M, Nagamitsu R, Matoba S
Development of a Prediction Model for Healthy Life Years Without Activity Limitation: National Cross-sectional Study
JMIR Public Health Surveill 2023;9:e46634
URL: https://publichealth.jmir.org/2023/1/e46634
doi:10.2196/46634
PMID:37195737

Original Paper

Disease Burden and the Accumulation of Multimorbidity of Noncommunicable Diseases in a Rural Population in Henan, China: Cross-sectional Study

Ying Chen[1], MS; Mingming Pan[1], PhD; Yaling He[1], PhD; Xiaokang Dong[1], PhD; Ze Hu[1], MS; Jian Hou[1], PhD; Yining Bao[2], PhD; Jing Yang[1], MS; Yinghao Yuchi[1], PhD; Ruiying Li[1], PhD; Linghui Zhu[1], MS; Ning Kang[1], MS; Wei Liao[1], PhD; Shuoyi Li[1], MS; Chongjian Wang[1*], PhD; Lei Zhang[2,3,4*], PhD

[1]Department of Epidemiology and Biostatistics, College of Public Health, Zhengzhou University, Zhengzhou, Henan, China

[2]China-Australia Joint Research Center for Infectious Diseases, School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi, China

[3]Artificial Intelligence and Modelling in Epidemiology Program, Melbourne Sexual Health Centre, Alfred Health, Melbourne, Australia

[4]Central Clinical School, Faculty of Medicine, Monash University, Melbourne, Australia

[*]these authors contributed equally

Corresponding Author:
Lei Zhang, PhD
China-Australia Joint Research Center for Infectious Diseases
School of Public Health
Xi'an Jiaotong University Health Science Center
No. 76, Yanta West Road
Xi'an, Shaanxi, 710061
China
Phone: 86 29 82655135
Email: lei.zhang1@monash.edu

Abstract

Background: Multimorbidity causes substantial disease and economic burdens on individuals and the health care system.

Objective: This study aimed to explore the disease burden of multimorbidity and the potential correlations among chronic noncommunicable diseases (NCDs) in a rural population in Henan, China.

Methods: A cross-sectional analysis was performed using the baseline survey of the Henan Rural Cohort Study. Multimorbidity was defined as the simultaneous occurrence of at least two NCDs in a participant. This study examined the multimorbidity pattern of 6 NCDs, including hypertension, dyslipidemia, type 2 diabetes mellitus, coronary heart disease, stroke, and hyperuricemia.

Results: From July 2015 to September 2017, a total of 38,807 participants (aged 18-79 years; 15,354 men and 23,453 women) were included in this study. The overall population prevalence of multimorbidity was 28.1% (10,899/38,807), and the multimorbidity of hypertension and dyslipidemia was the most common (8.1%, 3153/38,807). Aging, higher BMI, and unfavorable lifestyles were significantly associated with a higher risk of multimorbidity (multinomial logistic regression, all $P<.05$). The analysis of the mean age at diagnosis suggested a cascade of interrelated NCDs and their accumulation over time. Compared with participants without 2 conditional NCDs, participants with 1 conditional NCD would have higher odds of another NCD (1.2-2.5; all $P<.05$), and those with 2 conditional NCDs would elevate the odds of the third NCD to 1.4-3.5 (binary logistic regression, all $P<.05$).

Conclusions: Our findings indicate a plausible tendency for the coexistence and accumulation of NCDs in a rural population in Henan, China. Early prevention of multimorbidity is essential to reduce the NCD burden in the rural population.

(JMIR Public Health Surveill 2023;9:e43381) doi:10.2196/43381

KEYWORDS

multimorbidity; prevalence; associating factors; noncommunicable disease accumulation; NCD accumulation; public health

Introduction

Multimorbidity is the coexistence of multiple disease conditions over time. Many of these conditions shared similar age-associated structural, physiologic, and biological changes in predispose patients with multimorbidity, resulting in compound risks and adversities [1-4]. The occurrence of multimorbidity is often a powerful predictor of poor health outcomes in the future. Recent studies have demonstrated that people with multimorbidity were at significantly increased risks of depression, disability, and premature death. The physiological interaction between multiple disease conditions increases the difficulty of treating individual conditions [5]. Multimorbidity often leads to a substantial financial burden on individuals, families, and the health care system [6-10].

In China, this prevalence ranged from 6.4% to 76.5%, as reported in 2015 [11]. Previous studies have reported that aging, physical inactivity, and higher BMI were contributors to noncommunicable diseases (NCDs) and their multimorbidity [12]. However, as the study population and the targeted NCDs varied, the prevalence and associating factors of multimorbidity also substantially varied across different studies [11-13]. Further, in China, most studies have focused on older adults and were conducted among the community-dwelling population [14,15], whereas research on multimorbidity in rural adults remains limited. With an aging population and high levels of risk factors, multimorbidity is becoming a public health issue in China. However, understanding the trend and characteristics of the coexistence and accumulation of NCDs is limited.

The United States has emphasized the importance of timely diagnosis of multiple NCDs to support the development of effective clinical guidelines for multimorbidity [16]. In addition, previous studies found that NCDs can mutually interact with each other [17-22]. However, to our knowledge, the study of the accumulation of NCDs based on the temporal sequence of NCD onset was limited. A better understanding of the accumulation of NCDs can help explore complex interactions between NCDs and develop interventions for managing multimorbidity to diminish the public health burden.

Our study aimed to explore the disease burden of multimorbidity and determine the pattern of NCD accumulation among rural adults in Henan, China. We recruited 39,259 participants aged 18-79 years in Henan province for this study.

Methods

Study Population

The baseline survey of the Henan Rural Cohort Study enrolled a total of 39,259 participants aged from 18 to 79 years between July 2015 and September 2017 in the counties of Yuzhou, Xinxiang, Tongxu, Yima, and Suiping in Henan Province by using the stratified cluster sampling method. Detail description of the cohort had been previously published [23]. In this study, participants were excluded if they (1) had cancer or renal failure (n=350; for abnormal metabolism [protein, glucose, and lipid]) [24,25] or (2) did not have complete NCD information on hypertension, dyslipidemia, type 2 diabetes mellitus (T2DM), coronary heart disease (CHD), stroke, or hyperuricemia (n=102). Ultimately, 38,807 participants were included in this study (see the study flowchart about the data process in Multimedia Appendix 1).

Ethics Approval

This study was approved by the Zhengzhou University Life Science Ethics Committee (ethics approval 2015 MEC [S128]). Informed consent was signed by all participants.

Data Collection and Measurements

Information was collected through face-to-face interviews using a standardized questionnaire. Demographic characteristics included age (18-44, 45-59, and ≥60 years, according to the classification provided by the World Health Organization [WHO] [26]), gender, educational level (elementary school or below, junior high school, and senior high school or above), marital status (married or cohabiting and living alone), and average monthly individual income (<500 RMB [<US $72.4], 500-1000 RMB [US $72.4-144.8], and ≥1000 RMB [≥US $144.8], according to the average monthly income of Chinese rural populations [27]). Lifestyle factors included smoking status (never, former, and current), drinking status (never, former, and current), and physical activity (low, moderate, and high) [28]. Dietary status included vegetable and fruit intake (low [<500 g/day] and high [≥500 g/day]), salt diet (low and high according to dietary flavor habits), and fat diet (low [<75 g/day] and high [≥75 g/day] according to an average taking of meat from livestock and poultry). Family history of hypertension, dyslipidemia, diabetes, CHD, stroke, and gout were also recorded.

BMI was calculated as weight in kilograms divided by height in meters squared, and further categorized as underweight (<18.5 kg/m^2), normal (18.5-23.9 kg/m^2), overweight (24.0-27.9 kg/m^2), and obesity (≥28.0 kg/m^2) [29]. Systolic blood pressure and diastolic blood pressure were obtained by using an electronic sphygmomanometer (HEM-770afuzzy, Omron). Venous blood samples were obtained from individuals after at least eight hours of overnight fasting. Total cholesterol, triglyceride, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, fasting blood glucose, and serum uric acid (SUA) level were measured by a chemistry analyzer (Cobas C501, Roche Diagnostics GmbH).

Definitions of NCDs and Multimorbidity

In light of the lack of a standard approach to measure multimorbidity, the selection of morbidities to be included is inevitably subjective and relies on the previous studies. We have selected 6 NCDs—hypertension, dyslipidemia, T2DM, CHD, stroke, and hyperuricemia—that have been reported to be the core diseases in multimorbidity studies [30,31]. These 6 NCDs have been defined in our original survey for the cohort [23]. The definition of hypertension was average systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or self-reported hypertension diagnosed by a physician and receiving antihypertensive therapy within the past 2 weeks [32]. Dyslipidemia was defined as one of the following conditions: elevated total cholesterol level (≥6.2 mmol/L); elevated low-density lipoprotein cholesterol (≥4.1 mmol/L); low high-density lipoprotein cholesterol (<1.0 mmol/L); elevated triglyceride (≥2.3 mmol/L); or self-reported dyslipidemia diagnosed by a physician and the use of anti-dyslipidemia medications in the past 2 weeks [33]. T2DM was defined as fasting blood glucose ≥7.0 mmol/L or self-reported T2DM diagnosed by a physician and the use of antidiabetic medications in the past 2 weeks [33]. CHD and stroke were defined as self-reports of a previous diagnosis by specialist physicians according to criteria recommended by the WHO [34,35]. The definition of hyperuricemia was SUA >7.0 mg/dL (417.0 μmol/L) in men and SUA >6.0 mg/dL (357.0 μmol/L) in women. Multimorbidity was defined as the simultaneous occurrence of at least two NCDs in a single patient [36-39].

Statistical Analysis

The participants were divided into healthy individuals, individuals living with 1 NCD, those with multimorbidities of 2 NCDs, and those with multimorbidities of ≥3 NCDs. Continuous variables were expressed as mean (SD), and ANOVA was used for statistical analysis. Categorical variables were presented as counts and percentages, and the chi-square test was used for statistical analysis.

The prevalence of multimorbidity was assessed for each characteristic group, and we quantified the number of events for each NCD separately and visualized all observed combinations of NCDs with an intersection diagram. Multinomial logistic regression analysis was used to analyze the relationship between demographic factors, lifestyle, BMI, or family history of NCDs and 1 NCD or multimorbidity. When one factor was analyzed, all other factors were adjusted as covariance.

Based on the mean age at diagnosis, the temporal order of NCDs was determined, and binary logistic regression was used to explore the associations between NCD pairs. An NCD with an earlier onset was regarded as a predictor of an NCD with a later onset [40]. Following the cascade of NCD onset, binary logistic regression was used to assess the combined association of 2 conditional NCDs with a third, later NCD by creating a new combined variable with 4 groups of both conditional NCDs, in which the combined group without both conditional NCDs served as the reference. Demographic characteristics, lifestyle factors, dietary status, BMI, and family history of NCDs were adjusted in the regression models. In addition, in the regression of assessing the combined association, when 3 of the conditions were analyzed, the others were adjusted.

Results from the logistic regression models were presented as odds ratio (OR) with a 95% CI. The level of statistical significance was set at *P*<.05. Statistical analysis was conducted using SPSS (version 21.0; IBM Corp) and R (version 4.0.4; R Foundation for Statistical Computing).

Our study was implemented according to the Strengthening of the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [41] (see the study STROBE checklist for cross-sectional studies in Multimedia Appendix 2).

Results

Characteristics of Study Participants

Table 1 showed the characteristics of the participants of the baseline survey from July 2015 to September 2017. The mean age was 55.6 (SD 12.2) years among the 38,807 participants. Across the NCD groups, age, gender, educational level, average monthly individual income, marital status, physical activity, smoking status, drinking status, high-salt diet, high vegetable and fruit intake, high-fat diet, and BMI were all significantly different (ANOVA and chi-square test, all *P* values were <.001, except high-salt diet [*P*=.008]). Compared with the healthy individuals, participant groups with 1, 2, or 3 or more NCDs were more likely to be older and women and have lower education and income, higher BMI, poorer marital status, a higher-salt diet, and lower physical activity level (ANOVA and chi-square test, all *P*<.001). They were also less likely to have a high-fat diet and adequate fruit and vegetable intake than healthy individuals (chi-square test, all *P*<.001).

Table 1. Demographic characteristics of study participants according to multimorbidity.

Variables	Total participants (n=38,807)	No NCDs[a] (n=14,918)	1 NCD (n=12,990)	2 NCDs (n=7285)	≥3 NCDs (n=3614)	*P* value
Age (year), mean (SD)	55.6 (12.2)	52.2 (12.7)	52.3 (11.7)	58.6 (11.0)	60.8 (9.9)	<.001
Gender, women, n (%)	23,453 (60.4)	9552 (64)	7623 (58.7)	4200 (57.6)	2078 (57.5)	<.001
Educational level, n (%)						<.001
Elementary school or below	17,366 (44.8)	5912 (39.6)	5869 (45.2)	3592 (49.3)	1993 (55.2)	
Junior high school	15,457 (39.8)	6408 (43)	5241 (40.3)	2643 (36.3)	1165 (32.2)	
Senior high school or above	5984 (15.4)	2598 (17.4)	1880 (14.5)	1050 (14.4)	456 (12.6)	
Marital status, living alone, n (%)	3979 (10.3)	1335 (8.9)	1314 (10.1)	884 (12.1)	446 (12.3)	<.001
Average monthly income, n (%)						<.001
<500 RMB (<US $72.4)	13,825 (35.6)	4949 (33.2)	4693 (36.1)	2705 (37.1)	1478 (40.9)	
500-1000 RMB (US $72.4-144.8)	12,778 (32.9)	4966 (33.3)	4251 (32.7)	2482 (34.1)	1079 (29.9)	
≥1000 RMB (≥US $144.8)	12,204 (31.5)	5003 (33.5)	4046 (31.2)	2098 (28.8)	1057 (29.2)	
Physical activity, n (%)						<.001
Low	12,544 (32.3)	4126 (27.7)	4132 (31.8)	2699 (37.1)	1587 (43.9)	
Moderate	14,631 (37.7)	5899 (39.5)	4905 (37.8)	2587 (35.5)	1240 (34.3)	
High	11,632 (30)	4893 (32.8)	3953 (30.4)	1999 (27.4)	787 (21.8)	
Smoking, n (%)						<.001
Never	28,225 (72.7)	11,180 (75)	9319 (71.7)	5147 (70.6)	2579 (71.4)	
Former	3131 (8.1)	901 (6)	1019 (7.9)	726 (10)	485 (13.4)	
Current	7451 (19.2)	2837 (19)	2652 (20.4)	1412 (19.4)	550 (15.2)	
Drinking, n (%)						<.001
Never	29,981 (77.3)	11,860 (79.5)	9958 (76.7)	5463 (75)	2700 (74.7)	
Former	1794 (4.6)	523 (3.5)	549 (4.2)	432 (5.9)	290 (8)	
Current	7032 (18.1)	2535 (17)	2483 (19.1)	1390 (19.1)	624 (17.3)	
High-salt diet, n (%)	6937 (17.9)	2609 (17.5)	2319 (17.9)	1351 (18.6)	658 (18.2)	.008
More vegetable and fruit intake, n (%)	16,183 (41.7)	6541 (43.8)	5367 (41.3)	2880 (39.5)	1395 (38.6)	<.001
High-fat diet, n (%)	7421 (19.1)	3136 (21)	2447 (18.8)	1263 (17.3)	575 (15.9)	<.001
BMI (kg/m^2), mean (SD)	24.8 (3.6)	23.51 (3.2)	25.03 (3.4)	26.2 (3.6)	27.0 (3.5)	<.001

[a] NCD: noncommunicable disease.

Prevalence of NCDs and Their Multimorbidities

Among all participants, 33.5% (12,990/38,807) were living with 1 NCD, among which dyslipidemia was the most common (5858/38,807, 15.1%). In contrast, 18.8% (7285/38,807) of participants were living with 2 NCDs, and the most frequent cluster of NCDs was dyslipidemia and hypertension (3153/38,807, 8.1%). Moreover, 9.3% (3614/38,807) of participants were living with 3 or more NCDs, and the cluster of dyslipidemia, hypertension, and hyperuricemia (818/38,807, 2.1%) was the most common in participants (Figure 1 and Table 2).

Overall, the prevalence of multimorbidity (2 or ≥3 NCDs) was 28.1% (10,899/38,807), and it was significantly higher in participants who were aged over 60 years (1330/3982, 36%), men (4621/15,354, 30.1%), educated in elementary school or below (5585/17,366, 32.2%), living alone (1330/3982, 33.4%), had low physical activity (4286/12,544, 34.2%), and obese (3280/6878, 47.7%; chi-square test, all *P*<.001; Table 2).

Figure 1. NCDs multimorbidity. The left panel displayed bars for each NCD separately that quantify the total number of events per disease, and the line chart showed the cumulative prevalence of NCDs. CHD: coronary heart disease; DYS: dyslipidemia; HTN: hypertension; HUA: hyperuricemia; NCDs: noncommunicable diseases; T2DM: type 2 diabetes mellitus.

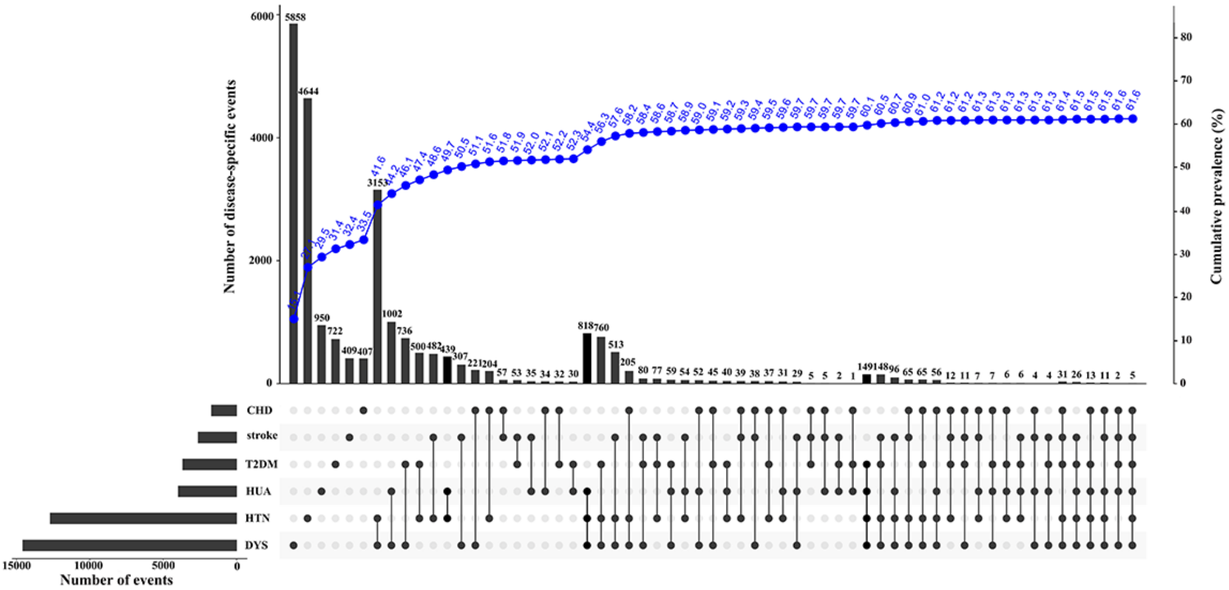


Table 2. The prevalence of 1 noncommunicable disease (NCD) and multimorbidity.

Variables	1 NCD, n (%)	2 NCDs, n (%)	≥3 NCDs, n (%)
Total (n=38,807)	12,990 (33.5)	7285 (18.8)	3614 (9.3)
Age (years)			
18-44 (n=6805)	1963 (28.6)	728 (10.7)	222 (3.3)
45-59 (n=15,181)	5225 (34.4)	2734 (18)	1156 (7.6)
≥60 (n=16,821)	5802 (34.5)	3823 (22.7)	2236 (13.3)
Gender, women (n=23,453)	7623 (32.5)	4200 (17.9)	2078 (8.9)
Educational level			
Elementary school or below (n=17,366)	5869 (33.8)	3592 (20.7)	1993 (11.5)
Junior high school (n=15,457)	5241 (33.9)	2643 (17.1)	1165 (7.5)
Senior high school or above (n=5984)	1880 (31.4)	1050 (17.6)	456 (7.6)
Marital status, living alone (n=3979)	1314 (33)	884 (22.2)	446 (11.2)
Average monthly income			
<500 RMB (<US $72.4; n=13,825)	4693 (34)	2705 (19.6)	1478 (10.7)
500-1000 RMB (US $72.4-144.8; n=12,778)	4251 (33.3)	2482 (19.4)	1079 (8.4)
≥1000 RMB (≥US $144.8; n=12,204)	4046 (33.2)	2098 (17.2)	1057 (8.7)
Physical activity			
Low (n=12,544)	4132 (32.9)	2699 (21.5)	1587 (12.7)
Moderate (n=14,631)	4905 (33.5)	2587 (17.7)	1240 (8.5)
High (n=11,632)	3953 (34)	1999 (17.2)	787 (6.8)
Smoking			
Never (n=28,225)	9319 (33)	5147 (18.2)	2579 (9.1)
Former (n=3131)	1019 (32.6)	726 (23.2)	485 (15.5)
Current (n=7451)	2652 (35.6)	1412 (19)	550 (7.4)
Drinking			
Never (n=29,981)	9958 (33.2)	5463 (18.2)	2700 (9)
Former (n=1794)	549 (30.6)	432 (24.1)	290 (16.2)
Current (n=7032)	2483 (35.3)	1390 (19.8)	624 (8.9)
High-salt diet (n=6937)	2319 (33.4)	1351 (19.5)	658 (9.5)
More vegetable and fruit intake (n=16,183)	5367 (33.2)	2880 (17.8)	1395 (8.6)
High-fat diet (n=7421)	2447 (33)	1263 (17)	575 (7.8)
BMI			
Underweight (n=938)	233 (24.8)	67 (7.1)	19 (2)
Normal (n=15,567)	4867 (31.3)	1938 (12.5)	687 (4.4)
Overweight (n=15,306)	5546 (36.2)	3243 (21.2)	1611 (10.5)
Obesity (n=6878)	2309 (33.6)	2012 (29.3)	1268 (18.4)
Family history of hypertension (n=7522)	2490 (33.1)	1657 (22)	1023 (13.6)
Family history of diabetes (n=1618)	523 (32.3)	339 (21)	194 (12)
Family history of CHD[a] (n=3106)	1053 (33.9)	534 (17.2)	306 (9.9)
Family history of hyperlipemia (n=1378)	420 (30.5)	249 (18.1)	152 (11)
Family history of stroke (n=3313)	1089 (32.9)	582 (17.6)	360 (10.9)
Family history of gout (n=80)	26 (32.5)	13 (16.3)	7 (8.8)

[a] CHD: coronary heart disease.

Associating Factors of Multimorbidity

Overall, participants with older age (all *P*<.001), those who are living alone (*P*<.001 and *P*=.007), those who are overweight and obese (all *P*<.001), former drinkers (all *P*<.001), current drinkers (*P*=.006 and *P*=.047), and those with a history of family hypertension and diabetes (all *P*<.001) were at a significantly higher risk of multimorbidity (2 or ≥3 NCDs, respectively; multinomial logistic regression). Although women (all *P*<.001) and participants with moderate and high physical activity (all *P*<.001), high-fat diet (*P*=.004 and *P*=.008), and who are underweight (all *P*<.001) were less likely to report multimorbidity (2 or ≥3 NCDs, respectively), individuals who currently smoke were less likely to report multimorbidity (≥3 NCDs; multinomial logistic regression, *P*<.001; Table 3).

Table 3. The associating factors of 1 noncommunicable disease (NCD) and multimorbidity.

Variables	1 NCD		2 NCDs		≥3 NCDs	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age (years)						
18-44	1.0		1.0		1.0	
45-59	1.7 (1.6-1.8)	<.001	2.5 (2.3-2.7)	<.001	3.5 (3.0-4.1)	<.001
≥60	2.6 (2.4-2.8)	<.001	5.2 (4.7-5.8)	<.001	10.7 (9.1-12.7)	<.001
Gender, women	0.8 (0.7-0.8)	<.001	0.8 (0.7-0.8)	<.001	0.7 (0.6-0.8)	<.001
Educational level						
Elementary school or below	1.0		1.0		1.0	
Junior high school	1.0 (0.9-1.0)	.53	0.9 (0.8-0.9)	.005	0.8 (0.7-0.9)	<.001
Senior high school or above	1.0 (0.9-1.0)	.24	1.0 (0.9-1.1)	.64	0.9 (0.8-1.0)	.08
Marital status, living alone	1.1 (1.0-1.2)	.08	1.3 (1.2-1.4)	<.001	1.2 (1.1-1.4)	.007
Average monthly income						
<500 RMB (<US $72.4)	1.00		1.0		1.0	
500-1000 RMB (US $72.4-144.8)	1.0 (0.9-1.0)	.08	1.0 (0.9-1.1)	.75	0.8 (0.7-0.9)	<.001
≥1000 RMB (≥US $144.8)	1.0 (0.9-1.0)	.28	0.9 (0.9-1.0)	.12	1.0 (0.9-1.1)	.42
Physical activity						
Low	1.0		1.0		1.0	
Moderate	0.9 (0.8-0.9)	<.001	0.8 (0.7-0.8)	<.001	0.6 (0.6-0.7)	<.001
High	0.8 (0.7-0.8)	<.001	0.6 (0.6-0.7)	<.001	0.4 (0.4-0.5)	<.001
Smoking						
Never	1.0		1.0		1.0	
Former	1.0 (0.9-1.1)	.52	1.1 (0.9-1.2)	.26	1.2 (1.1-1.4)	.02
Current	1.0 (0.9-1.1)	.51	1.0 (0.9-1.1)	.62	0.8 (0.7-0.9)	<.001
Drinking						
Never	1.0		1.0		1.0	
Former	1.0 (0.9-1.2)	.75	1.4 (1.2-1.7)	<.001	1.8 (1.5-2.1)	<.001
Current	1.1 (1.0-1.2)	.07	1.2 (1.1-1.3)	.006	1.1 (1.1-1.3)	.047
High-salt diet	1.0 (1.0-1.1)	.70	1.1 (1.0-1.1)	.26	1.0 (0.9-1.1)	.55
More vegetable and fruit intake	1.0 (0.9-1.0)	.10	1.0 (0.9-1.0)	.13	1.0 (0.9-1.1)	.43
High-fat diet	0.9 (0.9-0.9)	.03	0.9 (0.8-0.9)	.004	0.9 (0.8-0.9)	.009
BMI						
Underweight	0.6 (0.5-0.7)	<.001	0.4 (0.3-0.5)	<.001	0.3 (0.2-0.5)	<.001
Normal	1.0		1.0		1.0	
Overweight	2.0 (1.9-2.1)	<.001	3.0 (2.8-3.2)	<.001	4.4 (4.0-4.9)	<.001
Obesity	3.3 (3.1-3.6)	<.001	7.8 (7.1-8.6)	<.001	14.8 (13.1-16.6)	<.001
Family history of hypertension	1.4 (1.3-1.5)	<.001	1.9 (1.8-2.1)	<.001	2.7 (2.4-2.9)	<.001
Family history of diabetes	1.2 (1.1-1.3)	.04	1.4 (1.2-1.6)	<.001	1.6 (1.3-1.9)	<.001
Family history of CHD[b]	1.0 (1.0-1.1)	.41	0.9 (0.8-1.1)	.29	1.1 (0.9-1.3)	.27
Family history of hyperlipemia	0.9 (0.8-1.1)	.20	1.0 (0.8-1.2)	.77	1.1 (0.9-1.4)	.33
Family history of stroke	1.0 (0.9-1.0)	.25	0.9 (0.8-0.9)	.02	1.1 (0.9-1.2)	.36
Family history of gout	0.9 (0.5-1.5)	.61	0.8 (0.4-1.5)	.47	0.8 (0.3-1.9)	.55

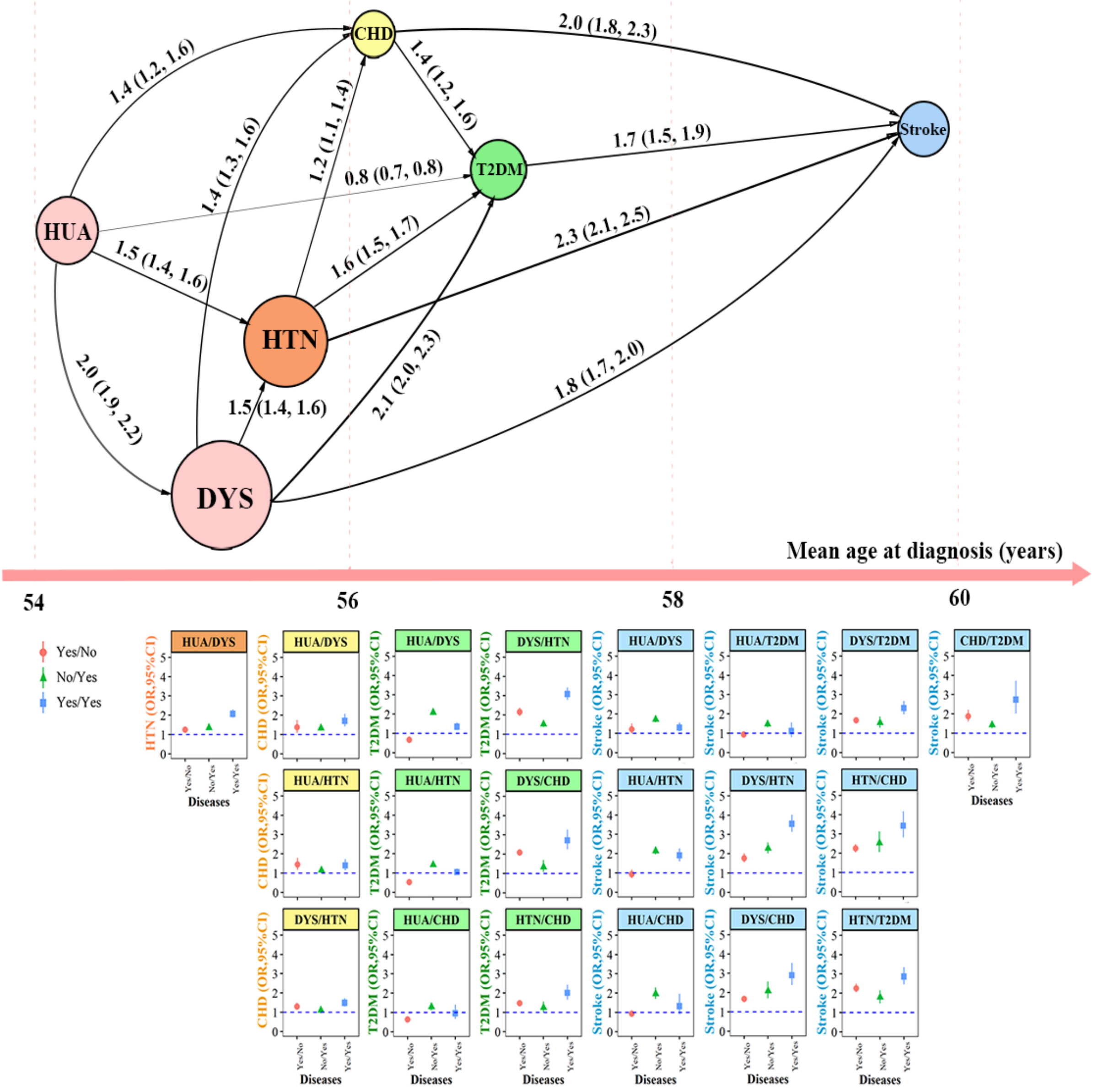
[a]OR: odds ratio.

[b]CHD: coronary heart disease.

Age Gaps Between NCDs

A cascade of NCDs was depicted by the mean age of participants at diagnosis. Among the study participants, 10.2% (3973/38,807) were diagnosed with hyperuricemia at a mean age of 54.0 years, followed by dyslipidemia (14,569/38,807, 37.5% at 55.1 years), hypertension (12,692/38,807, 32.7% at 55.4 years), CHD (1708/38,807, 4.4% at 56.1 years), T2DM (3664/38,807, 9.4% at 56.7 years), and stroke (2613/38,807, 6.7% at 59.1 years). In addition, we observed additional connections between NCD pairs (Figure 2 and Multimedia Appendix 3).

Figure 2. Age gaps between NCDs and the accumulation of NCDs in individuals based on mean age at diagnosis. NCDs along the x-axis were in order of increasing mean age at diagnosis. The number of the connected edge reflected the OR (95% CI) of the NCD pairs. The forest plot displayed the independent and combined effects of NCDs. The dots and lines exhibited OR and 95% CI, respectively. The model reference was "no/no," which meant without 2 conditional NCDs; "yes/no" meant having the first NCD of the 2 conditional NCDs, "no/yes" meant having the second NCD, and "yes/yes" meant having both conditional NCDs. CHD: coronary heart disease; DYS: dyslipidemia; HTN: hypertension; HUA: hyperuricemia; NCD: noncommunicable disease; OR: odds ratio; STR: stroke; T2DM: type 2 diabetes mellitus.



Impact of Multimorbidity on Hypertension Occurrence

Compared with individuals without 2 conditional NCDs, individuals with 1 conditional NCD (hyperuricemia or dyslipidemia) would have significantly higher odds of hypertension (1.3-1.4, logistic regression; both *P*<.001). Individuals with 2 conditional NCDs (hyperuricemia and dyslipidemia) would further elevate their odds of hypertension to 2.1 (*P*<.001; Figure 2 and Multimedia Appendix 4).

Impact of Multimorbidity on CHD Occurrence

Compared with individuals without 2 conditional NCDs, individuals with one of the conditional NCDs (hyperuricemia, dyslipidemia, or hypertension) would have significantly higher odds of CHD (1.2-1.5; the combination of hyperuricemia without dyslipidemia: *P*=.008; the combination of no hyperuricemia but dyslipidemia: *P*<.001; the combination of hyperuricemia without hypertension: *P*=.001; and the combination of no hyperuricemia but hypertension: *P*=.02). Individuals with 2 conditional NCDs (hyperuricemia and dyslipidemia combination: *P*<.001; or hyperuricemia and hypertension combination: *P*=.002) would further elevate their odds of CHD to 1.4-1.7 (Figure 2 and Multimedia Appendix 4).

Impact of Multimorbidity on T2DM Occurrence

Compared with individuals without 2 conditional NCDs, individuals with 1 conditional NCDs (dyslipidemia, hypertension, or CHD) would have a significantly higher odd of T2DM (1.3-2.2; all *P* values were <.001 except the combination of no dyslipidemia but CHD [*P*=.01]), and individuals with these 2 conditional NCDs (dyslipidemia and hypertension combination or dyslipidemia and CHD combination) would further elevate their odds of T2DM to 2.7-3.1 (all *P*<.001; Figure 2 and Multimedia Appendix 4).

Impact of Multimorbidity on Stroke Occurrence

Compared with individuals without 2 conditional NCDs, individuals with one of the conditional NCDs (dyslipidemia, hypertension, CHD, or T2DM) would have significantly higher odds of stroke (1.4-2.5; all *P*<.001). Individuals with these 2 conditional NCDs (6 NCD combinations) would further elevate their odds of a stroke to 2.3-3.5 (all *P*<.001; Figure 2 and Multimedia Appendix 4).

Discussion

Principal Findings

Our findings indicated a strong tendency for the coexistence and accumulation of NCDs in a rural population in Henan, China. The prevalence of multimorbidity was 28.1% (10,899/38,807) and increased with age, unfavorable lifestyles, and a higher BMI among Henan rural adults. The multimorbidity of dyslipidemia and hypertension was the most prevalent in study participants of all age groups. In addition, the onset of NCDs was associated, and a later NCD is often predictable by 1 or more preceding NCDs. The tendency of NCD accumulation indicates that compared with participants without conditional NCDs, participants with 1 existing NCD are more likely to develop a second NCD, whereas participants with 2 NCDs have an even greater likelihood to develop a third NCD.

Our study finding demonstrated a lower multimorbidity prevalence than previous studies [14]. A younger age, existing recall and reporting biases, and a fewer number of NCDs included may have contributed to the lower prevalence observed in our study. Further, consistent with previous studies [2,42,43], the prevalence of multimorbidity increases with age, and approximately one-third of older adults aged 60-79 years develop multimorbidity. Nevertheless, multimorbidity in the young and middle-aged population should not be neglected, given the increasing trend of risk factors in the young and middle-aged population [13,44]. Multimorbidity research and preventive strategies should not focus on older adults only but should instead recognize its importance throughout the whole life cycle.

Our study found that unfavorable lifestyles and a higher BMI were associated with a higher risk of multimorbidity. Our finding is supported by evidence in other settings. A multinational cohort study found that a prediagnostic healthy lifestyle, such as a healthy diet, lower BMI, and higher physical activity, could reduce the risk of multimorbidity [45]. A pooled study from 16 cohort studies demonstrates that a high BMI significantly contributes to the risk of cardiometabolic multimorbidity [46]. Further, a systematic review identifies that physical activity may improve immunity and reduce systemic inflammation and thereby ameliorate multimorbidity [47]. Overall, many modifiable risk factors associated with multimorbidity, including drinking, overweight and obesity, and physical inactivity, could be targets of behavioral prevention and interventions. Differing from previous studies [48-50], our study found that a high-fat diet, current smoking status, and being underweight may reduce the risk of multimorbidity. Given our study is a cross-sectional study with certain biases, the power to explain the causative effects of risk factors on diseases is limited. Further investigation is necessary to confirm this particular finding. Our findings support the WHO’s recommendation to implement evidence-based strategies to reduce poor use of alcohol, encourage physical activity, and maintain a healthy weight.

Our study suggested a tendency for NCD accumulation among the participants. The clustering of dyslipidemia and hypertension was the most prevalent multimorbidity observed in our study, consistent with previous studies [30,51]. A Japanese study further describes hypertension as the most common comorbidity among individuals with dyslipidemia [52]. In addition, the temporal sequence reflects the order of NCDs onset and suggests a potential causal association between the occurrence of these NCDs and the development of other comorbidities [40]. The accumulation of NCDs reflects that participants who lived with 1 conditional NCD are more likely to develop a second NCD, whereas individuals with 2 conditional NCDs are more likely to develop a third NCD than otherwise. Similar results have also been reported in a large-series prospective study: participants with the multimorbidity of T2DM and hypertension have a higher risk of CHD or stroke than participants with hypertension or T2DM [53]. This observation might be partly explained by obesity and the variations of metabolites that have accumulating side effects on health over time [54,55]. Kivimaki et al [54] found that obesity-related disease predicted or was predicted by 1 or more other obesity-related diseases, leading to the accumulation of NCDs. A study of 11,000 participants to examine metabolic pathways associated with 27 NCDs shows that two-thirds of metabolites are associated with more than 1 NCD. In addition, 420 metabolites shared between multimorbidity have been found, revealing several key common pathways in NCDs [55]. Our study suggests that taking reasonable health management and therapies of early NCDs can reduce the incidence or slow down the progression of multimorbidities effectively and ultimately improve the quality of life of the population [56,57].

Our findings have important implications for research and public health policy for multimorbidity. This study indicates that multimorbidity is an important health concern in rural Henan. In light of the WHO Sustainable Development Goals for preventing and controlling NCDs [58], existing health care needs to be improved to cope with the burden of multimorbidity. First, the prevalence of multimorbidity is higher in those older than 45 years. A life cycle approach to multimorbidity and to the challenges it poses to public health is vitally important [59]. Second, there is a clear need to increase the awareness of the importance of a healthy lifestyle in the rural population in Henan. Measures of propagating publicity and education on relevant health knowledge, improving health literacy, and changing unfavorable lifestyles should be implemented. Third, disease-specific guidelines are insufficient to effectively manage patients with multimorbidity, and new detailed guidelines for multimorbidity need to be developed [6]. Given the tendency of the coexistence and accumulation of NCDs, preventing multimorbidity will become difficult if the NCDs are treated separately. Integrated strategies for multimorbidity should be developed, such as monitoring at the individual level with electronic health records and increasing the screening management of high-risk populations.

Limitations

This study has several limitations. First, the causal relationship between NCDs may not be reliable on account of the data being from the baseline survey of a prospective cohort study in Henan Province. Therefore, prospective studies are needed to confirm the results of this study. Second, we define hypertension, dyslipidemia, hyperuricemia, and T2DM by self-report with laboratory measurements, whereas CHD and stroke are defined only by self-report, which might underestimate their prevalence. Finally, although several potential confounders are controlled, there may be other unknown confounders affecting the results of this study.

Conclusions

In conclusion, our study indicates that multimorbidity is an important health issue in rural Henan. Older age, unfavorable lifestyles, and a higher BMI are associated with a higher risk of multimorbidity. In addition, our findings highlight a plausible explanation for the coexistence and accumulation of NCDs. Care for multimorbidity is complex, and concerted efforts and early interventions to prevent the occurrence of multimorbidity are essential to improve health in the rural population in Henan, China.

Acknowledgments

The authors would like to give great gratitude to the participants and all research staff of this study.

Authors' Contributions

L Zhang and CW designed this study. YC, MP, YH, XD, ZH, JH, JY, YY, RL, L Zhu, SL, NK, and WL directed the data collection. YC analyzed the data and wrote the manuscript. JH and YB provided writing assistance. All authors listed have read, corrected, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study flowchart about the data process.
[PNG File , 47 KB - publichealth_v9i1e43381_app1.png]

Multimedia Appendix 2
The Strengthening of the Reporting of Observational Studies in Epidemiology (STROBE) statement checklist for cross-sectional studies.
[DOCX File , 29 KB - publichealth_v9i1e43381_app2.docx]

Multimedia Appendix 3
The association between noncommunicable disease pairs based on mean age at diagnosis.
[DOC File , 40 KB - publichealth_v9i1e43381_app3.doc]

Multimedia Appendix 4
The accumulation of noncommunicable diseases in individuals based on mean age at diagnosis.
[DOC File , 100 KB - publichealth_v9i1e43381_app4.doc]

References

1. Nobili A, Garattini S, Mannucci PM. Multiple diseases and polypharmacy in the elderly: challenges for the internist of the third millennium. J Comorb 2011 Dec 27;1:28-44 [FREE Full text] [doi: 10.15256/joc.2011.1.4] [Medline: 29090134]

2. Chen H, Cheng M, Zhuang Y, Broad JB. Multimorbidity among middle-aged and older persons in urban China: prevalence, characteristics and health service utilization. Geriatr Gerontol Int 2018 Oct;18(10):1447-1452. [doi: 10.1111/ggi.13510] [Medline: 30178629]
3. Zou S, Wang Z, Bhura M, Zhang G, Tang K. Prevalence and associated socioeconomic factors of multimorbidity in 10 regions of China: an analysis of 0.5 million adults. J Public Health (Oxf) 2022 Mar 07;44(1):36-50. [doi: 10.1093/pubmed/fdaa204] [Medline: 33300571]
4. Barnes PJ. Mechanisms of development of multimorbidity in the elderly. Eur Respir J 2015 Mar;45(3):790-806 [FREE Full text] [doi: 10.1183/09031936.00229714] [Medline: 25614163]
5. Whitty CJM, MacEwen C, Goddard A, Alderson D, Marshall M, Calderwood C, et al. Rising to the challenge of multimorbidity. BMJ 2020 Jan 06;368:l6964 [FREE Full text] [doi: 10.1136/bmj.l6964] [Medline: 31907164]
6. Kernick D, Chew-Graham CA, O'Flynn N. Clinical assessment and management of multimorbidity: NICE guideline. Br J Gen Pract 2017 May;67(658):235-236 [FREE Full text] [doi: 10.3399/bjgp17X690857] [Medline: 28450343]
7. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007 Dec;22 Suppl 3(Suppl 3):391-395 [FREE Full text] [doi: 10.1007/s11606-007-0322-1] [Medline: 18026807]
8. Hu RH, Hsiao FY, Chen LJ, Huang PT, Hsu WWY. Increasing age- and gender-specific burden and complexity of multimorbidity in Taiwan, 2003-2013: a cross-sectional study based on nationwide claims data. BMJ Open 2019 Jun 09;9(6):e028333 [FREE Full text] [doi: 10.1136/bmjopen-2018-028333] [Medline: 31182451]
9. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev 2011 Sep;10(4):430-439. [doi: 10.1016/j.arr.2011.03.003] [Medline: 21402176]
10. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. Ann Fam Med 2009 Jul;7(4):357-363 [FREE Full text] [doi: 10.1370/afm.983] [Medline: 19597174]
11. Hu X, Huang J, Lv Y, Li G, Peng X. Status of prevalence study on multimorbidity of chronic disease in China: systematic review. Geriatr Gerontol Int 2015 Jan;15(1):1-10. [doi: 10.1111/ggi.12340] [Medline: 25163532]
12. Khan N, Rahman M, Mitra D, Afsana K. Prevalence of multimorbidity among Bangladeshi adult population: a nationwide cross-sectional study. BMJ Open 2019 Nov 28;9(11):e030886 [FREE Full text] [doi: 10.1136/bmjopen-2019-030886] [Medline: 31784434]
13. The Lancet. Making more of multimorbidity: an emerging priority. Lancet 2018 Apr 28;391(10131):1637. [doi: 10.1016/S0140-6736(18)30941-3] [Medline: 29726322]
14. Yao SS, Cao GY, Han L, Chen ZS, Huang ZT, Gong P, et al. Prevalence and Patterns of Multimorbidity in a Nationally Representative Sample of Older Chinese: Results From the China Health and Retirement Longitudinal Study. J Gerontol A Biol Sci Med Sci 2020 Sep 25;75(10):1974-1980. [doi: 10.1093/gerona/glz185] [Medline: 31406983]
15. Zhang R, Lu Y, Shi L, Zhang S, Chang F. Prevalence and patterns of multimorbidity among the elderly in China: a cross-sectional study using national survey data. BMJ Open 2019 Aug 18;9(8):e024268 [FREE Full text] [doi: 10.1136/bmjopen-2018-024268] [Medline: 31427309]
16. Parekh AK, Goodman RA, Gordon C, Koh HK, HHS Interagency Workgroup on Multiple Chronic Conditions. Managing multiple chronic conditions: a strategic framework for improving health outcomes and quality of life. Public Health Rep 2011 Jul 1;126(4):460-471 [FREE Full text] [doi: 10.1177/003335491112600403] [Medline: 21800741]
17. Wu Y, Hu H, Cai J, Chen R, Zuo X, Cheng H, et al. Association of hypertension and incident diabetes in Chinese adults: a retrospective cohort study using propensity-score matching. BMC Endocr Disord 2021 Apr 29;21(1):87 [FREE Full text] [doi: 10.1186/s12902-021-00747-0] [Medline: 33926442]
18. Rho YH, Lu N, Peloquin CE, Man A, Zhu Y, Zhang Y, et al. Independent impact of gout on the risk of diabetes mellitus among women and men: a population-based, BMI-matched cohort study. Ann Rheum Dis 2016 Jan;75(1):91-95 [FREE Full text] [doi: 10.1136/annrheumdis-2014-205827] [Medline: 25277955]
19. Mortada I. Hyperuricemia, type 2 diabetes mellitus, and hypertension: an emerging association. Curr Hypertens Rep 2017 Sep;19(9):69. [doi: 10.1007/s11906-017-0770-x] [Medline: 28770533]
20. Wang J, Qin T, Chen J, Li Y, Wang L, Huang H, et al. Hyperuricemia and risk of incident hypertension: a systematic review and meta-analysis of observational studies. PLoS One 2014;9(12):e114259 [FREE Full text] [doi: 10.1371/journal.pone.0114259] [Medline: 25437867]
21. Kopin L, Lowenstein C. Dyslipidemia. Ann Intern Med 2017 Dec 05;167(11):ITC81-ITC96. [doi: 10.7326/AITC201712050] [Medline: 29204622]
22. Wang L, Qiu H, Luo L, Zhou L. Age- and sex-specific differences in multimorbidity patterns and temporal trends on assessing hospital discharge records in Southwest China: network-based study. J Med Internet Res 2022 Feb 25;24(2):e27146 [FREE Full text] [doi: 10.2196/27146] [Medline: 35212632]
23. Liu X, Mao Z, Li Y, Wu W, Zhang X, Huo W, et al. Cohort profile: The Henan Rural Cohort: a prospective study of chronic non-communicable diseases. Int J Epidemiol 2019 Dec 01;48(6):1756-1756j. [doi: 10.1093/ije/dyz039] [Medline: 30915440]
24. Gyamfi J, Kim J, Choi J. Cancer as a metabolic disorder. Int J Mol Sci 2022 Jan 21;23(3):1155 [FREE Full text] [doi: 10.3390/ijms23031155] [Medline: 35163079]

25. Rubenfeld S, Garber AJ. Abnormal carbohydrate metabolism in chronic renal failure. the potential role of accelerated glucose production, increased gluconeogenesis, and impaired glucose disposal. J Clin Invest 1978 Jul;62(1):20-28 [FREE Full text] [doi: 10.1172/JCI109107] [Medline: 659634]
26. Chen Y, Bao J, Yan Q, Wu C, Yang H, Zou J. Distribution of Modic changes in patients with low back pain and its related factors. Eur J Med Res 2019 Oct 09;24(1):34 [FREE Full text] [doi: 10.1186/s40001-019-0393-6] [Medline: 31597571]
27. China Statistical Yearbook 2018. National Bureau of Statistics. 2018. URL: http://www.stats.gov.cn/sj/ndsj/2018/indexeh.htm [accessed 2020-10-24]
28. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011 Aug;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
29. Wang Y, Li Y, Liu X, Zhang H, Abdulai T, Tu R, et al. Prevalence and influencing factors of coronary heart disease and stroke in Chinese rural adults: The Henan Rural Cohort Study. Front Public Health 2019 Jan 21;7:411 [FREE Full text] [doi: 10.3389/fpubh.2019.00411] [Medline: 32039127]
30. Gu J, Chao J, Chen W, Xu H, Wu Z, Chen H, et al. Multimorbidity in the community-dwelling elderly in urban China. Arch Gerontol Geriatr 2017 Jan;68:62-67. [doi: 10.1016/j.archger.2016.09.001] [Medline: 27654809]
31. He L, Biddle SJH, Lee JT, Duolikun N, Zhang L, Wang Z, et al. The prevalence of multimorbidity and its association with physical activity and sleep duration in middle aged and elderly adults: a longitudinal analysis from China. Int J Behav Nutr Phys Act 2021 Jun 10;18(1):77 [FREE Full text] [doi: 10.1186/s12966-021-01150-7] [Medline: 34112206]
32. Li N, Chen G, Liu F, Mao S, Liu Y, Hou Y, et al. Associations of long-term exposure to ambient PM with hypertension and blood pressure in rural Chinese population: The Henan Rural Cohort Study. Environ Int 2019 Jul;128:95-102 [FREE Full text] [doi: 10.1016/j.envint.2019.04.037] [Medline: 31035115]
33. Dong X, Zhang H, Wang F, Liu X, Yang K, Tu R, et al. Epidemiology and prevalence of hyperuricemia among men and women in Chinese rural population: The Henan Rural Cohort Study. Mod Rheumatol 2020 Sep;30(5):910-920. [doi: 10.1080/14397595.2019.1660048] [Medline: 31442098]
34. Joint International Society and Federation of Cardiology/World Health Organization Task Force. Nomenclature and criteria for diagnosis of ischemic heart disease. report of the Joint International Society and Federation of Cardiology/World Health Organization task force on standardization of clinical nomenclature. Circulation 1979 Mar;59(3):607-609. [doi: 10.1161/01.cir.59.3.607] [Medline: 761341]
35. Aho K, Harmsen P, Hatano S, Marquardsen J, Smirnov VE, Strasser T. Cerebrovascular disease in the community: results of a WHO collaborative study. Bull World Health Organ 1980;58(1):113-130 [FREE Full text] [Medline: 6966542]
36. Forman DE, Maurer MS, Boyd C, Brindis R, Salive ME, Horne FM, et al. Multimorbidity in older adults with cardiovascular disease. J Am Coll Cardiol 2018 May 15;71(19):2149-2161 [FREE Full text] [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]
37. Multiple long-term conditions (Multimorbidity): a priority for global health research. The Academy of Medical Sciences. 2018. URL: https://acmedsci.ac.uk/policy/policy-projects/multimorbidity [accessed 2020-10-24]
38. Ho IS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. Lancet Public Health 2021 Aug;6(8):e587-e597 [FREE Full text] [doi: 10.1016/S2468-2667(21)00107-9] [Medline: 34166630]
39. Shi X, Lima SMDS, Mota CMDM, Lu Y, Stafford RS, Pereira CV. Prevalence of multimorbidity of chronic noncommunicable diseases in Brazil: population-based study. JMIR Public Health Surveill 2021 Nov 25;7(11):e29693 [FREE Full text] [doi: 10.2196/29693] [Medline: 34842558]
40. Kivimäki M, Batty GD, Pentti J, Shipley MJ, Sipilä PN, Nyberg ST, et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. Lancet Public Health 2020 Mar;5(3):e140-e149 [FREE Full text] [doi: 10.1016/S2468-2667(19)30248-8] [Medline: 32007134]
41. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296 [FREE Full text] [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
42. Zhang L, Ma L, Sun F, Tang Z, Chan P. A multicenter study of multimorbidity in older adult inpatients in China. J Nutr Health Aging 2020 Jan 10;24(3):269-276 [FREE Full text] [doi: 10.1007/s12603-020-1311-x] [Medline: 32115607]
43. Keetile M, Navaneetham K, Letamo G. Prevalence and correlates of multimorbidity among adults in Botswana: a cross-sectional study. PLoS One 2020;15(9):e0239334 [FREE Full text] [doi: 10.1371/journal.pone.0239334] [Medline: 32976484]
44. Pefoyo AJK, Bronskill SE, Gruneir A, Calzavara A, Thavorn K, Petrosyan Y, et al. The increasing burden and complexity of multimorbidity. BMC Public Health 2015 Apr 23;15:415 [FREE Full text] [doi: 10.1186/s12889-015-1733-2] [Medline: 25903064]
45. Freisling H, Viallon V, Lennon H, Bagnardi V, Ricci C, Butterworth AS, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. BMC Med 2020 Jan 10;18(1):5 [FREE Full text] [doi: 10.1186/s12916-019-1474-7] [Medline: 31918762]

46. Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120 813 adults from 16 cohort studies from the USA and Europe. Lancet Public Health 2017 Jun;2(6):e277-e285 [FREE Full text] [doi: 10.1016/S2468-2667(17)30074-9] [Medline: 28626830]
47. Duggal NA, Niemiro G, Harridge SDR, Simpson RJ, Lord JM. Can physical activity ameliorate immunosenescence and thereby reduce age-related multi-morbidity? Nat Rev Immunol 2019 Sep;19(9):563-572. [doi: 10.1038/s41577-019-0177-9] [Medline: 31175337]
48. Ng R, Sutradhar R, Yao Z, Wodchis WP, Rosella LC. Smoking, drinking, diet and physical activity-modifiable lifestyle risk factors and their associations with age to first chronic disease. Int J Epidemiol 2020 Feb 01;49(1):113-130 [FREE Full text] [doi: 10.1093/ije/dyz078] [Medline: 31329872]
49. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet 2017 Dec 16;390(10113):2627-2642 [FREE Full text] [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]
50. Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Houston Miller N, Hubbard VS, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines. Circulation 2014 Jun 24;129(25 Suppl 2):S76-S99 [FREE Full text] [doi: 10.1161/01.cir.0000437740.48606.d1] [Medline: 24222015]
51. Mitsutake S, Ishizaki T, Teramoto C, Shimizu S, Ito H. Patterns of co-occurrence of chronic disease among older adults in Tokyo, Japan. Prev Chronic Dis 2019 Jan 31;16:E11 [FREE Full text] [doi: 10.5888/pcd16.180170] [Medline: 30703000]
52. Otsuka T, Takada H, Nishiyama Y, Kodani E, Saiki Y, Kato K, et al. Dyslipidemia and the risk of developing hypertension in a working-age male population. J Am Heart Assoc 2016 Mar 25;5(3):e003053 [FREE Full text] [doi: 10.1161/JAHA.115.003053] [Medline: 27016576]
53. Yen FS, Wei JC, Chiu L, Hsu CC, Hwu CM. Diabetes, hypertension, and cardiovascular disease development. J Transl Med 2022 Jan 03;20(1):9 [FREE Full text] [doi: 10.1186/s12967-021-03217-2] [Medline: 34980154]
54. Kivimäki M, Strandberg T, Pentti J, Nyberg ST, Frank P, Jokela M, et al. Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. Lancet Diabetes Endocrinol 2022 Apr;10(4):253-263 [FREE Full text] [doi: 10.1016/S2213-8587(22)00033-X] [Medline: 35248171]
55. Pietzner M, Stewart ID, Raffler J, Khaw K, Michelotti GA, Kastenmüller G, et al. Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. Nat Med 2021 Mar;27(3):471-479 [FREE Full text] [doi: 10.1038/s41591-021-01266-0] [Medline: 33707775]
56. Chudasama YV, Khunti KK, Zaccardi F, Rowlands AV, Yates T, Gillies CL, et al. Physical activity, multimorbidity, and life expectancy: a UK Biobank longitudinal study. BMC Med 2019 Jun 12;17(1):108 [FREE Full text] [doi: 10.1186/s12916-019-1339-0] [Medline: 31186007]
57. Muth C, Blom JW, Smith SM, Johnell K, Gonzalez-Gonzalez AI, Nguyen TS, et al. Evidence supporting the best clinical management of patients with multimorbidity and polypharmacy: a systematic guideline review and expert consensus. J Intern Med 2019 Mar 10;285(3):272-288 [FREE Full text] [doi: 10.1111/joim.12842] [Medline: 30357955]
58. Global action plan for the prevention and control of NCDs 2013–2020. World Health Organization. 2013 Nov 14. URL: https://www.who.int/publications/i/item/9789241506236 [accessed 2023-05-09]
59. Zhang Y, Chen C, Huang L, Liu G, Lian T, Yin M, et al. Associations among multimorbid conditions in hospitalized middle-aged and older adults in China: statistical analysis of medical records. JMIR Public Health Surveill 2022 Nov 24;8(11):e38182 [FREE Full text] [doi: 10.2196/38182] [Medline: 36422885]

Abbreviations

CHD: coronary heart disease
NCD: noncommunicable disease
OR: odds ratio
STROBE: Strengthening of the Reporting of Observational Studies in Epidemiology
SUA: serum uric acid
T2DM: type 2 diabetes mellitus
WHO: World Health Organization

Edited by Y Khader; submitted 11.10.22; peer-reviewed by M He, R Bai; comments to author 03.02.23; revised version received 07.03.23; accepted 18.03.23; published 22.05.23.

Please cite as:
Chen Y, Pan M, He Y, Dong X, Hu Z, Hou J, Bao Y, Yang J, Yuchi Y, Li R, Zhu L, Kang N, Liao W, Li S, Wang C, Zhang L
Disease Burden and the Accumulation of Multimorbidity of Noncommunicable Diseases in a Rural Population in Henan, China: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e43381
URL: https://publichealth.jmir.org/2023/1/e43381
doi:10.2196/43381
PMID:37213192

Original Paper

Evaluation of Diabetes Care Performance in Cambodia Through the Cascade-of-Care Framework: Cross-Sectional Study

Vannarath Te[1,2,3], MA, MPH, DDS; Srean Chhim[1,4], MSc, MPhil; Veerle Buffel[5], PhD; Wim Van Damme[2], MD, PhD; Josefien van Olmen[3], MD, PhD; Por Ir[1,4], MD, PhD; Edwin Wouters[5,6], PhD

[1]School of Public Health, National Institute of Public Health, Phnom Penh, Cambodia
[2]Health Policy Unit, Department of Public Health, Institute of Tropical Medicine (Antwerp), Antwerp, Belgium
[3]Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium
[4]Technical Office, National Institute of Public Health, Phnom Penh, Cambodia
[5]Centre for Population, Family & Health, Department of Sociology, University of Antwerp, Antwerp, Belgium
[6]Centre for Health Systems Research & Development, University of the Free State, Bloemfontein, South Africa

Corresponding Author:
Vannarath Te, MA, MPH, DDS
School of Public Health
National Institute of Public Health
Lot 80, 289 Samdach Penn Nouth St. (289)
Phnom Penh, 12152
Cambodia
Phone: 855 16524518
Email: vannarath_te@yahoo.com

Abstract

Background: Cambodia has seen an increase in the prevalence of type 2 diabetes (T2D) over the last 10 years. Three main care initiatives for T2D are being scaled up in the public health care system across the country: *hospital-based care*, *health center–based care*, and *community-based care*. To date, no empirical study has systematically assessed the performance of these care initiatives across the T2D care continuum in Cambodia.

Objective: This study aimed to assess the performance of the 3 care initiatives—individually or in coexistence—and determine the factors associated with the failure to diagnose T2D in Cambodia.

Methods: We used a cascade-of-care framework to assess the T2D care continuum. The cascades were generated using primary data from a cross-sectional population-based survey conducted in 2020 with 5072 individuals aged ≥40 years. The survey was conducted in 5 operational districts (ODs) selected based on the availability of the care initiatives. Multiple logistic regression analysis was used to identify the factors associated with the failure to diagnose T2D. The significance level of $P<.05$ was used as a cutoff point.

Results: Of the 5072 individuals, 560 (11.04%) met the definition of a T2D diagnosis (fasting blood glucose level ≥126 mg/dL and glycated hemoglobin level ≥6.5%). Using the 560 individuals as the fixed denominator, the cascade displayed substantial drops at the testing and control stages. Only 63% (353/560) of the participants had ever tested their blood glucose level in the last 3 years, and only 10.7% (60/560) achieved blood glucose level control with the cutoff point of glycated hemoglobin level <8%. The OD hosting the coexistence of care displayed the worst cascade across all bars, whereas the OD with hospital-based care had the best cascade among the 5 ODs. Being aged 40 to 49 years, male, and in the poorest category of the wealth quintile were factors associated with the undiagnosed status.

Conclusions: The unmet needs for T2D care in Cambodia were large, particularly in the testing and control stages, indicating the need to substantially improve early detection and management of T2D in the country. Rapid scale-up of T2D care components at public health facilities to increase the chances of the population with T2D of being tested, diagnosed, retained in care, and treated, as well as of achieving blood glucose level control, is vital in the health system. Specific population groups susceptible to being undiagnosed should be especially targeted for screening through active community outreach activities. Future research should incorporate digital health interventions to evaluate the effectiveness of the T2D care initiatives longitudinally with more diverse population groups from various settings based on routine data vital for integrated care.

Trial Registration: International Standard Randomized Controlled Trials Number (ISRCTN) ISRCTN41932064; https://www.isrctn.com/ISRCTN41932064

International Registered Report Identifier (IRRID): RR2-10.2196/36747

(JMIR Public Health Surveill 2023;9:e41902) doi:10.2196/41902

KEYWORDS

diabetes; cascade of care; implementation research; population-based survey; care continuum; mobile phone

Introduction

Background

In 2021, globally, 1 in 10 adults aged 20 to 79 years was living with type 2 diabetes (T2D) [1]. Similar to other countries in the Western Pacific region—the World Health Organization (WHO) region with the highest number of adults living with T2D (206 million) [1]—Cambodia has been severely affected by the T2D epidemic. In 2016, the prevalence rate of T2D in this lower–middle-income country was 9.6% among adults aged 18 to 69 years, signifying a significant increase compared with the prevalence rate in 2010 (2.9% among adults aged 25-64 years) [2]. As a response to the increased burden of chronic conditions, including T2D, across the globe, the WHO adopted the innovative care for chronic conditions (ICCC) framework as a road map for countries, regardless of income level, to transform their health systems toward better care for chronic conditions [3].

The incurable nature of T2D, along with its chronicity and silent progression, requires the condition to be diagnosed as early as possible and managed properly and promptly on a regular basis by patients, caregivers, and health care professionals to prevent or delay complications [3]. Care for T2D relies not only on medical interventions provided by health care professionals but also on high-quality and continuous self-management [4]. A systematic review shows that the quality of T2D care in low- and middle-income countries (LMICs) in Asia and the Middle East has been reported to be limited, with the care goals recommended in the evidence-based guidelines not being met [5]. The WHO promotes the adoption of integrated care for disease management in the health system as outlined in the ICCC framework [3], which is evidence based in improving blood glucose level control [6,7], to fill the gaps.

Likewise, care for T2D in Cambodia has been limited among the population at risk and people living with T2D. Many adults (more than two-thirds of the population) have never had their blood glucose level tested, and more than half of those living with T2D are not receiving treatment [2].

Cambodia’s health system is pluralistic—both public providers and private providers (including nonprofit organizations) provide care for T2D in the country. The ministry of health is in charge of the public health providers, which are organized on a district health system model and guided by the primary health care approach [8]. In this model, an operational district (OD) usually comprises a referral hospital providing secondary care and 10 to 25 health centers providing primary care with support from community health workers. The public providers mainly cover health prevention activities by providing primary health care for people with infectious diseases (such as HIV infection and AIDS, tuberculosis, and malaria) and focusing on maternal and child health, leaving care for chronic conditions, including T2D, to be provided mainly by private providers [9]. A self-reported survey on availability of T2D services at primary care facilities indicated that only approximately 1 in 5 health centers reported providing T2D services [10].

To improve the availability of integrated care for T2D, three main care initiatives for T2D are currently being scaled up across the 103 ODs in Cambodia [11]: (1) *hospital-based care*, (2) *health center–based care*, and (3) *community-based care*. In 2019, the ministry of health approved a national standard operating procedure for the management of T2D and hypertension in primary care in an attempt to integrate these 3 care initiatives for the T2D and hypertension care continuum in which health centers provide continuity and coordination across the care levels in the OD [11]. This standard operating procedure was adapted from the WHO package of essential noncommunicable disease interventions (WHO PEN) [12]—we refer to health centers implementing the standard operating procedure as the WHO PEN health centers.

Hospital-based care is standard care at the referral hospitals that provide ambulatory care for serious or complicated T2D cases. In 2018, with support from the ministry of health, 29 of the 117 referral hospitals provided exclusive care for T2D and hypertension at separate noncommunicable disease (NCD) clinics [13]. Health center–based care is implemented at the WHO PEN health centers. They are allowed to take care of mild or stable T2D cases without complications, with the diagnosis confirmed and treatment initiated at the referral hospital. In early 2020, only 86 of the 1221 health centers had implemented the WHO PEN program [13]. Community-based care is implemented through peer educator networks run by a Cambodian nongovernmental organization called MoPoTsyo. The networks offer (1) self-management support to patients through peer educators who have been diagnosed with T2D themselves, (2) laboratory tests, (3) physician consultations, and (4) low-cost medicines through a revolving drug fund program [14]. Each peer educator is responsible for a health center’s catchment area, with populations ranging from 10,000 to 20,000 [8]. By 2019, MoPoTsyo had 255 peer educators trained to serve >40,000 patients [14]. Detailed descriptions of the 3 care initiatives have been provided in a study protocol [15].

The coexistence of the 3 care initiatives—combining hospital-, health center–, and community-based care components—in an OD could potentially produce the ideal context for integrated care as described in the ICCC framework [3]. However, for the care initiatives to be integrated and thus for these 3 care initiatives to strengthen each other, necessary information has to be shared, and resources have to be coordinated in an effective and efficient manner [16].

To our knowledge, no empirical study has assessed the performance of the aforementioned 3 care initiatives—either individually or in coexistence—across the T2D care continuum in Cambodia. We used the test-treat-retain cascade of care as adapted from the HIV program to assess the T2D care continuum [17]. This method allowed us to document how many patients were lost along the care continuum with regard to testing, diagnosis, retention in care, receiving treatment, and achieving good control of their health condition. In other LMICs, limited existing studies have pooled secondary data from cross-sectional surveys to generate countrywide cascades of care [18-20] as an approach to assess health system performance to meet the T2D care continuum goals. A systematic assessment of the performance of the different care initiatives—either individually or in coexistence—currently being scaled up in Cambodia is not yet available.

Objectives

This study aimed to address the research gap by assessing the performance of the aforementioned care initiatives either individually or in coexistence with the cascade-of-care framework using primary data from a population-based survey and determining the factors associated with the undiagnosed status of T2D among the population.

Methods

Study Design

The study was part of a population-based survey conducted in 2020. It was a cross-sectional study involving 5072 individuals aged ≥40 years [15]. A detailed explanation of the study design was included in the study protocol [15].

Study Setting

Five ODs were purposively selected to assess the performance of the aforementioned care initiatives—individually or in coexistence. OD Samrong in Oddar Meanchey province provided hospital-based care at the NCD clinic of the referral hospital—the only public provider for T2D care in the OD at the time of the study. People with T2D visited the physician for a medical consultation (prescriptions and medicines were provided) on a monthly appointment basis. The second selected OD was OD Pearaing in Prey Veng province. This OD began implementing health center–based care in 2015. Assigned staff at the WHO PEN health centers receive training to perform screening, provide follow-up care for mild and stable T2D cases, and offer health education and counseling on healthy behavior as part of cardiovascular disease risk factors management [11]. By structural design, the NCD clinic at the referral hospital is required to support the WHO PEN health centers. In this OD, at the time of the study, 8 of the 9 health centers were WHO PEN health centers (high coverage). The third selected OD was OD Sotr Nikum in Siem Reap province. This OD has been historically and substantially influenced by financial aid from various development partners and nongovernmental organizations. At the time of the study, 5 of the 25 health centers in the OD were WHO PEN health centers (low coverage), supported by a chronic disease clinic that provided treatment and care to both people with T2D and hypertension and those with HIV infection [21]—the clinic was essentially the NCD clinic of the referral hospital. Therefore, we consider this OD the host of health center–based care (with context). The fourth selected OD was OD Kong Pisey in Kampong Speu province. In this OD, the peer educator network provided community-based care. At the time of the study, none of the public providers in this OD formally offered care for people with T2D. MoPoTsyo made arrangements with the referral hospital to provide physician consultations for people with T2D in the network once a week. The fifth selected OD was OD Daunkeo in Takeo province—the only OD found to host all 3 care initiatives together across the 103 ODs in Cambodia. At the time of the study, hospital-based care in this OD was provided at the NCD clinic of the referral hospital, whereas 8 of the 15 WHO PEN health centers in this OD provided health center–based care. The peer educator network provided community-based care, but the network had already been handed over to the OD health authorities for governance. Table 1 summarizes the study settings.

Table 1. Selected provinces and operational districts (ODs) with different types of care initiatives.

Name of OD	Province	Existing care provision	Care initiative
Samrong	Oddar Meanchey	NCD[a] clinic at referral hospital	Hospital-based care
Pearaing	Prey Veng	NCD clinic+WHO PEN[b] (high coverage)	Health center–based care
Sotr Nikum	Siem Reap	NCD clinic+WHO PEN (low coverage)	Health center–based care with context
Kong Pisey	Kampong Speu	Peer educator network	Community-based care
Daunkeo	Takeo	NCD clinic+WHO PEN+peer educator network	Coexistence of care

[a] NCD: noncommunicable disease.

[b] WHO PEN: World Health Organization package of essential noncommunicable disease interventions.

Study Participants and Recruitment

The target study participants were adults aged ≥40 years. This age group was targeted for T2D screening according to the national standard operating procedure [11]. The recruitment was processed via a 3-level procedure. First, within each OD, a list of villages affected by the care initiative was drawn up, and 44 villages were randomly selected. Second, 24 eligible households (ie, those containing at least 1 household member aged ≥40 years) were randomly selected from a list of all eligible households in the selected villages. Within the selected households, potential participants had to be (1) usual members of the household either staying in the house the night before the interview or not being absent for >6 months, (2) physically and mentally capable of answering questions, and (3) well-informed regarding the consent procedure for participation in the study. In the third step, 1 household member meeting the aforementioned eligibility criteria from each randomly selected household was randomly recruited into the study. Each selected participant was interviewed based on a preset questionnaire and their anthropometric measurements (blood pressure, body weight, height, and waist and hip circumferences) and biochemical measurements (fasting blood glucose [FBG] level, glycated hemoglobin [HbA$_{1c}$] level, and creatinine level) taken. Data were digitally collected using the KoboToolbox system developed by the Harvard Humanitarian Initiative [22].

Measures and Analytical Strategy

Primary Outcome of Interest

The main outcome of interest in this study was the cascade of care consisting of six bars: (1) the *prevalence* bar, (2) the *ever tested or screened* bar, (3) the *ever diagnosed* bar, (4) the *in care* bar, (5) the *in treatment* bar, and (6) the *under control* bar. A fixed denominator approach was used for constructing the cascades of care to identify the leakages between the stages of the care continuum [23]. Table 2 shows the definitions of each bar and describes the sources of the data extracted for analysis.

Table 2. Definitions of the cascade bars for type 2 diabetes (T2D).

Bars of the cascade of care for T2D and definitions	Source of data extracted for analysis
Prevalence of the target population living with T2D	
Participants having biochemical measurement of FBG[a] (capillary plasma value) level ≥126 mg/dL (7 mmol/L) and HbA$_{1c}$[b] level ≥6.5% [18,24,25]	• Measurement of FBG level • Measurement of HbA$_{1c}$ level
Participants reporting the use of drugs for T2D, irrespective of their biomarker values	• Response to the question, *Have you ever been told by a physician or other health worker that you have T2D?*
Number of people in the target population with T2D ever tested for T2D	
Patients classified as living with T2D having had FBG level tested in the last 3 years	• Response to the question, *Have you ever had your blood glucose level tested in the last 3 years?*
Number of those tested ever diagnosed for T2D	
Tested patients with T2D reporting ever being told by a physician or other health worker that they have T2D	• Response to the question, *Have you ever been told by a physician or other health worker that you have T2D?*
Number of those diagnosed in care	
Patients diagnosed with T2D reporting receiving treatment or care for their conditions at least once in the past 12 months	• Response to the question, *Did you get treatment or care for your T2D condition in the past 12 months?*
Number of those in care receiving treatment	

Patients with T2D in care reporting using drugs for T2D or insulin in the past 2 weeks	• Response to the question, *Are you currently receiving any of the following treatments for your T2D condition prescribed by a physician or other health care worker?*
	  • *Insulin*
	  • *Drugs (medication) that you have taken in the past 2 weeks*

Patients with T2D in care reporting following advice to lose weight, stop smoking, perform physical exercise, and be on a special prescribed diet	• Response to the question, *Are you currently receiving all of the following advice for your T2D condition prescribed by a physician or other health care worker?*
	  • *Special prescribed diet*
	  • *Advice to lose weight*
	  • *Advice to stop smoking*
	  • *Advice to start or perform more physical exercise*

Number of those receiving treatment with T2D under control	
Patients with T2D in treatment having HbA$_{1c}$ level <8% [18]	• Measurement of HbA$_{1c}$ level for known T2D

[a] FBG: fasting blood glucose.

[b] HbA$_{1c}$: glycated hemoglobin.

Secondary Outcomes of Interest

The secondary outcomes of interest were the factors associated with the undiagnosed status of participants living with T2D. We defined *person with undiagnosed status* as a person having biochemical measurements of FBG level ≥126 mg/dL and HbA_{1c} level ≥6.5% in our study but never being told by a physician or other health worker that they had T2D.

Explanatory Variables

The explanatory variables for this analysis included demographic characteristics, socioeconomic status, and the care initiatives (either individually or in coexistence). The demographic characteristics consisted of (1) age in years (40-49, 50-59, or ≥60); (2) sex (male or female); (3) educational attainment (none, primary school, secondary school, or higher); and (4) socioeconomic status (poorest, poor, medium, rich, or richest), which was measured using a household wealth index. To obtain the household wealth index, each household was interviewed using a 20-item questionnaire adapted from the 2014 Cambodia Demographic Health Survey [26]. This tool has been validated and widely used to classify household socioeconomic class [27]. Finally, the care initiative settings included (1) hospital-based care, (2) health center-based care, (3) health center-based care with context, (4) community-based care, and (5) the coexistence of the 3 care initiatives.

Statistical Analysis

We produced bar charts of the T2D cascades of care in accordance with the definitions provided in Table 2. We used bivariate analyses to compare the proportion of participants living with T2D without a diagnosis by participant characteristics. Subsequently, a multiple logistic regression analysis was used to identify the factors associated with the undiagnosed status. As we had only a limited number of explanatory variables in the bivariate analysis, we included all these variables in our initial multiple logistic regression analysis, regardless of the significance level. We additionally used a backward elimination method. Variables with the highest *P* value were eliminated from the model one by one. We retained all variables with a significance level of *P*<.05 in the final model. The statistical software Stata (version 14.2; StataCorp LLC) was used to perform the statistical analyses [28].

Ethics Approval

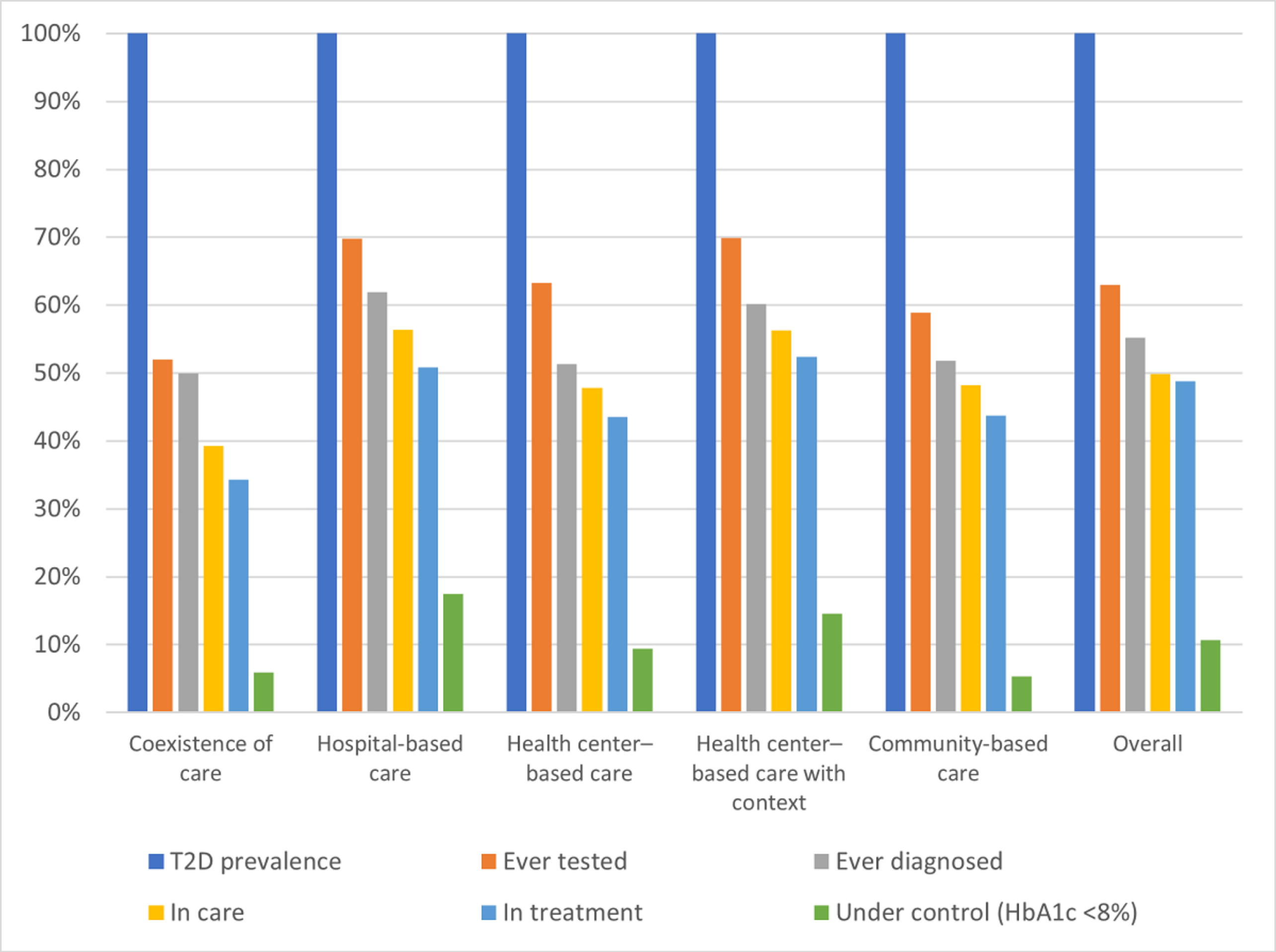
This study was approved by the National Ethics Committee for Health Research in Cambodia (NECHR; 105 NECHR) and by the institutional review board of the Institute of Tropical Medicine (Antwerp, Belgium; 1323/19). The study has also been registered as part of the Scale-up Integrated Care for Diabetes and Hypertension (SCUBY) project protocol at the ISRCTN Registry (ISRCTN41932064).

Results

Cascade of Care

Of the 5072 individuals participating in this study, 614 (12.11%) had raised blood glucose level (FBG level ≥126 mg/dL), and 560 (11.04%) met the definition of having T2D (Multimedia Appendix 1). Using the 560 individuals as the fixed denominator, we observed that 2 bars—*ever tested* and *under control*—had a substantial drop. Of the 560 individuals with T2D, only 353 (63%) had ever undergone a blood glucose level test in the last 3 years, 309 (55.2%) had ever been diagnosed as having T2D, 279 (49.8%) had received care in the past 12 months, and 273 (48.8%) had received insulin or antidiabetic medication in the past 2 weeks. In addition, only 130 (47.6%) of the 273 treated participants also received advice regarding a prescribed diet, weight loss, smoking cessation, and physical exercise. Only 10.7% (60/560) achieved blood glucose level control with the cutoff point of HbA_{1c} level <8% (Multimedia Appendix 1). Figure 1 shows the comparison of the cascades of care by setting with the overall cascade of care.

Figure 1. Type 2 diabetes (T2D) cascades of care by setting in 2020 in Cambodia. HbA_{1c}: glycated hemoglobin.



Participant Characteristics

As can be seen in Table 3, the majority of our study participants with T2D were those aged ≥50 years (476/560, 85%), female (418/560, 74.6%), and with low educational level (primary school level or lower: 490/560, 87.5%). Of the 560 participants living with T2D, 251 (44.8%) were undiagnosed. In this bivariate analysis, we observed that age (*P*<.001) and sex (*P*=.03) were associated with the undiagnosed status of participants living with T2D.

Table 3. Characteristics of the diagnosed and undiagnosed type 2 diabetes cases in 2020 in Cambodia.

Characteristics	Overall (N=560), n (%)	Diagnosed cases (n=309), n (%)	Undiagnosed cases (n=251), n (%)	*P* value
Age (years)				*<.001*[a]
40-49	84 (15)	29 (9.4)	55 (21.9)	
50-59	219 (39.1)	125 (40.4)	94 (37.5)	
≥60	257 (45.9)	155 (50.2)	102 (40.6)	
Sex				*.03*
Male	142 (25.4)	67 (21.7)	75 (29.9)	
Female	418 (74.6)	242 (78.3)	176 (70.1)	
Marital status				.53
Married or living with spouse	373 (66.6)	204 (66)	169 (67.3)	
Widowed or not living with spouse	179 (32)	102 (33)	77 (30.7)	
Never married and never lived together	8 (1.4)	3 (1)	5 (2)	
Educational level				.53
No formal education or less than primary education	164 (29.3)	95 (30.8)	69 (27.5)	
Primary education	326 (58.2)	179 (57.9)	147 (58.6)	
Secondary education or higher	70 (12.5)	35 (11.3)	35 (13.9)	
Household wealth quintile				.45
1 (poorest)	100 (17.9)	48 (15.5)	52 (20.8)	
2	102 (18.2)	54 (17.5)	48 (19.1)	
3	114 (20.3)	67 (21.7)	47 (18.7)	
4	113 (20.2)	67 (21.7)	46 (18.3)	
5 (richest)	131 (23.4)	73 (23.6)	58 (23.1)	
Care initiative				.22
Coexistence of care	102 (18.2)	51 (16.5)	51 (20.3)	
Community-based care	112 (20)	58 (18.8)	54 (21.5)	
Health center–based care	117 (20.9)	60 (19.4)	57 (22.7)	
Health center–based care with context	103 (18.4)	62 (20.1)	41 (16.3)	
Hospital-based care	126 (22.5)	78 (25.2)	48 (19.2)	

[a] Italicization indicates values that met the significance threshold (*P*<.05).

In the multiple logistic regression analysis (Table 4), being aged 40 to 49 years was associated with higher odds of not receiving the T2D diagnosis (adjusted odds ratio [AOR] 3.2, 95% CI 1.9-5.5; *P*<.001) compared with those aged ≥60 years. Male participants with T2D displayed higher odds of not being diagnosed (AOR 1.7, 95% CI 1.1-2.5; *P*<.001) than female participants living with T2D.

We also observed that being in the poorest category of the wealth quintile was associated with having higher odds of not being diagnosed with T2D (AOR 2.3, 95% CI 1.3-4.2; *P*=.005) than those in the richest category. Finally, the care initiative setting was also associated with the undiagnosed status of participants with T2D. Compared with those in the hospital-based care setting, higher odds of not being diagnosed were observed in the coexistence of care setting (AOR 1.9, 95% CI 1.1-3.3; *P*=.03), community-based care setting (AOR 1.9, 95% CI 1.1-3.3; *P*=.02), and health center–based care setting (AOR 2.1, 95% CI 1.2-3.6; *P*=.01).

It was observed that among the 309 participants diagnosed with T2D, 177 (57.3%) were diagnosed by a private provider, 121 (39.2%) by a public provider, and 11 (3.6%) by others. Table 5 compares public providers with private providers in each setting in terms of the proportion of participants with diagnosed T2D status and that of those with T2D control status. No statistical significance was observed.

Table 4. Factors associated with the undiagnosed status of participants with type 2 diabetes in 2020 in Cambodia.

Characteristics	Adjusted odds ratio (95% CI)	*P* value
Age (years)		
40-49	3.2 (1.9-5.5)	*<.001*[a]
50-59	1.2 (0.8-1.8)	.34
≥60	Reference	N/A[b]
Sex		
Male	1.7 (1.1-2.5)	*<.001*
Female	Reference	N/A
Household wealth quintile		
1 (poorest)	2.3 (1.3-4.2)	*.005*
2	1.5 (0.9-2.6)	.14
3	1.1 (0.6-1.8)	.84
4	1.0 (0.6-1.8)	.92
5 (richest)	Reference	N/A
Care initiative setting		
Coexistence of care	1.9 (1.1-3.3)	*.03*
Community-based care	1.9 (1.1-3.3)	*.02*
Health center–based care	2.1 (1.2-3.6)	*.01*
Health center–based care with context	1.2 (0.7-2.0)	.60
Hospital-based care	Reference	N/A

[a] Italicization indicates values that met the significance threshold (*P*<.05).

[b] N/A: not applicable.

Table 5. Distributions of participants with diagnosed type 2 diabetes (T2D) status and those with T2D control status categorized by health care provider in each setting.

Setting	Public provider, n (%)	Private provider, n (%)	*P* value
Participants with diagnosed T2D status (n=298[a])			.10
Coexistence of care	26 (8.7)	22 (7.4)	
Community-based care	16 (5.4)	41 (13.8)	
Health center–based care	22 (7.4)	37 (12.4)	
Health center–based care with context	27 (9.1)	35 (11.7)	
Hospital-based care	30 (10.1)	42 (14)	
Participants with T2D control status			
Coexistence of care (n=38)			.57
HbA_{1c}[b] level <8%	3 (7.9)	3 (7.9)	
HbA_{1c} level ≥8%	12 (31.6)	20 (52.6)	
Community-based care (n=53)			.88
HbA_{1c} level <8%	3 (5.7)	3 (5.7)	
HbA_{1c} level ≥8%	25 (47.1)	22 (41.5)	
Health center–based care (n=55)			.20
HbA_{1c} level <8%	1 (1.8)	10 (18.2)	
HbA_{1c} level ≥8%	12 (21.8)	32 (58.2)	
Health center–based care with context (n=57)			.23
HbA_{1c} level <8%	8 (14)	7 (12.3)	
HbA_{1c} level ≥8%	15 (26.3)	27 (47.4)	
Hospital-based care (n=67)			.19
HbA_{1c} level <8%	11 (16.4)	11 (16.4)	
HbA_{1c} level ≥8%	15 (22.4)	30 (44.8)	

[a] The category of *other provider* was removed from the analysis owing to its small proportion, which made the statistical test unreliable.

[b] HbA_{1c}: glycated hemoglobin.

Discussion

Principal Findings

This study used primary data from a cross-sectional survey to generate the cascade of care for the T2D care continuum in 5 purposively selected ODs in Cambodia. Overall, the cascade displayed substantial drops at the testing stage (207/560, 37%, loss from the *prevalence* bar) and at the control stage (213/560, 38%, loss from the *in treatment* bar), indicating that all selected settings, regardless of the care initiatives present, have limited capacity to detect people with T2D and control the condition (blood glucose level control) in those with T2D despite being in receipt of treatment. The findings were consistent with the T2D cascade analyses in other LMICs that displayed significant losses at the testing stage (also 37%) [18]. However, the drop between the treatment stage and the control stage observed in this study was much larger than that in the studies in other LMICs—only 15% in other LMICs compared with 38% (213/560) in this study [18]. With the cutoff point of HbA_{1c} level <8%, the proportion of those with T2D under control was 23% in other LMICs [18]; in this study, 10.7% (60/560) were considered as having achieved blood glucose level control. This is an exceptionally low rate, indicating that T2D in Cambodia is not being treated properly and adequately.

We disaggregated the cascades of care by study setting to observe the influence of the care initiatives. Unexpectedly, the coexistence of care setting displayed the worst cascade across all bars, whereas the hospital-based care setting had the best cascade among the 5 settings. This discovery was unexpectedly contradictory to the ICCC theoretical framework [3], calling into question the assumption underlying the ICCC framework that the combined care initiatives of health care organization and community represent an ideal context for integrated care for T2D and thereby would reduce leakages in the cascade. This suggests that the presence of health care infrastructure is not directly translated into improved care performance [29]. Implementation fidelity that focuses on the process of care implementation has to be taken into account [30]. Working mechanisms such as integrated care management across care levels and actors, the use of shared disease registries, and coordinated resources for self-management support and community education have to be in place for the coexistence of care to represent the ideal ICCC framework [3,16]. An investigation of the actual implementation of the care initiatives in these ODs was conducted in another study (Te V et al, unpublished data, July 2022). The investigation found that the 3 care initiatives were not implemented in an integrated way as intended in the written guideline [11] but in isolation, with limited interaction among them. The working mechanisms that facilitate integrated care for T2D in terms of shared necessary information and coordinated resources [16] were not observed. There was no proper system for following up patients for the continuity of care. The referral system among the communities, health centers, and referral hospitals was dysfunctional. The peer educator network in the OD with coexistence of care was not functioning optimally. The network had been handed over to the OD health authorities for governance, and technical or financial support from MoPoTsyo disappeared, rendering the network dysfunctional.

It should be noted that the care initiatives were not solely responsible for the provision of care for T2D in the selected ODs. On the basis of the same survey data, we found that, in general, health care use occurred dominantly in the private sector (78% among those seeking care in the 3 months preceding the survey), and referral hospitals were the common public health care facilities used by those with T2D and hypertension [31]. Therefore, our findings may not be fully attributed to the care initiatives. In the community-based care setting of OD Kong Pisey, only 12% (7/58) of the study participants were people living with T2D who were connected to the peer educator network, and only 4% (2/51) were connected to the peer educator network in the coexistence of care setting of OD Daunkeo. This may potentially undermine the effectiveness of community-based care. In a study based on MoPoTyso’s routine data, 43% of the people in the network achieved the median HbA_{1c} level of 7.1% [32]. In the hospital-based care setting of OD Samrong, only 5% (4/78) of the participants were seeking T2D care or treatment at the NCD clinic of the referral hospital in the 3 months preceding the survey, whereas at WHO PEN health centers in the health center–based care setting of OD Pearaing with high WHO PEN coverage, 3% (2/60) of the participants were identified seeking care for T2D.

Further statistical analysis, although not statistically significant, found that in all settings, except for the coexistence of care, private providers—who could not be fully incorporated into our study design owing to a lack of trustworthy information system in this sector—played a dominant role in diagnosing people with T2D. This suggests that the coexistence of care would increase the role of public health providers in the care continuum. In another study based on the same survey data, we found that the proportion of people with T2D seeking care at public health care facilities was higher than that of those with only hypertension or no condition [31]. This increased use of public health care facilities was also associated with a reduction in health care expenditure among patients, especially those in the poorest category of the wealth quintile who benefit from Health Equity Fund membership [31]. In a health system–level study, financial constraints have been found to be one of the main barriers to the T2D care continuum [6].

In this study, we found that 11.04% (560/5072) of the participants aged ≥40 years were identified as having T2D—of whom almost half (251/560, 44.8%) had not been diagnosed. This is a high prevalence rate because the overall prevalence rate of undiagnosed T2D in other LMICs has been reported to be 4.8% [18]. Predictors of being undiagnosed were being aged 40 to 49 years, being male, or falling in the poorest category of the wealth quintile. This suggests that more testing efforts are needed from the health system to reach people at risk for T2D, especially those from the aforementioned groups. A systematic review found that targeted screening was more cost-effective than universal screening [33]. A more convenient implementation arrangement for immediate diagnosis after testing should be put in place so that avoidable loss between these stages can be further minimized. In the national standard operating procedure [11], the WHO PEN health center staff are only allowed to perform the screening, whereas the diagnosis needs to be confirmed by the physician at the NCD clinic of the referral hospital. If the people who have been screened cannot have access to the diagnosis procedure at the referral hospital for some reason, the chances of not receiving prompt care or treatment increase. This requires strong coordination between the health centers and the NCD clinics, which has to be robust and supportive.

Limitations

First, despite using the primary data collected intentionally for the construction of the cascades of care, the sample size was not large enough to yield a sufficiently large number of patients with T2D who had achieved T2D control to enable us to assess the determinants of this particular bar. Second, the care initiatives, either individually or in coexistence, were not exclusively responsible for the provision of T2D care in each OD, thereby resulting in a weak connection between the presence of care initiatives and the cascade of care results of each study setting. We used the OD as a proxy variable to measure the effect of the care initiative, which in fact could mask a number of potential confounding contextual factors such as the dominant use of private services. In addition, in the Cambodian health system, the population is not confined to a particular public health facility in the catchment area. People can shop around freely, which means that patients may use services outside the catchment area of the facility. Third, the cross-sectional design did not allow us to determine the causal pathways leading to diabetes care outcomes, and the use of self-reported data in related sections could have produced biased results. A longitudinal study design with the collection of routine cohort data would enable us to address the limitations and evaluate the effectiveness of the different T2D care initiatives over time. This can be supported by digital health interventions. Systematic reviews have demonstrated the effectiveness of telemedicine via smartphone functions to provide self-care education, facilitate self-monitoring, produce the required treatment reminders, and collect feedback for health care professionals, which facilitates informed treatment recommendations [34,35]. In Cambodia, a study assessing the potential use of a wearable health monitor in the prevention and control of NCDs revealed that this health technology had the potential to support activities related to health promotion, patient follow-up and monitoring, and surveys of NCD risk factors, with positive user experiences and high levels of acceptance [36,37]. A digital health intervention that was tried among the MoPoTsyo networks produced valuable knowledge on pathways to address barriers to successful adoption in the Cambodian context [38].

Conclusions

This study provided an updated estimate of T2D prevalence among people aged ≥40 years (approximately 1 in 10 people) in Cambodia. The findings revealed that the unmet need for T2D care was large, particularly in the testing and control stages, indicating the need to substantially improve early detection and management of T2D in Cambodia. With almost half of the study participants with T2D undiagnosed (251/560, 44.8%) and thus unaware of their condition, early detection of people with T2D is an important first step that the health system needs to achieve to improve the T2D care continuum. We recommend rapid scale-up of T2D care components at public health facilities to increase the chances of the population with T2D of being tested, diagnosed, retained in care, and treated, as well as of achieving blood glucose control. At the same time, raising awareness and encouraging testing among the population at risk through a broad public health campaign should be one of the priorities. With advanced technology, a social media campaign has the potential to reach large parts of the population at low cost. Public health care use can reduce financial constraints among the population, particularly among those in the poorest category of the wealth quintile. We also recommend that within the context of resource constraints, specific groups considered susceptible (being male, being aged 40-49 years, or falling in the poorest category of the wealth quintile) should be especially targeted for testing through active community outreach activities because these groups are more likely to be unaware of their T2D condition. Adding care during off-hours for chronic conditions, including T2D, at public health facilities could increase access to care for male patients who are employed or busy during working hours. Future research should focus on evaluating the effectiveness of the different T2D care initiatives longitudinally with more diverse population groups from various settings. Given that digital health interventions have the potential to improve the prevention and control of NCDs while, at the same time, collecting longitudinal routine data vital for integrated care, feasibility and effectiveness studies of digital health interventions, such as telemedicine and mobile health, should be prioritized as a promising means to enable improvements along the T2D care continuum in Cambodia.

Acknowledgments

The authors would like to kindly acknowledge the Belgian government for funding this work through the VLIR-UOS project (KH2019TEA481A102) and the Fourth Funding Agreement Program of the Belgian Directorate-General for Development Cooperation and Humanitarian Aid and the Institute of Tropical Medicine (Antwerp). This study is part of the Scale-up Integrated Care for Diabetes and Hypertension (SCUBY) project, which received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825432.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors took part in conceptualizing the study design. VT and SC prepared the original draft. PI, JVO, VB, WVD, and EW provided feedback on the drafts. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Cascade of care for type 2 diabetes (T2D) in 2020 in Cambodia. HbA_{1c}: glycated hemoglobin.

[PNG File , 71 KB - publichealth_v9i1e41902_app1.png]

References

1. Diabetes around the world in 2021. International Diabetes Federation. 2021. URL: https://diabetesatlas.org/idfawp/resource-files/2021/11/IDFDA10-global-fact-sheet.pdf [accessed 2022-02-19]
2. University of Health Sciences. National Noncommunicable Disease Risk Factor Survey: STEPS Survey. Phnom Penh, Cambodia: University of Health Sciences; 2016.
3. Innovative care of chronic conditions: building blocks for action. World Health Organization. 2002. URL: https://apps.who.int/iris/bitstream/handle/10665/42500/WHO_NMC_CCH_02.01.pdf?sequence=1&isAllowed=y [accessed 2022-02-19]

4. Captieux M, Pearce G, Parke HL, Epiphaniou E, Wild S, Taylor SJ, et al. Supported self-management for people with type 2 diabetes: a meta-review of quantitative systematic reviews. BMJ Open 2018 Dec 14;8(12):e024262 [FREE Full text] [doi: 10.1136/bmjopen-2018-024262] [Medline: 30552277]
5. Shivashankar R, Kirk K, Kim WC, Rouse C, Tandon N, Narayan KM, et al. Quality of diabetes care in low- and middle-income Asian and Middle Eastern countries (1993-2012): 20-year systematic review. Diabetes Res Clin Pract 2015 Feb;107(2):203-223. [doi: 10.1016/j.diabres.2014.11.004] [Medline: 25529849]
6. Ong SE, Koh JJ, Toh SE, Chia KS, Balabanova D, McKee M, et al. Assessing the influence of health systems on type 2 diabetes mellitus awareness, treatment, adherence, and control: a systematic review. PLoS One 2018 Mar 29;13(3):e0195086 [FREE Full text] [doi: 10.1371/journal.pone.0195086] [Medline: 29596495]
7. Elissen AM, Steuten LM, Lemmens LC, Drewes HW, Lemmens KM, Meeuwissen JA, et al. Meta-analysis of the effectiveness of chronic care management for diabetes: investigating heterogeneity in outcomes. J Eval Clin Pract 2013 Oct;19(5):753-762. [doi: 10.1111/j.1365-2753.2012.01817.x] [Medline: 22372830]
8. Health information system master plan 2016-2020. Department of Planning and Health Information. Phnom Penh, Cambodia: Ministry of Health, Cambodia; 2017. URL: http://hismohcambodia.org/public/fileupload/carousel/HIS-MasterPlan-Nov2017.pdf [accessed 2019-12-17]
9. Annear PL, Grundy J, Ir P, Jacobs B, Men C, Nachtnebel M, et al. The kingdom of Cambodia: health system review. In: Annear PL, editor. World Health Organization. Manila, Philippines: Asia Pacific Observatory on Health Systems and Policies; 2015.
10. Te V, Ma S, Hok P, Long S, Chhea C, Ir P. The availability of diabetes services in Cambodian primary care facilities: an analysis of self-reported survey with health center staff. Cambodia J Public Health 2022 Jul 28;3(03):1-6 [FREE Full text]
11. National standard operating procedure for diabetes and hypertension management in primary care. Department of Preventive Medicine. Phnom Penh, Cambodia: Ministry of Health, Cambodia; 2019. URL: https://niph.org.kh/niph/uploads/library/pdf/GL240_PEN-SOP_DM_and_HBP_1_APR_2019-EN.pdf [accessed 2020-04-10]
12. Implementation tools: package of essential noncommunicable (PEN) disease interventions for primary health care in low-resource settings. World Health Organization. Geneva, Switzerland: WHO; 2013. URL: https://tinyurl.com/4tkabmye [accessed 2020-04-10]
13. Department of Preventive Medicine. Report of All Diabetes Clinics in Provinces. Phnom Penh, Cambodia: Ministry of Health, Cambodia; 2020.
14. Rao M, van Pelt M, LoGerfo J, Steinman LE, Heang H, Fitzpatrick AL. Utilization of diabetes management health care services and its association with glycemic control among patients participating in a peer educator-based program in Cambodia. PLoS One 2020 Jun 25;15(6):e0235037 [FREE Full text] [doi: 10.1371/journal.pone.0235037] [Medline: 32584854]
15. Te V, Wouters E, Buffel V, Van Damme W, van Olmen J, Ir P. Generation of cascades of care for diabetes and hypertension care continuum in Cambodia: protocol for a population-based survey protocol. JMIR Res Protoc 2022 Sep 02;11(9):e36747 [FREE Full text] [doi: 10.2196/36747] [Medline: 36053576]
16. Grover A, Joshi A. An overview of chronic disease models: a systematic literature review. Glob J Health Sci 2014 Oct 29;7(2):210-227 [FREE Full text] [doi: 10.5539/gjhs.v7n2p210] [Medline: 25716407]
17. HIV test-treat-retain cascade analysis: guide and tools. World Health Organization. Geneva, Switzerland: WHO; 2014. URL: https://applications.emro.who.int/docs/EMROPUB_2017_EN_20222.pdf?ua=1 [accessed 2019-12-17]
18. Manne-Goehler J, Geldsetzer P, Agoudavi K, Andall-Brereton G, Aryal KK, Bicaba BW, et al. Health system performance for people with diabetes in 28 low- and middle-income countries: a cross-sectional study of nationally representative surveys. PLoS Med 2019 Mar;16(3):e1002751 [FREE Full text] [doi: 10.1371/journal.pmed.1002751] [Medline: 30822339]
19. Ali MK, Bullard KM, Gregg EW, Del Rio C. A cascade of care for diabetes in the United States: visualizing the gaps. Ann Intern Med 2014 Nov 18;161(10):681-689. [doi: 10.7326/M14-0019] [Medline: 25402511]
20. Stokes A, Berry KM, Mchiza Z, Parker WA, Labadarios D, Chola L, et al. Prevalence and unmet need for diabetes care across the care continuum in a national sample of South African adults: evidence from the SANHANES-1, 2011-2012. PLoS One 2017 Oct 02;12(10):e0184264 [FREE Full text] [doi: 10.1371/journal.pone.0184264] [Medline: 28968435]
21. Janssens B, Van Damme W, Raleigh B, Gupta J, Khem S, Soy Ty K, et al. Offering integrated care for HIV/AIDS, diabetes and hypertension within chronic disease clinics in Cambodia. Bull World Health Organ 2007 Nov;85(11):880-885 [FREE Full text] [doi: 10.2471/blt.06.036574] [Medline: 18038079]
22. Kobo toolbox. Humanitarian Response. URL: https://www.humanitarianresponse.info/en/applications/kobotoolbox [accessed 2020-04-10]
23. Haber N, Pillay D, Porter K, Bärnighausen T. Constructing the cascade of HIV care: methods for measurement. Curr Opin HIV AIDS 2016 Jan;11(1):102-108. [doi: 10.1097/COH.0000000000000212] [Medline: 26545266]
24. Clinical practice guidelines for type 2 diabetes: a continuum of care for diabetes patients both with and without complications at NCD clinics/referral hospitals. Department of Preventive Medicine. Phnom Penh, Cambodia: Ministry of Health, Cambodia; 2015. URL: https://niph.org.kh/niph/uploads/library/pdf/GL236_DM_guideline_moh_en.pdf [accessed 2020-04-10]

25. Type 2 diabetes: practical targets and treatments. 3rd edition. Asian-Pacific Type 2 Diabetes Policy Group. Sydney, NSW, Australia: Health Communications Australia Pty Limited and In Vivo Communications Pty Limited; 2002. URL: https://apps.who.int/iris/bitstream/handle/10665/206942/0958073813_eng.pdf?sequence=1&isAllowed=y [accessed 2019-12-17]
26. National Institute of Statistics, Directorate General for Health, ICF International. Cambodia Demographic and Health Survey 2014. Phnom Penh, Cambodia/Rockville, MD: National Institute of Statistics, Directorate General for Health, Cambodia/ICF International, United States; 2015 Sep. URL: https://dhsprogram.com/pubs/pdf/fr312/fr312.pdf [accessed 2020-04-10]
27. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan 2006 Nov;21(6):459-468. [doi: 10.1093/heapol/czl029] [Medline: 17030551]
28. Stata/SE 14.2 for window. StataCorp LP. College Station, TX: StataCorp LP; 2016. URL: https://www.stata.com/products/windows/ [accessed 2020-04-10]
29. Service availability and readiness assessment (SARA): an annual monitoring system for service delivery. World Health Organization. Geneva, Switzerland: WHO; 2015. URL: https://apps.who.int/iris/handle/10665/149025 [accessed 2020-04-10]
30. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implement Sci 2007 Nov 30;2:40 [FREE Full text] [doi: 10.1186/1748-5908-2-40] [Medline: 18053122]
31. Chhim S, Te V, Buffel V, van Olmen J, Chham S, Long S, et al. Healthcare usage and expenditure among people with type 2 diabetes and/or hypertension in Cambodia: results from a cross-sectional survey. BMJ Open 2023 Jan 12;13(1):e061959 [FREE Full text] [doi: 10.1136/bmjopen-2022-061959] [Medline: 36635032]
32. Van Olmen J, Marie KG, Christian D, Clovis KJ, Emery B, Maurits VP, et al. Content, participants and outcomes of three diabetes care programmes in three low and middle income countries. Prim Care Diabetes 2015 Jun;9(3):196-202 [FREE Full text] [doi: 10.1016/j.pcd.2014.09.001] [Medline: 25281167]
33. Einarson TR, Bereza BG, Acs A, Jensen R. Systematic literature review of the health economic implications of early detection by screening populations at risk for type 2 diabetes. Curr Med Res Opin 2017 Feb;33(2):331-358. [doi: 10.1080/03007995.2016.1257977] [Medline: 27819150]
34. Aminuddin HB, Jiao N, Jiang Y, Hong J, Wang W. Effectiveness of smartphone-based self-management interventions on self-efficacy, self-care activities, health-related quality of life and clinical outcomes in patients with type 2 diabetes: a systematic review and meta-analysis. Int J Nurs Stud 2021 Apr;116:103286. [doi: 10.1016/j.ijnurstu.2019.02.003] [Medline: 30827741]
35. Correia JC, Meraj H, Teoh SH, Waqas A, Ahmad M, Lapão LV, et al. Telemedicine to deliver diabetes care in low- and middle-income countries: a systematic review and meta-analysis. Bull World Health Organ 2021 Mar 01;99(3):209-19B [FREE Full text] [doi: 10.2471/BLT.19.250068] [Medline: 33716343]
36. Liverani M, Ir P, Perel P, Khan M, Balabanova D, Wiseman V. Assessing the potential of wearable health monitors for health system strengthening in low- and middle-income countries: a prospective study of technology adoption in Cambodia. Health Policy Plan 2022 Sep 13;37(8):943-951 [FREE Full text] [doi: 10.1093/heapol/czac019] [Medline: 35262172]
37. Liverani M, Ir P, Wiseman V, Perel P. User experiences and perceptions of health wearables: an exploratory study in Cambodia. Glob Health Res Policy 2021 Sep 23;6(1):33 [FREE Full text] [doi: 10.1186/s41256-021-00221-3] [Medline: 34556184]
38. Steinman L, van Pelt M, Hen H, Chhorvann C, Lan CS, Te V, et al. Can mHealth and eHealth improve management of diabetes and hypertension in a hard-to-reach population? -Lessons learned from a process evaluation of digital health to support a peer educator model in Cambodia using the RE-AIM framework. Mhealth 2020 Oct 05;6:40 [FREE Full text] [doi: 10.21037/mhealth-19-249] [Medline: 33437836]

Abbreviations

AOR: adjusted odds ratio
FBG: fasting blood glucose
HbA$_{1c}$: glycated hemoglobin
ICCC: innovative care for chronic conditions
LMICs: low- and middle-income countries
NCD: noncommunicable disease
NECHR: National Ethics Committee for Health Research in Cambodia
OD: operational district
SCUBY: Scale-up Integrated Care for Diabetes and Hypertension
T2D: type 2 diabetes
WHO PEN: World Health Organization package of essential noncommunicable disease interventions
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 14.08.22; peer-reviewed by J LoGerfo, N Gandrakota; comments to author 07.03.23; revised version received 28.03.23; accepted 08.05.23; published 22.06.23.

Please cite as:
Te V, Chhim S, Buffel V, Van Damme W, van Olmen J, Ir P, Wouters E
Evaluation of Diabetes Care Performance in Cambodia Through the Cascade-of-Care Framework: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e41902
URL: https://publichealth.jmir.org/2023/1/e41902
doi:10.2196/41902
PMID:37347529

Original Paper

Dashboard With Bump Charts to Visualize the Changes in the Rankings of Leading Causes of Death According to Two Lists: National Population-Based Time-Series Cross-Sectional Study

Shu-Yu Tai[1], MD, DPhil; Ying-Chen Chi[2], DPhil; Yu-Wen Chien[3], MD, DPhil; Ichiro Kawachi[4], MD, DPhil; Tsung-Hsueh Lu[3], MD, DPhil

[1]Department of Family Medicine, Kaohsiung Municipal Ta-Tung Hospital and Kaohsiung Medical University Hospital; School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

[2]Department of Healthcare Information and Management, School of Health Technology, Ming Chuan University, Taoyuan, Taiwan

[3]Department of Public Health, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[4]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, United States

Corresponding Author:
Tsung-Hsueh Lu, MD, DPhil
Department of Public Health
College of Medicine
National Cheng Kung University
No. 1, Dah Hsueh Road, East District
Tainan, 701
Taiwan
Phone: 886 0928389971
Fax: 886 6 2359033
Email: robertlu@mail.ncku.edu.tw

Related Article:

This is a corrected version. See correction statement: https://publichealth.jmir.org/2024/1/e56014

Abstract

Background: Health advocates and the media often use the rankings of the leading causes of death (CODs) to draw attention to health issues with relatively high mortality burdens in a population. The National Center for Health Statistics (NCHS) publishes “Deaths: leading causes” annually. The ranking list used by the NCHS and statistical offices in several countries includes broad categories such as cancer, heart disease, and accidents. However, the list used by the World Health Organization (WHO) subdivides broad categories (17 for cancer, 8 for heart disease, and 6 for accidents) and classifies Alzheimer disease and related dementias and hypertensive diseases more comprehensively compared to the NCHS list. Regarding the data visualization of the rankings of leading CODs, the bar chart is the most commonly used graph; nevertheless, bar charts may not effectively reveal the changes in the rankings over time.

Objective: The aim of this study is to use a dashboard with bump charts to visualize the changes in the rankings of the leading CODs in the United States by sex and age from 1999 to 2021, according to 2 lists (NCHS vs WHO).

Methods: Data on the number of deaths in each category from each list for each year were obtained from the Wide-ranging Online Data for Epidemiologic Research system, maintained by the Center for Disease Control and Prevention. Rankings were based on the absolute number of deaths. The dashboard enables users to filter by list (NCHS or WHO) and demographic characteristics (sex and age) and highlight a particular COD.

Results: Several CODs that were only on the WHO list, including brain, breast, colon, hematopoietic, lung, pancreas, prostate, and uterus cancer (all classified as cancer on the NCHS list); unintentional transport injury; poisoning; drowning; and falls (all classified as accidents on the NCHS list), were among the 10 leading CODs in several sex and age subgroups. In contrast, several CODs that appeared among the 10 leading CODs according to the NCHS list, such as pneumonia, kidney disease, cirrhosis, and sepsis, were excluded from the 10 leading CODs if the WHO list was used. The rank of Alzheimer disease and related dementias and hypertensive diseases according to the WHO list was higher than their ranks according to the NCHS list. A marked increase in the ranking of unintentional poisoning among men aged 45-64 years was noted from 2008 to 2021.

Conclusions: A dashboard with bump charts can be used to improve the visualization of the changes in the rankings of leading CODs according to the WHO and NCHS lists as well as demographic characteristics; the visualization can help users make informed decisions regarding the most appropriate ranking list for their needs.

(JMIR Public Health Surveill 2023;9:e42149) doi:10.2196/42149

KEYWORDS

COVID-19; dashboard; data visualization; leading causes of death; mortality/trend; ranking; surveillance; cause of mortality; cause of death; monitoring; surveillance indicator; health statistics; mortality data

Introduction

Vital statistics can be used to monitor the vital signs of a nation [1]. Health advocates and the media often use the rankings of the leading causes of death (CODs) to draw attention to major health issues with relatively high mortality burdens in a population. The National Center for Health Statistics (NCHS) publishes “Deaths: leading causes” annually [2]. However, the ranking list used by the NCHS and statistical offices in several countries includes broad categories such as cancer, heart disease, and accidents [3-7]. The International Classification of Diseases and Health Related Problem Tenth Revision (ICD-10) has 89 subcategories in the broad cancer category, 41 subcategories in the broad heart disease category, and more than 100 subcategories in the broad accidents category [8]. Different subcategories within a broad category may have different levels of preventability, and therefore, may require different prevention strategies. For example, the 2019 version of the Organisation for Economic Co-operation and Development/Eurostat list of preventable and treatable CODs includes only 16 out of 89 cancer sites [9]. Stakeholders in public health policy require more specific and actionable information to help them set priorities.

To address the limitations of using broad categories for ranking, the World Health Organization (WHO) proposed a list in which cancer was divided into 17 cancer sites, heart disease into 8 specific forms of heart disease, and accidents into 6 unintentional injuries. Furthermore, more comprehensive COD categories are defined in the WHO list. For example, in the WHO list, the Alzheimer disease and related dementias (ADRD) category has ICD-10 codes F01 (vascular dementia), F03 (unspecified dementia), and G30 (Alzheimer disease), whereas in the NCHS list, ICD-10 code G30 is the only code available to classify ADRD-related mortality. In the WHO list, the hypertensive diseases category includes ICD-10 codes I10 (essential [primary] hypertension), I11 (hypertensive heart disease), I12 (hypertensive renal disease), I13 (hypertensive heart and renal disease), and I15 (secondary hypertension); however, in the NCHS list, ICD-10 codes I10, I12, and I15 are composed as one rankable category (essential hypertension and hypertensive renal disease) and I11 as well as I13 are included in the broad heart disease category. The WHO list has been routinely used to present the leading CODs by the WHO, the United Kingdom, Australia, and Italy [10-13]. However, the ranking profiles of leading CODs in the United States according to the WHO list remain unexplored.

Regarding the data visualization of the rankings of the leading CODs, the bar chart is the most commonly used chart [10-12,14]. However, it is difficult to reveal the changes in the rankings over time through bar charts. For example, the stacked bar chart by year was used by the UK Office for National Statistics; nevertheless, it was difficult to detect the changes in the ranking of a particular COD across years [11]. Bump charts have been recommended to represent the changes in the position of a given number of competing entities over a fixed duration [15]. Accordingly, in this study, we sought to design a dashboard with bump charts to visualize the changes in the rankings of the leading CODs in the Unites States by sex and age from 1999 to 2021 according to 2 lists (NCHS vs WHO).

Methods

The data on the number of deaths in each category were obtained from the Wide-ranging Online Data for Epidemiologic Research system maintained by the Center for Disease Control and Prevention for the period from 1999 to 2021 [16]. The NCHS list had 52 categories, and the WHO list had 65 categories. The name and ICD-10 codes of each category in the WHO list mapped to the NCHS list are presented in Table S1 in Multimedia Appendix 1. We followed the method used by the NCHS to determine the rank of each category. The NCHS determines the rank of each category on the basis of the number of deaths in each category.

We used Tableau to create a dashboard with a bump chart that can be filtered by list and demographic characteristics (ie, sex and age), which allows the viewers to select the dimension in which they are interested [15]. In the bump chart, each COD has a specific color, and the abbreviated name of each COD category appears in each circle. When the cursor points to a circle in the bump chart, a tooltip appears and indicates the number of deaths for that category, the proportion of total deaths that category accounts for, and the full name of the category. We also designed a filter to highlight the particular COD.

Ethical Considerations

This study used publicly available mortality data. All data used in this study were anonymized. This study was approved by the Institutional Review Board of National Cheng Kung University Hospital (B-EX-112-016) and the requirement of informed consent was waived by the abovementioned ethics committee.

Results

According to the NCHS list, the first and second leading CODs for both sexes of all ages were heart disease and cancer throughout the 23-year study period (Figure 1B). However, according to the WHO list, the first and second leading CODs were ischemic heart disease (IHD) and cerebrovascular disease from 1999 to 2002, IHD and lung cancer from 2003 to 2007, IHD and ADRD from 2008 to 2019, IHD and COVID-19 in 2020, and COVID-19 and IHD in 2021 (Figure 1A). In the WHO list, IHD and heart failure were each identified in the 10 leading CODs, whereas in the NCHS list, these 2 conditions were both classified as heart diseases and are not distinguished. Likewise, lung cancer and hematopoietic cancer were among the 10 leading CODs in the WHO list but were not distinguished in the NCHS list because in the NCHS list they were both classified into the broader cancer category (Figure 1A).

The 10 leading CODs for men of all ages according to the WHO and NCHS lists are presented in Figure 2. We noted an abrupt increase in the rank of unintentional poisoning according to the WHO list, from ninth in 2019 to seventh in 2020 and third in 2021 (Figure 2A). In contrast, the accidents category was the third leading COD from 2001 to 2019 and the fourth leading COD in 2020 and 2021 (Figure 2B). Several CODs such as cirrhosis, kidney disease, and pneumonia, which appeared among the 10 leading CODs according to the NCHS list (Figure 2B), were squeezed out from the 10 leading CODs if the WHO list was used.

The rankings of leading CODs among women of all ages in the WHO and NCHS lists are compared in Figure 3. According to the WHO list, ADRD ranked first from 2014 to 2020 and third in 2021 (Figure 3A). However, the Alzheimer disease category used in the NCHS list ranked fifth from 2002 to 2018, fourth from 2019 to 2020, and fifth in 2021 (Figure 3B). According to the WHO list, hypertensive diseases persistently increased in ranking, from eighth in 2009-2014 to seventh in 2015-2020, and to sixth in 2021 (Figure 3A). However, the category “hypertension” used in the NCHS list did not appear in the 10 leading CODs during the study period, except in 2019 and 2021, when it was ranked tenth (Figure 3B). By contrast, several CODs including pneumonia, kidney disease, and sepsis that appeared among the 10 leading CODs according to the NCHS list did not appear among the 10 leading CODs if the WHO list was used.

Figure 4 contrasts the rankings of the 10 leading CODs among men aged 45-64 years according to the WHO list with and without using a highlight filter. For example, by selecting unintentional poisoning in the highlight filter, we can better visualize the marked increase in the rankings of unintentional poisoning (Figure 4B).

Figure 1. Ranking of the 10 leading causes of death for both sexes of all ages in the United States according to (A) the World Health Organization (WHO) list and (B) the National Center for Health Statistics (NHCS) list. This dashboard is accessible [17]. Acci: accidents; AD: Alzheimer disease; ADRD: Alzheimer disease and related dementias; Ca: cancer; Cirr: chronic liver disease and cirrhosis; CLRD: chronic lower respiratory disease; CO19: COVID-19; CVD: cerebrovascular disease (stroke); DM: diabetes mellitus; HD: heart disease; HeaF: heart failure; HemC: hematopoietic cancer; Hypt: hypertensive diseases; IHD: ischemic heart disease; KidD: kidney diseases; LunC: lung cancer; Pneu: pneumonia; Pois: unintentional poisoning; Seps: sepsis; Suic: suicide; Urin: urinary diseases.

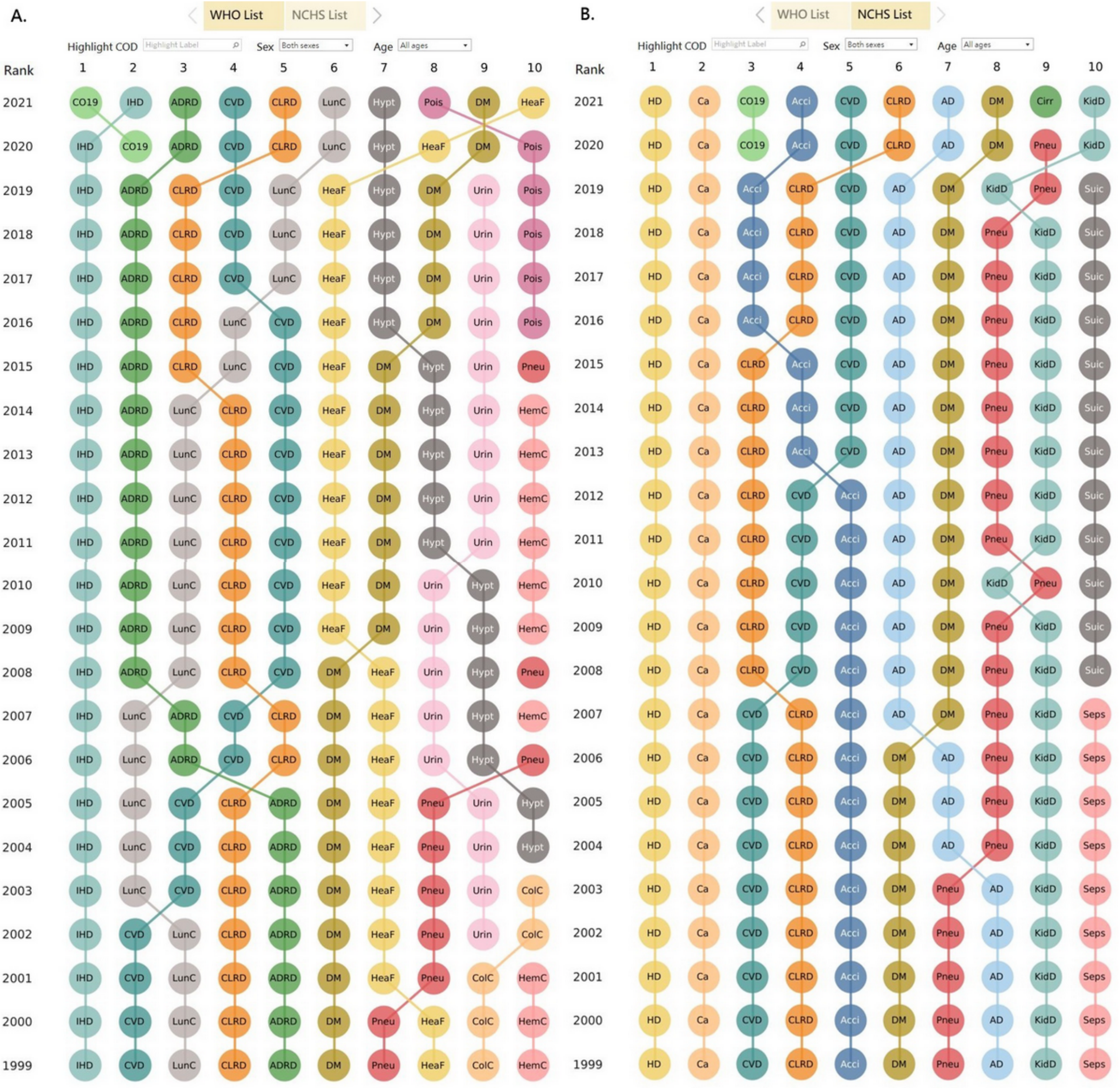
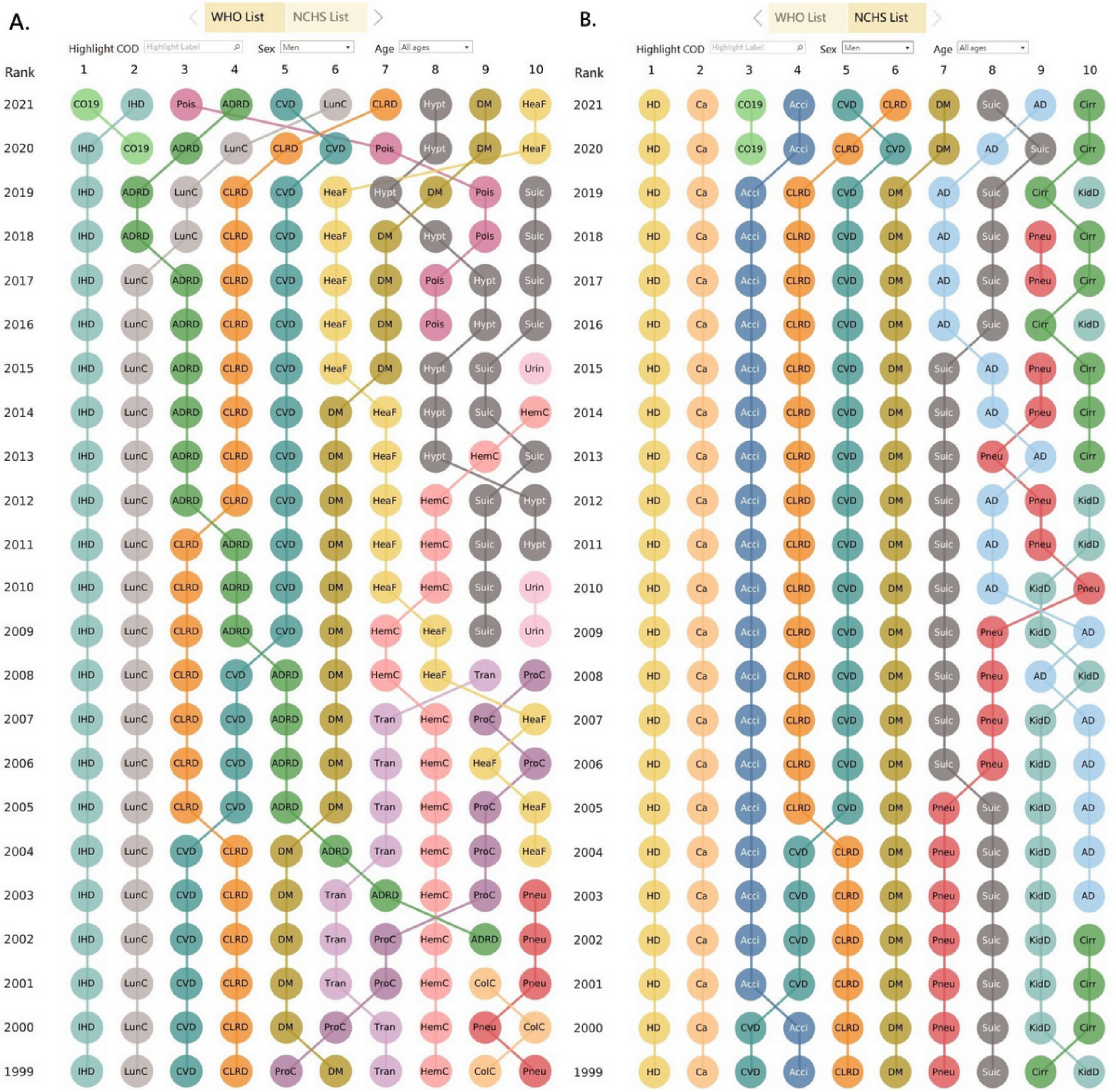


Figure 2. Ranking of the 10 leading causes of death for men of all ages in the United States according to (A) the World Health Organization (WHO) list and (B) the National Center for Health Statistics (NCHS) list. This dashboard is accessible [17]. Acci: accidents; AD: Alzheimer disease; ADRD: Alzheimer disease and related dementias; Ca: cancer; Cirr: chronic liver disease and cirrhosis; ColC: colon cancer; CLRD: chronic lower respiratory disease; CO19: COVID-19; CVD: cerebrovascular disease (stroke); DM: diabetes mellitus; HD: heart disease; HeaF: heart failure; HemC: hematopoietic cancer; Hypt: hypertensive diseases; IHD: ischemic heart disease; KidD: kidney diseases; LunC: lung cancer; Pneu: pneumonia; Pois: unintentional poisoning; ProC: prostate cancer; Suic: suicide; Tran: transportation accidents; Urin: urinary diseases.



A. WHO List

Highlight COD: Highlight Label; Sex: Men; Age: All ages

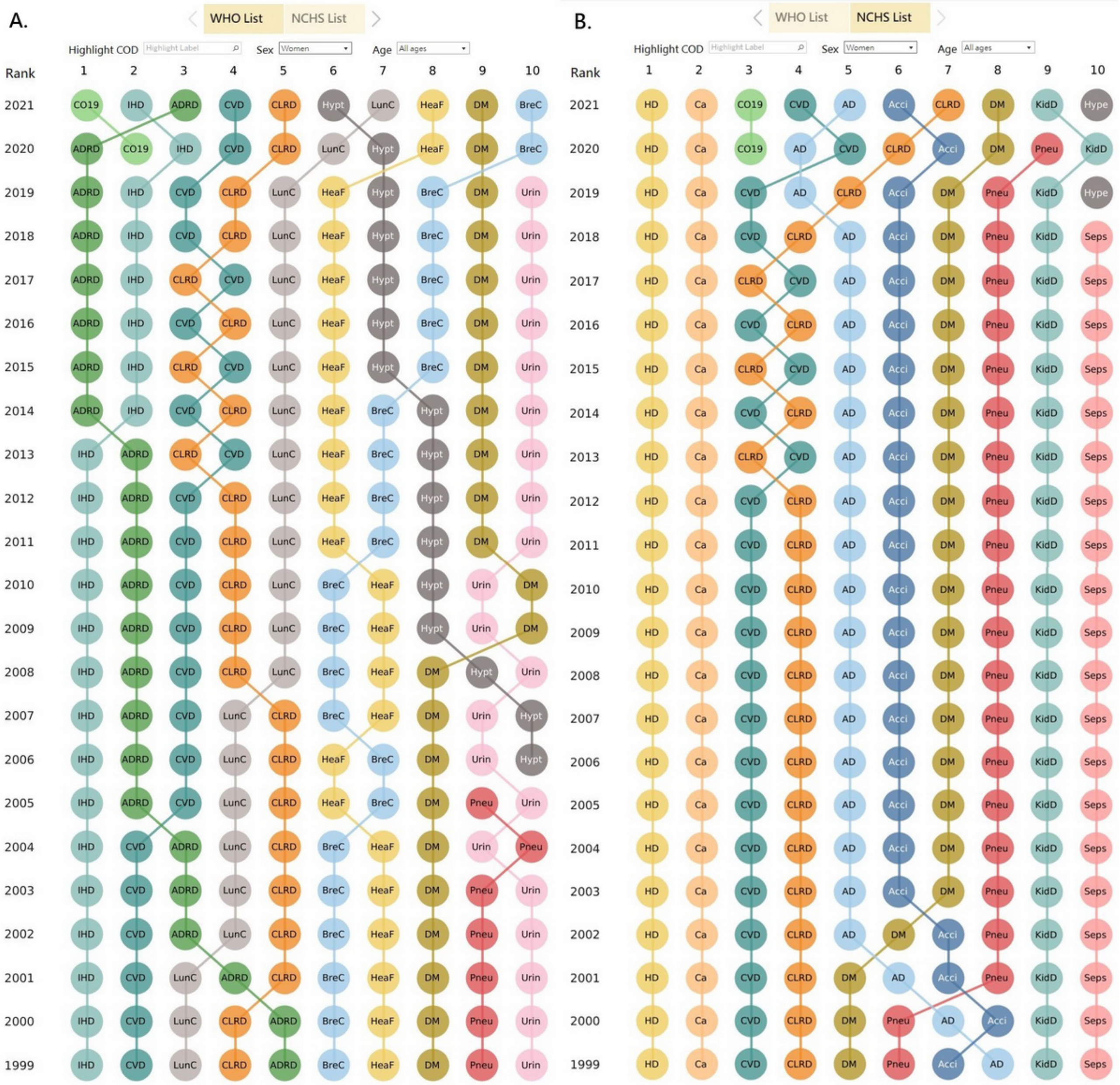
Rank	1	2	3	4	5	6	7	8	9	10
2021	CO19	IHD	Pois	ADRD	CVD	LunC	CLRD	Hypt	DM	HeaF
2020	IHD	CO19	ADRD	LunC	CLRD	CVD	Pois	Hypt	DM	HeaF
2019	IHD	ADRD	LunC	CLRD	CVD	HeaF	Hypt	DM	Pois	Suic
2018	IHD	ADRD	LunC	CLRD	CVD	HeaF	DM	Hypt	Pois	Suic
2017	IHD	LunC	ADRD	CLRD	CVD	HeaF	DM	Pois	Hypt	Suic
2016	IHD	LunC	ADRD	CLRD	CVD	HeaF	DM	Pois	Hypt	Suic
2015	IHD	LunC	ADRD	CLRD	CVD	HeaF	DM	Hypt	Suic	Urin
2014	IHD	LunC	ADRD	CLRD	CVD	DM	HeaF	Hypt	Suic	HemC
2013	IHD	LunC	ADRD	CLRD	CVD	DM	HeaF	Hypt	HemC	Suic
2012	IHD	LunC	ADRD	CLRD	CVD	DM	HeaF	HemC	Suic	Hypt
2011	IHD	LunC	CLRD	ADRD	CVD	DM	HeaF	HemC	Suic	Hypt
2010	IHD	LunC	CLRD	ADRD	CVD	DM	HeaF	HemC	Suic	Urin
2009	IHD	LunC	CLRD	ADRD	CVD	DM	HemC	HeaF	Suic	Urin
2008	IHD	LunC	CLRD	CVD	ADRD	DM	HemC	HeaF	Tran	ProC
2007	IHD	LunC	CLRD	CVD	ADRD	DM	Tran	HemC	ProC	HeaF
2006	IHD	LunC	CLRD	CVD	ADRD	DM	Tran	HemC	HeaF	ProC
2005	IHD	LunC	CLRD	CVD	ADRD	DM	Tran	HemC	ProC	HeaF
2004	IHD	LunC	CVD	CLRD	DM	ADRD	Tran	HemC	ProC	HeaF
2003	IHD	LunC	CVD	CLRD	DM	Tran	ADRD	HemC	ProC	Pneu
2002	IHD	LunC	CVD	CLRD	DM	Tran	ProC	HemC	ADRD	Pneu
2001	IHD	LunC	CVD	CLRD	DM	Tran	ProC	HemC	ColC	Pneu
2000	IHD	LunC	CVD	CLRD	DM	ProC	Tran	HemC	Pneu	ColC
1999	IHD	LunC	CVD	CLRD	ProC	DM	Tran	HemC	ColC	Pneu

B. NCHS List

Highlight COD: Highlight Label; Sex: Men; Age: All ages

Rank	1	2	3	4	5	6	7	8	9	10
2021	HD	Ca	CO19	Acci	CVD	CLRD	DM	Suic	AD	Cirr
2020	HD	Ca	CO19	Acci	CLRD	CVD	DM	AD	Suic	Cirr
2019	HD	Ca	Acci	CLRD	CVD	DM	AD	Suic	Cirr	KidD
2018	HD	Ca	Acci	CLRD	CVD	DM	AD	Suic	Pneu	Cirr
2017	HD	Ca	Acci	CLRD	CVD	DM	AD	Suic	Pneu	Cirr
2016	HD	Ca	Acci	CLRD	CVD	DM	AD	Suic	Cirr	KidD
2015	HD	Ca	Acci	CLRD	CVD	DM	Suic	AD	Pneu	Cirr
2014	HD	Ca	Acci	CLRD	CVD	DM	Suic	AD	Pneu	Cirr
2013	HD	Ca	Acci	CLRD	CVD	DM	Suic	Pneu	AD	Cirr
2012	HD	Ca	Acci	CLRD	CVD	DM	Suic	AD	Pneu	KidD
2011	HD	Ca	Acci	CLRD	CVD	DM	Suic	AD	Pneu	KidD
2010	HD	Ca	Acci	CLRD	CVD	DM	Suic	AD	KidD	Pneu
2009	HD	Ca	Acci	CLRD	CVD	DM	Suic	Pneu	KidD	AD
2008	HD	Ca	Acci	CLRD	CVD	DM	Suic	Pneu	AD	KidD
2007	HD	Ca	Acci	CLRD	CVD	DM	Suic	Pneu	KidD	AD
2006	HD	Ca	Acci	CLRD	CVD	DM	Suic	Pneu	KidD	AD
2005	HD	Ca	Acci	CLRD	CVD	DM	Pneu	Suic	KidD	AD
2004	HD	Ca	Acci	CVD	CLRD	DM	Pneu	Suic	KidD	AD
2003	HD	Ca	Acci	CVD	CLRD	DM	Pneu	Suic	KidD	AD
2002	HD	Ca	Acci	CVD	CLRD	DM	Pneu	Suic	KidD	Cirr
2001	HD	Ca	Acci	CVD	CLRD	DM	Pneu	Suic	KidD	Cirr
2000	HD	Ca	CVD	Acci	CLRD	DM	Pneu	Suic	KidD	Cirr
1999	HD	Ca	CVD	Acci	CLRD	DM	Pneu	Suic	Cirr	KidD

Figure 3. Ranking of the 10 leading causes of death for women of all ages in the United States according to (A) the World Health Organization (WHO) list and (B) the National Center for Health Statistics (NCHS) list. This dashboard is accessible [17]. Acci: accidents; AD: Alzheimer disease; ADRD: Alzheimer disease and related dementias; BreC: breast cancer; Ca: cancer; CLRD: chronic lower respiratory disease; CO19: COVID-19; CVD: cerebrovascular disease (stroke); DM: diabetes mellitus; HD: heart disease; HeaF: heart failure; Hypt: hypertensive diseases; IHD: ischemic heart disease; KidD: kidney diseases; LunC: lung cancer; Pneu: pneumonia; Seps: sepsis.



A.

Sex: Women
Age: All ages

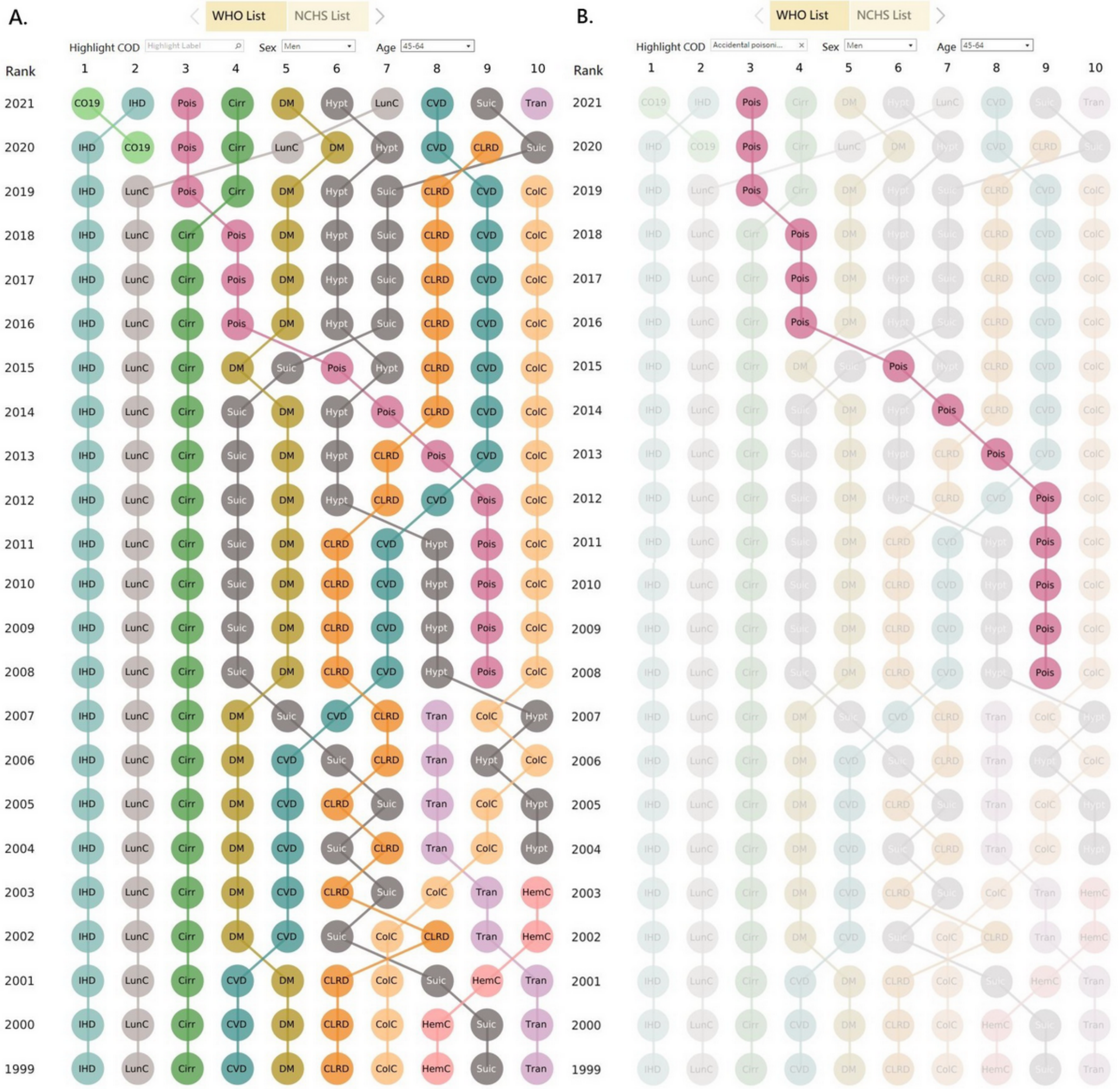
Rank	1	2	3	4	5	6	7	8	9	10
2021	CO19	IHD	ADRD	CVD	CLRD	Hypt	LunC	HeaF	DM	BreC
2020	ADRD	CO19	IHD	CVD	CLRD	LunC	Hypt	HeaF	DM	BreC
2019	ADRD	IHD	CVD	CLRD	LunC	HeaF	Hypt	BreC	DM	Urin
2018	ADRD	IHD	CVD	CLRD	LunC	HeaF	Hypt	BreC	DM	Urin
2017	ADRD	IHD	CLRD	CVD	LunC	HeaF	Hypt	BreC	DM	Urin
2016	ADRD	IHD	CVD	CLRD	LunC	HeaF	Hypt	BreC	DM	Urin
2015	ADRD	IHD	CLRD	CVD	LunC	HeaF	Hypt	BreC	DM	Urin
2014	ADRD	IHD	CVD	CLRD	LunC	HeaF	BreC	Hypt	DM	Urin
2013	IHD	ADRD	CLRD	CVD	LunC	HeaF	BreC	Hypt	DM	Urin
2012	IHD	ADRD	CVD	CLRD	LunC	HeaF	BreC	Hypt	DM	Urin
2011	IHD	ADRD	CVD	CLRD	LunC	HeaF	BreC	Hypt	DM	Urin
2010	IHD	ADRD	CVD	CLRD	LunC	BreC	HeaF	Hypt	Urin	DM
2009	IHD	ADRD	CVD	CLRD	LunC	BreC	HeaF	Hypt	Urin	DM
2008	IHD	ADRD	CVD	CLRD	LunC	BreC	HeaF	DM	Hypt	Urin
2007	IHD	ADRD	CVD	LunC	CLRD	BreC	HeaF	DM	Urin	Hypt
2006	IHD	ADRD	CVD	LunC	CLRD	HeaF	BreC	DM	Urin	Hypt
2005	IHD	ADRD	CVD	LunC	CLRD	HeaF	BreC	DM	Pneu	Urin
2004	IHD	CVD	ADRD	LunC	CLRD	BreC	HeaF	DM	Urin	Pneu
2003	IHD	CVD	ADRD	LunC	CLRD	BreC	HeaF	DM	Pneu	Urin
2002	IHD	CVD	ADRD	LunC	CLRD	BreC	HeaF	DM	Pneu	Urin
2001	IHD	CVD	LunC	ADRD	CLRD	BreC	HeaF	DM	Pneu	Urin
2000	IHD	CVD	LunC	CLRD	ADRD	BreC	HeaF	DM	Pneu	Urin
1999	IHD	CVD	LunC	CLRD	ADRD	BreC	HeaF	DM	Pneu	Urin

B.

Sex: Women
Age: All ages

Rank	1	2	3	4	5	6	7	8	9	10
2021	HD	Ca	CO19	CVD	AD	Acci	CLRD	DM	KidD	Hype
2020	HD	Ca	CO19	AD	CVD	CLRD	Acci	DM	Pneu	KidD
2019	HD	Ca	CVD	AD	CLRD	Acci	DM	Pneu	KidD	Hype
2018	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2017	HD	Ca	CLRD	CVD	AD	Acci	DM	Pneu	KidD	Seps
2016	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2015	HD	Ca	CLRD	CVD	AD	Acci	DM	Pneu	KidD	Seps
2014	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2013	HD	Ca	CLRD	CVD	AD	Acci	DM	Pneu	KidD	Seps
2012	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2011	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2010	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2009	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2008	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2007	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2006	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2005	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2004	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2003	HD	Ca	CVD	CLRD	AD	Acci	DM	Pneu	KidD	Seps
2002	HD	Ca	CVD	CLRD	AD	DM	Acci	Pneu	KidD	Seps
2001	HD	Ca	CVD	CLRD	DM	AD	Acci	Pneu	KidD	Seps
2000	HD	Ca	CVD	CLRD	DM	Pneu	AD	Acci	KidD	Seps
1999	HD	Ca	CVD	CLRD	DM	Pneu	Acci	AD	KidD	Seps

Figure 4. Ranking of 10 leading causes of death for men aged 45-64 years in the United States according to the World Health Organization (WHO) list (A) without using and (B) with using the highlight filter. This dashboard is accessible [17]. Cirr: chronic liver disease and cirrhosis; ColC: colon cancer; CLRD: chronic lower respiratory disease; CO19: COVID-19; CVD: cerebrovascular disease (stroke); DM: diabetes mellitus; HemC: hematopoietic cancer; Hypt: hypertensive diseases; IHD: ischemic heart disease; LunC: lung cancer; Pois: unintentional poisoning; Suic: suicide; Tran: transportation accidents.



Discussion

Principal Findings

The findings of this study indicate that profiles of the rankings of the 10 leading CODs differ depending on which list is used. Several CODs were among the 10 leading CODs in several sex and age subgroups when the WHO list was used, including brain, breast, colon, hematopoietic, lung, pancreas, prostate, and uterus cancer (all classified as cancer on the NCHS list); unintentional transport injuries; poisoning; drowning; and falls (which were all classified as accidents on the NCHS list). ADRD and hypertensive diseases, which have more comprehensive definitions in the WHO list, were ranked higher in the WHO list than similar categories in the NCHS list. By contrast, some CODs that appeared among the 10 leading CODs according to the NCHS list, such as cirrhosis, kidney disease, sepsis, and pneumonia, were squeezed out of the 10 leading CODs according to the WHO list. Through the use of a dashboard with bump charts, users can select a list, stratify the data by age and sex, and use a highlight filter to visualize the changes in the rankings of a particular COD across years.

The main reason for the differences between the lists in the COD rankings was the number of categories used to rank CODs. The NCHS list had 52 categories, and the WHO list had 65 categories (Table S1 in Multimedia Appendix 1). Selecting which list to use involves a trade-off. Some new specific CODs that existed in the 10 leading CODs according to the WHO list could not be revealed if the NCHS list was used. On the contrary, some CODs that appeared in the 10 leading CODs according to the NCHS would be squeezed out if the WHO list were used. The use of a more comprehensive definition for ADRD and hypertensive diseases might sacrifice the specificity, meaning that some individuals with ADRD diagnosis might be false positives. To obtain the most comprehensive information for health policy decision-making, we recommend using both lists.

Several studies have examined the changes in the rankings of the leading CODs in the United States [18-20]. Because the categories in the NCHS list are relatively broad, prominent changes cannot be detected, except for accidents, which increased in rank from fifth in 2011 to third in 2018 [19]. The pattern of changes in the leading CODs in the United States according to the WHO list differed from that in Japan, Korea, and Taiwan [7]. According to the WHO list, the leading COD from 1999 to 2020 in the United States was IHD, whereas in Japan, Korea, and Taiwan, the leading COD in most of the years during this period was stroke [7].

Strengths and Limitations

This study has several strengths. First, this study was the first to use the WHO list to illustrate changes in the ranking of the leading CODs in the United States. Using the WHO list enabled us to detect several drastic increases in the ranks of some leading CODs, which could not be revealed according to the NCHS list. The second strength of this study was the use of bump charts, which aid in visualizing the changes in rankings of specific CODs in a ladder-like manner. The third strength was the creation of a dashboard, which enables viewers to select the specific demographic group they wish to examine using either the NCHS or the WHO list (or another list can be added) and highlight the changes of a particular COD.

However, several limitations should be considered when interpreting the findings of this study. First, the changes in the relative position of each rank on the bump charts did not correspond to the extent of the changes in the proportions of total deaths accounted for by each category. For example, ADRD increased in ranking between 2013 and 2020 (from second to first) but accounted for fewer proportion of deaths in 2020 than it did in 2013 (10.9% vs 12.1%, respectively). Therefore, an increase (or decrease) in ranking does not indicate an increase (or decrease) in either burden or risk. Second, although the WHO list uses more specific categories compared to the NCHS list, several categories are still relatively broad. For example, unintentional poisoning comprises ten 3-digit ICD-10 codes (X40-X49) that each correspond to a different drug or chemical. Third, both ranking lists are limited in terms of detail and are not comprehensive, meaning that not all deaths can be categorized and ranked. The increase in the number of deaths in some emergent diseases within the residual category could not be detected using rankings of leading CODs.

Conclusions

The rankings of the leading CODs constitute a highly popular health statistic used to convey relative mortality burdens and are widely used by media and health advocates. A dashboard with bump charts can be used to improve visualizing the changes in the rankings of the leading CODs according to the WHO and NCHS lists as well as the demographic characteristics; it can also help users make informed decisions regarding the most appropriate ranking list for their needs.

Acknowledgments

The authors would like to thank Ms Pai-Huan Lin for the data analysis. This study was funded by the Ministry of Science and Technology, Taiwan (grant 111-2314-B-006-030). The funders had no role in the design of the study, collection and analysis of data, decision to publish, or preparation of the manuscript.

Data Availability

The data that support the findings of this study are publicly available [17].

Authors' Contributions

SYT and THL conceived the study design. YCC, YWC, and IK participated in the collection and analysis of data. SYT and THL drafted the manuscript. All authors contributed to the interpretation of the data and critically revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
International Classification of Diseases, Tenth Revision codes for the World Health Organization ranking list mapped to the National Center for Health Statistics ranking list.
[DOCX File , 36 KB - publichealth_v9i1e42149_app1.docx]

References

1. Gerberding JL. Measuring pandemic impact: vital signs from vital statistics. Ann Intern Med 2020 Dec 15;173(12):1022-1023 [FREE Full text] [doi: 10.7326/M20-6348] [Medline: 32915634]
2. Heron M. National Vital Statistics Reports. Deaths: Leading causes for 2019. 2021 Jun 21. URL: https://www.cdc.gov/nchs/data/nvsr/nvsr70/nvsr70-09-508.pdf [accessed 2023-04-03]
3. Griffiths C, Rooney C, Brock A. Leading causes of death in England and Wales--how should we group causes? Health Stat Q 2005(28):6-17. [Medline: 16315552]
4. Becker R, Silvi J, Ma Fat D, L'Hours A, Laurenti R. A method for deriving leading causes of death. Bull World Health Organ 2006 Apr 01;84(4):297-304 [FREE Full text] [doi: 10.2471/blt.05.028670] [Medline: 16628303]
5. Anderson R. Coding and classifying causes of death: Trends and international differences. In: International Handbook of Adult Mortality. Dordrecht, Netherlands: Springer; 2011:467-489.
6. Tai S, Lu T. Why was COVID-19 not the first leading cause of death in the United States in 2020? Rethinking the ranking list. Am J Public Health 2021 Dec;111(12):2096-2099. [doi: 10.2105/ajph.2021.306545]
7. Tai S, Cheon S, Yamaoka Y, Chien Y, Lu T. Changes in the rankings of leading causes of death in Japan, Korea, and Taiwan from 1998 to 2018: a comparison of three ranking lists. BMC Public Health 2022 May 10;22(1):926 [FREE Full text] [doi: 10.1186/s12889-022-13278-7] [Medline: 35538508]
8. International Statistical Classification of Diseases and Related Health Problems 10th Revision. World Health Organization. URL: https://icd.who.int/browse10/2019/en [accessed 2023-04-03]
9. Avoidable mortality: OECD/Eurostat lists of preventable and treatable causes of death. OECD/Eurostat. URL: https://tinyurl.com/ye6c3sfw [accessed 2023-04-03]
10. The top 10 causes of death. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [accessed 2023-04-03]
11. Leading causes of death, UK 2001 to 2018, 2020. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/articles/leadingcausesofdeathuk/2001to2018 [accessed 2023-04-03]
12. Deaths in Australia. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/life-expectancy-death/deaths-in-australia/contents/leading-causes-of-death [accessed 2023-04-03]
13. Istat. Leading causes of death in Italy, 2012. Istituto Nazionale di Statistica. 2014. URL: https://www.istat.it/it/archivio/140877 [accessed 2023-04-03]
14. Tejada-Vera B. Leading causes of death: United States, 19–2017. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/nchs/data-visualization/mortality-leading-causes/index.htm [accessed 2023-04-03]
15. Acharya S, Chellappan S. Pro Tableau: A Step-by-Step Guide. Pune, India: Apress; 2017.
16. Wide-ranging ONline Data for Epidemiologic Research (CDC WONDER). Centers for Disease Control and Prevention. URL: https://wonder.cdc.gov/ [accessed 2023-02-22]
17. Tableau Public. Leading causes of death in the United States. URL: https://public.tableau.com/app/profile/robert.lu/viz/LeadingCausesofDeathintheUnitedStates_16792025673420/bumpchart [accessed 2023-06-20]
18. Chang M, Moonesinghe R, Athar H, Truman B. Trends in disparity by sex and race/ethnicity for the leading causes of death in the United States-1999-2010. J Public Health Manag Pract 2016;22 Suppl 1:S13-S24. [doi: 10.1097/PHH.0000000000000267] [Medline: 25946701]
19. Rana JS, Khan SS, Lloyd-Jones DM, Sidney S. Changes in mortality in top 10 causes of death from 2011 to 2018. J Gen Intern Med 2021 Aug 23;36(8):2517-2518 [FREE Full text] [doi: 10.1007/s11606-020-06070-z] [Medline: 32705476]
20. Shiels MS, Haque AT, Berrington de González A, Freedman ND. Leading causes of death in the US during the COVID-19 pandemic, March 2020 to October 2021. JAMA Intern Med 2022 Aug 01;182(8):883-886. [doi: 10.1001/jamainternmed.2022.2476] [Medline: 35788262]

Abbreviations

ADRD: Alzheimer disease and related dementias
COD: cause of death
ICD-10: International Classification of Diseases, Tenth Revision
IHD: ischemic heart disease
NCHS: National Center for Health Statistics
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 24.08.22; peer-reviewed by R Anderson, M Kapsetaki; comments to author 10.02.23; revised version received 03.04.23; accepted 24.05.23; published 27.06.23.

Please cite as:
Tai SY, Chi YC, Chien YW, Kawachi I, Lu TH
Dashboard With Bump Charts to Visualize the Changes in the Rankings of Leading Causes of Death According to Two Lists: National Population-Based Time-Series Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e42149
URL: https://publichealth.jmir.org/2023/1/e42149
doi:10.2196/42149
PMID:37368475

Original Paper

Effect of Rapid Urbanization in Mainland China on the Seasonal Influenza Epidemic: Spatiotemporal Analysis of Surveillance Data From 2010 to 2017

Hao Lei[1*], PhD; Nan Zhang[2*], PhD; Beidi Niu[1], BS; Xiao Wang[3], MPH; Shenglan Xiao[3], PhD; Xiangjun Du[3], PhD; Tao Chen[4], PhD; Lei Yang[4], MASc; Dayan Wang[4], PhD; Benjamin Cowling[5], PhD; Yuguo Li[6], PhD; Yuelong Shu[3*], PhD

[1]School of Public Health, Zhejiang University, Hangzhou, China
[2]Key Laboratory of Green Built Environment and Energy Efficient Technology, Beijing University of Technology, Beijing, China
[3]School of Public Health, Shenzhen Campus, Sun Yat-sen University, Shenzhen, China
[4]Key Laboratory for Medical Virology, Chinese Center for Disease Control and Prevention, National Health Commission, Beijing, China
[5]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong
[6]Department of Mechanical Engineering, The University of Hong Kong, Pokfulam, Hong Kong
[*]these authors contributed equally

Corresponding Author:
Hao Lei, PhD
School of Public Health
Zhejiang University
866 Yu-hang-tang Road
Hangzhou, 310058
China
Phone: 86 0571 8707 6051
Email: leolei@zju.edu.cn

Abstract

Background: The world is undergoing an unprecedented wave of urbanization. However, the effect of rapid urbanization during the early or middle stages of urbanization on seasonal influenza transmission remains unknown. Since about 70% of the world population live in low-income countries, exploring the impact of urbanization on influenza transmission in urbanized countries is significant for global infection prediction and prevention.

Objective: The aim of this study was to explore the effect of rapid urbanization on influenza transmission in China.

Methods: We performed spatiotemporal analyses of province-level influenza surveillance data collected in Mainland China from April 1, 2010, to March 31, 2017. An agent-based model based on hourly human contact–related behaviors was built to simulate the influenza transmission dynamics and to explore the potential mechanism of the impact of urbanization on influenza transmission.

Results: We observed persistent differences in the influenza epidemic attack rates among the provinces of Mainland China across the 7-year study period, and the attack rate in the winter waves exhibited a U-shaped relationship with the urbanization rates, with a turning point at 50%-60% urbanization across Mainland China. Rapid Chinese urbanization has led to increases in the urban population density and percentage of the workforce but decreases in household size and the percentage of student population. The net effect of increased influenza transmission in the community and workplaces but decreased transmission in households and schools yielded the observed U-shaped relationship.

Conclusions: Our results highlight the complicated effects of urbanization on the seasonal influenza epidemic in China. As the current urbanization rate in China is approximately 59%, further urbanization with no relevant interventions suggests a worrisome increasing future trend in the influenza epidemic attack rate.

(JMIR Public Health Surveill 2023;9:e41435) doi:10.2196/41435

KEYWORDS

seasonal influenza; attack rate; urbanization; urban population; human contact; agent-based model; influenza; seasonal flu; spatiotemporal; epidemic; disease transmission; disease spread; epidemiology; influenza transmission; epidemics

Introduction

The influenza virus is mainly spread within close ranges rather than over long distances [1,2]. The distance from an influenza-infected individual may significantly affect the potential of influenza transmission. Social distancing is a well-recognized nonpharmaceutical intervention that could reduce the transmission of viruses from infected to susceptible individuals because of the increased physical distance between these individuals or reduced frequency of congregation in socially dense community settings such as schools or workplaces [3]. Urbanization is a complex phenomenon that is associated with population density in an urban area. In this study, urbanization rate, defined as the percentage of population living in an urban area, was used to characterize the urbanization process. Urbanization can worsen an epidemic through a variety of mechanisms [4]. Studies on low-income nations have clearly documented the association between epidemic severity and the combination of high-density living conditions, extreme poverty, and poor sanitation [5]. With respect to communicable diseases such as influenza, which is mainly spread within close ranges [1,2], the population concentrations in an urban area can lead to more rapid proliferation of infectious diseases [6]. Urbanization affects the seasonal epidemic transmission in highly urbanized countries [4,7]. In the United States with an urbanization rate of 83% in 2021 [8], cities in urban areas with higher population densities had a higher influenza transmission potential [7]. In Australia with an urbanization rate of 88% in 2021 [8], the peak influenza infection rate and peak prevalence increased steadily as the urban population density increased [4]. However, the impact of urbanization on the influenza epidemic in urban countries such as China with urbanization rates at 63% and India with urbanization rates at 35% in 2021, where about 70% of the population live, is less well-understood. In addition, in highly urbanized countries, urban living has been the norm in the last century. Therefore, the direct effects of further increases in the urban population density on the potential seasonal influenza epidemic might be less obvious [4]. In this study, attack rate, defined as the cumulative incidence rates of seasonal influenza during the study period, was used to characterize the seasonal influenza activities. Since urbanization will continue in urban countries in the near future, understanding the impact of urbanization on influenza transmission is essential for precise influenza prediction and prevention and for building a healthy society.

As a low-income country, China has experienced an unprecedented wave of rapid urbanization, as reported by the rapid increase in the urbanization rate from 17.9% in 1978 to 58.5% in 2017 [9]. The urbanization rate also varies significantly between provinces in China, thereby providing a unique opportunity to study the impact of urbanization on the influenza epidemic at different levels of urbanization within a single health care system. In China, urbanization has led not only to a rapid increase in the urban population density but also to a sharp decrease in the household size [10]. As households, schools (workplaces), and communities have been identified as the primary contexts of influenza transmission [11,12], the impact of urbanization on seasonal influenza in China may be complicated.

Dynamic models such as the susceptible-infected-recovered models have been developed to understand the dynamics of influenza transmission [7]. However, these models do not study the dynamics of influenza transmission at the individual level [13]. In highly urbanized countries, urban living has been the norm in the last century; thus, the susceptible-infected-recovered model could be used to study the impact of urbanization on influenza transmission [7]. However, the impact of urbanization on the influenza epidemic in urbanizing countries is complicated. In this study, to characterize the complicated impact of urbanization on human contact behaviors and population age structures in China, we used an agent-based model in which each individual or group of people was defined as an agent. The agent-based model based on human contact behavior has been used previously to model infectious disease transmission [11,14,15]. We also used similar models previously to successfully simulate influenza transmission in China [12,16].

In this study, we aimed to explore the effects of urbanization on seasonal influenza epidemic intensities in urban Mainland China (except Hong Kong, Macao, and Taiwan) according to the spatiotemporal analyses and modelling study of influenza surveillance data collected in Mainland China from 2010 to 2017.

Methods

Data Sources

We used weekly reports of influenza surveillance data from the Chinese National Influenza Center. The study data set was based on the number of specimens tested at 554 sentinel hospitals located in the urban areas of 31 provinces of Mainland China (including autonomous regions and municipalities). After excluding Tibet from the analysis due to the incompleteness of data, 30 provinces were evaluated. Although the surveillance data collection began in 2004, we focused on the period from April 1, 2010, to March 31, 2017, as the surveillance network was improved and expanded after the 2009 influenza A/H1N1 pandemic [17]. Shu et al [18] provide additional details about the Chinese influenza surveillance system. The data set of the influenza surveillance system in China included the number of visits at each hospital, number of cases of influenza-like illness (ILI), number of specimens tested, and number of laboratory-confirmed cases of influenza A (H1N1, H3N2, and pdmH1) and B (Yamagata and Victoria). ILI cases were identified based on a standard case definition: body temperature >38 °C, either cough or sore throat, and the absence of an alternative diagnosis. Several indicators were defined to characterize influenza activity in China. First, the ILI rate was defined as the number of ILI cases divided by the number of visits. Second, the influenza viral positive rate was defined as the number of laboratory-confirmed influenza cases divided by the number of specimens tested. In accordance with earlier studies [18-20], in this study, we defined a proxy for the weekly incidence rate (henceforth, incidence rate) of influenza, which is the multiplication of the ILI rate and the influenza viral positive rate. The incidence rate more precisely represents the incidence of influenza infection.

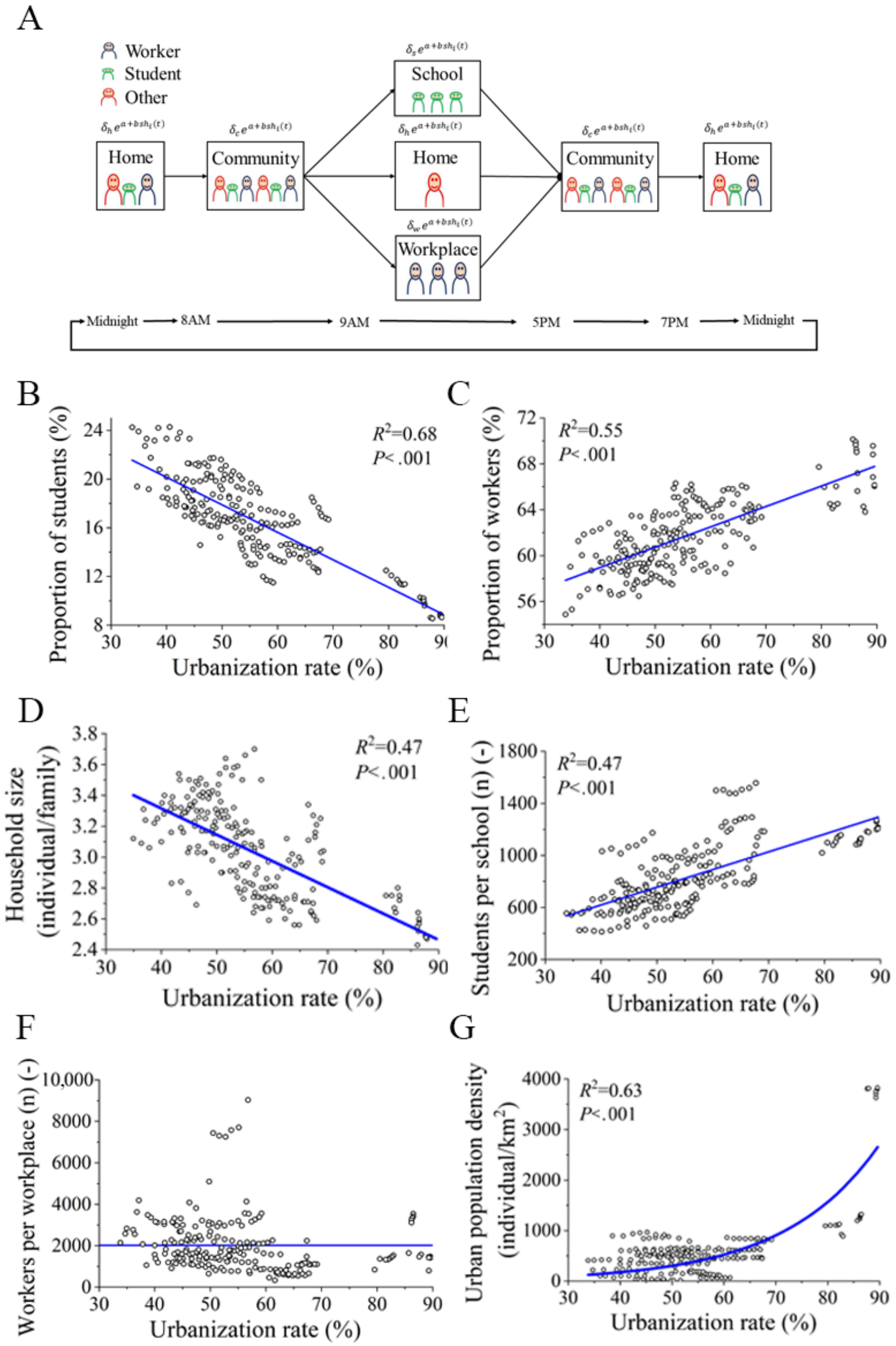
The influenza epidemic in Northern China intensifies in winter-spring months, while that in Southern China shows a semiannual cyclic pattern with clear peaks in both summer and winter [21,22]. We therefore defined an epidemiological annual cycle as the period from April 1 to March 31 of the following year [21], and we divided the annual cycle into summer months and winter-spring months, which included calendar weeks 14-39 (26 weeks in total) and weeks 40-13 of the following year (26 or 27 weeks in total), respectively [21]. The attack rate of the winter-spring or summer influenza epidemic in each province is defined as the cumulative weekly incidence rates during the half-year time (25 weeks in total): 12 weeks within the peak weeks since the median duration of an influenza epidemic wave is 15.6-22.5 weeks [21]. Province-level data on urbanization, urban areas, populations, household and school sizes, and number of students, schools, and factories from year 2010 to year 2016 were obtained from the China Statistical Yearbook [9], China City Statistical Yearbook [23], and National Bureau of Statistics of China [24].

Modelling

We used an agent-based model to simulate the influenza transmission dynamics in provinces with different urbanization rates. Our simulation strategy is similar to that reported previously in the literature [11]. Similar models have been used to simulate influenza transmission in Beijing [25] and Hong Kong [16]. In this study, we simplified the model and considered 4 types of indoor spaces: homes, schools, workplaces, and communities (ie, public spaces). These 4 types of indoor spaces are known as the primary contexts of influenza transmission [11,12]. Populations were divided into 3 groups: workers, students, and others.

Figure 1A presents the assumed typical daily movements of 3 groups of people. Although real human movement patterns are much more complicated, we considered the simulated movement pattern to represent a general daily schedule. The time step in the simulation was set to 1 hour. For each case, the simulation was performed with 100 replications to account for randomness in the status of each individual (infectious, susceptible/removed). Differences in human movement behaviors between weekdays and weekends were not considered. The percentage of students decreased significantly with urbanization (Figure 1B). The number of workers (N_W) was estimated based on the number of people aged between 18 and 60 years ($N_{18\text{-}60}$), the number of students in universities (N_{SU}), and the unemployment rate (R_u) as $N_W = (N_{18\text{-}60} - N_{SU}) * (1 - R_u)$. The percentage of workers increased significantly with urbanization (Figure 1C). The remaining populations were classified as others. Influenza transmission is affected by the number of students, workers, and others, as well as by the distributions of these populations in each indoor space. The province-level distribution of household sizes at different urbanization levels is shown in Figure 1D. Notably, the school size increased with urbanization (Figure 1E). As we did not observe a significant linear relationship between urbanization and the number of workers per workplace (Figure 1F), we adopted a mean value of 2020 workers per workplace at all urbanization levels. With urbanization, the urban population density increased exponentially (Figure 1G), and it is reasonable to believe that population contact in the community would also increase.

Figure 1. Individual routines and population distribution data in Mainland China between 2010 and 2016. (A) Flowchart of individual hourly routines. (B) Changes in the proportion of students with increasing urbanization. (C) Changes in the proportion of the workforce with increasing urbanization. (D) Changes in the household size with increasing urbanization. (E) Changes in the school size distribution with increasing urbanization. (F) Changes in the workplace size distribution with increasing urbanization. (G) Changes in the urban population density with increasing urbanization rates.



Here, we defined *Ii* as 1 if individual *i* were infectious and 0 otherwise. The force of infection per hour experienced by individual *i* by 1 infected person in a home, community, school, and workplace was then calculated as $I_k\delta_h e^{a+bshi(t)}$, $I_k\delta_c e^{a+bshi(t)}$, $I_k\delta_s e^{a+bshi(t)}$, and $I_k\delta_w e^{a+bshi(t)}$, respectively, where n_{hi} is the household size of individual *I*; n_{ci} is the mean number of people in the community that individual *i* would contact per hour, which was assumed to be in proportion to the urban population density [25]; n_{si} (n_{wi}) is the mean number of individuals that individual *i* would contact in school (workplace) per hour if he/she were a student (worker), which was assumed to be constant in different provinces and n_{si}=9, n_{wi}=4 [26]; and *a* and *b* are parameters used to adjust the impact of specific humidity, *sh*(t), at time t on influenza transmission. For this study, *a*=0.788 and *b*=–180 [27], while δ_h, δ_c, δ_s, and δ_w indicate the influenza transmission coefficients in a home, community, school, and workplace, respectively. We assumed that $\delta_h:\delta_c:\delta_s:\delta_w = 6:1:12:6$ [11], as children are known to have high clinical attack rates of influenza [28,29]. The absolute influenza transmission coefficients in homes, communities, schools, and workplaces (δ_h, δ_c, δ_s, and δ_w, respectively) are always unknown, and these are mainly determined by the antigenic evolution of the influenza virus and thus vary across years [7]. Therefore, we relied on our model approach to back-calculate the transmission coefficients. For each year, we selected the surveillance data in 1 province and used the least squares method to determine the influenza transmission coefficient parameter value in that year. Then, we used the back-calculated transmission coefficients to simulate the influenza transmission in other provinces and compared them with surveillance data to test our model. The simulation duration of influenza transmission was set to 8 months from November 1, 2013, to June 30, 2014. The time step was set to 1 hour. Each simulation under different settings was conducted 100 times to improve the reliability of the results. The latent and infectious periods were set to 2 and 4 days [30], respectively.

There were 2 types of settings in our simulation. When we simulated infection spread in different provinces, the real parameters from each province were used. When we simulated infection spread at different urbanization levels, the ideal parameters obtained by data fitting based on real parameters from all provinces were used. In this setting, Beijing climate data were used to eliminate the impact of climate differences. The detailed model structure and parameter estimation are shown in Multimedia Appendix 1 [11,12,27,30-34].

Spatial Analysis of the Influenza Surveillance Data by Province

The spread of influenza is influenced by climate [35]. As China is a climatologically diverse country, we cannot directly compare the province-level intensities of the influenza epidemic. To minimize the influence of climate on the analysis, we attempted to compare the epidemic attack rates in provinces with similar climates by using the following 2 methods. First, we analyzed the attack rates of the influenza epidemic in 15 provinces in Northern China that are situated in the temperate climate zone. Second, there is no significant latitude gradient with respect to province-level urbanization in China. Therefore, we classified the provinces according to the urbanization rate. The province urbanization rates ranged from 33.8% to 89.6%, and therefore, we classified the provinces according to the following urbanization rate ranges: <40%, 40%-50%, 50%-60%, 60%-70%, and >70%. The urbanization within each range was defined as the mean urbanization level of the provinces within this range. The influenza epidemic attack rate was also defined as the mean attack rate in the provinces within each urbanization range. Using year 2013 as an example, the provinces within different urbanization ranges are shown in Figure S2 of Multimedia Appendix 1. The mean latitudes of the provinces with urbanization levels <50%, 50%-60%, 60%-70%, and >70% were 31.8, 35.7, 30.4, and 36.8 °N, respectively; the variance in the latitude was narrow within each urbanization range. Therefore, the influence of climate could be minimized. Last, binomial fitting was used to quantify the observed U-shaped relationships between the province-level influenza epidemic attack rates and urbanization rates by the following equation: $y = ax^2+bx+c$, where y represents the influenza epidemic attack and x represents the urbanization rate. The least squares method was used to estimate the unknown parameters a, b, and c. The turning point of the urbanization rate was $-b/2a$.

Ethical Considerations

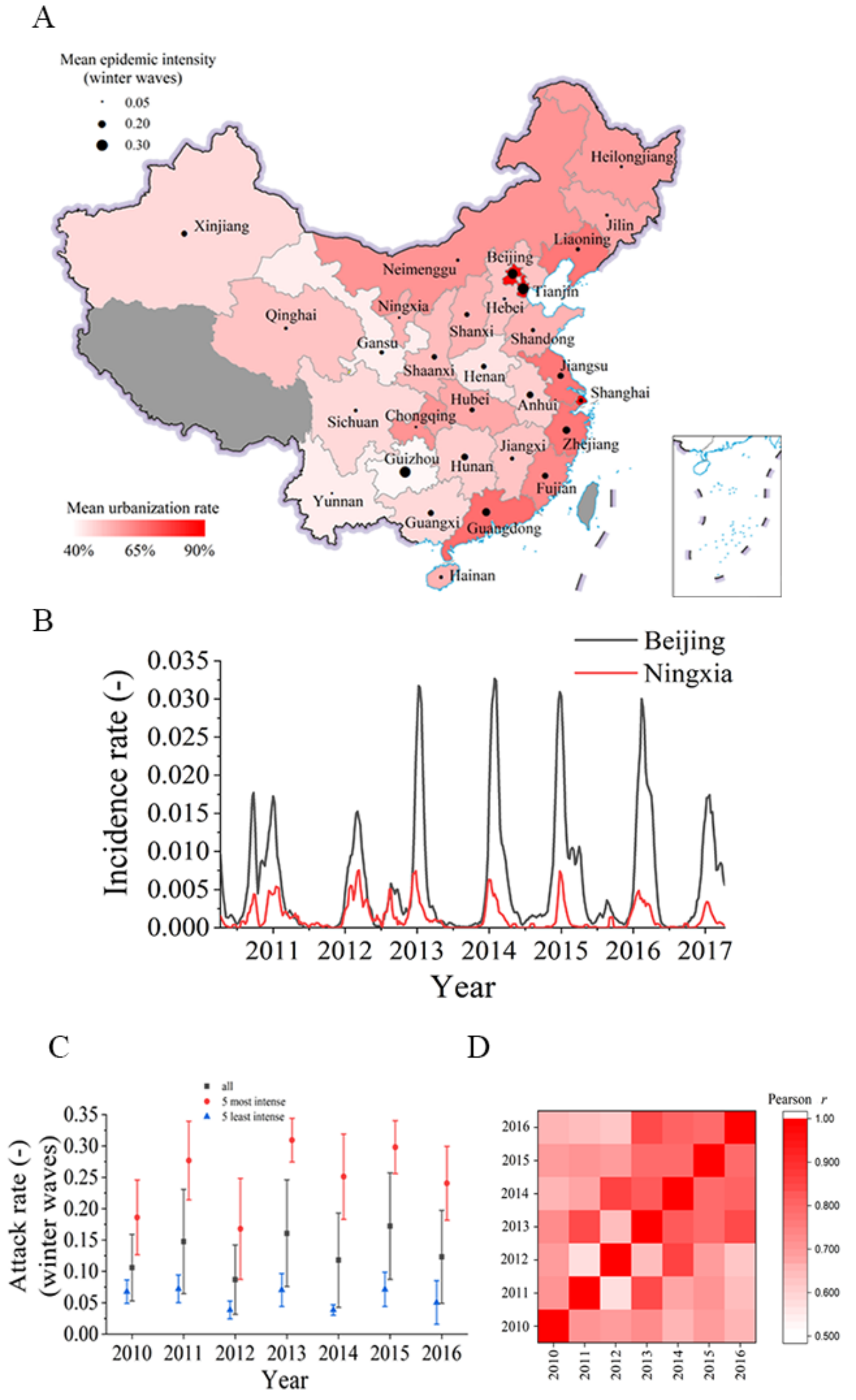
Data in this study are from public database, ie, Chinese National Influenza Centre [36]. We did not make any interventions in the study.

Results

Spatiotemporal Analysis of the Influenza Surveillance Data

During the study period between 2010 and 2016, differences in the mean epidemic attack rate followed a geographic pattern, with the highest intensive epidemic detected in both the most and least urban regions of Beijing (mean urbanization, 86%) and Guizhou (mean urbanization, 39%), respectively. The lowest epidemic attack rate was observed in moderate urban regions such as Ningxia (mean urbanization, 53%) and Chongqing (mean urbanization, 58%; Figure 2A). In addition, the differences persisted over time, that is, some provinces had consistently more intense epidemic than others year after year. The influenza incidence rates in Beijing were always higher than those in Ningxia (Figure 2B). For these 5 provinces with the most or least intense epidemic, the mean epidemic attack rates followed a similar trend over time as the mean epidemic attack rates in all 30 provinces (Figure 2C). The Pearson correlation coefficients of the epidemic attack rates in the 30 provinces between any 2 epidemiological years were higher than 0.5 (Figure 2D), indicating that the differences in the epidemic attack rates among provinces persisted over time. There was no significant linear relationship between the urbanization rate and influenza surveillance intensity (*P*=.64) (Figure S3 of Multimedia Appendix 1).

Figure 2. Systematic differences in the seasonal influenza epidemic attack rates in the winter-spring waves across provinces in Mainland China according to surveillance data between 2010 and 2016. (A) Mean influenza epidemic attack rates (black circles) and urbanization rates (colored bars) in 30 Chinese provinces; grey areas indicate a lack of data. (B) Time-series influenza incidence rates in Beijing and Ningxia provinces. (C) Influenza epidemic attack rates in 5 provinces with the most or least intense epidemic and in all provinces across epidemiological years, where error lines represent 95% CIs. (D) Pearson correlation coefficients of the epidemic attack rates in 30 provinces between epidemiological years.



Modelling Study

We hypothesize that these patterns of epidemic attack rates are due to urbanization. The categorization of the provinces by urbanization rates revealed that the correlation between the influenza epidemic attack rate and urbanization rate yielded a U-shaped curve throughout the study period, with a turning point at about 55% urbanization (Figure 3A). From our binomial-fitter curves, we also found U-shaped relationships between province-level influenza epidemic attack rates and urbanization rates in 15 provinces in Northern China, and we consistently observed the lowest attack rates in provinces with urbanization rates of 50%-60% (Figure 3B and Table 1). The agent-based model that we built in this study also revealed U-shaped relationships between province-level influenza epidemic attack rates and urbanization rates (red line in Figure 3B). The sensitivity analyses of the key parameters in the agent-based model are shown in Figure S4 of Multimedia Appendix 1.

Figure 3. Determination of a U-shaped relationship between the influenza epidemic attack rates and urbanization rate in China. Data are presented by (A) clustering of 30 provinces by the urbanization rate and (B) province-level data in 15 provinces in Northern China. The curves are the binomial fitting of the data in 2010-2016. The red line indicates the agent-based model simulation results, and the error bars indicate the 95% CIs of the 100 replications.

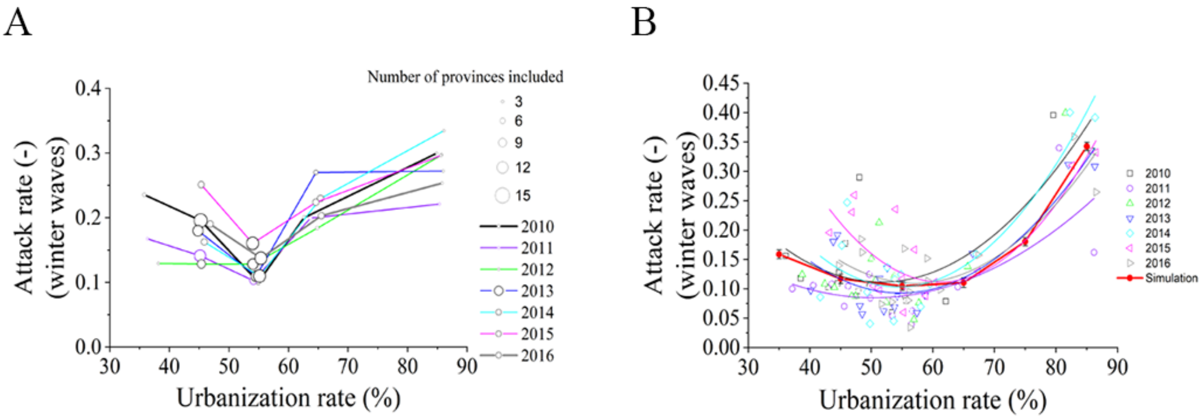
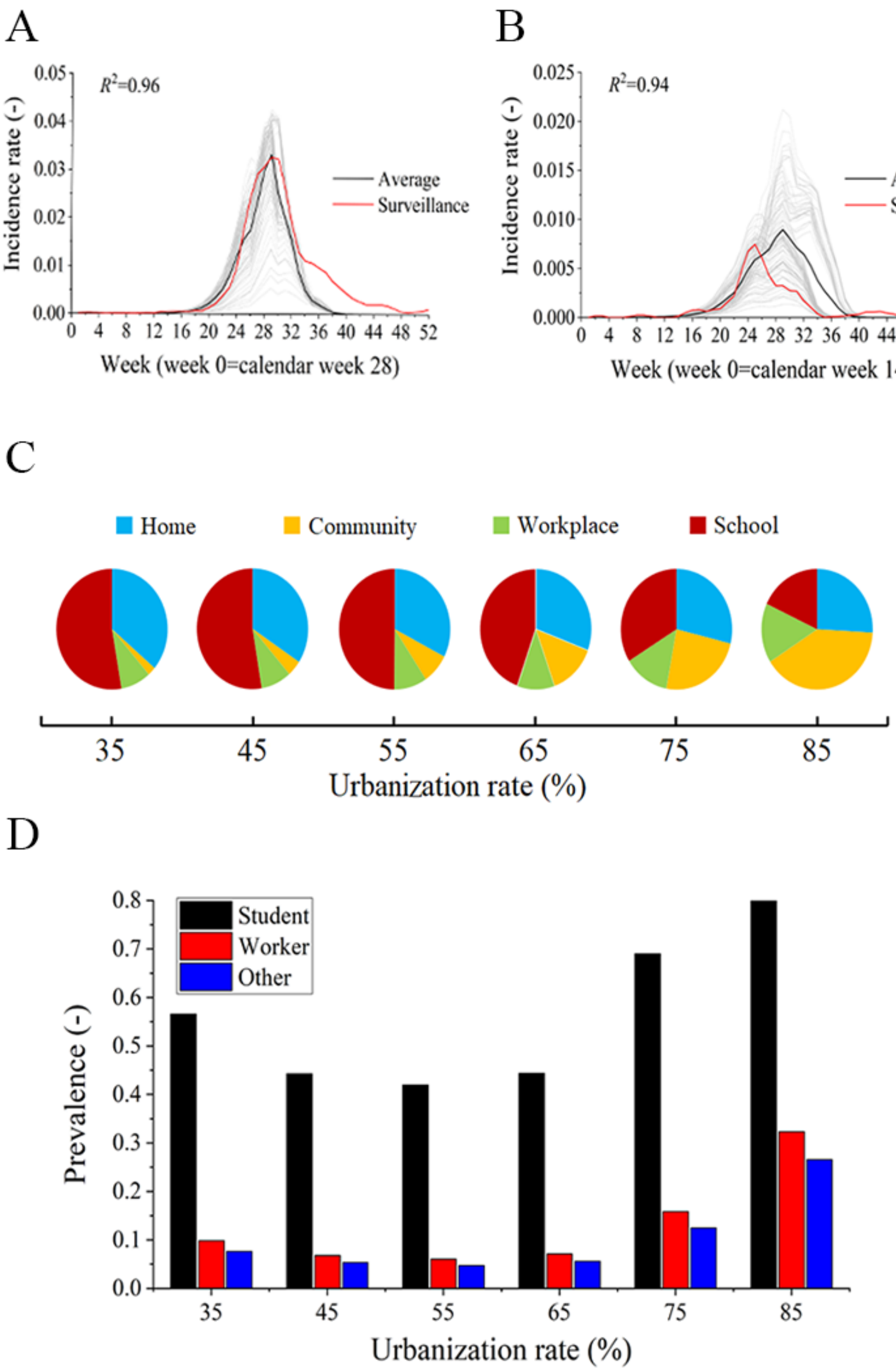


Table 1. Binomial fitting of the province-level epidemic attack rates in the 15 provinces in Northern China.

Epidemiological year	R^2	*P* value	Urbanization rate turning point (%)
2010	0.61	.004	52
2011	0.51	.01	50
2012	0.76	<.001	50
2013	0.77	<.001	55
2014	0.79	<.001	55
2015	0.67	.001	61
2016	0.68	.001	57

Next, we simulated the influenza transmission to quantify the U-shaped relationship observed in Figure 3. As demonstrated using 2013 as an example, we selected surveillance data from Beijing for back-calculation of the unknown transmission coefficients (Figure 4A, R^2=0.96). The back-calculated transmission coefficients and our model simulated the influenza transmission dynamics in Jilin well (Figure 4B, R^2=0.94), demonstrating that the models could be used to simulate influenza transmission in China. Next, we used the model to simulate influenza transmission in provinces with different urbanization rates. The model also yielded a U-shaped association between the urbanization rate and influenza epidemic attack rate (Figure 3B), with a turning point at 55% urbanization.

Figure 4. Simulated influenza epidemic in provinces with different urbanization rates. Simulated incidence rates in (A) Beijing and (B) Jilin. (C) Simulated contributions of each setting to influenza transmission according to the urbanization rate. (D) Influenza prevalence among students, workers, and others according to the urbanization rate.



Taking the simulation in Beijing and Jilin as an example, the agent-based model built in this study well-fitted the influenza surveillance data (Figure 4A and 4B). Thus, using the agent-based model built in this study, we simulated the contributions of different locations and people in influenza transmission with different urbanization rates. Figure 4C presents the percentage contributions of the 4 studied locations with different urbanization rates to influenza transmission. As the urbanization rate increased from 35% to 85%, the contributions of each location to influenza transmission increased from 2.7% to 41.8%: the contribution of workplaces increased from 8.3% to 13.4%, but that of homes decreased from 38.9% to 27.4% and that of schools decreased from 50.2% to 17.4%. Given the relatively frequent number of contacts and the high infection risk per contact in schools, our simulation results demonstrate that students face a much higher infection risk than workers (relative risk 5.4, SD 1.7) and others (relative risk 6.8, SD 2.2; Figure 4D).

Discussion

To the best our knowledge, this study is the first to explore the effect of urbanization on the attack rate of the seasonal influenza epidemic in a country experiencing an unprecedented rapid rate of urbanization. This study was made possible by the availability of census data and influenza surveillance data. Our analyses revealed a U-shaped relationship between the urbanization rate and influenza attack rate, with a turning point at around 55% urbanization rate. We used an agent-based mathematical model based on hourly human contact behaviors to explore the 2 main primary mechanisms responsible for the observed trend. We found that an increase in the urban population density (Figure 1G) increased the number of close contacts between people in communities and thus enhanced the transmission of influenza in public areas. Further, decreases in household sizes (Figure 1D) and percentages of students (Figure 1B) reduced the rates of influenza transmission in homes and schools, respectively. The net effects of these 2 mechanisms provide a simple but reasonable explanation for the observed U-shaped trend in the influenza epidemic attack rate over the 7-year study period. Further confirmation of this U-shaped trend would have significant implications for the influenza epidemic in China. In addition, although this study analyzes the influenza data in China, the methods may also be applicable in other urbanizing countries and for other respiratory infections such as COVID-19, measles, and Middle East respiratory syndrome because indoor environments such as households, schools (workplaces), and communities have always been identified as the primary context of infection transmission [37,38].

The epidemic attack rate is an important parameter for the development of medical surge capacities and public health systems, including primary care facilities. This is particularly important for influenza, as a vaccine specific for a new pandemic virus might not be widely available for up to 6 months based on the current vaccine production technology [3]. Previous studies have proposed that socioeconomic factors may affect both the influenza epidemic and pandemic [4,7,39]. Our study also demonstrates regional differences in the intensity of the influenza epidemic across China, which suggest that seasonal influenza interventions should vary among provinces. We found that the most intense influenza epidemic occurred in the most and least urbanized provinces, suggesting that increased attention should be directed to influenza control in the latter provinces, which are also economically underdeveloped. Previous analyses of the history of infectious diseases and human populations have suggested that the host population density is a critical determinant of the establishment and persistent endemicity of a pathogen in a population. Accordingly, urban aggregations play an important role in maintaining infectious disease endemics [40]. As people have become increasingly concentrated in urban areas, and particularly in large cities, influenza viruses could potentially be spread more easily nowadays. Moreover, large cities always serve as hubs in intercity travel networks, which are linked by high-speed rail in China. These hubs could also influence the spread of the influenza epidemic among cities. These potential trends associated with urbanization pose a challenge to future influenza interventions.

There are at least 2 policy implications to build healthy cities in the aspect of infection control. The first is that influenza vaccine recommendations in China should speed up since vaccine coverage is only approximately 2% in China [41], especially in provinces with urbanization rates greater than 60%. In this study, we found a U-shaped relationship between the influenza epidemic attack rate and urbanization rate, with a turning point between 50% and 60% urbanization. In China, the recent urbanization rate is approximately 60%, and a large number of provinces are currently at the turning point in the U-shaped trend curve. As further urbanization is expected, the trajectory of the influenza epidemic intensity in those provinces in the near future would be worrisome if no interventions were made because of the nonlinear increase in the strain placed on the medical infrastructure [4]. The second is that nonpharmaceutical interventions implemented to control the influenza epidemic and even pandemic should vary by province since this study demonstrates regional differences in the intensity of the influenza epidemic across China. This study shows that in low urbanized regions, influenza transmission mainly occurs in the schools, while in highly urbanized regions, influenza transmission mainly occurs in the community. These findings are well supported by the observed decreasing effectiveness of school closure with increasing urbanization in China [42]. Thus, in highly urbanized regions, ventilation and surface cleaning/disinfection in public places may be effective, while in low urbanized regions, school closures should be recommended.

Parameters in the simulation determined the infectious disease transmission (eg, attack rate). Some parameters such as the time schedule of workers and students and number of students per class were obtained from the literature or real statistical data. However, other parameters such as the daily numbers of the contacted students in each class and the number of workers in each office were estimated and assumed. We adjusted these parameters to minimize the difference between the simulation results and the real infection spread. Moreover, we performed a sensitivity analysis for 10 parameters to show how these parameters influence the infection spread (Figure S4 of Multimedia Appendix 1).

This study suggests a number of limitations and possibilities for future research. First, our model did not consider differences in the medical service levels. Highly urbanized areas have access to better medical services, which would enable a more rapid cure of infectious cases and better control of the influenza epidemic. However, approximately 90% of the seasonal influenza cases remain unreported [43], and infectious cases are mostly transmissible during the incubation period and during 2-3 days with clinical symptoms. Therefore, the medical service level would have a very limited impact on the influenza epidemic. Second, our model did not consider differences in the influenza vaccine coverage because this rate has increased only gradually to approximately 2% of the Chinese population over the past 15 years [41]. This low influenza vaccination rate would have a very limited influence on an epidemic. Third, the results of this study cannot be generalized to rural areas since there is a lack of rural-related data. Fourth, this study assumes that all types of agents (eg, worker, student) have a constant time schedule and all offices and classrooms in a province have the same setting. Therefore, the random scenarios, which could be referred to as a dynamic network, were ignored. Finally, higher resolution analyses of the influenza epidemic, such as those conducted at a city level, would provide more accurate information, as cities are the principal locations of influenza transmission between humans [7]. However, our relatively coarse data set did not allow us to evaluate the influenza patterns at the city or hospital level, even though the surveillance data from 4 of the 30 studied provinces were based on single cities (Beijing, Tianjin, Chongqing, and Shanghai).

In conclusion, we found a U-shaped relationship between the influenza attack rates and urbanization rate in provinces of the urbanizing mainland China, with a turning point between 50% and 60% urbanization. This finding is different from that reported in other studies in urbanized countries. Our study poses the potential challenges of the expected further urbanization process in China on population health and health infrastructure. Further, our results will be critical for precise influenza prediction and prevention.

Acknowledgments

This study was funded by grants from the Shenzhen Science and Technology Program (grant KQTD20180411143323605), the Collaborative Research Fund from the Research Grants Council of Hong Kong (grant C7025-16G), the National Natural Science Foundation of China (grant 82003509), and the Natural Science Foundation of Zhejiang Province (grant LQ20H260009).

Data Availability

Descriptions of the model structure and the parameters included in the model are available in Multimedia Appendix 1. Detailed influenza surveillance data are available from the corresponding author on reasonable request. The model code is available from the author’s GitHub site [44].

Authors' Contributions

YS, YL, and HL conceived and designed the study. YS and YL supervised the study. TC, LY, DW, and XW collected all the data. HL, XD, and XW cleaned the data. HL and NZ analyzed the data. NZ performed the simulation. HL wrote the drafts of the manuscript. YL, YS, NZ, SX, and BC commented on and revised the drafts of the manuscript. YS, YL, and HL interpreted the findings. All authors read and approved the final report.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary data.
[DOCX File , 621 KB - publichealth_v9i1e41435_app1.docx]

References

1. Brankston G, Gitterman L, Hirji Z, Lemieux C, Gardam M. Transmission of influenza A in human beings. Lancet Infect Dis 2007 Apr;7(4):257-265. [doi: 10.1016/S1473-3099(07)70029-4] [Medline: 17376383]
2. Lei H, Li Y, Xiao S, Lin C, Norris SL, Wei D, et al. Author Correction: Routes of transmission of influenza A H1N1, SARS CoV and norovirus in air cabin: Comparative analyses. Indoor Air 2021 Jan;31(1):275 [FREE Full text] [doi: 10.1111/ina.12767] [Medline: 33474780]
3. Qualls N, Levitt A, Kanade N, Wright-Jegede N, Dopson S, Biggerstaff M, CDC Community Mitigation Guidelines Work Group. Community Mitigation Guidelines to Prevent Pandemic Influenza - United States, 2017. MMWR Recomm Rep 2017 Apr 21;66(1):1-34 [FREE Full text] [doi: 10.15585/mmwr.rr6601a1] [Medline: 28426646]
4. Zachreson C, Fair KM, Cliff OM, Harding N, Piraveenan M, Prokopenko M. Urbanization affects peak timing, prevalence, and bimodality of influenza pandemics in Australia: Results of a census-calibrated model. Sci Adv 2018 Dec;4(12):eaau5294 [FREE Full text] [doi: 10.1126/sciadv.aau5294] [Medline: 30547086]
5. Our cities, our health, our future: acting on social determinants for health equity in urban settings. World Health Organization. URL: https://cdn.who.int/media/docs/default-source/documents/social-determinants-of-health/urban-settings-knowledge-network-final-report-2007.pdf?sfvrsn=90e5da72_3 [accessed 2021-07-15]
6. Eubank S, Guclu H, Kumar VSA, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling disease outbreaks in realistic urban social networks. Nature 2004 May 13;429(6988):180-184. [doi: 10.1038/nature02541] [Medline: 15141212]

7. Dalziel BD, Kissler S, Gog JR, Viboud C, Bjørnstad ON, Metcalf CJE, et al. Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities. Science 2018 Oct 05;362(6410):75-79 [FREE Full text] [doi: 10.1126/science.aat6030] [Medline: 30287659]
8. Urban population (% of total population). The World Bank. URL: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?view=chart [accessed 2023-03-01]
9. China Statistical Yearbook. URL: http://www.stats.gov.cn/sj/ndsj/2018/indexeh.htm [accessed 2021-07-15]
10. China's population development shows new features and trends. China National Bureau of Statistics. URL: http://www.stats.gov.cn/xxgk/jd/sjjd2020/202105/t20210513_1817408.html [accessed 2023-03-01]
11. Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 2005 Sep 08;437(7056):209-214 [FREE Full text] [doi: 10.1038/nature04017] [Medline: 16079797]
12. Zhang N, Huang H, Su B, Ma X, Li Y. A human behavior integrated hierarchical model of airborne disease transmission in a large city. Build Environ 2018 Jan;127:211-220 [FREE Full text] [doi: 10.1016/j.buildenv.2017.11.011] [Medline: 32287976]
13. Karimi E, Schmitt K, Akgunduz A. Effect of individual protective behaviors on influenza transmission: an agent-based model. Health Care Manag Sci 2015 Sep;18(3):318-333. [doi: 10.1007/s10729-014-9310-2] [Medline: 25578039]
14. Staffini A, Svensson AK, Chung U, Svensson T. An Agent-Based Model of the Local Spread of SARS-CoV-2: Modeling Study. JMIR Med Inform 2021 Apr 06;9(4):e24192 [FREE Full text] [doi: 10.2196/24192] [Medline: 33750735]
15. Nasrinpour HR, Reimer AA, Friesen MR, McLeod RD. Data Preparation for West Nile Virus Agent-Based Modelling: Protocol for Processing Bird Population Estimates and Incorporating ArcMap in AnyLogic. JMIR Res Protoc 2017 Jul 17;6(7):e138 [FREE Full text] [doi: 10.2196/resprot.6213] [Medline: 28716770]
16. Gao X, Wei J, Lei H, Xu P, Cowling BJ, Li Y. Building Ventilation as an Effective Disease Intervention Strategy in a Dense Indoor Contact Network in an Ideal City. PLoS One 2016;11(9):e0162481 [FREE Full text] [doi: 10.1371/journal.pone.0162481] [Medline: 27611368]
17. Liu X, Li Y, Zhu Y, Zhang J, Li X, Zhang J, et al. Seasonal pattern of influenza activity in a subtropical city, China, 2010-2015. Sci Rep 2017 Dec 13;7(1):17534 [FREE Full text] [doi: 10.1038/s41598-017-17806-z] [Medline: 29235535]
18. Shu Y, Song Y, Wang D, Greene CM, Moen A, Lee CK, et al. A ten-year China-US laboratory collaboration: improving response to influenza threats in China and the world, 2004-2014. BMC Public Health 2019 May 10;19(Suppl 3):520 [FREE Full text] [doi: 10.1186/s12889-019-6776-3] [Medline: 32326921]
19. Goldstein E, Cobey S, Takahashi S, Miller JC, Lipsitch M. Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: a statistical method. PLoS Med 2011 Jul;8(7):e1001051 [FREE Full text] [doi: 10.1371/journal.pmed.1001051] [Medline: 21750666]
20. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. Real-time influenza forecasts during the 2012-2013 season. Nat Commun 2013;4:2837 [FREE Full text] [doi: 10.1038/ncomms3837] [Medline: 24302074]
21. Yu H, Alonso WJ, Feng L, Tan Y, Shu Y, Yang W, et al. Characterization of regional influenza seasonality patterns in China and implications for vaccination strategies: spatio-temporal modeling of surveillance data. PLoS Med 2013 Nov;10(11):e1001552 [FREE Full text] [doi: 10.1371/journal.pmed.1001552] [Medline: 24348203]
22. Shu Y, Fang L, de Vlas SJ, Gao Y, Richardus JH, Cao W. Dual seasonal patterns for influenza, China. Emerg Infect Dis 2010 Apr;16(4):725-726 [FREE Full text] [doi: 10.3201/eid1604.091578] [Medline: 20350403]
23. China City Statistical Yearbook. URL: https://www.mohurd.gov.cn/gongkai/fdzdgknr/sjfb/tjxx/index.html [accessed 2023-05-09]
24. National Bureau of Statistics of China. URL: http://www.stats.gov.cn/ [accessed 2023-05-09]
25. Zhang N, Huang H, Duarte M, Zhang JJ. Dynamic population flow based risk analysis of infectious disease propagation in a metropolis. Environ Int 2016 Sep;94:369-379. [doi: 10.1016/j.envint.2016.03.038] [Medline: 27107973]
26. Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. Science 2020 Jun 26;368(6498):1481-1486 [FREE Full text] [doi: 10.1126/science.abb8001] [Medline: 32350060]
27. Pei S, Kandula S, Yang W, Shaman J. Forecasting the spatial transmission of influenza in the United States. Proc Natl Acad Sci USA 2018 Mar 13;115(11):2752-2757 [FREE Full text] [doi: 10.1073/pnas.1708856115] [Medline: 29483256]
28. Olson DR, Heffernan RT, Paladini M, Konty K, Weiss D, Mostashari F. Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City. PLoS Med 2007 Aug;4(8):e247 [FREE Full text] [doi: 10.1371/journal.pmed.0040247] [Medline: 17683196]
29. Schanzer D, Vachon J, Pelletier L. Age-specific differences in influenza A epidemic curves: do children drive the spread of influenza epidemics? Am J Epidemiol 2011 Jul 01;174(1):109-117 [FREE Full text] [doi: 10.1093/aje/kwr037] [Medline: 21602300]
30. Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza. Nature 2004 Dec 16;432(7019):904-906 [FREE Full text] [doi: 10.1038/nature03063] [Medline: 15602562]
31. Anderson M. The Population Dynamics of Infectious Diseases: Theory and Applications. Berlin, Germany: Springer; 1982.
32. Anderson RM, May RM. Infectious Diseases of Humans Dynamics and Control. London: Oxford University Press; 1992.

33. Bolton D. The Computation of Equivalent Potential Temperature. URL: https://journals.ametsoc.org/view/journals/mwre/108/7/1520-0493_1980_108_1046_tcoept_2_0_co_2.xml?tab_body=pdf [accessed 2023-05-09]
34. Zhang N, Cheng P, Jia W, Dung CH, Liu L, Chen W, et al. Impact of intervention methods on COVID-19 transmission in Shenzhen. Build Environ 2020 Aug;180:107106 [FREE Full text] [doi: 10.1016/j.buildenv.2020.107106] [Medline: 32834417]
35. Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog 2007 Oct 19;3(10):1470-1476 [FREE Full text] [doi: 10.1371/journal.ppat.0030151] [Medline: 17953482]
36. Chinese National Influenza Centre. Home page. Chinese National Influenza Centre. URL: https://ivdc.chinacdc.cn/cnic/en/
37. Kim D. Exploratory study on the spatial relationship between emerging infectious diseases and urban characteristics: Cases from Korea. Sustain Cities Soc 2021 Mar;66:102672 [FREE Full text] [doi: 10.1016/j.scs.2020.102672] [Medline: 33520608]
38. Lei H, Xu X, Xiao S, Wu X, Shu Y. Household transmission of COVID-19-a systematic review and meta-analysis. J Infect 2020 Dec;81(6):979-997 [FREE Full text] [doi: 10.1016/j.jinf.2020.08.033] [Medline: 32858069]
39. Grantz KH, Rane MS, Salje H, Glass GE, Schachterle SE, Cummings DAT. Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. Proc Natl Acad Sci U S A 2016 Nov 29;113(48):13839-13844 [FREE Full text] [doi: 10.1073/pnas.1612838113] [Medline: 27872284]
40. Dobson AP, Carper ER. Infectious Diseases and Human Population History. BioScience 1996 Feb;46(2):115-126. [doi: 10.2307/1312814]
41. Li L, Liu Y, Wu P, Peng Z, Wang X, Chen T, et al. Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. Lancet Public Health 2019 Sep;4(9):e473-e481 [FREE Full text] [doi: 10.1016/S2468-2667(19)30163-X] [Medline: 31493844]
42. Lei H, Jiang H, Zhang N, Duan X, Chen T, Yang L, et al. Increased urbanization reduced the effectiveness of school closures on seasonal influenza epidemics in China. Infect Dis Poverty 2021 Oct 21;10(1):127 [FREE Full text] [doi: 10.1186/s40249-021-00911-7] [Medline: 34674754]
43. Axelsen JB, Yaari R, Grenfell BT, Stone L. Multiannual forecasting of seasonal influenza dynamics reveals climatic and evolutionary drivers. Proc Natl Acad Sci U S A 2014 Jul 01;111(26):9538-9542 [FREE Full text] [doi: 10.1073/pnas.1321656111] [Medline: 24979763]
44. U-shape between influenza and urbanisation. GitHub. URL: https://github.com/Lei-hao-Group/Ushape-between-influenza-and-urbanisation [accessed 2023-05-09]

Abbreviations

ILI: influenza-like illness

Edited by Y Khader; submitted 08.10.22; peer-reviewed by A El Rifay, A Couture, A Staffini; comments to author 07.02.23; revised version received 06.04.23; accepted 11.04.23; published 07.07.23.

Please cite as:
Lei H, Zhang N, Niu B, Wang X, Xiao S, Du X, Chen T, Yang L, Wang D, Cowling B, Li Y, Shu Y
Effect of Rapid Urbanization in Mainland China on the Seasonal Influenza Epidemic: Spatiotemporal Analysis of Surveillance Data From 2010 to 2017
JMIR Public Health Surveill 2023;9:e41435
URL: https://publichealth.jmir.org/2023/1/e41435
doi:10.2196/41435
PMID:37418298

Original Paper

Effects of Tobacco Versus Electronic Cigarette Usage on Nonsuicidal Self-Injury and Suicidality Among Chinese Youth: Cross-Sectional Self-Report Survey Study

Yinzhe Wang[1,2*], MPhil, MSc; Shicun Xu[3,4*], PhD; Xiaoqian Zhang[5*], MD; Yanwen Zhang[1*], MSc; Yi Feng[6,7], PhD; Yuanyuan Wang[8], PhD; Runsen Chen[1,9], PhD

[1]Department of Human Development & Quantitative Methods, Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States

[2]Vanke School of Public Health, Tsinghua University, Beijing, China

[3]Northeast Asian Research Center, Jilin University, Changchun, Jilin, China

[4]China Center for Aging Studies and Social-Economic Development, Jilin University, Changchun, Jilin, China

[5]Department of Psychiatry, Tsinghua University Yuquan Hospital, Beijing, China

[6]Mental Health Center, Central University of Finance and Economics, Beijing, China

[7]Psychology Application Center for Enterprise and Society, School of Sociology and Psychology, Central University of Finance and Economics, Beijing, China

[8]Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, Beijing, China

[9]Institute for Healthy China, Tsinghua University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Shicun Xu, PhD
China Center for Aging Studies and Social-Economic Development
Jilin University
Kuang Yaming Bldg, 6th Fl
Changchun, Jilin
China
Phone: 86 43185166393
Email: xushicun@jlu.edu.cn

Abstract

Background: The increase in tobacco/conventional cigarette (CC) and electronic cigarette (EC) usage among Chinese youth has become a growing public health concern. This is the first large-scale study to compare the impact of CC and EC usage on risk for nonsuicidal self-injury (NSSI) and suicidality in cis-heterosexual and sexual and gender minority (SGM) youth populations in China.

Objective: This study examines the CC and EC risks for NSSI and suicidality among Chinese youth and compares the extent to which SGM and cis-heterosexual youth's risks for NSSI and suicidality are influenced by their CC and EC usage and dependence.

Methods: A total of 89,342 Chinese participants completed a cross-sectional self-report survey in 2021. Sociodemographic information, sexual orientations, gender identities, CC and EC usage, CC and EC dependence, and risks for suicidality and NSSI were assessed. The Mann-Whitney *U* test and chi-square test were performed for nonnormally distributed continuous variables and categorical variables, respectively. The multivariable linear regression model was used to examine both the influence of CC and EC usage and CC and EC dependence on NSSI and suicidality as well as the interaction effects of CC and EC usage and CC and EC dependence on NSSI and suicidality by group.

Results: The prevalence of CC usage ($P<.001$) and dependence ($P<.001$) among SGM participants was lower than that among their cis-heterosexual counterparts. However, the prevalence of EC usage ($P=.03$) and EC dependence ($P<.001$) among SGM participants was higher than that among their cis-heterosexual counterparts. The multivariable linear regression model showed that CC dependence and EC dependence had a unique effect on NSSI and suicidality (CCs: B=0.02, $P<.001$; B=0.09, $P<.001$; ECs: B=0.05, $P<.001$; B=0.14, $P<.001$, respectively). The interaction effects of (1) CC usage and group type on NSSI and suicidality (B=0.34, $P<.001$; B=0.24, $P=.03$, respectively) and dual usage and group type on NSSI and suicidality (B=0.54, $P<.001$; B=0.84, $P<.001$, respectively) were significant, (2) CC dependence and group type on NSSI were significant (B=0.07, *P*<.001), and (3) EC dependence and group type on NSSI and suicidality were significant (B=0.04, *P*<.001; B=0.09, *P*<.001, respectively). No significant interaction effect was observed between EC usage and group type on NSSI and suicidality (B=0.15, *P*=.12; B=0.33, *P*=.32, respectively) and between CC dependence and group type on suicidality (B=–0.01, *P*=.72).

Conclusions: Our study shows evidence of intergroup differences in NSSI and suicidality risks between SGM and cis-heterosexual youth related to CC and EC usage. These findings contribute to the growing literature on CC and EC in cis-heterosexual and SGM populations. Concerted efforts are necessary at a societal level to curb the aggressive marketing strategies of the EC industry and media coverage and to maximize the impact of educational campaigns on EC prevention and intervention among the youth population.

(JMIR Public Health Surveill 2023;9:e47058) doi:10.2196/47058

KEYWORDS

electronic cigarettes; tobacco; conventional cigarettes; SGM; nonsuicidal self-injury; suicidality; suicidal; cigarette; cigarettes; suicide; self-harm; mental health; sexual minority; gender minority; sexual and gender minority; transgender; youth; cross-sectional; survey; smoker; smoking

Introduction

Electronic cigarettes or e-cigarettes (ECs) are battery-based vaping devices simulating tobacco/conventional cigarette (CC) smoking by aerosolizing artificially manufactured liquid solutions containing nicotine for users' inhalation [1]. The association between the usage of CCs and suicidality is well-established, with previous studies showing that regular usage of CCs is associated with an increased risk of suicidality (ie, suicidal ideation, plan, and attempt) [2-4]. Several studies have reported the neurotoxicity of CC usage on adolescents' brain structure, increasing the risk of oxidative stress–related neurological disorders while promoting emotional dysregulation [4-6]. Subsequently, the emotional dysregulation brought about by CC usage could further aggravate individuals' pre-existing mental health conditions (eg, depressive thinking featured by hopelessness and worthlessness). Combined with relevant social factors (eg, isolation), the complex interplay between CC usage and mental health risks may heighten the risk of suicidality among CC users [4-6].

Similar to CC usage, EC usage can impose not only physical risks (ie, nicotine dependence, cardiovascular injury, respiratory system damage, and noncancer pulmonary disease) but also mental health threats, including but not limited to depression, anxiety, and nonsuicidal self-injury (NSSI) [5,6]. In addition, recent studies suggest that nicotine usage, including CC and EC usage, has a bivariate relationship with NSSI, underscoring the potential relationship between NSSI and both CC and EC usage [7,8]. Nevertheless, although findings from the abovementioned studies have shown the impact of CC and EC usage on youth NSSI and suicidality, their differences in chemical composition, method of consumption, and sociocultural perceptions are drastically different from each other. First, regarding the chemical composition, compared to tobacco in CC, EC is featured with a liquid containing both nicotine and flavoring, with a high percentage of it being glycerol or propylene glycerol [9]. Second, concerning the method of consumption, compared to the lighting process of CC, EC often uses a power button to activate an atomizer to generate an aerosol from its liquid, which simulates smoke from CC and requires no fire and burning, thereby making its usage more prevalent in places where CC is banned [9,10]. Third, with regard to sociocultural perceptions, EC has been marketed as less harmful than CC in the Chinese market, with relatively limited regulations on the minimum age to purchase compared to CC [10]. Due to the abovementioned differences, EC and CC may influence youth NSSI and suicidality differently.

Findings from past research postulate that college students' high susceptibility to mental health risks, exemplified in forms of anxiety, depression, and stress, when facing concerns resulting from sudden environmental changes (ie, moving from one's hometown to another city), social isolation, social alienation, worries of financial independence, high academic performance expectations, and learning style change corresponds to their high risk of suicidality [11,12]. Previous research argues that college students often begin to explore their gender and sexual identity in their college lives, during which the disparity between cis-heterosexual and sexual and gender minority (SGM) youth in suicidality may surface [13,14]. The 2016 National Health Interview Survey showed that 20.5% of sexual minority young adults smoke CCs, which is much higher than that reported among their heterosexual counterparts (ie, 15.3%) [15]. The prevalence of CC usage in the transgender and gender-diverse population is also higher than that in their cisgender counterparts [15,16]. Similarly, recent studies have also found that the prevalence of EC usage in the SGM population is higher than that in their cis-heterosexual counterpart [17-20]. Specifically, there is burgeoning evidence from previous studies that suggest SGM youth's current usage (SGM: 9.4%; cis-hetero: 4.9%), lifetime usage (SGM: 25.1%; cis-hetero: 14.3%), and past 30-day usage of ECs (SGM: 28.5%; cis-hetero: 23.4%) are all higher than those among cis-heterosexual youth, indicating this group's vulnerability toward EC usage [18-20].

Such a disparity in EC and CC usage between SGM and cis-heterosexual youth could be potentially explained by the minority stress model. Minority stressors, including but not limited to heterosexist discrimination, social rejection, social isolation, and homo/bi/transphobic harassment, could further worsen the stigmatization of SGM youth's identities while inducing higher stress levels among them, contributing to their increased risk of substance usage [21]. Furthermore, besides the minority stress experienced by SGM youth, the tobacco industry has further exploited their vulnerability since they have historically been targeting SGM youth as a major consumer crowd and pushed forward relevant campaigns such as the Project Subculture Urban Marketing for gay individuals in San Francisco [22]. Nevertheless, to date, no study has examined the relationship between the usage of CCs and ECs on college students' NSSI and suicidality while investigating potential intergroup differences existing between cis-heterosexual and SGM youth, leaving a critical gap in the literature. Most of the current CC and EC research [4-22] examining the increasing risks of youth for NSSI and suicidality focused on studying CC and EC usage separately and did not further investigate the different extents to which the SGM and cis-heterosexual youth's risks for NSSI and suicidality were affected by CC and EC usage, ignoring the potentially existing intergroup differences. Our cross-sectional study aims to fill this gap by investigating the effects of CC and EC on NSSI and suicidality and the difference in the effects between SGM and cis-heterosexual youth. This study's hypotheses postulate that (1) cigarette usage (including CC and EC usage) and cigarette dependence (including CC and EC dependence) have a unique and positive impact on both NSSI and suicidality, (2) cigarette usage has a more pronounced impact on NSSI and suicidality among SGM youth than on NSSI and suicidality among their cis-heterosexual peers, and (3) cigarette dependence has a more severe impact on NSSI and suicidality among SGM youth than on NSSI and suicidality among their cis-heterosexual peers.

Methods

Participant Recruitment

University students from 63 universities and colleges in the Jilin province, China, were invited to participate in a cross-sectional web-based self-report survey. Recruitment was performed via universities' and colleges' staff teachers and targeted WeChat groups (similar to WhatsApp groups). Students who were interested in participation entered the secure and anonymous questionnaire through the QR code mentioned in the study poster that was distributed by school staff teachers. Students were first asked to thoroughly read the study description, and those who selected the commensurate button and continued to answer the questionnaire were deemed as consented to this study. Individuals who successfully completed this questionnaire were offered the opportunity to join a cash prize lottery by accessing a separate link. The description of the cash prize lottery was listed on the front page of the web-based survey, and the amount was carefully calibrated to strike a balance between being sufficient to encourage participation and minimize dropout without being excessively large to create an undue influence on participants' behavior or reporting. Respondents' data were collected between October 26 and November 18, 2021, via Credamo [23].

For statistical quality control of this study, we excluded respondents who (1) were younger than 15 years (ie, below the lowest threshold of the regular age range of college or university students in China), (2) failed 2 or more of the 4 attention check questions (ie, this question is an attention check question. There are many colors of the sky, which can be green, blue, red and black, but, for this question, please choose "green" for your answer) that were designed to assess whether survey respondents were focused during their responding process, and (3) showed presence of logical contradictions or inconsistencies within responses (eg, male participants choosing "male" for sexual preference while choosing "heterosexual" for sexual orientation), omission of answers (eg, leaving blanks in their answers instead of choosing options such as "none of the above" or "prefer not to say"), or their answers showed evidence of patterned or nondiscriminatory responses (eg, consistently choosing the first option across multiple questions). The final sample consisted of 89,342 participants, among which 8853 were SGM respondents and 80,489 were cis-heterosexual respondents.

Study Measures

Gender Identity

Participants' sociodemographic information, including age, ethnicity (Han ethnicity or non-Han ethnicity), educational background (undergraduate, master, or doctoral students), residence areas prior to university/college enrollment (urban or rural), only-child status, and annual family income, was collected. This study's 3 gender identity assessment questions were based on the GenIUSS published report from the Williams Institute, University of California, Los Angeles [24]. First, participants were asked, "What was your sex assigned at birth? (gender specified at birth or the one listed on your original birth certificate)" with options being male and female. Second, participants were asked, "What is your gender identity? (your personal thoughts and understanding of your own gender)" with options being cisgender, transgender, nonbinary/gender nonconforming/agender/others, unsure, and cannot understand the question. Third, participants were asked, "If you can only choose from 1 option below, which gender option best describes you?" with options being male, female, transgender male (female-to-male or longing to become male), transgender female (male-to-female or longing to become female), genderqueer/nonbinary/gender nonconforming, cross-dresser, unsure, and not listed (please write down your answer in the blank). The third question served the function of discerning self-identified transgender male and female participants in our sample, including male-to-female, female-to-male, and participants who desired to live their lives as the opposite sex. Written answers from participants who chose "not listed (please write down your answer in the blank)" were manually inspected and reviewed by 2 research assistants who then recoded those participants' gender identities after reaching a consensus in thorough discussions. Due to the sociocultural unfamiliarity with the gender identity glossary, a validity check was performed for the first and third questions. Transgender and gender nonconforming participants, including transgender males, transgender females, and participants who were genderqueer/nonbinary/gender nonconforming, who answered consistently across all 3 questions were included in the analyses.

Sexual Orientation

This study's 3 sexual orientation assessment questions were also based on the SMART published report from the Williams Institute, University of California, Los Angeles [25]. First, participants were asked, "Which one of the following best describes your sexual orientation?" with choices being asexual (not sexually attracted to others or have low interest in sexual activity), homosexual, heterosexual, bisexual (can be romantically and sexually attracted to both males and females), pansexual (can be romantically and sexually attracted to anyone, regardless of their sex or gender identity), and not listed. Participants were then asked to answer the second and third questions, "What kind of person are you sexually attracted to?" and "What kind of person are you romantically attracted to?" with the same set of choices ranging from male, female, both male and female, any gender (including transgender, genderqueer, and nonbinary), unsure, and not listed (please write down your answer in the blank). For participants who chose "not listed" for the first question, 2 aforementioned research assistants manually inspected their written answers, confirmed their answers' consistency by comparing their answers to the second and third questions, and eventually recoded their sexual orientations after reaching consensus in thorough discussions. Participants who sexually identified as homosexual, bisexual, and pansexual were included in the final analyses. Upon coding completion, all participants in this study were successfully labeled as SGM and cis-heterosexual.

CC and EC Usage

We employed 2 screening questions to assess participants' CC and EC usage. Participants were asked to answer the first question, "Do you currently have smoking habits?" with options including yes and no. Those who chose yes continued to answer the remaining questions related to CC and EC usage, while others directly skipped to the next section of the questionnaire. Those who chose yes were then asked the second question, "Do you usually smoke e-cigarettes or conventional cigarettes?" with options including e-cigarettes, conventional cigarettes, and both. First, those who chose conventional cigarettes were then assessed by answering the 6-item validated Chinese version of the Fagerström Test of Nicotine Dependence (FTND) [26,27]. Their scores, ranging from 0 to 10, were calculated by summing up their responses, with higher scores indicating higher levels of nicotine dependence. Scores from 0 to 2 demonstrate very low dependence; scores from 3 to 4 demonstrate low dependence; scores at 5 demonstrate medium dependence; scores from 6 to 7 demonstrate high dependence; and scores from 8 to 10 demonstrate very high dependence. The validated Chinese version has been used in previous research for similar purposes, demonstrating good internal consistency and validity [28,29]. Second, for those who chose ECs, we revised the validated Chinese version of the FTND to measure their EC dependence. Although the revised scale also contains 6 items, we expanded the measurement ranges of both the first and fourth items by adding 2 additional options for each item, increasing the score range from 0-10 to 0-14. Nevertheless, participants still scored on a response scale, with a higher score indicating a higher level of EC dependence. Third, those who chose "Both" were asked to answer both sets of questions. For those who did not answer the corresponding part of the questionnaire, scores were set to 0. In this study, the reliability values of FTND (ω=0.83) and e-FTND (ω=0.80) were acceptable.

Suicidality

The Suicidal Behaviors Questionnaire-Revised (SBQ-R) was designed to measure individuals' suicide-related behaviors from 4 different dimensions, that is, individuals' history of suicide attempt(s), individuals' frequency of suicidal ideation, individuals' suicide threats, and individuals' future suicide attempts likelihood [30]. The validated Chinese version of the SBQ-R contains 4 items, with the total score ranging from 3 to 18 and higher scores indicating higher levels of suicidality [31]. A total score of 7 or higher was considered a high suicide risk, while a total score below 7 was considered a moderate or low suicide risk. In this study, the ω of SBQ-R was 0.81.

NSSI Analysis

Two items adapted based on the Clinician-Rated Severity of Nonsuicidal Self-Injury were used in this study [32]. The first item asked participants whether they have ever engaged in NSSI, with options including "Yes. I engaged in NSSI more than one year ago," "Yes, I engaged in NSSI within the past year," and "no." The second item asked how many days during the past year did participants engage in NSSI, with options including none, 1 to 4 days, 5 to 7 days, 8 to 11 days, and 12 days or more. We combined these 2 items as 1 item for participants to answer, with options ranging from "None," "I engaged in NSSI more than one year ago," "I have had NSSI 1 to 4 days in the past year," "I have had NSSI 5 to 7 days in the past year," "I have had NSSI 8 to 11 days in the past year," to "I have had NSSI 12 days or more in the past year," with participants' scores ranging from 0 (none) to 6 (12 days or more in the past year). Since NSSI for 5 or more days in the past year was the clinically recommended cutoff, as suggested by the American Psychiatric Association [32], participants in this study were classified into 3 groups (ie, group with no NSSI history, group with less than 5 days of NSSI within the past year, and group with more than 5 days of NSSI within the past year) for further analyses.

Data Analysis

First, we performed a descriptive analysis of the sociodemographic variables of the SGM and the cis-heterosexual groups. The Mann-Whitney *U* test and chi-square test were performed for nonnormally distributed continuous variables (eg, scores of CC dependence, EC dependence, NSSI, suicidality) and categorical variables (eg, classifications of cigarette usage, NSSI, and suicidality), respectively. Our primary analyses using the multivariable linear regression model involved examinations of both the influence of CC usage and EC usage on NSSI and suicidality and the interaction effects of CC usage and EC usage on NSSI and suicidality by group (ie, SGM and cis-heterosexual). Both models were controlled for other sociodemographic variables. Results were reported at 95% CIs and were only considered significant when *P* values were less than .05. For variables with a small number of missing values (eg, age, education background), we used mean substitution to input the missing data. We used ω as the measure of the tools' reliability index [33]. All statistical analyses were performed using SPSS (version 28, IBM Corp).

Ethics Approval

This cross-sectional survey received approval from the ethics committee at Jilin University (approval 2021-9-29). Informed consent for primary data collection and secondary data analyses from participants was automatically collected from students who selected the commensurate button and continued to answer the questionnaire. The privacy and confidentiality of the participants in this study were protected during data collection and analysis through anonymization and deidentification. None of the personally identifiable information was collected, and all data were stored in a personal computer in the laboratory with password protection. As mentioned earlier, compensation for participants in the study was provided with a link on the postsurvey completion page to the cash reward lottery, with the amount being meticulously calculated to be rewarding while not generating potential bias.

Results

Table 1 summarizes the sociodemographic characteristics of the participants by group type (ie, SGM or cis-heterosexual). Cigarette usage prevalence was different between SGM participants and cis-heterosexual participants (χ^2_3[N=89,342]=149.8; *P*<.001). Specifically, the proportion of SGM participants (454/8853, 5.1%) who only used CC was lower than that of their cis-heterosexual counterparts (7161/80,489, 8.9%; *P*<.05); the proportion of SGM participants (53/8853, 0.6%) who only used EC was higher than that of their cis-heterosexual counterpart (348/80,489, 0.4%; *P*<.05). Moreover, the proportion of SGM participants who used CCs was lower than that of their cis-heterosexual counterpart (873/8853, 9.9% vs 10,685/80,489, 13.3%; χ^2_1[N=89,342]=82.5; *P*<.001, respectively). In contrast, the proportion of SGM participants who used ECs was higher than that of their cis-heterosexual counterpart (472/8853, 5.3% vs 3872/80,489, 4.8%, χ^2_1[N=89,342]=4.7; *P*=.03, respectively).

Table 2 summarizes the intergroup differences between SGM and cis-heterosexual individuals in this study. The CC dependence score of the SGM group was lower than that of their cis-heterosexual counterpart (*z*=–9.25, *P*<.001). Furthermore, the SGM participants had lower dependence levels on CCs than the cis-heterosexual participants (χ^2_4[N=89,432]=37.6; *P*<.001). However, compared to the cis-heterosexual group, the SGM group scored higher in EC dependence, suicidality, and NSSI (*z*=3.20, *P*=.001; *z*=69.91, *P*<.001; *z*=54.32, *P*<.001, respectively). At the same time, the SGM group had higher suicide risk and individuals with more NSSI compared to the cis-heterosexual group (χ^2_1[N=89,432]=4507.5; *P*<.001; χ^2_2[N=89,432]=3070.9; *P*<.001, respectively). In addition, we analyzed the intergroup difference between the SGM and the cis-heterosexual groups in the revised 6-item scale measuring EC usage. As shown in Table S1 of Multimedia Appendix 1, results from the chi-square test analysis highlighted that compared with the cis-heterosexual group, the SGM group showed a greater degree of EC dependence on each item of the revised scale.

Table 1. Sociodemographic and cigarette usage data of the sample population in this study.

Variables	Sexual and gender minority participants (n=8853)	Cis-heterosexual participants (n=80,489)
Age[a] (years), mean (SD)	19.57 (1.74)	19.60 (1.75)
Sex assigned at birth, n (%)		
Male	2322 (26.2)	35,582 (44.2)
Female	6531 (73.8)	44,907 (55.8)
Ethnicity, n (%)		
Han ethnic group	7850 (88.7)	72,179 (89.7)
Others	1003 (11.3)	8310 (10.3)
Education background[b], n (%)		
Undergraduate	8361 (94.4)	75,780 (94.2)
Masters	477 (5.4)	4516 (5.6)
Doctoral	15 (0.2)	193 (0.2)
Residence areas prior to university/college enrollment, n (%)		
City	5304 (59.9)	40,060 (49.8)
Rural/suburban	3549 (40.1)	40,029 (50.2)
Only-child status, n (%)		
Yes	4506 (50.9)	37,829 (47)
No	4347 (49.1)	42,660 (53)
Annual family income (¥)[c], n (%)		
<¥6000	2287 (25.8)	24,020 (29.8)
¥6000-¥14,000	2831 (32)	26,135 (32.5)
¥14,000-¥23,000	1606 (18.1)	13,399 (16.6)
¥23,000-¥36,000	948 (10.7)	7894 (9.8)
¥36,000-¥70,000	715 (8.1)	5383 (6.7)
>¥70,000	466 (5.3)	3658 (4.5)
Cigarette usage, n (%)		
No	7927 (89.5)	69,456 (86.3)
Conventional cigarette	454 (5.1)	7161 (8.9)
Electronic cigarette	53 (0.6)	348 (0.4)
Dual usage	419 (4.7)	3524 (4.4)

[a] 48 individuals did not answer their age.

[b] 13 individuals did not answer their education background.

[c] ¥1=US $0.14.

Table 2. Intergroup differences between sexual and gender minority and cis-heterosexual participants.[a]

Variables	Sexual and gender minority participants (n=8853)	Cis-heterosexual participants (n=80,489)	*z* score	χ^2 (*df*)	*P* value
Conventional cigarettes					
Dependance, mean (SD)	0.17 (0.82)	0.24 (0.95)	–9.25	N/A[b]	<.001
Classification, n (%)				37.6 (4)	<.001
Very low (0-2)	8590 (97)	77,045 (95.7)			
Low (3-4)	155 (1.8)	2189 (2.7)			
Medium (5)	53 (0.6)	607 (0.8)			
High (6-7)	44 (0.5)	562 (0.7)			
Very high (8-10)	11 (0.1)	86 (0.1)			
e-Cigarettes					
Dependance, mean (SD)	0.20 (1.21)	0.14 (0.95)	3.20	N/A	.001
Suicidality					
Mean (SD)	6.19 (3.44)	4.11 (2.12)	69.91	N/A	<.001
Classification, n (%)				4507.5 (1)	<.001
Low risk (3-6)	5487 (62)	71,073 (88.3)			
High risk (7-18)	3366 (38)	9416 (11.7)			
Nonsuicidal self-injury					
Mean (SD)	0.53 (1.20)	0.13 (0.58)	54.32	N/A	.001
Classification, n (%)				3070.9 (2)	<.001
Never	6869 (77.6)	75,460 (93.8)			
Less than 4 days in the past year	1322 (14.9)	3885 (4.8)			
5 or more days in the past year	662 (7.5)	1144 (1.4)			

[a] The Mann-Whitney *U* test was performed for nonnormally distributed continuous variables (eg, scores) and chi-square test was performed for categorical variables (eg, classifications).

[b] N/A: not applicable.

Table 3 summarizes the effects of cigarette usage and cigarette dependence on NSSI. When only cigarette usage was included in the model (model 1), CC usage, EC usage, and dual usage were associated with increased risks of NSSI when compared to no usage (B=0.12, *P*<.001; B=0.24, *P*<.001; B=0.29, *P*<.001, respectively). However, when cigarette usage and cigarette dependence were both included in the model (model 3), EC usage was not associated with increased risks of NSSI when compared to no usage (B=0.05, *P*=.13). Regardless of whether cigarette dependence was included alone (model 2) or in conjunction with cigarette usage (model 3), EC dependence and CC dependence were associated with increased risks of NSSI (CCs: B=0.04, *P*<.001; B=0.02, *P*<.001; ECs: B=0.06, *P*<.001; B=0.05, *P*<.001, respectively).

Table 3. Effects of cigarette usage and dependence on nonsuicidal self-injury.[a]

Variables	Model 1		Model 2		Model 3	
	B (95% CI)	*P* value	B (95% CI)	*P* value	B (95% CI)	*P* value
Cigarette usage						
No cigarette usage	1 (Reference)	Reference	N/A[b]	N/A	1 (Reference)	Reference
Conventional cigarette	0.12 (0.10 to 0.13)	<.001	N/A	N/A	0.08 (0.06 to 0.10)	<.001
Electronic cigarette	0.24 (0.17 to 0.30)	<.001	N/A	N/A	0.05 (–0.02 to 0.12)	.13
Dual usage	0.29 (0.26 to 0.31)	<.001	N/A	N/A	0.10 (0.07 to 0.12)	<.001
Conventional cigarette dependence	N/A	N/A	0.04 (0.03 to 0.04)	<.001	0.02 (0.02 to 0.03)	<.001
Electronic cigarette dependence	N/A	N/A	0.06 (0.05 to 0.06)	<.001	0.05 (0.04 to 0.06)	<.001

[a] Model adjusted for age, sex assigned at birth, ethnicity, educational background, residence areas prior to university/college enrollment, only-child status, and annual family income.

[b] N/A: not applicable.

Table 4 summarizes the effects of cigarette usage and cigarette dependence on suicidality. When only cigarette usage was included in the model (model 1), CC usage, EC usage, and dual usage were associated with increased risks of suicidality when compared to no usage (B=0.22, *P*<.001; B=0.75, *P*<.001; B=0.82, *P*<.001, respectively). However, when both cigarette usage and cigarette dependence were included in the model (model 3), CC usage was not associated with increased risks of suicidality when compared to no usage (B=0.07, *P*=.07). Regardless of whether cigarette dependence was included alone (model 2) or in conjunction with cigarette usage (model 3), CC dependence and EC dependence were associated with increased risks of suicidality (CCs: B=0.11, *P*<.001; B=0.09, *P*<.001; ECs: B=0.17, *P*<.001; B=0.14, *P*<.001, respectively).

Table 4. Effects of cigarette usage and dependence on suicidality.[a]

Variables	Model 1		Model 2		Model 3	
	B (95% CI)	*P* value	B (95% CI)	*P* value	B (95% CI)	*P* value
Cigarette usage						
No cigarette usage	1 (Reference)	Reference	N/A[b]	N/A	1 (Reference)	Reference
Conventional cigarette	0.22 (0.17 to 0.28)	<.001	N/A	N/A	0.07 (–0.01 to 0.13)	.07
Electronic cigarette	0.75 (0.52 to 0.98)	<.001	N/A	N/A	0.25 (0.01 to 0.49)	.045
Dual usage	0.82 (0.74 to 0.89)	<.001	N/A	N/A	0.24 (0.14 to 0.34)	<.001
Conventional cigarette dependence	N/A	N/A	0.11 (0.09 to 0.13)	<.001	0.09 (0.07 to 0.11)	<.001
Electronic cigarette dependence	N/A	N/A	0.17 (0.15 to 0.18)	<.001	0.14 (0.12 to 0.16)	<.001

[a] Model adjusted for age, sex assigned at birth, ethnicity, educational background, residence areas prior to university/college enrollment, only-child status, and annual family income.

[b] N/A: not applicable.

Besides, the multivariable linear regression models showed that cigarette usage has a unique effect on NSSI and suicidality (Table 5). Specifically, the results demonstrated that CC usage, EC usage, and dual usage increased risks for both NSSI and suicidality when compared to no usage (CC: B=0.09, *P*<.001; B=0.22, *P*<.001; EC: B=0.20, *P*<.001; B=0.64, *P*<.001; dual usage: B=0.22, *P*<.001; B=0.68, *P*<.001, respectively). The results also showed that SGM participants had increased risks of NSSI and suicidality when compared to cisgender heterosexual participants (B=0.35, *P*<.001; B=1.94, *P*<.001, respectively). Table 5 underscores the significant interactions of (1) CC usage and group type on NSSI and suicidality (B=0.34, *P*<.001; B=0.24, *P*=.03, respectively) and (2) dual usage and group type on NSSI and suicidality (B=0.54, *P*<.001; B=0.84, *P*<.001, respectively). No significant interaction effect was observed between EC and group type on NSSI and suicidality (B=0.15, *P*=.12; B=0.33, *P*=.32, respectively).

Table 5. Interaction effects of cigarette usage with group models.[a]

Variables	Nonsuicidal self-injury		Suicidality	
	B (95% CI)	*P* value	B (95% CI)	*P* value
Main effects				
Cigarette usage				
No	1 (Reference)	Reference	1 (Reference)	Reference
Conventional cigarette	0.09 (0.08 to 0.11)	<.001	0.22 (0.16 to 0.28)	<.001
Electronic cigarette	0.20 (0.13 to 0.27)	<.001	0.64 (0.40 to 0.88)	<.001
Dual usage	0.22 (0.19 to 0.24)	<.001	0.68 (0.60 to 0.75)	<.001
Group[b]	0.35 (0.34 to 0.37)	<.001	1.94 (1.89 to 1.99)	<.001
Interaction effects				
Cigarette usage × group				
Conventional cigarette × group	0.34 (0.28 to 0.41)	<.001	0.24 (0.02 to 0.47)	.03
Electronic cigarette × group	0.15 (–0.04 to 0.35)	.12	0.33 (–0.33 to 0.99)	.32
Dual usage × group	0.54 (0.47 to 0.61)	<.001	0.84 (0.60 to 1.07)	<.001

[a] Model adjusted for age, sex assigned at birth, ethnicity, educational background, residence areas prior to university/college enrollment, only-child status, and annual family income.

[b] Group (0=cis-heterosexual participants; 1=sexual and gender minority participants).

In addition, the multivariable linear regression models showed that CC dependence and EC dependence have a unique effect on NSSI and suicidality (Table 6). Specifically, the results demonstrated that participants' CC and EC dependence simultaneously increased risks for both NSSI and suicidality (CCs: B=0.03, *P*<.001; B=0.12, *P*<.001; ECs: B=0.04, *P*<.001; B=0.13, *P*<.001, respectively). The results also showed that SGM participants had higher levels of NSSI and suicidality when compared to the cisgender heterosexual participants (B=0.38, *P*<.001; B=1.97, *P*<.001, respectively). Table 6 underscores the significant interactions of (1) CC dependence and group type (ie, SGM or cis-heterosexual) on NSSI (B=0.07, *P*<.001) and (2) EC dependence and group type on NSSI and suicidality (B=0.04, *P*<.001; B=0.09, *P*<.001, respectively). No significant interaction effect between CC dependence and group type on suicidality was found (B=–0.01, *P*=.72).

Table 6. Interaction effects of cigarette dependence with group models.[a]

Variables	Nonsuicidal self-injury		Suicidality	
	B (95% CI)	*P* value	B (95% CI)	*P* value
Main effects				
Conventional cigarette dependence	0.03 (0.03 to 0.04)	<.001	0.12 (0.10 to 0.13)	<.001
Electronic cigarette dependence	0.04 (0.04 to 0.05)	<.001	0.13 (0.12 to 0.15)	<.001
Group[b]	0.38 (0.36 to 0.39)	<.001	1.97 (1.92 to 2.02)	<.001
Interaction effects				
Conventional cigarette dependence × group	0.07 (0.05 to 0.09)	<.001	–0.01 (–0.08 to 0.06)	.72
Electronic cigarette dependence × Group	0.04 (0.03 to 0.06)	<.001	0.09 (0.04 to 0.14)	<.001

[a] Model adjusted for age, sex assigned at birth, ethnicity, educational background, residence areas prior to university/college enrollment, only-child status, and annual family income.

[b] Group (0=cis-heterosexual participants; 1=sexual and gender minority participants).

Discussion

Principal Findings

To our knowledge, this is the first study examining the usage of different cigarette types (ie, CCs and ECs) and their relationship with NSSI and suicidality while investigating intergroup differences between Chinese cis-heterosexual and SGM youth by using data from a large-scale cross-sectional survey. We found that CC usage and dependence were more prevalent among cis-heterosexual individuals, whereas EC usage and dependence were more prevalent among SGM individuals. We found that only CC usage or EC usage increased individuals' risks for both NSSI and suicidality; in particular, dual usage increased these risks. We also found that CC dependence and EC dependence could simultaneously impact individuals' risks for NSSI and suicidality independently, and the impacts of each could overlay with the usage of CCs and ECs.

Our findings shed light on the existing intergroup differences in susceptibility to NSSI and suicidality risks. Although CC usage, EC usage, and dual usage increased an individual's risk of NSSI and suicidality compared with no usage, we found that only CC usage and dual usage increased the risks of NSSI and suicidality in the SGM population compared to those in the cis-heterosexual population. These results are supported by previous research [34-42]. SGM youth are susceptible to the disproportionately high risk of CC usage due to societal, peer, and community influence, which are the major risk factors for NSSI and suicidality. Such societal, peer, and community influences could also be applied to EC usage with its increasing advertisements on social media, and EC usage could be a significant risk factor for mental health risks. Furthermore, recent studies [34-43] underline the fact that there is a significant overlap between CC and EC users in the SGM population, supporting our finding that individuals with dual usage have the highest risks of NSSI and suicidality compared to individuals with no usage of CC and EC.

We found that both CC dependence and EC dependence increased an individual's risk of NSSI and suicidality. Specifically, we found that both CC and EC dependence increased the risks of NSSI in the SGM population compared to that in the cis-heterosexual population. However, only EC dependence increased the risk of suicidality in the SGM population compared to that in the cis-heterosexual population. Although only a relatively limited amount of research has focused on investigating EC dependence, findings showed that youth with EC usage could experience dependence symptoms (ie, higher usage frequency compared to CC) unique to ECs, creating more barriers to EC cessation [43]. Considering that SGM youth could be more influenced by EC due to societal, peer, and community influence and EC has been deemed as the major risk factor for SGM population's mental health, we believe our findings on EC dependence and its impact on SGM youth could be valuable for further research. Thus, our findings add to the emerging breadth of evidence emphasizing the imminent risks posed by CC and EC usage and dependence among youth.

The existing intergroup difference between SGM and cis-heterosexual college students in NSSI and suicidality risks resulting from CC and EC usage and dependence can be explained by SGM individuals' less stable mental health and their vulnerability toward CC usage and EC dependence. First, findings from US and international studies have consistently underlined the continuing worsening situation of SGM youth's mental health [34-37]. Such a disparity arises from a wide range of structural vulnerabilities such as unavoidable systemic discrimination and prejudice in education, employment, and health care opportunities [38,39]. Further, these structural vulnerabilities stem from the imbalanced hierarchical social power structure that is responsible for generating and widening the mental health gap between SGM and cis-heterosexual individuals while limiting SGM individuals' accessibility to determinants of health, leaving the SGM population defenseless against risky behaviors such as substance use [39-45].

As previously mentioned, a number of explanations have been proposed in past research, including being impacted by the marketing target of the tobacco industry on the media scale, community norms on a group scale, and minority stressors on a personal scale. First, it is well-established that SGM youth are at disproportionately high risk of CC usage [41,42]. Past research also suggests that the SGM population has been the major marketing target of the tobacco industry, resulting in increased difficulties for SGM youth to resist CC usage when such behavior is normalized by the public media [44]. With the increase in EC usage, recent research underlines that the same aggressive marketing strategy has also been employed by the EC industry and public media. To be more precise, studies indicate that SGM youth were more likely to report being exposed to, searching for, sharing, or being shared with EC–related content on the news or social media platforms compared to their non-SGM peers [46-50]. Second, from a community standpoint, past research has pointed out that the community context could be an influential factor in SGM youth's smoking behavior [47-50]. Under the influence of increasing exposure to EC-related content among the SGM population, it becomes reasonable to assume that, in social situations with SGM peers (ie, one of the potential triggers for EC usage and combined CC and EC usage), SGM youth would become more susceptible to the rising trend of EC usage among their peers [48]. Third, on a personal scale, as mentioned earlier, the minority stress model underscores that unique stressors experienced by the SGM population, including but not limited to discrimination and internalized homo/bi/transphobia, could result in higher stress levels among this stigmatized group [40-45]. Such stress could then contribute to the thought pattern of EC usage being a way for individuals to express or assert their gender identity or to rebel against traditional gender norms [40-45]. Due to all the aforementioned reasons, concerted efforts need to be made on a societal level in terms of educating SGM youth about the misconceptions, safety concerns, and, most importantly, adverse mental health outcomes such as increased risks for NSSI and suicidality that could be potentially induced by EC usage.

From a broader perspective, the rise of EC usage among Chinese youth has become a public health concern in recent years. Unlike the easy accessibility of the adverse effects of CC usage from various reliable sources, youth often receive inaccurate information about EC usage, resulting in a higher likelihood of maintaining their positive perception of ECs. Under the massive impact created by extensive marketing campaigns, ECs have been promoted as less harmful than CCs and, therefore, have rapidly gained popularity among the younger generation [45]. Compared to the downhill journey of CC usage among Chinese youth (ie, a significant reduction from 12.9% in 2014 to 3.9% in 2019), the rise in EC usage (ie, from 1.2% in 2014 to 2.7% in 2019) reveals that smoking among youth is still a major public health concern in China, many of whom are not fully aware of the physical and mental health risks induced by EC usage [46]. Although camouflaged by the utility for smoking cessation, nonsecondhand smoke exposure, and easy accessibility, ECs' negative impact on youth's physical and psychological well-being cannot be ignored. Considering EC's relationship with NSSI and suicidality, efforts to prevent youth from EC addiction must be made on a societal level [50].

It is worth noting that Chinese government officials and policy makers have promulgated the *Law on Youth Protection* and published the *Notice of the "Guarding Youth" Special Action Plan for Youth Protection from Tobacco Abuse* in 2021 while further strengthening the supervision of EC products, regulating the market order, and standardizing industrial governance by putting into effect the Measures for the Administration of Electronic Cigarettes as of May 1, 2022 [50]. The implementation of these measures has prohibited youth younger than 18 years from purchasing EC products, which could correspondingly reduce their NSSI and suicidality risks. Nevertheless, to prevent the situation from further worsening, future efforts are still required from a societal level to regulate the aggressive marketing of the EC industry and to expand the media and school coverage of educational campaigns in formats such as nonprofit advertisements, school lectures, and short videos on social media platforms on EC usage prevention and its adverse mental health effects on youth.

Limitations

The findings from this study must be considered in light of a few limitations. First, findings from this study are limited to Chinese SGM college students and, thus, may not be generalizable to SGM youth of other cultural backgrounds or who reside in other countries. Nevertheless, our findings provide crucial insights into the disparity between SGM and cis-heterosexual youth with regard to EC usage, propelling future studies to further investigate the extent and nature of such disparity. Second, although the effect size found in this study was relatively small, our large sample size indicates that such an effect was still stable on a large population scale. Specifically, the impact of cigarette usage and dependence on suicidality and NSSI in SGM youth is more pronounced than that on their cis-heterosexual peers, highlighting the need from a public health standpoint to pay more attention to the impact of EC and CC on suicidality and NSSI among the SGM population. Third, due to the cross-sectional nature of this study, only associations were examined. Longitudinal data are required to investigate these associations over time. Fourth, the sex assigned at birth of the SGM participants in this study was mostly female (6531/8853, 73.8%), which could potentially incur sampling bias and limit the study findings' generalizability.

Conclusions and Implications

Overall, our findings have crucial implications for understanding the disparity of NSSI and suicidality risks between SGM and cis-heterosexual youth. First, this study contributes to the growing body of literature highlighting the mental health risks induced by CC usage and EC usage, emphasizing that SGM youth are more susceptible to EC usage. Second, since the sample in this study consisted of Chinese college students, these results could suggest that relevant school policies (eg, CC-free and EC-free campus policies that prohibit EC usage in school public areas and dormitories) and other school-level activities (eg, informative workshops directed by student health centers on EC usage, SGM susceptibility to EC usage, and its physical and mental health risks) could be enacted and hosted to address the health disparity among SGM students. Feedback surveys could be distributed for those policies and school-level interventions to ensure that these interventions attenuate SGM students' needs rather than perpetuate and exacerbate SGM stereotypical views. Third, findings from this study could inform governmental officials and policy makers on the vulnerability of SGM youth to EC usage. For instance, policy makers could further develop comprehensive EC control policies that are similar to the *Notice of the "Guarding Youth" Special Action Plan for Youth Protection from the Tobacco Abuse* in 2021 but with an emphasis on reducing EC usage and content exposure among SGM youth. Overall, our study provides clear guidance for influential parties to develop effective strategies to resolve the rising imminent public health concern of EC usage among SGM youth.

Acknowledgments

We would like to thank and appreciate all the college students in the Jilin province, China, who participated in this study, for taking the effort to participate in the survey, and be willing to share their information; the Vanke School of Public Health at Tsinghua University for the support; and all the research assistants and students at the Youth Mental Health Lab for their dedication.

Authors' Contributions

SX and RC contributed equally as co-corresponding authors. Yinzhe W, SX, and RC contributed to study conception and study design. SX performed data collection. Yinzhe W and YZ conducted data analysis. XZ, YF, and Yinzhe W interpreted the data. Yinzhe W, XZ, Yuanyuan W, YZ, SX, and RC contributed to paper writing and revision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Electronic cigarette dependence in sexual and gender minority and cis-heterosexual participants.
[DOCX File , 16 KB - publichealth_v9i1e47058_app1.docx]

References

1. Centers for Disease ControlPrevention (CDC). Notes from the field: electronic cigarette use among middle and high school students - United States, 2011-2012. MMWR Morb Mortal Wkly Rep 2013 Sep 06;62(35):729-730 [FREE Full text] [Medline: 24005229]
2. Xie W, Tackett AP, Berlowitz JB, Harlow AF, Kathuria H, Galiatsatos P, et al. Association of electronic cigarette use with respiratory symptom development among U.S. young adults. Am J Respir Crit Care Med 2022 Jun 01;205(11):1320-1329 [FREE Full text] [doi: 10.1164/rccm.202107-1718oc]
3. Livingston J, Chen CH, Kwon M, Park E. Physical and mental health outcomes associated with adolescent E-cigarette use. J Pediatr Nurs 2022;64:1-17. [doi: 10.1016/j.pedn.2022.01.006] [Medline: 35121206]
4. Griffin K, Twynstra J, Gilliland JA, Seabrook JA. Correlates of self-harm in university students: A cross-sectional study. J Am Coll Health 2023 Apr;71(3):959-966. [doi: 10.1080/07448481.2021.1909049] [Medline: 34242122]
5. Turner B, Helps CE, Ames ME. Stop self-injuring, then what? Psychosocial risk associated with initiation and cessation of nonsuicidal self-injury from adolescence to early adulthood. J Psychopathol Clin Sci 2022 Jan;131(1):45-57. [doi: 10.1037/abn0000718] [Medline: 34843270]
6. Li D, Yang X, Ge Z, Hao Y, Wang Q, Liu F, et al. Cigarette smoking and risk of completed suicide: a meta-analysis of prospective cohort studies. J Psychiatr Res 2012 Oct;46(10):1257-1266. [doi: 10.1016/j.jpsychires.2012.03.013] [Medline: 22889465]
7. Breslau N, Schultz LR, Johnson EO, Peterson EL, Davis GC. Smoking and the risk of suicidal behavior: a prospective study of a community sample. Arch Gen Psychiatry 2005 Mar;62(3):328-334. [doi: 10.1001/archpsyc.62.3.328] [Medline: 15753246]
8. Han B, Compton WM, Blanco C. Tobacco use and 12-month suicidality among adults in the United States. Nicotine Tob Res 2017 Jan;19(1):39-48 [FREE Full text] [doi: 10.1093/ntr/ntw136] [Medline: 27190402]
9. Cheng T. Chemical evaluation of electronic cigarettes. Tob Control 2014 May;23 Suppl 2(Suppl 2):ii11-ii17 [FREE Full text] [doi: 10.1136/tobaccocontrol-2013-051482] [Medline: 24732157]
10. Wang W, He Z, Feng N, Cai Y. Electronic cigarette use in China: awareness, prevalence and regulation. Tob Induc Dis 2019;17:30 [FREE Full text] [doi: 10.18332/tid/105393] [Medline: 31516473]
11. You Z, Chen M, Yang S, Zhou Z, Qin P. Childhood adversity, recent life stressors and suicidal behavior in Chinese college students. PLoS One 2014;9(3):e86672 [FREE Full text] [doi: 10.1371/journal.pone.0086672] [Medline: 24681891]
12. Cao C, Zhu C, Meng Q. Chinese international students' coping strategies, social support resources in response to academic stressors: does heritage culture or host context matter? Curr Psychol 2018 Jul 13;40(1):242-252 [FREE Full text] [doi: 10.1007/s12144-018-9929-0]
13. Wang Y, Feng Y, Han M, Duan Z, Wilson A, Fish J, et al. Methods of attempted suicide and risk factors in LGBTQ+ youth. Child Abuse Negl 2021 Dec;122:105352. [doi: 10.1016/j.chiabu.2021.105352] [Medline: 34655994]
14. Hottes T, Bogaert L, Rhodes A, Brennan D, Gesink D. Lifetime prevalence of suicide attempts among sexual minority adults by study sampling strategies: a systematic review and meta-analysis. Am J Public Health 2016 May;106(5):e1-e12 [FREE Full text] [doi: 10.2105/ajph.2016.303088]
15. Buchting F, Emory KT, Kim Y, Fagan P, Vera LE, Emery S. Transgender use of cigarettes, cigars, and e-cigarettes in a national study. Am J Prev Med 2017 Jul;53(1):e1-e7 [FREE Full text] [doi: 10.1016/j.amepre.2016.11.022] [Medline: 28094133]
16. LGBTQ+ people and commercial tobacco: health disparities and ways to advance health equity. Centers for Disease Control and Prevention. 2022 Jun 27. URL: https://www.cdc.gov/tobacco/disparities/lgbt/index.htm [accessed 2022-03-01]
17. Nayak P, Salazar LF, Kota KK, Pechacek TF. Prevalence of use and perceptions of risk of novel and other alternative tobacco products among sexual minority adults: Results from an online national survey, 2014-2015. Prev Med 2017 Nov;104:71-78 [FREE Full text] [doi: 10.1016/j.ypmed.2017.05.024] [Medline: 28579496]
18. Huang J, Kim Y, Vera L, Emery SL. Electronic cigarettes among priority populations: role of smoking cessation and tobacco control policies. Am J Prev Med 2016 Feb;50(2):199-209 [FREE Full text] [doi: 10.1016/j.amepre.2015.06.032] [Medline: 26410185]
19. Caputi T, Smith LR, Strathdee SA, Ayers JW. Substance use among lesbian, gay, bisexual, and questioning adolescents in the United States, 2015. Am J Public Health 2018 Aug;108(8):1031-1034 [FREE Full text] [doi: 10.2105/ajph.2018.304446]
20. Ganz O, Johnson AL, Cohn AM, Rath J, Horn K, Vallone D, et al. Tobacco harm perceptions and use among sexual and gender minorities: findings from a national sample of young adults in the United States. Addict Behav 2018 Jun;81:104-108 [FREE Full text] [doi: 10.1016/j.addbeh.2018.01.032] [Medline: 29454177]
21. Meyer I. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003 Sep;129(5):674-697 [FREE Full text] [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
22. Smith E, Thomson K, Offen N, Malone R. "If You Know You Exist, It's Just Marketing Poison": meanings of tobacco industry targeting in the lesbian, gay, bisexual, and transgender community. Am J Public Health 2008 Jun;98(6):996-1003 [FREE Full text] [doi: 10.2105/ajph.2007.118174]
23. Credamo. URL: https://www.credamo.com/#/aboutus [accessed 2021-10-26]

24. Best practices for asking questions to identify transgender and other gender minority respondents on population-based surveys (GenIUSS). UCLA Williams Institute. URL: https://williamsinstitute.law.ucla.edu/publications/geniuss-trans-pop-based-survey/ [accessed 2021-09-01]
25. Best practices for asking questions about sexual orientation on surveys (SMART). UCLA Williams Institute. URL: https://williamsinstitute.law.ucla.edu/publications/smart-so-survey/ [accessed 2021-09-01]
26. Huang C, Lin HH, Wang HH. The psychometric properties of the Chinese version of the Fagerstrom Test for Nicotine Dependence. Addict Behav 2006 Dec;31(12):2324-2327. [doi: 10.1016/j.addbeh.2006.02.024] [Medline: 16567055]
27. Fagerström KO. Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. Addict Behav 1978;3(3-4):235-241. [doi: 10.1016/0306-4603(78)90024-2] [Medline: 735910]
28. Lam T, Abdullah ASM, Chan SSC, Hedley AJ, Hong Kong Council on SmokingHealth Smoking Cessation Health Centre (SCHC) Steering Group. Adherence to nicotine replacement therapy versus quitting smoking among Chinese smokers: a preliminary investigation. Psychopharmacology (Berl) 2005 Feb;177(4):400-408. [doi: 10.1007/s00213-004-1971-y] [Medline: 15289997]
29. Zhang X, Al Jurdi RK, Zoghbi AW, Chen DC, Xiu MH, Tan YL, et al. Prevalence, demographic and clinical correlates of suicide attempts in Chinese medicated chronic inpatients with schizophrenia. J Psychiatr Res 2013 Oct;47(10):1370-1375. [doi: 10.1016/j.jpsychires.2013.05.024] [Medline: 23791457]
30. Osman A, Bagge CL, Gutierrez PM, Konick LC, Kopper BA, Barrios FX. The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. Assessment 2001 Dec;8(4):443-454. [doi: 10.1177/107319110100800409] [Medline: 11785588]
31. Huen JMY, Yip PSF, Osman A, Leung ANM. The Suicidal Behaviors Questionnaire-Revised (SBQ-R) and its Chinese version (C-SBQ-R): Further validity testing using the culture, comprehension, and translation bias procedure. Psychol Assess 2022 Jul;34(7):704-710. [doi: 10.1037/pas0001134] [Medline: 35467908]
32. Online assessment measures: clinician-rated severity of nonsuicidal self-injury. American Psychiatric Association. URL: https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM5_Clinician-Rated-Severity-of-Non-Suicidal-Self-Injury.pdf [accessed 2021-09-01]
33. Hayes A, Coutts JJ. Use omega rather than Cronbach's alpha for estimating reliability. But…. Communication Methods and Measures 2020 Feb 11;14(1):1-24 [FREE Full text] [doi: 10.1080/19312458.2020.1718629]
34. Marshal M, Dietz LJ, Friedman MS, Stall R, Smith HA, McGinley J, et al. Suicidality and depression disparities between sexual minority and heterosexual youth: a meta-analytic review. J Adolesc Health 2011 Aug;49(2):115-123 [FREE Full text] [doi: 10.1016/j.jadohealth.2011.02.005] [Medline: 21783042]
35. Eskin M, Kaynak-Demir H, Demir S. Same-sex sexual orientation, childhood sexual abuse, and suicidal behavior in university students in Turkey. Arch Sex Behav 2005 Apr;34(2):185-195. [doi: 10.1007/s10508-005-1796-8] [Medline: 15803252]
36. Wang Y, Hu Z, Peng K, Xin Y, Yang Y, Drescher J, et al. Discrimination against LGBT populations in China. The Lancet Public Health 2019 Sep;4(9):e440-e441 [FREE Full text] [doi: 10.1016/s2468-2667(19)30153-7]
37. Sun S, Pachankis JE, Li X, Operario D. Addressing minority stress and mental health among men who have sex with men (MSM) in China. Curr HIV/AIDS Rep 2020 Feb;17(1):35-62 [FREE Full text] [doi: 10.1007/s11904-019-00479-w] [Medline: 31950336]
38. Bourgois P, Holmes SM, Sue K, Quesada J. Structural vulnerability. Academic Medicine 2017;92(3):299-307 [FREE Full text] [doi: 10.1097/acm.0000000000001294]
39. Hatzenbuehler M, Phelan J, Link B. Stigma as a fundamental cause of population health inequalities. Am J Public Health 2013 May;103(5):813-821 [FREE Full text] [doi: 10.2105/ajph.2012.301069]
40. Watson RJ, Fish JN, Denary W, Caba A, Cunningham C, Eaton LA. LGBTQ state policies: A lever for reducing SGM youth substance use and bullying. Drug Alcohol Depend 2021 Apr 01;221:108659 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108659] [Medline: 33676075]
41. Wheldon C, Wiseman K. Tobacco use among transgender and gender non-conforming adults in the United States. Tob Use Insights 2019 May 23;12:1179173X1984941 [FREE Full text] [doi: 10.1177/1179173x19849419]
42. Goodin A, Fallin-Bennett K, Anderson-Hoagland E, Fallin-Bennett A. Tobacco use and mental health disparities in LGB youth. Public Health Nurs 2022 May;39(3):659-663. [doi: 10.1111/phn.12997] [Medline: 34672011]
43. Case K, Mantey DS, Creamer MR, Harrell MB, Kelder SH, Perry CL. E-cigarette-specific symptoms of nicotine dependence among Texas adolescents. Addict Behav 2018 Sep;84:57-61 [FREE Full text] [doi: 10.1016/j.addbeh.2018.03.032] [Medline: 29627634]
44. Williams JR, Housman JM, McDonald JD. If we build it, will they come? challenges of adapting and implementing a smoking cessation program for the LGBTQ community in Southcentral Texas. Health Promot Pract 2020 Jan;21(1_suppl):118S-123S. [doi: 10.1177/1524839919882385] [Medline: 31908192]
45. Grana R, Benowitz N, Glantz S. E-Cigarettes. Circulation 2014 May 13;129(19):1972-1986 [FREE Full text] [doi: 10.1161/circulationaha.114.007667]

46. Liu X, Xiao L, Zeng X, et al. Preplanned studies: tobacco use and exposure among secondary school students—China, 2019. Chinese Center for Disease Control and Prevention. URL: https://www.chinacdc.cn/jkzt/sthd_3844/slhd_12885/202005/t20200531_216942.html [accessed 2022-03-21]
47. Emory K, Buchting FO, Trinidad DR, Vera L, Emery SL. Lesbian, gay, bisexual, and transgender (LGBT) view it differently than non-LGBT: exposure to tobacco-related couponing, e-cigarette advertisements, and anti-tobacco messages on social and traditional media. Nicotine Tob Res 2019 Mar 30;21(4):513-522 [FREE Full text] [doi: 10.1093/ntr/nty049] [Medline: 29546337]
48. Hatzenbuehler ML, Wieringa NF, Keyes KM. Community-level determinants of tobacco use disparities in lesbian, gay, and bisexual youth: results from a population-based study. Arch Pediatr Adolesc Med 2011 Jun;165(6):527-532 [FREE Full text] [doi: 10.1001/archpediatrics.2011.64] [Medline: 21646585]
49. Roberts L, Heyworth B, Gilliver A, Mackereth P. Smoking and vaping among lesbian, gay, bisexual and trans people: results of a Proud2BSmokefree survey. Cancer Nursing Practice 2017 Jul 07;16(10):1-3 [FREE Full text] [doi: 10.7748/cnp.2017.e1435]
50. Zhong S, Wang Y, Chen R, Zhou L. E-cigarette use among youth in China. Lancet Reg Health West Pac 2022 Jul;24:100504 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100504] [Medline: 35789735]

Abbreviations

CC: conventional cigarette
EC: electronic cigarette
FTND: Fagerström Test of Nicotine Dependence
NSSI: nonsuicidal self-injury
SBQ-R: Suicidal Behaviors Questionnaire-Revised
SGM: sexual and gender minority

Edited by A Mavragani, T Sanchez; submitted 07.03.23; peer-reviewed by F Lu, Y Yang; comments to author 12.04.23; revised version received 11.05.23; accepted 24.05.23; published 07.07.23.

Please cite as:
Wang Y, Xu S, Zhang X, Zhang Y, Feng Y, Wang Y, Chen R
Effects of Tobacco Versus Electronic Cigarette Usage on Nonsuicidal Self-Injury and Suicidality Among Chinese Youth: Cross-Sectional Self-Report Survey Study
JMIR Public Health Surveill 2023;9:e47058
URL: https://publichealth.jmir.org/2023/1/e47058
doi:10.2196/47058
PMID:37418293

Original Paper

A Multidimensional Assessment of Activities of Daily Living, Mental Status, Communication, and Social Abilities Among Older Adults in Shenzhen, China: Cross-Sectional Study

Jing Wang[1], BM; Patrick Kwan[2,3,4], Prof Dr; Gong Zhang[5], Prof Dr; Mingwang Shen[1,6], PhD; Loretta Piccenna[2,3], PhD; Terence J O'Brien[2,3,4], Prof Dr; Lei Zhang[1,7,8], Prof Dr

[1]China-Australia Joint Research Centre for Infectious Diseases, School of Public Health, Xi'an Jiaotong University Health Science Centre, Xi'an, Shaanxi, China

[2]Department of Neuroscience, The Central Clinical School, Monash University, Melbourne, Victoria, Australia

[3]Department of Neurology, Alfred Health, Melbourne, Victoria, Australia

[4]Departments of Medicine and Neurology, The Royal Melbourne Hospital, The University of Melbourne, Parkville, Victoria, Australia

[5]Ministry of Education Key Laboratory of Tumour Molecular Biology and Key Laboratory of Functional Protein Research of Guangdong Higher Education Institutes, Institute of Life and Health Engineering, College of Life Science and Technology, Jinan University, Guangzhou, China

[6]Key Laboratory for Disease Prevention and Control and Health Promotion of Shaanxi Province, Xi'an, Shaanxi, China

[7]Artificial Intelligence and Modelling in Epidemiology Program, Melbourne Sexual Health Centre, Alfred Health, Melbourne, Australia

[8]Central Clinical School, Faculty of Medicine, Monash University, Melbourne, Australia

Corresponding Author:
Lei Zhang, Prof Dr
China-Australia Joint Research Centre for Infectious Diseases
School of Public Health
Xi'an Jiaotong University Health Science Centre
76 Yanta West Road
Xi'an, Shaanxi, 710061
China
Phone: 86 15910593477
Email: Lei.Zhang1@monash.edu

Abstract

Background: China is facing a rapidly expanding aging population. Insights into the health status of older adults are of great significance for health resource allocation and health care provision to this population.

Objective: With the goal of providing a comprehensive understanding of the health status of older adults and to inform potential interventions, we investigated the level of disability and identified risk factors associated with disability among the older population (aged ≥60 years) living in China.

Methods: A total of 8467 older adults living in the Chinese city of Shenzhen were enrolled in this cross-sectional study. We used a multidimensional ability assessment survey, which assessed their activities of daily living (ADL; including eating, bathing, grooming, dressing, defecation control, urination control, using a toilet unaided, transfer, flat-ground walking, stair activity), mental status (including cognitive function, aggressive behavior, depression symptoms), sensory and communication (including consciousness level, vision, hearing, communication), and social participation (including living, working, time/space orientation, distinguish persons, social communication) abilities. The impact of demographic risk factors on ability levels was analyzed using ordinal logistic regression. The correlations between the four dimensions of ability mentioned above were analyzed using Spearman correlation analysis.

Results: A total of 7766 participants were effectively assessed. The participants' average age was 70.64 (SD 8.46) years comprising 56.53% females. The overall ability level was classified as mildly, moderately, and severely impaired for 27.57% (n=2141), 2.83% (n=220), and 4.28% (n=332) of the 7766 participants, respectively. With increasing age, the proportion of impaired participants increased from 17.62% (365/2071) in the age group 60-64 years to 91.3% (253/277) in the age group above 90 years (*P*<.001), corresponding to an approximate 10% rise for every 5-year age increment. The odds of having more severe overall ability impairment in females was 1.15 times that in males (odds ratio [OR] 1.15, 95% CI 1.04-1.28). Participants who were divorced or widowed had a higher risk of more severe overall ability impairment than those currently married (OR 1.98, 95% CI 1.68-2.33). Participants living with nonrelatives had an increased risk of more severe overall ability impairment than those living alone (OR 2.38, 95% CI 1.46-3.91). Higher education level was a protective factor of overall ability impairment (college degree or above: OR 0.32, 95% CI 0.24-0.42). The four dimensions of ability assessed were significantly correlated; a low score for ADL was significantly correlated with poorer mental status, sensory and communication, and social participation (all *P*<.001).

Conclusions: The proportion of disability among Chinese older adults increases with age, being female, having lower education levels, being divorced or widowed, and living with nonrelatives. Impairment in ADL ability is significantly correlated with poor mental status, social participation, and sensory and communication abilities. A holistic approach to improving the health of the older population is recommended in China.

(JMIR Public Health Surveill 2023;9:e43612) doi:10.2196/43612

KEYWORDS

older adults; ability assessment; activities of daily living; mental status; sensory and communication; social participation

Introduction

Population aging is accelerating worldwide due to declining birth rates, economic development, and improved medical and health services that extend human life longevity [1]. It is estimated that the world's population aged 60 years and over will rise by 56% from 901 million to 1.4 billion during 2015-2030 [2]. In China, an estimated 176 million individuals were aged 65 years or above in 2019, accounting for 12.6% of the Chinese population. This indicates that China has entered a fast track of population aging [3]. By 2050, the older population aged 65 years or above in China is expected to reach 330 million, accounting for more than 20% of the world's older population [4]. Notably, China is currently in a stage of rapid development and the necessary resources required to address the aging population, including financial, material, and human resources, have not been fully prepared [5]. The rapidly aging population has brought substantial challenges to China's health care system, in particular with respect to the allocation of limited health resources and pension services required to secure a healthy life in older age [6].

Aging is accompanied by multiple chronic diseases such as hypertension, diabetes, cardiovascular diseases, and dementia. These conditions are often due to chronic, low-grade inflammation; macromolecular and organelle dysfunction; stem cell and progenitor changes; and cellular senescence [7,8]. According to the Fourth National Health Data Survey in China, the prevalence of chronic diseases was 64.5% in people aged above 65 years [9]. Moreover, mental health has also become a prevalent issue among the older Chinese population. Li et al [10] reported a depression prevalence of 11.6% among community-dwelling older adults and of 18.1% among older hospital inpatients in China. Physical and mental disabilities are known to have a profound impact on the quality of life experienced by older adults. This may be attributed to the challenges faced by this population in creating wealth due to their disabilities, which often place them in a lower socioeconomic profile rendering them highly susceptible to social isolation [11]. China has the largest population of disabled older adults in the world [12]: approximately 41 million older people were considered disabled and semidisabled in 2015, accounting for approximately 18.3% of people aged above 60 years in the country [13]. This number is expected to increase to 62 million in 2030 and to 98 million in 2050 [14]. In this context, to improve the quality of life of older people, it is important to adopt personalized multicomponent prevention and intervention measures to improve or maintain their overall abilities [15]. Hence, multidimensional ability assessments for older adults would provide the necessary evidence to inform prevention and intervention strategies.

To date, numerous studies have been conducted to assess the multidimensional abilities of older adults in developed countries with comprehensive assessment tools used in community settings, such as the Resident Assessment Instrument for Home Care and the EASYcare instrument, which have been widely used to evaluate physical functioning, including activities of daily living (ADL), mental functioning, and social functioning [16,17]. In comparison, there have been limited studies related to multidimensional ability assessment in the older population in China. Previous studies on ability assessment in older adults mainly focused on the ability to fulfill ADL [18,19]. Yet, as mentioned, a more comprehensive approach that assesses various aspects, including physical, mental, and social function, among the older population would provide more holistic insights into this population's overall health status [20-24]. In addition, previous studies in China have reported that the overall ability of the older population was largely related to their sex, education, and living condition; however, the findings were not consistent across different settings [25,26].

In this study, we investigated a large group of older adults in the Chinese city of Shenzhen. We aimed to comprehensively assess the older population's overall ability across multiple dimensions, including ADL, mental status, sensory and communication, and social participation. We further explored the potential demographic factors associated with the impairment of overall ability and the correlation between the four dimensions of ability. The study findings will provide evidence to inform interventions, health policy, and health resources allocation for the older population in a Chinese setting.

Methods

Setting

The study was conducted across two districts (57 community centers) in the city of Shenzhen, China. The rollout of the survey and data collection was coordinated by a community-based nursing home privately founded in 2016. The nursing home provides aged care, nursing services, rehabilitation training, and mass health education for older adults living in the community. The nursing home has employed a smart information cloud platform that can generate an electronic version of the questionnaire used to assess the overall ability of the older population. The information system enabled a sanity check on the input information and reported errors if the input information is inappropriate, providing a certain accuracy guarantee for the data. Data were securely backed up and stored on the cloud platform.

Study Design

A cross-sectional study design was used and the study was conducted according to STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (see Multimedia Appendix 1). From April 2017 to December 2019, a total of 8467 older adults aged ≥60 years and owning a local household registration were recruited through the nursing home mentioned above. The nursing home's catchment area covers a total of 57 local communities and 8467 older adults. Of the 8467 recruited older adults, 7965 completed assessments that were ultimately collected, excluding those who were absent at the time of the assessment or chose not to participate. During the assessment, the participants were routinely asked to provide demographic information and complete a standardized ability assessment survey. The survey was conducted using the "Ability Assessment for Older Adults" questionnaire, which is the standardized questionnaire published by the Ministry of Civil Affairs of the People's Republic of China. This questionnaire was jointly drafted by members of the China Social Welfare Association, School of Nursing of Peking University, China Medical Women's Association, and other units regarding the existing tools for the ability assessment of older adults in China and abroad [27]. A team of 30-50 staff members from the nursing home was trained to conduct face-to-face interviews to administer the survey. After they explained the purpose and significance of the survey, the participants provided informed consent. For the participants who were unable to complete the survey independently due to physical or cognitive disabilities, family members or primary caregivers answered the survey on their behalf. The collected data were then entered into the smart information cloud platform to inform the future provision of services. In this study, we retrospectively examined the information provided by the participants at enrollment.

Inclusion and Exclusion Criteria

The inclusion criteria for participants were as follows: (1) aged 60 years or older, (2) have a local household registration and have lived there for more than 6 months, (3) willing to participate in the survey for ability assessment, and (4) completed the demographic information collection form and the overall ability assessment questionnaire. Note that criteria 1-3 were utilized for selecting participants during the field survey, whereas the fourth criterion was employed during the data analysis phase.

Participants were excluded if there was missing information on the (1) variable "accidents over the last 30 days" in the demographic information collection form, (2) ADL assessment part of the overall ability assessment questionnaire, (3) mental status assessment part of the overall ability assessment questionnaire, (4) sensory and communication assessment part of the overall ability assessment questionnaire, and (5) social participation assessment part of the overall ability assessment questionnaire. All five criteria were used for excluding participants during the data analysis phase.

Ethics Approval

Ethics approval was obtained from the Biomedical Ethics Committee of the Medical Department of Xi'an Jiaotong University (Human Ethics Approval Number 2020-10). The data collection process was carried out using a deidentified approach to ensure the confidentiality of the study participants. During the study period, participants were offered certain services, including meal delivery and room cleaning, by the nursing home mentioned above, without any form of compensation. Prior to participation in the study, all participants provided their oral informed consent, in line with established ethical guidelines and regulatory requirements governing human research studies. This crucial step ensured that participants were fully informed of the study's objectives, procedures, and potential risks, and made an informed decision to voluntarily participate in the study.

Data Collection

We collected demographic and ability assessment data from the information service platform. To fulfill one of our research objectives, we selected demographic variables that were commonly collected from all study participants and had been previously shown to contribute to the ability level of older adults [25,26]. Ultimately, we selected five variables (age, sex, education, marital status, and living situation) as the demographic variables of interest for this analysis.

The participants' overall ability assessment consisted of four main dimensions. First, the ADL assessment included 10 daily activity items (eating, bathing, grooming, dressing, defecation control, urination control, using a toilet unaided, transfer, flat-ground walking, stair activity) and had a total score of 100 points (see Multimedia Appendix 2 for details). The score was graded according to four levels (100 points, unimpaired; 65-95 points, mildly impaired; 45-60 points, moderately impaired; ≤40 points, severely impaired). Second, the mental status assessment included three items (cognitive function, aggressive behavior, depression symptoms) with 6 points in total, which were divided into four levels (0 points, unimpaired; 1 point, mildly impaired; 2-3 points, moderately impaired; 4-6 points, severely impaired). Third, the sensory and communication assessment included four items (consciousness level, vision, hearing, and communication). The assessment results were classified according to certain conditions and were divided into four levels (unimpaired, mildly impaired, moderately impaired, and severely impaired). Fourth, the social participation assessment included five items (living ability, working ability, time/space orientation, distinguish persons, social communication ability) with a total score of 20 points divided into four levels (0-2 points, unimpaired; 3-7 points, mildly impaired; 8-13 points, moderately impaired; 14-20 points, severely impaired).

The classification of the overall ability level was predetermined by the survey according to the standard of the Chinese Ministry of Civil Affairs [27]. In brief, if all ADL, mental status, and sensory and communication dimensions were classified as "unimpaired" and only social participation was classified as "unimpaired" or "mildly impaired," then the overall ability result was classified as "unimpaired." The overall ability level was defined as "mildly impaired" if it satisfied one of the following: (1) the result of ADL was "unimpaired" but at least one of mental status or sensory and communication was classified as "mildly impaired" or above or the social participation was classified as "moderately impaired"; (2) the ADL was classified as "mildly impaired" but at least one of the other three dimensions was classified as "unimpaired" or "mildly impaired." The overall ability was classified as "moderately impaired" if it satisfied one of the following: (1) the result of ADL was "mildly impaired" but the other three dimensions were all classified as "moderately impaired" or at least one dimension was classified as "severely impaired"; (2) the result of ADL was "moderately impaired" and one or two of the other three dimensions were classified as "mildly impaired" or "moderately impaired." The overall ability was defined as "severely impaired" if it satisfied one of the following: (1) the result of ADL was "severely impaired," (2) the results of four dimensions were all "moderately impaired," and (3) the result of ADL was "moderately impaired" and at least one of the other three dimensions was classified as "severely impaired." Notably, participants with cognitive impairment/dementia or mental illness would raise the original level by one level. Participants with two or more falls, choking, suicide, and missing in the past 30 days would raise the original level by one level. Participants in a coma would be assessed as "severely impaired."

Statistical Analysis

Data management and statistical analysis were completed using Microsoft Excel 2016 and R 4.1.1; statistical significance was set at *P*<.05. Continuous variables are described by mean and SD, whereas categorical variables are described by frequency (n) and percentage (%). Data between different groups were compared with the χ^2 or Fisher test for categorical data and then the proportion test was used for posthoc multiple comparisons; *P* values were adjusted using Bonferroni correction. Ordinal logistic regression was used to analyze the association between demographic characteristics and the severity of the overall ability impairment. In this model, the assessment results of the overall ability of participants were used as the dependent variable and the demographic variables that exhibited statistical significance based on the χ^2 or Fisher test were used as independent variables. The η coefficient, SE, and *t* value were estimated to determine the statistical significance of each independent variable in the model. Associations between the severity of the overall ability impairment and demographic characteristics were also assessed by odds ratios and corresponding 95% CIs. The correlation between the various assessment items was analyzed using Spearman correlation analysis.

Results

Demographic Characteristics

According to the inclusion and exclusion criteria, 7766 participants out of 8467 individuals were selected for the analysis (Figure 1). Among the 7766 participants, the average age was 70.64 (SD 8.46) years, ranging from 60 to 109 years (Table 1). There were 3376 (43.47%) males and 4390 (56.53%) females. Among the 7766 participants, 45.58% (n=3540) had primary or junior high school education levels and 13.67% (n=1062) were divorced or widowed. Most participants (3745/7766, 48.22%) were living with their children before enrollment.

Figure 1. Flow chart for the selection of participants.

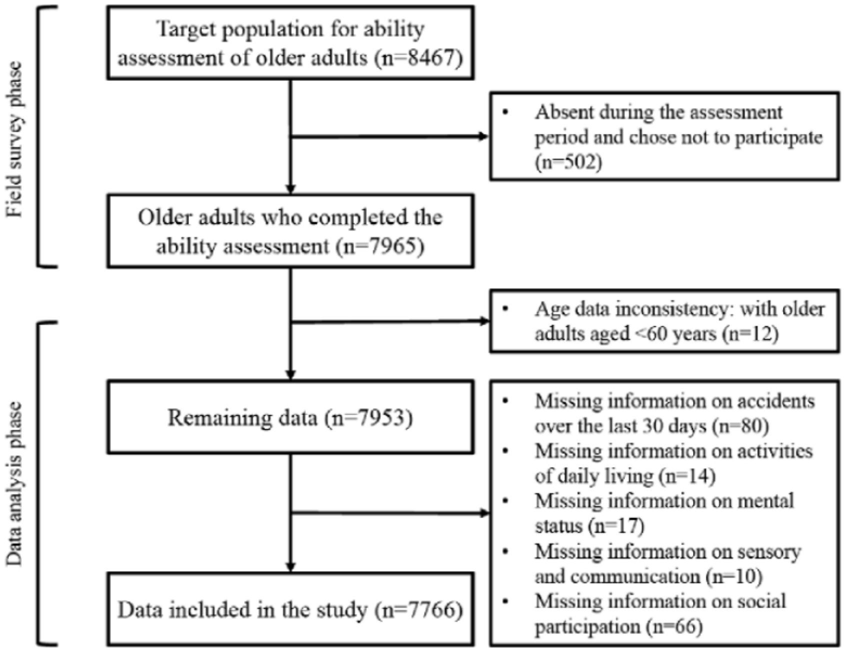


Table 1. Demographic characteristics of study participants according to ability impairment based on the Ability Assessment for Older Adults questionnaire (N=7766).

Variable	Responses, n (%)	Overall ability level, n (%)				Pearson χ^2 or Fisher exact test	*P* value
		Unimpaired (n=5073, 65.32%)	Mildly impaired (n=2141, 27.57%)	Moderately impaired (n=220, 2.83%)	Severely impaired (n=332, 4.28%)		
Age group (years)						1875.61	<.001
60-64	2071 (26.67)	1706 (82.38)	327 (15.79)	22 (1.06)	16 (0.77)		
65-69	2267 (29.19)	1705 (75.21)	493 (21.75)	33 (1.45)	36 (1.59)		
70-74	1400 (18.03)	935 (66.79)	406 (29.00)	29 (2.07)	30 (2.14)		
75-79	720 (9.27)	395 (54.86)	272 (37.78)	21 (2.92)	32 (4.44)		
80-84	597 (7.69)	210 (35.18)	287 (48.07)	35 (5.86)	65 (10.89)		
85-89	434 (5.59)	98 (22.58)	228 (52.53)	40 (9.22)	68 (15.67)		
≥90	277 (3.56)	24 (8.66)	128 (46.21)	40 (14.44)	85 (30.69)		
Sex						91.10	<.001
Male	3376 (43.47)	2403 (71.18)	782 (23.16)	75 (2.22)	116 (3.44)		
Female	4390 (56.53)	2670 (60.82)	1359 (30.96)	145 (3.30)	216 (4.92)		
Education						716.27	<.001
Illiterate/semi-illiterate	500 (6.44)	120 (24.00)	247 (49.40)	45 (9.00)	88 (17.60)		
Primary or junior high school	3540 (45.58)	2268 (64.07)	1027 (29.01)	116 (3.28)	129 (3.64)		
Senior high or technical school	1436 (18.49)	1139 (79.32)	246 (17.13)	17 (1.18)	34 (2.37)		
College degree or above	633 (8.15)	516 (81.52)	100 (15.80)	7 (1.10)	10 (1.58)		
Missing	1657 (21.34)	1030 (62.16)	521 (31.44)	35 (2.11)	71 (4.29)		
Marital status						641.38	<.001
Never married	23 (0.30)	10 (43.48)	6 (26.09)	4 (17.39)	3 (13.04)		
Currently married	5098 (65.65)	3694 (72.46)	1186 (23.26)	101 (1.98)	117 (2.30)		
Divorced or widowed	1062 (13.67)	403 (37.95)	442 (41.62)	79 (7.44)	138 (12.99)		
Missing	1583 (20.38)	966 (61.02)	507 (32.03)	36 (2.27)	74 (4.68)		
Living situation						394.89	<.001
Live alone	360 (4.64)	139 (38.61)	168 (46.67)	21 (5.83)	32 (8.89)		
Live with spouse	2342 (30.16)	1611 (68.79)	628 (26.82)	46 (1.96)	57 (2.43)		
Live with children	3745 (48.22)	2611 (69.72)	883 (23.58)	106 (2.83)	145 (3.87)		
Live with other relatives	64 (0.82)	35 (54.69)	21 (32.81)	4 (6.25)	4 (6.25)		
Live with nonrelatives	73 (0.94)	6 (8.22)	30 (41.09)	11 (15.07)	26 (35.62)		
Nursing home	16 (0.21)	4 (25.00)	4 (25.00)	4 (25.00)	4 (25.00)		
Other means and missing	1166 (15.01)	667 (57.20)	407 (34.91)	28 (2.40)	64 (5.49)		

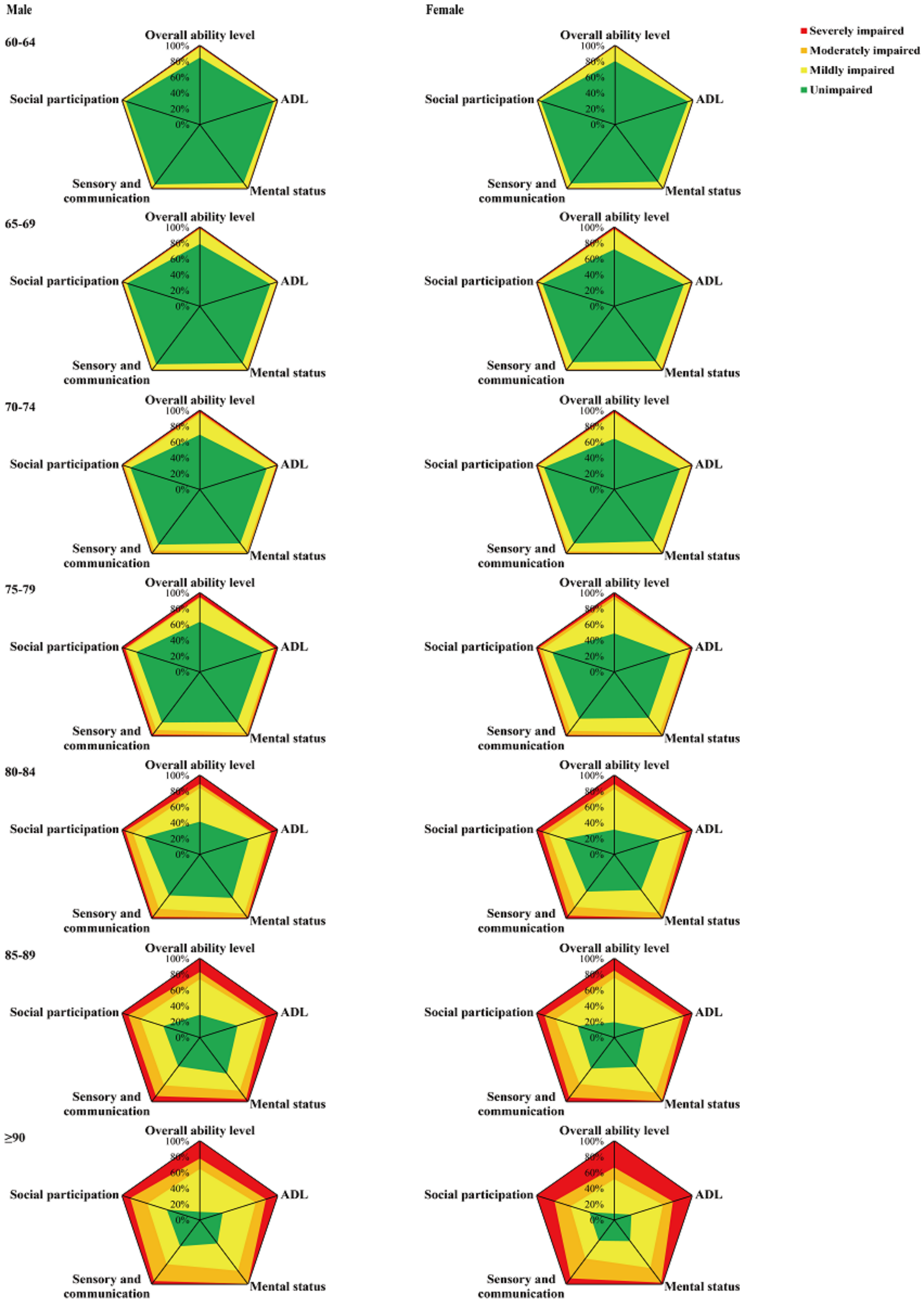
Ability Assessment

In brief, 65.32% of participants' overall ability was assessed as unimpaired, 27.57% as mildly impaired, 2.83% as moderately impaired, and 4.28% as severely impaired (Table 1). For ADL, 81.97% of participants were classified as unimpaired, whereas 13.80%, 1.47%, and 2.76% were classified as mildly, moderately, and severely impaired, respectively. For social participation, 85.63% of participants were classified as unimpaired, whereas 8.63%, 3.23%, and 2.51% were classified as mildly, moderately, and severely impaired, respectively. For sensory and communication, 81.73% of participants were classified as unimpaired, whereas 11.79%, 5.02%, and 1.46% were classified as mildly, moderately, and severely impaired, respectively. For mental status, 80.03% of participants were classified as unimpaired, whereas 15.75%, 3.78%, and 0.44% were classified as mildly, moderately, and severely impaired, respectively.

Impact of Demographic Characteristics on Overall Ability

The association of ability levels with demographic characteristics is summarized in Table 1; detailed results of pairwise comparisons are presented in Multimedia Appendix 3. Age was a key factor that influenced the overall ability of the participants (Table 1, Figure 2). With increasing age, the proportion of impaired participants gradually increased from 17.62% (365/2071) in the age group 60-64 years to 91.34% (253/277) in the age group above 90 years (*P*<.001), corresponding to an approximately 10% rise for every 5-year age increment. In particular, the respective proportion of moderately and severely impaired participants increased from 1.06% and 0.77% in the age group 60-64 years to 14.44% and 30.69% in the age group above 90 years (Table 1). Further, the proportion of overall ability impairment was significantly higher in female participants than in male participants (39.18% vs 28.82%, *P*<.001). The proportion of overall ability impairment was much higher in illiterate or semi-illiterate participants (76.00%) than in those with a college degree or above (18.48%, *P*<.001). The proportion of overall ability impairment was higher in divorced or widowed (62.05%) or never-married (56.52%) participants than in those currently married (27.54%, *P*<.001 and *P*=.02 respectively). A higher proportion of participants who lived alone (61.39%) had an ability impairment compared with those who lived with their spouse (31.21%) or children (30.28%) (all *P*<.001).

Figure 2. Radar maps showing the distribution of various impaired levels of ability dimensions stratified by sex and age groups. ADL: activities of daily living.



Ordinal Logistic Regression Analysis for the Severity of Overall Ability Impairment

The analysis results revealed that age, sex, education, marital status, and living situation were significantly associated with the severity of the overall ability impairment (Table 2). Specifically, the odds of having more severe overall ability impairment increased with age; for participants aged above 90 years, the odds of having more severe overall ability impairment was 23.92 times that of participants aged 60-64 years. The odds of having more severe overall ability impairment in females was 1.15 times that in males. Participants who were divorced or widowed were associated with a higher risk of more severe overall ability impairment than those currently married. In addition, participants living with nonrelatives had an increased risk of more severe overall ability impairment than those living alone. In contrast, higher education levels (college degree or above) emerged as a protective factor of the overall ability impairment.

Table 2. Ordinal logistic regression analysis of demographic characteristics associated with the severity of the overall ability impairment.

Variable	η coefficient (SE)	*t* value	*P* value	OR[a] (95% CI)
Age group (years)				
60-64 (reference)	N/A[b]	N/A	N/A	1.00
65-69	0.40 (0.08)	5.23	<.001	1.49 (1.28-1.73)
70-74	0.76 (0.08)	9.22	<.001	2.13 (1.82-2.51)
75-79	1.19 (0.10)	12.49	<.001	3.30 (2.74-3.98)
80-84	1.92 (0.10)	18.93	<.001	6.81 (5.59-8.31)
85-89	2.30 (0.12)	19.69	<.001	9.94 (7.91-12.50)
≥90	3.17 (0.14)	22.65	<.001	23.92 (18.18-31.50)
Sex				
Male (reference)	N/A	N/A	N/A	1.00
Female	0.14 (0.05)	2.64	.01	1.15 (1.04-1.28)
Education				
Illiterate or semi-illiterate (reference)	N/A	N/A	N/A	1.00
Primary or junior high school	–0.40 (0.10)	–3.88	<.001	0.67 (0.55-0.82)
Senior high or technical school	–0.91 (0.12)	–7.57	<.001	0.40 (0.32-0.51)
College degree or above	–1.15 (0.15)	–7.84	<.001	0.32 (0.24-0.42)
Missing	–0.83 (0.17)	–4.96	<.001	0.43 (0.31-0.60)
Marital status				
Never married	1.28 (0.44)	2.93	<.001	3.60 (1.51-8.45)
Currently married (reference)	N/A	N/A	N/A	1.00
Divorced or widowed	0.68 (0.08)	8.26	<.001	1.98 (1.68-2.33)
Missing	0.34 (0.15)	2.36	.02	1.41 (1.06-1.88)
Living situation				
Live alone (reference)	N/A	N/A	N/A	1.00
Live with spouse	0.15 (0.13)	1.17	.24	1.16 (0.91-1.49)
Live with children	–0.18 (0.12)	–1.52	.13	0.84 (0.67-1.05)
Live with other relatives	–0.09 (0.29)	–0.30	.77	0.92 (0.51-1.61)
Live with nonrelatives	0.87 (0.25)	3.46	<.001	2.38 (1.46-3.91)
Nursing home	1.20 (0.47)	2.58	.01	3.33 (1.32-8.31)
Other means and missing	0.24 (0.14)	1.77	.08	1.27 (0.98-1.66)

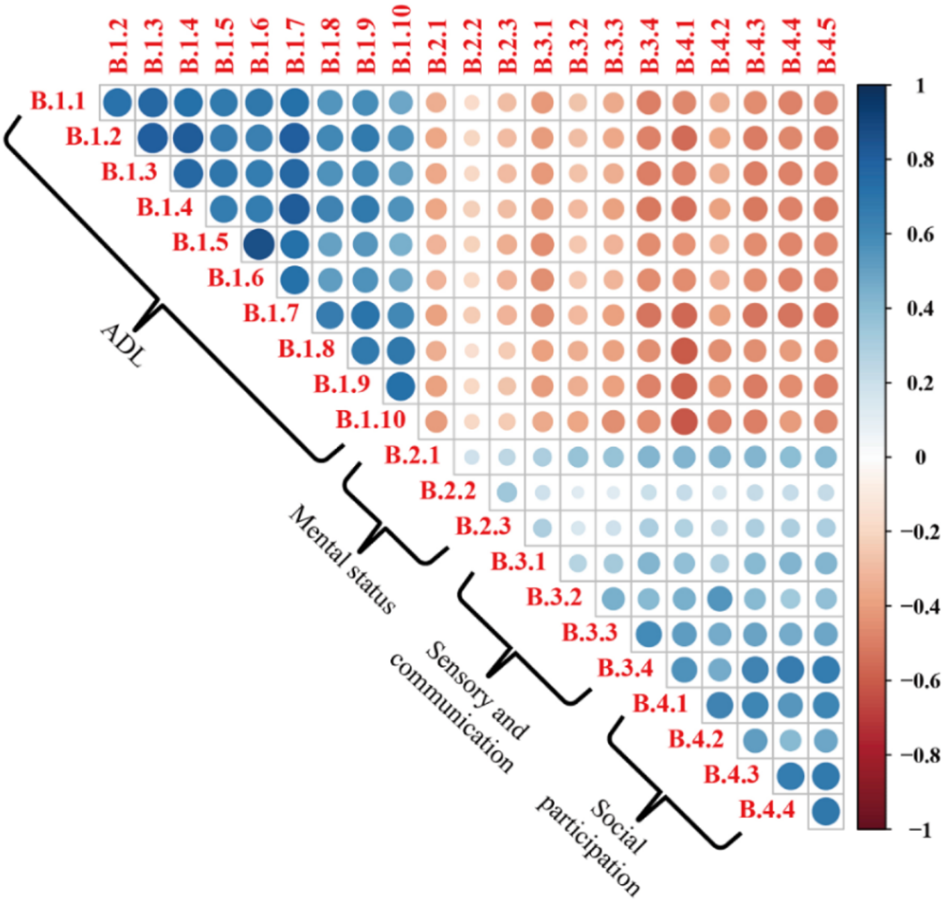
[a] OR: odds ratio.

[b] N/A: not applicable.

Correlations of Ability Assessment Scores With Various Assessment Dimensions

Figure 3 demonstrates that the four assessment dimensions were significantly correlated. Items within each dimension were positively correlated with other items in the same dimension (all *P*<.001). Overall, poor ADL was significantly correlated with poor mental status, sensory and communication, and social participation (all *P*<.001).

Figure 3. Correlation between the scores of individual items in the ability assessment. ADL: activities of daily living. See Multimedia Appendix 2 for the specific questionnaire items.



Discussion

Principal Findings

Our study assessed the health status of a large older population in the Chinese city of Shenzhen. Since 95% of Shenzhen citizens migrated from other places all over China, our study provides certain generality of the Chinese population. We found that nearly 34.7% of participants' overall ability was impaired, which is higher than the reported older adult disability rate of 28.5% (25.9%-31.2%) in a recently published meta-analysis of other relevant studies from the Chinese population [28]. The proportion of moderately and severely impaired participants (7.11%) is also slightly higher than the findings from the Sixth Population Census of China (2.95%) in 2010 [29]. The differences in population size, the definition of disability, and the time of the investigation may all contribute to the difference in this result. We found that older age, being female, having lower education levels, being divorced or widowed, and living with nonrelatives were important factors associated with ability impairment. In particular, with every 5-year increment in age, the proportion of ability impairment increased by approximately 10%. Our study found that the four dimensions of ADL, mental status, sensory and communication, and social participation were significantly correlated. Poor ability in one dimension also indicated poor abilities in the other three dimensions.

Our study demonstrated that age was a strong risk factor for overall ability impairment in the Chinese community-dwelling older population and its impact increased with age. Our finding is consistent with a previous study indicating that the disability rate in older adults doubles with every 5-year age increment [30]. With aging, physical function deteriorates gradually, and nutrient intake and assimilation may be deficient and imbalanced. The prevalence of multimorbidity increases substantially with age; the proportion of adults over the age of 65 years and 85 years with multimorbidity is more than 60% and 80%, respectively [31]. Multimorbidity such as hypertension, diabetes, and cardiovascular disease can considerably reduce ability levels during a late-life stage [32,33]. Further, the risk of neurodegenerative diseases such as Alzheimer disease and Parkinson disease also increases with age, leading to cognitive dysfunction and a reduced ability level.

Our study found a significant 15% higher risk of more severe overall ability impairment in females than in males. This is in agreement with previous findings [25,26]. A similar sex difference has also been documented in various income-developed countries, including the United States and Korea, and in developing countries such as Mexico and Indonesia [34]. The sex difference may be due to two reasons. First, females generally have a longer life expectancy than males [35]. At the same time, since males are more likely to suffer from common fatal diseases, the number of disability cases in males is lower than that for females [36]. Second, major diseases that lead to disability, such as arthritis and depression, are more common in females [37].

Our finding that a lower education level was associated with a more severe overall ability impairment risk also echoes previous studies [38,39]. Several factors may contribute to this association. First, older adults with a higher education level may have more opportunities in engaging in employment after retirement, leading to a wider social network and better social support [40]. Second, a higher education level also contributes to a better socioeconomic status, giving older adults access to wider and better health care when suffering from diseases [41]. Third, older adults with a higher education level usually have a greater awareness of self-care and self-protection, which results in a healthier lifestyle and lower disability risk.

We identified that older adults who were divorced or widowed were more likely to have severe overall ability impairment compared with those currently married as the reference condition. This finding also echoes the findings of other researchers [42,43]. Marriage is a core relationship for Chinese adults, and in this setting, remaining married in general implies greater satisfaction in the social relationship and well-being of older adults [44]. Further, Chinese couples in stable marriages are often able to accumulate more wealth to secure better health care and medical services than others [45].

Our study found that four dimensions of the ability assessment were highly correlated. Consistently, previous studies also found that ADL ability had a significant impact on cognitive function [46,47] and that more social participation helped maintain the instrumental daily living ability of older adults [48-50]. Our research indicates the need for a comprehensive approach to care for older adults, supporting their physical, mental, and social well-being. By administering the questionnaire employed in this study to assess older adults’ abilities across these dimensions, we can determine and follow their individual status and identify aspects that should be improved, allowing for tailored interventions. To this end, smart devices are becoming a viable option to support the health and independence of the older population. These devices have become increasingly prevalent with rapid advancements in photoelectronics and information technology, which can assist with daily activities, promote mental acuity, enhance sensory function, and facilitate social engagement [51].

The main contribution of our study lies in the use of a more comprehensive assessment tool to assess the ability of older adults. With the increase in aging, there has been a growing emphasis on healthy aging. Achieving healthy aging entails ensuring that older adults are healthy in multiple dimensions, including physical, mental, and social participation [52]. Therefore, as previous studies in China have assessed ability in only one dimension, they failed to provide a comprehensive reflection of the overall ability of older adults. Our study reinforces the importance of a multidimensional assessment of older adults’ ability as a means to develop comprehensive interventions that support healthy aging. Our study provides important evidence to inform the government and health providers about the health profile of community-dwelling older adults, and serves as a valuable reference for the development of services and resource allocation for the older population.

Limitations and Future Implications

Our study has several limitations. First, the data collected were mainly from one district in a southern Chinese city and hence may not be representative of other Chinese settings with different climates. Second, although the sample size of this study was large, there were relatively fewer participants in categories with certain demographic characteristics, such as the number of never-married (single) participants with moderately impaired overall ability. Third, due to missing data on other potential confounders such as disease and lifestyle factors, we were unable to include them in our study. This may affect our results; for instance, disease complications are more likely to affect the ability of older adults at a higher age and therefore the effect of age on the ability assessment may have been overestimated in our study. Fourth, our study followed a cross-sectional design; therefore, evidence for causal associations will need to be further investigated. Nevertheless, our multidimensional assessment has provided valuable evidence to inform the health status of the older population. As a part of future investigations, we plan to integrate these community-based data with hospital-based electronic medical records. This integration may form a much larger database and enable us to investigate the complex factors that affect the health of the older population in greater detail to provide additional evidence to support healthy aging initiatives.

Conclusions

In conclusion, this study found a high level of disability among the older population living in Shenzhen, China. The proportion of disability increased with age, being female, having lower education levels, being divorced or widowed, and living with nonrelatives. Impairment in ADL ability was significantly correlated with poor mental status, social participation, and sensory and communication ability. Holistic interventions targeting these factors should be implemented to prevent or delay the development of disability in the Chinese older population.

Acknowledgments

We thank all of the participants for their contribution in participating in this project and the nursing home staff for helping with data collection. LZ is supported by the National Natural Science Foundation of China (grant number 81950410639), Outstanding Young Scholars Support Program (grant number 3111500001), Xi'an Jiaotong University Basic Research and Profession Grant (grant numbers xtr022019003, xzy032020032), Epidemiology Modeling and Risk Assessment (grant number 20200344), and Xi'an Jiaotong University Young Scholar Support Grant (grant number YX6J004). MS was supported by the National Key R&D Program of China (grant number 2022YFC2505100), National Natural Science Foundation of China (grant number 12171387), China Postdoctoral Science Foundation (grant numbers 2018M631134, 2020T130095ZX), and Young Talent Support Program of Shaanxi University Association for Science and Technology (grant number 20210307).

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

JW contributed to the data analysis and drafted the manuscript. PK contributed to critical revision of the manuscript. GZ and MS contributed to the data collection and revision of the manuscript. LP and TJB contributed to the revision of the manuscript. LZ contributed to the study design, data analysis, and critical revision of the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of cross-sectional studies.
[DOC File , 31 KB - publichealth_v9i1e43612_app1.doc]

Multimedia Appendix 2
English translation of the Ability Assessment for Older Adults questionnaire.
[DOC File , 155 KB - publichealth_v9i1e43612_app2.doc]

Multimedia Appendix 3
P values for pairwise differences in the proportion of overall ability impairment between different demographic groups.
[DOC File , 53 KB - publichealth_v9i1e43612_app3.doc]

References

1. Lunenfeld B. An aging world--demographics and challenges. Gynecol Endocrinol 2008 Jan 07;24(1):1-3. [doi: 10.1080/09513590701718364] [Medline: 18224537]
2. World Population Ageing 2015: Highlights. United Nations. 2015. URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Highlights.pdf [accessed 2021-09-15]
3. Total population 2020. National Bureau of Statistics of China. URL: http://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817187.html [accessed 2021-09-15]
4. World Population Ageing 2019: Highlights. United Nations. 2019. URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf [accessed 2021-09-20]
5. Terada-Hagiwara A, Park D, Van Rijn H. Three features of aging in the People's Republic of China and implications for development agenda. ADB. 2018. URL: https://www.adb.org/publications/aging-peoples-republic-china-development [accessed 2021-09-15]
6. Gu S. A comparative study of increasing demand for health care for older people in China and the United Kingdom. World Sci Res J 2020;6(4):218-251 [FREE Full text] [doi: 10.6911/WSRJ.202004_6(4).0023]
7. Tchkonia T, Kirkland JL. Aging, cell senescence, and chronic disease: emerging therapeutic strategies. JAMA 2018 Oct 02;320(13):1319-1320. [doi: 10.1001/jama.2018.12440] [Medline: 30242336]
8. Wu J, Xiao J, Li T, Li X, Sun H, Chow EP, et al. A cross-sectional survey on the health status and the health-related quality of life of the elderly after flood disaster in Bazhong city, Sichuan, China. BMC Public Health 2015 Feb 19;15(1):163 [FREE Full text] [doi: 10.1186/s12889-015-1402-5] [Medline: 25884807]
9. Output table of the fourth national health service survey in 2008. National Health Commission of the People's Republic of China. 2010. URL: http://www.nhc.gov.cn/mohwsbwstjxxzx/s8211/201009/49143.shtml [accessed 2021-09-16]
10. Li N, Chen G, Zeng P, Pang J, Gong H, Han Y, et al. Prevalence of depression and its associated factors among Chinese elderly people: A comparison study between community-based population and hospitalized population. Psychiatry Res 2016 Sep 30;243:87-91. [doi: 10.1016/j.psychres.2016.05.030] [Medline: 27376667]
11. Rahman MHU, Srivastava S, Kumar P, Singh A, Gupta D, Kaur V. Effect of disability on high quality of life among older adults in low and middle-income countries. Ageing Int 2021 Jul 24;47(3):373-391. [doi: 10.1007/s12126-021-09425-x]
12. China. AARP International. 2019. URL: https://www.aarpinternational.org/initiatives/aging-readiness-competitiveness-arc/china [accessed 2021-09-20]
13. Chen N, Li X, Yuan N, Zhou C, Wang C. Utilization willingness of institutional care between disabled and non-disabled seniors: evidence from Jiangsu, China. BMC Health Serv Res 2019 Jun 24;19(1):410 [FREE Full text] [doi: 10.1186/s12913-019-4259-y] [Medline: 31234849]
14. Liu S, Yu P. Current status of disability in the elderly and its influencing factors. Chin J Geriatrics 2019;38(10):1079-1081.

15. Bevilacqua R, Soraci L, Stara V, Riccardi GR, Corsonello A, Pelliccioni G, et al. A systematic review of multidomain and lifestyle interventions to support the intrinsic capacity of the older population. Front Med 2022 Jul 15;9:929261 [FREE Full text] [doi: 10.3389/fmed.2022.929261] [Medline: 35911409]
16. Stoop A, Lette M, van Gils PF, Nijpels G, Baan CA, de Bruin SR. Comprehensive geriatric assessments in integrated care programs for older people living at home: A scoping review. Health Soc Care Community 2019 Sep 21;27(5):e549-e566 [FREE Full text] [doi: 10.1111/hsc.12793] [Medline: 31225946]
17. Morley JE. Rapid geriatric assessment: secondary prevention to stop age-associated disability. Clin Geriatr Med 2017 Aug;33(3):431-440. [doi: 10.1016/j.cger.2017.03.006] [Medline: 28689573]
18. Zhang Y, Xiong Y, Yu Q, Shen S, Chen L, Lei X. The activity of daily living (ADL) subgroups and health impairment among Chinese elderly: a latent profile analysis. BMC Geriatr 2021 Jan 07;21(1):30 [FREE Full text] [doi: 10.1186/s12877-020-01986-x] [Medline: 33413125]
19. Qian J, Wu K, Luo H, Cao P, Ren X. Prevalence of loss of activities of daily living and influencing factors in elderly population in China. Zhonghua Liu Xing Bing Xue Za Zhi 2016 Sep 10;37(9):1272-1276. [doi: 10.3760/cma.j.issn.0254-6450.2016.09.018] [Medline: 27655577]
20. Yang M, Ding X, Dong B. The measurement of disability in the elderly: a systematic review of self-reported questionnaires. J Am Med Dir Assoc 2014 Feb;15(2):150-159. [doi: 10.1016/j.jamda.2013.10.004] [Medline: 24314698]
21. Shang X, Peng W, Hill E, Szoeke C, He M, Zhang L. Incidence of medication-treated depression and anxiety associated with long-term cancer, cardiovascular disease, diabetes and osteoarthritis in community-dwelling women and men. EClinicalMedicine 2019 Oct;15:23-32 [FREE Full text] [doi: 10.1016/j.eclinm.2019.08.010] [Medline: 31709411]
22. Shang X, Hodge AM, Peng W, He M, Zhang L. Are leading risk factors for cancer and mental disorders multimorbidity shared by these two individual conditions in community-dwelling middle-aged adults? Cancers 2020 Jun 26;12(6):1700 [FREE Full text] [doi: 10.3390/cancers12061700] [Medline: 32604790]
23. Shang X, Peng W, Hill E, Szoeke C, He M, Zhang L. Incidence, progression, and patterns of multimorbidity in community-dwelling middle-aged men and women. Front Public Health 2020 Aug 18;8:404 [FREE Full text] [doi: 10.3389/fpubh.2020.00404] [Medline: 33014956]
24. Shang X, Peng W, Wu J, He M, Zhang L. Leading determinants for multimorbidity in middle-aged Australian men and women: a nine-year follow-up cohort study. Prev Med 2020 Dec;141:106260. [doi: 10.1016/j.ypmed.2020.106260] [Medline: 33017600]
25. Chen S, Qin J, Li Y, Wei Y, Long B, Cai J, et al. Disability and its influencing factors among the elderly in a county, Guangxi Province, China. Int J Environ Res Public Health 2018 Sep 09;15(9):1967 [FREE Full text] [doi: 10.3390/ijerph15091967] [Medline: 30205622]
26. Ma L, Li Z, Tang Z, Sun F, Diao L, Li J, et al. Prevalence and socio-demographic characteristics of disability in older adults in China: findings from China Comprehensive Geriatric Assessment Study. Arch Gerontol Geriatr 2017 Nov;73:199-203. [doi: 10.1016/j.archger.2017.07.029] [Medline: 28822922]
27. Wang S, Feng X, Liu Y, Wang H, Xie H, Wang Z. Ability assessment for older adults. Ministry of Civil Affairs of the People's Republic of China. 2013. URL: https://xxgk.mca.gov.cn:8445/gdnps/n2445/n2575/n2580/n2582/n2590/c117035/attr/272893.pdf [accessed 2023-07-23]
28. Zhenjie W, Bei L, Zhanyuan G, Hanmo Y. Meta analysis on disability of the elderly in China. Chin J Gerontol 2020;40(08):1671-1674. [doi: 10.3969/j.issn.1005-9202.2020.08.031]
29. Sixth population census data. National Bureau of Statistics of China. URL: http://www.stats.gov.cn/sj/pcsj/rkpc/6rp/indexch.htm [accessed 2023-07-23]
30. Yoshida D, Ninomiya T, Doi Y, Hata J, Fukuhara M, Ikeda F, et al. Prevalence and causes of functional disability in an elderly general population of Japanese: the Hisayama study. J Epidemiol 2012;22(3):222-229 [FREE Full text] [doi: 10.2188/jea.je20110083] [Medline: 22343328]
31. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul 07;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
32. Koller D, Schön G, Schäfer I, Glaeske G, van den Bussche H, Hansen H. Multimorbidity and long-term care dependency--a five-year follow-up. BMC Geriatr 2014 May 28;14(1):70 [FREE Full text] [doi: 10.1186/1471-2318-14-70] [Medline: 24884813]
33. Freedman VA, Carr D, Cornman JC, Lucas RE. Aging, mobility impairments and subjective wellbeing. Disabil Health J 2017 Oct;10(4):525-531 [FREE Full text] [doi: 10.1016/j.dhjo.2017.03.011] [Medline: 28385571]
34. Wheaton FV, Crimmins EM. Female disability disadvantage: a global perspective on sex differences in physical function and disability. Ageing Soc 2016 Jul;36(6):1136-1156 [FREE Full text] [doi: 10.1017/S0144686X15000227] [Medline: 27453613]
35. Modig K, Andersson T, Vaupel J, Rau R, Ahlbom A. How long do centenarians survive? Life expectancy and maximum lifespan. J Intern Med 2017 Aug 24;282(2):156-163 [FREE Full text] [doi: 10.1111/joim.12627] [Medline: 28470872]
36. Leveille SG, Resnick HE, Balfour J. Gender differences in disability: evidence and underlying reasons. Aging 2000 Apr 7;12(2):106-112. [doi: 10.1007/BF03339897] [Medline: 10902052]

37. Newman AB, Brach JS. Gender gap in longevity and disability in older persons. Epidemiol Rev 2001 Jan 01;23(2):343-350. [doi: 10.1093/oxfordjournals.epirev.a000810] [Medline: 12192741]
38. Tsai Y. Education and disability trends of older Americans, 2000-2014. J Public Health 2017 Sep 01;39(3):447-454 [FREE Full text] [doi: 10.1093/pubmed/fdw082] [Medline: 27679658]
39. Chiu C, Hayward M, Saito Y. A comparison of educational differences on physical health, mortality, and healthy life expectancy in Japan and the United States. J Aging Health 2016 Oct 01;28(7):1256-1278. [doi: 10.1177/0898264316656505] [Medline: 27590801]
40. Nurrika D, Zhang S, Tomata Y, Sugawara Y, Tanji F, Tsuji I. Education level and incident functional disability in elderly Japanese: The Ohsaki Cohort 2006 study. PLoS One 2019 Mar 12;14(3):e0213386 [FREE Full text] [doi: 10.1371/journal.pone.0213386] [Medline: 30861035]
41. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front Public Health 2020 Jun 18;8:231 [FREE Full text] [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
42. Jasser JJB, Harsha NAY, Ziq LSA, Shalash AA, Ghandour R, Giacaman R. Prevalence of disability among elderly people in the occupied Palestinian territory: a cross-sectional study. Lancet 2017 Aug;390:S15. [doi: 10.1016/S0140-6736(17)32016-0]
43. Hosseinpoor AR, Bergen N, Kostanjsek N, Kowal P, Officer A, Chatterji S. Socio-demographic patterns of disability among older adult populations of low-income and middle-income countries: results from World Health Survey. Int J Public Health 2016 Apr 4;61(3):337-345 [FREE Full text] [doi: 10.1007/s00038-015-0742-3] [Medline: 26537634]
44. Robles TF, Kiecolt-Glaser JK. The physiology of marriage: pathways to health. Physiol Behav 2003 Aug;79(3):409-416. [doi: 10.1016/s0031-9384(03)00160-4] [Medline: 12954435]
45. Wood N, McMunn A, Webb E, Stafford M. Marriage and physical capability at mid to later life in England and the USA. PLoS One 2019 Jan 23;14(1):e0209388 [FREE Full text] [doi: 10.1371/journal.pone.0209388] [Medline: 30673714]
46. Sun Q, Jiang N, Lu N, Lou VWQ. Bidirectional relationship between cognitive function and loss hierarchy of activities of daily living among older adults with disabilities in urban China: a cross-lagged analysis. BMJ Open 2022 Sep 02;12(9):e057211 [FREE Full text] [doi: 10.1136/bmjopen-2021-057211] [Medline: 36691162]
47. Sposito G, Neri AL, Yassuda MS. Advanced Activities of Daily Living (AADLs) and cognitive performance in community-dwelling elderly persons: Data from the FIBRA Study - UNICAMP. Rev Bras Geriatr Gerontol 2016 Feb;19(1):7-20. [doi: 10.1590/1809-9823.2016.15044]
48. Tomioka K, Kurumatani N, Hosoi H. Social participation and cognitive decline among community-dwelling older adults: a community-based longitudinal study. J Gerontol B Psychol Sci Soc Sci 2018 Jun 14;73(5):799-806. [doi: 10.1093/geronb/gbw059] [Medline: 27194753]
49. Tomioka K, Kurumatani N, Hosoi H. Age and gender differences in the association between social participation and instrumental activities of daily living among community-dwelling elderly. BMC Geriatr 2017 Apr 28;17(1):99 [FREE Full text] [doi: 10.1186/s12877-017-0491-7] [Medline: 28454521]
50. Tomioka K, Kurumatani N, Hosoi H. Association between social participation and 3-year change in instrumental activities of daily living in community-dwelling elderly adults. J Am Geriatr Soc 2017 Jan 27;65(1):107-113. [doi: 10.1111/jgs.14447] [Medline: 27673582]
51. Papetti A, Iualé M, Ceccacci S, Bevilacqua R, Germani M, Mengoni M. Smart objects: an evaluation of the present state based on user needs. 2014 Presented at: Second International Conference of Distributed, Ambient, and Pervasive Interactions; Jun 22-27, 2014; Heraklion, Greece. [doi: 10.1007/978-3-319-07788-8_34]
52. Chan ASW, Ho JMC, Tam HL, Tang PMK. Book Review: Successful aging: a neuroscientist explores the power and potential of our lives. Front Psychol 2021 Jun 18;12:705368. [doi: 10.3389/fpsyg.2021.705368]

Abbreviations

ADL: activities of daily living
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 18.10.22; peer-reviewed by X Zou, X Ma, Y Lin, R Bevilacqua, ASW Chan ; comments to author 06.02.23; revised version received 09.03.23; accepted 21.06.23; published 10.08.23.

Please cite as:
Wang J, Kwan P, Zhang G, Shen M, Piccenna L, O'Brien TJ, Zhang L
A Multidimensional Assessment of Activities of Daily Living, Mental Status, Communication, and Social Abilities Among Older Adults in Shenzhen, China: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e43612
URL: https://publichealth.jmir.org/2023/1/e43612
doi:10.2196/43612
PMID:37561566

Original Paper

Quantifying Adverse Childhood Experiences in Oklahoma With the Oklahoma Adversity Surveillance Index System (OASIS): Development and Cross-Sectional Study

Jason Walter Beaman[1], DO, MS, MPH; Cherie Josephine Miner[2], BA; Cadence Bolinger[3], BA, MSc

[1]School of Forensic Sciences, Center for Health Sciences, Oklahoma State University, Tulsa, OK, United States
[2]Department of Psychiatry and Behavioral Sciences, Center for Health Sciences, Oklahoma State University, Tulsa, OK, United States
[3]School of Community Health Sciences, Counseling and Counseling Psychology, Oklahoma State University, Stillwater, OK, United States

Corresponding Author:
Jason Walter Beaman, DO, MS, MPH
School of Forensic Sciences
Center for Health Sciences
Oklahoma State University
1111 W. 17th St.
Tulsa, OK, 74107
United States
Phone: 1 918 561 1108
Email: jason.beaman@okstate.edu

Abstract

Background: Developmental trauma depending on several factors may lead to later adult health risks and is an increasing public health concern, especially in states with predominantly rural populations. Oklahoma remains one of the states in America with the highest count of adverse childhood experiences (ACEs); therefore, more refined research methods for quantifying ACEs are vital for ensuring proper statewide interventions.

Objective: While data sets already exist at the state level measuring specific ACEs like divorce or child abuse, the state currently lacks a single source for specific ACEs that can incorporate regions to allow for the identification of counties where ACEs are especially high. This county identification will allow for assessing trends in adversity prevalence over time to indicate where targeted interventions should be done and which counties experience amplified long-term consequences of high ACE rates. Thus, the model for the Oklahoma Adversity Surveillance Index System (OASIS) was born—a public health tool to map ACEs at the county level and grade them by severity over time.

Methods: County-level data for 6 ACEs (mental illness, divorce, neglect, child abuse, domestic violence, and substance use) were collected from the Oklahoma Department of Human Services, Oklahoma State Department of Health, and Oklahoma Community Mental Health Centers for the years 2010 to 2018. First, a potential ACEs score (PAS) was created by standardizing and summing county rates for each ACE. To examine the temporal change in the PAS, a bivariate regression analysis was conducted. Additionally, an ACEs severity index (ASI) was created as a standardized measure of ACE severity across time. This included scoring counties based on severity for each ACE individually and summing the scores to generate an overall ASI for each county, capturing the severity of all ACEs included in the analysis.

Results: Mental illness and substance use showed the highest rates at the state level. Results from the regression were significant ($F_{1,76}$=5.269; P=.02), showing that county PAS showed an increase over years. The ASI scores ranged from 0 to 6, and 4 Oklahoma counties (Adair, McCurtain, Muskogee, and Pittsburg) received a score of 6.

Conclusions: OASIS involves the identification of counties where ACEs are most prevalent, allowing for the prioritization of interventions in these “hot spot” counties. In addition, regression analysis showed that ACEs increased in Oklahoma from 2010 to 2018. Future efforts should center on adding additional ACEs to the ASI and correlating adverse outcome rates (such as violence and medical disorder prevalence) at the county level with high ASI scores.

(JMIR Public Health Surveill 2023;9:e45891) doi:10.2196/45891

KEYWORDS

statewide intervention; adverse childhood experiences; public health; surveillance system; trauma; rural; mental illness; childhood trauma; surveillance; developmental trauma; adult

Introduction

Background

The Adverse Childhood Experiences Study, published in 1998, is a sentinel publication in modern health care [1]. This study demonstrated several important factors. First, childhood adversity is associated with poor outcomes as an adult, including cardiovascular disease, diabetes, obesity, liver and digestive diseases, and cancer [2-4]. Second, this correlation is dose-dependent—the more adversity, the higher the likelihood of poor outcomes [1]. Third, adversity in childhood extends beyond personal factors such as abuse and neglect and includes household dysfunction [1]. Since this study was first published, adverse childhood experiences (ACEs) have become common terminology in medicine [4,5]. With the strong correlation with adverse health outcomes, measuring the burden of ACEs can be important for public health and prevention. Secondary to the burden of disease and dysfunction, it is important to quantify the effect of ACEs on a community. Understanding this burden is essential for the deployment of limited resources and the implementation of evidence-based primary, secondary, and tertiary prevention programs.

A definitive list of ACEs does not exist. Those listed in the original Adverse Childhood Experiences Study questionnaire included physical, emotional, and sexual abuse; physical and emotional neglect; violence toward one's mother; mental illness; substance abuse; and incarceration of a household member [1]. Other ACEs that should be considered according to the Behavioral Risk Factor Surveillance System (BRFSS) and Child Trends include educational and financial opportunities, chronic and infectious diseases, maternal health, risky behavior, injury, violence outside of the home, neighborhood safety, homelessness, bullying, racial and ethnic discrimination, and parent or guardian death [6,7].

Screening for childhood adversity and trauma is gaining acceptance and is recommended by the American Academy of Pediatrics [8]. It is unclear whether or not assessing ACEs at the patient level can alter health outcomes [9,10]. However, it has been demonstrated that patients are not averse to discussing the topic [11]. The more uptake there is in the use of patient instruments such as the ACEs questionnaire, the more useful this method can be in public health surveillance.

Assessing ACEs Through Instruments

There are many different instruments available for assessing the number of ACEs an individual experiences. Most of these instruments will score the presence of an ACE in a dichotomous fashion, such as yes or no for the exposure. This method does not account for the intensity or duration of a particular traumatic experience. Further, this method does not allow for the experience to be evaluated in context. For example, a patient who is the child of a divorce may become separated from an abusive caregiver. Such circumstances underline the problems in attempting to understand and contextualize social dysfunction and the need for further development in methods for measuring and analyzing ACEs.

Perhaps the most common method of assessing ACEs is what was used in the Adverse Childhood Experiences Study, which was a retrospective survey [1]. This method can be used for public health programming and also for basic research. In the public health field, an ACE module was added to collect information on child abuse and neglect and household challenges through the BRFSS [12]. The ACE module was included at least once in the BRFSS of every state by the end of 2020 [13]. The BRFSS ACE module asks an individual to reflect on their childhood and answer a series of questions surrounding abuse, violence, drug use, and mental illness in the home which are adapted from the original Adverse Childhood Experiences Study [14]. Since 2009, the BRFSS has been collecting data on ACEs, and the module was updated in 2019 to include questions about neglect. Most recently, it was updated in 2021 with questions about childhood support and caregiver involvement.

Another important survey is the National Survey of Children's Health (NSCH). This is an annual survey conducted in the United States that began in 2003 and has been reported as recently as 2021 [7,15]. One child under the age of 18 in each household is selected to be the focus of the survey. Specific ACE questions in the survey include parent or guardian divorce, separation, death, or incarceration; mental illness; suicide; illicit drug use; alcohol use; depression of a household member; domestic violence; neighborhood violence; income; and housing or food security [7]. In 2016, data collected by the NSCH indicated that 45% of children experienced at least one ACE, and 10% experienced 3 or more ACEs [7,16]. The most common ACEs reported were parental separation and low socioeconomic status [7].

Another method used to measure ACEs is by directly asking an individual as they intersect with the health care system [4,17]. This includes having the patient complete a screening instrument, such as the Adverse Childhood Experiences Questionnaire (ACE-Q), which can then be analyzed at a larger clinical level [11,18,19]. However, patient questionnaires also have shortcomings. First, they measure individuals presenting to a health care setting [18-20]. This is important because of well-known disparities in individuals accessing health care, including lack of a provider (such as in rural areas) or lack of ability to pay [21]. And while there is research indicating some patients feel comfortable discussing childhood trauma or ACEs with a health care provider [22], there are also patients with concerns about medical record storage in general, as issues with privacy arise in the increasingly digital world [23].

Knowledge of ACEs in a community could be important for primary, secondary, and tertiary prevention planning. Improved knowledge of prevalence rates could lead to larger gains in positive outcomes and more efficient use of precious resources. This is especially important in high-burden, low-resource areas, such as Oklahoma [24-27]. Oklahoma has consistently been recognized as leading the nation in childhood adversity, with 28.5% of Oklahoma children experiencing 2 or more ACEs [28]. Oklahoma also experiences poor social determinants of health (SDOH), such as health care shortages, violence, and food insecurity, among others [29]. Further, there is a lack of access to care due to a lack of primary care providers outside of metropolitan areas in rural communities, which make up the majority of Oklahoma [24-27]. These upstream deficiencies lead to high levels of adverse health outcomes such as diabetes, obesity, substance use disorders, and overdose mortality [24,30].

While it is known that Oklahomans have high rates of childhood adversity, there is no condensed source of data to which stakeholders can turn to receive aggregated, targeted information on ACEs at a granular (county) level. While there are great data sets like the BRFSS and NSCH, these lack county-level statistics for all of Oklahoma [31]. Currently, data is spread out across various state agencies, such as the Oklahoma State Courts Network (OSCN), the Oklahoma Department of Health Services (OKDHS), and Oklahoma community mental health centers (CMHCs). With multiple agencies doing different parts of the work in silos, Oklahoma does not have a coordinated trauma-response system in place to capture and publish all ACE data in one location.

To address this lack of a coordinated system and to combat this epidemic of childhood adversity and subsequent negative health consequences, the Oklahoma Adversity Surveillance Index System (OASIS) was created. OASIS was initiated in 2021 through a collaboration between the Oklahoma State University Center for Health Sciences and multiple state agencies. Data-sharing agreements were used to obtain data that were not publicly available. This system relies on several concepts. First, the adversity burden shared by a population should have a footprint in interaction with public agencies that collect or could collect data. Second, while there may be inaccuracies in data collection, these inaccuracies should be equally inaccurate over space and time. This allows for hot spot determination and comparisons among similar geographical entities (ie, county, zip code) and time (ie, years). Third, the data should be regularly collected and analyzed to monitor emerging areas and trends. Identifying trends can allow for the examination of intercounty differences and temporal variations between geographical entities. These trends can indicate where further analyses should be done to determine the factors that may contribute to the observed variations in ACE burden across counties and years. Such factors could include socioeconomic conditions, access to health care and support services, community resources, education levels, and public health initiatives. Finally, all efforts should be made to review and refine the system so that it becomes more accurate and insightful over time.

With the implementation of OASIS, the state of Oklahoma now has access to a public health tool that encapsulates county-level statewide data on ACEs that were previously unavailable or difficult to obtain. Additionally, with OASIS in place, the state will be able to carry out targeted interventions in the counties that need it the most. Without this surveillance system, high-burden counties would have been previously unidentifiable, and it would not have been possible to specifically target them with additional resources and prevention strategies. We hope that other states will replicate our methodology and identify their own hot spots to ultimately better serve areas that need more ACE-related support, resources, and services.

Methods

Overview

Data were obtained from various state agency sources for the years 2010 to 2018. Counts of abuse and neglect were obtained from the OKDHS. The OKDHS reports these data as a combination of physical, sexual, and psychological abuse, listed as a single "abuse" category. Ideally, these data would be separated, but currently, a county-level data set with these distinctions is unavailable. Neglect is listed separately. Data were available at the zip code and county level. OKDHS receives data from reports substantiated by child protective services (CPS). Data for divorce were obtained from the Oklahoma State Department of Health (OSDH). The OSDH collects divorce data for all counties in Oklahoma from county clerks. Unfortunately, Oklahoma does not inquire about whether or not children are involved when an individual files for divorce.

Data for parental mental illness and substance use were obtained at the zip code level for patients treated at Oklahoma CMHCs. CMHCs aggregate their data from appointments and patient files. This data are then submitted to the Oklahoma Department of Mental Health and Substance Abuse Services (ODMHSAS). Rates of mental illness and substance use were calculated at the population level. For divorce, mental illness, and substance use, information is not obtained as to whether an individual has children, so only raw numbers were used. Domestic violence rates were calculated through the number of charges filed as listed in the OSCN. Through the use of a publicly available application programming interface, records were searched at the county level. Again, this information may not be specific to parents. Data from the US Census Bureau were used to calculate population data for all counties. Counties were coded as metro or nonmetro following US Department of Agriculture Economic Research Service Rural Urban Continuum Codes.

Potential ACE Score

To calculate additive totals, county crude rates for each ACE were standardized to 100 (average county crude rates per year are reported in the Results section). The available ACEs were added to create a potential ACE score (PAS). The PAS ranges from 0 to 600, representing the potential for children in each county to experience ACEs. Rates of determinants like divorce and substance use at the county level were used as proxies for the potential ACEs experienced by the children in that county.

Statistical Analysis

Data were uploaded into SPSS (version 28.0.1.1; IBM Corp) for analysis. To examine the temporal changes in ACEs, a bivariate regression analysis was conducted. The PAS was used as the dependent variable, and the year was treated as the independent variable. This analysis allowed for the assessment of how the PAS changed over time. Results were visualized using Tableau (version 2021.4.5; Tableau Software, LLC).

ACE Severity Index

In addition to the PAS, we developed the ACE severity index (ASI) as a standardized measure to assess the burden and severity of ACEs over time. The ASI aimed to identify counties more heavily impacted by ACEs, considering both the depth and duration of ACEs. To determine the impact of ACEs on specific counties, we scored each county based on ACE severity. Severity was measured for each ACE individually, assessing the extent of its occurrence over time. Counties were classified as “severe” for a particular ACE if its rate was above the state average. Severity scores ranged from 0 (indicating the county had an average or below-average rate for that ACE in all years) to 9 (indicating the county had an above-average rate for that ACE in all years). We calculated severity scores for each county for all 6 ACEs included in the analysis. Counties with severity scores above the average were given a point for that ACE, contributing toward their total ASI. This process was repeated for all 6 ACEs analyzed. The outcome was an ASI for each county, encompassing the severity of all ACEs included in the analysis and ranging from 0 to 6.

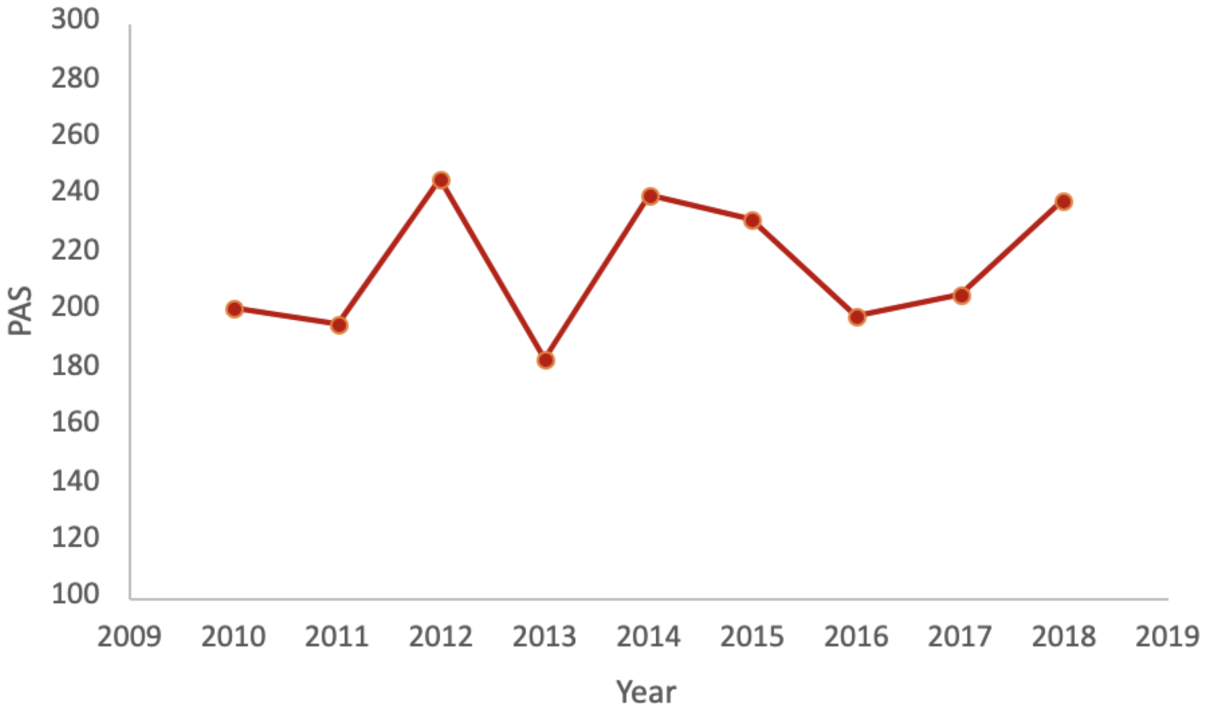
Ethical Considerations

Data-sharing agreements were obtained before the project start. The project was approved as non–human subjects research by the institutional review board at Oklahoma State University Center for Health Sciences (#2021009). The requirement for informed consent was waived due to the retrospective nature and minimal risk of this study.

Results

Over the study period (2010-2018), the average PAS was 215.08, with an overall increasing trend from 2010 to 2018 (Figure 1). An analysis of the data set revealed notable trends in several Oklahoma counties regarding their PAS from 2010 to 2018. Specifically, Choctaw and Greer counties showed an increasing trend over the study period, while Cimarron and Harper counties consistently remained below the average PAS. Adair, Woodward, Coal, Muskogee, Pittsburg, and Brian counties consistently maintained a PAS above the average.

Figure 1. Yearly Oklahoma average PAS. PAS: potential adverse childhood event score.



Choctaw County exhibited a consistent upward trend in PAS over the years, with a score of 274.637 in 2010 and 322.841 in 2011, reaching a peak of 439.225 in 2012, and still maintaining a score in the top 10 of counties in 2018 (335.35). Greer County also experienced an increasing trend, starting at 171.454 in 2010 and steadily rising to 302.128 in 2018. On the other hand, Cimarron County consistently remained below the average PAS throughout the study period. The county had a relatively low PAS of 83.137 in 2010 and showed slight fluctuations over the years, but remained consistently below average, reaching 108.26 in 2018. Similarly, Harper County started with a PAS of 163.285 in 2010 and remained consistently below average, with a score of 114.229 in 2018.

In contrast, Adair County consistently displayed remarkable PAS well above the average throughout the study period. Starting with a high PAS of 408.354 in 2010, Adair County maintained consistently elevated scores, reaching its peak of 433.736 in 2018. Similarly, Woodward County demonstrated a pattern of consistently scoring above the average. Beginning at 295.013 in 2010, Woodward County steadily increased its PAS over the years, culminating in a peak of 369.031 in 2018. Joining them in consistently exhibiting higher-than-average PAS were several other notable counties. Coal County, despite experiencing some fluctuations, maintained relatively high scores throughout the study period, with its highest PAS of 404.271 recorded in 2010 and remaining significantly above average at 309.158 in 2018. Muskogee County followed a similar trajectory, with a starting PAS of 277.018 in 2010 and a peak of 322.252 in 2018. Pittsburg County consistently scored above the average as well, reaching a peak PAS of 319.14 in 2018. Notably, Bryan County had a consistent above-average pattern, starting at 281.216 in 2010 and reaching its peak of 242.676 in 2018. These counties, including Adair and Woodward, stood out for their remarkable and consistent above-average PAS throughout the study period. Notably, the top 10 counties in 2018 had higher PAS values compared to the top ten counties in 2010 (Table 1).

Table 1. Top 10 Oklahoma counties for PAS[a] in 2010 and 2018.

Ranking	2010		2018	
	County	PAS	County	PAS
1	Adair	408.354	Adair	433.74
2	Coal	404.271	Woodward	369.03
3	Beckham	331.141	Kay	347.2
4	Marshall	307.987	Seminole	345.48
5	Seminole	304.522	Choctaw	335.35
6	Woodward	295.013	Muskogee	322.25
7	Okfuskee	283.525	Pittsburg	319.14
8	Bryan	281.216	Ottawa	319.08
9	Muskogee	277.018	Blaine	316.2
10	Pittsburg	275.589	Latimer	312.23

[a] PAS: potential adverse childhood event score.

Our analysis revealed a significant relationship between year and PAS ($F_{1,76}$=5.269; P=.02) (Table 2). For each 1 unit increase in the year variable, there was a corresponding 2.193 increase in the PAS (SE 0.955, $t_{1,76}$=2.295; P=.02). The model summary indicates that the regression model explains a small proportion of the variance in the PAS.

Table 2. Model summary of regression analysis assessing the relationship between year and potential adverse childhood event score. The dependent variable was potential adverse childhood event score.

Model	R	R^2	Adjusted R^2	Estimate SE	Change in statistics		
					Change in R^2	Change in F (df)	P value
1	0.087[a]	0.008	0.006	64.94422332	0.008	5.269 (1,691)	.02

[a] Predictors were (constant), year.

When visualized, the difference between county PAS in 2010 and 2018 shows an overall increase (Figures 2 and 3).

Figure 2. Oklahoma PAS county distribution in 2010. PAS: potential adverse childhood event score.

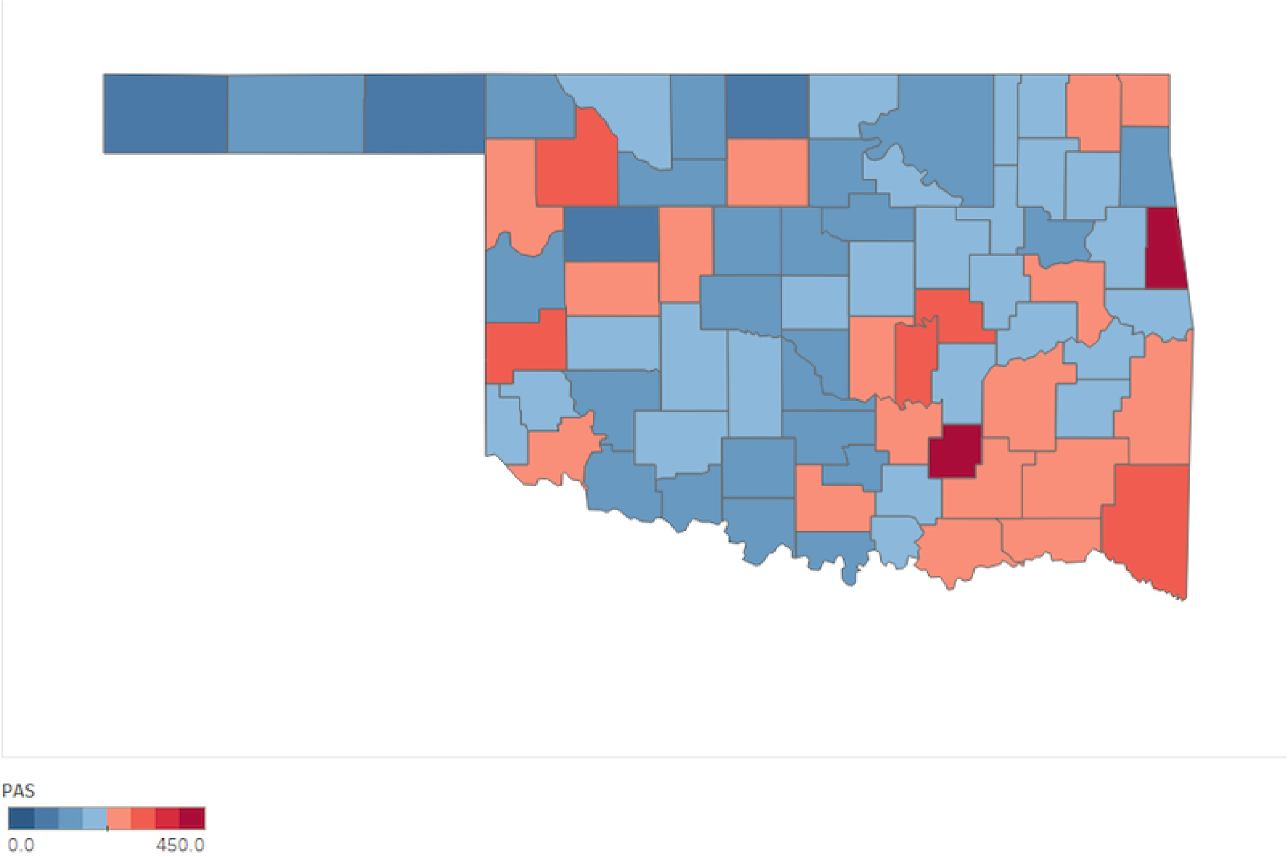
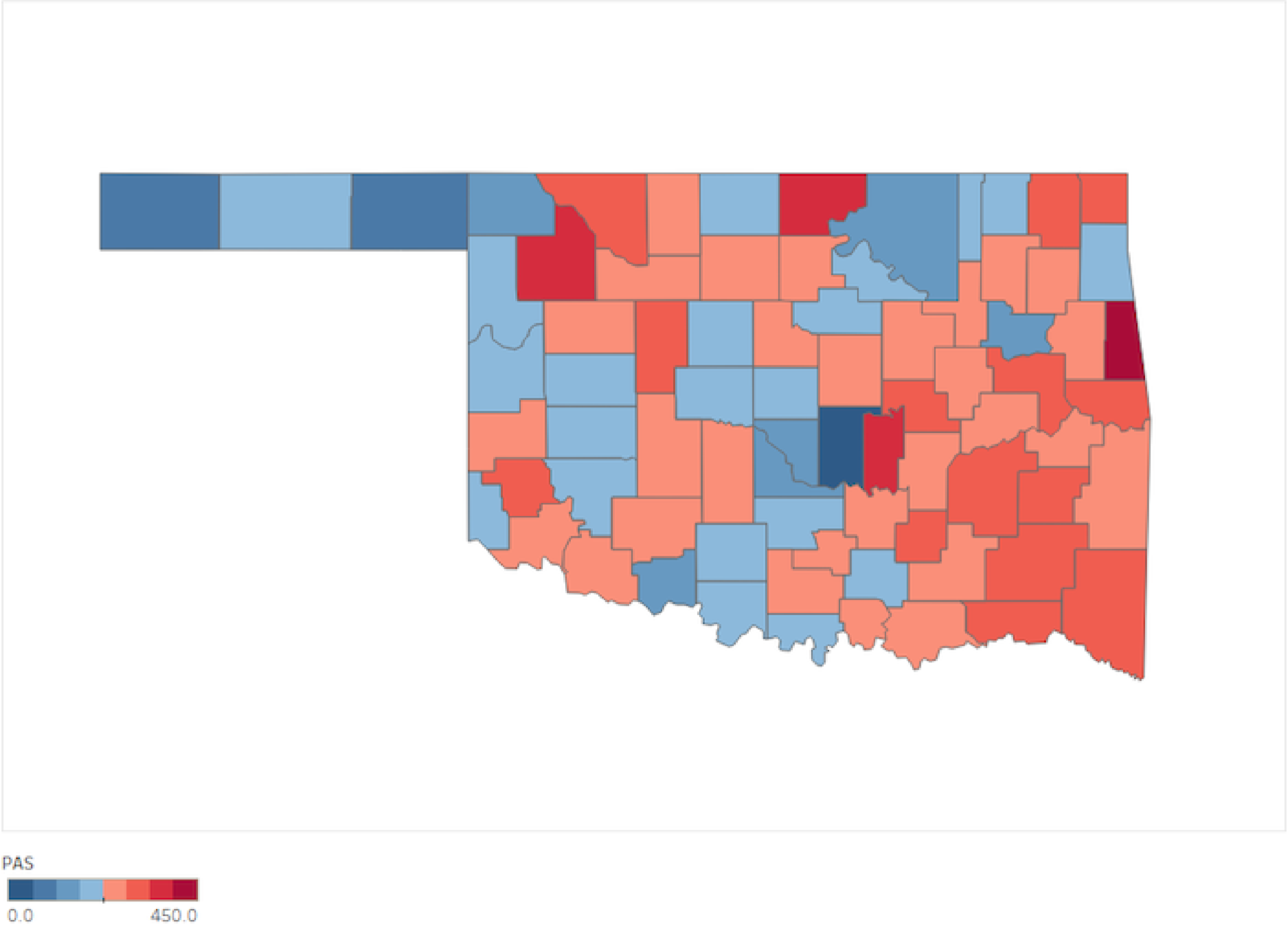
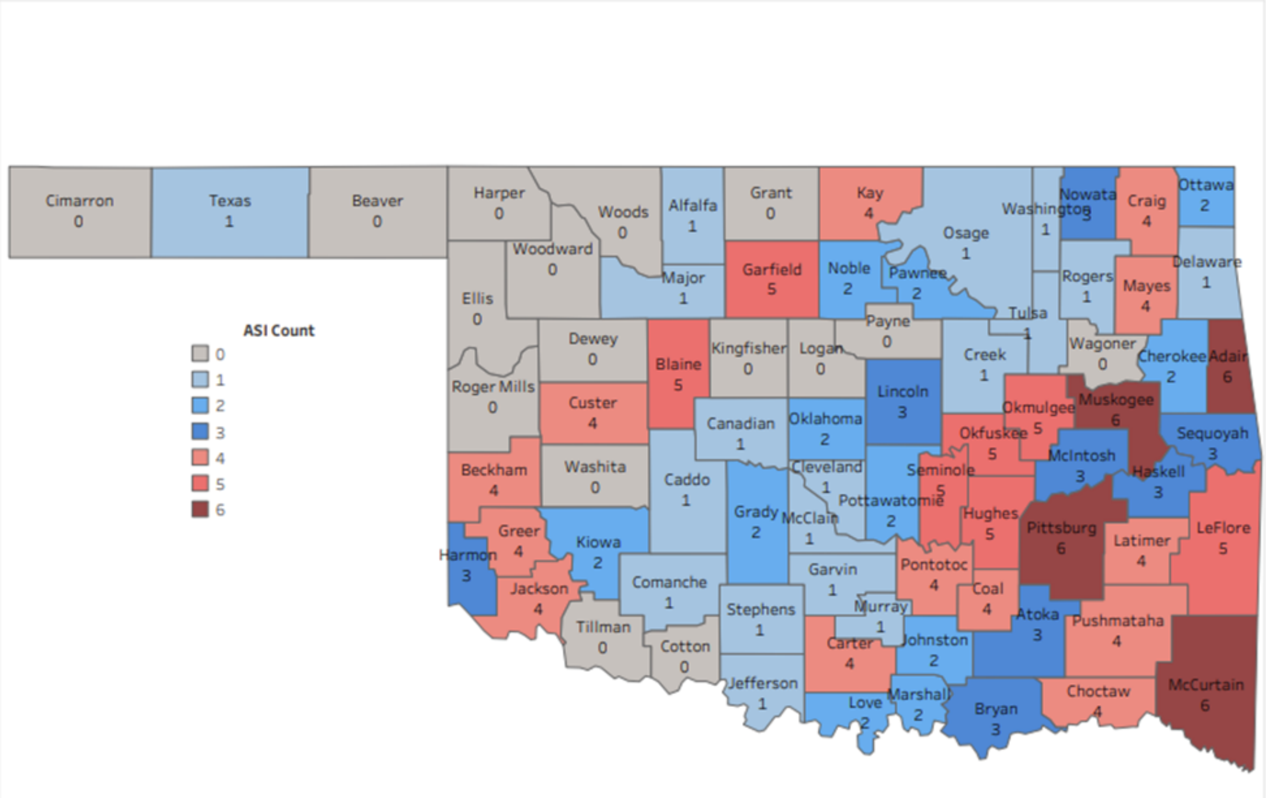


Figure 3. Oklahoma PAS county distribution in 2018. PAS: potential adverse childhood event score.



The results for the ASI showed that 4 nonmetro Oklahoma counties—Adair, McCurtain, Muskogee, and Pittsburg—had scores of 6 (Figure 4). Garfield, Blaine, Seminole, Okfuskee, Okmulgee, Hughes, and LeFlore counties received scores of 5.

Figure 4. Map of Oklahoma Adversity Surveillance Index System (OASIS) showing the ASI for each Oklahoma county. Zero represents the lowest burden and 6 the highest burden encompassing all 6 adverse childhood events included in our analysis. ASI: adverse childhood event severity index.



Discussion

Principal Findings

In this study, we aimed to examine the prevalence of ACEs and their variations across counties in Oklahoma using OASIS. Our findings provide valuable insights into the burden of ACEs in the state and highlight the need for targeted interventions and resource allocation. The analysis of the PAS revealed an increasing trend in ACE burden from 2010 to 2018, and our regression analysis demonstrated a significant relationship between the year and PAS, indicating a continuous rise in ACE burden over time.

The ASI allowed us to identify counties more heavily impacted by ACEs. After indexing counties on the ASI, 4 counties stood out, with scores of 6 of 6. All 4 counties are rural, nonmetro, located in Indian Country, are significantly below the national median income, have a greater percentage of persons living in poverty, and have higher unemployment rates than the national average [32]. All of these poor SDOH are known upstream risk factors for ACEs [7]. Much more in-depth analysis of these counties, along with others with high ASI scores, should be done to ascertain how their poor SDOH are driving childhood adversity.

The average county in Oklahoma exhibited rates of ACEs that were consistently above the state average for 4 or more years. This increasing trend is concerning and shows the need for further expansion of OASIS with the addition of other ACEs, for example, incarceration of a family member; Oklahoma consistently ranks in the top 10 for incarceration rate in the United States over the past decade [33]. Our analysis of the PAS showed that there was a significant relationship between year and PAS ($F_{1,691}$=5.269; P=.02). For every 1 unit increase by year, there was a 2.193 (SE 0.955) unit increase in PAS ($t_{1,76}$=2.295; P=.02).

OASIS has several strengths and limitations. The strengths include the creation of a framework that can be built upon and improved over time. With OASIS, we have designed and implemented a framework for data sharing and analysis across state data sets, an accomplishment that is not necessarily easy to do given the sensitivity around information on ACE variables. Having created a proof of concept, legislators and agency heads can understand how the way they collect data intersects with other agencies and allows judicious spending of resources. Another strength is the ability to compare geographical entities together. Comparing zip codes in large urban areas could be valuable to understanding population disparities, as survey studies of ACE prevalence have known weaknesses [34]. To generalize the results of a survey to the general population, it should be administered to a sufficiently large and randomized sample of the general population. This can be difficult to achieve. For example, the population mixture of the original Adverse Childhood Events Study was primarily White, health-insured individuals [35]. The BRFSS has similar concerns. Analysis has shown that individuals answering the ACEs module are more likely to have higher education than the US average [6,35]. These differences are notable in health care because it is well documented that there are significant health disparities among races, insured and uninsured people, and different levels of education [36-40]. Thus, with OASIS, the state of Oklahoma now has a way to easily access county-level ACE-related data that are generalizable to the whole state. This public health tool allows stakeholders to better identify areas that need more ACE-related support, resources, and services. Understanding why is the first step toward prevention.

Another strength is the ability to identify hot spots by depth and time. The PAS allows for counties to be compared across time to see if their rates of childhood adversity are increasing or decreasing. In addition, the PAS allows for counties to be compared to each other and trends identified between counties. We found that some counties, such as Adair, ranked consistently higher on the PAS compared to other counties. On the other hand, some counties, such as Harper and Cimarron, consistently ranked below the average PAS. These counties showing lower rates over time should be studied further; they can serve as models for successful prevention strategies that could be replicated in other areas.

OASIS is not without limitations, however. First is the lack of data on all known ACEs. For example, at one time, Oklahoma was the second most prevalent place in the world to incarcerate women and incarcerated more men and women combined than any other state in the United States [41,42]. Yet the Oklahoma Department of Corrections does not systematically collect data on inmates and their children. This prevents the understanding of this trauma as it relates to other ACEs and also prevents understanding the damage mass female incarceration can have on a community. One additional limitation of the surveillance system is the absence of validity testing, which means that the extent of error or inaccuracy within the system is unknown. Consequently, there is a risk of incorrectly identifying hot spots. However, it should be noted that since all data were collected using the same method, any inaccuracies should be consistent across different areas. Nonetheless, it is crucial for future efforts to develop methods that can test the validity of the system and address its identified weaknesses. By doing so, improvements can be made to enhance accuracy and ensure more reliable identification of hot spots. Another weakness lies in the extrapolation of population data without accounting for the number of children involved. The collected data for divorce, for example, are based on raw divorce numbers, without considering divorces that specifically affect children, as that information is not available from ODMHSAS court records. Consequently, there is a possibility of overcounting divorces involving children. However, it is important to note that this overcounting issue is consistent across all counties, as the same calculation method is used. Therefore, while the accuracy may be compromised, it is uniformly inaccurate across all counties.

A final limitation are the data themselves. OASIS relies on individuals accessing services and does not include those that do not. For example, if a mother accessed mental health treatment outside of a CMHC, that would not be measured. This is less of a problem in rural areas that have limited providers. However, in the urban areas of Oklahoma City and Tulsa, this likely leads to undercounting. This limitation is highlighted again in divorce. Individuals do not have to be married to raise children together. Therefore, when unmarried parents separate, that would not be reflected in OSDH divorce data, and this would lead to an undercounting of events.

Conclusions

The findings of our study have significant implications for the field of public health and can shape future interventions and strategies. The identification of counties with a high prevalence of ACEs through OASIS allows for targeted interventions and prioritization of resources in these hot spot areas. Tailored interventions at the county level are crucial, considering the variations in ACE burden across different regions.

Our study highlights the value of the ASI in revealing geographic disparities in the severity and persistence of ACEs. Four counties in particular stood out, with an ASI score of 6, indicating an elevated prevalence of ACEs and consistently above-average rates across all ACE types examined. This signifies a pressing need for urgent and coordinated action from policymakers, health practitioners, social care providers, and researchers in addressing this public health issue. Counties with high ASI scores are likely to experience amplified long-term consequences associated with ACEs, including increased risks for chronic diseases, reduced educational achievement, and a higher likelihood of involvement with the criminal justice system [43,44]. It is imperative to respond promptly with both ACE prevention and effective management strategies to mitigate these outcomes. The ASI also serves as a valuable guide for targeted resource allocation. By identifying counties burdened by ACEs, resources can be efficiently prioritized, focusing on services such as trauma-informed care, community health initiatives, and educational interventions tailored to address ACEs. Future research should explore community-level factors (eg, poverty, community violence, service accessibility) as well as individual and family-level dynamics to design effective prevention and intervention strategies.

While acknowledging the limitations of this study, OASIS provides a foundation that can be refined and improved over time. This surveillance system facilitates collaboration and coordination across sectors, states, and agencies, enabling a comprehensive response to high-need areas. Future efforts should focus on expanding the range of ACEs, improving data collection, adding adverse outcome rates, and conducting validity testing. The ultimate objective of OASIS should be to develop effective strategies that not only address the present situation in counties with high childhood adversity indicated by the ASI but also prevent future generations from experiencing similar levels of adversity.

Acknowledgments

Marin White is acknowledged for their efforts in data retrieval and cleaning.

Data Availability

The data sets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Felitti VJ, Anda RF, Nordenberg D, Williamson DF, Spitz AM, Edwards V, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. Am J Prev Med 1998 May;14(4):245-258. [doi: 10.1016/s0749-3797(98)00017-8] [Medline: 9635069]
2. Vig KD, Paluszek MM, Asmundson GJG. ACEs and physical health outcomes. In: Adverse Childhood Experiences: Using Evidence to Advance Research, Practice, Policy, and Prevention. Cambridge, MA: Academic Press; 2020:71-90.
3. Bethell CD, Carle A, Hudziak J, Gombojav N, Powers K, Wade R, et al. Methods to assess adverse childhood experiences of children and families: Toward approaches to promote child well-being in policy and practice. Acad Pediatr 2017;17(7S):S51-S69 [FREE Full text] [doi: 10.1016/j.acap.2017.04.161] [Medline: 28865661]
4. McLennan JD, MacMillan HL, Afifi TO. Questioning the use of adverse childhood experiences (ACEs) questionnaires. Child Abuse Negl 2020 Mar;101:104331 [FREE Full text] [doi: 10.1016/j.chiabu.2019.104331] [Medline: 31887655]
5. Onigu-Otite E, Idicula S. Introducing ACEs (adverse childhood experiences) and resilience to first-year medical students. MedEdPORTAL 2020 Sep 15;16:10964 [FREE Full text] [doi: 10.15766/mep_2374-8265.10964] [Medline: 32964120]
6. Behavioral risk factor surveillance system ACE data. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/violenceprevention/aces/ace-brfss.html [accessed 2021-12-13]
7. Sacks V, Murphey D. The prevalence of adverse childhood experiences, nationally, by state, and by race or ethnicity. Child Trends. 2018. URL: https://www.childtrends.org/publications/prevalence-adverse-childhood-experiences-nationally-state-race-ethnicity [accessed 2021-12-20]
8. Garner AS, Shonkoff JP, Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, Section on Developmental and Behavioral Pediatrics. Early childhood adversity, toxic stress, and the role of the pediatrician: translating developmental science into lifelong health. Pediatrics 2012 Jan;129(1):e224-e231. [doi: 10.1542/peds.2011-2662] [Medline: 22201148]
9. Gilgoff R, Singh L, Koita K, Gentile B, Marques SS. Adverse childhood experiences, outcomes, and interventions. Pediatr Clin North Am 2020 Apr;67(2):259-273. [doi: 10.1016/j.pcl.2019.12.001] [Medline: 32122559]
10. Cdc: Preventing ACEs may mitigate chronic health issues. American Academy of Family Physicians. URL: https://www.aafp.org/news/health-of-the-public/20191113mmwraces.html [accessed 2022-07-11]
11. Rariden C, SmithBattle L, Yoo JH, Cibulka N, Loman D. Screening for adverse childhood experiences: Literature review and practice implications. J Nurse Pract 2021 Jan;17(1):98-104 [FREE Full text] [doi: 10.1016/j.nurpra.2020.08.002] [Medline: 32963502]
12. Harris NB. Screening for adverse childhood experiences. JAMA 2020 Nov 03;324(17):1788-1789. [doi: 10.1001/jama.2020.16452] [Medline: 33141203]
13. ACEs Modules in BRFSS by state. Google Docs. URL: https://docs.google.com/spreadsheets/d/1yO2aIhGFKXaFKanRj7irIz7D_-Lwr4OPqI5hq8HwnRs/edit [accessed 2022-07-05]
14. BRFSS adverse childhood experience module. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/violenceprevention/pdf/acestudy/BRFSS_ACEModuleQuestions_2021_508.pdf [accessed 2023-08-04]
15. Adverse childhood experiences NSCH Data brief. Maternal and Child Health Bureau. URL: https://mchb.hrsa.gov/sites/default/files/mchb/data-research/nsch-ace-databrief.pdf [accessed 2023-08-04]
16. NSCH 2016 17: Adverse childhood experiences, nationwide. National Survey of Children’s Health. URL: https://www.nschdata.org/browse/survey/results?q=5545&r=1 [accessed 2022-03-10]
17. Liss DT, Chubak J, Anderson ML, Saunders KW, Tuzzio L, Reid RJ. Patient-reported care coordination: associations with primary care continuity and specialty care use. Ann Fam Med 2011;9(4):323-329 [FREE Full text] [doi: 10.1370/afm.1278] [Medline: 21747103]
18. Mendel WE, Sperlich M, Fava NM. "Is there anything else you would like me to know?": Applying a trauma-informed approach to the administration of the adverse childhood experiences questionnaire. J Community Psychol 2021 Jul;49(5):1079-1099 [FREE Full text] [doi: 10.1002/jcop.22562] [Medline: 33792050]

19. Purewal SK, Marques SS, Koita K, Bucci M. Assessing the integration of the Center for Youth Wellness Adverse Childhood Experiences Questionnaire (CYW ACE-Q) in a pediatric primary care setting. J Adolesc Health 2016 Feb;58(2):S47. [doi: 10.1016/j.jadohealth.2015.10.106]
20. Bifulco A, Schimmenti A. Assessing child abuse: "We need to talk!". Child Abuse Negl 2019 Dec;98:104236. [doi: 10.1016/j.chiabu.2019.104236] [Medline: 31683250]
21. Hargreaves MK, Mouton CP, Liu J, Zhou YE, Blot WJ. Adverse childhood experiences and health care utilization in a low-income population. J Health Care Poor Underserved 2019;30(2):749-767 [FREE Full text] [doi: 10.1353/hpu.2019.0054] [Medline: 31130549]
22. Goldstein E, Athale N, Sciolla AF, Catz SL. Patient preferences for discussing childhood trauma in primary care. Perm J 2017;21:16-055 [FREE Full text] [doi: 10.7812/TPP/16-055] [Medline: 28333604]
23. Vodicka E, Mejilla R, Leveille SG, Ralston JD, Darer JD, Delbanco T, et al. Online access to doctors' notes: patient concerns about privacy. J Med Internet Res 2013 Sep 26;15(9):e208 [FREE Full text] [doi: 10.2196/jmir.2670] [Medline: 24072335]
24. Kinney S, Darden A, Konrad KML, Darden PM, Madden C. Community engagement: Why Oklahoma? Why now? J Okla State Med Assoc 2018 Oct;111(8):790-794 [FREE Full text] [Medline: 31289410]
25. Cooper MT, Campbell J, Dileki N, Darden P. Access to care for children under five in Oklahoma: a geographic imputation application. J Okla State Med Assoc 2018 Oct;111(8):784-789 [FREE Full text] [Medline: 31289409]
26. Wheeler DL, Hackler JB. Preparing physicians for rural-based primary care practice: a preliminary evaluation of rural training initiatives at OSU-COM. J Am Osteopath Assoc 2017 May 01;117(5):315-324 [FREE Full text] [doi: 10.7556/jaoa.2017.057] [Medline: 28459472]
27. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: Interrelated drivers for healthy aging. Front Public Health 2020;8:231 [FREE Full text] [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
28. Child and adolescent health “Oklahoma Fact Sheet 2021: Strong Roots Grow a Strong Nation”. Child and Adolescent Health Measurement Initiative. URL: https://www.cahmi.org [accessed 2021-12-02]
29. Oklahoma health improvement plan 2020. Oklahoma State Department of Health. URL: https://oklahoma.gov/content/dam/ok/en/health/health2/documents/ohip-2020.pdf [accessed 2021-12-08]
30. Oklahoma primary care health care workforce gap analysis. Oklahoma State University.: Oklahoma State Department of Health; 2015. URL: https://oklahoma.gov/content/dam/ok/en/health/health2/documents/draft-osdh-gap-analysis-07012015.pdf [accessed 2021-12-02]
31. NSCH 2017 18: Adverse childhood experiences - 9 items, Nationwide vs. Oklahoma. National Survey of Children’s Health. URL: https://www.nschdata.org/browse/survey/results?q=7205&r=1&r2=38 [accessed 2023-02-27]
32. QuickFacts: Oklahoma County, Oklahoma. United States Census Bureau. URL: https://www.census.gov/quickfacts/fact/table/adaircountyoklahoma,oklahomacountyoklahoma/PST045221 [accessed 2022-10-18]
33. State statistics information. National Institute of Corrections. URL: https://nicic.gov/resources/nic-library/state-statistics [accessed 2023-06-21]
34. Pierannunzi C, Hu SS, Balluz L. A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004-2011. BMC Med Res Methodol 2013 Mar 24;13:49 [FREE Full text] [doi: 10.1186/1471-2288-13-49] [Medline: 23522349]
35. Cronholm PF, Forke CM, Wade R, Bair-Merritt MH, Davis M, Harkins-Schwarz M, et al. Adverse childhood experiences: Expanding the concept of adversity. Am J Prev Med 2015 Sep;49(3):354-361. [doi: 10.1016/j.amepre.2015.02.001] [Medline: 26296440]
36. Hunt B, Whitman S. Black: White health disparities in the United States and Chicago: 1990-2010. J Racial Ethn Health Disparities 2015 Mar;2(1):93-100. [doi: 10.1007/s40615-014-0052-0] [Medline: 26863247]
37. Lum TY, Vanderaa JP. Health disparities among immigrant and non-immigrant elders: the association of acculturation and education. J Immigr Minor Health 2010 Oct;12(5):743-753. [doi: 10.1007/s10903-008-9225-4] [Medline: 19184599]
38. Orsi JM, Margellos-Anast H, Whitman S. Black-White health disparities in the United States and Chicago: a 15-year progress analysis. Am J Public Health 2010 Feb;100(2):349-356. [doi: 10.2105/AJPH.2009.165407] [Medline: 20019299]
39. Volkers AC, Westert GP, Schellevis FG. Health disparities by occupation, modified by education: a cross-sectional population study. BMC Public Health 2007 Aug 08;7:196 [FREE Full text] [doi: 10.1186/1471-2458-7-196] [Medline: 17686141]
40. Ward-Peterson M, Wagner E. Health disparities. In: Gulliver SB, Cohen LM, editors. The Wiley Encyclopedia of Health Psychology. New York, NY: Wiley; 2020:59.
41. de Vera A. Tulsa program making it possible for women facing a life of incarceration to transform their future. FOX23 News. URL: https://www.fox23.com/news/women-recovery-is-making-it-possible-female-inmates-transform-their-lives/WC6IEAYPYRCGRJ5LQQUGOVLCPI/ [accessed 2022-07-12]
42. State-by-state data. The Sentencing Project. 2015. URL: https://www.sentencingproject.org/the-facts/ [accessed 2022-07-12]
43. Liming KW, Grube WA. Wellbeing outcomes for children exposed to multiple adverse experiences in early childhood: A systematic review. Child Adolesc Soc Work J 2018 Jan 19;35(4):317-335. [doi: 10.1007/s10560-018-0532-x]
44. Merrick MT, Ford DC, Ports KA, Guinn AS, Chen J, Klevens J, et al. Vital signs: Estimated proportion of adult health problems attributable to adverse childhood experiences and implications for prevention - 25 states, 2015-2017. MMWR

Morb Mortal Wkly Rep 2019 Nov 08;68(44):999-1005 [FREE Full text] [doi: 10.15585/mmwr.mm6844e1] [Medline: 31697656]

Abbreviations

ACE-Q: Adverse Childhood Experiences Questionnaire
ACE: adverse childhood experience
ASI: adverse childhood experience severity index
BRFSS: Behavioral Risk Factor Surveillance System
CMHC: community mental health center
CPS: child protective services
NSCH: National Survey of Children’s Health
OASIS: Oklahoma Adversity Surveillance Index System
ODMHSAS: Oklahoma Department of Mental Health and Substance Abuse Services
OKDHS: Oklahoma Department of Human Services
OSDH: Oklahoma State Department of Health
PAS: potential adverse childhood experience score
SDOH: social determinants of health

Edited by A Mavragani, T Sanchez; submitted 20.01.23; peer-reviewed by R Gilgoff, L Lafrado; comments to author 15.02.23; revised version received 07.03.23; accepted 18.07.23; published 22.08.23.

Please cite as:
Beaman JW, Miner CJ, Bolinger C
Quantifying Adverse Childhood Experiences in Oklahoma With the Oklahoma Adversity Surveillance Index System (OASIS): Development and Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e45891
URL: https://publichealth.jmir.org/2023/1/e45891
doi:10.2196/45891
PMID:37467063

Original Paper

Association Between Urinary Bisphenols and Body Composition Among American Adults: Cross-Sectional National Health and Nutrition Examination Survey Study

Jiakun Li[1,2*], MD; Zilong Zhang[1*], MD; Chichen Zhang[1*], MD; Qiyu Zhu[1]; Jing Zhao[2], BA; Hui Zong[2], PhD; Qi Deng[1]; Jiaming Zheng[1], MD; Erman Wu[2], MD; Rongrong Wu[2], PhD; Tong Tang[2], BA; Yihang Zhang[1], BA; Qiaosen Dong[1], MD; Yifan Li[1], MD; Jiao Wang[2], BA; Lu Yang[1], MD; Shi Qiu[1], MD; Bairong Shen[2], PhD; Qiang Wei[1], MD

[1]Department of Urology, West China Hospital, Sichuan University, Chengdu, China

[2]Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, China

[*]these authors contributed equally

Corresponding Author:
Qiang Wei, MD
Department of Urology
West China Hospital
Sichuan University
No. 37, Guoxue Alley, Chengdu, Sichuan, P.R.
Chengdu
China
Phone: 86 18980601425
Email: weiqiang933@126.com

Abstract

Background: Bisphenol A (BPA), bisphenol S (BPS), and bisphenol F (BPF) are widely used in various consumer products. They are environmental contaminants with estrogenic properties that have been linked to various health outcomes. Understanding their impact on body composition is crucial for identifying potential health risks and developing preventive strategies. However, most current studies have only focused on their relationship with BMI.

Objective: This study aimed to investigate the association between urinary levels of BPA, BPS, and BPF and body composition, including BMI, lean mass, and fat mass, in a large population-based sample.

Methods: We conducted a cross-sectional analysis using data from the National Health and Nutrition Examination Survey 2003-2016. Body composition data were assessed using dual-energy X-ray absorptiometry, which provided precise measurements of lean mass, fat mass, and other indicators. We used multivariate linear regression models to estimate the associations, adjusting for potential confounders such as age, gender, race, socioeconomic factors, and lifestyle variables.

Results: The results revealed significant associations between bisphenol exposure and body composition. After adjusting for covariates, BPS showed a positive association with BMI, with quartiles 3 and 4 having 0.91 (95% CI 0.34-1.48) and 1.15 (95% CI 0.55-1.74) higher BMI, respectively, compared with quartile 1 ($P<.001$). BPA was negatively associated with total lean mass (TLM) and appendicular lean mass, with quartiles 2, 3, and 4 having –7.85 (95% CI –11.44 to –4.25), –12.33 (95% CI –16.12 to –8.54), and –11.08 (95% CI –15.16 to –7.01) lower TLM, respectively, compared with quartile 1 ($P<.001$). BPS was negatively associated with TLM, with quartiles 3 ($\beta=-10.53$, 95% CI –16.98 to –4.08) and 4 ($\beta=-11.14$, 95% CI –17.83 to –4.45) having significantly lower TLM ($P=.005$). Both BPA and BPS showed a positive dose-response relationship with trunk fat (BPA: $P=.002$; BPS: $P<.001$) and total fat (BPA: $P<.001$; BPS: $P=.01$). No significant association was found between BPF and any body composition parameter.

Conclusions: This large-sample study highlights the associations between urinary levels of BPA and BPS and alterations in body composition, including changes in lean mass, fat mass, and regional fat distribution. These findings underscore the importance of understanding the potential health risks associated with bisphenol exposure and emphasize the need for targeted interventions to mitigate adverse effects on body composition.

(JMIR Public Health Surveill 2023;9:e49652) doi:10.2196/49652

KEYWORDS

bisphenols; body composition; environmental pollutants; public health; medical informatics

Introduction

Background

Bisphenol A (BPA) is an endocrine disrupter used globally in the manufacture of various plastic products including food containers, baby bottles, toys, and medical supplies [1,2]. It not only exhibits estrogen-like effects but also antagonizes thyroid hormone receptors, which may lead to adverse metabolic effects and diseases such as cardiovascular conditions, tumors, and other metabolism-related ailments [2,3].

Countries including Canada, the European Union, the United States, and China have prohibited the use of BPA in certain plastic products due to these health risks. The US Food and Drug Administration has set strict acceptable daily intake levels for BPA, but studies have indicated that even levels below the acceptable daily intake can contribute to obesity and related metabolic disturbances. Consequently, BPA substitutes, bisphenol S (BPS) and bisphenol F (BPF), were introduced [4].

These BPA derivatives (BPS and BPF), obtained by replacing the propane group in the molecular structure of BPA, are found to have estrogenic effects similar to those of BPA, indicating that they could also be unsafe [4]. They are prevalent in everyday packaging products and can be ingested through diet, air inhalation, or skin contact [5]. These chemicals can interfere with the nuclear hormone receptor signaling pathway, thereby affecting adipocyte proliferation and differentiation, which could potentially lead to a variety of metabolic diseases [4].

Body composition, including not only BMI but also fat mass and lean body mass, is a critical aspect of health [6]. Excessive body fat has been linked to adverse health outcomes including breast cancer [7,8] and cardiovascular risk factors [9]. Abdominal fat distribution is associated with an increased risk of type 2 diabetes [10], arterial stiffness [11], and cardiovascular disease [12,13], whereas lean body mass, predominantly composed of skeletal muscles, influences bone health and optimal aging [14,15]. Skeletal muscle is the largest and most plastic component of lean body mass (accounting for approximately 50% of lean body mass). Therefore, changes in lean body mass are primarily attributed to skeletal muscle mass. The maintenance of skeletal muscle is the basis of exercise, energy balance, and overall quality of life [16,17] and is critical for the risk of heart metabolic diseases [18].

Studies have shown that BPA promotes adipocyte differentiation and insulin resistance [19,20]. Previous studies have explored the relationship between urinary BPA and body composition, but these studies have been largely limited to specific populations, such as women who are nonobese and in the premenopausal stage or older adults [21,22]. At present, research on the association between BPA analogs (BPS and BPF) and body composition is scarce.

Objective

We aimed to bridge this gap by assessing the relationship between urinary BPA, BPS, BPF, and body composition in adult National Health and Nutrition Examination Survey (NHANES) participants using dual-energy X-ray absorptiometry (DXA) scans. By doing so, we hope to provide valuable insights into the health implications of BPA and its analogs in a broader population.

Methods

Patients Selection

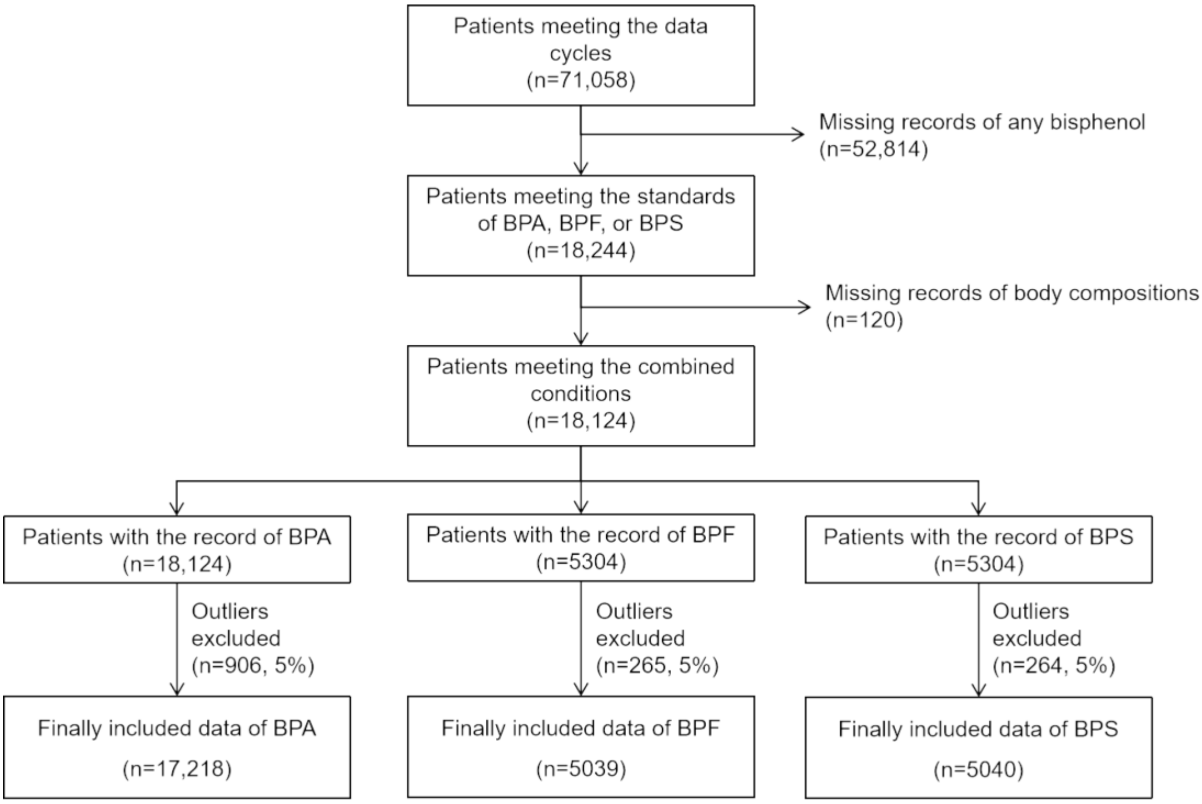
The NHANES database is a research program designed to assess the health and nutritional status of nonhospitalized adults and children in the United States. The unique feature of this survey is that it combines interviews and physical examination [23]. More specifically, nearly 7000 residents are randomly selected and invited to the NHANES interviews every year. The interviews include population-, socioeconomic status–, diet-, and health-related issues. The examination section includes medical, dental, and physiological measurements as well as laboratory tests conducted by trained medical personnel.

Because NHANES used the DXA method to measure body composition, only people aged 8 to 59 years were measured in this database. In addition, pregnant women, people who have used radiocontrast agent (barium) in the past 7 days, and people who weigh >204 kg or are taller than 195.58 cm have also been initially excluded from the NHANES surveys. Our study included and analyzed data from 7 NHANES survey cycles (2003-2004, 2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, and 2015-2016). Missing data on bisphenols (BPA, BPS, and BPF); BMI; or body compositions (total lean mass [TLM], appendicular lean mass [ALM], trunk fat [TRF], bone mineral content [BMC], and total fat [TOF]) were excluded.

Exclusion criteria were as follows: (1) data regarding BPA, BPS, or BPA were missing; (2) body composition data were missing; (3) the percentage of bisphenols (BPA, BPS, or BPF) >95% was defined as an outlier, and outliers were excluded.

We included 71,058 original records at first. After excluding the missing data and outliers, 17,218 data were finally included in the BPA analysis study, whereas 5039 and 5040 data were finally included in the BPF and BPS analysis studies, respectively. Detailed inclusion and exclusion criteria are shown in Figure 1.

Figure 1. Flowchart of inclusion and exclusion criteria among American adults from the National Health and Nutrition Examination Survey 2003 to 2016. BPA: bisphenol A; BPF: bisphenol F; BPS: bisphenol S.



Ethical Considerations

All study participants signed written informed consent forms, and the NHANES data collection was approved by the Research Ethics Review Board of the National Center for Health Statistics [24]. The NHANES has stringent consent protocols and procedures for ensuring confidentiality and protection against identification. This study was a secondary data analysis, which lacked personal identifiers, and the NHANES data are free for public use and available on the web [23]. Therefore, this study did not require an institutional review board review [24].

Measurement of Bisphenols

A sensitive phenol-detection method was used in this study [25]. This method is a web-based solid-phase extraction method coupled with high-performance liquid chromatography and tandem mass spectrometry. In addition, urine creatinine in each urine sample was measured and used to adjust the urine dilution. The specifics of the bisphenol measurements were provided by NHANES [26]. We sought to access the NHANES, gleaned the relevant data successfully, and recorded them in their database duly, corresponding to the measurements undertaken.

Measurement of Body Composition

The body composition included TLM, ALM, TRF, BMC, and TOF. DXA is the most widely accepted method for measuring body composition, partly because of its speed, ease of use, and low radiation exposure [27,28]. The NHANES DXA examination provides data on body composition (bone and soft tissue) as well as data on age, gender, and race and ethnic groups, which are generally representative of the United States, to study the association between body composition and other health conditions and risk factors, such as cardiovascular disease, diabetes, hypertension, and patterns of physical activity and dietary intake. DXA scans provide bone and soft tissue measurements of the entire body, arms and legs, trunk, and head.

The whole-body scan was obtained using the Hologic Discovery model A densitometer (version Apex 3.2; Hologic) [29]. The radiation exposure of whole-body DXA scanning was extremely low, less than 20 μSv. DXA examinations were performed by trained and certified radiologists. Further details of the DXA examination protocol are recorded in the Body Composition Procedure Manual on the NHANES website.

Covariates

Age, race, education level, marital status, and income:poverty ratio were among the demographic factors recorded. The income:poverty ratio was determined by dividing household income by the poverty threshold, specifically the size of the household, in the corresponding year and the state. BMI, which was determined by dividing weight by height squared, was one of the variables examined. Laboratory test results included serum calcium and phosphorus levels. The variables in the questionnaire included the coronary artery disease (CAD) score, alcohol use, physical activity, and smoking status. The CAD score was obtained from the Cardiovascular Disease and Health section of the NHANES database defined by the Rose questionnaire [30]. Energy (kcal) was calculated as the energy intake in food and beverages consumed within 24 hours before the interview (midnight to midnight), based on the energy of each food item in the US Department of Agriculture Food and Nutrient Database for Dietary Studies.

Data Analyses

We used appropriate survey procedures to explain the complex sampling design and weights of the NHANES [31]. The weights of the NHANES data consider several survey features, including the differential probabilities of selection for the sampling domains, survey nonresponse, and differences between the final sample distribution and the target population distribution. Each of the 3 levels of data collection for the NHANES (screening, interview, and examination) had a response rate. Subsequently, the sample weights were calculated for each level of data collection by the statistician of the National Center for Health Statistics. The selection of weights was carried out by the researchers based on the NHANES analysis protocol and the study content. This study selected the full sample of the 2-year mobile examination center exam weight. When we combined the weights of 7 cycles of the NHANES study, we divided the raw weight data by 7. Continuous variables were expressed as mean (95% CI), whereas categorical variables were expressed as percentages. The Student 2-tailed *t* test or Mann-Whitney *U* test was used to compare continuous variables, and the chi-square test or Fisher exact test was used to compare categorical variables.

A multivariate linear regression model was used to estimate the association between the body composition of phenolic substances, including an unadjusted model (nonadjusted); a minimally adjusted model (adjust I; adjusted only for gender, age, race, energy [kcal], and creatinine, urine [mg/100 mL]); and a fully adjusted model (adjust II; adjusted for gender, age, race, ratio of family income to poverty, education level, marital status, CAD score, smoking status, alcohol intake, physical activity, energy [kcal], and creatinine, urine (mg/100 mL]).

In the aforementioned models, body composition and body weight were used for adjustment in all analyses, except BMI and body weight. In addition, for missing covariate data, the method of missing value interpolation was supplemented. The interpolation threshold was set to 5%. Phenolic substances (BPA, BPS, or BPF) were considered continuous independent variables and categorical variables (divided into quartiles), and the lowest quartile was used as the reference. As the detection rate of BPF was low, quartiles 1 and 2 of the original data were combined to form a new quartile.

We performed tests for linear trends by entering the median value of each category as a continuous variable in the models. Interactive and stratified analyses were performed. Each stratification was adjusted for all factors (gender, age, alcohol intake, race, smoking status, and physical activity) except for the stratification factor itself.

All statistical analyses were performed using the software packages R (version 4.2.0; The R Foundation) and Empower (X&Y Solutions Inc) [32]. Double-tailed *P*<.05 was considered statistically significant.

Results

In total, 18,124 participants were selected for our study. Among them, 8988 were male and 9133 were female, the median age was 39.49 (IQR 38.94-40.04) years, and the median BMI was 27.22 (IQR 27.03-27.41) kg/m^2. Among the 18,124 participants, we conducted statistical analysis on 17,218 (95%) participants with BPA data, 5039 (27.8%) people with BPS data, and 5040 (27.8%) people with BPF data according to the information from the NHANES database. In addition, urinary BPA, BPS, and BPF were 2.27 (95% CI, 2.22-2.33), 0.70 (95% CI, 0.65-0.75), and 0.72 (95% CI, 0.66-0.78) ng/mL, respectively. Detailed characteristics of the participants are presented in Table 1.

Table 1. Characteristics, including baseline, urinary bisphenols and body compositions, of American adults from the National Health and Nutrition Examination Survey 2003-2016.

Characteristics[a]	Total (n=18,124)	BPA[b] (n=17,218)	BPS[c] (n=5039)	BPF[d] (n=5040)
Age (years), median (IQR)	39.49 (38.94-40.04)	39.65 (39.10-40.21)	40.13 (38.99-41.27)	40.18 (39.02-41.35)
Energy (kcal), mean (95% CI)	2171.8 (2149.81-2193.79)	2172 (2150.07-2193.94)	2120.14 (2078.95-2161.34)	2110.17 (2067.34-2152.99)
Creatinine, urine (mg/100 mL), mean (95% CI)	123.65 (121.61-125.69)	120.65 (118.55-122.76)	122.32 (118.30-126.34)	122.1 (117.95-126.24)
Weight (kg), mean (95% CI)	75.67 (75.09-76.25)	75.58 (74.99-76.17)	76.23 (74.99-77.46)	75.92 (74.73-77.11)
BMI (kg/m^2), mean (95% CI)	27.22 (27.03-27.41)	27.18 (26.99-27.38)	27.64 (27.20-28.07)	27.61 (27.19-28.03)
TLM[e] (g), mean (95% CI)	48,965.91 (48,524.82-49,406.99)	48,954.92 (48,507.70-49,402.14)	49,114.27 (48,203.47-50,025.07)	48,934.37 (48,007.14-49,861.59)
ALM[f] (g), mean (95% CI)	21,163.06 (20,942.60-21,383.52)	21,148.1 (20,926.20-21,370.00)	20,985.1 (20,546.07-21,424.14)	20,932.29 (20,486.99-21,377.58)
TRF[g] (g), mean (95% CI)	12,231.43 (11,976.08-12,486.77)	12,176.15 (11,920.52-12,431.78)	12,095.54 (11,660.14-12,530.94)	12,072.51 (11,652.44-12,492.58)
BMC[h] (g), mean (95% CI)	2260.29 (2242.46-2278.12)	2260.52 (2242.28-2278.76)	2191.89 (2164.24-2219.53)	2185.9 (2156.96-2214.85)
TOF[i] (g), mean (95% CI)	25,522.09 (25,072.09-25,972.09)	25,442.36 (25,001.33-25,883.38)	25,447.67 (24,684.67-26,210.67)	25,319.4 (24,577.47-26,061.33)
Urinary bisphenols (ng/mL), mean (95% CI)	N/A[j]	2.27 (2.22-2.33)	0.7 (0.65-0.75)	0.72 (0.66-0.78)
Gender, %				
Male	49.10	48.91	48.08	47.54
Female	50.90	51.09	51.92	52.46
Race, %				
Mexican American	9.40	9.48	10.18	10.65
Other Hispanic	5.46	5.41	6.34	6.56
Non-Hispanic Black	12.12	11.93	12.16	12.33
Non-Hispanic White	65.68	65.81	62.21	61.04
Other race	7.35	7.37	9.11	9.42
Ratio of family income to poverty, %				
<1.3	23.49	23.13	24.51	25.54
1.3-3.5	35.73	35.45	35.28	35
>3.5	40.79	41.42	40.21	39.47
Education level, %				
Less than high school	18.15	18.12	16.36	16.41
High school or general educational development	22.20	21.89	20.09	20.45
Above high school	59.65	59.99	63.55	63.13
Marital status, %				
Married or living with partner	69.05	69.30	69.35	68.99
Living alone	30.95	30.70	30.65	31.01
CAD[k] score, %				
0	62.07	62.08	59.20	60.07
1	24.15	24.18	24.92	24.12
2	13.79	13.74	15.88	15.80

Characteristics[a]	Total (n=18,124)	BPA[b] (n=17,218)	BPS[c] (n=5039)	BPF[d] (n=5040)
Smoking, %				
Never	54.54	54.76	57.76	57.78
Former	28.03	27.96	27.98	27.73
Current	17.43	17.29	14.25	14.48
Alcohol intake (per day), %				
None	77	76.86	78.26	78.17
Moderate	7.20	7.19	6.27	6.05
Heavy	15.80	15.95	15.47	15.79
Physical activity, %				
Less than moderate	39.11	38.91	42.73	42.96
Moderate	15.69	15.66	10.19	10.46
Vigorous	45.20	45.43	47.09	46.59

[a] The data showed for continuous variables were survey-weighted mean (95% CI) and for categorical variables, survey-weighted percentages (the absolute values cannot be provided due to being adjusted by weighting).

[b] BPA: bisphenol A.

[c] BPS: bisphenol S.

[d] BPF: bisphenol F.

[e] TLM: total lean mass.

[f] ALM: appendicular lean mass.

[g] TRF: trunk fat.

[h] BMC: bone mineral content.

[i] TOF: total fat.

[j] N/A: not applicable.

[k] CAD: coronary artery disease.

Associations between bisphenols (BPA, BPS, and BPF) and body compositions are presented in Table 2. A dose-response relationship was observed between BPA and BMI in the crude model (*P*<.001). The relationship did not exist after full adjustment; however, there was a positive association between BPS and BMI, with 0.91 and 1.15 higher BMI in quartiles 3 and 4, respectively, compared with quartile 1 (*P*<.001). In addition, there was also a dose-response relationship between BPA and TLM in the minimally and fully adjusted models (all *P*<.001), whereas a negative association was found between BPS and TLM after adjusting for all potential confounders, compared with quartile 1; participants in quartiles 3 and 4 had significantly lower TLM (quartile 3, model 2: β=–10.53, 95% CI –16.98 to –4.08; quartile 4, model 2: β=–11.14, 95% CI –17.83 to –4.45; *P*=.005). After adjusting for all confounders, a dose-response relationship was also found between BPA and ALM, with those in quartiles 2, 3, and 4 having significantly lower ALM than those in quartile 1 (quartile 2: β=–3.31, 95% CI –5.73 to –.90; quartile 3: β=–5.46, 95% CI –8.01 to –2.90; quartile 4: β=–4.80, 95% CI –7.55 to –2.05; *P*=.01). However, there was no substantial association between BPS and ALM. In addition, there was a dose-response relationship between BPA and TRF or TOF. For TRF, participants in quartiles 2, 3, and 4 (model 2: β=3.55, 95% CI 1.17-5.93; β=8.29, 95% CI 5.78-10.79; β=5.68, 95% CI 2.98-8.37, respectively; *P*=.002) had a significantly higher TRF than those at quartile 1. Similarly, there was a dose-response relationship between BPA and TOF (model 2: *P*<.001). For BPS, a positive association was identified between BPS and TRF or TOF; those in quartiles 3 and 4 had 5.16 and 8.41 higher TRF, 9.92 and 10.34 higher TOF when compared with quartile 1, respectively (for TRF, model 2: *P*<.001; for TOF, model 2: *P*=.01). Furthermore, although there was no significant association between BPA and BMC, there was a negative association between BPS and TOF (model 2: *P*<.001). After complete adjustment, there was no substantially association between BPF and body composition indicators (Table 2 and Multimedia Appendix 1).

Table 2. Association between bisphenols and body composition among American adults from the National Health and Nutrition Examination Survey 2003-2016.

	Nonadjusted		Adjusted model I[a]		Adjusted model II[b]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Urinary bisphenol A						
BMI (kg/m²)		<.001[c]		.23[c]		.11[c]
0.14-0.7	Ref[d]		Ref		Ref	
0.8-1.5	.62 (.31 to .92)	<.001	.31 (.02 to .59)	.04	.19 (–.09 to .46)	.19
1.6-3.1	1.19 (.89 to 1.49)	<.001	.46 (.16 to .76)	.002	.35 (.06 to .64)	.02
3.2-11	1.01 (.70 to 1.31)	<.001	–.01 (–.34 to .32)	.95	–.11 (–.43 to .20)	.48
TLM[e]		.11[c]		<.001[c]		<.001[c]
0.14-0.7	Ref		Ref		Ref	
0.8-1.5	–1.65 (–6.58 to 3.28)	.51	–8.96 (–12.66 to –5.26)	<.001	–7.85 (–11.44 to –4.25)	<.001
1.6-3.1	–2.15 (–7.05 to 2.76)	.39	–13.17 (–17.06 to –9.28)	<.001	–12.33 (–16.12 to –8.54)	<.001
3.2-11	3.00 (–1.98 to 7.97)	.24	–11.62 (–15.79 to –7.45)	<.001	–11.08 (–15.16 to –7.01)	<.001
ALM[f]		<.001[c]		.005[c]		.01[c]
0.14-0.7	Ref		Ref		Ref	
0.8-1.5	1.47 (–1.63 to 4.58)	.35	–4.07 (–6.53 to –1.61)	.001	–3.31 (–5.73 to –.90)	.007
1.6-3.1	2.93 (–.17 to 6.02)	.06	–5.85 (–8.44 to –3.25)	<.001	–5.46 (–8.01 to –2.90)	<.001
3.2-11	6.33 (3.19 to 9.47)	<.001	–5.43 (–8.22 to –2.65)	<.001	–4.80 (–7.55 to –2.05)	<.001
TRF[g]		.85[c]		<.001[c]		.002[c]
0.14-0.7	Ref		Ref		Ref	
0.8-1.5	1.63 (–1.37 to 4.63)	.29	4.53 (2.06 to 7.00)	<.001	3.55 (1.17 to 5.93)	.004
1.6-3.1	5.78 (2.79 to 8.76)	<.001	9.03 (6.44 to 11.63)	<.001	8.29 (5.78 to 10.79)	<.001
3.2-11	.76 (–2.27 to 3.79)	.62	6.66 (3.88 to 9.45)	<.001	5.68 (2.98 to 8.37)	<.001
TOF[h]		.23[c]		<.001[c]		<.001[c]
0.14-0.7	Ref		Ref		Ref	
0.8-1.5	2.13 (–2.99 to 7.25)	.42	9.66 (5.76 to 13.57)	<.001	8.58 (4.78 to 12.39)	<.001
1.6-3.1	3.73 (–1.37 to 8.83)	.15	15.07 (10.96 to 19.18)	<.001	14.19 (10.18 to 18.20)	<.001
3.2-11	–2.04 (–7.21 to 3.12)	.44	13.08 (8.68 to 17.48)	<.001	12.51 (8.20 to 16.82)	<.001
Urinary bisphenol S						
BMI (kg/m²)		<.001[c]		<.001[c]		<.001[c]
0.07-0.1	Ref		Ref		Ref	
0.2-0.3	.73 (.13 to 1.34)	.02	.27 (–.30 to .84)	.35	.26 (–.29 to .81)	.35
0.4-0.8	2.02 (1.42 to 2.62)	<.001	1.04 (.45 to 1.62)	<.001	.91 (.34 to 1.48)	.002
0.9-4.7	2.70 (2.10 to 3.31)	<.001	1.31 (.70 to 1.91)	<.001	1.15 (.55 to 1.74)	<.001
TLM		.006[c]		<.001[c]		.005[c]
0.07-0.1	Ref		Ref		Ref	
0.2-0.3	–2.53 (–10.95 to 5.88)	.56	–4.46 (–10.82 to 1.89)	.17	–3.81 (–10.05 to 2.44)	.23
0.4-0.8	–6.42 (–14.64 to 1.80)	.13	–12.79 (–19.32 to –6.25)	<.001	–10.53 (–16.98 to –4.08)	.001
0.9-4.7	–11.26 (–19.59 to –2.92)	.008	–13.51 (–20.28 to –6.74)	<.001	–11.14 (–17.83 to –4.45)	.001
TRF		<.001[c]		<.001[c]		<.001[c]

	Nonadjusted		Adjusted model I[a]		Adjusted model II[b]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
0.07-0.1	Ref		Ref		Ref	
0.2-0.3	1.92 (–3.19 to 7.03)	.46	1.79 (–2.41 to 5.99)	.40	1.00 (–3.08 to 5.08)	.63
0.4-0.8	7.53 (2.58 to 12.48)	.003	6.97 (2.70 to 11.24)	.001	5.16 (.99 to 9.34)	.02
0.9-4.7	13.05 (8.00 to 18.10)	<.001	10.45 (6.00 to 14.91)	<.001	8.41 (4.05 to 12.76)	<.001
TOF		.01[c]		.002[c]		.01[c]
0.07-0.1	Ref		Ref		Ref	
0.2-0.3	1.29 (–7.48 to 10.05)	.77	3.27 (–3.46 to 10.01)	.34	2.64 (–3.99 to 9.27)	.44
0.4-0.8	6.06 (–2.50 to 14.63)	.17	12.24 (5.31 to 19.18)	<.001	9.92 (3.07 to 16.78)	.005
0.9-4.7	10.09 (1.39 to 18.79)	.02	12.71 (5.53 to 19.90)	<.001	10.34 (3.23 to 17.45)	.004
BMC[i]		<.001[c]		<.001[c]		<.001[c]
0.07-0.1	Ref		Ref		Ref	
0.2-0.3	–.04 (–.66 to .58)	.90	–.10 (–.72 to .51)	.74	–.01 (–.62 to .60)	.97
0.4-0.8	–1.21 (–1.82 to –.60)	<.001	–1.21 (–1.85 to –.58)	<.001	–1.00 (–1.63 to –.38)	.002
0.9-4.7	–1.75 (–2.37 to –1.13)	<.001	–1.70 (–2.36 to –1.05)	<.001	–1.46 (–2.11 to –.81)	<.001

[a] Model I was adjusted for gender, age, race, energy, creatinine, and urine.

[b] Model II was adjusted for gender, age, race, ratio of family income to poverty, education level, marital status, CAD score, smoking, alcohol intake per day, physical activity, energy, and creatinine, urine.

[c] *P* value for trend.

[d] Ref: reference.

[e] TLM: total lean mass.

[f] ALM: appendicular lean mass.

[g] TRF: trunk fat.

[h] TOF: total fat.

[i] BMC: bone mineral content.

The results of the interaction and stratified analyses are shown in Table 3. The associations between BPA, BPF, and body compositions were not significantly modified by gender. However, the association between BPS and body composition parameters (TLM, TRF, and TOF) was stronger among female participants (for interaction: *P*=.002, *P*=.007, and *P*=.003, respectively). A dose-response relationship was found between BPS and body composition parameters (TLM, TRF, and TOF; for trend *P*<.001, *P*<.001, and *P*=.001, respectively). In addition, the association of BPA and BPS with BMI was significantly modified by age (for interaction, *P*=.02 and *P*=.04, respectively). BPA was inversely associated with BMI in patients aged ≤20 years (for trend, *P*<.001). Compared with quartile 1, the BMI level decreased by 1.11 in quartile 4 (95% CI –1.61 to –.62; *P*<.001), whereas there was a dose-response association between BPA and BMI in people aged 20 to 50 years (*P*=.01, *P*<.001, and *P*=.003, respectively). For BPS, the association between BPS and BMI was stronger in people aged 20 to 50 years, and a dose-response association was found (for trend, *P*=.02). Furthermore, alcohol intake was a potential confounder in the association between BPF and ALM (for interaction, *P*=.001). Finally, the observed associations between bisphenols (BPA, BPS, and BPF) and body composition parameters were not significantly modified by race, smoking status, and physical activity (data not shown; Multimedia Appendix 2).

Table 3. Interaction and stratified analyses among American adults from the National Health and Nutrition Examination Survey 2003-2016.

	Quartile 1	Quartile 2		Quartile 3		Quartile 4		*P* value (trend)	*P* value (interaction)
		β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value		
Urinary bisphenol A[a]									
BMI (kg/m^2; per age in years)									.02
Age ≤20	0	–.69 (–1.13 to –.25)	.002	–.87 (–1.33 to –.41)	<.001	–1.11 (–1.61 to –.62)	<.001	<.001	
Age 20-50	0	.64 (.13 to 1.16)	.01	1.19 (.64 to 1.74)	<.001	.90 (.31 to 1.50)	.003	.04	
Age ≥50	0	.12 (–.34 to .58)	.61	.23 (–.25 to .71)	.35	–.28 (–.79 to .24)	.30	.13	
Urinary bisphenol S[b]									
BMI (kg/m^2; per age in years)									.04
Age ≤20	0	–.53 (–1.31 to .26)	.19	–.73 (–1.58 to .11)	.09	–.36 (–1.18 to .46)	.39	.89	
Age 20-50	0	.26 (–.77 to 1.29)	.62	1.45 (.36 to 2.55)	.009	1.34 (.21 to 2.48)	.02	.02	
Age ≥50	0	.37 (–.55 to 1.29)	.43	1.11 (.18 to 2.05)	.02	.69 (–.27 to 1.66)	.16	.39	
TLM[c]									.002
Male	0	–4.91 (–15.02 to 5.20)	.34	–5.49 (–15.61 to 4.63)	.29	–9.11 (–19.68 to 1.45)	.09	.15	
Female	0	3.47 (–5.16 to 12.10)	.43	–3.19 (–12.06 to 5.68)	.48	–11.26 (–20.43 to –2.08)	.02	<.001	
TRF[d]									.007
Male	0	2.16 (–3.72 to 8.04)	.47	3.16 (–2.68 to 9.00)	.29	4.87 (–1.24 to 10.99)	.12	.16	
Female	0	–2.59 (–8.40 to 3.22)	.38	1.67 (–4.24 to 7.59)	.58	9.93 (3.79 to 16.08)	.002	<.001	
TOF[e]									.003
Male	0	5.11 (–5.61 to 15.83)	.35	5.45 (–5.31 to 16.21)	.32	9.42 (–1.81 to 20.65)	.10	.16	
Female	0	–4.78 (–13.98 to 4.42)	.31	3.98 (–5.46 to 13.42)	.41	11.19 (1.40 to 20.98)	.03	.001	
Urinary bisphenol F[f]									
ALM[g]									.01
Alcohol intake									
None	0	–1.21 (–5.55 to 3.13)	.58	–.96 (–5.17 to 3.25)	.65	N/A[h]	N/A	.65	
Moderate	0	–1.25 (–23.21 to 20.71)	.91	1.49 (–20.47 to 23.45)	.89	N/A	N/A	.69	
Heavy	0	–9.23 (–24.19 to 5.73)	.23	–15.31 (–30.72 to .09)	.05	N/A	N/A	.28	

[a] Quartile ranges of urinary bisphenol A: quartile 1=0.14 to 0.7; quartile 2=0.8 to 1.5; quartile 3=1.6 to 3.1; quartile 4=3.2 to 11.

[b] Quartile ranges of urinary bisphenol S: quartile 1=0.07 to 0.1; quartile 2=0.2 to 0.3; quartile 3=0.4 to 0.8; quartile 4=0.9 to 4.7.

[c] TLM: total lean mass.

[d] TRF: trunk fat.

[e] TOF: total fat.

[f] Quartile ranges of urinary bisphenol F: quartile 1=0.3 to 0.6; quartile 2=0.7 to 7.5; quartile 3>7.5.

[g] ALM: appendicular lean mass.

[h] N/A: not applicable.

Discussion

Principal Findings

In this study, we discovered that there was no significant association between BPA and BMI, whereas BPA was negatively associated with TLM and ALM and had a positive dose-response with TOF and TRF. BPS was positively associated with BMI, TRF, and TOF and negatively associated with TLM, and there was no discernible relationship between BPS and ALM. However, no significant association was observed between BPF and body composition. Although previous studies have revealed that BPA is positively associated with human BMI [33-35], and urinary BPA content and phthalate metabolites are positively associated with body fat content in older adults and premenopausal women [1], it is unclear whether BPA and its derivatives (BPS and BPF) have similar effects on the specific body composition of the general population.

Eating habits may have influenced the study results. Wang et al [36] conducted a survey that included 360 primary and junior high school students in Shanghai as participants and found a positive association between BMI and BPA [36]. Regarding the possible effects of BPA exposure on adult BMI, the study by Carwile and Michels [33] that included 2747 adults participating in the 2003 to 2006 NHANES found that despite the nonlinear association between BPA and BMI, the BMI of the participants in the second, third, and fourth BPA quartiles were higher than those in the primary BPA quartile. Other studies investigated the relationship between BPA and BMI in young Indian women (age: mean 21.2, SD 2.37 years) and revealed that there was a significant association between BPA and BMI (P=.03) [37]. None of the 3 studies included diet as a possible covariate. In contrast, Braun et al [38] found no significant association between urinary BPA levels and BMI of prenatal women and 2-to 5-year-old children in the Cincinnati region after adjusting for diet, exercise, and di(2-ethylhexyl) phthalate. Thus, the differences among these findings may lie in whether diet was selected as a possible covariate and was well adjusted. Diet is an important source of exposure to BPA [39-41], whereas obesity is associated with certain dietary patterns, such as the Mediterranean diet [42] or ketogenic diet [43]. Previous studies indicated that consuming a substantial amount of plant-based foods in the Mediterranean diet may potentially reduce the severity of nonalcoholic fatty liver disease–related outcomes [44] and relate to favorable measures of adiposity [45]. Furthermore, a highly restrictive ketogenic diet is considered an effective and safe therapeutic intervention for patients who are obese [46,47], favoring their body composition [48]. In our study, total calorie intake, TOF intake, and alcohol intake were included as possible dietary covariates; the participants in the 2003 to 2004 and 2005 to 2006 NHANES were selected as participants of this study, and their body composition and urinary BPA level were well measured. No significant association was found between BMI and BPA in the fully adjusted model. This was consistent with the trend described by Braun et al [38].

There was a positive association between BPS and BMI, which is consistent with the findings of Jala et al [37]. In the study of the effect of BPS on male sperm quality, Ghayda et al [49] also found that men with high BMI have correspondingly higher urine BPS levels. For lean mass, both BPA and BPS showed a negative dose-response relationship with TLM, and there was a negative association between BPA and ALM in the fully adjusted model. Corbasson et al [1] found a negative association between BPA and TLM in the male group after adjusting for the covariates of waist circumference and BMI. However, this association was not observed in the female group. For body fat, urinary BPA and BPS levels in the NHANES participants were positively associated with TRF and TOF. Inconsistent with our findings, there was no significant relationship between BPA and fat mass in the study by Corbasson et al [1]. However, because of the use of NHANES as a data source for us and this study, our study incorporated an extended range of data periods, which may engender more inferences. Zhang et al [50] found that abnormal liver metabolism in mice, glycolysis, trichloroacetic acid metabolic pathway effects, and decreased succinic acid levels were related to visceral fat synthesis, indicating that BPS may also lead to an increase in TRF and TOF [51].

BPF was not significantly associated with TRF or TOF, possibly because current environmental BPF exposure is insufficient as a decisive factor for obesity. For example, Liu et al [52] used the data of 1521 participants in the NHANES 2013 to 2014 to evaluate the relationship between BPF and obesity. They also found that BPF and obesity were not significantly associated with obesity after adjusting for factors such as demographic, socioeconomic, lifestyle, and urinary creatinine concentration; however, a significant association was observed in the nonadjusted model. However, this might be an adjustment for sensitive populations, such as low-income individuals who are prone to overexposure. In their study, the exposure of BPF before and after adjustment was 0.4 versus 0.3 ng/mL, respectively, which was much less than the BPA exposure of 1.5 versus 1.1 ng/mL [52], and studies have shown that BPF can interfere with the IRS-1/PI3K/AKT signal transduction pathway and subsequently affect the insulin-dependent glucose metabolism pathway of adipocytes (3T3-L1 adipocytes), resulting in obesity [53].

According to the stratified analyses, the association between bisphenols and body composition was influenced by gender, age, and the potential impact of alcohol. Previous research has reported gender differences in the effects of BPA on obesity [54], and our study has also demonstrated such differences. On the basis of previous studies, this may be attributed to the differences in BPA metabolism and estrogen receptor expression, as well as gender-related BPA exposure [55,56].

Gender-related hormonal variations can lead to different responses to bisphenol exposure [57,58]. The Environmental Obesogen Hypothesis suggests that prenatal or early life exposure to estrogen-like environmental pollutants may increase the risk of fat accumulation and obesity [59]. For instance, BPA may have the greatest impact during early developmental stages, potentially due to epigenetic modifications related to stem cells [60]. In addition, alcohol consumption is associated with obesity, as it can affect liver metabolism and function [61]. The interaction between alcohol consumption and bisphenol compounds requires further research and validation.

Estrogen receptors play a crucial role in regulating lipid accumulation and improving insulin sensitivity and are therefore regarded as therapeutic targets for preventing obesity and metabolic disorders [57]. In vitro studies have previously observed that BPS and BPF exhibit similar toxicity and estrogenic effects as BPA, exerting an influence on cellular metabolism through pathways such as growth inhibition, oxidative stress, and gene expression [62]. Another study demonstrated the estrogenic agonistic effects of BPA, BPS, and BPF, individually, and the synergistic effect of these 3 bisphenols when combined, exhibiting a more potent action [63]. By contrast, obesity and metabolic disorders have been associated with oxidative stress [64]. Insulin resistance, a key factor in the pathogenesis of obesity and diabetes, is intricately linked to oxidative stress [65]. BPA and its derivatives, BPS and BPF, can disrupt the endogenous antioxidant system, leading to a severe imbalance in enzymes, such as superoxide dismutase and catalase, in liver cells exposed to BPS [66]. Furthermore, studies have indicated that adult exposure to BPF results in a significant imbalance in catalase activity [67]. Regarding the findings of this study, it is evident that both BPA and BPS have a significant impact on body composition, particularly on the content and distribution of adipose tissue. The potential mechanisms of BPS toxicity may involve interactions with estrogen receptors [68], DNA, and protein binding [69], as well as oxidative stress effects [70].

In this study, both BPA and BPS exhibited significant associations with body composition, displaying trends similar to those of BMI, TLM, TRF, and TOF. Furthermore, BPA was correlated with ALM, whereas BPS demonstrated an additional association with BMC. However, no significant correlation was observed between BPF and body composition. Studies suggest that BPS has a lesser obesogenic effect compared with BPA, with its ability to induce triglyceride accumulation being only 1% of that of BPA [71]. Nevertheless, for BPF exposure, some studies indicate an association with developing obesity, whereas others yield inconclusive results [72]; thus, further research is needed to elucidate its role.

In this cross-sectional study, although we were unable to establish a causal relationship between BPA and its derivatives, BPS and BPF, and body composition, we identified an association between them. The analysis conducted in this study not only provides a reference value for understanding the mechanisms of bisphenols but also holds potential value in guiding individuals to reduce adverse changes in body composition and mitigate metabolic-related reactions by proper exposure management or avoidance of bisphenols. Furthermore, it presents potential references for exploring the pathogenic mechanisms and preventive measures for obesity and other metabolic-related disorders. These findings underscore the importance of understanding the potential health risks associated with bisphenol exposure and emphasize the need for targeted interventions to mitigate the potential adverse effects on body composition.

Furthermore, to our knowledge, this is an innovative study to evaluate the association between specific body compositions, including ALM, TLM, TOF, TRF, and urinary BPA derivatives, such as BPS and BPF. Compared with previous studies, we included alcohol, calorie, and fat intake as possible covariates in the analysis to exclude the possible effects of diet on obesity based on the association between obesity and BPA as well as its derivatives. Thus, this study’s findings more accurately reflect the association between endocrine disrupters and body components. Another advantage of our study is that the American NHANES cohort has a large population base and ethnic diversity. Statistics was used for weighted analysis to represent a larger population. We also evaluated the association between BPS and TLM and found that BPS was negatively associated with TLM, which has not been previously reported. Lean body mass is significant for bone health and optimal aging [1]. The skeletal muscle is the largest and most malleable component of lean body mass (accounting for approximately 50% of lean body mass). Therefore, the change in lean body mass is mainly attributed to skeletal muscle mass. The maintenance of skeletal muscle is fundamental to locomotion, energy homeostasis, and overall quality of life [73]. This suggests that exposure to BPS may lead to a decline in exercise capacity, as the substitution of BPA also has safety issues. In fact, some studies have shown that BPS and BPA are positively associated with oxidative stress, which is associated with an increase in triglycerides and indirectly affects the physical fitness of healthy young adults (age: mean 23.5, SD 2 years). Jing et al [74] also found that the myotube diameter and distribution area of C2C12 myoblasts that were exposed to 10 to 4 M BPS decreased significantly compared with the control group.

Limitations

Our study has several limitations. First, bisphenol exposure was measured using the urine of the NHANES participants. However, it is widely acknowledged that urinary BPA and its derivatives do not fully reflect bisphenol exposure. In fact, their levels change greatly over a short period, which may lead to uncertain relationships between BPA, BPS, BPF, and body composition [1]. Therefore, it is important to conduct repeated sampling measurements of the population. In addition, our covariates were limited and did not include other possible disrupters, such as exposure to other endocrine disrupters or drug intake, which can affect metabolism during the study period and the menopausal state of female individuals. Moreover, the data we used were collected from the urine of each participant only once, so it could not fully reflect the exposure to BPA and its derivatives. However, repeated sample collection may also have affected the enthusiasm of the participants, resulting in bias in the results. Nevertheless, urine collection is the most effective and widely adopted method for detecting BPA and

BPS (bisphenol derivatives). The NHANES and other studies have generally used indirect methods to detect BPA levels. However, some studies have suggested that the bisphenol content in the human body may be disputable due to different detection methods. In 2020, a study proposed that direct methods measure BPA levels nearly 20 times higher than indirect methods, which may change our understanding of BPA levels in humans [75]. Nevertheless, the trend in the results obtained using mainstream detection methods is unquestionable. We found a negative dose-response between BPF, BPA, and TLM, which provides a basis for future research. However, the specific mechanism remains unclear and should be further explored in future studies. Additionally, the characteristics of human exposure to bisphenols are mixed. In this study, we analyzed the independent records of bisphenol compounds in the database, and we will analyze the effects of all bisphenol mixtures together by experimental methods for future research.

Conclusions

Our study found that BPA, as a representative bisphenols, is related to changes in body composition. Although people have begun to pay more attention to BPA, the impact of its derivatives on humans cannot be ignored. As a derivative of BPA, BPS is related to the changes in various body components and has a stronger association with women and changes in age. BPF was not significantly associated with various body components in this study. We speculate that the degree of environmental exposure to BPF is not sufficient to cause a significant impact, but this conjecture needs to be elucidated through additional experiments.

Acknowledgments

The authors thank Dr Chi Chen, Changzhong Chen, and Xing-Lin Chen for providing the statistical methodology consultation.

This work was supported by Sichuan University (No. 2023SCU12057).

Data Availability

The data used for this study were obtained from the National Health and Nutrition Examination Survey (NHANES) database, 2003 to 2016 [23], and only publicly available data were used in this study. The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

JL, ZZ, and CZ contributed equally as the first authors of this manuscript. SQ and QW were responsible for the conception and design of this study. QZ, JW, HZ, TT, Q Deng, YZ, Q Dong, and YL interpreted the data. CZ, JL, ZZ, JZ, RW, EW, BS, and SQ were responsible for the data acquisition. JL, ZZ, and CZ wrote the first draft, interpreted the data, and wrote the final version of the manuscript. JZ was committed to the language editing of the manuscript. All authors critically revised the manuscript for important intellectual content and approved the final version of the manuscript. The authors have read and approved the final manuscript. SQ, BS, and QW contributed equally as the corresponding authors of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
All results of the associations between bisphenols and body composition among American adults that were obtained from the National Health and Nutrition Examination Survey 2003 to 2016.
[DOC File , 69 KB - publichealth_v9i1e49652_app1.doc]

Multimedia Appendix 2
All results of the interaction and stratified analyses between bisphenols and body composition among American adults from the National Health and Nutrition Examination Survey 2003 to 2016.
[DOC File , 281 KB - publichealth_v9i1e49652_app2.doc]

References

1. Corbasson I, Hankinson SE, Stanek EJ, Reeves KW. Urinary bisphenol-A, phthalate metabolites and body composition in US adults, NHANES 1999-2006. Int J Environ Health Res 2016 Sep 19;26(5-6):606-617. [doi: 10.1080/09603123.2016.1233524] [Medline: 27643383]
2. Carlsson A, Sørensen K, Andersson AM, Frederiksen H, Juul A. Bisphenol A, phthalate metabolites and glucose homeostasis in healthy normal-weight children. Endocr Connect 2018 Jan;7(1):232-238 [FREE Full text] [doi: 10.1530/EC-17-0344] [Medline: 29237763]

3. Yang J, Link C, Henderson YO, Bithi N, Hine C. Peripubertal bisphenol a exposure imparts detrimental age-related changes in body composition, cognition, and hydrogen sulfide production capacities. Antioxid Redox Signal 2021 Oct 05;36(16-18):1246-1267 [FREE Full text] [doi: 10.1089/ars.2020.8226] [Medline: 34314248]
4. Liao C, Liu F, Moon HB, Yamashita N, Yun S, Kannan K. Bisphenol analogues in sediments from industrialized areas in the United States, Japan, and Korea: spatial and temporal distributions. Environ Sci Technol 2012 Nov 06;46(21):11558-11565. [doi: 10.1021/es303191g] [Medline: 23020513]
5. Liu Y, Tang W, Ao J, Zhang J, Feng L. Transcriptomics integrated with metabolomics reveals the effect of Bisphenol F (BPF) exposure on intestinal inflammation. Sci Total Environ 2022 Apr 10;816:151644. [doi: 10.1016/j.scitotenv.2021.151644] [Medline: 34774955]
6. Ellis KJ, Abrams SA, Wong WW. Monitoring childhood obesity: assessment of the weight/height index. Am J Epidemiol 1999 Nov 01;150(9):939-946. [doi: 10.1093/oxfordjournals.aje.a010102] [Medline: 10547139]
7. Dallal CM, Brinton LA, Bauer DC, Buist DS, Cauley JA, Hue TF, B~FIT Research Group. Obesity-related hormones and endometrial cancer among postmenopausal women: a nested case-control study within the B~FIT cohort. Endocr Relat Cancer 2013 Feb 18;20(1):151-160 [FREE Full text] [doi: 10.1530/ERC-12-0229] [Medline: 23222000]
8. Rohan TE, Heo M, Choi L, Datta M, Freudenheim JL, Kamensky V, et al. Body fat and breast cancer risk in postmenopausal women: a longitudinal study. J Cancer Epidemiol 2013;2013:754815-754813 [FREE Full text] [doi: 10.1155/2013/754815] [Medline: 23690776]
9. Liu PY, Hornbuckle LM, Panton LB, Kim J, Ilich JZ. Evidence for the association between abdominal fat and cardiovascular risk factors in overweight and obese African American women. J Am Coll Nutr 2012 Apr;31(2):126-132. [doi: 10.1080/07315724.2012.10720018] [Medline: 22855918]
10. Mason C, Craig CL, Katzmarzyk PT. Influence of central and extremity circumferences on all-cause mortality in men and women. Obesity (Silver Spring) 2008 Dec;16(12):2690-2695 [FREE Full text] [doi: 10.1038/oby.2008.438] [Medline: 18927548]
11. Carey VJ, Walters EE, Colditz GA, Solomon CG, Willett WC, Rosner BA, et al. Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women. The Nurses' Health Study. Am J Epidemiol 1997 Apr 01;145(7):614-619. [doi: 10.1093/oxfordjournals.aje.a009158] [Medline: 9098178]
12. Lee M, Choh AC, Demerath EW, Towne B, Siervogel RM, Czerwinski SA. Associations between trunk, leg and total body adiposity with arterial stiffness. Am J Hypertens 2012 Oct 12;25(10):1131-1137 [FREE Full text] [doi: 10.1038/ajh.2012.92] [Medline: 22785405]
13. Fantin F, Rossi AP, Cazzadori M, Comellato G, Mazzali G, Gozzoli MP, et al. Central and peripheral fat and subclinical vascular damage in older women. Age Ageing 2013 May 28;42(3):359-365. [doi: 10.1093/ageing/aft005] [Medline: 23449599]
14. Evans W. Functional and metabolic consequences of sarcopenia. J Nutr 1997 May;127(5 Suppl):998S-1003S. [doi: 10.1093/jn/127.5.998S] [Medline: 9164283]
15. Evans WJ. Skeletal muscle loss: cachexia, sarcopenia, and inactivity. Am J Clin Nutr 2010 Apr;91(4):1123S-1127S [FREE Full text] [doi: 10.3945/ajcn.2010.28608A] [Medline: 20164314]
16. Baskin KK, Winders BR, Olson EN. Muscle as a "mediator" of systemic metabolism. Cell Metab 2015 Feb 03;21(2):237-248 [FREE Full text] [doi: 10.1016/j.cmet.2014.12.021] [Medline: 25651178]
17. Argilés JM, Campos N, Lopez-Pedrosa JM, Rueda R, Rodriguez-Mañas L. Skeletal muscle regulates metabolism via interorgan crosstalk: roles in health and disease. J Am Med Dir Assoc 2016 Sep 01;17(9):789-796 [FREE Full text] [doi: 10.1016/j.jamda.2016.04.019] [Medline: 27324808]
18. DeFronzo RA, Jacot E, Jequier E, Maeder E, Wahren J, Felber JP. The effect of insulin on the disposal of intravenous glucose. Results from indirect calorimetry and hepatic and femoral venous catheterization. Diabetes 1981 Dec;30(12):1000-1007. [doi: 10.2337/diab.30.12.1000] [Medline: 7030826]
19. Vom Saal FS, Nagel SC, Coe BL, Angle BM, Taylor JA. The estrogenic endocrine disrupting chemical bisphenol A (BPA) and obesity. Mol Cell Endocrinol 2012 May 06;354(1-2):74-84 [FREE Full text] [doi: 10.1016/j.mce.2012.01.001] [Medline: 22249005]
20. Boucher JG, Boudreau A, Atlas E. Bisphenol A induces differentiation of human preadipocytes in the absence of glucocorticoid and is inhibited by an estrogen-receptor antagonist. Nutr Diabetes 2014 Jan 13;4(1):e102 [FREE Full text] [doi: 10.1038/nutd.2013.43] [Medline: 24418828]
21. Zhao HY, Bi YF, Ma LY, Zhao L, Wang TG, Zhang LZ, et al. The effects of bisphenol A (BPA) exposure on fat mass and serum leptin concentrations have no impact on bone mineral densities in non-obese premenopausal women. Clin Biochem 2012 Dec;45(18):1602-1606. [doi: 10.1016/j.clinbiochem.2012.08.024] [Medline: 22981830]
22. Lind PM, Roos V, Rönn M, Johansson L, Ahlström H, Kullberg J, et al. Serum concentrations of phthalate metabolites are related to abdominal fat distribution two years later in elderly women. Environ Health 2012 Apr 2;11(1):21. [doi: 10.1186/1476-069x-11-21]
23. National Health and Nutrition Examination Survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2023-06-15]

24. NCHS Ethics Review Board (ERB) approval. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm [accessed 2023-06-15]
25. Zhou X, Ye X, Calafat AM. Automated on-line column-switching HPLC-MS/MS method for the quantification of triclocarban and its oxidative metabolites in human urine and serum. J Chromatogr B Analyt Technol Biomed Life Sci 2012 Jan 15;881-882:27-33. [doi: 10.1016/j.jchromb.2011.11.024] [Medline: 22192874]
26. National Health and Nutrition Examination Survey 2003-2004 data documentation, codebook, and frequencies, environmental phenols (L24EPH_C). Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/Nchs/Nhanes/2003-2004/L24EPH_C.htm#URXBPH [accessed 2023-09-12]
27. Genant HK, Engelke K, Fuerst T, Glüer CC, Grampp S, Harris ST, et al. Noninvasive assessment of bone mineral and structure: state of the art. J Bone Miner Res 1996 Jun;11(6):707-730. [doi: 10.1002/jbmr.5650110602] [Medline: 8725168]
28. Baran DT, Faulkner KG, Genant HK, Miller PD, Pacifici R. Diagnosis and management of osteoporosis: guidelines for the utilization of bone densitometry. Calcif Tissue Int 1997 Dec;61(6):433-440. [doi: 10.1007/s002239900362] [Medline: 9383266]
29. Hologic discovery model a densitometer. Hologic Inc. URL: http://www.hologic.com [accessed 2023-06-15]
30. Rose GA, Blackburn H. Cardiovascular survey methods. Monogr Ser World Health Organ 1968;56:1-188. [Medline: 4972212]
31. NHANES survey methods and analytical guidelines. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx [accessed 2023-06-15]
32. Home page. EmpowerStats. URL: https://www.empowerstats.com [accessed 2023-06-15]
33. Carwile JL, Michels KB. Urinary bisphenol A and obesity: NHANES 2003-2006. Environ Res 2011 Aug;111(6):825-830 [FREE Full text] [doi: 10.1016/j.envres.2011.05.014] [Medline: 21676388]
34. Wang T, Li M, Chen B, Xu M, Xu Y, Huang Y, et al. Urinary bisphenol A (BPA) concentration associates with obesity and insulin resistance. J Clin Endocrinol Metab 2012 Feb;97(2):E223-E227. [doi: 10.1210/jc.2011-1989] [Medline: 22090277]
35. Oppeneer SJ, Robien K. Bisphenol A exposure and associations with obesity among adults: a critical review. Public Health Nutr 2015 Jul;18(10):1847-1863 [FREE Full text] [doi: 10.1017/S1368980014002213] [Medline: 25311796]
36. Wang HX, Zhou Y, Tang CX, Wu JG, Chen Y, Jiang QW. Association between bisphenol A exposure and body mass index in Chinese school children: a cross-sectional study. Environ Health 2012 Oct 19;11(1):79. [doi: 10.1186/1476-069x-11-79]
37. Jala A, Varghese B, Dutta R, Adela R, Borkar RM. Levels of parabens and bisphenols in personal care products and urinary concentrations in Indian young adult women: implications for human exposure and health risk assessment. Chemosphere 2022 Jun;297:134028. [doi: 10.1016/j.chemosphere.2022.134028] [Medline: 35218786]
38. Braun JM, Lanphear BP, Calafat AM, Deria S, Khoury J, Howe CJ, et al. Early-life bisphenol a exposure and child body mass index: a prospective cohort study. Environ Health Perspect 2014 Nov;122(11):1239-1245 [FREE Full text] [doi: 10.1289/ehp.1408258] [Medline: 25073184]
39. Wilson NK, Chuang JC, Morgan MK, Lordo RA, Sheldon LS. An observational study of the potential exposures of preschool children to pentachlorophenol, bisphenol-A, and nonylphenol at home and daycare. Environ Res 2007 Jan;103(1):9-20. [doi: 10.1016/j.envres.2006.04.006] [Medline: 16750524]
40. Albanito L, Lappano R, Madeo A, Chimento A, Prossnitz ER, Cappello AR, et al. G-protein-coupled receptor 30 and estrogen receptor-alpha are involved in the proliferative effects induced by atrazine in ovarian cancer cells. Environ Health Perspect 2008 Dec;116(12):1648-1655 [FREE Full text] [doi: 10.1289/ehp.11297] [Medline: 19079715]
41. Shelby MD. NTP-CERHR monograph on the potential human reproductive and developmental effects of bisphenol A. NTP CERHR MON 2008 Sep(22):v, vii-vix, 1 [FREE Full text] [Medline: 19407859]
42. Lotfi K, Saneei P, Hajhashemy Z, Esmaillzadeh A. Adherence to the Mediterranean diet, five-year weight change, and risk of overweight and obesity: a systematic review and dose-response meta-analysis of prospective cohort studies. Adv Nutr 2022 Feb 01;13(1):152-166 [FREE Full text] [doi: 10.1093/advances/nmab092] [Medline: 34352891]
43. Bruci A, Tuccinardi D, Tozzi R, Balena A, Santucci S, Frontani R, et al. Very low-calorie ketogenic diet: a safe and effective tool for weight loss in patients with obesity and mild kidney failure. Nutrients 2020 Jan 27;12(2):333 [FREE Full text] [doi: 10.3390/nu12020333] [Medline: 32012661]
44. Del Bo' C, Perna S, Allehdan S, Rafique A, Saad S, AlGhareeb F, et al. Does the Mediterranean diet have any effect on lipid profile, central obesity and liver enzymes in Non-Alcoholic Fatty Liver Disease (NAFLD) Subjects? A systematic review and meta-analysis of randomized control trials. Nutrients 2023 May 09;15(10):2250 [FREE Full text] [doi: 10.3390/nu15102250] [Medline: 37242133]
45. Rai SK, Gortmaker SL, Hu FB, Kanaya AM, Kandula NR, Sun Q, et al. A South Asian Mediterranean-style diet is associated with favorable adiposity measures and lower diabetes risk: the MASALA cohort. Obesity (Silver Spring) 2023 Jun 19;31(6):1697-1706. [doi: 10.1002/oby.23759] [Medline: 37203330]
46. Barrea L, Verde L, Vetrani C, Marino F, Aprano S, Savastano S, et al. VLCKD: a real time safety study in obesity. J Transl Med 2022 Jan 08;20(1):23 [FREE Full text] [doi: 10.1186/s12967-021-03221-6] [Medline: 34998415]
47. Muscogiuri G, El Ghoch M, Colao A, Hassapidou M, Yumuk V, Busetto L, Obesity Management Task Force (OMTF) of the European Association for the Study of Obesity (EASO). European guidelines for obesity management in adults with a

very low-calorie Ketogenic diet: a systematic review and meta-analysis. Obes Facts 2021;14(2):222-245 [FREE Full text] [doi: 10.1159/000515381] [Medline: 33882506]

48. Caprio M, Infante M, Moriconi E, Armani A, Fabbri A, Mantovani G, Cardiovascular Endocrinology Club of the Italian Society of Endocrinology. Very-low-calorie ketogenic diet (VLCKD) in the management of metabolic diseases: systematic review and consensus statement from the Italian Society of Endocrinology (SIE). J Endocrinol Invest 2019 Nov;42(11):1365-1386. [doi: 10.1007/s40618-019-01061-2] [Medline: 31111407]
49. Ghayda RA, Williams PL, Chavarro JE, Ford JB, Souter I, Calafat AM, et al. Urinary bisphenol S concentrations: potential predictors of and associations with semen quality parameters among men attending a fertility center. Environ Int 2019 Oct;131:105050 [FREE Full text] [doi: 10.1016/j.envint.2019.105050] [Medline: 31376593]
50. Zhang W, Gao Y, Chen C, Li A, Ma X, Liang Y, et al. Bisphenol S promotes the Formation of Visceral Fat in mice. Environ Sci Technol Lett 2021;8(8):699-704 [FREE Full text] [doi: 10.1021/acs.estlett.1c00419]
51. Menni C, Migaud M, Glastonbury CA, Beaumont M, Nikolaou A, Small KS, et al. Metabolomic profiling to dissect the role of visceral fat in cardiometabolic health. Obesity (Silver Spring) 2016 Jun;24(6):1380-1388 [FREE Full text] [doi: 10.1002/oby.21488] [Medline: 27129722]
52. Liu B, Lehmler HJ, Sun Y, Xu G, Liu Y, Zong G, et al. Bisphenol A substitutes and obesity in US adults: analysis of a population-based, cross-sectional study. Lancet Planet Health 2017 Jun;1(3):e114-e122 [FREE Full text] [doi: 10.1016/S2542-5196(17)30049-9] [Medline: 29308453]
53. Chen H, Li J, Zhang Y, Zhang W, Li X, Tang H, et al. Bisphenol F suppresses insulin-stimulated glucose metabolism in adipocytes by inhibiting IRS-1/PI3K/AKT pathway. Ecotoxicol Environ Saf 2022 Feb;231:113201 [FREE Full text] [doi: 10.1016/j.ecoenv.2022.113201] [Medline: 35051757]
54. Robles-Aguilera V, Gálvez-Ontiveros Y, Rodrigo L, Salcedo-Bellido I, Aguilera M, Zafra-Gómez A, et al. Factors associated with exposure to dietary bisphenols in adolescents. Nutrients 2021 May 05;13(5):1553 [FREE Full text] [doi: 10.3390/nu13051553] [Medline: 34062990]
55. Mackay H, Patterson ZR, Khazall R, Patel S, Tsirlin D, Abizaid A. Organizational effects of perinatal exposure to bisphenol-A and diethylstilbestrol on arcuate nucleus circuitry controlling food intake and energy expenditure in male and female CD-1 mice. Endocrinology 2013 Apr;154(4):1465-1475. [doi: 10.1210/en.2012-2044] [Medline: 23493373]
56. Takeuchi T, Tsutsumi O, Nakamura N, Ikezuki Y, Takai Y, Yano T, et al. Gender difference in serum bisphenol A levels may be caused by liver UDP-glucuronosyltransferase activity in rats. Biochem Biophys Res Commun 2004 Dec 10;325(2):549-554. [doi: 10.1016/j.bbrc.2004.10.073] [Medline: 15530427]
57. Mauvais-Jarvis F. Estrogen and androgen receptors: regulators of fuel homeostasis and emerging targets for diabetes and obesity. Trends Endocrinol Metab 2011 Jan;22(1):24-33 [FREE Full text] [doi: 10.1016/j.tem.2010.10.002] [Medline: 21109497]
58. Rochester JR, Bolden AL. Bisphenol S and F: a systematic review and comparison of the hormonal activity of bisphenol a substitutes. Environ Health Perspect 2015 Jul;123(7):643-650 [FREE Full text] [doi: 10.1289/ehp.1408989] [Medline: 25775505]
59. Grün F, Blumberg B. Environmental obesogens: organotins and endocrine disruption via nuclear receptor signaling. Endocrinology 2006 Jun;147(6 Suppl):S50-S55. [doi: 10.1210/en.2005-1129] [Medline: 16690801]
60. Janesick A, Blumberg B. Endocrine disrupting chemicals and the developmental programming of adipogenesis and obesity. Birth Defects Res C Embryo Today 2011 Mar;93(1):34-50 [FREE Full text] [doi: 10.1002/bdrc.20197] [Medline: 21425440]
61. Barouki R, Samson M, Blanc EB, Colombo M, Zucman-Rossi J, Lazaridis KN, et al. The exposome and liver disease - how environmental factors affect liver health. J Hepatol 2023 Aug;79(2):492-505 [FREE Full text] [doi: 10.1016/j.jhep.2023.02.034] [Medline: 36889360]
62. Qian L, Chen C, Guo L, Deng J, Zhang X, Zheng J, et al. Developmental and reproductive impacts of four bisphenols in daphnia magna. Int J Mol Sci 2022 Nov 23;23(23):14561 [FREE Full text] [doi: 10.3390/ijms232314561] [Medline: 36498889]
63. Park C, Song H, Choi J, Sim S, Kojima H, Park J, et al. The mixture effects of bisphenol derivatives on estrogen receptor and androgen receptor. Environ Pollut 2020 May;260:114036 [FREE Full text] [doi: 10.1016/j.envpol.2020.114036] [Medline: 31995776]
64. Hopps E, Noto D, Caimi G, Averna MR. A novel component of the metabolic syndrome: the oxidative stress. Nutr Metab Cardiovasc Dis 2010 Jan;20(1):72-77. [doi: 10.1016/j.numecd.2009.06.002] [Medline: 19747805]
65. Jakubiak GK, Osadnik K, Lejawa M, Osadnik T, Goławski M, Lewandowski P, et al. "Obesity and insulin resistance" is the component of the metabolic syndrome most strongly associated with oxidative stress. Antioxidants (Basel) 2021 Dec 29;11(1):79 [FREE Full text] [doi: 10.3390/antiox11010079] [Medline: 35052583]
66. Huang M, Liu S, Fu L, Jiang X, Yang M. Bisphenol A and its analogues bisphenol S, bisphenol F and bisphenol AF induce oxidative stress and biomacromolecular damage in human granulosa KGN cells. Chemosphere 2020 Aug;253:126707. [doi: 10.1016/j.chemosphere.2020.126707] [Medline: 32289607]
67. Meng Z, Tian S, Yan J, Jia M, Yan S, Li R, et al. Effects of perinatal exposure to BPA, BPF and BPAF on liver function in male mouse offspring involving in oxidative damage and metabolic disorder. Environ Pollut 2019 Apr;247:935-943. [doi: 10.1016/j.envpol.2019.01.116] [Medline: 30823348]

68. Cao H, Wang F, Liang Y, Wang H, Zhang A, Song M. Experimental and computational insights on the recognition mechanism between the estrogen receptor α with bisphenol compounds. Arch Toxicol 2017 Dec;91(12):3897-3912. [doi: 10.1007/s00204-017-2011-0] [Medline: 28616630]
69. Wang YQ, Zhang HM, Cao J, Tang BP. Binding of a new bisphenol analogue, bisphenol S to bovine serum albumin and calf thymus DNA. J Photochem Photobiol B 2014 Sep 05;138:182-190. [doi: 10.1016/j.jphotobiol.2014.06.002] [Medline: 24972352]
70. Zhang T, Xue J, Gao CZ, Qiu RL, Li YX, Li X, et al. Urinary concentrations of bisphenols and their association with biomarkers of oxidative stress in people living near e-waste recycling facilities in China. Environ Sci Technol 2016 Apr 05;50(7):4045-4053. [doi: 10.1021/acs.est.6b00032] [Medline: 26974222]
71. Masuno H, Iwanami J, Kidani T, Sakayama K, Honda K. Bisphenol a accelerates terminal differentiation of 3T3-L1 cells into adipocytes through the phosphatidylinositol 3-kinase pathway. Toxicol Sci 2005 Apr;84(2):319-327. [doi: 10.1093/toxsci/kfi088] [Medline: 15659569]
72. Alharbi HF, Algonaiman R, Alduwayghiri R, Aljutaily T, Algheshairy RM, Almutairi AS, et al. Exposure to Bisphenol A substitutes, Bisphenol S and Bisphenol F, and its association with developing obesity and diabetes mellitus: a narrative review. Int J Environ Res Public Health 2022 Nov 29;19(23):15918 [FREE Full text] [doi: 10.3390/ijerph192315918] [Medline: 36497992]
73. Frampton J, Murphy KG, Frost G, Chambers ES. Higher dietary fibre intake is associated with increased skeletal muscle mass and strength in adults aged 40 years and older. J Cachexia Sarcopenia Muscle 2021 Dec;12(6):2134-2144 [FREE Full text] [doi: 10.1002/jcsm.12820] [Medline: 34585852]
74. Jing J, Pu Y, Veiga-Lopez A, Lyu L. In vitro effects of emerging bisphenols on myocyte differentiation and insulin responsiveness. Toxicol Sci 2020 Nov 01;178(1):189-200 [FREE Full text] [doi: 10.1093/toxsci/kfaa130] [Medline: 32750123]
75. Gerona R, Vom Saal FS, Hunt PA. BPA: have flawed analytical techniques compromised risk assessments? Lancet Diabetes Endocrinol 2020 Jan;8(1):11-13. [doi: 10.1016/S2213-8587(19)30381-X] [Medline: 31813841]

Abbreviations

ALM: appendicular lean mass
BMC: bone mineral content
BPA: bisphenol A
BPF: bisphenol F
BPS: bisphenol S
CAD: coronary artery disease
DXA: dual-energy X-ray absorptiometry
NHANES: National Health and Nutrition Examination Survey
TLM: total lean mass
TOF: total fat
TRF: trunk fat

Edited by A Mavragani, T Sanchez; submitted 04.06.23; peer-reviewed by LH Duntas, J Luo; comments to author 14.06.23; revised version received 21.06.23; accepted 24.08.23; published 19.09.23.

Please cite as:
Li J, Zhang Z, Zhang C, Zhu Q, Zhao J, Zong H, Deng Q, Zheng J, Wu E, Wu R, Tang T, Zhang Y, Dong Q, Li Y, Wang J, Yang L, Qiu S, Shen B, Wei Q
Association Between Urinary Bisphenols and Body Composition Among American Adults: Cross-Sectional National Health and Nutrition Examination Survey Study
JMIR Public Health Surveill 2023;9:e49652
URL: https://publichealth.jmir.org/2023/1/e49652
doi:10.2196/49652
PMID:37615638

Original Paper

Structural Equation Model for Social Support and Quality of Life Among Individuals With Mental Health Disorders During the COVID-19 Pandemic

Bach Xuan Tran[1,2*], PhD; Vu Anh Trong Dam[1*], MPH; Pascal Auquier[2], PhD; Laurent Boyer[2], PhD; Guillaume Fond[2], PhD; Hung Manh Nguyen[3], PhD; Hung Tuan Nguyen[4], PhD; Huong Thi Le[1], PhD; Ha Nhi Thi Tran[5], PhD; Giang Thu Vu[6], MA; Manh Duc Nguyen[1], BPH; Duong Anh Thi Nguyen[1], BPH; Bang Viet Ly[1], BPH; Carl A Latkin[7], PhD; Melvyn WB Zhang[8], PhD; Roger CM Ho[9,10], PhD; Cyrus SH Ho[9], PhD

[1]Institute for Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam
[2]CEReSS, Research Centre on Health Services and Quality of Life, Aix Marseille University, Marseille, France
[3]Mai Huong Daycare Psychiatric Hospital, Hanoi, Vietnam
[4]National Psychiatric Hospital No 1, Hanoi, Vietnam
[5]Hanoi Department of Health, Hanoi, Vietnam
[6]School of Psychology, The University of Queensland, Brisbane, Australia
[7]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
[8]Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, Singapore
[9]Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
[10]Institute for Health Innovation and Technology, National University of Singapore, Singapore, Singapore
[*]these authors contributed equally

Corresponding Author:
Bach Xuan Tran, PhD
Institute for Preventive Medicine and Public Health
Hanoi Medical University
1 Ton That Tung Street, Dong Da district
Hanoi, 100000
Vietnam
Phone: 84 888288770
Email: bach.ipmph2@gmail.com

Abstract

Background: In light of the COVID-19 pandemic, the distribution of social support for mental health problems has likely become unequal. Family- and community-based social support has been recognized as a promising approach for mental disorders; however, limited global frameworks have been applied to developing countries such as Vietnam.

Objective: The aim of this study was to evaluate the quality of life and social support among patients with mental health disorders in Vietnam and to investigate the factors associated with quality of life among these patients.

Methods: A cross-sectional study was conducted on 222 psychiatric patients in Hanoi from 2020 to 2022. A structured questionnaire was developed based on four standardized scales: Mental Well-Being-5 scale, Multidimensional Scale of Perceived Social Support, EuroQoL-visual analog scale (EQ-VAS), and EuroQoL-5 dimensions-5 levels (EQ-5D-5L) scale. Tobit regression was used to identify factors associated with the EQ-5D-5L and EQ-VAS scores. Structural equation modeling was applied to verify the relationship between quality of life and social support.

Results: The results showed that perceived support from family scored the highest compared to support from friends and significant others. Patients with depression reported the lowest quality of life and perceived social support. Structural equation modeling showed a root mean square error of approximation of 0.055 (90% CI 0.006-0.090), comparative fit index of 0.954, Tucker-Lewis index of 0.892, and standardized root mean squared error of 0.036 (*P*<.001). The hypothetical model indicated statistically significant correlations between EQ-VAS score and social support (*P*=.004), EQ-5D-5L and mental well-being (*P*<.001), and social support and mental well-being (*P*<.001). Critical deterioration of quality of life and inconsistency in social support for patients with mental illness were also recorded.

Conclusions: There is a need to enhance social support and service delivery in Vietnam, focusing on occupation and quality of life. The correlations between social support, quality of life, and mental health issues suggest the potential of a clinical-social integrated intervention model of care.

(JMIR Public Health Surveill 2023;9:e47239) doi:10.2196/47239

KEYWORDS

contextual; social support; quality of life; mental health disorders; Vietnam; mental health; mental illness; cross-sectional; association

Introduction

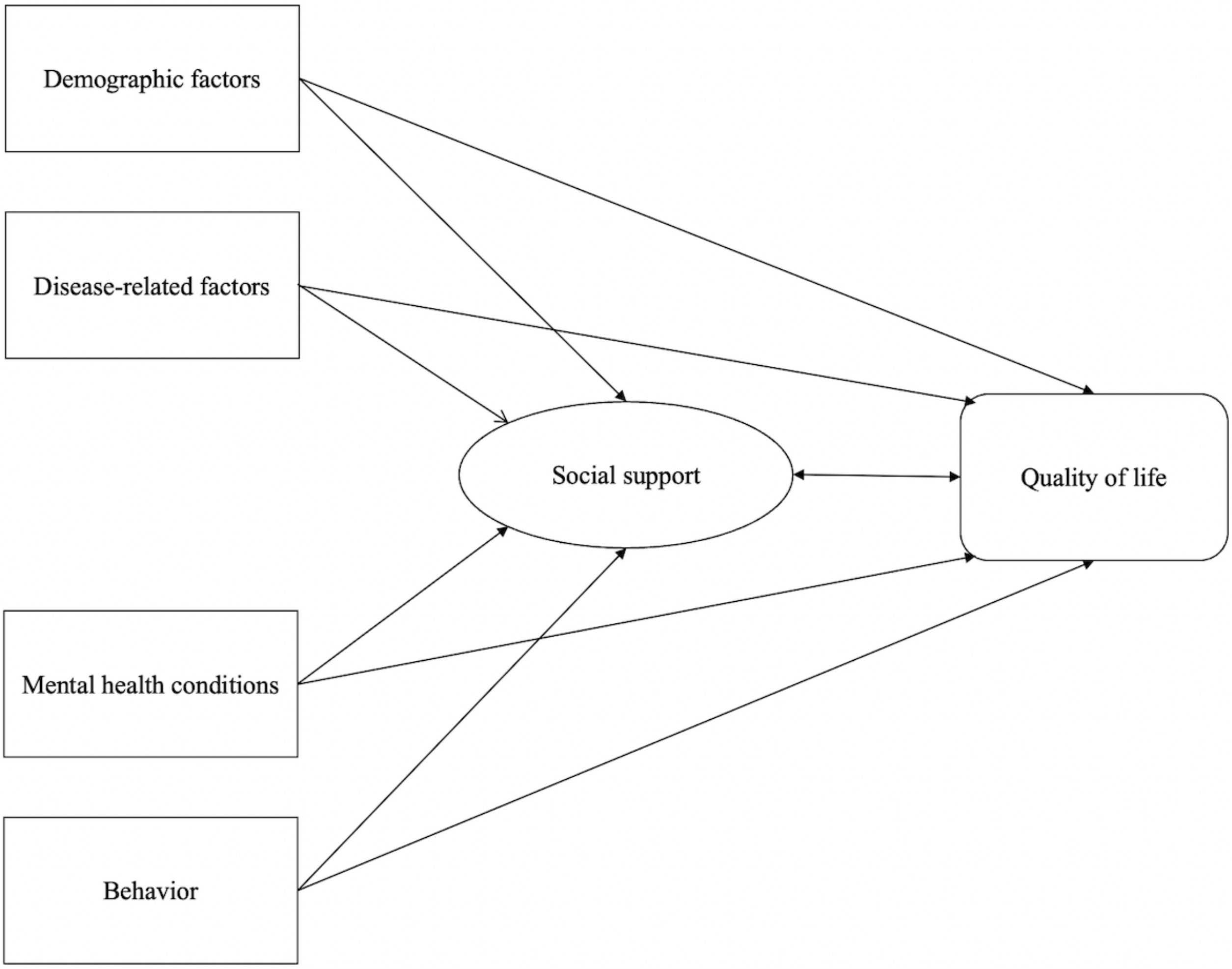
Awareness of the burden posed by mental health disorders is steadily growing on a global scale, particularly in low- and middle-income countries [1]. According to a report by the World Health Organization (WHO), mental health disorders affected approximately 970 million people worldwide, constituting 12.8% of the population. These disorders are measured in terms of disability-adjusted life years (DALYs) and years lived with disability (YLDs), with mental health disorders accounting for over 40% of YLDs [2] and 15% of DALYs [3] globally. The prevalence of mental health disorders saw a sharp increase in 2020 due to the COVID-19 pandemic. Depressive symptoms were reported in 48.3% of the Chinese population, 32.7% of the Italian population, and 23.6% of the Spanish population [4]. In Vietnam, 45.8% of the population reported mild to moderate symptoms of stress following an extended period of social restrictions [5]. COVID-19 significantly impacted patients with pre-existing mental health disorders [6,7]. Strict social restrictions led to heightened stress levels and feelings of isolation among hospitalized patients [8]. Patients with mental illness were profoundly affected by changes in daily routines, social rhythms, and their sense of stability [9]. Additionally, socioeconomic factors such as remote working, wage cuts, and increased household expenditures were linked to the worsening of pre-existing disorders [10]. Given that mental health outcomes are influenced by a variety of biological, social, cultural, economic, and religious factors, it is crucial to approach mental health from a multidimensional perspective. The COVID-19 pandemic emphasizes the urgent need to shift the approach to mental health from a biomedical and therapeutic standpoint to a psychosocial perspective that maximizes family- and community-based interventions.

An increasing body of research has been dedicated to examining the influence of social determinants on health care outcomes. The WHO has officially acknowledged the significant impact of social factors on the prevalence and duration of mental disorders, attributing them to be responsible for shaping, sustaining, and enhancing overall health quality [11,12]. The UK National Health Service has emphasized the vital role of community connectors and advocated for an integrated approach that blends clinical treatment, mental health care, and social support [13]. Both frameworks acknowledge social support as a potent public health intervention, encompassing multitiered collaborations, including personal-interpersonal, personal-community, and interpersonal-community interactions. Furthermore, social support can function as a consistent source of care and a therapeutic approach for individuals with mental illnesses, without relying heavily on facility resources as is the case for many clinical interventions. Nevertheless, standardized models encounter the challenge of not fully capturing the influence of social support on diverse determinants of health, such as societal, cultural, and religious factors. Furthermore, since current frameworks are primarily rooted in Western contexts, there exists limited evidence regarding the effectiveness of the social support model in developing countries and resource-constrained settings [14-16].

As COVID-19 presented unique challenges to individuals with mental health disorders, it is critical to understand the role of social support in their quality of life (QoL). The standardized social support frameworks have been found to be limited in their applicability to the context of resource-scarce settings. While previous studies have identified unfavorable social outcomes for patients with mental health disorders, including social dysfunction, lack of social networks, severe stigma, and interpersonal challenges, contextualized social support models have neither been implemented nor studied for this population in Vietnam [17,18]. To contribute to the growing body of research on mental health care and social support in Vietnam, shedding light on the unique challenges faced by this population and identifying potential areas for intervention and improvement, we conducted this study with two main objectives: (1) to measure social support and QoL among patients with mental health disorders in Vietnam, and (2) develop and verify a structural model linking various factors that influence the QoL of patients with mental health disorders, particularly the social support aspect. Furthermore, we sought to confirm the bidirectional association between social support and QoL, where patients with mental health disorders suffer from low QoL, which in turn leads to a lack of social support; conversely, a lack of social support is identified as an underlying cause of mental disorders, resulting in lower QoL. Particularly in this time of COVID-19, where the pandemic has even further magnified the challenges faced by individuals with mental health disorders, such information can be invaluable in the development of tailored interventions that can strengthen social support and improve the QoL of patients.

After reviewing the existing literature, we identified demographic, behavioral, illness-related, mental health–related, and social support factors that contribute to influencing the QoL of individuals with mental health conditions [19-21]. The conceptual framework employed in this study is presented in Figure 1.

Figure 1. Conceptual framework of the study.



Methods

Study Setting and Participants

From January 2020 to June 2022, a cross-sectional study was conducted in Hanoi, Vietnam. The eligibility criteria included patients who were 18 years old or above, currently being treated at the Mai Huong Daycare Psychiatric, willing to participate in the study through providing informed consent, and had the physical and mental capacity to answer the interviewer's questions. Patients who had serious cognitive impairment or were unable to answer the data collectors' questions were excluded from the recruitment process. A day psychiatric hospital model is currently being applied by Mai Huong Daycare Psychiatric, making the participants outpatients. The day-hospital model allows patients to spend less time at the hospital, reducing the potential negative effects of prolonged hospital stays, while facilitating faster recovery and improved QoL [22].

There were five investigators who were members of the research team and well-trained medical staff from Hanoi Medical University. The interview was conducted in a separate room where patient privacy was guaranteed. The patient was interviewed in a face-to-face manner using the questionnaire. The questionnaire required approximately 10-15 minutes to complete.

Measure and Instruments

A structured questionnaire was used consisting of five main components: (1) demographic characteristics, (2) clinical history (chronic diseases, family psychiatric history, number of outpatient medical examinations, and number of mental health disorders examinations), (3) mental well-being (Mental Health Inventory-5 [MHI-5] scale), (4) social support (Multidimensional Scale of Perceived Social Support [MSPSS]), and (5) QoL (EuroQoL-5 dimensions-5 levels [EQ-5D-5L] scale and EuroQoL-visual analog scale [EQ-VAS]).

Variables

Outcome Variables

The EQ-5D-5L scale was used to measure the QoL of participants through five domains: Mobility, Self-care, Usual Activities, Pain/Discomfort, and Anxiety/Depression. Each question had a score from 1 ("extreme problems") to 5 ("no problems"), which resulted in 3125 possible health states from 1111 (worst health) to 55,555 (full health) [23]. One single "utility" score was defined for each health state and could be transformed using the interim scoring for EQ-5D-5L. This study used the validated Vietnamese version of the EQ-5D-5L with a score ranging from –0.5115 to 1 [24]. Higher scores indicated better QoL. The EQ-VAS is a vertical visual analog scale to self-assess health states from a score of 100 (best imaginable health) to 0 (worst imaginable health) [25].

The MSPSS was used with 12 items to measure the social support from the following subscales: family (4 items), friends (4 items), and significant others (4 items) [26]. Participants reported the MSPSS using a 7-point Likert scale ranging from 1 "strongly disagree" to 7 "strongly agree." The score of subscales is the average score of the corresponding 4 items. Higher scores indicated higher support from family, friends, or significant others [26]. The maximum total score was 84, with a higher score indicating greater perceived social support. The MSPSS has been used widely for several groups such as those with infertility, people living with HIV/AIDS, and adolescents [27-29].

Covariates

Socioeconomic Status

Respondents were asked about their sociodemographic background, including age, gender, marital status, and occupation.

Substance Use Behaviors

We asked participants one question each about their current smoking and alcohol consumption behavior.

Clinical Characteristics

Participants were asked about their clinical characteristics, including comorbidities, chronic diseases, acute symptoms, family history of mood disorders, type of disorders (schizophrenia, bipolar disorder, depression, anxiety disorders, mixed anxiety-depressive disorder, sleep disorders, emotional disturbance, epilepsy, and other), family history of mental health disorders, the number of outpatient medical examinations, and the number of mental health disorders examinations.

The MHI-5 scale was used to measure the general mental health status of participants. There was a total of five items with each item scored from 1 to 5. The total score was measured by adding the total score of 5 questions and multiplying by 4. The total score ranged from 20 to 100 [30]. Higher scores indicated more severe mental health problems. The MHI-5 scale has been used for several groups such as adolescents and the general population [31,32].

Sample Size and Sampling Method

Total population sampling was used in this study. The sample size was calculated using the formula to estimate the mean score of EQ-VAS among participants according to EQ-VAS scores reported from another study conducted in Vietnam [33]. This was selected for two main reasons. First, the study was conducted on adult patients quite similar to our study group. In Vietnam, there is still very little research on patients with mental illness. Second, the previous study used the same tool to measure the QoL of patients (EQ-VAS). Hence, we used α=.05, the mean score of EQ-VAS from the previous study of 66.3 (SD 12.5) [33], and a relative error of 0.025 to calculate the required sample size, resulting in a sample size of 219 patients. It was expected that 5% of those recruited would refuse to participate; therefore, we set the final sample size to N=230. At the end of data collection, 222 participants were recruited for the study.

Statistical Analysis

The data analysis was conducted using Stata version 16 and missing data were managed using the listwise deletion method. Descriptive statistics are presented as the mean (SD) for continuous variables and frequencies and proportions for categorical variables. The Wilcoxon-Mann-Whitney test and Kruskal-Wallis test were used to test for differences in perceived social support and QoL scores across individual characteristics.

Individual characteristics, clinical history, mental well-being, and social support were considered potential covariates for the full models' QoL. Multivariate Tobit regression was utilized to identify factors related to the EQ-VAS and EQ-5D-5L scores. Stepwise forward techniques were employed to create reduced models, with a significance level of P<.20 as the inclusion threshold. Statistical significance was set at P<.05. In contrast, the inferential statistics applied in this study utilized structural equation modeling, which involved employing maximum-likelihood estimation and path analysis to establish the implicit connection between the variables as originally hypothesized in Figure 1. Goodness-of-fit tests such as the goodness of fit index, comparative fit index (CFI), root mean square error approximation (RMSEA), Tucker-Lewis index (TLI), and standard root mean squared residual (SRMR) were used to evaluate the models.

Ethical Considerations

This analysis used a part of the data set of a longitudinal study on individuals with and without mental health disorders in Vietnam, Vietnam's Brain and Behavior Cohort, led by authors BXT and RCMH. The study monitors changes in functional near-infrared data and self-reported health outcomes using portable, functional near-infrared spectroscopy and clinical-behavioral assessments among Vietnamese patients, in an effort toward developing an artificial intelligence–based diagnosis system for psychiatric disorders in Vietnam.

The protocol of this study was approved by the Institutional Review Board of Hanoi Medical University (58/GCN-HĐĐNCYSH-ĐHYHN). We confirmed that all procedures were performed in accordance with relevant guidelines and regulations. Written informed consent was required from participants before participating in the study. Patients invited to participate were fully explained the content, purpose, and benefits of the study. The collected information was kept confidential and only used for research purposes. The data were encrypted to ensure confidentiality of the information.

Results

Table 1 provides demographic information of the respondents, with a mean age of 43.1 (SD 13.4) years and a majority being men, single, having a high school education, and working as freelancers. The prevalence of chronic diseases was moderate (45%), with schizophrenia and bipolar affective disorder being the most commonly reported mental disorders. Participants reported an average of 8.9 outpatient medical examinations and 8.2 mental health disorder examinations in the past year.

Table 1. Individual characteristics of participants (N=222).

Characteristics	Value
Gender, n (%)	
Man	122 (55.0)
Woman	100 (45.0)
Marital status, n (%)	
Single	106 (47.8)
Married	96 (43.2)
Divorced/separated/widow	20 (9.0)
Education, n (%)	
Below high school	75 (33.8)
High school	78 (35.1)
University	69 (31.1)
Occupation, n (%)	
Unemployed	50 (22.5)
White-collar worker	18 (8.1)
Blue-collar worker	22 (9.9)
Student	28 (12.6)
Retired	20 (9.0)
Freelancer	84 (37.8)
Chronic diseases, n (%)	
None	122 (55.0)
One disease	62 (27.9)
Two or more diseases	38 (17.1)
Family history of mental health disorders, n (%)	28 (12.7)
Drinking alcohol, n (%)	35 (15.8)
Smoking, n (%)	48 (21.6)
Type of mental health disorder, n (%)	
Other	45 (20.5)
Sleep disorders	27 (12.3)
Schizophrenia	44 (20.0)
Schizoaffective disorder	15 (6.8)
Anxiety disorders	10 (4.6)
Bipolar affective disorder	29 (13.2)
Mixed anxiety disorder-depression	18 (8.2)
Emotional disturbance	7 (3.2)
Depression	14 (6.4)
Psychosis	11 (5.0)
Age (years), mean (SD)	43.1 (15.4)
Number of outpatient medical examinations, mean (SD)	8.9 (11.6)
Number of mental health disorders examinations, mean (SD)	8.2 (10.9)

Table 2 presents the perceived social support scores, with mean scores of 5.08 (SD 1.09) for the family group, 4.32 (SD 1.52) for the friend group, 4.18 (SD 1.68) for significant others, and 54.33 (SD 13.52) for total perceived social support. Married participants reported higher scores for significant other support and total perceived social support than single participants

(*P*=.03). Unemployed participants had the lowest perceived social support total score, while blue-collar workers had the highest perceived social support total score.

Table 2. Perceived social support regarding individual characteristics (N=222).

Characteristics	Family group support		Friend group support		Significant other support		Total perceived support	
	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value
Total	5.08 (1.09)		4.32 (1.52)		4.18 (1.68)		54.33 (13.52)	
Gender		.64		.96		.45		.53
Man	5.12 (1.09)		4.34 (1.52)		4.10 (1.70)		54.27 (13.74)	
Woman	5.03 (1.10)		4.30 (1.52)		4.27 (1.67)		54.40 (13.31)	
Marital status		.06		.87		.03		.03
Single	4.95 (1.23)		4.32 (1.63)		4.01 (1.70)		53.11 (14.76)	
Married	5.28 (0.91)		4.33 (1.46)		4.47 (1.62)		56.32 (12.36)	
Other	4.83 (1.01)		4.31 (1.27)		3.66 (1.72)		51.20 (10.96)	
Education		.47		.90		.53		.89
Below high school	5.15 (1.08)		4.27 (1.53)		4.33 (1.66)		54.99 (12.76)	
High school	5.00 (1.01)		4.35 (1.49)		4.10 (1.48)		53.77 (12.71)	
University	5.10 (1.20)		4.36 (1.56)		4.11 (1.92)		54.25 (15.27)	
Occupation		.41		.28		.35		.47
Unemployed	5.08 (1.21)		3.99 (1.75)		3.75 (1.82)		51.22 (15.59)	
White-collar worker	5.29 (1.19)		4.51 (1.18)		4.06 (2.04)		55.44 (13.78)	
Blue-collar worker	5.16 (1.10)		4.35 (1.83)		4.47 (1.78)		55.91 (15.86)	
Student	4.65 (1.10)		4.32 (1.45)		4.03 (1.36)		52.00 (10.36)	
Retired	5.19 (0.93)		3.88 (1.58)		4.43 (1.82)		53.95 (12.24)	
Freelancer	5.13 (1.02)		4.58 (1.32)		4.38 (1.54)		56.39 (12.59)	
Chronic diseases		.27		.37		.44		.39
None	5.18 (1.07)		4.38 (1.54)		4.26 (1.67)		55.25 (13.19)	
One disease	4.85 (1.21)		4.22 (1.46)		3.94 (1.69)		52.08 (14.53)	
Two or more diseases	5.14 (0.93)		4.32 (1.59)		4.31 (1.73)		55.05 (12.76)	
Family history of mental health disorders		.45		.69		.62		.65
No	5.11 (1.10)		4.32 (1.55)		4.15 (1.72)		54.33 (13.50)	
Yes	4.89 (1.09)		4.31 (1.39)		4.39 (1.49)		54.39 (14.12)	
Drinking alcohol		.16		.60		.43		.32
No	5.12 (1.08)		4.34 (1.53)		4.21 (1.68)		54.64 (13.46)	
Yes	4.89 (1.17)		4.25 (1.48)		4.04 (1.72)		52.69 (13.92)	
Smoking		.69		.50		.23		.34
No	5.05 (1.12)		4.37 (1.48)		4.25 (1.65)		54.72 (13.33)	
Yes	5.18 (1.01)		4.14 (1.67)		3.91 (1.79)		52.92 (14.22)	

Table 3 shows the QoL scores of the respondents, with mean scores of 67.99 (SD 20.51) for the EQ-VAS and 0.88 (SD 0.17) for EQ-5D-5L. Women had lower EQ-VAS scores than men (*P*=.01). Patients who were white-collar workers had the highest EQ-VAS and EQ-5D-5L scores, retired patients had the lowest EQ-VAS scores, and students had the lowest EQ-5D-5L scores. No significant differences were found between groups in EQ-5D-5L scores.

Table 3. Quality of life–related scores regarding individual characteristics (N=222).

Characteristics	EQ-VAS[a]		EQ-5D-5L[b] index	
	Mean (SD)	*P* value	Mean (SD)	*P* value
Total	67.99 (20.51)		0.88 (0.17)	
Gender		.01		.09
Man	71.80 (18.31)		0.89 (0.18)	
Woman	63.34 (22.14)		0.86 (0.16)	
Marital status		.96		.79
Single	68.08 (22.22)		0.88 (0.16)	
Married	68.14 (18.75)		0.87 (0.20)	
Other	66.75 (20.15)		0.88 (0.14)	
Education		.41		.47
Below high school	66.13 (21.31)		0.86 (0.22)	
High school	68.05 (18.09)		0.87 (0.16)	
University	69.93 (22.26)		0.90 (0.12)	
Occupation		.01		.09
Unemployed	67.04 (20.59)		0.86 (0.18)	
White-collar worker	78.83 (15.88)		0.90 (0.14)	
Blue-collar worker	67.41 (23.16)		0.86 (0.11)	
Student	65.04 (24.49)		0.84 (0.19)	
Retired	57.75 (17.51)		0.86 (0.14)	
Freelancer	69.80 (18.98)		0.90 (0.19)	
Chronic diseases		.14		.32
None	70.02 (20.30)		0.88 (0.20)	
One disease	66.71 (21.67)		0.86 (0.16)	
Two or more than two diseases	63.53 (18.83)		0.89 (0.11)	
Family history of mental health disorders		.76		.11
No	68.02 (20.97)		0.88 (0.18)	
Yes	68.04 (17.76)		0.87 (0.11)	
Drinking alcohol		.15		.31
No	68.87 (20.44)		0.88 (0.17)	
Yes	63.29 (20.58)		0.85 (0.16)	
Smoking		.51		.47
No	67.26 (21.44)		0.87 (0.18)	
Yes	70.60 (16.65)		0.89 (0.16)	

[a] EQ-VAS: EuroQoL-visual analog scale.

[b] EQ-5D-5L: EuroQoL-5 dimensions-5 levels.

Table 4 shows the characteristics of QoL and perceived social support based on the type of mental health disorder. Patients with psychosis had the highest EQ-5D-5L score, followed by schizophrenia and emotional disturbance (*P*=.03). Patients with sleep disorders had the highest total perceived social support score, followed by those with anxiety disorders (*P*=.01). In addition, patients with anxiety disorders had the highest family support scores (*P*=.06), whereas patients with sleep disorders had the highest support scores for the friend group (*P*=.04) and significant others group (*P*=.02).

Table 4. Characteristics of quality of life and perceived social support regarding the type of mental health disorders (N=222).

Mental health disorder	Quality of life, mean (SD)		Perceived social support, mean (SD)			
	EQ-VAS[a]	EQ-5D-5L[b]	Family group	Friend group	Other groups	Total
Total	67.90 (20.59)	0.88 (0.17)	5.07 (1.09)	4.31 (1.52)	4.16 (1.68)	54.18 (13.44)
Other	69.36 (20.07)	0.90 (0.12)	5.31 (1.04)	4.28 (1.56)	3.96 (1.87)	54.18 (10.16)
Sleep disorders	66.26 (21.79)	0.81 (0.30)	5.46 (0.88)	5.06 (1.03)	4.94 (0.86)	61.85 (9.45)
Schizophrenia	71.59 (18.94)	0.93 (0.10)	4.86 (0.97)	3.81 (1.58)	3.64 (1.62)	49.23 (13.97)
Schizoaffective disorder	65.67 (25.97)	0.80 (0.28)	5.18 (0.97)	4.07 (1.64)	4.00 (1.70)	53.00 (14.93)
Anxiety disorders	71.50 (17.00)	0.86 (0.13)	5.48 (0.89)	5.03 (1.42)	4.48 (1.96)	59.90 (14.88)
Bipolar affective disorder	68.07 (22.76)	0.86 (0.18)	4.80 (1.45)	4.40 (1.38)	4.54 (1.60)	54.97 (15.12)
Mixed anxiety disorder-depression	63.89 (18.44)	0.85 (0.13)	4.71 (1.04)	4.42 (1.60)	4.60 (1.78)	54.89 (14.00)
Emotional disturbance	72.71 (19.53)	0.91 (0.13)	4.86 (0.99)	4.93 (0.66)	4.39 (1.27)	56.71 (5.99)
Depression	57.14 (21.99)	0.83 (0.14)	4.77 (1.40)	3.50 (1.60)	3.14 (1.86)	45.64 (15.28)
Psychosis	67.73 (18.35)	0.96 (0.07)	5.32 (0.90)	4.48 (1.65)	4.59 (1.40)	57.55 (14.11)

[a] EQ-VAS: EuroQoL-visual analog scale.

[b] EQ-5D-5L: EuroQoL-5 dimensions-5 levels.

Table 5 displays the factors related to the QoL of participants. Older age, being a woman, being married, and drinking alcohol were negative factors that decreased the QoL score. Greater mental well-being was a positive factor that increased both the EQ-VAS and EQ-5D-5L scores. Patients who had higher perceived social support had a higher EQ-VAS score. The full multivariate tobit regression models to identify factors associated with the QoL of participants are presented in Multimedia Appendix 1.

Table 5. Factors related to the quality of life of participants (N=222).

Factors	EQ-VAS[a] score, coefficient (95% CI)	EQ-5D-5L[b] index, coefficient (95% CI)
Socioeconomic factors		
Age (years)	–0.23 (–0.39 to –0.06)	N/A[c]
Gender (woman vs man=reference)	–10.00 (–15.10 to –4.90)	N/A
Marital status (vs single=reference)		
Married	N/A	–0.10 (–0.17 to –0.03)
Other	N/A	0.04 (–0.08 to 0.16)
Number of outpatient medical examinations (times)	–0.50 (–1.01 to 0.01)	N/A
Number of mental health disorders examinations (times)	0.41 (–0.13 to 0.96)	N/A
Drinking alcohol (yes vs no=reference)	–7.57 (–14.51 to –0.63)	N/A
Mental health		
MHI-5[d] score	0.34 (0.23 to 0.45)	0.005 (0.004 to 0.007)
Perceived social support score	0.31 (0.11 to 0.50)	N/A

[a] EQ-VAS: EuroQoL-visual analog scale.

[b] EQ-5D-5L: EuroQoL-5 dimensions-5 levels.

[c] N/A: not applicable.

[d] MHI-5: Mental Health Inventory-5.

Figure 2 presents the results of the structural model and standardized path coefficients of the hypothetical model. The analysis of the structural equation model using the study variables in the hypothetical model resulted in an RMSEA of 0.055 (90% CI 0.006-0.090), CFI of 0.954, TLI of 0.892, and SRMR of 0.036 (*P*<.001). The structural equation model showed that mental well-being was significantly related (*P*<.001) to the QoL score (coefficient=0.37, 95% CI 0.27-0.47 for EQ-VAS; coefficient=0.003, 95% CI 0.002-0.004 for EQ-5D-5L) and with social support (coefficient=0.01, 95% CI 0.004-0.013). Additionally, the model revealed a bidirectional relationship between social support and EQ-VAS (coefficient=2.27, 95% CI 0.74-3.81; *P*=.004). Table 6 illustrates the full models, including the outcome and both tier-1 and tier-2 predictors.

Figure 2. Structural model and standardized path coefficients (N=222). EQ-5D-5L: EuroQol-5 dimensions-5 levels; EQ-VAS: EuroQoL-visual analog scale.

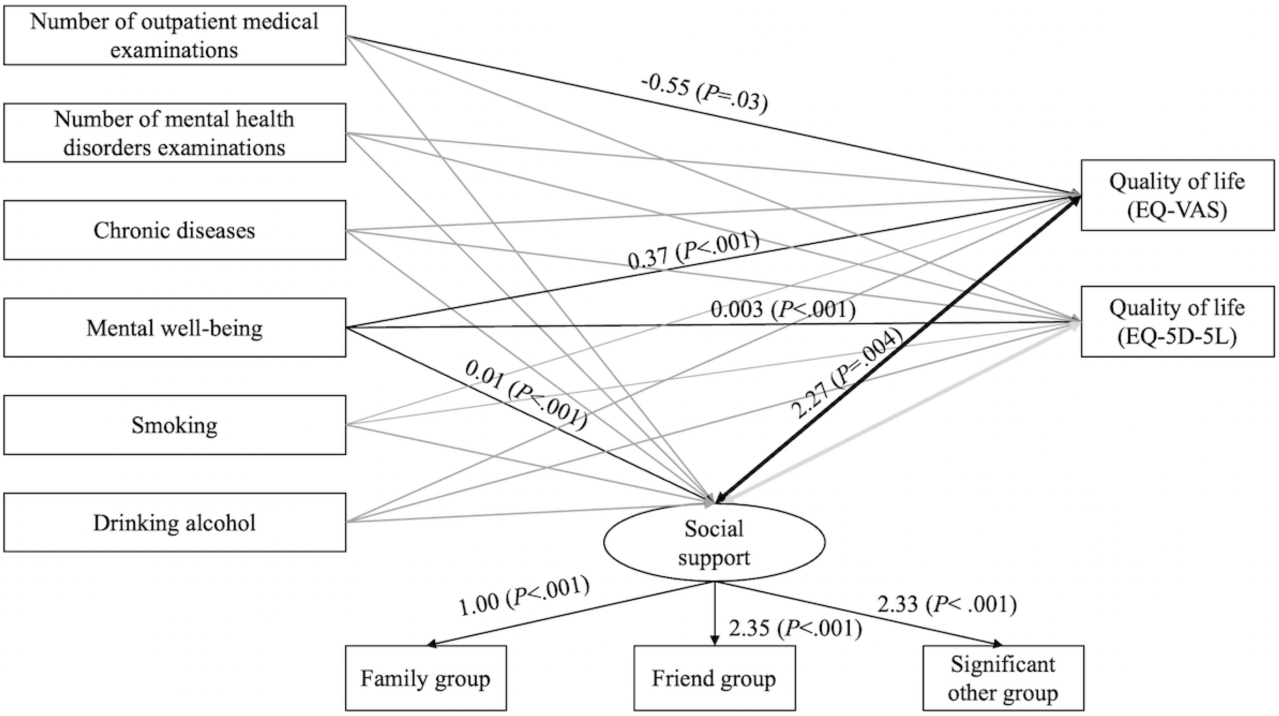


Table 6. Full models, including the outcome and both tier-1 and tier-2 predictors.

Structural component	Standardized path coefficient	*P* value of unstandardized estimate
Direct paths: tier 2 to tier 1		
Number of outpatient medical examinations to Social support	–0.014	.11
Number of mental health disorders examination to Social support	0.011	.21
Chronic diseases to Social support	–0.013	.80
Mental well-being to Social support	0.009	<.001
Smoking to Social support	–0.136	.18
Drinking alcohol to Social support	0.056	.63
Direct paths: tier 2 to the outcome		
Number of outpatient medical examinations to EQ-VAS[a]	–0.554	.03
Number of mental health disorders examination to EQ-VAS	0.413	.14
Chronic diseases to EQ-VAS	–3.172	.05
Mental well-being to EQ-VAS	0.371	<.001
Smoking to EQ-VAS	4.312	.18
Drinking alcohol to EQ-VAS	–5.620	.12
Number of outpatient medical examinations to EQ-5D-5L[b]	–0.002	.50
Number of mental health disorders examinations to EQ-5D-5L	0.001	.59
Chronic diseases to EQ-5D-5L	0.002	.91
Mental well-being to EQ-5D-5L	0.003	<.001
Smoking to EQ-5D-5L	0.018	.52
Drinking alcohol to EQ-5D-5L	–0.022	.50
Measurements		
Social support to family group	1.000	Constrained
Social support to friend group	2.352	<.001
Social support to significant others groups	3.056	<.001
Bidirectional relationships		
Social support and EQ-VAS	2.274	.004
Social support and EQ-5D-5L	–0.001	.90

[a] EQ-VAS: EuroQoL-visual analog scale.

[b] EQ-5D-5L: EuroQoL-5 dimensions-5 levels.

Discussion

Principal Results

This study revealed a decline in QoL and unequal distribution of social support among patients with mental health disorders, as students experienced the lowest social support and workers perceived the highest levels of social support. Depression, sleeping disorders, and bipolar-related problems were the most significant mental health concerns recorded. The current resource allocation scheme was problematic, as all common mental health issues had critically lower levels of social support than other disorders. This study found that QoL decreased with age, among women, and for those with alcohol consumption. The most prevalent source of support across all patient groups was family support, followed by friend groups, implying a potential design of integrated community- and family-based interventions for mental health issues in resource-scarce settings.

Comparison With Prior Work

This study did not measure any COVID-19–related variables; however, since the study was conducted during the severe COVID-19 epidemic, the results can provide a comprehensive picture of mental health disorders in Vietnam during COVID-19. Overall, the QoL among patients with mental health disorders was lower than that of the general Vietnamese population [34]. Patients with mental health disorders experience impairments in daily tasks; isolation; and a decrease in self-sufficiency, self-confidence, and self-esteem.

In this study, participants only received outpatient treatments. Therefore, in addition to attending medical appointments, many patients still work and carry out daily tasks. Our study revealed that patients with mental health disorders had the lowest scores related to perceived support from friend groups. Amal et al [35] conducted a study on psychiatric patients in Egypt and reported similar findings. This finding is likely related, in part, to stigma and prejudice, which have a direct impact on the social opportunities for individuals with mental illnesses. Furthermore, the general population is unaware of the consequences of mental illness and many are often terrified by people with these health conditions. By contrast, this study revealed that the patients had the highest social support from family. This result can be explained by the fact that since our participants were outpatients, they spent most of their time with their family members. Family relationships are strong in Southeast Asia, which may be beneficial if they are used as social support rather than social coercion. Many patients with major mental illnesses either live with their families (parents, spouses, siblings, and children) or have regular contact with them. Goldberg et al [36] evaluated the impact of social networks among individuals with mental impairments in the United States, and discovered that the closest relatives were the most commonly utilized source of support. Furthermore, Brunt and Hansson [37] discovered that patients in in-patient and assisted community settings in Sweden had a large proportion of family members in their social networks.

Our study revealed inconsistent patterns in QoL and support availability among different occupations and mental disorders. Students experienced the lowest social support and EQ-5D-5L scores. Blue-collar workers received the highest levels of social support but experienced the lowest QoL compared to other occupation groups. Understandably, stable employment tends to provide certain benefits such as secure incomes, health insurance (though not yet covering mental disorders), and support systems [38]. The work characteristics of many blue-collar workers in Vietnam and other countries involve intensive and heavy manual labor tasks and appalling working environments, which could lead to a low QoL, even in the usual context that is amplified in the presence of mental disorders. Likewise, for treated and returning-to-work patients, sudden re-exposure to hectic work routines could also lead to loss of job satisfaction or even to relapse of previous conditions [39]. This study highlights the urgent need to develop workplace-based management standard practices that support and distribute the workload to help address the health status of treated and returning-to-work patients.

According to several studies, depression, a mix of anxiety-depression, and sleeping disorders are common determinants of low QoL, not only among mental health patients but across all population groups [40-42]. Patients with these conditions have been found to have the lowest QoL [40-42]. However, our study found that the level of social support provided for depression in particular was critically lower than that for other illnesses. This finding suggests that the distribution of social support is not linked to the needs of patients with mental health disorders [43]. The social support system focuses on illness severity as an indicator rather than on QoL. Although the impairment caused by most cases of depression, anxiety, and sleeping disorders is generally not as severe as that associated with other mental health conditions such as schizophrenia or psychosis, the symptomology of conditions is constantly present in daily life, hardly eradicated, and often only treated with outpatient care [43]. This means that patients diagnosed with these conditions have to bear the burden of their illnesses while maintaining normal social functioning such as working, caregiving, or housekeeping. As a result, patients with mental health disorders are subject to induced stress and feelings of inadequacy or become overwhelmed when their responsibilities are not met, which amplifies their mental health disorders and can create a vicious cycle that significantly impacts their holistic well-being. Instead of relying on illness severity to allocate social support resources, health care professionals should focus on QoL as an indicator for more effective distribution. Moreover, patients may deliberately avoid social support due to insecurity about their health conditions and fear of the social stigma associated with such conditions [44,45]. Interian et al [46] suggested that social stigma not only prohibits patients from seeking social support but also from seeking care altogether. Therefore, additional studies should be conducted on the correlation between social stigma and social support among patients with mental health conditions to better understand and promote help-seeking behavior among this vulnerable population.

Patients with mental disorders have lower QoL and lack social support. The structural equation model analysis confirmed the bidirectional relationship between social support and EQ-VAS scores. Previous research has linked the lack of social support to social isolation, loneliness, and higher risks of physical health problems [47], which can lead to the development of mental health issues. However, various studies, such as the meta-analysis performed by Harandi et al [48] or the cross-sectional study by Connell et al [49] have confirmed the relationships between social support, QoL, and mental health disorders, but only separately. The present study suggests that interventions to improve each of these domains can be achieved through the two other domains, which are rarely included in the literature. For example, strengthening family support has major impacts on reducing mental health symptoms immediately and in the long term [50,51], while social support availability and QoL should be integrated into the screening standard for early diagnosis and prevention. A longitudinal study conducted over the course of 23 years with 1-year, 4-year, 10-year, and 23-year follow-ups indicated the sustainability of family members as a source of support during treatment and recovery [52]. However, further clinical studies are needed to determine the extent to which social support and QoL can serve as indicators for diagnoses of mental health problems. Given the bidirectional association between social support and QoL and the impacts of their interplay on mental health issues, efforts should be directed to both ends to optimize treatment outcomes and prevent long-term impacts on well-being [52].

Implications

This study offers several implications for mental health treatment and interventions. First, given the outpatient nature of treatment where patients continue to work and carry out daily tasks alongside their therapy appointments, and considering the long-term and often lifelong nature of mental health treatment, it is crucial to provide stable and lifelong sources of social support such as family-based support, which has been shown to have a substantial impact on treatment outcomes [53,54]. Family-based interventions should be standardized to suit the Vietnamese context, based on successes and challenges in countries that have adopted social support using prior standardized frameworks [14,55,56]. Guidelines for family engagement should be provided for the care and service delivery of patients with mental health conditions. Additionally, constant feedback and follow-ups between clinicians and patients are essential to optimize treatment quality and measure progress. Second, a model of "work therapy" should be implemented to provide treatments as work tasks in the workplace, especially for labor workers who experience low levels of social support, to prevent detachment from the working routine and help patients quickly readapt to work after treatment. Finally, mental health treatment should be delivered through a combination of approaches, including social support, psychological therapy, family-based interventions, and work-related treatments.

Strengths and Limitations

The main strength of this study is that the majority of participants only received outpatient treatments, allowing for the investigation of variables in general settings such as daily life and the workplace.

However, our study had several limitations. First, we were unable to analyze causal relationships due to the cross-sectional study design. As a result, we cannot be certain whether the capacity of the patients was impacted by social support or other reasons. Second, self-reported data may suffer from a limitation of recall bias, and social desirability bias may cause certain answers to be underestimated or overestimated. Furthermore, because the time of data collection varied, recall bias might be a concern that leads to inaccuracies, especially for patients with mental health disorders. The population sampling technique may have limited the generalizability of the results for Vietnamese patients. Finally, our study was conducted during the COVID-19 pandemic, limiting the generalizability of the study results in a postpandemic period. Despite these limitations, the study identified significant determinants and trends, enabling a proposal of multilevel interventions for mental health issues and social support.

Conclusions

Our study sheds light on the challenges faced by individuals with mental health disorders in Vietnam, including a decrease in QoL and barriers to accessing social support. The QoL model was used to fit a structural equation model to systematically verify and analyze the relationship between QoL and other variables with social support as a mediator. We found that mental well-being, QoL, and social support are interconnected in a complementary manner, suggesting that improving one factor can positively impact the others. However, further research is needed to generalize these findings to other populations before implementing a standardized framework. Our study identified several issues within the current social support system, including inadequate support for labor workers with mental health issues and a severity-based approach to resource allocation rather than one that prioritizes QoL. Finally, our study highlights the need for a shift in the mental health intervention approach, and we propose an integrated model of family-based and therapeutic interventions to address mental health disorders.

Acknowledgments

BXT is the principal investigator of this project, which is a ministerial-level study managed by the Vietnam Ministry of Health (decision number 850/QĐ-BYT) and implemented by the Institute for Preventive Medicine and Public Health, Hanoi Medical University. This research is supported by Gia Lam Urban Development And Investment Company Limited, Vingroup, and by the Vingroup Innovation Foundation (VINIF) (grant number VINIF.2019.DA14); National University of Singapore (NUS) Department of Psychological Medicine (R-177-000-100-001/R-177-000-003-001); and NUS iHeathtech Other Operating Expenses (R-722-000-004-731). The authors would also like to thank the Vietnam Young Physician Association, Hanoi Medical University, and the research collaborators.

Data Availability

All data generated or analyzed during this study are included in this published article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full multivariate tobit regression models to identify factors associated with the quality of life of participants (N=222).
[DOCX File , 18 KB - publichealth_v9i1e47239_app1.docx]

References

1. Patel V. Mental health in low- and middle-income countries. Br Med Bull 2007;81-82:81-96. [doi: 10.1093/bmb/ldm010] [Medline: 17470476]
2. GBD 2017 Disease Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic

analysis for the Global Burden of Disease Study 2017. Lancet 2018 Dec 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]

3. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry 2016 Feb;3(2):171-178. [doi: 10.1016/S2215-0366(15)00505-2] [Medline: 26851330]
4. Xiong J, Lipsitz O, Nasri F, Lui LMW, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord 2020 Dec 01;277:55-64 [FREE Full text] [doi: 10.1016/j.jad.2020.08.001] [Medline: 32799105]
5. Bui N, Thi HV, Ngoc Suong MV, Chu D. Mental health due to COVID-19 in students in Vietnam: an initial assessment. J Loss Trauma 2022 Mar 17;28(3):276-278. [doi: 10.1080/15325024.2022.2046675]
6. Dong L, Bouey J. Public mental health crisis during COVID-19 pandemic, China. Emerg Infect Dis 2020 Jul;26(7):1616-1618. [doi: 10.3201/eid2607.200407] [Medline: 32202993]
7. Wu T, Jia X, Shi H, Niu J, Yin X, Xie J, et al. Prevalence of mental health problems during the COVID-19 pandemic: a systematic review and meta-analysis. J Affect Disord 2021 Feb 15;281:91-98 [FREE Full text] [doi: 10.1016/j.jad.2020.11.117] [Medline: 33310451]
8. Chatterjee SS, Barikar C, Mukherjee A. Impact of COVID-19 pandemic on pre-existing mental health problems. Asian J Psychiatr 2020 Jun;51:102071 [FREE Full text] [doi: 10.1016/j.ajp.2020.102071] [Medline: 32334407]
9. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
10. Zandifar A, Badrfam R. Iranian mental health during the COVID-19 epidemic. Asian J Psychiatr 2020 Jun;51:101990 [FREE Full text] [doi: 10.1016/j.ajp.2020.101990] [Medline: 32163908]
11. Wilkinson R, Marmot M. Social determinants of health: the solid facts. Geneva: World Health Organization; 2003.
12. Willems S, Van Roy K, De Maeseneer J. Educating Health Professionals to Address the Social Determinants of Health. In: A Framework for Educating Health Professionals to Address the Social Determinants of Health. Washington, DC: National Academies Press; 14, 2016. URL: https://www.ncbi.nlm.nih.gov/books/NBK395989 [doi: https://doi.org/10.17226/21923]
13. The community mental health framework for adults and older adults. NHS National Collaborating Centre for Mental Health. 2019. URL: https://www.england.nhs.uk/wp-content/uploads/2019/09/community-mental-health-framework-for-adults-and-older-adults.pdf [accessed 2023-08-25]
14. Bjørlykhaug KI, Karlsson B, Hesook SK, Kleppe LC. Social support and recovery from mental health problems: a scoping review. Nordic Social Work Res 2021 Jan 04;12(5):666-697. [doi: 10.1080/2156857x.2020.1868553]
15. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
16. Terry R, Townley G. Exploring the role of social support in promoting community integration: an integrated literature review. Am J Community Psychol 2019 Dec;64(3-4):509-527. [doi: 10.1002/ajcp.12336] [Medline: 31116874]
17. Seivewright H, Tyrer P, Johnson T. Persistent social dysfunction in anxious and depressed patients with personality disorder. Acta Psychiatr Scand 2004 Feb;109(2):104-109. [doi: 10.1046/j.1600-0447.2003.00241.x] [Medline: 14725590]
18. Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatry 2018 May 29;18(1):156 [FREE Full text] [doi: 10.1186/s12888-018-1736-5] [Medline: 29843662]
19. Huang H, Shen H, Ning K, Zhang R, Sun W, Li B, et al. Quality of life and its correlates in alcohol use disorder patients with and without depression in China. Front Psychiatry 2020;11:627338 [FREE Full text] [doi: 10.3389/fpsyt.2020.627338] [Medline: 33551885]
20. Moon JR, Cho YA, Huh J, Kang I, Kim D. Structural equation modeling of the quality of life for patients with marfan syndrome. Health Qual Life Outcomes 2016 Jun 02;14:83 [FREE Full text] [doi: 10.1186/s12955-016-0488-5] [Medline: 27249938]
21. Prochaska JJ, Das S, Young-Wolff KC. Smoking, mental illness, and public health. Annu Rev Public Health 2017 Mar 20;38:165-185 [FREE Full text] [doi: 10.1146/annurev-publhealth-031816-044618] [Medline: 27992725]
22. What is a day hospital? Mai Huong Day Psychiatric Hospital. 2017. URL: http://www.maihuong.gov.vn/vi/gioi-thieu/benh-vien-ban-ngay-la-gi.html [accessed 2023-08-25]
23. EQ-5D User Guides. EuroQol. URL: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf [accessed 2023-08-25]
24. Mai VQ, Sun S, Minh HV, Luo N, Giang KB, Lindholm L, et al. An EQ-5D-5L value set for Vietnam. Qual Life Res 2020 Jul;29(7):1923-1933 [FREE Full text] [doi: 10.1007/s11136-020-02469-7] [Medline: 32221805]
25. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. Qual Life Res 2014 Apr;23(3):977-989 [FREE Full text] [doi: 10.1007/s11136-013-0537-z] [Medline: 24081873]
26. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Person Assess 1988 Mar;52(1):30-41 [FREE Full text] [doi: 10.1207/s15327752jpa5201_2]

27. Lam TQ, Linh TT, Thuy LB. The impact of social support on infertility-related stress: a study in the Vietnamese context. J Soc Sci 2021;09(12):259-273. [doi: 10.4236/jss.2021.912017]
28. Nguyen TTP, Nguyen TT, Do HN, Vu TBT, Vu KL, Do HM, et al. Parent-child relationship quality and internet use in a developing country: adolescents' perspectives. Front Psychol 2022;13:847278 [FREE Full text] [doi: 10.3389/fpsyg.2022.847278] [Medline: 35295385]
29. Kieu PT, Vuong NL, Dung DV. Validation of Multidimensional Scale of Perceived Social Support (MSPSS) in Vietnamese among people living with HIV/AIDS. AIDS Behav 2023 Aug 09;27(8):2488-2496 [FREE Full text] [doi: 10.1007/s10461-022-03974-1] [Medline: 36622488]
30. Berwick DM, Murphy JM, Goldman PA, Ware JE, Barsky AJ, Weinstein MC. Performance of a five-item mental health screening test. Med Care 1991 Feb;29(2):169-176. [doi: 10.1097/00005650-199102000-00008] [Medline: 1994148]
31. Le HT, Lai AJX, Sun J, Hoang MT, Vu LG, Pham HQ, et al. Anxiety and depression among people under the nationwide partial lockdown in Vietnam. Front Public Health 2020 Oct 29;8:589359 [FREE Full text] [doi: 10.3389/fpubh.2020.589359] [Medline: 33194995]
32. Nguyen TTP, Nguyen TT, Dam VTA, Vu TTM, Do HT, Vu GT, et al. Mental wellbeing among urban young adults in a developing country: a latent profile analysis. Front Psychol 2022;13:834957 [FREE Full text] [doi: 10.3389/fpsyg.2022.834957] [Medline: 36118453]
33. Truong T, Kim N, Nguyen MT, Do D, Nguyen HT, Le T, et al. Quality of life and health status of hospitalized adults with congenital heart disease in Vietnam: a cross-sectional study. BMC Cardiovasc Disord 2021 May 05;21(1):229 [FREE Full text] [doi: 10.1186/s12872-021-02026-1] [Medline: 33947343]
34. Nguyen LH, Tran BX, Hoang Le QN, Tran TT, Latkin CA. Quality of life profile of general Vietnamese population using EQ-5D-5L. Health Qual Life Outcomes 2017 Oct 11;15(1):199 [FREE Full text] [doi: 10.1186/s12955-017-0771-0] [Medline: 29020996]
35. Gabal S. Relationship between social support and the quality of life among psychiatric patients. Port Said Sci J Nurs 2017 Jun 01;4(1):182-201. [doi: 10.21608/pssjn.2017.33095]
36. Goldberg RW, Rollins AL, Lehman AF. Social network correlates among people with psychiatric disabilities. Psychiatr Rehabil J 2003;26(4):393-402. [doi: 10.2975/26.2003.393.402] [Medline: 12739910]
37. Brunt D, Hansson L. The social networks of persons with severe mental illness in in-patient settings and supported community settings. J Mental Health 2009 Jul 06;11(6):611-621. [doi: 10.1080/09638230021000058175]
38. Conn C, Walford V. An introduction to health insurance for low income countries. Heart Resources. URL: https://www.heart-resources.org/wp-content/uploads/2012/10/An-introduction-to-Health-Insurance.pdf [accessed 2023-08-25]
39. Lysaght R, Fabrigar L, Larmour-Trode S, Stewart J, Friesen M. Measuring workplace social support for workers with disability. J Occup Rehabil 2012 Sep 15;22(3):376-386. [doi: 10.1007/s10926-012-9357-1] [Medline: 22350141]
40. Szentkirályi A, Madarász CZ, Novák M. Sleep disorders: impact on daytime functioning and quality of life. Expert Rev Pharmacoecon Outcomes Res 2009 Feb;9(1):49-64. [doi: 10.1586/14737167.9.1.49] [Medline: 19371179]
41. Rezaei E, Moghadam ZB, Saraylu K. Quality of life in pregnant women with sleep disorder. J Family Reprod Health 2013 Jun;7(2):87-93 [FREE Full text] [Medline: 24971108]
42. Visser MR, Smets EM. Fatigue, depression and quality of life in cancer patients: how are they related? Support Care Cancer 1998 Mar;6(2):101-108. [doi: 10.1007/s005200050142] [Medline: 9540167]
43. Snippe E, Simons CJP, Hartmann JA, Menne-Lothmann C, Kramer I, Booij SH, et al. Change in daily life behaviors and depression: within-person and between-person associations. Health Psychol 2016 May;35(5):433-441 [FREE Full text] [doi: 10.1037/hea0000312] [Medline: 26690641]
44. Vogel DL, Wade NG, Wester SR, Larson L, Hackler AH. Seeking help from a mental health professional: the influence of one's social network. J Clin Psychol 2007 Mar;63(3):233-245. [doi: 10.1002/jclp.20345] [Medline: 17211874]
45. Eisenberg D, Downs MF, Golberstein E, Zivin K. Stigma and help seeking for mental health among college students. Med Care Res Rev 2009 Oct;66(5):522-541. [doi: 10.1177/1077558709335173] [Medline: 19454625]
46. Interian A, Ang A, Gara MA, Link BG, Rodriguez MA, Vega WA. Stigma and depression treatment utilization among Latinos: utility of four stigma measures. Psychiatr Serv 2010 Apr;61(4):373-379 [FREE Full text] [doi: 10.1176/ps.2010.61.4.373] [Medline: 20360276]
47. Lack of social support (indicator). OECD. 2022. URL: https://data.oecd.org/healthrisk/lack-of-social-support.htm#:~:text=Definition%20of,on%20in%20times%20of%20trouble [accessed 2023-08-25]
48. Harandi TF, Taghinasab MM, Nayeri TD. The correlation of social support with mental health: a meta-analysis. Electron Physician 2017 Sep 25;9(9):5212-5222 [FREE Full text] [doi: 10.19082/5212] [Medline: 29038699]
49. Connell J, O'Cathain A, Brazier J. Measuring quality of life in mental health: are we asking the right questions? Soc Sci Med 2014 Nov;120:12-20 [FREE Full text] [doi: 10.1016/j.socscimed.2014.08.026] [Medline: 25194472]
50. van Harmelen A, Gibson JL, St Clair MC, Owens M, Brodbeck J, Dunn V, et al. Friendships and family support reduce subsequent depressive symptoms in at-risk adolescents. PLoS One 2016;11(5):e0153715 [FREE Full text] [doi: 10.1371/journal.pone.0153715] [Medline: 27144447]

51. Lindsey MA, Joe S, Nebbitt V. Family matters: the role of mental health stigma and social support on depressive symptoms and subsequent help seeking among African American boys. J Black Psychol 2010 Nov 01;36(4):458-482 [FREE Full text] [doi: 10.1177/0095798409355796] [Medline: 20953336]
52. Kamen C, Cosgrove V, McKellar J, Cronkite R, Moos R. Family support and depressive symptoms: a 23-year follow-up. J Clin Psychol 2011 Mar;67(3):215-223. [doi: 10.1002/jclp.20765] [Medline: 21254050]
53. Cano A, Scaturo DJ, Sprafkin RP, Lantinga LJ, Fiese BH, Brand F. Family support, self-rated health, and psychological distress. Prim Care Companion J Clin Psychiatry 2003 Jun;5(3):111-117 [FREE Full text] [doi: 10.4088/pcc.v05n0302] [Medline: 15154021]
54. Reed M, Bedard C, Perlman CM, Browne DT, Ferro MA. Family functioning and health-related quality of life in parents of children with mental illness. J Child Fam Stud 2023 Feb 25:1-12 [FREE Full text] [doi: 10.1007/s10826-023-02556-6] [Medline: 37362627]
55. Kopinak JK. Mental health in developing countries: challenges and opportunities in introducing Western mental health system in Uganda. Int J MCH AIDS 2015;3(1):22-30 [FREE Full text] [Medline: 27621983]
56. National Service Frameworks. Modern standards and service models: mental health. National Health Service. 1999. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/198051/National_Service_Framework_for_Mental_Health.pdf [accessed 2023-08-25]

Abbreviations

CFI: comparative fit index
DALYs: disability-adjusted life years
EQ-5D-5L: EuroQoL-5 dimensions-5 levels
EQ-VAS: EuroQoL-visual analog scale
MHI-5: Mental Health Inventory-5
MSPSS: Multidimensional Scale of Perceived Social Support
QoL: quality of life
RMSEA: root mean square error approximation
SRMR: standard root mean squared residual
TLI: Tucker-Lewis index
WHO: World Health Organization
YLDs: years lived with disability

Edited by A Mavragani, T Sanchez; submitted 13.03.23; peer-reviewed by YJ Bao, F Venturelli, A AL-Asadi; comments to author 18.05.23; revised version received 13.07.23; accepted 08.08.23; published 11.10.23.

Please cite as:
Tran BX, Dam VAT, Auquier P, Boyer L, Fond G, Nguyen HM, Nguyen HT, Le HT, Tran HNT, Vu GT, Nguyen MD, Nguyen DAT, Ly BV, Latkin CA, Zhang MWB, Ho RCM, Ho CSH
Structural Equation Model for Social Support and Quality of Life Among Individuals With Mental Health Disorders During the COVID-19 Pandemic
JMIR Public Health Surveill 2023;9:e47239
URL: https://publichealth.jmir.org/2023/1/e47239
doi:10.2196/47239
PMID:37819706

Original Paper

Burden of and Trends in Urticaria Globally, Regionally, and Nationally from 1990 to 2019: Systematic Analysis

Xiaoli Liu[1,2], MD; Yuan Cao[3], MD; Wenhui Wang[3], MD

[1]Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing, China
[2]Department of Hernia and Abdominal Wall Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[3]Department of Dermatology, Peking University Third Hospital, Beijing, China

Corresponding Author:
Wenhui Wang, MD
Department of Dermatology
Peking University Third Hospital
49 North Garden Road, Haidian District
Beijing, 100191
China
Phone: 86 15611908751
Email: wwh0608@126.com

Abstract

Background: Urticaria presents a significant global health challenge due to its sudden onset and potential for severe allergic reactions. Past data on worldwide prevalence and incidence is inconsistent due to differing study methodologies, regional differences, and evolving diagnostic criteria. Past studies have often provided broad ranges instead of specific figures, underscoring the necessity for a cohesive global perspective to inform public health strategies.

Objective: We aimed to assess the global burden of urticaria using the 2019 Global Burden of Disease (GBD) study data and systematically analyze urticaria prevalence, incidence, and disability-adjusted life years (DALYs) at global, regional, and national levels, thereby informing more effective prevention and treatment strategies.

Methods: We analyzed the global, regional, and national burden of urticaria from 1990 to 2019 using the 2019 GBD study coordinated by the Institute for Health Metrics and Evaluation. Estimations of urticaria prevalence, incidence, and DALYs were derived using DisMod-MR 2.1, a Bayesian meta-regression tool. The Socio-demographic Index (SDI) was used to correlate development status with health outcomes. The GBD's division of the world into 21 regions and 204 countries and territories facilitated a comprehensive assessment. Age-standardized estimated annual percentage changes were generated for urticaria metrics to quantify temporal trends, with age standardization adjusting for potential confounding from age structure.

Results: From 1990 to 2019, the global age-standardized prevalence, incidence, and DALY rates of urticaria showed marginal changes. In 2019, 65.14 million individuals were affected, with a prevalence rate of 841.88 per 100,000 population. The DALY rate was 50.39 per 100,000 population. Compared to 1990, the global age-standardized prevalence, incidence, and DALY rates saw increases of 2.92, 4.84, and 0.31 per 100,000 population, respectively. Women persistently had higher rates than men. At a regional level in 2019, low-middle SDI regions exhibited the highest age-standardized metrics, whereas high SDI regions reported the lowest. Central Europe showed the highest rates, contrasting with Western Europe's lowest rates. Nationally, urticaria prevalence in 2019 varied dramatically, from a low of 27.1 per 100,000 population in Portugal to a high of 92.0 per 100,000 population in Nepal. India reported the most DALYs at 749,495.9, followed by China, Pakistan, and the United States. Agewise data showed higher rates in younger age groups, which diminished with age and then experienced a slight resurgence in the oldest populations. This pattern was pronounced in women and younger populations, with the largest rises seen in those aged less than 40 years and the smallest in those aged more than 70 years.

Conclusions: Urticaria remains a significant global health issue, with considerable variation across regions, countries, and territories. The increased burden among women, the rising burden in younger populations, and the regional differences in disease burden call for tailored interventions and policies to tackle this emerging public health issue.

(JMIR Public Health Surveill 2023;9:e50114) doi:10.2196/50114

KEYWORDS

urticaria; burden of disease; prevalence; incidence; disability-adjusted life years

Introduction

Urticaria, commonly known as hives, is a dermatologic condition characterized by the sudden onset of itchy, red, and raised welts or bumps on the skin [1,2]. These hives can appear anywhere on the body and vary in size and shape, often causing discomfort and distress to patients [2,3]. In addition to the physical symptoms, urticaria can also have a significant impact on patients' quality of life, affecting their social and psychological well-being [4,5]. The global burden of urticaria has been a growing concern in recent years, with its high prevalence and potential to cause severe allergic reactions and even anaphylaxis in some cases [1,6]. Understanding the global epidemiology of urticaria is crucial for improving prevention and treatment strategies and reducing its impact on public health [7-9].

Despite the high prevalence of urticaria, there is limited information on its global burden [10]. A systematic review and meta-analysis of 52 studies from different regions of the world found that the prevalence of chronic urticaria ranged from 0.1% to 8.9%, with a median of 0.8% [11]. The variation in reported prevalence could be due to differences in study design, sample size, and diagnostic criteria. Furthermore, the burden of urticaria in terms of incidence has not been well documented. A study from the United States estimated that the annual incidence of acute urticaria was 119 per 100,000 population, while the incidence of chronic urticaria was 20 per 100,000 population [12]. Another study from Germany reported a higher incidence of chronic urticaria: 0.15% among a total of 3.53 million individuals [9]. However, these estimates may not be generalizable to other regions of the world.

Understanding the burden of urticaria is crucial for guiding the allocation of health care resources and developing effective prevention and treatment strategies. For instance, a study from Brazil found that adults with chronic urticaria have substantially worse outcomes than people living without chronic urticaria, including health-related quality of life decrements, anxiety, and sleep difficulties [13]. Chronic urticaria was also associated with significant impairments in work and nonwork activities and greatly elevated health care resource use [13]. Another study found that the quality of life of patients with chronic urticaria was significantly lower than that of the general population, underscoring the impact of urticaria on patients' well-being [4].

More research is needed to understand the epidemiology and burden of urticaria worldwide. The Global Burden of Disease (GBD) study aims to provide a systematic analysis of the burden of various human diseases, including urticaria, from 1990 to 2019, offering insights into their prevalence, incidence, disability-adjusted life years (DALYs), and trends. The study can identify areas with the highest burden and inform public health policies and resource allocation. Our study's purpose is to assess the global burden of urticaria using the 2019 GBD study data, examining prevalence, incidence, and DALYs at global, regional, and national levels. The significance lies in informing prevention and treatment strategies, reducing urticaria's impact on human health, and improving quality of life worldwide.

Methods

Data Sources and Study Design

The 2019 GBD study was used to analyze the global, regional, and country urticaria burden from 1990 to 2019. The Institute for Health Metrics and Evaluation (IHME) coordinates the GBD study, which assesses illness, injury, and risk factor burden in 204 nations and territories. The GBD Results tool, an online data repository, provided the latest GBD estimations for our investigation.

To estimate the prevalence, incidence, and DALYs of urticaria, we used DisMod-MR 2.1, a Bayesian meta-regression tool developed by the IHME for modeling disease epidemiology. This tool allows for the integration of various data sources, including surveys, administrative records, and published studies, while accounting for potential biases and heterogeneity in the data.

Socio-Demographic Index

The Socio-demographic Index (SDI) is a composite indicator expressed on a scale of 0 to 1 of development status that is strongly correlated with health outcomes. It is the geometric mean of indices of lag-distributed income per capita, mean years of schooling for individuals aged 15 years or more, and total fertility rate for individuals aged less than 25 years. A location with an SDI of 0 indicates a theoretical minimum level of development status relevant to health outcomes, while a location with an SDI of 1 indicates a theoretical maximum level.

The SDI is stratified into quintiles to provide nuanced distinctions in sociodemographic development. Specifically, a low SDI spans from 0 to 0.454743, followed by a low-middle SDI ranging from 0.454743 to 0.607679. The middle SDI encompasses values between 0.607679 and 0.689504. Subsequently, the high-middle SDI covers the range of 0.689504 to 0.805129, and finally, a high SDI is defined from values of 0.805129 up to 1.

Regions, Countries, and Territories

The GBD study divides the world into 21 regions based on geographical location and epidemiological similarity. These regions are further subdivided into 204 countries and territories. In our analysis, we assessed the burden of urticaria for each region and country, enabling us to identify areas with the highest disease burden and potential targets for intervention.

Statistical Analysis

Urticaria prevalence, incidence, and DALYs were computed for each SDI, area, and nation. We generated age-standardized estimated annual percentage changes (EAPCs) to quantify urticaria incidence, prevalence, and DALYs. Age-standardized incidence, prevalence, and DALY rates of urticaria were compared across populations after they were adjusted for potential age structure confounding by applying the age-specific rates for each location, gender, and year to a GBD world standard population.

The EAPC is a widely used measure of the age-standardized rate trend over a specified time interval [14]. We fitted a regression line to the natural logarithm of the age-standardized rate to calculate the EAPC: $y = \alpha + \beta x + \varepsilon$, where $y = \ln$ (age-standardized rate) and x = calendar year. This was then expressed as a percentage: $100 \times (e\beta - 1)$. The 95% CI of the EAPC was calculated to reflect the temporal trend in the age-standardized rate. An upward trend in the age-standardized rate was indicated when the EAPC and the lower boundary of the 95% CI were positive, while a downward trend was indicated when the EAPC and the upper boundary of the 95% CI were negative. We calculated the EAPCs for age-standardized incidence rates, age-standardized prevalence rates, and age-standardized DALY rates of urticaria to reflect their temporal trends. All data analyses were conducted using R (version 4.2.1; R Foundation for Statistical Computing) and Origin (2022 version; OriginLab).

Ethical Considerations

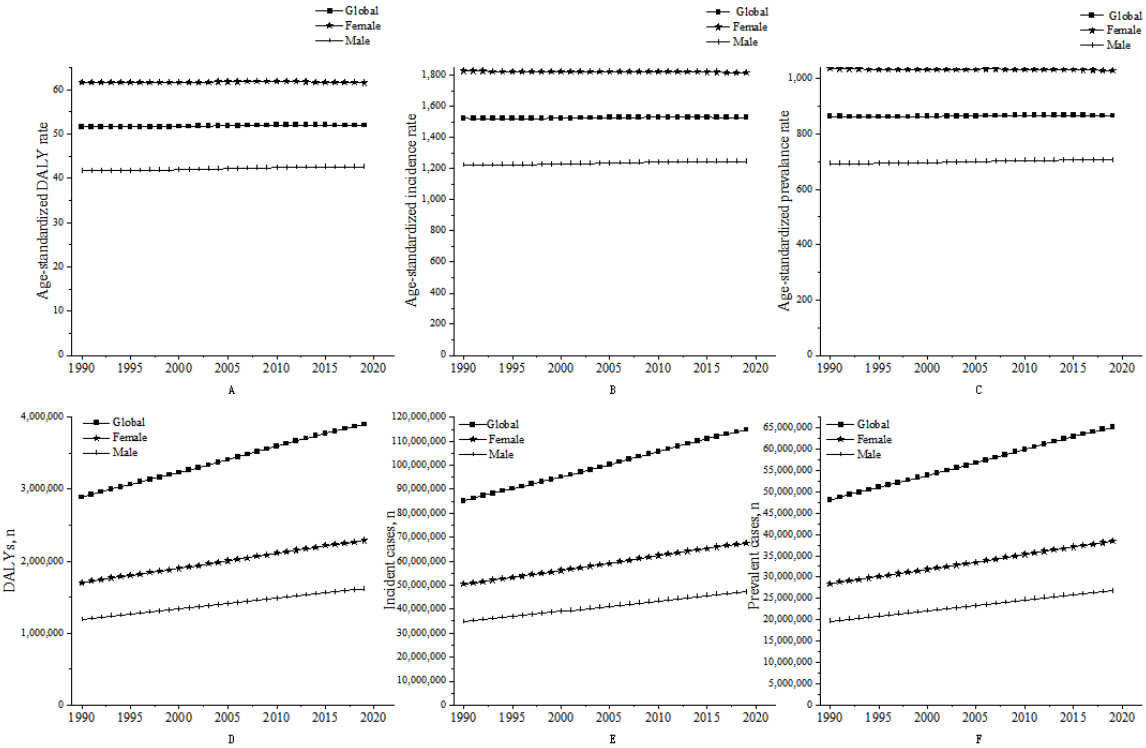
Our research is based on the secondary analysis of data derived from the 2019 GBD study. This comprehensive study integrated data from a diverse array of primary sources, including various surveys, censuses, vital statistics, and other health-related data sets that previously received ethical approval from their respective institutional review boards. Given that our study leverages secondary data, which is an aggregation and anonymization of these primary data sets, our analysis did not necessitate a separate institutional review board approval or exemption [14].

Results

Global Level

From 1990 to 2019, the global age-standardized prevalence, incidence, and DALY rates of urticaria remained stable, with women having higher rates than men. The number of cases, DALYs, and incidence increased during this period. In 2019, 65.14 million people had urticaria, with a prevalence of 841.88 per 100,000 population and an age-standardized prevalence rate of 865.51 per 100,000 population. The global DALY rate was 50.39 per 100,000 population, with an age-standardized rate of 51.91 per 100,000 population. Compared to 1990, the global age-standardized prevalence, incidence, and DALY rates increased by 2.92, 4.84, and 0.31 per 100,000 population, respectively, in 2019 (Figure 1; Multimedia Appendix 1, Tables S1 and S2).

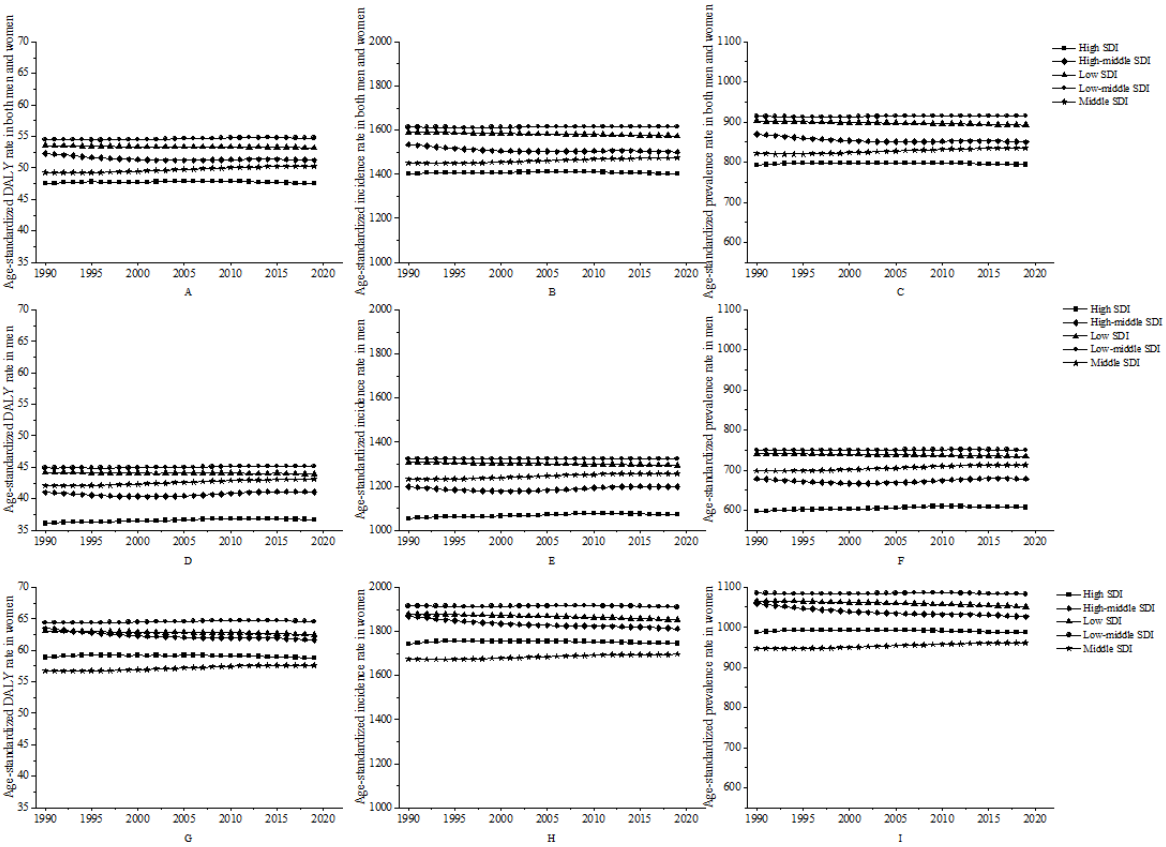
Figure 1. Global, female, and male age-standardized urticaria DALY rate (A), incidence rate (B), prevalence rate (C), DALY number (D), incident cases (E), and prevalent cases (F) from 1990 to 2019. DALY: disability-adjusted life year.



Regional Level

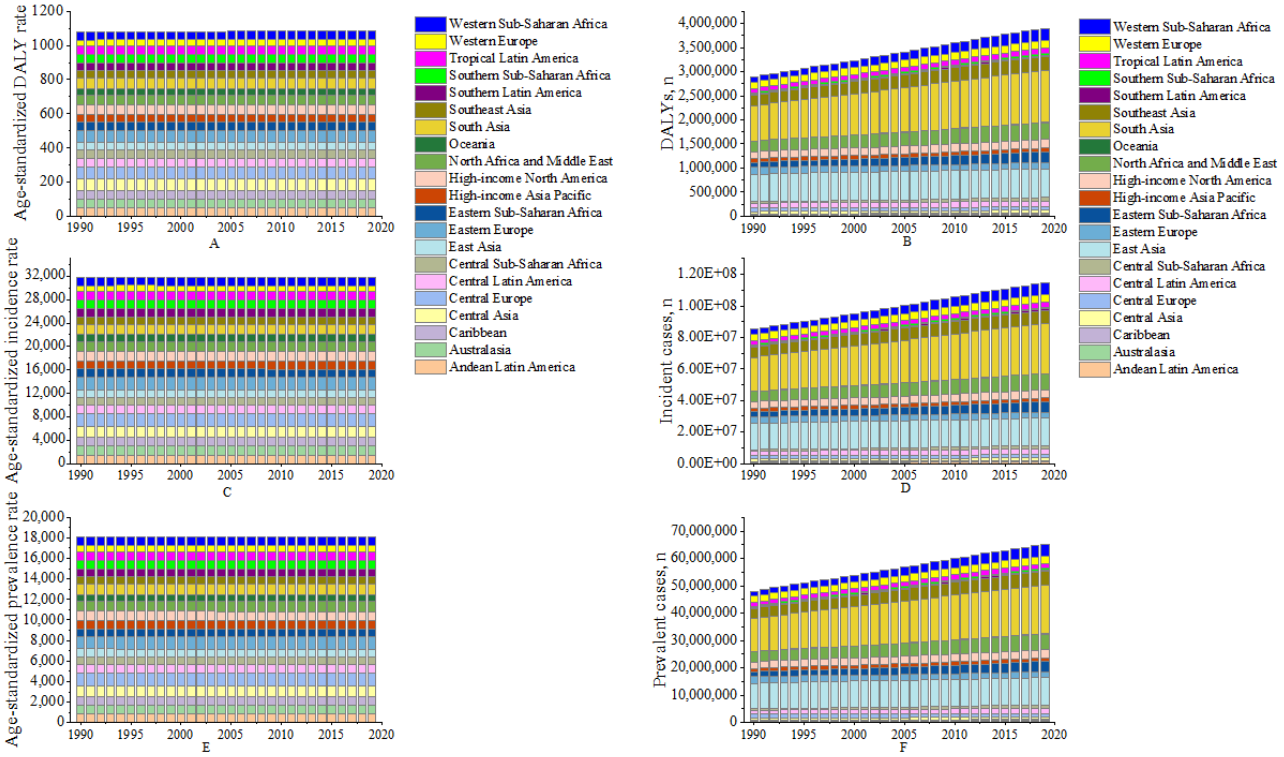
In 2019, low-middle SDI regions had the highest age-standardized prevalence, incidence, and DALY rates, while high SDI regions had the lowest. Middle SDI regions experienced the most significant percentage changes from 1990 to 2019 in all 3 health indicators. All levels of SDI showed a decline in age-standardized DALY rates over the past 30 years, especially for the high-middle SDI group, which decreased from 52.25 per 100,000 population in 1990 to 51.18 per 100,000 population in 2019. However, the decline in the low-middle SDI group was smaller, from 54.44 per 100,000 population in 1990 to 54.72 per 100,000 population in 2019. Overall, the age-standardized DALY rates for women were higher than those for men in all SDI groups. In terms of different SDI groups, the high SDI group consistently had the lowest age-standardized DALY rates, while the low SDI group consistently had the highest rates. The age-standardized DALY rates for women were consistently higher than those for men, especially in the high SDI and high-middle SDI groups (Figure 2; Multimedia Appendix 1, Table S1).

Figure 2. Urticaria age-standardized DALYs, incidence rate, and prevalence rate for both men and women (A, B, C), women (D, E, F), and men (G, H, I) in different SDI areas from 1990 to 2019. DALY: disability-adjusted life year; SDI: Socio-demographic Index.



Incidence rates declined for all SDI groups, with the greatest decrease observed in the high-middle SDI group, from 1534.30 per 100,000 population in 1990 to 1499.49 per 100,000 population in 2019. However, the change in incidence in the low-middle SDI group was not obvious, from 1613.38 per 100,000 population in 1990 to 1613.91 per 100,000 population in 2019. Women had higher incidence rates than men, especially in high SDI and high-middle SDI groups. Prevalence rates remained stable across different SDI levels, with women having higher rates than men. In 2019, Central Europe had the highest age-standardized prevalence, incidence, and DALY rates, while Western Europe had the lowest. From 1990 to 2019, all 21 regions showed stable age-standardized rates but increasing DALY, prevalence, and incidence numbers, with South Asia and East Asia having significantly higher numbers (Figure 2-3; Multimedia Appendix 1, Table S1).

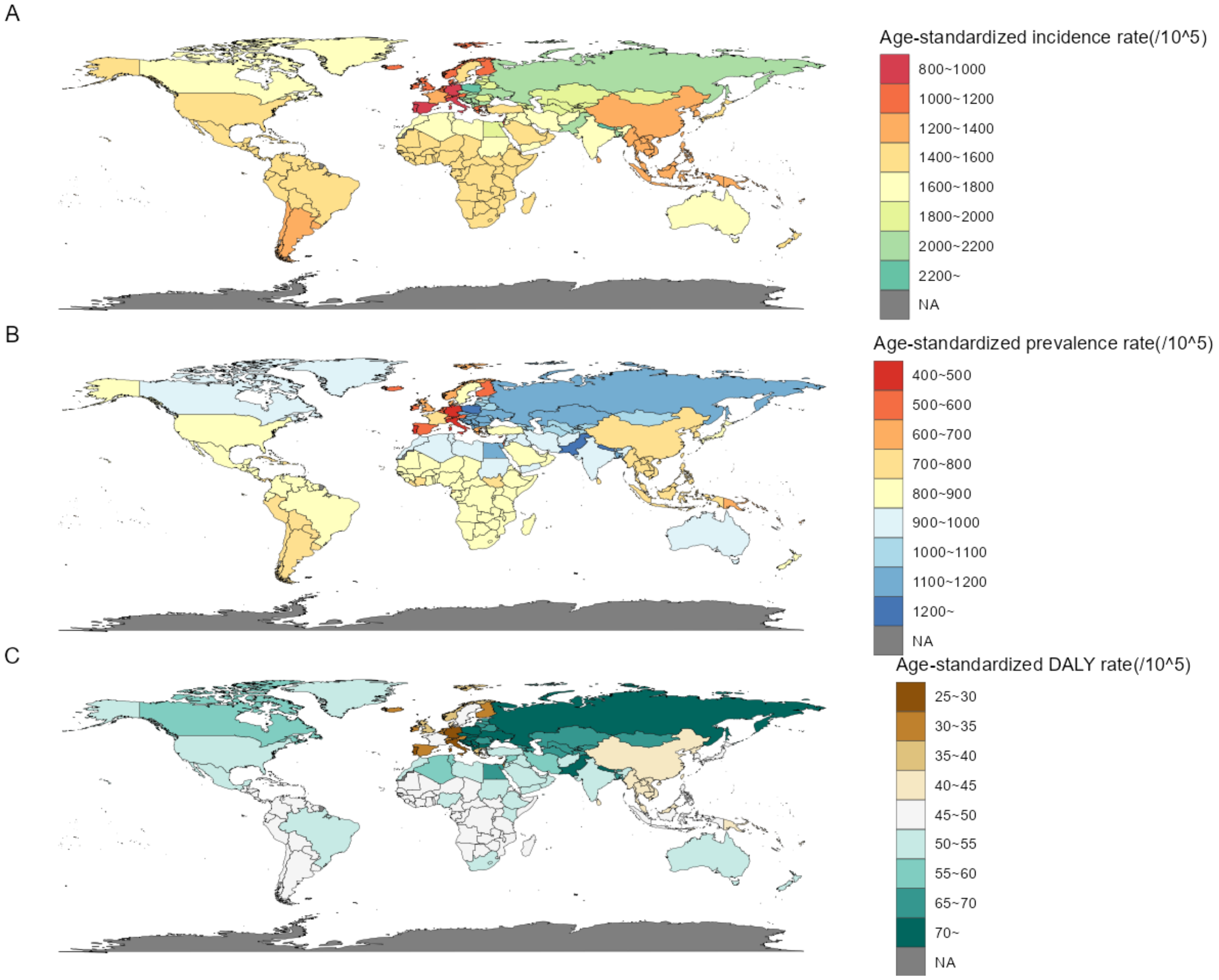
Figure 3. Age-standardized DALY rate (A), number of DALYs (B), incidence rate (C), incident cases (D), prevalence rate (E), and prevalent cases (F) for urticaria in 21 disease burden regions from 1990 to 2019. DALY: disability-adjusted life year.



National Level

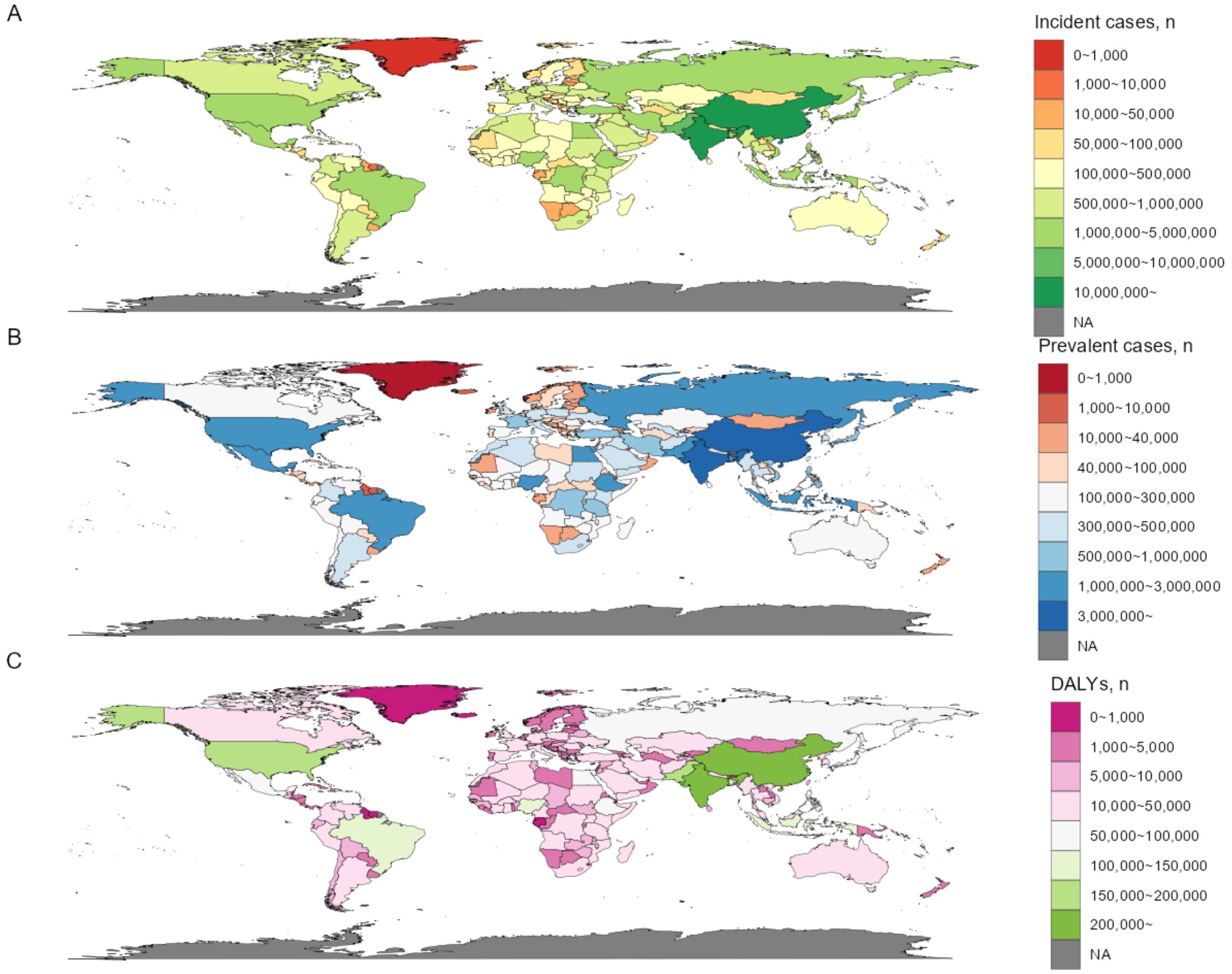
The DALYs of urticaria varied significantly between countries and territories. In 2019, the lowest rate, 27.1 (95% uncertainty interval [UI] 17.7-38.1) per 100,000 population, was reported for Portugal, while the highest rate, 92.0 (95% UI 60.3-130.2) per 100,000 population, was reported for Nepal. The data include the DALY numbers for each country and territory in 2019, with India having the highest number at 749,495.9 (95% UI 490,301.3-1,074,876.0), followed by China at 576,544.5 (95% UI 379,886.1-818,597), Pakistan at 180,580.9 (95% UI 117,789.7-263,500.0), and the United States at 153,479.1 (95% UI 102,752.9-213,697.0), among others. The countries and territories with the lowest DALY numbers included Tokelau (0.60, 95% UI 0.39-0.87), Niue (0.67, 95% UI 0.43-0.96), and Nauru (4.74, 95% UI 3.09-6.93). From 1990 to 2019, the annual percentage change in age-standardized DALY rate differed, with the highest decrease in Qatar at –0.0719% (EAPC –0.0719, 95%CI –0.0906 to –0.0532) (Figure 4-5; Multimedia Appendix 1, Table S3).

Figure 4. Age-standardized incidence rate (A), age-standardized prevalence rate (B), and age-standardized DALY rate (C) for urticaria per 100,000 population in 2019 in 204 countries and territories. DALY: disability-adjusted life year; NA: not available.



For incidence, the age-standardized rate and number of new cases also varied widely across countries and territories. In 2019, out of 204 countries and territories, Portugal reported the lowest rate, 818.2 (95% UI 728.6-917.3) per 100,000 population, while Nepal had the highest rate, 2665.5 (95% UI 2904.2-2445.1) per 100,000 population. Incident cases in the 204 countries and territories varied greatly, ranging from as low as 17.8 (95% UI 15.6-20.5) in Tokelau to as high as 22,065,409 (95% UI 19,265,922-25,287,089) in India. Between 1990 and 2019, the annual percentage change in age-standardized incidence rate varied across countries and territories, with Qatar having the highest decrease at –0.0774% (EAPC –0.0774, 95% CI –0.0961 to –0.0587) (Figure 4-5; Multimedia Appendix 1, Table S3).

Figure 5. Incident cases (A), prevalent cases (B), and DALY numbers (C) for urticaria in 2019 in 204 countries and territories. DALY: disability-adjusted life year.



In terms of prevalence, there was considerable variation in age-standardized rates and the number of cases across countries and territories. Portugal reported the lowest age-standardized prevalence rate, 456.4 (95% UI 406.2-512.2) per 100,000 people in 2019, while Nepal reported the highest rate, 1534.6 (95% UI 1416.2-1662.5) per 100,000 people. Prevalence numbers for the 204 countries and territories ranged widely, from as low as 10.0 (95% UI 8.7-11.5) in Tokelau to as high as 12,552,793 (95% UI 10,916,085-14,379,341) in India. From 1990 to 2019, annual percentage changes in age-standardized prevalence rate varied across countries and territories. Qatar had the highest decrease, with a decline of 0.0795% (EAPC –0.0795, 95% CI –0.0988 to –0.0603) (Figure 4-5; Multimedia Appendix 1, Table S3).

Age

We analyzed global data on rates and numbers for DALYs, prevalence, and incidence across age groups between 1990 and 2019. Rates were higher in younger age groups, decreased with age, and slightly increased in the oldest groups. Women had higher rates than men. A general increase in rates and numbers was observed across age groups between 1990 and 2019 that was more pronounced in younger populations and women. The largest increases occurred in age groups younger than 40 years, while the smallest increases were in groups older than 70 years. In both age groups, women consistently had higher numbers for DALYs, prevalence, and incidence, indicating an overall rising trend in disease burden, prevalence, and incidence (Multimedia Appendix 2, Figure S1).

Regional and SDI-level urticaria disease burden were also analyzed, focusing on rates and numbers for DALYs, prevalence, and incidence across age groups and genders. The youngest age group generally had the highest rates and numbers, decreasing with age. Women consistently had higher rates and numbers across most age groups and regions. Significant regional and SDI-level variations existed, with regions like South Asia, Western sub-Saharan Africa, North Africa, and the Middle East having higher burdens. Central and Eastern Europe had higher rates and numbers than Western Europe. DALY numbers decreased with age across SDI regions and genders, with women having higher numbers. DALY rate trends were similar, but differences between SDI regions and genders were less pronounced (Multimedia Appendix 2, Figures S2-S5).

At the SDI level, urticaria prevalence rates were analyzed across age groups, genders, and SDI regions. The rates were highest in the low-middle SDI group and generally decreased with age, with women consistently having higher rates. High and high-middle SDI groups had lower prevalence rates than low and low-middle groups. The prevalence number also decreased with age and was higher in women. High SDI and high-middle SDI regions had lower prevalence numbers (Multimedia Appendix 2, Figure S6).

Urticaria incidence rates varied across age groups and SDI regions, with women generally having higher rates. The highest incidence rate was in children younger than 5 years, and rates decreased with age. Low-middle SDI regions had the highest incidence rates, while middle SDI regions had the lowest. Incidence numbers followed similar trends, with higher numbers in low SDI and low-middle SDI regions and a decrease with age. The incidence was consistently higher in women across all groups (Multimedia Appendix 2, Figure S7).

Discussion

Our study examines global, regional, and national urticaria trends by age, gender, and SDI regions. Women had higher prevalence, incidence, and DALYs from 1990 to 2019. Urticaria rates were highest in low-middle SDI regions and lowest in high SDI regions. Central Europe had the highest age-standardized rates, while Western Europe had the lowest. Portugal had the lowest rate in 2019 and Nepal the highest. Annual age-standardized rates fluctuated over 30 years. Women had higher urticaria rates across all age categories. Between 1990 and 2019, disease burden, prevalence, and incidence increased across most age categories, particularly in younger people and women.

One of the main findings in our study is the higher burden of urticaria among women compared to men. This observation is consistent with previous studies that have reported a higher prevalence of autoimmune diseases in women [15,16]. The higher rates of urticaria in women may be attributed to several factors, including hormonal differences, genetic predispositions, and environmental exposures [17-19]. Hormonal fluctuations, particularly estrogen levels, have been implicated in the regulation of immune responses and may contribute to the increased susceptibility of women to autoimmune diseases [20,21]. Genetic factors may also play a role, as certain genes related to immune function are located on the X chromosome, potentially resulting in a higher risk of autoimmune diseases in women [22]. Further research is needed to better understand the underlying mechanisms and develop targeted interventions to reduce the burden of urticaria in affected individuals [23-25].

The observed increase in the global burden of urticaria from 1990 to 2019 is a concerning trend that warrants further investigation. Several factors could potentially contribute to this increase, such as changes in environmental exposures, lifestyle factors, and diagnostic practices. Increased exposure to allergens and irritants, such as air pollution and chemical substances, may have contributed to the rise in urticaria prevalence [1,26]. Additionally, changes in lifestyle factors, such as diet, stress, and sedentary behavior, may also have contributed to the increase in urticaria cases [27,28]. Lastly, improvements in diagnostic practices and increased awareness of the condition may have led to a higher detection rate of urticaria cases over time [29,30].

Our findings also highlight substantial regional variations in the burden of urticaria, with low-middle SDI regions exhibiting the highest rates and high SDI regions displaying the lowest. These disparities can be attributed to differences in socioeconomic conditions, health care infrastructure, and access to care, which may influence the overall disease burden in these regions [31-34]. Moreover, regional variations in environmental exposures, such as air pollution and allergen prevalence, could contribute to the observed differences in urticaria burden across regions [35-38].

At the national level, we found significant variation in urticaria's disease burden, with age-standardized rates and numbers of DALYs, incidence, and prevalence differing substantially across countries and territories. These findings highlight the importance of understanding the unique factors that contribute to the burden of urticaria in different countries and territories and the need for tailored interventions to address these factors. For instance, countries and territories with a high burden of urticaria may benefit from increased investment in health care infrastructure, improved access to care, and targeted public health interventions to reduce exposure to environmental risk factors [39-44].

Our study also revealed that the burden of urticaria is generally higher in younger age groups and decreases with age, with a slight increase observed in the oldest age groups. This pattern may be partially explained by differences in immune system function and regulation across different age groups, with younger individuals potentially being more susceptible to immune-mediated conditions such as urticaria [45,46]. Additionally, the increased burden of urticaria in the youngest age groups may be related to the high frequency of infections in children, which can trigger urticaria in susceptible individuals [47,48]. The slight increase in urticaria burden among the oldest age groups may be attributed to age-related changes in the immune system, known as immunosenescence, which can result in a higher risk of autoimmune diseases [1,49].

Our study has several strengths, including the use of a comprehensive data set spanning 3 decades and covering multiple aspects of urticaria burden at the global, regional, and national levels. However, several limitations should also be acknowledged. First, the quality and availability of data on urticaria burden may vary across countries and territories and time periods, potentially affecting the accuracy of our estimates. Second, our study primarily focused on a descriptive analysis of trends in urticaria burden and did not explore the underlying causal factors contributing to these trends. Future research should aim to identify the specific risk factors and mechanisms that contribute to the observed disparities in urticaria burden across different populations and age groups.

In conclusion, our findings highlight the increasing burden of urticaria at the global, regional, and national levels, with substantial disparities observed across age, gender, and SDI groups. The higher burden of urticaria among women, the increasing burden among younger populations, and the regional variations in disease burden underscore the need for targeted interventions and policies to address this growing public health concern. Further research is required to identify the specific risk factors and mechanisms underlying the observed disparities in urticaria burden and inform the development of effective interventions and strategies to reduce the global burden of this condition.

Data Availability

The data sets generated during and/or analyzed during this study are available in the Global Burden of Disease (GBD) study’s public database [50].

Authors' Contributions

WW had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors contributed to concept and design; critical revision of the manuscript for important intellectual content; administrative, technical, or material support; and supervision. XL contributed to acquisition, analysis, and interpretation of data and drafting of the manuscript. XL and WW contributed to statistical analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Tables S1-S3.
[DOCX File , 96 KB - publichealth_v9i1e50114_app1.docx]

Multimedia Appendix 2
Supplementary Figures S1-S7.
[DOCX File , 440 KB - publichealth_v9i1e50114_app2.docx]

References

1. Zuberbier T, Aberer W, Asero R, Abdul Latiff AH, Baker D, Ballmer-Weber B, et al. The EAACI/GA(2)LEN/EDF/WAO guideline for the definition, classification, diagnosis and management of urticaria. Allergy 2018 Jul;73(7):1393-1414. [doi: 10.1111/all.13397] [Medline: 29336054]
2. Kolkhir P, Giménez-Arnau AM, Kulthanan K, Peter J, Metz M, Maurer M. Urticaria. Nat Rev Dis Primers 2022 Sep 15;8(1):61. [doi: 10.1038/s41572-022-00389-z] [Medline: 36109590]
3. Antia C, Baquerizo K, Korman A, Bernstein JA, Alikhan A. Urticaria: A comprehensive review: Epidemiology, diagnosis, and work-up. J Am Acad Dermatol 2018 Oct;79(4):599-614. [doi: 10.1016/j.jaad.2018.01.020] [Medline: 30241623]
4. Gonçalo M, Gimenéz-Arnau A, Al-Ahmad M, Ben-Shoshan M, Bernstein J, Ensina L, et al. The global burden of chronic urticaria for the patient and society. Br J Dermatol 2021 Feb;184(2):226-236. [doi: 10.1111/bjd.19561] [Medline: 32956489]
5. Stadler P, Marsela E, Kämmerer T, Frommherz LH, Clanner-Engelshofen B, French LE, et al. Impact of allergic reactions and urticaria on mental health and quality of life. Allergol Immunopathol (Madr) 2022;50(2):124-130. [doi: 10.15586/aei.v50i2.525] [Medline: 35257555]
6. Kim M, Lee YW, Kim SR, Kim J, Min TK, Park H, et al. COVID-19 vaccine-associated anaphylaxis and allergic reactions: consensus statements of the KAAACI Urticaria/Angioedema/Anaphylaxis Working Group. Allergy Asthma Immunol Res 2021 Jul;13(4):526-544 [FREE Full text] [doi: 10.4168/aair.2021.13.4.526] [Medline: 34212542]
7. Kim BR, Yang S, Choi JW, Choi CW, Youn SW. Epidemiology and comorbidities of patients with chronic urticaria in Korea: A nationwide population-based study. J Dermatol 2018 Jan;45(1):10-16. [doi: 10.1111/1346-8138.14075] [Medline: 28983950]
8. Seo J, Kwon J. Epidemiology of urticaria including physical urticaria and angioedema in Korea. Korean J Intern Med 2019 Mar;34(2):418-425 [FREE Full text] [doi: 10.3904/kjim.2017.203] [Medline: 29742892]
9. Weller K, Maurer M, Bauer A, Wedi B, Wagner N, Schliemann S, et al. Epidemiology, comorbidities, and healthcare utilization of patients with chronic urticaria in Germany. J Eur Acad Dermatol Venereol 2022 Jan;36(1):91-99. [doi: 10.1111/jdv.17724] [Medline: 34622498]
10. Hay RJ, Johns NE, Williams HC, Bolliger IW, Dellavalle RP, Margolis DJ, et al. The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions. J Invest Dermatol 2014 Jun;134(6):1527-1534 [FREE Full text] [doi: 10.1038/jid.2013.446] [Medline: 24166134]
11. Maurer M, Weller K, Bindslev-Jensen C, Giménez-Arnau A, Bousquet PJ, Bousquet J, et al. Unmet clinical needs in chronic spontaneous urticaria. A GA(2)LEN task force report. Allergy 2011 Mar;66(3):317-330. [doi: 10.1111/j.1398-9995.2010.02496.x] [Medline: 21083565]
12. Greaves M. Chronic urticaria. J Allergy Clin Immunol 2000 Apr;105(4):664-672. [doi: 10.1067/mai.2000.105706] [Medline: 10756214]

13. Balp M, Lopes da Silva N, Vietri J, Tian H, Ensina LF. The burden of chronic urticaria from Brazilian patients' perspective. Dermatol Ther (Heidelb) 2017 Dec;7(4):535-545 [FREE Full text] [doi: 10.1007/s13555-017-0191-4] [Medline: 28748405]
14. Cao G, Liu J, Liu M. Global, regional, and national incidence and mortality of neonatal preterm birth, 1990-2019. JAMA Pediatr 2022 Aug 01;176(8):787-796 [FREE Full text] [doi: 10.1001/jamapediatrics.2022.1622] [Medline: 35639401]
15. Fairweather D, Rose NR. Women and autoimmune diseases. Emerg Infect Dis 2004 Nov;10(11):2005-2011 [FREE Full text] [doi: 10.3201/eid1011.040367] [Medline: 15550215]
16. Whitacre CC. Sex differences in autoimmune disease. Nat Immunol 2001 Sep;2(9):777-780. [doi: 10.1038/ni0901-777] [Medline: 11526384]
17. Bingham CO. Immunomodulatory approaches to the management of chronic urticaria: an immune-mediated inflammatory disease. Curr Allergy Asthma Rep 2008 Jul;8(4):278-287. [doi: 10.1007/s11882-008-0046-2] [Medline: 18606079]
18. Nie Y, Liu L, Xue S, Yan L, Ma N, Liu X, et al. The association between air pollution, meteorological factors, and daily outpatient visits for urticaria in Shijiazhuang, Hebei Province, China: a time series analysis. Environ Sci Pollut Res Int 2023 Jan;30(4):10664-10682. [doi: 10.1007/s11356-022-22901-4] [Medline: 36076138]
19. Choo KJL, Ho AFW, Gui H, Tay PJM, Lee HY, Koh MS, et al. Relationship between local weather, air pollution and hospital attendances for urticaria in children: Time stratified analysis of 12,002 cases. Clin Exp Allergy 2022 Jan;52(1):180-182. [doi: 10.1111/cea.14015] [Medline: 34549849]
20. Kasperska-Zajac A, Brzoza Z, Rogala B. Sex hormones and urticaria. J Dermatol Sci 2008 Nov;52(2):79-86. [doi: 10.1016/j.jdermsci.2008.04.002] [Medline: 18485675]
21. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol 2016 Oct;16(10):626-638. [doi: 10.1038/nri.2016.90] [Medline: 27546235]
22. Libert C, Dejager L, Pinheiro I. The X chromosome in immune functions: when a chromosome makes the difference. Nat Rev Immunol 2010 Aug;10(8):594-604. [doi: 10.1038/nri2815] [Medline: 20651746]
23. Toppe E, Haas N, Henz BM. Neutrophilic urticaria: clinical features, histological changes and possible mechanisms. Br J Dermatol 1998 Feb;138(2):248-253. [doi: 10.1046/j.1365-2133.1998.02069.x] [Medline: 9602869]
24. Helaskoski E, Suojalehto H, Kuuliala O, Aalto-Korte K. Occupational contact urticaria and protein contact dermatitis: causes and concomitant airway diseases. Contact Dermatitis 2017 Dec;77(6):390-396. [doi: 10.1111/cod.12856] [Medline: 28795430]
25. Larenas-Linnemann DES, Parisi CAS, Ritchie C, Cardona-Villa R, Cherrez-Ojeda I, Cherrez A, et al. Update on omalizumab for urticaria: What's new in the literature from mechanisms to clinic. Curr Allergy Asthma Rep 2018 May 09;18(5):33. [doi: 10.1007/s11882-018-0787-5] [Medline: 29744661]
26. Gilles S, Akdis C, Lauener R, Schmid-Grendelmeier P, Bieber T, Schäppi G, et al. The role of environmental factors in allergy: A critical reappraisal. Exp Dermatol 2018 Nov;27(11):1193-1200. [doi: 10.1111/exd.13769] [Medline: 30099779]
27. Weldon D. Quality of life in patients with urticaria and angioedema: assessing burden of disease. Allergy Asthma Proc 2014;35(1):4-9. [doi: 10.2500/aap.2014.35.3713] [Medline: 24433591]
28. Milán T, Verkasalo P, Kaprio J, Koskenvuo M. Lifestyle differences in twin pairs discordant for basal cell carcinoma of the skin. Br J Dermatol 2003 Jul;149(1):115-123. [doi: 10.1046/j.1365-2133.2003.05352.x] [Medline: 12890204]
29. Maurer M, Zuberbier T, Metz M. The classification, pathogenesis, diagnostic workup, and management of urticaria: An update. Handb Exp Pharmacol 2022;268:117-133. [doi: 10.1007/164_2021_506] [Medline: 34247278]
30. Frigas E, Park MA. Acute urticaria and angioedema: diagnostic and treatment considerations. Am J Clin Dermatol 2009;10(4):239-250. [doi: 10.2165/00128071-200910040-00004] [Medline: 19489657]
31. Caraballo L, Zakzuk J, Lee BW, Acevedo N, Soh JY, Sánchez-Borges M, et al. Particularities of allergy in the tropics. World Allergy Organ J 2016;9:20 [FREE Full text] [doi: 10.1186/s40413-016-0110-7] [Medline: 27386040]
32. McBain R, Salhi C, Morris JE, Salomon JA, Betancourt TS. Disease burden and mental health system capacity: WHO Atlas study of 117 low- and middle-income countries. Br J Psychiatry 2012 Dec;201(6):444-450. [doi: 10.1192/bjp.bp.112.112318] [Medline: 23137730]
33. Taneja N, Mewara A. Shigellosis: Epidemiology in India. Indian J Med Res 2016 May;143(5):565-576 [FREE Full text] [doi: 10.4103/0971-5916.187104] [Medline: 27487999]
34. Bauer UE, Briss PA, Goodman RA, Bowman BA. Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. Lancet 2014 Jul 05;384(9937):45-52. [doi: 10.1016/S0140-6736(14)60648-6] [Medline: 24996589]
35. Weinmayr G, Weiland SK, Björkstén B, Brunekreef B, Büchele G, Cookson WOC, et al. Atopic sensitization and the international variation of asthma symptom prevalence in children. Am J Respir Crit Care Med 2007 Sep 15;176(6):565-574. [doi: 10.1164/rccm.200607-994OC] [Medline: 17575099]
36. Nkurunungi G, Lubyayi L, Versteeg SA, Sanya RE, Nassuuna J, Kabagenyi J, et al. Do helminth infections underpin urban-rural differences in risk factors for allergy-related outcomes? Clin Exp Allergy 2019 May;49(5):663-676 [FREE Full text] [doi: 10.1111/cea.13335] [Medline: 30633850]
37. Lule SA, Mpairwe H, Nampijja M, Akello F, Kabagenyi J, Namara B, et al. Life-course of atopy and allergy-related disease events in tropical sub-Saharan Africa: A birth cohort study. Pediatr Allergy Immunol 2017 Jun;28(4):377-383 [FREE Full text] [doi: 10.1111/pai.12719] [Medline: 28339128]

38. Thalayasingam M, Allameen NA, Soh JY, Bigliardi P, Van Bever H, Shek LPC. Wheat-dependent exercise-induced anaphylaxis: a retrospective case review from a tertiary hospital. Postgrad Med J 2014 Sep;90(1067):488-492. [doi: 10.1136/postgradmedj-2013-132257] [Medline: 24982201]
39. Peprah E, Wonkam A. Biomedical research, a tool to address the health issues that affect African populations. Global Health 2013 Oct 21;9(1):50 [FREE Full text] [doi: 10.1186/1744-8603-9-50] [Medline: 24143865]
40. Lyerly HK, Fawzy MR, Aziz Z, Nair R, Pramesh CS, Parmar V, et al. Regional variation in identified cancer care needs of early-career oncologists in China, India, and Pakistan. Oncologist 2015 May;20(5):532-538 [FREE Full text] [doi: 10.1634/theoncologist.2014-0213] [Medline: 25888267]
41. Sherif YA, Hassan MA, Thuy Vu M, Rosengart TK, Davis RW. Twelve Tips on enhancing global health education in graduate medical training programs. Medical Teacher 2020 May 12;43(2):142-147. [doi: 10.1080/0142159x.2020.1762033]
42. Ratnayake R, Rawashdeh F, AbuAlRub R, Al-Ali N, Fawad M, Bani Hani M, et al. Access to care and prevalence of hypertension and diabetes among Syrian Refugees in Northern Jordan. JAMA Netw Open 2020 Oct 01;3(10):e2021678 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.21678] [Medline: 33052405]
43. Phung KT, Chaaya M, Prince M, Atweh S, El Asmar K, Karam G, et al. Dementia prevalence, care arrangement, and access to care in Lebanon: A pilot study. Alzheimers Dement 2017 Dec;13(12):1317-1326 [FREE Full text] [doi: 10.1016/j.jalz.2017.04.007] [Medline: 28579315]
44. Rajeev M, Guis H, Edosoa GT, Hanitriniaina C, Randrianarijaona A, Mangahasimbola RT, et al. How geographic access to care shapes disease burden: The current impact of post-exposure prophylaxis and potential for expanded access to prevent human rabies deaths in Madagascar. PLoS Negl Trop Dis 2021 Apr;15(4):e0008821 [FREE Full text] [doi: 10.1371/journal.pntd.0008821] [Medline: 33901194]
45. Radonjic-Hoesli S, Hofmeier KS, Micaletto S, Schmid-Grendelmeier P, Bircher A, Simon D. Urticaria and angioedema: An update on classification and pathogenesis. Clin Rev Allergy Immunol 2018 Feb;54(1):88-101 [FREE Full text] [doi: 10.1007/s12016-017-8628-1] [Medline: 28748365]
46. Hennino A, Bérard F, Guillot I, Saad N, Rozières A, Nicolas J. Pathophysiology of urticaria. Clin Rev Allergy Immunol 2006 Feb;30(1):3-11. [doi: 10.1385/CRIAI:30:1:003] [Medline: 16461989]
47. Saini S, Shams M, Bernstein JA, Maurer M. Urticaria and angioedema across the ages. J Allergy Clin Immunol Pract 2020 Jun;8(6):1866-1874. [doi: 10.1016/j.jaip.2020.03.030] [Medline: 32298850]
48. Wang EA, Chan SK. Chronic urticaria in children: An update on diagnosis and treatment. Curr Allergy Asthma Rep 2020 Jun 06;20(8):31. [doi: 10.1007/s11882-020-00929-1] [Medline: 32506151]
49. Ventura MT, Scichilone N, Gelardi M, Patella V, Ridolo E. Management of allergic disease in the elderly: key considerations, recommendations and emerging therapies. Expert Rev Clin Immunol 2015;11(11):1219-1228. [doi: 10.1586/1744666X.2015.1081564] [Medline: 26483075]
50. GBD Results. Global Burden of Disease study. URL: https://vizhub.healthdata.org/gbd-results/ [accessed 2023-10-20]

Abbreviations

DALY: disability-adjusted life year
EAPC: age-standardized estimated annual percentage change
GBD: Global Burden of Disease
IHME: Institute for Health Metrics and Evaluation
SDI: Socio-demographic Index
UI: uncertainty interval

Edited by A Mavragani, T Sanchez; submitted 19.06.23; peer-reviewed by X Lin, H Guo; comments to author 20.08.23; revised version received 04.09.23; accepted 03.10.23; published 26.10.23.

Please cite as:
Liu X, Cao Y, Wang W
Burden of and Trends in Urticaria Globally, Regionally, and Nationally from 1990 to 2019: Systematic Analysis
JMIR Public Health Surveill 2023;9:e50114
URL: https://publichealth.jmir.org/2023/1/e50114
doi:10.2196/50114
PMID:37883176

Original Paper

Trend of the Tuberculous Pleurisy Notification Rate in Eastern China During 2017-2021: Spatiotemporal Analysis

Ying Zhou[1*], BD; Dan Luo[2*], BD; Kui Liu[3,4*], MD; Bin Chen[3], MD; Songhua Chen[3], MD; Junhang Pan[3], PhD; Zhengwei Liu[3], MD; Jianmin Jiang[3,5], PhD

[1]School of Public Health, Hangzhou Normal University, Hangzhou, China

[2]School of Public Health, Hangzhou Medical College, Hangzhou, China

[3]Department of Tuberculosis Control and Prevention, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

[4]National Centre for Tuberculosis Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[5]Key Laboratory of Vaccine, Prevention and Control of Infectious Disease of Zhejiang Province, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Jianmin Jiang, PhD
Department of Tuberculosis Control and Prevention
Zhejiang Provincial Center for Disease Control and Prevention
3399 Binsheng Road
BinJiang District
Hangzhou, 310000
China
Phone: 86 057187115009
Fax: 86 057187115100
Email: jmjiang@cdc.zj.cn

Abstract

Background: Tuberculous pleurisy (TP) presents a serious allergic reaction in the pleura caused by *Mycobacterium tuberculosis*; however, few studies have described its spatial epidemiological characteristics in eastern China.

Objective: This study aimed to determine the epidemiological distribution of TP and predict its further development in Zhejiang Province.

Methods: Data on all notified cases of TP in Zhejiang Province, China, from 2017 to 2021 were collected from the existing tuberculosis information management system. Analyses, including spatial autocorrelation and spatial-temporal scan analysis, were performed to identify hot spots and clusters, respectively. The prediction of TP prevalence was performed using the seasonal autoregressive integrated moving average (SARIMA), Holt-Winters exponential smoothing, and Prophet models using R (The R Foundation) and Python (Python Software Foundation).

Results: The average notification rate of TP in Zhejiang Province was 7.06 cases per 100,000 population, peaking in the summer. The male-to-female ratio was 2.18:1. In terms of geographical distribution, clusters of cases were observed in the western part of Zhejiang Province, including parts of Hangzhou, Quzhou, Jinhua, Lishui, Wenzhou, and Taizhou city. Spatial-temporal analysis identified 1 most likely cluster and 4 secondary clusters. The Holt-Winters model outperformed the SARIMA and Prophet models in predicting the trend in TP prevalence.

Conclusions: The western region of Zhejiang Province had the highest risk of TP. Comprehensive interventions, such as chest x-ray screening and symptom screening, should be reinforced to improve early identification. Additionally, a more systematic assessment of the prevalence trend of TP should include more predictors.

(JMIR Public Health Surveill 2023;9:e49859) doi:10.2196/49859

KEYWORDS

tuberculous pleurisy; spatio-temporal; epidemiology; prediction; time series

Introduction

Humanity has been affected by tuberculosis (TB) for centuries, which is an infectious agent–caused disease considered the leading cause of mortality worldwide since the beginning of recorded history. Among people who contract TB each year, the majority are adults, with the incidence of TB being higher in men than in women. According to the Global Tuberculosis Report of 2022 released by the World Health Organization, the estimated TB incidence rate in 2021 increased by 3.6% from 2020, and China accounted for 7.4% of TB cases worldwide. Furthermore, the number of TB deaths worldwide successively increased from 1.4 million in 2019 to 1.6 million in 2021, only after the onset of the COVID-19 pandemic [1]. *Mycobacterium tuberculosis* generally affects not only the lungs but also other sites such as the pleura [2].

Tuberculosis pleurisy (TP) is a common clinical manifestation of TB post primary infection or due to the reactivation of latent TB [3]. Available documentation shows that TP accounts for about 4% of total tuberculosis cases in the United States and Brazil, 20% of cases in South Africa, and 7.3% of cases in Korea [4]. In the United States, the incidence of both TP and TB has been declining, but to a lesser degree for TP [5]. In 2018, it was categorized as pulmonary tuberculosis (PTB) in China [6]. TP can be divided into dry and exudative pleurisy. Dry pleurisy is an early inflammatory reaction of the pleura, usually without obvious imaging abnormalities. Exudative pleurisy is characterized by unilateral or bilateral pleural effusions, which often occur as thickened pleural adhesions with slow absorption [7]. The recommended standard regimen for TP is up to 8 months, which is longer than the usual treatment period for TB [8]. Previous studies have shown that the prevalence of TP is low in women, while the number of patients with TP is higher in eastern China, among farmers and patients with AIDS [9-11]. However, the temporal periodicity and spatial clustering characteristics of TP have not been explored in eastern China.

Zhejiang Province is located in the eastern region of China and has a developed economy. In recent years, certain parts of the province have experienced a high burden of TB. Understanding the spatial and temporal patterns of TB incidence can provide insights into the clusters of this disease [12]. Spatial-temporal analysis has been widely used to describe the epidemiological distribution and aggregation of infectious diseases in both space and time; however, few studies have explored the spatial epidemiological characteristics of TP. In comparison to traditional epidemiological techniques, tools using geographic information systems have significant advantages in tracking disease, identifying areas of clustering, and analyzing the spread of disease across communities [13]. Spatiotemporal cluster detection has played a crucial role in understanding infectious disease spread, with its dynamic course and dimensions of time and place, and a combination of both [14]. In addition, time-series models were selected to provide scientific clues for predicting the prevalence of diseases such as desiccation syndrome and AIDS [15,16]. However, there are few such studies on TP.

Therefore, this study aimed to use spatial-temporal analysis to identify the epidemiological characteristics of TP. Furthermore, the spatial differentiation of TP can provide a visual map depicting the diversity of TP occurrence across different regions. Additionally, this study used time-series models to provide scientific clues for the control of and interventions targeting TP infections in the province.

Methods

Overview of the Study Area

Zhejiang Province is located on the eastern coast of China, with a land area of 101,800 km^2, accounting for 1.06% of the country’s land area [17]. The province encompasses plains, mountains, coasts, islands, and lakes. Zhejiang Province comprises 90 counties and cities, including 2 subprovincial cities and 9 prefecture-level cities [17,18].

Data Source

All recorded cases of TP in Zhejiang Province from 2017 to 2021 were obtained from the TB Information Management System (TBIMS), which is a first-generation web-based information system designed and built by the National Tuberculosis Control and Prevention Center in 2005 [19]. The system collected data on demographics, diagnosis, laboratory results, and treatment outcomes of all patients with TB during the specified period. In addition, demographic data for Zhejiang Province were obtained from the China Information System for Disease Control and Prevention.

Definition

Diagnostic criteria for TP were as follows: (1) confirmed patients: cultures of pleural biopsy specimens were positive, and the strain was identified as *M tuberculosis*, changes in tuberculosis confirmed through pleural tissue biopsy; and (2) clinically diagnosed patients: imaging examination showing pleural effusion with (i) a moderately positive or strongly positive outcome on the tuberculin pure protein derivative skin test, (ii) a positive outcome on the γ-interferon release test, or (iii) a positive outcome on the *M tuberculosis* antibody test [20]. Delays in seeking medical treatment were indicated when the interval between the appearance of the first symptom and the visit to the designated hospital was > 2 weeks, and the delay in diagnosis was the time between the first consultation and a confirmed diagnosis [21].

General Epidemiological Characteristics of Patients With TP

The features of patients with TP in Zhejiang Province from 2017 to 2021 are described in the context of sex, age group, city, antituberculosis treatment, mode of case finding, treatment history, delay in access to care and final diagnosis, and treatment outcomes.

Spatial Autocorrelation Analysis

In this study, the Moran I index and local indicators of spatial association (LISA) were used to analyze global and local spatial autocorrelation. The global Moran I index is a comprehensive measure of spatial autocorrelation in the whole study area, and it can indicate the average degree of spatial differences between each unit and the surrounding area. If partial positive and partial negative spatial correlations coexist in the overall study area, LISA are needed to reveal the possible spatial variability [22].

Space-Time Scan Statistics

For the spatiotemporal analysis, we used the Kulldorff spatial-temporal scan statistic to identify clusters based on specific time and location combinations. This statistic uses a Poisson probability distribution mode. The Kulldorff method uses a moving cylindrical window to scan the study area. The height of the cylinder represents the possible clustering time, the bottom represents the clustering area, and the radius serves as the scanning risk population [23]. The significance of the clusters identified at a 95% confidence level was tested using Monte Carlo simulations to identify the clusters.

Prediction Models

Three models were used in this study. The Holt-Winters exponential smoothing model eliminated some random fluctuations, which can assign different weights to the data of each period and reasonably predict the future development trend [16]. The seasonal autoregressive integrated moving average (SARIMA) model divided the observed value into 3 parts: residuals, seasonal feature, and true trend [24]. The Holt-Winters exponential smoothing model and SARIMA model were introduced more specifically in a previous study [16]. The Prophet model, developed by Meta, allowed using time as a regressor to fit both linear and nonlinear functions of time [25], and a more specific introduction to this model is provided in Xie et al’s [26] study.

Statistical Analysis

Descriptive analysis was conducted using Excel (Microsoft Corp) and SPSS (version 28; IBM Corp). The Moran I index was calculated using GeoDa (version 1.20), and SaTScan software (version 10.1; Harvard Medical School) was used to determine the spatial and temporal aggregation of TP. All visual results are presented using the ArcGIS software (version 10.8; SERI Inc). The prediction of TP prevalence was model-fitted using R (version 4.2.0; The R Foundation) and Python (version 3.8; Python Software Foundation).

Ethical Considerations

This study was approved by the Ethics Committee of the Zhejiang Provincial Center for Disease Control and Prevention (audit number 2022-032-01), which waived the requirement for informed consent. All personal information obtained in this study was treated and kept confidential in accordance with the required protocols.

Results

General Epidemiological Characteristics of Patients With TP

During the study period, a total of 11,531 patients with TP were identified, with a notification rate of 7.8 cases per 100,000 population in 2017, 8.2 cases per 100,000 population in 2018, 7.7 cases per 100,000 population in 2019, 6.6 cases per 100,000 population in 2020, and 5.0 cases per 100,000 population in 2021, respectively (Figure 1).The male to female ratio was 2.18:1. Among all cases of TP, the top notified cases and notification rate were noted in 2018 (n=2594; 8.2 cases per 100,000 population). The number of cases peaked in the second quarter (n=3354) and plummeted in the fourth quarter (n=2558) during the study period, indicating seasonal distribution characteristics. In terms of age distribution, 74 (0.64%) individuals were aged 0-18 years, 6359 (55.15%) were aged 19-59 years, and 5098 (44.21%) were aged ≥60 years. In terms of geographical distribution, 2319 (20%) cases were notified in Hangzhou, while 1294 (11.22%), 1393 (12.08%), 1452 (12.59%), and 1259 (10.92%) cases were reported in Wenzhou, Ningbo, Jinhua, and Taizhou, respectively. Passive findings was the main source of TP identification in Zhejiang Province. Nearly all patients (n=11,326, 98.22%) were treated for the first time, with nearly 90% of them having completed treatment. Delayed access to medical care was observed among 6317 (54.78%) patients with TP, and diagnosis was delayed among 1462 (12.68%) patients. In addition, 86.26% (n=9949) of patients with TP were treated with a duration of ≥9 months (Table 1).

Figure 1. Change in the trend of tuberculous pleurisy notification rate and case numbers by month.

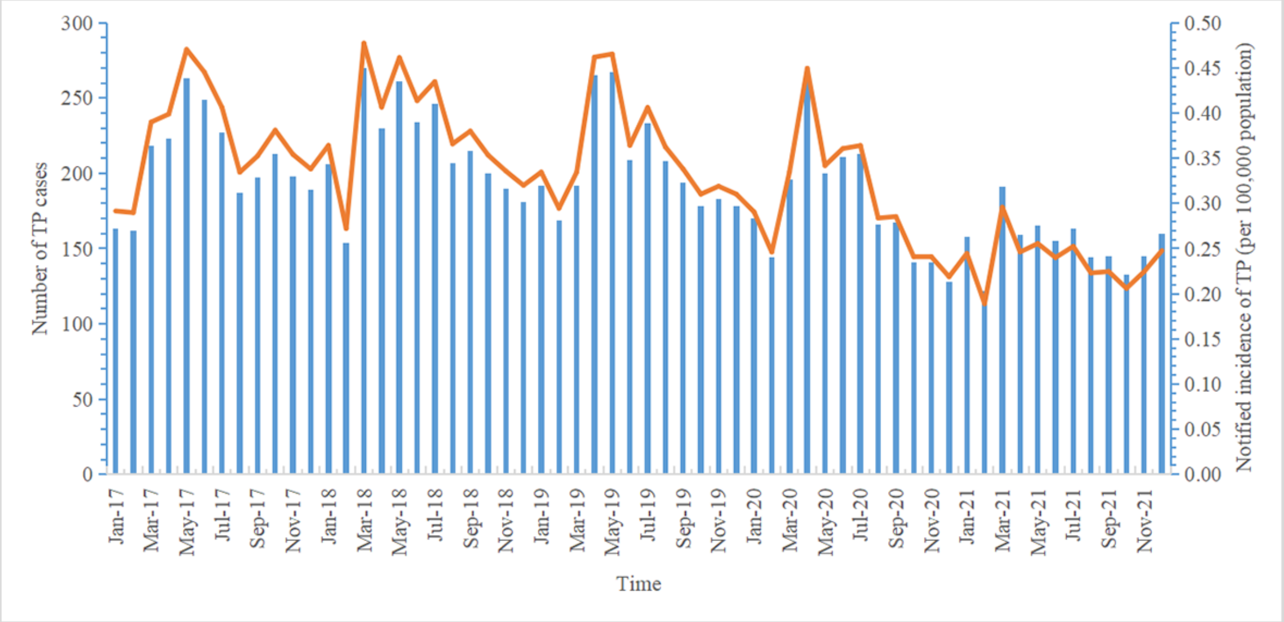


Table 1. Epidemiological characteristics of patients with tuberculosis pleurisy in Zhejiang Province from 2017 to 2021 (N=11,531).

Characteristic	Patients, n (%)
Sex	
Male	7901 (68.52)
Female	3630 (31.48)
Age group (years)	
0-18	74 (0.64)
19-59	6359 (55.15)
≥60	5098 (44.21)
City	
Hangzhou	2319 (20.11)
Jinhua	1452 (12.59)
Ningbo	1393 (12.08)
Wenzhou	1294 (11.22)
Taizhou	1259 (10.92)
Shaoxing	1078 (9.35)
Quzhou	858 (7.44)
Jiaxing	674 (5.85)
Huzhou	582 (5.05)
Lishui	506 (4.39)
Zhoushan	116 (1.00)
Case finding	
Passive finding	11,513 (99.84)
Active finding	18 (0.16)
Treatment history	
Initial treatment	11,326 (98.22)
Retreatment	205 (1.78)
Anti – tuberculosis pleurisy treatment	
Yes	11,523 (99.93)
No	8 (0.07)
Time of delay in seeking medical treatment (days)	
0-14	5214 (45.22)
15-29	2968 (25.74)
30-44	1569 (13.61)
45-59	503 (4.36)
≥60	1261 (10.94)
Unknown	16 (0.14)
Time of delay in diagnosis in the health system (days)	
0-14	10,069 (87.32)
15-29	773 (6.70)
30-44	302 (2.62)
45-59	122 (1.06)
≥60	237 (2.06)
Unknown	28 (0.24)

Characteristic	Patients, n (%)
Interval between diagnosis of tuberculosis pleurisy and end of treatment (days)	
0-180	734 (6.37)
181-270	795 (6.89)
≥271	9949 (86.28)
Unknown	53 (0.46)
Treatment outcome	
Treatment completed	10,422 (90.38)
Death	203 (1.76)
Cure	242 (2.10)
Changes in diagnosis	204 (1.77)
Failure	14 (0.13)
Adverse reactions	122 (1.06)
Transfer to multidrug-resistant tuberculosis pleurisy treatment	10 (0.09)
Others	314 (2.73)

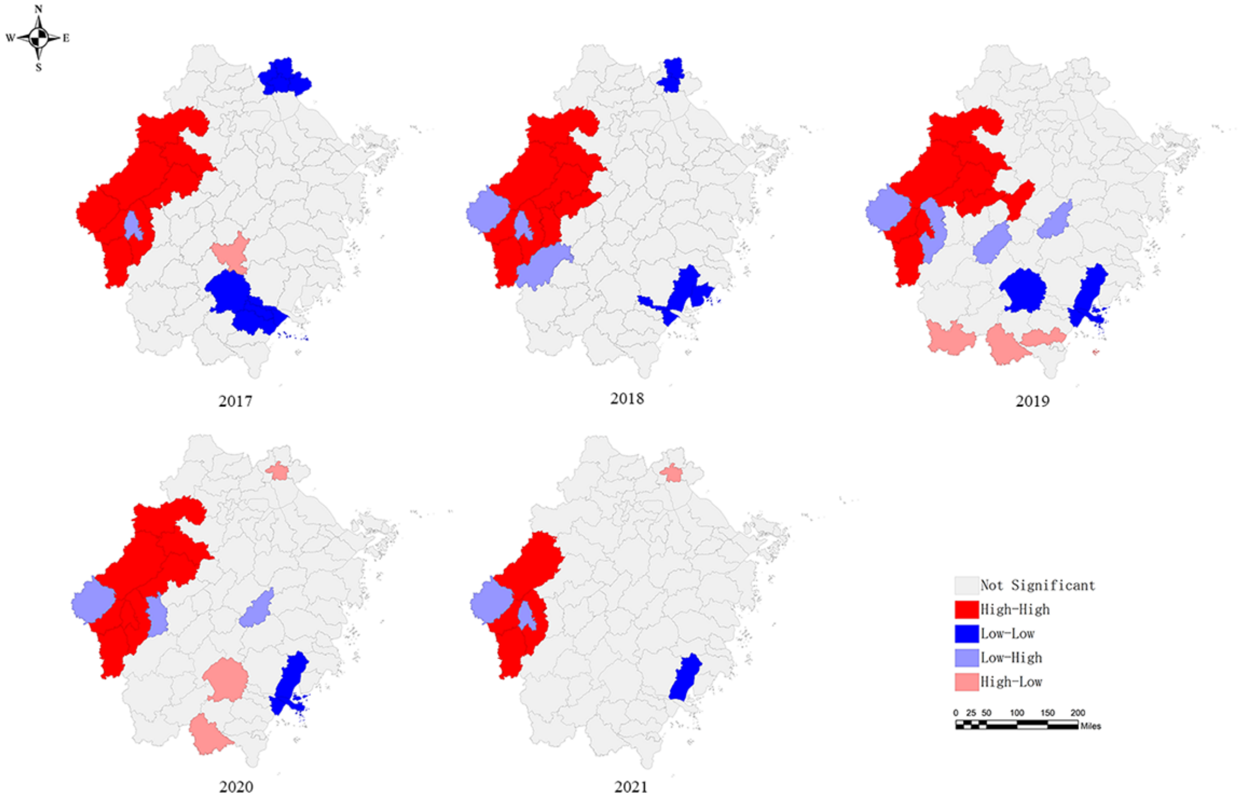
Spatial Autocorrelation Analysis

On spatial autocorrelation analysis, the global Moran I index ranged from 0.187 to 0.362, indicating spatial heterogeneity in the distribution of patients with TP in Zhejiang Province from 2017 to 2021. Furthermore, there was a positive correlation between the notification rate and spatial distribution ($P<.01$; Table 2). The local autocorrelation LISA cluster plot showed that in Zhejiang Province, hot spot areas were located in parts of Quzhou, Jinhua, and Hangzhou, and cold spot areas were clustered in Jiaxing, Wenzhou, and Lishui. Although the hot and cold spots changed every year, Jiangshan, Qujiang, Changshan in Quzhou City, and Chun'an in Hangzhou City emerged as hot spots throughout the study period. Furthermore, Longwan and Yueqing in Wenzhou City were identified as cold spots in 4 out of the 5 years (Figure 2).

Table 2. Spatial autocorrelation analysis of the tuberculosis pleurisy notification rate.

Year	Moran I index	*z* score	*P* value
2017	0.362	5.275	.001
2018	0.296	4.099	.001
2019	0.187	2.632	.007
2020	0.209	3.026	.006
2021	0.241	3.864	.003

Figure 2. Local indicators of spatial association of the tuberculosis pleurisy notification rate in Zhejiang Province.



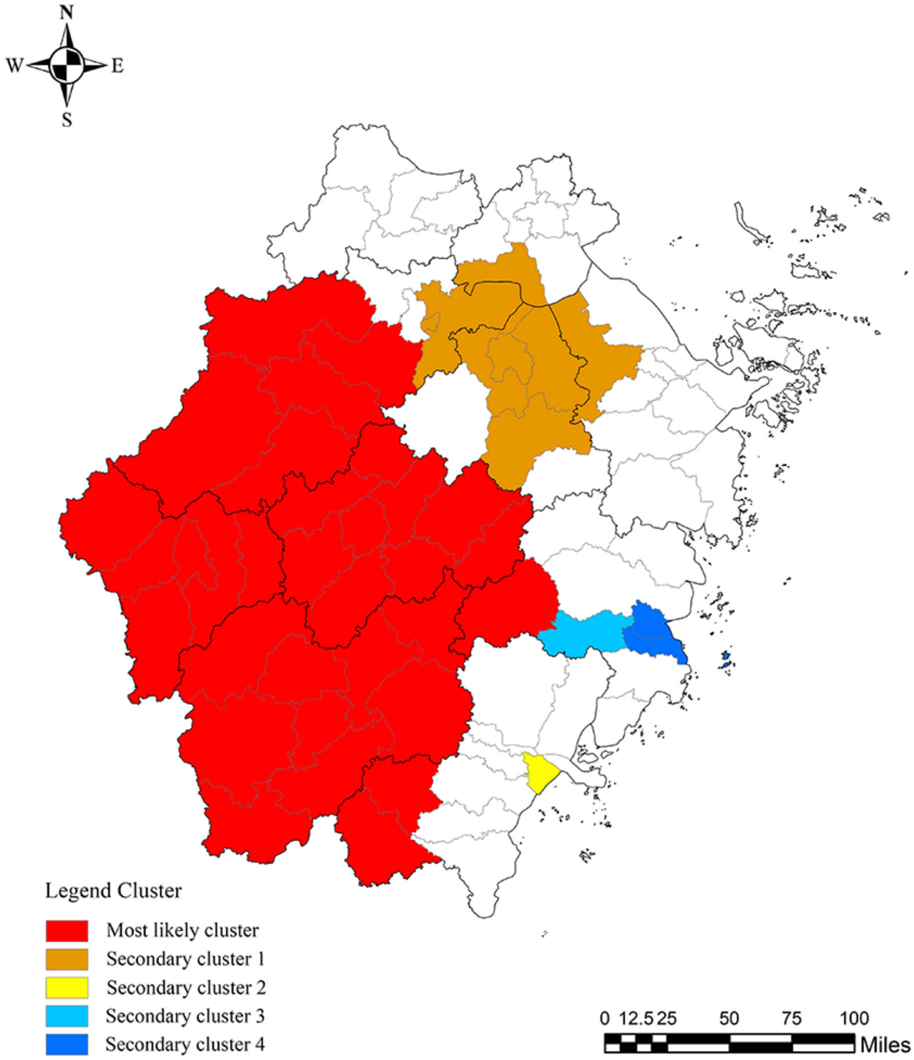
Spatial-Temporal Scan Statistics

Spatial-temporal scan statistics of TP in Zhejiang Province from 2010 to 2021 identified 1 most likely cluster and 4 secondary clusters. The most likely cluster was the high-risk cluster, which encompassed 32 counties in western and southwestern Zhejiang Province, including all counties and districts of Quzhou, Jinhua, and Lishui; Wencheng and Taishun in Wenzhou; and Xianju in Lishui. The most likely cluster period ranged from April 2017 to September 2019, with 2383 cases of TP and a relative risk of 1.98 (log likelihood ratio=382.55, *P*<.01; Table 3 and Figure 3).

Table 3. Spatial-temporal scanning of the tuberculosis pleurisy notification rate in Zhejiang Province, 2017-2021.

Cluster type	Cluster period	Coordinates/radius	Counties, n	Notified (expected) cases, n (n)	Log likelihood ratio	Relative risk	*P* value
Most likely cluster	January 20, 2017, to September 30, 2019	(28.578400° N, 118.595300° E)/199.19 km	32	2383 (1341.84)	382.55	1.98	<.01
Secondary cluster 1	January 1, 2017, to June 30, 2018	(30.042100° N, 120.663400° E)/51.20 km	8	785 (488.22)	80.06	1.65	<.01
Secondary cluster 2	February 1, 2018, to July 31, 2020	(27.885900° N, 120.815300° E)/0 km	1	48 (3.87)	76.84	12.46	<.01
Secondary cluster 3	February 1, 2017, to November 30, 2018	(28.602000° N, 121.076900° E)/0 km	1	124 (48.34)	41.40	2.58	<.01
Secondary cluster 4	April 1, 2017, to August 31, 2018	(28.652100° N, 121.475800° E)/12.06 km	2	129 (74.78)	16.25	1.73	<.01

Figure 3. Spatial-temporal clustering of the tuberculosis pleurisy notification rate in Zhejiang Province.



Model Predictions

On comparing the fitness of predicting TP prevalence trends in 2021, the Holt-Winters exponential smoothing model showed better predictive performance than the SARIMA and Prophet models, with lower mean-square error, root-mean-square error, mean absolute error, mean absolute percentage error, and symmetric mean absolute percentage error values (Table 4). The composition of the original sequence and the predictive results determined using the Holt-Winters exponential smoothing model are presented in Table 4. The exponential smoothing model was used to predict the number of notified cases of TP, and its actual values in 2021 are shown in Table 5.

Table 4. Critical indices of the Holt-Winters exponential smoothing, SARIMA[a], and Prophet models.

Indices	SARIMA	Holt-Winters	Prophet
MSE[b]	978.593	479.170	579.114
RMSE[c]	31.282	21.890	24.065
MAE[d]	25.247	17.380	17.501
MAPE[e]	16.270	11.130	N/A[f]
SMAPE[g]	16.883	12.080	N/A

[a] SARIMA: seasonal autoregressive integrated moving average.

[b] MSE: mean-square error.

[c] RMSE: root-mean-square error.

[d] MAE: mean absolute error.

[e] MAPE: mean absolute percentage error.

[f] N/A: not applicable.

[g] SMAPE: symmetric mean absolute percentage error.

Table 5. Holt-Winters exponential smoothing model used to predict the number of notified cases of tuberculosis pleurisy and its actual values in 2021.

Month	Actual value	Estimate	95% CI
January	158	141	104-178
February	122	119	105-133
March	191	162	161-163
April	159	176	134-217
May	165	175	140-210
June	155	156	116-197
July	163	156	132-181
August	144	129	99-159
September	145	128	116-140
October	133	120	82-157
November	145	115	108-121
December	160	108	107-108

Discussion

Principal Findings

TP, as a subgroup of PTB in China, has attracted increasing attention in recent years due to its severity. According to a bibliometric study, there has been an increasing number of TP-related studies in the past decade, and China had the most studies in this field [27]; furthermore, Mexico, the Philippines, Vietnam, China, and India had the highest prevalences of TP worldwide [5]. Zhejiang Province is a developed coastal area of China, and the notification rate of TP accounts for a certain proportion. Therefore, it is necessary to understand cluster epidemics at the spatial and temporal levels.

In this study, epidemiological characteristics showed that men were more likely to develop TP than women, which is in line with the results of previous studies [28,29]. Given that TP occurs secondary to PTB, we speculated that possible explanations for sex differences were attributed to lifestyle and genetic susceptibility, such as smoking, alcohol use, drug abuse, sex hormones, and genetic variations [30,31]. Therefore, it is necessary to consider men as the key population for prevention and control, such as by strengthening screening and improving the awareness of TP in this specific population. Furthermore, we found that the TP notification rate peaked mainly in the summer. As shown in the literature, when the mean ambient temperature increased from 15.1 °C to 24.5 °C, the cumulative adjusted relative risk of hospitalization for TB was 1.07 (95% CI 1.00-1.14) [32]. Although there is no clear evidence showing a direct correlation between TP and temperature, we postulate that due to the increased temperature, there is a likely increase in the synthesis of mycolic acids, which are vital for the growth of *M tuberculosis* and contribute to its subverting of and escaping from the immune system [33].

In our study, 6317 (54.78%) patients with TP experienced delayed access to medical care. A literature analysis of TB cases from 32 provinces in China showed that patient delays were also common among TB cases, with a median of 20 (IQR 6-46) days [21]. We speculated that a comparatively long patient delay may result in further invasion of the pleura. Furthermore, the median delay in PTB diagnosis in China was 1 (IQR 0-8) day, while a delay of TP diagnosis of more than 14 days accounted for 12.68% of cases in our study [21]. To our knowledge, a definite diagnosis of TP was made by detecting *M tuberculosis* from pleural effusion or pleural tissue [34]. However, culturing *M tuberculosis* takes 2-8 weeks to obtain results in some regions without access to rapid diagnosis tools, which could influence effective medical interventions [35]. Consequently, further delay in anti-TB treatment for TP might result in pleural thickening or TB empyema that requires surgical resolution [36]. Considering the long duration of the disease, insidious onset, and lack of specificity of diagnostic methods for TP [37,38], it is prone to both abovementioned delays, which may further lead to increased transmission of TB [39]. Therefore, timely detection, such as using x-ray imaging or computed tomography, and adequate health education are needed to identify TP in the early stages and improve awareness of TB information [40]. For patients with normal immune function, the current recommended treatment duration is 6 months for isoniazid and rifampicin with 2 months for ethambutol and pyrazinamide [8]. In China, the National Tuberculosis Program guidelines specifically recommend that the treatment period for TP be ≥9 months, while nearly 14% of patients had received insufficient treatment. Therefore, more attention must be paid to the treatment duration of TP.

Spatial-temporal analysis is commonly used to explore disease clusters and identify areas at high risk of infectious disease [41]. Spatial analysis indicated spatial heterogeneity among patients with TP, with hot spots located mainly in the western cities of Zhejiang Province. The relatively large agricultural population and relatively low economic status in these areas, along with limited knowledge related to TP, are the common risk factors for the development of TP [42]. Therefore, the early detection strategy of TP in this population should be strengthened by implementing modern approaches, such as large-scale active screening. Furthermore, the spatial-temporal analysis indicated that the most likely clusters were mainly concentrated in the western part of Zhejiang Province, which were similar to the results of spatial analysis. This result unveils the distribution pattern of TP incidence in Zhejiang Province and provides a theoretical basis for the regional rational allocation of health resources related to TP and effective implementation of comprehensive prevention and control measures [41]. Therefore, exploring the causes and mechanisms leading to the regional clustering of TP should be carried out in the future.

Each predictive model has its own limitations and merits due to the different prerequisites and mechanisms involved in predicting epidemics. In this study, our exponential smoothing model has a good prediction effect, and the actual value of 66.7% falls within the 95% CI of the predicted value. Part of the actual value that is not within the prediction range may be unsatisfactory for clinical applications. The possible reason was the impact of the COVID-19 pandemic, resulting in the deviation of the model prediction.

Limitations

However, this study also has some limitations. First, TP is not easy to diagnose, and it was not mandatory to include TBIMS before 2018, which could cause an underestimation of TP prevalence from our available surveillance system. Second, the combination and breakdown of different areas of the panel data makes it impossible to carry out detailed analysis in some areas. Third, the prediction model did not include other meteorological factors such as humidity and wind. Last, some potential influencing factors were not considered, such as HIV coinfection, which would be considered in the future.

Conclusions

In this study, the male population had a high notification rate of TP, and a high notification rate of this disease was noted during summer. The western region of Zhejiang Province has the highest relative risk of TP. Comprehensive interventions, such as chest x-ray screening and symptom screening, should be strengthened to improve early identification. Additionally, a more systematic assessment of the prevalence trend of TP should include more predictors.

Acknowledgments

We appreciate the local centers for disease control and prevention, community health care centers, and tuberculosis (TB)-designated hospitals in Zhejiang Province to implement TB, including tuberculosis pleurisy, reporting and notification during daily work. This study was supported by the National-Zhejiang Health Commission Major S&T Project (grant WKJ-ZJ-2118), Zhejiang Provincial Medical and Health Project (2021KY618 and 2020KY520).

Data Availability

All data and materials were included in this paper. The corresponding author (JJ) can provide data upon reasonable request after all studies and substudies have been completed.

Authors' Contributions

JP, BC, and KL conceptualized and designed the study. ZL and JJ managed the study. DL and YZ analyzed the data. SC and BC interpreted the study results. YZ, JJ, and DL drafted the manuscript for important intellectual content. KL, JP, and JJ reviewed and edited the manuscript. YZ, DL, KL, BC, SC, JP, ZL, and JJ approved the final version of the manuscript to be published.

Conflicts of Interest

None declared.

References

1. Global Tuberculosis Report 2022. World Health Organization. 2022. URL: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 [accessed 2023-10-05]
2. Boonsarngsuk V, Mangkang K, Santanirand P. Prevalence and risk factors of drug-resistant extrapulmonary tuberculosis. Clin Respir J 2018 Jun;12(6):2101-2109. [doi: 10.1111/crj.12779] [Medline: 29424965]
3. Ryan H, Yoo J, Darsini P. Corticosteroids for tuberculous pleurisy. Cochrane Database Syst Rev 2017 Mar 14;3(3):CD001876 [FREE Full text] [doi: 10.1002/14651858.CD001876.pub3] [Medline: 28290161]
4. Jeon D. Tuberculous pleurisy: an update. Tuberc Respir Dis (Seoul) 2014 Apr;76(4):153-159 [FREE Full text] [doi: 10.4046/trd.2014.76.4.153] [Medline: 24851127]
5. Baumann MH, Nolan R, Petrini M, Lee YCG, Light RW, Schneider E. Pleural tuberculosis in the United States: incidence and drug resistance. Chest 2007 Apr;131(4):1125-1132. [doi: 10.1378/chest.06-2352] [Medline: 17426219]
6. People’s Republic of China state health and Family Planning Commission. Tuberculosis classification (WS196—2017). Electronic J Emerg Infect Dis 2018 Aug 31;3:191-192. [doi: 10.19871/j.cnki.xfcrbzz.2018.03.018]
7. National Health and Family Planning Commission of the People's Republic of China. Diagnostic criteria for tuberculosis. Electronic J Emerg Infect Dis 2018 Feb 28;3:59-61. [doi: 10.19871/j.cnki.xfcrbzz.2018.01.017]
8. Shaw JA, Koegelenberg CFN. Pleural Tuberculosis. Clin Chest Med 2021 Dec;42(4):649-666. [doi: 10.1016/j.ccm.2021.08.002] [Medline: 34774172]
9. Frye MD, Pozsik CJ, Sahn SA. Tuberculous pleurisy is more common in AIDS than in non-AIDS patients with tuberculosis. Chest 1997 Aug;112(2):393-397. [doi: 10.1378/chest.112.2.393] [Medline: 9266874]
10. Wen P, Wei M, Han C, He Y, Wang M. Risk factors for tuberculous empyema in pleural tuberculosis patients. Sci Rep 2019 Dec 20;9(1):19569 [FREE Full text] [doi: 10.1038/s41598-019-56140-4] [Medline: 31863026]
11. Chen S, Wang Y, Zhan Y, Liu C, Wang Q, Feng J, et al. The incidence of tuberculous pleurisy in mainland China from 2005 to 2018. Front Public Health 2023;11:1180818 [FREE Full text] [doi: 10.3389/fpubh.2023.1180818] [Medline: 37397728]
12. Mohamed AS, Levine M, Camp JW, Lund E, Yoder JS, Glickman LT, et al. Temporal patterns of human and canine Giardia infection in the United States: 2003-2009. Prev Vet Med 2014 Feb 01;113(2):249-256. [doi: 10.1016/j.prevetmed.2013.11.006] [Medline: 24309130]
13. Gao S, Mioc D, Anton F, Yi X, Coleman DJ. Online GIS services for mapping and sharing disease information. Int J Health Geogr 2008 Feb 25;7:8 [FREE Full text] [doi: 10.1186/1476-072X-7-8] [Medline: 18298859]
14. Rogers DJ, Randolph SE. Studying the global distribution of infectious diseases using GIS and RS. Nat Rev Microbiol 2003 Dec;1(3):231-237 [FREE Full text] [doi: 10.1038/nrmicro776] [Medline: 15035027]
15. Luo Z, Jia X, Bao J, Song Z, Zhu H, Liu M, et al. A combined model of SARIMA and prophet models in forecasting AIDS incidence in Henan Province, China. Int J Environ Res Public Health 2022 May 12;19(10) [FREE Full text] [doi: 10.3390/ijerph19105910] [Medline: 35627447]
16. Wang S, Wei F, Li H, Wang Z, Wei P. Comparison of SARIMA model and Holt-Winters model in predicting the incidence of Sjögren's syndrome. Int J Rheum Dis 2022 Nov;25(11):1263-1269. [doi: 10.1111/1756-185X.14417] [Medline: 35962522]
17. Liu K, Li T, Vongpradith A, Wang F, Peng Y, Wang W, et al. Identification and prediction of tuberculosis in Eastern China: analyses from 10-year population-based notification data in Zhejiang Province, China. Sci Rep 2020 May 04;10(1):7425 [FREE Full text] [doi: 10.1038/s41598-020-64387-5] [Medline: 32367050]
18. Liu K, Chen S, Zhang Y, Li T, Xie B, Wang W, et al. Tuberculosis burden caused by migrant population in Eastern China: evidence from notification records in Zhejiang Province during 2013-2017. BMC Infect Dis 2022 Jan 31;22(1):109 [FREE Full text] [doi: 10.1186/s12879-022-07071-5] [Medline: 35100983]
19. Li T, Du X, Shewade HD, Soe KT, Zhang H. What happens to migrant tuberculosis patients who are transferred out using a web-based system in China? PLoS One 2018;13(11):e0206580 [FREE Full text] [doi: 10.1371/journal.pone.0206580] [Medline: 30408131]
20. Huang Y, Zuo L, Ultrasound Professional Committee of Tuberculosis Branch of Chinese Medical Association, Interventional Ultrasound Professional Committee of Interventional Physician Branch of Chinese Medical Doctor Association. Expert consensus on ultrasound diagnosis, classification and interventional therapy of tuberculous pleurisy (2022 Edition). Chinese Journal of Antituberculosis 2022 Sep 10:880-897. [doi: 10.19982/j.issn.1000-6621.20220257]
21. Li T, Du X, Kang J, Luo D, Liu X, Zhao Y. Patient, diagnosis, and treatment delays among tuberculosis patients before and during COVID-19 epidemic - China, 2018-2022. China CDC Wkly 2023 Mar 24;5(12):259-265 [FREE Full text] [doi: 10.46234/ccdcw2023.047] [Medline: 37138894]
22. Li M, OU J. Spatio-temporal analysis of influenza A(H1N1) in China during 2009-2013 based on GIS. Geogr Res 2016 Nov 18:2139-2152. [doi: 10.11821/dlyj201611011]

23. Yamamura M, de Freitas IM, Santo Neto M, Chiaravalloti Neto F, Popolin MAP, Arroyo LH, et al. Spatial analysis of avoidable hospitalizations due to tuberculosis in Ribeirao Preto, SP, Brazil (2006-2012). Rev Saude Publica 2016;50:20 [FREE Full text] [doi: 10.1590/S1518-8787.2016050006049] [Medline: 27191156]
24. Ndeh NT, Tesfaldet YT, Budnard J, Chuaicharoen P. The secondary outcome of public health measures amidst the COVID-19 pandemic in the spread of other respiratory infectious diseases in Thailand. Travel Med Infect Dis 2022;48:102348 [FREE Full text] [doi: 10.1016/j.tmaid.2022.102348] [Medline: 35523394]
25. Ogallo W, Wanyana I, Tadesse GA, Wanjiru C, Akinwande V, Kabwama S, et al. Quantifying the impact of COVID-19 on essential health services: a comparison of interrupted time series analysis using Prophet and Poisson regression models. J Am Med Inform Assoc 2023 Mar 16;30(4):634-642. [doi: 10.1093/jamia/ocac223] [Medline: 36534893]
26. Xie C, Wen H, Yang W, Cai J, Zhang P, Wu R, et al. Trend analysis and forecast of daily reported incidence of hand, foot and mouth disease in Hubei, China by Prophet model. Sci Rep 2021 Jan 14;11(1):1445 [FREE Full text] [doi: 10.1038/s41598-021-81100-2] [Medline: 33446859]
27. Bian Y, Deng M, Zhang Q, Hou G. Global trends of research on tuberculous pleurisy over the past 15 years: a bibliometric analysis. Front Cell Infect Microbiol 2022;12:937811 [FREE Full text] [doi: 10.3389/fcimb.2022.937811] [Medline: 36111237]
28. Tahseen S, Khanzada FM, Baloch AQ, Abbas Q, Bhutto MM, Alizai AW, et al. Extrapulmonary tuberculosis in Pakistan-A nation-wide multicenter retrospective study. PLoS One 2020;15(4):e0232134 [FREE Full text] [doi: 10.1371/journal.pone.0232134] [Medline: 32343714]
29. Luo Y, Yan F, Xue Y, Mao L, Lin Q, Tang G, et al. Diagnostic utility of pleural fluid T-SPOT and interferon-gamma for tuberculous pleurisy: a two-center prospective cohort study in China. Int J Infect Dis 2020 Oct;99:515-521 [FREE Full text] [doi: 10.1016/j.ijid.2020.08.007] [Medline: 32777584]
30. Khan AH, Sulaiman SAS, Hassali MA, Khan KU, Ming LC, Mateen O, et al. Effect of smoking on treatment outcome among tuberculosis patients in Malaysia; a multicenter study. BMC Public Health 2020 Jun 04;20(1):854 [FREE Full text] [doi: 10.1186/s12889-020-08856-6] [Medline: 32498682]
31. Kirenga BJ, Ssengooba W, Muwonge C, Nakiyingi L, Kyaligonza S, Kasozi S, et al. Tuberculosis risk factors among tuberculosis patients in Kampala, Uganda: implications for tuberculosis control. BMC Public Health 2015 Jan 21;15:13 [FREE Full text] [doi: 10.1186/s12889-015-1376-3] [Medline: 25604986]
32. Chong KC, Yeoh EK, Leung CC, Lau SYF, Lam HCY, Goggins WB, et al. Independent effect of weather, air pollutants, and seasonal influenza on risk of tuberculosis hospitalization: an analysis of 22-year hospital admission data. Sci Total Environ 2022 Sep 01;837:155711. [doi: 10.1016/j.scitotenv.2022.155711] [Medline: 35523336]
33. Takayama K, Armstrong EL, Davidson LA, Kunugi KA, Kilburn JO. Effect of low temperature on growth, viability, and synthesis of mycolic acids of Mycobacterium tuberculosis strain H37Ra. Am Rev Respir Dis 1978 Jul;118(1):113-117. [doi: 10.1164/arrd.1978.118.1.113] [Medline: 98081]
34. Jeon D. Tuberculous pleurisy: an update. Tuberc Respir Dis (Seoul) 2014 Apr;76(4):153-159 [FREE Full text] [doi: 10.4046/trd.2014.76.4.153] [Medline: 24851127]
35. Ritchie SR, Harrison AC, Vaughan RH, Calder L, Morris AJ. New recommendations for duration of respiratory isolation based on time to detect Mycobacterium tuberculosis in liquid culture. Eur Respir J 2007 Sep;30(3):501-507 [FREE Full text] [doi: 10.1183/09031936.00131406] [Medline: 17537768]
36. Barbas CS, Cukier A, de Varvalho CR, Barbas Filho JV, Light RW. The relationship between pleural fluid findings and the development of pleural thickening in patients with pleural tuberculosis. Chest 1991 Nov;100(5):1264-1267. [doi: 10.1378/chest.100.5.1264] [Medline: 1935279]
37. Liu YJ, Jiang Z, Chen H, Jing H, Cao X, Coia JE, et al. Description of demographic and clinical characteristics of extrapulmonary tuberculosis in Shandong, China. Hippokratia 2020;24(1):27-32 [FREE Full text] [Medline: 33364736]
38. Lee C, Chiu L, Chang C, Chung F, Li S, Chou C, et al. The clinical experience of Mycobacterial culture yield of pleural tissue by pleuroscopic pleural biopsy among tuberculous pleurisy patients. Medicina (Kaunas) 2022 Sep 15;58(9) [FREE Full text] [doi: 10.3390/medicina58091280] [Medline: 36143957]
39. Zhongwei J, Zuhong L. Effect of delayed diagnosis and treatment on transmission of tuberculosis. J Syst Sci Math Sci 2016 Jan 15;36(1):28.
40. Jiang J, Zhang X, Wang F. Influencing factors for the delay in seeking health care of pulmonary tuberculosis in the elderly. Prev Med 2020;6:612-614. [doi: 10.19485/j.cnki.issn2096-5087.2020.06.018]
41. Qian Y, Kun C. Spatial distribution patterns of pulmonary tuberculosis incidence in Zhejiang province: spatial autocorrelation analysis. Chin J Public Health 2013 Apr 30;29(4):485-487.
42. Gui JJ, Zhang TF, Liu ZF. Epidemiological characteristics and spatial clusters of pulmonary tuberculosis in Zhejiang province,2005-2011. Chin J Public Health 2016;32(1):11-14. [doi: 10.11847/zgggws2016-32-01-04]

Abbreviations

LISA: local indicators of spatial association
PTB: pulmonary tuberculosis

SARIMA: seasonal autoregressive integrated moving average
TB: tuberculosis
TP: tuberculous pleurisy

Edited by A Mavragani, T Sanchez; submitted 12.06.23; peer-reviewed by T Li, W Tang, Y Liu; comments to author 19.08.23; revised version received 31.08.23; accepted 19.09.23; published 30.10.23.

Please cite as:
Zhou Y, Luo D, Liu K, Chen B, Chen S, Pan J, Liu Z, Jiang J
Trend of the Tuberculous Pleurisy Notification Rate in Eastern China During 2017-2021: Spatiotemporal Analysis
JMIR Public Health Surveill 2023;9:e49859
URL: https://publichealth.jmir.org/2023/1/e49859
doi:10.2196/49859
PMID:37902822

Original Paper

Measurement of Health-Related Quality of Life in Individuals With Rare Diseases in China: Nation-Wide Online Survey

Richard Huan Xu[1], PhD; Shamay S M Ng[1], Prof Dr; Nan Luo[2], PhD; Dong Dong[3,4], PhD; Shuyang Zhang[5], Prof Dr

[1]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Kowloon, China (Hong Kong)
[2]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore
[3]JC School of Public Health and Primary Care, The Chinese University of Hong Kong, New Territories, China (Hong Kong)
[4]Shenzhen Research Institute, The Chinese University of Hong Kong, Sehnzhen, China
[5]Department of Cardiology, Peking Union Medical College and Chinese Academy of Medical Science, Beijing, China

Corresponding Author:
Dong Dong, PhD
JC School of Public Health and Primary Care
The Chinese University of Hong Kong
Shatin
New Territories,
China (Hong Kong)
Phone: 852 22528488
Email: dongdong@cuhk.edu.hk

Abstract

Background: Rare diseases (RDs) affect millions of people worldwide, and these diseases can severely impact the health-related quality of life (HRQoL) of those affected. Despite this, there is a lack of research measuring HRQoL using the EQ-5D-5L, which is one of the most widely used generic preference-based instruments to measure HRQoL in populations living with RDs.

Objective: This study aimed to measure HRQoL using the EQ-5D-5L in a large number of patients with various types of RDs in China, and to examine the relationship between respondents' socioeconomic characteristics and preference-based health utility scores.

Methods: The data used in this study were obtained from a web-based survey conducted in China. The survey aimed to explore and understand the health and socioeconomic status of patients with RDs in China. We recruited registered and eligible members, including patients or their primary caregivers, from 33 RD patient associations to complete the questionnaires via their internal social networks. HRQoL was measured using the EQ-5D-5L utility score, which was calculated based on an established Chinese value set. Utility scores have been presented based on demographics and disease-related information. Univariate linear regression analysis was used to assess the differences in the EQ-5D-5L utility scores between subgroups.

Results: A total of 12,502 respondents completed the questionnaire and provided valid responses, including 6919 self-completed respondents and 5583 proxy-completed respondents. Data from 10,102 participants over the age of 12 years were elicited for analysis. Among patients with RDs, 65.3% (6599/10,102), 47.5% (4799/10,102), 47.0% (4746/10,102), 24.8% (2506/10,102), and 18.4% (1855/10,102) reported no problems for "self-care," "usual activities," "mobility," "pain/discomfort," and "anxiety/depression," respectively. A full health state was reported by 6.0% (413/6902) and 9.2% (295/3200) of self- and proxy-completed patients, respectively. Among self-completed patients, 69.9% (4826/6902) and 50.4% (3478/6902) reported no problems for "self-care" and "usual activities," respectively, whereas only 17.7% (1223/6902) reported problems for "anxiety/depression." Proxy-completed respondents showed a higher proportion of reporting extreme problems than self-completed respondents in all 5 dimensions. The mean utility scores reported by self- and proxy-completed respondents were 0.691 and 0.590, respectively. Different types of caregivers reported different utility scores, and among them, proxy-completed (mother) respondents reported the highest mean utility score.

Conclusions: The establishment of a normative profile for RD patients can facilitate patients' adaptation and assess the effectiveness of interventions to improve the HRQoL and well-being of this population. Differences between self- and proxy-completed HRQoL assessed by the EQ-5D-5L have been identified in this study. This finding highlights the importance of incorporating perspectives from both patients and their proxies in clinical practice. Further development of the patient cohort is necessary to assess long-term changes in HRQoL in the RD population.

(JMIR Public Health Surveill 2023;9:e50147) doi:10.2196/50147

KEYWORDS

EQ-5D-5L; rare disease; normative profile; utility score; caregiver

Introduction

Rare diseases (RDs) are conditions that affect a very small proportion of the population worldwide. Currently, there are approximately 7000 types of RDs, of which over 70% are genetic, and most of them show a chronic course. The prevalence threshold and definition of RDs vary across countries. In the United States and European Union, approximately 25-30 million and 45-60 million people experience RDs, respectively [1,2]. Worldwide, the population prevalence of RDs is estimated to be 3.5%-5.9%, which amounts to approximately 260-440 million people [3]. Since effective treatments for most RDs do not exist, RD patients typically require complex care, resulting in poor health status [4]. Poor access to information, high treatment expenses, and lack of social support can have a significant negative impact on their health-related quality of life (HRQoL) [5].

The EQ-5D is a well-known and widely used health status instrument worldwide. It was developed by the EuroQol Group as a concise, generic, preference-based instrument for measuring, comparing, and evaluating health status in the population aged 12 years or above [6]. It describes a person’s health as a multi-dimensional profile and provides a single utility value for it, which can be used as a HRQoL measure to facilitate the cost-utility analysis of health care and monitoring of population health [7]. The performance of the EQ-5D has been assessed in numerous health conditions and treatments, and it has shown acceptable psychometric properties. Currently, 2 versions of the EQ-5D exist: the 3-level option (3L) and the 5-level option (5L). The EQ-5D-5L is recommended owing to its better performance in reducing the “ceiling effect” compared to that of the EQ-5D-3L [8]. Currently, there is limited research on the use of the EQ-5D to measure HRQoL in the RD population as a whole. For example, Serrano-Aguilar et al used the EQ-5D for assessing HRQoL in approximately 3000 European patients and caregivers with 10 RDs [9]. Efthymiadou et al assessed HRQoL in more than 600 patients with RDs who were mostly from Europe and the United States [10], and Ng et al evaluated HRQoL in 286 patients with RDs in Hong Kong [11].

There are several gaps in the existing literature on HRQoL among individuals with RDs. First, there is a lack of research on HRQoL among Chinese patients with RDs. Currently, there are approximately 70 million Chinese patients with RDs [12], and without sufficient data and research, providing proper treatment and care for these patients is challenging. Second, although studies have recruited relatively large samples of patients with RDs in Europe, no studies with a similar sample size have been conducted in Asian countries. Random variation can have a larger impact on results in the case of small samples, potentially leading to spurious findings that might not be supported by the data. Third, no studies used the EQ-5D-5L to assess HRQoL in patients with RDs as a whole in the Chinese population. Lastly, no studies have reported the inclusion of patients with advanced-stage RDs (either subjective or objective rating). This may have generated obvious selection bias and significantly affected the quality and reliability of the findings. Therefore, the objective of this study was to measure the HRQoL of individuals or caregivers with various types of RDs using a large sample in China. Specifically, we aimed to establish a normative reference of the EQ-5D-5L for the RD population, including individuals who are older than 12 years, as a whole.

Methods

Sample and Data Collection

The data used in this study were obtained from a nationwide, cross-sectional, web-based survey conducted between August 2019 and January 2020. The survey aimed to explore and understand the health and socioeconomic status of patients with RDs in China. The research team comprised members from the China Alliance for Rare Disease, Chinese University of Hong Kong, China Illness Challenge Foundation, and Peking Union Medical College Hospital. In total, 32 patient associations (PAs) for RDs participated in the study. These PAs were selected from a list provided by the National Rare Diseases Registry System of China and the National Network to Collaborate on Diagnosis and Treatment of Rare Diseases. These 2 organizations were established by the Peking Union Medical College Hospital and supported by the National Health Commission of the People’s Republic of China to manage patients with RDs in China. The content of the questionnaire and logic of the survey were confirmed by the research team and representatives of all the PAs via 2 rounds of face-to-face conferences and more than 10 rounds of web-based meetings between June and August 2019.

All participants were registered members of a PA, which indicates that they were either patients with a formal diagnosis or primary caregivers of patients with a formal diagnosis. In this study, a primary caregiver was defined as a person aged ≥18 years who has been designated as a necessary caretaker responsible for managing the patient’s well-being. This may include, but is not limited to, parents, legal guardians, and paid caregivers. Since the EQ-5D-5L is not appropriate for measuring HRQoL in individuals younger than 12 years, only data from patients with RDs who were 12 years or older were elicited for analysis. Patients without cognitive impairments, who could read Chinese and provide informed consent, were encouraged to complete the questionnaire on their own. However, for all eligible participants, the option of proxy completion by caregivers was also available.

Invitations were disseminated to all eligible patients or caregivers via the PAs’ internal social networks. Eligible members who were interested in participating in the survey were invited to join an online survey group, where the study aims, process, and expected results were provided via group communication. The survey link was sent to all survey group members. The research team worked with the PA’s staff to manage the survey and remind the participants to complete it within 10 days. They sent reminders to participants on the second, third, fifth, and seventh days after the survey link was sent. The first page of the questionnaire was the informed consent form. All participants were required to read through it and click the “Agree” button at the end of the page before starting the survey. Information about participants’ background characteristics, HRQoL, symptoms, social support, and medication was collected. The research team could not access personally identifiable information of the participants; only the PAs’ staff could access such information.

Ethics Approval

The study protocol and informed consent form were approved by the Institutional Review Board of the Chinese University of Hong Kong (Reference ID: SBRE-18-268).

Instrument

Both the self- and proxy-completed versions of the EQ-5D-5L were used in this study [13]. The first section of the EQ-5D-5L is a descriptive system, which comprises 5 health-related dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) rated on a 5-option Likert scale ranging from no problems to extreme problems. All health states described by the descriptive system could be converted into a single utility score using a scoring algorithm based on public preferences. Utility scores are anchored between 1 (full health) and 0 (a state as bad as being dead) as required by their use in economic evaluation. A utility score of less than 0 represents health states regarded as worse than a state that is as bad as being dead. In this study, the EQ-5D-5L China value set and scoring algorithm were used [14]. The second section of the EQ-5D-5L is a visual analog scale (EQ-VAS). It is a scale that takes values between 100 (best imaginable health) and 0 (worst imaginable health), on which patients provide a global assessment of their health. The main difference between the self- and proxy-completed versions of the EQ-5D-5L is the adopted perspective. The proxy version asked respondents to report how the care receiver would rate his or her own health.

Statistical Analysis

A descriptive analysis was applied for the participants’ demographic and health characteristics. Because self- and proxy-completed HRQoL may be systematically different, self-completed and proxy-completed EQ-5D-5L data were analyzed separately. Continuous variables (eg, EQ-5D-5L utility score) have been presented as mean, median, and SD. Categorical variables (eg, response on EQ-5D-5L dimensions) have been presented as frequency (n) and proportion (%). Pediatric patients were defined as those aged ≤18 years, whereas older patients were defined as those aged ≥60 years. Self-completed patients were defined as those who completed the questionnaire by themselves (must be ≥12 years), and proxy-completed patients (<12 years or ≥12 years but with very poor health status) were defined as those whose primary caregivers completed the questionnaire. The EQ-5D-5L profile was presented as responses on each dimension of the health classification descriptive system. The EQ-5D-5L utility score was presented for subgroups of patients stratified by their characteristics, including sex, age group, educational level, employment, family registry, family annual income, disease duration, number of children, assistive devices used in daily life, disability status, and number of family members living together, as well as the types of RDs. The univariate linear regression analysis was used to assess differences in the EQ-5D-5L utility score between subgroups. In addition, the differences in the mean EQ-5D-5L utility scores for different caregiver types (father, mother, children, spouse, grandparents/relatives, or others) stratified by background characteristics have been presented. All statistical analyses were performed using R software. Statistical significance was considered if $P<.05$ was obtained.

Results

Background Characteristics of Patients With RDs

A total of 20,802 responses were collected. Of these, 8300 were excluded because the respondents either dropped out of the survey or provided incomplete responses. The remaining 12,502 respondents included 6919 self-completed patients and 5583 proxy-completed patients. Table 1 shows their background characteristics stratified by response type. Of the patients, 53.2% (6653/12,502) were of the male sex, 16.2% (2016/12,502) were aged <10 years, and 3.8% (469/12,502) were aged ≥61 years, and more than half (6488/12,502, 51.9%) were urban residents. Table 1 provides a detailed breakdown of the differences. Most of the RD types were never reported previously worldwide, for example, Kallmann syndrome, Dravet syndrome, and Niemann–Pick disease. Among them, 2249 patients had myasthenia gravis, 1504 had hemophilia, and 994 had scleroderma (Multimedia Appendix 1). The self-completed and proxy-completed patients significantly differed in all the background characteristics.

Table 1. Patient background characteristics.

Characteristic	Full sample (N=12,502)	Self-completed sample (n=6919)	Proxy-completed sample (n=5583)	*P* value
Gender, n (%)				<.001
Male	6653 (53.2)	3224 (46.6)	3429 (61.4)	
Female	5849 (46.7)	3695 (53.4)	2154 (38.6)	
Age (years), n (%)				<.001
≤10	2020 (16.2)	4 (0.1)	2016 (36.3)	
11-20	2111 (16.9)	390 (5.6)	1721 (31.0)	
21-30	2570 (20.6)	2128 (30.8)	442 (8.0)	
31-40	2728 (21.9)	2367 (34.3)	361 (6.5)	
41-50	1621 (13.0)	1319 (19.1)	302 (5.4)	
51-60	943 (7.6)	543 (7.9)	400 (7.2)	
≥61	469 (3.8)	146 (2.3)	313 (5.6)	
Missing	40 (0.3)	12 (0.1)	28 (0.2)	
Employment, n (%)				<.001
Active	3578 (28.6)	3019 (43.6)	559 (10.0)	
Nonactive	8924 (71.4)	3900 (56.4)	5024 (90.0)	
Family registry, n (%)				<.001
Urban	6488 (51.9)	3783 (54.7)	2705 (48.5)	
Rural	5991 (48.0)	3120 (45.1)	2871 (51.4)	
Missing	23 (0.1)	7 (0.2)	16 (0.3)	
Family income per year (CNY[a]), n (%)				<.001
≤5000	705 (5.6)	433 (6.3)	272 (4.9)	
5001-10,000	782 (6.3)	427 (6.2)	355 (6.4)	
10,001-30,000	2584 (20.7)	1367 (19.8)	1217 (21.8)	
30,001-50,000	2923 (23.4)	1459 (21.1)	1464 (26.2)	
50,001-100,000	3228 (25.8)	1823 (26.3)	1405 (25.2)	
100,001-200,000	1554 (12.4)	955 (13.8)	599 (10.7)	
200,001-300,000	417 (3.3)	263 (3.8)	154 (2.8)	
300,001-500,000	196 (1.6)	125 (1.8)	71 (1.3)	
≥500,001	113 (0.9)	67 (1.0)	46 (0.8)	
Duration of RDs[b] (years), n (%)				<.001
≤10	4709 (37.6)	2681 (38.7)	2028 (36.3)	
11-20	4909 (39.2)	2214 (32.0)	2695 (48.3)	
21-30	1545 (12.3)	1145 (16.5)	400 (7.2)	
31-40	696 (5.5)	593 (8.6)	1032 (18.5)	
≥41	302 (2.4)	260 (3.8)	42 (0.8)	
Missing	341 (2.7)	26 (0.4)	315 (5.6)	
Number of children, n (%)				<.001
0	7230 (57.8)	2968 (42.9)	4262 (76.3)	
1	3516 (28.1)	2816 (40.7)	700 (12.5)	
2	1438 (11.5)	989 (14.3)	449 (8.0)	
≥3	318 (2.5)	146 (2.1)	172 (3.1)	

Characteristic	Full sample (N=12,502)	Self-completed sample (n=6919)	Proxy-completed sample (n=5583)	*P* value
Using assistive devices in daily life (eg, wheelchair, ventilator, etc), n (%)				<.001
No	7200 (57.6)	4074 (58.9)	3126 (58.9)	
Rarely	1887 (15.1)	1134 (16.4)	753 (16.4)	
Sometimes	1490 (11.9)	852 (12.3)	639 (11.4)	
Often	950 (7.6)	458 (6.6)	492 (8.8)	
Always	958 (7.7)	396 (5.7)	562 (10.1)	
Disability (either physical or psychological), n (%)				<.001
Yes	3551 (28.4)	1880 (27.2)	1671 (29.9)	
No	8951 (71.6)	5039 (72.8)	3912 (71.1)	
Number of family members living together, n (%)				<.001
0	480 (3.8)	271 (3.9)	209 (3.7)	
1	1855 (14.8)	1140 (16.5)	715 (12.8)	
2	3457 (27.6)	1975 (28.5)	1482 (26.5)	
3	2725 (21.7)	1304 (18.8)	1421 (25.5)	
≥4	3155 (25.2)	1585 (22.9)	1570 (28.1)	
Missing	830 (6.6)	644 (9.3)	186 (3.3)	
Perceived disease severity (score 1-10), mean (SD)	8.2 (2.3)	7.8 (2.3)	8.7 (2.1)	<.001

[a] A currency exchange rate of CNY 1=US $0.14 is applicable.

[b] RD: rare disease.

EQ-5D-5L Profile for Patients With RDs

Data from 10,102 participants over the age of 12 years were elicited for analysis. Their background characteristics are presented in Multimedia Appendix 2. Table 2 displays the proportion of reported health states for each level of the EQ-5D-5L dimensions. Among patients with RDs, 65.3% (6599/10,102), 47.5% (4799/10,102), 47.0% (4746/10,102), 24.8% (2506/10,102), and 18.4% (1855/10,102) reported no problems for "self-care," "usual activities," "mobility," "pain/discomfort," and "anxiety/depression," respectively. In the case of pediatric patients with RDs, 49.3% (704/1428), 40.2% (574/1428), and 28.6% (408/1428) reported no problems for "usual activities," "pain/discomfort," and "anxiety/depression," respectively, and these proportions were around 10% higher than the proportions in the overall sample. The proportions of having no problems in all 5 dimensions were lower in older patients with RDs than in the overall sample.

Among self-completed respondents, 69.9% (4826/6902) and 50.4% (3478/6902) reported no problems for "self-care" and "usual activities," respectively, whereas only 17.7% (1223/6902) reported problems for "anxiety/depression." Moreover, 6.0% (413/6902) and 0.3% (24/6902) of patients reported full and worst health states, respectively.

Table 2. Health status reported by using the EQ-5D-5L descriptive system.

Variable	All patients (N=10,102)	Pediatric (12-18 years) (n=1428)	Older (≥60 years) (n=503)	Self-completed (n=6902)	Proxy-completed (n=3200)
Mobility, n (%)					
No problem	4746 (47.0)	802 (56.2)	134 (26.6)	3311 (48.0)	1435 (44.8)
Slight problems	2495 (24.7)	219 (15.3)	135 (26.8)	1844 (26.7)	651 (20.3)
Moderate problems	1311 (13.0)	125 (8.8)	84 (16.7)	917 (13.3)	394 (12.3)
Severe problems	744 (7.4)	63 (4.4)	70 (13.9)	477 (6.9)	267 (8.3)
Extreme problems	806 (8.0)	219 (15.3)	80 (15.9)	353 (5.1)	453 (14.2)
Self-care, n (%)					
No problem	6599 (65.3)	872 (61.1)	212 (42.1)	4826 (69.9)	1773 (55.4)
Slight problems	1751 (17.3)	190 (13.3)	113 (22.5)	1221 (17.7)	530 (16.6)
Moderate problems	746 (7.4)	98 (6.9)	67 (13.3)	466 (6.8)	280 (8.8)
Severe problems	345 (3.4)	45 (3.2)	39 (7.8)	190 (2.8)	155 (4.8)
Extreme problems	661 (6.5)	223 (15.6)	72 (14.3)	199 (2.9)	462 (14.4)
Usual activities, n (%)					
No problem	4799 (47.5)	704 (49.3)	127 (25.2)	3478 (50.4)	1321 (41.3)
Slight problems	2862 (28.3)	322 (22.5)	162 (32.2)	2028 (29.4)	834 (26.1)
Moderate problems	1222 (12.1)	140 (9.8)	89 (17.7)	826 (12.0)	396 (12.4)
Severe problems	638 (6.3)	83 (5.8)	66 (13.1)	374 (5.4)	264 (8.3)
Extreme problems	581 (5.8)	179 (12.5)	59 (11.7)	196 (2.8)	385 (12.0)
Pain/discomfort, n (%)					
No problem	2506 (24.8)	574 (40.2)	77 (15.3)	1606 (23.3)	900 (28.1)
Slight problems	4540 (44.9)	533 (37.3)	198 (39.4)	3292 (47.7)	1248 (39.0)
Moderate problems	2055 (20.3)	196 (13.7)	132 (26.2)	1424 (20.6)	631 (19.7)
Severe problems	647 (6.4)	72 (5.0)	62 (12.3)	399 (5.8)	248 (7.8)
Extreme problems	354 (3.5)	53 (3.7)	34 (6.8)	181 (2.6)	173 (5.4)
Anxiety/depression, n (%)					
No problem	1855 (18.4)	408 (28.6)	72 (14.3)	1223 (17.7)	632 (19.8)
Slight problems	4407 (43.6)	656 (45.9)	178 (35.4)	3066 (44.4)	1341 (41.9)
Moderate problems	2379 (23.5)	220 (15.4)	140 (27.8)	1695 (24.6)	684 (21.4)
Severe problems	888 (8.8)	80 (5.6)	64 (12.7)	571 (8.3)	317 (9.9)
Extreme problems	573 (5.7)	64 (4.5)	49 (9.7)	347 (5.0)	226 (7.1)
Full health[a], n (%)	708 (7.0)	214 (15.0)	21 (4.2)	413 (6.0)	295 (9.2)
Worst health[b], n (%)	74 (0.7)	13 (0.9)	17 (3.4)	24 (0.3)	50 (1.6)

[a] Full health indicates that the respondent selected “no problem” for all 5 dimensions of the EQ-5D descriptive system (1, 1, 1, 1, and 1).

[b] Worst health indicates that the respondent selected “severe/extreme problems” for all 5 dimensions of the EQ-5D descriptive system (5, 5, 5, 5, and 5).

Proxy-completed respondents showed a higher proportion of reporting extreme problems than self-completed respondents in all 5 dimensions. Among proxy-completed respondents, 55.4% (1773/3200) and 44.8% (1435/3200) reported no problems for “self-care” and “mobility,” respectively. Regarding the types of caregivers, fathers reported the highest proportion of care receivers experiencing problems related to extreme mobility and self-care (91/535, 17.0% and 84/535, 15.7%, respectively), whereas children caregivers reported a higher proportion of care receivers experiencing extreme problems with usual activities and pain/discomfort (69/595, 11.6% and 47/595, 7.9%) than other types of caregivers (Table 3). Meanwhile, grandparents/relatives and children caregivers reported similarly high proportions of extreme problems with anxiety/depression (approximately 11%) compared to other types of caregivers.

Table 3. Proxy-completed health status using the EQ-5D-5L descriptive system.

Variable	Father (n=535)	Mother (n=1084)	Children (n=595)	Spouse (n=369)	Grandparents/relatives (n=158)	Others (n=459)
Mobility, n (%)						
No problem	258 (48.2)	570 (52.6)	208 (35.0)	143 (38.8)	56 (35.4)	200 (43.6)
Slight problems	99 (18.5)	188 (17.3)	137 (23.0)	93 (25.2)	35 (22.2)	99 (21.6)
Moderate problems	56 (10.5)	91 (8.4)	100 (16.8)	50 (13.6)	30 (19.0)	67 (14.6)
Severe problems	31 (5.8)	64 (5.9)	76 (12.8)	39 (10.6)	16 (10.1)	41 (8.9)
Extreme problems	91 (17.0)	171 (15.8)	74 (12.4)	44 (11.9)	21 (13.3)	52 (11.3)
Self-care, n (%)						
No problem	293 (54.8)	634 (58.5)	283 (47.6)	205 (55.6)	77 (48.7)	281 (61.2)
Slight problems	88 (16.4)	152 (14.0)	110 (18.5)	78 (21.1)	33 (20.9)	69 (15.0)
Moderate problems	43 (8.0)	80 (7.4)	71 (11.9)	23 (6.2)	20 (12.7)	43 (9.4)
Severe problems	27 (5.0)	43 (4.0)	41 (6.9)	19 (5.1)	10 (6.3)	15 (3.3)
Extreme problems	84 (15.7)	175 (16.1)	90 (15.1)	44 (11.9)	18 (11.4)	51 (11.1)
Usual activities, n (%)						
No problem	225 (42.1)	500 (46.1)	199 (33.4)	143 (38.8)	48 (30.4)	206 (44.9)
Slight problems	132 (24.7)	257 (23.7)	172 (28.9)	107 (29.0)	48 (30.4)	118 (25.7)
Moderate problems	64 (12.0)	105 (9.7)	89 (15.0)	50 (13.6)	29 (18.4)	59 (12.9)
Severe problems	38 (7.1)	75 (6.9)	66 (11.1)	28 (7.6)	17 (10.8)	40 (8.7)
Extreme problems	76 (14.2)	147 (13.6)	69 (11.6)	41 (11.1)	16 (10.1)	36 (7.8)
Pain/discomfort, n (%)						
No problem	175 (32.7)	391 (36.1)	121 (20.3)	82 (22.2)	25 (15.8)	106 (23.1)
Slight problems	212 (39.6)	424 (39.1)	211 (35.5)	166 (45.0)	65 (41.1)	170 (37.0)
Moderate problems	85 (15.9)	163 (15.0)	153 (25.7)	70 (19.0)	45 (28.5)	115 (25.1)
Severe problems	30 (5.6)	66 (6.1)	63 (10.6)	32 (8.7)	12 (7.6)	45 (9.8)
Extreme problems	33 (6.2)	40 (3.7)	47 (7.9)	19 (5.1)	11 (7.0)	23 (5.0)
Anxiety/depression, n (%)						
No problem	127 (23.7)	272 (25.1)	79 (13.3)	59 (16.0)	16 (10.1)	79 (17.2)
Slight problems	234 (43.7)	485 (44.7)	219 (36.8)	167 (45.3)	63 (39.9)	173 (37.7)
Moderate problems	98 (18.3)	196 (18.1)	144 (24.2)	82 (22.2)	47 (29.7)	117 (25.5)
Severe problems	35 (6.5)	86 (7.9)	88 (14.8)	38 (10.3)	15 (9.5)	55 (12.0)
Extreme problems	41 (7.7)	45 (4.2)	65 (10.9)	23 (6.2)	17 (10.8)	35 (7.6)
Full health[a], n (%)	55 (10.3)	136 (12.5)	36 (6.1)	20 (5.4)	5 (3.2)	43 (9.4)
Worst health[b], n (%)	9 (1.7)	7 (0.6)	17 (2.9)	6 (1.6)	4 (2.5)	7 (1.5)

[a] Full health indicates that the respondent selected “no problem” for all 5 dimensions of the EQ-5D descriptive system (1, 1, 1, 1, and 1).

[b] Worst health indicates that the respondent selected “severe/extreme problems” for all 5 dimensions of the EQ-5D descriptive system (5, 5, 5, 5, and 5).

EQ-5D-5L Utility Score for Patients With RDs

Tables 4-6 show the means and SDs of the EQ-5D-5L utility scores stratified by different background characteristics and types of RDs. The mean utility scores were 0.659 (SD 0.324), 0.691 (SD 0.284), and 0.590 (SD 0.388) for the overall, self-completed, and proxy-completed samples, respectively. In the overall sample, the EQ-5D-5L utility score was significantly associated with belonging to the female sex, being aged ≤10 years, being actively employed, residing in urban areas, having a high family annual income, having a short duration of disease, and not having children. Similar trends were observed in the self- and proxy-completed samples, except for insignificant differences in the utility score between patients with different numbers of children in the self-completed sample and between urban and rural residents in the proxy-completed sample.

Regarding the mean utility score for specific RDs, patients with amyotrophic lateral sclerosis reported a significantly lower utility score than those with other RDs; however, in the self-completed sample, the mean utility score for spinal muscular atrophy was the lowest (Multimedia Appendix 3).

Regarding the EQ-5D-5L utility scores reported by different types of caregivers, we found that mothers reported a higher score (0.633) whereas children caregivers reported a lower score (0.513) compared to other caregivers (0.552). Detailed comparisons of the utility scores of caregivers are provided in Table 7.

Table 4. EQ-5D utility scores and the associations with patients' background characteristics (full sample).

Variable	Utility score, mean (SD)	Coefficient (95% CI)	*P* value
Overall	0.659 (0.324)	N/A[a]	N/A
Gender			
Male	0.633 (0.339)	N/A	N/A
Female	0.685 (0.308)	0.051 (0.039 to 0.064)	<.001
Age (years)			
12-20	0.660 (0.362)	N/A	N/A
21-30	0.709 (0.289)	0.049 (0.030 to 0.068)	<.001
31-40	0.698 (0.278)	0.038 (0.019 to 0.057)	<.001
41-50	0.625 (0.319)	−0.035 (−0.057 to 0.014)	.001
51-60	0.557 (0.364)	−0.104 (−0.129 to 0.079)	<.001
≥61	0.476 (0.412)	−0.185 (−0.217 to 0.152)	<.001
Employment			
Active	0.787 (0.206)	N/A	N/A
Nonactive	0.589 (0.355)	−0.198 (−0.211 to −0.186)	<.001
Family registry			
Urban	0.669 (0.322)	N/A	N/A
Rural	0.647 (0.327)	−0.021 (−0.034 to −0.008)	.001
Family annual income (CNY[b])			
≤5000	0.505 (0.381)	N/A	N/A
5001-10,000	0.543 (0.367)	0.038 (0.003 to 0.074)	.03
10,001-30,000	0.602 (0.334)	0.098 (0.069 to 0.127)	<.001
30,001-50,000	0.658 (0.316)	0.153 (0.125 to 0.182)	<.001
50,001-100,000	0.702 (0.299)	0.197 (0.169 to 0.225)	<.001
100,001-200,000	0.746 (0.277)	0.241 (0.211 to 0.272)	<.001
200,001-300,000	0.751 (0.287)	0.246 (0.204 to 0.288)	<.001
300,001-500,000	0.755 (0.283)	0.250 (0.195 to 0.305)	<.001
≥500,001	0.753 (0.315)	0.249 (0.180 to 0.317)	<.001
Duration of RDs[c] (years)			
≤10	0.710 (0.291)	N/A	N/A
11-20	0.644 (0.340)	−0.066 (−0.081 to −0.052)	<.001
21-30	0.617 (0.340)	−0.093 (−0.112 to −0.074)	<.001
31-40	0.591 (0.317)	−0.119 (−0.145 to −0.093)	<.001
≥41	0.499 (0.348)	−0.211 (−0.249 to −0.174)	<.001
Number of children			
0	0.669 (0.329)	N/A	N/A
1	0.664 (0.311)	−0.005 (−0.019 to 0.009)	.48
2	0.641 (0.323)	−0.028 (−0.047 to −0.009)	.005
≥3	0.530 (0.369)	−0.139 (−0.175 to −0.102)	<.001
Using assistive devices in daily life (eg, wheelchair, ventilator, etc)			
No	0.822 (0.177)	N/A	N/A
Rarely	0.616 (0.242)	−0.206 (−0.219 to −0.193)	<.001

Variable	Utility score, mean (SD)	Coefficient (95% CI)	*P* value
Sometimes	0.518 (0.285)	−0.304 (−0.318 to −0.290)	<.001
Often	0.316 (0.338)	−0.506 (−0.524 to −0.488)	<.001
Always	0.089 (0.336)	−0.733 (−0.751 to −0.716)	<.001
Disability (either physical or psychological)			
Yes	0.450 (0.377)	N/A	N/A
No	0.742 (0.258)	0.293 (0.280 to 0.306)	<.001
Number of family members living together			
0	0.584 (0.372)	N/A	N/A
1	0.629 (0.343)	0.045 (0.011 to 0.078)	.009
2	0.674 (0.318)	0.089 (0.057 to 0.122)	<.001
3	0.662 (0.322)	0.078 (0.045 to 0.111)	<.001
≥4	0.662 (0.316)	0.077 (0.045 to 0.110)	<.001

[a] N/A: not applicable.

[b] A currency exchange rate of CNY 1=US $0.14 is applicable.

[c] RD: rare disease.

Table 5. EQ-5D utility scores of the self-completed sample.

Variable	Utility score, mean (SD)	Coefficient (95% CI)	*P* value
Overall	0.691 (0.284)	N/A[a]	N/A
Gender			
Male	0.664 (0.298)	N/A	N/A
Female	0.715 (0.270)	0.051 (0.037 to 0.064)	<.001
Age (years)			
12-20	0.709 (0.294)	N/A	N/A
21-30	0.719 (0.274)	0.009 (−0.021 to 0.040)	.53
31-40	0.708 (0.266)	−0.002 (−0.032 to 0.028)	.91
41-50	0.649 (0.298)	−0.061 (−0.092 to −0.029)	<.001
51-60	0.631 (0.308)	−0.078 (−0.115 to −0.042)	<.001
≥61	0.584 (0.358)	−0.126 (−0.178 to −0.073)	<.001
Employment			
Active	0.793 (0.195)	N/A	N/A
Nonactive	0.612 (0.316)	−0.181 (−0.194 to −0.168)	<.001
Family registry			
Urban	0.700 (0.281)	N/A	N/A
Rural	0.680 (0.288)	−0.020 (−0.034 to −0.007)	.003
Family income per year (CNY[b])			
≤5000	0.530 (0.353)	N/A	N/A
5001-10,000	0.577 (0.330)	0.047 (0.010 to 0.084)	.01
10,001-30,000	0.627 (0.300)	0.097 (0.068 to 0.127)	<.001
30,001-50,000	0.680 (0.275)	0.150 (0.120 to 0.179)	<.001
50,001-100,000	0.738 (0.251)	0.208 (0.179 to 0.237)	<.001
100,001-200,000	0.786 (0.220)	0.256 (0.225 to 0.287)	<.001
200,001-300,000	0.790 (0.230)	0.260 (0.218 to 0.302)	<.001
300,001-500,000	0.806 (0.223)	0.276 (0.221 to 0.331)	<.001
≥500,001	0.793 (0.276)	0.263 (0.192 to 0.333)	<.001
Duration of RDs[c] (years)			
≤10	0.750 (0.238)	N/A	N/A
11-20	0.693 (0.286)	−0.057 (−0.072 to −0.410)	<.001
21-30	0.640 (0.311)	−0.109 (−0.129 to −0.090)	<.001
31-40	0.602 (0.310)	−0.148 (−0.172 to −0.123)	<.001
≥41	0.498 (0.347)	−0.251 (−0.287 to −0.216)	<.001
Number of children			
0	0.686 (0.298)	N/A	N/A
1	0.697 (0.276)	0.011 (−0.004 to 0.026)	.14
2	0.694 (0.269)	0.008 (−0.013 to 0.028)	.45
≥3	0.670 (0.265)	−0.016 (−0.063 to 0.032)	.52
Using assistive devices in daily life (eg, wheelchair, ventilator, etc)			
No	0.824 (0.159)	N/A	N/A
Rarely	0.634 (0.222)	−0.190 (−0.204 to −0.176)	<.001

Variable	Utility score, mean (SD)	Coefficient (95% CI)	*P* value
Sometimes	0.533 (0.269)	−0.291 (−0.307 to −0.275)	<.001
Often	0.381 (0.327)	−0.443 (−0.464 to −0.423)	<.001
Always	0.187 (0.348)	−0.638 (−0.660 to −0.615)	<.001
Disability (either physical or psychological)			
Yes	0.510 (0.347)	N/A	N/A
No	0.759 (0.222)	0.249 (0.235 to 0.263)	<.001
Number of family members living together			
0	0.683 (0.277)	N/A	N/A
1	0.672 (0.307)	−0.010 (−0.048 to 0.027)	.58
2	0.689 (0.284)	0.006 (−0.030 to 0.043)	.73
3	0.692 (0.281)	0.009 (−0.028 to 0.046)	.63
≥4	0.699 (0.275)	0.016 (−0.021 to 0.053)	.40

[a] N/A: not applicable.

[b] A currency exchange rate of CNY 1=US $0.14 is applicable.

[c] RD: rare disease.

Table 6. EQ-5D utility scores of the proxy-completed sample.

Variable	Utility score, mean (SD)	Coefficient (95% CI)	*P* value
Overall	0.590 (0.388)	N/A[a]	N/A
Gender			
Male	0.579 (0.394)	N/A	N/A
Female	0.603 (0.381)	0.024 (−0.003 to 0.051)	.09
Age (years)			
12-20	0.646 (0.378)	N/A	N/A
21-30	0.661 (0.345)	0.014 (−0.026 to 0.055)	.49
31-40	0.637 (0.337)	−0.010 (−0.053 to 0.034)	.67
41-50	0.521 (0.381)	−0.125 (−0.173 to −0.078)	<.001
51-60	0.456 (0.408)	−0.191 (−0.233 to −0.149)	<.001
≥61	0.422 (0.427)	−0.225 (−0.271 to −0.178)	<.001
Employment			
Active	0.755 (0.255)	N/A	N/A
Nonactive	0.556 (0.402)	−0.199 (−0.234 to −0.164)	<.001
Family registry			
Urban	0.599 (0.390)	N/A	N/A
Rural	0.578 (0.388)	−0.02 (−0.047 to 0.007)	.14
Family income per year (CNY[b])			
≤5000	0.442 (0.439)	N/A	N/A
5001-10,000	0.475 (0.425)	0.034 (0.384 to 0.499)	.39
10,001-30,000	0.551 (0.391)	0.109 (−0.043 to 0.111)	.001
30,001-50,000	0.616 (0.378)	0.175 (0.045 to 0.174)	<.001
50,001-100,000	0.625 (0.370)	0.183 (0.111 to 0.238)	<.001
100,001-200,000	0.642 (0.368)	0.200 (0.121 to 0.246)	<.001
200,001-300,000	0.638 (0.387)	0.196 (0.130 to 0.270)	<.001
300,001-500,000	0.591 (0.381)	0.150 (0.099 to 0.294)	.03
≥500,001	0.655 (0.384)	0.214 (0.016 to 0.283)	.007
Duration of RDs[c] (years)			
≤10	0.617 (0.372)	N/A	N/A
11-20	0.569 (0.398)	−0.066 (−0.095 to −0.036)	<.001
21-30	0.552 (0.405)	−0.083 (−0.126 to −0.039)	<.001
31-40	0.530 (0.348)	−0.104 (−0.182 to −0.027)	.009
≥41	0.504 (0.359)	−0.130 (−0.249 to −0.011)	.03
Number of children			
0	0.643 (0.372)	N/A	N/A
1	0.531 (1.400)	−0.111 (−0.144 to −0.078)	<.001
2	0.526 (2.393)	−0.116 (−0.156 to −0.077)	<.001
≥3	0.411 (0.402)	−0.232 (−0.292 to −0.172)	<.001
Using assistive devices in daily life (eg, wheelchair, ventilator, etc)			
No	0.817 (0.215)	N/A	N/A
Rarely	0.573 (0.279)	−0.244 (−0.270 to −0.217)	<.001

Variable	Utility score, mean (SD)	Coefficient (95% CI)	*P* value
Sometimes	0.485 (0.315)	−0.332 (−0.360 to −0.303)	<.001
Often	0.207 (0.329)	−0.610 (−0.643 to −0.576)	<.001
Always	−0.015 (0.289)	−0.832 (−0.861 to −0.803)	<.001
Disability (either physical or psychological)			
Yes	0.336 (0.403)	N/A	N/A
No	0.705 (0.322)	0.369 (0.343 to 0.395)	<.001
Number of family members living together			
0	0.439 (0.442)	N/A	N/A
1	0.540 (0.393)	0.102 (0.037 to 0.166)	.002
2	0.639 (0.382)	0.201 (0.139 to 0.262)	<.001
3	0.606 (0.382)	0.167 (0.104 to 0.230)	<.001
≥4	0.580 (0.378)	0.141 (0.079 to 0.204)	<.001

[a] N/A: not applicable.

[b] A currency exchange rate of CNY 1=US $0.14 is applicable.

[c] RD: rare disease.

Table 7. EQ-5D-5L utility scores stratified by the types of caregivers.

Variable	Utility score, mean (SD)					
	Father	Mother	Children	Spouse	Grandparents/relatives	Others
Overall	0.603 (0.396)	0.633 (0.377)	0.513 (0.409)	0.588 (0.382)	0.530 (0.375)	0.595 (0.371)
Gender						
Male	0.559 (0.411)	0.618 (0.383)	0.552 (0.401)	0.573 (0.409)	0.541 (0.362)	0.564 (0.383)
Female	0.732 (0.316)	0.653 (0.369)	0.483 (0.414)	0.602 (0.356)	0.515 (0.392)	0.645 (0.345)
Age (years)						
12-20	0.628 (0.390)	0.653 (0.376)	N/A[a]	N/A	0.631 (0.354)	0.548 (0.416)
21-30	0.588 (0.415)	0.666 (0.347)	N/A	0.791 (0.186)	0.601 (0.302)	0.669 (0.337)
31-40	0.695 (0.320)	0.536 (0.441)	N/A	0.677 (0.331)	0.584 (0.361)	0.630 (0.320)
41-50	0.562 (0.306)	0.480 (0.314)	0.583 (0.339)	0.498 (0.416)	0.435 (0.383)	0.529 (0.413)
51-60	N/A	0.468 (0.408)	0.460 (0.406)	0.458 (0.408)	0.234 (0.395)	0.443 (0.402)
≥61	N/A	N/A	0.405 (0.437)	0.393 (0.467)	N/A	0.443 (0.466)
Employment						
Active	0.755 (0.285)	0.762 (0.245)	0.685 (0.290)	0.775 (0.230)	0.722 (0.237)	0.782 (0.252)
Nonactive	0.589 (0.402)	0.623 (0.384)	0.483 (0.420)	0.470 (0.412)	0.463 (0.391)	0.502 (0.385)
Family registry						
Urban	0.634 (0.388)	0.643 (0.371)	0.515 (0.424)	0.574 (0.396)	0.489 (0.375)	0.603 (0.364)
Rural	0.566 (0.405)	0.617 (0.387)	0.509 (0.394)	0.608 (0.362)	0.559 (0.373)	0.588 (0.377)
Family income per year (CNY[b])						
≤5000	0.391 (0.451)	0.547 (0.412)	0.329 (0.469)	0.396 (0.488)	0.449 (0.395)	0.446 (0.427)
5001-10,000	0.492 (0.450)	0.502 (0.437)	0.493 (0.420)	0.395 (0.419)	0.357 (0.414)	0.495 (0.415)
10,001-30,000	0.573 (0.431)	0.579 (0.384)	0.491 (0.400)	0.484 (0.410)	0.510 (0.362)	0.571 (0.347)
30,001-50,000	0.606 (0.386)	0.661 (0.360)	0.524 (0.427)	0.670 (0.325)	0.636 (0.331)	0.605 (0.369)
50,001-100,000	0.627 (0.370)	0.670 (0.364)	0.571 (0.370)	0.604 (0.373)	0.504 (0.414)	0.643 (0.353)
100,001-200,000	0.709 (0.350)	0.693 (0.350)	0.481 (0.416)	0.617 (0.356)	0.588 (0.328)	0.728 (0.301)
200,001-300,000	0.787 (0.244)	0.594 (0.412)	0.553 (0.407)	0.763 (0.377)	N/A	0.679 (0.365)
300,001-500,000	0.494 (0.346)	0.609 (0.434)	0.450 (0.430)	0.803 (0.143)	N/A	0.710 (0.410)
≥500,001	0.775 (0.169)	0.660 (0.326)	0.330 (0.488)	0.877 (0.111)	N/A	0.928 (0.144)
Duration of RDs[c] (years)						
≤10	0.620 (0.386)	0.663 (0.362)	0.536 (0.397)	0.629 (0.367)	0.599 (0.299)	0.675 (0.336)
11-20	0.595 (0.391)	0.602 (0.390)	0.508 (0.414)	0.522 (0.419)	0.444 (0.390)	0.553 (0.396)
21-30	0.544 (0.464)	0.633 (0.347)	0.383 (0.442)	0.553 (0.349)	0.542 (0.467)	0.570 (0.391)
31-40	0.318 (0.364)	0.501 (0.374)	0.453 (0.447)	0.592 (0.276)	0.717 (0.093)	0.571 (0.325)
≥41	N/A	0.625 (0.330)	0.644 (0.275)	0.385 (0.544)	0.436 (0.422)	0.395 (0.357)
Number of children						
0	0.624 (0.393)	0.651 (0.373)	0.699 (0.341)	0.757 (0.218)	0.601 (0.337)	0.608 (0.367)
1	0.511 (0.426)	0.544 (0.384)	0.469 (0.412)	0.557 (0.396)	0.457 (0.410)	0.596 (0.375)
2	0.572 (0.365)	0.494 (0.372)	0.476 (0.417)	0.608 (0.383)	0.493 (0.397)	0.588 (0.350)
≥3	0.403 (0.396)	0.417 (0.418)	0.386 (0.400)	0.622 (0.355)	0.310 (0.341)	0.364 (0.463)

Variable	Utility score, mean (SD)					
	Father	Mother	Children	Spouse	Grandparents/relatives	Others
Using assistive devices in daily life (eg, wheelchair, ventilator, etc)						
No	0.838 (0.218)	0.835 (0.210)	0.786 (0.222)	0.804 (0.193)	0.756 (0.259)	0.806 (0.215)
Rarely	0.636 (0.240)	0.588 (0.297)	0.501 (0.298)	0.605 (0.236)	0.527 (0.238)	0.586 (0.280)
Sometimes	0.440 (0.354)	0.498 (0.322)	0.431 (0.342)	0.485 (0.293)	0.484 (0.270)	0.552 (0.264)
Often	0.202 (0.321)	0.244 (0.337)	0.173 (0.317)	0.159 (0.318)	0.070 (0.274)	0.269 (0.361)
Always	0.007 (0.288)	0.059 (0.283)	−0.115 (0.245)	−0.133 (0.246)	−0.030 (0.255)	0.009 (0.343)
Disability (either physical or psychological)						
Yes	0.336 (0.400)	0.381 (0.393)	0.237 (0.405)	0.246 (0.419)	0.343 (0.409)	0.350 (0.403)
No	0.732 (0.324)	0.777 (0.280)	0.597 (0.372)	0.679 (0.315)	0.607 (0.332)	0.732 (0.268)
Number of family members living together						
0	0.563 (0.402)	0.533 (0.400)	0.333 (0.442)	0.491 (0.490)	0.423 (0.340)	0.401 (0.470)
1	0.507 (0.462)	0.603 (0.381)	0.471 (0.400)	0.585 (0.352)	0.405 (0.429)	0.594 (0.360)
2	0.663 (0.403)	0.672 (0.370)	0.607 (0.410)	0.549 (0.389)	0.531 (0.331)	0.652 (0.357)
3	0.641 (0.360)	0.616 (0.382)	0.601 (0.383)	0.557 (0.419)	0.566 (0.402)	0.577 (0.378)
≥4	0.530 (0.391)	0.607 (0.383)	0.505 (0.384)	0.676 (0.332)	0.679 (0.284)	0.561 (0.373)

[a] N/A: not applicable.

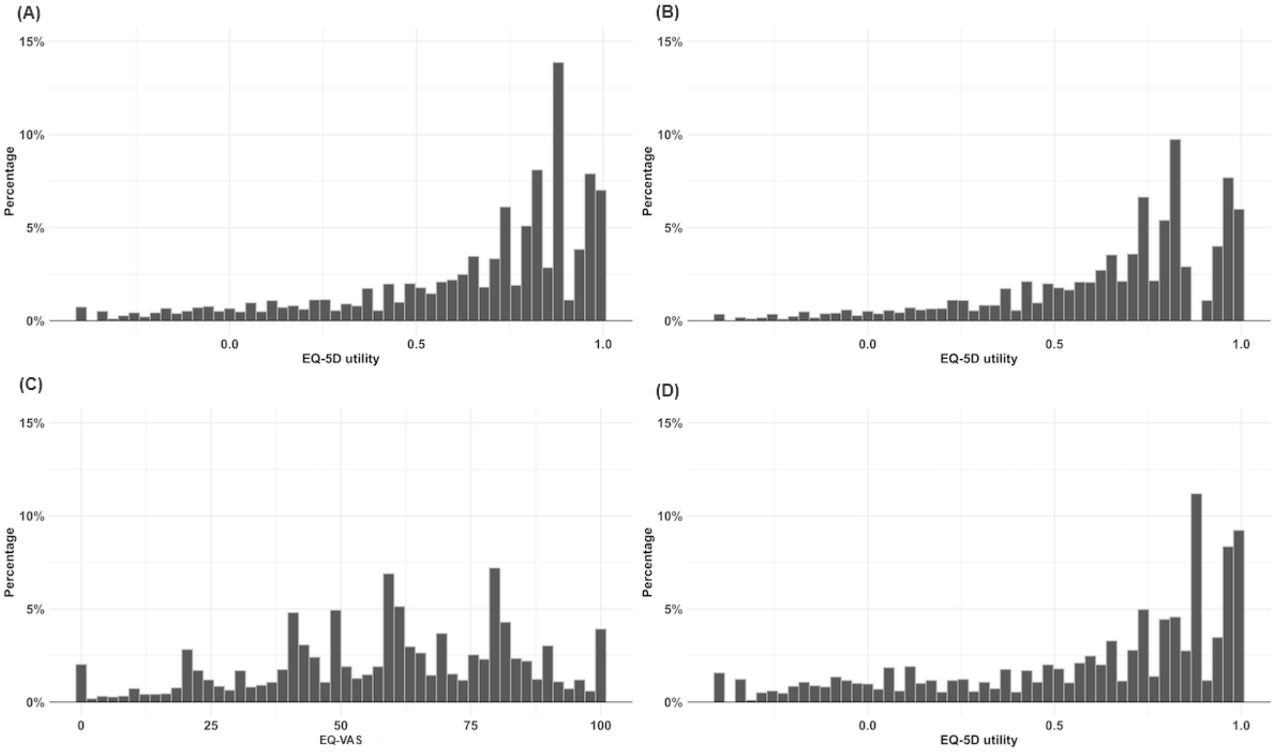
[b] A currency exchange rate of CNY 1=US $0.14 is applicable.

[c] RD: rare disease.

EQ-VAS Profile

The mean and median EQ-VAS scores of the full sample were 59.22 and 61, respectively, with scores ranging from 0 to 100. The mean and median EQ-VAS scores of the self- and proxy-completed samples were 60.19 and 61, and 57.13 and 60, respectively. There was a strong and significant association between the EQ-5D-5L utility score and the EQ-VAS score (r=0.52; *P*<.001). Additionally, Figure 1 presents the distribution of the EQ-VAS and EQ-5D-5L utility scores.

Figure 1. Distribution of the EQ-5D-5L utility scores and EQ-VAS scores. (A) EQ-5D utility score for the full sample; (B) EQ-5D utility score for the self-completed sample; (C) EQ-VAS score for the full sample; (D) EQ-5D utility score for the proxy-completed sample.



Discussion

Principal Findings

This study presents a population norm of HRQoL for patients with RDs in China, using the preference-based value set of the EQ-5D-5L. This is also the largest study estimating the utility scores of the EQ-5D-5L in patients with RDs worldwide, even including studies using the EQ-5D-3L or any other preference-based measures. Although we only included 33 types of RDs, it is still the largest number of RDs included in any EQ-5D-5L–related study worldwide. Additionally, we estimated the EQ-5D-5L profile and utility scores for both pediatric and older patients with RDs, as well as self-completed and proxy-completed patients with RDs, and evaluated the differences between caregiver types. Our study provides valuable information as a reference for health evaluation and comparisons of different health care interventions in RDs, both locally and globally.

Comparisons With Previous Studies

The mean utility score for patients with RDs (0.659) in our sample, as measured by the EQ-5D-5L, was significantly lower than that for the general Chinese population in China (0.957) [15] and Hong Kong (0.92) [16]. This score is comparable to the findings of López-Bastida et al [17], who utilized a sample of 1544 patients with RDs from several European countries. Our mean utility score is similar to theirs, suggesting that our results are consistent with those of other large-scale studies conducted in the RD population. However, López-Bastida et al did not report the respondents’ selections on the dimensions of the EQ-5D. This makes it difficult to compare dimension-level data between the 2 studies. Moreover, given that the value set for China and European countries may be different, the index score between this study and the Europeans study may not be comparable. Compared with another study using the EQ-5D-3L in Chinese patients with RDs in Hong Kong [11], there were higher proportions of “no problems” for the dimensions of “mobility,” “self-care,” and “usual activities,” but lower proportions of “no problems” for “pain/discomfort” and “anxiety/depression.” Additionally, regarding the utility score, the Hong Kong study reported a lower mean score (0.53) than our study. However, the Hong Kong study had a very small sample size, which may have limited the reliability of their findings. Compared to studies assessing HRQoL in Chinese patients with chronic diseases in general [18-22], the utility scores in patients with RDs were significantly lower. It reflects the large negative impact of RDs on HRQoL and the greater health care needs of patients with RDs in China.

Similar to the general population norm [15,16,23-25], demographics and socioeconomics also significantly impacted patients with RDs. For example, older patients with RDs living in rural areas or having lower family annual incomes reported significantly lower utility scores than their counterparts. Regarding disease-related factors, we found a very large difference (0.733) between patients with RDs who did not use assistive devices and those who used them in daily life. Despite the global need and recognized benefits of using assistive products, access to them remains limited [26]. Our findings quantify the importance of assistive devices for improving the HRQoL of patients with RDs and promoting their well-being. Moreover, we have further presented the EQ-5D-5L utility score for all 33 types of RDs. We found that patients with neuromuscular diseases (amyotrophic lateral sclerosis and spinal muscular disease) tended to report lower utility scores compared to the scores in patients with other types of RDs. This is in line with previous findings. For example, Ng et al [11] and Sequeira et al [27] indicated that patients with neurologic diseases reported the lowest utility scores among patients of all RD categories. Despite acknowledging that some studies have measured HRQoL based on disease category (eg, neurologic disease) rather than a single disease, we decided not to categorize the 33 RDs into different categories, because even though some RDs share similar pathologic and physiologic mechanisms, each RD has its own characteristics. Our list can help researchers in flexibly comparing HRQoL in their research based on either a single RD or a category of RDs.

The EQ-5D-5L proxy questionnaire has been shown to be feasible and valid in various populations. However, most studies have focused on people with mental health problems, such as dementia [28,29]. The EQ-5D-5L profile in patients with RDs is limited. Given that a high proportion of RD patients are children or adolescents, owing to the genetic nature of RDs, proxy-completed data are very useful in RD research. Our study provides valuable information about EQ-5D-5L utility scores completed by proxies. We found that the mean utility score of proxy-completed patients was lower than that of self-completed patients, which is consistent with previous findings [30]. This is reasonable as most RD patients who cannot complete the questionnaire themselves have poor physical and mental health. However, for this study, we only used the proxy-patient version of the EQ-5D-5L, and HRQoL measured by the proxy-proxy version may be assessed in the future.

In addition, to provide a more detailed analysis, we have further stratified the EQ-5D utility score based on the type of caregiver. Our findings are similar to those of a previous study reporting that different proxies provide different ratings of patient health [31]. Mothers reported a higher proportion of “no problems” for nearly all dimensions of the EQ-5D-5L than fathers and grandparents or relatives. However, fathers reported a higher proportion of “extreme problems” for 4 out of 5 dimensions than mothers. This may be because when the mother is the primary caregiver, the health condition of the child is usually not very severe, allowing the father to go out to work. However, if the father is the primary caregiver, both parents may have caregiving responsibilities and the child’s health condition is often much more serious. Additionally, most child caregivers reported a lower proportion of “no problems” than other types of caregivers, as the care receivers of these caregivers are usually older patients with RDs. These data are useful because, compared to patients with normal disease, more adult patients with RDs cannot complete the EQ-5D by themselves owing to their poor health status. Measuring HRQoL by caregivers is therefore important to examine the effectiveness of interventions that are designed to improve the practice of RD care.

Strengths and Limitations

This study presents a comprehensive analysis of a normative profile of the EQ-5D-5L for patients with RDs in China. We used a large sample size and included both self- and proxy-completed data to provide a more complete picture of HRQoL for this population than that provided by previous studies worldwide. Our findings can serve as a baseline for future comparisons of HRQoL with RD populations, both locally and globally. Moreover, the results can be used to evaluate the efficacies of various policies, strategies, or interventions aimed at improving the HRQoL of the RD population. Furthermore, the study provides EQ-5D-5L utility values for the future cost-effectiveness analysis of interventions for RDs. By providing a detailed and nuanced understanding of the EQ-5D-5L normative profile for patients with RDs, this study can help inform future research and policy initiatives designed to address the unique health needs of this population.

Despite these strengths, several limitations should be addressed. The primary limitation of this study is that all data were collected voluntarily, which may have led to a lack of information from patients with a poor health status or caregivers with low willingness to participate in the survey, generating selection bias. Additionally, all data were collected via a web-based survey, which could exclude patients not familiar with or not having easy access to the internet, potentially leading to selection bias. Moreover, in this study, we used the EQ-5D-5L instead of the EQ-5D-Y to collect HRQoL data from children and adolescents with RDs. Given that the EQ-5D-Y is a youth-friendly version that has demonstrated high feasibility and reliability in young patients, the estimation of the utility score in this population in our sample might be less reliable. In the future, it is highly encouraged to conduct studies that use EQ-5D-Y to assess HRQoL in young patients with RDs. Finally, although the sample size in our study is larger than that in previous studies, it is important to note that there may be over 20 million people living with RDs in China. Therefore, our sample may not be fully representative, and the generalizability of our findings could be limited.

Conclusion

This study assessed the HRQoL of Chinese patients with RDs using the preference-based EQ-5D-5L. Our findings contribute new knowledge to the existing literature on the relationship among HRQoL, demographics, and health status in patients with RDs. The established normative profile of HRQoL reveals disparities and heterogeneities existing in the health status, as measured by the EQ-5D-5L, across different socioeconomic groups and disease categories. Additionally, our study revealed important differences between self- and proxy-completed HRQoL assessments in the RD population. This finding underscores the need to incorporate perspectives from both patients and their proxies in clinical practice. Further development of the patient cohort is necessary to assess the long-term changes in HRQoL in the RD population.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Grant Number 2016YFC0901500) and the Center for Rare Diseases Research, Chinese Academy of Medical Sciences, Beijing, China (Grant Number 2016ZX310174-4).

Data Availability

Derived data supporting the findings of this study are available from the corresponding author on request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient characteristics stratified by the types of rare diseases.
[DOCX File , 18 KB - publichealth_v9i1e50147_app1.docx]

Multimedia Appendix 2
Background characteristics of patients aged 12 years or older.
[DOCX File , 25 KB - publichealth_v9i1e50147_app2.docx]

Multimedia Appendix 3
EQ-5D-5L utility scores reported by specific rare diseases.
[DOCX File , 26 KB - publichealth_v9i1e50147_app3.docx]

References

1. Bogart KR, Irvin VL. Health-related quality of life among adults with diverse rare disorders. Orphanet J Rare Dis 2017 Dec 07;12(1):177 [FREE Full text] [doi: 10.1186/s13023-017-0730-1] [Medline: 29212508]
2. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet 2020 Feb 16;28(2):165-173 [FREE Full text] [doi: 10.1038/s41431-019-0508-0] [Medline: 31527858]
3. Bruckner-Tuderman L. Epidemiology of rare diseases is important. J Eur Acad Dermatol Venereol 2021 Apr 10;35(4):783-784. [doi: 10.1111/jdv.17165] [Medline: 33689185]
4. Kaufmann P, Pariser AR, Austin C. From scientific discovery to treatments for rare diseases - the view from the National Center for Advancing Translational Sciences - Office of Rare Diseases Research. Orphanet J Rare Dis 2018 Nov 06;13(1):196 [FREE Full text] [doi: 10.1186/s13023-018-0936-x] [Medline: 30400963]
5. Bogart K, Hemmesch A, Barnes E, Blissenbach T, Beisang A, Engel P, Chloe Barnes Advisory Council on Rare Diseases. Healthcare access, satisfaction, and health-related quality of life among children and adults with rare diseases. Orphanet J Rare Dis 2022 May 12;17(1):196 [FREE Full text] [doi: 10.1186/s13023-022-02343-4] [Medline: 35549731]
6. Brooks R. The EuroQol Group after 25 years. Berlin, Germany: Springer; 2013.
7. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. Appl Health Econ Health Policy 2017 Apr 13;15(2):127-137 [FREE Full text] [doi: 10.1007/s40258-017-0310-5] [Medline: 28194657]
8. Buchholz I, Janssen MF, Kohlmann T, Feng Y. A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. Pharmacoeconomics 2018 Jun 23;36(6):645-661 [FREE Full text] [doi: 10.1007/s40273-018-0642-5] [Medline: 29572719]
9. Serrano-Aguilar P, Linertová R, Posada-de-la-Paz M, López-Bastida J, González-Hernández N, Taruscio D. Recruitment procedures for descriptive socio-economic studies in rare diseases. The BURQOL-RD project. Expert Opinion on Orphan Drugs 2015 Jun 17;3(7):759-765. [doi: 10.1517/21678707.2015.1057499]
10. Efthymiadou O, Mossman J, Kanavos P. Differentiation of health-related quality of life outcomes between five disease areas: Results from an international survey of patients. Int J Technol Assess Health Care 2018 Sep 25;34(5):498-506. [doi: 10.1017/S0266462318000557]
11. Ng YN, Ng NY, Fung JL, Lui AC, Cheung NY, Wong WH, et al. Evaluating the Health-Related Quality of Life of the Rare Disease Population in Hong Kong Using EQ-5D 3-Level. Value Health 2022 Sep;25(9):1624-1633 [FREE Full text] [doi: 10.1016/j.jval.2022.04.1725] [Medline: 35568675]
12. Dong D, Chung RY, Chan RHW, Gong S, Xu RH. Why is misdiagnosis more likely among some people with rare diseases than others? Insights from a population-based cross-sectional study in China. Orphanet J Rare Dis 2020 Oct 28;15(1):307 [FREE Full text] [doi: 10.1186/s13023-020-01587-2] [Medline: 33115515]

13. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
14. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L Value Set for China. Value Health 2017 Apr;20(4):662-669 [FREE Full text] [doi: 10.1016/j.jval.2016.11.016] [Medline: 28408009]
15. Yang Z, Busschbach J, Liu G, Luo N. EQ-5D-5L norms for the urban Chinese population in China. Health Qual Life Outcomes 2018 Nov 08;16(1):210 [FREE Full text] [doi: 10.1186/s12955-018-1036-2] [Medline: 30409137]
16. Wong EL, Cheung AW, Wong AY, Xu RH, Ramos-Goñi J, Rivero-Arias O. Normative Profile of Health-Related Quality of Life for Hong Kong General Population Using Preference-Based Instrument EQ-5D-5L. Value Health 2019 Aug;22(8):916-924 [FREE Full text] [doi: 10.1016/j.jval.2019.02.014] [Medline: 31426933]
17. López-Bastida J, Oliva-Moreno J, Linertová R, Serrano-Aguilar P. Social/economic costs and health-related quality of life in patients with rare diseases in Europe. Eur J Health Econ 2016 Apr 29;17 Suppl 1(S1):1-5. [doi: 10.1007/s10198-016-0780-7] [Medline: 27023708]
18. Liang Z, Zhang T, Lin T, Liu L, Wang B, Fu AZ, et al. Health-related quality of life among rural men and women with hypertension: assessment by the EQ-5D-5L in Jiangsu, China. Qual Life Res 2019 Aug 4;28(8):2069-2080. [doi: 10.1007/s11136-019-02139-3] [Medline: 30830645]
19. Wong ELY, Xu RH, Cheung AWL. Measurement of health-related quality of life in patients with diabetes mellitus using EQ-5D-5L in Hong Kong, China. Qual Life Res 2020 Jul 05;29(7):1913-1921 [FREE Full text] [doi: 10.1007/s11136-020-02462-0] [Medline: 32140920]
20. Pan C, Zhang R, Luo N, He J, Liu R, Ying X, et al. How the EQ-5D utilities are derived matters in Chinese diabetes patients: a comparison based on different EQ-5D scoring functions for China. Qual Life Res 2020 Nov 12;29(11):3087-3094. [doi: 10.1007/s11136-020-02551-0] [Medline: 32533422]
21. Wong ELY, Xu RH, Cheung AWL. Health-related quality of life among patients with hypertension: population-based survey using EQ-5D-5L in Hong Kong SAR, China. BMJ Open 2019 Sep 27;9(9):e032544 [FREE Full text] [doi: 10.1136/bmjopen-2019-032544] [Medline: 31562165]
22. Weng G, Hong Y, Luo N, Mukuria C, Jiang J, Yang Z, et al. Comparing EQ-5D-3L and EQ-5D-5L in measuring the HRQoL burden of 4 health conditions in China. Eur J Health Econ 2023 Mar 10;24(2):197-207 [FREE Full text] [doi: 10.1007/s10198-022-01465-7] [Medline: 35538178]
23. McNamara S, Schneider PP, Love-Koh J, Doran T, Gutacker N. Quality-Adjusted Life Expectancy Norms for the English Population. Value Health 2023 Feb;26(2):163-169 [FREE Full text] [doi: 10.1016/j.jval.2022.07.005] [Medline: 35965226]
24. Gautier L, Azzi J, Saba G, Bonnelye G, de Pouvourville G. Population norms in France with EQ-5D-5L: health states, value indexes, and VAS. Eur J Health Econ 2023 Dec 10;24(9):1517-1530. [doi: 10.1007/s10198-022-01559-2] [Medline: 36625971]
25. McCaffrey N, Kaambwa B, Currow DC, Ratcliffe J. Health-related quality of life measured using the EQ-5D-5L: South Australian population norms. Health Qual Life Outcomes 2016 Sep 20;14(1):133 [FREE Full text] [doi: 10.1186/s12955-016-0537-0] [Medline: 27644755]
26. Pontikas C, Tsoukalas E, Serdari A. A map of assistive technology educative instruments in neurodevelopmental disorders. Disabil Rehabil Assist Technol 2022 Oct 30;17(7):738-746. [doi: 10.1080/17483107.2020.1839580] [Medline: 33125855]
27. Sequeira AR, Mentzakis E, Archangelidi O, Paolucci F. The economic and health impact of rare diseases: A meta-analysis. Health Policy and Technology 2021 Mar;10(1):32-44. [doi: 10.1016/j.hlpt.2021.02.002]
28. Michalowsky B, Hoffmann W, Mohr W, Rädke A, Xie F. Comparing the psychometric properties of EQ-5D-3L and EQ-5D-5L proxy ratings by informal caregivers and a health professional for people with dementia. Health Qual Life Outcomes 2022 Oct 05;20(1):140 [FREE Full text] [doi: 10.1186/s12955-022-02049-y] [Medline: 36199138]
29. Martin A, Meads D, Griffiths AW, Surr CA. How Should We Capture Health State Utility in Dementia? Comparisons of DEMQOL-Proxy-U and of Self- and Proxy-Completed EQ-5D-5L. Value Health 2019 Dec;22(12):1417-1426 [FREE Full text] [doi: 10.1016/j.jval.2019.07.002] [Medline: 31806199]
30. Tol M, Kuipers J, Willigenburg N, Willems H, Poolman R. How are you doing in the eyes of your spouse? Level of agreement between the self-completed EQ-5D-5L and two proxy perspectives in an orthopaedic population: a randomized agreement study. Health Qual Life Outcomes 2021 Jan 27;19(1):35-16 [FREE Full text] [doi: 10.1186/s12955-021-01679-y] [Medline: 33499876]
31. Hounsome N, Orrell M, Edwards RT. EQ-5D as a quality of life measure in people with dementia and their carers: evidence and key issues. Value Health 2011 Mar;14(2):390-399 [FREE Full text] [doi: 10.1016/j.jval.2010.08.002] [Medline: 21402307]

Abbreviations

HRQoL: health-related quality of life
PA: patient association
RD: rare disease

VAS: visual analog scale

Edited by A Mavragani, T Sanchez; submitted 20.06.23; peer-reviewed by C Guo, M Boarini; comments to author 05.07.23; revised version received 11.07.23; accepted 06.10.23; published 31.10.23.

Please cite as:
Xu RH, Ng SSM, Luo N, Dong D, Zhang S
Measurement of Health-Related Quality of Life in Individuals With Rare Diseases in China: Nation-Wide Online Survey
JMIR Public Health Surveill 2023;9:e50147
URL: https://publichealth.jmir.org/2023/1/e50147
doi:10.2196/50147
PMID:37906229

Original Paper

Sex and Age Differences in the Association Between Social Determinants of Health and Cardiovascular Health According to Household Income Among Mongolian Adults: Cross-Sectional Study

Sun Young Shim[1], MPH, RN; Hyeonkyeong Lee[2], RN, PhD

[1]College of Nursing, Yonsei University, Seoul, Republic of Korea
[2]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Hyeonkyeong Lee, RN, PhD
Mo-Im Kim Nursing Research Institute
College of Nursing
Yonsei University
50 Yonsei-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 22283373
Fax: 82 22278303
Email: hlee39@yuhs.ac

Abstract

Background: Although social determinants of health (SDH) are an underlying cause of poor cardiovascular health (CVH), there is insufficient evidence for the association between SDH and CVH, which varies by sex and age among Mongolian adults.

Objective: We aimed to explore whether education, household income, and health insurance were associated with CVH according to sex and age among Mongolian adults.

Methods: The final sample included data on 5691 participants (male: n=2521. 44.3% and female: n=3170, 55.7%) aged 18-69 years from the 2019 World Health Organization STEPwise approach to noncommunicable disease risk-factor surveillance. CVH was measured using a modified version of Life’s Simple 7 with 4 health behaviors (cigarette smoking, BMI, physical activity, and a healthy diet) and 3 biological factors (blood pressure, fasting glucose, and total cholesterol blood levels) and classified into poor, intermediate, and ideal levels as recommended by the American Heart Association. Multinomial logistic regression analyses examined the associations between SDH and CVH by monthly equivalized household income after adjusting for age, sex, work status, area, history of myocardial infarction or stroke, use of aspirin, and use of statin. Subgroup analyses were conducted to examine the associations between SDH and CVH based on sex and age, considering monthly equivalized household income as a key variable.

Results: Using the ideal level of CVH as a reference, among those with the lowest household income, having less than 12 years of education, and not having health insurance were associated with poor CVH (education level: odds ratio [OR] 2.42, 95% CI 1.30-4.51; *P*=.006; health insurance: OR 2.17, 95% CI 1.13-4.18; *P*=.02). These associations were more profound among female individuals (education level: OR 2.99, 95% CI 1.35-6.63; *P*=.007; health insurance: OR 2.54, 95% CI 1.09-5.90; *P*=.03) and those aged 18-44 years (education level: OR 3.22, 95% CI 1.54-6.72; *P*=.002; health insurance: OR 2.03, 95% CI 0.98-4.18; *P*=.06).

Conclusions: Participants in the lowest household income group with lower educational levels and without health insurance were more likely to have poor CVH, and these results were more pronounced in female individuals and young adults. These findings suggest the need to develop strategies for CVH equity in Mongolian female individuals and young adults that consider income levels, education levels, and health insurance.

(JMIR Public Health Surveill 2023;9:e44569) doi:10.2196/44569

KEYWORDS

social determinants of health; cardiovascular health; education; household income; health insurance; association; risk factors; cardiovascular; cardiovascular disease; cross-sectional study

Introduction

The trend toward universal health coverage in Mongolia seems to have stagnated. Service coverage related to cardiovascular health (CVH) risk factors such as hypertension, diabetes, and smoking is also below the global average; there was no improvement between 2000 and 2019 [1]. Of the 4 noncommunicable diseases (NCDs) most common among Mongolian adults aged 30-70 years, cardiovascular disease (CVD) accounted for the largest proportion (17.3%) of deaths [2]. In a cross-sectional study that examined a nationally representative sample of 70,380 people aged 35-75 years in Inner Mongolia, 25% of the study participants were at high-risk for CVD, with more than 10% at risk of developing CVD after 10 years [3]. However, Mongolian adults' low rates of hypertension control treatment, along with high rates of smoking, sodium intake, and obesity, indicate that CVH is not adequately managed in Mongolian adults [4-7].

There is substantial evidence for the effectiveness of individual lifestyle changes to reduce CVH risk [8-10]; nevertheless, it is necessary to examine situations such as the structural social determinants of health (SDH) in Mongolia, where essential service coverage is not guaranteed. SDH, such as education and health insurance coverage, can be more important than lifestyle choices in influencing population health and health equity because of their connection to CVD pathogenesis. Chronic psychosocial stress can be induced among individuals who are susceptible to SDH, and this stress can lead to chronic inflammation by activating the sympathetic nervous system linked to altered stress hormone. For example, glucocorticoid receptor resistance due to chronic activation of the sympathetic nervous system blunts the anti-inflammatory response; it could contribute to enabling CVD development and progression. [11]. Further, the Denmark study showed that a low education level was independently associated with a higher risk of CVD among 1.6 million Danish employees [12]. In Colombia, people with lower education levels have a higher risk of dying prematurely from CVD, and these education inequalities affect female individuals more than male individuals [13]. A prospective cohort study that examined 390,881 Chinese adults aged 18-64 years with a mean follow-up of 10.4 years revealed that having private health insurance coverage was associated with a 21% lower risk of CVD mortality compared to being uninsured [14].

However, there are insufficient studies showing the influence of SDH on CVH among Mongolian adults. Thus, the identification of SDH should be the foundation for developing strategies that promote CVH and reduce long-term health inequality [15]. Research on this topic can provide critical information that can be used for evidence-based planning and decision-making based on priorities regarding CVH equity. As prevalence of CVD risk differed according to sex and age among Mongolian adults [3,4], the population group characteristics should be considered when confirming the influence of SDH on CVH. Thus, this study aimed to explore the sex and age differences in the independent associations between SDH (educational level and health insurance) and CVH among Mongolian adults.

Methods

Data and Study Participants

Data were obtained from the 2019 Mongolia STEPwise approach to noncommunicable disease risk-factor surveillance (STEPS). The World Health Organization (WHO) initiated STEPS in various countries, recognizing a need for standardized worldwide risk-factor data on NCD risk factors for surveillance systems [16,17]. STEPS was conducted in Mongolia to assess the impact and effect of the integrated national program on "NCD Prevention and Control" [5]. STEPS is a repeated household survey that is population based and cross-sectional. It uses multistage cluster sampling to extract and investigate a nationally representative sample [18]. The fourth Mongolia STEPS survey was carried out using 377 sampling clusters selected from 21 provinces and 9 districts of Ulaanbaatar in 2019 [5]. Of the 6654 people who participated in the 2019 Mongolia STEPS, participants aged younger than 18 years (n=173, 2.6%) and those who had missing data on key variables (n=790, 11.9%) were excluded and 5691 (85.5%) participants were finally included in this secondary data analysis.

Ethical Considerations

All participants provided written informed consent prior to participating in the 2019 Mongolia STEPS. This study was granted an exemption by the institutional review board (4-2021-1341) of Severance Hospital at Yonsei University College of Medicine, Republic of Korea, which allows secondary analyses without additional consent.

SDH Factor

Information about education level, possessing health insurance, and average household earnings in the past year were assessed through face-to-face interviews [16]. Educational attainment was categorized into 2 groups: <12 years of education and ≥12 years of education (reference group). Current health insurance was assessed using the question "Do you currently have any kind of health insurance or health care coverage?" [16]. The average monthly household income was calculated by dividing the average household income in the previous year by 12. Equivalized household income was defined as the average monthly household income divided by the square root of the number of household members [19]. The study participants were divided into 4 groups according to the quartiles of equivalized household income.

CVH Factor

We used the modified Life's Simple 7 by the American Heart Association to measure CVH. It addresses 4 health behaviors, including BMI, smoking, physical activity, and diet, and 3 health risk factors, including blood pressure, fasting glucose levels, and total cholesterol [20]. A cutoff of 23 kg/m^2 for BMI, which has been validated for the Asian population, was used in this study [21]. Smoking, physical activity, and diet were assessed through face-to-face interviews using standardized questionnaires [16]. Salt consumption for diet was used instead of the American Heart Association's guidelines [22]. Blood pressure was measured 3 times for 3-minute intervals. Each measurement was taken on the left arm, with the participants in a seated posture [18]. The mean of the last 2 measurements was used in the analysis [18]. Overnight fasting blood samples were collected at a local clinic or health center [18]. CVH metrics were scored as poor (0 points), intermediate (1 point), and ideal (2 points). The total CVH scores were calculated as the sum of the scores for each component for a possible range of 0-14 (Table S1 in Multimedia Appendix 1). Scores of 0-7 were considered as poor, 8-11 as intermediate, and 12-14 as ideal for CVH [23].

Covariates

We adjusted for sex, age (18-39, 40-59, and ≥60 years), work status (employee, self-employed, or others), area (rural or urban), history of myocardial infarction or stroke (yes or no), use of aspirin (yes or no), and use of statin (yes or no). The history of myocardial infarction or stroke was investigated using the question "Have you ever had a heart attack or chest pain from heart disease (angina) or a stroke (cerebrovascular accident or incident)?" The use of aspirin or statin was assessed using the questions "Are you currently taking aspirin regularly to prevent or treat heart disease?" and "Are you currently taking statins (Lovastatin/Simvastatin/Atorvastatin or any other statin) regularly to prevent or treat heart disease?" respectively. The covariate data were collected by self-reporting.

Statistical Analysis

Descriptive values were calculated for the total participant group as well as the male and female groups, represented as a weighted mean (95% CI) and number (weighted proportion). The differences in variables according to CVH categories in male and female participants were calculated using ANOVA for continuous variables and the chi-square test for categorical variables. Independent associations between SDH (education level and health insurance) and CVH were assessed through multinomial logistic regression analyses between the sexes after adjusting for covariates. Subgroup analyses regarding associations between SDH and CVH according to monthly equivalized household income were conducted among the total participants. This subgroup analysis according to monthly equivalized household income was also performed by sex and age to identify different sex and age associations between SDH and CVH by household income. The period of public education in Mongolia gradually changed from 10 years in 2004 to 12 years in 2013; students who graduated in 2013 were the first to have received 12 years of public education. Therefore, sensitivity analyses were conducted by defining low education level as less than 11 years of education and examining the association between education level and CVH according to monthly equivalized household income. Adjusted odds ratios (ORs) and 95% CIs were calculated using the PROC SURVEYLOGISTIC procedures to apply population weights, and all statistical tests were performed using SAS (version 9.4; SAS Institute Inc).

Results

The general characteristics of study participants by sex according to CVH are described in Multimedia Appendix 2. Of the total 5691 participants, 29.6% (n=1685), 56.4% (n=3208), and 14% (n=798) showed poor, intermediate, and ideal CVH, respectively. The mean CVH score of study participants was 8.80 (SD 2.46). The weighted mean age of study participants was 37.5 (SE 0.26) years, and more than two-thirds (2888/5691, 69.3%) were aged 18-44 years. Approximately half (n=3075, 49.6%) of the participants had less than 12 years of education, and one-fifth (n=888, 19.5%) had no health insurance. In male participants, 41.1% (n=1036), 50.2% (n=1265), and 8.7% (n=220) showed poor, intermediate, and ideal CVH, respectively. The mean age, the proportion of employees, self-employment, and use of aspirin were likely to be higher in participants with poor CVH (all *P*<.001). In female participants, 20.5% (n=649), 61.3% (n=1943), and 18.2% (n=578) showed poor, intermediate, and ideal CVH, respectively. The mean age and the proportion of aspirin and of statin use were likely to be higher in those with poor CVH (all *P*<.001). Those with less than 12 years of education were more likely to have poorer CVH (*P*=.001).

Table 1 shows the results of multinomial logistic regression regarding the association between education, health insurance, and CVH categories. Of the total participants, using ideal CVH as a reference, those with less than 12 years of education were not associated with intermediate and poor CVH after adjusting for potential covariates (intermediate CVH: OR 1.12, 95% CI 0.91-1.40; poor CVH: OR 1.00, 95% CI 0.78-1.29). Similarly, of the total participants, those without health insurance were not associated with intermediate and poor CVH, respectively (intermediate CVH: OR 0.97, 95% CI 0.73-1.28; poor CVH: OR 1.19, 95% CI 0.85-1.67). For both male and female participants, low education level and absence of health insurance were not associated with intermediate and poor CVH.

However, as a result of stratification according to quartiles of monthly household income, participants with less than 12 years of education in the Q1 (lowest household income) and Q2 groups were associated with a 2.42 (95% CI 1.30-4.51) and 1.90 (95% CI 1.08-3.33) times higher likelihood of poor CVH, respectively (Figure 1 and Table S2 in Multimedia Appendix 1). These associations were consistently found when defining low education level as less than 11 years (data not shown). Participants without health insurance in the Q1 group were also associated with a 2.17 (95% CI 1.13-4.18) times higher likelihood of poor CVH.

Table 1. Results of multinomial logistic regression of the association between education, health insurance, and CVH[a] by modified Life's Simple 7[b,c].

Variables	Intermediate CVH		Poor CVH	
	n (%)	OR[d] (95% CI)	n (%)	OR (95% CI)
Education				
All participants (N=5691)				
≥12 years	1458 (55.7)	1.00 (N/A[e])	721 (27.6)	1.00 (N/A)
<12 years	1750 (56.9)	1.12 (0.91-1.40)	964 (31.4)	1.00 (0.78-1.29)
Male participants (n=2521)				
≥12 years	497 (50.0)	1.00 (N/A)	405 (40.7)	1.00 (N/A)
<12 years	768 (50.3)	0.98 (0.6-1.40)	631 (41.4)	0.95 (0.64-1.40)
Female participants (n=3170)				
≥12 years	961 (59.3)	1.00 (N/A)	316 (19.5)	1.00 (N/A)
<12 years	982 (63.4)	1.24 (0.94-1.64)	333 (21.5)	0.97 (0.68-1.38)
Health insurance				
All participants (N=5691)				
Yes	2718 (56.6)	1.00 (N/A)	1415 (29.5)	1.00 (N/A)
No	490 (55.2)	0.97 (0.73-1.28)	270 (30.4)	1.19 (0.85-1.67)
Male participants (n=2521)				
Yes	1014 (50.1)	1.00 (N/A)	847 (41.8)	1.00 (N/A)
No	251 (50.7)	0.86 (0.56-1.33)	189 (38.2)	1.00 (0.63-1.58)
Female participants (n=3170)				
Yes	1704 (61.4)	1.00 (N/A)	568 (20.5)	1.00 (N/A)
No	239 (60.8)	1.05 (0.75-1.49)	81 (20.6)	1.47 (0.95-2.27)

[a] CVH: cardiovascular health.

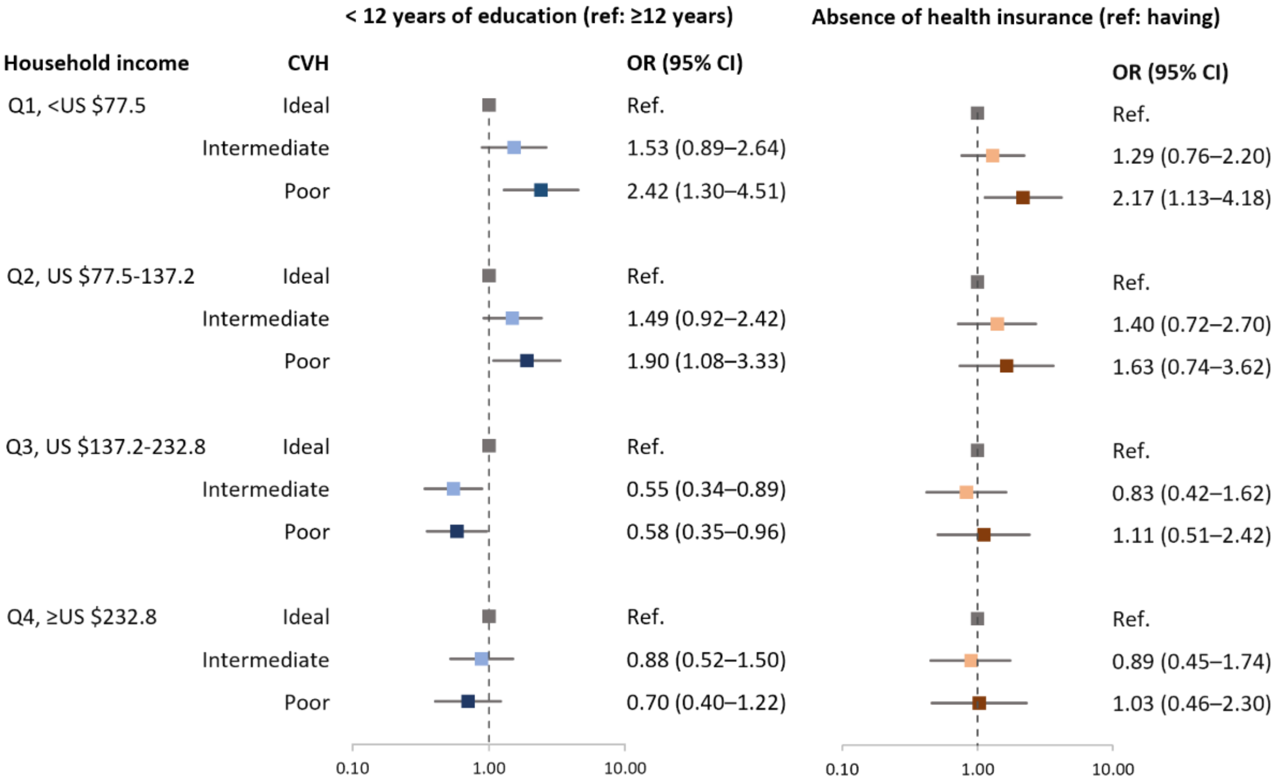
[b] The estimates are in reference to the ideal CVH group.

[c] Age, sex, work status, area, history of myocardial infarction or stroke, use of aspirin, and use of statin were adjusted.

[d] OR: odds ratio.

[e] N/A: not available.

Figure 1. Results of multinomial logistic regression of the association between social determinants of health and CVH according to household income. Age, sex, work status, area, history of myocardial infarction or stroke, use of aspirin, and use of statin were adjusted. CVH: cardiovascular health; OR: odds ratio; Ref: reference.



The associations between education level and CVH according to monthly household income quartile by sex and age are illustrated in Figure 2. Female participants in the Q1 (lowest household income) group with less than 12 years of education were associated with a 2.99 (95% CI 1.35-6.63) times higher likelihood of poor CVH compared to female participants with more than 12 years of education (Table S3 in Multimedia Appendix 1). However, male participants with less than 12 years of education were not associated with intermediate and poor CVH in the Q1 group (intermediate CVH, OR 0.78, 95% CI 0.31-1.96; poor CVH, OR 1.44, 95% CI 0.56-3.72). Participants in the Q1 group who were aged 18-44 years and had less than 12 years of education were also associated with a 3.22 (95% CI 1.54-6.72) times higher likelihood of poor CVH. Participants in the Q3 and Q4 groups, aged 45-69 years, with less than 12 years of education were associated with a lower likelihood of poor CVH (Q3, OR 0.25, 95% CI 0.10-0.60; Q4, OR 0.36, 95% CI 0.13-0.97).

Similarly, female participants in the Q1 (lowest household income) group without health insurance were associated with a 2.54 (95% CI 1.09-5.90) times higher likelihood of poor CVH compared to female participants with health insurance (Figure 3 and Table S4 in Multimedia Appendix 1). However, male participants in the Q1 group without health insurance were not associated with poor CVH (OR 2.02, 95% CI 0.73-5.57). Participants in the Q1 group, aged 18-44 years, without health insurance were marginally associated with a 2.03 (95% CI 0.98-4.18) times higher likelihood of poor CVH. Participants in the Q1 group, aged 45-69 years, without health insurance were not associated with poor CVH (OR 2.67, 95% CI 0.60-11.95). The OR for the Q3 group of participants aged 45-69 years was not accounted for as the model did not converge due to quasi-complete separation.

Figure 2. Results of multinomial logistic regression of the association between low education level and CVH according to household income by sex and age. The reference is ≥12 years of education. Adjusted for all other factors not involving the subgroup, including age, sex, work status, area, history of myocardial infarction or stroke, use of aspirin, and use of statin were adjusted. CVH: cardiovascular health.

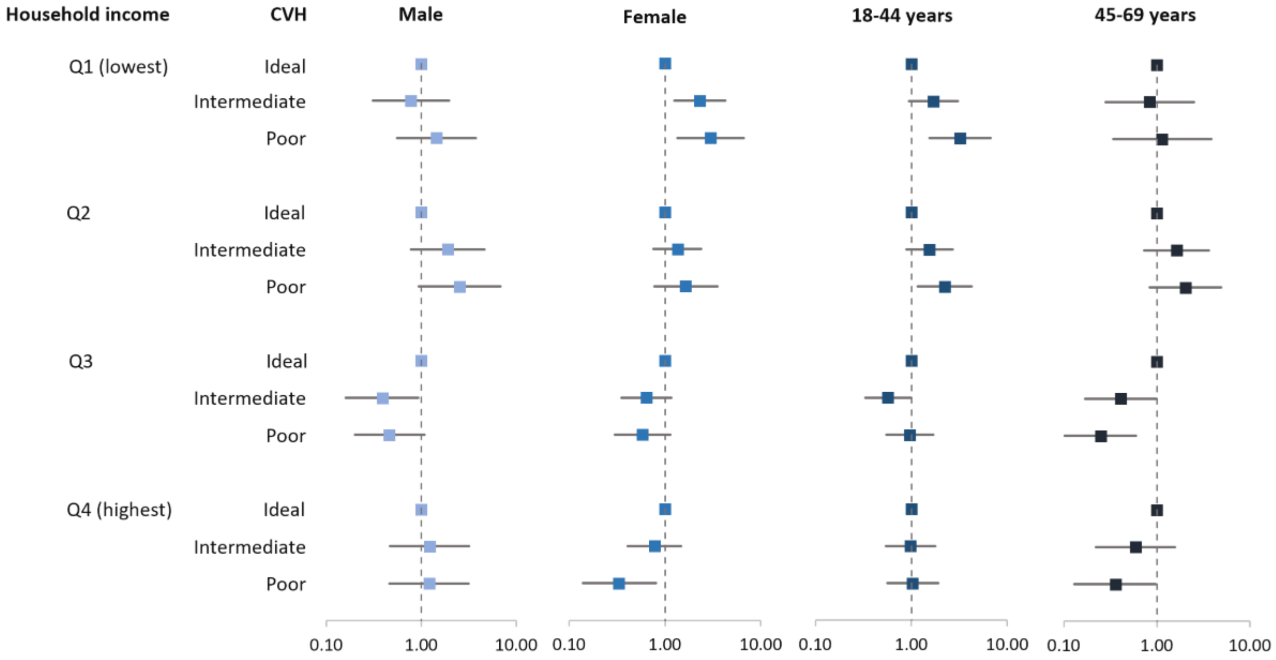
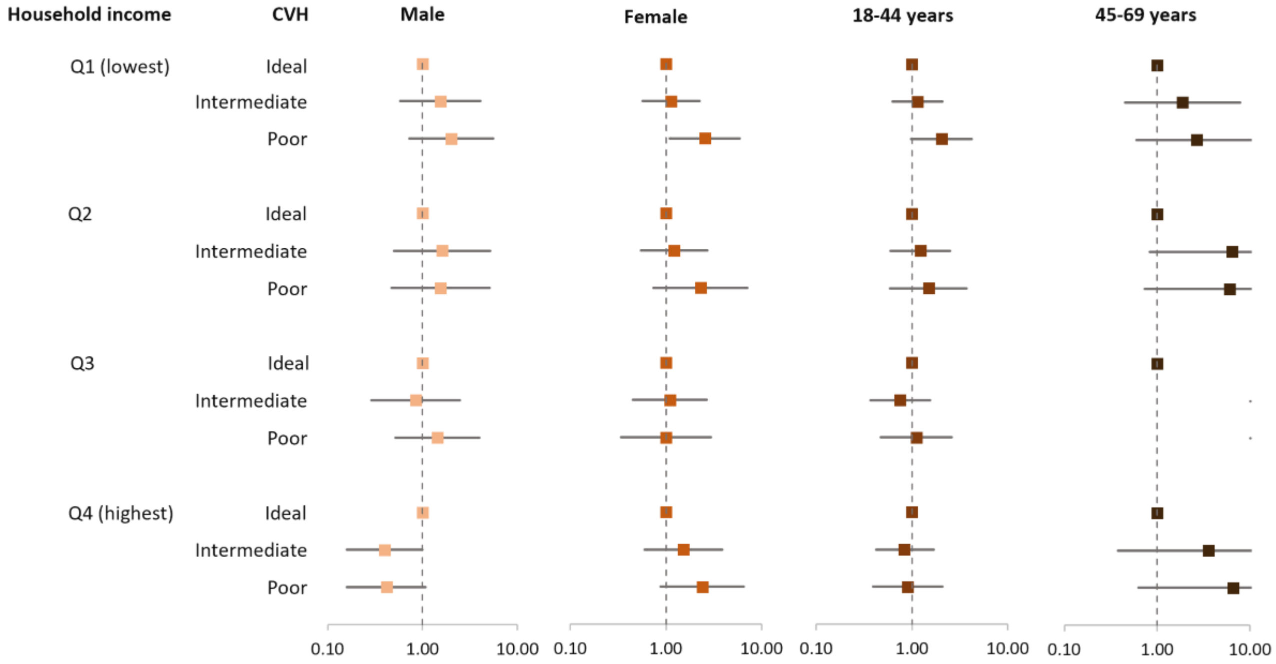


Figure 3. Results of multinomial logistic regression of the association between absence of health insurance and CVH according to household income by sex and age. The reference is having health insurance. Adjusted for all other factors not involving the subgroup, including age, sex, work status, area, history of myocardial infarction or stroke, use of aspirin, and use of statin were adjusted. CVH: cardiovascular health.



Discussion

Principal Findings

For all participants, having less education or no health insurance was not associated with CVH; however, stratified analyses by monthly household income quartiles showed that having less education or no health insurance was associated with poor CVH in the lowest household income groups. These CVH disparities were more pronounced among female individuals and young adults. This study contributes to establishing policies implementation that reduces the disparity of CVH, the most notable cause of disease burden, by presenting the basis for prioritizing a vulnerable population. For example, health service delivery guidelines should include consideration of the link between SDH and CVD, and training for health care providers should emphasize this. Low levels of education may be associated with low levels of health literacy for CVD prevention and management practices [24,25]. Therefore, health care providers should assess the health literacy levels of their target population and provide interventions that take this into account.

There was a gap in the prevalence of CVD risk factors according to the WHO region or income level; Mongolia is a country with a high prevalence of CVD risk factors [26]. For example, nonelevated blood pressure, one of the tracer indicators for calculating the universal health coverage service coverage index, was 29, indicating the low coverage of blood pressure management in Mongolia [26]. Further, the linear trends in cardiovascular risk factors such as smoking, obesity, and elevated blood pressure among Mongolian adults are all projected to be significantly higher than the global target by 2025 [27]. Thus, there is a need for the adoption and dissemination of cardiovascular risk-factor management guidelines. The WHO-initiated STEPS in 2002 is practical and helpful in monitoring the status of CVH and can identify the achievement rate for each cardiovascular risk factor, thereby aiding the formulation of effective health policies for managing risk factors at a country level [16]. Establishing a high-vulnerability priority population in population health is essential for ensuring universal health in which no one is excluded.

The association between education level and CVH is inconsistent in previous studies. A cross-sectional study of 7771 nationally representative participants aged ≥25 years in the United States reported that participants with at least a college degree were 4.12 times more likely to have ideal CVH status than participants with less than a high school education [28]. However, a recent cross-sectional study of 1634 Asian American adults aged ≥20 years in the United States revealed that low educational attainment was associated with a low likelihood of having ideal CVH compared to high educational attainment after adjusting for age and sex; however, this association became nonsignificant after additionally adjusting for income category and nativity status [26]. Further studies are needed to comprehensively understand the relationship between educational attainment and CVH in the contexts of including other SDH, such as income and race, for health equity in populations that are vulnerable to CVH problems.

Mongolia implemented a social health insurance system in 1994 to maintain equity in access to health care. Consequently, 95.6% of the population has been covered by social health insurance since 2017 [29,30]. Despite this high health insurance coverage rate, health expenses are significantly related to poverty in Mongolia, with pronounced consequences in households with patients with NCD [31]. Our study also showed that the population without health insurance had a higher risk of poor CVH status in the lowest income group than in the higher income groups, which is consistent with the previous study [32].

In this study, the CVH disparities according to SDH were predominant in female individuals and young adults. Mongolia is ranked 69th out of 156 countries in the global gender gap index, and the score of estimated earned income among the subitems measuring this index was 0.67 (0.00=imparity, 1.00=parity) [33]. Numerous previous studies support the sex differences in cardiovascular risk factors and diseases [34-36]. Our study revealed that low education level or lack of health insurance contributed more to poor CVH, especially among low-income female individuals. Furthermore, a history of CVD in Mongolians currently appears in relatively young population groups; an increase in CVD at a young age is a remarkable health issue [5]. Therefore, efforts to promote CVH that consider sex and age, such as interventions for modifying CVH for young adults or creating strategies that target gender equality, might be needed to achieve health equity in Mongolia. In countries with limited universal health coverage such as Mongolia, a strategy of using text messages to target younger populations may be useful in managing hypertension to prevent CVD, as identified in a previous research review [37].

The strength of this study was that, to the best of our knowledge, this study is the first to identify the sex and age differences in the association between SDH and CVH according to income among Mongolian adults. Further, we used nationally representative Mongolian data to minimize external validity threats.

Limitations

This study has some limitations. First, it is possible that participants who had less than 12 years of education but completed all public education courses were grouped together with those who did not complete public education courses. This can be attributed to the gradual change of public education in Mongolia from 10 years in 2004 to 12 years in 2013. Thus, we performed the sensitivity analysis by dividing the education level based on 11 years and confirmed that the results were similar to our primary results. Second, we measured salt consumption as the diet component of CVH, which is not a guideline in Life’s Simple 7. Third, there is a possibility of misclassification of self-reported variables. Fourth, we used the most recent fourth survey to determine the latest findings on the association between SDH and CVH. The STEPS Mongolia cross-sectional survey began in 2005 and has been conducted 4 times every 4 to 6 years. Further research is needed to identify trends in CVH following structural SDH in order to prioritize risk factors that have not been improved over time. Fifth, the STEPS data did not provide information on individual residence, limiting our ability to determine the impact of residence on health. Although we found that living in a rural area was associated with lower CVH compared with urban area, we suggest including variables related to residence because even within the same rural area, there may be differences in the environment in which people live (Table S5 in Multimedia Appendix 1). Lastly, although we adjusted potential cofounders, there might still be a chance of residual confounding we did not control for.

Conclusions

In conclusion, we confirmed that in Mongolia, where universal health coverage is low, the population group that has low structural socioeconomic levels and health care coverage is more vulnerable to CVH problems. As in other low-income countries, female individuals are more vulnerable to poor CVH than male individuals, suggesting that this vulnerability can only be improved when fundamental problems such as structural health determinants are addressed. Strategies and interventions aimed at reducing the CVH disparity in the Mongolian population should be prioritized for female individuals and young adults (aged 18-44 years) with low socioeconomic status and without health insurance.

Acknowledgments

This paper uses data from the Mongolia 2019 STEPwise approach to noncommunicable disease risk-factor surveillance survey, implemented by the Minister of Health, Mongolia, with the support of the World Health Organization. This research was supported by the Brain Korea 21 FOUR Project funded by National Research Foundation of Kora, Yonsei University College of Nursing.

Data Availability

The data sets generated or analyzed during this study are available in The World Health Organization noncommunicable disease microdata repository (World Health Organization, June 20, 2023, STEPwise approach to noncommunicable disease risk-factor surveillance 2019 Mongolia).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary table 1-5.
[DOCX File , 42 KB - publichealth_v9i1e44569_app1.docx]

Multimedia Appendix 2
General characteristics of study participants according to cardiovascular health by the modified Life's Simple 7.
[DOCX File , 23 KB - publichealth_v9i1e44569_app2.docx]

References

1. Global health observatory data repository. World Health Organization. 2022. URL: https://apps.who.int/gho/data/view.main.INDEXOFESSENTIALSERVICECOVERAGEv [accessed 2023-11-12]
2. Health indicators-2019. Center for Health Development. URL: http://hdc.gov.mn/media/uploads/2021-05/Health_Indicator_2019_ENG.pdf [accessed 2023-11-12]
3. Xi Y, Cao N, Niu L, Zhu H, Bao H, Qiao L, et al. Prevalence and treatment of high cardiovascular disease risk in Inner Mongolia, China. Rev Cardiovasc Med 2021;22(2):521-529 [FREE Full text] [doi: 10.31083/j.rcm2202060] [Medline: 34258921]
4. Potts H, Baatarsuren U, Myanganbayar M, Purevdorj B, Lkhagvadorj BU, Ganbat N, et al. Hypertension prevalence and control in Ulaanbaatar, Mongolia. J Clin Hypertens (Greenwich) 2020;22(1):103-110 [FREE Full text] [doi: 10.1111/jch.13784] [Medline: 31913578]
5. World Health Organization. Fourth national steps survey on the prevalence of noncommunicable disease and injury risk factors-2019-2020. WHO Representative Office for Mongolia. 2019. URL: https://cdn.who.int/media/docs/default-source/ncds/ncd-surveillance/data-reporting/mongolia/mongolia-steps-survey---2019_brief-summary_english.pdf [accessed 2023-11-12]
6. Tuvdendorj A, Feenstra T, Tseveen B, Buskens E. Smoking-attributable burden of lung cancer in Mongolia a data synthesis study on differences between men and women. PLoS One 2020;15(2):e0229090 [FREE Full text] [doi: 10.1371/journal.pone.0229090] [Medline: 32059049]
7. Chimeddamba O, Gearon E, Stevenson C, Ng WL, Baasai B, Peeters A. Trends in adult overweight and obesity prevalence in Mongolia, 2005-2013. Obesity (Silver Spring) 2016;24(10):2194-2201 [FREE Full text] [doi: 10.1002/oby.21595] [Medline: 27580924]
8. Chung MK, Eckhardt LL, Chen LY, Ahmed HM, Gopinathannair R, Joglar JA, et al. Lifestyle and risk factor modification for reduction of atrial fibrillation: a scientific statement from the American Heart Association. Circulation 2020;141(16):e750-e772 [FREE Full text] [doi: 10.1161/CIR.0000000000000748] [Medline: 32148086]
9. Lee LL, Mulvaney CA, Wong YKY, Chan ES, Watson MC, Lin HH. Walking for hypertension. Cochrane Database Syst Rev 2021;2(2):CD008823 [FREE Full text] [doi: 10.1002/14651858.CD008823.pub2] [Medline: 33630309]
10. Blumenthal JA, Hinderliter AL, Smith PJ, Mabe S, Watkins LL, Craighead L, et al. Effects of lifestyle modification on patients with resistant hypertension: results of the TRIUMPH randomized clinical trial. Circulation 2021;144(15):1212-1226 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.121.055329] [Medline: 34565172]
11. Powell-Wiley TM, Baumer Y, Baah FO, Baez AS, Farmer N, Mahlobo CT, et al. Social determinants of cardiovascular disease. Circ Res 2022;130(5):782-799 [FREE Full text] [doi: 10.1161/circresaha.121.319811]

12. Framke E, Sørensen JK, Andersen PK, Svane-Petersen AC, Alexanderson K, Bonde JP, et al. Contribution of income and job strain to the association between education and cardiovascular disease in 1.6 million Danish employees. Eur Heart J 2020;41(11):1164-1178 [FREE Full text] [doi: 10.1093/eurheartj/ehz870] [Medline: 31844881]
13. Jimenez M, Arroyave I. How educational inequalities in cardiovascular mortality evolve while healthcare insurance coverage grows: Colombia, 1998 to 2015. Value Health Reg Issues 2020;23:112-121 [FREE Full text] [doi: 10.1016/j.vhri.2020.08.001] [Medline: 33217714]
14. Song L, Wang Y, Chen B, Yang T, Zhang W, Wang Y. The association between health insurance and all-cause, cardiovascular disease, cancer and cause-specific mortality: a prospective cohort study. Int J Environ Res Public Health 2020;17(5):1525 [FREE Full text] [doi: 10.3390/ijerph17051525] [Medline: 32120888]
15. Marmot M, Friel S, Bell R, Houweling TAJ, Taylor S, Commission on Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health. Lancet 2008;372(9650):1661-1669. [doi: 10.1016/S0140-6736(08)61690-6] [Medline: 18994664]
16. Riley L, Guthold R, Cowan M, Savin S, Bhatti L, Armstrong T, et al. The World Health Organization STEPwise approach to noncommunicable disease risk-factor surveillance: methods, challenges, and opportunities. Am J Public Health 2016;106(1):74-78 [FREE Full text] [doi: 10.2105/AJPH.2015.302962] [Medline: 26696288]
17. STEPwise approach to NCD risk factor surveillance (STEPS). World Health Organization. 2022. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps [accessed 2023-11-12]
18. WHO STEPS surveillance manual: the WHO STEPwise approach to chronic disease risk factor surveillance. Report no.: 9241593830. World Health Organization. 2005. URL: https://www.who.int/docs/default-source/ncds/ncd-surveillance/steps/steps-manual.pdf?sfvrsn=c281673d_5 [accessed 2023-11-12]
19. What are equivalence scales? Organisation for Economic Copperation and Development (OECD). 2005. URL: https://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf [accessed 2023-11-12]
20. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic impact goal through 2020 and beyond. Circulation 2010;121(4):586-613 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192703] [Medline: 20089546]
21. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
22. Liu H, Yao Y, Wang Y, Ma L, Liu X, Guo S, et al. Ideal cardiovascular health metrics and the risk of non-alcoholic fatty liver disease: a cross-sectional study in Northern China. Liver Int 2019;39(5):950-955. [doi: 10.1111/liv.14060] [Medline: 30712318]
23. Corlin L, Short MI, Vasan RS, Xanthakis V. Association of the duration of ideal cardiovascular health through adulthood with cardiometabolic outcomes and mortality in the Framingham offspring study. JAMA Cardiol 2020;5(5):549-556 [FREE Full text] [doi: 10.1001/jamacardio.2020.0109] [Medline: 32159731]
24. Kanejima Y, Shimogai T, Kitamura M, Ishihara K, Izawa KP. Impact of health literacy in patients with cardiovascular diseases: a systematic review and meta-analysis. Patient Educ Couns 2022;105(7):1793-1800. [doi: 10.1016/j.pec.2021.11.021] [Medline: 34862114]
25. Lindahl B, Norberg M, Johansson H, Lindvall K, Ng N, Nordin M, et al. Health literacy is independently and inversely associated with carotid artery plaques and cardiovascular risk. Eur J Prev Cardiol 2020;27(2):209-215 [FREE Full text] [doi: 10.1177/2047487319882821] [Medline: 31615294]
26. Alam MT, Echeverria SE, DuPont-Reyes MJ, Vasquez E, Murillo R, Gonzalez T, et al. Educational attainment and prevalence of cardiovascular health (Life's Simple 7) in Asian Americans. Int J Environ Res Public Health 2021;18(4):1480 [FREE Full text] [doi: 10.3390/ijerph18041480] [Medline: 33557415]
27. Noncommunicable diseases Mongolia. World Health Organization. 2018. URL: https://cdn.who.int/media/docs/default-source/country-profiles/ncds/mng_en.pdf [accessed 2023-11-12]
28. Johnson AE, Herbert BM, Stokes N, Brooks MM, Needham BL, Magnani JW. Educational attainment, race, and ethnicity as predictors for ideal cardiovascular health: from the National Health and Nutrition Examination Survey. J Am Heart Assoc 2022;11(2):e023438 [FREE Full text] [doi: 10.1161/JAHA.121.023438] [Medline: 34984911]
29. Mongolia health indicators 2017. World Health Organization Regional Office for the Western Pacific (WPRO-WHO). Mongolia: Center for Health Development (Mongolia) URL: http://hdc.gov.mn/media/files/2017%20eng.pdf [accessed 2023-11-12]
30. Dorjdagva J, Batbaatar E, Svensson M, Dorjsuren B, Togtmol M, Kauhanen J. Does social health insurance prevent financial hardship in Mongolia? inpatient care: a case in point. PLoS One 2021;16(3):e0248518 [FREE Full text] [doi: 10.1371/journal.pone.0248518] [Medline: 33788865]
31. Dugee O, Sugar B, Dorjsuren B, Mahal A. Economic impacts of chronic conditions in a country with high levels of population health coverage: lessons from Mongolia. Trop Med Int Health 2019;24(6):715-726 [FREE Full text] [doi: 10.1111/tmi.13231] [Medline: 30870575]

32. McClurkin MA, Yingling LR, Ayers C, Cooper-McCann R, Suresh V, Nothwehr A, et al. Health insurance status as a barrier to ideal cardiovascular health for U.S. adults: data from the National Health and Nutrition Examination Survey (NHANES). PLoS One 2015;10(11):e0141534 [FREE Full text] [doi: 10.1371/journal.pone.0141534] [Medline: 26535890]
33. Crotti R, Pal KK, Ratcheva RZV. Global gender gap report 2021. World Economic Forum. 2021. URL: https://www3.weforum.org/docs/WEF_GGGR_2021.pdf [accessed 2023-11-12]
34. Appelman Y, van Rijn BB, Haaf MET, Boersma E, Peters SAE. Sex differences in cardiovascular risk factors and disease prevention. Atherosclerosis 2015;241(1):211-218. [doi: 10.1016/j.atherosclerosis.2015.01.027] [Medline: 25670232]
35. Colafella KMM, Denton KM. Sex-specific differences in hypertension and associated cardiovascular disease. Nat Rev Nephrol 2018;14(3):185-201. [doi: 10.1038/nrneph.2017.189] [Medline: 29380817]
36. EUGenMed Cardiovascular Clinical Study Group, Regitz-Zagrosek V, Oertelt-Prigione S, Prescott E, Franconi F, Gerdts E, et al. Gender in cardiovascular diseases: impact on clinical manifestations, management, and outcomes. Eur Heart J 2016;37(1):24-34 [FREE Full text] [doi: 10.1093/eurheartj/ehv598] [Medline: 26530104]
37. Tam HL, Leung LYL, Wong EML, Cheung K, Chan ASW. Integration of text messaging interventions into hypertension management among older adults: a systematic review and meta-analysis. Worldviews Evid Based Nurs 2022;19(1):16-27. [doi: 10.1111/wvn.12549] [Medline: 35014147]

Abbreviations

CVH: cardiovascular health
CVD: cardiovascular disease
NCD: noncommunicable disease
OR: odds ratio
SDH: social determinants of health
STEPS: STEPwise approach to noncommunicable disease risk-factor surveillance
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 24.11.22; peer-reviewed by HL Tam, Z Chen; comments to author 09.06.23; revised version received 11.08.23; accepted 03.10.23; published 01.12.23.

Please cite as:
Shim SY, Lee H
Sex and Age Differences in the Association Between Social Determinants of Health and Cardiovascular Health According to Household Income Among Mongolian Adults: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e44569
URL: https://publichealth.jmir.org/2023/1/e44569
doi:10.2196/44569
PMID:38039072

Original Paper

Hepatitis C Knowledge and Self-Reported Testing Behavior in the General Population in China: Online Cross-Sectional Survey

Yin Liu[1], PhD; Juan Su[2], BA; Xiaoyang Wang[1], PhD; Huifang Xu[1], PhD; Hong Wang[1], PhD; Ruihua Kang[1], PhD; Liyang Zheng[1], BA; Yixian Wang[1], BA; Chunya Liu[1], BA; Yiping Jing[1], BA; Shaokai Zhang[1], PhD

[1]Department of Cancer Epidemiology, The Affiliated Cancer Hospital of Zhengzhou University (Henan Cancer Hospital), Zhengzhou, China
[2]Yinchuan Hospital of Stomatology, Yinchuan, China

Corresponding Author:
Shaokai Zhang, PhD
Department of Cancer Epidemiology
The Affiliated Cancer Hospital of Zhengzhou University (Henan Cancer Hospital)
127 Dongming Road
Zhengzhou, 450008
China
Phone: 86 65587361
Email: shaokaizhang@126.com

Abstract

Background: The World Health Organization has proposed a worldwide target of eliminating hepatitis C virus (HCV) by 2030. A better understanding of HCV, testing behaviors, and associated factors in the general population is essential.

Objective: This study aimed to assess HCV knowledge, self-reported HCV testing behavior, and willingness to undergo HCV screening in the general Chinese population.

Methods: A cross-sectional online survey of the general Chinese population aged ≥15 years was conducted from November 2021 to May 2023. Participant characteristics were assessed based on their knowledge level and uptake of HCV testing. Participants ever having heard of HCV were recognized as being aware of HCV and asked additional HCV knowledge questions using a brief, validated 9-item scale. Participants with 0-3 points and who were unaware of HCV were categorized as having poor knowledge, and those with 4-6 points and ≥7 points were categorized as having fair and good knowledge, respectively. Participant uptake of HCV testing, testing results, reasons for undergoing or not undergoing HCV testing, and willingness to undergo HCV screening were collected through self-reports. Ordinal and binary logistic regression analyses were used to assess factors associated with the HCV knowledge level and the uptake of HCV testing, respectively.

Results: A total of 1491 valid participants' questionnaires were included. Of these, 714 (47.6%) participants were aware of HCV. The proportion of participants with poor, fair, and good HCV knowledge was 63.4% (945/1491), 9.3% (139/1491), and 27.3% (407/1491), respectively. A total of 465 (31.2%) participants reported ever undergoing HCV testing, and 4 (0.9%) were anti-HCV antibody positive. Most participants were tested for HCV following blood donation (353/465, 75.9%). The most common reasons for not undergoing HCV screening were a lack of HCV awareness (665/1026, 64.8%), followed by a low self-perceived risk of infection (176/1026, 17.2%). Of 1026 participants who had never undergone HCV testing, 937 (91.3%) were willing to undergo HCV screening if universal screening was provided at no cost. The HCV knowledge level was positively associated with the HCV testing rate. Participants who were less educated, lived in rural areas, resided in West China, and were currently alcohol drinkers had lower HCV knowledge and reduced odds of having undergone HCV testing. In contrast, participants with a blood donation history and a family history of hepatitis B virus or HCV infection had higher HCV knowledge and increased odds of prior testing. Participants aged ≥60 years had lower knowledge, and women had reduced odds of having undergone previous HCV testing.

Conclusions: The general population of China has low HCV knowledge and testing rate. There is an urgent need for enhanced HCV awareness and scaled-up HCV screening and treatment. Individuals who are less well educated, reside in less-developed areas, currently drink alcohol, and are female should be prioritized for health education and interventions.

(JMIR Public Health Surveill 2023;9:e39472) doi:10.2196/39472

KEYWORDS

HCV; hepatitis C virus; knowledge; testing behaviors; general population; cross-sectional; online survey; testing; screening; patient education; China; Chinese; patient education; health education; hepatitis; viral disease; viral infection; communicable disease; liver

Introduction

Hepatitis C virus (HCV) infection is a major cause of acute and chronic liver diseases (eg, liver cirrhosis and hepatocellular carcinoma) worldwide and is linked to an estimated 29,000 deaths each year [1]. About 58 million people have HCV, with approximately 1.5 million new infections occurring annually [1]. China has the largest number of patients with HCV infection, accounting for approximately 14% of worldwide HCV infections [2,3].

In 2016, the World Health Organization (WHO) responded to the high worldwide burden of HCV by aiming to reduce new infections by 80% and deaths by 65% by 2030 [4]. Given the lack of an effective vaccine, WHO prioritized improving HCV knowledge and scaling up screening, care, and treatment to meet its targets. Adequate HCV knowledge is critical for the prevention of HCV infection and transmission and for individuals to seek appropriate testing, care, and treatment [5]. In China, studies have explored HCV knowledge among groups at high risk of HCV infection and transmission, including people receiving methadone maintenance treatment (MMT) [6] and men who have sex with men (MSM) [7]. However, no studies have assessed HCV knowledge in the general population. Although most new HCV infections are among high-risk groups, many people are either not aware of or refuse to report risk behaviors because of potential discrimination and therefore remain hidden in the general population. This puts additional people at risk of HCV exposure through contaminated blood, unsafe sexual contact, and mother-to-child transmission [8,9]. A meta-analysis found that the national prevalence of HCV infection in the general Chinese population was about 0.91% [10], constituting a large subgroup with the ability to further spread HCV. Thus, an assessment of HCV knowledge and associated factors in the general Chinese population is important to inform HCV elimination plans.

Scaling up screening, care, and treatment is also critical to achieve WHO targets. People with HCV infection can be easily screened with an anti-HCV antibody test, and more than 95% of these individuals can be cured using highly effective oral direct acting antiviral (DAA) treatment regimens [11]. In China, DAAs have been approved for use against HCV infection since 2017 and have been listed in the medical insurance directory since 2020. However, because HCV is often asymptomatic, only 18% of people with HCV infection are diagnosed, leading to a low treatment rate [12]. As a result, a low HCV testing rate is the main barrier to HCV elimination. In China, HCV screening is primarily limited to high-risk groups [13]. Economic analyses have revealed that screening alongside DAA treatment is cost-effective for both the high-risk population and the general population [14-17]. Thus, this strategy may be necessary to achieve the WHO target. However, there is a lack of information about HCV testing behavior and willingness to undergo HCV screening in the general Chinese population.

This study was conducted to assess HCV knowledge, self-reported HCV testing behavior, and willingness to undergo HCV screening in the general Chinese population to inform policy decision-making aimed at eliminating HCV. The characteristics of patients who have less knowledge and are less likely to undergo testing were also explored.

Methods

Design and Setting

A cross-sectional, online survey was conducted in China from November 2021 to May 2023. The convenience sampling method was used to recruit participants. Individuals who lived in China and were ≥15 years of age were eligible for the study. Those who were health care workers, including physicians, nurses, nurse assistants, pharmacists and pharmacy technicians, radiographers, community health workers, health officers, hospital orderlies, and other health care professionals, or health care students, were excluded.

Participant Characteristics

The following information was collected: (1) sociodemographic characteristics, including age, gender, marital status (married, single/divorced/widowed), years of education (<7 years, 7-12 years, >12 years), ethnicity (Han, minority), occupation (student, not working/unemployed, others), residence (rural, urban), and geographic region (east, central, west) [17]; (2) behavioral characteristics, including alcohol consumption (never, current, ever) and history of blood donation (yes, no); and (3) family history of hepatitis B virus (HBV) or HCV (yes, no) infection.

HCV Knowledge

Initially, awareness of HCV was determined based on participant responses to the following question: “Have you ever heard of hepatitis C disease or hepatitis C virus (HCV)?” Individuals who replied yes were considered as being aware of HCV and asked additional questions to assess the depth of their HCV knowledge based on surveys conducted previously [18,19]. After meeting the standards for reliability and validity, a brief 9-item scale including questions about the nature of hepatitis C disease and modes of transmission was developed for this study. Correct answers scored 1, and incorrect or “don’t know” answers scored 0. Overall HCV knowledge scores ranged from 0 to 9. Participants were categorized into 3 groups based on their knowledge scores. Those scoring 0-3 points were categorized as having poor knowledge, those scoring 4-6 points were categorized as having fair knowledge, and those scoring 7-9 points were categorized as having good knowledge of HCV. Participants who had never heard of hepatitis C disease or HCV were recognized as being unaware of HCV and categorized as having poor knowledge.

Participant HCV Testing Behavior and Willingness to Undergo HCV Screening

Participant uptake of HCV testing, testing results, reasons for undergoing or not undergoing HCV testing, and willingness to undergo HCV screening were collected through self-reports.

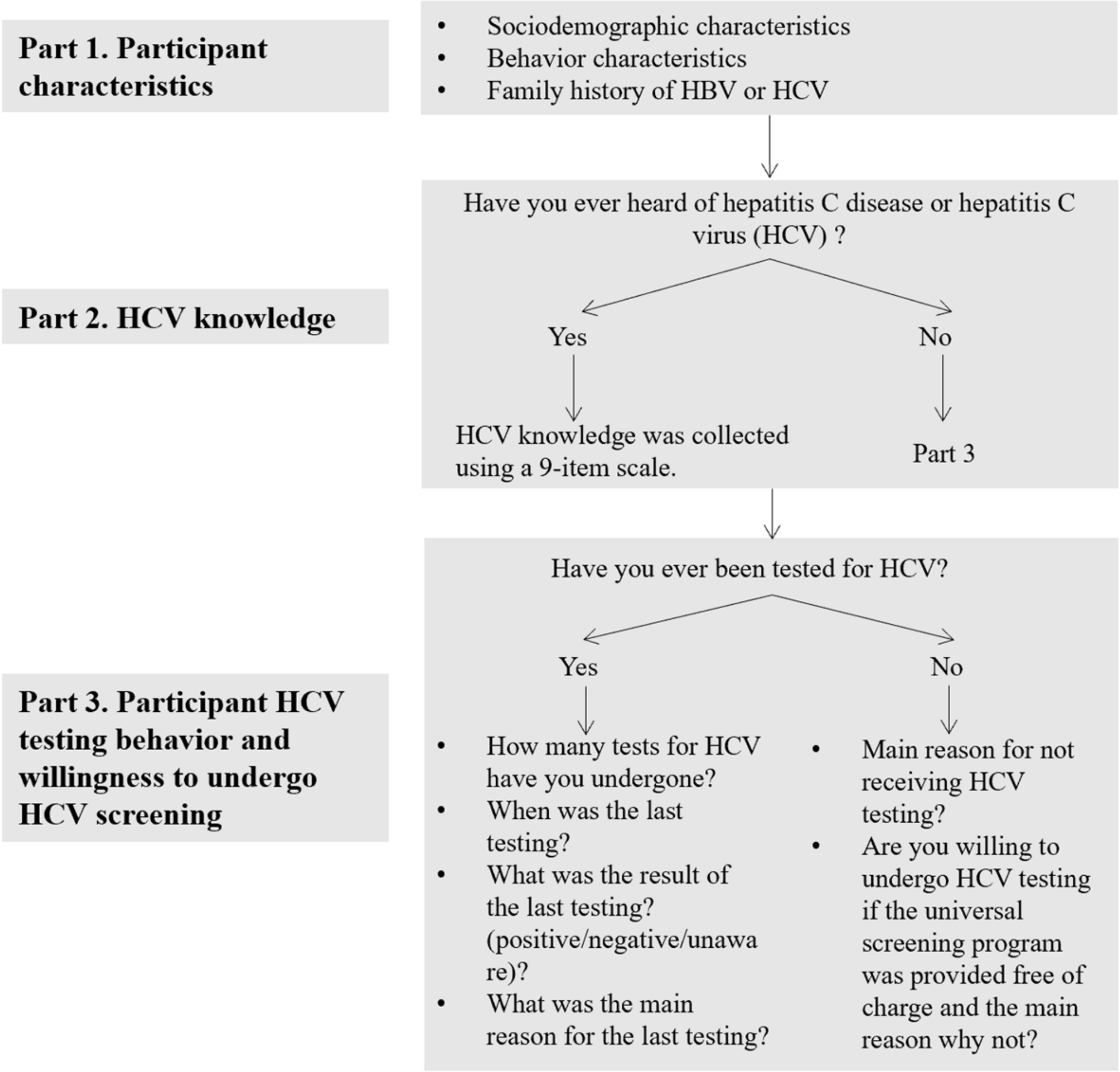
The uptake of HCV testing was determined from responses to the following question: “Have you ever been tested for HCV? (yes/no)”. Participants who reported yes were asked the following questions: “How many tests for HCV have you undergone?,” “When was the last testing?,” “What was the result of the last testing? (positive/negative/unaware),” and “What was the reason for the last testing?” Participants who reported no were further asked about their main reason for not undergoing HCV testing and whether they were willing to undergo HCV testing if universal screening was provided free of charge. Those not willing to undergo screening were asked to provide the main reason why not.

In China, potential blood donors are required to undergo HCV testing since 1993 [20]. All participants in this study with a history of blood donation after 1993 were informed in advance about undergoing HCV testing in order to reduce any self-reporting biases.

Data Collection

WeChat and Weibo were the primary method of recruiting eligible participants. An anonymous, self-administered questionnaire was developed that comprised 3 parts. The structure of the questionnaire is shown in Figure 1. Wenjuanxing, the most popular social media platform in China, was used to disseminate the electronic questionnaire.

Figure 1. The structure of the questionnaire. HCV: hepatitis C virus.



Sample Size

The level of HCV knowledge (poor, fair, good) was used as the main outcome to evaluate the sample size, and the following formula was used to determine the multinomial proportions [21]:

It was assumed that the proportion of participants with poor knowledge (p_1), fair knowledge (p_2), and good knowledge (p_3) was 60%, 15%, and 25%, respectively; the confidence level ($1 - \alpha$) was 95%, and precision (d) was 5%. Thus, the target sample size was calculated as 551 respondents. Considering a nonresponse rate of 10%, it was determined that at least 613 participants should be enrolled in the study.

Statistical Analysis

The data collected in Wenjuanxing were exported into Microsoft Excel and then imported into SAS V9.4 for data analysis. The reliability and validity of the HCV knowledge scale were determined. The Cronbach coefficient was calculated to determine reliability. The average variance extracted (AVE) and composite reliability (CR) were calculated for convergent validity. The square roots of the AVE and the absolute values of correlation coefficients between factors were measured for discriminant validity. Confirmatory factor analysis (CFA) was performed for the goodness of fit.

Mean (SD) values were used for normally distributed continuous variables, and medians, the first quartile (Q1), and the third quartile (Q3) were used for nonnormally distributed continuous variables. The frequency distribution and percentage were used to describe categorical variables.

Factors independently associated with HCV knowledge were assessed using univariate and multivariate ordinal logistic regression analyses, with the HCV knowledge level (poor, fair, good) as the dependent variable and participant characteristics as potential independent variables. The proportional odds assumption was assessed, with $P>.05$ being considered statistically significant. Factors independently associated with the uptake of HCV testing were assessed using univariate and multivariate binary logistic regression analyses, with the uptake of HCV testing (yes, no) as the dependent variable and participant characteristics as potential independent variables. A correlation matrix was used to exclude the confounding variables with strong collinearity. Variables with $P<.1$ in univariate models were used in multivariate regression models. Statistical significance was set as $P<.05$. A crude odds ratio (cOR) with 95% CI was reported for univariate models, and an adjusted odds ratio (aOR) with 95% CI wAs reported for multivariate models. Sensitivity analyses were conducted by limiting the data set to participants without a history of blood donation to control for the impact of blood donation on HCV testing behavior.

Ethical Considerations

Ethical approval was obtained from Henan Cancer Hospital (2023-KY-0139). The electronic questionnaire was written in standard Chinese. The participants received a cover letter informing them of the study aims and procedures and that their participation was anonymous and voluntary. It was assumed that individuals who completed the survey had read the cover letter and consented to participate. To ensure the effectiveness of data collection, participants were required to answer a question to receive the next question. All completed questionnaires were kept private and were accessible only to the study staff responsible for data analysis.

Results

Participant Characteristics

A total of 1929 complete questionnaires were obtained, and after excluding those from health care workers or health care students, 1491 (77.3%) participants from 26 provinces were included in the final analysis. The geographical distribution of the included participants is shown in Figure S1 in Multimedia Appendix 1. Characteristics of the included participants are shown in Table 1. Of the 1491 participants, 476 (31.9%), 364 (24.4%), and 651 (43.7%) were from East, West, and Central China, respectively. In total, the average age of the participants was 41.4 (SD 15.4) years. Slightly over half of the participants were female (n=819, 54.9%) and located in urban areas (n=843, 56.5%). Most participants were married (n=1060, 71.1%), educated for >7 years (n=1160, 77.8%), of Han nationality (n=1405, 94.2%), and employed/self-employed (n=1286, 86.3%). About 26.4% (n=394) of the participants currently drank alcohol, 25.3% (n=377) had ever donated blood, and 12.4% (n=185) had a family history of HBV or HCV.

Table 1. Characteristics of the included participants (N=1491).

Characteristics	Total	East China	West China	Central China
Age (years)				
<60, n (%)	1230 (82.5)	433 (91.0)	310 (85.2)	487 (74.8)
≥60, n (%)	261 (17.5)	43 (9.0)	54 (14.8)	164 (25.2)
Mean (SD)	41.4 (15.4)	38.1 (13.9)	38.8 (15.8)	45.2 (15.4)
Gender, n (%)				
Male	672 (45.1)	204 (42.9)	178 (48.9)	290 (44.5)
Female	819 (54.9)	272 (57.1)	186 (51.1)	361 (55.5)
Marital status, n (%)				
Single/divorced/widowed	431 (28.9)	168 (35.3)	140 (38.5)	123 (18.9)
Married	1060 (71.1)	308 (64.7)	224 (61.5)	528 (81.1)
Years of education, n (%)				
<7	331 (22.2)	81 (17.0)	22 (6.0)	228 (35.0)
7-12	440 (29.5)	103 (21.6)	150 (41.2)	187 (28.7)
>12	720 (48.3)	292 (61.3)	192 (52.7)	236 (36.3)
Ethnicity, n (%)				
Han	1405 (94.2)	448 (94.1)	316 (86.8)	641 (98.5)
Minority	86 (5.8)	28 (5.9)	48 (13.2)	10 (1.5)
Occupation, n (%)				
Students	139 (9.3)	32 (6.7)	72 (19.8)	35 (5.4)
Employed/self-employed	1286 (86.3)	420 (88.2)	280 (76.9)	586 (90.0)
Not working/unemployed	66 (4.4)	24(5.0)	12 (3.3)	30 (4.6)
Residence, n (%)				
Urban	843 (56.5)	340 (71.4)	264 (72.5)	239 (36.7)
Rural	648 (43.5)	136 (28.6)	100 (27.5)	412 (63.3)
Alcohol consumption, n (%)				
Never	961 (64.5)	318 (66.8)	210 (57.7)	433 (66.5)
Current	394 (26.4)	122 (25.6)	106 (29.1)	166 (25.5)
Ever	136 (9.1)	36 (7.6)	48 (13.2)	52 (8.0)
History of blood donation, n (%)				
Yes	377 (25.3)	147 (30.9)	70 (19.2)	160 (24.6)
No	1114 (74.7)	329 (69.1)	294 (80.8)	491 (75.4)
Family history of HBV[a] or HCV[b] infection, n (%)				
No	1306 (87.6)	425 (89.3)	322 (88.5)	559 (85.9)
Yes	185 (12.4)	51 (10.7)	42 (11.5)	92 (14.1)

[a] HBV: hepatitis B virus.

[b] HCV: hepatitis C virus.

Reliability and Validity of the HCV Knowledge Scale

The HCV knowledge scale was found to have good reliability and validity. The total Cronbach coefficient was 0.941(95% CI 0.936-0.945) and the Cronbach coefficients of disease nature and modes of transmission were 0.824 (95% CI 0.798-0.848), and 0.932 (95% CI 0.927-0.937), respectively.

The factor loading of each item of disease nature and modes of transmission was 0.779-0.985 and 0.648-0.967, respectively. The AVE values of the 2 factors were 0.514 and 0.640, respectively, and the CR values were all >0.7, demonstrating strong convergent validity of the questionnaire constructs (Table 2).

Table 2. Factor loading of each item and convergent validity and composite reliability measurement results.

Construct and item	Factor loading
Disease nature (AVE[a]=0.514, CR[b]=0.769)	
HCV[c] infection is usually asymptomatic and must be detected with testing.	0.822
Chronic HCV infection may develop into cirrhosis and liver cancer.	0.985
Hepatitis C can be cured.	0.779
Modes of transmission (AVE=0.640, CR=0.923)	
Can HCV be transmitted via kissing a person with HCV infection?	0.697
Can HCV be transmitted by sharing water and food with people with HCV infection?	0.648
Sexual contact with multiple partners will increase the risk of HCV transmission.	0.707
Can HCV be transmitted during pregnancy or birth in the case of a mother with HCV infection?	0.914
Can HCV be transmitted by using unsterilized syringes, needles, or surgical instruments?	0.967
Can HCV be transmitted by getting a tattoo/piercing?	0.768

[a] AVE: average variance extracted.

[b] CR: composite reliability.

[c] HCV: hepatitis C virus.

The square roots of the AVE values were higher than the absolute values of the correlation coefficients between the disease nature and modes of transmission, indicating that the questionnaire had good discriminant validity (Table 3).

CFA results indicated that the 2-factor model with 9 items had a good fit: root mean square error of approximation (RMSEA)=0.034, comparative fit index (CFI)=0.953, Tucker-Lewis index (TLI)=0.954, and standardized root mean square residual (SRMR)=0.093.

Table 3. Correlation matrix of disease nature and modes of transmission.

Construct	Disease nature	Modes of transmission
Disease nature	0.717[a]	N/A[b]
Modes of transmission	0.675[c]	0.800[a]

[a] Square roots of the average variance extracted (AVE).

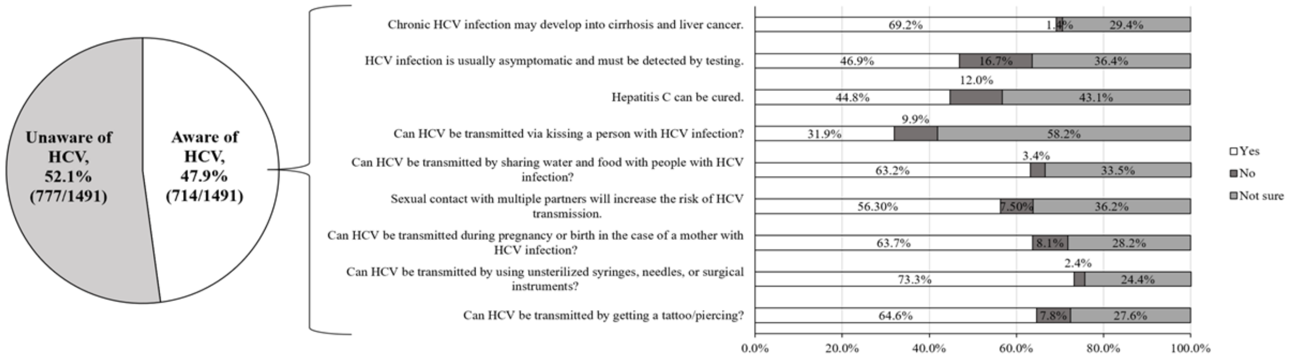
[b] N/A: not applicable.

[c] *P*<.05.

Participant Knowledge of HCV

Less than half (714/1491, 47.9%) of the participants were aware of HCV. Of these, the proportions of participants who correctly answered the items related to disease nature ranged from 44.8% (320/714) to 69.2% (494/714) and of those who correctly answered the items related to modes of transmission ranged from 33.5% (239/714) to 73.3% (523/714). A detailed description of each item is presented in Figure 2.

Figure 2. Awareness of HCV and knowledge regarding hepatitis C disease's nature and modes of transmission. HCV: hepatitis C virus.



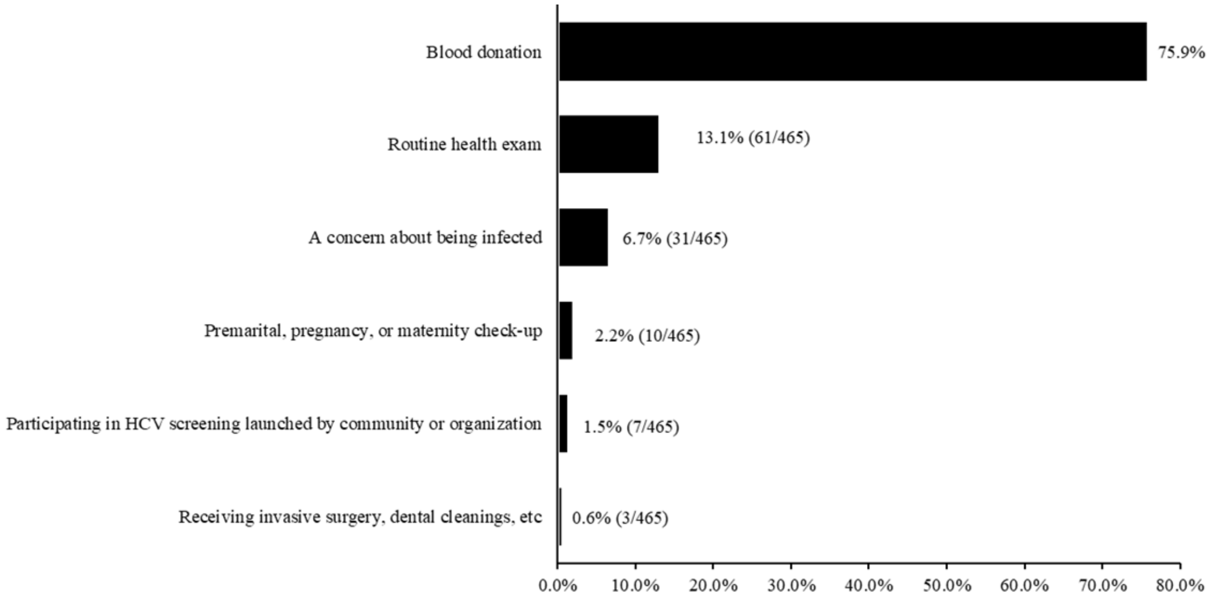
Participants who were unaware of HCV were regarded as having poor knowledge. Overall, the proportion of participants who had poor, fair, and good knowledge was 63.4% (945/1491), 9.3% (139/1491), and 27.3% (407/1491), respectively.

Participant HCV Testing Behavior and Willingness to Undergo HCV Screening

A total of 465 (31.2%) participants had ever undergone HCV testing. Of these, 4 (0.9%) were anti-HCV antibody positive and 288 (61.9%) were unaware of their results. Most participants underwent 1-time testing (419/465, 90.1%) and had undergone their last HCV test at a median of 16.0 (Q1, Q3: 10.0, 63.0) months previously.

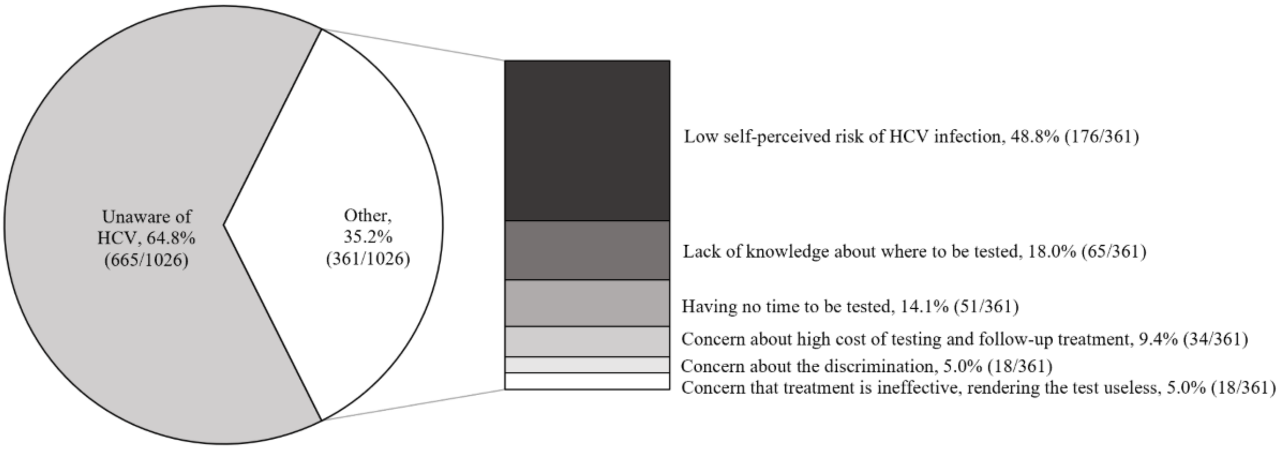
The reasons for the last testing are shown in Figure 3. Most participants (353/465, 75.9%) were tested for HCV following blood donation, 13.1% (61/465) were tested during a routine health exam, and 6.7% (31/465) were tested due to a concern about being infected.

Figure 3. Participants' reasons for undergoing their last HCV test. HCV: hepatitis C virus.



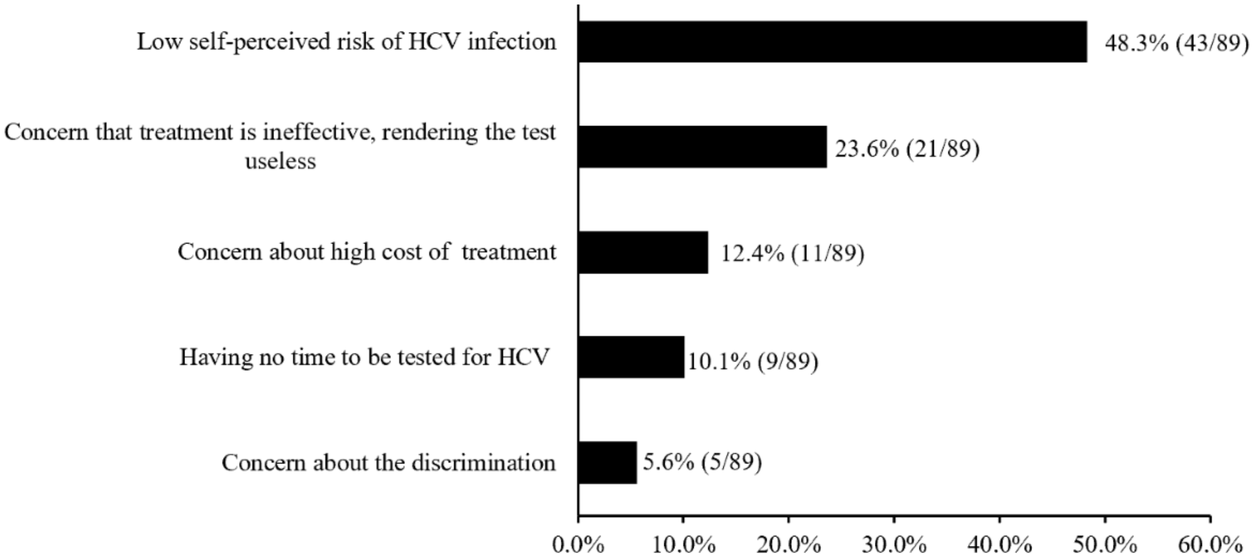
The reasons the participants never underwent an HCV test are listed in Figure 4. The most common barrier against HCV testing was a lack of awareness about the virus (665/1026, 64.8%). Of the participants who were aware of HCV, the main reasons for not undergoing HCV testing included a low self-perceived risk of HCV infection (176/361, 48.8%), a lack of knowledge of where to be tested (65/361, 18.0%), and having no time to be tested (51/361, 14.1%).

Figure 4. Participants' reasons for never undergoing an HCV test. HCV: hepatitis C virus.



Of the participants who never underwent HCV testing, 91.3% (937/1026) were willing to undergo HCV screening if universal screening was provided free of charge. The most common reasons participants were unwilling to undergo testing were a low self-perceived risk of HCV infection (43/89, 48.3%), followed by a concern that treatment is ineffective, rendering the test useless (21/89, 23.6%); see Figure 5.

Figure 5. Participants' reasons for being unwilling to undergo HCV screening. HCV: hepatitis C virus.



Factors Associated With HCV Knowledge

The correlation matrix did not show the existence of high collinearity between variables (Table S1 in Multimedia Appendix 2). Factors associated with HCV knowledge are shown in Tables 4 and 5. In the final multivariate ordinal logistic regression, participants aged ≥60 years (aOR=0.374, 95% CI 0.245-0.572), less educated (<7 years: aOR=0.607, 95% CI 0.401-0.919; 7-12 years: aOR=0.757, 95% CI 0.561-0.953), living in rural areas (aOR=0.625, 95% CI 0.472-0.830), located in West China (aOR=0.452, 95% CI 0.330-0.620), and currently drinking alcohol (aOR=0.680, 95% CI 0.520-0.890) had lower HCV knowledge scores. Participants with a blood donation history (aOR=1.758, 95% CI 1.376-2.245) and a family history of HBV or HCV infection (aOR=2.624, 95% CI 1.815-3.792) were more likely to have higher HCV knowledge scores.

Table 4. Factors associated with HCV[a] knowledge (univariate analysis).

Characteristics	HCV knowledge level			Univariate analysis	
	Poor, n (%)	Fair, n (%)	Good, n (%)	cOR[b] (95%CI)	*P* value[c]
Age (years)					
<60	730 (59.3)	130 (10.6)	370 (30.1)	Reference	N/A[d]
≥60	215 (82.4)	9 (3.4)	37 (14.2)	0.322 (0.231-0.450)	<.001
Gender					
Male	430 (64.0)	62 (9.2)	180 (26.8)	Reference	N/A
Female	515 (62.9)	77 (9.4)	227 (27.7)	1.048 (0.851-1.290)	.655
Marital status					
Single/divorced/widowed	255 (59.2)	45 (10.4)	131 (30.4)	Reference	N/A
Married	690 (65.1)	94 (8.9)	276 (26.0)	0.786 (0.627-0.984)	.036
Years of education					
<7	257 (77.6)	18 (5.4)	56 (16.9)	0.339 (0.253-0.454)	<0.001
7-12	309 (70.2)	26 (5.9)	105 (23.9)	0.508 (0.397-0.650)	<0.001
>12	379 (52.6)	95 (13.2)	246 (34.2)	Reference	N/A
Ethnicity					
Han	894 (63.6)	126 (9.0)	385 (27.4)	Reference	N/A
Minority	51 (59.3)	13 (15.1)	22 (25.6)	1.101 (0.711-1.705)	.665
Occupation					
Students	76 (54.7)	19 (13.7)	44 (31.7)	Reference	N/A
Employed/self-employed	812 (63.1)	117 (9.1)	357 (27.8)	0.747 (0.530-1.053)	.097
Not working/unemployed	57 (86.4)	3 (4.5)	6 (9.1)	0.201 (0.092-0.435)	<.001
Residence					
Urban	472 (56.0)	99 (11.7)	272 (32.3)	Reference	N/A
Rural	473 (73.0)	40 (6.2)	135 (20.8)	0.492 (0.396-0.61)	<.001
Geographic region					
East	258 (54.2)	64 (13.4)	154 (32.4)	Reference	N/A
West	270 (74.2)	24 (6.6)	70 (19.2)	0.438 (0.327-0.585)	<.001
Central	417 (64.1)	51 (7.8)	183 (28.1)	0.714 (0.564-0.903)	.005
Alcohol consumption					
Never	569 (59.2)	110 (11.4)	282 (29.3)	Reference	N/A
Current	290 (73.6)	20 (5.1)	84 (21.3)	0.551 (0.427-0.710)	<.001
Ever	86 (63.2)	9 (6.6)	41 (30.1)	0.903 (0.628-1.296)	.579
History of blood donation					
Yes	181 (48.0)	54 (14.3)	142 (37.7)	2.189 (1.738-2.750)	<.001
No	764 (68.6)	85 (7.6)	265 (23.8)	Reference	N/A
Family history of HBV[e] or HCV infection					
No	837 (64.1)	132 (10.1)	337 (25.8)	Reference	N/A
Yes	108 (58.4)	7 (3.8)	70 (37.8)	1.432 (1.058-1.938)	.020

[a] HCV: hepatitis C virus.

[b] cOR: crude odds ratio.

[c] The proportional odds assumption was satisfied for each variable, with *P*>.05.

[d] N/A: not applicable.

[e] HBV: hepatitis B virus.

Table 5. Factors associated with HCV[a] knowledge (multivariate analysis).

Characteristics	HCV knowledge level			Multivariate analysis	
	Poor, n (%)	Fair, n (%)	Good, n (%)	aOR[b] (95%CI)	*P* value[c]
Age (years)					
<60	730 (59.3)	130 (10.6)	370 (30.1)	Reference	N/A[d]
≥60	215 (82.4)	9 (3.4)	37 (14.2)	0.374 (0.245-0.572)	<.001
Gender					
Male	430 (64.0)	62 (9.2)	180 (26.8)	N/A	N/A
Female	515 (62.9)	77 (9.4)	227 (27.7)	N/A	N/A
Marital status					
Single/divorced/widowed	255 (59.2)	45 (10.4)	131 (30.4)	Reference	N/A
Married	690 (65.1)	94 (8.9)	276 (26.0)	0.999 (0.759-1.316)	.978
Years of education					
<7	257 (77.6)	18 (5.4)	56 (16.9)	0.607 (0.401-0.919)	.007
7-12	309 (70.2)	26 (5.9)	105 (23.9)	0.757 (0.561-0.953)	.022
>12	379 (52.6)	95 (13.2)	246 (34.2)	Reference	N/A
Ethnicity					
Han	894 (63.6)	126 (9.0)	385 (27.4)	N/A	N/A
Minority	51 (59.3)	13 (15.1)	22 (25.6)	N/A	N/A
Occupation					
Students	76 (54.7)	19 (13.7)	44 (31.7)	Reference	N/A
Employed/self-employed	812 (63.1)	117 (9.1)	357 (27.8)	0.828 (0.550-1.246)	.351
Not working/unemployed	57 (86.4)	3 (4.5)	6 (9.1)	0.612 (0.096-1.128)	.111
Residence					
Urban	472 (56.0)	99 (11.7)	272 (32.3)	Reference	N/A
Rural	473 (73.0)	40 (6.2)	135 (20.8)	0.625 (0.472-0.830)	.004
Geographic region					
East	258 (54.2)	64 (13.4)	154 (32.4)	Reference	N/A
West	270 (74.2)	24 (6.6)	70 (19.2)	0.452 (0.330-0.620)	<.001
Central	417 (64.1)	51 (7.8)	183 (28.1)	0.974 (0.744-1.275)	.848
Alcohol consumption					
Never	569 (59.2)	110 (11.4)	282 (29.3)	Reference	N/A
Current	290 (73.6)	20 (5.1)	84 (21.3)	0.680 (0.520-0.890)	.015
Ever	86 (63.2)	9 (6.6)	41 (30.1)	0.844 (0.574-1.243)	.485
History of blood donation					
Yes	181 (48.0)	54 (14.3)	142 (37.7)	1.758 (1.376-2.245)	<.001
No	764 (68.6)	85 (7.6)	265 (23.8)	Reference	N/A
Family history of HBV[e]or HCV infection					
No	837 (64.1)	132 (10.1)	337 (25.8)	Reference	N/A
Yes	108 (58.4)	7 (3.8)	70 (37.8)	2.624 (1.815-3.792)	<.001

[a] HCV: hepatitis C virus.

[b] aOR: adjusted odds ratio.

[c] The proportional odds assumption was satisfied for each variable, with *P*>.05.

[d] N/A: not applicable.

[e]HBV: hepatitis B virus.

Factors Associated With the Uptake of HCV Testing

The correlation matrix did not show the existence of high collinearity between variables (Table S1 in Multimedia Appendix 2). Factors associated with the uptake of HCV testing are presented as Tables 6 and 7. In the final multivariate logistic regression, participants who were female (aOR=0.623, 95% CI 0.478-0.813), were less educated (<7 years: aOR=0.210, 95% CI 0.139-0.316; 7-12 years: aOR=0.693, 95% CI 0.508-0.946), lived in rural areas (aOR=0.690, 95% CI 0.548-0.868), were located in West China (aOR=0.434, 95% CI 0.306-0.617), currently drank alcohol (aOR=0.441, 95% CI 0.318-0.613), and had poor HCV knowledge (aOR=0.353, 95% CI 0.270-0.461) had a lower odds of being previously tested for HCV. Participants with a family history of HBV or HCV infection (aOR=5.126, 95% CI 3.457-7.600) had higher odds of having undergone HCV testing. After limiting participants without a history of blood donation, factors associated with the uptake of HCV testing remained similar, except that a few variables lost their statistical significance due to the reduced sample size (Table S2 in Multimedia Appendix 3).

Table 6. Factors associated with the uptake of HCV[a] testing (univariate analysis).

Characteristics	Uptake of HCV testing		Univariate analysis	
	No, n (%)	Yes, n (%)	cOR[b] (95%CI)	*P* value
Age (years)				
<60	838(68.1)	392(31.9)	Reference	N/A[c]
≥60	188(72.0)	73(28.0)	0.830(0.617–1.115)	.217
Gender				
Male	447 (66.5)	225 (33.5)	Reference	N/A
Female	579 (70.7)	240 (29.3)	0.823 (0.661-1.025)	.083
Marital status				
Single/divorced/widowed	285 (66.1)	146 (33.9)	Reference	N/A
Married	741 (69.9)	319 (30.1)	0.840 (0.661-1.067)	.153
Years of education				
<7	274 (82.8)	57 (17.2)	0.321 (0.232-0.443)	<.001
7-12	315 (71.6)	125 (28.4)	0.613 (0.474-0.791)	<.001
>12	437 (60.7)	283 (39.3)	Reference	N/A
Ethnicity				
Han	969 (69.0)	436 (31.0)	Reference	N/A
Minority	57 (66.3)	29 (33.7)	1.131 (0.71-1.793)	.601
Occupation				
Students	86 (61.9)	53 (38.1)	Reference	N/A
Employed/self-employed	888 (69.1)	398 (30.9)	0.727 (0.506-1.044)	.085
Not working/unemployed	52 (78.8)	14 (21.2)	0.437 (0.221-0.864)	.017
Residence				
Urban	554 (65.7)	289 (34.3)	Reference	N/A
Rural	472 (72.8)	176 (27.2)	0.715 (0.571-0.894)	.003
Geographic region				
East	301 (63.2)	175 (36.8)	Reference	N/A
West	284 (78.0)	80 (22.0)	0.484 (0.355-0.660)	<.001
Central	441 (67.7)	210 (32.3)	0.819 (0.638-1.05)	.115
Alcohol consumption				
Never	626 (65.1)	335 (34.9)	Reference	N/A
Current	320 (81.2)	74 (18.8)	0.432 (0.324-0.574)	<.001
Ever	80 (58.8)	56 (41.2)	1.308 (0.906-1.886)	.151
Family history of HBV[d] or HCV infection				
Yes	937 (71.7)	369 (28.3)	Reference	N/A
No	89 (48.1)	96 (51.9)	2.738 (2.003-3.7)	<.001
HCV knowledge level				
Poor	752 (79.6)	193 (20.4)	0.248 (0.193-0.318)	<.001
Fair	74 (53.2)	65 (46.8)	0.849 (0.577-1.248)	.404
Good	200 (49.1)	207 (50.9)	Reference	N/A

[a] HCV: hepatitis C virus.

[b] cOR: crude odds ratio.

[c] N/A: not applicable.

[d] HBV: hepatitis B virus.

Table 7. Factors associated with the uptake of HCV[a] testing (multivariate analysis).

Characteristics	Uptake of HCV testing		Multivariate analysis	
	No, n (%)	Yes, n (%)	aOR[b] (95%CI)	*P* value
Age (years)				
<60	838(68.1)	392(31.9)	N/A[c]	N/A
≥60	188(72.0)	73(28.0)	N/A	N/A
Gender				
Male	447 (66.5)	225 (33.5)	Reference	N/A
Female	579 (70.7)	240 (29.3)	0.623 (0.478-0.813)	<.001
Marital status				
Single/divorced/widowed	285 (66.1)	146 (33.9)	N/A	N/A
Married	741 (69.9)	319 (30.1)	0.811 (0.635-1.037)	.152
Years of education				
<7	274 (82.8)	57 (17.2)	0.210 (0.139-0.316)	<.001
7-12	315 (71.6)	125 (28.4)	0.693 (0.508-0.946)	.021
>12	437 (60.7)	283 (39.3)	Reference	N/A
Ethnicity				
Han	969 (69.0)	436 (31.0)	N/A	N/A
Minority	57 (66.3)	29 (33.7)	N/A	N/A
Occupation				
Students	86 (61.9)	53 (38.1)	Reference	N/A
Employed/self-employed	888 (69.1)	398 (30.9)	0.831 (0.364-1.896)	.205
Not working/unemployed	52 (78.8)	14 (21.2)	0.726 (0.457-1.154)	.517
Residence				
Urban	554 (65.7)	289 (34.3)	Reference	N/A
Rural	472 (72.8)	176 (27.2)	0.690 (0.548-0.868)	<.001
Geographic region				
East	301 (63.2)	175 (36.8)	Reference	N/A
West	284 (78.0)	80 (22.0)	0.434 (0.306-0.617)	<.001
Central	441 (67.7)	210 (32.3)	0.999 (0.753-1.325)	.996
Alcohol consumption				
Never	626 (65.1)	335 (34.9)	Reference	N/A
Current	320 (81.2)	74 (18.8)	0.441 (0.318-0.613)	<.001
Ever	80 (58.8)	56 (41.2)	1.313 (0.868-1.984)	.197
Family history of HBV[d] or HCV infection				
Yes	937 (71.7)	369 (28.3)	Reference	N/A
No	89 (48.1)	96 (51.9)	5.126 (3.457-7.600)	<.001
HCV knowledge level				
Poor	752 (79.6)	193 (20.4)	0.353 (0.270-0.461)	<.001
Fair	74 (53.2)	65 (46.8)	0.920 (0.611-1.385)	.689
Good	200 (49.1)	207 (50.9)	Reference	N/A

[a] HCV: hepatitis C virus.

[b] aOR: adjusted odds ratio.

[c] N/A: not applicable.

[d] HBV: hepatitis B virus.

Discussion

Principal Findings

HCV is a serious public health problem in both China and worldwide. To the best of our knowledge, this is the first study to assess the HCV knowledge and testing behavior of the general Chinese population. The findings inform the development and improvement of HCV-related health education and screening strategies.

HCV knowledge, even at a basic level, was low in the general Chinese population included in this survey. Less than half of the participants had ever heard of HCV, which is lower than that reported in high-income countries. In South Korea, for example, 56.4% of the general population was aware of HCV [22]. The lack of health education in the general population of China is an important reason for the low rate of disease awareness [23]. Although China issued "Guidelines for the Prevention and Treatment of Hepatitis C" in 2004, this document only emphasizes the need for HCV health education in high-risk groups [13,24]. After evaluating HCV knowledge among participants in this study, several concerns were found to be worth noting. First, participants had particularly low awareness of treatment outcomes, a finding that could delay or prevent treatment initiation. Only 44.8% of survey respondents who were aware of HCV believed that the disease could be cured. This may be because prior to the approval of DAAs, the traditional treatment for HCV was a combined injection of peg-interferon-α and ribavirin, and the sustained virologic response (SVR) to this medication was only about 52% (95% CI 34%-69%) [25-27]. Although the more recent approval of DAAs has meant that >95% of patients can now achieve an SVR [11] , the success of this treatment is not widely known. Second, participants had low knowledge of sexual or household transmission of HCV. For example, of the participants who had ever heard of HCV, >60% were aware that HCV can be transmitted via blood or from mother to child. In contrast, 63.2% believed incorrectly that people could become infected by sharing water or food with a person with HCV infection. This misunderstanding could result in discrimination against individuals with HCV infection, thereby preventing them from undergoing testing and engaging in normal social and work activities. In addition, slightly over half of the participants who were aware of HCV recognized that sexual contact with multiple partners can increase the risk of transmission. Although sexual transmission in monogamous heterosexual relationships is rare, we still included it as a question on HCV knowledge. This is mainly because unsafe sexual contact is an important route of HCV transmission among MSM or people with HIV infection [28,29]. In this study, MSM or HIV infection was not included as a variable, so it was not possible to assess knowledge of the sexual transmission of HCV in this population. In addition, unsafe sexual contact increases the risk of other infectious diseases, including HIV, HBV, and syphilis. Thus, we hope that the negative public attitude toward unsafe sexual contact is beneficial for the prevention of multiple infections, including HCV. These findings indicate that public health education is urgently needed in China, especially regarding HCV treatment outcomes and modes of transmission.

About one-third of the general population (31.2%) included in this study reported ever undergoing HCV testing, which was lower than that reported in high-income countries. In Italy, for example, 43.4% of the general population declared that it has undergone HCV testing in the past [30]. Moreover, this study found that 75.9% of the survey respondents who had undergone HCV testing did so during blood donation. This might be related to the increased focus on the safety of blood products. Since 1993, the Chinese government has had a policy to screen blood donors and has enacted related laws to screen all donors for anti-HCV antibodies; in 2015, China began screening blood donors who were anti-HCV negative for HCV RNA [20]. However, of participants without a history of blood donation, only 7.9% had ever been tested for HCV, most often during a routine health exam. In addition, most of those tested for HCV were unaware of their results, which could lead to delayed diagnosis and treatment. In fact, most anti-HCV antibody–positive patients often get no opportunity to undergo early treatment due to limited options for HCV consultation and referral in China. According to a survey conducted among patients with HCV in 76 hospitals in China, only 34.9% underwent further HCV RNA testing and 12.2% underwent antiviral treatment [31]. Furthermore, a lack of knowledge of where to be tested and having no time to be tested were the important reasons participants who were aware of HCV did not undergo HCV testing. Self-testing for HCV antibodies, where people collect their own specimen, perform a simple rapid test, and interpret the result, has been recommended by WHO as an additional testing approach to increase HCV testing coverage since 2021 [32]. Self-testing offers privacy and convenience. A systematic review and meta-analysis also suggested that HCV self-testing is feasible and well accepted [33]. In China, evaluation of the cost-effectiveness of HCV self-testing is needed to implement and scale-up this program nationwide. Nevertheless, it is critically important to expand HCV knowledge in the general population and strengthen access to testing and standardized treatment in China.

Despite the overall prevalence of HCV antibody–positive individuals in the general Chinese population being reported as only 0.91%, prevalence was shown to increase with advancing age, peaking at 3.95% among people >65 years of age [10]. Thus, older people may be the key population for HCV testing in China. However, this study found that the HCV testing rate is not significantly higher among the elderly, suggesting that targeted interventions are required for the older population. Low awareness of HCV and a low self-perceived risk of HCV infection were the most reported barriers against HCV testing, and the elderly were also shown to have low HCV knowledge. In addition to strengthening publicity and education, birth cohort screening is recommended. WHO suggests that universal screening be implemented in populations with an HCV antibody seroprevalence of >2%-5% and indicates that birth cohort screening of older adults who have high seroprevalence (eg, people born between 1945 and 1965 in the United States) in settings with a low general prevalence is cost-effective [34]. This study found that people more than 60 years old had a high willingness (162/188, 86.2%) to undergo HCV testing if universal screening was provided free of charge. Thus, it may be feasible to perform universal birth cohort screening. However, cost-effectiveness analysis is needed to determine the optimal birth cohort for HCV screening in order to achieve the goal of eliminating HCV in China.

This study assumed that all participants with a history of blood donation were tested for HCV. Under this assumption, we found that those who were less educated, living in rural areas, and located in West China had less HCV knowledge and testing uptake. West China is a less developed region compared to East and Central China [35]. These findings are in line with previous studies showing that individuals with poor education and living in less developed areas have less access to health care innovations [36]. In addition to sociodemographic characteristics, the findings of this study revealed that people who currently drink alcohol have low HCV knowledge and a low likelihood of undergoing testing and those with a history of blood donation and a family history of viral hepatitis have higher knowledge and testing. Although alcohol use is not a risk factor for HCV, it is believed to have a deleterious impact on the course of hepatitis C disease [37], for example, to promote HCV infection persistence, cytotoxicity, and oxidative stress. Thus, compared to nondrinkers, alcohol drinkers may be in greater need for HCV education and testing interventions to prevent and promote the early treatment of HCV. Blood donors tend to be well educated and may receive HCV education in blood donation centers [38]. People with a family history of HBV or HCV may be prompted to seek health care due to a concern about exposure. This study also showed that participants with low HCV knowledge were less likely to initiate HCV testing, suggesting that improving HCV knowledge may increase testing rates. In summary, HCV-related health education should be prioritized for those who are less educated, located in less-developed areas, drink alcohol, and may otherwise have limited health awareness.

It is worth noting that the HCV knowledge of women was not significantly lower than that of men but their testing rate was significantly lower. Further analysis revealed that in this survey, compared to men, women had a higher proportion of low self-perceived risk of HCV infection (men: 50/447, 11.8%; women: 126/579, 22.8%) and a lack of time to be tested (men: 11/447, 2.5%; women: 40/579, 6.9%). However, women are at higher risk of both transmitting and acquiring HCV through sexual and vertical (mother to child during childbirth) transmission [39,40]. Thus, suitable and effective strategies should be implemented for women, including help in accessing health care services.

Limitations

This study has several limitations. First, since WeChat was used to disseminate the questionnaire, only those with access to this social media platform were included in the study. This may have resulted in an insufficient representation of the general population and limited the conclusions. Generally, people without access to this social media platform are more likely to be less educated and located in low-income areas [41]; thus, it can be inferred that HCV knowledge and the testing rate of the general Chinese population would be lower than that of this survey. Second, the data presented in this study were self-reported and thus might be subject to recall and social bias. Third, this was a cross-sectional study; thus, no causal relationship could be concluded from the findings. Finally, only some indicators of HCV knowledge and the uptake of HCV testing were investigated.

Conclusion

This study highlights that the general population of China has limited HCV knowledge and a low HCV testing rate. There is an urgent need to design and distribute education information to enhance HCV knowledge and provide targeted information regarding treatment outcomes and all possible modes of transmission. Increased referral to standardized treatment and scaled-up HCV screening, including birth cohort screening for older adults at higher risk of HCV transmission and infection, is also recommended. In addition, health education and interventions should be prioritized for women, those with less education, those living in less-developed areas, and those who currently drink alcohol.

Acknowledgments

This study was funded by the Science and Technology Department of Henan Province (RKX202102011). The authors thank Fu-gui Liu, Jian-yong Fu, Pan-pan Mao, Xi-jia Tang, and Shan-zei Wang for their assistance with data collection.

Authors' Contributions

YL and SKZ designed and conducted the study. JS, XYW, and HW analyzed the data and drafted the manuscript. All authors have read and agreed to the drafted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The geographical distribution of the included participants.
[PNG File , 220 KB - publichealth_v9i1e39472_app1.png]

Multimedia Appendix 2
Correlation matrix of the included variables.
[PDF File (Adobe PDF File), 101 KB - publichealth_v9i1e39472_app2.pdf]

Multimedia Appendix 3
Factors associated with the uptake of HCV testing in participants without a history of blood donation.
[PDF File (Adobe PDF File), 172 KB - publichealth_v9i1e39472_app3.pdf]

References

1. Hepatitis C. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/hepatitis-c [accessed 2022-05-06]
2. Polaris Observatory HCV Collaborators. Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study. Lancet Gastroenterol Hepatol 2017 Mar;2(3):161-176. [doi: 10.1016/S2468-1253(16)30181-9] [Medline: 28404132]
3. Polaris Observatory Collaborators. HCV. CDA Foundation. URL: https://cdafound.org/dashboard/polaris/dashboard.html [accessed 2022-08-06]
4. Global health sector strategy on viral hepatitis 2016–2021. World Health Organization. 2016 May 17. URL: https://www.who.int/publications/i/item/WHO-HIV-2016.06 [accessed 2023-11-27]
5. Maqsood S, Iqbal S, Zakar R, Zakar MZ, Fischer F. Determinants of overall knowledge and health behaviours in relation to hepatitis B and C among ever-married women in Pakistan: evidence based on Demographic and Health Survey 2017-18. BMC Public Health 2021 Dec 30;21(1):2328 [FREE Full text] [doi: 10.1186/s12889-021-12406-z] [Medline: 34969379]
6. Hser Y, Du J, Li J, Zhao M, Chang YJ, Peng CY, et al. Hepatitis C among methadone maintenance treatment patients in Shanghai and Kunming, China. J Public Health (Oxf) 2012 Mar 02;34(1):24-31 [FREE Full text] [doi: 10.1093/pubmed/fdr098] [Medline: 22138488]
7. Wang R, Cui N, Long M, Mu L, Zeng H. Barriers to uptake of hepatitis C virus (HCV) health intervention among men who have sex with men in Southwest China: a qualitative study. Health Soc Care Community 2020 Jul 15;29(2):445-452. [doi: 10.1111/hsc.13104]
8. Rabaan AA, Al-Ahmed SH, Bazzi AM, Alfouzan WA, Alsuliman SA, Aldrazi FA, et al. Overview of hepatitis C infection, molecular biology, and new treatment. J Infect Public Health 2020 May;13(5):773-783 [FREE Full text] [doi: 10.1016/j.jiph.2019.11.015] [Medline: 31870632]
9. Spearman CW, Dusheiko GM, Hellard M, Sonderup M. Hepatitis C. Lancet 2019 Oct;394(10207):1451-1466. [doi: 10.1016/s0140-6736(19)32320-7]
10. Gao Y, Yang J, Sun F, Zhan S, Fang Z, Liu X, et al. Prevalence of anti-HCV antibody among the general population in Mainland China between 1991 and 2015: a systematic review and meta-analysis. Open Forum Infect Dis 2019 Mar;6(3):ofz040 [FREE Full text] [doi: 10.1093/ofid/ofz040] [Medline: 30863789]
11. Gao Y, Kong F, Li G, Li C, Zheng S, Lin J, et al. Coblopasvir and sofosbuvir for treatment of chronic hepatitis C virus infection in China: a single-arm, open-label, phase 3 trial. Liver Int 2020 Nov 13;40(11):2685-2693 [FREE Full text] [doi: 10.1111/liv.14633] [Medline: 33047868]
12. Li M, Zhuang H, Wei L. How would China achieve WHO's target of eliminating HCV by 2030? Expert Rev Anti Infect Ther 2019 Oct 12;17(10):763-773. [doi: 10.1080/14787210.2019.1675509] [Medline: 31578079]
13. Chinese Society of Hepatology and Chinese Society of Infectious Diseases, Chinese Medical Association. Guideline of prevention and treatment for hepatitis C (2019 version). Chin J Infect Dis 2020;38(1):9-28.
14. Deuffic-Burban S, Huneau A, Verleene A, Brouard C, Pillonel J, Le Strat Y, et al. Assessing the cost-effectiveness of hepatitis C screening strategies in France. J Hepatol 2018 Oct;69(4):785-792 [FREE Full text] [doi: 10.1016/j.jhep.2018.05.027] [Medline: 30227916]
15. Assoumou S, Tasillo A, Leff J, Schackman BR, Drainoni ML, Horsburgh CR, et al. Cost-effectiveness of one-time hepatitis C screening strategies among adolescents and young adults in primary care settings. Clin Infect Dis 2018 Jan 18;66(3):376-384 [FREE Full text] [doi: 10.1093/cid/cix798] [Medline: 29020317]
16. Wong WW, Tu H, Feld JJ, Wong T, Krahn M. Cost-effectiveness of screening for hepatitis C in Canada. CMAJ 2015 Feb 17;187(3):E110-E121 [FREE Full text] [doi: 10.1503/cmaj.140711] [Medline: 25583667]
17. Girardin F, Hearmon N, Castro E, Negro F, Eddowes L, Gétaz L, et al. Modelling the impact and cost-effectiveness of extended hepatitis C virus screening and treatment with direct-acting antivirals in a Swiss custodial setting. Clin Infect Dis 2019 Nov 13;69(11):1980-1986. [doi: 10.1093/cid/ciz088] [Medline: 30715266]

18. Balfour L, Kowal J, Corace KM, Tasca GA, Krysanski V, Cooper CL, et al. Increasing public awareness about hepatitis C: development and validation of the brief hepatitis C knowledge scale. Scand J Caring Sci 2009 Dec 10;23(4):801-808. [doi: 10.1111/j.1471-6712.2008.00668.x] [Medline: 19500309]
19. Tronina O, Gotlib J, Małkowski P, Jaworski M, Panczyk M. Translation and validation study of the Polish version of the Brief Hepatitis C Knowledge Scale. PLoS One 2020 Jul 9;15(7):e0235764 [FREE Full text] [doi: 10.1371/journal.pone.0235764] [Medline: 32645071]
20. Mei X, Lu H. Prevalence, diagnosis, and treatment of hepatitis C in Mainland China. Glob Health Med 2021 Oct 31;3(5):270-275 [FREE Full text] [doi: 10.35772/ghm.2021.01080] [Medline: 34782868]
21. Adcock CJ. Sample size determination: a review. J Royal Statistical Soc D 1997 Jul;46(2):261-283. [doi: 10.1111/1467-9884.00082]
22. Choi GH, Jang ES, Kim J, Jeong S. A survey of the knowledge of and testing rate for hepatitis C in the general population in South Korea. Gut Liver 2020 Nov 15;14(6):808-816 [FREE Full text] [doi: 10.5009/gnl19296] [Medline: 32066209]
23. Li J, Pang L, Liu Z. Interpretation of the National Action Plan for eliminating hepatitis C as a public health threat (2021-2030). China CDC Wkly 2022 Jul 22;4(29):627-630 [FREE Full text] [doi: 10.46234/ccdcw2022.139] [Medline: 35919827]
24. Chinese Society of Hepatology and Chinese Society of Infectious Diseases, Chinese Medical Association. Guideline of prevention and treatment for hepatitis C (2004 version). Chin J Hepatol 2004;12(4):194-198.
25. Zanini B, Covolo L, Donato F, Lanzini A. Effectiveness and tolerability of combination treatment of chronic hepatitis C in illicit drug users: meta-analysis of prospective studies. Clin Ther 2010 Dec;32(13):2139-2159. [doi: 10.1016/s0149-2918(11)00021-x]
26. Hilsden R, Macphail G, Grebely J, Conway B, Lee S. Directly observed pegylated interferon plus self-administered ribavirin for the treatment of hepatitis C virus infection in people actively using drugs: a randomized controlled trial. Clin Infect Dis 2013 Aug;57(Suppl 2):S90-S96. [doi: 10.1093/cid/cit327] [Medline: 23884072]
27. Dimova R, Zeremski M, Jacobson I, Hagan H, Des Jarlais DC, Talal A. Determinants of hepatitis C virus treatment completion and efficacy in drug users assessed by meta-analysis. Clin Infect Dis 2013 Mar;56(6):806-816 [FREE Full text] [doi: 10.1093/cid/cis1007] [Medline: 23223596]
28. Xie J, Han Y, Qiu Z, Li Y, Li Y, Song X, et al. Prevalence of hepatitis B and C viruses in HIV‐positive patients in China: a cross‐sectional study. J Int AIDS Soc 2016 Mar 14;19(1):20659. [doi: 10.7448/ias.19.1.20659]
29. Irvin R, Gamble T, Malone J, Wang Z, Wilson E, Hughes JP, et al. HIV high prevalence of hepatitis C virus antibodies among urban US men who have sex with men, independent of human immunodeficiency virus status. Clin Infect Dis 2021 Oct 05;73(7):e2205-e2210 [FREE Full text] [doi: 10.1093/cid/ciaa1869] [Medline: 33346798]
30. Lo Moro G, Scaioli G, Vola L, Guastavigna L, Frattin R, De Vito E, et al. Exploring knowledge and awareness of HCV infection and screening test: a cross-sectional survey among an Italian sample. J Community Health 2023 Oct 28;48(5):769-783 [FREE Full text] [doi: 10.1007/s10900-023-01218-4] [Medline: 37115378]
31. Feng XD, Yu H, Pan L, Gao Y, Tang Y, Mao Y, et al. Current status of diagnosistreatment of hepatitis C in medical institutions. Chin J AIDS STD 2020;2020(26):247-249.
32. Recommendations and guidance on hepatitis C virus self-testing. World Health Organization. URL: https://apps.who.int/iris/handle/10665/342803 [accessed 2023-08-01]
33. Perazzo H, Castro R, Villela-Nogueira C, Torres M, Silva SL, Cardoso SW, et al. Acceptability and usability of oral fluid HCV self-testing for hepatitis C diagnosis: a systematic review and meta-analysis. J Viral Hepat 2023 Nov 23;30(11):838-847. [doi: 10.1111/jvh.13876] [Medline: 37485619]
34. World Health Organization. Guidelines on Hepatitis B and C Testing. Geneva: World Health Organization; 2017.
35. Shi Y, Feng C, Yu Q, Guo L. Integrating supply and demand factors for estimating ecosystem services scarcity value and its response to urbanization in typical mountainous and hilly regions of south China. Sci Total Environ 2021 Nov 20;796:149032 [FREE Full text] [doi: 10.1016/j.scitotenv.2021.149032] [Medline: 34280627]
36. Zhang X, Yu P, Yan J, Ton A M Spil I. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv Res 2015 Feb 21;15(1):71 [FREE Full text] [doi: 10.1186/s12913-015-0726-2] [Medline: 25885110]
37. Siu L, Foont J, Wands J. Hepatitis C virus and alcohol. Semin Liver Dis 2009 May 22;29(2):188-199 [FREE Full text] [doi: 10.1055/s-0029-1214374] [Medline: 19387918]
38. Hu W, Meng H, Hu Q, Feng L, Qu X. Blood donation from 2006 to 2015 in Zhejiang Province, China: annual consecutive cross-sectional studies. BMJ Open 2019 May 19;9(5):e023514 [FREE Full text] [doi: 10.1136/bmjopen-2018-023514] [Medline: 31110081]
39. Yelemkoure ET, Yonli AT, Montesano C, Ouattara AK, Diarra B, Zohoncon TM, et al. Prevention of mother-to-child transmission of hepatitis B virus in Burkina Faso: screening, vaccination and evaluation of post-vaccination antibodies against hepatitis B surface antigen in newborns. J Public Health Afr 2018 Dec 21;9(3):816 [FREE Full text] [doi: 10.4081/jphia.2018.816] [Medline: 30687485]
40. Nankya-Mutyoba J, Aizire J, Makumbi F, Ocama P, Kirk GD. Hepatitis B virus perceptions and health seeking behaviors among pregnant women in Uganda: implications for prevention and policy. BMC Health Serv Res 2019 Oct 26;19(1):760 [FREE Full text] [doi: 10.1186/s12913-019-4516-0] [Medline: 31655575]

41. Anikeeva O, Bywood P. Social media in primary health care: opportunities to enhance education, communication and collaboration among professionals in rural and remote locations: did you know? Practical practice pointers. Aust J Rural Health 2013 Apr 15;21(2):132-134. [doi: 10.1111/ajr.12020] [Medline: 23586577]

Abbreviations

aOR: adjusted odds ratio
AVE: average variance extracted
CFA: confirmatory factor analysis
cOR: crude odds ratio
CR: composite reliability
DAA: direct acting antiviral
HBV: hepatitis B virus
HCV: hepatitis C virus
MSM: men who have sex with men
SVR: sustained virologic response
WHO: World Health Organization

Edited by A Mavragani, S Travis; submitted 11.05.22; peer-reviewed by S Facente, Y Qiu, A Boyd; comments to author 22.09.22; revised version received 08.08.23; accepted 09.11.23; published 11.12.23.

Please cite as:
Liu Y, Su J, Wang X, Xu H, Wang H, Kang R, Zheng L, Wang Y, Liu C, Jing Y, Zhang S
Hepatitis C Knowledge and Self-Reported Testing Behavior in the General Population in China: Online Cross-Sectional Survey
JMIR Public Health Surveill 2023;9:e39472
URL: https://publichealth.jmir.org/2023/1/e39472
doi:10.2196/39472
PMID:38079213

Original Paper

Association of Uncontrolled Hypertension or Diabetes Mellitus With Major Adverse Cardiovascular Events and Mortality in South Korea: Population-Based Cohort Study

Sung-Hee Oh[1], MPH, PhD; Dohyang Kim[2], MS; Jinseub Hwang[2], PhD; Jae-Heon Kang[3], MD, PhD; Yeongkeun Kwon[4], MD, PhD; Jin-Won Kwon[1], MPH, PhD

[1]BK21 FOUR Community-Based Intelligent Novel Drug Discovery Education Unit, College of Pharmacy and Research Institute of Pharmaceutical Sciences, Kyungpook National University, Daegu, Republic of Korea

[2]Department of Statistics, Daegu University, Gyeongbuk, Republic of Korea

[3]Department of Family Medicine, Kangbuk Samsung Hospital, College of Medicine, Sungkyunkwan University, Seoul, Republic of Korea

[4]Centre for Obesity and Metabolic Diseases, Korea University Anam Hospital, Seoul, Republic of Korea

Corresponding Author:
Jin-Won Kwon, MPH, PhD
BK21 FOUR Community-Based Intelligent Novel Drug Discovery Education Unit
College of Pharmacy and Research Institute of Pharmaceutical Sciences
Kyungpook National University
80, Daehak-ro, Buk-gu
Daegu, 41566
Republic of Korea
Phone: 82 53 950 8580
Fax: 82 53 950 8557
Email: jwkwon@knu.ac.kr

Abstract

Background: Managing hypertension (HT) and diabetes mellitus (DM) is crucial to preventing cardiovascular diseases. Few studies have investigated the incidence and risk of cardiovascular diseases or mortality in uncontrolled HT or DM in the Asian population. Epidemiological studies of cardiovascular disease should be conducted with continuous consideration of the changing disease risk profiles, lifestyles, and socioeconomic status over time.

Objective: We aimed to examine the association of uncontrolled HT or DM with the incidence of cardiovascular events or deaths from any cause.

Methods: This population-based retrospective study was conducted using data from the Korean National Health Insurance Service–National Health Screening Cohort, including patients aged 40-79 years who participated in national screening from 2002 to 2003 and were followed up until 2015. The health screening period from 2002 to 2013 was stratified into 6 index periods in 2-year cycles, and the follow-up period from 2004 to 2015 was stratified accordingly into 6 subsequent 2-year periods. The incidence rates and hazard ratio (HR) for major adverse cardiovascular events (MACE) and death from any cause were estimated according to HT or DM control status. Extended Cox models with time-dependent variables updated every 2 years, including sociodemographic characteristics, blood pressure (BP), fasting blood glucose (FBG), medication prescription, and adherence, were used.

Results: Among the total cohort of 440,249 patients, 155,765 (35.38%) were in the uncontrolled HT or DM group. More than 60% of the patients with HT or DM who were prescribed medications did not achieve the target BP or FBG. The incidence of MACE was 10.8-15.5 and 9.6-13.3 per 1000 person-years in the uncontrolled DM and uncontrolled HT groups, respectively, and increased with age. In the uncontrolled HT and DM group, the incidence of MACE was high (15.2-17.5 per 1000 person-years) at a relatively young age and showed no age-related trend. Adjusted HR for MACE were 1.28 (95% CI 1.23-1.32) for the uncontrolled DM group, 1.32 (95% CI 1.29-1.35) for the uncontrolled HT group, and 1.54 (95% CI 1.47-1.60) for the uncontrolled HT and DM group. Adjusted HR for death from any cause were 1.05 (95% CI 1.01-1.10) for the uncontrolled DM group, 1.13 (95% CI 1.10-1.16) for the uncontrolled HT group, and 1.17 (95% CI 1.12-1.23) for the uncontrolled HT and DM group.

Conclusions: This up-to-date evidence of cardiovascular epidemiology in South Korea serves as the basis for planning public health policies to prevent cardiovascular diseases. The high uncontrolled rates of HT or DM, regardless of medication prescription, have led us to suggest the need for a novel system for effective BP or glycemic control, such as a community-wide management program using mobile health technology.

(JMIR Public Health Surveill 2023;9:e42190) doi:10.2196/42190

KEYWORDS

prevention and control; cardiovascular diseases; diabetes mellitus; hypertension; extended Cox regression; cohort study

Introduction

Cardiovascular diseases impose enormous health and economic burden globally [1]. The World Health Organization reported that approximately 17.9 million people died of cardiovascular diseases in 2019, accounting for 32% of deaths worldwide [2]. In South Korea, 17.8% of the total deaths reported in 2019 were caused by heart diseases (10.5%) and cerebrovascular diseases (7.3%) [3]. The substantial burden of cardiovascular diseases in the recent years is attributable to the increase in the number of patients living with cardiovascular diseases, prolonged duration of diseases, and accompanying disabilities [4,5]. The global burden of cardiovascular diseases, evaluated as disability-adjusted life years, has increased by 46% for ischemic heart disease (IHD) and by 32% for stroke from 1990 to 2019, with IHD ranking first and stroke ranking second in the global burden of disease among adults aged 50 to 74 years in 2019 [6].

To prevent cardiovascular diseases, it is important to manage hypertension (HT) and diabetes mellitus (DM), known as the most common contributing risk factors [7]. However, approximately half of the adults with high blood pressure (BP) or blood glucose level remain undiagnosed [8,9]. Furthermore, the target BP and blood glucose levels may not be achieved even with treatment owing to nonadherence to therapy or unhealthy lifestyles [10]. The uncontrolled rate among people with HT in 2019 was reported to be approximately 80% globally, and approximately 50% of the patients undergoing treatment did not achieve BP control [8]. In addition, a study in the United States found that most patients with diabetes were treated but the uncontrolled rate was approximately 50% [11].

A few studies have reported the incidence and risk of complications according to the control status of patients diagnosed with HT or DM. In a recent study, patients with uncontrolled DM had 68% higher odds of developing micro- or macroangiopathy than those with controlled DM [12]. Furthermore, individuals with undiagnosed DM have poor long-term glycemic control and can die within a short time after the onset of complication-related symptoms [13]. The incidence of cardiovascular and cerebrovascular complications was higher in patients with refractory HT on treatment than in those with controlled HT [14]. The data on recent population-based epidemiological studies conducted in the Asian population are relatively scarce [15]. Uncontrolled HT or DM may increase the incidence of cardiovascular events or death regardless of the medication prescription or diagnosis of HT or DM.

The incidence, mortality, and risk of cardiovascular disease vary substantially between populations and over time [16], and it is important to investigate the epidemiology of cardiovascular disease by continuously considering the changing disease risk profiles, lifestyles, and socioeconomic status over time [6]. Therefore, the aim of this nationwide study was to determine the incidence of cardiovascular events according to HT or DM control status in South Korea and to examine the association between disease control and cardiovascular events or death from any cause by adjusting for time-varying characteristics.

Methods

Data Sources and Study Population

This retrospective population-based cohort study was based on data from the National Health Insurance Service–National Health Screening Cohort (NHIS-HEALS) in South Korea. The NHIS-HEALS sample data included 514,866 individuals (aged 40-79 years), who comprised a random selection of 10% of all national health screening program participants covered under the mandatory National Health Insurance or Medical Aid system in 2002 and 2003. The cohort was followed up to 2015 annually for information on death and health care use and biennially for the health screening information. The NHIS-HEALS data provided the following information: demographic variables, income, and date of death from the eligibility database; variables for specific health problems and risk factors evaluated using a self-administered questionnaire such as smoking status, frequency per week of alcohol consumption, days per week of physical activity, medical history and family history, and bioclinical laboratory results from the national health screening database; records of inpatient and outpatient use, including procedures, prescribed drugs, treatments, and diagnostic information using the International Classification of Diseases, 10th revision code from the health care use database based on data collected during the process of claiming health care services; and the types of health care institutions and health care human resources from the health care provider database. These 4 databases were linked using unique personal identification numbers created for the NHIS-HEALS [17]. Among the 514,866 individuals, 489,276 (95.03%) with at least 1 claim of health care services use between 2002 and 2003 were included in the study to define the baseline clinical information. Study participants with cardiovascular disease or all-cause death from 2002 to 2003 were excluded (n=48,217), and 810 participants with missing data on BP and fasting blood glucose (FBG) from 2002 to 2003 were excluded. After exclusion, a total of 440,249 participants were included in the study (Figure S1 in Multimedia Appendix 1).

Ethical Consideration

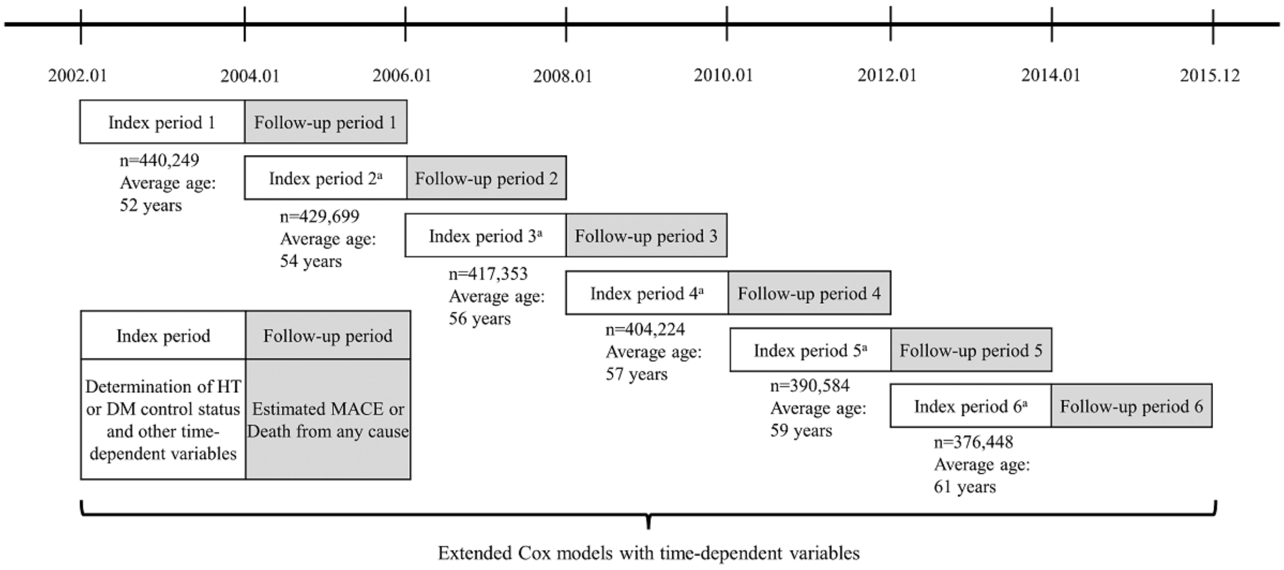
This study was exempted from the Institutional Review Board of Kyungpook National University (number 2020-0157) because the data used in this study were anonymized and deidentified.

Informed consent was not required owing to the retrospective nature of the study.

Study Design and Measurements

The health screening period from 2002 to 2013 was stratified into 6 index periods in 2-year cycles, and the bioclinical laboratory results and various confounding variables were determined during the index period. The follow-up period from 2004 to 2015 was stratified accordingly into six 2-year periods. The subsequent 2-year periods after each index period were defined as follow-up periods 1 to 6. Cardiovascular disease and death from any cause were estimated for each group based on the control status of HT or DM during the 6 follow-up periods (Figure 1). The bioclinical laboratory tests (BP, FBG, BMI, etc) were conducted on participants after an overnight fast at the biennial health screening examination. BP was measured in the seated participants for approximately 5 minutes, and HT was defined as systolic BP ≥140 mm Hg and diastolic BP ≥90 mm Hg. DM was defined as an FBG level ≥126 mg/dL. Well-controlled status was defined as FBG level <126 mg/dL for DM and systolic BP <140 mm Hg and diastolic BP <90 mm Hg for HT. As confounding variables associated with an increased risk of cardiovascular diseases, this study considered age; sex; type of national health security program; income level; BMI; smoking; drinking; exercise; Charlson Comorbidity Index (CCI); disability severity; number of outpatient visits and hospitalization days; use of medication; and medication possession ratio (MPR) of antihypertensive, antidiabetic, and antihyperlipidemic medications [4]. All drugs listed in the national formulary in South Korea as antidiabetic, antihypertensive, and antihyperlipidemic medications were included. Medication users were defined as those who were prescribed medication at least once. Age was classified as middle age (40-64 years) and older age (65-69 years, 70-79 years, and ≥80 years). The income levels were divided into grades 0 to 2 (lowest), 3 to 5, 6 to 8, and 9 to 10 (highest). A person with a grade of 0 covered under the medical aid system was defined as having an income below the minimum cost of living. Study participants were categorized according to BMI based on the World Health Organization criteria for the Asian population: <18.5 kg/m^2 (underweight), 18.5 to 24.9 kg/m^2 (normal to overweight), 25.0 to 29.9 kg/m^2 (class 1 obese), or ≥30 kg/m^2 (class 2 obese) [18]. Disability severity was defined as severe (grades 1 to 3), mild (grades 4 to 6), and no disability, based on the disability grade system under the Welfare of Disabled Persons Act. MPR was defined as the sum of the days of supply from the first to the last prescription divided by the time between the last prescription date plus the days of supply and the first prescription date [19].

Figure 1. Schematic description of the study design. The cohort of national health screening program participants from 2002 to 2003 was followed up to 2015. The health screening period of a 2-year cycle from 2002 to 2013 was defined as 6 index periods, and the follow-up period from 2004 to 2015 was stratified according to 6 subsequent 2-year periods after each index period. [a]Participants with MACE or death from any cause during the previous follow-up period were excluded from the participants for the subsequent index period. From index periods 1 to 6, the number of participants and the average age were presented. DM: diabetes mellitus; HT: hypertension; MACE: major adverse cardiovascular events.



Outcomes

Study outcomes included (1) any major adverse cardiovascular events (MACE), that is, hospitalization for cardiovascular disease defined as a composite of IHD (myocardial infarction and angina), cerebrovascular diseases (cerebral infarction, hemorrhagic stroke, and transient ischemic attack), and heart failure; (2) death from any cause; and (3) MACE or death from any cause [20]. We defined MACE using the primary and secondary International Classification of Diseases, 10th revision diagnostic codes listed in Table S1 in Multimedia Appendix 1. All study participants underwent follow-up examinations until the occurrence of either of the following 2 outcomes: incidence of MACE or death from any cause or the end of follow-up in 2015, whichever occurred first.

Statistical Analysis

The characteristics of the 2 groups were compared using the chi-square test for categorical variables, indicated as count with percentage, and 2-tailed *t* test for continuous variables, indicated as mean with SD.

Incidence rates over each period were calculated as the number of events per 1000 person-years of observation for each outcome; the rates were standardized to the age distribution from each period. Numerators were the number of first events in a specified follow-up period, and denominators were the number of persons at risk during the corresponding index period. Data on individuals who died were censored at the time of death. Extended Cox models with time-dependent variables were used to estimate the hazard ratio (HR) for the association of HT or DM control with MACE or death from any cause. The estimated risk factor measurements, such as BP, FBG, age, type of national health security program, income level, BMI, smoking, drinking, exercise, CCI, disability severity, number of outpatient visits and hospitalization days, and MPR, were used as time-dependent variables. These variables were updated every 2 years. If participants had a missing value in an index period, it was replaced by that participant's previously observed value using the last observation carried forward imputation method. Statistical analyses were conducted using SAS software (version 9.4; SAS Institute Inc) at a significance level of .05.

Results

Characteristics

Among the total cohort of 440,249 individuals, 284,484 (64.61%) individuals were in the well-controlled group and 155,765 (35.38%) individuals were in the uncontrolled group. The average age at entry was 52 (SD 9.4) years for all participants, which increased from 52 to 61 years during index periods 1 to 6 (Figure 1). The characteristics of the study population during the first index period are presented in Table 1. There were substantial differences between the well-controlled and uncontrolled groups in age, sex, income level, national health security program, BMI, smoking, drinking frequency, exercise frequency, CCI, disability severity, number of outpatient visits and hospitalization days, and medication use ($P<.001$). The uncontrolled group had a higher proportion of men, obesity, current smokers, and patients with CCI ≥3; higher drinking frequency; and lower income levels than those in the well-controlled group. Only 23.24% (36,192/155,765) of all uncontrolled patients were prescribed medication; the proportion of patients with antihypertensive medication use was 18.89% (23,169/122,646) in the uncontrolled HT group, and the proportion of patients with antidiabetic medication use was 28.41% (5126/18,042) in the uncontrolled DM group. Among the uncontrolled patients, despite being prescribed medications, 34.92% (10,380/29,523) of patients who had been prescribed antihypertensive medications and 39.12% (4784/12,229) of patients who had been prescribed antidiabetic medications had an MPR <70%. Meanwhile, >60% of medication users had uncontrolled HT or DM: 64.05% (27,318/42,652) of patients who had been prescribed antihypertensive medications had uncontrolled HT and 60.85% (9388/15,429) of patients who had been prescribed antidiabetic medications had uncontrolled DM.

Table 1. Characteristics by the control status of hypertension (HT) or diabetes mellitus (DM) of study participants in the first index period in 2002 to 2003.

Characteristics	Well-controlled group (n=284,484)	Uncontrolled group			
		HT (n=122,646)	DM (n=18,042)	HT and DM (n=15,077)	Total (n=155,765)[a]
Age (years), n (%)					
40 to <65	257,700 (90.59)	99,589 (81.2)	14,920 (82.70)	11,642 (77.22)	126,151 (80.99)
65 to <70	16,147 (5.68)	12,985 (10.59)	1817 (10.07)	1877 (12.45)	16,679 (10.71)
70 to <80	10,637 (3.74)	10,072 (8.21)	1305 (7.23)	1558 (10.33)	12,935 (8.3)
Sex, n (%)					
Male	142,312 (50.02)	71,289 (58.13)	11,246 (62.33)	9793 (64.95)	92,328 (59.27)
Female	142,172 (49.98)	51,357 (41.87)	6796 (37.67)	5284 (35.05)	63,437 (40.73)
Income level, n (%)					
Lowest	42,429 (14.91)	21,226 (17.31)	3034 (16.82)	2823 (18.72)	27,083 (17.39)
Low-middle	60,505 (21.27)	28,554 (23.28)	4129 (22.89)	3571 (23.69)	36,254 (23.27)
Middle-high	81,798 (28.75)	35,662 (29.08)	5318 (29.48)	4469 (29.64)	45,449 (29.18)
Highest	99,752 (35.06)	37,204 (30.33)	5561 (30.82)	4214 (27.95)	46,979 (30.16)
National health security program, n (%)					
National Health Insurance	284,284 (99.93)	122,517 (99.89)	18,012 (99.83)	15,062 (99.9)	155,591 (99.89)
Medical Aid	200 (0.07)	129 (0.11)	30 (0.17)	15 (0.1)	174 (0.11)
BMI (kg/m^2), n (%)					
<18.5	7657 (2.69)	2073 (1.69)	401 (2.22)	214 (1.42)	2688 (1.73)
≥18.5 to <25	193,630 (68.06)	66,960 (54.6)	10,724 (59.44)	7447 (49.39)	85,131 (54.65)
≥25 to <30	78,138 (27.47)	48,352 (39.42)	6319 (35.02)	6507 (43.16)	61,178 (39.28)
≥30	5059 (1.78)	5261 (4.29)	598 (3.31)	909 (6.03)	6768 (4.35)
Smoking, n (%)					
Nonsmoker	188,771 (66.36)	78,327 (63.86)	10,489 (58.14)	8952 (59.38)	97,768 (62.77)
Ex-smoker	22,882 (8.04)	11,157 (9.1)	1607 (8.91)	1367 (9.07)	14,131 (9.07)
Current smoker	61,608 (21.66)	28,019 (22.85)	5220 (28.93)	4149 (27.52)	37,388 (24)
Drinking frequency, n (%)					
Never	163,840 (57.59)	63,675 (51.92)	9844 (54.56)	7381 (48.96)	80,900 (51.94)
2-3 per month	44,889 (15.78)	17,217 (14.04)	2505 (13.88)	1888 (12.52)	21,610 (13.87)
1-4 per week	60,578 (21.29)	32,324 (26.36)	4379 (24.27)	4389 (29.11)	41,092 (26.38)
Every day	9929 (3.49)	7219 (5.89)	941 (5.22)	1134 (7.52)	9294 (5.97)
Exercise frequency, n (%)					
Never	158,372 (55.67)	68,534 (55.88)	9959 (55.2)	8297 (55.03)	86,790 (55.72)
1-4 times per week	93,821 (32.98)	38,416 (31.32)	5711 (31.65)	4643 (30.8)	48,770 (31.31)
5-7 times per week	24,834 (8.73)	12,048 (9.82)	1934 (10.72)	1778 (11.79)	15,760 (10.12)
Charlson Comorbidity Index scores, n (%)					
0	146,346 (51.44)	61,699 (50.31)	6616 (36.67)	5931 (39.34)	74,246 (47.67)
1	86,439 (30.38)	36,653 (29.89)	4919 (27.26)	4092 (27.14)	45,664 (29.32)
2	34,140 (12)	15,385 (12.54)	2878 (15.95)	2360 (15.65)	20,623 (13.24)
≥3	17,559 (6.17)	8909 (7.26)	3629 (20.11)	2694 (17.87)	15,232 (9.78)
Disability severity, n (%)					

Characteristics	Well-controlled group (n=284,484)	Uncontrolled group			
		HT (n=122,646)	DM (n=18,042)	HT and DM (n=15,077)	Total (n=155,765)[a]
No disability	283,347 (99.6)	121,870 (99.37)	17,914 (99.29)	14,952 (99.17)	154,736 (99.34)
Mild	728 (0.26)	499 (0.41)	66 (0.37)	65 (0.43)	630 (0.4)
Severe	409 (0.14)	277 (0.23)	62 (0.34)	60 (0.40)	399 (0.26)
Number of outpatient visits and hospitalization days, mean (SD)	14.2 (12.6)	16.6 (14.13)	20.0 (15.65)	20.4 (15.7)	17.3 (14.5)
Use of medication, n (%)					
Yes	15,245 (5.36)	24,226 (19.75)	5939 (32.92)	6027 (39.97)	36,192 (23.24)
Antidiabetic medication	3200 (1.12)	2841 (2.32)	5126 (28.41)	4262 (28.27)	12,229 (7.85)
Antihypertensive medication	13,129 (4.62)	23,169 (18.89)	2205 (12.22)	4149 (27.52)	29,523 (18.95)
Antidiabetic and antihypertensive medication	1084 (0.38)	1784 (1.45)	1392 (7.72)	2384 (15.81)	5560 (3.57)
Use of antidiabetic medication, n (%)					
Yes	3200 (1.12)	2841 (2.32)	5126 (28.41)	4262 (28.27)	12,229 (7.85)
MPR[b], n/N (%)					
<70%	1210/3200 (37.81)	1028/2841 (36.18)	2042/5126 (39.84)	1714/4262 (40.22)	4784/12,229 (39.12)
≥70%	1990/3200 (62.19)	1813/2841 (63.82)	3084/5126 (60.16)	2548/4262 (59.78)	7445/12,229 (60.88)
Use of antihypertensive medication, n (%)					
Yes	13,129 (4.62)	23,169 (18.89)	2205 (12.22)	4149 (27.52)	29,523 (18.95)
MPR, n/N (%)					
<70%	4404/13,129 (33.54)	8066/23,169 (34.81)	774/2205 (35.1)	1468/4149 (35.38)	10,308/29,523 (34.92)
≥70%	8725/13,129 (66.46)	15,103/23,169 (65.19)	1431/2205 (64.9)	2681/4149 (64.62)	19,215/29,523 (65.08)
Use of antihyperlipidemic medication, n (%)					
Yes	2839 (1)	2569 (2.09)	818 (4.53)	742 (4.92)	4129 (2.65)
MPR, n/N (%)					
<70%	1181/2839 (41.6)	1050/2569 (40.87)	343/818 (41.93)	334/742 (45.01)	1727/4129 (41.83)
≥70%	1658/2839 (58.4)	1519/2569 (59.13)	475/818 (58.07)	408/742 (54.99)	2402/4129 (58.17)

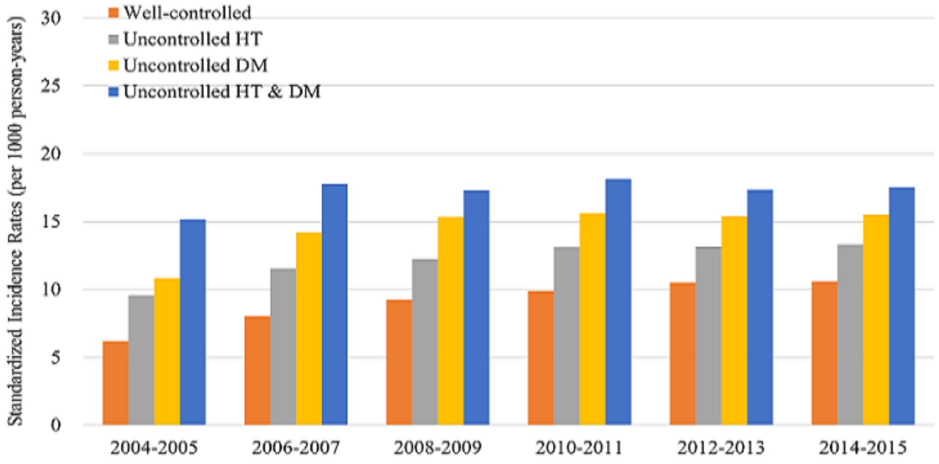
[a] Except for the mean medication possession ratio, all variables significantly differed between the well-controlled group and the total uncontrolled groups at *P*<.001.

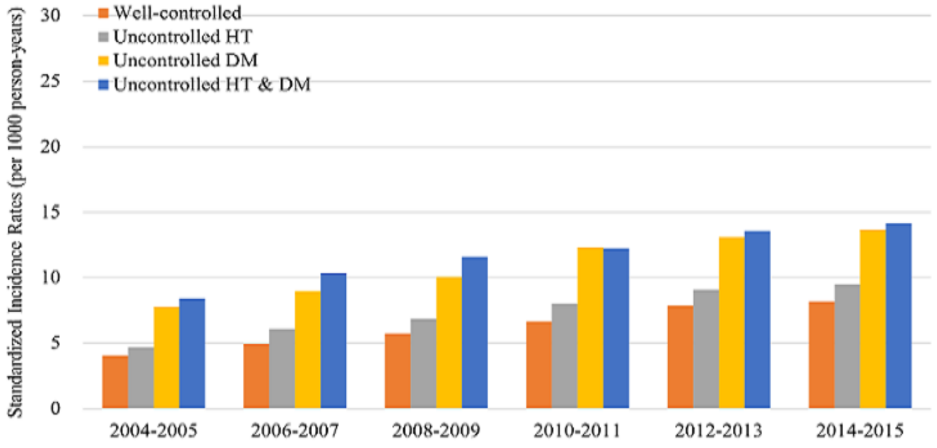
[b] MPR: medication possession ratio.

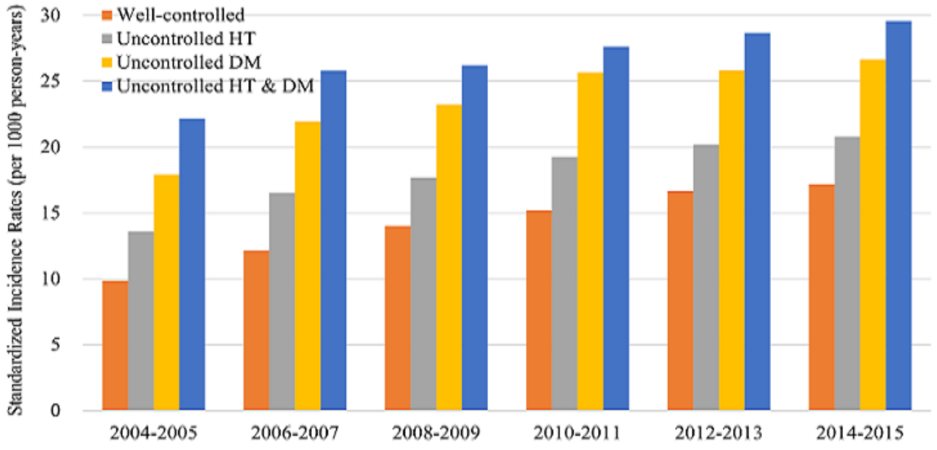
Incidence of MACE or Death From Any Cause

The incidence of MACE increased in the order of the uncontrolled HT group, uncontrolled DM group, and uncontrolled HT and DM group (Figure 2A; Table S2 in Multimedia Appendix 1). Moreover, during the follow-up periods from index periods 1 to 6 with increasing age, the incidence continued to grow; the absolute change in the incidence of MACE was +4.4 (range 6.2-10.6), +3.7 (range 9.6-13.3), and +4.7 (range 10.8-15.5) cases per 1000 person-years among the well-controlled group, uncontrolled HT group, and uncontrolled DM group, respectively. However, in the uncontrolled HT and DM group, the incidence of MACE was high at a relatively young age (15.2 at the age of 52 years on average in 2004-2005 and 17.5 at the age of 61 years on average in 2014-2015) and showed no increasing trend (15.2, 17.8, 17.3, 18.1, 17.3, and 17.5 at every 2 years from 2004-2005 to 2014-2015).

Figure 2. (A) Incidence rates of major adverse cardiovascular event (MACE); (B) death from any cause; and (C) MACE or death from any cause per 1000 person-years by the control status of hypertension (HT) or diabetes mellitus (DM).


(A) MACE


(B) death from any cause


(C) MACE or death from any cause

With respect to death from any cause, the incidence rates per 1000 person-years were higher in the uncontrolled group than in the well-controlled group, especially in the uncontrolled DM group (Figure 2B; Table S2 in Multimedia Appendix 1). In all groups, the increase with age was more pronounced in deaths from any cause than in MACE. Between the follow-up periods 1 and 6, the absolute change in the incidence of death from any cause was +4.1 (range 4.0-8.2), +4.8 (range 4.7-9.5), +5.9 (range 7.8-13.7), and +5.7 (range 8.4-14.1) cases per 1000 person-years among the well-controlled group, the uncontrolled HT group, uncontrolled DM group, and the uncontrolled HT and DM group, respectively.

During each follow-up period, the incidence rates of MACE or death from any cause in the uncontrolled HT and DM group were approximately twice of that in the well-controlled group.

The incidence of MACE or death from any cause increased with age (Figure 2C; Table S2 in Multimedia Appendix 1).

Risk of MACE or Death From Any Cause

Univariable analyses showed that higher age, lower income level, poor adherence (MPR <70%), poor health behaviors (underweight or obese, smoking, drinking every day, and no exercising), or more frequent outpatient visits and hospital stays may be associated with an increased cardiovascular or mortality risk. These variables with statistically significant association on univariable analyses were included in the multivariable model as time-varying covariates for MACE and death from any cause (Table S3 in Multimedia Appendix 1). Uncontrolled DM or HT status was associated with an increased risk of MACE and all-cause mortality (Table 2). Compared with the well-controlled group, the event-rate increase of MACE was 28% in the uncontrolled DM group (HR 1.28, 95% CI 1.23-1.32), 32% in the uncontrolled HT group (HR 1.32, 95% CI 1.29-1.35), and 54% in the uncontrolled HT and DM group (HR 1.54, 95% CI 1.47-1.60). With respect to death from any cause, the event-rate increase was 5% in the uncontrolled DM group (HR 1.05, 95% CI 1.01-1.10), 13% in the uncontrolled HT group (HR 1.13, 95% CI 1.10-1.16), and 17% in the uncontrolled HT and DM group (HR 1.17, 95% CI 1.12-1.23).

Table 2. Adjusted hazard ratios (HRs) for major adverse cardiovascular event (MACE), death from any cause, and MACE or death from any cause associated with risk factors (multivariable analysis[a]).

Control status of HT[b] or DM[c]	MACE		Death from any cause		MACE or death from any cause	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Well-controlled	1.00	Reference	1.00	Reference	1.00	Reference
Uncontrolled HT	1.32 (1.29-1.35)	<.001	1.13 (1.10-1.16)	<.001	1.24 (1.22-1.27)	<.001
Uncontrolled DM	1.28 (1.23-1.32)	<.001	1.05 (1.01-1.10)	<.001	1.18 (1.14-1.21)	<.001
Uncontrolled HT and DM	1.54 (1.47-1.60)	<.001	1.17 (1.12-1.23)	<.001	1.36 (1.31-1.40)	<.001

[a] The analysis based on extended Cox models was adjusted for time-dependent variables, such as age, sex, type of national health security program, income level, BMI, smoking, drinking frequency, exercise frequency, Charlson Comorbidity Index scores, disability severity, number of outpatient visits and hospitalization days, and medication possession ratio.

[b] HT: hypertension.

[c] DM: diabetes mellitus.

Discussion

Principal Findings

This is the first nationwide retrospective cohort study to explore the latest incidence of MACE or death from any cause by HT or DM control status and their association, reflecting the time-varying sociodemographic and clinical characteristics of the population in South Korea.

We observed that 35.38% (155,765/440,249) of the total cohort aged 40 to 79 years belonged to the uncontrolled group for HT or DM in the first index period, and the proportion of uncontrolled HT or DM appeared to decrease over time. However, this decreasing trend may be explained by the exclusion of uncontrolled patients who died or developed MACE over time in the cohort setting. Medication use is important for BP or FBG control [7,10]. Approximately 23.24% (36,192/155,765) of the patients with uncontrolled DM or HT were prescribed medication. Similar to our study, previous studies have shown low prescription rates in patients with HT and DM [3]. Among medication users, uncontrolled HT or DM accounted for >60% of patients, which was comparable with the BP and blood glucose control rate of 50% for the treated patients reported in foreign real-world clinical practice [21,22]. These results may be because possessing a prescribed medication does not necessarily mean adherence to the treatment. Nonetheless, our findings emphasize the need for novel BP- and blood glucose–lowering strategies, along with managing modifiable risk factors for cardiovascular diseases. The implementation of an efficient system such as mobile health, which has been proven effective in controlling HT and DM by continuously monitoring diseases through improving the health care delivery process, may be considered [23,24]. Along with this, sustained, community-wide programs targeting cardiovascular risk factors and behavioral changes to improve health would be beneficial [5].

Our study also showed the association between medication use and MACE or death from any cause (Table S3 in Multimedia Appendix 1). The risk of MACE or death from any cause was higher in the poor adherence group (MPR <70%) than in the adherence group (MPR ≥70%). It is well known that poor adherence to antihypertensive medication and antidiabetic medication may increase the risk of cardiovascular events [25,26]. Therefore, understanding the factors related to adherence to medication could be the key to preventing MACE and death. Strategies tailored to the economical, geographical, sociological, educational, and health care system contexts of patients are effective in improving adherence to medication [27]. Further studies are needed to identify the barriers that affect adherence to medication. In addition, the study showed that nonmedication users may have a lower risk of MACE or death from any cause than medication users. A lower risk for nonmedication users was expected because most of them achieved controlled BP or glycemic status without medication use. According to the guidelines on the prevention of cardiovascular diseases, patients with HT or DM requiring medication treatment have a higher cardiovascular risk than those who do not require medication treatment [7].

The study indicated that poor health behaviors (underweight or obese, smoking, drinking every day, and no exercise) or more frequent outpatient visits and hospital stays may be associated with an increased cardiovascular or mortality risk. These findings are consistent with previously reported results [28]. In terms of drinking behavior, a trend toward higher mortality or cardiovascular risk was observed in nondrinkers than in moderate drinkers. The causal role of alcohol consumption in cardiovascular diseases remains unclear [29].

We found that the incidence of MACE in South Korean adults aged 40 to 79 years was 10.8 to 15.5 and 9.6 to 13.3 per 1000 person-years during the follow-up periods in the uncontrolled DM and HT groups, respectively. After adjusting for medication use and other risks, the uncontrolled HT and DM groups had a 1.32-fold and 1.28-fold higher risk of MACE than the well-controlled group, respectively. In particular, the uncontrolled HT and DM group had a 1.54-fold higher risk of MACE than the well-controlled group. The higher incidence in the uncontrolled group may have been influenced by the relatively poor health behaviors and low economic status of the uncontrolled group. These clustered factors are likely to mediate the impact on ill health and premature mortality [30] and interact synergistically, increasing the individual's total risk of cardiovascular diseases [4]. Most previous studies have estimated the incidence of cardiovascular disease in newly diagnosed patients, regardless of the control status of BP or FBG [2,31]. In an earlier study including patients in South Korea aged 20 to 89 years with type 2 diabetes, the incidence of hospitalization for coronary artery disease and cerebrovascular disease per 1000 person-years was reported as 6.11 and 4.50, respectively, during 2009 to 2011. These results, which are lower than those in our study, could be because the study was conducted on patients with DM who have prescribed medication and included some patients with controlled DM.

In this study, the patients were older, and the mean age of the total cohort from 2002 to 2003 was 52 years, and the mean age of the cohort in 2012 to 2013, excluding patients in whom the outcomes occurred, was 61 years. Similar to a previous study [32], the incidence of MACE tended to increase with age in the uncontrolled HT group and uncontrolled DM group. This trend was particularly marked for the incidence of death from any cause. However, in the uncontrolled HT and DM group, the incidence of MACE was high at a relatively early age, and no age-related trends were observed. There were quite a few participants in the uncontrolled HT and DM group in the study, accounting for 9.68% (15,077/155,765) of the total uncontrolled group. The percentage of uncontrolled HT and DM group may be underestimated owing to the masking HT in patients with high FBG levels [33,34]. HT and DM are common comorbidities linked by various vascular mechanisms and risk factors and accelerate each other's development [35]. The coexistence of HT and DM contributes to the development of cardiovascular disease by amplifying vascular damage and endothelial dysfunction [36]. Therefore, managing these patients at a young age can substantially reduce the incidence of cardiovascular diseases. For example, if patients with uncontrolled HT and DM control their BP and FBG levels at the age of 52 years, the incidence of MACE would be reduced by 60% (6.2 and 15.2 per 1000 person-years in the well-controlled group and the uncontrolled HT and DM group, respectively; Table S2 in Multimedia Appendix 1). These findings indicate that the uncontrolled HT and DM group requires strict management as early as possible. For example, more aggressive treatment with a BP target of <130/80 mm Hg could be considered in adults with DM and HT, consistent with the American Society's recommendation for managing high BP [37].

In traditional Cox regression models, risk factors measured at baseline are usually associated with clinical outcomes occurring after a period. However, things may change during follow-up; either the effect of a fixed baseline risk factor may vary over time, resulting in a weakening or strengthening of associations over time, or the risk factor itself may vary over time [38]. Thus, a more detailed analysis of lifetime risks would take repeated measures into consideration for potential changes and would adjust for the time-dependent effect of contributing risk factors [39]. Lifestyle factors such as medication use, smoking, exercise, and obesity have changed significantly over the past 20 years among adults in South Korea [6], and BP and FBG levels are also factors that change over time. However, few studies have been conducted taking this into account to examine the association of control of HT or DM with cardiovascular events or mortality. This study accounted for changes in these factors over time. A previous study using traditional Cox analysis conducted in South Korea found that HT and DM were more strongly associated with hospitalization for coronary heart disease than that reported in our results [40]. The conservative results in our study may be explained by the fact that most important clinical risk factors such as medication prescription and adherence were considered and reflected as time-dependent manners. Our results using extended Cox analysis with adjusting various time-dependent factors indicated that uncontrolled HT and uncontrolled DM were independent risk factors for MACE.

This study has several strengths. First, the study used a large national cohort with a low rate of follow-up loss over a period of >10 years [17]. Second, when potential risk factors change over time, it may be inappropriate from a clinical perspective to correlate all future survival rates with the factors evaluated at one moment [38]. Thus, time-varying variables such as BP, FBG, medication prescription, and adherence were adjusted using extended Cox analysis. It allows for the accurate attribution of risk factors for each patient at the time of MACE or death events. Third, as clinical laboratory tests were performed for all national health screening program participants, our study included not only diagnosed patients but also patients who were not aware of HT or DM or were diagnosed but untreated patients. Diagnosis based on these clinical data decreases the likelihood of misclassification of disease status, which causes a bias of estimates of association [41].

Limitations

This study had some limitations. First, BP was measured using a device validated in various hospitals and clinics certified by the National Health Insurance Service. However, the BP measurement devices used at each institution were not uniform, and the measured values may differ depending on the measurement method. BP measured using automated office BP measurements is generally 5 to 10 mm Hg lower than the typical manual office BP [42,43]. Second, we used FBG criteria for the diagnosis of DM, not hemoglobin A_{1c} (HbA_{1c}) criteria, because the NHIS-HEALS database did not include the HbA_{1c} variable. However, FBG levels were closely associated with specific HbA_{1c} levels in the South Korean population, and FBG criteria were reported to be valid in the South Korean population [44]. Third, although our study controlled for various measured confounders, including medication use and patient characteristics by multivariable analysis, additional or unmeasured confounders may exist. Fourth, as our data were not linked to the cause of death data provided by Statistics Korea, we could not estimate the cause-specific mortality. Fifth, although we considered time-varying cardiovascular risk profiles and lifestyle using extended Cox analysis, the trajectories of BP or FBG changes over the observation period were not reflected. Trajectory modeling assessing the association of health outcomes with subgroups of individuals that share similar patterns of laboratory assessments is a subject worthy of further research [45]. Despite these limitations, the screening cohort data, including laboratory results and all health care use information used in this study, provide a powerful resource for an up-to-date population-based evaluation.

Conclusions

This study highlights the importance of controlling BP or FBG levels to reduce the risk of MACE or death. This up-to-date understanding of the association between BP and FBG control and cardiovascular events serves as the basic data for planning public health policies to prevent cardiovascular diseases. Moreover, the high uncontrolled rates of HT or DM, regardless of medication prescription, have led us to suggest the need for novel efficient systems such as a community-wide management program using mobile health technology that is effective in continuously controlling high BP and blood glucose levels.

Acknowledgments

The authors would like to thank the National Health Insurance Service for providing data (National Health Insurance Service–National Health Screening Cohort) for this study. This research was supported by a grant from the Korea Health Promotion Research and Development Project funded by the Ministry of Health and Welfare, South Korea (grant HS20C0041).

Data Availability

The data underlying the results presented in the study are available from the National Health Insurance Data Sharing Service after the committee of the National Health Insurance Service reviews the application for approval of the use of National Health Insurance Service–National Health Screening Cohort data. Researchers receive access privileges to the requested database after paying the fee. Interested researchers will be able to access the data in the same manner and replicate the findings of our study by following the protocol outlined in the *Methods* section.

Authors' Contributions

JWK conceived the idea and achieved funding for this study. SHO contributed to data acquisition and analysis. DK and JH contributed to statistical analysis. JWK, SHO, JHK, and YK contributed to data interpretation. SHO wrote the first draft of the manuscript, and JWK, JHK, and YK revised the manuscript. JWK is the guarantor of this work and has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Selection flowchart of study population, definition of outcomes, and incidence rates and hazard ratios of major adverse cardiovascular events (MACE), death from any cause, and MACE or death from any cause.
[DOCX File , 81 KB - publichealth_v9i1e42190_app1.docx]

References

1. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2020 update: a report from the American Heart Association. Circulation 2020 Mar 03;141(9):e139-e596 [FREE Full text] [doi: 10.1161/CIR.0000000000000757] [Medline: 31992061]
2. Quan H, Chen G, Walker RL, Wielgosz A, Dai S, Tu K, Hypertension Outcome and Surveillance Team. Incidence, cardiovascular complications and mortality of hypertension by sex and ethnicity. Heart 2013 May;99(10):715-721. [doi: 10.1136/heartjnl-2012-303152] [Medline: 23403406]

3. Jeong H, Kim H, Lee K, Lee JH, Ahn HM, Shin SA, et al. Medical visits, antihypertensive prescriptions and medication adherence among newly diagnosed hypertensive patients in Korea. Environ Health Prev Med 2017 Mar 17;22(1):10 [FREE Full text] [doi: 10.1186/s12199-017-0619-6] [Medline: 29165108]
4. Dahlöf B. Cardiovascular disease risk factors: epidemiology and risk assessment. Am J Cardiol 2010 Jan 04;105(1 Suppl):3A-9A. [doi: 10.1016/j.amjcard.2009.10.007] [Medline: 20102968]
5. Vasan RS, Benjamin EJ. The future of cardiovascular epidemiology. Circulation 2016 Jun 21;133(25):2626-2633 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.116.023528] [Medline: 27324358]
6. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
7. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 2019 Sep 10;140(11):e596-e646 [FREE Full text] [doi: 10.1161/CIR.0000000000000678] [Medline: 30879355]
8. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980 [FREE Full text] [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
9. Ogurtsova K, Guariguata L, Barengo NC, Ruiz PL, Sacre JW, Karuranga S, et al. IDF diabetes Atlas: global estimates of undiagnosed diabetes in adults for 2021. Diabetes Res Clin Pract 2022 Jan;183:109118. [doi: 10.1016/j.diabres.2021.109118] [Medline: 34883189]
10. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.120.15026] [Medline: 32370572]
11. Fang M, Wang D, Coresh J, Selvin E. Trends in diabetes treatment and control in U.S. adults, 1999-2018. N Engl J Med 2021 Jun 10;384(23):2219-2228 [FREE Full text] [doi: 10.1056/NEJMsa2032271] [Medline: 34107181]
12. Łukasiewicz A, Cichoń E, Kostecka B, Kiejna A, Jodko-Modlińska A, Obrębski M, et al. Association of higher rates of type 2 diabetes (T2DM) complications with psychological and demographic variables: results of a cross-sectional study. Diabetes Metab Syndr Obes 2022 Oct 28;15:3303-3317 [FREE Full text] [doi: 10.2147/DMSO.S369809] [Medline: 36329807]
13. Sandy JL, Besançon S, Sidibé AT, Minkailou M, Togo A, Ogle GD. Rapid increases in observed incidence and prevalence of type 1 diabetes in children and youth in Mali, 2007-2016. Pediatr Diabetes 2021 Jun;22(4):545-551. [doi: 10.1111/pedi.13191] [Medline: 33586301]
14. Denisova AR, Solntseva TD, Zarmanbetova AS, Tkacheva AA, Sivakova OA, Chazova IE. [The incidence of cardiovascular and cerebrovascular complications in patients with uncontrolled hypertension]. Ter Arkh 2022 Jan 15;94(1):94-99. [doi: 10.26442/00403660.2022.01.201395] [Medline: 36286923]
15. Shin JI, Oh J, Kim HC, Choi D, Yoon YS. Current state of cardiovascular research in Korea. Circ Res 2019 Dec 06;125(12):1141-1145 [FREE Full text] [doi: 10.1161/CIRCRESAHA.119.310859] [Medline: 31804914]
16. Hajifathalian K, Ueda P, Lu Y, Woodward M, Ahmadvand A, Aguilar-Salinas CA, et al. A novel risk score to predict cardiovascular disease risk in national populations (Globorisk): a pooled analysis of prospective cohorts and health examination surveys. Lancet Diabetes Endocrinol 2015 May;3(5):339-355. [doi: 10.1016/S2213-8587(15)00081-9] [Medline: 25819778]
17. Seong SC, Kim YY, Park SK, Khang YH, Kim HC, Park JH, et al. Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. BMJ Open 2017 Sep 24;7(9):e016640 [FREE Full text] [doi: 10.1136/bmjopen-2017-016640] [Medline: 28947447]
18. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
19. Kozma CM, Dickson M, Phillips AL, Meletiche DM. Medication possession ratio: implications of using fixed and variable observation periods in assessing adherence with disease-modifying drugs in patients with multiple sclerosis. Patient Prefer Adherence 2013 Jun 12;7:509-516 [FREE Full text] [doi: 10.2147/PPA.S40736] [Medline: 23807840]
20. Bosco E, Hsueh L, McConeghy KW, Gravenstein S, Saade E. Major adverse cardiovascular event definitions used in observational analysis of administrative databases: a systematic review. BMC Med Res Methodol 2021 Nov 06;21(1):241 [FREE Full text] [doi: 10.1186/s12874-021-01440-5] [Medline: 34742250]
21. Jurić D, Pavličević I, Marušić A, Malički M, Buljan I, Šarotić V, et al. Effectiveness of treatment of newly diagnosed hypertension in family medicine practices in South Croatia. BMC Fam Pract 2019 Jan 14;20(1):10 [FREE Full text] [doi: 10.1186/s12875-019-0902-2] [Medline: 30642264]
22. Karagiannis T, Paschos P, Paletas K, Matthews DR, Tsapas A. Dipeptidyl peptidase-4 inhibitors for treatment of type 2 diabetes mellitus in the clinical setting: systematic review and meta-analysis. BMJ 2012 Mar 12;344:e1369. [doi: 10.1136/bmj.e1369] [Medline: 22411919]

23. So CF, Chung JW. Telehealth for diabetes self-management in primary healthcare: a systematic review and meta-analysis. J Telemed Telecare 2018 Jun;24(5):356-364. [doi: 10.1177/1357633X17700552] [Medline: 28463033]
24. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776 [FREE Full text] [doi: 10.2196/17776] [Medline: 32217503]
25. Fukuda H, Mizobe M. Impact of nonadherence on complication risks and healthcare costs in patients newly-diagnosed with diabetes. Diabetes Res Clin Pract 2017 Jan;123:55-62. [doi: 10.1016/j.diabres.2016.11.007] [Medline: 27940390]
26. Shin S, Song H, Oh SK, Choi KE, Kim H, Jang S. Effect of antihypertensive medication adherence on hospitalization for cardiovascular disease and mortality in hypertensive patients. Hypertens Res 2013 Nov;36(11):1000-1005. [doi: 10.1038/hr.2013.85] [Medline: 23966057]
27. Macquart de Terline D, Kane A, Kramoh KE, Ali Toure I, Mipinda JB, Diop IB, et al. Factors associated with poor adherence to medication among hypertensive patients in twelve low and middle income Sub-Saharan countries. PLoS One 2019 Jul 10;14(7):e0219266 [FREE Full text] [doi: 10.1371/journal.pone.0219266] [Medline: 31291293]
28. Park D, Lee JH, Han S. Underweight: another risk factor for cardiovascular disease?: a cross-sectional 2013 Behavioral Risk Factor Surveillance System (BRFSS) study of 491,773 individuals in the USA. Medicine (Baltimore) 2017 Dec;96(48):e8769 [FREE Full text] [doi: 10.1097/MD.0000000000008769] [Medline: 29310352]
29. Larsson SC, Burgess S, Mason AM, Michaëlsson K. Alcohol consumption and cardiovascular disease: a Mendelian randomization study. Circ Genom Precis Med 2020 Jun;13(3):e002814 [FREE Full text] [doi: 10.1161/CIRCGEN.119.002814] [Medline: 32367730]
30. Syddall HE, Westbury LD, Simmonds SJ, Robinson S, Cooper C, Sayer AA. Understanding poor health behaviours as predictors of different types of hospital admission in older people: findings from the Hertfordshire Cohort Study. J Epidemiol Community Health 2016 Mar;70(3):292-298 [FREE Full text] [doi: 10.1136/jech-2015-206425] [Medline: 26481495]
31. Koo BK, Lee CH, Yang BR, Hwang SS, Choi NK. The incidence and prevalence of diabetes mellitus and related atherosclerotic complications in Korea: a National Health Insurance Database Study. PLoS One 2014 Oct 16;9(10):e110650 [FREE Full text] [doi: 10.1371/journal.pone.0110650] [Medline: 25329714]
32. Kim H, Kim S, Han S, Rane PP, Fox KM, Qian Y, et al. Prevalence and incidence of atherosclerotic cardiovascular disease and its risk factors in Korea: a nationwide population-based study. BMC Public Health 2019 Aug 14;19(1):1112 [FREE Full text] [doi: 10.1186/s12889-019-7439-0] [Medline: 31412823]
33. Franklin SS, Thijs L, Li Y, Hansen TW, Boggia J, Liu Y, International Database on Ambulatory blood pressure in Relation to Cardiovascular Outcomes Investigators. Masked hypertension in diabetes mellitus: treatment implications for clinical practice. Hypertension 2013 May;61(5):964-971 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.00289] [Medline: 23478096]
34. Ben-Dov IZ, Ben-Ishay D, Mekler J, Ben-Arie L, Bursztyn M. Increased prevalence of masked blood pressure elevations in treated diabetic subjects. Arch Intern Med 2007 Oct 22;167(19):2139-2142. [doi: 10.1001/archinte.167.19.2139] [Medline: 17954811]
35. Li X, Wang J, Shen X, An Y, Gong Q, Li H, et al. Higher blood pressure predicts diabetes and enhances long-term risk of cardiovascular disease events in individuals with impaired glucose tolerance: twenty-three-year follow-up of the Daqing diabetes prevention study. J Diabetes 2019 Jul;11(7):593-598 [FREE Full text] [doi: 10.1111/1753-0407.12887] [Medline: 30556339]
36. Petrie JR, Guzik TJ, Touyz RM. Diabetes, hypertension, and cardiovascular disease: clinical insights and vascular mechanisms. Can J Cardiol 2018 May;34(5):575-584 [FREE Full text] [doi: 10.1016/j.cjca.2017.12.005] [Medline: 29459239]
37. Whelton PK, Carey RM, Aronow WS, Casey Jr DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2018 May 15;71(19):e127-e248 [FREE Full text] [doi: 10.1016/j.jacc.2017.11.006] [Medline: 29146535]
38. Dekker FW, de Mutsert R, van Dijk PC, Zoccali C, Jager KJ. Survival analysis: time-dependent effects and time-varying risk factors. Kidney Int 2008 Oct;74(8):994-997 [FREE Full text] [doi: 10.1038/ki.2008.328] [Medline: 18633346]
39. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. Lancet 2014 May 31;383(9932):1899-1911 [FREE Full text] [doi: 10.1016/S0140-6736(14)60685-1] [Medline: 24881994]
40. Jee SH, Jang Y, Oh DJ, Oh BH, Lee SH, Park SW, et al. A coronary heart disease prediction model: the Korean Heart Study. BMJ Open 2014 May 21;4(5):e005025 [FREE Full text] [doi: 10.1136/bmjopen-2014-005025] [Medline: 24848088]
41. Chen Q, Galfalvy H, Duan N. Effects of disease misclassification on exposure-disease association. Am J Public Health 2013 May;103(5):e67-e73. [doi: 10.2105/AJPH.2012.300995] [Medline: 23488509]
42. SPRINT Research Group, Wright JT, Williamson JD, Whelton PK, Snyder JK, Sink KM, et al. A randomized trial of intensive versus standard blood-pressure control. N Engl J Med 2015 Nov 26;373(22):2103-2116 [FREE Full text] [doi: 10.1056/NEJMoa1511939] [Medline: 26551272]

43. American Diabetes Association. 9. Cardiovascular disease and risk management: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018 Jan;41(Suppl 1):S86-104. [doi: 10.2337/dc18-S009] [Medline: 29222380]
44. Hong S, Kang JG, Kim CS, Lee SJ, Lee CB, Ihm SH. Fasting plasma glucose concentrations for specified HbA1c goals in Korean populations: data from the Fifth Korea National Health and Nutrition Examination Survey (KNHANES V-2, 2011). Diabetol Metab Syndr 2016 Aug 30;8(1):62 [FREE Full text] [doi: 10.1186/s13098-016-0179-8] [Medline: 27579145]
45. Nguena Nguefack HL, Pagé MG, Katz J, Choinière M, Vanasse A, Dorais M, et al. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. Clin Epidemiol 2020 Oct 30;12:1205-1222 [FREE Full text] [doi: 10.2147/CLEP.S265287] [Medline: 33154677]

Abbreviations

BP: blood pressure
CCI: Charlson Comorbidity Index
DM: diabetes mellitus
FBG: fasting blood glucose
HbA_{1c}: hemoglobin A_{1c}
HR: hazard ratio
HT: hypertension
IHD: ischemic heart disease
MACE: major adverse cardiovascular events
MPR: medication possession ratio
NHIS-HEALS: National Health Insurance Service–National Health Screening Cohort

Edited by A Mavragani; submitted 26.08.22; peer-reviewed by HS Sohn, EJ Park, T Wan; comments to author 26.10.22; revised version received 27.11.22; accepted 06.01.23; published 03.02.23.

Please cite as:
Oh SH, Kim D, Hwang J, Kang JH, Kwon Y, Kwon JW
Association of Uncontrolled Hypertension or Diabetes Mellitus With Major Adverse Cardiovascular Events and Mortality in South Korea: Population-Based Cohort Study
JMIR Public Health Surveill 2023;9:e42190
URL: https://publichealth.jmir.org/2023/1/e42190
doi:10.2196/42190
PMID:36735297

Original Paper

Multimorbidity and SARS-CoV-2–Related Outcomes: Analysis of a Cohort of Italian Patients

Alberto Catalano[1], MSc; Lucia Dansero[1], MSc; Winston Gilcrease[1,2], MSc; Alessandra Macciotta[1], MSc; Carlo Saugo[3], MD; Luca Manfredi[1], MSc; Roberto Gnavi[4], MD; Elena Strippoli[4], MSc; Nicolás Zengarini[4], MSc; Valeria Caramello[5], MD; Giuseppe Costa[1,4], MD; Carlotta Sacerdote[6], MD, PhD; Fulvio Ricceri[1,4], PhD

[1]Centre for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Sciences, University of Turin, Orbassano (TO), Italy
[2]UNESCO Chair in Sustainable Development and Territory Management, University of Turin, Turin, Italy
[3]Klinik für Innere Medizin - Gastroenterologie, Hepatologie & Infektiologie, Helios Klinikum Berlin-Buch, Berlin, Germany
[4]Unit of Epidemiology, Regional Health Service ASLTO3, Grugliasco (TO), Italy
[5]Emergency Department and High Dependency Unit, San Luigi Gonzaga University Hospital, Orbassano (TO), Italy
[6]Unit of Cancer Epidemiology, Città della Salute e della Scienza University Hospital, Turin, Italy

Corresponding Author:
Alberto Catalano, MSc
Centre for Biostatistics, Epidemiology, and Public Health
Department of Clinical and Biological Sciences
University of Turin
Regione Gonzole 10
Orbassano (TO), 10043
Italy
Phone: 39 0116705443
Email: alberto.catalano@unito.it

Abstract

Background: Since the outbreak of the COVID-19 pandemic, identifying the main risk factors has been imperative to properly manage the public health challenges that the pandemic exposes, such as organizing effective vaccination campaigns. In addition to gender and age, multimorbidity seems to be 1 of the predisposing factors coming out of many studies investigating the possible causes of increased susceptibility to SARS-CoV-2 infection and adverse outcomes. However, only a few studies conducted have used large samples.

Objective: The objective is to evaluate the association between multimorbidity, the probability to be tested, susceptibility, and the severity of SARS-CoV-2 infection in the Piedmont population (Northern Italy, about 4 million inhabitants). For this purpose, we considered 5 main outcomes: access to the swab, positivity to SARS-CoV-2, hospitalization, intensive care unit (ICU) admission, and death within 30 days from the first positive swab.

Methods: Data were obtained from different Piedmont health administrative databases. Subjects aged from 45 to 74 years and infections diagnosed from February to May 2020 were considered. Multimorbidity was defined both with the Charlson Comorbidity Index (CCI) and by identifying patients with previous comorbidities, such as diabetes and oncological, cardiovascular, and respiratory diseases. Multivariable logistic regression models (adjusted for age and month of infection and stratified by gender) were performed for each outcome. Analyses were also conducted by separating 2 age groups (45-59 and 60-74 years).

Results: Of 1,918,549 subjects, 85,348 (4.4%) performed at least 1 swab, of whom 12,793 (14.9%) tested positive for SARS-CoV-2. Of these 12,793 subjects, 4644 (36.3%) were hospitalized, 1508 (11.8%) were admitted to the ICU, and 749 (5.9%) died within 30 days from the first positive swab. Individuals with a higher CCI had a higher probability of being swabbed but a lower probability of testing positive. We observed the same results when analyzing subjects with previous oncological and cardiovascular diseases. Moreover, especially in the youngest group, we identified a greater risk of being hospitalized and dying. Among comorbidities considered in the study, respiratory diseases seemed to be the most likely to increase the risk of having a positive swab and worse disease outcomes.

Conclusions: Our study shows that patients with multimorbidity, although swabbed more frequently, are less likely to get infected with SARS-CoV-2, probably due to greater attention on protective methods. Moreover, a history of respiratory diseases is a risk factor for a worse prognosis of COVID-19. Nonetheless, whatever comorbidities affect the patients, a strong dose-response effect was observed between an increased CCI score and COVID-19 hospitalization, ICU admission, and death. These results are important in terms of public health because they help in identifying a group of subjects who are more prone to worse SARS-CoV-2 outcomes. This information is important for promoting targeted prevention and developing policies for the prioritization of public health interventions.

(JMIR Public Health Surveill 2023;9:e41404) doi:10.2196/41404

KEYWORDS

multimorbidity; SARS-CoV-2; mortality; intensive care unit; epidemiology; COVID-19; pandemic; severity; cardiovascular; respiratory; disease; risk; public health; intervention

Introduction

Background

Since the outbreak of the COVID-19 pandemic in Italy and the world, identifying the risk factors for the development of a severe prognosis in patients with SARS-CoV-2 infection has been imperative to properly manage public health challenges that the pandemic exposes. For instance, organizing effective vaccination campaigns when just a limited number of vaccines were available was essential, and it is still important at the moment to decide on the prioritization of new campaigns for subsequent vaccine doses [1,2]. Moreover, the availability of accurate information about the possible evolution of the disease in the presence of risk factors can support the population in the choice of prudent behaviors aimed at preventing the spread of the contagion chain, with particular regard to the most vulnerable individuals.

Many studies in the literature have focused on the sociodemographic and clinical characteristics that are most associated with the development of a severe prognosis. Age and gender play an important role in determining the prognosis of patients with SARS-CoV-2. The elderly are especially at higher risk of becoming severely ill [3-5], and males are also a greater risk of a severe prognosis [5,6]. Among the several mechanisms proposed to explain the severity of COVID-19 in older adults, the increased burden of multimorbidity is 1 of the most relevant [7]. Moreover, several studies also have shown that multimorbidity affects COVID-19 severity in a way that is age independent, underscoring the need for an extensive study of the relationship between multimorbidity and SARS-CoV-2 outcomes [7].

Available evidence related to the identification of comorbidities that can negatively impact the prognosis in patients with SARS-CoV-2 refers essentially to observational, cross-sectional, retrospective cohort studies or case-control studies. There is general agreement on the role of hypertension, cardiovascular disease, and diabetes in increasing the risk of severe disease [5,6,8-10]. Therefore, comorbidities that affect the vascular system may play a key role in the worsening of clinical conditions in people affected by COVID-19. Moreover, there is some consensus on the role that obesity, cancer, chronic obstructive pulmonary disease (COPD), chronic kidney disease, and immunodepression could play in aggravating the prognosis [4-6,9]. However, these associations have been highlighted in a more sporadic and less uniform way.

In an initial phase, the correlation between the presence and the number of comorbidities with a more unfavorable outcome was pointed out to evaluate the prognosis of patients with SARS-CoV-2 [9]. This tool has been nevertheless effective and useful in individuating the key feature of the population at risk, which was crucial during the first phase of the pandemic.

It was in fact observed that subjects with some comorbidities (measured as a Charlson Comorbidity Index [CCI] [11]) equal to or greater than 1 are at higher risk of severe COVID-19 outcomes compared to subjects with no comorbidity [12,13]. Moreover, an innovative approach to the SARS-CoV-2 pandemic (the so-called “syndemic approach”) suggests that both bio-bio and biosocial interactions can act as complex drivers that increase subjects’ susceptibility to worsen COVID-19 outcomes [14-17].

Consequently, it is crucial to identify multifactorial profiles, including social determinants of health, and multimorbidity patterns associated with COVID-19 outcomes in order to recognize a broader syndemic health burden in specific subgroups of subjects with those characteristics [18,19].

Study Design and Aim

This large population-based region-wide study based on administrative health databases aims to evaluate the association between multimorbidity and the susceptibility and severity of SARS-CoV-2 infection in the Piedmont population (Northern Italy, about 4 million inhabitants) aged from 45 to 74 years, considering COVID-19 infections diagnosed from February to May 2020.

Specifically, 5 different main outcomes were considered: access to the swab, positivity to SARS-CoV-2, hospitalization, admission to the intensive care unit (ICU), and death within 30 days from the first positive swab.

Methods

Study Population

For the analyses, data were obtained from the Piedmont Longitudinal Study (PLS), an administrative cohort based on the anonymous record linkage at the individual level of different social, health, and administrative databases. This study includes 2011 census data, hospital discharges, a mortality register, an outpatients register, an exemption register (in Italy, subjects with at least 1 chronic disease are exempt from paying for examinations related to their disease), and a drug prescriptions database. In addition, starting from February 2020, this study was enriched by the regional platform about COVID-19, in which infection data about subjects who had at least 1 contact with the regional health system related to SARS-CoV-2 are collected.

The study population comprised all assisted and domiciled subjects in Piedmont aged from 45 to 74 years. These 2 age classes were chosen in order to exclude young and elderly subjects, as they could have influenced the results. Furthermore, this choice made it possible to exclude the majority of patients residing in nursing homes in order to avoid any bias. With regard to SARS-CoV-2 infection, we considered those patients who developed the infection in Piedmont from February 22, 2020 (when the first case of SARS-CoV-2 was recorded in Italy), to May 31, 2020.

Variables' Definitions

Multimorbidity was defined using the CCI [11] retrieved by the record linkage with data of the hospital discharges and the drug prescription register between 2015 and 2019. In Multimedia Appendix 1, Table S1a, we present the algorithm's definition.

The presence of prevalent oncological, cardiovascular, and respiratory diseases was also retrieved, as well as prevalent previous myocardial infarctions, heart failures, cerebrovascular diseases, and diabetes (algorithms presented in Multimedia Appendix 1, Table S1b).

Five different outcomes were considered: access to a SARS-CoV-2 test (nasal swab in almost all cases), positivity to the test, hospitalization, admission to the ICU, and mortality within 30 days after testing positive for SARS-CoV-2 infection. The 4 latter outcomes were compared to both the general population and to the total tested subjects or the subjects positive for SARS-CoV-2, as appropriate.

Statistical Analysis

All the variables were described using absolute frequencies and percentages. For each outcome and exposure, we fitted a multivariable logistic regression model, in which the odds ratio (OR) estimates, together with their 95% Cls, were adjusted for age (used as a continuous variable) and for the month of infection or the first negative swab, where appropriate.

Moreover, to assess the probability of transition into various states on the basis of the CCI, multistate models were implemented. The possible transitions that were considered in the study involved 3 different states: from positivity to COVID-19 to hospitalization, from positivity to death within 30 days of the initial positive swab, and from hospitalization to 30-day mortality (Multimedia Appendix 1, Figure S1). In 1 case, the multistate models were stratified by the CCI score and adjusted for age and the month of infection; in the second case, the multistate models were adjusted for age, the month of infection, and the CCI.

All analyses were stratified by gender and 2 different age groups, 45-59 and 60-74 years, because we observed the differences between gender and age as effect modifiers.

Analyses were performed using SAS (V9.4) and R (V4.2.1).

Ethical Considerations

All analyses were conducted according to the World Medical Association's Declaration of Helsinki. In fact, the study is included in the National Statistical Plan and did not need approvals or permits from the ethics committee. For privacy purposes, the data used for analysis were anonymized.

Results

Study Population

Of 1,918,549 assisted and domiciled subjects in Piedmont aged from 45 to 74 years, 85,348 (4.4%) performed at least 1 swab during the observation period, of whom 12,793 (14.9%) tested positive for SARS-CoV-2. Of these 12,793 patients, 4644 (36.3%) were hospitalized due to COVID-19, 1508 (11.8%) were admitted to the ICU, and 749 (5.9%) died within 30 days from the first positive swab.

The descriptive statistics related to the exposures and variables used in the study are shown in Table 1. Among other results, the most interesting was related to gender. Overall, despite more women testing positive for SARS-CoV-2 (n=6832, 53.4%), higher percentages of men were hospitalized (3074/4644, 66.2%), were admitted to the ICU (1118/1508, 74.1%), and died (558/749, 74.5%). The same trend was observed when considering age: although more subjects aged 45-59 years tested positive for SARS-CoV-2 (n=7324, 57.2%), more individuals aged 60-74 years were hospitalized (2927/4644, 63%), were admitted to the ICU (1005/1508, 66.6%), and died (629/749, 84%).

Descriptive analyses also showed that multimorbidity (measured as CCI>1) increased the probability of being swabbed, hospitalized, admitted to the ICU, and dying from COVID-19.

Table 2 shows the OR estimates, together with their 95% Cls, related to accessing swabs for the entire population. The results showed that as the CCI increased, the likelihood of undergoing swab testing increased for both the male and the female population. In addition, other estimates indicated that being affected by 1 of the diseases considered in the study increased the likelihood of undergoing at least 1 swab test, especially in the case of heart failure, respiratory diseases, and cerebrovascular diseases (males aged 45-59 years: OR 2.60, 95% CI 2.30-2.94; males aged 60-74 years: OR 2.71, 95% CI 2.53-2.90; females aged 45-59 years: OR 1.63, 95% CI 1.45-1.85; females aged 60-74 years: OR 3.31, 95% CI 3.04-3.60).

The relationship between multimorbidity and SARS-CoV-2 positivity among those who performed at least 1 swab is shown in Table 3. The OR estimates indicated an inverse trend: the likelihood of testing positive for SARS-CoV-2 infection decreased as the CCI value increased. Moreover, those who were affected by oncological diseases had a significantly lower likelihood of testing positive for SARS-CoV-2, especially men (45-59 years old: OR 0.72, 95% CI 0.55-0.95; 60-74 years old: OR 0.63, 95% CI 0.53-0.73). Other specific diseases were less associated with the results of the SARS-CoV-2 infection test. On the contrary, when compared to the general population, multimorbid patients had an increased likelihood of testing positive (Multimedia Appendix 1, Table S2).

Table 4 shows the results related to the subjects who were admitted to the hospital from among those who tested positive for SARS-CoV-2. The results indicated that particularly for men from the ages of 45 to 59 years, the probability of hospital admission rose significantly as the CCI score increased (CCI=1: OR 1.50, 95% CI 1.20-1.88; CCI=2-3: OR 2.14, 95% CI 1.51-3.01; CCI=4+: OR 4.77, 95% CI 2.28-9.99). This also included females aged 60-74 years (CCI=1: OR 1.22, 95% CI 0.98-1,52; CCI=2-3: OR 2.28, 95% CI 1.74-2.99; CCI=4+: OR 4.22, 95% CI 2.56-6.97). Regarding the specific diseases considered in the study, the estimates suggested that except in the case of cerebrovascular diseases, patients affected by comorbidities had a significantly higher risk of being hospitalized. The results were similar when considering the entire population (Multimedia Appendix 1, Table S3).

The results related to admission to the ICU among those who tested positive with at least 1 swab are shown in Table 5. Unexpectedly, the association with multimorbidity was weak except for older women, while only a few associations were found with the specific diseases investigated. In contrast, compared to the general population, it emerged that patients with comorbidities and a higher CCI value had a higher risk of being admitted to intensive care for SARS-CoV-2 infection, especially in the case of respiratory diseases (Multimedia Appendix 1, Table S4).

Table 6 shows the OR estimates related to death within 30 days from the first positive swab among those who tested positive for SARS-CoV-2. Results suggested that as the CCI value increased, the risk of dying from COVID-19 significantly increased. When analyzing specific comorbidities, the estimates showed that the risk is significantly higher, especially among the younger population (eg, oncological diseases: OR 6.03, 95% CI 3.00-12.12 for males and OR 11.03, 95% CI 3.93-30.96 for females). Results were confirmed, considering the comparison with the general population (Multimedia Appendix 1, Table S5).

Table 7 and Multimedia Appendix 1, Figure S2, illustrate the outcomes of the multistate models. The probability of being hospitalized after testing positive for SARS-CoV-2 increased as the CCI increased, notably in the younger population, according to the hazard ratio (HR) estimates shown in the table. Furthermore, it appears that multimorbidity increased the risk of dying within 30 days of the first infection both without being hospitalized and after being hospitalized, with the exception of a small number of cases where some estimates were not statistically significant due to the low number of cases transiting among states.

Finally, stratifying the multistate models for the CCI revealed that both in the younger and the older population, the likelihood of COVID-19 severe outcomes considerably increased as the CCI score increased.

Table 1. Descriptive statistics related to the demographic and clinical characteristics of patients.

Variable and category	Entire population (N=1,918,549), n (%)	Access to the swab (n=85,348), n (%)	Positivity to SARS-Cov-2 (n=12,793), n (%)	Hospitalization (n=4644), n (%)	Admitted to the ICU[a] (n=1508), n (%)	Death within 30 days (n=749), n (%)
Gender						
Male	938,610 (48.9)	32,129 (37.6)	5961 (46.6)	3075 (66.2)	1118 (74.1)	558 (74.5)
Female	979,939 (51.1)	53,219 (62.4)	6832 (53.4)	1569 (33.8)	390 (25.9)	191 (25.5)
Age (years)						
45-59	1,062,861 (55.4)	54,672 (64.1)	7324 (57.2)	1717 (37.0)	503 (33.4)	120 (16.0)
60-74	855,688 (44.6)	30,676 (35.9)	5469 (42.8)	2927 (63.0)	1005 (66.6)	629 (84.0)
Month of first swab						
February-March	N/A[b]	14,134 (16.6)	5364 (41.9)	3088 (66.5)	1079 (71.5)	502 (67.0)
April-May	N/A	71,214 (83.4)	7429 (58.1)	1556 (33.5)	429 (28.5)	247 (33.0)
CCI[c]						
0	1,442,666 (75.2)	58,775 (68.8)	8,620 (67.4)	2,599 (56.0)	837 (55.5)	280 (37.4)
1	335,685 (17.5)	16,180 (19.0)	2540 (19.8)	1091 (23.5)	361 (23.9)	175 (23.3)
2-3	121,723 (6.3)	8099 (9.5)	1277 (10.0)	710 (15.3)	233 (15.5)	199 (26.6)
4+	18,475 (1.0)	2294 (2.7)	356 (2.8)	244 (5.2)	77 (5.1)	95 (12.7)
Oncological disease	58,742 (3.1)	3993 (4.7)	525 (4.1)	303 (6.5)	83 (5.5)	82 (11.0)
Cardiovascular disease	136,335 (7.1)	9239 (10.8)	1480 (11.6)	797 (17.2)	252 (16.7)	217 (29.0)
Respiratory disease	50,328 (2.6)	4792 (5.6)	751 (5.9)	426 (9.2)	124 (8.2)	129 (17.2)
Myocardial infarction	40,288 (2.1)	2416 (2.8)	420 (3.3)	278 (6.0)	99 (6.6)	78 (10.4)
Heart failure	7,817 (0.4)	762 (0.9)	117 (0.9)	79 (1.7)	22 (1.5)	28 (3.7)
Cerebrovascular disease	23,753 (1.2)	2197 (2.6)	378 (2.9)	186 (4.0)	53 (3.5)	72 (9.6)
Diabetes	157,214 (8.2)	8335 (9.8)	1547 (12.1)	896 (19.3)	352 (23.3)	239 (31.9)

[a] ICU: intensive care unit.

[b] N/A: not applicable.

[c] CCI: Charlson Comorbidity Index.

Table 2. OR[a] estimates related to access to the swab for the entire population (N=85,348), stratified by gender and age group (1 model for each variable).

Variable and category	Male		Female	
	Age 45-59 years (n=16,978, 19.9%), OR (95% CI)[b]	Age 60-74 years (n=15,151, 17.8%), OR (95% CI)[b]	Age 45-59 years (n=37,694, 44.2%), OR (95% CI)[b]	Age 60-74 years (n=15,525, 18.1%), OR (95% CI)[b]
CCI[c]				
0	Reference	Reference	Reference	Reference
1	1.37 (1.32-1.43)	1.31 (1.26-1.37)	1.22 (1.19-1.26)	1.39 (1.34-1.45)
2-3	2.41 (2.26-2.56)	2.31 (2.21-2.42)	1.37 (1.31-1.45)	2.19 (2.09-2.30)
4+	4.41 (3.93-4.95)	4.93 (4.62-5.26)	2.22 (1.91-2.57)	5.18 (4.73-5.68)
Oncological disease	2.31 (2.10-2.53)	2.04 (1.93-2.15)	1.18 (1.10-1.27)	1.80 (1.69-1.92)
Cardiovascular disease	1.94 (1.84-2.05)	2.11 (2.03-2.19)	1.45 (1.37-1.52)	2.19 (2.09-2.29)
Respiratory disease	2.41 (2.24-2.59)	3.20 (3.03-3.37)	1.68 (1.57-1.79)	3.34 (3.13-3.55)
Myocardial infarction	1.57 (1.42-1.74)	1.71 (1.62-1.81)	1.53 (1.32-1.77)	1.94 (1.76-2.14)
Heart failure	2.45 (2.00-3.00)	3.07 (2.78-3.40)	1.71 (1.27-2.29)	3.28 (2.81-3.83)
Cerebrovascular disease	2.60 (2.30-2.94)	2.71 (2.53-2.90)	1.63 (1.45-1.85)	3.31 (3.04-3.60)
Diabetes	1.46 (1.37-1.55)	1.46 (1.40-1.52)	1.21 (1.15-1.27)	1.60 (1.52-1.67)

[a] OR: odds ratio.

[b] Estimates adjusted for age. All estimates (95% confidence level) were significant.

[c] CCI: Charlson Comorbidity Index.

Table 3. OR[a] estimates related to positivity to SARS-CoV-2 among those who performed at least 1 swab (N=12,793), stratified by gender and age group (1 model for each variable).

Variable and category	Male		Female	
	Age 45-59 years (n=2847, 22.3%), OR (95% CI)[b]	Age 60-74 years (n=3114, 24.3%), OR (95% CI)[b]	Age 45-59 years (n=4477, 35.0%), OR (95% CI)[b]	Age 60-74 years (n=2355, 18.4%), OR (95% CI)[b]
CCI[c]				
0	Reference	Reference	Reference	Reference
1	0.97 (0.87-1.10)	1.07 (0.96-1.19)	1.01 (0.93-1.10)	1.05 (0.94-1.17)
2-3	0.79 (0.66-0.95)[d]	0.86 (0.77-0.97)[d]	0.98 (0.84-1.14)	0.93 (0.81-1.07)
4+	0.60 (0.42-0.86)[d]	0.79 (0.67-0.95)[d]	0.67 (0.41-1.09)	0.91 (0.71-1.17)
Oncological disease	0.72 (0.55-0.95)[d]	0.63 (0.53-0.73)[d]	0.82 (0.66-1.02)	0.70 (0.57-0.87)[d]
Cardiovascular disease	0.80 (0.68-0.94)[d]	0.91 (0.82-1.01)	0.89 (0.77-1.04)	1.02 (0.89-1.16)
Respiratory disease	0.89 (0.73-1.09)	0.88 (0.76-1.01)	0.82 (0.67-1.02)	1.05 (0.89-1.25)
Myocardial infarction	0.83 (0.62-1.12)	0.84 (0.72-0.98)	0.58 (0.34-0.99)[d]	1.10 (0.84-1.43)
Heart failure	0.46 (0.22-0.98)[d]	0.78 (0.59-1.04)	0.84 (0.33-2.14)	1.10 (0.74-1.65)
Cerebrovascular disease	0.95 (0.67-1.35)	1.18 (0.99-1.41)	0.83 (0.56-1.22)	1.12 (0.90-1.41)
Diabetes	1.02 (0.86-1.20)	1.25 (1.13-1.39)[d]	1.16 (1.01-1.34)[d]	1.11 (0.98-1.27)

[a] OR: odds ratio.

[b] Estimates adjusted for age and the month of the first swab.

[c] CCI: Charlson Comorbidity Index.

[d] Significant estimates (95% confidence level).

Table 4. OR[a] estimates related to admission to the hospital among those who tested positive for SARS-CoV-2 (N=4644), stratified by gender and age group (1 model for each variable).

Variable and category	Male		Female	
	Age 45-59 years (n=1101, 23.7%), OR (95% CI)[b]	Age 60-74 years (n=1974, 42.5%), OR (95% CI)[b]	Age 45-59 years (n=616, 13.3%), OR (95% CI)[b]	Age 60-74 years (n=953, 20.5%), OR (95% CI)[b]
CCI[c]				
0	Reference	Reference	Reference	Reference
1	1.50 (1.20-1.88)[d]	1.51 (1.23-1.84)[d]	1.50 (1.20-1.88)[d]	1.22 (0.98-1.52)
2-3	2.14 (1.51-3.01)[d]	1.40 (1.12-1.75)[d]	2.49 (1.77-3.50)[d]	2.28 (1.74-2.99)[d]
4+	4.77 (2.28-9.99)[d]	1.98 (1.41-2.78)[d]	2.45 (0.86-6.94)	4.22 (2.56-6.97)[d]
Oncological disease	2.62 (1.52-4.51)[d]	1.27 (0.93-1.73)	2.26 (1.39-3.68)[d]	1.97 (1.34-2.92)[d]
Cardiovascular disease	1.53 (1.13-2.08)[d]	1.30 (1.01-1.49)[d]	1.69 (1.17-2.44)[d]	1.52 (1.19-1.95)[d]
Respiratory disease	2.17 (1.48-3.19)[d]	1.47 (1.12-1.92)[d]	2.13 (1.31-3.45)[d]	2.16 (1.57-2.96)[d]
Myocardial infarction	1.84 (1.03-3.30)[d]	1.47 (1.09-1.97)[d]	2.05 (0.61-6.94)	2.36 (1.44-3.89)[d]
Heart failure	7.90 (1.53-40.70)[d]	2.18 (1.23-3.86)[d]	1.25 (0.12-12.56)	1.88 (0.90-3.93)
Cerebrovascular disease	1.39 (0.69-2.81)	0.92 (0.68-1.26)	1.30 (0.47-3.63)	1.11 (0.73-1.71)
Diabetes	1.65 (1.20-2.28)[d]	1.61 (1.31-1.96)[d]	1.89 (1.35-2.63)[d]	2.10 (1.64-2.69)[d]

[a] OR: odds ratio.

[b] Estimates adjusted for age and the month of the first swab.

[c] CCI: Charlson Comorbidity Index.

[d] Significant estimates (95% confidence level).

Table 5. OR[a] estimates related to admission to the ICU[b] among those who tested positive for SARS-CoV-2 (N=1508), stratified by gender and age group (1 model for each variable).

Variable and category	Male		Female	
	Age 45-59 years (n=370, 24.5%), OR (95% CI)[c]	Age 60-74 years (n=748, 49.6%), OR (95% CI)[c]	Age 45-59 years (n=133, 8.8%), OR (95% CI)[c]	Age 60-74 years (n=257, 17.1%), OR (95% CI)[c]
CCI[d]				
0	Reference	Reference	Reference	Reference
1	1.39 (1.04-1.86)[e]	1.15 (0.94-1.42)	1.06 (0.66-1.70)	1.25 (0.91-1.73)
2-3	1.32 (0.85-2.05)	1.04 (0.82-1.32)	2.46 (1.39-4.36)[e]	1.89 (1.32-2.71)[e]
4+	1.91 (0.83-4.39)	1.09 (0.77-1.54)	1.33 (0.17-10.35)	2.35 (1.32-4.19)[e]
Oncological disease	1.44 (0.74-2.80)	0.86 (0.62-1.20)	0.91 (0.28-2.94)	0.97 (0.55-1.71)
Cardiovascular disease	1.58 (1.09-2.30)[e]	0.84 (0.68-1.03)	2.02 (1.08-3.76)[e]	1.23 (0.87-1.75)
Respiratory disease	1.53 (0.96-2.45)	0.84 (0.63-1.14)	1.92 (0.81-4.51)	1.53 (1.01-2.32)[e]
Myocardial infarction	2.05 (1.08-3.86)[e]	1.04 (0.77-1.40)	2.00 (0.25-15.85)	1.75 (0.95-3.22)
Heart failure	N/A[f]	1.16 (0.66-2.06)	N/A	1.39 (0.52-3.69)
Cerebrovascular disease	1.39 (0.55-3.50)	0.74 (0.50-1.08)	1.08 (0.14-8.21)	0.94 (0.49-1.80)
Diabetes	1.43 (0.98-2.11)	1.58 (1.31-1.92)[e]	2.58 (1.51-4.42)[e]	2.14 (1.57-2.92)[e]

[a] OR: odds ratio.

[b] ICU: intensive care unit.

[c] Estimates adjusted for age and the month of the first swab.

[d] CCI: Charlson Comorbidity Index.

[e] Significant estimates (95% confidence level).

[f] N/A: not applicable.

Table 6. OR[a] estimates related to death within 30 days from the first positive swab among those who tested positive to SARS-CoV-2 (N=749), stratified by gender and age group (1 model for each variable).

Variable and category	Male		Female	
	Age 45-59 years (n=370, 24.5%), OR (95% CI)[b]	Age 60-74 years (n=748, 49.6%), OR (95% CI)[b]	Age 45-59 years (n=133, 8.8%), OR (95% CI)[b]	Age 60-74 years (n=257, 17.1%), OR (95% CI)[b]
CCI[c]				
0	Reference	Reference	Reference	Reference
1	1.56 (0.89-2.75)	1.33 (1.02-1.74)[d]	2.26 (0.68-7.56)	1.50 (0.99-2.27)
2-3	5.33 (3.06-9.27)[d]	2.24 (1.72-2.91)[d]	16.80 (6.35-44.45)[d]	2.43 (1.58-3.73)[d]
4+	17.32 (7.91-37.90)[d]	3.26 (2.29-4.64)[d]	31.88 (5.96-170.43)[d]	5.90 (3.31-10.52)[d]
Oncological disease	6.03 (3.00-12.12)[d]	1.46 (1.04-2.05)[d]	11.03 (3.93-30.96)[d]	1.38 (0.77-2.47)
Cardiovascular disease	3.52 (2.12-5.86)[d]	1.56 (1.24-1.95)[d]	2.81 (0.82-9.62)	2.04(1.41-2.94)[d]
Respiratory disease	3.89 (2.16-7.02)[d]	2.07 (1.54-2.78)[d]	5.88 (1.70-20.39)[d]	2.92 (1.91-4.45)[d]
Myocardial infarction	2.80 (1.15-6.81)[d]	1.50 (1.09-2.06)[d]	N/A[e]	2.16 (1.41-2.94)[d]
Heart failure	8.42 (1.66-42.77)[d]	1.79 (0.99-3.23)	N/A	6.04 (2.70-13.53)[d]
Cerebrovascular disease	5.68 (2.24-14.36)[d]	1.96 (1.38-2.79)[d]	N/A	2.06 (1.17-3.65)[d]
Diabetes	2.36 (1.34-4.15)[d]	1.79 (1.43-2.23)[d]	5.72 (2.21-14.79)[d]	2.35 (1.64-3.35)[d]

[a] OR: odds ratio.

[b] Estimates adjusted for age and the month of the first swab.

[c] CCI: Charlson Comorbidity Index.

[d] Significant estimates (95% confidence level).

[e] N/A: not applicable.

Table 7. HR[a] estimates related to multistate models for each possible transition, stratified by gender and age group.

CCI[b]	Male (age 45-59 years), HR (95% CI)[c]	Male (age 60-74 years), HR (95% CI)[c]	Female (age 45-59 years), HR (95% CI)[c]	Female (age 60-74 years), HR (95% CI)[c]
Positive --> hospitalization				
0	Reference	Reference	Reference	Reference
1	1.34 (1.15-1.56)	1.19 (1.06-1.33)	1.43 (1.17-1.75)	1.19 (1.01-1.39)
2-3	1.48 (1.18-1.86)	1.09 (0.96-1.23)	2.49 (1.90-3.26)	1.61 (1.34-1.93)
4+	2.71 (1.85-3.97)	1.25 (1.04-1.49)	2.42 (1.09.5.33)	2.39 (1.81-3.16)
Positive --> death within 30 days from first infection				
0	Reference	Reference	Reference	Reference
1	3.87 (0.01-infinity)	0.73 (0.29-1.87)	N/A[d]	2.67 (1.05-6.80)
2-3	1.20 (0.01-infinity)	2.10 (1.03-4.28)	22.57 (2.32-219.69)	4.13 (1.53-11.12)
4+	220750.22 (0.01-infinity)	3.33 (1.37-8.11)	N/A	7.69 (2.00-29.52)
Hospitalization --> death within 30 days from first infection				
0	Reference	Reference	Reference	Reference
1	1.42 (0.30-6.61)	1.22 (0.94-1.59)	8.29 (2.81-24.38)	1.40 (0.89-2.21)
2-3	5.10 (1.34-19.44)	1.92 (1.49-2.49)	0.03 (0.01-37674.39)	1.61 (0.99-2.61)
4+	728.13 (355.06-1493.24)	2.38 (1.71-3.31)	0.41 (0.01-93403.78)	3.46 (1.98-6.04)

[a] HR: hazard ratio.

[b] CCI: Charlson Comorbidity Index.

[c] Estimated adjusted for age and the month of the first swab.

[d] N/A: not applicable.

Discussion

Principal Findings

In this study, we analyzed the association between multimorbidity and SARS-CoV-2 outcomes in the population of the large Italian region of Piedmont. It emerged that multimorbidity is a strong risk factor for a worse prognosis of COVID-19, especially in the younger population. In addition, results highlighted that although subjects with previous diseases were more likely to be swabbed, they had a general lower risk of being infected.

Regarding access to the swab during the first wave of the COVID-19 pandemic, the estimates showed that the likelihood of being swabbed was greater for patients with previous diseases, regardless of the kind of disease, and this risk increased with the increase in multimorbidity (measured as the CCI). This is consistent with what was observed during the outbreak of the COVID-19 pandemic in early 2020, when only few laboratories were equipped for performing the SARS-CoV-2 test from nasal swabs and the Italian government published strict clinical and epidemiological criteria for accessing the tests (among those, subjects with a chronic disease are considered at higher risk), which were also limited in availability [20]. In addition, from April 2020, patients with chronic diseases were swabbed to access outpatient services, such as dialysis or cancer treatments.

Focusing on positivity to SARS-CoV-2, when considering the entire population, it emerged that the risk of infection is higher for subjects with previous comorbidities, in line with other studies [21,22]. This occurs because subjects affected by comorbidities, due to poor clinical conditions, generally perform more swabs compared to healthy subjects; thus, their probability of testing positive for COVID-19 is higher. In contrast, when investigating the likelihood of testing positive only among those who performed at least 1 swab, we identified that patients with comorbidities are less likely to test positive for SARS-CoV-2. This result could be due to the fact that subjects with poorer clinical conditions pay more attention to protective methods, such as social distancing, wearing masks, handwashing, and avoiding overcrowded places, which significantly reduce the risk of SARS-CoV-2 infection [23].

In our investigation of hospitalization, we also found that multimorbidity, which in our study was measured through the CCI, is a strong risk factor for a worse prognosis of COVID-19. This is consistent with a number of studies conducted on the topic [12,13,24].

Among all possible comorbidities, Chudasama et al [24] identified in their study that pre-existing hypertension is the most prevalent condition in subjects affected by severe SARS-CoV-2 infection, and it mainly coexists with other previous comorbidities: stroke, diabetes, and chronic kidney disease. However, the risk of severe COVID-19 is highest in patients affected by both previous diabetes and pre-existing chronic kidney disease. One of the possible reasons patients with multiple comorbidities have an increased risk of developing severe SARS-CoV-2 is that they generally use inhibitors of the renin-angiotensin system (RAS) to limit the effect of their comorbidities. These inhibitors cause the overexpression of angiotensin-converting enzyme 2 (ACE2), which in turn facilitates the entry of SARS-CoV-2 into human target cells. In addition, high levels of some biomarkers, such as C-reactive protein, D-dimer, procalcitonin, and ferritin, in individuals with multiple comorbidities may lead or contribute to a worse prognosis of COVID-19 [13]. In fact, it has been found that these biomarkers are often elevated in subjects who contract severe infection.

In our study, we also identified that the risk of being admitted to the ICU among those who tested positive to the virus was not significantly higher for subjects affected by multimorbidity, except in a few cases, such as cardiovascular diseases and diabetes. Given the findings that emerged from our research in relation to hospital admissions and death, this result would seem to be counterintuitive. This finding may be attributable to the decisions made by health care professionals during the first pandemic wave about which patients to admit to intensive care and which not to admit. In fact, health care professionals had to choose which subjects to admit to the ICU on the basis of their clinical conditions, the number and severity of comorbidities, age, and possible benefits of admission, also due in the peak weeks of the pandemic to the limited number of beds available. Only subjects with the highest clinical outcomes and potential benefits were therefore admitted to the ICU. The mortality data support this explanation; in fact, as expected, the probability of dying from COVID-19 increased dramatically as the CCI value increased.

Considering specific comorbidities, we found that subjects with almost all previous comorbidities have a greater risk of developing a worse prognosis of COVID-19, considering both the general population and only patients with SARS-CoV-2. This is in line with other studies that have highlighted this topic. According to some research that investigated the relationship between oncological diseases and SARS-CoV-2–related outcomes [25-27], people with cancer have a greater risk of developing a worse prognosis of COVID-19. This finding could be due to the fact that patients with cancer are particularly susceptible to the immunosuppressive state caused by antitumor therapies received, such as radiotherapy and chemotherapy [26]. In addition, the risk of COVID-19 severity has been found to be higher for individuals who received their last chemotherapy within 14 days of admission [28].

Further studies that have investigated the association between cardiovascular or cerebrovascular diseases and COVID-19 outcomes found that the risk of severe SARS-CoV-2 is significantly higher for subjects affected by these comorbidities [8,29]. Furthermore, according to a meta-analysis conducted on 56 studies [30], it was shown that the risk of developing severe COVID-19 is greater, considering patients with SARS-CoV-2 and any pre-existing cardiovascular diseases and also when considering specific cardiovascular comorbidities separately, such as acute cardiac injury or heart failure, as shown in our study. This association could be due to the fact that drugs used to limit cardiovascular and cerebrovascular risk, such as ACE inhibitors and angiotensin II receptor blockers (ARBs), have numerous effects that could influence the susceptibility to or the severity of COVID-19. In fact, it was demonstrated that ACE inhibitors and ARBs increase the expression of ACE2, which is the viral receptor for SARS-COV-2 and facilitates the virus entry into pneumocytes [29,31].

In relation to diabetes, several studies have also shown that it is a risk factor for the mortality and severity of COVID-19. In fact, according to different studies and meta-analyses [32-35], the risk of contracting severe SARS-CoV-2 or dying from the virus infection has been found to be significantly higher for patients with diabetes. One possible reason for this is that subjects with diabetes have a higher risk of uncontrolled inflammatory response, higher levels of tissue injury–related enzymes, a higher hypercoagulable state, and higher serum levels of inflammatory biomarkers, such as C-reactive protein, D-dimer, interleukin-6 (IL-6), serum ferritin, and coagulation index. This greater susceptibility to an inflammatory status could lead to a worse prognosis of COVID-19, especially in patients with poor glycemic control, since hyperglycemia is a powerful antagonist of the immune response [32,36]. In addition, these subjects also have an immune system downregulated by impairing the function of innate immunity, such as chemotaxis and the activity of neutrophils and macrophages, that could lead to severe COVID-19 outcomes or mortality [35].

Other studies have instead found a relationship between pre-existing respiratory comorbidities and the risk of developing a worse prognosis of COVID-19, especially in the case of COPD, asthma, and obstructive sleep apnea (OSA) [37-39]. This association, on the one hand, is due to the fact that previous respiratory diseases could worsen lung function, could make the airways hypersensitive, and could cause immune alteration in the patients, possibly leading to subjects contracting more severe SARS-CoV-2 [38]. On the other hand, especially in the case of pre-existing OSA, hypercapnia and hypoxemia, surges of sympathetic activation, and increased inflammatory markers may contribute to contracting more severe SARS-CoV-2 [39].

Although in our study, we only investigated diabetes and oncological, cardiovascular, cerebrovascular, and respiratory diseases, previous research has shown that other comorbidities, such as chronic liver and chronic kidney diseases, are also associated with a more severe prognosis of COVID-19 [40,41].

Strengths and Limitations

This study represents an advance on what is already present in the literature. In fact, compared to what has been investigated on this topic to date, in this research, (1) a population study was conducted instead of a clinical study, which made it possible both to have a much larger number of people available and to investigate what happened in an entire region and not only in hospitals or health care institutions, whose studies are generally conducted in more advanced facilities (eg, university hospitals); (2) the period analyzed in the study (ie, the first wave of the COVID-19 pandemic) made it possible to obtain results that are not influenced by the various organized prevention/vaccination strategies implemented subsequently;

and (3) the assessment of the probability to be tested could be a bias in susceptibility evaluation.

The main limitation of this study is that data derived from record linkage of health administrative databases, where information bias (although not differential) is present, and depth of information were limited. A further limitation is that data were based only on the first wave of the pandemic. Moreover, the administrative nature of the sample was not able to capture the social elements that could be used to fully conceive a syndemic approach; however, the obtained results identified a disease-disease interaction that could be the basis for further research in this framework.

Conclusion

In a sample of nearly 2 million subjects, our study is 1 of the first to assess the association between multimorbidity and all SARS-CoV-2–related outcomes. Our findings show that during the first wave of the pandemic, patients with multimorbidity were closely monitored, as proven by a high frequency of tests for SARS-CoV-2. However, although swabbed more frequently, they appeared to be less likely to become infected with SARS-CoV-2, probably due to greater attention paid to protective methods. However, a history of respiratory diseases is a risk factor for a worse prognosis of COVID-19. Nonetheless, whatever comorbidities affect the patients, a strong dose-response effect was observed between an increased CCI score and COVID-19 hospitalization, ICU admission, and death.

These results are critical to public health policy and planning as they help in identifying a group of subjects who are more prone to worse SARS-CoV-2 outcomes. This information is particularly important for the current pandemic scenario, where the emergency has given way to the SARS-CoV-2 pandemic's embeddedness into daily life. In fact, these results suggest that future clinical and public health interventions (eg, vaccination prioritization, early monoclonal antibody treatment, prevention measurement and campaigns) should be centered on the multimorbid patient category because they are more likely to need to be protected from COVID-19 severe outcomes. Furthermore, the early response to the COVID-19 pandemic provided a framework for our observations that might be applicable to future health challenges. It will be crucial for future research to investigate the biosocial relationship in this context.

Acknowledgments

Contributors AC and FR conceptualized the study, were responsible for data curation and investigation, and conducted the formal data analysis. AC, WG, RG, GC, CS, and FR oversaw the project. AC, LD, AM, LM, RG, ES, NZ, and FR accessed and verified the data. AC, WG, CS, and FR wrote the initial draft of the manuscript. LD, AM, LM, RG, ES, NZ, VC, GC, and CS reviewed and edited the manuscript. The corresponding author had full access to all the data in the study and had the final responsibility for the decision to submit for publication.

Moreover, the authors thank the Consorzio per il Sistema Informativo del Piemonte for its contribution in prompt data gathering and data management.

Data Availability

Raw data cannot be made freely available because of restrictions imposed by the ethical committees, which do not allow open/public sharing of data on individuals. However, aggregated data are available for other researchers upon request to the corresponding author.

Conflicts of Interest

The authors have no financial or nonfinancial interests to disclose. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Multimedia Appendix 1
Supplementary material.
[DOCX File , 259 KB - publichealth_v9i1e41404_app1.docx]

References

1. Williams J, Degeling C, McVernon J, Dawson A. How should we conduct pandemic vaccination? Vaccine 2021 Mar 05;39(6):994-999 [FREE Full text] [doi: 10.1016/j.vaccine.2020.12.059] [Medline: 33423839]
2. Forman R, Shah S, Jeurissen P, Jit M, Mossialos E. COVID-19 vaccine challenges: what have we learned so far and what remains to be done? Health Policy 2021 May;125(5):553-567 [FREE Full text] [doi: 10.1016/j.healthpol.2021.03.013] [Medline: 33820678]
3. Chilimuri S, Sun H, Alemam A, Mantri N, Shehi E, Tejada J, et al. Predictors of mortality in adults admitted with covid-19: retrospective cohort study from New York City. West J Emerg Med 2020 Jul 08;21(4):779-784 [FREE Full text] [doi: 10.5811/westjem.2020.6.47919] [Medline: 32726241]

4. Cummings MJ, Baldwin MR, Abrams D, Jacobson SD, Meyer BJ, Balough EM, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. Lancet 2020 Jun 06;395(10239):1763-1770 [FREE Full text] [doi: 10.1016/S0140-6736(20)31189-2] [Medline: 32442528]
5. España PP, Bilbao A, García-Gutiérrez S, Lafuente I, Anton-Ladislao A, Villanueva A, COVID-19-Osakidetza Working Group. Predictors of mortality of COVID-19 in the general population and nursing homes. Intern Emerg Med 2021 Oct 05;16(6):1487-1496 [FREE Full text] [doi: 10.1007/s11739-020-02594-8] [Medline: 33400164]
6. Ng WH, Tipih T, Makoah NA, Vermeulen J, Goedhals D, Sempa JB, et al. Comorbidities in SARS-CoV-2 patients: a systematic review and meta-analysis. mBio 2021 Mar 09;12(1):e03647-20 [FREE Full text] [doi: 10.1128/mBio.03647-20] [Medline: 33563817]
7. Gallo A, Pero E, Pellegrino S, Macerola N, Murace CA, Ibba F, et al. How can biology of aging explain the severity of covid-19 in older adults. Clin Geriatr Med 2022 Aug;38(3):461-472 [FREE Full text] [doi: 10.1016/j.cger.2022.04.002] [Medline: 35868666]
8. Li B, Yang J, Zhao F, Zhi L, Wang X, Liu L, et al. Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. Clin Res Cardiol 2020 May;109(5):531-538 [FREE Full text] [doi: 10.1007/s00392-020-01626-9] [Medline: 32161990]
9. Guan W, Liang W, Zhao Y, Liang H, Chen Z, Li Y, China Medical Treatment Expert Group for COVID-19. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. Eur Respir J 2020 May;55(5):2000547 [FREE Full text] [doi: 10.1183/13993003.00547-2020] [Medline: 32217650]
10. Scheen A, Marre M, Thivolet C. Prognostic factors in patients with diabetes hospitalized for COVID-19: findings from the CORONADO study and other recent reports. Diabetes Metab 2020 Oct;46(4):265-271 [FREE Full text] [doi: 10.1016/j.diabet.2020.05.008] [Medline: 32447101]
11. Charlson ME, Pompei P, Ales KL, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987 Jan;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8]
12. Christensen DM, Strange JE, Gislason G, Torp-Pedersen C, Gerds T, Fosbøl E, et al. Charlson Comorbidity Index score and risk of severe outcome and death in Danish covid-19 patients. J Gen Intern Med 2020 Oct;35(9):2801-2803 [FREE Full text] [doi: 10.1007/s11606-020-05991-z] [Medline: 32583345]
13. Tuty Kuswardhani RA, Henrina J, Pranata R, Anthonius Lim M, Lawrensia S, Suastika K. Charlson Comorbidity Index and a composite of poor outcomes in COVID-19 patients: a systematic review and meta-analysis. Diabetes Metab Syndr 2020;14(6):2103-2109 [FREE Full text] [doi: 10.1016/j.dsx.2020.10.022] [Medline: 33161221]
14. Horton R. Offline: COVID-19 is not a pandemic. Lancet 2020 Sep 26;396(10255):874 [FREE Full text] [doi: 10.1016/S0140-6736(20)32000-6] [Medline: 32979964]
15. Rod MH, Hulvej Rod N. Towards a syndemic public health response to COVID-19. Scand J Public Health 2021 Mar;49(1):14-16. [doi: 10.1177/1403494820982862] [Medline: 33407002]
16. Jakovljevic M, Samarzija M, Milicic D, Reiner Z, Sartorius N. Comorbidities and syndemics in the covid-19 age: challenges and opportunities for bringing separated branches of medicine closer to each other. Psychiatr Danub 2021;33(Suppl 4):402-413 [FREE Full text] [Medline: 34718258]
17. Courtin E, Vineis P. COVID-19 as a syndemic. Front Public Health 2021;9:763830 [FREE Full text] [doi: 10.3389/fpubh.2021.763830] [Medline: 34568273]
18. Fronteira I, Sidat M, Magalhães JP, de Barros FPC, Delgado AP, Correia T, et al. The SARS-CoV-2 pandemic: a syndemic perspective. One Health 2021 Jul;12:100228 [FREE Full text] [doi: 10.1016/j.onehlt.2021.100228] [Medline: 33614885]
19. Gravlee CC. Systemic racism, chronic health inequities, and COVID-19: a syndemic in the making? Am J Hum Biol 2020 Oct;32(5):e23482 [FREE Full text] [doi: 10.1002/ajhb.23482] [Medline: 32754945]
20. Trevisan C, Pedone C, Maggi S, Noale M, Di Bari M, Sojic A, Epicovid Working Group. Accessibility to SARS-CoV-2 swab test during the covid-19 pandemic: did age make the difference? Health Policy 2021 Dec;125(12):1580-1586 [FREE Full text] [doi: 10.1016/j.healthpol.2021.10.002] [Medline: 34649753]
21. McQueenie R, Foster HME, Jani BD, Katikireddi SV, Sattar N, Pell JP, et al. Multimorbidity, polypharmacy, and COVID-19 infection within the UK Biobank cohort. PLoS One 2020;15(8):e0238091 [FREE Full text] [doi: 10.1371/journal.pone.0238091] [Medline: 32817712]
22. de Lusignan S, Dorward J, Correa A, Jones N, Akinyemi O, Amirthalingam G, et al. Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study. Lancet Infect Dis 2020 Sep;20(9):1034-1042 [FREE Full text] [doi: 10.1016/S1473-3099(20)30371-6] [Medline: 32422204]
23. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 01;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
24. Chudasama YV, Zaccardi F, Gillies CL, Razieh C, Yates T, Kloecker DE, et al. Patterns of multimorbidity and risk of severe SARS-CoV-2 infection: an observational study in the U.K. BMC Infect Dis 2021 Oct 04;21(1):908 [FREE Full text] [doi: 10.1186/s12879-021-06600-y] [Medline: 34481456]

25. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. Lancet Oncol 2020 Mar;21(3):335-337 [FREE Full text] [doi: 10.1016/S1470-2045(20)30096-6] [Medline: 32066541]
26. Zhang L, Zhu F, Xie L, Wang C, Wang J, Chen R, et al. Clinical characteristics of COVID-19-infected cancer patients: a retrospective case study in three hospitals within Wuhan, China. Ann Oncol 2020 Jul;31(7):894-901 [FREE Full text] [doi: 10.1016/j.annonc.2020.03.296] [Medline: 32224151]
27. Roel E, Pistillo A, Recalde M, Fernández-Bertolín S, Aragón M, Soerjomataram I, et al. Cancer and the risk of coronavirus disease 2019 diagnosis, hospitalisation and death: a population-based multistate cohort study including 4 618 377 adults in Catalonia, Spain. Int J Cancer 2022 Mar 01;150(5):782-794 [FREE Full text] [doi: 10.1002/ijc.33846] [Medline: 34655476]
28. Tian J, Yuan X, Xiao J, Zhong Q, Yang C, Liu B, et al. Clinical characteristics and risk factors associated with COVID-19 disease severity in patients with cancer in Wuhan, China: a multicentre, retrospective, cohort study. Lancet Oncol 2020 Jul;21(7):893-903 [FREE Full text] [doi: 10.1016/S1470-2045(20)30309-0] [Medline: 32479790]
29. Pranata R, Huang I, Lim MA, Wahjoepramono EJ, July J. Impact of cerebrovascular and cardiovascular diseases on mortality and severity of COVID-19-systematic review, meta-analysis, and meta-regression. J Stroke Cerebrovasc Dis 2020 Aug;29(8):104949. [doi: 10.1016/j.jstrokecerebrovasdis.2020.104949] [Medline: 32927523]
30. Hessami A, Shamshirian A, Heydari K, Pourali F, Alizadeh-Navaei R, Moosazadeh M, et al. Cardiovascular diseases burden in COVID-19: systematic review and meta-analysis. Am J Emerg Med 2021 Aug;46:382-391 [FREE Full text] [doi: 10.1016/j.ajem.2020.10.022] [Medline: 33268238]
31. Nishiga M, Wang DW, Han Y, Lewis DB, Wu JC. COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. Nat Rev Cardiol 2020 Oct;17(9):543-558 [FREE Full text] [doi: 10.1038/s41569-020-0413-9] [Medline: 32690910]
32. Kumar A, Arora A, Sharma P, Anikhindi SA, Bansal N, Singla V, et al. Is diabetes mellitus associated with mortality and severity of COVID-19? A meta-analysis. Diabetes Metab Syndr 2020;14(4):535-545 [FREE Full text] [doi: 10.1016/j.dsx.2020.04.044] [Medline: 32408118]
33. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. Diabetes Metab Res Rev 2020 Mar 31;36(7):e3319 [FREE Full text] [doi: 10.1002/dmrr.3319] [Medline: 32233013]
34. Hartmann-Boyce J, Rees K, Perring JC, Kerneis SA, Morris EM, Goyder C, et al. Risks of and from SARS-CoV-2 infection and covid-19 in people with diabetes: a systematic review of reviews. Diabetes Care 2021 Dec;44(12):2790-2811 [FREE Full text] [doi: 10.2337/dc21-0930] [Medline: 34711637]
35. Wu J, Zhang J, Sun X, Wang L, Xu Y, Zhang Y, et al. Influence of diabetes mellitus on the severity and fatality of SARS-CoV-2 (COVID-19) infection. Diabetes Obes Metab 2020 Oct;22(10):1907-1914 [FREE Full text] [doi: 10.1111/dom.14105] [Medline: 32496012]
36. Giorda CB, Picariello R, Tartaglino B, Nada E, Doglio M, Romeo F, et al. From swab testing to health outcomes within the T2DM population: impact of diabetes background on COVID19 progression. Diabetes Res Clin Pract 2021 Oct;180:109021 [FREE Full text] [doi: 10.1016/j.diabres.2021.109021] [Medline: 34437941]
37. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, ISARIC4C Investigators. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ 2020 May 22;369:m1985 [FREE Full text] [doi: 10.1136/bmj.m1985] [Medline: 32444460]
38. Huang BZ, Chen Z, Sidell MA, Eckel SP, Martinez MP, Lurmann F, et al. Asthma disease status, COPD, and covid-19 severity in a large multiethnic population. J Allergy Clin Immunol Pract 2021 Oct;9(10):3621-3628.e2 [FREE Full text] [doi: 10.1016/j.jaip.2021.07.030] [Medline: 34389242]
39. Lohia P, Sreeram K, Nguyen P, Choudhary A, Khicher S, Yarandi H, et al. Preexisting respiratory diseases and clinical outcomes in COVID-19: a multihospital cohort study on predominantly African American population. Respir Res 2021 Mar 05;22(1):37 [FREE Full text] [doi: 10.1186/s12931-021-01647-6] [Medline: 33546658]
40. Gao Y, Ding M, Dong X, Zhang J, Kursat Azkur A, Azkur D, et al. Risk factors for severe and critically ill COVID-19 patients: a review. Allergy 2021 Mar;76(2):428-455. [doi: 10.1111/all.14657] [Medline: 33185910]
41. Oyelade T, Alqahtani J, Canciani G. Prognosis of COVID-19 in patients with liver and kidney diseases: an early systematic review and meta-analysis. Trop Med Infect Dis 2020 May 15;5(2):E80 [FREE Full text] [doi: 10.3390/tropicalmed5020080] [Medline: 32429038]

Abbreviations

ACE: angiotensin-converting enzyme
ARB: angiotensin II receptor blocker
CCI: Charlson Comorbidity Index
COPD: chronic obstructive pulmonary disease
HR: hazard ratio
ICU: intensive care unit

OR: odds ratio
OSA: obstructive sleep apnea

Edited by T Sanchez, A Mavragani; submitted 25.07.22; peer-reviewed by V Possenti, V Simeon, S Tang; comments to author 03.11.22; revised version received 24.11.22; accepted 22.12.22; published 09.02.23.

Please cite as:
Catalano A, Dansero L, Gilcrease W, Macciotta A, Saugo C, Manfredi L, Gnavi R, Strippoli E, Zengarini N, Caramello V, Costa G, Sacerdote C, Ricceri F
Multimorbidity and SARS-CoV-2–Related Outcomes: Analysis of a Cohort of Italian Patients
JMIR Public Health Surveill 2023;9:e41404
URL: https://publichealth.jmir.org/2023/1/e41404
doi:10.2196/41404
PMID:36626821

Original Paper

Prediction of Suicidal Behaviors in the Middle-aged Population: Machine Learning Analyses of UK Biobank

Junren Wang[1,2*], MPH; Jiajun Qiu[1,2*], PhD; Ting Zhu[1,2], PhD; Yu Zeng[1,2], MS; Huazhen Yang[1,2], MS; Yanan Shang[1,2], PhD; Jin Yin[1,2], MS; Yajing Sun[1,2], MS; Yuanyuan Qu[1,2], MA; Unnur A Valdimarsdóttir[3,4,5], PhD; Huan Song[1,2,3], MD, PHD

[1]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China
[2]Med-X Center for Informatics, Sichuan University, Chengdu, China
[3]Center of Public Health Sciences, Faculty of Medicine, University of Iceland, Reykjavík, Iceland
[4]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[5]Department of Epidemiology, Harvard T H Chan School of Public Health, Harvard University, Boston, MA, United States
[*]these authors contributed equally

Corresponding Author:
Huan Song, MD, PHD
West China Biomedical Big Data Center
West China Hospital
Sichuan University
Guo Xue Lane 37
Chengdu, 610021
China
Phone: 86 28 85164176
Email: songhuan@wchscu.cn

Abstract

Background: Suicidal behaviors, including suicide deaths and attempts, are major public health concerns. However, previous suicide models required a huge amount of input features, resulting in limited applicability in clinical practice.

Objective: We aimed to construct applicable models (ie, with limited features) for short- and long-term suicidal behavior prediction. We further validated these models among individuals with different genetic risks of suicide.

Methods: Based on the prospective cohort of UK Biobank, we included 223 (0.06%) eligible cases of suicide attempts or deaths, according to hospital inpatient or death register data within 1 year from baseline and randomly selected 4460 (1.18%) controls (1:20) without such records. We similarly identified 833 (0.22%) cases of suicidal behaviors 1 to 6 years from baseline and 16,660 (4.42%) corresponding controls. Based on 143 input features, mainly including sociodemographic, environmental, and psychosocial factors; medical history; and polygenic risk scores (PRS) for suicidality, we applied a bagged balanced light gradient-boosting machine (LightGBM) with stratified 10-fold cross-validation and grid-search to construct the full prediction models for suicide attempts or deaths within 1 year or between 1 and 6 years. The Shapley Additive Explanations (SHAP) approach was used to quantify the importance of input features, and the top 20 features with the highest SHAP values were selected to train the applicable models. The external validity of the established models was assessed among 50,310 individuals who participated in UK Biobank repeated assessments both overall and by the level of PRS for suicidality.

Results: Individuals with suicidal behaviors were on average 56 years old, with equal sex distribution. The application of these full models in the external validation data set demonstrated good model performance, with the area under the receiver operating characteristic (AUROC) curves of 0.919 and 0.892 within 1 year and between 1 and 6 years, respectively. Importantly, the applicable models with the top 20 most important features showed comparable external-validated performance (AUROC curves of 0.901 and 0.885) as the full models, based on which we found that individuals in the top quintile of predicted risk accounted for 91.7% (n=11) and 80.7% (n=25) of all suicidality cases within 1 year and during 1 to 6 years, respectively. We further obtained comparable prediction accuracy when applying these models to subpopulations with different genetic susceptibilities to suicidality. For example, for the 1-year risk prediction, the AUROC curves were 0.907 and 0.885 for the high (>2nd tertile of PRS) and low (<1st) genetic susceptibilities groups, respectively.

Conclusions: We established applicable machine learning–based models for predicting both the short- and long-term risk of suicidality with high accuracy across populations of varying genetic risk for suicide, highlighting a cost-effective method of identifying individuals with a high risk of suicidality.

(JMIR Public Health Surveill 2023;9:e43419) doi:10.2196/43419

KEYWORDS

suicide; suicidal behaviors; risk prediction; machine learning approach; genetic susceptibility; machine learning; behavior; data; model; sex; risk; cost-effective

Introduction

According to the estimation of the Global Burden of Disease study, approximately 800,000 people die by suicide every year [1], which translates to the astonishing number of 1 person dying by suicide every 40 seconds. In the United Kingdom, there were 5691 deaths by suicide registered in England and Wales in 2019, which corresponded to an age-standardized rate of 11 deaths per 100,000 people [2]. Importantly, behind the number of suicidal deaths, there is a much higher incidence of suicide attempts requiring further research. From 2000 to 2010, a prospective study using data from 5 emergency departments in the United Kingdom identified 38,415 individuals who presented at an emergency department following a suicide attempt [3], among which only 261 (0.7%) died. This finding implies that the population targeted for suicide prevention, such as timely psychological support, is considerably larger. However, only 28% of people who attempt suicide in the United Kingdom have previously received psychiatric services [4]. Therefore, it is urgent to improve the identification of individuals at high risk for suicidality to improve suicide prevention.

The previous research suggests that the mechanisms of suicidality are complex and multifactorial [5], likely involving interactions between genetic, psychological (including traumatic experiences), and socioeconomic or other environmental factors [6,7]. This report might explain the suboptimal accuracy of suicidality prediction based on traditional statistical models, for example, with the area under the receiver operating characteristic (AUROC) curve reported to be 0.58 in a meta-analysis of 367 studies, which was only slightly better than a prediction of chance [8].

Alternatively, as tools that can deal with multidimensional data, artificial intelligence techniques (including machine learning) that have been widely used to uncover predictions of multiple diseases [9-11] might have the potential to improve the prediction of suicidality. Indeed, based on data from electronic medical records and mental health questionnaires, as well as sociodemographic factors, researchers have constructed machine learning models that obtained good performance (AUROC=0.590-0.930) for suicidality prediction in the high-risk population [12]. Likewise, more recent efforts to predict suicide attempts or deaths in the general population using this approach have yielded promising results, showing AUROC curves of 0.80 and 0.88 among men and women, respectively, in a Danish population and an AUROC curve of 0.857 among participants in the National Alcohol Epidemiological Survey in the United States [13,14]. However, prior studies did not consider several important factors, such as genetic background [7] and lifestyle factors (eg, diet, physical activity, and sleep) [15,16]. In addition, all these existing models require many input variables (2554 and 2978 inputted features for the Danish and US study, respectively), which have limited implications for daily practice.

Taking advantage of enriched information about suicidality and environmental factors, as well as the available individual-level genotyping data in UK Biobank, we aimed to construct applicable models using a machine learning approach (ie, with limited features) to predict suicidal behavior over both the short and long term. To test the robustness of our models, we validate them among individuals with different genetic risks of suicide.

Methods

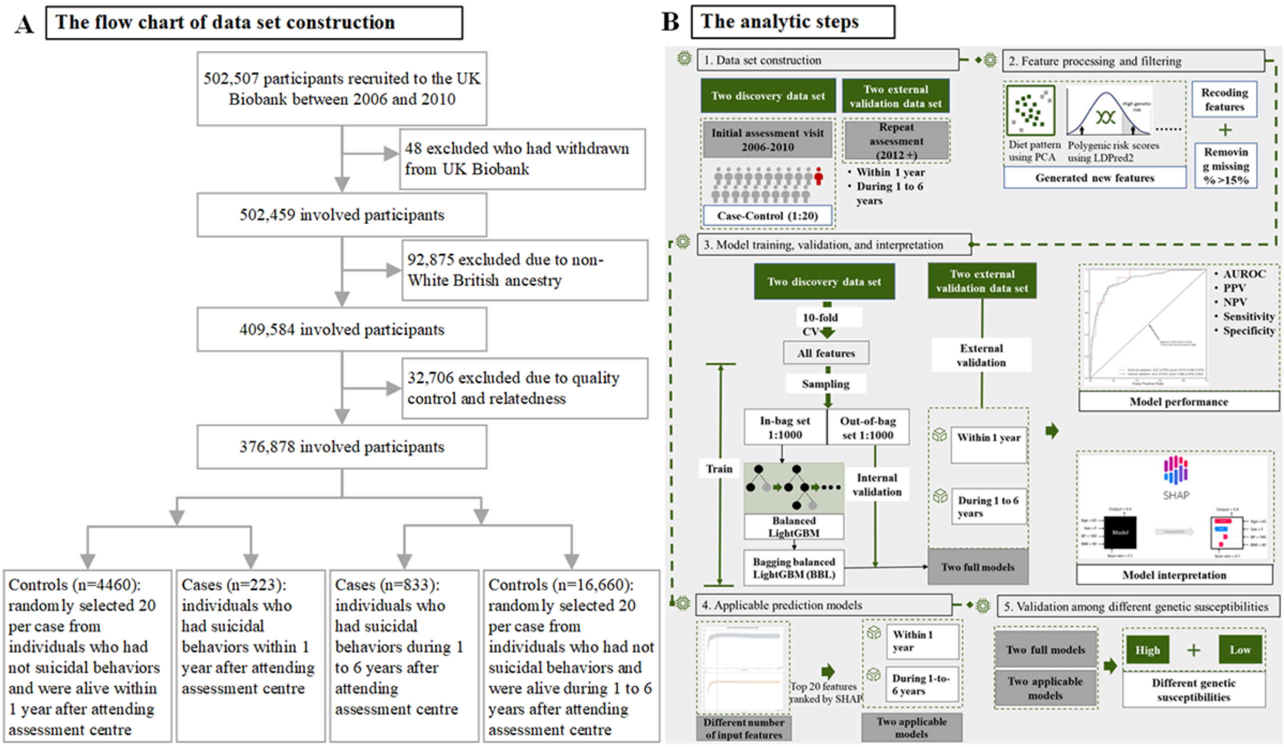
Data Source

A prospective UK Biobank cohort recruited 502,507 participants aged 40 to 69 years across the United Kingdom between 2006 and 2010 [17], which coincides with a high-risk age group of suicide among men and women [18]. At recruitment, all participants filled out questionnaires covering information on sociodemographic, lifestyle, and health–related factors, with a physical examination and collection of biological samples performed during the initial assessment. After recruitment, a proportion was invited several years later to repeat the assessment. In that study, 20,334 participants received a first repeated assessment in 2012 and 2013 and 51,131 received a second repeated assessment visit in 2014.

To track health-related outcomes, UK Biobank data have been linked periodically to multiple national registries with the participants' consent [17]. The inpatient hospital data were obtained through linked hospital records in England, Scotland, and Wales, which were mapped from the Hospital Episode Statistics in England, the Scottish Morbidity Record, and the Patient Episode Database in Wales [19]. Primary care data were obtained from multiple data suppliers, including the Phoenix Partnership and Egton Medical Information Systems, which cover approximately 45% of UK Biobank participants [20]. The mortality data were obtained from national death registers, such as the National Health Services (NHS) Digital Registry and the NHS Central Registry [21].

In this study, among the 502,507 UK Biobank participants, we excluded 48 individuals who had withdrawn from the UK Biobank. To ensure the measurement of genetic susceptibility for suicidality, 376,878 individuals with White ancestry and eligible genotyping data were included in the analysis (Figure 1A). Specifically, the polygenic risk score (PRS) was used as an index of genetic susceptibility, which was generated based on the genome-wide association study (GWAS) summary statistics (ie, effect sizes and standard errors for the variants) from an independent sample of 50,264 Danish residents involving 6,024 cases with an incidence of suicide attempt and 44,240 controls [22]. In addition to removing individuals with nonhomogenous European ancestry, this GWAS study applied principal components of genetic ancestry to take into account the effect of population stratification. We computed the PRS using LDPred2, a method of PRS calculation based on a matrix of correlations between genetic variants, which is faster, more accurate, and more robust than the LDPred14 [23]. In a validation step, the calculated PRS showed a high consistency with the studied phenotype (ie, suicidal behaviors) in our study population, yielding a mean area under the curve of 0.550 and an odds ratio of 2.34 (95% CI 1.66-3.29) by a unit increase in the PRS. During the analysis, we defined the genetic risk levels of suicidality as low (<1st tertile of the PRS), moderate (1st-2nd tertile), and high (>2nd tertile).

Figure 1. Flowchart of the study. AUROC: area under the receiver operator curve; CV: corss validation; LightGBM: light gradient-boosting machine; NPV: negative predictive value; PCA: principal component analysis; PPV: positive predictive value.



Ethics Approval

UK Biobank has full ethical approval from the NHS National Research Ethics Service (16/NW/0274), and informed consent was obtained before data collection from each participant. This study was also approved by the biomedical research ethics committee of West China Hospital (2019-1171).

Ascertainment of Suicidal Behaviors

To expand the application of our models to suicide prevention, both suicide attempts and deaths identified during the study period were considered suicidal behaviors of interest, which is consistent with previous studies [24,25]. Specifically, death by suicide was defined as death with suicide as the underlying cause of death and documented by its correspondence to the International Classification of Diseases 9th revision (ICD-9) and 10th revision (ICD-10) codes (ie, ICD-10: X60-84 and Y10-34; ICD-9: E950-958) [24,25] in the death register. Suicide attempts were considered as hospital admissions with a diagnosis of intentional self-harm (ICD 10: X60-84 and ICD-9: E950-958) or self-harm of undetermined intent (ICD-10: Y10-34) [24,25]. With relatively stable age- and sex-standardized incidence rates, the absolute number of suicide attempts and deaths was high within the first year of enrollment and dropped gradually to half that number in 6th year (Figure S1 of Multimedia Appendix 1 [26-28]). Thus, the outcomes of interest were suicidal behaviors occurring within 1 year (ie, short term) and 1 to 6 years (ie, long term) after the recruitment. We considered individuals with suicide attempts before the recruitment as those having a history of suicide attempts.

Data Set Construction

We constructed separate data sets for predicting suicidal behaviors within 1 year and 1 to 6 years. For the short-term risk prediction, we identified cases of suicide attempts or deaths at least 1 time within 1 year after recruitment (n=223). Controls (n=4460) were randomly selected (1:20 allocation ratio) from the remaining participants who were eligible, alive, and free of suicidal behaviors 1 year after the recruitment, resulting in a data set consisting of 4683 participants (Figure 1A). The same strategies were applied to constructing data sets for long-term (ie, 1 to 6 years) suicide risk prediction, yielding a full data set of 17,493 participants, with 833 (4.8%) and 16,660 (95.2%) cases and controls, respectively.

The 2 aforementioned data sets were then used as discovery data sets for model training and the assessment of internal validity. We additionally used a subsample comprising 50,310 participants of White ancestry from UK Biobank who participated in the repeat assessments. Among this subsample, there were 12 (0.02%) and 31 (0.06%) individuals who attempted or died by suicide within 1 year or during 1 to 6 years after their repeat measurements, respectively, as the validation data set for assessing external validity.

Feature Processing and Filtering

Taking full advantage of the diversity of variables in UK Biobank, we generated a feature list involving multidimensional factors. Due to difficulties obtaining individual genetic data in the real world, we did not involve the PRS in the construction of the prediction models, but we subsequently validated the suicide prediction models with the subgroups of varying (ie, high and low) genetic susceptibility to suicidality to demonstrate their robustness. Information regarding sociodemographic, environmental, and psychosocial factors was derived from the data collected at recruitment using the touchscreen or face-to-face interview questionnaires. For categorical variables (eg, "In general, how would you rate your overall health?"), UK Biobank assigns negative values to categories denoting missingness (ie, −1 refers to "Prefer not to answer," and −3 refers to "Do not know"). Therefore, we recorded those negative values as "NA." Specifically, instead of directly using variables collected through a generic diet questionnaire, we identified dietary patterns based on the results of principal component analysis with varimax rotation (Figure S2 of Multimedia Appendix 1). They were referred to as the prudent, western, and open-sandwich patterns [26], yielding variables with top factor loadings in each component (Table S1 in Multimedia Appendix 1). Medical data included the physical examinations (eg, pulse rate, blood pressure, and grip strength of both hands) conducted during the initial medical center visit, and we calculated mean values when multiple records existed. Additionally, a history of psychiatric disorders was defined as any previous diagnosis of psychiatric disorders before baseline (ICD-10: any F), which was identified through self-reported, hospital inpatient, and primary care data. To consider the influence of somatic fitness, we generated time-varying (0 to 1 and 1 to 4 years before the recruitment) dichotomous variables for each subtype of severe somatic diseases [29]. For the analyses of the total study population, the level of genetic susceptibility to suicidality (low, moderate, or high) was also considered a candidate feature.

After excluding variables with over 15% of missing or irrelevant data (eg, device ID, seated boxing height, and hair color), we included a total of 143 features. The coding book of the included features is shown in Table S2 of Multimedia Appendix 1.

Model Training and Validation

We constructed prediction models using all eligible features. The balanced bagging algorithm is proven to have good performance for classification models with class-imbalanced data [14]. Moreover, the light gradient-boosting machine (LightGBM) [30], as a gradient-boosting algorithm, has been widely applied in machine learning research due to its fast computational speed, high accuracy, and ability to handle missing values [11]. Therefore, considering the data imbalance and the existence of missing values, we used the balanced bagging LightGBM approach to achieve high classification accuracy and fast computation speed, which bagged 1000 balanced LightGBM classifiers (ie, using "class_weight" ="balanced") after random downsampling [31]. We tuned the parameters by using stratified 10-fold cross-validation and grid-search, with the best combination of hyperparameters shown in the Methods section of Multimedia Appendix 1. Each of the 1000 balanced LightGBM classifiers randomly selected subsamples from the group of the minority class (ie, those who had suicidal behavior) and matched samples with the same size from the group of the majority class (ie, those who had no suicidal behavior) to construct case-control samples (ie, the in-bag set). The randomly selected case-control samples were applied to train balanced LightGBM classifiers, and the remaining sample, referred to as the out-of-bag (OOB) set, was used to estimate the prediction of the suicide risk score of the OOB set.

We defined the OOB set as the internal validation set. Specifically, we aggregated the predicted suicide risk scores of the OOB set from the 1000 balanced LightGBM classifiers to estimate the internal validated predicted error [32], and we regarded the models with the highest OOB AUROCs as optimal. Then, we computed the predicted suicide risk scores of the externally validated data sets from the repeated assessments for the optimal model. Due to the lack of agreement regarding which of the risk thresholds of classification provides the most sufficient clinical utility, we computed the AUROC [13,24], sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at different suicide risk score thresholds.

Model Explanation

Interpretations of the models were measured using the Shapley Additive Explanations (SHAP) approach, which quantifies the relationship of the input features with the outcome [33]. Specifically, we computed the contribution of all the features to the studied suicidal behaviors for each participant and assigned each feature an importance score (ie, a SHAP value) after considering its interactions with the remaining features. The absolute values of the average SHAP values were presented as a bar plot illustrating the relative importance of these input features for the models' predictions at the population level.

Applicable Prediction Models

To facilitate the application of the prediction models, we conducted feature reduction by illustrating the changes in the prediction accuracy of the models with different numbers of input features (ie, those with top 10, 20, 50, and 100 SHAP values) [34,35]. As shown in Figure S3 of Multimedia Appendix 1, the models for predicting suicidal behaviors within 1 year and from 1 to 6 years both achieved overall good performance when the input feature dimension with the highest SHAP value was increased to 20, so we considered the models with 20 input features as the applicable prediction models which might facilitate the future implication.

Model Validation Among Individuals With Different Genetic Susceptibilities

To illustrate the robustness of the suicide prediction models, we validated both full and applicable models in the whole population as well as subgroups of varying (ie, high and low) genetic susceptibility to suicidality by computing the OOB performance of these models.

We performed the data set construction and calculation of the PRS using R software version 3.6.1 (Lucent Technologies Co). The machine learning model development was achieved using Python software version 3.6 (Software Foundation), imbalanced-learn 0.9.0, and lightgbm version 3.2.1. We conducted the model interpretation analysis using SHAP version 0.38.1. We then analyzed the models' performance and plot creation using scikit-learn version 1.0.2 and matplotlib version 3.3.2, respectively.

Results

Study Population Characteristics

The data sets for the prediction of suicidal behavior prediction within 1 year and for 1 to 6 years showed largely comparable characteristics at baseline (Table 1). We obtained similar ages, with mean ages of 56.75 (SD 8.03) and 56.65 (SD 7.99) years, respectively, and female-to-male sex distributions of 1:1.13 and 1:1.20, respectively. However, the characteristics of the validation sample for external validity (ie, individuals involved in the repeat assessments) were different from the discovery sample (ie, individuals recruited in the initial assessment visit), characterized by older age, more likely to have a history of psychiatric disorders, and lived in their own accommodation at time of data collection (Table 1).

Table 1. Basic characteristics of analytic samples for the construction of prediction models for 1 year and 1 to 6 years.

Characteristics	Discovery		External validation (n=50,310)
	Within 1 year (n=4683)	1 to 6 years (n=17,493)	
Age (years), mean (SD)	56.75 (8.03)	56.65 (7.99)	63.24 (7.49)
Gender, n (%)			
Female	2480 (53)	9537 (54.5)	25,675 (51)
Male	2203 (47)	7956 (45.5)	24,635 (49)
History of psychiatric disorders, n (%)			
No	3760 (80.3)	14,371 (82.2)	39,226 (78)
Yes	923 (19.7)	3122 (17.8)	11,084 (22)
History of suicide attempt n (%)			
No	4574 (97.7)	17,270 (98.7)	50,108 (99.6)
Yes	109 (2.3)	223 (1.3)	202 (0.4)
Have you ever seen a psychiatrist for nerves, anxiety, tension, or depression? n (%)			
No	3999 (85.4)	15,188 (86.8)	45,372 (90.2)
Yes	660 (14.1)	2248 (12.9)	4556 (9.1)
Missing	24 (0.5)	57 (0.3)	382 (0.8)
Have you ever seen a general practitioner for nerves, anxiety, tension, or depression? n (%)			
No	2882 (61.5)	11,181 (63.9)	34,168 (67.9)
Yes	1768 (37.8)	6215 (35.5)	15,685 (31.2)
Missing	33 (0.7)	97 (0.6)	457 (0.9)
In the past, how often have you smoked tobacco? n (%)			
Smoked on most or all days	1212 (25.9)	4237 (24.2)	11,900 (23.7)
Smoked occasionally	573 (12.2)	2231 (12.8)	6061 (12)
Just tried once or twice	685 (14.6)	2634 (15.1)	7983 (15.9)
I have never smoked	1806 (38.6)	6897 (39.4)	22,789 (45.3)
Missing	407 (8.7)	1494 (8.5)	1577 (3.1)
Do you live in your own accommodation? n (%)			
No	2194 (46.9)	8136 (46.5)	11,670 (23.2)
Yes	2414 (51.5)	9155 (52.3)	38,114 (75.8)
Missing	75 (1.6)	202 (1.2)	526 (1)
Average annual total household income before tax[a], n (%)			
Less than £18,000 (US $16,676)	945 (20.2)	3410 (19.5)	6502 (12.9)
£18,000 to £30,999 (US $16,676 to $28,718)	1045 (22.3)	3802 (21.7)	13,250 (26.3)
£31,000 to £51,999 (US $28,719 to $48,173)	1073 (22.9)	3958 (22.6)	13,577 (27)
£52,000 to £100,000 (US $48,174 to $92,642)	816 (17.4)	3113 (17.8)	9529 (18.9)
Greater than £100,000 (US $92,642)	194 (4.1)	797 (4.6)	2614 (5.2)
Missing	610 (13)	2413 (13.8)	4838 (9.6)
Cases, n (%)	223 (47.62)	833 (47.62)	12 (0.02)[b] and 31 (0.06)[c]

[a] Note that income data were collected between 2008 and 2010. Income was converted to US dollars according to the current exchange range.

[b] Refers to the number of cases of suicidal behaviors within 1 year.

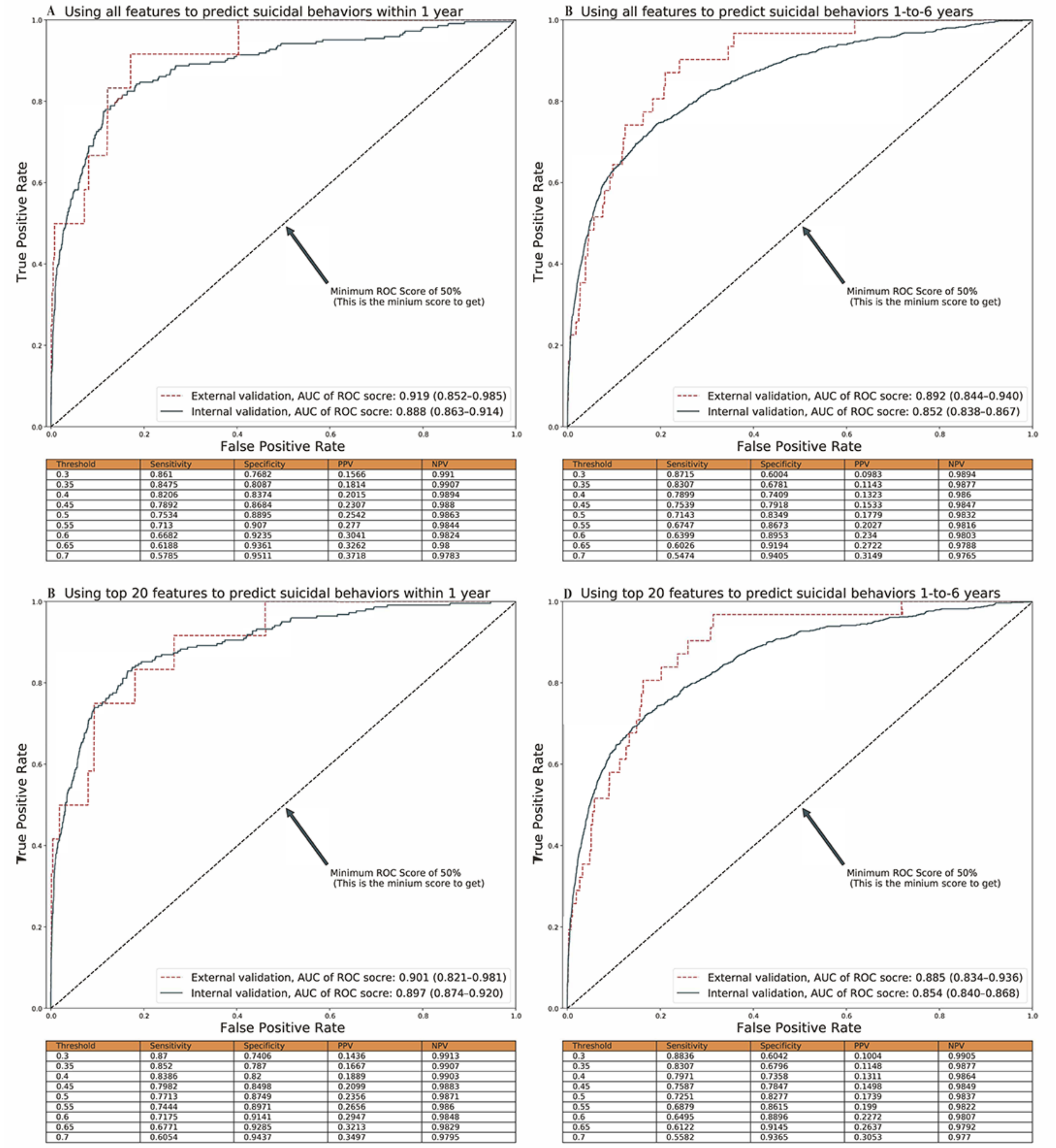
[c] Refers to the number of cases of suicidal behaviors for 1 to 6 years.

Prediction Models Involving All Features

The internal validated AUROC of the prediction models involving all features was 0.888 (95% CI 0.863-0.914) for the prediction of suicidal behaviors within 1 year and 0.852 (95% CI 0.838-0.867) for 1 to 6 years (Figure 2). Figure 2 shows values of sensitivity, specificity, and predictive indices over a series of risk thresholds. For instance, at the 0.70 risk threshold, the short- and long-term sensitivities were, respectively, 57.85% and 54.74%, the specificities were 95.11% and 94.05%, the PPVs were 37.18% and 31.49%, and the NPVs were 97.83% and 97.65%. Furthermore, the externally validated AUROC curves were 0.919 (95% CI 0.852-0.985) for the model predicting suicidal behaviors within 1 year and 0.892 (95% CI 0.844-0.940) for the model predicting suicidal behaviors between 1 and 6 years, indicating the robustness of the prediction models.

Regarding the importance of features measured using SHAP values, similar to age, family income, and body fat percentage, mental health–related factors (eg, history of psychiatric disorders, history of suicide attempt, etc) were top ranked in both models within 1 year (7 among the top 20 features) and during 1 to 6 years (8 among the top 20 features). However, notable differences were observed with respect to some lifestyle and social factors (eg, “How many years of using a mobile phone at least once per week to make or receive calls?” and “Age you first had sexual intercourse”), as these seemed to have greater importance for only the prediction models within 1 year but not 1 to 6 years (Figure 3). In contrast, some general health–related factors (ie, “In general how would you rate your overall health?” and “Compared with 1 year ago, has your weight changed?”) seemed only important for the 1-to-6 years prediction model. Detailed information on the included features is shown in Table S2 of Multimedia Appendix 1.

Figure 2. The performance of prediction models using all input features and top 20 features. The area under the receiver operating characteristic (AUROC) curve. The tables showed the internal validation performance (ie, sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]) of suicide prediction models at different classified thresholds.



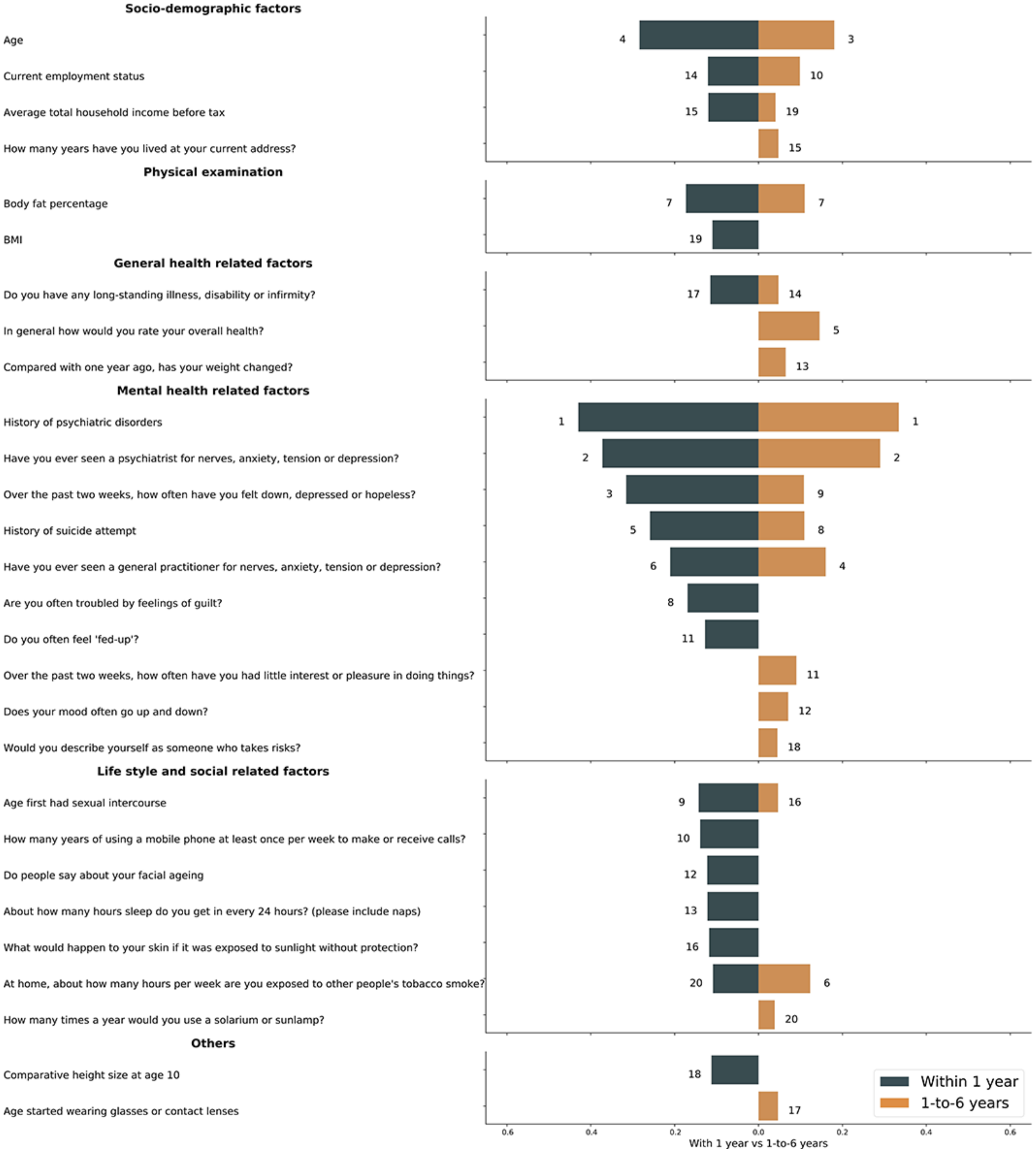
Threshold	Sensitivity	Specificity	PPV	NPV
0.3	0.861	0.7682	0.1566	0.991
0.35	0.8475	0.8087	0.1814	0.9907
0.4	0.8206	0.8374	0.2015	0.9894
0.45	0.7892	0.8684	0.2307	0.988
0.5	0.7534	0.8895	0.2542	0.9863
0.55	0.713	0.907	0.277	0.9844
0.6	0.6682	0.9235	0.3041	0.9824
0.65	0.6188	0.9361	0.3262	0.98
0.7	0.5785	0.9511	0.3718	0.9783

Threshold	Sensitivity	Specificity	PPV	NPV
0.3	0.8715	0.6004	0.0983	0.9894
0.35	0.8307	0.6781	0.1143	0.9877
0.4	0.7899	0.7409	0.1323	0.986
0.45	0.7539	0.7918	0.1533	0.9847
0.5	0.7143	0.8349	0.1779	0.9832
0.55	0.6747	0.8673	0.2027	0.9816
0.6	0.6399	0.8953	0.234	0.9803
0.65	0.6026	0.9194	0.2722	0.9788
0.7	0.5474	0.9405	0.3149	0.9765

Threshold	Sensitivity	Specificity	PPV	NPV
0.3	0.87	0.7406	0.1436	0.9913
0.35	0.852	0.787	0.1667	0.9907
0.4	0.8386	0.82	0.1889	0.9903
0.45	0.7982	0.8498	0.2099	0.9883
0.5	0.7713	0.8749	0.2356	0.9871
0.55	0.7444	0.8971	0.2656	0.986
0.6	0.7175	0.9141	0.2947	0.9848
0.65	0.6771	0.9285	0.3213	0.9829
0.7	0.6054	0.9437	0.3497	0.9795

Threshold	Sensitivity	Specificity	PPV	NPV
0.3	0.8836	0.6042	0.1004	0.9905
0.35	0.8307	0.6796	0.1148	0.9877
0.4	0.7971	0.7358	0.1311	0.9864
0.45	0.7587	0.7847	0.1498	0.9849
0.5	0.7251	0.8277	0.1739	0.9837
0.55	0.6879	0.8615	0.199	0.9822
0.6	0.6495	0.8896	0.2272	0.9807
0.65	0.6122	0.9145	0.2637	0.9792
0.7	0.5582	0.9365	0.3053	0.977

Figure 3. The comparison of top 20 features identified in suicide risk prediction full models for within 1 year and 1-to-6 years. The dark blue and yellow bar represent the relatively importance of these input features for the prediction, respectively. And the numbers next to the bars are corresponding to the ranking of top 20 features. The detailed information of the included features is shown in Table S2 in Multimedia Appendix 1.



Prediction Models Involving the Top 20 Features

Figure S3 of Multimedia Appendix 1 displays the indices of model performance for the models involving different numbers of the top features (ie, top 20, 40, 60, and 100). Accordingly, the 2 models with the top 20 input features were considered optimal (Figure 2). The AUROC curves for their internal and external validations for the within 1-year suicide prediction were 0.897 (95% CI 0.874-0.920) and 0.901 (95% CI 0.821-0.981), respectively. For the 1-to-6 years prediction, the corresponding estimate was 0.854 (95% CI 0.840-0.868) and 0.885 (95% CI 0.834-0.936), respectively. Based on the applicable models, we found individuals in the top quintile of predicted risk accounting for 91.7% (n=11) and 80.7% (n=25) of all cases of suicide attempts or deaths within 1 year and during 1 to 6 years, respectively.

Models for Individuals With Different Genetic Susceptibilities

Using both full and simplified prediction models, we obtained a comparable prediction accuracy for individuals with low and high genetic susceptibilities to suicidality (Figures S4 and S5 of Multimedia Appendix 1). For instance, for short-term risk prediction, the AUROC curves for models with the top 20 involved features were 0.907 and 0.885 for the high and low genetic susceptibility groups, respectively. The corresponding numbers for the long-term risk prediction were 0.869 and 0.822, respectively.

Discussion

Principal Findings

In this study on a community-based UK Biobank cohort of over 0.5 million UK residents aged 40 to 69 years (covering the age group with a high risk of suicide [18]), we established machine learning–based models to accurately predict both short- and long-term risks of suicide attempts and deaths (AUROC=0.892-0.919). Importantly, our applicable models achieved high predictive accuracy across populations with varying genetic susceptibility to suicide with a limited number (ie, 20) of phenotypic features that could be accessed easily through practice. Specifically, we found that individuals with the top 20% of predicted risks comprised over 80% of real cases of suicide attempts or deaths, suggesting that our approach may be a cost-effective way to identify high-risk middle-aged individuals who should be targeted for suicide prevention. In addition, besides some well-known suicide risk factors (ie, mental health–related conditions), these established models provide novel insights into factors driving suicidal behaviors, revealing that some lifestyle and social factors (eg, cell phone use frequency, etc) may be risk factors for suicidal behaviors in the short-term, while self-reported general health ratings are more important for the prediction of long-term suicidal risk.

In line with 2 previous studies focusing on machine learning–based suicide risk prediction in the general population using data from Danish health registers [13] and the National Alcohol Epidemiological Survey of the United States [14], our results identified mental health–related factors (ie, prior suicide attempt, history of psychiatric disorders, and past emotion) and sociodemographic factors (ie, age and family income) as top features for suicide risk prediction. However, benefiting from the enriched data in UK Biobank, particularly items related to neuroticism, lifestyle, social contacts, and self-rated general health, our prediction models achieved improved performance. In addition, the comparison of features that matter for short- versus longer-term suicide risk was not addressed in prior investigations. Similar efforts have been made in some specific populations (eg, patients receiving psychiatric [24] or other medical care [36] and soldiers [37]), though with only comparable predictive accuracy (ie, the AUROC curves ranged between 0.77 and 0.93) with more homogeneous clinical populations.

Consistent with our findings, neuroticism was reported as a risk factor for suicidal behaviors in a previous study, with plausible mechanisms of shared genetic components [38]. Likewise, severe somatic diseases, disabilities, or physical weakness have consistently been reported to be associated with higher suicide risk, which is possibly due to the chronic stress associated with these diagnoses and living with these diseases [7]. Previous efforts exploring the association between BMI and suicidality have led to inconsistent results [39], and the association between body fat and suicidality has remained largely unexplored. Nevertheless, our findings of the association between body fat percentage and suicidality gain support from a Mendelian randomization analysis, which revealed a causal link between a high percentage of body fat and depression [40].

Our attempts to construct separate models for the prediction of both short- and long-term suicide risks indicated that the models generally achieved better prediction accuracy for the more immediate period before the suicide attempt or death, which is in line with the findings of prior studies concerning time-varying suicide risk assessments [24,41]. While factors directly reflecting mental health impairment show consistent importance for both short- and long-term suicidal risk prediction, the significance of lifestyle and social factors (eg, the frequency of using a cell phone to make or receive calls) was mainly observed for short-term risk (ie, within 1 year), indicating the role of lower social support and social relations among individuals with suicide risk [7]. Additionally, our findings on the association between self-reported health ratings and long-term suicide risk are in line with the results of the Danish study, which also found that medical diagnoses and medications related to some somatic illnesses (eg, infection and respiratory diseases) measured 48 months before suicide were more important indicators of suicide risk than those measured 6 months earlier [13].

Strengths and Weaknesses

The major merits of our study include the use of multidimensional data (including individual-level genotyping data) from a large community-based cohort of UK Biobank. The application of the machine learning approach, together with the use of SHAP values for feature interpretation, enabled us to identify the most informative variables that maximized the efficiency of the data for an accurate prediction of suicide risk. The imbalance in the sample sizes of the cases and the controls was mitigated by randomly downsampling and setting class weights for imbalanced classes in LightGBM during the training step [30,42]. Further, we improved the feasibility of our prediction models by using the feature reduction process, where accurate classification was achieved with only 20 features. Although no similar data from independent samples could be used for external validation, the validity of our models was demonstrated in a subgroup of UK Biobank participants who repeated surveys many years after the baseline measurement (showing different basic characteristics compared to the discovery data set), as well as the subpopulations stratified by their level of genetic susceptibility to suicidality.

A notable limitation of this study is the absence of data from emergency care departments, which were the main source for suicide case identification in previous studies [13,43]. Therefore, our study focused on suicidal behaviors resulting in hospitalization or death, and those with less severe consequences require further investigation. In addition, it is difficult to distinguish suicide attempts from nonsuicidal intentional self-harm based on ICD codes, as clinical diagnoses tended to be consequence oriented (ie, leading to life-threatening harm or not) or dependent on self-reported reasoning on intent. Moreover, such outcome ascertainment strategies have been demonstrated to suffer from poor sensitivity, resulting in a risk of underestimation of suicidal cases, as well as attenuated associations between studied exposures and suicidal outcomes [44]. Nevertheless, as this is the most feasible method to identify suicidal behavior, similar definitions and ascertainment of suicidal behaviors have been widely used in other large community- or population-based studies with a similar focus [24,25]. Furthermore, we only used the LightGBM as the base estimator for bagging, mainly due to its capability to handle missing values and achieve high discrimination accuracy [30]. It is possible that other machine learning approaches (eg, deep neural network), with some common methods of feature engineering (eg, standardization, one-hot encoding), might obtain better performance at the price of model interpretability. Finally, the UK Biobank study recruited only 5.5% of the invited individuals in the age range of 40 to 69 years, leading to a selection bias of the study population compared to the entirety of the population in the United Kingdom [45]. Consequently, the generalization of our findings to the total UK population and other populations cannot be made.

Conclusions

In conclusion, based on a UK Biobank cohort, we established clinically applicable machine learning–based models for accurately predicting both short- and long-term risks of suicidal behaviors. The good performance of the models for subgroups with different genetic susceptibilities to suicidality highlights the possibility of applying these models to high-risk individual identification in the general middle-aged population, which may facilitate the development of cost-effective suicide prevention.

Acknowledgments

This research was conducted using UK Biobank (application 54803). This work uses data provided by patients and collected by the National Health Services (NHS) as part of their care and support. This research used data assets made available by National Safe Haven as part of the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation. We thank the team members of the West China Biomedical Big Data Center for Disease Control and Prevention for their support.

This work was supported by the 1.3.5 Project for Disciplines of Excellence of the West China Hospital at Sichuan University (ZYYC21005) and by the National Natural Science Foundation of China (81971262) to author HS.

Data Availability

Data from the UK Biobank are available to all researchers upon submitted application. All codes associated with the current submission are available and can be requested by contacting the corresponding authors.

Authors' Contributions

HS and JW were responsible for the study's concept and design. JW, HY, and YQ were responsible for data and project management. JW, YZ, and HY performed the data cleaning and analysis. JW, JQ, TZ, UAV, and HS interpreted the data. JW, JQ, TZ, YZ, HY, YS, JY, UAV, and HS drafted and revised the manuscript. All the authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods, tables, and figures.
[DOCX File , 5323 KB - publichealth_v9i1e43419_app1.docx]

References

1. Roth GA. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Jan;385(9963):117-171. [doi: 10.1016/s0140-6736(14)61682-2]
2. Suicides in England and Wales. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/suicidesintheunitedkingdomreferencetables [accessed 2022-06-01]
3. Kapur N, Steeg S, Turnbull P, Webb R, Bergen H, Hawton K, et al. Hospital management of suicidal behaviour and subsequent mortality: a prospective cohort study. Lancet Psychiatry 2015 Sep;2(9):809-816. [doi: 10.1016/s2215-0366(15)00169-8]
4. Houston K, Haw C, Townsend E, Hawton K. General practitioner contacts with patients before and after deliberate self harm. Br J Gen Pract 2003 May;53(490):365-370 [FREE Full text] [Medline: 12830563]
5. Turecki G. The molecular bases of the suicidal brain. Nat Rev Neurosci 2014 Dec;15(12):802-816 [FREE Full text] [doi: 10.1038/nrn3839] [Medline: 25354482]

6. Nock MK, Borges G, Bromet EJ, Alonso J, Angermeyer M, Beautrais A, et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. Br J Psychiatry 2008 Mar 02;192(2):98-105 [FREE Full text] [doi: 10.1192/bjp.bp.107.040113] [Medline: 18245022]
7. Turecki G, Brent DA, Gunnell D, O'Connor RC, Oquendo MA, Pirkis J, et al. Suicide and suicide risk. Nat Rev Dis Primers 2019 Oct 24;5(1):74. [doi: 10.1038/s41572-019-0121-0] [Medline: 31649257]
8. Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. Psychol Bull 2017 Mar;143(2):187-232. [doi: 10.1037/bul0000084] [Medline: 27841450]
9. Qiu J, Yin J, Qian W, Liu J, Huang Z, Yu H, et al. A novel multiresolution-statistical texture analysis architecture: radiomics-aided diagnosis of PDAC based on plain CT images. IEEE Trans Med Imaging 2021 Jan;40(1):12-25. [doi: 10.1109/tmi.2020.3021254]
10. Mi X, Zou B, Zou F, Hu J. Permutation-based identification of important biomarkers for complex diseases via machine learning models. Nat Commun 2021 May 21;12(1). [doi: 10.1038/s41467-021-22756-2]
11. Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. Nat Med 2020 Jan 13;26(1):71-76. [doi: 10.1038/s41591-019-0724-8] [Medline: 31932807]
12. Belsher BE, Smolenski DJ, Pruitt LD, Bush NE, Beech EH, Workman DE, et al. Prediction models for suicide attempts and deaths: a systematic review and simulation. JAMA Psychiatry 2019 Jun 01;76(6):642-651. [doi: 10.1001/jamapsychiatry.2019.0174] [Medline: 30865249]
13. Gradus JL, Rosellini AJ, Horváth-Puhó E, Street AE, Galatzer-Levy I, Jiang T, et al. Prediction of sex-specific suicide risk using machine learning and single-payer health care registry data from Denmark. JAMA Psychiatry 2020 Jan 01;77(1):25-34 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.2905] [Medline: 31642880]
14. García de la Garza Á, Blanco C, Olfson M, Wall MM. Identification of suicide attempt risk factors in a national US survey using machine learning. JAMA Psychiatry 2021 Apr 01;78(4):398-406 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4165] [Medline: 33404590]
15. Berardelli I, Corigliano V, Hawkins M, Comparelli A, Erbuto D, Pompili M. Lifestyle interventions and prevention of suicide. Front Psychiatry 2018 Nov 6;9:567 [FREE Full text] [doi: 10.3389/fpsyt.2018.00567] [Medline: 30459660]
16. Bernert RA, Turvey CL, Conwell Y, Joiner TE. Association of poor subjective sleep quality with risk for death by suicide during a 10-year period: a longitudinal, population-based study of late life. JAMA Psychiatry 2014 Oct 01;71(10):1129-1137 [FREE Full text] [doi: 10.1001/jamapsychiatry.2014.1126] [Medline: 25133759]
17. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018 Oct 10;562(7726):203-209 [FREE Full text] [doi: 10.1038/s41586-018-0579-z] [Medline: 30305743]
18. Iacobucci G. Suicide rates continue to rise in England and Wales. BMJ 2020 Sep 03;370:m3431. [doi: 10.1136/bmj.m3431] [Medline: 32883661]
19. Mapping inpatient hospital data across England, Scotland and Wales. UK Biobank. URL: http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/inpatient_mapping.pdf [accessed 2020-10-01]
20. Data providers and dates of data availability. UK Biobank. URL: https://biobank.ndph.ox.ac.uk/showcase/exinfo.cgi?src=Data_providers_and_dates [accessed 2021-03-01]
21. Mortality data: linkage to death registries. UK Biobank. URL: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/DeathLinkage.pdf [accessed 2020-10-01]
22. Erlangsen A, Appadurai V, Wang Y, Turecki G, Mors O, Werge T, et al. Genetics of suicide attempts in individuals with and without mental disorders: a population-based genome-wide association study. Mol Psychiatry 2020 Oct 16;25(10):2410-2421 [FREE Full text] [doi: 10.1038/s41380-018-0218-y] [Medline: 30116032]
23. Privé F, Arbel J, Vilhjálmsson BJ. LDpred2: better, faster, stronger. Bioinformatics 2020 Dec 16;36(22-23):5424-5431 [FREE Full text] [doi: 10.1093/bioinformatics/btaa1029] [Medline: 33326037]
24. Chen Q, Zhang-James Y, Barnett EJ, Lichtenstein P, Jokinen J, D'Onofrio BM, et al. Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: A machine learning study using Swedish national registry data. PLoS Med 2020 Nov 6;17(11):e1003416 [FREE Full text] [doi: 10.1371/journal.pmed.1003416] [Medline: 33156863]
25. Simon GE, Johnson E, Lawrence JM, Rossom RC, Ahmedani B, Lynch FL, et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. Am J Psychiatry 2018 Oct 01;175(10):951-960 [FREE Full text] [doi: 10.1176/appi.ajp.2018.17101167] [Medline: 29792051]
26. Wei X, Zhu C, Ji M, Fan J, Xie J, Huang Y, et al. Diet and risk of incident lung cancer: a large prospective cohort study in UK Biobank. Am J Clin Nutr 2021 Dec 01;114(6):2043-2051. [doi: 10.1093/ajcn/nqab298] [Medline: 34582556]
27. Tyrrell J, Wood AR, Ames RM, Yaghootkar H, Beaumont RN, Jones SE, et al. Gene-obesogenic environment interactions in the UK Biobank study. Int J Epidemiol 2017 Apr 01;46(2):559-575 [FREE Full text] [doi: 10.1093/ije/dyw337] [Medline: 28073954]

28. Niarchou M, Byrne EM, Trzaskowski M, Sidorenko J, Kemper KE, McGrath JJ, et al. Genome-wide association study of dietary intake in the UK biobank study and its associations with schizophrenia and other traits. Transl Psychiatry 2020 Feb 03;10(1):51-171 [FREE Full text] [doi: 10.1038/s41398-020-0688-y] [Medline: 32066663]
29. Morgan C, Webb RT, Carr MJ, Kontopantelis E, Chew-Graham CA, Kapur N, et al. Self-harm in a primary care cohort of older people: incidence, clinical management, and risk of suicide and other causes of death. Lancet Psychiatry 2018 Nov;5(11):905-912. [doi: 10.1016/s2215-0366(18)30348-1]
30. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W. Lightgbm: A highly efficient gradient boosting decision tree. 2017 Presented at: 31st Conference on Neural Information Processing Systems; December 8; Long Beach, CA p. A.
31. Xia H, Wei X, Gao Y, Lv H. Traffic prediction based on ensemble machine learning strategies with bagging and LightGBM. 2019 Presented at: 53rd IEEE International Conference on Communications Workshops (ICC Workshops); May 20-24; Shanghai, China. [doi: 10.1109/ICCW.2019.8757058]
32. Breiman L. Out-of-bag estimation. University of California Berkeley. URL: https://www.stat.berkeley.edu/~breiman/OOBestimation.pdf [accessed 2021-11-01]
33. Lundberg S. A unified approach to interpreting model predictions. 2017 Presented at: 31st Conference on Neural Information Processing Systems; December 8; Long Beach, CA.
34. Bugaj M, Wrobel K, Iwaniec J. Model explainability using SHAP values for LightGBM predictions. 2021 Presented at: 2021 IEEE XVIIth International Conference on the Perspective Technologies and Methods in MEMS Design (MEMSTECH); May 12-16; Polyana, Ukraine. [doi: 10.1109/MEMSTECH53091.2021.9468078]
35. Bi Y, Xiang D, Ge Z, Li F, Jia C, Song J. An interpretable prediction model for identifying N-Methylguanosine sites based on XGBoost and SHAP. Mol Ther Nucleic Acids 2020 Dec 04;22:362-372 [FREE Full text] [doi: 10.1016/j.omtn.2020.08.022] [Medline: 33230441]
36. Barak-Corren Y, Castro VM, Javitt S, Hoffnagle AG, Dai Y, Perlis RH, et al. Predicting suicidal behavior from longitudinal electronic health records. Am J Psychiatry 2017 Mar 01;174(2):154-162. [doi: 10.1176/appi.ajp.2016.16010077] [Medline: 27609239]
37. Nock MK, Millner AJ, Joiner TE, Gutierrez PM, Han G, Hwang I, et al. Risk factors for the transition from suicide ideation to suicide attempt: results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). J Abnorm Psychol 2018 Mar;127(2):139-149 [FREE Full text] [doi: 10.1037/abn0000317] [Medline: 29528668]
38. Peters EM, John A, Bowen R, Baetz M, Balbuena L. Neuroticism and suicide in a general population cohort: results from the UK Biobank Project. BJPsych Open 2018 Mar 16;4(2):62-68 [FREE Full text] [doi: 10.1192/bjo.2017.12] [Medline: 29971148]
39. Perera S, Eisen RB, Dennis BB, Bawor M, Bhatt M, Bhatnagar N, et al. Body Mass Index is an important predictor for suicide: results from a systematic review and meta-analysis. Suicide Life Threat Behav 2016 Dec 20;46(6):697-736. [doi: 10.1111/sltb.12244] [Medline: 27094229]
40. Speed MS, Jefsen OH, Børglum AD, Speed D, Østergaard SD. Investigating the association between body fat and depression via Mendelian randomization. Transl Psychiatry 2019 Aug 05;9(1):184 [FREE Full text] [doi: 10.1038/s41398-019-0516-4] [Medline: 31383844]
41. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. Clin Psychol Sci 2017 Apr 11;5(3):457-469. [doi: 10.1177/2167702617691560]
42. Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst, Man, Cybern C 2012 Jul;42(4):463-484. [doi: 10.1109/tsmcc.2011.2161285]
43. Sanderson M, Bulloch AG, Wang J, Williams KG, Williamson T, Patten SB. Predicting death by suicide following an emergency department visit for parasuicide with administrative health care system data and machine learning. EClinicalMedicine 2020 Mar;20:100281 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100281] [Medline: 32300738]
44. Swain RS, Taylor LG, Braver ER, Liu W, Pinheiro SP, Mosholder AD. A systematic review of validated suicide outcome classification in observational studies. Int J Epidemiol 2019 Oct 01;48(5):1636-1649. [doi: 10.1093/ije/dyz038] [Medline: 30907424]
45. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 2017 Nov 01;186(9):1026-1034 [FREE Full text] [doi: 10.1093/aje/kwx246] [Medline: 28641372]

Abbreviations

AUROC: area under the receiver operating characteristic
GWAS: genome-wide association study
ICD-9/ICD-10: International Classification of Diseases 9th revision/10th revision
LightGBM: light gradient-boosting machine
NHS: National Health Services
NPV: negative predictive value

OOB: out of bag
PPV: positive predictive value
PRS: polygenic risk score
SHAP: Shapley Additive Explanations

Edited by A Mavragani, T Sanchez; submitted 11.10.22; peer-reviewed by D Bowen, A Teles; comments to author 01.12.22; revised version received 21.12.22; accepted 12.01.23; published 20.02.23.

Please cite as:
Wang J, Qiu J, Zhu T, Zeng Y, Yang H, Shang Y, Yin J, Sun Y, Qu Y, Valdimarsdóttir UA, Song H
Prediction of Suicidal Behaviors in the Middle-aged Population: Machine Learning Analyses of UK Biobank
JMIR Public Health Surveill 2023;9:e43419
URL: https://publichealth.jmir.org/2023/1/e43419
doi:10.2196/43419
PMID:36805366

Original Paper

Community-Level Social Support to Mitigate the Impact of Combined Frailty and Multimorbidity on Psychological Distress Among Rural Chinese Older Adults During the COVID-19 Pandemic: Multilevel Modeling Study

Yi Wang[1], PhD; Peipei Fu[1], PhD; Jie Li[1], PhD; Tingting Gao[1], PhD; Zhengyue Jing[1], PhD; Qiong Wang[1], PhD; Dan Zhao[1], PhD; Chengchao Zhou[1,2], PhD

[1]Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[2]National Health Commission Key Lab of Health Economics and Policy Research, Shandong University, Jinan, China

Corresponding Author:
Chengchao Zhou, PhD
Centre for Health Management and Policy Research
School of Public Health
Cheeloo College of Medicine, Shandong University
44 Wen-hua-xi Road
Jinan, 250012
China
Phone: 86 531 8838 1567
Fax: 86 531 8838 2553
Email: zhouchengchao@sdu.edu.cn

Abstract

Background: Accumulating research provides evidence that the psychological health of older people deteriorated from before to during the COVID-19 pandemic. Unlike robust individuals, coexisting frailty and multimorbidity expose older adults to more complicated and wide-ranging stressors. Community-level social support (CSS) is also an important impetus for age-friendly interventions, and it is 1 of the components of social capital that is seen as an ecological-level property. To date, we have not found research that examines whether CSS buffered the adverse impacts of combined frailty and multimorbidity on psychological distress in a rural setting during COVID-19 in China.

Objective: This study explores the combined effect of frailty and multimorbidity on psychological distress in rural Chinese older adults during the COVID-19 pandemic and examines whether CSS would buffer the aforementioned association.

Methods: Data used in this study were extracted from 2 waves of the Shandong Rural Elderly Health Cohort (SREHC), and the final analytic sample included 2785 respondents who participated in both baseline and follow-up surveys. Multilevel linear mixed effects models were used to quantify the strength of the longitudinal association between frailty and multimorbidity combinations and psychological distress using 2 waves of data for each participant, and then, cross-level interactions between CSS and combined frailty and multimorbidity were included to test whether CSS would buffer the adverse impact of coexisting frailty and multimorbidity on psychological distress.

Results: Frail older adults with multimorbidity reported the most psychological distress compared to individuals with only 1 or none of the conditions (β=.68, 95% CI 0.60-0.77, *P*<.001), and baseline coexisting frailty and multimorbidity predicted the most psychological distress during the COVID-19 pandemic (β=.32, 95% CI 0.22-0.43, *P*<.001). Further, CSS moderated the aforementioned association (β=–.16, 95% CI –0.23 to –0.09, *P*<.001), and increased CSS buffered the adverse effect of coexisting frailty and multimorbidity on psychological distress during the COVID-19 pandemic (β=–.11, 95% CI –0.22 to –0.01, *P*=.035).

Conclusions: Our findings suggest that more public health and clinical attention should be paid to psychological distress among multimorbid older adults with frailty when facing public health emergencies. This research also suggests that community-level interventions prioritizing social support mechanisms, specifically improving the average levels of social support within communities, may be an effective approach to alleviate psychological distress for rural older adults who concurrently manifest frailty and multimorbidity.

(JMIR Public Health Surveill 2023;9:e43762) doi:10.2196/43762

KEYWORDS

psychological distress; frailty; multimorbidity; community-level social support; COVID-19 pandemic; psychological; rural; older adults; community; support; effectiveness

Introduction

Mental health is an important component of healthy aging, and reducing psychological distress is vital for both physical health and well-being in later life [1]. Accumulating research [2-4] provides evidence that the psychological health of older people deteriorated from before to during the COVID-19 pandemic. It is important to identify groups that are at greater risk of psychological distress in response to the COVID-19 pandemic so as to develop appropriate interventions, especially for those who live in rural areas, because rural older persons have lower socioeconomic status, have less access to social support and health services, and report worse mental health compared with those in urban areas [5]. Frailty and multimorbidity are common and major age-related clinical syndromes and well-established independent risk factors for psychological health in older people [6-8]. Separate studies have reported that frailty or multimorbidity can independently increase the levels of psychological distress in older people [7,9,10]. However, an individual can be multimorbid but not frail or frail while having multimorbidity. More importantly, frailty can be more modifiable than chronic conditions [11], and this distinction is important when assessing relationships with psychological health because they may inform different management strategies for prognosis and planning interventions. Both the National Institute for Health and Care Excellence (NICE) and the British Geriatrics Society emphasize the importance of combined frailty and multimorbidity in older adults who are at greater risk of adverse outcomes and might benefit more from treatment optimization [12,13]. For example, according to the statement document of NICE, not all older adults with multimorbidity require additional support beyond standard health care but frailty should be considered an important condition to keep in mind in the management of older adults with multimorbidity [13,14]. Unlike robust individuals, coexisting frailty and multimorbidity would expose older adults to more complicated and wide-ranging stressors. Nevertheless, the combined effect of frailty and multimorbidity on psychological distress has not been previously investigated in community-dwelling older adults in rural China. By analyzing whether combined multimorbidity and frailty contribute to psychological distress can inform the development of interventions for those at heightened risk and aid policy decisions concerning the subsequent management of the COVID-19 pandemic response.

The social determinants of the health framework [15] believe that health inequality cannot be attributed solely to differences in individual characteristics. The environmental contexts that reflect the social resources and opportunities can affect individuals’ access to healthy lives and choices strongly and ultimately shape our health and well-being [16]. Understanding the contextual determinants of psychological distress has important public health implications for developing population-scale interventions. Just as biologists have distinguished levels of organization to describe living organisms, social epidemiologists have advanced levels to delineate different scales of environmental contexts that affect an individual’s health, such as the community level or state level [17,18]. Among these different levels of contexts, the community is 1 of the most common areas of focus in China. The concept of “community” was adopted by the Chinese government through the community-building project since the 1980s, which can be understood as a social group comprising people living in the same area, sharing similar values or culture, and having direct or indirect continuous interaction. Communities in rural China exert important functionalities in providing primary health care services, advocating government policies, managing social life affairs, and improving quality and cultural accomplishment for community residents [19]. In China’s response to COVID-19, the Chinese government implemented social distancing and lockdown measures at the community level, which has become the intersection and bridge between government institutions, businesses, social organizations, volunteer groups, families, and individuals [20]. The psychological health of older people living in rural areas may be particularly sensitive to community social environments as they tend to be less mobile and rely more on locally provided services, as well as social support and connection. Community-level social support (CSS), defined as the extent to which a community leads someone to be cared for, loved, respected, and be a member of a network of mutual obligations [21], has been associated prominently with overall health and longevity [22,23]. Thus, CSS may serve as a potentially important target of intervention on psychological health.

Social capital theory [24] provides a useful framework for understanding the potentially protective mechanisms of increased CSS on psychological distress theoretically. CSS, as 1 of the components of social capital that is seen as an ecological-level property, can give rise to social capital that reflects the levels of reciprocity and connectedness within a community [24], and thus exerts a contextual effect on individual health via several possible pathways, such as dissemination of health-promoting knowledge, maintenance of healthy norms of behavior through informal social control, promotion of access to local services and amenities, and provision of psychological processes of emotional support and mutual respect [25,26]. Although individual-level studies of social support have yielded useful insights [27-29], an examination of social support as a contextual effect on reducing psychological distress remains understudied. Previous studies have highlighted the protective effects of CSS on the mental health and well-being of older people [30]; however, no such research has been conducted in rural China. People living in rural areas usually have more difficulties in accessing public services and lack of medical attention compared with those in urban areas [31] and thus may experience more psychological distress. Further, CSS may not only affect mental health directly but also have indirect effects by serving as a buffer against the adverse impacts of different stressors, such as frailty and multimorbidity. To date, we have not found research that examines whether CSS buffered the adverse effects of combined frailty and multimorbidity on psychological distress in a rural setting during the COVID-19 in China. A better understanding of the cross-level interaction between CSS and combined frailty and multimorbidity is necessary to design more appropriate and precise interventions aimed at reducing psychological distress.

Taken together, the first aim of our study is to explore the longitudinal association between coexisting frailty and multimorbidity and psychological distress in rural older adults, as well as whether baseline combined frailty and multimorbidity predict psychological distress during the COVID-19 pandemic. The second aim of this study is to examine whether improving CSS would buffer the adverse impacts of coexisting frailty and multimorbidity on psychological distress.

Methods

Study Design and Data Sources

Data used in this study were extracted from 2 waves of the Shandong Rural Elderly Health Cohort (SREHC), an ongoing longitudinal study of community-dwelling older people in rural Shandong, China. The baseline survey was completed in June 2019, which was considered prior to the COVID-19 pandemic on the basis of the declaration from the World Health Organization on March 11, 2020, that COVID-19 can be characterized as a pandemic [32]. The detailed study design, sampling, and data collection have been described elsewhere [2,33,34]. Briefly, we used the multistage stratified random sampling method to select participants, excluding those who had a clinical diagnosis of dementia or psychiatric diseases, could not complete the interview due to severe physical conditions, or were unwilling to cooperate with the interviewers. There were 3243 respondents aged 60 years and above who participated in the baseline survey. Next, a follow-up survey was conducted during the COVID-19 pandemic from August to September 2020. The final analytic sample included 2785 (85.88%) respondents who participated in both baseline and follow-up surveys. We restricted the analysis to these 2 data points for the following reasons: First, we were specifically interested in examining the short-term consequences of combined frailty and multimorbidity on psychological distress so as to identify the most vulnerable group with the greatest risk of worse psychological outcomes and inform early interventions during the pandemic. Second, previous research has shown that social support would have an immediate effect on a scaled measure of psychological health [27]; therefore, we assessed the moderating role of CSS between combined frailty and multimorbidity and psychological distress with just a 1-year lag based on the assumption that there may be a limited time lag for CSS to mitigate one's psychological distress.

Ethical Considerations

The study purpose, significance, methods, and risks were explained to all participants, and written informed consent was obtained from each participant before the 2 surveys. Ethical approval for the study was obtained from the Ethics Committee of Shandong University (approval no. 20181228).

Measurements

We used the 10-item Kessler Psychological Distress Scale (K10) [35] to assess the psychological distress in this study. K10 is a commonly global measure for screening mental health, and its reliability and validity have been confirmed among older people in China [36,37]. K10 contains 10 items using a 5-point Likert scale from "none of the time" to "all of the time" to evaluate psychological distress, including depression, anxiety, nervousness, hopelessness, restlessness, and worthlessness, in the past 4 weeks. K10 in this study showed good internal consistency reliability, with Cronbach α=.91 for both baseline and follow-up surveys. The raw score of K10 ranges from 10 to 50 points, with higher scores indicating a higher level of psychological distress. We transformed the K10 score into z-scores based on the mean and SD.

We used the 10-item Social Support Rating Scale (SSRS) [38] to assess the level of an individual's social support, which is the most prevalent tool for measuring social support in China [39]. The SSRS contains objective support (3 items about visible, practical, and direct support), subjective support (4 items about the perceived level of support from family members, neighbors, and friends), and social support (3 items about the level of social support used). The SSRS has been shown to have good reliability and validity [40] in China, with Cronbach α>.7 for both baseline and follow-up surveys in our study. The raw SSRS score ranges from 12 to 66 points, with higher scores indicating higher levels of social support. In this study, we transformed raw SSRS scores into z-scores. A CSS score was computed by aggregating the standardized measures (z-scores) of the individual-level social support within each community, which reflected the average levels of social support for each community. This method has been proven to be valid and widely applied in previous studies related to social support [30,41,42].

In this study, frailty was assessed using the criteria of the frailty phenotype, which was proposed and validated by Fried et al [43]. It consists of 5 components: shrinking (unintentional weight loss), slowness (a walking time of 4.6 m adjusted by gender and height; individuals who met the criteria in the walking test or were unable to perform the test due to physical limitations were considered positive for slowness) [44], self-reported exhaustion, weakness (grip strength), and low physical activity. Older people who met 3-5 criteria were considered frail, while those who met 0-2 criteria were considered nonfrail.

Multimorbidity was defined as the coexistence of 2 or more chronic health conditions based on the Chinese Centers for Disease Control and Prevention (CDC) recommendations [45] and previous studies [46,47], including hypertension, diabetes, chronic lung disease, heart disease, asthma, liver disease, stroke, dyslipidemia, cancer, digestive disease, kidney disease, and arthritis. To validate the accuracy of this information, trained interviewers asked for help from village doctors to confirm the self-reported chronic condition information in the chronic disease case management system. We categorized multimorbidity into 3 groups: no chronic condition, 1 chronic condition, and multimorbidity.

The control variables included sex (male, female), age, marital status (divorced/widowed, married), educational attainment (illiteracy, primary school, junior school, high school or above), economic status (household income per capita; quartile 1 was the poorest and quartile 4 the richest), sedentary behavior (hours/day), smoking status (never/past, current smoking), drinking status (never/past, current drinking), and individual-level social support. We selected these variables as potential confounders based on existing studies [48-50].

Statistical Analysis

First, we compared the baseline characteristics between respondents and nonrespondents using cross-tabulation with *t* tests for continuous variables and chi-square tests for categorical variables (see Table 1). To examine the combined effects of frailty and multimorbidity, we created a *categorical indicator* with the following 6 groups: (1) nonfrail without a chronic condition, (2) nonfrail with 1 chronic condition, (3) nonfrail with multimorbidity, (4) frail without a chronic condition, (5) frail with 1 chronic condition, and (6) frail with multimorbidity. Multilevel linear mixed effects models were used to quantify the strength of the longitudinal association between frailty and multimorbidity combinations and psychological distress using 2 waves of data for each participant (see model 1 in Table 2), and then, cross-level interactions between CSS and combined frailty and multimorbidity (see model 2 in Table 2) were included. The selected 3-level linear mixed effects modeling strategy was appropriate for repeated measures to account for the hierarchical structure of the data set that included survey time nested within individuals and individuals nested within communities. A random intercept was included in each model. Both model 1 and model 2 were adjusted for sex, age, education, economic status, marital status, sedentary time, smoking status, drinking status, individual-level social support, and survey time (to capture any individual-level idiosyncratic disturbances over time).

Table 1. Attrition analysis of selected baseline characteristics.

Characteristics	Analytical sample (N=2785)	Dropouts (N=458)
Sex, n (%); *t* (*df*)/χ^2 (*df*)=0.030 (1), *P* =.86		
Male	1015 (36.45)	165 (36.03)
Female	1770 (63.55)	293 (63.97)
Age (years), mean (SD); *t* (*df*)/χ^2 (*df*)=1.345 (1), *P*=.18	69.19 (6.16)	68.78 (6.34)
Education, n (%) ; *t* (*df*)/χ^2 (*df*)=0.271 (3), *P* =.97		
Illiteracy	1164 (41.80)	189 (41.27)
Primary school	1076 (38.64)	182 (39.74)
Junior school	407 (14.61)	66 (14.41)
High school or above	138 (4.96)	21 (4.59)
Economic status, n (%) ; *t* (*df*)/χ^2 (*df*)=1.323 (3), *P* =.72		
Quartile 1	692 (24.85)	110 (24.02)
Quartile 2	675 (24.24)	120 (26.20)
Quartile 3	720 (25.85)	110 (24.02)
Quartile 4	698 (25.06)	118 (25.76)
Marital status, n (%) ; *t* (*df*)/χ^2 (*df*)=3.451 (1), *P* =.06		
Divorced/widowed	695 (24.96)	133 (29.04)
Married	2090 (75.04)	325 (70.96)
Sedentary time, mean (SD); *t* (*df*)/χ^2 (*df*)=0.114 (1), *P*=.91 (hours/day)	4.36 (2.01)	4.34 (1.90)
Smoking status, n (%) ; *t* (*df*)/χ^2 (*df*)=0.009 (1), *P* =.93		
Never/past	2202 (79.07)	363 (79.26)
Current	583 (20.93)	95 (20.74)
Drinking status, n (%) ; *t* (*df*)/χ^2 (*df*)=1.473 (1), *P* =.22		
Never/past	2161 (77.59)	367 (80.13)
Current	624 (22.41)	91 (19.87)
Combined frailty and multimorbidity , n (%) ; *t* (*df*)/χ^2 (*df*)=7.935 (5), *P* =.16		
Nonfrail, no chronic condition	684 (24.56)	129 (28.17)
Nonfrail, 1 chronic condition	856 (30.74)	142 (31.00)
Nonfrail, multimorbidity	745 (26.75)	104 (22.71)
Frail, no chronic condition	71 (2.55)	13 (2.84)
Frail, 1 chronic condition	171 (6.14)	36 (7.86)
Frail, multimorbidity	258 (9.26)	34 (7.42)
K10[a], mean (SD); *t* (*df*)/χ^2 (*df*)=0.640 (1), *P*=.52	16.63 (7.44)	16.39 (7.60)
Individual social support, mean (SD); *t* (*df*)/χ^2 (*df*)=0.515 (1), *P*=.61	43.10 (6.27)	42.94 (6.45)
CSS[b], mean (SD); *t* (*df*)/χ^2 (*df*)=–1.190 (1), *P*=.23	43.05 (2.94)	43.23 (2.78)

[a] K10: 10-item Kessler Psychological Distress Scale.

[b] CSS: community-level social support.

Table 2. Longitudinal associations between combined frailty and multimorbidity, CSS[a], and psychological distress before and during the COVID-19 pandemic[b].

Effects	Model 1		Model 2	
	β (95% CI)	*P* value	β (95% CI)	*P* value
Fixed effects: combined frailty and multimorbidity				
Nonfrail, no chronic condition	Reference	N/A[c]	Reference	N/A
Nonfrail, 1 chronic condition	0.13 (0.07 to 0.19)	<.001	0.13 (0.07 to 0.18)	<.001
Nonfrail, multimorbidity	0.25 (0.19 to 0.31)	<.001	0.24 (0.18 to 0.31)	<.001
Frail, no chronic condition	0.40 (0.27 to 0.53)	<.001	0.40 (0.26 to 0.53)	<.001
Frail, 1 chronic condition	0.48 (0.38 to 0.57)	<.001	0.46 (0.37 to 0.56)	<.001
Frail, multimorbidity	0.68 (0.60 to 0.77)	<.001	0.67 (0.59 to 0.75)	<.001
Fixed effects: CSS (z-scores)	–0.06 (–0.11 to –0.02)	.008	–0.02 (–0.08 to 0.03)	.42
Nonfrail, no chronic condition × CSS	N/A	N/A	Reference	N/A
Nonfrail, 1 chronic condition × CSS	N/A	N/A	–0.04 (–0.09 to 0.01)	.14
Nonfrail, multimorbidity × CSS	N/A	N/A	–0.03 (–0.08 to 0.03)	.33
Frail, no chronic condition × CSS	N/A	N/A	–0.03 (–0.15 to 0.09)	.67
Frail, 1 chronic condition × CSS	N/A	N/A	–0.11 (–0.19 to –0.02)	.011
Frail, multimorbidity × CSS	N/A	N/A	–0.16 (–0.23 to –0.09)	<.001
Fixed effects: Baseline	Reference	N/A	Reference	N/A
Fixed effects: Follow-up	0.09 (0.02 to 0.15)	.008	0.08 (0.02 to 0.15)	.012
Fixed effects: Intercept	0.63 (0.19 to 1.07)	.005	0.64 (0.20 to 1.08)	.005
Random effects				
Community level	0.02 (0.01 to 0.04)	<.001	0.03 (0.01 to 0.05)	<.001
Individual level	0.54 (0.50 to 0.58)	<.001	0.54 (0.51 to 0.58)	<.001
Residual	0.30 (0.29 to 0.32)	<.001	0.30 (0.28 to 0.31)	<.001

[a] CSS: community-level social support.

[b] In total, 41 observations were excluded due to having missing data. All models were adjusted for sex, age, education, economic status, marital status, sedentary time, smoking status, and drinking status.

[c] N/A: not applicable.

Next, to minimize the possible reverse causality, 2-level mixed effects models with a lagged dependent variable (LDV) were conducted to account for the cluster of participants within communities, and we also controlled the levels of psychological distress at baseline (see Table 3). Model 1 in Table 3 included follow-up psychological distress as the dependent variable and the baseline categorical indicator and the change in CSS between the 2 waves as the independent variable to examine whether baseline frailty and multimorbidity combinations predicted psychological distress during the COVID-19 pandemic. Model 2 in Table 3 included model 3 variables plus a 2-way cross-level interaction term between baseline frailty and multimorbidity combinations and the increased CSS. Both model 3 and model 4 were adjusted for baseline confounding variables, including sex, education, economic status, marital status, sedentary time, smoking status, drinking status, individual-level social support, and K10 scores, to test whether the increased CSS from baseline to follow-up mitigated the adverse impact of combined frailty and multimorbidity on psychological distress.

Table 3. Baseline combined frailty and multimorbidity and changes in CSS[a] for predicting psychological distress at follow-up[b].

Effects	Model 1		Model 2	
	β (95% CI)	*P* value	β (95% CI)	*P* value
Fixed effects: combined frailty and multimorbidity				
Nonfrail, no chronic condition	Reference	N/A[c]	Reference	N/A
Nonfrail, 1 chronic condition	0.08 (0.01 to 0.15)	.033	0.08 (0.01 to 0.15)	.027
Nonfrail, multimorbidity	0.13 (0.06 to 0.21)	<.001	0.14 (0.06 to 0.21)	<.001
Frail, no chronic condition	–0.03 (–0.20 to 0.15)	.76	–0.02 (–0.20 to 0.16)	0.84
Frail, 1 chronic condition	0.26 (0.14 to 0.38)	<.001	0.26 (0.14 to 0.39)	<.001
Frail, multimorbidity	0.32 (0.22 to 0.43)	<.001	0.33 (0.22 to 0.44)	<.001
Fixed effects: Δ CSS (z-scores)	–0.07 (–0.12 to –0.02)	.003	–0.03 (–0.10 to 0.03)	.34
Nonfrail, no chronic condition × ΔCSS	N/A	N/A	Reference	N/A
Nonfrail, 1 chronic condition × ΔCSS	N/A	N/A	–0.06 (–0.13 to 0.01)	.08
Nonfrail, multimorbidity × ΔCSS	N/A	N/A	–0.02 (–0.09 to 0.05)	.62
Frail, no chronic condition × ΔCSS	N/A	N/A	–0.07 (–0.21 to 0.08)	.38
Frail, 1 chronic condition × ΔCSS	N/A	N/A	–0.05 (–0.16 to 0.05)	.32
Frail, multimorbidity × ΔCSS	N/A	N/A	–0.11 (–0.22 to –0.01)	.035
Fixed effects: Intercept	0.71 (0.30 to 1.12)	<.001	0.72 (0.31 to 1.13)	<.001
Random effects				
Community level	0.02 (0.01 to 0.04)	<.001	0.02 (0.01 to 0.04)	<.001
Residual	0.47 (0.44 to 0.49)	<.001	0.47 (0.44 to 0.49)	<.001

[a] CSS: community-level social support.

[b] In total, 41 observations were excluded due to having missing data. All models were adjusted for baseline covariates (sex, education, economic status, marital status, sedentary time, smoking status, drinking status, and the 10-item Kessler Psychological Distress Scale [K10] score), as well as concern about COVID-19 and the likelihood of contracting COVID-19 at follow-up.

[c] N/A: not applicable.

We ran additional analyses under the same multilevel framework to see how confounders adjusted the models. From model 0 (the unadjusted model) through model 2 (fully adjusted model), the variance partition coefficient for the community level decreased from 3% to 2% and from 6% to 4% for longitudinal association and predictive models, respectively, which implies that our confounders explained about one-third of the community-level variance in worsened psychological distress. Next, we added the separate variable of frailty status and multimorbidity in a full model to show the independent effects of frailty and multimorbidity as a comparison. All analyses were performed using Stata 14.2 (Stata Corp LP).

Results

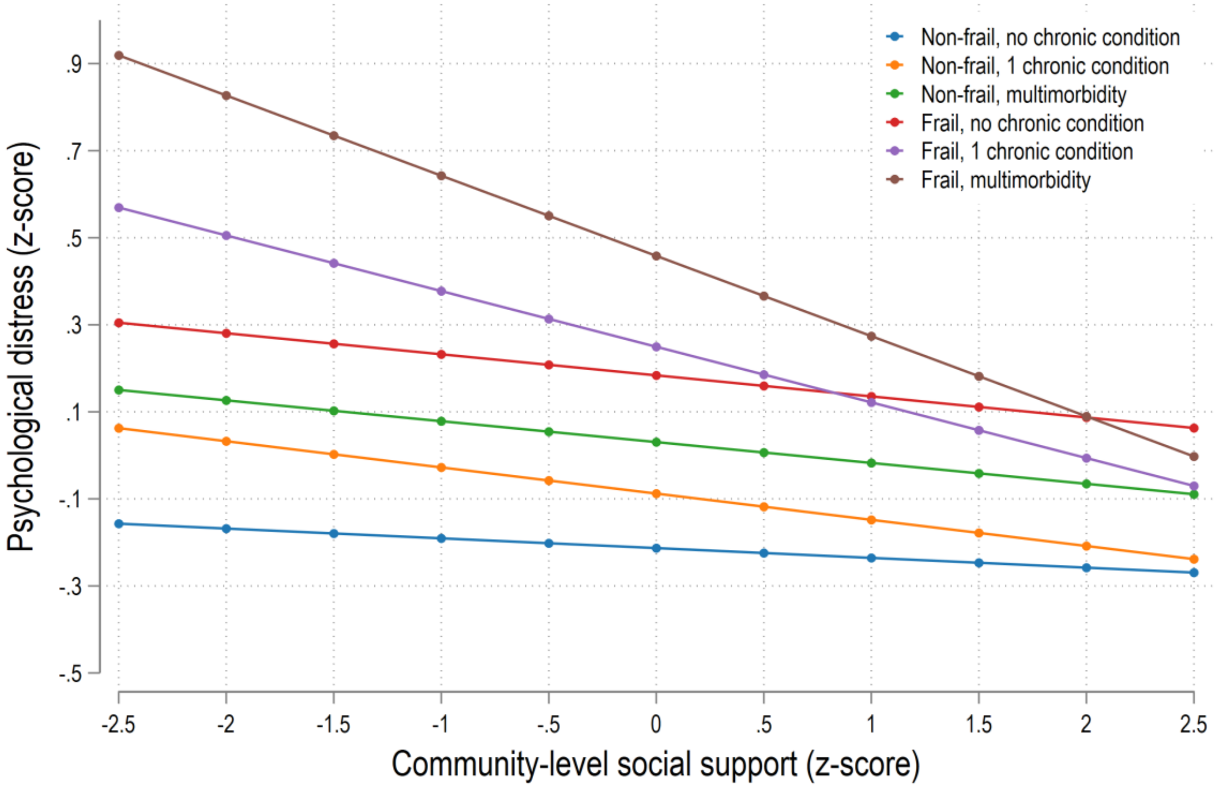
Descriptive Analyses

The baseline survey included 3243 older people from 58 communities. Of the 3243 respondents at baseline, 458 (14.12%) individuals were lost to follow-up because of the following reasons: outmigration with children (n=121, 26.4%), dropping out (n=295, 64.4%), and deaths (n=42, 9.2%). This yielded an analytic sample of 2785 (follow-up rate=85.88%) who participated in the 2 surveys. All characteristics (sex, age, education, economic status, marital status, sedentary time, smoking and drinking status, chronic condition, frailty, and social support) were similar between respondents (n=2785, 85.88%) and nonrespondents (n=458, 14.12%) at the follow-up survey (*P*>.05). Of the 2785 respondents, 1770 (63.55%) were female, the mean age was 69.19 (SD 6.16) years, 1164 (41.80%) had no education, and 2062 (74.04%) were married. The characteristics of the selected sample and dropouts are provided in Table 1. Psychological distress score values increased from a K10 score of 16.60 (SD 7.46) at baseline to 18.23 (SD 8.00) at the follow-up survey. The combination of frailty and multimorbidity was observed in 258 (9.26%) respondents, while 684 (24.56%) respondents were nonfrail without a chronic condition. Both individual-level social support and CSS decreased during follow-up (see Table 1).

Multilevel Mixed Effects Models

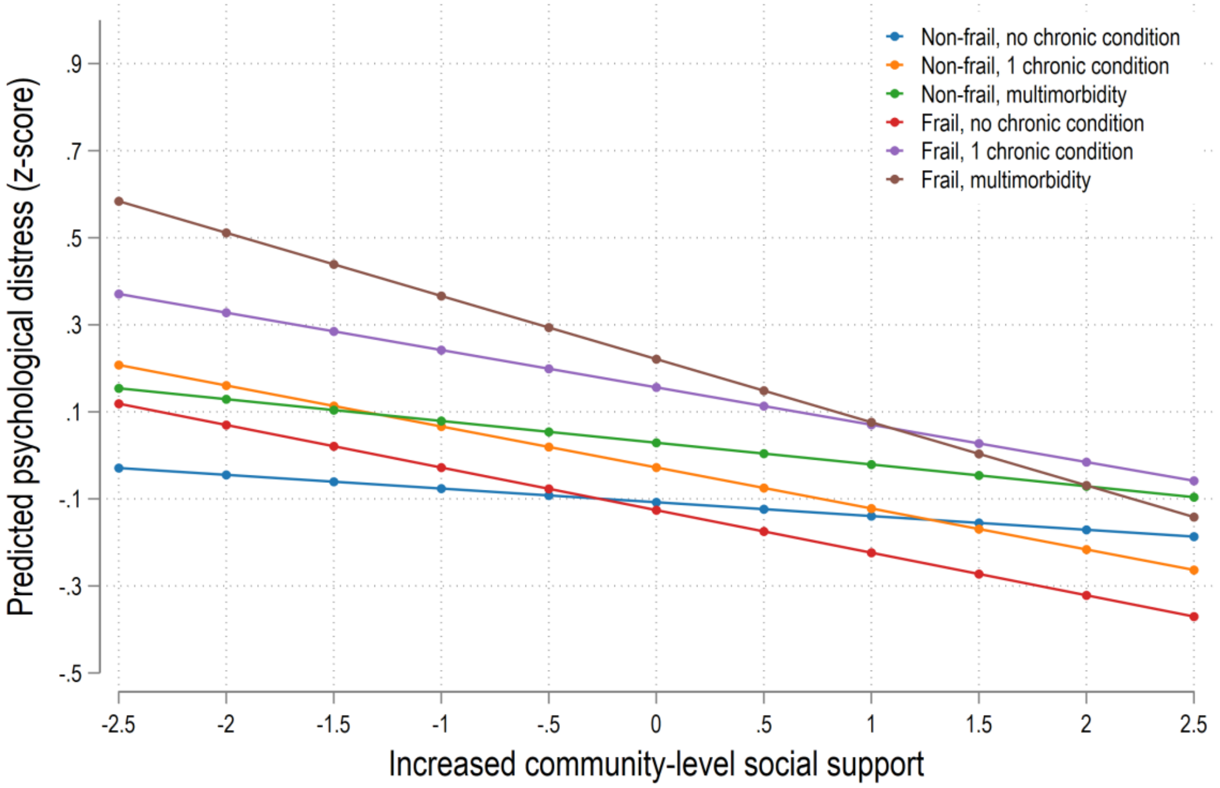
Table 2 shows the results of multilevel mixed effects models with random intercepts at the individual and community levels. Model 1 showed that psychological distress significantly deteriorated during the COVID-19 pandemic compared with before the COVID-19 (β=.09, 95% CI 0.02-0.15, *P*=.008), while CSS (β=–.06, 95% CI –0.11 to –0.02, *P*=.008) was protectively associated with psychological distress. After adjustment for potential confounders, older adults who were nonfrail with 1 chronic condition (β=.13, 95% CI 0.07-0.19, *P*<.001), nonfrail with multimorbidity (β=.25, 95% CI 0.19-0.31, *P*<.001), frail without a chronic condition (β=.40, 95% CI 0.27-0.53, *P*<.001), and frail with 1 chronic condition (β=.48, 95% CI 0.38-0.57, *P*<.001) were at higher risk of psychological distress compared with older adults without any of the conditions, but the highest risk was observed in those with the combined presence of frailty and multimorbidity (β=.68, 95% CI 0.60-0.77, *P*<.001). As shown in model 2, the main effect of combined frailty and multimorbidity on psychological distress remained significant, and there was a significant cross-level interaction between CSS and frail individuals with multimorbidity (β=–.16, 95% CI –0.23 to –0.09, *P*<.001). To help the interpretations, the plots in Figure 1 show the estimated values of levels of psychological distress based on the estimates in model 2. Within the group of respondents who had combined frailty and multimorbidity, those with lower levels of CSS reported significantly higher psychological distress than respondents with higher levels of CSS, indicating the moderating role of CSS between frail individuals with multimorbidity and psychological distress.

Figure 1. Longitudinal associations between CSS and psychological distress over 2 waves by combinations of frailty and multimorbidity (fully adjusted models). CSS: community-level social support.



Model 1 in Table 3 shows the effects of baseline frailty and multimorbidity on wave 2 psychological distress. Those with coexisting frailty and multimorbidity (β=.32, 95% CI 0.22-0.43, *P*<.001) showed the strongest association with greater psychological distress during the COVID-19 pandemic compared with nonfrail individuals without a chronic condition, while increased CSS (β=–.07, 95% CI –0.12 to –0.02, *P*=.003) from wave 1 to wave 2 was significantly associated with lower subsequent levels of psychological distress. In model 2, we examined whether increased CSS moderated the adverse impact of baseline coexisting frailty and multimorbidity on psychological distress during the COVID-19 pandemic. We found a significant interaction between increased CSS from wave 1 to wave 2 and coexisting frailty and multimorbidity (β=–.11, 95% CI –0.22 to –0.01, *P*=.035) for wave 2 psychological distress. As Figure 2 demonstrates, for individuals with coexisting frailty and multimorbidity at baseline, those with higher increased levels of CSS reported significantly lower psychological distress than respondents with lower increased levels of CSS, while such effect was less substantial for older adults who were nonfrail without a chronic condition.

Figure 2. Increased CSS and psychological distress during the COVID-19 pandemic by baseline combinations of frailty and multimorbidity (fully adjusted models). CSS: community-level social support.



Discussion

Principal Findings

Using data from a cohort of community-dwelling older adults in rural China, this study identified that coexisting frailty and multimorbidity are significantly associated with greater levels of psychological distress and that baseline coexisting frailty and multimorbidity predicted higher levels of psychological distress during the COVID-19 pandemic. Further, residing in communities with a higher atmosphere of social support moderates the relationship of combined frailty and multimorbidity and psychological distress. We also emphasized the role of increased CSS in mitigating the adverse effects of combined frailty and multimorbidity on psychological distress during the COVID-19 pandemic. These results highlight the community social environment, more specifically CSS, as an important buffer that may mitigate the adverse impacts of combined frailty and multimorbidity on psychological distress.

Multimorbidity and frailty are 2 major adverse health conditions among older adults [14], and the 2 conditions are also risk factors for mental health [6,7,51]. However, there is no previous research that has explicitly examined their combined effects on psychological distress in older adults. Only 1 similar study has shown that the combination of frailty and multimorbidity is associated with physical limitation and mortality [52], and our study extends the literature for psychological health. In our sample, the results of longitudinal association showed that the magnitude of the coefficient of psychological distress was strongest for coexisting frailty and multimorbidity compared to only 1 or none of the conditions. This is where this study makes the first novel contribution to previous work: we identified a vulnerable group of older adults at increased risk of psychological distress. It is plausible that frail older adults already face a higher risk of poor resolution of homeostasis [53], plus the worry that results from multimorbidity, making them more vulnerable when coping with stressful situations and increasing the risk of psychological distress. Results from this prospective study also add evidence to the literature by identifying baseline combined frailty and multimorbidity as a predictor of greater psychological distress in older adults during the COVID-19 pandemic, emphasizing the importance of identifying and preventing coexisting frailty and multimorbidity in reducing psychological distress in older adults when facing major public health issues. We should interpret our findings incorporating the context of the pandemic. Unlike normal daily circumstances, the COVID-19 pandemic would expose people to more complicated and wide-ranging stressors. In addition to the primary stressor of the fear of infection, the pandemic has brought a plethora of secondary stressors, such as social isolation and resource scarcity [51]. These sources of stress could lead rural residents to face the potential disruption of basic medical services, especially the possibly reduced access to non–COVID-19 health care, which therefore adversely affects their psychological health. For example, older persons with frailty and multimorbidity had greater needs for basic medical services; however, the lockdown and polymerase chain reaction (PCR) testing policies led to worries about access to health care services, which in turn led to greater psychological distress. One evidence is that while the rural health care system did not break down during the quarantine, most of the rural residents chose to delay health care seeking due to worry about being infected [54]. Our findings suggested that older adults with frailty and multimorbidity needed more professional psychological support during the pandemic; however, the availability of psychological therapy or support was limited, especially in rural areas. Moreover, the current mass quarantines and restrictions to public transport have inevitably become a major barrier to accessing psychological therapy or support for older adults [55], which may also cause their increased psychological distress.

Importantly, our study identified that residing in a community with higher social support mitigates the relationship between coexisting frailty and multimorbidity and psychological distress. This finding adds to the literature documenting the stress-buffering role of CSS in psychological health. The moderating role of CSS remains significant even after controlling for individual-level social support, suggesting that there are contextual effects that alleviate psychological distress regardless of individual-level social support. In other words, older adults with lower individual-level social support could also benefit from their communities with a higher social support atmosphere. One possible explanation is social contagion [56]. CSS is meaningful in shaping community cohesion and connectedness [30], which is important for individuals to obtain the resources they need. Compared with living in communities with low levels of social support, health-related information may be disseminated faster in communities with higher levels of social support [57], which may increase the likelihood of people taking up healthy standards of behavior. Older adults with coexisting frailty and multimorbidity are more likely to collect and notice health-related information and resources that are provided by social contagion (health-related information, behavioral norms, etc) and thereby benefit more from social contagion, while those without any of the conditions may not pay much attention to the information and thus may not be sensitive to the community atmosphere. We noticed that the interaction between nonfrail older individuals with multimorbidity and CSS was not significant, while that between frail older individuals with multimorbidity and CSS was significant. This finding may be due to the different severity of disease within the multimorbidity group across the frailty state (ie, frail people may have more advanced diseases than nonfrail people). Older people with multimorbidity who have not yet developed a frail state may be resilient to the negative effect of multiple comorbidities. Thus, frail individuals with multimorbidity are more sensitive and pay more attention to their own health than nonfrail individuals with multimorbidity. Concerning the dynamic role of CSS, we observed that increased CSS from baseline to follow-up mitigated the adverse impact of baseline coexisting frailty and multimorbidity on psychological distress during the COVID-19 pandemic. Practically, our findings imply that interventions targeted toward improving CSS might alleviate the effects of frailty and multimorbidity on the psychological well-being of older adults, especially when facing major public health issues.

We also observed an increase in psychological distress in the context of the COVID-19 pandemic. In addition to the fear of being infected with COVID-19, widespread secondary stressors associated with the pandemic may include decreased opportunities for interpersonal contact and excursions and potential economic loss [51]. Given the lack of control groups that were unaffected by the COVID-19 pandemic in our study, more complex statistical design is needed in future studies, and we will not discuss more at this point.

Strengths and Limitations

This study has several strengths and limitations. The novelties and strengths of this study include the study design and the measures we used. First, we used a large sample of community-dwelling older adults and followed up the psychological health outcomes with little missing data. Moreover, we used repeated measures of the key predictors and outcomes, including frailty, chronic conditions, social support, and psychological distress, and adjudicated K10 and CSS using standardized criteria. Furthermore, we not only used 2 waves of the data to show the longitudinal association between coexisting frailty and multimorbidity and psychological distress but also used baseline values to predicted psychological distress during the COVID-19 pandemic, which showed the robustness of the results.

Some limitations need to be considered. First, the observed relationships between combined frailty and multimorbidity and psychological distress cannot be constructed as causal relations, because this was an observational study. However, we took a set of measures to minimize the possible reverse causality, such as using lagged dependent variable models, controlling psychological distress at baseline, and adding a wave variable. Second, this study was conducted only in rural areas; whether the results are applicable to urban areas needs further study. In addition, our sample may not be representative of all rural older adults in China, as our participants were selected from 3 counties. As such, more nationally representative studies should be conducted to corroborate our findings. Third, as social support was assessed based on questionnaires, the results are subject to measurement bias: individuals with severe psychological distress might perceive themselves as having poor social support. However, we speculate that the aggregating social support at the community level could be influenced less. Fourth, some variables, such as chronic conditions and sedentary time, were self-reported, which may result in some recall bias. Finally, we cannot completely rule out residual confounding because of unmeasured variables; however, our results were robust after adjustment for multiple confounding variables.

Conclusion

In summary, the findings of this study indicate that community-dwelling rural older adults with both frailty and multimorbidity have higher levels of psychological distress compared to those with either condition alone or none of the conditions. Baseline coexisting frailty and multimorbidity predict the most subsequent psychological distress during the COVID-19 pandemic. Thus, frail rural older adults with multimorbidity constitute a high-risk group for psychological distress. Moreover, residing in communities with highly average levels of social support moderates the relationship between combined frailty and multimorbidity and psychological distress, and we confirmed the role of increased CSS in buffering the adverse effects of combined frailty and multimorbidity on psychological distress during the COVID-19 pandemic. Our findings suggest that more public health and clinical attention should be paid to frail older adults with multimorbidity in response to psychological distress. In addition, improving the average levels of social support within communities may be an effective approach to alleviating psychological distress in older adults who concurrently manifest frailty and multimorbidity.

Acknowledgments

This work was supported by the National Science Foundation of China (grant nos. 72274109, 72204147, 71774104, 71974117); the China Medical Board (grant no. 16-257) the Cheeloo Youth Scholar Grant, Shandong University (grant nos. IFYT1810, 2012DX006); and the China Population Welfare Foundation (grant no. GJB-6-20211028-556).

We thank the officials of health agencies, all participants, and staff at the study sites for their cooperation.

Data Availability

The data sets used in this study are not publicly available due to them containing information that could compromise research participant privacy, but they are available from the corresponding author upon reasonable request.

Authors' Contributions

YW was responsible for drafting the manuscript, study design, data analysis, and interpretation; PF, JL, TG, ZJ, QW, and DZ for data acquisition, study concept, and methodology; and CZ for conceptualization, writing—review and editing, and funding acquisition. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Saadeh M, Welmer A, Dekhtyar S, Fratiglioni L, Calderón-Larrañaga A. The role of psychological and social well-being on physical function trajectories in older adults. J Gerontol A Biol Sci Med Sci 2020 Jul 13;75(8):1579-1585 [FREE Full text] [doi: 10.1093/gerona/glaa114] [Medline: 32384140]
2. Wang Y, Fu P, Li J, Jing Z, Wang Q, Zhao D, et al. Changes in psychological distress before and during the COVID-19 pandemic among older adults: the contribution of frailty transitions and multimorbidity. Age Ageing 2021 Jun 28;50(4):1011-1018 [FREE Full text] [doi: 10.1093/ageing/afab061] [Medline: 33710264]
3. Fujita K, Inoue A, Kuzuya M, Uno C, Huang CH, Umegaki H, et al. Mental health status of the older adults in Japan during the COVID-19 pandemic. J Am Med Dir Assoc 2021 Jan;22(1):220-221 [FREE Full text] [doi: 10.1016/j.jamda.2020.11.023] [Medline: 33321080]
4. Cao WJ, Fang ZW, Hou GQ, Han M, Xu XR, Dong JX, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287(12):112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 0005281912000036]
5. Jiang H, Burström B, Chen J, Burström K. Rural-urban inequalities in poor self-rated health, self-reported functional disabilities, and depression among Chinese older adults: evidence from the China Health and Retirement Longitudinal Study 2011 and 2015. Int J Environ Res Public Health 2021 Jun 18;18(12):6557 [FREE Full text] [doi: 10.3390/ijerph18126557] [Medline: 34207132]
6. Calderón-Larrañaga A, Vetrano DL, Welmer A, Grande G, Fratiglioni L, Dekhtyar S. Psychological correlates of multimorbidity and disability accumulation in older adults. Age Ageing 2019 Nov 01;48(6):789-796 [FREE Full text] [doi: 10.1093/ageing/afz117] [Medline: 31579908]
7. Feng L, Nyunt MSZ, Feng L, Yap KB, Ng TP. Frailty predicts new and persistent depressive symptoms among community-dwelling older adults: findings from Singapore longitudinal aging study. J Am Med Dir Assoc 2014 Jan;15(1):76.e7-76.e12. [doi: 10.1016/j.jamda.2013.10.001] [Medline: 24314697]
8. Nguyen QD, Wu C, Odden MC, Kim DH. Multimorbidity patterns, frailty, and survival in community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2019 Jul 12;74(8):1265-1270 [FREE Full text] [doi: 10.1093/gerona/gly205] [Medline: 30169580]
9. Read JR, Sharpe L, Modini M, Dear BF. Multimorbidity and depression: a systematic review and meta-analysis. J Affect Disord 2017 Oct 15;221:36-46. [doi: 10.1016/j.jad.2017.06.009] [Medline: 28628766]
10. Borges MK, Oude Voshaar RC, Romanini CFDV, Oliveira FM, Lima NA, Petrella M, et al. Could frailty be an explanatory factor of the association between depression and other geriatric syndromes in later life? Clin Gerontol 2021 Oct 25;44(2):143-153. [doi: 10.1080/07317115.2020.1836106] [Medline: 33100186]

11. Hanlon P, Nicholl BI, Jani BD, Lee D, McQueenie R, Mair FS. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. Lancet Public Health 2018 Jul;3(7):e323-e332. [doi: 10.1016/s2468-2667(18)30091-4]
12. Turner G, Clegg A, British Geriatrics Society, Age UK, Royal College of General Practioners. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age Ageing 2014 Nov;43(6):744-747 [FREE Full text] [doi: 10.1093/ageing/afu138] [Medline: 25336440]
13. Multimorbidity: clinical assessment and management (NICE Guideline 56). National Institute for Health and Care Excellence. 2016 Sep 21. URL: https://www.nice.org.uk/guidance/ng56 [accessed 2023-02-22]
14. Vetrano DL, Palmer K, Marengoni A, Marzetti E, Lattanzio F, Roller-Wirnsberger R, Joint Action ADVANTAGE WP4 Group. Frailty and multimorbidity: a systematic review and meta-analysis. J Gerontol A Biol Sci Med Sci 2019 Apr 23;74(5):659-666. [doi: 10.1093/gerona/gly110] [Medline: 29726918]
15. A conceptual framework for action on the social determinants of health. World Health Organization. 2010. URL: https://apps.who.int/iris/handle/10665/44489 [accessed 2023-02-22]
16. Brazil N. The multidimensional clustering of health and its ecological risk factors. Soc Sci Med 2022 Feb;295:113772. [doi: 10.1016/j.socscimed.2021.113772] [Medline: 33637329]
17. Datta GD, Subramanian S, Colditz GA, Kawachi I, Palmer JR, Rosenberg L. Individual, neighborhood, and state-level predictors of smoking among US Black women: a multilevel analysis. Soc Sci Med 2006 Aug;63(4):1034-1044. [doi: 10.1016/j.socscimed.2006.03.010] [Medline: 16650514]
18. Kawachi I, Subramanian S. Social epidemiology for the 21st century. Soc Sci Med 2018 Jan;196:240-245. [doi: 10.1016/j.socscimed.2017.10.034] [Medline: 29113687]
19. Ko P, Yeung WJ. An ecological framework for active aging in China. J Aging Health 2018 Dec 30;30(10):1642-1676. [doi: 10.1177/0898264318795564] [Medline: 30160571]
20. Li C, Tan C, Wu A, Gao C. COVID-19: the role of community in China's response. J R Soc Med 2020 Jul;113(7):280-281. [doi: 10.1177/0141076820935344] [Medline: 32663427]
21. Zadravec Šedivy N, Podlogar T, Kerr DCR, De Leo D. Community social support as a protective factor against suicide: a gender-specific ecological study of 75 regions of 23 European countries. Health Place 2017 Nov;48:40-46. [doi: 10.1016/j.healthplace.2017.09.004] [Medline: 28934635]
22. Norris FH, Stevens SP, Pfefferbaum B, Wyche KF, Pfefferbaum RL. Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. Am J Community Psychol 2008 Mar;41(1-2):127-150. [doi: 10.1007/s10464-007-9156-6] [Medline: 18157631]
23. Liddle J, Pitcher N, Montague K, Hanratty B, Standing H, Scharf T. Connecting at local level: exploring opportunities for future design of technology to support social connections in age-friendly communities. Int J Environ Res Public Health 2020 Jul 31;17(15):5544 [FREE Full text] [doi: 10.3390/ijerph17155544] [Medline: 32751898]
24. Riumallo-Herl CJ, Kawachi I, Avendano M. Social capital, mental health and biomarkers in Chile: assessing the effects of social capital in a middle-income country. Soc Sci Med 2014 Mar;105:47-58 [FREE Full text] [doi: 10.1016/j.socscimed.2013.12.018] [Medline: 24495808]
25. Murayama H, Fujiwara Y, Kawachi I. Social capital and health: a review of prospective multilevel studies. J Epidemiol 2012;22(3):179-187 [FREE Full text] [doi: 10.2188/jea.je20110128] [Medline: 22447212]
26. Berkman L, Kawachi I, Glymour M. Social Epidemiology (2 ed.). New York, NY: Oxford University Press; 2014.
27. Milner A, Krnjacki L, Butterworth P, LaMontagne AD. The role of social support in protecting mental health when employed and unemployed: a longitudinal fixed-effects analysis using 12 annual waves of the HILDA cohort. Soc Sci Med 2016 Mar;153:20-26. [doi: 10.1016/j.socscimed.2016.01.050] [Medline: 26867208]
28. Lakey B, Orehek E. Relational regulation theory: a new approach to explain the link between perceived social support and mental health. Psychol Rev 2011 Jul;118(3):482-495. [doi: 10.1037/a0023477] [Medline: 21534704]
29. Yates JA, Clare L, Woods RT, Cognitive Function and Ageing Study: Wales. "You've got a friend in me": can social networks mediate the relationship between mood and MCI? BMC Geriatr 2017 Jul 13;17(1):144 [FREE Full text] [doi: 10.1186/s12877-017-0542-0] [Medline: 28705185]
30. Matsuyama Y, Aida J, Hase A, Sato Y, Koyama S, Tsuboya T, et al. Do community- and individual-level social relationships contribute to the mental health of disaster survivors?: a multilevel prospective study after the Great East Japan Earthquake. Soc Sci Med 2016 Feb;151:187-195. [doi: 10.1016/j.socscimed.2016.01.008] [Medline: 26808337]
31. Jia Z, Xu S, Zhang Z, Cheng Z, Han H, Xu H, et al. Association between mental health and community support in lockdown communities during the COVID-19 pandemic: evidence from rural China. J Rural Stud 2021 Feb;82:87-97 [FREE Full text] [doi: 10.1016/j.jrurstud.2021.01.015] [Medline: 33437114]
32. Coronavirus disease 2019 (Covid-19): situation report-51. World Health Organization. 2020. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10 [accessed 2023-02-22]
33. Wang Y, Li J, Fu P, Jing Z, Zhou C. Social engagement and sleep quality among older adults in rural China: is there a sex difference? J Am Med Dir Assoc 2021 Oct;22(10):2140-2146. [doi: 10.1016/j.jamda.2020.11.014] [Medline: 33309649]

34. Wang Y, Li J, Fu P, Jing Z, Zhao D, Zhou C. Social support and subsequent cognitive frailty during a 1-year follow-up of older people: the mediating role of psychological distress. BMC Geriatr 2022 Feb 28;22(1):162 [FREE Full text] [doi: 10.1186/s12877-022-02839-5] [Medline: 35227216]
35. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
36. Zhou C, He J, Xu L, Sun H, Zheng W, Wang X. First application of Kessler 10 scale to measuring mental health status of the aged in China (in Chinese). Chin J Clin Psychol 2009;17(006):761-763. [doi: 10.16128/j.cnki.1005-3611.2009.06.003]
37. Zhou C, Chu J, Wang T. Reliability and validity of 10-item Kessler scale (K10) Chinese version in evaluation of mental health status of Chinese population (in Chinese). Chin J Clin Psychol 2008;16(6):9. [doi: 10.16128/j.cnki.1005-3611.2008.06.026]
38. Xiao S. Theoretical basis and research applications of social support rating scale (in Chinese). J Clin Psychiatry 1994;4(2):98-100.
39. Zhao Z, Jing S, Yan Z, Yu L. Social change and birth cohort decrease in social support for older adults in China: a cross-temporal meta-analysis, 1994-2018. Health Soc Care Community 2020 Sep 06;28(5):1438-1447. [doi: 10.1111/hsc.13004] [Medline: 32378280]
40. Lin J, Su Y, Lv X, Liu Q, Wang G, Wei J, et al. Perceived stressfulness mediates the effects of subjective social support and negative coping style on suicide risk in Chinese patients with major depressive disorder. J Affect Disord 2020 Mar 15;265:32-38. [doi: 10.1016/j.jad.2020.01.026] [Medline: 31959583]
41. Kawachi I, Subramanian SV. Measuring and modeling the social and geographic context of trauma: a multilevel modeling approach. J Trauma Stress 2006 Apr;19(2):195-203. [doi: 10.1002/jts.20108] [Medline: 16612828]
42. Hikichi H, Aida J, Matsuyama Y, Tsuboya T, Kondo K, Kawachi I. Community-level social capital and cognitive decline after a natural disaster: a natural experiment from the 2011 Great East Japan Earthquake and Tsunami. Soc Sci Med 2020 Jul;257:111981. [doi: 10.1016/j.socscimed.2018.09.057] [Medline: 30293854]
43. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar 01;56(3):M146-M157. [doi: 10.1093/gerona/56.3.M146] [Medline: 5637250]
44. Ensrud KE, Kats AM, Schousboe JT, Taylor BC, Vo TN, Cawthon PM, Osteoporotic Fractures in Men Study (MrOS). Frailty phenotype and healthcare costs and utilization in older men. J Am Geriatr Soc 2020 Sep;68(9):2034-2042 [FREE Full text] [doi: 10.1111/jgs.16522] [Medline: 32402097]
45. Chronic non-communicable diseases. Chinese Center for Disease Control and Prevention. 2020. URL: http://www.chinacdc.cn/jkzt/mxfcrjbhsh/ [accessed 2023-02-22]
46. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev 2011 Sep;10(4):430-439. [doi: 10.1016/j.arr.2011.03.003] [Medline: 21402176]
47. Yao S, Cao G, Han L, Chen Z, Huang Z, Gong P, et al. Prevalence and patterns of multimorbidity in a nationally representative sample of older Chinese: results from the China Health and Retirement Longitudinal Study. J Gerontol A Biol Sci Med Sci 2020 Sep 25;75(10):1974-1980. [doi: 10.1093/gerona/glz185] [Medline: 31406983]
48. Jin Y, Si H, Qiao X, Tian X, Liu X, Xue Q, et al. Relationship between frailty and depression among community-dwelling older adults: the mediating and moderating role of social support. Gerontologist 2020 Nov 23;60(8):1466-1475. [doi: 10.1093/geront/gnaa072] [Medline: 32556208]
49. Kikuchi H, Machida M, Nakamura I, Saito R, Odagiri Y, Kojima T, et al. Changes in psychological distress during the COVID-19 pandemic in Japan: a longitudinal study. J Epidemiol 2020 Nov 05;30(11):522-528 [FREE Full text] [doi: 10.2188/jea.JE20200271] [Medline: 32963212]
50. López J, Perez-Rojo G, Noriega C, Carretero I, Velasco C, Martinez-Huertas J, et al. Psychological well-being among older adults during the COVID-19 outbreak: a comparative study of the young-old and the old-old adults. Int Psychogeriatr 2020 Nov 22;32(11):1365-1370 [FREE Full text] [doi: 10.1017/S1041610220000964] [Medline: 32438934]
51. Zheng J, Morstead T, Sin N, Klaiber P, Umberson D, Kamble S, et al. Psychological distress in North America during COVID-19: the role of pandemic-related stressors. Soc Sci Med 2021 Feb;270:113687 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113687] [Medline: 33465600]
52. Woo J, Leung J. Multi-morbidity, dependency, and frailty singly or in combination have different impact on health outcomes. Age (Dordr) 2014 Apr 3;36(2):923-931 [FREE Full text] [doi: 10.1007/s11357-013-9590-3] [Medline: 24091565]
53. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 02;381(9868):752-762 [FREE Full text] [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
54. Wang H, Zhang M, Li R, Zhong O, Johnstone H, Zhou H, et al. Tracking the effects of COVID-19 in rural China over time. Int J Equity Health 2021 Jan 14;20(1):35 [FREE Full text] [doi: 10.1186/s12939-020-01369-z] [Medline: 33446205]
55. Silver A. Could rural China's healthcare deal with covid-19? BMJ 2021 Nov 25;375:n2759. [doi: 10.1136/bmj.n2759] [Medline: 34824132]
56. Nakagomi A, Shiba K, Hanazato M, Kondo K, Kawachi I. Soc Sci Med 2020 Aug;259:113140. [doi: 10.1016/j.socscimed.2020.113140] [Medline: 32645611]

57. Ransome Y, Ojikutu BO, Buchanan M, Johnston D, Kawachi I. Neighborhood social cohesion and inequalities in COVID-19 diagnosis rates by area-level Black/African American racial composition. J Urban Health 2021 Apr 23;98(2):222-232 [FREE Full text] [doi: 10.1007/s11524-021-00532-3] [Medline: 33759068]

Abbreviations

CSS: community-level social support
K10: 10-item Kessler Psychological Distress Scale
NICE: National Institute for Health and Care Excellence
SSRS: Social Support Rating Scale

Edited by A Mavragani, T Sanchez; submitted 23.10.22; peer-reviewed by C Li, Z Huo; comments to author 25.01.23; revised version received 07.02.23; accepted 14.02.23; published 09.03.23.

Please cite as:
Wang Y, Fu P, Li J, Gao T, Jing Z, Wang Q, Zhao D, Zhou C
Community-Level Social Support to Mitigate the Impact of Combined Frailty and Multimorbidity on Psychological Distress Among Rural Chinese Older Adults During the COVID-19 Pandemic: Multilevel Modeling Study
JMIR Public Health Surveill 2023;9:e43762
URL: https://publichealth.jmir.org/2023/1/e43762
doi:10.2196/43762
PMID:36811848

Original Paper

The Impact of the COVID-19 Pandemic on the Registration and Care Provision of Mental Health Problems in General Practice: Registry-Based Study

Jan Vandamme[1*], MD; Simon Gabriël Beerten[1*], MD, MSc; Jonas Crèvecoeur[2,3], PhD; Steve Van den Bulck[1], MD, PhD; Bert Aertgeerts[1], MD, PhD; Nicolas Delvaux[1], MD, PhD; Gijs Van Pottelbergh[1], MD, PhD; Mieke Vermandere[1], MD, PhD; Laura Tops[1], MSc; Thomas Neyens[2,3], PhD; Bert Vaes[1], MD, PhD

[1]Academic Center for General Practice, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium
[2]Leuven Biostatistics and Statistical Bioinformatics Centre, Faculty of Medicine, KU Leuven, Leuven, Belgium
[3]Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Data Science Institute, UHasselt, Hasselt, Belgium
[*]these authors contributed equally

Corresponding Author:
Simon Gabriël Beerten, MD, MSc
Academic Center for General Practice
Department of Public Health and Primary Care
KU Leuven
Kapucijnenvoer 7 blok h
bus 7001
Leuven, 3000
Belgium
Phone: 32 16 19 40 50
Email: simon.beerten@kuleuven.be

Abstract

Background: The impact of the COVID-19 pandemic on mental health in general practice remains uncertain. Several studies showed an increase in terms of mental health problems during the pandemic. In Belgium, especially during the first waves of the pandemic, access to general practice was limited. Specifically, it is unclear how this impacted not only the registration of mental health problems itself but also the care for patients with an existing mental health problem.

Objective: This study aimed to know the impact of the COVID-19 pandemic on (1) the incidence of newly registered mental health problems and (2) the provision of care for patients with mental health problems in general practice, both using a pre–COVID-19 baseline.

Methods: The prepandemic volume of provided care (*care provision*) for patients with mental health problems was compared to that from 2020-2021 by using INTEGO, a Belgian general practice morbidity registry. Care provision was defined as the total number of new registrations in a patient’s electronic medical record. Regression models evaluated the association of demographic factors and care provision in patients with mental health problems, both before and during the pandemic.

Results: During the COVID-19 pandemic as compared to before the COVID-19 pandemic, the incidence of registered mental health problems showed a fluctuating course, with a sharp drop in registrations during the first wave. Registrations for depression and anxiety increased, whereas the incidence of registered eating disorders, substance abuse, and personality problems decreased. During the 5 COVID-19 waves, the overall incidence of registered mental health problems dropped during the wave and rose again when measures were relaxed. A relative rise of 8.7% and 40% in volume of provided care, specifically for patients with mental health problems, was seen during the first and second years of the COVID-19 pandemic, respectively. Care provision for patients with mental health problems was higher in older patients, male patients, patients living in center cities (*centrumsteden*), patients with lower socioeconomic status (SES), native Belgian patients, and patients with acute rather than chronic mental health problems. Compared to prepandemic care provision, a reduction of 10% was observed in people with a low SES.

Conclusions: This study showed (1) a relative overall increase in the registrations of mental health problems in general practice and (2) an increase in care provision for patients with mental health problems in the first 2 years of the COVID-19 pandemic. Low SES remained a determining factor for more care provision, but care provision dropped significantly in people with mental health problems with a low SES. Our findings suggest that the pandemic in Belgium was also largely a “syndemic,” affecting different layers of the population disproportionately.

(JMIR Public Health Surveill 2023;9:e43049) doi:10.2196/43049

KEYWORDS

COVID-19; mental health; care provision; general practice; socioeconomic status

Introduction

Similar to most European countries during the COVID-19 pandemic, Belgium has experienced multiple COVID-19 waves in the period from 2020-2021. Measures were taken to reduce the COVID-19 burden during these waves. The reduction of social contacts, self-isolation, and quarantine were the main nonpharmaceutical interventions [1].

Several studies showed an increase in mental distress during the pandemic [2-8]. This increase has, among others, been attributed to decreased social contact [9], financial uncertainty [10], and the fear of catching COVID-19 [11].

Surveys in Belgium showed a similar trend [12,13], as did multiple studies in the United Kingdom [14-16]. Interestingly, some studies showed a significant correlation between the incidence of depressive symptoms and the measures that were taken [17]. Government measures implemented during waves of infections tended to ease depressive symptoms [17].

During surges of COVID-19 infections, assessing mental health status in general practice was challenging. The imposed mitigation measures decreased access to routine care, and the increased burden of care for patients with COVID-19 left little room for treating other conditions such as mental health problems. Although some studies showed an increase in the number of anxiety-related visits [18], most studies reported a reduction in overall mental health consultations [19-22]. Additionally, there was evidence for reduced access to mental health care for patients with preexisting mental health problems [23].

Most studies on mental health during the COVID-19 pandemic were cross-sectional in nature and often showed a disparity between different nations and risk groups [24-26].

Thus, longitudinal data and a pre–COVID-19 baseline are often missing in international literature. These are important tools for countries to monitor their population’s mental health during a pandemic [27].

We aimed to fill this gap and provide a longitudinal overview of how the COVID-19 pandemic was experienced in Belgian general practice in terms of registered mental health problems (both during and before the pandemic) and what factors led to an increased or decreased volume of care for these patients.

Therefore, we evaluated the impact of the COVID-19 pandemic on (1) the registration of mental health problems and (2) care provision for patients with existing mental health problems in general practice. Our hypotheses were that the number of registrations of mental health problems would increase during the pandemic, relative to the years before, and that care provision for patients with existing mental health problems would drop, given the more limited access to care.

Methods

INTEGO

This was a retrospective cohort study using INTEGO, a general practice morbidity registry in Flanders, the Dutch-speaking region of Belgium [28]. The study methods have been described earlier, also in the context of COVID-19, and were adjusted for the current study [29]. The data from INTEGO is largely representative for the Flemish population [28].

Our study population was defined as all patients (of all ages) who visited 1 of the 105 INTEGO practices between February 1, 2018, and January 31, 2022, and consisted of 397,489 patients. Data were automatically collected via the electronic health record of the participating practices and included patient characteristics, medical history, and socioeconomic status (SES). Available demographic characteristics included the year of birth, sex, nationality (native Belgian or nonnative patients), and place of residence (postal code). Diagnoses were coded using the International Classification of Primary Care, second edition (ICPC-2). Low SES was defined as the presence of an increased reimbursement under the national health care insurance policy. In Belgium, patients effectively only pay a small, fixed amount of the total fees for medical care and medication (copay or *remgeld* in Dutch); the rest of the cost is reimbursed by health insurance. The reimbursement for patients with a low SES is higher, thus making this fixed amount lower.

Collection of Data and Intervention Period

We evaluated the impact of the COVID-19 pandemic by (1) comparing the incidence of registered mental health diagnoses and (2) comparing care provision for patients with mental health problems 2 years before the pandemic (ie, February 1, 2018, to January 31, 2020) with the first 2 years of the pandemic (ie, February 1, 2020, to January 31, 2022). In this period, Belgium experienced 5 COVID-19 waves, approximately around the following dates: first wave (March 1 to May 31, 2020), second wave (October 1 to November 30, 2020), third wave (March 1 to May 31, 2021); fourth wave (October 1 to the end of 2021); and fifth (Omicron) wave (started December 2021) [30].

There was an issue with some data for care provision at the end of 2021 due to a data flow error from the electronic medical software to our database. This is indicated in the corresponding figure below.

Outcome Measures

Registration of Mental Health Diagnoses

For the first part of the study, we examined the difference in the registration of incident mental health problems during and before the COVID-19 pandemic. “Incident” here means registered for the first time since the start of the study period.

All psychological ICPC-2 codes (P-codes) were categorized into 7 subgroups related to depression; anxiety; psychosis; eating problems, substance abuse, and personality problems; sexuality problems; sleeping problems; and other mental health problems. Eating problems, substance abuse, and personality problems were pooled because of the insufficiently high patient counts necessary for statistical analysis. For each of these subgroups, a separate analysis was carried out.

We then further distinguished between acute and chronic mental health problems based on the registered ICPC-2 code [31]. This information can be found in Table S1 in Multimedia Appendix 1.

Evolution of Care Provision

The second part of the study examined the evolution of care provision during the pandemic, compared to the 2 years before.

Care provision was defined as the number of new data entries added to one’s medical history over a predefined time window. These entries included medication prescriptions, laboratory test results, measurements of medical parameters, diagnoses, physical therapy referrals, and radiology orders. In the computation of care provision, multiple laboratory tests or medication prescriptions registered on the same date were counted as one entry.

Care provision was investigated in patients with an acute mental problem, patients with an incident chronic mental problem, and patients with a prevalent chronic mental problem. An acute mental problem was defined as an acute P-code that was present at the beginning of the defined period or incident during this period. An incident chronic mental problem described a chronic P-code that was newly registered during the defined period or had its initial registration no more than 16 weeks before this period. A prevalent chronic mental problem implied the registration of a chronic P-code that was made >16 weeks before the defined period.

Statistical Analysis

Relative care provision was computed as the ratio between care provision in the first 2 years of the COVID-19 pandemic and care provision during the previous 2 years. More granular, biweekly care data were obtained by comparing the same calendar weeks within both periods. CIs for relative care provisions were computed by assuming a normal distribution for the logarithm of the relative care provision.

Generalized linear models analyzed care provision per patient with an existing mental health problem (ie, with a registered P-code) per year as a function of the covariates age, sex, nationality, center city, and SES, with the following reference categories as a basis for comparison in the model: aged 18-35 years, female, Belgian, living outside center cities, and high SES. Center cities (*centrumsteden*) are urban areas in Flanders with high population density and a central economic and cultural function in its surrounding area.

We compared care provision for the 3 predefined P-code groups as described above (acute, incident chronic, and prevalent chronic). Besides these demographic covariates, we included in the model the general practice that the patients visited to correct for registration differences across practices. We chose a fixed practice effect over a random effect as the distribution of these effects is strongly nonnormal and the high number of patients per practice allows for an accurate estimation. We modelled care provision using a Poisson distribution with log link. The model estimated the change in care provision for the reference category in the first and second years of the COVID-19 pandemic and the multiplicative effect of demographic covariates on care provision both in the period before the pandemic and during the first 2 years of the pandemic. This approach allowed us to assess the overall impact of the pandemic as well as how this impact differs for various demographic groups in the population. Statistical analysis was conducted in R statistical software (version 4.0.3; R Foundation for Statistical Computing).

Ethical Approval

The INTEGO procedures were approved by the Ethics Committee Research UZ/KU Leuven (nr. ML1723) and by the Belgian National Privacy Commission’s Sectoral Committee (decision nr. 13.026 of March 19, 2013, last modified on April 17, 2018).

INTEGO was waived the need for individual informed consent, instead using an opt-out procedure for patients who do not want their data to be included in the database. This procedure was approved by the aforementioned ethical review board. There was no compensation or financial incentive provided for inclusion in the INTEGO database.

The procedures outlined in this study are in accordance with current Belgian guidelines and regulations. All study data were pseudonymized before use.

Results

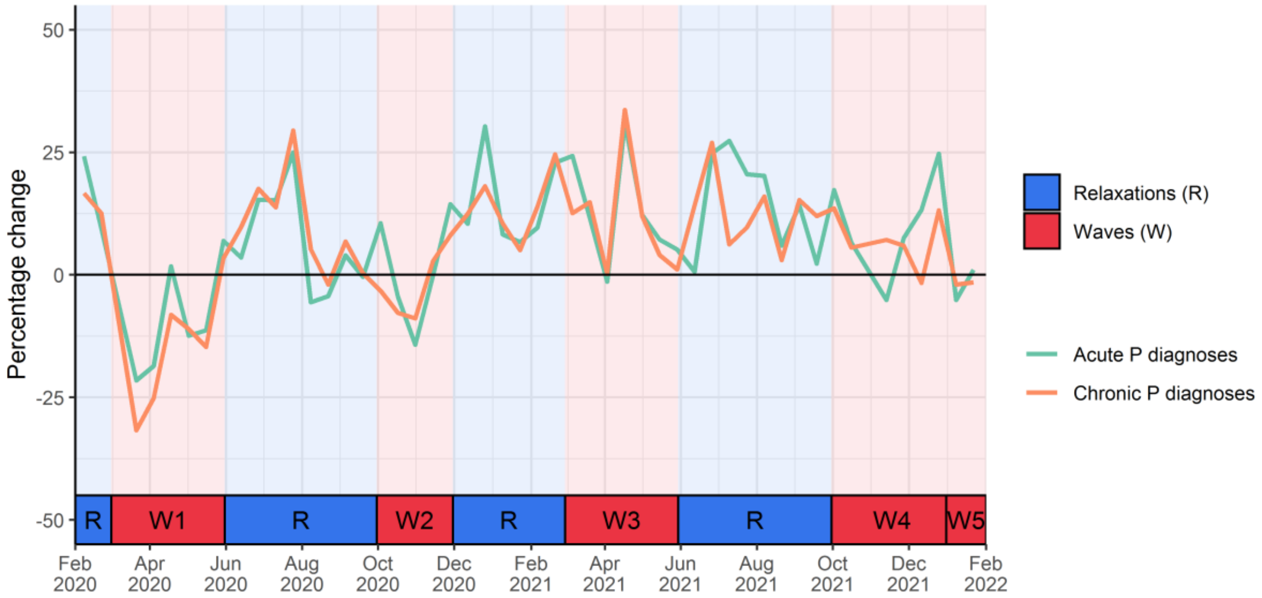
Evolution of the Registration of Incident Acute and Chronic Mental Health Problems

General

The incidence of registered acute and chronic mental health disorders showed a fluctuating course (Figure 1). This analysis is split into acute and chronic mental health diagnoses based on the ICPC-P codes (see Table S1 in Multimedia Appendix 1). Shortly before the first wave, the incidence of both acute and chronic mental health disorders was between 20% and 25% higher compared to the 2 years before, and this dropped to almost 25% less at the beginning of the first wave. When measures were relaxed, registrations rose again compared to the previous years until the end of July and normalized from August to October. During the second wave, the incidence of registered acute and chronic mental health disorders dropped again but less compared to the first wave. After the second wave, the incidence of registered disorders increased, especially for acute disorders. This fluctuating course became less pronounced in the second COVID-19 year, but especially for the acute disorders, the increase during relaxations can still be seen.

The evolution of the registration for the different subgroups during the study period can be seen in Figure S1 in Multimedia Appendix 1. The patient characteristics are summarized in Table S2 in Multimedia Appendix 1.

Figure 1. Evolution of incident acute and chronic mental health diagnoses (registered P-codes) during the first 2 years of the COVID-19 pandemic (February 1, 2020, to January 31, 2022) relative to the 2 previous years. P: psychological; P-code: psychological International Classification of Primary Care, second edition, code.



Mood Problems and Anxiety

During the first year of the COVID-19 pandemic, the registration of mood problems and anxiety increased significantly by 4% (95% CI 1%-8%) and 9% (95% CI 6%-13%), respectively (Table 1). During the first wave, the registration of anxiety rose significantly by 22% (95% CI 15%-30%). After the second wave, both mood problems and other mental health problems showed a significant incidence rise of 14% (95% CI 6%-22%) and 9% (95% CI 4%-14%).

For the second COVID-19 year, the results were similar. Registrations of anxiety and mood disorders rose by the same amount, although the incidence rates declined significantly around late 2021.

Table 1. The average number of registered mental health diagnoses per 1000 patient-years during the first and second year of the COVID-19 pandemic and percentage change in registered incidence relative to the 2 prepandemic years along with 95% CIs. Results are stratified by calendar period and P-code[a] groups.

Group and year	February		March to May		June to September		October to November		December to January		Total	
	Mean	Percentage change (%; 95% CI)	Mean	Percentage change (%; 95% CI)	Mean	Percentage change (%; 95% CI)	Mean	Percentage change (%; 95% CI)	Mean	Percentage change (%; 95% CI)	Mean	Percentage change (%; 95% CI)
Mood problems												
1	20.48	10 (0 to 22)	17.99	–11 (–16 to –5)	21.13	14 (8 to 20)	22.07	0 (–6 to 7)	25.14	14 (6 to 22)	20.78	4 (1 to 8)
2	20.82	12 (1 to 24)	22.29	9 (3 to 15)	21.32	15 (9 to 21)	20.4	–2 (–8 to 3)	21.81	–12 (–20 to –3)	21	6 (3 to 9)
Anxiety												
1	17.53	3 (–7 to 16)	21.28	22 (15 to 30)	18.38	11 (5 to 17)	18.63	0 (–8 to 7)	19.84	2 (–4 to 10)	19.04	9 (6 to 13)
2	18.36	8 (–2 to 21)	20.59	19 (12 to 26)	18.49	12 (6 to 18)	16.8	–7 (–12 to –1)	18.7	–13 (–22 to –4)	18.35	5 (2 to 9)
Psychosis												
1	2.36	19 (–13 to 63)	2.03	2 (–15 to 24)	2.33	20 (2 to 40)	2.38	9 (–11 to 36)	2.58	0 (–18 to 22)	2.3	9 (0 to 20)
2	2.15	9 (–20 to 50)	2.23	12 (–5 to 35)	2.27	17 (0 to 37)	2.16	–2 (–18 to 16)	2.26	–21 (–41 to 5)	2.22	6 (–3 to 15)
Sleeping problems												
1	15.39	22 (8 to 38)	10.99	–16 (–23 to –9)	13.95	4 (–2 to 10)	15.25	–5 (–12 to 3)	17.99	5 (–2 to 14)	14.07	0 (–4 to 2)
2	14.28	14 (0 to 29)	14.8	11 (4 to 20)	14.49	8 (2 to 15)	13.71	–9 (–15 to –2)	16.01	–22 (–30 to –13)	14.33	1 (–2 to 4)
Eating problems, substance abuse, and personality problems												
1	7.05	6 (–11 to 27)	4.47	–33 (–41 to –24)	5.79	–20 (–27 to –11)	6.91	–10 (–21 to 1)	6.8	–13 (–23 to –1)	5.89	–18 (–22 to –13)
2	5.79	–12 (–27 to 6)	6.61	0 (–10 to 9)	6.42	–10 (–18 to –2)	5.15	–30 (–37 to –21)	5.93	–33 (–44 to –20)	6.03	–15 (–20 to –11)
Sexuality problems												
1	2.27	–4 (–29 to 30)	2.57	15 (–2 to 37)	2.84	15 (0 to 33)	2.8	–4 (–21 to 16)	2.9	1 (–16 to 23)	2.72	7 (0 to 16)
2	3	26 (–3 to 66)	2.63	19 (0 to 40)	2.61	6 (–7 to 23)	2.56	–8 (–22 to 7)	2.79	–10 (–31 to 16)	2.65	4 (–3 to 13)
Other mental health problems												
1	45.28	23 (14 to 33)	32.18	–21 (–24 to –17)	39.66	9 (5 to 13)	46.38	–4 (–9 to 0)	48.89	9 (4 to 14)	39.66	0 (–2 to 2)
2	47.32	29 (20 to 39)	49.6	21 (17 to 26)	45.25	25 (20 to 29)	44.38	–1 (–5 to 2)	47.6	–6 (–12 to 0)	44.92	13 (11 to 15)

[a] P-code: psychological International Classification of Primary Care, second edition, code.

Eating Disorders, Substance Abuse, and Personality Problems

The registration of eating problems, substance abuse, and personality problems showed a significant decrease of 18% (95% CI 13%-22%) in the first COVID-19 year and a decrease of 15% (95% CI 11%-20%) in the second year. During these 2 years, there were no significant fluctuations but rather a steady decrease in registrations each period, regardless of wave.

Psychosis, Sexuality, and Sleeping Problems

The incidence of registered psychosis and sexuality problems increased nonsignificantly during both the first and second COVID-19 years. The month before the first wave was characterized by an increased registration for all groups, except sexuality problems. The registration of sleeping problems then was 22% (95% CI 8%-35%) higher compared to the same period in the 2 previous years.

Evolution of Care Provision for Patients With Mental Health Problems and the Effect of Demographic Factors

Before the start of the pandemic, almost no difference in care provision for the 3 groups (acute, incident chronic, and prevalent chronic mental health problems) was observed compared to the 2 years before COVID-19. In this analysis, patients with a recent chronic P-code diagnosis were split from those who received their diagnoses longer ago to see whether additional care was mainly given close to the onset of diagnosis or whether older chronic P-code diagnoses also resulted in continued care. Care provision almost stayed the same at the start of the first wave, but provision rose for all 3 groups during the first relaxation period. Starting from the end of May 2020, overall care provision remained higher than in previous years (Figure 2).

In general, there was 8.7% (95% CI 8.4%-9.1%) more care provided to patients in the reference category during the first year of COVID-19 compared to the 2 years before the pandemic. In the second year, this increased to 38.3% (95% CI 37.2%-39.5%).

In general, more care was provided for acute mental health problems and in older patients, female patients, native Belgian patients, patients living in center cities, and patients with a low SES (Figure 3).

Care provision for foreign patients somewhat increased compared to the 2 years before COVID-19 (although lower in 2021 than in 2020) but stayed slightly less than native Belgian patients. In patients with a low SES, we observed more care provision than for those with a high SES during the study period, although care provision dropped 10% for people with a low SES during the pandemic compared to the 2 years before.

Compared to patients with acute mental health problems, patients with incident chronic and prevalent chronic mental health problems had less care provision (around 10% and 30% less, respectively). This was similar to the period before COVID-19.

Figure 2. Evolution of care provision for patients with a registered acute, incident chronic, and prevalent chronic mental problem during the first 2 years of the COVID-19 pandemic (February 1, 2020, to January 31, 2022) relative to the 2 years before. Gray bar: no data due to technical issues with data flow.

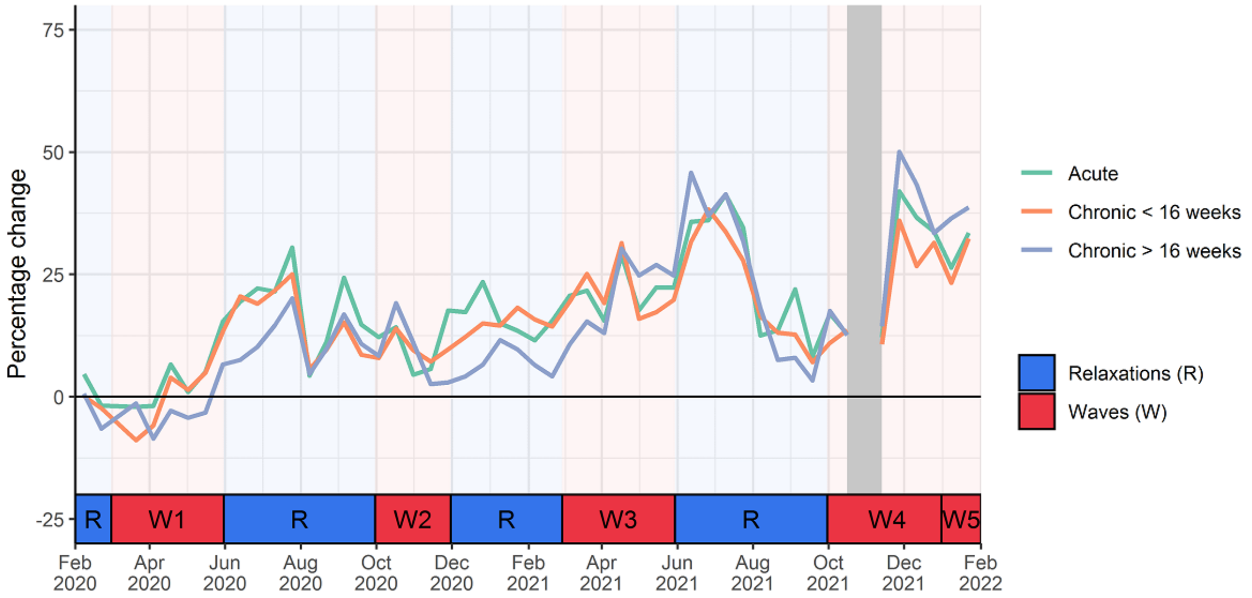
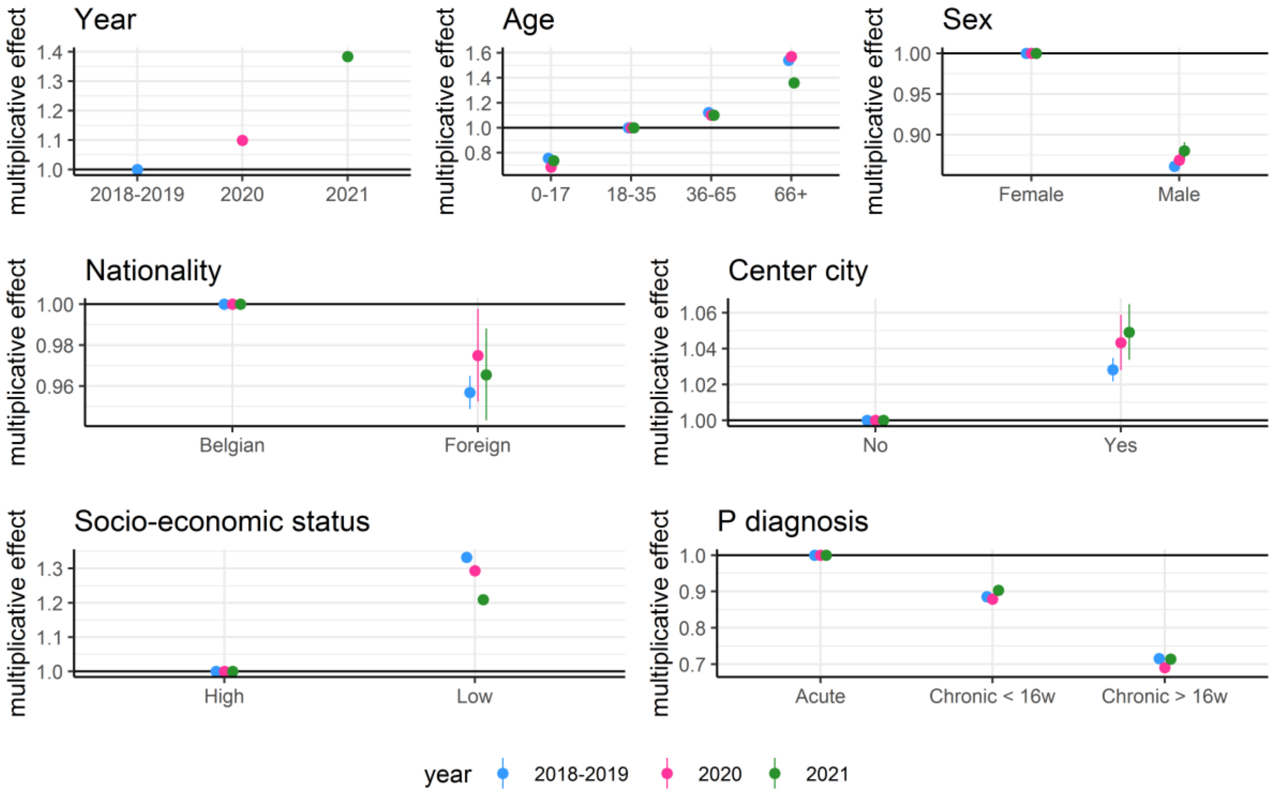


Figure 3. Multiplicative effect and associated 95% CIs of covariates on care provision for mental health problems per patient per year. Separate parameters were estimated for the period February 1, 2018, to January 31, 2020, and the period including the 2 COVID-19 years (February 1, 2020, to January 31, 2022), which are labelled 2018-2019, 2020 and 2021, respectively.



Discussion

Principal Findings

This registry-based study showed that the first 2 years of the COVID-19 pandemic had significant repercussions on the registration of new mental health problems and care provision for patients with mental health problems in general practice.

Waves of infections resulted in a drop in registered mental health problems and an increase of up to 25% during relaxation periods when compared to the previous years. This phenomenon was less pronounced in the second COVID-19 year.

With a rise of around 9% in the first COVID-19 year and almost 40% in the second year, overall care provision for patients with mental health problems increased during the first 2 years of the pandemic. In general, care provision was 50% higher among those aged 66 years and over, around 4% higher for patients living in center cities, and approximately 30% higher for those with a lower SES; it was lower among male (–10%) and nonnative Belgian (–4%) patients. Patients with acute mental health problems had the most care provided to them with 10% to 30% more registrations than patients with chronic mental health problems.

During the pandemic, care provision dropped by 10% in patients with a lower SES.

Comparison With Literature

There currently is little evidence on the care provision for patients with preexisting mental health problems during COVID-19, especially in general practice [23]. A similar UK-based primary care cohort study showed the same overall drop in the registration of mental health problems at the start of the pandemic and the same normalization of these rates after April 2020 [19]. We extended this study period and saw a continuing relative increase in registrations of new mental health problems during the pandemic as compared to before. This increase was fluctuating and corresponded to the specific COVID-19 waves in Belgium.

The multivariate analysis for our definition of care provision that we conducted in this study was new. We provided more evidence to the hypothesis that low SES affected the outcomes disproportionately, despite these patients having the highest recorded diagnoses of depression, anxiety, and self-harm before the pandemic [19]. The overall increase in registered mood problems and anxiety matched the rise in mental distress seen in surveys [12,14-16].

Annual reports of Sciensano (Belgium’s national public health institute) and Belgium’s Superior Health Council signaled a high prevalence of sleep disturbance, anxiety, and depressive disorders during the COVID-19 pandemic [12,30,32]. Our finding that the registration of mood disorders rose when strict measures were taken is confirmed by the report of the Superior Health Council [32]. The impact was more pronounced during stringent or long-lasting measures, for instance, due to the existing uncertainty and lack of social interaction. However, during waves with high viral spread, it is not inconceivable that general practitioners were overwhelmed by COVID-19–related care, leaving little room for other clinical work. This could have led to an underestimation of the actual burden of mental health problems during such periods.

Additionally, both Belgian reports also noted different groups being at risk for more mental health problems, such as adolescents, single individuals, people entitled to social security payments, people with low SES, or people with known mental health problems [12,32], which was in keeping with our findings.

Interestingly, we found that care provision for people with lower SES *decreased* during the pandemic.

Strengths and Limitations

The main strength of this study was the use of longitudinal data from a large general practice registry, which allowed us to investigate a large sample of the Flemish population. We calculated a proxy for the SES of the patients in INTEGO, and the correlation of this social determinant could be further assessed. Since INTEGO weekly and automatically collects data via the electronic health record of the participating general practices, further monitoring is easily accessible.

There are however some limitations to this study. Data were collected from 105 practices that are individually responsible for delivering qualitative and complete data. Possible registration errors were minimized by accounting for interpractice differences, and steps have been taken to improve data completeness [33]. However, there might be a general underregistration of mental health problems in general practice (especially subclinical problems) that we were unable to pick up due to them not being coded.

Our figures for registrations of mental health problems, especially during the first wave, might be underestimations because of patients simply not visiting general practices due to a fear of COVID-19 in general.

Furthermore, we defined care provision as the number of contacts added to a patients' medical record, but this does not necessarily reflect the quality of each contact in regard to the depth of care provision. Additionally, a decrease in care provision does not necessarily imply a proportional change in the content of care (eg, a decrease might be a good thing if there are fewer unnecessary prescriptions but a bad thing if it leads to fewer referrals to specialized care for more severe problems). Finally, although the grouping of P-codes was done after consensus, the classification may be somewhat arbitrary, since some mental health problems could not be studied separately, as we opted to pool them due to insufficiently high patient counts.

It should also be noted that teleconsultations may be a useful and viable alternative to a live consultation for mental health problems in times of pandemic [34]. Our data could not distinguish between a live consultation and a teleconsultation.

Implications for Practice and Future Research

Obviously, the measures taken in the first 2 years of the COVID-19 pandemic have had a noticeable effect on mental health and care provision for mental health problems in Flanders. The negative impact may also have been compounded by what the Superior Health Council labelled as the "infodemic"—in which negative topics about COVID-19 were disseminated in an unmotivated and almost autocratic way [32]. Governmental policies with a commitment to improve nationwide resilience and social security, using efficient and accessible health care resources, must be a priority [35]. It is clear that communication about the pandemic must be given clearly and at the right time, particularly to the most vulnerable groups of the population [32].

The decreasing trend in care provision for socioeconomically vulnerable patients during the COVID-19 pandemic is a notable finding, deserving of further study. Similar to the infodemic discussed earlier, COVID-19 has also been regarded by some experts as a "syndemic" [36,37]. A syndemic in this context refers to a pandemic exacerbated by—previously disregarded or undervalued—social determinants of health: housing, income, neighborhood deprivation, and other social risk factors for disease. Although COVID-19 might not be regarded as a syndemic everywhere in the world [38], based on our results, we would argue that this might be the case in Belgium.

Conclusions

Our study found that the incidence of new registrations for mental health problems was generally higher during the COVID-19 pandemic than in the 2 years before. There was a slight decrease during the first wave, after which the incidence continued to climb, following a fluctuating course.

In addition, we studied care provision for patients with established mental health problems in general practice during the first 2 years of the COVID-19 pandemic, compared to the 2 years before. Care provision was generally higher during the pandemic than before. Low SES remained a determining factor for more care provision, but care provision dropped significantly in people with mental health problems with a low SES.

Based on our findings, we would suggest that the pandemic in Belgium was also largely a "syndemic," affecting different layers of the population disproportionately.

Acknowledgments

The authors wish to thank Sophie Wen, MD, for critically proofreading the manuscript.

This work was supported by a research project from the Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek [FWO]). TN gratefully acknowledges funding by the Internal Funds KU Leuven (project number 3M190682).

These funding sources had no role in the study design, collection, analysis, or interpretation of the data, nor in writing the manuscript or the decision for submission.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the inclusion of protected health information but can be made available after deidentification upon reasonable request to the corresponding author (SGB).

Conflicts of Interest

None declared.

Multimedia Appendix 1

The list of ICPC-2 groups, evolution of registered P-codes, and distribution of patient characteristics. ICPC-2: International Classification of Primary Care, second edition; P-code: psychological ICPC-2 code.

[PDF File (Adobe PDF File), 410 KB - publichealth_v9i1e43049_app1.pdf]

References

1. Dzúrová D, Květoň V. How health capabilities and government restrictions affect the COVID-19 pandemic: cross-country differences in Europe. Appl Geogr 2021 Oct;135:102551 [FREE Full text] [doi: 10.1016/j.apgeog.2021.102551] [Medline: 34456395]
2. Hossain MM, Sultana A, Purohit N. Mental health outcomes of quarantine and isolation for infection prevention: a systematic umbrella review of the global evidence. Epidemiol Health 2020 Jun 2;42:e2020038 [FREE Full text] [doi: 10.4178/epih.e2020038] [Medline: 32512661]
3. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
4. Farooq S, Tunmore J, Wajid Ali M, Ayub M. Suicide, self-harm and suicidal ideation during COVID-19: a systematic review. Psychiatry Res 2021 Dec;306:114228 [FREE Full text] [doi: 10.1016/j.psychres.2021.114228] [Medline: 34670162]
5. Brenna CT, Links PS, Tran MM, Sinyor M, Heisel MJ, Hatcher S. Innovations in suicide assessment and prevention during pandemics. Public Health Res Pract 2021 Sep 08;31(3):3132111 [FREE Full text] [doi: 10.17061/phrp3132111] [Medline: 34494071]
6. Schmidt RA, Genois R, Jin J, Vigo D, Rehm J, Rush B. The early impact of COVID-19 on the incidence, prevalence, and severity of alcohol use and other drugs: a systematic review. Drug Alcohol Depend 2021 Nov 01;228:109065 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.109065] [Medline: 34600257]
7. Henssler J, Stock F, van Bohemen J, Walter H, Heinz A, Brandt L. Mental health effects of infection containment strategies: quarantine and isolation-a systematic review and meta-analysis. Eur Arch Psychiatry Clin Neurosci 2021 Mar;271(2):223-234 [FREE Full text] [doi: 10.1007/s00406-020-01196-x] [Medline: 33025099]
8. Généreux M, Schluter PJ, Landaverde E, Hung KK, Wong CS, Mok CPY, et al. The evolution in anxiety and depression with the progression of the pandemic in adult populations from eight countries and four continents. Int J Environ Res Public Health 2021 May 01;18(9):4845 [FREE Full text] [doi: 10.3390/ijerph18094845] [Medline: 34062769]
9. Marroquín B, Vine V, Morgan R. Mental health during the COVID-19 pandemic: Effects of stay-at-home policies, social distancing behavior, and social resources. Psychiatry Res 2020 Nov;293:113419 [FREE Full text] [doi: 10.1016/j.psychres.2020.113419] [Medline: 32861098]
10. Tham WW, Sojli E, Bryant R, McAleer M. Common mental disorders and economic uncertainty: evidence from the COVID-19 pandemic in the U.S. PLoS One 2021 Dec 2;16(12):e0260726 [FREE Full text] [doi: 10.1371/journal.pone.0260726] [Medline: 34855850]
11. Şimşir Z, Koç H, Seki T, Griffiths MD. The relationship between fear of COVID-19 and mental health problems: a meta-analysis. Death Stud 2022;46(3):515-523. [doi: 10.1080/07481187.2021.1889097] [Medline: 33641626]
12. Berete F, Braekman E, Charafeddine R, Demarest S, Drieskens S, Gisle L, et al. Vijfde COVID-19 gezondheidsenquête: eerste resultaten. Article in Dutch. Sciensano. 2020. URL: https://www.sciensano.be/sites/default/files/report5_covid-19his_nl_1.pdf [accessed 2023-03-02]
13. UAntwerpen, UHasselt, KU Leuven. De Grote Corona-studie. 2020. URL: https://corona-studie.shinyapps.io/corona-studie/ [accessed 2023-03-02]
14. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892 [FREE Full text] [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]
15. O'Connor RC, Wetherall K, Cleare S, McClelland H, Melson AJ, Niedzwiedz CL, et al. Br J Psychiatry 2021 Jun;218(6):326-333 [FREE Full text] [doi: 10.1192/bjp.2020.212] [Medline: 33081860]
16. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, COVID-19 Suicide Prevention Research Collaboration. Suicide risk and prevention during the COVID-19 pandemic. Lancet Psychiatry 2020 Jun;7(6):468-471 [FREE Full text] [doi: 10.1016/S2215-0366(20)30171-1] [Medline: 32330430]

17. Lee Y, Lui LMW, Chen-Li D, Liao Y, Mansur RB, Brietzke E, et al. Government response moderates the mental health impact of COVID-19: a systematic review and meta-analysis of depression outcomes across countries. J Affect Disord 2021 Jul 01;290:364-377 [FREE Full text] [doi: 10.1016/j.jad.2021.04.050] [Medline: 34052584]
18. Stephenson E, Butt DA, Gronsbell J, Ji C, O'Neill B, Crampton N, et al. Changes in the top 25 reasons for primary care visits during the COVID-19 pandemic in a high-COVID region of Canada. PLoS One 2021 Aug 12;16(8):e0255992 [FREE Full text] [doi: 10.1371/journal.pone.0255992] [Medline: 34383844]
19. Carr MJ, Steeg S, Webb RT, Kapur N, Chew-Graham CA, Abel KM, et al. Effects of the COVID-19 pandemic on primary care-recorded mental illness and self-harm episodes in the UK: a population-based cohort study. Lancet Public Health 2021 Feb;6(2):e124-e135 [FREE Full text] [doi: 10.1016/S2468-2667(20)30288-7] [Medline: 33444560]
20. Mansfield KE, Mathur R, Tazare J, Henderson AD, Mulick AR, Carreira H, et al. Indirect acute effects of the COVID-19 pandemic on physical and mental health in the UK: a population-based study. Lancet Digit Health 2021 Apr;3(4):e217-e230 [FREE Full text] [doi: 10.1016/S2589-7500(21)00017-0] [Medline: 33612430]
21. Williams R, Jenkins DA, Ashcroft DM, Brown B, Campbell S, Carr MJ, et al. Diagnosis of physical and mental health conditions in primary care during the COVID-19 pandemic: a retrospective cohort study. Lancet Public Health 2020 Oct;5(10):e543-e550 [FREE Full text] [doi: 10.1016/S2468-2667(20)30201-2] [Medline: 32979305]
22. Rachamin Y, Senn O, Streit S, Dubois J, Deml MJ, Jungo KT. Impact of the COVID-19 pandemic on the intensity of health services use in general practice: a retrospective cohort study. Int J Public Health 2021 May 7;66:635508 [FREE Full text] [doi: 10.3389/ijph.2021.635508] [Medline: 34744588]
23. Murphy L, Markey K, O' Donnell C, Moloney M, Doody O. The impact of the COVID-19 pandemic and its related restrictions on people with pre-existent mental health conditions: a scoping review. Arch Psychiatr Nurs 2021 Aug;35(4):375-394 [FREE Full text] [doi: 10.1016/j.apnu.2021.05.002] [Medline: 34176579]
24. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. Sci Rep 2021 May 13;11(1):10173 [FREE Full text] [doi: 10.1038/s41598-021-89700-8] [Medline: 33986414]
25. Kahil K, Cheaito MA, El Hayek R, Nofal M, El Halabi S, Kudva KG, et al. Suicide during COVID-19 and other major international respiratory outbreaks: a systematic review. Asian J Psychiatr 2021 Feb;56:102509 [FREE Full text] [doi: 10.1016/j.ajp.2020.102509] [Medline: 33418284]
26. Zürcher SJ, Kerksieck P, Adamus C, Burr CM, Lehmann AI, Huber FK, et al. Prevalence of mental health problems during virus epidemics in the general public, health care workers and survivors: a rapid review of the evidence. Front Public Health 2020 Nov 11;8:560389 [FREE Full text] [doi: 10.3389/fpubh.2020.560389] [Medline: 33262966]
27. Lin Y, Chen C, Wu S. Efficiency and quality of data collection among public mental health surveys conducted during the COVID-19 pandemic: systematic review. J Med Internet Res 2021 Feb 10;23(2):e25118 [FREE Full text] [doi: 10.2196/25118] [Medline: 33481754]
28. Truyers C, Goderis G, Dewitte H, Akker MV, Buntinx F. The Intego database: background, methods and basic results of a Flemish general practice-based continuous morbidity registration project. BMC Med Inform Decis Mak 2014 Jun 06;14:48 [FREE Full text] [doi: 10.1186/1472-6947-14-48] [Medline: 24906941]
29. Van den Bulck S, Crèvecoeur J, Aertgeerts B, Delvaux N, Neyens T, Van Pottelbergh G, et al. The impact of the COVID-19 pandemic on the incidence of diseases and the provision of primary care: A registry-based study. PLoS One 2022 Jul 6;17(7):e0271049 [FREE Full text] [doi: 10.1371/journal.pone.0271049] [Medline: 35793324]
30. COVID-19. Sciensano. 2022. URL: https://epistat.wiv-isp.be/covid/ [accessed 2023-03-02]
31. Nielen MMJ, Spronk I, Davids R, Korevaar JC, Poos R, Hoeymans N, et al. Estimating morbidity rates based on routine electronic health records in primary care: observational study. JMIR Med Inform 2019 Jul 26;7(3):e11929 [FREE Full text] [doi: 10.2196/11929] [Medline: 31350839]
32. De "Belgian Mental Health Monitor COVID-19", een methode om de overheidsrepons op de impact van COVID-19 op de mentale gezondheid van de Belgische bevolking op te volgen. Hoge Gezondheidsraad. 2021 Jul. URL: https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/20210713_hgr-9640_belgian_mental_health_monitor_vweb.pdf [accessed 2023-03-02]
33. Van den Bulck S, De Burghgraeve T, Raat W, Mamouris P, Coursier P, Vankrunkelsven P, et al. The effect of automated audit and feedback on data completeness in the electronic health record of the general physician: protocol for a cluster randomized controlled trial. Trials 2021 May 04;22(1):325 [FREE Full text] [doi: 10.1186/s13063-021-05259-9] [Medline: 33947448]
34. Carrillo de Albornoz S, Sia K, Harris A. The effectiveness of teleconsultations in primary care: systematic review. Fam Pract 2022 Jan 19;39(1):168-182. [doi: 10.1093/fampra/cmab077] [Medline: 34278421]
35. Van Hoof E, De Laet H, Résibois M, Gérard S, Dekeyser S, Loix E, et al. Living document I: Belgian mental health (care) data repository. Superior Health Council. 2022 Jan 11. URL: https://zenodo.org/record/5838409 [accessed 2023-03-02]
36. Abrams EM, Greenhawt M, Shaker M, Pinto AD, Sinha I, Singer A. The COVID-19 pandemic: adverse effects on the social determinants of health in children and families. Ann Allergy Asthma Immunol 2022 Jan;128(1):19-25 [FREE Full text] [doi: 10.1016/j.anai.2021.10.022] [Medline: 34699969]

37. Courtin E, Vineis P. COVID-19 as a Syndemic. Front Public Health 2021 Sep 9;9:763830 [FREE Full text] [doi: 10.3389/fpubh.2021.763830] [Medline: 34568273]
38. Mendenhall E. The COVID-19 syndemic is not global: context matters. Lancet 2020 Nov 28;396(10264):1731 [FREE Full text] [doi: 10.1016/S0140-6736(20)32218-2] [Medline: 33256917]

Abbreviations

ICPC-2: International Classification of Primary Care, second edition
P-code: psychological International Classification of Primary Care, second edition, code
SES: socioeconomic status

Edited by A Mavragani, T Sanchez; submitted 28.09.22; peer-reviewed by N Nazari, S Tian, T Magnée; comments to author 26.10.22; revised version received 17.11.22; accepted 16.12.22; published 10.03.23.

Please cite as:
Vandamme J, Beerten SG, Crèvecoeur J, Van den Bulck S, Aertgeerts B, Delvaux N, Van Pottelbergh G, Vermandere M, Tops L, Neyens T, Vaes B
The Impact of the COVID-19 Pandemic on the Registration and Care Provision of Mental Health Problems in General Practice: Registry-Based Study
JMIR Public Health Surveill 2023;9:e43049
URL: https://publichealth.jmir.org/2023/1/e43049
doi:10.2196/43049
PMID:36599160

Original Paper

The Bidirectional Association Between Cognitive Function and Gait Speed in Chinese Older Adults: Longitudinal Observational Study

Haibin Li[1,2*], PhD; Jiajia Zhang[3*], BSc; Xinye Zou[4], PhD; Xiuqin Jia[3], PhD; Deqiang Zheng[5], PhD; Xiuhua Guo[5], PhD; Wuxiang Xie[6], PhD; Qi Yang[3,7,8], MD, PhD

[1]Department of Cardiac Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[2]Heart Center and Beijing Key Laboratory of Hypertension, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[3]Department of Radiology, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[4]Department of Education, University of Cambridge, Cambridge, United Kingdom
[5]Department of Epidemiology and Health Statistics, School of Public Health, Capital Medical University, Beijing, China
[6]Peking University Clinical Research Institute, Peking University First Hospital, Beijing, China
[7]Beijing Laboratory for Cardiovascular Precision Medicine, Beijing, China
[8]Key Laboratory of Medical Engineering for Cardiovascular Disease, Ministry of Education, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Qi Yang, MD, PhD
Department of Radiology
Beijing Chaoyang Hospital
Capital Medical University
Number 8, Gongti South Road
Chaoyang District
Beijing, 100020
China
Phone: 86 01085231928
Email: yangyangqiqi@gmail.com

Abstract

Background: Cognitive and gait speed decline are common conditions in older adults and are often associated with future adverse consequences. Although an association between cognitive function and gait speed has been demonstrated, its temporal sequence remains unclear, especially in older Chinese adults. Clarifying this could help identify interventions to improve public health in older adults.

Objective: This study aims to examine the longitudinal reciprocal association between gait speed and cognitive function and the possible temporal sequence of changes in both factors in a national longitudinal cohort.

Methods: Data were derived from 2 waves (2011 baseline and 2015 follow-up) of the China Health and Retirement Longitudinal Study (CHARLS). Participants 60 years or older, without dementia or Parkinson disease at baseline, and with completed data on gait speed and cognition at both baseline and follow-up were included. Usual gait speed was measured over two 2.5-m walks. Mental intactness and episodic memory were used to assess global cognitive function. Cross-lagged panel models and linear mixed-effects models were used to examine the association between cognition and gait speed over time. Standardized coefficients were reported.

Results: A total of 3009 participants (mean age 66.4 years, SD 5.4 years; 1422/3009, 47.26%, female participants) were eligible for inclusion in our analyses. Cross-lagged panel analyses revealed that after accounting for baseline gait speed, cognition, and potential confounders, baseline global cognition (β=.117, 95% CI 0.082-0.152; *P*<.001), mental intactness (β=.082, 95% CI 0.047-0.118; *P*<.001), and episodic memory (β=.102, 95% CI 0.067-0.137; *P*<.001) were associated with subsequent gait speed. Simultaneously, baseline gait speed was also associated with subsequent global cognition (β=.056, 95% CI 0.024-0.087; *P*=.001), mental intactness (β=.039, 95% CI 0.008-0.069; *P*=.01), and episodic memory (β=.057, 95% CI 0.023-0.092; *P*=.001). The comparison of standardized cross-lagged coefficients suggested that the effect size of baseline global cognition on subsequent gait speed was significantly larger than the reverse effect (χ_1^2=6.50, *P* for difference=.01). However, the effects of both mental intactness and episodic memory on subsequent gait speed were not significantly stronger than those of the reverse pathway (χ_1^2=3.33, *P* for difference=.07 and χ_1^2=3.21, *P* for difference=.07). Linear mixed-effects analyses further supported these bidirectional relationships, revealing that lower baseline cognitive scores predicted steeper declines in gait speed trajectory, and slower baseline gait speed predicted more declines in cognitive trajectory over time.

Conclusions: There is a longitudinal bidirectional association between usual gait speed and both global cognitive function and specific domains of mental intactness and episodic memory among Chinese older adults. Baseline global cognition is likely to have a stronger association with subsequent gait speed than the reverse pathway. This interlinkage is noteworthy and may have implications for public health. Maintaining normal cognitive function may be an important interventional strategy for mitigating age-related gait speed reduction.

(JMIR Public Health Surveill 2023;9:e44274) doi:10.2196/44274

KEYWORDS

aging; cognitive function; gait speed; cross-lagged panel models; longitudinal study

Introduction

Population aging is a global challenge, with China being one of Asia's most rapidly aging countries. The seventh population census in China reported that adults aged 60 years or older account for 18.70% of the total population, meaning that the prevalence of age-related diseases in China is set to increase dramatically [1-3]. Cognitive function is a vital dimension of healthy aging [4,5]. Cognitive function declines with age, which subsequently leads to many age-related diseases. Gait speed, an indicator that is easier to measure than other parameters, has been used as a simple, reliable, and sensitive indicator to assess and reflect the functional capacity and overall health [6,7]. Diminished gait speed is considered a strong predictor of the onset of negative outcomes in older adults [7]. Interestingly, declines in cognitive and gait speed often occur concurrently in older adults and are often related [8,9]. Moreover, these 2 common geriatric symptoms are associated with future adverse events, including falls, sarcopenia, disability, dementia, and even death, leading to a heavy burden on public health [10-13]. Therefore, cognitive function and gait speed have drawn widespread attention from public health and clinical researchers as health-related factors leading to age-related diseases.

Previous cross-sectional studies have reported an association between cognitive function and gait speed in older adults [14-19]. Several longitudinal studies have further investigated unidirectional temporal relationships, but the results have been inconsistent. For instance, longitudinal data from the Tasmanian Study of Cognition and Gait [18] revealed that cognitive decline predicted decreases in gait speed. A multicenter randomized controlled trial from European countries also showed that poorer baseline cognition was linked to an incident slow gait [20]. Evidence from older American adults further supported this unidirectional relationship [21,22]. However, a recent prospective study in Sweden suggested that slow gait speed predicts poor cognition [23]. Similar results were also found in older Australian and Japanese adults [24-26]. Furthermore, the currently available studies examining the longitudinal bidirectional association between cognition and gait speed have been conducted mainly in Western countries. Among these, 2 studies from the United States and 1 from the United Kingdom reported that cognition and gait speed were associated with each other over time [9,27,28]. While 2 longitudinal American studies also reported a reciprocal relationship, one [29] found that slower gait speed may be the driving factor in older men and women, whereas the other [30] found that cognitive decline was more likely to predict slower gait speed in older women. However, findings from a sample of 1478 older adults in the Mayo Clinic Study of Aging suggested that slow gait speed is likely a precursor to later cognitive decline, but not vice versa [31].

In summary, it appears that there is a relationship between cognitive function and gait speed, but the temporal sequence remains unclear. Moreover, due to methodological limitations (eg, linear mixed-effects models) [28,30,31], most previous studies could only unidirectionally examine the effect of gait speed on cognition or the effect of cognition on gait speed, but failed to simultaneously examine bidirectional associations between the 2 factors or examine temporal precedence [29,32]. Finally, relevant studies were primarily conducted among Western people [9,27-31], which differed greatly from China in terms of socioeconomic development, culture, and nutritional intake. To the best of our knowledge, the longitudinal reciprocal association between cognition and gait speed and its temporal sequence have not been examined in China.

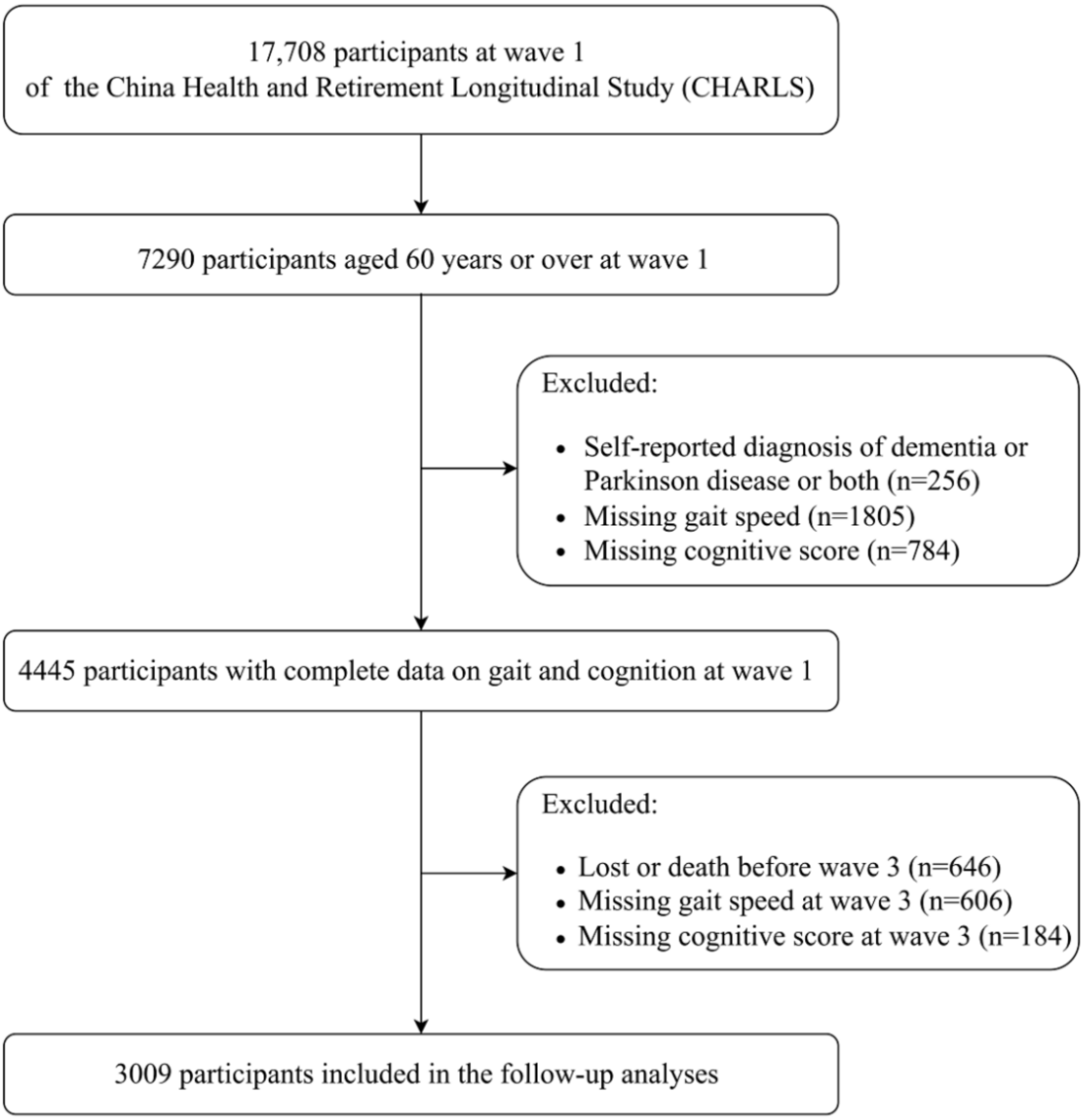
Therefore, to fill these research gaps, this study aims to investigate the longitudinal bidirectional relationship between cognition and gait speed using a cross-lagged panel design, and determine the possible temporal sequence of the changes in the 2 factors in older Chinese community-dwelling adults. The clarity of the temporal sequence of cognitive decline and decreased gait speed might help identify the underlying causal predominance and early interventions for preventing age-related diseases.

Methods

Participants

Data were obtained from the China Health and Retirement Longitudinal Study (CHARLS), a nationwide population-based prospective cohort survey of adults aged 45 years or older. Further details on the study design and sampling strategies of the CHARLS have been described previously [33]. Briefly, in the baseline survey, a multistage stratified probability-proportionate-to-size sampling was adopted, and 17,708 adults from 450 villages or urban communities in 150 counties across 28 provinces in China were recruited. The baseline survey was implemented during the 2011-2012 period (wave 1), with subsequent waves completed approximately every 2 years, and data from wave 3 were collected from 2015 to 2016. This study included 7290 participants aged 60 years or older at baseline. Among these, participants were excluded due to a self-reported diagnosis of dementia or Parkinson disease or both at wave 1 (n=256), loss or death before wave 3 (n=646), or not having finished the gait speed test (n=1805 at wave 1 and 606 at wave 3) or cognitive tests (n=784 at wave 1 and 184 at wave 3). Ultimately, 3009 individuals were included in this study (Figure 1).

Figure 1. Flowchart of participants' selection.



Assessment of Cognitive Function

Two well-established composite measures were used to assess global cognitive function: mental intactness and episodic memory, which have been validated in Chinese people [34]. These metrics are similar to those used in the American Health and Retirement Study and in previous studies [35,36]. Mental intactness was based on components from the Telephone Interview of Cognitive Status, which described time orientation (naming the month, day, year, week, and season), numerical ability (serial subtraction of 7 from 100, 5 times), and visual and spatial abilities (redrawing task of 2 overlapping pentagons) [36,37]. Mental intactness scores were equal to the sum of the correct answers, with numbers ranging from 0 to 11. Episodic memory was based on immediate and delayed word recall tests. After the interviewer randomly read a list of 10 Chinese nouns, the respondents were asked to memorize as many words as they could immediately (immediate recall) and several minutes later (delayed recall). Episodic memory scores were the sum of correctly repeated words, ranging from 0 to 20. Global cognition scores were the aggregate of mental intactness and episodic memory scores ranging from 0 to 31, with higher scores indicating better cognitive performance.

Assessment of Gait Speed

Gait speed was computed over 2 straight 2.5-m path walking courses. Those who had difficulty stopping walking (eg, recent surgery, injury, or health problems) were excluded from this test. Eligible older participants were instructed to walk 2.5 m along a noncarpeted walking course (there and back) at their usual speed. Timing began when one of the participants' feet crossed the start line and landed completely, and ended when the participant crossed the finish line completely. The average velocity (m/s; accurate to 2 decimal places) was calculated, with higher scores indicating faster speeds.

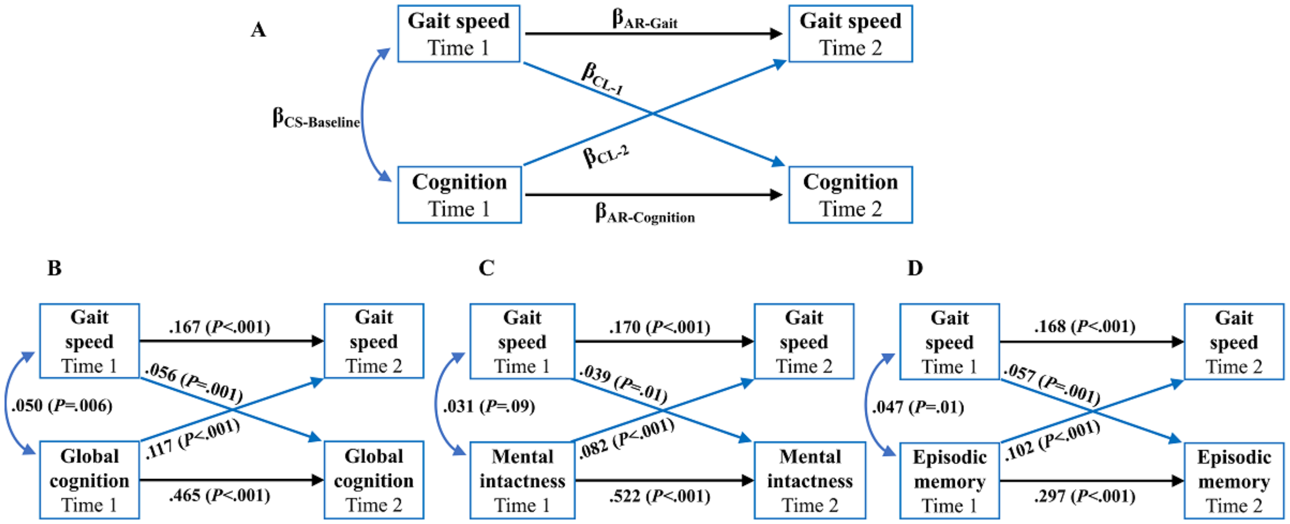
Covariates

The following baseline measurements were included as potential covariates: age, gender, residence region, marital status, educational level, self-reported health, current smoking and drinking status, self-reported visual and hearing impairments, depressive symptoms, restrictions on activities of daily living (ADL), and BMI. Information on chronic diseases included hypertension, diabetes mellitus, dyslipidemia, heart disease, stroke, cancer, lung disease, arthritis, kidney disease, digestive disease, and asthma. The number of diseases was used as a covariate. Restriction was considered as 1 or more dependency in ADL, including dressing, bathing, eating, getting into/out of bed, and using the toilet [38]. Depressive symptoms were measured using 10 items from the Center for Epidemiologic Studies Depression Scale, and a score of 10 or more was defined as positive [39].

Statistical Analysis

Continuous variables are presented as means and SDs and categorical variables are shown as numbers (proportions). Cross-lagged path models were first performed to examine the possible longitudinal reciprocal association between cognition and gait speed measured at 2 time points 4 years apart [32,40]. Figure 2A demonstrates this study's general modeling strategy, including 5 paths and their corresponding coefficients. Cross-lagged paths concurrently include 2 cross-lagged coefficients, $\beta_{CL\text{-}1}$ and $\beta_{CL\text{-}2}$. $\beta_{CL\text{-}1}$ represents the effect of gait speed at time 1 on cognitive scores at time 2, whereas $\beta_{CL\text{-}2}$ implies the regression of cognitive scores at time 1 on gait speed at time 2. The temporal sequence was determined by comparing the estimated standardized cross-lagged coefficients [40]. Cross-sectional paths between gait speed and cognitive scores were also modeled, and the coefficient $\beta_{CS\text{-}Baseline}$ at time 1 represented the baseline correlation between gait speed and cognitive scores. Finally, 2 autoregressive coefficients, $\beta_{AR\text{-}Gait}$ and $\beta_{AR\text{-}Cognition}$, were obtained from the paths between times 1 and 2 for gait speed and cognitive scores, respectively, accounting for with-person stability in each measure from the baseline to the subsequent time.

Figure 2. Cross-lagged panel models applied to assess gait speed and cognition among Chinese older adults. (A) The general modeling strategy used for the cross-lagged panel models. (B) The cross-lagged panel model where gait speed was associated with global cognition. (C) The cross-lagged panel model where gait speed was associated with mental intactness. (D) The cross-lagged panel model where gait speed was associated with episodic memory. Numeric values are standardized structural regression coefficients. AR: autoregressive; CL: cross-lagged; CS: cross sectional.



Before the cross-lagged path analysis, the baseline and subsequent values of gait speed and cognitive scores were adjusted for a series of covariates by regression residual analyses, and then standardized with Z-transformation (mean 0, SD 1). Standardized coefficients were reported from the 4 models. A crude model and 3 multivariable-adjusted models were built: model 1 was built without any adjustments; model 2 was adjusted for age and sex; age, sex, residence, marital status, educational level, and smoking and drinking status were further controlled in model 3; and finally, model 4 was comprehensively adjusted, including the covariates in model 3 plus self-reporting of health scores, number of diseases, depressive symptoms, visual and hearing impairments, restriction on ADL, and BMI. A comparative fit index (CFI) >0.90 and standardized root-mean-square residual (SRMR) <0.05 were used to suggest the good fitness of the models [41]. Linear mixed-effects models were further implemented to visualize the longitudinal associations between cognition and gait speed with aging.

Several sensitivity analyses were performed. First, we controlled for the specific chronic diseases (11 diseases in total) instead of the number of diseases in model 4. Second, in addition to the covariates in model 4, we further adjusted for living environmental factors to control for their potential confounding, including the type of accommodation (multistory or 1 story), water sources (running water or no running water), room temperature (comfortable or uncomfortable), and household air pollution (clean fuels for both heating and cooking, clean fuel for heating or cooking, solid fuels for both heating and cooking) [42]. Finally, we excluded individuals with extremely low gait speed or global cognitive function scores or both at baseline (≤mean – 2 SD) to reduce the potential influence of the reverse temporal sequence. All analyses in this study were conducted in Stata 16.0 (StataCorp) and $P \leq .05$ (2-tailed) was considered statistically significant.

Ethical Considerations

The Biomedical Ethics Review Committee of Peking University approved the ethical application for the collection of human participant data for the CHARLS (IRB00001052-11015). The study data were anonymous. All the respondents provided written informed consent before the survey.

Results

Ultimately, a national sample of 3009 adults aged 60 years or older was included, around half of whom were female (1422/3009, 47.26%). The mean age of the participants at baseline was 66.4 (SD 5.4) years. The baseline characteristics of the participants are summarized in Table 1.

Table 2 presents the results of the cross-lagged panel models. In models 1, 2, and 3, the standardized structural regression coefficients for all paths were strongly significant (Table 2). Fully adjusting for potential confounders diminished the effect size of all associations, but mostly remained statistically significant (Table 2); therefore, we mainly presented results from model 4. Moreover, autoregressive paths in all models for gait speed and cognition were highly remarkable, with coefficients not close to 0, capturing relative stability and influence from the baseline to the subsequent time point [40].

Figure 2B illustrates the cross-lagged path analysis of the gait speed and global cognitive scores in model 4. Cross sectionally, gait speed was positively associated with global cognitive scores (β=.05, 95% CI 0.014-0.086; P=.006). In addition, there was a significant cross-lagged pathway, and the standardized path coefficient from global cognitive scores at time 1 to gait speed at time 2 (β=.117, 95% CI 0.082-0.152; P<.001) was significantly stronger than that from gait speed at time 1 to global cognitive scores at time 2 (β=.056, 95% CI 0.024-0.087; P=.001; χ_1^2=6.50, P for difference=.01). Fitting indicators (CFI=0.952 and SRMR=0.034) suggested a good model fit based on the criteria of CFI>0.90 and SRMR<0.05 (Table 2).

Figure 2C shows the cross-lagged model of gait speed and mental intactness scores in model 4. At baseline, the correlation was not significant (β=.031, 95% CI 0.005-0.067; P=.09). However, there was a positive association between mental intactness scores at time 1 and gait speed at time 2 (β=.082, 95% CI 0.047-0.118; P<.001), and gait speed at time 1 was also associated with mental intactness scores at time 2 (β=.039, 95% CI 0.008-0.069; P=.01). The standardized path coefficient from mental intactness to gait speed was larger than that of the reverse pathway, but the difference was not significant (χ_1^2=3.33, P for difference=.07). The overall fit of this model was good (CFI=0.969 and SRMR=0.029; Table 2).

Figure 2D presents the results of gait speed and episodic memory scores in model 4. Gait speed was associated with episodic memory scores at baseline (β=.047, 95% CI 0.012-0.083; P=.01). For cross-lagged effects, episodic memory scores at time 1 were positively and significantly associated with gait speed at time 2 (β=.102, 95% CI 0.067-0.137; P<.001) and vice versa (β=.057, 95% CI 0.023-0.092; P=.001); however, the difference between the 2 standardized cross-lagged path coefficients was not significant (χ_1^2=3.21, P for difference=.07). The fitting parameters were CFI=0.928 and SRMR=0.031, indicating that the final model fits the data well (Table 2).

Multimedia Appendix 1 shows the baseline characteristics of living environmental factors of the study population. The first 2 sensitivity analyses showed similar bidirectional associations with model 4 (Multimedia Appendix 2). Specifically, when controlling for specific chronic diseases at baseline instead of the number of diseases or after additionally adjusting for living environmental factors, all cross-lagged path coefficients were still strongly significant (P<.05 in all cases; see Multimedia Appendix 2 for precise P values). In addition, after excluding 97 participants with extremely low gait speed or global cognitive scores or both, the association of gait speed with global cognition and episodic memory did not differ from model 4 (Multimedia Appendix 2). While the effect of baseline gait speed on subsequent mental intactness was slightly attenuated and became marginally significant (β=.030, 95% CI 0.002-0.061; P=.06), the reverse pathway still strongly held (β=.084, 95% CI 0.049-0.120; P<.001; Multimedia Appendix 2).

Trajectories from Figure 3A further supported the temporal associations from cross-lagged path models, with slower baseline gait speed (1 SD below the mean baseline speed) predicting faster decline rates in global cognitive, mental intactness, and episodic memory scores over time. In Figure 3B, the trajectories were also consistent with the current results; older adults with lower (1 SD below the mean baseline score) baseline cognitive scores appeared to have steeper declines in gait speed than those with higher scores across the follow-up period.

Table 1. Baseline characteristics of the study population (N=3009).

Characteristic	Values
Age (years), mean (SD)	66.4 (5.4)
Female, n (%)	1422 (47.26)
Rural residence, n (%)	2020 (67.13)
Married, n (%)	2412 (80.16)
Educational level, n (%)	
No formal education	1609 (53.47)
Primary school	853 (28.35)
Middle school	390 (12.96)
High school or above	157 (5.22)
Current smoker, n (%)	977 (32.47)
Current drinker, n (%)	969 (32.20)
Self-report of health score, mean (SD)	3.1 (0.9)
History of diseases, n (%)	
Hypertension	968 (32.17)
Diabetes	202 (6.71)
Dyslipidemia	303 (10.07)
Heart diseases	444 (14.76)
Stroke	69 (2.29)
Cancer	23 (0.76)
Lung disease	394 (13.09)
Arthritis	1082 (35.96)
Kidney disease	178 (5.92)
Digestive disease	682 (22.67)
Asthma	200 (6.65)
Number of diseases, mean (SD)	1.6 (1.5)
Comorbidity, n (%)	
0	809 (26.89)
1	875 (29.08)
≥2	1325 (44.03)
Visual impairment, n (%)	252 (8.37)
Hearing impairment, n (%)	351 (11.67)
Depressive symptoms, n (%)	1165 (38.72)
Restriction on ADL[a], n (%)	529 (17.58)
BMI (kg/m^2), mean (SD)	23.0 (3.7)
Weight status, n (%)	
Underweight	240 (7.98)
Normal	1353 (44.97)
Overweight	600 (19.94)
Obesity	789 (26.22)
Gait speed (m/s), mean (SD)	0.6 (0.2)
Global cognitive scores, mean (SD)	13.9 (5.1)

Characteristic	Values
Mental intactness scores, mean (SD)	7.4 (3.0)
Episodic memory scores, mean (SD)	6.5 (3.1)

[a] ADL: activities of daily living.

Table 2. Cross-lagged panel model results for gait speed and cognition scores in a population-based study of older adults.[a]

Gait and cognition results	Gait→cognition[b]	Cognition→gait[b]	Cross sectional[b]	Autoregressive[b]		Fit indices			
	β_{CL-1}[c]	β_{CL-2}[d]	$\beta_{CS-Baseline}$[e]	$\beta_{AR-Gait}$[f]	$\beta_{AR-Cognition}$[g]	CFI[h]	TLI[i]	RMSEA[j]	SRMR[k]
Global cognitive scores									
Model 1[l]	.098[m]	.244[m]	.175[m]	.227[m]	.630[m]	0.944	0.719	0.204	0.045
Model 2[n]	.066[m]	.175[m]	.112[m]	.187[m]	.594[m]	0.963	0.817	0.142	0.034
Model 3[o]	.060[m]	.124[m]	.077[m]	.183[m]	.479[m]	0.956	0.782	0.122	0.033
Model 4[p]	.056[q]	.117[m]	.050[q]	.167[m]	.465[m]	0.952	0.758	0.122	0.034
Mental intactness scores									
Model 1[l]	.076[m]	.216[m]	.165[m]	.235[m]	.630[m]	0.962	0.812	0.178	0.037
Model 2[n]	.049[m]	.147[m]	.094[m]	.193[m]	.644[m]	0.975	0.874	0.126	0.029
Model 3[o]	.042[q]	.091[m]	.055[q]	.188[m]	.535[m]	0.972	0.862	0.105	0.028
Model 4[p]	.039[r]	.082[m]	.031	.170[m]	.522[m]	0.969	0.846	0.107	0.029
Episodic memory scores									
Model 1[l]	.118[m]	.186[m]	.127[m]	.246[m]	.404[m]	0.894	0.470	0.195	0.052
Model 2[n]	.081[m]	.143[m]	.091[m]	.194[m]	.376[m]	0.927	0.634	0.135	0.038
Model 3[o]	.065[m]	.107[m]	.067[m]	.185[m]	.305[m]	0.927	0.634	0.135	0.038
Model 4[p]	.057[q]	.102[m]	.047[r]	.168[m]	.297[m]	0.928	0.638	0.104	0.031

[a] See Figure 2 for a cross-lagged panel model diagram as a reference.

[b] Standardized structural regression coefficients.

[c] β_{CL-1} is cross-lagged path 1, where gait speed at time 1 predicts cognition scores at time 2.

[d] β_{CL-2} is cross-lagged path 2, where cognition scores at time 1 predict gait speed at time 2.

[e] $\beta_{CS-Baseline}$ is the cross-sectional association between gait speed and cognition scores within time 1.

[f] $\beta_{AR-Gait}$ is the autoregressive coefficient for gait speed.

[g] $\beta_{AR-Cognition}$ is the autoregressive coefficient for cognition scores.

[h] CFI: comparative fit index.

[i] TLI: Tucker-Lewis Index.

[j] RMSEA: root-mean-square error of approximation.

[k] SRMR: standardized root-mean-square residual.

[l] Model 1 was a crude model.

[m] $P<.001$.

[n] Model 2 was adjusted for age and sex.

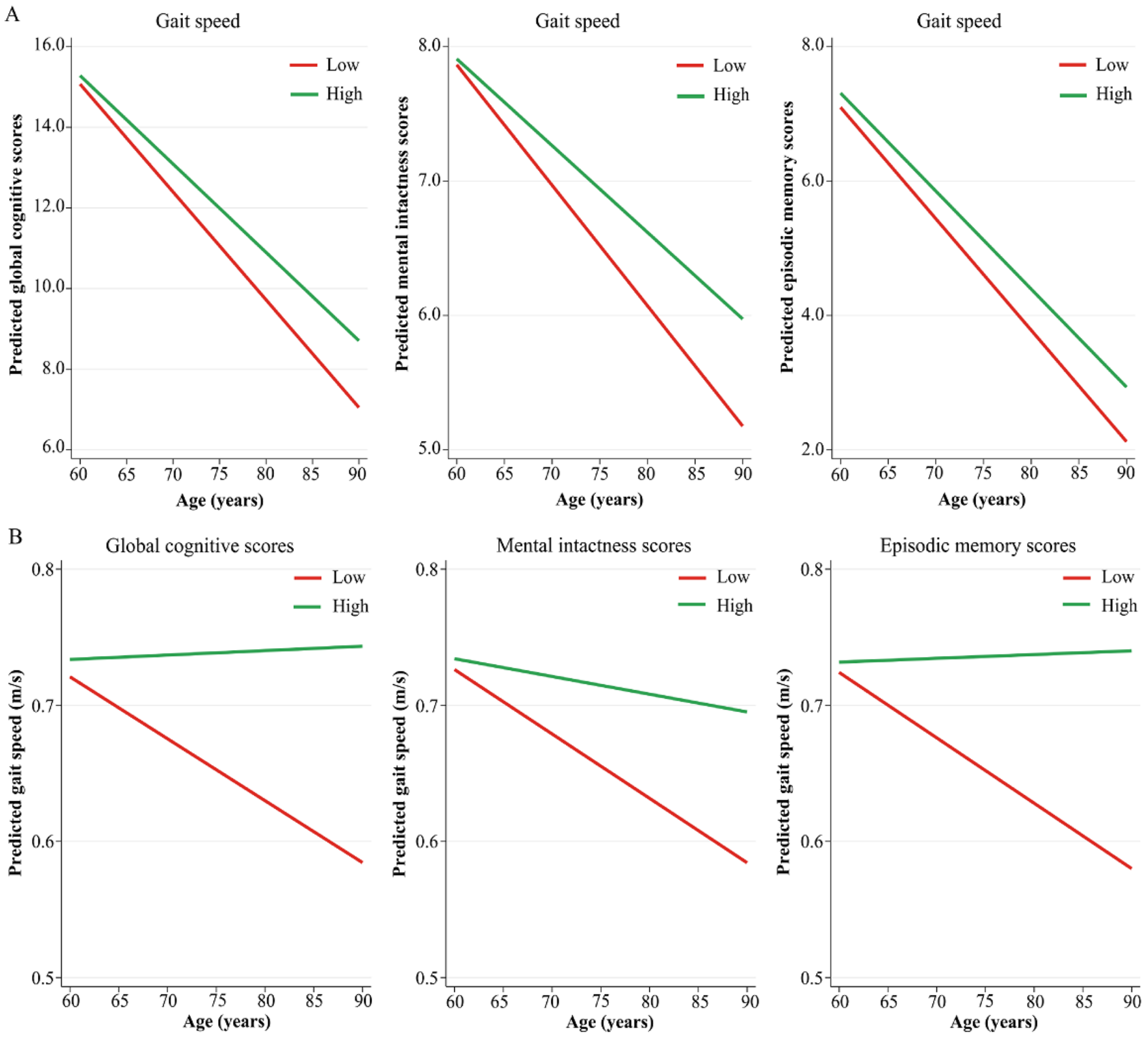
[o] Model 3 was adjusted for age, sex, residence, marital status, educational level, smoking, and drinking status.

[p] Model 4 was adjusted for covariates in model 3 plus self-reported health score, number of diseases, depressive symptoms, visual and hearing impairment, restriction on activities of daily living, and BMI.

[q] $P<.01$.

[r] $P<.05$.

Figure 3. Trajectories of gait speed and cognition scores in a population-based study of Chinese older adults. (A) The predicted model estimates derived from linear mixed-effects models for global cognition, mental intactness, and episodic memory scores. The y-axis represents the predicted cognition scores based on model estimates, and separate lines indicate 1 SD above and below the mean gait speed ("high" and "low," respectively). (B) Predicted model estimates derived from linear mixed-effects models of gait speed. The y-axis represents the predicted gait speed based on model estimates, and separate lines indicate 1 SD above and below the mean cognition scores ("high" and "low," respectively).



Discussion

Principal Findings

During the 4-year follow-up among Chinese older adults, there was a longitudinal bidirectional association between global cognitive function, mental intactness, episodic memory, and gait speed after accounting for baseline levels and possible covariates. In addition, the comparison of standardized cross-lagged coefficients suggested that the temporal association from baseline poor global cognition to subsequent slow gait speed may be stronger than vice versa. Trajectories derived from linear mixed-effects models added further evidence on the dynamic associations, with poorer baseline cognition and slower gait speed predicting faster decreases in gait speed and declines in cognition, respectively.

Similar to previous cross-sectional studies [8,16,17], we also found that gait speed was reciprocally and positively associated with cognitive function at baseline among older adults. However, longitudinal studies may provide stronger evidence in the temporal direction. Two previous longitudinal studies, both using linear mixed-effects models, have shown different results, with one [28] reporting a bidirectional association between cognition and gait speed among older individuals in the United Kingdom, and the other suggesting that faster initial gait speed was associated with less later cognitive decline among older American adults, but not vice versa [31]. This inconsistency may be attributable to differences in the mean age, sample size, and heterogeneity of walking speed measures. Nevertheless, owing to a unidirectional preset, this statistical method cannot simultaneously examine bidirectional associations of 2 time-varying factors or temporal precedence [29,32]. Therefore, we fitted cross-lagged path models based on structural equation models and visualized temporal association using linear mixed-effects models, as suggested by a previous study in this area [29].

Our findings provided new insights into the nature of age-related changes in gait speed and cognition. To the best of our knowledge, only a few studies have elucidated the longitudinal reciprocal association between cognition and gait speed in older adults. Two such studies, also based on structural equation models, have also found bidirectional relationships between poor cognition and slow gait speed among older American people [9,27]. Specifically, 1 study included 412 participants aged 60 or older, finding that executive function, but not memory, predicted subsequent usual gait speed, and vice versa [9]. However, another study conducted in women aged 70-79 years reported a mutual association between gait speed and global cognitive function and particular memory. In our study, there was a bidirectional relationship between global cognition and gait speed and between specific domains (mental intactness and episodic memory) and gait speed. The mental status in the CHARLS describes time orientation, numerical ability, and visual and spatial abilities, and is similar to executive function [43]. Compared with these 2 studies, the sample size in our study was relatively larger. Our study had a 4-year follow-up duration that was shorter than the 9- and 6-year follow-up periods. Different cognitive tests on specific domains and discrepancies in the gender of the participants may also lead to this inconsistency. Therefore, more similar studies in older Chinese individuals are needed to verify whether the results hold across a longer follow-up interval or in sex-specific older adults. Another notable study among older American adults reported that early gait speed decline may be a precursor to later global cognitive decline and that the reciprocal association was weaker [29]. Although this study also revealed reciprocal relationships between the 2 factors, its temporal sequence contradicts ours. In our analysis, the standardized cross-lagged coefficient of baseline global cognition on follow-up gait speed was nearly 2 times that of baseline gait speed on subsequent cognition, indicating that global cognitive decline might be the cause of decreased gait speed than its result. The reasons for this discrepancy may be the difference in participants' baseline average age and the length of the follow-up period, given that both cognition and gait speed decline over time in older adults [8]. Our study participants were over 60 years old, while the older adults were 70-79 years in their study. Moreover, their design had a longer follow-up time of 9 years than our 4-year interval. Thus, the current temporal sequence of cognitive decline to gait slowing requires further evidence, especially in older Chinese individuals.

Consistent with prior unidirectional longitudinal studies [18,21,22,44], our results suggest that poor baseline cognition predicted gait speed decline after 4 years, irrespective of the exclusion of those with very low baseline gait speed and global cognition. Moreover, poor specific domains of cognition, including mental intactness and episodic memory, are predictive of decreased gait speed, which supports the results from previous studies [28,45]. A small pilot randomized controlled trial showed that improvements in domains of cognition, including executive function, attention, visuospatial skills, and processing speed, were associated with increases in walking speed [45]. Memory has also been reported to be positively and longitudinally associated with gait speed in older adults, supporting our results [28].

Although we found a stronger directionality of early cognitive change in subsequent gait speed change than the opposite pathway, we do not discount the importance of early gait change. As with prior results [24,26], although weak, we again demonstrated the predictive role of baseline faster gait speed in later cognition. A recent study also revealed that decreased gait speed precedes global cognitive decline during a longer follow-up duration than ours [23]. Evidence from a prospective study in 70-89-year older adults has also shown results similar to ours [31], with baseline faster gait speed associated with better global cognition, executive function, and memory. In addition, both gait speed and cognitive function were important when assessing the risk of dementia in older adults [29,46].

Neuroimaging studies also support this reciprocal relationship. Age-related changes in gait and cognition have been reported to share common underlying structural alterations in the prefrontal and temporal lobes [47]. Furthermore, several mechanisms may explain the stronger effect of cognition on gait speed. First, adverse structural alterations in the brain, such as hippocampal atrophy, white matter, and an increase in white matter damage, are associated with decreasing gait speed in older adults [48]. Second, walking is a complex task that requires energy, balance, and coordination of multiple systems, including the nervous system, and thus disruption or deterioration of any system tends to decrease gait speed [6,49]. Finally, specific cognitive capacities may be vital for the initiation and maintenance of walking [22]. A functional magnetic resonance imaging study has shown that some areas with greater activation in the brain reflect increased cognitive monitoring of movements in older participants compared with that in younger people [50]. Overall, cognitive decline may precede or co-occur with gait speed decrease among older Chinese adults; therefore, early arrangements should be ensured to maintain both cognitive function and gait status to avoid further adverse outcomes. Meanwhile, cognitive decline might be a sensitive early marker of gait deceleration, indicating that cognition might be an essential factor to be considered along with gait speed to identify risks for gait speed decrease or mobility deficits among older adults.

Strengths and Limitations

This study has several strengths. Participants aged 60 years or older were from the CHARLS, which is a nationally representative, population-based prospective cohort in which cognitive functioning was assessed with well-validated measures. Relatively comprehensive information is available in the CHARLS, and related covariates were adjusted as much as possible in this study's final model. In addition, this study expands upon cross-sectional and traditional longitudinal analyses by performing cross-lagged panel models, which allowed for the capture of a temporal sequence by simultaneously evaluating the strengths of temporal bidirectional associations [32,40]. In addition, compared with similar studies [27,29], our study added linear mixed-effects models to visualize the temporal bidirectional relationships of the 2 factors. Finally, to our knowledge, this is the first study to examine the possible temporal sequence of changes in gait speed and cognition in older Chinese people, and our findings deserve further verification.

Several limitations also need to be considered. First, our analysis involved an approximately 4-year follow-up interval due to the limited availability of gait speed information at the later wave mark; therefore, it could not be determined whether this association continued over a longer period. Second, while the reciprocal associations between cognition and gait speed were statistically significant after accounting for a variety of potential confounding factors, there may still be some related but uncontrolled variables such as the apolipoprotein E4 (ApoE4) genotype. Slow walking speed in those with ApoE4 preceded lower memory scores [24], but unfortunately, these data are not available in the CHARLS. Based on a genomic study, the *e4* allele has a 6.9% frequency rate in the Chinese population; therefore, this genotype may have slightly affected our results [35]. Third, the older participants were allowed walking aids for this study, which may obscure the true association between gait and cognition to some extent. As one recent study has shown, the cross-sectional correlation between cognition and gait may be attenuated when considering the interaction of walking-aid usage [8]. Eventually, gait speed was measured at a usual pace, with this study solely focused on this walking condition; therefore, we were unable to determine whether the relationships we reported agree with the results obtained under other walking conditions, such as maximum walking speed, and this requires further research to identify.

Conclusions

This study demonstrates a longitudinal reciprocal association of gait speed with global cognitive function, mental intactness, and episodic memory among Chinese older adults. Cross-lagged path coefficients suggest that baseline global cognition is likely to have a stronger association with subsequent gait speed than the reverse pathway. These findings further reinforce the importance of promoting both cognition and gait conditions early, and provide ideas for improving the clarity of the causal direction of cognition and gait speed changes. Maintaining normal cognitive function may be an important interventional strategy for mitigating age-related gait speed decrease.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (grants 81961128030 and 62076169), Beijing Hospitals Authority Youth Program (grants QML20200304), and Golden Seed Program of Beijing Chaoyang Hospital (CYJZ202101). We thank the China Health and Retirement Longitudinal Study (CHARLS) research team and the field team for their contributions to data collection. This analysis uses data or information from the Harmonized CHARLS data set and Codebook, Version D, as of June 2021, developed by the Gateway to Global Aging Data. The development of the Harmonized CHARLS was funded by the National Institute on Aging (grants R01 AG030153, RC2 AG036619, R03 AG043052). For further information, please refer to [51].

Data Availability

The China Health and Retirement Longitudinal Study (CHARLS) data set is publicly available at [52]. The data analyzed in our study can be obtained from the corresponding author upon request.

Authors' Contributions

HL designed the study; QY, XJ, and HL obtained funding; HL and JZ conducted the data analysis; HL, JZ, and XZ drafted and critically revised the manuscript. HL, JZ, XZ, XJ, DZ, XG, and WX substantially interpreted the data. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics (living environmental factors) of the study population.
[DOCX File , 16 KB - publichealth_v9i1e44274_app1.docx]

Multimedia Appendix 2
Cross-lagged panel model results for gait speed and cognition scores in a population-based study of older adults (sensitivity analyses).
[DOCX File , 21 KB - publichealth_v9i1e44274_app2.docx]

References

1. Chen R, Xu P, Li F, Song P. Internal migration and regional differences of population aging: An empirical study of 287 cities in China. Biosci Trends 2018 May 13;12(2):132-141 [FREE Full text] [doi: 10.5582/bst.2017.01246] [Medline: 29607873]

2. Fang EF, Scheibye-Knudsen M, Jahn HJ, Li J, Ling L, Guo H, et al. A research agenda for aging in China in the 21st century. Ageing Res Rev 2015 Nov;24(Pt B):197-205 [FREE Full text] [doi: 10.1016/j.arr.2015.08.003] [Medline: 26304837]
3. Ren R, Qi J, Lin S, Liu X, Yin P, Wang Z, et al. The China Alzheimer Report 2022. Gen Psychiatr 2022 Mar 11;35(1):e100751 [FREE Full text] [doi: 10.1136/gpsych-2022-100751] [Medline: 35372787]
4. Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. Lancet 2011 Mar 19;377(9770):1019-1031 [FREE Full text] [doi: 10.1016/S0140-6736(10)61349-9] [Medline: 21371747]
5. Lu N, Wu B, Pei Y. Exploring the reciprocal relationship between cognitive function and edentulism among middle-aged and older adults in China. Age Ageing 2021 May 05;50(3):809-814 [FREE Full text] [doi: 10.1093/ageing/afaa173] [Medline: 32931547]
6. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. J Aging Phys Act 2015 Apr;23(2):314-322 [FREE Full text] [doi: 10.1123/japa.2013-0236] [Medline: 24812254]
7. Verghese J, Wang C, Holtzer R. Relationship of clinic-based gait speed measurement to limitations in community-based activities in older adults. Arch Phys Med Rehabil 2011 May;92(5):844-846 [FREE Full text] [doi: 10.1016/j.apmr.2010.12.030] [Medline: 21530734]
8. Öhlin J, Ahlgren A, Folkesson R, Gustafson Y, Littbrand H, Olofsson B, et al. The association between cognition and gait in a representative sample of very old people - the influence of dementia and walking aid use. BMC Geriatr 2020 Jan 31;20(1):34 [FREE Full text] [doi: 10.1186/s12877-020-1433-3] [Medline: 32005103]
9. Tian Q, An Y, Resnick SM, Studenski S. The relative temporal sequence of decline in mobility and cognition among initially unimpaired older adults: Results from the Baltimore Longitudinal Study of Aging. Age Ageing 2017 May 01;46(3):445-451 [FREE Full text] [doi: 10.1093/ageing/afw185] [Medline: 27744302]
10. Hackett RA, Davies-Kershaw H, Cadar D, Orrell M, Steptoe A. Walking Speed, Cognitive Function, and Dementia Risk in the English Longitudinal Study of Ageing. J Am Geriatr Soc 2018 Sep;66(9):1670-1675 [FREE Full text] [doi: 10.1111/jgs.15312] [Medline: 29508385]
11. Artaud F, Singh-Manoux A, Dugravot A, Tzourio C, Elbaz A. Decline in Fast Gait Speed as a Predictor of Disability in Older Adults. J Am Geriatr Soc 2015 Jun;63(6):1129-1136. [doi: 10.1111/jgs.13442] [Medline: 26096387]
12. Sabia S, Dumurgier J, Tavernier B, Head J, Tzourio C, Elbaz A. Change in fast walking speed preceding death: results from a prospective longitudinal cohort study. J Gerontol A Biol Sci Med Sci 2014 Mar;69(3):354-362 [FREE Full text] [doi: 10.1093/gerona/glt114] [Medline: 23913931]
13. Kitamura A, Seino S, Abe T, Nofuji Y, Yokoyama Y, Amano H, et al. Sarcopenia: prevalence, associated factors, and the risk of mortality and disability in Japanese older adults. J Cachexia Sarcopenia Muscle 2021 Feb;12(1):30-38 [FREE Full text] [doi: 10.1002/jcsm.12651] [Medline: 33241660]
14. Toots ATM, Taylor ME, Lord SR, Close JCT. Associations Between Gait Speed and Cognitive Domains in Older People with Cognitive Impairment. J Alzheimers Dis 2019;71(s1):S15-S21. [doi: 10.3233/JAD-181173] [Medline: 31256121]
15. Marino FR, Lessard DM, Saczynski JS, McManus DD, Silverman-Lloyd LG, Benson CM, et al. Gait Speed and Mood, Cognition, and Quality of Life in Older Adults With Atrial Fibrillation. J Am Heart Assoc 2019 Nov 19;8(22):e013212 [FREE Full text] [doi: 10.1161/JAHA.119.013212] [Medline: 31735113]
16. George KM, Gilsanz P, Peterson RL, Glymour MM, Mayeda ER, Mungas DM, et al. Physical Performance and Cognition in a Diverse Cohort: Kaiser Healthy Aging and Diverse Life Experiences (KHANDLE) Study. Alzheimer Dis Assoc Disord 2021;35(1):23-29 [FREE Full text] [doi: 10.1097/WAD.0000000000000428] [Medline: 33629977]
17. Sousa NMF, Macedo RC, Brucki SMD. Cross-sectional associations between cognition and mobility in Parkinson's disease. Dement Neuropsychol 2021;15(1):105-111 [FREE Full text] [doi: 10.1590/1980-57642021dn15-010011] [Medline: 33907603]
18. Callisaya ML, Blizzard CL, Wood AG, Thrift AG, Wardill T, Srikanth VK. Longitudinal Relationships Between Cognitive Decline and Gait Slowing: The Tasmanian Study of Cognition and Gait. J Gerontol A Biol Sci Med Sci 2015 Oct;70(10):1226-1232. [doi: 10.1093/gerona/glv066] [Medline: 26009641]
19. Chen X, Han P, Yu X, Zhang Y, Song P, Liu Y, et al. Relationships between sarcopenia, depressive symptoms, and mild cognitive impairment in Chinese community-dwelling older adults. J Affect Disord 2021 May 01;286:71-77. [doi: 10.1016/j.jad.2021.02.067] [Medline: 33714172]
20. Dyer AH, Lawlor B, Kennelly SP, NILVAD Study Group. Gait speed, cognition and falls in people living with mild-to-moderate Alzheimer disease: data from NILVAD. BMC Geriatr 2020 Mar 30;20(1):117 [FREE Full text] [doi: 10.1186/s12877-020-01531-w] [Medline: 32228468]
21. Gothe NP, Fanning J, Awick E, Chung D, Wójcicki TR, Olson EA, et al. Executive function processes predict mobility outcomes in older adults. J Am Geriatr Soc 2014 Feb;62(2):285-290 [FREE Full text] [doi: 10.1111/jgs.12654] [Medline: 24521364]
22. Atkinson HH, Rosano C, Simonsick EM, Williamson JD, Davis C, Ambrosius WT, Health ABC study. Cognitive function, gait speed decline, and comorbidities: the health, aging and body composition study. J Gerontol A Biol Sci Med Sci 2007 Aug;62(8):844-850. [doi: 10.1093/gerona/62.8.844] [Medline: 17702875]

23. Skillbäck T, Blennow K, Zetterberg H, Skoog J, Rydén L, Wetterberg H, et al. Slowing gait speed precedes cognitive decline by several years. Alzheimers Dement 2022 Sep;18(9):1667-1676 [FREE Full text] [doi: 10.1002/alz.12537] [Medline: 35142034]
24. Jayakody O, Breslin M, Srikanth VK, Callisaya ML. Gait Characteristics and Cognitive Decline: A Longitudinal Population-Based Study. J Alzheimers Dis 2019;71(s1):S5-S14. [doi: 10.3233/JAD-181157] [Medline: 30958358]
25. Taylor ME, Lasschuit DA, Lord SR, Delbaere K, Kurrle SE, Mikolaizak AS, et al. Slow gait speed is associated with executive function decline in older people with mild to moderate dementia: A one year longitudinal study. Arch Gerontol Geriatr 2017 Nov;73:148-153. [doi: 10.1016/j.archger.2017.07.023] [Medline: 28818760]
26. Chou M, Nishita Y, Nakagawa T, Tange C, Tomida M, Shimokata H, et al. Role of gait speed and grip strength in predicting 10-year cognitive decline among community-dwelling older people. BMC Geriatr 2019 Jul 05;19(1):186 [FREE Full text] [doi: 10.1186/s12877-019-1199-7] [Medline: 31277579]
27. Krall JR, Carlson MC, Fried LP, Xue Q. Examining the dynamic, bidirectional associations between cognitive and physical functioning in older adults. Am J Epidemiol 2014 Oct 15;180(8):838-846 [FREE Full text] [doi: 10.1093/aje/kwu198] [Medline: 25205829]
28. Gale CR, Allerhand M, Sayer AA, Cooper C, Deary IJ. The dynamic relationship between cognitive function and walking speed: the English Longitudinal Study of Ageing. Age (Dordr) 2014;36(4):9682 [FREE Full text] [doi: 10.1007/s11357-014-9682-8] [Medline: 24997019]
29. Best JR, Liu-Ambrose T, Boudreau RM, Ayonayon HN, Satterfield S, Simonsick EM, Health, AgingBody Composition Study. An Evaluation of the Longitudinal, Bidirectional Associations Between Gait Speed and Cognition in Older Women and Men. J Gerontol A Biol Sci Med Sci 2016 Dec;71(12):1616-1623 [FREE Full text] [doi: 10.1093/gerona/glw066] [Medline: 27069098]
30. Atkinson HH, Rapp SR, Williamson JD, Lovato J, Absher JR, Gass M, et al. The relationship between cognitive function and physical performance in older women: results from the women's health initiative memory study. J Gerontol A Biol Sci Med Sci 2010 Mar;65(3):300-306 [FREE Full text] [doi: 10.1093/gerona/glp149] [Medline: 19789197]
31. Mielke MM, Roberts RO, Savica R, Cha R, Drubach DI, Christianson T, et al. Assessing the temporal relationship between cognition and gait: slow gait predicts cognitive decline in the Mayo Clinic Study of Aging. J Gerontol A Biol Sci Med Sci 2013 Aug;68(8):929-937 [FREE Full text] [doi: 10.1093/gerona/gls256] [Medline: 23250002]
32. Selig J, Little T. Autoregressive and Cross-lagged Panel Analysis for Longitudinal Data. Handbook of Developmental Research Methods. New York, NY: The Guilford Press; 2012:265-278.
33. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
34. Meng Q, Wang H, Strauss J, Langa KM, Chen X, Wang M, et al. Validation of neuropsychological tests for the China Health and Retirement Longitudinal Study Harmonized Cognitive Assessment Protocol. Int Psychogeriatr 2019 Dec;31(12):1709-1719 [FREE Full text] [doi: 10.1017/S1041610219000693] [Medline: 31309907]
35. Li H, Li C, Wang A, Qi Y, Feng W, Hou C, et al. Associations between social and intellectual activities with cognitive trajectories in Chinese middle-aged and older adults: a nationally representative cohort study. Alzheimers Res Ther 2020 Sep 25;12(1):115 [FREE Full text] [doi: 10.1186/s13195-020-00691-6] [Medline: 32977839]
36. Luo Y, Zhong Y, Pang L, Zhao Y, Liang R, Zheng X. The effects of indoor air pollution from solid fuel use on cognitive function among middle-aged and older population in China. Sci Total Environ 2021 Feb 01;754:142460. [doi: 10.1016/j.scitotenv.2020.142460] [Medline: 33254849]
37. Qin T, Yan M, Fu Z, Song Y, Lu W, Fu A, et al. Association between anemia and cognitive decline among Chinese middle-aged and elderly: evidence from the China health and retirement longitudinal study. BMC Geriatr 2019 Nov 12;19(1):305 [FREE Full text] [doi: 10.1186/s12877-019-1308-7] [Medline: 31718564]
38. Quinn TJ, McArthur K, Ellis G, Stott DJ. Functional assessment in older people. BMJ 2011 Aug 22;343:d4681 [FREE Full text] [doi: 10.1136/bmj.d4681] [Medline: 21859792]
39. Lei X, Sun X, Strauss J, Zhang P, Zhao Y. Depressive symptoms and SES among the mid-aged and elderly in China: evidence from the China Health and Retirement Longitudinal Study national baseline. Soc Sci Med 2014 Nov;120:224-232 [FREE Full text] [doi: 10.1016/j.socscimed.2014.09.028] [Medline: 25261616]
40. Kearney M. Cross-lagged panel analysis. In: Allen MR, editor. Sage Encyclopedia of Communication Research Methods. London, UK: Sage; 2017.
41. Zhang T, Zhang H, Li Y, Sun D, Li S, Fernandez C, et al. Temporal Relationship Between Childhood Body Mass Index and Insulin and Its Impact on Adult Hypertension: The Bogalusa Heart Study. Hypertension 2016 Sep;68(3):818-823 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.116.07991] [Medline: 27432860]
42. Yang Y, Cao L, Xia Y, Li J. The Effect of Living Environmental Factors on Cardiovascular Diseases in Chinese Adults: Results from a Cross-sectional and Longitudinal Study. Eur J Prev Cardiol 2022 Dec 20:zwac304. [doi: 10.1093/eurjpc/zwac304] [Medline: 36537654]
43. Lin L, Cao B, Chen W, Li J, Zhang Y, Guo VY. Association of Adverse Childhood Experiences and Social Isolation With Later-Life Cognitive Function Among Adults in China. JAMA Netw Open 2022 Nov 01;5(11):e2241714 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.41714] [Medline: 36367722]

44. Soumaré A, Tavernier B, Alpérovitch A, Tzourio C, Elbaz A. A cross-sectional and longitudinal study of the relationship between walking speed and cognitive function in community-dwelling elderly people. J Gerontol A Biol Sci Med Sci 2009 Oct;64(10):1058-1065. [doi: 10.1093/gerona/glp077] [Medline: 19561146]
45. Verghese J, Mahoney J, Ambrose AF, Wang C, Holtzer R. Effect of cognitive remediation on gait in sedentary seniors. J Gerontol A Biol Sci Med Sci 2010 Dec;65(12):1338-1343. [doi: 10.1093/gerona/glq127] [Medline: 20643703]
46. Verghese J, Annweiler C, Ayers E, Barzilai N, Beauchet O, Bennett DA, et al. Motoric cognitive risk syndrome: multicountry prevalence and dementia risk. Neurology 2014 Aug 19;83(8):718-726 [FREE Full text] [doi: 10.1212/WNL.0000000000000717] [Medline: 25031288]
47. Callisaya ML, Beare R, Phan TG, Blizzard L, Thrift AG, Chen J, et al. Brain structural change and gait decline: a longitudinal population-based study. J Am Geriatr Soc 2013 Jul;61(7):1074-1079. [doi: 10.1111/jgs.12331] [Medline: 23796055]
48. Holtzer R, Epstein N, Mahoney JR, Izzetoglu M, Blumen HM. Neuroimaging of mobility in aging: a targeted review. J Gerontol A Biol Sci Med Sci 2014 Nov;69(11):1375-1388 [FREE Full text] [doi: 10.1093/gerona/glu052] [Medline: 24739495]
49. Stringhini S, Carmeli C, Jokela M, Avendaño M, McCrory C, d'Errico A, LIFEPATH Consortium. Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: multi-cohort population based study. BMJ 2018 Mar 23;360:k1046 [FREE Full text] [doi: 10.1136/bmj.k1046] [Medline: 29572376]
50. Heuninckx S, Wenderoth N, Debaere F, Peeters R, Swinnen SP. Neural basis of aging: the penetration of cognition into action control. J Neurosci 2005 Jul 20;25(29):6787-6796 [FREE Full text] [doi: 10.1523/JNEUROSCI.1263-05.2005] [Medline: 16033888]
51. Gateway to Global Aging Data. URL: https://g2aging.org/ [accessed 2023-02-26]
52. China Health and Retirement Longitudinal Study. URL: http://charls.pku.edu.cn/en [accessed 2023-02-26]

Abbreviations

ADL: activities of daily living
CFI: comparative fit index
CHARLS: China Health and Retirement Longitudinal Study
SRMR: standardized root-mean-square residual

Edited by A Mavragani, T Sanchez; submitted 14.11.22; peer-reviewed by Y An, Y Zhao; comments to author 19.12.22; revised version received 08.01.23; accepted 19.01.23; published 14.03.23.

Please cite as:
Li H, Zhang J, Zou X, Jia X, Zheng D, Guo X, Xie W, Yang Q
The Bidirectional Association Between Cognitive Function and Gait Speed in Chinese Older Adults: Longitudinal Observational Study
JMIR Public Health Surveill 2023;9:e44274
URL: https://publichealth.jmir.org/2023/1/e44274
doi:10.2196/44274
PMID:36917163

Original Paper

Electronic Health Record–Based Absolute Risk Prediction Model for Esophageal Cancer in the Chinese Population: Model Development and External Validation

Yuting Han[1*], PhD; Xia Zhu[2,3*], PhD; Yizhen Hu[1], PhD; Canqing Yu[1,4], PhD; Yu Guo[5], MSc; Dong Hang[2,3], PhD; Yuanjie Pang[1], DPhil; Pei Pei[5], MSc; Hongxia Ma[2,3], PhD; Dianjianyi Sun[1,4], PhD; Ling Yang[6,7], PhD; Yiping Chen[6,7], DPhil; Huaidong Du[6,7], PhD; Min Yu[8], MSc; Junshi Chen[9], MD; Zhengming Chen[7], DPhil; Dezheng Huo[10], PhD; Guangfu Jin[2,3], PhD; Jun Lv[1,4], PhD; Zhibin Hu[2,3], PhD; Hongbing Shen[2,3], PhD; Liming Li[1,4], MPH

[1]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China

[2]Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China

[3]Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Medicine, China International Cooperation Center for Environment and Human Health, Nanjing Medical University, Nanjing, China

[4]Peking University Center for Public Health and Epidemic Preparedness and Response, Beijing, China

[5]Chinese Academy of Medical Sciences, Beijing, China

[6]Medical Research Council Population Health Research Unit, University of Oxford, Oxford, United Kingdom

[7]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[8]Zhejiang Center for Disease Control and Prevention, Hangzhou, China

[9]China National Center for Food Safety Risk Assessment, Beijing, China

[10]Department of Public Health Sciences, The University of Chicago, Chicago, IL, United States

[*]these authors contributed equally

Corresponding Author:
Jun Lv, PhD
Department of Epidemiology and Biostatistics
School of Public Health
Peking University
No 38 Xueyuan Rd, Haidian District
Beijing, 100191
China
Phone: 86 10 82801528
Fax: 86 10 82801528
Email: lvjun@bjmu.edu.cn

Abstract

Background: China has the largest burden of esophageal cancer (EC). Prediction models can be used to identify high-risk individuals for intensive lifestyle interventions and endoscopy screening. However, the current prediction models are limited by small sample size and a lack of external validation, and none of them can be embedded into the booming electronic health records (EHRs) in China.

Objective: This study aims to develop and validate absolute risk prediction models for EC in the Chinese population. In particular, we assessed whether models that contain only EHR-available predictors performed well.

Methods: A prospective cohort recruiting 510,145 participants free of cancer from both high EC-risk and low EC-risk areas in China was used to develop EC models. Another prospective cohort of 18,441 participants was used for validation. A flexible parametric model was used to develop a 10-year absolute risk model by considering the competing risks (full model). The full model was then abbreviated by keeping only EHR-available predictors. We internally and externally validated the models by using the area under the receiver operating characteristic curve (AUC) and calibration plots and compared them based on classification measures.

Results: During a median of 11.1 years of follow-up, we observed 2550 EC incident cases. The models consisted of age, sex, regional EC-risk level (high-risk areas: 2 study regions; low-risk areas: 8 regions), education, family history of cancer (simple model), smoking, alcohol use, BMI (intermediate model), physical activity, hot tea consumption, and fresh fruit consumption (full model). The performance was only slightly compromised after the abbreviation. The simple and intermediate models showed good calibration and excellent discriminating ability with AUCs (95% CIs) of 0.822 (0.783-0.861) and 0.830 (0.792-0.867) in the external validation and 0.871 (0.858-0.884) and 0.879 (0.867-0.892) in the internal validation, respectively.

Conclusions: Three nested 10-year EC absolute risk prediction models for Chinese adults aged 30-79 years were developed and validated, which may be particularly useful for populations in low EC-risk areas. Even the simple model with only 5 predictors available from EHRs had excellent discrimination and good calibration, indicating its potential for broader use in tailored EC prevention. The simple and intermediate models have the potential to be widely used for both primary and secondary prevention of EC.

(JMIR Public Health Surveill 2023;9:e43725) doi:10.2196/43725

KEYWORDS

esophageal cancer; prediction model; absolute risk; China; prospective cohort; screening; primary prevention; development; external validation; electronic health record

Introduction

China has the largest burden of esophageal cancer (EC), accounting for around half of the global incident cases and deaths in 2018 [1,2]. The prevalence, disability-adjusted life years, and direct medical expenditures are projected to continue to increase [3]. Upper endoscopy has been widely performed for screening and diagnosing EC, but the cost and potential harm of invasive procedures as well as the need for expertise and endoscopy skills training preclude a population-wide application, which may partially explain the poor prognosis of EC. Thus, identifying a high-risk population for endoscopy through prediction models would be more feasible and effective.

In China, 4 diagnostic models (ie, estimating the probability of prevalent EC) have been developed to act as a prescreening tool for endoscopy, with an area under the receiver operating characteristic curve (AUC) ranging from 0.681 to 0.843 [4-7]. However, these models were all developed from populations in high-risk rural areas and may not apply to low-risk rural and urban areas, where a large proportion of cases countrywide occurs [8]. Only 1 model was externally validated [5]. Besides diagnostic models, a few prognostic models (ie, predicting the absolute risk of EC in 5 or more years) have recently been developed from European cohorts [9-12]. These models can be used not only for early detection but also for primary prevention such as intensive lifestyle interventions. However, such models may not perform well for the Chinese population since the risk factor profile is different and the predominant subtype of EC is esophageal squamous cell carcinoma (ESCC) among the Chinese, while esophageal adenocarcinoma is the predominant subtype in the western population [13]. To the best of our knowledge, there is only 1 prognostic model in China, which was developed in a high-risk rural area [14]; this model was based on a case-control design, which was prone to selection bias and recall bias. Moreover, the limited EC cases (n=244) and the lack of external validation could induce overfitting and encroach generalizability.

The national Basic Public Health Service program in China requires establishing health records for all residents [15]. The efficiency and extensive use of population risk stratification for tailored prevention can be greatly improved by embedding prediction models within the electronic health record (EHR) system, that is, by directly estimating the risk of EC and identifying high-risk individuals for EC based on information from EHRs. However, some predictors in the existing models (eg, food temperature, eating speed) are not available in EHRs and need to be additionally collected even in high-risk areas of EC.

To address the above limitations, we used a large prospective cohort of 0.5 million people from both high EC-risk and low EC-risk areas of China for model development and another prospective cohort for external validation. We first constructed a 10-year absolute risk prediction model for EC with the inclusion of established and probable EC risk factors to maximize model performance. Then, we simplified the model by keeping predictors that are available in the Basic Public Health Service health records. We assessed whether the simple but potentially widely applicable model showed acceptable performance in both cohorts.

Methods

Data for Model Development

Data from the China Kadoorie Biobank (CKB), a large-scale nationwide prospective cohort of 512,725 participants aged 30-79 years, were used for model development. The baseline survey was performed between 2004 and 2008 in 10 geographically defined regions (5 urban and 5 rural). The details of the study design and survey methods have been reported previously [16]. Incident cases of EC and all-cause mortality were identified through linkage with the mortality and disease registries and national health insurance claims database, supplemented with local residential records and annual active confirmation. The International Classification of Diseases, 10th revision was used to code all EC (C15) by trained staff who were blinded to the baseline information. The adjudication of the incident cancer cases is ongoing, with medical records of 1283 EC cases having been retrieved, in which 1246 (97.1%) were confirmed as EC and 830 had pathological diagnoses. After excluding 41 cases with subtype reported as unknown, 92.7% (731/789) of the cases were classified as ESCC.

Data for External Validation

An independent prospective cohort from Changzhou of the Jiangsu province, a low EC-risk rural area in China, was used for external validation. In brief, 20,803 participants aged 30 years and older were recruited from 23 villages in 2004-2005. Incident EC cases and all-cause mortality were identified through active follow-up in 2008-2009, 2012-2013, and 2018-2019, and through linkage with the disease and mortality registries. Trained staff who were blinded to baseline information further confirmed suspected cases of nonfatal cancer by reviewing local medical records or visiting village doctors.

Ethics Approval

The study protocol for CKB was approved by the ethics review committee of the Chinese Center for Disease Control and Prevention (Beijing, China: 005/2004) and the Oxford Tropical Research Ethics Committee, University of Oxford (UK: 025–04). The Changzhou cohort was approved by the ethical review committee of the Nanjing Medical University (Nanjing, China), and written informed consent was collected from all the participants.

Predictor Variables

At baseline, all participants in the CKB and Changzhou cohort completed a questionnaire and had physical measurements taken. Candidate predictors were identified based on established risk factors for EC and factors that have been included in previous EC prediction models [17,18]. Candidate predictors included age, sex, smoking, alcohol use, education, household income, marital status, family history of EC, BMI, waist circumference, physical activity, hot food consumption, and consumption of fresh vegetables, fresh fruit, red meat, and preserved vegetables. To model the large geographic disparity in EC incidence in China, we created a variable to denote living in a high-risk or low-risk area. Of the 10 study regions of CKB, we assigned Hui county in Henan province and Pengzhou in Sichuan province to high-risk areas, according to the most recent guideline for EC in China [19,20]. The criteria for defining high-risk areas are described in Multimedia Appendix 1 [19,21-26]. The details of baseline prevalence and incidence of EC by study region are shown in Multimedia Appendix 2. Because data on the family history of EC and hot food consumption were not recorded at baseline in the CKB, we used family history of cancer and hot tea consumption as surrogates for the above-established risk factors. The details of the assessment of predictors are shown in Multimedia Appendix 1.

Statistical Methods

In the CKB cohort, participants who were previously diagnosed with cancer (n=2578) or had missing data on BMI (n=2) were excluded, leaving 510,145 participants for development. In the Changzhou cohort, participants who were previously diagnosed with cancer (n=239), out of the age range of 30-79 years (n=1902), had a recorded implausible censoring date for loss to follow-up (n=5), or had missing data on candidate predictors (n=216) were excluded, leaving 18,441 participants for external validation. Participants were considered at risk from enrollment to the first date of diagnosis of EC, death, loss to follow-up, or end of follow-up (CKB: December 31, 2017; Changzhou cohort: January 31, 2019).

Model Development

Based on the whole CKB data set, we separately fitted a model for EC and a model for all-cause mortality. For the EC model, a flexible parametric model on the cumulative hazard scale was used to estimate the baseline hazards and hazard ratios of the predictors for EC, with age as the time scale [27]. Age was modeled using restricted cubic splines with boundary knots at 30 and 90 and internal knots at 60 and 70. The established risk factors of EC (age, sex, smoking, and alcohol use) and regional risk level (high-risk/low-risk areas) entered the model directly. Two strategies were employed for the selection of the other predictors. First, other candidate predictors were included in the full model and kept if *P*<.05. Second, the predictor selection was repeated using stepwise backward elimination. Two strategies selected the same set of predictors. The variable grouping was determined using the Bayesian information criteria. All 2-way interactions were tested, but none of those significantly improved model performance. Further, we simplified the full model by keeping only predictors available in the health records. As age is the most important predictor, we also constructed an age-only model for comparison. Therefore, 4 models were constructed, with predictors included in the model: (1) age-only: age; (2) simple model: age, sex, regional risk level, education, and family history of cancer, which are available for all residents in the health records; (3) intermediate model: simple model plus smoking, alcohol use, and BMI, which are additionally available for residents aged 65 years and older, and diabetic or hypertensive patients in the health records; (4) full model: intermediate model plus physical activity, hot tea consumption, and fresh fruit consumption, which go beyond the available health records but have the potential to improve the risk prediction. We then used the same settings of the flexible parametric model to model the hazards of all-cause mortality, with sex, residence area (urban/rural), and regional risk level in the model. We used cause-specific hazard models to account for the competing risks. Briefly, the 10-year absolute risk (AR) of EC for a participant who is age a is calculated as

Model Validation

The methods for model validation are detailed in Multimedia Appendix 1. In brief, we externally validated the age-only, simple, and intermediate models, but not the full model, because data on physical activity in metabolic equivalent of task-hours and hot tea consumption were unavailable in the Changzhou cohort. We also conducted an internal validation in the CKB by using data splitting and 500-sample bootstrapping. Cancer-free participants whose retention in the cohorts was less than 10 years were included to test calibration but were excluded from other validation measures, since it was unknown whether they could have experienced an EC if they had been followed up to 10 years.

Discrimination was quantified by calculating the AUC. Calibration was assessed by plotting the observed risk obtained using Kaplan-Meier analyses against the predicted risk by decile. Because of the large geographical variation in the incidence of EC in China, we recalibrated the models by using the method proposed by the World Health Organization Cardiovascular Disease Risk Chart Working Group with a slight modification [21]. Further, continuous Net Reclassification Improvement and Integrated Discrimination Improvement were used to evaluate the added predictive ability of additional predictors [28,29]. In the internal validation using data splitting, calibration and discrimination were also assessed in subgroups defined by regional risk level, residence area, sex, age group, and special population aged 65 years and older or with diabetes or hypertension who are of particular concern to the Basic Public Health Service. To offer a reference for primary care practices, we estimated a range of performance indices corresponding to a series of cutoffs.

Several sensitivity analyses were conducted. First, we separately developed 2 models for high EC-risk (high-risk model) and low EC-risk (low-risk model) areas by using the same strategy as the primary analyses and assessed their discrimination and calibration in the corresponding areas. Second, we restricted EC cases to (1) pathologically confirmed cases, (2) cases that were pathologically confirmed as ESCC, (3) cases that were pathologically confirmed but not as ESCC, and (4) cases that were pathologically confirmed but not as ESCC (scenario 3) or that were not pathologically confirmed. In the above 4 scenarios, we excluded EC cases that did not meet the corresponding criteria and examined the discriminating ability of the models (Multimedia Appendix 3). Third, since some asymptomatic EC cases might be undiagnosed, we excluded the EC cases documented in the first year of follow-up and used the same strategy to develop and validate the models.

Results

The mean age of the 510,145 participants in the CKB and 18,441 participants in the Changzhou cohort was 52.0 (SD 10.7) years and 51.2 (SD 12.1) years, respectively. The details of the baseline characteristics of the predictors are described in Table 1. During a median of 11.1 (IQR 10.2-12.1) years of follow-up of the CKB, we identified 2550 EC cases, with an incidence (per 100,000 person-years) of 46.2. High EC-risk areas had a significantly higher incidence than low EC-risk areas (132.2 vs 20.2, respectively). In the Changzhou cohort, 114 EC cases were identified during a median follow-up of 13.6 (IQR 13.5-14.4) years, with an incidence of 47.1.

Table 2 and Multimedia Appendix 4 list the hazard ratios and 95% CIs for predictors of EC and all-cause mortality in the CKB. Male, living in high-risk areas, less educated, having a family history of cancer, smoking, alcohol use, underweight, less physical activity, preferring burning hot tea, and rare intake of fresh fruits were associated with a greater risk of EC.

In the external validation, the simple and intermediate models exhibited similar and excellent discriminating ability with AUCs (95% CIs) of 0.822 (0.783-0.861) and 0.830 (0.792-0.867), respectively (Figure 1). In the internal validation, the AUCs (95% CIs) of the simple, intermediate, and full models were 0.871 (0.858-0.884), 0.879 (0.867-0.892), and 0.883 (0.871-0.895), respectively (Figure 1). Although there were only limited increases in the AUCs with more predictors included in the models, continuous Net Reclassification Improvement and Integrated Discrimination Improvement indicated improved accuracy of the predicted risks for both cases and those that were not cases (Multimedia Appendix 5). In the less biased internal validation method of bootstrapping, the above results were not greatly altered (Multimedia Appendix 6). The original simple and intermediate models significantly underestimated the risk of EC in the Changzhou cohort. The recalibration parameters, *b* and *k*, were 1.22 and 1.97, respectively. Age-specific observed risks of EC used to calculate *b* and *k* are shown in Multimedia Appendix 7. After recalibration, the calibration plot showed excellent agreement between the observed and predicted risks for the simple and intermediate models (Figure 2). In the internal validation, the predicted risk of the simple, intermediate, and full models agreed well with the observed risk by a tenth of the predicted risk, except for the top 2 deciles where slight underestimations seemed to have occurred (Figure 3).

The density of the predicted risks of models in cases was greater than that in those that were not cases (Multimedia Appendix 8 and Multimedia Appendix 9). The performance of the models across a series of cutoffs is presented in Multimedia Appendix 10. Compared with their counterparts, the models discriminated better in low-risk areas, rural areas, women, or middle-aged adults younger than 65 years without diabetes and hypertension in the internal validation (Multimedia Appendix 11). The predicted risks agreed well with the observed risks in all subgroups.

In the sensitivity analysis, we separately developed 2 models for high-risk and low-risk areas. The included predictors and the hazard ratios (95% CIs) are listed in Multimedia Appendix 12. When these 2 models were applied in their corresponding validation set, the model for low-risk areas performed better than the models in the primary analyses (Multimedia Appendix 5 and Multimedia Appendix 13). When we took the availability and results of pathology reports into consideration, models had excellent discriminating ability in all scenarios (Multimedia Appendix 3 and Multimedia Appendix 14). Excluding EC cases occurring in the first year of follow-up did not alter the performance of the models (Multimedia Appendix 15).

Table 1. Baseline characteristics of the participants by disease status in the China Kadoorie Biobank and Changzhou cohort.

	China Kadoorie Biobank			Changzhou cohort		
	EC[a] case (n=2550)	Not an EC case (n=507,595)	Total (N=510,145)	EC case (n=114)	Not an EC case (n=18,327)	Total (N=18,441)
Age (years), mean (SD)	60.7 (8.7)	52.0 (10.7)	52.0 (10.7)	60.9 (9.1)	51.1 (12.1)	51.2 (12.1)
Male, n (%)	1757 (68.9)	207,477 (40.9)	209,234 (41)	77 (67.5)	7611 (41.5)	7688 (41.7)
Urban, n (%)	468 (18.4)	224,300 (44.2)	224,768 (44.1)	0 (0)	0 (0)	0 (0)
High-risk area[b], n (%)	1692 (66.4)	116,715 (23)	118,407 (23.2)	0 (0)	0 (0)	0 (0)
Family history of cancer, n (%)	720 (28.2)	84,948 (16.7)	85,668 (16.8)	26 (22.8)	3354 (18.3)	3380 (18.3)
High level of physical activity[c], n (%)	565 (22.2)	126,739 (25)	127,304 (25)	—[d]	—	—
Highest education, n (%)						
Illiterate or primary school	1917 (75.2)	257,088 (50.6)	259,005 (50.8)	70 (61.4)	8372 (45.7)	8442 (45.8)
Middle or high school	598 (23.5)	220,780 (43.5)	221,378 (43.4)	43 (37.7)	9841 (53.7)	9884 (53.6)
College or university	35 (1.4)	29,727 (5.9)	29,762 (5.8)	1 (0.9)	114 (0.6)	115 (0.6)
Current smoking (cigarettes or equivalent per day), n (%)						
<30	979 (38.4)	114,815 (22.6)	115,794 (22.7)	49 (43)	4230 (23.1)	4279 (23.2)
≥30	235 (9.2)	19,155 (3.8)	19,390 (3.8)	8 (7)	713 (3.9)	721 (3.9)
Daily alcohol use (grams of pure alcohol per day), n (%)						
<30	70 (2.7)	11,503 (2.3)	11,573 (2.3)	9 (7.9)	1123 (6.1)	1132 (6.1)
30-59	161 (6.3)	14,884 (2.9)	15,045 (2.9)	9 (7.9)	1049 (5.7)	1058 (5.7)
≥60	386 (15.1)	19,085 (3.8)	19,471 (3.8)	31 (29)	1956 (10.7)	1989 (10.8)
BMI (kg/m^2), n (%)						
<18.5	175 (6.9)	21,965 (4.3)	22,140 (4.3)	4 (3.5)	984 (5.4)	988 (5.4)
18.5-23.9	1482 (58.1)	263,169 (51.8)	264,651 (51.9)	76 (66.7)	9967 (54.4)	10,043 (54.4)
≥24.0	893 (35)	222,461 (43.8)	223,354 (43.8)	34 (29.8)	7376 (40.3)	7410 (40.2)
Tea temperature preference, n (%)						
Not daily drinker/warm tea drinker	2090 (82)	426,628 (84)	428,718 (84)	—	—	—
Hot tea	311 (12.2)	59,425 (11.7)	59,736 (11.7)	—	—	—
Burning hot tea	149 (5.8)	21,542 (4.2)	21,691 (4.3)	—	—	—
Fresh fruit consumption[e], n (%)						
Daily	165 (6.5)	95,715 (18.9)	95,880 (18.8)	4 (3.5)	813 (4.4)	817 (4.4)
Weekly	658 (25.8)	207,716 (40.9)	208,374 (40.8)	98 (86)	15,921 (86.9)	16,019 (86.9)
Less than weekly	1727 (67.7)	204,164 (40.2)	205,891 (40.4)	12 (10.5)	1591 (8.7)	1603 (8.7)

[a] EC: esophageal cancer.

[b] High-risk area denotes Hui county in Henan province and Pengzhou in Sichuan province in our study.

[c] High-level physical activity was defined as age-specific and sex-specific upper quarter of total physical activity level measured by metabolic equivalent of task-hours per day.

[d] Not available.

[e] Data on the fresh fruit consumption of 2 participants in the Changzhou cohort were missing.

Table 2. Hazard ratios (95% CIs) for the predictor variables of esophageal cancer in the China Kadoorie Biobank.

	Cases (n)	Cases/person years (1/100,000)	Age-only model, HR[a] (95% CI)	Simple model, HR (95% CI)	Intermediate model, HR (95% CI)	Full model, HR (95% CI)
Spline basis of age (knots: 30, 60, 70, 90)						
First	N/A[b]	N/A	3.56 (3.29-3.86)	3.28 (3.03-3.56)	3.30 (3.04-3.59)	3.28 (3.02-3.56)
Second	N/A	N/A	1.17 (1.11-1.23)	1.15 (1.09-1.21)	1.14 (1.09-1.20)	1.14 (1.09-1.20)
Third	N/A	N/A	1.02 (1.00-1.04)	1.00 (0.98-1.02)	1.00 (0.98-1.02)	1.00 (0.98-1.02)
Sex						
Male	1757	79.12	N/A	Reference	Reference	Reference
Female	793	24.02	N/A	0.31 (0.28-0.34)	0.40 (0.36-0.44)	0.42 (0.37-0.46)
High-risk area[c]						
No	858	20.22	N/A	Reference	Reference	Reference
Yes	1692	132.22	N/A	6.31 (5.81-6.86)	6.07 (5.58-6.60)	5.61 (5.14-6.13)
Highest education						
Illiterate or primary school	1917	69.06	N/A	Reference	Reference	Reference
Middle or high school	598	24.68	N/A	0.60 (0.55-0.66)	0.65 (0.59-0.72)	0.68 (0.62-0.76)
College or university	35	10.82	N/A	0.32 (0.23-0.44)	0.37 (0.26-0.52)	0.45 (0.32-0.63)
Family history of cancer						
No	1830	39.88	N/A	Reference	Reference	Reference
Yes	720	77.14	N/A	1.71 (1.57-1.86)	1.78 (1.63-1.94)	1.74 (1.59-1.89)
Current smoking						
No	1336	32.72	N/A	N/A	Reference	Reference
Cigarettes or equivalent per day among smokers						
<30	979	79.48	N/A	N/A	1.15 (1.05-1.27)	1.12 (1.02-1.24)
≥30	235	112.95	N/A	N/A	1.25 (1.07-1.47)	1.24 (1.06-1.45)
Daily alcohol use						
No	1933	38.43	N/A	N/A	Reference	Reference
Grams of pure alcohol per day among alcohol consumers						
<30	70	56.84	N/A	N/A	0.97 (0.77-1.24)	1.03 (0.81-1.30)
30-59	161	100.03	N/A	N/A	1.39 (1.18-1.64)	1.43 (1.21-1.69)
≥60	386	185.48	N/A	N/A	2.01 (1.78-2.26)	2.06 (1.82-2.32)
BMI (kg/m^2)						
<18.5	175	77.77	N/A	N/A	Reference	Reference
18.5-23.9	1482	51.70	N/A	N/A	0.73 (0.62-0.85)	0.76 (0.65-0.88)
≥24.0	893	36.73	N/A	N/A	0.61 (0.52-0.72)	0.64 (0.54-0.76)
Physical activity						
Low	1985	48.10	N/A	N/A	N/A	Reference
High[d]	565	40.49	N/A	N/A	N/A	0.80 (0.73-0.88)
Tea temperature preference						
Not a daily drinker or warm tea drinker	2090	44.96	N/A	N/A	N/A	Reference

	Cases (n)	Cases/person years (1/100,000)	Age-only model, HR[a] (95% CI)	Simple model, HR (95% CI)	Intermediate model, HR (95% CI)	Full model, HR (95% CI)
Hot tea	311	48.27	N/A	N/A	N/A	1.03 (0.91-1.17)
Burning hot tea	149	64.89	N/A	N/A	N/A	1.49 (1.25-1.77)
Fresh fruit consumption						
Daily	165	15.74	N/A	N/A	N/A	Reference
Weekly	658	28.97	N/A	N/A	N/A	1.07 (0.90-1.27)
Less than weekly	1727	78.38	N/A	N/A	N/A	1.79 (1.51-2.12)

[a] HR: hazard ratio.

[b] N/A: not applicable.

[c] High-risk area denotes Hui county in Henan province and Pengzhou in Sichuan province in our study.

[d] High-level physical activity was defined as age-specific and sex-specific upper quarter of total physical activity level measured by metabolic equivalent of task-hours per day.

Figure 1. Receiver operating characteristic curves and corresponding areas under the receiver operating characteristic curve for the esophageal cancer prediction models. (A) Internal validation in the China Kadoorie Biobank using data splitting. (B) External validation in the Changzhou cohort. The models included age (age-only model), sex, regional risk level, education, family history of cancer (simple model), smoking, alcohol use, BMI (intermediate model), physical activity, hot tea consumption, and fresh fruit consumption (full model). AUC: area under the receiver operating characteristic curve.

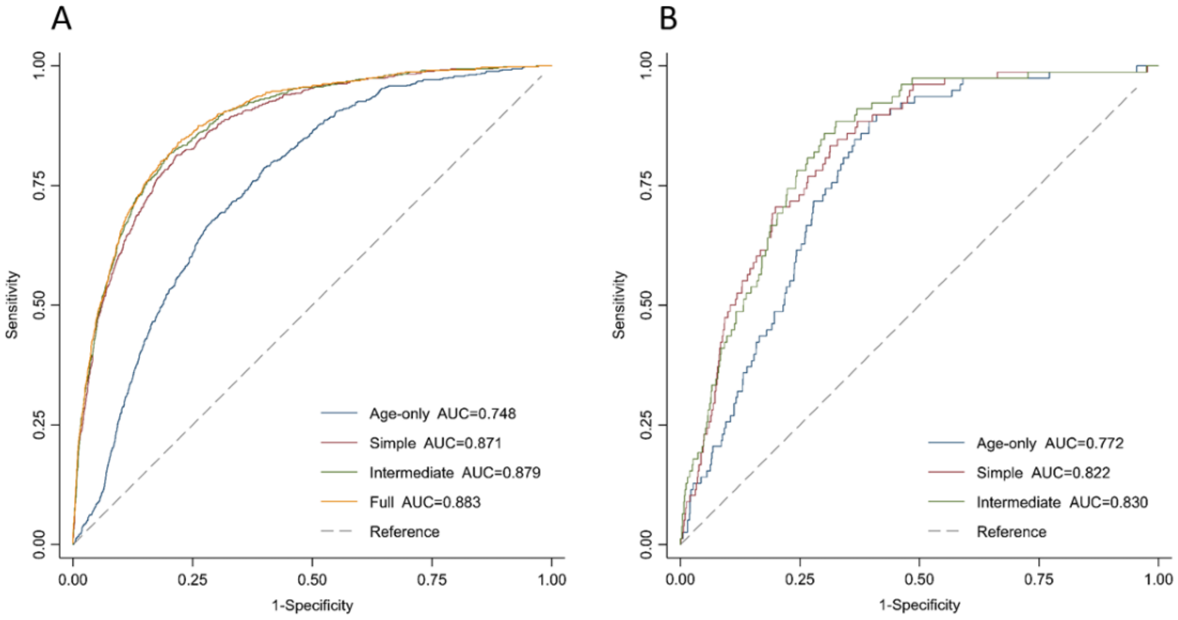


Figure 2. Calibration plot of the esophageal cancer prediction models in the Changzhou cohort. Calibration of the original (A) age-only, (C) simple, and (E) intermediate models. Calibration of the recalibrated (B) age-only, (D) simple, and (F) intermediate models. The observed 10-year risk was estimated by Kaplan-Meier analysis and plotted against model-predicted risk by decile. Models were recalibrated using the method proposed by the World Health Organization Cardiovascular Disease Risk Chart Working Group with a slight modification. For details, see Multimedia Appendix 1.

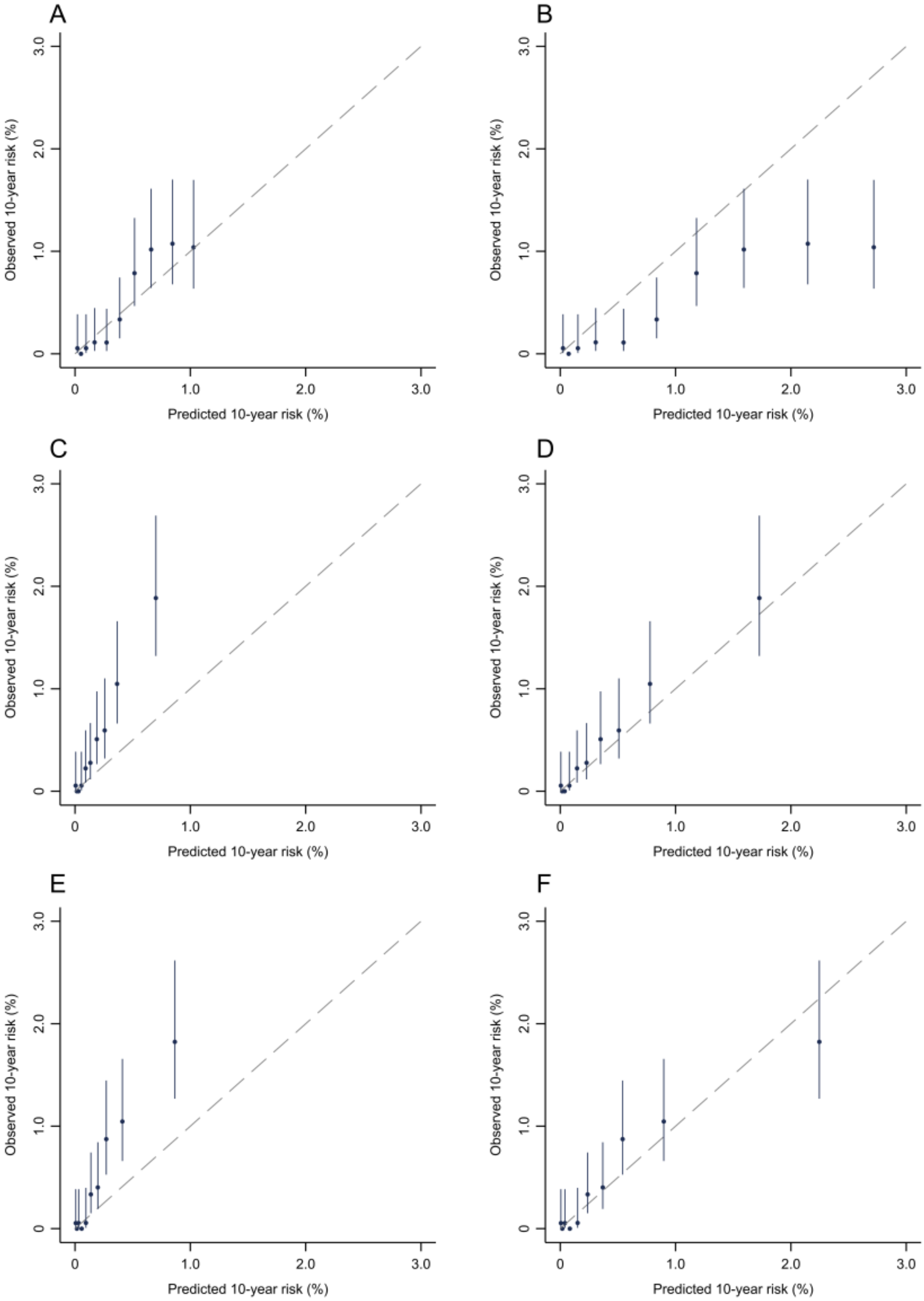
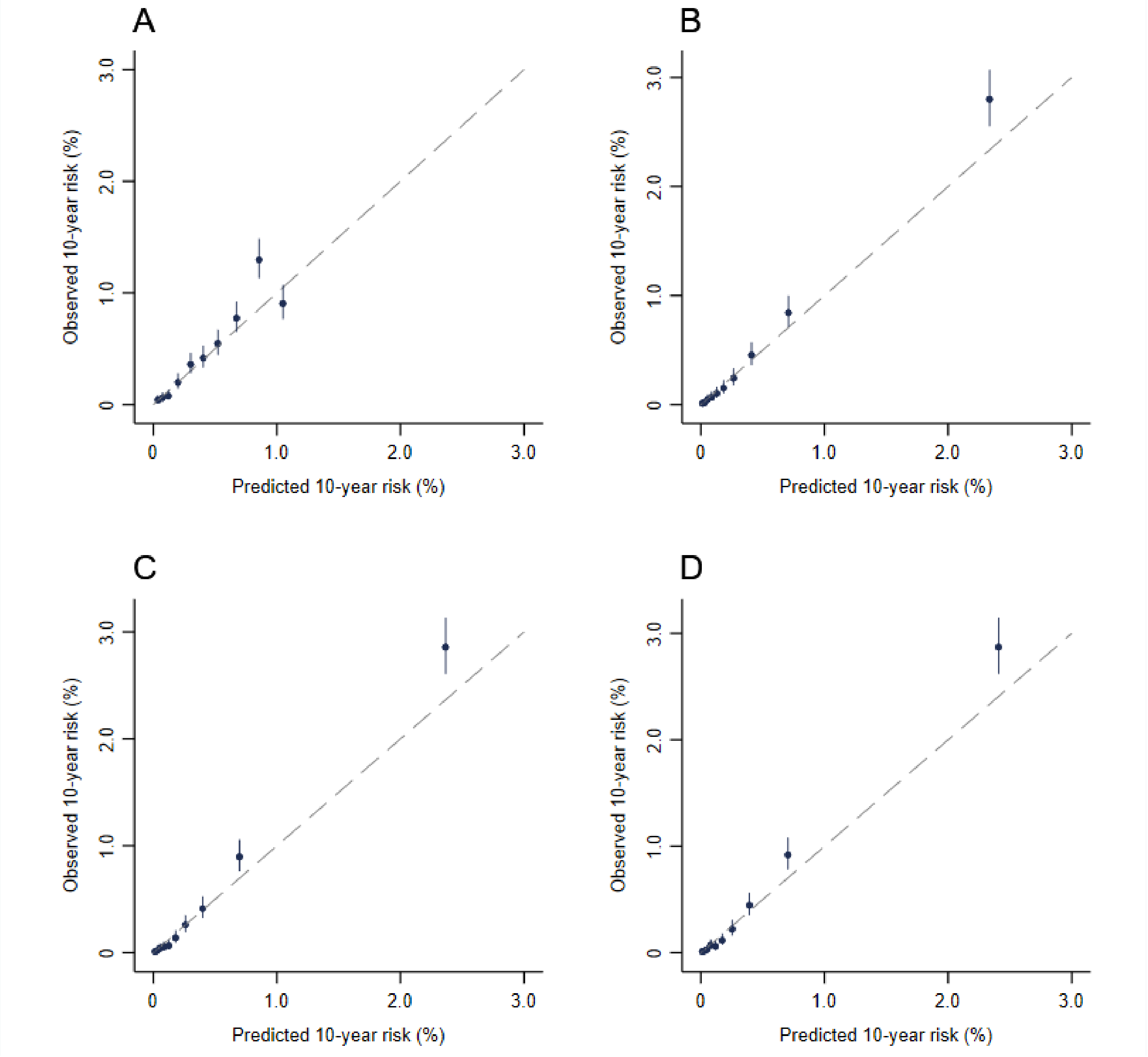


Figure 3. Calibration plot of the esophageal cancer prediction models in the China Kadoorie Biobank by using data splitting. (A) Age-only model. (B) Simple model. (C) Intermediate model. (D) Full model. Models were fitted to a random two-thirds of the China Kadoorie Biobank data and evaluated on the remaining one-third. The calibration plots in the validation set were plotted. The observed 10-year risk was estimated by Kaplan-Meier analyses and plotted against model-predicted risk by decile.



Discussion

In a large prospective cohort study, we developed 3 nested 10-year EC absolute risk prediction models for Chinese adults aged 30-79 years. The models included age, sex, regional EC-risk level, education, family history of cancer (simple model), smoking, alcohol use, BMI (intermediate model), physical activity, hot tea consumption, and fresh fruit consumption (full model). The simple and intermediate models were externally validated in an independent prospective cohort and they exhibited excellent discrimination and good calibration. The performance of these models was compromised by keeping only predictors available in the health records but only to a small and acceptable extent.

The models that we constructed included established risk factors for EC (eg, age, smoking, alcohol use) and factors associated with increased EC risk in the CKB and in previous studies [17,18]. A previous review attributed the geographical variation in the incidence of EC in China to some unique factors in high-risk areas, such as exposure to carcinogens (eg, nitrosamines, their precursors) via water, food, and other sources [30]. To capture this variation as well as to denote some unmeasured unique factors in high-risk areas, we included regional risk level in our models. Although this predictor contributed the most to the model performance with a hazard ratio of around 6 (Table 2), our stratified validation showed that the other predictors still maintained excellent performance in both high-risk and low-risk areas (Multimedia Appendix 11).

Previous models included clinical symptoms such as dysphagia and poststernal pain to identify high-risk individuals with prevalent EC for further endoscopies [4-6]. In contrast, our models, which were intended to identify individuals at risk for developing EC in the next 10 years, did not include clinical symptoms. In a prior model developed based on a hospital-based case-control study, 25 single-nucleotide polymorphisms, in addition to age, smoking, and alcohol use, resulted in an increased AUC from 0.639 to 0.707 [7]. Some other factors such as exposure to cooking fumes, pesticides, or salty foods were also included in previous models. To develop a parsimonious model that can be potentially used widely, we did not consider genetic variants and less well-established risk factors. Nevertheless, the AUCs of our models were still higher than those of most of the previous Chinese models (range 0.681-0.843) [4-7,14].

As expected, our finding that the simple and intermediate models retained similar performance as the full model despite the fewer predictors included is reasonable since the lost information due to the removal of the predictors was more or less supplemented by other correlated predictors. A previous study showed that the discriminatory information needed for the same unit of increase of AUC exponentially increased with AUC [31]. Thus, an already high AUC of >0.8 for the simple model can only be improved by highly informative predictors. Given the similar performance and excellent discriminating ability, it is acceptable to use the simple or intermediate model in situations where the EHRs are complete and up-to-date and easily implemented in a lower-cost way than an organized screening program. Further, we noticed that the same predictors could contribute differently in subpopulations. For example, the inclusion of lifestyle factors barely improved the discriminating ability in women (Multimedia Appendix 11), which may be explained by the low prevalence and dosage intensity of smoking and alcohol use in Chinese women.

For most prediction models, underestimations or overestimations are commonly observed in an external validation, which were also observed in our study. However, across the groups defined by the deciles of predicted risks, the observed risks proportionally increased with increased predicted risk rather than an irregular misestimation. More importantly, the underestimation disappeared after recalibration. Such results implied that the predictors in our models are predictive, the coefficients estimated in the CKB are robust and generalizable, and the underestimation was mainly caused by the mismatch of EC incidence between the CKB and Changzhou cohort.

Unlike the models in previous studies, our models calculated the absolute risk of EC instead of the relative risk and could facilitate primary prevention of EC. The essence of intuition of the absolute risk can not only raise population awareness and motivate adherence to lifestyle changes but also enhance effective communication between health professionals and individuals and help health professionals identify high-risk populations for intensive lifestyle interventions. Further, several predictors in our models are modifiable, such as smoking, alcohol use, and BMI, which could be treated as targets of intervention.

Our study has several strengths. We used a large prospective cohort with the largest number of EC cases from urban, rural, high EC-risk, and low EC-risk areas in China for model development and used another prospective cohort for external validation. This method ensures that our models are robust and potentially generalizable to a wide range of areas. To the best of our knowledge, ours is the first study to develop and externally validate EC models by using 2 independent prospective Chinese cohorts. Last but not the least, we developed and validated an abbreviated version of the risk prediction model that could be easily embedded within the EHR system and enable an efficient and automatic population risk stratification. To facilitate the usage of our models, we provide an easy-to-use Stata code and example in Multimedia Appendix 16 (Stata calculator, which is a modified version of the code shared by Dr Muller) [22].

Some limitations of our study merit consideration. First, some EC cases in the CKB had only clinical diagnoses but no pathological diagnoses for various reasons. Therefore, we could not exclude cases of esophageal adenocarcinoma. However, more than 90% of the EC cases are ESCC in China [32], which was confirmed by our ongoing adjudication of incident EC cases in the CKB. More importantly, models maintained high discriminating abilities when we restricted EC cases to those with a pathological diagnosis of ESCC or those without a pathological diagnosis of ESCC. Second, we did not collect information on the family history of EC specifically and their preference for hot foods and drinks in the baseline survey. Therefore, we used the family history of any cancer and preference for very hot tea consumption as surrogates. Third, although we found limited improvement in AUC by including more predictors in the model (full model), whether other established risk factors of EC, such as disease history of the esophagus and genetic predisposition of EC could further improve the AUC warrants future research. Fourth, we only externally validated our models in low EC-risk rural areas. Further validations in other areas are warranted.

In summary, using data from 2 prospective cohorts, we developed and validated 3 nested 10-year EC absolute risk prediction models for Chinese adults, which may be particularly useful for populations in low EC-risk areas. Even the simple model with only 5 predictors available from residents' EHRs showed excellent discrimination and good calibration, indicating its potential for broader use in tailored EC prevention. Further research is needed to assess the real-world performance in aiding population-wide stratification, identify optimal risk cutoffs for initiating intensive lifestyle interventions and endoscopy screening, and establish an optimal screening protocol (including multistage screening) for individuals or regions with different risks.

Acknowledgments

We thank the participants in this study and the members of the survey teams in each of the 10 regional centers as well as the project development and management teams based at Beijing, Oxford, and the 10 regional centers. This work was supported by the National Natural Science Foundation of China (82192904, 82192901, 82192900, 82125033, 81922061). Y Han is supported by the China Postdoctoral Science Foundation (BX20220018, 2022M720007). The China Kadoorie Biobank baseline survey and the first resurvey were supported by a grant from the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up is supported by grants from the UK Wellcome Trust (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), grants from the National Key R&D Program of China (2016YFC0900500), National Natural Science Foundation of China (81390540, 91846303, 81941018), and Chinese Ministry of Science and Technology (2011BAI09B01). The funders had no role in the study design, data collection, data analysis and interpretation, writing of the report, or the decision to submit the paper for publication.

Data Availability

Details of how to access the China Kadoorie Biobank data and details of the data release schedule are available from [33].

Authors' Contributions

JL and LL conceived and designed the study with equal contribution to this work. LL, ZC, and JC, as members of the China Kadoorie Biobank (CKB) steering committee, designed and supervised the conduct of the CKB study, obtained funding, and together with CY, YG, YP, PP, DS, LY, YC, HD, and MY acquired the data for the CKB study. D Hang, HM, GJ, ZH, and HS designed and supervised the conduct of the Changzhou cohort study. Y Han and Y Hu analyzed the CKB data, and XZ analyzed the Changzhou cohort data. Y Han wrote the first draft of the manuscript. JL, D Huo, and LL contributed to the interpretation of the results and critical revision of the manuscript for important intellectual content and approved the final version of the manuscript. All authors reviewed and approved the final manuscript. JL and LL are the guarantors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods.
[DOCX File , 24 KB - publichealth_v9i1e43725_app1.docx]

Multimedia Appendix 2
Prevalence and incidence of esophageal cancer by study region.
[DOCX File , 30 KB - publichealth_v9i1e43725_app2.docx]

Multimedia Appendix 3
Design of the sensitivity analysis in consideration of pathology reports.
[DOCX File , 379 KB - publichealth_v9i1e43725_app3.docx]

Multimedia Appendix 4
Hazard ratios (95% CIs) for predictor variables of all-cause mortality in the China Kadoorie Biobank.
[DOCX File , 24 KB - publichealth_v9i1e43725_app4.docx]

Multimedia Appendix 5
Comparison of the area under the receiver operating characteristic curve, continuous Net Reclassification Improvement, and Integrated Discrimination Improvement of esophageal cancer prediction models in the China Kadoorie Biobank and Changzhou cohort.
[DOCX File , 26 KB - publichealth_v9i1e43725_app5.docx]

Multimedia Appendix 6
Comparison of the area under the receiver operating characteristic curve, continuous Net Reclassification Improvement, and Integrated Discrimination Improvement of esophageal cancer prediction models in the China Kadoorie Biobank by using bootstrap.
[DOCX File , 25 KB - publichealth_v9i1e43725_app6.docx]

Multimedia Appendix 7
Age-specific observed risk of esophageal cancer in low-risk areas of the China Kadoorie Biobank and Changzhou cohort.
[DOCX File , 24 KB - publichealth_v9i1e43725_app7.docx]

Multimedia Appendix 8
Discriminating ability of the recalibrated prediction models in the Changzhou cohort.

[DOCX File , 395 KB - publichealth_v9i1e43725_app8.docx]

Multimedia Appendix 9
Discriminating ability of the prediction models in the China Kadoorie Biobank using data splitting.
[DOCX File , 394 KB - publichealth_v9i1e43725_app9.docx]

Multimedia Appendix 10
Performance of the esophageal cancer prediction model across different predicted risk cutoffs in the China Kadoorie Biobank and Changzhou cohort.
[DOCX File , 45 KB - publichealth_v9i1e43725_app10.docx]

Multimedia Appendix 11
Discrimination and calibration of the intermediate model in subcohorts of the China Kadoorie Biobank by data splitting.
[DOCX File , 569 KB - publichealth_v9i1e43725_app11.docx]

Multimedia Appendix 12
Hazard ratios (95% CIs) for predictor variables of esophageal cancer prediction models developed separately in high-risk and low-risk areas of the derivation subcohort of the China Kadoorie Biobank.
[DOCX File , 31 KB - publichealth_v9i1e43725_app12.docx]

Multimedia Appendix 13
Model performance of 2 esophageal cancer prediction models developed separately in high-risk and low-risk areas of the derivation subcohort of the China Kadoorie Biobank and applied in the corresponding validation subcohort.
[DOCX File , 74 KB - publichealth_v9i1e43725_app13.docx]

Multimedia Appendix 14
Discriminating ability of the esophageal cancer prediction models in the China Kadoorie Biobank by data splitting in consideration of pathology reports.
[DOCX File , 195 KB - publichealth_v9i1e43725_app14.docx]

Multimedia Appendix 15
Model performance of the esophageal cancer prediction models in the China Kadoorie Biobank using data splitting after excluding esophageal cancer cases occurring in the first year of follow-up.
[DOCX File , 1658 KB - publichealth_v9i1e43725_app15.docx]

Multimedia Appendix 16
Stata calculator.
[ZIP File (Zip Archive), 62 KB - publichealth_v9i1e43725_app16.zip]

References

1. Arnold M, Ferlay J, van Berge Henegouwen MI, Soerjomataram I. Global burden of oesophageal and gastric cancer by histology and subsite in 2018. Gut 2020 Sep;69(9):1564-1571. [doi: 10.1136/gutjnl-2020-321600] [Medline: 32606208]
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
3. Li Y, Xu J, Gu Y, Sun X, Dong H, Chen C. The Disease and Economic Burdens of Esophageal Cancer in China from 2013 to 2030: Dynamic Cohort Modeling Study. JMIR Public Health Surveill 2022 Mar 02;8(3):e33191 [FREE Full text] [doi: 10.2196/33191] [Medline: 34963658]
4. Chen W, Li H, Ren J, Zheng R, Shi J, Li J, et al. Selection of high-risk individuals for esophageal cancer screening: A prediction model of esophageal squamous cell carcinoma based on a multicenter screening cohort in rural China. Int J Cancer 2021 Jan 15;148(2):329-339 [FREE Full text] [doi: 10.1002/ijc.33208] [Medline: 32663318]
5. Liu Z, Guo C, He Y, Chen Y, Ji P, Fang Z, et al. A clinical model predicting the risk of esophageal high-grade lesions in opportunistic screening: a multicenter real-world study in China. Gastrointest Endosc 2020 Jun;91(6):1253-1260.e3. [doi: 10.1016/j.gie.2019.12.038] [Medline: 31911077]
6. Liu M, Liu Z, Cai H, Guo C, Li X, Zhang C, et al. A Model To Identify Individuals at High Risk for Esophageal Squamous Cell Carcinoma and Precancerous Lesions in Regions of High Prevalence in China. Clin Gastroenterol Hepatol 2017 Oct;15(10):1538-1546.e7. [doi: 10.1016/j.cgh.2017.03.019] [Medline: 28342951]

7. Chang J, Huang Y, Wei L, Ma B, Miao X, Li Y, et al. Risk prediction of esophageal squamous-cell carcinoma with common genetic variants and lifestyle factors in Chinese population. Carcinogenesis 2013 Aug;34(8):1782-1786. [doi: 10.1093/carcin/bgt106] [Medline: 23536576]
8. Chen R, Zheng RS, Zhang SW, Zeng HM, Wang SM, Sun KX, et al. [Analysis of incidence and mortality of esophageal cancer in China, 2015]. Zhonghua Yu Fang Yi Xue Za Zhi 2019 Nov 06;53(11):1094-1097. [doi: 10.3760/cma.j.issn.0253-9624.2019.11.004] [Medline: 31683393]
9. Kunzmann AT, Thrift AP, Cardwell CR, Lagergren J, Xie S, Johnston BT, McMenamin, et al. Model for Identifying Individuals at Risk for Esophageal Adenocarcinoma. Clin Gastroenterol Hepatol 2018 Aug;16(8):1229-1236.e4. [doi: 10.1016/j.cgh.2018.03.014] [Medline: 29559360]
10. Xie S, Ness-Jensen E, Medefelt N, Lagergren J. Assessing the feasibility of targeted screening for esophageal adenocarcinoma based on individual risk assessment in a population-based cohort study in Norway (The HUNT Study). Am J Gastroenterol 2018 Jun;113(6):829-835. [doi: 10.1038/s41395-018-0069-9] [Medline: 29748563]
11. Wang Q, Lagergren J, Xie S. Prediction of individuals at high absolute risk of esophageal squamous cell carcinoma. Gastrointest Endosc 2019 Apr;89(4):726-732.e2. [doi: 10.1016/j.gie.2018.10.025] [Medline: 30616974]
12. Wang Q, Ness-Jensen E, Santoni G, Xie S, Lagergren J. Development and Validation of a Risk Prediction Model for Esophageal Squamous Cell Carcinoma Using Cohort Studies. Am J Gastroenterol 2021 Apr;116(4):683-691. [doi: 10.14309/ajg.0000000000001094] [Medline: 33982937]
13. Gao QY, Fang JY. Early esophageal cancer screening in China. Best Pract Res Clin Gastroenterol 2015 Dec;29(6):885-893. [doi: 10.1016/j.bpg.2015.09.018] [Medline: 26651250]
14. Shen Y, Xie S, Zhao L, Song G, Shao Y, Hao C, et al. Estimating Individualized Absolute Risk for Esophageal Squamous Cell Carcinoma: A Population-Based Study in High-Risk Areas of China. Front Oncol 2020;10:598603. [doi: 10.3389/fonc.2020.598603] [Medline: 33489898]
15. Yuan B, Balabanova D, Gao J, Tang S, Guo Y. Strengthening public health services to achieve universal health coverage in China. BMJ 2019 Jun 21;365:l2358 [FREE Full text] [doi: 10.1136/bmj.l2358] [Medline: 31227480]
16. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, China Kadoorie Biobank (CKB) collaborative group. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011 Dec;40(6):1652-1666 [FREE Full text] [doi: 10.1093/ije/dyr120] [Medline: 22158673]
17. World Cancer Research Fund, American Insititute for Cancer Research. Diet, nutrition, physical activity and oesophageal cancer. Continuous Update Project Expert Report. URL: https://www.wcrf.org/wp-content/uploads/2021/02/oesophageal-cancer-report.pdf [accessed 2023-02-28]
18. Yu C, Tang H, Guo Y, Bian Z, Yang L, Chen Y, China Kadoorie Biobank Collaborative Group. Hot Tea Consumption and Its Interactions With Alcohol and Tobacco Use on the Risk for Esophageal Cancer: A Population-Based Cohort Study. Ann Intern Med 2018 Apr 03;168(7):489-497 [FREE Full text] [doi: 10.7326/M17-2000] [Medline: 29404576]
19. He J, Chen WQ, Li ZS, Li N, Ren JS, Tian JH, Expert Group of China Guideline for the Screening, Early Detection and Early Treatment of Esophageal Cancer, Work Group of China Guideline for the Screening, Early Detection and Early Treatment of Esophageal Cancer. [China guideline for the screening, early detection and early treatment of esophageal cancer (2022, Beijing)]. Zhonghua Zhong Liu Za Zhi 2022 Jun 23;44(6):491-522. [doi: 10.3760/cma.j.cn112152-20220517-00348] [Medline: 35754225]
20. Bai Y, Yang F, Liu C, Li DF, Wang S, Lin R, National Clinical Research Center for Digestive Diseases (Shanghai), Chinese Society of Digestive Endoscopy, Digestive Endoscopy Professional Committee of Chinese Endoscopist Association, Cancer Endoscopy Professional Committee of China Anti-Cancer Association. Expert consensus on the clinical application of high-frequency electrosurgery in digestive endoscopy (2020, Shanghai). J Dig Dis 2022 Jan;23(1):2-12. [doi: 10.1111/1751-2980.13074] [Medline: 34953023]
21. WHO CVD Risk Chart Working Group. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health 2019 Oct;7(10):e1332-e1345 [FREE Full text] [doi: 10.1016/S2214-109X(19)30318-3] [Medline: 31488387]
22. Muller DC, Johansson M, Brennan P. Lung Cancer Risk Prediction Model Incorporating Lung Function: Development and Validation in the UK Biobank Prospective Cohort Study. J Clin Oncol 2017 Mar 10;35(8):861-869 [FREE Full text] [doi: 10.1200/JCO.2016.69.2467] [Medline: 28095156]
23. Millwood IY, Li L, Smith M, Guo Y, Yang L, Bian Z, China Kadoorie Biobank collaborative group. Alcohol consumption in 0.5 million people from 10 diverse regions of China: prevalence, patterns and socio-demographic and health-related correlates. Int J Epidemiol 2013 Jun;42(3):816-827 [FREE Full text] [doi: 10.1093/ije/dyt078] [Medline: 23918852]
24. Du H, Bennett D, Li L, Whitlock G, Guo Y, Collins R, China Kadoorie Biobank Collaborative Group. Physical activity and sedentary leisure time and their associations with BMI, waist circumference, and percentage body fat in 0.5 million adults: the China Kadoorie Biobank study. Am J Clin Nutr 2013 Mar;97(3):487-496 [FREE Full text] [doi: 10.3945/ajcn.112.046854] [Medline: 23364014]
25. Qin C, Guo Y, Pei P, Du H, Yang L, Chen Y, et al. The Relative Validity and Reproducibility of Food Frequency Questionnaires in the China Kadoorie Biobank Study. Nutrients 2022 Feb 14;14(4):794 [FREE Full text] [doi: 10.3390/nu14040794] [Medline: 35215443]

26. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015 Jan 06;162(1):W1-73 [FREE Full text] [doi: 10.7326/M14-0698] [Medline: 25560730]
27. Lambert PC, Royston P. Further Development of Flexible Parametric Models for Survival Analysis. The Stata Journal 2009 Aug 01;9(2):265-290. [doi: 10.1177/1536867x0900900206]
28. Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008 Jan 30;27(2):157-72; discussion 207. [doi: 10.1002/sim.2929] [Medline: 17569110]
29. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011 Jan 15;30(1):11-21 [FREE Full text] [doi: 10.1002/sim.4085] [Medline: 21204120]
30. Lin Y, Totsuka Y, Shan B, Wang C, Wei W, Qiao Y, et al. Esophageal cancer in high-risk areas of China: research progress and challenges. Ann Epidemiol 2017 Mar;27(3):215-221. [doi: 10.1016/j.annepidem.2016.11.004] [Medline: 28007352]
31. Gail MH, Pfeiffer RM. Breast Cancer Risk Model Requirements for Counseling, Prevention, and Screening. J Natl Cancer Inst 2018 Sep 01;110(9):994-1002 [FREE Full text] [doi: 10.1093/jnci/djy013] [Medline: 29490057]
32. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. Gut 2015 Mar;64(3):381-387. [doi: 10.1136/gutjnl-2014-308124] [Medline: 25320104]
33. China Kadoorie Biobank. URL: https://www.ckbiobank.org/ [accessed 2023-02-28]

Abbreviations

AUC: area under the receiver operating characteristic curve
CKB: China Kadoorie Biobank
EC: esophageal cancer
EHR: electronic health record
ESCC: esophageal squamous cell carcinoma

Edited by Y Khader; submitted 21.10.22; peer-reviewed by S He, P Han; comments to author 20.12.22; revised version received 09.01.23; accepted 03.02.23; published 15.03.23.

Please cite as:
Han Y, Zhu X, Hu Y, Yu C, Guo Y, Hang D, Pang Y, Pei P, Ma H, Sun D, Yang L, Chen Y, Du H, Yu M, Chen J, Chen Z, Huo D, Jin G, Lv J, Hu Z, Shen H, Li L
Electronic Health Record–Based Absolute Risk Prediction Model for Esophageal Cancer in the Chinese Population: Model Development and External Validation
JMIR Public Health Surveill 2023;9:e43725
URL: https://publichealth.jmir.org/2023/1/e43725
doi:10.2196/43725
PMID:36781293

Original Paper

Investigation of the Trajectory of Muscle and Body Mass as a Prognostic Factor in Patients With Colorectal Cancer: Longitudinal Cohort Study

Dongjin Seo[1*], MD; Han Sang Kim[2,3,4*], MD, PhD; Joong Bae Ahn[2], MD, PhD; Yu Rang Park[3,4,5], PhD

[1]Department of Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

[2]Yonsei Cancer Center, Division of Medical Oncology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

[3]Brain Korea 21 FOUR Project, Graduate School of Medical Science, Yonsei University College of Medicine, Seoul, Republic of Korea

[4]Institute for Innovation in Digital Healthcare, Severance Hospital, Seoul, Republic of Korea

[5]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics
Yonsei University College of Medicine
50-1 Yonsei-ro
Seodaemungu
Seoul, 03722
Republic of Korea
Phone: 82 10 5240 3434
Email: yurangpark@yuhs.ac

Abstract

Background: Skeletal muscle and BMI are essential prognostic factors for survival in colorectal cancer (CRC). However, there is a lack of understanding due to scarce studies on the continuous aspects of these variables.

Objective: This study aimed to evaluate the prognostic impact of the initial status and trajectories of muscle and BMI on overall survival (OS) and assess whether these 4 profiles within 1 year can represent the profiles 6 years later.

Methods: We analyzed 4056 newly diagnosed patients with CRC between 2010 to 2020. The volume of the muscle with 5-mm thickness at the third lumbar spine level was measured using a pretrained deep learning algorithm. The skeletal muscle volume index (SMVI) was defined as the muscle volume divided by the square of the height. The correlation between BMI status at the first, third, and sixth years of diagnosis was analyzed and assessed similarly for muscle profiles. Prognostic significances of baseline BMI and SMVI and their 1-year trajectories for OS were evaluated by restricted cubic spline analysis and survival analysis. Patients were categorized based on these 4 dimensions, and prognostic risks were predicted and demonstrated using heat maps.

Results: Trajectories of SMVI were categorized as decreased (812/4056, 20%), steady (2014/4056, 49.7%), or increased (1230/4056, 30.3%). Similarly, BMI trajectories were categorized as decreased (792/4056, 19.5%), steady (2253/4056, 55.5%), or increased (1011/4056, 24.9%). BMI and SMVI values in the first year after diagnosis showed a statistically significant correlation with those in the third and sixth years (P<.001). Restricted cubic spline analysis showed a nonlinear relationship between baseline BMI and SMVI change ratio and OS; BMI, in particular, showed a U-shaped correlation. According to survival analysis, increased BMI (hazard ratio [HR] 0.83; P=.02), high baseline SMVI (HR 0.82; P=.04), and obesity stage 1 (HR 0.80; P=.02) showed a favorable impact, whereas decreased SMVI trajectory (HR 1.31; P=.001), decreased BMI (HR 1.23; P=.02), and initial underweight (HR 1.38; P=.02) or obesity stages 2-3 (HR 1.79; P=.01) were negative prognostic factors for OS. Considered simultaneously, BMI >30 kg/m^2 with a low SMVI at the time of diagnosis resulted in the highest mortality risk. We observed improved survival in patients with increased muscle mass without BMI loss compared to those with steady muscle mass and BMI.

Conclusions: Profiles within 1 year of both BMI and muscle were surrogate indicators for predicting the later profiles. Continuous trajectories of body and muscle mass are independent prognostic factors of patients with CRC. An automatic algorithm provides a unique opportunity to conduct longitudinal evaluations of body compositions. Further studies to understand the complicated natural courses of muscularity and adiposity are necessary for clinical application.

(JMIR Public Health Surveill 2023;9:e43409) doi:10.2196/43409

KEYWORDS

body mass index; BMI; colorectal cancer; deep neural network model; skeletal muscle; skeletal muscle volume index; SMVI

Introduction

Colorectal cancer (CRC) is the third most common cancer, accounting for approximately 10% of newly diagnosed cancers and cancer-related mortality worldwide [1]. Epidemiologic studies indicate that increasing age and resource-rich countries are associated with CRC development [2]. Modifiable lifestyle factors such as smoking, alcohol intake, physical activity, and obesity possibly impact CRC development [3,4].

Given the rapid global increase in obesity in recent years, understanding the effects of obesity on cancer outcomes is important [5,6]. Previous studies suggest that increased BMI is associated with poor prognosis and resistance to anticancer treatment [7-9]. Patients with stage II or III disease with high BMI at the time of diagnosis had decreased overall survival (OS) and disease-free survival [10]. Moreover, studies have described a relationship between muscle and CRC prognosis, suggesting that low muscle mass (ie, sarcopenia) is a prognostic factor for poor survival, increased postoperative complications, and decreased treatment response [11,12]. Despite the well-established prognostic impact of baseline profiles of obesity and muscle mass, personalized risk assessments or intervention decisions regarding these body compositions are limited in clinical practice.

Previously, skeletal muscle was manually measured using computed tomography (CT) images at the level of the third lumbar spine vertebra (L3), which includes an intense localization procedure and inevitable operator errors in the manual identification of the L3 vertebrae [13,14]. Recently, the medical application of artificial intelligence (AI) using machine and deep learning algorithms has increased, enabling the resolution of medical problems with ease of use, robustness, and precision [15]. Moreover, studies have demonstrated that automated CT imaging evaluations using deep learning models are feasible for extracting complicated body composition parameters, and the prognostic effects of these measurements on mortality in CRC have been confirmed [16-18]. Although previous automated body profile evaluations have partially fulfilled the gap for clinical application, the evaluation of adiposity and muscularity is limited to one time point, and thus, not considered alternatively during various treatment or interventions.

To comprehend the natural course of adiposity and muscularity of patients with CRC, we continuously tracked skeletal muscle and BMI profiles over 1 year and evaluated their representativeness for longer prospective profiles. Moreover, we evaluated the prognostic effects of the trajectories of changes in skeletal muscle and BMI to identify their longitudinal effects on CRC prognosis and demonstrated predicted mortality risks for each body composition trajectory, specified by initial status and changes in BMI and muscularity. Our study highlights the continuous characteristics of muscle and BMI in OS and supports the clinical applicability of artificial intelligence–powered automated evaluation for CRC risk modification and management.

Methods

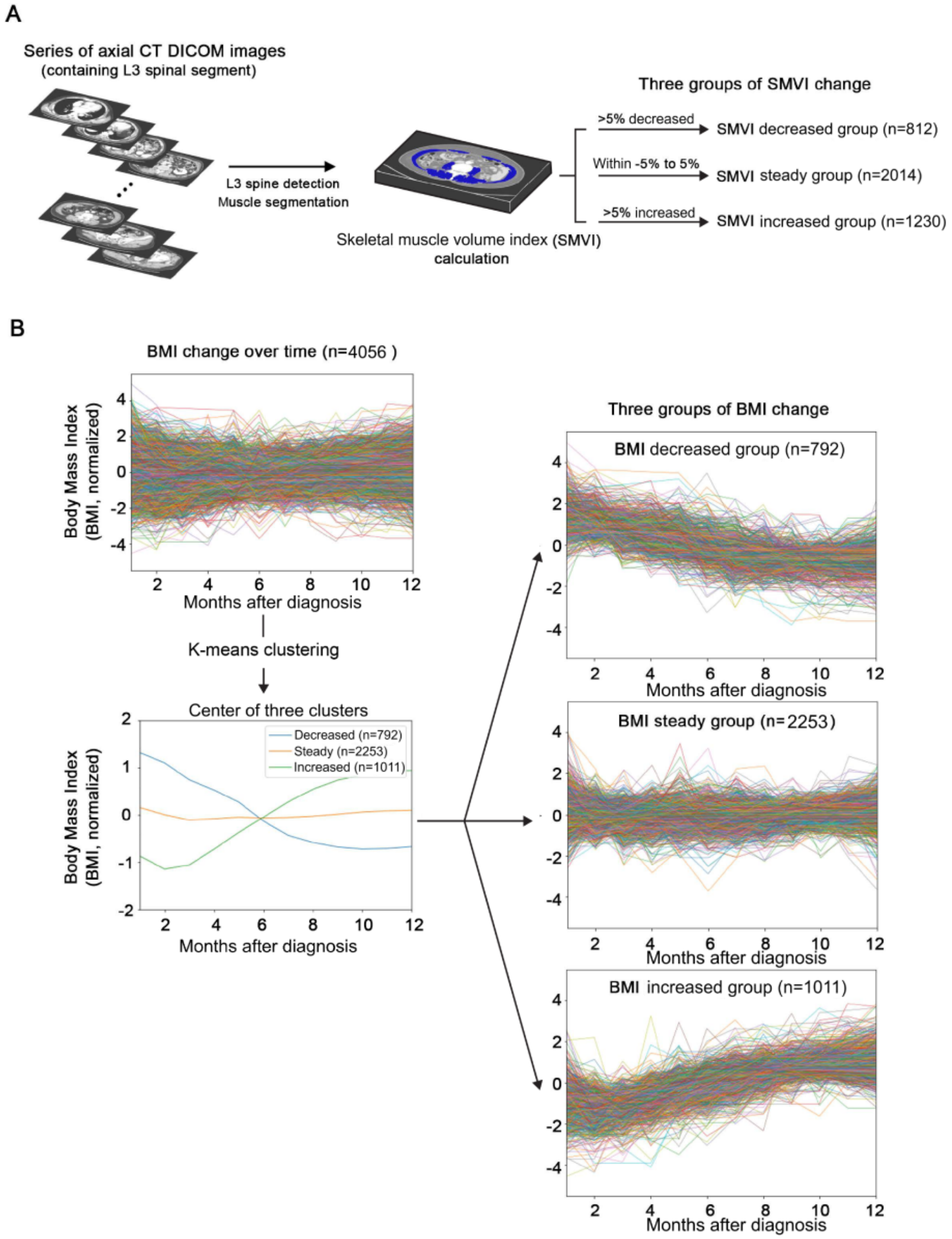
Study Population and Data Collection

A total of 4056 patients with newly diagnosed CRC were enrolled in the Yonsei Cancer Registry Database between January 1, 2010, and September 30, 2020. Patients with abdominal CT images and BMI information up to 1 year after diagnosis were eligible. CT images within 1 year of diagnosis were obtained from the medical database of Severance Hospital. BMI was measured as the patient's weight in kilograms divided by the square of the patient's height in meters (kg/m^2). Skeletal muscle volume index (SMVI) outliers and SMVI change ratios were detected using median absolute deviation and were excluded. Demographic factors, such as age, sex, weight, height, variables related to diagnosis, progression, the treatment of CRC, and the date of death or follow-up loss, were collected.

Automated CT-Derived Skeletal Muscle Mass Measurement

Using the UNet architecture–based automated skeletal muscle measurement algorithm proposed by Islam et al [19], skeletal muscle was analyzed with a series of axial CT images. The overall process of muscle assessment is summarized in Figure 1A. Automated CT-derived skeletal muscle mass measurement involved three steps: (1) axial CT images within 2.5 mm superior and inferior to L3 level were detected, (2) areas of the muscle were calculated, and (3) L3-level muscle volume was calculated. SMVI was defined as L3-level muscle volume divided by the square of the patient's height in meters (cm^3/m^2).

Figure 1. Study design. (A) The process of skeletal muscle segmentation and patient classification according to SMVI patterns. Axial CT image series were inputted (left). The blue area indicates the segmented muscle area and the volume of the muscle with 5-mm thickness was calculated (middle). Patients were classified into SMVI decreased, steady, and increased groups with the SMVI change ratio (right). (B) The process of patient classification according to 1-year BMI patterns. Lines indicate the 1-year BMI pattern of individual patients. The centers of 3 clusters by k-means clustering are presented. The 1-year BMI trajectories of the patients are demonstrated within each BMI pattern. CT: computed tomography; DICOM: Digital Imaging and Communications in Medicine; L3: third lumbar spine vertebra; SMVI: skeletal muscle volume index.



Patient Classification According to SMVI and BMI Patterns

Baseline SMVI was defined as the mean SMVI within 3 months after diagnosis. SMVI change ratios were calculated as the difference between the baseline and last SMVI values. SMVI patterns were divided into 3 groups according to the SMVI change ratio: below −5% as the decreased group, over 5% as the increased group, and between −5% to 5% as the steady group (Figure 1A). Patients' 1-year BMI patterns were defined as the trajectory of the 12 monthly mean BMI values within 1 year after diagnosis. Missing monthly mean BMI was imputed assuming a linear change of monthly BMI. Using the k-means clustering method, patients were divided into 3 BMI pattern groups: decreased, steady, and increased (Figure 1B).

Statistical Analysis

Patient characteristics were compared between BMI pattern groups. The normality of variables was determined using the quantile-quantile plot. Parametric and nonparametric continuous variables were compared using ANOVA and Kruskal-Wallis tests, respectively. Categorical variables were compared using the chi-square test. The effect sizes for ANOVA, Kruskal-Wallis test, and chi-square test were individually calculated using partial eta squared [20], eta squared [21], and Cramer V [22], respectively. The baseline BMI group was classified according to Asian cutoffs determined by the World Health Organization (WHO; BMI <18.5 kg/m^2, underweight; 18.5-23 kg/m^2, normal range; 23-25 kg/m^2, preobese; 25-30 kg/m^2, obese stage 1; and >30 kg/m^2, obese stages 2-3) [23]. The SMVI was categorized into tertiles according to SMVI values (SMVI <18.68 cm^3/m^2 [lower third], low SMVI; 18.68-22.33 cm^3/m^2 [middle third], normal SMVI; and >22.33 cm^3/m^2 [upper third], high SMVI).

The correlation between the BMI of the first, third, and sixth years after diagnosis was evaluated with Pearson correlation, among patients with available BMI data for a sufficient period. Identical correlation analysis was conducted with SMVI among 300 random patients with available CT images for a sufficient period. The association of 6-year overall mortality with baseline and change ratio of both SMVI and BMI was analyzed using restricted cubic spline analysis. The nonlinearity was assessed by Wald statistics. Patients' 6-year OS was analyzed for survival analysis, and survival curves were plotted using the Kaplan-Meier method and compared between SMVI pattern groups using the log-rank test. Cox proportional hazard regression was conducted to predict 6-year OS. Similar survival analyses were performed for each BMI pattern group. All statistical analyses were conducted using JupyterLab (version 1.2.6; Project Jupyter) [24], Python (version 3.6.8; Python Software Foundation) [25], and R (version 4.1.3; R Foundation for Statistical Computing) [26]. Two-sided *P* values of <.05 were considered statistically significant.

Risk Assessment and Heat Map Generation

Three hazard ratio (HR) heat maps were designed with HRs of the survival analysis, with the *seaborn* package (v.0.11.2) [27]. One heat map demonstrated the HRs of 15 baseline conditions, defined by 5 baseline BMI and 3 baseline SMVI groups. The other heat map demonstrated the HRs of 9 body composition change patterns, defined by 3 BMI patterns and 3 SMVI pattern groups. Another heat map demonstrated the HRs of each baseline condition specified by the body composition change groups.

Ethics Approval

The study design was approved by the institutional review board of Severance Hospital, Seoul, South Korea (IRB 4-2020-1304). The need for informed consent was waived by the ethics committee, as this study used routinely collected log data managed anonymously at all stages, including data cleaning and statistical analyses. The study protocol was performed following the guidelines of the International Conference on Harmonization of Good Clinical Practice, the Declaration of Helsinki, and relevant legislation for observational studies.

Results

Study Population

The median follow-up time was 45.5 (range 12.0-129.5) months. The median age was 61.0 (IQR 52.0-69.0) years, and 56.7% (2299/4056) of patients were men. At the time of diagnosis, the average baseline BMI was 23.2 (SD 3.0) kg/m^2. Of the 4056 patients, 4.7% (n=192) were classified as underweight, 43.9% (n=1780) as normal weight, 25.1% (n=1019) as preobese, 23.9% (n=969) as obese stage 1, and 2.4% (n=96) as obese stages 2-3. The mean SMVI at diagnosis was 20.7 (SD 4.1) cm^3/m^2.

Automated CT-Derived Skeletal Muscle Mass Measurement and Patient Classification Based on Longitudinal BMI and SMVI Changes

Patients were classified by SMVI patterns into 3 groups: decreased (812/4056, 20%), steady (2014/4056, 49.7%), and increased (1230/4056, 30.3%) groups. Using the k-means clustering method, clustering BMI patterns into 3 groups—decreased (792/4056, 19.5%), steady (2253/4056, 55.5%), and increased (1011/4056, 24.9%) groups—was the most optimized clustering option (Figure 1B and Multimedia Appendix 1) [28].

The baseline characteristics of each group according to BMI patterns are summarized in Table 1. The median age at diagnosis was 61.0 (IQR 52.0-69.0), 61.0 (IQR 53.0-69.0), and 60.0 (IQR 52.0-67.5) years in the decreased, steady, and increased BMI groups, respectively (Table 1). Baseline BMI was 24.4 (SD 3.1), 23.2 (SD 3.0), and 22.5 (SD 2.9) kg/m^2 in the decreased, steady, and increased BMI groups, respectively. Baseline SMVI was 21.4 (SD 4.1), 20.8 (SD 4.1), and 20.0 (SD 3.8) cm^3/m^2 in the decreased, steady, and increased BMI groups, respectively.

Table 1. Patient characteristics according to 1-year BMI trajectory group: decreased, steady, and increased.

	1-year BMI trajectory pattern group			*P* value[a]	Effect size[b]
	Decreased (n=792)	Steady (n=2253)	Increased (n=1011)		
Age (years), median (IQR)	61.0 (52.0-69.0)	61.0 (53.0-69.0)	60.0 (52.0-67.5)	.01	.002
Sex, n (%)				.02	.044
Male	474 (59.8)	1285 (57)	540 (53.4)		
Female	318 (40.2)	968 (43)	471 (46.6)		
Baseline, mean (SD)					
BMI (kg/m^2)	24.4 (3.1)	23.2 (3.0)	22.5 (2.9)	<.001	.043
SMVI[c] (cm^3/m^2)	21.4 (4.1)	20.8 (4.1)	20 (3.8)	<.001	.013
BMI group, n (%)				<.001	.141
Underweight	12 (1.5)	107 (4.7)	73 (7.2)		
Normal	262 (33.1)	979 (43.5)	539 (53.3)		
Preobese	206 (26)	604 (26.8)	209 (20.7)		
Obesity stage 1	277 (35)	515 (22.9)	177 (17.5)		
Obesity stages 2-3	35 (4.4)	48 (2.1)	13 (1.3)		
1-year SMVI trajectory pattern group, n (%)				<.001	.213
Decreased	311 (39.3)	398 (17.7)	103 (10.2)		
Steady	356 (44.9)	1219 (54.1)	439 (43.4)		
Increased	125 (15.8)	636 (28.2)	469 (46.4)		
Change ratio (%), mean (SD)					
BMI (kg/m^2)	−9.1 (4.1)	0.4 (2.8)	11.0 (5.4)	<.001	.752
SMVI (cm^3/m^2)	−2.8 (9.1)	1.4 (7.8)	4.9 (8.7)	<.001	.088
Cancer stage, n (%)				<.001	.116
I	83 (10.5)	273 (12.1)	26 (2.6)		
II	118 (14.9)	462 (20.5)	172 (17)		
III	320 (40.4)	924 (41)	447 (44.2)		
IV	271 (34.2)	594 (26.4)	366 (36.2)		
Recurrence or metastasis, n (%)	374 (47.2)	851 (37.8)	507 (50.1)	<.001	.113
Death, n (%)	228 (28.8)	538 (23.9)	261 (25.8)	.02	.043
Follow-up duration (years), mean (SD)[d]	3.7 (3.3)	4.0 (4.1)	3.8 (3.6)	<.001	.008
Treatment, n (%)					
Surgery	737 (93.1)	2071 (91.9)	902 (89.2)	.008	.049
CTx[e]	686 (86.6)	1822 (80.9)	961 (95.1)	<.001	.168
RTx[f]	377 (47.6)	613 (27.2)	256 (25.3)	<.001	.181

[a] Categorical variables were compared using the chi-square test. Continuous variables were compared using ANOVA and Kruskal-Wallis tests.

[b] Cramer V effect size: 0.1=small, 0.3=medium, and 0.5=large (for 2 subvariable comparisons) and 0.07=small, 0.21=medium, and 0.35=large (for more than 3 subvariable comparisons). Partial eta-squared effect size: 0.01=small, 0.06=medium, and 0.14=large. Eta-squared effect size: 0.01=small, 0.06=medium, and 0.14=large.

[c] SMVI: skeletal muscle volume index.

[d] The follow-up duration of patients with a follow-up period longer than 6 years was considered as 6 years.

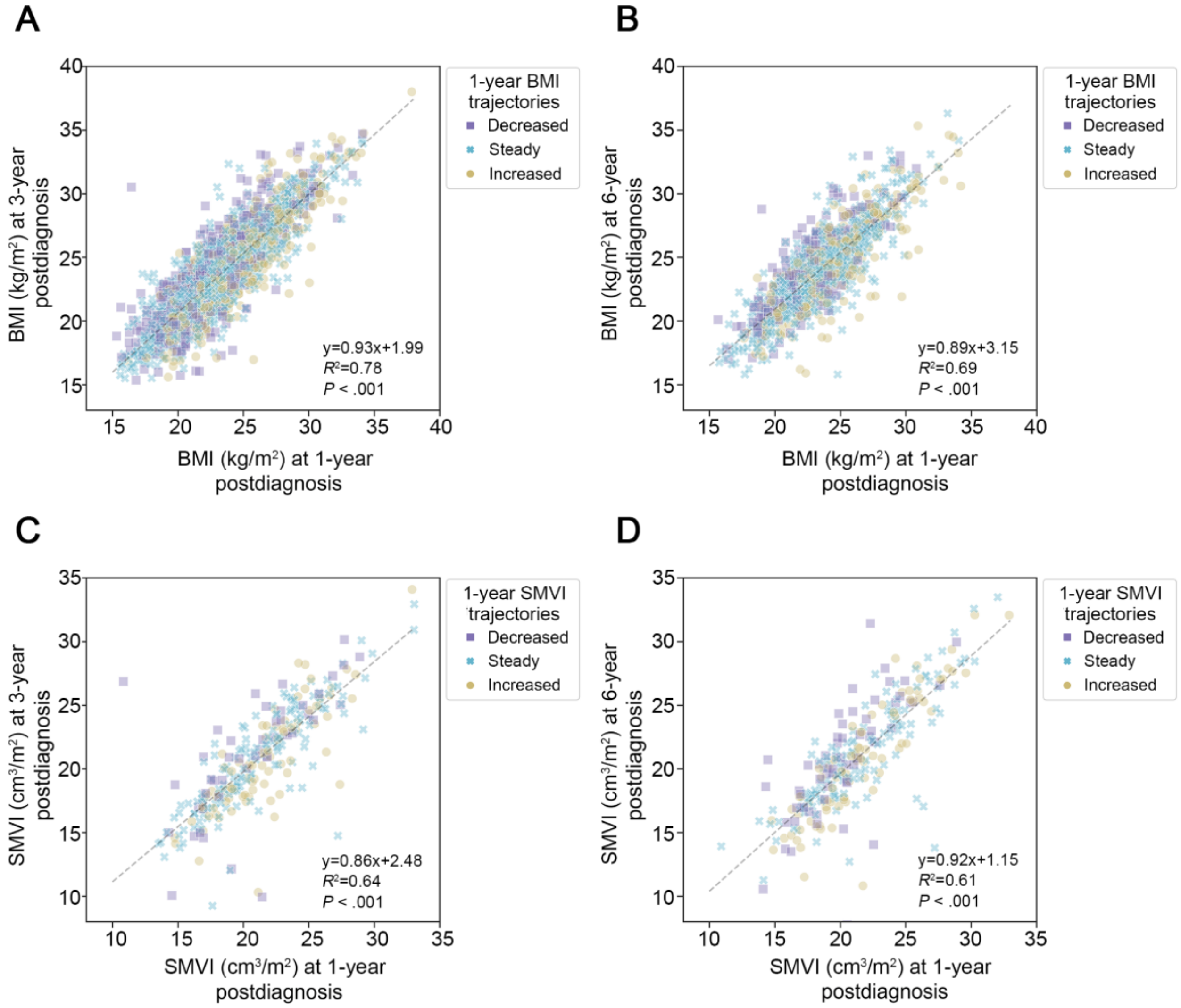
[e] CTx: chemotherapy.

[f] RTx: radiotherapy.

Correlations Between First-Year and Subsequent Profiles of Obesity and Muscle

Among the 4056 patients, BMI data were available for 3217 patients 3 years after diagnosis and for 1318 patients 6 years after diagnosis. Among these patients, BMI in the first and third years showed a statistically significant positive correlation (*P*<.001), whereas the trajectory groups were evenly distributed (Figure 2A). BMI in the first year was also significantly correlated with that in the sixth year with a nonspecific distribution pattern of trajectory groups (*P*<.001; Figure 2B). Among the randomly selected 300 patients for correlation analysis between first- versus third-year SMVI, data from 278 patients were available for CT analysis, and the correlation was found to be positive (*P*<.001; Figure 2C). Among the randomly selected 300 patients for correlation analysis between first- versus sixth-year SMVI, 269 patients were analyzed, and the correlation was positive (*P*<.001; Figure 2D).

Figure 2. Correlation between BMI or SMVI profiles at the first year and 3 and 6 years after diagnosis. Scatter plots of patients' BMI or SMVI profiles are presented, and results of trend line formula, R-square, and *P* value of Pearson correlation are described at the lower right in each graph. Magenta square, cyan cross, and yellow dot indicate decreased, steady, and increased groups, respectively, for the 1-year trajectory of BMI or SMVI. Black dashed lines represent the trend line of each scatter plots. SMVI: skeletal muscle volume index.

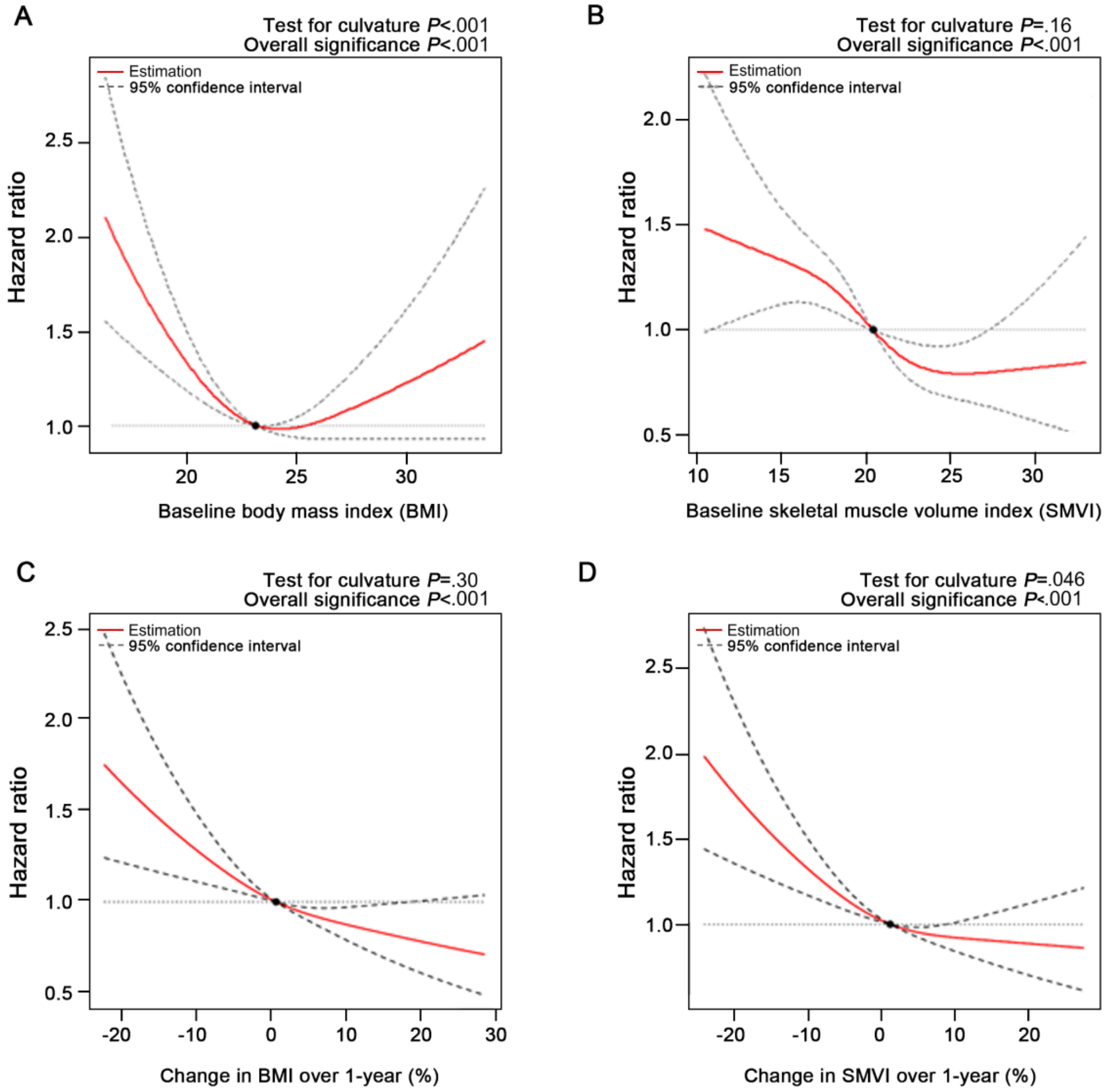


Relationship of the Baseline and Change Ratio of Obesity and Muscle Mass With Mortality

According to restricted cubic spline analysis, baseline BMI and mortality risk exhibited a U-shaped relationship (*P*<.001; Figure 3A). The lowest mortality risk was observed among patients with a normal BMI. Baseline SMVI exhibited an inverse relationship with the risk of death, with an L-shaped pattern, whereby patients with the lowest muscle mass had the highest risk (*P*<.001; Figure 3B).

Change ratios of BMI demonstrated an inverse correlation with overall mortality (*P*<.001; Figure 3C), with the lowest risk observed for an increased BMI of >20%. Patients with BMI loss >20% had a 65% increased risk of death. Similarly, an increase in SMVI over time was associated with improved OS (*P*<.001; Figure 3D). A decrease in SMVI by 20% increased the mortality risk by 78%, suggesting that the restoration of body mass and muscle is critical for improving OS.

Figure 3. The relationship of the baseline and trajectories of obesity and muscle mass with mortality using restricted cubic spline analysis. Relationship between overall mortality and (A) baseline BMI, (B) baseline SMVI, (C) BMI change ratio, and (D) SMVI change ratio. *P* values of nonlinearity and significance of the relationship are described on top of each graph. Red lines represent restricted cubic spline curves, and black dashed lines represent 95% CIs. The reference is the median of each variable. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and patterns of BMI and SMVI (decreased, steady, or increased). Baseline BMI, baseline SMVI, BMI pattern, and SMVI pattern were excluded from an adjustment in the analysis against baseline BMI, baseline SMVI, BMI change ratio, and SMVI change ratio, respectively. SMVI: skeletal muscle volume index.

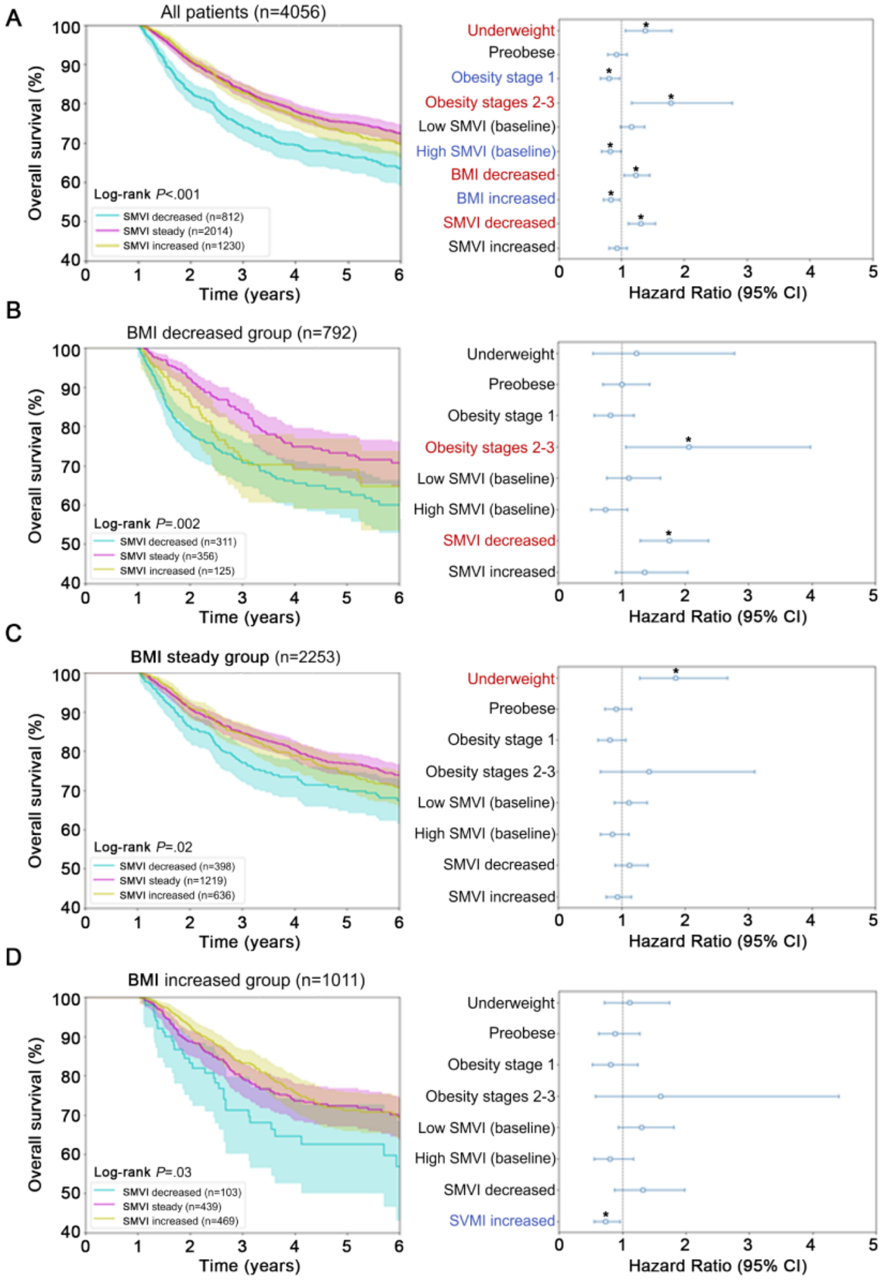


BMI and SMVI as Prognostic Factors for Patient Outcomes in Multivariate Analysis

The trajectory of the decreased SMVI group was associated with shorter OS (6-year OS rate: 63.4% in decreased SMVI vs 72.4% in steady SMVI vs 69.6% in increased SMVI; *P*<.001). Multivariate analysis indicated that obesity stage 1 (HR 0.80, 95% CI 0.66-0.97; *P*=.02), high baseline SMVI (HR 0.82, 95% CI 0.68-0.99; *P*=.04), and increased BMI (HR 0.83, 95% CI 0.71-0.97; *P*=.02) were favorable prognostic factors for OS. Negative prognostic factors included being underweight (HR 1.38, 95% CI 1.06-1.80; *P*=.02), obesity stages 2-3 (HR 1.79, 95% CI 1.16-2.76; *P*=.008), decreased BMI (HR 1.23, 95% CI 1.04-1.45; *P*=.02), and decreased SMVI (HR 1.31, 95% CI 1.11-1.54; *P*=.001), consistent with the restricted cubic spline analysis (Figure 4A and Multimedia Appendix 2).

Based on the trajectories of BMI changes affecting OS, we performed a subgroup analysis according to 3 BMI trajectories (Figure 4B-D). The trajectory of SMVI changes impacted OS in the 3 BMI change groups (log-rank test: decreased BMI, *P*=.002; steady BMI, *P*=.02; and increased BMI, *P*=.03; Figure 4B-D). Of note, in patients with a decrease in BMI, the analysis revealed that obesity stages 2-3 at diagnosis (HR 2.06, 95% CI 1.07-3.97; *P*=.03) and decreased SMVI (HR 1.75, 95% CI 1.30-2.37; *P*<.001) had a negative prognostic impact on OS, suggesting that increased body fat at diagnosis or muscle wasting was associated with unfavorable outcomes in CRC (Figure 4B and Multimedia Appendix 3). In the steady BMI group, being underweight at the time of CRC diagnosis (HR 1.85, 95% CI 1.29-2.67; *P*=.001) had an adverse prognostic impact on OS (Figure 4C and Multimedia Appendix 4). In the increased BMI group, an increase in SMVI was associated with better OS (HR 0.73, 95% CI 0.55-0.97; *P*=.03), suggesting that increased body mass, predominantly due to skeletal muscle mass, positively impacted cancer survival (Figure 4D and Multimedia Appendix 5).

Figure 4. Kaplan-Meier curve and Cox proportional hazard regression analysis. Results for (A) total patients, (B) decreased BMI group, (C) steady BMI group, and (D) increased BMI group are presented. Kaplan-Meier curve results are shown on the left. Cyan, magenta, and yellow lines indicate decreased, steady, and increased SMVI groups, respectively. The hazard ratio and 95% CIs are shown on the right. Red and blue marked variables are statistically significant (*P*<.05) with hazard ratios of >1 and <1, respectively. Variables denoted with stars showed statistically significant *P* values (<.05). Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and patterns of BMI and SMVI (decreased, steady, or increased). SMVI: skeletal muscle volume index.

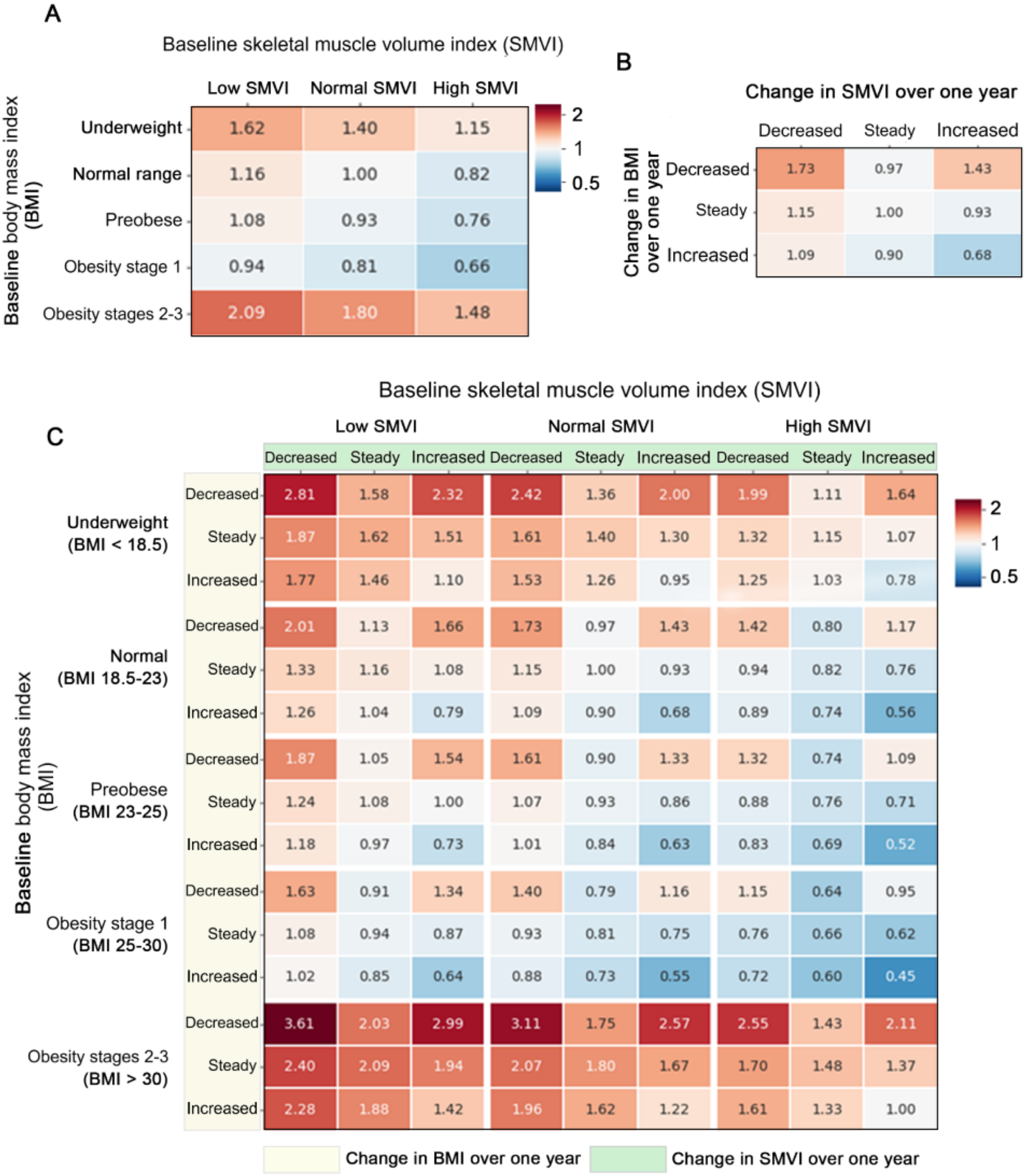


HR Heatmap Representing Associations Between BMI, SMVI, and OS

In consideration of baseline BMI and SMVI at the time of diagnosis, the highest mortality risk was observed in patients with high body mass (BMI >30 kg/m^2) and low SMVI (HR 2.09; Figure 5A), although high muscle mass at the time of CRC diagnosis had a positive impact on OS (HR range 0.66-0.82; BMI 18.5-30 kg/m^2). Correlation analysis between the trajectories of BMI and SMVI revealed that increased body mass and muscle mass were associated with the lowest mortality risk (HR 0.68; *P*=.001; Figure 5B), whereas decreased body mass and muscle wasting were associated with the highest mortality risk (HR 1.73; *P*<.001; Figure 5B and Multimedia Appendices 6 and 7).

Finally, we generated an HR heat map depicting the baseline and trajectories of both BMI and SMVI (Figure 5C). High baseline BMI (>30 kg/m^2) was associated with increased mortality risk regardless of baseline muscle mass or muscle changes (Figure 5C). High baseline SMVI was associated with improved survival in patients with a BMI range from normal to obesity stage 1. There was a trend toward improved OS in patients with increased muscle without BMI loss.

Figure 5. Predicted mortality risk heat map representing associations between BMI, skeletal muscle volume index (SMVI), and overall survival (OS). Patient groups of each heat map were classified by (A) baseline status classified by 5 baseline BMI and 3 baseline SMVI groups, (B) pattern groups classified by 3 BMI and 3 SMVI pattern groups, and (C) all 4 dimensions: baseline BMI and SMVI profiles and trajectories of BMI and SMVI. The predicted mortality risk of each specific patient group is described in each corresponding square, and colors represent the natural logarithm of the predicted mortality risk.



Baseline skeletal muscle volume index (SMVI)

	Low SMVI	Normal SMVI	High SMVI
Underweight	1.62	1.40	1.15
Normal range	1.16	1.00	0.82
Preobese	1.08	0.93	0.76
Obesity stage 1	0.94	0.81	0.66
Obesity stages 2-3	2.09	1.80	1.48

Change in SMVI over one year

Change in BMI over one year	Decreased	Steady	Increased
Decreased	1.73	0.97	1.43
Steady	1.15	1.00	0.93
Increased	1.09	0.90	0.68

Baseline skeletal muscle volume index (SMVI)

		Low SMVI			Normal SMVI			High SMVI		
		Decreased	Steady	Increased	Decreased	Steady	Increased	Decreased	Steady	Increased
Underweight (BMI < 18.5)	Decreased	2.81	1.58	2.32	2.42	1.36	2.00	1.99	1.11	1.64
	Steady	1.87	1.62	1.51	1.61	1.40	1.30	1.32	1.15	1.07
	Increased	1.77	1.46	1.10	1.53	1.26	0.95	1.25	1.03	0.78
Normal (BMI 18.5-23)	Decreased	2.01	1.13	1.66	1.73	0.97	1.43	1.42	0.80	1.17
	Steady	1.33	1.16	1.08	1.15	1.00	0.93	0.94	0.82	0.76
	Increased	1.26	1.04	0.79	1.09	0.90	0.68	0.89	0.74	0.56
Preobese (BMI 23-25)	Decreased	1.87	1.05	1.54	1.61	0.90	1.33	1.32	0.74	1.09
	Steady	1.24	1.08	1.00	1.07	0.93	0.86	0.88	0.76	0.71
	Increased	1.18	0.97	0.73	1.01	0.84	0.63	0.83	0.69	0.52
Obesity stage 1 (BMI 25-30)	Decreased	1.63	0.91	1.34	1.40	0.79	1.16	1.15	0.64	0.95
	Steady	1.08	0.94	0.87	0.93	0.81	0.75	0.76	0.66	0.62
	Increased	1.02	0.85	0.64	0.88	0.73	0.55	0.72	0.60	0.45
Obesity stages 2-3 (BMI > 30)	Decreased	3.61	2.03	2.99	3.11	1.75	2.57	2.55	1.43	2.11
	Steady	2.40	2.09	1.94	2.07	1.80	1.67	1.70	1.48	1.37
	Increased	2.28	1.88	1.42	1.96	1.62	1.22	1.61	1.33	1.00

Discussion

Principal Findings

This study evaluated the characteristics of baseline BMI and SMVI and their 1-year trajectories along with the prognostic impacts of these 4 parameters on OS in patients with CRC. The BMI and SMVI profiles of the first year after diagnosis were highly correlated with the third- and sixth-year profiles, implying that these are surrogate indicators for representing subsequent body composition profiles. Moreover, our survival analysis results indicated that high muscularity positively impacted OS, whereas both depletion and excess adiposity adversely impacted survival. Alterations in adiposity and muscle mass significantly affected 6-year overall mortality. Reciprocal compensation of these 2 factors indicated that changes in BMI had a superior prognostic impact than changes in muscle.

Comparison With Prior Work

According to our restricted cubic spline analysis, baseline SMVI did not exhibit a nonlinear relationship with mortality, whereas baseline BMI demonstrated a U-shaped relationship with mortality. In addition, high muscle mass at baseline exhibited a protective effect against survival, and extreme BMI (underweight or severe obesity) predicted lower survival. These findings suggest that muscularity at the time of diagnosis may have protective effects, whereas both adiposity depletion and excess may have disadvantages for survival. The results of the survival analysis of our study also support the independent prognostic effects of body composition profiles at the time of diagnosis. The mechanisms underscoring the impact of body composition on CRC mortality at the time of diagnosis have yet to be elucidated because completely eliminating methodological biases is challenging [8,9]. A widely accepted explanation is that an adequate amount of adiposity acts as a metabolic reservoir and allows patients to tolerate cachectic situations during cancer progression and treatment [29]. However, excess body fat is correlated with higher mortality via various pathways. Dysregulated adipose tissue can increase inflammatory adipokines, leading to systemic inflammation and a tumor-friendly microenvironment [30], thereby exacerbating catabolic pathways [31]. Patients with low muscularity and high adiposity were associated with higher inflammatory-related serum proteins and cytokines [17]. In addition, the specification of adipose tissue into subcutaneous, visceral, and intra- or intermuscular adipose tissue revealed distinct relationships with CRC mortality: subcutaneous fat exhibits a U-shaped association, whereas visceral and intermuscular fats exhibit a positive relationship, implying context-dependent roles of fat [9].

In this study, an increase in muscle mass or BMI demonstrated higher OS, highlighting the beneficial effects of increasing muscle and body mass. Survival analysis within subgroups of each BMI trajectory group showed a combined prognostic influence of muscularity and adiposity. HR results of the survival analysis conducted within steady and increased BMI groups described that survival improved in the order of increased, steady, and decreased SMVI. However, within the decreased BMI group, an increase in SMVI was associated with poor survival compared with steady SMVI, which contradicts general findings. Although several findings were statistically insignificant, this implies that the protective role of increased muscle differs regarding the characteristics of BMI changes, since the prognostic effects of changes in muscle and adiposity may reciprocally compensate. In addition, according to the survival analysis considering all 9 patterns of BMI and SMVI change, patients in the increased SMVI with decreased BMI group had statistically significant increased mortality risk. The superior decrease of adiposity that leads to decreased BMI despite the improvement of muscularity may override the protective effects of increased muscle mass.

Although BMI is known to have limitations in the direct representation of adiposity [9], BMI profiles within 1 year could be an efficient parameter for predicting the prognosis of patients with CRC in most primary care settings. First, continuously repeated measurements of BMI can be performed, which enables longitudinal evaluation with ease. Second, a 1-year BMI profile highly correlates with future BMI profiles. Third, the trajectory of BMI is a critical independent prognostic factor, where it is a parameter containing longitudinal information of multiple compositions. Adiposity-related [17] and muscle-related indexes such as SMVI, muscle-to-bone ratio, or skeletal muscle radiodensity [16,18] are also essential prognostic factors, yet frequent assessments are limited since these parameters are products of occasional CT imaging studies that depend on cancer treatment progress, patient's medical condition, health insurance, etc. Thus, synergizing CT-driven information and BMI data will promote the future clinical application.

Previous studies support the association between physical activity and cancer prognosis. A prospective cohort study revealed that sedentary behavior is associated with cancer mortality risk and replacing sedentary time with physical activities may improve OS [32]. According to a meta-analysis of nonrandomized trials, higher-intensity physical activity was significantly associated with lower mortality risk [33]. In addition, resistance training in patients with cancer resulted in improvement in muscle strength, body composition, and benefits for cancer survival [34,35]. Exercise is safe and feasible for patients with CRC and results in improvements in various health-related outcomes [36]. Collectively, appropriate interventions through physical activities could improve the outcomes of patients with CRC by improving body composition profiles. Studies evaluating the role of an exercise intervention to improve OS are rare; however, a phase-2 randomized clinical trial proved that high-intensity interval training delayed prostate cancer progression [37]. Thus, further studies evaluating whether exercise intervention improves long-term outcomes should be conducted.

Strengths

Our study has the following strengths. We analyzed a large size of patients and assessed various natural courses of patients with CRC who did not receive additional interventions besides clinical practices. To comprehend the role of adiposity and muscularity in CRC progression, understanding its alternating nature is necessary for evaluating time-dependent prognostic impacts. Therefore, longitudinal patterns of BMI and muscle were analyzed simultaneously. Given that adiposity and muscle commonly change during cancer treatments [9], we longitudinally analyzed BMI and muscle for up to 1 year, and our study shows that the 1-year profile of 2 variables can aptly represent the course over 6 years. As these are proven to be independent prognostic factors, analyzing the 1-year course of adiposity and muscularity will be an efficient and effective method to determine the mortality risk. As a part of continuous body composition analysis, the deep learning–based assessment was implemented to minimize manual processes and enable rapid and accurate analysis of a large number of CT images. The application of automated analysis processes in clinical practice will highly promote the improvement of personalized care and risk management.

Limitations

Our study has several limitations. We evaluated the 6-year OS as a survival outcome using 1-year trajectories of BMI and muscle mass. However, according to the correlation analysis, trajectory analysis up to 1 year was sufficient for representing the courses beyond a year. Since physical examinations and imaging studies are frequent during the first year as a part of the diagnosis or treatment process, we assumed that 1 year was sufficient for classifying body composition profiles. Prolonged longitudinal trajectorial analysis may provide additional insight into the longer-term effects of adiposity or muscle in cancer survival, thereby facilitating personalized risk assessment and management.

Conclusions

In conclusion, our findings demonstrated the natural courses of BMI and muscle and highlight the 1-year trajectory as a surrogate indicator for representing further progress, wherein deep learning–based automated muscle assessment enabled longitudinal analysis of muscle. Thus, this study can be ground evidence for further research regarding prognostic improvements of patients with cancer by interventions that improve body composition profiles. In addition, our results provided crucial insights into the prognostic roles of adiposity and muscle mass in CRC survival. Body composition profiles at the time of diagnosis and alterations of these parameters are independent prognostic factors of CRC survival. The combination of BMI, which is frequently gathered but crude, with muscle-related parameter, which is automatically analyzed but occasionally obtained, could be a precise risk assessment tool. For further clinical application, supporting studies evaluating the natural course of patients with CRC will be necessary and more personalized prognosis prediction methods are required. To fully understand the role of adiposity and muscularity in cancer progression, impacts on other clinical outcomes should also be studied.

Acknowledgments

This research was supported by the Bio-Industrial Technology Development Program (20014841) funded by the Ministry of Trade, Industry & Energy (MOTIE; South Korea) and a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (Ministry of Science and ICT [MSIT]; grant 2022R1A2C4001879). We also thank the Medical Illustration & Design, part of the Medical Research Support Services of Yonsei University College of Medicine, South Korea, for revising illustrations for this work. Additionally, this study was supported by the Division of Digital Health of the Yonsei University Health System, Seoul, South Korea, with regard to data curation.

Data Availability

Raw data for this study were generated at Severance Hospital. Derived data supporting the findings of this study are available from the corresponding author upon request.

Authors' Contributions

DS and HSK contributed equally as co–first authors. DS, HSK, and YRP contributed to conceptualization. DS, HSK, JBA, and YRP contributed to data curation. DS, HSK, and YRP contributed to methodology. DS contributed to formal analysis, software, and writing—original draft. YRP and HSK contributed to writing—review and editing. YRP, HSK, and JBA contributed to supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The inertia result of the k-means clustering of patients’ 1-year trajectory according to cluster numbers from 1 to 10. The most efficient clustering result was 3 clusters using the elbow method, and its inertia is denoted by the red circle.

[PNG File , 56 KB - publichealth_v9i1e43409_app1.png]

Multimedia Appendix 2

Cox proportional hazard regression result within total patients. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the

administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and patterns of BMI and SMVI (decreased, steady, or increased). SMVI: skeletal muscle volume index.
[DOCX File , 21 KB - publichealth_v9i1e43409_app2.docx]

Multimedia Appendix 3
Cox proportional hazard regression result within the decreased BMI group. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and patterns of SMVI (decreased, steady, or increased). SMVI: skeletal muscle volume index.
[DOCX File , 21 KB - publichealth_v9i1e43409_app3.docx]

Multimedia Appendix 4
Cox proportional hazard regression result within the steady BMI group. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and patterns of SMVI (decreased, steady, or increased). SMVI: skeletal muscle volume index.
[DOCX File , 21 KB - publichealth_v9i1e43409_app4.docx]

Multimedia Appendix 5
Cox proportional hazard regression result within the increased BMI group. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and patterns of SMVI (decreased, steady, or increased). SMVI: skeletal muscle volume index.
[DOCX File , 21 KB - publichealth_v9i1e43409_app5.docx]

Multimedia Appendix 6
Cox proportional hazard regression result for hazard ratio heat map. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and 9 patient groups divided by three 1-year BMI trajectories (decreased, steady, or increased) and three 1-year SMVI trajectories (decreased, steady, or increased). SMVI: skeletal muscle volume index.
[DOCX File , 22 KB - publichealth_v9i1e43409_app6.docx]

Multimedia Appendix 7
Survival analysis results against total patients for hazard ratio heat map. (A) Kaplan-Meier curve and survival curve comparison by log-rank test. Patient groups were classified with 1-year trajectories of BMI and SMVI and are represented in different colors. (B) Cox proportional hazard regression result for hazard ratio heat map. Adjusted variables were age at diagnosis (above or below 65 years); sex; stage; primary cancer location (colon or rectum); histology (adenocarcinoma or others); recurrence or metastasis; the administration of surgery, chemotherapy, or radiotherapy; baseline BMI (underweight, normal, preobese, obesity stage 1, or obesity stages 2-3); baseline SMVI (low, normal, or high); and 9 patient groups divided by three 1-year BMI trajectories (decreased, steady, or increased) and three 1-year SMVI trajectories (decreased, steady, or increased). SMVI: skeletal muscle volume index.
[PNG File , 152 KB - publichealth_v9i1e43409_app7.png]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet 2019 Oct 19;394(10207):1467-1480. [doi: 10.1016/S0140-6736(19)32319-0] [Medline: 31631858]
3. Khil H, Kim SM, Hong S, Gil HM, Cheon E, Lee DH, et al. Time trends of colorectal cancer incidence and associated lifestyle factors in South Korea. Sci Rep 2021 Jan 28;11(1):2413 [FREE Full text] [doi: 10.1038/s41598-021-81877-2] [Medline: 33510236]

4. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. Nat Rev Dis Primers 2015 Nov 05;1:15065 [FREE Full text] [doi: 10.1038/nrdp.2015.65] [Medline: 27189416]
5. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiol Biomarkers Prev 2007 Dec;16(12):2533-2547. [doi: 10.1158/1055-9965.EPI-07-0708] [Medline: 18086756]
6. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. Lancet 2014 Aug 30;384(9945):755-765 [FREE Full text] [doi: 10.1016/S0140-6736(14)60892-8] [Medline: 25129328]
7. Caan BJ, Meyerhardt JA, Kroenke CH, Alexeeff S, Xiao J, Weltzien E, et al. Explaining the obesity paradox: the association between body composition and colorectal cancer survival (C-SCANS study). Cancer Epidemiol Biomarkers Prev 2017 Jul;26(7):1008-1015 [FREE Full text] [doi: 10.1158/1055-9965.EPI-17-0200] [Medline: 28506965]
8. Cespedes Feliciano EM, Kroenke CH, Caan BJ. The obesity paradox in cancer: how important is muscle? Annu Rev Nutr 2018 Aug 21;38:357-379. [doi: 10.1146/annurev-nutr-082117-051723] [Medline: 29727593]
9. Caan BJ, Cespedes Feliciano EM, Kroenke CH. The importance of body composition in explaining the overweight paradox in cancer-counterpoint. Cancer Res 2018 Apr 15;78(8):1906-1912 [FREE Full text] [doi: 10.1158/0008-5472.CAN-17-3287] [Medline: 29654153]
10. Renfro LA, Loupakis F, Adams RA, Seymour MT, Heinemann V, Schmoll H, et al. Body mass index is prognostic in metastatic colorectal cancer: pooled analysis of patients from first-line clinical trials in the ARCAD database. J Clin Oncol 2016 Jan 10;34(2):144-150 [FREE Full text] [doi: 10.1200/JCO.2015.61.6441] [Medline: 26503203]
11. Baracos VE, Arribas L. Sarcopenic obesity: hidden muscle wasting and its impact for survival and complications of cancer therapy. Ann Oncol 2018 Feb 01;29(suppl_2):ii1-ii9 [FREE Full text] [doi: 10.1093/annonc/mdx810] [Medline: 29506228]
12. Abbass T, Dolan R, McSorley ST, Horgan PG, McMillan DC. Skeletal muscle index (SMI) status and survival in patients undergoing surgery for colorectal cancer (CRC): A longitudinal study. J Clin Oncol 2020 Feb 01;38(4_suppl):175-175. [doi: 10.1200/jco.2020.38.4_suppl.175]
13. Derstine BA, Holcombe SA, Ross BE, Wang NC, Su GL, Wang SC. Optimal body size adjustment of L3 CT skeletal muscle area for sarcopenia assessment. Sci Rep 2021 Jan 11;11(1):279 [FREE Full text] [doi: 10.1038/s41598-020-79471-z] [Medline: 33431971]
14. Vangelov B, Bauer J, Kotevski D, Smee RI. The use of alternate vertebral levels to L3 in computed tomography scans for skeletal muscle mass evaluation and sarcopenia assessment in patients with cancer: a systematic review. Br J Nutr 2022 Mar 14;127(5):722-735. [doi: 10.1017/S0007114521001446] [Medline: 33910664]
15. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019 Apr 04;380(14):1347-1358. [doi: 10.1056/NEJMra1814259] [Medline: 30943338]
16. Keyl J, Hosch R, Berger A, Ester O, Greiner T, Bogner S, et al. Deep learning-based assessment of body composition and liver tumour burden for survival modelling in advanced colorectal cancer. J Cachexia Sarcopenia Muscle 2023 Feb;14(1):545-552 [FREE Full text] [doi: 10.1002/jcsm.13158] [Medline: 36544260]
17. Fleming CA, O'Connell EP, Kavanagh RG, O'Leary DP, Twomey M, Corrigan MA, et al. Body composition, inflammation, and 5-year outcomes in colon cancer. JAMA Netw Open 2021 Aug 02;4(8):e2115274 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.15274] [Medline: 34459908]
18. Xiao J, Caan BJ, Cespedes Feliciano EM, Meyerhardt JA, Peng PD, Baracos VE, et al. Association of low muscle mass and low muscle radiodensity with morbidity and mortality for colon cancer surgery. JAMA Surg 2020 Oct 01;155(10):942-949 [FREE Full text] [doi: 10.1001/jamasurg.2020.2497] [Medline: 32805015]
19. Islam S, Kanavati F, Arain Z, Da Costa OF, Crum W, Aboagye EO, et al. Fully automated deep-learning section-based muscle segmentation from CT images for sarcopenia assessment. Clin Radiol 2022 May;77(5):e363-e371. [doi: 10.1016/j.crad.2022.01.036] [Medline: 35260232]
20. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013 Nov 26;4:863 [FREE Full text] [doi: 10.3389/fpsyg.2013.00863] [Medline: 24324449]
21. Tomczak M, Tomczak E. The need to report effect size estimates revisited: an overview of some recommended measures of effect size. Trends Sport Sci 2014;21(1):19-25 [FREE Full text]
22. Kim HY. Statistical notes for clinical researchers: chi-squared test and Fisher's exact test. Restor Dent Endod 2017 May;42(2):152-155 [FREE Full text] [doi: 10.5395/rde.2017.42.2.152] [Medline: 28503482]
23. WHO expert consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
24. Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, et al. Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Positioning and Power in Academic Publishing: Players, Agents and Agendas.: IOS Press; 2016 Presented at: 20th International Conference on Electronic Publishing; June 7-9, 2016; Göttingen, Germany p. 87-90. [doi: 10.3233/978-1-61499-649-1-87]
25. VanRossum G, Drake FL. The Python Language Reference. Amsterdam, Netherlands: Python Software Foundation; 2010.
26. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2013. URL: https://www.r-project.org/ [accessed 2023-03-10]

27. Waskom M, Botvinnik O, Hobson P, Halchenko Y, Lukauskas S, Cole JB, et al. seaborn: v0.7.1 (June 2016). Zenodo. 2016 Jun 5. URL: https://zenodo.org/record/54844#.ZAt5zezMJzQ [accessed 2023-03-10]
28. Marutho D, Handaka SH, Wijaya E, Muljono. The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. 2018 Presented at: 2018 International Seminar on Application for Technology of Information and Communication; September 21-22, 2018; Semarang, Indonesia p. 533-538. [doi: 10.1109/isemantic.2018.8549751]
29. Ngeow J, Eng C. Oh GxE! the complexity of body mass index and colon cancer risk. J Natl Cancer Inst 2021 Jan 04;113(1):5-6 [FREE Full text] [doi: 10.1093/jnci/djaa059] [Medline: 32324885]
30. Campbell PT, Lin Y, Bien SA, Figueiredo JC, Harrison TA, Guinter MA, et al. Association of body mass index with colorectal cancer risk by genome-wide variants. J Natl Cancer Inst 2021 Jan 04;113(1):38-47 [FREE Full text] [doi: 10.1093/jnci/djaa058] [Medline: 32324875]
31. Feliciano EMC, Kroenke CH, Meyerhardt JA, Prado CM, Bradshaw PT, Kwan ML, et al. Association of systemic inflammation and sarcopenia with survival in nonmetastatic colorectal cancer: results from the C SCANS study. JAMA Oncol 2017 Dec 01;3(12):e172319 [FREE Full text] [doi: 10.1001/jamaoncol.2017.2319] [Medline: 28796857]
32. Gilchrist SC, Howard VJ, Akinyemiju T, Judd SE, Cushman M, Hooker SP, et al. Association of sedentary behavior with cancer mortality in middle-aged and older US adults. JAMA Oncol 2020 Aug 01;6(8):1210-1217 [FREE Full text] [doi: 10.1001/jamaoncol.2020.2045] [Medline: 32556069]
33. Takemura N, Chan SL, Smith R, Cheung DST, Lin C. The effects of physical activity on overall survival among advanced cancer patients: a systematic review and meta-analysis. BMC Cancer 2021 Mar 07;21(1):242 [FREE Full text] [doi: 10.1186/s12885-021-07988-1] [Medline: 33678180]
34. Hardee JP, Porter RR, Sui X, Archer E, Lee I, Lavie CJ, et al. The effect of resistance exercise on all-cause mortality in cancer survivors. Mayo Clin Proc 2014 Aug;89(8):1108-1115 [FREE Full text] [doi: 10.1016/j.mayocp.2014.03.018] [Medline: 24958698]
35. Strasser B, Steindorf K, Wiskemann J, Ulrich CM. Impact of resistance training in cancer survivors: a meta-analysis. Med Sci Sports Exerc 2013 Nov;45(11):2080-2090. [doi: 10.1249/MSS.0b013e31829a3b63] [Medline: 23669878]
36. Singh B, Hayes SC, Spence RR, Steele ML, Millet GY, Gergele L. Exercise and colorectal cancer: a systematic review and meta-analysis of exercise safety, feasibility and effectiveness. Int J Behav Nutr Phys Act 2020 Sep 24;17(1):122 [FREE Full text] [doi: 10.1186/s12966-020-01021-7] [Medline: 32972439]
37. Kang DW, Fairey AS, Boulé NG, Field CJ, Wharton SA, Courneya KS. Effects of exercise on cardiorespiratory fitness and biochemical progression in men with localized prostate cancer under active surveillance: the ERASE randomized clinical trial. JAMA Oncol 2021 Oct 01;7(10):1487-1495 [FREE Full text] [doi: 10.1001/jamaoncol.2021.3067] [Medline: 34410322]

Abbreviations

CRC: colorectal cancer
CT: computed tomography
HR: hazard ratio
L3: third lumbar spine vertebra
OS: overall survival
SMVI: skeletal muscle volume index
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 11.10.22; peer-reviewed by M Kapsetaki, X Xu, K Gupta; comments to author 26.01.23; revised version received 16.02.23; accepted 23.02.23; published 22.03.23.

Please cite as:
Seo D, Kim HS, Ahn JB, Park YR
Investigation of the Trajectory of Muscle and Body Mass as a Prognostic Factor in Patients With Colorectal Cancer: Longitudinal Cohort Study
JMIR Public Health Surveill 2023;9:e43409
URL: https://publichealth.jmir.org/2023/1/e43409
doi:10.2196/43409
PMID:36947110

Original Paper

The Effect of Activity Participation in Middle-Aged and Older People on the Trajectory of Depression in Later Life: National Cohort Study

Yan Gao[1], PhD; Zhihao Jia[1], MEd; Liangyu Zhao[1], MEd; Suyue Han[1], MEd

[1] School of Physical Education, Shandong University, Jinan, Shandong, China

Corresponding Author:
Yan Gao, PhD
School of Physical Education
Shandong University
17922 Jingshi Road, Lixia District
Jinan, Shandong, 250061
China
Phone: 86 18663709793
Email: gaoyanluck@sdu.edu.cn

Abstract

Background: More activity participation is an important means of handling depression and promoting positive aging, but the impact of changes in activity participation on the developmental trajectory of depression has not been fully studied.

Objective: The purpose of this study is to study the relationship between current activity participation and depression in middle-aged and older people (≥45 years old) and the relationship between activity participation and the developmental trajectory of depression in later life in China.

Methods: This study used data from the China Health and Retirement Longitudinal Study (CHARLS) across 7 years and included a total of 4818 middle-aged and older people (≥45 years old). Controlling for relevant control variables, the latent growth curve model and the cross-lag model were used to assess the effect of changes in activity participation on the depression trajectory in later life and the main lag relationship between activity participation and depression. Activity participation as well as depression were measured using the self-reported activity and health status based on the CHARLS questionnaire.

Results: Among the 4818 respondents, the mean values of physical activity participation, social leisure activity participation, and depression ranged from 76.98 (SD 15.16) to 83.95 (SD 5.72), from 7.43 (SD 8.67) to 9.25 (SD 10.16), and from 7.61 (SD 5.72) to 8.82 (SD 6.51), respectively. Our findings revealed that activity participation could be related to depression. Physical activity participation predicted initial depression (β=–0.631, *P*<.001) and its trajectory (β=0.461, *P*<.001). However, social leisure activity participation predicted initial depression (β=–0.223, *P*<.001) but did not predict its trajectory (β=0.067, *P*=.159). Finally, cross-lag regression analysis further demonstrated the predictive effect of activity participation on depression.

Conclusions: This study demonstrates the prediction of activity participation for future depression in the Chinese middle-aged and older populations. The data showed that activity participation is significantly associated with changes in depression and future depression among middle-aged and older people in China. The Chinese government should encourage middle-aged and older people to participate in various activities, which can effectively prevent the aggravation of depression and also have a positive significance for positive aging.

(JMIR Public Health Surveill 2023;9:e44682) doi:10.2196/44682

KEYWORDS

activity participation; middle-aged and older people; depression; cohort study; developmental

Introduction

Better activity participation not only has a positive impact on the physical and mental health of middle-aged and older people [1-3] but also improves the health-related quality of life (HRQoL) [4,5]. Specifically, active participation in various activities can provide middle-aged and older people with a series of positive feelings so that they can better cope with the pressure of daily life and maintain good social interaction. This will make them feel happier and more satisfied with their lives, which will further result in improved HRQoL and mental health [3].

At present, activity participation has become a hot topic in depression research. If the elderly cannot adapt to changes brought about by physical function (PF) decline or the change in social roles (eg, changes in activity participation), anxiety and depression will be an inevitable consequence [6]. Depression is a common mental disorder. Worldwide, an estimated 5% of adults have depression. According to statistics, after the COVID-19 pandemic, there were about 374 million cases of anxiety (4802 cases per 100,000 people) and 246 million cases of major depression (3153 cases per 100,000 people), up by 26% and 28%, respectively, compared to the previous years [7]. Therefore, mental health problems have become a big problem worldwide. At the same time, the prevalence of depression in China is also a big issue. It is estimated that the prevalence of depression in the elderly in China is as high as 23.6%, and its prevalence will increase not only with age [8,9] but also with a trend of "rejuvenation" [10]. It is clear that depression in the middle-aged and the older population has become a major public health problem in China. Exploring the developmental trajectory of depression in the middle-aged and older populations and its relationship with related factors is of great significance for actively coping with healthy aging.

Although some studies have shown that activity participation levels in middle-aged and older adults increase or stabilize over time [11,12], in general, the academic community generally believes that it becomes increasingly difficult for middle-aged and older people to maintain active participation as they grow older [13,14]. The concept of life cycle is widely used. In psychology, it mainly refers to the life cycle of people and the life cycle of the family and also refers to the process of birth, growth, aging, illness, and death. Life cycle theory holds that after middle-aged people enter old age, their physiological and psychological processes experience a series of changes [14], which leads to changes in activity participation [6]. In a nutshell, the activity participation of the elderly (eg, leisure activities) gradually becomes simpler and decreases over time [11,14]. Numerous studies have shown that older adults generally have lower levels of social activity participation, and this lower level may further decrease with age [15-17]. Approximately 39.0% of respondents showed low social activity engagement at baseline and decreased engagement within 10 years [15]. Especially among the elderly in China, they spend more time sedentary than the young [18]. In addition, life cycle theory [14] believes that when entering the final stage of the life cycle, both physical and mental functions decline sharply, suggesting that the trends in activity participation and depression should be nonlinear.

Currently, there are a large number of studies focusing on the relationship between activity participation and depression. Activity theory is commonly used to explain the relationship between social participation and mental health in older adults. Activity theory believes that middle-aged and older people can establish new roles to cope with the loss of their original roles so as to adapt to society and obtain psychological satisfaction [19]. In other words, older people can establish new roles through activity participation to alleviate the depressive symptoms caused by role changes [6]. This theory is also supported by the results of some clinical studies, which suggest that activity participation can prevent depression [20] and is becoming an effective treatment for depression [21]. This shows that leisure activities are actually a protective factor for depression in the elderly and that participation in activities may reduce the rate of depression [22]. Additionally, studies have shown that changes in PF and changes in depression are reciprocal [23-25]. Given the direct link between activity participation and PF, activity participation changes may also be associated with changes in depression. Therefore, changes in activity participation in middle-aged and older adults may predict changes in depression. Finally, behavioral activation (BA) theory suggests that engaging in activities that make individuals feel pleasure improves depression and promotes activity participation after depression improves [26]. This suggests that there may be a temporal sequence between activity participation and depression. Therefore, this study further determines the temporal order of the relationship between activity participation and depression.

Previous studies have often included activity participation as a time-invariant dependent variable or a simple control variable and have simply verified the correlation between the initial levels of the 2 variables [22,27]. The effect of changes in activity participation on depression was rarely considered, and this led to a lack of discussion of the relationship between the initial levels of the 2 variables and the rate of change. Therefore, this study selected data from the China Health and Retirement Longitudinal Study (CHARLS) spanning 7 years using a latent growth curve model (LGCM) and a cross-lag model to explore the developmental trajectories of activity participation and depression and the patterns of interaction between their trajectories, further revealing the dynamic development relationship between variables. Based on the existing literature and theories, this study has a few overall hypotheses: (1) middle-aged and elderly activity participation shows a nonlinear downward trend with age, (2) middle-aged and elderly depression shows a nonlinear upward trend with age, (3) the activity participation of middle-aged and older adults predicts their depression during the same period, (4) the initial level and rate of change of activity participation in middle-aged and older adults could predict the development rate of the depression trajectory, and (5) the middle-aged and older adults' early activity participation at 4 points in time predicts later depression.

Methods

Sample and Data Collection

Data were obtained from CHARLS. Our national baseline survey was launched in 2011, covering 150 county-level units, 450 village-level units, and about 17,000 people in 10,000 households, and these samples were tracked every 2-3 years thereafter. More details about CHARLS can be found in previous publications [28]. This database uses data from 4 waves in 2011 (wave I), 2013 (wave II), 2015 (wave III), and 2018 (wave IV), excluding respondents younger than 45 years old.

The detailed sample-screening process is shown in Multimedia Appendix 1. This study first focused on all respondents in wave

I. Since the purpose of the study was to study the relationship between activity participation and depression, plus with the control of relevant time-varying and invariant variables, 3908 respondents of the final wave I were excluded from our study. The specific reasons were as follows: they were younger than 45 years or their age was not recorded (n=367, 9.4%); they had self-reported memory-related disorders and brain damage, intellectual disability, missing data on brain damage, or intellectual disability at baseline (n= 528, 13.5%); they had no baseline data on depressive symptoms (n=2101, 53.7%); they had no activity participation data at baseline (n=1943, 49.7%); and they had missing values for relevant covariates (gender, marriage, education, self-rated health, chronic disease, smoking, drinking; n=1943, 49.7%). Next, this study performed the same first process on wave II, wave III, and wave IV data as wave I, and in the second step, the study excluded respondents who had not participated in all of them for 4 years. Finally, after these operations, individuals with missing values on any variable were excluded, and a total of 4818 valid individuals were included.

Ethical Considerations

This study was organized by the National Development Institute of Peking University, complied with the ethical guidelines of the 1975 Declaration of Helsinki, and was approved by the Ethics Committee of Peking University (approval number: IRB00001052-13074), hosted by the National Development Institute of Peking University. The studies involving human participants were reviewed and approved by the Research Ethics Committees of Peking University (IRB00001052-11015). The patients/participants provided their written informed consent to participate in this study.

Activity Participation

Combining the definitions of activity participation by previous scholars [29] and the characteristics of activity participation of Chinese seniors [30], this study believed that the activity participation of Chinese seniors can be divided into these categories: exercise activities, social activities, and interaction with friends. Based on the characteristics of CHARLS, this study further summarized the activity participation of middle-aged and older people as physical activity participation (participation in exercise-related activities) and social leisure activity participation (participation in activities related to socializing and friend-making).

In addition, previous physical activity participation has mostly been measured only in terms of intensity, duration, and days of physical activity participation [31,32]. To measure physical activity participation more comprehensively, based on previous studies and the HRQoL scale of CHARLS [6,33], this study chose to comprehensively measure the physical activity participation of middle-aged and older people (as shown in Multimedia Appendix 2) from the aspects of PF, role-body (RP), body pain (BP), general health (GH).

In this study, PF measured whether health conditions prevent normal physical activity participation, the physiological functioning dimension measures whether physical health problems lead to limited participation in physical activity, the somatic pain dimension measures BP and its impact on participation in daily physical activities, and the overall health dimension measures their own evaluation of health status and trends [33]. In terms of measuring social leisure activity participation, based on a previous study [30], 2 questions were asked (as shown in Multimedia Appendix 2): What social activities have been carried out in the past month? What is the frequency of participating in these activities?

This study rescored these 2 types of activity parameters in reverse, ensuring that higher scores indicated more active activity participation. Then, based on the scoring rules of the Medical Outcomes Study (MOS) 36-item Short-Form (SF-36) Health Survey scale, the raw scores of the 2 activities' participation in their respective components were calculated. The conversion scores of daily physical activity participation and social leisure activity participation were calculated using the range method between 0 and 100. Finally, the Cronbach α coefficient was used to measure the reliability of the above method. The results showed that the Cronbach α value of both physical activity participation and social leisure activity participation was higher than 0.6; hence, this scoring method is reliable [6,33].

Depression

Depression was screened using the Chinese version of the Center for Epidemiologic Study Depression Scale (CES-D), with 10 short-form scale items [33], and results illustrated the magnitude of depression. The CES-D in CHARLS is a simplified version of the depression scale developed by Radolff at the National Institute of Mental Health [34]. The CES-D consists of 10 questions asking respondents about depression experienced 1 week before, and existing research has demonstrated its applicability in middle-aged and older Chinese populations [35]. Of these 10 questions, this study reverse-coded questions 5 and 8, and the total score was the sum of all questions. Each question is scored on a 4-point scale, ranging from 0 to 3, and the total score ranges from 0 to 30; therefore, the higher the score, the more severe the depression [36].

Control Variable

Age, gender, education level, marital status, chronic disease, smoking, drinking, self-reported health, and number of chronic diseases were included as covariates in the analysis. The results of the 4 surveys of these variables were sorted out. This study took the baseline measurement results of gender and education level as time-invariant variables, leaving the remaining variables as time-varying variables. The education level was divided into "no formal education," "primary school," "junior high school," and "high school education or above" [23,30]. Marital status included "married" and "other." Self-perceived health status was reported as "very good," "good," "fair," "poor," or "bad." The current smoking and drinking status was assessed by self-reporting based on the questions "Are you currently smoking?" and "Are you currently drinking?" Chronic diseases were all self-reported and marked as 0, 1, or at least 2 by the number of chronic diseases.

Statistical Analysis

Descriptive analysis was performed using IBM SPSS (version 24.0) to determine the basic characteristics of all respondents.

The association of activity participation and depression was explored using Mplus (version 8.0). First, to study the respective trends in activity participation and depression in the middle-aged and older populations, an unconditional linear growth model and an unconditional quadratic growth model were constructed (see Multimedia Appendices 3 and 4). Second, a growth model with time-invariant and time-varying variables was constructed to study the interrelationship between depression and physical and leisure activities. Participation and depression were added to the model as time-varying variables, mainly looking at the relationship between activity participation and depression at the same time. Next, to avoid measurement errors and more accurately examine the relationship between depression and physical and social activities, a parallel latent growth model was developed (see Figure 1). Finally, after studying the dynamic properties of the variables using this model, the main lag relationship between depression and physical and social activities in the older population was further investigated using a cross-lag model (see Figure 2).

The hypothetical model fit was tested using multiple fit indices, including the standardized root mean square residual (SRMR), the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI). Specifically, if SRMR≤0.08, the model is acceptable; if RMSEA≤0.08, the model is acceptable. For CFI and TLI, values >0.90 are considered acceptable [37].

Figure 1. (A) LGCM of activity participation and depression. (B) LGCM of social leisure activity participation and depression. The activity participation intercept, activity participation slope, depression intercept, and depression slope were regressed on the covariates simultaneously. Valid paths and nonsalient paths are plotted as solid and dashed lines, respectively. I: intercept; LGCM: latent growth curve model; S: slope.

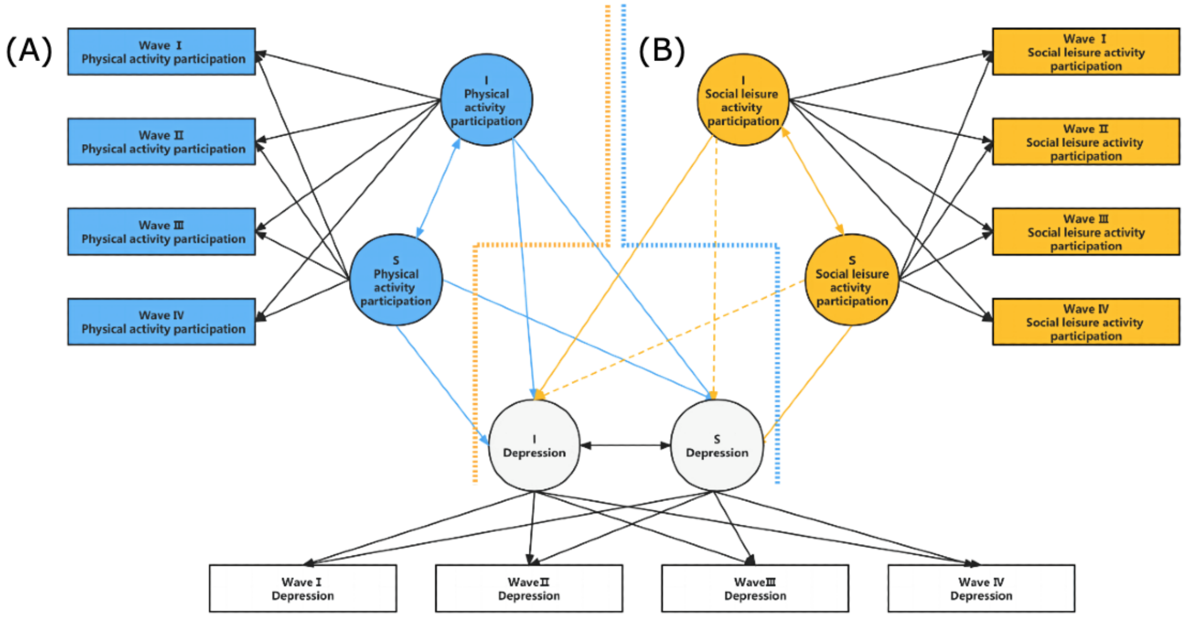
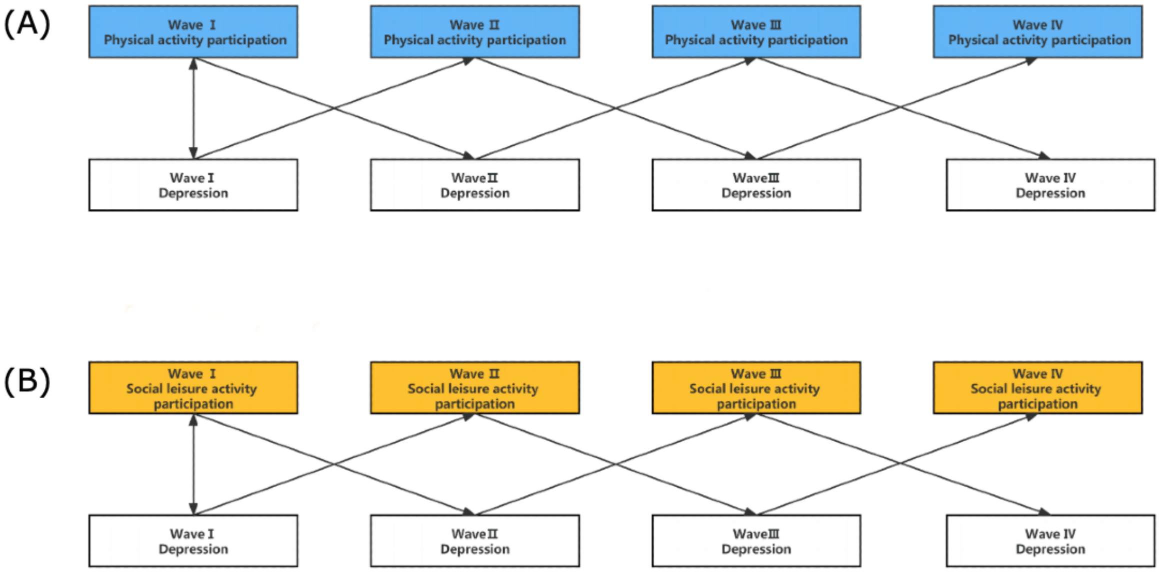


Figure 2. (A) Cross-lag model of physical activity participation and depression. (B) Cross-lag model of social leisure activity participation and depression. Note: activity participation and depression in all 4 waves were regressed on the covariates simultaneously.



Results

Common Method Bias

The Harman single-factor test is a diagnostic technique to evaluate the severity of common method bias. In this study, the Harman single-factor test was performed on all the data from the 4 surveys. The test results showed that in these 4 surveys, the variance explained by the first factor was 21.7%, which is significantly less than 40.0%. This suggests that there is no common methodological bias in this study [38].

Descriptive Statistics

Table 1 shows the descriptive statistics of all variables at baseline. The results showed that from 2011 to 2018, the average values of physical activity participation, social leisure activity participation, and depression ranged from 76.98 (SD 15.16) to 83.95 (SD 5.72), from 7.43 (SD 8.67) to 9.25 (SD 10.16), and from 7.61 (SD 5.72) to 8.82 (SD 6.51), respectively, respectively. Multimedia Appendices 5 and 6 are a clearer demonstration of the scores of and linear trends in depression, physical activity participation, and social leisure activity participation among middle-aged and older people during the 4 measurement periods. Multimedia Appendix 7 shows the binary correlations of key variables across 4 surveys. The results showed a negative correlation between physical activity participation, social leisure activity participation, and depression from 2011 to 2018. That is, the higher the scores of physical activity participation and social leisure activity participation, the lower the score of depression.

Table 1. Descriptive analysis of participants (N=4818).

Variable	Wave I, mean (SD)	Wave II, mean (SD)	Wave III, mean (SD)	Wave IV, mean (SD)
Age (years)	57.13 (7.95)	59.13 (7.94)	61.13 (7.95)	64.13 (7.95)
Marital status	1.36 (1.09)	1.41 (1.17)	1.45 (1.20)	1.64 (1.38)
Education level	3.65 (1.90)	2.06 (0.95)	2.05 (0.96)	1.95 (0.91)
Self-reported health	3.41 (0.10)	2.60 (1.01)	3.41 (1.06)	2.99 (1.01)
Chronic disease	1.31 (1.36)	1.58 (1.50)	1.79 (1.80)	2.37 (1.92)
Smoking	0.32 (0.47)	0.2 (0.40)	0.28 (0.45)	0.27 (0.44)
Drinking	0.35 (0.48)	0.29 (0.46)	0.37 (0.48)	0.34 (0.47)
Depression	7.93 (6.11)	7.61 (5.72)	7.75 (6.29)	8.62 (6.51)
Physical activity participation	83.95 (5.72)	79.25 (11.06)	77.43 (13.93)	76.98 (15.16)
Social leisure activity participation	7.43 (8.67)	9.25 (10.16)	8.93 (10.56)	8.25 (10.20)

Developmental Trajectories of Depression Levels Among Seniors

To investigate the potential relationship between depressive symptoms and physical activity participation as well as social leisure activity participation, unconditional linear and unconditional quadratic latent growth curve models (LGCMs) were initially used to assess initial depression, physical activity participation, social leisure activity participation, and their trajectories over time in older adults. The results are shown in Multimedia Appendices 3 and 4, respectively. Multimedia Appendix 8 summarizes the fit indices for these models. Table 2 provides the within-subject mean for repeated measures (intercept/baseline), the between-subject variance in the intercept (intercept variance), the between-subject mean change (slope), the between-subject mean change (slope variance), and the quadratic slope variance and quadratic intercept variance.

With regard to the depression degree, both the linear growth model (model 1) and the quadratic growth model (model 2) of the depression degree were suitable (model 1: CFI=0.981, TCL=0.977, RMSEA=0.057, SRMR=0.031; model 2: CFI=1.000, TCL=1.001, RMSEA=1.001, SRMR=0.013). Although the fitness was acceptable, the quadratic slope variance of depression in the quadratic growth model of depression was not statistically significant (B=0.011, *P*=.045), as shown in Table 2. Therefore, the study chose to analyze the data from model 1. It was found that the variance estimates of the intercept factor and slope factor of the depression degree were 18.705 (*P*<.001) and 0.163 *P*<.001), respectively, indicating that there were significant interindividual differences in initial levels and growth rates. In addition, the annual mean slope of the depression level was 0.110 (SD 0.013, *P*<.001), meaning the depression level increased linearly during the 4 survey periods.

Table 2. Intercept and slope estimates for unconditional linear and nonlinear latent growth models.

Trajectory and models	Intercept (SE)	Intercept variance (SE)	Slope (SE)	Slope variance (SE)	Quadratic (SE)	Quadratic variance (SE)
Trajectory of depression						
Model 1	7.595 (0.079)[a]	18.705 (0.653)[a]	0.110 (0.013)[a]	0.163 (0.026)[a]	N/A[b]	N/A
Model 2	7.910 (0.086)[a]	16.305 (1.340)[a]	–0.243 (0.042)[a]	–0.074 (0.379)	0.049 (0.006)[a]	0.011 (0.005)
Trajectory of physical activity participation						
Model 3	82.620 (0.112)[a]	17.351 (1.377)[a]	–0.913 (0.032)[a]	2.557 (0.132)[a]	N/A	N/A
Model 4	83.817 (0.113)[a]	1.673 (3.334)[a]	–2.594 (0.090)[a]	9.874 (1.240)[a]	0.234 (0.011)[a]	0.046 (0.023)
Trajectory of physical activity participation						
Model 5	7.848 (0.121)[a]	37.665 (1.575)[a]	0.134 (0.022)[a]	0.592 (0.077)[a]	N/A	N/A
Model 6	7.265 (0.124)[a]	34.239 (3.580)[a]	1.028 (0.073)[a]	3.655 (1.058)[a]	–0.128 (0.010)[a]	0.049 (0.017)[a]

[a] The corresponding variable was statistically significant.

[b] N/A: not applicable.

Developmental Trajectories of Physical Activity Participation and Leisure Activity Participation Among Older Adults

In terms of physical activity, the quadratic growth model (model 4) for their latent growth was better than the linear growth model (model 3); see Multimedia Appendix 8. With regard to the quadratic growth model for physical activity participation, the quadratic slope variance of physical activity participation was 0.046 (*P*=.048); see Table 2. Model 4 data were used as a result. Specifically, the initial level (intercept) of physical activity participation was 83.817 (*P*<.001). Physical activity participation decreased over the 4 test periods (slope=–2.594, *P*<.001), and the rate of decline increased year by year (curve slope=0.234, *P*<.001), indicating a nonlinear decline in physical activity participation over the 4 test periods. In addition, the estimated variances of the intercept factor and slope factor for physical activity participation were 1.673 and 9.874, respectively, and the variance of the slope factor was significant at the 0.001 level, indicating interindividual differences in growth rates.

Based on the fitting degree of the latent growth model of social leisure activity participation, the fitting degree of the quadratic growth model (model 6) of social leisure activity participation was better than that of the linear growth model (model 5) of social leisure activity participation (see Multimedia Appendix 8, Table S3). In addition, the quadratic slope variance of social leisure activity participation was 0.049 (*P*<.001); see Table 2. Therefore, model 6 data were used for this study. Specifically, the initial level of social leisure activity participation (intercept) was 7.265 (*P*<.001). The slope of social leisure activity participation during the 4 survey periods was 1.028 (*P*<.001), indicating that social leisure activity participation showed a linear upward trend in the 4 test periods (slope=1.028, *P*<.001), but the rate of increase decreased year by year (curve slope=–0.128, *P*<.001). In addition, the estimated variances of the intercept factor and the slope factor of social leisure activity participation were 34.239 (*P*<.001) and 3.655 (*P*<.001), respectively, indicating that there were significant interindividual differences in the initial level and growth rate.

Relationship Between Physical Activity Participation and Social Leisure Activity Ability and Depression

To explore the potential relationship between the depression level and participation in physical and social leisure activities, a model was built with time-invariant and time-varying variables. This model examined whether taking physical activity participation and social leisure activity participation as time-varying variables has an effect on the depression level and whether the depression level is time-varying on physical activity participation when these endogenous variables are considered at the same time. The results of the model are shown in Table 3.

The results showed that at any time point, the lower the physical activity participation and social leisure activity participation, the higher the degree of depression, which means that physical activity participation and social leisure activity participation did affect the depression level of middle-aged and older people at corresponding time points, and based on the results, physical activity participation has a greater effect on depression than social leisure activity participation.

In addition, the effect of depression level on physical activity participation and social leisure activity participation was also considered. The results showed that at any time point, the higher the degree of depression, the lower the ability to participate in physical and social leisure activities, which means that the degree of depression does affect physical activity participation and social leisure activity participation for middle-aged and older people at corresponding time points.

Table 3. Standardized coefficients[a] in LGCMs[b] with depression, physical activity participation, and social leisure activity participation as time-varying variables.

Variables	Wave III depression	Wave II depression	Wave III depression	Wave IV epression
Physical activity participation (depression				
Wave I physical activity participation (depression)	–0.167 (–0.024)	N/A[c]	N/A	N/A
Wave II physical activity participation (depression)	N/A	–0.284 (–0.05)	N/A	N/A
Wave III physical activity participation (depression)	N/A	N/A	–0.368 (–0.06)	N/A
Wave IV physical activity participation (depression)	N/A	N/A	N/A	–0.396 (–0.052)
Social leisure activity participation (depression)				
Wave I social leisure activity participation (depression)	–0.064 (–0.092)	N/A	N/A	N/A
Wave II social leisure activity participation (depression)	N/A	–0.058 (–0.058)	N/A	N/A
Wave III social leisure activity participation (depression)	N/A	N/A	–0.056 (–0.068)	N/A
Wave IV social leisure activity participation (depression)	N/A	N/A	N/A	–0.048 (–0.060)

[a] All *P*<.001.

[b] LGCM: latent growth curve model.

[c] N/A: not applicable.

Parallel LGCM

To avoid measurement errors and to more accurately examine whether depression is associated with physical activity participation and whether the trajectories change over time, a parallel LGCM was developed, as shown in Figure 1. This parallel LGCM examined the relationship between the intercepts and slopes of physical activity participation and social leisure activity participation and the intercepts and slopes of depression, and further examined the influence process between the 2 by setting regression equations between growth factors. This study, however, only focused on the interaction between the 2.

The parallel LGCM of physical activity participation for depression was well fitted (CFI=0.908, TLI=0.873, RMSEA=0.056). In Figure 1, the effective paths and nonsalient paths are drawn as solid and dashed lines, respectively. Detailed results are shown in Table 4. At the initial level, physical activity participation was negatively correlated with depression (r=–0.631, *P*<.001), that is, the worse the individual's initial physical activity participation, the higher the depression. Initial physical activity participation positively predicted the rate of change in depression (β=0.461, *P*<.001), indicating that individuals with higher initial physical activity participation scores have a higher rate of depression growth. Similarly, the slope of physical activity participation in the middle-aged and older populations affected not only the initial depression level (β=–0.261, *P*<.001) but also the rate of change in depression (β=–0.69, *P*<.001), indicating that the growth rate of physical activity participation suppresses the increase in the initial level of depression and its growth rate.

In addition, another parallel LGCM was built for social leisure activity participation and depression (see Table 4). The parallel LGCM fit well (CFI=0.951, TLI=0.933, RMSEA=0.033, RSMR=0.036). In Figure 1, the effective and nonsalient paths are plotted as solid and dashed lines, respectively. At the initial level, the degree of depression was negatively correlated with social leisure activity participation (r=–0.223, *P*<.001), indicating that the higher the individual's initial social leisure activity participation, the lower the depression. Initial social leisure activity participation did not predict the rate of change in depression (β=0.067, *P*=.159), indicating that initial social leisure activity participation is not associated with the rate of increase in depression. Similarly, although the slope of social leisure activity participation in the middle-aged and older populations did not affect the initial level of depression (β=0.08, *P*=.087), it did affect the rate of change in depression (β=–0.303, *P*<.001), indicating that the growth of social leisure activity participation suppresses the growth rate of depression.

Table 4. Results of LGCMs[a] for activity participation and depression.

Path	Estimate (SE)	*P* value
Physical activity participation → depression		
I[b] physical activity participation → I depression	–0.631 (0.031)	<.001
S[c] physical activity participation → I depression	–0.261 (0.025)	<.001
I physical activity participation → S depression	0.461 (0.065)	<.001
S physical activity participation → S depression	–0.69 (0.061)	<.001
Social leisure activity participation → depression		
I social leisure activity participation → I depression	–0.23 (0.027)	<.001
S social leisure activity participation → I depression	0.08 (0.047)	.087
I social leisure activity participation → S depression	0.067 (0.047)	.159
S social leisure activity participation → S depression	–0.303 (0.084)	<.001

[a] LGCM: latent growth curve model.

[b] I: intercept.

[c] S: slope.

Cross-Lag Model

The dynamic characteristics of the variables were studied using the LGCM. To further investigate the main lag relationship between activity participation and depression in the middle-aged and older populations, cross-lag regression analysis was performed across 4 measures. Cross-lag regression analysis can reveal complex relationships between 2 variables. The autoregressive effect of each variable is controlled by setting a coefficient of stability, which is the best way to test for "pure" effects between variables and to understand to what extent one variable predicts the other. In addition, to achieve a more compact model, the crossing paths across the waves were restricted to achieve equality, as shown in Figure 2. This procedure sets autoregressive effects as fixed effects, which avoids convergence problems. The relationship between variables was not considered stable across time, because only the lag effect was set as a random effect.

The results are shown in Table 5. Both activity participation and depression maintained high stability across the 4 measurements. Physical activity participation in the older population in wave I significantly and negatively predicted depression in wave II (β=–0.125, *P*<.001). Physical activity participation in the middle-aged and older populations in wave II significantly and negatively predicted depression in wave III (β=–0.216, *P*<.001). Physical activity participation in the older population in wave III significantly and positively predicted depression in wave IV (β=–0.233, *P*<.001). Likewise, prior depression also had a significant negative impact on later physical activity participation.

Table 5. Cross-lag regression path coefficients.[a]

Regression path	Wave I to wave II	Wave II to wave III	Wave III to wave IV
Physical activity participation and depression			
Physical activity participation → depression	–0.125	–0.216	–0.233
Depression → physical activity participation	–0.874	–0.737	–0.685
Depression → depression	0.456	0.414	0.414
Physical activity participation → physical activity participation	0.09	–0.049	–0.049
Social leisure activity participation and depression			
Social leisure activity participation → depression	–0.04	–0.045	–0.044
Depression → social leisure activity participation	–0.058	–0.054	–0.058
Depression → depression	0.515	0.475	0.481
Social leisure activity participation → social leisure activity participation	0.399	0.430	0.446

[a] All *P*<.001.

Similar results were shown in the cross-lag model of social leisure activity participation and depression as in the cross-lag model of physical activity participation and depression. The social leisure activity participation of the older population in wave I significantly and negatively predicted depression in wave II (β=–0.04, *P*<.001). Social leisure activity participation in the middle-aged and older populations in wave II significantly and negatively predicted depression in wave III (β=–0.045, *P*<.001).

The social leisure activity participation of the older population in wave III significantly and predicted depression in wave IV (β=–0.044, *P*<.001). Similarly, previous depression also had a significant negative impact on later social leisure activity participation.

Discussion

Principal Results

Based on a CHARLS survey spanning 7 years, this study found that activity participation has a long-term effect on depression. First, physical activity participation decreased significantly between 2011 and 2018 among middle-aged and older Chinese populations, suggesting that the physical activity participation of middle-aged and older adults decreases with age [14,18]. Studies have shown that chronic diseases related to aging can lead to the risk of arthritis and disability and that impairment of activities of daily living gradually increase with age [39,40], which in turn leads to a lifestyle change in middle-aged and older people. A series of chain reactions end up with a reduction in physical activity participation [41]. Life cycle theory can explain this change well. This study selected respondents over the age of 45 years. As the survey continued, the life cycle of these respondents gradually entered a decline period, so their PFs also began to decline. We can even speculate that this phenomenon might be more pronounced if the initial age restriction is controlled to over 60 years [14]. However, it is worth mentioning that it is precisely because of our selection of the age of ≥45 years (the average age in wave I was 57 years) that as these people get older, they start to withdraw from work (the Chinese government stipulates that 60 years is the retirement age for citizens) and have more time for social leisure activity. Therefore , there was a significant increase in leisure activity participation from wave I to wave II. However, after wave II, most of them (the average age in wave II was 60 years) also entered old age, and social leisure activity began to decline continuously, which is also consistent with previous studies [14]. Finally, based on the analysis of the trajectory of depression, the level of depression in middle-aged and older people in China increased significantly between 2011 and 2018, and updated evidence was provided again to prove that the trajectory of depression is affected by age [8,23,42].

Second, the study found a horizontal correlation between activity participation and depression within the same period and further demonstrated that this relationship persisted during different periods. Specifically, models incorporating time-varying and invariant variables proved that between 2011 and 2018, activity participation in middle-aged and older adults significantly predicted depression in the same period and that depression in middle-aged and older adults also significantly predicted participation in activities in the same period. This suggests that the trajectory of depression is affected by activity participation, and the developmental trajectory of activity participation is similarly affected by depression [42-44]. The behavioral theory of depression and BA theory can successfully explain the mutual influence of activity participation on depression. The behavioral theory of depression states that low-response-conditioned positive reinforcement (RCPR) rates can directly trigger depression, so depression can be improved by re-establishing RCPR rates through altering activity participation and other activities that make individuals feel pleasant [45] and that improvements in depression in turn will lead to greater activity participation [46]. Previous studies have also found that the activity participation of middle-aged and older people may be plagued by poor physical and mental health (eg, depression, anxiety) [47] and that poor mental health and related symptoms (eg, irritability) result in reducing activity participation through perceptions of stress, perceptions of health, and some physical symptoms [27].

Additionally, a parallel LGCM showed that the higher the initial level of activity participation, the slower the rise in depression, which is consistent with previous analyses that showed that activity participation is a protective factor for depression among older adults [48,49]. Active participation in various activities as an active lifestyle attitude in middle-aged and older people can inhibit the increase in depression [50]. Moreover, the parallel LGCM also found that the slope of activity participation was significantly correlated with the slope of depression, suggesting that the rate of decline in activity participation among older adults could predict the rate of decline in depression, meaning that the faster the decline in activity participation, the sooner the depression rate increase. In addition, our study suggests that the development of depression partly influences the development of activity participation. One possible explanation is that depression seriously affects the physical health of patients [51], and physical health is directly related to the activity participation of middle-aged and elderly people [52].

Finally, the study performed a cross-lag regression analysis on activity participation and depression in middle-aged and elderly populations to observe the time series between activity participation and depression. Results showed that activity participation in middle-aged and older people can positively predict subsequent depression and that depression can also predict subsequent activity participation; however, no 1-way time series of activity participation and depression was found. This result further proves the bidirectional correlation between activity participation and depression in middle-aged and older people [53].

Limitations

This study has some limitations. First, the participation data of this research activity came from the subjective answers of middle-aged and older people. As much as the study team tried to control some biases caused by subjective answers, such as excluding respondents with cognitive impairment, intellectual disability, and memory problems from the study, there would still be problems posed by subjective answers. Second, physical activity participation and social leisure activity participation were categorized to represent the activity participation of middle-aged and older people. The frequency of participation and the quality of activities were included in the selection of indicators; however, with regard to the different types of activity participation, the classification is still not detailed. For future studies, a detailed division of different types of activity participation should be made to compare the effects of different types of activity participation on the physical and mental health of middle-aged and older people. Finally, the focus of the study was the effect of activity participation on depression. Although the study found that the development of depression also affects the development of activity participation, it did not focus on explaining this relationship, and future research should pay more attention to depression's impact on activity participation.

Conclusion

Study results support a bidirectional association between activity participation and depression. Activity participation affects not only the initial level of depression but also the trend in changes in depression. In the future, more attention should be paid to depression in middle-aged and older people with less active participation. The interconnectedness of early activity participation and subsequent depression is noteworthy and may have implications for public health. Overall, this study highlights the relationship between activity participation and depression level changes, bringing positive implications for preventing depression aggravation in later life.

Acknowledgments

This work was supported by the National Social Science Foundation of China (grant number: 21BTY054), the China Postdoctoral Science Foundation (grant number: 2017M622169), and the Future Project for Youth Scholar of Shandong University (grant number: 2017WLJH17). The content of the paper is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

We thank all the interviewees for their voluntary participation in the China Health and Retirement Longitudinal Study (CHARLS).

Data Availability

The data sets generated and analyzed during the study are available from the corresponding author upon reasonable request.

Authors' Contributions

YG was responsible for the study design, data analysis, interpretation of the data, and writing of the manuscript. ZJ was responsible for the study design, data analysis, and writing of the manuscript. LZ and SH were responsible for the study design and writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Flowchart of study respondents.
[DOCX File , 351 KB - publichealth_v9i1e44682_app1.docx]

Multimedia Appendix 2
HRQoL variable selection. HRQoL: health-related quality of life.
[DOCX File , 13 KB - publichealth_v9i1e44682_app2.docx]

Multimedia Appendix 3
Unconditional linear LGCM. i: intercept; LGCM: latent growth curve model; s: slope.
[DOCX File , 91 KB - publichealth_v9i1e44682_app3.docx]

Multimedia Appendix 4
Unconditional nonlinear LGCM. i: intercept; LGCM: latent growth curve model; s: slope.
[DOCX File , 137 KB - publichealth_v9i1e44682_app4.docx]

Multimedia Appendix 5
Change trajectory of depression.
[DOCX File , 3542 KB - publichealth_v9i1e44682_app5.docx]

Multimedia Appendix 6
Change trajectory of physical activity participation and social leisure activity participation.
[DOCX File , 3542 KB - publichealth_v9i1e44682_app6.docx]

Multimedia Appendix 7

Correlation coefficient matrix.
[DOCX File , 15 KB - publichealth_v9i1e44682_app7.docx]

Multimedia Appendix 8
Fit statistics for LGCMs. LGCM: latent growth curve model.
[DOCX File , 14 KB - publichealth_v9i1e44682_app8.docx]

References

1. Yu W, Chen R, Zhang M, Li Z, Gao F, Yu S, et al. Cognitive decline trajectories and influencing factors in China: a non-normal growth mixture model analysis. Arch Gerontol Geriatr 2021 Jul;95:104381. [doi: 10.1016/j.archger.2021.104381] [Medline: 33657489]
2. Bath PA, Deeg D. Social engagement and health outcomes among older people: introduction to a special section. Eur J Ageing 2005 Mar 9;2(1):24-30 [FREE Full text] [doi: 10.1007/s10433-005-0019-4] [Medline: 28794713]
3. Martinez HN, Santaella RE, Rodriguez-Garcia AM. Benefits of physical activity for the promotion of active aging in elderly. Bibliographic review. Retos-Nuevas Tendencias En Educacion Fisica Deporte Y Recreacion 2021;2021(39):829-834. [Medline: 000577598600001]
4. Meyer S, Rosenblum S. Children with celiac disease: health-related quality of life and leisure participation. Am J Occup Ther 2016;70(6):7006220010p1-7006220010p8. [doi: 10.5014/ajot.2016.020594] [Medline: 27767940]
5. Lemon BW, Bengtson VL, Peterson JA. An exploration of the activity theory of aging: activity types and life satisfaction among in-movers to a retirement community. J Gerontol 1972 Oct 01;27(4):511-523. [doi: 10.1093/geronj/27.4.511] [Medline: 5075497]
6. Wang L, Liu W. Effects of family doctor contract services on the health-related quality of life among individuals with diabetes in China: evidence from the CHARLS. Front Public Health 2022;10:865653 [FREE Full text] [doi: 10.3389/fpubh.2022.865653] [Medline: 35602141]
7. Herrman H, Patel V, Kieling C, Berk M, Buchweitz C, Cuijpers P, et al. Time for united action on depression: a Lancet–World Psychiatric Association Commission. Lancet 2022 Mar;399(10328):957-1022. [doi: 10.1016/s0140-6736(21)02141-3]
8. Yang Y. Is old age depressing? Growth trajectories and cohort variations in late-life depression. J Health Soc Behav 2007 Mar 24;48(1):16-32. [doi: 10.1177/002214650704800102] [Medline: 17476921]
9. Li D, Zhang D, Shao J, Qi X, Tian L. A meta-analysis of the prevalence of depressive symptoms in Chinese older adults. Arch Gerontol Geriatr 2014;58(1):1-9. [doi: 10.1016/j.archger.2013.07.016] [Medline: 24001674]
10. Zhang Y, Zhao M. Gender disparities and depressive symptoms over the life course and across cohorts in China. J Affect Disord 2021 Dec 01;295:620-627. [doi: 10.1016/j.jad.2021.08.134] [Medline: 34509776]
11. Agahi N, Lennartsson C, Kåreholt I, Shaw BA. Trajectories of social activities from middle age to old age and late-life disability: a 36-year follow-up. Age Ageing 2013 Nov 24;42(6):790-793. [doi: 10.1093/ageing/aft140] [Medline: 24064237]
12. Park S, Kwon E, Lee H. Life course trajectories of later-life cognitive functions: does social engagement in old age matter? Int J Environ Res Public Health 2017 Apr 07;14(4):393 [FREE Full text] [doi: 10.3390/ijerph14040393] [Medline: 28387725]
13. Sluzki CE. Personal social networks and health: conceptual and clinical implications of their reciprocal impact. Fam Syst Health 2010 Mar;28(1):1-18. [doi: 10.1037/a0019061] [Medline: 20438199]
14. Zhu C, Zhou L, Zhang X. Effects of leisure activities on the cognitive ability of older adults: a latent variable growth model analysis. Front Psychol 2022 Apr 13;13:838878 [FREE Full text] [doi: 10.3389/fpsyg.2022.838878] [Medline: 35496137]
15. Dong X, Li Y, Simon MA. Social engagement among U.S. Chinese older adults--findings from the PINE Study. J Gerontol A Biol Sci Med Sci 2014 Nov 05;69 Suppl 2(Suppl 2):S82-S89 [FREE Full text] [doi: 10.1093/gerona/glu152] [Medline: 25378453]
16. Kim DE, Yoon JY. Trajectory classes of social activity and their effects on longitudinal changes in cognitive function among older adults. Arch Gerontol Geriatr 2022 Jan;98:104532. [doi: 10.1016/j.archger.2021.104532] [Medline: 34592681]
17. Kelly ME, Duff H, Kelly S, McHugh Power JE, Brennan S, Lawlor BA, et al. The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. Syst Rev 2017 Dec 19;6(1):259 [FREE Full text] [doi: 10.1186/s13643-017-0632-2] [Medline: 29258596]
18. Feng Q, Fong JH, Zhang W, Liu C, Chen H. Leisure activity engagement among the oldest old in China, 1998–2018. Am J Public Health 2020 Oct;110(10):1535-1537. [doi: 10.2105/ajph.2020.305798]
19. Wu CP, Jiang XQ. An Introduction to Gerontology. Beijing: Renmin University of China Press; 2006.
20. Glass TA, De Leon CFM, Bassuk SS, Berkman LF. Social engagement and depressive symptoms in late life: longitudinal findings. J Aging Health 2006 Aug 30;18(4):604-628. [doi: 10.1177/0898264306291017] [Medline: 16835392]
21. Greaves CJ, Farbus L. Effects of creative and social activity on the health and well-being of socially isolated older people: outcomes from a multi-method observational study. J R Soc Promot Health 2006 May 07;126(3):134-142. [doi: 10.1177/1466424006064303] [Medline: 16739619]

22. Hong S, Hasche L, Bowland S. Structural relationships between social activities and longitudinal trajectories of depression among older adults. Gerontologist 2009 Feb 18;49(1):1-11 [FREE Full text] [doi: 10.1093/geront/gnp006] [Medline: 19362999]
23. Huang J, Wong R, Chen C, Mao I, Huang C, Chang W, et al. Trajectory of depression symptoms and related factors in later life--a population based study. J Affect Disord 2011 Oct;133(3):499-508. [doi: 10.1016/j.jad.2011.04.048] [Medline: 21616542]
24. Niti M, Ng T, Kua EH, Ho RCM, Tan CH. Depression and chronic medical illnesses in Asian older adults: the role of subjective health and functional status. Int J Geriatr Psychiatry 2007 Nov;22(11):1087-1094. [doi: 10.1002/gps.1789] [Medline: 17407107]
25. Wang L, van Belle G, Kukull WB, Larson EB. Predictors of functional change: a longitudinal study of nondemented people aged 65 and older. J Am Geriatr Soc 2002 Sep;50(9):1525-1534. [doi: 10.1046/j.1532-5415.2002.50408.x] [Medline: 12383150]
26. Carvalho JP, Hopko DR. Behavioral theory of depression: reinforcement as a mediating variable between avoidance and depression. J Behav Ther Exp Psychiatry 2011 Jun;42(2):154-162. [doi: 10.1016/j.jbtep.2010.10.001] [Medline: 21315876]
27. Dragioti E, Gerdle B, Levin L, Bernfort L, Dong H. Association between participation activities, pain severity, and psychological distress in old age: a population-based study of Swedish older adults. Int J Environ Res Public Health 2021 Mar 10;18(6):2795 [FREE Full text] [doi: 10.3390/ijerph18062795] [Medline: 33801881]
28. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb 12;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
29. Deeg DJH, Bath PA. Self-rated health, gender, and mortality in older persons: introduction to a special section. Gerontologist 2003 Jun 01;43(3):369-371. [doi: 10.1093/geront/43.3.369] [Medline: 12810900]
30. Wang R, Chen Z, Zhou Y, Shen L, Zhang Z, Wu X. Melancholy or mahjong? Diversity, frequency, type, and rural-urban divide of social participation and depression in middle- and old-aged Chinese: a fixed-effects analysis. Soc Sci Med 2019 Oct;238:112518. [doi: 10.1016/j.socscimed.2019.112518] [Medline: 31473574]
31. Rahman M, Liang C, Gu D, Ding Y, Akter M. Understanding levels and motivation of physical activity for health promotion among Chinese middle-aged and older adults: a cross-sectional investigation. J Healthc Eng 2019;2019:9828241 [FREE Full text] [doi: 10.1155/2019/9828241] [Medline: 31534647]
32. Lau PWC, Wang JJ, Ransdell LL, Shi L. The effectiveness of Facebook as a social network intervention to increase physical activity in Chinese young adults. Front Public Health 2022 Jul 22;10:912327 [FREE Full text] [doi: 10.3389/fpubh.2022.912327] [Medline: 35937270]
33. Hao X, Yang Y, Gao X, Dai T. Evaluating the effectiveness of the health management program for the elderly on health-related quality of life among elderly people in China: findings from the China Health and Retirement Longitudinal Study. Int J Environ Res Public Health 2019 Jan 03;16(1):113 [FREE Full text] [doi: 10.3390/ijerph16010113] [Medline: 30609852]
34. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Measur 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
35. Zhou L, Ma X, Wang W. Relationship between cognitive performance and depressive symptoms in Chinese older adults: the China Health and Retirement Longitudinal Study (CHARLS). J Affect Disord 2021 Feb 15;281:454-458. [doi: 10.1016/j.jad.2020.12.059] [Medline: 33360747]
36. Amtmann D, Kim J, Chung H, Bamer AM, Askew RL, Wu S, et al. Comparing CESD-10, PHQ-9, and PROMIS depression instruments in individuals with multiple sclerosis. Rehabil Psychol 2014 May;59(2):220-229 [FREE Full text] [doi: 10.1037/a0035919] [Medline: 24661030]
37. Wang X, Lu J, Liu Q, Yu Q, Fan J, Gao F, et al. Childhood experiences of threat and deprivation predict distinct depressive symptoms: a parallel latent growth curve model. J Affect Disord 2022 Dec 15;319:244-251. [doi: 10.1016/j.jad.2022.09.061] [Medline: 36162654]
38. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
39. Rosso AL, Eaton CB, Wallace R, Gold R, Curb JD, Stefanick ML, et al. Combined impact of geriatric syndromes and cardiometabolic diseases on measures of functional impairment. J Gerontol A Biol Sci Med Sci 2011 Mar 11;66(3):349-354 [FREE Full text] [doi: 10.1093/gerona/glq230] [Medline: 21317242]
40. Chen C, Mullan J, Griffiths D, Kreis IA, Lan T, Chiu H. Trajectories of depression and their relationship with health status and social service use. Arch Gerontol Geriatr 2011 Sep;53(2):e118-e124. [doi: 10.1016/j.archger.2010.07.006] [Medline: 20810178]
41. Zimmer Z, Hickey T, Searle MS. The pattern of change in leisure activity behavior among older adults with arthritis. Gerontologist 1997 Jun 01;37(3):384-392. [doi: 10.1093/geront/37.3.384] [Medline: 9203762]
42. Tan KL, Yadav H. Depression among the urban poor in Peninsular Malaysia: a community based cross-sectional study. J Health Psychol 2013 Jan 09;18(1):121-127. [doi: 10.1177/1359105311433908] [Medline: 22322990]

43. Wassink-Vossen S, Collard RM, Oude Voshaar RC, Comijs HC, de Vocht HM, Naarding P. Physical (in)activity and depression in older people. J Affect Disord 2014 Jun;161:65-72. [doi: 10.1016/j.jad.2014.03.001] [Medline: 24751309]
44. Lindwall M, Larsman P, Hagger MS. The reciprocal relationship between physical activity and depression in older European adults: a prospective cross-lagged panel design using SHARE data. Health Psychol 2011 Jul;30(4):453-462. [doi: 10.1037/a0023268] [Medline: 21480713]
45. Abreu PR, Santos CE. Behavioral models of depression: a critique of the emphasis on positive reinforcement. Int J Behav Consult Ther 2008;4(2):130-145. [doi: 10.1037/h0100838]
46. van Genugten CR, Schuurmans J, Hoogendoorn AW, Araya R, Andersson G, Baños R, et al. Examining the theoretical framework of behavioral activation for major depressive disorder: smartphone-based ecological momentary assessment study. JMIR Ment Health 2021 Dec 06;8(12):e32007 [FREE Full text] [doi: 10.2196/32007] [Medline: 34874888]
47. Lordan G, Pakrashi D. Make time for physical activity or you may spend more time sick!. Soc Indic Res 2014 Jan 17;119(3):1379-1391. [doi: 10.1007/s11205-013-0545-y]
48. Schuch FB, Vancampfort D, Firth J, Rosenbaum S, Ward PB, Silva ES, et al. Physical activity and incident depression: a meta-analysis of prospective cohort studies. Am J Psychiatry 2018 Jul 01;175(7):631-648. [doi: 10.1176/appi.ajp.2018.17111194] [Medline: 29690792]
49. Wang X, Guo J, Liu H, Zhao T, Li H, Wang T. Impact of social participation types on depression in the elderly in China: an analysis based on counterfactual causal inference. Front Public Health 2022 Apr 1;10:792765 [FREE Full text] [doi: 10.3389/fpubh.2022.792765] [Medline: 35433576]
50. Siv G, Ove H, Eystein S. Active social participation and depression among older persons. J Clin Nurs 2010 Oct;19(s1):151.
51. Luo Y, Wang Z, Zheng X. [Association between the change of daily living activities and symptoms of depression in Chinese middle-aged and elderly people]. Zhonghua Liu Xing Bing Xue Za Zhi 2017 Aug 10;38(8):1055-1059. [doi: 10.3760/cma.j.issn.0254-6450.2017.08.011] [Medline: 28847053]
52. Yi Y, Park Y. Structural equation model of the relationship between functional ability, mental health, and quality of life in older adults living alone. PLoS One 2022 Aug 3;17(8):e0269003 [FREE Full text] [doi: 10.1371/journal.pone.0269003] [Medline: 35921330]
53. Bae S. Autoregressive cross-lagged modelling of the relationship between social activity, depressive symptoms, and cognitive function in Korean elderly. Psychogeriatrics 2021 May 09;21(3):350-358. [doi: 10.1111/psyg.12674] [Medline: 33690949]

Abbreviations

BA: behavioral activation
BP: body pain
CES-D: Center for Epidemiologic Study Depression Scale
CFI: comparative fit index
CHARLS: China Health and Retirement Longitudinal Study
HRQoL: health-related quality of life
LGCM: latent growth curve model
PF: physical function
RCPR: response-conditioned positive reinforcement
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by A Mavragani, T Sanchez; submitted 29.11.22; peer-reviewed by X Hou, T Martino; comments to author 11.01.23; revised version received 29.01.23; accepted 02.02.23; published 23.03.23.

Please cite as:
Gao Y, Jia Z, Zhao L, Han S
The Effect of Activity Participation in Middle-Aged and Older People on the Trajectory of Depression in Later Life: National Cohort Study
JMIR Public Health Surveill 2023;9:e44682
URL: https://publichealth.jmir.org/2023/1/e44682
doi:10.2196/44682
PMID:36951932

Original Paper

Comorbidity Differences by Trajectory Groups as a Reference for Identifying Patients at Risk for Late Mortality in Childhood Cancer Survivors: Longitudinal National Cohort Study

Hyery Kim[1*], MD, PhD; Hae Reong Kim[2*], PhD; Sung Han Kang[1], MD; Kyung-Nam Koh[1], MD, PhD; Ho Joon Im[1], MD, PhD; Yu Rang Park[2], PhD

[1]Division of Pediatric Hematology and Oncology, Department of Pediatrics, Asan Medical Center Children's Hospital, University of Ulsan College of Medicine, Seoul, Republic of Korea

[2]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics
Yonsei University College of Medicine
50-1 Yonsei-ro
Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2228 2493
Email: yurangpark@yuhs.ac

Abstract

Background: Childhood cancer has a high long-term morbidity and mortality rate. Five years after the initial cancer diagnosis, approximately two-thirds of childhood cancer survivors experience at least one late complication, with one-quarter experiencing severe, life-threatening complications. Chronic health conditions can impact survivors' life planning and daily activities, reducing their health-related quality of life. Comprehensive and longitudinal data are required for investigations of national claims data.

Objective: This study aimed to address clinical and health policy interventions and improved survival rates. A comprehensive categorization of the long-term morbidities associated with childhood cancer survivorship is required. We analyzed the trajectory groups associated with long-term mortality among childhood cancer survivors.

Methods: We collected data from a nationwide claims database of the entire Korean population. Between 2003 and 2007, patients diagnosed with and treated for cancer before the age of 20 years were included. With 8119 patients who survived >10 years, 3 trajectory groups were classified according to yearly changes in the number of diagnoses (the lowest in group 1 and the highest in group 3).

Results: The patterns of most comorbidities and survival rates differed significantly between the trajectory groups. Group 3 had a higher rate of mental and behavioral disorders, neoplasms, and blood organ diseases than the other two groups. Furthermore, there was a difference in the number of diagnoses by trajectory groups over the entire decade, and the disparity increased as the survival period increased. If a patient received more than four diagnoses, especially after the fourth year, the patient was likely to be assigned to group 3, which had the worst prognosis. Group 1 had the highest overall survival rate, and group 3 had the lowest (P<.001). Group 3 had the highest hazard ratio of 4.37 (95% CI 2.57-7.42; P<.001) in a multivariate analysis of late mortality.

Conclusions: Our findings show that the pattern of comorbidities differed significantly among trajectory groups for late death, which could help physicians identify childhood cancer survivors at risk for late mortality. Patients with neoplasms, blood organ diseases, or mental and behavioral disorders should be identified as having an increased risk of late mortality. Furthermore, vigilance and prompt action are essential to mitigate the potential consequences of a child cancer survivor receiving four or more diagnoses within a year.

(JMIR Public Health Surveill 2023;9:e41203) doi:10.2196/41203

KEYWORDS

cancer survivor; childhood cancer; mortality; morbidity; survival; cancer; children; pediatrics

Introduction

Childhood cancer survival rates have increased significantly over the last 4 decades [1]. Most patients survive for a long time, with over 500,000 in the United States and 25,000 in Korea [2,3]. However, prolonged survival after treatment is associated with significant long-term morbidity and premature mortality [4]. Approximately two-thirds of childhood cancer survivors develop at least one late complication that occurs or persists 5 years after the initial cancer diagnosis, with one-quarter developing severe, life-threatening late complications [5,6] that have significantly improved over the past 4 decades [1]. The Childhood Cancer Survivor Study (CCSS) found that 54% of long-term childhood cancer survivors developed at least one severe, disabling, life-threatening, or fatal chronic health condition by the age of 50 years, which is five times higher than that of their healthy peers [7]. Chronic health conditions, such as heart failure, secondary neoplasms, and pulmonary dysfunction can be life-threatening, and other health conditions, such as infertility and hearing impairment, can affect survivors' life planning and daily activities, decreasing their health-related quality of life (HRQoL) [8].

Therefore, monitoring complications and assessing long-term health risks in pediatric and adolescent cancer survivors is essential. In long-term survivors, early intervention may reduce the risk of morbidity due to coexisting conditions. Numerous research groups on a global scale monitor the chronic health condition of childhood cancer survivors, manage their HRQoL [8-11], and develop follow-up strategies using web-based questionnaires [12].

Although previous cohort studies of pediatric cancer survivors elaborated on long-term comorbidities and chronic health conditions, data collection was limited [9,13,14]. First, these studies mainly relied on self-reported questionnaire data and, therefore, identified false-negative outcome events, especially in minor or rare complications that survivors do not report. Second, unlike Korea's single-payer health care system, health care providers in other countries are subject to changes in individual patients, making the collection of complete medical records difficult [15]. Third, existing cohorts have limitations covering a long period or accurately representing the general population. Lastly, previous studies primarily described the first occurrence of complications and did not show morbidity trends over time.

The Korean National Health Insurance Service (NHIS) is South Korea's universal health insurance coverage system [16]. It includes detailed treatment practices and prescriptions based on the fee-for-service payment model and the health information of all South Korean citizens enrolled in national medical insurance [16]. Patients with cancer can receive special government support for medical expenses in South Korea; therefore, once diagnosed, patients with cancer must be registered with the NHIS with a special code and cancer diagnosis so that we can identify their claims data [17]. The NHIS has constructed and operated large-capacity health care big data systems, providing researchers with a wide range of secure health care data for research purposes. The NHIS collects longitudinal patient data that can be used for longitudinal temporal studies, such as the analysis of long-term cancer survivors. Consequently, active longitudinal and epidemiological studies of patients with cancer are being conducted using NHIS data [18-20].

We analyzed long-term health-related conditions using the nationwide claims database of Korea, which contains medical records for approximately 97% of the South Korean population. This study aimed to investigate childhood cancer survivors' temporal diagnostic patterns and assess late mortality using a trajectory group-based method.

Methods

Data Source

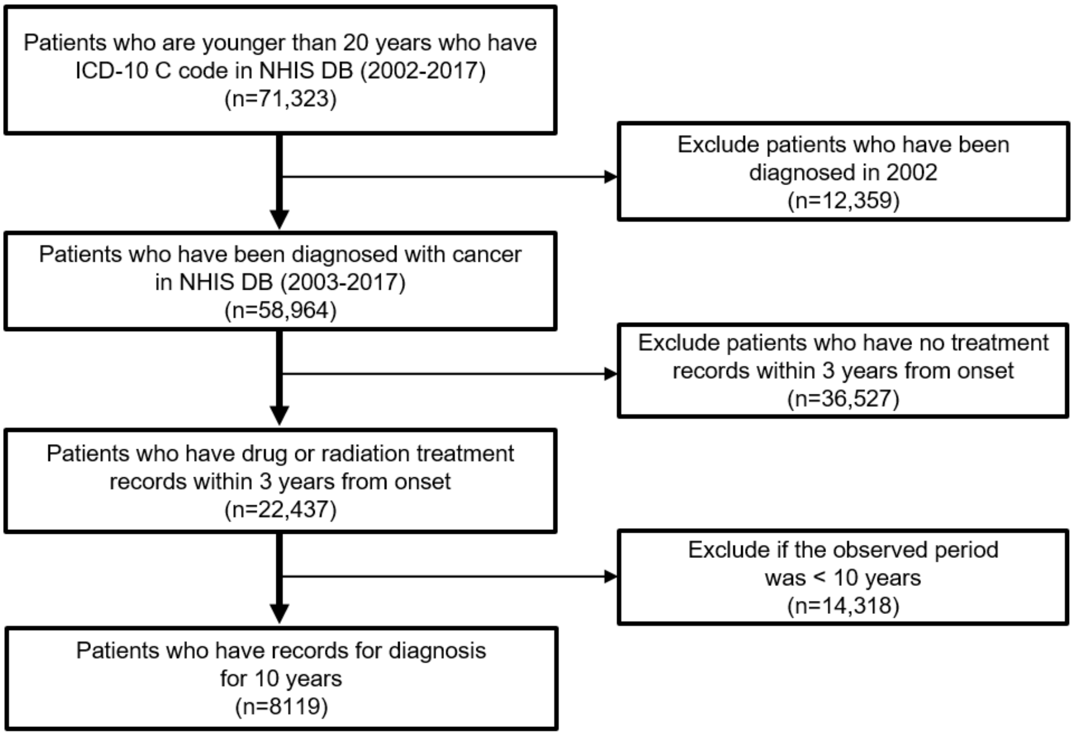
We obtained data from the NHIS, a government insurer in Korea that provides universal health insurance to approximately 97% of Koreans [17]. The NHIS is a universal health insurance system that includes detailed treatment practices and prescriptions based on the fee-for-service payment model and the health information of all Korean citizens who have signed up for national medical insurance [16]. Data included public data on health care use, such as disease diagnoses, drug prescriptions, and procedures. National health examination results for the entire Korean population, including smoking habits, alcohol consumption, physical measurements, body measurements, and demographics as well as socioeconomic variables such as age, sex, income rank, household location, disability, and mortality. Diagnoses were coded in accordance with the *International Statistical Classification of Disease, Tenth Revision* (*ICD-10*) within health insurance claims data [16,21]. The NHIS data were linked to death statistics from Statistics Korea based on the resident numbers of patients [22]. We collected death information from all childhood cancer survivors for this study. Public users can access all data from NHIS and Korean Statistics Information Services for research purposes.

Study Population

We identified 71,323 patients from the NHIS database who were younger than 20 years and had any cancer codes (based on the *ICD-10*) between 2002 and 2017. Clinically representative diagnostic groups were defined for analysis (Multimedia Appendix 1). There were 549 claim codes for treatment (ie, chemotherapy, radiotherapy, hematopoietic stem cell transplantation [HCT]; Multimedia Appendix 2). Surgery was not considered treatment because of this ambiguity. The time of earliest cancer diagnosis was defined as the onset of the earliest diagnosis in patients who were prescribed any treatment within 1 month of the initial diagnosis.

A total of 12,359 patients diagnosed in 2002 and 36,527 patients without treatment records within 3 years from the onset date were excluded (Figure 1). Subsequently, 8119 patients who had survived for >10 years were selected for the final analysis.

Figure 1. Patient selection flowchart. DB: database; ICD-10: International Classification of Diseases, 10th Revision; NHIS: National Health Insurance Service.



Analyzing Trajectory Groups

For trajectory analysis, we used each patient's total number of diagnoses per year as an input variable. We used a square root transformation to calculate the total number of diagnoses. Age, sex, year of onset and diagnosis, duration of treatment exposure, total number and type of treatment, type of care, and socioeconomic status were used to identify differences between the trajectory groups. The most significant diagnosis was entered into the "Main Diagnosis" field, while other diagnoses were entered into the "Subdiagnosis" field. Only those entered in the "Main Diagnosis" section were used for analysis. We pooled all main diagnostic codes from each patient's hospital visits during the survival period and analyzed the diagnoses by excluding primary cancer as a comorbidity.

We used group-based latent trajectory modeling (GBTM) on longitudinal data of childhood cancer survivors to identify homogenous subgroups of a population with similar patterns of outcome change over time [23]. Response values were estimated using finite mixture models in GBTM. The GBTM was fitted with a censored normal distribution, assuming that the total number of diagnoses per year in the claim data follows a normal distribution with a limited scale [24]. The following criteria were used for selecting the optimal number of groups: the lowest value in the Bayesian information criteria, the average posterior probability of group assignment (≥0.7), and group size such that ≥5% of the study sample was assigned to one trajectory group [25]. The trajectory group was assessed using the Proc Traj procedure in SAS 9.4 [26] (SAS Institute).

Statistical Analysis

Trajectory group characteristics are presented as means with SDs for normally distributed characteristics or as medians with IQRs for skewed distribution. The chi-square tests for categorical variables and one-way ANOVA for continuous variables were used to compare the characteristics of the trajectory groups. Kaplan-Meier methods and log-rank tests were used for survival analysis. Cox proportional hazard models with 95% CIs were estimated to investigate differences [27]. The absolute standardized difference (ASD) is a standard measure for comparing values before and after propensity matching. However, pairwise comparisons were performed for the three study groups to identify intergroup differences based on the maximum value. According to the guidelines, the two groups are different from each other by 0.1 [28]. We used the Stddiff package in R (version 3.5.3; R Foundation for Statistical Computing) to calculate ASD. *P* values <.05 on both sides were considered statistically significant.

Ethics Approval

This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the institutional review board of the Asan Medical Center in Seoul, Republic of Korea (IRB 2019-0915). All KNHIS patient records were anonymized to ensure patient confidentiality. Because the analysis used anonymous data, the institutional review board waived the requirement for informed consent.

Results

Trajectory Groups for Pediatric Patients With Cancer

In the 10-year survivor cohort (n=8119; Multimedia Appendices 3 and 4), 3 trajectory groups were selected with the best fit. The first group (n=3104, 38.2%) had relatively smaller changes in the number of diagnoses over time and was thus referred to as the "relatively stable group (group 1)" (red line; Multimedia Appendix 4). The second group (n=4148, 51.1%) was referred to as the "moderately decreasing group (group 2)" (green line), and the remaining population (n=867, 10.7%) was referred to as the "rapidly decreasing group (group 3)" (blue line), which had the most diagnoses and a rapidly decreasing trend over time.

Characteristics by Trajectory Groups

The median age at onset significantly decreased in groups 1 to 3 (*P*<.001; Table 1). Lymphoid leukemia was found in 158 of 867 (18.2%) patients of group 3 patients and 838 of 3104 (20.2%) patients in group 2; however, only 264 of 3104 (8.5%) patients of group 1 (*P*<.001). The top five onset diagnoses were "secondary and unspecified malignant neoplasm of lymph nodes." The median duration of treatment exposure and the total number of prescriptions increased significantly in groups 1 to 3 (*P*<.001).

The use of radiotherapy and the proportion of patients with any HCT increased significantly from group 1 to 3 (radiotherapy: 52.2%, 70.4%, 76.2%; *P*<.001, HCT: 4.1%, 11.2%, 21.6%; *P*<.001, respectively).

In terms of the onset year, 46.1% of group 1 and 21.5% of group 3 were diagnosed between 2003 and 2004 (*P*<.001). In terms of the temporal patterns of each categorical variable, chemotherapy use decreased, while chemotherapy with radiotherapy or HCT steadily increased (Multimedia Appendix 5).

Table 1. Characteristics by trajectory groups for 10 years of follow-up with childhood cancer survivors.

Category	Group 1 (n=3104)	Group 2 (n=4148)	Group 3 (n=867)	Total (n=8119)	*P* value[a]
Onset age (years), median (IQR)	12 (7.0-16.0)	7 (3.0-13.0)	3 (1.0-10.0)	9 (4.0-14.0)	<.001
Sex, n (%)					.04
Male	1818 (58.6)	2308 (55.6)	496 (57.2)	4622 (56.9)	
Female	1286 (41.4)	1840 (44.4)	371 (42.8)	3497 (43.1)	
Top 5 onset diagnoses in total, n (%)					
Lymphoid leukemia	264 (8.5)	838 (20.2)	158 (18.2)	1260 (15.5)	.05
CNS[b] tumor	363 (11.7)	501 (12.1)	161 (18.6)	1025 (12.6)	.29
NHL[c]	347 (11.2)	402 (9.7)	76 (8.8)	825 (10.2)	.85
Myeloid leukemia	208 (6.7)	339 (8.2)	56 (6.5)	603 (7.4)	.88
Secondary and unspecified malignant neoplasm of the lymph nodes	296 (9.5)	239 (5.8)	33 (3.8)	568 (7.0)	.25
Claims records, n (%)					<.001
Inpatient	17,923 (5.0)	59,236 (6.0)	26,007 (6.1)	103,166 (5.8)	
Outpatient	342,558 (94.6)	932,153 (93.7)	399,001 (93.4)	1,673,712 (93.8)	
Emergency	1579 (0.4)	3106 (0.3)	2190 (0.5)	6875 (0.4)	
Social economy status[d], median (IQR)	14 (9.0-17.0)	14 (9.0-17.0)	13 (9.0-16.0)	14 (9.0-17.0)	<.001
Treatment exposure[e] (days), median (IQR)	249 (0-1027)	935 (240-1835)	1506 (464-3266)	651 (120-1655)	<.001
Total number of prescriptions for treatment, median (IQR)	7 (2-43)	45 (6-114)	74 (15-168)	27 (3-87)	<.001
Types of treatment, n (%)					
CT[f] + RT[g]	1268 (40.9)	2278 (54.9)	450 (51.9)	3996 (49.2)	.11
CT only	1479 (47.6)	1223 (29.5)	197 (22.7)	2899 (35.7)	<.001
RT only	231 (7.4)	178 (4.3)	25 (2.9)	434 (5.3)	.32
CT + RT + allo[h] HCT[i]	69 (2.2)	288 (6.9)	94 (10.8)	451 (5.6)	.05
CT + RT + auto[j] HCT	53 (1.7)	175 (4.2)	92 (10.6)	320 (3.9)	.02
CT + allo HCT	2 (0.1)	0 (0.0)	0 (0.0)	2 (0.02)	.91
Allo HCT only	2 (0.1)	0 (0.0)	1 (0.1)	3 (0.04)	.95
Summary of treatment, n (%)					
Any RT	1621 (52.2)	2919 (70.4)	661 (76.2)	5201 (64.1)	<.001
Any HCT	126 (4.1)	463 (11.2)	187 (21.6)	776 (9.6)	<.001
Year of onset, n (%)					<.001
2003-2004	1431 (46.1)	1407 (33.9)	186 (21.5)	3024 (37.2)	
2005-2006	937 (30.2)	1437 (34.6)	347 (40.0)	2721 (33.5)	
2007-2008	736 (23.7)	1304 (31.4)	334 (38.5)	2374 (29.2)	

[a] Calculated using the chi-square test (categorical variables) or one-way ANOVA (numerical variables).

[b] CNS: central nervous system.

[c] NHL: non-Hodgkin lymphoma.

[d] Socioeconomic status refers to the 20th quartile by insurance fee, and 1255 patients did not have the quartile information.

[e] Treatment exposure was defined as the duration from the first date to the last date when any chemotherapeutic agents or radiotherapy codes were prescribed.

[f] CT: chemotherapy.

[g] RT: radiotherapy.

[h] Allo: allogeneic.

[i] HCT: hematopoietic stem cell transplantation.

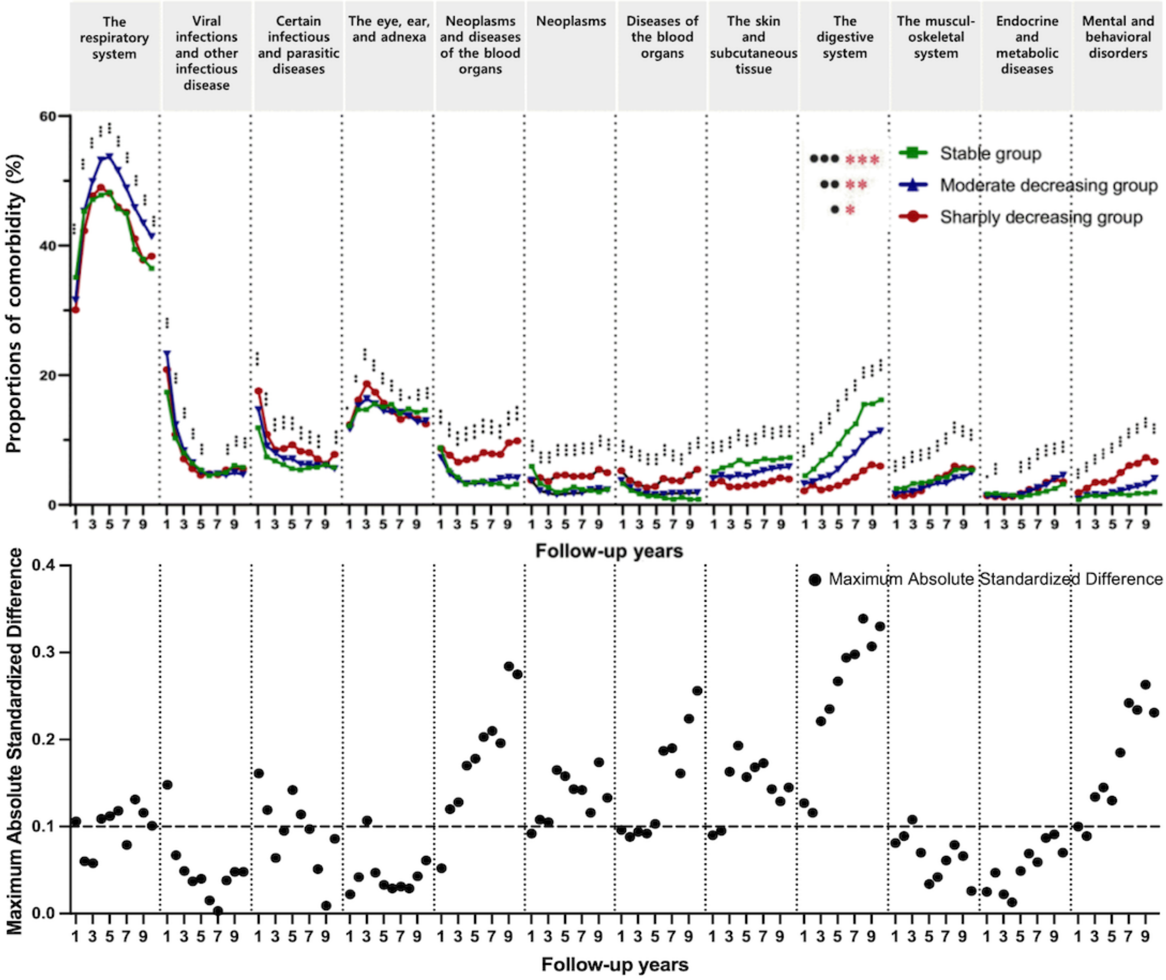
[j] Auto: autologous.

Comorbidity by Trajectory Groups

Figure 2 shows the proportion of each diagnosis in relation to the total yearly numbers of the top 10 comorbidities (Multimedia Appendix 6). The annual incidence showed three major trends: early and decreased midterm surge and continuously decreasing and growing proportions. Infection-related disorders (eg, certain infections, parasitic diseases, viral infections, and other infectious diseases) have declined since the beginning (Figure 2 and Multimedia Appendix 7). Although respiratory disorders were the most common comorbidities at all times, the proportion of diseases in the eye, ear, adnexa, or respiratory system showed a midterm surge at approximately 3-5 years after diagnosis. However, endocrine or metabolic, mental or behavioral, digestive, musculoskeletal, and skin-related disorders have shown increasing trends over the last 10 years. The proportions of comorbidities were statistically different among the three groups at most follow-up periods in the top 10 comorbidities. Specifically, mental and behavioral disorders predominantly increased in group 3 within the 10 years and were statistically significant among trajectory groups for 10 years. Furthermore, in group 3, neoplasms or diseases of the blood organs (D00-D89) showed a unique trend, decreasing in the first 4 years and increasing thereafter. In groups 1 and 2, there was a decreasing pattern, and the differences were statistically significant at all times. When the D code was divided into neoplasm (D00-D49) and disease of the blood organs (D50-D89) for statistical analysis, group 3 was statistically substantially higher during the entire period.

When ASD was used to evaluate the judgment, it was more sensitive than the chi-square test results. By the follow-up year, ASD of mental and behavioral disorders, neoplasms, diseases of the blood organs, skin and subcutaneous tissue, and the digestive system showed increasing trends above the threshold of 0.1, demonstrating *overt* differences in the incidence of comorbidities between each group (Figure 2 and Multimedia Appendix 7). Over 10 years, the prevalence of mental and behavioral disorders, neoplasms, and diseases of the blood organs (also known as neoplasm and blood organ disease) gradually increased in group 3. The prevalence of ASD in these diseases was relatively high in groups 3 and 1. For the past 10 years, skin, subcutaneous tissue, and digestive system diseases increased, particularly in group 1. There was a significant difference between groups 3 and 1.

Figure 2. The prevalence of the top 10 comorbidities of pediatric cancer survivors by follow-up year and a plot of the maximum absolute standardized difference between two groups by follow-up year in the top 10 comorbidities, presented pairwise. Pairwise *P* values for two of the three groups are presented as follows: *<.05, **<.01, and ***<.001. No *P* values were corrected.

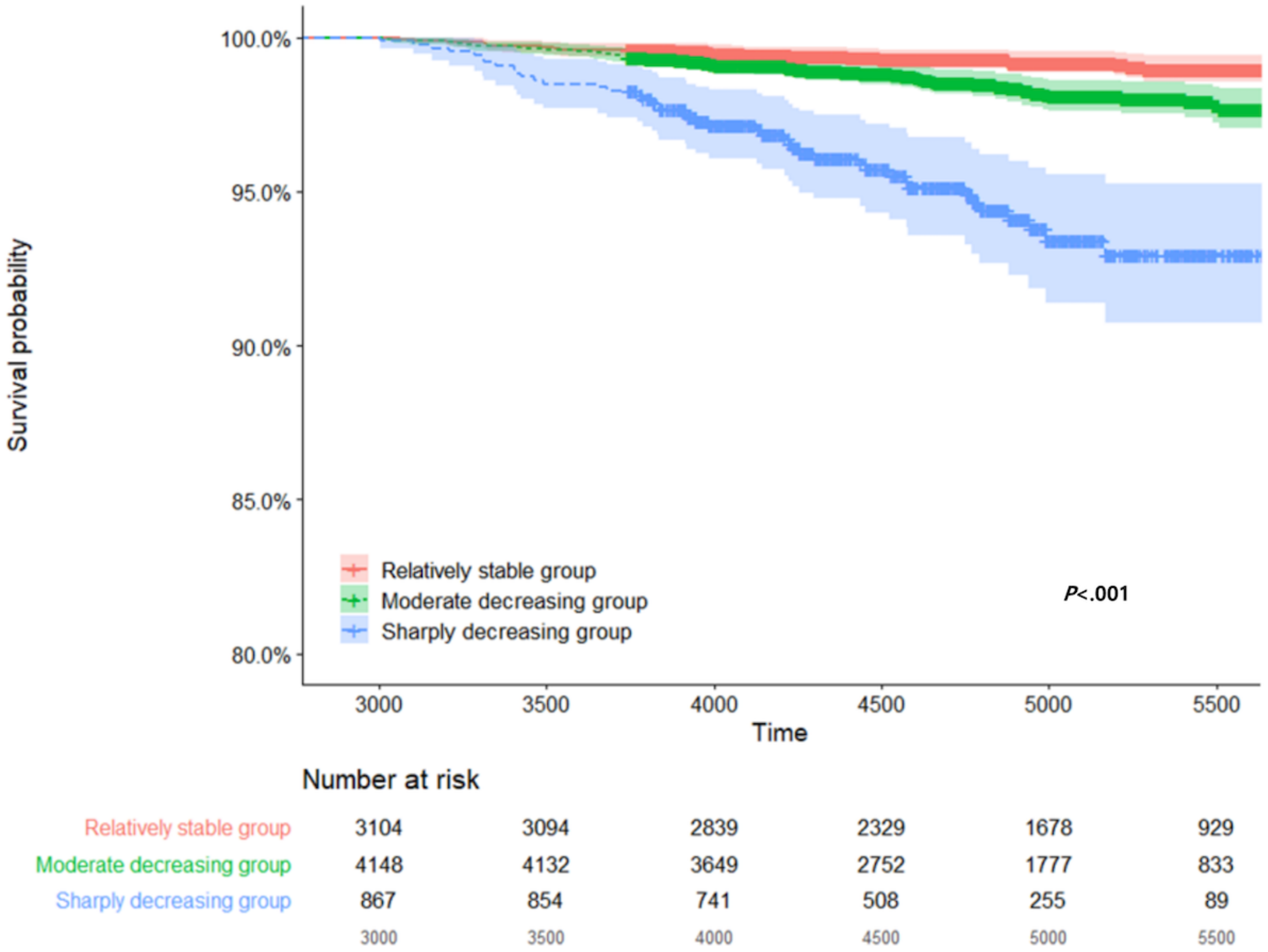


Survival Outcome by Trajectory Groups

In terms of overall survival, it was confirmed that the mortality rate increased more rapidly in group 3 than in group 1. In particular, it was observed that the difference in mortality rate increased more rapidly from 3700 days in group 3 than in groups 1 and 2, which was a statistically significant result (*P*<.001; Figure 3). Multivariate analysis was performed after adjusting for sex, age at onset (5 years), and treatment exposure duration.

Group 3 had the highest risk as determined by the trajectory method. The hazard ratio (HR) was 4.37 (95% CI 2.57-7.42; *P*<.001; Table 2). Group 2 had an HR of 1.68 (95% CI 1.05-2.7; *P*=.03) compared with group 1. The age was calculated by dividing it by 5 years. As a result, it was confirmed that the patient's age when compared to 5 years of age increased the risk of overall survival. The HR for overall survival was 1.53 (95% CI 0.91-2.59; *P*=.11) when compared to the group younger than 5 years (reference group). However, it was not statistically significant. The HRs or overall survival were 2.68 (95% CI 1.64-4.39; *P*<.001) and 2.57 (95% CI 1.57-4.22; *P*<.001), respectively, and had a significant effect on survival when compared to the reference group. The result of the multivariate survival analysis was adjusted for sex and treatment duration. Central nervous system tumor was the most common cause of death at the time of death (Multimedia Appendix 8).

Figure 3. Kaplan-Meier plots for overall survival by trajectory groups for 10-year follow-up. Overall survival was highest in group 1 and lowest in group 3 (*P*<.001).



Number at risk

	3000	3500	4000	4500	5000	5500
Relatively stable group	3104	3094	2839	2329	1678	929
Moderate decreasing group	4148	4132	3649	2752	1777	833
Sharply decreasing group	867	854	741	508	255	89

Table 2. HRs for all-cause death by trajectory groups.

Cohort and variables	Univariate		Multivariate[a]	
	HR[b] (95% CI)	*P* value	HR (95% CI)	*P* value
10-year				
Group 1	1 (reference)	N/A[c]	1 (reference)	N/A
Group 2	2.11 (1.33-3.36)	.001	1.68 (1.05-2.7)	.03
Group 3	7.32 (4.47-12.01)	<.001	4.37 (2.57-7.42)	<.001
Onset age (<5 years old)	1 (reference)	N/A	1 (reference)	N/A
Onset age (≥5 years and <10 years)	1.2 (0.72-2.02)	.48	1.53 (0.91-2.59)	.11
Onset age (≥10 years and <15 years)	1.66 (1.03-2.69)	.04	2.68 (1.64-4.39)	<.001
Onset age (≥15 years)	1.67 (1.04-2.69)	.04	2.57 (1.57-4.22)	<.001
Sex (reference: male)	0.57 (0.39-0.82)	.003	N/A	N/A
Treatment exposure duration	1.001 (1.001-1.001)	<.001	N/A	N/A

[a] Adjusted for sex and duration of treatment.

[b] HR: hazard ratio.

[c] N/A: not applicable.

Discussion

This study evaluated the 10-year trajectories of childhood cancer survivors nationwide. Using group-based trajectory analysis, we analyzed all hospital visits due to any health impairments. We demonstrated that the risk groups for long-term mortality could be well classified using the entire spectrum of diagnoses as a surrogate for the cumulative burden of late complications. Our study found that the pattern of cumulative comorbidities during the survival periods significantly differed according to the trajectory group; physicians could use this to identify patients at risk for late mortality in childhood cancer survivors.

The NHIS database was used in this study, which can complement the limitations of previous studies on cancer survivors. The NHIS database is one of the world's largest claims data sets. It centralizes all claims data from the Korean population regardless of age or region, distinguishing it from the Medicare and Medicaid programs in the United States. This reduces the possibility of selection bias [29]. This universal health care provider system facilitates the collection of populationwide longitudinal data, which makes it helpful in analyzing the long-term outcomes of childhood cancer survivors.

The St Jude Lifetime Cohort Study (SJLIFE) recently reported the cumulative disease burden of 168 chronic health conditions grouped into 48 condition-specific categories for childhood cancer survivors [9]. SJLIFE reported that the survivors advancing age changed the landscape of the dominant comorbid conditions. In this report, the cumulative burden of chronic health conditions per individual was differentiated by the initial cancer types of the survivors. This report from SJLIFE could provide general health practitioners and clinical investigators with references for the expected cumulative burden of comorbidities in specific survivor groups.

Although late mortality is as important as a chronic disease in childhood cancer survivors, there are no direct references to late mortality. Additionally, psychiatric diseases were not analyzed.

Our findings suggest that the pattern of comorbidities can predict high-risk groups for late mortality. Childhood cancer survivors are at risk for adverse psychological outcomes according to the CCSS study of 6199 patients, depression had a prevalence of 11.4% after a median follow-up duration of 7.8 years [30]. Survivors have an 80% higher risk of developing clinically significant mental health symptoms than their siblings [31].

According to the results of this study, high-risk groups with low survival rates are likely to include childhood cancer survivors who receive more than a certain number of diagnoses annually during long-term follow-up. Our study showed that mental and behavioral disorders could increase the risk of late mortality due to other severe late complications. Therefore, survivors who often visit hospitals for mental or behavioral disorders should be closely monitored for late mortality. In addition, neoplasms or blood organ diseases were common in group 3, and these conditions were associated with abnormal hematologic findings unrelated to the malignancy. Therefore, such blood organ diseases might result from other severe health conditions or complications of ongoing treatment.

Furthermore, there was a difference in the number of diagnoses by trajectory group not only in the first year but throughout the decade (Multimedia Appendix 4). As the duration of survival increased, so did the disparity. For the first 1-3 years, there was no significant difference in the number of average diagnoses between groups; however, if the patient received more than four diagnoses, especially after the fourth year, the patient was strongly suspected to be in group 3, which had the worst prognosis. Consequently, if a childhood cancer survivor receives four or more diagnoses in a year, close monitoring and early intervention for complications are necessary.

Factors associated with the high-risk trajectory group in this study were younger age at primary cancer diagnosis, longer treatment duration, and a recent treatment period with radiotherapy or HCT. The duration of treatment and the number of treatment codes were proportional to the overall loading of the treatment; therefore, a higher loading of the treatments was invariably related to a higher risk of mortality.

Reports on the age at primary cancer diagnosis and its relationship to subsequent cancers are contradictory [32-34]. Generally, younger age at the time of radiotherapy or HCT is associated with an increased risk of complications and subsequent cancers [35,36]. Particularly, testicular cancer and Hodgkin disease are risk factors for developing cardiovascular disease, secondary malignancies, and death from noncancerous causes at a younger age [37].

Treatment timing distinguished the trajectory groups in this study. Other studies have found that previously treated patients had a higher risk of long-term complications and mortality than those receiving treatment [14,38,39]. Contrary to previous studies, our study found that more recently treated survivors had poorer survival outcomes than those previously treated, possibly because our study included data from 2003 rather than the 1990s. Moreover, since 2003, radiotherapy and HCT have been increasingly used with other treatment modalities. The increasing use of alternative donors [40] and the introduction of tandem autologous HCT in patients with high-risk neuroblastoma and medulloblastoma may have contributed to the recent increase in HCT [41,42]. Therefore, more recent survivors are likely to have received intensive treatment.

Although this study showed a novel approach to examining nationwide longitudinal data using trajectory analysis, it had several limitations. First, because this data set was established for claims and reimbursements, there may be differences between the actual disease and the claimed diagnosis. Second, this study did not include a healthy control group. Third, it could be argued that the difference in survival between the trajectory groups is due to the high frequency of recurrence or secondary cancer in the high-risk group rather than comorbidities. Finally, claims databases are incapable of distinguishing between cancer remission and recurrence because the recurrence diagnosis and treatment discontinuation needs to be clearly defined. This may have led to skewed results in identifying trajectory groups with increased morbidity and mortality due to protracted chemotherapy-related comorbidities.

In conclusion, our study is the first to show that the risk groups for long-term mortality in childhood cancer survivors can be grouped using the entire spectrum of diagnosed diseases as a surrogate for the cumulative burden of late complications. We demonstrated that clinically severe health conditions and hospital visits due to health impairments could significantly impact long-term survival. Even without a primary cancer diagnosis, medical practitioners should pay attention to the subpopulation of patients with increased or specific comorbidities.

Acknowledgments

This work was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (grant HI21C2046), and a grant (2019IT0005) from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea.

Data Availability

The data are not publicly available due to privacy or ethical restrictions.

Authors' Contributions

All authors contributed to the conception and design of the study. Material preparation, data collection, and analysis were performed by HK, HRK, and YRP, respectively. The first draft of the manuscript was written by HK and HRK, and all authors commented on previous versions of the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Representative diagnostic groups for the International Classification of Diseases codes.

[DOCX File , 15 KB - publichealth_v9i1e41203_app1.docx]

Multimedia Appendix 2

List of claiming codes in the Korean National Health Insurance Program for cancer treatments.

[DOCX File , 25 KB - publichealth_v9i1e41203_app2.docx]

Multimedia Appendix 3

Tabulated Bayesian information criteria (BIC) and 2ΔBIC.

[DOCX File , 15 KB - publichealth_v9i1e41203_app3.docx]

Multimedia Appendix 4

Group-based trajectory graphs based on the transformed numbers of diagnosis by follow-up year.

[DOCX File , 153 KB - publichealth_v9i1e41203_app4.docx]

Multimedia Appendix 5

Comparison of onset diagnosis and the types of treatment every year after diagnosis.

[DOCX File , 22 KB - publichealth_v9i1e41203_app5.docx]

Multimedia Appendix 6

International Classification of Diseases codes for the top 10 comorbidities.

[DOCX File , 16 KB - publichealth_v9i1e41203_app6.docx]

Multimedia Appendix 7

Absolute standardized differences between comorbidities of the three groups by follow-up year (for 10 years).

[DOCX File , 50 KB - publichealth_v9i1e41203_app7.docx]

Multimedia Appendix 8

Diagnosis at the time of death.

[DOCX File , 20 KB - publichealth_v9i1e41203_app8.docx]

References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014;64(1):9-29. [doi: 10.3322/caac.21208] [Medline: 24399786]
2. Robison LL, Hudson MM. Survivors of childhood and adolescent cancer: life-long risks and responsibilities. Nat Rev Cancer 2014 Jan;14(1):61-70 [FREE Full text] [doi: 10.1038/nrc3634] [Medline: 24304873]
3. Ghim TT. Time to establish multidisciplinary childhood cancer survivorship programs in Korea. Korean J Hematol 2010 Jun;45(2):84-87 [FREE Full text] [doi: 10.5045/kjh.2010.45.2.84] [Medline: 21120183]
4. Mertens AC, Liu Q, Neglia JP, Wasilewski K, Leisenring W, Armstrong GT, et al. Cause-specific late mortality among 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. J Natl Cancer Inst 2008 Oct 01;100(19):1368-1379 [FREE Full text] [doi: 10.1093/jnci/djn310] [Medline: 18812549]
5. Oeffinger KC, Eshelman DA, Tomlinson GE, Buchanan GR, Foster BM. Grading of late effects in young adult survivors of childhood cancer followed in an ambulatory adult setting. Cancer 2000 Apr 01;88(7):1687-1695. [Medline: 10738228]
6. Stevens MC, Mahler H, Parkes S. The health status of adult survivors of cancer in childhood. Eur J Cancer 1998 Apr;34(5):694-698. [doi: 10.1016/s0959-8049(98)00012-4] [Medline: 9713276]
7. Armstrong GT, Kawashima T, Leisenring W, Stratton K, Stovall M, Hudson MM, et al. Aging and risk of severe, disabling, life-threatening, and fatal events in the childhood cancer survivor study. J Clin Oncol 2014 Apr 20;32(12):1218-1227 [FREE Full text] [doi: 10.1200/JCO.2013.51.1055] [Medline: 24638000]
8. Calaminus G, Baust K, Berger C, Byrne J, Binder H, Casagranda L, et al. Health-related quality of life in European childhood cancer survivors: protocol for a study within PanCareLIFE. JMIR Res Protoc 2021 Jan 25;10(1):e21851 [FREE Full text] [doi: 10.2196/21851] [Medline: 33492237]
9. Bhakta N, Liu Q, Ness KK, Baassiri M, Eissa H, Yeo F, et al. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). Lancet 2017 Dec 09;390(10112):2569-2582 [FREE Full text] [doi: 10.1016/S0140-6736(17)31610-0] [Medline: 28890157]
10. Fardell JE, Wakefield CE, De Abreu Lourenco R, Signorelli C, McCarthy M, McLoone J, ANZCHOG Survivorship Group. Long-term health-related quality of life in young childhood cancer survivors and their parents. Pediatr Blood Cancer 2021 Dec;68(12):e29398. [doi: 10.1002/pbc.29398] [Medline: 34606168]
11. Dixon SB, Li N, Yasui Y, Bhatia S, Casillas JN, Gibson TM, et al. Racial and ethnic disparities in neurocognitive, emotional, and quality-of-life outcomes in survivors of childhood cancer: A report from the Childhood Cancer Survivor Study. Cancer 2019 Oct 15;125(20):3666-3677 [FREE Full text] [doi: 10.1002/cncr.32370] [Medline: 31291014]
12. Kilsdonk E, van Dulmen-den Broeder E, van der Pal HJ, Hollema N, Kremer LC, van den Heuvel-Eibrink MM, et al. Effect of web-based versus paper-based questionnaires and follow-up strategies on participation rates of Dutch childhood cancer survivors: a randomized controlled trial. JMIR Cancer 2015 Nov 24;1(2):e11 [FREE Full text] [doi: 10.2196/cancer.3905] [Medline: 28410161]
13. Hudson MM, Oeffinger KC, Jones K, Brinkman TM, Krull KR, Mulrooney DA, et al. Age-dependent changes in health status in the Childhood Cancer Survivor cohort. J Clin Oncol 2015 Feb 10;33(5):479-491 [FREE Full text] [doi: 10.1200/JCO.2014.57.4863] [Medline: 25547510]
14. Anderson C, Nichols HB. Trends in late mortality among adolescent and young adult cancer survivors. J Natl Cancer Inst 2020 Oct 01;112(10):994-1002 [FREE Full text] [doi: 10.1093/jnci/djaa014] [Medline: 32123906]
15. Leisenring WM, Mertens AC, Armstrong GT, Stovall MA, Neglia JP, Lanctot JQ, et al. Pediatric cancer survivorship research: experience of the Childhood Cancer Survivor Study. J Clin Oncol 2009 May 10;27(14):2319-2327 [FREE Full text] [doi: 10.1200/JCO.2008.21.1813] [Medline: 19364957]
16. Kim MK, Han K, Lee S. Current trends of big data research using the Korean National Health Information Database. Diabetes Metab J 2022 Jul;46(4):552-563 [FREE Full text] [doi: 10.4093/dmj.2022.0193] [Medline: 35929173]
17. Shin DW, Lee J, Jung JH, Han K, Kim SY, Choi KS, et al. Disparities in cervical cancer screening among women with disabilities: a national database study in South Korea. J Clin Oncol 2018 Sep 20;36(27):2778-2786. [doi: 10.1200/JCO.2018.77.7912] [Medline: 30074846]
18. Chung IY, Lee J, Park S, Lee JW, Youn HJ, Hong JH, Study of Multi-disciplinARy Teamwork of breast cancer survivorSHIP (SMARTSHIP) Group. Nationwide analysis of treatment patterns for Korean breast cancer survivors using National Health Insurance Service data. J Korean Med Sci 2018 Oct 29;33(44):e276 [FREE Full text] [doi: 10.3346/jkms.2018.33.e276] [Medline: 30369858]
19. Woo A, Lee SW, Koh HY, Kim MA, Han MY, Yon DK. Incidence of cancer after asthma development: 2 independent population-based cohort studies. J Allergy Clin Immunol 2021 Jan;147(1):135-143. [doi: 10.1016/j.jaci.2020.04.041] [Medline: 32417133]
20. Kim YA, Kang D, Moon H, Sinn D, Kang M, Woo SM, et al. Survival in untreated hepatocellular carcinoma: a national cohort study. PLoS One 2021;16(2):e0246143 [FREE Full text] [doi: 10.1371/journal.pone.0246143] [Medline: 33539397]
21. Kim HK, Song SO, Noh J, Jeong I, Lee B. Data configuration and publication trends for the Korean National Health Insurance and health insurance review and assessment database. Diabetes Metab J 2020 Oct;44(5):671-678 [FREE Full text] [doi: 10.4093/dmj.2020.0207] [Medline: 33115211]

22. Shin H, Lee J, Song J, Lee S, Lee J, Lim B, et al. Cause-of-death statistics in the Republic of Korea, 2014. J Korean Med Assoc 2016;59(3):221. [doi: 10.5124/jkma.2016.59.3.221]
23. Klijn SL, Weijenberg MP, Lemmens P, van den Brandt PA, Lima Passos V. Introducing the fit-criteria assessment plot - a visualisation tool to assist class enumeration in group-based trajectory modelling. Stat Methods Med Res 2017 Oct;26(5):2424-2436. [doi: 10.1177/0962280215598665] [Medline: 26265768]
24. Nguena Nguefack HL, Pagé MG, Katz J, Choinière M, Vanasse A, Dorais M, et al. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. Clin Epidemiol 2020;12:1205-1222 [FREE Full text] [doi: 10.2147/CLEP.S265287] [Medline: 33154677]
25. Xiang X. Seven-year trajectories of depressive symptoms and their predictors among older americans. J Aging Health 2020;32(7-8):795-806. [doi: 10.1177/0898264319852835] [Medline: 31169060]
26. Allen NB, Siddique J, Wilkins JT, Shay C, Lewis CE, Goff DC, et al. Blood pressure trajectories in early adulthood and subclinical atherosclerosis in middle age. JAMA 2014 Feb 05;311(5):490-497 [FREE Full text] [doi: 10.1001/jama.2013.285122] [Medline: 24496536]
27. Mayland CR, Ingarfield K, Rogers SN, Dey P, Thomas S, Waylen A, et al. Disease trajectories, place and mode of death in people with head and neck cancer: findings from the 'Head and Neck 5000' population-based prospective clinical cohort study. Palliat Med 2020 May;34(5):639-650 [FREE Full text] [doi: 10.1177/0269216320904313] [Medline: 32103703]
28. Stuart EA, Lee BK, Leacy FP. Prognostic score-based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. J Clin Epidemiol 2013 Aug;66(8 Suppl):S84-S90.e1 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.01.013] [Medline: 23849158]
29. Choi EK. Cardiovascular research using the Korean National Health Information database. Korean Circ J 2020 Sep;50(9):754-772 [FREE Full text] [doi: 10.4070/kcj.2020.0171] [Medline: 32725984]
30. Tonorezos ES, Ford JS, Wang L, Ness KK, Yasui Y, Leisenring W, et al. Impact of exercise on psychological burden in adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. Cancer 2019 Sep 01;125(17):3059-3067 [FREE Full text] [doi: 10.1002/cncr.32173] [Medline: 31067357]
31. Hudson MM, Mertens AC, Yasui Y, Hobbie W, Chen H, Gurney JG, Childhood Cancer Survivor Study Investigators. Health status of adult long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. JAMA 2003 Sep 24;290(12):1583-1592. [doi: 10.1001/jama.290.12.1583] [Medline: 14506117]
32. Neglia JP, Robison LL, Stovall M, Liu Y, Packer RJ, Hammond S, et al. New primary neoplasms of the central nervous system in survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. J Natl Cancer Inst 2006 Nov 01;98(21):1528-1537. [doi: 10.1093/jnci/djj411] [Medline: 17077355]
33. Henderson TO, Whitton J, Stovall M, Mertens AC, Mitby P, Friedman D, et al. Secondary sarcomas in childhood cancer survivors: a report from the Childhood Cancer Survivor Study. J Natl Cancer Inst 2007 Feb 21;99(4):300-308. [doi: 10.1093/jnci/djk052] [Medline: 17312307]
34. Friedman DL, Whitton J, Leisenring W, Mertens AC, Hammond S, Stovall M, et al. Subsequent neoplasms in 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. J Natl Cancer Inst 2010 Jul 21;102(14):1083-1095 [FREE Full text] [doi: 10.1093/jnci/djq238] [Medline: 20634481]
35. Oeffinger KC, Hudson MM. Long-term complications following childhood and adolescent cancer: foundations for providing risk-based health care for survivors. CA Cancer J Clin 2004;54(4):208-236 [FREE Full text] [doi: 10.3322/canjclin.54.4.208] [Medline: 15253918]
36. Baker KS, Leisenring WM, Goodman PJ, Ermoian RP, Flowers ME, Schoch G, et al. Total body irradiation dose and risk of subsequent neoplasms following allogeneic hematopoietic cell transplantation. Blood 2019 Jun 27;133(26):2790-2799 [FREE Full text] [doi: 10.1182/blood.2018874115] [Medline: 30992266]
37. Woodward E, Jessop M, Glaser A, Stark D. Late effects in survivors of teenage and young adult cancer: does age matter? Ann Oncol 2011 Dec;22(12):2561-2568 [FREE Full text] [doi: 10.1093/annonc/mdr044] [Medline: 21427066]
38. Gibson TM, Mostoufi-Moab S, Stratton KL, Leisenring WM, Barnea D, Chow EJ, et al. Temporal patterns in the risk of chronic health conditions in survivors of childhood cancer diagnosed 1970-99: a report from the Childhood Cancer Survivor Study cohort. Lancet Oncol 2018 Dec;19(12):1590-1601. [doi: 10.1016/S1470-2045(18)30537-0] [Medline: 30416076]
39. Yeh JM, Ward ZJ, Chaudhry A, Liu Q, Yasui Y, Armstrong GT, et al. Life expectancy of adult survivors of childhood cancer over 3 decades. JAMA Oncol 2020 Mar 01;6(3):350-357 [FREE Full text] [doi: 10.1001/jamaoncol.2019.5582] [Medline: 31895405]
40. Annual the number of BMT in Korea 1983~2021. Korean Blood & Marrow Transplantation Nursing Society. URL: http://www.bmtnurse.org/default/mp6/mp6_sub6.php?sub=01
41. Sung KW, Ahn HS, Cho B, Choi Y, Chung NG, Hwang TJ, et al. Efficacy of tandem high-dose chemotherapy and autologous stem cell rescue in patients over 1 year of age with stage 4 neuroblastoma: the Korean Society of Pediatric Hematology-Oncology experience over 6 years (2000-2005). J Korean Med Sci 2010 May;25(5):691-697 [FREE Full text] [doi: 10.3346/jkms.2010.25.5.691] [Medline: 20436703]
42. Sung KW, Lim DH, Son MH, Lee SH, Yoo KH, Koo HH, et al. Reduced-dose craniospinal radiotherapy followed by tandem high-dose chemotherapy and autologous stem cell transplantation in patients with high-risk medulloblastoma. Neuro Oncol 2013 Mar;15(3):352-359 [FREE Full text] [doi: 10.1093/neuonc/nos304] [Medline: 23258845]

Abbreviations

ASD: absolute standardized difference
CCSS: Childhood Cancer Survivor Study
GBTM: group-based latent trajectory modeling
HCT: hematopoietic stem cell transplantation
HR: hazard ratio
HRQoL: health-related quality of life
ICD-10: International Statistical Classification of Disease, Tenth Revision
NHIS: National Health Insurance Service
SJLIFE: St Jude Lifetime Cohort Study

Edited by A Mavragani; submitted 19.07.22; peer-reviewed by HJ Kang, J Finkelstein; comments to author 21.10.22; revised version received 30.11.22; accepted 06.01.23; published 24.03.23.

Please cite as:
Kim H, Kim HR, Kang SH, Koh KN, Im HJ, Park YR
Comorbidity Differences by Trajectory Groups as a Reference for Identifying Patients at Risk for Late Mortality in Childhood Cancer Survivors: Longitudinal National Cohort Study
JMIR Public Health Surveill 2023;9:e41203
URL: https://publichealth.jmir.org/2023/1/e41203
doi:10.2196/41203
PMID:36754630

Original Paper

Investigation of the Relationship Between Psychiatry Visit and Suicide After Deliberate Self-harm: Longitudinal National Cohort Study

Hye Hyeon Kim[1*], PhD; Chanyoung Ko[1*], MD; Ji Ae Park[1], MS; In Han Song[2,3], PhD; Yu Rang Park[1], PhD

[1] Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea
[2] Health & Mental Health Lab, Yonsei University Graduate School of Social Welfare, Seoul, Republic of Korea
[3] Institute for Convergence Science Academy, Yonsei University, Seoul, Republic of Korea
[*] these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics, Yonsei University College of Medicine
50 Yonsei-Ro, Seodaemun-Gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 2494
Email: YURANGPARK@yuhs.ac

Abstract

Background: Deliberate self-harm (DSH) along with old age, physical disability, and low socioeconomic status are well-known contributors to suicide-related deaths. In recent years, South Korea has the highest suicide death rate among all Organization for Economic Co-operation and Development countries. Owing to the difficulty of accessing data of individuals with DSH behavior who died by suicide, the factors associated with suicide death in these high-risk individuals have not been sufficiently explored. There have been conflicting findings with regard to the relationship between previous psychiatric visits and suicidal death.

Objective: We aimed to address the following 3 questions: Are there considerable differences in demographics, socioeconomic status, and clinical features in individuals who received psychiatric diagnosis (either before DSH or after DSH event) and those who did not? Does receiving a psychiatric diagnosis from the Department of Psychiatry, as opposed to other departments, affect survival? and Which factors related to DSH contribute to deaths by suicide?

Methods: We used the Korean National Health Insurance Service Database to design a cohort of 5640 individuals (3067/5640, 54.38% women) who visited the hospital for DSH (International Classification of Diseases codes X60-X84) between 2002 and 2020. We analyzed whether there were significant differences among subgroups of individuals with DSH behavior based on psychiatric diagnosis status (whether they had received a psychiatric diagnosis, either before or after the DSH event) and the department from which they had received the psychiatric diagnosis. Another main outcome of the study was death by suicide. Cox regression models yielded hazard ratios (HRs) for suicide risk. Patterns were plotted using Kaplan-Meier survival curves.

Results: There were significant differences in all factors including demographic, health-related, socioeconomic, and survival variables among the groups that were classified according to psychiatric diagnosis status ($P<.001$). The group that did not receive a psychiatric diagnosis had the lowest survival rate (867/1064, 81.48%). Analysis drawn using different departments from where the individual had received a psychiatric diagnosis showed statistically significant differences in all features of interest ($P<.001$). The group that had received psychiatric diagnoses from the Department of Psychiatry had the highest survival rate (888/951, 93.4%). These findings were confirmed using the Kaplan-Meier survival curves ($P<.001$). The severity of DSH (HR 4.31, 95% CI 3.55-5.26) was the most significant contributor to suicide death, followed by psychiatric diagnosis status (HR 1.84, 95% CI 1.47-2.30).

Conclusions: Receiving psychiatric assessment from a health care professional, especially a psychiatrist, reduces suicide death in individuals who had deliberately harmed themselves before. The key characteristics of individuals with DSH behavior who die by suicide are male sex, middle age, comorbid physical disabilities, and higher socioeconomic status.

(JMIR Public Health Surveill 2023;9:e41261) doi:10.2196/41261

KEYWORDS

deliberate self-harm; suicide; psychiatry; suicidal; death; mortality; psychiatric; Korea

Introduction

Deliberate self-harm (DSH) is a behavior in which one intentionally harms oneself either by means of nonfatal methods (wrist cutting or self-poisoning via medication with minimal adverse side effects) or fatal methods (the use of firearms or pesticides or hanging) [1,2]. It is a global health problem and is clinically important because individuals who commit DSH are at an increased risk of death by suicide [3-8]. A history of DSH is not only a marker of short-term suicide and subsequent self-harm but is also associated with poor treatment prognosis and noncompliance [2].

Every suicide case can be viewed as a DSH case, but not all DSH cases necessarily lead to suicide [2]. Therefore, it is difficult to measure DSH rates from nationwide statistical polls, which only report the number of deaths by suicide [9]. Previous studies on DSH used cohorts of small to moderate sizes based on surveys geared toward a particular population or extracted from national medical service use data using the International Classification of Diseases, 10th Revision (ICD-10) codes for DSH diagnoses and the cause of death information derived from linkage to the National Death Index [1,10,11]. Furthermore, acquiring a nationwide self-harm cohort assembly is difficult because not every individual with DSH behavior uses medical services or receives clinical diagnoses for DSH [12,13]. Although studying these individuals at the national level would be technically improbable, attempting to do so using patient information collected from every hospital in the respective country may still reveal valuable information. This information could be generalized to better understand these high-risk individuals, allowing for more suicide prevention opportunities.

Previous literature informs us of the factors contributing to deaths by suicide. These include sociodemographic factors such as poor income class; old age; male sex; and clinical data such as having received a psychiatric diagnosis, previous suicide attempts, substance abuse history, and having physical disabilities [3,5,6,12,14-19]. Among various physical disabilities and medical conditions, disease states affecting activities of daily living and overall function such as malignancies, musculoskeletal disorders, pulmonary conditions (ie, chronic obstructive pulmonary disease and asthma), and neurological conditions (ie, stroke) were not only associated with severe mental illnesses and increased health service usage [20] but also act as important risk factors for suicide in both men and women [16,21]. According to one study, older adults with physical disabilities who die by suicide often consult their physicians within weeks of their death [12]. Physical symptoms such as pain are often the focus of these visits, and mental distress and suicidal and self-harming ideations are often missed [15].

Although there is much evidence for this important connection between physical disabilities and deaths by suicide, nonpsychiatric health care professionals tend to take ambivalent stances on how to address mental health issues with their patients, as suggested by a few previous studies [22-24]. This ambivalence most likely stems from anticipated difficulties in dealing with patients with mental health conditions [25], and the scarcity of protocols for the assessment and management of mental health crises in nonpsychiatric patients is thought to be fueling the discomfort of actively screening for mental health issues and suicide risk in this population. Currently, questionnaires and scales for screening depression and anxiety, such as the Patient Health Questionnaire-9 [26], and suicidal ideation, such as the Columbia-Suicide Severity Rating Scale [27], and protocols that not only assess but also offer suicide-specific interventions, such as the Collaborative Assessment and Management of Suicidality [28], are clinical tools exclusively used in psychiatric emergency settings, such as a psychiatric ward or emergency room [29]. Unless receiving extensive training for how to administer these tools, it is very unlikely that nonpsychiatric health care professionals would be able to successfully make use of them, which was a point addressed in one study that investigated the possibility of implementing mental health risk assessment tools using digital resources [30]. In all parts of the world, there appears to be a lack of evidence-based, comprehensive guides for nonpsychiatric health care professionals in mental health and suicide risk assessment in nonpsychiatric patients; therefore, this area requires much more attention and investigation. By studying individuals with DSH behavior who failed to receive psychiatric assessment, whom we expect to also have physical disabilities, we may be able to find important clues for developing a suicide risk assessment in a nonpsychiatric patient protocol geared toward nonpsychiatric health care professionals.

Suicide is a greater issue for South Korea, which has been rated as the leading country for suicide among all Organization for Economic Co-operation and Development countries for several years [31]. To date, 3 studies have used Korean National Death Registration data or National Health Insurance (NHI) service data, from 2002 to 2013, to study the factors associated with suicide rate trends in South Korea [8,18,32]. Choi et al [8] and Chen et al [18], using 2 different data sets that were both still representative of the entire South Korean population, reported that old age, low socioeconomic status, and visiting medical facilities with ICD-10 F00-F99 codes (psychiatric diagnoses and mental and behavioral disorders) were associated with an increased risk of suicide completion. A study conducted by Kim et al [32] investigated the prevalence of depression in South Korea and provided evidence that an increased prevalence of depression contributed to an increased risk of suicide. Previous literature on suicide-related deaths in South Korea neither investigated nor used nationwide all-hospital patient data on self-harm. Moreover, the authors did not use a strict operational definition of death by suicide to study individuals with DSH behavior in South Korea. Therefore, a more thorough investigation of individuals with DSH behavior and the factors contributing to the increased rate of suicide is warranted.

Objectives

This study aimed to address the following questions: (1) Are there significant differences in demographics, socioeconomic status, and clinical features in individuals who received a psychiatric diagnosis (either before DSH or after the DSH event) and those who did not? (2) Does receiving a psychiatric diagnosis from the Department of Psychiatry as opposed to other departments affect survival? and (3) Which factors related to DSH contribute to death by suicide?

Methods

Data Sources

We used the Korean National Health Insurance Service (KNHIS) database [33], an administrative database based on health insurance claims from the entire South Korean population, to analyze the characteristics of patients with DSH and suicidal behavior. In this retrospective study, we constructed a DSH cohort to identify patients who visited the hospital for more than one DSH episode from January 2002 to December 2020. DSH was defined using ICD-10 codes for intentional self-harm (X60-X84).

The KNHIS is a universal coverage health insurance system, which includes detailed treatment practices and prescriptions based on the fee-for-service payment model as well as the health information of all citizens who have signed up for national medical insurance in South Korea [34]. Almost all South Koreans (97%) are enrolled in the KNHIS and most receive medical treatment at least once a year [33]. Data included public data on health care use such as disease diagnoses; drug prescriptions and procedures; national health examination results, including smoking, drinking, physical measurements, and body measurements; and demographic and socioeconomic variables such as age, sex, income rank, household location, disability, and mortality of the entire Korean population. In particular, within the health insurance claim data, diagnoses were coded in compliance with the ICD-10 [34,35]. However, because DSH is a very sensitive subject, some data related to personal information, such as personal data of clinics and household location, detailed drug prescriptions, and treatments, were not available. Statistics Korea links and provides the cause of death statistics [36]. We asked Statistics Korea to link the cause and date of death of all patients with DSH. The KNHIS data were linked to the cause of death statistics based on the resident number for patients.

Study Design

This study was designed to identify and analyze how differences in suicide survival and demographic, socioeconomic, and health-related variables vary depending on psychiatric diagnosis for patients who were admitted to the hospital after DSH. We chose certain factors, previously shown to have moderate to high correlation with DSH or death by suicide, as target variables for analysis using the KNHIS database [2,4,11,14,37]. Through this process, a total of 11 variables from the KNHIS database were extracted for in-depth analyses: age, sex, drinking, smoking, cancer, physical disability, insurance type, psychiatric disorders or psychiatric diagnosis, self-harm treatment setting, self-harm method, and somatic disorders or Charlson comorbidity index (CCI; Multimedia Appendix 1). Before performing the main analyses, we preprocessed the data using the following 4 steps.

First, we included suicide survival status and DSH severity in the DSH-related factors. Death status was classified by confirming the date of death, and suicide death was classified by confirming the diagnostic code with X60-X84 from the ICD-10. We also classified the severity of DSH using the DSH method. DSH attempts with drugs (X60-X64) and cutting (X78) were defined as nonfatal DSH. In addition, we defined DSH attempts as fatal DSH including hanging (X70), pesticides (X68), or jumping from heights (X80), according to previous studies [38,39].

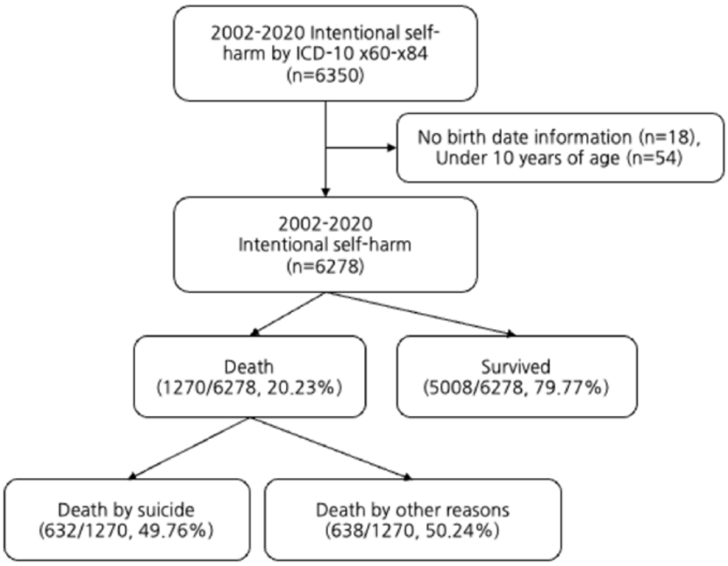
Second, we included age and sex in the demographic factors. We excluded patients under the age of 10 years because the cause of death could be accidental, and death by suicide was perceived as intentional after the age of 10 years [40]. We analyzed the differences between the groups by grouping the ages by 10 years.

Third, we included CCI, cancer, psychiatric visits, and disability in the health-related factors [36]. CCI is a widely used comorbidity correction method that involves assigning a specific weight of 1 to 6 points to 19 diseases identified by medical record surveys and then adjusting the sum of these weights [36]. This was used to confirm the risk of DSH according to the severity of the disease. We also categorized the patients into 2 groups: those who had at least one cancer diagnosis and those who had none, regardless of the type of cancer. For psychiatric visits, we included all events when patients had been given a psychiatric diagnosis at least once, not only in the psychiatric department but also in other departments. Disability was classified into 6 grades and was used as a substitute indicator of mobility. These grades were also classified into 3 categories: severe (grades 1-3), mild (grades 4-6), and no disability [41].

Finally, we included smoking, drinking, and NHI as the socioeconomic factors. For smoking and drinking, although the units of measurement were different according to the type of alcohol and tobacco, we categorized patients into 3 groups according to its severity: (1) *severe*, when substantial use level required health care provider or expert help, (2) *moderate*, when substantial use level was less severe but abstinence was recommended, and (3) *NA*, when a related item was not measured because of nonsubstantial users or unknown reasons. For NHI, we classified people into 2 groups: Medical Aid beneficiaries (Medical benefit) who are covered by the Medical Aid program as a low-income population and NHI members with employee health insurance and local-subscriber health insurance (employee or local) above the threshold of 50% of the median income who are covered by the NHI program [42].

Of the 6350 patients who committed DSH, 18 (0.28%) patients with missing birth dates and 54 (0.85%) patients aged <10 years were excluded. The cohort database in this study comprised 6278 patients who had experienced more than one deliberate episode from January 2002 to December 2020, and among the cohort participants, 5008 survivors and 632 suicide deaths were selected for the analysis (Figure 1).

Figure 1. Patient selection flow. ICD-10: International Classification of Diseases, 10th Revision.



To determine how psychiatry and its treatment affect DSH, we divided the patients into the following groups: (1) those who self-harmed before receiving a psychiatric diagnosis, (2) those who self-harmed after receiving a psychiatric diagnosis, and (3) those who did not receive a psychiatric diagnosis. We then analyzed whether these groups showed significant differences in their features.

A psychiatric diagnosis can be received in other departments through practices such as the prescription of psychiatric drugs. We also analyzed in detail whether there were significant differences between the groups classified according to the department providing the psychiatric diagnosis. All patients were classified into five groups as follows: (1) psychiatric diagnosis received only from psychiatry, (2) psychiatric diagnosis received from other departments and then from psychiatry, (3) psychiatric diagnosis received from psychiatry and then from other departments, (4) psychiatric diagnosis received from other departments only, and (5) no psychiatric diagnosis. Next, we analyzed whether there were significant differences between the groups.

Statistical Analysis

The mean and SD or the median and range were used to express continuous variables. ANOVA was used to compare categorical independent variables (with 2 or more categories) and a normally distributed interval-dependent variable. The Kruskal-Wallis test was used when 1 independent variable had 2 or more levels and an ordinal dependent variable. The cumulative incidence of events was calculated using the Kaplan-Meier estimator of the relevant survival function and graphed with 95% CIs. Hazard ratios (HRs) were calculated using a Cox proportional hazards model. To identify the risk factors possibly affecting suicide after DSH, we first conducted a univariate analysis using a logistic regression model procedure for 11 variables. In the multivariate linear regression analysis, only variables that were statistically significant ($P<.05$) through univariate regression analysis were selected, and their effect on suicidal risk was evaluated. Variables found to be significant in the univariate analysis were included in the multivariate analysis using the Cox proportional hazards model [43]. Among the Cox regression analysis results, the HRs and 95% CIs were reported. The threshold for statistical significance was set at $P<.05$. The R software was used to conduct all statistical analyses (version 3.6.3; R Foundation for Statistical Computing).

Ethics Approval

This retrospective study was approved by the institutional review board of Severance Hospital, Yonsei University College of Medicine, and the requirement for informed consent was waived (institutional review board number 4-2021-0950). As this was a retrospective study, the need for informed consent was waived because it used collected data that were anonymously managed at all stages, including data cleansing and statistical analyses.

Results

Patient Characteristics

A total of 5640 patients were included in this study, of which 2573 (45.62%) were male, and the remaining were female. Of the 5640 patients, 4576 (81.13%) patients received a psychiatric diagnosis between January 2002 and December 2020, with 3821 (67.75%) patients receiving a psychiatric diagnosis before self-harm. In total, 1064 (18.87%) patients did not receive any psychiatric diagnosis or treatment.

There were significant differences in all features, including demographic, health-related, socioeconomic, and survival variables between the groups ($P<.001$; Table 1). Most notably, there was a significant difference in survival rates sequentially in the group that did not receive a psychiatric diagnosis (81.48%), followed by one group that received psychiatric diagnosis and later self-harmed (89.06%), and one group that self-harmed before receiving the psychiatric diagnosis (97.7%). As each group had distinct characteristics, we described the characteristics of each group in the order of the suicide survival rate as follows.

Table 1. Characteristics of individuals with deliberate self-harm (DSH) behavior based on psychiatric diagnosis status (receiving at least one psychiatric diagnosis either before or after DSH event or not receiving any at all).

	Self-harm before receiving psychiatric diagnosis (n=755, 13.39%)	Self-harm after receiving psychiatric diagnosis (n=3821, 67.75%)	No psychiatric diagnosis (n=1064, 18.87%)	Total (N=5640)	*P* value
Suicide survival status, n (%)					<.001
Survived	738 (97.7)	3403 (89.06)	867 (81.48)	5008 (88.79)	
Death by suicide	17 (2.3)	418 (10.94)	197 (18.52)	632 (11.21)	
Self-harm severity, n (%)					<.001
Fatal	364 (48.2)	1393 (36.46)	600 (56.39)	2357 (41.79)	
Nonfatal	391 (51.8)	2428 (63.54)	464 (43.61)	3283 (58.21)	
Demographic factors					
Sex, n (%)					<.001
Male	399 (52.8)	1505 (39.39)	669 (62.88)	2573 (45.62)	
Female	356 (47.2)	2316 (60.61)	395 (37.12)	3067 (54.38)	
Age, median (range)	40.0 (25.0-53.0)	48.0 (32.0-63.0)	43.0 (29.0-54.0)	46.0 (30.0-59.0)	<.001
Age group, n (%)					<.001
10-19	104 (13.8)	264 (6.91)	123 (11.56)	491 (8.71)	
20-29	139 (18.4)	562 (14.71)	154 (14.47)	855 (15.16)	
30-39	134 (17.7)	523 (13.69)	192 (18.05)	849 (15.05)	
40-49	133 (17.6)	670 (17.53)	213 (20.02)	1016 (18.01)	
50-59	132 (17.5)	686 (17.95)	200 (18.80)	1018 (18.05)	
60-69	61 (8.1)	430 (11.25)	88 (8.27)	579 (10.27)	
70-79	31 (4.1)	405 (10.60)	73 (6.86)	509 (9.02)	
>80	21 (2.8)	281 (7.35)	21 (1.97)	323 (5.73)	
Health-related factor					
Charlson comorbidity index score, median (range)	0.0 (0.0-2.0)	1.0 (0.0-3.0)	1.0 (0.0-2.0)	1.0 (0.0-3.0)	<.001
Cancer, n (%)	298 (39.5)	1827 (47.81)	363 (34.12)	2488 (44.11)	<.001
Psychiatric visit, n (%)	540 (71.5)	3261 (85.34)	0 (0.00)	3801 (67.39)	<.001
Disability, n (%)					<.001
Mild	43 (5.7)	353 (9.24)	52 (4.89)	448 (7.94)	
Severe	32 (4.2)	255 (6.67)	21 (1.97)	308 (5.46)	
No	680 (90.1)	3213 (84.09)	991 (93.14)	4884 (86.60)	
Socioeconomic factor, n (%)					
Smoking					<.001
Moderate	283 (37.5)	1731 (45.30)	403 (37.88)	2417 (42.85)	
Heavy	171 (22.6)	702 (18.37)	252 (23.68)	1125 (19.95)	
N/A[a]	301 (39.9)	1388 (36.33)	409 (38.44)	2098 (37.20)	
Drinking					<.001
Moderate	265 (35.1)	1697 (44.41)	376 (35.34)	2338 (41.45)	
Heavy	190 (25.2)	730 (19.11)	278 (26.13)	1198 (21.24)	
N/A	300 (39.7)	1394 (36.48)	410 (38.53)	2104 (37.31)	
National Health Insurance					<.001
Employee or local	718 (95.1)	3343 (87.49)	1029 (96.71)	5090 (90.25)	

	Self-harm before receiving psychiatric diagnosis (n=755, 13.39%)	Self-harm after receiving psychiatric diagnosis (n=3821, 67.75%)	No psychiatric diagnosis (n=1064, 18.87%)	Total (N=5640)	*P* value
Medical benefit	37 (4.9)	478 (12.51)	35 (3.29)	550 (9.75)	

[a] N/A: not applicable.

First, the group without a psychiatric diagnosis had the lowest suicide survival rate (867/1064, 81.48%) and highest rate of fatal self-harm (600/1064, 56.39%). A total of 1064 patients were included in the group, 197 (18.52%) of whom died by suicide. The median age was 43.0 (range 29.0-54.0) years, and 669 (62.88%) patients were male. This group was mainly composed of middle-age men (605/1064, 66.87%) in their 30s and 50s. Regarding health-related factors, none of them had a history of psychiatric visits, and many (991/1064, 93.14%) had no disability. Regarding socioeconomic factors, most had employee or local insurance (1029/1064, 96.71%) and had the highest rate of heavy smokers and drinkers among the 3 groups.

Second, the group that self-harmed after receiving a psychiatric diagnosis had the highest rate of DSH using the nonfatal method (2428/3821, 63.54%). A total of 3821 patients were included in this group, 418 of whom died by suicide. Regarding demographic factors, there was a high proportion of older women (aged >60 years) and the median age was 48.0 (range 32.0-63.0) years. They had the highest rates of cancer (1827/3821, 47.81%) and psychiatric visit history (3261/3821, 85.34%) among the 3 groups. A history of disability was noted in 608 (608/3821, 15.91%) patients. Socioeconomic factors had the highest rates for medical aid benefit insurance (478/3821, 12.51%), moderate smoking (1731/3821, 45.30%), and drinking (1697/3821, 44.41%).

Finally, the group that self-harmed before receiving a psychiatric diagnosis, that is, the group that self-harmed before clinical visits, had the highest suicide survival rate (738/755, 97.7%), with a high rate of nonfatal self-harm (391/755, 51.8%). A total of 755 patients were included in this group, 17 of whom died by suicide. The median age of the patients was 40.0 (range 25.0-53.0) years, and 399 (52.8%) patients were male. They had a higher rate of people in their 10s and 20s (32.2%) than other age groups. Cancer, disability history, and psychiatric visit history was noted in 298 (39.5%), 37 (4.8%), 75 (9.8%), and 540 (71.5%) patients, respectively. Regarding socioeconomic factors, most participants had employee or local insurance (718/755, 95.1%), and there were more smokers (454/755, 60.1%) and drinkers (455/755, 60.3%).

Characteristics of Individuals With DSH Behavior

We also classified all patients into 5 subgroups based on the medical department they had received a psychiatric diagnosis from (Department of Psychiatry vs other departments) to determine whether psychiatric evaluation from the Department of Psychiatry as opposed to a nonpsychiatric department affected the outcome (Table 2). All variables were significantly different among the 5 groups (*P*<.001). In particular, these groups significantly differed in suicide rates from the group treated at the psychiatric clinic only, with the highest suicide survival rate (888/951, 93.4%) in the group with no psychiatric diagnosis and the lowest suicide survival rate (867/1064, 81.48%) in the following order.

Table 2. Comparison of individuals with deliberate self-harm by the different departments through which one received a psychiatric diagnosis.

Variable	Psychiatric diagnosis from psychiatry (n=951, 16.86%)	Psychiatric diagnosis from other departments first, then psychiatry (n=1330, 23.58%)	Psychiatric diagnosis from psychiatry first, then other departments (n=1520, 26.95%)	Psychiatric diagnosis from other departments (n=775, 13.74%)	No psychiatric diagnosis (n=1064, 18.87%)	Total (N=5640)	*P* value
Survival status, n (%)							<.001
Survived	888 (93.4)	1212 (91.13)	1366 (89.87)	675 (87.1)	867 (81.48)	5008 (88.79)	
Death by suicide	63 (6.6)	118 (8.87)	154 (10.13)	100 (12.9)	197 (18.52)	632 (11.21)	
Self-harm severity, n (%)							<.001
Fatal	615 (64.7)	877 (65.94)	946 (62.24)	381 (49.2)	464 (43.61)	2357 (41.79)	
Nonfatal	336 (35.3)	453 (34.06)	574 (37.76)	394 (50.8)	600 (56.39)	3283 (58.21)	
Demographic factors							
Sex, n (%)							<.001
Male	432 (45.4)	504 (37.89)	562 (36.97)	406 (52.4)	669 (62.88)	2573 (45.62)	
Female	519 (54.6)	826 (62.11)	958 (63.03)	369 (47.6)	395 (37.12)	3067 (54.38)	
Age (years), median (range)	32.0 (21.0-47.0)	45.0 (30.0-59.0)	52.0 (38.0-68.0)	53.0 (40.0-66.0)	43.0 (29.0-54.0)	51.0 (40.0-64.0)	<.001
Age groups, n (%)							<.001
10-19	178 (18.6)	100 (7.52)	64 (4.21)	26 (3.3)	123 (11.56)	491 (8.71)	
20-29	247 (25.8)	227 (17.07)	163 (10.72)	64 (8.2)	154 (14.47)	855 (15.16)	
30-39	170 (17.7)	200 (15.04)	191 (12.57)	96 (12.4)	192 (18.05)	849 (15.05)	
40-49	153 (16.0)	241 (18.12)	266 (17.50)	143 (18.4)	213 20.02)	1016 (18.01)	
50-59	101 (10.5)	230 (17.29)	311 (20.46)	176 (22.7)	200 (18.80)	1018 (18.05)	
60-69	54 (5.6)	155 (11.65)	180 (11.84)	102 (13.1)	88 (8.27)	579 (10.27)	
70-79	32 (3.3)	113 (8.50)	202 (13.29)	89 (11.5)	73 (6.86)	509 (9.02)	
>80	16 (1.7)	64 (4.81)	143 (9.41)	79 (10.1)	21 (1.97)	323 (5.73)	
Medical history–related factor							
Cancer, n (%)	322 (33.9)	641 (48.20)	812 (53.42)	350 (45.2)	363 (34.12)	2488 (44.11)	<.001
Charlson comorbidity index score, median (range)	0.0 (0.0-1.0)	1.0 (0.0-2.0)	2.0 (0.0-3.0)	2.0 (0.0-3.0)	1.0 (0.0-2.0)	2.0 (0.0-3.0)	<.001
Disability status, n (%)							<.001
Mild	24 (2.5)	106 (7.97)	173 (11.38)	83 (10.7)	52 (4.89)	438 (7.77)	
Severe	32 (3.3)	103 (7.74)	118 (7.76)	35 (4.5)	21 (1.97)	309 (5.48)	
No	895 (94.1)	1121 (84.29)	1229 (80.86)	657 (84.8)	991 (93.14)	4893 (86.76)	
Socioeconomic factors, n (%)							
Smoking							<.001
Moderate	278 (29.2)	574 (43.16)	772 (50.79)	390 (50.3)	403 (37.88)	2417 (42.85)	

Variable	Psychiatric diagnosis from psychiatry (n=951, 16.86%)	Psychiatric diagnosis from other departments first, then psychiatry (n=1330, 23.58%)	Psychiatric diagnosis from psychiatry first, then other departments (n=1520, 26.95%)	Psychiatric diagnosis from other departments (n=775, 13.74%)	No psychiatric diagnosis (n=1064, 18.87%)	Total (N=5640)	*P* value
Heavy	163 (17.1)	250 (18.80)	291 (19.14)	169 (21.8)	252 (23.68)	1125 (19.95)	
N/A[a]	510 (53.6)	506 (38.04)	457 (30.07)	216 (27.9)	409 (38.44)	2098 (37.20)	
Drinking							<.001
Moderate	250 (26.3)	582 (43.76)	775 (50.99)	355 (45.8)	376 (35.34)	2338 (41.45)	
Heavy	192 (20.2)	241 (18.12)	284 (18.68)	203 (26.2)	278 (26.13)	1198 (21.24)	
N/A	509 (53.5)	507 (38.12)	461 (30.33)	217 (28.0)	410 (38.53)	2104 (37.30)	
National Health Insurance							<.001
Employee or local	928 (97.6)	1160 (87.22)	1273 (83.75)	700 (90.3)	1029 (96.71)	5090 (90.25)	
Medical benefit	23 (2.4)	170 (12.78)	247 (16.25)	75 (9.7)	35 (3.29)	550 (9.75)	

[a] N/A: not applicable.

First, the group of patients who received a psychiatric diagnosis from the psychiatric department had the highest suicide survival rate (888/951, 93.4%). A total of 951 (16.86%) patients were included in this group, 63 (6.6%) of whom died by suicide. The median age was 32 (range21.0-47.0) years, and 432 (45.4%) patients were male. The number of people in their 10s and 20s was higher than that of people aged >60 years. A total of 322 (33.9%) patients were diagnosed with cancer, and the median CCI score was 0.0 (range 0.0-1.0). Among the 5 groups, they had the lowest rates of disability (56/951, 5.8%), medical benefit insurance (23/951, 2.4%), moderate/heavy smokers (441/951, 46.3%), and drinkers (442/951, 46.3%).

Second, the group of patients who received a psychiatric diagnosis from other departments and then from psychiatry had the highest rate of self-harm with fatal methods (877/1330, 65.94%). A total of 1330 (23.58%) patients were included in this group, 118 (8.87%) of whom died by suicide. The median age was 45.0 (range 30.0-59.0) years, and 504 (37.9%) patients were male. The proportion of people in their 40s (241/1330, 18.12%) was higher than in other age groups. Cancer and disability history were noted in 641 (48.20%) and 209 (15.73%) patients, respectively. The median CCI score was 1.0 (range 0.0-2.0). Regarding socioeconomic factors, 824 (61.96%) and 823 (61.88%) patients were moderate or heavy smokers and drinkers, respectively.

Third, the group of patients who received a psychiatric diagnosis from psychiatry and then from other departments included 1520 (26.95%) patients, 154 (10.13%) of whom died by suicide. The median patient age was 52 (range 38-68) years, and 562 (36.97%) patients were male. They reported a higher fatality rate (946/1520, 62.24%) in their DSH. The proportion of people in their 50s (311/1520, 20.46%) was higher than that of other age groups. They had the highest rates of cancer with 812 (53.42%) patients, disability with 291 (19.14%) patients, moderate smokers (772/1520, 50.79%), moderate drinkers (775/1520, 50.99%), and medical benefit insurance (247/1520, 16.25%) among the 5 groups.

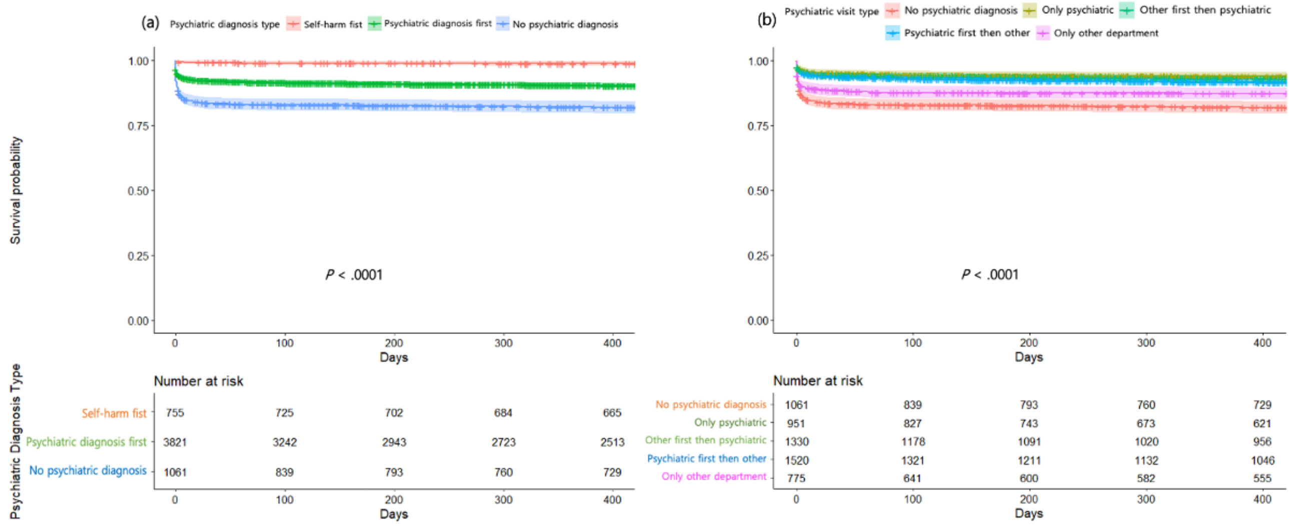
Fourth, the group of patients who received a psychiatric diagnosis only from departments other than psychiatry included 775 (13.74%) patients, 100 (12.9%) of whom died by suicide. The median age was 53.0 (range 40.0-66.0) years, and 406 (52.4%) patients were male. The highest proportion of people in their 50s (22.7%) was among the 5 groups. A history of cancer and disability was noted in 350 (45.2%) and 118 (15.2%) patients, respectively. The median CCI score was 2.0 (range 0.0-3.0). Most of them had employee or local insurance (700/775, 90.3%), and they had the highest number of heavy drinkers (203/775, 26.2%) among the 5 groups.

Finally, the group of patients who did not have a psychiatric diagnosis from any department had the lowest suicide survival rate (81.5%). A total of 1064 (18.87%) patients were included in this group, 197 (18.52%) of whom died by suicide. The median age of the patients was 43.0 (range 29.0-54.0) years, and 669 (62.88%) patients were men. The rate of people in their 40s and 50s (413/1064, 38.82%) was higher than that in other age groups. A history of cancer and disability was noted in 363 (34.12%) and 73 (6.86%) patients, respectively. The median CCI score was 1.0 (range 0.0-2.0). Most of them had employee or local insurance (1029/1064, 96.71%) and had the highest number of heavy smokers (252/1064, 23.68%) among the 5 groups.

Survival Rates of Individuals With DSH Behavior by Various Subgroups

Temporal patterns were plotted using Kaplan-Meier survival curves, calculated separately for groups according to the time between self-harm and receiving a psychiatric diagnosis (3 groups) and for groups according to the departments providing psychiatric diagnosis (5 groups) in Figure 2. These survival curves indicated that there were significantly more incidences of death by suicide over time by days in the 1-year follow-up among the participants with *no psychiatric diagnosis* than in those with *self-harm first* (Figure 2A) and among the participants with *only other departments* than in those with *only psychiatric department* (Figure 2B).

Figure 2. Kaplan-Meier curves displaying the estimated suicide survival probability for (a) 3 groups of patients whether they either did or did not receive a psychiatric diagnosis and whether they did receive a psychiatric diagnosis either before or after deliberated self-harm, as well as for (b) 5 groups of patients whether they did or did not receive a psychiatric diagnosis from psychiatric or other departments.



Risk Factors of Completed Suicide in Individuals With DSH Behavior

In the univariate Cox proportional hazards analysis, all variables showed a statistically significant impact on the suicide survival rate, except for the type of NHI (*P*=.39) and cancer history (*P*=.07; Table 3). Univariate Cox proportional hazards model indicated that higher severity of self-harm (HR 5.25, 95% CI 4.35-6.34; *P*<.001), male sex (HR 0.49, 95% CI 0.42-0.57; *P*<.001), older age (HR 1.04, 95% CI 1.03-1.04; *P*<.001), severe CCI (HR 1.51, 95% CI 1.43-1.59; *P*<.001), mild disability (HR 0.49, 95% CI 0.39-0.61; *P*<.001), severe smoking (HR 0.88, 95% CI 0.79-0.98; *P*=.02), severe drinking (HR 0.86, 95% CI 0.77-0.95; *P*=.004), no psychiatric diagnosis (HR 1.72, CI 1.45-2.04; *P*<.001), and no psychiatric visit (HR 0.66, 95% CI 1.47-2.25; *P*=.005) were significantly and independently associated with increased risk of suicide of intensive self-harm (Table 3).

Table 3. Cox proportional hazards regression of the association between self-harm-related potential risk factors and suicide-related mortality.

	Univariate analysis			Multivariate analysis		
	Coef[a]	HR[b] (95% CI)	*P* value	Coef	HR (95% CI)	*P* value
Self-harm severity	1.659	5.25 (4.35-6.34)	<.001[c]	1.461	4.31 (3.55-5.26)	<.001[c]
Demographic factors						
Sex						
Female	N/A[d]	1 (reference)	N/A	N/A	1 (reference)	N/A
Male	0.716	0.49 (0.42-0.57)	<.001[c]	0.362	0.70 (0.59-0.83)	<.001[c]
Age	0.034	1.04 (1.03-1.04)	<.001[c]	0.042	1.04 (1.03-1.06)	<.001[c]
Medical history–related factor						
Cancer	−0.149	0.86 (0.74-1.01)	.07	N/A	N/A	N/A
Charlson comorbidity index	0.411	1.51 (1.43-1.59)	<.001[c]	0.087	0.92 (0.77-1.10)	.34
Disability status						
No	N/A	1 (reference)	N/A	N/A	1 (reference)	N/A
Mild	0.722	0.49 (0.39-0.61)	<.001[c]	0.160	0.85 (0.68-1.07)	.17
Severe	−1.024	0.36 (0.23-0.56)	<.001[c]	−0.563	0.57 (0.37-0.89)	.01[e]
Socioeconomic factor						
Smoking	0.129	0.88 (0.79-0.98)	.02[e]	0.272	0.76 (0.62-0.93)	.009[f]
Drinking	0.156	0.86 (0.77-0.95)	.004[f]	0.288	0.75 (0.61-0.92)	.005[f]
National Health Insurance						
Medical benefit	N/A	1 (reference)	N/A	N/A	1 (reference)	N/A
Employee or local	−1.024	0.89 (0.67-1.17)	.39	N/A	N/A	N/A
Psychiatric diagnosis and psychiatry department–related factor						
Psychiatric diagnosis type						
Psychiatric diagnosis first	N/A	1 (reference)	N/A	N/A	1 (reference)	N/A
Self-harm first	−1.729	0.18 (0.11-0.29)	<.001[c]	−1.865	0.15 (0.10-0.25)	<.001[c]
No psychiatric diagnosis	0.541	1.72 (1.45-2.04)	<.001[c]	0.608	1.84 (1.47-2.30)	<.001[c]
Psychiatry visit type						
Psychiatry first, then other	N/A	1 (reference)	N/A	N/A	1 (reference)	N/A
Only psychiatry	−0.418	0.66 (0.49-0.88)	.005[f]	−0.380	1.46 (1.08-1.99)	.02[e]
Other first, then psychiatry	−0.143	0.87 (0.68-1.10)	.24	−0.104	1.11 (0.87-1.41)	.40
Only other departments	0.247	1.28 (1.00-1.65)	.05	0.262	1.30 (1.01-1.68)	.04[e]
No psychiatric diagnosis	N/A	1.82 (1.47-2.25)	<.001[c]	N/A	N/A	N/A

[a] Coef: coefficient.

[b] HR: hazard ratio.

[c] *P*<.001.

[d] N/A: not applicable.

[e] *P*<.05.

[f] *P*<.01.

To jointly determine the impact on survival, 9 statistically significant variables were incorporated into the multivariate model. The factors that affected suicide are given in the following order based on the HR value: higher severity of self-harm (HR 4.31, 95% CI 3.55-5.26; *P*<.001), never visited a psychiatrist (HR 1.84, 95% CI 1.47-2.30; *P*<.001), psychiatric diagnosis from departments other than psychiatry (HR 1.30, 95% CI 1.01-1.68; *P*=.04), older age (HR 1.04, 95% CI

1.03-1.06; *P*<.001), severe smoking (HR 0.76, 95% CI 0.62-1.07; *P*=.009), severe drinking habits (HR 0.75; 95% CI 0.61-0.92; *P*=.005), male sex (HR 0.70, 95% CI 0.59-0.83; *P*<.001), and no severe disability (HR 0.57, 95% CI 0.37-0.89; *P*=.01). All these factors may have statistically significant effects on the high risk of suicide.

Psychiatric diagnosis and psychiatry department–related variables, as defined in our study, also had a statistically significant effect on the high risk of suicide. *No psychiatric diagnosis* indicates that *no clinical visit* had a statistically significant effect on low suicide survival (*P*<.001). *Only psychiatry visit* (HR 1.46, 95% CI 0.49-0.87; *P*=.02) and *only other departments* (HR 1.30, 95% CI: 1.01-1.68; *P*=.04) indicates that the proper treatment for self-harm had statistically significant effect on high suicide survival.

Discussion

Principal Findings

This study aimed to investigate the characteristics of individuals with DSH behavior using nationwide cohort data and to determine the factors that contribute to suicide-related deaths in this population. Such efforts could provide valuable information about individuals with DSH behavior who are especially susceptible to subsequent death by suicide in South Korea, the nation with the highest suicide rate among Organization for Economic Co-operation and Development countries [31]. We predicted that having a psychiatric diagnosis would act as a protective factor against suicide completion. This prediction was based on previous studies describing the relationship between mental health care access and suicide rates. On the basis of a recent study exploring national trends in mental health care among adults who previously attempted suicide, despite the increasing rate of suicide attempts, the use of services among those who attempted suicide did not increase; therefore, there is a great need to expand service accessibility [9]. Several studies have suggested that greater access to mental health services was associated with a reduced incidence of suicide. More access was measured in terms of living at a range of distances from the nearest mental health provider, the availability of specific services, or the density of mental health providers [44-46]. However, these studies did not differentiate between psychiatrists and nonpsychiatric physicians in their analyses. A study exploring short-term suicide risk after psychiatric hospital discharge showed a trend in which patients who had severe psychiatric conditions for which they were seen by psychiatrists and then admitted to the psychiatric ward committed suicide within a few months [47]. Despite conflicting previous findings, it is still unclear whether being seen by a mental health care professional reduces the rate of suicide-related death in individuals with DSH behavior.

The main finding of this study is that missing a psychiatric diagnosis and therefore missing an appointment with a psychiatrist may contribute to increased HRs independent of already known contributing factors such as male sex, old age, the use of fatal form of DSH, and physical disability. Unsurprisingly, the most important contributor to suicide was the use of a fatal method of self-harm, which has been repeatedly mentioned as a very important risk factor for suicide death in previous literature [3,5,6,11,14-19]. Having a psychiatric diagnosis, whether given by a psychiatrist or a nonpsychiatrist physician, was the second most important contributing factor for suicide-related deaths. In alignment with findings from other studies, having access to a hospital where the physician can evaluate one’s psychiatric conditions may reduce the risk of suicide in individuals who had or would engage in DSH [7]. More importantly, whether someone had an interview or a check-up with a psychiatrist may be more informative than any demographic or socioeconomic status information in predicting one’s chance of committing fatal forms of DSH, which could potentially lead to suicide death.

As there was a strong correlation between receiving a psychiatric diagnosis and suicide completion, it was deemed imperative to identify the characteristics of individuals who were at a very high risk of suicide death—individuals who had never received a psychiatric diagnosis. Among the 3 subgroups of DSH, the subgroup that had never encountered a medical professional for their mental health condition showed the highest percentage of deaths by suicide, and as expected, more than half of them used fatal methods of DSH (Table 1). From a socioeconomic standpoint, this subgroup consisted mostly of men in their 40s and 50s, who consumed heavy amounts of alcohol and received health insurance through Medicare (employment/regional health insurance). Some of these characteristics, such as male sex and substance abuse, are known risk factors for increased suicide rates [2]; some other characteristics, such as being middle-age and receiving Medicare, have not been associated with suicide deaths. The KNHIS provides mandatory public health insurance, offering coverage of health care services to almost 100% of South Koreans; 97% of South Koreans are covered by Medicare and 3% are covered by Medicaid [8]. Medicaid is provided to people whose income is insufficient to meet their needs and those of their families, and they are exempted from health insurance fees, whereas those with Medicare pay approximately 10% to 30% of their total medical expenses when using medical facilities. Hence, the type of health insurance is an indirect measure of socioeconomic status, where receiving Medicaid is associated with lower socioeconomic status. Surprisingly, compared with individuals who visited the hospital for psychiatric assessment, those who did not were mostly on Medicare. Our study’s finding is opposite to what previous studies have suggested: higher socioeconomic status, including higher levels of educational achievement, higher income, and employment, is associated with better health and lower risks of all-cause mortality, including deaths by suicide [14,48]. Medicare use may be a primary contributing factor of suicide death in these individuals with DSH behavior, as having to pay for additional medical expenses to be able to see a medical professional for mental health concern may have discouraged such pursuit. Higher socioeconomic status may also have contributed to not seeking psychiatric care; South Korea being notorious for its competitive work environment and demanding high performance and efficiency has shaped such an atmosphere where middle-age employees must hide weaknesses, such as having psychological difficulties.

Table 2 displays how seeing a psychiatrist as opposed to a nonpsychiatrist physician affected survival. For our subanalysis, we categorized individuals who had received a psychiatric diagnosis by the different hospital departments they paid visit to. Suicide-related death rates increased in a stepwise fashion. Notably, individuals who had at least 1 visit to the psychiatrist were mostly women, and this sex difference in psychiatric visits has been noted in previous literature [18]. In addition, individuals in their 20s were the largest age group for most visits to the psychiatric department. Although seeing a psychiatrist continues to be taboo for most older South Korean adults and older adults, the younger generations appear to be more willing to seek professional help and voice concerns about their mental health. This finding may also reflect the recent rise in the youth and young adult depression and anxiety cases in South Korea; this phenomenon has already been hinted in previous studies [32,49]. Another noteworthy finding is that individuals who had mild to severe physical disabilities belonged to subgroups that had received psychiatric diagnoses from psychiatry departments as well as other departments and showed a tendency to engage in more severe forms of DSH. Having a physical disability significantly increased one's risk of DSH as well as death by suicide, in part by making them more susceptible to the deterioration of mental health [14,50]. It may be in the best interest of nonpsychiatric physicians to perform routine psychiatric evaluations and to actively consult with psychiatrists if their patients show signs of heightened anxiety and depression during treatment for their physical health conditions.

The total number of individuals with DSH behavior and suicide-related deaths was considerably small despite using NHI cohort data, which included all health-related information of the entire Korean population. According to the 2020 national statistics, 25.7 per 100,000 persons died of suicide attempts [51]; hence, the data set we used in our analyses did not account for all DSH and deaths by suicide cases in South Korea. We infer that this shortage of data could be accounted for by 2 main reasons: (1) suicide and self-harming behavior are major taboos in South Korean society and (2) suicide or self-harm–related diagnosis would disqualify one from reaping medical insurance benefits. It is suspected that most bereaved families as well as the attempters who survived for the reasons mentioned above would strongly refuse to receive a diagnosis for DSH. This would then put pressure on physicians to forgo, including self-harm–related ICD-10 diagnosis codes in the official hospital records. Thus, although high suicide rates in South Korea are readily recognized by the media and the Korean Statistical Information Service, it is not possible to acquire patient information from hospital electronic medical records and the NHI database for every DSH case. Nevertheless, the data set used for our analyses was technically gathered from every hospital in South Korea; therefore, it should represent the DSH population in South Korea to a reasonable extent. To that end, although our study results should be interpreted with careful consideration, our analyses have important clinical value.

Conclusions

In conclusion, we investigated and described the characteristics of individuals with DSH behavior using South Korea's nationwide cohort data and found that receiving proper assessment for psychiatric conditions, especially from a psychiatrist, reduced suicide deaths. In this modern age, we believe that the pattern we observed in individuals with DSH behavior whose deaths were suicide related is not a phenomenon solely pertinent to South Korea but also applicable in other parts of the world. Physicians, irrespective of their culture and language, may take this into consideration when screening for individuals at high risk of DSH and suicide. We also found some important clinical features of individuals with physical disabilities associated with suicidality, requiring mental health condition or suicide risk assessment in nonpsychiatric clinical settings; such information could be used to develop guidelines to assist nonpsychiatric health care professionals. Future studies using larger cohort data are necessary to further explore and replicate the findings of this study.

Acknowledgments

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP 2022-0-00064) grant funded by the Korean government (Ministry of Science and ICT [MSIT]). The funding was mainly used for data use and research labor costs. The authors used and paid for using Korean National Health Insurance Service (KNHIS) data of patients with experience of self-harm at least once for the entire nation and Statistics Korea (KOSTAT) data of deceased patients with the causes of death.

Authors' Contributions

HHK, CK, HS, and YRP contributed to study design. HHK managed the data acquisition and performed data analysis. CK contributed to data interpretation and critical revision of the manuscript. CK and HHK wrote the first draft of the manuscript. JHP performed minor data analyses. All the authors discussed the results and commented on the manuscript at all stages.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Variable selection based on literature review.
[DOCX File , 21 KB - publichealth_v9i1e41261_app1.docx]

References

1. Moran P, Coffey C, Romaniuk H, Olsson C, Borschmann R, Carlin JB, et al. The natural history of self-harm from adolescence to young adulthood: a population-based cohort study. Lancet 2012 Jan 21;379(9812):236-243 [FREE Full text] [doi: 10.1016/S0140-6736(11)61141-0] [Medline: 22100201]
2. Olfson M, Wall M, Wang S, Crystal S, Gerhard T, Blanco C. Suicide following deliberate self-harm. Am J Psychiatry 2017 Aug 01;174(8):765-774. [doi: 10.1176/appi.ajp.2017.16111288] [Medline: 28320225]
3. Tidemalm D, Beckman K, Dahlin M, Vaez M, Lichtenstein P, Långström N, et al. Age-specific suicide mortality following non-fatal self-harm: national cohort study in Sweden. Psychol Med 2015 Jun;45(8):1699-1707. [doi: 10.1017/S0033291714002827] [Medline: 25425148]
4. Hawton K, Harriss L. Deliberate self-harm in people aged 60 years and over: characteristics and outcome of a 20-year cohort. Int J Geriatr Psychiatry 2006 Jun;21(6):572-581. [doi: 10.1002/gps.1526] [Medline: 16645937]
5. Zahl DL, Hawton K. Repetition of deliberate self-harm and subsequent suicide risk: long-term follow-up study of 11,583 patients. Br J Psychiatry 2004 Jul;185:70-75. [doi: 10.1192/bjp.185.1.70] [Medline: 15231558]
6. Tidemalm D, Långström N, Lichtenstein P, Runeson B. Risk of suicide after suicide attempt according to coexisting psychiatric disorder: Swedish cohort study with long term follow-up. BMJ 2008 Nov 18;337:a2205 [FREE Full text] [doi: 10.1136/bmj.a2205] [Medline: 19018040]
7. Runeson B, Haglund A, Lichtenstein P, Tidemalm D. Suicide risk after nonfatal self-harm: a national cohort study, 2000-2008. J Clin Psychiatry 2016 Feb;77(2):240-246 [FREE Full text] [doi: 10.4088/JCP.14m09453] [Medline: 26301588]
8. Choi SB, Lee W, Yoon JH, Won JU, Kim DW. Ten-year prediction of suicide death using Cox regression and machine learning in a nationwide retrospective cohort study in South Korea. J Affect Disord 2018 Apr 15;231:8-14. [doi: 10.1016/j.jad.2018.01.019] [Medline: 29408160]
9. Bommersbach TJ, Rosenheck RA, Rhee TG. National trends of mental health care among US adults who attempted suicide in the past 12 months. JAMA Psychiatry 2022 Mar 01;79(3):219-231 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.3958] [Medline: 35044428]
10. Beckman K, Mittendorfer-Rutz E, Lichtenstein P, Larsson H, Almqvist C, Runeson B, et al. Mental illness and suicide after self-harm among young adults: long-term follow-up of self-harm patients, admitted to hospital care, in a national cohort. Psychol Med 2016 Dec 01;46(16):3397-3405 [FREE Full text] [doi: 10.1017/S0033291716002282] [Medline: 27644850]
11. Troya MI, Babatunde O, Polidano K, Bartlam B, McCloskey E, Dikomitis L, et al. Self-harm in older adults: systematic review. Br J Psychiatry 2019 Apr 01;214(4):186-200 [FREE Full text] [doi: 10.1192/bjp.2019.11] [Medline: 30789112]
12. Ahmedani BK, Simon GE, Stewart C, Beck A, Waitzfelder BE, Rossom R, et al. Health care contacts in the year before suicide death. J Gen Intern Med 2014 Jun;29(6):870-877 [FREE Full text] [doi: 10.1007/s11606-014-2767-3] [Medline: 24567199]
13. Tang S, Reily NM, Arena AF, Batterham PJ, Calear AL, Carter GL, et al. People who die by suicide without receiving mental health services: a systematic review. Front Public Health 2021 Jun;9(6):736948-736947 [FREE Full text] [doi: 10.3389/fpubh.2021.736948] [Medline: 35118036]
14. Crump C, Sundquist K, Sundquist J, Winkleby MA. Sociodemographic, psychiatric and somatic risk factors for suicide: a Swedish national cohort study. Psychol Med 2014 Jan;44(2):279-289 [FREE Full text] [doi: 10.1017/S0033291713000810] [Medline: 23611178]
15. Waern M, Rubenowitz E, Runeson B, Skoog I, Wilhelmson K, Allebeck P. Burden of illness and suicide in elderly people: case-control study. BMJ 2002 Jun 08;324(7350):1355-1389 [FREE Full text] [doi: 10.1136/bmj.324.7350.1355] [Medline: 12052799]
16. Fässberg MM, Cheung G, Canetto SS, Erlangsen A, Lapierre S, Lindner R, et al. A systematic review of physical illness, functional disability, and suicidal behaviour among older adults. Aging Ment Health 2016 Jun 08;20(2):166-194 [FREE Full text] [doi: 10.1080/13607863.2015.1083945] [Medline: 26381843]
17. Edwards AC, Ohlsson H, Mościcki E, Crump C, Sundquist J, Kendler KS, et al. Alcohol use disorder and non-fatal suicide attempt: findings from a Swedish National Cohort Study. Addiction 2022 Jan 08;117(1):96-105 [FREE Full text] [doi: 10.1111/add.15621] [Medline: 34159695]
18. Chen YY, Yang CT, Cha ES, Sha F, Yip PSF. Quantifying the contributions of age, sex, methods, and urbanicity to the changing suicide rate trends in South Korea, 2001-2016. Soc Psychiatry Psychiatr Epidemiol 2020 Sep;55(9):1121-1132 [FREE Full text] [doi: 10.1007/s00127-020-01855-3] [Medline: 32189039]
19. Guseva Canu I, Bovio N, Mediouni Z, Bochud M, Wild P, Swiss National Cohort (SNC). Suicide mortality follow-up of the Swiss National Cohort (1990-2014): sex-specific risk estimates by occupational socio-economic group in working-age population. Soc Psychiatry Psychiatr Epidemiol 2019 Dec;54(12):1483-1495. [doi: 10.1007/s00127-019-01728-4] [Medline: 31127347]
20. Ronaldson A, Elton L, Jayakumar S, Jieman A, Halvorsrud K, Bhui K. Severe mental illness and health service utilisation for nonpsychiatric medical disorders: a systematic review and meta-analysis. PLoS Med 2020 Sep;17(9):e1003284-e1003295 [FREE Full text] [doi: 10.1371/journal.pmed.1003284] [Medline: 32925912]

21. Hubers AAM, Moaddine S, Peersmann SHM, Stijnen T, van Duijn E, van der Mast RC, et al. Suicidal ideation and subsequent completed suicide in both psychiatric and non-psychiatric populations: a meta-analysis. Epidemiol Psychiatr Sci 2018 Apr;27(2):186-198 [FREE Full text] [doi: 10.1017/S2045796016001049] [Medline: 27989254]
22. Schurman RA, Kramer PD, Mitchell JB. The hidden mental health network. treatment of mental illness by nonpsychiatrist physicians. Arch Gen Psychiatry 1985 Jan 14;42(1):89-94 [FREE Full text] [doi: 10.1001/archpsyc.1985.01790240091010] [Medline: 3966857]
23. Wang J, Wang Q, Wimalaratne I, Menkes DB, Wang X. Chinese non-psychiatric hospital doctors' attitudes toward management of psychological/psychiatric problems. BMC Health Serv Res 2017 Aug 22;17(1):576-594 [FREE Full text] [doi: 10.1186/s12913-017-2521-8] [Medline: 28830543]
24. AlSalem M, Alamri R, Hejazi S. Nonpsychiatric healthcare professionals' attitudes toward patients with mental illnesses in Makkah City, Saudi Arabia: a cross-sectional study. Neuropsychiatr Dis Treat 2020 Aug 22;16(1):341-348 [FREE Full text] [doi: 10.2147/NDT.S236148] [Medline: 32099371]
25. Joung J, Jang MY, Shim J, Ko Y, Shin SH. Difficulties in caring for psychiatric patient as experienced by non-psychiatric nurses. Article in Korean. J Korean Acad Nurs 2017 Feb;47(1):49-59 [FREE Full text] [doi: 10.4040/jkan.2017.47.1.49] [Medline: 28262654]
26. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. primary care evaluation of mental disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
27. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec 10;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
28. Jobes D, Drozd J. The CAMS Approach to Working with Suicidal Patients. Journal of Contemporary Psychotherapy 2004 ;34(1):73-85 [FREE Full text] [doi: 10.1023/B:JOCP.0000010914.98781.6a] [Medline: 32099371]
29. Boudreaux ED, Larkin C, Camargo CA, Miller IW. Validation of a secondary screener for suicide risk: results from the Emergency Department Safety Assessment and Follow-up Evaluation (ED-SAFE). Jt Comm J Qual Patient Saf 2020 Jun 30;46(6):342-352 [FREE Full text] [doi: 10.1016/j.jcjq.2020.03.008] [Medline: 32417230]
30. Aref-Adib G, Landy G, Eskinazi M, Sommerlad A, Morant N, Johnson S, et al. Assessing digital risk in psychiatric patients: mixed methods study of psychiatry trainees' experiences, views, and understanding. JMIR Ment Health 2020 Jul 29;7(7):e19008-e19052 [FREE Full text] [doi: 10.2196/19008] [Medline: 32726288]
31. Jang H, Lee W, Kim YO, Kim H. Suicide rate and social environment characteristics in South Korea: the roles of socioeconomic, demographic, urbanicity, general health behaviors, and other environmental factors on suicide rate. BMC Public Health 2022 Mar 01;22(1):410 [FREE Full text] [doi: 10.1186/s12889-022-12843-4] [Medline: 35227243]
32. Kim GE, Jo MW, Shin YW. Increased prevalence of depression in South Korea from 2002 to 2013. Sci Rep 2020 Oct 12;10(1):16979 [FREE Full text] [doi: 10.1038/s41598-020-74119-4] [Medline: 33046758]
33. Kyoung DS, Kim HS. Understanding and utilizing claim data from the Korean National Health Insurance Service (NHIS) and Health Insurance Review and Assessment (HIRA) database for research. J Lipid Atheroscler 2022 May 12;11(2):103-110 [FREE Full text] [doi: 10.12997/jla.2022.11.2.103] [Medline: 35656154]
34. Kim MK, Han K, Lee SH. Current trends of big data research using the Korean National Health Information Database. Diabetes Metab J 2022 Jul;46(4):552-563 [FREE Full text] [doi: 10.4093/dmj.2022.0193] [Medline: 35929173]
35. Kim HK, Song SO, Noh J, Jeong IK, Lee BW. Diabetes Metab J 2020 Oct;44(5):671-678 [FREE Full text] [doi: 10.4093/dmj.2020.0207] [Medline: 33115211]
36. Shin HY, Lee JY, Song J, Lee S, Lee BW, Lim B, et al. Cause-of-death statistics in the Republic of Korea, 2014. J Korean Med Assoc 2016 Oct 01;59(3):221-228. [doi: 10.5124/jkma.2016.59.3.221]
37. Skegg K. Self-harm. Lancet 2005 Apr 15;366(9495):1471-1483 [FREE Full text] [doi: 10.1016/S0140-6736(05)67600-3] [Medline: 16243093]
38. Spicer RS, Miller TR. Suicide acts in 8 states: incidence and case fatality rates by demographics and method. Am J Public Health 2000 Dec 12;90(12):1885-1891. [doi: 10.2105/ajph.90.12.1885] [Medline: 11111261]
39. Yeum TS, Kim B, Kim EY, Kim SH, Ha K, Ahn YM. Factors affecting suicide method lethality among suicide attempters in the Korea National Suicide Survey. J Nerv Ment Dis 2018 Mar;206(3):202-210. [doi: 10.1097/NMD.0000000000000735] [Medline: 28902068]
40. Lee MJ, Romero S, Hong I, Park HY. Evaluating Korean Personal Assistance Services Classification System. Ann Rehabil Med 2018 Oct;42(5):758-766 [FREE Full text] [doi: 10.5535/arm.2018.42.5.758] [Medline: 30404425]
41. Oh MK, Jang H, Kim YI, Jo B, Kim Y, Park JH, et al. Differences in obesity rates between people with and without disabilities and the association of disability and obesity: a nationwide population study in South Korea. J Prev Med Public Health 2012 Jul;45(4):211-218 [FREE Full text] [doi: 10.3961/jpmph.2012.45.4.211] [Medline: 22880152]
42. Park S. Medical service utilization and out-of-pocket spending among near-poor National Health Insurance members in South Korea. BMC Health Serv Res 2021 Aug 28;21(1):886-888 [FREE Full text] [doi: 10.1186/s12913-021-06881-8] [Medline: 34454499]

43. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. Br J Cancer 2003 Aug 04;89(3):431-436 [FREE Full text] [doi: 10.1038/sj.bjc.6601119] [Medline: 12888808]
44. McCarthy JF, Blow FC, Ignacio RV, Ilgen MA, Austin KL, Valenstein M. Suicide among patients in the Veterans Affairs health system: rural-urban differences in rates, risks, and methods. Am J Public Health 2012 Mar 04;102 Suppl 1(Suppl 1):S111-S117 [FREE Full text] [doi: 10.2105/AJPH.2011.300463] [Medline: 22390583]
45. Price JH, Mrdjenovich AJ, Dake JA. Prevalence of state firearm mortality and mental health care resources. J Community Health 2009 Oct;34(5):383-391. [doi: 10.1007/s10900-009-9163-7] [Medline: 19536645]
46. Shumway M, Alvidrez J, Leary M, Sherwood D, Woodard E, Lee EK, et al. Impact of capacity reductions in acute public-sector inpatient psychiatric services. Psychiatr Serv 2012 Feb 01;63(2):135-141. [doi: 10.1176/appi.ps.201000145] [Medline: 22302330]
47. Olfson M, Wall M, Wang S, Crystal S, Liu SM, Gerhard T, et al. Short-term suicide risk after psychiatric hospital discharge. JAMA Psychiatry 2016 Nov 01;73(11):1119-1126 [FREE Full text] [doi: 10.1001/jamapsychiatry.2016.2035] [Medline: 27654151]
48. Denney JT, Rogers RG, Krueger PM, Wadsworth T. Adult suicide mortality in the United States: marital status, family size, socioeconomic status, and differences by sex. Soc Sci Q 2009 Dec 01;90(5):1167 [FREE Full text] [doi: 10.1111/j.1540-6237.2009.00652.x] [Medline: 20161640]
49. Yun JY, Chung H, Sim JA, Yun YH. Prevalence and associated factors of depression among Korean adolescents. PLoS One 2019 Dec 01;14(10):e0223176 [FREE Full text] [doi: 10.1371/journal.pone.0223176] [Medline: 31618232]
50. Yook V, Kim H, Kim EJ, Kim Y, Lee G, Choi JH, et al. Psychological autopsy study comparing suicide decedents with and without a history of suicide attempts in a nationwide sample of South Korea. Suicide Life Threat Behav 2022 Apr;52(2):190-198. [doi: 10.1111/sltb.12750] [Medline: 33811661]
51. Vital Statistics of Korea: Deaths (persons). KOSIS (Korean Statistical Information Service). URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B8000F&language=en [accessed 2023-03-24]

Abbreviations

CCI: Charlson comorbidity index
DSH: deliberate self-harm
HR: hazard ratio
ICD-10: International Classification of Diseases, 10th Revision
KNHIS: Korean National Health Insurance Service
NHI: National Health Insurance

Edited by A Mavragani; submitted 20.07.22; peer-reviewed by V Kumar, MM Patwary; comments to author 23.08.22; revised version received 20.09.22; accepted 29.09.22; published 12.04.23.

Please cite as:
Kim HH, Ko C, Park JA, Song IH, Park YR
Investigation of the Relationship Between Psychiatry Visit and Suicide After Deliberate Self-harm: Longitudinal National Cohort Study
JMIR Public Health Surveill 2023;9:e41261
URL: https://publichealth.jmir.org/2023/1/e41261
doi:10.2196/41261
PMID:37043262

Original Paper

Consultation Rate and Mode by Deprivation in English General Practice From 2018 to 2022: Population-Based Study

Emma Maria Vestesson[1,2], MA; Kaat Lieve An De Corte[1], MSc; Elizabeth Crellin[1], MSc; Jean Ledger[3], PhD; Minal Bakhai[3], MBBS; Geraldine M Clarke[1], PhD

[1]The Health Foundation, London, United Kingdom
[2]Great Ormond Street Institute of Child Health, University College London, London, United Kingdom
[3]National Health Service England, London, United Kingdom

Corresponding Author:
Emma Maria Vestesson, MA
The Health Foundation
8 Salisbury Square
London, EC4Y 8AP
United Kingdom
Phone: 44 020 7257 8000
Email: emma.vestesson@gmail.com

Abstract

Background: The COVID-19 pandemic has had a significant impact on primary care service delivery with an increased use of remote consultations. With general practice delivering record numbers of appointments and rising concerns around access, funding, and staffing in the UK National Health Service, we assessed contemporary trends in consultation rate and modes (ie, face-to-face versus remote).

Objective: This paper describes trends in consultation rates in general practice in England for key demographics before and during the COVID-19 pandemic. We explore the use of remote and face-to-face consultations with regard to socioeconomic deprivation to understand the possible effect of changes in consultation modes on health inequalities.

Methods: We did a retrospective analysis of 9,429,919 consultations by general practitioners, nurses, or other health care professionals between March 2018 and February 2022 for patients registered at 397 general practices in England. We used routine electronic health records from Clinical Practice Research Datalink Aurum with linkage to national data sets. Negative binomial models were used to predict consultation rates and modes (ie, remote versus face-to-face) by age, sex, and socioeconomic deprivation over time.

Results: Overall consultation rates increased by 15% from 4.92 in 2018-2019 to 5.66 in 2021-2022 with some fluctuation during the start of the COVID-19 pandemic. The breakdown into face-to-face and remote consultations shows that the pandemic precipitated a rapid increase in remote consultations across all groups, but the extent varies by age. Consultation rates increased with increasing levels of deprivation. Socioeconomic differences in consultation rates, adjusted for sex and age, halved during the pandemic (from 0.36 to 0.18, indicating more consultations in the most deprived), effectively narrowing relative differences between deprivation quintiles. This trend remains when stratified by sex, but the difference across deprivation quintiles is smaller for men. The most deprived saw a relatively larger increase in remote and decrease in face-to-face consultation rates compared to the least deprived.

Conclusions: The substantial increases in consultation rates observed in this study imply an increased pressure on general practice. The narrowing of consultation rates between deprivation quintiles is cause for concern, given ample evidence that health needs are greater in more deprived areas.

(JMIR Public Health Surveill 2023;9:e44944) doi:10.2196/44944

KEYWORDS

primary care; deprivation; England; remote consultations; pandemic; COVID-19; CPRD Aurum; Clinical Practice Research Datalink; general practice; health inequalities; consultation; electronic health records; age; sex; care delivery; database; information management; population data; Index of Multiple Deprivation; IMD; cohort; longitudinal; data linkage; data link; general practice; random; consult; person-year

Introduction

In the last few years, digital technology has enabled new ways of working in general practice. Before the COVID-19 pandemic, remote consultations (by telephone, video, text based, and web based) were steadily increasing on a background of supporting policy. The National Health Service Long Term Plan 2019 committed to every patient having the right to digital-first primary care by 2023-2024, and the 2019/20 general practitioner (GP) contract reform framework set out a requirement for all practices to offer web-based consultation systems by 2021 [1,2]. The COVID-19 pandemic served as a catalyst for the uptake of remote consultations, as guidance was issued to triage patient contacts, wherever possible, and encourage the use of remote consultations, if clinically appropriate, to avoid the risks of COVID-19 exposure to patients and staff [3]. Before the pandemic, it was reported that less than 30% of consultations were carried out remotely; within weeks, it increased to 77% [4].

Both remote and face-to-face consultation modes have advantages and disadvantages. Remote consultations can offer potential benefits through expanding access to services and appointment flexibility—in particular, for patients in rural areas, patients who find face-to-face consultations difficult, and those with substantial barriers for travelling to their general practice (eg, having mobility issues and employment or caring commitments). Faster access to care and a more cost-effective alternative to face-to-face appointments have also been highlighted as potential benefits to remote consultations [5]. In this way, remote consultations could decrease some of the current inequalities observed in the use of primary care.

On the other hand, there are concerns that remote consultations could exclude certain patient groups and exacerbate health inequalities [6]. Although the evidence around remote consultations and inequalities is limited, there is ample evidence that primary care is under more pressure in more deprived areas, so understanding the impact of a rapid increase in remote consultations is important [7]. Factors such as age, disability, digital exclusion, communication needs, data poverty, and lack of trust can impact people's willingness or ability to access and benefit from remote consultations. A systematic review conducted at the start of the pandemic collated evidence on remote versus face-to-face consultations with a focus on inequalities and observed variations in the use of remote consultations, but the evidence was not conclusive [8]. A study of remote primary care for people experiencing homelessness during the pandemic found that the shift to remote consultations created barriers to accessing care due to factors such as the lack of funds to make calls or access to a telephone [9]. A cross-sectional study observed that there are differences in the proportion of consultations delivered remotely by category of the Index of Multiple Deprivation (IMD) [10]. However, a longitudinal study [5] did not observe a difference in the change of proportion of remote consultations over time by deprivation. Many of these studies are from before or the inception of the pandemic and the use of remote consultations has substantially changed. Studies from outside the United Kingdom do observe inequalities; but due to the difference in health systems, these results are unlikely to generalize to the English population [11,12].

In this study, we use person-level data from a large, nationally representative sample to describe contemporary patterns in consultation rates and modes in English general practice before and during the COVID-19 pandemic. We further investigate the inequalities in consultation rates and modes between individuals grouped by age, sex, and deprivation.

Methods

Study Design and Data

We performed a cohort study using person-level data from the Clinical Practice Research Datalink (CPRD) Aurum between January 2018 and February 2022. CPRD Aurum is a database with routinely collected data from general practices in England that use the EMIS Web information management system. CPRD contains data for over 40 million patients from 1332 practices in England as of May 2022 [13]. Patients are broadly representative of the English population based on age, sex, and deprivation [14]. CPRD provided linkage to the 2015 IMD at the patient level based on the patient's Lower Layer Super Output Area and to the 2011 urban-rural classification based on the practice Lower Layer Super Output Area. The study protocol was approved by CPRD's Research Data Governance (protocol number 21_000357).

Ethics Approval

The study protocol was approved by CPRD's Research Data Governance (protocol number 21_000357).

Eligibility Criteria

We studied patients registered at general practices that participated in CPRD Aurum. Eligibility criteria were applied at both practice and patient level. A total of 400 practices located in England were sampled at random. Eligible patients were those with acceptable data quality (verified by CPRD); registered at one of the 400 practices at any point between January 2018 and May 2021; recorded as either male or female sex; and eligible for area-level linkage to the IMD 2015. From this cohort, 600,000 patients were randomly sampled. Three GP practices were identified by CPRD as having duplication issues, and therefore, were excluded. In addition, 2 patients were excluded as they no longer met the inclusion criteria after the final data set was extracted.

Consultation Type and Consultation Mode

The primary information source on consultations was the CPRD Aurum "Consultation" table, which contains information on the type (eg, clinical or admin) and mode (eg, telephone or video) of the consultation. The consultation table can be linked to the "Staff" table to gain information about the staff member and the observations that occurred during the consultation [14]. We included consultations carried out by GPs, nurses, and other general practice care providers. Clinical consultations were identified. Consultations that were not attended were excluded. This builds on methods used on CPRD Gold and Aurum [15,16].

Consultations were further classified by mode of delivery as either remote (by telephone, video, SMS, and through the internet) or face-to-face (at the GP surgery or at home) consultations based on information in the consultation table for consultation mode or observations recorded during the consultation (Table S1 in Multimedia Appendix 1). Where a consultation's mode was unclear, it was assumed to be face-to-face. A patient could have multiple consultations per day with a mix of modes. The final data set included consultations for patients between March 2018 and February 2022. For year-on-year comparisons, we grouped consultations in 12-month periods (March 2018-February 2019, March 2019-February 2020, March 2020-February 2021, and March 2021-February 2022), covering 2 years before the pandemic and 2 years during the pandemic.

Statistical Analysis

We calculated person-years of observation over the study period based on the patient's registration dates and when the practices submitted data to CPRD. Crude consultation rates and the proportion of remote consultations were calculated for each month and for 12-month periods. To compare consultation rates over time, the number of consultations, remote consultations and face-to-face consultations per patient were modelled using a negative binomial regression with person-years as an offset. The models included age, sex, deprivation, and pandemic year as well as interactions between these terms. A small number of patients without a deprivation quintile were excluded from the models. From these models, we estimated adjusted consultation rates for the 12-month periods for patients grouped by deprivation quintile, age, and sex to allow absolute estimates of the differences between these groups. These rates were calculated for the "average" patient with respect to all the other variables in the model.

All analyses were carried out on a secure analysis server at the Health Foundation using R (version 4.0.3; R Core Team), with the ggeffects package for estimating predicted values [17,18].

Results

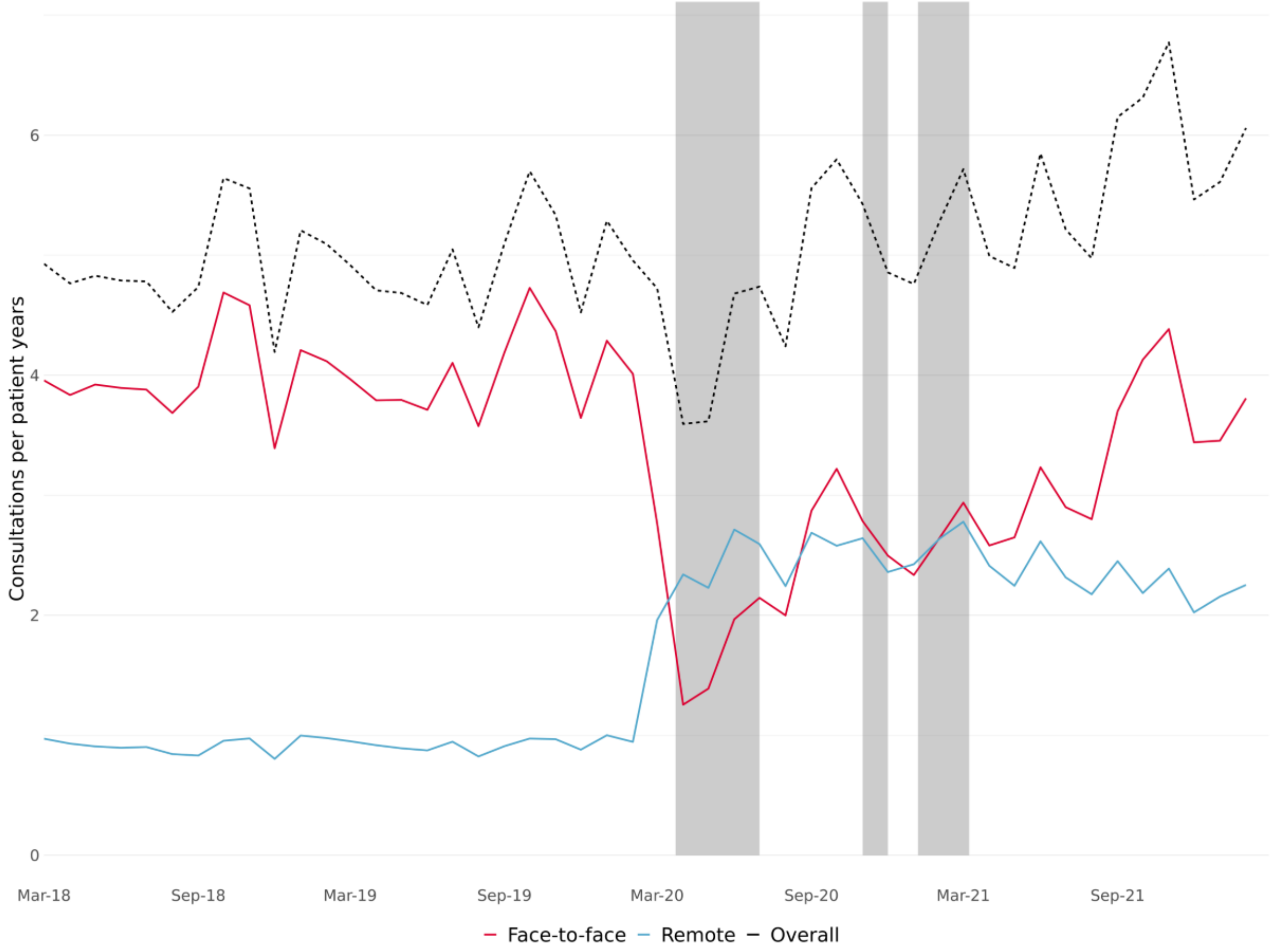
Consultation Rate

The study included 9,429,919 consultations over 1,863,507 person-years of observation from 397 practices located across all regions in England between March 2018 and February 2022. Overall, face-to-face and remote consultation rates did not change materially between 2018-2019 and 2019-2020 (Table 1; Figure 1). Consultations averaged 4.92 and 4.94 per person-year in 2018-2019 and 2019-2020, respectively, with approximately 4 times as many face-to-face (4.00 and 4.02 per person-year in 2018-2019 and 2018-2020, respectively) as remote (0.91 and 0.92 per person-year in 2018-2019 and 2018-2020, respectively) consultations. During the pandemic, consultation rates initially dropped from 4.92 per person-year in 2018-2019 to 4.76 per person-year in 2020-2021; but by 2021-2022, it had increased to 5.66 per person-year, 15% higher than 2018-2019 (Table 1; Figure 1).

Table 1. Temporal trends in general practitioner (GP), nurse, and other health care professional (HCP) consultations by subtype of face-to-face and remote consultation modes. Crude rates are per person-year (N=1,863,507).

All GP, nurse, and other HCP consultations per person-year	2018-2019, crude rate (%)	2019-2020, crude rate (%)	2020-2021, crude rate (%)	2021-2022, crude rate (%)	Change between 2018-2019 and 2021-2022 (%)
All modes	4.92 (100)	4.94 (100)	4.76 (100)	5.66 (100)	15.0
Face-to-face	4.00 (81.3)	4.02 (81.4)	2.32 (48.7)	3.32 (58.7)	–17.1
Remote	0.91 (18.5)	0.92 (18.6)	2.45 (51.5)	2.34 (41.3)	155.3

Figure 1. Overall consultation rates per patient-year and split by consultation delivery mode (ie, face-to-face and remote), calculated monthly between March 2018 and February 2022. The greyed-out areas correspond to periods of England-wide COVID-19 lockdowns.



Consultation Mode

Over our study period, there were dramatic shifts in consultation modes. In 2020-2021, compared with 2018-2019, the rate of face-to-face consultations almost halved from 4.00 per person-year to 2.34 per person-year, while the rate of remote consultations more than doubled from 0.91 to 2.45 per person-year. During 2021-2022, rates of face-to-face consultations steadily recovered, whereas the rate of remote consultations remained stable; and by the end of our study period in February 2022, more consultations were delivered face-to-face (63%) than remotely (37%; Figure S1 in Multimedia Appendix 1). In the last year of our study, 2021-2022, rates of face-to-face consultations were 17.1% lower, and rates of remote consultations were 155.3% higher, compared with prepandemic 2018-2019.

Variation by Age

Age-specific consultation rates had a consistently J-shaped distribution across all years with the highest rates in the youngest patients aged 0-4 years, decreasing to lowest levels in children aged 5-10 years, before steadily increasing in each age group to a peak in patients aged 75 years and older (Table 2).

Table 2. Consultations by mode and age per patient-years. All values for remote and face-to-face consultations are crude rates followed by percentages.

Age (years)	Consultations											
	2018-2019			2019-2020			2020-2021			2021-2022		
	Remote	Face-to-face	All modes[a]	Remote	Face-to-face	All modes[a]	Remote	Face-to-face	All modes[a]	Remote	Face-to-face	All modes[a]
0-4	0.95 (17.1)	4.61 (82.9)	5.56	0.95 (17)	4.64 (83)	5.60	1.96 (42.5)	2.64 (57.5)	4.60	2.3 (43.7)	2.96 (56.3)	5.26
5-10	0.4 (21)	1.49, (79)	1.89	0.39 (20.4)	1.51 (79.6)	1.89	0.87 (57.3)	0.65 (42.7)	1.51	0.98 (48.8)	1.03 (51.2)	2.00
11-17	0.38 (19.3)	1.58 (80.7)	1.96	0.37 (18.7)	1.62 (81.3)	2.00	0.99 (55.2)	0.8 (44.8)	1.79	1.08 (46.3)	1.25 (53.7)	2.33
18-24	0.63 (19.3)	2.63 (80.7)	3.26	0.63 (19.4)	2.63 (80.6)	3.26	1.83 (55)	1.5 (45)	3.33	1.76 (43.9)	2.25 (56.1)	4.01
25-34	0.73 (19.9)	2.94 (80.1)	3.67	0.72 (19.8)	2.93 (80.2)	3.66	2.03 (52.5)	1.83 (47.5)	3.86	1.91 (42)	2.63 (58)	4.53
35-44	0.74 (19.2)	3.12 (80.8)	3.86	0.74 (19.1)	3.13 (80.9)	3.87	2.07 (52.4)	1.88 (47.6)	3.94	1.98 (41)	2.86 (59)	4.84
45-54	0.84 (18.1)	3.81 (81.9)	4.65	0.86 (18.1)	3.92 (81.9)	4.78	2.44 (50.9)	2.35 (49.1)	4.79	2.35 (39.6)	3.59 (60.4)	5.95
55-64	0.97 (17)	4.72 (83)	5.69	0.98 (17)	4.8 (83)	5.78	2.81 (49.3)	2.89 (50.7)	5.70	2.64 (38.7)	4.18 (61.3)	6.82
65-74	1.29 (16)	6.76 (84)	8.05	1.32 (16.6)	6.65 (83.4)	7.97	3.55 (48.3)	3.8 (51.7)	7.35	3.27 (38.6)	5.21 (61.4)	8.48
≥75	2.5 (20.8)	9.52 (79.2)	12.02	2.51 (21.2)	9.34 (78.8)	11.84	5.92 (53.9)	5.06 (46.1)	10.99	5.06 (43.4)	6.61 (56.6)	11.67

[a] The percentage values are 100.

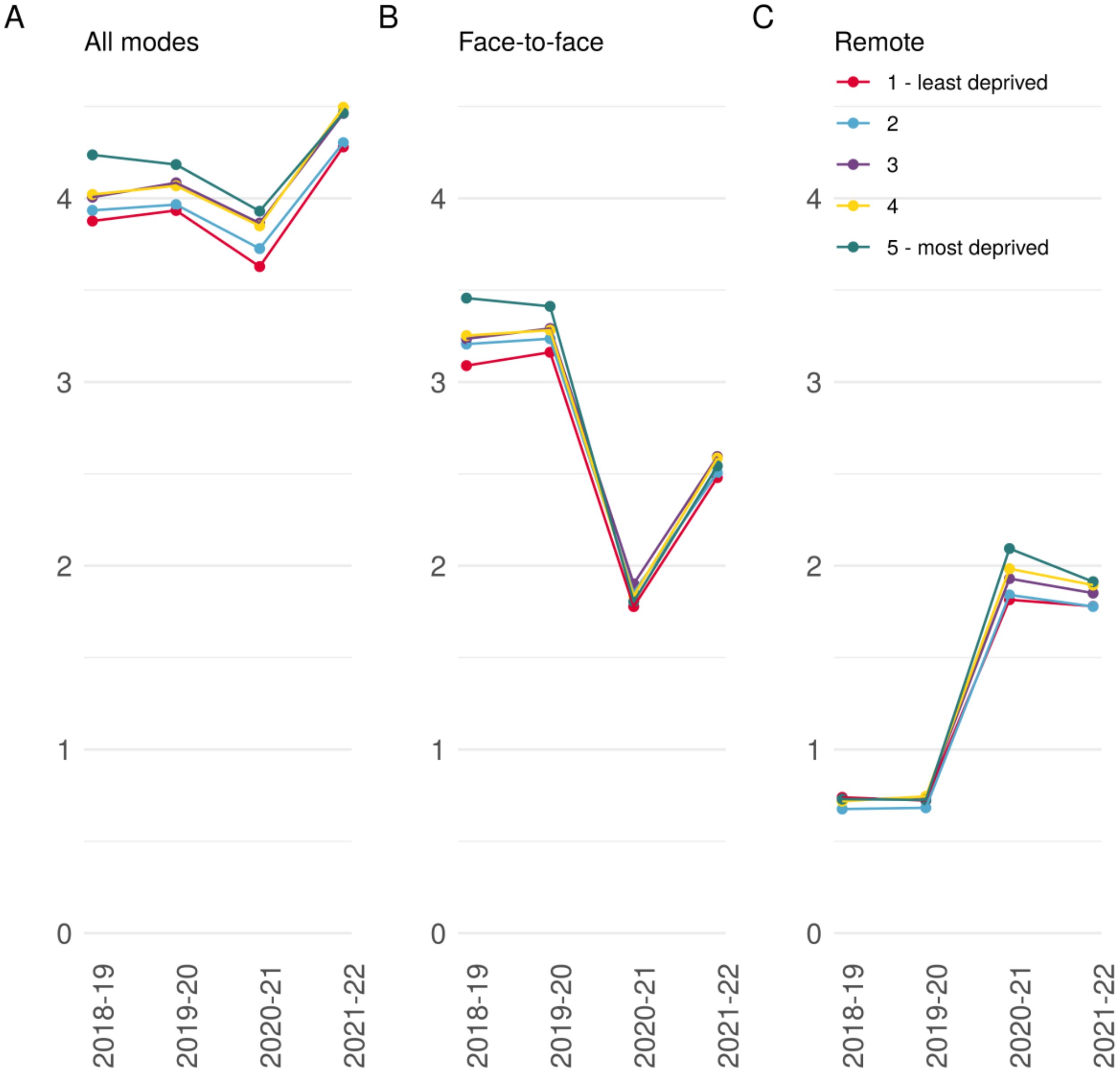
Between 2018 and 2020, consultation rates remained fairly steady within each age group with little variation in the proportion of consultations delivered remotely. In 2020-2021, a year-on-year decrease in the overall consultation rate was driven by those in the youngest (aged 0-17 years) and oldest (aged over 55 years) age groups. The decline was most pronounced in infants (aged 0-4 years), although this group sustained the highest proportion of face-to-face consultations (57.5%) during the first year of the pandemic.

In 2021-2022, consultation rates recovered to higher than prepandemic levels for all patients except for those in the youngest (aged 0-4 years) and oldest (aged ≥75 years) age groups. There was markedly less variation in consultation modes across the age groups in 2021-2022 compared with 2020-2021.

Variation by Deprivation

From the multivariable analysis, we present predicted age- and sex-adjusted consultation rates by deprivation quintiles over time (Figure 2 and Table S2 in Multimedia Appendix 1). Each quintile represents 20% of local areas in England.

Figure 2. Yearly consultation rates stratified by deprivation quintiles. (A) Overall consultation rate per person per year. (B) Face-to-face consultation rate per person per year. (C) Remote consultation rate per person per year.



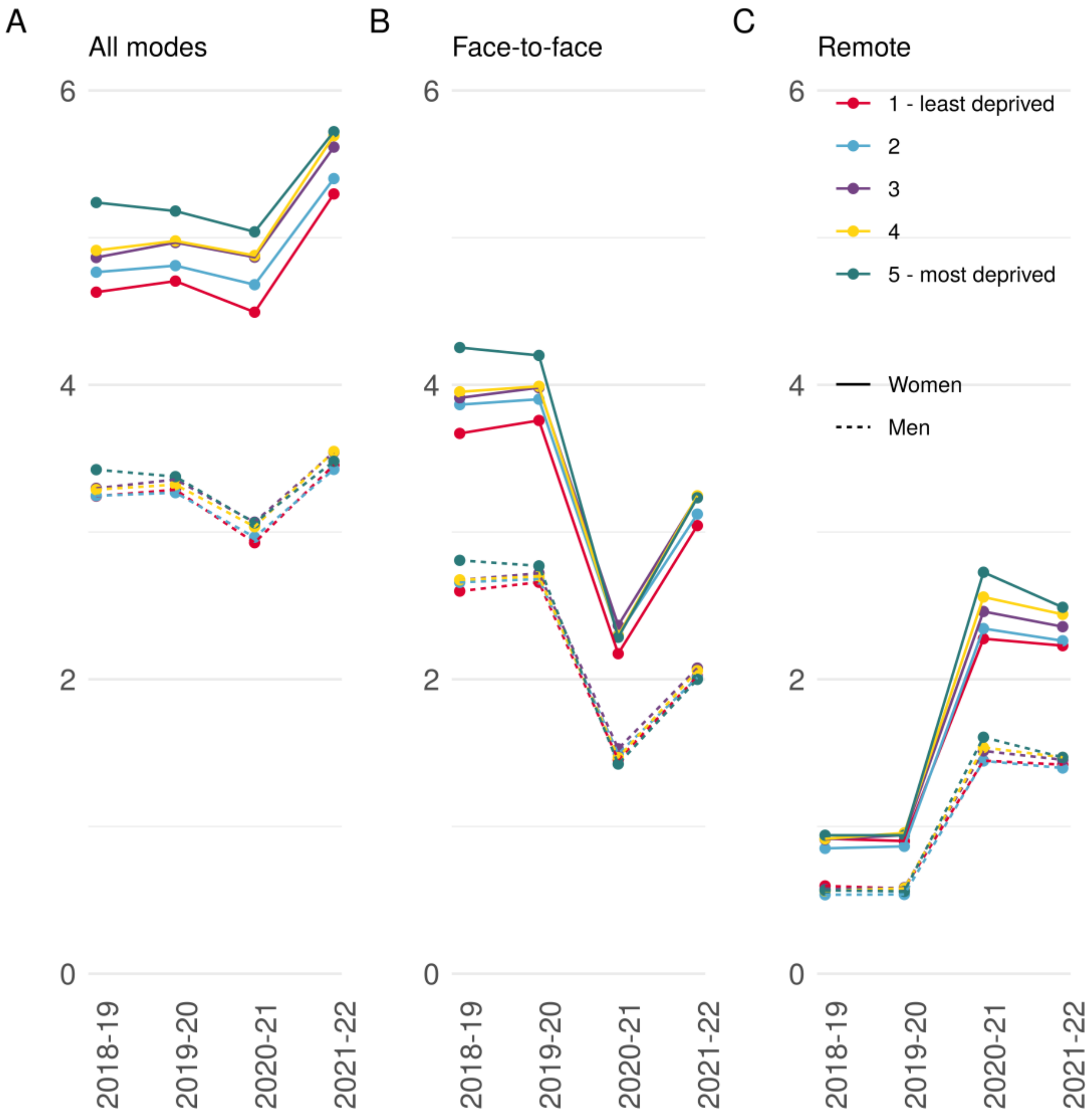
Consultation rates increased with increasing deprivation in each study period and were at their highest in the second year of the pandemic (2021-2022) for all deprivation quintiles (Figure 2A). Over the study period, consultation rates increased by 10% in the least deprived quintile, compared with 5% in the most deprived quintile, effectively narrowing relative differences between deprivation quintiles.

Rates of face-to-face consultations also tracked deprivation quintiles, with patients living in higher deprivation quintiles consistently having a greater rate of face-to-face consultations than those living in lower deprivation quintiles (Figure 2B). Inequalities in rates of face-to-face consultations also narrowed during the pandemic with differences between the most and least deprived quintiles decreasing from 0.37 per person-year in 2018-2019 to 0.06 per person-year in 2021-2022.

There was little difference in the rate of remote consultations across deprivation quintiles before the pandemic (Figure 2C). During the pandemic, inequalities in rates of remote consultations widened, and patients in the higher deprivation quintiles had consistently higher rates than those in lower deprivation quintiles. In 2021-2022, remote consultation rates ranged from 1.91 per person-year in the most deprived quintile to 1.78 per person-year in the least deprived quintile.

Figure 3 presents consultation rates by deprivation quintile and sex over time (Table S3 in Multimedia Appendix 1). Notably, men have consistently lower consultation rates than women. Trends in the rate of face-to-face and remote consultation rates across deprivation quintiles for women and men are similar to that observed for the overall population except that differences between the quintiles are narrower in men compared with women. For example, in 2018-2019, the difference in the consultation rate for men in the most deprived compared with the least deprived quintile was 0.18 per person-year, whereas it was 0.61 per person-year for women.

Figure 3. Predicted consultation rates over time by sex. (A) Overall consultations per year by deprivation. (B) Face-to-face consultations per year by deprivation. (C) Remote consultations per year by deprivation.



Discussion

We provided contemporary, nationally representative data on the rate and mode of clinical consultations by GP, nurse, and other health care professional staff in general practice, analyzing trends across patient age, sex, and deprivation over 1 million patient-years. These are the first comprehensive data to break down consultations by mode. Consultation rates in primary cares were higher in 2021-2022 than before the pandemic; the pandemic has led to a dramatic reconfiguration of consultation modes in English general practices, with substantial heterogeneity across patient age, sex, and deprivation.

The consultation rates calculated in this analysis are lower than those observed by Hobbs et al [19] between 2007 and 2014. However, they are consistent with national data, and due to variations in methodology, direct comparisons can be difficult. Nevertheless, patterns across age and sex are consistent with those reported by Hobbs et al.

Before the pandemic, more than 75% of consultations were face-to-face. After a sharp drop at the start of the pandemic, the proportion of face-to-face consultations has slowly increased but remains below prepandemic levels at the end of our study. This reflects a significant shift in general practice processes with the adoption of a hybrid approach with a mix of appointment types. These trends in consultation modes have been observed elsewhere but not in a large nationally representative sample [5,8,9,20-22].

The dip in the proportion of face-to-face consultations at the start of the pandemic likely reflects operational guidance at the time. The pandemic created a drive for change [23]. Before the pandemic, the implementation of remote consultations was slow due to a combination of technological difficulties, confidence, and concerns about quality of care and low uptake from patients [24]. The model of general practice changed as a result, with more practices using doctor-led remote triage as the access point for services [4]. The subsequent increase in face-to-face consultations, despite repeated waves of high COVID-19 infection rates, likely reflects the embedding of processes to manage infection control risk, a better understanding of the risks of the virus, as well as changing government guidance on lockdowns and restrictions, although it is expected face-to-face consultation will remain below prepandemic levels to reflect new ways of working [25,26].

In the first year of the pandemic, the overall drop in the year-on-year consultation rates was driven by children (aged 0-17 years) and older patients (aged ≥65 years). COVID-19 restrictions, such as home schooling, social distancing, and shielding recommendations, reduced the spread of some diseases especially non–COVID-19 respiratory infections, which may be reflected in the lower consultation rates [27]. On the other hand, some age groups saw an increase in consultation rates in the first year of the pandemic, which could be COVID-19–related consultations. Despite the drop in overall consultations for these groups, they sustained higher proportions of face-to-face consultations. This reflects an active prioritization of groups by health need, as they are more likely to present with issues that are more complex or difficult to assess remotely.

We observed consistently higher consultation rates in patients living in more deprived compared with less deprived areas. This tallies with evidence that patients living in more deprived areas are more likely to have higher health needs [28]. The pandemic likely had a negative impact on everybody's health; this is reflected in all quintiles of deprivation experiencing higher consultation rates in 2021-2022 compared to the prepandemic years. However, over the course of the pandemic, differences between the most and least deprived quintiles reduced by 5 percentage points, reflecting overall larger increases in consultation rates for patients living in the least deprived populations. During the pandemic, the health of patients in more deprived areas worsened faster compared with those patients in less deprived areas. Therefore, it seems unlikely that inequalities in health care needs have decreased over the pandemic.

Moreover, our findings may indicate that the demand-capacity gap has widened for patients living in more deprived compared to less deprived areas. Previous research highlighted that practices in more deprived areas manage 10% more needs adjusted for population size and receive 7% less funding adjusted for needs compared with practices in more affluent areas [29]. In addition, they faced greater workforce challenges with lower staff to patient ratios, lower recruitment rates, and more staff absences due to illness during the pandemic [7]. Overall, the evidence points to existing significant pressure limiting the capacity to stretch services further in response to pandemic- driven demand, exacerbating health inequalities. This indicates a need for greater support and investment in services in deprived areas.

At a patient level, reasons why patients in deprived areas had relatively fewer consultations during the pandemic may include a higher risk of being infected or getting severely ill with COVID-19 [30], a greater likelihood of lower levels of health and digital literacy [31], and an uneven economic impact [32]. These factors may have combined to result in both higher anxiety levels and greater concerns around contacting health care services due to risk of exposure to infection alongside greater difficulty in taking time off work to attend to health care needs. These findings correlate with reported patient overall experience of making an appointment where the fall in satisfaction rates was greater for practices in more deprived areas [33]. A greater reliance on technology could have also been a barrier to access for some patients in more deprived areas [34].

During the pandemic, the increase in remote consultations was reflected across all deprivation quintiles but was relatively larger for those living in the most deprived quintile—the difference in consultation rates between the most and least deprived populations increased by 13 percentage points for remote consultations in contrast to a decrease of 10 percentage points for face-to-face consultations. The reasons for the relatively greater use of remote consultations in more deprived populations during the pandemic are likely many and complex, relating to both practice- and patient-level effects. At the practice level, staff at practices in more deprived areas are more likely to be older or from ethnic minority backgrounds known to have been more impacted by COVID-19 [30,35-37]. Consequently, more staff might have been working from home in more deprived practices with more care being delivered remotely. At the patient level, socioeconomic variations in need could have played a role, as patients in more deprived quintiles are more likely to have long-term conditions and multimorbidity [38]. Patients living in deprived areas are both at higher risk of being infected and getting severely ill with COVID-19, which could lead to more remote consultations. Individual circumstances, such as ability to travel to the practice, could also impact the choice of modality [39].

Women having higher consultation rates than men has been widely reported [11]. Interestingly, sex differences in consulting rates varied by deprivation status, reflecting a greater socioeconomic gradient in consulting rates among women than men. This finding has been observed elsewhere [40] and could be a consequence of differences in the number of comorbidities between women and men, which also increases with increasing deprivation [41].

Digital technology will play an increasing role in general practice, and a deeper understanding of what mode of consultation suits particular patients' needs and preferences is continuing to develop [42]. There is limited evidence on differences in clinical outcomes and patient satisfaction between remote and face-to-face consultations. However, a recent systematic review [43] suggests that overall system design, including the triage process and technology interface, is key to good outcomes especially staff workload, but more research is needed on a broad range of outcomes. Most remote consultations in England were telephone based and easier to access than other more technology-heavy remote consultation options, such as video calls. Despite the potential of digital tools to improve service delivery, it is vital to continuously improve their accessibility and usability and monitor the impact on health inequalities.

This study has several strengths. Using a rich clinical nationally representative patient-level data set that includes more than 1 million patient-years allows us to explore consultation modes by different patient groups. Comparisons between deprivation quintiles were adjusted for age and sex, making comparisons more accurate. We developed a novel method to derive consultation mode that combines information from the time of booking with what happened during the consultation to improve the accuracy of our results.

Data quality is one of the main limitations of this study, as the consultation mode is not consistently recorded in the data. The consultation mode was derived using a new method that has not been externally validated. The default assumption was that the consultation was face-to-face; we are therefore more likely to have underestimated the number of remote consultations. However, we found a strong increase in the use of remote consultations during the pandemic, with a peak around 75% of consultations being remote, which is consistent with other sources giving us confidence in our methods [14,16].

In summary, we found that general practice is delivering more consultations than ever, but that trends in consultation rates and modes across deprivation quintiles indicate exacerbation of existing socioeconomic inequalities. This general increase in the rate of consultations matches other evidence and reports from GPs, but further research is needed to ensure that consultation rates match health needs across deprivation groups.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary files.
[DOCX File , 205 KB - publichealth_v9i1e44944_app1.docx]

References

1. The NHS long term plan. NHS. URL: https://www.longtermplan.nhs.uk/ [accessed 2023-03-13]
2. GP Contract. NHS. 2022. URL: https://www.england.nhs.uk/gp/investment/gp-contract/gp-contract-documentation-2018-19/ [accessed 2022-10-15]
3. Bakhai M, Croney L, Waller O, Henshall NF. Using online consultations in primary care implementation toolkit. NHS England. 2020. URL: https://www.england.nhs.uk/publication/using-online-consultations-in-primary-care-implementation-toolkit/ [accessed 2023-04-17]
4. Marshall M, Howe A, Howsam G, Mulholland M, Leach J. COVID-19: a danger and an opportunity for the future of general practice. Br J Gen Pract 2020 Jun;70(695):270-271 [FREE Full text] [doi: 10.3399/bjgp20X709937] [Medline: 32393503]
5. Murphy M, Scott LJ, Salisbury C, Turner A, Scott A, Denholm R, et al. Implementation of remote consulting in UK primary care following the COVID-19 pandemic: a mixed-methods longitudinal study. Br J Gen Pract 2021;71(704):e166-e177 [FREE Full text] [doi: 10.3399/BJGP.2020.0948] [Medline: 33558332]
6. Velasquez D, Mehrotra A. Ensuring the growth of telehealth during COVID-19 does not exacerbate disparities in care. Health Affairs Blog. 2020. URL: https://www.healthaffairs.org/do/10.1377/forefront.20200505.591306/full/ [accessed 2023-04-13]
7. Nussbaum C, Massou E, Fisher R, Morciano M, Harmer R, Ford J. Inequalities in the distribution of the general practice workforce in England: a practice-level longitudinal analysis. BJGP Open 2021 Oct;5(5):10 [FREE Full text] [doi: 10.3399/BJGPO.2021.0066] [Medline: 34404634]
8. Parker RF, Figures EL, Paddison CA, Matheson JI, Blane DN, Ford JA. Inequalities in general practice remote consultations: a systematic review. BJGP Open 2021 Jun;5(3):12 [FREE Full text] [doi: 10.3399/BJGPO.2021.0040] [Medline: 33712502]
9. Howells K, Amp M, Burrows M, Brown J, Brennan R, Dickinson J, et al. Remote primary care during the COVID-19 pandemic for people experiencing homelessness: a qualitative study. Br J Gen Pract 2022 Jul;72(720):e492-e500 [FREE Full text] [doi: 10.3399/BJGP.2021.0596] [Medline: 35379604]
10. Joy M, McGagh D, Jones N, Liyanage H, Sherlock J, Parimalanathan V, et al. Reorganisation of primary care for older adults during COVID-19: a cross-sectional database study in the UK. Br J Gen Pract 2020 Aug;70(697):e540-e547 [FREE Full text] [doi: 10.3399/bjgp20X710933] [Medline: 32661009]

11. Jaffe D, Lee L, Huynh S, Haskell T. Health Inequalities in the Use of Telehealth in the United States in the Lens of COVID-19. Popul Health Manag 2020 Oct;23(5):368-377 [FREE Full text] [doi: 10.1089/pop.2020.0186] [Medline: 32816644]
12. Chang JE, Lai AY, Gupta A, Nguyen AM, Berry CA, Shelley DR. Rapid transition to telehealth and the digital divide: implications for primary care access and equity in a post-COVID era. Milbank Q 2021 Jun;99(2):340-368 [FREE Full text] [doi: 10.1111/1468-0009.12509] [Medline: 34075622]
13. CPRD Aurum May 2022 dataset. CPRD. 2022. URL: https://cprd.com/cprd-aurum-may-2022-dataset [accessed 2022-11-15]
14. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. Int J Epidemiol 2019 Dec 01;48(6):1740-1740g [FREE Full text] [doi: 10.1093/ije/dyz034] [Medline: 30859197]
15. Watt T, Sullivan R, Aggarwal A. Primary care and cancer: an analysis of the impact and inequalities of the COVID-19 pandemic on patient pathways. BMJ Open 2022 Mar 24;12(3):e059374 [FREE Full text] [doi: 10.1136/bmjopen-2021-059374] [Medline: 35332047]
16. Carey I, Hosking F, Harris T, DeWilde S, Beighton C, Cook D. An evaluation of the effectiveness of annual health checks and quality of health care for adults with intellectual disability: an observational study using a primary care database. HSDR 2017:1-169 [FREE Full text] [doi: 10.3310/hsdr05250] [Medline: 28930382]
17. The R Project for Statistical Computing. 2003. URL: https://www.r-project.org/ [accessed 2022-10-26]
18. Lüdecke D. Ggeffects: tidy data frames of marginal effects from regression models. J Open Source Softw 2018;3(26):1-5 [FREE Full text] [doi: 10.21105/joss.00772]
19. Baral SD, Poteat T, Strömdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. Lancet Infect Dis 2013 Mar;13(3):214-222. [doi: 10.1016/s1473-3099(12)70315-8]
20. Green M, McKee M, Katikireddi S. Remote general practitioner consultations during COVID-19. Lancet Digit Health 2022 Jan;4(1):e7 [FREE Full text] [doi: 10.1016/S2589-7500(21)00279-X] [Medline: 34952678]
21. Foley K, Maile E, Bottle A, Neale F, Viner R, Kenny S, et al. Impact of COVID-19 on primary care contacts with children and young people in England: longitudinal trends study 2015–2020. Br J Gen Pract 2022 Apr 04;72(720):e464-e471 [FREE Full text] [doi: 10.3399/BJGP.2021.0643] [Medline: 35667683]
22. Geraldine C, Dias A, Wolters A. Access to and delivery of general practice services. The Health Foundation. 2022. URL: https://www.health.org.uk/publications/access-to-and-delivery-of-general-practice-services [accessed 2023-04-13]
23. Ledger J, Bakhai M. The temporal dimensions of health technology adoption during the covid-19 pandemic: revisiting Roger's diffusionist innovation. In: Waring J, Denis JL, Reff Pedersen A, Tenbensel T, editors. Organising Care in a Time of Covid-19: Implications for Leadership, Governance and Policy. Cambridge, UK: Springer International Publishing; 2021:245.
24. Baines R, Tredinnick-Rowe J, Jones R, Chatterjee A. Barriers and enablers in implementing electronic consultations in primary care: scoping review. J Med Internet Res 2020 Nov 12;22(11):e19375 [FREE Full text] [doi: 10.2196/19375] [Medline: 33035177]
25. Brown J. Coronavirus: a history of English lockdown laws. House of Commons Library. URL: https://commonslibrary.parliament.uk/research-briefings/cbp-9068/ [accessed 2021-05-09]
26. Turner A, Scott A, Horwood J, Salisbury C, Denholm R, Scott L, et al. Maintaining face-to-face contact during the COVID-19 pandemic: a longitudinal qualitative investigation in UK primary care. BJGP Open 2021 Jul 13;5(5):BJGPO.2021.0036 [FREE Full text] [doi: 10.3399/bjgpo.2021.0036]
27. Rezel-Potts E, L'Esperance V, Gulliford MC. Antimicrobial stewardship in the UK during the COVID-19 pandemic: a population-based cohort study and interrupted time-series analysis. Br J Gen Pract 2021 May;71(706):e331-e338 [FREE Full text] [doi: 10.3399/BJGP.2020.1051] [Medline: 33690150]
28. Cookson R, Propper C, Asaria M, Raine R. Socio-Economic Inequalities in Health Care in England. Fiscal Studies 2016 Nov 21;37(3-4):371-403. [doi: 10.1111/j.1475-5890.2016.12109]
29. Level or not? The Health Foundation. 2020 Sep. URL: https://www.health.org.uk/publications/reports/level-or-not [accessed 2022-11-15]
30. Phiri P, Delanerolle G, Al-Sudani A, Rathod S. COVID-19 and Black, Asian, and minority ethnic communities: a complex relationship without just cause. JMIR Public Health Surveill 2021 Feb 01;7(2):e22581 [FREE Full text] [doi: 10.2196/22581] [Medline: 33481752]
31. Simpson R, Knowles E, O'Cathain A. Health literacy levels of British adults: a cross-sectional survey using two domains of the Health Literacy Questionnaire (HLQ). BMC Public Health 2020 Nov 30;20(1):1819 [FREE Full text] [doi: 10.1186/s12889-020-09727-w] [Medline: 33256670]
32. Francis-Devine B. Poverty in the UK: statistics. House of Commons Library. 2023 May 06. URL: https://researchbriefings.files.parliament.uk/documents/SN07096/SN07096.pdf [accessed 2022-11-15]
33. Public perceptions of health and social care (November–December 2021). The Health Foundation. 2022. URL: https://www.health.org.uk/publications/public-perceptions-of-health-and-social-care-november-december-2021 [accessed 2022-10-27]

34. Davies A, Honeyman M, Gann B. Addressing the digital inverse care law in the time of COVID-19: Potential for digital technology to exacerbate or mitigate health inequalities. J Med Internet Res 2021 Apr 07;23(4):e21726 [FREE Full text] [doi: 10.2196/21726] [Medline: 33735096]
35. Morales DR, Ali SN. COVID-19 and disparities affecting ethnic minorities. Lancet 2021 May 08;397(10286):1684-1685. [doi: 10.1016/S0140-6736(21)00949-1] [Medline: 33939952]
36. Chattopadhyay I, Davies G, Adhiyaman V. The contributions of NHS healthcare workers who are shielding or working from home during COVID-19. Future Healthc J 2020 Oct 16;7(3):e57-e59 [FREE Full text] [doi: 10.7861/fhj.2020-0096] [Medline: 33094257]
37. Ethnic minority-led GP practices: impact and experience of CQC regulation. CQC. URL: https://www.cqc.org.uk/publications/themed-work/ethnic-minority-led-gp-practices-impact-experience-cqc-regulation [accessed 2022-07-18]
38. Head A, Fleming K, Kypridemos C, Schofield P, Pearson-Stuttard J, O'Flaherty M. Inequalities in incident and prevalent multimorbidity in England, 2004-19: a population-based, descriptive study. Lancet Healthy Longev 2021 Aug;2(8):e489-e497 [FREE Full text] [doi: 10.1016/S2666-7568(21)00146-X] [Medline: 36097998]
39. Rosen R, Wieringa S, Greenhalgh T, Leone C, Rybczynska-Bunt S, Hughes G. Clinical risk in remote consultations in general practice: findings from in-COVID-19 pandemic qualitative research. BJGP Open 2022;6(3):1-11 [FREE Full text] [doi: 10.3399/BJGPO.2021.0204]
40. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? An analysis of routinely collected UK general practice data. BMJ Open 2013 Aug 19;3(8):e003320 [FREE Full text] [doi: 10.1136/bmjopen-2013-003320] [Medline: 23959757]
41. Watt T, Raymond A, Rachet-Jacquet L. Quantifying health inequalities in England. The Health Foundation. 2022. URL: https://www.health.org.uk/news-and-comment/charts-and-infographics/quantifying-health-inequalities [accessed 2023-04-17]
42. Beech J. What do the public think are the biggest challenges facing general practice? BMJ Internet. BMJ 2022:377 [FREE Full text] [doi: 10.1136/bmj.o1149]
43. Darley S, Coulson T, Peek N, Moschogianis S, van der Veer SN, Wong D, et al. Understanding how the design and implementation of online consultations affect primary care quality: systematic review of evidence with recommendations for designers, providers, and researchers. J Med Internet Res 2022 Oct 24;24(10):e37436 [FREE Full text] [doi: 10.2196/37436] [Medline: 36279172]

Abbreviations

CPRD: Clinical Practice Research Datalink
GP: general practitioner
IMD: Index of Multiple Deprivation

Edited by A Mavragani, T Sanchez; submitted 12.12.22; peer-reviewed by C Pearce, A Yazdanian; comments to author 23.01.23; revised version received 03.02.23; accepted 14.03.23; published 02.05.23.

Please cite as:
Vestesson EM, De Corte KLA, Crellin E, Ledger J, Bakhai M, Clarke GM
Consultation Rate and Mode by Deprivation in English General Practice From 2018 to 2022: Population-Based Study
JMIR Public Health Surveill 2023;9:e44944
URL: https://publichealth.jmir.org/2023/1/e44944
doi:10.2196/44944
PMID:37129943

Original Paper

The Role of Metabolic Dysfunction–Associated Fatty Liver Disease in Developing Chronic Kidney Disease: Longitudinal Cohort Study

Suosu Wei[1], MM; Jian Song[2], PhD; Yujie Xie[3], MM; Junzhang Huang[4], MM; Jianrong Yang[5], BM

[1]Department of Scientific Cooperation of Guangxi Academy of Medical Sciences, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[2]Institute of Cardiovascular Diseases of Guangxi Academy of Medical Sciences, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[3]Department of Breast and Thyroid Surgery, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[4]Department of Hepatobiliary, Pancreas and Spleen Surgery, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[5]Institute of Health Management of Guangxi Academy of Medical Sciences, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

Corresponding Author:
Jianrong Yang, BM
Institute of Health Management of Guangxi Academy of Medical Sciences
People's Hospital of Guangxi Zhuang Autonomous Region
No 6 Taoyuan Road
Nanning, 530021
China
Phone: 86 13978630882
Email: yjr@gxams.org.cn

Abstract

Background: The association between metabolic dysfunction–associated fatty liver disease (MAFLD) and chronic kidney disease (CKD) is unclear.

Objective: This longitudinal cohort study aimed to test whether MAFLD plays an important role in the development of CKD.

Methods: This cohort study included 41,246 participants who had undergone 3 or more health examinations from 2008 to 2015 at the People's Hospital of Guangxi Zhuang Autonomous Region, China. Participants were categorized into 2 groups according to whether they presented with or without MAFLD. The occurrence of new-onset CKD was stated as either an estimated glomerular filtration rate of <60 mL/min per 1.73 m^2 or a higher level of albuminuria during their follow-up appointment. The association between MAFLD and CKD was evaluated using a Cox regression method.

Results: Of the 41,246 participants, 11,860 (28.8%) were diagnosed with MAFLD. Over the course of the 14-year follow-up (median 10.0 years), 5347 (13%) participants experienced a new incident of CKD (135.73 per 10,000 person-years). MAFLD was discovered as an important risk factor for new incidents of CKD (hazard ratio 1.18, 95% CI 1.11-1.26) by using the multivariable Cox proportional hazard regression model. When stratified by gender, the adjusted hazard ratio for the incidence of CKD in men and women with MAFLD were 1.16 (95% CI 1.07-1.26) and 1.32 (95% CI 1.18-1.48), respectively. According to the subgroup analysis results, after adjusting for confounding factors, the MAFLD-related CKD risk was greater in men aged <60 years ($P_{\text{interaction}}$=.001) and in those with combined dyslipidemia ($P_{interaction}$=.02), but this relationship was not found in women (all $P_{interaction}$>.05).

Conclusions: MAFLD plays an important role in the development of new incidents of CKD in the long run.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2200058543; https://www.chictr.org.cn/showproj.html?proj=153109

(JMIR Public Health Surveill 2023;9:e45050) doi:10.2196/45050

KEYWORDS

chronic kidney disease; risk factor; kidney; renal; cohort study; metabolic dysfunction–associated fatty liver disease; MAFLD; incidence rate; incidence; liver; chronic disease; prevalence; association

Introduction

Nowadays, kidney disease poses a considerable burden on global health, and it affects global morbidity and mortality. In particular, there are about 10% of adults worldwide who experience chronic kidney disease (CKD) directly, which constitutes a big public health issues [1]. Global health policy makers should pay more attention to CKD, which is a preventable and treatable condition [1]. Currently, more than 2.5 million people receive kidney transplants, and this will reach 5.4 million by 2030. CKD can develop into end-stage renal disease, which can be fatal if renal replacement therapy is not an option [2]. The prevalence of CKD is 10.8% in China [3]. Therefore, it is of great public health importance to explore the related risk factors of CKD and take active preventive measures.

Nonalcoholic fatty liver disease (NAFLD) is characterized by hepatic steatosis without excessive alcohol consumption or other competing factors and fat accumulation in more than 5% of hepatocytes [4]. About 25% of adults worldwide currently have NAFLD [5]. Approximately 30% of Chinese adults have NAFLD, and rural areas are less likely to have cases than urban areas. In addition, men are more likely to have it than women [6]. Obesity, hypertension, and diabetes are metabolic risk factors for both NAFLD and CKD [7]. In patients who have NAFLD, the prevalence rate of CKD is 20% to 55%, and this is reduced to 5% to 35% for patients who do not have NAFLD. Some research results have found a correlation between NAFLD and CKD, which showed that NAFLD increases the risk of developing CKD on its own [8-12].

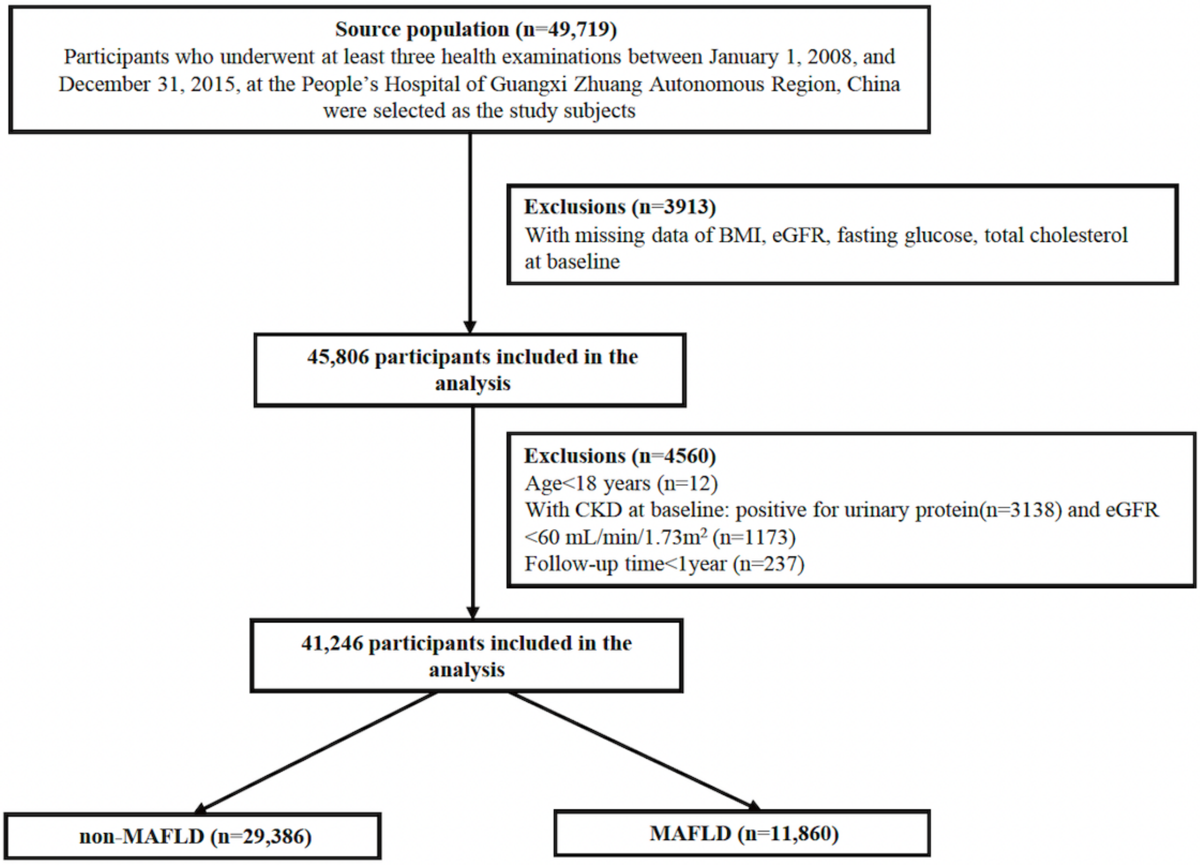
Due to the limitations in the definition for NAFLD in the clinical setting, in 2020, an international group of experts established the conception of metabolic dysfunction–associated fatty liver disease (MAFLD). This incorporated a diagnostic criteria of MAFLD that was based on fatty liver disease combined type 2 diabetes mellitus (T2DM) as well as either overweight or obesity or metabolic dysfunction [13,14]. The definition of MAFLD reflects that this type of liver disease is a series of complex metabolic disorders that do not rule out either excessive drinking or other chronic liver diseases [15,16]. It has been shown that patients with MAFLD have more metabolic complications than those with NAFLD and they have increased risks of advanced liver fibrosis and CKD [17,18]. However, due to several inconsistent reports regarding the correlation between MAFLD and CKD, this needs to be clarified. Therefore, we analyzed the relationship between MAFLD and new-onset CKD by conducting a 14-year retrospective cohort study.

Methods

Study Design and Participants

We used an active health management platform to construct a retrospective cohort study, and the protocol of study was registered on the Chinese Clinical Trial Registry (ChiCTR2200058543) [19]. Briefly, the active health management platform used an advanced medical data management system to manage patients and connected and indexed all the diagnostic and treatment records held at the hospital. It contained the outpatient, inpatient, and physical examination data with respect to the diagnosis of the patients. In addition, the treatment data, test reports, examination reports, electronic medical records, and other medical data of the outpatient, inpatient, and physical examinations were also recorded. All the medical information could be accessed from this platform, and when a patient comes to the clinic, the information is automatically integrated into this platform. This study enrolled participants who had received at least 3 health examinations and had an abdominal ultrasonography performed from 2008 to 2015 at the People’s Hospital of Guangxi Zhuang Autonomous Region, China (n=49,719). The date of the first health examination was defined as the index date. The exclusion criteria in this study were participants with (1) an estimated glomerular filtration rate (eGFR) <60 mL/min per 1.73 m^2 or a high level of proteinuria at baseline, (2) missing either demographic or laboratory data, and (3) follow-up time <1 year. After exclusion, 41,246 participants were recruited into this study (Figure 1).

Figure 1. Flow chart of the selection of participants in the active health management platform. CKD: chronic kidney disease; eGFR: estimate glomerular filtration rate; MAFLD: metabolic dysfunction–associated fatty liver disease.



Ethics Approval

The guidelines of the Declaration of Helsinki (6th revision, 2008) were followed in this study. The People's Hospital of Guangxi Zhuang Autonomous Region Ethics Committee approved this study (KY-QT-202203). As the data were anonymized, individual informed consent was not needed for this study.

Ascertainment of MAFLD

All enrolled participants were diagnosed with a fatty liver by hepatic ultrasound using the Asia-Pacific guidelines as a diagnostic guideline [20]. MAFLD was diagnosed if participants had a fatty liver and at least one the following criteria was met:

1. BMI ≥23 kg/m^2
2. Diagnosed with T2DM
3. BMI <23 kg/m^2 and 2 or more metabolic abnormalities for Asian men and women, including a waist circumference ≥90/80 cm, blood pressure ≥130/85 mmHg, plasma triglyceride (TG) ≥1.70 mmol/L, plasma high-density lipoprotein cholesterol (HDL-C) <1.0 mmol/L for men and <1.3 mmol/L for women, prediabetes (fasting glucose level of 5.6-6.9 mmol/L, 2-hour postload glucose level of 7.8-11.1 mmol/L, or hemoglobin A_{1c} of 5.7%-6.4%), homeostasis model assessment of insulin resistance score ≥2.5 (this study did not include these data), and a plasma high-sensitivity C-reactive protein level >2 mg/L (this study not include these data) [14].

Measurements

Baseline data including sociodemographic, anthropometric, and clinical information and laboratory results were retrieved from the active health management platform. If any data were missing for the index date, then the data relating to the nearest index date were obtained. The indicator of renal function in this study was eGFR, which was based on the CKD Epidemiology Collaboration equation [21]. The covariates assessed in this study included BMI, T2DM, hypertension, and dyslipidemia. BMI was stratified into normal weight (<23 kg/m^2) and overweight or obese (≥23 kg/m^2). Participants with systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg were considered to have hypertension. Participants with fasting blood glucose ≥7.0mmol/L or hemoglobin A_{1c} ≥ 6.5% were considered to have T2DM. Participants with total cholesterol ≥5.2 mmol/L, TG ≥1.7 mmol/L, low-density lipoprotein cholesterol (LDL-C) ≥3.4 mmol/L, or HDL-C <1.0mmol/L were considered to have dyslipidemia. All the laboratory tests were conducted within the health examination hospital.

Study Outcome

All participants were followed prospectively until their last health examinations and medical visit in the period from 2016 to 2022. As a clinical end point, participants who have eGFR <60 mL/min per 1.73 m^2 or 2 incidents of albuminuria were considered to be a new incident of CKD during the follow-up period.

Statistical Analysis

For categorical variables, we presented counts and percentages, and for continuous variables, we presented medians and IQRs. For baseline data between participants with and without MAFLD, we used Mann-Whitney *U* and chi-square tests to compare continuous data and categorical data, respectively. The incidence of CKD was calculated by dividing the total number of newly diagnosed CKD cases by the total number of person-years contributed by participants during follow-up. We used Kaplan-Meier analysis to estimate the cumulative incidence of CKD between patients with and without MAFLD and log-rank tests to compare the differences between the 2 groups. We used Cox hazard models to calculation the hazard ratios (HRs) and 95% CIs for the risk between MAFLD and CKD. We used Schoenfeld residuals to evaluate the proportional hazards assumption. We selected covariates according to the reverse selection procedure and confounding factors reported in the previous literature [18]. Age and gender were adjusted in model 1. In addition, comorbidities of hypertension, overweight or obesity, dyslipidemia, and T2DM were adjusted in model 2. Furthermore, LDL-C, alanine aminotransferase (ALT), serum aspartate aminotransferase (AST), and serum creatinine were adjusted in model 3. Finally, stratified analysis was conducted in predesignated subgroups according to gender; age; overweight or obesity; and the presence of hypertension, dyslipidemia, and T2DM. We used R software packages (version 3.6.3; R Foundation for Statistical Computing) to conduct statistical analysis in this study. All test for *P* values were 2-tailed, and *P*<.05 was considered statistically significant.

Results

Baseline Characteristics

After applying the exclusion criteria, 41,246 participants were finally included for analysis in this cohort study. Among the included participants, 11,860 (28.8%) were diagnosed with MAFLD. These participants were more likely to be older; be men; have T2DM, hypertension, or dyslipidemia; and be either overweight or obese (all *P*<.001). The participants with MALFD also had higher BMI, fasting glucose, total cholesterol, LDL-C levels, and TG and lower HDL-C levels, gamma-glutamyl transferase, ALT, AST, and eGFR than participants without MAFLD (all *P*<.001; Table 1).

Table 1. Baseline characteristics of the participants in this study.

Characteristic	Total (N=41,246)			Men (n=22,417)			Women (n=18,829)		
	Non-MAFLD[a] (n=29,386)	MAFLD (n=11,860)	*P* value	Non-MAFLD (n=13,355)	MAFLD (n=9062)	*P* value	Non- MAFLD (n=16,031)	MAFLD (n=2798)	*P* value
Age (years), median (IQR)	37.0 (29.0-46.0)	45.0 (37.0-53.0)	<.001	37.0 (29.0-47.0)	43.0 (35.0-50.0)	<.001	37.0 (29.0-45.0)	50.0 (43.0-57.0)	<.001
Gender, men, n (%)	13,355 (45.4)	9062 (76.4)	<.001	N/A[b]	N/A	N/A	N/A	N/A	N/A
BMI (kg/m^2), median (IQR)	23.5 (21.6-25.4)	26.7 (25.0-28.6)	<.001	24.3 (22.6-26.1)	26.9 (25.3-28.7)	<.001	22.7 (21.0-24.5)	26.1 (24.4-28.4)	<.001
Waist circumference (cm), median (IQR)	83.0 (77.0-89.0)	93.0 (88.0-98.0)	<.001	87.0 (83.0-92.0)	94.0 (90.0-99.0)	<.001	79.0 (74.0-85.0)	87.0 (84.0-93.0)	<.001
Systolic blood pressure (mmHg), median (IQR)	116.0 (107.0-127.0)	129.0 (119.0-140.0)	<.001	122.0 (114.0-132.0)	129.0 (120.0-140.0)	<.001	111.0 (103.0-121.0)	128.0 (116.0-141.0)	<.001
Diastolic blood pressure (mmHg), median (IQR)	71.0 (65.0-79.0)	80.0 (73.0-87.0)	<.001	74.0 (68.0-82.0)	81.0 (74.0-88.0)	<.001	69.0 (62.0-76.0)	76.0 (70.0-85.0)	<.001
Diabetes mellitus, n (%)	1393 (4.7)	2234 (18.8)	<.001	831 (6.2)	1652 (18.2)	<.001	562 (3.5)	582 (20.8)	<.001
Hypertension, n (%)	2986 (10.2)	3665 (30.9)	<.001	1935 (14.5)	2822 (31.1)	<.001	1051 (6.6)	843 (30.1)	<.001
Dyslipidemia, n (%)	13,321 (45.3)	9110 (76.8)	<.001	7346 (55)	7228 (79.8)	<.001	5,975 (37.3%)	1882 (67.3)	<.001
Overweight or obesity, n (%)	16,699 (56.8)	11,362 (95.8)	<.001	9382 (70.3)	8807 (97.2)	<.001	7317 (45.6)	2555 (91.3)	<.001
Fasting glucose (mmol/L), median (IQR)	5.0 (4.6-5.3)	5.3 (4.9-5.8)	<.001	5.0 (4.7-5.4)	5.3 (4.9-5.8)	<.001	5.0 (4.6-5.3)	5.4 (5.0-5.9)	<.001
ALT[c] (unit/L), median (IQR)	16.0 (12.0-23.0)	26.0 (19.0-38.0)	<.001	20.0 (15.0-28.0)	29.0 (21.0-41.0)	<.001	14.0 (11.0-19.0)	20.0 (15.0-27.0)	<.001
AST[d] (unit/L), median (IQR)	20.0 (17.0-24.0)	23.0 (20.0-28.0)	<.001	22.0 (18.0-26.0)	24.0 (20.0-29.0)	<.001	19.0 (16.0-22.0)	21.0 (18.0-25.0)	<.001
GGT[e] (unit/L), median (IQR)	18.0 (13.0-27.0)	34.0 (23.0-54.0)	<.001	25.0 (18.0-37.0)	38.0 (26.0-60.0)	<.001	14.0 (11.0-18.0)	21.0 (16.0-29.0)	<.001
TC[f] (mmol/L), median (IQR)	4.8 (4.2-5.4)	5.3 (4.7-5.9)	<.001	4.8 (4.3-5.5)	5.3 (4.7-5.9)	<.001	4.7 (4.2-5.3)	5.3 (4.7-6.0)	<.001
HDL[g] cholesterol (mmol/L), median (IQR)	1.4 (1.2-1.7)	1.2 (1.0-1.4)	<.001	1.3 (1.1-1.5)	1.1 (1.0-1.3)	<.001	1.5 (1.3-1.8)	1.3 (1.1-1.5)	<.001
LDL[h] cholesterol (mmol/L), median (IQR)	3.0 (2.5-3.5)	3.4 (3.0-4.0)	<.001	3.2 (2.6-3.6)	3.4 (3.0-4.0)	<.001	2.9 (2.4-3.4)	3.4 (3.0-4.0)	<.001
TG[i] (mmol/L), median (IQR)	1.0 (0.7-1.4)	1.8 (1.3-2.7)	<.001	1.1 (0.8-1.6)	1.9 (1.3-2.8)	<.001	0.8 (0.6-1.1)	1.5 (1.1-2.2)	<.001
Creatinine (µmol/L), median (IQR)	70.0 (59.0-84.0)	80.0 (69.0-90.0)	<.001	85.0 (78.0-93.0)	84.0 (77.0-92.0)	<.001	60.0 (55.0-66.0)	61.0 (55.0-67.8)	<.001
eGFR[j] (mL/min per 1.73 m^2), median (IQR)	104.7 (92.5-115.5)	97.0 (86.1-106.6)	<.001	98.3 (87.4-109.1)	96.2 (85.4-106.1)	<.001	109.7 (98.7-119.1)	99.8 (88.8-108.2)	<.001

[a] MAFLD: metabolic dysfunction–associated fatty liver disease.

[b] N/A: not applicable.

[c] ALT: alanine aminotransferase.

[d] AST: aspartate transaminase.

[e] GGT: gamma-glutamyl transferase.

[f] TC: total cholesterol.

[g] HDL: high-density lipoprotein.

[h] LDL: low-density lipoprotein.

[i] TG: triglyceride.

[j] eGFR: estimated glomerular filtration rate.

Incidence of CKD Between 14-year Follow-up

The median follow-up of this cohort study was 10.0 (IQR 7.7-12.0) years. During 393,933.1 person-years of follow-up, 5347 (13%) of the 41,246 participants developed new incidents of CKD (135.73 per 10,000 person-years, 95% CI 132.1-139.4). In the MAFLD group, the rate of new incidents of CKD was 177.9 (95% CI 170.2-185.8) per 10,000 person-years, whereas in non-MAFLD group, the rate of new incidents of CKD was 118.8 (95% CI 114.9-122.9) per 10,000 person-years. When stratified by gender, in the men, the new incidence of CKD was 167.0 (95% CI 158.4-175.8) per 10,000 person-years in the MAFLD group and 114.6 (95% CI 108.8-120.6) per 10,000 person-years in non-MAFLD group. In women, the new incidence of CKD was 212.5 (95% CI 195.5-230.6) per 10,000 person-years in the MAFLD group and 122.4 (95% CI 116.9-128.1) per 10,000 person-years in non-MAFLD group (Table 2). Kaplan-Meier analysis indicated that the cumulative incidence rate of CKD in the MAFLD group was significantly higher than that in the non-MAFLD group (*P*<.001). This result was similar to observed when the participants were stratified by gender (Figure 2).

Table 2. The incidence chronic kidney disease in participants with and without metabolic dysfunction–associated fatty liver disease.

Group	Events, n	Person-years	Incidence (per 10,000 person-years; 95% CI)	Hazard ratio (95% CI)					
				Model 1[a]	*P* value	Model 2[b]	*P* value	Model 3[c]	*P* value
Total									
Non-MAFLD[d]	3342	281,207.0	118.8 (114.9-122.9)	Reference		Reference		Reference	
MAFLD	2005	112,726.1	177.9 (170.2-185.8)	1.27 (1.19-1.34)	<.001	1.16 (1.09-1.24)	<.001	1.18 (1.11-1.26)	<.001
Men									
Non-MAFLD	1473	128,539.8	114.6 (108.8-120.6)	Reference		Reference		Reference	
MAFLD	1431	85,715.8	167.0 (158.4-175.8)	1.28 (1.19-1.38)	<.001	1.14 (1.05-1.24)	<.001	1.16 (1.07-1.26)	<.001
Woman									
Non-MAFLD	1869	152,667.2	122.4 (116.9-128.1)	Reference		Reference		Reference	
MAFLD	574	27,010.3	212.5 (195.5-230.6)	1.42 (1.28-1.57)	<.001	1.30 (1.16-1.45)	<.001	1.32 (1.18-1.48)	<.001

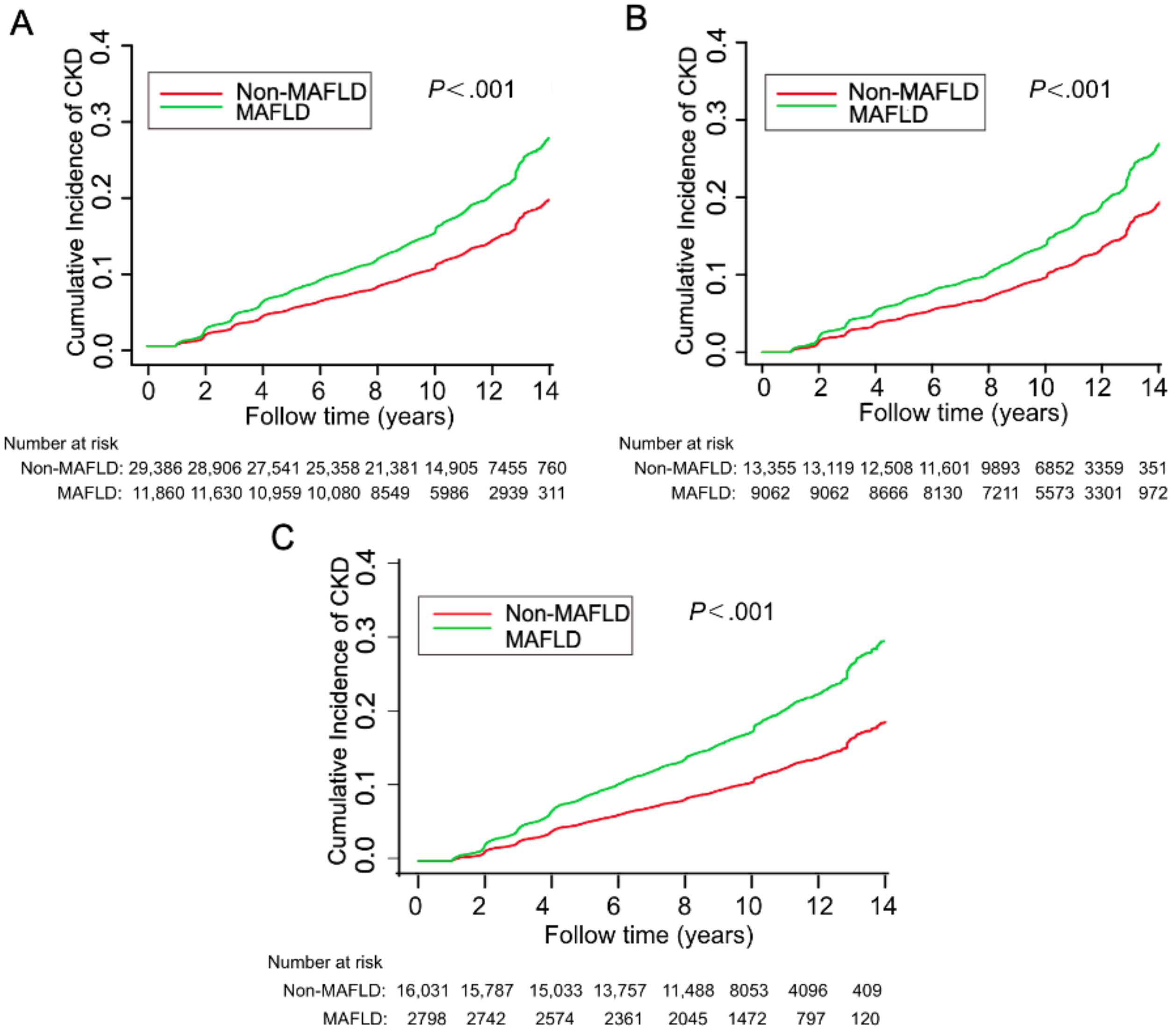
[a] Model 1: The total group was adjusted for age and gender. Men and woman groups were adjusted for age.

[b] Model 2: Model 1 with additional adjustments for hypertension, dyslipidemia, overweight or obesity, and diabetes mellitus.

[c] Model 3: Model 2 with additional adjustments for low-density lipoprotein, aspartate aminotransferase, alanine aminotransferase, and creatinine.

[d] MAFLD: metabolic dysfunction–associated fatty liver disease.

Figure 2. The cumulative incidence of chronic kidney disease (CDK) based on (A) the total number of participants, (B) men, and (C) women. MAFLD: metabolic dysfunction–associated fatty liver disease.



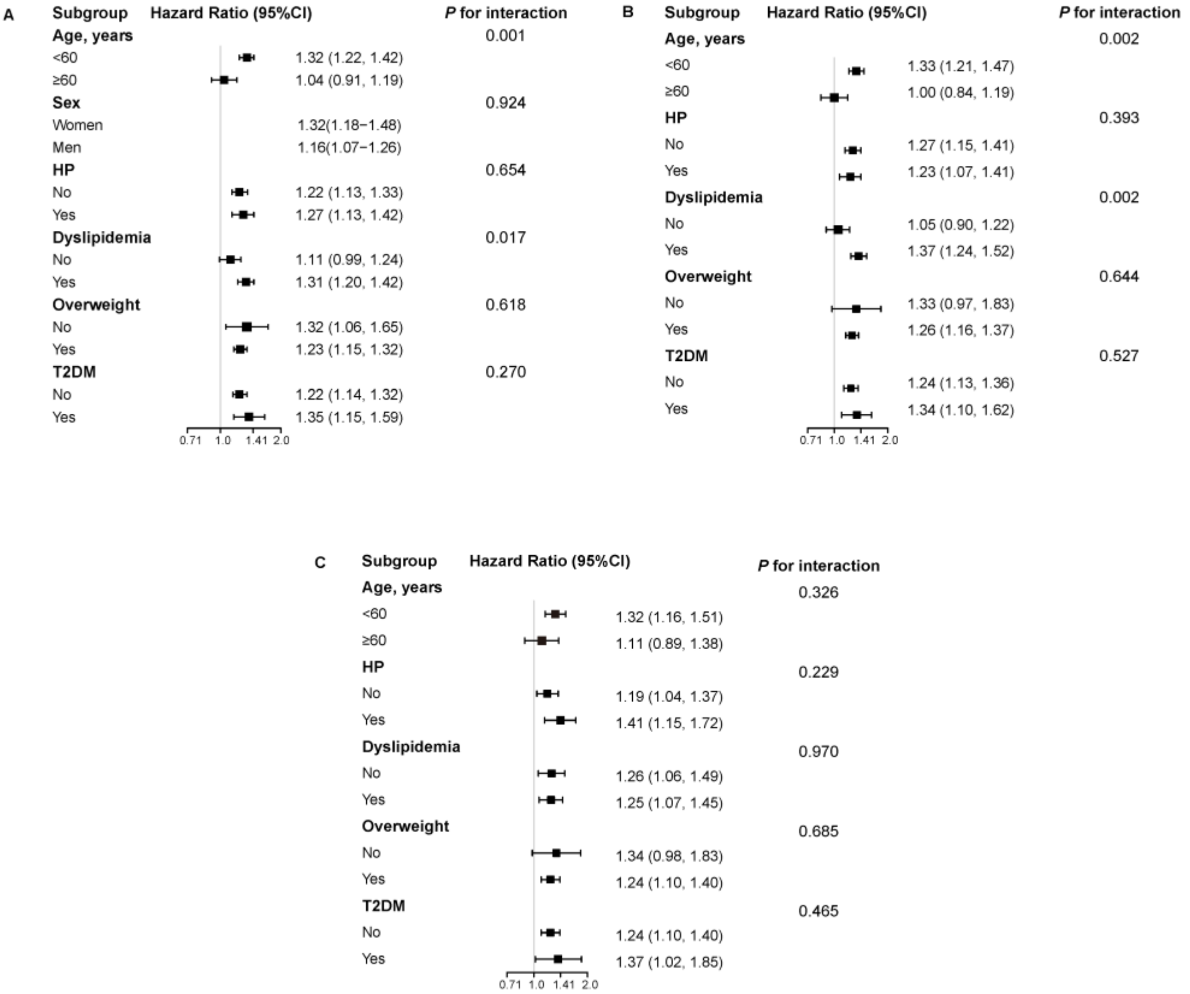
MAFLD Is a Risk Factor in the Development of New Incidence of CKD

When age and gender were adjusted using multivariable Cox proportional hazard model, the result showed that the participants with MAFLD had a 1.27-fold higher HR for the incidence of CKD than those without MAFLD (adjusted HR 1.27, 95% CI 1.19-1.34). Even after the additional adjustments for hypertension, dyslipidemia, overweight or obesity, T2DM, LDL, AST, ALT, and creatinine, participants with MAFLD still had a higher risk in the development of new incidents of CKD (adjusted HR 1.18, 95% CI 1.11-1.26). When stratified by gender and compared with those without MAFLD, based on the fully adjusted Cox proportional hazard model, in men, the HR for the incidence of CKD in participants with MAFLD was 1.16 (95% CI 1.07-1.26), and in women, it was 1.32 (95% CI 1.18-1.48; Table 2).

Subgroup Analysis

In the subgroup analyses, we observed that in the total and men groups, the MAFLD-related CKD risk was stronger in patients aged <60 years and in those with combined dyslipidemia ($P_{interaction}$=.001; Figure 3A-B), but this was not seen in the women group ($P_{interaction}$=.02; Figure 3C). However, the other subgroup factors did not show any significant interaction discrepancies (all *P*>.05; Figure 3).

Figure 3. Forest plots showing the effect of metabolic-dysfunction associated fatty liver disease on the incidence of chronic kidney disease based on (A) the total number of participants, (B) men, and (C) women. Each stratification was adjusted for all the factors (age, gender, hypertension, dyslipidemia, overweight or obesity, type 2 diabetes mellitus [T2DM], low-density lipoprotein, aspartate aminotransferase, alanine aminotransferase, and creatinine) except the stratification factor itself.



Discussion

Principal Findings

In this large, 14-year longitudinal cohort study, we found that participants with MAFLD have a higher risk for developing CKD than those without MAFLD. This was similar to the results found when the patients were stratified by gender. After adjusting for confounding factors such as demographics, laboratory parameters, and complications, these findings remained unchanged. Interestingly, we found that the MAFLD-related CKD risk was stronger in men aged <60 years and men with combined dyslipidemia but not in women.

This study provides further evidence to support the observations from several other studies regarding the relationship between MAFLD and CKD [18,22-24]. A cohort study with a median of 5.1 years of follow-up showed that participants with MAFLD had a 39% higher risk of new-onset CKD when compared to participants without MAFLD [18]. A cohort study with a median of 6.3 years of follow-up using a health examination database from a Japanese population similarly reported that compared to participants without MAFLD, participants with MAFLD had a 12% higher risk of new-onset CKD [22]. A cohort study with a median of 4.6 years of follow-up from China showed that participants with MAFLD had a 64% higher risk of incidents of CKD than those without MAFLD [24]. Another retrospective cohort study with a median of 4.6 years of follow-up from Japan also showed that those with MAFLD carried a 24% higher risk of incidents for CKD [23]. The results of this 14-year longitudinal study found that participants with MAFLD had a 18% higher risk of incidents of CKD when compared to those without MAFLD. For men and women, the risk of incidents for CKD changed to 16% and 32%, respectively. These findings are consistent with the aforementioned studies. In this study, as shown in a further subgroup analysis when metabolic factors were adjusted, the relationship between MAFLD and CKD remained unchanged. This may be due to the definition of MAFLD, which focuses more on the metabolic aspects of the disease.

The development of CKD can cause irreversible decline of kidney function, and this will lead to a need for dialysis and an increase in the number of deaths [1]. However, studies have shown that the epidemiology of CKD varies according to gender, and more women than men were affected by this disease, especially when stage G3 CKD is taken into consideration [25]. Therefore, when discussing the relationship between MAFLD and CKD, the data should be considered separately according to gender. The reason why women have a higher incidence of CKD may be related to the overdiagnosis of CKD by using the eGFR equation. It may also reflect the longer life expectancy observed in women [25,26]. Our results show that men have a higher prevalence of MAFLD, however, women have a higher risk of incidents of CKD. In addition, this study showed that men aged <60 years and men with combined dyslipidemia had a particular higher risk of CKD. However, this phenomenon has not been observed in women.

The gender difference observed may be due to the fact that there are some variables in men that may contribute to obesity, diabetes, and metabolic syndrome that are different in women. A previous study had shown that different lifestyle factors such as a high lipid diet and smoking are potential factors for gender differences in the progression of CKD [25]. On the other hand, this may be related to the lower age of men and higher age of women in the participants with MAFLD in this study. In addition, this study shown that the prevalence of dyslipidemia in men with MAFLD was also higher than that in women. Dyslipidemia can lead to inflammation, oxidative stress, and lipid toxicity, resulting in an impaired glomerular filtration barrier and proteinuria. The mechanism by which lipids can affect eGFR may be that a fatty liver damages the kidney through excessive very low–density lipoprotein secretion, which would induce atherosclerotic dyslipidemia [27]. The increased levels of circulating TG-rich lipoproteins and oxidized LDL-C can promote glomerular damage and mesangial cell proliferation. However, the potential mechanism of gender differences in the risk of CKD onset with dyslipidemia needs further study. This study shows the importance for the future development of more gender-specific interventions in medical practice. In particular, interventions should be given as early as possible, which might reduce the risk of incidents of CKD in men with MAFLD, especially in younger men with comorbid dyslipidemia. It is suggested that the renal function of patients with MAFLD should be evaluated regularly in clinical practice. A multidisciplinary and person-centered approach can be taken to manage patients with both MAFLD and CKD, as most of these patients have common metabolic comorbidities such as obesity, hypertension, atherosclerotic dyslipidemia, or T2DM. In addition, some drugs that target metabolic risk factors, such as glucagon-like peptide-1 receptor agonists and sodium-glucose cotransporter-2 inhibitors, can be given on a case-by-case basis and may benefit both the liver and kidneys in patients with MAFLD and CKD.

According to the latest definition, metabolic disorders are a key feature of MAFLD. Therefore, when compared with patients with NAFLD, those with MAFLD have a greater likelihood of complications in metabolism-associated diseases, but they are also more prone to liver fibrosis [17,28]. Since fibrosis-associated NAFLD correlates with CKD [29], the relationship between MAFLD and CKD may be closer than that with NAFLD. At present, there are only a few studies on the pathophysiological mechanism of MAFLD and CKD, but the factors that lead to kidney-liver crosstalk can be better explained than those in the pathogenesis that occurs between NAFLD and CKD. These may include genetic, internal environmental risk, and metabolic factors. First, many studies have shown that polymorphisms in genes such as *PNPLA3*, *HSD17B13*, *TM6SF2*, *MBOAT7*, and *GCKR* might play the important roles in the occurrence and development of NAFLD and CKD [17]. Second, a study showed that microbiota from the intestines as well as the intestinal barrier integrity might affect the correlation between NAFLD and CKD through the gut-liver-kidney signaling axis [30]. Studies have shown that the dysregulation of adipokines in patients with MAFLD, including decreased adiponectin and increased fatty acid–binding protein 4 [31], may lead to renal dysfunction [32], glomerular injury [33], tubule-interstitial injury [34], and a decline of eGFR [35]. Finally, the definition of MAFLD includes metabolic risk factors, such as obesity, hypertension, diabetes, and dyslipidemia, that have also been linked to increasing the risk of CKD incidents [36,37]. A study showed that, compared with participants without metabolic syndrome, participants with metabolic syndrome have a 2.5-fold higher risk of developing CKD [38]. Therefore, metabolic dysfunction may play a crucial role in the connection between MAFLD and CKD.

Limitations

The highlight of this study is that it consisted of a large cohort of participants and had a long-term follow-up to observe the incidence of CKD. However, there are some limitations to this study. (1) The participants inhabited urban districts who received health examinations yearly in 1 hospital, and we excluded 1927 participants who had no data regarding MAFLD; therefore, the potential for selection bias was unavoidable. (2) Liver biopsies is the optimum way to diagnose fatty liver disease. However, this is an invasive method and ultrasonography was used instead to diagnose fatty liver disease in our health examinations. A study had shown that the sensitivity and specificity of ultrasonography were high for the diagnosis of fatty liver [39], and it is frequently used in large-scale epidemiological studies [40]. (3) Although 2 incidents of albuminuria were used as the clinical end point in this study, as kidney function was evaluated through an annual health examination, only qualitative data could be obtained. Therefore, at the end of the study, some of the participants may have had acute kidney injury. (4) The health examinations used in this study did not routinely measure either insulin or high-sensitivity C-reactive protein levels. Therefore, it is possible that MAFLD may have been misclassified in some participants. (5) Data relating to the pathological severity of hepatic steatosis were not collected in this study. (6) This study has a long follow-up period, and the metabolic status of participants may have changed over time. Therefore, this study only analyzes the metabolic status at baseline, which may not accurately reflect the true metabolic status of participants. (7) This study was not able to adequately collect participants' smoking, drinking, and eating habits and concomitant therapy, such as the use of angiotensin-converting enzyme inhibitors or sartans, diuretics, statins, and antidiabetic drugs; therefore, the impact of living habits and concomitant therapy on outcomes could not be assessed. In addition, changes in baseline characteristics of the participants at follow-up were not included in the statistical analysis; therefore, the impact of changes in baseline characteristics over time on outcomes could not be assessed either. (8) The results of this study come from a population in China who underwent annual health examinations. It is likely that different ethnic groups are associated with different metabolic risk factors and further studies on a range of ethnic populations are required to verify this study.

Conclusions

This retrospective cohort study found that there was a modest and independent correlation between MAFLD and CKD. It is necessary to explore the relationship between MAFLD and CKD through a large-scale cohort study and basic research and to develop new intervention strategies so that we can prevent the future occurrence of CKD in the population with MAFLD.

Acknowledgments

This work was supported by the Major Project of Science and Technology of Guangxi Zhuang Autonomous Region (grant Guike-AA22096018) and the Specific Research Project of Guangxi for Research Bases and Talents (grant Guike-AD21220042).

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the inclusion of protected health information but can be made available after deidentification upon reasonable request to the corresponding author (JY).

Authors' Contributions

SW and JY contributed to conceptualization. SW, JS, YX, and JH contributed to methodology. SW, JS, and JH contributed to accessing and verifying the underlying data. SW, JS, YX, and JH contributed to investigation. JS, YX, and JH contributed to visualization. SW, JS, and JY contributed to supervision. SW, JS, and JY contributed to writing—original draft. SW, JS, YX, JH, and JY contributed to writing—review and editing.

Conflicts of Interest

None declared.

References

1. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020 Feb 29;395(10225):709-733 [FREE Full text] [doi: 10.1016/S0140-6736(20)30045-3] [Medline: 32061315]
2. Liyanage T, Ninomiya T, Jha V, Neal B, Patrice HM, Okpechi I, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet 2015 May 16;385(9981):1975-1982. [doi: 10.1016/S0140-6736(14)61601-9] [Medline: 25777665]
3. Zhang L, Wang F, Wang L, Wang W, Liu B, Liu J, et al. Prevalence of chronic kidney disease in China: a cross-sectional survey. Lancet 2012 Mar 03;379(9818):815-822. [doi: 10.1016/S0140-6736(12)60033-6] [Medline: 22386035]
4. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, American Gastroenterological Association, American Association for the Study of Liver Diseases, American College of Gastroenterologyh. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. Gastroenterology 2012 Jun;142(7):1592-1609. [doi: 10.1053/j.gastro.2012.04.001] [Medline: 22656328]
5. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease: meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016 Jul;64(1):73-84. [doi: 10.1002/hep.28431] [Medline: 26707365]
6. Sarin SK, Kumar M, Eslam M, George J, Al Mahtab M, Akbar SMF, et al. Lancet Gastroenterol Hepatol 2020 Feb;5(2):167-228 [FREE Full text] [doi: 10.1016/S2468-1253(19)30342-5] [Medline: 31852635]
7. Targher G, Chonchol MB, Byrne CD. CKD and nonalcoholic fatty liver disease. Am J Kidney Dis 2014 Oct;64(4):638-652. [doi: 10.1053/j.ajkd.2014.05.019] [Medline: 25085644]
8. Musso G, Gambino R, Tabibian JH, Ekstedt M, Kechagias S, Hamaguchi M, et al. Association of non-alcoholic fatty liver disease with chronic kidney disease: a systematic review and meta-analysis. PLoS Med 2014 Jul;11(7):e1001680 [FREE Full text] [doi: 10.1371/journal.pmed.1001680] [Medline: 25050550]
9. Mantovani A, Zaza G, Byrne CD, Lonardo A, Zoppini G, Bonora E, et al. Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: a systematic review and meta-analysis. Metabolism 2018 Feb;79:64-76. [doi: 10.1016/j.metabol.2017.11.003] [Medline: 29137912]
10. Wijarnpreecha K, Thongprayoon C, Boonpheng B, Panjawatanan P, Sharma K, Ungprasert P, et al. Nonalcoholic fatty liver disease and albuminuria: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol 2018 Sep;30(9):986-994. [doi: 10.1097/MEG.0000000000001169] [Medline: 29787418]

11. Mantovani A, Petracca G, Beatrice G, Csermely A, Lonardo A, Schattenberg JM, et al. Non-alcoholic fatty liver disease and risk of incident chronic kidney disease: an updated meta-analysis. Gut 2022 Jan;71(1):156-162. [doi: 10.1136/gutjnl-2020-323082] [Medline: 33303564]
12. Sun D, Ye F, Kani HT, Yang J, Zheng KI, Zhang H, et al. Higher liver stiffness scores are associated with early kidney dysfunction in patients with histologically proven non-cirrhotic NAFLD. Diabetes Metab 2020 Sep;46(4):288-295. [doi: 10.1016/j.diabet.2019.11.003] [Medline: 31786360]
13. Wang T, Wang R, Bu Z, Targher G, Byrne CD, Sun D, et al. Association of metabolic dysfunction-associated fatty liver disease with kidney disease. Nat Rev Nephrol 2022 Apr;18(4):259-268. [doi: 10.1038/s41581-021-00519-y] [Medline: 35013596]
14. Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, et al. A new definition for metabolic dysfunction-associated fatty liver disease: an international expert consensus statement. J Hepatol 2020 Jul;73(1):202-209. [doi: 10.1016/j.jhep.2020.03.039] [Medline: 32278004]
15. Brunt EM, Ramrakhiani S, Cordes BG, Neuschwander-Tetri BA, Janney CG, Bacon BR, et al. Concurrence of histologic features of steatohepatitis with other forms of chronic liver disease. Mod Pathol 2003 Jan;16(1):49-56. [doi: 10.1097/01.MP.0000042420.21088.C7] [Medline: 12527713]
16. Eslam M, Sanyal AJ, George J, International Consensus Panel. MAFLD: a consensus-driven proposed nomenclature for metabolic associated fatty liver disease. Gastroenterology 2020 May;158(7):1999-2014.e1. [doi: 10.1053/j.gastro.2019.11.312] [Medline: 32044314]
17. Sun DQ, Jin Y, Wang TY, Zheng KI, Rios RS, Zhang H, et al. MAFLD and risk of CKD. Metabolism 2021 Feb;115:154433. [doi: 10.1016/j.metabol.2020.154433] [Medline: 33212070]
18. Jung CY, Koh HB, Park KH, Joo YS, Kim HW, Ahn SH, et al. Metabolic dysfunction-associated fatty liver disease and risk of incident chronic kidney disease: a nationwide cohort study. Diabetes Metab 2022 Jul;48(4):101344. [doi: 10.1016/j.diabet.2022.101344] [Medline: 35346856]
19. Chinese Clinical Trial Registry. URL: http://www.chictr.org.cn/ [accessed 2023-03-08]
20. Farrell GC, Chitturi S, Lau GKK, Sollano JD, Asia-Pacific Working Party on NAFLD. Guidelines for the assessment and management of non-alcoholic fatty liver disease in the Asia-Pacific region: executive summary. J Gastroenterol Hepatol 2007 Jun;22(6):775-777. [doi: 10.1111/j.1440-1746.2007.05002.x] [Medline: 17565629]
21. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009 May 05;150(9):604-612 [FREE Full text] [doi: 10.7326/0003-4819-150-9-200905050-00006] [Medline: 19414839]
22. Tanaka M, Mori K, Takahashi S, Higashiura Y, Ohnishi H, Hanawa N, et al. Metabolic dysfunction-associated fatty liver disease predicts new onset of chronic kidney disease better than fatty liver or nonalcoholic fatty liver disease. Nephrol Dial Transplant 2023 Feb 28;38(3):700-711. [doi: 10.1093/ndt/gfac188] [Medline: 35612975]
23. Hashimoto Y, Hamaguchi M, Okamura T, Nakanishi N, Obora A, Kojima T, et al. Metabolic associated fatty liver disease is a risk factor for chronic kidney disease. J Diabetes Investig 2022 Feb;13(2):308-316 [FREE Full text] [doi: 10.1111/jdi.13678] [Medline: 34561962]
24. Liang Y, Chen H, Liu Y, Hou X, Wei L, Bao Y, et al. Association of MAFLD with diabetes, chronic kidney disease, and cardiovascular disease: a 4.6-year cohort study in China. J Clin Endocrinol Metab 2022 Jan 01;107(1):88-97 [FREE Full text] [doi: 10.1210/clinem/dgab641] [Medline: 34508601]
25. Carrero JJ, Hecking M, Chesnaye NC, Jager KJ. Sex and gender disparities in the epidemiology and outcomes of chronic kidney disease. Nat Rev Nephrol 2018 Mar 22;14(3):151-164. [doi: 10.1038/nrneph.2017.181] [Medline: 29355169]
26. Glassock R, Delanaye P, El Nahas M. An age-calibrated classification of chronic kidney disease. JAMA 2015 Aug 11;314(6):559-560 [FREE Full text] [doi: 10.1001/jama.2015.6731] [Medline: 26023760]
27. DeFilippis AP, Blaha MJ, Martin SS, Reed RM, Jones SR, Nasir K, et al. Nonalcoholic fatty liver disease and serum lipoproteins: the Multi-Ethnic Study of Atherosclerosis. Atherosclerosis 2013 Apr;227(2):429-436 [FREE Full text] [doi: 10.1016/j.atherosclerosis.2013.01.022] [Medline: 23419204]
28. Yamamura S, Eslam M, Kawaguchi T, Tsutsumi T, Nakano D, Yoshinaga S, et al. MAFLD identifies patients with significant hepatic fibrosis better than NAFLD. Liver Int 2020 Dec;40(12):3018-3030. [doi: 10.1111/liv.14675] [Medline: 32997882]
29. Yeung MW, Wong GL, Choi KC, Luk AO, Kwok R, Shu SS, et al. Advanced liver fibrosis but not steatosis is independently associated with albuminuria in Chinese patients with type 2 diabetes. J Hepatol 2017 Oct 06;68(1):147-156. [doi: 10.1016/j.jhep.2017.09.020] [Medline: 28989092]
30. Raj D, Tomar B, Lahiri A, Mulay SR. The gut-liver-kidney axis: novel regulator of fatty liver associated chronic kidney disease. Pharmacol Res 2020 Feb;152:104617. [doi: 10.1016/j.phrs.2019.104617] [Medline: 31881272]
31. Tanaka M, Takahashi S, Higashiura Y, Sakai A, Koyama M, Saitoh S, et al. Circulating level of fatty acid-binding protein 4 is an independent predictor of metabolic dysfunction-associated fatty liver disease in middle-aged and elderly individuals. J Diabetes Investig 2022 May;13(5):878-888 [FREE Full text] [doi: 10.1111/jdi.13735] [Medline: 34889064]
32. Saito N, Furuhashi M, Koyama M, Higashiura Y, Akasaka H, Tanaka M, et al. Elevated circulating FABP4 concentration predicts cardiovascular death in a general population: a 12-year prospective study. Sci Rep 2021 Feb 17;11(1):4008 [FREE Full text] [doi: 10.1038/s41598-021-83494-5] [Medline: 33597568]

33. Tanaka M, Furuhashi M, Moniwa N, Maeda T, Takizawa H, Matsumoto M, et al. Significance of urinary fatty acid-binding protein 4 level as a possible biomarker for the identification of minimal change disease in patents with nephrotic-range proteinuria. BMC Nephrol 2020 Nov 03;21(1):459 [FREE Full text] [doi: 10.1186/s12882-020-02122-y] [Medline: 33143633]
34. Liu J, Huang R, Li X, Guo F, Li L, Zeng X, et al. Genetic inhibition of FABP4 attenuated endoplasmic reticulum stress and mitochondrial dysfunction in rhabdomyolysis-induced acute kidney injury. Life Sci 2021 Mar 01;268:119023. [doi: 10.1016/j.lfs.2021.119023] [Medline: 33434534]
35. Okazaki Y, Furuhashi M, Tanaka M, Mita T, Fuseya T, Ishimura S, et al. Urinary excretion of fatty acid-binding protein 4 is associated with albuminuria and renal dysfunction. PLoS One 2014;9(12):e115429 [FREE Full text] [doi: 10.1371/journal.pone.0115429] [Medline: 25506691]
36. Kurella M, Lo JC, Chertow GM. Metabolic syndrome and the risk for chronic kidney disease among nondiabetic adults. J Am Soc Nephrol 2005 Jul;16(7):2134-2140. [doi: 10.1681/ASN.2005010106] [Medline: 15901764]
37. Singh AK, Kari JA. Metabolic syndrome and chronic kidney disease. Curr Opin Nephrol Hypertens 2013 Mar;22(2):198-203. [doi: 10.1097/MNH.0b013e32835dda78] [Medline: 23340215]
38. Chen J, Muntner P, Hamm LL, Jones DW, Batuman V, Fonseca V, et al. The metabolic syndrome and chronic kidney disease in U.S. adults. Ann Intern Med 2004 Feb 03;140(3):167-174. [doi: 10.7326/0003-4819-140-3-200402030-00007] [Medline: 14757614]
39. Hernaez R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Guallar E, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. Hepatology 2011 Sep 02;54(3):1082-1090 [FREE Full text] [doi: 10.1002/hep.24452] [Medline: 21618575]
40. Chang Y, Ryu S, Kim Y, Cho YK, Sung E, Kim H, et al. Low levels of alcohol consumption, obesity, and development of fatty liver with and without evidence of advanced fibrosis. Hepatology 2020 Mar;71(3):861-873. [doi: 10.1002/hep.30867] [Medline: 31325180]

Abbreviations

ALT: alanine aminotransferase
AST: aspartate aminotransferase
CKD: chronic kidney disease
eGFR: estimated glomerular filtration rate
HDL-C: high-density lipoprotein cholesterol
HR: hazard ratio
LDL-C: low-density lipoprotein cholesterol
MAFLD: metabolic dysfunction–associated fatty liver disease
NAFLD: nonalcoholic fatty liver disease
TG: triglyceride
T2DM: type 2 diabetes mellitus

Edited by A Mavragani, T Sanchez; submitted 14.12.22; peer-reviewed by CY Jung, R Armstrong Junior, J Wu, L Zhang; comments to author 23.02.23; revised version received 29.03.23; accepted 18.04.23; published 04.05.23.

Please cite as:
Wei S, Song J, Xie Y, Huang J, Yang J
The Role of Metabolic Dysfunction–Associated Fatty Liver Disease in Developing Chronic Kidney Disease: Longitudinal Cohort Study
JMIR Public Health Surveill 2023;9:e45050
URL: https://publichealth.jmir.org/2023/1/e45050
doi:10.2196/45050
PMID:37140958

Original Paper

Prediction of Multimorbidity in Brazil: Latest Fifth of a Century Population Study

Xi-liang Li[1], PhD; Hang Huang[2], MR; Ying Lu[3], PhD; Randall S Stafford[4], MD, PhD; Simone Maria Lima[5], PhD; Caroline Mota[5], PhD; Xin Shi[1,6,7], PhD

[1]School of Mathematics and Information Science, Shandong Technology and Business University, Yantai, China
[2]School of Statistics, Shandong Technology and Business University, Yantai, China
[3]Department of Biomedical Data Science, School of Medicine, Stanford University, Stanford, CA, United States
[4]Department of Medicine, School of Medicine, Stanford University, Stanford, CA, United States
[5]Management Engineering Department, Universidade Federal de Pernambuco, Recife, Brazil
[6]Institute of Health Science, China Medical University, Shenyang, China
[7]Manchester Metropolitan University, Manchester Metropolitan University, United Kingdom

Corresponding Author:
Xin Shi, PhD
School of Mathematics and Information Science
Shandong Technology and Business University
Binhai Road
Yantai, 264005
China
Phone: 86 18059892450
Email: jasonshi510@hotmail.com

Abstract

Background: Multimorbidity is characterized by the co-occurrence of 2 or more chronic diseases and has been a focus of the health care sector and health policy makers due to its severe adverse effects.

Objective: This paper aims to use the latest 2 decades of national health data in Brazil to analyze the effects of demographic factors and predict the impact of various risk factors on multimorbidity.

Methods: Data analysis methods include descriptive analysis, logistic regression, and nomogram prediction. The study makes use of a set of national cross-sectional data with a sample size of 877,032. The study used data from 1998, 2003, and 2008 from the Brazilian National Household Sample Survey, and from 2013 and 2019 from the Brazilian National Health Survey. We developed a logistic regression model to assess the influence of risk factors on multimorbidity and predict the influence of the key risk factors in the future, based on the prevalence of multimorbidity in Brazil.

Results: Overall, females were 1.7 times more likely to experience multimorbidity than males (odds ratio [OR] 1.72, 95% CI 1.69-1.74). The prevalence of multimorbidity among unemployed individuals was 1.5 times that of employed individuals (OR 1.51, 95% CI 1.49-1.53). Multimorbidity prevalence increased significantly with age. People over 60 years of age were about 20 times more likely to have multiple chronic diseases than those between 18 and 29 years of age (OR 19.6, 95% CI 19.15-20.07). The prevalence of multimorbidity in illiterate individuals was 1.2 times that in literate ones (OR 1.26, 95% CI 1.24-1.28). The subjective well-being of seniors without multimorbidity was 15 times that among people with multimorbidity (OR 15.29, 95% CI 14.97-15.63). Adults with multimorbidity were more than 1.5 times more likely to be hospitalized than those without (OR 1.53, 95% CI 1.50-1.56) and 1.9 times more likely need medical care (OR 1.94, 95% CI 1.91-1.97). These patterns were similar in all 5 cohort studies and remained stable for over 21 years. A nomogram model was used to predict multimorbidity prevalence under the influence of various risk factors. The prediction results were consistent with the effects of logistic regression; older age and poorer participant well-being had the strongest correlation with multimorbidity.

Conclusions: Our study shows that multimorbidity prevalence varied little in the past 2 decades but varies widely across social groups. Identifying populations with higher rates of multimorbidity prevalence may improve policy making around multimorbidity prevention and management. The Brazilian government can create public health policies targeting these groups, and provide more medical treatment and health services to support and protect the multimorbidity population.

(JMIR Public Health Surveill 2023;9:e44647) doi:10.2196/44647

KEYWORDS

Brazil; demographic factors; logistic regression analysis; multimorbidity; nomogram prediction; prevalence

Introduction

The co-occurrence of 2 or more chronic diseases in an individual is called multimorbidity [1,2]. Chronic diseases include hypertension, diabetes, cancer, cardiovascular disease, and chronic kidney disease [3]. Chronic diseases pose the greatest threat to human health in the modern world [4,5]. Chronic diseases usually occur together; this multimorbidity will cause greater harm to the body than any chronic disease on its own. The prevalence of multimorbidity has elevated significantly in recent years [6]. The health care needs of patients with multimorbidity differ from those of patients with a single condition and require a complex structured care plan for improved treatment [7]. This will be challenging in contexts where health care resources are scarce, and more medicines will be required to maintain the patient's health. The vulnerability of these patients to safety issues is also elevated [8-10]. Increased prevalence of multimorbidity increases drug overconsumption, the complexity of disease management, the burden on health care services, and the rate of repeat hospitalizations, leading to increased health spending in the country [11,12]. An accurate estimation of multimorbidity prevalence is thus critical to assessing the public health impact of multimorbidity and predicting the medical needs of patients with multimorbidity [13].

The multimorbidity burden can be attributed to various causes, including sociodemographic and behavioral factors; gender is also associated with multimorbidity [14]. Countries differ in the ways they are affected by multimorbidity. Some developing countries do not have access to basic medical care or have sociocultural factors that affect some portions of the population. For example, women are socially, culturally, economically, and educationally disadvantaged in India, making them more vulnerable to chronic diseases [14-16]. In the United States, the prevalence of chronic disease in Black people is significantly higher than that in White people. However, some US studies suggest that gender is not a factor affecting multimorbidity [17,18]. As the population ages, the prevalence of multimorbidity gradually increases, because older adults are more susceptible to multimorbidity [19]. Although multiple studies have examined the relationship between sociodemographic factors and multimorbidity, few have explored the way these relationships change over time or attempted to predict future changes [20-22].

This study aimed to assess the dynamic distribution of 9 chronic diseases in the Brazilian population using cross-sectional survey data obtained from the Brazilian National Household Sample Survey (PNAD) and the Brazilian National Health Survey (PNS) in 1998, 2003, 2008, 2013, and 2019. PNAD and PNS are population-based surveys representing urban and rural Brazilians residing in private households, not in institutions. Data analysis was used to study whether sociodemographic factors, health care needs, and health insurance were associated with the prevalence of multimorbidity. Nomogram plots were used to predict the prevalence of multimorbidity in the future. Compared with traditional prediction methods, the use of nomograms provides excellent graphic visualization. The prediction model was validated in calibration curves by combining different influencing factors of participants. By predicting multimorbidity prevalence within groups, it may be possible to accurately identify the main affected population and ultimately reduce the burden on the medical system.

Methods

Data Availability

The following two data sources were combined in this study: (1) PNAD and (2) PNS data. The PNAD and PNS are complex multistage surveys conducted by the Brazilian Institute of Geography and Statistics to assess the situation of households in Brazil, and are available on the internet [23,24]. The PNAD study is a national survey. A 3-stage self-weighted cluster sampling technique was used. In the first stage, areas with larger populations or cities were selected. Other cities belonging to the same area were then divided into roughly the same level, and the selection was made through a system of size ratios. In the second stage, the contents of the 2010 census were systematically selected. The third stage identified households to survey on an area-specific basis [23]. PNS uses a 3-stage stratified sampling method. In the first stage, participants were selected by simple random sampling. The 2013 PNS study selection was based on households in the second stage, whereas the third stage involved a random selection of residents aged 18 years and above to answer the survey. The 2019 edition of the study involved individuals aged 15 years and above [24].

Ethics Approval

The Brazilian Institute of Geography and Statistics received approval from the National Research Ethics Committee in Brazil and obtained consent from all PNAD and PNS participants. In case of nonagreement, the interviewees could always refuse to participate in the research. No additional data were collected for this study; we used only data presented by the Brazilian Institute of Geography and Statistics.

Study Design

PNAD and PNS data were obtained from 5-year national surveys conducted in 1998, 2003, 2008, 2013, and 2019. Variables with more than 25% missing values were removed from our analysis to ensure the accuracy of the study. We included participants older than 18 years and excluded 6770 participants who did not answer all the questions about chronic diseases. The remaining 877,032 participants constitute a large pool of research participants.

In this study, binary logistic regression was used to compare multimorbidity group differences (gender, race, age, etc) and to calculate odds ratios (OR)—the ratio between the chance that an event occurs in 1 group and the chance that it occurs in another group. Prediction models can be used to aid health care providers' decision-making by estimating the probability that a specific disease or condition is present (diagnostic models) or that a specific event will occur in the future (prognostic models). This model can be used to aid such efforts by predicting the strength of correlation between multimorbidity and other variables, as well as to determine if our logistic regression results are same with nomogram model, further confirming the consistency of all results in this study [25].

Variables

This paper focuses on the prevalence of multimorbidity in Brazil. First, we selected 9 chronic diseases: back disease, rheumatism or arthritis, cancer, diabetes, asthma, hypertension, heart disease, chronic kidney disease, and depression. Chronic disease data are based on self-reported information in response to PNAD and PNS questions, such as "Has a physician or health care professional ever told you that you have diabetes?" or "Has any physician ever given you diabetes diagnosis?" Back diseases included chronic back problems, such as chronic back or neck pain, low back pain, sciatica, and vertebrae or disc problems, but the survey did not delimit the period of the pain occurrence [26]. Among the participants, 165,759 had 2 or more chronic diseases. Second, we selected 11 independent variables to assess the extent to which they were correlated with prevalence of multimorbidity: gender, race, age group, education, employment, region, health insurance, participant well-being, health service accessibility, health service need, and hospitalization. Third, we performed data analysis to estimate the future prevalence of multimorbidity under different influencing factors, such as gender and age. In the survey, subjective well-being was scored on a 5-point scale (1 being the best and 5 being the worst) based on participants' perceptions their own health. The literacy levels of participants were assessed using literacy rate, and employment status was evaluated by asking whether they were employed during the survey period. This study made use of data on whether participants had been hospitalized within the past 12 months or had medical service needs within the past 2 weeks to analyze the effect of multimorbidity on public health, understand the use of medical services, and assist policy makers in formulating public health policies to reduce health spending. We divided the country into 5 regions (South, Southeast, Midwest, Northeast, and North) according to the official Brazilian territorial division to analyze the regional prevalence of frequently occurring diseases. This paper derives some new variables from existing data in databases, such as age group within official regions.

Statistical Analysis

This study separately analyzed 3 aspects of multimorbidity separately: sociodemographic factors, health services, and subjective well-being. Descriptive analysis and data visualization were used to identify trends in changes among participants between years. The association between predictors and prevalence of multimorbidity was analyzed using hypothesis testing. Frequency analysis explored prevalence and the main prevalent population. The effect of variables on multimorbidity was studied using the logistic regression models. OR values were calculated with 95% CI. First, sociodemographic factors, including gender, race, age group, region, literacy, employment status, and health insurance status were considered to compute the probability of multimorbidity (with 1, without 0). Next, logistic regressions were used to investigate the prevalence of multimorbidity in participants with different levels of subjective well-being, health service availability, demand for health services, and hospitalization duration. This allowed us to analyze all the variables in the model independently to determine the OR of the predictor variables. Finally, a nomogram was developed to calculate risk scores using the model; these risk scores were then used to predict prevalence and determine the likelihood that individuals in certain sociodemographic categories will develop multimorbidity.

All statistical analyses were performed using SPSS Statistics, and R (version 4.0.5) was used to visualize the frequency analysis and model prediction.

Results

Descriptive Analysis

A total of 877,032 participants were included in the analysis. Presence of multimorbidity was self-reported by participants aged 18 years or older. Data from 1998, 2003, 2008, 2013, and 2019 were used to analyze the association between 11 risk factors on the prevalence of multimorbidity. Total multimorbidity prevalence was 18.9%. An upward trend in multimorbidity prevalence was present in all years except for the 1998 to 2003 time period. The largest change occurred between 2013 and 2019, when prevalence increased by 5.4%, a significant change (Figure 1). We compared sociodemographic factors, subjective well-being, and health insurance status in different years to the prevalence of multimorbidity (Figure 2). Except for purchasing health insurance and ethnic factors, the change trends of other factors were the same in 5 years. The prevalence of multimorbidity in females was significantly higher than that in males and could reach 11.5% when influence of gender factors on the prevalence of multimorbidity. It was significantly less prevalent in literate than in illiterate participants, and in employed than in unemployed participants. The rates of health care use and hospitalization were significantly higher for participants with multimorbidity, and this difference became larger over time. Regarding self-assessment of health, people who felt their health was "bad" or "very bad" were more likely to have multimorbidity. The prevalence of people who have a better sense of well-being is rising, in contrast to the bad or very bad group. There is little difference in multimorbidity prevalence between people who have health insurance and those who do not, or between different racial or ethnic populations. However, the overall prevalence increased gradually, with the highest rate being seen in the most recent year of data (2019). In the study, the main incidence patterns of 9 diseases in 27 Brazilian federal units were studied. Hypertension and back disease are always the diseases with the highest prevalence in Brazil and are significantly more prevalent in all regions of Brazil.

Figure 1. Prevalence of multimorbidity.

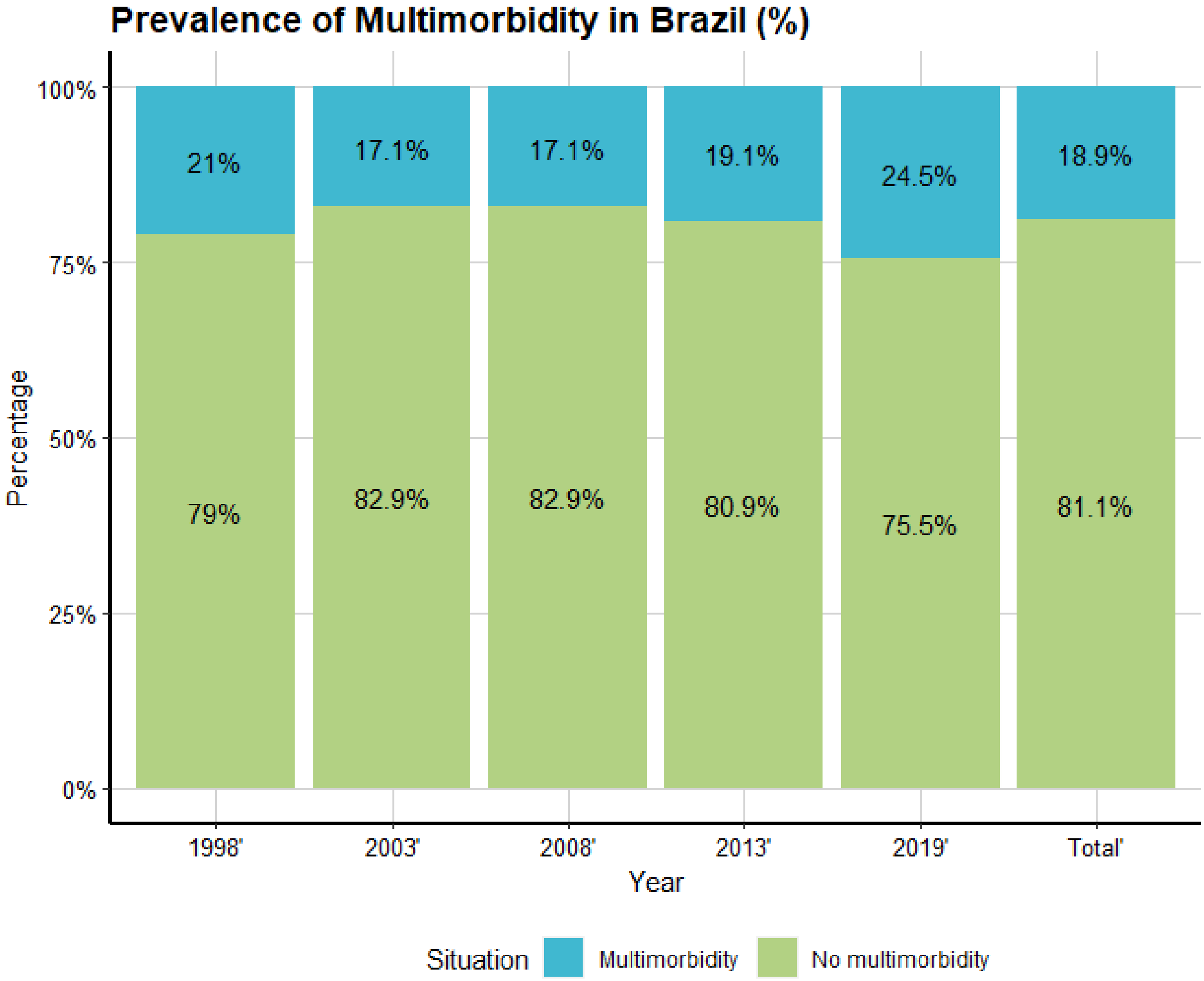
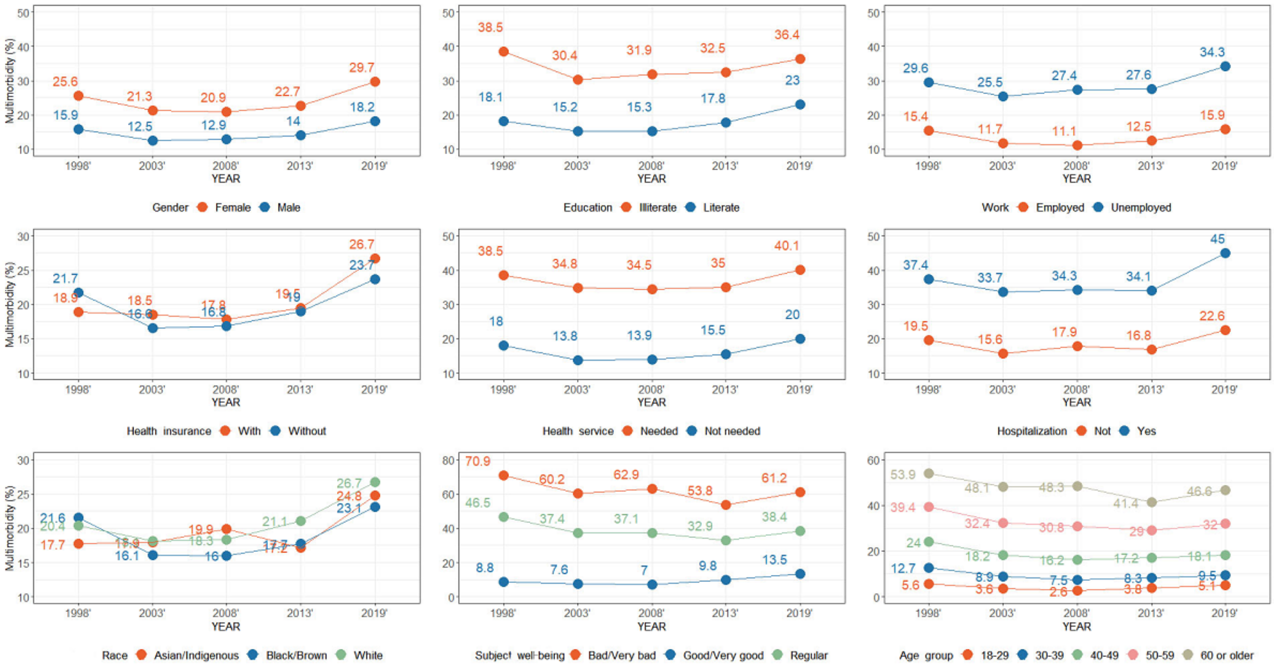


Figure 2. Percentage of multimorbidity by year with sociodemographic, subjective well-being, and health service characteristics.



Logistic Regression and Model Prediction

Tables 1 and 2 indicate the relative degree of multimorbidity risk among participants from 1998 to 2019. Regression analysis revealed that in different cohorts, the changes in the risk of multimorbidity according to demographic variable did not change significantly. However, the magnitude of influence was consistent with the descriptive analysis. Among the differences found, the prevalence of multimorbidity in females was 1.72 times that in males (OR 1.72, 95% CI 1.69-1.74), consistent with the frequency analysis results in the 5-year cohort study. The Asian, Black or Brown, and White groups showed the same degree of prevalence of multimorbidity. The older the age, the higher the risk of multimorbidity. Participants aged 60 years and older were 20 times more likely than those aged 18 to 29 years to have multiple conditions (OR 19.6, 95% CI 19.15-20.07). In addition, being unemployed was positively associated with multimorbidity prevalence, and illiterate participants were 1.3 times more likely than literate participants to have multimorbidity (OR 1.26, 95% CI 1.24-1.28). Multimorbidity presence was higher in the Southern region (OR 1.31, 95% CI 1.28-1.34) than in the Northern region. Individuals having multimorbidity were mostly uninsured (OR 1.05, 95% CI 1.04-1.07).

Table 1. Results of the logistic regressiona examining the association of sociodemographic factors with multimorbidity (odds ratio [OR], 95% CI).

	1998	2003	2008	2013	2019	Overall
Gender						
Male, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Female, OR (95% CI)	1.76 (1.72-1.81)	1.76 (1.72-1.81)	1.60 (1.57-1.64)	1.76 (1.68-1.86)	1.89 (1.82-1.96)	1.72 (1.69-1.74)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Race						
White, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Asian/Indigenous, OR (95% CI)	0.73 (0.63-0.85)	0.89 (0.78-1.02)	1.03 (0.92-1.16)	0.91 (0.75-1.12)	1.06 (0.92-1.23)	0.90 (0.84-0.96)
Brown or Black, OR (95% CI)	1.17 (1.14-1.20)	1.09 (1.06-1.12)	1.09 (1.06-1.11)	1.01 (0.96-1.06)	1.06 (1.02-1.10)	1.08 (1.07-1.09)
P value	<.001	<.001	<.001	.289	.225	<.001
Age group (years)						
18 to 29, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
30 to 39, OR (95% CI)	2.58 (2.48-2.69)	2.77 (2.64-2.90)	3.12 (2.96-3.29)	2.45 (2.17-2.75)	2.11 (1.91-2.31)	2.68 (2.62-2.75)
40 to 49, OR (95% CI)	5.72 (5.49-5.95)	6.25 (5.98-6.53)	7.37 (7.02-7.74)	5.59 (5-6.26)	4.49 (4.10-4.90)	5.93 (5.79-6.07)
50 to 59, OR (95% CI)	11.22 (10.76-11.70)	12.88 (12.32-13.47)	16.06 (15.30-16.86)	10.52 (9.42-11.76)	9.17 (8.41-10)	12 (11.71-12.28)
60 or older, OR (95% CI)	17.91 (17.18-18.67)	21.71 (20.77-22.69)	28.11 (26.78-29.50)	15.45 (13.84-17.24)	14.46 (13.29-15.74)	19.60 (19.15-20.07)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Education						
Literate, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Illiterate, OR (95% CI)	1.37 (1.33-1.41)	1.17 (1.13-1.21)	1.16 (1.13-1.20)	1.09 (1.01-1.18)	1.04 (0.98-1.10)	1.26 (1.24-1.28)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Work						
Employed, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Unemployed, OR (95% CI)	1.38 (1.32-1.39)	1.51 (1.47-1.55)	1.66 (1.62-1.70)	1.51 (1.43-1.60)	1.58 (1.52-1.64)	1.51 (1.49-1.53)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Region						
North, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Northeast, OR (95% CI)	0.71 (0.68-0.75)	0.78 (0.75-0.82)	0.91 (0.88-0.95)	1.04 (0.97-1.21)	1.12 (1.06-1.18)	0.93 (0.91-0.95)
Southeast, OR (95% CI)	0.66 (0.63-0.69)	0.93 (0.89-0.97)	1.13 (1.08-1.18)	1.14 (1.05-1.23)	1.31 (1.24-1.39)	1.05 (1.02-1.07)
South, OR (95% CI)	0.83 (0.79-0.87	1.22 (1.16-1.28)	1.37 (1.31-1.44)	1.63 (1.49-1.79)	1.49 (1.40-1.60)	1.31 (1.28-1.34)
Midwest, OR (95% CI)	0.86 (0.82-0.91)	1.15 (1.09-1.21)	1.21 (1.15-1.27)	1.27 (1.17-1.40)	1.19 (1.11-1.28)	1.19 (1.16-1.22)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Health insurance						
With insurance, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Without insurance, OR (95% CI)	1.25 (1.22-1.29)	1.0 (0.97-1.03)	1.03 (1-1.05)	1 (0.95-1.06)	0.90 (0.86-0.94)	1.05 (1.04-1.07)
P value	<.001	<.001	<.001	<.001	<.001	<.001

[a] Binary logistic regression model by gender, race, age group, literacy, work, region, and health insurance.

Table 2. Results of the logistic regressiona examining the association of subjective well-being and use of health care with multimorbidity (odds ratio [OR], 95% CI).

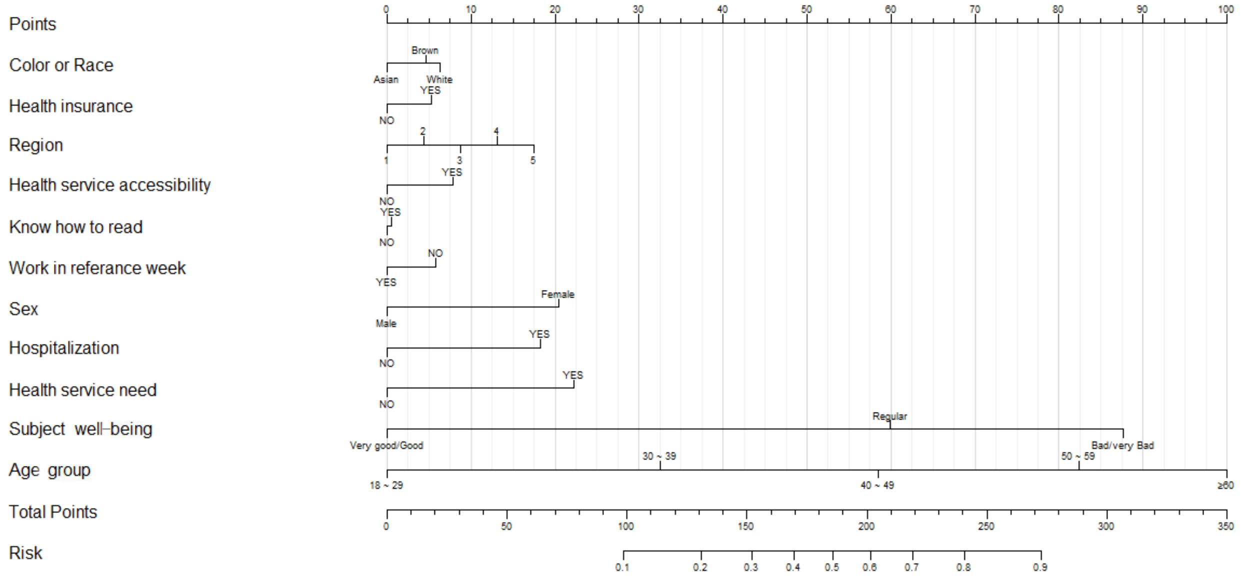
	1998	2003	2008	2013	2019	Overall
Subjective well-being						
Very good or good, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Regular, OR (95% CI)	8.31 (8.10-8.52)	6.46 (6.30-6.62)	7.05 (6.88-7.22)	4.12 (3.92-4.34)	3.66 (3.53-3.80)	6.43 (6.35-6.51)
Bad or very bad, OR (95% CI)	21.204 (20.28-22.17)	14.33 (13.75-14.94)	17.93 (17.23-18.66)	8.92 (8.25-9.66)	8.25 (7.77-8.77)	15.293 (14.97-15.63)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Health service accessibility						
No, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Yes, OR (95% CI)	1.19 (1.16-1.22)	1.39 (1.35-1.44)	1.45 (1.41-1.49)	1.46 (1.38-1.55)	1.33 (1.28-1.39)	1.32 (1.31-1.34)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Health service need						
No, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Yes, OR (95% CI)	1.68 (1.63-1.74)	2.03 (1.98-2.09)	1.99 (1.94-2.05)	2.10 (1.96-2.20)	2.03 (1.95-2.11)	1.94 (1.91-1.97)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Hospitalization						
No, OR (95% CI)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)	1.0 (reference)
Yes, OR (95% CI)	1.35 (1.30-1.40)	1.52 (1.47-1.58)	1.57 (1.51-1.62)	1.54 (1.42-1.66)	1.86 (1.76-1.97)	1.53 (1.50-1.56)
P value	<.001	<.001	<.001	<.001	<.001	<.001

[a] Binary logistic regression model by Subjective well-being, Health service accessibility, Health service need and Hospitalization, Yes: With multimorbidity, No: Without multimorbidity.

Table 2 shows a correlation of subjective well-being and medical assistance needs with multimorbidity. The correlation of "bad" and "very bad" subjective well-being with multimorbidity was 15 times (OR 15.29, 95% CI 14.97-15.63) that of "good" and "very good" in all cohorts of participants. Participants with multimorbidity were 1.3 times more likely to need health care services than those without multimorbidity (OR 1.32, 95% CI 1.31-1.34). In addition, patients with multiple diseases required 2 times more medical services (OR 1.94, 95% CI 1.91-1.97). Moreover, participants hospitalized during the study had a 1.5-fold prevalence of multimorbidity than those not hospitalized (OR 1.53, 95% CI 1.50-1.56).

Nomograms have been developed as user-friendly prediction tools owing to their easy-to-interpret visualization interface. As shown in Figure 3, we constructed nomograms based on each category, with each variable exhibiting its axis corresponding to points representing their significance in the model. The corresponding points were summed up to obtain the total points, and the final probability was obtained by projecting the total points on the risk axis. The plot includes all significant variables based on logistic regression, incorporating sociodemographic factors, subjective well-being, and medical assistance within the nomogram. When calculating the scores for each risk factor, we found that risk scores were significantly higher for individuals aged 60 years or above, those with poorer subjective well-being, those in need of health services and hospitalization, and female and unemployed individuals. We analyzed the multimorbidity prevalence of these risk factors over 21 years to predict the likelihood of multimorbidity among members of these sociodemographic groups in the future. The prediction accuracy of the nomogram can be evaluated by the c-index value: a c-index value of 0.5-0.7 indicates poor prediction, whereas 0.7-0.9 indicates good prediction. The nomogram model developed in this study had a c-index value of 0.854, indicating that the model could correctly predict the results.

Figure 3. Nomograms using sociodemographic, health service, and subject-wellbeing factors selected for predicting multimorbidity risk. Region 1: North; Region 2: Northeast; Region 3: Southeast; Region 4: South; Region 5: Midwest. Given Points are independent points for each factor and Total Points are the sum of all points.



Discussion

Overview

Multimorbidity is a growing global challenge that places a strain on population health and imposes a large economic burden on society. Globally, multimorbidity is estimated to affect 65%-98% of those older than 65 years; the prevalence of multimorbidity is increasing, especially among older people, those with low language proficiency, and women [27,28]. Our findings indicate that the correlation of sociodemographic factors (specifically gender), health insurance status, and subjective well-being with multimorbidity was broadly consistent throughout a 21-year period. These findings are in line with those of other research. With aging, the prevalence of multimorbidity gradually increases, as reported in many studies [13,29,30]. Patients with multimorbidity take more medication, have more complicated disease management, use more health care services, and are more likely to undergo repeat hospitalizations, ultimately increasing overall national health expenditure [15,16].

From 1998 to 2019, the prevalence of multimorbidity first decreased, then increased. The decrease in prevalence could be because government expenditures on health care increased, which improved the health care system and reduced the rate of chronic disease. Particularly noteworthy is the change between 2013 and 2019, during which prevalence increased by 5.4%; in other time periods the increase was more gradual. Declining fertility has led to demographic changes in Brazil since the early 2000s, with an increase in the aging population, causing increased multimorbidity prevalence overall [31].

Studies in the United States show that Black people have a significantly higher rate of multimorbidity compared with White people [18]. Black or Brown populations in Brazil had lower incomes, less access to health insurance, and poorer living conditions [32], reflecting that living environment and income can be closely correlated with multimorbidity. This study examined the effect of educational attainment on multimorbidity, suggesting that illiterate adults had a higher prevalence of multimorbidity. Along with Brazil, this is the case in other countries, including India, South Africa, and Spain [33].

Subjective well-being was significantly related to prevalence of multimorbidity over the 21-year period analyzed in our study. In future research, targeted medical policies could be formulated for individuals with worse subjective well-being to reduce unnecessary medical expenditures [34].

Illness was assessed by whether or not the individual had purchased health insurance, although the time of purchase relative to disease occurrence is unknown. However, having a medical health plan increases access to medical services, reducing physical health risks [30].

In this study, only 19.1% of the participants had health insurance. Although some have chronic illnesses, they may be unaffordable for health insurance. In Brazil, patients have free access through the health care system to some essential medicines, but 1 research found low availability of drugs in all population strata. Drugs not provided by health care system can lead users to abandon prescribed treatments for not being able to buy them in the private sector with their own resources [35,36]. Previous studies have shown a gender difference in patterns of multimorbidity. A systematic review of most previous studies indicated that females had a greater prevalence of multimorbidity than males. This difference might be related to region, social, environmental, or economic factors. As these factors vary globally, their associations with multimorbidity might differ across populations [29,37,38].

Age has the most significant correlation with prevalence of multimorbidity, increasing significantly as the aging population increases. This has implications for countries, such as Brazil with aging populations. Subjective well-being also has a strong correlation with disease status, followed by health insurance status, hospitalization status, and gender. Based on the predicted situation, targeted medical policies should be implemented for individuals in groups with high rates of multimorbidity, reducing medical resource waste, and further improving the efficiency of medical services.

Limitations

Our study analyzed the prevalence of multimorbidity in Brazil over the past 2 decades. Our study used self-reporting to determine whether participants had 1 or more chronic diseases; such self-reporting could be inaccurate if the participants did not remember their diagnoses, or the physician did not accurately diagnose a disease. Self-report has been widely used in public health and is considered a valid approach [31]. Another limitation is that some survey questions have changed over the course of the research. However, the content of the questions has not significantly changed; hence, these changes may not significantly affect the results. In addition, methodological modifications of data collection could have implications for calculating multimorbidity prevalence based on region.

Conclusions

The distribution patterns of multimorbidity indicate that prevalence varies across social groups, but the differences remain largely consistent over the years in Brazil. This study also confirmed that the differences in demographic factors, including gender, education, employment status, and regional factors, significantly impair the public health situation in Brazil. The lack of change in multimorbidity throughout groups over the years offers an opportunity for epidemiologists and the public health sector to develop effective ways to prevent multimorbidity. Our findings may support the development of more effective public resource management. Among possible solutions, public health departments can provide more health services, particularly for groups with high prevalence of multimorbidity, such as women and older people. The government can also anticipate future patterns of multimorbidity and develop policies targeting individuals at high risk for multimorbidity and explore ways to reduce their risks. Additionally, by focusing on individuals in groups with the highest prevalence of multimorbidity, the government can implement public health policies to reduce the harm of multimorbidity while avoiding waste of public resources and assist the public health department in effectively dealing with the problems related to multimorbidity.

Acknowledgments

The authors would like to thank Ms. Emily Bartlett Hines for her editorial expertise. This study was supported by the funders of the National Natural Science Foundation of China (11971273, 12126426) and the Natural Science Foundation of Shandong Province (ZR2018MA004); Immersion Technology and Evaluation Shandong Engineering Research Center (2022); Immersive Smart Devices for Health Care System R&D and Industrial Application Innovation Platform (2022).

Conflicts of Interest

None declared.

References

1. Beard JR, Officer AM, Cassels AK. The world report on ageing and health. Gerontologist 2016;56(suppl 2):S163-S166. [doi: 10.1093/geront/gnw037] [Medline: 26994257]
2. Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, Pons-Vigués M, Foguet-Boreu Q, Muñoz MA, et al. Multimorbidity patterns in the elderly: a prospective cohort study with cluster analysis. BMC Geriatr 2018;18(1):16 [FREE Full text] [doi: 10.1186/s12877-018-0705-7] [Medline: 29338690]
3. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
4. Zeng Q, Yang J, Wang Z, Liu H, Wang J, Yang T, et al. The epidemiological characteristics of noncommunicable diseases and malignant tumors in Guiyang, China: cross-sectional study. JMIR Public Health Surveill 2022;8(10):e36523 [FREE Full text] [doi: 10.2196/36523] [Medline: 36306160]
5. Patel P, Sabin K, Godfrey-Faussett P. Approaches to improve the surveillance, monitoring, and management of noncommunicable diseases in HIV-infected persons: viewpoint. JMIR Public Health Surveill 2018;4(4):e10989 [FREE Full text] [doi: 10.2196/10989] [Medline: 30573446]
6. Kingston A, Robinson L, Booth H, Knapp M, Jagger C, MODEM project. Projections of multi-morbidity in the older population in England to 2035: estimates from the population ageing and care simulation (PACSim) model. Age Ageing 2018;47(3):374-380 [FREE Full text] [doi: 10.1093/ageing/afx201] [Medline: 29370339]
7. Salisbury C, Johnson L, Purdy S, Valderas JM, Montgomery AA. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract 2011;61(582):e12-e21 [FREE Full text] [doi: 10.3399/bjgp11X548929] [Medline: 21401985]

8. Mercer S, Furler J, Moffat K, Fischbacher-Smith D, Sanci L. Multimorbidity: technical series on safer primary care. World Health Organization. 2016. URL: https://apps.who.int/iris/bitstream/handle/10665/252275/9789241511650-eng.pdf [accessed 2024-04-26]
9. Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. BMJ 2015;350:h176 [FREE Full text] [doi: 10.1136/bmj.h176] [Medline: 25646760]
10. Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: an overview of systematic reviews. Ageing Res Rev 2017;37:53-68. [doi: 10.1016/j.arr.2017.05.003] [Medline: 28511964]
11. Bezerra de Souza DL, Oliveras-Fabregas A, Espelt A, Bosque-Prous M, de Camargo Cancela M, Teixidó-Compañó E, et al. Multimorbidity and its associated factors among adults aged 50 and over: a cross-sectional study in 17 European countries. PLoS One 2021;16(2):e0246623 [FREE Full text] [doi: 10.1371/journal.pone.0246623] [Medline: 33571285]
12. Katon WJ, Lin EH, Von Korff M, Ciechanowski P, Ludman EJ, Young B, et al. Collaborative care for patients with depression and chronic illnesses. N Engl J Med 2010;363(27):2611-2620 [FREE Full text] [doi: 10.1056/NEJMoa1003955] [Medline: 21190455]
13. Nguyen H, Manolova G, Daskalopoulou C, Vitoratou S, Prince M, Prina AM. Prevalence of multimorbidity in community settings: a systematic review and meta-analysis of observational studies. J Comorb 2019;9:2235042X19870934 [FREE Full text] [doi: 10.1177/2235042X19870934] [Medline: 31489279]
14. Blümel JE, Carrillo-Larco RM, Vallejo MS, Chedraui P. Multimorbidity in a cohort of middle-aged women: risk factors and disease clustering. Maturitas 2020;137:45-49 [FREE Full text] [doi: 10.1016/j.maturitas.2020.04.016] [Medline: 32498936]
15. Kshatri JS, Palo SK, Bhoi T, Barik SR, Pati S. Prevalence and patterns of multimorbidity among rural elderly: findings of the AHSETS study. Front Public Health 2020;8:582663 [FREE Full text] [doi: 10.3389/fpubh.2020.582663] [Medline: 33251177]
16. Nour NM. An introduction to global women's health. Rev Obstet Gynecol 2008;1(1):33-37. [Medline: 18701928]
17. Johnson-Lawrence V, Zajacova A, Sneed R. Education, race/ethnicity, and multimorbidity among adults aged 30-64 in the national health interview survey. SSM Popul Health 2017;3:366-372 [FREE Full text] [doi: 10.1016/j.ssmph.2017.03.007] [Medline: 29349230]
18. King DE, Xiang J, Pilkerton CS. Multimorbidity trends in United States adults, 1988-2014. J Am Board Fam Med 2018;31(4):503-513 [FREE Full text] [doi: 10.3122/jabfm.2018.04.180008] [Medline: 29986975]
19. Laires PA, Dias S, Gama A, Moniz M, Pedro AR, Soares P, et al. The association between chronic disease and serious COVID-19 outcomes and its influence on risk perception: survey study and database analysis. JMIR Public Health Surveill 2021;7(1):e22794 [FREE Full text] [doi: 10.2196/22794] [Medline: 33433397]
20. Rzewuska M, de Azevedo-Marques JM, Coxon D, Zanetti ML, Zanetti AC, Franco LJ, et al. Epidemiology of multimorbidity within the Brazilian adult general population: evidence from the 2013 national health survey (PNS 2013). PLoS One 2017;12(2):e0171813 [FREE Full text] [doi: 10.1371/journal.pone.0171813] [Medline: 28182778]
21. Carvalho JN, Roncalli AG, Cancela MC, Souza DL. Prevalence of multimorbidity in the Brazilian adult population according to socioeconomic and demographic characteristics. PLoS One 2017;12(4):e0174322 [FREE Full text] [doi: 10.1371/journal.pone.0174322] [Medline: 28384178]
22. Nunes BP, Chiavegatto Filho ADP, Pati S, Cruz Teixeira DS, Flores TR, Camargo-Figuera FA, et al. Contextual and individual inequalities of multimorbidity in Brazilian adults: a cross-sectional national-based study. BMJ Open 2017;7(6):e015885 [FREE Full text] [doi: 10.1136/bmjopen-2017-015885] [Medline: 28601836]
23. Instituto Brasileiro de Geografia e Estatística [Brazilian Institute of Geography and Statistics]. IBGE. 2020. URL: https://www.ibge.gov.br [accessed 2020-05-08]
24. Damacena GN, Szwarcwald CL, Malta DC, Souza-Júnior PRBD, Vieira MLFP, Pereira CA, et al. The development of the national health survey in Brazil, 2013. Epidemiol e Serv saude 2015;24(2):197-206.
25. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015;162(1):55-63 [FREE Full text] [doi: 10.7326/M14-0697] [Medline: 25560714]
26. Malta DC, Oliveira MM, Andrade SSCA, Caiaffa WT, Souza MFM, Bernal RTI. Factors associated with chronic back pain in adults in Brazil. Rev Saude Publica 2017;51(suppl 1):9s [FREE Full text] [doi: 10.1590/S1518-8787.2017051000052] [Medline: 28591350]
27. Hoffman C, Rice D, Sung HY. Persons with chronic conditions. Their prevalence and costs. JAMA 1996;276(18):1473-1479. [Medline: 8903258]
28. Fortin M, Bravo G, Hudon C, Vanasse A, Lapointe L. Prevalence of multimorbidity among adults seen in family practice. Ann Fam Med 2005;3(3):223-228 [FREE Full text] [doi: 10.1370/afm.272] [Medline: 15928225]
29. Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. PLoS One 2014;9(7):e102149 [FREE Full text] [doi: 10.1371/journal.pone.0102149] [Medline: 25048354]

30. Garin N, Koyanagi A, Chatterji S, Tyrovolas S, Olaya B, Leonardi M, et al. Global multimorbidity patterns: a cross-sectional, population-based, multi-country study. J Gerontol A Biol Sci Med Sci 2016;71(2):205-214 [FREE Full text] [doi: 10.1093/gerona/glv128] [Medline: 26419978]
31. Kalache A, Veras RP, Ramos LR. The ageing of the world's population: a new challenge. Revista de saude publica 1987;21:200-210.
32. Kiil A. What characterises the privately insured in universal health care systems? A review of the empirical evidence. Health Policy 2012;106(1):60-75. [doi: 10.1016/j.healthpol.2012.02.019] [Medline: 22459052]
33. Shi X, Lima SMDS, Mota CMM, Lu Y, Stafford RS, Pereira CV. Prevalence of multimorbidity of chronic noncommunicable diseases in Brazil: population-based study. JMIR Public Health Surveill 2021;7(11):e29693 [FREE Full text] [doi: 10.2196/29693] [Medline: 34842558]
34. Wu CS, Lai MS, Gau SS, Wang SC, Tsai HJ. Concordance between patient self-reports and claims data on clinical diagnoses, medication use, and health system utilization in Taiwan. PLoS One 2014;9(12):e112257 [FREE Full text] [doi: 10.1371/journal.pone.0112257] [Medline: 25464005]
35. Mendes LV, Campos MR, Chaves GC, Silva RM, Freitas PS, Costa KS, et al. Disponibilidade de medicamentos nas unidades básicas de saúde e fatores relacionados: uma abordagem transversal. Saude Debate 2014;38(spe):109-123. [doi: 10.5935/0103-1104.2014s009]
36. Tavares NU, Bertoldi AD, Thumé E, Facchini LA, França GV, Mengue SS. Fatores associados à baixa adesão ao tratamento medicamentoso em idosos [Factors associated with low adherence to medication in older adults]. Rev Saude Publica 2013;47(6):1092-1101 [FREE Full text] [doi: 10.1590/s0034-8910.2013047004834] [Medline: 24626547]
37. Health topics: women's health. World Health Organization. 2015. URL: https://www.who.int/health-topics/women-s-health [accessed 2023-04-26]
38. Verdonk P, Klinge I. Mainstreaming sex and gender analysis in public health genomics. Gend Med 2012;9(6):402-410. [doi: 10.1016/j.genm.2012.10.006] [Medline: 23164527]

Abbreviations

OR: odds ratio
PNAD: Brazilian National Household Sample Survey
PNS: Brazilian National Health Survey

Edited by A Mavragani, T Sanchez; submitted 28.11.22; peer-reviewed by J Hadlock, Q Zeng; comments to author 22.12.22; revised version received 11.01.23; accepted 06.04.23; published 30.05.23.

Please cite as:
Li XL, Huang H, Lu Y, Stafford RS, Lima SM, Mota C, Shi X
Prediction of Multimorbidity in Brazil: Latest Fifth of a Century Population Study
JMIR Public Health Surveill 2023;9:e44647
URL: https://publichealth.jmir.org/2023/1/e44647
doi:10.2196/44647
PMID:37252771

Original Paper

Epidemiology and a Predictive Model of Prognosis Index Based on Machine Learning in Primary Breast Lymphoma: Population-Based Study

Yushuai Yu[1*], MM; Zelin Xu[1*], MM; Tinglei Shao[1*], BMed; Kaiyan Huang[1], MM; Ruiliang Chen[1], BMed; Xiaoqin Yu[1], BMed; Jie Zhang[1], PhD; Hui Han[1], MD; Chuangui Song[1], PhD

Department of Breast Surgery, Fujian Medical University Union Hospital, Fuzhou, China

[*]these authors contributed equally

Corresponding Author:
Chuangui Song, PhD
Department of Breast Surgery
Fujian Medical University Union Hospital
29 Xin Quan Road, Gulou District
Fuzhou, 350001
China
Phone: 86 13960709993
Fax: 86 59183357896 801
Email: Songcg1971@hotmail.com

Abstract

Background: Primary breast lymphoma (PBL) is a rare disease whose epidemiological features, treatment principles, and factors used for the patients' prognosis remain controversial.

Objective: The aim of this study was to explore the epidemiology of PBL and to develop a better model based on machine learning to predict the prognosis for patients with primary breast lymphoma.

Methods: The annual incidence of PBL was extracted from the surveillance, epidemiology, and end results database between 1975 and 2019 to examine disease occurrence trends using Joinpoint software (version 4.9; National Cancer Institute). We enrolled data from 1251 female patients with primary breast lymphoma from the surveillance, epidemiology, and end results database for survival analysis. Univariable and multivariable analyses were performed to explore independent prognostic factors for overall survival and disease-specific survival of patients with primary breast lymphoma. Eight machine learning algorithms were developed to predict the 5-year survival of patients with primary breast lymphoma.

Results: The overall incidence of PBL increased drastically between 1975 and 2004, followed by a significant downward trend in incidence around 2004, with an average annual percent change (AAPC) of −0.8 (95% CI −1.1 to −0.6). Disparities in trends of PBL exist by age and race. The AAPC of the 65 years or older cohort was about 1.2 higher than that for the younger than 65 years cohort. The AAPC of White patients is 0.9 (95% CI 0.0-1.8), while that of Black patients was significantly higher at 2.1 (95% CI −2.5 to 6.9). We also identified that the risk of death from PBL is multifactorial and includes patient factors and treatment factors. Survival analysis revealed that the patients diagnosed between 2007 and 2015 had a significant risk reduction of mortality compared to those diagnosed between 1983 and 1990. The gradient booster model outperforms other models, with 0.752 for sensitivity and 0.817 for area under the curve. The important features established with the gradient booster model were the year of diagnosis, age, histologic type, and primary site, which were the 4 most relevant variables to explain 5-year survival status.

Conclusions: The incidence of PBL started demonstrating a tendency to decrease after 2004, which varied by age and race. In recent years, the prognosis of patients with primary breast lymphoma has been remarkably improved. The gradient booster model had a promising performance. This model can help clinicians identify the early prognosis of patients with primary breast lymphoma and therefore improve the clinical outcome by changing management strategies and patient health care.

(JMIR Public Health Surveill 2023;9:e45455) doi:10.2196/45455

KEYWORDS

primary breast lymphoma; epidemiology; prognosis; machine learning; disparities

Introduction

Primary breast lymphoma (PBL) is defined as a kind of lymphoma that is only located in the breast, as well as ipsilateral lymph nodes at the time of initial diagnosis [1,2]. It represents no more than 1% of breast malignancies and less than 3% of extranodal lymphomas [3,4]. Nevertheless, the incidence rate of PBL is rising in recent years and deserves attention [5].

Since PBL is a rare disease, its epidemiological characteristics, treatment, and prognosis remain controversial. In particular, the wide variations of its prognoses, which have been reported in different studies, challenged us to evaluate the prognosis of PBL. The 5-year survival rates ranged from 50% to 90% in the previous reports [6-11]. Certainly, the diversity may be due to different case series, a small sample available, different subtypes, clinical stages, treatment methods, and so on. Research is inconclusive about what will affect the outcome of PBL and how much of an effect change can bring. The 5-year survival rate varies according to different clinical stages: 89% for stage I and 50% for stage II [12]. A huge difference exists in the long-term prognosis in recent years, showing improvement with the development of modern therapy [13]. Age has also been reported as an independent prognostic factor, and cases of elderly patients were complicated by more comorbidities that caused a poor prognosis [14,15]. In some studies, chemotherapy and radiotherapy were associated with longer survival, and there was no benefit from mastectomy [2,16]. However, compared with the above factors, histological subtypes play a more important role. Diffuse large B-cell lymphoma (DLBCL) is the most common subtype, which is more aggressive, followed by follicular and mucosa-associated lymphoid tissue (MALT) lymphoma with indolent behavior [13,17]. Picasso et al [18] found that tumors in 50% of patients with primary breast lymphoma were located in the upper outer quadrant and 25% were in the upper inner quadrant; however, there are no studies that try to analyze prognosis between different primary sites. Consequently, the factor of the primary site was also incorporated into our study. There are many other factors that may contribute to the outcome of PBL, but it is not clear about the role of each one.

In order to build up a reliable way to predict the prognosis of patients with primary breast lymphoma, we need to combine all potential prognostic factors with different weight ratios in 1 model. Since it is difficult to set up an effective model in the traditional way under complex interference factors, for example, Nomogram, we use machine learning and the Surveillance, Epidemiology, and End Results (SEER) database to conduct our study. First, we investigated the epidemiology, clinicopathologic features, treatment modalities, and outcomes of PBL. Second, we tried to establish a predictive model with the assistance of machine learning including 11 prognostic factors (age, race, year of diagnosis, marital status, laterality, primary tumor site, histology, Ann Arbor stage, surgery status, radiation status, and chemotherapy status). We believe our work may help with the evaluation of patients with primary breast lymphoma in the future.

Methods

Data Source and Study Population

The annual incidence of PBL was extracted from the SEER database between 1975 and 2019 to examine national trends, and all incidence rates were age adjusted. Since the Ann Arbor staging was not available until 1983, patients diagnosed between 1975 and 1982 were not included in the survival analysis and the establishment of the machine learning model. Finally, we enrolled 1251 patients using SEER*Stat (version 8.3.9; National Cancer Institute) for survival analysis, according to the following inclusion criteria: female, year of diagnosis from 1983 to 2015, the age of diagnosis more than 15 years, breast lymphoma as the only primary malignant cancer diagnosis, and Ann Arbor stage I-II. The exclusion criteria were as follows: Ann Arbor stage III-IV (because these were considered unlikely to be in accordance with extranodal disease) or unknown information, younger than 15 years old, multiple tumors, male cases, and patients who died within 30 days. This study tracked the duration of follow-up starting from the day of diagnosis to December 31, 2019, or the date of death, which can provide follow-up data for more than 5 years. Patient characteristics and treatment courses in our study were identified. The data related to age, race, year of diagnosis, marital status, laterality, primary tumor site, histology, Ann Arbor stage, surgery status, radiation status, and chemotherapy status. Surgery was divided into mastectomy and breast-conserving surgery. We cannot further classify chemotherapy and radiotherapy as the SEER database does not provide detailed chemotherapy and radiotherapy data, such as the regimen, dose, and duration. However, anthracycline-based chemotherapy regimens and radiotherapy of extranodal lesions are the primary treatment options for patients with primary breast lymphoma [2,19,20], so these limitations did not influence our results much.

Ethics Approval

Our primary data were extracted from the SEER database, which is publicly available. We got permission for data extraction and usage after signing a data-use agreement for the SEER 1975-2019 research data file. Consequently, human subject research ethics review and informed consent were exempted from this study. We confirm that the information of enrolled patients was anonymous or deidentified. In addition, all statistical analyses were conducted in accordance with the regulations of the SEER Program.

Outcome Measurement

The primary outcome of the study is overall survival (OS). It was defined as from the date of initial diagnosis to the date of death by any cause including PBL. Patients who were alive on the date of the last follow-up were censored. Disease-specific survival (DSS), which served as a secondary study outcome in our study, was defined as from the date of diagnosis to the date of death due to PBL.

Statistical Analysis

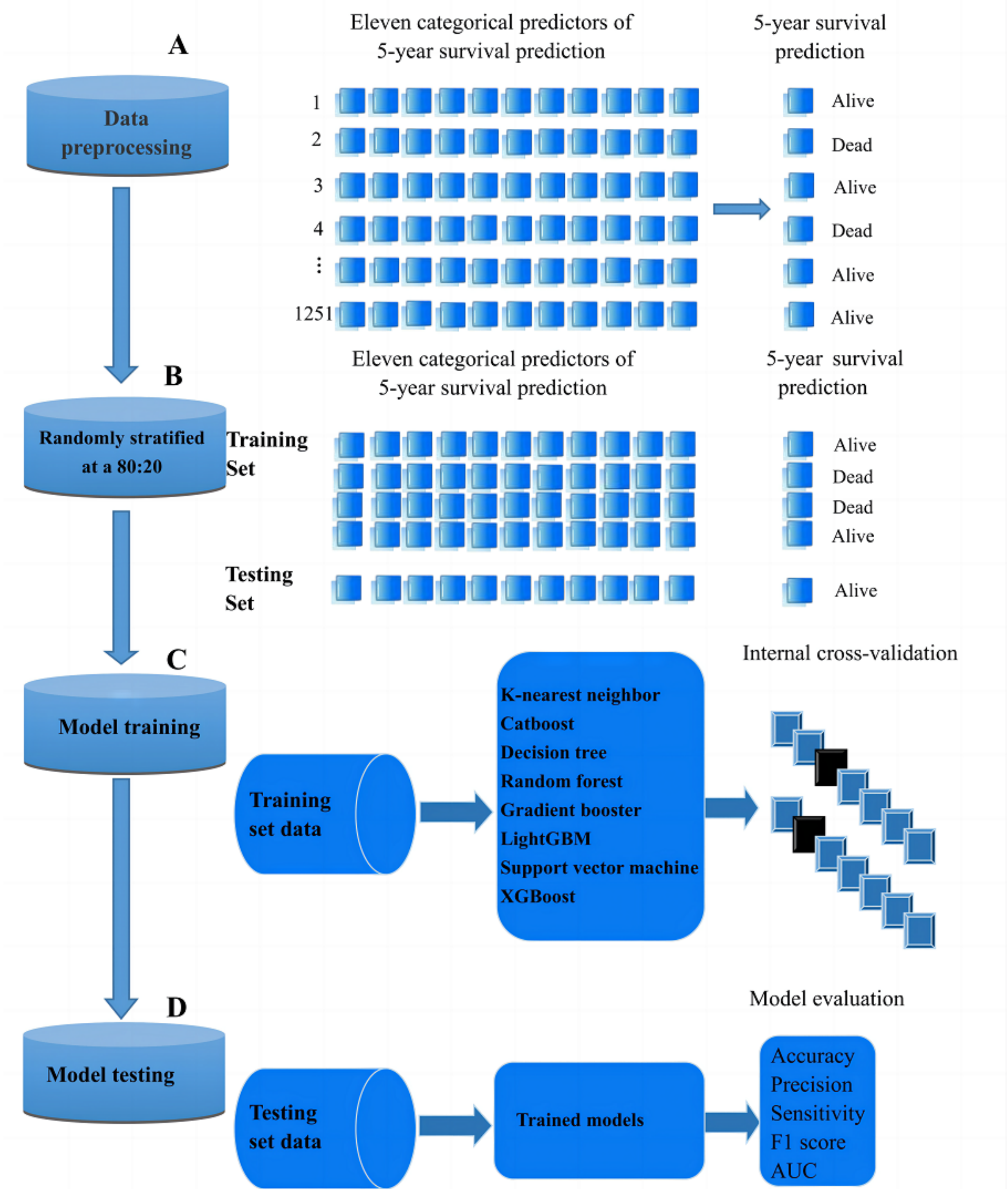
The incidence of PBL for trend analysis was retrieved from the SEER database. The time trends in incidence for PBL were assessed and fit using Joinpoint software (version 4.9; National

Cancer Institute) based on log-linear models. Annual percentage change and average annual percent change (AAPC) were calculated to indicate the direction and magnitude of the trends. In order to explore the influences of demographic differences for PBL incidence, the overall population was stratified into different groups, including age and race.

The Kaplan-Meier method was used to generate survival curves. The log-rank test was performed to determine the differences between different demographic and clinical characteristics of PBL patients. Hazard ratio (HR) with 95% CI was identified by using a Cox proportional hazard regression model to determine the factors associated with the outcomes. These statistical analyses were conducted by using SPSS (version 26.0; IBM Corp), and a *P* value of less than .05 was considered as a statistical difference.

Eleven categorical predictors including age, race, year of diagnosis, marital status, laterality, primary tumor site, histology, Ann Arbor stage, surgery status, radiation status, and chemotherapy status were collected to build a machine learning model for 5-year survival prediction (Figure 1). The package of “MissForest” was used to impute missing values in the data set. Of all enrolled patients, 35.1% (n=439) of patients did not have information on the primary site, 11.2% (n=140) of patients were missing information on the histologic type, and 18.1% (n=226) of patients had no treatment information regarding surgery. The portions of missing values were far less than the cutoff of heavy missingness (75%), which promises good performance of the “MissForest” algorithm [21]. Before building machine learning models, all patients with primary breast lymphoma were randomly divided into a training set and a testing set, at an 80:20 ratio. In our study, 8 machine learning algorithms were used, including K-nearest neighbor, CatBoost, decision tree, random forest method, Gradient Boost, LightGBM, support vector machine, and XGBoost models. A 10-fold internal cross-validation was used to identify the optimal parameters, which provided the highest degree of accuracy in each model. Subsequently, the performance of all machine learning algorithms was evaluated in a testing set, and evaluation measures consisted of accuracy, precision, sensitivity, F1 score, and area under the receiver operating characteristic curve (AUC). The contribution of each element to the machine learning model was evaluated by using feature importance based on the package of “partial_dependence.” Python (version 3.8; Python Software Foundation) was used in these processes.

Figure 1. The flowchart of machine learning development process. AUC: area under the curve.



Results

Baseline Characteristics of Patients

The baseline clinical characteristics of the patients are shown in Table 1. Overall, a total of 1251 eligible patients were enrolled in our study. Among the patients, 540 (43.2%) were <65 years and 711 (56.8%) were ≥65 years. A total of 31 (2.5%) patients had tumors in the bilateral, and 1220 (97.5%) patients had tumors in the unilateral. The Ann Arbor stages were distributed as follows: 976 (78%) cases were stage I and 275 (22%) cases were stage II. DLBCL was the most common histologic type of PBL followed by MALT and follicular lymphoma (FL), accounting for 43.4%. In addition, the enrolled patients were not inclined to accept local therapy, including surgery (no surgery vs breast-conserving and mastectomy: n=656, 52.4% vs n=369, 29.5%) and radiotherapy (no radiation vs radiation: n=1031, 82.4% vs n=220, 17.6%), while the percentage of patients in no chemotherapy and chemotherapy was about the same (no chemotherapy vs chemotherapy: n=656, 52.4% vs n=595, 47.6%).

Table 1. Baseline characteristics of primary breast lymphoma.

Characteristics	Patients (N=1251), n (%)
Age (years)	
<65	540 (43.2)
≥65	711 (56.8)
Race	
White	1016 (81.2)
Black	104 (8.3)
Other[a]	131 (10.5)
Marital status	
Married	633 (50.6)
Not married[b]	618 (49.4)
Laterality	
Unilateral	1220 (97.5)
Bilateral	31 (2.5)
Year of diagnosis	
1983-1990	71 (6)
1991-1998	166 (13.3)
1999-2006	401 (32.1)
2007-2015	613 (49)
Primary site	
Axillary tail	19 (1.5)
Central portion	74 (6)
Inner quadrant	77 (6.2)
Lower-inner quadrant	49 (3.9)
Lower-outer quadrant	52 (4.2)
Nipple	9 (1)
Overlapping lesion	223 (17.8)
Upper-outer quadrant	309 (24.7)
NA[c]	439 (35.1)
Histologic type	
DLBCL[d]	543 (43.4)
MALT[e]	241 (19.3)
CLL/SLL[f]	58 (4)
FL[g]	192 (15.3)
ALCL[h]	27 (2)
Other[i]	50 (4)
NA	140 (11.2)
Ann arbor stage	
I	976 (78)
II	275 (22)
Surgery approach	

Characteristics	Patients (N=1251), n (%)
No surgery	656 (52.4)
BCS[j]	333 (26.6)
Mastectomy	36 (3)
NA	226 (18.1)
Radiation status	
No	1031 (82.4)
Yes	220 (17.6)
Chemotherapy status	
No	656 (52.4)
Yes	595 (47.6)

[a] Other includes American Indian/Alaskan Native, Asian/Pacific Islander, and unknown.

[b] Not married includes divorced, separated, single (never married), unmarried or domestic partner, and widowed.

[c] NA: not available.

[d] DLBCL: diffuse large B-cell lymphoma.

[e] MALT: mucosa-associated lymphoid tissue.

[f] CLL/SLL: chronic lymphocytic leukemia/small lymphocytic lymphoma.

[g] FL: follicular lymphoma.

[h] ALCL: anaplastic large cell lymphoma.

[i] Other includes anaplastic large cell lymphoma, angioimmunoblastic T-cell lymphoma, Burkitt lymphoma, extranodal NK-/T-cell lymphoma, nasal type, Mantle cell lymphoma, peripheral T-cell lymphoma, precursor B-lymphoblastic lymphoma, subcutaneous panniculitis-like T-cell lymphoma, T lymphoblastic leukemia/lymphoma.

[j] BCS: breast-conserving surgery.

Incidence of Breast Lymphoma

The annual percentage change and AAPC for patients with primary breast lymphoma by age and race from 1975 to 2019 are reported in Table 2 and Figure 2. The results demonstrate a remarkable AAPC growth trend of 0.8 (95% CI 0.1-1.5) of patients with primary breast lymphoma during the period 1975 to 2019. The incidence of PBL appears to have a turning point around 2004. From 1975 to 2004, an upward trend was observed, followed by a decline from 2004 to 2019 (AAPC=−0.8; 95% CI −1.1 to −0.6). The AAPC of the ≥65 years cohort was about 1.2 higher than that for the <65 years cohort, which revealed that the incidence of PBL increased slowly with increasing age. The AAPC of White patients is 0.9 (95% CI 0.0-1.8), while that of Black patients was significantly higher at 2.1 (95% CI −2.5 to 6.9). Generally, the PBL incidence substantially increased for the White population between 1975 and 2004 followed by a downward trend for the period between 2004 and 2019. The Black patient cohort also has a similar trend; however, the peak was in 2002 and the incidence has declined noticeably slower than for White patients.

Table 2. Trends in age-standardized incidence rates of primary breast lymphoma in 1975-2019.

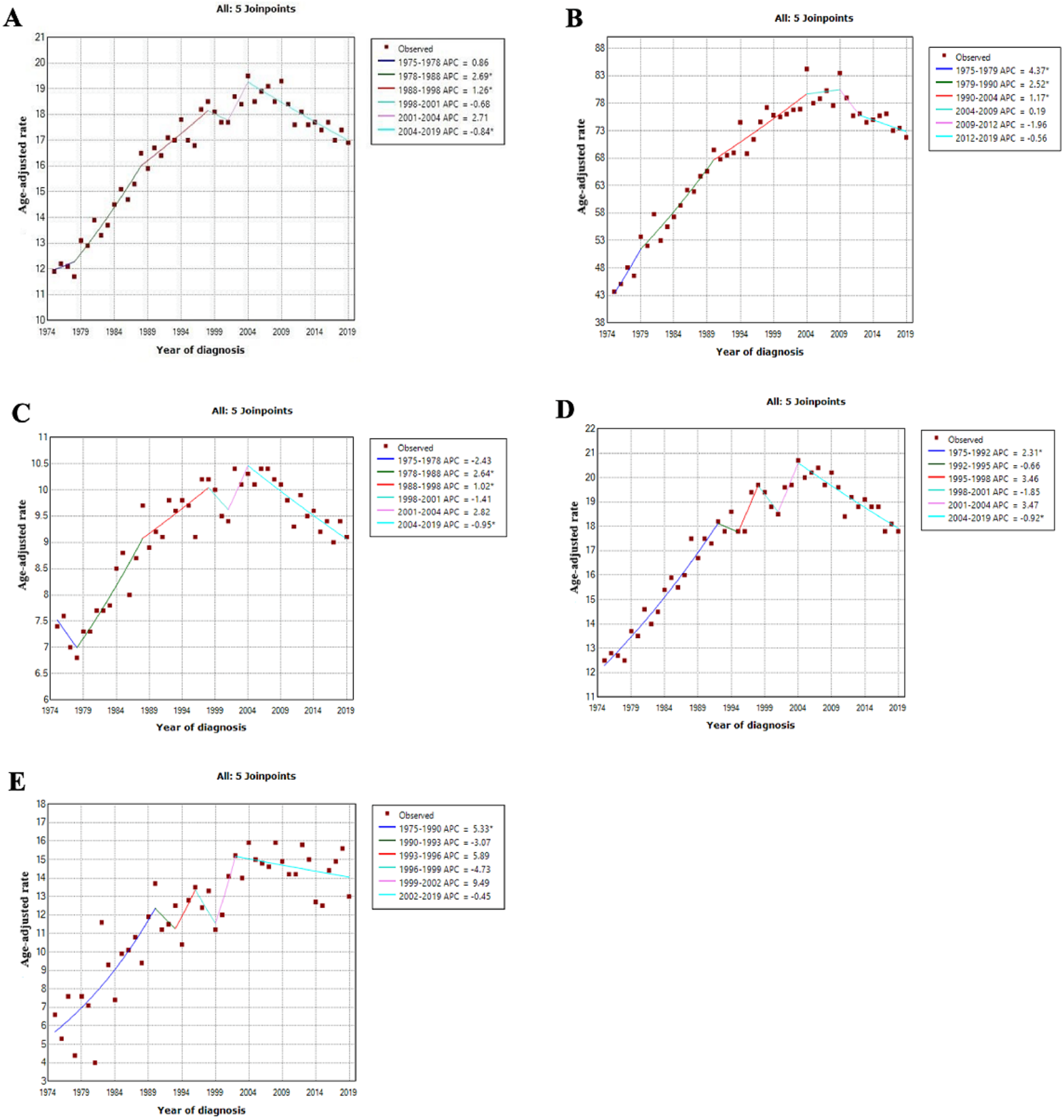
Model	Age ≥65 (years)		Age <65 (years)		White		Black		Overall	
	Year	APC[a] (95% CI)	Year	APC (95% CI)	Year	APC (95% CI)	Year	APC (95% CI)	Year	APC (95% CI)
Trend 1	1975-1979	4.4[b] (0.8 to 8.0)	1975-1978	−2.4 (−7.9 to 3.3)	1975-1992	2.3[b] (2.0 to 2.6)	1975-1990	5.3[b] (2.8 to 7.9)	1975-1978	0.9 (−4.0 to 5.9)
Trend 2	1979-1990	2.5[b] (1.9 to 3.2)	1978-1988	2.6[b] (1.7 to 3.6)	1992-1995	−0.7 (−7.5 to 6.7)	1990-1993	−3.1 (−34.3 to 43.1)	1978-1988	2.7[b] (1.9 to 3.5)
Trend 3	1990-2004	1.2[b] (0.8 to 1.5)	1988-1998	1.0[b] (0.2 to 1.8)	1995-1998	3.5 (−3.5 to 10.9)	1993-1996	5.9 (−25.8 to 51.1)	1988-1998	1.3[b] (0.6 to 1.9)
Trend 4	2004-2009	0.2 (−1.8 to 2.2)	1998-2001	−1.4 (−9.4 to 7.3)	1998-2001	−1.9 (−8.3 to 5.0)	1996-1999	−4.7 (−31.6 to 32.7)	1998-2001	−0.7 (−7.0 to 6.1)
Trend 5	2009-2012	−2.0 (−7.9 to 4.3)	2001-2004	2.8 (−5.1 to 11.4)	2001-2004	3.5 (−2.4 to 9.7)	1999-2002	9.5 (−20.0 to 49.9)	2001-2004	2.7 (−3.6 to 9.4)
Trend 6	2012-2019	−0.6 (−1.3 to 0.2)	2004-2019	−1.0[b] (−1.3 to 0.6)	2004-2019	−0.9[b] (−1.2 to 0.7)	2002-2019	−0.5 (−1.3 to 0.4)	2004-2019	−0.8[b] (−1.1 to 0.6)
AAPC[c]	1975-2019	1.2[b] (0.6 to 1.8)	1975-2019	0.4 (−0.5 to 1.3)	1975-2019	0.9 (0.0 to 1.8)	1975-2019	2.1 (−2.5 to 6.9)	1975-2019	0.8[b] (0.1 to 1.5)

[a] APC: annual percentage change.

[b] Significantly different from 0 at α=.05 (*P*<.05). There are 5 joinpoints for each model.

[c] AAPC: average annual percent change.

Figure 2. Trend in the incidence of primary breast lymphoma from 1975 to 2019: (A) Overall; (B) Age ≥65 (years); (C) Age <65 (years); (D) White; and (E) Black. *Indicates that the APC is significantly different from 0 at α=.05 level. Final selected model: 2 joinpoints. APC: annual percent change.



Survival Analysis

The median follow-up time for enrolled patients was 106 months. The Kaplan-Meier curves of OS and DSS based on different baselines in demographic and clinical characteristics are shown in Figures 3 and 4. In our analysis, patients older than 65 years had a noticeably poorer prognosis than younger, which indicated age was an important prognosis factor. The Kaplan-Meier curves for the time period of disease diagnosis are shown in Figure 3E and Figure 4E, the period between 2007 and 2015 was far superior to other time periods. Patients with diffuse large B-cell lymphoma have shorter survival periods compared with other histologic types. Lower Ann Arbor stage (stage I) at diagnosis patients demonstrated a distinctive survival benefit over those with higher Ann Arbor stage (stage II). The primary site in the breast could also influence the prognosis of PBL, as patients with a central portion or nipple neoplasm had a poorer prognosis than patients whose primary tumor sites were in other breast quadrants. The actual laterality of the primary site (bilateral or unilateral) and race did not appear to be related to the prognosis. In terms of therapeutic approaches, breast-conserving surgery and radiotherapy had better OS and DSS.

The univariate Cox regression analysis for each variable is shown in Multimedia Appendix 1. The result of multivariate analysis shown in Table 3 revealed that age, marital status, year of diagnosis, histologic type, Ann Arbor stage, and radiation status were independent prognosis factors. Patients who are older at the time of diagnosis have a higher risk of death than those who are younger (OS: HR 3.458, 95% CI 2.766-4.323, *P*<.001; DSS: HR 1.997, 95% CI 1.511-2.639, *P*<.001, respectively). In terms of marital status, married women had a significant survival advantage (OS: HR 1.549, 95% CI 1.294-1.854, *P*<.001; DSS: HR 1.462, 95% CI 1.140-1.874, *P*=.003, respectively). The patients diagnosed between 2007 and 2015 had a significant risk reduction in mortality than those who were diagnosed between 1983 and 1990 (OS: HR 0.536, 95% CI 0.312-0.919, *P*=.02; DSS: HR 0.411, 95% CI 0.199-0.849, *P*=.02, respectively). Histological type is one of the fundamental features to describe PBL. Patients with mucosa-associated lymphoid tissue, chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL), and FL had a significant OS and DSS advantage than DLBCL (MALT vs DLBCL: OS, HR 0.396, 95% CI 0.287-0.546, *P*<.001; DSS: HR 0.197, 95% CI 0.111-0.348, *P*<.001; CLL/SLL vs DLBCL: OS, HR 0.448, 95% CI 0.290-0.694, *P*<.001; DSS, HR 0.125, 95% CI 0.045-0.345, *P*<.001; FL vs DLBCL: OS, HR 0.519, 95% CI 0.392-0.686, *P*<.001; DSS, HR 0.396, 95% CI 0.265-0.593, *P*<.001). Patients with a higher Ann Arbor stage (stage II) at diagnosis had a higher hazard of death than those with a lower stage (stage I) at diagnosis, which yielded an HR of 1.414 (95% CI 1.146-1.744) in OS analysis. Radiation lowered the risk of disease-specific mortality and all-cause mortality (OS: HR 0.709, 95% CI 0.551-0.913, *P*=.008; DSS: HR 0.620, 95% CI 0.430-0.893, *P*=.01, respectively). However, no significant difference both in OS and DSS level was detected in patients who received surgery and chemotherapy compared with those who did not (breast-conserving surgery vs no surgery: OS, HR 1.050, 95% CI 0.808-1.364, *P*=.72; DSS, HR 0.948, 95% CI 0.640-1.404, *P*=.79; mastectomy vs no surgery: OS, HR 1.036, 95% CI 0.647-1.658, *P*=.88; DSS, HR 1.196, 95% CI 0.661-2.163, *P*=.55; chemotherapy vs no chemotherapy: OS, HR 0.820, 95% CI 0.661-1.018, *P*=.07; DSS, HR 0.869, 95% CI 0.652-1.159, *P*=.34).

Figure 3. Kaplan-Meier estimate of overall survival by subgroup analysis: (A) age, (B) race, (C) marital status, (D) laterality, (E) year of diagnosis; (F) primary tumor site, (G) histology; (H) Ann Arbor stage, (I) surgery status, (J) radiation status, and (K) chemotherapy status. OS: overall survival.

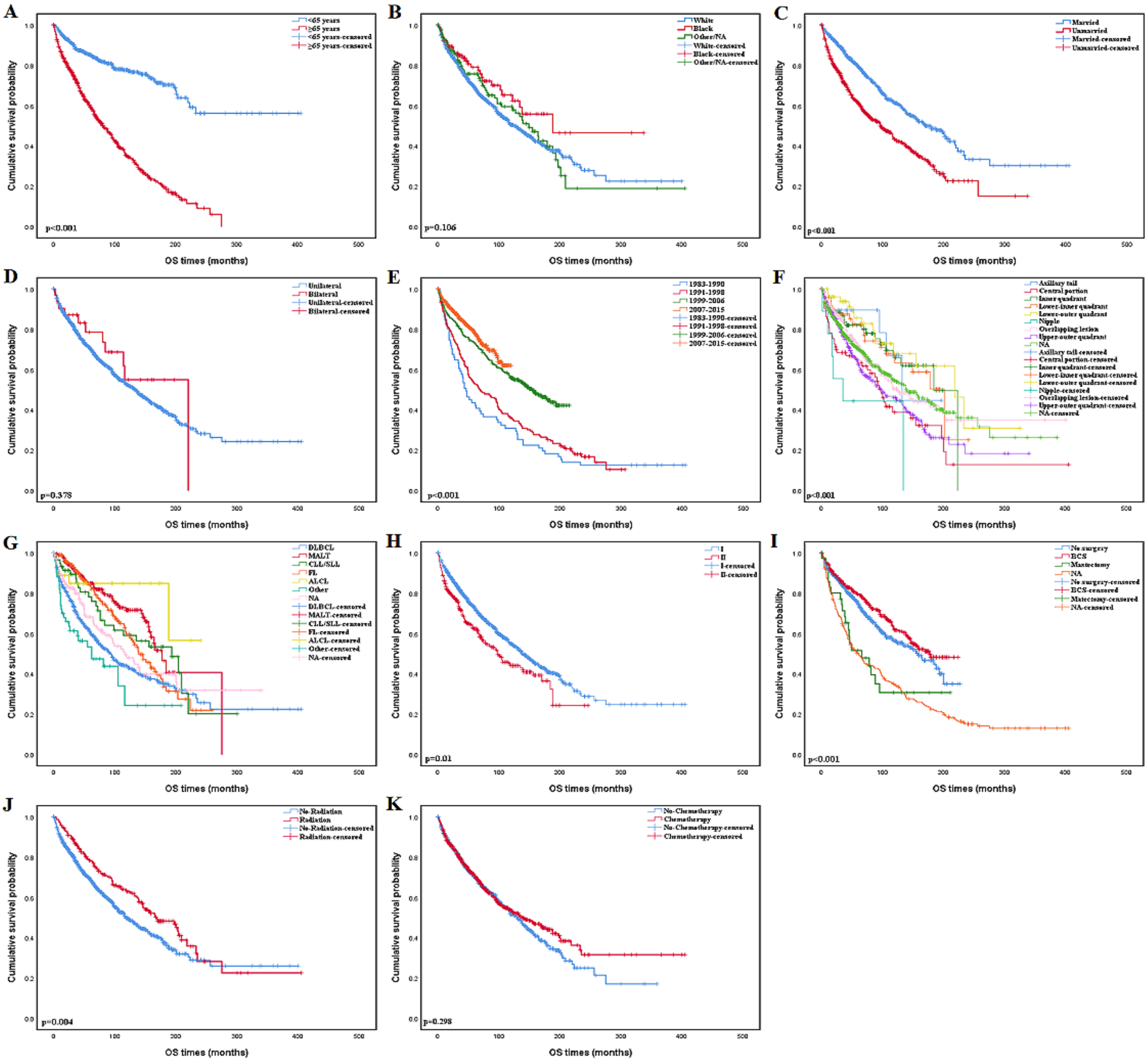


Figure 4. Kaplan-Meier estimate of disease-specific survival by subgroup analysis: (A) age, (B) race, (C) marital status; (D) laterality; (E) year of diagnosis, (F) primary tumor site; (G) histology, (H) Ann Arbor stage, (I) surgery status, (J) radiation status; and (K) chemotherapy status. DSS: disease-specific survival.

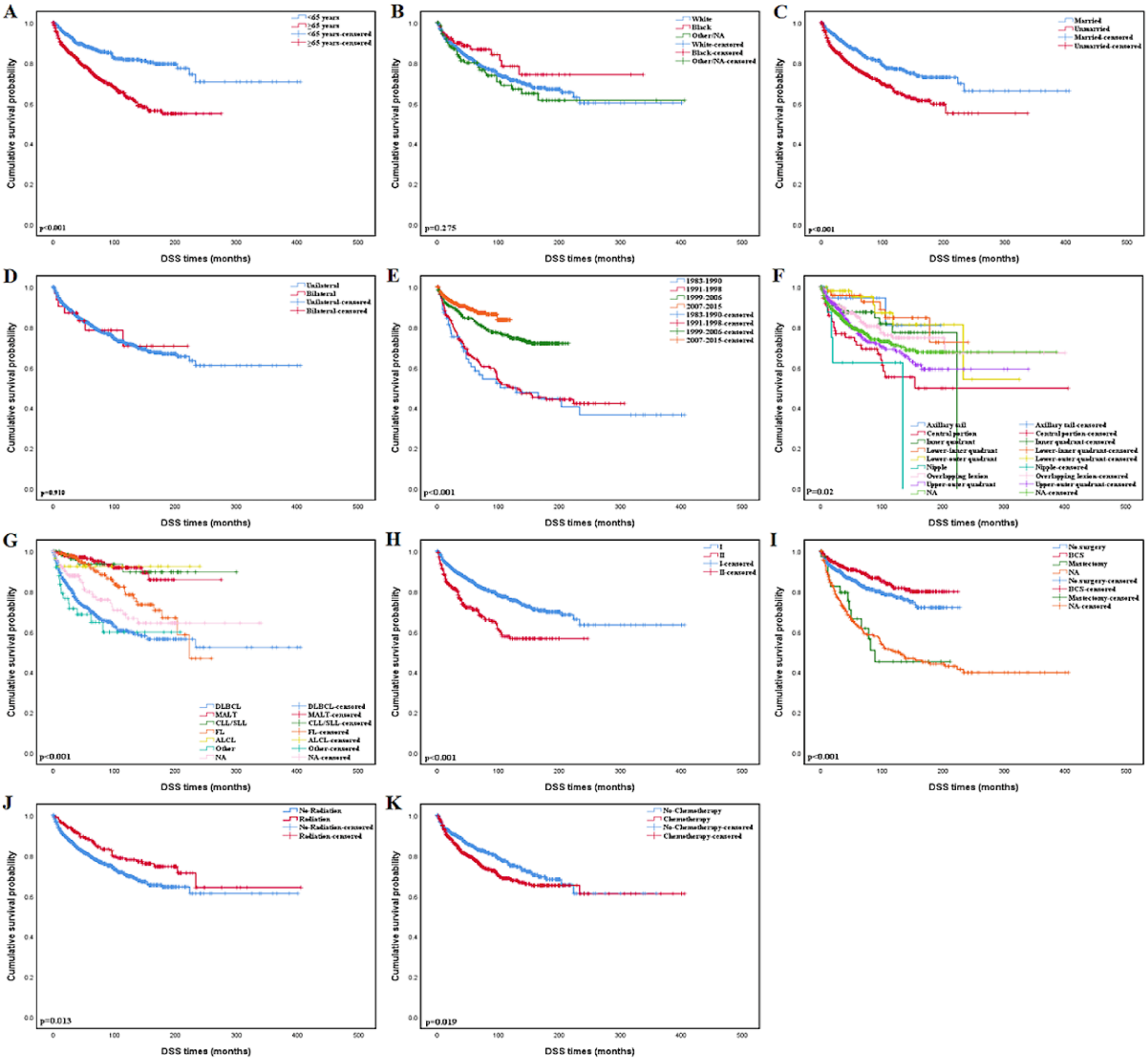


Table 3. Multivariate Cox proportional hazard model of disease-specific survival and overall survival in all patients.

Variables	DSS[a]		OS[b]	
	HR[c] (95% CI)	*P* value	HR (95% CI)	*P* value
Age (years)				
<65	Reference	—[d]	Reference	—
≥65	1.997 (1.511-2.639)	*<.001* [e]	3.458 (2.766-4.323)	*<.001*
Race				
White	Reference	—	Reference	—
Black	0.978 (0.575-1.662)	.93	0.932 (0.632-1.375)	.72
Other[f]	1.044 (0.720-1.515)	.82	0.999 (0.751-1.329)	>.99
Marital status				
Married	Reference	—	Reference	—
Not married[g]	1.462 (1.140-1.874)	*.003*	1.549 (1.294-1.854)	*<.001*
Laterality				
Unilateral	Reference	—	Reference	—
Bilateral	1.179 (0.537-2.592)	.68	0.863 (0.464-1.603)	.64
Year of diagnosis				
1983-1990	Reference	—	Reference	—
1991-1998	0.975 (0.645-1.474)	.90	0.810 (0.592-1.107)	.19
1999-2006	0.604 (0.299-1.216)	.16	0.610 (0.361-1.032)	.07
2007-2015	0.411 (0.199-0.849)	*.02*	0.536 (0.312-0.919)	*.02*
Primary site				
Axillary tail	Reference	—	Reference	—
Central portion	1.812 (0.425-7.730)	.42	1.371 (0.538-3.496)	.51
Inner quadrant	1.370 (0.305-6.145)	.68	0.866 (0.325-2.306)	.77
Lower-inner quadrant	0.927 (0.185-4.642)	.93	1.027 (0.376-2.807)	.96
Lower-outer quadrant	0.718 (0.143-3.596)	.69	0.770 (0.275-2.157)	.62
Nipple	2.123 (0.375-12.010)	.40	1.848 (0.547-6.249)	.32
Overlapping lesion	1.022 (0.244-4.272)	.98	1.018 (0.411-2.525)	.97
Upper-outer quadrant	1.398 (0.339-5.764)	.64	1.301 (0.529-3.199)	.57
NA[h]	1.222 (0.298-5.012)	.78	0.995 (0.405-2.441)	.99
Histologic type				
DLBCL[i]	Reference	—	Reference	—
MALT[j]	0.197 (0.111-0.348)	*<.001*	0.396 (0.287-0.546)	*<.001*
CLL/SLL[k]	0.125 (0.045-0.345)	*<.001*	0.448 (0.290-0.694)	*<.001*
FL[l]	0.396 (0.265-0.593)	*<.001*	0.519 (0.392-0.686)	*<.001*
ALCL[m]	0.234 (0.057-0.967)	*.045*	0.409 (0.165-1.013)	.05
Other[n]	1.133 (0.673-1.908)	.64	1.357 (0.919-2.005)	.13
NA	0.556 (0.373-0.830)	*.004*	0.659 (0.492-0.882)	*.005*
Ann Arbor stage				
I	Reference	—	Reference	—
II	1.847 (1.412-2.416)	*<.001*	1.414 (1.146-1.744)	*.001*

Variables	DSS[a]		OS[b]	
	HR[c] (95% CI)	*P* value	HR (95% CI)	*P* value
Surgery approach				
No surgery	Reference	—	Reference	—
BCS[o]	0.948 (0.640-1.404)	.79	1.050 (0.808-1.364)	.72
Mastectomy	1.196 (0.661-2.163)	.55	1.036 (0.647-1.658)	.88
NA	1.386 (0.760-2.528)	.29	1.319 (0.839-2.075)	.23
Radiation status				
No	Reference	—	Reference	—
Yes	0.620 (0.430-0.893)	*.01*	0.709 (0.551-0.913)	*.008*
Chemotherapy status				
No	Reference	—	Reference	—
Yes	0.869 (0.652-1.159)	.34	0.820 (0.661-1.018)	.07

[a] DSS: disease-specific survival.

[b] OS: overall survival.

[c] HR: hazard ratio.

[d] Not applicable.

[e] Italics indicate statistical significance.

[f] Other includes American Indian/Alaskan native, Asian/Pacific Islander, and unknown.

[g] Not married includes divorced, separated, single (never married), unmarried or domestic partner, and widowed.

[h] NA: not available.

[i] DLBCL: diffuse large B-cell lymphoma.

[j] MALT: mucosa-associated lymphoid tissue.

[k] CLL/SLL: chronic lymphocytic leukemia/small lymphocytic lymphoma.

[l] FL: follicular lymphoma.

[m] ALCL: anaplastic large cell lymphoma.

[n] Other includes anaplastic large cell lymphoma; angioimmunoblastic T-cell lymphoma; Burkitt lymphoma; extranodal NK-/T-cell lymphoma, nasal type; Mantle cell lymphoma; Peripheral T-cell lymphoma; Precursor B-lymphoblastic lymphoma; Subcutaneous panniculitis-like T-cell lymphoma; T lymphoblastic leukemia/lymphoma.

[o] BCS: breast-conserving surgery.

Machine Learning–Based 5-Year Survival Prediction in Patients With Primary Breast Lymphoma

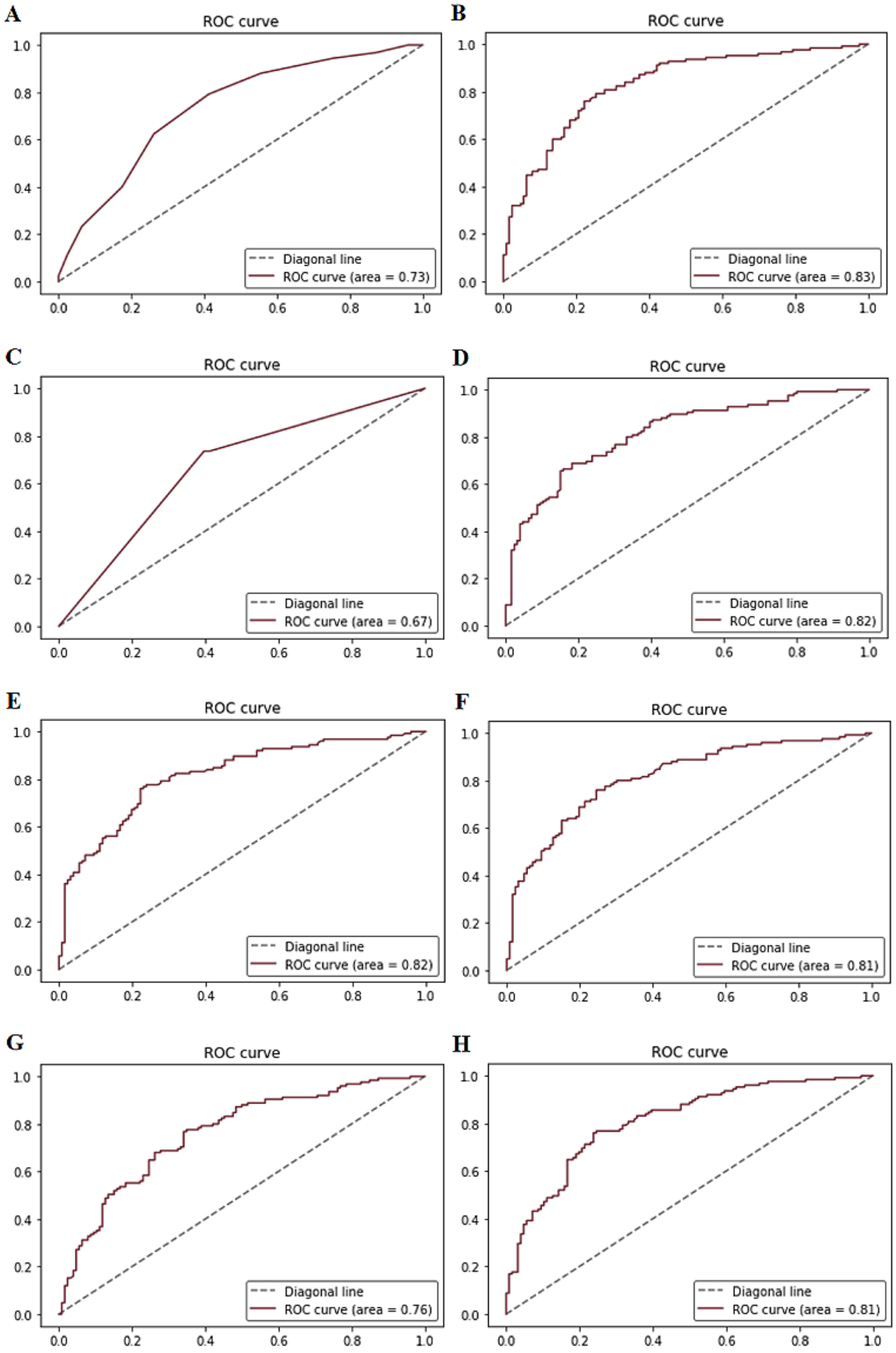
We used a 1251-patient data set for training 8 machine learning models to predict the 5-year survival after PBL diagnosis. The performance of these 8 algorithms is presented in detail in Table 4. The resulting confusion matrix is shown in Multimedia Appendix 2. For the test data set, the sensitivities were K-nearest neighbor model (0.624), Catboost model (0.736), decision tree model (0.736), random forest model (0.720), gradient booster model (0.752), LightGBM model (0.712), support vector machine model (0.696), and XGBoost model (0.728). The AUCs were K-nearest neighbor model (0.735), Catboost model (0.829), decision tree model (0.667), random forest model (0.817), gradient booster model (0.817), LightGBM model (0.814), support vector machine model (0.761), and XGBoost model (0.811). The receiver operating characteristic curves of the 8 models are shown in Figure 5. Due to the design of our study, we focused primarily on testing the sensitivity of patients at high risk of experiencing death in the fifth year. The gradient booster model demonstrated the highest accuracy, precision, sensitivity, and F1 score of all these 8 models; the model also shows a high AUC. Accordingly, the gradient booster algorithm proved to be the most appropriate model for this study. Multimedia Appendix 3 indicates the importance scores for each variable used in the gradient booster, which suggested that year of diagnosis, age, histologic type, and primary site were the 4 most relevant variables to explain 5-year survival status.

Table 4. Model performance for the 5-year survival.

Algorithms	Accuracy	Precision	Sensitivity	F1 score	AUC[a]
K-nearest neighbor	0.721	0.703	0.624	0.661	0.735
Catboost	0.757	0.767	0.736	0.751	0.829
Decision tree	0.669	0.648	0.736	0.689	0.667
Random forest	0.721	0.720	0.720	0.720	0.817
Gradient booster	0.765	0.770	0.752	0.761	0.817
LightGBM	0.745	0.761	0.712	0.736	0.814
Support vector machine	0.685	0.680	0.696	0.688	0.761
XGBoost	0.745	0.752	0.728	0.740	0.811

[a] AUC: area under the curve.

Figure 5. Receiver operating characteristic curves for all models. ROCs for all models: (A) K-nearest neighbor; (B) Catboost; (C) decision tree; (D) random Forest; (E) gradient booster; (F) LightGBM; (G) support vector machine; (H) XGBoost. ROC: receiver operating characteristic curve.



Discussion

Principal Findings

To understand the changes in the incidence trend and survival of patients with primary breast lymphoma over the last 40 years, we conducted a large population-based epidemiological study with data from the SEER database. The results of our study showed the overall incidence of PBL appeared to start a downward trend around 2004. Our study is the first one to report this encouraging phenomenon. We also developed and validated novel models based on machine learning algorithms for predicting 5-year survival. In particular, the gradient booster model achieved the most promising results in terms of AUC, accuracy, precision, sensitivity, and F1 score.

PBL is a comparatively rare form of extranodal lymphoma, and global reports about the incidence of PBL remain limited.

According to the literature from 1975 to 2000, the overall incidence of PBL had increased dramatically but has lately stabilized [13]. Our results indicate a similar trend. By extracting data from 1975 to 2019 from the SEER database, we identified an overall increase in the incidence of PBL during 1975-2004, with a peak incidence rate in 2004. In comparison, a previous article that estimated the non-Hodgkin lymphoma (NHL) incidence for 185 countries in 2018 using the IARC's GLOBOCAN database also showed an increasing trend in the incidence of PBL until the 1990s and the stabilization or decrease thereafter [22]. Genetic determinants, lifestyle, and environmental factors have been considered as causes for NHL [23]. Infections with hepatitis C virus, Epstein-Barr virus, *Helicobacter pylori*, and HIV increase the risk of NHL, and the reduced incidence of lymphoma can be partly explained by the decrease in the incidence of virus infections and advances in antiretroviral therapy [24-26]. In our study, we focused on the possible impact of age and race on the incidence of disease. Our data show that the AAPC in an older population was 3 times higher than in a younger population. Consistent with our results, a report from Cancer Research UK indicated that a high proportion of NHL diagnoses occurred in older people, with the highest incidence in people aged 80 to 84 years [27]. Our data revealed that the incidence of PBL increased slowly with increasing age. In addition, race was also significantly associated with the incidence rate, as the AAPC of Black people was much higher than that of White people. The apparent differences in the incidence of PBL by race may be related to the levels of access to health care, availability of diagnostic services, endemic infections, and environmental factors [28]. In summary, this preliminary finding suggests that future biological and epidemiological analyses on PBL should be stratified by age and ethnic background.

Building up a reliable way to predict the prognosis of patients with primary breast lymphoma plays a crucial role in the early determination of the treatment of patients with primary breast lymphoma. Currently, we determine the prognosis for patients with primary breast lymphoma mainly based on clinicopathological characteristics; however, the limited data cannot provide adequate information for clinicians to deal with this extremely complex disease. Although a study developed a nomogram to predict the survival of PBL, Ann Arbor stage III and IV patients were all involved in their study, which may affect the effectiveness of the model [1,2,15]. Additionally, treatment approaches that have a critical impact on the prognosis of the disease including surgery, chemotherapy, and radiotherapy were not incorporated into the model due to the lack of statistical significance [15]. These omissions make any conclusions highly controversial. Machine learning is being widely used in the medical field for disease diagnosis, prognosis, therapeutic modality selection, and so on [29-31]. A machine learning model can automatically adjust the weight of the factors to make the best use of the data. Our study used the 5-year survival of patients with primary breast lymphoma as the predictive end point, an important point for early determination of prognosis. The results showed that the performance of the gradient booster model was superior to that of all the other models and is regarded as a promising model. Machine learning techniques have also been used to predict the 5- and 10-year recurrence of invasive breast cancer. Massafra et al [32] enrolled 529 patients with breast cancer from Italy, reaching good AUC values of 0.771 and 0.763 for the recurrence prediction at 5 and 10 years. There are 28 features associated with primary breast cancer clinicopathological characteristics and treatment programs that were used to train models, which are more detailed than our cohort. Twenty-eight features were used because breast cancer can provide more clinicopathological characteristics and have more treatment methods than PBL. However, our established gradient booster model shows a higher AUC value, which reflects that it is still possible to train promising prediction models, even with limited predictors.

The important features established with the gradient booster model were the year of diagnosis, age, histologic type, and primary site as the 4 most relevant variables to explain the 5-year survival status. The year of diagnosis was considered as the most meaningful predictive prognostic factor, which suggested that current treatments, probably in combination with newer systemic treatments (likely rituximab), have improved the control of this disease [13]. Rituximab was approved by the US Food and Drug Administration for marketing in the United States on November 26, 1997, and was a milestone for the treatment of PBL. This may be an important factor in why our results demonstrate a significant improvement in the prognosis of patients with primary breast lymphoma after 1999. Age and histologic type were confirmed as important prognosis factors in the machine learning model, like other investigations [3,15,33,34]. Undoubtedly, elderly PBL patients become complicated by more comorbidities and poor drug tolerance or physical condition, which may have a direct negative impact on survival time [35,36]. In accordance with the past research, different histologic types demonstrated dramatic survival differences, which can be partly due to some cell phenotypes with a high proliferation feature and association with a poor treatment response [33,34,37]. Interestingly, the primary site of PBL is thought to be an important predictive prognostic feature in the gradient booster model, which was never reported or analyzed by other series. There are significant differences in the primary site of PBL, a cohort found PBL predisposes to locate in the upper outer quadrant, which is also confirmed in our study [18]. Our Kaplan-Meier survival curves across different primary sites suggested that the central portion and nipple site have worse outcomes. This may be explained by the anatomy of breast lymphatic drainage. The subareolar plexus collects lymphatics originating from breast parenchyma, and then they drain to the lymph nodes of the axilla. Additionally, the deep lymphatic channels connect to the superficial cutaneous lymphatic plexus, especially in the subareolar plexus around the nipple [38,39]. The primary site of central portion or nipple may be prone to be present with lymphatic vessel invasion, causing a poor prognosis; however, this needs further study in future trials or experimental research.

Our study provides information on the incidence and prognostic factors over the last 40 years involving a sufficient sample size. In addition, the first prognostic model for patients with primary breast lymphoma based on a machine learning algorithm was performed for clinical use. We are confident that we have built a predictive model with a good performance, and it can provide physicians with an easy-to-access predictive tool and facilitate a more personalized follow-up strategy, management strategies, and patient care for patients with primary breast lymphoma. The model may help to identify patients who are at a higher risk of a poor outcome and will require more aggressive treatment. However, there are some limitations to our study. Due to limited information from the SEER database, we did not include variables such as biomarkers, chemotherapy regimen, radiotherapy dosing, and targeted drugs in our model. Therefore, when interpreting the results, caution should be used. Further, the database may not capture all relevant patient outcomes including the frequency of central nervous system recurrence, which might potentially affect the interpretation of our results. In addition, the models generated in this study have not been verified in an external validation cohort. In order to achieve this objective, we are collecting related case information to establish a database.

Conclusions

The incidence of PBL started demonstrating a tendency to decrease after 2004, which varied by age and race. In recent years, the prognosis of PBL has been remarkably improved. The gradient booster model had a promising performance. This model can help clinicians identify the prognosis of patients with primary breast lymphoma early and therefore improve clinical outcomes by changing management strategies and patient health care.

Acknowledgments

This study was sponsored by Fujian provincial health technology project (2021GG01010074).

Data Availability

The data sets generated and analyzed during this study are available in the Surveillance, Epidemiology, and End Results (SEER) database [40].

Authors' Contributions

CS and YY conceived the study design. YY, ZX, and TS participated in performing data cleaning and statistical analysis. YY, KH, and RC wrote the original draft. JZ, HH, and CS were responsible for resources and supervision. XY and CS edited the manuscript. JZ, HH, and CS are the corresponding authors for this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Univariate Cox proportional hazard model of disease-specific survival and overall survival in all patients.
[DOCX File , 23 KB - publichealth_v9i1e45455_app1.docx]

Multimedia Appendix 2
Confusion matrix of 8 algorithms for the 5-year survival status.
[DOCX File , 17 KB - publichealth_v9i1e45455_app2.docx]

Multimedia Appendix 3
Feature importance of the gradient booster model.
[PNG File , 15 KB - publichealth_v9i1e45455_app3.png]

References

1. Wiseman C, Liao KT. Primary lymphoma of the breast. Cancer 1972;29(6):1705-1712 [FREE Full text] [doi: 10.1002/1097-0142(197206)29:6<1705::aid-cncr2820290640>3.0.co;2-i] [Medline: 4555557]
2. Ryan G, Martinelli G, Kuper-Hommel M, Tsang R, Pruneri G, Yuen K, International Extranodal Lymphoma Study Group. Primary diffuse large B-cell lymphoma of the breast: prognostic factors and outcomes of a study by the International Extranodal Lymphoma Study Group. Ann Oncol 2008;19(2):233-241 [FREE Full text] [doi: 10.1093/annonc/mdm471] [Medline: 17932394]
3. Aviv A, Tadmor T, Polliack A. Primary diffuse large B-cell lymphoma of the breast: looking at pathogenesis, clinical issues and therapeutic options. Ann Oncol 2013;24(9):2236-2244 [FREE Full text] [doi: 10.1093/annonc/mdt192] [Medline: 23712546]
4. Cheah CY, Campbell BA, Seymour JF. Primary breast lymphoma. Cancer Treat Rev 2014;40(8):900-908 [FREE Full text] [doi: 10.1016/j.ctrv.2014.05.010] [Medline: 24953564]

5. Jia Y, Sun C, Liu Z, Wang W, Zhou X. Primary breast diffuse large B-cell lymphoma: a population-based study from 1975 to 2014. Oncotarget 2018;9(3):3956-3967 [FREE Full text] [doi: 10.18632/oncotarget.23285] [Medline: 29423097]
6. Guo HY, Zhao XM, Li J, Hu XC. Primary non-Hodgkin's lymphoma of the breast: eight-year follow-up experience. Int J Hematol 2008;87(5):491-497 [FREE Full text] [doi: 10.1007/s12185-008-0085-4] [Medline: 18414980]
7. Giardini R, Piccolo C, Rilke F. Primary non-Hodgkin's lymphomas of the female breast. Cancer 1992;69(3):725-735 [FREE Full text] [doi: 10.1002/1097-0142(19920201)69:3<725::aid-cncr2820690320>3.0.co;2-c] [Medline: 1730123]
8. Liu MT, Hsieh CY, Wang AY, Pi CP, Chang TH, Huang CC, et al. Primary breast lymphoma: a pooled analysis of prognostic factors and survival in 93 cases. Ann Saudi Med 2005;25(4):288-293 [FREE Full text] [doi: 10.5144/0256-4947.2005.288] [Medline: 16212120]
9. Pérez FF, Lavernia J, Aguiar-Bujanda D, Miramón J, Gumá J, Álvarez R, et al. Primary breast lymphoma: analysis of 55 cases of the Spanish lymphoma oncology group. Clin Lymphoma Myeloma Leuk 2017;17(3):186-191 [FREE Full text] [doi: 10.1016/j.clml.2016.09.004] [Medline: 27847267]
10. Kewan T, Covut F, Ahmed R, Haddad A, Daw H. Clinical characteristics and outcomes of primary breast lymphoma: the Cleveland clinic experience. Cureus 2020;12(6):e8611 [FREE Full text] [doi: 10.7759/cureus.8611] [Medline: 32676248]
11. Takahashi H, Sakai R, Sakuma T, Matsumura A, Miyashita K, Ishii Y, et al. Comparison of clinical features between primary and secondary breast diffuse large B cell lymphoma: a Yokohama cooperative study group for hematology multicenter retrospective study. Indian J Hematol Blood Transfus 2021;37(1):60-66 [FREE Full text] [doi: 10.1007/s12288-020-01307-7] [Medline: 33707836]
12. Moura C, Leite MI, Parreira R, Medeiros A. Primary breast lymphoma. J Surg Case Rep 2020;2020(1):rjz405 [FREE Full text] [doi: 10.1093/jscr/rjz405] [Medline: 31976067]
13. Thomas A, Link BK, Altekruse S, Romitti PA, Schroeder MC. Primary breast lymphoma in the United States: 1975-2013. J Natl Cancer Inst 2017;109(6):djw294 [FREE Full text] [doi: 10.1093/jnci/djw294] [Medline: 28376147]
14. Lyou CY, Yang SK, Choe DH, Lee BH, Kim KH. Mammographic and sonographic findings of primary breast lymphoma. Clin Imaging 2007;31(4):234-238 [FREE Full text] [doi: 10.1016/j.clinimag.2007.02.028] [Medline: 17599616]
15. Peng F, Li J, Mu S, Cai L, Fan F, Qin Y, et al. Epidemiological features of primary breast lymphoma patients and development of a nomogram to predict survival. Breast 2021;57:49-61 [FREE Full text] [doi: 10.1016/j.breast.2021.03.006] [Medline: 33774459]
16. Jennings WC, Baker RS, Murray SS, Howard CA, Parker DE, Peabody LF, et al. Primary breast lymphoma: the role of mastectomy and the importance of lymph node status. Ann Surg 2007;245(5):784-789 [FREE Full text] [doi: 10.1097/01.sla.0000254418.90192.59] [Medline: 17457172]
17. Foo MY, Lee WP, Seah CMJ, Kam C, Tan SM. Primary breast lymphoma: a single-centre experience. Cancer Rep (Hoboken) 2019;2(1):e1140 [FREE Full text] [doi: 10.1002/cnr2.1140] [Medline: 32721093]
18. Picasso R, Tagliafico A, Calabrese M, Martinoli C, Pistoia F, Rossi A, et al. Primary and secondary breast lymphoma: focus on epidemiology and imaging features. Pathol Oncol Res 2020;26(3):1483-1488 [FREE Full text] [doi: 10.1007/s12253-019-00730-0] [Medline: 31428995]
19. Goda JS, Gospodarowicz M, Pintilie M, Wells W, Hodgson DC, Sun A, et al. Long-term outcome in localized extranodal mucosa-associated lymphoid tissue lymphomas treated with radiotherapy. Cancer 2010;116(16):3815-3824 [FREE Full text] [doi: 10.1002/cncr.25226] [Medline: 20564130]
20. Wilder RB, Jones D, Tucker SL, Fuller LM, Ha CS, McLaughlin P, et al. Long-term results with radiotherapy for stage I-II follicular lymphomas. Int J Radiat Oncol Biol Phys 2001;51(5):1219-1227 [FREE Full text] [doi: 10.1016/s0360-3016(01)01747-3] [Medline: 11728680]
21. Tang F, Ishwaran H. Random forest missing data algorithms. Stat Anal Data Min 2017;10(6):363-377 [FREE Full text] [doi: 10.1002/sam.11348] [Medline: 29403567]
22. Miranda-Filho A, Piñeros M, Znaor A, Marcos-Gragera R, Steliarova-Foucher E, Bray F. Global patterns and trends in the incidence of non-Hodgkin lymphoma. Cancer Causes Control 2019;30(5):489-499 [FREE Full text] [doi: 10.1007/s10552-019-01155-5] [Medline: 30895415]
23. Morton LM, Slager SL, Cerhan JR, Wang SS, Vajdic CM, Skibola CF, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes: the InterLymph non-Hodgkin lymphoma subtypes project. J Natl Cancer Inst Monogr 2014;2014(48):130-144 [FREE Full text] [doi: 10.1093/jncimonographs/lgu013] [Medline: 25174034]
24. Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob Health 2016;4(9):e609-e616 [FREE Full text] [doi: 10.1016/S2214-109X(16)30143-7] [Medline: 27470177]
25. Gibson TM, Morton LM, Shiels MS, Clarke CA, Engels EA. Risk of non-Hodgkin lymphoma subtypes in HIV-infected people during the HAART era: a population-based study. AIDS 2014;28(15):2313-2318 [FREE Full text] [doi: 10.1097/QAD.0000000000000428] [Medline: 25111081]
26. Marcos-Gragera R, Pollán M, Chirlaque MD, Gumà J, Sanchez MJ, Garau I, Non-Hodgkin's Lymphoma Working Group. Attenuation of the epidemic increase in non-Hodgkin's lymphomas in Spain. Ann Oncol 2010;21(Suppl 3):iii90-iii96 [FREE Full text] [doi: 10.1093/annonc/mdq088] [Medline: 20427366]

27. Non-Hodgkin's lymphoma statistics. Cancer Research UK. 2018. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/non-hodgkin-lymphoma [accessed 2019-01-21]
28. Ali MKM, Samhouri Y, Law JY, Khan C, Wegner R, Lee ST, et al. Patterns of treatment and their outcomes in primary breast lymphoma; a comprehensive population-based analysis. Breast Cancer 2022;29(6):1022-1031 [FREE Full text] [doi: 10.1007/s12282-022-01382-w] [Medline: 35788963]
29. Huang W, Shang Q, Xiao X, Zhang H, Gu Y, Yang L, et al. Raman spectroscopy and machine learning for the classification of esophageal squamous carcinoma. Spectrochim Acta A Mol Biomol Spectrosc 2022;281:121654 [FREE Full text] [doi: 10.1016/j.saa.2022.121654] [Medline: 35878494]
30. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. Artif Intell Med 2005;34(2):113-127 [FREE Full text] [doi: 10.1016/j.artmed.2004.07.002] [Medline: 15894176]
31. Demircioğlu A. The effect of preprocessing filters on predictive performance in radiomics. Eur Radiol Exp 2022;6(1):40 [FREE Full text] [doi: 10.1186/s41747-022-00294-w] [Medline: 36045274]
32. Massafra R, Comes MC, Bove S, Didonna V, Diotaiuti S, Giotta F, et al. A machine learning ensemble approach for 5- and 10-year breast cancer invasive disease event classification. PLoS One 2022;17(9):e0274691 [FREE Full text] [doi: 10.1371/journal.pone.0274691] [Medline: 36121822]
33. Caon J, Wai ES, Hart J, Alexander C, Truong PT, Sehn LH, et al. Treatment and outcomes of primary breast lymphoma. Clin Breast Cancer 2012;12(6):412-419 [FREE Full text] [doi: 10.1016/j.clbc.2012.07.006] [Medline: 23018097]
34. Talwalkar SS, Miranda RN, Valbuena JR, Routbort MJ, Martin AW, Medeiros LJ. Lymphomas involving the breast: a study of 106 cases comparing localized and disseminated neoplasms. Am J Surg Pathol 2008;32(9):1299-1309 [FREE Full text] [doi: 10.1097/PAS.0b013e318165eb50] [Medline: 18636016]
35. Edwards BK, Noone AM, Mariotto AB, Simard EP, Boscoe FP, Henley SJ, et al. Annual report to the nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. Cancer 2014;120(9):1290-1314 [FREE Full text] [doi: 10.1002/cncr.28509] [Medline: 24343171]
36. Mohile SG, Dale W, Somerfield MR, Hurria A. Practical assessment and management of vulnerabilities in older patients receiving chemotherapy: ASCO guideline for geriatric oncology summary. J Oncol Pract 2018;14(7):442-446 [FREE Full text] [doi: 10.1200/JOP.18.00180] [Medline: 29932846]
37. Yoshida S, Nakamura N, Sasaki Y, Yoshida S, Yasuda M, Sagara H, et al. Primary breast diffuse large B-cell lymphoma shows a non-germinal center B-cell phenotype. Mod Pathol 2005;18(3):398-405 [FREE Full text] [doi: 10.1038/modpathol.3800266] [Medline: 15492762]
38. Jesinger RA. Breast anatomy for the interventionalist. Tech Vasc Interv Radiol 2014;17(1):3-9 [FREE Full text] [doi: 10.1053/j.tvir.2013.12.002] [Medline: 24636325]
39. Tanis PJ, Nieweg OE, Olmos RAV, Kroon BB. Anatomy and physiology of lymphatic drainage of the breast from the perspective of sentinel node biopsy. J Am Coll Surg 2001;192(3):399-409 [FREE Full text] [doi: 10.1016/s1072-7515(00)00776-6] [Medline: 11245383]
40. Surveillance, Epidemiology, and End Results Program. URL: https://seer.cancer.gov/ [accessed 2023-05-31]

Abbreviations

AAPC: average annual percent change
AUC: area under the curve
DLBCL: diffuse large B-cell lymphoma
DSS: disease-specific survival
HR: hazard ratio
MALT: mucosa-associated lymphoid tissue
NHL: non-Hodgkin lymphoma
OS: overall survival
PBL: primary breast lymphoma
SEER: Surveillance, Epidemiology, and End Results

Edited by A Mavragani, T Sanchez; submitted 02.01.23; peer-reviewed by F Gomez, S Han; comments to author 15.03.23; revised version received 02.04.23; accepted 08.05.23; published 08.06.23.

Please cite as:
Yu Y, Xu Z, Shao T, Huang K, Chen R, Yu X, Zhang J, Han H, Song C
Epidemiology and a Predictive Model of Prognosis Index Based on Machine Learning in Primary Breast Lymphoma: Population-Based Study
JMIR Public Health Surveill 2023;9:e45455
URL: https://publichealth.jmir.org/2023/1/e45455
doi:10.2196/45455
PMID:37169516

Original Paper

Temporal and Bidirectional Association Between Blood Pressure Variability and Arterial Stiffness: Cross-Lagged Cohort Study

Zhiyuan Wu[1,2*], PhD; Haiping Zhang[1*], PhD; Yutao Wang[3], MD; Zhiwei Li[1], PhD; Xia Li[4], PhD; Lixin Tao[1], PhD; Xiuhua Guo[1], PhD

[1]School of Public Health, Capital Medical University, Beijing, China
[2]Center of Precision Health, Edith Cowan University, Perth, Australia
[3]Shanghai Fufan Information Technology Ltd Co, Shanghai, China
[4]Department of Mathematics and Statistics, La Trobe University, Melbourne, Australia
[*]these authors contributed equally

Corresponding Author:
Xiuhua Guo, PhD
School of Public Health
Capital Medical University
No.10 Xitoutiao, Youanmen Street
Beijing, 100069
China
Phone: 86 83911508
Email: statguo@ccmu.edu.cn

Abstract

Background: The causal relationship between blood pressure variability (BPV) and arterial stiffness remains debated.

Objective: This study aimed to explore the temporal and bidirectional associations between long-term BPV and arterial stiffness using a cohort design with multiple surveys.

Methods: Participants from the Beijing Health Management Cohort who underwent health examinations from visit 1 (2010-2011) to visit 5 (2018-2019) were enrolled in this study. Long-term BPV was defined as intraindividual variation using the coefficient of variation (CV) and SD. Arterial stiffness was measured by brachial-ankle pulse wave velocity (baPWV). The bidirectional relationship between BPV and arterial stiffness was explored using cross-lagged analysis and linear regression models, with records before and after visit 3 categorized as phase 1 and phase 2, respectively.

Results: Of the 1506 participants, who were a mean of 56.11 (SD 8.57) years old, 1148 (76.2%) were male. The cross-lagged analysis indicated that the standardized coefficients of BPV at phase 1 directing to the baPWV level at phase 2 were statistically significant but not vice-versa. The adjusted regression coefficients of the CV were 4.708 (95% CI 0.946-8.470) for systolic blood pressure, 3.119 (95% 0.166-6.073) for diastolic pressure, and 2.205 (95% CI 0.300-4.110) for pulse pressure. The coefficients of the SD were 4.208 (95% CI 0.177-8.239) for diastolic pressure and 4.247 (95% CI 0.448-8.046) for pulse pressure. The associations were predominant in the subgroup with hypertension, but we did not observe any significant association of baPWV level with subsequent BPV indices.

Conclusions: The findings supported a temporal relationship between long-term BPV and arterial stiffness level, especially among people with hypertension.

(JMIR Public Health Surveill 2023;9:e45324) doi:10.2196/45324

KEYWORDS

blood pressure variability; brachial-ankle pulse wave velocity; arterial stiffness; cross-lagged panel analysis

Introduction

Hypertension is one of the most significant risk factors for heart failure, atrial fibrillation, stroke, chronic kidney disease, and overall cardiovascular disease (CVD), with a largely increasing health burden [1,2]. According to the Chinese Hypertension Survey, between 2012 and 2015, the prevalence of hypertension was 23.2% (approximately 244.5 million) among adults and over 55% in people aged 65 and older [3], imposing a significant life loss [4,5]. Arterial stiffness is an important pathological mechanism of high blood pressure that leads to various adverse outcomes [4,6]. Arterial stiffness, as an independent predictor of CVD, could cause microvascular damage to organs rich in low-resistance vascular tissue (kidney, brain, etc), affecting the myocardium and vascular function through afterload [7,8]. Both arterial stiffness and blood pressure are aggravated with aging and can lead to cardiovascular events. In addition, dietary and lifestyle factors are also important causes of arterial stiffness progression, and studies have shown that a plant-based diet can significantly reduce arterial stiffness and blood pressure levels [9-14]. Therefore, the precise identification of arterial stiffness risk markers for primary prevention and personalized treatment based on modifiable factors is essential to prevent the progression of arterial stiffness and reduce the risk of CVD.

Previous studies have shown that elevated blood pressure precedes arterial stiffness [15]. In addition to blood pressure, blood pressure variability (BPV), such as visit-to-visit BPV, day-to-day home BPV, and 24-hour ambulatory BPV, is related to CVD [16-19]. Numerous studies have shown that all types of BPV are positively associated with adverse cardiovascular events (coronary heart disease, stroke, CVD, mortality, etc) in people with hypertension [20-22], indicating the potential predictive capacity of BPV for an unfavorable arterial stiffness level. On the other hand, several studies have shown that age-dependent arterial stiffness could also cause exaggerated BPV [23,24]. An assessment of the temporal association between BPV and arterial stiffness can facilitate the early identification of cardiovascular risk and subsequent intervention [25]. However, the cross-sectional design of most existing studies cannot indicate the temporal relationship between BPV and arterial stiffness [24,26,27]; thus, the causal association between BPV and arterial stiffness needs further research [28].

In this study, we aimed to comprehensively investigate the temporal and bidirectional relationship between long-term BPV and arterial stiffness using a cross-lagged panel design and further explore the interaction effect of hypertension status in a large cohort.

Methods

Study Design and Population

This study included participants from the Beijing Health Management Cohort, which is a large-scale longitudinal cohort study investigating the risk factors and biomarkers of cardiometabolic disorders. Our study is a secondary analysis of the Beijing Health Management Cohort under a longitudinal cohort design with multiple consecutive surveys. In total, 2381 individuals aged 18 years or above underwent multiple comprehensive biennial health examinations at visit 1 (2010-2011), visit 2 (2012-2013), visit 3 (2014-2015), visit 4 (2016-2017), and visit 5 (2018-2019), and arterial stiffness data were collected. We excluded 760 individuals with missing blood pressure data at any health visit. To minimize the possible effect of reverse causality, 115 participants with CVD (coronary heart disease, myocardial infarction, and stroke) at visit 3 (as the baseline) were further excluded. Finally, this study was restricted to a subset of 1506 participants with complete data. The period from visit 1 to visit 3 was defined as phase 1, and the period from visit 3 to visit 5 was defined as phase 2, as illustrated in Figure S1 of Multimedia Appendix 1.

Ethics Approval

This study was approved by the Ethics Committee of Capital Medical University (2020SY031). The study data were anonymous. All participants provided written informed consent before taking part in this study.

Measurement and Data Collection

The participants' sociodemographic characteristics and lifestyle and health-related information were collected via a standard questionnaire by our trained staff, including age, sex, education, smoking status, drinking status, regular physical activity, diagnosis history of diseases, and medication information. Educational level was categorized as no or some primary school, middle school or high school, and bachelor's degree or above. Smoking was defined as "current," "former," and "never." Drinking status was defined as "currently drinking" and "not currently drinking." Regular physical activity was defined as having moderate or intense activity at work or during leisure time more than 4 times and 80 minutes per week. Vigorous activities make breathing much harder than normal and include heavy lifting, digging, plowing, aerobics, fast bicycling, and cycling with a heavy load. Moderate activities make breathing somewhat harder than normal and include carrying light loads, bicycling at a regular pace, mopping the floor, doing Taiji, and walking briskly. Mild activity refers to waking at work, at home, walking to travel from place to place, and any other walking that one might do solely for recreation, sport, exercise, or leisure [29].

The physical and biochemical examination data were acquired from the electronic medical record system. BMI was calculated as weight (in kg)/height (in meters squared). Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were presented as the average of 2 measurements on the right arm using a sphygmomanometer after resting for at least 10 minutes. The pulse pressure (PP) was calculated as SBP minus DBP. Hypertension status was defined as SBP ≥140 mmHg or DBP ≥90 mmHg or use of any antihypertensive medication or self-reported history of hypertension according to the Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC-7) criteria [30]. Data on serum total cholesterol (TC), triglyceride, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), fasting blood glucose (FBG), glycated hemoglobin (HbA1c), and plasma high-sensitivity C-reactive protein (hs-CRP) were also collected in this study. Diabetes was defined as FBG ≥7.0 mmol/l, HbA1c ≥6.5%, or the use of any glucose-lowering medication or self-reported history of diabetes according to the American Diabetes Association [31]. Moreover, according to the Guidelines on Prevention and Treatment of Dyslipidemia for Chinese Adults [32], dyslipidemia was defined as triglyceride ≥2.3 mmol/l, TC ≥6.2 mmol/l, LDL-C ≥4.1 mmol/l, HDL-C <1.0 mmol/l, any lipid-lowering medication or self-reported history of dyslipidemia.

Assessment of BPV and Arterial Stiffness

BPV was assessed at phase 1 (visit 1 to visit 3) and phase 2 (visit 3 to visit 5). Long-term BPV was defined as the intraindividual variation in SBP, DBP, and PP at 3 visits and measured by the coefficient of variation (CV) and SD. The mean value of 3 blood pressure indices at each phase was also calculated.

The arterial stiffness level was measured by brachial-ankle pulse wave velocity (baPWV), which is a simple, noninvasive, automatic method widely used in population-based studies. The baPWV was measured with an Omron Colin BP-203RPE III device (Omron Health Care). The time interval between the wave front of the brachial waveform and the ankle waveform is expressed as the time interval between the brachium and ankle (ΔTba). La indicates the path length from the suprasternal notch to the ankle, and Lb refers to the path length from the suprasternal notch to the brachium based on the height of the subjects. The final baPWV was calculated as (La-Lb)/ΔTba, which has been previously described [33]. The arterial stiffness level was defined as the mean of baPWV measurements at phases 1 and 2.

Statistical Analyses

Baseline characteristics were presented as the mean (SD), median (IQR P_{25}-P_{75}), or number (percentage), as appropriate. We used the cross-lagged panel model to analyze the temporal relationship between BPV and baPWV levels. The lagged effect between the observed variables, the correlation between the variables at each phase, and the residual variance of the observed variables were estimated in the cross-lagged panel analysis, which elucidates the temporal relationship between correlated variables and indicates the precursor variable [34,35].

Then, a multivariable linear regression model was used to estimate the bidirectional associations between BPV and baPWV levels. We calculated the regression coefficients of BPV indices at phase 1, with baPWV at phase 2 as the independent variable, and the coefficients of baPWV level at phase 1, with BPV indices at phase 2 as the independent variable. To adjust for potential confounding factors, three models were established: (1) model 1 was adjusted for the individual-level mean blood pressure; (2) model 2 was adjusted for mean blood pressure, age, and sex; and (3) model 3 was further adjusted for BMI, education level, smoking, drinking, physical activity, hypertension, dyslipidemia, diabetes, triglyceride, TC, FBG, HbA1c, and hs-CRP. In total, there were 17 variables in the regression model. Thus, a group of 1506 participants in this study could satisfy the sample size requirement given the principle of 10 samples per variable in the regression model [36]. Furthermore, the blood pressure at visit 3 was adjusted in the models in place of the mean blood pressure value in the sensitivity analysis. The difference was considered statistically significant at 2-sided $P<.05$, and a 95% CI was provided.

In addition, we analyzed the association of BPV with arterial stiffness levels in the subgroups with and without hypertension. The use of antihypertensive medication was additionally adjusted for in the subgroup with hypertension, considering that antihypertensive agents are an important determinant of BPV and are related to arterial stiffness levels. The difference was considered statistically significant at 2-sided $P<.1$, and a 90% CI was provided in the subgroup analysis [37].

All the analyses were conducted using R software (version 4.1.0; R Foundation for Statistical Computing), using the packages *lavaan* for cross-lagged panel analysis and *ggplot2* for visualization.

Results

Characteristics

Among the 1506 participants, the mean age was 56.11 (SD 8.57) years, and 661 (43.9%) were diagnosed with hypertension at visit 3. The median value of the cumulative average level of baPWV was 1395.50 (IQR 1298.12-1544.34) cm/s at phase 1 and 1438.25 (IQR 1344.00-1588.67) cm/s at phase 2. Detailed characteristics of the study population and the variability indices of blood pressure are summarized in Tables 1 and 2, respectively.

Table 1. Characteristics of the study population (N=1506) at baseline (visit 3).

Characteristics	Variables
Age (years), mean (SD)	56.11 (8.57)
Sex, n (%)	
Male	1148 (76.2)
Female	358 (23.8)
BMI[a] (kg/m^2), mean (SD)	25.67 (2.90)
Education level, n (%)	
Primary	122 (8.1)
Secondary	794 (52.7)
Postsecondary	590 (39.2)
Physical activity, n (%)	751 (49.9)
Smoking, n (%)	
None	851 (56.5)
Former	218 (14.5)
Current	437 (29.0)
Current drinking, n (%)	931 (61.8)
Hypertension, n (%)	661 (43.9)
Antihypertensive medication, n (%)	507 (33.7)
Diabetes, n (%)	250 (16.6)
Dyslipidemia, n (%)	667 (44.3)
SBP[b] (mmHg), mean (SD)	122.09 (14.59)
DBP[c] (mmHg), mean (SD)	71.05 (9.68)
Fasting plasma glucose[d] (mmol/L), mean (SD)	5.58 (1.04)
Glycated hemoglobin, mean (SD)	5.71 (0.63)
Triglyceride[d] (mmol/L), mean (SD)	1.63 (0.99)
TC[d] (mmol/L), mean (SD)	4.80 (0.86)
hs-CRP[e] (mg/L), median (IQR)	0.69 (0.28-1.27)

[a] Calculated as weight in kilograms divided by height in meters squared.

[b] SBP: systolic blood pressure.

[c] DBP: diastolic blood pressure.

[d] TC: total cholesterol. To convert fasting plasma glucose to mg/dl, multiply by 18; to convert triglyceride to mg/dl, multiply by 88.6; to convert TC to mg/dl, multiply by 38.66.

[e] hs-CRP: high-sensitivity C-reactive protein.

Table 2. BPV[a] and arterial stiffness parameters.

BPV measurements	Phase 1	Phase 2
CV[b] of SBP[c], median (IQR)	5.57 (2.99-8.74)	5.81 (3.51-8.85)
CV of DBP[d], median (IQR)	7.06 (3.67-10.95)	6.60 (3.98-10.18)
CV of PP[e], median (IQR)	10.88(6.02-16.62)	10.70(6.29-15.71)
SD of SBP, median (IQR)	7.00 (3.54-10.99)	7.07 (4.24-10.61)
SD of DBP, median (IQR)	4.95 (2.83-7.98)	4.62 (2.83-7.09)
SD of PP, median (IQR)	5.29(2.83-8.14)	5.29(3.06-7.78)
baPWV[f] level, median (IQR), cm/s	1395.50 (1298.12-1544.34)	1438.25 (1344.00-1588.67)

[a] BPV: blood pressure variability.

[b] CV: coefficient of variation.

[c] SBP: systolic blood pressure.

[d] DBP: diastolic blood pressure.

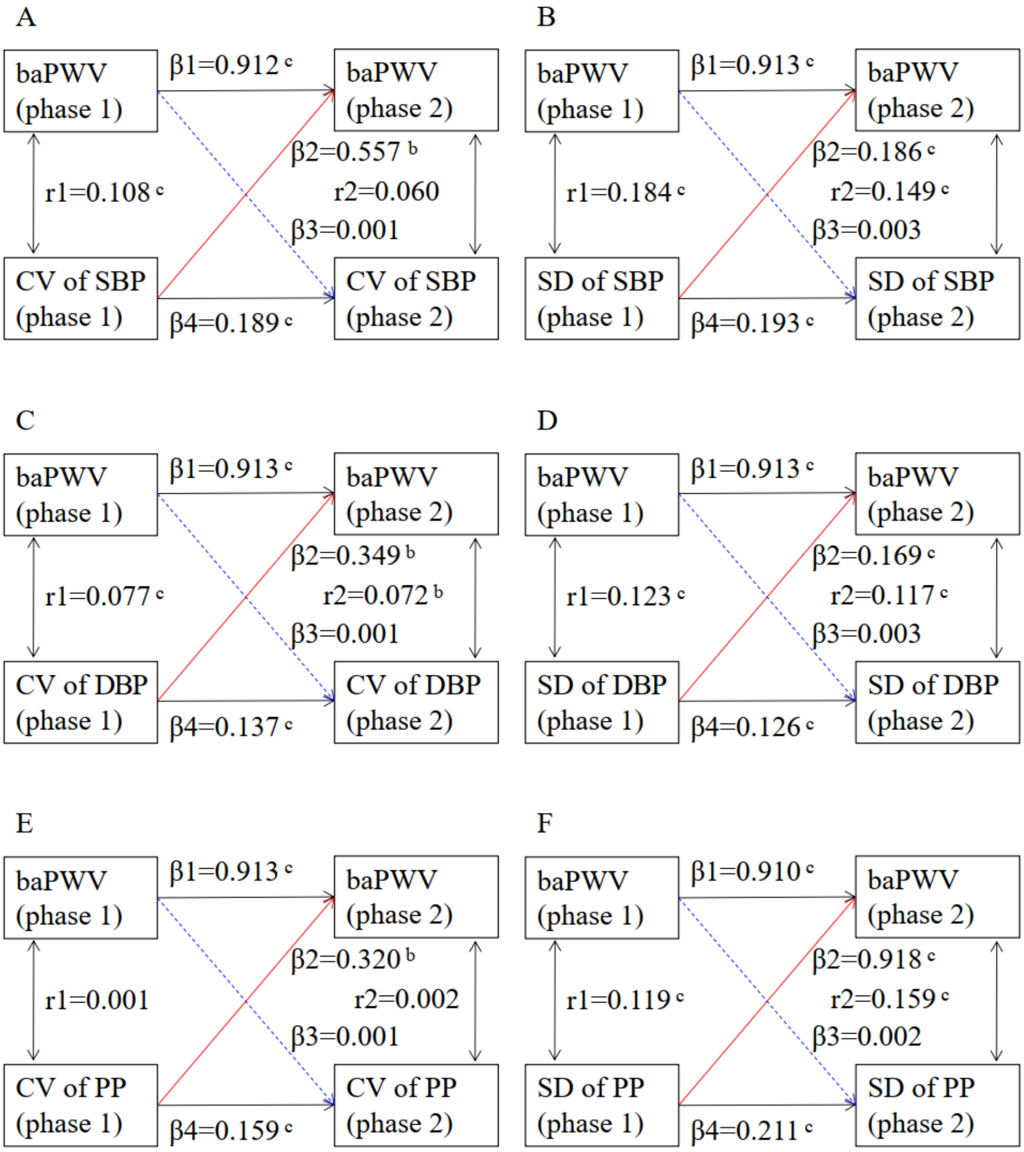
[e] PP: pulse pressure.

[f] baPWV: brachial-ankle pulse wave velocity.

Temporal Relationship Between BPV and baPWV

Assessments of BPV indices and baPWV at phase 1 and phase 2 constituted a typical cross-lagged panel design. Figure 1 presents the association of BPV indices at phase 1 with baPWV level at phase 2 and the association of baPWV at phase 1 with the BPV indices at phase 2. The standardized coefficients of the CV of SBP, DBP, and PP were 0.557, 0.349, and 0.320, respectively, and the SDs of SBP, DBP, and PP were 0.186, 0.169, and 0.918, respectively, to the follow-up baPWV level (all *P*<.01), while the standardized coefficients of baPWV level at phase 1 to the subsequent BPV indices were all weak and insignificant.

Figure 1. Bidirectional analysis of blood pressure variability (BPV) indices and brachial-ankle pulse wave velocity (baPWV) level by cross-lagged panel analysis. β1: baPWV at phase 1→baPWV at phase 2; β2: BPV at phase 1→baPWV at phase 2; β3: baPWV at phase 1→BPV at phase 2; β4: BPV at phase 1→BPV at phase 2; r1: correlation coefficient between BPV and baPWV at phase 1; r2: correlation coefficient between BPV and baPWV at phase 2. A: BPV was measured by coefficient of variation (CV) of systolic blood pressure (SBP); B: BPV was measured by SD of SBP; C: BPV was measured by CV of diastolic blood pressure (DBP); D: BPV was measured by SD of DBP; E: BPV was measured by CV of pulse pressure (PP); F: BPV was measured by SD of PP. a means *P*<.05; b means *P*<.01; c means *P*<.001.



Bidirectional Relationship Between BPV and baPWV

In the fully adjusted model (model 3), a 1-unit increase in the CV of SBP, DBP, and PP at phase 1 was associated with a 4.708 (95% CI 0.946-8.470), 3.119 (95% CI 0.166-6.073), and 2.205 (95% CI 0.300-4.110) increase, respectively, in the baPWV level at phase 2. Similarly, a 1-unit increase in the SD of DBP and PP at phase 1 was associated with a 4.208 (95% CI 0.177-8.239) and 4.247 (95% CI 0.448-8.046) increase in the baPWV level at phase 2 (Table 3). The positive linear regression lines of the BPV indices at phase 1 and the baPWV level at phase 2 are presented in Figure S2 of Multimedia Appendix 1. In contrast, we did not observe any significant association between the baPWV level at phase 1 and the subsequent BPV indices (Table S1 of Multimedia Appendix 1).

Table 3. Association between BPV[a] at phase 1 and the cumulative average level of baPWV[b] at phase 2[c].

BPV measurements	Model 1[d]			Model 2[e]			Model 3[f]		
	Coefficient	95% CI	*P* value	Coefficient	95% CI	*P* value	Coefficient	95% CI	*P* value
SBP[g] at phase 1									
CV[h]	4.748	2.617 to 6.880	<.001	2.320	0.559 to 4.080	.01	4.708	0.946 to 8.470	.02
SD	3.714	2.001 to 5.428	<.001	1.800	0.385 to 3.215	.01	1.666	−0.870 to 4.202	.20
DBP[i] at phase 1									
CV	4.571	2.662 to 6.480	<.001	−0.246	−1.721 to 1.230	.74	3.119	0.166 to 6.073	.04
SD	6.099	3.499 to 8.699	<.001	−0.133	−2.138 to 1.872	.90	4.208	0.177 to 8.239	.04
PP[j] at phase 1									
CV	2.088	0.983 to 3.193	<.001	1.309	0.335 to 2.283	.009	2.205	0.300 to 4.110	.02
SD	4.416	2.234 to 6.599	<.001	2.777	0.852 to 4.703	.005	4.247	0.448 to 8.046	.03

[a] BPV: blood pressure variability.

[b] baPWV: brachial-ankle pulse wave velocity.

[c] Phase 1 refers to the period from visit 1 to visit 3, and phase 2 refers to the period from visit 3 to visit 5.

[d] Adjusted for the mean blood pressure level during phase 1.

[e] Adjusted for the mean blood pressure level during phase 1, age, and sex.

[f] Adjusted for BMI, education level, smoking, drinking, physical activity, hypertension or not, dyslipidemia or not, diabetes or not, triglyceride, total cholesterol (TC), fasting glucose, glycated hemoglobin (HbA1c), and high-sensitivity C-reactive protein (hs-CRP), in addition to the covariates in model 2.

[g] SBP: systolic blood pressure.

[h] CV: coefficient of variation.

[i] DBP: diastolic blood pressure.

[j] PP: pulse pressure.

Sensitivity Analysis

The results remained almost consistent when the blood pressure level at visit 3 was alternatively adjusted in the model, apart from PP (Table S2 of Multimedia Appendix 1). Spearman correlation coefficients among CV, SD, mean, maximum, and minimum SBP, DBP, and PP at the 2 phases are shown in Table S3 in Multimedia Appendix 1. There were strong correlations between the CV and SD of blood pressure indices and weak correlations of BPV indices with the mean, maximum, and minimum blood pressure, which indicated that BPV indices could reflect a novel feature beyond blood pressure itself.

Subgroup Analysis

We found that the relationship between BPV indices at phase 1 and baPWV level at phase 2 was affected by hypertension status (Figure 2). The CV and SD of SBP and PP were suggested to be associated with the subsequent baPWV level in participants with hypertension. A 1-unit increase in the CV of SBP and PP at phase 1 was associated with a 4.861 (90% CI 0.642-9.081) and 2.733 (90% CI 0.335-5.131) increase in baPWV level at phase 2, respectively, and a 1-unit increase in the SD of SBP and PP was associated with a 3.283 (90% CI 0.083-6.484) and 4.740 (90% CI 0.299-9.182) increase in baPWV, respectively. Table 4 summarizes the impact of the mean blood pressure, BPV indices, and antihypertensive medication (only in participants with hypertension) on the follow-up baPWV level. The direct blood pressure level was positively associated with the baPWV level in participants with both hypertension and normal blood pressure, and the use of antihypertensive medication was significantly associated with a lowered baPWV level in the subgroup with hypertension.

Figure 2. Associations of blood pressure variability (BPV) indices and the subsequent artery stiffness level in participants with both hypertension and normal blood pressure.The coefficient was adjusted for the mean blood pressure, age, sex, BMI, education level, smoking, drinking, physical activity, hypertension or not, dyslipidemia or not, diabetes or not, triglyceride, total cholesterol, fasting glucose, glycated hemoglobin (HbA1c), hs-CRP (high-sensitivity C-reactive protein), and antihypertensive medication use (in participants with hypertension). DBP: diastolic blood pressure; PP: pulse pressure; SBP: systolic blood pressure.

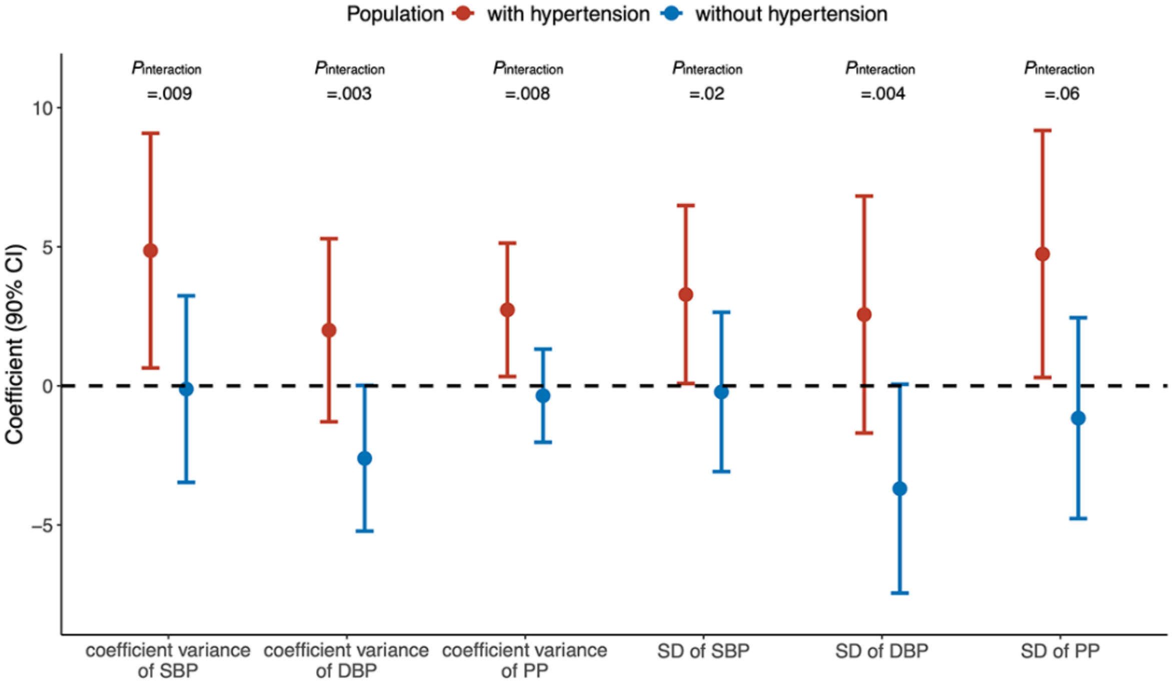


Table 4. Association of BPV[a], mean blood pressure, and use of antihypertensive medication with the subsequent baPWV[b] levels among participants with both hypertension and normal blood pressure[c].

Measurements	Hypertension			Normal blood pressure		
	Coefficient[d]	90% CI	*P* value	Coefficient[d]	90% CI	*P* value
CV[e] of SBP[f]	4.861	0.642 to 9.081	.06	−0.117	−3.471 to 3.237	.95
Mean of SBP	4.273	2.651 to 5.895	<.001	4.554	3.227 to 5.882	<.001
Antihypertensive medication	−54.524	−102.105 to −6.944	.06	N/A[g]	N/A	N/A
SD of SBP	3.283	0.083 to 6.484	.09	−0.219	−3.083 to 2.645	.90
Mean of SBP	4.029	2.373 to 5.685	<.001	4.569	3.227 to 5.911	<.001
Antihypertensive medication	−54.617	−102.292 to −6.942	.06	N/A	N/A	N/A
CV of DBP	2.000	−1.293 to 5.293	.32	−2.609	−5.224 to 0.013	.11
Mean of DBP	4.653	2.078 to 7.227	.003	6.208	4.104 to 8.313	<.001
Antihypertensive medication	−62.483	−111.44 to −13.525	.04	N/A	N/A	N/A
SD of DBP[h]	2.564	−1.698 to 6.825	.32	−3.698	−7.452 to 0.056	.11
Mean of DBP	4.411	1.762 to 7.059	.007	6.492	4.335 to 8.649	<.001
Antihypertensive medication	−62.168	−111.086 to −13.25	.04	N/A	N/A	N/A
CV of PP[i]	2.733	0.335 to 5.131	.06	−0.353	−2.028 to 1.321	.73
Mean of PP	4.116	1.964 to 6.268	.002	4.505	2.45 to 6.56	<.001
Antihypertensive medication	−56.565	−105.21 to −7.921	.06	N/A	N/A	N/A
SD of PP	4.740	0.299 to 9.182	.08	−1.162	−4.774 to 2.449	.60
Mean of PP	3.492	1.319 to 5.665	.009	4.603	2.569 to 6.638	<.001
Antihypertensive medication	−57.520	−106.175 to −8.865	.049	N/A	N/A	N/A

[a] BPV: blood pressure variability.

[b] baPWV: brachial-ankle pulse wave velocity.

[c] The associations between BPV indices and baPWV measurement were examined in separate regression models.

[d] Model was adjusted for the mean blood pressure, age, sex, BMI, education level, smoking, drinking, physical activity, hypertension or not, dyslipidemia or not, diabetes or not, triglyceride, total cholesterol (TC), fasting glucose, HbA1c, hs-CRP, and antihypertensive medication use (in participants with hypertension).

[e] CV: coefficient of variation.

[f] SBP: systolic blood pressure.

[g] N/A: not applicable.

[h] DBP: diastolic blood pressure.

[i] PP: pulse pressure.

Discussion

Principal Findings

This study explored the temporal and bidirectional associations between long-term visit-to-visit BPV and arterial stiffness levels based on a Chinese general population. In the cross-lagged analyses, the BPV indices were shown to have a temporal relationship with subsequent arterial stiffness, indicating that BPV is an independent predictor of unfavorable arterial stiffness levels and precedes arteriosclerosis progression, especially in people with hypertension. The classification of the temporal relationship in this study provides a scientific basis for identifying individuals at a high risk of arterial stiffness, further offering opportunities to advance primary prevention and personalized interventions for arterial stiffness.

Blood pressure usually stabilizes within a certain range under the complex interaction of environmental factors and normal cardiovascular regulation function. Due to baroreflex dysfunction associated with arteriole remodeling, BPV, and arterial stiffness levels increase with age, but the mutual associations between BPV and arterial stiffness remain unclear [38]. Previous studies have shown that BPV, such as visit-to-visit BPV, day-by-day home BPV, or 24-hour ambulatory BPV, is an independent risk factor for CVD [24,39]. In a post hoc patient-level analysis of 7 clinical trials, Donald et al [40] assessed the association between visit-to-visit BPV and percent atheroma volume and found that in patients with medically treated coronary artery disease, higher BPV, especially systolic BPV, is significantly associated with atherosclerotic progression and adverse cardiovascular outcomes such as death, myocardial infarction, stroke, emergency revascularization for acute coronary syndrome, or hospitalization for unstable angina. In terms of short-term BPV, Sander et al [41] found that diurnal systolic BPV is a significant predictor of atherosclerosis progression in people with hypertension via long-term continuous monitoring of circadian BPV and measurement of atherosclerosis progression using common carotid artery intima-media thickness. In the progression of CVD, arterial stiffness is an important pathological basis and early lesion signal for adverse events [6]. Our study supplements the evidence about the temporal and bidirectional relationships between the visit-to-visit BPV and arterial stiffness level using a cross-lagged cohort design and indicates that BPV could precede arterial stiffness and arteriosclerosis, especially in people with hypertension.

The pathological mechanism between BPV and arterial stiffness or CVD has not been completely determined [17,28]. According to the previous literature, endothelial dysfunction caused by vascular remodeling plays an important role in BPV-induced arterial stiffness, subsequently leading to the development of CVD [28,42-44]. Keith et al [45] found that BPV is positively correlated with endothelial dysfunction, as reflected in a significant association between higher BPV with lower endothelium-dependent flow-mediated vasodilation (FMD) and lower FMD/non–endothelium-dependent nitroglycerin-mediated vasodilation (NMD) ratio, which suggests that endothelium-specific vasodilation is impaired in individuals with higher BPV. Moreover, a significant association between endothelial dysfunction and arterial stiffness has been observed in clinical studies. For example, Carmel et al [44] demonstrated that brachial artery FMD is negatively correlated with aortic pulse wave velocity (PWV) in healthy individuals, which implies that endothelial dysfunction mediates the effect of BPV on arterial stiffness.

Furthermore, the hemodynamic environment changes caused by BPV may be an important reason for vascular endothelial dysfunction, which is closely related to the progression of arterial stiffness and atherosclerotic diseases [42,46]. The shear stress (SS) generated by blood flow and vascular endothelium triggers the release of various vasoactive mediators from endothelial cells, among which nitric oxide (NO), produced by endothelial-type nitric oxide synthase (eNOS or NOS3), is considered the main endothelial diastolic factor [47]. Under physiological conditions, a dynamic balance is maintained between blood flow stimulation and endothelial cell responses to maintain a stable internal environment [47]. Stable blood flow results in higher endothelial SS, lower intimal hyperplasia, and reduced atherosclerosis risk [25,48]. In contrast, complicated blood flow patterns produce low endothelial SS in the vascular wall, which facilitates the transshipment and accumulation of atherosclerotic risk factors, such as low-density lipoprotein to the intima of the vascular wall, and facilitates the inflammatory response of the vascular wall [48,49]. In addition, oscillatory SS induced by complex flow patterns can induce the expression of leukocyte adhesion molecules in endothelial cells, recruit inflammatory cells (monocytes and T cells) to the arterial wall, and initiate local vascular inflammation [49,50]. Oscillatory SS could also promote the activation of the oxidation process by increasing nicotinamide adenine dinucleotide-reduced (NADH) oxidase activity [51]. In the physiological state, the continuous release of NO from endothelial cells contributes to the functional regulation of arterial elasticity, and the combined effect of oscillatory SS and pressure could downregulate the expression level of endothelial constitutive nitric oxide synthase (EcNOS) and reduce the synthesis of NO [52]. Increased BPV might adversely affect the stability of hemodynamics in a physiological setting, thereby affecting vascular endothelial function.

In addition, previous population-based studies have shown that high BPV is strongly associated with biomarkers of vascular inflammation, independent of the mean levels of blood pressure [53]. In short, BPV could be both a cause and an indicator of arterial stiffness and atherosclerosis through hemodynamic environment changes and vascular endothelial dysfunction. In another prospective cohort, researchers found that baseline common carotid arterial intima-media thickness was not significantly associated with BPV during follow-up [41], which suggests that BPV is the initial cause of arteriosclerosis and arterial stiffness from another aspect [42] and could be the underlying pathological mechanism and therapeutic target of arterial stiffness and related atherosclerotic vascular diseases. Interventions in diet, lifestyle habits, and pharmacological treatment in high-risk groups can effectively reduce arterial stiffness levels [54,55]. The temporal association between increased BPV and arterial stiffness identified in this study further emphasizes the role of BPV levels as an early warning for the risk of arterial stiffness. Our findings are potentially valuable for delaying or avoiding the progression of arterial stiffness at a population level and promoting early and active cardiovascular health in the public health setting.

Based on the vicious cycle of hemodynamic stress and vascular disease, Kario et al [24] proposed the concept of systemic hemodynamic atherothrombotic syndrome (SHATS). Patients were assessed with scales that synergistically consider hemodynamic stress (various forms of blood pressure and BPV) and vascular disease to identify SHATS early and provide effective intervention to reduce adverse cardiovascular events and target organ damage [24]. In this study, we further demonstrated the potential role of maintaining blood pressure stability and reducing BPV to control arterial stiffness status. On the one hand, we found that the use of antihypertensive medication is associated with decreased arterial stiffness levels. Combination therapy with long-acting calcium channel blockers and angiotensin II receptor blockers has demonstrated to be effective in reducing both long-term BPV and arterial stiffness [56]. On the other hand, the association between BPV and arterial stiffness is still significant after adjusting for the use of antihypertensive medication. Previous studies on antihypertensive medication adherence also found that low adherence resulted in only a small increase in BPV compared with high adherence, and BPV was significantly associated with CVD independent of medication adherence [57,58], which underlines the comprehensive and novel strategy for controlling BPV levels. Strict BP control by taking medication, together with leading a healthy lifestyle, may play a pivotal role in suppressing arterial stiffening via BPV reduction in people with hypertension [59].

In addition, this study found that BPV preceded arterial stiffness, suggesting a high risk of CVD in people with higher BPV. Previous studies have shown that plant-based diets have protective effects against CVD. In particular, the Mediterranean diet significantly reduces arterial stiffness and blood pressure levels [9], and Dietary Approaches to Stop Hypertension (DASH) combined with exercise can significantly reduce blood pressure in people with intractable hypertension [13,14]. Additionally, moderate nut intake can also help control CVD risk factors such as weight by reducing appetite and fat absorption [11]. Enhanced monitoring of BPV and individualized interventions in diet, lifestyle habits, and pharmacotherapy for those at a high risk of increased BPV are significant in reducing the risk of arterial stiffness and the occurrence of subsequent cardiovascular events [60-62]. Based on these study findings, the integration of long-term monitoring of blood pressure levels into community-level health examinations is essential to optimize the prevention efficacy of CVD. Long-term BPV complements the performance of early CVD risk assessment at the population level compared to a single-point level of blood pressure, providing an opportunity to shift the focus from treatment to population-level risk prevention, which emphasizes the potential effectiveness of community-level continuous monitoring of blood pressure levels and early identification of increased BPV. Among people with hypertension, enhancing medication adherence and adhering to healthy dietary habits to maintain stable blood pressure over a long period are important for preventing subsequent adverse cardiovascular outcomes.

Limitations

Several limitations of this study should be acknowledged. First, different antihypertensive drugs have distinct effects on BPV and arterial stiffness levels, but detailed information on medication use is lacking in our study. The interaction effect between different antihypertensive medications and BPV on arterial stiffness levels needs further investigation. Second, this study is based on a Chinese cohort, and the observed results warrant further validation in other populations with large sample sizes. Data on household socioeconomic status and food intake were not collected in the cohort. The estimated effect size could be biased given the potential effects of socioeconomic status and dietary patterns on the arterial stiffness level. Third, the specific effect of maintaining long-term stable blood pressure on suppressing arterial stiffening and promoting cardiovascular health warrants further research.

Conclusions

Our findings suggested a temporal relationship between visit-to-visit BPV and arterial stiffness during 10-year surveys, especially in people with hypertension. Maintaining long-term stable blood pressure should be emphasized for suppressing arterial stiffness. This study further demonstrates that long-term BPV provides a new target for the early prevention of CVD in the public health setting, suggesting the potential benefit of incorporating long-term monitoring and assessment of BPV into community-level health examinations.

Acknowledgments

We thank all the staff and participants of the Beijing Health Management Cohort for their invaluable contributions. This study was funded by the National Natural Science Foundation of China (81872708 and 82073668) to author LT.

Data Availability

Data collected for this study are available on request with publication; requests should be directed to the corresponding author.

Authors' Contributions

Authors XG (statguo@ccmu.edu.cn) and LT (taolixin.2008@163.com) are co-corresponding authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental tables and figures.
[DOC File , 270 KB - publichealth_v9i1e45324_app1.doc]

References

1. Olsen MH, Angell SY, Asma S, Boutouyrie P, Burger D, Chirinos JA, et al. A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on hypertension. Lancet 2016 Dec 26;388(10060):2665-2712. [doi: 10.1016/S0140-6736(16)31134-5] [Medline: 27671667]
2. Fuchs FD, Whelton PK. High blood pressure and cardiovascular disease. Hypertension 2020 Feb;75(2):285-292. [doi: 10.1161/HYPERTENSIONAHA.119.14240] [Medline: 31865786]
3. Wang Z, Chen Z, Zhang L, Wang X, Hao G, Zhang Z, China Hypertension Survey Investigators. Status of hypertension in China: results from the China Hypertension Survey, 2012-2015. Circulation 2018 May 29;137(22):2344-2356. [doi: 10.1161/CIRCULATIONAHA.117.032380] [Medline: 29449338]

4. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980 [FREE Full text] [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
5. Zhou B, Perel P, Mensah GA, Ezzati M. Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. Nat Rev Cardiol 2021 Nov 28;18(11):785-802 [FREE Full text] [doi: 10.1038/s41569-021-00559-8] [Medline: 34050340]
6. Chirinos JA, Segers P, Hughes T, Townsend R. Large-artery stiffness in health and disease: JACC state-of-the-art review. J Am Coll Cardiol 2019 Sep 03;74(9):1237-1263 [FREE Full text] [doi: 10.1016/j.jacc.2019.07.012] [Medline: 31466622]
7. Chirinos JA, Segers P. Noninvasive evaluation of left ventricular afterload: part 2: arterial pressure-flow and pressure-volume relations in humans. Hypertension 2010 Oct;56(4):563-570. [doi: 10.1161/HYPERTENSIONAHA.110.157339] [Medline: 20733088]
8. Kaess BM, Rong J, Larson MG, Hamburg NM, Vita JA, Levy D, et al. Aortic stiffness, blood pressure progression, and incident hypertension. JAMA 2012 Sep 05;308(9):875-881 [FREE Full text] [doi: 10.1001/2012.jama.10503] [Medline: 22948697]
9. Jennings A, Berendsen AM, de Groot LCPGM, Feskens EJM, Brzozowska A, Sicinska E, et al. Mediterranean-style diet improves systolic blood pressure and arterial stiffness in older adults. Hypertension 2019 Mar;73(3):578-586 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.118.12259] [Medline: 30636547]
10. Noce A, Romani A, Bernini R. Dietary intake and chronic disease prevention. Nutrients 2021 Apr 19;13(4) [FREE Full text] [doi: 10.3390/nu13041358] [Medline: 33921568]
11. Pahlavani N, Rostami D, Ebrahimi F, Azizi-Soleiman F. Nuts effects in chronic disease and relationship between walnuts and satiety: review on the available evidence. Obesity Medicine. URL: https://doi.org/10.1016/j.obmed.2019.100173 [accessed 2023-06-13]
12. Kolahdouz-Mohammadi R, Malekahmadi M, Clayton ZS, Sadat SZ, Pahlavani N, Sikaroudi MK, et al. Effect of egg consumption on blood pressure: a systematic review and meta-analysis of randomized clinical trials. Curr Hypertens Rep 2020 Feb 29;22(3):24 [FREE Full text] [doi: 10.1007/s11906-020-1029-5] [Medline: 32114646]
13. Pahlavani N, Khayyatzadeh SS, Banazadeh V, Bagherniya M, Tayefi M, Eslami S, et al. Adherence to a Dietary Approach to Stop Hypertension (DASH)-style in relation to daytime sleepiness. Nat Sci Sleep 2020;12:325-332 [FREE Full text] [doi: 10.2147/NSS.S246991] [Medline: 32607032]
14. Blumenthal JA, Hinderliter AL, Smith PJ, Mabe S, Watkins LL, Craighead L, et al. Effects of lifestyle modification on patients with resistant hypertension: results of the TRIUMPH randomized clinical trial. Circulation 2021 Oct 12;144(15):1212-1226 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.121.055329] [Medline: 34565172]
15. Chen W, Li S, Fernandez C, Sun D, Lai C, Zhang T, et al. Temporal relationship between elevated blood pressure and arterial stiffening among middle-aged Black and White adults: the Bogalusa Heart Study. Am J Epidemiol 2016 Apr 01;183(7):599-608 [FREE Full text] [doi: 10.1093/aje/kwv274] [Medline: 26960706]
16. Yano Y, Reis JP, Lewis CE, Sidney S, Pletcher MJ, Bibbins-Domingo K, et al. Association of blood pressure patterns in young adulthood with cardiovascular disease and mortality in middle age. JAMA Cardiol 2020 Apr 01;5(4):382-389 [FREE Full text] [doi: 10.1001/jamacardio.2019.5682] [Medline: 31968050]
17. Gosmanova EO, Mikkelsen MK, Molnar MZ, Lu JL, Yessayan LT, Kalantar-Zadeh K, et al. Association of systolic blood pressure variability with mortality, coronary heart disease, stroke, and renal disease. J Am Coll Cardiol 2016 Sep 27;68(13):1375-1386 [FREE Full text] [doi: 10.1016/j.jacc.2016.06.054] [Medline: 27659458]
18. Alaarag A, Amin O. Blood pressure variability in patients with angina and non-obstructive coronary artery disease. J Hum Hypertens 2021 Dec;35(12):1074-1080. [doi: 10.1038/s41371-020-00475-3] [Medline: 33414504]
19. Chang TI, Tabada GH, Yang J, Tan TC, Go AS. Visit-to-visit variability of blood pressure and death, end-stage renal disease, and cardiovascular events in patients with chronic kidney disease. J Hypertens 2016 Feb;34(2):244-252 [FREE Full text] [doi: 10.1097/HJH.0000000000000779] [Medline: 26599220]
20. Diaz KM, Tanner RM, Falzon L, Levitan EB, Reynolds K, Shimbo D, et al. Visit-to-visit variability of blood pressure and cardiovascular disease and all-cause mortality: a systematic review and meta-analysis. Hypertension 2014 Nov;64(5):965-982 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.114.03903] [Medline: 25069669]
21. Kario K. Prognosis in relation to blood pressure variability: pro side of the argument. Hypertension 2015 Jun;65(6):1163-9; discussion 1169. [doi: 10.1161/HYPERTENSIONAHA.115.04800] [Medline: 25916727]
22. Wei F, Li Y, Zhang L, Xu T, Ding F, Wang J, et al. Beat-to-beat, reading-to-reading, and day-to-day blood pressure variability in relation to organ damage in untreated Chinese. Hypertension 2014 Apr;63(4):790-796 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.113.02681] [Medline: 24396027]
23. Hallock P, Benson IC. Studies on the Elastic Properties of Human Isolated Aorta. J Clin Invest 1937 Jul;16(4):595-602 [FREE Full text] [doi: 10.1172/JCI100886] [Medline: 16694507]
24. Ishiyama Y, Hoshide S, Kanegae H, Kario K. Increased arterial stiffness amplifies the association between home blood pressure variability and cardiac overload: the J-HOP study. Hypertension 2020 Jun;75(6):1600-1606. [doi: 10.1161/HYPERTENSIONAHA.119.14246] [Medline: 32336235]

25. Bencivenga L, De Souto Barreto P, Rolland Y, Hanon O, Vidal J, Cestac P, et al. Blood pressure variability: A potential marker of aging. Ageing Res Rev 2022 Sep;80:101677. [doi: 10.1016/j.arr.2022.101677] [Medline: 35738476]
26. Li C, Liu R, Wang J, Yang J. Relationship between blood pressure variability and target organ damage in elderly patients. Eur Rev Med Pharmacol Sci 2017 Dec;21(23):5451-5455 [FREE Full text] [doi: 10.26355/eurrev_201712_13934] [Medline: 29243789]
27. Schillaci G, Bilo G, Pucci G, Laurent S, Macquin-Mavier I, Boutouyrie P, et al. Relationship between short-term blood pressure variability and large-artery stiffness in human hypertension: findings from 2 large databases. Hypertension 2012 Aug;60(2):369-377. [doi: 10.1161/HYPERTENSIONAHA.112.197491] [Medline: 22753222]
28. Miyauchi S, Nagai M, Dote K, Kato M, Oda N, Kunita E, et al. Visit-to-visit blood pressure variability and arterial stiffness: which came first: the chicken or the egg? Curr Pharm Des 2019;25(6):685-692. [doi: 10.2174/1381612825666190329122024] [Medline: 30931845]
29. Wu Z, Zhang H, Miao X, Li H, Pan H, Zhou D, et al. High-intensity physical activity is not associated with better cognition in the elder: evidence from the China Health and Retirement Longitudinal Study. Alzheimers Res Ther 2021 Nov 03;13(1):182 [FREE Full text] [doi: 10.1186/s13195-021-00923-3] [Medline: 34732248]
30. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. JAMA 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
31. American Diabetes Association. Standards of medical care in diabetes-2020 abridged for primary care providers. Clin Diabetes 2020 Jan;38(1):10-38 [FREE Full text] [doi: 10.2337/cd20-as01] [Medline: 31975748]
32. Joint Committee for Developing Chinese Guidelines on Prevention and Treatment of Dyslipidemia in Adults. Chinese guidelines on prevention and treatment of dyslipidemia in adults. Zhonghua Xin Xue Guan Bing Za Zhi 2007 May;35(5):390-419. [Medline: 17711682]
33. Yamashina A, Tomiyama H, Takeda K, Tsuda H, Arai T, Hirose K, et al. Validity, reproducibility, and clinical significance of noninvasive brachial-ankle pulse wave velocity measurement. Hypertens Res 2002 May;25(3):359-364 [FREE Full text] [doi: 10.1291/hypres.25.359] [Medline: 12135313]
34. Han T, Meng X, Shan R, Zi T, Li Y, Ma H, et al. Temporal relationship between hyperuricemia and obesity, and its association with future risk of type 2 diabetes. Int J Obes (Lond) 2018 Jul;42(7):1336-1344. [doi: 10.1038/s41366-018-0074-5] [Medline: 29717279]
35. Wu S, Jin C, Li S, Zheng X, Zhang X, Cui L, et al. Aging, arterial stiffness, and blood pressure association in Chinese adults. Hypertension 2019 Apr;73(4):893-899. [doi: 10.1161/HYPERTENSIONAHA.118.12396] [Medline: 30776974]
36. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
37. Fan X, Liu Z, Miyata T, Dasarathy S, Rotroff DM, Wu X, et al. Effect of acid suppressants on the risk of COVID-19: a propensity score-matched study using UK Biobank. Gastroenterology 2021 Jan;160(1):455-458.e5 [FREE Full text] [doi: 10.1053/j.gastro.2020.09.028] [Medline: 32980342]
38. Donato AJ, Machin DR, Lesniewski LA. Mechanisms of dysfunction in the aging vasculature and role in age-related disease. Circ Res 2018 Sep 14;123(7):825-848 [FREE Full text] [doi: 10.1161/CIRCRESAHA.118.312563] [Medline: 30355078]
39. Rothwell PM, Howard SC, Dolan E, O'Brien E, Dobson JE, Dahlöf B, et al. Prognostic significance of visit-to-visit variability, maximum systolic blood pressure, and episodic hypertension. Lancet 2010 Mar 13;375(9718):895-905. [doi: 10.1016/S0140-6736(10)60308-X] [Medline: 20226988]
40. Clark D, Nicholls SJ, St John J, Elshazly MB, Ahmed HM, Khraishah H, et al. Visit-to-visit blood pressure variability, coronary atheroma progression, and clinical outcomes. JAMA Cardiol 2019 May 01;4(5):437-443 [FREE Full text] [doi: 10.1001/jamacardio.2019.0751] [Medline: 30969323]
41. Sander D, Kukla C, Klingelhöfer J, Winbeck K, Conrad B. Relationship between circadian blood pressure patterns and progression of early carotid atherosclerosis: A 3-year follow-up study. Circulation 2000 Sep 26;102(13):1536-1541. [doi: 10.1161/01.cir.102.13.1536] [Medline: 11004145]
42. Palombo C, Kozakova M. Arterial stiffness, atherosclerosis and cardiovascular risk: Pathophysiologic mechanisms and emerging clinical indications. Vascul Pharmacol 2016 Feb;77:1-7. [doi: 10.1016/j.vph.2015.11.083] [Medline: 26643779]
43. Shimbo D, Shea S, McClelland RL, Viera AJ, Mann D, Newman J, et al. Associations of aortic distensibility and arterial elasticity with long-term visit-to-visit blood pressure variability: the Multi-Ethnic Study of Atherosclerosis (MESA). Am J Hypertens 2013 Jul;26(7):896-902 [FREE Full text] [doi: 10.1093/ajh/hpt040] [Medline: 23537891]
44. McEniery CM, Wallace S, Mackenzie IS, McDonnell B, Yasmin, Newby DE, et al. Endothelial function is associated with pulse pressure, pulse wave velocity, and augmentation index in healthy humans. Hypertension 2006 Oct;48(4):602-608. [doi: 10.1161/01.HYP.0000239206.64270.5f] [Medline: 16940223]
45. Diaz KM, Veerabhadrappa P, Kashem MA, Feairheller DL, Sturgeon KM, Williamson ST, et al. Relationship of visit-to-visit and ambulatory blood pressure variability to vascular function in African Americans. Hypertens Res 2012 Jan;35(1):55-61 [FREE Full text] [doi: 10.1038/hr.2011.135] [Medline: 21814215]

46. Hisamatsu T, Miura K, Ohkubo T, Arima H, Fujiyoshi A, Satoh A, SESSA Research Group. Home blood pressure variability and subclinical atherosclerosis in multiple vascular beds: a population-based study. J Hypertens 2018 Nov;36(11):2193-2203. [doi: 10.1097/HJH.0000000000001810] [Medline: 29939942]
47. Wang S, Chennupati R, Kaur H, Iring A, Wettschureck N, Offermanns S. Endothelial cation channel PIEZO1 controls blood pressure by mediating flow-induced ATP release. J Clin Invest 2016 Dec 01;126(12):4527-4536 [FREE Full text] [doi: 10.1172/JCI87343] [Medline: 27797339]
48. Chatzizisis YS, Coskun AU, Jonas M, Edelman ER, Feldman CL, Stone PH. Role of endothelial shear stress in the natural history of coronary atherosclerosis and vascular remodeling: molecular, cellular, and vascular behavior. J Am Coll Cardiol 2007 Jun 26;49(25):2379-2393 [FREE Full text] [doi: 10.1016/j.jacc.2007.02.059] [Medline: 17599600]
49. Chappell DC, Varner SE, Nerem RM, Medford RM, Alexander RW. Oscillatory shear stress stimulates adhesion molecule expression in cultured human endothelium. Circ Res 1998 Mar 23;82(5):532-539. [doi: 10.1161/01.res.82.5.532] [Medline: 9529157]
50. Shintani Y, Kikuya M, Hara A, Ohkubo T, Metoki H, Asayama K, et al. Ambulatory blood pressure, blood pressure variability and the prevalence of carotid artery alteration: the Ohasama study. J Hypertens 2007 Aug;25(8):1704-1710. [doi: 10.1097/HJH.0b013e328172dc2e] [Medline: 17620969]
51. De Keulenaer GW, Chappell DC, Ishizaka N, Nerem RM, Alexander RW, Griendling KK. Oscillatory and steady laminar shear stress differentially affect human endothelial redox state: role of a superoxide-producing NADH oxidase. Circ Res 1998 Jun 01;82(10):1094-1101. [doi: 10.1161/01.res.82.10.1094] [Medline: 9622162]
52. Ziegler T, Bouzourène K, Harrison VJ, Brunner HR, Hayoz D. Influence of oscillatory and unidirectional flow environments on the expression of endothelin and nitric oxide synthase in cultured endothelial cells. Arterioscler Thromb Vasc Biol 1998 May;18(5):686-692. [doi: 10.1161/01.atv.18.5.686] [Medline: 9598825]
53. Tatasciore A, Zimarino M, Renda G, Zurro M, Soccio M, Prontera C, et al. Awake blood pressure variability, inflammatory markers and target organ damage in newly diagnosed hypertension. Hypertens Res 2008 Dec;31(12):2137-2146. [doi: 10.1291/hypres.31.2137] [Medline: 19139603]
54. Boutouyrie P, Chowienczyk P, Humphrey JD, Mitchell GF. Arterial stiffness and cardiovascular risk in hypertension. Circ Res 2021 Apr 02;128(7):864-886. [doi: 10.1161/CIRCRESAHA.121.318061] [Medline: 33793325]
55. Petersen KS, Blanch N, Keogh JB, Clifton PM. Effect of Weight Loss on Pulse Wave Velocity. ATVB 2015 Jan;35(1):243-252. [doi: 10.1161/atvbaha.114.304798] [Medline: 25414255]
56. Zhang Y, Agnoletti D, Safar ME, Blacher J. Effect of antihypertensive agents on blood pressure variability: the Natrilix SR versus candesartan and amlodipine in the reduction of systolic blood pressure in hypertensive patients (X-CELLENT) study. Hypertension 2011 Aug;58(2):155-160. [doi: 10.1161/HYPERTENSIONAHA.111.174383] [Medline: 21747047]
57. Muntner P, Levitan EB, Joyce C, Holt E, Mann D, Oparil S, et al. Association between antihypertensive medication adherence and visit-to-visit variability of blood pressure. J Clin Hypertens (Greenwich) 2013 Feb;15(2):112-117 [FREE Full text] [doi: 10.1111/jch.12037] [Medline: 23339729]
58. Kronish IM, Lynch AI, Oparil S, Whittle J, Davis BR, Simpson LM, et al. The association between antihypertensive medication nonadherence and visit-to-visit variability of blood pressure: findings From the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. Hypertension 2016 Jul;68(1):39-45 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.115.06960] [Medline: 27217410]
59. Parati G, Ochoa JE, Lombardi C, Bilo G. Assessment and management of blood-pressure variability. Nat Rev Cardiol 2013 Mar;10(3):143-155. [doi: 10.1038/nrcardio.2013.1] [Medline: 23399972]
60. D'Elia L, Galletti F, La Fata E, Sabino P, Strazzullo P. Effect of dietary sodium restriction on arterial stiffness: systematic review and meta-analysis of the randomized controlled trials. J Hypertens 2018 Apr;36(4):734-743. [doi: 10.1097/HJH.0000000000001604] [Medline: 29084085]
61. Laurent S, Boutouyrie P, Vascular Mechanism Collaboration. Dose-dependent arterial destiffening and inward remodeling after olmesartan in hypertensives with metabolic syndrome. Hypertension 2014 Oct;64(4):709-716. [doi: 10.1161/HYPERTENSIONAHA.114.03282] [Medline: 25001274]
62. Lopes S, Afreixo V, Teixeira M, Garcia C, Leitão C, Gouveia M, et al. Exercise training reduces arterial stiffness in adults with hypertension: a systematic review and meta-analysis. J Hypertens 2021 Feb 01;39(2):214-222. [doi: 10.1097/HJH.0000000000002619] [Medline: 32833924]

Abbreviations

ΔTba: time interval between the brachium and ankle
baPWV: brachial-ankle pulse wave velocity
BPV: blood pressure variability
CV: coefficient of variation
CVD: cardiovascular disease
DASH: Dietary Approaches to Stop Hypertension
DBP: diastolic pressure

EcNOS: endothelial constitutive nitric oxide synthase
eNOS: endothelial-type nitric oxide synthase
FBG: fasting blood glucose
FMD: flow-mediated vasodilation
HbA1c: glycated hemoglobin
HDL-C: high-density lipoprotein cholesterol
hs-CRP: high-sensitivity C-reactive protein
JNC-7: Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure
LDL-C: low-density lipoprotein cholesterol
NADH: nicotinamide adenine dinucleotide-reduced
NMD: non–endothelium-dependent nitroglycerin-mediated vasodilation
NO: nitric oxide
PWV: pulse wave velocity
SBP: systolic blood pressure
SHATS: systemic hemodynamic atherothrombotic syndrome
SS: shear stress
TC: total cholesterol

Edited by A Mavragani, T Sanchez; submitted 24.12.22; peer-reviewed by A Higaki, N Pahlavani; comments to author 12.04.23; revised version received 20.04.23; accepted 24.05.23; published 04.07.23.

Please cite as:
Wu Z, Zhang H, Wang Y, Li Z, Li X, Tao L, Guo X
Temporal and Bidirectional Association Between Blood Pressure Variability and Arterial Stiffness: Cross-Lagged Cohort Study
JMIR Public Health Surveill 2023;9:e45324
URL: https://publichealth.jmir.org/2023/1/e45324
doi:10.2196/45324
PMID:37402142

Original Paper

Association of Socioeconomic Status With Long-Term Outcome in Survivors After Out-of-Hospital Cardiac Arrest: Nationwide Population-Based Longitudinal Study

Kyung Hun Yoo[1*], MD; Yongil Cho[1,2*], MD, PhD; Jaehoon Oh[1,2], MD, PhD; Juncheol Lee[1,2], MD, PhD; Byuk Sung Ko[1,2], MD, PhD; Hyunggoo Kang[1,2], MD, PhD; Tae Ho Lim[1,2], MD, PhD; Sang Hwan Lee[1], MD

[1]Department of Emergency Medicine, Hanyang University Hospital, Seoul, Republic of Korea
[2]Department of Emergency Medicine, College of Medicine, Hanyang University, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Jaehoon Oh, MD, PhD
Department of Emergency Medicine
College of Medicine
Hanyang University
Seoul, 04763
Republic of Korea
Phone: 82 2 2290 9829
Fax: 82 2 2290 9280
Email: ojjai@hanmail.net

Abstract

Background: Out-of-hospital cardiac arrest (OHCA) is a major public health problem and a leading cause of death worldwide. Previous studies have focused on improving the survival of people who have had OHCA by analyzing short-term survival outcomes, such as the return of spontaneous circulation, 30-day survival, and survival to discharge. Research has been conducted on prehospital prognostic factors to improve the survival of patients with OHCA, among which the association between socioeconomic status (SES) and survival has been reported. SES could affect bystander cardiopulmonary resuscitation rates and whether OHCA is witnessed, and low cardiopulmonary resuscitation education rates are associated with low SES. It has been reported that areas with high SES have shorter hospital transfer times and more public defibrillators per person. Previous studies have shown the impact of SES disparities on the short-term survival of patients with OHCA. However, understanding the impact of SES on the long-term prognosis of OHCA survivors remains limited. As long-term outcomes are more indicative of a patient's ongoing health care needs and the burden on public health than short-term outcomes, understanding the long-term prognosis of OHCA survivors is important.

Objective: This study aimed to identify whether SES influenced the long-term outcomes of OHCA.

Methods: Using health claims data obtained from the National Health Insurance (NHI) service in Korea, we included OHCA survivors who were hospitalized between January 2005 and December 2015. The patients were divided into 2 groups: NHI and Medical Aid (MA) groups, with the MA group defined as having a low SES. Cumulative mortality was estimated using the Kaplan-Meier method, and a Cox proportional hazards model was used to evaluate the impact of SES on long-term mortality. A subgroup analysis was performed based on whether cardiac procedures were performed.

Results: We followed 4873 OHCA survivors for up to 14 years (median of 3.3 years). The Kaplan-Meier survival curve showed that the MA group had a significantly decreased long-term survival rate compared to the NHI group. With an adjusted hazard ratio (aHR) of 1.52 (95% CI 1.35-1.72), low SES was associated with increased long-term mortality. The overall mortality rate of the patients who underwent cardiac procedures in the MA group was significantly higher than that of the NHI group (aHR 1.72, 95% CI 1.05-2.82). The overall mortality rate of patients without cardiac procedures was also increased in the MA group compared to the NHI group (aHR 1.39, 95% CI 1.23-1.58).

Conclusions: OHCA survivors with low SES had an increased risk of poor long-term outcomes compared with those with higher SES. OHCA survivors with low SES who have undergone cardiac procedures need considerable care for long-term survival.

(JMIR Public Health Surveill 2023;9:e47156) doi:10.2196/47156

KEYWORDS

out-of-hospital cardiac arrest; OHCA; socioeconomic status; SES; long-term outcome; survivor; public health; cardiac arrest; socioeconomic disparities; hospital discharge; clinical outcomes

Introduction

Out-of-hospital cardiac arrest (OHCA) is a leading cause of death worldwide [1,2]. With an annual incidence among adults of 55 per 100,000 people worldwide, OHCA is a major global health concern [3,4]. Despite efforts to resuscitate people who have had cardiac arrest in communities and medical facilities, the poor prognosis for OHCA survival imposes an enormous burden on public health [2,5]. Even after being discharged from the hospital, cardiac arrest survivors face significant morbidity and mortality [6].

To achieve the best possible outcomes for adults who have had OHCA, a multifaceted approach is required. This includes early detection of cardiac arrest, activation of an emergency response, prompt administration of high-quality cardiopulmonary resuscitation (CPR), early defibrillation for shockable rhythms, and post–return of spontaneous circulation (ROSC) care [7]. However, disparities in geography, ethnicity, and socioeconomic status (SES) may impact the provision of these processes [8]. The racial and socioeconomic composition of neighborhoods can affect bystander CPR rates, and lower CPR education rates may be associated with low SES [9-11]. Patients with OHCA residing in areas with a high SES have shorter hospital transfer times and higher availability of public access defibrillators per person [12,13]. Previous studies have shown the potential impact of SES disparities on the short-term survival of patients with OHCA. A recent study reported that SES may also influence the long-term outcomes of OHCA survivors, emphasizing the need for further investigation [14]. In addition, the Institute of Medicine in the United States called for more investigation into socioeconomic differences in the outcomes of patients with OHCA [15]. However, most studies have only evaluated clinical outcomes during a short-term follow-up, such as survival to hospital discharge. Therefore, understanding the impact of SES on the long-term prognosis of OHCA survivors remains limited. Moreover, although SES is a strong predictor of cardiovascular disease prognosis, the association between SES and long-term cardiac-associated outcomes of OHCA survivors has not been investigated [16,17].

Since 1990, the Utstein Resuscitation Registry, an international standard guideline for data collection and analysis aimed at improving the outcomes of patients with OHCA, has primarily focused on short-term outcome measures: ROSC, 30-day survival, and survival to discharge [18]. However, long-term outcomes are more indicative of a patient’s ongoing health care needs and the burden on public health than short-term outcomes. With the emphasis placed on long-term prognosis, investigations are being conducted among OHCA survivors who have survived 30 days or more after OHCA, following the completion of their prehospital experience [19-21].

Therefore, this nationwide study aimed to explore the potential effect of SES on long-term outcomes among patients who survived 30 days or more after OHCA. In addition, the association between SES and long-term outcomes among OHCA survivors with and without cardiac procedures was evaluated.

Methods

Data Sources and Setting

We conducted a population-based study using health claims data obtained from the National Health Insurance (NHI) service. The NHI service database includes inpatient and outpatient medical information, such as patient demographics, diagnoses, procedures, drug prescriptions, date of death, and type of patient insurance (ie, NHI and Medical Aid [MA]). The diagnoses were classified based on International Classification of Diseases, 10th Revision (ICD-10) codes. All patients who visit the emergency department (ED) in South Korea for emergency symptoms, such as cardiac arrest, are charged an ED management fee. The NHI service covers a portion of the ED management fee for patients with health insurance and pays the entire fee for those with MA.

Study Design and Population

This was a retrospective longitudinal cohort study to analyze long-term mortality. Patients who had a primary diagnosis of cardiac arrest (ICD-10 code I46.x) and were hospitalized on the day of the initial diagnosis between January 2005 and December 2015 were included. To exclude patients with in-hospital cardiac arrest, those without a primary diagnosis code for cardiac arrest or a code for ED management fees were excluded. To validate the definition of OHCA, we investigated the medical records of 252 patients who visited a tertiary hospital and had a primary diagnosis code for cardiac arrest (I46.x) and a code for ED management fees. This OHCA definition had a 92.1% (232/252) positive predictive value [22]. This study included only adult patients aged 18 years or older. Patients with OHCA who survived for 30 days or more were defined as cardiac arrest survivors, and those who survived for less than 30 days were excluded. In addition, we excluded patients with cardiac arrest resulting from nonmedical causes or trauma (ICD-10 S and T codes).

Due to the absence of medical records in NHI service data, we were unable to determine whether the cause of cardiac arrest was cardiac in origin. However, OHCA survivors who had undergone cardiac procedures could be assumed to have cardiac arrest due to a cardiac origin. Therefore, we examined whether OHCA survivors had undergone cardiac procedures such as percutaneous coronary intervention (PCI), pacemaker placement, implantable cardioverter defibrillator placement, cardiac resynchronization therapy with pacemakers, cardiac resynchronization therapy with defibrillators, and coronary artery bypass grafting during hospitalization. According to this definition, patients with OHCA were divided into cardiac procedure and no cardiac procedure groups.

Health Insurance Systems in South Korea

In South Korea, the NHI system was introduced in 1977, and by July 1989, the entire population was covered [23]. In the past, NHI in South Korea had multiple insurance societies that covered employees and self-employed individuals separately. In 2000, however, all insurance societies were merged into a single-payer system. The NHI program covers almost 98% of the total population, which in 2014 numbered approximately 50 million people [24]. This system has been maintained until the present. The NHI service is divided into the NHI program and the MA program. The MA program is a public assistance program for low-income people who are recipients of the National Basic Livelihood Security System as part of social welfare programs, which is comparable to the Medicaid program in the United States [25]. Beneficiaries of the MA program are divided into 2 categories, types 1 and 2, based on their inability to work (including those younger than 18 years or older than 65 years and those who are disabled) and their ability to work, respectively [26]. The NHI program is divided into employed and self-employed insured groups. The insurance premium for employed and insured individuals is determined according to income and is paid by the employer. The insurance premium for self-employed individuals is based on household income, property, income, vehicles owned, age, and sex [27]. In 2020, the MA program covered 2.9% of the population, while the NHI program covered 97.1% [28]. Since the NHI service conducts compulsory collection from insured individuals in accordance with the law, the citizens of South Korea are obligated to pay insurance premiums [29].

The NHI service data include patient demographics, general specifications (eg, department, date of visit, and state), in-hospital treatment (eg, medical expenses, prescription fees, examination fees, procedure codes, and operation codes), out-of-hospital prescriptions, diagnoses, death records, and socioeconomic variables such as income decile [27,28]. Cosmetic surgery and unproven therapies are not covered by insurance [29]. The greatest value of these data is that they encompass practically the entire population, making them the closest to real-world data, which are referred to as big data [28]. These government-run national health care claims data are available to researchers for public research purposes [30]. They enable researchers to investigate all prescriptions, procedures, and operations performed by domestic medical institutions [31,32].

SES Indicators

Individual-level indicators of SES are generally defined by income, education, and occupation [33]. The proportion of individual-level indicators differs significantly between MA and NHI beneficiaries: ordinary income level (low income: 95% vs 34%), educational level (≤elementary: 61% vs 40%), and proportion of economic activity (no economic status: 82% vs 57%) [34]. A recent study used the measurement of individual-level SES by Medicaid eligibility, which is comparable to the MA program, as a reliable indicator [35]. In addition, several previous studies used insurance type as an individual-level SES indicator [36,37]. Comparing the insurance types across NHI and MA groups can provide insights into the most vulnerable populations. Correspondingly, we categorized the patients into 2 groups: NHI and MA groups, with the MA group defined as having a low SES.

Outcomes

The follow-up data obtained for up to 14 years (until 2018) were analyzed, and the primary outcome was long-term cumulative mortality according to SES. The subgroups were divided into OHCA with a cardiac procedure and OHCA without cardiac procedure, and the long-term cumulative mortality for each group was analyzed. Follow-up was initiated on the index date and continued until the date of death or December 31, 2018 (the last day of the study), whichever occurred first. In addition, 1-year, 3-year, 5-year, 10-year, and overall mortality were analyzed for each group. In addition, we performed an analysis by dividing the NHI group into 4 quartiles based on the insurance premium level.

Ethics Approval

This study was approved by the Institutional Review Board of Hanyang University Hospital (HYUH-2023-01-021) and the Health Insurance Review and Assessment Service (NHIS-2023-1-122). The need for written informed consent was waived because data analyses were performed retrospectively using anonymized data derived from the NHI service. The study was conducted in accordance with the Helsinki Declaration.

Statistical Analysis

Continuous variables are presented as the mean and SD for variables with a normal distribution and median and IQR for variables with a nonnormal distribution. Normality was evaluated using the Shapiro-Wilk test. Categorical variables are presented as frequencies and percentages. The categorical variables of 2 groups were compared using the Fisher exact test, and continuous variables were compared using the 2-tailed *t* test or Mann-Whitney *U* test as appropriate. Cumulative mortality was estimated by the Kaplan-Meier method. A Cox proportional hazards model was used to identify the impact of SES on long-term mortality, and the results are presented as hazard ratios and 95% CIs. The covariables included age, sex, and Charlson Comorbidity Index (CCI) score. The CCI score was calculated at the time of diagnosis within 1 year before the index date using the Quan algorithm [38]. By correcting the covariates, the adjusted hazard ratio (aHR) was obtained. Statistical significance was determined using 2-sided tests with a *P* value <.05. All statistical analyses were conducted using R software (version 4.2.0; R Foundation for Statistical Computing) and SAS software (version 9.4; SAS Institute Inc).

Results

Baseline Characteristics of the Study Population

We enrolled 4937 patients with OHCA aged 18 years or older who had survived for 30 days or more. Finally, a total of 4873 patients were selected as the study population after excluding 4 patients whose insurance type information was missing and

60 patients whose insurance eligibility and date of death information was missing (NHI: 57 of 4537 patients, MA: 3 of 396 patients). The median follow-up duration was 1214 (IQR 140-2112) days, and the maximum follow-up duration was 5059 days. The number of patients with OHCA who underwent cardiac procedures during hospitalization was 1121 (23%) (Figure 1). The cardiac procedure group's median follow-up duration was 1768 (IQR 1.261-2.458) days, and the maximum follow-up duration was 4935 days. The no cardiac procedure group's median follow-up duration was 716.5 (IQR 96.0-1926.8) days, and the maximum follow-up duration was 5059 days. The baseline characteristics of the study population are summarized in Tables 1 and 2.

Figure 1. Flow diagram of the study population selection process. OHCA: out-of-hospital cardiac arrest; NHI: National Health Insurance; MA: Medical Aid.

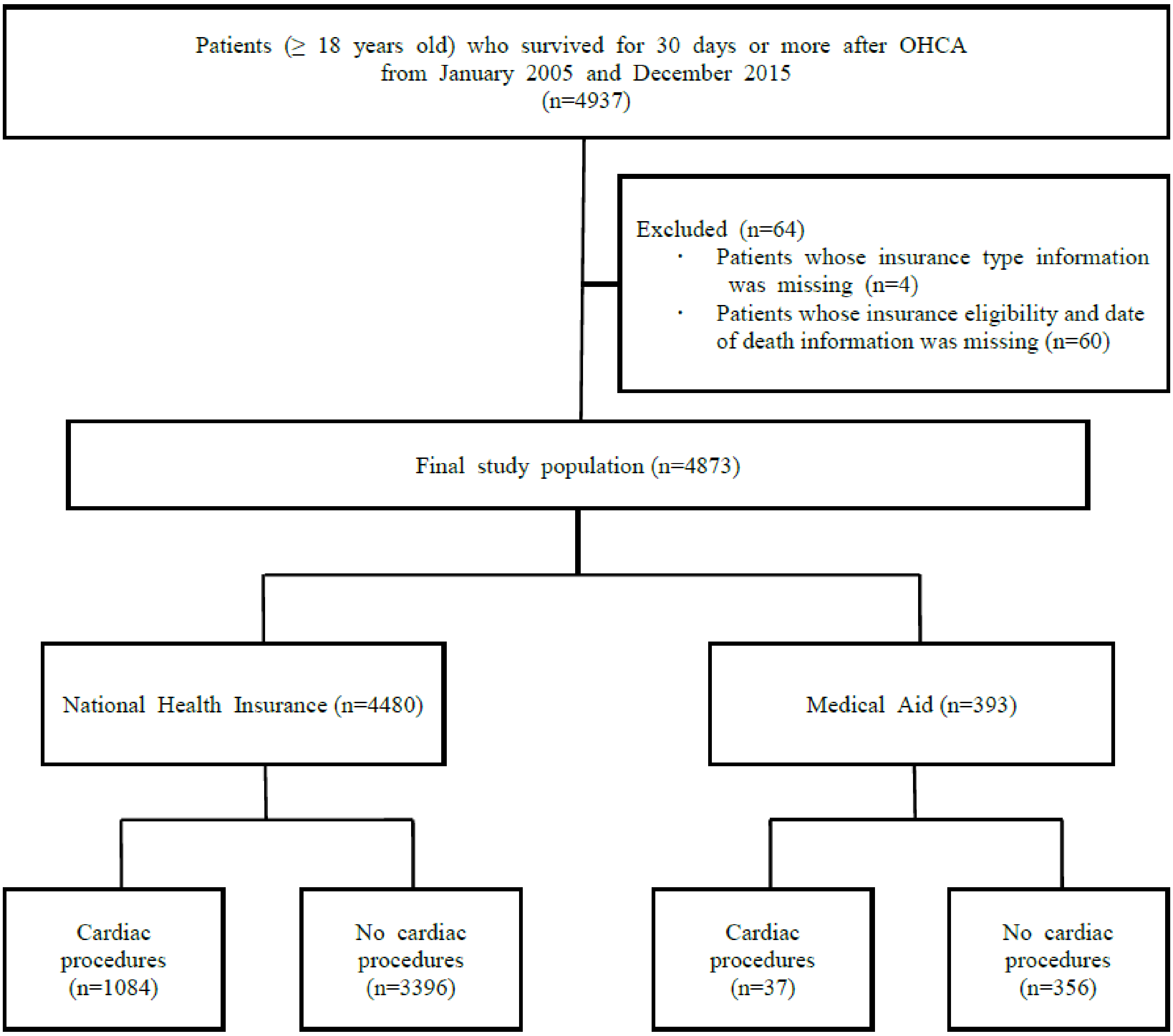


Table 1. Baseline characteristics of the study population.

Characteristics	NHI[a] group (n=4480)	MA[b] group (n=393)
Age (years), median (IQR)	59 (48-70)	63 (52-74)
Age (years), n (%)		
18-39	529 (11.8)	25 (6)
40-49	701 (15.6)	58 (15)
50-59	1130 (25.2)	89 (23)
60-69	951 (21.2)	83 (21)
70-79	802 (17.9)	84 (21)
≥80	367 (8.2)	54 (14)
Sex, n (%)		
Male	3190 (71.2)	211 (53.7)
Female	1290 (28.8)	182 (46.3)
CCI[c] score, n (%)		
0	753 (16.8)	37 (9)
1	1016 (22.7)	49 (12)
2	913 (20.4)	70 (18)
3	710 (15.8)	65 (16)
≥4	1088 (24.3)	172 (43)

[a] NHI: National Health Insurance.

[b] MA: Medical Aid.

[c] CCI: Charlson Comorbidity Index.

Table 2. Baseline characteristics of the subgroup populations.

Characteristics	Cardiac procedures (N=1129)		No cardiac procedures (N=3804)	
	NHI[a] group (n=1084)	MA[b] group (n=37)	NHI[a] group (n=3396)	MA[b] group (n=356)
Age (years), median (IQR)	54 (45-64)	60 (50-72)	60 (50-72)	63 (52-74)
Age (years), n (%)				
18-39	173 (16)	5 (14)	356 (10.5)	20 (5.6)
40-49	215 (19.8)	4 (11)	486 (14.3)	54 (15.2)
50-59	320 (29.5)	9 (24)	810 (23.9)	80 (22.5)
60-69	230 (21.2)	9 (24)	721 (21.2)	74 (20.8)
70-79	124 (11.4)	5 (14)	678 (20)	79 (22.2)
≥80	22 (2)	5 (14)	345 (10.2)	49 (13.8)
Sex, n (%)				
Male	893 (82.4)	19 (51)	2297 (67.6)	192 (53.9)
Female	191 (17.6)	18 (49)	1099 (32.4)	164 (46.1)
CCI[c] score, n (%)				
0	198 (18.3)	2 (5)	555 (16.3)	35 (9.8)
1	293 (27)	5 (14)	723 (21.3)	44 (12.4)
2	230 (21.2)	7 (19)	683 (20.1)	63 (17.7)
3	175 (16.1)	4 (11)	535 (15.8)	61 (17.1)
≥4	188 (17.3)	19 (51)	900 (26.5)	153 (43)

[a] NHI: National Health Insurance.

[b] MA: Medical Aid.

[c] CCI: Charlson Comorbidity Index.

Clinical Outcomes

The Kaplan-Meier curve showed that the MA group had significantly decreased long-term survival compared with the NHI group (Figure 2). The results of Cox regression analysis (Table 3, all patients) showed that a low SES was associated with increased long-term mortality (aHR 1.52, 95% CI 1.35-1.72; *P*<.001).

In subgroups (Table 3, cardiac procedures and no cardiac procedures), Cox regression analysis showed that a low SES was associated with increased long-term mortality in both the cardiac procedure (aHR 1.72, 95% CI 1.05-2.82; *P*=.03) and no cardiac procedure groups (aHR 1.39, 95% CI 1.23-1.58; *P*<.001). The Kaplan-Meier curves for patients with OHCA with cardiac procedures and without cardiac procedures are presented in Multimedia Appendices 1 and 2.

The 1-year, 3-year, 5-year, and 10-year mortality rates in the MA group compared with the NHI group are presented in Table 4. The 1-year, 3-year, and 5-year mortality rates of the OHCA with cardiac procedure group were not statistically significant. However, the aHR of 10-year mortality rate in the MA group was significantly higher than that in the NHI group (aHR 1.72, 95% CI 1.05-2.82). Among all patients and in the no cardiac procedure group, the aHR of 1-year mortality rate was the highest (aHR 1.64, 95% CI 1.42-1.90; aHR 1.54, 95% CI 1.34-1.79). The forest plot based on the results of the 1-year, 3-year, 5-year, and 10-year mortality rates of the MA group compared with those of the NHI group is illustrated in Figure 3.

The baseline characteristics of the study population, divided into quartiles based on the level of insurance premiums for NHI beneficiaries, are summarized in Multimedia Appendices 3-5. Cox regression analysis revealed that the lowest quartile group and the MA group had significantly higher mortality rates compared to the highest quartile group (aHR 1.12, 95% CI 1-1.26; aHR 1.60, 95% CI 1.40-1.83, respectively). For the other quartile groups, no statistically significant differences were observed. In particular, patients who underwent cardiac procedures showed a significantly higher risk of long-term mortality (lowest quartile of NHI: aHR 1.54, 95% CI 1.08-2.18; MA: aHR 2.07, 95% CI 1.23-3.51). The MA group showed higher long-term mortality rates regardless of whether the patients had undergone cardiac procedures or not (Multimedia Appendix 6).

Figure 2. Kaplan-Meier survival curves for long-term mortality of out-of-hospital cardiac arrest survivors according to the type of patient insurance. NHI: National Health Insurance; MA: Medical Aid.

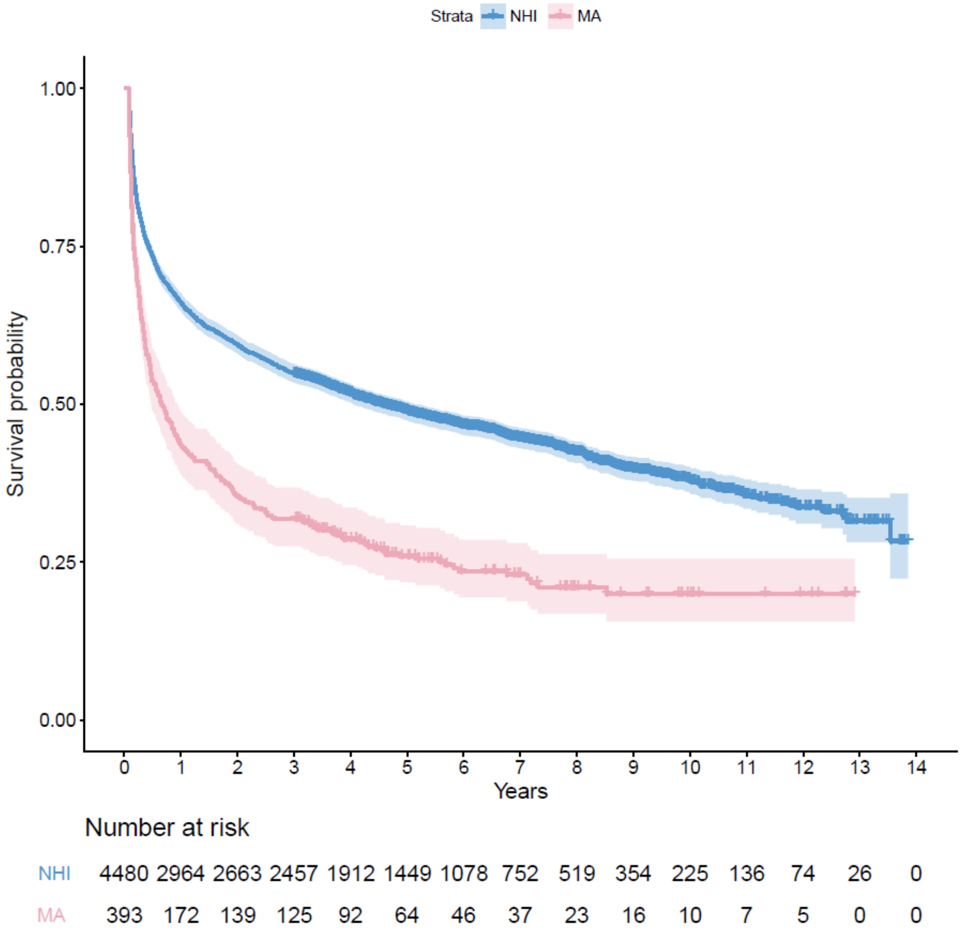


Table 3. Cox regression analysis to identify the impact of low socioeconomic status on long-term mortality among survivors of out-of-hospital cardiac arrest.

	Crude HR[a] (95% CI)	*P* value	Adjusted HR[a] (95% CI	*P* value
All patients	1.87 (1.65-2.11)	<.001	1.52 (1.35-1.72)	<.001
Cardiac procedures	2.67 (1.65-4.32)	<.001	1.72 (1.05-2.82)	.03
No cardiac procedures	1.57 (1.38-1.78)	<.001	1.39 (1.23-1.58)	<.001

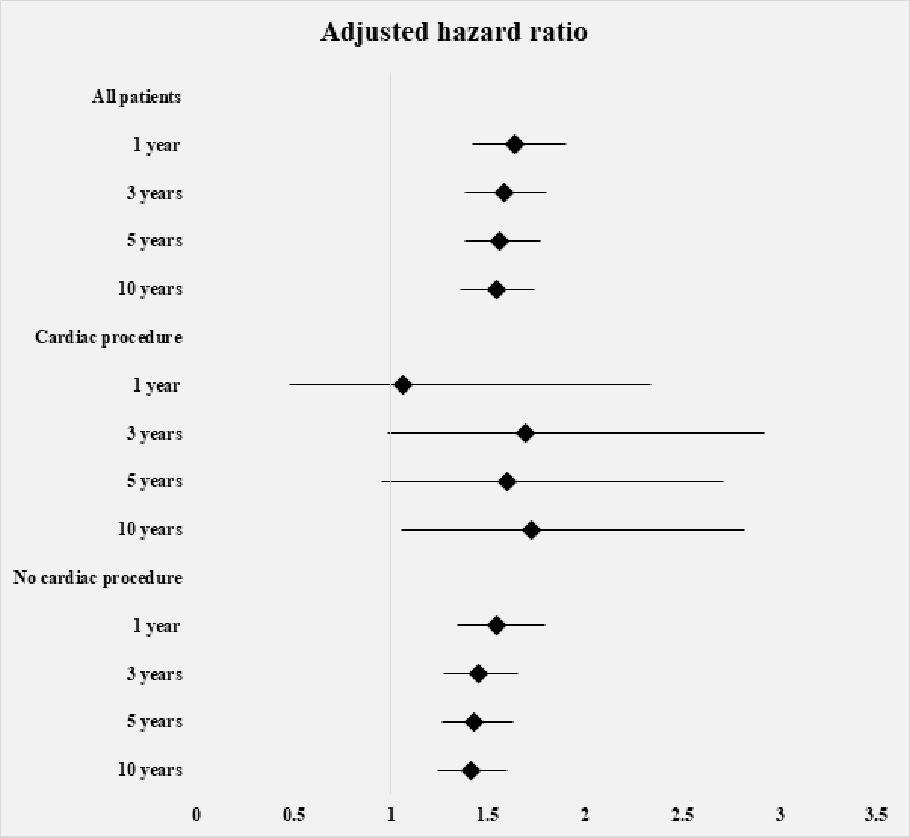
[a] HR: hazard ratio.

Table 4. Cox regression analysis to identify the impact of low socioeconomic status on 1-year, 3-year, 5-year, and 10-year mortality among survivors after out-of-hospital cardiac arrest.

	Crude HR[a] (95% CI)	Adjusted HR[a] (95% CI
All patients		
1-year mortality	1.97 (1.71-2.27)	1.64 (1.42-1.90)
3-year mortality	1.92 (1.69-2.18)	1.58 (1.38-1.80)
5-year mortality	1.91 (1.69-2.16)	1.56 (1.38-1.77)
10-year mortality	1.88 (1.67-2.12)	1.54 (1.36-1.74)
Cardiac procedures		
1-year mortality	1.91 (0.89-4.09)	1.06 (0.48-2.34)
3-year mortality	2.82 (1.66-4.78)	1.69 (0.98-2.92)
5-year mortality	2.57 (1.55-4.28)	1.60 (0.95-2.71)
10-year mortality	2.67 (1.65-4.32)	1.72 (1.05-2.82)
No cardiac procedures		
1-year mortality	1.70 (1.47-1.96)	1.54 (1.34-1.79)
3-year mortality	1.61 (1.41-1.84)	1.45 (1.27-1.65)
5-year mortality	1.61 (1.42-1.83)	1.43 (1.26-1.63)
10-year mortality	1.58 (1.39-1.79)	1.41 (1.24-1.60)

[a]HR: hazard ratio.

Figure 3. Forest plot based on the 1-year, 3-year, 5-year, and 10-year mortality results of the Medical Aid group compared with the National Health Insurance group. Adjusted hazard ratios and 95% CIs were analyzed using a Cox proportional hazards model.



Discussion

Principal Findings

This population-based study analyzed the association between long-term mortality and SES among survivors after OHCA by following these patients for up to 14 years. The long-term mortality rate of patients with a low SES was approximately 52% greater than that of patients with a higher SES. The 10-year mortality rate of patients who had undergone a cardiac procedure among OHCA survivors with a low SES was approximately >72% than that of OHCA survivors with a higher SES. With the exception of OHCA survivors in the cardiac procedure group, the aHR of 1-year mortality was highest for patients with a low SES compared to patients with a higher SES.

The majority of studies on OHCA have focused on short-term clinical outcomes, such as the ROSC, 30-day survival, and survival to discharge; however, the long-term clinical outcomes beyond 1 year are poorly understood [21]. In addition to the interest in short-term clinical outcomes, the long-term clinical outcomes and recovery of OHCA survivors have drawn attention [39]. The recovery from cardiac arrest continues long after the patient is discharged from the hospital [40,41]. As the importance of long-term prognosis has become apparent, "recovery" has been added as a new link in the chain of survival in the 2020 American Heart Association guidelines [7].

To the best of our knowledge, this is the first study to investigate the association between SES and the long-term prognoses (greater than 10 years) of OHCA survivors using data from the NHI service. In the past, the focus was on the survival of patients with OHCA, but now both the European Resuscitation Council guidelines and the American Heart Association guidelines also emphasize the importance of survivors' recovery and rehabilitation [42,43]. From a research standpoint, this study can provide evidence for establishing a long-term strategy for OHCA survivors with a low SES.

In public health research, SES is typically measured using individual-level or area-level indicators depending on the study objectives [44,45]. Area-level SES indicators are used when a geographical area's SES is the object of analysis rather than an individual's SES. Area-level indicators are also often used as proxies for missing individual-level indicators in public health research, in which case they pose a risk of ecological fallacy when drawing conclusions about individuals based on area-level measurements [10,46,47]. When area-level indicators are used as proxies for individual-level SES, the SES-mortality association could be underestimated [46,48]. Long-term cumulative mortality according to SES among OHCA survivors was the primary outcome of our study. Moreover, to minimize the influence of prehospital factors associated with SES that may have area-level characteristics, such as witnessed arrest, bystander CPR, defibrillators, and the EMS system, we included patients who survived 30 days or more after OHCA. Therefore, it was considered appropriate in this study to measure SES as an individual-level indicator.

Recent studies have noted the lack of research measuring SES at the individual level [10,46,49]. Income, education, and occupation are the most traditional individual-level indicators of SES and have proven to be highly useful in describing and evaluating health disparities [45,50]. Measuring SES at the individual level usually reveals greater disparities in health outcomes and may increase awareness of the most vulnerable individuals [48,49]. However, when considered in isolation, these disparities only provide a limited view of SES inequalities in health [45]. In our additional analysis, we divided the NHI data by insurance premium level as a measure of income. There were statistically significant differences between patients in the lowest and highest quartiles who had undergone cardiac procedures. However, in the MA group, in which all 3 indicators of SES were vulnerable, significant differences were observed regardless of whether patients underwent cardiac procedures. There were significant differences between NHI beneficiaries and MA beneficiaries in terms of income, education, and economic activity. Thus, one of the strengths of our study is that it presented results of the comparison of SES as an individual-level measurement. We also provided the opportunity to compare results on a national scale by using nationwide population data.

It has been reported that a low SES is associated with poor long-term outcomes in a variety of diseases, such as cardiovascular disease, stroke, and surgery [51-53]. However, the association between long-term prognosis and SES among OHCA survivors remains unclear. Recently, Møller et al [19] reported that patients with a high income had a higher survival rate after OHCA than patients with a low income. The proportion of patients who survived 1 year after OHCA was 96.4% in the high-income group and 84.2% in the low-income group. Of a total of 1785 patients who survived 5 years after OHCA, the survival rate was 87.6% in the high-income group and 64.1% in the low-income group [19]. Our study showed that low SES had an adverse effect on the long-term outcomes of OHCA survivors, which is in line with results from a previous study. It is well known that a high SES is associated with several favorable prognostic factors for OHCA survival. However, a previous study showed that SES is also associated with the long-term outcomes of OHCA survivors, even after adjusting for witnessed arrest, bystander CPR, and initial shockable rhythm, which are well-known prehospital positive prognostic factors [19].

Although the exact etiology explaining why survivors with low SES have poor long-term outcomes after OHCA is not well understood, several hypotheses can be proposed. The usual explanation for health disparities is differences in lifestyle [54]. Individuals with a lower SES are more likely to engage in harmful alcohol use, tobacco use, unhealthy diets, and physical inactivity [55]. Medhekar et al [56] reported a similar pattern among patients with lower SES among survivors after cardiac arrest. Differences in general health due to the accumulation of these various risk factors can affect prognosis. Unfortunately, such data were not available for this study. Previous studies have shown that a low SES is associated with an increased risk of multimorbidity [57]. Likewise, our study showed that the MA group had a higher proportion of patients with multimorbidity (CCI score ≥4, 43.3% vs 24.3%). However, MA itself was independently associated with long-term mortality after adjusting for comorbidities.

It is also possible that patients with low SES receive less adequate hospital care. Post–cardiac arrest care, including critical care interventions, is another important link in the chain of survival, and post arrest interventions such as targeted temperature management and PCI are associated with improved cardiac arrest survival and neurologic recovery [42,58]. Casey and Mumma [59] showed an association between government insurance (odds ratio [OR] 0.56, 95% CI 0.51-0.61) and Medicare insurance (OR 0.44, 95% CI 0.40-0.48) and lower rates of cardiac catheterization, indicating that lower income patients may experience lower rates of PCI. Huebinger et al [60] reported that lower income, education, and occupation were all associated with lower rates of targeted temperature management and that lower income was also associated with lower rates of PCI. Although we did not investigate hospital care in this study, patients with a low SES are likely to have received fewer interventions, as the rate of MA beneficiaries among patients who underwent cardiac procedures was approximately 3.3%, and the rate of MA beneficiaries among those who did not undergo cardiac procedures was approximately 9.4%.

It is well known that individuals with a low SES have poor medication adherence [61]. Among patients with cardiovascular disease, there is a strong association between medication nonadherence and adverse clinical outcomes [62]. Therefore, adherence to medication is particularly important for OHCA survivors who have undergone cardiovascular procedures. Liang et al [63] showed that pharmacists who provided pharmaceutical services via home visits were effective in reducing medication problems among socioeconomically disadvantaged individuals. As such, efforts are needed to improve drug compliance among patients with low SES, which can also improve the long-term prognosis of OHCA survivors. Nielsen et al [64] showed that patients with low income (OR 0.2, 95% CI 0.1-0.5) and patients who lived alone (OR 0.4, 95% CI 0.2-0.9) were less likely to participate in cardiac rehabilitation (CR) programs. Meillier et al [65] suggested that inequalities in the recruitment and participation of socially vulnerable and low-educated patients in CR programs could be reduced through systematic screening and social differentiation.

Recent studies have reported the need for national financial assistance to help cardiovascular disease patients with a low SES participate in CR [66,67]. Likewise, to improve the long-term prognosis of patients with low SES among OHCA survivors, a strategy to increase rehabilitation participation and compliance with medication should be established. In addition, research on the need for national support to aid the recovery of cardiac arrest survivors is necessary. The disparity in SES will become an increasingly important public health issue [68]. The newly added link, recovery, highlights the need for the system of care to support the recovery (both short- and long-term) of cardiac arrest survivors as they return to social function [7,42]. In terms of the system of care to support the successful recovery of OHCA survivors, efforts to narrow the socioeconomic gap are important. Particularly, OHCA survivors who have undergone cardiac procedures may need public attention to improve their long-term survival for more than 10 years.

Our study has several limitations. First, the diagnosis of OHCA was defined using ICD-10 codes, and diagnostic inaccuracies could not be ruled out. Defining patients with OHCA and in-hospital cardiac arrest can be particularly imprecise. We validated the definition of OHCA, but it was based on a small sample size. Second, although this study included patients who survived 30 days or more after OHCA to minimize the influence of factors that could have affected the survival of patients with OHCA, such as initial cardiac arrest rhythm, witnessed status, bystander CPR, and duration of cardiac arrest, the potential impact could not be completely discounted. Third, because our study was based on data from a single nation, it is difficult to generalize the findings to other populations. Fourth, we only analyzed mortality as a long-term prognosis for OHCA survivors. Although mortality is an important issue in patients' long-term outcomes, health-related quality of life (HRQOL) also has a substantial impact on long-term survival [69]. HRQOL is a multifaceted concept that comprises biological function, psychological function, and social function, which are influenced by a variety of factors [70]. Therefore, SES can influence not only the long-term mortality rate but also the HRQOL of patients. Unfortunately, it was challenging to evaluate HRQOL using NHI data. The association between SES and HRQOL among OHCA survivors needs additional investigation.

Conclusions

This nationwide population-based study showed an association between SES and long-term outcomes among survivors after OHCA. Survivors with low SES, especially vulnerable individuals such as those eligible for MA, were at higher risk of unfavorable long-term outcomes after OHCA. Particularly, OHCA survivors with low SES who have undergone cardiac procedures need greater attention for long-term survival.

Acknowledgments

This study used the National Health Insurance Service database (NHIS-2023-1-122). The interpretations and conclusions reported herein do not represent those of the National Health Insurance Service.

This study was supported by the National Research Foundation of Korea grant NRF-2022R1A2C1012627. The National Research Foundation of Korea had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

National Health Insurance Service data are third-party data that are not owned by the authors. The raw data can be accessed with permission from the National Health Insurance Service in Korea.

Authors' Contributions

JO conceived the study and contributed to critically revising the manuscript. KHY contributed to data analysis, data interpretation, and manuscript writing. YC contributed to the study design, data analysis, and assisted in drafting. JL and SHL contributed to data collection and extraction. BSK, HK, and THL were responsible for the manuscript review. All authors reviewed the drafted manuscript for critical content and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Kaplan-Meier survival curves for long-term mortality of out-of-hospital cardiac arrest survivors with cardiac procedures according to the type of patient insurance. NHI: National Health Insurance; MA: Medical Aid.
[PNG File , 43 KB - publichealth_v9i1e47156_app1.png]

Multimedia Appendix 2
Kaplan-Meier survival curves for long-term mortality of out-of-hospital cardiac arrest survivors without cardiac procedures according to the type of patient insurance. NHI: National Health Insurance; MA: Medical Aid.
[PNG File , 51 KB - publichealth_v9i1e47156_app2.png]

Multimedia Appendix 3
Baseline characteristics of the study population categorized into quartiles based on insurance premium level.
[DOCX File , 14 KB - publichealth_v9i1e47156_app3.docx]

Multimedia Appendix 4
Baseline characteristics of the study population that underwent cardiac procedures categorized into quartiles based on insurance premium level.
[DOCX File , 14 KB - publichealth_v9i1e47156_app4.docx]

Multimedia Appendix 5
Baseline characteristics of the study population that did not undergo cardiac procedures categorized into quartiles based on insurance premium level.
[DOCX File , 14 KB - publichealth_v9i1e47156_app5.docx]

Multimedia Appendix 6
Cox regression analysis of the impact of insurance premium level on long-term mortality among survivors after out-of-hospital cardiac arrest.
[DOCX File , 13 KB - publichealth_v9i1e47156_app6.docx]

References

1. Berdowski J, Berg RA, Tijssen JGP, Koster RW. Global incidences of out-of-hospital cardiac arrest and survival rates: systematic review of 67 prospective studies. Resuscitation 2010;81(11):1479-1487. [doi: 10.1016/j.resuscitation.2010.08.006] [Medline: 20828914]
2. Hwang SO, Cha KC, Jung WJ, Roh YI, Kim TY, Chung SP, Steering Committee of the 2020 Korean Guidelines for Cardiopulmonary ResuscitationEmergency Cardiovascular Care. 2020 Korean guidelines for cardiopulmonary resuscitation. part 1. update process and highlights. Clin Exp Emerg Med 2021;8(S):S1-S7 [FREE Full text] [doi: 10.15441/ceem.21.021] [Medline: 34034445]
3. Kim JW, Ha J, Kim T, Yoon H, Hwang SY, Jo IJ, et al. Developing a time-adaptive prediction model for out-of-hospital cardiac arrest: nationwide cohort study in Korea. J Med Internet Res 2021;23(7):e28361 [FREE Full text] [doi: 10.2196/28361] [Medline: 36260382]
4. Yan S, Gan Y, Jiang N, Wang R, Chen Y, Luo Z, et al. The global survival rate among adult out-of-hospital cardiac arrest patients who received cardiopulmonary resuscitation: a systematic review and meta-analysis. Crit Care 2020;24(1):61 [FREE Full text] [doi: 10.1186/s13054-020-2773-2] [Medline: 32087741]

5. Stecker EC, Reinier K, Marijon E, Narayanan K, Teodorescu C, Uy-Evanado A, et al. Public health burden of sudden cardiac death in the United States. Circ Arrhythm Electrophysiol 2014;7(2):212-217 [FREE Full text] [doi: 10.1161/CIRCEP.113.001034] [Medline: 24610738]
6. Shuvy M, Morrison LJ, Koh M, Qiu F, Buick JE, Dorian P, Rescu Epistry Investigators. Long-term clinical outcomes and predictors for survivors of out-of-hospital cardiac arrest. Resuscitation 2017;112:59-64. [doi: 10.1016/j.resuscitation.2016.12.026] [Medline: 28104428]
7. Panchal AR, Bartos JA, Cabañas JG, Donnino MW, Drennan IR, Hirsch KG, Adult BasicAdvanced Life Support Writing Group. Part 3: adult basic and advanced life support: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020;142(16_suppl_2):S366-S468 [FREE Full text] [doi: 10.1161/CIR.0000000000000916] [Medline: 33081529]
8. Berg DD, Bobrow BJ, Berg RA. Key components of a community response to out-of-hospital cardiac arrest. Nat Rev Cardiol 2019;16(7):407-416. [doi: 10.1038/s41569-019-0175-4] [Medline: 30858511]
9. Sasson C, Magid DJ, Chan P, Root ED, McNally BF, Kellermann AL, CARES Surveillance Group. Association of neighborhood characteristics with bystander-initiated CPR. N Engl J Med 2012;367(17):1607-1615 [FREE Full text] [doi: 10.1056/NEJMoa1110700] [Medline: 23094722]
10. Chamberlain RC, Barnetson C, Clegg GR, Halbesma N. Association of measures of socioeconomic position with survival following out-of-hospital cardiac arrest: a systematic review. Resuscitation 2020;157:49-59. [doi: 10.1016/j.resuscitation.2020.09.025] [Medline: 33010372]
11. Huebinger R, Vithalani V, Osborn L, Decker C, Jarvis J, Dickson R, et al. Community disparities in out of hospital cardiac arrest care and outcomes in Texas. Resuscitation 2021;163:101-107. [doi: 10.1016/j.resuscitation.2021.03.021] [Medline: 33798624]
12. Hsia RY, Huang D, Mann NC, Colwell C, Mercer MP, Dai M, et al. A US national study of the association between income and ambulance response time in cardiac arrest. JAMA Netw Open 2018;1(7):e185202 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.5202] [Medline: 30646394]
13. Lee D, Stiepak JK, Pommerenke C, Poloczek S, Grittner U, Prugger C. Public access defibrillators and socioeconomic factors on the small-scale spatial level in Berlin. Dtsch Arztebl Int 2022;119(22):393-399 [FREE Full text] [doi: 10.3238/arztebl.m2022.0180] [Medline: 35477511]
14. Pemberton K, Franklin RC, Bosley E, Watt K. Long-term outcomes of adult out-of-hospital cardiac arrest in Queensland, Australia (2002-2014): incidence and temporal trends. Heart 2021;107(16):1310-1319. [doi: 10.1136/heartjnl-2020-317333] [Medline: 33219108]
15. Institute of Medicine, Board on Health Sciences Policy, Committee on the Treatment of Cardiac Arrest: Current Status and Future Directions. In: Schultz AM, McCoy MA, Graham R, editors. Strategies to Improve Cardiac Arrest Survival: A Time to Act. Washington, DC: National Academies Press; 2015.
16. Rosengren A, Smyth A, Rangarajan S, Ramasundarahettige C, Bangdiwala SI, AlHabib KF, et al. Socioeconomic status and risk of cardiovascular disease in 20 low-income, middle-income, and high-income countries: the prospective urban rural epidemiologic (PURE) study. Lancet Glob Health 2019;7(6):e748-e760 [FREE Full text] [doi: 10.1016/S2214-109X(19)30045-2] [Medline: 31028013]
17. Tousoulis D, Oikonomou E, Vogiatzi G, Vardas P. Cardiovascular disease and socioeconomic status. Eur Heart J 2020;41(34):3213-3214 [FREE Full text] [doi: 10.1093/eurheartj/ehaa405] [Medline: 32383766]
18. Perkins GD, Jacobs IG, Nadkarni VM, Berg RA, Bhanji F, Biarent D, Utstein Collaborators. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update of the utstein resuscitation registry templates for out-of-hospital cardiac arrest: a statement for healthcare professionals from a task force of the international liaison committee on resuscitation (American Heart Association, European Resuscitation Council, Australian and New Zealand Council on Resuscitation, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Council of Southern Africa, Resuscitation Council of Asia); and the American heart association emergency cardiovascular care committee and the council on cardiopulmonary, critical care, perioperative and resuscitation. Circulation 2015;132(13):1286-1300 [FREE Full text] [doi: 10.1161/CIR.0000000000000144] [Medline: 25391522]
19. Møller S, Wissenberg M, Søndergaard K, Kragholm K, Folke F, Hansen CM, et al. Long-term outcomes after out-of-hospital cardiac arrest in relation to socioeconomic status. Resuscitation 2021;167:336-344 [FREE Full text] [doi: 10.1016/j.resuscitation.2021.07.015] [Medline: 34302925]
20. Ho AFW, Lim MJR, Earnest A, Blewer A, Graves N, Yeo JW, Singapore PAROS Investigators. Long term survival and disease burden from out-of-hospital cardiac arrest in Singapore: a population-based cohort study. Lancet Reg Health West Pac 2023;32:100672 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100672] [Medline: 36785853]
21. Chin YH, Yaow CYL, Teoh SE, Foo MZQ, Luo N, Graves N, et al. Long-term outcomes after out-of-hospital cardiac arrest: a systematic review and meta-analysis. Resuscitation 2022;171:15-29. [doi: 10.1016/j.resuscitation.2021.12.026] [Medline: 34971720]
22. Cho Y, Oh J, Shin JH, Kim BS, Park JK, Lee JH, et al. Long-term prognosis and causes of death among survivors after out-of-hospital cardiac arrest: a population-based longitudinal study. Resuscitation 2022;173:31-38. [doi: 10.1016/j.resuscitation.2022.02.005] [Medline: 35157998]

23. Jeon B, Kwon S. Effect of private health insurance on health care utilization in a universal public insurance system: a case of South Korea. Health Policy 2013;113(1-2):69-76 [FREE Full text] [doi: 10.1016/j.healthpol.2013.05.007] [Medline: 23786992]
24. Kwon S. Thirty years of national health insurance in South Korea: lessons for achieving universal health care coverage. Health Policy Plan 2009;24(1):63-71 [FREE Full text] [doi: 10.1093/heapol/czn037] [Medline: 19004861]
25. Kong NY, Kim DH. Factors influencing health care use by health insurance subscribers and medical aid beneficiaries: a study based on data from the Korea welfare panel study database. BMC Public Health 2020;20(1):1133 [FREE Full text] [doi: 10.1186/s12889-020-09073-x] [Medline: 32689960]
26. Lee DW, Jang J, Choi DW, Jang SI, Park EC. The effect of shifting medical coverage from national health insurance to medical aid type I and type II on health care utilization and out-of-pocket spending in South Korea. BMC Health Serv Res 2020;20(1):979 [FREE Full text] [doi: 10.1186/s12913-020-05778-2] [Medline: 33109176]
27. Kim HK, Song SO, Noh J, Jeong IK, Lee BW. Data configuration and publication trends for the Korean national health insurance and health insurance review & assessment database. Diabetes Metab J 2020;44(5):671-678 [FREE Full text] [doi: 10.4093/dmj.2020.0207] [Medline: 33115211]
28. Kyoung DS, Kim HS. Understanding and Utilizing Claim Data from the Korean National Health Insurance Service (NHIS) and Health Insurance Review and Assessment (HIRA) Database for Research. J Lipid Atheroscler 2022 May;11(2):103-110 [FREE Full text] [doi: 10.12997/jla.2022.11.2.103] [Medline: 35656154]
29. Jung CH, Seo GH, Suh S, Bae JC, Kim MK, Hwang Y, et al. The population-based risk of need for coronary revascularization according to the presence of type 2 diabetes mellitus and history of coronary heart disease in the Korean population. PLoS One 2015;10(6):e0128627 [FREE Full text] [doi: 10.1371/journal.pone.0128627] [Medline: 26053222]
30. Kim S, Kim MS, You SH, Jung SY. Conducting and reporting a clinical research using Korean healthcare claims database. Korean J Fam Med 2020;41(3):146-152 [FREE Full text] [doi: 10.4082/kjfm.20.0062] [Medline: 32456382]
31. Ahn HY, Chae JE, Moon H, Noh J, Park YJ, Kim SG. Trends in the diagnosis and treatment of patients with medullary thyroid carcinoma in Korea. Endocrinol Metab (Seoul) 2020;35(4):811-819 [FREE Full text] [doi: 10.3803/EnM.2020.709] [Medline: 33212545]
32. Seong SC, Kim YY, Khang YH, Heon Park JH, Kang HJ, Lee H, et al. Data resource profile: the national health information database of the national health insurance service in South Korea. Int J Epidemiol 2017;46(3):799-800 [FREE Full text] [doi: 10.1093/ije/dyw253] [Medline: 27794523]
33. Marshall IJ, Wang Y, Crichton S, McKevitt C, Rudd AG, Wolfe CDA. The effects of socioeconomic status on stroke risk and outcomes. Lancet Neurol 2015;14(12):1206-1218. [doi: 10.1016/S1474-4422(15)00200-8] [Medline: 26581971]
34. Kim JH, Lee KS, Yoo KB, Park EC. The differences in health care utilization between medical aid and health insurance: a longitudinal study using propensity score matching. PLoS One 2015;10(3):e0119939 [FREE Full text] [doi: 10.1371/journal.pone.0119939] [Medline: 25816234]
35. Cheng E, Soulos PR, Irwin ML, Feliciano EMC, Presley CJ, Fuchs CS, et al. Neighborhood and individual socioeconomic disadvantage and survival among patients with nonmetastatic common cancers. JAMA Netw Open 2021;4(12):e2139593 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.39593] [Medline: 34919133]
36. Cho Y, Lim TH, Kang H, Lee Y, Lee H, Kim H. Socioeconomic status and depression as combined risk factors for acute myocardial infarction and stroke: a population-based study of 2.7 million Korean adults. J Psychosom Res 2019;121:14-23 [FREE Full text] [doi: 10.1016/j.jpsychores.2019.01.016] [Medline: 30712815]
37. Kim HL, Lee JY, Lim WH, Seo JB, Kim SH, Zo JH, et al. Relationship of socioeconomic status to arterial stiffness: comparison between medical aid beneficiaries and national health insurance beneficiaries. Am J Hypertens 2020;33(8):718-725 [FREE Full text] [doi: 10.1093/ajh/hpaa075] [Medline: 32391565]
38. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JE, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
39. Amacher SA, Bohren C, Blatter R, Becker C, Beck K, Mueller J, et al. Long-term survival after out-of-hospital cardiac arrest: a systematic review and meta-analysis. JAMA Cardiol 2022;7(6):633-643 [FREE Full text] [doi: 10.1001/jamacardio.2022.0795] [Medline: 35507352]
40. Schnaubelt S, Mayr FB, Losert H, Domanovits H, Sulzgruber P, Holzer M, et al. Very long-term survivors of in-hospital and out-of-hospital cardiac arrest show considerable impairment of daily life. Resuscitation 2022;173:192-200 [FREE Full text] [doi: 10.1016/j.resuscitation.2022.01.023] [Medline: 35131405]
41. Kim YM, Jeung KW, Kim WY, Park YS, Oh JS, You YH, Steering Committee of 2020 Korean Guidelines for Cardiopulmonary ResuscitationEmergency Cardiovascular Care. 2020 Korean guidelines for cardiopulmonary resuscitation. Part 5. Post-cardiac arrest care. Clin Exp Emerg Med 2021;8(S):S41-S64 [FREE Full text] [doi: 10.15441/ceem.21.025] [Medline: 34034449]
42. Berg KM, Cheng A, Panchal AR, Topjian AA, Aziz K, Bhanji F, Adult Basic Advanced Life Support, Pediatric Basic Advanced Life Support, Neonatal Life Support, Resuscitation Education Science Writing Groups. Part 7: systems of care: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020;142(16_suppl_2):S580-S604 [FREE Full text] [doi: 10.1161/CIR.0000000000000899] [Medline: 33081524]

43. Nolan JP, Sandroni C, Böttiger BW, Cariou A, Cronberg T, Friberg H, et al. European Resuscitation Council and European Society of Intensive Care Medicine guidelines 2021: post-resuscitation care. Intensive Care Med 2021;47(4):369-421 [FREE Full text] [doi: 10.1007/s00134-021-06368-4] [Medline: 33765189]
44. Golembiewski E, Allen KS, Blackmon AM, Hinrichs RJ, Vest JR. Combining nonclinical determinants of health and clinical data for research and evaluation: rapid review. JMIR Public Health Surveill 2019;5(4):e12846 [FREE Full text] [doi: 10.2196/12846] [Medline: 31593550]
45. Galobardes B, Lynch J, Smith GD. Measuring socioeconomic position in health research. Br Med Bull 2007;81-82:21-37 [FREE Full text] [doi: 10.1093/bmb/ldm001] [Medline: 17284541]
46. Moss JL, Johnson NJ, Yu M, Altekruse SF, Cronin KA. Comparisons of individual- and area-level socioeconomic status as proxies for individual-level measures: evidence from the mortality disparities in American communities study. Popul Health Metr 2021;19(1):1 [FREE Full text] [doi: 10.1186/s12963-020-00244-x] [Medline: 33413469]
47. Geronimus AT, Bound J, Neidert LJ. On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. J Am Stat Assoc 1996;91(434):529-537 [FREE Full text] [doi: 10.3386/t0189]
48. Steenland K, Henley J, Calle E, Thun M. Individual- and area-level socioeconomic status variables as predictors of mortality in a cohort of 179,383 persons. Am J Epidemiol 2004;159(11):1047-1056 [FREE Full text] [doi: 10.1093/aje/kwh129] [Medline: 15155289]
49. van Nieuwenhuizen BP, Oving I, Kunst AE, Daams J, Blom MT, Tan HL, et al. Socio-economic differences in incidence, bystander cardiopulmonary resuscitation and survival from out-of-hospital cardiac arrest: a systematic review. Resuscitation 2019;141:44-62 [FREE Full text] [doi: 10.1016/j.resuscitation.2019.05.018] [Medline: 31199944]
50. Al-Dhahir I, Reijnders T, Faber JS, van den Berg-Emons RJ, Janssen VR, Kraaijenhagen RA, et al. The barriers and facilitators of ehealth-based lifestyle intervention programs for people with a low socioeconomic status: scoping review. J Med Internet Res 2022;24(8):e34229 [FREE Full text] [doi: 10.2196/34229] [Medline: 36001380]
51. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, American Heart Association Stroke Council, Council on CardiovascularStroke Nursing, Council on Clinical Cardiology,Council on Quality of CareOutcomes Research. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016;47(6):e98-e169 [FREE Full text] [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
52. Shishehbor MH, Litaker D, Pothier CE, Lauer MS. Association of socioeconomic status with functional capacity, heart rate recovery, and all-cause mortality. JAMA 2006;295(7):784-792 [FREE Full text] [doi: 10.1001/jama.295.7.784] [Medline: 16478901]
53. Jerath A, Austin PC, Ko DT, Wijeysundera HC, Fremes S, McCormack D, et al. Socioeconomic status and days alive and out of hospital after major elective noncardiac surgery: a population-based cohort study. Anesthesiology 2020;132(4):713-722 [FREE Full text] [doi: 10.1097/ALN.0000000000003123] [Medline: 31972656]
54. Marmot M. Inequalities in health. N Engl J Med 2001;345(2):134-136. [doi: 10.1056/NEJM200107123450210] [Medline: 11450663]
55. Allen L, Williams J, Townsend N, Mikkelsen B, Roberts N, Foster C, et al. Socioeconomic status and non-communicable disease behavioural risk factors in low-income and lower-middle-income countries: a systematic review. Lancet Glob Health 2017;5(3):e277-e289 [FREE Full text] [doi: 10.1016/S2214-109X(17)30058-X] [Medline: 28193397]
56. Medhekar AN, Adhikari S, Abdul-Al AS, Matinrazm S, Kancharla K, Bhonsale A, et al. Lower socioeconomic status is associated with increased long-term mortality after sudden cardiac arrest. Clin Cardiol 2019;42(8):735-740 [FREE Full text] [doi: 10.1002/clc.23211] [Medline: 31165498]
57. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
58. Callaway CW, Schmicker RH, Brown SP, Albrich JM, Andrusiek DL, Aufderheide TP, ROC Investigators. Early coronary angiography and induced hypothermia are associated with survival and functional recovery after out-of-hospital cardiac arrest. Resuscitation 2014;85(5):657-663 [FREE Full text] [doi: 10.1016/j.resuscitation.2013.12.028] [Medline: 24412161]
59. Casey SD, Mumma BE. Sex, race, and insurance status differences in hospital treatment and outcomes following out-of-hospital cardiac arrest. Resuscitation 2018;126:125-129 [FREE Full text] [doi: 10.1016/j.resuscitation.2018.02.027] [Medline: 29518439]
60. Huebinger R, Abella BS, Chavez S, Luber S, Al-Araji R, Panczyk M, et al. Socioeconomic status and post-arrest care after out-of-hospital cardiac arrest in Texas. Resuscitation 2022;176:107-116. [doi: 10.1016/j.resuscitation.2022.03.036] [Medline: 35439577]
61. Alter DA, Naylor CD, Austin P, Tu JV. Effects of socioeconomic status on access to invasive cardiac procedures and on mortality after acute myocardial infarction. N Engl J Med 1999;341(18):1359-1367 [FREE Full text] [doi: 10.1056/NEJM199910283411806] [Medline: 10536129]
62. Simon ST, Kini V, Levy AE, Ho PM. Medication adherence in cardiovascular medicine. BMJ 2021;374:n1493 [FREE Full text] [doi: 10.1136/bmj.n1493] [Medline: 34380627]

63. Liang YH, Wang KH, Huang HM, Shia BC, Chan SY, Ho CW, et al. Reducing medication problems among minority individuals with low socioeconomic status through pharmacist home visits. Int J Environ Res Public Health 2022;19(7):4234 [FREE Full text] [doi: 10.3390/ijerph19074234] [Medline: 35409914]
64. Nielsen KM, Faergeman O, Foldspang A, Larsen ML. Cardiac rehabilitation: health characteristics and socio-economic status among those who do not attend. Eur J Public Health 2008;18(5):479-483 [FREE Full text] [doi: 10.1093/eurpub/ckn060] [Medline: 18614608]
65. Meillier LK, Nielsen KM, Larsen FB, Larsen ML. Socially differentiated cardiac rehabilitation: can we improve referral, attendance and adherence among patients with first myocardial infarction? Scand J Public Health 2012;40(3):286-293 [FREE Full text] [doi: 10.1177/1403494812443600] [Medline: 22637368]
66. Gaalema DE, Elliott RJ, Savage PD, Rengo JL, Cutler AY, Pericot-Valverde I, et al. Financial incentives to increase cardiac rehabilitation participation among low-socioeconomic status patients: a randomized clinical trial. JACC Heart Fail 2019;7(7):537-546 [FREE Full text] [doi: 10.1016/j.jchf.2018.12.008] [Medline: 31078475]
67. Ritchey MD, Maresh S, McNeely J, Shaffer T, Jackson SL, Keteyian SJ, et al. Tracking cardiac rehabilitation participation and completion among medicare beneficiaries to inform the efforts of a national initiative. Circ Cardiovasc Qual Outcomes 2020;13(1):e005902 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.119.005902] [Medline: 31931615]
68. Chandola T, Ferrie J, Sacker A, Marmot M. Social inequalities in self reported health in early old age: follow-up of prospective cohort study. BMJ 2007;334(7601):990 [FREE Full text] [doi: 10.1136/bmj.39167.439792.55] [Medline: 17468119]
69. Elliott VJ, Rodgers DL, Brett SJ. Systematic review of quality of life and other patient-centred outcomes after cardiac arrest survival. Resuscitation 2011;82(3):247-256 [FREE Full text] [doi: 10.1016/j.resuscitation.2010.10.030] [Medline: 21216080]
70. Haraldstad K, Wahl A, Andenæs R, Andersen JR, Andersen MH, Beisland E, LIVSFORSK network. A systematic review of quality of life research in medicine and health sciences. Qual Life Res 2019;28(10):2641-2650 [FREE Full text] [doi: 10.1007/s11136-019-02214-9] [Medline: 31187410]

Abbreviations

aHR: adjusted hazard ratio
CCI: Charlson Comorbidity Index
CPR: cardiopulmonary resuscitation
CR: cardiac rehabilitation
ED: emergency department
HRQOL: health-related quality of life
ICD: International Classification of Diseases
MA: Medical Aid
NHI: National Health Insurance
OHCA: out-of-hospital cardiac arrest
OR: odds ratio
PCI: percutaneous coronary intervention
ROSC: return of spontaneous circulation
SES: socioeconomic status

Edited by A Mavragani; submitted 09.03.23; peer-reviewed by AFW Ho, S Schnaubelt; comments to author 23.03.23; revised version received 12.04.23; accepted 14.06.23; published 11.07.23.

Please cite as:
Yoo KH, Cho Y, Oh J, Lee J, Ko BS, Kang H, Lim TH, Lee SH
Association of Socioeconomic Status With Long-Term Outcome in Survivors After Out-of-Hospital Cardiac Arrest: Nationwide Population-Based Longitudinal Study
JMIR Public Health Surveill 2023;9:e47156
URL: https://publichealth.jmir.org/2023/1/e47156
doi:10.2196/47156
PMID:37432716

Original Paper

Intentional Self-Harm Among US Veterans With Traumatic Brain Injury or Posttraumatic Stress Disorder: Retrospective Cohort Study From 2008 to 2017

Bhanu Pratap Singh Rawat[1], PhD; Joel Reisman[2], AB; Terri K Pogoda[3,4], PhD; Weisong Liu[5], PhD; Subendhu Rongali[1], PhD; Robert H Aseltine Jr[6], PhD; Kun Chen[7], PhD; Jack Tsai[5], PhD; Dan Berlowitz[5], MD; Hong Yu[1,3,5], PhD; Kathleen F Carlson[8,9], PhD

[1]Manning College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA, United States
[2]Center for Healthcare Organization & Implementation Research, VA Boston Healthcare System, Bedford, MA, United States
[3]Center for Healthcare Organization & Implementation Research, VA Boston Healthcare System, Boston, MA, United States
[4]Boston University School of Public Health, Boston, MA, United States
[5]Center of Biomedical and Health Research in Data Sciences, University of Massachusetts Lowell, Lowell, MA, United States
[6]Division of Behavioral Sciences and Community Health, UConn Health, Farmington, CT, United States
[7]Department of Statistics, University of Connecticut, Storrs, CT, United States
[8]Center to Improve Veteran Involvement in Care, VA Portland Health Care System, Portland, OR, United States
[9]Oregon Health & Science University-Portland State University School of Public Health, Portland, OR, United States

Corresponding Author:
Bhanu Pratap Singh Rawat, PhD
Manning College of Information and Computer Sciences
University of Massachusetts Amherst
140 Governors Dr
Amherst, MA, 01002
United States
Phone: 1 4135452744
Email: brawat@umass.edu

Abstract

Background: Veterans with a history of traumatic brain injury (TBI) and/or posttraumatic stress disorder (PTSD) may be at increased risk of suicide attempts and other forms of intentional self-harm as compared to veterans without TBI or PTSD.

Objective: Using administrative data from the US Veterans Health Administration (VHA), we studied associations between TBI and PTSD diagnoses, and subsequent diagnoses of intentional self-harm among US veterans who used VHA health care between 2008 and 2017.

Methods: All veterans with encounters or hospitalizations for intentional self-harm were assigned "index dates" corresponding to the date of the first related visit; among those without intentional self-harm, we randomly selected a date from among the veteran's health care encounters to match the distribution of case index dates over the 10-year period. We then examined the prevalence of TBI and PTSD diagnoses within the 5-year period prior to veterans' index dates. TBI, PTSD, and intentional self-harm were identified using International Classification of Diseases diagnosis and external cause of injury codes from inpatient and outpatient VHA encounters. We stratified analyses by veterans' average yearly VHA utilization in the 5-year period before their index date (low, medium, or high). Variations in prevalence and odds of intentional self-harm diagnoses were compared by veterans' prior TBI and PTSD diagnosis status (TBI only, PTSD only, and comorbid TBI/PTSD) for each VHA utilization stratum. Multivariable models adjusted for age, sex, race, ethnicity, marital status, Department of Veterans Affairs service-connection status, and Charlson Comorbidity Index scores.

Results: About 6.7 million veterans with at least two VHA visits in the 5-year period before their index dates were included in the analyses; 86,644 had at least one intentional self-harm diagnosis during the study period. During the periods prior to veterans' index dates, 93,866 were diagnosed with TBI only; 892,420 with PTSD only; and 102,549 with comorbid TBI/PTSD. Across all three VHA utilization strata, the prevalence of intentional self-harm diagnoses was higher among veterans diagnosed with TBI, PTSD, or TBI/PTSD than among veterans with neither diagnosis. The observed difference was most pronounced among veterans in the high VHA utilization stratum. The prevalence of intentional self-harm was six times higher among those with comorbid TBI/PTSD (6778/58,295, 11.63%) than among veterans with neither TBI nor PTSD (21,979/1,144,991, 1.92%). Adjusted odds ratios suggested that, after accounting for potential confounders, veterans with TBI, PTSD, or comorbid TBI/PTSD had higher odds of self-harm compared to veterans without these diagnoses. Among veterans with high VHA utilization, those with comorbid TBI/PTSD were 4.26 (95% CI 4.15-4.38) times more likely to receive diagnoses for intentional self-harm than veterans with neither diagnosis. This pattern was similar for veterans with low and medium VHA utilization.

Conclusions: Veterans with TBI and/or PTSD diagnoses, compared to those with neither diagnosis, were substantially more likely to be subsequently diagnosed with intentional self-harm between 2008 and 2017. These associations were most pronounced among veterans who used VHA health care most frequently. These findings suggest a need for suicide prevention efforts targeted at veterans with these diagnoses.

(JMIR Public Health Surveill 2023;9:e42803) doi:10.2196/42803

KEYWORDS

self-harm; suicide; suicide attempt; suicidal ideation; veteran; suicidal; brain injury; trauma; posttraumatic stress disorder; PTSD; big data; prevalence; correlation; risk factor; traumatic brain injury; TBI

Introduction

For decades, suicide has been one of the leading causes of death [1,2] and a major health concern among veterans in the United States [3]. The most recent Department of Veterans Affairs (VA) Office of Mental Health and Suicide Prevention's National Veteran Suicide Prevention Annual Report stated that the annual suicide rate for veterans (31.6 per 100,000) was twice the rate of that for nonveteran US adults (16.8 per 100,000) [4].

Suicide prevention is a top clinical priority for the VA and its nationally integrated health care system, overseen by the Veterans Health Administration (VHA) [4]. Central to this effort is the identification of individuals at risk for suicide or suicidal behavior [5,6]. A substantial body of research has used electronic health record data to identify individuals likely to make suicide attempts and other forms of intentional self-harm [7-10]. Research has shown that adults and youths with a history of nonfatal self-harm are at a markedly elevated risk of suicide [11-13].

Other important risk factors for suicidal behavior include sociodemographic, psychological, and physical health characteristics [14-17]. In particular, two health conditions relatively common among veterans—traumatic brain injury (TBI) and posttraumatic stress disorder (PTSD)—have been identified in multiple studies as risk factors for suicide [2,18-20]. For example, Bullman and Kang [21] reported that Vietnam veterans with PTSD were at elevated risk of suicide death when compared with those without PTSD. Another study by Jakupcak et al [22] observed similar patterns for post-9/11 veterans, in which veterans with PTSD were reported to have a four times greater likelihood of suicidal ideation as compared to veterans without PTSD. In a 2011 study, Brenner et al [18] reported that PTSD was associated with 2.8 times greater odds of a documented suicide attempt compared to veterans without PTSD; the authors suggested that future analyses compare those with and without comorbid TBI history. In another study, Brenner et al [20] conducted an analysis to examine the associations between TBI diagnosis and suicide among individuals receiving care at VHA and reported that veterans with TBI were 1.55 times more likely to complete suicide as compared to those without TBI. The mechanisms by which TBI or PTSD lead to an elevated risk of suicide or suicidal behaviors are not well understood, but the associations are consistent with constructs of risk described in the interpersonal theory of suicide (ie, thwarted belongingness, perceived burdensomeness, and acquired capability) [23]. Notably, individuals with TBI or PTSD often have co-occurring physical health (eg, chronic pain), mental health (eg, depression and anxiety), and behavioral conditions (eg, substance use disorder) that can lead to functional limitations and other vulnerabilities (eg, unemployment and homelessness) [24-28], which together or individually, can elevate the factors comprising each of these constructs [29].

Previous studies examining associations between TBI and PTSD diagnoses and suicidal behavior have mostly been conducted with relatively small or select samples or with veterans from a single war era [2,18,19]. These studies have suggested that TBI and PTSD are important risk factors for suicidal behavior. Moreover, previous studies have tended to only analyze the associations of TBI [30] and PTSD [31] separately with suicidal behavior and have not examined the potential interactive effects of comorbid TBI and PTSD.

In this study, we analyzed data from approximately 6.7 million veterans who received VHA health care between 2008 and 2017 to examine population-level variations in the prevalence of intentional self-harm among those with or without prior TBI and PTSD diagnosis. We focused on suicide attempts as well as other intentional self-harm events because previous research [11,12] has shown that adults and youths were at a markedly elevated risk of suicide completion after nonfatal self-harm events. Our analysis was stratified by level of VHA utilization to account for differential access to VHA care and the opportunity to receive the diagnoses of interest. Our objectives were to evaluate variation in the prevalence of intentional self-harm by veterans' prior TBI and PTSD diagnosis status across different VHA utilization levels, and to examine the odds of veterans' intentional self-harm by prior TBI and PTSD diagnosis status across different VHA utilization levels while controlling for potentially confounding variables. The overarching goal of this work is to inform suicide and other intentional self-harm prevention efforts.

Methods

Ethics Approval

This study was approved by the VA Bedford Healthcare System Institutional Review Board. We identified 6,693,275 veterans who visited a VHA facility at least twice during the 10-year period between January 1, 2008, and December 31, 2017. We extracted veterans' demographic and health care data from the VHA Corporate Data Warehouse (CDW) database. All preprocessing and analyses were conducted using the VA Informatics and Computing Infrastructure (VINCI) accessed from Bedford, Virginia.

Dependent Variable

This study defined intentional self-harm as suicide attempts or other self-harm using diagnosis codes specified for those behaviors supplemented with additional evidence that an injury incurred by a veteran was due to self-harm. Suicide attempt or intentional self-harm was defined as a health care encounter (outpatient visit, including urgent care, or hospitalization) that met any of three conditions: presence of a code that explicitly indicated suicide attempt or intentional self-harm, presence of codes identifying suicidal ideation *and* injury or poisoning assigned for the same encounter, or presence of codes identifying injury or poisoning *and* diagnoses for a mental disorder assigned for the same encounter. Patrick et al [32] used this definition based on diagnosis and external cause of injury codes from the *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM*) manual. For this study, we derived a set of *International Statistical Classification of Diseases, Tenth Revision* (*ICD-10*) codes that corresponded with the *ICD-9* codes used by Patrick et al [32]. The algorithms and ICD codes used are provided in Table S1 in Multimedia Appendix 1.

Independent Variable

For veterans with one or more intentional self-harm events, we chose the first coded incident of intentional self-harm as the "index date." For veterans with no intentional self-harm events, we randomly selected a date from among the veterans' encounters to match the distribution of dates among those with self-harm events in the 10-year period. Each veteran was classified into one of four mutually exclusive groups based on the presence or absence of TBI and PTSD diagnoses during the 5 years prior to the veteran's index date:

- No TBI or PTSD diagnoses
- TBI diagnosis only
- PTSD diagnosis only
- Comorbid TBI/PTSD diagnoses

The codes used to identify both TBI and PTSD are included in Table S2 in Multimedia Appendix 1.

To account for "informed presence" bias based on the varying opportunity to receive the diagnosis codes of interest arising from differential health care utilization [33], we classified veterans into 3 groups according to the average annual VHA health care visits in the 5-year period before their "index date." The *low* VHA utilization group included those at the 50th percentile and below (≤6.6 visits per year), the *medium* utilization group included those above the 50th but below and at the 75th percentiles (>6.6 and ≤15.6 visits), and the *high* utilization group included those above the 75th percentile (>15.6 visits). Analyses were stratified by VHA utilization group.

Covariates

Veterans' age, sex, race, ethnicity, marital status, and VA service-connection status—a measure of military service–related disability assigned by the Veterans Benefits Administration—were extracted from the CDW. Age was categorized as <30, 30-44, 45-69, or ≥70 years. Sex was categorized as a binary variable, while race was categorized as White, Black or African American, Asian, Native Hawaiian or other Pacific Islander, American Indian or Alaska Native, or other and ethnicity as Hispanic versus non-Hispanic. Marital status was categorized as single, married, divorced, or widowed, and VA service-connection as no service-connection, service-connected <50%, or service-connected ≥50%. We also computed the Charlson Comorbidity Index (CCI) scores using diagnoses from inpatient and outpatient visits, and categorized veterans as having low (<5), medium (5-15), or high (>15) Charlson Comorbidity scores [34].

Analyses

We calculated the prevalence of veterans' intentional self-harm (yes/no) according to TBI and PTSD status for each VHA health care utilization group. We modeled the likelihood of intentional self-harm among veterans by TBI and PTSD status using logistic regression to estimate odds ratios (ORs) and their corresponding 95% CIs. Adjusted ORs (AORs) were calculated using two multivariable logistic regression models: minimally adjusted models that controlled for age, sex, and race, and fully adjusted models that controlled for age, sex, race, ethnicity, marital status, VA service-connection status, and CCI scores.

Results

Overview

Among the 6,693,275 veterans who used VHA health care between 2008 and 2017 at least twice, 93.1% (n=6,231,400) were coded as male and 6.9% (n=461,874) as female; 76.74% (n=5,136,343) were coded as being White, 15.85% (n=1,060,951) Black or African American, 0.98% (n=65,402) Asian, 0.84% (n=56,167) American Indian or Alaska Native, 0.96% (n=63,943) Native Hawaiian or other Pacific Islander, and 4.64% (n=310,469) as other (Table 1). More than one-third of veterans (n=2,318,144, 34.63%) were service-connected at 50% or higher. The VHA health care utilization strata had the following composition of veterans: *low* (n=3,327,615, 49.72%); *medium* (n=1,679,954, 25.09%), and *high* (n=1,686,066, 25.19%). Overall, we identified 86,644 (1.29%) veterans with one or more encounters involving intentional self-harm. TBI and PTSD were relatively prevalent among VHA users, with 1.4% (n=93,866) having been diagnosed with TBI only, 13.33% (n=892,420) with PTSD only, and 1.53% (n=102,549) with comorbid TBI/PTSD.

Table 1. Prevalence of intentional self-harm by veterans’ sociodemographic and health variables, stratified by VHA utilization.

	Low VHA[a] utilization		Medium VHA utilization		High VHA utilization	
	Self-harm, n (%)	Total veterans, n	Self-harm, n (%)	Total veterans, n	Self-harm, n (%)	Total veterans, n
Age (years)						
<30	2256 (0.93)	243,318	3007 (5.76)	52,241	4150 (15.85)	26,188
30-44	2712 (0.54)	497,756	4841 (2.61)	185,421	12,119 (8.17)	148,312
45-69	5107 (0.35)	1,472,935	9838 (1.15)	854,474	36,067 (3.63)	993,893
≥70	669 (0.06)	1,113,606	1228 (0.21)	587,458	4650 (0.9)	517,673
Sex						
Female	1031 (0.49)	212,124	2097 (1.85)	113,257	6772 (4.96)	136,493
Male	9713 (0.31)	3,115,490	16,817 (1.07)	1,566,337	50,214 (3.24)	1,549,573
Race						
White	7958 (0.31)	2,608,591	14,277 (1.1)	1,292,525	42,159 (3.41)	1,235,227
Black or African American	1801 (0.38)	469,890	3114 (1.18)	263,253	10,856 (3.31)	327,808
Asian	170 (0.42)	40,872	176 (1.24)	14,187	363 (3.51)	10,343
American Indian or Alaska Native	129 (0.46)	27,919	231 (1.71)	13,482	734 (4.97)	14,766
Native Hawaiian or other Pacific Islander	140 (0.45)	31,152	256 (1.58)	16,157	503 (3.02)	16,634
Other	546 (0.37)	149,191	860 (1.08)	79,990	2371 (2.92)	81,288
Ethnicity						
Hispanic	705 (0.4)	175,414	1272 (1.35)	94,218	3667 (3.56)	102,962
Not Hispanic	10,039 (0.31)	3,152,201	17,642 (1.11)	1,585,376	53,319 (3.36)	1,583,104
Marital status						
Single	2437 (0.57)	427,381	3838 (2.04)	188,239	11,463 (5.55)	206,491
Married	3839 (0.2)	1,940,210	6572 (0.71)	928,639	17,178 (2.11)	814,864
Widowed	377 (0.15)	253,548	755 (0.56)	135,848	2767 (2.05)	134,697
Divorced	4007 (0.59)	677,616	7663 (1.81)	422,514	25,495 (4.83)	527,684
Unknown	84 (0.29)	28,860	86 (1.98)	4354	83 (3.56)	2330
Veterans Affairs service-connection status						
No service-connection	4141 (0.25)	1,659,561	6184 (0.81)	759,889	16,890 (2.79)	605,966
Service-connection <50%	1806 (0.24)	751,492	2940 (0.88)	333,972	7918 (3.0)	264,251
Service-connection ≥50%	4797 (0.52)	916,562	9790 (1.67)	585,733	32,178 (3.94)	815,849
Charlson Comorbidity Index score						
≤5	7889 (0.38)	2,092,840	12,078 (1.52)	792,553	26,207 (4.72)	555,414
5-15	2787 (0.23)	1,221,470	6662 (0.77)	869,826	29,407 (2.72)	1,079,467
≥15	68 (0.51)	13,305	174 (1.01)	17,215	1372 (2.68)	51,185

[a] VHA: Veterans Health Administration.

Prevalence of Intentional Self-Harm by TBI and PTSD Diagnosis Status

The prevalence of intentional self-harm by TBI and PTSD diagnosis status is presented in Table 2, stratified by *low*, *medium*, and *high* VHA utilization groups.

Between 2008 and 2017, the overall prevalence of intentional self-harm was highest for veterans in the *high* VHA utilization group (56,986/1,743,052, 3.38%). In this stratum, the prevalence of intentional self-harm was highest for veterans in the comorbid TBI/PTSD diagnosis group (6778/58,295, 11.63%) and lowest among those with no TBI or PTSD diagnosis (21,979/1,144,991, 1.92%). Those in the TBI only and PTSD only diagnosis groups also had elevated prevalence levels, at 5.21% (2335/44,833) and 5.91% (25,894/437,947), respectively. This pattern of elevated prevalence was similar for veterans in the *medium* and *low* VHA utilization groups; however, the disparity between diagnosis groups in the *medium* and *low* VHA utilization groups was considerably lower than in the *high* VHA utilization group. Across all three VHA utilization groups, the prevalence of self-harm diagnosis among veterans with PTSD only was consistently higher by about 0.7% than those in the TBI only diagnosis group.

Table 2. Prevalence of intentional self-harm by veterans' TBI and PTSD diagnosis status, stratified by Veterans Health Administration utilization.

	Intentional self-harm	
	Yes	No
Veteran diagnosis status, n	86,644	6,606,631
Low health care utilization (less than 50th percentile), n (%)		
TBI[a] and PTSD[b] diagnoses		
Neither	6792 (0.22)	3,065,664 (99.78)
TBI only	215 (0.84)	25,409 (99.16)
PTSD only	3407 (1.6)	209,468 (98.40)
Comorbid TBI/PTSD	330 (1.98)	16,330 (98.02)
Total	10,744 (0.32)	3,316,871 (99.68)
Medium health care utilization (50th-75th percentile), n (%)		
TBI and PTSD diagnoses		
Neither	9466 (0.68)	1,377,527 (99.32)
TBI only	558 (2.38)	22,851 (97.62)
PTSD only	7548 (3.12)	234,050 (96.88)
Comorbid TBI/PTSD	1342 (4.86)	26,252 (95.14)
Total	18,914 (1.13)	1,660,680 (98.77)
High health care utilization (75th percentile and higher), n (%)		
TBI and PTSD diagnoses		
Neither	21,979 (1.92)	1,123,012 (98.08)
TBI only	2335 (5.21)	42,498 (94.79)
PTSD only	25,894 (5.91)	412,053 (94.09)
Comorbid TBI/PTSD	6778 (11.63)	51,517 (88.37)
Total	56,986 (3.38)	1,629,080 (96.62)

[a] TBI: traumatic brain injury.

[b] PTSD: posttraumtic stress disorder.

Odds of Intentional Self-Harm by TBI and PTSD Diagnosis Status

Table 3 presents ORs and AORs based on bivariable and multivariable logistic regression models, again stratified by VHA utilization group. (ORs for model covariates are provided in Table S3 in Multimedia Appendix 1.)

Among veterans in the *high* VHA utilization group, and compared to veterans with no TBI or PTSD diagnoses, those with comorbid TBI/PTSD had the highest odds of intentional self-harm (OR 6.72, 95% CI 6.54-6.90), followed by those in the PTSD only group (OR 3.21, 95% CI 3.16-3.26) and then those in the TBI only group (OR 2.81, 95% CI 2.69-2.93). These patterns were similar in the fully adjusted model, where the AORs for comorbid TBI/PTSD group were the highest (AOR 4.26, 95% CI 4.15-4.38), followed by the PTSD only group (AOR 2.90, 95% CI 2.85-2.95) and then by the TBI only group (OR 2.44, 95% CI 2.34-2.55).

Among veterans in the *medium* VHA utilization group, the odds of intentional self-harm were highest among those with comorbid TBI/PTSD (OR 7.42, 95% CI 7.01-7.85), followed by the PTSD only group (OR 4.68, 95% CI 4.54-4.82) and then the TBI only group (OR 3.54, 95% CI 3.25-3.85). This pattern changed slightly when potential confounders were added to the models: AORs using the fully adjusted model were highest among those with PTSD only (AOR 3.23, 95% CI 3.14-3.33) and lowest among those with TBI only (AOR 2.32, 95% CI 2.13-2.53). The AOR for the comorbid TBI/PTSD group decreased from 7.42 (95% CI 7.01-7.85) to 3.17 (95% CI 2.99-3.36) after accounting for all available confounding variables.

Among veterans in the *low* VHA utilization group, the odds of intentional self-harm were also highest among those with comorbid TBI/PTSD (OR 9.05, 95% CI 8.10-10.11), followed by the PTSD only group (OR 7.33, 95% CI 7.04-7.64) and then the TBI only group (OR 3.79, 95% CI 3.31-4.34). Similar to the *medium* VHA utilization group, the pattern changed when potential confounders were added to the fully adjusted models: AORs for the PTSD only group (AOR 5.14, 95% CI 4.93-5.35) were highest, closely followed by the comorbid TBI/PTSD group (AOR 4.83, 95% CI 4.32-5.40) and then the TBI only group (AOR 2.48, 95% CI 2.17-2.84).

Table 3. Odds of intentional self-harm by veterans' TBI and PTSD diagnosis status, stratified by Veterans Health Administration utilization.

Veteran diagnosis status	Logistic regression		
	Bivariable, OR[a] (95% CI)	Multivariable, AOR[b] (95% CI)	
	Model	Minimally adjusted model[c]	Fully adjusted model[d]
Low health care utilization (less than 50th percentile)			
TBI[e] and PTSD[f] diagnoses			
Neither	Reference	Reference	Reference
TBI only	3.79 (3.31-4.34)	2.60 (2.27-2.98)	2.48 (2.17-2.84)
PTSD only	7.33 (7.04-7.64)	5.22 (5.01-5.44)	5.14 (4.93-5.35)
Comorbid TBI/PTSD	9.05 (8.10-10.11)	4.85 (4.34-5.42)	4.83 (4.32-5.40)
Medium health care utilization (50th-75th percentile)			
TBI and PTSD diagnoses			
Neither	Reference	Reference	Reference
TBI only	3.54 (3.25-3.85)	2.33 (2.14-2.54)	2.32 (2.13-2.53)
PTSD only	4.68 (4.54-4.82)	2.91 (2.83-3.00)	3.23 (3.14-3.33)
Comorbid TBI/PTSD	7.42 (7.01-7.85)	2.81 (2.65-2.97)	3.17 (2.99-3.36)
High health care utilization (75th percentile and higher)			
TBI and PTSD diagnoses			
Neither	Reference	Reference	Reference
TBI only	2.81 (2.69-2.93)	2.50 (2.40-2.61)	2.44 (2.34-2.55)
PTSD only	3.21 (3.16-3.26)	2.35 (2.31-2.39)	2.90 (2.85-2.95)
Comorbid TBI/PTSD	6.72 (6.54-6.90)	3.40 (3.31-3.49)	4.26 (4.15-4.38)

[a] OR: odds ratio.

[b] AOR: adjusted odds ratio.

[c] Minimally adjusted model included age, sex, and race.

[d] Fully adjusted model included age, sex, race, ethnicity, marital status, Veterans Affairs service-connection status, and Charlson Comorbidity Index scores.

[e] TBI: traumatic brain injury.

[f] PTSD: posttraumtic stress disorder.

Discussion

Our analysis of approximately 6.7 million VHA-using veterans found that those with a documented diagnosis of TBI or PTSD were significantly more likely to have subsequent VHA documentation of a suicide attempt or other form of intentional self-harm relative to those with neither diagnosis between 2008 and 2017. For all three VHA utilization groups, the odds of intentional self-harm for veterans in the comorbid TBI/PTSD group were more than three times as high as those with neither condition, even after controlling for multiple potential confounders. For those in the PTSD only group, the adjusted odds of intentional self-harm were approximately three times that of those with neither condition. For the TBI only group, the adjusted odds of intentional self-harm were nearly double relative to those with neither condition. For all three VHA utilization groups, the prevalence of intentional self-harm for veterans in the comorbid TBI/PTSD group was higher than in the TBI only and PTSD only groups.

For the *medium* and *low* VHA utilization groups, adjusted odds were highest for the PTSD only group closely followed by the comorbid TBI/PTSD group. However, for the *high* VHA utilization group, adjusted odds for the comorbid TBI/PTSD group were consistently higher than those of the PTSD only group. Previous studies have suggested that symptom severity may play an important role in this association [35,36]. Our study did not consider symptom severity because this measure is difficult to ascertain in a large administrative data set such as the VHA's CDW. Future research should consider how TBI and PTSD symptom severity drive functional outcomes and, in turn, may lead to intentional self-harm, as addressing functional limitations may be an important focus for prevention programs.

Psychological and functional impairments can be common for veterans with co-occurring PTSD and TBI [37]. PTSD is a known contributor to disability [38], resulting in more than US $40 billion in costs related to direct health care and unemployment among military populations [39]. Despite experiencing long-term symptoms of PTSD and the availability of various therapies, many individuals with PTSD wait years to decades before seeking professional treatment [40,41]. Patients may be skeptical toward therapy [42] and believe treatment to be ineffective, harmful [43], and only for extreme problems. Some may also believe treatment involves a loss of control or autonomy, or as something for people who are "weak," "crazy," or "incompetent" [40]. Some individuals with PTSD do not seek help at all, as only approximately one-quarter of individuals with PTSD symptoms receive mental health services [44,45]. A significant number of veterans drop out of their PTSD treatment altogether due to reasons such as work interference, stigma related to the ailment, confidentiality concerns, and perceived treatment ineffectiveness, thus contributing to a population of veterans with unresolved symptoms related to PTSD [46,47].

Similarly, veterans with TBI, regardless of severity, may have a variety of co-occurring psychological and physical health challenges for years following injury [48], including affective, cognitive, somatosensory, and vestibular postconcussive symptoms [49]. Other impairments, especially documented for post-9/11 veterans, may include social/family dysfunction, unemployment, and general difficulty with community reintegration [25]. Individuals with TBI also may not be completely aware of their limitations [50] and, therefore, unable to fully engage in treatment [51]. Given the observed difficulties of veterans with documented PTSD, TBI, or both, VHA suicide prevention services should increase efforts to educate veterans on the importance and nature of therapeutic options to better engage at-risk veterans with VA health care services.

Veterans with TBI or PTSD also face different challenges in seeking treatment such as problems with access, sociocultural environment, past medical experiences, and illness burden [42,45,52]. Access barriers are mainly organizational, due to lack of knowledge about VHA, and logistical, and include time-consuming and complex enrollment processes in VHA, veterans not knowing which VHA services they are eligible for, and time- and distance-related expenses [53]. VHA has made extensive efforts to mitigate these logistical problems, including the implementation of the community care program, which allows veterans to seek care in the community if VHA facilities are unable to furnish services or if long travel distances or appointment wait times pose a burden [54].

A study performed by McCarthy et al [5] suggests that the inclusion of the Recovery Engagement and Coordination for Health–Veterans Enhanced Treatment (REACH VET) program at the VHA has resulted in reduced prevalence of nonfatal suicide attempts along with other greater treatment engagement and care provided to veterans. The algorithm used in REACH VET to identify individuals in the facilities' top 0.1% suicide risk tiers uses TBI and PTSD diagnoses as variables among the 381 measures used for prediction [6,17]. Our study strengthens the case for using TBI and PTSD as risk factors for suicide risk prediction in the algorithm. Our findings can further help in the identification of high suicide risk veterans by flagging patients with comorbid TBI/PTSD diagnosis using the already available identifiers used by REACH VET [2,18,55].

There are several limitations to take into account in the context of these results. First, this study only included veterans who received health care in VHA settings. Motives for veterans not seeking VHA care are complex, and thus these veterans likely have different TBI-, PTSD-, and intentional self-harm–related profiles. Second, this study relied on diagnoses in VHA administrative data to identify both TBI and PTSD diagnosis status and intentional self-harm outcomes. The *high* VHA utilization group, while comprising only 25.19% (1,686,066/6,693,275) of the study population, comprised 65.77% (56,986/86,644) of those with detected suicide attempts and other forms of intentional self-harm, 47.76% (44,833/93,866) of the TBI only group, 49.07% (437,947/892,420) of the PTSD only group, and 56.85% (58,295/102,549) of the comorbid TBI/PTSD group. It is likely that underdiagnosis was present [56]. For example, veterans within the *low* and *medium* VHA utilization groups may have experienced TBI, PTSD, or both but received care for them outside the VHA or not at all. Such misclassification of true exposure or outcomes status can bias results. However, we observed similar patterns across all three levels of VHA health care utilization, suggesting that results may not be explained entirely by informed presence bias. Third, our study used a broad range of diagnosis codes for intentional self-harm, per Patrick et al [32], to potentially decrease the likelihood that a true self-harm event would be missed (eg, if only suicide attempt codes were used). Future work that further examines and establishes the validity of more ICD codes to detect intentional self-harm would strengthen this line of research. Fourth, the analysis represented TBI and PTSD simply as present or absent according to diagnosis codes. A more refined analysis would incorporate the nature of trauma exposure, symptom severity, barriers to care, and social risk factors. Fifth, our work studied the associations of different TBI and PTSD diagnosis groups with intentional self-harm while adjusting for age, sex, race, ethnicity, marital status, VA service-connection status, and different comorbidities. This analysis did not directly examine the effects of these covariates on intentional self-harm. Further in-depth analysis is required to understand the associations between these sociodemographic and health variables and intentional self-harm while controlling for confounder sets relevant to those specific associations.

Our study adds to the literature on associations of TBI and PTSD diagnoses with intentional self-harm. It provides generalized results across a large veteran population of varied demographics and VHA utilization. Our findings indicate that veterans with a TBI and/or PTSD diagnosis are more likely to have an intentional self-harm event than veterans without either of these diagnoses. These results may contribute to VHA clinical decision-making regarding the frequency and intensity of follow-up or treatment for veterans with PTSD and TBI diagnoses.

Future research on the risk of intentional self-harm among veterans with TBI and/or PTSD can benefit from both qualitative and quantitative efforts. Qualitatively exploring veterans' experiences and perceptions of health and functional status, as well as facilitators and barriers to health care services, can help to better identify health and access needs. Quantitatively exploring TBI and PTSD severity based on the nature of trauma exposure, current or previous treatment, and the intensity of posttraumatic symptoms can provide more granular associations between TBI and PTSD diagnosis and intentional self-harm. These efforts may help to inform clinicians in targeting preventative and treatment strategies toward a high-risk veteran cohort.

This study examined 6.7 million veteran VHA users and found that those with TBI and/or PTSD diagnoses were more likely to have documented intentional self-harm compared to veterans with neither diagnosis. Veterans with TBI, PTSD, or comorbid TBI/PTSD should be considered high-priority recipients of VA suicide prevention services and may need special engagement efforts given various access and treatment barriers experienced among those with these conditions.

Acknowledgments

Research reported in this study was supported by the National Institute of Mental Health of the National Institutes of Health under award R01MH125027. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 38 KB - publichealth_v9i1e42803_app1.docx]

References

1. Kang HK, Bullman TA. Is there an epidemic of suicides among current and former U.S. military personnel? Ann Epidemiol 2009 Oct;19(10):757-760. [doi: 10.1016/j.annepidem.2009.05.004] [Medline: 19628411]
2. Barnes SM, Walter KH, Chard KM. Does a history of mild traumatic brain injury increase suicide risk in veterans with PTSD? Rehabil Psychol 2012 Feb;57(1):18-26. [doi: 10.1037/a0027007] [Medline: 22369114]
3. Kaplan MS, Huguet N, McFarland BH, Newsom JT. Suicide among male veterans: a prospective population-based study. J Epidemiol Community Health 2007 Jul;61(7):619-624 [FREE Full text] [doi: 10.1136/jech.2006.054346] [Medline: 17568055]
4. 2021 national veteran suicide prevention: annual report. US Department of Veterans Affairs: Mental Health. 2021 Sep. URL: https://www.mentalhealth.va.gov/docs/data-sheets/2021/2021-National-Veteran-Suicide-Prevention-Annual-Report-FINAL-9-8-21.pdf [accessed 2023-05-05]
5. McCarthy JF, Cooper SA, Dent KR, Eagan AE, Matarazzo BB, Hannemann CM, et al. Evaluation of the recovery engagement and coordination for health-veterans enhanced treatment suicide risk modeling clinical program in the Veterans Health Administration. JAMA Netw Open 2021 Oct 01;4(10):e2129900 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.29900] [Medline: 34661661]
6. Kessler RC, Hwang I, Hoffmire CA, McCarthy JF, Petukhova MV, Rosellini AJ, et al. Developing a practical suicide risk prediction model for targeting high-risk patients in the Veterans health Administration. Int J Methods Psychiatr Res 2017 Sep;26(3):e1575 [FREE Full text] [doi: 10.1002/mpr.1575] [Medline: 28675617]
7. Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ, Army STARRS Collaborators. Predicting suicides after psychiatric hospitalization in US Army soldiers: the Army Study To Assess Risk and rEsilience in Servicemembers (Army STARRS). JAMA Psychiatry 2015 Jan;72(1):49-57 [FREE Full text] [doi: 10.1001/jamapsychiatry.2014.1754] [Medline: 25390793]
8. Barak-Corren Y, Castro VM, Javitt S, Hoffnagle AG, Dai Y, Perlis RH, et al. Predicting suicidal behavior from longitudinal electronic health records. Am J Psychiatry 2017 Feb 01;174(2):154-162. [doi: 10.1176/appi.ajp.2016.16010077] [Medline: 27609239]
9. Poulin C, Shiner B, Thompson P, Vepstas L, Young-Xu Y, Goertzel B, et al. Predicting the risk of suicide by analyzing the text of clinical notes. PLoS One 2014;9(1):e85733 [FREE Full text] [doi: 10.1371/journal.pone.0085733] [Medline: 24489669]
10. Ben-Ari A, Hammond K. Text mining the EMR for modeling and predicting suicidal behavior among US veterans of the 1991 Persian Gulf War. 2015 Presented at: 48th Hawaii International Conference on System Sciences; January 5-8, 2015; Kauai, HI. [doi: 10.1109/hicss.2015.382]

11. Olfson M, Wall M, Wang S, Crystal S, Bridge J, Liu S, et al. Suicide after deliberate self-harm in adolescents and young adults. Pediatrics 2018 Apr;141(4):e20173517. [doi: 10.1542/peds.2017-3517] [Medline: 29555689]
12. Olfson M, Wall M, Wang S, Crystal S, Gerhard T, Blanco C. Suicide following deliberate self-harm. Am J Psychiatry 2017 Aug 01;174(8):765-774. [doi: 10.1176/appi.ajp.2017.16111288] [Medline: 28320225]
13. Klonsky ED, May AM, Glenn CR. The relationship between nonsuicidal self-injury and attempted suicide: converging evidence from four samples. J Abnorm Psychol 2013 Feb;122(1):231-237. [doi: 10.1037/a0030278] [Medline: 23067259]
14. Simon GE, Johnson E, Lawrence JM, Rossom RC, Ahmedani B, Lynch FL, et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. Am J Psychiatry 2018 Oct 01;175(10):951-960 [FREE Full text] [doi: 10.1176/appi.ajp.2018.17101167] [Medline: 29792051]
15. Walkup JT, Townsend L, Crystal S, Olfson M. A systematic review of validated methods for identifying suicide or suicidal ideation using administrative or claims data. Pharmacoepidemiol Drug Saf 2012 Jan;21 Suppl 1:174-182. [doi: 10.1002/pds.2335] [Medline: 22262604]
16. Tran T, Luo W, Phung D, Harvey R, Berk M, Kennedy RL, et al. Risk stratification using data from electronic medical records better predicts suicide risks than clinician assessments. BMC Psychiatry 2014 Mar 14;14:76 [FREE Full text] [doi: 10.1186/1471-244X-14-76] [Medline: 24628849]
17. McCarthy JF, Bossarte RM, Katz IR, Thompson C, Kemp J, Hannemann CM, et al. Predictive modeling and concentration of the risk of suicide: implications for preventive interventions in the US Department of Veterans Affairs. Am J Public Health 2015 Sep;105(9):1935-1942. [doi: 10.2105/AJPH.2015.302737] [Medline: 26066914]
18. Brenner L, Betthauser L, Homaifar B, Villarreal E, Harwood JEF, Staves PJ, et al. Posttraumatic stress disorder, traumatic brain injury, and suicide attempt history among veterans receiving mental health services. Suicide Life Threat Behav 2011 Aug;41(4):416-423. [doi: 10.1111/j.1943-278X.2011.00041.x] [Medline: 21599727]
19. Pompili M, Sher L, Serafini G, Forte A, Innamorati M, Dominici G, et al. Posttraumatic stress disorder and suicide risk among veterans: a literature review. J Nerv Ment Dis 2013 Sep;201(9):802-812. [doi: 10.1097/NMD.0b013e3182a21458] [Medline: 23995037]
20. Brenner L, Ignacio R, Blow F. Suicide and traumatic brain injury among individuals seeking Veterans Health Administration services. J Head Trauma Rehabil 2011;26(4):257-264. [doi: 10.1097/HTR.0b013e31821fdb6e] [Medline: 21734509]
21. Bullman TA, Kang HK. Posttraumatic stress disorder and the risk of traumatic deaths among Vietnam veterans. J Nerv Ment Dis 1994 Nov;182(11):604-610. [doi: 10.1097/00005053-199411000-00002] [Medline: 7964667]
22. Jakupcak M, Cook J, Imel Z, Fontana A, Rosenheck R, McFall M. Posttraumatic stress disorder as a risk factor for suicidal ideation in Iraq and Afghanistan War veterans. J Trauma Stress 2009 Aug;22(4):303-306. [doi: 10.1002/jts.20423] [Medline: 19626682]
23. Joiner TE. Why People Die by Suicide. Cambridge, MA: Harvard University Press; 2005.
24. Seal KH, Bertenthal D, Barnes DE, Byers AL, Strigo I, Yaffe K, Chronic Effects of Neurotrauma Consortium Study Group. Association of traumatic brain injury with chronic pain in Iraq and Afghanistan veterans: effect of comorbid mental health conditions. Arch Phys Med Rehabil 2017 Aug;98(8):1636-1645. [doi: 10.1016/j.apmr.2017.03.026] [Medline: 28455190]
25. Pugh MJ, Swan AA, Carlson KF, Jaramillo CA, Eapen BC, Dillahunt-Aspillaga C, Trajectories of Resilience and Complex Comorbidity Study Team. Traumatic brain injury severity, comorbidity, social support, family functioning, and community reintegration among veterans of the Afghanistan and Iraq wars. Arch Phys Med Rehabil 2018 Feb;99(2S):S40-S49. [doi: 10.1016/j.apmr.2017.05.021] [Medline: 28648681]
26. Amick M, Meterko M, Fortier CB, Fonda JR, Milberg WP, McGlinchey RE. The deployment trauma phenotype and employment status in veterans of the wars in Iraq and Afghanistan. J Head Trauma Rehabil 2018;33(2):E30-E40 [FREE Full text] [doi: 10.1097/HTR.0000000000000308] [Medline: 28422901]
27. Greer N, Ackland P, Sayer N, Spoont M, Taylor B, MacDonald R, et al. Relationship of Deployment-Related Mild Traumatic Brain Injury to Posttraumatic Stress Disorder, Depressive Disorders, Substance Use Disorders, Suicidal Ideation, and Anxiety Disorders: A Systematic Review. Washington, DC: Department of Veterans Affairs (US); Mar 2019.
28. Twamley EW, Hays CC, Van Patten R, Seewald PM, Orff HJ, Depp CA, et al. Neurocognition, psychiatric symptoms, and lifetime homelessness among veterans with a history of traumatic brain injury. Psychiatry Res 2019 Jan;271:167-170. [doi: 10.1016/j.psychres.2018.11.049] [Medline: 30481694]
29. Chu C, Buchman-Schmitt JM, Stanley IH, Hom MA, Tucker RP, Hagan CR, et al. The interpersonal theory of suicide: a systematic review and meta-analysis of a decade of cross-national research. Psychol Bull 2017 Dec;143(12):1313-1345 [FREE Full text] [doi: 10.1037/bul0000123] [Medline: 29072480]
30. Greer N, Sayer NA, Spoont M, Taylor BC, Ackland PE, MacDonald R, et al. Prevalence and severity of psychiatric disorders and suicidal behavior in service members and veterans with and without traumatic brain injury: systematic review. J Head Trauma Rehabil 2020;35(1):1-13. [doi: 10.1097/HTR.0000000000000478] [Medline: 31033741]
31. Holliday R, Borges LM, Stearns-Yoder KA, Hoffberg AS, Brenner LA, Monteith LL. Posttraumatic stress disorder, suicidal ideation, and suicidal self-directed violence among U.S. military personnel and veterans: a systematic review of the literature from 2010 to 2018. Front Psychol 2020;11:1998 [FREE Full text] [doi: 10.3389/fpsyg.2020.01998] [Medline: 32982838]

32. Patrick AR, Miller M, Barber CW, Wang PS, Canning CF, Schneeweiss S. Identification of hospitalizations for intentional self-harm when E-codes are incompletely recorded. Pharmacoepidemiol Drug Saf 2010 Dec;19(12):1263-1275 [FREE Full text] [doi: 10.1002/pds.2037] [Medline: 20922709]
33. Goldstein BA, Bhavsar NA, Phelan M, Pencina MJ. Controlling for informed presence bias due to the number of health encounters in an electronic health record. Am J Epidemiol 2016 Dec 01;184(11):847-855 [FREE Full text] [doi: 10.1093/aje/kww112] [Medline: 27852603]
34. Charlson ME, Pompei P, Ales KL, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
35. Bryant RA. Posttraumatic stress disorder and traumatic brain injury: can they co-exist? Clin Psychol Rev 2001 Aug;21(6):931-948. [doi: 10.1016/s0272-7358(00)00074-x] [Medline: 11497213]
36. Tsai J, Whealin JM, Scott JC, Harpaz-Rotem I, Pietrzak RH. Examining the relation between combat-related concussion, a novel 5-factor model of posttraumatic stress symptoms, and health-related quality of life in Iraq and Afghanistan veterans. J Clin Psychiatry 2012 Aug;73(8):1110-1118. [doi: 10.4088/JCP.11m07587] [Medline: 22781019]
37. Lippa SM, Fonda JR, Fortier CB, Amick MA, Kenna A, Milberg WP, et al. Deployment-related psychiatric and behavioral conditions and their association with functional disability in OEF/OIF/OND veterans. J Trauma Stress 2015 Feb;28(1):25-33 [FREE Full text] [doi: 10.1002/jts.21979] [Medline: 25703936]
38. Outcalt SD, Kroenke K, Krebs EE, Chumbler NR, Wu J, Yu Z, et al. Chronic pain and comorbid mental health conditions: independent associations of posttraumatic stress disorder and depression with pain, disability, and quality of life. J Behav Med 2015 Jun;38(3):535-543. [doi: 10.1007/s10865-015-9628-3] [Medline: 25786741]
39. Davis LL, Schein J, Cloutier M, Gagnon-Sanschagrin P, Maitland J, Urganus A, et al. The economic burden of posttraumatic stress disorder in the United States from a societal perspective. J Clin Psychiatry 2022 Apr 25;83(3):21m14116 [FREE Full text] [doi: 10.4088/JCP.21m14116] [Medline: 35485933]
40. Sayer NA, Clothier B, Spoont M, Nelson DB. Use of mental health treatment among veterans filing claims for posttraumatic stress disorder. J Trauma Stress 2007 Feb;20(1):15-25. [doi: 10.1002/jts.20182] [Medline: 17345650]
41. Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):629-640. [doi: 10.1001/archpsyc.62.6.629] [Medline: 15939840]
42. Hundt NE, Mott JM, Miles SR, Arney J, Cully JA, Stanley MA. Veterans' perspectives on initiating evidence-based psychotherapy for posttraumatic stress disorder. Psychol Trauma 2015 Nov;7(6):539-546. [doi: 10.1037/tra0000035] [Medline: 25915648]
43. Possemato K, Wray LO, Johnson E, Webster B, Beehler GP. Facilitators and barriers to seeking mental health care among primary care veterans with posttraumatic stress disorder. J Trauma Stress 2018 Oct;31(5):742-752. [doi: 10.1002/jts.22327] [Medline: 30338576]
44. Iversen AC, van Staden L, Hughes JH, Browne T, Greenberg N, Hotopf M, et al. Help-seeking and receipt of treatment among UK service personnel. Br J Psychiatry 2010 Aug;197(2):149-155. [doi: 10.1192/bjp.bp.109.075762] [Medline: 20679269]
45. Stecker T, Shiner B, Watts BV, Jones M, Conner KR. Treatment-seeking barriers for veterans of the Iraq and Afghanistan conflicts who screen positive for PTSD. Psychiatr Serv 2013 Mar 01;64(3):280-283. [doi: 10.1176/appi.ps.001372012] [Medline: 23450385]
46. Schottenbauer MA, Glass CR, Arnkoff DB, Tendick V, Gray SH. Nonresponse and dropout rates in outcome studies on PTSD: review and methodological considerations. Psychiatry 2008;71(2):134-168. [doi: 10.1521/psyc.2008.71.2.134] [Medline: 18573035]
47. Hoge CW, Grossman SH, Auchterlonie JL, Riviere LA, Milliken CS, Wilk JE. PTSD treatment for soldiers after combat deployment: low utilization of mental health care and reasons for dropout. Psychiatr Serv 2014 Aug 01;65(8):997-1004. [doi: 10.1176/appi.ps.201300307] [Medline: 24788253]
48. Elliott JE, Balba NM, McBride AA, Callahan ML, Street KT, Butler MP, et al. Different methods for traumatic brain injury diagnosis influence presence and symptoms of post-concussive syndrome in United States veterans. J Neurotrauma 2021 Nov 15;38(22):3126-3136 [FREE Full text] [doi: 10.1089/neu.2021.0031] [Medline: 34382417]
49. Meterko M, Baker E, Stolzmann KL, Hendricks AM, Cicerone KD, Lew HL. Psychometric assessment of the Neurobehavioral Symptom Inventory-22: the structure of persistent postconcussive symptoms following deployment-related mild traumatic brain injury among veterans. J Head Trauma Rehabil 2012;27(1):55-62. [doi: 10.1097/HTR.0b013e318230fb17] [Medline: 22190009]
50. Kelley E, Sullivan C, Loughlin JK, Hutson L, Dahdah MN, Long MK, et al. Self-awareness and neurobehavioral outcomes, 5 years or more after moderate to severe brain injury. J Head Trauma Rehabil 2014;29(2):147-152. [doi: 10.1097/HTR.0b013e31826db6b9] [Medline: 23249770]
51. Meulenbroek P, Ness B, Lemoncello R, Byom L, MacDonald S, O'Neil-Pirozzi TM, et al. Social communication following traumatic brain injury part 2: identifying effective treatment ingredients. Int J Speech Lang Pathol 2019 Apr;21(2):128-142. [doi: 10.1080/17549507.2019.1583281] [Medline: 30955383]

52. Seedat S, Lochner C, Vythilingum B, Stein DJ. Disability and quality of life in post-traumatic stress disorder: impact of drug treatment. Pharmacoeconomics 2006;24(10):989-998. [doi: 10.2165/00019053-200624100-00006] [Medline: 17002481]
53. Sayer NA, Friedemann-Sanchez G, Spoont M, Murdoch M, Parker LE, Chiros C, et al. A qualitative study of determinants of PTSD treatment initiation in veterans. Psychiatry 2009;72(3):238-255. [doi: 10.1521/psyc.2009.72.3.238] [Medline: 19821647]
54. Massarweh NN, Itani KMF, Morris MS. The VA MISSION Act and the future of veterans' access to quality health care. JAMA 2020 Jul 28;324(4):343-344. [doi: 10.1001/jama.2020.4505] [Medline: 32602896]
55. McIntire KL, Crawford KM, Perrin PB, Sestak JL, Aman K, Walter LA, et al. Factors increasing risk of suicide after traumatic brain injury: a state-of-the-science review of military and civilian studies. Brain Inj 2021 Jan 18;35(2):151-163. [doi: 10.1080/02699052.2020.1861656] [Medline: 33460350]
56. Carlson KF, Barnes JE, Hagel EM, Taylor BC, Cifu DX, Sayer NA. Sensitivity and specificity of traumatic brain injury diagnosis codes in United States Department of Veterans Affairs administrative data. Brain Inj 2013 Jun;27(6):640-650. [doi: 10.3109/02699052.2013.771795] [Medline: 23514276]

Abbreviations

AOR: adjusted odds ratio
CCI: Charlson Comorbidity Index
CDW: Corporate Data Warehouse
ICD-10: International Classification of Diseases, Tenth Revision
ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification
OR: odds ratio
PTSD: posttraumatic stress disorder
REACH VET: Recovery Engagement and Coordination for Health–Veterans Enhanced Treatment
TBI: traumatic brain injury
VA: Veterans Affairs
VHA: Veterans Health Administration
VINCI: VA Informatics and Computing Infrastructure

Edited by Y Khader; submitted 19.09.22; peer-reviewed by S Bannon, A AL-Asadi; comments to author 31.12.22; revised version received 06.03.23; accepted 12.04.23; published 24.07.23.

Please cite as:
Rawat BPS, Reisman J, Pogoda TK, Liu W, Rongali S, Aseltine Jr RH, Chen K, Tsai J, Berlowitz D, Yu H, Carlson KF
Intentional Self-Harm Among US Veterans With Traumatic Brain Injury or Posttraumatic Stress Disorder: Retrospective Cohort Study From 2008 to 2017
JMIR Public Health Surveill 2023;9:e42803
URL: https://publichealth.jmir.org/2023/1/e42803
doi:10.2196/42803
PMID:37486751

Original Paper

Depression as a Mediator and Social Participation as a Moderator in the Bidirectional Relationship Between Sleep Disorders and Pain: Dynamic Cohort Study

Si Fan[1], MA; Qianning Wang[1], MA; Feiyang Zheng[1], PhD; Yuanyang Wu[1], PhD; Tiantian Yu[1], MA; Yanting Wang[1], MA; Xinping Zhang[1*], PhD; Dexing Zhang[2*], PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[2]School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, Hong Kong
[*]these authors contributed equally

Corresponding Author:
Xinping Zhang, PhD
School of Medicine and Health Management
Tongji Medical College
Huazhong University of Science and Technology
Hangkong Road 13
Wuhan, Hubei
China
Phone: 86 18062643970
Email: xpzhang602@hust.edu.cn

Abstract

Background: Chronic pain, sleep disorders, and depression are major global health concerns. Recent studies have revealed a strong link between sleep disorders and pain, and each of them is bidirectionally correlated with depressive symptoms, suggesting a complex relationship between these conditions. Social participation has been identified as a potential moderator in this complex relationship, with implications for treatment. However, the complex interplay among sleep disorders, pain, depressive symptoms, and social participation in middle- and old-aged Asians remains unclear.

Objective: This study aimed to examine the bidirectional relationship between sleep disorders and pain in middle- and old-aged Chinese and measure the role of depression as a mediator and social participation as a moderator in this bidirectional relationship through a dynamic cohort study.

Methods: We used data from the China Health and Retirement Longitudinal Study across 5 years and included a total of 7998 middle- and old-aged people (≥45 years old) with complete data in 2011 (T1), 2015 (T2), and 2018 (T3). The cross-lag model was used to assess the interplay among sleep disorders, pain, depressive symptoms, and social participation. Depressive symptoms were assessed by the 10-item Centre for Epidemiological Studies Depression scale. Sleep disorders were assessed by a single-item sleep quality scale and nighttime sleep duration. The pain score was the sum of all pain locations reported. Social participation was measured using self-reported activity.

Results: Our results showed significant cross-lagged effects of previous sleep disorders on subsequent pain at T2 (β=.141; P<.001) and T3 (β=.117; P<.001) and previous pain on subsequent poor sleep at T2 (β=.080; P<.001) and T3 (β=.093; P<.001). The indirect effects of previous sleep disorders on pain through depressive symptoms (β=.020; SE 0.004; P<.001; effect size 21.98%), as well as previous pain on sleep disorders through depressive symptoms (β=.012; SE 0.002; P<.001; effect size 20.69%), were significant across the 3 time intervals. Among participants with high levels of social participation, there were no statistically significant effects of previous sleep disorders on subsequent pain at T2 (β=.048; P=.15) and T3 (β=.085; P=.02), nor were there statistically significant effects of previous pain on subsequent sleep disorders at T2 (β=.037; P=.15) and T3 (β=.039; P=.24). Additionally, the mediating effects of depressive symptoms on the sleep disorders-to-pain pathway (P=.14) and the pain-to-sleep disorders pathway (P=.02) were no longer statistically significant.

Conclusions: There is a bidirectional relationship between sleep disorders and pain in middle- and old-aged Asians; depression plays a longitudinal mediating role in the bidirectional relationship between them; and social participation moderates the bidirectional relationship between them directly and indirectly by affecting depression. Future interventions may consider the complex relationship between these conditions and adopt a comprehensive treatment regime.

(JMIR Public Health Surveill 2023;9:e48032) doi:10.2196/48032

KEYWORDS

depression; dynamic cohort; longitudinal mediation; pain; sleep disorders; social participation

Introduction

Chronic pain, sleep disorders, and depression are prevalent worldwide and lead to a serious disease burden. For instance, chronic pain affects hundreds of millions of people globally, with estimates of its prevalence ranging from 10% to 40% in different countries [1-4]. It also attracted widespread attention as a global health issue. For example, Healthy People 2030 has put reducing the proportion of adults with high-impact chronic pain as one of its main goals in the United States [5]. Additionally, sleep disorders, particularly insomnia, affect over 300 million people just in China, and of those, about 30% experience one or more symptoms of insomnia, seriously affecting individuals' physical, emotional, social, and occupational functions, and those health risks are on the rise [6]. Especially during the COVID-19 outbreak, the prevalence of insomnia increased significantly [7]. Third, depression ranks fourth in disease burden worldwide, with nearly 350 million people affected by it, which can lead to a 12.7% mortality rate [8], and other health problems often caused by depression, such as type 2 diabetes [9], cardiovascular disease [10], and suicide [11], with great burdens on families and communities [8]. Overall, chronic pain, sleep disorders, and depression often coexist and are prevalent in the general population, posing more serious health risks [12,13].

A strong link between sleep disorders and pain has been examined first in the past 2 decades. For instance, research indicates that at least 50% of insomniacs also have chronic pain [14]. Additionally, experimental studies on healthy participants have also suggested that pain levels increase after sleep restriction or interruption [15,16]. Conversely, pain has also been found to be a strong predictor of sleep disorders. For example, individuals with chronic pain often experience poor sleep quality [17,18], and it has been estimated that chronic pain patients are 18 times more likely to reach the clinical diagnostic criteria for insomnia than those without pain [19].

Second, studies have found that depressive symptoms might be bidirectionally correlated with sleep disorders and pain, respectively. Specifically, the link between depressive symptoms and sleep disorders might be bidirectional. Sleep disorders are related to maladaptive coping mechanisms that can lead to depression, especially in chronic pain patients [20]. Conversely, depressive symptoms may potentially induce sleep disorders, as evidenced by the fact that sleep disorders are the most common residual symptoms after the remission of a depressive episode [21,22]. Furthermore, there is a relationship between depressive symptoms and pain, which may be mutually reinforcing, leading to a vicious circle between them. Depression and pain symptoms follow the same descending pathways of the central nervous system. Depression may reduce the regulatory role of the descending pain system and amplify pain perception by depleting 5-hydroxytryptamine and norepinephrine [23]. Chronic pain can also induce depression through various molecular mechanisms, such as midbrain dopamine, inflammatory factors, and epigenetic modifications [24]. Thus, depressive symptoms might play a mediating effect in considering the relationship between sleep disorders and pain.

However, there have been limited high-quality studies using large-scale longitudinal data to explore the trilateral relationship among depression, sleep disorders, and pain because the existing research on the relationship between them has mainly been based on cross-sectional designs [25,26], especially in the Asian population [27]. Given the differences in geographical environment, lifestyle, race, and cultural affiliation, the prevalence of sleep deficiency [28,29] and the perception of pain [30,31] as well as the willingness to seek support for depression [32,33] may differ between older people in Asia and Western countries. Therefore, the results of Western studies may not be generalizable to Asian populations, and longitudinal studies in Asia are desperately needed to answer the question of whether sleep disorders are causally related to pain (or vice versa) and to explore the potential mediating effect of depressive symptoms on this relationship.

In addition, the primary therapeutic regimes on pain are meeting a huge challenge, and society around the globe is seeking aggressive complementary treatments. First, chronic pain might be best considered a disease. Despite clinical trials and guidelines recommending personalized, multimodal, interdisciplinary treatment approaches, chronic pain remains active [34]. Second, the primary therapeutic guidelines for sleep disorders are cognitive behavioral therapy and pharmacological treatment [35]. However, meta-analyses have shown that although cognitive behavioral therapy can improve subjective sleep quality [36], it does not improve sleep parameters defined by polysomnography [37]. Furthermore, cognitive behavioral therapy is often not available due to the limited number of trained professionals. And patients often prefer nonpharmacological treatments due to the unsatisfactory benefit-risk ratio of hypnotic drugs [38-40]. Third, it is difficult to unravel unique causal pathways due to the potential bidirectional effects between pain and sleep disorders. Measures to improve sleep may be a useful addition to pain management plans [41], which means that combined treatment of sleep disorders and chronic pain may be better than single treatment. Given that, multidisciplinary interventions that attempt to interrupt the possible vicious cycle between sleep disorders and pain may be helpful. Therefore, new explorations are needed to delve into novel approaches with better accessibility and economic effectiveness to moderate the bidirectional pathways of sleep disorders and pain.

Social participation may be a potential moderator. Previous literature suggested the potential benefits of social participation in alleviating depression, pain, and sleep disorders. Social interaction and communication during social participation may motivate mutual support, provide a sense of belonging to an individual, and to a large extent reduce social isolation [42,43], which therefore may prevent depression [44]. The biopsychosocial model of pain hypothesizes that pain is influenced by biological, psychological, and social factors. Satisfaction with one's social role and self-perceived ability to engage in social roles during social participation are related to pain [45]. Berkman's social relations model [46] suggests that social participation, as a downstream factor of social relations, can have a positive effect on mental health and health behaviors to promote sleep [47]. Finally, this study hypothesizes that social participation, as an accessible and cost-effective method, can moderate the bidirectional pathways of sleep disorders and pain through direct and indirect pathways.

Therefore, three main objectives of this dynamic cohort study were proposed: (1) to explore the bidirectional relationship between sleep disorders and pain; (2) to explore the longitudinal mediating role of depressive symptoms on the bidirectional relationship between sleep disorders and pain given the potentially important role of depressive symptoms; and (3) to explore the moderating effect of social participation on the bidirectional relationship between sleep disorders and pain. This study aims to provide higher levels of evidence regarding the above objectives based on a national prospective cohort study, contributing to a better understanding of the complex interplay between sleep disorders, pain, depressive symptoms, and social participation.

Methods

Sample and Data Collection

The data used in this study were obtained from the follow-up survey of the China Health and Retirement Longitudinal Study (CHARLS). The participants are middle- and old-aged residents. The baseline data collection for CHARLS was conducted in 2011 and included approximately 10,000 households and 17,500 individuals from 150 counties and districts and 450 villages and resident committees in 28 provinces. The follow-up surveys were conducted every 2 to 3 years, and all survey data were made public 1 year after the end of data collection.

CHARLS adopted a multistage stratified probability proportional to size sampling technique, and the detailed sampling technique can be found elsewhere [48]. All data were collected using face-to-face interviews at 3 time points, covering participants' demographic background, family characteristics, health behavior and status, and retirement information. More detailed information about CHARLS can be found on its official website [49].

This study used the follow-up data from 2013 (T1), 2015 (T2), and 2018 (T3). The sample size is 7998 participants who are selected with a criterion of providing complete information on sleep disorders, pain, depression, and social participation in all 3 waves.

Ethics Approval

The studies involving human participants were reviewed and approved by the Research Ethics Committees of Peking University (IRB00001052-11015). The patients or participants provided their written informed consent to participate in this study. The data for analysis were deidentified without revealing any personal identity.

Depression

The 10-item Center for Epidemiologic Studies Depression Scale (CES-D 10) was used to measure depressive symptoms. Covering the respondent's positive feelings, negative emotions, and physical symptoms in the past week. The item of CES-D 10 was scaled from 0 to 3, with the total score ranging from 0 to 30. Higher scores indicate more depressive symptoms. CES-D 10 has shown good internal reliability, with a Cronbach α of 0.815 [50].

Sleep Disorders

Sleep disorders were assessed by a single-item sleep quality scale and nighttime sleep duration, which were both commonly used in previous studies [51-53]. The total score for sleep disorders was calculated as the sum of the scores for the single-item sleep quality scale and nighttime sleep duration.

Participants were asked to rate their sleep quality using the single-item sleep quality scale. The single item in the scale was "My sleep was restless," which had a 4-point response scale ranging from "rarely or none of the time" to "most or all of the time." The self-reported measure of sleep quality as a single item has been widely used in previous studies and has been shown to be highly correlated with the multi-item versions of the sleep quality scale [51]. Furthermore, it has been reported that the test-retest reliability coefficients of the single-item sleep quality scale reached 0.90 [54].

Nighttime sleep duration was assessed based on self-reported data in response to the question, "How many hours did you usually sleep per day on average in the past month?" According to the classification on sleep duration from the National Sleep Foundation [55] and the research from Gu et al [56], the responses were divided into 4 groups based on the number of hours slept per night and scored from 0 to 3. A score of 3 was assigned to participants who reported sleeping less than 5 hours or more than 10 hours per night; a score of 2 was assigned to those who reported sleeping between 5 and 6 hours or between 9 and 10 hours per night; a score of 1 was assigned to those who reported sleeping between 6 and 7 hours or between 8 and 9 hours per night; and a score of 0 was assigned to those who reported sleeping between 7 and 8 hours per night.

Pain

In this study, participants were asked to report whether they experienced pain. Those who did not report experiencing pain were assigned a value of 0, while those who reported experiencing pain were asked to report their pain location using a validated body map [57], including the head, shoulders, arms, wrists, fingers, chest, stomach, back, waist, buttocks, legs, knees, ankles, toes, neck, and other parts. Each pain location was counted as 1 point, and the total pain score was the sum of all pain locations reported. These methods to determine the extent and distribution of pain are standard in population-based studies of pain and have been shown to be valid and reliable [58].

Social Participation

Combining the definitions of social participation by previous scholars [59] and the characteristics of activity participation of Chinese seniors [60], this study adapted the definition of social participation from previous literature that emphasizes participation in activities that provide interactions with others in society or the community. Based on this adapted definition, the following types of activities were included: (1) interacting with friends; (2) playing mahjong, chess, cards, or going to the community club; (3) providing help to relatives, friends, or neighbors who do not live with you; (4) dancing, fitness, qigong, etc; (5) participating in club activities; (6) participating in volunteer activities or charity activities; (7) taking care of patients or disabled people who do not live with you; (8) going to school or attending training courses; (9) stock trading; (10) surfing the internet; and (11) other social participation that meets our definition. One social participation counted as 1 point. If the respondent answered "yes" to any of the above social participation, they were asked about the corresponding frequency (3=almost every day, 2=almost every week, and 1=not often). The total score of social participation was the cumulative score of the corresponding frequency of each individual social participation.

Control Variable

To minimize the possibility of a third variable influence on the sleep disorders-pain relationship and to keep the parsimony of our model, a limited number of covariates related to sleep and pain were controlled. According to previous studies, all covariates are based on baseline data [61]. First, demographic control variables included age, gender (coded as 0=male and 1=female), and educational level (coded as 0=illiterate, 1=junior high school or below, and 2=high school or above). Then, lifestyle covariates included current smoking and current drinking status. The respondents were asked if they currently smoke (coded as 1=yes and 0=no) and drink (coded as 1=yes and 0=no). Interpersonal relationship covariates included marital status (coded as 1=married and 0=never married, separated, divorced, or widowed) and residence (coded as 1=rural and 0=urban). In addition, the covariates closely related to pain included "have measures been taken to relieve pain?" (coded as 1=yes and 0=no).

Statistical Analysis

In order to maximize statistical power and minimize bias, full-information maximum likelihood estimation was conducted in Mplus (version 8.3; Muthén & Muthén). This estimation method is referred to as principled missing data handling approach, as it does not directly replace missing values but uses existing information to estimate population parameters without bias [62], and it is considered suitable for attributing missing longitudinal data, such as the data analyzed here.

The following variables were inputted into the model as continuous variables: sleep disorders, pain, depressive symptoms, social participation, and age, while the remaining variables were inputted as categorical variables. Structural equation modeling with observed variables was done to evaluate the association between sleep disorders and pain at all time points (model 1), the mediating role of depressive symptoms (model 2), and the moderating role of social participation (model 3). Thus, a full mediation structural equation model was constructed [63]. Following the suggestion of Hoyle [64], 2 fit indices were used to evaluate the model fit, including the standardized root mean square residual (SRMR) and the comparative fit index (CFI). A CFI value greater than 0.95 was considered to indicate a good model fit, while an SRMR value below 0.06 was considered to indicate a good fit, although values below 0.08 were considered to be acceptable. All mediations were tested with 5000 bootstrap iterations.

Results

Descriptive Statistics

The sample characteristics based on baseline data and the distribution of key variables by period are presented in Table 1. The average scores of key variables from T1 to T3 were as follows: sleep disorders (2.18, 2.20, and 2.45); depressive symptoms (7.89, 8.07, and 8.87); pain (0.94, 1.73, and 3.10); and social participation (2.01, 1.89, and 1.83), respectively.

Table 1. The characteristics of covariates at baseline and depressive symptoms, sleep disorders, pain, and social participation from T1 to T3 (N=7998).

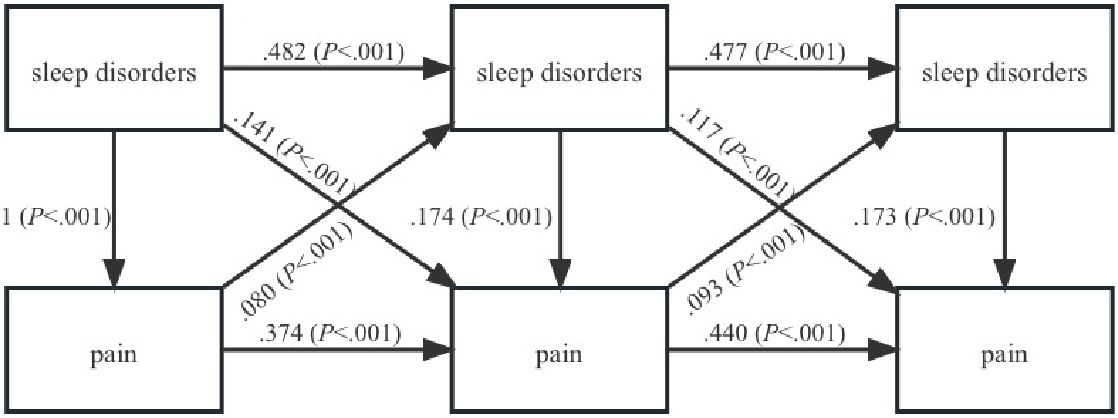
Variable	2013 (T1)	2015 (T2)	2018 (T3)
Age (years), mean (SD)	58.47 (8.54)	—[a]	—
Sex, n (%)			
Female	3865 (48.32)	—	—
Male	4133 (51.68)	—	—
Education, n (%)			
Illiteracy	1738 (21.73)	—	—
Junior high school or less	5172 (64.67)	—	—
High school or more	1088 (13.6)	—	—
Marital status, n (%)			
Married	7258 (90.75)	—	—
Not in marriage	740 (9.25)	—	—
Residence, n (%)			
Urban	2991 (37.4)	—	—
Rural	5007 (62.6)	—	—
Smoke, n (%)			
Yes	2176 (27.21)	—	—
No	5822 (72.79)	—	—
Drink, n (%)			
Yes	2706 (33.83)	—	—
No	5292 (66.17)	—	—
Taken measures to relieve pain, n (%)			
Yes	1508 (18.85)	—	—
No	6490 (81.15)	—	—
Pain, mean (SD)	0.94 (2.05）	1.73 (3.41）	3.10 (3.79）
Depressive symptoms, mean (SD)	7.89 (5.75）	8.07 (6.38）	8.87 (6.52）
Sleep disorders, mean (SD)	2.18 (1.93）	2.20 (1.95）	2.45 (1.98）
Social participation, mean (SD)	2.01 (2.37)	1.89 (2.41)	1.83 (2.38)

[a] Not available.

The Relationship Between Sleep Disorders and Pain

Figure 1 depicts the cross-lagged model of the bidirectional relationship between sleep disorders and pain (model 1). After controlling for covariates, model 1 fitted the data adequately (CFI=0.919; SRMR=0.045). Sleep disorders at each time point were positively related to themselves over time, as was pain, as would be expected. The cross-lagged effects of previous sleep disorders on pain at T2 (β=.141; *P*<.001) and T3 (β=.117; *P*<.001) as well as pain on subsequent sleep disorders at T2 (β=.080; *P*<.001) and T3 (β=.093; *P*<.001) were significant. These results support our hypothesis that sleep disorders and pain are bidirectionally related.

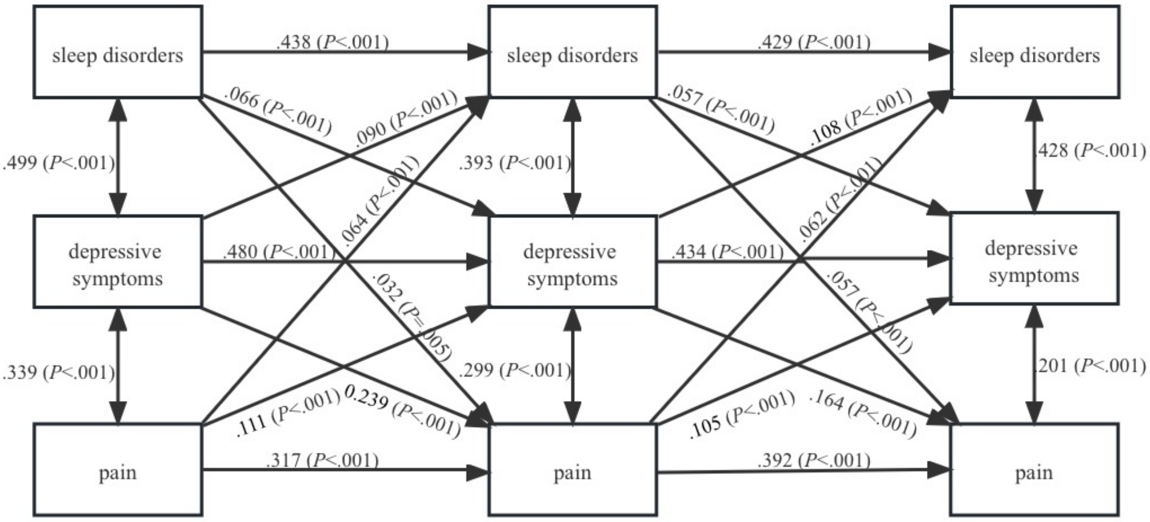
Figure 1. Cross-lagged panel model of the bidirectional relationship between sleep disorders and pain.



The Longitudinal Mediating Role of Depressive Symptoms

As Figure 2 presents, after adding 2 indirect paths that shared depressive symptoms as a potential mediator, model 2 still had a good fit to the data (CFI=0.983; SRMR=0.045). At a significance level of 0.01, we found that the cross-lagged effects of sleep disorders on subsequent pain at T2 (β=.032; *P*=.005) and T3 (β=.057; *P*<.001), as well as previous pain on subsequent sleep disorders at T2 (β=.064; *P*<.001) and T3 (β=.062; *P*<.001), were significant but slightly reduced in size compared to those in model 1.

Figure 2. Longitudinal mediating effect of depressive symptoms in the bidirectional relationship between sleep disorders and pain.



We found 2 specific indirect paths from the depressive symptoms. First, higher previous sleep disorders predicted worse subsequent depressive symptoms (β=.066; *P*<.001), and worse previous depressive symptoms predicted higher amounts of subsequent pain (β=.164; *P*<.001). Second, higher previous pain predicted worse subsequent depressive symptoms (β=.111; *P*<.001), and worse previous depressive symptoms predicted higher amounts of subsequent sleep disorders (β=.108; *P*<.001). To further test the longitudinal mediation, we used a 95% CI in Mplus (Table 2). Effect sizes of the longitudinal mediations were computed using MacKinnon's formula for calculating the mediated percentage, which is the indirect effect divided by the total effect (Table 2). The indirect effects of previous sleep disorders on pain through depressive symptoms (β=.020; SE 0.004; *P*<.001; effect size 21.98%), as well as previous pain on sleep disorder through depressive symptoms (β=.012; SE 0.002; *P*<.001; effect size 20.69%), were significant across the 3-time intervals. Both effect sizes fell in the medium range on the basis of Cohen's guidelines, suggesting that these mediations captured significant covariation over time.

Table 2. Statistical results of longitudinal mediating effect of depressive symptoms in the bidirectional relationship between sleep disorders and pain.

Model 2	β	SE	*P* value	Effect size (%)[a]
T1 sleep disorders to T2 depressive symptoms to T3 pain	.020	0.004	<.001	21.98
T1 pain to T2 depressive symptoms to T3 sleep disorders	.012	0.002	<.001	20.69

[a] Effect size is the proportion mediated, which is calculated by dividing the indirect effect by the total effect.

The Moderating Effect of Social Participation

To test our third hypothesis, we examined whether social participation could moderate the bidirectional relationship between sleep disorders and pain in model 3. The results of model 3 are shown in Figure 3 and Table 3 and fit the data well (CFI=0.983; SRMR=0.045). We found support for this prediction in participants with high levels of social participation (greater than the mean). As shown in Figure 3, social participation directly moderates the bidirectional relationship between sleep disorders and pain. This was demonstrated by the lack of significant cross-lagged effects between sleep disorders and pain in participants with high levels of social participation, at a significance level of .01. Specifically, we did not observe statistically significant effects of sleep disorders on subsequent pain at T2 (β=.048; *P*=.15) and T3 (β=.085; *P*=.02), nor did we observe statistically significant effects of previous pain on subsequent sleep disorders at T2 (β=.037; *P*=.15) and T3 (β=.039; *P*=.24).

Figure 3. Moderating effect of social participation in the relationship among sleep disorders, pain, and depressive symptoms.

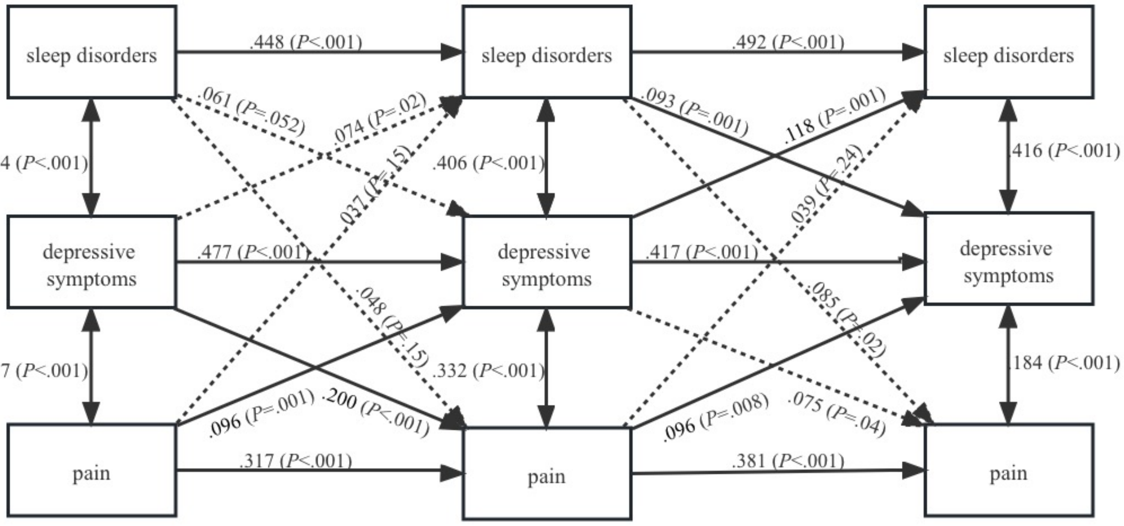


Table 3. Statistical results of longitudinal mediating effect of depressive symptoms in the bidirectional relationship between sleep disorders and pain in participants with high levels of social participation.

Model 3	β	SE	*P* value	Effect size (%)[a]
T1 sleep disorders to T2 depressive symptoms to T3 pain	.010	0.007	.14	4.95
T1 pain to T2 depressive symptoms to T3 sleep disorders	.016	0.007	.02	28.57

[a] Effect size is the proportion mediated, which is calculated by dividing the indirect effect by the total effect.

Using a 95% CI in Mplus (Table 3), we found that depressive symptoms no longer mediated the sleep disorders-to-pain pathway (*P*=.14) and the pain-to-sleep disorders pathway (*P*=.02) in participants with high levels of social participation. This indicated that social participation indirectly moderated the bidirectional relationship between sleep disorders and pain by affecting depression.

Discussion

Principal Findings

In summary, the dynamic cohort study based on the longitudinal data in middle- and old-aged Asians demonstrates a reciprocal relationship between sleep disorders and pain, and the prospective reciprocal relationship is partially mediated by depressive symptoms. The reciprocal relationship between sleep disorders and pain is also moderated by social participation directly and indirectly by affecting depression.

Previous research already showed that sleep disorders, pain, and depressive symptoms are cross-sectionally related. Although there are longitudinal studies and even a randomized controlled study [65] in Western countries, they only suggest that sleep disorders and pain interact in a bidirectional manner. Despite some cross-sectional studies demonstrating the indirect path from pain to sleep disorders through depression symptoms [26] and the reverse influence [25], the data are cross-sectional and correlational, meaning that causality cannot be inferred and firm conclusions about mediation effects cannot be drawn. To our knowledge, there has been 1 cohort study [66], but depressive symptoms were assessed during the same wave as sleep disorders, making it uncertain that they were indeed a consequence of sleep disorders. Our study explores the basic mechanism of the mutual relationship between depression, sleep disorders, and pain and provides higher-level evidence by large-scale longitudinal data. Our findings not only strengthen the bidirectional relationship between sleep disorders and pain and the longitudinal mediating effect of depression symptoms in this bidirectional relationship, but also explore the social participation as an accessible and feasible moderation measure to interrupt the vicious cycle of sleep disorders and pain directly and indirectly by affecting depression, which will provide inspiration for the treatment of sleep disorders and pain. Our study contributes to the literature by examining the prospective relationship between sleep disorders, pain, depression, and social participation in a large-scale cohort study. The advantages of our study include large sample size, prospective design, and nationally representative data [67].

Collectively, our findings from a large cohort of middle- and old-aged Chinese population support the bidirectional relationship between sleep disorders and pain, with a stronger effect of sleep disorders on pain at T2 (β=.141; *P*<.001) and T3 (β=.117; *P*<.001) compared with the effect of pain on sleep disorders at T2 (β=.080; *P*<.001) and T3 (β=.093; *P*<.001). These findings are in accordance with previous research conducted in Western countries. For instance, a study found that poor baseline sleep quality predicted higher pain intensity at 6 months (β=.18; 95% CI 0.07-0.30), and vice versa (β=.14; 95% CI 0.01-0.26) [68]. Additionally, a randomized controlled trial revealed that changes in sleep complaints at 3 months significantly predicted changes in pain at 12 months (β=.29; *P*<.001), and to a lesser extent, changes in pain predicted changes in sleep (β=.15; *P*<.05) [65], which is consistent with our study in terms of the direction and strength of the bidirectional relationship between sleep disorders and pain. However, a study of Western emerging adults (18-25 years old) partly argued against the long-term bidirectional relationship, suggesting that sleep disorders significantly increased chronic pain (β=.15; *P*<.05) [66]. In comparison, we further found a slightly weaker but significant longitudinal effect of pain on sleep disorders, which may be due to the different age stages and races of the population.

There is evidence that emotional disturbances and negative emotions such as depression can largely explain the relationship between pain and sleep disorders. For instance, a study found that depression partially mediated the effects of pain, substantiating an indirect path from pain to sleep disturbance through depressive symptoms, and this indirect effect (β=.20; SE 0.07) accounted for 37.46% of the total effects in the model tested [26]. Another study showed that sleep disorders have a strong positive association with depressive symptoms (β=.67; 95% CI 0.57-0.75; *P*<.001), which in turn has a moderate positive association with pain severity (β=.36; 95% CI 0.19-0.51; *P*<.001), and the mediating effect of depressive symptoms was significant (β=.24; SE 0.052) [25]. In addition, a cohort study found that sleep disorders indirectly affect the severity of abdominal pain through depressive symptoms (β=.05; 95% CI 0.01-0.09; k^2=0.06) [66]. In agreement with these studies, our results support the mediating role for depression symptoms in the bidirectional relationship between sleep disorders and pain, confirming an indirect pathway from pain to sleep disorders through depressive symptoms and the reverse influence. Therefore, previous findings suggesting a causal link between sleep disorders and pain may be, in part, due to the effects of depressive symptoms. Our study contributes to the current understanding of the relationship between sleep disorders, pain, and depression in the Asian population, providing valuable insights where there was previously a gap in knowledge.

In the middle- and old-aged population with high levels of social participation, we identified that social participation moderated the bidirectional relationship between sleep disorders and pain directly and indirectly by affecting depression. This suggests that for individuals with pain, higher levels of social participation are associated with a reduced likelihood of experiencing sleep disorders. Similarly, for individuals with sleep disorders, higher levels of social participation are associated with a reduced likelihood of experiencing pain. This is a new finding, as previous related studies only suggested that social participation has potential benefits in alleviating depression, pain, and sleep disorders. For instance, social participation has a significant negative effect on the depressive symptoms of the older adults, whether they are empty-nest older adults (β=–2.404; *P*<.001) or non–empty-nest older adults (β=–1.957; *P*<.001) [44]. Additionally, high-frequency social participation, such as participating in interest groups (β=–.59; *P*<.01), community-related organizations (β=–.41; *P*<.05), sports (β=–.39; *P*<.05) can have a positive effect on mental health and health behaviors to promote sleep [47]. Moreover, satisfaction with their social roles during social participation (β=–.16; *P*<.01) and self-perceived ability to participate in social roles are related to pain (β=–.19; *P*<.001) [45]. However, there is little systematic study on the role of social participation in the cyclical relationship between sleep disorders, depression, and pain. This study reveals that social participation can serve as a practical and cost-effective strategy to moderate the bidirectional relationship between sleep disorders and pain directly and indirectly by affecting depression, which has potential implications for their treatment. Additionally, social participation may also moderate the relationship between sleep disorders and pain by affecting circadian rhythms, as suggested by the literature. For example, social participation in activities such as dancing and fitness can modulate the molecular clock in skeletal muscle, affecting both the amplitude and phase of circadian rhythms [69]. Disrupted circadian rhythms have direct effects on both sleep disorders and pain, as multiple clinical and foundational science studies have reported that circadian rhythm disruption can directly alter pain thresholds [70]. Moreover, disrupted circadian rhythms can lead to undesirable or irregular timing of sleep, exacerbating sleep disorders [71].

Future research could investigate the effects of various types of social participation with different durations, frequencies, and timings on circadian rhythms. Future research could compare the impacts of overnight activities such as playing mahjong or caring for disabled family members with daytime activities like helping neighbors or engaging in volunteer work. Future research could also investigate the effects of various types of social participation with different durations, frequencies, and timings on the relationship. Such research could shed light on the intricate relationship between social participation, circadian rhythms, sleep disorders, and pain.

Limitations

Although this study indicates some interesting relationship, there are limitations to the interpretation of the results. First, the self-reported nature of the data means that results must be interpreted with caution. There may be a tendency to over- or underestimate symptoms. Second, data were not collected on specific pain diagnoses. Different potential diseases and causes may play a role in the relationship between pain and sleep, so it would be useful in the future to collect such data for subgroup analysis. Third, medical information such as the disease itself and other comorbidities, as well as the severity of symptoms and medication records, was not collected and therefore could not be controlled for in our analysis. For example, some analgesic compounds can lead to sleep fragmentation and even insomnia [72]. Future research should consider the effects of medications on sleep structure as well as emotional and physical states. Fourth, due to our inability to incorporate potential confounders into the analysis or adjust for them, such as stressful life events, family history of mental disorders or genetic characteristics, as well as family characteristics, health behaviors and status, and retirement information, residual confounding may exist. Future studies should consider these limitations and attempt to address them to provide a more comprehensive understanding of this complex relationship. In addition, categorizing risk factors for sleep, such as susceptibility, precipitating, and persistent factors, to better identify associated factors will provide more targeted recommendations for the management of pain, sleep disorders, and other symptoms.

Conclusions

In conclusion, our longitudinal cohort found a bidirectional relationship between sleep disorders and pain in middle- and old-aged Asians; depression plays a longitudinal mediating role in the bidirectional relationship between them; and social participation moderated the bidirectional relationship between them directly and indirectly by affecting depression. In future research, the moderation mechanism of social participation will be a focus with high interest, which will provide further insights into the nonpharmacological treatment and care of sleep disorders and pain. Overall, future interventions may consider the complex relationship of these conditions and adopt a comprehensive treatment regime.

Acknowledgments

This work was supported by the Hubei Liberal Arts Major and Cross-Cutting Project (grant 2017WKZDJC006). We would like to express our gratitude to the CHARLS research team and field team for their contributions to data collection. We would also like to thank all the interviewees who voluntarily participated in the CHARLS.

Data Availability

The CHARLS data set is publicly available on its official website [49]. The data analyzed in our study can be obtained from the corresponding author upon request.

Authors' Contributions

SF was responsible for the study design, data analysis, interpretation of the data, and writing of the manuscript. XZ contributed to the study design, data interpretation, manuscript revision, approval, and funding acquisition. DZ contributed to data interpretation, manuscript revision and approval. The other authors were responsible for data interpretation.

Conflicts of Interest

None declared.

References

1. Todd A, McNamara CL, Balaj M, Huijts T, Akhter N, Thomson K, et al. The European epidemic: pain prevalence and socioeconomic inequalities in pain across 19 European countries. Eur J Pain 2019 Sep;23(8):1425-1436 [FREE Full text] [doi: 10.1002/ejp.1409] [Medline: 31038816]
2. de Souza JB, Grossmann E, Perissinotti DMN, de Oliveira Junior JO, da Fonseca PRB, Posso IDP. Prevalence of chronic pain, treatments, perception, and interference on life activities: Brazilian population-based survey. Pain Res Manag 2017;2017:4643830 [FREE Full text] [doi: 10.1155/2017/4643830] [Medline: 29081680]

3. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open 2016 Jun 20;6(6):e010364 [FREE Full text] [doi: 10.1136/bmjopen-2015-010364] [Medline: 27324708]
4. Yong RJ, Mullins PM, Bhattacharyya N. Prevalence of chronic pain among adults in the United States. Pain 2022 Feb;163(2):e328-e332. [doi: 10.1097/j.pain.0000000000002291] [Medline: 33990113]
5. health.gov. Reduce the proportion of adults with chronic pain that frequently limits life or work activities - CP-01 - Healthy People 2030. URL: https://health.gov/healthypeople/objectives-and-data/browse-objectives/chronic-pain/reduce-proportion-adults-chronic-pain-frequently-limits-life-or-work-activities-cp-01 [accessed 2023-06-01]
6. Chung KF, Yeung WF, Ho FYY, Yung KP, Yu YM, Kwok CW. Cross-cultural and comparative epidemiology of insomnia: the Diagnostic and Statistical Manual (DSM), International Classification of Diseases (ICD) and International Classification of Sleep Disorders (ICSD). Sleep Med 2015 Apr;16(4):477-482. [doi: 10.1016/j.sleep.2014.10.018] [Medline: 25761665]
7. Li Y, Qin Q, Sun Q, Sanford LD, Vgontzas AN, Tang X. Insomnia and psychological reactions during the COVID-19 outbreak in China. J Clin Sleep Med 2020 Aug 15;16(8):1417-1418 [FREE Full text] [doi: 10.5664/jcsm.8524] [Medline: 32351206]
8. Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. JAMA Psychiatry 2015 Apr;72(4):334-341 [FREE Full text] [doi: 10.1001/jamapsychiatry.2014.2502] [Medline: 25671328]
9. Vancampfort D, Correll CU, Galling B, Probst M, De Hert M, Ward PB, et al. Diabetes mellitus in people with schizophrenia, bipolar disorder and major depressive disorder: a systematic review and large scale meta-analysis. World Psychiatry 2016 Jun;15(2):166-174 [FREE Full text] [doi: 10.1002/wps.20309] [Medline: 27265707]
10. Seldenrijk A, Vogelzangs N, Batelaan NM, Wieman I, van Schaik DJF, Penninx BJWH. Depression, anxiety and 6-year risk of cardiovascular disease. J Psychosom Res 2015 Feb;78(2):123-129 [FREE Full text] [doi: 10.1016/j.jpsychores.2014.10.007] [Medline: 25454680]
11. Ribeiro JD, Huang X, Fox KR, Franklin JC. Depression and hopelessness as risk factors for suicide ideation, attempts and death: meta-analysis of longitudinal studies. Br J Psychiatry 2018 May;212(5):279-286 [FREE Full text] [doi: 10.1192/bjp.2018.27] [Medline: 29587888]
12. Reeves WC, Lin JMS, Nater UM. Mental illness in metropolitan, urban and rural Georgia populations. BMC Public Health 2013 Apr 30;13:414 [FREE Full text] [doi: 10.1186/1471-2458-13-414] [Medline: 23631737]
13. Rodríguez-Blázquez C, Damián J, Andrés-Prado MJ, Almazán-Isla J, Alcalde-Cabero E, Forjaz MJ, DISCAP-ARAGON Research Group. Associations between chronic conditions, body functions, activity limitations and participation restrictions: a cross-sectional approach in Spanish non-clinical populations. BMJ Open 2016 Jun 14;6(6):e010446 [FREE Full text] [doi: 10.1136/bmjopen-2015-010446] [Medline: 27301483]
14. Taylor DJ, Mallory LJ, Lichstein KL, Durrence HH, Riedel BW, Bush AJ. Comorbidity of chronic insomnia with medical problems. Sleep 2007 Feb;30(2):213-218 [FREE Full text] [doi: 10.1093/sleep/30.2.213] [Medline: 17326547]
15. Chang JR, Fu SN, Li X, Li SX, Wang X, Zhou Z, et al. The differential effects of sleep deprivation on pain perception in individuals with or without chronic pain: a systematic review and meta-analysis. Sleep Med Rev 2022 Dec 1;66:101695 [FREE Full text] [doi: 10.1016/j.smrv.2022.101695] [Medline: 36334461]
16. Neverdahl JP, Uglem M, Matre D, Hansen JO, Engstrøm M, Tronvik E, et al. Pain thresholds and suprathreshold pain after sleep restriction in migraine - a blinded crossover study. Cephalalgia 2022 May;42(6):466-480 [FREE Full text] [doi: 10.1177/03331024211056565] [Medline: 34786965]
17. Mathias JL, Cant ML, Burke ALJ. Sleep disturbances and sleep disorders in adults living with chronic pain: a meta-analysis. Sleep Med 2018 Dec 1;52:198-210. [doi: 10.1016/j.sleep.2018.05.023] [Medline: 30314881]
18. Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. J Pain 2013 Dec;14(12):1539-1552 [FREE Full text] [doi: 10.1016/j.jpain.2013.08.007] [Medline: 24290442]
19. Tang NKY, Wright KJ, Salkovskis PM. Prevalence and correlates of clinical insomnia co-occurring with chronic back pain. J Sleep Res 2007 Mar;16(1):85-95. [doi: 10.1111/j.1365-2869.2007.00571.x] [Medline: 17309767]
20. Call-Schmidt TA, Richardson SJ. Prevalence of sleep disturbance and its relationship to pain in adults with chronic pain. Pain Manag Nurs 2003 Sep;4(3):124-133. [doi: 10.1016/s1524-9042(02)54212-0] [Medline: 14566710]
21. Hertenstein E, Trinca E, Wunderlin M, Schneider CL, Züst MA, Fehér KD, et al. Cognitive behavioral therapy for insomnia in patients with mental disorders and comorbid insomnia: a systematic review and meta-analysis. Sleep Med Rev 2022 Apr;62:101597 [FREE Full text] [doi: 10.1016/j.smrv.2022.101597] [Medline: 35240417]
22. van Mill JG, Hoogendijk WJG, Vogelzangs N, van Dyck R, Penninx BWJH. Insomnia and sleep duration in a large cohort of patients with major depressive disorder and anxiety disorders. J Clin Psychiatry 2010 Mar;71(3):239-246. [doi: 10.4088/JCP.09m05218gry] [Medline: 20331928]
23. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. Arch Intern Med 2003 Nov 10;163(20):2433-2445 [FREE Full text] [doi: 10.1001/archinte.163.20.2433] [Medline: 14609780]
24. Meda RT, Nuguru SP, Rachakonda S, Sripathi S, Khan MI, Patel N. Chronic pain-induced depression: a review of prevalence and management. Cureus 2022 Aug;14(8):e28416 [FREE Full text] [doi: 10.7759/cureus.28416] [Medline: 36171845]

25. Harrison L, Wilson S, Heron J, Stannard C, Munafò MR. Exploring the associations shared by mood, pain-related attention and pain outcomes related to sleep disturbance in a chronic pain sample. Psychol Health 2016 May;31(5):565-577 [FREE Full text] [doi: 10.1080/08870446.2015.1124106] [Medline: 26726076]
26. Nicassio PM, Ormseth SR, Kay M, Custodio M, Irwin MR, Olmstead R, et al. The contribution of pain and depression to self-reported sleep disturbance in patients with rheumatoid arthritis. Pain 2012 Jan;153(1):107-112 [FREE Full text] [doi: 10.1016/j.pain.2011.09.024] [Medline: 22051047]
27. Cao X, Chen Z, Wu L, Zhou J. Co-occurrence of chronic pain, depressive symptoms, and poor sleep quality in a health check-up population in China: a multicenter survey. J Affect Disord 2021 Feb 15;281:792-798 [FREE Full text] [doi: 10.1016/j.jad.2020.11.060] [Medline: 33229026]
28. Nandagiri V, Vannemreddy S, Spector A. Sleep disparities in Asian Americans: a comprehensive review. J Clin Sleep Med 2023 Feb 1;19(2):393-402. [doi: 10.5664/jcsm.10330] [Medline: 36239044]
29. Whinnery J, Jackson N, Rattanaumpawan P, Grandner MA. Short and long sleep duration associated with race/ethnicity, sociodemographics, and socioeconomic position. Sleep 2014 Mar 1;37(3):601-611 [FREE Full text] [doi: 10.5665/sleep.3508] [Medline: 24587584]
30. Rowell LN, Mechlin B, Ji E, Addamo M, Girdler SS. Asians differ from non-hispanic whites in experimental pain sensitivity. Eur J Pain 2011 Aug;15(7):764-771 [FREE Full text] [doi: 10.1016/j.ejpain.2010.11.016] [Medline: 21561793]
31. Watson PJ, Latif RK, Rowbotham DJ. Ethnic differences in thermal pain responses: a comparison of South Asian and White British healthy males. Pain 2005 Nov;118(1-2):194-200. [doi: 10.1016/j.pain.2005.08.010] [Medline: 16202529]
32. Bailey RK, Mokonogho J, Kumar A. Racial and ethnic differences in depression: current perspectives. Neuropsychiatr Dis Treat 2019;15:603-609 [FREE Full text] [doi: 10.2147/NDT.S128584] [Medline: 30863081]
33. Eken HN, Dee EC, Powers AR, Jordan A. Racial and ethnic differences in perception of provider cultural competence among patients with depression and anxiety symptoms: a retrospective, population-based, cross-sectional analysis. Lancet Psychiatry 2021 Nov;8(11):957-968 [FREE Full text] [doi: 10.1016/S2215-0366(21)00285-6] [Medline: 34563316]
34. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. Lancet 2021 May 29;397(10289):2082-2097. [doi: 10.1016/S0140-6736(21)00393-7] [Medline: 34062143]
35. Medalie L, Cifu AS. Management of chronic insomnia disorder in adults. JAMA 2017 Feb 21;317(7):762-763. [doi: 10.1001/jama.2016.19004] [Medline: 28241342]
36. Wu JQ, Appleman ER, Salazar RD, Ong JC. Cognitive behavioral therapy for insomnia comorbid with psychiatric and medical conditions: a meta-analysis. JAMA Intern Med 2015 Sep;175(9):1461-1472 [FREE Full text] [doi: 10.1001/jamainternmed.2015.3006] [Medline: 26147487]
37. Mitchell LJ, Bisdounis L, Ballesio A, Omlin X, Kyle SD. The impact of cognitive behavioural therapy for Insomnia on objective sleep parameters: a meta-analysis and systematic review. Sleep Med Rev 2019 Oct;47:90-102. [doi: 10.1016/j.smrv.2019.06.002] [Medline: 31377503]
38. Lai MM, Lin CC, Lin CC, Liu CS, Li TC, Kao CH. Long-term use of zolpidem increases the risk of major injury: a population-based cohort study. Mayo Clin Proc 2014 May;89(5):589-594. [doi: 10.1016/j.mayocp.2014.01.021] [Medline: 24684782]
39. Morin CM, Gaulier B, Barry T, Kowatch RA. Patients' acceptance of psychological and pharmacological therapies for Insomnia. Sleep 1992 Aug;15(4):302-305 [FREE Full text] [doi: 10.1093/sleep/15.4.302] [Medline: 1519003]
40. Omvik S, Pallesen S, Bjorvatn B, Sivertsen B, Havik OE, Nordhus IH. Patient characteristics and predictors of sleep medication use. Int Clin Psychopharmacol 2010 Mar;25(2):91-100. [doi: 10.1097/YIC.0b013e328334e5e6] [Medline: 20071997]
41. Vitiello MV, Rybarczyk B, Von Korff M, Stepanski EJ. Cognitive behavioral therapy for insomnia improves sleep and decreases pain in older adults with co-morbid insomnia and osteoarthritis. J Clin Sleep Med 2009 Aug 15;5(4):355-362 [FREE Full text] [Medline: 19968014]
42. Hikichi H, Aida J, Matsuyama Y, Tsuboya T, Kondo K, Kawachi I. Community-level social capital and cognitive decline after a natural disaster: a natural experiment from the 2011 great east Japan Earthquake and Tsunami. Soc Sci Med 2020 Jul;257:111981 [FREE Full text] [doi: 10.1016/j.socscimed.2018.09.057] [Medline: 30293854]
43. Lin N, Ye X, Ensel WM. Social support and depressed mood: a structural analysis. J Health Soc Behav 1999 Dec;40(4):344-359. [Medline: 10643160]
44. Su D, Chen Z, Chang J, Gong G, Guo D, Tan M, et al. Effect of social participation on the physical functioning and depression of empty-nest elderly in China: evidence from the China Health and Retirement Longitudinal Survey (CHARLS). Int J Environ Res Public Health 2020 Dec;17(24):9438 [FREE Full text] [doi: 10.3390/ijerph17249438] [Medline: 33339258]
45. Solé E, Racine M, Tomé-Pires C, Galán S, Jensen MP, Miró J. Social factors, disability, and depressive symptoms in adults with chronic pain. Clin J Pain 2020 May;36(5):371-378. [doi: 10.1097/AJP.0000000000000815] [Medline: 32040011]
46. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep 15;51(6):843-857 [FREE Full text] [doi: 10.1016/s0277-9536(00)00065-4] [Medline: 10972429]
47. Wang Y, Li J, Fu P, Jing Z, Zhou C. Social engagement and sleep quality among older adults in rural China: is there a sex difference? J Am Med Dir Assoc 2021 Oct;22(10):2140-2146. [doi: 10.1016/j.jamda.2020.11.014] [Medline: 33309649]

48. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China health and retirement longitudinal study. Bull World Health Organ 2014 Jan 1;92(1):29-41 [FREE Full text] [doi: 10.2471/BLT.13.124495] [Medline: 24391298]
49. China Health and Retirement Longitudinal Study. URL: http://charls.pku.edu.cn/ [accessed 2023-07-04]
50. Boey KW. Cross-validation of a short form of the CES-D in Chinese elderly. Int J Geriatr Psychiatry 1999 Aug;14(8):608-617. [doi: 10.1002/(sici)1099-1166(199908)14:8<608::aid-gps991>3.0.co;2-z] [Medline: 10489651]
51. Snyder E, Cai B, DeMuro C, Morrison MF, Ball W. A new single-item sleep quality scale: results of psychometric evaluation in patients with chronic primary insomnia and depression. J Clin Sleep Med 2018 Nov 15;14(11):1849-1857 [FREE Full text] [doi: 10.5664/jcsm.7478] [Medline: 30373688]
52. Zhu Q, Fan H, Zhang X, Ji C, Xia Y. Changes in sleep duration and 3-year risk of mild cognitive impairment in Chinese older adults. Aging (Albany NY) 2020 Jan 3;12(1):309-317 [FREE Full text] [doi: 10.18632/aging.102616] [Medline: 31899685]
53. Hou C, Lin Y, Zimmer Z, Tse LA, Fang X. Association of sleep duration with risk of all-cause mortality and poor quality of dying in oldest-old people: a community-based longitudinal study. BMC Geriatr 2020 Sep 21;20(1):357 [FREE Full text] [doi: 10.1186/s12877-020-01759-6] [Medline: 32958037]
54. Cappelleri JC, Bushmakin AG, McDermott AM, Sadosky AB, Petrie CD, Martin S. Psychometric properties of a single-item scale to assess sleep quality among individuals with fibromyalgia. Health Qual Life Outcomes 2009 Jun 17;7:54 [FREE Full text] [doi: 10.1186/1477-7525-7-54] [Medline: 19534799]
55. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health 2015 Mar 1;1(1):40-43 [FREE Full text] [doi: 10.1016/j.sleh.2014.12.010] [Medline: 29073412]
56. Gu D, Sautter J, Pipkin R, Zeng Y. Sociodemographic and health correlates of sleep quality and duration among very old Chinese. Sleep 2010 May 1;33(5):601-610 [FREE Full text] [doi: 10.1093/sleep/33.5.601] [Medline: 20469802]
57. Savedra MC, Tesler MD, Holzemer WL, Wilkie DJ, Ward JA. Pain location: validity and reliability of body outline markings by hospitalized children and adolescents. Res Nurs Health 1989 Oct;12(5):307-314. [doi: 10.1002/nur.4770120506] [Medline: 2798951]
58. Margolis RB, Chibnall JT, Tait RC. Test-retest reliability of the pain drawing instrument. Pain 1988 Apr;33(1):49-51. [doi: 10.1016/0304-3959(88)90202-3] [Medline: 3380550]
59. Levasseur M, Richard L, Gauvin L, Raymond E. Inventory and analysis of definitions of social participation found in the aging literature: proposed taxonomy of social activities. Soc Sci Med 2010 Dec;71(12):2141-2149 [FREE Full text] [doi: 10.1016/j.socscimed.2010.09.041] [Medline: 21044812]
60. Wang R, Chen Z, Zhou Y, Shen L, Zhang Z, Wu X. Melancholy or mahjong? Diversity, frequency, type, and rural-urban divide of social participation and depression in middle- and old-aged Chinese: a fixed-effects analysis. Soc Sci Med 2019 Oct 1;238:112518 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112518] [Medline: 31473574]
61. Santini ZI, Jose PE, Koyanagi A, Meilstrup C, Nielsen L, Madsen KR, et al. Formal social participation protects physical health through enhanced mental health: a longitudinal mediation analysis using three consecutive waves of the Survey of Health, Ageing and Retirement in Europe (SHARE). Soc Sci Med 2020 Apr 1;251:112906 [FREE Full text] [doi: 10.1016/j.socscimed.2020.112906] [Medline: 32151885]
62. Dong Y, Peng CYJ. Principled missing data methods for researchers. Springerplus 2013 Dec;2(1):222 [FREE Full text] [doi: 10.1186/2193-1801-2-222] [Medline: 23853744]
63. Jose PE. The merits of using longitudinal mediation. Educational Psychologist 2016 Dec;51(3-4):331-341. [doi: 10.1080/00461520.2016.1207175]
64. Hoyle RH. Structural Equation Modeling: Concepts, Issues, and Applications. Thousand Oaks, United States: SAGE Publications, Incorporated; 1995.
65. Koffel E, Kroenke K, Bair MJ, Leverty D, Polusny MA, Krebs EE. The bidirectional relationship between sleep complaints and pain: analysis of data from a randomized trial. Health Psychol 2016 Jan;35(1):41-49 [FREE Full text] [doi: 10.1037/hea0000245] [Medline: 26076002]
66. Bonvanie IJ, Oldehinkel AJ, Rosmalen JGM, Janssens KAM. Sleep problems and pain: a longitudinal cohort study in emerging adults. Pain 2016 Apr;157(4):957-963. [doi: 10.1097/j.pain.0000000000000466] [Medline: 26683236]
67. Lin L, Wang HH, Lu C, Chen W, Guo VY. Adverse childhood experiences and subsequent chronic diseases among middle-aged or older adults in China and associations with demographic and socioeconomic characteristics. JAMA Netw Open 2021 Oct 25;4(10):e2130143 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.30143] [Medline: 34694390]
68. Morelhão PK, Gobbi C, Christofaro DGD, Damato TM, Grande GD, Frange C, et al. Bidirectional association between sleep quality and low back pain in older adults: a longitudinal observational study. Arch Phys Med Rehabil 2022 Aug;103(8):1558-1564. [doi: 10.1016/j.apmr.2021.11.009] [Medline: 34968438]
69. Gabriel BM, Zierath JR. Circadian rhythms and exercise - re-setting the clock in metabolic disease. Nat Rev Endocrinol 2019 Apr;15(4):197-206. [doi: 10.1038/s41574-018-0150-x] [Medline: 30655625]
70. Bumgarner JR, Walker WH, Nelson RJ. Circadian rhythms and pain. Neurosci Biobehav Rev 2021 Oct;129:296-306 [FREE Full text] [doi: 10.1016/j.neubiorev.2021.08.004] [Medline: 34375675]

71. Meyer N, Harvey AG, Lockley SW, Dijk DJ. Circadian rhythms and disorders of the timing of sleep. Lancet 2022 Sep 24;400(10357):1061-1078. [doi: 10.1016/S0140-6736(22)00877-7] [Medline: 36115370]
72. Puterflam JM, Comis JJ, Lan Q, Liu C, Lipschitz AJ, Grunstein RR, et al. The short-term effects of opioid and non-opioid pharmacotherapies on sleep in people with chronic low back pain: a systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev 2022 Oct 1;65:101672. [doi: 10.1016/j.smrv.2022.101672] [Medline: 36087458]

Abbreviations

CES-D 10: 10-item Center for Epidemiologic Studies Depression Scale
CFI: comparative fit index
CHARLS: China Health and Retirement Longitudinal Study
SRMR: standardized root mean square residual

Edited by A Mavragani, T Sanchez; submitted 11.04.23; peer-reviewed by J Xiao, SY Wang; comments to author 18.05.23; revised version received 07.06.23; accepted 20.06.23; published 26.07.23.

Please cite as:
Fan S, Wang Q, Zheng F, Wu Y, Yu T, Wang Y, Zhang X, Zhang D
Depression as a Mediator and Social Participation as a Moderator in the Bidirectional Relationship Between Sleep Disorders and Pain: Dynamic Cohort Study
JMIR Public Health Surveill 2023;9:e48032
URL: https://publichealth.jmir.org/2023/1/e48032
doi:10.2196/48032
PMID:37494109

Original Paper

Joint Effects of Long-Term Exposure to Ambient Fine Particulate Matter and Ozone on Asthmatic Symptoms: Prospective Cohort Study

Jiahong Xu[1*], MPH; Yan Shi[2*], PhD; Gongbo Chen[3*], PhD; Yanfei Guo[2], PhD; Weiling Tang[1], MPH; Cuiling Wu[1], MPH; Shuru Liang[1], MPH; Zhongguo Huang[1], MPH; Guanhao He[1], PhD; Xiaomei Dong[1], PhD; Ganxiang Cao[4], MPH; Pan Yang[1], PhD; Ziqiang Lin[1], PhD; Sui Zhu[1], PhD; Fan Wu[5], PhD; Tao Liu[1], PhD; Wenjun Ma[1], PhD

[1]Department of Public Health and Preventive Medicine, School of Medicine, Jinan University, Guangzhou, China
[2]Shanghai Municipal Centre for Disease Control and Prevention, Shanghai, China
[3]School of Public Health and Preventive Medicine, Monash University, Victoria, Australia
[4]School of Public Health, Guangdong Pharmaceutical University, Guangzhou, China
[5]Shanghai Medical College, Fudan University, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Tao Liu, PhD
Department of Public Health and Preventive Medicine
School of Medicine
Jinan University
No.601 West, Huangpu Road, Tianhe District
Guangzhou, 510632
China
Phone: 86 13826026605
Email: gztt_2002@163.com

Abstract

Background: The associations of long-term exposure to air pollutants in the presence of asthmatic symptoms remain inconclusive and the joint effects of air pollutants as a mixture are unclear.

Objective: We aimed to investigate the individual and joint associations of long-term exposure to ambient fine particulate matter ($PM_{2.5}$) and daily 8-hour maximum ozone concentrations (MDA8 O_3) in the presence of asthmatic symptoms in Chinese adults.

Methods: Data were derived from the World Health Organization Study on Global Ageing and Adult Health (WHO SAGE) cohort study among adults aged 50 years or older, which was implemented in 1 municipality and 7 provinces across China during 2007-2018. Annual average MDA8 O_3 and $PM_{2.5}$ at individual residential addresses were estimated by an iterative random forest model and a satellite-based spatiotemporal model, respectively. Participants who were diagnosed with asthma by a doctor or taking asthma-related therapies or experiencing related conditions within the past 12 months were recorded as having asthmatic symptoms. The individual associations of $PM_{2.5}$ and MDA8 O_3 with asthmatic symptoms were estimated by a Cox proportional hazards regression model, and the joint association was estimated by a quantile g-computation model. A series of subgroup analyses was applied to examine the potential modifications of some characteristics. We also calculated the population-attributable fraction (PAF) of asthmatic symptoms attributed to $PM_{2.5}$ and MDA8 O_3.

Results: A total of 8490 adults older than 50 years were included, and the average follow-up duration was 6.9 years. During the follow-up periods, 586 (6.9%) participants reported asthmatic symptoms. Individual effect analyses showed that the risk of asthmatic symptoms was positively associated with MDA8 O_3 (hazard ratio [HR] 1.12, 95% CI 1.01-1.24, for per quantile) and $PM_{2.5}$ (HR 1.18, 95% CI 1.05-1.31, for per quantile). Joint effect analyses showed that per equal quantile increment of MDA8 O_3 and $PM_{2.5}$ was associated with an 18% (HR 1.18, 95% CI 1.05-1.33) increase in the risk of asthmatic symptoms, and $PM_{2.5}$ contributed more (68%) in the joint effects. The individual PAFs of asthmatic symptoms attributable to $PM_{2.5}$ and MDA8 O_3 were 2.86% (95% CI 0.17%-5.50%) and 4.83% (95% CI 1.42%-7.25%), respectively, while the joint PAF of asthmatic symptoms attributable to exposure mixture was 4.32% (95% CI 1.10%-7.46%). The joint associations were greater in participants with obesity, in urban areas, with lower family income, and who used unclean household cooking fuel.

Conclusions: Long-term exposure to $PM_{2.5}$ and MDA8 O_3 may individually and jointly increase the risk of asthmatic symptoms, and the joint effects were smaller than the sum of individual effects. These findings informed the importance of joint associations of long-term exposure to air pollutants with asthma.

(JMIR Public Health Surveill 2023;9:e47403) doi:10.2196/47403

KEYWORDS

asthma; joint association; ozone; PM2.5; population attributable fraction; air pollution; long-term exposure

Introduction

Asthma is a heterogeneous disease defined by a history of respiratory symptoms (eg, wheezing, shortness of breath, chest tightness, and cough) that vary over time and in intensity, accompanied by variable expiratory airflow limitation [1]. Asthma could affect people of all ages worldwide. The Global Burden of Disease (GBD) study in 2019 estimated that there were 262 million people worldwide affected by asthma, with a 15.7% increase in prevalence since 2010 [2].

The complexity of asthma is particularly challenging, and increasing evidence has linked air pollution exposures to the incidence of asthma. There is ample evidence that supports the association of short-term exposure to air pollution and morbidity of asthma [3,4]. However, the associations between long-term exposure to air pollution and asthma remain inconclusive [5]. Some studies reported positive associations of long-term exposure to air pollutants with the incidence of asthma [6-11], while others reported null or negative associations [12-15]. This inconsistency may be related to different populations, exposure assessment approaches for air pollution, and study designs among previous studies. For example, several studies were conducted in children, in which the role of chronic air pollution exposure in the development of childhood asthma has been well demonstrated [4,16], while in some studies conducted in adults, the role of air pollution in adult-onset asthma was inconclusive [6,7,10,11]. It was suggested that childhood-onset and adult-onset asthma are 2 distinct asthma phenotypes that have different clinical, biological, and genetic characteristics [17]. The associations of air pollution with adult-onset asthma should not be extrapolated from studies in children and should be particularly investigated. In addition, more studies that investigated the chronic effects of air pollution on asthma were cross-sectional studies, and few were prospective studies [18]. Those results from cross-sectional studies may be biased due to recall bias and confounding bias. Therefore, we believe more prospective studies on adults are needed.

Although studies on the long-term effects of air pollution on asthma in adults are few and inconsistent, several previous studies showed a positive association between air pollution and other chronic respiratory diseases such as chronic obstructive pulmonary disease (COPD). For example, a 15-year population cohort study in Canada found positive associations of COPD with ambient fine particulate matter ($PM_{2.5}$) and O_3 [19]. A systemic review and meta-analysis showed that per 10 $\mu g/m^3$ increase in $PM_{2.5}$ is associated with an increased incidence of COPD (hazard ratio [HR] 1.18, 95% CI 1.13-1.23) [20]. These studies suggested the plausibility of the association between air pollution and asthma.

It is increasingly recognized that people in their daily lives do not get exposed to a single air pollutant but rather to multiple air pollutants as a mixture that altogether may impact health. The idea of coexposures to multiple pollutants has been articulated as part of the exposome concept [21], and the scientific community and the US Environmental Protection Agency are moving toward a multipollutant approach to quantify the health consequences of air pollution mixtures as a whole [22]. A prospective cohort study from Australia reported a significant joint association of 5 air pollutants (PM_{10}, $PM_{2.5}$, CO, NO_x, and SO_2) with COPD, although they did not find significant associations of COPD with several pollutants in the multiple pollutant model [12], suggesting the importance of assessing joint effects of air pollutants as a whole. However, to date, very few studies have estimated the joint effects of air pollutants on the risk of asthma.

China is one of the most popular countries with a large number of patients with asthma and has serious air pollution. The GBD 2019 study estimated that there were about 24.8 million patients with asthma across China in 2019 [2]. The annual average $PM_{2.5}$ concentration across China was as high as 77.05 $\mu g/m^3$ in 2013 [23]. To tackle severe air pollution, Chinese governments have implemented a series of rigorous air pollution control policies in the past decade [24]. Although the national average $PM_{2.5}$ concentration across China in 2021 has been significantly reduced to 30 $\mu g/m^3$, it is still 6 times higher than the recommended air quality guidelines (5 $\mu g/m^3$) by World Health Organization (WHO) (2021). In addition, these control policies in air pollution inappropriately reduce NO_x rather than the volatile organic compounds, which may lead to an increase in the ambient ozone (O_3) [25]. The national population-weighted mean concentrations of maximum daily 8-hour average ozone (MDA8 O_3) increased from 89.34 $\mu g/m^3$ in 2013 to 100.96 $\mu g/m^3$ in 2019 [26]. Currently, both ambient $PM_{2.5}$ and O_3 have become the major air pollutants across China. However, the individual and joint associations of long-term exposure to $PM_{2.5}$ and O_3 with asthma in adults remain unknown.

To fill in these research gaps, this study implemented a national cohort study in China to estimate both individual and joint associations of long-term exposure to $PM_{2.5}$ and O_3 with risk of asthmatic symptoms and calculated the population-attributable fraction (PAF) of asthmatic symptoms attributable to $PM_{2.5}$ and O_3 in Chinese adults aged older than 50 years.

Methods

Study Design and Population

The World Health Organization Study on Global Ageing and Adult Health (WHO SAGE) is a cohort study that was conducted in 1 municipality (Shanghai) and 7 provinces (Guangdong, Hubei, Jilin, Shaanxi, Shandong, Yunnan, and Zhejiang) in China by using a stratified multistage random cluster sampling among adults aged older than 18 years and older from 2007 to 2018 [27,28]. A total of 18,673 participants were recruited at baseline in the first wave (2007-2010) conducted through a face-to-face household interview and were followed up during 2014-2015 and 2017-2018, respectively. In each round of follow-up investigation, new participants were added to enlarge the sample size. Eventually, a total of 18,673 participants were included in the research. Participants were excluded from this study if they were lost to follow-up (n=8053), reported asthmatic symptoms at baseline (n=420), unavailability of data for follow-up (n=33), without information about asthma (n=676), were younger than 50 years (n=728), and missing information on key covariates (n=273). Finally, a total of 8490 participants were included in the analysis (Figure S1 in Multimedia Appendix 1).

Ethics Approval

This study was approved by the Ethics Committee of the Chinese Centre for Disease Control and Prevention. Informed consent was obtained from each participant before the interview. Individually identifiable information was anonymous. WHO SAGE was approved by the World Health Organization's Ethical Review Board (RPC146), in China, SAGE wave1 was approved by the Chinese Center for Disease Control and Prevention (200601), and SAGE wave 2-3 was approved by Shanghai Municipal Center for Disease Control and Prevention (2014-8, 2018-1).

Assessment of Environmental Exposures

Daily 8-hour maximum ozone concentrations (MDA8 O_3) with a spatial resolution of 7×7 km across China during 2008-2018 were estimated by an iterative random forest model, which were derived from ground monitoring ozone data, satellite-derived ozone column amount, various meteorological information, normalized difference vegetation index, fractional urban cover, and elevation data. The random forest model showed sample-based and site-based cross-validation R^2 of 0.84 and 0.79, respectively, indicating the high accuracy of estimation of daily MDA8 O_3 [29]. The annual average MDA8 O_3 was calculated using the daily data in each grid. Annual mean $PM_{2.5}$ concentration was measured by a satellite-based spatiotemporal model with a spatial resolution of 1×1 km across China, which was established by the Atmospheric Composition Analysis Group. The R^2 value for the spatiotemporal model was 0.81 [30]. The annual average temperature was obtained from ERA5 [31].

We first calculated the annual mean exposure to MDA8 O_3, $PM_{2.5}$, and temperature in each calendar year (from the first day to the last day of a year) at a spatial grid across China. Then, the environment exposure data were assigned to each participant based on latitude and longitude of residence address, which ensured the accuracy of exposure information if the participant changed address. One-year exposure was regarded as an indicator of long-term exposure level and treated as a time-varying variable [32,33].

Outcome Measurement

Participants who were diagnosed with asthma by a doctor prior to the interview or taking asthma-related therapies within the past 12 months prior to the interview were recorded as having asthmatic symptoms. Participants were also defined as cases with asthmatic symptoms based on an algorithm about self-reported symptom-based questions within the past 12 months prior to the survey that is listed in Table S1 in Multimedia Appendix 1 [34]. The outcome was collected at baseline and at each round of follow-up, and the new asthmatic symptom was defined as the first ever met the condition listed above during follow-up.

Covariates

Covariates in this study included age (50-64 years and ≥65 years), sex (male and female), urbanicity (urban and rural), region (southern and northern), smoking status (ever and never), drinking status (ever, never), BMI (kg/m^2) (normal [≤23.9], overweight [24.0-27.9], and obese [≥28]), marriage status (married and unmarried), family income (low and high), major type of household cooking fuel (clean fuel and unclean fuel), educational level (no formal education, primary school, and middle school or higher), physical activity (low, middle, and high), fruit intake (sufficient and insufficient), vegetable intake (sufficient and insufficient), air pollution–related occupations (related and not related), and ambient temperature (Table S1 in Multimedia Appendix 1).

Statistical Analysis

Individual Associations of Air Pollutants With the Presence of Asthmatic Symptoms

A Cox proportional hazards model using time-varying covariates and coefficients [35] was used to estimate the HRs and 95% CI on the presence of asthmatic symptoms for each quartile increment in the pollutant concentration in the single-pollutant model. We also explored the nonlinear effects by applying a natural spline with 3 degrees of freedom in the model. We applied a series of subgroups analysis to examine the potential modifications of some characteristics.

Based on the above estimates, we used the following formula (equation 1) to calculate the individual PAF of asthmatic symptoms attributable to $PM_{2.5}$ and MDA8 O_3 [36]:

where P_i indicates the proportion of the population in exposure category "*i*" (using each $PM_{2.5}$ and MDA8 O_3 concentration quantile as the category) and RR_i represents HR of exposure category "*i*."

Joint Associations of Air Pollutants With the Presence of Asthmatic Symptoms

We first tested the collinearity among air pollutants and covariates using the Variance Inflation Score, and the results illustrated that the low collinearity with the Variance Inflation Scores of the multipollutant model among variables was less than 10 (Table S2 in Multimedia Appendix 1).

A quantile g-computation model was used by using a Cox proportional hazards model as an underlying model to estimate the joint effects of $PM_{2.5}$ and MDA8 O_3. We estimated the joint effect of air pollutants for an equal quartile increment of each pollutant in the model. The quantile g-computation model first transformed each pollutant into a categorical variable, coded as 0, 1, 2, 3, and then estimated associations with health outcomes when all pollutants change 1 unit. The model also calculated the weight of each pollutant, if all the exposure pollutants have the same direction, interpreting as the proportion of the effect and sum to 1. If the exposure pollutants have different directions of effect, the weights are interpreted as a proportion of the positive or negative effect, and sum to 2 [37]. Moreover, the quantile g-computation model was also conducted in subgroups. The between-group variability was tested using a 2-sample *z* test [38], and *P* values calculated by 2-by-2 comparisons between multiple groups were corrected using the Bonferroni method. The formula (equation 2) for the between-group variability test is as follows:

where b_1 and b_2 are the estimated effects of 2 groups, and SE_1^2 and SE_2^2 are the standard errors of the estimated effects of 2 groups [38].

Based on the above estimates, we calculated the individual and joint PAFs of asthmatic symptoms attributable to $PM_{2.5}$ and MDA8 O_3 [36]. We used the following formula (equation 3) to calculate the joint PAF:

where $PAF_{overall}$ represents the effect values of all risk factors, PAF_i is the PAF for exposure category "*i*." A variance-covariance matrix of the parameters from the regression model was used to estimate the 95% CI of joint PAF by randomly generating 10,000 PAFs from the normal distribution.

Sensitivity Analysis

Sensitivity analysis adjustment for different confounders was used to test the robustness of our findings. For some confounders with serious missing such as physical activity, fruit intake, and vegetable intake, sensitivity analyses were conducted only in participants with completed data.

All analyses were conducted by R software (version 4.1.3; R Development Core Team). All tests were 2-tailed, and *P*<.05 was considered statistically significant.

Results

Characteristics of Study Participants

A total of 8490 participants were finally included in this analysis, with a mean of 6.9 (SD 2.3) years of follow-up and an overall 58,737 person-years of follow-up. Among the total participants, 2950 (34.75%) were older than 64 years, 4066 were (47.89%) males, 4867 (57.43%) resided in rural areas, 5259 (61.94%) lived in Southern China, and 4641(54.66%) used unclean energy as the major household cooking fuel. At the follow-up, 586 (6.9%) participants were diagnosed with asthmatic symptoms (Table 1).

The mean $PM_{2.5}$ concentration was 42.85 (SD 12.74) μg/m^3, and the 25%, 50%, and 75% quantiles were 31.70 μg/m^3, 41.30 μg/m^3, and 53.10 μg/m^3, respectively. The mean MDA8 O_3 concentration was 94.19 (SD 5.62) μg/m^3, and the 25%, 50%, and 75% quantiles were 89.12 μg/m^3, 93.73 μg/m^3, and 97.29 μg/m^3, respectively. $PM_{2.5}$ was positively correlated with ambient MDA8 O_3 (*r*=0.663, *P*<.001) and ambient temperature (*r*=0.090, *P*<.001). Ambient MDA8 O_3 was also positively correlated with ambient temperature (*r*=0.428, *P*<.001; Table S3 in Multimedia Appendix 1).

Table 1. Characteristics of study participants in the World Health Organization Study on Global Ageing and Adult Health cohort study in China.

Categories	Total participants (N=8490)	Participants with asthmatic symptoms during the follow-up	
		No (n=7904)	Yes (n=586)
MDA8 O_3 (μg/m^3), mean (SD)	94.19 (5.62)	94.18 (5.64)	94.12 (5.46)
$PM_{2.5}$ (μg/m^3), mean (SD)	46.20 (12.74)	46.02 (12.64)	46.64 (14.25)
Age (years), n (%)	61.89 (8.57)	61.71 (8.52)	64.31 (8.9)
50-64	5540 (65.25)	5225 (66.11)	315 (53.8)
≥65	2950 (34.75)	2679 (33.89)	271 (46.2)
Gender, n (%)			
Male	4066 (47.89)	3780 (47.82)	286 (48.8)
Female	4424 (52.11)	4124 (52.18)	300 (51.2)
Urbanicity, n (%)			
Rural	4876 (57.43)	4606 (58.27)	270 (46.1)
Urban	3614 (42.57)	3298 (41.73)	316 (53.9)
Region, n (%)			
Southern China	5259 (61.94)	4943 (62.54)	316 (53.9)
Northern China	3231 (38.06)	2961 (37.46)	270 (46.1)
Marital status, n (%)			
Married	7340 (86.45)	6861 (86.80)	479 (81.7)
Unmarried	1150 (13.55)	1043 (13.20)	107 (18.3)
BMI (kg/m^2), n (%)			
Normal weight (≤23.9)	4683 (55.16)	4359 (55.2)	324 (55.3)
Overweight (24-27.9)	2772 (32.65)	2587 (32.73)	185 (31.6)
Obese (≥28)	1035 (12.19)	958 (12.12)	77 (13.1)
Major type of household cooking fuel, n (%)			
Clean	4641 (54.66)	4308 (54.50)	333 (56.8)
Unclean	3849 (45.34)	3596 (45.50)	253 (43.2)
Smoking status, n (%)			
Ever	2819 (33.20)	2609 (33.01)	210 (35.8)
Never	5671 (66.80)	5295 (66.99)	376 (64.2)
Drinking status, n (%)			
Ever	2716 (31.99)	2526 (31.96)	190 (32.4)
Never	5774 (68.01)	5378 (68.04)	396 (67.6)
Family income, n (%)			
High	4306 (50.72)	4008 (50.71)	298 (50.8)
Low	4184 (49.28)	3896 (49.29)	288 (49.2)
Educational level, n (%)			
No formal education	3393 (39.97)	3154 (39.90)	239 (40.8)
Primary school	1802 (21.22)	1662 (21.03)	140 (23.9)
Middle school or higher	3295 (38.81)	3088 (39.07)	207 (35.3)
Fruit intake, n (%)			
Sufficient	1259 (14.83)	1151 (14.56)	108 (18.4)

Categories	Total participants (N=8490)	Participants with asthmatic symptoms during the follow-up	
		No (n=7904)	Yes (n=586)
Insufficient	6699 (78.90)	6253 (79.11)	446 (76.1)
Missing	532 (6.27)	500 (6.33)	32 (5.5)
Vegetable intake, n (%)			
Sufficient	5828 (68.64)	5407 (68.41)	421 (71.8)
Insufficient	2417 (28.47)	2266 (28.67)	151 (25.8)
Missing	245 (2.89)	231 (2.92)	14 (2.4)
Air pollution–related occupations, n (%)			
Related	1045 (12.31)	950 (12.02)	95 (16.2)
Not related	6358 (74.89)	5945 (75.21)	413 (70.5)
Missing	1087 (12.80)	1009 (12.77)	78 (13.3)
Physical activity, n (%)			
High	4025 (47.41)	3749 (47.43)	276 (47.1)
Middle	1737 (20.46)	1625 (20.56)	112 (19.1)
Low	1611 (18.98)	1495 (18.91)	116 (19.8)
Missing	1117 (13.15)	1035 (13.10)	82 (14)

Individual Effects of Long-Term Exposure to MDA8 O_3 and $PM_{2.5}$ in the Presence of Asthmatic Symptoms

We observed positive nonlinear exposure-response curves of MDA8 O_3 and $PM_{2.5}$ in the presence of asthmatic symptoms (Figures 1 and 2). Linear analyses showed that per quantile increment of MDA8 O_3 (HR 1.12, 95% CI 1.01-1.24) and $PM_{2.5}$ (HR 1.18, 95% CI 1.05-1.31) were positively associated with the risk of asthmatic symptoms after adjusted for confounders. Subgroup analyses suggested that the associations were modified by several individual characteristics. For example, the associations between $PM_{2.5}$ and asthmatic symptoms were stronger in rural individuals (HR 1.54, 95% CI 1.30-1.83) than in urban individuals (HR 0.89, 95% CI 0.76-1.05), in individuals with low family income (HR 1.43, 95% CI 1.21-1.69) than in individual with high family income (HR 0.98, 95% CI 0.84-1.14), in individuals using unclean household cooking fuel (HR 1.66, 95% CI 1.38-1.99) than individuals using clean fuel (HR 0.90, 95% CI 0.78-1.04), and in individuals exposed to higher temperatures (HR 1.73, 95% CI 1.27-2.37) than in individuals exposed to lower temperatures (HR 0.92, 95% CI 0.75-1.12; Table 2).

Figure 1. Exposure-response curves of long-term exposure to MDA8 O_3 concentration with the presence of asthmatic symptoms in China. Adjustment for age, sex, urbanicity, region, smoking status, drinking status, BMI, marriage status, educational level, household income, major type of indoor cooking fuel, and ambient temperature. HR: hazard ratio; MDA8 O_3: daily 8-hour maximum ozone concentrations.

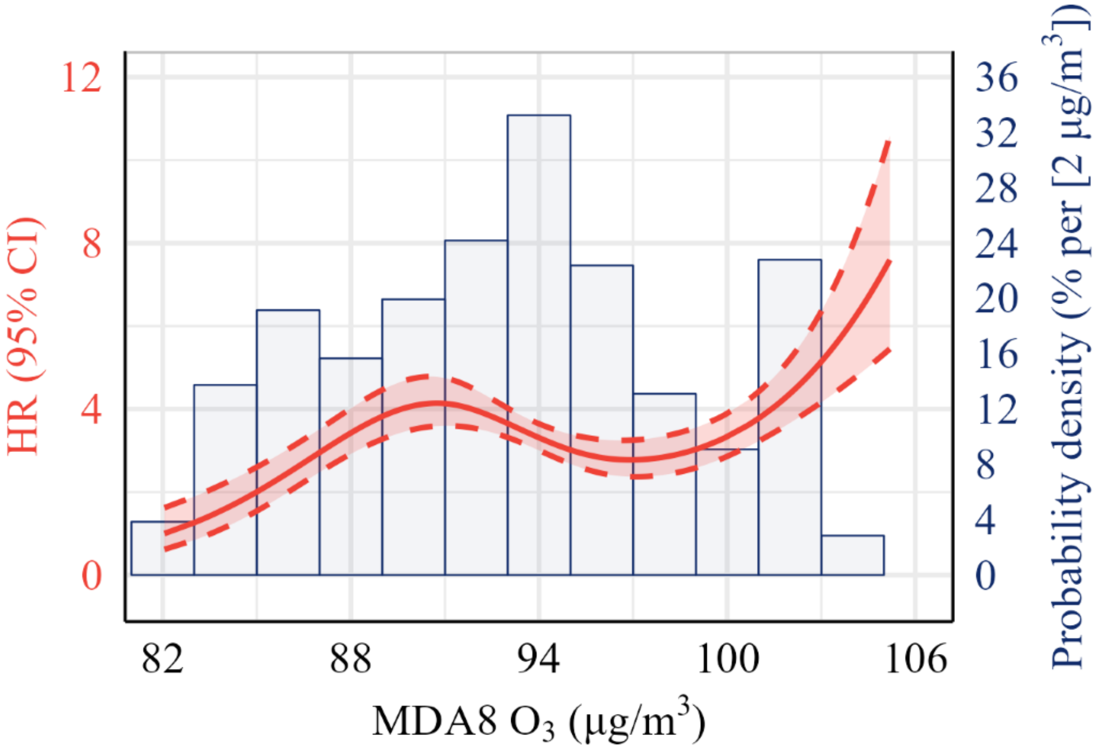


Figure 2. Exposure-response curves of long-term exposure to $PM_{2.5}$ concentration with the presence of asthmatic symptoms in China. Adjustment for age, sex, urbanicity, region, smoking status, drinking status, BMI, marriage status, educational level, household income, major type of indoor cooking fuel, and ambient temperature. HR: hazard ratio; $PM_{2.5}$: particulate matter.

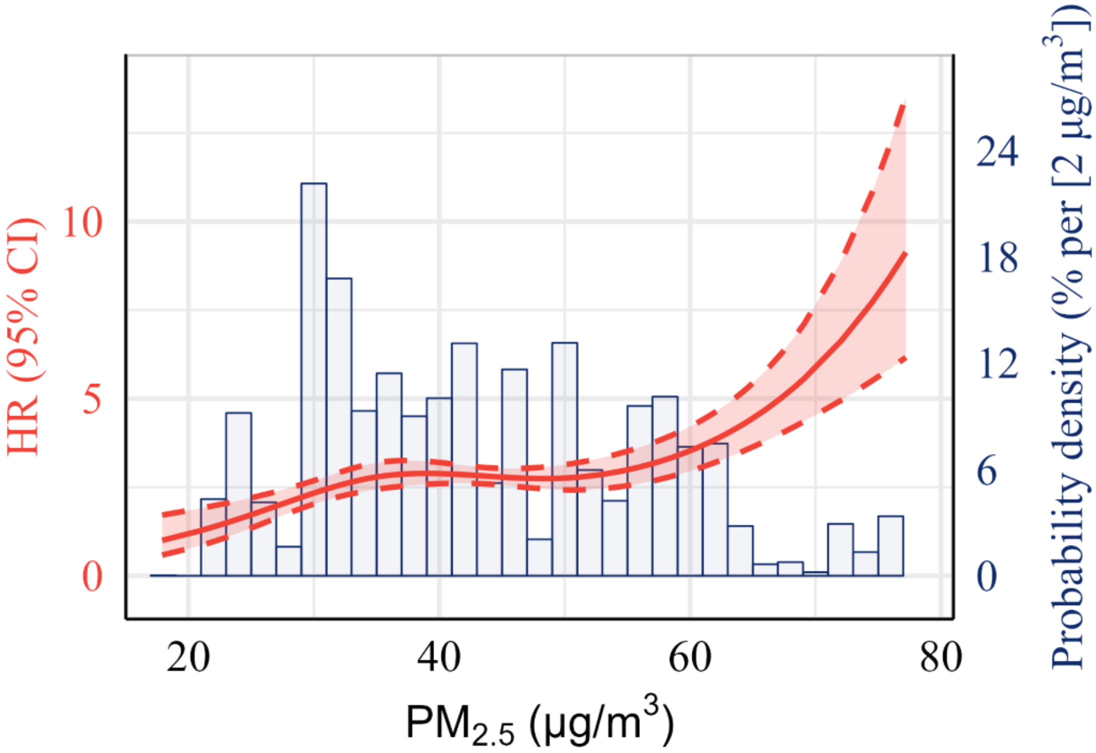


Table 2. The individual and joint associations of long-term exposures to MDA8 O_3 and $PM_{2.5}$ (per quantile increment) in the presence of asthmatic symptoms in China.[a]

Characteristics	Participants (n)	MDA8 O_3, HR (95% CI)	*P* for difference test	$PM_{2.5}$, HR (95% CI)	*P* for difference test	Joint association, HR (95% CI)	*P* for difference test
Total	8490	1.12 (1.01-1.24)	—[b]	1.18 (1.05-1.31)	—	1.18 (1.05-1.33)	—
Age (years)							
50-64	5540	1.09 (0.94-1.26)	—	1.24 (1.06-1.45)	—	1.28 (1.07-1.54)	—
≥65	2950	1.16 (0.99-1.35)	.58	1.11 (0.95-1.29)	.30	1.04 (0.84-1.27)	.13
Gender							
Male	4066	1.06 (0.92-1.23)	—	1.08 (0.93-1.27)	—	1.19 (1.00-1.42)	—
Female	4424	1.18 (1.01-1.37)	.35	1.27 (1.08-1.48)	.17	1.17 (0.90-1.51)	.88
BMI							
Normal weight	4683	1.07 (0.94-1.23)	—	1.19 (1.03-1.38)	—	1.14 (0.95-1.37)	—
Overweight	2772	1.09 (0.90-1.33)	.87	0.98 (0.81-1.19)	.95	1.02 (0.79-1.30)	.37
Obese	1035	1.87 (1.30-2.69)	.005	1.94 (1.33-2.82)	.20	2.05 (1.32-3.19)	.017
Urbanicity							
Rural	4876	1.15 (1.00-1.32)	—	1.54 (1.30-1.83)	—	1.47 (1.21-1.79)	—
Urban	3614	1.10 (0.91-1.29)	.69	0.89 (0.76-1.05)	<.001	0.87 (0.70-1.09)	<.001
Family income							
High	4306	1.12 (0.96-1.30)	—	0.98 (0.84-1.14)	—	0.98 (0.80-1.21)	—
Low	4184	1.14 (0.98-1.32)	.87	1.43 (1.21-1.69)	.001	1.39 (1.13-1.72)	.02
Major type of household cooking fuel							
Clean	4641	1.07 (0.93-1.24)	—	0.90 (0.78-1.04)	—	0.91 (0.77-1.09)	—
Unclean	3849	1.16 (0.99-1.36)	.47	1.66 (1.38-1.99)	<.001	1.65 (1.29-2.11)	<.001
Temperature (°C)							
Q[c]1 (4.41-12.68)	1931	1.16 (0.95-1.42)	—	1.11 (0.81-1.51)	—	1.16 (0.91-1.48)	—
Q2 (12.69-15.80)	2140	0.98 (0.81-1.20)	.26	1.18 (0.97-1.44)	.73	1.06 (0.82-1.38)	.49
Q3 (15.81-16.41)	2071	0.79 (0.67-0.93)	.004	0.92 (0.75-1.12)	.31	0.86 (0.68-1.08)	.10
Q4 (16.42-22.94)	2348	1.18 (0.87-1.58)	.94	1.73 (1.27-2.37)	.05	1.71 (1.03-2.86)	.30

[a] Adjustment for age, sex, urbanicity, region, smoking status, drinking status, BMI, marriage status, educational level, household income, indoor fuel type, and temperature. In subgroup analyses, other confounders except for the subgroup category variable, analyzed as an independent variable, were adjusted for.

[b] Not applicable.

[c] Q: Quartile.

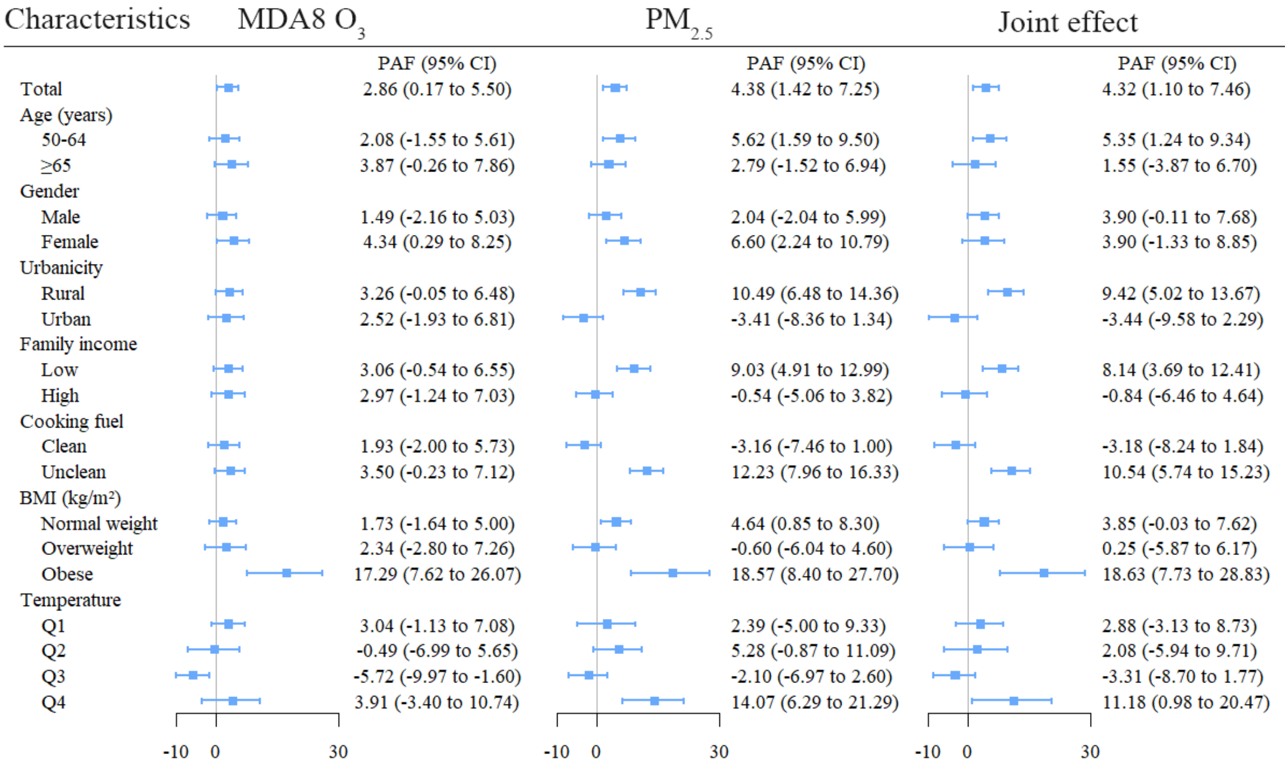
Joint Effects of Long-Term Exposure to MDA8 O_3 and $PM_{2.5}$ With Asthmatic Symptoms

Joint effect analyses showed that per quantile increment of MDA8 O_3 and $PM_{2.5}$ were associated with an 18% (HR 1.18, 95% CI 1.05-1.33) increase in the risk of asthmatic symptoms. In the joint effect, the contribution of $PM_{2.5}$ and MDA8 O_3 was 68% and 32%, respectively. We also found greater joint HRs estimated by subgroup analyses in rural residents (HR 1.47, 95% CI 1.21-1.79) than in urban residents (HR 0.87, 95% CI 0.70-1.09), in low-income level (HR 1.39, 95% CI 1.13-1.72) than high-income level (HR 0.98, 95% CI 0.80-1.21), in unclean indoor fuel type (HR 1.65, 95% CI 1.29-2.11) than clean indoor fuel type (HR 0.91, 95% CI 0.77-1.09). Moreover, the highest HR was found in participants with obesity (HR 2.05, 95% CI 1.32-3.19; Table 2).

PAFs of Asthmatic Symptoms Attributable to MDA8 O_3 and $PM_{2.5}$

Figure 3 shows the PAFs of asthmatic symptoms attributable to MDA8 O_3 and $PM_{2.5}$, estimated by the individual and joint associations. In the total participants, the individual PAFs caused by MDA8 O_3 and $PM_{2.5}$ were 2.86% (95% CI 0.17%-5.50%) and 4.38% (95% CI 1.42%-7.25%), and the joint PAFs by MDA8 O_3 and $PM_{2.5}$ were 4.32% (95% CI 1.10%-7.46%). Subgroup analyses showed a large variation of PAFs among different groups. For example, the PAFs of asthmatic symptoms were the largest in obese individuals. The individual PAF caused by MDA8 O_3 and $PM_{2.5}$ was 17.29% (95% CI 7.62%-26.07%) and 18.57% (95% CI 8.40%-27.70%), and the joint PAFs by MDA8 O_3 and $PM_{2.5}$ was 18.63% (95% CI 7.73%-28.83%). More detailed information is shown in Figure 3.

Figure 3. The population-attributable fraction of the presence of asthmatic symptoms attributable to MDA8 O_3 and $PM_{2.5}$ exposures, estimated by the individual and joint associations. Q: Quartile; MDA8 O_3: daily 8-hour maximum ozone concentrations; $PM_{2.5}$: particulate matter.



Sensitivity Analyses

Results indicated that HRs did not substantially change after adjustment for different confounders (Table S4 in Multimedia Appendix 1).

Discussion

Principal Findings

In this national cohort study, we found a positive correlation between individual and joint associations on long-term exposure to ambient $PM_{2.5}$ and MDA8 O_3 with the risk of asthmatic symptoms in a Chinese population older than 50 years. Based on the joint associations, 4.32% of asthmatic symptoms could be attributable to $PM_{2.5}$ and MDA8 O_3 exposures, and ambient $PM_{2.5}$ contributed much more to the joint effect. In addition, the associations were modified by obesity, urbanicity, family income level, household cooking fuel, and ambient temperature. Our findings are significant for estimating the disease burden of air pollution and making policies for air pollution control.

In the literature, relatively fewer studies have investigated the associations of long-term exposure to $PM_{2.5}$ and MDA8 O_3 with asthma incidence in adults compared to those studies on short-term exposures and on childhood asthma. Some studies reported consistent results with this study that long-term exposure to $PM_{2.5}$ and MDA8 O_3 were separately associated with an increased risk of asthma incidence [6-8]. For example, the Health Effects Institute reported that long-term exposure to higher $PM_{2.5}$ was associated with an increased risk of asthma, which was from the European Cohorts in the ELAPSE Project [6]. The Epidemiological Study of the Genetic and Environmental Factors of Asthma reported that the risk of asthma increased with O_3 exposure [7]. McDonnell et al [39] also reported that long-term exposure to O_3 was positively associated with new-onset asthma in adult males. However, several previous studies did not find positive associations of $PM_{2.5}$ and MDA8 O_3 with the incidence of asthma [12,13,40]. For example, a prospective cohort from Australia reported a null association between long-term exposure to $PM_{2.5}$ with adult-onset asthma [12]. The results from the ELAPSE project reported a negative association (HR 0.90, 95% CI 0.81-0.99 for each 10 $\mu g/m^3$) between long-term exposure to O_3 and the incidence of asthma [6]. Although it may be difficult to directly compare the results among these studies due to different study populations, this interstudy variation of associations suggests that the associations of long-term exposure to $PM_{2.5}$ and O_3 with asthma incidence remain inconclusive, and more studies are needed particularly in low-and middle-income countries. For example, we did not find a prospective cohort that has investigated the associations of long-term exposure to $PM_{2.5}$ and O_3 with asthma in China.

Exposure to air pollution can trigger inflammatory and immune responses, oxidative stress, and airway remodeling in the development and exacerbation of asthma. Exposure to $PM_{2.5}$ induces the release of inflammatory mediators from alveolar macrophages, which is an important pathogenesis of pulmonary inflammation in the development of asthma. $PM_{2.5}$ induces inflammatory responses, associated with Th1/Th2 pathway imbalance and leads to Th2-oriented inflammation [41]. Apart from that, $PM_{2.5}$ induces the production of reactive oxygen species. Reactive oxygen species will further enhance the oxidative stress response, resulting in DNA, protein, lipid, and other cellular and molecular damage, leading to respiratory disease [42,43]. In this study, $PM_{2.5}$ also induced a significant upregulation of vascular endothelial growth factor A production, a signaling event that controls vascular remodeling [44]. Furthermore, exposure to O_3 is associated with an increase in free radicals and biomarkers associated with oxidative stress [45]. Ozone exposure also induced airway inflammation with increased numbers of neutrophils [46].

More importantly, we quantified the joint associations of long-term exposure to $PM_{2.5}$ and MDA8 O_3 in the presence of asthmatic symptoms in a Chinese population. Our findings had 2 major implications. First, $PM_{2.5}$ (68%) contributed more to the joint associations than O_3 (32%), suggesting the more important role of $PM_{2.5}$ in the development of asthmatic symptoms. Second, the PAFs of asthmatic symptoms based on the joint associations were lower than the sum of PAFs based on the individual associations, which suggests that the calculation based on a single-pollutant model might overestimate the effects of air pollutants due to the potential mutual confounding [47]. Several studies reported similar findings to this study [47-49]. For example, Winquist et al [47] reported smaller joint effect estimates from multipollutant models than estimates from single-pollutant models. In practice, people are usually exposed to multiple pollutants as air pollution complex mixture. Therefore, our findings combined with previous studies suggested that estimating the disease burden of air pollution calculated using a joint effect model may be more appealing because it transparently and explicitly specifies the contribution of various pollutants of mixtures.

PAF is a statistical indicator that quantitatively describes the health impact of a certain risk factor on a population. Specifically, it represents the proportion of total disease (or mortality) in the population that can be attributed to a certain factor, so we describe that a total of 4.32% of the presence of asthmatic symptoms could be attributable to joint $PM_{2.5}$ and MDA8 O_3 exposures. Although this figure is not very large, air pollution exposure is ubiquitous, and the large number of patients with asthmatic symptoms in China may indicate important public health issues. It was estimated that there were about 848,000 incident cases with asthmatic symptoms in 2019 in Chinese people older than 50 years [2]. As a result, there were about 37,000 cases with asthmatic symptoms, which may have been caused by exposure to $PM_{2.5}$ and O_3.

The results of stratified analyses showed greater effects of $PM_{2.5}$ and O_3 on the risk of asthmatic symptoms in participants in rural areas than in urban areas. The stronger associations in rural areas may be because the rural population is more likely to smoke and use unclean cooking fuel. Usage of unclean cooking fuel could also increase the risk of asthmatic symptoms through similar mechanisms of air pollution [50,51], and hence synergize the effects of ambient air pollutants. Compared with urban residents, rural residents usually have harder physical activities, which may increase their exposure to air pollution. Moreover, the rural population may be more susceptible to air pollution due to the disadvantaged sanitation and health care systems, dwelling environments, poor health status, and so forth [52]. These potential mechanisms were also confirmed by our findings of greater associations in participants with lower family income and using unclean energy as the major household cooking fuel.

The stratified analyses also showed greater associations in obese participants than in lean individuals. The modification of obesity on the respiratory health impacts of air pollution has been reported in other studies [53,54]. For example, a cross-sectional study in northeastern Chinese cities reported that obesity enhanced the respiratory health effects caused by air pollution in children [53]. Obesity is a proinflammatory state, and the adipose tissue propagates inflammation by recruitment of macrophages via chemokines such as monocyte chemoattractant protein-1, and via elaboration of cytokines and chemokines such as leptin, interleukin-6, and tumor necrosis factor α. Meanwhile, inhaled air pollutants could also produce these cytokines and chemokines, and hence increase inflammation [53,54]. As a result, obesity may amplify the inflammatory response induced by air pollution and increase the vulnerability of obese participants to harmful respiratory effects caused by air pollution exposures.

There are a few limitations in this study. First, many participants were lost to follow-up due to reasons such as immigration, urbanization, and so forth, which may lead to selection bias. Second, the presence of asthmatic symptoms was based on self-reported questionnaires, which may lead to misclassification bias. However, all reported patients must be diagnosed by a doctor or by clinical symptoms, which could partially reduce the bias. Third, some potential confounders such as physical activity, occupational exposure to air pollution, vegetable and fruit intake, and the higher rate of missing information were not adjusted for in the main models. However, the results of sensitivity analyses conducted only in those participants with complete information showed that the results are robust to these factors (Table S4 in Multimedia Appendix 1). Fourth, we did not include other air pollutants such as NO_2, SO_2, and CO due to the unavailability of data in the earlier years. In particular, NO_2 was involved in the production of atmospheric ozone, and the effects of MDA8 O_3 were biased due to the unavailability of NO_2. Therefore, more studies are needed in the future. Fifth, 2 different methods were applied to assess the participants' exposure to ambient MDA8 O_3 and $PM_{2.5}$, which may affect the result. However, the 2 methods are commonly adopted at present in air pollution exposure assessment, and the prediction results of the 2 models have been verified with high accuracy, which suggests that the impacts of using different assessment methods on our findings may be limited.

Conclusions

In conclusion, this study provided novel evidence that long-term exposures to ambient $PM_{2.5}$ and O_3 were individually and jointly associated with a higher presence of asthmatic symptoms, and ambient $PM_{2.5}$ contributed more to the joint effects. The combined effects of air pollutants based on a single-pollutant model might be overestimated. The joint effects were more pronounced in participants with obesity, from rural areas, with lower income levels, and who use unclean household cooking fuel.

Acknowledgments

We thank all participants who took part in this program, and those investigators who contributed to this program. This work was supported by the National Natural Science Foundation of China (42075173, 42175181, and 42275187) and the National Foreign Expert Project (G2022199006L).

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Authors' Contributions

TL and FW conceptualized and design this study. YG and YS contributed to data collection. JX, GC, and SL contributed to data interpretation and methodology; JX, YS, and GC contributed to data analysis and writing of the original draft. TL, FW, WM, GH, XD, PY, and ZL reviewed and edited the manuscript. All authors had full access to all data and approved the final manuscript as submitted. FW and TL are co-corresponding authors and have contributed equally to this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The joint effects of long-term exposure to ambient $PM_{2.5}$ and ozone on incident asthma: prospective cohort study.

[DOCX File , 107 KB - publichealth_v9i1e47403_app1.docx]

References

1. Reddel HK, Bacharier LB, Bateman ED, Brightling CE, Brusselle GG, Buhl R, et al. Global initiative for asthma strategy 2021: executive summary and rationale for key changes. J Allergy Clin Immunol Pract 2022;10(1S):S1-S18 [FREE Full text] [doi: 10.1016/j.jaip.2021.10.001] [Medline: 34718211]
2. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
3. Zhang S, Li G, Tian L, Guo Q, Pan X. Short-term exposure to air pollution and morbidity of COPD and asthma in East Asian area: a systematic review and meta-analysis. Environ Res 2016;148:15-23. [doi: 10.1016/j.envres.2016.03.008] [Medline: 26995350]
4. Guarnieri M, Balmes JR. Outdoor air pollution and asthma. Lancet 2014;383(9928):1581-1592 [FREE Full text] [doi: 10.1016/S0140-6736(14)60617-6] [Medline: 24792855]
5. Anderson HR, Favarato G, Atkinson RW. Long-term exposure to outdoor air pollution and the prevalence of asthma: meta-analysis of multi-community prevalence studies. Air Qual Atmos Health 2011;6(1):57-68. [doi: 10.1007/s11869-011-0145-4]
6. Brunekreef B, Strak M, Chen J, Andersen ZJ, Atkinson R, Bauwelinck M, et al. Mortality and morbidity effects of long-term exposure to low-level PM, BC, NO, and O: an analysis of European cohorts in the ELAPSE project. Res Rep Health Eff Inst 2021(208):1-127. [Medline: 36106702]
7. Havet A, Zerimech F, Sanchez M, Siroux V, Le Moual N, Brunekreef B, et al. Outdoor air pollution, exhaled 8-isoprostane and current asthma in adults: the EGEA study. Eur Respir J 2018;51(4):1702036 [FREE Full text] [doi: 10.1183/13993003.02036-2017] [Medline: 29618600]
8. Hwang BF, Lee YL, Lin YC, Jaakkola JJK, Guo YL. Traffic related air pollution as a determinant of asthma among Taiwanese school children. Thorax 2005;60(6):467-473 [FREE Full text] [doi: 10.1136/thx.2004.033977] [Medline: 15923246]
9. Knibbs LD, Cortés de Waterman AM, Toelle BG, Guo Y, Denison L, Jalaludin B, et al. The Australian Child Health and Air Pollution Study (ACHAPS): a national population-based cross-sectional study of long-term exposure to outdoor air pollution, asthma, and lung function. Environ Int 2018;120:394-403. [doi: 10.1016/j.envint.2018.08.025] [Medline: 30125857]

10. Ai S, Qian ZM, Guo Y, Yang Y, Rolling CA, Liu E, et al. Long-term exposure to ambient fine particles associated with asthma: a cross-sectional study among older adults in six low- and middle-income countries. Environ Res 2019;168:141-145. [doi: 10.1016/j.envres.2018.09.028] [Medline: 30316099]
11. Cai Y, Zijlema WL, Doiron D, Blangiardo M, Burton PR, Fortier I, et al. Ambient air pollution, traffic noise and adult asthma prevalence: a BioSHaRE approach. Eur Respir J 2017;49(1):1502127 [FREE Full text] [doi: 10.1183/13993003.02127-2015] [Medline: 27824608]
12. Hendryx M, Luo J, Chojenta C, Byles JE. Air pollution exposures from multiple point sources and risk of incident chronic obstructive pulmonary disease (COPD) and asthma. Environ Res 2019;179(Pt A):108783. [doi: 10.1016/j.envres.2019.108783] [Medline: 31590000]
13. Young MT, Sandler DP, DeRoo LA, Vedal S, Kaufman JD, London SJ. Ambient air pollution exposure and incident adult asthma in a nationwide cohort of U.S. women. Am J Respir Crit Care Med 2014;190(8):914-921 [FREE Full text] [doi: 10.1164/rccm.201403-0525OC] [Medline: 25172226]
14. Lazarevic N, Dobson AJ, Barnett AG, Knibbs LD. Long-term ambient air pollution exposure and self-reported morbidity in the Australian longitudinal study on women's health: a cross-sectional study. BMJ Open 2015;5(10):e008714 [FREE Full text] [doi: 10.1136/bmjopen-2015-008714] [Medline: 26503387]
15. Solomon C, Poole J, Jarup L, Palmer K, Coggon D. Cardio-respiratory morbidity and long-term exposure to particulate air pollution. Int J Environ Health Res 2003;13(4):327-335. [doi: 10.1080/09603120310001616100] [Medline: 14594699]
16. Anderson HR, Favarato G, Atkinson RW. Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies. Air Qual Atmos Health 2011;6(1):47-56. [doi: 10.1007/s11869-011-0144-5]
17. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. Nat Med 2012;18(5):716-725. [doi: 10.1038/nm.2678] [Medline: 22561835]
18. Lipfert FW. Long-term associations of morbidity with air pollution: a catalog and synthesis. J Air Waste Manag Assoc 2018;68(1):12-28 [FREE Full text] [doi: 10.1080/10962247.2017.1349010] [Medline: 28679072]
19. Shin S, Bai L, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, et al. Air pollution as a risk factor for incident chronic obstructive pulmonary disease and asthma. A 15-year population-based cohort study. Am J Respir Crit Care Med 2021;203(9):1138-1148 [FREE Full text] [doi: 10.1164/rccm.201909-1744OC] [Medline: 33147059]
20. Park J, Kim HJ, Lee CH, Lee CH, Lee HW. Impact of long-term exposure to ambient air pollution on the incidence of chronic obstructive pulmonary disease: a systematic review and meta-analysis. Environ Res 2021;194:110703 [FREE Full text] [doi: 10.1016/j.envres.2020.110703] [Medline: 33417909]
21. Wild CP. The exposome: from concept to utility. Int J Epidemiol 2012;41(1):24-32 [FREE Full text] [doi: 10.1093/ije/dyr236] [Medline: 22296988]
22. Dominici F, Peng RD, Barr CD, Bell ML. Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. Epidemiology 2010;21(2):187-194 [FREE Full text] [doi: 10.1097/EDE.0b013e3181cc86e8] [Medline: 20160561]
23. Ma Z, Hu X, Sayer AM, Levy R, Zhang Q, Xue Y, et al. Satellite-based spatiotemporal trends in PM2.5 concentrations: China, 2004-2013. Environ Health Perspect 2016;124(2):184-192 [FREE Full text] [doi: 10.1289/ehp.1409481] [Medline: 26220256]
24. Xie Z. China’s historical evolution of environmental protection along with the forty years’ reform and opening-up. Environ Sci Ecotechnology 2020;1:100001 [FREE Full text] [doi: 10.1016/j.ese.2019.100001]
25. Tobías A, Carnerero C, Reche C, Massagué J, Via M, Minguillón MC, et al. Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic. Sci Total Environ 2020;726:138540 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138540] [Medline: 32302810]
26. Meng X, Wang W, Shi S, Zhu S, Wang P, Chen R, et al. Evaluating the spatiotemporal ozone characteristics with high-resolution predictions in mainland China, 2013-2019. Environ Pollut 2022;299:118865 [FREE Full text] [doi: 10.1016/j.envpol.2022.118865] [Medline: 35063542]
27. Kowal P, Chatterji S, Naidoo N, Biritwum R, Fan W, Lopez Ridaura R, SAGE Collaborators. Data resource profile: the World Health Organization study on global AGEing and adult health (SAGE). Int J Epidemiol 2012;41(6):1639-1649 [FREE Full text] [doi: 10.1093/ije/dys210] [Medline: 23283715]
28. Wu F, Guo Y, Kowal P, Jiang Y, Yu M, Li X, et al. Prevalence of major chronic conditions among older Chinese adults: the Study on Global AGEing and adult health (SAGE) wave 1. PLoS One 2013;8(9):e74176 [FREE Full text] [doi: 10.1371/journal.pone.0074176] [Medline: 24069278]
29. Chen G, Chen J, Dong GH, Yang BY, Liu Y, Lu T, et al. Improving satellite-based estimation of surface ozone across China during 2008–2019 using iterative random forest model and high-resolution grid meteorological data. Sustain Cities Soc 2021;69:102807. [doi: 10.1016/j.scs.2021.102807]
30. Hammer MS, van Donkelaar A, Li C, Lyapustin A, Sayer AM, Hsu NC, et al. Global estimates and long-term trends of fine particulate matter concentrations (1998-2018). Environ Sci Technol 2020;54(13):7879-7890 [FREE Full text] [doi: 10.1021/acs.est.0c01764] [Medline: 32491847]
31. Hersbach H, Bell B, Berrisford P, Horányi A, Muñoz-Sabater J, Nicolas J, et al. Global reanalysis: goodbye ERA-Interim, hello ERA5. ECMWF Newsletter 2019;159:17-24. [doi: 10.21957/vf291hehd7]

32. Niu Y, Yang T, Gu X, Chen R, Meng X, Xu J, China Pulmonary Health Study Group. Long-term ozone exposure and small airway dysfunction: the China Pulmonary Health (CPH) study. Am J Respir Crit Care Med 2022;205(4):450-458. [doi: 10.1164/rccm.202107-1599OC] [Medline: 34813411]
33. Hvidtfeldt UA, Severi G, Andersen ZJ, Atkinson R, Bauwelinck M, Bellander T, et al. Long-term low-level ambient air pollution exposure and risk of lung cancer - a pooled analysis of 7 European cohorts. Environ Int 2021;146:106249 [FREE Full text] [doi: 10.1016/j.envint.2020.106249] [Medline: 33197787]
34. Arokiasamy P, Uttamacharya U, Jain K, Biritwum RB, Yawson AE, Wu F, et al. The impact of multimorbidity on adult physical and mental health in low- and middle-income countries: what does the study on global ageing and adult health (SAGE) reveal? BMC Med 2015;13:178 [FREE Full text] [doi: 10.1186/s12916-015-0402-8] [Medline: 26239481]
35. Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CGM. Time-varying covariates and coefficients in Cox regression models. Ann Transl Med 2018;6(7):121 [FREE Full text] [doi: 10.21037/atm.2018.02.12] [Medline: 29955581]
36. Thacher JD, Poulsen AH, Hvidtfeldt UA, Raaschou-Nielsen O, Brandt J, Geels C, et al. Long-term exposure to transportation noise and risk for type 2 diabetes in a nationwide cohort study from Denmark. Environ Health Perspect 2021;129(12):127003 [FREE Full text] [doi: 10.1289/EHP9146] [Medline: 34855467]
37. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 2020;128(4):47004 [FREE Full text] [doi: 10.1289/EHP5838] [Medline: 32255670]
38. Schenker N, Gentleman JF. On judging the significance of differences by examining the overlap between confidence intervals. Am Stat 2001;55(3):182-186. [doi: 10.1198/000313001317097960]
39. McDonnell WF, Abbey DE, Nishino N, Lebowitz MD. Long-term ambient ozone concentration and the incidence of asthma in nonsmoking adults: the AHSMOG study. Environ Res 1999;80(2 Pt 1):110-121. [doi: 10.1006/enrs.1998.3894] [Medline: 10092402]
40. Jacquemin B, Siroux V, Sanchez M, Carsin AE, Schikowski T, Adam M, et al. Ambient air pollution and adult asthma incidence in six European cohorts (ESCAPE). Environ Health Perspect 2015;123(6):613-621 [FREE Full text] [doi: 10.1289/ehp.1408206] [Medline: 25712593]
41. Pang L, Yu P, Liu X, Fan Y, Shi Y, Zou S. Fine particulate matter induces airway inflammation by disturbing the balance between Th1/Th2 and regulation of GATA3 and Runx3 expression in BALB/c mice. Mol Med Rep 2021;23(5):378 [FREE Full text] [doi: 10.3892/mmr.2021.12017] [Medline: 33760131]
42. Liu K, Hua S, Song L. PM2.5 exposure and asthma development: the key role of oxidative stress. Oxid Med Cell Longev 2022;2022:3618806 [FREE Full text] [doi: 10.1155/2022/3618806] [Medline: 35419163]
43. Sijan Z, Antkiewicz DS, Heo J, Kado NY, Schauer JJ, Sioutas C, et al. An in vitro alveolar macrophage assay for the assessment of inflammatory cytokine expression induced by atmospheric particulate matter. Environ Toxicol 2015;30(7):836-851. [doi: 10.1002/tox.21961] [Medline: 24497439]
44. Xu X, Wang H, Liu S, Xing C, Liu Y, Aodengqimuge, et al. TP53-dependent autophagy links the ATR-CHEK1 axis activation to proinflammatory VEGFA production in human bronchial epithelial cells exposed to fine particulate matter (PM2.5). Autophagy 2016;12(10):1832-1848 [FREE Full text] [doi: 10.1080/15548627.2016.1204496] [Medline: 27463284]
45. Corradi M, Alinovi R, Goldoni M, Vettori M, Folesani G, Mozzoni P, et al. Biomarkers of oxidative stress after controlled human exposure to ozone. Toxicol Lett 2002 05;134(1-3):219-225. [doi: 10.1016/s0378-4274(02)00169-8] [Medline: 12191881]
46. Xue Y, Zhou Y, Bao W, Fu Q, Hao H, Han L, et al. STAT3 and IL-6 contribute to corticosteroid resistance in an OVA and ozone-induced asthma model with neutrophil infiltration. Front Mol Biosci 2021;8:717962 [FREE Full text] [doi: 10.3389/fmolb.2021.717962] [Medline: 34760922]
47. Winquist A, Kirrane E, Klein M, Strickland M, Darrow LA, Sarnat SE, et al. Joint effects of ambient air pollutants on pediatric asthma emergency department visits in Atlanta, 1998-2004. Epidemiology 2014;25(5):666-673 [FREE Full text] [doi: 10.1097/EDE.0000000000000146] [Medline: 25045931]
48. Schildcrout JS, Sheppard L, Lumley T, Slaughter JC, Koenig JQ, Shapiro GG. Ambient air pollution and asthma exacerbations in children: an eight-city analysis. Am J Epidemiol 2006;164(6):505-517 [FREE Full text] [doi: 10.1093/aje/kwj225] [Medline: 16798793]
49. Liu T, Jiang Y, Hu J, Li Z, Li X, Xiao J, et al. Joint associations of short-term exposure to ambient air pollutants with hospital admission of ischemic stroke. Epidemiology 2023;34(2):282-292. [doi: 10.1097/EDE.0000000000001581] [Medline: 36722811]
50. Jie Y, Ismail NH, Jie X, Isa ZM. Do indoor environments influence asthma and asthma-related symptoms among adults in homes?: a review of the literature. J Formos Med Assoc 2011;110(9):555-563 [FREE Full text] [doi: 10.1016/j.jfma.2011.07.003] [Medline: 21930065]
51. Sinha D, Ray MR. Health effects of indoor air pollution due to cooking with biomass fuel. In: Roberts SM, Kehrer JP, Klotz LO, editors. Studies on Experimental Toxicology and Pharmacology. AG, Switzerland: Humana Press; 2015:267-302.

52. Liu T, Meng H, Yu M, Xiao Y, Huang B, Lin L, et al. Urban-rural disparity of the short-term association of PM with mortality and its attributable burden. Innovation (Camb) 2021;2(4):100171 [FREE Full text] [doi: 10.1016/j.xinn.2021.100171] [Medline: 34778857]
53. Dong GH, Qian Z, Liu MM, Wang D, Ren WH, Fu Q, et al. Obesity enhanced respiratory health effects of ambient air pollution in Chinese children: the Seven Northeastern Cities study. Int J Obes (Lond) 2013;37(1):94-100. [doi: 10.1038/ijo.2012.125] [Medline: 22846775]
54. Limaye S, Salvi S. Obesity and asthma: the role of environmental pollutants. Immunol Allergy Clin North Am 2014;34(4):839-855. [doi: 10.1016/j.iac.2014.07.005] [Medline: 25282295]

Abbreviations

COPD: chronic obstructive pulmonary disease
GBD: Global Burden of Disease
HR: hazard ratio
MDA8 O_3: daily 8-hour maximum ozone concentrations
PAF: population-attributable fraction
$PM_{2.5}$: particulate matter
WHO SAGE: World Health Organization Study on Global Ageing and Adult Health

Edited by A Mavragani, T Sanchez; submitted 18.03.23; peer-reviewed by HW Lee, X Deng; comments to author 17.04.23; revised version received 08.05.23; accepted 21.06.23; published 03.08.23.

Please cite as:
Xu J, Shi Y, Chen G, Guo Y, Tang W, Wu C, Liang S, Huang Z, He G, Dong X, Cao G, Yang P, Lin Z, Zhu S, Wu F, Liu T, Ma W
Joint Effects of Long-Term Exposure to Ambient Fine Particulate Matter and Ozone on Asthmatic Symptoms: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e47403
URL: https://publichealth.jmir.org/2023/1/e47403
doi:10.2196/47403
PMID:37535415

Original Paper

Association of Catastrophic Health Expenditure With the Risk of Depression in Chinese Adults: Population-Based Cohort Study

Yaping Wang[1], PhD; Wannian Liang[2,3], PhD; Min Liu[1,4], PhD; Jue Liu[1,4,5,6], PhD

[1]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[2]Vanke School of Public Health, Tsinghua University, Beijing, China
[3]Institute for Healthy China, Tsinghua University, Beijing, China
[4]Key Laboratory of Epidemiology of Major Diseases, Peking University, Ministry of Education, Beijing, China
[5]Institute for Global Health and Development, Peking University, Beijing, China
[6]Peking University Health Science Center-Weifang Joint Research Center for Maternal and Child Health, Peking University, Beijing, China

Corresponding Author:
Min Liu, PhD
Department of Epidemiology and Biostatistics
School of Public Health
Peking University
No. 38 Xueyuan Road
Haidian District
Beijing, 100191
China
Phone: 86 10 8 2805146
Fax: 86 10 8 2805146
Email: liumin@bjmu.edu.cn

Abstract

Background: Depression is one of the most common mental illnesses, and it may have a lasting effect on one’s whole life. As a form of financial hardship, catastrophic health expenditure (CHE) may be associated with depression. However, current evidence about the relationship between CHE and the risk of depression is insufficient.

Objective: This study aimed to explore the relationship between CHE and the risk of depression among Chinese adults.

Methods: In this study, we used 3 waves of the China Family Panel Studies (CFPS) from 2012, 2016, and 2018. The CFPS are a nationally representative study covering 25 of 31 provinces in Chinese mainland and representing nearly 94.5% of the total population. We selected eligible household heads as participants, divided them into 2 groups by CHE events at baseline (exposed group: with CHE; unexposed group: without CHE), and followed them up. Households with CHE were defined as having out-of-pocket medical expenditures exceeding 40% of the total household nonfood expenditure, and people with depression were identified by the 8-item Centre for Epidemiological Studies Depression Scale (CES-D). We first described the baseline characteristics and used logistical regression to estimate their effects on CHE events. Then, we used Cox proportional hazard models to estimate adjusted hazard ratios and 95% CIs of depression among participants with CHE compared with those without CHE. Finally, we analyzed the subgroup difference in the association between CHE and depression.

Results: Of a total of 13,315 households, 9629 were eligible for analysis. Among them, 6824 (70.9%) were men. The mean age was 50.15 (SD 12.84) years. Only 987 (10.3%) participants had no medical insurance. The prevalence of CHE at baseline was 12.9% (1393/9629). Participants with a higher family economic level (adjusted odds ratio [aOR] 1.15, 95% CI 1.02-1.31) and with the highest socioeconomic development level (aOR 1.18, 95% CI 1.04-1.34) had a higher prevalence of CHE than reference groups. During a median of 71 (IQR 69-72) person-months of follow-up, the depression incidence of participants with CHE (1.41 per 1000 person-months) was higher than those without CHE (0.73 per 1000 person-months). Multivariable models revealed that the adjusted hazard ratio for the incidence of depression in participants with CHE was 1.33 (95% CI 1.08-1.64), and this association appeared to be greater in participants without outpatient services (for interaction, P=.048).

Conclusions: CHE was significantly associated with increased risk of depression among Chinese adults. Concentrated work should be done to monitor CHE, and more efforts to ensure financial protection need to be made to prevent depression, especially for people with high health care needs.

(JMIR Public Health Surveill 2023;9:e42469) doi:10.2196/42469

KEYWORDS

catastrophic health expenditure; depression; universal health coverage; economic burden; socioeconomic status

Introduction

There is a strong bidirectional linkage between health and poverty. To cut off this linkage, the United Nation's Sustainable Development Goals include target 3.8, which aims to achieve universal health coverage (UHC) by 2030 [1]. UHC means that all people can receive the health services they need without experiencing financial hardship [1]. According to the World Health Organization, globally in 2017, almost 1.4 billion people experienced financial hardship due to out-of-pocket (OOP) health payments, among whom nearly 1 billion people were pushed into extreme poverty [2]. Therefore, there is still a long way to go to achieve UHC goals by 2030 considering the additional impact of COVID-19.

One of the 2 essential parts of UHC is providing financial protection for people to pay for health services [3], and the government of China has made great progress on this front. For example, as early as 2009, China reformed a series of health care reforms including three national basic medical insurance programs: (1) Urban Employee Basic Medical Insurance (UEBMI) designed for employed urban residents; (2) the Urban Resident Basic Medical Insurance (URBMI) covering the unemployed, retired, older adults, students, and children in urban areas; and (3) the New Rural Cooperative Medical Scheme (NRCMS) for rural residents [4]. Currently, China's national basic insurance programs cover over 1.35 billion people, about 97% of the total population [5]. Additionally, to eliminate poverty, the Decision on Winning the Battle Against Poverty policy proposed by China in 2015 achieved substantial results in education, primary medical care, and the basic living needs of people living in poverty. In fact, in 2020, the government of China announced that all low-income counties in China had been lifted out of poverty [6]. Benefitting from policies and measures on financial protection of health, the incidence of catastrophic health expenditure (CHE) in China declined from 14.7% in 2010 to 8.7% in 2018 [7]. However, there are still a few people encountering financial hardship, and the impact on mental and physical health caused by CHE is still worth studying.

Depression is a prevalent mental illness globally that contributes to the global burden of diseases as the leading mental cause of mortality for all ages [8]. Once depression appears, it can have a lasting and profound influence on one's whole life. The direct outcome of depression is a poor quality of life [9]. A study conducted on patients with schizophrenia indicated that depression has a strong negative effect on all 8 domains of subjective quality of life [10]. Furthermore, a meta-analysis showed that compared to controls, patients with depression had significant moderate cognitive deficits in executive function, memory, and attention (Cohen *d* effect sizes ranging from −0.34 to −0.65) [11]. Moreover, depression is a negative factor in cardiovascular disease (CVD) incidence, severity, and outcomes [12]. Rajan et al [13] found that depression was associated with CVD incidence (hazard ratio [HR] 1.14, 95% CI 1.05-1.24) and myocardial infarction (HR 1.14, 95% CI 1.05-1.24). Meng et al [14] also found that depression was associated with a higher risk of CVD mortality (HR 1.22, 95% CI 1.04-1.44). Therefore, the prevention and management of depression is a crucial and urgent public health issue.

A series of social, psychological, and biological factors and their complex interactions can play a role in depression occurrence [15,16]. Some study results posit that lower social support, a lower socioeconomic position, and economic difficulties are associated with a higher risk of depression [17,18], and the association is stronger between financial hardship and depression than other socioeconomic variables [19]. CHE, in theory, may have an impact on the mental health of family members because of reduced necessary expenditures. Nevertheless, current studies mostly focus on the likelihood of CHE events among people with depression, not the impact of CHE on depression occurrence [20,21]. In this study, we used 3 waves (2012, 2016, and 2018) of nationally representative data from the China Family Panel Studies (CFPS) to analyze the association of CHE with the risk of depression.

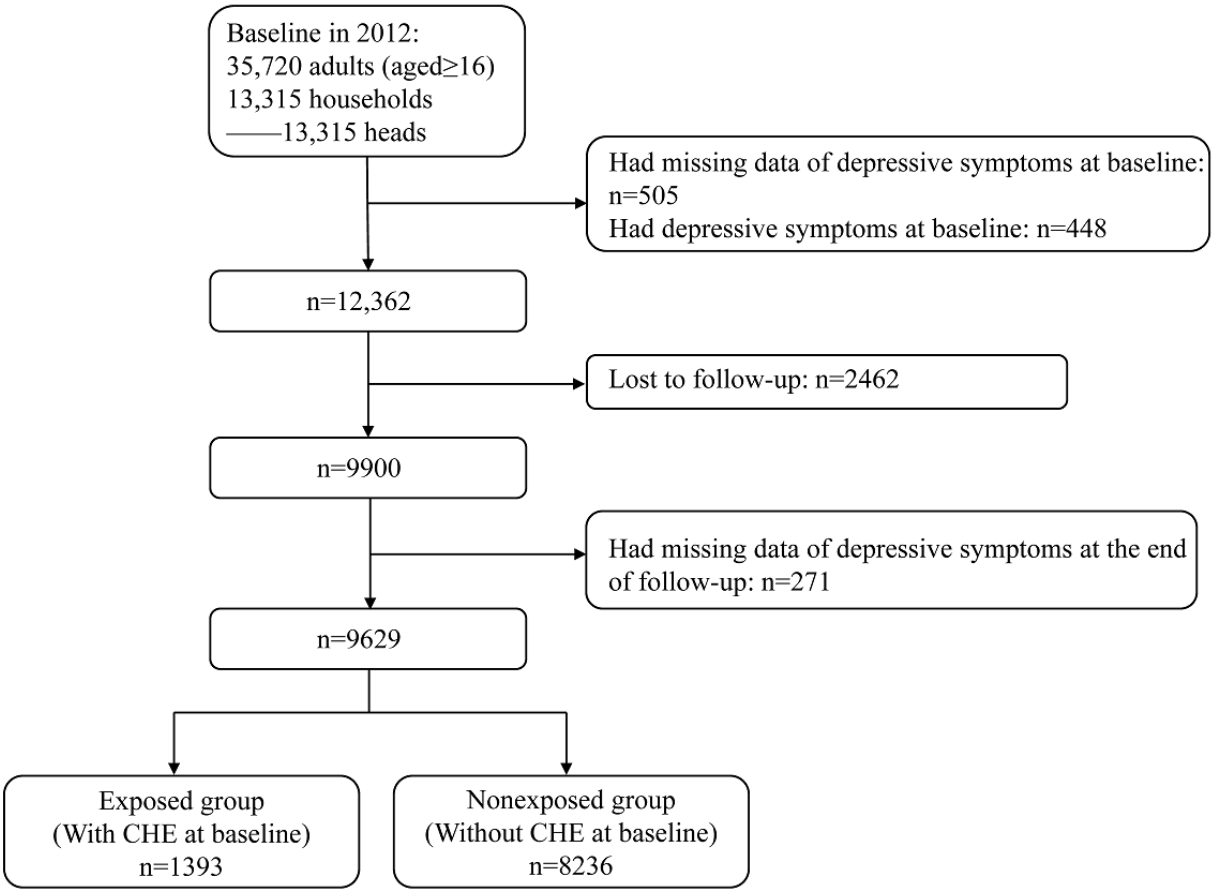
Methods

Study Design and Participants

Data in this study were obtained from CFPS, which is almost a nationally representative longitudinal study covering 25 of 31 provinces/municipalities in Chinese mainland (not including Xinjiang, Tibet, Inner Mongolia, Ningxia autonomous region, and Qinghai and Hainan provinces), representing nearly 94.5% of the total population in the Chinese mainland [22]. The CFPS were implemented by the Institute of Social Science Survey of Peking University to collect individual-, household-, and community-level data every 2 years. A baseline survey was conducted in 2010, and follow-up data from 2012, 2014, 2016, and 2018 were available for download from the official CFPS website [23].

As the information on depression in some of the CFPS waves was deficient, we used data from 2012, 2016, and 2018. The survey in 2012 included 35,720 adults (aged ≥16 years) and 13,315 households with valid interview responses. Individuals who had missing data at baseline (n=505), had depression at baseline (n=448), were lost to follow-up (n=2462), and had missing data of depression at the end of follow-up (n=271) were excluded. Finally, a total of 9629 households (household heads) were included in this study (Figure 1).

Figure 1. Flowchart of participants selected from China Family Panel Studies (CFPS). CHE: catastrophic health expenditure.



Ethics Approval

The CFPS, which involved human participants, were approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-14010). All written informed consent was provided by participants aged over 15 years or their parents (for those aged 15 years and under). The participants' personal information and privacy were strictly protected by the CFPS according to the rules set by Peking University's Biomedical Ethics Review Committee.

Measurements of Catastrophic Health Expenditure

The measurement of catastrophic health expenditure (CHE) was based on households, where a household member was defined by marriage, blood, or adoptive relationship, as well as an ongoing economic tie [24]. The head of the household was identified as the key decision-maker when the household faced important matters and decisions. Household OOP health payments were measured as the medical expenditure of all family members excluding reimbursed expenses in the past 12 months. The yearly household food expenditure was estimated as the monthly meal expenses multiplied by 12, and the total household expenditure in the past 12 months was calculated as the sum of monthly daily expenditures (food, daily used commodities and necessities, transportation, etc) multiplied by 12 plus yearly special expenditures (electricity, medical care, clothing, etc).

Households who experienced CHE were defined as those with OOP medical expenditures exceeding 40% of the household's capacity to pay (calculated as total household expenditure minus household food expenditure) [25]. CHE prevalence at baseline was measured as the percentage of the number of household heads incurring CHE to total participants. The formula was:

where N is the total number of participants, and CHE_i is 1 when the i_{th} household incurred CHE and 0 otherwise.

Measurements of Depression

Depression was measured by the 8-item Center for Epidemiological Studies Depression Scale (CES-D), which is a shortened version of the original 20-item CES-D including 8 depressive symptoms (Table S1 in Multimedia Appendix 1). CES-D was not designed as a diagnostic tool but is widely used to identify individuals at high risk of depression in general populations and various subpopulations [26]. The score of each item in the 20-item CES-D ranges from 0 to 3 according to the frequency response of depressive symptoms in the past week, with 0 indicating rarely or none of the time (<1 day), 1 indicating some or a little of the time (1-2 days), 2 indicating occasionally or a moderate amount of time (3-4 days), and 3 indicating most or all of the time (5-7 days) [26]. The commonly used 8-item CES-D asked respondents whether they experienced any of the 8 depressive symptoms "most of time during the last week"

[27], which was scored according to the answer (yes=1; no=0). The 8-item CES-D used in CFPS included the same depressive symptoms but was scored by frequency response. As "most or all of the time" in frequency response is equal to the "yes" answer to "most of time last week," we transferred the 4-score level of frequency response to the yes/no response [28]. Hence, the score of 8-item CES-D ranged from 0-8, and a total score ≥3 was deemed to indicate depression [28]. As the 8-item CES-D cannot diagnose depressive disorders, people with an 8-item CES-D score ≥3 in our study were identified as having depressive symptoms [29].

Covariates

Covariates in this study included (1) demographic characteristics: gender (male, female), age group (16-39, 40-49, 50-49, ≥60 years), marital status (married/partnered, other), education (no or some formal education, primary school, middle school, high school and above), and insurance (without any insurance, UEBMI, URBMI, NRCMS, other); (2) health-related characteristics: self-reported health (good, medium, poor), chronic diseases (yes, no), BMI (normal, lower, overweight, obese), outpatient services (yes, no), inpatient services (yes, no), current smoking (yes, no), and drinking (yes, no); and (3) socioeconomic characteristics: residence (urban, rural), family economic level (lowest, lower, higher, highest), family size (1-2, 3-4, ≥5), and socioeconomic development level (lowest, lower, higher, highest).

The family economic level was classified by the quartiles of household annual income, which was converted into 2012 US dollars based on purchasing power parities published by the World Bank [30] (lowest: <$4775.56, lower: $4775.56 to <$9831.46, higher: $9831.46 to <$17173.60, highest: ≥$17173.60). Numerous studies have reported that nighttime light intensity is highly associated with socioeconomic development [31,32]. Therefore, the socioeconomic development level in this study was identified by the quartiles of provinces' mean nighttime light intensity in 2012 (lowest: <0.089, lower: 0.089-0.31, higher: 0.32-0.68, highest: ≥0.69), which was obtained from the Harvard Dataverse [33].

Statistical Analyses

Baseline characteristics of the participants were presented as mean (SD) for continuous variables or frequencies and percentages for categorical variables. The Pearson χ^2 test was used to compare distributions of CHE according to different baseline characteristics. A multivariate logistical regression model was used to analyze the determinants of CHE. We calculated the incidence rates (number of events divided by accumulated person-month) and used the univariate and multivariate Cox proportional hazard models to estimate the HRs and 95% CIs of depression among participants with CHE compared with those without CHE. First, the model adjusted demographic characteristics in multivariate model 1, including gender, age group, marital status, education, and insurance. Next, the model adjusted health-related factors in model 2 based on model 1, including self-reported health, chronic diseases, BMI, outpatient services, inpatient services, current smoking, and current drinking. Finally, the model adjusted socioeconomic factors in model 3 based on model 2, including residence, family economic level, family size, and socioeconomic development level.

To examine the robustness of our findings, we conducted 3 sensitivity analyses. First, we defined households with CHE as those having OOP medical expenditure exceeding 25% of the total household expenditure according to the World Bank. [25] Second, we transferred the categorical variables of age group and family economic level into continuous variables and conducted the same analysis in the final model. Third, we changed the variable nighttime light intensity into the gross regional product (GRP, divided into 4 groups based on the quartiles of 2012 per capita GRP [34]) to indicate the socioeconomic development level in the final model.

Finally, the analysis was stratified by gender, age group, insurance, chronic diseases, self-reported health, outpatient services, inpatient services, residence, socioeconomic development level, and family economic level, and a multiplicative interaction term was included in the final model with the stratification variable removed for subgroup analysis.

All data were analyzed in R software (version 4.2.1; R Foundation for Statistical Computing). A 2-sided *P* value <.05 was considered to be significant.

Results

Baseline Characteristics

A total of 9629 participants were enrolled in this study, of which 6824 (70.9%) were male. The mean age was 50.15 (SD 12.84) years. Among the participants, 89.5% (8622/9629) were married or partnered. Only 987 (10.3%) participants had no medical insurance, and 1341 (13.9%), 603 (6.3%), and 6317 (65.6%) participants had UEBMI, URBMI, and NRCMS, respectively (Table 1).

Table 1. Distribution of CHE[a] by baseline characteristics.

Characteristics	Total (N=9629)	Nonexposed group[b] (n=8236), n (%)	Exposed group[c] (n=1393), n (%)	Chi-square (*df*)	*P* value
Demographic characteristics					
Gender				3.6 (1)	.06
Male	6824	5867 (86)	957 (14)		
Female	2805	2369 (84.5)	436 (15.5)		
Age group (years)				346.3 (3)	<.001
16-39	1931	1753 (90.8)	178 (9.2)		
40-49	3016	2750 (91.2)	266 (8.8)		
50-59	2320	1969 (84.9)	351 (15.1)		
≥60	2362	1764 (74.7)	598 (25.3)		
Marital status				9.1 (1)	.003
Married/partnered	8622	7407 (85.9)	1215 (14.1)		
Other	1007	829 (82.3)	178 (17.7)		
Education				97.3 (3)	<.001
No or some formal education	2506	2005 (80)	501 (20)		
Primary school	2322	1980 (85.3)	342 (14.7)		
Middle school	2789	2475 (88.7)	314 (11.3)		
High school and above	2012	1776 (88.3)	236 (11.7)		
Insurance				5.5 (4)	.24
None	987	856 (86.7)	131 (13.3)		
UEBMI[d]	1341	1169 (87.2)	172 (12.8)		
URBMI[e]	603	511 (84.7)	92 (15.3)		
NRCMS[f]	6317	5378 (85.1)	939 (14.9)		
Other	381	322 (84.5)	59 (15.5)		
Health-related characteristics					
Self-reported health				245.5 (2)	<.001
Good	5874	5241 (89.2)	633 (10.8)		
Medium	2001	1692 (84.6)	309 (15.4)		
Poor	1754	1303 (74.3)	451 (25.7)		
Currently smoking				12.6 (1)	<.001
No	5355	4519 (84.4)	836 (15.6)		
Yes	4274	3717 (87)	557 (13)		
Drinking				16.4 (1)	<.001
No	7282	6168 (84.7)	1114 (15.3)		
Yes	2347	2068 (88.1)	279 (11.9)		
Chronic diseases				104 (1)	<.001
No	8268	7195 (87)	1073 (13)		
Yes	1361	1041 (76.5)	320 (23.5)		
BMI				25.8 (3)	<.001
Normal	5596	4776 (85.3)	820 (14.7)		
Lower	711	569 (80)	142 (20)		

Characteristics	Total (N=9629)	Nonexposed group[b] (n=8236), n (%)	Exposed group[c] (n=1393), n (%)	Chi-square (*df*)	*P* value
Overweight	2658	2301 (86.6)	357 (13.4)		
Obese	664	590 (88.9)	74 (11.1)		
Outpatient services				101.8 (1)	<.001
No	7577	6624 (87.4)	953 (12.6)		
Yes	2052	1612 (78.6)	440 (21.4)		
Inpatient services				164.4 (1)	<.001
No	8747	7644 (87.4)	1103 (12.6)		
Yes	882	592 (67.1)	290 (32.9)		
Socioeconomic characteristics					
Residence				11.7 (1)	<.001
Urban	4315	3750 (86.9)	565 (13.1)		
Rural	5314	4486 (84.4)	828 (15.6)		
Family economic level				55.8 (3)	<.001
Lowest	2988	2439 (81.6)	549 (18.4)		
Lower	2336	2021 (86.5)	315 (13.5)		
Higher	2366	2084 (88.1)	282 (11.9)		
Highest	1939	1692 (87.3)	247 (12.7)		
Family size				151.5 (2)	<.001
1-2	2158	1677 (77.7)	481 (22.3)		
3-4	4302	3833 (89.1)	469 (10.9)		
≥5	3169	2726 (86)	443 (14)		
Socioeconomic development level				6.4 (3)	.09
Lowest	2491	2158 (86.6)	333 (13.4)		
Lower	1492	1258 (84.3)	234 (15.7)		
Higher	3726	3199 (85.9)	527 (14.1)		
Highest	1920	1621 (84.4)	299 (15.6)		

[a] CHE: catastrophic health expenditure.

[b] People without CHE.

[c] People with CHE.

[d] UEBMI: Urban Employee Basic Medical Insurance.

[e] URBMI: Urban Resident Basic Medical Insurance.

[f] NRCMS: New Rural Cooperative Medical Scheme.

CHE Prevalence

At baseline, the prevalence of CHE was 12.9% (1393/9629) among the participants. Except for gender, insurance, and socioeconomic development level, the distribution of baseline characteristics was significantly different between households with CHE and those without CHE (Table 1). The logistical regression analysis revealed that participants with poor (adjusted odds ratio [aOR] 1.64, 95% CI 1.39-1.94) and medium self-reported health (aOR 1.17, 95% CI 1-1.37), chronic diseases (aOR 1.30, 95% CI 1.11-1.52), outpatient services (aOR 1.16, 95% CI 1-1.35) and inpatient services (aOR 2.40, 95% CI 2.03-2.84), rural residence (aOR 1.21, 95% CI 1.05-1.39), family size ≥5 people (aOR 1.22, 95% CI 1.09-1.36), higher family economic level (aOR 1.15, 95% CI 1.02-1.31), and the highest socioeconomic development level (aOR 1.18, 95% CI 1.04-1.34) had a higher prevalence of CHE than reference groups (Table S2 in Multimedia Appendix 1).

Risk of Depression

During a median of 71 (interquartile range: 69-72) person-months of follow-up, 532 (5.53%) of 9629 participants developed depression, of which 403 (75.75%) cases had no CHE and 129 (24.25%) cases had CHE. The incidence rate of depression among participants without and with CHE was 0.73 and 1.41 per 1000 person-months, respectively (Table 2). In the unadjusted analysis, participants who had CHE at baseline were associated with a 99% higher risk of depression (crude HR 1.99, 95% CI 1.63-2.42; Table 2). All multivariable–adjusted analyses showed a significant association of CHE with the risk of depression (Table S3 in Multimedia Appendix 1). In the fully adjusted model, participants with CHE had a 33% increased risk of developing depression compared to those without CHE (adjusted HR [aHR] 1.33, 95% CI 1.08-1.64; Table 2).

In the sensitivity analyses, the association between CHE and risk of depression was stable when we (1) defined CHE as OOP medical expenditure exceeding 25% of the total household expenditure, (2) transferred the categorical variables age group and family economic level to continuous variables, and (3) changed the variable nightlight time intensity into GRP to indicate socioeconomic development level in the final model (Table S4 in Multimedia Appendix 1).

Table 2. Association of CHE[a] with risk of depression in the univariate and multivariate Cox proportional hazard models.

Characteristics	Outcome (N=532), n (%)	Incidence (per 1000 person-months)	Univariate model		Multivariate model[b]	
			cHR[c] (95% CI)	*P* value	aHR[d] (95% CI)	*P* value
CHE						
Without	403 (75.8)	0.73	Reference		Reference	
With	129 (24.2)	1.41	1.99 (1.63-2.42)	<.001	1.33 (1.08-1.64)	.008
Gender						
Male	318 (59.8)	0.70	Reference		Reference	
Female	214 (40.2)	1.15	1.69 (1.42-2.01)	<.001	1.51 (1.19-1.91)	.001
Age group (years)						
16-39	50 (9.4)	0.39	Reference		Reference	
40-49	144 (27.1)	0.71	2.39 (1.93-2.97)	<.001	1.61 (1.27-2.06)	<.001
50-59	154 (28.9)	0.99	0.84 (0.69-1.02)	.08	0.78 (0.64-0.96)	.02
≥60	184 (34.6)	1.20	1.03 (0.87-1.22)	.70	0.97 (0.82-1.15)	.74
Marital status						
Married/partnered	433 (81.4)	0.75	Reference		Reference	
Other	99 (18.6)	1.54	2.22 (1.78-2.76)	<.001	1.34 (1.05-1.70)	.02
Education						
No or some formal education	255 (47.9)	1.56	Reference		Reference	
Primary school	121 (22.7)	0.78	0.36 (0.29-0.44)	<.001	0.63 (0.50-0.80)	<.001
Middle school	100 (18.8)	0.53	1.27 (1.04-1.54)	.02	1.18 (0.96-1.44)	.12
High school and above	56 (10.5)	0.42	0.95 (0.79-1.15)	.61	1.01 (0.83-1.22)	.95
Insurance						
None	71 (13.3)	1.11	Reference		Reference	
UEBMI[e]	35 (6.6)	0.39	0.35 (0.23-0.52)	<.001	0.52 (0.34-0.79)	.002
URBMI[f]	23 (4.3)	0.57	0.51 (0.32-0.81)	.004	0.53 (0.33-0.85)	.009
NRCMS[g]	394 (74.1)	0.93	0.82 (0.64-1.06)	.13	0.68 (0.52-0.89)	.004
Other	9 (1.7)	0.35	0.30 (0.15-0.61)	.001	0.43 (0.21-0.88)	.02
Self-reported health						
Good	201 (37.8)	0.51	Reference		Reference	
Medium	110 (20.7)	0.83	1.61 (1.28-2.03)	<.001	1.32 (1.04-1.68)	.02
Poor	221 (41.5)	1.93	3.89 (3.21-4.71)	<.001	2.10 (1.67-2.64)	<.001
Currently smoking						
No	314 (59.0)	0.88	Reference		Reference	
Yes	218 (41.0)	0.77	0.86 (0.72-1.02)	.08	1.14 (0.92-1.42)	.24
Drinking						
No	435 (81.8)	0.90	Reference		Reference	
Yes	97 (18.2)	0.62	0.67 (0.54-0.84)	.001	0.88 (0.69-1.11)	.29
Chronic diseases						
No	419 (78.8)	0.76	Reference		Reference	
Yes	113 (21.2)	1.25	1.62 (1.31-1.99)	<.001	1.17 (0.94-1.46)	.15

Characteristics	Outcome (N=532), n (%)	Incidence (per 1000 person-months)	Univariate model		Multivariate model[b]	
			cHR[c] (95% CI)	*P* value	aHR[d] (95% CI)	*P* value
BMI						
Normal	311 (58.5)	0.84	Reference		Reference	
Lower	65 (12.2)	1.41	1.75 (1.34-2.28)	<.001	1.13 (0.86-1.48)	.40
Overweight	118 (22.2)	0.66	0.79 (0.64-0.98)	.03	0.92 (0.74-1.13)	.42
Obese	38 (7.1)	0.85	1.05 (0.75-1.47)	.78	1.13 (0.80-1.58)	.50
Outpatient services						
No	326 (61.3)	0.64	Reference		Reference	
Yes	206 (38.7)	1.52	2.36 (1.98-2.81)	<.001	1.41 (1.16-1.72)	.001
Inpatient services						
No	477 (89.7)	0.82	Reference		Reference	
Yes	55 (10.3)	0.95	1.18 (0.89-1.56)	.24	0.70 (0.52-0.94)	.02
Residence						
Urban	173 (32.5)	0.61	Reference		Reference	
Rural	359 (67.5)	1.01	1.65 (1.37-1.98)	<.001	1.23 (1-1.52)	.06
Family economic level						
Lowest	240 (45.1)	1.21	Reference		Reference	
Lower	109 (20.5)	0.70	0.55 (0.46-0.67)	<.001	0.75 (0.61-0.93)	.008
Higher	116 (21.8)	0.74	1.11 (0.92-1.34)	.28	1.03 (0.85-1.24)	.79
Highest	67 (12.6)	0.52	0.79 (0.66-0.96)	.02	0.83 (0.69-1)	.049
Family size						
1-2	165 (31)	1.17	Reference		Reference	
3-4	200 (37.6)	0.69	0.72 (0.62-0.84)	<.001	0.96 (0.81-1.15)	.69
≥5	167 (31.4)	0.79	1.34 (1.16-1.55)	<.001	1.02 (0.88-1.19)	.78
Socioeconomic development level						
Lowest	174 (32.7)	1.05	Reference		Reference	
Lower	92 (17.3)	0.93	0.69 (0.57-0.83)	.001	1.01 (0.78-1.32)	.92
Higher	186 (35.0)	0.75	0.92 (0.77-1.11)	.39	0.86 (0.69-1.06)	.16
Highest	80 (15.0)	0.63	1.01 (0.84-1.20)	.94	0.73 (0.55-0.97)	.03

[a] CHE: catastrophic health expenditure.

[b] Adjusted for demographic characteristics (gender, age group, education, marital status, and insurance), health-related characteristics (self-reported health, smoking status, drinking, chronic disease, BMI, and outpatient and inpatient services), and socioeconomic characteristics (residence, family economic level, family size, and socioeconomic development level).

[c] cHR: crude hazard ratio.

[d] aHR: adjusted hazard ratio.

[e] UEBMI: Urban Employee Basic Medical Insurance.

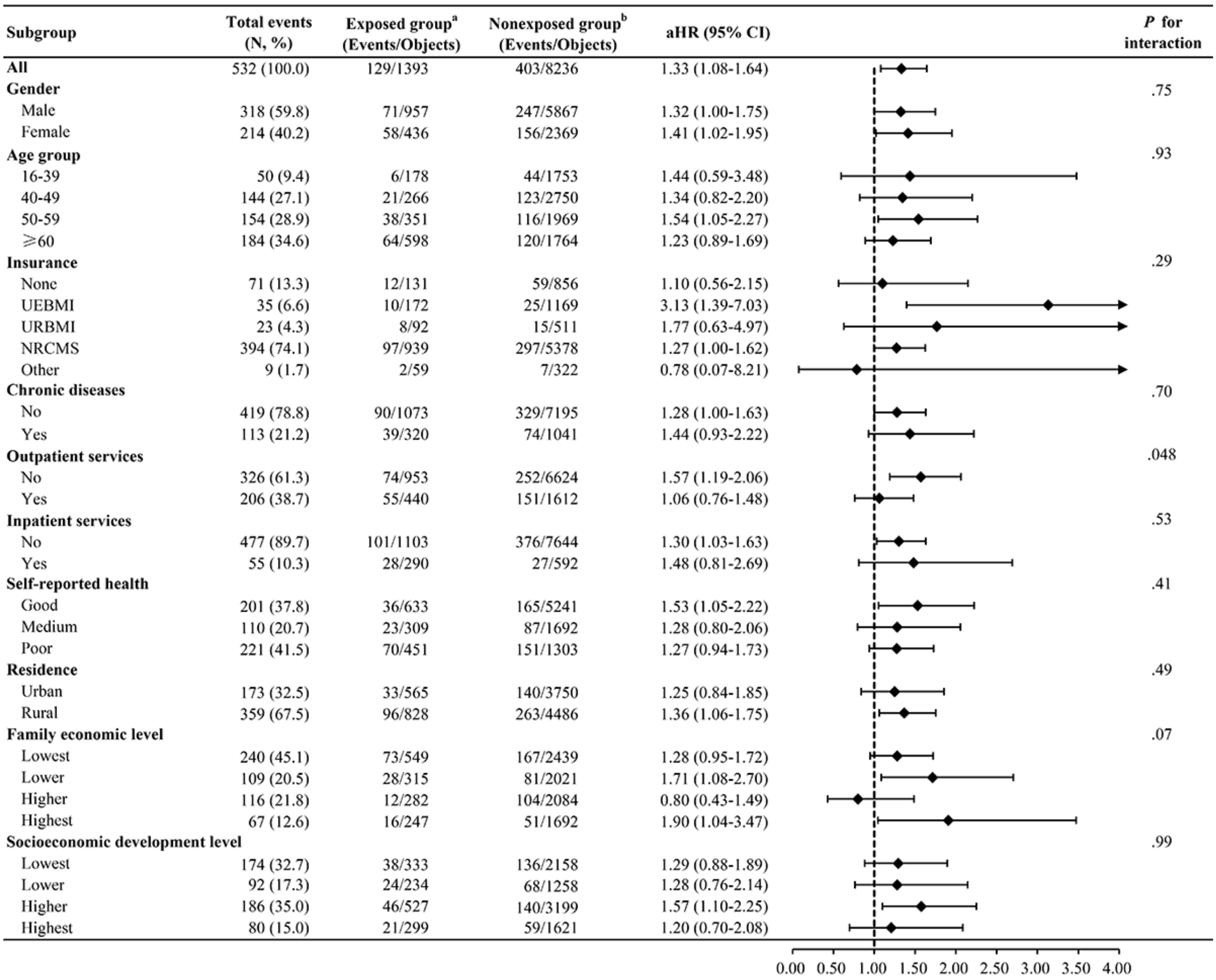
[f] URBMI: Urban Resident Basic Medical Insurance.

[g] NRCMS: New Rural Cooperative Medical Scheme.

Subgroup Analyses

In the subgroup analyses, the association of CHE with the risk of depression did not appear to be modified by most selected baseline characteristics, such as gender, age group, insurance, or chronic diseases (Figure 2). However, the association between CHE and the risk of depression was found to be somewhat greater among participants without outpatient services (for interaction, *P*=.048). The risk of depression was higher among participants with CHE who did not have outpatient services (aHR 1.57, 95% CI 1.10-2.25) than those who had outpatient services (aHR 1.06, 95% CI 0.76-1.48).

Figure 2. Association of catastrophic health expenditure (CHE) with risk of depression in univariate and multivariate Cox proportional hazard models. Superscripted a refers to people with CHE. Superscripted b refers to people without CHE. aHR: adjusted hazard ratio; NRCMS: New Rural Cooperative Medical Scheme; UEBMI: Urban Employee Basic Medical Insurance; URBMI: Urban Resident Basic Medical Insurance.



Subgroup	Total events (N, %)	Exposed group[a] (Events/Objects)	Nonexposed group[b] (Events/Objects)	aHR (95% CI)	P for interaction
All	532 (100.0)	129/1393	403/8236	1.33 (1.08-1.64)	
Gender					.75
Male	318 (59.8)	71/957	247/5867	1.32 (1.00-1.75)	
Female	214 (40.2)	58/436	156/2369	1.41 (1.02-1.95)	
Age group					.93
16-39	50 (9.4)	6/178	44/1753	1.44 (0.59-3.48)	
40-49	144 (27.1)	21/266	123/2750	1.34 (0.82-2.20)	
50-59	154 (28.9)	38/351	116/1969	1.54 (1.05-2.27)	
⩾60	184 (34.6)	64/598	120/1764	1.23 (0.89-1.69)	
Insurance					.29
None	71 (13.3)	12/131	59/856	1.10 (0.56-2.15)	
UEBMI	35 (6.6)	10/172	25/1169	3.13 (1.39-7.03)	
URBMI	23 (4.3)	8/92	15/511	1.77 (0.63-4.97)	
NRCMS	394 (74.1)	97/939	297/5378	1.27 (1.00-1.62)	
Other	9 (1.7)	2/59	7/322	0.78 (0.07-8.21)	
Chronic diseases					.70
No	419 (78.8)	90/1073	329/7195	1.28 (1.00-1.63)	
Yes	113 (21.2)	39/320	74/1041	1.44 (0.93-2.22)	
Outpatient services					.048
No	326 (61.3)	74/953	252/6624	1.57 (1.19-2.06)	
Yes	206 (38.7)	55/440	151/1612	1.06 (0.76-1.48)	
Inpatient services					.53
No	477 (89.7)	101/1103	376/7644	1.30 (1.03-1.63)	
Yes	55 (10.3)	28/290	27/592	1.48 (0.81-2.69)	
Self-reported health					.41
Good	201 (37.8)	36/633	165/5241	1.53 (1.05-2.22)	
Medium	110 (20.7)	23/309	87/1692	1.28 (0.80-2.06)	
Poor	221 (41.5)	70/451	151/1303	1.27 (0.94-1.73)	
Residence					.49
Urban	173 (32.5)	33/565	140/3750	1.25 (0.84-1.85)	
Rural	359 (67.5)	96/828	263/4486	1.36 (1.06-1.75)	
Family economic level					.07
Lowest	240 (45.1)	73/549	167/2439	1.28 (0.95-1.72)	
Lower	109 (20.5)	28/315	81/2021	1.71 (1.08-2.70)	
Higher	116 (21.8)	12/282	104/2084	0.80 (0.43-1.49)	
Highest	67 (12.6)	16/247	51/1692	1.90 (1.04-3.47)	
Socioeconomic development level					.99
Lowest	174 (32.7)	38/333	136/2158	1.29 (0.88-1.89)	
Lower	92 (17.3)	24/234	68/1258	1.28 (0.76-2.14)	
Higher	186 (35.0)	46/527	140/3199	1.57 (1.10-2.25)	
Highest	80 (15.0)	21/299	59/1621	1.20 (0.70-2.08)	

Discussion

CHE can introduce great financial pressure to a household, as household members may have to borrow money, use savings, and even sell assets [35], which may lead to a longtime negative effect on their mental health [36]. To our knowledge, this is the first cohort study on the association between CHE and the risk of depression. We found that health-related characteristics (chronic diseases, self-reported health, and outpatient and inpatient services) and socioeconomic characteristics (residence, socioeconomic development level, family size, and family economic level) had an impact on CHE. We also found that prior CHE events were strongly associated with depression experienced by household heads and that there was an interaction between CHE and outpatient services for depression.

Our results revealed that people with chronic diseases, medium and poor self-reported health, outpatient services and inpatient services, rural, family size ≥5 people, and the highest socioeconomic development level had a higher prevalence of CHE, and these findings are consistent with those of other studies [7,37]. Poor health directly leads to a higher frequency of accessing health services, further leads to a heavy financial burden of medical expenditure, and finally, leads to CHE. Nighttime light intensity is an indicator of socioeconomic development, and areas with higher economic growth may have access to diagnostic technology for most illnesses. In addition, people in areas undergoing rapid socioeconomic development may pay more attention to the early detection of diseases and other preventive health care services [38]. Consequently, health-related characteristics and socioeconomic characteristics were the main factors behind CHE.

The results of our study revealed that people with CHE had a 33% higher risk of developing depression (aHR 1.33, 95% CI 1.08-1.64) than people without CHE. Socioeconomic variables, especially financial hardship, have a great influence on depression [39]. As a form of financial hardship, CHE plays a role in depression incidence. Once a household incurred CHE, family members had to take various measures to cope with financial loss, which could take a long time to recover [35]. Given this, household heads had a higher risk of developing depression with CHE than those without CHE under financial pressure.

As CHE could trigger depression, which may further lead to adverse outcomes [40,41], it is crucial to eliminate CHE to prevent depression. The first step to eliminating CHE is based on reducing OOP health expenditures. Because poor health is associated with an increased need and use of outpatient and inpatient care, and this negative correlation can be intensified by the number of chronic diseases [37], our results suggest that relevant departments should continue to expand health promotion, especially for chronic diseases. Additionally, governments should provide sufficient and targeted financial protection, including medical insurance and poverty subsidies, to reduce OOP payments and mitigate impoverishment caused by poor health. Most studies on CHE are interested in subpopulations that have specialized diseases like cancers, HIV, and tuberculosis [42-44]. The common features of these diseases are long duration, which may even coexist with people and cause a higher risk of CHE. To decrease individual OOP health expenditure, inclined policies and measures should be introduced for these diseases. For example, to relieve the financial burden faced by people living with chronic diseases, China launched catastrophic medical insurance (critical illness insurance) in 2012 and implemented it nationwide in 2016 after city-based testing, with the aim of reimbursing patients whose OOP health expenditure exceeded a predetermined basic medical insurance level [45]. In addition, for eligible people facing unaffordable medical expenses, China provided medical assistance and even offered treatment for free or at a reduced price for some priority diseases [46]. Of course, these policies are mostly aimed at financially vulnerable people. However, our study found that people with a higher family economic level and the highest socioeconomic development level were more likely to occur CHE, which may be explained by the concentrated distribution and utilization of health services. [47] In view of this, more financial protection policies should be implemented for the whole population regardless of their economic capacity. Additionally, as the current health care delivery in China is still fragmented and treatment-based [48], with increasing incidence of age-related diseases, integrated establishment and improvement of primary health services should be intensified in health-related policies.

To cut off the linkage between CHE and depression, the second point is to halt the progression to depression. As the evidence shows in studies on financial hardship and depression, not all people experiencing financial hardship will develop depression [39]. These differing responses to CHE may be explained by various socioeconomic and psychological variables in the stress process model, such as social support, self-esteem, personal agency, and personal ability to manage difficulties [39,49,50]. Therefore, measures to prevent and regulate depression should be extended widely. Additionally, timely financial assistance should be implemented for vulnerable populations, such as people with chronic diseases and households under the minimum living guarantee.

In this study, we found that the risk of depression was higher among participants with CHE who did not have outpatient services (aHR 1.57, 95% CI 1.10-2.25) than those who had outpatient services (aHR 1.06, 95% CI 0.76-1.48). In CFPS, the variable “outpatient services” was measured by whether participants had used outpatient services in the past 2 weeks when they felt uncomfortable. Those without outpatient services included people who did not feel uncomfortable and who felt uncomfortable but did not use health services. The latter group may have a lower socioeconomic status. Moreover, heads of households with lower social support, a lower socioeconomic position, and economic difficulties were more likely to experience depression due to life pressures [17,18]. Consequently, not accessing outpatient services intensified the association between CHE and the risk of depression.

There are several limitations to our study. First, because the 8-item CES-D is not equal to clinical diagnosis, our study could only estimate the association between CHE and the risk of depressive symptoms, which calls for future research on CHE and depressive disorders. Second, due to the limitation of the original data, households with CHE did not include those facing extreme poverty who could not seek health services and whose household medical expenditure was zero. Third, expenditure data for calculating CHE were mainly based on the participants’ memory, which may be prone to recall bias.

In conclusion, people with poor health and a higher socioeconomic position had a higher prevalence of CHE and CHE was significantly associated with a higher risk of depression. To prevent depression induced by CHE, concentrated work should be made to monitor CHE, and more efforts to ensure financial protection need to be introduced and strengthened, especially for people with higher health care needs.

Acknowledgments

We appreciate the China Family Panel Studies (CFPS) for conducting the total survey and providing data. We also thank all the interviewers, respondents, and volunteers in the survey. Moreover, we acknowledge the helpful advice of the editors and reviewers. This study was supported by the National Natural Science Foundation of China (72122001, 71934002).

Data Availability

Data from all are available to download from the official China Family Panel Studies (CFPS) website.

Authors' Contributions

JL conceptualized and designed the study. YW carried out the literature search, data analysis, and interpretation; compiled the tables and figures; and drafted the manuscript. WL, ML, and JL wrote, reviewed, and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables.
[DOCX File , 41 KB - publichealth_v9i1e42469_app1.docx]

References

1. Universal health coverage. World Health Organization. URL: https://www.who.int/health-topics/universal-health-coverage#tab=tab_ [accessed 2022-08-06]
2. Tracking universal health coverage: global monitoring report. World Health Organization and World Bank. URL: https://cdn.who.int/media/docs/default-source/world-health-data-platform/events/tracking-universal-health-coverage-2021-global-monitoring-report_uhc-day.pdf?sfvrsn=fd5c65c6_5&download=true [accessed 2022-08-06]
3. Universal health coverage (UHC): fact sheets. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-(uhc) [accessed 2022-08-06]
4. Sun J, Lyu S. The effect of medical insurance on catastrophic health expenditure: evidence from China. Cost Eff Resour Alloc 2020 Feb 27;18:10 [FREE Full text] [doi: 10.1186/s12962-020-00206-y] [Medline: 32127784]
5. Yin W. To accelerate the development of people’s health and promote the innovation of medical insurance system. People's Daily. URL: http://health.people.com.cn/n1/2016/1010/c398004-28764621.html [accessed 2022-08-06]
6. Poverty alleviation: China's experience and contribution. The State Council Information Office of the People's Republic of China. URL: http://www.scio.gov.cn/ztk/dtzt/44689/45216/index.htm [accessed 2022-08-07]
7. Liu C, Liu Z, Nicholas S, Wang J. Trends and determinants of catastrophic health expenditure in China 2010-2018: a national panel data analysis. BMC Health Serv Res 2021 May 29;21(1):526 [FREE Full text] [doi: 10.1186/s12913-021-06533-x] [Medline: 34051762]
8. GBD 2019 DiseasesInjuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
9. Ruo B, Rumsfeld JS, Hlatky MA, Liu H, Browner WS, Whooley MA. Depressive symptoms and health-related quality of life: the Heart and Soul Study. JAMA 2003 Jul 9;290(2):215-221 [FREE Full text] [doi: 10.1001/jama.290.2.215] [Medline: 12851276]
10. Reine G, Lançon C, Di Tucci S, Sapin C, Auquier P. Depression and subjective quality of life in chronic phase schizophrenic patients. Acta Psychiatr Scand 2003 Oct;108(4):297-303. [doi: 10.1034/j.1600-0447.2003.00132.x] [Medline: 12956831]
11. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. Psychol Med 2014 Jul;44(10):2029-2040. [doi: 10.1017/S0033291713002535] [Medline: 24168753]
12. Elderon L, Whooley MA. Depression and cardiovascular disease. Prog Cardiovasc Dis 2013 May;55(6):511-523. [doi: 10.1016/j.pcad.2013.03.010] [Medline: 23621961]
13. Rajan S, McKee M, Rangarajan S, Bangdiwala S, Rosengren A, Gupta R, Prospective Urban Rural Epidemiology (PURE) Study Investigators. Association of symptoms of depression with cardiovascular disease and mortality in low-, middle-, and high-income countries. JAMA Psychiatry 2020 Oct 01;77(10):1052-1063 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.1351] [Medline: 32520341]
14. Meng R, Yu C, Liu N, He M, Lv J, Guo Y, China Kadoorie Biobank Collaborative Group. association of depression with all-cause and cardiovascular disease mortality among adults in China. JAMA Netw Open 2020 Feb 05;3(2):e1921043 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.21043] [Medline: 32049295]
15. Heim C, Binder EB. Current research trends in early life stress and depression: review of human studies on sensitive periods, gene-environment interactions, and epigenetics. Exp Neurol 2012 Jan;233(1):102-111. [doi: 10.1016/j.expneurol.2011.10.032] [Medline: 22101006]
16. Malhi GS, Mann JJ. Depression. Lancet 2018 Nov;392(10161):2299-2312. [doi: 10.1016/s0140-6736(18)31948-2]
17. Langer, Crockett MA, Bravo-Contreras M, Carrillo-Naipayan C, Chaura-Marió M, Gómez-Curumilla B, et al. Social and Economic Factors Associated With Subthreshold and Major Depressive Episode in University Students During the COVID-19 Pandemic. Front Public Health 2022;10:893483 [FREE Full text] [doi: 10.3389/fpubh.2022.893483] [Medline: 35664111]
18. Wei D, Au Yeung SL, He J, Xiao W, Lu J, Tu S, et al. The role of social support in family socio-economic disparities in depressive symptoms during early pregnancy: Evidence from a Chinese birth cohort. J Affect Disord 2018 Oct 01;238:418-423. [doi: 10.1016/j.jad.2018.06.014] [Medline: 29913379]
19. Butterworth P, Olesen SC, Leach LS. The role of hardship in the association between socio-economic position and depression. Aust N Z J Psychiatry 2012 Apr 05;46(4):364-373. [doi: 10.1177/0004867411433215] [Medline: 22508596]

20. Hailemichael Y, Hanlon C, Tirfessa K, Docrat S, Alem A, Medhin G, et al. Catastrophic health expenditure and impoverishment in households of persons with depression: a cross-sectional, comparative study in rural Ethiopia. BMC Public Health 2019 Jul 11;19(1):930 [FREE Full text] [doi: 10.1186/s12889-019-7239-6] [Medline: 31296207]
21. Yan R, Li L, Duan X, Zhao J. Association of depressive symptoms with health service use and catastrophic health expenditure among middle-aged and older Chinese adults: analysis of population-based panel data. J Am Med Dir Assoc 2023 May;24(5):664-671.e7. [doi: 10.1016/j.jamda.2022.11.018] [Medline: 36574953]
22. Xie Y, Lu P. The sampling design of the China Family Panel Studies (CFPS). Chin J Sociol 2015 Dec;1(4):471-484 [FREE Full text] [doi: 10.1177/2057150X15614535] [Medline: 29854418]
23. China Family Panel Studies. URL: http://www.isss.pku.edu.cn/cfps/ [accessed 2022-08-06]
24. Xie Y, Hu J. An Introduction to the China Family Panel Studies (CFPS). Chinese Sociological Review Sep 2014;47(1):3-29. [doi: 10.2753/CSA2162-0555470101.2014.11082908]
25. Cylus J, Thomson S, Evetovits T. Catastrophic health spending in Europe: equity and policy implications of different calculation methods. Bull World Health Organ 2018 Sep 01;96(9):599-609 [FREE Full text] [doi: 10.2471/BLT.18.209031] [Medline: 30262941]
26. Smarr KL, Keefer AL. Measures of depression and depressive symptoms. Arthritis Care Res (Hoboken) 2020 Oct;72 Suppl 10:608-629. [doi: 10.1002/acr.24191] [Medline: 33091258]
27. Turvey CL, Wallace RB, Herzog R. A revised CES-D measure of depressive symptoms and a DSM-based measure of major depressive episodes in the elderly. Int Psychogeriatr 1999 Jun;11(2):139-148. [doi: 10.1017/s1041610299005694] [Medline: 11475428]
28. Wallace R, Herzog A, Ofstedal M, Steffick D, Fonda S, Langa K. Documentation of affective functioning measures in the health and retirement study. HRS Online. URL: https://hrsonline.isr.umich.edu/sitedocs/userg/dr-005.pdf [accessed 2022-08-05]
29. Litwin H. The association between social network relationships and depressive symptoms among older Americans: what matters most? Int Psychogeriatr 2011 Aug;23(6):930-940. [doi: 10.1017/S1041610211000251] [Medline: 21356159]
30. PPP conversion factor, GDP (LCU per international $) - China. World Bank. URL: https://data.worldbank.org/indicator/PA.NUS.PPP?locations=CN&most_recent_year_desc=true&view=chart [accessed 2023-05-20]
31. Beyer RCM, Franco-Bedoya S, Galdo V. Examining the economic impact of COVID-19 in India through daily electricity consumption and nighttime light intensity. World Dev 2021 Apr;140:105287 [FREE Full text] [doi: 10.1016/j.worlddev.2020.105287] [Medline: 34305264]
32. Zhang G, Guo X, Li D, Jiang B. Evaluating the potential of LJ1-01 nighttime light data for modeling socio-economic parameters. Sensors (Basel) 2019 Mar 26;19(6) [FREE Full text] [doi: 10.3390/s19061465] [Medline: 30917491]
33. Chen Z, Yu B, Yang C, Zhou Y, Yao S, Qian X. An extended time-series (2000-2018) of global NPP-VIIRS-like nighttime light data. Harvard Dataverse. 2020. URL: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/YGIVCD [accessed 2022-08-04]
34. National Bureau of Statistics of the People's Republic of China. China Statistical Yearbook 2012. Beijing, China: China Statistical Publishing House; 2013:57.
35. Kasahun GG, Gebretekle GB, Hailemichael Y, Woldemariam AA, Fenta TG. Catastrophic healthcare expenditure and coping strategies among patients attending cancer treatment services in Addis Ababa, Ethiopia. BMC Public Health 2020 Jun 22;20(1):984 [FREE Full text] [doi: 10.1186/s12889-020-09137-y] [Medline: 32571275]
36. Kiely KM, Leach LS, Olesen SC, Butterworth P. How financial hardship is associated with the onset of mental health problems over time. Soc Psychiatry Psychiatr Epidemiol 2015 Jun 17;50(6):909-918. [doi: 10.1007/s00127-015-1027-0] [Medline: 25683473]
37. Zhao Y, Atun R, Oldenburg B, McPake B, Tang S, Mercer SW, et al. Physical multimorbidity, health service use, and catastrophic health expenditure by socioeconomic groups in China: an analysis of population-based panel data. Lancet Glob Health 2020 Jun;8(6):e840-e849. [doi: 10.1016/s2214-109x(20)30127-3]
38. Tang K, Zhang Y, Wang H, Tan SH, Bai L, Liu Y. Regional economic development, household income, gender and hypertension: evidence from half a million Chinese. BMC Public Health 2020 Jun 10;20(1):901 [FREE Full text] [doi: 10.1186/s12889-020-09002-y] [Medline: 32522178]
39. Frankham C, Richardson T, Maguire N. Psychological factors associated with financial hardship and mental health: A systematic review. Clin Psychol Rev 2020 Apr;77:101832. [doi: 10.1016/j.cpr.2020.101832] [Medline: 32088498]
40. Yoshikawa H, Aber JL, Beardslee WR. The effects of poverty on the mental, emotional, and behavioral health of children and youth: implications for prevention. Am Psychol 2012;67(4):272-284. [doi: 10.1037/a0028015] [Medline: 22583341]
41. Ridley M, Rao G, Schilbach F, Patel V. Poverty, depression, and anxiety: causal evidence and mechanisms. Science 2020 Dec 11;370(6522). [doi: 10.1126/science.aay0214] [Medline: 33303583]
42. Sun C, Shi J, Fu W, Zhang X, Liu G, Chen W, et al. Catastrophic health expenditure and its determinants among households with breast cancer patients in China: a multicentre, cross-sectional survey. Front Public Health 2021;9:704700 [FREE Full text] [doi: 10.3389/fpubh.2021.704700] [Medline: 34291034]

43. Jung H, Kwon YD, Noh J. Financial burden of catastrophic health expenditure on households with chronic diseases: financial ratio analysis. BMC Health Serv Res 2022 Apr 27;22(1):568 [FREE Full text] [doi: 10.1186/s12913-022-07922-6] [Medline: 35477404]
44. Assebe LF, Negussie EK, Jbaily A, Tolla MTT, Johansson KA. Financial burden of HIV and TB among patients in Ethiopia: a cross-sectional survey. BMJ Open 2020 Jun 01;10(6):e036892 [FREE Full text] [doi: 10.1136/bmjopen-2020-036892] [Medline: 32487582]
45. Li H, Jiang L. Catastrophic medical insurance in China. Lancet 2017 Oct 14;390(10104):1724-1725. [doi: 10.1016/S0140-6736(17)32603-X] [Medline: 29047433]
46. Cao W, Hsieh E, Li T. Optimizing treatment for adults with HIV/AIDS in China: successes over two decades and remaining challenges. Curr HIV/AIDS Rep 2020 Feb;17(1):26-34 [FREE Full text] [doi: 10.1007/s11904-019-00478-x] [Medline: 31939111]
47. Ni X, Li Z, Li X, Zhang X, Bai G, Liu Y, et al. Socioeconomic inequalities in cancer incidence and access to health services among children and adolescents in China: a cross-sectional study. Lancet 2022 Sep 24;400(10357):1020-1032. [doi: 10.1016/S0140-6736(22)01541-0] [Medline: 36154677]
48. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 years of health-care reform in China: progress and gaps in Universal Health Coverage. Lancet 2019 Sep 28;394(10204):1192-1204. [doi: 10.1016/S0140-6736(19)32136-1] [Medline: 31571602]
49. Wickrama KAS, Surjadi FF, Lorenz FO, Conger RD, Walker C. Family economic hardship and progression of poor mental health in middle-aged husbands and wives. Fam Relat 2012 Apr 01;61(2):297-312 [FREE Full text] [doi: 10.1111/j.1741-3729.2011.00697.x] [Medline: 22577243]
50. Drentea P, Reynolds JR. Where does debt fit in the stress process model? Soc Ment Health 2015 Mar;5(1):16-32. [doi: 10.1177/2156869314554486] [Medline: 31106006]

Abbreviations

aHR: adjusted hazard ratio
aOR: adjusted odds ratio
CES-D: Center for Epidemiological Studies Depression Scale
CFPS: China Family Panel Studies
CHE: catastrophic health expenditure
CVD: cardiovascular disease
GRP: gross regional product
HR: hazard ratio
NRCMS: New Rural Cooperative Medical Scheme
OOP: out-of-pocket
UEBMI: Urban Employee Basic Medical Insurance
UHC: universal health coverage
URBMI: Urban Resident Basic Medical Insurance

Edited by A Mavragani, T Sanchez; submitted 05.09.22; peer-reviewed by T Liyuan, J Ma, S Pesälä, J Arias De La Torre; comments to author 25.04.23; revised version received 18.05.23; accepted 26.05.23; published 15.08.23.

Please cite as:
Wang Y, Liang W, Liu M, Liu J
Association of Catastrophic Health Expenditure With the Risk of Depression in Chinese Adults: Population-Based Cohort Study
JMIR Public Health Surveill 2023;9:e42469
URL: https://publichealth.jmir.org/2023/1/e42469
doi:10.2196/42469
PMID:37581926

Original Paper

Association Between the Dietary Inflammatory Index and the Risk of Fracture in Chinese Adults: Longitudinal Study

Lu Wang[1], MSc; Chen Ye[2], MSc; Fanghong Zhao[3], BS; Hongjing Wu[1], MSc; Ruoyu Wang[2], MSc; Zhaofeng Zhang[2,4*], MD, PhD; Jie Li[1,5*], MSc

[1]Department of Epidemiology and Health Statistics, Hebei Medical University, Shijiazhuang, China
[2]Department of Nutrition and Food Hygiene, School of Public Health, Peking University Health Science Center, Beijing, China
[3]Beijing Center for Disease Control and Prevention, Beijing, China
[4]Beijing's Key Laboratory of Food Safety Toxicology Research and Evaluation, Beijing, China
[5]Beijing Fengtai District Center for Disease Control and Prevention, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Zhaofeng Zhang, MD, PhD
Department of Nutrition and Food Hygiene
School of Public Health, Peking University Health Science Center
No.38 Xueyuan Road,Haidian District
Beijing, 100191
China
Phone: 86 10 82801575
Email: zhangzhaofeng@bjmu.edu.cn

Abstract

Background: Chronic inflammation plays a crucial role in tissue injury, osteoporosis, and fracture. The dietary inflammatory index (DII) is a tool for assessing the potential for inflammation in the diet. However, the association between the DII and fractures remains controversial from previous studies.

Objective: We aimed to explore the correlation between the DII and fracture risk in Chinese adults.

Methods: We included 11,999 adults (5519 men and 6480 women) who were a part of the China Health and Nutrition Survey (1997-2015) prospective cohort. A 3-day, 24-hour meal review method was used to calculate the DII score. The fractures were identified using a questionnaire. Cox proportional hazards models were used to estimate the hazard ratios (HRs) and 95% CIs for fractures. Subgroup, sensitivity, and restricted cubic spline analyses were performed.

Results: During the 18 years of follow-up (median follow-up 9.0 years), 463 men and 439 women developed fractures. The median DII score was 0.64 (IQR −1.74 to 1.46) for the total sample, 0.75 (IQR −1.68 to 1.50) for men, and 0.53 (IQR −1.79 to 1.42) for women. The DII score had a positive correlation with the risk of fracture among women but not among men. For men, after adjusting for covariates, the HRs for quintiles of DII were 1, 0.96 (95% CI 0.66-1.41), 1.05 (95% CI 0.74-1.49), 0.89 (95% CI 0.62-1.26), and 0.94 (95% CI 0.67-1.34; trend: *P*=.62). The HRs for women were 1, 1.13 (95% CI 0.72-1.79), 1.24 (95% CI 0.83-1.86), 1.51 (95% CI 1.02-2.22), and 1.62 (95% CI 1.10-2.39; trend: *P*=.004). The restricted cubic spline analysis showed a significant association between fracture risk and DII score in women (overall association: *P*=.01); as the DII scores were >0.53, HRs showed a significant upward trend. Women aged <50 years or who are nonsmokers, who are nondrinkers, or with nonabdominal obesity had a positive association between fracture risk and the DII score. In sensitivity analyses, after excluding people with diabetes or hypertension, there was still a positive association between fracture risk and the DII score in women. Among the DII components, the DII scores of protein (trend: *P*=.03), niacin (trend: *P*=.002), and iron (trend: *P*=.02) showed significant associations with the risk of fracture in women.

Conclusions: Proinflammatory diet consumption increased the fracture risk in Chinese women aged <50 years. The high consumption of anti-inflammatory foods and low consumption of proinflammatory foods may be an important strategy to prevent fractures in women.

(JMIR Public Health Surveill 2023;9:e43501) doi:10.2196/43501

KEYWORDS

dietary inflammatory index; fractures; diet; risk

Introduction

Fractures Hazard

Globally, fractures pose a serious economic burden and public health issue [1,2]. Fractures can result in work absence, decreased productivity, disability, impaired quality of life, health loss, and high health care costs and are a major burden on health care systems at large [3-5], especially in people with osteoporosis [6,7]. The aging of the population is associated with an increasing burden of fractures around the world [8]. A systematic review including 22 studies from high-income nations or regions and 3 from low- or intermediate-income nations indicated that hip fracture incidence still increased rapidly in low- or intermediate-income nations [9]. Hip fracture is the most devastating type of osteoporotic fracture. It is projected that the number of hip fractures around the world will increase from 1.26 million in 1990 to 4.5 million by 2050, about half of which are likely to occur in Asia, particularly in China [10]. China, the largest middle-income country, has the largest population in the world and has been experiencing unprecedented rapid population aging. Fractures will be a substantial health burden in China.

Fragility fractures caused by low-energy trauma (slips, trips, and falls from standing height) were studied extensively [11]. Fragility fractures are closely associated with bone density (osteopenia and osteoporosis) [12,13]. Osteoporosis is the leading cause of fragility fractures, with an osteoporotic fracture occurring every 3 seconds [14]. The study of Global Burden (1990-2019) indicated that China, with the high disease burden of disability-adjusted life years number in low bone mass–related fractures, ranks second in 204 countries and territories [15]. Injuries have recently been recognized in China’s long-term development agenda, *Healthy China 2030* [16]. Fractures are an important contributor to the injury burden in China. We need to provide specific fracture prevention strategies to support the development of the Healthy China policy.

Fractures and Chronic Inflammation

Currently, the mechanism underlying osteoporosis and fractures remains unclear, but chronic inflammation is widely regarded as an important cause of osteoporosis. Several observational studies have examined the potential association between serum inflammatory markers, such as interleukin (IL)–6, tumor necrosis factor (TNF)–α, and C-reactive protein (CRP), and fractures [17-21]. Proinflammatory cytokines such as TNF-α and IL-6 in osteoclasts can lead to bone erosion [22,23]. Diet is an important source of inflammation [24]. Inflammation levels in the body can be mediated by dietary components [25-27], some nutrients (eg, saturated fatty acids and trans fatty acids) can increase inflammation levels, and some nutrients (eg, fiber and some vitamins) can decrease inflammation levels [23,28]. To quantify the inflammatory potential of diets in various population groups [24], Cavicchia [29] first proposed the dietary inflammatory index (DII) in 2009, and Shivappa [24] developed the evaluation indexes further in 2014. They included 6500 references; studied the effects of food parameters on 6 inflammatory markers (IL1-β, IL-4, IL-6, IL-10, TNF-α, and CRP); rated the proinflammatory, anti-inflammatory, and noninflammatory effects of each food parameter; and finally determined 45 dietary components to calculate DII. A higher DII score indicates that a diet is more likely to enhance the body’s inflammatory response, whereas a lower DII score indicates that the diet reduces inflammation levels in the body [28,30-32]. The DII is a good indicator to estimate dietary quality associated with inflammation and has also been used to explore relationships with various health outcomes, including hypertension, diabetes, hyperuricemia, metabolic syndrome, asthma, breast cancer, and colorectal cancer [33-39].

Association Between the DII and Fractures

In recent years, several studies have been conducted to examine the association between the DII score and fractures [23,32,40-43]. However, the association between DII and fracture risk showed inconsistent results, which may be attributed to differences in age, gender, ethnicity, and time of observation. A study involving postmenopausal women reported that a high DII score was associated with an increased hip fracture risk [32]. Moreover, results from North American studies showed that a higher DII score was associated with a higher incidence of fractures [40]. However, these studies were limited to White women. In addition, in the Tasmanian Older Adult Cohort Study, a higher DII score was associated with lower bone density and an increased incidence of fractures in community-dwelling older men but decreased fracture incidence in women [41]. The Brazilian Osteoporosis Study showed a lack of association between DII and low-impact fractures in the Brazilian population [42]. These foreign studies, because of racial differences, are not specific to China. In addition, populations in Western countries tend to consume animal-based diets, whereas populations in Eastern countries, such as China, tend to consume plant-based diets, characterized by higher quantities of vegetables and fruits [44]. Depending on the diet, the level of inflammation also varies [29]. Results from participants based in the Guangdong province and Hong Kong cohort study showed that a proinflammatory diet was positively associated with hip fracture risk in men and women [23,43]. Because the participants in the 2 Chinese studies either were from an economically advanced region of China or had settled in the southern part of the country, we believe that these results are not representative of a typical Chinese diet [23,43].

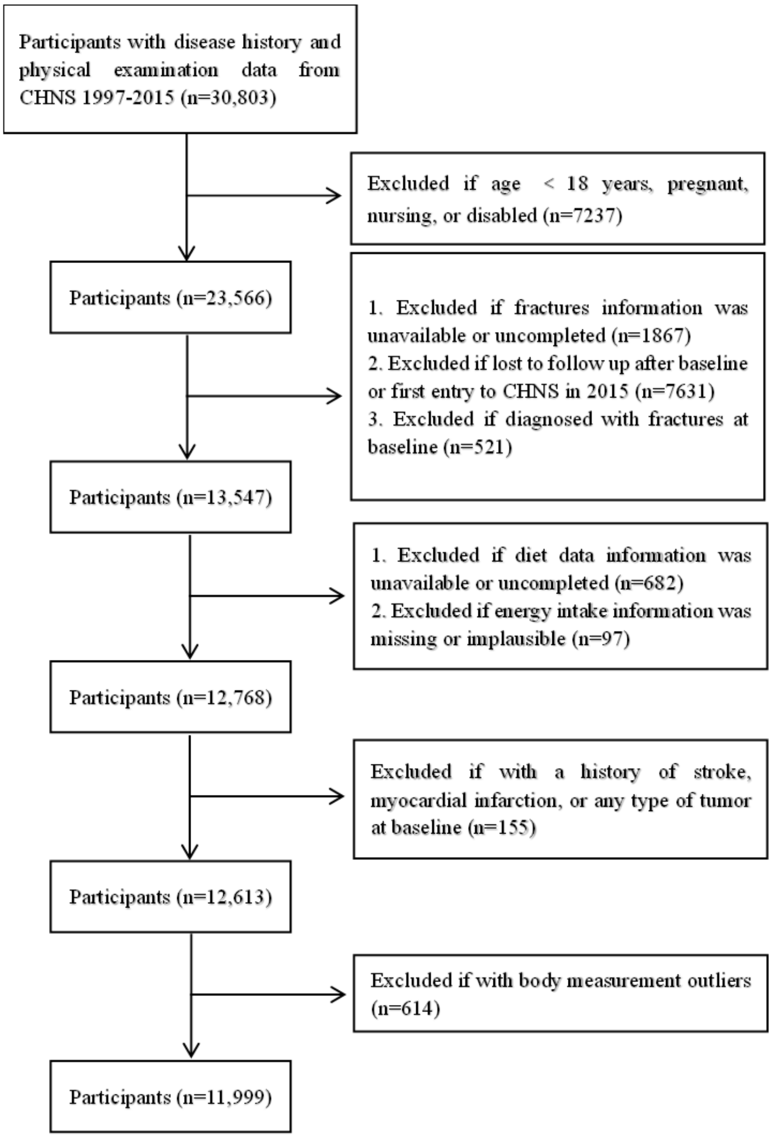
Currently, the pattern of food consumption has changed with the rapid development of the Chinese society. Despite the improvement in diet quality, the current dietary pattern is still less than ideal. The diet in China is gradually being westernized, with an increased intake of red meat, processed meats, and sugary drinks, which are associated with higher levels of inflammation [45]. Currently, China proposed the suggestion of “reduced salt, reduced oil, reduced sugar, healthy oral cavity, healthy weight, and healthy bone (three reductions and three healthy),” which was the action of national health promotion. Therefore, this study, which prospectively examined the association between the DII and fracture risk in a large sample in China, aimed to provide theoretical support and guidance for the use of nutrition and diet to reduce the risk of fractures. We hope to use practical actions to help the Healthy China development agenda.

Methods

Study Population

This study used data from a subcohort of the China Health and Nutrition Survey (CHNS). Founded in 1989, the CHNS is a prospective cohort study that has conducted 10 follow-up surveys until 2015. In our study, we used the CHNS data obtained between 1997 and 2015. During this time, 30,803 participants with a disease history and physical examination data were enrolled. This study included adult participants aged ≥18 years. We excluded participants who were pregnant, nursing, or disabled and those without available or complete fracture data or follow-up data. We also excluded participants with missing or implausible energy intake information (>5000 or <700 kcal/d); a baseline diagnosis of fractures; or a history of myocardial infarction, stroke, or any type of tumor at baseline. We also excluded participants with body measurement outliers. Finally, we included 11,999 study participants, including 5519 men and 6480 women (Figure 1).

Figure 1. Study population flowchart. CHNS: China Health and Nutrition Survey.



Assessment of Nutrient Intake and the DII

Dietary intake assessment in the CHNS involved the 3-day, 24-hour meal review method for participating individuals and a household food inventory, which involved the weighing and measuring of products (used to obtain information on the consumption of edible oils and condiments) over the same 3 days (2 weekdays and 1 weekend day). The China Food Composition Tables Standard Edition provides general nutrition data for more than 3000 types of foods and ingredients in China. We combined the dietary data obtained from the 3-day, 24-hour meal review method with the China Food Composition Tables Standard Edition to obtain the data used to calculate the DII scores.

The mean intake of every food variable was transformed with standardized values from a world database into a *z* score, converted to a percentile, and centered by doubling and subtracting 1. Finally, the centered percentile score for each food variable was multiplied by its associated literature-derived inflammatory effect score and these scores were summed across the 45 dietary variables, thus providing an individual DII score. The higher the DII score, the more proinflammatory the diet; more negative values indicate more anti-inflammatory diets [24].

The energy-adjusted DII was calculated using the following steps, which were performed before the transformed *z* score. Energy adjustment was performed using the density method [46]; all the food variables and the world database were converted to units per 1000 kcal [47].

In total, 28 of the 45 possible food parameters were analyzed to obtain the overall DII scores. These included energy, carbohydrate, protein, total fat, alcohol, β-carotene, cholesterol, fiber, folic acid, niacin, iron, magnesium, selenium, zinc, monounsaturated fatty acids, polyunsaturated fatty acids, isoflavones, thiamine, riboflavin, saturated fat, vitamin A, vitamin C, garlic, ginger, onion, green or black tea, pepper, and thyme or oregano.

Assessment of Covariates

On the basis of the Cox proportional hazards model of fractures, previous studies have analyzed demographic characteristics, lifestyle factors, physical measures, and disease information [41,48]. Information on demographic characteristics, obtained from a questionnaire, included age, gender, area of residence, marital status, level of educational attainment, and household income per capita (categorized into quartiles). Lifestyle factors obtained via a questionnaire included physical activity level (PAL; grouped into 3 levels), smoking status, and drinking status. Physical measures included the BMI, midarm muscle circumference (MAMC), and waist-to-hip ratio (WHR). The BMI was calculated as weight (kg) / height (m)^2. MAMC was calculated as midupper arm circumference (cm) − π × (triceps skinfold thickness / 10). WHR was calculated as waist circumference (cm) divided by hip circumference (cm). According to the World Health Organization Asian WHR standard, a man with WHR ≥0.90 and a woman with WHR ≥0.85 are considered to have abdominal obesity [49]. The disease information including diabetes and hypertension was obtained using questionnaires. The Osteoporosis Self-assessment Tool for Asians (OSTA) index has been shown to distinguish between different degrees of osteoporosis. OSTA was calculated as 0.2 × (weight [kg] − age [years]); the following classification was used: OSTA <−4, high risk for osteoporosis and fracture; −4< OSTA <−1, medium risk for osteoporosis and fracture; and OSTA >−1, low risk for osteoporosis and fracture [50-52]. For all covariates, we used the baseline year measure.

Outcome Identification

The outcome was defined based on the questionnaire survey. The participants were asked to report their history of fractures using a questionnaire-based interview at each follow-up since 1997. The questions were posed as follows: “(1) History of bone fracture? If yes, (2) age (years) at 1st bone fracture, and (3) number of times bone fracture?” We used the date when participants first entered the survey as the baseline date for participants according to the questionnaire. The follow-up person-time for each participant was calculated as the interval between baseline and the occurrence of fractures, the survey day in the last survey round before the participant was lost to follow-up, or the latest survey in 2015, whichever came first.

Statistical Analysis

All statistical analyses were performed separately for men and women. We also divided the participants of each gender into 5 groups according to the quintiles of the DII. Participant baseline characteristics were described as a number (percentage) for categorical variables and means (SD) or as a median (IQR 25%-75%) for continuous variables. The chi-square test and rank-sum test were used to compare categorical and continuous variables between the men and women.

Cox proportional hazards models were used to estimate the hazard ratios (HRs) and 95% CIs for fractures. We also adjusted for multiple covariates before exploring the relationship between DII and fracture risk. To test the proportional hazards assumption, we conducted likelihood ratio tests. The results met the HR assumption. To calculate HRs among quintiles, the lowest intake quintiles were used as a reference. Three models were established. Model 1 was adjusted for demographic characteristics, including age, gender, residence area, highest education level, household income, and marital status. Model 2 was adjusted for lifestyle factors and physical measures: model 1+smoking status, drinking status, PAL, BMI, MAMC, and WHR. Model 3 adjusted for disease information: model 2+hypertension, diabetes, and OSTA index level. Tests for trends were performed for continuous variables using categorical DII scores by quintiles.

Stratified analyses were performed according to age, BMI, smoking status, drinking status, obesity status, OSTA index level, and MAMC. The *P* value for interaction was calculated with multiplicative terms by multiplying the quintiles of the DII by categorical variables used in the multivariable model.

Several sensitivity analyses were conducted in this study. Participants with hypertension or diabetes mellitus were excluded from the study. Participants were divided into 4 groups: nonhypertension, nondiabetes, nonhypertension and nondiabetes, as well as nonhypertension or nondiabetes.

Carbohydrate, total fat, protein, thiamine, riboflavin, niacin, folic acid, vitamin A, iron, selenium, zinc, and magnesium of the DII component were divided into 5 levels in men and women to estimate the HRs and 95% CIs for fractures. To calculate the DII scores for these nutrients, we controlled for energy consumption.

We used restricted cubic splines (RCSs) to test for linearity, and RCSs were used for DII in multivariable-adjusted Cox regression analyses (model 3) for women separately. The median values were used as references in the RCS analyses.

All statistical analyses were performed using SPSS (version 27.0; IBM Corp) and R (version 4.1.0; R Foundation for Statistical Computing). Two-sided *P* values <.05 were considered as statistically significant.

Ethics Approval

The CHNS, is an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health (NINH, former National Institute of Nutrition and Food Safety) at the Chinese Center for Disease Control and Prevention (CCDC) with an ethical approval number of 2015017 [53].

Participation and Informed Consent

The CHNS is an ongoing open cohort. All participants provided informed consent for inclusion before participating in the study and allowed for secondary analysis without additional consent. The study data are anonymized or deidentified.

Results

Baseline Characteristics

We included 11,999 adults (5519 men and 6480 women) who were part of the CHNS (1997-2015) prospective cohort. The average age of the participants was 44.0 (SD 14.6) years. Of the total 11,999 participants, 5519 (46%) were men and 6480 (54%) were women. The median DII score was 0.64 (IQR −1.74 to 1.46) for the total sample, 0.75 (IQR −1.68 to 1.50) for men, and 0.53 (IQR −1.79 to 1.42) for women. A histogram of the DII in men and women is presented in Figure S1 in Multimedia Appendix 1. During the 18 years of follow-up (median follow-up 9.0 years), we ascertained that 463 men and 439 women developed fractures. The baseline characteristics of the study population are described across the quintiles of DII in Table 1. The population baseline characteristics by gender are also presented in Table S1 in Multimedia Appendix 1. At baseline, marital status, education level, smoking status, drinking status, PAL, BMI, MAMC, WHR, abdominal obesity status, OSTA index level, carbohydrate, total fat, and protein levels were statistically different between men and women (all *P*<.001).

Table 1. Baseline characteristics of the study population in men and women[a].

Characteristics	Qs[b] of DII[c] in men			Qs of DII in women		
	Q1 (n=1104)	Q3 (n=1104)	Q5 (n=1103)	Q1 (n=1296)	Q3 (n=1296)	Q5 (n=1296)
Age (years), mean (SD)	47.9 (14.0)	41.5 (14.5)	43.3 (14.3)	44.8 (14.8)	42.8 (13.9)	42.9 (14.0)
<50	464 (42.03)	658 (59.6)	626 (56.75)	656 (50.62)	753 (58.1)	749 (57.79)
≥50	640 (57.97)	446 (40.4)	477 (43.25)	640 (49.38)	543 (41.9)	547 (42.21)
Marital status, n (%)						
Never married	79 (7.16)	178 (16.12)	140 (12.69)	44 (3.4)	60 (4.63)	84 (6.48)
Married	985 (89.22)	895 (81.07)	934 (84.68)	1138 (87.81)	1143 (88.19)	1123 (86.65)
Divorced	19 (1.72)	7 (0.63)	9 (0.82)	28 (2.16)	11 (0.85)	4 (0.31)
Widowed	21 (1.9)	24 (2.17)	20 (1.81)	86 (6.64)	82 (6.33)	85 (6.56)
Residence, n (%)						
Rural	419 (37.95)	815 (73.82)	733 (66.46)	554 (42.75)	958 (73.92)	880 (67.9)
Urban	685 (62.05)	289 (26.18)	370 (33.51)	742 (57.25)	338 (26.08)	416 (32.1)
Household income level (yuan ¥; yuan ¥1=US $0.14), n (%)						
Low (<8532)	254 (23.01)	299 (27.08)	285 (25.84)	308 (23.77)	365 (28.16)	339 (26.16)
Medium (8532-15,576)	235 (21.29)	308 (27.9)	305 (27.65)	293 (22.61)	330 (25.46)	329 (25.39)
High (15,577-30,500)	274 (24.82)	259 (23.46)	264 (23.93)	313 (24.15)	361 (27.85)	326 (25.15)
Very high (>30,500)	341 (30.89)	238 (21.56)	249 (22.57)	382 (29.48)	240 (18.52)	302 (23.3)
Education level, n (%)						
None	63 (5.71)	197 (17.84)	165 (14.96)	172 (13.27)	484 (37.35)	458 (35.34)
Graduate from primary school	146 (13.22)	291 (26.36)	266 (24.12)	213 (16.44)	316 (24.38)	265 (20.45)
Lower middle school degree	380 (34.42)	373 (33.79)	384 (34.81)	407 (31.4)	315 (24.31)	344 (26.54)
Upper middle school degree or above	515 (46.65)	243 (22.01)	288 (26.11)	504 (38.89)	181 (13.97)	229 (17.67)
Former or current smoker, n (%)	662 (59.96)	710 (64.31)	682 (61.83)	32 (2.47)	54 (4.17)	60 (4.63)
Former or current drinker, n (%)	707 (64.04)	717 (64.95)	743 (67.36)	159 (12.27)	130 (10.03)	130 (10.03)
Physical activity level, n (%)						
Low	710 (64.31)	376 (34.06)	434 (39.35)	981 (75.69)	543 (41.9)	609 (47)
Medium	203 (18.39)	197 (17.84)	213 (19.31)	142 (10.96)	113 (8.72)	165 (12.73)
High	191 (17.3)	531 (48.1)	456 (41.34)	173 (13.35)	640 (49.38)	522 (40.28)
BMI (kg/m^2), n (%)						
≤18.4	29 (2.63)	97 (8.79)	57 (5.17)	79 (6.1)	92 (7.1)	75 (5.79)
18.5-23.9	531 (48.1)	777 (70.38)	709 (64.28)	667 (51.47)	817 (63.04)	806 (62.19)
24.0-27.9	422 (38.22)	191 (17.3)	280 (25.39)	394 (30.4)	322 (24.85)	335 (25.85)
≥28.0	122 (11.05)	39 (3.53)	57 (5.17)	156 (12.04)	65 (5.02)	80 (6.17)
MAMC[d] (cm), median (IQR)	23.1 (20.8-25.3)	22.4 (20.7-24.1)	22.9 (21.3-24.9)	20.5 (18.7-22.5)	20.0 (18.6-21.6)	20.4 (18.7-22.5)
WHR[e], median (IQR)	0.90 (0.85-0.93)	0.86 (0.82-0.90)	0.87 (0.82-0.91)	0.84 (0.80-0.89)	0.84 (0.79-0.88)	0.82 (0.79-0.87)
Abdominal obesity[f], n (%)	526 (47.64)	258 (23.37)	319 (28.92)	608 (46.91)	532 (41.05)	440 (33.95)
Hypertension, n (%)	151 (13.68)	27 (2.45)	58 (5.26)	160 (12.35)	52 (4.01)	56 (4.32)
Diabetes, n (%)	50 (4.53)	6 (0.54)	18 (1.63)	34 (2.62)	13 (1)	23 (1.77)
Fractures, n (%)	52 (4.71)	128 (11.59)	117 (10.61)	37 (2.85)	99 (7.64)	135 (10.42)

Characteristics	Qs[b] of DII[c] in men			Qs of DII in women		
	Q1 (n=1104)	Q3 (n=1104)	Q5 (n=1103)	Q1 (n=1296)	Q3 (n=1296)	Q5 (n=1296)
OSTA[g] index level[h], n (%)						
>−1	984 (89.13)	955 (86.5)	995 (90.21)	1088 (83.95)	1028 (79.32)	1055 (81.4)
−1 to −4	106 (9.6)	132 (11.96)	100 (9.07)	172 (13.27)	214 (16.51)	201 (15.51)
<−4	14 (1.27)	17 (1.54)	8 (0.73)	36 (2.78)	54 (4.17)	40 (3.09)
Carbohydrate (g/d), median (IQR)	274.3 (201.7-349.4)	390.3 (322.3-467.5)	360.8 (271.7-465.6)	237.7 (175.3-303.4)	332.5 (261.7-405.1)	320.4 (246.6-402.0)
Total fat (g/d), median (IQR)	71.0 (50.4-95.9)	61.2 (41.8-86.0)	72.9 (50.3-98.9)	60.2 (41.7-83.0)	51.4 (34.0-72.8)	66.7 (45.0-89.0)
Protein (g/d), median (IQR)	75.3 (61.4-92.2)	69.6 (57.4-85.0)	75.8 (62.2-92.2)	63.9 (51.6-80.0)	59.0 (48.0-72.0)	65.5 (53.3-80.1)

[a] The data are presented as mean (SD) or median (IQR) for continuous variables and as n (%) for categorical variables.

[b] Q: quintile.

[c] DII: dietary inflammatory index.

[d] MAMC: midarm muscle circumference.

[e] WHR: waist-to-hip ratio.

[f] A man WHR ≥0.90 and a women WHR ≥0.85 is abdominal obesity.

[g] OSTA: Osteoporosis Self-assessment Tool for Asians.

[h] The OSTA index level distinguishes between different degrees of osteoporosis.

Association of DII With Fractures

The results of the Cox proportional hazards models suggested that the DII was significantly associated with fracture risk in people (Table 2). However, this correlation exists only in women rather than men. For men, after adjusting for covariates, the HRs for quintiles of DII were 1, 0.96 (95% CI 0.66-1.41), 1.05 (95% CI 0.74-1.49), 0.89 (95% CI 0.62-1.26), and 0.94 (95% CI 0.67-1.34; trend: *P*=.62); for women, after adjusting for covariates, the HRs for quintiles of DII were 1, 1.13 (95% CI 0.72-1.79), 1.24 (95% CI 0.83-1.86), 1.51 (95% CI 1.02-2.22), and 1.62 (95% CI 1.10-2.39; trend: *P*=.004). When DII was treated as a continuous variable, it was also significantly associated with fracture risk in women rather than in men. The 2 results were coinciding with those of the man and woman quintiles.

Table 2. Hazard ratios (95% CIs) of fractures according to the quintiles (Qs) of the dietary inflammatory index (DII)[a].

	Qs of DII, hazard ratio (95% CI)					Trend, *P* value[b]	Continuous DII, hazard ratio (95% CI)
	Q1	Q2	Q3	Q4	Q5		
Men							
Model 1[c]	1 (ref[d])	0.96 (0.66-1.41)	1.04 (0.73-1.46)	0.89 (0.62-1.24)	0.95 (0.67-1.34)	.63	0.99 (0.92-1.05)
Model 2[e]	1 (ref)	0.96 (0.66-1.41)	1.05 (0.74-1.49)	0.88 (0.62-1.25)	0.94 (0.67-1.34)	.63	0.99 (0.92-1.05)
Model 3[f]	1 (ref)	0.96 (0.66-1.41)	1.05 (0.74-1.49)	0.89 (0.62-1.26)	0.94 (0.67-1.34)	.62	0.99 (0.92-1.05)
Women							
Model 1[c]	1 (ref)	1.11 (0.71-1.76)	1.17 (0.78-1.75)	1.45 (0.99-2.13)	1.57 (1.07-2.31)	.005	1.12 (1.04-1.20)
Model 2[e]	1 (ref)	1.12 (0.71-1.77)	1.21 (0.81-1.80)	1.48 (1.00-2.18)	1.59 (1.08-2.34)	.049	1.12 (1.04-1.20)
Model 3[f]	1 (ref)	1.13 (0.72-1.79)	1.24 (0.83-1.86)	1.51 (1.02-2.22)	1.62 (1.10-2.39)	.004	1.12 (1.04-1.21)

[a] Q1 indicates participants having the lowest DII values, the least proinflammatory level; Q5 indicates participants having the highest DII values, the most proinflammatory level. Data were hazard ratios (95% CIs), calculated using Cox proportional hazards analyses.

[b] *P* value for trend: tests for trends were performed for continuous variables using categorical DII scores by Qs.

[c] Adjusted for age, gender, residence area, highest education level, household income, and marital status.

[d] ref: reference.

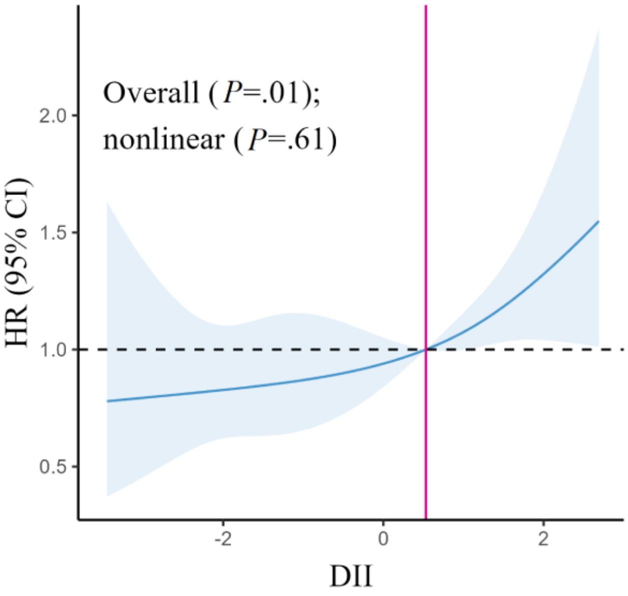
[e] Model 1+smoking status, drinking status, physical activity level, BMI, midarm muscle circumference, and waist-hip ratio.

[f] Model 2+hypertension, diabetes, and Osteoporosis Self-assessment Tool for Asians index level.

RCS Analysis

To further understand the correlation between the DII and fracture risk, we performed an RCS analysis. The RCS analysis showed a significant association between fracture risk and the DII score in women (overall association: *P*=.01; Figure 2). As the DII score was >0.53, HR showed a significant upward trend. However, the spline variable confirmed no significant departure from the nonlinear relationship (nonlinear: *P*=.61) between fracture risk and the DII score.

Figure 2. Multivariable-adjusted hazard ratios (HRs; blue solid lines) and 95% CIs (blue shadow) for risk of fracture according to the dietary inflammatory index (DII) score among women in model 3. The median intakes were set as references (black dotted line; HR=1.00). The solid pink line represents the line where the point corresponding to the value of the DII in the curve was located when HR=0.



Subgroup Analyses

The association between fracture risk and the DII score, as stratified by age, BMI, smoking status, drinking status, obesity status, OSTA index level, and MAMC, was further investigated. Women aged <50 years had a positive association between fracture risk and the DII score (trend: *P*=.003; HR 2.15, 95% CI 1.16-3.97), whereas no association was found among those aged ≥50 years (trend: *P*=.38; HR 1.15, 95% CI 0.69-1.92). Women who are nonsmokers, who are nondrinkers, and with nonabdominal obesity had a positive association between fracture risk and the DII score (trend: *P*=.004; HR 1.64, 95% CI 1.10-2.44; trend: *P*=.02; HR 1.51, 95% CI 1.01-2.27; and trend: *P*=.03; HR 1.92, 95% CI 1.11-3.32, respectively), whereas no association was found among smokers, drinkers, and abdominal obesity (trend: *P*=.37; HR 0.61, 95% CI 0.13-2.97; trend: *P*=.06; HR 3.12, 95% CI 0.86-16.84; and trend: *P*=.07; HR 1.31, 95% CI 0.77-2.23, respectively). Women with an OSTA index level >−1 had a positive association between fracture risk and the DII score (trend: *P*=.001; HR 2.02, 95% CI 1.27-3.21), whereas no association was found among those with an OSTA index level ≤−1 (trend: *P*=.67; HR 0.79, 95% CI 0.40-1.57) and no association was found among those with BMI <24 kg/m^2, ≥24 kg/m^2, MAMC <21.46, or MAMC ≥21.46 (Table 3). However, there was no statistically significant association between fracture risk and the DII score in the man subgroup (all *P*>.05; Table S2 in Multimedia Appendix 1). When DII was treated as a continuous variable, these subgroup analyses results were coinciding with the men and women quintiles, which are also shown in the forest plots in Figure S2 in Multimedia Appendix 1. None of the subgroup analyses had interactions (Table 3).

Table 3. Hazard ratios (95% CIs) of fractures according to the quintiles (Qs) of the dietary inflammatory index (DII) by age, BMI, smoking status, drinking status, obesity status, Osteoporosis Self-assessment Tool for Asians (OSTA) index level, and midarm muscle circumference (MAMC) in women[a].

Subgroups	Qs of DII in women, hazard ratio (95% CI)					Trend, *P* value[b]	Interaction, *P* value[c]	Continuous DII, hazard ratio (95% CI)
	Q1	Q2	Q3	Q4	Q5			
Age (years)							.50	
<50	1 (ref[d])	1.15 (0.55-2.40)	1.38 (0.73-2.61)	1.62 (0.86-3.01)	2.15 (1.16-3.97)	.003		1.19 (1.06-1.34)
≥50	1 (ref)	1.02 (0.57-1.81)	0.90 (0.52-1.54)	1.16(0.70-1.94)	1.15 (0.69-1.92)	.38		1.05 (0.92-1.16)
BMI (kg/m^2)							.07	
<24	1 (ref)	1.08 (0.59-1.97)	1.36 (0.79-2.32)	1.45 (0.87-2.44)	1.64 (0.92-2.77)	.03		1.11 (1.01-1.23)
≥24	1 (ref)	0.89 (0.45-1.76)	1.14 (0.63-2.07)	1.29 (0.73-2.28)	1.35 (0.76-2.37)	.13		1.11 (1.00-1.24)
Smoking status							.60	
Former or current smoker	1 (ref)	1.10 (0.28-4.91)	0.74 (0.18-3.46)	0.61 (0.14-2.88)	0.61 (0.13-2.97)	.37		0.88 (0.64-1.24)
Nonsmoker	1 (ref)	1.12 (0.70-1.78)	1.23 (0.81-1.86)	1.45 (0.97-2.17)	1.64 (1.10-2.44)	.004		1.13 (1.05-1.22)
Drinking status							.37	
Former or current drinker	1 (ref)	1.12 (0.17-7.31)	2.61 (0.68-14.40)	2.20 (0.59-12.05)	3.12 (0.86-16.84)	.06		1.26 (0.99-1.65)
Nondrinker	1 (ref)	1.12 (0.70-1.79)	1.13 (0.75-1.73)	1.42 (0.95-2.12)	1.51 (1.01-2.27)	.02		1.10 (1.02-1.19)
OSTA[e] index level							.29	
>−1	1 (ref)	0.98 (0.55-1.76)	1.44 (0.89-2.34)	1.67 (1.05-2.67)	2.02 (1.27-3.21)	.001		1.20 (1.09-1.31)
≤−1	1 (ref)	0.80 (0.38-1.67)	1.14 (0.59-2.21)	0.84 (0.43-1.65)	0.79 (0.40-1.57)	.67		0.96 (0.84-1.10)
Obesity status[f]							.36	
Abdominal obesity	1 (ref)	0.98 (0.51-1.88)	0.77 (0.42-1.40)	1.23 (0.71-2.10)	1.31 (0.77-2.23)	.07		1.09 (0.98-1.21)
Nonabdominal obesity	1 (ref)	1.27 (0.69-2.34)	1.79 (1.03-3.11)	1.30 (0.74-2.29)	1.92 (1.11-3.32)	.03		1.13 (1.02-1.26)
MAMC							.72	
<21.46	1 (ref)	0.81 (0.51-1.27)	1.16 (0.79-1.70)	1.06 (0.73-1.55)	1.25 (0.86-1.82)	.07		1.07 (0.99-1.15)
≥21.46	1 (ref)	1.07 (0.66-1.73)	1.34 (0.52-3.44)	1.38 (0.43-4.43)	1.51 (0.39-5.83)	.54		1.04 (0.97-1.11)

[a] Values were multivariable-adjusted hazard ratios (95% CIs) for risk of fractures according to Qs of the DII stratified by age, BMI, smoking status, drinking status, obesity status, OSTA index level, and MAMC in model 3. Q1 indicates participants having the lowest DII values, the least proinflammatory level; Q5 indicates participants having the highest DII values, the most proinflammatory level.

[b] *P* value for trend: tests for trends were performed for continuous variables using categorical DII score by Qs.

[c] *P* value for interaction was calculated by contrasting the coefficients of the cross-product of stratified values and DII Qs in the model.

[d] ref: reference.

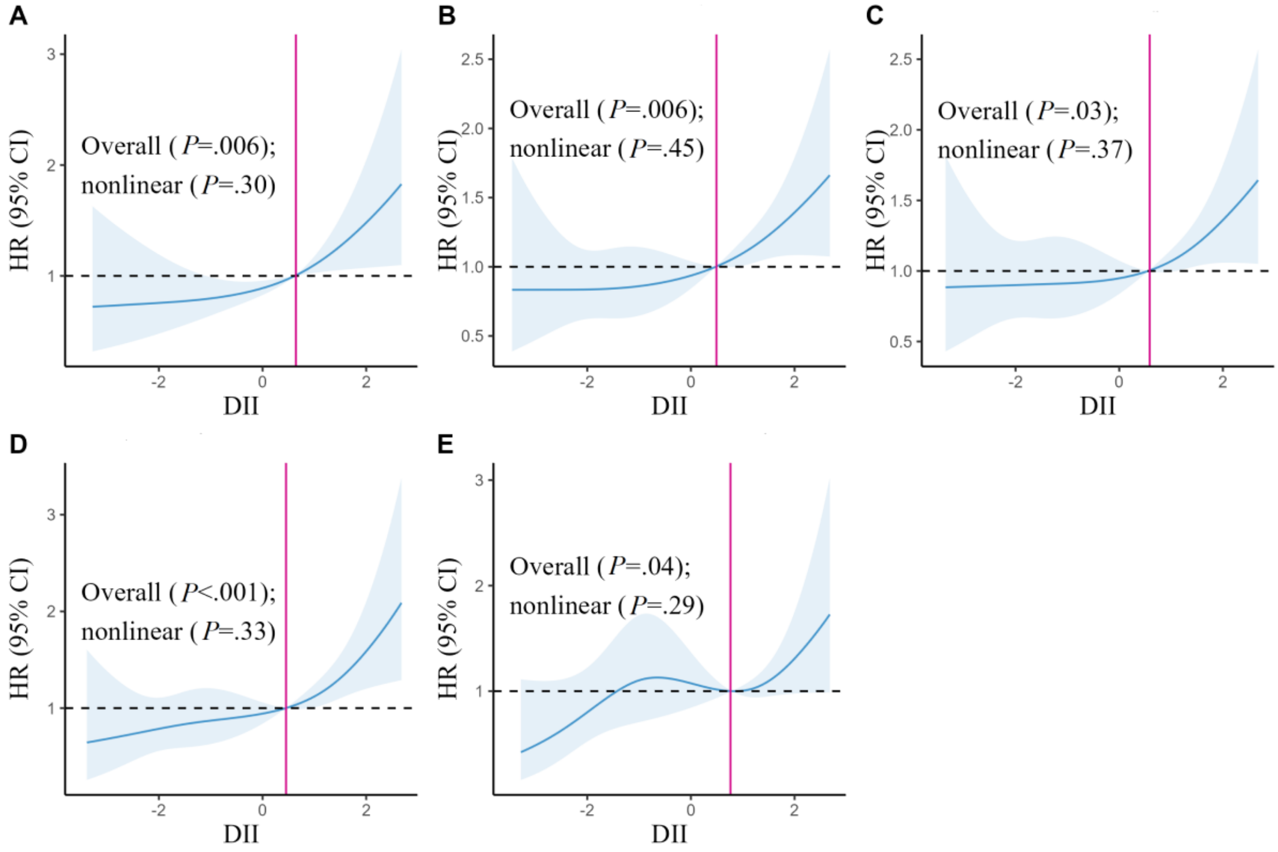
[e] The OSTA index level distinguishes different degrees of osteoporosis.

[f] A man waist-to-hip ratio ≥0.90 and a woman waist-to-hip ratio ≥0.85 is abdominal obesity.

RCS Analysis of Subgroup

To further understand the correlation between the DII and fracture risk among women in the subgroup, we performed an RCS analysis. We showed the statistically significant subgroups in the Cox regression results of subgroup analysis according to subgroup analysis in Table 3, such as age <50 years (trend: *P*=.003), nonsmoker (trend: *P*=.004), nondrinker (trend: *P*=.02), OSTA index level >−1 (trend: *P*=.001), and nonabdominal obesity (trend: *P*=.03). The RCS analysis results showed a significant association between fracture risk and the DII score among women in the subgroup for age <50 years (overall association: *P*=.006; Figure 3A); nonsmoker (overall association: *P*=.006; Figure 3B); nondrinker (overall association: *P*=.03; Figure 3C); OSTA index level >−1 (overall association: *P*<.001; Figure 3D), and nonabdominal obesity (overall association: *P*=.04; Figure 3E). The results of the RCS analysis of the remaining subgroups are presented in Figure S3 in Multimedia Appendix 1. In general, the curves for age <50 years (Figure 3A), nonsmoker (Figure 3B), nondrinker (Figure 3C), and OSTA index level >−1 (Figure 3D), except for nonabdominal obesity (Figure 3E), showed a trend of rapid rise after a slow rise, and the corresponding nodes were 0.64, 0; 0.49, 0; 0.59, 0; and 0.45, 0; respectively. The curve corresponding to nonabdominal obesity (Figure 3E) also showed a rapid upward trend at node 0.76, 0. These results suggest that the higher the inflammation score of the diet, the greater the risk of fracture and that the cumulative effect of inflammatory foods greatly increases the risk of fractures.

Figure 3. Multivariable-adjusted hazard ratios (HRs; blue solid lines) and 95% CIs (blue shadow) for risk of fracture according to the dietary inflammatory index (DII) score among (A) women with age <50 years, (B) nonsmoker, (C) nondrinker, (D) Osteoporosis Self-assessment Tool for Asians (OSTA) index level >−1, and (E) nonabdominal obesity in model 3. The median intakes were set as references (black dotted line; HR=1.00). The solid pink line represents the line where the point corresponding to the value of the DII in the curve was located when HR=0.



Sensitivity Analyses

In sensitivity analyses (Table 4), after excluding people with diabetes or hypertension, there was still a positive association between fracture risk and the DII score in women. For women with nonhypertension, the HRs for quintiles of the DII were 1, 0.99 (95% CI 0.59-1.68), 1.44 (95% CI 0.93-2.24), 1.75 (95% CI 1.14-2.68), and 1.91 (95% CI 1.25-2.92; trend: *P*<.001). For women with nondiabetes, the HRs for quintiles of DII were 1, 1.09 (95% CI 0.68-1.74), 1.24 (95% CI 0.82-1.87), 1.50 (95% CI 1.03-2.22), and 1.62 (95% CI 1.09-2.40; trend: *P*=.003). For women with nonhypertension and diabetes, the HRs for quintiles of the DII were 1, 1.16 (95% CI 0.68-1.99), 1.42 (95% CI 0.90-2.25), 1.78 (95% CI 1.15-2.78), and 1.97 (95% CI 1.26-3.06; trend: *P*<.001). For women with nonhypertension or diabetes, the HRs for quintiles of the DII were 1, 1.09 (95% CI 0.69-1.73), 1.20 (95% CI 0.80-1.79), 1.46 (95% CI 0.99-2.14), and 1.56 (95% CI 1.06-2.30; trend: *P*=.006). When the DII was treated as a continuous variable, it was significantly associated with fracture risk in women with nonhypertension, nondiabetes, nonhypertension and diabetes, and nonhypertension or diabetes. These results are consistent with those of the quintiles.

Table 4. Sensitivity analyses of dietary inflammatory index (DII) in association with fracture risk in women[a].

Groups	Qs[b] of DII in women, hazard ratio (95% CI)					Trend, *P* value[c]	Continuous DII, hazard ratio (95% CI)
	Q1	Q2	Q3	Q4	Q5		
Nonhypertension	1 (ref[d])	0.99 (0.59-1.68)	1.44 (0.93-2.24)	1.75 (1.14-2.68)	1.91 (1.25-2.92)	<.001	1.19 (1.06-1.34)
Nondiabetes	1 (ref)	1.09 (0.68-1.74)	1.24 (0.82-1.87)	1.50 (1.03-2.22)	1.62 (1.09-2.40)	.003	1.05 (0.92-1.16)
Nonhypertension and diabetes	1 (ref)	1.16 (0.68-1.99)	1.42 (0.90-2.25)	1.78 (1.15-2.78)	1.97 (1.26-3.06)	<.001	1.11 (1.01-1.23)
Nonhypertension or diabetes	1 (ref)	1.09 (0.69-1.73)	1.20 (0.80-1.79)	1.46 (0.99-2.14)	1.56 (1.06-2.30)	.006	1.11 (1.00-1.24)

[a] Values are multivariable-adjusted hazard ratios (95% CIs) for the risk of fractures according to quintiles of the DII stratified by hypertension and diabetes status in model 3. Q1 indicates participants having the lowest DII values, the least proinflammatory level; Q5 indicates participants having the highest DII values, the most proinflammatory level.

[b] Q: quintile.

[c] *P* value for trend: tests for trends were performed for continuous variables using categorical DII scores by Qs.

[d] ref: reference.

Association of DII Components With Fractures

The association of macronutrients, some micronutrient DII scores, and fracture risk in men and women is shown in Table S3 in Multimedia Appendix 1. The DII scores for protein (trend: *P*=.03), niacin (trend: *P*=.002), and iron (trend: *P*=.02) showed significant associations with fracture risk in women. The higher the intake of protein, the higher the risk of fracture. The higher the intake of niacin and iron, the lower the risk of fracture. The DII scores for carbohydrate, total fat, and protein were associated with fracture risk in men. The association between other macronutrients and micronutrient DII scores and fracture risk is shown in Table S3 in Multimedia Appendix 1.

Discussion

Principal Findings

This study, which used a large-scale sample in China, aimed to prospectively explore the association between dietary inflammatory potential and fracture risk. This study found that a higher DII score (more proinflammatory diets) was associated with a higher risk of fracture in women than in men. To confirm this conclusion, we used Cox proportional hazards models to demonstrate a positive correlation between DII and fracture risk in women and used an RCS analysis to further illustrate the trend of fracture risk with the DII score. In addition, we performed subgroup and sensitivity analyses to further verify the stability of the results. This reliable and robust result can provide dietary interventions and strategies for fracture prevention, promote bone health, provide a specific basis for China’s Healthy Lifestyle Action, and promote the Healthy China development agenda.

Comparison With Prior Work

In this study, the association between the DII and fracture risk differed by gender. Statistically significant results were observed among woman participants rather than man participants. This finding is consistent with previous studies [40,54]. Estrogen is hypothesized to have a protective anti-inflammatory effect [55,56]. Other studies have reported that inflammation has different roles and mechanisms in men and women (eg, for cardiovascular diseases) [57]. In addition, compared with women, men have worse habits, which offsets the positive effect of diet on the prevention of fractures. Tobacco smoke is a proinflammatory agent. It may be that the effect of tobacco swamps the effect of dietary sources of inflammation [58]. In this study, men smoked at about 16 times the rate of women (62.6% of men were past or present smokers vs only about 3.8% of women). However, 2 studies from Guangdong and Hong Kong in China showed an association between the DII and fractures in both genders [23,43]. It may be difficult to directly compare the findings of a case-control study with those of a longitudinal study. In addition, 2 studies considered the older adult population as the research object, and this gender difference became less pronounced with advancing age [59]. We believe that for women, reducing the intake of a proinflammatory diet or increasing the intake of an anti-inflammatory diet, such as fruits, vegetables, and tea, can decrease fracture risk. Consuming ≤5 servings per day of fruit and vegetables is associated with a higher hip fracture risk [60]. Tea and tea extract flavonoids with antioxidant and anti-inflammatory properties have osteoprotective effects on bone biology. Huang et al [61] showed that high tea consumption versus no tea consumption reduced fracture risk by 31%. In postmenopausal women, estrogen levels decrease significantly with age and the protective anti-inflammatory effect is weakened, which can easily cause osteoporosis and increase the risk of fractures.

The RCS analysis showed a changing trend that fracture risk increased with increasing DII scores. Notably, as the DII score was >0.53, HR had shown a significant upward trend. The cumulative effect of inflammatory foods greatly increases the risk of fractures. Therefore, the intake of inflammatory foods should be reduced as early as possible to avoid the accumulation of inflammatory effects of the diet, thereby reducing the risk of fractures.

To our surprise, the research results are present only in women aged <50 years rather than those aged ≥50 years. However, a study of postmenopausal White women confirmed that an inflammatory diet is associated with fractures [32]. This may be due to a significant change in estrogen levels after menopause in women who are prone to osteoporosis and an increase in the risk of fractures [62,63]. However, our results still need to be discussed because of population characteristics, dietary patterns, and differences in the instruments used to assess diet quality. There are 2 possible explanations for this finding. On the one hand, with increasing age, especially in older adults, due to altered body function and physiological and pathological reasons, there is less food intake and food use than in the young. On the other hand, it could be that the benefits of a less inflammatory dietary pattern for bone health are overshadowed by the much greater risk for fracture produced by aging [32,64]. Therefore, our study highlights the importance of early dietary intervention. Health gains are predicted to be larger when earlier dietary changes are initiated in life. Sustained changes from a typical to an optimized diet from an early age could translate into an increase in life expectancy of ≥10 years [65]. In addition, we recommend that women aged ≥50 years should be the focus of the fracture population and take measures to enhance their intake and use of anti-inflammatory diets.

We observed that a less inflammatory diet reduced the risk of fractures in women with nonabdominal obesity. However, this result was not observed in women with abdominal obesity. Therefore, the effect of the regional accumulation of body fat on fractures needs to be considered. Intraabdominal obesity is an important risk factor for low-grade inflammation [66]. A preferential accumulation of visceral adipose tissue and adipose tissue promotes proinflammatory factor production, which drives chronic low-grade inflammation to cause diabetes, cardiovascular disease, and metabolic disease [66-69]. Abdominal obesity is highly prevalent in China [70,71]. Maintaining a good body weight and body shape is essential for the prevention of fractures.

Furthermore, we did not find an association between DII and fractures in women who smoke and drink. Smoking and excessive alcohol consumption induce oxidative stress to weaken the antioxidant effects and activate a variety of transcription factors such as inflammatory cytokines [72]. Smoking affects the intake of some antioxidants in food, such as vitamins C and E, to reduce the effect of an anti-inflammatory diet [73-76]. In addition, if people consume high amounts of alcohol, they may also consume high amounts of carbohydrates and energy, which have a proinflammatory score on the DII [39]. A cohort study in an Italian population showed that a healthy lifestyle, that is, adherence to a healthy dietary pattern, abstinence from smoking and regular engagement in physical activity, and has a positive effect on reducing the risk of all-cause mortality, especially when the combined effect of all 3 lifestyle behaviors was considered [77]. Therefore, we advocate smoking cessation and alcohol restriction and eat more anti-inflammatory foods to maximize the positive effect of diet on fracture prevention.

In addition, our study found that the DII scores of protein, niacin, and iron showed significant associations with fracture risk in women. Protein intake increases the fracture risk in women. Excessive protein intake can lead to dietary acid load, and an acidic diet can lead to osteoporosis and an increased risk of fractures [78,79]. However, recent systematic reviews have generally stated that protein intake reduces fracture risk [60]. Wallace et al [80] showed that a high dietary protein intake was associated with a 16% reduction in hip fracture risk compared with a low protein intake. In addition, Rizzoli et al [60] showed that inadequate protein intake in the diet may cause more disease problems than excessive protein intake. The niacin intake reduced the fracture risk in women. The potential beneficial mechanisms of action of niacin on bone are through decreased inflammation. Niacin decreases CRP levels, and higher CRP levels are associated with fractures [81]. B-vitamins appear to influence the development of collagen and alter the metabolism of osteoblasts in a dose-dependent manner [82,83]. Carbone et al [84] showed that dietary niacin intake was significantly associated with an increased risk of hip fracture per HRs (95% CI 1.01-1.24) with spline models, suggesting a U-shaped association. Clinical studies have shown that the incidence of osteoporosis and fractures in iron metabolism disorders is significantly increased and that iron deficiency affects bone metabolism. In healthy menopausal women, dietary iron is positively correlated with bone mineral density and may decrease the risk of fractures [85]. However, excess iron increases oxidative stress, causing inflammatory changes that destroy bones. Therefore, the relationship between iron and fracture needs further research [86]. It is worth noting that the intake of macronutrients (carbohydrates, total fat, and protein) was associated with fracture risk in men. However, this result was not observed when considering the overall diet. Therefore, it is reasonable to believe that a single nutrient does not have a significant impact on fracture risk. Because of the complexity of the interactions between diets, we believe that overall dietary optimization is an important way to improve bone health.

Strengths and Limitations

Our study has several strengths. First, we used a large-scale, nationally representative, prospective cohort design study to examine the association with dietary inflammatory potential and fractures. Second, we controlled for multiple confounding factors and performed subgroup and sensitivity analyses to ensure that the results were robust and reliable. Our results can provide data support for subsequent studies. Third, our research results, which demonstrated the relationship between diet and fracture risk in women aged 18 to 49 years, can supplement diet and fracture relationships among low age groups.

There are several limitations to this study that should be considered in with regard to its results. First, possible confounding variables existed and could have affected the outcome. There may be variables that can affect fractures such as osteoporosis or bone mineral density. Second, the fractures were self-reported and not clinically confirmed. This study did not distinguish between the cause and specific site of the fracture. Third, this study did not consider the probability of patients being lost to follow-up due to the fractures themselves, which may underestimate the incidence of fractures. Fourth, the correlation between the DII and fracture risk varies between men and women, which may be due to the incompatibility of the mechanisms of action of inflammatory effects in men and women, and further research is needed. Whether lowering the DII score in the future reduces the risk of fractures will require higher-quality randomized controlled trials to be validated.

Conclusions

In conclusion, proinflammatory diet consumption increased the fracture risk in Chinese women aged <50 years. The high consumption of anti-inflammatory foods and low consumption of proinflammatory foods may be an important strategy to prevent fractures in women. Future randomized controlled trials with diets rich in anti-inflammatory components are needed to confirm causality and to consider whether such interventions can reduce the incidence of fractures.

Acknowledgments

The authors wish to acknowledge funding from the National Natural Science Foundation of China (grant 81472970) and Beijing Municipal Commission of Health of China (grant 2022012).

The authors thank the staff and participants of the China Health and Nutrition Survey Study for their important contribution.

Data Availability

The data that support the findings of this study are available from the China Health and Nutrition Survey [87].

Authors' Contributions

ZZ and JL had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. ZZ, LW, and CY contributed to the acquisition, analysis, and interpretation of the data. LW contributed to manuscript writing. ZZ, JL, and FZ contributed to critical revision of the manuscript for important intellectual content. LW, CY, HW, and RW analyzed the data. All authors have approved the submission of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data tables and figures used to support the findings reported in the paper.
[DOCX File , 491 KB - publichealth_v9i1e43501_app1.docx]

References

1. Court-Brown CM, Caesar B. Epidemiology of adult fractures: a review. Injury 2006 Aug;37(8):691-697. [doi: 10.1016/j.injury.2006.04.130] [Medline: 16814787]
2. Polinder S, Haagsma J, Panneman M, Scholten A, Brugmans M, Van Beeck E. The economic burden of injury: health care and productivity costs of injuries in the Netherlands. Accid Anal Prev 2016 Aug;93:92-100. [doi: 10.1016/j.aap.2016.04.003] [Medline: 27177394]
3. Pike C, Birnbaum HG, Schiller M, Sharma H, Burge R, Edgell ET. Direct and indirect costs of non-vertebral fracture patients with osteoporosis in the US. Pharmacoeconomics 2010;28(5):395-409. [doi: 10.2165/11531040-000000000-00000] [Medline: 20402541]
4. Borgström F, Karlsson L, Ortsäter G, Norton N, Halbout P, Cooper C, International Osteoporosis Foundation. Fragility fractures in Europe: burden, management and opportunities. Arch Osteoporos 2020 Apr 19;15(1):59 [FREE Full text] [doi: 10.1007/s11657-020-0706-y] [Medline: 32306163]
5. Tatangelo G, Watts J, Lim K, Connaughton C, Abimanyi-Ochom J, Borgström F, et al. The cost of osteoporosis, osteopenia, and associated fractures in Australia in 2017. J Bone Miner Res 2019 Apr;34(4):616-625 [FREE Full text] [doi: 10.1002/jbmr.3640] [Medline: 30615801]
6. Hernlund E, Svedbom A, Ivergård M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos 2013;8(1):136 [FREE Full text] [doi: 10.1007/s11657-013-0136-1] [Medline: 24113837]
7. GBD 2019 Fracture Collaborators. Global, regional, and national burden of bone fractures in 204 countries and territories, 1990-2019: a systematic analysis from the global burden of disease study 2019. Lancet Healthy Longev 2021 Sep;2(9):e580-e592 [FREE Full text] [doi: 10.1016/S2666-7568(21)00172-0] [Medline: 34723233]

8. Wang L, Yu W, Yin X, Cui L, Tang S, Jiang N, et al. Prevalence of osteoporosis and fracture in China: the China osteoporosis prevalence study. JAMA Netw Open 2021 Aug 02;4(8):e2121106 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.21106] [Medline: 34398202]
9. Ballane G, Cauley JA, Luckey MM, Fuleihan GE. Secular trends in hip fractures worldwide: opposing trends East versus West. J Bone Miner Res 2014 Aug;29(8):1745-1755 [FREE Full text] [doi: 10.1002/jbmr.2218] [Medline: 24644018]
10. Zhang C, Feng J, Wang S, Gao P, Xu L, Zhu J, et al. Incidence of and trends in hip fracture among adults in urban China: a nationwide retrospective cohort study. PLoS Med 2020 Aug 06;17(8):e1003180 [FREE Full text] [doi: 10.1371/journal.pmed.1003180] [Medline: 32760065]
11. Dare AJ, Hu G. China's evolving fracture burden. Lancet Glob Health 2017 Aug;5(8):e736-e737 [FREE Full text] [doi: 10.1016/S2214-109X(17)30254-1] [Medline: 28666815]
12. Kanis JA, Cooper C, Rizzoli R, Reginster JY, Scientific Advisory Board of the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO) and the Committees of Scientific Advisors and National Societies of the International Osteoporosis Foundation (IOF). Executive summary of the European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Calcif Tissue Int 2019 Mar;104(3):235-238 [FREE Full text] [doi: 10.1007/s00223-018-00512-x] [Medline: 30796490]
13. Fang Y, Zhu J, Fan J, Sun L, Cai S, Fan C, et al. Dietary inflammatory index in relation to bone mineral density, osteoporosis risk and fracture risk: a systematic review and meta-analysis. Osteoporos Int 2021 Apr;32(4):633-643. [doi: 10.1007/s00198-020-05578-8] [Medline: 32740669]
14. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int 2006 Dec;17(12):1726-1733. [doi: 10.1007/s00198-006-0172-4] [Medline: 16983459]
15. Shen Y, Huang X, Wu J, Lin X, Zhou X, Zhu Z, et al. The global burden of osteoporosis, low bone mass, and its related fracture in 204 countries and territories, 1990-2019. Front Endocrinol (Lausanne) 2022 May 20;13:882241 [FREE Full text] [doi: 10.3389/fendo.2022.882241] [Medline: 35669691]
16. Ning P, Schwebel DC, Hu G. Healthy China 2030: a missed opportunity for injury control. Inj Prev 2017 Dec;23(6):363 [FREE Full text] [doi: 10.1136/injuryprev-2017-042314] [Medline: 28232400]
17. Barbour KE, Lui LY, Ensrud KE, Hillier TA, LeBlanc ES, Ing SW, Study of Osteoporotic Fractures (SOF) Research Group. Inflammatory markers and risk of hip fracture in older white women: the study of osteoporotic fractures. J Bone Miner Res 2014 Sep;29(9):2057-2064 [FREE Full text] [doi: 10.1002/jbmr.2245] [Medline: 24723386]
18. Barbour KE, Boudreau R, Danielson ME, Youk AO, Wactawski-Wende J, Greep NC, et al. Inflammatory markers and the risk of hip fracture: the women's health initiative. J Bone Miner Res 2012 May;27(5):1167-1176 [FREE Full text] [doi: 10.1002/jbmr.1559] [Medline: 22392817]
19. Eriksson AL, Movérare-Skrtic S, Ljunggren Ö, Karlsson M, Mellström D, Ohlsson C. High-sensitivity CRP is an independent risk factor for all fractures and vertebral fractures in elderly men: the MrOS Sweden study. J Bone Miner Res 2014 Feb;29(2):418-423 [FREE Full text] [doi: 10.1002/jbmr.2037] [Medline: 23857741]
20. Ishii S, Cauley JA, Greendale GA, Crandall CJ, Danielson ME, Ouchi Y, et al. C-reactive protein, bone strength, and nine-year fracture risk: data from the Study of Women's Health Across the Nation (SWAN). J Bone Miner Res 2013 Jul;28(7):1688-1698 [FREE Full text] [doi: 10.1002/jbmr.1915] [Medline: 23456822]
21. Cauley JA, Barbour KE, Harrison SL, Cloonan YK, Danielson ME, Ensrud KE, et al. Inflammatory markers and the risk of hip and vertebral fractures in men: the osteoporotic fractures in men (MrOS). J Bone Miner Res 2016 Dec;31(12):2129-2138 [FREE Full text] [doi: 10.1002/jbmr.2905] [Medline: 27371811]
22. Redlich K, Hayer S, Ricci R, David JP, Tohidast-Akrad M, Kollias G, et al. Osteoclasts are essential for TNF-alpha-mediated joint destruction. J Clin Invest 2002 Nov;110(10):1419-1427 [FREE Full text] [doi: 10.1172/JCI15582] [Medline: 12438440]
23. Zhang ZQ, Cao WT, Shivappa N, Hebert JR, Li BL, He J, et al. Association between diet inflammatory index and osteoporotic hip fracture in elderly Chinese population. J Am Med Dir Assoc 2017 Aug 01;18(8):671-677. [doi: 10.1016/j.jamda.2017.02.011] [Medline: 28411093]
24. Shivappa N, Steck SE, Hurley TG, Hussey JR, Hébert JR. Designing and developing a literature-derived, population-based dietary inflammatory index. Public Health Nutr 2014 Aug;17(8):1689-1696 [FREE Full text] [doi: 10.1017/S1368980013002115] [Medline: 23941862]
25. Zhao G, Etherton TD, Martin KR, West SG, Gillies PJ, Kris-Etherton PM. Dietary alpha-linolenic acid reduces inflammatory and lipid cardiovascular risk factors in hypercholesterolemic men and women. J Nutr 2004 Nov;134(11):2991-2997 [FREE Full text] [doi: 10.1093/jn/134.11.2991] [Medline: 15514264]
26. Esmaillzadeh A, Kimiagar M, Mehrabi Y, Azadbakht L, Hu FB, Willett WC. Dietary patterns and markers of systemic inflammation among Iranian women. J Nutr 2007 Apr;137(4):992-998 [FREE Full text] [doi: 10.1093/jn/137.4.992] [Medline: 17374666]
27. Ma Y, Hébert JR, Li W, Bertone-Johnson ER, Olendzki B, Pagoto SL, et al. Association between dietary fiber and markers of systemic inflammation in the women's health initiative observational study. Nutrition 2008 Oct;24(10):941-949 [FREE Full text] [doi: 10.1016/j.nut.2008.04.005] [Medline: 18562168]

28. Shivappa N, Steck SE, Hurley TG, Hussey JR, Ma Y, Ockene IS, et al. A population-based dietary inflammatory index predicts levels of C-reactive protein in the seasonal variation of blood cholesterol study (SEASONS). Public Health Nutr 2014 Aug;17(8):1825-1833 [FREE Full text] [doi: 10.1017/S1368980013002565] [Medline: 24107546]
29. Cavicchia PP, Steck SE, Hurley TG, Hussey JR, Ma Y, Ockene IS, et al. A new dietary inflammatory index predicts interval changes in serum high-sensitivity C-reactive protein. J Nutr 2009 Dec;139(12):2365-2372 [FREE Full text] [doi: 10.3945/jn.109.114025] [Medline: 19864399]
30. Shivappa N, Hébert JR, Rietzschel ER, De Buyzere ML, Langlois M, Debruyne E, et al. Associations between dietary inflammatory index and inflammatory markers in the Asklepios Study. Br J Nutr 2015 Feb 28;113(4):665-671 [FREE Full text] [doi: 10.1017/S000711451400395X] [Medline: 25639781]
31. Tabung FK, Steck SE, Zhang J, Ma Y, Liese AD, Agalliu I, et al. Construct validation of the dietary inflammatory index among postmenopausal women. Ann Epidemiol 2015 Jun;25(6):398-405 [FREE Full text] [doi: 10.1016/j.annepidem.2015.03.009] [Medline: 25900255]
32. Orchard T, Yildiz V, Steck SE, Hébert JR, Ma Y, Cauley JA, et al. Dietary inflammatory index, bone mineral density, and risk of fracture in postmenopausal women: results from the women's health initiative. J Bone Miner Res 2017 May;32(5):1136-1146 [FREE Full text] [doi: 10.1002/jbmr.3070] [Medline: 28019686]
33. Tabung FK, Steck SE, Liese AD, Zhang J, Ma Y, Caan B, et al. Association between dietary inflammatory potential and breast cancer incidence and death: results from the women's health initiative. Br J Cancer 2016 May 24;114(11):1277-1285 [FREE Full text] [doi: 10.1038/bjc.2016.98] [Medline: 27100730]
34. Wirth MD, Burch J, Shivappa N, Violanti JM, Burchfiel CM, Fekedulegn D, et al. Association of a dietary inflammatory index with inflammatory indices and metabolic syndrome among police officers. J Occup Environ Med 2014 Sep;56(9):986-989 [FREE Full text] [doi: 10.1097/JOM.0000000000000213] [Medline: 25046320]
35. Tabung FK, Steck SE, Ma Y, Liese AD, Zhang J, Caan B, et al. The association between dietary inflammatory index and risk of colorectal cancer among postmenopausal women: results from the Women's Health Initiative. Cancer Causes Control 2015 Mar;26(3):399-408 [FREE Full text] [doi: 10.1007/s10552-014-0515-y] [Medline: 25549833]
36. Wood LG, Shivappa N, Berthon BS, Gibson PG, Hebert JR. Dietary inflammatory index is related to asthma risk, lung function and systemic inflammation in asthma. Clin Exp Allergy 2015 Jan;45(1):177-183 [FREE Full text] [doi: 10.1111/cea.12323] [Medline: 24708388]
37. Hariharan R, Odjidja EN, Scott D, Shivappa N, Hébert JR, Hodge A, et al. The dietary inflammatory index, obesity, type 2 diabetes, and cardiovascular risk factors and diseases. Obes Rev 2022 Jan;23(1):e13349. [doi: 10.1111/obr.13349] [Medline: 34708499]
38. Vissers LE, Waller M, van der Schouw YT, Hébert JR, Shivappa N, Schoenaker DA, et al. A pro-inflammatory diet is associated with increased risk of developing hypertension among middle-aged women. Nutr Metab Cardiovasc Dis 2017 Jun;27(6):564-570. [doi: 10.1016/j.numecd.2017.03.005] [Medline: 28446366]
39. Kim HS, Kwon M, Lee HY, Shivappa N, Hébert JR, Sohn C, et al. Higher pro-inflammatory dietary score is associated with higher hyperuricemia risk: results from the case-controlled Korean genome and epidemiology Study_Cardiovascular disease association study. Nutrients 2019 Aug 05;11(8):1803 [FREE Full text] [doi: 10.3390/nu11081803] [Medline: 31387233]
40. Veronese N, Stubbs B, Koyanagi A, Hébert JR, Cooper C, Caruso MG, et al. Pro-inflammatory dietary pattern is associated with fractures in women: an eight-year longitudinal cohort study. Osteoporos Int 2018 Jan;29(1):143-151 [FREE Full text] [doi: 10.1007/s00198-017-4251-5] [Medline: 29018920]
41. Cervo MM, Shivappa N, Hebert JR, Oddy WH, Winzenberg T, Balogun S, et al. Longitudinal associations between dietary inflammatory index and musculoskeletal health in community-dwelling older adults. Clin Nutr 2020 Feb;39(2):516-523. [doi: 10.1016/j.clnu.2019.02.031] [Medline: 30852031]
42. Morimoto M, Shivappa N, de Souza Genaro P, Martini LA, Schuch NJ, Hebert JR, et al. Lack of association between dietary inflammatory index and low impact fractures in the Brazilian population: the Brazilian Osteoporosis Study (BRAZOS). Adv Rheumatol 2019 Apr 11;59(1):16. [doi: 10.1186/s42358-019-0059-7] [Medline: 30971320]
43. Su Y, Yeung SS, Chen YM, Leung JC, Kwok TC. The associations of dietary inflammatory potential with musculoskeletal health in Chinese community-dwelling older people: the Mr. OS and Ms. OS (Hong Kong) cohort study. J Bone Miner Res 2022 Jun;37(6):1179-1187 [FREE Full text] [doi: 10.1002/jbmr.4556] [Medline: 35416312]
44. He J, Fang A, Yu S, Shen X, Li K. Dietary nonheme, Heme, and total iron intake and the risk of diabetes in adults: results from the China health and nutrition survey. Diabetes Care 2020 Apr;43(4):776-784 [FREE Full text] [doi: 10.2337/dc19-2202] [Medline: 32033991]
45. He Y, Li Y, Yang X, Hemler EC, Fang Y, Zhao L, et al. The dietary transition and its association with cardiometabolic mortality among Chinese adults, 1982-2012: a cross-sectional population-based study. Lancet Diabetes Endocrinol 2019 Jul;7(7):540-548 [FREE Full text] [doi: 10.1016/S2213-8587(19)30152-4] [Medline: 31085143]
46. Yang Y, Hozawa A, Kogure M, Narita A, Hirata T, Nakamura T, et al. Dietary inflammatory index positively associated with high-sensitivity C-reactive protein level in Japanese from NIPPON DATA2010. J Epidemiol 2020 Feb 05;30(2):98-107 [FREE Full text] [doi: 10.2188/jea.JE20180156] [Medline: 30745493]

47. Qin Y, Chen Y, Zhang Z, Huang S, Jiao C, Zhang Z, et al. Associations of dietary inflammatory potential with postpartum weight change and retention: results from a cohort study. Obesity (Silver Spring) 2021 Oct;29(10):1689-1699. [doi: 10.1002/oby.23238] [Medline: 34490743]
48. Tomata Y, Shivappa N, Zhang S, Nurrika D, Tanji F, Sugawara Y, et al. Dietary inflammatory index and disability-free survival in community-dwelling older adults. Nutrients 2018 Dec 03;10(12):1896 [FREE Full text] [doi: 10.3390/nu10121896] [Medline: 30513971]
49. Danielsson O, Nissinen MJ, Jula A, Salomaa V, Männistö S, Lundqvist A, et al. Waist and hip circumference are independently associated with the risk of liver disease in population-based studies. Liver Int 2021 Dec;41(12):2903-2913. [doi: 10.1111/liv.15053] [Medline: 34510711]
50. Hu FJ, Liu H, Liu XL, Jia SL, Hou LS, Xia X, et al. Mid-upper arm circumference as an alternative screening instrument to appendicular skeletal muscle mass index for diagnosing sarcopenia. Clin Interv Aging 2021 Jun 15;16:1095-1104 [FREE Full text] [doi: 10.2147/CIA.S311081] [Medline: 34163153]
51. Koh LK, Sedrine WB, Torralba TP, Kung A, Fujiwara S, Chan SP, Osteoporosis Self-Assessment Tool for Asians (OSTA) Research Group. A simple tool to identify asian women at increased risk of osteoporosis. Osteoporos Int 2001;12(8):699-705. [doi: 10.1007/s001980170070] [Medline: 11580084]
52. Marin RV, Pedrosa MA, Moreira-Pfrimer LD, Matsudo SM, Lazaretti-Castro M. Association between lean mass and handgrip strength with bone mineral density in physically active postmenopausal women. J Clin Densitom 2010 Jan;13(1):96-101. [doi: 10.1016/j.jocd.2009.12.001] [Medline: 20171571]
53. Zhang B, Zhai FY, Du SF, Popkin BM. The China Health and Nutrition Survey, 1989-2011. Obes Rev 2014 Jan;15 Suppl 1(0 1):2-7 [FREE Full text] [doi: 10.1111/obr.12119] [Medline: 24341753]
54. Cawthon PM. Gender differences in osteoporosis and fractures. Clin Orthop Relat Res 2011 Jul;469(7):1900-1905 [FREE Full text] [doi: 10.1007/s11999-011-1780-7] [Medline: 21264553]
55. Vegeto E, Benedusi V, Maggi A. Estrogen anti-inflammatory activity in brain: a therapeutic opportunity for menopause and neurodegenerative diseases. Front Neuroendocrinol 2008 Oct;29(4):507-519 [FREE Full text] [doi: 10.1016/j.yfrne.2008.04.001] [Medline: 18522863]
56. Eskes T, Haanen C. Why do women live longer than men? Eur J Obstet Gynecol Reprod Biol 2007 Aug;133(2):126-133. [doi: 10.1016/j.ejogrb.2007.01.006] [Medline: 17324494]
57. Pandey AS, Shreevastva NK, Neupane DP. Nicotine exposure, blood pressure, and inflammation in tobacco smokers and chewers in a rural community in Nepal. Subst Use Misuse 2014 Jun;49(7):798-803. [doi: 10.3109/10826084.2014.880483] [Medline: 24491150]
58. Dahl K, Ahmed LA, Joakimsen RM, Jørgensen L, Eggen AE, Eriksen EF, et al. High-sensitivity C-reactive protein is an independent risk factor for non-vertebral fractures in women and men: the Tromsø study. Bone 2015 Mar;72:65-70. [doi: 10.1016/j.bone.2014.11.012] [Medline: 25460573]
59. Cummings SR, Cawthon PM, Ensrud KE, Cauley JA, Fink HA, Orwoll ES, Osteoporotic Fractures in Men (MrOS) Research Groups, Study of Osteoporotic Fractures Research Groups. BMD and risk of hip and nonvertebral fractures in older men: a prospective study and comparison with older women. J Bone Miner Res 2006 Oct;21(10):1550-1556 [FREE Full text] [doi: 10.1359/jbmr.060708] [Medline: 16995809]
60. Rizzoli R, Biver E, Brennan-Speranza TC. Nutritional intake and bone health. Lancet Diabetes Endocrinol 2021 Sep;9(9):606-621. [doi: 10.1016/S2213-8587(21)00119-4] [Medline: 34242583]
61. Huang YP, Chen LS, Feng SH, Liang YS, Pan SL. Tea consumption and the risks of osteoporosis and hip fracture: a population-based longitudinal follow-up study. Osteoporos Int 2023 Jan;34(1):101-109 [FREE Full text] [doi: 10.1007/s00198-022-06569-7] [Medline: 36241848]
62. Pacifici R. Estrogen, cytokines, and pathogenesis of postmenopausal osteoporosis. J Bone Miner Res 1996 Aug;11(8):1043-1051. [doi: 10.1002/jbmr.5650110802] [Medline: 8854239]
63. Wu D, Cline-Smith A, Shashkova E, Perla A, Katyal A, Aurora R. T-cell mediated inflammation in postmenopausal osteoporosis. Front Immunol 2021 Jun 30;12:687551 [FREE Full text] [doi: 10.3389/fimmu.2021.687551] [Medline: 34276675]
64. Robbins J, Aragaki AK, Kooperberg C, Watts N, Wactawski-Wende J, Jackson RD, et al. Factors associated with 5-year risk of hip fracture in postmenopausal women. JAMA 2007 Nov 28;298(20):2389-2398. [doi: 10.1001/jama.298.20.2389] [Medline: 18042916]
65. Fadnes LT, Økland JM, Haaland Ø, Johansson KA. Estimating impact of food choices on life expectancy: a modeling study. PLoS Med 2022 Feb 08;19(2):e1003889 [FREE Full text] [doi: 10.1371/journal.pmed.1003889] [Medline: 35134067]
66. Strasser B, Arvandi M, Siebert U. Resistance training, visceral obesity and inflammatory response: a review of the evidence. Obes Rev 2012 Jul;13(7):578-591. [doi: 10.1111/j.1467-789X.2012.00988.x] [Medline: 22385646]
67. Tchernof A, Després JP. Pathophysiology of human visceral obesity: an update. Physiol Rev 2013 Jan;93(1):359-404 [FREE Full text] [doi: 10.1152/physrev.00033.2011] [Medline: 23303913]
68. Engström G, Hedblad B, Stavenow L, Jonsson S, Lind P, Janzon L, et al. Incidence of obesity-associated cardiovascular disease is related to inflammation-sensitive plasma proteins: a population-based cohort study. Arterioscler Thromb Vasc Biol 2004 Aug;24(8):1498-1502. [doi: 10.1161/01.ATV.0000134293.31512.be] [Medline: 15178569]

69. Wan H, Wang Y, Xiang Q, Fang S, Chen Y, Chen C, et al. Associations between abdominal obesity indices and diabetic complications: Chinese visceral adiposity index and neck circumference. Cardiovasc Diabetol 2020 Jul 31;19(1):118 [FREE Full text] [doi: 10.1186/s12933-020-01095-4] [Medline: 32736628]
70. Zhang L, Wang Z, Wang X, Chen Z, Shao L, Tian Y, China Hypertension Survey investigators. Prevalence of abdominal obesity in China: results from a cross-sectional study of nearly half a million participants. Obesity (Silver Spring) 2019 Nov;27(11):1898-1905. [doi: 10.1002/oby.22620] [Medline: 31549787]
71. Zhang X, Zhang M, Zhao Z, Huang Z, Deng Q, Li Y, et al. Geographic variation in prevalence of adult obesity in China: results from the 2013-2014 national chronic disease and risk factor surveillance. Ann Intern Med 2020 Feb 18;172(4):291-293. [doi: 10.7326/M19-0477] [Medline: 31658469]
72. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative stress, inflammation, and cancer: how are they linked? Free Radic Biol Med 2010 Dec 01;49(11):1603-1616 [FREE Full text] [doi: 10.1016/j.freeradbiomed.2010.09.006] [Medline: 20840865]
73. Dyer AR, Elliott P, Stamler J, Chan Q, Ueshima H, Zhou BF, INTERMAP Research Group. Dietary intake in male and female smokers, ex-smokers, and never smokers: the INTERMAP study. J Hum Hypertens 2003 Sep;17(9):641-654 [FREE Full text] [doi: 10.1038/sj.jhh.1001607] [Medline: 13679954]
74. Phillips EL, Arnett DK, Himes JH, McGovern PG, Blackburn H, Luepker RV. Differences and trends in antioxidant dietary intake in smokers and non-smokers, 1980-1992: the Minnesota Heart Survey. Ann Epidemiol 2000 Oct;10(7):417-423. [doi: 10.1016/s1047-2797(00)00064-8] [Medline: 11018344]
75. Thuesen B, Husemoen L, Fenger M, Jakobsen J, Schwarz P, Toft U, et al. Determinants of vitamin D status in a general population of Danish adults. Bone 2012 Mar;50(3):605-610. [doi: 10.1016/j.bone.2011.12.016] [Medline: 22227435]
76. Brot C, Jorgensen NR, Sorensen OH. The influence of smoking on vitamin D status and calcium metabolism. Eur J Clin Nutr 1999 Dec;53(12):920-926. [doi: 10.1038/sj.ejcn.1600870] [Medline: 10602348]
77. Prinelli F, Yannakoulia M, Anastasiou CA, Adorni F, Di Santo SG, Musicco M, et al. Mediterranean diet and other lifestyle factors in relation to 20-year all-cause mortality: a cohort study in an Italian population. Br J Nutr 2015 Mar 28;113(6):1003-1011. [doi: 10.1017/S0007114515000318] [Medline: 25746109]
78. Papageorgiou M, Merminod F, Chevalley T, van Rietbergen B, Ferrari S, Rizzoli R, et al. Associations between age-related changes in bone microstructure and strength and dietary acid load in a cohort of community-dwelling, healthy men and postmenopausal women. Am J Clin Nutr 2020 Oct 01;112(4):1120-1131 [FREE Full text] [doi: 10.1093/ajcn/nqaa191] [Medline: 32678420]
79. Pedone C, Napoli N, Pozzilli P, Lauretani F, Bandinelli S, Ferrucci L, et al. Quality of diet and potential renal acid load as risk factors for reduced bone density in elderly women. Bone 2010 Apr;46(4):1063-1067 [FREE Full text] [doi: 10.1016/j.bone.2009.11.031] [Medline: 20005315]
80. Wallace TC, Frankenfeld CL. Dietary protein intake above the current RDA and bone health: a systematic review and meta-analysis. J Am Coll Nutr 2017 Aug;36(6):481-496. [doi: 10.1080/07315724.2017.1322924] [Medline: 28686536]
81. Berglundh S, Malmgren L, Luthman H, McGuigan F, Åkesson K. C-reactive protein, bone loss, fracture, and mortality in elderly women: a longitudinal study in the OPRA cohort. Osteoporos Int 2015 Feb;26(2):727-735. [doi: 10.1007/s00198-014-2951-7] [Medline: 25410434]
82. Kim GS, Kim CH, Park JY, Lee KU, Park CS. Effects of vitamin B12 on cell proliferation and cellular alkaline phosphatase activity in human bone marrow stromal osteoprogenitor cells and UMR106 osteoblastic cells. Metabolism 1996 Dec;45(12):1443-1446. [doi: 10.1016/s0026-0495(96)90171-7] [Medline: 8969275]
83. Bailey RL, van Wijngaarden JP. The role of B-vitamins in bone health and disease in older adults. Curr Osteoporos Rep 2015 Aug;13(4):256-261. [doi: 10.1007/s11914-015-0273-0] [Medline: 26017584]
84. Carbone LD, Bůžková P, Fink HA, Raiford M, Le B, Isales CM, et al. Association of dietary niacin intake with incident hip fracture, BMD, and body composition: the cardiovascular health study. J Bone Miner Res 2019 Apr;34(4):643-652 [FREE Full text] [doi: 10.1002/jbmr.3639] [Medline: 30659655]
85. Lu M, Liu Y, Shao M, Tesfaye GC, Yang S. Associations of iron intake, serum iron and serum ferritin with bone mineral density in women: the national health and nutrition examination survey, 2005-2010. Calcif Tissue Int 2020 Mar;106(3):232-238. [doi: 10.1007/s00223-019-00627-9] [Medline: 31754762]
86. Harris MM, Houtkooper LB, Stanford VA, Parkhill C, Weber JL, Flint-Wagner H, et al. Dietary iron is associated with bone mineral density in healthy postmenopausal women. J Nutr 2003 Nov;133(11):3598-3602. [doi: 10.1093/jn/133.11.3598] [Medline: 14608080]
87. Home page. China Health and Nutrition Survey. URL: https://www.cpc.unc.edu/projects/china [accessed 2023-07-21]

Abbreviations

CHNS: China Health and Nutrition Survey
CRP: C-reactive protein
DII: dietary inflammatory index
HR: hazard ratio

IL: interleukin
MAMC: midarm muscle circumference
OSTA: Osteoporosis Self-assessment Tool for Asians
PAL: physical activity level
RCS: restricted cubic spline
TNF: tumor necrosis factor
WHR: waist-to-hip ratio

Edited by A Mavragani, T Sanchez; submitted 13.10.22; peer-reviewed by J Delgado-Ron, R Ciptaningtyas; comments to author 19.05.23; revised version received 06.06.23; accepted 21.06.23; published 17.08.23.

Please cite as:
Wang L, Ye C, Zhao F, Wu H, Wang R, Zhang Z, Li J
Association Between the Dietary Inflammatory Index and the Risk of Fracture in Chinese Adults: Longitudinal Study
JMIR Public Health Surveill 2023;9:e43501
URL: https://publichealth.jmir.org/2023/1/e43501
doi:10.2196/43501
PMID:37590048

Original Paper

Patterns and Trends in Mortality Associated With and Due to Diabetes Mellitus in a Transitioning Region With 3.17 Million People: Observational Study

Xiaopan Li[1*], PhD; Ru Liu[1,2*], MA; Yichen Chen[3,4*], PhD; Yan Han[1], MA; Qizhe Wang[1,2], MA; Yaxin Xu[1,2], MA; Jing Zhou[1], MA; Sunfang Jiang[1,2], MD

[1]Department of Health Management Center, Zhongshan Hospital, Fudan University, Shanghai, China
[2]Department of General Practice, Zhongshan Hospital, Fudan University, Shanghai, China
[3]Office of Scientific Research and Information Management, Center for Disease Control and Prevention, Pudong New Area, Shanghai, China
[4]School of Public Health, Fudan University, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Sunfang Jiang, MD
Department of General Practice
Zhongshan Hospital
Fudan University
180 Fenglin Rd
Shanghai, 200233
China
Phone: 86 13817597264
Fax: 86 021 31298409
Email: jiang.sunfang@zs-hospital.sh.cn

Abstract

Background: Diabetes mellitus (DM) imposes a significant disease burden in economically transitioning regions. Most transitioning regions share similar experience in urbanization processes. Shanghai's Pudong district serves as a representative area of such regions.

Objective: We aimed to assess the burden of and trends in DM mortality in Shanghai's Pudong district and analyze the impact of aging and multimorbidity.

Methods: A longitudinal, population-based study was conducted to analyze DM mortality in Pudong from 2005 to 2020. We used joinpoint regression to analyze epidemiological features and long-term trends in crude mortality rate (CMR), age-standardized mortality rate worldwide (ASMRW), and years of life lost (YLL). Furthermore, the decomposition method was used to evaluate the contribution of demographic and nondemographic factors associated with mortality.

Results: There were 49,414 deaths among individuals with DM, including 15,512 deaths due to DM. The CMR and ASMRW were 109.55/10^5 and 38.01/10^5 person-years, respectively. Among the mortality associated with and due to DM, the total annual ASMRW increased by 3.65% (95% CI 3.25%-4.06%) and 1.38% (95% CI 0.74%-2.02%), respectively. Additionally, the total annual YLL rate increased by 4.98% (95% CI 3.92%-6.05%) and 2.68% (95% CI 1.34%-4.04%). The rates of YLL increase in persons aged 30 to 44 years (3.98%, 95% CI 0.32%-7.78%) and 45 to 59 years (4.31%, 95% CI 2.95%-5.69%) were followed by the increase in persons aged 80 years and older (10.53%, 95% CI 9.45%-11.62%) for deaths associated with DM. The annual CMR attributable to demographic factors increased by 41.9% (95% CI 17.73%-71.04%) and 36.72% (95% CI 16.69%-60.2%) for deaths associated with and due to DM, respectively. Hypertension, cerebrovascular disease, and ischemic heart disease were the top 3 comorbidities.

Conclusions: Aging and multimorbidity played essential roles in changing the burden of DM in an urbanizing and transitioning region. There is an increasing disease burden among young and middle-aged people, emphasizing the need for greater attention to these groups. Health management is an emerging method that holds important implications for alleviating the future burden of DM.

(JMIR Public Health Surveill 2023;9:e43687) doi:10.2196/43687

KEYWORDS

diabetes mellitus; mortality; years of life lost; multimorbidity; trend analysis; diabetes; disease; urbanization; aging; epidemiology

Introduction

Diabetes mellitus (DM) is a state of metabolic disorder characterized by hyperglycemia that can lead to damage in the heart, blood vessels, kidneys, eyes, and nerves [1]. Worldwide, the number of people with DM has quadrupled in the last 3 decades, and DM ranks ninth among the major causes of death [2]. In 2021, according to the *IDF Atlas*, *Tenth Edition*, an estimated 1 of 10 people aged 20 to 79 years had diabetes, and approximately 75% of people with DM lived in developing countries [3]. In China, the prevalence of DM increases with age; about 11.2% of people aged 18 years and older were living with DM from 2015 to 2017 [4]. Additionally, most patients with DM experience at least one complication that can be life-threatening [5]. DM is associated with an 8-year reduction in lifespan in the United States and has a negative effect on the quality of life, given the high incidence of complications. Approximately 12% of worldwide health expenditure in 2015 was spent on the treatment of DM and related complications [6].

With the strengthening of reforms and opening up, Shanghai's Pudong district, with a population of 3.17 million, has become a center for trade, economic development, scientific advancements, and technological innovations; it has also witnessed notable improvements in public services, including enhanced social security coverage, long-term care, and education and health services. Pudong serves as a representative transitioning region that has experienced dramatic improvements in living standards. However, with these improvements, many risk factors for the incidence and progression of DM have increased, such as unhealthy eating habits, sedentary lifestyles, psychosocial stress, depression, and overweight or obesity [7]. The urbanization process in Pudong has substantially impacted the disease spectrum. Considering that other transitioning regions are likely to undergo similar urbanization processes, insights into the burden of DM and trends in DM-related mortality in Pudong can inform the management and planning of health care services in these regions in the future [8].

Without sufficient studies on disease-related and specific deaths, estimating the true burden of a disease is challenging [9]. It is also crucial to differentiate mortality associated with DM and due to DM. Such differentiation is essential for assessing the effectiveness of DM treatment and obtain valuable insights for policymakers in developing regions, enabling them to devise new health care strategies that address unmet needs. Therefore, our study aimed to assess the burden of DM mortality, examine its trends, and investigate the impacts of population aging and multimorbidity in Pudong, a transitional region undergoing urbanization.

Methods

Data Source

The study population consisted of residents of Pudong, Shanghai, with DM who had died and for whom data were available in the Mortality Registration System from 2005 to 2020. We selected Pudong as the study site for 3 reasons. First, Pudong encompasses both urban and rural subregions, covering an area of 1210.41 km^2 with a population of 3.17 million [10]. Second, people older than 65 years account for 23.73% of the total registered population in Pudong. The large population base and significant proportion of hyper-aged residents allow for a more realistic evaluation of the epidemiological features of deaths associated with DM. Third, the gross domestic product in Pudong increased from RMB 210.8 billion (US $29.4 billion) in 2005 to RMB 1320.7 billion (US $184.2 billion) in 2020, with an average annual growth rate of 13.01% [11]. Therefore, with the fastest economic development among Shanghai districts, Pudong is recognized as a microcosm of China's urbanization. The registration system is administered by the Public Security Bureau and the Statistics Bureau of Pudong. The system contains comprehensive population demographic characteristics, including age, sex, and cause of death. To maintain the integrity of the registration system, regular evaluations and data cleaning are carried out. DM was defined according to the *International Classification of Diseases, Tenth revision* (*ICD-10*) codes E10 to E14, and DM-related deaths were classified according to whether DM was a direct or indirect cause of death. More specifically, we screened all causes of death that included DM and defined them as death associated with DM. There were 4 listed direct causes of death on each certificate, one of which was used to define death due to DM. Death due to DM was determined by clinical doctors according to the patients' situation and relevant examination. The causes of death were independently reviewed by rigorously trained clinicians from the local Center for Disease Control and Prevention (CDC) based on the actual circumstances of each patient.

Ethics Approval

This surveillance protocol was approved by the Ethics Committee of the Shanghai Pudong CDC (2016-04-0586). The study did not involve any intervention among participants, and all data were maintained in strict confidentiality.

ICD-10 Codes

This study used the *ICD-10* codes E10 to E14 to identify causes of death related to DM. These codes include insulin-dependent DM (E10), non–insulin-dependent diabetes of the young (E11), malnutrition-related DM (E12), other specified DM (E13), and unspecified DM (E14). In addition, other diseases were classified as complications according to the *ICD-10*. To account for potential variation in data quality across different years, we followed a recommended method to evaluate data quality [12].

The method helps to improve the reliability of the results by defining and categorizing “garbage codes.”

Statistical Analyses

Specific-cause mortality was calculated based on deaths with DM as the underlying cause of death (ie, deaths due to DM), while all-cause mortality was calculated based on deaths with DM among all causes of death (deaths associated with DM).

The crude mortality rate (CMR) and age-standardized mortality rate worldwide (ASMRW) for DM were calculated per 100,000 persons (/10^5) and stratified by sex and year. We used the Poisson approximation method and the Mantel-Haenszel test to compare the CMR and ASMRW between sexes, respectively. Years of life lost (YLL) was used as a key measure to quantify the burden of disease and was calculated according to the original method described by Murray and Lopez that was detailed previously by Ye et al [13]. The YLL formula adopted by the World Health Organization (WHO) [14,15] was applied. Mortality rates were calculated for the following age groups: 0 to 4 years, 5 to 14 years, 15 to 29 years, 30 to 44 years, 45 to 59 years, 60 to 69 years, 70 to 79 years, and ≥80 years. Due to there being few DM-related deaths before the age of 30, trends in age-specific CMR, ASMRW, and YLL were calculated for the following age groups: 0 to 29 years, 30 to 44 years, 45 to 59 years, 60 to 69 years, 70 to 79 years, and ≥80 years.

Joinpoint regression, a popular approach for analyzing changes in trends, was used in this study. We modeled the time series using continuous linear segments and minimized the weighted sum of squared errors and the number of joinpoints using permutation tests [16]. We used the Joinpoint Regression Program (version 4.3.1.0; National Cancer Institute) to calculate trends in CMR, ASMRW, and YLL over time. The results were presented as the average annual percent change (AAPC) with corresponding 95% CI. Using the *z* test, we determined whether the annual percent change (APC) significantly differed from zero [17]. The difference decomposition method was used to quantify determinants of demographic and nondemographic factors. The rates of DM distribution in different groups were calculated and presented as percentages. In the results section, we report statistically significant ($P<.05$) changes as increases and decreases in AAPC and nonsignificant trends as stable. All statistical analyses were performed using SPSS (version 21.0; IBM Inc) and R (version 3.4.3; R Core Team). The significance level was set at $P<.05$.

Results

Baseline Characteristics

In this study, we extracted 45,107,809 records at the person-year level from residents in Pudong from 2005 to 2020; during this period, 336,823 deaths occurred. Among these deaths, 49,414 were associated with DM, including 15,512 deaths due to DM (Figure 1 and Multimedia Appendix 1, Figure S1). Approximately equal representation of men and women was observed among deaths associated with DM, with women accounting for 51.01% of deaths. The average age at death was 77.83 (SD 10.72) years, and the median age at death was 79.83 years. The total CMR and ASMRW of deaths associated with DM were 109.55/10^5 and 38.01/10^5 person-years, respectively. Among men, the CMR and ASMRW were 107.54/10^5 person-years and 43.84/10^5 person-years, respectively; among women, the corresponding figures were 111.55/10^5 person-years and 32.61/10^5 person-years. Among the individuals who died due to DM, 53.59% were women; the CMR and ASMRW in women were 36.79/10^5 person-years and 10.79/10^5 person-years, respectively. The YLL due to premature death was 425,938.83 years, and the YLL rate was 944.27/10^5 person-years for deaths associated with DM. The YLL and YLL rates in men (217,230.26 years and 964.92/10^5 person-years) were higher than in women (208,708.56 years and 923.69/10^5 person-years; Table 1). The YLL and YLL rates increased as age exceeded ≥80 years. The top 3 age groups in terms of YLL and YLL rates of deaths associated with DM were 70 to 79 years, ≥80 years, and 60 to 69 years, with YLL of 133,784.59, 125,400.22, and 100,335.50, respectively, and YLL rates of 4088.18/10^5, 6668.12/10^5, and 1614.79/10^5, respectively; Multimedia Appendix 1, Table S1).

Figure 1. Flow chart of deaths selected from among all deaths in Pudong, Shanghai, from 2005 to 2020. DM: Diabetes mellitus.

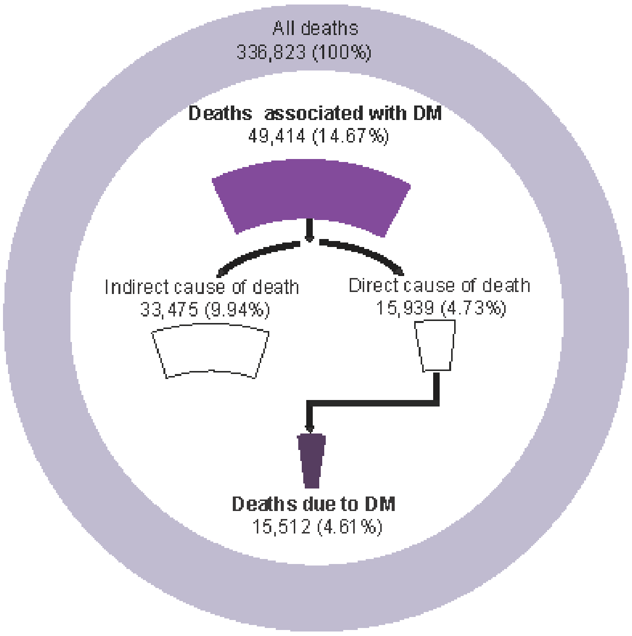


Table 1. Baseline characteristics of people who died of causes associated with and due to diabetes mellitus from 2005 to 2020.

Characteristic	Deaths, n (%)	Age at death (years)		CMR[a], /10^5	ASMRW[b], /10^5	YLL[c], years	YLL rate, /10^5
		Mean (SD)	Median (range)				
Deaths associated with DM[d]							
Gender							
Male	24,210 (48.99)	75.61 (11.19)	77.54 (17.22-105.90)	107.54	43.84	217,230.26	964.92
Female	25,204 (51.01)	79.96 (9.78)	81.69 (21.76-105.65)	111.55	32.61	208,708.56	923.69
Total	49,414 (100)	77.83 (10.72)	79.83 (17.22-105.90)	109.55	38.01	425,938.83	944.27
Top 3 comorbidities among deaths associated with DM							
Ischemic heart disease (*ICD-10*[e] codes I20-I25)	9075 (18.37)	80.82 (9.40)	82.38 (27.15-105.19)	20.12	6.39	67,544.01	149.74
Cerebrovascular disease (*ICD-10* codes I60-I69)	8695 (17.60)	78.94 (9.69)	80.75 (30.70-105.65)	19.28	6.39	70,585.83	156.48
Lung cancer (*ICD-10* codes C33-C34)	2369 (4.79)	73.34 (9.37)	74.00 (32.70-98.59)	5.25	2.08	24,181.69	53.61
Deaths due to DM							
Gender							
Male	7199 (46.41)	75.38 (12.01)	77.69 (20.36-101.86)	31.98	13.20	65,481.23	290.86
Female	8313 (53.59)	80.01 (10.26)	81.79 (21.76-105.07)	36.79	10.79	68,850.69	304.72
Total	15,512 (100)	77.86 (11.34)	80.13 (20.36-105.07)	34.39	11.98	134,331.92	297.80
Top 3 comorbidities among deaths due to DM							
Other diseases of the respiratory system (*ICD-10* codes J95-J99)	4062 (26.19)	79.53 (10.00)	81.37 (21.76-103.35)	9.01	2.96	32,091.78	71.14
Cerebrovascular disease (*ICD-10* codes I60-I69)	3841 (24.76)	78.39 (9.51)	79.95 (32.34-101.97)	8.52	2.87	31,956.80	70.85
Ischemic heart diseases (*ICD-10* codes I20-I25)	2953 (19.04)	79.86 (9.38)	81.39 (27.69-103.48)	6.55	2.13	23,054.70	51.11

[a] CMR: crude mortality rate.

[b] ASMRW: age-standardized mortality rate worldwide.

[c] YLL: years of life lost.

[d] DM: diabetes mellitus.

[e] *ICD-10*: *International Classification of Diseases, Tenth revision.*

Main Comorbidities

Among the deaths associated with DM, the top 3 comorbidities were ischemic heart disease (*ICD-10* codes I20-I25; 9075/49,414, 18.37%), cerebrovascular disease (*ICD-10* codes I60-I69; 8695/49,414, 17.6%) and lung cancer (*ICD-10* codes C33-C34; 2369/49,414, 4.79%). The YLL and YLL rates due to cerebrovascular disease were the highest (70,585.83 years and $156.48/10^5$, respectively), followed by ischemic heart disease (67,544.01 years and $149.74/10^5$; Table 1). The age at death, CMR, and ASMRW of these comorbidities are presented in Table 1.

The top 10 comorbidities among deaths associated with DM were the same for both men and women, with only the ranking differing; the sex-specific ranking is presented in Table 2. The ranking of comorbidities among deaths due to DM was also similar between men and women except for the ninth comorbidity, as shown in Table 3. On average, each person who died due to DM had 2.09 different comorbidities throughout their life. The disease spectrum of deaths associated with DM and due to DM is shown in Multimedia Appendix 1, Table S2.

Table 2. Top 10 comorbidities among men and women who died of causes associated with diabetes mellitus.

Type and cause of death (*ICD-10*[a] code)	Men (n=24,210), n (%)	Women (n=25,204), n (%)
Diabetes mellitus (E10-14)	7199 (29.74)	8313 (32.98)
Cerebrovascular disease (I60-69)	4056 (16.75)	4639 (18.41)
Ischemic heart disease (I20-25)	3993 (16.49)	5082 (20.16)
Lung cancer (C33-34)	1669 (6.89)	700 (2.78)
Chronic lower respiratory disease (J40-47)	1459 (6.03)	419 (1.66)
Liver cancer (C22)	711 (2.94)	419 (1.66)
Colorectal cancer (C18-21)	582 (2.40)	473 (1.88)
Malignant neoplasm of the pancreas (C25)	549 (2.27)	503 (2.00)
Stomach cancer (N40-51)	530 (2.19)	357 (1.42)
Falls (W00-19)	256 (1.06)	326 (1.29)
Other diseases[b]	3206 (13.24)	4639 (18.41)

[a] *ICD-10*: *International Classification of Diseases, Tenth revision*.

[b] Other diseases are listed in Multimedia Appendix 1, Table S2.

Table 3. Top 10 comorbidities among men and women who died due to diabetes mellitus.

Type and cause of death (*ICD-10*[a] code)	Men (n=15,419), n (%)	Women (n=16,958), n (%)
Hypertensive diseases (I10-15)	3520 (22.81)	4151 (24.48)
Other diseases of the respiratory system (J95-99)	2627 (17.03)	2627 (15.49)
Cerebrovascular disease (I60-69)	1869 (12.11)	2044 (12.05)
Ischemic heart disease (I20-25)	1343 (8.70)	1701 (10.03)
Heart disease (I05-52)	1307 (8.47)	1407 (8.30)
Renal failure (N17-19)	1290 (8.36)	1237 (7.29)
Chronic lower respiratory disease (J10-18)	740 (4.80)	651 (3.84)
Metabolic disorders (E70-90)	524 (3.40)	587 (3.46)
Ill-defined and unknown cause of mortality (R95-99)	362 (2.35)	352 (2.08)
General symptoms and signs (R50-69)	356 (2.31)	404 (2.38)
Other diseases[b]	1491 (9.66)	1733 (10.22)

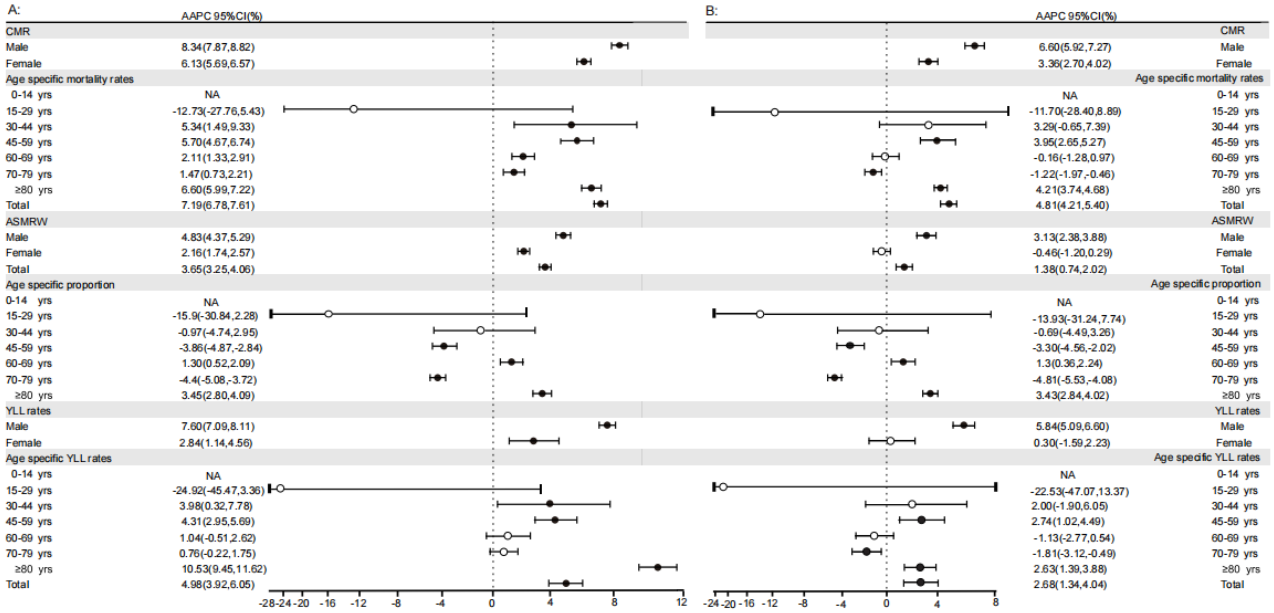
[a] *ICD-10*: *International Classification of Diseases, Tenth revision*.

[b] Other diseases are listed in Multimedia Appendix 1, Table S2.

DM-Specific Premature Death

The total YLL rate increased by 4.98% (95% CI 3.92%-6.05%; *P*<.001) per year and 2.68% (95% CI 1.34%-4.04%; *P*<.001) per year in terms of deaths associated with and deaths due to DM, respectively. The YLL rates of men and women who died due to DM or causes associated with DM increased (all *P*<.001), except for women (*P*=.74) who died due to DM. However, the YLL rates in people that died of causes associated with DM were inconsistent with increasing age; they increased by 3.98% (95% CI 0.32%-7.78%; *P*=.04), 4.31% (95% CI 2.95%-5.69%; *P*<.001), and 10.53% (95% CI 9.45%-11.62%; *P*<.001) per year in the people aged 30 to 44 years, 45 to 59 years, and ≥80 years, respectively. Regarding deaths due to DM, the YLL rate increased by 2.74% (95% CI 1.02%-4.49%; *P*=.004) and 2.63% (95% CI 1.39%-3.88%; *P*<.001) per year in people aged 45 to 59 years and ≥80 years, respectively, while they decreased by 1.81% (95% CI –3.12% to –0.49%; *P*=.01) per year in people aged 70 to 79 years. The YLL rates in the remaining age groups were stable (Figure 2).

Figure 2. Trends in CMR, ASMRW, age-specific proportion of mortality, and YLL among people who died of causes associated with and due to DM according to sex and age group in Pudong, Shanghai, China, from 2005 to 2020. (A) Deaths associated with DM; (B) deaths due to DM. AAPC: average annual percent change; ASMRW: age-standardized mortality rate worldwide; CMR: crude mortality rate; DM: diabetes mellitus; N/A: not applicable; YLL: years of life lost.



Trends in the DM Burden

The long-term trend for total CMR among people who died of causes associated with DM or due to DM increased by 7.19% and 4.81%, respectively. In terms of deaths associated with DM, the CMR also differed by sex (all *P*<.05) and by age over 30 years (all *P*<.05). For deaths due to DM, people aged 30 to 44 years (*P*=.09) and 60 to 69 years (*P*=.76) did not show significant differences in the CMR. The CMR of different age groups among people who died of causes associated with and due to DM are presented in Multimedia Appendix 1, Table S1. The tendencies of modeled CMR, ASMRW, and YLL rates according to sex and age group are shown in Figure 2. The observed CMR, ASMRW, and YLL rates are shown in Multimedia Appendix 1, Tables S3-S6, and the age-specific proportions of death are shown in Multimedia Appendix 1, Table S7. The ASMRW significantly increased among people who died of causes associated with and due to DM, with increases of 3.65% (male AAPC=4.83%; *P*<.001; female AAPC=2.16%; *P*<.001) and 1.38% (male AAPC=3.13%; *P*<.001; female AAPC=–0.46%; *P*=.21), respectively.

The proportion of deaths associated with DM and deaths due to DM varied across age groups; the proportion increased in people aged 60 to 69 years (1.3%, 95% CI 0.52%-2.09%; *P*=.003; 1.3%, 95% CI 0.36%-2.24%; *P*=.009, respectively) and ≥80 years (3.45%, 95% CI 2.8%-4.09%; *P*<.001; 3.43%, 95% CI 2.84%-4.02%; *P*<.001, respectively). However, for deaths associated with DM, the proportion of deaths decreased by 3.86% (95% CI –4.87% to –2.84%; *P*<.001) and 4.4% (95% CI –5.08% to –3.72%; *P*<.001) in people aged 45 to 59 years and 70 to 79 years, respectively. There were no significant changes in the proportion of deaths among people younger than 45 years (all *P*>.05; Figure 2).

Quantitative Impacts of Demographic and Nondemographic Factors on Mortality Rates

Among people who died of causes associated with DM, the CMR attributable to nondemographic factors increased in men (AAPC=17.24%; *P*<.001), women (AAPC=24.16%; *P*=.002) and the total population (AAPC=17.82%; *P*<.001); the CMR attributable to demographic factors also increased by 42.53%, 41.47%, and 41.9% in these groups, respectively, from 2006 to 2020 (all *P*=.001). In terms of deaths due to DM, the CMR caused by nondemographic factors increased in women (AAPC=22.26%; *P*<.001) and the total population (AAPC=20.38%; *P*=.002); there were no significant changes in men (*P*=.10). In contrast, the CMR caused by demographic factors increased in men (AAPC=36.11%; *P*=.001), women (AAPC=37.5%; *P*<.001), and the total population (AAPC=36.72%; *P*<.001). The trends for increased CMR caused by demographic and nondemographic factors relative to the CMR for DM in 2005 are shown in Multimedia Appendix 1, Figure S2. The increased CMR caused by demographic and nondemographic factors is shown in Multimedia Appendix 1, Table S8.

Discussion

Principal Findings

DM and its complications have made considerable contributions to global mortality and disability burdens, particularly in regions undergoing rapid epidemiological transitions [1]. Several studies have demonstrated a decline in DM mortality in developed regions, such as the United States [18], Canada, United Kingdom [19], and Hong Kong [20]. Over the past 40 years of reform and opening up, Pudong has experienced the fastest modernization in China and serves as a representative area for studying the effects of urbanization. In this study, we used population-level data to provide a comprehensive overview of the secular trends in mortality associated with and due to DM among people in Pudong, Shanghai, from 2005 to 2020. We found that DM mortality persistently increased with age, despite annual fluctuations. The magnitude of the increase in DM mortality was similar for both men and women, with more quickly increasing rates of death associated with DM observed for people aged 30 to 44 years and 45 to 59 years. The top 3 comorbidities contributing to deaths associated with and due to DM were hypertension, cerebrovascular disease, and ischemic heart diseases. In addition, population aging and nondemographic factors also played an important role in the burden of DM.

The total CMR trend of death associated with DM and due to DM increased with age, and the CMR levels were higher among men. Similarly, for both measures, the total ASMRW and YLL rates increased with age, and men had more quickly increasing rates. In addition, these rates were higher among people who died of causes associated with DM compared to those who died due to DM. These results indicate that individuals with concomitant DM, especially men, are more likely to experience higher mortality rates and YLL than people who died due to DM alone.

The mortality trend observed in our study is consistent with a previous study that analyzed 22 prospective cohorts in Asia from 1963 to 2006 and found a consistent increase in DM mortality, with higher all-cause mortality among women [21]. Another study that estimated DM mortality in different provinces also showed that the increase in age-standardized mortality among men from 2005 to 2020 was higher than among their female counterparts [22]. In contrast, our results are inconsistent with studies conducted in developed countries, which reported a decrease in DM mortality. For example, in the United States, all-cause mortality declined by 20% every 10 years, and specific-cause mortality due to DM complications also decreased from 1988 to 1994 and from 2010 to 2015 [18]. Such a decrease in DM mortality was also found in England [23] and Australia [24]. Such discrepancies in findings between our study and those from other developed regions may be attributed to differences in population aging and urbanization, as supported by our analysis of demographic and nondemographic factors. First, Shanghai has experienced rapid growth in the older population during the 21st century. Pudong has been classified as an aging society since 1982 (7.72% of the population was aged >65 years), an aged society since 2008 (14.21% aged >65 years), and a hyper-aged society since 2018 (21.81% aged >65 years; Multimedia Appendix 1, Figure S3) [25]. The accelerating process of population aging is accompanied by increases in the prevalence and mortality of DM with age, as well as increases in the number of more chronic complications, which are associated with adverse prognoses. Kirkman et al [26] showed that more than 25% of the population aged ≥65 years in the United States had DM, and that population aging drives the epidemic of DM. They also systematically analyzed the impact of population aging on people with DM and the recommended measures to address this issue. Second, as urbanization continues to progress (accompanied by economic and cultural development), people in Pudong have changed their dietary and lifestyle habits, resulting in substantial health impacts [27]. These changes have contributed to an increasing prevalence of metabolism-related diseases, such as obesity, metabolic syndrome, DM, hypertension, hyperlipidemia, and heart disease, which are more prevalent among regions undergoing urbanization and transition. Many studies have shown that the prevalence of DM is higher among more affluent and better educated populations, particularly in low-income countries [28,29]. In addition, as the pace of life speeds up, fewer people find time to engage in regular exercise, despite physical activity playing a crucial role in managing diabetes. Numerous studies have emphasized the importance of increased physical activity and fitness in alleviating DM, as well as calorie restriction and weight loss [30,31].

In addition, people aged ≥80 years exhibited the highest mortality and increase in YLL rate, probably due to their increased age. However, we also observed substantial mortality and increased YLL rates in people aged 30 to 44 years and 45 to 59 years that were associated with DM. These two age groups constitute a significant portion of the social workforce, and an increased disease burden in these groups exerts a significant adverse impact on economic development. Several potential explanations may account for these findings. First, patients with DM in these age groups may not have a comprehensive understanding of the severity of DM and its complications, resulting in low medication adherence and limited self-care behaviors. A study conducted in Ghana confirmed that patients aged 70 years and older were more likely to adhere to medication compared to those younger than 50 years [32]. Moreover, the decreasing YLL rate among deaths due to DM in people aged 70 to 79 years further supports this finding. Second, a younger age at diagnosis of type 2 DM is associated with higher mortality [33]. A previous study demonstrated that young DM patients exhibit worse glycemic control and face greater difficulty meeting glycohemoglobin and low-density lipoprotein cholesterol treatment goals [34]. Finally, people in these age ranges often shoulder the responsibility of supporting their households and have demanding work schedules, leaving limited time for complying with healthy lifestyle and medication regimens. As discussed in the previous section, urbanization has contributed to the increased burden of diseases.

In this study, we observed a high prevalence of multimorbidity among patients with DM, particularly in relation to cardiovascular disease, which includes hypertension, cerebrovascular disease, and ischemic heart disease. Hypertension emerged as the most prevalent comorbidity among deaths associated with DM. This finding aligns with the research conducted by Safar et al [35], who reported significantly higher cardiovascular risk in subjects with DM receiving treatment for hypertension, despite adequate glycemic and blood pressure control. Numerous studies have established that DM and cardiovascular disease have similar risk factors and pathogenic mechanisms, including a combination of genetic and metabolic factors, such as insulin resistance, contributing to their prevalence [36]. As mentioned earlier, hyperglycemia plays a pivotal role in the development of cardiovascular and cerebrovascular complications. DM stands as the leading risk factor for cardiovascular mortality due to its prominent involvement in cardiovascular pathogenesis [37]. Moreover, previous studies have shown that increased cancer incidence [38] and mortality are associated with DM [39]. Consistent with these findings, our study found a higher prevalence of various cancers among people who died of causes associated with DM. These cancers included liver cancer, colorectal cancer, malignant neoplasm of the pancreas, and stomach cancer. Patients with DM may be at increased risk of cancer due to shared risk factors, such as age, obesity, physical inactivity, and smoking. Therefore, the presence of multimorbidity of DM underscores the significant disease burden and the necessity of the management of glycemia and its associated complications, particularly cardiovascular disease and malignant tumors. Regular testing and follow-ups at community health centers should be prioritized to enhance overall care.

With the significant burden of DM, health policies play a crucial role in promoting education and support for the management of chronic disease. In 2005, Shanghai first piloted a community-based diabetes management policy and offered free health examinations for the elderly. Although we have not yet observed an inflection point in the mortality rate, we noted a stabilization of CMR and ASMRW among deaths due to DM in the past 5 years, following the revision and improvement of the policy (Multimedia Appendix 1, Figure S4). Furthermore, the Chinese government approved the Healthy China 2030 Strategy in 2016. In accordance with the national plan, Shanghai implemented its own plan to achieve the objectives in the Health Shanghai 2030 Planning Outline, which called for comprehensive interventions for health issues, lifelong health care, prevention and control of major diseases, and improvement of the health service system. Health management was emphasized in this plan and actively encouraged in community health centers and health management centers in district and municipal hospitals [40]. With improved individual socioeconomic status and health awareness, an increasing number of people are willing to undergo regular medical check-ups that screen for chronic diseases and malignant tumors, as well as undergo regular tests and follow-ups at community health centers, particularly for primary prevention of chronic diseases. With continual implementation of these policies, we anticipate an inflection point in DM mortality in the near future.

To our knowledge, this study is the first to analyze the disease burden of deaths associated with and due to DM in mainland China based on a large population (3.17 million people) and over a relatively long time frame (15 years.) The study area is representative and provides insights into the changing disease spectrum in transitioning regions. However, this study still has some limitations. First, a comprehensive life cycle analysis of the disease burden is currently unavailable, as we are still in the process of determining all causes of death, extrapolating relevant deaths during quality assessments, and refining the data. Second, although Shanghai is a transitional region in China, it is possible that its demographic and economic characteristics differ from those of other developing countries. Economic development will result in similar changes in lifestyle habits, making the data valuable as a reference. Third, our data lacked some details on DM, such as the clinical phenotype, duration, severity, treatment efficacy, and use of community health centers. Furthermore, other relevant factors, such as unhealthy lifestyle, low personal income, and adverse family circumstances, were also not captured in the data. Nonetheless, it is important to note that our study used complete and accurate public data from the government surveillance system, ensuring a high level of data quality.

Conclusion

Population aging and multimorbidity have played essential roles in changing the burden of DM in an urbanizing and transitioning region. Young and middle-aged people have an increasing disease burden and should be given more attention. DM, an important disease associated with aging, remains an understudied and unaddressed medical threat in an aging society. Health management as an emerging method has important implications for alleviating the future burden of DM.

Acknowledgments

We thank all the staff in the Vital Statistics System of Pudong New Area from 2005 to 2020 for their great work in data collection and assuring data of high quality. This study was funded by grants from the Shanghai Municipal Program for Medical Leading Talents (2019LJ15; to author SJ), the National Key Research and Development Program (2022YFC3600901), and the Shanghai Public Health System Construction Three-year Action Plan Outstanding Youth Talent Training Program (GWV-10.2-YQ43; to author YC).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

XL and RL drafted and revised the manuscript. XL, YC, and YX participated in the collection, analysis, and interpretation of data. YX, JZ, XL, and SJ contributed to data collection and made suggestions during the analysis. JZ and SJ conceived the study, participated in its design, and coordinated production and critical revision of the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material and data on deaths associated with and due to diabetes mellitus from 2005 to 2020 in Pudong, Shanghai, China.
[PDF File (Adobe PDF File), 735 KB - publichealth_v9i1e43687_app1.pdf]

References

1. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol 2018 Feb;14(2):88-98. [doi: 10.1038/nrendo.2017.151] [Medline: 29219149]
2. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
3. Branca Vergano L, Monesi M, Vicenti G, Bizzoca D, Solarino G, Moretti B. Posterior approaches in malleolar fracture: when, why and how. J Biol Regul Homeost Agents 2020;34(3 Suppl. 2):89-95. ADVANCES IN MUSCULOSKELETAL DISEASES AND INFECTIONS. [Medline: 32856446]
4. Morley JE, Abbatecola AM, Woo J. Management of comorbidities in older persons with type 2 diabetes. J Am Med Dir Assoc 2017 Aug 01;18(8):639-645. [doi: 10.1016/j.jamda.2017.05.006] [Medline: 28625388]
5. Cheng S, Wang C, Ko Y. Costs and length of stay of hospitalizations due to diabetes-related complications. J Diabetes Res 2019 Sep 08;2019:2363292-2363296 [FREE Full text] [doi: 10.1155/2019/2363292] [Medline: 31583247]
6. Coles B, Zaccardi F, Hvid C, Davies MJ, Khunti K. Cardiovascular events and mortality in people with type 2 diabetes and multimorbidity: A real-world study of patients followed for up to 19 years. Diabetes Obes Metab 2021 Jan;23(1):218-227. [doi: 10.1111/dom.14218] [Medline: 33026165]
7. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 years of health-care reform in China: progress and gaps in universal health coverage. Lancet 2019 Sep 28;394(10204):1192-1204. [doi: 10.1016/S0140-6736(19)32136-1] [Medline: 31571602]
8. Gassasse Z, Smith D, Finer S, Gallo V. Association between urbanisation and type 2 diabetes: an ecological study. BMJ Glob Health 2017;2(4):e000473 [FREE Full text] [doi: 10.1136/bmjgh-2017-000473] [Medline: 29104770]
9. Zhu D, Ozaki A, Virani SS. Disease-specific excess mortality during the COVID-19 pandemic: an analysis of weekly US death data for 2020. Am J Public Health 2021 Aug;111(8):1518-1522. [doi: 10.2105/AJPH.2021.306315] [Medline: 34185569]
10. Li X, Qian M, Zhao G, Yang C, Bao P, Chen Y, et al. The performance of a community-based colorectal cancer screening program: Evidence from Shanghai Pudong New Area, China. Prev Med 2019 Jan;118:243-250. [doi: 10.1016/j.ypmed.2018.11.002] [Medline: 30412744]
11. Shanghai Pudong Statistical Yearbook 2021. People's Government of Pudong New Area, Shanghai. URL: https://www.pudong.gov.cn/008006035003/20220630/703404.html [accessed 2023-08-05]
12. Chen H, Hao L, Yang C, Yan B, Sun Q, Sun L, et al. Understanding the rapid increase in life expectancy in Shanghai, China: a population-based retrospective analysis. BMC Public Health 2018 Feb 14;18(1):256 [FREE Full text] [doi: 10.1186/s12889-018-5112-7] [Medline: 29444657]
13. Ye W, Lu W, Li X, Chen Y, Wang L, Zeng G, et al. Long-term changes in the premature death rate in lung cancer in a developed region of China: Population-based study. JMIR Public Health Surveill 2022 Apr 20;8(4):e33633-e33633. [doi: 10.2196/33633] [Medline: 35442209]
14. GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1160-1203 [FREE Full text] [doi: 10.1016/S0140-6736(20)30977-6] [Medline: 33069325]
15. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Jan 10;385(9963):117-171 [FREE Full text] [doi: 10.1016/S0140-6736(14)61682-2] [Medline: 25530442]
16. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
17. Li X, Zhang M, Chen Y, Lv H, Du Y. Longitudinal analysis of ovarian cancer death patterns during a rapid transition period (2005-2020) in Shanghai, China: A population-based study. Front Oncol 2022 Sep 03;12(1):1003297. [doi: 10.3389/fonc.2022.1003297] [Medline: 36263229]
18. Gregg EW, Cheng YJ, Srinivasan M, Lin J, Geiss LS, Albright AL, et al. Trends in cause-specific mortality among adults with and without diagnosed diabetes in the USA: an epidemiological analysis of linked national survey and vital statistics

data. Lancet 2018 Jun 16;391(10138):2430-2440 [FREE Full text] [doi: 10.1016/S0140-6736(18)30314-3] [Medline: 29784146]

19. Lind M, Garcia-Rodriguez LA, Booth GL, Cea-Soriano L, Shah BR, Ekeroth G, et al. Mortality trends in patients with and without diabetes in Ontario, Canada and the UK from 1996 to 2009: a population-based study. Diabetologia 2013 Dec;56(12):2601-2608. [doi: 10.1007/s00125-013-3063-1] [Medline: 24114114]
20. Wu H, Lau ESH, Ma RCW, Kong APS, Wild SH, Goggins W, et al. Secular trends in all-cause and cause-specific mortality rates in people with diabetes in Hong Kong, 2001-2016: a retrospective cohort study. Diabetologia 2020 Apr;63(4):757-766. [doi: 10.1007/s00125-019-05074-7] [Medline: 31942668]
21. Yang JJ, Yu D, Wen W, Saito E, Rahman S, Shu X, et al. Association of diabetes with all-cause and cause-specific mortality in Asia: A pooled analysis of more than 1 million participants. JAMA Netw Open 2019 Apr 05;2(4):e192696 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.2696] [Medline: 31002328]
22. Wang F, Wang W, Yin P, Liu Y, Liu J, Wang L, et al. Mortality and years of life lost in diabetes mellitus and its subcategories in China and its provinces, 2005-2020. J Diabetes Res 2022;2022:1609267 [FREE Full text] [doi: 10.1155/2022/1609267] [Medline: 35493611]
23. Pearson-Stuttard J, Bennett J, Cheng YJ, Vamos EP, Cross AJ, Ezzati M, et al. Trends in predominant causes of death in individuals with and without diabetes in England from 2001 to 2018: an epidemiological analysis of linked primary care records. Lancet Diabetes Endocrinol 2021 Mar;9(3):165-173 [FREE Full text] [doi: 10.1016/S2213-8587(20)30431-9] [Medline: 33549162]
24. Harding J, Shaw JE, Peeters A, Davidson S, Magliano DJ. Age-specific trends from 2000-2011 in all-cause and cause-specific mortality in type 1 and type 2 diabetes: A cohort study of more than one million people. Diabetes Care 2016 Jun;39(6):1018-1026. [doi: 10.2337/dc15-2308] [Medline: 27208325]
25. Luo Y, Su B, Zheng X. Trends and challenges for population and health during population aging - China, 2015-2050. China CDC Wkly 2021 Jul 09;3(28):593-598 [FREE Full text] [doi: 10.46234/ccdcw2021.158] [Medline: 34594944]
26. Kirkman M, Briscoe VJ, Clark N, Florez H, Haas LB, Halter JB, et al. Diabetes in older adults. Diabetes Care 2012 Dec;35(12):2650-2664 [FREE Full text] [doi: 10.2337/dc12-1801] [Medline: 23100048]
27. Popkin BM. Synthesis and implications: China's nutrition transition in the context of changes across other low- and middle-income countries. Obes Rev 2014 Jan;15 Suppl 1(0 1):60-67 [FREE Full text] [doi: 10.1111/obr.12120] [Medline: 24341759]
28. Kehlenbrink S, Smith J, Ansbro, Fuhr DC, Cheung A, Ratnayake R, et al. The burden of diabetes and use of diabetes care in humanitarian crises in low-income and middle-income countries. Lancet Diabetes Endocrinol 2019 Aug;7(8):638-647. [doi: 10.1016/S2213-8587(19)30082-8] [Medline: 30878268]
29. Popkin B, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev 2012 Jan;70(1):3-21 [FREE Full text] [doi: 10.1111/j.1753-4887.2011.00456.x] [Medline: 22221213]
30. Magkos F, Hjorth MF, Astrup A. Diet and exercise in the prevention and treatment of type 2 diabetes mellitus. Nat Rev Endocrinol 2020 Oct;16(10):545-555. [doi: 10.1038/s41574-020-0381-5] [Medline: 32690918]
31. Vranic M, Berger M. Exercise and diabetes mellitus. Diabetes 1979 Feb;28(2):147-163. [doi: 10.2337/diab.28.2.147] [Medline: 369929]
32. Afaya RA, Bam V, Azongo TB, Afaya A, Kusi-Amponsah A, Ajusiyine JM, et al. Medication adherence and self-care behaviours among patients with type 2 diabetes mellitus in Ghana. PLoS One 2020;15(8):e0237710 [FREE Full text] [doi: 10.1371/journal.pone.0237710] [Medline: 32822381]
33. Magliano DJ, Sacre JW, Harding JL, Gregg EW, Zimmet PZ, Shaw JE. Young-onset type 2 diabetes mellitus - implications for morbidity and mortality. Nat Rev Endocrinol 2020 Jun;16(6):321-331. [doi: 10.1038/s41574-020-0334-z] [Medline: 32203408]
34. Yeung RO, Zhang Y, Luk A, Yang W, Sobrepena L, Yoon K, et al. Metabolic profiles and treatment gaps in young-onset type 2 diabetes in Asia (the JADE programme): a cross-sectional study of a prospective cohort. Lancet Diabetes Endocrinol 2014 Dec;2(12):935-943. [doi: 10.1016/S2213-8587(14)70137-8] [Medline: 25081582]
35. Safar ME, Gnakaméné JB, Bahous SA, Yannoutsos A, Thomas F. Longitudinal study of hypertensive subjects with type 2 diabetes mellitus: Overall and cardiovascular risk. Hypertension 2017 Jun;69(6):1029-1035. [doi: 10.1161/HYPERTENSIONAHA.116.08962] [Medline: 28396537]
36. Jia G, Whaley-Connell A, Sowers JR. Diabetic cardiomyopathy: a hyperglycaemia- and insulin-resistance-induced heart disease. Diabetologia 2018 Jan;61(1):21-28 [FREE Full text] [doi: 10.1007/s00125-017-4390-4] [Medline: 28776083]
37. Cole JB, Florez JC. Genetics of diabetes mellitus and diabetes complications. Nat Rev Nephrol 2020 Jul;16(7):377-390 [FREE Full text] [doi: 10.1038/s41581-020-0278-5] [Medline: 32398868]
38. Suh S, Kim K. Diabetes and cancer: Cancer should be screened in routine diabetes assessment. Diabetes Metab J 2019 Dec;43(6):733-743 [FREE Full text] [doi: 10.4093/dmj.2019.0177] [Medline: 31902143]
39. Barone BB, Yeh HC, Snyder CF, Peairs KS, Stein KB, Derr RL, et al. Long-term all-cause mortality in cancer patients with preexisting diabetes mellitus: a systematic review and meta-analysis. JAMA 2008 Dec 17;300(23):2754-2764 [FREE Full text] [doi: 10.1001/jama.2008.824] [Medline: 19088353]

40. Schmittdiel JA, Gopalan A, Lin MW, Banerjee S, Chau CV, Adams AS. Population health management for diabetes: Health care system-level approaches for improving quality and addressing disparities. Curr Diab Rep 2017 May;17(5):31 [FREE Full text] [doi: 10.1007/s11892-017-0858-3] [Medline: 28364355]

Abbreviations

AAPC: average annual percent change
APC: annual percent change
ASMRW: age-standardized mortality rate worldwide
CDC: Center for Disease Control and Prevention
CMR: crude mortality rate
DM: diabetes mellitus
***ICD-10*:** *International Classification of Diseases Tenth version*
WHO: World Health Organization
YLL: years of life lost

Edited by A Mavragani, T Sanchez; submitted 20.10.22; peer-reviewed by F Lami, N Trehan; comments to author 01.06.23; revised version received 20.06.23; accepted 25.07.23; published 04.09.23.

Please cite as:
Li X, Liu R, Chen Y, Han Y, Wang Q, Xu Y, Zhou J, Jiang S
Patterns and Trends in Mortality Associated With and Due to Diabetes Mellitus in a Transitioning Region With 3.17 Million People: Observational Study
JMIR Public Health Surveill 2023;9:e43687
URL: https://publichealth.jmir.org/2023/1/e43687
doi:10.2196/43687
PMID:37665630

Original Paper

Association Between Metabolic Syndrome and Mortality: Prospective Cohort Study

Wenzhen Li[1*], MD, PhD; Dajie Chen[2], PhD; Ying Peng[3], MS; Zuxun Lu[4], MD; Mei-Po Kwan[5,6,7], PhD; Lap Ah Tse[1,6,8*], MD, PhD

[1]Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
[2]Department of Health Service and Management, Wuhan Polytechnic University, Wuhan, China
[3]Department of Communicable Diseases Control and Prevention, Wuhan Center for Disease Control and Prevention, Wuhan, China
[4]Department of Social Medicine and Health Management, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[5]Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
[6]Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
[7]Institute of Future Cities, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
[8]Shenzhen Municipal Key Laboratory for health Risk Analysis, Shenzhen Research Institute of the Chinese University of Hong Kong, Shenzhen, China
[*]these authors contributed equally

Corresponding Author:
Lap Ah Tse, MD, PhD
Jockey Club School of Public Health and Primary Care
The Chinese University of Hong Kong
Hong Kong
China (Hong Kong)
Phone: 852 22528791
Fax: 852 2606 3500
Email: shelly@cuhk.edu.hk

Abstract

Background: Metabolic syndrome (MetS) is a common metabolic disorder that results from the increasing prevalence of obesity, which has been an increasing concern in recent years. Previous evidence indicated that MetS was associated with mortality; however, different definitions of MetS were used. In 2005, the National Cholesterol Education Program (NCEP) Adult Treatment Panel (ATP) III updated the definition of MetS, which has since been widely adopted. Therefore, it is necessary to conduct a novel study among other populations and countries with a larger sample size using the updated definition of MetS and death code to examine the association of MetS with all-cause and cause-specific mortality.

Objective: We aimed to examine the associations of MetS with all-cause and cause-specific mortality.

Methods: A total of 36,414 adults were included in this study, using data from the National Health and Nutrition Examination Survey (NHANES) III (1988-1994) and the continuous NHANES (1999-2014) in the United States. Death outcomes were ascertained by linkage to National Death Index records through December 31, 2015. MetS was defined by the NCEP ATP III-2005 criterion. Complex survey design factors including sample weights, clustering, and stratification were considered for all analyses with instructions for using NHANES data. Cox proportional hazards models were used to estimate hazard ratios (HRs) and 95% CIs for mortality from all causes, heart disease, diabetes, and cancer.

Results: We observed 8494 deaths during the 16.71 years of follow-up. Compared with those without MetS, individuals with MetS were associated with a significantly elevated multiadjusted HR of 1.24 (95% CI 1.16-1.33), 1.44 (95% CI 1.25-1.66), and 5.15 (95% CI 3.15-8.43) for all cause, heart diseases, and diabetes mellitus, respectively, whereas no significant association was found for cancer mortality (HR 1.17, 95% CI 0.95-1.43).

Conclusions: Our study provides additional evidence that MetS and its components are significantly associated with all-cause, heart disease, and diabetes mortality, but not with cancer mortality. Health care professionals should pay more attention to MetS and its individual component.

(JMIR Public Health Surveill 2023;9:e44073) doi:10.2196/44073

KEYWORDS

metabolic syndrome; mortality; heart disease; diabetes mellitus; cancer

Introduction

Metabolic syndrome (MetS) is a common metabolic disorder that results from the increasing prevalence of obesity [1], which has been an increasing concern in the past few years [2]. A large number of studies have been conducted to explore the definition, prevalence, and associated factors of MetS [3-5], as well as to examine the relationship between MetS and cardiovascular disease (CVD) [6], diabetes mellitus (DM) [7], and other disease such as skin symptoms [8,9]. MetS is highly prevalent in many countries such as the United States [10], China [2], and India [11], despite the definitions of MetS being varied. MetS is a well-known, health-related problem.

Previous evidence indicated that MetS was associated with mortality; however, different definitions of MetS were used. One study [12] was conducted from 1984 to 1989 among men aged 42-60 years using the World Health Organization (WHO) and modified National Cholesterol Education Program (NCEP) definitions for MetS; it consistently showed that CVD and all-cause mortality increased in 1209 men with MetS. Another study [13] indicated that the WHO definition of MetS identified adults with increased CVD morbidity and mortality, but the study was conducted among middle-aged adults about 30 years ago in Europe with a small sample size. A large-sample study [14] conducted in 1978 and 1987 examined the relationship between syndrome X and mortality; however, syndrome X was the initial concept of MetS, which was different from MetS. Besides, similar to the 2 studies above, the deaths were also coded according to the International Classification of Diseases (ICD), Ninth Revision. It is well known that MetS is becoming prevalent among young adults due to great changes in lifestyle, diet, and socioeconomic environment, and the definition of MetS has been updated or modified several times and changed drastically. In 2005, the Third Report of the US NCEP Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel [ATP] III) updated again the definition of MetS according to the modified American Diabetes Association criteria for impaired fasting glucose, which has since been widely adopted in the United States and elsewhere due to its simplicity to use in a clinical setting and its advantage of avoiding emphasis on a single cause [15].

Therefore, it is necessary to conduct a novel study among other population and countries with a large sample size using the updated and widely used definition of MetS and death code to examine the association of MetS with all-cause and cause-specific mortality.

Methods

Study Design and Participants

The study population were obtained from the National Health and Nutrition Examination Survey (NHANES) III (1988-1994) and the continuous NHANES (1999-2014), and data were obtained by questionnaire and interview, mobile physical examination, and laboratory tests with a complex, multistage, and probability sampling method. Details of NHANES have been described on the web [16]. In this analysis, 33,994 participants from the NHANES III (1988-1994) and 82,091 from the continuous NHANES (1999-2014) data sets were first enrolled. After excluding those without high-density lipoprotein cholesterol (HDL-C), blood pressure, plasma fasting glucose, triglycerides (TGs), and waist circumference data, 15,530 participants from the NHANES III and 25,371 from the continuous NHANES remained. Furthermore, we excluded participants aged <18 years and those without mortality data. Finally, a total of 36,414 participants were retained in our cohort for analysis. The flowchart of the study is presented in Figure S1 in Multimedia Appendix 1.

Ascertainment of MetS

MetS was defined by the NCEP ATP III-2005 criterion [15], that is, a person who has 3 or more of the following criteria: (1) elevated waist circumference (EWC): waist circumference ≥102 cm in men and ≥88 cm in women; (2) elevated blood pressure: blood pressure ≥130/85 mm Hg or drug treatment of previously diagnosed hypertension; (3) reduced HDL-C: <40 mg/dL in men and <50 mg/dL in women or specific treatment for reduced HDL-C; (4) elevated TGs: TG level ≥150 mg/dL or drug treatment for elevated TG; and (5) elevated fasting glucose: fasting glucose level of ≥100 mg/L or drug treatment for elevated glucose and previously diagnosed type 2 diabetes.

Ascertainment of Death

Mortality status was ascertained by probabilistic matching to the National Death Index through December 31, 2015, using a unique study identifier. Details of the matching method are available from the National Center for Health Statistics [17]. Causes of deaths were classified according to the codes of ICD-10. Primary outcomes in this study were mortality from all causes, heart diseases (codes I00-I09, I11, I13, and I20-I51), cancer (codes C00-C97), and diabetes (codes E10-E14).

Ascertainment of Covariates

Gender, age, race and ethnicity (Mexican American, non-Hispanic Black, non-Hispanic White, or others), educational level (less than high school, high school or equivalent, or college or above), marital status (married; separated, including widowed and divorced groups; or never married), smoking, physical activity, and BMI were obtained by interviews and physical examinations. BMI was calculated as weight (kg) divided by height squared (m^2) and was categorized into 3 groups: <25, 25-30, and ≥30 kg/m^2 [18]. Family poverty-to-income ratio (PIR) level was grouped into 3 categories: 0-1.0, 1.1-3.0, and >3.0 [19]. Current smokers were defined as those who smoked at least 100 cigarettes in their lifetime and smoked at the time of survey. Baseline moderate-to-vigorous leisure physical activity level was defined as active physical activity such as brisk walking, carrying or lifting light or heavy loads, running, basketball, bicycling, swimming, or volleyball for at least 10 minutes continuously. Self-reported general health condition was classified 3 groups: very good to excellent, good, and poor to fair. Multiple imputation was used for missing values [20].

Statistical Analysis

Complex survey design factors including sample weights, clustering, and stratification were considered for all analyses with instructions for using NHANES data. We compared baseline characteristics by MetS in the 2 intervals by using the Rao-Scott chi-square test for categorical variables and ANOVA and the Kruskal-Wallis test adjusted for sampling weights for continuous variables. The Kaplan-Meier method was used to plot the survival curves associated with MetS and the number of MetS components. Cox proportional hazards models were used to estimate the hazard ratios (HRs) with 95% CIs for all-cause and cause-specific morality for MetS and its components.

The baseline age (years, continuous), gender, and race and ethnicity (Mexican American, non-Hispanic Black, non-Hispanic White, or others) were adjusted in model 2. Furthermore, education level (less than high school, high school or equivalent, or college or above), family PIR level (0-1.0, 1.1-3.0, or >3.0), marital status (married, separated, or never married), BMI (<25, 25-30, or ≥30 kg/m^2), smoking status (yes or no), active physical activity level (yes or no), self-reported health status (very good to excellent, good, or poor to fair) were adjusted in model 3. For analyses of one of the components of MetS and mortality, we further adjusted the other 4 components in model 4. We also conducted stratified analyses according to each covariate. We conducted several sensitivity analyses to test the robustness of the results. First, we removed participants with missing values for covariates and ran complete case analyses. Second, we excluded individuals with prevalent CVD, cancer, or DM to minimize potential reverse causation due to severe illness. Third, we excluded individuals who had a follow-up time of less than 3 years (including those who died within 3 years of follow-up).

All statistical analyses were performed using SAS (version 9.4; SAS Institute Inc) and R (version 4.0.5; R Foundation for Statistical Computing), with 2-sided *P*<.05 considered statistically significant.

Ethical Considerations

This study proposal was approved by the Ethics and Human Subject Committee of Tongji Medical College, Huazhong University of Science and Technology (2019 IEC (S342)).

Written informed consent was obtained from all study participants in the NHANES, and the NHANES was approved by the Ethics Review Board of the National Centers for Health Statistics [21]. The NHANES data are free for public use and available on the web [22].

Results

Baseline characteristic of the population are shown in Table 1. The median follow-up for survivors was 16.71 (95% CI 15.17-18.25) years. Among the 36,414 individuals, the weighted mean age was 48.07 (range 18-90) years: 57.22 (SD 16.637) years for those with MetS and 42.18 (SD 18.496) years for those without MetS. A total of 18,887 (51.87%) were women, 14,109 (38.75%) were younger adults aged <40 years, and 17,079 (46.9%) had good or excellent general health. Among all the participants, 39.17% (n=14,265) had MetS. Significant differences were observed in all groups except for the group of gender (all *P*<.05).

There were 8494 deaths during the follow-up: 1428 (16.81%) deaths from heart diseases, 1280 (15.07%) deaths from cancer, and 220 (2.59%) deaths from DM. Table 2 demonstrates the associations of MetS and its components with all-cause and cause-specific mortality. Furthermore, we explored the relationships between the number of MetS components and mortality; similar findings were observed, and significant positive relationships were found between the different number of MetS components and all-cause, heart disease, and DM mortality, not for cancer mortality. The hazard of mortality increases with an increase in the number of MetS components. Besides, except for elevated TG, the other 4 components were associated with an increased hazard of all-cause mortality after adjusting for all covariates. Elevated blood pressure and EWC were not associated with heart disease mortality, whereas elevated TG and EWC were not associated with DM mortality. Interestingly, reduced HDL-C was associated with the hazard of cancer mortality (HR 1.36, 95% CI 1.12-1.64). Furthermore, the results of the sensitivity analyses did not change substantially (Tables S1-3 in Multimedia Appendix 1). In addition, Figure 1 shows that the cumulative hazard gradually increased with MetS and the number of MetS components during the follow-up time (all *P*<.001), and cumulative hazard was higher among adults with MetS and an increased number of MetS components (*P*<.001).

The subgroup analyses of MetS with all-cause, heart disease, and DM mortality are showed in Table 3. In the stratified analyses, significant associations between MetS and all-cause mortality were found among most groups, especially for groups of marital status and education levels. The relationships only existed for the separated marital status and did not exist among those with high school or equivalent educational levels. Similar findings were observed in the associations of MetS with heart disease and DM mortality. In addition, the positive relationship between MetS and heart disease mortality was found among non-Hispanic White individuals (HR 1.21 95% CI 1.11-1.32), those with PIR >3.0 (HR 1.82; 95% CI 1.46-2.28), and those with active physical activity (HR 1.53, 95% CI 1.30-1.81). Besides, individuals aged <40 years and those with PIR from 0 to 1.0 were not associated with elevated hazard of DM mortality. Men with MetS had higher hazard of heart disease mortality (1.54, 95% CI 1.23-1.93 for men; 1.38, 95% CI 1.16-1.65 for women), whereas women with MetS had higher hazard of DM mortality (8.75, 95% CI 4.20-18.21 for women; 3.60, 95% CI 1.63-7.98 for men).

Table 1. Baseline characteristics of participants from the National Health and Nutrition Examination Survey (NHANES) according to metabolic syndrome (MetS)[a].

Characteristic	All participants (n=36,414)	MetS (n=14,265, 39.17%)	Non-MetS (n=22,149, 60.83%)	Rao-Scott χ^2 (*df*)	*P* value
Gender, n (%)				5.55 (1)	.02
Man	17,527 (48.13)	6746 (47.29)	10,781 (48.67)		
Woman	18,887 (51.87)	7519 (52.17)	11,368 (51.33)		
Age (years), mean (SD)	48.07 (19.25)	57.22 (16.64)	42.18 (18.50)	1798.875 (2)	<.001
<40, n (%)	14,109 (38.75)	2440 (17.1)	11,669 (52.68)		
40-60, n (%)	10,409 (28.59)	4542 (31.84)	5867 (26.49)		
≥60, n (%)	11,896 (32.67)	7283 (51.06)	4613 (20.83)		
Race and ethnicity, n (%)				12.94 (3)	.005
Mexican American	8279 (22.74)	3259 (22.85)	5020 (22.66)		
Non-Hispanic Black	8446 (23.19)	2905 (20.36)	5541 (25.02)		
Non-Hispanic White	16,036 (44.04)	6734 (47.21)	9302 (42)		
Others	3653 (10.03)	1367 (9.58)	2286 (10.32)		
Educational levels, n (%)				64.47 (2)	<.001
Less than high school	17,243 (47.35)	7208 (50.53)	10,035 (45.31)		
High school or equivalent	7053 (19.45)	2778 (19.47)	4305 (19.44)		
College or above	12,088 (33.2)	4279 (30)	7809 (35.26)		
Married status, n (%)				314.18 (2)	<.001
Married	11,193 (30.74)	5086 (35.65)	6107 (27.57)		
Separated	16,809 (46.16)	7345 (51.49)	9464 (42.73)		
Never married	8412 (23.1)	1834 (12.86)	6578 (29.7)		
Family poverty-to-income ratio level, n (%)				32.39 (2)	<.001
0-1.0	7420 (20.38)	2934 (20.57)	4486 (20.25)		
1.1-3.0	15,846 (43.52)	6435 (45.11)	9411 (42.49)		
>3.0	13,148 (36.11)	4896 (34.32)	8252 (37.26)		
BMI (kg/m^2), n (%)				4642.97 (2)	<.001
<25	12,750 (35.01)	1697 (11.9)	11,053 (49.9)		
25-30	12,441 (34.17)	5072 (35.56)	7369 (33.27)		
≥30	11,223 (30.82)	7496 (52.55)	3727 (16.83)		
Smoking status, n (%)				22.59 (1)	<.001
Yes	17,606 (48.35)	7353 (51.55)	10,253 (46.29)		
No	18,808 (51.65)	6912 (48.45)	11,896 (53.71)		
Active physical activity, n (%)				190.66 (1)	<.001
Yes	15,726 (43.19)	4858 (34.06)	10,868 (49.07)		
No	20,688 (56.81)	9407 (65.94)	11,281 (50.93)		
General health condition, n (%)				938.15 (2)	<.001
Very good to excellent	17,079 (46.9)	4523 (31.71)	12,556 (56.69)		
Good	13,292 (36.5)	6052 (42.43)	7240 (32.69)		
Poor to fair	6043 (16.60)	3690 (25.87)	2353 (10.62)		
Follow-up time (years), median (95% CI)	16.71 (15.17-18.25)	12.48 (11.93-13.03)	21.63 (21.30-21.97)	990.1294[b]	<.001
Self-reported chronic diseases (including CVD[c], cancer, and DM[d]), n (%)				2083.75 (1)	<.001

Characteristic	All participants (n=36,414)	MetS (n=14,265, 39.17%)	Non-MetS (n=22,149, 60.83%)	Rao-Scott χ^2 (*df*)	*P* value
Yes	8324 (22.86)	5677 (39.8)	2647 (11.95)		
No	28,090 (77.14)	8588 (60.2)	19,502 (88.05)		

[a] All estimates accounted for complex survey designs, and *P* values were calculated using ANOVA adjusting for sampling weights and Rao-Scott chi-square test for continuous and categorical variables.

[b] Kruskal-Wallis test was used.

[c] CVD: cardiovascular disease.

[d] DM: diabetes mellitus.

Table 2. Associations of metabolic syndrome (MetS) and its components and all-cause and specific cause mortality.

Association	Hazard ratio (95% CI)			
	All causes (n=8494)	Heart disease (n=1428)	Cancer (n=1280)	DM[a] (n=220)
MetS				
Model 1[b]	2.90 (2.66-3.16)	3.56 (3.16-4.01)	2.35 (1.97-2.80)	12.19 (7.48-19.84)
Model 2[c]	1.30 (1.21-1.40)	1.49 (1.33-1.67)	1.49 (0.98-1.43)	6.34 (3.70-10.87)
Model 3[d]	1.24 (1.16-1.33)	1.44 (1.25-1.66)	1.17 (0.95-1.43)	5.15 (3.15-8.43)
Number of MetS[e] (reference=0)				
1	1.29 (1.12-1.48)	1.31 (0.82-2.09)	1.04 (0.75-1.43)	11.65 (1.99-68.17)
2	1.41 (1.21-1.63)	1.70 (1.10-2.64)	1.11 (0.80-1.53)	10.76 (2.08-55.77)
3	1.51 (1.28-1.78)	1.82 (1.27-2.61)	1.26 (0.88-1.81)	29.53 (5.63-154.93)
4	1.70 (1.48-1.96)	2.51 (1.63-3.88)	1.20 (0.86-1.68)	56.62 (10.20-314.33)
5	2.01 (1.68-2.40)	2.77 (1.68-4.56)	1.32 (0.93-1.89)	138.08 (24.35-783.13)
Components of MetS[e]				
EBP[f]	1.15 (1.07-1.24)	1.15 (0.96-1.38)	0.91 (0.77-1.08)	2.23 (1.29-3.85)
EGLU[g]	1.18 (1.10-1.27)	1.38 (1.17-1.63)	1.03 (0.87-1.21)	6.66 (3.49-12.71)
ETG[h]	1.14 (1.09-1.21)	1.43 (1.24-1.65)	1.05 (0.89-1.25)	1.76 (1.01-3.08)
Reduced HDL-C[i]	1.24 (1.16-1.32)	1.39 (1.17-1.64)	1.35 (1.13-1.60)	3.05 (2.00-4.64)
EWC[j]	1.17 (1.08-1.27)	1.12 (0.87-1.43)	1.09 (0.89-1.34)	1.89 (1.17-3.06)
Components of MetS[k]				
EBP	1.12 (1.04-1.21)	1.09 (0.91-1.30)	0.91 (0.77-1.08)	1.79 (1.04-3.10)
EGLU	1.14 (1.06-1.23)	1.32 (1.11-1.57)	1.01 (0.85-1.20)	5.77 (3.00-11.10)
ETG	1.05 (0.99-1.12)	1.29 (1.09-1.53)	0.96 (0.80-1.16)	1.07 (0.61-1.89)
Reduced HDL-C	1.20 (1.11-1.29)	1.25 (1.04-1.50)	1.36 (1.12-1.64)	2.56 (1.63-4.03)
EWC	1.12 (1.03-1.22)	1.04 (0.81-1.32)	1.07 (0.88-1.31)	1.42 (0.90-2.23)

[a] DM: diabetes mellitus.

[b] Model 1: unadjusted.

[c] Model 2: adjusted for age (continuous), gender, and race.

[d] Model 3: model 2 + educational levels, marriage status, family poverty-to-income ratio level, BMI (category), smoking, physical activity, and general health condition.

[e] Model 3.

[f] EBP: elevated blood pressure.

[g] EGLU: elevated fasting glucose.

[h] ETG: elevated triglycerides.

[i] HDL-C: high-density lipoprotein cholesterol.

[j] EWC: elevated waist circumference.

[k] Model 3+adjusted for other 4 components.

Figure 1. Unadjusted Kaplan-Meier hazard curves: (A) metabolic syndrome (MetS) and heart disease; (B) MetS and diabetes mellitus; (C) MetS and all causes; (D) the number of MetS components and heart disease; (E) the number of MetS components and diabetes mellitus; and (F) the number of MetS components and all causes.

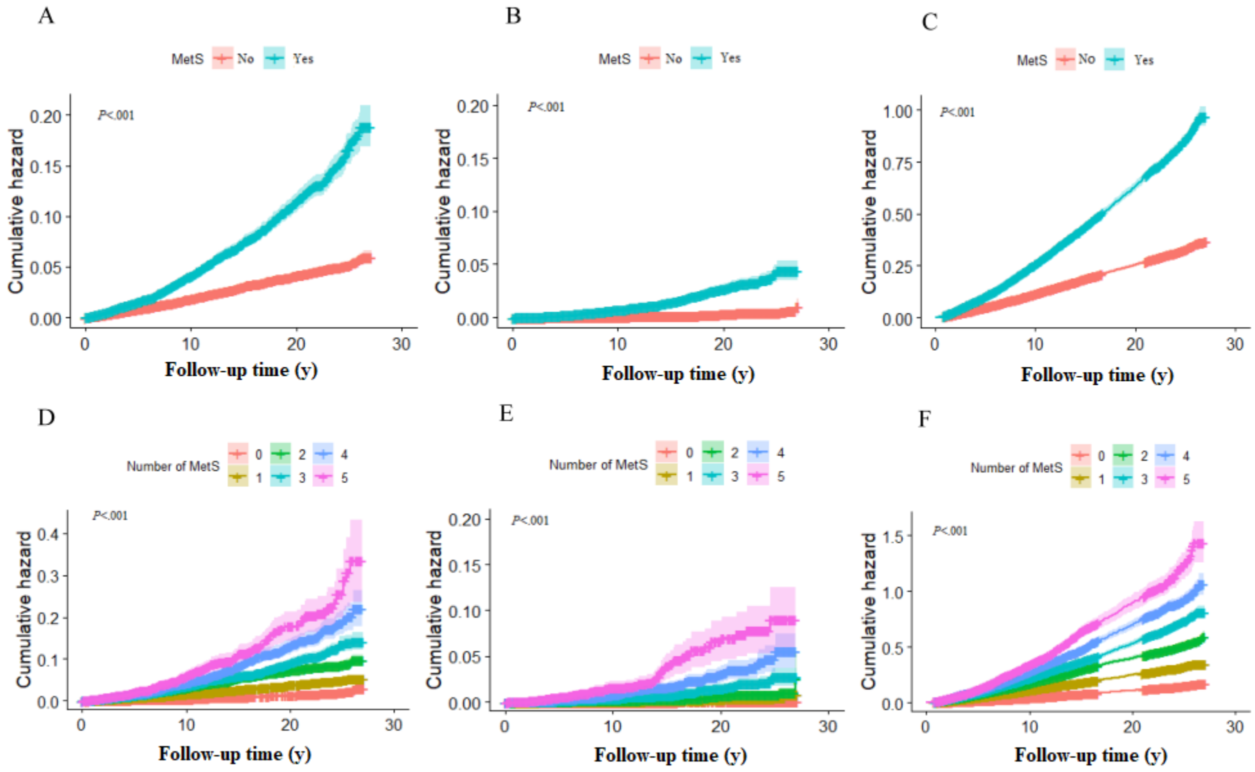


Table 3. Stratified analyses on associations of metabolic syndrome (MetS) and its components and all-cause and specific cause mortality (adjusted for other covariates).

Characteristic	Hazard ratio (95% CI)		
	All causes	Heart disease	DM[a]
Gender			
Man	1.23 (1.13-1.34)	1.54 (1.23-1.93)	3.60 (1.63-7.98)
Woman	1.26 (1.14-1.39)	1.38 (1.16-1.65)	8.75 (4.20-18.21)
Age group (years)			
<40	2.02 (1.49-2.74)	3.55 (1.01-12.46)	0.69 (0.18-2.63)
40-60	1.43 (1.18-1.73)	1.67 (1.06-2.63)	9.99 (3.87-25.80)
≥60	1.17 (1.08-1.27)	1.33 (1.15-1.54)	4.55 (2.46-8.43)
Race and ethnicity			
Mexican American	1.20 (1.07-1.34)	1.27 (0.83-1.93)	2.89 (1.30-6.45)
Non-Hispanic Black	1.34 (1.18-1.52)	1.25 (0.91-1.71)	4.62 (2.59-8.26)
Non-Hispanic White	1.21 (1.11-1.32)	1.45 (1.22-1.73)	5.54 (2.78-11.02)
Others	1.59 (1.07-2.36)	2.30 (0.67-7.93)	N/A[b]
Educational levels			
Less than high school	1.26 (1.18-1.36)	1.39 (1.17-1.65)	4.94 (3.11-7.84)
High school or equivalent	1.19 (0.98-1.44)	1.07 (0.51-2.24)	1.00 (0.28-3.52)
College or above	1.18 (1.01-1.38)	1.99 (1.27-3.10)	31.27 (8.22-118.85)
Married status			
Married	1.18 (0.92-1.50)	1.57 (0.25-9.97)	0.93 (0.19-4.54)
Separated	1.23 (1.15-1.32)	1.40 (1.20-1.64)	5.42 (3.17-9.30)
Never married	1.00 (0.67-1.50)	1.44 (0.60-3.47)	9.22 (0.53-162.16)
Family poverty-to-income ratio level			
0-1.0	1.30 (1.07-1.57)	1.19 (0.74-1.90)	1.61 (0.67-3.87)
1.1-3.0	1.12 (1.03-1.23)	1.26 (0.99-1.60)	8.38 (3.00-23.46)
>3.0	1.36 (1.21-1.53)	1.82 (1.46-2.28)	6.03 (2.96-12.29)
BMI (kg/m^2)			
<25	1.22 (1.11-1.35)	1.40 (1.07-1.83)	5.57 (2.65-11.73)
25-30	1.14 (1.02-1.27)	1.33 (1.06-1.68)	2.91 (1.35-6.27)
≥30	1.45 (1.18-1.77)	1.72 (1.11-2.66)	13.83 (4.58-41.76)
Smoking status			
Yes	1.28 (1.15-1.42)	1.58 (1.25-1.99)	5.60 (3.01-10.44)
No	1.18 (1.08-1.30)	1.30 (1.08-1.57)	4.65 (1.92-11.29)
Active physical activity			
Yes	1.28 (1.15-1.42)	1.53 (1.30-1.81)	4.17 (2.12-8.21)
No	1.17 (1.05-1.31)	1.27 (0.92-1.75)	7.71 (3.32-17.91)
General health condition			
Very good to excellent	1.25 (1.12-1.40)	1.76 (1.37-2.25)	3.58 (1.72-7.45)
Good	1.26 (1.13-1.40)	1.24 (1.02-1.51)	8.94 (2.60-30.73)
Poor to fair	1.14 (1.01-1.28)	1.44 (1.10-1.89)	5.23 (1.84-14.82)

[a] DM: diabetes mellitus.

[b] N/A: not applicable.

Discussion

Principal Findings

Our study suggested that MetS was associated with a significantly elevated hazard of all-cause, heart disease, and DM mortality, and the associations were also significant after adjusting for sociodemographic factors, lifestyle factors, and health status. Existing studies have reported consistent results regarding the relationship between MetS and the risk of CVD mortality [13], in which MetS was defined using the WHO definition, and our study provided further evidence of the association between MetS and mortality hazard, using the NCEP ATP Ⅲ definition. However, the nonsignificant association of MetS and all-cause mortality was observed in another study of men with the NCEP ATP III definition (relative risk 1.67, 95% CI 0.91-3.08) [12], and our study showed that men with MetS had an increased hazard of all-cause mortality. Different definitions of MetS may be a main cause. In addition, population differences and sample size may also be the important factors. Moreover, the previous study only included 1029 men aged 42-60 years. Our study was conducted with a large sample size among adults aged 18-90 years, which could provide more information. The findings may have important public health implications. MetS is prevalent in many areas around the world and has been reported to be associated with several diseases. Our study provides supportive evidence on a positive association between MetS and mortality, which indicates an urgent need of the prevention and control of MetS.

Previous studies have revealed that the number of MetS components was related to DM and CVD in a dose-response way [6,7], which may explain why individuals with more MetS components had higher hazards of all-cause, heart disease, and DM mortality in this study. Furthermore, there were great differences between the components of MetS and mortality, especially for heart disease and DM mortality, which may be due to the different pathological mechanisms and the interactive effects of the components [23-25]. In addition, our study suggested that MetS was not associated with cancer mortality, but a reduced HDL-C level may increase the hazard of cancer mortality. Previous studies showed that metabolic abnormalities or the underlying insulin resistance; specific cancer sites such as breast and uterine cancer [26,27]; and cancer mortality risk tended to increase with an increasing number of metabolic abnormalities in women but not in men [14]. More epidemiological studies and pathological mechanisms on the issue should be conducted and discussed.

Furthermore, our study indicated that there were significant associations of MetS with all-cause, heart disease, and DM mortality among individuals with the separated marital status, including the widowed and divorced groups, but not among those with other marital statuses. This suggests that health care professionals need pay more attention to these individuals. Besides, we also found significant associations of MetS with heart disease and DM mortality among the population with high PIR and the non-Hispanic White population. High prevalence of unhealthy lifestyle and risk factors of cardiometabolic diseases in high-income populations [28] and the fact that obesity prevalence varied by racial and ethnic group [29] may result in this phenomenon. In addition, men with MetS had higher risk of heart disease mortality, whereas women with MetS had higher risk of DM mortality. Gender differences in various disease and related risk factors has been observed [30,31], which may be due to the biology of gender; however, variation in cultural, societal, and historical contexts could also result in different life experiences of men and women and variation in the mortality [32]. Therefore, public health policies should recognize variations across genders as well as incorporate cultural and societal factors within and across countries [33]. Besides, a higher HRs was observed in all-cause, heart disease, and DM mortality among smoking individuals, which is easy to understand considering the obvious and well-known hazards of smoking to human health [34,35]. Our study once again suggested that intervention in smokers may contribute to reducing mortality.

This study had several strengths. First, the prospective cohort was well designed with a large and nationally representative sample and a reliable assessment of cause of deaths. Second, we conducted detailed analyses of MetS and its components and mortality, and we stratified the analyses. Besides, we adjusted for a large number of potential confounders, including socioeconomic status and lifestyle factors. Some limitations should also be acknowledged. First, the limited number of DM deaths may affect the validity; however, it does not seem to have a significant influence on the main findings of the study. Second, many covariates were mainly self-reported and were available only at baseline. Thus, it was impossible to use time-varying covariates to capture changes in the possible confounders over time. Third, our findings may only be representative of US residents, and further study should be conducted to validate the generalizability to other populations. Fourth, people with diabetes often die from complications of diabetes, such as heart and kidney diseases, and the cause of death may be reported as heart disease, not diabetes. Besides, lifestyle changes might have occurred since the survey, which may affect the results. Furthermore, people without a diagnosis of diabetes, with parameters in the normal or prediabetes range, may progress to type 2 diabetes along with its complications, which also might influence the results. Last, but not least, due to the nature of observational studies, residual confounding was still possible.

Conclusion

In conclusion, MetS is associated with higher hazard of all-cause, heart disease, and DM mortality among US adults. Future studies are needed to reveal the mechanisms underlying the association between MetS and mortality. Both the identification of MetS using simple and efficient criteria and the early prevention and treatment are of great importance to improving the healthy life expectancy of a population, which should be of concern to health care professionals.

Acknowledgments

We thank the study group of the National Health and Nutrition Examination Survey.

The study was supported by grants from the Research Committee—Research Assistant Professorship Scheme (2021-22; reference SMH/NL/JF/RAP2122/0752/22en), Research Committee—Strategic Seed Funding for Collaborative Research Scheme (2022-23; reference TL/SSFCRS2223/0953/23ww), the Chinese University of Hong Kong, Hong Kong Research Grants Council—General Research Fund (grant 165056653), Health and Medical Research Fund (reference 18190471), Vice-Chancellor Discretionary Fund (Round III; grant 136366853), Hong Kong Research Grants Council (Research Matching Grants 8601219 and 8601242), and a grant from the Research Committee on Research Sustainability of Major Research Grants Council Funding Schemes (3133235) of the Chinese University of Hong Kong. The funders did not play any role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; nor in the preparation, review, or approval of the manuscript.

Data Availability

All the data used in this study are available on the web [36].

Authors' Contributions

WL wrote the manuscript and performed statistical analysis. WL and LAT conceived and designed the study and interpreted the data. DC, YP, ZL, and MPK critically revised the manuscript and approved the final version. LAT and WL are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analyses, flowchart of the study, and supplementary methods.
[DOCX File , 186 KB - publichealth_v9i1e44073_app1.docx]

References

1. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet 2005 Apr;365(9468):1415-1428. [doi: 10.1016/S0140-6736(05)66378-7] [Medline: 15836891]
2. Li W, Song F, Wang X, Wang L, Wang D, Yin X, et al. Prevalence of metabolic syndrome among middle-aged and elderly adults in China: current status and temporal trends. Ann Med 2018 Jun;50(4):345-353. [doi: 10.1080/07853890.2018.1464202] [Medline: 29637787]
3. Huang PL. A comprehensive definition for metabolic syndrome. Dis Model Mech 2009 Apr 30;2(5-6):231-237 [FREE Full text] [doi: 10.1242/dmm.001180] [Medline: 19407331]
4. Yu Y, Paul K, Arah OA, Mayeda ER, Wu J, Lee E, et al. Air pollution, noise exposure, and metabolic syndrome - a cohort study in elderly Mexican-Americans in Sacramento area. Environ Int 2020 Jan;134:105269 [FREE Full text] [doi: 10.1016/j.envint.2019.105269] [Medline: 31778933]
5. Gu D, Reynolds K, Wu X, Chen J, Duan X, Reynolds RF, InterASIA Collaborative Group. Prevalence of the metabolic syndrome and overweight among adults in China. Lancet 2005 Apr;365(9468):1398-1405. [doi: 10.1016/S0140-6736(05)66375-1] [Medline: 15836888]
6. Li W, Song F, Wang X, Wang D, Chen D, Yue W, et al. Relationship between metabolic syndrome and its components and cardiovascular disease in middle-aged and elderly Chinese population: a national cross-sectional survey. BMJ Open 2019 Aug 18;9(8):e027545 [FREE Full text] [doi: 10.1136/bmjopen-2018-027545] [Medline: 31427317]
7. Li W, Wang D, Wang X, Gong Y, Cao S, Yin X, et al. The association of metabolic syndrome components and diabetes mellitus: evidence from China National Stroke Screening and Prevention Project. BMC Public Health 2019 Feb 14;19(1):192 [FREE Full text] [doi: 10.1186/s12889-019-6415-z] [Medline: 30764803]
8. Aalemi AK, Bahain MB, Hamdard AG. Metabolic syndrome and psoriasis: a case-control study in Kabul, Afghanistan. Diabetes Metab Syndr Obes 2021 Mar 30;14:1465-1471 [FREE Full text] [doi: 10.2147/DMSO.S305806] [Medline: 33833537]
9. Adibi N, Robati RM. Skin and metabolic syndrome: a review of the possible associations. J Res Med Sci 2021 Feb 27;26:16 [FREE Full text] [doi: 10.4103/jrms.JRMS_585_20] [Medline: 34084195]
10. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. JAMA 2002 Jan 16;287(3):356-359. [doi: 10.1001/jama.287.3.356] [Medline: 11790215]

11. Gupta R, Deedwania PC, Gupta A, Rastogi S, Panwar RB, Kothari K. Prevalence of metabolic syndrome in an Indian urban population. Int J Cardiol 2004 Nov;97(2):257-261. [doi: 10.1016/j.ijcard.2003.11.003] [Medline: 15458693]
12. Lakka HM, Laaksonen DE, Lakka TA, Niskanen LK, Kumpusalo E, Tuomilehto J, et al. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. JAMA 2002 Dec 04;288(21):2709-2716. [doi: 10.1001/jama.288.21.2709] [Medline: 12460094]
13. Isomaa B, Almgren P, Tuomi T, Forsén B, Lahti K, Nissén M, et al. Cardiovascular morbidity and mortality associated with the metabolic syndrome. Diabetes Care 2001 Apr;24(4):683-689. [doi: 10.2337/diacare.24.4.683] [Medline: 11315831]
14. Trevisan M, Liu J, Bahsas FB, Menotti A, Risk Factor and Life Expectancy Research Group. Syndrome X and mortality: a population-based study. Am J Epidemiol 1998 Nov 15;148(10):958-966. [doi: 10.1093/oxfordjournals.aje.a009572] [Medline: 9829867]
15. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, American Heart Association, National Heart, Lung, and Blood Institute. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement. Circulation 2005 Oct 25;112(17):2735-2752. [doi: 10.1161/CIRCULATIONAHA.105.169404] [Medline: 16157765]
16. NHANES survey methods and analytic guidelines. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx [accessed 2023-08-28]
17. 2019 public-use linked mortality files. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/data-linkage/mortality-public.htm# [accessed 2023-08-28]
18. Expert Panel on the Identification, Evaluation, and Treatment of Overweight, Obesity in Adults, National Heart, Lung, and Blood Institute, National Institute of Diabetes and Digestive and Kidney Diseases. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: the Evidence Report. Bethesda, MD: National Institutes of Health; Sep 1998.
19. Brown AF, Liang L, Vassar SD, Escarce JJ, Merkin SS, Cheng E, et al. Trends in racial/ethnic and nativity disparities in cardiovascular health among adults without prevalent cardiovascular disease in the United States, 1988 to 2014. Ann Intern Med 2018 Apr 17;168(8):541-549 [FREE Full text] [doi: 10.7326/M17-0996] [Medline: 29554692]
20. Chen C, Ye Y, Zhang Y, Pan X, Pan A. Weight change across adulthood in relation to all cause and cause specific mortality: prospective cohort study. BMJ 2019 Oct 16;367:l5584 [FREE Full text] [doi: 10.1136/bmj.l5584] [Medline: 31619383]
21. NCHS ethics review board (ERB) approval. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm [accessed 2023-08-28]
22. NHANES - National Health and Nutritional Examination Survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2023-08-28]
23. Grundy SM. Metabolic syndrome: connecting and reconciling cardiovascular and diabetes worlds. J Am Coll Cardiol 2006 Mar 21;47(6):1093-1100 [FREE Full text] [doi: 10.1016/j.jacc.2005.11.046] [Medline: 16545636]
24. Liu SJ, Guo ZR, Hu XS, Wu M, Chen FM, Kang GD, et al. Risks for type-2 diabetes associated with the metabolic syndrome and the interaction between impaired fasting glucose and other components of metabolic syndrome the study from Jiangsu, China of 5 years follow-up. Diabetes Res Clin Pract 2008 Jul;81(1):117-123. [doi: 10.1016/j.diabres.2008.03.008] [Medline: 18485514]
25. Kahn R, Buse J, Ferrannini E, Stern M, American Diabetes Association, European Association for the Study of Diabetes. The metabolic syndrome: time for a critical appraisal: joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. Diabetes Care 2005 Sep;28(9):2289-2304. [doi: 10.2337/diacare.28.9.2289] [Medline: 16123508]
26. Kazer RR. Insulin resistance, insulin-like growth factor I and breast cancer: a hypothesis. Int J Cancer 1995 Aug 09;62(4):403-406. [doi: 10.1002/ijc.2910620408] [Medline: 7543456]
27. Nyholm H, Djursing H, Hagen C, Agner T, Bennett P, Svenstrup B. Androgens and estrogens in postmenopausal insulin-treated diabetic women. J Clin Endocrinol Metab 1989 Nov;69(5):946-949. [doi: 10.1210/jcem-69-5-946] [Medline: 2677038]
28. Abdalla SM, Yu S, Galea S. Trends of biomarkers of cardiovascular disease in the United States by income: disparities between the richest 20% and the poorest 80%,1999-2018. SSM Popul Health 2021 Mar;13:100745 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100745] [Medline: 33604447]
29. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA 2010 Jan 20;303(3):235-241. [doi: 10.1001/jama.2009.2014] [Medline: 20071471]
30. Wu YT, Niubo AS, Daskalopoulou C, Moreno-Agostino D, Stefler D, Bobak M, et al. Sex differences in mortality: results from a population-based study of 12 longitudinal cohorts. CMAJ 2021 Mar 15;193(11):E361-E370 [FREE Full text] [doi: 10.1503/cmaj.200484] [Medline: 33722827]
31. Aaberg ML, Burch DM, Hud ZR, Zacharias MP. Gender differences in the onset of diabetic neuropathy. J Diabetes Complications 2008 Mar;22(2):83-87. [doi: 10.1016/j.jdiacomp.2007.06.009] [Medline: 18280437]
32. Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. Lancet 2011 Oct 08;378(9799):1297-1305. [doi: 10.1016/S0140-6736(11)60781-2] [Medline: 21839503]

33. Nelson JP, McNall AD. What happens to drinking when alcohol policy changes? a review of five natural experiments for alcohol taxes, prices, and availability. Eur J Health Econ 2017 May;18(4):417-434. [doi: 10.1007/s10198-016-0795-0] [Medline: 27055901]
34. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: an update. J Am Coll Cardiol 2004 May 19;43(10):1731-1737 [FREE Full text] [doi: 10.1016/j.jacc.2003.12.047] [Medline: 15145091]
35. Maddatu J, Anderson-Baucum E, Evans-Molina C. Smoking and the risk of type 2 diabetes. Transl Res 2017 Jun;184:101-107 [FREE Full text] [doi: 10.1016/j.trsl.2017.02.004] [Medline: 28336465]
36. NHANES 2017-2018. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017&tdsourcetag=s_pcqq_aiomsg [accessed 2023-08-28]

Abbreviations

ATP: Adult Treatment Panel
CVD: cardiovascular disease
DM: diabetes mellitus
EWC: elevated waist circumference
HDL-C: high-density lipoprotein cholesterol
HR: hazard ratio
ICD: International Classification of Diseases
MetS: metabolic syndrome
NCEP: National Cholesterol Education Program
NHANES: National Health and Nutrition Examination Survey
PIR: poverty-to-income ratio
TG: triglyceride
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 05.11.22; peer-reviewed by D Wang, M Yang, L Quintana, M Nomali; comments to author 29.01.23; revised version received 30.04.23; accepted 16.08.23; published 05.09.23.

Please cite as:
Li W, Chen D, Peng Y, Lu Z, Kwan MP, Tse LA
Association Between Metabolic Syndrome and Mortality: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e44073
URL: https://publichealth.jmir.org/2023/1/e44073
doi:10.2196/44073
PMID:37669100

Original Paper

Trends in Cause-Specific Injury Mortality in China in 2005-2019: Longitudinal Observational Study

Zixiang Ji[1*], BMed; Hengjing Wu[1*], PhD; Rongyu Zhu[1], BMed; Lu Wang[1], BSc; Yuzhu Wang[1], BMed; Lijuan Zhang[1], PhD

[1] Clinical Center for Intelligent Rehabilitation Research, Shanghai YangZhi Rehabilitation Hospital, Tongji University School of Medicine, Tongji University, Shanghai, China

[*] these authors contributed equally

Corresponding Author:
Lijuan Zhang, PhD
Clinical Center for Intelligent Rehabilitation Research, Shanghai YangZhi Rehabilitation Hospital
Tongji University School of Medicine, Tongji University
50 Chifeng Road, Yangpu
Shanghai, 201619
China
Phone: 86 13817934887
Email: zhangxiaoyi@tongji.edu.cn

Abstract

Background: Over the last few decades, although the age-standardized mortality rate (ASMR) of injury has shown a significant declining trend in China, this pattern has dramatically reversed recently.

Objective: We aimed to elucidate the geographical, demographic, and temporal trends of cause-specific injuries, the reversal phenomenon of these trends, and the fluctuations of injury burden from 2005 to 2019 in China.

Methods: A longitudinal observational study was performed using the raw data of injury deaths in the National Cause-of-Death surveillance data provided by the disease surveillance points system in 2005-2019. The cause-specific injuries were divided into disparate subgroups by sex, age, urban/rural region, and eastern/central/western areas of China. The burden of injury was assessed using potential years of life lost (PYLL), average years of life lost (AYLL), and PYLL rate (PYLLR). Temporal trends of mortality rates and burden were evaluated using best-fitting joinpoint models.

Results: Injury deaths accounted for 7.51% (1,156,504/15,403,835) of all-cause deaths in China in 2005-2019. The crude mortality rate of all-cause injury was 47.74 per 100,000 persons. The top 3 injury types (traffic accident, falls, and suicide) accounted for 70.57% (816,145/1,156,504) of all injury-related deaths. The ASMR of all-cause injury decreased (P=.003), while the crude mortality rate remained unchanged (P=.52) during 2005-2019. A significant reverse trend in ASMR of all-cause injury was observed in urban older adults since 2013, mainly due to the inverted trend in injuries from falls. A reverse trend in ASMR of suicide was observed among individuals aged 10-24 years, with notable increases by 35.18% (annual percentage change 15.4%, 95% CI 4.1%-28.0%) in men since 2017. The AYLL and PYLLR of all-cause injury among older adults showed consistent ascending trends from 2005 to 2019 (average annual percentage change [AAPC] 6.1%, 95% CI 5.4%-6.9%, 129.04% increase for AYLL; AAPC 5.4%, 95% CI 2.4%-8.4%, 105.52% increase for PYLLR). The AYLL due to suicide for individuals aged 10-24 years showed a considerable upswing tendency (AAPC 0.5%, 95% CI 0.4%-0.7%, 8.02% increase).

Conclusions: Although the ASMR of all-cause injury decreased in China from 2005 to 2019, the trend in suicide among adolescents and young adults and falls among older adults has been on the rise in recent years. Interventions should be encouraged to mitigate the cause-specific burdens of injury death.

(JMIR Public Health Surveill 2023;9:e47902) doi:10.2196/47902

KEYWORDS

reverse; age-standardized mortality rate; injury; suicide; trend; potential years of life lost; average years of life lost; crude mortality rate; falls; older adults; young adults

Introduction

Injury continues to pose a significant public health risk factor worldwide with a high mortality rate, and it contributes to more than 5 million deaths annually [1], accounting for about 9% of global mortality [2]. In China, injury is the fifth leading cause of death [3], accounting for almost 500,000 deaths and 300 million injuries per year [4], and the massive cost of health care amounts to US $9.6 billion annually [5]. Injury-related long-term disability sequelae [6] and deaths could have important negative effects on families, societies, and governments, including decline of the economy, destruction of social production, and restriction of the growth of the country's overall strengths to some extent.

With its reform and opening-up policy as well as its membership in the World Health Organization, China has experienced an unprecedented rise in motorized vehicles and a major transformation in the mode of travel and transportation [7]. The incidence of injuries from motor vehicle crashes has substantially increased, particularly in the rural regions of China, where the conflict between the proliferation of cars and the inability to keep up with road traffic management measures has greatly contributed to the frequency of motor vehicle crashes [5]. Meanwhile, drowning has become the leading cause of childhood injury mortality [5] due to the high percentage of children left without supervision and care because of increased working opportunities and income for parents in urban regions. The advancement of ideologies has lagged economic growth, and suicide among rural women has become the leading cause of injury death [8]. However, with the introduction of the "Health China 2030 plan," the concept of health has gradually been reinforced, and cause-specific injury has shown a meaningful declining trajectory [9] with the strengthening of health education, improvement in the health service system, and optimization of related policies and measures. Despite advancements in health policy initiatives in China, differences in injury mortalities in the underprivileged communities require further investigations. The origins of injury-related deaths in rural regions, particularly those caused by suicide, may alter significantly as time goes on, globalization increases, and ideologies shift drastically.

This study investigates the association of the geographical, demographic, and temporal trends of cause-specific injuries with the magnitude and patterns of injury mortality in mainland China during 2005-2019. Temporal and age-specific trends were analyzed to explore the intricacies of injury mortality in various periods, geographies, and populations. Our findings can serve as a foundation for further research and provide the government and administration with scientific evidence for the formulation of relevant government policies and execution of judgments to lower the risks of injury and death.

Methods

Source of Data

Injury death data were derived from the National Cause-of-Death Surveillance data set, which was recorded and rechecked for a fixed period by the disease surveillance points (DSP) system established in 1978. By 2013, the number of points had increased to 605 with a population of more than 300 million Chinese residents, accounting for more than 24% of the whole national population [7]. The strict procedures [10] involved in collecting and evaluating data were implemented between upper and lower administrative regions to ensure the data's completeness, authenticity, and reliability. Information on individual deaths in all population catchment areas has been reported in real time by using an internet-based reporting system by the local Centers for Disease Control and Prevention (CDC) since 2008. Local CDCs also check the coding and internal logic of items reported on death certificates, and then death data are subsequently reported to the national CDC. Following a review and correction during the previous quarter, the Chronic Disease Center of the CDC finally receives all death numbers from the monitoring sites. Any unreliable reports are addressed by thoroughly analyzing medical records or by conducting verbal autopsies. Additionally, every 3 years, a regular survey of a nationally representative sample is conducted at each disease monitoring point location to collect data that correct underreporting. Together, the DSPs provided a nationally representative picture of mortality in China. The International Classification of Diseases tenth revision was used to code injuries based on their underlying causes of mortality. In general, the injuries (V0-Y89) are split into 2 categories: intentional injuries and unintentional injuries. The formed category consists of suicide and other intentional injuries, while the latter consists of traffic accidents, poisoning, falls, fires, drowning, and other unintentional injuries. We disaggregated the data by subgroups such as sex (men and women), region (urban and rural), area (eastern, central, and western), and age (≤39 years, 40-64 years, and ≥65 years).

Data Analysis

The crude mortality rate (CMR), age-standardized mortality rate (ASMR) and 95% CIs, and rates of rank and proportion of death causes were calculated every year from 2005 to 2019. For overall and cause-specific injury, risk ratio (RR) values for CMR and ASMR were calculated separately for different sexes, regions, and areas. The age-period-cohort model was implemented to develop independent effect estimates of age, period, and birth cohort on injury and suicide mortality to analyze period- and cohort-related risks and longitudinal and cross-sectional age trends. Populations were divided into 5-year intervals from 2005 to 2019. The longitudinal age curve comprised consecutive 5-year age segments from 0-4 years to 85+ years for all-cause injury and from 5-9 years to 85+ years for suicide due to virtually no suicide deaths among those younger than 5 years. The RR values and the estimable parameters and functions concerning the age-period-cohort model should be reanalyzed by Fisher exact test if the Wald chi-square test did not meet the criteria. Temporal and age-specific trends both reflected changes in injury mortality. Joinpoint regression analysis was implemented to log-transform the CMR and ASMR, and the optimum yearly trend curve and rate of variation were then exhibited using the Grid Search method, Monte Carlo permutation tests, and Modified Bayesian Information Criterion. Annual percentage change (APC) showed the rate for certain segments within the study period, while average APC (AAPC) represented the average change rate from

2005 to 2019, and 95% CI suggested statistical significance. The potential years of life lost (PYLL), average years of life lost (AYLL), and PYLL rate (PYLLR) (Multimedia Appendix 1) were indicators of the injury burden that quantified the impact of premature mortality on PYLL. Epidata 3.1 (The Epidata Association, Odense) and Excel (version 2019; Microsoft) were utilized for data entry and management. Joinpoint (Joinpoint Regression Program version 4.9.1.0; National Cancer Institute) was applied to present the trends of injury mortality changes under best-fitting joinpoint models. The age-period-cohort online tool [11] (National Cancer Institute) provided the calculated parameters of the age-period-cohort model. R software (version 4.2.2; MathSoft) was used for statistical analysis. The results were considered significant when *P* values were less than .05 in 2-sided tests.

Ethics Approval

This research received ethics approval from the ethics committee of the School of Medicine, Tongji University, Shanghai (approval 2023tjdxsy037).

Results

Basic Information on Injury Deaths in China During 2005-2019

A total of 2,414,759,166 person-years recorded by DSPs were included in this investigation from 2005 to 2019. The number of injury deaths was 1,156,504 (men:women=2.1:1.0), accounting for 7.51% of all-cause deaths (n=15,403,835) within the study period. The details of the annual demographic data are presented in Table S2 of Multimedia Appendix 2, and the entire aggregate of information concerning injuries in these 15 years is shown in Tables S3 and S4 of Multimedia Appendix 2. The CMR of all injuries was 47.74 per 100,000 (95% CI 47.65-47.82 per 100,000). Traffic accidents (16.57 per 100,000) was the leading cause of injury-related deaths, followed by falls (9.66 per 100,000), suicide (7.46 per 100,000), drowning (3.75 per 100,000), poisoning (3.00 per 100,000), and fires (0.64 per 100,000). The 3 major causes of death accounted for 70.57% (1,156,504/15,403,835) of all injury-related deaths, with ASMR following the same pattern. In the whole population, the CMRs and ASMRs of all-cause injury in men were significantly higher than those in women ($RR_{\text{men vs women}}$ 2.024, 95% CI 2.016-2.032 for CMR; $RR_{\text{men vs women}}$ 2.397, 95% CI 2.389-2.407 for ASMR). The mortality rates of all-cause injury in rural residents were generally higher than those in urban residents ($RR_{\text{rural vs urban}}$ 1.492, 95% CI 1.486-1.499 for CMR; $RR_{\text{rural vs urban}}$ 1.577, 95% CI 1.570-1.584 for ASMR). Western area residents had considerably higher mortality rates of all-cause injury than eastern area residents ($RR_{\text{western vs eastern}}$ 1.179, 95% CI 1.174-1.185 for CMR; $RR_{\text{western vs eastern}}$ 1.408, 95% CI 1.401-1.415 for ASMR) and central area residents ($RR_{\text{western vs central}}$ 1.146, 95% CI 1.140-1.151 for CMR; $RR_{\text{western vs central}}$ 1.214, 95% CI 1.208-1.219 for ASMR). We also ranked each injury type according to disparate geographical and economic regions (Figures S1 and S2 of Multimedia Appendix 2). Traffic accidents and suicide were the top 2 injury deaths in the central areas of China, while traffic accidents and falls were the leading causes of injury in other regions.

Temporal Trends of Injuries in China From 2005 to 2019

The ASMR of all-cause injury decreased by 33.42% (AAPC –2.6%, 95% CI –4.3% to –0.9%), while the CMR remained stable (AAPC –0.7%, 95% CI –2.8% to 1.4%) from 2005 to 2019 (Figure 1). The ASMRs of suicide decreased by 61.40% (AAPC –6.3%, 95% CI –7.8% to –4.9%), drowning decreased by 48.26% (AAPC –4.6%, 95% CI –7.4% to –1.7%), fires decreased by 50.73% (AAPC –4.0%, 95% CI –7.5% to –0.3%), and poisoning decreased by 35.12% (AAPC –2.9%, 95% CI –4.9% to –0.8%), while traffic accidents and falls exhibited no significant change during this period (both *P*>.05). Trends of injury ASMR by sex, region, area, age, and type are shown in Table S7 of Multimedia Appendix 2. Briefly, the ASMR of all-cause injury decreased by 31.85% (AAPC –2.4%, 95% CI –4.2% to –0.5%) in men and by 36.67% in women (AAPC –3.1%, 95% CI –4.4% to –1.7%). The ASMR of all-cause injury in rural residents decreased by 33.85% (AAPC –2.8%, 95% CI –5.4% to –0.1%), while the ASMR of urban residents remained unchanged (AAPC –2.3%, 95% CI –6.7% to 2.2%). The ASMR of all injuries decreased by 31.79% in the eastern areas (AAPC –2.4%, 95% CI –4.9% to 0.1%), 38.90% in the central areas (AAPC –3.3%, 95% CI –4.8% to –1.6%), and 28.68% in the western areas (AAPC –2.0%, 95% CI –3.9% to 0.0%). Despite a substantial downtrend in overall suicides between 2005 and 2019, various subgroups showed distinct patterns (Figure 2, Table S8 of Multimedia Appendix 2). The overall ASMR of suicide decreased in men by 57.19% (AAPC –5.5%, 95% CI –7.3% to –3.7%) and by 66.62% in women (AAPC –7.4%, 95% CI –8.6% to –6.1%). The ASMR of suicide in men was significantly higher than that in women (RR 1.380, 95% CI 1.367-1.394) and the trend of RR increased by 28.25% (AAPC 1.8%, 95% CI 1.1%-2.5%) during the whole study period, particularly after 2009 (APC 2.4%, 95% CI 1.8%-3.1%). Although the ASMR of suicide decreased in the rural regions by 63.09% (AAPC –6.6%, 95% CI –7.4% to –5.8%) and by 56.35% in the urban regions (AAPC –5.5%, 95% CI –8.3% to –2.7%), the ASMR of suicide in rural residents was higher than that in urban residents (RR 1.767, 95% CI 1.747-1.788). The ratio of urban region to rural region suicide ASMRs showed a generally increasing tendency, inching closer to 1, especially after 2008 (36.93% increase, APC 2.8%, 95% CI 1.3%-4.3%). Further analysis revealed that the highest ASMR of suicide was continuously reported in rural men, while the ASMR of suicide in urban women was consistently minimum. Contrary to all-cause injury, suicide presented the highest ASMR in the central areas of China, where there was also the steepest decline in the slope in 2005-2019 (64.94% decrease in the central areas, AAPC –7.0%, 95% CI –8.4% to –5.5%), followed by the eastern areas (60.25% decrease, AAPC –6.2%, 95% CI –8.4% to –3.9%) and the western areas (57.79% decrease, AAPC –5.7%, 95% CI –8.4% to –2.8%).

Figure 1. Trends in the mortality rates of all-cause injury and disparate types of injury in China in 2005-2019. A. Trends in crude mortality rates and age-standardized mortality rates per 100,000 cases of all-cause injury in 2005-2019. B. Trends in age-standardized mortality rates per 100,000 cases of disparate types of injury in 2005-2019. ASMR: age-standardized mortality rate; CMR: crude mortality rate. **P*<.05, ***P*<.01, ****P*<.001.

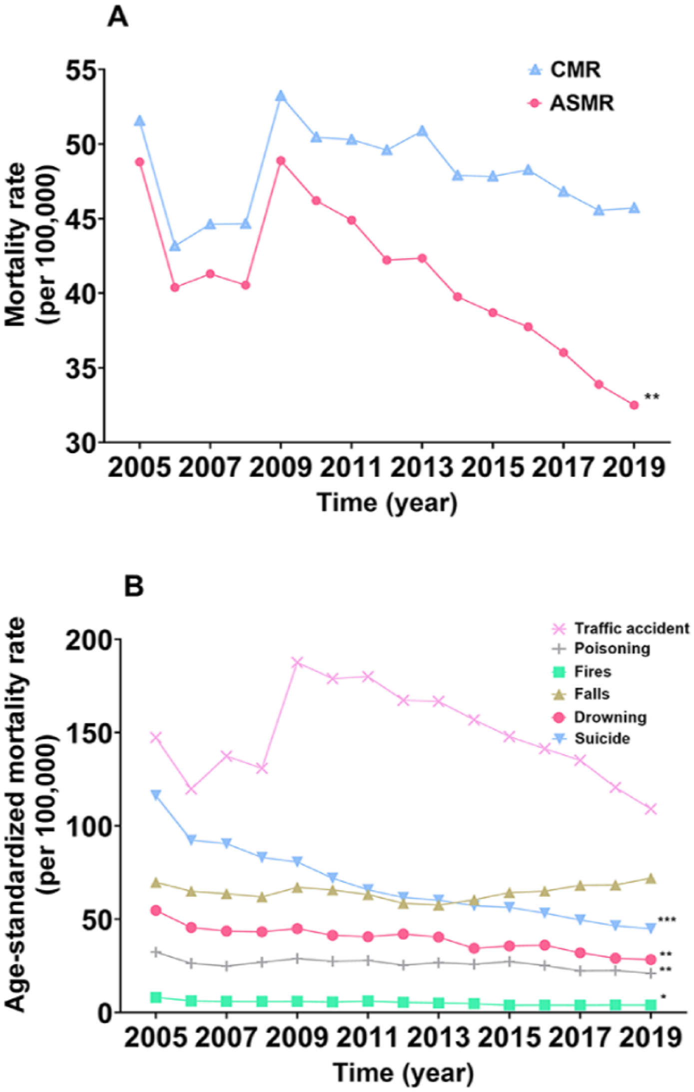
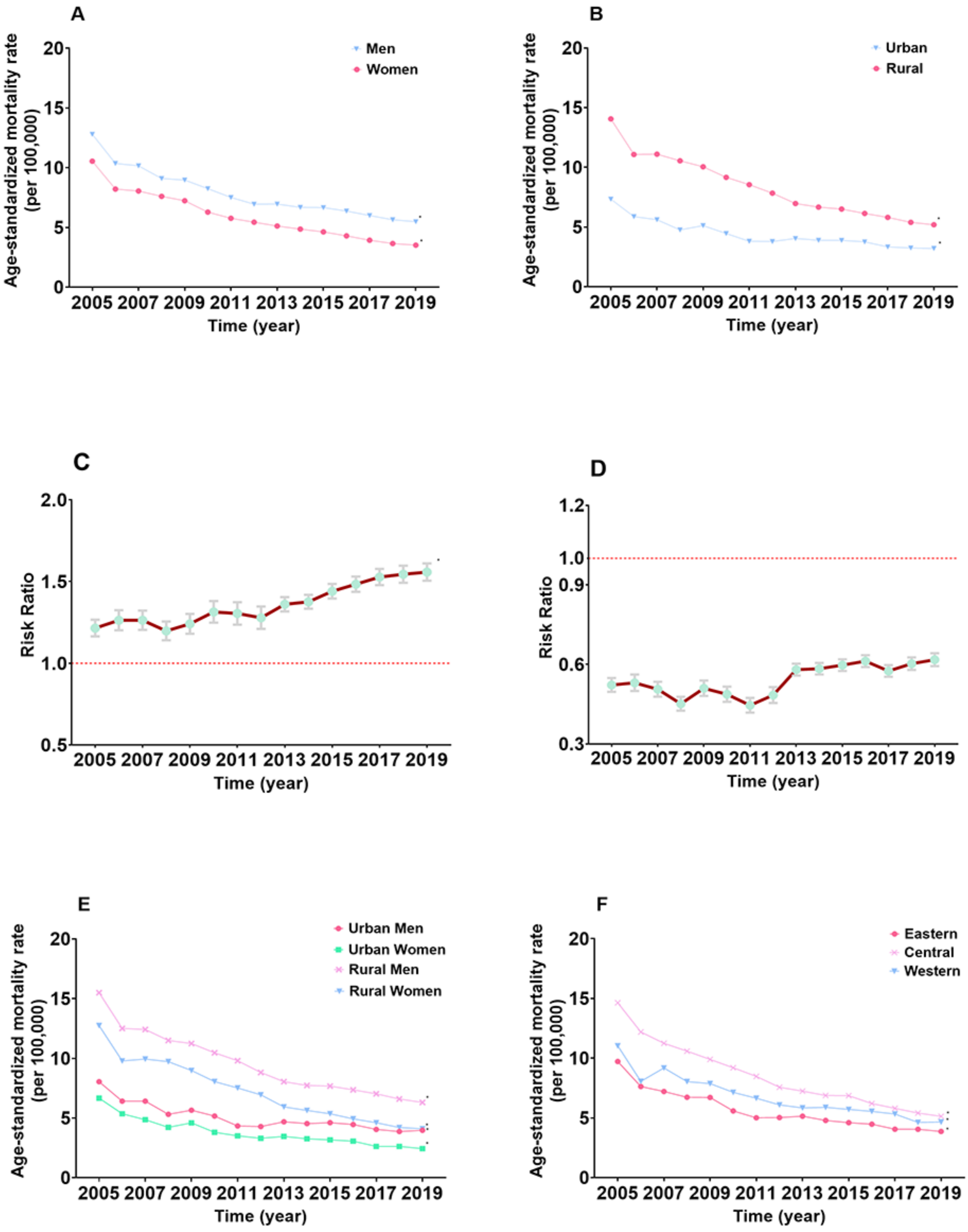


Figure 2. Trends in age-standardized mortality rates per 100,000 cases and risk ratio of suicide in the whole population in China in 2005-2019. A. The trend in age-standardized mortality rate by sex. B. The trend in age-standardized mortality rate by region. C. The trend in age-standardized mortality rate risk ratio of men to women. D. The trend in age-standardized mortality rate risk ratio of urban regions to rural regions. E. The trend in age-standardized mortality rate by sex and region. F. The trend in age-standardized mortality rate by area. *P<.001.



Temporal Reverse Trends of Injuries in China From 2005 to 2019

Several age groups showed reversals in the temporal trends for all-cause injury and suicide ASMRs in contrast to the overall population’s patterns. We divided the participants into 3 age groups: ≥65 years old (older adults), 40-64 years old, and ≤39 years old. The ASMRs of all-cause injury in the ≤39 years and 40-64 years populations decreased by 46.69% and 31.40%, respectively, but the decrease was less obvious in the older adult populations (*P*=.93). Among the older adult populations, a reverse trend of all-cause injury was observed since 2013. Therefore, we further analyzed the underlying situations and present the results in Figure 3. These reverse trends occurred since 2013, especially among men in the urban regions (32.38% increase, APC 1.7%, 95% CI 0.6%-2.8%) and women in the western areas of China (41.73% increase, APC 1.9%, 95% CI 0.8%-2.9%), although there were increasing tendencies across all areas, as well as all men and women except for the central area residents. Among older adults, falls were the leading cause of the reversal trends. The ASMRs of falls showed significantly increasing tendencies since 2013 among all subgroups, even among the central area residents (All *P*<.05, Figure S3 of Multimedia Appendix 2). After stratification by age, we found the trend in ASMR of suicide among adolescents and young adults aged 10-24 years had reversed recently (Figure 4, Figure S4 of Multimedia Appendix 2). Detailed analysis revealed a noticeable increase by 35.18% (APC 15.4%, 95% CI 4.1%-28.0%) in males aged 10-24 years since 2017 (Figure 4). Remarkable increases in the urban (73.98% increase, APC 32.4%, 95% CI 9.9%-59.3%) and eastern area (78.32% increase, APC 34.7%, 95% CI 6.8%-69.9%) residents were found in this study, which might be contributed by men (58.58% increase, APC 27.5%, 95% CI 2.4%-58.8% for urban men; 81.80% increase, APC 29.9%, 95% CI 4.0%-62.4% for eastern area men). Suicide was the third leading cause of injury among adolescents and young adults aged 10-24 years, accounting for 13.32% of all deaths caused by overall injuries, second only to traffic accidents (41.02%) and drowning (19.58%). There were significant differences in CMRs of suicide for both sexes within the disparate subgroups (Figure 5). The CMR of suicide increased with age, and the 20-24 years age group showed the greatest difference (RR 1.375, 95% CI 1.308-1.447), followed by the 10-14 years age group (RR 1.220, 95% CI 1.091-1.365) and the 15-19 years age group (RR 1.192, 95% CI 1.118-1.271). Although the CMR in rural regions was considerably higher than that in urban regions, the difference in the CMR between both sexes in urban regions (RR 1.297, 95% CI 1.204-1.397) indicated more significance than that between both sexes in the rural regions (RR 1.262, 95% CI 1.208-1.318). The CMR of suicide increased from the eastern area to the western area of China with statistically significant disparities, but the differences in CMR of suicide between men and women decreased from the eastern to western areas of China.

Figure 3. Trends in age-standardized mortality rates per 100,000 older adults in China in 2005-2019. A. The trend in age-standardized mortality rate among older adults by sex. B. The trend in age-standardized mortality rate of older adults by urban region and sex. C. The trend in age-standardized mortality rate of older adults by rural region and sex. D. The trend in age-standardized mortality rate of older adults by eastern area and sex. E. The trend in age-standardized mortality rate of older adults by central area and sex. F. The trend in age-standardized mortality rate of older adults by western area and sex. **P*<.01 (for trends in 2013-2019).

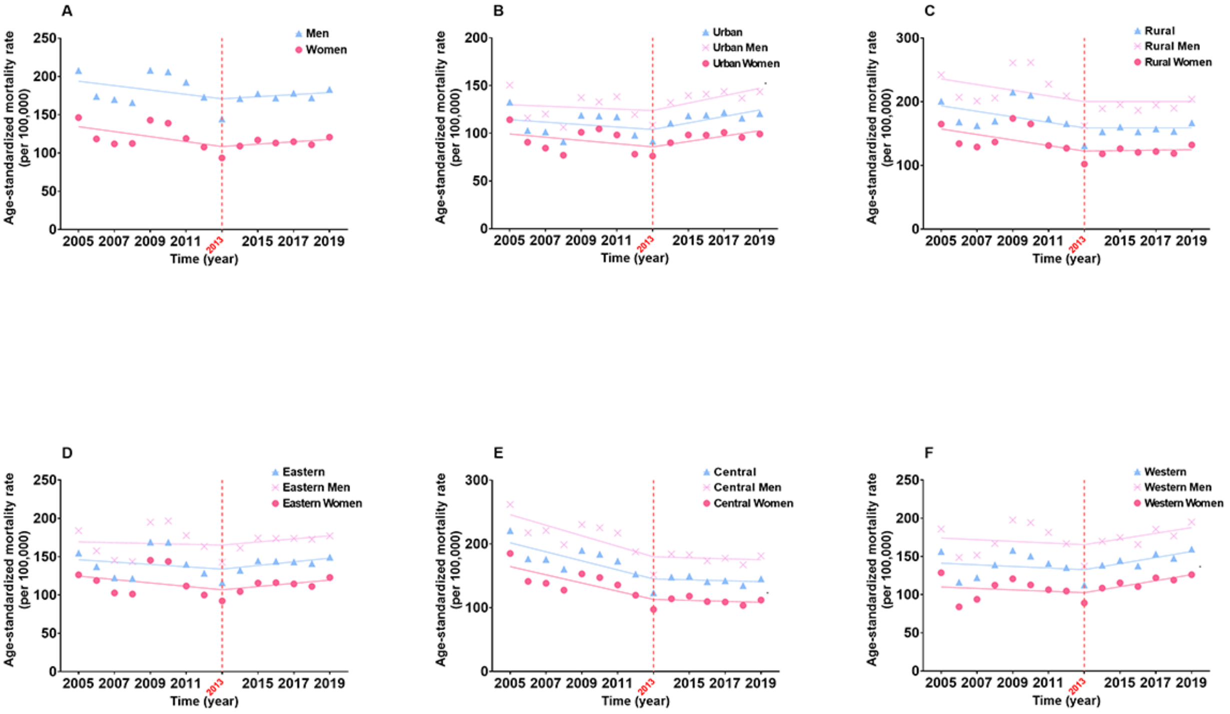


Figure 4. Trends of suicide among young individuals aged 10-24 years in the whole population as well as urban and eastern area populations of China and the ranks of the leading causes of injury among young individuals aged 10-24 years in 2005-2019. A. The trend of suicide among males aged 10-24 years. B. The trend of suicide among females aged 10-24 years. C. The ranks of the leading causes of injury among young individuals aged 10-24 years. D. The trend of suicide among young individuals aged 10-24 years in urban regions. E. The trend of suicide among males aged 10-24 years in urban regions. F. The trend of suicide among females aged 10-24 years in urban regions. G. The trend of suicide among young individuals aged 10-24 years in eastern areas. H. The trend of suicide among males aged 10-24 years in eastern areas. I. The trend of suicide among females aged 10-24 years in eastern areas. **P*<.05, ***P*<.01 (for trends in 2017-2019).

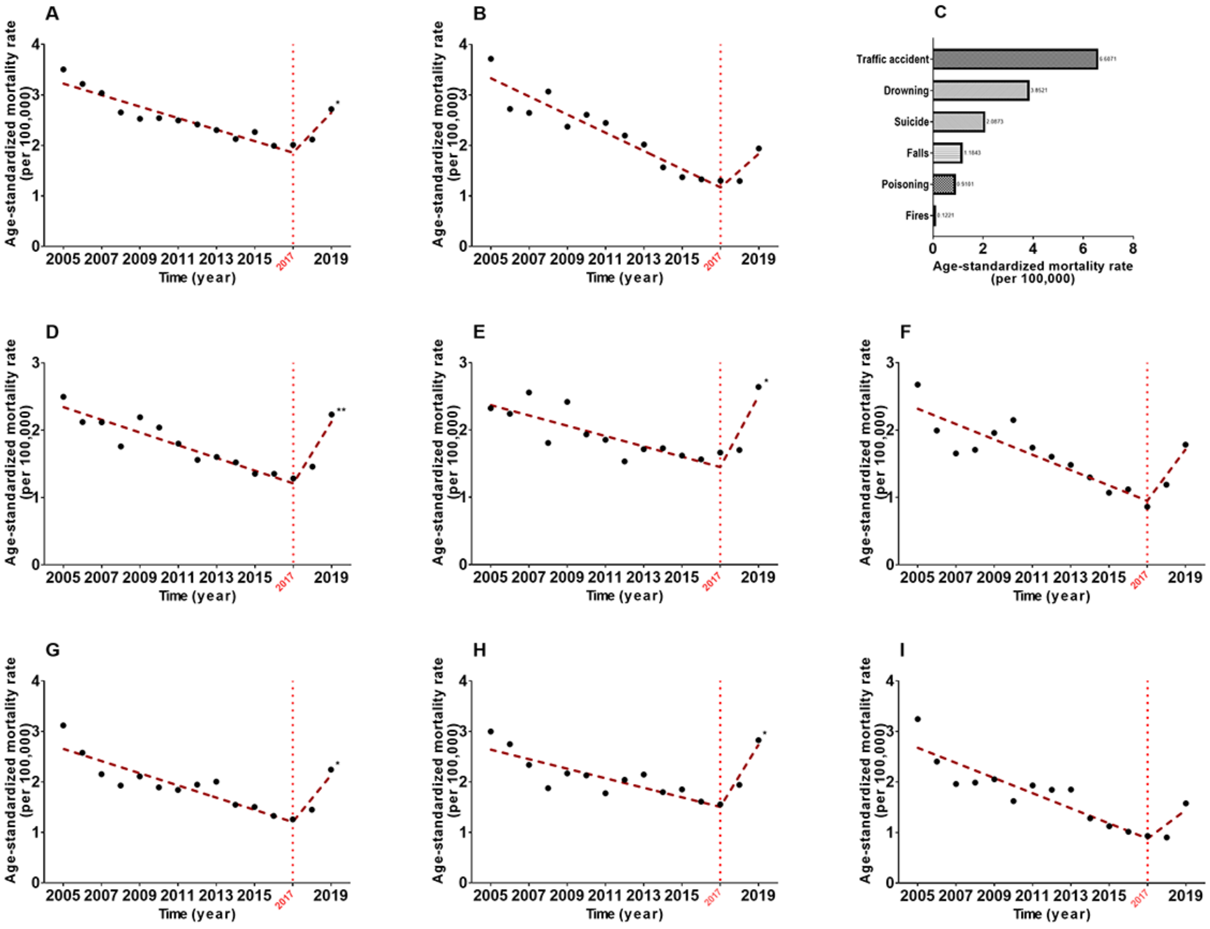
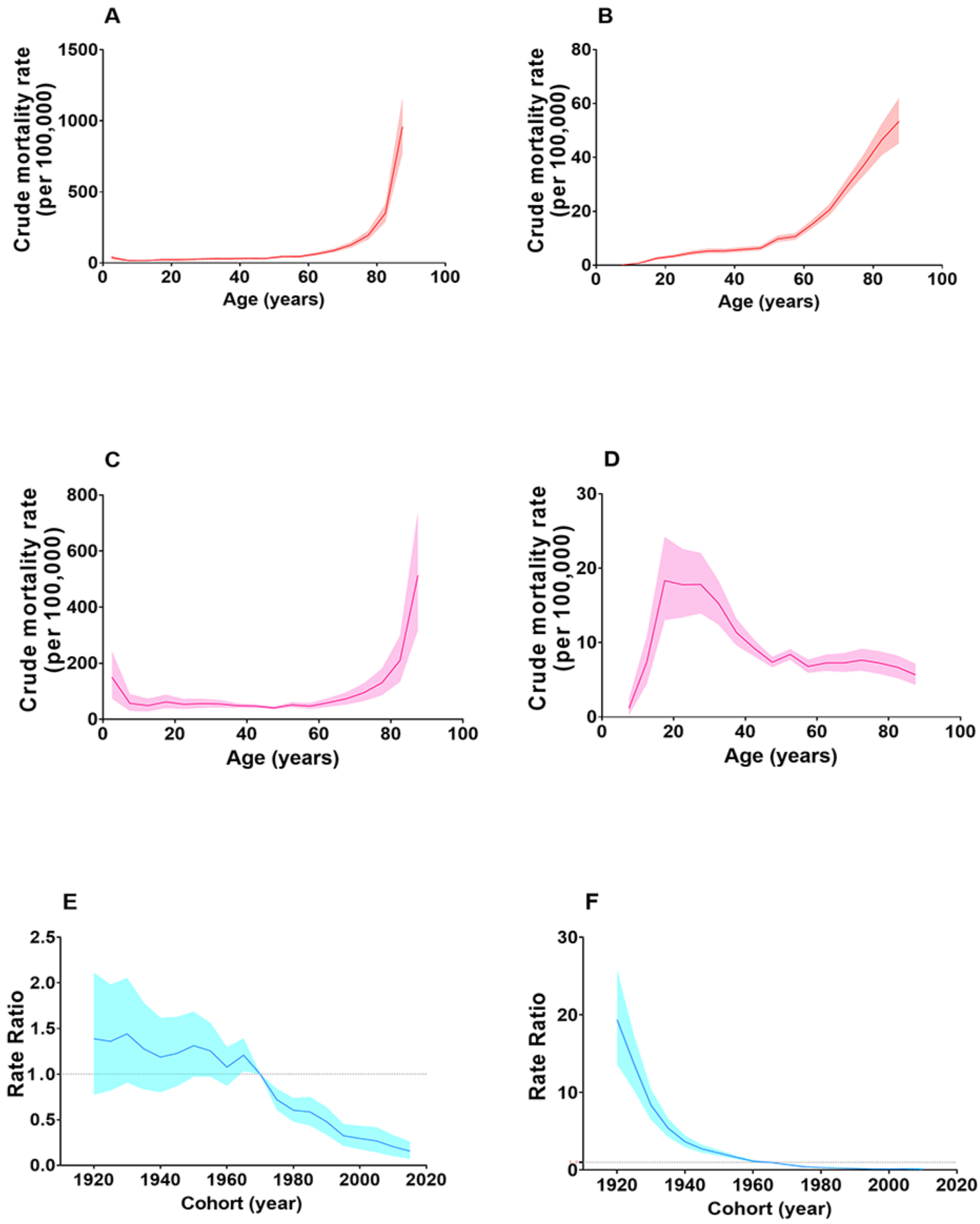


Figure 5. The crude mortality rate per 100,000 suicide cases of both sexes within the disparate subgroups among adolescents and young adults aged 10-24 years in China in 2005-2019. A. Crude mortality rates in the different age groups in China in 2005-2019. B. Crude mortality rates in urban and rural China in 2005-2019. C. Crude mortality rates in eastern, central, and western areas of China in 2005-2019. **P*<.001.

Age-Specific and Cohort Trends of All-Cause Injury and Suicide During 2005-2019

The trends of all-cause injury as well as age-specific suicide are demonstrated by cross-sectional age curves and longitudinal age curves, respectively (Figure 6). For all-cause injury, the pattern of the cross-sectional age change curves was identical to that of the longitudinal age curves, with 0-5 to 10-14 years age group showing a downward trend and ≥15 years group generally exhibiting a progressive increase with age, notably among older adults. For suicide, the age-specific cross-sectional suicide curve indicated that mortality rates increased with age, whereas the longitudinal age curve showed an inverted "U" skewed peak shape, with a gradual increase in the age groups from 5-9 to 15-19 years, reaching a maximum among the 15-19 to 25-29 years, a pronounced decrease with increasing age among the 25-29 to 45-49 years age groups, and a stable tendency over 50 years. Moreover, the net drift in suicide from 2005 to 2019 was more than SD 1% per year, at –5.75% per year (95% CI –6.42 to –5.28) (all age deviations=0, Wald χ^2_{15}=122.6; P<.001). Compared to that of suicide (99.77% decrease, AAPC –27.2%, 95% CI –30.1% to –24.1%), the cohort-related risk of all-cause injury (89.24% decrease, AAPC –11.1%, 95% CI –15.7% to –6.3%) exhibited a significant downward trend among the 1920-1924 to 2015-2019 birth cohorts (Figure 6).

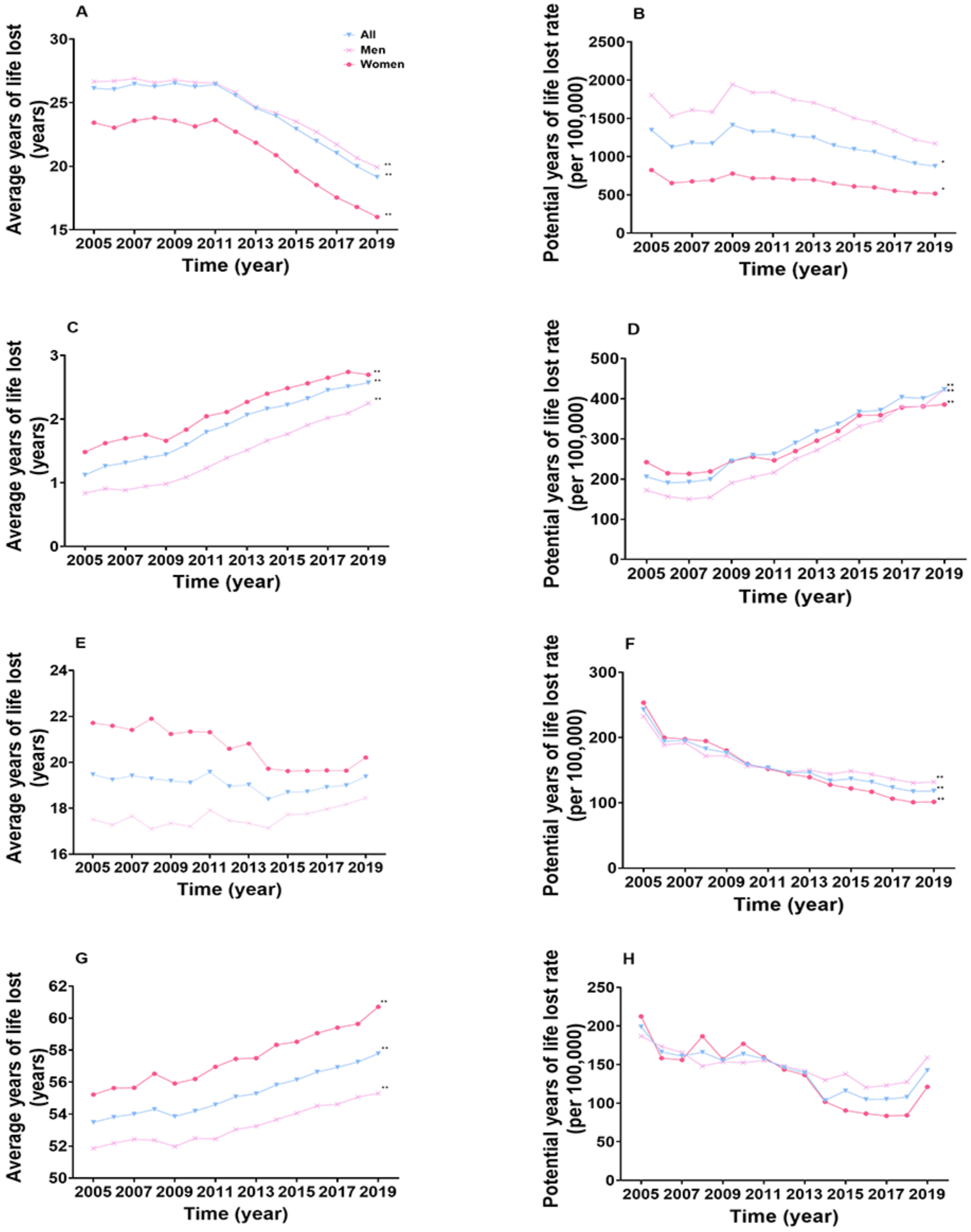
Figure 6. Age-specific and cohort trends of all-cause injury and suicide in China in 2005-2019. A. Cross-sectional age curve of all-cause injury among the 0-4 to 85+ years age group. B. Cross-sectional age curve of suicide among the 5-9 to 85+ years age group. C. Longitudinal age curve of all-cause injury among the 0-4 to 85+ years age group. D. Longitudinal age curve of suicide among the 5-9 to 85+ years group. E. The trend in the risk ratio of all-cause injury crude mortality rates among birth cohorts in 1920-1924 and 2015-2019 using the birth cohort of 1970-1974 as the reference group. F. The trend in the risk ratio of suicide crude mortality rates among the birth cohort in 1920-1924 and 2010-2014 by using the birth cohort of 1965-1969 as the reference group.



Temporal Trends of Burden in Injuries in China During 2005-2019

In terms of all-cause injury in the whole population during 2005-2019, AYLL decreased by 26.71% (AAPC –2.2%, 95% CI –2.5% to –2.0%) and PYLLR decreased by 35.01% (AAPC –2.8%, 95% CI –4.7% to –0.8%), and both these trends were found to decrease faster in women (31.62% decrease, AAPC –2.7%, 95% CI –3.3% to –2.1% for AYLL; 37.29% decrease, AAPC –3.1%, 95% CI –5.1% to –1.0% for PYLLR) than in men (25.25% decrease, AAPC –2.1%, 95% CI –2.4% to –1.8% for AYLL; 34.93% decrease, AAPC –2.8%, 95% CI –5.6% to 0.1% for PYLLR) (Figure 7). For all-cause injury among older adults, AYLL (129.04% increase, AAPC 6.1%, 95% CI 5.4%-6.9%) and PYLLR (105.52% increase, AAPC 5.4%, 95% CI 2.4%-8.4%) increased consistently throughout the 2005-2019 period, with the same patterns occurring in both men (168.49% increase, AAPC 7.2%, 95% CI 6.1%-8.4% for AYLL; 145.85% increase, AAPC 6.6%, 95% CI 3.3%-9.9% for PYLLR) and women (81.81% increase, AAPC 4.2%, 95% CI 2.9%-5.6% for AYLL; 59.31% decrease, AAPC 4.1%, 95% CI 2.0%-6.2% for PYLLR). In the whole population, AYLL due to suicide, with higher AYLL among women, demonstrated an increasing tendency since 2014 (5.34% increase, APC 0.8%, 95% CI 0.5%-1.1%), especially for men (7.68% increase, APC 1.3%, 95% CI 0.9%-1.8%) (Figure 7). Distinctly, PYLLR of suicide in the whole population decreased by 51.34% (AAPC –4.9%, 95% CI –6.9% to –2.8%), while women (59.95% decrease, AAPC –6.2%, 95% CI –7.5% to –4.9%) showed a larger reduction than men (43.19% decrease, AAPC –3.7%, 95% CI –5.5% to –2.0%). For youths aged 10-24 years, a continuous upward trend occurred in AYLL in 2005-2019 (8.02% increase, AAPC 0.5%, 95% CI 0.4%-0.7%) similar to that among men, which increased by 6.63% (AAPC 0.5%, 95% CI 0.3%-0.7%) while that among women increased by 9.92% (AAPC 0.6%, 95% CI 0.4%-0.9%). PYLLR of suicide among adolescents and young adults aged 10-24 years remained stable in 2005-2019.

Figure 7. Trends in the average years of life lost and potential years of life lost rate in all-cause injury and suicide by sex and age in China in 2005-2019. A. The trend of average years of life lost in all-cause injury in the whole study population in 2005-2019. B. The trend of potential years of life lost rate in all-cause injury in the whole study population in 2005-2019. C. The trend of average years of life lost in all-cause injury among older adults in 2005-2019. D. The trend of potential years of life lost rate in all-cause injury among older adults in 2005-2019. E. The trend of average years of life lost in suicide in the study population in 2005-2019. F. The trend of potential years of life lost rate in suicide in the whole study population in 2005-2019. G. The trend of average years of life lost due to suicide among adolescents and young adults aged 10-24 years in 2005-2019. H. The trend of potential years of life lost rate in suicide among adolescents and young adults aged 10-24 years in 2005-2019. **P*<.01, ***P*<.001 (for trends in 2005-2019).



Discussion

Principal Findings

In this study, our main findings were that (1) the ASMR of all-cause injury decreased while the CMR remained unchanged in 2005-2019; (2) significant increases in ASMRs of all-cause injury among urban male older adults and western area female older adults were detected since 2013; (3) the ASMR of suicide among males aged 10-24 years in the whole population as well as among the urban and the eastern area residents reversed remarkably since 2017; (4) in the older adult group, AYLL and PYLL of all-cause injury showed a substantial increase in the past 15 years; and (5) in the 10-24 years age group, AYLL due to suicide showed a significantly increasing trend (Figure S6 of Multimedia Appendix 2). The slope of the decreasing trend of ASMR in men was smaller than that in women, which is in line with that reported in other studies [8,12]. This phenomenon may be caused by inequalities in the sex ratio at birth [13], personalities [14], coping manners, attitudes [15], diverse roles, identities in the society, pressure from different walks, possibilities, degrees of experiencing various hazards [8,15,16], and lifestyles [4]. Our study shows that the ASMR of injury in rural residents was higher than that in urban residents, which was consistent with that reported in some previous studies performed in low-income or high-income countries [5,17]. The western area in China had the highest ASMR among all the areas, which may be linked to the multiple environmental constructions, inequality in medical resources, enforcement of pertinent laws and regulations, and disparate prehospital treatments due to different economic levels in those locations [18].

The leading cause of injury death in mainland China is traffic accidents. The roads in rural regions and western areas of China are more rugged, coupled with low awareness of safety (less use of seat belts and helmets, drunk driving [19], breaking traffic laws, etc), lack of health resources, and incomplete road constructions and facilities, resulting in a significantly higher mortality rate than that in other places. Given China's growing number of motor vehicles [20], effective policy development and stringent implementation such as the installation of a speed-limiting device [21] are particularly necessary. However, the slope decline in the rural regions was sharper than that in the urban regions, which might be due to the substantial resources and labor invested by the government [3] and comprehensive development that is being strengthened to bring about greater environmental transformations in rural regions.

With the aging of China's society [22], it is imperative to minimize injury deaths of older adults. The factors for the fall-oriented reversal trend might be associated with physical health, psychological condition, environmental elements, and medical resources [23,24]. The physical capabilities of aging populations continue to deteriorate, and diseases such as osteoporosis, cardiovascular conditions, and weakened reactions may dramatically impair daily activities, obviously increasing the risk of injury. Owing to urbanization and economic development, the majority of the young people in rural regions migrate to urban regions, leaving the older adults alone (empty nests). Consequently, there is a lack of timely assistance when injuries from falls occur [25]. Meanwhile, as the older adults are typically eager to spend time with family, the absence of care will eventually lead to psychological illness, which might be another potential risk factor. The complexity of the terrain and poor public transportation, particularly in mountainous areas, increases the risk of falls among older adults, and the shortage of medical aid supplies or inefficient use of available resources prevents the older adults injured in falls from receiving timely and effective treatment [25], frequently increasing the possibility of death. Therefore, the government should allocate more funds to alleviate the environmental issues and lack of medical resources resulting from economic backwardness, improve medical technology to treat older adults' illness more effectively, and pay more attention to older adults' psychological health with more public welfare activities to bring comfort in their loneliness.

A considerable decline in the ASMR of overall suicide across all patterns of injuries was detected in our study, unlike that in the United States and several east-Asian countries with rapidly increasing tendency [26,27]. Consistent with other researches [28,29], our findings revealed that suicide death in men was significantly higher than that in women, which was lower than that reported in the United States [30] and some other Asian countries [31]. The role of women in the family has undergone a dramatic change in Chinese traditions, wherein women are required to stay at home to raise children. It is well known that the high risk for suicide is a consequence of individuals living in communities where there is lack of regular interpersonal contact and communication. Women now have more opportunities to leave the home to pursue their own goals and engage in more social activities, which relieves the psychological strain and helps reduce the possibility of suicidal thoughts. When confronted with stress, women are generally more motivated than men to communicate and release to alleviate the stress; thus, men have a greater potential to commit suicide than women under the same high-pressure environments. Moreover, women are keener to care about the quality of life [32,33], which invariably increases the sense of well-being and contributes to the decline in suicide mortality among women. Therefore, it is critical to pay more attention to men's psychological health, relieve the societal pressure on men, enhance safety awareness education, and improve the overall satisfaction and quality of life. With a backward and undeveloped economy, more children in the family, and arduous and stressful life, many individuals in rural regions commit suicide to relieve themselves. However, as the economy of the country develops, people's standards of living will rise, urbanization will continue to increase, more rural residents will relocate to urban regions, and their incomes will increase, which can result in a decline in the suicide rate.

Notably, the ASMR of suicide in the youth group aged 10-24 years has reversed in recent years, which is identical to the trend observed in Shanghai [34], and college students are reported to have higher risks of suicidality [35]. The uprising tendency in more high-income areas (the urban and eastern), especially among men, was significant, as reported in Japan [36]. Additionally, age patterns indicated that young people faced a higher suicide risk (when the net drift is more than or equal to SD 1% per year, it suggests that there is a more severe bias when the cross-sectional age trend is used to explain the natural history pattern rather than the longitudinal age curve to a great extent [11]). Economic development has pros and cons. With urbanization and economic growth's positive effects on suicide steadily decreasing [37], competitive pressure has increased in all walks, and young people today undertake more stress in this fast-paced world, which is especially apparent among men in high-income regions [38]. On the one hand, young adults aged 20-24 years who have recently graduated and entered the society find it difficult to adjust to the pressures of the workplace and directly encounter economic issues, which differ from those in school. On the other hand, those who are studying are constantly under pressure from their peers and instructors. According to the interpersonal theory of suicide, perceived burdensomeness and a lack of belonging raise the potential of fatal suicide [39]. Consequently, youth resort to suicide deaths because they overlook their mental health and are not able to effectively manage their stress [40]. Based on the diathesis-stress model, vulnerability factors and stressful life events together increase the incidence of suicidality [41]. Due to the one-child policy adoption [42], the majority of the young Chinese generations are singletons, which increases the possibility of psychological strain since there are fewer adequate individuals with whom they can share when they run into difficulties. A concern that cannot be ignored is that the risk of suicidal ideation might also be worsened by an irregular lifestyle [43], including staying up too late and eating irregularly. Apart from the above, it is crucial to pay more attention to youths' emotional issues as they are not capable of handling stress on their own, which might negatively affect their psychological well-being if they are not addressed properly. Further, a thwarted sensation of belonging may be caused by single parents or reconstituted families, which, along with the rising divorce rate in contemporary society, might leave certain psychological scars that could later result in negative emotional outbursts in youths. Therefore, surveillance procedures in dormitories and campuses should be reinforced so that youths who exhibit aberrant behavior can receive active psychological therapy and treatment. Additionally, mental health education classes and counseling services could be offered to improve advice on dealing with psychological strain. Moreover, the media and other platforms could be used to spread awareness about mental health to the public, which can reduce the negative forms of therapy for mental health issues resulting from shyness or narrow-mindedness and enable adolescents and young adults to develop the right concept of stress relief or positive psychotherapy. The burden of injury among youth has been increasing over years, which might be linked to the escalating mental health issues resulting from increased stress in the contemporary society. Youths are the cornerstone and the future of family and society; thus, AYLL of youth will contribute to a deeper negative impact on the social economy and structure. The government in China should be urged to pay greater attention to young people's suicide as soon as possible in consideration of the sustained increase in AYLL.

Limitations

Our study has several limitations. First, the DSP system had a high rate of missing data, which could cause bias in the mortality rate computation. Second, the increase in DSP from 161 in 2005 to 605 in 2013 could make the data skewed. Third, there was no in-depth study on the underlying causative factors of assorted patterns of injuries in this study, such as suicide, in which pesticide poisoning and hanging [44] were the major causes of death. One study [45] indicated that the major way to commit suicide in China has changed from pesticide poisoning and hanging to falls and jumping from heights. The risk factors for falls vary among older adults, for example, some older adults are more likely to fall due to an underlying disease while others may have a physical disability. Fourth, further investigations on the before-and-after comparison of injury mortality rates in different locations and age groups could not be performed after the government and key departments had implemented corresponding preventative and solution measures to investigate appropriate policies and strategies. Fifth, our study primarily addresses the loss of life-years due to premature mortality without addressing the loss of life-years owing to health loss from nonfatal injuries. Sixth, the trend of all-cause injury mortality varied largely in 2005-2009, which mainly resulted from underreporting of deaths (the entirety of the nationwide crude mortality underreporting rate was 16.68% [6271/37,603], with a weighted adjusted underreporting rate of 17.44% [46]), changing in reporting methods (from filling out case report cards to reported online), and external factors (such as epidemics, natural disasters, or socioeconomic events, especially the 2008 Wenchuan earthquake that resulted in 18,000 people missing [5]).

Conclusion

Despite a markedly declining trend from 2005 to 2019, injury still accounts for a major part of deaths in China. The ASMR of suicide in adolescents and young adults has significantly reversed recently, while the ASMR of falls among older adults has increased since 2007. Our results provide an up-to-date scientific foundation for the establishment of policies and initiatives to prevent mortality due to cause-specific injury.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant 81872720) and the Shanghai Municipal Health Commission (grant 201840066) grants for LZ.

Data Availability

The data that support the findings of this study are available from the cause-of-death data set made by the Chinese Centers for Disease Control and Prevention, but some restrictions apply to the accessibility to the data.

Authors' Contributions

ZJ and HW are co-first authors. All authors contributed to the conception and methodology of this research. ZJ, RZ, LW, HW, and YW performed data entry and management. Data analysis was completed by ZJ and LZ. The preliminary draft was finished by ZJ, and all authors participated in the discussions and revision of the previous manuscript version. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Formulae and methods of injury-related indices of mortality and burden.
[DOCX File , 23 KB - publichealth_v9i1e47902_app1.docx]

Multimedia Appendix 2
Supplementary data on deaths and burdens of cause-specific injury.
[DOCX File , 2227 KB - publichealth_v9i1e47902_app2.docx]

References

1. Sun S, Yang L, Hu X, Zhu Y, Liu B, Yang Y, et al. The burden of the current curative expenditure of injury in Dalian, China-a study based on the "system of health accounts 2011". BMC Public Health 2021 Jan 19;21(1):157 [FREE Full text] [doi: 10.1186/s12889-021-10164-6] [Medline: 33468078]
2. GBD 2017 DALYsHALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1859-1922 [FREE Full text] [doi: 10.1016/S0140-6736(18)32335-3] [Medline: 30415748]
3. Ozanne-Smith J, Li Q. A social change perspective on injury prevention in China. Inj Prev 2018 Jun;24(Suppl 1):i25-i31. [doi: 10.1136/injuryprev-2017-042712] [Medline: 29730599]
4. Yuan X, Kou C, Zhang M, Ma W, Tang Z, Sun H, et al. Injury and poisoning mortality trends in urban and rural China from 2006 to 2020 based on age-period-cohort analysis. Int J Environ Res Public Health 2022 Jun 11;19(12):7172 [FREE Full text] [doi: 10.3390/ijerph19127172] [Medline: 35742421]
5. Wang SY, Li YH, Chi GB, Xiao SY, Ozanne-Smith J, Stevenson M, et al. Injury-related fatalities in China: an under-recognised public-health problem. Lancet 2008 Nov 15;372(9651):1765-1773. [doi: 10.1016/S0140-6736(08)61367-7] [Medline: 18930527]
6. GBD 2016 Traumatic Brain InjurySpinal Cord Injury Collaborators. Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Jan;18(1):56-87 [FREE Full text] [doi: 10.1016/S1474-4422(18)30415-0] [Medline: 30497965]
7. Wang L, Ning P, Yin P, Cheng P, Schwebel DC, Liu J, et al. Road traffic mortality in China: analysis of national surveillance data from 2006 to 2016. Lancet Public Health 2019 May;4(5):e245-e255 [FREE Full text] [doi: 10.1016/S2468-2667(19)30057-X] [Medline: 31054642]
8. Zhang L, Li Z, Li X, Zhang J, Zheng L, Jiang C, et al. Study on the trend and disease burden of injury deaths in Chinese population, 2004-2010. PLoS One 2014;9(1):e85319 [FREE Full text] [doi: 10.1371/journal.pone.0085319] [Medline: 24465534]
9. Leilei D, Pengpeng Y, Haagsma JA, Ye J, Yuan W, Yuliang E, et al. The burden of injury in China, 1990–2017: findings from the Global Burden of Disease Study 2017. The Lancet Public Health 2019 Sep;4(9):e449-e461. [doi: 10.1016/s2468-2667(19)30125-2]
10. Liu S, Wu X, Lopez AD, Wang L, Cai Y, Page A, et al. An integrated national mortality surveillance system for death registration and mortality surveillance, China. Bull World Health Organ 2016 Jan 01;94(1):46-57 [FREE Full text] [doi: 10.2471/BLT.15.153148] [Medline: 26769996]
11. Rosenberg PS, Check DP, Anderson WF. A web tool for age-period-cohort analysis of cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2296-2302 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0300] [Medline: 25146089]
12. Cai P, Wu X, Liu Z, Deng Y, Chen X, Yi G, et al. Analysis of the burden and trend of injury in Sichuan, China, from 2006 to 2015: results from the national injury surveillance system. BMJ Open 2019 Nov 18;9(11):e031184 [FREE Full text] [doi: 10.1136/bmjopen-2019-031184] [Medline: 31740468]
13. Jiang Q, Zhang C. Recent sex ratio at birth in China. BMJ Glob Health 2021 May;6(5):e005438 [FREE Full text] [doi: 10.1136/bmjgh-2021-005438] [Medline: 34006519]
14. Marusic A, Musek J, Gudjonsson G. Injury proneness and personality. Nord J Psychiatry 2001;55(3):157-161. [doi: 10.1080/08039480152036029] [Medline: 11827609]

15. Pant PR, Deave T, Banstola A, Bhatta S, Joshi E, Adhikari D, et al. Home-related and work-related injuries in Makwanpur district, Nepal: a household survey. Inj Prev 2021 Oct;27(5):450-455 [FREE Full text] [doi: 10.1136/injuryprev-2020-043986] [Medline: 33148799]
16. Hu X, Tang C, Wang D. Social media support and funding assistance for psychological injuries in social work. Front Public Health 2022;10:796769 [FREE Full text] [doi: 10.3389/fpubh.2022.796769] [Medline: 35602134]
17. He C, Liu L, Chu Y, Perin J, Dai L, Li X, et al. National and subnational all-cause and cause-specific child mortality in China, 1996-2015: a systematic analysis with implications for the Sustainable Development Goals. Lancet Glob Health 2017 Feb;5(2):e186-e197 [FREE Full text] [doi: 10.1016/S2214-109X(16)30334-5] [Medline: 28007477]
18. Jiang B, Liang S, Peng Z, Cong H, Levy M, Cheng Q, et al. Transport and public health in China: the road to a healthy future. The Lancet 2017 Oct;390(10104):1781-1791. [doi: 10.1016/s0140-6736(17)31958-x]
19. Rosen HE, Bari I, Paichadze N, Peden M, Khayesi M, Monclús J, et al. Global road safety 2010-18: An analysis of Global Status Reports. Online ahead of print. Injury 2022 Jul 20:504-506. [doi: 10.1016/j.injury.2022.07.030] [Medline: 35906119]
20. Wang L, Yu C, Zhang G, Zhang Y, Luo L. Comparison of secular trends in road injury mortality in China and the United States: an age-period-cohort analysis. Int J Environ Res Public Health 2018 Nov 09;15(11):2508 [FREE Full text] [doi: 10.3390/ijerph15112508] [Medline: 30423957]
21. Forbes G, Gardner T, Mcgee H, Srinivasan R. Methods and practices for setting speed limits: an informational report. United States Department of Transportation. URL: https://highways.dot.gov/safety/speed-management/methods-and-practices-setting-speed-limits-informational-report [accessed 2012-04-01]
22. The Lancet. Population ageing in China: crisis or opportunity? Lancet 2022 Nov 26;400(10366):1821. [doi: 10.1016/S0140-6736(22)02410-2] [Medline: 36436518]
23. WHO global report on falls prevention in older age. WHO. 2007. URL: https://www.who.int/publications/i/item/9789241563536 [accessed 2008-03-17]
24. Fan J, Yu C, Guo Y, Bian Z, Sun Z, Yang L, China Kadoorie Biobank Collaborative Group. Frailty index and all-cause and cause-specific mortality in Chinese adults: a prospective cohort study. Lancet Public Health 2020 Dec;5(12):e650-e660 [FREE Full text] [doi: 10.1016/S2468-2667(20)30113-4] [Medline: 33271078]
25. Nasr A, Tomasich F, Collao I, Abreu P, Marttos A. The Trauma Golden Hour. Switzerland: Springer Nature; 2020.
26. Hedegaard H, Curtin SC, Warner M. Suicide rates in the United States continue to increase. NCHS Data Brief 2018 Jun(309):1-8. [Medline: 30312151]
27. Chen Y, Wu KC, Yousuf S, Yip PSF. Suicide in Asia: opportunities and challenges. Epidemiol Rev 2012;34:129-144. [doi: 10.1093/epirev/mxr025] [Medline: 22158651]
28. Liu Y, Lan Z, Yin Y, Liu NH, Tong Y. Trends in suicide rates and the case-fatality of pesticide self-poisoning in an agricultural county in China, 2009 to 2014. J Affect Disord 2021 Mar 15;283:52-59. [doi: 10.1016/j.jad.2021.01.024] [Medline: 33517228]
29. Phillips MR, Li X, Zhang Y. Suicide rates in China, 1995-99. Lancet 2002 Mar 09;359(9309):835-840. [doi: 10.1016/S0140-6736(02)07954-0] [Medline: 11897283]
30. Garnett MF, Curtin SC, Stone DM. Suicide mortality in the United States, 2000-2020. NCHS Data Brief 2022 Mar(433):1-8 [FREE Full text] [Medline: 35312475]
31. Sinyor M, Tse R, Pirkis J. Global trends in suicide epidemiology. Curr Opin Psychiatry 2017 Jan;30(1):1-6. [doi: 10.1097/YCO.0000000000000296] [Medline: 27845946]
32. Lee RLT, Chien WT, Ligot J, Nailes JM, Tanida K, Takeuchi S, et al. Associations between quality of life, psychosocial well-being and health-related behaviors among adolescents in Chinese, Japanese, Taiwanese, Thai and the Filipino populations: a cross-sectional survey. Int J Environ Res Public Health 2020 Apr 01;17(7):2402 [FREE Full text] [doi: 10.3390/ijerph17072402] [Medline: 32244727]
33. Rinaldo N, Gualdi-Russo E, Khyatti M, Lakhoua C, Toselli S. Psychosocial health and quality of life among North African women. Health Soc Care Community 2022 Sep;30(5):1827-1837 [FREE Full text] [doi: 10.1111/hsc.13562] [Medline: 34519114]
34. Qiao J, Xia T, Fang B, Cai R, Chen L, Qian N, et al. The reversing trend in suicide rates in Shanghai, China, from 2002 to 2020. J Affect Disord 2022 Jul 01;308:147-154 [FREE Full text] [doi: 10.1016/j.jad.2022.04.056] [Medline: 35429532]
35. Lew B, Kõlves K, Osman A, Abu Talib M, Ibrahim N, Siau CS, et al. Suicidality among Chinese college students: A cross-sectional study across seven provinces. PLoS One 2020;15(8):e0237329 [FREE Full text] [doi: 10.1371/journal.pone.0237329] [Medline: 32822365]
36. Yoshioka E, Hanley SJ, Kawanishi Y, Saijo Y. Time trends in method-specific suicide rates in Japan, 1990-2011. Epidemiol Psychiatr Sci 2016 Feb;25(1):58-68 [FREE Full text] [doi: 10.1017/S2045796014000675] [Medline: 25373686]
37. Cai Z, Chen M, Ye P, Yip PS. Socio-economic determinants of suicide rates in transforming China: A spatial-temporal analysis from 1990 to 2015. Lancet Reg Health West Pac 2022 Feb;19:100341 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100341] [Medline: 35024666]
38. Zhang J, Li N, Tu X, Xiao S, Jia C. Risk factors for rural young suicide in China: a case-control study. J Affect Disord 2011 Mar;129(1-3):244-251 [FREE Full text] [doi: 10.1016/j.jad.2010.09.008] [Medline: 20880593]

39. Becker SP, Foster JA, Luebbe AM. A test of the interpersonal theory of suicide in college students. J Affect Disord 2020 Jan 01;260:73-76 [FREE Full text] [doi: 10.1016/j.jad.2019.09.005] [Medline: 31493642]
40. Oh HY, Marinovich C, Jay S, Zhou S, Kim JHJ. Abuse and suicide risk among college students in the United States: Findings from the 2019 Healthy Minds Study. J Affect Disord 2021 Mar 01;282:554-560. [doi: 10.1016/j.jad.2020.12.140] [Medline: 33433385]
41. Carballo JJ, Llorente C, Kehrmann L, Flamarique I, Zuddas A, Purper-Ouakil D, STOP Consortium. Psychosocial risk factors for suicidality in children and adolescents. Eur Child Adolesc Psychiatry 2020 Jun;29(6):759-776 [FREE Full text] [doi: 10.1007/s00787-018-01270-9] [Medline: 30684089]
42. Yang B, Ollendick TH, Dong Q, Xia Y, Lin L. Only children and children with siblings in the People's Republic of China: levels of fear, anxiety, and depression. Child Dev 1995 Oct;66(5):1301-1311. [Medline: 7555217]
43. Yu Y, Yao Z, Wei Y, Kou C, Yao B, Sun W, et al. Depressive symptoms as a mediator between excessive daytime sleepiness and suicidal ideation among Chinese college students. Int J Environ Res Public Health 2022 Dec 06;19(23):16334 [FREE Full text] [doi: 10.3390/ijerph192316334] [Medline: 36498403]
44. Yang G, Phillips MR, Zhou M, Wang L, Zhang Y, Xu D. Understanding the unique characteristics of suicide in China: national psychological autopsy study. Biomed Environ Sci 2005 Dec;18(6):379-389. [Medline: 16544520]
45. Preventing suicide: a global imperative. WHO. URL: https://www.who.int/publications/i/item/9789241564779 [accessed 2014-08-17]
46. Wang L, Wang L, Cai Y, Ma L, Zhou M. [Analysis of under-reporting of mortality surveillance from 2006 to 2008 in China]. Zhonghua Yu Fang Yi Xue Za Zhi 2011 Dec;45(12):1061-1064. [Medline: 22336336]

Abbreviations

AAPC: average annual percentage change
APC: annual percentage change
ASMR: age-standardized mortality rate
AYLL: average years of life lost
CDC: Centers for Disease Control and Prevention
CMR: crude mortality rate
DSP: disease surveillance points
PYLL: potential years of life lost
PYLLR: potential years of life lost rate
RR: risk ratio

Edited by A Mavragani, T Sanchez; submitted 05.04.23; peer-reviewed by Y Wang, Y Tong; comments to author 01.06.23; revised version received 17.06.23; accepted 25.07.23; published 15.09.23.

Please cite as:
Ji Z, Wu H, Zhu R, Wang L, Wang Y, Zhang L
Trends in Cause-Specific Injury Mortality in China in 2005-2019: Longitudinal Observational Study
JMIR Public Health Surveill 2023;9:e47902
URL: https://publichealth.jmir.org/2023/1/e47902
doi:10.2196/47902
PMID:37713250

Original Paper

Secular Trends in Gastric and Esophageal Cancer Attributable to Dietary Carcinogens From 1990 to 2019 and Projections Until 2044 in China: Population-Based Study

Hui-Xiao Guo[1*], MSc; Qi Wang[1*], MSc; Can Wang[1], MSc; Qing Chen Yin[1], BA; Hua Zhi Huo[1], BA; Bing Hua Yin[1], BA

Handan Central Hospital, Handan, China
*these authors contributed equally

Corresponding Author:
Bing Hua Yin, BA
Handan Central Hospital
Num 59, Congtai Road
Congtai District
Handan, 056001
China
Phone: 86 03102118383
Email: 553050711@qq.com

Abstract

Background: Little is known about trends in or projections of the disease burden of dietary gastric and esophageal cancer (GEC) in China.

Objective: We aim to report GEC deaths and disability-adjusted life years (DALYs) from 1990 to 2019, predict them through 2044, and decompose changes in terms of population growth, population aging, and epidemiological changes.

Methods: We retrieved dietary GEC data from the Global Burden of Disease (GBD) online database and used joinpoint regression and age-period-cohort models to analyze trends in dietary GEC deaths and DALYs from 1990 to 2019 in China. We used a Bayesian age period cohort model of integrated nested Laplace approximations to predict the disease burden of GEC through 2044 and obtained the estimated population of China from 2020 to 2050 from the Global Health Data Exchange website. Finally, we applied a recently developed decomposition method to attribute changes between 2019 and 2044 to population growth, population aging, and epidemiological changes.

Results: The summary exposure values and age-standardized rates decreased significantly from 1990 to 1999, with percentage changes of –0.06% (95% CI –0.11% to –0.02%) and –0.05% (95% CI –0.1% to –0.02%), respectively. From 1990 to 2019, for dietary esophageal cancer, the percentage change in age-standardized mortality rate (ASMR) was –0.79% (95% CI –0.93% to –0.58%) and the percentage change in age-standardized DALY rate (ASDR) was –0.81% (95% CI –0.94% to –0.61%); these were significant decreases. For dietary stomach cancer, significant decreases were also observed for the percentage change in ASMR (–0.43%, 95% CI –0.55% to –0.31%) and the percentage change in ASDR (–0.47%, 95% CI –0.58% to –0.35%). In addition, data from both the joinpoint regression and annual percentage change analyses demonstrated significantly decreasing trends for the annual percentage change in ASMR and ASDR for GEC attributable to dietary carcinogens. The overall annual percentage change (net drift) was –5.95% (95% CI –6.25% to –5.65%) for dietary esophageal cancer mortality and –1.97% (95% CI –2.11% to –1.83%) for dietary stomach cancer mortality. Lastly, in 2044, dietary esophageal cancer deaths and DALYs were predicted to increase by 192.62% and 170.28%, respectively, due to age structure (121.58% and 83.29%), mortality change (76.81% and 92.43%), and population size (–5.77% and –5.44%). In addition, dietary stomach cancer deaths and DALYs were predicted to increase by 118.1% and 54.08%, with age structure, mortality rate change, and population size accounting for 96.71% and 53.99%, 26.17% and 3.97%, and –4.78% and –3.88% of the change, respectively.

Conclusions: Although the predicted age-standardized rates of mortality and DALYs due to dietary GEC show downward trends, the absolute numbers are still predicted to increase in the next 25 years due to rapid population aging in China.

(JMIR Public Health Surveill 2023;9:e48449) doi:10.2196/48449

KEYWORDS

esophageal cancer; gastric cancer; dietary carcinogens; prediction; temporal trends; China

Introduction

Gastric and esophageal cancer (GEC) are common cancers of the upper gastrointestinal tract, with estimates of 1.69 million new cases and 1.31 million deaths in 2020 worldwide, and 47.42% of them occur in China [1]. Primary preventive measures, such as limiting tobacco and alcohol use and maintaining a healthy lifestyle and diet, have important public health implications for reducing the disease burden of GEC and thus remain a high priority [2]. For dietary factors, habitually consuming fruits and vegetables is associated with a lower risk of esophageal cancer [3], and high sodium consumption is a well-established risk factor for gastric cancer [4].

Diet is considered an important modifiable factor for preventing cancer, and understanding and planning for the burden of cancers attributable to dietary carcinogens is of great public health significance. Thus, we searched PubMed and the China National Knowledge Infrastructure database with the terms (“gastric cancer”) OR (“stomach cancer”) AND (“esophagus cancer”) AND (“disease burden”) AND (“China”) AND (“diet”) OR (“dietary factors”) for previous studies published between January 1, 1990, and April 17, 2023, about trends and projections of deaths and disability-adjusted life years (DALYs) due to gastric and esophageal cancer attributable to dietary factors in China. We found previous research that reported current trends in the disease burden and attributable risk factors of gastric and esophageal cancer in China [5-7]. The research also predicted future changes and trends for the above two cancers separately [8-10]. However, there were no available data on trends and projections for the disease burden of dietary gastric and esophageal cancer in the Chinese population. Finally, summarizing a comprehensive time trend of the GEC burden caused by specific carcinogenic factors in China is of great significance for developing precise population prevention strategies.

The 2019 Global Burden of Disease (GBD) study was a systematic scientific effort to provide a unique resource on the burden of diseases across causes of disability and death in 204 countries and territories [11]; thus, it also provides a unique opportunity to assess and project long-term trends in the disease burden attributable to various risk factors. Our study aims to present recent trends in GEC deaths and DALYs attributable to dietary factors between 1990 and 2019 and short-term predictions through to the year 2044 in China. We also decompose the estimates in terms of population growth, population aging, and epidemiological changes. This study will provide useful information to optimize the allocation of health resources to reduce the dietary GEC burden in China.

Methods

Study Design and Time Frame

The data used in this study were derived from the 2019 GBD, specifically focusing on the disease burden of dietary esophageal and stomach cancers in China. The detailed methodology of the GBD estimation process has been fully described elsewhere [11]. In this study, we present the secular trends in gastric and esophageal cancer attributable to dietary carcinogens between 1990 and 2019 and also project the disease burden until 2044 in China.

Location and Population

In this study, we chose *China* from the database as the location, *esophageal cancer* and *stomach cancer* as the causes, *dietary risks* as the risk, and *death* and *DALYs* as the measures. Data were downloaded from the Global Health Data Exchange website [12].

Dietary Factors and Disease Information

A total of 11 dietary factors associated with 5 cancer types were assessed in the GBD. In this study, we focused on the disease burden of dietary GEC. For esophageal cancer, dietary factors included low consumption of fruits and vegetables; stomach cancer was only attributed to a diet high in sodium. The indices of disease burden included summary exposure values (SEV), deaths, DALYs, age-standardized mortality rate (ASMR), age-standardized DALY rate (ASDR), and population-attributable fraction (PAF). In addition, attributable deaths and DALYs were estimated using the total resulting death rate or DALYs multiplied by the PAF for the risk-outcome pair for age, sex, cause, and location. The SEV and the number, rate, and PAF for dietary GEC were also obtained directly from the Global Health Data Exchange website.

Statistical Analysis

Descriptive analyses were conducted and the temporal trends were expressed as annual percentage change (APC) and average annual percentage change (AAPC). Trends for age-standardized SEV, ASMR, ASDR, and age-standardized PAF (ASPAF) were further assessed by joinpoint regression analysis on a log scale since these measures followed a Poisson distribution.

Furthermore, we present the age period cohort results as follows [13]: (1) net drift, which represents the APC of the expected age-adjusted rates over time; (2) local drifts, which assess the age period cohort of the expected age-specific rates over time; (2) the longitudinal age curve, which shows the fitted longitudinal age-specific rates in the reference cohort, adjusted for period deviations; (3) period rate ratios (RRs), the ratio of age-specific rates in each period relative to the reference period; and (4) cohort RRs, the ratio of age-specific rates in each cohort relative to the reference cohort.

Then, we used a Bayesian age period cohort model of integrated nested Laplace approximations (with the R packages *BAPC* and *INLA*) to predict the number of deaths, DALYs, ASMR, and ASDR of dietary carcinogen–attributable GEC cases from 2019 to 2044 [14]. Briefly, we assumed the inverse gamma prior and applied a second-order random walk (RW2) to adjust for excessive dispersion [15]. Additionally, we obtained the estimated population of China from 2020 to 2050. This approach used inputs of population by age and sex in 2017, sex ratios at birth in 2017, and forecasted for age-specific fertility rates, age-specific mortality rates, and net migration for all locations through 2100. This analysis complied with the Guidelines on Accurate and Transparent Health Estimate Reporting [16,17].

Finally, we also used a recently developed decomposition method to attribute changes in the total number of GEC deaths and DALYs from dietary factors due to population growth, population aging, and age-specific changes between 2019 and 2044 [18,19]. Compared with previous studies [20-22], the recently developed method was insensitive to the decomposition order and the choice of reference group [19], enhancing the stability of our results. More details are provided in Multimedia Appendix 1 [13-22]. All statistical analyses were performed using R (version 3.6.0; R Core Team).

Ethical Considerations

This study used data from the GBD study, which was approved by the institutional review board of the University of Washington School of Medicine. The original data collection obtained informed consent from study participants or was granted exemptions by the institutional review board. As this is a secondary analysis of existing data, no additional human participant research ethics review or informed consent was required. Study data were anonymized and deidentified to protect the privacy and confidentiality of study participants. Our analysis complied with the Guidelines on Accurate and Transparent Health Estimate Reporting [16].

Results

Temporal Trends in GEC Attributable to Dietary Factors Between 1990 and 2019

As shown in Table 1, the SEV and age-standardized rates both decreased significantly from 1990 to 1999, with percentage changes of –0.06% (95% CI –0.11% to –0.02%) and –0.05% (95% CI –0.1% to –0.02%). The indices of cancer burden included the estimated number, all-age PAF and ASPAF, mortality, and DALYs. For dietary esophageal cancer, except for the estimated number of deaths, other indices showed a significantly decreasing trend. For dietary stomach cancer, although the percentage changes for the number of deaths increased significantly from 1990 to 2019 (0.36%, 95% CI 0.08% to 0.68%), significant decreases were observed for ASMR (–0.43%, 95% CI –0.55% to –0.31%) and ASDR (–0.47%, 95% CI –0.58% to –0.35%). There were no significant changes in other disease burden indices. More details are shown in Table 1.

Table 2 shows the results of the joinpoint regression model. Between 1990 and 2019, for esophageal cancer attributable to dietary carcinogens, the AAPC for ASMR and ASDR and the ASPAF for death rates and ASPAF for DALYs were –5.28% (95% CI –5.48% to 5.08%), –5.63% (95% CI –5.84% to –5.43%), –3.51% (95% CI –3.59% to –3.42%), and –3.58% (95% CI –3.67% to –3.5%), respectively. Furthermore, for dietary stomach cancer, the ASMR decreased by an average of 1.95% (–2.18% to –1.72%) per year. The average decline in ASDR was 2.21% (–2.46% to –1.96%). AAPCs in the ASPAF for death rates and ASPAF for DALYs were –0.04% (95% CI –0.04% to –0.03%) and –0.02% (95% CI –0.03% to –0.02%), respectively.

In addition, the declining trends in above 4 indices differed between dietary esophageal cancer and stomach cancer. For dietary esophageal cancer, although the magnitude of the decline in these 5 trends varied across different periods, all of them demonstrated a significantly decreasing trend. A rapid downward trend appeared during 2004 to 2014 in both ASMR (–9.38%, 95% CI –9.61% to –9.14%) and ASDR (–9.99%, 95% CI –10.22% to –9.75%). In general, the ASMR and ASDR for dietary stomach cancer showed declining trends. However, during 1998 to 2004, there were significant upward trends in ASMR (2.92%, 95% CI 2.37% to 3.46%) and ASDR (2.63%, 95% CI 2.15% to 3.12%). The ASPAF for both mortality and DALYs of dietary stomach cancer showed a slow and smooth decline.

Table 1. The summary exposure values and burden of esophageal cancer and stomach cancer attributable to dietary carcinogens.

Metric/measure	1990 values (95% CI)	2019 values (95% CI)	Percentage change (95% CI)
Summary exposure values			
All ages, %	82.76 (76.67 to 85.98)	77.67 (69.95 to 82.22)	–0.06 (–0.11 to –0.02)
Age-standardized, %	81.86 (75.44 to 85.15)	77.37 (69.58 to 81.93)	–0.05 (–0.10 to –0.02)
Esophageal cancer			
Deaths			
Deaths, n	40,514 (17,911 to 66,618)	20,509 (4338 to 52,321)	–0.49 (–0.84 to 0.01)
PAF[a] (all ages), %	0.23 (0.10 to 0.36)	0.08 (0.02 to 0.20)	–0.65 (–0.89 to –0.35)
ASPAF,[b] %	0.23 (0.10 to 0.36)	0.08 (0.02 to 0.21)	–0.64 (–0.88 to –0.34)
Mortality (all ages) $1/10^5$	3.42 (1.51 to 5.63)	1.44 (0.31 to 3.68)	–0.57 (–0.87 to –0.16)
ASMR,[c] $1/10^5$	5.07 (2.23 to 8.32)	1.07 (0.24 to 2.73)	–0.79 (–0.93 to –0.58)
DALYs[d]			
DALYs, n	1,034,489 (449,013 to 1,696,655)	455,090 (90,540 to 1,163,817)	–0.56 (–0.87 to –0.08)
PAF (all ages), %	0.23 (0.10 to 0.36)	0.08 (0.02 to 0.20)	–0.66 (–0.89 to –0.35)
ASPAF, %	0.23 (0.10 to 0.36)	0.08 (0.02 to 0.21)	–0.65 (–0.89 to –0.35)
All-age rate, $1/10^5$	87.40 (37.93 to 143.34)	32.00 (6.37 to 81.82)	–0.63 (–0.89 to –0.35)
ASDR,[e] $1/10^5$	116.66 (51.12 to 190.9)	22.22 (4.54 to 56.38)	–0.81 (–0.94 to –0.61)
Stomach cancer			
Deaths			
Numbers, n	27,227 (613 to 101,649)	37,131 (833 to 138,479)	0.36 (0.08 to 0.68)
PAF (all ages), %	0.09 (0.01 to 0.33)	0.09 (0.01 to 0.33)	–0.01 (–0.12 to 0.01)
ASPAF, %	0.09 (0.01 to 0.33)	0.09 (0.01 to 0.33)	–0.01 (–0.11 to 0.01)
Mortality (all ages), $1/10^5$	2.30 (0.05 to 8.59)	2.61 (0.06 to 9.74)	0.13 (–0.10 to 0.4)
ASMR, $1/10^5$	3.34 (0.08 to 12.56)	1.90 (0.04 to 7.12)	–0.43 (–0.55 to –0.31)
DALYs			
DALYs, n	734,448 (16,388 to 2,731,936)	873,813 (19,283 to 3,220,231)	0.19 (–0.06 to 0.48)
PAF (all ages), %	0.09 (0.01 to 0.33)	0.09 (0.01 to 0.33)	0.01 (–0.10 to 0.01)
ASPAF, %	0.09 (0.01 to 0.33)	0.09 (0.01 to 0.33)	–0.01 (–0.09 to 0.01)
All-age rate, $1/10^5$	62.05 (1.38 to 230.80)	61.43 (1.36 to 226.40)	–0.01 (–0.22 to 0.23)
ASDR, $1/10^5$	80.72 (1.81 to 301.47)	42.52 (0.94 to 157.03)	–0.47 (–0.58 to –0.35)

[a] PAF: population-attributable fraction.

[b] ASPAF: age-standardized population-attributable fraction.

[c] ASMR: age-standardized mortality rate.

[d] DALY: disability-adjusted life year.

[e] ASDR: age-standardized disability-adjusted life year rate.

Table 2. Joinpoint analysis of trends in disease burden of dietary esophageal cancer and stomach cancer in China, 1990 to 2019.

Metric	Trend 1		Trend 2		Trend 3		Trend 4		Trend 5		1990-2019
	Period	APC,[a] % (95% CI)	Period	APC, % (95% CI)	Period	APC, % (95% CI)	Period	APC, % (95% CI)	Period	APC, % (95% CI)	AAPC,[b] % (95% CI)
Age-standardized mortality rate											
Esophageal cancer	1990-1993	–2.56 (–3.38 to –1.73)	1993-1998	–3.9 (–4.47 to –3.32)	1998-2004	–1.91 (–2.41 to –1.42)	2004-2014	–9.38 (–9.61 to –9.14)	2014-2019	–3.87 (–4.56 to –3.18)	–5.28 (–5.48 to –5.08)
Stomach cancer	1990-1998	–2.10 (–2.36 to –1.85)	1998-2004	2.92 (2.37 to 3.46)	2004-2011	–4.02 (–4.41 to –3.64)	2011-2015	–4.43 (–5.56 to –3.29)	2015-2019	–2.59 (–3.32 to –1.85)	–1.95 (–2.18 to –1.72)
Age-standardized population attributable proportion of death rates											
Esophageal cancer	1990-1995	–2.33 (–2.45 to –2.21)	1995-2002	–3.79 (–3.90 to –3.69)	2002-2010	–4.63 (–4.74 to –4.52)	2010-2014	–3.78 (–4.27 to –3.29)	2014-2019	–2.24 (–2.48 to –2.00)	–3.51 (–3.59 to –3.42)
Stomach cancer	1990-1995	–0.06 (–0.07 to –0.05)	1995-1999	–0.09 (–0.11 to –0.07)	1999-2010	–0.03 (–0.03 to –0.03)	2010-2013	0.00 (–0.04 to 0.04)	2013-2019	–0.02 (–0.02 to –0.01)	–0.04 (–0.04 to –0.03)
Age-standardized DALY[c] rate											
Esophageal cancer	1990-1993	–2.89 (–3.73 to –2.04)	1993-1998	–4.20 (–4.79 to –3.60)	1998-2004	–2.18 (–2.69 to –1.67)	2004-2014	–9.99 (–10.22 to –9.75)	2014-2019	–3.83 (–4.55 to –3.11)	–5.63 (–5.84 to –5.43)
Stomach cancer	1990-1995	–2.50 (–2.96 to –2.03)	1995-1998	–1.99 (–4.04 to 0.10)	1998-2004	2.63 (2.15 to 3.12)	2004-2015	–4.52 (–4.68 to –4.37)	2015-2019	–2.67 (–3.32 to –2.02)	–2.21 (–2.46 to –1.96)
Age-standardized population attributable proportion of DALY rates											
Esophageal cancer	1990-1995	–2.43 (–2.55 to –2.31)	1995-2002	–3.96 (–4.07 to –3.85)	2002-2010	–4.78 (–4.89 to –4.67)	2010-2014	–3.79 (–4.29 to –3.29)	2014-2019	–2.10 (–2.34 to –1.86)	–3.58 (–3.67 to –3.50)
Stomach cancer	1990-2000	–0.04 (–0.04 to –0.04)	2000-2004	0.02 (0.00 to 0.04)	2004-2010	–0.03 (–0.03 to –0.02)	2010-2013	–0.01 (–0.04 to 0.03)	2013-2019	–0.03 (–0.03 to –0.02)	–0.02 (–0.03 to –0.02)

[a] APC: annual percentage change.

[b] AAPC: average annual percentage change.

[c] DALY: disability-adjusted life year.

Effects of Age, Time Period, and Cohort on Cancer Burden Between 1990 and 2019

Figure 1 demonstrates the age-period-cohort analysis results on deaths and DALYs from dietary esophageal cancer. The ranges for age and period were 25 to ≥95 years and 1990 to 2019, respectively. For dietary esophageal cancer deaths, the overall APC (net drift) was –5.95% (95% CI –6.25% to –5.65%). According to the results for local drifts and net drift (Figure 1A), the data that had a percentage value below 0 suggests that there was a significantly decreasing trend in all age groups. The longitudinal age curve for dietary esophageal cancer mortality showed that the rates increased from the age of 25 years and reached the highest level at about the age of 57 years, then gradually decreased with age (Figure 1B). In addition, there were persistent downward trends in the estimated period and cohort RRs, as illustrated in Figure 1C and D. Similar patterns were observed for the age, time period, and cohort effects on the DALYs of esophageal cancer associated with dietary carcinogens (Figure 1E-H).

The results of age period cohort models on death and DALYs from dietary stomach cancer are shown in Figure 2. In general, the net drift was –1.97% (95% CI –2.11% to –1.83%) for dietary stomach cancer mortality. The results for local drifts and net drift show that the percentage declined from the age of 25 years and reached the lowest level at about the age of 43 years, then gradually increased with age; the peak point occurred at about the age of 93 years, after which the percentage decreased with age (Figure 2A). For the longitudinal age curve, the maximum value occurred at about the age of 88 years, after which there was a short and rapid declining trend (Figure 2B). As illustrated in Figure 2C, the time-period RRs decreased from 1992 to 1997, increased until 2002, and then showed a continuous downward trend. Finally, there was a persistent downward trend in the estimated cohort RRs (Figure 2D). Similar patterns were observed for the age period cohort analysis of the DALYs for dietary stomach cancer (Figure 2E to H).

Figure 1. Effects of age, time period, and cohort on deaths and DALYs due to dietary esophageal cancer from 1990 to 2019 in China. Local drifts with net drift values of (A) deaths and (E) DALYs; fitted longitudinal age curves of (B) deaths and (F) DALYs; relative risks of (C) death and (G) DALYs in each period compared with the reference period (2000-2004); and relative risks of (D) deaths and (H) DALYs in each cohort compared with the reference cohort (1940-1944 cohort). DALY: disability-adjusted life year.

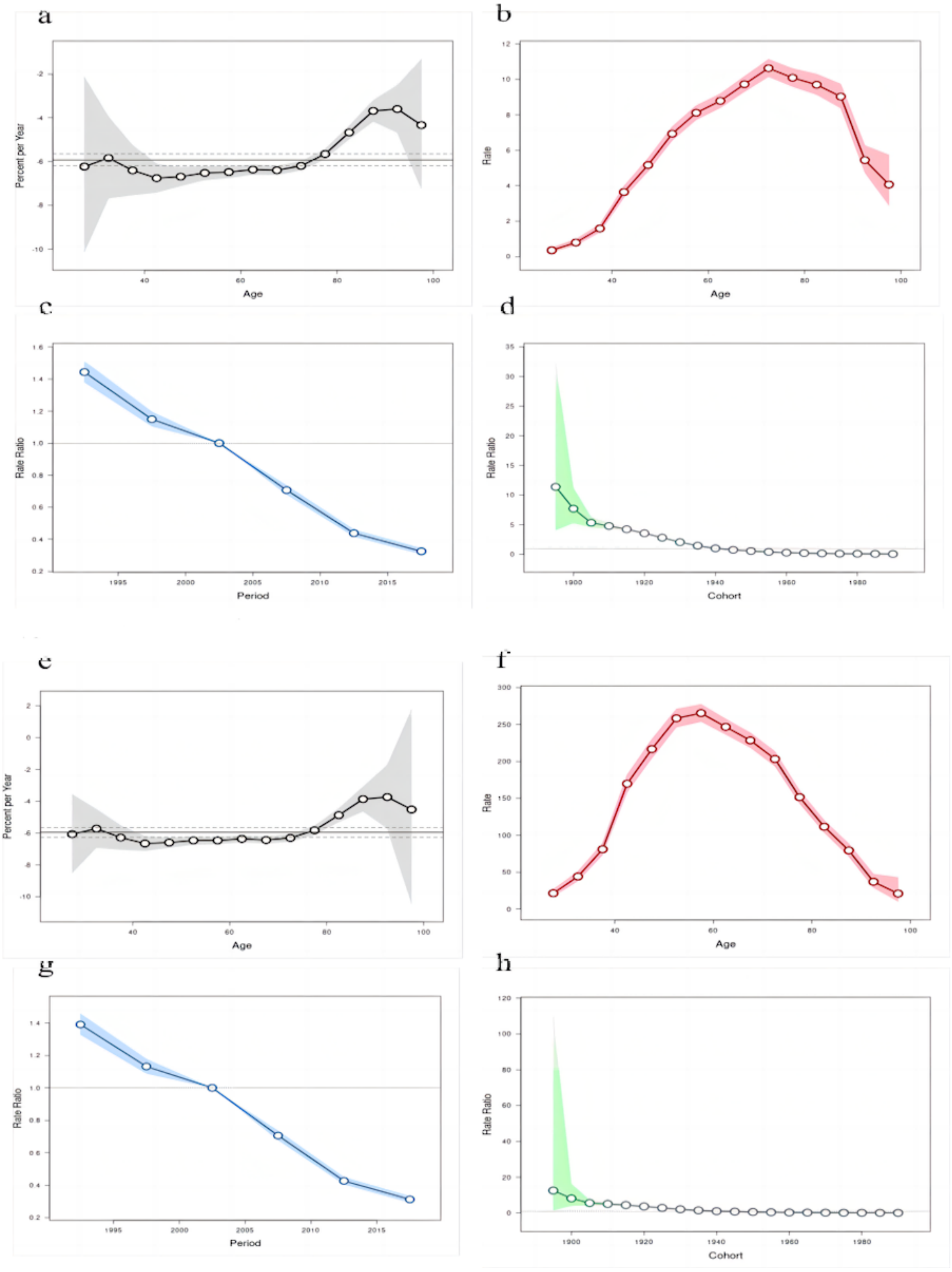
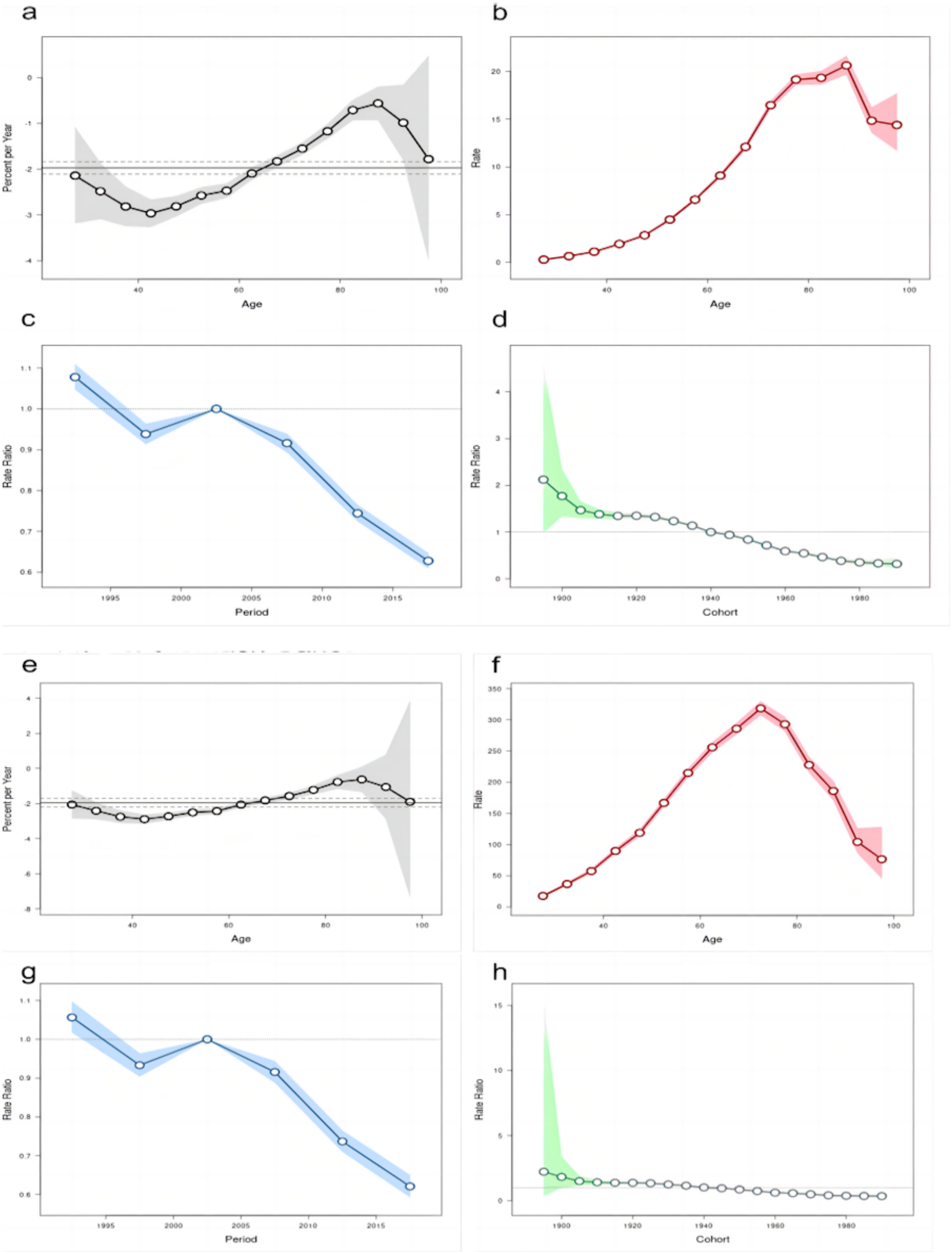


Figure 2. Effects of age, time period, and cohort on deaths and DALYs due to dietary stomach cancer from 1990 to 2019 in China. Local drifts with net drift values of (A) deaths and (E) DALYs; fitted longitudinal age curves of (B) deaths and (F) DALYs; relative risks of (C) death and (G) DALYs in each period compared with the reference period (2000-2004); and relative risks of (D) deaths and (H) DALYs in each cohort compared with the reference cohort (1940-1944 cohort). DALY: disability-adjusted life year.



Predictions and Changes Between 2019 and 2044

Figure 3 shows the future trends for the disease burden of GEC attributable to dietary carcinogens from 2020 to 2044 in China. For dietary esophageal cancer, there are predicted to be gradually decreasing trends in the ASMR (Figure 3A) and ASDR (Figure 3B). For dietary stomach cancer, the ASMR is predicted to decrease from 1990 to 1999, increase until 2004, and then show a continuous downward trend (Figure 3C). Similar patterns are seen in the ASDR (Figure 3D).

Figure 3. Temporal trends of for ASMR and ASDR for esophageal cancer and stomach cancer attributable to dietary carcinogens between 1990 and 2044 in China. The dotted lines represent the observational values from the Global Burden of Disease data set. The predictive mean values are shown as black solids, and the fan is the predictive distribution between the 5% and 95% quintiles. (A) ASMR for esophageal cancer; (B) ASMR for stomach cancer; (C) ASDR for esophageal cancer; (D) ASDR for stomach cancer. ASDR: age-standardized disability-adjusted life year rate; ASMR: age-standardized mortality rate; DALY: disability-adjusted life year.

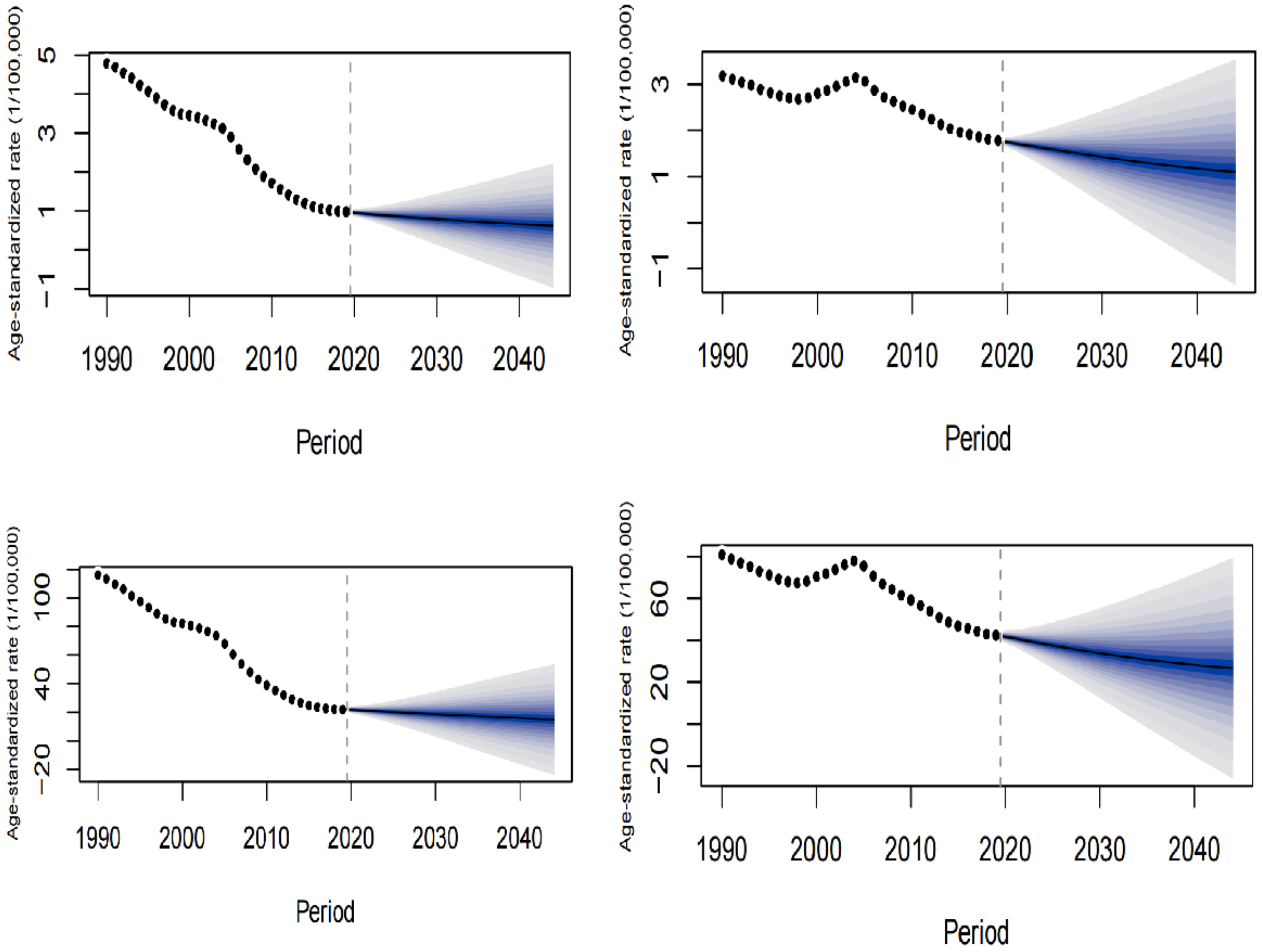


Table 3 presents decomposition analysis results for the disease burden of dietary esophageal and stomach cancer. In 2044, dietary esophageal cancer ASMR was predicted to increase by 192.62% compared with 2019, which was due to age structure (121.58%), mortality change (76.81%), and population size (–5.77%). Dietary esophageal cancer ASDR was predicted to increase by 170.28% because of age structure (83.29%), mortality change (92.43%), and population size (–5.44%). Furthermore, dietary stomach cancer ASMR in 2044 was predicted to increase by 118.1%, with age structure, mortality change and population size contributing 96.71%, 26.17%, and –4.78%, respectively. Dietary stomach cancer ASDR was predicted to increase by 54.08%, with age structure, mortality rate changes, and population size accounting for 53.99%, 3.97%, and –3.88%, respectively.

Table 3. Predicted number and contribution changes for dietary esophageal cancer and stomach cancer deaths and disability-adjusted life years from 1990 to 2044.

	Values, n		Changes between 2019 and 2044, %				
	2019	2044	Total change	Change due to mortality change	Change due to demographics	Change due to age structure	Change due to population size
Deaths							
Esophageal cancer	20,492	59,965	192.62	76.81	115.81	121.58	–5.77
Stomach cancer	37,112	80,942	118.1	26.17	91.93	96.71	–4.78
Disability-adjusted life years							
Esophageal cancer	455,089	1,230,011	170.28	92.43	77.85	83.29	–5.44
Stomach cancer	873,813	1,346,403	54.08	3.97	50.11	53.99	–3.88

Discussion

The GEC burden attributable to dietary risk factors in China from 1990 to 2044 is systematically summarized in this study. Deaths and DALYs due to dietary esophageal cancer declined by 49% and 56%, respectively, but deaths and DALYs due to dietary stomach cancer grew by 36% and 19%, respectively, over the last 30 years. For esophageal and stomach cancers attributable to dietary carcinogens, although both the ASMR and the ASDR are predicted to decline in the coming 25 years, the numbers of both deaths and DALYs are predicted to continue to increase. The results of decomposition methods supported the finding that the aging population is the main reason for future changes in the GEC burden.

Dietary modifications are an effective strategy to reduce the GEC burden in terms of deaths and DALYs. In our study, dietary carcinogenic factors were a diet low in fruits and vegetables for esophageal cancer and a diet high in sodium for stomach cancer. Epidemiological studies have found that low intake of fresh fruits and vegetables is associated with a higher risk of esophageal squamous cell carcinoma, which is the dominant histological subtype in China [23]. In addition, accumulating evidence suggests that people who prefer salty food and salt-preserved meat and fish generally have a higher risk of gastric cancer [24]. The World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) report also supports the strong association between high-salt foods and gastric cancer [25]. However, the SEV for dietary carcinogens from 1990 to 2019 only slightly declined, from 82.76% to 77.67%, which is consistent with previous findings from the Chinese Health and Nutrition Survey (CHNS). Nevertheless, data from the CHNS show that intake did not meet the recommended levels (an average of 269.4 g of fresh vegetables and 40.7 g of fruit per day in 2012 vs recommend daily intakes of 300-500 g of vegetables and 200-350 g of fruit for adults in 2016) [26-28]. According to the latest data from the 2020 Report on Chinese Residents' Chronic Diseases and Nutrition, the consumption levels of fruits and vegetables were lower than the recommended intake levels, while the intake of salt was much higher than the standard [29]. These unhealthy dietary habits might be related to a lower awareness of basic knowledge of cancer in China, especially the level of knowledge of primary prevention [30]. Thus, our findings suggest that people should still attach importance to a nutritionally balanced diet and the primary prevention of cancer in the future.

In general, the temporal trends for deaths and DALYs for dietary GEC decreased from 1990 to 2019, even though these trends fluctuated in different time segments over the last 30 years. The joinpoint method describes short-term trends more accurately by dividing the long-term trend line into several segments and analyzing each segment. For dietary stomach cancer, the proportion of deaths and DALYs remained stable from 1990 to 2019, which is also consistent with a previous study [10]. But deaths and DALY rates for stomach cancer attributable to dietary carcinogens declined over the past 30 years, which is similar to the changes for stomach cancer in China reported by other studies [8,9]. This might be related to the increasing coverage of population-based screening projects. From 2001 to 2020, a total of 5 cancer screening programs were implemented in high-risk areas of China focusing on screening for stomach cancer [31]. Notably, a significantly increasing trend of dietary stomach cancer deaths and DALY rates occurred in the period from 1998 to 2004, which might be explained by the establishment of the national cancer registry system and the improvement of data quality. In 1982, the Reporting Manual of Cancer Registration was first released by the Office of National Central Cancer Registry. In 2002, the National Central Cancer Registry was launched under the supervision of the former Ministry of Health and was responsible for systematically collecting, analyzing, and interpreting cancer data [23]. Moreover, esophageal and gastric cancer have always been treated as upper gastrointestinal tumors for primary and secondary prevention; thus, initial gastric registry data might derive in part from upper gastrointestinal tumor data.

Our findings show that the rates and percentages of deaths and DALYs due to dietary esophageal cancer declined in China over the past 30 years, which is similar to the trends for esophageal cancer associated with all risk factors [2,5,6]. A previous study also showed that the burden of esophageal cancer associated with a low-fruit diet declined in China from 1990 to 2017 [7]; however, there are no prior studies focusing on low-vegetable diets. These declining trends were associated with long-term, extensive efforts in esophageal cancer etiological prevention and screening in China. Chinese researchers launched a series of prevention projects starting in 1985, including the Linxian general population nutrition intervention trial, a dysplasia population nutrition intervention trial, a chemoprevention trial [32], and a cancer screening program in high-risk areas [31], because the bulk of the burden of esophageal cancer worldwide comes from East Asia, particularly from China [24]. Considering the success of dietary interventions, future efforts should focus on improving dietary habits to reduce the burden of GEC.

We further revealed trends for age, period, and cohort effects for deaths and DALYs due to dietary GEC. Regarding the age effect, the top points of the longitudinal age curves for deaths and DALYs for esophageal cancer were 60 years and 72 years, which is similar to the peak ages for deaths and incidence [10]. For dietary stomach cancer deaths and DALYs, the peak of the age curves shifted to 85 years and 70 years, which is also consistent with the highest death and incidence ages reported in a previous study [6]. These age-specific trends in both esophageal and stomach cancer related to dietary carcinogens might be due to both age-related biological factors and the increased level of dietary exposure, as explained by a similar study of occupational lung cancer [33]. Furthermore, the period effect reflected changes in social and economic policy, as well as the impact of major events at specific points in time [34]. In this study, we found that the change trends in period RRs for dietary GEC mortality and DALYs were similar to the results of the joinpoint analysis and were explained by the implementation of the cancer registry system, primary prevention, and screening programs, as mentioned above. Lastly, birth cohort effects represent the influence of physical and social exposures that appear earlier in the life process and accumulate with time [35]. In our study, the cohort RRs for dietary carcinogens associated with GEC mortality and DALYs showed gradually decreasing trends for cohorts born since 1895.

We first projected the trends of dietary GEC death and DALYs through to 2044 and used a recently developed decomposition method to attribute changes to population growth, population aging, and age-specific changes [18,19]. Previous studies predicted the burden of GEC attributable to all risk factors and found that there was a persistently decreasing trend in terms of ASMR and ASDR, but the number of new cases was expected to increase in the next decades [8-10,36,37]. Moreover, population aging contributed to most of the additional dietary GEC deaths and DALYs, which could be explained by the continuous increase of the older population and the higher incidence in this age group. As shown in our study, change due to risk factors also deserves attention, as it contributes to increased disease burden. Therefore, considering the persistent increase in case numbers, dietary GEC might be one of the main public health concerns in the near future in China, and it is necessary to increase the promotion of healthy diets.

Our study has several limitations, as mentioned in earlier GBD study reports [38,39]. First, the main histological subtypes of GEC (esophageal squamous cell carcinoma and esophageal adenocarcinoma, cardia and noncardia stomach cancer) have distinct incidence trends and geographical distributions [40,41], but data for these subtypes are not currently captured independently in the GBD database. Second, although other dietary factors, such as high-sugar drinks and alcohol, are recognized to be related to deaths and DALYs due to GEC in China, the disease burden caused by these dietary carcinogens was not estimated [42]. Lastly, we only estimated the burden of GEC attributable to dietary factors; the combined effects of dietary and other risk factors may increase or complicate the burden of GEC [43,44]. Despite these limitations, using the most up-to-date data and advanced modeling strategies, our study is the first to provide a comprehensive understanding of the burden of dietary GEC from 1990 to 2044.

Although the predicted ASMR and ASDR for dietary GEC generally show downward trends, absolute numbers will still increase in the next 25 years due to the rapid population aging seen in China. China still needs to strengthen the targeted interventions and address modifiable risk factors, such as advocating for a balanced diet.

Acknowledgments

We appreciate the work done by the 2019 Global Burden of Disease study collaborators. This work was supported by Hebei Province Medical Science Research Funding (20210943).

Data Availability

The data sets generated for this study can be found in the Global Burden of Disease repository [12].

Authors' Contributions

HXG and BHY contributed to conceptualization and design, QW and QCY contributed to data extraction, HXG and QW contributed to formal analysis, HXG and QW contributed to writing of the original draft, and QCY and HZH contributed to review and editing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental information on statistical analysis.
[DOCX File , 49 KB - publichealth_v9i1e48449_app1.docx]

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Fan J, Liu Z, Mao X, Tong X, Zhang T, Suo C, et al. Global trends in the incidence and mortality of esophageal cancer from 1990 to 2017. Cancer Med 2020 Sep;9(18):6875-6887 [FREE Full text] [doi: 10.1002/cam4.3338] [Medline: 32750217]
3. Qin X, Jia G, Zhou X, Yang Z. Diet and esophageal cancer risk: An umbrella review of systematic reviews and meta-analyses of observational studies. Adv Nutr 2022 Dec 22;13(6):2207-2216 [FREE Full text] [doi: 10.1093/advances/nmac087] [Medline: 36041184]
4. Thrift AP, Wenker TN, El-Serag HB. Global burden of gastric cancer: epidemiological trends, risk factors, screening and prevention. Nat Rev Clin Oncol 2023 May;20(5):338-349. [doi: 10.1038/s41571-023-00747-0] [Medline: 36959359]
5. Li R, Sun J, Wang T, Huang L, Wang S, Sun P, et al. Comparison of secular trends in esophageal cancer mortality in China and Japan during 1990-2019: An age-period-cohort analysis. Int J Environ Res Public Health 2022 Aug 18;19(16):10302 [FREE Full text] [doi: 10.3390/ijerph191610302] [Medline: 36011937]

6. He Y, Wang Y, Luan F, Yu Z, Feng H, Chen B, et al. Chinese and global burdens of gastric cancer from 1990 to 2019. Cancer Med 2021 May;10(10):3461-3473 [FREE Full text] [doi: 10.1002/cam4.3892] [Medline: 33931958]
7. Yang S, Lin S, Li N, Deng Y, Wang M, Xiang D, et al. Burden, trends, and risk factors of esophageal cancer in China from 1990 to 2017: an up-to-date overview and comparison with those in Japan and South Korea. J Hematol Oncol 2020 Nov 02;13(1):146. [doi: 10.1186/s13045-020-00981-4] [Medline: 33138852]
8. Zhou J, Zheng R, Zhang S, Chen R, Wang S, Sun K, et al. Gastric and esophageal cancer in China 2000 to 2030: Recent trends and short-term predictions of the future burden. Cancer Med 2022 Apr;11(8):1902-1912 [FREE Full text] [doi: 10.1002/cam4.4586] [Medline: 35148032]
9. Yang X, Zhang T, Zhang H, Sang S, Chen H, Zuo X. Temporal trend of gastric cancer burden along with its risk factors in China from 1990 to 2019, and projections until 2030: comparison with Japan, South Korea, and Mongolia. Biomark Res 2021 Nov 16;9(1):84 [FREE Full text] [doi: 10.1186/s40364-021-00340-6] [Medline: 34784961]
10. Li S, Chen H, Man J, Zhang T, Yin X, He Q, et al. Changing trends in the disease burden of esophageal cancer in China from 1990 to 2017 and its predicted level in 25 years. Cancer Med 2021 Mar;10(5):1889-1899 [FREE Full text] [doi: 10.1002/cam4.3775] [Medline: 33586344]
11. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
12. GBD Results. VizHub. URL: https://ghdx.healthdata.org/gbd-results-tool [accessed 2023-09-05]
13. Tzeng I, Chen J. Exploring hepatocellular carcinoma mortality using weighted regression estimation for the cohort effect in Taiwan from 1976 to 2015. Int J Environ Res Public Health 2022 May 04;19(9):5573 [FREE Full text] [doi: 10.3390/ijerph19095573] [Medline: 35564966]
14. Barboza GE. The geography of child maltreatment: A spatiotemporal analysis using Bayesian hierarchical analysis with integrated nested Laplace approximation. J Interpers Violence 2019 Jan;34(1):50-80. [doi: 10.1177/0886260516639583] [Medline: 27036155]
15. Riebler A, Held L. Projecting the future burden of cancer: Bayesian age-period-cohort analysis with integrated nested Laplace approximations. Biom J 2017 May;59(3):531-549. [doi: 10.1002/bimj.201500263] [Medline: 28139001]
16. Stevens GA, Alkema L, Black RE, Boerma JT, Collins GS, Ezzati M, (The GATHER Working Group). Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement. Lancet 2016 Dec 10;388(10062):e19-e23 [FREE Full text] [doi: 10.1016/S0140-6736(16)30388-9] [Medline: 27371184]
17. Vollset SE, Goren E, Yuan C, Cao J, Smith AE, Hsiao T, et al. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. Lancet 2020 Oct 17;396(10258):1285-1306 [FREE Full text] [doi: 10.1016/S0140-6736(20)30677-2] [Medline: 32679112]
18. Cheng X, Yang Y, Schwebel DC, Liu Z, Li L, Cheng P, et al. Population ageing and mortality during 1990-2017: A global decomposition analysis. PLoS Med 2020 Jun;17(6):e1003138 [FREE Full text] [doi: 10.1371/journal.pmed.1003138] [Medline: 32511229]
19. Cheng X, Tan L, Gao Y, Yang Y, Schwebel DC, Hu G. A new method to attribute differences in total deaths between groups to population size, age structure and age-specific mortality rate. PLoS One 2019;14(5):e0216613 [FREE Full text] [doi: 10.1371/journal.pone.0216613] [Medline: 31075117]
20. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1736-1788 [FREE Full text] [doi: 10.1016/S0140-6736(18)32203-7] [Medline: 30496103]
21. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: A systematic analysis for the Global Burden of Disease study. JAMA Oncol 2017 Apr 01;3(4):524-548 [FREE Full text] [doi: 10.1001/jamaoncol.2016.5688] [Medline: 27918777]
22. GBD 2015 MortalityCauses of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 08;388(10053):1459-1544 [FREE Full text] [doi: 10.1016/S0140-6736(16)31012-1] [Medline: 27733281]
23. Maomao C, He L, Dianqin S, Siyi H, Xinxin Y, Fan Y, et al. Current cancer burden in China: epidemiology, etiology, and prevention. Cancer Biol Med 2022 Aug 30;19(8):1121-1138 [FREE Full text] [doi: 10.20892/j.issn.2095-3941.2022.0231] [Medline: 36069534]
24. GBD 2017 Oesophageal Cancer Collaborators. The global, regional, and national burden of oesophageal cancer and its attributable risk factors in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol 2020 Jun;5(6):582-597 [FREE Full text] [doi: 10.1016/S2468-1253(20)30007-8] [Medline: 32246941]
25. Clinton SK, Giovannucci EL, Hursting SD. The World Cancer Research Fund/American Institute for Cancer Research Third Expert Report on Diet, Nutrition, Physical Activity, and Cancer: impact and future directions. J Nutr 2020 Apr 01;150(4):663-671 [FREE Full text] [doi: 10.1093/jn/nxz268] [Medline: 31758189]

26. Li L, Ouyang Y, Wang H, Huang F, Wang Y, Zhang J, et al. Disparities in fresh fruit and vegetable intake by sociodemographic and behavioural factors among adults in China. Public Health Nutr 2022 Mar;25(3):649-656 [FREE Full text] [doi: 10.1017/S1368980020003274] [Medline: 32924898]
27. Li YC, Jiang B, Zhang M, Huang ZJ, Deng Q, Zhou MG, et al. Vegetable and fruit consumption among Chinese adults and associated factors: A nationally representative study of 170,847 adults. Biomed Environ Sci 2017 Dec;30(12):863-874 [FREE Full text] [doi: 10.3967/bes2017.117] [Medline: 29335056]
28. Morris ME, Aguilera A. Mobile, social, and wearable computing and the evolution of psychological practice. Prof Psychol Res Pr 2012 Dec;43(6):622-626 [FREE Full text] [doi: 10.1037/a0029041] [Medline: 25587207]
29. National Health and Family Planning Commission of People's Republic of China. Report on Nutrition and Chronic Diseases in China. Acta Nutrimenta Sinica 2020;42(6):521.
30. He S, Li H, Cao M, Sun D, Yang F, Yan X, et al. Geographic, demographic, and socioeconomic disparities and factors associated with cancer literacy in China: National cross-sectional study. JMIR Public Health Surveill 2023 Feb 17;9:e43541 [FREE Full text] [doi: 10.2196/43541] [Medline: 36800218]
31. Cao M, Li H, Sun D, He S, Yu Y, Li J, et al. Cancer screening in China: The current status, challenges, and suggestions. Cancer Lett 2021 May 28;506:120-127. [doi: 10.1016/j.canlet.2021.02.017] [Medline: 33684533]
32. Wang S, Abnet CC, Qiao Y. What have we learned from Linxian esophageal cancer etiological studies? Thorac Cancer 2019 May;10(5):1036-1042 [FREE Full text] [doi: 10.1111/1759-7714.13058] [Medline: 30925028]
33. Fan Y, Jiang Y, Li X, Li X, Li Y, Wu H, et al. Burden of lung cancer attributable to occupational carcinogens from 1990 to 2019 and projections until 2044 in China. Cancers (Basel) 2022 Aug 11;14(16):3883 [FREE Full text] [doi: 10.3390/cancers14163883] [Medline: 36010878]
34. Okui T. An age-period-cohort analysis for prevalence of common psychiatric disorders in Japan, 1999-2017. Soc Psychiatry Psychiatr Epidemiol 2021 Apr;56(4):639-648 [FREE Full text] [doi: 10.1007/s00127-020-01988-5] [Medline: 33221927]
35. Janssen F, Kunst AE, Netherlands EpidemiologyDemography Compression of Morbidity research group. Cohort patterns in mortality trends among the elderly in seven European countries, 1950-99. Int J Epidemiol 2005 Oct;34(5):1149-1159. [doi: 10.1093/ije/dyi123] [Medline: 15964911]
36. Morgan E, Arnold M, Camargo MC, Gini A, Kunzmann AT, Matsuda T, et al. The current and future incidence and mortality of gastric cancer in 185 countries, 2020-40: A population-based modelling study. EClinicalMedicine 2022 May;47:101404 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101404] [Medline: 35497064]
37. Li Y, Xu J, Gu Y, Sun X, Dong H, Chen C. The disease and economic burdens of esophageal cancer in China from 2013 to 2030: Dynamic cohort modeling study. JMIR Public Health Surveill 2022 Mar 02;8(3):e33191 [FREE Full text] [doi: 10.2196/33191] [Medline: 34963658]
38. Yuan R, Ren F, Xie Y, Li K, Tong Z. The global, regional, and national burdens of cervical cancer attributable to smoking from 1990 to 2019: Population-based study. JMIR Public Health Surveill 2022 Dec 23;8(12):e40657 [FREE Full text] [doi: 10.2196/40657] [Medline: 36563035]
39. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
40. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev 2014 May;23(5):700-713 [FREE Full text] [doi: 10.1158/1055-9965.EPI-13-1057] [Medline: 24618998]
41. Kamangar F, Chow W, Abnet CC, Dawsey SM. Environmental causes of esophageal cancer. Gastroenterol Clin North Am 2009 Mar;38(1):27-57, vii [FREE Full text] [doi: 10.1016/j.gtc.2009.01.004] [Medline: 19327566]
42. Tay SW, Li JW, Fock KM. Diet and cancer of the esophagus and stomach. Curr Opin Gastroenterol 2021 Mar 01;37(2):158-163. [doi: 10.1097/MOG.0000000000000700] [Medline: 33315794]
43. Kim SA, Kwak JH, Eun CS, Han DS, Kim YS, Song KS, et al. Association of dietary antioxidant vitamin intake and gastric cancer risk according to smoking status and histological subtypes of gastric cancer: A case-control study in Korea. Nutr Cancer 2023;75(2):652-661. [doi: 10.1080/01635581.2022.2147274] [Medline: 36453620]
44. Ghosh NR, Jones LA. Dietary risk factors for esophageal cancer based on World Health Organization regions. Nutrition 2022 Mar;95:111552. [doi: 10.1016/j.nut.2021.111552] [Medline: 34999383]

Abbreviations

AAPC: average annual percentage change
APC: annual percentage change
ASDR: age-standardized disability-adjusted life year rate
ASMR: age-standardized mortality rate
ASPAF: age-standardized population-attributable fraction
CHNS: Chinese Health and Nutrition Survey
DALY: disability-adjusted life year

GBD: Global Burden of Disease
GEC: gastric and esophageal cancer
PAF: population-attributable fraction
RR: rate ratio
RW2: second-order random walk
SEV: summary exposure value
WCRF/AICR: World Cancer Research Fund/American Institute for Cancer Research

---

Edited by A Mavragani, T Sanchez; submitted 24.04.23; peer-reviewed by Z Su, J Luo, M Cao; comments to author 31.05.23; revised version received 11.06.23; accepted 10.08.23; published 21.09.23.

Please cite as:
Guo HX, Wang Q, Wang C, Yin QC, Huo HZ, Yin BH
Secular Trends in Gastric and Esophageal Cancer Attributable to Dietary Carcinogens From 1990 to 2019 and Projections Until 2044 in China: Population-Based Study
JMIR Public Health Surveill 2023;9:e48449
URL: https://publichealth.jmir.org/2023/1/e48449
doi:10.2196/48449
PMID:37560940

Original Paper

Evaluation of Primary Allied Health Care in Patients Recovering From COVID-19 at 6-Month Follow-up: Dutch Nationwide Prospective Cohort Study

Anne I Slotegraaf[1*], MSc; Marissa H G Gerards[2,3*], PhD; Arie C Verburg[4], PhD; Marian A E de van der Schueren[1,5], PhD; Hinke M Kruizenga[6], PhD; Maud J L Graff[4,7], PhD; Edith H C Cup[7], PhD; Johanna G Kalf[7], PhD; Antoine F Lenssen[3], PhD; Willemijn M Meijer[8], PhD; Renée A Kool[9]; Rob A de Bie[2], PhD; Philip J van der Wees[4,7], PhD; Thomas J Hoogeboom[4], PhD; Dutch Consortium Allied Healthcare COVID-19[10*]

[1]Division of Human Nutrition and Health, Wageningen University and Research, Wageningen, Netherlands

[2]Department of Epidemiology, Care and Public Health Institute, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

[3]Department of Physical Therapy, Maastricht University Medical Centre, Maastricht, Netherlands

[4]IQ Healthcare, Radboud Institute for Health Sciences, Radboud University Medical Centre, Nijmegen, Netherlands

[5]Department of Nutrition, Dietetics and Lifestyle, HAN University of Applied Sciences, Nijmegen, Netherlands

[6]Department of Nutrition and Dietetics, Amsterdam University Medical Centers, Amsterdam, Netherlands

[7]Department of Rehabilitation, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, Netherlands

[8]Netherlands Institute for Health Services Research, Nivel, Utrecht, Netherlands

[9]Lung Foundation Netherlands, Amersfoort, Netherlands

[10]See Acknowledgments

[*]these authors contributed equally

Corresponding Author:
Thomas J Hoogeboom, PhD
IQ Healthcare
Radboud Institute for Health Sciences
Radboud University Medical Centre
Kapittelweg 54
Nijmegen, 6525 EP
Netherlands
Phone: 31 243615305
Email: Thomas.hoogeboom@radboudumc.nl

Abstract

Background: Patients recovering from COVID-19 often experience persistent problems in their daily activities related to limitations in physical, nutritional, cognitive, and mental functioning. To date, it is unknown what treatment is needed to support patients in their recovery from COVID-19.

Objective: This study aimed to evaluate the primary allied health care of patients recovering from COVID-19 at 6-month follow-up and to explore which baseline characteristics are associated with changes in the scores of outcomes between baseline and 6-month follow-up.

Methods: This Dutch nationwide prospective cohort study evaluated the recovery of patients receiving primary allied health care (ie, dietitians, exercise therapists, occupational therapists, physical therapists, and speech and language therapists) after COVID-19. All treatments offered by primary allied health professionals in daily practice were part of usual care. Patient-reported outcome measures on participation, health-related quality of life, fatigue, physical functioning, and psychological well-being were assessed at baseline and at 3- and 6-month follow-up. Linear mixed model analyses were used to evaluate recovery over time, and uni- and multivariable linear regression analyses were used to examine the association between baseline characteristics and recovery.

Results: A total of 1451 adult patients recovering from COVID-19 and receiving treatment from 1 or more primary allied health professionals were included. For participation (Utrecht Scale for Evaluation of Rehabilitation—Participation range 0-100), estimated mean differences of at least 2.3 points were observed at all time points. For the health-related quality of life (EuroQol Visual Analog Scale, range 0-100), the mean increase was 12.3 (95% CI 11.1-13.6) points at 6 months. Significant improvements were found for fatigue (Fatigue Severity Scale, range 1-7): the mean decrease was –0.7 (95% CI –0.8 to –0.6) points at 6 months. However, severe fatigue was reported by 742/929 (79.9%) patients after 6 months. For physical functioning (Patient-Reported Outcomes Measurement Information System—Physical Function Short Form 10b, range 13.8-61.3), the mean increase was 5.9 (95% CI 5.9-6.4) points at 6 months. Mean differences of –0.8 (95% CI –1.0 to –0.5) points for anxiety (Hospital Anxiety and Depression Scale range 0-21) and –1.6 (95% CI –1.8 to –1.3) points for depression were found after 6 months. A worse baseline score, hospital admission, and male sex were associated with greater improvement between baseline and 6-month follow-up, whereas age, the BMI, comorbidities, and smoking status were not associated with mean changes in any outcome measures.

Conclusions: Patients recovering from COVID-19 who receive primary allied health care make progress in recovery but still experience many limitations in their daily activities after 6 months. Our findings provide reference values to health care providers and health care policy makers regarding what to expect from the recovery of patients who receive health care from 1 or more primary allied health professionals.

Trial Registration: ClinicalTrials.gov NCT04735744; https://tinyurl.com/3vf337pn

International Registered Report Identifier (IRRID): RR2-10.2340/jrm.v54.2506

(JMIR Public Health Surveill 2023;9:e44155) doi:10.2196/44155

KEYWORDS

COVID-19; allied health care; primary care; care; patient; physical; nutritional; cognitive; mental functioning; support; recovery; diet; exercise; exercise therapist; physical therapist; speech therapist; language; descriptive statistics; regression; linear mixed model; statistics; statistician; statistical

Introduction

An estimated 32%-57% of patients recovering from a COVID-19 infection experience severe and long-term problems in daily functioning and participation [1-3]. It is becoming increasingly clear that both patients with mild symptoms and those with serious symptoms during an acute COVID-19 infection are at risk of developing the post–COVID-19 condition [1,2,4,5]. Post–COVID-19, also referred to as "long COVID," is defined as "signs and symptoms that develop during or after a COVID-19 infection, continuing for more than 12 weeks, and that are not explained by an alternative diagnosis" [6-8]. To date, it is unknown what treatment is needed to support patients in their recovery from COVID-19.

Patients recovering from COVID-19 often experience persistent problems in their daily activities related to limitations in physical, nutritional, cognitive, and mental functioning [3,5,9-11]. Fatigue is the most prevalent and persistent symptom, irrespective of the severity of the initial infection [3,5,10,12,13]. Longitudinal data suggest that fatigue does not resolve over time in many patients, even if they receive health care [3,9,10,13,14]. Increased levels of fatigue can result in lower levels of physical activity [15] and limit patients in activities of daily living (eg, housekeeping and grocery shopping) and outdoor pursuits [16]. Mental problems, such as anxiety and depression, are common in patients recovering from COVID-19. A study by Huang et al [17] showed that anxiety and depression were present in approximately 23% of patients 6 months after the onset of COVID-19 symptoms. Sisó-Almirall et al [10] showed that 36% of patients still reported mental problems after 3 months, and no significant associations were found with COVID-19 severity. Furthermore, previous studies have observed a worsened health-related quality of life (HRQoL) in patients recovering from COVID-19, both hospitalized and nonhospitalized, who did not recuperate after a follow-up period of several months [9,18-21].

The World Health Organization (WHO) suggests that rehabilitation for patients with the post–COVID-19 condition requires person-centered care that recommends multidisciplinary collaboration among health care professionals. These multidisciplinary rehabilitation interventions may include breathing techniques, physical exercise therapy, cognitive behavioral therapy, occupational therapy, nutritional support, and improving swallowing physiology [22]. In the Netherlands, mono- and multidisciplinary best-practice recommendations for primary allied health professionals have been developed for the treatment of patients recovering from COVID-19 [23-26]. Based on the overall effects of primary allied health care, it is expected that primary allied health professionals (ie, dietitians, exercise therapists, occupational therapists, physical therapists, and speech and language therapists) can play a role in the recovery of patients with COVID-19 who experience persistent limitations in daily physical functioning and participation. In July 2020, the Dutch Ministry of Health, Welfare and Sports instated a temporary regulation in primary allied health care to facilitate the treatment of patients recovering from COVID-19 and to stimulate research. This regulation enables the reimbursement of primary allied health care for every patient from basic health insurance coverage. With a referral from a general practitioner (GP) or medical specialist, primary allied health care treatment is reimbursed for a period of 6 months. If recovery during this period is insufficient, an extension by a second 6-month period is possible upon referral by a medical specialist. As COVID-19 is still a novel condition and the evidence base for allied health treatment in patients with post–COVID-19 syndrome is small, it is vital that new data and insights be shared as soon as they are available; therefore, the aim of this paper is to present the results of recovery of patients receiving primary allied health care after a COVID-19 infection.

We provide outcomes at 3- and 6-month follow-up regarding participation, the HRQoL, physical functioning, fatigue, and psychological well-being. In addition, we explore which baseline characteristics are associated with changes in these outcomes between baseline and 6-month follow-up.

Methods

Study Design and Setting

As part of a nationwide project to evaluate the recovery of patients receiving primary allied health care after a COVID-19 infection, a prospective cohort study was set up in collaboration with various patient organizations (ie, the Lung Foundation Netherlands, the Netherlands Patient Federation, and Harteraad) and with input from patients contacted through these organizations [27]. In this prospective cohort study, patients were included at the start of their treatment with 1 or more primary allied health professionals. All treatments offered by primary allied health professionals in daily practice were part of usual care and were preferably based on recommendations and guidelines published by the professional bodies of the respective care providers, as available at the start of the research [23-26]. The inclusion period for the cohort study was between March 29 and June 19, 2021. Primary outcome measures were assessed at baseline (T0) and again after 3 months (T1) and 6 months (T2). The full study protocol with timelines is published elsewhere [27]. In this paper, we report the results of our primary outcome measures at baseline, 3-month follow-up, and 6-month follow-up.

Ethical Considerations

The study protocol was approved by the Medical Ethics Committee of the Radboud University Medical Centre (registration #2020-7278). The study has been registered in the ClinicalTrials registry (NCT04735744). Informed consent was obtained from all patients before enrollment in the study, and all procedures were conducted in accordance with the Declaration of Helsinki.

Participants

Adult patients (age≥18 years) were eligible for inclusion in the cohort if they were recovering from COVID-19 and started treatment with 1 or more primary allied health professionals (ie, a dietitian, exercise therapist, occupational therapist, physical therapist, or speech and language therapist). Patients may have received treatment from 1 or more primary allied health professionals during the course of the study. Patients were included regardless of their hospital admission status during the acute phase of COVID-19. Patients who were unable to complete questionnaires in Dutch and patients who were receiving palliative care were excluded from the study.

Data Collection

Patients could enroll in the study by (1) signing up after an invitation from their treating primary allied health professional or (2) signing up on their own initiative, upon which the research team also invited the treating primary allied health professional to participate. The enrollment procedure of this study is described in detail in the published study protocol [27]. Both patients and primary allied health professionals reported data via the specifically designed Your Research app. Patients were asked to download the app on their smartphones or make use of the web version. Questionnaires were sent out through this app at the start of the treatment (baseline) and after 3 and 6 months. Patients unable to participate via digital methods were provided with an opportunity to complete the questionnaires on paper and return them by post. Primary allied health professionals were asked to use the web version of the app.

Outcome Measures

Data on patient characteristics were collected by the treating primary allied health professionals at the start of the treatment. Patient-reported outcome domains (participation, HRQoL, fatigue, and physical functioning) were assessed at baseline and after 3 and 6 months. Data on psychological well-being were collected at baseline and after 6 months.

Patient Characteristics

Patient characteristics were collected via an online record form and contained the following items on demographics: age, sex, height (in cm), weight (in kg) both at the start of treatment and before COVID-19 infection, living status (whether the patient had an informal caregiver), and referring physician. Furthermore, data on symptom severity at the onset of treatment (ie, mild to moderate [mild symptoms up to mild pneumonia], severe [dyspnea, hypoxia, or <50% lung involvement on imaging], or critical [respiratory failure, shock, or multiorgan system dysfunction], as described in Ref. [28]) as well as hospital admissions during the acute phase of COVID-19 (ie, no hospital admission, admission to hospital ward or intensive care unit [ICU]) were recorded. Additionally, data on comorbidities (ie, cardiovascular disease, chronic lung disease, diabetes mellitus, kidney disease, liver disease, immune disease, oncological disease, chronic neuromuscular disorders) and smoking status were collected. The body weight and height were used to calculate each patient’s BMI (weight/height^2) and categorized as defined by WHO [29].

Participation

Participation was assessed with the Utrecht Scale for Evaluation of Rehabilitation—Participation (USER-P). The USER-P is a 31-item self-administered questionnaire reflecting a patient’s participation in daily life, divided over 3 subscales: frequencies, restrictions, and satisfaction. The total scores range from 0 to 100 for each subscale, with higher scores indicating better participation (higher frequency, fewer restrictions, and higher satisfaction) [30]. We arbitrarily assumed a 5-point difference on 1 of these USER-P scales to be clinically relevant for patients recovering from COVID-19 [31,32].

Health-Related Quality of Life

The HRQoL was assessed with the EuroQoL 5 Dimensions 5 Level (EQ-5D-5L) tool, a 5-item questionnaire measuring a person’s status on 5 dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [33]. Furthermore, the EuroQoL Visual Analog Scale (EQ-VAS) was recorded by the patients. The EQ-VAS provides a quantitative measure of a patient’s perception of their overall health, with a score ranging from 0 to 100, with higher scores indicating a higher HRQoL. A difference of 8 points on the EQ-VAS was considered clinically relevant [34].

Fatigue

Fatigue was assessed with the Fatigue Severity Scale (FSS), a 9-item scale measuring the severity of fatigue and its effect on patients' activities and lifestyle. The score of each item ranges from 1 to 7, where 1 indicates strong disagreement and 7 indicates strong agreement. The total score is calculated using the mean value of the 9 items, with a score of 4 or more indicating severe fatigue [35]. A difference of 0.45 points on the FSS mean score was considered clinically relevant [36].

Physical Functioning

Limitations in physical functioning were assessed with the Patient-Reported Outcomes Measurement Information System Physical Functioning Short Form 10b (PROMIS-PF-10b), a 10-item questionnaire measuring the self-reported ability to perform activities of daily life. Items reflect 4 subcategories: upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck and back), as well as instrumental activities of daily living, such as running errands [37]. Total scores range from 13.8 (severely physically impaired) to 61.3 (not physically impaired), with a mean score of 50 (SD 10) representing the mean score of a reference population [38]. A difference of 3.6 points was considered clinically relevant [39].

Psychological Well-Being

The Hospital Anxiety and Depression Scale (HADS) was used to assess psychological well-being. This 14-item self-administered questionnaire describes symptoms of anxiety and depression. The HADS is divided into an anxiety score (HADS-A) and a depression score (HADS-D), each containing 7 items. The total score ranges from 0 to 21 for both subscales, where a total score of 11 or more indicates a probable clinical diagnosis of depression or anxiety [40,41]. A difference of 1.7 points was considered clinically relevant [42].

Statistical Analysis

Descriptive statistics were used to describe the patient population and to analyze the primary outcome measures at baseline and after 3 and 6 months using numbers and proportions for categorical variables, means (SDs), and medians (IQRs) for continuous variables. Linear mixed model analyses were used to evaluate recovery over time for participation, the HRQoL, fatigue, physical functioning, and psychological well-being. This analysis accounts for correlation between repeated measures on the same subject and uses all available data from this subject. A model with a random intercept and all other variables fixed was also generated. The primary outcomes were used as dependent variables, while time (categorical: baseline, 3 months, and 6 months) was used as a fixed factor.

Uni- and multivariate regression analyses were used to explore which baseline characteristics were associated with changes in the scores of the main outcome measures between baseline (T0) and 6-month follow-up (T2). This analysis used data from complete cases, and missingness at random (MAR) was tested (Table S1 in Multimedia Appendix 1). Univariate analyses were performed to determine which baseline characteristics (ie, age, sex, BMI, hospitalization, comorbidities, baseline score, and smoking status) were associated with the mean change in each outcome measure. Comorbidities were coded into 3 categories: none, 1, and 2 or more comorbidities. Variables with *P*<.157 in the univariate regression were included in the multivariate model [43]. The backward elimination of variables was then performed in order of statistical significance until only factors that were significantly associated with the outcome remained. A sensitivity analysis was performed by forcing age, sex, and hospital admission into the models as these factors are known to be related with recovery over time. Results of this sensitivity analysis are presented in Table S2 in Multimedia Appendix 1. *P*<.05 was considered statistically significant for all analyses based on 2-sided testing. All data were analyzed using SPSS Statistics version 25 (IBM Corp).

Patient and Public Involvement

During the development of this study, we involved patients to provide feedback on the readability and appropriateness of proposed measures. The usability of the smartphone and web-based app versions was also tested by patients. Participating patients received updates on the status of the study via their smartphone or web app. Furthermore, various patient organizations (ie, the Lung Foundation Netherlands, the Netherlands Patient Federation, and Harteraad) participated during routine research meetings.

Results

Patient Characteristics

In total, 1451 patients were included in this study (Figure 1), receiving 1708 different allied health care treatments. Their mean age was 49 (SD 13) years, and 63.8% (848/1330) of the patient population was female (Table 1). The majority (1015/1315, 77.2%) had not been hospitalized for COVID-19, and 1002/1311 (76.4%) patients had experienced mild-to-moderate severity of symptoms during the infection period. The mean BMI was 28 (SD 6) kg/m^2, and 68.9% (738/1071) of the patient population was classified as being overweight or obese (BMI>25 kg/m^2). In addition, 1 comorbidity was reported by 410/1331 (30.8%) patients, and 2 or more comorbidities were reported by 155/1331 (11.6%) patients. Cardiovascular disease (193/1331, 14.5%) and chronic lung disease (183/1331, 13.7%) were the most prevalent comorbidities. Most patients (1086/1331, 81.6%) had been referred for primary allied health care by their GP.

Figure 1. Flow diagram for inclusion of patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study, with follow-up after 3 and 6 months from the start of treatment. EQ-5D-5L: EuroQoL 5 Dimensions 5 Level; FSS: Fatigue Severity Scale; HADS: Hospital Anxiety and Depression Scale; PROMIS: Patient-Reported Outcomes Measurement Information System; USER-P: Utrecht Scale for Evaluation of Rehabilitation–Participation.

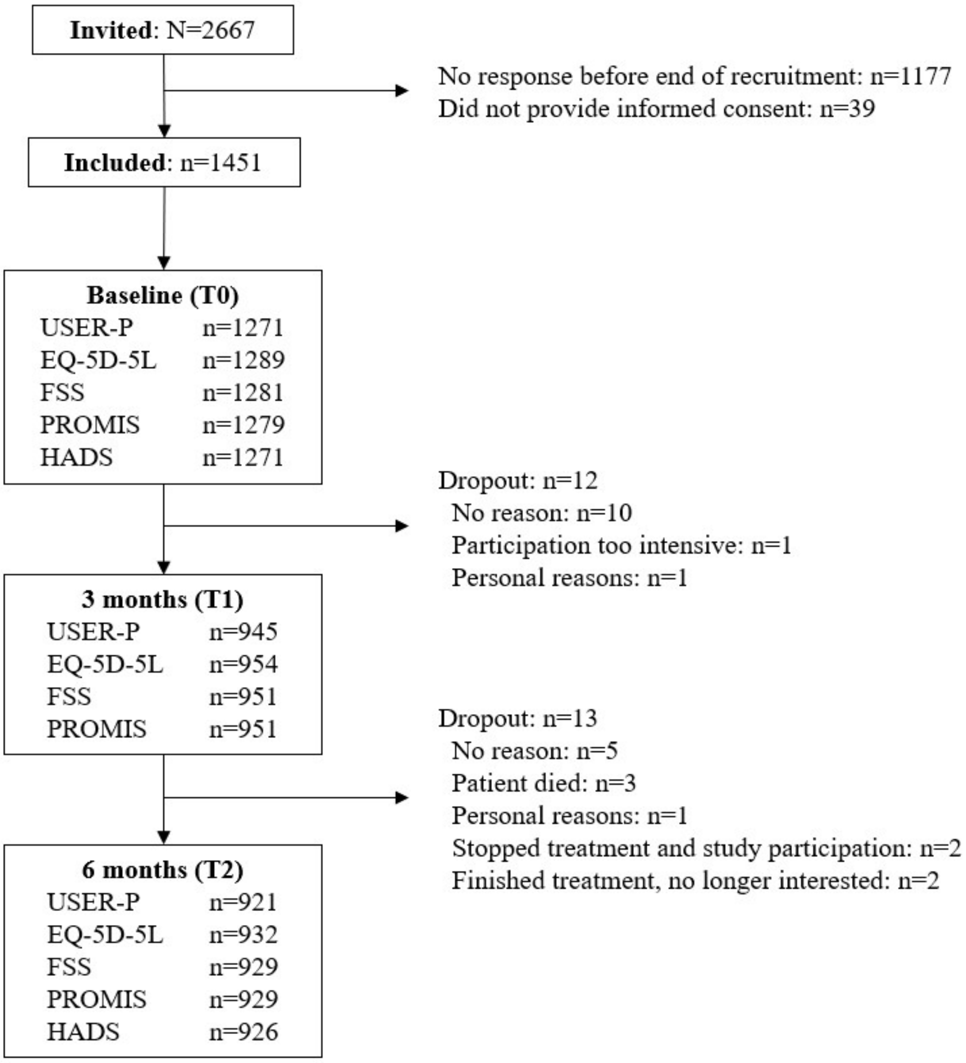


Table 1. General characteristics of patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (N=1451).

General characteristics	Participants
Treatments (n=1708)[a], n (%)	
Physical therapy/exercise therapy	1005 (58.8)
Occupational therapy	364 (21.3)
Dietary care	224 (13.1)
Speech and language therapy	115 (6.7)
Sex (n=1330)[b], n (%)	
Male	482 (36.2)
Female	848 (63.8)
Age (years; n=1331), mean (SD)	49 (13)
COVID-19 severity (n=1311)[b], n (%)	
Mild/moderate	1002 (76.4)
Serious	271 (20.7)
Very serious	38 (2.9)
Admission to hospital for COVID-19 infection (n=1315)[b], n (%)	
Hospitalized, including the ICU[c]	87 (6.6)
Hospitalized	213 (16.2)
Not hospitalized	1015 (77.2)
BMI (kg/m^2; n=1071)[b]; mean 28 (SD 6)	
Underweight (<18.5)	10 (0.9)
Normal weight (18.5-25.0)	323 (30.2)
Overweight (25.0-30.0)	404 (37.7)
Obese (>30.0)	334 (31.2)
Smoking status (n=1305)[b], n (%)	
Current	63 (4.8)
Former	166 (12.7)
Never	1076 (82.5)
Living status (n=1322)[b], n (%)	
Alone	212 (16.0)
Cohabiting	1110 (84.0)
Informal caregiver (n=1319)[b], n (%)	
Yes	526 (39.9)
No	793 (60.1)
Comorbidities (n=1331), n (%)	
0	766 (57.6)
1	410 (30.8)
≥2	155 (11.6)

[a] Some participants received multiple treatments from multiple primary allied health professionals. Therefore, the number of treatments exceeded the number of participants.

[b] Data were not fully available for all patients.

[c] ICU: intensive care unit.

Primary Outcome Measures

Table 2 presents data on the outcome measures at baseline and at 3- and 6-month follow-up. Additionally, clinically relevant improvements at 6-month follow-up are presented in Table S3 in Multimedia Appendix 1. After 6 months, the majority of patients showed a clinically relevant improvement on the USER-P restrictions and satisfaction subscales (576/890, 64.7%, patients and 543/891, 60.9%, patients, respectively), while 540/908 (59.5%) patients showed a clinically relevant improvement on the EQ-VAS (mean 67.4, SD 19.1 points) compared to baseline (mean 55.5, SD 17.8 points). Severe fatigue was reported by 1205/1281 (94.1%) patients at baseline, persisting after 6 months in 742/929 (79.9%) patients. A clinically relevant improvement on the FSS mean score was found in 490/904 (54.2%) patients. Based on PROMIS-PF-10b scores, over two-thirds of the patients reported being more than 60% impaired, limited, or restricted in physical functioning at baseline, which decreased to 37.8% (351/929) after 6 months; 517/902 (57.3%) patients experienced a clinically relevant improvement in physical functioning. The majority of patients scored less than 7 points on the HADS anxiety and depression scores both at baseline and at 6 months, which indicates no anxiety disorder or depression. At baseline, the HADS anxiety score indicated a probable clinical diagnosis of anxiety disorder in 292/1271 (23%) patients, which decreased slightly to 170/926 (18.4%) patients after 6 months. A probable clinical diagnosis of depression was indicated by the HADS depression score in 285/1271 (22.4%) patients at baseline, decreasing to 137/926 (14.8%) patients at 6-month follow-up.

Table 2. General outcome measures at baseline and after 3 and 6 months in patients[a] recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study.

General outcome measures	Baseline (T0)	3 months (T1)	6 months (T2)
Participation			
Sample population, n/N (%)	1271/1451 (87.6)	945/1451 (65.1)	921/1451 (63.5)
USER-P[b] frequencies subscale score, mean (SD)	27.5 (10.3)	30.5 (10.5)	31.3 (10.1)
USER-P restrictions subscale score, mean (SD)	65.8 (20.1)	73.6 (19.4)	77.3 (19.8)
USER-P satisfaction subscale score, mean (SD)	48.6 (17.8)	54.8 (19.5)	58.1 (19.9)
HRQoL[c]			
Sample population, n/N (%)	1289/1451 (88.8)	954/1451 (65.7)	932/1451 (64.2)
EQ-VAS[d] score, mean (SD)	55.5 (17.8)	64.3 (18.2)	67.4 (19.1)
Fatigue			
Sample population, n/N (%)	1281/1451 (88.3)	951/1451 (65.5)	929/1451 (64.0)
FSS[e] score, mean (SD)	5.6 (1.0)	5.2 (1.2)	4.9 (1.3)
≥4 points, n (%)	1205 (94.1)	815 (85.7)	742 (79.9)
Physical functioning			
Sample population, n/N (%)	1279/1451 (88.1)	951/1451 (65.5)	929/1451 (64.0)
PROMIS-PF-10b[f] score, mean (SD)	37.7 (6.0)	41.5 (7.6)	43.5 (8.50)
100% impaired, limited, or restricted, n (%)	2 (0.2)	2 (0.2)	140 (15.1)
80%-99% impaired, limited, or restricted, n (%)	427 (33.4)	176 (18.5)	211 (22.7)
60%-79% impaired, limited, or restricted, n (%)	452 (35.3)	276 (29.0)	202 (21.7)
40%-59% impaired, limited, or restricted, n (%)	259 (20.3)	227 (23.9)	157 (16.9)
20%-39% impaired, limited, or restricted, n (%)	97 (7.6)	133 (14.0)	150 (16.1)
1%-19% impaired, limited, or restricted, n (%)	39 (3.0)	95 (10.0)	72 (7.8)
0% impaired, limited, or restricted, n (%)	4 (0.3)	42 (4.4)	140 (15.1)
Psychological well-being[g] (anxiety)			
Sample population, n/N (%)	1271/1451 (87.6)	N/A[h]	926/1451 (63.8)
HADS[i] anxiety score, mean (SD)	7.1 (4.5)	N/A	6.3 (4.7)
≤7 points, n (%)	746 (58.7)	N/A	613 (66.2)
8-10 points, n (%)	233 (18.3)	N/A	143 (15.4)
≥11 points, n (%)	292 (23.0)	N/A	170 (18.4)
Psychological well-being (depression)			
Sample population, n/N (%)	1271/1451 (87.6)	N/A	926/1451 (63.8)
HADS depression score, mean (SD)	7.3 (4.2)	N/A	5.7 (4.3)
≤7 points, n (%)	689 (54.2)	N/A	638 (68.9)
8-10 points, n (%)	297 (23.4)	N/A	151 (16.3)
≥11 points, n (%)	285 (22.4)	N/A	137 (14.8)

[a] Data were not fully available for all patients.

[b] USER-P: Utrecht Scale for Evaluation of Rehabilitation—Participation.

[c] HRQoL: health-related quality of life.

[d] EQ-VAS: EuroQol Visual Analog Scale.

[e] FSS: Fatigue Severity Scale.

[f] PROMIS-PF-10b: Patient-Reported Outcomes Measurement Information System Physical Functioning Short Form 10b.

[g] Psychological well-being was only assessed at baseline (T0) and at 6 months (T2).

[h] N/A: not applicable.

[i] HADS: Hospital Anxiety and Depression Scale.

Patient-Reported Recovery Over Time

Table 3 shows the effect of time on the outcome measures. For all dependent variables, a random intercept model was the best-fitting model. No variables were significantly related to missing values in the outcome measures at any point in time. A significant effect of time was observed for all outcome measures at 3- and 6-month follow-up (*P*<.001). For participation, estimated mean differences of at least 2.9 points (*P*<.001) were observed for all 3 subscales at all time points. For the HRQoL, the mean increase was 9.0 points (95% CI 7.8-10.2) at 3 months and 12.3 points (95% CI 11.1-13.6) after 6 months. Furthermore, significant improvements were found for fatigue and physical functioning at all time points. The greatest improvements were seen after just 3 months for all outcome measures measured at both 3 and 6 months. Mean differences of –0.8 (95% CI –1.0 to –0.5) on the HADS anxiety score and –1.6 (95% CI –1.8 to –1.3) on the HADS depression score were observed.

Table 3. Results of linear mixed model analysis for the outcome measures participation, HRQoL[a], fatigue, physical functioning, and psychological well-being in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study.

General outcome measures	Baseline (T0), mean (SE)	3 months (T1), mean (SE)	6 months (T2), mean (SE)	At 3 months		At 6 months	
				Estimate (95% CI)	*P* value	Estimate (95% CI)	*P* value
Participation							
USER-P[b] frequencies subscale	27.5 (0.3)	30.5 (0.3)	31.5 (0.3)	2.9 (2.3 to 3.7)	<.001	3.9 (3.3 to 4.7)	<.001
USER-P restrictions subscale	64.6 (0.6)	73.7 (0.6)	77.6 (0.6)	9.1 (7.9 to 10.3)	<.001	13.0 (11.8 to 14.2)	<.001
USER-P satisfaction subscale	48.7 (0.5)	54.7 (0.6)	58.4 (0.6)	5.9 (4.8 to 7.2)	<.001	9.7 (8.5 to 10.9)	<.001
HRQoL (EQ-VAS[c] score)	55.6 (0.5)	64.6 (0.6)	67.9 (0.7)	9.0 (7.8 to 10.2)	<.001	12.3 (11.1 to 13.6)	<.001
Fatigue (FSS[d] score)	5.6 (0.03)	5.2 (0.04)	4.9 (0.04)	–0.4 (–0.5 to –0.4)	<.001	–0.7 (–0.8 to –0.6)	<.001
Physical functioning (PROMIS-PF-10b[e] score)	37.7 (0.2)	41.6 (0.2)	43.7 (0.2)	3.9 (3.5 to 4.3)	<.001	5.9 (5.6 to 6.4)	<.001
Psychological well-being[f]							
HADS[g] anxiety score	7.1 (0.1)	N/A[h]	6.3 (0.1)	N/A	N/A	–0.8 (–1.0 to –0.5)	<.001
HADS depression score	7.3 (0.1)	N/A	5.7 (0.1)	N/A	N/A	–1.6 (–1.8 to –1.3)	<.001

[a] HRQoL: health-related quality of life.

[b] USER-P: Utrecht Scale for Evaluation of Rehabilitation—Participation.

[c] EQ-VAS: EuroQol Visual Analog Scale.

[d] FSS: Fatigue Severity Scale.

[e] PROMIS-PF-10b: Patient-Reported Outcomes Measurement Information System Physical Functioning Short Form 10b.

[f] Psychological well-being was only assessed at baseline (T0) and at 6 months (T2).

[g] HADS: Hospital Anxiety and Depression Scale.

[h] N/A: not applicable.

Factors Associated With Changes in the Scores of the Main Outcome Measures

Multivariable regression models were estimated to identify factors associated with changes in scores between baseline and 6-month follow-up for each outcome measure. Tables 4-10 provide an overview of the final regression models. All univariable and multivariable regression models are shown in Tables S4 and S5 in Multimedia Appendix 1. Having a worse baseline score was related to greater improvements for all outcome measures. For all 3 subscales of the USER-P and physical functioning, patients admitted to the hospital during the infection period of COVID-19 showed greater improvements in scores than nonhospitalized patients, even when correcting for baseline scores. In terms of the HRQoL, patients admitted to a hospital ward showed greater improvements than patients who had not been hospitalized, although no associations were found with ICU admissions. Male participants showed greater improvements than female participants in all outcome measures, except for psychological well-being, for which no association was found for sex. The baseline age, BMI, comorbidities, and smoking status were not significantly associated with the mean change in any of the outcome measures in our patient population. In a sensitivity analysis (Table S2 in Multimedia Appendix 1) where age, sex, and hospital admission were forced into the model, additional associations were found between the male sex and satisfaction in participation and between age and frequencies of participation, physical functioning, and symptoms of anxiety.

Table 4. Multivariable linear regression models on the outcome measure participation (USER-P[a] frequencies subscale) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (R2 overall model=0.272, *P*<.001).

Outcome measure	β	95% CI	*P* value
Hospital admission			.001
No	Reference	N/A[b]	N/A
Hospital ward	2.556	0.851 to 4.262	.003
ICU[c]	3.079	0.611 to 5.547	.015
Baseline score	–.496	–0.558 to –0.435	<.001

[a] USER-P: Utrecht Scale for Evaluation of Rehabilitation—Participation.

[b] N/A: not applicable.

[c] ICU: intensive care unit.

Table 5. Multivariable linear regression models on the outcome measure participation (USER-P[a] restrictions subscale) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (R2 overall model=0.277, *P*<.001).

Outcome measure	β	95% CI	*P* value
Sex			.001
Male	Reference	N/A[b]	N/A
Female	–5.337	–7.813 to –2.861	.001
Hospital admission			<.001
No	Reference	N/A	N/A
Hospital ward	3.581	0.316 to 6.845	.032
ICU[c]	9.165	4.522 to 13.809	<.001
Baseline score	–.462	–0.520 to –0.405	<.001

[a] USER-P: Utrecht Scale for Evaluation of Rehabilitation—Participation.

[b] N/A: not applicable.

[c] ICU: intensive care unit.

Table 6. Multivariable linear regression models on the outcome measure participation (USER-P[a] satisfaction subscale) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (R2 overall model=0.159, *P*<.001).

Outcome measure	β	95% CI	*P* value
Hospital admission			.003
No	Reference	N/A[b]	N/A
Hospital ward	3.577	0.356 to 6.798	.030
ICU[c]	6.728	2.144 to 11.311	.004
Baseline score	–.402	–0.467 to –0.338	<.001

[a] USER-P: Utrecht Scale for Evaluation of Rehabilitation—Participation.

[b] N/A: not applicable.

[c] ICU: intensive care unit.

Table 7. Multivariable linear regression models on the outcome measure HRQoL[a] (EQ-VAS[b] score) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (R2 overall model=0.245, *P*<.001).

Outcome measure	β	95% CI	*P* value
Sex			<.001
Male	Reference	N/A[c]	N/A
Female	–4.855	–7.378 to –2.333	<.001
Hospital admission			.10
No	Reference	N/A	N/A
Hospital ward	3.594	0.231 to 6.957	.036
ICU[d]	2.106	–2.615 to 6.827	.38
Baseline score	–.524	–0.589 to –0.459	<.001

[a] HRQoL: health-related quality of life.

[b] EQ-VAS: EuroQol Visual Analog Scale.

[c] N/A: not applicable.

[d] ICU: intensive care unit.

Table 8. Multivariable linear regression models on the outcome measure fatigue (FSS[a] score) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (R2 overall model=0.064, *P*<.001).

Outcome measure	β	95% CI	*P* value
Sex			<.001
Male	Reference	N/A[b]	N/A
Female	.284	0.130 to 0.438	<.001
Baseline score	–.301	–0.381 to –0.222	<.001

[a] FSS: Fatigue Severity Scale.

[b] N/A: not applicable.

Table 9. Multivariable linear regression models on the outcome measure physical functioning (PROMIS-PF-10b[a] score) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study (R2 overall model=0.064, *P*<.001).

Outcome measure	β	95% CI	*P* value
Sex			<.001
Male	Reference	N/A[b]	N/A
Female	–2.342	–3.341 to –1.343	<.001
Hospital admission			
No	Reference	N/A	.004
Hospital ward	1.149	–0.165 to 2.463	.09
ICU[c]	2.917	1.064 to 4.771	.002
Baseline score	–.125	–0.203 to –0.046	<.001

[a] PROMIS-PF-10b: Patient-Reported Outcomes Measurement Information System Physical Functioning Short Form 10b.

[b] N/A: not applicable.

[c] ICU: intensive care unit.

Table 10. Multivariable linear regression models on the outcome measure psychological well-being (HADS[a] anxiety and depression scores; all baseline scores) in patients recovering from COVID-19 receiving primary allied health care in a Dutch prospective cohort study.

Outcome measure	β	R^2 overall model	95% CI	*P* value
HADS anxiety	–.354	0.160 (*P*<.001)	–0.407 to –0.301	<.001
HADS depression	–.392	0.179 (*P*<.001)	–0.447 to –0.337	<.001

[a] HADS: Hospital Anxiety and Depression Scale.

Discussion

Principal Findings

This study presents the first results of our evaluation of the recovery of our unique cohort of patients with COVID-19 receiving primary allied health care until their 6-month follow-up. We explored which baseline characteristics were associated with changes in the scores of the main outcome measures over this 6-month period. Most patients showed a clinically relevant improvement in all outcome measures; however, despite improvement, many patients still experienced persistent problems in their daily lives, with limitations in physical and mental functioning. A worse baseline score, hospital admission, and, for some outcome measures, the male sex were associated with greater improvement between baseline and 6-month follow-up; however, age, the BMI, comorbidities, and smoking status were not associated with the mean change in any of the outcome measures.

Comparison With Other Studies

The majority of our patient population showed a clinically relevant improvement 6 months after starting treatment provided by 1 or more primary allied health professionals; nevertheless, a large group of patients experienced persistent problems in their daily lives. The mean EQ-VAS score of our patient population (67 points) remained well below the population norm in the Netherlands, which is 82 points [44]. These results are consistent with previous findings that the HRQoL is impaired in the majority of patients post–COVID-19 [12,15,17,45-48]. Persistent fatigue was highly prevalent among the patients included in our study, with 79.9% still reporting severe fatigue (measured with the FSS) after 6 months. These results are consistent with previous studies on patients recovering from COVID-19, showing that fatigue is the most common complaint [5,10,14,47,49-51], even after 6 months [13,15,52-54]. The mean PROMIS-PF-10b score of our population (mean 43.5, SD 8.5) remained well below the population norm in the Netherlands (mean 50, SD 10). These results are also consistent with previous studies [15,47] and indicate that persistent symptoms due to COVID-19 may lead to patients experiencing limitations in physical functioning.

Relative to other outcome measures, a smaller percentage of patients showed a clinically relevant improvement in psychological well-being. This was due to an observed ceiling effect, as 58.7% and 54.2% of patients showed no indication of an anxiety disorder or depression at baseline, respectively. Data of these patients are still informative, however, as they could also have deteriorated throughout the follow-up period. With scores indicating a probable clinical diagnosis of anxiety disorder or depression in 18.4% and 14.8% of patients, respectively, after 6 months, our findings are similar to those reported in previous studies, which showed prevalence rates ranging from 11% to 40% [8,17,46,48,51,54-57]. Furthermore, we performed an additional subgroup analysis to explore the differences in changes in the scores of the outcome measures between patients who showed indications of depression or anxiety disorder at baseline and patients who did not (data not shown). Based on this analysis, we conclude that whether a patient shows indications of depression or anxiety disorder at baseline has little effect on the change in their scores over time.

We found that male participants showed greater improvements than female participants in participation, the HRQoL, fatigue, and physical functioning. These results are consistent with previous studies showing that female participants experience more persistent symptoms after a COVID-19 infection [10,14,21,45,49,51]. Furthermore, patients admitted to the hospital for COVID-19 showed greater improvements than nonhospitalized patients in terms of participation, the HRQoL, and physical functioning, which is in line with previous studies [12,21,45,51]. We observed no associations between fatigue and hospital admission, age, the BMI, comorbidities, or smoking status, which is also consistent with other studies [10,14,51,53,54], indicating that fatigue is highly prevalent in patients recovering from COVID-19, irrespective of the severity of initial infection and patient characteristics. We found that having a worse baseline score is related to greater improvement in anxiety and depressive symptoms; however, no associations with any patient characteristics were found. Similar to our results, previous studies have found no associations between the frequency of anxiety or depressive symptoms and disease severity or hospital admission [10,46,48,49,57]. In contrast, other studies have found the female sex [54,58,59] and older age [54,56] to be predictors of anxiety or depressive symptoms in patients with COVID-19. Although the association between age and symptoms of anxiety was not evident in our primary analysis, it did show up in our sensitivity analysis (Table S2 in Multimedia Appendix 1).

Limitations

This study did not include a control group to determine the potential effects of primary allied health care by comparing outcome measures with patients who did not receive this type of care. In addition, with a lack of available pre–COVID-19 data for our population, it was difficult to draw conclusions about the impact of pre-existing conditions versus problems in the daily activities and participation of these patients due to their COVID-19 infection.

For the interpretation of results, it is important to consider that the baseline measurement in this study was taken at the start of the treatment by 1 or more primary allied health professionals. It is possible that a patient had already experienced symptoms for some time and only consulted a primary allied health professional at a later stage. Additionally, it should be considered that not all patients received treatment from 1 or more primary allied health professionals during the entire 6-month follow-up period of this study. Some patients received short-term treatment, while others were still receiving treatment at 6 months.

The characteristics of included patients are comparable to COVID-19 populations described by others [3], and therefore, we assume that the study population is representative of the population in 2021, our inclusion period. Based on the inclusion period, which was between March and July 2021, our population most likely had the Wuhan or Alpha variant of SARS-CoV-2 [60]. Different variants may cause different symptoms, and the recovery of patients infected with other variants (eg, Delta or Omicron) may differ from that of our population. A total of 25 patients dropped out during this study (Figure 1). Although a proportion of the patients did not complete all questionnaires, the response rates were still sufficient: 93% at baseline, 68% after 3 months, and 67% after 6 months [61]. There was no selective missingness of data based on patient characteristics (including disease severity) and scores of the outcome measures (Table S1 in Multimedia Appendix 1).

Implications and Future Perspectives

Future research and in-depth analyses of our data are needed to gain more insight into the outcome measures and recovery of patients after COVID-19 who visit 1 or more primary allied health professionals. Future papers will include the results after a 12-month follow-up, determining the related health care costs and profession-specific outcomes per allied health discipline.

Conclusion

The results of this study show that patients recovering from COVID-19 and receiving primary allied health care make progress in recovery, but many still experience limitations in their daily activities and participation after 6 months. The findings of our study provide reference values for health care providers and health care policy makers about what to expect from the recovery of patients who receive or have received health care from 1 or more primary allied health professionals.

Acknowledgments

The authors would like to thank all patients and primary allied health professionals for reporting their data. We would like to thank the professional bodies of the primary allied health professionals in optimizing patient inclusion and the advisory board of the project for their input from their stakeholders' perspective. We would also like to thank all research staff for support with the collection of data, data cleaning, and statistical advice. Finally, we would like to thank all remaining collaborators from the Dutch Consortium Allied Healthcare COVID-19 for their contribution: Carla Agasi-Idenburg, Johanna M van Dongen, Edwin Geleijn, Ron van Heerde, Anja de Kruif, Marike van der Leeden, Raymond W J G Ostelo, Amber Ronteltap, Marike van der Schaaf, Sonja van Oers, and Cindy Veenhof.

This project was funded by ZonMw Efficiency Studies (10390062010001) and received additional funding for setting up the data collection tool by the Royal Dutch Society for Physiotherapy, the Association for Quality in Physical Therapy, the Nivel Netherlands Institute for Health Services Research, Stichting Revalidatie en Wetenschap, and Maastricht University.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

The Dutch Consortium Allied Healthcare COVID-19 designed the research. AIS, MHGG, and ACV conducted the research. AIS and MHGG analyzed the data. ACV, HMK, EHCC, JGK, AFL, WMM, and TJH contributed to the analyses and interpretation of data. All authors critically reviewed the manuscript and approved the final version of the manuscript submitted for publication. The corresponding author attests that all listed authors meet the authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

All authors have completed the International Committee of Medical Journal Editors uniform disclosure form and declare that all authors received financial support from ZonMw for the submitted work, had no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years, and had no other relationships or activities that could appear to have influenced the submitted work.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 64 KB - publichealth_v9i1e44155_app1.docx]

References

1. Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: a 6-month retrospective cohort study of 273,618 survivors of COVID-19. PLoS Med 2021 Sep 28;18(9):e1003773 [FREE Full text] [doi: 10.1371/journal.pmed.1003773] [Medline: 34582441]
2. Webber S, Tittlemier BJ, Loewen HJ. Apparent discordance between the epidemiology of COVID-19 and recommended outcomes and treatments: a scoping review. Phys Ther 2021 Nov 01;101(11):pzab155 [FREE Full text] [doi: 10.1093/ptj/pzab155] [Medline: 34160029]
3. Ballering AV, van Zon SKR, olde Hartman TC, Rosmalen JGM. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. Lancet 2022 Aug;400(10350):452-461. [doi: 10.1016/s0140-6736(22)01214-4]
4. Augustin M, Schommers P, Stecher M, Dewald F, Gieselmann L, Gruell H, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. Lancet Reg Health Eur 2021 Jul;6:100122 [FREE Full text] [doi: 10.1016/j.lanepe.2021.100122] [Medline: 34027514]
5. Goërtz YMJ, Van Herck M, Delbressine JM, Vaes AW, Meys R, Machado FVC, et al. Persistent symptoms 3 months after a SARS-CoV-2 infection: the post-COVID-19 syndrome? ERJ Open Res 2020 Oct;6(4):00542-2020 [FREE Full text] [doi: 10.1183/23120541.00542-2020] [Medline: 33257910]
6. COVID-19 rapid guideline: managing the long-term effects of COVID-19. National Institute for Health and Care Excellence. 2020. URL: https://www.nice.org.uk/guidance/ng188 [accessed 2023-09-15]
7. Shah W, Hillman T, Playford ED, Hishmeh L. Managing the long term effects of covid-19: summary of NICE, SIGN, and RCGP rapid guideline. BMJ 2021 Jan 22;372:n136. [doi: 10.1136/bmj.n136] [Medline: 33483331]
8. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr 22;27(4):601-615 [FREE Full text] [doi: 10.1038/s41591-021-01283-z] [Medline: 33753937]
9. Carfì A, Bernabei R, Landi F, Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent symptoms in patients after acute COVID-19. JAMA 2020 Aug 11;324(6):603-605 [FREE Full text] [doi: 10.1001/jama.2020.12603] [Medline: 32644129]
10. Sisó-Almirall A, Brito-Zerón P, Conangla Ferrín L, Kostov B, Moragas Moreno A, Mestres J, et al. Long Covid-19: proposed primary care clinical guidelines for diagnosis and disease management. Int J Environ Res Public Health 2021 Apr 20;18(8):4350 [FREE Full text] [doi: 10.3390/ijerph18084350] [Medline: 33923972]
11. Maxwell E. Living with Covid19. National Institute for Health Research. 2020. URL: https://evidence.nihr.ac.uk/collection/living-with-covid19/ [accessed 2023-09-15]
12. Malik P, Patel K, Pinto C, Jaiswal R, Tirupathi R, Pillai S, et al. Post-acute COVID-19 syndrome (PCS) and health-related quality of life (HRQoL)—a systematic review and meta-analysis. J Med Virol 2022 Jan 07;94(1):253-262 [FREE Full text] [doi: 10.1002/jmv.27309] [Medline: 34463956]
13. Van Herck M, Goërtz YMJ, Houben-Wilke S, Machado FVC, Meys R, Delbressine JM, et al. Severe fatigue in long COVID: web-based quantitative follow-up study in members of online long COVID support groups. J Med Internet Res 2021 Sep 21;23(9):e30274 [FREE Full text] [doi: 10.2196/30274] [Medline: 34494964]
14. Townsend L, Dyer AH, Jones K, Dunne J, Mooney A, Gaffney F, et al. Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection. PLoS One 2020;15(11):e0240784 [FREE Full text] [doi: 10.1371/journal.pone.0240784] [Medline: 33166287]
15. Tabacof L, Tosto-Mancuso J, Wood J, Cortes M, Kontorovich A, McCarthy D, et al. Post-acute COVID-19 syndrome negatively impacts physical function, cognitive function, health-related quality of life, and participation. Am J Phys Med Rehabil 2022;101(1):48-52. [doi: 10.1097/phm.0000000000001910]
16. Humphreys H, Kilby L, Kudiersky N, Copeland R. Long COVID and the role of physical activity: a qualitative study. BMJ Open 2021 Mar 10;11(3):e047632 [FREE Full text] [doi: 10.1136/bmjopen-2020-047632] [Medline: 33692189]
17. Huang C, Li X, Gu X, Zhang H, Ren L, Guo L, et al. Health outcomes in people 2 years after surviving hospitalisation with COVID-19: a longitudinal cohort study. Lancet Respir Med 2022 Sep 16;10(9):863-876 [FREE Full text] [doi: 10.1016/S2213-2600(22)00126-6] [Medline: 35568052]
18. Nguyen HC, Nguyen MH, Do BN, Tran CQ, Nguyen TTP, Pham KM, et al. People with suspected COVID-19 symptoms were more likely depressed and had lower health-related quality of life: the potential benefit of health literacy. J Clin Med 2020 Mar 31;9(4):965 [FREE Full text] [doi: 10.3390/jcm9040965] [Medline: 32244415]
19. Zhang Y, Ma ZF. Impact of the COVID-19 pandemic on mental health and quality of life among local residents in Liaoning Province, China: a cross-sectional study. Int J Environ Res Public Health 2020 Mar 31;17(7):2381 [FREE Full text] [doi: 10.3390/ijerph17072381] [Medline: 32244498]
20. Qu G, Zhen Q, Wang W, Fan S, Wu Q, Zhang C, et al. Health-related quality of life of COVID-19 patients after discharge: a multicenter follow-up study. J Clin Nurs 2021 Jun 17;30(11-12):1742-1750 [FREE Full text] [doi: 10.1111/jocn.15733] [Medline: 33656210]
21. Arab-Zozani M, Hashemi F, Safari H, Yousefi M, Ameri H. Health-related quality of life and its associated factors in COVID-19 patients. Osong Public Health Res Perspect 2020 Oct;11(5):296-302 [FREE Full text] [doi: 10.24171/j.phrp.2020.11.5.05] [Medline: 33117634]

22. Clinical management of COVID-19. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2023.1 [accessed 2023-09-15]
23. KNGF-standpunt Fysiotherapie bij COVID-19: aanbevelingen voor fysiotherapeutisch handelen tijdens de ziekenhuisopname. Royal Dutch Society for Physical Therapy (KNGF). 2020. URL: https://www.kngf.nl/binaries/content/assets/kennisplatform/onbeveiligd/coronavirus/kngf_standpunt_fysiotherapie_covid-19_v3.0_22022022.pdf [accessed 2023-09-15]
24. Position statment COVID-19 version 1.4. Dutch Society for Speech-Language Therapy (NVLF). 2020. URL: https://www.nvlf.nl/wp-content/uploads/sites/2/2020/06/Factsheet-Intramurale-instellingen.pdf [accessed 2023-09-15]
25. Handreiking ergotherapie bij COVID-19 cliënten in de herstelfase. Dutch Society for Occupational Therapy (EN). 2021. URL: https://tinyurl.com/wzv35nca [accessed 2023-09-15]
26. Treatment plan of dietitian at COVID-19 after hospital discharge. The recovery phase after discharge from hospital: dietetics treatment in (Post) COVID-19 patients in a rehabilitation center. Dutch Society for Dietitians (NVD). 2021. URL: https://nvdietist.nl/artikelen/behandelplan-van-dietist-binnen-paramedische-herstelzorg-covid-19/ [accessed 2023-09-15]
27. De Bie RA, Verburg AC, Agasi-Idenburg C, Cup EHC, Dekker C, Van Dongen JM, et al. Evaluation of allied healthcare in patients recovering from Covid-19: study protocol and baseline data of a national prospective cohort study. J Rehabil Med 2022 Aug 26;54:jrm00309 [FREE Full text] [doi: 10.2340/jrm.v54.2506] [Medline: 35735900]
28. Interim clinical guidance for management of patients with confirmed coronavirus disease (COVID-19). Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html [accessed 2023-09-15]
29. Body mass index - BMI. World Health Organization. 2010 May 6. URL: https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations [accessed 2023-09-15]
30. Post MWM, van der Zee CH, Hennink J, Schafrat CG, Visser-Meily JM, van Berlekom SB. Validity of the Utrecht Scale for Evaluation of Rehabilitation-Participation. Disabil Rehabil 2012 Oct 06;34(6):478-485. [doi: 10.3109/09638288.2011.608148] [Medline: 21978031]
31. van der Zee CH, Baars-Elsinga A, Visser-Meily JM, Post MW. Responsiveness of two participation measures in an outpatient rehabilitation setting. Scand J Occup Ther 2013 May 14;20(3):201-208. [doi: 10.3109/11038128.2012.754491] [Medline: 23312022]
32. van der Zee CH, Kap A, Rambaran Mishre R, Schouten EJ, Post MWM. Responsiveness of four participation measures to changes during and after outpatient rehabilitation. J Rehabil Med 2011 Nov;43(11):1003-1009 [FREE Full text] [doi: 10.2340/16501977-0879] [Medline: 22031346]
33. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec 9;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
34. Zanini A, Aiello M, Adamo D, Casale S, Cherubino F, Della Patrona S, et al. Estimation of minimal clinically important difference in EQ-5D Visual Analog Scale score after pulmonary rehabilitation in subjects with COPD. Respir Care 2015 Jan 21;60(1):88-95 [FREE Full text] [doi: 10.4187/respcare.03272] [Medline: 25336531]
35. Valko P, Bassetti CL, Bloch KE, Held U, Baumann CR. Validation of the Fatigue Severity Scale in a Swiss cohort. Sleep 2008 Nov;31(11):1601-1607 [FREE Full text] [doi: 10.1093/sleep/31.11.1601] [Medline: 19014080]
36. Rooney S, McFadyen DA, Wood DL, Moffat DF, Paul PL. Minimally important difference of the Fatigue Severity Scale and modified Fatigue Impact Scale in people with multiple sclerosis. Mult Scler Relat Disord 2019 Oct;35:158-163. [doi: 10.1016/j.msard.2019.07.028] [Medline: 31400557]
37. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007 May;45(5 Suppl 1):S3-S11 [FREE Full text] [doi: 10.1097/01.mlr.0000258615.42478.55] [Medline: 17443116]
38. Terwee CB, Roorda LD, de Vet HCW, Dekker J, Westhovens R, van Leeuwen J, et al. Dutch–Flemish translation of 17 item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS). Qual Life Res 2014 Aug 9;23(6):1733-1741. [doi: 10.1007/s11136-013-0611-6] [Medline: 24402179]
39. Sandvall B, Okoroafor UC, Gerull W, Guattery J, Calfee RP. Minimal clinically important difference for PROMIS physical function in patients with distal radius fractures. J Hand Surg Am 2019 Jun;44(6):454-459.e1. [doi: 10.1016/j.jhsa.2019.02.015] [Medline: 30954311]
40. Snaith RP. The Hospital Anxiety and Depression Scale. Health Qual Life Outcomes 2003 Aug 01;1(1):29 [FREE Full text] [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
41. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
42. Lemay K, Tulloch HE, Pipe AL, Reed JL. Establishing the minimal clinically important difference for the Hospital Anxiety and Depression Scale in patients with cardiovascular disease. J Cardiopulm Rehabil Prev 2019 Nov;39(6):E6-E11. [doi: 10.1097/HCR.0000000000000379] [Medline: 30489438]
43. Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York, NY: Springer Sciences + Business Media, LLC; 2009.

44. Janssen B, Szende A. Population norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, editors. Self-Reported Population Health: An International Perspective based on EQ-5D. Dordrecht, the Netherlands: Springer; 2014:19-30.
45. Garrigues E, Janvier P, Kherabi Y, Le Bot A, Hamon A, Gouze H, et al. Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. J Infect 2020 Dec;81(6):e4-e6 [FREE Full text] [doi: 10.1016/j.jinf.2020.08.029] [Medline: 32853602]
46. Rass V, Beer R, Schiefecker AJ, Kofler M, Lindner A, Mahlknecht P, et al. Neurological outcome and quality of life 3 months after COVID-19: a prospective observational cohort study. Eur J Neurol 2021 Oct 03;28(10):3348-3359 [FREE Full text] [doi: 10.1111/ene.14803] [Medline: 33682276]
47. Vaes AW, Goërtz YMJ, Van Herck M, Machado FV, Meys R, Delbressine JM, et al. Recovery from COVID-19: a sprint or marathon? 6-Month follow-up data from online long COVID-19 support group members. ERJ Open Res 2021 Apr 26;7(2):00141-2021 [FREE Full text] [doi: 10.1183/23120541.00141-2021] [Medline: 34041295]
48. van den Borst B, Peters JB, Brink M, Schoon Y, Bleeker-Rovers CP, Schers H, et al. Comprehensive health assessment 3 months after recovery from acute coronavirus disease 2019 (COVID-19). Clin Infect Dis 2021 Sep 07;73(5):e1089-e1098 [FREE Full text] [doi: 10.1093/cid/ciaa1750] [Medline: 33220049]
49. Ceban F, Ling S, Lui LM, Lee Y, Gill H, Teopiz KM, et al. Fatigue and cognitive impairment in post-COVID-19 syndrome: a systematic review and meta-analysis. Brain Behav Immun 2022 Mar;101:93-135 [FREE Full text] [doi: 10.1016/j.bbi.2021.12.020] [Medline: 34973396]
50. Pavli A, Theodoridou M, Maltezou HC. Post-COVID syndrome: Incidence, clinical spectrum, and challenges for primary healthcare professionals. Arch Med Res 2021 Aug;52(6):575-581 [FREE Full text] [doi: 10.1016/j.arcmed.2021.03.010] [Medline: 33962805]
51. Shanbehzadeh S, Tavahomi M, Zanjari N, Ebrahimi-Takamjani I, Amiri-Arimi S. Physical and mental health complications post-COVID-19: scoping review. J Psychosom Res 2021 Aug;147:110525 [FREE Full text] [doi: 10.1016/j.jpsychores.2021.110525] [Medline: 34051516]
52. Alkodaymi MS, Omrani OA, Fawzy NA, Shaar BA, Almamlouk R, Riaz M, et al. Prevalence of post-acute COVID-19 syndrome symptoms at different follow-up periods: a systematic review and meta-analysis. Clin Microbiol Infect 2022 May;28(5):657-666 [FREE Full text] [doi: 10.1016/j.cmi.2022.01.014] [Medline: 35124265]
53. Ezzat MM, MD DAE, Elsherif AA. Prevalence of fatigue in patients post COVID-19. Eur J Mol Clin Med 2021;8(3):1330-1340.
54. Menges D, Ballouz T, Anagnostopoulos A, Aschmann HE, Domenghino A, Fehr JS, et al. Burden of post-COVID-19 syndrome and implications for healthcare service planning: A population-based cohort study. PLoS One 2021 Jul 12;16(7):e0254523 [FREE Full text] [doi: 10.1371/journal.pone.0254523] [Medline: 34252157]
55. González J, Benítez ID, Carmona P, Santisteve S, Monge A, Moncusí-Moix A, CIBERESUCICOVID Project (COV20/00110, ISCIII). Pulmonary function and radiologic features in survivors of critical COVID-19: a 3-month prospective cohort. Chest 2021 Jul;160(1):187-198 [FREE Full text] [doi: 10.1016/j.chest.2021.02.062] [Medline: 33676998]
56. Writing Committee for the COMEBAC Study Group, Morin L, Savale L, Pham T, Colle R, Figueiredo S, et al. Four-month clinical status of a cohort of patients after hospitalization for COVID-19. JAMA 2021 Apr 20;325(15):1525-1534 [FREE Full text] [doi: 10.1001/jama.2021.3331] [Medline: 33729425]
57. Renaud-Charest O, Lui LM, Eskander S, Ceban F, Ho R, Di Vincenzo JD, et al. Onset and frequency of depression in post-COVID-19 syndrome: a systematic review. J Psychiatr Res 2021 Dec;144:129-137 [FREE Full text] [doi: 10.1016/j.jpsychires.2021.09.054] [Medline: 34619491]
58. Mazza MG, Palladini M, De Lorenzo R, Magnaghi C, Poletti S, Furlan R, COVID-19 BioB Outpatient Clinic Study Group, et al. Persistent psychopathology and neurocognitive impairment in COVID-19 survivors: effect of inflammatory biomarkers at three-month follow-up. Brain Behav Immun 2021 May;94:138-147 [FREE Full text] [doi: 10.1016/j.bbi.2021.02.021] [Medline: 33639239]
59. Righi E, Mirandola M, Mazzaferri F, Dossi G, Razzaboni E, Zaffagnini A, et al. Determinants of persistence of symptoms and impact on physical and mental wellbeing in Long COVID: a prospective cohort study. J Infect 2022 Apr;84(4):566-572 [FREE Full text] [doi: 10.1016/j.jinf.2022.02.003] [Medline: 35150765]
60. Coronadashboard - varianten van het coronavirus. Rijksoverheid. 2022. URL: https://coronadashboard.rijksoverheid.nl/landelijk/varianten [accessed 2023-09-15]
61. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? Assess Eval High Educ 2008 Jun;33(3):301-314. [doi: 10.1080/02602930701293231]

Abbreviations

EQ-5D-5L: EuroQoL 5 Dimensions 5 Level
EQ-VAS: EuroQol Visual Analog Scale
FSS: Fatigue Severity Scale
HADS: Hospital Anxiety and Depression Scale
HRQoL: health-related quality of life

ICU: intensive care unit
PROMIS-PF-10b: Patient-Reported Outcomes Measurement Information System Physical Functioning Short Form 10b
USER-P: Utrecht Scale for Evaluation of Rehabilitation—Participation

Edited by A Mavragani, T Sanchez; submitted 08.11.22; peer-reviewed by J Dros, V Brahmbhatt; comments to author 06.06.23; revised version received 10.07.23; accepted 31.07.23; published 20.10.23.

Please cite as:
Slotegraaf AI, Gerards MHG, Verburg AC, de van der Schueren MAE, Kruizenga HM, Graff MJL, Cup EHC, Kalf JG, Lenssen AF, Meijer WM, Kool RA, de Bie RA, van der Wees PJ, Hoogeboom TJ, Dutch Consortium Allied Healthcare COVID-19
Evaluation of Primary Allied Health Care in Patients Recovering From COVID-19 at 6-Month Follow-up: Dutch Nationwide Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e44155
URL: https://publichealth.jmir.org/2023/1/e44155
doi:10.2196/44155
PMID:37862083

Original Paper

Association Among Household Water, Sanitation, and Hygiene (WASH) Status and Typhoid Risk in Urban Slums: Prospective Cohort Study in Bangladesh

Birkneh Tilahun Tadesse[1*], MD, PhD; Farhana Khanam[2*], MBBS; Faisal Ahmmed[2], MSc; Xinxue Liu[3], PhD; Md Taufiqul Islam[2], MBBS; Deok Ryun Kim[1], PhD; Sophie SY Kang[1], MSc; Justin Im[1], PhD; Fahima Chowdhury[2], MBBS; Tasnuva Ahmed[2], MBBS; Asma Binte Aziz[1], MBBS; Masuma Hoque[2], MBBS; Juyeon Park[1], MSc; Gideok Pak[1], MSc; Hyon Jin Jeon[1], MSc; Khalequ Zaman[2], PhD; Ashraful Islam Khan[2], PhD; Jerome H Kim[1], MD; Florian Marks[1,4,5], PhD; Firdausi Qadri[2], PhD; John D Clemens[1,2], MD

[1]Epidemiology, Public Health, Impact Unit, International Vaccine Institute, Seoul, Republic of Korea

[2]Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh

[3]Oxford Vaccine Group, Department of Pediatrics, University of Oxford, Oxford, United Kingdom

[4]Cambridge Institute of Therapeutic Immunology and Infectious Disease, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom

[5]Madagascar Institute for Vaccine Research, University of Antananarivo, Antananarivo, Madagascar

[*]these authors contributed equally

Corresponding Author:
Birkneh Tilahun Tadesse, MD, PhD
Epidemiology, Public Health, Impact Unit
International Vaccine Institute
SNU Research Park
1 Gwanak-ro, Gwanak-gu
Seoul, 08826
Republic of Korea
Phone: 82 1098041348
Email: birkneh.tadesse@ivi.int

Abstract

Background: Typhoid fever, or enteric fever, is a highly fatal infectious disease that affects over 9 million people worldwide each year, resulting in more than 110,000 deaths. Reduction in the burden of typhoid in low-income countries is crucial for public health and requires the implementation of feasible water, sanitation, and hygiene (WASH) interventions, especially in densely populated urban slums.

Objective: In this study, conducted in Mirpur, Bangladesh, we aimed to assess the association between household WASH status and typhoid risk in a training subpopulation of a large prospective cohort (n=98,087), and to evaluate the performance of a machine learning algorithm in creating a composite WASH variable. Further, we investigated the protection associated with living in households with improved WASH facilities and in clusters with increasing prevalence of such facilities during a 2-year follow-up period.

Methods: We used a machine learning algorithm to create a dichotomous composite variable (“Better” and “Not Better”) based on 3 WASH variables: private toilet facility, safe drinking water source, and presence of water filter. The algorithm was trained using data from the training subpopulation and then validated in a distinct subpopulation (n=65,286) to assess its sensitivity and specificity. Cox regression models were used to evaluate the protective effect of living in “Better” WASH households and in clusters with increasing levels of “Better” WASH prevalence.

Results: We found that residence in households with improved WASH facilities was associated with a 38% reduction in typhoid risk (adjusted hazard ratio=0.62, 95% CI 0.49-0.78; *P*<.001). This reduction was particularly pronounced in individuals younger than 10 years at the first census participation, with an adjusted hazard ratio of 0.49 (95% CI 0.36-0.66; *P*<.001). Furthermore, we observed an inverse relationship between the prevalence of “Better” WASH facilities in clusters and the incidence of typhoid, although this association was not statistically significant in the multivariable model. Specifically, the adjusted hazard of typhoid decreased by 0.996 (95% CI 0.986-1.006) for each percent increase in the prevalence of “Better” WASH in the cluster (*P*=.39).

Conclusions: Our findings demonstrate that existing variations in household WASH are associated with differences in the risk of typhoid in densely populated urban slums. This suggests that attainable improvements in WASH facilities can contribute to enhanced typhoid control, especially in settings where major infrastructural improvements are challenging. These findings underscore the importance of implementing and promoting comprehensive WASH interventions in low-income countries as a means to reduce the burden of typhoid and improve public health outcomes in vulnerable populations.

(JMIR Public Health Surveill 2023;9:e41207) doi:10.2196/41207

KEYWORDS

water; sanitation; sanitary; contaminated; contamination; hygiene; hygienic; WASH; water, sanitation and hygiene; typhoid fever; enteric fever; typhus; typhoid; enteric; salmonella; protection; recursive partitioning; Bangladesh; low- and middle-income countries; LMIC; bacteria; bacterial; bacterial infection; machine learning; algorithm; algorithms; model; low income; slum; slums; risk; infection control; incidence; prevalence; epidemiology; epidemiological; poverty

Introduction

Typhoid fever, also known as enteric fever, is a deadly febrile illness caused by the *Salmonella enterica* serovar Typhi (*Salmonella* Typhi) [1,2]. Based on 2019 data, approximately 9 million people contract typhoid annually, leading to 110,000 deaths [3]. Common symptoms include persistent fever, fatigue, headache, nausea, abdominal pain, and digestive issues, and occasionally a rash [2,3]. Severe cases of typhoid fever can have life-threatening complications and result in fatality [4]. The effectiveness of treating typhoid has been compromised by the growing resistance to various antibiotics [5,6]. To address this challenge, the World Health Organization advocates for the use of the typhoid conjugate vaccine (TCV) in children aged 6 months and older [3]. Nevertheless, the implementation of vaccination programs alone will not be enough to effectively control and eliminate the transmission of typhoid fever, particularly in urban slum areas where water, sanitation, and hygiene (WASH) facilities and practices are inadequate [7].

The primary mode of transmission for *Salmonella* Typhi is through the fecal-oral route via consumption of food and water that has been contaminated [2]. Consequently, inadequate access to clean water, improved sanitary facilities, appropriate waste management, and proper hand hygiene practices can contribute to an increased risk of typhoid fever [7,8]. For instance, successful elimination of typhoid fever in high-income countries is primarily attributed to major improvements in WASH facilities and practices [7,9]. However, substantial WASH modernization requires major financial, capital, and political commitment, and continues to be an unmet need in resource-constrained settings where unsafe drinking water and lack of sanitary facilities, augmented by unplanned urbanization, significantly contribute to increased typhoid transmission [10,11].

The need for multidimensional interventions, which now include vaccination using available TCVs, to prevent exposure to and transmission of typhoid fever demands implementation of feasible approaches to improving WASH, especially within densely populated urban settings. The emphasis on implementing practical enhancements in WASH is significant, yet remains relevant until a comprehensive upgrade of WASH practices and infrastructure is implemented in settings with high typhoid burden. Addressing the problem of typhoid fever in urban slums is particularly important as the highest burden has been reported in these settings including the occurrence of enhanced environmental transmission [12,13]. For example, in urban slums in Bangladesh, a high burden of typhoid fever (incidence of 200 per 100,000 person years of observation; PYO) and paratyphoid fever (incidence of 40 per 100,000 PYO) was reported [14]. The implementation of WASH practices has the potential to yield critical benefits, including the reduction of *Salmonella* Typhi transmission within communities. By mitigating the prevalence of this bacterium, the burden of bloodstream infections caused by *Salmonella* Typhi can be diminished. As a result, the consumption of antibiotics may be significantly reduced, contributing to the ongoing battle against antimicrobial resistance.

Our earlier work demonstrated that existing improvements in WASH facilities in a Kolkata slum were associated with reduced risk of typhoid (hazard ratio [HR]=0.57, 95% CI 0.37-0.90; *P*=.015) [8]. In this work, we developed a decision tree defining “Good” *versus* “Not Good” household WASH using 4 variables: source of drinking water, hand washing with soap, presence of flush toilet, and source of daily water use. The decision tree successfully discriminated between households at higher versus lower risk of typhoid (area under the curve [AUC] of 58%, 95% CI 54-61; sensitivity of 90.4%, 95% CI 84.2-94.8; and specificity of 19.8%, 95.5 CI 19.3-20.3). Notably, though, the data used for this analysis were accrued more than 20 years ago, possibly limiting their generalizability to circumstances in impoverished urban slums with endemic typhoid today.

Here we investigate whether existing improvements in WASH in a contemporary slum in Dhaka were also associated with a reduced risk of typhoid. We reexamined the hypothesis that variations in household WASH already present in urban slums can successfully predict the risk of typhoid fever. We test this hypothesis using the data from prospective, population-based, surveillance for typhoid fever in the control clusters of a cluster randomized trial of TCV recently completed in urban slums of Mirpur, Dhaka, Bangladesh [15].

Methods

Cluster Randomized Trial on the Vi-Tetanus Toxoid Vaccine

A cluster randomized trial (CRT) was conducted between May 2018 and March 2020 with the aim to evaluate the protection, including herd protection, conferred by Vi-tetanus toxoid conjugate vaccine (Vi-TT, Typbar-TCV, Bharat Biotech International Limited) when administered in a mass vaccination campaign. This participant- and observer-blinded CRT included children 9 months to <16 years within a censused total population of 205,760, where 150 clusters were randomized to receive Vi-TT or Japanese encephalitis (JE) vaccines (control clusters) in a 1:1 ratio. These clusters were approximately of comparable geographic size with mean area of 0.02 (SD 0.01) km^2 and included contiguous households with approximately equal total populations separated by natural borders wherever possible [7]. Age-eligible, nonpregnant residents who provided informed consent received a single dose of the assigned vaccine. The CRT was implemented in the Mirpur area in Bangladesh located in the north-east of Dhaka city. Mirpur is a densely populated settlement with a total surface area of 58.66 km^2 and approximately 3.5 million residents, serving as a field study site since 1987.

A baseline census was performed between February 14 and March 25, 2018, to enumerate the entire population in this study's area, except nonpermanent residents and individuals who planned to move out within a month. During the census, individual- and household-level socioeconomic and demographic data, details on water-sanitation-hygiene facilities and practices, and geo-positioning coordinates of each household were collected after obtaining informed consent. A "household" was defined as persons sharing the same cooking pot. The census was subsequently updated at approximately 6-month intervals to capture all births, deaths, and migrations. At each census, 13 nonbinary variables characterizing household WASH were collected. Because there were no significant differences between the household WASH status at first census and the subsequent updates by household, WASH characteristics at first census participation were used for this analysis. Accordingly, WASH variables collected during the baseline census were used to characterize household WASH status for individuals present in this study's area during the baseline census; the date of participation in the baseline census was taken as the start date of follow-up. For individuals who moved into the clusters after the baseline census and for births, the WASH variables collected during the first census update after the onset of their first participation were used to characterize household WASH status. The date of migration in and the date of birth were considered as the starting date of follow-up of these participants.

Typhoid Fever Surveillance

Surveillance for typhoid fever started on February 26, 2018, in all age groups using a common protocol across all 8 clinical facilities (Mirpur Field Clinic; Suhrawardy Hospital; Radda Barnen; Adhunik Hospital; Shishu hospital; Kurmitola General Hospital; Kingston Hospital; Islami Bank Hospital) serving this study's population. Biweekly home visits of the entire population were performed by community health workers to promote care seeking of all household members for febrile illnesses at one of this study's facilities. Blood samples for microbiological culture (3 mL for ≤17 years, and 5 mL for >17 years old) were collected for individuals presenting with history of fever ≥2 days or with axillary temperature of ≥38 °C, and clinical findings were systematically recorded by a study physician. Identity cards distributed during the censuses enabled identification of participants when presenting for febrile illness related care and treatment. For individuals not having the identification cards, computerized censuses on electronic tablets available at each study site were used to confirm identity. Blood specimens for culture were obtained and transported on the same day of collection to laboratories at the icddr,b. Blood cultures were monitored in an automated BacT/ALERT system, and subculturing was done on MacConkey, chocolate, and blood agar plates following a positive signal by the machine. Colony morphology, microscopic examination of Gram stain, biochemical tests, and slide agglutination tests were used for identification of *Salmonella* isolates. *Salmonella*-specific antisera including polyvalent O and O1, serogroup "O" *Salmonella* antigens [D], and flagellar "H" antigens [d] (Denka Sieken) were used for agglutination [16,17]. Treatment for patients with typhoid fever was based on the physicians' discretion, guided by antimicrobial susceptibility results. A single febrile episode was considered for febrile visits with the onset of fever ≤14 days after discharge for a previous febrile visit. A home visit was performed within 14 days of discharge for all participants with positive blood cultures to confirm that the person whose name was given on the day of presentation to the treatment center had indeed sought treatment on that date. A typhoid fever episode was defined as a febrile episode with a confirmed identity during which at least 1 blood culture yielded *S enterica* serotype Typhi (*Salmonella* Typhi).

Variable Selection to Define "Better" Versus "Not Better" WASH

As a first step, we partitioned each WASH variable collected during the censuses into a binary variable defining "Better" *versus* "Not Better" household WASH. The variables were partitioned based on substantive judgment aligned with the local context, and without prior knowledge of the typhoid incidence rates in participants with different categories of these variables. We then divided the households of total population in the JE clusters (75 clusters) at random into 2 mutually exclusive subpopulations: 98,087 (60%) of households constituting a "training" subpopulation and the other 65,286 (40%) a "validation" subpopulation. Variables associated with typhoid risk in bivariate Cox regression models at $P≤.20$ in the training subpopulation were selected for model development. Inclusion of all 13 WASH variables regardless of associations of the individual variables with the risk of typhoid, did not reveal any improvement in model performance. Proportionality assumptions for each variable were confirmed before inclusion in the bivariate models.

Development of the WASH Decision Tree to Predict Typhoid in the Training Subpopulation

As described elsewhere, we developed a composite WASH variable based on the selected candidate WASH variables to predict typhoid risk using a recursive partitioning-based algorithm in the training subpopulation [8,18]. In brief, we developed a binary WASH rule that predicts typhoid fever risk using individual WASH variables that were individually associated with the risk of typhoid fever and then conjoined as specified by the decision tree. Ultimately, the decision tree represents a composite WASH rule based on a combination of these WASH variables (Better versus Not Better WASH) predicting a classification of no typhoid versus typhoid, respectively. The decision tree then yielded several terminal nodes predictive of “higher” or “lower” risk of typhoid fever, which then served as alternative cutoffs used to create a dichotomous composite rule for “higher” or “lower” risk of typhoid, using the Youden index [12]. The WASH features in the decision tree predictive of “higher” risk of typhoid in this dichotomous typhoid variable then defined “Not Better” WASH and the features predictive of “lower” risk of typhoid in the dichotomous variable defined “Better” WASH.

The prediction rule was designed to discriminate the risk of typhoid fever in a dynamic population over the 2-year period of postvaccination follow-up in the trial. The algorithm was first evaluated in the training subpopulation assuming 1:1 ratio for a default loss of function of the cost of false positive and false negative classifications, with at least 300 observations at each terminal node. Subsequently, the algorithm was cross validated with estimation of the cross-validation error in 1 of 10 randomly assembled partitions of the training subpopulation. The 10-fold cross-validation technique was employed with the “*rpart*” package in the R software (R Foundation for Statistical Computing). We first constructed a tree using the training data, including terminal nodes. Then, we pruned the tree to obtain the smallest tree with the lowest misclassification error. The minimal complexity parameter was used to prune the model in finding the optimal prediction rule providing the minimum error with at least 2 terminal nodes in the tree. A receiver operating characteristic (ROC) curve was used to define the AUC for different cutoffs for dichotomizing the predictive rule, with the optimal cutoff for typhoid prediction selected as that which maximized the Youden Index, which includes sensitivity plus specificity, in relation to the proportion of participants in the terminal nodes developing typhoid fever [19].

Evaluation of the Composite WASH Variable in a Validation Subpopulation and Prediction of Typhoid Risk in the Total Population

To evaluate reproducibility of sensitivity and specificity of the composite WASH variable developed in the training subpopulation, we tested the variable in a distinct *validation* subpopulation. After ascertaining concordance of sensitivity (the proportion of participants developing typhoid fever who lived in households with “Not Better” WASH) and specificity (the proportion of participants not developing typhoid fever who lived in households with “Better” WASH) in predicting typhoid risk between the training and validation subpopulations, we then evaluated the ability of the variable to predict typhoid risk in the total population in the JE clusters.

Data Analysis

We constructed Cox proportional hazard regression models to evaluate the association between household WASH status and risk of typhoid in household inhabitants as a function of time to the first typhoid episode. Follow-up of persons in the baseline census began at the time of the census, and follow-up of those entering later as births or in-migrants began on the date of entry into the population. Follow-up was right censored by death, outmigration, and the end of this study. Before inclusion in the models, proportionality assumptions for all independent variables were assessed by a bivariate Cox model. The models were adjusted for randomization stratifying variables including geographic ward, longer than median distance to this study’s clinics, and number of eligible children residing in each cluster at baseline. As well, other baseline covariates, including age in years, number of participants in clusters, religion (Muslim), and monthly household expenditure (in Bangladeshi taka), were fitted as independent variables if they were associated with the hazard of typhoid in bivariate Cox models at $P<.15$. The average expenditure of households in a cluster was used as the monthly expenditure of households with unspecified monthly expenditure. The model coefficient for living in a “Better” WASH household was exponentiated to estimate the HR of typhoid fever, and the 95% CI for the HR was estimated using a robust sandwich method to adjust for the design effect of the clusters. Protection by Better WASH against typhoid fever was assessed as $[(1 – \text{HR}) \times 100\%)]$ with corresponding 95% interval limits for percentage protection. We also used Cox models to evaluate the protective association between the cluster prevalence of “Better” WASH households and the risk of typhoid in the total population during the follow-up period. The “Better” WASH cluster prevalence during the 2 years follow-up was calculated as the proportion of PYO contributed by members of households with “Better” WASH in the cluster divided by the total PYO for members of all households in the same cluster. We also assessed the protective association between increasing levels of “Better” WASH cluster prevalence, expressed in Cox models dimensionally, and the risk of typhoid after adjustment for the above-mentioned covariates. In these models, “Better” WASH coverage of the cluster was assigned as the “Better” WASH coverage datapoint for each resident in the respective cluster.

The analysis was performed using the *rpart* package for decision tree modelling, *rpart.plot* package for tree plotting, *pROC* package for the ROC curve, *survival* package for the Cox model, and *dplyr* package for data management under R Studio analytical software (R Foundation for Statistical Computing) [20-22]. The presence of collinearity in the multivariate models was assessed by inspection of variance inflation factors, with high variance inflation factors indicative of collinearity. For all the statistical analyses, a $P<.05$ (2-tailed) was taken as the margin of statistical significance.

Ethical Considerations

The trial has received ethical approval from the Research and Ethical Review Committees of icddr,b (PR-17115), as well as the institutional review boards of Oxford University and the University of Maryland. Informed written consent from all participants was obtained, with parental or guardian consent for those younger than the age of 16 years, and assent from participants aged 11 years to younger than 16 years. The trial was monitored by a local Data and Safety Monitoring Board established by icddr,b, as well as an international Data and Safety Monitoring Board established by the University of Maryland. This study's data are anonymous and were securely stored in access-controlled cabinets. The trial has been registered in the ISRCTN Registry under the identifier ISRCTN11643110.

Results

Assembly and Characteristics of Training and Validation Subpopulations

A total of 103,064 individuals living in 25,478 households within the control clusters for the trial were enumerated during the baseline census. In these households, 51,211 (49.7%) individuals were men, and the mean age was 26.8 (SD 17.3) years. During the 2 years of follow-up, 57,091 persons migrated into this study's area, 3218 were born, and 43,034 migrated out of this study's area or died, making the population residing in the JE clusters at the last census 120,339 persons, and the population ever followed in these clusters during the 2 years 163,373 persons (Figure 1). The training subpopulation comprised 98,087 individuals with a mean age 25.9 (SD 17.3) years, while the validation subpopulation included 65,286 individuals with a mean age 25.9 (SD 17.3) years (Table 1). Men were 48,638 (49.6%) in the training subpopulation and 32,496 (49.8%) in the validation subpopulation. During the 2-years of follow-up, a total of 435 first episodes of typhoid cases were detected, yielding an incidence of 209/100,000 PYO in the total population residing in the control clusters. The numbers of typhoid cases in the training and validation subpopulations were 260 and 175, respectively.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram of population dynamics during April 15, 2018, and March 15, 2020. JE: Japanese encephalitis; N: number; *S* Typhi: *Salmonella enterica* serovar Typhi.

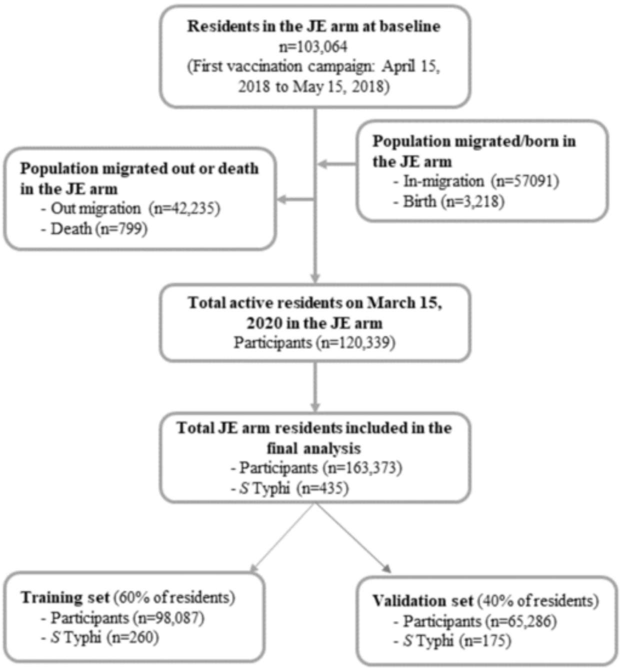


Table 1. Baseline characteristics of the total, training, and validation populations in the control clusters of the Vi-TT[a] cluster randomized trial, Dhaka, Bangladesh.

Characteristics	Full set	Training set	Validation set
Age (years), mean (SD)	25.9 (17.3)	25.9 (17.3)	25.9 (17.3)
Gender: men, n (%)	81,134 (49.7)	48,638 (49.6)	32,496 (49.8)
Religion: Muslim, n (%)	161,591 (98.9)	97,004 (98.9)	64,587 (98.9)
Monthly expenditure (BDT[b])[c], mean (SD)	17,420.3 (9053.2)	17,397.5 (7793)	17,454.6 (10,670.3)
Toilet facility: private, n (%)	68838 (42.1)	41155 (42)	27683 (42.4)
Adult toilet: flush toilet, n (%)	7804 (4.8)	4617 (4.7)	3187 (4.9)
Child toilet: flush toilet, n (%)	1000 (0.6)	580 (0.6)	420 (0.6)
Source of drinking water: private tap, well, or pump; water vendor, n (%)	63,263 (38.7)	37,854 (38.6)	25,409 (38.9)
Treated drinking water, n (%)	122,258 (74.8)	73,527 (75)	48,731 (74.6)
Treated cleaning water, n (%)	475 (0.3)	300 (0.3)	175 (0.3)
WASH[d] before meals, n (%)	121,595 (74.4)	73,019 (74.4)	48,576 (74.4)
Hand wash after defecation, n (%)	159,390 (97.6)	95,698 (97.6)	63,692 (97.6)
Handwashing water available in HH[e], n (%)	159,193 (97.5)	95,616 (97.5)	63,577 (97.4)
Handwashing soap available in HH, n (%)	161,443 (98.8)	96,985 (98.9)	64,458 (98.7)
Waste disposal place: fixed disposal, n (%)	155,458 (95.2)	93,406 (95.2)	62,052 (95)
Distance to drinking water source: shorter than median distance, n (%)	108,129 (66.2)	64,772 (66)	43,357 (66.4)
Water filter available in HH, n (%)	19,271 (11.8)	11,628 (11.9)	7643 (11.7)

[a] Vi-TT: Vi-tetanus toxoid.

[b] BDT: Bangladeshi taka; currency conversion rate: US $1 = BDT 110.48

[c] In total, 75 households had missing household expenditures, which were then substituted with the average household expenditure.

[d] WASH: waster, sanitation, and hygiene.

[e] HH: household.

Model Development, Training, and Validation

In the training subpopulation, having a private toilet (HR 0.57, 95% CI 0.43-0.77; *P*<.001); a safe source of drinking water, defined as a private tap, private well, bottled water, or water vendor (HR 0.65, 95% CI 0.49-0.87; *P*=.003); and availability of a water filter in the household during observation (HR 0.73, 95% CI 0.45-1.18; *P*=.20) were associated with typhoid risk among members of the household (Table 2). These 3 variables were used to build a binary composite prediction rule defining “Better” *versus* “Not Better” household WASH using the recursive partitioning algorithm (Figure 2).

Table 2. Bivariate associations of binary WASH[a] variables with typhoid risk in the training subpopulation of the control clusters of the Vi-TT[b] cluster randomized trial, Dhaka, Bangladesh.

Variables	Yes				No				Hazard ratio (HR[c])[d]	
	n	Cases	PY[e]	IR[f]/100,000 PY	n	Cases	PY	IR/100,000 PY	HR (95% CI)[g]	*P* value[g]
Toilet facility: private	41,155	86	57,718	149	56,932	174	67,381	258	0.57 (0.43-0.77)	<.001
Adult toilet: flush toilet	4617	11	6223	177	93,470	249	11,8876	209	1.07 (0.35-3.30)	.91
Child toilet: flush toilet	580	2	749	267	97,507	258	124,350	207	1.65 (0.41-6.69)	.48
Source of drinking water: private tap, well, or pump; water vendor	37,854	82	52,216	157	60,233	178	72,882	244	0.65 (0.49-0.87)	.003
Treated drinking water	73,527	183	92,350	198	24,560	77	32,748	235	0.91 (0.65-1.27)	.57
Treated cleaning water	300	1	420	238	97,787	259	124,678	208	1.13 (0.22-5.78)	.89
Hand wash before meals	73,019	191	92,584	206	25,068	69	32,514	212	1.15 (0.87-1.52)	.33
Hand wash after defecation	95,698	251	122,048	206	2389	9	3051	295	0.82 (0.48-1.40)	.47
Handwashing water available in HH[h]	95,616	251	121,736	206	2424	9	3312	272	0.94 (0.33-2.70)	.91
Handwashing soap available in HH	96,985	257	123,590	208	1102	3	1508	199	1.3 (0.56-3.03)	.537
Waste disposal place: fixed disposal	93,406	240	118,789	202	4681	20	6309	317	0.77 (0.43-1.37)	.37
Distance to drinking water source: shorter than median distance	64,772	175	83,175	210	33,315	85	41,923	203	1.11 (0.85-1.45)	.46
Water filter available in HH	11,628	24	15,966	150	86,459	236	109,132	216	0.73 (0.45-1.18)	.20

[a] WASH: water, sanitation, and hygiene.

[b] Vi-TT: Vi-tetanus toxoid.

[c] HR: hazard ratio.

[d] Estimated from extended Cox proportional hazards model.

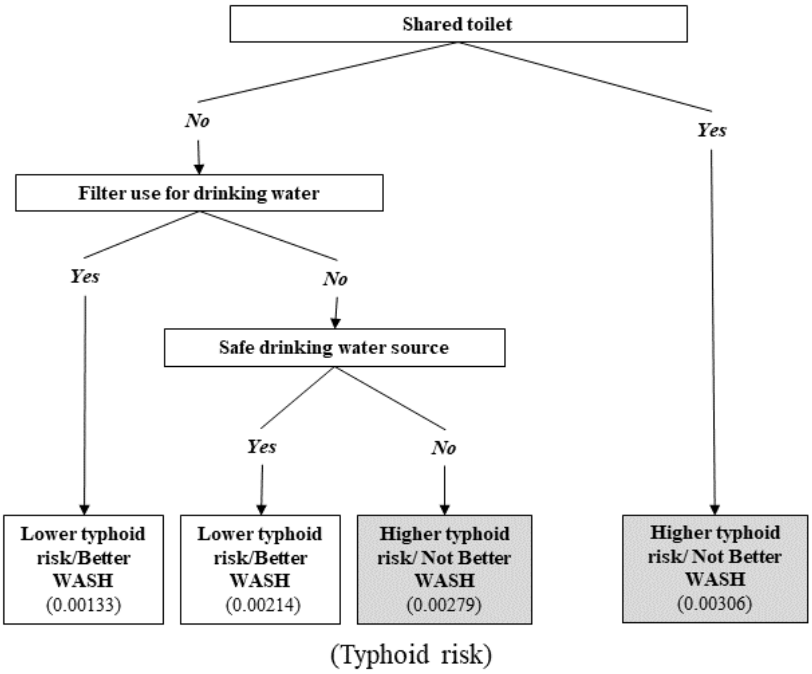
[e] PY: person years.

[f] IR: incidence rate.

[g] Calculated using robust SE assuming risk of typhoid is correlated within clusters.

[h] HH: household.

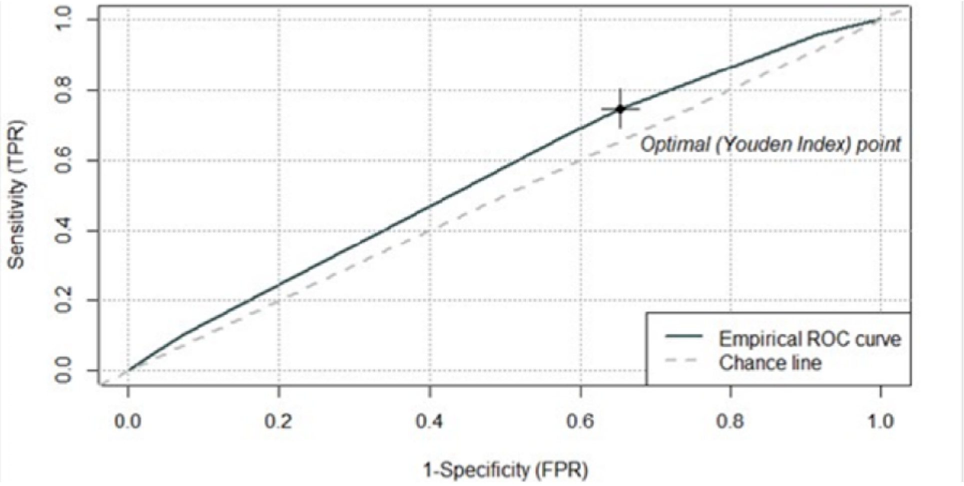
Figure 2. Decision tree defining "Better" and "Not Better" household WASH using 3 binary WASH variables collected at first census. The figures within the parentheses in the final boxes of the decision tree indicate the proportion of individuals who tested positive for typhoid fever. Out of 40,329 households, 18 (0.0446%) experienced 2 typhoid episodes. Among these, 13 out of 28,230 (0.046%) were from households categorized as Not Better WASH, while 5 out of 12,099 (0.0413%) were from households categorized as Better WASH. WASH: water, sanitation, and hygiene.



Performance of the Composite WASH Variable Predicting Typhoid Fever

The decision tree is shown in Figure 2. The dominant bifurcation for the prediction of typhoid in the tree was having a private toilet. In the training subpopulation, an optimal cutoff value of 0.0025 maximized the Youden index using the ROC curve, with an AUC of 56% (95% CI 53-59; Figure 3). With this threshold, the rule had a sensitivity of 74.6% (95% CI 68.9-80.0) and specificity of 34.7% (95% CI 34.4-35.0). Applying this rule to the validation subpopulation yielded comparable characteristics to those observed in the training subpopulation with a sensitivity and specificity of 74.9% (95% CI 67.8-81.1) and 35.2% (95% CI 34.9-35.6), respectively. The positive and negative predictive values were 0.3% (95% CI 0.26-0.35) and 99.8% (95% CI 99.7-99.9), respectively.

Figure 3. Receiver operator characteristic curve used to define the best cut off value to classify households as having "Better" versus "Not Better" WASH. AUC=56% (95% CI 53-59); best cutoff using the Youden index is 0.0025. AUC: area under the receiver operator characteristic curve; FPR: false positive rate; ROC: receiver operator characteristic; TPR, true positive rate; WASH: water, sanitation, and hygiene.



Prediction of Typhoid Incidence by Household WASH Status in the Total Population

We next applied the WASH prediction rule to the overall population residing in the control clusters to predict typhoid risk. The incidence of typhoid fever in all age groups living in "Better" WASH households was 139/100,000 PYO compared to 252/100,000 PYO in those living in "Not Better" WASH households, a reduction in typhoid risk of 38% (95% CI 22-51; *P*<.001) after adjusting for covariates. The reduction in typhoid risk was largely attributed to the protection associated with living in "Better" WASH households among children <10 years old (adjusted hazard ratio [aHR] 0.49, 95% CI 0.36-0.66; *P*<.001). Protective associations were aHR 0.35 (95% CI 0.20-0.61; *P*<.001) in children 2-4 years old, and aHR 0.58 (95% CI 0.36-0.92; *P*=.02) in 5-9 years old (Table 3 and Figure S1 in Multimedia Appendix 1).

The level of “Better” WASH cluster prevalence appeared inversely related to the log incidence of typhoid (Figure 4). However, the association between an increasing prevalence of “Better” WASH households in a cluster and typhoid risk was not significant in the adjusted multivariable model (aHR 0.996; 95% CI 0.986-1.006; *P*=.39) for each percent increase in the prevalence of “Better WASH” households in the cluster. The linear “Better” WASH cluster prevalence exhibited collinearity with household WASH status in the multivariable model, which precluded simultaneous estimation of the protective effects associated with living in “Better” WASH households and those associated with living in clusters with increasing prevalence of “Better” WASH households.

Table 3. Individual-level protection against typhoid associated with “Better” household WASH[a] in the total population residing in the control clusters of the Vi-TT[b] cluster randomized trial, Dhaka, Bangladesh, stratified by age at onset of follow-up.

Years	“Better” WASH				“Not Better” WASH				Hazard ratio (HR[c])[d]			
	n[e]	Cases	PY[f]	IR[g]/100,000 PY	N	Cases	PY	IR/100,000 PY	Crude HR[h]	*P* value[h]	Adjusted HR[h,i]	*P* value[h]
All	56,969	110	79,326	139	106,404	325	129,088	252	0.56 (0.44,0.71)	<.001	0.62 (0.49-0.78)	<.001
<2 years	3356	13	3883	335	6979	45	7211	624	0.54 (0.28,1.06)	.08	0.65 (0.32-1.28)	.21
2-4 years	3131	18	4424	407	6438	93	7919	1174	0.35 (0.22,0.57)	<.001	0.35 (0.20-0.61)	<.001
5-9 years	4894	26	7086	367	9819	84	12,586	667	0.57 (0.35,0.94)	.03	0.58 (0.36-0.92)	.02
10-14 years	4997	20	7278	275	9265	38	12,123	313	0.88 (0.5,1.56)	.66	0.92 (0.53-1.59)	.77
>15 years	40,591	33	56,655	58	73,903	65	89,250	73	0.76 (0.48,1.21)	.25	0.79 (0.50-1.24)	.31

[a] WASH: water, sanitation, and hygiene.

[b] Vi-TT: Vi-tetanus toxoid.

[c] HR: hazard ratio.

[d] Hazard ratios are estimated from extended Cox proportional hazards model.

[e] n: number; 100K: 100,000.

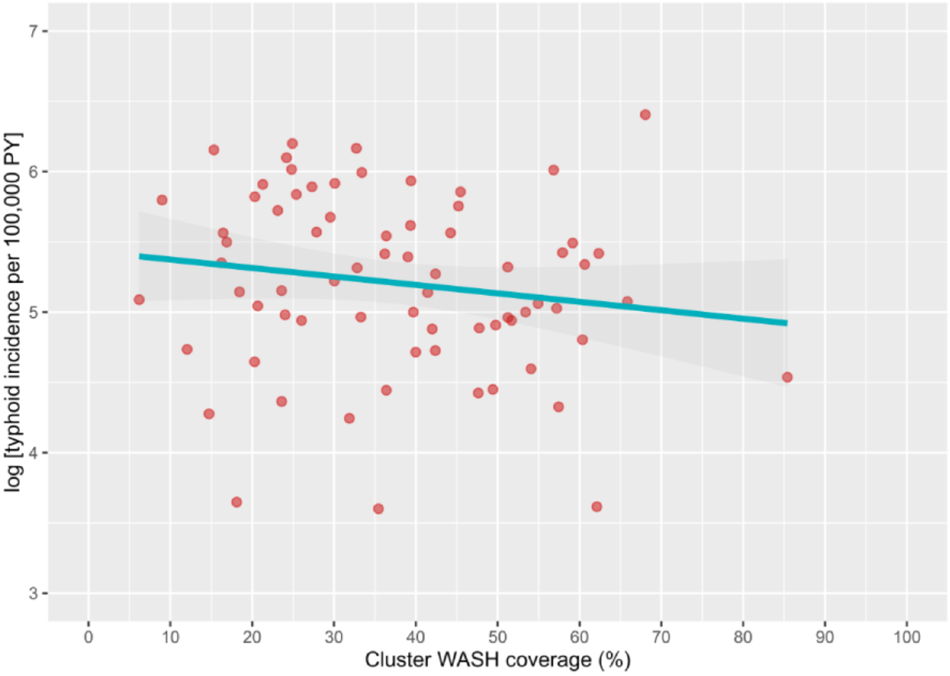
[f] PY: person years.

[g] IR: incidence rate.

[h] *P* values and CI are calculated using robust SE assuming risk of typhoid is correlated within clusters.

[i] Hazard ratio adjusted for the stratifying variables for randomization, including geographic ward, longer distance to study clinics than median distance, number of eligible children at baseline, and other baseline covariates age, Muslim religion; total monthly income; number of participants in a cluster.

Figure 4. Increasing levels of "Better" WASH cluster prevalence and the risk of typhoid in the total population residing in the control clusters of the Vi-TT cluster randomized trial. Out of 40,329 households, 28,230 (70%) were categorized as "Not Better" WASH households, and the remaining 12,099 (30%) fell under the "Better" WASH category. PY: person years; Vi-TT: Vi-tetanus toxoid; WASH: water, sanitation, and hygiene.



Discussion

We developed a composite WASH variable for a poor urban population in Dhaka using a machine learning approach, where 3 binary WASH variables associated with typhoid risk were combined to discriminate "Better" *versus* "Not Better" household WASH status. The variable predicted the risk of typhoid fever after adjusting for confounding variables including household socioeconomic status, with individuals in households with "Better" WASH having a 38% lower risk of typhoid than those in households with "Not Better" WASH, a reduction in risk that appeared primarily attributable to reduced risk in children younger than the age of 10 years. The higher risk reduction in the younger age group with improved WASH practices can be potentially explained by the elevated incidence of typhoid among this demographic [23,24]. Further, there may be other behavioral disparities between adults and children, such as children primarily staying at home and adults being more frequently outdoors for work-related activities, which could contribute to this discrepancy.

The WASH variables defining the composite household WASH rule in this analysis are somewhat different from those found to predict the risk of typhoid in an earlier analysis in Kolkata [8], except for quality of drinking water, which contributed to the definition of the composite WASH variable in both analyses. These differences are, however, anticipated owing to the differences between the 2 settings including sociocultural characteristics, data collection time, and the different data instruments used for characterizing WASH in the 2 studies. Furthermore, both studies were primarily designed to evaluate typhoid vaccine protection, and neither study attempted in-depth characterization of WASH with the data collection instruments, so component criteria of the composite variables are very likely to be markers of WASH facilities and practices, rather than the specific facilities and practices per se that account for a reduced risk of typhoid. However, associations with typhoid in both settings were independent of socioeconomic status, underscoring that it was likely that better WASH rather than greater affluence was responsible for the protective associations with typhoid.

Some limitations should be noted while interpreting the results. First, as noted earlier, we used WASH variables collected as part of the randomized clinical trial of vaccines and were not ascertained for determination of the balance of the randomized intervention arms rather than as in a specific study of WASH. Second, our initial dichotomization of WASH variables in the population was based on perceived clinical and public health significance, although such judgements were made without knowledge of typhoid outcomes. Third, our creation of a composite WASH variable for households in a dichotomous fashion was done to create a simplified assessment of household level WASH and cluster prevalence of "Better" WASH, but this could have led to a potential loss of information on certain WASH variables, which might make our analysis conservative. An additional factor potentially influencing data validity is the Hawthorne effect, arising from alterations in observed behaviors.

Despite these limitations, our analysis has several strengths. First, typhoid incidence data used in this analysis were collected using a prospective cohort with comprehensive surveillance of typhoid fever with repeated censuses and WASH assessments that allowed analysis of a dynamic cohort in a highly mobile population. The intensified community engagement and proactive house-to-house visits by community health workers have played a role in ensuring that the proportion of participants across all age groups who refrained from seeking care for febrile episodes did not result in notable variations in care-seeking practices. Second, the detailed demarcation of the clusters using natural structures, when possible, enabled unambiguous estimates of the “Better” WASH coverage in the clusters [6]. Third, external validation of the WASH prediction rule using a separate validation subpopulation revealed similar levels of sensitivity and specificity, suggesting that the prediction of typhoid by the composite variable was not exaggerated by overfitting the rule to the data used for derivation.

Simplified prediction rules, like the one presented in our study, offer great potential for various stakeholders involved in combating typhoid fever, including public health decision makers, researchers, and clinicians. These rules highlight the effectiveness of existing WASH adaptations, which can be implemented by most urban slum households to reduce susceptibility to typhoid in such environments. They can also be used to guide vaccine roll-out in high-risk populations and for community engagement initiatives aimed at increasing awareness of typhoid risk. Since the WASH data was collected from a diverse urban slum in Dhaka, the results can be extrapolated to other slums in Dhaka and similar regions with high typhoid burdens. Although further investigation is needed to address implementation barriers of “Better” WASH parameters, well-designed risk communication strategies can facilitate the adoption of such adaptations [25]. While not directly related to the clinical diagnosis and treatment of individual typhoid cases, WASH interventions play a crucial role in the broader strategy by potentially reducing antibiotic resistance. By decreasing antibiotic usage, these interventions contribute to preserving the efficacy of antibiotics, thus expanding the range of treatment options available for combating typhoid fever.

In conclusion, our findings suggest that improvements in WASH that already exist in this slum population without specific external intervention are associated with a reduced risk of typhoid and lend optimism to the notion that simple and inexpensive improvements of WASH to protect against typhoid are possible. Research to develop and evaluate such interventions should be pursued. As well, it will be important to assess how better existing WASH interacted with and perhaps modified protection by Vi-TT typhoid vaccine tested in this CRT. We are carrying out these analyses and will report them in future publications.

Acknowledgments

Funding for this study was from Bill and Melinda Gates Foundation (INV-025386).

Data Availability

The data and code will be available upon reasonable request to the corresponding author.

Authors' Contributions

JDC, FQ, FM, BTT, and FK conceptualized this study. Data analysis was done by FA, XL, JDC, BTT, FK, DRK, GP, JP, and MTI. Development of the first draft of this paper was by BTT, FK, MTI, XL, and JDC. JDC, FQ, FM, AIK, JHK, KZ, and DRK supervised for this paper. JI, SSYK, FC, TA, ABA, HJJ, and MH read, reviewed, and revised this paper. All the authors contributed to subsequent drafts. All the authors read and approved the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Kaplan-Meier curve displaying time to onset of typhoid cases in Better versus Not Better WASH households.

[PNG File , 83 KB - publichealth_v9i1e41207_app1.png]

References

1. GBD 2017 Typhoid and Paratyphoid Collaborators. The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis 2019;19(4):369-381 [FREE Full text] [doi: 10.1016/S1473-3099(18)30685-6] [Medline: 30792131]
2. Parry CM, Hien TT, Dougan G, White NJ, Farrar JJ. Typhoid fever. N Engl J Med 2002;347(22):1770-1782 [FREE Full text] [doi: 10.1056/NEJMra020201] [Medline: 12456854]
3. Typhoid. World Health Organization. Geneva; 2023. URL: https://www.who.int/news-room/fact-sheets/detail/typhoid [accessed 2023-11-07]
4. Marchello CS, Birkhold M, Crump JA. Complications and mortality of typhoid fever: a global systematic review and meta-analysis. J Infect 2020;81(6):902-910 [FREE Full text] [doi: 10.1016/j.jinf.2020.10.030] [Medline: 33144193]
5. Fatima M, Kumar S, Hussain M, Memon NM, Vighio A, Syed MA, et al. Morbidity and mortality associated with typhoid fever among hospitalized patients in hyderabad district, Pakistan, 2017-2018: retrospective record review. JMIR Public Health Surveill 2021;7(5):e27268 [FREE Full text] [doi: 10.2196/27268] [Medline: 33999000]

6. Vighio A, Syed MA, Hussain I, Zia SM, Fatima M, Masood N, et al. Risk factors of extensively drug resistant typhoid fever among children in Karachi: case-control study. JMIR Public Health Surveill 2021;7(5):e27276 [FREE Full text] [doi: 10.2196/27276] [Medline: 33973861]
7. Stanaway JD, Atuhebwe PL, Luby SP, Crump JA. Assessing the feasibility of typhoid elimination. Clin Infect Dis 2020;71(Suppl 2):S179-S184 [FREE Full text] [doi: 10.1093/cid/ciaa585] [Medline: 32725226]
8. Im J, Islam MT, Ahmmed F, Kim DR, Khan AI, Zaman K, et al. Can existing improvements of water, sanitation, and hygiene (WASH) in urban slums reduce the burden of typhoid fever in these settings? Clin Infect Dis 2021;72(11):e720-e726 [FREE Full text] [doi: 10.1093/cid/ciaa1429] [Medline: 32964216]
9. Brockett S, Wolfe MK, Hamot A, Appiah GD, Mintz ED, Lantagne D. Associations among water, sanitation, and hygiene, and food exposures and typhoid fever in case-control studies: a systematic review and meta-analysis. Am J Trop Med Hyg 2020;103(3):1020-1031 [FREE Full text] [doi: 10.4269/ajtmh.19-0479] [Medline: 32700668]
10. Breiman RF, Cosmas L, Njuguna H, Audi A, Olack B, Ochieng JB, et al. Population-based incidence of typhoid fever in an urban informal settlement and a rural area in Kenya: implications for typhoid vaccine use in Africa. PLoS One 2012;7(1):e29119 [FREE Full text] [doi: 10.1371/journal.pone.0029119] [Medline: 22276105]
11. Mogasale VV, Ramani E, Mogasale V, Park JY, Wierzba TF. Estimating typhoid fever risk associated with lack of access to safe water: a systematic literature review. J Environ Public Health 2018;2018:9589208 [FREE Full text] [doi: 10.1155/2018/9589208] [Medline: 30174699]
12. Luby SP. Urban slums: a supportive ecosystem for typhoidal salmonellae. J Infect Dis 2018;218(suppl_4):S250-S254 [FREE Full text] [doi: 10.1093/infdis/jiy324] [Medline: 30060082]
13. Akullian A, Ng'eno E, Matheson AI, Cosmas L, Macharia D, Fields B, et al. Environmental transmission of typhoid fever in an urban slum. PLoS Negl Trop Dis 2015;9(12):e0004212 [FREE Full text] [doi: 10.1371/journal.pntd.0004212] [Medline: 26633656]
14. Naheed A, Ram PK, Brooks WA, Hossain MA, Parsons MB, Talukder KA, et al. Burden of typhoid and paratyphoid fever in a densely populated urban community, Dhaka, Bangladesh. Int J Infect Dis 2010;14(Suppl 3):e93-e99 [FREE Full text] [doi: 10.1016/j.ijid.2009.11.023] [Medline: 20236850]
15. Theiss-Nyland K, Qadri F, Colin-Jones R, Zaman K, Khanam F, Liu X, et al. Assessing the impact of a Vi-polysaccharide conjugate vaccine in preventing typhoid infection among Bangladeshi children: a protocol for a phase IIIb trial. Clin Infect Dis 2019;68(Suppl 2):S74-S82 [FREE Full text] [doi: 10.1093/cid/ciy1107] [Medline: 30845333]
16. Khanam F, Sheikh A, Sayeed MA, Bhuiyan MS, Choudhury FK, Salma U, et al. Evaluation of a typhoid/paratyphoid diagnostic assay (TPTest) detecting anti-Salmonella IgA in secretions of peripheral blood lymphocytes in patients in Dhaka, Bangladesh. PLoS Negl Trop Dis 2013;7(7):e2316 [FREE Full text] [doi: 10.1371/journal.pntd.0002316] [Medline: 23951368]
17. Khanam F, Sayeed MA, Choudhury FK, Sheikh A, Ahmed D, Goswami D, et al. Typhoid fever in young children in Bangladesh: clinical findings, antibiotic susceptibility pattern and immune responses. PLoS Negl Trop Dis 2015;9(4):e0003619 [FREE Full text] [doi: 10.1371/journal.pntd.0003619] [Medline: 25849611]
18. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods 2009;14(4):323-348 [FREE Full text] [doi: 10.1037/a0016973] [Medline: 19968396]
19. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation 2007;115(5):654-657 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.105.594929] [Medline: 17283280]
20. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77 [FREE Full text] [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
21. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, US: Springer; 2000.
22. Wickham H, François R, Henry L, Müller K, Vaughan D, Posit Software, PBC. dplyr: a grammar of data manipulation. R package version 0.7.6. 2018. URL: https://cran.r-project.org/web/packages/dplyr/index.html [accessed 2023-11-07]
23. Garrett DO, Longley AT, Aiemjoy K, Yousafzai MT, Hemlock C, Yu AT, et al. Incidence of typhoid and paratyphoid fever in Bangladesh, Nepal, and Pakistan: results of the Surveillance for enteric fever in Asia project. Lancet Glob Health 2022;10(7):e978-e988 [FREE Full text] [doi: 10.1016/S2214-109X(22)00119-X] [Medline: 35714648]
24. Marks F, von Kalckreuth V, Aaby P, Adu-Sarkodie Y, El Tayeb MA, Ali M, et al. Incidence of invasive salmonella disease in sub-Saharan Africa: a multicentre population-based surveillance study. Lancet Glob Health 2017;5(3):e310-e323 [FREE Full text] [doi: 10.1016/S2214-109X(17)30022-0] [Medline: 28193398]
25. Barac R, Als D, Radhakrishnan A, Gaffey MF, Bhutta ZA, Barwick M. Implementation of interventions for the control of typhoid fever in low- and middle-income countries. Am J Trop Med Hyg 2018;99(3_Suppl):79-88 [FREE Full text] [doi: 10.4269/ajtmh.18-0110] [Medline: 30047365]

Abbreviations

aHR: adjusted hazard ratio
AUC: area under the curve
CRT: cluster randomized trial
HR: hazard ratio
JE: Japanese encephalitis
PYO: person years of observation
ROC: receiver operating characteristic
TCV: typhoid conjugate vaccine
Vi-TT: Vi-tetanus toxoid
WASH: water, sanitation, and hygiene

Edited by A Mavragani, T Sanchez; submitted 18.07.22; peer-reviewed by B Padhi, S Bhattacharya, M Raimi, A Allam, W Klement, K Robb; comments to author 25.07.23; revised version received 30.08.23; accepted 05.10.23; published 20.11.23.

Please cite as:
Tadesse BT, Khanam F, Ahmmed F, Liu X, Islam MT, Kim DR, Kang SSY, Im J, Chowdhury F, Ahmed T, Aziz AB, Hoque M, Park J, Pak G, Jeon HJ, Zaman K, Khan AI, Kim JH, Marks F, Qadri F, Clemens JD
Association Among Household Water, Sanitation, and Hygiene (WASH) Status and Typhoid Risk in Urban Slums: Prospective Cohort Study in Bangladesh
JMIR Public Health Surveill 2023;9:e41207
URL: https://publichealth.jmir.org/2023/1/e41207
doi:10.2196/41207
PMID:37983081

Original Paper

Effects of the Implementation of Transport-Driven Poverty Alleviation Policy on Health Care–Seeking Behavior and Medical Expenditure Among Older People in Rural Areas: Quasi-Experimental Study

Yuanyang Wu[1], PhD; Qianning Wang[1], MA; Feiyang Zheng[1], PhD; Tiantian Yu[1], MA; Yanting Wang[1], MA; Si Fan[1], MA; Xinping Zhang[1*], Prof Dr; Lianping Yang[2,3,4*], Prof Dr

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2]School of Public Health, Sun Yat-sen University, Guangzhou, China
[3]Sun Yat-Sen Global Health Institute, Institute of State Governance, Sun Yat-Sen University, Guangzhou, China
[4]Institute for Global Health and Development, Peking University, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Xinping Zhang, Prof Dr
School of Medicine and Health Management
Tongji Medical College
Huazhong University of Science and Technology
No.13 Hangkong Road
Wuhan, 430030
China
Phone: 86 18071509979
Email: xpzhang602@hust.edu.cn

Abstract

Background: Improving the rural residents' accessibility to and affordability of health care is recognized as a common target globally. The Health in All Policies approach, from the Declaration of Helsinki to the United Nations' *Decade Of Healthy Ageing*, strengthened the far-reaching effect of large-scale public policies on health care–seeking behavior; however, the effects of national transport policy on health care–seeking behavior is unclear.

Objective: This quasi-experimental study aimed to examine the effects of the implementation of transport-driven poverty alleviation (TPA) policy on health care–seeking behavior and medical expenditure among older adults in rural areas and the mechanism underlying these effects.

Methods: We designed a quasi-experiment to estimate the effects of TPA policy implementation on health care–seeking behavior and medical expenditure among older adults in rural areas through a difference-in-differences (DID) analysis based on data from the China Health and Retirement Longitudinal Study in 2011, 2013, 2015, and 2018. The underlying mechanism was analyzed and effect modification patterns were further investigated by poor households, health status, and age.

Results: Our findings validated a positive contribution of TPA policy on health care–seeking behavior among older adults in rural areas. After the implementation of TPA policy, the number of inpatient visits increased by annually 0.35 times per person, outpatient medical expenditure increased by 192% per month, and inpatient medical expenditure increased by 57% annually compared with those of older adults in rural areas without the implementation of TPA policy. Further, there was a significant modification effect, with a positive effect among poor households, healthier older adults, and those aged 60-80 years. Additionally, the policy improved the patients' capabilities to seek long-distance care (β=23.16, 95% CI –0.99 to 45.31) and high-level hospitals (β=.08, 95% CI –0.02 to 0.13), and increased individual income to acquire more medical services (β=4.57, 95% CI –4.46 to 4.68).

Conclusions: These findings validate the positive contribution of TPA policy on health care–seeking behavior among older adults in rural areas; however, the medical expenditure incurred was also high. Concerted efforts are needed to address health care–seeking dilemmas in rural areas, and attention must be paid to curbing medical expenditure growth for older adults in rural areas during TPA policy implementation.

(JMIR Public Health Surveill 2023;9:e49603) doi:10.2196/49603

KEYWORDS

transport-driven poverty alleviation; health care–seeking behavior; medical expenditures; difference-in-differences; quasi-experimental study

Introduction

Poor health care access and high medical expenditure in rural areas are major causes of inequalities and global health burden [1]. Growing inequalities between rural and urban areas in health care use and health outcomes have been reported in globally—for example, rural residents who make up about 80% of the population of 1.4 billion individuals enjoyed about 20% of the national health resources in China—and high-level hospitals are scarce in rural areas [2,3]. Improving the rural residents' accessibility to and affordability of health care is recognized as a common target in both high-income and transitional counties [4].

It was a milestone that the Health in All Policies (HiAP) approach proposed by the Declaration of Helsinki placed emphasis on social determinants of health, in which transport was recognized as a determinant in health care–seeking behavior by influencing geographic accessibility and spatial behavior [5]. Previous studies paid substantial attention to emergency medical transportation services and road traffic mortality [6,7].

Andersen's health behavior model pointed out that the use of medical services was affected by individual predisposing characteristics, enabling resources, and needs [8]. Arcury et al [5] further added geographic accessibility and spatial behavior in the health behavior model, including distance, transportation availability, and activity space. The distance decay effect showed that the interaction between the 2 sites decreased with an increase in the distance between them, and this effect of distance was also observed on the use of medical services [9]. Zielinski et al [10] found that patients living further than 40 km of a hospital had lower use of secondary health care services based on the data from the Care Data Warehouse in Östergötland, Sweden [10]. From the perspective of transportation availability, Arcury et al [5], using survey data from a sample of 1059 households located in 12 western North Carolina counties, found that those who had a driver's license had 2.29-fold more health care visits for chronic care and 1.92-fold more visits for regular checkup care than those without a driver's license [11]. In terms of activity space, the development of transportation expanded the scope of residents' activities and facilitated access to remote medical services. Liu et al [12] found that after the opening of high-speed railway, Chinese patients were more likely to go to areas with a high density of health care resources. Rural residents were willing to travel long distances to receive quality services, which was not available locally or when their primary care facilities were located in urban centers [13].

Poor transport infrastructure contributes profoundly to poverty and economic burdens. Taking into account the needs of transport infrastructure in impoverished areas, the Ministry of Transport of China issued the "13th Five-Year Plan for Transport Poverty Alleviation" (hereinafter called "The Plan"), which is committed to strengthening the transport infrastructure in impoverished areas in 26 provinces of China [14]. The Plan aimed to further improve the construction of transport infrastructure in impoverished areas, enhance the capacity and level of transport services, and strengthen the capacity for transport safety and security, with a view to comprehensively resolve the backwardness of transport infrastructure in impoverished areas by 2020. A detailed description of the Transport-driven Poverty Alleviation (TPA) policy is showed in Multimedia Appendix 1.

Better transport has been associated with low transport costs and individual income growth, which are key factors affecting health care–seeking behavior. On the one hand, better transport mitigated the distance decay effect in medical service usage. Time, money, and energy spent for seeking medical treatment were regarded in determining the transport cost, and increased the price of medical services [15]. Alam et al [16] found that transport costs were an important part of medical expenditures by comparing different access to health services during pregnancy, delivery, and the postpartum period among women. In addition to direct economic costs, long-distance travel led to greater pain and discomfort and unexpected physical costs [17]. For example, when the distance to a hospital increased, residents' health care usage significantly decreased, and the probability of outpatient care also decreased [9]. In contrast, convenient transport generated many social benefits, including lesser travel time and medical expenses and high-quality medical resources.

On the other hand, transport promoted economic development and increased personal income in rural areas due to the fundamental role of transport infrastructure. Income growth in rural areas after the implementation of TPA policy improved the rural residents' ability to receive health care services, such as higher-income groups in China having a better health status and using more health care services [18].

In summary, high-credibility empirical evidence is needed urgently to implement the HiAP approach to achieve Millennium Development Goals, even though few studies have verified the relationship between transport and health care–seeking behavior. TPA policy was implemented in China, where the fight against poverty made a comprehensive success with 98.99 million poor households, 832 poor counties, and 128,000 poor villages being lifted out of poverty. This study aimed to explore the effect of TPA policy on health care–seeking behavior and medical expenditure among older adults in rural areas by using data from the China Health and Retirement Longitudinal Study (CHARLS).

Methods

Study Design

This is a quasi-experimental design study, and the policy's effect was examined using a difference-in-differences (DID) method with the 2011 and 2018 CHARLS panel data. Differences in health care–seeking behavior and medical expenditure before and after TPA policy implementation were analyzed.

Participants

For randomization, participants aged 45 years and older were selected randomly from 28 provinces, 150 counties, 450 communities (villages), and 12,400 households. Within each selected household, one resident aged 45 years or older was selected randomly for the sample and their spouse was automatically included. If nobody aged 45 years or older lived in the selected household, the household was skipped. Participants aged 60 years and older or those who had all variables in our study based on the 4 waves of the CHARLS were included.

Intervention and Control Groups

The intervention was the TPA policy implemented in 2016 [14], which was measured by a dummy variable indicating whether the TPA policy is implemented or not and rural road mileages. The treatment group comprised participants in provinces covered by the policy, such as Yunnan and Xinjiang Provinces, and the control group comprised participants in provinces not covered by the policy, such as Shandong Province (details of the national policy are shown in Multimedia Appendix 1).

Outcomes

The primary outcomes were health care–seeking behaviors and medical expenditure. Health care–seeking behaviors included the number of outpatient visits in last month and inpatient visits in the last year. Medical expenditure measures included total expenditures for outpatient care in the last month and total expenditures for inpatient care in last year, measured in Chinese Renminbi (Yuan).

Data Sources

The data were mainly obtained from the 2011, 2013, 2015, and 2018 waves of CHARLS, which is an ongoing prospective cohort study on the determinants of healthy ageing in members of the population aged 45 years and older. The CHARLS was led by the National Development Research Institute of Peking University and jointly executed by the China Social Sciences Survey Center of Peking University and the Youth League Committee of Peking University. Its baseline survey encompassed 450 villages and communities nationwide, with a sample size of 17,708 individuals in 10,257 households. The CHARLS uses a systematic random sampling strategy to select respondents. First, district and county units are implicitly stratified by region, rural or urban areas, and gross domestic product per capita. Based on the probability proportional to population size, 150 district and county units were randomly chosen among all county units, and 3 village or community units were further randomly selected within each county unit. The team developed a special mapping software to help draw a sample frame of all households in each unit. Within each household selected from the mapping frame, 1 resident aged 45 years or older was randomly chosen to be the main respondent, and this person's spouse was automatically included in the sample. To avoid human error and manipulation, each sampling stage was computerized, and all interviews were conducted using computer-aided personal interview technology [19]. The 4 waves used the same ascertainment and assessment protocols. The rural road mileages measuring TPA policy implementation were obtained from 2011 to 2018 from the *China Transport Statistical Yearbook*.

Information collected mainly included basic sociodemographic status (ie, age, sex, marital status, education, and economic status), health status (ie, self-reported ability to perform basic activities of daily living and the number of chronic diseases), daily behavior habits (ie, smoking and drinking), access to health care services (ie, distance to outpatient or inpatient care, and hospital level of outpatient or inpatient care), health care–seeking behavior (ie, outpatient and inpatient visits), and medical expenditure (ie, outpatient or inpatient expenditure).

The basic sociodemographic status, health status, and daily habit factors were selected as covariates. A detailed definition of covariates is shown in Table 1.

Table 1. Characteristics of the participants in the intervention and control groups from 2011 to 2018.

Variables	Total sample	Control group	Intervention group
Age (years), mean (SD)	68.38 (6.89)	68.82 (7.27)	68.28 (6.80)
Sex (male), n (%)	3333 (49.71)	538 (48.59)	2778 (49.65)
Female, n (%)	3372 (50.29)	570 (51.41)	2819 (50.35)
Marital status (married), n (%)	5475 (81.67)	926 (83.60)	4545 (81.22)
Divorced or never married, n (%)	1230 (18.33)	182 (16.40)	1052 (18.79)
Years of education, mean (SD)	4.18 (4.11)	4.01 (4.19)	4.21 (4.09)
Number of chronic diseases, mean (SD)	1.17 (1.31)	1.03 (1.24)	1.20 (1.33)
Disability status, n (%)			
Yes	780 (11.64)	112 (10.16)	668 (11.95)
No	5925 (88.36)	996 (89.83)	4929 (88.05)
Smoking status, n (%)			
Yes	1139 (16.99)	176 (15.90)	965 (17.25)
No	5566 (83.01)	932 (84.10)	4632 (82.75)
Drinking status, n (%)			
Yes	2156 (32.16)	379 (34.26)	1772 (31.66)
No	4549 (67.84)	729 (65.74)	3825 (68.34)
Economic status[a], mean (SD)	5.26 (3.44)	6.03 (3.53)	5.08 (3.39)
Distance to outpatient care (km), mean (SD)	15.31 (104.93)	10.68 (31.64)	16.11 (112.85)
Hospital level for outpatients, n (%)			
County	5070 (75.62)	760 (68.66)	4337 (77.50)
Prefectural	1175 (17.53)	244 (22.05)	912 (16.31)
Provincial	460 (6.85)	104 (9.29)	348 (6.19)
Distance to inpatient care (km), mean (SD)	41.85 (173.93)	27.18 (54.18)	44.42 (187.06)
Hospital level for inpatients, n (%)			
Country	5083 (75.82)	788 (71.17)	4295 (76.75)
Prefectural	1186 (17.70)	229 (20.72)	963 (17.21)
Provincial	436 (6.39)	91 (8.11)	339 (6.04)
Individual income (yuan; log-transformed), mean (SD)	2.82 (3.58)	3.07 (3.85)	2.76 (3.51)
Observations, n	6705	1108	5597

[a]Logarithmic value of annual per capita income.

Statistical Analysis

Overview

The effect of TPA policy on health care–seeking behavior, including outpatient visits (model 1) and inpatient visits (model 3), and medical expenditure, including outpatient expenditure (model 2) and inpatient expenditure (model 4), were estimated using 4 DID models [20] in which outpatient and inpatient expenditure were in logarithmic form.

A binary policy indicator variable *treat* was used to decide between intervention and control, and a binary time variable *post* was used to decide between pre- and postintervention. The coefficient of product the terms *treat* and *post* was the policy effect. The specific model was set as follows:

$$Healthservice\_utilization_{it} = \alpha_0 + \beta_1 treat_{it} + \beta_2 post_{it} + \beta_3 treat \times post_{it} + \beta_4 control_{it} + \varepsilon_{it} \quad (1)$$

where $treat_{it}$ is the policy indicator dummy variable, indicating that the participant is covered by the TPA policy, during period *t* (the treatment group), valued at 1, and those not covered by TPA policy (the control group), valued at 0. $post_{it}$ is a time dummy variable, and the value is 1 after issuing the TPA policy (the 2018 wave) and 0 before issuing of the TPA policy (the 2011, 2013, and 2015 waves). $treat_{it} \times post_{it}$ is the product term of $treat_{it}$ and $post_{it}$. $control_{it}$ is a series of control variables, β_1,

β_2, β_3, and β_4 are the regression coefficients, where β_3 is the effect of TPA policy, and ε_{it} is the random error.

The mechanism underlying the effects of TPA policy implementation on health care–seeking behavior and medical expenditure was analyzed using 5 DID models: regression models to determine the association between traveling distance to outpatient visits and TPA policy implementation (model 5), hospital level of outpatient visits and TPA policy implementation (model 6), traveling distance to inpatient care and TPA policy implementation (model 7), hospital level of inpatient visits and TPA policy implementation (model 8), and individual income and TPA policy implementation (model 9).

Effect Modification Analysis

To explore possible effect modifications by economic status (ie, poor household: yes vs no), self-reported health status (ie, healthy and unhealthy), and age (ie, around 60-70, 70-80, and ≥80 years), we performed effect modification analysis stratified by these potential modifiers (Multimedia Appendix 2). According to the Andersen’s health behavior model, age was the individual predisposing characteristic, economic status was the enabling resource, and health status was the need factor, which impact health care usage behavior [21,22].

Sensitivity Analyses

First, the independent variable of TPA policy implementation was replaced by rural road mileage, and the effects of rural road mileage on health care–seeking behavior and on medical expenditure were examined by using a fixed-effects model. Second, the validity of the identification assumption of the DID model was verified by temporal trends in outpatient visits, inpatient visits, and expenditure of outpatients and inpatients, respectively, from 2011 to 2018. Simultaneously, a propensity score matching method was used to overcome the differences between the control and intervention groups, and then a weighted DID estimation was performed after matching to adjust the sampling. Finally, the dependent variables were replaced with out-of-pocket medical expenditure to exclude the effect of health insurance, because out-of-pocket medical expenditure was not reimbursed by health insurance. Furthermore, the regression analysis was restricted to older adults without pension to avoid the estimation bias from pensions. We further performed a placebo test to best attempt to overcome interferences from other poverty alleviation policies by adjusting the policy time to 2015.

Ethical Considerations

The studies involving human participants were reviewed and approved by the Research Ethics Committees of Peking University (IRB00001052-11015) [23]. This survey was anonymous and the answers are protected by privacy laws. Written informed consent clarifying the study purposes was obtained from each participant before completing the interview.

Results

In total, 6705 participants (mean age 68.38 years; 3333, 49.71% male; 5475, 81.67% married) were included. Of them, 5597 (83.48%) and 1108 (16.52%) participants were in intervention and control groups, respectively. Participants in both groups had similar covariate distributions.

Their mean number of outpatient and inpatient visits was 2.37 (SD 2.66) and 1.57 (SD 1.34) in 1 month and 1 year, respectively. The mean log-transformed outpatient and inpatient expenditures were 5.84 (SD 1.49) and 9.20 (SD 1.28), respectively. Mean inpatient visits increased from 1.47 (SD 1.11) to 1.61 (SD 1.79) in the control group and from 1.53 (SD 1.24) to 1.70 (SD 1.52) in intervention group. Mean inpatient expenditure increased from 9.72 (SD 1.14) to 9.82 (SD 1.31) in the control group and from 8.96 (SD 1.27) to 9.39 (SD 1.23) in treatment group (Multimedia Appendix 3).

Table 2 shows the effect of TPA policy implementation on outpatient and inpatient visits of older adults in rural areas with controlling potential confounders. After the implementation of TPA policy, the number of inpatient visits increased by 0.35 times annually, the outpatient medical expenditure increased by 192% per month, and the inpatient medical expenditure increased by 57% annually among the older adults in rural areas. Other control variables also influenced health care–seeking behavior and medical expenditure. Women had higher outpatient expenditures but lower inpatient visits than men. Higher the prevalence of chronic diseases among older people, greater the number of outpatient visits, outpatient expenditures, inpatient visits, and inpatient expenditures. Similarly, disabled older adults or those with a better economic status had more outpatient and inpatient visits. However, older adults who smoked and drank had lower outpatient expenditures and fewer inpatient visits.

Table 2. Regression models to analyze the associations of Transport-driven Policy Alleviation (TPA) policy implementation with health care–seeking behavior and medical expenditure among older adults from 2011 to 2018. All models have time-fixed effects.

Variables	Model 1 (2779 observations; R^2=0.007)		Model 2 (5131 observations; R^2=0.233)		Model 3 (2235 observations; R^2=0.024)		Model 4 (699 observations; R^2=0.075)	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
TPA	–.11 (–0.38 to 0.15)	.40	1.92 (1.58 to 2.26)	<.001	.35 (0.21 to 0.50)	<.001	.57 (0.34 to 0.79)	<.001
Age	–.003 (–0.02 to 0.01)	.73	.01 (–0.01 to 0.03)	.34	.003 (–0.01 to 0.01)	.45	–.01 (–0.03 to 0.01)	.20
Sex (female)	.04 (–0.22 to 0.31)	.75	.50 (0.20 to 0.80)	<.001	–.18 (–0.34 to 0.03)	.05	–.08 (–0.32 to 0.16)	.52
Marital status (married)	–.05 (–0.31 to 0.21)	.70	–.03 (–0.32 to 0.27)	.87	–.05 (–0.20 to 0.11)	.16	.10 (–0.14 to 0.35)	.41
Years of education	–.02 (–0.05 to 0.01)	.11	.13 (0.09 to 0.16)	<.001	.003 (–0.01 to 0.02)	.48	.02 (–0.01 to 0.04)	.13
Chronic diseases	.08 (0.01 to 0.15)	.03	.02 (–0.06 to 0.10)	.60	.07 (0.03 to 0.11)	.01	.004 (–0.06 to 0.07)	.88
Disability (yes)	.34 (0.04 to 0.63)	.03	.26 (–0.07 to 0.58)	.13	.14 (–0.01 to 0.30)	.01	.50 (0.26 to 0.75)	<.001
Smoking (yes)	–.01 (–0.34 to 0.32)	.95	–.87 (–1.22 to –0.52)	<.001	–.27 (–0.49 to –0.04)	.05	–.21 (–0.60 to 0.18)	.30
Drinking (yes)	–.21 (–0.49 to 0.06)	.13	–.24 (–0.55 to 0.07)	.13	–.20 (–0.37 to –0.03)	.64	–.13 (–0.41 to 0.15)	.36
Economic status[a]	.01 (–0.02 to 0.04)	.61	.05 (0.01 to 0.09)	.01	–.002 (–0.02 to 0.02)	.67	.01 (–0.02 to 0.04)	.34
Cons[b]	2.57 (1.27 to 3.88)	<.001	2.22 (0.73 to 3.72)	.01	1.33 (0.55 to 2.12)	.06	9.15 (7.85 to 10.44)	<.001

[a] Logarithmic value of the annual per capita income.

[b] Cons: constant term.

On sensitivity analyses, first, rural road mileage had a significant positive effect on the number of inpatient visits (*P*<.001) and outpatient expenditure (*P*<.001) for older adults in rural areas, which was generally consistent with results shown in Table 2. Second, Figures 1 and 2 show that outpatient visits, outpatient expenditure, and inpatient expenditure had a similar temporal trend between the intervention and control groups, which verified the validity of the parallel trend assumption. After matching, there was still a positive effect of TPA policy implementation on outpatient and inpatient expenditures (*P*<.001). Third, the estimates of medical expenditure and the outcomes among older adults without pension supported the positive effect of TPA policy implementation on inpatient visits and medical expenditure (Multimedia Appendix 4).

Figure 1. Parallel trend of older adults' outpatient (A) and inpatient (B) visit times from 2011 to 2018.

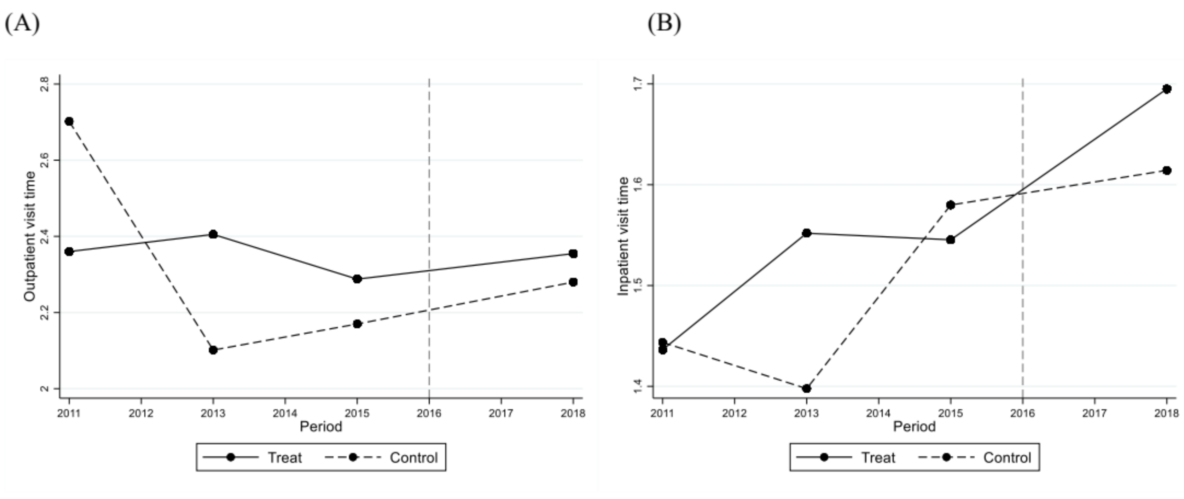
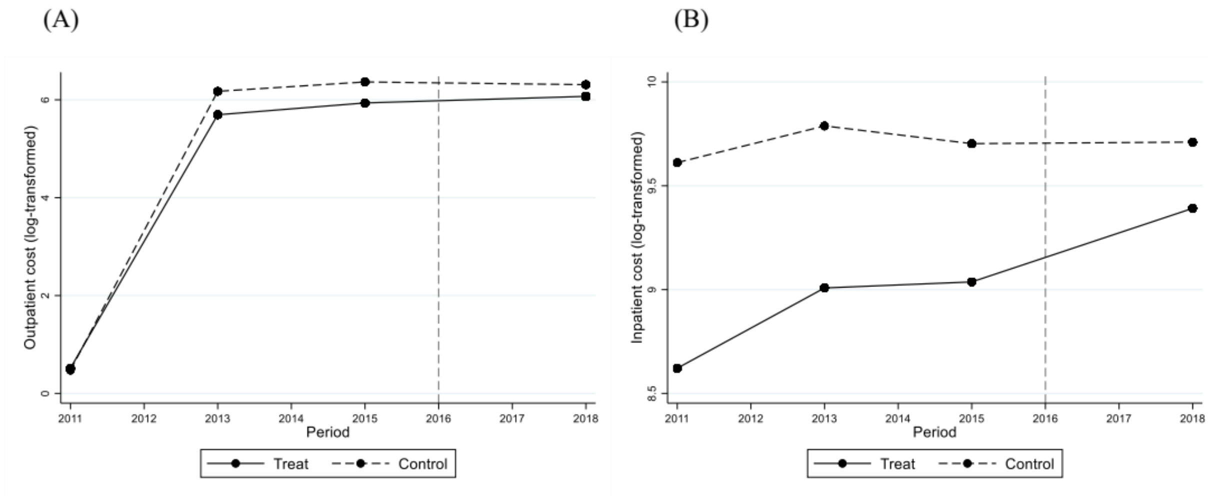


Figure 2. Parallel trend of the older adults' outpatient (A) and inpatient (B) expenditures from 2011 to 2018.



The effects of TPA policy on health care–seeking behavior were also examined and are summarized in Table 3. Models 5-9 show the effect of TPA policy implementation on travelling distance and hospital level of outpatient or inpatient visits. TPA policy implementation significantly increased patients' traveling distance to hospitals by 23 km compared to that of patients without a TPA policy (*P*=.04). The choice of hospital level was increased by 0.08 points after TPA policy implementation (*P*=.01). Model 9 shows that TPA policy implementation increased individual incomes of older adults in rural areas by 457% (*P*<.001).

Table 3. Regression models analyzing the association between Transport-driven Poverty Alleviation (TPA) policy implementation on travelling distance, hospital level, and individual income among the older adults in rural areas from 2011 to 2018. All models had control variables.

Variables	Model 5 (1620 observations)		Model 6 (770 observations)		Model 7 (1788 observations)		Model 8 (1313 observations)		Model 9 (12,600 observations)	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
TPA	9.66 (–4.38 to 23.70)	.18	–.01 (–0.08 to 0.06)	.81	23.16 (0.99 to 45.31)	.04	.08 (0.02 to 0.13)	.01	4.57 (4.46 to 4.68)	<.001
Cons[a]	24.678 (–50.36 to 99.71)	.52	1.23 (0.86 to 1.60)	<.001	165.16 (43.14 to 287.19)	.008	1.45 (1.13 to 1.77)	<.001	–1.20 (–1.77 to –0.62)	<.001

[a] Cons: constant term.

The potential effect modifications were examined in subgroups by poor household, self-reported health status, and age and are summarized in Multimedia Appendix 2. In the overall sample, after the implementation of TPA policy, the outpatient expenditure increased by 469% and the inpatient expenditure increased by 85.9% among poor households. In addition, TPA policy implementation increased medical expenditures of older adults with better self-rated health status significantly compared to the medical expenditures of those with poor health. When considering differences in age, TPA policy implementation increased outpatient and inpatient expenditures among 60-80–year-old adults significantly (*P*<.001).

Discussion

Principal Findings

In this quasi-experimental study of a nationwide sample of older adults in rural areas, we observed a positive effect of the implementation of a national transport policy on health care–seeking behavior. Empirical results show that TPA policy implementation increased inpatient visits and outpatient and inpatient expenditures significantly. After conducting sensitivity analyses, the promoting effect of TPA policy implementation on inpatient visit times and medical expenditure among older adults in rural areas was still supported.

With rapid growth of the economy, differences in the usage of medical services and health inequality between urban and rural residents in China have been increasingly consequential, and rural residents have been at a disadvantage [24]. In China, due to the disparity between urban and rural areas, medical resources are concentrated in the large urban hospitals, while there are relatively poor resources in rural health care institutions, such as rural and community-level hospitals. Rural residents living in remote areas have faced many limitations when using medical services due to long distances to hospitals and inconvenient transport in rural China. Transport barriers in rural areas may be an important factor affecting medical service usage among older adults in rural areas [25]. When the distance to hospitals increased, residents' use of health care significantly decreased, and the probability of outpatient care also decreased [26]. Rural patients were more likely to travel longer distances for access to high-quality medical resources when there was a lack of such resources in their areas of residence; therein, the availability of transportation played an important role.

Previous studies have reported the positive role of transport in health care–seeking behavior [11-13,27]. For example, Badji et al [27] found that when public transport availability was high, people with disabilities visited their general practitioners on average 0.5 more times per year without considering the associated medical expenditure. In particular, it was worth noting that TPA policy implementation also increased high-level health care usage, such as inpatient visits in high-level hospitals, which was consistent with Liu et al's [12] findings that 6.9% of outpatients in the Sichuan province of China were transferred to high-level medical institutions via high-speed rail. Nevertheless, our conclusions were of more universal significance because of a nation-level data analysis and paying attention to medical expenditure in rural areas [11-13,27]. TPA policy implementation not only improved the accessibility of older adults in rural areas to acquire medical services, including travelling longer distances and going to higher-level hospitals for inpatient visits, but also incurred greater medical expenditure along with more inpatient visits, which was in line with the fact that rural patients' preferences for high-level hospitals and them increasingly choosing to access higher-level hospitals resulted in greater medical expenditures [28,29].

Possible explanations for the effects of TPA policy implementation were as follows. First, traveling costs decreased after implementing TPA policy. The traveling time and expenditure of seeking medical services were considered medical service expenditure [15]. While better transport conditions and better road quality reduced the time expenditure for seeking health care, more choices and increased access to public transport also reduced the travel-related expenditure. Especially for older adults in rural areas, there was sensitive price elasticity due to their low income. Hence, lower transport expenditure meant lower prices of medical services, which stimulated more inpatient visits. Second, TPA policy implementation resulted in more high-level hospital–seeking behaviors. TPA policy implementation made it convenient for older adults in rural areas to go to higher-level hospitals, and patients' traveling distance increased after policy implementation. Convenient transport changed the "distance decay effect" to a "distance enhancement effect" by removing the geographic barrier and improving patients' accessibility to medical services, showing that rural patients were more inclined to go to areas with high-quality medical resources at the expense of a longer traveling distance. Additionally, patients were more likely to receive inpatient care in higher-level hospitals after TPA policy implementation. Inequalities in medical resource allocation in China have remained serious [18]. There were 1308 tertiary-level hospitals in China, mainly distributed in central cities. TPA policy implementation accelerated access to high-level medical services and thus increased the medical expenditure. For example, after the opening of high-speed railway, patients were more likely to go to areas with a high density of health care resources [12]. Third, transport increased individual income valuable for seeking inpatient services, which was consistent with previous findings that higher income also was associated with greater access to medical services, and social and economic inequality resulted in health inequality [30,31].

Effect modification analysis revealed that a poor household status, self-rated health status, and age were potential modifiers, and there were modified effects on inpatient expenditure by poor household status, self-rated health status, and age. According to Andersen's health behavior model, age and self-rated health status were individual predisposing characteristics, and income was an enabling resource; these were closely associated with health care usage. First, inpatient expenditure increased further after TPA policy implementation in poor households. Vulnerable groups are more likely to face barriers when they need health services, including physical overdraft, deficient social security, and low income. In addition, a long traveling time was an additional vulnerability faced by poorer individuals in remote areas when seeking health care [32-35]. However, TPA policy implementation increased their seeking behavior through raising their income. Second, compared to those with a poor self-rated health status, inpatient expenditure was higher among healthy older adults. In contrast with previous findings that self-rated health was an important factor in outpatients' health care usage [36], we further found that self-rated health status was associated with inpatient care and expenditure. Third, compared with older adults aged >80 years, inpatient expenditure increased further among older adults aged 60-80 years. It could be possible that older adults aged >80 years have a high demand for health care usage due to poor health, regardless of whether the policy was implemented.

Our findings provide new evidence on the association of transport policy implementation with health care–seeking behavior among older adults in rural areas and medical expenditure based on evidence from China. Inspired by the HiAP approach, our findings provide some new implications about improving accessibility and affordability of health services for vulnerable populations in high-income or transitional countries. Especially for rural areas worldwide, in addition to increasing medical resources, improving transport conditions and eliminating barriers to medical service usage are also meaningful measures to promote medical service usage.

Limitations

This study had several limitations. First, the use of self-reported measures of health care–seeking behavior and medical expenditures might have been underestimated, particularly among older people and those from lower socioeconomic and educational backgrounds, who might be more likely to underreport these factors. Second, the CHARLS questionnaire does not disclose the county-level information of respondents, and its sampling and data collection were beyond our control. As a result, we could not conduct multilevel model analysis and assign the intervention and control groups on the basis of county-level information. But, during the implementation process, TPA policy was actually planned and implemented by provincial governments in a unified manner; hence, using provincial-level information as an identifier for policy intervention could be reliable in this case. Finally, health care–seeking behavior includes a visit rate of twice a week among patients, a nonvisit rate of twice a week, hospitalization rate, and per capita hospital stay. However, these data are not available. Therefore, outpatient and inpatient health care–seeking behavior measures at the individual level were selected to explore their association with TPA policy implementation as microscopic evidence for transport policy and health care usage. Future studies could further explore the complex association between transport policy implementation and other health care–seeking behavior.

Conclusions

Our findings verify the beneficial effects of TPA policy implementation on inpatient visits and found an increase in medical expenditure among older adults in rural areas. Our findings further provide empirical evidence regarding the relationships of TPA policy implementation with health care–seeking behavior and medical expenditure, which are valuable for addressing difficulties in acquiring health care services and advancing health equality of rural vulnerable groups. A package of policies is further suggested to be integrated into health policies to improve rural patients’ poor access to medical services and enhance their ability to pay for medical expenditure. In addition, concerted efforts are needed to balance the high-quality medical resources and narrow the great gap between urban and rural areas, and finally contribute to healthy ageing and health equity worldwide.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant 72274066). We would like to express our gratitude to the China Health and Retirement Longitudinal Study (CHARLS) research team and field team for their contributions to data collection. We would also like to thank all the interviewees who voluntarily participated in the CHARLS.

Data Availability

The original data set is publicly available at the China Health and Retirement Longitudinal Study (CHARLS) website [23]. The data analyzed in our study can be obtained from the corresponding author upon reasonable request.

Authors' Contributions

XZ and LY were prominent contributors to the development of the manuscript. XZ conceptualized and designed the study. YW contributed to the literature search, figure development, study design, data analyses, data curation, data interpretation, and drafting of the manuscript. XZ, LY, and YW contributed to methodology and reviewed and edited the manuscript. XZ and LY supervised the study, reviewed the manuscript draft, and revised and proofread the manuscript. YW, QW, FZ, TY, YW, and SF proofread the manuscript. All authors reviewed the final version of the manuscript and agreed on its contents and submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Implementation scope of transport-driven poverty alleviation policy.
[DOCX File , 17 KB - publichealth_v9i1e49603_app1.docx]

Multimedia Appendix 2
Subgroup analysis of TPA policy on medical expenditure among poor- households elderly (yes vs. no).
[DOCX File , 15 KB - publichealth_v9i1e49603_app2.docx]

Multimedia Appendix 3
Outcome distribution of pre-and post-intervention within group from 2011~2018.
[DOCX File , 15 KB - publichealth_v9i1e49603_app3.docx]

Multimedia Appendix 4
Sensitive analysis of rural road mileage on healthcare seeking behavior and medical expenditures among the elderly.
[DOCX File , 16 KB - publichealth_v9i1e49603_app4.docx]

References

1. Shrime MG, Dare AJ, Alkire BC, O'Neill K, Meara JG. Catastrophic expenditure to pay for surgery worldwide: a modelling study. The Lancet Global Health 2015 Apr;3:S38-S44. [doi: 10.1016/s2214-109x(15)70085-9]
2. Li J, Zhang K, Tian L. Multiple facets of China's health inequality. The Lancet 2006 Apr;367(9520):1397. [doi: 10.1016/s0140-6736(06)68611-x]

3. Liu Y, Rao K, Wu J, Gakidou E. China's health system performance. The Lancet 2008 Nov;372(9653):1914-1923. [doi: 10.1016/s0140-6736(08)61362-8]
4. Mills A, Ataguba JE, Akazili J, Borghi J, Garshong B, Makawia S, et al. Equity in financing and use of health care in Ghana, South Africa, and Tanzania: implications for paths to universal coverage. The Lancet 2012 Jul;380(9837):126-133. [doi: 10.1016/s0140-6736(12)60357-2]
5. Arcury TA, Gesler WM, Preisser JS, Sherman J, Spencer J, Perin J. The effects of geography and spatial behavior on health care utilization among the residents of a rural region. Health Serv Res 2005 Feb;40(1):135-155 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00346.x] [Medline: 15663706]
6. Wang L, Ning P, Yin P, Cheng P, Schwebel DC, Liu J, et al. Road traffic mortality in China: analysis of national surveillance data from 2006 to 2016. Lancet Public Health 2019 May;4(5):e245-e255 [FREE Full text] [doi: 10.1016/S2468-2667(19)30057-X] [Medline: 31054642]
7. Solomon E, Wing H, Steiner J, Gottlieb L. Impact of transportation interventions on health care outcomes: a systematic review. Med Care 2020 Apr;58(4):384-391. [doi: 10.1097/MLR.0000000000001292] [Medline: 31985588]
8. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1. [doi: 10.2307/2137284]
9. Nemet GF, Bailey AJ. Distance and health care utilization among the rural elderly. Soc Sci Med 2000 May;50(9):1197-1208. [doi: 10.1016/s0277-9536(99)00365-2] [Medline: 10728841]
10. Zielinski A, Borgquist L, Halling A. Distance to hospital and socioeconomic status influence secondary health care use. Scand J Prim Health Care 2013 Jun 10;31(2):83-88 [FREE Full text] [doi: 10.3109/02813432.2012.759712] [Medline: 23301541]
11. Arcury TA, Preisser JS, Gesler WM, Powers JM. Access to transportation and health care utilization in a rural region. J Rural Health 2005 Jan;21(1):31-38. [doi: 10.1111/j.1748-0361.2005.tb00059.x] [Medline: 15667007]
12. Liu G, Tang C, Liu Y, Bu T, Tang D. Will high-speed railway influence the healthcare seeking behaviour of patients? Quasi-experimental evidence from China. Health Policy Plan 2021 Nov 11;36(10):1633-1643. [doi: 10.1093/heapol/czab018] [Medline: 34058009]
13. Horner SD, Ambrogne J, Coleman MA, Hanson C, Hodnicki D, Lopez SA, et al. Traveling for care: factors influencing health care access for rural dwellers. Public Health Nurs 1994 Jun;11(3):145-149. [doi: 10.1111/j.1525-1446.1994.tb00393.x] [Medline: 8898552]
14. 13th Five Year Plan for Transportation Poverty Alleviation. Ministry of Transport. 2016. URL: https://www.gov.cn/xinwen/2016-08/11/content_5098829.htm [accessed 2023-11-17]
15. Chandra A, Finkelstein A, Sacarny A, Syverson C. Health care exceptionalism? Performance and allocation in the US health care sector. Am Econ Rev 2016 Aug;106(8):2110-2144 [FREE Full text] [doi: 10.1257/aer.20151080] [Medline: 27784907]
16. Alam N, Chowdhury ME, Kouanda S, Seppey M, Alam A, Savadogo JR, et al. The role of transportation to access maternal care services for women in rural Bangladesh and Burkina Faso: A mixed methods study. Int J Gynaecol Obstet 2016 Nov 08;135 Suppl 1(S1):S45-S50. [doi: 10.1016/j.ijgo.2016.09.003] [Medline: 27836084]
17. Kelly J, Dwyer J, Willis E, Pekarsky B. Travelling to the city for hospital care: access factors in country Aboriginal patient journeys. Aust J Rural Health 2014 Jun 11;22(3):109-113. [doi: 10.1111/ajr.12094] [Medline: 25039844]
18. Fang P, Dong S, Xiao J, Liu C, Feng X, Wang Y. Regional inequality in health and its determinants: evidence from China. Health Policy 2010 Jan;94(1):14-25. [doi: 10.1016/j.healthpol.2009.08.002] [Medline: 19735959]
19. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb 12;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
20. Yao Y, Lv X, Qiu C, Li J, Wu X, Zhang H, et al. The effect of China's Clean Air Act on cognitive function in older adults: a population-based, quasi-experimental study. The Lancet Healthy Longevity 2022 Feb;3(2):e98-e108. [doi: 10.1016/s2666-7568(22)00004-6]
21. Atella V, Piano Mortari A, Kopinska J, Belotti F, Lapi F, Cricelli C, et al. Trends in age-related disease burden and healthcare utilization. Aging Cell 2019 Feb 29;18(1):e12861 [FREE Full text] [doi: 10.1111/acel.12861] [Medline: 30488641]
22. Zhang S, Chen Q, Zhang B. Understanding healthcare utilization in China through the Andersen Behavioral Model: review of evidence from the China Health And Nutrition Survey. RMHP 2019 Nov;Volume 12:209-224. [doi: 10.2147/rmhp.s218661]
23. Harmonized CHARLS datasets. URL: https://charls.charlsdata.com/pages/data/111/zh-cn.html [accessed 2023-11-17]
24. Tang S, Meng Q, Chen L, Bekedam H, Evans T, Whitehead M. Tackling the challenges to health equity in China. The Lancet 2008 Oct;372(9648):1493-1501. [doi: 10.1016/s0140-6736(08)61364-1]
25. Mattson J. Transportation, distance, and health care utilization for older adults in rural and small urban areas. Transp Res Rec 2011 Jan 01;2265(1):192-199. [doi: 10.3141/2265-22]
26. Li L, Wang J, Yuan J. The Impact of Hospital Distance on the Chinese Inpatient Service Demand in Rural Areas: The Application of Discrete Choice Model. Chinese Health Economics 2014;33(1):11-13 [FREE Full text] [doi: 10.7664/CHE20140102]
27. Badji S, Badland H, Rachele JN, Petrie D. Public transport availability and healthcare use for Australian adults aged 18–60 years, with and without disabilities. J Transp Health 2021 Mar;20:101001. [doi: 10.1016/j.jth.2020.101001]

28. Liu Y, Zhong L, Yuan S, van de Klundert J. Why patients prefer high-level healthcare facilities: a qualitative study using focus groups in rural and urban China. BMJ Glob Health 2018 Sep 19;3(5):e000854 [FREE Full text] [doi: 10.1136/bmjgh-2018-000854] [Medline: 30258653]
29. Yu W, Li M, Ye F, Xue C, Zhang L. Patient preference and choice of healthcare providers in Shanghai, China: a cross-sectional study. BMJ Open 2017 Nov 01;7(10):e016418 [FREE Full text] [doi: 10.1136/bmjopen-2017-016418] [Medline: 29092898]
30. Mackenbach JP, Bakker MJ. Tackling socioeconomic inequalities in health: analysis of European experiences. The Lancet 2003 Oct;362(9393):1409-1414. [doi: 10.1016/s0140-6736(03)14639-9]
31. Elgar FJ, Pförtner T, Moor I, De Clercq B, Stevens GWJM, Currie C. Socioeconomic inequalities in adolescent health 2002–2010: a time-series analysis of 34 countries participating in the Health Behaviour in School-aged Children study. The Lancet 2015 May;385(9982):2088-2095. [doi: 10.1016/s0140-6736(14)61460-4]
32. McDonough P, Sacker A, Wiggins RD. Time on my side? Life course trajectories of poverty and health. Soc Sci Med 2005 Oct;61(8):1795-1808. [doi: 10.1016/j.socscimed.2005.03.021] [Medline: 16029777]
33. Weiss DJ, Nelson A, Vargas-Ruiz CA, Gligorić K, Bavadekar S, Gabrilovich E, et al. Global maps of travel time to healthcare facilities. Nat Med 2020 Dec;26(12):1835-1838. [doi: 10.1038/s41591-020-1059-1] [Medline: 32989313]
34. Choi JW, Park E, Chun S, Han K, Han E, Kim TH. Health care utilization and costs among medical-aid enrollees, the poor not enrolled in medical-aid, and the near poor in South Korea. Int J Equity Health 2015 Nov 14;14(1):128 [FREE Full text] [doi: 10.1186/s12939-015-0257-9] [Medline: 26572490]
35. Liao P, Zhang X, Zhang W. Endogenous health risks, poverty traps, and the roles of health insurance in poverty alleviation. Health Econ Rev 2022 Apr 19;12(1):25 [FREE Full text] [doi: 10.1186/s13561-022-00370-2] [Medline: 35438342]
36. Pu C, Bai YM, Chou YJ. The impact of self-rated health on medical care utilization for older people with depressive symptoms. Int J Geriatr Psychiatry 2013 May;28(5):479-486. [doi: 10.1002/gps.3849] [Medline: 22700167]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
DID: difference-in-differences
HiAP: Health in All Policies
TPA: Transport-driven Poverty Alleviation

Edited by A Mavragani, T Sanchez; submitted 03.06.23; peer-reviewed by J Han, L Lafrado; comments to author 03.08.23; revised version received 23.08.23; accepted 07.11.23; published 28.11.23.

Please cite as:
Wu Y, Wang Q, Zheng F, Yu T, Wang Y, Fan S, Zhang X, Yang L
Effects of the Implementation of Transport-Driven Poverty Alleviation Policy on Health Care–Seeking Behavior and Medical Expenditure Among Older People in Rural Areas: Quasi-Experimental Study
JMIR Public Health Surveill 2023;9:e49603
URL: https://publichealth.jmir.org/2023/1/e49603
doi:10.2196/49603
PMID:38015603

Original Paper

Patient-Centered Economic Burden of Exudative Age-Related Macular Degeneration: Retrospective Cohort Study

Kyungseon Choi[1,2*], MSc, PharmD; Sang Jun Park[3*], MSc, MD, PhD; Sola Han[4], MSc, PharmD, PhD; Yongseok Mun[5], MD; Da Yun Lee[3], MS; Dong-Jin Chang[6], MSc, MD; Seok Kim[7], MPH; Sooyoung Yoo[7], PhD; Se Joon Woo[3], MD, PhD; Kyu Hyung Park[8], MD, PhD; Hae Sun Suh[1,2,9], MA, MPharm, PhD

[1]Department of Regulatory Science, Graduate School, Kyung Hee University, Seoul, Republic of Korea

[2]Institute of Regulatory Innovation Through Science, Kyung Hee University, Seoul, Republic of Korea

[3]Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

[4]Health Outcomes Division, College of Pharmacy, University of Texas at Austin, Austin, TX, United States

[5]Department of Ophthalmology, Kangnam Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Republic of Korea

[6]Department of Ophthalmology and Visual Science, Yeouido St Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

[7]Healthcare ICT Research Center, Office of eHealth Research and Businesses, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

[8]Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Hospital, Seoul, Republic of Korea

[9]College of Pharmacy, Kyung Hee University, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Hae Sun Suh, MA, MPharm, PhD
College of Pharmacy
Kyung Hee University
26, Kyunghee-daero, Dongdaemun-gu
Seoul, 02447
Republic of Korea
Phone: 82 29619492
Email: haesun.suh@khu.ac.kr

Abstract

Background: Exudative age-related macular degeneration (AMD), one of the leading causes of blindness, requires expensive drugs such as anti–vascular endothelial growth factor (VEGF) agents. The long-term regular use of effective but expensive drugs causes an economic burden for patients with exudative AMD. However, there are no studies on the long-term patient-centered economic burden of exudative AMD after reimbursement of anti-VEGFs.

Objective: This study aimed to evaluate the patient-centered economic burden of exudative AMD for 2 years, including nonreimbursement and out-of-pocket costs, compared with nonexudative AMD using the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).

Methods: This retrospective cohort study was conducted using the OMOP CDM, which included 2,006,478 patients who visited Seoul National University Bundang Hospital from June 2003 to July 2019. We defined the exudative AMD group as patients aged >50 years with a diagnosis of exudative AMD and a prescription for anti-VEGFs or verteporfin. The control group was defined as patients aged >50 years without a diagnosis of exudative AMD or a prescription for anti-VEGFs or verteporfin. To adjust for selection bias, controls were matched by propensity scores using regularized logistic regression with a Laplace prior. We measured any medical cost occurring in the hospital as the economic burden of exudative AMD during a 2-year follow-up period using 4 categories: total medical cost, reimbursement cost, nonreimbursement cost, and out-of-pocket cost. To estimate the average cost by adjusting the confounding variable and overcoming the positive skewness of costs, we used an exponential conditional model with a generalized linear model.

Results: We identified 931 patients with exudative AMD and matched 783 (84.1%) with 2918 patients with nonexudative AMD. In the exponential conditional model, the total medical, reimbursement, nonreimbursement, and out-of-pocket incremental costs were estimated at US $3426, US $3130, US $366, and US $561, respectively, in the first year and US $1829, US $1461, US $373, and US $507, respectively, in the second year. All incremental costs in the exudative AMD group were 1.89 to 4.25 and 3.50 to 5.09 times higher in the first and second year, respectively, than those in the control group ($P<.001$ in all cases).

Conclusions: Exudative AMD had a significantly greater economic impact ($P<.001$) for 2 years on reimbursement, nonreimbursement, and out-of-pocket costs than nonexudative AMD after adjusting for baseline demographic and clinical characteristics using the OMOP CDM. Although economic policies could relieve the economic burden of patients with exudative AMD over time, the out-of-pocket cost of exudative AMD was still higher than that of nonexudative AMD for 2 years. Our findings support the need for expanding reimbursement strategies for patients with exudative AMD given the significant economic burden faced by patients with incurable and fatal diseases both in South Korea and worldwide.

(JMIR Public Health Surveill 2023;9:e49852) doi:10.2196/49852

KEYWORDS

blindness; age-related macular degeneration; economic burden; cost of illness; retrospective cohort study; common data model

Introduction

Background

Age-related macular degeneration (AMD) is a progressive retinal disorder that can lead to irreversible loss of central vision. AMD can be classified as dry or exudative (wet or neovascular). Exudative AMD is less common; however, it causes faster and more catastrophic vision loss than dry AMD. Anti–vascular endothelial growth factors (VEGFs) have demonstrated remarkable efficacy in averting blindness, converting once unavoidably blinding conditions into chronic diseases [1]. Anti-VEGFs are currently suggested as the primary treatment for exudative AMD [2], and 5 anti-VEGF agents (ie, ranibizumab, aflibercept, brolucizumab, bevacizumab, and faricimab) are clinically accessible. Photodynamic therapy with verteporfin, which was the first-line treatment in the past, is now used as adjuvant therapy for some subtypes of exudative AMD because of the lack of expectation of improvement in visual acuity [3]. All approved anti-VEGFs and modalities are quite expensive, at least US $500, and should be used repeatedly; however, bevacizumab, the only one used off-label, has demonstrated equivalent efficacy to ranibizumab [4] and is significantly inexpensive and widely used in patients [5].

In early randomized clinical trials, patients needed monthly injections in the first year of treatment under a fixed regimen. However, in real-world clinical practice, because of expensive out-of-pocket costs and strict reimbursement criteria, the fixed regimen was not affordable; thus, the treatment-and-extend regimen was introduced, showing similar efficient treatment outcomes with fewer injections [6]. Despite the introduction of more affordable treatment regimens [7], patients with exudative AMD still need 7 to 8 injections in the first year of treatment, which causes a substantial economic burden.

Although there are several studies [8-15] that have examined the economic burden of exudative AMD, they have some limitations that need to be addressed. First, most previous studies were conducted before the introduction of effective drugs such as anti-VEGFs, which have significantly affected the treatment landscape for exudative AMD. Second, previous studies were conducted through interviews, which may not provide an accurate evaluation of the true economic burden. Third, although a few studies analyzed real-world data, they only included reimbursed interventions and had a limited follow-up period of 1 year. In addition, exudative AMD is characterized by the common use of bevacizumab, an off-label drug, and nonreimbursed measurements, which add to the patients' burden of disease. Given the intractable nature of exudative AMD, which requires long-term treatment and management, there is a critical unmet need to assess the long-term economic burden of patients with exudative AMD using real-world data, including nonreimbursed interventions.

Objectives

Therefore, to fill this gap and generate real-world evidence to understand clinical, economic, and policy implications, we aimed to estimate the economic burden on patients with exudative AMD, including nonreimbursement costs up to the second year, using the real-world data from the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). Using electronic medical records (EMRs) transformed into the OMOP CDM of a major university hospital with a large number of patients with AMD is the best way to define real-world patients with exudative AMD and encompass drugs that are not reimbursed by the national health insurance in South Korea. This approach will facilitate the estimation of the economic burden, including nonreimbursed and out-of-pocket costs, from the patients' perspective with a high quality of data.

Methods

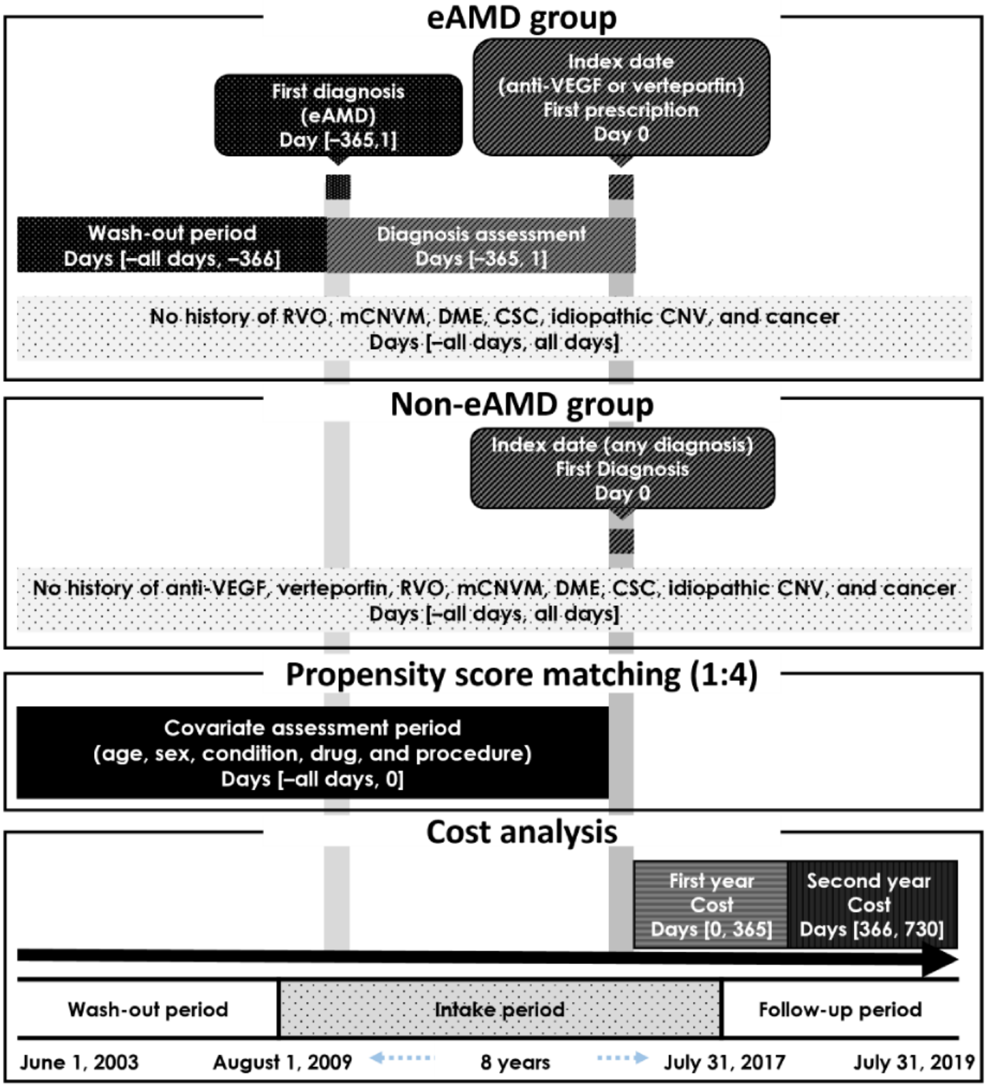
Study Design and Data Source

This retrospective cohort study used the Observational Health Data Sciences and Informatics (OHDSI) OMOP CDM from Seoul National University Bundang Hospital (SNUBH), which is one of the largest and leading hospitals in South Korea. The OMOP CDM is a structured data model combined with standardized vocabulary. All medical and cost data were extracted, transformed, and loaded into the OMOP CDM version 5.3 following the OHDSI extract, transform, and load process and contained not only the diagnosis, drug exposure, and procedure occurrence for reimbursement but also all records for medical practice with costs by order and receipt regardless of reimbursement [16]. All data were verified using Automated Characterization of Health Information at Large-Scale Longitudinal Evidence Systems and double-checked by data analysts and clinicians [17,18].

Study Population

We analyzed the OMOP CDM database, which included 2,006,478 patients who visited SNUBH from June 2003 to July 2019. To ensure a 1-year wash-out period and a 2-year follow-up period, the intake period started in August 2009 and ended in July 2017. We established the eligibility criteria for the exudative and nonexudative AMD groups based on the opinions of ophthalmologists. For the exudative AMD group, the index date was the first prescription date for anti-VEGFs or verteporfin. The exudative AMD group was defined as follows: (1) a patient who was prescribed anti-VEGFs or verteporfin by an ophthalmologist for the first time between August 1, 2009, the day ranibizumab was approved for reimbursement by the National Health Insurance Service in South Korea, and July 31, 2017; and (2) a patient aged >50 years who had been diagnosed with exudative AMD within 1 year before the index date plus 1 day. Among these patients, we excluded those with cancer or other ophthalmic diseases using anti-VEGFs (choroidal neovascularization membrane, central serous chorioretinopathy, diabetic macular edema, and retinal vein occlusion). The nonexudative AMD group was defined as patients aged >50 years who had been diagnosed with any disease except exudative AMD between August 1, 2009, and July 31, 2017. We defined the index date for the nonexudative AMD group as the date of the first diagnosis of any disease to compare the outcomes of the 2 groups over time and investigate potential differences in disease progression and treatment efficacy. We also excluded patients with any of the following conditions: (1) those who were prescribed anti-VEGFs, (2) those diagnosed with cancer, and (3) those diagnosed with other ophthalmic diseases using anti-VEGFs (Figure 1).

Figure 1. Analysis scheme. Data were included from June 2003 to July 2019. The intake period was defined as August 2009 to July 2017. The follow-up period was 2 years from the index date. CNV: choroidal neovascularization; CSC: central serous chorioretinopathy; DME: diabetic macular edema; eAMD: exudative age-related macular degeneration; mCNVM: myopic choroidal neovascular membrane; RVO: retinal vein occlusion; VEGF: vascular endothelial growth factor.



Outcomes

The economic burden of exudative AMD was measured by estimating the medical costs caused by any reason in the hospital in the first and second years after the index date (Figure 1). Medical costs were grouped into 4 categories (Textbox 1): total medical costs, reimbursement costs, nonreimbursement costs, and out-of-pocket costs. We estimated the incremental costs of the exudative AMD group in comparison with those of the nonexudative AMD group. The total medical cost was calculated as the sum of the reimbursement and nonreimbursement costs. The reimbursement cost includes all expenses only limited to reimbursed items, which comprise the costs of health care benefits from the National Health Insurance Service and copayment, in which patients bear part of the expenses of reimbursement. Similarly, the nonreimbursement cost includes all expenses for any nonreimbursed items. The out-of-pocket cost was estimated as the sum of the medical costs paid by the patient irrespective of reimbursement. All costs were calculated in Korean won and converted to US dollars (US $1=1200 Korean won).

Textbox 1. Definitions of cost categories for the patient-centered economic burden outcome.

- Total medical cost: all expenses associated with each patient at the hospital, encompassing both reimbursed and nonreimbursed expenditures for 1 year (sum of the reimbursement and nonreimbursement costs)
- Reimbursement cost: all expenses incurred by patients restricted to items eligible for reimbursement for 1 year
- Nonreimbursement cost: all expenses incurred by patients restricted to items eligible for nonreimbursement for 1 year
- Out-of-pocket cost: all expenses borne directly by the patient irrespective of any reimbursement for 1 year

Statistical Analysis

Data analyses were conducted using ATLAS (version 2.10.1; Observational Health Data Science and Informatics), the R software (version 4.0.3; R Foundation for Statistical Computing), the Health Resources Econometric Analysis Tool (HERMES), which we developed for cost analysis with the OMOP CDM vocabulary in R environment, and Health Analytics Data-to-Evidence Suite, formerly known as the OHDSI Methods Library. The HERMES code and algorithm are available on GitHub [19]. We evaluated the baseline characteristics of the exudative and nonexudative AMD groups, including demographics and clinical characteristics, drugs, and conditions. To reduce selection bias in the quasi-experimental design, we performed propensity score matching using regularized logistic regression with a Laplace prior (least absolute shrinkage and selection operator; LASSO) [20,21]. We used demographic data, conditions, drugs, procedures, and Charlson Comorbidity Index scores as covariates in the propensity score matching (Figure 1). We considered covariates as unmatched variables if the standardized difference was >0.10 after propensity score matching. We included single concepts and concept groups within the parent categories for conditions and drugs in the propensity score matching. We excluded highly correlated covariates related to exudative AMD, including visual- and retinal-related disorders, anti-VEGFs, verteporfin, and ophthalmological drugs. We used an exponential conditional model (ECM) with a generalized linear model to estimate the average cost by adjusting the confounding variable and overcoming the positive skewness of costs [22]. To determine a suitable distribution and link function, we performed modified Park tests and the Box-Cox test. The goodness of fit of the ECM was evaluated using the Akaike information criterion. Patients with 0 costs who did not incur costs during 1 year after the index date were excluded. We considered demographic data; 1-year costs before the index date, which were defined as the preindex costs; and unmatched variables from the propensity score matching as covariates to adjust for remaining confounders. To interpret the results, we exponentiated the coefficient. The exponentiated coefficient represents the expected change in the response variable associated with a 1-unit increase in the predictor variable when the other predictor variables are at their mean values.

Ethical Considerations

This study was approved by the institutional review board of SNUBH (X-2012-657-902). The data were deidentified and anonymized; thus, informed consent was not required.

Results

Patient Characteristics

Of the 2,006,478 patients who visited SNUBH, 931 (0.05%) were newly diagnosed with exudative AMD, and 373,676 (18.62%) had nonexudative AMD. After 1:4 propensity score matching, 84.1% (783/931) of patients with exudative AMD and 0.78% (2918/373,676) of patients with nonexudative AMD were matched. After excluding patients with 0 costs in the first year, of the 2918 patients with nonexudative AMD, there were 2702 (92.6%) left (Figure 2).

In the propensity score model, 1079 covariates were matched. The standardized mean difference was <0.10 for all variables that significantly affected medical costs, such as cardiovascular disease and opioids. In total, 4 covariates (“other antiepileptics”; “bisphosphonates, combinations”; “corticosteroids, potent, other combinations”; and “drug affecting bone structure and mineralization”) had a standardized mean difference of >0.10; thus, we included these variables in the ECM to adjust for remaining confounders (Table 1).

Figure 2. Selection flow of patients with exudative age-related macular degeneration (AMD) and nonexudative AMD from June 2003 to July 2019. VEGF: vascular endothelial growth factor.

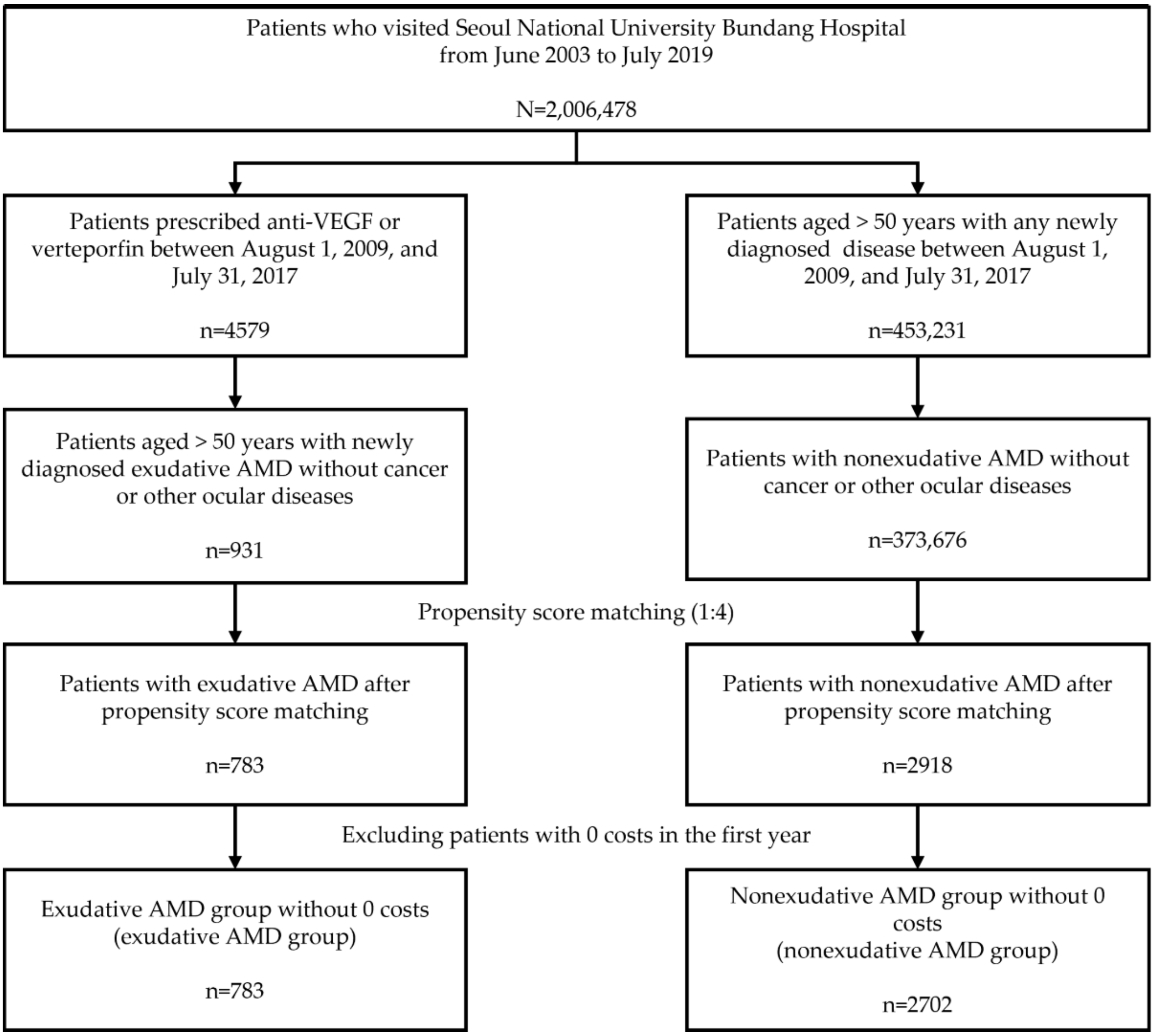


Table 1. Baseline demographic and clinical characteristics for the exudative age-related macular degeneration (AMD) group and nonexudative AMD group in the index period from August 2009 to July 2017 before and after propensity score (PS) matching.

Characteristic[a]	Before PS matching			After PS matching		
	Exudative AMD group (n=931)	Nonexudative AMD group (n=373,676)	SMD[b]	Exudative AMD group (n=783)	Nonexudative AMD group (n=2918)	SMD
Age (years), mean (SD)	72.47 (8.44)	62.48 (9.87)	1.01	71.52 (8.28)	72.16 (10.53)	−0.06
Sex, n (%)						
Female	425 (45.65)	215,184 (57.59)	−0.24	348 (44.44)	1205 (40.55)	0.08
Medical history—general, n (%)[c]						
Acute respiratory disease	3 (0.32)	1144 (0.31)	0.00	1 (0.13)	7 (0.22)	−0.02
Chronic liver disease	6 (0.64)	3499 (0.94)	−0.03	3 (0.38)	6 (0.42)	0.00
Chronic obstructive lung disease	8 (0.86)	306 (0.22)	0.09	4 (0.51)	7 (0.22)	0.05
Dementia	16 (1.72)	4572 (1.22)	0.04	8 (1.02)	21 (0.67)	0.04
Depressive disorder	8 (0.85)	3542 (0.95)	−0.01	4 (0.51)	9 (0.29)	0.04
Diabetes mellitus	63 (6.77)	11,633 (3.11)	0.17	31 (4)	105 (4)	0.00
Gastroesophageal reflux disease	11 (1.18)	5942 (1.59)	−0.04	6 (0.77)	24 (0.82)	−0.01
Gastrointestinal hemorrhage	3 (0.32)	1751 (0.47)	−0.02	1 (0.13)	3 (0.11)	0.01
Hyperlipidemia	32 (3.44)	8838 (2.37)	0.06	19 (2.43)	63 (2.54)	−0.01
Hypertensive disorder	63 (6.77)	18,839 (5.04)	0.07	29 (3.7)	90 (3.61)	0.00
Lesion of liver	4 (0.43)	3123 (0.84)	−0.05	1 (0.13)	6 (0.33)	−0.04
Obesity	2 (0.21)	1155 (0.31)	−0.02	2 (0.26)	6 (0.19)	0.01
Osteoarthritis	16 (1.72)	13,706 (3.67)	−0.12	6 (0.77)	30 (1.4)	−0.06
Pneumonia	16 (1.72)	2883 (0.77)	0.09	10 (1.28)	23 (0.77)	0.05
Renal impairment	11 (1.18)	1904 (0.51)	0.07	6 (0.77)	13 (0.56)	0.03
Rheumatoid arthritis	3 (0.32)	1959 (0.52)	−0.03	2 (0.26)	11 (0.36)	−0.02
Viral hepatitis C	3 (0.32)	596 (0.16)	0.03	1 (0.13)	1 (0.03)	0.03
Medical history—cardiovascular disease, n (%)[c]						
Atrial fibrillation	13 (1.4)	2702 (0.72)	0.07	6 (0.77)	26 (0.84)	−0.01
Cerebrovascular disease	41 (4.4)	21,165 (5.66)	−0.06	26 (3.32)	67 (2.72)	0.04
Coronary arteriosclerosis	11 (1.18)	3866 (1.03)	0.01	3 (0.38)	14 (0.55)	−0.03
Heart disease	52 (5.59)	16,624 (4.45)	0.05	22 (2.81)	82 (3.22)	−0.02
Heart failure	5 (0.54)	1501 (0.4)	0.02	2 (0.26)	6 (0.29)	−0.01
Ischemic heart disease	17 (1.83)	6067 (1.62)	0.02	10 (1.28)	36 (1.54)	−0.02
Peripheral vascular disease	1 (0.11)	193 (0.05)	0.02	1 (0.13)	1 (0.03)	0.03
Medication use, n (%)[c]						
Agents acting on the renin-angiotensin system	83 (8.92)	18,055 (4.83)	0.16	39 (4.98)	138 (5.23)	−0.01
Antibacterials for systemic use	11 (1.18)	7935 (2.12)	−0.07	10 (1.28)	29 (1.09)	0.02
Antidepressants	44 (4.73)	15,864 (4.25)	0.02	24 (3.07)	79 (2.89)	0.01
Antiepileptics	40 (4.3)	11,917 (3.19)	0.06	19 (2.43)	72 (2.61)	−0.01
Anti-inflammatory and antirheumatic products	134 (14.39)	47,361 (12.67)	0.05	61 (7.79)	243 (10.39)	−0.09
Antithrombotic agents	124 (13.32)	31,233 (8.36)	0.16	62 (7.92)	238 (9.16)	−0.04
Beta-blocking agents	51 (5.48)	14,149 (3.79)	0.08	25 (3.19)	82 (3.16)	0.00
Calcium channel blockers	92 (9.88)	20,166 (5.4)	0.17	46 (5.87)	170 (6.69)	−0.03

Characteristic[a]	Before PS matching			After PS matching		
	Exudative AMD group (n=931)	Nonexudative AMD group (n=373,676)	SMD[b]	Exudative AMD group (n=783)	Nonexudative AMD group (n=2918)	SMD
Diuretics	67 (7.2)	13,486 (3.61)	0.16	31 (3.96)	111 (4.04)	0.00
Drugs for acid-related disorders	113 (12.14)	60,875 (16.29)	−0.12	60 (7.66)	207 (8.15)	−0.02
Drugs for obstructive airway diseases	29 (3.11)	7911 (2.12)	0.06	17 (2.17)	50 (2)	0.01
Drugs used in diabetes	54 (5.8)	11,107 (2.97)	0.14	26 (3.32)	110 (4.38)	−0.06
Immunosuppressants	3 (0.32)	1161 (0.31)	0.00	1 (0.13)	6 (0.29)	−0.04
Lipid-modifying agents	93 (9.99)	23,533 (6.3)	0.14	49 (6.26)	164 (6.15)	0.00
Opioids	67 (7.2)	36,821 (9.85)	−0.10	36 (4.6)	138 (5.23)	−0.03
Psycholeptics	102 (10.96)	45,363 (12.14)	−0.04	57 (7.28)	192 (7.95)	−0.03
Psychostimulants, agents used for ADHD[d], and nootropics	8 (0.86)	2093 (0.56)	0.04	4 (0.51)	14 (0.54)	0.00
Other antiepileptics[e,f]	51 (5.48)	9697 (2.6)	0.15	19 (2.43)	116 (4.34)	−0.11
Drugs affecting bone structure and mineralization[f]	87 (9.34)	15,650 (4.19)	0.21	46 (5.87)	181 (8.63)	−0.11
Bisphosphonates (combinations)[f]	87 (9.34)	15,640 (4.19)	0.21	46 (5.87)	181 (8.63)	−0.11
Corticosteroids (potent; other combinations)[f]	64 (6.87)	11,736 (3.14)	0.178	33 (4.21)	127 (6.83)	−0.11

[a] These covariates are a subset of the 1079 covariates used for PS matching.

[b] SMD: standardized mean difference.

[c] Medical history and medication use were identified through diagnosis and prescription within 1 year before the index date.

[d] ADHD: attention-deficit/hyperactivity disorder.

[e] Medication use was identified through prescription any time before the index date.

[f] The covariates were considered unmatched covariates based on standardized differences, which were adjusted in the following exponential conditional model.

The Incremental Health Care Cost of Exudative AMD After Propensity Score Matching

The total medical costs observed in the first year were US $4565 in the exudative AMD group, US $3253 higher than those in the nonexudative AMD group. As a result of the ECM adjusting for remaining confounders, the total medical costs in the exudative AMD group were estimated to be US $4738, which was US $3426 more than those of the nonexudative AMD group. The reimbursement costs in the exudative AMD group were US $4038, which was approximately US $3130 more than the reimbursement costs incurred in the nonexudative AMD group. Regarding the nonreimbursement costs, those incurred in the exudative AMD group were US $776, which was US $366 more than the nonreimbursement costs in the nonexudative AMD group. The out-of-pocket costs were US $1241 in the exudative AMD group, which was US $680 higher than those in the other group. In the second year, the exudative AMD group also had higher medical costs than the nonexudative AMD group. In the ECM estimation, the total medical costs were US $1829 higher in the exudative AMD group. The exudative AMD group had higher reimbursement, nonreimbursement, and out-of-pocket costs of US $1461, US $373, and US $507, respectively (Table 2).

Table 2. The observed and estimated incremental health care costs of the exudative age-related macular degeneration (AMD) group compared with the nonexudative AMD group for 2 years after the index date.

	Observed costs after propensity score matching (US $)			Estimated costs from exponential conditional model[a] (US $)		
	Exudative AMD group (n=783[b]), mean (SD)	Nonexudative AMD group (n=2702[b]), mean (SD)	Incremental costs	Exudative AMD group (n=783[b]), mean (SE)	Nonexudative AMD group (n=2702[b]), mean (SE)	Incremental costs
First year						
Total costs	4564.86 (2702.37)	1311.86 (3201.04)	3253	4738.49 (127.42)	1312.22 (62.14)	3426.27
Reimbursement costs	3784.40 (2313.04)	903.28 (2591.63)	2881.12	4038.44 (195.37)	908.42 (54.08)	3130.02
Nonreimbursement costs	780.46 (731.74)	408.58 (853.27)	371.88	776.46 (23.31)	410.52 (15.73)	365.94
Out-of-pocket costs	1244.21 (1470.13)	677.80 (1233.55)	566.41	1240.54 (46.56)	680 (22.77)	560.53
Second year						
Total costs	1814.35 (2804.95)	418.57 (2548.41)	1395.79	2489.11 (1845.74)	659.91 (538.96)	1829.20
Reimbursement costs	1385.39 (2336.94)	311.03 (2240.33)	1074.36	2036.23 (747.24)	575.50 (520.54)	1460.73
Nonreimbursement costs	428.96 (107.54)	107.54 (525.1)	321.42	491.74 (92.01)	119.08 (21.65)	372.66
Out-of-pocket costs	624.45 (1070.29)	194.36 (760.08)	430.09	736.81 (109.57)	229.91 (46.04)	506.90

[a] The exponential conditional model used a log link function with gamma distribution adjusting for sex, age, preindex cost, and unmatched covariates ("other antiepileptics"; "drugs affecting bone structure and mineralization"; and "corticosteroids, potent, other combination"). The SEs of the estimated costs were calculated using bootstrapping.

[b] After propensity score matching, patients with 0 costs in the first year were excluded.

The ECM to Estimate the Patient-Centered Economic Burden of Exudative AMD After Propensity Score Matching

In the ECM, the total medical costs, reimbursement costs, nonreimbursement costs, and out-of-pocket costs in the first year were significantly 3.52, 4.25, 2.05, and 1.89 times higher, respectively, in the exudative AMD group. There were more significant differences in all costs in the second year than in the first year. The total medical costs, reimbursement costs, nonreimbursement costs, and out-of-pocket costs were significantly 4.93, 5.09, 4.68, and 3.50 times higher, respectively, in the exudative AMD group. Age was found to be a significant predictor of health care costs in the first year (*P*<.001) but not in the second year (*P*=.14). Sex was not a significant factor for costs in the ECM. The preindex cost was always a positive estimate and showed significance except for the nonreimbursement costs and out-of-pocket costs in the first year. Most of the unmatched covariates did not show significance; however, there were cases of positive values with significance for antiepileptics and bone structure–affecting drugs and negative values with significance for corticosteroids (Table 3).

Table 3. The multivariable analysis of medical costs in the first and second years after the index date using the exponential conditional models with propensity score matching in patients with exudative age-related macular degeneration (AMD).

	Intercept[a]		Exudative AMD group[b]		Age[c]		Sex[d]		Preindex cost[e]		Drug group					
											Other antiepileptics[f]		Drugs affecting bone structure and mineralization[g]		Corticosteroids (potent; other combinations)[g]	
	β	*P* value	β	*P* value	β	*P* value	β	*P* value	β	*P* value	β	*P* value	β	*P* value	β	*P* value
First year																
Total costs	6.0145	<.001	1.2646	<.001	.0147	<.001	.0984	.16	.0001	.002	.1163	.54	.2520	.16	−.0563	.59
Reimbursement costs	5.4161	<.001	1.4464	<.001	.0172	<.001	.1395	.09	.0001	<.001	.0884	.68	.2408	.24	.0175	.88
Nonreimbursement costs	5.3879	<.001	.7172	<.001	.0084	.007	.0165	.79	<.0001	.63	.1528	.36	.4959	.002	−.3160	<.001
Out-of-pocket costs	5.7831	<.001	.6345	<.001	.0095	.004	.0621	.28	<.0001	.09	.2131	.16	.3154	.03	−.1780	.04
Second year																
Total costs	4.8186	<.001	1.5948	<.001	.0156	.14	−.1408	.45	.0002	.004	1.0024	.04	−.1456	.76	−.2252	.41
Reimbursement costs	4.3207	<.001	1.6278	<.001	.0182	.10	−.2146	.34	.0002	.01	1.0219	.09	−.2270	.69	−.1129	.73
Nonreimbursement costs	4.5257	<.001	1.5430	<.001	.0008	.92	.0424	.78	.0001	.01	.9547	.02	.17	.66	−.5841	.01
Out-of-pocket Costs	4.9482	<.001	1.2527	<.001	.0032	.66	−.0294	.82	.0002	<.001	.8716	.10	.1453	.65	−.3055	.10

[a] The exponential conditional model used a log link function with gamma distribution adjusting for sex, age, preindex cost, and unmatched covariates.

[b] The reference group was the nonexudative AMD group.

[c] Age was reported in years and is presented as a continuous variable in this table.

[d] Sex was represented as a binary variable, with a value of 1 indicating male and a value of 0 indicating female.

[e] Preindex cost was calculated for 1 year before the index date.

[f] Any time before index date.

[g] For −365 days to index date.

Discussion

Principal Findings

In this study, we found that patients with exudative AMD had a significantly higher economic burden than patients with nonexudative AMD over a 2-year follow-up period, particularly in terms of nonreimbursement and out-of-pocket costs. To address potential bias and confounding from real-world data, we used propensity score matching with LASSO and ECM. The findings of this study are notable as practically nothing is known about the long-term economic burden of exudative AMD from the perspective of patients (eg, nonreimbursement and out-of-pocket costs) using real-world data. The economic burden of patients with exudative AMD was significantly higher than that of patients with nonexudative AMD for 2 years. Specifically, the total medical costs adjusted for confounders after propensity score matching of the exudative AMD group were US $3426 and US $1829 in the first and second years, respectively. The incremental costs of reimbursement for exudative AMD were US $3130 in the first year and US $1461 in the second year. The nonreimbursement cost was US $366 higher in the exudative AMD group in the first year, and the incremental cost increased to US $373 in the second year. In addition, the out-of-pocket cost of exudative AMD was US $561 higher in the first year and continued to US $507 in the second year. To the best of our knowledge, this is the first study to derive real-world evidence of the patient-centered economic burden of exudative AMD. Our findings highlight the need for improved reimbursement strategies that reduce out-of-pocket costs for patients with exudative AMD, thereby reducing the financial burden associated with this disease.

Comparison With Previous Studies

Historically, the economic burden of exudative AMD has been reported using various methods and outcomes but has not adequately captured the patient perspective using real-world data (Table 4). The annual disease burden of AMD was reported to be €51.3 to €101.1 million (US $54.8-$108.01 million) in France, Germany, Italy, and the United Kingdom in the study by Bonastre et al [9] and €5300 to €12,445 (US $5662.47-$13,296.10) in Canada, Germany, France, Spain, and the United Kingdom in the study by Cruess et al [11]. The annual cost of exudative AMD was 7 times higher than that of the control group in the study by Lotery et al [14]. In addition, high societal costs associated with exudative AMD were reported in the United States in the study by Brown et al [10]. These 4 studies were conducted using a survey, in which the burden of exudative AMD was very high even before the anti-VEGF era. After anti-VEGF therapy appeared, 3 studies reported the burden of the disease. Spooner et al [15] reported that the direct medical cost of exudative AMD was Aus $199.20 (US $126.66) and the indirect cost was Aus $64.8 (US $41.20) using questionnaires. Previous studies using self-reported measures or surveys can provide valuable information, but they have several limitations [23]. Research based on human interaction is subject to recall bias. Respondents may not remember or respond accurately, and it can be difficult to derive representative values. These are limitations that can lead to inaccurate results. Depending on the nature of the survey, there may be a limited amount of time to respond, and some questions may be difficult to answer, resulting in missing or inaccurate information. Furthermore, as it is impossible to survey the entire population, a survey may not be representative and may be subject to selection and social desirability biases, where respondents may underreport certain behaviors or overreport others because of perceived social norms. Finally, survey results may not be generalizable to other populations or periods.

Table 4. Previous studies associated with the economic burden of exudative age-related macular degeneration.

Study, year	Economic analysis		Outcomes
	Price	Quantity	
Bonastre et al [9], 2002	Unit cost	Consultation with an ophthalmologist	Yearly budget impact on medical cost
Lotery et al [14], 2007	Unit cost	Survey	Annual health resource use cost
Cruess et al [11], 2008	Unit cost	Self-reported and medical chart	Annual societal cost
Brown et al [10], 2016[a]	Self-reported weighted average annual cost from cohort	Self-reported weighted average annual cost from cohort	Annual direct and indirect costs
Spooner et al [15], 2018	Interview	Interview	Annual direct and indirect costs
Kim et al [13], 2019[a]	Exponential conditional model with generalized linear model using South Korean claims data	Exponential conditional model with generalized linear model using South Korean claims data	Incremental reimbursement cost
Jeon et al [12], 2020[a]	Generalized linear model using South Korean claims data	Generalized linear model using South Korean claims data	1-year reimbursement cost
Almony et al [8], 2021[a]	Generalized linear model using US claims data	Generalized linear model using US claims data	1-year reimbursement cost

[a] Economic burden was estimated using a macro-costing approach.

In contrast, analyzing real-world data can provide a more accurate and comprehensive economic burden reflecting clinical practice. The claims, EMR, and other administrative data can be more representative of the entire population and less prone to biases. There were several previous studies analyzing the burden of exudative AMD using real-world data. In the study by Kim et al [13] analyzing administrative data from South Korea, the total reimbursement cost of exudative AMD was 2.13 times to 4.06 times higher than that of nonexudative AMD, and the incremental reimbursement cost was estimated to be US $3699 through the ECM in South Korea. They compared controls as a nonhealthy population with comparable comorbidities using a propensity score model. The methodology overcame the limitations of cost data through ECM and generalizability by using the general population, and a significant 1-year economic burden was identified despite applying a conservative approach through the propensity score matching technique. However, the limitations of using claims data are that they do not reflect uninsured treatment and, therefore, do not define patients with exudative AMD in the real clinical environment, do not reflect nonpayment items such as bevacizumab and optical coherence tomography in the economic burden, and do not estimate the economic burden of real patients. Jeon et al [12] compared exudative AMD with diabetic macular edema and reported only the reimbursement cost with a 1-year follow-up using claims data. Owing to the matching group and changes in reimbursement criteria with exudative AMD, the cost was estimated to be higher than that in the study by Kim et al [13]. The most recent study investigated annual reimbursement costs [8]. Almony et al [8] reported the economic burden of exudative AMD by choroidal neovascularization, and it is the only study to stratify the economic burden by severity, patient-eye level in commercially insured patients in the United States; however, it was a short-term study with 18 months of follow-up and did not compare exudative AMD with any other conditions. The lack of a control group makes comparisons difficult, and the presence of distorting outliers may have overestimated the economic burden. In addition, as in previous studies, the patient-centered economic burden, which comprises nonreimbursed items or out-of-pocket costs, was not calculated. As in previous studies, we also found a high economic burden of exudative AMD and verified its association with exudative AMD using real-world data by adjusting for selection bias and skewed data. Compared with previous studies that only included insurance-covered items and underestimated the patients' economic burden, we included nonreimbursement and out-of-pocket costs that directly affect patients' economic burden. Furthermore, as exudative AMD requires long-term management, we stratified the economic burden by the first and second year and found that there was still a considerable economic burden in the second year.

Policies for Exudative AMD

The economic burden of patients with exudative AMD is substantial worldwide. To address this issue, governments worldwide have expanded the insurance coverage for exudative AMD. In the United States, Medicare covers 80% of the cost of macular degeneration screening and anti-VEGF injections such as aflibercept, ranibizumab, bevacizumab, and verteporfin for Medicare beneficiaries. Similarly, in the United Kingdom, evidence-based medical recommendations are provided through organizations such as the National Institute for Health and Care Excellence, and cost-effectiveness analysis is used to cover anti-VEGFs such as aflibercept and brolucizumab. However, bevacizumab, which is as effective as ranibizumab [4], relatively inexpensive, and cost-effective [24], has not been approved and reimbursed in the United Kingdom or South Korea. In the United States, covering bevacizumab led to US $17.3-billion savings for Medicare and patients from 2008 to 2015 [25].

In South Korea, the pricing system operates on a fee-for-service structure, with the national health insurance acting as a third-party payer covering a portion of medical fees for items eligible for reimbursement [26]. Inpatient care typically involves an out-of-pocket rate of up to 20%. Outpatient services exhibit variations in out-of-pocket rates based on the hospital type (clinics at 30%, hospitals at 40%, general hospitals at 50%, and tertiary hospitals at up to 60% of the reimbursement costs [27]). To mitigate the financial burden on patients with exudative AMD in South Korea, the government has taken proactive measures. Sustained reimbursement has been established for optical coherence tomography measurements and anti-VEGF treatments. In addition, an extended benefit coverage policy has been introduced, reducing cost sharing to 10% for patients with rare and incurable conditions such as exudative AMD. However, as this study showed, the out-of-pocket cost for patients with exudative AMD remains significantly high, and this remains true even in the second year. Therefore, there is a need for a sustained, efficient, and sufficient reimbursement strategy for a high economic disease burden that efficiently uses limited national health insurance finances while reducing the burden on patients. Strategies include discovering new drugs via randomized clinical trials or using real-world data, and approving as well as reimbursing affordable and effective drugs, like bevacizumab, based on cost-effectiveness studies. Such a strategy could help reduce the out-of-pocket cost for patients with exudative AMD; improve their access to treatment; and, ultimately, establish the financial health of insurance by lowering the reimbursement cost. To establish this strategy, real-world data and evidence are needed for diseases with economic burden and rare and intractable diseases, including cost and clinical data.

Economic Burden Analysis Using the OMOP CDM

It is difficult to analyze claims or EMR data because of the sensitive patient information. However, the advent of the OMOP CDM has introduced an environment in which data can be analyzed without disclosing sensitive information. Therefore, many recent studies have analyzed and published EMR or claims data through the OMOP CDM. Previous studies [28-30] using the OMOP CDM have investigated mainly clinical outcomes. In this study, we conducted the first-ever analysis of economic data that had been transformed and validated into the OMOP CDM, and we found that the OMOP CDM could be used for economic analysis with HERMES. The OMOP CDM allowed us to verify the validity of the clinical and cost data and HERMES made it easy to conduct econometric analysis under the OMOP-CDM structure. Compared with previous studies, we could estimate the patient-centered economic burden in the long term by analyzing valid data on costs, including nonreimbursed items, such as bevacizumab, or measurements. By identifying patients who use nonreimbursable treatments and estimating the long-term patient-centered economic burden, our study has significant implications for policy and economic evaluation studies of expensive drugs for rare and incurable diseases such as exudative AMD. Overall, our research highlights the power of the OMOP CDM for unlocking new insights into health care economics and improving the lives of patients with serious medical conditions.

Limitations

This study has a number of limitations. First, the data source for this study was the medical records from a single tertiary hospital, which may not be structured and representative of the entire population of patients with exudative AMD. Nonetheless, to address this issue, the data were transformed into the OMOP CDM and validated by experts in the fields of medicine, data science, and economics. Furthermore, as patients with exudative AMD are more likely to seek treatment in tertiary hospitals, the exudative AMD group is less likely to be overestimated. In addition, to account for selection bias and confounding variables, a propensity score model using LASSO and the ECM were applied, and all the results were found to be robustly significant. Second, we systematically excluded specific medical conditions, including cancer and other ophthalmic diseases treated with anti-VEGF agents, from the cohorts under investigation. Our research focused on the meticulous assessment of the patient-centered economic implications related to exudative AMD. Given the pronounced economic impact associated with cancer, the substantial medical expenditures attributable to cancer could potentially exert an undue influence on the overall economic burden analysis. This influence has the potential to introduce statistical anomalies and distortions within the econometric model, thereby warranting their exclusion from our study cohorts as potential outliers. Moreover, our intention was to delineate the population of real-world patients with exudative AMD with utmost precision. The exclusion of other ophthalmic conditions served the purpose of averting potential misspecifications in our analytical framework. Third, other potential confounding variables such as the type and status of patients' insurance, lifestyle habits, and the stage of exudative AMD were absent from the data set. Therefore, we were unable to include unobserved confounding variables. As we adjusted for selection bias using propensity score matching with LASSO and assessed covariate balance through standardized mean differences, we presume that this limitation's influence on the study findings was negligible. In addition, most of the population in South Korea is covered by national health care insurance services, with just 3% of patients supported by medical aid programs.

Conclusions

In conclusion, exudative AMD had a significantly greater economic impact for 2 years on reimbursement, nonreimbursement, and out-of-pocket costs than nonexudative AMD after adjusting for baseline demographic and clinical characteristics. Although economic policies could relieve the economic burden of patients with exudative AMD over time, the out-of-pocket cost of exudative AMD was still higher than that of nonexudative AMD for 2 years. Our findings support the need for expanding reimbursement strategies for patients with exudative AMD given the significant economic burden faced by patients with incurable and fatal diseases both in South Korea and worldwide.

Acknowledgments

This research was supported by a grant from the Patient-Centered Clinical Research Coordinating Center funded by the Ministry of Health and Welfare, Republic of Korea (grant HC21C0163), and a grant (21153MFDS601) from the Ministry of Food and Drug Safety in 2023. The funding organization had no role in the design or conduct of this study.

Data Availability

As the Observational Medical Outcomes Partnership Common Data Model is a distributed research network, access to the data is restricted to internal private networks, and the data are not publicly available.

Conflicts of Interest

None declared.

References

1. Guymer RH, Campbell TG. Age-related macular degeneration. Lancet 2023 Apr 29;401(10386):1459-1472. [doi: 10.1016/S0140-6736(22)02609-5] [Medline: 36996856]
2. Han X, Chen Y, Gordon I, Safi S, Lingham G, Evans J, et al. A systematic review of clinical practice guidelines for age-related macular degeneration. Ophthalmic Epidemiol 2023 Jun;30(3):213-220. [doi: 10.1080/09286586.2022.2059812] [Medline: 35417274]
3. Studnička J, Říhová B, Rencová E, Rozsíval P, Dubská Z, Chrapek O, et al. Cost and effectiveness of therapy for wet age-related macular degeneration in routine clinical practice. Ophthalmologica 2013;230(1):34-42. [doi: 10.1159/000350802] [Medline: 23751929]
4. Nguyen CL, Oh LJ, Wong E, Wei J, Chilov M. Anti-vascular endothelial growth factor for neovascular age-related macular degeneration: a meta-analysis of randomized controlled trials. BMC Ophthalmol 2018 May 30;18(1):130 [FREE Full text] [doi: 10.1186/s12886-018-0785-3] [Medline: 29843663]
5. Kaiser SM, Arepalli S, Ehlers JP. Current and future anti-VEGF agents for neovascular age-related macular degeneration. J Exp Pharmacol 2021 Sep 29;13:905-912 [FREE Full text] [doi: 10.2147/JEP.S259298] [Medline: 34616189]
6. Rosenberg D, Deonarain DM, Gould J, Sothivannan A, Phillips MR, Sarohia GS, et al. Efficacy, safety, and treatment burden of treat-and-extend versus alternative anti-VEGF regimens for nAMD: a systematic review and meta-analysis. Eye (Lond) 2023 Jan;37(1):6-16. [doi: 10.1038/s41433-022-02020-7] [Medline: 35396574]
7. Wecker T, Grundel B, Reichl S, Stech M, Lange C, Agostini H, et al. Anti-VEGF injection frequency correlates with visual acuity outcomes in pro re nata neovascular AMD treatment. Sci Rep 2019 Mar 01;9(1):3301 [FREE Full text] [doi: 10.1038/s41598-019-38934-8] [Medline: 30824721]
8. Almony A, Keyloun KR, Shah-Manek B, Multani JK, McGuiness CB, Chen C, et al. Clinical and economic burden of neovascular age-related macular degeneration by disease status: a US claims-based analysis. J Manag Care Spec Pharm 2021 Sep;27(9):1260-1272 [FREE Full text] [doi: 10.18553/jmcp.2021.27.9.1260] [Medline: 34464210]
9. Bonastre J, Le Pen C, Anderson P, Ganz A, Berto P, Berdeaux G. The epidemiology, economics and quality of life burden of age-related macular degeneration in France, Germany, Italy and the United Kingdom. Eur J Health Econ 2002;3(2):94-102. [doi: 10.1007/s10198-002-0104-y] [Medline: 15609135]

10. Brown MM, Brown GC, Lieske HB, Tran I, Turpcu A, Colman S. Societal costs associated with neovascular age-related macular degeneration in the united states. Retina 2016 Feb;36(2):285-298. [doi: 10.1097/IAE.0000000000000717] [Medline: 26428606]
11. Cruess AF, Zlateva G, Xu X, Soubrane G, Pauleikhoff D, Lotery A, et al. Economic burden of bilateral neovascular age-related macular degeneration: multi-country observational study. Pharmacoeconomics 2008;26(1):57-73. [doi: 10.2165/00019053-200826010-00006] [Medline: 18088159]
12. Jeon H, Lee H, Yoon D, Lee Y, Kim JH, Jee D, et al. Burden of diabetic macular oedema in patients receiving antivascular endothelial growth factor therapy in South Korea: a healthcare resource use and cost analysis. BMJ Open 2020 Dec 29;10(12):e042484 [FREE Full text] [doi: 10.1136/bmjopen-2020-042484] [Medline: 33376178]
13. Kim S, Park SJ, Byun SJ, Park KH, Suh HS. Incremental economic burden associated with exudative age-related macular degeneration: a population-based study. BMC Health Serv Res 2019 Nov 12;19(1):828 [FREE Full text] [doi: 10.1186/s12913-019-4666-0] [Medline: 31718629]
14. Lotery A, Xu X, Zlatava G, Loftus J. Burden of illness, visual impairment and health resource utilisation of patients with neovascular age-related macular degeneration: results from the UK cohort of a five-country cross-sectional study. Br J Ophthalmol 2007 Oct;91(10):1303-1307 [FREE Full text] [doi: 10.1136/bjo.2007.116939] [Medline: 17504847]
15. Spooner KL, Mhlanga CT, Hong TH, Broadhead GK, Chang AA. The burden of neovascular age-related macular degeneration: a patient's perspective. Clin Ophthalmol 2018;12:2483-2491 [FREE Full text] [doi: 10.2147/OPTH.S185052] [Medline: 30584267]
16. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc 2012;19(1):54-60 [FREE Full text] [doi: 10.1136/amiajnl-2011-000376] [Medline: 22037893]
17. Huser V, DeFalco FJ, Schuemie M, Ryan PB, Shang N, Velez M, et al. Multisite evaluation of a data quality tool for patient-level clinical data sets. EGEMS (Wash DC) 2016 Nov 30;4(1):1239 [FREE Full text] [doi: 10.13063/2327-9214.1239] [Medline: 28154833]
18. Huser V, Kahn MG, Brown JS, Gouripeddi R. Methods for examining data quality in healthcare integrated data repositories. Pac Symp Biocomput 2018;23:628-633 [FREE Full text] [Medline: 29218922]
19. Choi K. HERMES: a health resources econometric analysis tool. GitHub. 2023. URL: https://github.com/kyseonchoi/HERMES/tree/main [accessed 2023-08-29]
20. Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D. Massive parallelization of serial inference algorithms for a complex generalized linear model. ACM Trans Model Comput Simul 2013 Jan;23(1):1-17 [FREE Full text] [doi: 10.1145/2414416.2414791] [Medline: 25328363]
21. Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. Int J Epidemiol 2018 Dec 01;47(6):2005-2014 [FREE Full text] [doi: 10.1093/ije/dyy120] [Medline: 29939268]
22. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? J Health Econ 2001 Jul;20(4):461-494. [doi: 10.1016/s0167-6296(01)00086-8] [Medline: 11469231]
23. Bauhoff S. Self-report bias in estimating cross-sectional and treatment effects. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research. Dordrecht, Netherlands: Springer; 2014:5798-5800.
24. van Asten F, Michels CT, Hoyng CB, van der Wilt GJ, Klevering BJ, Rovers MM, et al. The cost-effectiveness of bevacizumab, ranibizumab and aflibercept for the treatment of age-related macular degeneration-a cost-effectiveness analysis from a societal perspective. PLoS One 2018 May 17;13(5):e0197670 [FREE Full text] [doi: 10.1371/journal.pone.0197670] [Medline: 29772018]
25. Rosenfeld PJ, Windsor MA, Feuer WJ, Sun SJ, Frick KD, Swanson EA, et al. Estimating medicare and patient savings from the use of bevacizumab for the treatment of exudative age-related macular degeneration. Am J Ophthalmol 2018 Jul;191:135-139. [doi: 10.1016/j.ajo.2018.04.008] [Medline: 29655642]
26. Jung HW, Kwon YD, Noh J. How public and private health insurance coverage mitigates catastrophic health expenditures in Republic of Korea. BMC Health Serv Res 2022 Aug 16;22(1):1042 [FREE Full text] [doi: 10.1186/s12913-022-08405-4] [Medline: 35971176]
27. Park S. Medical service utilization and out-of-pocket spending among near-poor National Health Insurance members in South Korea. BMC Health Serv Res 2021 Aug 28;21(1):886 [FREE Full text] [doi: 10.1186/s12913-021-06881-8] [Medline: 34454499]
28. Li X, Burn E, Duarte-Salles T, Yin C, Reich C, Delmestri A, et al. Comparative risk of thrombosis with thrombocytopenia syndrome or thromboembolic events associated with different covid-19 vaccines: international network cohort study from five European countries and the US. BMJ 2022 Oct 26;379:e071594 [FREE Full text] [doi: 10.1136/bmj-2022-071594] [Medline: 36288813]
29. Seo SI, Park CH, You SC, Kim JY, Lee KJ, Kim J, et al. Association between proton pump inhibitor use and gastric cancer: a population-based cohort study using two different types of nationwide databases in Korea. Gut 2021 Nov 11;70(11):2066-2075. [doi: 10.1136/gutjnl-2020-323845] [Medline: 33975868]

30. You SC, Rho Y, Bikdeli B, Kim J, Siapos A, Weaver J, et al. Association of ticagrelor vs clopidogrel with net adverse clinical events in patients with acute coronary syndrome undergoing percutaneous coronary intervention. JAMA 2020 Oct 27;324(16):1640-1650 [FREE Full text] [doi: 10.1001/jama.2020.16167] [Medline: 33107944]

Abbreviations

AMD: age-related macular degeneration
CDM: Common Data Model
ECM: exponential conditional model
EMR: electronic medical record
HERMES: Health Resources Econometric Analysis Tool
LASSO: least absolute shrinkage and selection operator
OHDSI: Observational Health Data Sciences and Informatics
OMOP: Observational Medical Outcomes Partnership
SNUBH: Seoul National University Bundang Hospital
VEGF: vascular endothelial growth factor

Edited by A Mavragani, T Sanchez; submitted 12.06.23; peer-reviewed by G Silvestri, SC You, C Wang; comments to author 11.08.23; revised version received 01.09.23; accepted 12.10.23; published 08.12.23.

Please cite as:
Choi K, Park SJ, Han S, Mun Y, Lee DY, Chang DJ, Kim S, Yoo S, Woo SJ, Park KH, Suh HS
Patient-Centered Economic Burden of Exudative Age-Related Macular Degeneration: Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e49852
URL: https://publichealth.jmir.org/2023/1/e49852
doi:10.2196/49852
PMID:38064251

Original Paper

Association Between Serum Glycated Hemoglobin Levels at Early Gestation and the Risk of Subsequent Pregnancy Loss in Pregnant Women Without Diabetes Mellitus: Prospective Cohort Study

Xiaotian Chen[1], MSc; Yi Zhang[1], MPH; Hongyan Chen[1], PhD; Yalan Dou[1], PhD; Yin Wang[1], PhD; Wennan He[1], MSc; Xiaojing Ma[2,3], MD; Wei Sheng[3], PhD; Weili Yan[1,3,4*], PhD; Guoying Huang[2,3,4*], MD

[1]Department of Clinical Epidemiology & Clinical Trial Unit, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China

[2]Pediatric Heart Center, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai, China

[3]Shanghai Key Laboratory of Birth Defects, Children's Hospital of Fudan University, Shanghai, China

[4]Research Unit of Early Intervention of Genetically Related Childhood Cardiovascular Diseases (2018RU002), Chinese Academy of Medical Sciences, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Weili Yan, PhD
Department of Clinical Epidemiology & Clinical Trial Unit
Children's Hospital of Fudan University
National Children's Medical Center
399 Wan Yuan Road
Shanghai, 201102
China
Phone: 86 021 64931215
Fax: 86 021 64931215
Email: yanwl@fudan.edu.cn

Abstract

Background: As a severe morbidity during pregnancy, the etiology of spontaneous pregnancy loss (SPL) remains largely unknown. Serum glycated hemoglobin (HbA_{1c}) level is an established predictor of SPL risk among women with diabetes, but little is known about whether such an association exists among pregnant women without diabetes when glycemic levels are within the normal range.

Objective: This study aimed to quantify the association between maternal HbA_{1c} levels in early pregnancy and subsequent SPL risk in a cohort of pregnant women without diabetes.

Methods: This prospective cohort study involved 10,773 pregnant women without diabetes enrolled at their first antenatal care visit at a hospital's early pregnancy clinic from March 2016 to December 2018 in Shanghai, China. HbA_{1c} and fasting blood glucose (FBG) levels were examined at enrollment. Participants with diabetes before or pregnancy or those diagnosed with gestational diabetes were excluded. Diagnosis of SPL, defined as fetal death occurring before 28 gestational weeks, was derived from medical records and confirmed via telephone interviews. We used generalized linear models to quantify the associations of continuous and dichotomized maternal HbA_{1c} levels with SPL risk and reported crude and adjusted risk ratios (RRs) and 95% CIs. A restricted cubic spline (RCS) regression model was used to assess the potential nonlinear dose-response relationship. Adjusted covariates included maternal age, education level, preconception BMI, gestational weeks, gravidity, history of adverse pregnancy outcomes, family history of diabetes, folic acid supplementation, and smoking and drinking during the periconception period.

Results: In total, 273 (2.5%) SPL cases occurred. Every 0.5% increase in HbA_{1c} levels was linearly associated with a 23% increase in SPL risk (adjusted RR [aRR] 1.23; 95% CI 1.01-1.50). The RCS model revealed that this association was linear (*P*=.77 for the nonlinearity test). Analyses based on dichotomized HbA_{1c} levels showed a significantly increased risk of SPL when HbA_{1c} levels were ≥5.9% (aRR 1.67; 95% CI 0.67-3.67), and the significance threshold was ≥5.6% (aRR 1.60; 95% CI 1.01-2.54). Sensitivity analyses showed similar results when including the participants with missing SPL records or HbA_{1c} data. Linear associations of HbA_{1c} levels remained significant even in the subgroups without overweight, alcohol consumption, and a family history of diabetes and adverse pregnancy outcomes. Every 1 mmol/L increment in maternal FBG levels was associated with a >2-fold higher risk of SPL (aRR 2.12; 95% CI 1.61-2.80; $P<.001$).

Conclusions: Higher HbA_{1c} levels in early pregnant women without diabetes are associated with an increased SPL risk in a dose-response manner. Pregnant women with an HbA_{1c} level above 5.6% at early gestation need attention for its potentially increased risk for SPL. Our findings support the need to monitor HbA_{1c} levels to identify individuals at high risk of subsequent SPL in the general population of pregnant women.

Trial Registration: ClinicalTrials.gov NCT02737644; https://clinicaltrials.gov/study/NCT02737644

(JMIR Public Health Surveill 2023;9:e46986) doi:10.2196/46986

KEYWORDS

glycated hemoglobin; hemoglobin A1c; spontaneous pregnancy loss; gynecology; gynecological; obstetric; obstetrics; prospective cohort study; cohort; risk; risks; miscarriage; miscarriages; adverse outcome; adverse outcomes; risk ratio; pregnant women; pregnancy loss; gestational diabetes; fetal death; glycemic control; women; diabetes; diabetic; HbA1c; gestational; maternal; fetus; fetal; HbA1c levels; metabolic health; pregnant; pregnancy; association; associations; associated

Introduction

Spontaneous pregnancy loss (SPL), also known as spontaneous abortion, is one of the serious morbidities during pregnancy and precedes an increased risk of reduced fertility, long-term depression, and anxiety among pregnant women [1]. As the definition of SPL varies between countries and international organizations, estimates of the prevalence of SPL vary among previous studies [2]. Data from a large nationally representative survey from the United States showed that approximately 20% of clinically recognized pregnancies ended in SPL (including stillbirths and ectopic or tubal pregnancies) during the whole gestation period [3]. Pregnancy losses from the first antenatal care visit to that at 28 weeks' gestation are recorded in practice in China [4,5]. A nationwide study with 6.4 million medical records of Chinese pregnant women reported 2.8% of pregnancy losses before 28 weeks of gestation [6]. To date, the mechanism underlying the etiology of SPL remains largely unknown, and over 50% of women with SPL have no identified risk factors [7,8].

Poor glycemic control during pregnancy is an established independent predictor for adverse pregnancy outcomes [9]. Serum glycated hemoglobin (HbA_{1c}) levels are conventionally used for monitoring blood glucose control [10]. According to the American Diabetes Association's (ADA's) clinical practice guidelines, individuals with HbA_{1c} levels within 5.7%-6.4% and ≥6.5% are classified as having prediabetes and diabetes, respectively, in the general population, and HbA_{1c} levels of 5.9% and higher in pregnant women are considered an early indicator of abnormal glucose metabolism and a higher risk of adverse pregnancy and neonatal outcomes [11]. Previous studies have reported associations of HbA_{1c} levels of pregnant women with insulin-dependent diabetes and SPL risk [12-14]. However, little is known about these associations in pregnant women without diabetes.

An HbA_{1c} test at the first prenatal visit has been recommended for those at risk of developing gestational diabetes only, including having obesity and having a family history of gestational diabetes, but not yet for the general population of pregnant women [11,15,16]. In this prospective cohort study, we aimed to evaluate whether HbA_{1c} levels in pregnant women without diabetes are associated with a subsequent SPL risk.

Methods

Study Population

Pregnant women included in this study were a subcohort of the ongoing Shanghai Preconception Cohort Study (SPCC; ClinicalTrials.gov NCT02737644) [17], who were enrolled between March 2016 and December 2018 from one of the study sites—a tertiary maternity hospital where HbA_{1c} and fasting blood glucose (FBG) levels were routinely examined at the first antenatal visit for all pregnant women. This maternity hospital is one of the largest delivery hospitals (with >20,000 births per year) during the study period and accounts for over 20% of the annual deliveries in the city. Participants were enrolled at their first antenatal visit at early pregnancy clinics, and each woman had only one medical record for this study. Biochemistry and SPL diagnosis data were extracted from the hospital's electronic medical record system. Among all pregnant women during the study period (N=13,129), 10,773 were eligible for the primary analysis after the exclusion of those who met any one of the following criteria: missing medical records after the first antenatal visit, missing information regarding HbA_{1c} or FBG levels at entry, having received artificial abortions, self-reported diabetes before pregnancy (ie, an HbA_{1c} level of ≥6.5% or FBG level of ≥7.0 mmol/L at the first antenatal visit), or having received insulin treatment during pregnancy or taking oral hypoglycemic drugs before or during pregnancy. We defined this group as pregnant women without diabetes.

Ethical Considerations

Ethics approval for this study was sought from the institutional Ethics Committee of the Children's Hospital of Fudan University (2016–49). Written informed consent was obtained from all participants before recruitment. All data were anonymously analyzed.

Exposures and Covariates

We treated the fasting HbA_{1c} levels in early pregnancy as the main exposure in this study. Given that FBG levels are also an index of glycemic control, both HbA_{1c} and FBG levels were abstracted from the medical records for analyses. According to the routine practice of the hospital where this subcohort was recruited, HbA_{1c} and FBG levels were measured in 2 hours using a venous blood sample after overnight fasting (>8 hours of fasting) at their first antenatal care visit. Venous blood for the HbA_{1c} test was collected in an EDTA-containing tube and determined using high-performance liquid chromatography (Bio-Rad) in the hospital's certified standard clinical examination center following standard protocols.

A series of variables regarding known or suspected risk factors for SPL were considered covariates in the association analysis [18]. As described elsewhere, demographic characteristics and pregnancy history of the participants upon enrollment were collected through a prespecified standard self-administered questionnaire and an interview with the obstetric nurse during the first antenatal visit. Maternal preconception BMI (pre-BMI) was calculated using self-reported measures of prepregnancy body weight and categorized as normal weight or overweight. We used both the Chinese standard (≥24 kg/m^2) and the international standard (≥25 kg/m^2) to define the overweight status [19,20]. We defined a family history of diabetes as having at least 1 first-degree relative diagnosed with diabetes. Smoking exposure and alcohol drinking were defined as smoking cigarettes or having been exposed to second-hand cigarette smoke, and as consuming any alcoholic beverages within 3 months before or during the current pregnancy, respectively. Folic acid supplementation (FAS) was defined as having a regular intake of pure folic acid tablets or multivitamins containing folic acid before or during early pregnancy. Gestational weeks at enrollment were routinely determined by the last menstruation period and confirmed through a routine ultrasonographic examination. Participants were defined as having a history of adverse pregnancy outcomes if they had abortions, preterm delivery, stillbirths, or ectopic pregnancy in previous pregnancies.

Outcomes

The recorded SPL cases from the hospital's electronic medical system were fetal deaths occurring before 28 gestational weeks in accordance with the Chinese clinical guidelines [4]. Stillbirth (fetal deaths after 28 gestational weeks), and artificial abortions due to ectopic pregnancies, molar pregnancies, or any clinically recognized disorders were not considered SPL in this study. Trained staff verified these diagnoses through a personal telephone interview with pregnant women or their husbands before the analysis.

Statistical Analysis

Continuous variables were reported as mean (SD) values for a normal or approximate normal distribution and median (IQR) values for a skewed distribution, and 2-tailed unpaired Student *t* tests or Mann-Whitney *U* tests were used for comparisons between the SPL and non-SPL group, respectively. Normality was visually inspected using frequency histograms. Categorical variables were summarized as frequencies and percentages, and chi-square tests were used for the group comparisons.

Our primary aim was to investigate the associations of HbA_{1c} and FBG levels as continuous variables with SPL risk. We used generalized linear models with binomial family and log link functions treating HbA_{1c} (rescaled through dividing by 0.5) and FBG levels as continuous variables to estimate crude and adjusted risk ratios (aRRs) and 95% CIs. A 0.5% absolute increment was chosen for the HbA_{1c} level because it reflects a clinically important change [21]. Considering that a HbA_{1c} level of 5.9% in early pregnancy has been suggested as the cutoff for identifying women at increased risk of adverse pregnancy outcomes [11], we further assessed the associations based on dichotomous HbA_{1c} (coded as 1 for a HbA_{1c} level of ≥5.9% and 0 for an HbA_{1c} level of <5.9%). We used a restricted cubic spline (RCS) regression model with 3 knots (5th, 50th, and 95th percentile levels) fitted in R (*rms* package) to assess the potential nonlinear dose-response relationship of HbA_{1c} levels with SPL risk. Adjusted covariates included maternal age, education level, pre-BMI, gestational week at HbA_{1c} examination, gravidity, history of adverse pregnancy outcomes, family history of diabetes, FAS, and exposure to smoking and drinking during the periconception period.

We conducted an exploratory analysis to investigate the HbA_{1c} cutoff indicating pregnant women at an increased risk of SPL by using a series of regression models using HbA_{1c} levels from the median level of the study population (5.1%) to 6.0% in a 0.1% interval without adjustment for multiple testing.

As there were 2147 participants—1360 (10.5%) with missing medical records after the first antenatal visit and 787 (6.1%) with missing HbA_{1c} data—we conducted 2 sensitivity analyses to test the robustness of the main results after multiple imputation with chained equations based on a missing at random assumption. First, we added 1360 pregnant women with missing SPL data after imputation (sensitivity analysis 1: n=10,773+1360). Second, we included 787 pregnant women with missing HbA_{1c} levels at enrollment after imputation (sensitivity analysis 2: n=10,773+787). To ensure that our main results are free from potential bias from unmeasured confounders, we further repeated the primary association analyses within subgroups with a potentially low risk of SPL, including pregnant women younger than 35 years, without overweight, a family history of diabetes, a history of adverse pregnancy outcomes, or smoking or alcohol drinking exposures. Statistical analyses were performed by Stata (version 16.0; StataCorp) and R package (version 3.6.1; The R Foundation), and all statistical tests were 2-sided at a significance level of .05. A post hoc power analysis was performed for the main association analyses of continuous HbA_{1c} levels with SPL risk, which revealed that with the current sample size, associations with a risk ratio (RR) lower than 0.84 or greater than 1.19 for continuous exposures of interest will ensure a power of ≥80% at an α level of 5%.

Results

Characteristics of the Study Population

A total of 10,773 eligible pregnant women were included in the main analysis (Figure 1), with a mean age of 30.5 (SD 4.0) years and an average of 10.8 (SD 1.7) weeks of gestation at enrollment (Table 1). The majority of participants (n=9866, 91.6%) had a college or higher level of education, 4394 (40.9%) had more than 1 gravidity, 2056 (19.1%) had overweight before pregnancy, and 3874 (36.3%) had a family history of diabetes. Smoking and alcohol consumption were reported by 1043 (9.8%) and 857 (8.0%) pregnant women, respectively. Besides, 2817 (26.4%) pregnant women had a history of adverse pregnancy outcomes. The distribution of HbA_{1c} and FBG levels approximately met a normal distribution, with a mean value of 5.09% (SD 0.30%) and 4.49 (SD 0.78) mmol/L, respectively. The 2 markers were weakly correlated (r=.19; P<.001) and overlapped in quartiles (Figure S1 in Multimedia Appendix 1). We summarized the characteristics of the 10,773 eligible pregnant women and 2147 women with missing medical records or HbA_{1c} data (Table S1 in Multimedia Appendix 1), which were overall similar.

Figure 1. Flowchart of the study population. HbA_{1c}: serum glycated hemoglobin; SPCC: Shanghai Preconception Cohort Study; SPL: spontaneous pregnancy loss.

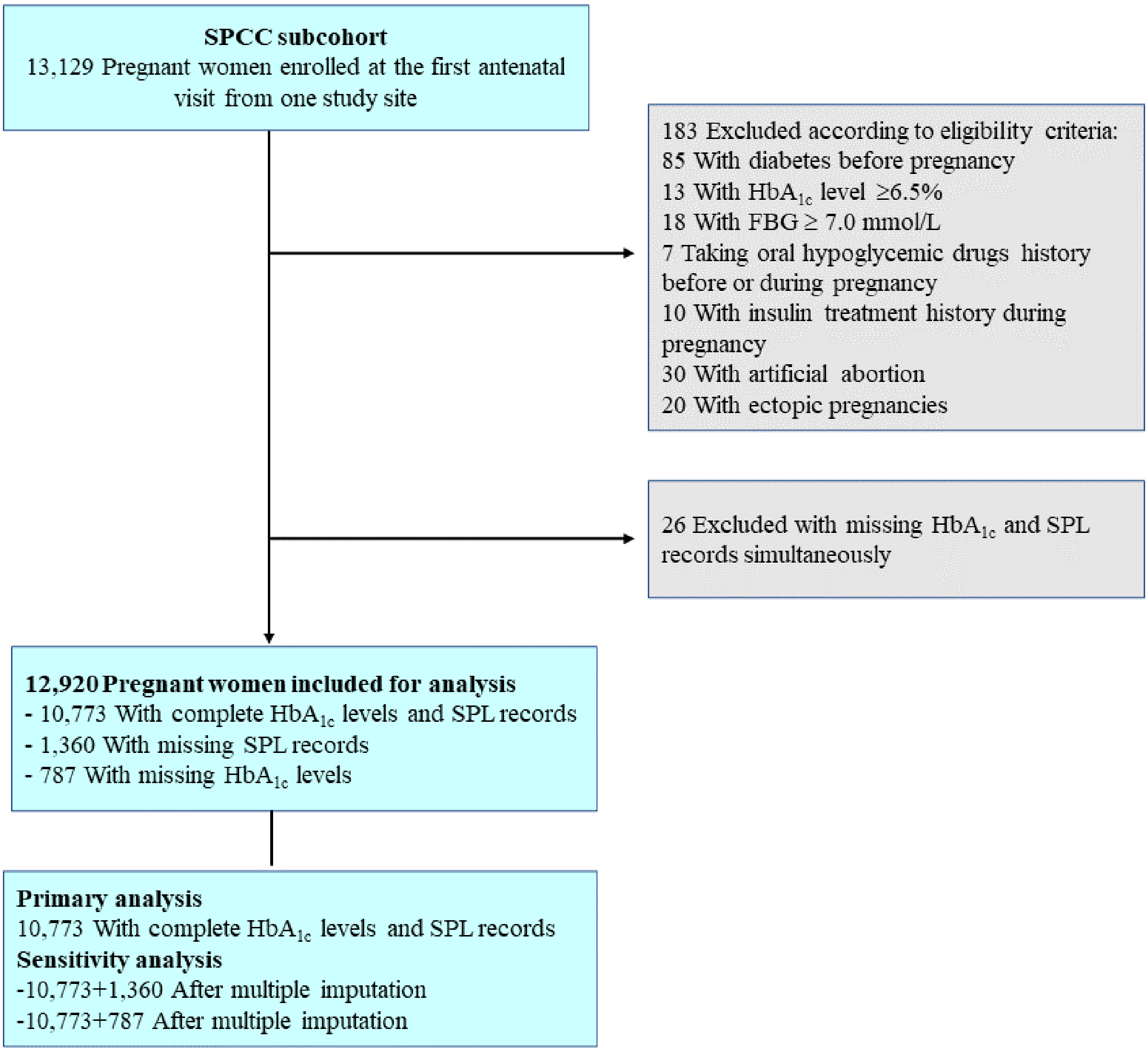


Table 1. Baseline characteristics of study population at the first antenatal care visit.

Variables	Total sample (N=10,773)	SPL[a] group (n=273)	Non-SPL group (n=10,500)	*P* value
Age[b] (years), mean (SD)	30.5 (4.0)	32.0 (4.4)	30.4 (3.9)	<.001
Pre-BMI (kg/m^2), mean (SD)	21.7 (2.9)	22.5 (3.0)	21.7 (2.9)	<.001
Pre-BMI categories, n (%)				
≥24 kg/m^2	2056 (19.1)	65 (23.8)	1991 (19.0)	.047
≥25 kg/m^2	1347 (12.5)	47 (17.2)	1300 (12.4)	.02
Missing	32 (0.3)	—[c]	32 (0.3)	N/A[d]
Gestational weeks at enrollment[e], mean (SD)	10.8 (1.7)	10.7 (1.6)	10.7 (1.7)	.48
Education level, n (%)				.24
College and above	9866 (91.6)	247 (90.5)	9619 (92.4)	
Below college	817 (7.6)	26 (9.5)	791 (7.6)	
Missing	90 (0.8)	—	90 (0.8)	
Gravidity, n (%)				.26
1	6357 (59.1)	149 (54.6)	6208 (59.2)	
2	2759 (25.7)	75 (27.5)	2684 (25.6)	
≥3	1635 (15.2)	49 (18.9)	1586 (15.1)	
Missing	22 (0.2)	2 (0.7)	20 (0.1)	
History of adverse pregnancy outcomes, n (%)	2817 (26.4)	93 (34.1)	2724 (26.2)	.003
Missing	93 (0.9)	1 (0.4)	92 (0.9)	N/A
Family history of diabetes, n (%)	3874 (36.3)	112 (41.0)	3762 (36.2)	.10
Missing	104 (1.0)	5 (1.8)	99 (0.9)	N/A
FAS[f] before or during early pregnancy, n (%)	8022 (75.1)	240 (87.9)	7804 (75.0)	.06
Missing	90 (0.8)	4 (1.5)	86 (0.8)	N/A
Smoking exposure, n (%)	1043 (9.8)	33 (12.1)	1010 (9.7)	.19
Missing	91 (0.8)	2 (0.7)	89 (0.8)	N/A
Alcohol drinking, n (%)	857 (8.0)	19 (7.0)	838 (8.1)	.51
Missing	91 (0.8)	3 (1.1)	88 (0.8)	N/A
FBG[g,h] (mmol/L), mean (SD)	4.49 (0.78)	4.64 (0.39)	4.48 (0.47)	<.001
FBG≥6.1 mmol/L, n (%)	29 (.3)	1 (.4)	28 (.3)	.75
HbA_{1c}[i] (%), mean (SD)	5.09 (0.30)	5.15 (0.33)	5.09 (0.30)	.002
HbA_{1c}≥5.9%, n (%)	126 (1.2)	6 (2.2)	120 (1.1)	.03

[a] SPL: spontaneous pregnancy loss.

[b] Age data were missing for 23 out of 10,773 (0.2%), 3 out of 273 (1.1%), and 20 out of 10,500 (0.2%) pregnant women in the total sample, SPL group, and non-SPL group, respectively.

[c] Not available.

[d] N/A: not applicable.

[e] Data for gestational weeks at enrollment were missing for 60 out of 10,773 (0.6%), 2 out of 273 (0.7%), and 58 out of 10,500 (0.5%) pregnant women in the total sample, SPL group, and non-SPL group, respectively.

[f] FAS: folic acid supplementation.

[g] FBG: fasting blood glucose.

[h] FBG levels were ≥6.1% for 29 out of 10,773 (0.3%), 1 out of 273 (0.4%), and 28 out of 10,500 (0.3%) pregnant women in the total sample, SPL group, and non-SPL group, respectively (*P*=.75).

[i] HbA_{1c}: serum glycated hemoglobin.

The incidence rate of SPL was 2.5% among 10,773 pregnant women in this study. Compared to pregnant women without SPL, those with SPL were 1.6 years older (mean ages of the SPL and non-SPL groups were 32.0, SD 4.4 vs 30.4, SD 3.9 years, respectively) during pregnancy, were more likely to have overweight before pregnancy (23.8% vs 19.0%), and had higher levels of HbA_{1c} (5.15% SD 0.33% vs 5.09% SD 0.30%; albeit in the clinically normal range) and FBG (4.64, SD 0.39 vs 4.48, SD 0.47 mmol/L) at the first antenatal visit.

Associations Between Maternal HbA1c and FBG Levels and SPL Risk

As shown in Table 2, maternal HbA_{1c} levels showed significant positive associations with SPL risk with an unadjusted RR of 1.34 (95% CI 1.10-1.63) per 0.5% increase in HbA_{1c} levels, and the association remained significant after adjusting for covariates (aRR 1.23, 95% CI 1.01-1.50; *P*=.04). The RCS model showed a linear association of SPL risk with increasing HbA_{1c} levels through the whole range of HbA_{1c} values (*P*=.77 for the nonlinearity test; Figure 2). No significant association was found between a HbA_{1c} level of ≥5.9% and SPL risk after adjusting for covariates. Every 1 mmol/L increment in maternal FBG levels was associated with a >2-fold higher risk of SPL (aRR 2.12, 95% CI 1.61-2.80; *P*<.001).

Table 2. Associations between maternal HbA_{1c}[a] and FBG[b] levels and SPL[c] risk during early pregnancy.

Analysis	SPL group/total sample, n/n	Crude model		Adjusted model[d]	
		Risk ratio (95% CI)	*P* value	Risk ratio (95% CI)	*P* value
Primary analysis					
HbA_{1c} (per 0.5% increment)	273/10,773	1.34 (1.10-1.63)	.004	1.23 (1.01-1.50)	.04
<5.9%	267/10,647	Reference	N/A[e]	Reference	N/A
≥5.9%	6/126	1.94 (.85-4.45)	.12	1.67 (0.76-3.67)	.20
FBG (mmol/L)	273/10,773	2.38 (1.82-3.11)	<.001	2.12 (1.61-2.80)	<.001
Sensitivity analysis 1 (Including 1360 pregnant women with missing medical diagnosis records based on imputation)					
HbA_{1c} (per 0.5% increment)	306/12,133	1.36 (1.11-1.67)	.003	1.25 (1.02-1.53)	.03
<5.9%	299/11,993	Reference	N/A	Reference	N/A
≥5.9%	7/140	2.02 (0.93-4.42)	.08	1.51 (0.68-3.34)	.31
Sensitivity analysis 2 (including 787 pregnant women with missing HbA_{1c} levels based on imputation)					
HbA_{1c} (per 0.5% increment)	373/11,560	1.36 (1.10-1.69)	.005	1.25 (1.01-1.55)	.04
<5.9%	366/11,430	Reference	N/A	Reference	N/A
≥5.9%	7/130	1.59 (0.70-3.65)	.27	1.42 (0.49-2.67)	.75

[a] HbA_{1c}: serum glycated hemoglobin.

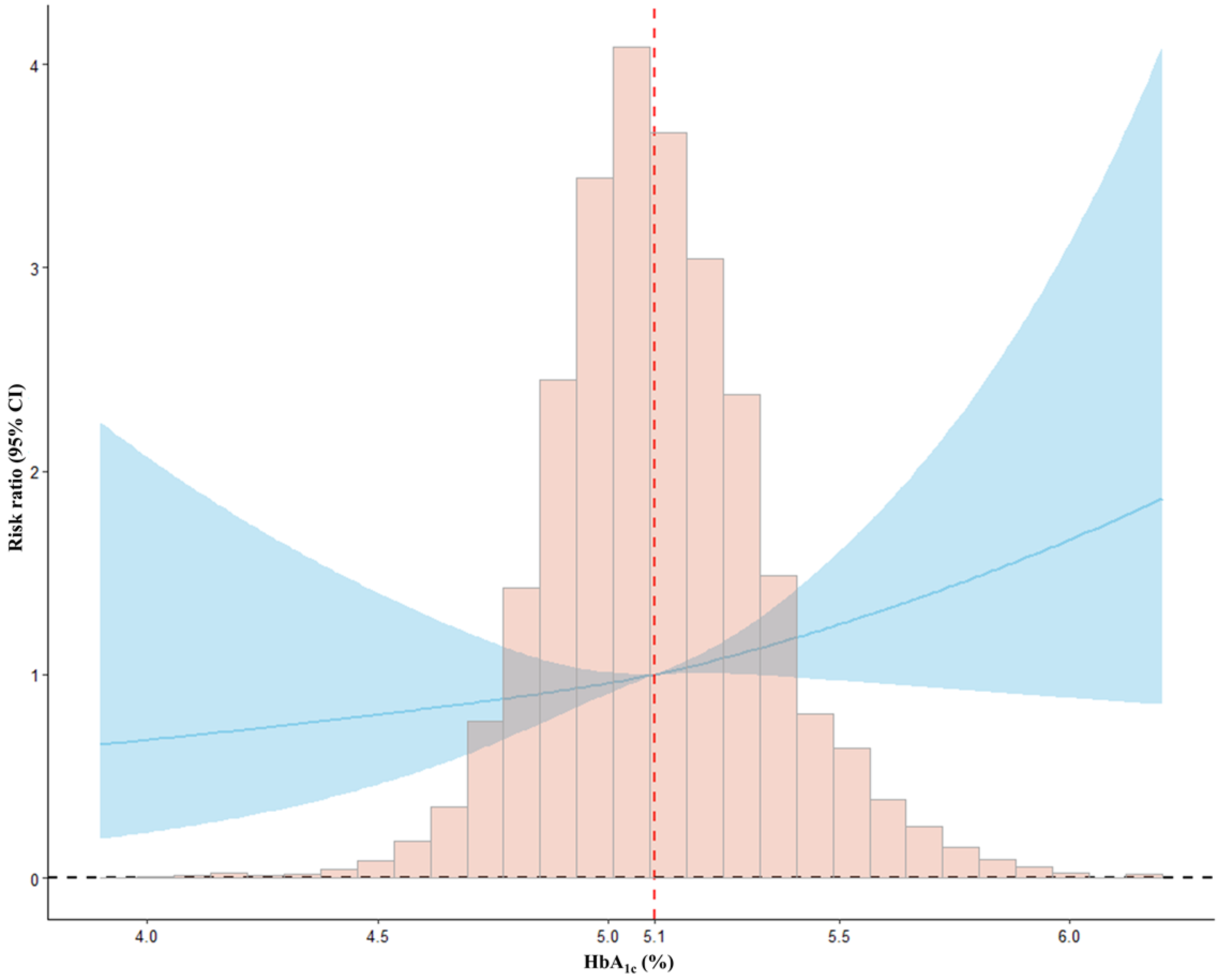
[b] FBG: fasting blood glucose.

[c] SPL: spontaneous pregnancy loss.

[d] Adjusted for age, pre-BMI, gestational weeks, education, gravidity, history of abnormal pregnancy, family history of diabetes, folic acid supplementation, and drinking and smoking status.

[e] N/A: not applicable.

Figure 2. Restricted cubic spline plots for the association between maternal serum glycated hemoglobin (HbA_{1c}) levels in early pregnancy with spontaneous pregnancy loss (SPL) risk.



Sensitivity and Subgroup Analyses

In the first sensitivity analysis including 1360 pregnant women with missing medical diagnosis records, the association between HbA_{1c} levels and SPL risk did not substantially change (aRR per 0.5% increment 1.25, 95% CI 1.02-1.53; *P*=.03; Table 2). Similar results were also observed in the second sensitivity analysis upon including pregnant women with missing data on HbA_{1c} levels (aRR per 0.5% increment 1.25, 95% CI 1.01-1.55; *P*=.04).

Our exploratory analyses showed that the strength of the associations increased markedly from below 1.2 to 1.6 at an HbA_{1c} level of 5.6% (aRR 1.60; 95% CI 1.01-2.54; *P*=.048) and increased further at higher cutoff levels, although significance was not achieved at HbA_{1c} levels of 5.9% and 6.0% (Figure S2 in Multimedia Appendix 1). In further subgroup analyses including low-SPL risk populations without overweight, alcohol drinking, a family history of diabetes, and a history of adverse pregnancy outcomes, the associations of every 0.5% increase in HbA_{1c} levels with SPL risk remained very similar compared to those in the main analysis (aRR 1.32, 95% CI 1.07-1.63; aRR 1.18, 95% CI 1.01-1.38; aRR 1.24, 95% CI 1.03-1.48; and aRR 1.73, 95% CI 1.26-2.36; respectively; Table S2 in Multimedia Appendix 1).

Discussion

Principal Findings

In this large prospective cohort study, we provide solid evidence that among pregnant women without diabetes, HbA_{1c} levels in early gestation within the clinically normal range were associated with an increased risk of SPL in a linear dose-response manner. Although far below the recommended threshold for diagnosing gestational diabetes, HbA_{1c} levels may indicate an increased subsequent SPL risk in pregnant women in general. Our findings in a prospective cohort are novel and deepen our understanding of the important pathophysiologic role of impaired maternal glycemic metabolism in the development of SPL.

Comparison With Prior Work

Several studies have investigated the adverse effect of elevated HbA_{1c} levels on SPL risk only among pregnant women with diabetes and have reported conflicting findings [12-14,22]. A case-control study including 432 control women and 386 women with type 1 diabetes found that the rate of SPL did not significantly differ between the 2 groups (16.1% vs 16.2%, respectively) [12]. However, another case-control study among women with type 1 diabetes reported that elevated HbA_{1c} levels were associated with SPL risk [13]. Another cohort study of 573 women with type 1 diabetes observed a linear association between HbA_{1c} levels and the risk of adverse pregnancy outcomes including SPL when HbA_{1c} levels were >7.0% [14]. This is the first prospective cohort study demonstrating the link between maternal HbA_{1c} levels in early pregnancy and SPL risk among pregnant women without diabetes. These associations were still significant in the population at a low risk of SPL. We also observed a highly overlapped distribution of FBG levels among HbA_{1c} categories, and consistent associations with SPL were also observed with regard to FBG levels. However, regarding their clinical application, FBG level is a less ideal marker than HbA_{1c} level because the former is less stable, requires a fasting state for examination, and has relatively greater intraindividual variability [23]. Our findings suggest that attention to glycemic control should not be limited to pregnant women with diabetes but should also include those with high HbA_{1c} levels within the clinical range during early gestation for the related increased risk of SPL.

Research Implications

Our study addresses 2 important issues with important clinical and research implications. First, our findings expand on the literature on risk factors for SPL, an adverse pregnancy outcome whose modifiable determinants remain poorly understood. Second, we address a risk factor that is known to be modifiable through lifestyle and pharmacological interventions. Measurement of HbA_{1c} levels during pregnancy is conventionally used to monitor glycemic control in pregnant women with diabetes [24]; this marker is superior to FBG levels, mainly for its greater stability but lesser variability among individuals and its nonrequirement of fasting [23]. Given the special physiological status of pregnancy, the target HbA_{1c} level for ideal glycemic control remains inconclusive. In China, based on recommendations for the general population, an HbA_{1c} level less than 6.5% is recommended to indicate ideal glycemic control before conception [25]. An HbA_{1c} level of ≥5.9% is recommended by the ADA's guidelines as an indicator for screening pregnant women with a higher risk of preeclampsia, macrosomia, shoulder dystocia, and perinatal death [11]. Our findings add new evidence for the linear relationship between glucose metabolism markers and SPL risk when HbA_{1c} levels are clinically normal [11], and the magnitudes of the associations are even stronger in those exposed to smoking or those who have a history of adverse pregnancy outcomes.

The mechanism underlying the association between maternal HbA_{1c} levels and SPL in women without diabetes remains unclear. Animal studies found that poor glycemic control may facilitate premature programmed cell death of key progenitor cells of the blastocyst and promote pregnancy loss [26]. In addition, poor glycemic control may interfere with implantation by inhibiting trophectoderm differentiation and increasing oxidative stress, thus affecting the expression of critical genes essential for embryogenesis [27,28]. Further population studies exploring the biological mechanisms underlying the association between HbA_{1c} levels and SPL in pregnant women without diabetes are warranted.

The findings of our exploratory analysis are clinically relevant, indicating that pregnant women with HbA_{1c} levels above 5.6% need attention for owing to a potentially increased risk of SPL. Given that HbA_{1c} levels are 0.5%-0.6% lower in early pregnancy than in nonpregnant women [29], HbA_{1c} levels over 5.6% at early gestation may reflect an elevated HbA_{1c} status that occurred before or at pregnancy (ie, prediabetes status), which accounts for over 13% of the population [6]. Another large cohort study and systematic review have reported associations between an HbA_{1c} level of <6.4% with increased severe maternal morbidity and subsequent gestational diabetes mellitus [30,31]. Taken together, based on the aforementioned evidence and our findings, we propose that monitoring of HbA_{1c} levels in early pregnancy is necessary general pregnant women, and HbA_{1c} levels exceeding 5.6% might be considered an indicator of high risk for SPL and additional medical care.

Strengths and Limitations

The merits of our study include the prospective nature of the data, the large sample size, and the consistent results from robust analyses, which make our findings convincing. However, several limitations exist. First, residual confounding, such as uterine abnormalities, chromosomal abnormalities, antiphospholipid syndrome, and thyroid disorders, cannot be completely ruled out due to the nature of observational studies. Data on important biomarkers such as insulin levels during early gestation were not available. Second, in the exploratory analyses, we did not correct for multiplicity. Prospective multicenter cohort studies investigating whether a maternal HbA_{1c} level of 5.6% in early pregnancy could indicate an increased risk of SPL are warranted in the future. Third, compared with the rate of clinically registered pregnant women ending in SPL (11%-20%), the SPL rate in this cohort was much lower. Given the nature of the study population, which was recruited at their first antenatal care visit (with a median of 10 weeks of gestation), SPL occurred earlier than our observation was inevitably missed, selection bias exists, and our findings are not generalizable to the entire general population. The study population we selected, comprising pregnant women without diabetes, may include some cases of mild gestational diabetes diagnosed midgestation and without insulin treatment. Moreover, potential biases such as outcome misclassification could not be ruled out from the association analyses based on imputation.

Conclusions

This study is the first to document that maternal HbA_{1c} levels in early pregnancy are associated with the subsequent risk of SPL in a dose-response manner among pregnant women without diabetes. Our findings support the need to monitor HbA_{1c} levels for identifying high risk of subsequent SPL in pregnant women in general and expand on the growing literature linking overall metabolic health to reproductive and pregnancy health among otherwise healthy women.

Acknowledgments

This study is supported by National key research and development program (grant 2021YFC2701004), Natural Science Foundation of China (grants 82070323, 82373584, and 82204051), Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (grant 2019-I2M-5-002), Shanghai Sailing Program (grant 22YF1403900), and Youth Project of Shanghai Municipal Commission of Health (grant 20214Y0439). The authors gratefully thank all participants in this study. We are grateful for the work of physicians, interviewers, laboratory technicians, and the whole Shanghai Preconception Cohort Study (SPCC) team. Members of the SPCC group are listed in Multimedia Appendix 1. We declare that artificial intelligence (AI) was not used in writing any portion of the manuscript.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

WY, GH, and XC conceived the hypothesis. XC performed the statistical analyses and drafted the manuscript. YZ, XM, and WS were involved in recruitment and data collection. YZ, XC, and YW examined biological samples and contributed to data management. HC, WY, and GH verified the underlying data and critically reviewed the manuscript. All authors are involved in the interpretation of the results and approve the manuscript before submission. XC, WY, and GH obtained funding and had full access to all the data. WY and GH are co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional figures, tables, and information.
[DOCX File , 442 KB - publichealth_v9i1e46986_app1.docx]

References

1. Blackmore ER, Côté-Arsenault D, Tang W, Glover V, Evans J, Golding J, et al. Previous prenatal loss as a predictor of perinatal depression and anxiety. Br J Psychiatry 2011 May 01;198(5):373-378 [FREE Full text] [doi: 10.1192/bjp.bp.110.083105] [Medline: 21372060]
2. Quenby S, Gallos ID, Dhillon-Smith RK, Podesek M, Stephenson MD, Fisher J, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. Lancet 2021 May 01;397(10285):1658-1667. [doi: 10.1016/S0140-6736(21)00682-6] [Medline: 33915094]
3. Rossen LM, Ahrens KA, Branum AM. Trends in risk of pregnancy loss among US women, 1990-2011. Paediatr Perinat Epidemiol 2018 Jan 01;32(1):19-29 [FREE Full text] [doi: 10.1111/ppe.12417] [Medline: 29053188]
4. Writing Group of Chinese Expert Consensus on Diagnosis and Treatment of Spontaneous Abortion. Chinese expert consensus on diagnosis and treatment of spontaneous abortion (2020 edition). Chin J Pract Gynecol Obstet 2020 Dec 10;11:1082-1090 [FREE Full text] [doi: 10.19538/j.fk2020110113]
5. Qiao J, Wang Y, Li X, Jiang F, Zhang Y, Ma J, et al. A Lancet Commission on 70 years of women's reproductive, maternal, newborn, child, and adolescent health in China. Lancet 2021 Jun 26;397(10293):2497-2536. [doi: 10.1016/S0140-6736(20)32708-2] [Medline: 34043953]
6. Wei Y, Xu Q, Yang H, Yang Y, Wang L, Chen H, et al. Preconception diabetes mellitus and adverse pregnancy outcomes in over 6.4 million women: a population-based cohort study in China. PLoS Med 2019 Oct 01;16(10):e1002926-e1002941 [FREE Full text] [doi: 10.1371/journal.pmed.1002926] [Medline: 31574092]
7. Alijotas-Reig J, Garrido-Gimenez C. Current concepts and new trends in the diagnosis and management of recurrent miscarriage. Obstet Gynecol Surv 2013 Jun 01;68(6):445-466. [doi: 10.1097/OGX.0b013e31828aca19] [Medline: 23942472]
8. van Dijk MM, Kolte AM, Limpens J, Kirk E, Quenby S, van Wely M, et al. Recurrent pregnancy loss: diagnostic workup after two or three pregnancy losses? A systematic review of the literature and meta-analysis. Hum Reprod Update 2020 Apr 15;26(3):356-367 [FREE Full text] [doi: 10.1093/humupd/dmz048] [Medline: 32103270]
9. HAPO Study Cooperative Research Group, Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 2008 May 08;358(19):1991-2002. [doi: 10.1056/NEJMoa0707943] [Medline: 18463375]
10. Weykamp C. HbA1c: a review of analytical and clinical aspects. Ann Lab Med 2013 Nov 01;33(6):393-400 [FREE Full text] [doi: 10.3343/alm.2013.33.6.393] [Medline: 24205486]

11. American Diabetes Association Professional Practice Committee. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2022. Diabetes Care 2022 Jan 01;45(Suppl 1):S17-S38. [doi: 10.2337/dc22-S002] [Medline: 34964875]
12. Mills JL, Simpson JL, Driscoll SG, Jovanovic-Peterson L, Van Allen M, Aarons JH, et al. Incidence of spontaneous abortion among normal women and insulin-dependent diabetic women whose pregnancies were identified within 21 days of conception. N Engl J Med 1988 Dec 22;319(25):1617-1623. [doi: 10.1056/NEJM198812223192501] [Medline: 3200277]
13. Hanson U, Persson B, Thunell S. Relationship between haemoglobin A1C in early type 1 (insulin-dependent) diabetic pregnancy and the occurrence of spontaneous abortion and fetal malformation in Sweden. Diabetologia 1990 Feb 01;33(2):100-104. [doi: 10.1007/BF00401047] [Medline: 2328844]
14. Nielsen GL, Møller M, Sørensen HT. HbA1c in early diabetic pregnancy and pregnancy outcomes: a Danish population-based cohort study of 573 pregnancies in women with type 1 diabetes. Diabetes Care 2006 Dec 01;29(12):2612-2616. [doi: 10.2337/dc06-0914] [Medline: 17130193]
15. Goyal A, Gupta Y, Singla R, Kalra S, Tandon N. American Diabetes Association "Standards of Medical Care-2020 for Gestational Diabetes Mellitus": a critical appraisal. Diabetes Ther 2020 Aug 01;11(8):1639-1644 [FREE Full text] [doi: 10.1007/s13300-020-00865-3] [Medline: 32564336]
16. Jia W, Weng J, Zhu D, Ji L, Lu J, Zhou Z, Chinese Diabetes Society. Standards of medical care for type 2 diabetes in China 2019. Diabetes Metab Res Rev 2019 Sep 01;35(6):e3158-e3184. [doi: 10.1002/dmrr.3158] [Medline: 30908791]
17. Wang D, Zhang Y, Jiang Y, Ye Y, Ji M, Dou Y, SPCC group. Shanghai Preconception Cohort (SPCC) for the association of periconceptional parental key nutritional factors with health outcomes of children with congenital heart disease: a cohort profile. BMJ Open 2019 Nov 24;9(11):e031076-e031085 [FREE Full text] [doi: 10.1136/bmjopen-2019-031076] [Medline: 31767586]
18. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Gynecology. ACOG Practice Bulletin No. 200: Early Pregnancy Loss. Obstet Gynecol 2018 Nov 01;132(5):e197-e207. [doi: 10.1097/AOG.0000000000002899] [Medline: 30157093]
19. Wang Y, Mi J, Shan X, Wang QJ, Ge K. Is China facing an obesity epidemic and the consequences? The trends in obesity and chronic disease in China. Int J Obes (Lond) 2007 Jan 01;31(1):177-188. [doi: 10.1038/sj.ijo.0803354] [Medline: 16652128]
20. Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, International Diabetes Federation Task Force on Epidemiology and Prevention, Hational Heart, Lung, and Blood Institute, American Heart Association, World Heart Federation, International Atherosclerosis Society, International Association for the Study of Obesity. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009 Oct 20;120(16):1640-1645. [doi: 10.1161/CIRCULATIONAHA.109.192644] [Medline: 19805654]
21. Little RR, Rohlfing CL. The long and winding road to optimal HbA1c measurement. Clin Chim Acta 2013 Mar 15;418:63-71 [FREE Full text] [doi: 10.1016/j.cca.2012.12.026] [Medline: 23318564]
22. Greene MF, Hare JW, Cloherty JP, Benacerraf BR, Soeldner JS. First-trimester hemoglobin A1 and risk for major malformation and spontaneous abortion in diabetic pregnancy. Teratology 1989 Mar 01;39(3):225-231. [doi: 10.1002/tera.1420390303] [Medline: 2727930]
23. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2010 Jan 01;33 Suppl 1(Suppl 1):S62-S69 [FREE Full text] [doi: 10.2337/dc10-S062] [Medline: 20042775]
24. Hughes RCE, Rowan J, Florkowski CM. Is there a role for HbA1c in pregnancy? Curr Diab Rep 2016 Jan 06;16(1):5-15. [doi: 10.1007/s11892-015-0698-y] [Medline: 26739347]
25. Chinese Diabetes Society, Dalong Z. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition). Chin J Diabetes Mellitus 2021 Apr 25;13(4):317-332 [FREE Full text] [doi: 10.3760/cma.j.cn311282-20210304-00142]
26. Moley KH, Chi MM, Knudson CM, Korsmeyer SJ, Mueckler MM. Hyperglycemia induces apoptosis in pre-implantation embryos through cell death effector pathways. Nat Med 1998 Dec 24;4(12):1421-1424. [doi: 10.1038/4013] [Medline: 9846581]
27. Saisho Y. β-cell dysfunction: Its critical role in prevention and management of type 2 diabetes. World J Diabetes 2015 Feb 15;6(1):109-124 [FREE Full text] [doi: 10.4239/wjd.v6.i1.109] [Medline: 25685282]
28. Robajac D, Masnikosa R, Miković Ž, Mandić V, Nedić O. Oxidation of placental insulin and insulin-like growth factor receptors in mothers with diabetes mellitus or preeclampsia complicated with intrauterine growth restriction. Free Radic Res 2015 Apr 25;49(8):984-989. [doi: 10.3109/10715762.2015.1020798] [Medline: 25812587]
29. Yu H, Qi X, Wang X. Application of glycated hemoglobin in the perinatal period. Int J Clin Exp Med 2014 Dec 30;7(12):4653-4659 [FREE Full text] [Medline: 25663962]
30. Davidson AJF, Park AL, Berger H, Aoyama K, Harel Z, Cook JL, et al. Risk of severe maternal morbidity or death in relation to elevated hemoglobin A1c preconception, and in early pregnancy: A population-based cohort study. PLoS Med 2020 May 19;17(5):e1003104-e1003122 [FREE Full text] [doi: 10.1371/journal.pmed.1003104] [Medline: 32427997]

31. Kattini R, Hummelen R, Kelly L. Early gestational diabetes mellitus screening with glycated hemoglobin: a systematic review. J Obstet Gynaecol Can 2020 Nov 01;42(11):1379-1384. [doi: 10.1016/j.jogc.2019.12.015] [Medline: 32268994]

Abbreviations

ADA: American Diabetes Association
aRR: adjusted risk ratio
FAS: folic acid supplementation
FBG: fasting blood glucose
HbA_{1c}: serum glycated hemoglobin
Pre-BMI: preconception BMI
RCS: restricted cubic spline
RR: risk ratio
SPCC: Shanghai Preconception Cohort Study
SPL: spontaneous pregnancy loss

Edited by A Mavragani, T Sanchez; submitted 04.03.23; peer-reviewed by C Shen, Z Liang, H Inadera; comments to author 18.09.23; revised version received 27.09.23; accepted 02.11.23; published 12.12.23.

Please cite as:
Chen X, Zhang Y, Chen H, Dou Y, Wang Y, He W, Ma X, Sheng W, Yan W, Huang G
Association Between Serum Glycated Hemoglobin Levels at Early Gestation and the Risk of Subsequent Pregnancy Loss in Pregnant Women Without Diabetes Mellitus: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e46986
URL: https://publichealth.jmir.org/2023/1/e46986
doi:10.2196/46986
PMID:38085559

Original Paper

Manifestations and Outcomes of Intracerebral Hemorrhage During the COVID-19 Pandemic in China: Multicenter, Longitudinal Cohort Study

Yan Wan[1*], PhD; Quan Wei He[1*], PhD; Shaoli Chen[1*], MD; Man Li[1], PhD; Yuanpeng Xia[1], PhD; Lei Zhang[1], MD; Zhou Sun[1], MD; Xiaolu Chen[1], MD; David Wang[2], DO; Jiang Chang[3], PhD; Bo Hu[1], PhD

[1]Department of Neurology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2]Barrow Neurological Institute, St. Joseph's Hospital and Medical Center, Phoenix, AZ, United States
[3]School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[*]these authors contributed equally

Corresponding Author:
Bo Hu, PhD
Department of Neurology, Union Hospital
Tongji Medical College
Huazhong University of Science and Technology
NO.1277, Jiefang Road
Wuhan, 430022
China
Phone: 86 13707114863
Email: hubo@mail.hust.edu.cn

Abstract

Background: The COVID-19 pandemic has inevitably affected the distribution of medical resources, and epidemic lockdowns have had a significant impact on the nursing and treatment of patients with other acute diseases, including intracerebral hemorrhage (ICH).

Objective: This study aimed to investigate how the COVID-19 pandemic affected the manifestations and outcomes of patients with ICH.

Methods: Patients with acute ICH before (December 1, 2018-November 30, 2019) and during (December 1, 2019-November 30, 2020) the COVID-19 pandemic at 31 centers in China from the Chinese Cerebral Hemorrhage: Mechanism and Intervention (CHEERY) study were entered into the analysis. Demographic information and clinical manifestations and outcomes were collected and compared between the 2 groups.

Results: From December 1, 2018, to November 30, 2020, a total of 3460 patients with ICH from the CHEERY study were enrolled and eventually analyzed. The results showed that during the COVID-19 pandemic, patients with ICH were more likely to be older (*P*<.001) with a history of ischemic stroke (*P*=.04), shorter time from onset to admission (*P*<.001), higher blood pressure (*P*<.001), higher fasting blood glucose (*P*=.003), larger hematoma volume (*P*<.001), and more common deep ICH (*P*=.01) and intraventricular hemorrhage (*P*=.02). These patients required more intensive care unit treatment (*P*<.001) and preferred to go to the hospital directly rather than call an ambulance (*P*<.001). In addition, the COVID-19 pandemic was associated with an increased risk of pulmonary infection during hospitalization (adjusted risk ratio [$RR_{adjusted}$] 1.267, 95% CI 1.065-1.509), lower probability of a 3-month good outcome ($RR_{adjusted}$ 0.975, 95% CI 0.956-0.995), and a higher probability of in-hospital ($RR_{adjusted}$ 3.103, 95% CI 2.156-4.465), 1-month ($RR_{adjusted}$ 1.064, 95% CI 1.042-1.087), and 3-month ($RR_{adjusted}$ 1.069, 95% CI 1.045-1.093) mortality.

Conclusions: Our study indicated that the cloud of COVID-19 has adversely impacted the presentation and outcomes of ICH. Medical workers may pay more attention to patients with ICH, while the public should pay more attention to hypertension control and ICH prevention.

Trial Registration: Chinese Clinical Trial Registry ChiCTR1900020872; https://www.chictr.org.cn/showprojEN.html?proj=33817

(JMIR Public Health Surveill 2023;9:e34386) doi:10.2196/34386

KEYWORDS

COVID-19; intracerebral hemorrhage; manifestation; outcome; cohort study

Introduction

COVID-19, the pneumonia caused by SARS-CoV-2 infection, emerged in Wuhan City, Hubei Province, China in December 2019 [1]. The World Health Organization (WHO) declared the COVID-19 outbreak a public health emergency of international concern on January 30, 2020, and declared a pandemic on March 11, 2020 [2]. In China, from January 3, 2020, to March 16, 2023, there have been 99,185,059 confirmed cases of COVID-19 with 120,576 deaths reported to the WHO [3]. Wuhan was the core area of the COVID-19 storm in China, and this city was locked down from January 23, 2020, to April 8, 2020 [4]. Various public health interventions were implemented by the Chinese government to suppress the big pandemic, including but not limited to intensive intracity and intercity traffic restrictions, social distancing measures, home isolation and centralized quarantine, and improvement of medical resources [5,6].

In addition, as a global health crisis, the COVID-19 pandemic has been the most serious challenge to delivering timely care to patients with other conditions [7]. Its outbreak proved to have a major impact on patients with cerebrovascular diseases [8,9], for example, significantly decreasing the rate of stroke admissions [10,11], intravenous thrombolysis and mechanical thrombectomy [12,13], and prolonging the time from symptom onset to treatment of ischemic stroke [14]. Intracerebral hemorrhage (ICH), one of the most devastating diseases worldwide with a mortality rate of up to twice that of ischemic stroke [15], requires emergency medical care in specialized neurological intensive care units [16]. Under the condition of tight medical resources, it was difficult to balance stroke treatment and COVID-19 prevention during the COVID-19 pandemic [8-10,12], and it is not clear whether the COVID-19 pandemic affected the manifestations and outcomes of patients with ICH.

Moreover, hypertension management during the COVID-19 pandemic may be a challenge, which is an important therapy for ICH [17,18]. There is an unproven notion that angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin receptor blockers (ARBs) might theoretically increase the risk of COVID-19 infection because increased angiotensin-converting enzyme-2 activity has been shown to lead to practical medicine restriction. Recently, there have been reports that discontinuation of ACEI/ARB in patients with COVID-19 is not related to the severity of COVID-19 [19,20]. However, discontinuation of ACEI/ARB was once not an uncommon phenomenon. Besides, during the COVID-19 pandemic, the public may have also suffered from social pressure, anxiety, the depressed economy, a lack of public health resources, and so on [21,22]. All of these factors may disturb the control of risk factors of ICH, resulting in increased occurrence and poor outcomes.

The Chinese Cerebral Hemorrhage: Mechanism and Intervention (CHEERY) study is a prospective multicenter cohort study that began on December 1, 2018, and recruits patients with spontaneous ICH within 7 days of onset in China [18]. This cohort covered periods before and after the COVID-19 outbreak. In this study, in order to investigate how the COVID-19 pandemic affected the manifestations and outcomes of patients with ICH, we retrospectively analyzed the data of patients with ICH in 1-year periods before (December 1, 2018-November 30, 2019) and during (December 1, 2019-November 30, 2020) the COVID-19 pandemic in the CHEERY study.

Methods

Study Design

The CHEERY study is a multicenter Chinese program that includes consecutive patients with ICH admitted to 31 stroke centers beginning in December 2018. It recruits patients with spontaneous ICH within 7 days of onset and aims to identify any major risk factors affecting the prognosis of ICH [18]. To explore the effects of the COVID-19 pandemic on the manifestations and outcomes of patients with ICH, we selected and analyzed data from December 1, 2018, to November 30, 2020, which happens to span 1 year before (December 1, 2018-November 30, 2019) and during (December 1, 2019-November 30, 2020) the epidemic. We followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guidelines.

Participants

All patients with ICH enrolled in the study must meet the following criteria: (1) older than 18 years; (2) have acute ICH with ≤7 days duration from onset to admission; and (3) have received either computerized tomography or magnetic resonance imaging of the brain to confirm the diagnosis of ICH. Due to the centralized isolation of patients with COVID-19, this study did not include patients with ICH complicated with COVID-19 infection.

Data Collection and Analysis

Demographic information was collected, including age, gender, residence, smoking and alcohol history, and medical history from the medical record. Gender was recorded from the medical record, which was copied from the patient’s resident ID card. The mode of transportation to the hospital and duration from onset to admission were collected shortly after admission from patients or escorts. Systolic blood pressure, diastolic blood pressure, hematoma location, hematoma volume, and laboratory data were copied from the medical record. The baseline National Institute of Health Stroke Scale and Glasgow coma scale were assessed by 2 trained neurologists. After admission, the first head computerized tomography scan was acquired, and 2 imageologists calculated the volume of bleeding with the ABC/2 formula and recorded the location of bleeding and intraventricular hemorrhage (IVH). The SMASH-U (structural vascular lesions, medication, amyloid angiopathy, systemic disease, hypertension, and undetermined) flowchart was adopted to classify the most likely cause for each ICH event [23].

Structural vascular lesions were defined as vascular structural abnormalities at the bleeding sites verified by imaging or pathological findings. Medication was defined as warfarin use with an international normalized ratio ≥2, novel oral anticoagulant use within 3 days, full-dose heparin, or thrombolytic agent use. Amyloid angiopathy was defined as lobar, cortical, or subcortical hemorrhage among patients aged ≥55 years, according to the Boston criteria. Systemic disease was conventionally defined by the presence of thrombocytopenia or liver cirrhosis or alternatively by the presence of non–drug-induced coagulopathy or renal failure. Hypertension was defined as deep or infratentorial hemorrhage with pre-existing hypertension history. All other causes and unknown causes were considered undetermined. Pulmonary infection was diagnosed by 2 well-trained and experienced neurologists according to the modified Centers for Disease Control and Prevention criteria in combination with the patient's clinical symptoms, laboratory results, and radiological examinations within 1 week after ICH. A modified ranking scale score was assessed centrally by face-to-face, telephone, or WeChat interview.

Definitions

ICH was defined as hemorrhage into the cerebral parenchyma that may also extend into ventricles, and rarely, subarachnoid spaces. A modified ranking scale score of 0-3 points was defined as a good outcome.

Statistical Analysis

Data were presented as the mean (SD), number (%), and median (IQR). The Kolmogorov-Smirnov test for normality was used to assess the data distribution of continuous variables. Continuous variables were tested with a *t* test or Mann-Whitney U test. Categorical variables were tested with the χ^2 test or Fisher exact test. The associations of the outcomes of interest between the 2 groups were analyzed using the modified Poisson regression model, and risk ratios (RRs) were reported along with 95% CIs. Age and gender were used as covariables. SPSS statistical software (version 26.0, SPSS Corp) was used to analyze the data, and the statistical significance was set at $P<.05$.

Ethical Considerations

The study protocol and data collection were conducted strictly following the Declaration of Helsinki and approved by the Research Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (approval no. 2018-S485) and registered in the Chinese Clinical Trial Registry (ChiCTR1900020872) [24]. Protocol training was conducted for all researchers prior to the implementation of the project. All participants signed a written informed consent form prior to enrollment. The informed consent should be signed by the participant, but for those participants who could not sign the informed consent form by themselves for any reason, their parents, legal guardians, or protectors signed the informed consent form.

Results

From December 1, 2018, to November 30, 2020, there were 3460 patients with acute ICH enrolled in the CHEERY study, and 3260 patients were included in the final analysis (Figure 1). The demographic data are listed in Table 1.

Figure 1. Flowchart of patient enrollment in the Chinese Cerebral Hemorrhage: Mechanism and Intervention study. ICH: intracerebral hemorrhage; mRS: modified ranking scale; SMASH-U: structural vascular lesions, medication, amyloid angiopathy, systemic disease, hypertension, and undetermined.

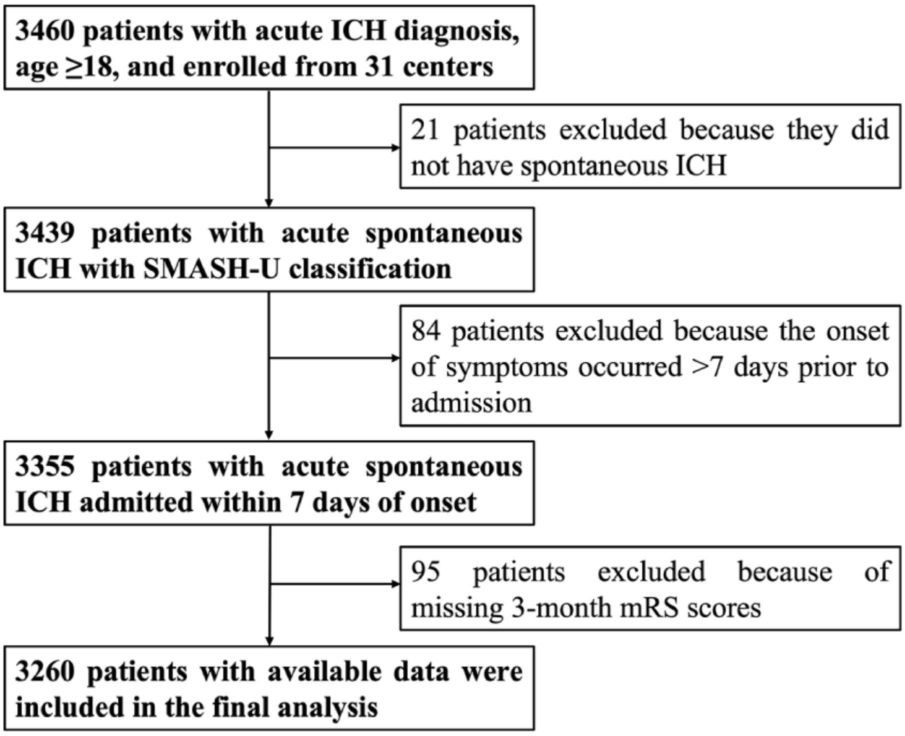


Table 1. Baseline and admission characteristics of patients admitted to hospitals in the year during COVID-19 pandemic (December 1, 2019-November 30, 2020) and the year before the COVID-19 outbreak (December 1, 2018-Novemver 30, 2019).

Characteristic	Time period		*P* value
	During the COVID-19 pandemic (n=1510)	Before the COVID-19 pandemic (n=1750)	
Age (years), median (IQR)	63 (54-71)	61 (52-70)	<.001
Age (years), n (%)			.001
<60	625 (41.4)	828 (47.3)	
≥60	885 (58.6)	922 (52.7)	
Gender, n (%)			.009
Men	1045 (69.2)	1136 (64.9)	
Women	465 (30.8)	614 (35.1)	
Residence, n (%)			
Rural	850 (56.3)	930 (53.1)	.18
Urban	639 (42.3)	797 (45.5)	
Unknown	21 (1.4)	23 (1.3)	
Smoking history, n (%)			.001
Yes	424 (28.1)	584 (33.4)	
No	1081 (71.6)	1159 (66.2)	
Unknown	5 (0)	7 (0)	
Alcohol history, n (%)			.02
Yes	346 (22.9)	464 (26.5)	
No	1159 (76.8)	1279 (73.1)	
Unknown	5 (0)	7 (0)	
Comorbidities, n (%)			
Ischemic heart disease	62 (4.1)	53 (3)	.25
Ischemic stroke	156 (10.4)	150 (8.6)	.04
Hypertension	970 (64.7)	1141 (66.1)	.40
Diabetes	137 (9.1)	143 (8.2)	.06
Antihypertension[a]	281/970 (29)	299/1141 (26.2)	.16
Pre-ACEI/ARB[b] usage	49/281 (17.4)	45/299 (15.1)	.44
Transportation to hospital, n (%)			<.001
Ambulance	660 (43.7)	1012 (57.8)	
Self-admission	838 (55.5)	686 (39.2)	
Unknown	12 (0.8)	52 (3)	
Duration from onset to admission (hours), median (IQR)	4 (2-12)	6 (2-24)	<.001
SBP[c] on admission (mm Hg), median (IQR)	170 (150-190)	162 (145-180)	<.001
DBP[d] on admission (mm Hg), median (IQR)	97 (87-108)	94 (84-106)	<.001
Fasting blood glucose (mmol/L), median (IQR)	6.3 (5.3-7.9)	6.2 (5.3-7.4)	.003
NIHSS[e] on admission, median (IQR)	9 (4-16)	8 (3-15)	.05

Characteristic	Time period		*P* value
	During the COVID-19 pandemic (n=1510)	Before the COVID-19 pandemic (n=1750)	
GCS[f] on admission, median (IQR)	14 (10-15)	14 (11-15)	.009
Location of hematoma, n (%)			.01
Nondeep	292 (19.3)	403 (23)	
Deep	1218 (80.7)	1347 (77)	
Hematoma volume (ml), median (IQR)	12.0 (5.0-28.0)	10.7 (4.9-22.0)	<.001
IVH[g], n (%)	249 (16.5)	238 (13.6)	.02
ICU[h] treatment, n (%)	960 (63.7)	938 (55.2)	<.001
Surgical intervention, n (%)	271 (17.9)	358 (20.5)	.07
SMASH-U[i] classification, n (%)			
Structural lesion	26 (1.9)	31 (1.8)	.83
Medication	7 (0.5)	4 (0.2)	.19
Amyloid angiopathy	196 (14.7)	235 (13.9)	.58
Systemic diseases	16 (1.2)	18 (1.1)	.74
Hypertension	725 (54.2)	919 (54.5)	.86
Undetermined	361 (27)	478 (28.4)	.41
Hospitalization length (days), median (IQR)	15 (8-21)	15 (10-21)	.02

[a] Percentage of pre-ICH antihypertension treatments in patients with hypertension.

[b] ACEI/ARB: angiotensin-converting enzyme inhibitor/angiotensin receptor blocker.

[c] SBP: systolic blood pressure.

[d] DBP: diastolic blood pressure.

[e] NIHSS: National Institute of Health Stroke Scale.

[f] GCS: Glasgow coma scale.

[g] IVH: intraventricular hemorrhage.

[h] ICU: intensive care unit.

[i] SMASH-U: structural vascular lesions, medication, amyloid angiopathy, systemic disease, hypertension, and undetermined.

The patients’ median age was 63 (IQR 54-71) years during the COVID-19 pandemic and 61 (IQR 52-70) years before the COVID-19 breakout (*P*<.001). The COVID-19 pandemic was associated with a higher percentage of male patients (*P*=.009) and patients with a history of ischemic stroke (*P*=.04). Fewer patients had a history of smoking (*P*=.001) and drinking (*P*=.02) during the pandemic (Table 1).

During the COVID-19 pandemic, patients were more likely to go to the hospital directly by themselves (during: n=838, 55.5%; before: n=686, 39.2%; *P*<.001) and took less time from onset to admission (during: median 4, IQR 2-12 hours; before: median 6, IQR 2-24 hours; *P*<.001) compared with before the pandemic. The COVID-19 pandemic was also associated with an increase in systolic blood pressure (during: median 170, IQR 150-190 mm HG; before: median 162, IQR 145-180 mm Hg; *P*<.001), diastolic blood pressure (during: median 97, IQR 87-108 mm HG; before: median 94, IQR 84-106 mm Hg; *P*<.001), fasting blood glucose (during: median 6.3, IQR 5.3-7.9 mmol/L; before: median 6.2, IQR 5.3-7.4 mmol/L; *P*=.003), deep ICH (during: n=1218, 80.7%; before: n=1347, 77%; *P*=.01), hematoma volume (during: median 12.0, IQR 5.0-28.0 mL; before: median 10.7, IQR 4.9-22.0 mL; *P*<.001), intraventricular hemorrhage (during: n=249, 16.5%; before: n=238, 13.6%; *P*=.02), and ICU treatment (during: n=960, 63.7%; before: n=938, 55.2%; *P*<.001) (Table 1).

We then comparatively evaluated the outcomes of ICH. Table 2 shows that the COVID-19 pandemic was associated with an increased risk of pulmonary infection during hospitalization ($RR_{unadjusted}$ 1.299, 95% CI 1.091-1.546; $RR_{adjusted}$ 1.267, 95% CI 1.065-1.509), lower probability of 3-month good outcome ($RR_{unadjusted}$ 0.970, 95% CI 0.950-0.989; $RR_{adjusted}$ 0.975, 95% CI 0.956-0.995), and higher in-hospital mortality ($RR_{unadjusted}$ 3.233, 95% CI 2.246-4.653; $RR_{adjusted}$ 3.103, 95% CI 2.156-4.465), 1-month mortality ($RR_{unadjusted}$ 1.070, 95% CI 1.048-1.093; $RR_{adjusted}$ 1.064, 95% CI 1.042-1.087), and 3-month mortality ($RR_{unadjusted}$ 1.075, 95% CI 1.051-1.099; $RR_{adjusted}$ 1.069, 95% CI 1.045-1.093).

Table 2. Outcomes in patients admitted to hospitals during (December 1, 2019-November 30, 2020) and before (December 1, 2018-Novemver 30, 2019) the COVID-19 pandemic.

Characteristics	Time period		Unadjusted analysis		Adjusted analysis[a]	
	During the COVID-19 pandemic (n=1510), n (%)	Before the COVID-19 pandemic (n=1750), n (%)	Risk ratio (95% CI)	*P* value	Risk ratio (95% CI)	*P* value
Pulmonary infection	232 (15.4)	207 (11.8)	1.299 (1.091-1.546)	.003	1.267 (1.065-1.509)	.008
1-month mRS[b] 0-3	797 (52.8)	977 (55.8)	0.980 (0.959-1.002)	.08	0.985 (0.963-1.007)	.17
3-month mRS 0-3	923 (61.1)	1158 (66.2)	0.970 (0.950-0.989)	.003	0.975 (0.956-0.995)	.01
In-hospital mortality	106 (7)	38 (2.2)	3.233 (2.246-4.653)	<.001	3.103 (2.156-4.465)	<.001
1-month mortality	282 (18.7)	190 (10.9)	1.070 (1.048-1.093)	<.001	1.064 (1.042-1.087)	<.001
3-month mortality	348 (23)	253 (14.5)	1.075 (1.051-1.099)	<.001	1.069 (1.045-1.093)	<.001

[a] Adjusted for age and gender.

[b] mRS: modified ranking scale.

Discussion

Since December 2019, the outbreak of COVID-19 infections has had a major impact on the occurrence, development, and treatment of cerebrovascular diseases [7,8]. Currently, despite the international open policy, the epidemic still occasionally erupts in small areas. In this study, in order to investigate how the COVID-19 pandemic affected the manifestations and outcomes of patients with ICH, we retrospectively analyzed the data of patients with ICH in 1-year periods before (December 1, 2018-November 30, 2019) and during (December 1, 2019-November 30, 2020) the COVID-19 pandemic from the CHEERY study, a prospective multicenter cohort study. Our study found that during the COVID-19 pandemic, patients with ICH were more likely to be older with a history of ischemic stroke, higher blood pressure and fasting blood glucose, larger hematoma volume, shorter time from onset to admission, more common deep ICH and IVH, more ICU treatment, and preferred to go to hospital directly rather than call an ambulance. In addition, during the COVID-19 pandemic, patients with ICH had a higher risk of pulmonary infection, lower probability of 3-month good outcomes, and higher in-hospital, 1-month, and 3-month mortality.

ICH is among the diseases with the highest mortality; in low-income countries, its 30-day mortality may be as high as 40% [25]. Early identification and timely treatment are crucial to reducing patient mortality [26]. Research has found that more than one-third of patients have hematoma enlargement within 24 hours of onset, and approximately 38%-76% experience hematoma enlargement within 3 hours of onset. Every 10% increase in hematoma volume increases the risk of mortality by 5%, and every 1 mL increase in hematoma volume increases the risk of poor prognosis by 7% [27,28]. On the other hand, brain edema that occurs within hours of a patient's onset and persists for days to weeks is also closely related to patient outcomes [29,30], with an average increase of 75% in perihematomal edema within the first 24 hours after ICH, which is also closely related to patient mortality [31,32]. In addition, persistent high systolic blood pressure in the early stages of ICH can lead to an increased risk of hematoma expansion, neurological deterioration, and even brain hernia formation, leading to patient death [33]. Therefore, early blood pressure reduction and intensive care for patients with ICH are crucial to reducing the mortality rate [34,35]. Unfortunately, such actions may have been difficult to implement during the pandemic.

During the pandemic, many studies reported a significant increase in the number of deaths from diseases other than COVID-19 [36,37]. Researchers found that in the early stages of the pandemic (March-July 2020), there was a marked increase in the number of patients dying from cerebrovascular disease [38]; this may have resulted from public health interventions [39] and limited medical care [40,41], patients refusing to enter the hospital for fear of infection [42,43], a decline in hospitals' emergency care capacity [44,45], and insufficient poststroke intensive care [46] and rehabilitation care [10,47].

The global burden of ICH is related to the inadequate management of chronic hypertension and other modifiable risk factors [48]. In China, nearly half of adults aged 35-75 years have hypertension, but only 30.1% of patients with hypertension are being treated, and only about 7.2% are under control [49]. During the COVID-19 pandemic, the situation was worse, and the higher blood pressure in patients with ICH could be attributed to social pressure, anxiety, the depressed economy, the lack of public health resources, inadequate control of risk factors, and unwillingness to seek medical treatment during this special period [40-47]. Furthermore, the use of ACEI/ARBs might have theoretically increased the risk of COVID-19 infection, though recent reports did not find any correlation between the discontinuation of ACEI/ARB in patients with COVID-19 and the severity of COVID-19 [19,20]. For the fear of infection of COVID-19, patients may have discontinued the use of ACEI/ARB. A previous study in China also showed a significant increase in excess deaths due to cardiovascular disease and diabetes in the early lockdown period [43], presumably due to poor hypertension control during lockdowns in populations due to COVID-19 infections conflicting with antihypertensive drugs, such as ACEI.

Our previous studies have found that age, deep ICH, fasting blood glucose, and hematoma volume were predicters of in-hospital neurological deterioration and 3-month poor outcomes in the SIGNALS and ADVISING scores [50,51]. In this study, we found these factors were also associated with poor outcomes in patients with ICH during the COVID-19 pandemic. Older people had more risk factors, such as hypertension, diabetes mellitus, ischemic stroke, ischemic heart disease, and other diseases, which could induce and aggravate ICH. COVID-19 may have restricted the control of risk factors of ICH, resulting in more older patients with ICH.

Our study found that more patients with ICH went to hospitals directly without calling an ambulance. We speculate that this phenomenon was related to limited public medical resources and anxiety.

Most important of all, our study found that patients with ICH demonstrated higher NIHSS scores on admission and larger hemorrhage volumes. They also needed more ICU intervention and had poorer outcomes and mortality at 3 months. This finding indicated that ICH was more severe during the COVID-19 pandemic; this could be related to the patients' higher blood pressure on admission, which may lead to a larger hematoma and poor prognosis. Our results were consistent with the report that older age and larger ICH volume are poor predictors of mortality [52].

One main limitation of our study was that we only included hospitalized patients. Those who were treated in an outpatient setting and died before reaching the hospital were not included. During the COVID-19 pandemic period corresponding to this study (December 1, 2019-November 30, 2020), more than 86,000 Chinese people were infected with COVID-19, as reported to the WHO. Vaccines, treatments, and minimally intrusive suppression measures are being developed for long-term prevention and control of COVID-19 in China [53]. As per government policy, infected patients were transferred to designated hospitals, which were not included in our study centers. Therefore, this study did not contain patients with COVID-19.

A major strength of this study is the use of the CHEERY study and its consecutive enrollment of patients within a defined study time, including the year before and after the COVID-19 outbreak. This excludes relevant selection bias and ensures that results from this cohort are fairly representative.

Our study indicated that the cloud of COVID-19 has adversely impacted the presentation and outcomes of ICH. Medical workers should pay more attention to patients with ICH, while the public should pay more attention to hypertension control and ICH prevention.

Acknowledgments

The authors would like to express their gratitude to all the participants and investigators in the 31 centers of the CHEERY study. This work was supported by the National Natural Science Foundation of China (grants 81901214 to YW, 82071335 to QWH, and 81820108010 to BH) and the National Key Research and Development Program of China (grant 2018YFC1312200 to BH).

Data Availability

Some of the data of this study are provided in Multimedia Appendix 1. Additional data are available upon reasonable request after approval by BH.

Authors' Contributions

BH was responsible for the concept and design of the study. YW and QWH did the literature search and wrote the manuscript. ML, YPX, LZ, ZS, and XLC acquired and interpreted the data. SLC and JC analyzed the data. DW and JC revised the manuscript. JC and YW provided administrative, technical, or statistical support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data set.
[XLSX File (Microsoft Excel File), 475 KB - publichealth_v9i1e34386_app1.xlsx]

References

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet 2020 Feb;395(10223):497-506. [doi: 10.1016/s0140-6736(20)30183-5] [Medline: 31986264]
2. Lau J, Yu Y, Xin M, She R, Luo S, Li L, et al. Adoption of preventive measures during the very early phase of the COVID-19 outbreak in China: national cross-sectional survey study. JMIR Public Health Surveill 2021 Oct 07;7(10):e26840 [FREE Full text] [doi: 10.2196/26840] [Medline: 34479184]
3. China. World Health Organization. URL: https://covid19.who.int/region/wpro/country/cn [accessed 2023-10-24]
4. Ke Y, Cui J, Wong Y. Ecological study on differences in COVID-19 fatality among Wuhan, rest of Hubei, and rest of China. J Epidemiol Glob Health 2021 Mar;11(1):42-45 [FREE Full text] [doi: 10.2991/jegh.k.200902.001] [Medline: 33009735]

5. Tang J, Abbasi K. What can the world learn from China's response to covid-19? BMJ 2021 Dec 01;375:n2806 [FREE Full text] [doi: 10.1136/bmj.n2806] [Medline: 34853017]
6. Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA 2020 May 19;323(19):1915-1923 [FREE Full text] [doi: 10.1001/jama.2020.6130] [Medline: 32275295]
7. Haldane V, De Foo C, Abdalla SM, Jung A, Tan M, Wu S, et al. Health systems resilience in managing the COVID-19 pandemic: lessons from 28 countries. Nat Med 2021 Jun;27(6):964-980. [doi: 10.1038/s41591-021-01381-y] [Medline: 34002090]
8. Tu W, Xu Y, Chen H, Li J, Du J. Impact of the COVID-19 pandemic lockdown on hospitalizations for cerebrovascular disease and related in-hospital mortality in China: a nationwide observational study. Arch Gerontol Geriatr 2023 Jan;104:104832 [FREE Full text] [doi: 10.1016/j.archger.2022.104832] [Medline: 36219895]
9. Kansagra AP, Goyal MS, Hamilton S, Albers GW. Collateral effect of Covid-19 on stroke evaluation in the United States. N Engl J Med 2020 Jul 23;383(4):400-401. [doi: 10.1056/nejmc2014816] [Medline: 32383831]
10. Rinkel LA, Prick JCM, Slot RER, Sombroek NMA, Burggraaff J, Groot AE, et al. Impact of the COVID-19 outbreak on acute stroke care. J Neurol 2021 Feb;268(2):403-408 [FREE Full text] [doi: 10.1007/s00415-020-10069-1] [Medline: 32691235]
11. Mariet A, Giroud M, Benzenine E, Cottenet J, Roussot A, Aho-Glélé LS, et al. Hospitalizations for stroke in France during the COVID-19 pandemic before, during, and after the national lockdown. Stroke 2021 Apr;52(4):1362-1369. [doi: 10.1161/strokeaha.120.032312] [Medline: 33626900]
12. Zhao J, Li H, Kung D, Fisher M, Shen Y, Liu R. Impact of the COVID-19 epidemic on stroke care and potential solutions. Stroke 2020 Jul;51(7):1996-2001. [doi: 10.1161/strokeaha.120.030225] [Medline: 32432997]
13. Zhou Y, Hong C, Chang J, Xia Y, Jin H, Li Y, et al. Intravenous thrombolysis for acute ischaemic stroke during COVID-19 pandemic in Wuhan, China: a multicentre, retrospective cohort study. J Neurol Neurosurg Psychiatry 2021 Feb;92(2):226-228. [doi: 10.1136/jnnp-2020-324014] [Medline: 33115934]
14. Tanaka K, Matsumoto S, Nakazawa Y, Yamada T, Sonoda K, Nagano S, et al. Delays in presentation time under the COVID-19 epidemic in patients with transient ischemic attack and mild stroke: a retrospective study of three hospitals in a Japanese prefecture. Front Neurol 2021;12:748316 [FREE Full text] [doi: 10.3389/fneur.2021.748316] [Medline: 34777220]
15. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
16. Hemphill JC, Greenberg SM, Anderson CS, Becker K, Bendok BR, Cushman M, et al. Guidelines for the management of spontaneous intracerebral hemorrhage. Stroke 2015 Jul;46(7):2032-2060. [doi: 10.1161/str.0000000000000069] [Medline: 26022637]
17. Qureshi AI, Palesch YY, Barsan WG, Hanley DF, Hsu CY, Martin RL, et al. Intensive blood-pressure lowering in patients with acute cerebral hemorrhage. N Engl J Med 2016 Sep 15;375(11):1033-1043. [doi: 10.1056/nejmoa1603460] [Medline: 27276234]
18. Wan Y, Guo H, Shen J, Chen S, Li M, Xia Y, et al. Association between preonset anti-hypertensive treatment and intracerebral hemorrhage mortality: a cohort study from CHEERY. Front Neurol 2022;13:794080 [FREE Full text] [doi: 10.3389/fneur.2022.794080] [Medline: 35321510]
19. Bauer A, Schreinlechner M, Sappler N, Dolejsi T, Tilg H, Aulinger BA, et al. Discontinuation versus continuation of renin-angiotensin-system inhibitors in COVID-19 (ACEI-COVID): a prospective, parallel group, randomised, controlled, open-label trial. Lancet Respir Med 2021 Aug;9(8):863-872 [FREE Full text] [doi: 10.1016/S2213-2600(21)00214-9] [Medline: 34126053]
20. Savarese G, Benson L, Sundström J, Lund LH. Association between renin-angiotensin-aldosterone system inhibitor use and COVID-19 hospitalization and death: a 1.4 million patient nationwide registry analysis. Eur J Heart Fail 2021 Mar;23(3):476-485 [FREE Full text] [doi: 10.1002/ejhf.2060] [Medline: 33222412]
21. Zhou M, Guo W. Social factors and worry associated with COVID-19: evidence from a large survey in China. Soc Sci Med 2021 May;277:113934 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113934] [Medline: 33878665]
22. Han L, Zhan Y, Li W, Xu Y, Xu Y, Zhao J. Associations between the perceived severity of the COVID-19 pandemic, cyberchondria, depression, anxiety, stress, and lockdown experience: cross-sectional survey study. JMIR Public Health Surveill 2021 Sep 16;7(9):e31052 [FREE Full text] [doi: 10.2196/31052] [Medline: 34478402]
23. Mosconi MG, Paciaroni M, Agnelli G, Marzano M, Alberti A, Venti M, et al. SMASH-U classification: a tool for aetiology-oriented management of patients with acute haemorrhagic stroke. Intern Emerg Med 2021 Jan;16(1):109-114. [doi: 10.1007/s11739-020-02330-2] [Medline: 32266689]
24. Chinese Clinical Trial Registry. URL: http://www.chictr.org.cn [accessed 2023-10-24]
25. van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. The Lancet Neurology 2010 Feb;9(2):167-176. [doi: 10.1016/s1474-4422(09)70340-0] [Medline: 20056489]

26. Lord AS, Gilmore E, Choi HA, Mayer SA. Time course and predictors of neurological deterioration after intracerebral hemorrhage. Stroke 2015 Mar;46(3):647-652. [doi: 10.1161/strokeaha.114.007704] [Medline: 25657190]
27. Chang GY. Hematoma growth is a determinant of mortality and poor outcome after intracerebral hemorrhage. Neurology 2007 Feb 06;68(6):471. [doi: 10.1212/01.wnl.0000256290.15120.37] [Medline: 17283327]
28. Wang W, Lu J, Liu L, Jia J, Zhao X. Ultraearly hematoma growth in acute spontaneous intracerebral hemorrhage predicts early and long-term poor clinical outcomes: a prospective, observational cohort study. Front Neurol 2021;12:747551 [FREE Full text] [doi: 10.3389/fneur.2021.747551] [Medline: 34975715]
29. Jiang C, Guo H, Zhang Z, Wang Y, Liu S, Lai J, et al. Molecular, pathological, clinical, and therapeutic aspects of perihematomal edema in different stages of intracerebral hemorrhage. Oxid Med Cell Longev 2022;2022:3948921 [FREE Full text] [doi: 10.1155/2022/3948921] [Medline: 36164392]
30. Lee KH, Lioutas V, Marchina S, Selim M, iDEF Investigators. The prognostic roles of perihematomal edema and ventricular size in patients with intracerebral hemorrhage. Neurocrit Care 2022 Oct;37(2):455-462. [doi: 10.1007/s12028-022-01532-0] [Medline: 35676589]
31. Lv X, Li Z, Deng L, Yang WS, Li YL, Huang YJ, et al. Early perihematomal edema expansion: definition, significance, and association with outcomes after intracerebral hemorrhage. Oxid Med Cell Longev 2021;2021:6249509 [FREE Full text] [doi: 10.1155/2021/6249509] [Medline: 34552686]
32. Grunwald Z, Beslow LA, Urday S, Vashkevich A, Ayres A, Greenberg SM, et al. Perihematomal edema expansion rates and patient outcomes in deep and lobar intracerebral hemorrhage. Neurocrit Care 2017 Apr;26(2):205-212 [FREE Full text] [doi: 10.1007/s12028-016-0321-3] [Medline: 27844466]
33. Mustanoja S, Putaala J, Koivunen R, Surakka I, Tatlisumak T. Blood pressure levels in the acute phase after intracerebral hemorrhage are associated with mortality in young adults. Eur J Neurol 2018 Aug;25(8):1034-1040. [doi: 10.1111/ene.13662] [Medline: 29683540]
34. Li Q, Warren AD, Qureshi AI, Morotti A, Falcone GJ, Sheth KN, et al. Ultra-early blood pressure reduction attenuates hematoma growth and improves outcome in intracerebral hemorrhage. Ann Neurol 2020 Aug;88(2):388-395 [FREE Full text] [doi: 10.1002/ana.25793] [Medline: 32453453]
35. Qureshi AI, Huang W, Lobanova I, Barsan WG, Hanley DF, Hsu CY, et al. Outcomes of intensive systolic blood pressure reduction in patients with intracerebral hemorrhage and excessively high initial systolic blood pressure: post hoc analysis of a randomized clinical trial. JAMA Neurol 2020 Nov 01;77(11):1355-1365 [FREE Full text] [doi: 10.1001/jamaneurol.2020.3075] [Medline: 32897310]
36. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet 2022 Apr 16;399(10334):1513-1536 [FREE Full text] [doi: 10.1016/S0140-6736(21)02796-3] [Medline: 35279232]
37. Islam N, Shkolnikov VM, Acosta RJ, Klimkin I, Kawachi I, Irizarry RA, et al. Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. BMJ 2021 May 19;373:n1137 [FREE Full text] [doi: 10.1136/bmj.n1137] [Medline: 34011491]
38. Lee W, Woo Park S, Weinberger D, Olson D, Simonsen L, Grenfell BT, et al. Direct and indirect mortality impacts of the COVID-19 pandemic in the United States, March 1, 2020 to January 1, 2022. Elife 2023 Feb 22;12:e77562 [FREE Full text] [doi: 10.7554/eLife.77562] [Medline: 36811598]
39. Teo K, Leung WC, Wong Y, Liu RK, Chan AH, Choi OM, et al. Delays in stroke onset to hospital arrival time during COVID-19. Stroke 2020 Jul;51(7):2228-2231. [doi: 10.1161/strokeaha.120.030105] [Medline: 32432998]
40. Onteddu SR, Nalleballe K, Sharma R, Brown AT. Underutilization of health care for strokes during the COVID-19 outbreak. Int J Stroke 2020 Jul;15(5):NP9-NP10. [doi: 10.1177/1747493020934362] [Medline: 32478606]
41. Sheng S, Wang X, Gil Tommee C, Arulprakash N, Kamran M, Shah V, et al. Continued underutilization of stroke care during the COVID-19 pandemic. Brain Behav Immun Health 2021 Aug;15:100274 [FREE Full text] [doi: 10.1016/j.bbih.2021.100274] [Medline: 34589777]
42. Lange S, Ritchey M, Goodman A, Dias T, Twentyman E, Fuld J, et al. Potential indirect effects of the COVID-19 pandemic on use of emergency departments for acute life-threatening conditions - United States, January-May 2020. Am J Transplant 2020 Sep;20(9):2612-2617 [FREE Full text] [doi: 10.1111/ajt.16239] [Medline: 32862556]
43. Liu J, Zhang L, Yan Y, Zhou Y, Yin P, Qi J, et al. Excess mortality in Wuhan city and other parts of China during the three months of the covid-19 outbreak: findings from nationwide mortality registries. BMJ 2021 Feb 24;372:n415 [FREE Full text] [doi: 10.1136/bmj.n415] [Medline: 33627311]
44. Laukkanen L, Lahtinen S, Liisanantti J, Kaakinen T, Ehrola A, Raatiniemi L. Early impact of the COVID-19 pandemic and social restrictions on ambulance missions. Eur J Public Health 2021 Oct 26;31(5):1090-1095 [FREE Full text] [doi: 10.1093/eurpub/ckab065] [Medline: 33856015]
45. Fitzpatrick D, Duncan EAS, Moore M, Best C, Andreis F, Esposito M, et al. Epidemiology of emergency ambulance service calls related to COVID-19 in Scotland: a national record linkage study. Scand J Trauma Resusc Emerg Med 2022 Jan 28;30(1):9 [FREE Full text] [doi: 10.1186/s13049-022-00995-6] [Medline: 35090527]

46. Wahlster S, Sharma M, Lewis A, Patel PV, Hartog CS, Jannotta G, et al. The coronavirus disease 2019 pandemic's effect on critical care resources and health-care providers: a global survey. Chest 2021 Feb;159(2):619-633 [FREE Full text] [doi: 10.1016/j.chest.2020.09.070] [Medline: 32926870]
47. Pop R, Quenardelle V, Hasiu A, Mihoc D, Sellal F, Dugay MH, et al. Impact of the COVID-19 outbreak on acute stroke pathways - insights from the Alsace region in France. Eur J Neurol 2020 Sep;27(9):1783-1787 [FREE Full text] [doi: 10.1111/ene.14316] [Medline: 32399995]
48. Schrag M, Kirshner H. Management of intracerebral hemorrhage: JACC focus seminar. J Am Coll Cardiol 2020 Apr 21;75(15):1819-1831 [FREE Full text] [doi: 10.1016/j.jacc.2019.10.066] [Medline: 32299594]
49. Lu J, Lu Y, Wang X, Li X, Linderman GC, Wu C, et al. Prevalence, awareness, treatment, and control of hypertension in China: data from 1·7 million adults in a population-based screening study (China PEACE Million Persons Project). The Lancet 2017 Dec;390(10112):2549-2558. [doi: 10.1016/s0140-6736(17)32478-9]
50. He Q, Guo H, Bi R, Chen S, Shen J, Long C, et al. Prediction of neurological deterioration after intracerebral hemorrhage: the SIGNALS score. J Am Heart Assoc 2022 Aug 02;11(15):e026379 [FREE Full text] [doi: 10.1161/JAHA.122.026379] [Medline: 35916347]
51. Wan Y, Guo H, Chen S, Chang J, Wang D, Bi R, et al. ADVISING score: a reliable grading scale based on injury and response for intracerebral haemorrhage. Stroke Vasc Neurol 2023 Apr;8(2):111-118 [FREE Full text] [doi: 10.1136/svn-2022-001707] [Medline: 36137597]
52. Poon MTC, Fonville AF, Al-Shahi Salman R. Long-term prognosis after intracerebral haemorrhage: systematic review and meta-analysis. J Neurol Neurosurg Psychiatry 2014 Jun;85(6):660-667. [doi: 10.1136/jnnp-2013-306476] [Medline: 24262916]
53. Li Z, Chen Q, Feng L, Rodewald L, Xia Y, Yu H, et al. Active case finding with case management: the key to tackling the COVID-19 pandemic. The Lancet 2020 Jul;396(10243):63-70. [doi: 10.1016/s0140-6736(20)31278-2]

Abbreviations

ACEI: angiotensin-converting enzyme inhibitor
ARB: angiotensin receptor blockers
CHEERY: Chinese cerebral hemorrhage: mechanism and intervention
ICH: intracerebral hemorrhage
IVH: intraventricular hemorrhage
RR: risk ratio
SMASH-U: structural vascular lesions, medication, amyloid angiopathy, systemic disease, hypertension, and undetermined
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by T Sanchez, A Mavragani; submitted 22.10.21; peer-reviewed by M Sotiropoulos, F Lami, S Hertling; comments to author 29.11.22; revised version received 19.03.23; accepted 05.09.23; published 13.12.23.

Please cite as:
Wan Y, He QW, Chen S, Li M, Xia Y, Zhang L, Sun Z, Chen X, Wang D, Chang J, Hu B
Manifestations and Outcomes of Intracerebral Hemorrhage During the COVID-19 Pandemic in China: Multicenter, Longitudinal Cohort Study
JMIR Public Health Surveill 2023;9:e34386
URL: https://publichealth.jmir.org/2023/1/e34386
doi:10.2196/34386
PMID:38090794

Original Paper

Pharmaceutical Agents as Potential Drivers in the Development of Early-Onset Colorectal Cancer: Case-Control Study

Irit Ben-Aharon[1*], DPhil, MD; Ran Rotem[2*], DPhil; Cheli Melzer-Cohen[3], MSci; Gilad Twig[4,5,6], MD; Andrea Cercek[7], MD; Elizabeth Half[1], MD; Tal Goshen-Lago[1], BA; Gabriel Chodik[5*], DPhil; David Kelsen[7,8*], MD

[1]Department of Gastroenterology, Rambam Healthcare Campus, Haifa, Israel
[2]Harvard T Chan School of Public Health, Boston, MA, United States
[3]KSM Research and Innovation Center, Maccabi Healthcare Services, Tel-Aviv, Israel
[4]The Institute of Endocrinology Diabetes and Metabolism, Sheba Medical Center, Ramat Gan, Israel
[5]Department of Preventive Medicine and Epidemiology, School of Public Health, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[6]Gertner Institute for Epidemiology & Health Policy Research, Sheba Medical Center, Ramat Gan, Israel
[7]Gastrointestinal Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States
[8]Weill Cornell Medical College, New York, NY, United States
*these authors contributed equally

Corresponding Author:
David Kelsen, MD
Gastrointestinal Oncology Service
Department of Medicine
Memorial Sloan Kettering Cancer Center
300 East 66th Street
Office 1027
New York, NY, 10065
United States
Phone: 1 646 888 4179
Email: kelsend@mskcc.org

Abstract

Background: The incidence of early-onset colorectal cancer (EOCRC) rose abruptly in the mid 1990s, is continuing to increase, and has now been noted in many countries. By 2030, 25% of American patients diagnosed with rectal cancer will be 49 years or younger. The large majority of EOCRC cases are not found in patients with germline cancer susceptibility mutations (eg, Lynch syndrome) or inflammatory bowel disease. Thus, environmental or lifestyle factors are suspected drivers. Obesity, sedentary lifestyle, diabetes mellitus, smoking, alcohol, or antibiotics affecting the gut microbiome have been proposed. However, these factors, which have been present since the 1950s, have not yet been conclusively linked to the abrupt increase in EOCRC. The sharp increase suggests the introduction of a new risk factor for young people. We hypothesized that the driver may be an off-target effect of a pharmaceutical agent (ie, one requiring regulatory approval before its use in the general population or an off-label use of a previously approved agent) in a genetically susceptible subgroup of young adults. If a pharmaceutical agent is an EOCRC driving factor, regulatory risk mitigation strategies could be used.

Objective: We aimed to evaluate the possibility that pharmaceutical agents serve as risk factors for EOCRC.

Methods: We conducted a case-control study. Data including demographics, comorbidities, and complete medication dispensing history were obtained from the electronic medical records database of Maccabi Healthcare Services, a state-mandated health provider covering 26% of the Israeli population. The participants included 941 patients with EOCRC (≤50 years of age) diagnosed during 2001-2019 who were density matched at a ratio of 1:10 with 9410 control patients. Patients with inflammatory bowel disease and those with a known inherited cancer susceptibility syndrome were excluded. An advanced machine learning algorithm based on gradient boosted decision trees coupled with Bayesian model optimization and repeated data sampling was used to sort through the very high-dimensional drug dispensing data to identify specific medication groups that were consistently linked with EOCRC while allowing for synergistic or antagonistic interactions between medications. Odds ratios for the identified medication classes were obtained from a conditional logistic regression model.

Results: Out of more than 800 medication classes, we identified several classes that were consistently associated with EOCRC risk across independently trained models. Interactions between medication groups did not seem to substantially affect the risk. In our analysis, drug groups that were consistently positively associated with EOCRC included beta blockers and valerian *(Valeriana officinalis*). Antibiotics were not consistently associated with EOCRC risk.

Conclusions: Our analysis suggests that the development of EOCRC may be correlated with prior use of specific medications. Additional analyses should be used to validate the results. The mechanism of action inducing EOCRC by candidate pharmaceutical agents will then need to be determined.

(JMIR Public Health Surveill 2023;9:e50110) doi:10.2196/50110

KEYWORDS

early onset colorectal cancer; pharmaceutical agents; increased risk; colorectal cancer; health provider; Israel; machine learning; decision tree; gradient boosting; risk factors; decision support; risk; risks; colorectal; cancer; oncology; internal medicine; gastroenterology; gastrointestinal; pharmaceutical; pharmaceuticals; drug; drugs

Introduction

An abrupt increase in the incidence of early-onset colorectal cancer (EOCRC), generally defined as the occurrence of colon or rectal cancer in people less than 50 years of age, began in the mid 1990s; this was first noted in the early 2000s [1-3]. The majority of EOCRC cases are not due to inherited cancer susceptibility genes or inflammatory bowel disease (IBD), which have long been known as risk factors for EOCRC [3-5]. This implies that the driving factors for the increase in EOCRC are environmental, although there is presumably an interplay between genes and the environment. The rise in the incidence of EOCRC has now been documented in many countries [3,6,7]. While an increasing prevalence of obesity, high processed foods diets, diabetes, smoking, alcohol consumption, and sedentary lifestyles have been implicated in EOCRC, these lifestyle and comorbidity factors have been present since the 1950s, and the marked increase in EOCRC was noted beginning in the 1990s [8,9]. Further, many patients with EOCRC are neither obese nor sedentary; data on obesity as an increased risk factor for EOCRC report both increased and decreased risk compared to controls [10-13]. The extensive use of antibiotics has also been suggested as a causative factor through the induction of changes to the spectrum of the gut microbiome, but data regarding the gut microbiome and metabolites in patients with EOCRC compared to unaffected patients are conflicting [14,15].

The observation that the increase in EOCRC incidence was abrupt rather than occurring gradually over decades suggested to us that a factor leading to colonic neoplasia that was not previously widely used was made available to young people over a relatively short period of time. We thus hypothesized that a driver for EOCRC could be an off-target effect of a medication that acts directly or indirectly on the colonic mucosa to result in the activation of oncogenic pathways or silencing of protective pathways in individuals who otherwise have a benign genetic makeup. These medications were either not previously available or not previously widely used and were prescribed by physicians or obtained over-the-counter (OTC) by young people. These medications could be new drug approvals by regulatory agencies through new drug applications (NDAs) or supplementary NDAs or drugs that began to be used for uses other than the approved indication (ie, off-label use). The pharmaceutical agent could act either directly on the colonic mucosa or indirectly (eg, via the gut microbiome or its metabolites). Even though EOCRC is increasing in incidence, it still affects only a small fraction of all young people. Young patients developing EOCRC may have a genetic predisposition, such as a genetic polymorphism (possibly resulting in altered drug metabolism), which translates into an off target carcinogenic effect.

To test our hypothesis, we sought to identify pharmaceutical agents used more frequently in patients with EOCRC than in their peer controls and determine which may be risk factors for EOCRC by using a novel machine learning method to analyze a large Israeli electronic medical records (EMRs) database that includes digitized pharmacy records for patients with EOCRC and matched controls.

Methods

Study Population

To identify pharmaceutical agents related to EOCRC, a nested case-control study was performed using data from Maccabi Healthcare Services (MHS), an Israeli state-mandated health provider that serves 2.6 million members (26% of the Israeli population). EMRs have been used in MHS since the mid 1990s. Each patient’s EMR contains individual-level demographic and clinical information, including clinical diagnoses, hospitalizations, medical procedures, laboratory test results, and filled prescriptions.

Case Ascertainment and Matching

MHS maintains a cancer registry though linkage with the Israel National Cancer Registry (INCR). The INCR was established in 1965 with the aim of continuously collecting data on newly diagnosed cancer cases from all medical institutions in Israel. Since 1981, all pathology results from diagnosed cancer cases must be submitted to the INCR. To further ensure complete case ascertainment, the MHS cancer registry supplements the national database with 2 additional sources of data, namely, histological findings from the MHS central laboratory and purchase authorization records for cancer-specific medications.

We used the MHS cancer registry to retrieve information on all patients with colorectal cases first diagnosed between January 1, 2001, to December 31, 2019, who received their first cancer diagnosis at age 50 years or younger (n=1461). To minimize misclassification of medication exposure, we restricted the analysis to patients who were continuously enrolled in MHS for at least 3 years prior to the index date (n=1174). We also excluded 237 cases who had EMR indications of IBD (Crohn disease or ulcerative colitis) or a personal or family history of suspicion for an inherited cancer susceptibility gene. Thus, the final case population included 941 patients with colorectal cancer.

For each case we matched 10 controls through density (risk-set) matching. Controls had to be continuously enrolled in MHS for at least 3 years before the index date, have no personal history of colorectal cancer, and have no evidence of colorectal cancer at the index date. Matching was done using sex, residential district, sociodemographic status (based on enumeration area, which is the smallest geostatistical unit of the Israeli central bureau of statistics), minority subpopulation at the residential area (Israeli Arab or Jewish Orthodox and Ultra Orthodox or Jewish secular), birth year (±2 years), and periphery index (proximity to a large urban center).

Exposure Assessment

MHS members can fill prescriptions at any of more than 700 affiliated pharmacies across the country with a minimum copayment. OTC medications can also be purchased through these pharmacies at a discount. For both cases and controls, we obtained information on all dispensing of medications (OTC and prescription) prior to the index date, implementing a 2-year lag prior to the index date to reduce the likelihood of reverse-causation. To reduce the dimensionality of the data and ease interpretation, individual medications were grouped according to the fifth level (chemical substance) of the Anatomical Therapeutic Chemical classification system. Drugs that were used by less than 0.1% (103/10,351) of the cohort were excluded.

Comorbidity Information

Our goal in this analysis was to identify medication classes that were linked with EOCRC to inform future studies. Thus, we did not aim to make casual interpretations. We collected information on several comorbidities recorded before the index state that have been linked with EOCRC [16-18]. These included diabetes (International Classification of Diseases, Ninth Revision [ICD-9] code 250.x); hypertension (ICD-9 code 401.x); cardiovascular conditions including ischemic heart disease (ICD-9 codes 410.x, 412, 429.7, and 429.79), non–myocardial infarction (ICD-9 codes 36.x, 411.x, 414.x, and 429.2), congestive heart failure (ICD-9 codes 404.x and 428.x), peripheral vascular disease (ICD-9 codes 440.x, 441.x, 442.x, and 443.x), stroke (ICD-9 codes 433.x1 and 438.x), transient ischemic attack (ICD-9 code 435.x), cerebrovascular atherosclerosis (ICD-9 code 434.x0), and atrial fibrillation (ICD-9 code 427.3x); and obesity (ICD-9 code 278.02). Data on comorbidities were obtained from the MHS automated patient registries, which apply case ascertainment algorithms to define patients with chronic diseases according to multiple data sources and disease-specific international guidelines. These include separate registries for cardiovascular disease, diabetes, and hypertension. MHS physicians are required to measure patients' body weight annually and document this in the EHR [17-20]. To assess potential surveillance bias, we also calculated the average frequency of monthly physician visits for each patient by summing all physician visits recorded for a given patient by each patient's follow-up time in the system up to 2 years before the index date, as this can affect the likelihood and timing of receiving an EOCRC diagnosis as well as the nature of medication prescription patterns.

Statistical Analysis

The very high dimensionality of the medication dispensing data limits the use of conventional statistical methods for the analysis since such methods cannot account for the frequent coexposure to many different medication classes. Thus, to conduct a high-throughput screening of medications and identify specific medication groups, individual or in combination, that were consistently linked with EOCRC, we used eXtreme Gradient Boosting (XGBoost), a highly efficient implementation of gradient tree boosting [21]. Briefly, gradient boosting is an ensemble technique that combines the outputs of many individual decision trees to capture complex associations with an outcome of interest. The combining of trees is done iteratively so that each new tree is fit on the residuals of the previous one. XGBoost uses parallel processing to reduce computing time and advanced regularization to enhance the stability of the predictions. The model requires defining several data-specific regularization parameters to optimize performance, which were selected though a 5-fold cross-validation coupled with a Bayesian model optimization [22].

To enhance the reproducibility of our findings, we trained 50 independent XGBoost models, each time setting aside a randomly selected 20% of the data. A list of medication groups associated with EOCRC was derived through global Shapley additive explanations (SHAP) values [23], focusing on medication groups that were associated with the outcome in at least 50% of the independent runs or those with global SHA*P* value in the top 2.5% of the distribution in at least 1 of the runs. A final XGBoost model was then trained with these selected medication groups using the full data to evaluate the presence of informative interactions, assessed through the tree depth hyperparameter following final model optimization. Since machine learning models, including XGBoost, do not provide specific effect estimates, the final list of medications was then included in a conditional logistic regression model to obtain odds ratios (ORs) with 95% CIs for each medication category while accounting for the matched risk sets and adjusting for the frequency of physician visits as well as the patients' history of diabetes, cardiovascular disease, and excess body weight at baseline that were found to be significant in the multivariable model. All statistical analyses were performed using R (version 4.0.2, R Foundation for Statistical Computing).

Ethical Considerations

The study protocol was approved by the MHS institutional review board (MHS IRB 0034-24). Informed consent was waived by the institutional review board because patients' identifying details were removed. There was no compensation provided to individual patients for participation in the study.

Results

Population characteristics are shown in Table 1. The average age at the time of EOCRC diagnosis was 43.6 (SD 6.7) years. The average number of follow-up months and the frequency of physician visits was similar for cases and controls. Cases and controls were well-matched.

Of the more than 800 medication classes used by the study participants, we identified 5 classes that were consistently associated (>50%) with EOCRC risk across the independently trained models. An additional 10 medication groups not consistently associated with EOCRC also had SHAP values in the top 2.5% in at least 1 run (Table 2). Interactions between medication groups did not seem to be informative for risk prediction.

An examination of the effect estimates for the identified medication groups from the conditional regression model adjusted for the aforementioned comorbidities suggested several medication classes whose use was associated with increased odds of EOCRC. These included a beta-blocker, an angiotensin converting enzyme (ACE) inhibitor, and valerian, an herbal medication. Some protective associations were also observed, including for paroxetine, a selective serotonin uptake inhibitor. Antibiotics were not consistently associated with increased risks (Table 3).

Table 1. Demographics of patients with colorectal cancer and controls from data covering the period 2001-2019.

Demographics	Colorectal cases (n=941)	Controls (n=9410)
Sex, n (%)		
Male	459 (48.8)	4590 (48.8)
Female	482 (51.2)	4820 (51.2)
Population subgroup, n (%)		
Secular Jewish	829 (88.1)	8290 (88.1)
Jewish Ultra Orthodox	64 (6.8)	640 (6.8)
Israeli Arab	48 (5.1)	480 (5.1)
Residential area, n (%)		
North	176 (18.7)	1760 (18.7)
Center	183 (19.4)	1830 (19.4)
Jerusalem	211 (22.4)	2110 (22.4)
Sharon	212 (22.5)	2120 (22.5)
South	159 (16.9)	1590 (16.9)
Sociodemographic status[a], mean (SD)	6.5 (1.9)	6.5 (1.9)
Weeks of follow-up before index date, mean (SD)[a]	641 (273)	646 (278)
Age at index date[b] (years), mean (SD)	43.6 (6.7)	43.6 (6.6)
Physician visits over follow-up period[c], yearly mean (SD)	7.2 (6.0)	7.2 (6.0)

[a] Sociodemographic status was measured on a 10-point scale (1: lowest; 10: highest).

[b] The index date was defined as the first date of colorectal cancer diagnosis.

[c] Used as a proxy measure of the frequency of contact with the medical system, which could affect the likelihood and timing of receiving a colorectal cancer diagnosis and medication prescription patterns.

Table 2. Prevalence of the use of pharmaceutical agents in patients with EOCRC and controls using eXtreme Gradient Boosting from data covering the period 2001-2019.

ATC[a] class	ATC description	Prevalence of use in cases, n (%)	Prevalence of use in controls, n (%)	Rate of runs in which an association with CRC[b] was observed[c] (%)	Rate of runs with a global SHAP[d] value ≥97.5% of the distribution[e] (%)
c07aa05	Propranolol	42 (4.5)	228 (2.4)	70	5
c09aa02	Enalapril	44 (4.7)	161 (2.8)	60	3
n05cm09	Valerian	47 (5)	291 (3.1)	58	2
n06ab05	Paroxetine	15 (1.6)	376 (4)	54	10
c10aa01	Simvastatin	99 (10.5)	724 (7.7)	50	4
a10ba02	Metformin	40 (4.3)	234 (2.5)	46	2
d04aa13	Dimetindene	67 (7.1)	906 (9.6)	40	1
s01ca01	Dexamethansone	175 (18.6)	1986 (21.1)	36	3
d07ac13	Mometasone	103 (10.9)	1249 (13.3)	32	1
d11ax18	Diclofenac	86 (9.1)	672 (7.1)	30	1
a02ba03	Famotidine	161 (17.1)	1381 (14.7)	28	1
s01aa01	Chloramphenicol	301 (32)	2782 (29.6)	24	2
j01ce02	Phenoxymethylpenicillin	335 (35.6)	3443 (36.6)	20	1
r01ba52	Pseuroephedrine, combination	243 (25.8)	2631 (28)	20	1
b03ad01	Ferrous amino acid complex	174 (18.5)	1590 (16.9)	16	1
g01aa10	Clindamycin	79 (8.4)	685 (7.3)	10	1
r03ac03	Terbutaline	51 (5.4)	451 (4.8)	2	1

[a] ATC: Anatomical Therapeutic Chemical classification system.

[b] CRC: colorectal cancer.

[c] Based on global SHA*P* values calculated for each of 50 runs. A global SHA*P* value >0 for a given feature was considered evidence for an association with the outcome.

[d] SHAP: Shapley additive explanation.

[e] Considering each of the 50 different global SHA*P* values distributions individually.

Table 3. Odds ratios considering each of the 50 different global Shapley additive explanation value distributions from data covering the period 2001-2019.

ATC[a] class	ATC description	Odds ratio[b] (95% CI)	*P* value
c07aa05	Propranolol	1.94 (1.37-2.77)	<.001
n05cm09	Valerian	1.61 (1.15-2.25)	.01
c09aa02	Enalapril	1.34 (0.9-1.98)	.15
d11ax18	Diclofenac	1.33 (1.04-1.71)	.03
c10aa01	Simvastatin	1.3 (1-1.69)	.05
g01aa10	Clindamycin	1.22 (0.94-1.59)	.14
s01aa01	Chloramphenicol	1.19 (1.02-1.4)	.03
a02ba03	Famotidine	1.17 (0.97-1.42)	.10
r03ac03	Terbutaline	1.17 (0.86-1.59)	.32
b03ad01	Ferrous amino acid complex	1.15 (0.95-1.4)	.16
a10ba02	Metformin	1.14 (0.66-1.96)	.65
j01ce02	Phenoxymethylpenicillin	0.96 (0.82-1.11)	.58
r01ba52	Pseuroephedrine, combination	0.86 (0.73-1.02)	.07
s01ca01	Dexamethansone	0.8 (0.67-0.97)	.02
d07ac13	Mometasone	0.79 (0.63-0.98)	.04
d04aa13	Dimetindene	0.71 (0.54-0.92)	.01
n06ab05	Paroxetine	0.33 (0.2-0.57)	.001

[a] ATC: Anatomical Therapeutic Chemical classification system.

[b] Effect estimates from conditional logistic regression accounting for the matched risk sets, with additional adjustment for the average frequency of physician visits and any history of diabetes, cardiovascular disease, or excess weight prior to the index date.

Discussion

Principal Results

Our analysis of the use of pharmaceutical agents based on data from a quarter of the Israeli population during the period from 2001 to 2019, shortly after the marked increase in EOCRC incidence began, identified several medications that were used significantly more frequently by patients affected by early-onset colorectal cancer than their age-matched controls. Antibiotics, proposed in other studies as the responsible agent for the increase in EOCRC cases by changing the gut microbiome, were not used more frequently in patients with EOCRC than in controls. Agents used more frequently in patients with EOCRC included beta-blockers, ACE inhibitors, and an herbal supplement, valerian.

Comparison With Other EOCRC Research Studies

The incidence of EOCRC has steadily increased since the mid 1990s; it is estimated that of the approximately 153,000 new cases of colorectal cancer diagnosed in 2023 in the United States, 19,550 will be in people younger than 50 years of age [24,25]. By 2030, it is estimated that 1 in 10 new cases of colon cancer and 1 in 4 cases of rectal cancer will occur in young people. EOCRC is on the rise in many countries [7,26].

The abrupt increase in the incidence of EOCRC suggests that the causative agent is an environmental factor, as neither IBD colitis-associated cancers nor inherited cancer susceptibility genes have increased in incidence, and neither are found in the large majority of patients with EOCRC [3,4,27,28]. The continued increasing incidence of EOCRC further suggests that the environmental factor is still present; while the absolute number of cases is small, age cohort analyses suggest the increasing incidence is higher in younger age cohorts (eg, 20-29 years of age) and that succeeding age cohorts have a higher incidence of EOCRC [29-31]. The large majority of EOCRC cases are located on the left side of the colon, especially in the rectum [32,33]. While some studies have shown little difference in the spectrum of tumor genomic alterations between EOCRC and sporadic cancer in older patients, others have not [4,34]. A recent analysis noted differences based on gender and ethnicity in nonhypermutated tumors in patients with EOCRC [4,34,35]. However, these somatic tumor genomic analyses did not focus on the initiating driving factors in “normal” colonic mucosae.

The suspicion that an environmental factor is responsible for the increasing incidence of EOCRC has led to analyses of comorbidities including obesity, diabetes mellitus, and lifestyle factors, such as a sedentary rather than active lifestyle, smoking and alcohol use, and the widespread use of antibiotics leading to a change in the gut microbiome [8,13,15]. For alcohol use, a recent population-based study from South Korea found an association between moderate and heavy alcohol use and an increased incidence of left-sided colon and rectal cancers, but not proximal cancers [36]. The effect was greater in men than in women. However, alcohol use was self-reported and assessed at a single point, not over time, and comorbidities including IBD and a family history of colorectal cancer were not excluded. For other environmental factors, such as obesity, some studies have found an increased EOCRC risk, but others have not [10-12]. Factors such as obesity, diabetes mellitus, more sedentary lifestyles, smoking, and alcohol use have increased gradually over decades and do not address the abrupt increase in the incidence of EOCRC.

The sharp increase in incidence suggests that a new inciting factor was introduced to a target population. This led us to hypothesize that a medication not previously widely used in young people had been introduced into general use and either directly or indirectly initiates neoplasia in colonic mucosa. A medication may have been introduced either through regulatory approval (NDAs), approval for an additional indication for an already approved drug (supplementary NDAs), a change in practice in which an already approved agent is more widely used by physicians to treat a condition other than the approved indication, or a new use of an OTC agent. While the incidence of EOCRC is increasing, it still affects only a small percentage of people under the age of 50 years. We therefore further hypothesize that in younger people who develop EOCRC, an otherwise benign genetic variant (not resulting in colonic neoplasia in the absence of exposure to the suspected medication) is potentiated by exposure to the medication to initiate or propagate neoplasia.

In order to obtain data for an initial evaluation of our hypothesis, it is necessary to have access to both detailed demographic and clinical information linked to detailed medication use data. In Israel, all citizens have access to health care through 4 health maintenance organizations. Detailed and extensive demographic, clinical, and pharmaceutical data are available to allow an analysis of pharmaceutical agents and clinical outcomes. The databases of Israeli health maintenance organizations have enabled the assessment of the risk of vaccine toxicities during the COVID-19 pandemic, evaluating, for example, the incidence and patients affected by myocarditis [37]. For our study, we used digitized records including pharmaceutical use both of prescribed and OTC agents from a large segment of the Israeli population as the source for our data. Our analysis indicated that, as we hypothesized, the use of several pharmaceutical agents was more common in young people diagnosed with EOCRC when compared to their control peers.

In regard to the agents that we have identified in our initial analysis, for propranolol, an agent one would not ordinarily expect to be prescribed to young people, we suspect its use was to treat anxiety or for side effects of antianxiety agents. Antihypertensives, such as ACE inhibitors, may be prescribed more commonly in young people as both stress and other conditions, such as obesity, increase the risk for hypertension. Both propranolol and valerian have been reported to have suppressive effects on established cancer cells; to our knowledge, there are no data as to their effects in inducing dysplasia and cancer in normal colon mucosae cells [38-40]. Data concerning ACE inhibitors and the risk for cancer are conflicting, with a recent report indicating null associations between use of ACE inhibitors and colorectal cancer risk [41]. Previous studies have suggested an association between pharmaceutical agents and cancer, including insulin use and breast cancer, antihypertensives and skin cancer, and sitagliptin and pancreas cancer, but these associations have not yet been confirmed. Yang et al [42] recently reported a web-based algorithm to assess cancer risk for 6 classes of pharmaceuticals (including antihypertensives, antihyperuricemics, antihyperlipidemics, nonsteroidal anti-inflammatory drugs, and antianxiety agents) in Taiwanese patients. However, the study only included agents used within 2 years of cancer diagnosis, which is a period during which the cancer may have already developed. For EOCRC, only 4 patients were between ages of 20-39 years, and the number of patients younger than 50 years was not provided.

The agents we have identified in our initial analysis do not include antibiotics. While it has been suggested that the widespread use of antibiotics during childhood and adolescence may underlie a carcinogenic effect on the gut microbiome, which may be translated into an increased incidence of EOCRC, the data are conflicting. In a recent British study, antibiotic consumption was associated with colorectal cancer in both older and EOCRC cohorts, while a population-based case-control study failed to find conclusive evidence that antibiotics were associated with EOCRC risk [43,44]. In a population-based analysis of the effect of cesarian section (removing exposure to the maternal vaginal flora), while a trend for a higher rate of EOCRC in women born by vaginal delivery was found, the overall analysis did not show a difference in the incidence of EOCRC in individuals born by cesarian section versus those undergoing vaginal delivery [45].

If our hypothesis is correct and validation studies confirm that a pharmaceutical agent or class of agents is a driver in the development of EOCRC, this would lead to regulatory bodies requiring strategies to mitigate risk [46]. Guidance to industry from the US Food and Drug Administration, for example, provides instructions for determining the level of risk, the patient populations at risk, and when available, biomarkers to determine if an individual patient is at risk. If a companion diagnostic test (eg, a genetic polymorphism found in patients with EOCRC but not in their unaffected peers) is developed, this would be included in the labeling indication. Examples of risk mitigation strategies for other illnesses include labeling indication instructions to identify patients with CYP2C19 alleles that poorly or almost totally do not metabolize clopidogel (Plavix) to minimize the risk of using an ineffective drug or drug dose [47]. For thalidomide, no safe dose has been identified for use of this agent in women who are pregnant, and the labeling indication includes a black box warning.

Strengths and Limitations

Strengths of our study include the use of an EMR including digitized pharmaceutical data covering a large percentage of an entire country's population, long-term follow-up, and the use of an advanced machine learning algorithm. We recognize that while we used extensive matching and adjustment for possible confounders, as well as repeated sampling of the data to enhance reproducibility, our results are still susceptible to confounding by indication. That is, rather than reflecting an association with the medications, the observed finding could reflect associations with the underlying conditions for which the medications were prescribed [48]. We also recognize the sample size limitations and have not drawn causality conclusions regarding the agents identified as being used more frequently in patients with EOCRC compared to controls. To assess the risk of publication bias, we systematically reviewed the literature with regard to our hypothesis that a pharmaceutical agent is a driving factor in the abrupt increase in the incidence of EOCRC and found no publications to date that have studied this hypothesis specifically in EOCRC. We regard our current data as supporting our hypothesis but note that validation studies from similar health care system databases are needed. Moreover, prospective studies to identify the mechanism of action of pharmaceutical agents used more frequently by patients with EOCRC than controls on normal colonic mucosae in initiating a neoplastic process in a susceptible population of younger adults are required. We have initiated such studies.

In conclusion, our analysis suggests that EOCRC may be associated with the prior use of specific medications. Additional studies of pharmaceutical agents as possible drivers of EOCRC and the potential mechanisms of action for this effect are underway.

Acknowledgments

This study was supported in part by grants from the Applebaum Foundation, the Lambert Foundation, and the Ginsberg Foundation.

Data Availability

According to Israel Ministry of Health regulations and Maccabi Healthcare Services policy, individual-level data cannot be shared openly (even if deidentified). Specific requests for remote access to deidentified community-level data will be reviewed by Kahn-Sagol-Maccabi, Maccabi Healthcare Services Research and Innovation Center.

Conflicts of Interest

AC declares research support from Seagen and GSK and an advisory role with Merck, Seagen, Pfizer, Bayer, Janssen, and GSK. The other authors have no conflicts of interest to disclose.

References

1. Akimoto N, Ugai T, Zhong R, Hamada T, Fujiyoshi K, Giannakis M, et al. Rising incidence of early-onset colorectal cancer - a call to action. Nat Rev Clin Oncol 2021 Apr;18(4):230-243 [FREE Full text] [doi: 10.1038/s41571-020-00445-1] [Medline: 33219329]
2. Cavestro GM, Mannucci A, Zuppardo RA, Di Leo M, Stoffel E, Tonon G. Early onset sporadic colorectal cancer: worrisome trends and oncogenic features. Dig Liver Dis 2018 Jun;50(6):521-532. [doi: 10.1016/j.dld.2018.02.009] [Medline: 29615301]
3. Spaander MCW, Zauber AG, Syngal S, Blaser MJ, Sung JJ, You YN, et al. Young-onset colorectal cancer. Nat Rev Dis Primers 2023 Apr 27;9(1):21 [FREE Full text] [doi: 10.1038/s41572-023-00432-7] [Medline: 37105987]
4. Cercek A, Chatila WK, Yaeger R, Walch H, Fernandes GDS, Krishnan A, et al. A comprehensive comparison of early-onset and average-onset colorectal cancers. J Natl Cancer Inst 2021 Nov 29;113(12):1683-1692 [FREE Full text] [doi: 10.1093/jnci/djab124] [Medline: 34405229]
5. Ben-Aharon I, van Laarhoven HWM, Fontana E, Obermannova R, Nilsson M, Lordick F. Early-onset cancer in the gastrointestinal tract is on the rise-evidence and implications. Cancer Discov 2023 Mar 01;13(3):538-551. [doi: 10.1158/2159-8290.CD-22-1038] [Medline: 36757194]
6. Siegel RL, Torre LA, Soerjomataram I, Hayes RB, Bray F, Weber TK, et al. Global patterns and trends in colorectal cancer incidence in young adults. Gut 2019 Dec;68(12):2179-2185. [doi: 10.1136/gutjnl-2019-319511] [Medline: 31488504]
7. Anugwom C, Braimoh G, Sultan A, Johnson WM, Debes JD, Mohammed A. Epidemiology and genetics of early onset colorectal cancer-African overview with a focus on Ethiopia. Semin Oncol 2023;50(1-2):28-33. [doi: 10.1053/j.seminoncol.2023.03.007] [Medline: 37032270]
8. Li H, Chen X, Hoffmeister M, Brenner H. Associations of smoking with early- and late-onset colorectal cancer. JNCI Cancer Spectr 2023 Jan 03;7(1):pkad004 [FREE Full text] [doi: 10.1093/jncics/pkad004] [Medline: 36759940]
9. Chen J, Zhang IL, Terry MB, Yang W. Dietary factors and early-onset colorectal cancer in the United States-an ecologic analysis. Cancer Epidemiol Biomarkers Prev 2023 Feb 06;32(2):217-225 [FREE Full text] [doi: 10.1158/1055-9965.EPI-22-0442] [Medline: 36129804]
10. Stoffel EM, Murphy CC. Epidemiology and mechanisms of the increasing incidence of colon and rectal cancers in young adults. Gastroenterology 2020 Jan;158(2):341-353 [FREE Full text] [doi: 10.1053/j.gastro.2019.07.055] [Medline: 31394082]
11. Liu P, Wu K, Ng K, Zauber AG, Nguyen LH, Song M, et al. Association of obesity with risk of early-onset colorectal cancer among women. JAMA Oncol 2019 Jan 01;5(1):37-44 [FREE Full text] [doi: 10.1001/jamaoncol.2018.4280] [Medline: 30326010]
12. Low EE, Demb J, Liu L, Earles A, Bustamante R, Williams CD, et al. Risk factors for early-onset colorectal cancer. Gastroenterology 2020 Aug;159(2):492-501.e7 [FREE Full text] [doi: 10.1053/j.gastro.2020.01.004] [Medline: 31926997]
13. Hua H, Jiang Q, Sun P, Xu X. Risk factors for early-onset colorectal cancer: systematic review and meta-analysis. Front Oncol 2023;13:1132306 [FREE Full text] [doi: 10.3389/fonc.2023.1132306] [Medline: 37213277]

14. Zhang J, Haines C, Watson AJM, Hart AR, Platt MJ, Pardoll DM, et al. Oral antibiotic use and risk of colorectal cancer in the United Kingdom, 1989-2012: a matched case-control study. Gut 2019 Nov;68(11):1971-1978. [doi: 10.1136/gutjnl-2019-318593] [Medline: 31427405]
15. Kong C, Liang L, Liu G, Du L, Yang Y, Liu J, et al. Integrated metagenomic and metabolomic analysis reveals distinct gut-microbiome-derived phenotypes in early-onset colorectal cancer. Gut 2023 Jun;72(6):1129-1142. [doi: 10.1136/gutjnl-2022-327156] [Medline: 35953094]
16. Zhuo X, Melzer Cohen C, Chen J, Chodick G, Alsumali A, Cook J. Validating the UK prospective diabetes study outcome model 2 using data of 94,946 Israeli patients with type 2 diabetes. J Diabetes Complications 2022 Jan;36(1):108086. [doi: 10.1016/j.jdiacomp.2021.108086] [Medline: 34799250]
17. Weitzman D, Chodick G, Shalev V, Grossman C, Grossman E. Prevalence and factors associated with resistant hypertension in a large health maintenance organization in Israel. Hypertension 2014 Sep;64(3):501-507. [doi: 10.1161/HYPERTENSIONAHA.114.03718] [Medline: 24958503]
18. Chodick G, Omer-Gilon M, Derazne E, Puris G, Rotem R, Tzur D, et al. Adolescent body mass index and changes in pre-pregnancy body mass index in relation to risk of gestational diabetes. EClinicalMedicine 2021 Dec;42:101211 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101211] [Medline: 34849479]
19. Shalev V, Chodick G, Goren I, Silber H, Kokia E, Heymann AD. The use of an automated patient registry to manage and monitor cardiovascular conditions and related outcomes in a large health organization. Int J Cardiol 2011 Nov 03;152(3):345-349. [doi: 10.1016/j.ijcard.2010.08.002] [Medline: 20826019]
20. Chodick G, Heymann AD, Shalev V, Kookia E. The epidemiology of diabetes in a large Israeli HMO. Eur J Epidemiol 2003;18(12):1143-1146. [doi: 10.1023/b:ejep.0000006635.36802.c8] [Medline: 14758871]
21. Chen T. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: KDD '16; August 13-17, 2016; San Francisco. [doi: 10.1145/2939672.2939785]
22. Bernd B, Bossek J, Horn D, Thomas J, Lang M. mlrMBO: a modular framework for model-based optimization of expensive black-box functions. arXiv. Preprint posted online on March 9, 2017 [FREE Full text]
23. Lundberg SM, Lee S. A unified approach to interpreting model predictions. 2017 Presented at: 31st Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach.
24. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. CA Cancer J Clin 2020 May;70(3):145-164 [FREE Full text] [doi: 10.3322/caac.21601] [Medline: 32133645]
25. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. CA Cancer J Clin 2023;73(3):233-254 [FREE Full text] [doi: 10.3322/caac.21772] [Medline: 36856579]
26. Huang J, Lucero-Prisno DE, Zhang L, Xu W, Wong SH, Ng SC, et al. Updated epidemiology of gastrointestinal cancers in East Asia. Nat Rev Gastroenterol Hepatol 2023 May;20(5):271-287. [doi: 10.1038/s41575-022-00726-3] [Medline: 36631716]
27. Mosher CA, Brown GR, Weideman RA, Crook TW, Cipher DJ, Spechler SJ, et al. Incidence of colorectal cancer and extracolonic cancers in veteran patients with inflammatory bowel disease. Inflamm Bowel Dis 2018 Feb 15;24(3):617-623. [doi: 10.1093/ibd/izx046] [Medline: 29390104]
28. Malham M, Jakobsen C, Paerregaard A, Virta LJ, Kolho K, Wewer V. The incidence of cancer and mortality in paediatric onset inflammatory bowel disease in Denmark and Finland during a 23-year period: a population-based study. Aliment Pharmacol Ther 2019 Jul;50(1):33-39. [doi: 10.1111/apt.15258] [Medline: 31069829]
29. Siegel RL, Fedewa SA, Anderson WF, Miller KD, Ma J, Rosenberg PS, et al. Colorectal cancer incidence patterns in the United States, 1974-2013. J Natl Cancer Inst 2017 Aug 01;109(8):djw322 [FREE Full text] [doi: 10.1093/jnci/djw322] [Medline: 28376186]
30. Dwyer AJ, Murphy CC, Boland CR, Garcia R, Hampel H, Limburg P, et al. A summary of the fight colorectal cancer working meeting: exploring risk factors and etiology of sporadic early-age onset colorectal cancer. Gastroenterology 2019 Aug;157(2):280-288 [FREE Full text] [doi: 10.1053/j.gastro.2019.04.049] [Medline: 31095950]
31. Tanaka LF, Figueroa SH, Popova V, Klug SJ, Buttmann-Schweiger N. The rising incidence of early-onset colorectal cancer. Dtsch Arztebl Int 2023 Feb 03;120(Forthcoming):59-64 [FREE Full text] [doi: 10.3238/arztebl.m2022.0368] [Medline: 36471648]
32. Tom CM, Mankarious MM, Jeganathan NA, Deutsch M, Koltun WA, Berg AS, et al. Characteristics and outcomes of right-versus left-sided early-onset colorectal cancer. Dis Colon Rectum 2023 Apr 01;66(4):498-510. [doi: 10.1097/DCR.0000000000002273] [Medline: 35001052]
33. Sinicrope FA. Increasing incidence of early-onset colorectal cancer. N Engl J Med 2022 Apr 21;386(16):1547-1558. [doi: 10.1056/NEJMra2200869] [Medline: 35443109]
34. Marx O, Mankarious M, Yochum G. Molecular genetics of early-onset colorectal cancer. World J Biol Chem 2023 Mar 27;14(2):13-27 [FREE Full text] [doi: 10.4331/wjbc.v14.i2.13] [Medline: 37034132]
35. Holowatyj AN, Wen W, Gibbs T, Seagle HM, Keller SR, Edwards DRV, et al. Racial/ethnic and sex differences in somatic cancer gene mutations among patients with early-onset colorectal cancer. Cancer Discov 2023 Mar 01;13(3):570-579 [FREE Full text] [doi: 10.1158/2159-8290.CD-22-0764] [Medline: 36520636]

36. Jin EH, Han K, Shin CM, Lee DH, Kang SJ, Lim JH, et al. Sex and tumor-site differences in the association of alcohol intake with the risk of early-onset colorectal cancer. J Clin Oncol 2023 Aug 01;41(22):3816-3825 [FREE Full text] [doi: 10.1200/JCO.22.01895] [Medline: 37315287]
37. Witberg G, Barda N, Hoss S, Richter I, Wiessman M, Aviv Y, et al. Myocarditis after Covid-19 vaccination in a large health care organization. N Engl J Med 2021 Dec 02;385(23):2132-2139 [FREE Full text] [doi: 10.1056/NEJMoa2110737] [Medline: 34614329]
38. Anselmino LE, Baglioni MV, Malizia F, Laluce NC, Etichetti CB, Marignac VLM, et al. Repositioning metformin and propranolol for colorectal and triple negative breast cancers treatment. Sci Rep 2021 Apr 14;11(1):8091 [FREE Full text] [doi: 10.1038/s41598-021-87525-z] [Medline: 33854147]
39. Liao P, Song K, Zhu Z, Liu Z, Zhang W, Li W, et al. Propranolol suppresses the growth of colorectal cancer through simultaneously activating autologous CD8 T cells and inhibiting tumor AKT/MAPK pathway. Clin Pharmacol Ther 2020 Sep;108(3):606-615. [doi: 10.1002/cpt.1894] [Medline: 32418204]
40. Shi F, Li Y, Han R, Fu A, Wang R, Nusbaum O, et al. Valerian and valeric acid inhibit growth of breast cancer cells possibly by mediating epigenetic modifications. Sci Rep 2021 Jan 28;11(1):2519 [FREE Full text] [doi: 10.1038/s41598-021-81620-x] [Medline: 33510252]
41. Brasky TM, Flores KF, Larson JC, Newton AM, Shadyab AH, Watanabe JH, et al. Associations of angiotensin-converting enzyme inhibitor or angiotensin receptor blocker use with colorectal cancer risk in the women's health initiative. Cancer Epidemiol Biomarkers Prev 2021 May;30(5):1029-1032 [FREE Full text] [doi: 10.1158/1055-9965.EPI-20-1401] [Medline: 33627382]
42. Yang H, Islam MM, Nguyen PAA, Wang C, Poly TN, Huang C, et al. Development of a web-based system for exploring cancer risk with long-term use of drugs: logistic regression approach. JMIR Public Health Surveill 2021 Feb 15;7(2):e21401 [FREE Full text] [doi: 10.2196/21401] [Medline: 33587043]
43. McDowell R, Perrott S, Murchie P, Cardwell C, Hughes C, Samuel L. Oral antibiotic use and early-onset colorectal cancer: findings from a case-control study using a national clinical database. Br J Cancer 2022 Apr;126(6):957-967 [FREE Full text] [doi: 10.1038/s41416-021-01665-7] [Medline: 34921228]
44. Nguyen LH, Cao Y, Batyrbekova N, Roelstraete B, Ma W, Khalili H, et al. Antibiotic therapy and risk of early-onset colorectal cancer: a national case-control study. Clin Transl Gastroenterol 2022 Jan 13;13(1):e00437 [FREE Full text] [doi: 10.14309/ctg.0000000000000437] [Medline: 35029165]
45. Cao Y, Nguyen LH, Tica S, Otegbeye E, Zong X, Roelstraete B, et al. Evaluation of birth by cesarean delivery and development of early-onset colorectal cancer. JAMA Netw Open 2023 Apr 03;6(4):e2310316 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.10316] [Medline: 37103933]
46. Guidance for industry: warnings and precautions, contraindications, boxed warning sections of labeling for human prescription drug and biological products – content and format. Food and Drug Administration. URL: https://www.fda.gov/downloads/drugs/guidances/ucm075096.pdf [accessed 2023-11-24]
47. FDA drug safety communication: reduced effectiveness of Plavix (clopidogrel) in patients who are poor metabolizers of the drug. Food and Drug Administration. URL: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm203888.htm [accessed 2023-11-24]
48. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. JAMA 2016 Nov 01;316(17):1818-1819. [doi: 10.1001/jama.2016.16435] [Medline: 27802529]

Abbreviations

ACE: angiotensin converting enzyme
EMR: electronic medical record
EOCRC: early-onset colorectal cancer
IBD: inflammatory bowel disease
ICD-9: International Classification of Diseases, Ninth Revision
INCR: Israel National Cancer Registry
MHS: Maccabi Healthcare Services
NDA: new drug application
OR: odds ratio
OTC: over-the-counter
SHAP: Shapley additive explanations
XGBoost: eXtreme Gradient Boosting

Edited by A Mavragani, T Sanchez; submitted 22.06.23; peer-reviewed by T Lee, S Matsuda, K Gupta; comments to author 17.07.23; revised version received 08.09.23; accepted 07.11.23; published 13.12.23.

Please cite as:
Ben-Aharon I, Rotem R, Melzer-Cohen C, Twig G, Cercek A, Half E, Goshen-Lago T, Chodik G, Kelsen D
Pharmaceutical Agents as Potential Drivers in the Development of Early-Onset Colorectal Cancer: Case-Control Study
JMIR Public Health Surveill 2023;9:e50110
URL: https://publichealth.jmir.org/2023/1/e50110
doi:10.2196/50110
PMID:37933755

Original Paper

Trends and Projection of the Incidence of Active Pulmonary Tuberculosis in Southwestern China: Age-Period-Cohort Analysis

Jinou Chen[1], MD; Yubing Qiu[1], BA; Wei Wu[1], MD; Rui Yang[1], MD; Ling Li[1], MD; Yunbin Yang[1], MD; Xing Yang[1], MD; Lin Xu[1], PhD

Division of Tuberculosis Control and Prevention, Yunnan Center for Disease Control and Prevention, Kunming, China

Corresponding Author:
Lin Xu, PhD
Division of Tuberculosis Control and Prevention
Yunnan Center for Disease Control and Prevention
No.158 Dongsi Road
Xishan District, Yunnan Province
Kunming, 650000
China
Phone: 86 871 63620068
Fax: 86 871 63638207
Email: xulinth@hotmail.com

Abstract

Background: The control of pulmonary tuberculosis (PTB) is critical for achieving the vision of World Health Organization's End TB goal.

Objective: This study analyzes the temporal trends in PTB incidence associated with age, period, and birth cohorts from 2006 to 2020 in Yunnan, China; projects the PTB burden till 2030; and explores the drivers of PTB incidence.

Methods: The aggregated PTB incidence rates between 2005 and 2020 were obtained from the National Notifiable Disease Reporting System. We used the age-period-cohort model to evaluate the age, period, and cohort effects on PTB incidence. We applied the Bayesian age-period-cohort model to project future PTB incidence from 2021 to 2030. We applied the decomposition algorithm to attribute the incidence trends to population aging, population growth, and age-specific changes from 2006 to 2030.

Results: From 2006 to 2020, the PTB incidence in Yunnan was relatively stable, although the absolute number showed an increase. The net drift was –1.56% (95% CI –2.41% to –0.70%). An M-shaped bimodal local drift and a longitudinal age curve were observed. The overall local drift was below zero for most age groups except for the age groups of 15-19 years (2.37%, 95% CI –0.28% to 5.09%) and 50-54 years (0.41%, 95% CI –1.78% to 2.64%). The highest risk of PTB incidence was observed in the age group of 65-69 years, and another peak was observed in the age group of 20-24 years. Downward trends were observed for both period and cohort effects, but the cohort effect trends were uneven. A higher risk was observed for the birth cohorts of 1961-1970 (rate ratio [RR]$_{1961\text{-}1965}$=1.10, 95% CI 0.88-1.38; RR$_{1966\text{-}1970}$=1.11, 95% CI 0.92-1.37) and 2001-2010 (RR$_{2001\text{-}2005}$=0.92, 95% CI 0.63-1.34; RR$_{2006\text{-}2010}$=0.84, 95% CI 0.45-1.58) than for the adjacent cohorts. The Bayesian age-period-cohort model projected that PTB incidence will continually increase from 2021 to 2030 and that PTB incidence in 2030 will be 2.28 times higher than that in 2006. The age-specific change was the leading cause for the growing PTB disease burden.

Conclusions: Although there are several levels and measures for PTB control, the disease burden is likely to increase in the future. To bridge the gap of TB-free vision, our study suggests that public health policies be put in place soon, including large-scale active case-finding, priority prevention policies for high-risk older adult and young adult populations, and reduction of possible grandparent-grandchildren transmission patterns.

(JMIR Public Health Surveill 2023;9:e48015) doi:10.2196/48015

KEYWORDS

pulmonary tuberculosis; age-period-cohort modeling; projection; decomposition of disease burden; bayesian age-period-cohort model

Introduction

Tuberculosis (TB) is a chronic communicable disease caused by *Mycobacterium tuberculosis*, which typically affects the pulmonary tissues and is therefore referred to as pulmonary tuberculosis (PTB). This disease can affect other sites as well. The TB epidemic is a major public health challenge. Globally, until the emergence of the COVID-19 pandemic, TB was the leading cause of death from a single infectious agent [1]. The World Health Organization's End TB strategy and the United Nation's Sustainable Development Goal to end the TB epidemic by 2035 has ambitious targets of 95% reduction in TB deaths and 90% reduction in TB new case incidence by 2035 globally [2].

China is one of the highest TB burden countries, with 7.4% of the estimated incident cases worldwide in 2022 [1]. Although a downward trend of TB prevalence has been observed in the recent decades [3], a large number of new PTB cases have caused a high disease burden in the country. Therefore, the estimation of the PTB burden in China is critical for achieving the vision of the End TB goal. Several studies [4-6] have reported on the PTB burden trend across time in China. These studies have shown the historical trend of TB incidence, but because those studies failed to differentiate the detailed contributions of age, calendar period, and birth cohort to the overall temporal trends, the effects of the TB control policy before 2005 were hard to evaluate. Additionally, research on prospective trends and projecting the future trends of TB burden is limited. It is necessary to evaluate the localized TB burden in specific regions to improve the control strategy. We addressed these knowledge gaps by examining the age-, period-, and cohort-associated TB incidence in the Yunnan province of China by age-period-cohort (APC) analysis.

The APC approach is a statistical framework for the trend estimation of disease incidence and mortality, and this approach has been extensively applied in cancer epidemiology [7-11], infectious disease modeling [12,13], and chronic disease research [14]. A wide range of research fields are required to understand the temporal changes in disease outcomes attributed to 3 components by applying the APC model. First, the age effects were the changes related to the biological and behavioral changes that affected the aging of individuals, representing the outcomes in terms of different age groups. Second, the period effects were the changes related to public health or social events in particular years; this component affected all age groups and birth cohorts simultaneously in a calendar period. Third, the cohort effects were the outcome changes of birth in the same calendar years, which reflected the specific epidemiological exposure of the different birth cohorts over the life course.

This study demonstrates the effects of age, period, and cohort on the incidence of PTB in the Yunnan province between 2006 and 2020. This study forecasts the PTB incidence in 2030 by the Bayesian APC (BAPC) model. The decomposition analysis attributes the future PTB incidence changes to population aging, population growth, and age-specific reasons. This study will elucidate the various trends of TB incidence in the future and identify the potential drivers of the changes in PTB incidence. To achieve the Sustainable Development Goal in 2035, this study may provide evidence for future control strategies and help in public health decision-making, design of PTB screening, and resource allocation.

Methods

Study Sites

Yunnan, a province in the southwest of China, has 16 prefectures and 129 counties. The population of Yunan was 44,499,123 in 2006, consisting of 22,957,243 males and 21,541,880 females [15,16]. By 2020, the population grew to 48,583,533 consisting of 25,204,786 males and 23,378,747 females [17]. The population census presents the population structure shifting and population aging in Yunnan in 2 decades. The ratio of the population younger than 15 years was 25.96%, 20.73%, and 19.57% and that of the population older than 65 years was 6.09%, 11.06%, and 10.75% in 2000, 2010, and 2020, respectively [15-17].

Ethical Considerations

This study was submitted to the ethics committee of the Yunnan Center for Disease Control and Prevention. This study does not involve human participants and animals. Ethics approval was not required for this study because we did not include any identifiable data of patients' personal information, including name, identity information, address, and telephone number. This study only reviewed secondary aggregated data at the population level; therefore, waiver of written informed consent was taken (institutional review board exemption: 2023-14).

Data Source

Active PTB has been notified through the National Notifiable Disease Reporting System (NNDRS) in China since 2004. Thereafter, the Tuberculosis Information Management System was built and applied to complete a standard TB case notification and follow-up in 2005. PTB incidence data between 2005 and 2020 were collected from NNDRS. Notifying PTB cases is considered as mandatory routine work and is enforced by law in China. When a patient with TB visits a health facility and is diagnosed with presumptive or confirmed PTB by physicians, his or her information must be notified to NNDRS by this health facility. Theoretically, NNDRS can capture the data of all the patients with PTB who seek medical advice in a region. After the presumptive cases are diagnosed further, the aggregated clinically diagnosed and laboratory confirmed cases are considered as PTB incidence. To guarantee the incidence data, the NNDRS includes only new PTB cases that have not been treated or reported earlier. This system can manually or automatically mark and delete the duplicate cases by searching and matching identical information of the full names and residential identification card numbers of historically notified cases. Whenever PTB cases are identified as duplicate status, these are excluded from the incidence statistics in NNDRS. The annual aggregated data contain the incidence numbers and incidence rates of PTB, and the data are stratified by the calendar year of notification, age, and sex. The annual population data between 2005 and 2020 were generated from NNDRS. The source of the projected population data was based on the United

Nations report on world population prospects [18]. Our research used the Chinese population projection between 2021 and 2030. We generated the Yunnan provincial population projection by a fixed proportion of the provincial population among the Chinese population.

Statistical Analysis

All statistical analyses were conducted with R software (version 4.0.2; R Core Team). *P* levels <.05 were considered statistically significant for 2-sided tests.

APC Model

In this study, we used the APC model to illustrate the age, period, and cohort effects on PTB incidence simultaneously. This approach attempts to resolve and decompose these effects and identify the potential triggers for the varying PTB incidence trends. Due to the perfect linear dependency of the age, period, and cohort variables, the linear effect of the APC model cannot be uniquely estimated, and the difficulty in solving the model parameters is called the APC identification problem [19,20]. The intrinsic estimator was developed based on the decomposition of the matrix singular value, which could fit a stable solution of parameter estimation [21]. The generic form of the APC model can be written as follows: $Y_{ijk} = \log(\lambda_{ijk}) = \mu + \alpha_i + \pi_j + \gamma_k + \in_{ijk}$

where Y_{ijk} is the outcome variable to be explained, assumed to follow a quasi-Poisson distribution. The Y_{ijk} or $\log(\lambda_{ijk})$ correspond to the PTB incidence rate for the *i*th age group, *j*th period, and *k*th birth cohort. The μ is the intercept, representing the overall average effect. The error $\in_{ijk}$ is assumed to follow the additive, independent, and identically normal distribution. The α, π, and γ are the log-linear model coefficients corresponding to age, period, and cohort effects, respectively. The usual constraints are the sum of the parameters to zero, presented as $\sum\alpha_i = \sum\pi_j = \sum\gamma_k = 0$.

The APC model requires equal intervals for period, age, and cohort. Therefore, we aggregated the data and then divided each component interval into widths of 5 years. We divided age into 18 age groups (0-4, 5-9, 10-14, …, 80-84, >85), period into 3 periods (2006-2010, 2011-2015, 2016-2020), and cohort into 20 birth cohorts (1921-1925, 1926-1930, …, 2016-2020). A tool developed by the National Cancer Institute was used for the APC model parameters and effects estimation [22]. The net drift, local drifts, longitudinal age curve, period, and cohort rate ratio (RR) were evaluated.

BAPC Model

We projected the future incidence of PTB cases from 2021 to 2030 by using the BAPC model. The BAPC model is based on the previously stated population projection. The Bayesian approach divides the deviance into age, period, and cohort effects and treats the unknown parameters as random with an appropriate prior distribution. The second-order random walk smoothing the prior and guaranteeing the second differences in the models were identifiable, which was a random variance assumed to follow an independent mean-zero normal distribution. The BAPC model used the integrated nested Laplace approximation methods to estimate the posterior marginal distributions directly, and then, integrated nested Laplace approximation made extrapolations attributed to age, period, and cohort effects [23,24]. The model fitting was free of Markov Chain Monte Carlo simulations—the reason was to avoid the convergence issues introduced by Markov Chain Monte Carlo sampling that was generally used in the Bayesian approach.

Decomposition of the Projected Burden

We further analyzed the triggers for the changes in PTB incidence trends from 2005 to 2030 by the decomposition approach. The decomposition method was developed recently [25]. Briefly, this method attributes changes in PTB incidence to 3 components: population aging, population growth, and age-specific changes between 2005 and each following year to 2030. The method of decomposition has been documented elsewhere, and details are given in Multimedia Appendix 1. The method of decomposition considers the 2-way and 3-way interaction of these 3 components, and decomposition is more robust than other methods [26]. We analyzed the absolute and relative contributions of the 3 components to the change incidence of PTB based on previous BAPC projections. The absolute contribution was the attributed incidence number, while the relative contribution was estimated as the attributed incidence divided by the total incidence in 2006 × 100%. The positive contributions correspond to an increase in PTB incidence, and the negative contributions correspond to a decrease in PTB incidence. The net change was the total PTB incidence comprising these 3 components. The age-specific changes in PTB incidence refer to all the differences in the incidence that cannot be explained by population aging and population growth.

Results

Temporal Trends in PTB Incidence

In 2006-2020, the annual mean PTB incidence was 58.57 per 100,000; the rate for males (95.71 per 100,000) was higher than that for females (46.77 per 100,000). The sex ratio (male:female ratio) for incidence showed a decline from 2.14 in 2006 to 1.98 in 2020. Between 2006 and 2020, the incidence of PTB increased from 18,336 in 2006 to 19,412 in 2020 for males and from 8552 in 2006 to 9768 in 2020 for females. However, the incidence rate was relatively stable, which was 60.42 per 100,000 in 2006 and 60.06 per 100,000 in 2020 (Figure 1). Across years, the age-standardized incidence rate decreased in most age groups for both males and females, which is presented along each row of the PTB Lexis diagram in Figure 2. A comparison across age groups along each column of the PTB Lexis diagram shows that the group younger than 15 years showed a lower PTB incidence. The highest PTB incidence was observed in the age group of 70-74 years, with the highest PTB incidence among males aged >85 years who were older than the females (age 70-74 years). The diagonals in the Lexis diagram present the different cohorts, overall, and for both sexes. The birth cohorts of 1946-1950, 1951-1955, and 1956-1960 showed higher PTB incidence than other adjacent cohorts, while the birth cohorts of 1981-1985 and 1986-1990 presented higher PTB incidence across periods.

Figure 1. Changes in the number of cases and incidence rate of pulmonary tuberculosis stratified by gender in Yunnan, China, from 2006 to 2020. (A) Number of pulmonary tuberculosis cases stratified by gender. (B) Incidence rate of pulmonary tuberculosis stratified by gender.

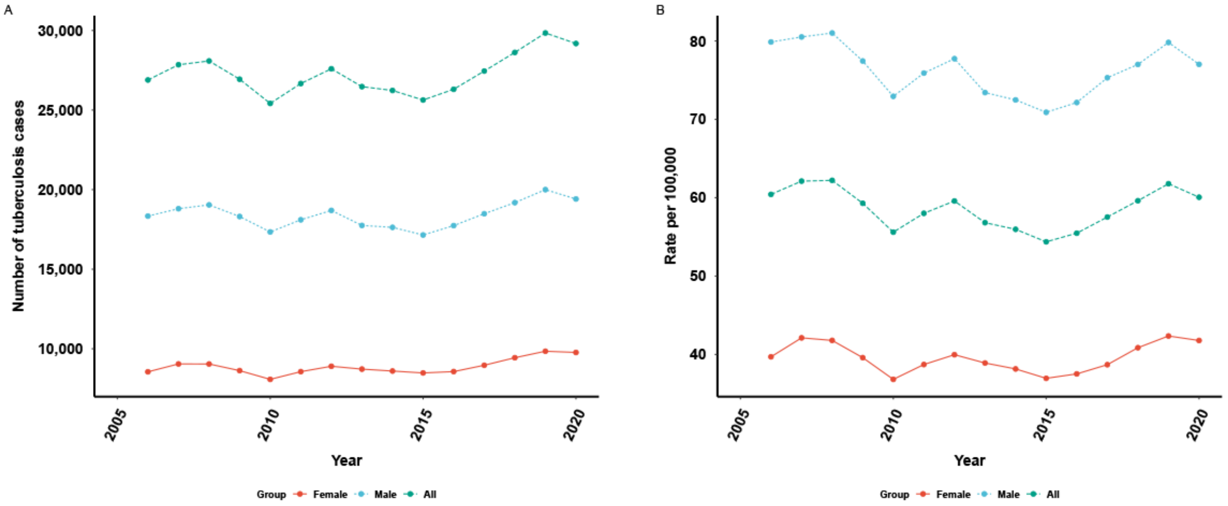
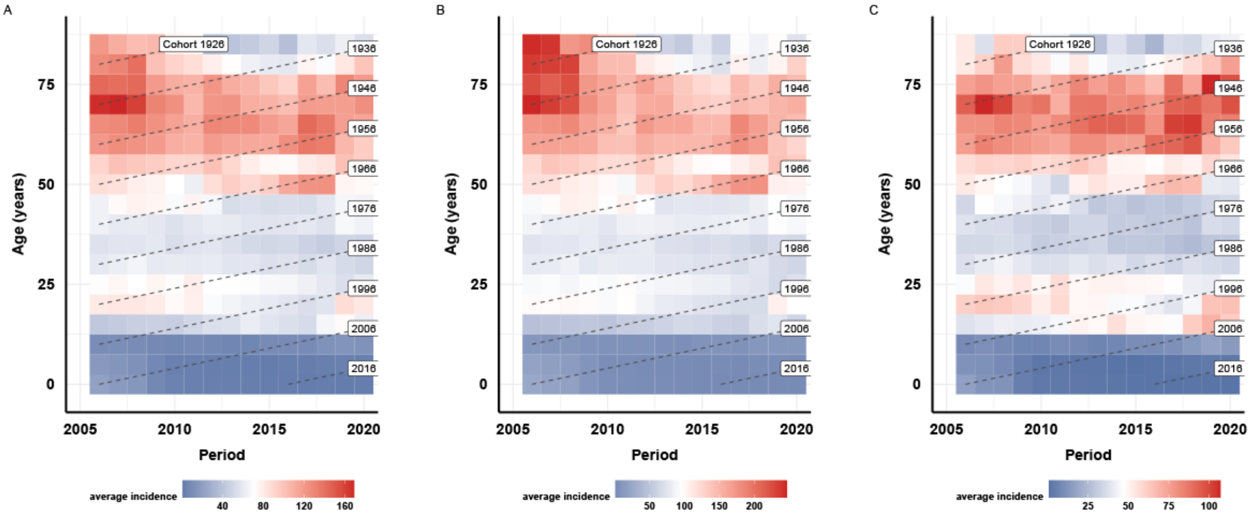


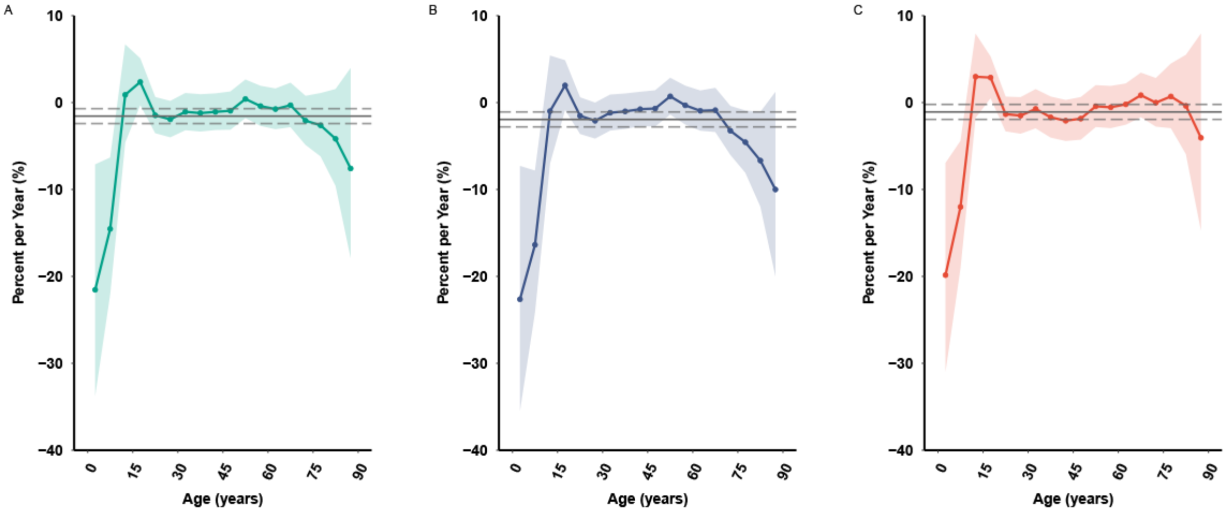
Figure 2. Pulmonary tuberculosis incidence matrix and Lexis diagram by year of diagnosis and age at diagnosis in Yunnan, China, from 2006 to 2020. (A) Incidence matrix of pulmonary tuberculosis. (B) Incidence matrix of pulmonary tuberculosis for males. (C) Incidence matrix of pulmonary tuberculosis for females. Blue represents lower incidence rate, while red represents higher incidence rate. The scales in each panel are different. For each panel, the rows represent the incidence across calendar years for the specific age group, the columns represent the incidence across the specific age group for that calendar year, and the dash lines on the diagonal represent the changes in the incidence for the birth cohort.



APC Model Analysis of PTB Incidence

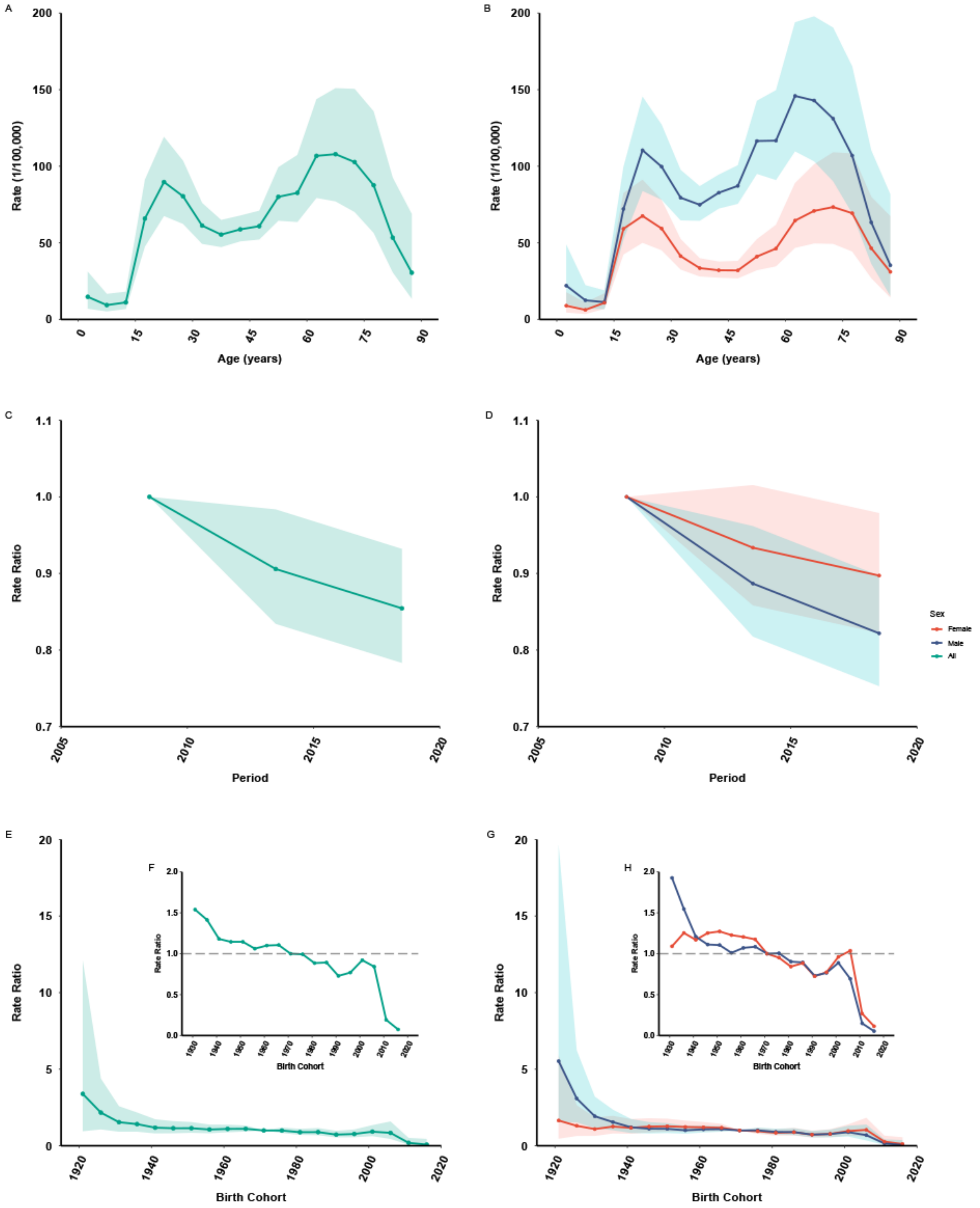
The estimated age, period, and cohort effects on PTB incidence are shown in Figure 3 and Figure 4. The net drift presented the overall log-linear annual percentage change of incidence across periods and birth cohorts. The overall net drift was –1.56% (95% CI –2.41% to –0.70%). Differences across sex were observed for males (–1.94%, 95% CI –2.79% to –1.07%) and females (–1.07%, 95% CI –1.94% to –0.21%). PTB incidence among females was lesser than that among males between 2006 and 2020. M-shaped local drift curves were observed, which refer to the difference in the age-specific annual percentage changes. The overall local drift was above zero for 2 age groups, that is, the group aged 15-19 years (local drift 2.37%, 95% CI –0.28% to 5.09%) and the group aged 50-54 years (local drift 0.41%, 95% CI –1.78% to 2.64%). The local drift was below zero for most of the other age groups. The trends of the estimated PTB incidence increased across years for these 2 age groups. Identical effects were observed in both sexes, although the female local drift was above zero for the group aged 65-69 years. The longitudinal age curves presented 2 peaks as well (Figure 4A). For both sexes, after adjusting the period and birth cohort effects, the PTB incidence presented an “up-down-up” trend across ages (Figure 4B). The highest risk of incidence was in the age group of 65-69 years, with another peak around the age groups of 15-19 years, 20-24 years, and 25-29 years.

Figure 3. Net drifts and local drifts of pulmonary tuberculosis incidence in Yunnan, China, from 2006 to 2020. (A) Net drifts and local drifts of pulmonary tuberculosis. (B) Net drifts and local drifts of pulmonary tuberculosis for males. (C) Net drifts and local drifts of pulmonary tuberculosis for females. The horizontal grey solid line represents the net drift values, while the grey dash line corresponds to the 95% CIs. The curve with a colored solid line represents the local drift values, and the colored ribbon shows the 95% CIs.



The period (or cohort) effects were the rate ratios of age-specific rates in each period (or cohort) relative to the reference. Downward trends (RR=0.85, 95% CI 0.78-0.93, 2016-2020 vs 2006-2010) were observed in the period effects for overall and for both sexes. After adjusting for the age and birth cohort effects, a faster decline for the PTB period effect was observed for males (RR=0.82, 95% CI 0.75-0.89, 2016-2020 vs 2006-2010) than for females (RR=0.89, 95% CI 0.82-0.97, 2016-2020 vs 2006-2010, Figure 4C and Figure 4D). The pattern for cohort effects was similar to that for the period effects. The cohort effects continuously declined from the earliest cohort (RR=3.38, 95% CI 0.94-12.08) to the latest cohort (RR=0.07, 95% CI 0.01-0.45). The gender-stratified trends for cohort effects showed a downward trend, with the effects for males declining more rapidly than those for females (Figure 4E and Figure 4G). Furthermore, the downward trend was uneven; the convex cohort effect curve presented a higher RR for the birth cohort 1961-1970 ($RR_{1961\text{-}1965}$=1.10, 95% CI 0.88-1.38; $RR_{1966\text{-}1970}$=1.11, 95% CI 0.92-1.37) and 2001-2010 ($RR_{2001\text{-}2005}$=0.92, 95% CI 0.63-1.34; $RR_{2006\text{-}2010}$=0.84, 95% CI 0.45-1.58) than the adjacent cohorts for overall and both sexes (Figure 4F and Figure 4H). Moreover, the net drift, local drifts, period, and cohort effects were all statistically significant ($P<.05$) checked by the Wald χ^2 test (Table S1 of Multimedia Appendix 1).

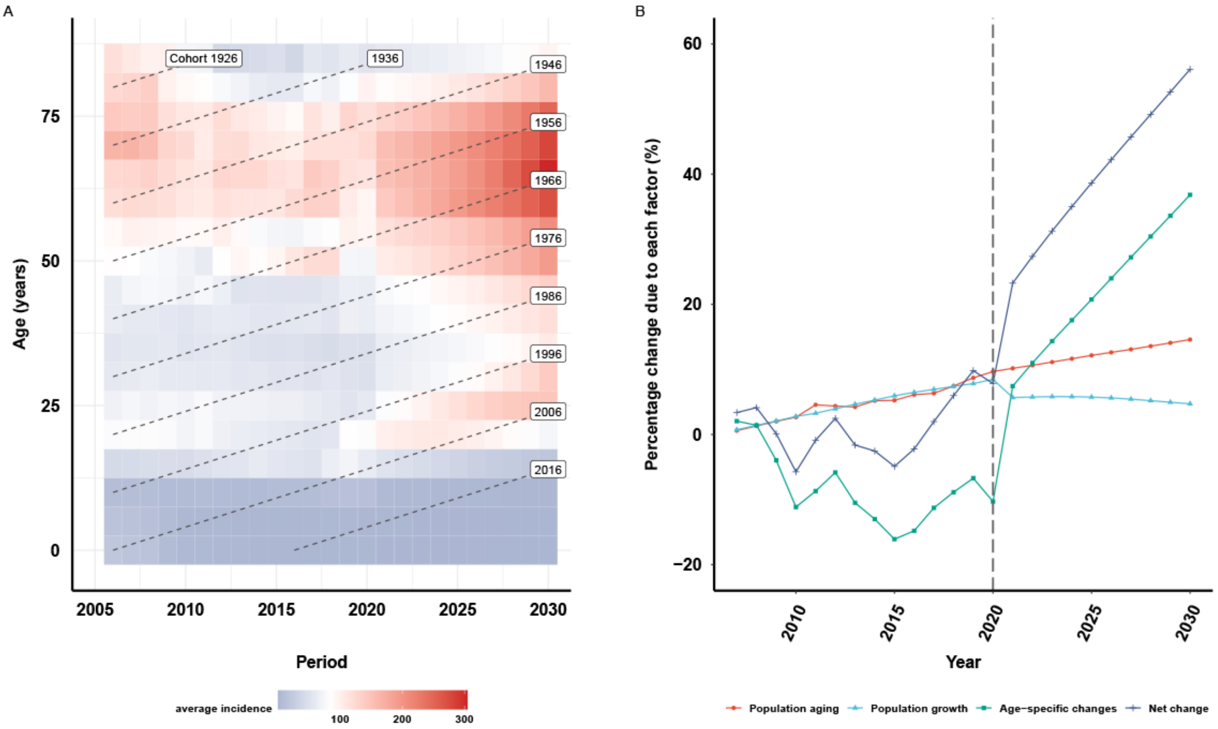
Figure 4. The longitudinal age curve, period, and birth cohort effects curve for pulmonary tuberculosis in Yunnan, China, from 2006 to 2020. (A) Overall longitudinal age curve and the 95% CIs. (B) Longitudinal age curve stratified by gender. (C) Overall period effect and the 95% CIs for pulmonary tuberculosis compared with reference period (2006-2010), adjusted with age and cohort effects. (D) Period effect stratified by gender. (E) Overall cohort effect and the 95% CIs for pulmonary tuberculosis compared with reference cohort (1971-1976), adjusted with age and period effects. (F) Detailed overall cohort effect of birth cohorts between 1931 and 2016. (G) Cohort effect stratified by gender. (H) Detailed cohort effect of birth cohorts between 1931 and 2016, stratified by gender.



BAPC-Projected PTB Incidence

The BAPC predicted that total PTB incidence would continue to increase, with 61,306 in 2030 in Yunnan, which is 2.28 times that of the year 2006 (Table S2 of Multimedia Appendix 1). However, the projection presented heterogeneity across age groups. Age groups younger than 15 years show a continually decreased PTB incidence, although the older adults and younger adults presented an increasing trend, especially in the birth cohorts of 1961-1970 and 2001-2010 (age range of 60-69 years and 20-29 years, respectively) in 2030 (Figure 5A).

Figure 5. The projected incidence for pulmonary tuberculosis and the decomposition to changes in the pulmonary tuberculosis incidence in Yunnan, China, from 2006 to 2030. (A) The Bayesian age-period-cohort model–projected incidence for pulmonary tuberculosis from 2006 to 2030. (B) Decomposition of the contribution of population aging, population growth, and age-specific changes to the projected incidence of pulmonary tuberculosis from 2006 to 2030.



Decomposition of the PTB Burden

The trigger of PTB incidence from 2006 to 2030 was analyzed by a demographic decomposition. The results showed that age-specific incidence was the primary factor for the incremental trend of PTB incidence (Table 1, Figure 5B). The projection suggested that there are appended 34,396 (56.11%) PTB cases in 2030 compared with the reference year of 2006. The ordered contribution of the 3 factors to the increased trend is the age-specific incidence (36.82% increase from 2006), population aging (14.58% increase from 2006), and population growth (4.71% increase from 2006).

Table 1. The projection, decomposition, and contribution of 3 factors to pulmonary tuberculosis incidence estimated by the Bayesian age-period-cohort model from 2007 to 2030 in Yunnan, China.

Year[a]	Total population (N)	Population aging, n (%)	Population growth, n (%)	Age-specific incidence, n (%)	Net changes, n (%)
2007	44,830,030	163 (0.59)	203 (0.73)	571 (2.05)	937 (3.36)
2008	45,139,978	367 (1.31)	393 (1.40)	397 (1.41)	1157 (4.12)
2009	45,430,010	542 (2.01)	558 (2.07)	–1069 (–3.97)	31 (0.12)
2010	45,710,006	675 (2.65)	704 (2.76)	–2839 (–11.15)	–1460 (–5.74)
2011	45,966,362	1216 (4.56)	870 (3.26)	–2321 (–8.70)	–235 (–0.88)
2012	46,310,029	1203 (4.36)	1088 (3.94)	–1610 (–5.83)	682 (2.47)
2013	46,590,006	1117 (4.22)	1231 (4.65)	–2779 (–10.50)	–431 (–1.63)
2014	46,866,000	1360 (5.18)	1385 (5.28)	–3412 (–13)	–667 (–2.54)
2015	47,139,000	1344 (5.24)	1524 (5.94)	–4127 (–16.09)	–1259 (–4.91)
2016	47,418,000	1607 (6.11)	1704 (6.47)	–3897 (–14.80)	–586 (–2.23)
2017	47,710,002	1737 (6.33)	1902 (6.93)	–3096 (–11.28)	543 (1.98)
2018	48,004,978	2132 (7.45)	2116 (7.39)	–2538 (–8.87)	1709 (5.97)
2019	48,299,907	2596 (8.70)	2334 (7.82)	–2003 (–6.72)	2926 (9.81)
2020	48,583,533	2821 (9.66)	2472 (8.47)	–3009 (–10.31)	2284 (7.82)
2021	47,465,047	3561 (10.16)	1988 (5.67)	2602 (7.42)	8152 (23.25)
2022	47,590,518	3934 (10.62)	2131 (5.75)	4067 (10.98)	10,132 (27.35)
2023	47,706,077	4354 (11.12)	2274 (5.81)	5611 (14.33)	12,239 (31.26)
2024	47,792,273	4817 (11.64)	2406 (5.81)	7265 (17.55)	14,488 (35)
2025	47,837,336	5330 (12.15)	2517 (5.74)	9099 (20.75)	16,946 (38.64)
2026	47,850,023	5869 (12.61)	2614 (5.61)	11,168 (23.99)	19,652 (42.21)
2027	47,822,401	6481 (13.07)	2690 (5.43)	13,496 (27.22)	22,667 (45.72)
2028	47,765,525	7182 (13.56)	2750 (5.19)	16,105 (30.42)	26,037 (49.18)
2029	47,697,145	7991 (14.07)	2812 (4.95)	19,080 (33.60)	29,883 (52.62)
2030	47,628,061	8939 (14.58)	2887 (4.71)	22,570 (36.82)	34,396 (56.11)

[a] The year 2006 is the reference.

Discussion

Principal Findings

Our study first describes and analyzes the age, period, and cohort effects on PTB incidence in Yunnan, China, between 2006 and 2020 by using the APC model. Then, we predicted the PTB incidence from 2021 to 2030 by using the integrated nested Laplace approximation technique to estimate the BAPC model. Our research further decomposes the drivers of trends and changes in PTB incidence from 2006 to 2030. We found reduction in PTB incidence coupled with negative net drifts in Yunnan between 2006 and 2020. For the age effect, a 2-peak pattern was observed for both local drifts and the longitudinal age curve, which showed that young adults (age 15-29 years) and older adults (age 60-74 years) were high-risk groups for the priority intervention. The APC analysis presented decreased period effects from 2006 to 2020, which suggests the effectiveness of local PTB strategies and other myriad public health implementations in the last few decades. Besides, an exceptional decline presented in cohort effects implies potential transmission patterns. The BAPC predicted a 2.2 times increment in PTB incidence in 2030 compared with that in 2006—the dominant factor that led to the dramatic growth was age-specific changes.

The descriptive analysis showed that from 2006 to 2020, the incidence rate of PTB was stable. However, the incidence increased for both sexes. This inconsistency in the incidence rate and the absolute number of incidence cases is attributed to the population growth in Yunnan in the last few decades [15-17]. The Lexis diagram illustrates a direct landscape of the cross-sectional, longitudinal, and diagonal representation of PTB incidence; the 3-directional depiction explicitly indicated that the PTB incidence among older adults was much higher than that among young adults (age 15-29 years). However, the APC model uncovered the true age pattern of PTB incidence, which was not in line with the Lexis diagram illustration. The APC model presented 2 critical indicators of net drift and local drift. The overall and sex-stratified net drift decreased and was negative in the study period between 2006 to 2020, suggesting that the PTB incidence declined across periods and birth cohorts.

The age effect represents the rate differences among age groups—this may be associated with complex factors such as biological aging, personal health condition, and lifestyle behavior. The age-specific temporal trends were illustrated by longitudinal age curves in this study. The curve showed quite a low incidence in populations younger than 15 years. One possible reason was the national neonatal Bacillus Calmette-Guérin (BCG) vaccination program. The BCG vaccination policy was activated during the 1960s in China. The neonatal BCG vaccination program has been performed nationwide since the 1970s, and the high vaccination coverage has not been interrupted until now. Another possible reason was the difficulty in the diagnosis of pediatric tuberculosis—there were cases of PTB unreported or unnotified due to misdiagnosis. The unreported rate was 58.3%, 20.3%, and 15.3% among populations younger than 15 years, populations aged 16-64 years, and populations older than 65 years, respectively. The risk of unreported tuberculosis in populations younger than 15 years was 7.3 times higher than that in populations older than 65 years [27]. A steep increase was observed in populations older than 15 years—the first incidence peaked in young populations in the age range of 15-29 years. The infection risk in this population may have peaked due to their behavioral changes with large social circles and communication networks. Moreover, the age group of 15-29 years may enjoy a higher degree of freedom and mobility than those younger than 15 years; this correlated with higher chance of exposure to the infectious source of patients with active PTB. The highest peak was observed in older adults aged 65-69 years likely due to immune deficiency diseases, comorbidities, and unhealthy lifestyle behaviors.

The sex differences in the curve may be caused by the different lifestyles of males and females—men have more broader social contacts and networks than women. Besides, men who experienced potential sex-different sociobehavioral risk factors of smoking and alcohol drinking were more likely to develop active PTB, though the difference in incidence between the sexes was not significant in specific age groups. In this study, the M-shaped bimodal curve was observed for PTB incidence rate. The "up-down-up" pattern for the longitudinal age curve was consistent with that reported in other studies in China [5,6]. These consistency trends revealed that young adults (aged 15-29 years) and older adults (aged 60-74 years) were high-risk groups for PTB incidence from a broader perspective.

The local drift distribution was also M-shaped. The definition of local drift is the estimated annual percentage change over time specific to age groups. Our results indicated that PTB incidence increased in populations aged 15-19 years and in those aged 50-54 years across the study period. This finding suggests that instant public health interventions should be prioritized for these high-risk group populations. The PTB burden must significantly decrease in the coming years in China to bridge the gap of the End TB strategy and to achieve the Sustainable Development Goals [28]. Although the active case-finding strategy has been partially implemented as pilot and focused on students and older adults in China [29], the screening policy and preventive therapy specific to high-risk populations should be endorsed as routine work of the national TB program. Case-finding practices in the physical examination screening policy should be implemented routinely for campuses and workplaces, especially for new students and new employees. The screening strategy should consider special sites such as military camps, jails, and nursing homes.

Period effects were modified by macrofactors such as shift in the socioeconomic and medical status, but the period effects on PTB incidence were equal across all age groups and birth cohorts within calendar years. This study presents the continuous decline of the period effect over the past 15 years. The decline is attributed to the elimination strategy, wherein patients with PTB are detected as early as possible and treated with antituberculosis drugs. In the last few decades, the key approach for managing the PTB epidemic in Yunnan has been to control the source of infection. First, Yunnan implemented the directly observed treatment short-course strategy in 2004, which promoted case detection and cured more patients. Second, the surveillance system of NNDRS and Tuberculosis Information Management System were applied simultaneously since 2004, wherein cases were notified and registered by a modernized web-based instant information system. Third, between 2015 and 2017, Yunnan reformed the TB delivery system to the "3-in-1" model. This model referred to the Center of Disease Control and Prevention responsible for comprehensive management, the TB-designated hospital responsible for diagnosis and treatment, and the community-level medical and health care institutions responsible for referral and directly observed treatment short-course management [30]. The tripartite cooperation promoted the accessibility and availability of PTB diagnosis and treatment. Fourth, because there was moderate prevalence of HIV in Yunnan [31], TB/HIV bidirectional screening lowered both the burden of TB and HIV disease, thereby reducing TB transmission among patients with high-risk HIV/AIDS. Fifth, a pilot study in Yunnan suggested that multiple rounds of case-finding enhanced the case-finding strategy, which contributed to a decrease in the TB incidence [29]. Further, the implementation of novel public health approaches such as large-scale application of modern molecular rapid diagnostic techniques of GeneXpert, advancement of the patient care cascade, and development of an advanced information system [32] helped reduce the disease burden in last 2 decades.

The cohort effect known as the generation effect is the varying disease incidence among various generations and birth cohorts associated with specific epidemiological exposure such as wars, famine, baby boom, and disease pandemic. This intergenerational effect could impact the course of life. Our study reveals the earliest birth cohort with the highest RR of PTB and the most recent birth cohort with the lowest RR. This finding is consistent with those reported in other studies in China [5,6,33]. The difference and decrease in the intergenerational effects may be related to the stable socioeconomic status since the establishment of China in 1949. More specifically, the rapid development and continuous economic improvement accompanied by improvement in the quality of life in the past 40 years has led to a reduction in the epidemiological risk exposures for PTB incidence in birth cohorts after 1949.

Another contribution of this study was the finding of higher risks of PTB incidence in the birth cohort of 1961-1970, regardless of the highest and lowest RR cohorts. A possible explanation could be that China endured the great famine from 1959 to 1961, which led to the loss of millions of lives and other health consequences [34-36]. Yunnan, being in the southwestern region of China, has a lower economic status, insufficient food supply, and inadequate transportation in the mountainous rural areas, which led to greater severity in the famine. Biologically, malnutrition during the growth and development stages of children leads to immune system deficiency, consequently causing a higher risk of infection [37,38]. The low nutritional level could be characterized by the lower BMI of the population, which was a definite risk factor for TB incidence [38,39]. Another noteworthy finding was the sudden increasing RR for the birth cohort of 2001-2010. The speculative reason was the intergenerational effect for these 2 birth cohorts. The famine cohort was during 1961-1970, with a supposed average of 20 years for population regeneration [40]—their second-generation offspring was exactly the birth cohort of 2001-2010. Currently, the most popular pattern of bringing up children is that the older adults look after their grandchildren in China. Thus, compared to other age groups, older adults have a significantly increased general contact with those aged 5-10 years [41]. Due to the high probability of PTB incidence among the grandparents, the grandchildren are more likely to contract PTB owing to their close contact. In particular, the speculated grandparent-grandchildren transmission pattern was illustrated by the female cohort curve and expressed as grandmother-granddaughter transmission. Furthermore, the latent tuberculosis infection prevalence increased with age in China, while the risk of active PTB incidence increased synchronously, especially for the generation during the great famine. The hypothetical grandparent-grandchildren transmission pattern would cause more disease burden due to increased latent tuberculosis infection. Regarding the prevention and control of TB, especially in older adult women, the health care policy should consider not only active case-finding but also latent tuberculosis infection surveillance and prevention. Moreover, there was also possibility of nosocomial infections from long-term care facilities and health care systems rather than by community or family member transmission. To sum up, an optimized control policy should involve a comprehensive intervention and screening strategy for older adults.

The BAPC model is particularly useful for predicting future PTB burden, as it involves no parametric assumptions, and it is the only method to achieve sensible projections [24]. Our study shows the striking forecast that PTB incidence will rapidly increase from 2021 to 2030, and by the end of 2030, PTB incidence can increase to 2.28 times higher than that in 2006. The decomposition analysis presents the double PTB incidence attributed to the sequential factors of age-specific changes, population aging, and population growth. The projection and decomposition validated the previous speculation of the grandparent-grandchildren transmission pattern. The famine cohort (birth in 1961-1970) would have the highest PTB incidence in 2030, and their grandchildren cohort (birth in 2001-2010) would have a higher risk than the adjacent cohorts. Although population aging and population growth contributed to only a small portion of PTB incidence, the cohort effect (age-specific changes) can be explained as the leading cause of incidence. Another noteworthy point is the accelerated descent of the Chinese population and fast aging [42-45]. Our findings offer possible explanation for the increasing PTB incidence to be related to the dominant factors of age-specific change and population aging, but the relationship between these factors and their interactions with the projected PTB incidence should be addressed by further evidence.

Limitations

This study has some limitations. First, the inferences in this study might be limited by ecological fallacy and false logical deduction; these inferences need to be examined by further field and population researches. Second, the intervals of the group for age, period, and cohort were 5 years; although aggregated data reduce potential overdispersion, it could neglect the effect changes in a fine scale. The PTB incidence data were available only of the recent 15 years, which added uncertainty to the long-term analysis and projection. Third, the reported data might underestimate the effects due to underreporting of patients with PTB. Fourth, the uncertainty of the United Nation's world population prospects projection to future Chinese and Yunnan provincial population sizes and structures and the lack of consideration of migration could unsettle and bias the population-structure–based results of BAPC and decomposition analysis. Considering the Chinese seventh population census conducted in 2020 [17], the effect of population structure shifting should be addressed by further studies. Fifth, the forecast decomposed factors into population aging and growth and age-specific changes. Further research is needed to examine other related factors such as the socioeconomic index, latent tuberculosis infection status, herd immunity, population dietary nutrition level, and the COVID-19 pandemic. Besides, there were residual effects of bias from the APC approach despite the use of the intrinsic estimator methodology. Field investigations are needed to verify the speculated grandparent-grandchildren PTB transmission pattern.

Conclusion

Our study presents the pattern of PTB incidence in Yunnan, China, and elucidates the underlying causes for the temporal trends by quantifying the age, period, and cohort effects with population-based surveillance data and the macroepidemiological APC model. Although there are several measures and levels for PTB control in China, the PTB burden is likely to increase in the future mainly due to age-specific reasons. The endeavors to achieve the End TB vision in the next decade should include large-scale active case-finding, priority prevention policies for the high-risk older adult and younger populations, and reduction in possible grandparent-grandchildren transmission patterns.

Acknowledgments

This study was supported by the Yunnan Center for Disease Control and Prevention. We thank our colleagues from the Department of Tuberculosis Control and Prevention for their assistance in data collection. We also appreciate the local centers for disease control and prevention, community health care centers, and tuberculosis-designated hospitals for their work on disease notification and data collection. This study was supported by the Yunnan Health Training Program for High-Level Talents (grant H-2019027) and Yunnan Provincial High-Level Talent Incubator Program.

Data Availability

The data sets analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JC and LX conceived and designed this study. JC, YQ, WW, RY, LL, YY, and XY provided the source data for this study. JC prepared the software and performed the statistical analysis. JC and LX prepared the manuscript and interpreted the data. YQ, WW, RY, LL, YY, and XY assisted with the editing of the paper and provided critical comments. JC and LX revised the paper critically for important intellectual content. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The disease burden decomposition method, parameters of the fitted age-period-cohort model, and projected age-specific number of pulmonary tuberculosis incidence by the Bayesian age-period-cohort model.

[DOCX File , 26 KB - publichealth_v9i1e48015_app1.docx]

References

1. Global tuberculosis report 2022. World Health Organization. URL: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 [accessed 2023-03-01]
2. Uplekar M, Weil D, Lonnroth K, Jaramillo E, Lienhardt C, Dias HM, for WHO's Global TB Programme. WHO's new end TB strategy. Lancet 2015 May 02;385(9979):1799-1801. [doi: 10.1016/S0140-6736(15)60570-0] [Medline: 25814376]
3. Wang L, Liu J, Chin DP. Progress in tuberculosis control and the evolving public-health system in China. Lancet 2007 Feb 24;369(9562):691-696 [FREE Full text] [doi: 10.1016/S0140-6736(07)60316-X] [Medline: 17321314]
4. Cui Y, Shen H, Wang F, Wen H, Zeng Z, Wang Y, et al. A long-term trend study of tuberculosis incidence in China, India and United States 1992-2017: A joinpoint and age-period-cohort analysis. Int J Environ Res Public Health 2020 May 11;17(9):1-19 [FREE Full text] [doi: 10.3390/ijerph17093334] [Medline: 32403353]
5. Wang L, Wang W. Temporal trends in notification and mortality of tuberculosis in China, 2004-2019: A joinpoint and age-period-cohort analysis. Int J Environ Res Public Health 2021 May 24;18(11):5607 [FREE Full text] [doi: 10.3390/ijerph18115607] [Medline: 34073943]
6. Dong Z, Wang Q, Yu S, Huang F, Liu J, Yao H, et al. Age-period-cohort analysis of pulmonary tuberculosis reported incidence, China, 2006-2020. Infect Dis Poverty 2022 Jul 28;11(1):85 [FREE Full text] [doi: 10.1186/s40249-022-01009-4] [Medline: 35902982]
7. Du J, Sun H, Sun Y, Du J, Cao W, Sun S. Assessment of age, period, and cohort effects of lung cancer incidence in Hong Kong and projection up to 2030 based on changing demographics. Am J Cancer Res 2021;11(12):5902-5916 [FREE Full text] [Medline: 35018232]
8. Cheng F, Xiao J, Shao C, Huang F, Wang L, Ju Y, et al. Burden of thyroid cancer from 1990 to 2019 and projections of incidence and mortality until 2039 in China: findings from global burden of disease study. Front Endocrinol (Lausanne) 2021;12:738213 [FREE Full text] [doi: 10.3389/fendo.2021.738213] [Medline: 34690931]
9. Luo Q, O'Connell DL, Yu XQ, Kahn C, Caruana M, Pesola F, et al. Cancer incidence and mortality in Australia from 2020 to 2044 and an exploratory analysis of the potential effect of treatment delays during the COVID-19 pandemic: a statistical modelling study. Lancet Public Health 2022 Jun;7(6):e537-e548 [FREE Full text] [doi: 10.1016/S2468-2667(22)00090-1] [Medline: 35660215]
10. Liu Y, Zhang C, Wang Q, Wu K, Sun Z, Tang Z, et al. Temporal trends in the disease burden of colorectal cancer with its risk factors at the global and national level from 1990 to 2019, and projections until 2044. Clin Epidemiol 2023;15:55-71 [FREE Full text] [doi: 10.2147/CLEP.S388323] [Medline: 36659904]
11. Luo G, Zhang Y, Etxeberria J, Arnold M, Cai X, Hao Y, et al. Projections of lung cancer incidence by 2035 in 40 countries worldwide: population-based study. JMIR Public Health Surveill 2023 Feb 17;9:e43651 [FREE Full text] [doi: 10.2196/43651] [Medline: 36800235]

12. Gao D, Zou Z, Zhang W, Chen T, Cui W, Ma Y. Age-period-cohort analysis of HIV mortality in China: data from the global burden of disease study 2016. Sci Rep 2020 Apr 27;10(1):7065 [FREE Full text] [doi: 10.1038/s41598-020-63141-1] [Medline: 32341364]
13. Ren N, Li Y, Wan Z, Wang R, Zhang W, Dzakah EE, et al. Patterns of HIV or AIDS mortality among older people from 1990 to 2019 in China: age-period-cohort analysis. JMIR Public Health Surveill 2022 Nov 17;8(11):e35785 [FREE Full text] [doi: 10.2196/35785] [Medline: 36394944]
14. Jha RP, Shri N, Patel P, Dhamnetiya D, Bhattacharyya K, Singh M. Trends in the diabetes incidence and mortality in India from 1990 to 2019: a joinpoint and age-period-cohort analysis. J Diabetes Metab Disord 2021 Dec;20(2):1725-1740 [FREE Full text] [doi: 10.1007/s40200-021-00834-y] [Medline: 34900822]
15. China PCO. Tabulation on the 2000 Population Census of The People's Republic of China. Beijing, China: China Statistics Press; Jul 1, 2001.
16. China PCO. Tabulation on the 2010 Population Census of The People's Republic of China. Beijing: China Statistics Press; Jul 01, 2011.
17. China PCO. Tabulation on the 2020 Population Census of The People's Republic of China. Beijing, China: China Statistics Press; Jul 01, 2021.
18. World population prospects 2020. United Nations Department of Economic and Social Affairs. URL: https://population.un.org/wpp [accessed 2023-02-01]
19. Rosenberg PS. A new age-period-cohort model for cancer surveillance research. Stat Methods Med Res 2019;28(10-11):3363-3391. [doi: 10.1177/0962280218801121] [Medline: 30306829]
20. Smith TR, Wakefield J. A review and comparison of age-period-cohort models for cancer incidence. Statist Sci 2016 Nov 1;31(4):591-610. [doi: 10.1214/16-sts580]
21. Yang Y, Schulhofer‐Wohl S, Fu W, Land K. The intrinsic estimator for age‐period‐cohort analysis: what it is and how to use it. American Journal of Sociology 2008 May;113(6):1697-1736. [doi: 10.1086/587154]
22. Rosenberg PS, Check DP, Anderson WF. A web tool for age-period-cohort analysis of cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2296-2302 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0300] [Medline: 25146089]
23. Knoll M, Furkel J, Debus J, Abdollahi A, Karch A, Stock C. An R package for an integrated evaluation of statistical approaches to cancer incidence projection. BMC Med Res Methodol 2020 Oct 15;20(1):257 [FREE Full text] [doi: 10.1186/s12874-020-01133-5] [Medline: 33059585]
24. Riebler A, Held L. Projecting the future burden of cancer: Bayesian age-period-cohort analysis with integrated nested Laplace approximations. Biom J 2017 May;59(3):531-549. [doi: 10.1002/bimj.201500263] [Medline: 28139001]
25. Cheng X, Yang Y, Schwebel DC, Liu Z, Li L, Cheng P, et al. Population ageing and mortality during 1990-2017: A global decomposition analysis. PLoS Med 2020 Jun;17(6):e1003138 [FREE Full text] [doi: 10.1371/journal.pmed.1003138] [Medline: 32511229]
26. Cheng X, Tan L, Gao Y, Yang Y, Schwebel DC, Hu G. A new method to attribute differences in total deaths between groups to population size, age structure and age-specific mortality rate. PLoS One 2019;14(5):e0216613 [FREE Full text] [doi: 10.1371/journal.pone.0216613] [Medline: 31075117]
27. Li T, Shewade HD, Soe KT, Rainey JJ, Zhang H, Du X, et al. Under-reporting of diagnosed tuberculosis to the national surveillance system in China: an inventory study in nine counties in 2015. BMJ Open 2019 Jan 28;9(1):e021529 [FREE Full text] [doi: 10.1136/bmjopen-2018-021529] [Medline: 30696665]
28. Liu K, Xie Z, Xie B, Chen S, Zhang Y, Wang W, et al. Bridging the gap in end tuberculosis targets in the elderly population in Eastern China: observational study from 2015 to 2020. JMIR Public Health Surveill 2022 Jul 29;8(7):e39142 [FREE Full text] [doi: 10.2196/39142] [Medline: 35904857]
29. Chen J, Qiu Y, Rueda ZV, Hou J, Lu K, Chen L, et al. Role of community-based active case finding in screening tuberculosis in Yunnan province of China. Infect Dis Poverty 2019 Oct 29;8(1):92 [FREE Full text] [doi: 10.1186/s40249-019-0602-0] [Medline: 31661031]
30. Long Q, Guo L, Jiang W, Huan S, Tang S. Ending tuberculosis in China: health system challenges. Lancet Public Health 2021 Dec;6(12):e948-e953 [FREE Full text] [doi: 10.1016/S2468-2667(21)00203-6] [Medline: 34838198]
31. Lu L, Jia M, Ma Y, Yang L, Chen Z, Ho DD, et al. The changing face of HIV in China. Nature 2008 Oct 02;455(7213):609-611. [doi: 10.1038/455609a] [Medline: 18833270]
32. Chen J, Ruan Y, Wang K, Li L, Yang Y, Yang R, et al. Development and application of a multidrug-resistant tuberculosis case management system - Yunnan Province, China, 2017-2020. China CDC Wkly 2022 Sep 23;4(38):855-861 [FREE Full text] [doi: 10.46234/ccdcw2022.177] [Medline: 36284689]
33. Cheng Q, Trangucci R, Nelson KN, Fu W, Collender PA, Head JR, et al. Prenatal and early-life exposure to the Great Chinese Famine increased the risk of tuberculosis in adulthood across two generations. Proc Natl Acad Sci U S A 2020 Nov 03;117(44):27549-27555 [FREE Full text] [doi: 10.1073/pnas.2008336117] [Medline: 33077583]
34. Xu H, Li L, Zhang Z, Liu J. Is natural experiment a cure? Re-examining the long-term health effects of China's 1959-1961 famine. Soc Sci Med 2016 Jan;148:110-122 [FREE Full text] [doi: 10.1016/j.socscimed.2015.11.028] [Medline: 26692092]

35. Li C, Lumey LH. Early-life exposure to the Chinese Famine of 1959-1961 and type 2 diabetes in adulthood: a systematic review and meta-analysis. Nutrients 2022 Jul 12;14(14):2855 [FREE Full text] [doi: 10.3390/nu14142855] [Medline: 35889812]
36. Tao T, Dai L, Ma J, Li Y, Guo Z. Association between early-life exposure to the Great Chinese Famine and poor physical function later in life: a cross-sectional study. BMJ Open 2019 Jul 19;9(7):e027450 [FREE Full text] [doi: 10.1136/bmjopen-2018-027450] [Medline: 31326928]
37. Cegielski JP, McMurray DN. The relationship between malnutrition and tuberculosis: evidence from studies in humans and experimental animals. Int J Tuberc Lung Dis 2004 Mar;8(3):286-298. [Medline: 15139466]
38. Lönnroth K, Williams BG, Cegielski P, Dye C. A consistent log-linear relationship between tuberculosis incidence and body mass index. Int J Epidemiol 2010 Feb;39(1):149-155. [doi: 10.1093/ije/dyp308] [Medline: 19820104]
39. Chen J, Zha S, Hou J, Lu K, Qiu Y, Yang R, et al. Dose-response relationship between body mass index and tuberculosis in China: a population-based cohort study. BMJ Open 2022 Mar 16;12(3):e050928 [FREE Full text] [doi: 10.1136/bmjopen-2021-050928] [Medline: 35296472]
40. Yin-feng Z, Jia-wei H. The Actual and Ideal Fertility Age of Chinese Population: 1994-2012. Popul Develop 2016;22(1):11-12.
41. Zheng B, Zhu W, Pan J, Wang W. Patterns of human social contact and mask wearing in high-risk groups in China. Infect Dis Poverty 2022 Jun 18;11(1):69 [FREE Full text] [doi: 10.1186/s40249-022-00988-8] [Medline: 35717198]
42. GBD 2017 PopulationFertility Collaborators. Population and fertility by age and sex for 195 countries and territories, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1995-2051 [FREE Full text] [doi: 10.1016/S0140-6736(18)32278-5] [Medline: 30496106]
43. Chen Y, Guo F, Wang J, Cai W, Wang C, Wang K. Provincial and gridded population projection for China under shared socioeconomic pathways from 2010 to 2100. Sci Data 2020 Mar 09;7(1):83 [FREE Full text] [doi: 10.1038/s41597-020-0421-y] [Medline: 32152299]
44. Vollset SE, Goren E, Yuan C, Cao J, Smith AE, Hsiao T, et al. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. Lancet 2020 Oct 17;396(10258):1285-1306 [FREE Full text] [doi: 10.1016/S0140-6736(20)30677-2] [Medline: 32679112]
45. Luo D, Yan X, Xu R, Zhang J, Shi X, Ma J, et al. Chinese trends in adolescent marriage and fertility between 1990 and 2015: a systematic synthesis of national and subnational population data. Lancet Glob Health 2020 Jul;8(7):e954-e964 [FREE Full text] [doi: 10.1016/S2214-109X(20)30130-3] [Medline: 32562651]

Abbreviations

APC: age-period-cohort
BAPC: Bayesian age-period-cohort
BCG: Bacillus Calmette-Guérin
NNDRS: National Notifiable Disease Reporting System
PTB: pulmonary tuberculosis
RR: rate ratio
TB: tuberculosis

Edited by A Mavragani, T Sanchez; submitted 08.04.23; peer-reviewed by T Li, R Mpofu, L Huang; comments to author 11.05.23; revised version received 19.06.23; accepted 31.10.23; published 29.12.23.

Please cite as:
Chen J, Qiu Y, Wu W, Yang R, Li L, Yang Y, Yang X, Xu L
Trends and Projection of the Incidence of Active Pulmonary Tuberculosis in Southwestern China: Age-Period-Cohort Analysis
JMIR Public Health Surveill 2023;9:e48015
URL: https://publichealth.jmir.org/2023/1/e48015
doi:10.2196/48015
PMID:38157236

Original Paper

Spatiotemporal Changes in Pulmonary Tuberculosis Incidence in a Low-Epidemic Area of China in 2005-2020: Retrospective Spatiotemporal Analysis

Qi Zhang[1], MPH; Huan Ding[1], MPH; Song Gao[1], PhD; Shipeng Zhang[1], MPH; Shiya Shen[1], BSc; Xiaoyan Chen[1], BSc; Zhuping Xu[1], BSc

Department of Chronic Communicable Disease, The Affiliated Wuxi Center for Disease Control and Prevention of Nanjing Medical University, Wuxi Center for Disease Control and Prevention, Wuxi, China

Corresponding Author:
Zhuping Xu, BSc
Department of Chronic Communicable Disease
The Affiliated Wuxi Center for Disease Control and Prevention of Nanjing Medical University, Wuxi Center for Disease Control and Prevention
499 Jin Cheng Road
Wuxi, 214023
China
Phone: 86 0510 82726760
Fax: 86 0510 85727331
Email: Xzp15951589696@163.com

Abstract

Background: In China, tuberculosis (TB) is still a major public health problem, and the incidence of TB has significant spatial heterogeneity.

Objective: This study aimed to investigate the temporal trends and spatial patterns of pulmonary tuberculosis (PTB) in a low-epidemic area of eastern China, Wuxi city, from 2005 to 2020.

Methods: The data of PTB cases from 2005 to 2020 were obtained from the Tuberculosis Information Management System. The joinpoint regression model was used to identify the changes in the secular temporal trend. Kernel density analysis and hot spot analysis were used to explore the spatial distribution characteristics and clusters of the PTB incidence rate.

Results: A total of 37,592 cases were registered during 2005-2020, with an average annual incidence rate of 34.6 per 100,000 population. The population older than 60 years had the highest incidence rate of 59.0 per 100,000 population. In the study period, the incidence rate decreased from 50.4 to 23.9 per 100,000 population, with an average annual percent change of –4.9% (95% CI –6.8% to –2.9%). The incidence rate of pathogen-positive patients increased during 2017-2020, with an annual percent change of 13.4% (95% CI 4.3%-23.2%). The TB cases were mainly concentrated in the city center, and the incidence of hot spots areas gradually changed from rural areas to urban areas during the study period.

Conclusions: The PTB incidence rate in Wuxi city has been declining rapidly with the effective implementation of strategies and projects. The populated urban centers will become key areas of TB prevention and control, especially in the older population.

(JMIR Public Health Surveill 2023;9:e42425) doi:10.2196/42425

KEYWORDS

pulmonary tuberculosis; spatial analysis; temporal analysis; epidemiology; China

Introduction

Tuberculosis (TB) is a chronic communicable disease caused by the bacillus *Mycobacterium tuberculosis*, which spreads across the population via the respiratory tract [1,2]. The pathogen most typically invades the lungs (pulmonary tuberculosis [PTB]), though it can affect any organs of the body (extrapulmonary TB) [3,4]. With high infection rates, morbidity, mortality, drug resistance, and HIV coinfection, TB still remains one of the most intractable public health issues, particularly in low- and middle-income countries [5]. According to the *Global Tuberculosis Report 2020* by the World Health Organization [1], there were approximately 10.0 million (range 8.9-11.0 million) people with TB worldwide, of whom 54.6% (n=5.5

million) were from 5 countries, namely India (26.0%), Indonesia (8.5%), China (8.4%), Philippines (6.0%), and Pakistan (5.7%).

With the second-largest TB burden, the Chinese government has taken effective national prevention and control strategies to combat TB. Since 1991, a pilot TB control strategy, the directly observed treatment and short-course (DOTS) strategy, was implemented in several provinces, and the government expanded the DOTS strategy nationwide by 2005 [6]. Moreover, the government revised the Law of the People’s Republic of China on the Prevention and Treatment of Infectious Diseases, and the Chinese Center for Disease Control and Prevention developed the nationwide Notifiable Infectious Disease Reporting Information System, which is the largest internet-based disease reporting system after the severe acute respiratory syndrome outbreak [7]. With the implementation of various TB prevention and control measures, the morbidity and mortality of TB in China have decreased dramatically [8,9]. The incidence of PTB and smear-positive TB in China dropped from 611 and 170 per 100,000 population in 1990 to 442 and 59 per 100,000 population in 2010, respectively [6]. However, due to differences in the population, climate, and socioeconomic factors, the prevalence of PTB in China has significant spatial heterogeneity [8,10,11]. According to the fifth national TB epidemiological survey in 2010, the prevalence of PTB in the eastern, central, and western regions of China was 291, 463, and 695 per 100,000 population, respectively [12]. In addition, the prevalence was significantly higher in rural areas (569 per 100,000 population) than in urban areas (307 per 100,000 population). Another study [13] based on the reported PTB cases in China showed that the reported incidence of PTB in 2019 was 55.6 per 100,000 population in China, and the incidence in Jiangsu Province was 31.3 per 100,000 population, which was one of the lowest incidences nationwide and much lower than that in the Tibet Autonomous Region (182.4 per 100,000 population). In previous studies, researchers have analyzed the spatial epidemiological characteristics of PTB, which are of great significance for the effective allocation of limited health resources [14-17]. However, previous studies [15-17] mainly focused on areas with high PTB incidence in China; however, they seriously ignored the evolution of epidemiological characteristics of PTB in low-epidemic areas. Nie et al [18] analyzed the spatial distribution characteristics of PTB cases in Hefei City from 2009 to 2020 and reported that the incidence of PTB in Hefei city decreased from 57.96 per 100,000 population in 2009 to 31.04 per 100,000 population in 2020 with the high-high incidence cluster gradually decreasing in rural areas. Through targeted prevention and control strategies of PTB conducted in low-epidemic areas, the incidence of PTB had been reduced from high to low incidence. The epidemiological characteristics and prevention and control strategies of PTB in low-epidemic areas can provide a reference for the formulation in high-epidemic areas in the future to accelerate the end of PTB epidemic.

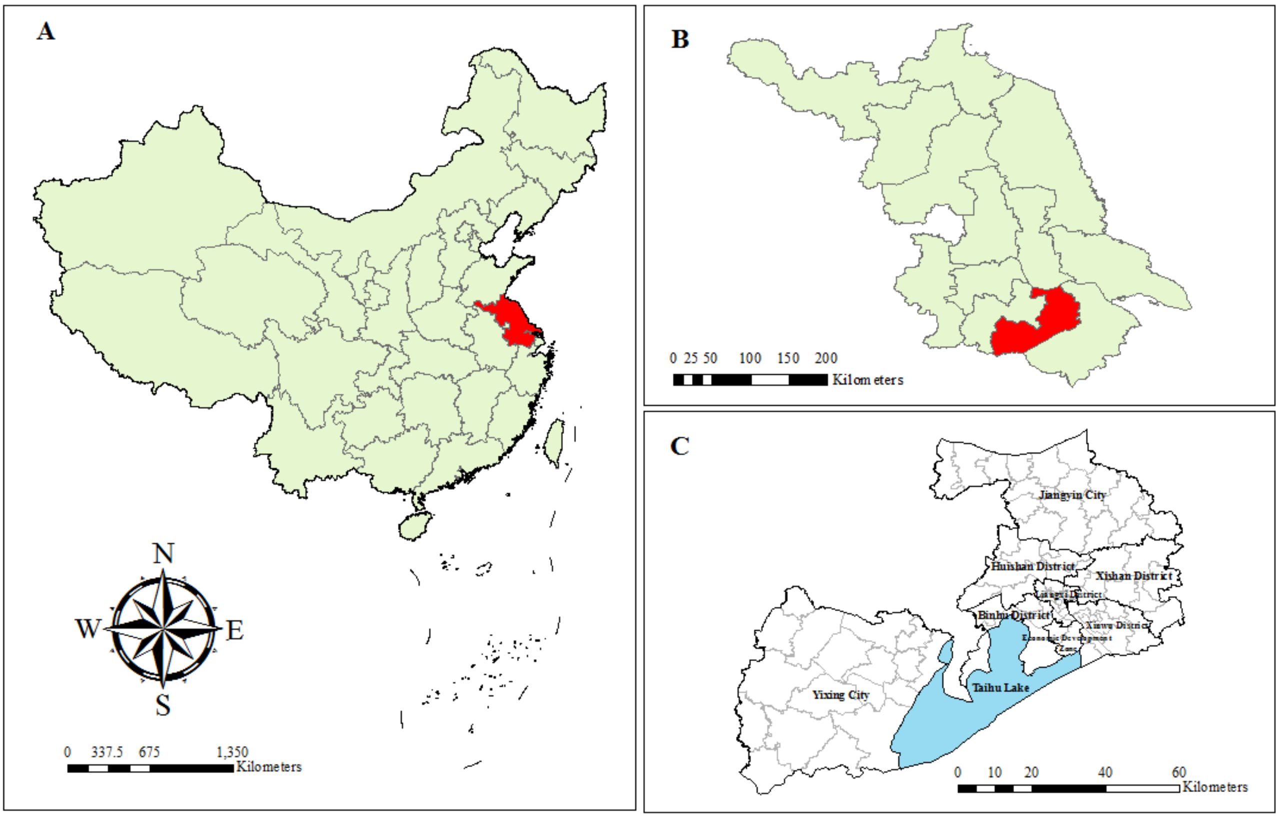
This study aimed to investigate the temporal trends and spatial patterns of PTB incidence in a low-epidemic are of China, Wuxi city in Jiangsu Province, from 2005 to 2020 through epidemic characteristics analysis and spatial analysis at the subdistrict level.

Methods

Study Area

Wuxi city, with monsoon and maritime climate, is located in southern Jiangsu Province on the eastern coast of China (Figure 1). It covers an area of 4627.5 km^2 and is composed of 6 districts and 2 county-level cities (Jiangyin city and Yixing city) with a total population of 7.5 million persons. It has a highly developed economy and emerged among the top-ranking cities in China in 2020 with a per capita gross domestic product exceeding US $20,000 [19].

Figure 1. Location of the study area. The location of (A) Jiangsu Province in China, (B) Wuxi city in Jiangsu Province, and (C) 6 districts and 2 county-level cities in Wuxi city.



Data Collection

The Tuberculosis Information Management System is one of the components of the Notifiable Infectious Disease Reporting Information System. Since 2004, the system was implemented in all cities of China, and every PTB case has been required to report to this system in real time. In this study, the data of PTB cases from 2005 to 2020 in Wuxi city were obtained from the Tuberculosis Information Management System. Information about PTB cases included age, gender, address, as well as medical information such as drug resistance, classification (primary treatment or retreatment), and outcome. Demographic data were collected from the Statistical Yearbook of Wuxi.

Statistical Analysis

Descriptive statistics were used to illustrate the characteristics of the population distribution. R software (version 4.0.3; The R Foundation) was used to investigate the level of statistical significance.

Temporal Analysis

The joinpoint regression model [20] was used to identify the changes in the secular temporal trend of PTB in Wuxi city from 2005 to 2020. In the model, log-transformed PTB incidence rates and notification years were identified as the response variable and independent variable, respectively. The annual percent change, the average annual percent change, and the corresponding 95% CIs were obtained through joinpoint regression analysis. The best-fitted model was identified using the Bayesian Information Criterion (BIC). The equation for computing the BIC for a k-joinpoint model is as follows:

Where *SSE* is the sum of squared errors of the k-joinpoint regression model, *Parm*(*k*)=2*(k+1) is the number of parameters of the k-joinpoint model, and *Obs* is the number of observations. The k-joinpoint model with the minimum value of BIC(k) is selected as the final model. The joinpoint regression model was developed by the Joinpoint Regression Program (version 4.9.0.0; March 2021) from the Surveillance Research Program of the US National Cancer Institute.

Spatial Analysis

Kernel density analysis [21] was used to explore the spatial distribution characteristics of PTB cases. The basic idea of Kernel density analysis is that geographical events have a high probability of occurrence in areas with high spatial point density and low probability in areas with low spatial point density, which can be used to explore hot spots of diseases. The density value was calculated on the basis of the following formulas:

Where $f_{(x,y)}$ represents the density value of the spatial coordinate position (x,y), n represents the number of PTB cases within the distance scale, h represents the bandwidth, k is the density function, $d_{i(x,y)}$ represents the distance from a PTB case i to (x,y). ArcGIS software (version 10.4; ESRI) was used for Kernel density analysis of the PTB spatial distribution characteristics in Wuxi city.

Cluster Analysis

Hot spot analysis (Getis-Ord Gi*) [22] was used to identify the clusters of PTB incidence rate in Wuxi city. First, the addresses were aggregated to the subdistricts, and the incidence rate of PTB was calculated. Then, the Getis-Ord Gi* statistic was analyzed for PTB incidence of subdistricts at different times. The Getis-Ord local statistic is given as follows:

where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features.

The resultant z scores and P values showed that high or low PTB incidence rates cluster spatially. A high z score with a small P value for the incidence indicated a spatial clustering of high incidence rates. A low negative z score with a small P value indicated a spatial clustering of low incidence rates. A z score near 0 indicated no apparent spatial clustering of PTB incidence rates. The Getis-Ord Gi* statistic was also analyzed using ArcGIS software (version 10.4; ESRI), and 999 permutations were conducted to ensure the statistical significance at 95% CIs.

Results

Demographic Characteristics

A total of 37,592 cases of PTB were registered during 2005 and 2020 in Wuxi city, with an average annual incidence rate of 34.6 per 100,000 population. Among the reported PTB cases, pathogen-positive PTB accounted for 50.5% (18,987/37,592) of cases, and the incidence was 17.5 per 100,000 population. As shown in Table 1, the incidence in men was more than twice that in women (χ^2_1= 6414.4, P<.001), with rates of 48.6 and 20.0 per 100,000 population, respectively. Although in the past 15 years, the largest number of notified PTB cases occurred in the age group of 20-39 years (14,754/37,592, 39.2%), the population older than 60 years had the highest incidence rate of 59.0 per 100,000 population, which was 1.4 times that of the 20-39–year age group.

Table 1. The number and incidence of pulmonary tuberculosis with different demographic characteristics in Wuxi city, 2005-2020.

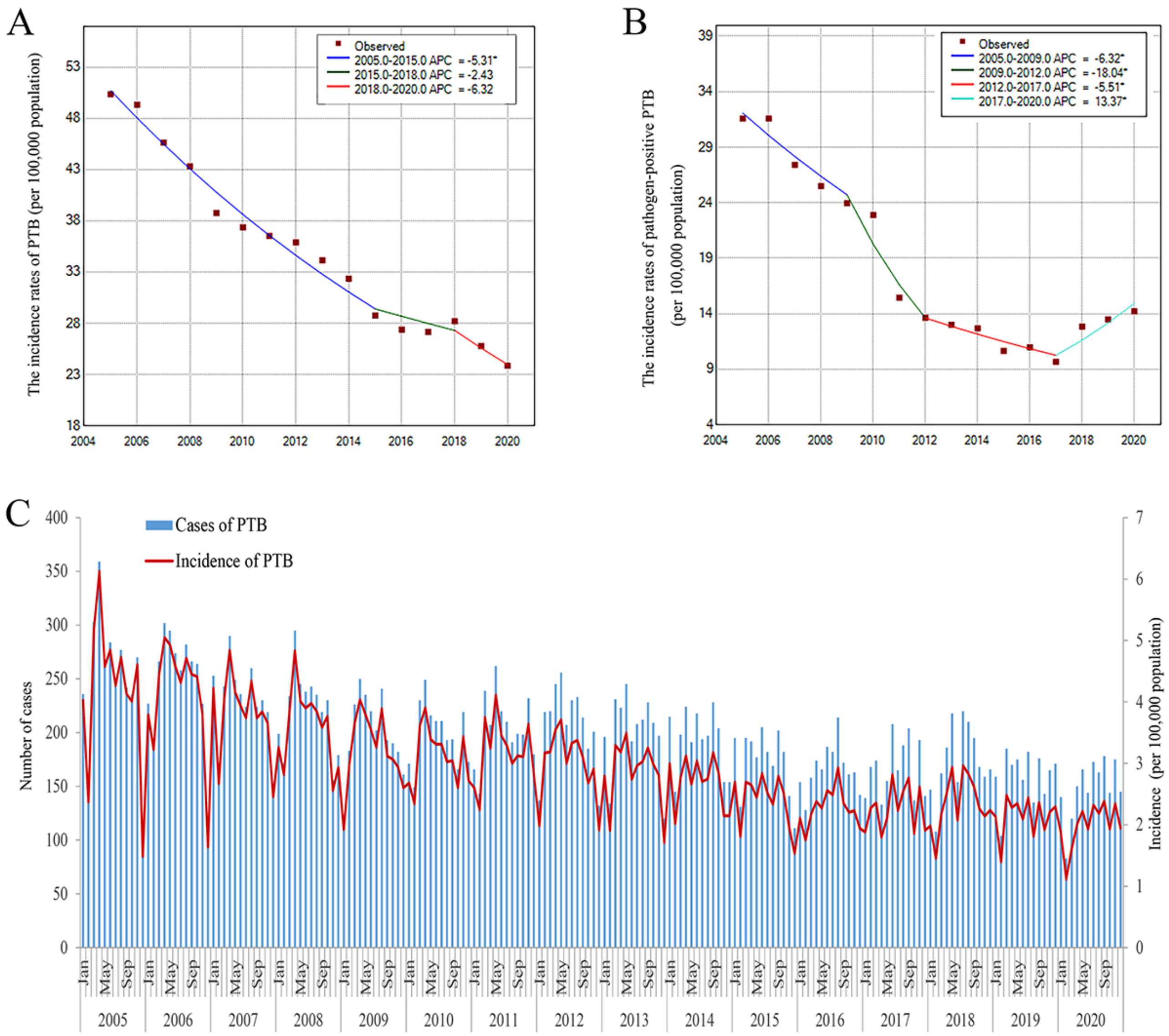
Characteristics	Males		Females		Total	
	Cases, n	Incidence[a]	Cases, n	Incidence	Cases, n	Incidence
Age group (years)						
0-19	1213	11.4	73	0.8	1946	9.6
20-39	9787	54.0	4967	29.4	14,754	42.1
40-59	7974	43.5	2578	14.8	10,552	29.5
≥60	8061	94.1	2279	25.4	10,340	59.0
Pathogenic classification						
Positive	13,865	24.9	5122	9.7	18,987	17.5
Negative	12,665	22.8	5169	9.8	17,834	16.4
Unknown	505	0.9	266	0.5	771	0.7
Total	27,035	48.6	10,557	20.0	37,592	34.6

[a] Average annual incidence rate per 100,000 population.

Temporal Distribution

As shown in Figure 2A, the incidence rate of PTB in Wuxi city decreased from 50.4 per 100,000 population in 2005 to 23.9 per 100,000 population in 2020, with an average annual percent change of –4.9% (95% CI –6.8% to –2.9%). PTB incidence, with the implementation of the DOTS strategy and the Floating Population TB Control Pilot Project of the Global Fund TB Control Program in Wuxi, drastically declined during 2005 and 2015 with an annual percent change of –5.3% (95% CI –6.0% to –4.6%). Subsequently, PTB incidence decreased slowly by –2.4% every year. However, since 2018, the incidence in Wuxi has once again entered a period of rapid decline, with an annual percent change of –6.4%. The pathogenic positive PTB incidence showed different temporal patterns in Figure 2B. The pathogenic positive PTB incidence decreased from 31.6 per 100,000 population in 2005 to 9.7 per 100,000 population in 2017. However, there was an ascending trend in the incidence of pathogen-positive PTB from 2017 to 2020, with an annual percent change of 13.4% (95% CI 4.3%-23.2%). The seasonal trend of PTB incidence in Wuxi city gradually changed from a single peak to a bimodal pattern, resulting in no obvious incidence peak. But, a trough of PTB incidence rates was observed during the entire study period (Figure 2C).

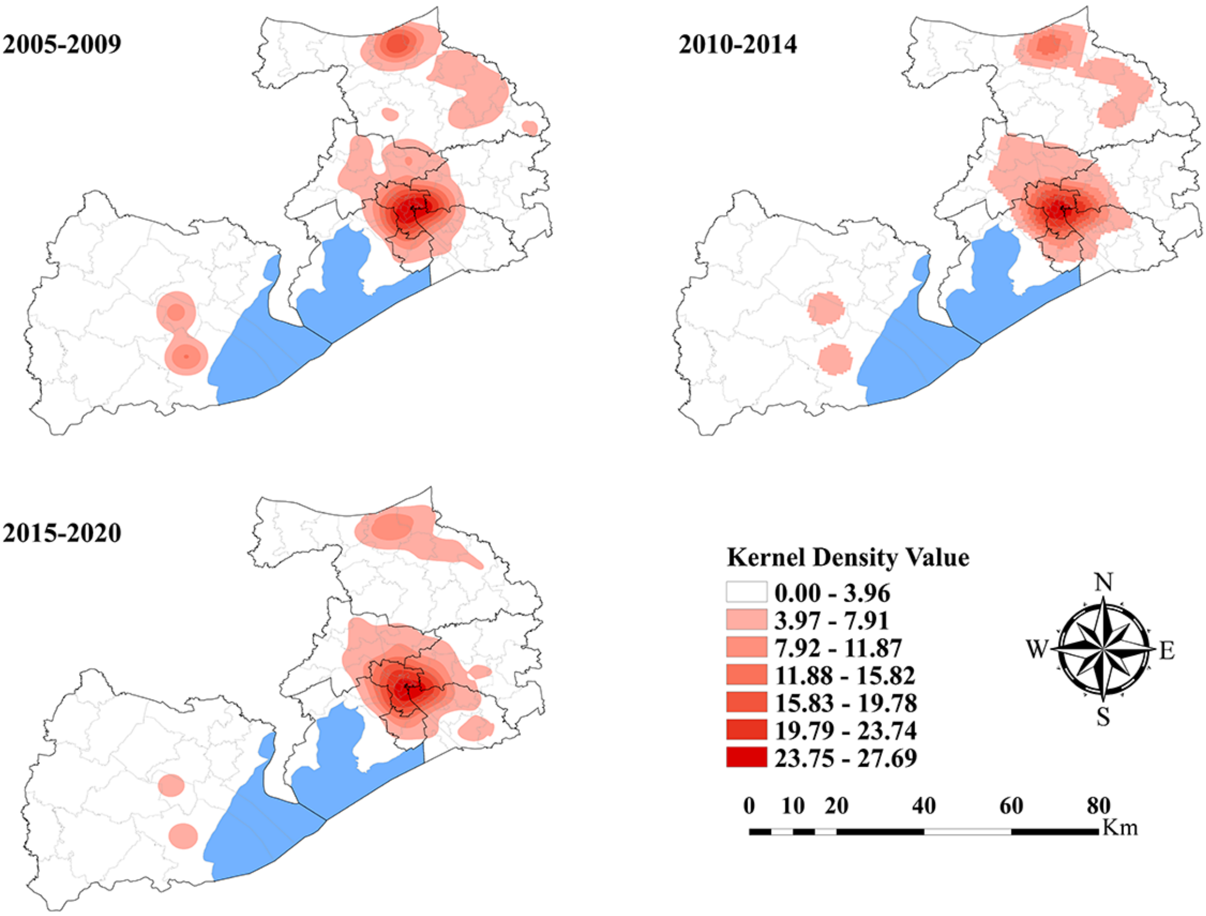
Figure 2. Temporal trend of pulmonary tuberculosis (PTB) and pathogen-positive PTB in Wuxi city, 2005-2020.



Spatial Distribution

The spatial distribution of PTB cases in Wuxi city at the county level during 2005-2020 is shown in Figure 3. The cases were mainly concentrated in the city center, namely Liangxi District and the surrounding subdistricts or towns. The same spatial characteristic also existed in 2 county-level cities, Jiangyin city and Yixing city, where PTB cases were concentrated in downtown areas. On comparing the Kernel density of cases in 2005-2009, 2010-2014, and 2015-2020, the high-PTB-density areas gradually shrank, and the Kernel density value also decreased.

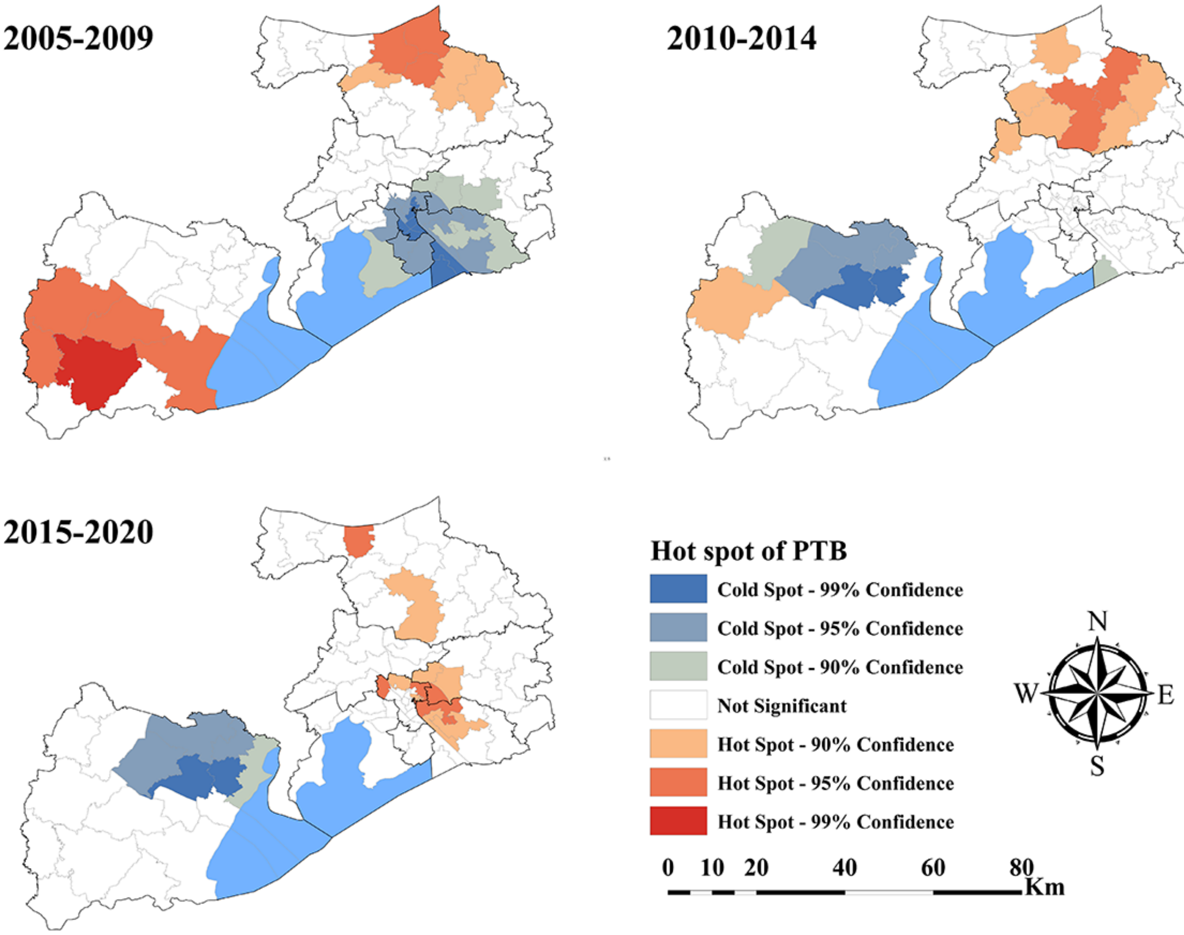
Figure 3. The spatial distribution of pulmonary tuberculosis cases in Wuxi city at the county level, 2005-2020.



Cluster Analysis

Figure 4 shows the results of PTB incidence cluster analysis of the Local Getis Gi* statistic. In 2005-2009, we found that the significant hot spots were in the northeast of Jiangyin city and southwest of Yixing city. Significant cold spots were located in Liangxi District, Economic Development Zone, and the surrounding subdistricts or towns. According to our results in 2010-2014, the significant hot spots were mainly located in the eastern and southern regions of Jiangyin city. However, the significant cold spots were mostly in northern Yixing city. During 2015-2020, the location of the PTB incidence cold spots had not changed much and was still in Yixing city. However, PTB incidence hot spots were mainly detected in the Wuxi city center and 2 subdistricts of Jiangyin city.

Figure 4. The spatial cluster of the pulmonary tuberculosis (PTB) incidence rate in Wuxi city at county level, 2005-2020.



Discussion

Principal Findings

During the 16-year period of 2005 to 2020, the PTB notification rates of Wuxi city decreased from 50.4 to 23.9 per 100,000 population at an average rate of –4.9% per year. The decline was higher than that in China (–3.8%) during the same period [8]. In addition to the implementation of national measures such as the DOTS strategy, national TB prevention strategies, control plans, and government commitments, 3 factors could have contributed to the difference. First, the lower latent infection rate of TB indicates a lower TB incidence. According to the results of a population-based, multicenter, prospective cohort study by Gao et al [23], the prevalence of latent TB infection in the eastern China plains was 13.5%, which was lower than that in the central and western regions of China. Further, the well-developed medical conditions and convenient access to medical care in the eastern plain regions have further halted PTB transmission [19]. The second factor was the implementation of the Floating Population Tuberculosis Control Project from 2007 to 2011, which was funded by the Fifth Round-China Global Fund TB program. The effects of the program were consistent with studies conducted in other cities [24,25]. By providing free inspection and grants, supervising the entire treatment process, the cure rate among patients treated for PTB in the floating population increased, and the transmission of PTB in the population reduced. The last factor was introducing and promoting the use of new technology in the diagnosis of TB. From 2018 to 2019, loop-mediated isothermal amplification was used free of charge to diagnose PTB in all TB-designated (specialized) hospitals in Wuxi city, mainly for further testing of sputum smear–negative specimens. However, the TB-designated (specialized) hospitals in Wuxi provide free Xpert MTB/RIF molecular tests for all individuals suspected with TB since 2020. With the application of molecular techniques, more pathogen-positive PTB cases were promptly detected and treated. It was essential for reducing the contagiousness of pathogen-positive PTB cases and controlling PTB transmission. Thus, the incidence of pathogenic positive PTB increased from 9.7 per 100,000 population in 2017 to 14.2 per 100,000 population in 2020, with an annual percent change of 13.4%; nonetheless, the overall incidence of PTB decreased at a rate of 6.3% per year in Wuxi city.

During the study period, the seasonal pattern was gradually unobvious—this may be explained by the lower incidence rate of PTB in Wuxi city. However, similar to other studies in China [9,22,26], our results show an incidence trough in February, especially in 2020. The reason for the decreasing rate is that Chinese people often prepare for the Lunar New Year in February, and citizens usually try to avoid seeking medical advice immediately. This is one of the reasons that an epidemic peak occurred in March and April. But in 2020, to control and prevent the outbreak of COVID-19, the Wuxi government also implemented a series of public interventions including community or home isolation, restriction of public traffic, health care staff reassignment [27-29], etc. It had seriously affected notification rates. In February 2020, the incidence rate of PTB in Wuxi city dropped by 35.3% compared to the average rate of the same period in 2017-2019.

According to our study, the population older than 60 years was the main target group of PTB with the highest incidence rate of 59.0 per 100,000 population, especially in males (94.1 per 100,000 populations). This finding was similar to those of other studies [14,16,22]. On the one hand, the older population harbors a high burden of the prevalence of latent TB infection, and on the other hand, the attenuation of immunity and underlying disease result in higher susceptibility to infection and reinfection in the older population [23,30,31]. With China's population aging rapidly, the prevention and control of TB in the older population will be a major challenge to eradicate TB.

Notably, spatial heterogeneity of PTB was observed in Wuxi city during 2005 and 2020. The results from the maps of Kernel density estimation analysis showed that PTB cases were mainly concentrated in the city center. This could be explained by the uneven distribution of the population in Wuxi city. With economic and social factors, the population is usually concentrated in the city center and the surrounding subdistricts or towns. Therefore, these densely populated areas were prone to the spread of PTB, which led to the accumulation of cases. According to PTB incidence cluster analysis, the hot spots of high-PTB-incidence areas gradually shifted from the suburbs to the city center of Wuxi city, which was similar to the spatial characteristics of PTB from 2009 to 2020 in Hefei city [18]. The results indicate that PTB control measures in the past decade have been very effective in rural areas. However, it is an urgent need for new PTB control strategies or projects to prevent PTB from spreading throughout the population in densely populated urban areas.

Limitations

This study has several limitations. First, because of death or treatment in other areas, some cases were not registered in the official system. Second, the incidence rate of PTB in this study was based on the resident population, so the incidence may have been underestimated in some areas with a large floating population. Those people were not registered in Wuxi city owing to their return to their registered permanent residence to begin PTB treatment. Third, we only conducted analyses to explore the spatial and temporal characteristics of PTB. Further studies should be carried out to detect the risk factors of PTB in low-epidemic areas in China.

Conclusions

This study identified temporal trends and spatial distribution of PTB incidence at the subdistrict level in a low-epidemic area of China, Wuxi city, from 2005 to 2020. The PTB incidence rate in Wuxi city has been declining rapidly with effective implementation of strategies and projects. The populated urban centers are expected to become key areas of PTB prevention and control, especially for the older population.

Acknowledgments

This work was supported by the Wuxi Development Medical Disciplines (FZXK2021010) and Wuxi Science and Technology Development Medical and Health Guiding Project.

Data Availability

The data sets used and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

QZ and ZX designed the study. HD and SS extracted the data and constructed the database. QZ and SZ analyzed the data. QZ drafted the manuscript. SG, XC, and ZX made critical revisions to the manuscript.

Conflicts of Interest

Not applicable.

References

1. Global tuberculosis report 2020. World Health Organization. 2020. URL: https://www.who.int/publications-detail-redirect/9789240013131 [accessed 2021-10-13]
2. Dheda K, Barry CE, Maartens G. Tuberculosis. Lancet 2016 Mar;387(10024):1211-1226. [doi: 10.1016/s0140-6736(15)00151-8]
3. Cruz-Knight W, Blake-Gumbs L. Tuberculosis: an overview. Prim Care 2013 Sep;40(3):743-756. [doi: 10.1016/j.pop.2013.06.003] [Medline: 23958367]
4. Golden MP, Vikram HR. Extrapulmonary tuberculosis: an overview. Am Fam Physician 2005 Nov 01;72(9):1761-1768 [FREE Full text] [Medline: 16300038]
5. Espindola A, Varughese M, Laskowski M, Shoukat A, Heffernan J, Moghadas S. Strategies for halting the rise of multidrug resistant TB epidemics: assessing the effect of early case detection and isolation. Int Health 2017 Mar 01;9(2):80-90. [doi: 10.1093/inthealth/ihw059] [Medline: 28338827]

6. Wang L, Zhang H, Ruan Y, Chin DP, Xia Y, Cheng S, et al. Tuberculosis prevalence in China, 1990-2010; a longitudinal analysis of national survey data. Lancet 2014 Jun 14;383(9934):2057-2064. [doi: 10.1016/S0140-6736(13)62639-2] [Medline: 24650955]
7. Wang L, Liu J, Chin DP. Progress in tuberculosis control and the evolving public-health system in China. Lancet 2007 Feb;369(9562):691-696. [doi: 10.1016/s0140-6736(07)60316-x]
8. Wang Y, Zhu W, Li T, Chen W, Wang W. Changes in newly notified cases and control of tuberculosis in China: time-series analysis of surveillance data. Infect Dis Poverty 2021 Feb 24;10(1):16 [FREE Full text] [doi: 10.1186/s40249-021-00806-7] [Medline: 33627191]
9. Guo Z, Xiao D, Wang X, Wang Y, Yan T. Epidemiological characteristics of pulmonary tuberculosis in mainland China from 2004 to 2015: a model-based analysis. BMC Public Health 2019 Feb 21;19(1):219 [FREE Full text] [doi: 10.1186/s12889-019-6544-4] [Medline: 30791954]
10. Li X, Wang L, Zhang H, Jiang S, Fang Q, Chen J, et al. Spatial variations of pulmonary tuberculosis prevalence co-impacted by socio-economic and geographic factors in People's Republic of China, 2010. BMC Public Health 2014 Mar 17;14(1):257 [FREE Full text] [doi: 10.1186/1471-2458-14-257] [Medline: 24629032]
11. Sun W, Gong J, Zhou J, Zhao Y, Tan J, Ibrahim A, et al. A spatial, social and environmental study of tuberculosis in China using statistical and GIS technology. Int J Environ Res Public Health 2015 Jan 27;12(2):1425-1448 [FREE Full text] [doi: 10.3390/ijerph120201425] [Medline: 25633032]
12. Zhu W, Wang Y, Li T, Chen W, Wang W. Gap to End-TB targets in eastern China: a joinpoint analysis from population-based notification data in Zhejiang Province, China, 2005-2018. Int J Infect Dis 2021 Mar;104:407-414 [FREE Full text] [doi: 10.1016/j.ijid.2021.01.007] [Medline: 33434670]
13. Wang Q, Li T, Du X, Ni N, Zhao YL, Zhang H. The analysis of national tuberculosis reported incidence and mortality, 2015—2019. Chin J Antituberc 2021;43(02):107-112. [doi: 10.3969/j.issn.1000-6621.2021.02.002]
14. Huang L, Li X, Abe EM, Xu L, Ruan Y, Cao C, et al. Spatial-temporal analysis of pulmonary tuberculosis in the northeast of the Yunnan province, People's Republic of China. Infect Dis Poverty 2017 Mar 24;6(1):53 [FREE Full text] [doi: 10.1186/s40249-017-0268-4] [Medline: 28335803]
15. Wang X, Yin S, Li Y, Wang W, Du M, Guo W, et al. Spatiotemporal epidemiology of, and factors associated with, the tuberculosis prevalence in northern China, 2010-2014. BMC Infect Dis 2019 Apr 30;19(1):365 [FREE Full text] [doi: 10.1186/s12879-019-3910-x] [Medline: 31039734]
16. Yu Y, Wu B, Wu C, Wang Q, Hu D, Chen W. Spatial-temporal analysis of tuberculosis in Chongqing, China 2011-2018. BMC Infect Dis 2020 Jul 22;20(1):531 [FREE Full text] [doi: 10.1186/s12879-020-05249-3] [Medline: 32698763]
17. Alene KA, Xu Z, Bai L, Yi H, Tan Y, Gray D, et al. Spatial clustering of drug-resistant tuberculosis in Hunan province, China: an ecological study. BMJ Open 2021 Apr 01;11(4):e043685 [FREE Full text] [doi: 10.1136/bmjopen-2020-043685] [Medline: 33795303]
18. Nie T, Chen W, Zhang J. Analysis of spatial distribution characteristics of pulmonary tuberculosis cases in Hefei City, 2009-2020. Chin J Antituberc 2022;44(9):947-953. [doi: 10.19982/j.issn.1000-6621.20220159]
19. United Nations Department of Economic and Social Affairs. World Statistics Pocketbook 2021. New York, NY: United Nations; 2021.
20. Kim H. Kim H-J, Fay MP, Feuer EJ, Midthune DN, ‘ Permutation tests for joinpoint regression with applications to cancer rates’. Statistics in Medicine 2000 19:335–351. Statist Med 2001 Feb 28;20(4):655-655. [doi: 10.1002/sim.811]
21. Diggle PJ. Statistical Analysis of Spatial Point Patterns (2nd edition). London: Hodder Education Publishers; 2003.
22. Ge E, Zhang X, Wang X, Wei X. Spatial and temporal analysis of tuberculosis in Zhejiang Province, China, 2009-2012. Infect Dis Poverty 2016 Feb 23;5(1):11 [FREE Full text] [doi: 10.1186/s40249-016-0104-2] [Medline: 26906041]
23. Gao L, Lu W, Bai L, Wang X, Xu J, Catanzaro A, LATENTTB-NSTM study team. Latent tuberculosis infection in rural China: baseline results of a population-based, multicentre, prospective cohort study. Lancet Infect Dis 2015 Mar;15(3):310-319. [doi: 10.1016/S1473-3099(14)71085-0] [Medline: 25681063]
24. Li X, Zhang H, Jiang S, Wang J, Liu X, Li W, et al. Active pulmonary tuberculosis case detection and treatment among floating population in China: an effective pilot. J Immigr Minor Health 2010 Dec 11;12(6):811-815. [doi: 10.1007/s10903-010-9336-6] [Medline: 20221695]
25. Zhang D, Pang X, Wang X, Chen S, Qi W, Li S. Evaluation of the effectiveness of tuberculosis control project among the floating population. Chinese Journal of Antituberculosis 2012;34(4):237-240.
26. Mao Q, Zeng C, Zheng D, Yang Y. Analysis on spatial-temporal distribution characteristics of smear positive pulmonary tuberculosis in China, 2004-2015. Int J Infect Dis 2019 Mar;80S:S36-S44 [FREE Full text] [doi: 10.1016/j.ijid.2019.02.038] [Medline: 30825654]
27. Liu Q, Lu P, Shen Y, Li C, Wang J, Zhu L, et al. Collateral impact of the coronavirus disease 2019 (COVID-19) pandemic on tuberculosis control in Jiangsu Province, China. Clin Infect Dis 2021 Aug 02;73(3):542-544 [FREE Full text] [doi: 10.1093/cid/ciaa1289] [Medline: 32857838]
28. Hogan AB, Jewell BL, Sherrard-Smith E, Vesga JF, Watson OJ, Whittaker C, et al. Potential impact of the COVID-19 pandemic on HIV, tuberculosis, and malaria in low-income and middle-income countries: a modelling study. Lancet Glob Health 2020 Sep;8(9):e1132-e1141. [doi: 10.1016/s2214-109x(20)30288-6]

29. Xu C, Li T, Hu D, Zhang H, Zhao Y, Liu J. Predicted impact of the COVID-19 responses on deaths of tuberculosis - China, 2020. China CDC Wkly 2021 Jan 08;3(2):21-24 [FREE Full text] [doi: 10.46234/ccdcw2021.004] [Medline: 34594898]
30. Gao L, Bai L, Liu J, Lu W, Wang X, Li X, LATENTTB-NSTM study team. Annual risk of tuberculosis infection in rural China: a population-based prospective study. Eur Respir J 2016 Jul 26;48(1):168-178 [FREE Full text] [doi: 10.1183/13993003.00235-2016] [Medline: 27230438]
31. Lu P, Liu Q, Zhou Y, Martinez L, Kong W, Ding X, et al. Predictors of discordant tuberculin skin test and QuantiFERON-TB Gold In-tube results in Eastern China: a population-based, cohort study. Clin Infect Dis 2021 Jun 01;72(11):2006-2015. [doi: 10.1093/cid/ciaa519] [Medline: 32369577]

Abbreviations

BIC: Bayesian Information Criterion
DOTS: directly observed treatment and short-course
PTB: pulmonary tuberculosis
TB: tuberculosis

Edited by Y Khader; submitted 03.09.22; peer-reviewed by A Waruru, V Dhikav; comments to author 22.10.22; revised version received 14.01.23; accepted 18.01.23; published 08.03.23.

Please cite as:
Zhang Q, Ding H, Gao S, Zhang S, Shen S, Chen X, Xu Z
Spatiotemporal Changes in Pulmonary Tuberculosis Incidence in a Low-Epidemic Area of China in 2005-2020: Retrospective Spatiotemporal Analysis
JMIR Public Health Surveill 2023;9:e42425
URL: https://publichealth.jmir.org/2023/1/e42425
doi:10.2196/42425
PMID:36884278

Original Paper

The Spatiotemporal Pattern and Its Determinants of Hemorrhagic Fever With Renal Syndrome in Northeastern China: Spatiotemporal Analysis

Yanding Wang[1,2*], MPH; Xianyu Wei[3*], MPH; Ruizhong Jia[2], BA; XingYu Peng[1], MPH; Xiushan Zhang[2], MPH; Meitao Yang[1,2], MPH; Zhiqiang Li[1,2], MPH; Jinpeng Guo[2], PhD; Yong Chen[2], PhD; Wenwu Yin[4], MPH; Wenyi Zhang[1,2,3], PhD; Yong Wang[1,2,3], PhD

[1]School of Public Health, China Medical University, Shenyang, China
[2]Chinese PLA Center for Disease Control and Prevention, Beijing, China
[3]School of Public Health, Anhui Medical University, Hefei, China
[4]Chinese Center for Disease Control and Prevention, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Yong Wang, PhD
Chinese PLA Center for Disease Control and Prevention
20 East street, Fengtai District
Beijing
China
Phone: 86 13693013596
Email: ywang7508@sina.com

Abstract

Background: Hemorrhagic fever with renal syndrome (HFRS) is a significant zoonotic disease mainly transmitted by rodents. However, the determinants of its spatiotemporal patterns in Northeast China remain unclear.

Objective: This study aimed to investigate the spatiotemporal dynamics and epidemiological characteristics of HFRS and detect the meteorological effect of the HFRS epidemic in Northeastern China.

Methods: The HFRS cases of Northeastern China were collected from the Chinese Center for Disease Control and Prevention, and meteorological data were collected from the National Basic Geographic Information Center. Times series analyses, wavelet analysis, Geodetector model, and SARIMA model were performed to identify the epidemiological characteristics, periodical fluctuation, and meteorological effect of HFRS in Northeastern China.

Results: A total of 52,655 HFRS cases were reported in Northeastern China from 2006 to 2020, and most patients with HFRS (n=36,558, 69.43%) were aged between 30-59 years. HFRS occurred most frequently in June and November and had a significant 4- to 6-month periodicity. The explanatory power of the meteorological factors to HFRS varies from $0.15 \leq q \leq 0.01$. In Heilongjiang province, mean temperature with a 4-month lag, mean ground temperature with a 4-month lag, and mean pressure with a 5-month lag had the most explanatory power on HFRS. In Liaoning province, mean temperature with a 1-month lag, mean ground temperature with a 1-month lag, and mean wind speed with a 4-month lag were found to have an effect on HFRS, but in Jilin province, the most important meteorological factors for HFRS were precipitation with a 6-month lag and maximum evaporation with a 5-month lag. The interaction analysis of meteorological factors mostly showed nonlinear enhancement. The SARIMA model predicted that 8,343 cases of HFRS are expected to occur in Northeastern China.

Conclusions: HFRS showed significant inequality in epidemic and meteorological effects in Northeastern China, and eastern prefecture-level cities presented a high risk of epidemic. This study quantifies the hysteresis effects of different meteorological factors and prompts us to focus on the influence of ground temperature and precipitation on HFRS transmission in future studies, which could assist local health authorities in developing HFRS-climate surveillance, prevention, and control strategies targeting high-risk populations in China.

(JMIR Public Health Surveill 2023;9:e42673) doi:10.2196/42673

KEYWORDS

HFRS; climate change; Northeastern China; spatiotemporal dynamic; Geodetector

Introduction

Hemorrhagic fever with renal syndrome (HFRS) is a significant zoonotic disease caused by hantavirus carried by rodent animals, such as *Apodemus agrarius* and *Rattus norvegicus* [1-3]. Hantavirus often leads to different clinical symptoms such as fever; headache; circulatory collapse with hypotension; gastrointestinal symptoms; and severe life-threatening damage, including bleeding or acute kidney injury [4-7]. Moreover, the latest research noted a 73% increased risk of lymphoma among patients with HFRS [8].

Currently, HFRS is found in more than 30 countries [9-13], such as Russia, the United States, China, Japan, and other countries, which poses a severe threat to human health and economic development. In China, the prevention and control of HFRS have been a prominent public health concern. It is estimated that about 100,000 cases of HFRS occur each year, with China being the most prevalent country [14-16], accounting for 70% to 90% of all HFRS cases and more than 200,000 cases between 2004 and 2019.

According to previous studies, Northeastern China is always the highest risk region of the HFRS epidemic in China [17-19], including Heilongjiang, Liaoning, and Jilin provinces. Since HFRS was first detected in Northeastern China in 1931, the epidemic area has gradually expanded, and cases have been reported in 31 provinces in China to date [16,20]. Although previous studies have mainly analyzed some provinces or specific cities with HFRS outbreaks in mainland China, there is also a lack of systematic evidence of the dynamics of HFRS occurrence in Northeastern China, which has the highest HFRS incidence in recent years.

In addition, the risk of HFRS is closely affected by climate factors [18,21], with temperature, wind speed, and humidity being essential drivers. Geodetector is an emerging technique in recent years and has become a reliable tool for detecting and exploiting spatial heterogeneity and determinants [22-25]. Geodetector could identify relationships between diseases and their associated factors by detecting similarities in the spatial distribution of independent and dependent variables, using the power of determinant as a metric. The traditional approach to interaction is to add the 2 factors to the regression model, but the interaction between the 2 factors is not necessarily multiplicative. Besides detecting spatial heterogeneity and the significant factors, Geodetector has the advantage of being able to detect the interaction between the 2 factors and analyze the effect of the strength, direction, and linearity or nonlinearity of the interaction in the distribution of the disease [26].

Therefore, this study combined epidemical statistics, geographic information system, and Geodetector modeling to investigate the characteristics of HFRS epidemic distribution patterns and spatiotemporal distribution and further quantify the meteorological effect of HFRS in Northeastern China, which could assist health departments in preparing targeted preventive and control strategies with a rational allocation of health resources.

Methods

The monthly HFRS cases from Northeastern China from January 2006 to December 2020 were collected from the Chinese Center for Disease Control and Prevention (CDC). HFRS was included on the list of Class B Notifiable Diseases in China in 1950, and HFRS cases were mandatory to report to the Chinese CDC by law according to the standard and unified protocol established by the Chinese CDC. In this study, all HFRS cases were confirmed according to the diagnostic criteria for HFRS from the Ministry of Health of the People’s Republic of China [27]. The case definition for HFRS was an individual who had traveled to an HFRS endemic area or who had contact with rodent feces, saliva, and urine within 2 months before the onset of illness, with clinical manifestations such as fever, chills, hemorrhage, headache, back pain, abdominal pain, acute renal dysfunction, and hypotension. In addition, the person had to meet at least one laboratory criteria for diagnosis: a positive result for hantavirus-specific immunoglobulin M, a 4-fold rise in titers of hantavirus-specific immunoglobulin G, a positive result for hantavirus-specific ribonucleic acid by reverse transcription polymerase chain reaction in clinical specimens, or having hantavirus isolated from clinical specimens.

Monthly meteorological factors of Northeastern China were obtained from the China Meteorological Data Sharing Service System from January 1, 2006, to December 31, 2020, including mean temperature (TEM), mean pressure (PRS), mean ground temperature (GST), sunshine duration, mean relative humidity, mean wind speed, precipitation, maximum evaporation, and minimum evaporation. Demographic information was obtained from the Chinese Statistical Yearbook. The geographic data were obtained from the National Basic Geographic Information Center.

Based on the previous studies, we used meteorological factors with a lag of 1-6 months [28-30] because meteorological factors showed a delayed effect on the risk of HFRS. The map is shown in Figure 1, and the mechanism route in this study is shown in Figure S1 in Multimedia Appendix 1. We analyzed the epidemiological characteristics and periodic fluctuation of HFRS with data on yearly and monthly scales. From the seasonal perspective, we classified March to May as spring, June to August as summer, September to November as autumn, and December to February as winter. Descriptive analysis of demographic characteristics of HFRS cases was mainly analyzed using Excel (version 2019; Microsoft Corp) and SPSS (version 21.0; IBM Corp). The annual incidence rate of HFRS in prefecture-level cities was calculated based on the number of new HFRS cases and the total population at the end of the year. Global spatial autocorrelation analysis, which describes the overall spatial distribution and evaluates whether the attribute has spatial aggregation [31], was used to analyze the distribution of HFRS cases using 36 cities in 3 provinces in Northeastern

China. The global Moran *I* and *Z* values were computed to evaluate the clustering pattern. A positive Moran *I* value ($I>0$) means there is a positive spatial correlation and HFRS has an aggregated distribution, a Moran *I* value close to 0 means there is no autocorrelation, and a negative Moran index ($I<0$) means there is negative spatial correlation and HFRS has a discrete distribution.

Wavelet analysis was used to explore the variation of HFRS incidence periodicity [32] and detect the shift of the periodic pattern of the HFRS epidemic in Northeastern China by using Matlab (version 2016b; MathWorks Inc).

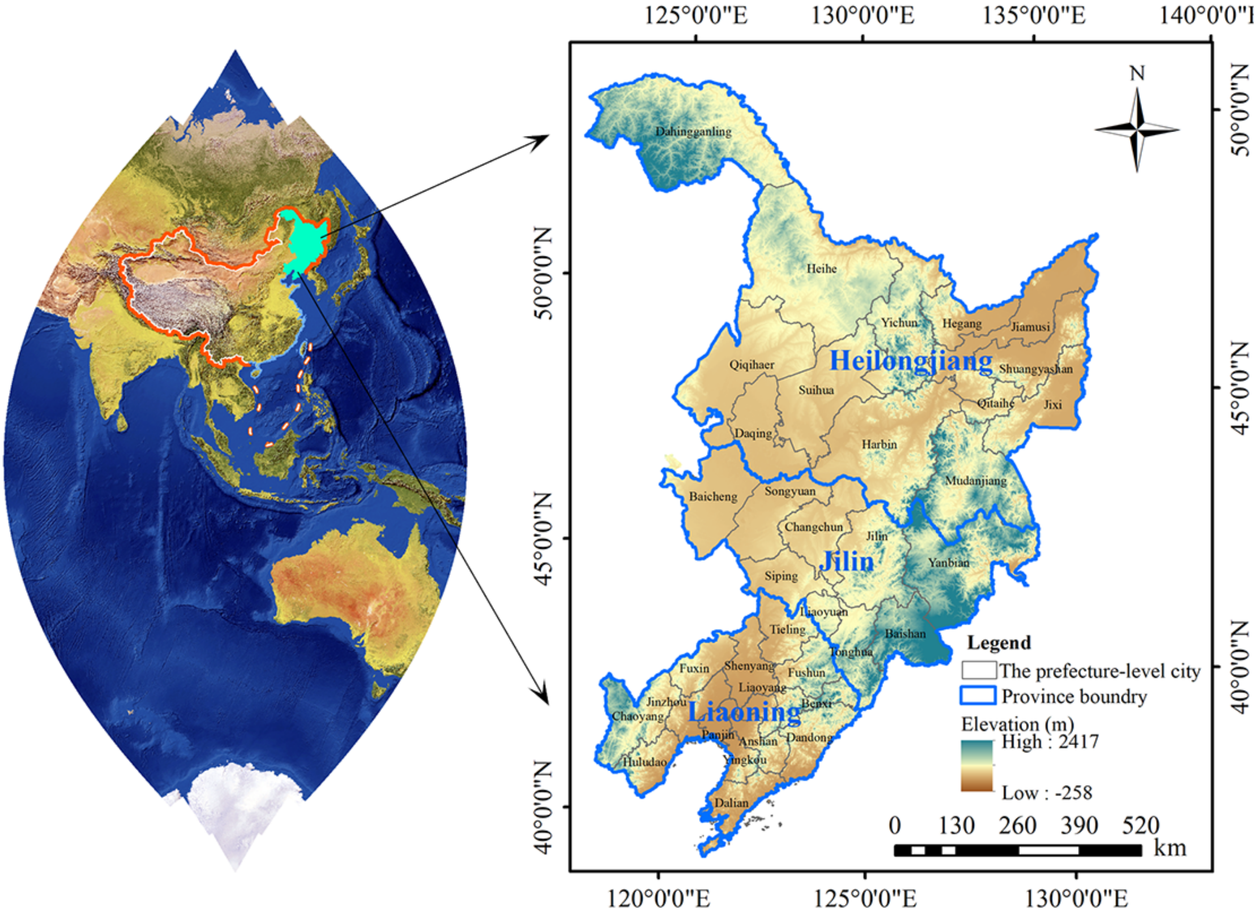
Subsequently, Geodetector modeling [22] was used to identify the governing force of a responding variable under the assumption that variable *X* is associated with variable *Y* if their spatial patterns are consistent. This study used the factor detector to detect the spatial heterogeneity of HFRS and how *X* explains the spatial patterns of HFRS using the *GD* package in R software (version 4.1.1; R Foundation for Statistical Computing). It is defined as the *q*-statistic:

(1)

where variable *X* is stratified into strata $h = 1, 2, \ldots, L$; the study area is composed of *N* units; stratum *h* is composed of N_h units; σ_h^2 is the stratum variance; σ^2 is the population variance; and *Y* means the HFRS cases. The *q* value indicates how the variable *X* explains $100 \times q\%$ of the dependent variable *Y*. For *q* [0,1], a higher value of *q* indicates a stronger spatially stratified heterogeneity (SSH) of the dependent variable *Y*. The *q* value can be used as a major factor to identify the meteorological effect of HFRS. SSH represents that the value of a certain attribute differs between different types or regions. The interaction detector could assess whether the explanatory powers of the 2 factors are enhanced, weakened, or independent of each other. The *q* values of 2 factors and their interaction were calculated. The new layer of interaction is formed by the tangent of overlay variables *X1* and *X2*. By comparing $q(X1)$ and $q(X2)$ with $q(X1 \cap X2)$, the interaction type between 2 factors can be identified: the interaction relationship between 2 factors is a bivariate enhancement when $\max(q(X1), q(X2)) < q(X1 \cap X2) < q(X1) + q(X2)$, whereas it is a nonlinear enhancement if $q(X1 \cap X2) > q(X1) + q(X2)$.

Finally, the HFRS cases data in Northeastern China from 2006 to 2020 were analyzed using the seasonal autoregressive integrated moving average (SARIMA) model to forecast the HFRS cases in the next 5 years (60 months). Mean absolute error, mean absolute percentage error (MAPE), and root mean square error were used to measure the performance of the SARIMA model.

Figure 1. The study area in Northeastern China, which was obtained from the National Catalogue Service For Geographic Information of China [33].



Results

Demographic Characteristics

A total of 52,655 cases of HFRS occurred in Northeastern China, of which 26,088, 11,386 and 15,181 cases were reported in Heilongjiang, Jilin, and Liaoning provinces, respectively, with an average of 293 new HFRS cases every month in Northeastern China from 2006 to 2020 (Table 1). Heilongjiang province has the highest incidence of HFRS in Northeastern China with 8.23 per 10,000 cases (26,088/31,710,000), followed by Jilin province with 4.75 per 10,000 cases (15,181/23,990,000) and Liaoning province with 3.57 per 10,000 cases (11,386/42,550,000).

In the analysis of population distribution characteristics of HFRS from 2006 to 2020 (Table 1), patients with HFRS were predominantly male; of the 52,655 cases, 40,664 (77.23%) were male and 11,991 (22.77%) were female, with a sex ratio of 3.4:1. Most patients with HFRS (n=36,558, 69.43%) were aged between 30-59 years, of which 38.55% (14,092/36,558) were in the 40-49 years age group.

Table 1. The characteristics of hemorrhagic fever with renal syndrome cases in Northeast China from 2006 to 2020.

	Heilongjiang province		Jilin province		Liaoning province		Northeastern China		
	Male, n	Female, n	Male, n	Female, n	Male, n	Female, n	Male, n	Female, n	Total, n
Age group (years)									
0-9	25	21	23	14	15	24	63	59	122
10-19	749	174	354	91	446	156	1549	421	1970
20-29	2945	591	1253	203	1370	333	5568	1127	6695
30-39	4831	930	2147	371	2162	604	9140	1905	11,045
40-49	5637	1395	2541	581	2937	1001	11,115	2977	14,092
50-59	3976	1449	1808	611	2435	1142	8219	3202	11,421
60-69	1811	751	749	319	1258	646	3818	1716	5534
70-79	483	236	182	93	368	183	1033	512	1545
≥80	61	23	30	16	68	33	159	72	231
Total	20,518	5570	9087	2299	11,059	4122	40,664	11,991	52,655

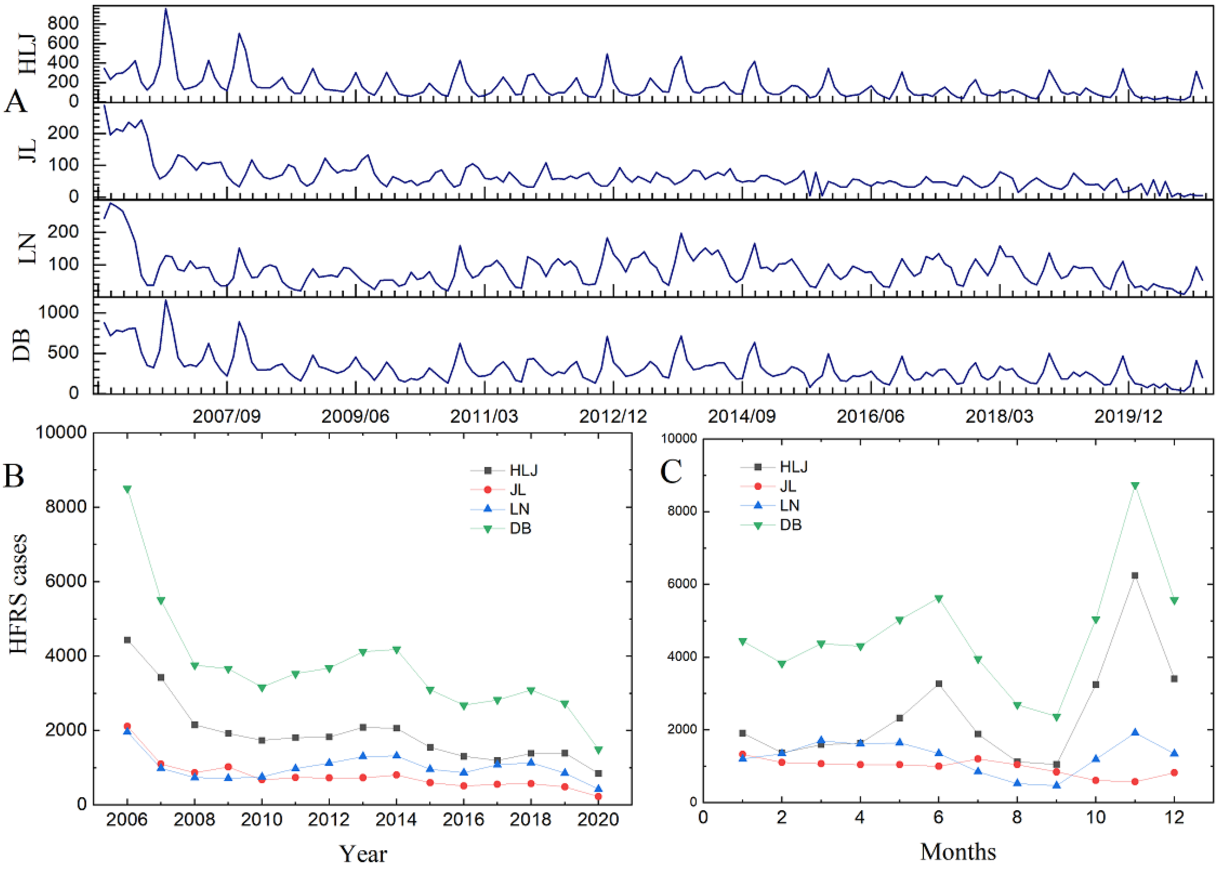
Temporal Distribution of HFRS

The analysis of the number of HFRS incidences over 15 years shows that the number of HFRS cases in Northeastern China generally showed a decreasing trend over years, mainly with a rapid decrease from 2006 to 2010. An increasing trend was demonstrated from 2011 to 2014, followed by a fluctuation reduction between 2011 and 2016. There was a rising trend between 2016 and 2018 and a decreasing trend between 2019 and 2020.

As for the monthly characteristics, this study also showed that HFRS cases were seasonal, including 2 HFRS epidemic peaks: summer and winter peaks. The epidemic months of Northeastern China were different in the 3 provinces.

In November, Northeastern China had the highest HFRS incidence with 0.88 per 10,000 cases (8676/98,250,000), followed by a lower incidence in June with 0.59 per 10,000 cases (5818/98,250,000). Specifically, Heilongjiang province had a prominent winter peak of HFRS from 2006 to 2020, and HFRS cases spiked in November, accounting for 1.75 per 10,000 cases (5539/31,710,000) in Heilongjiang province in November. However, the number of HFRS cases in Liaoning province peaked in late autumn, with cases concentrated in November with 0.45 per 10,000 cases (1921/42,550,000) in Liaoning province, which differed slightly from the second incidence peak of 0.40 per 10,000 cases (1706/42,550,000) in March. The monthly and seasonal variation trends of HFRS cases in the 3 provinces of Northeastern China are shown in Figure 2.

Figure 2. Yearly and monthly variations of HFRS in Northeastern China from 2006 to 2020. (A) Monthly time series change of HFRS cases from January 2006 to December 2020. (B) Yearly variation of HFRS cases in the 3 provinces and Northeastern China. (C) Seasonal variation of the HFRS cases in the 3 provinces and Northeastern China. DB: Northeastern China; HFRS: hemorrhagic fever with renal syndrome; HLJ: Heilongjiang province; JL: Jilin province; LN: Liaoning province.

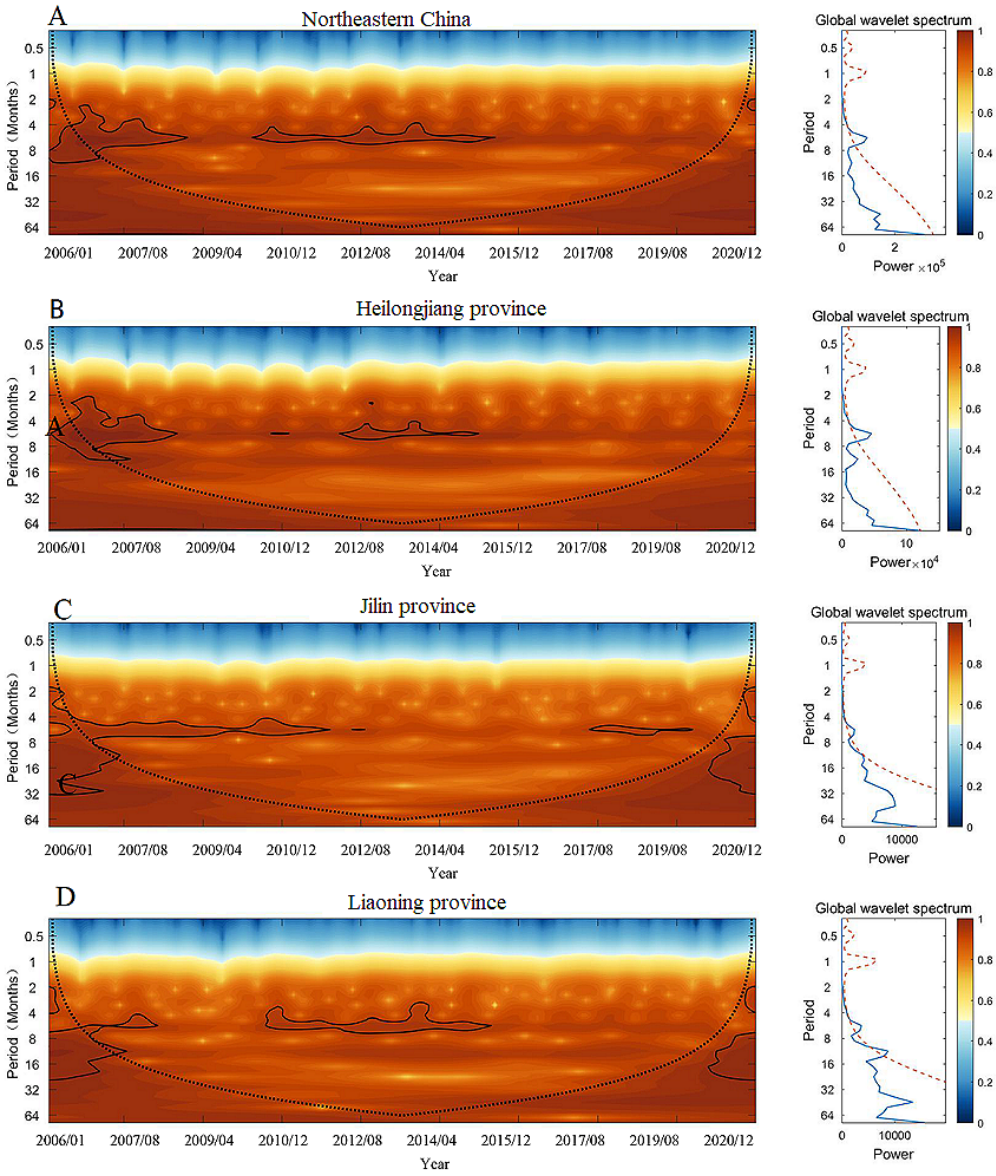


Wavelet Analysis

Wavelet analysis was performed on HFRS cases in Northeastern China for a total of 180 months from 2006 to 2020, and the main periodicities and variation characteristics of the 3 provinces were identified. The results are shown in Figure 3.

Wavelet power spectra detected a significant 4- to 6-month periodic mode for HFRS in Northeastern China from 2006 to 2020 (Figure 3A), which was characterized by a strong periodical fluctuation of 2 to 8 months from 2006 to 2009, while a dominant shift period of 4 to 6 months existed from 2010 to 2015. In Heilongjiang province (Figure 3B), a 2- to 8-month periodicity from 2006 to 2009 and an approximately 4- to 6-month periodicity from 2014 to 2015 were revealed. The HFRS periodicity in Jilin province was 4 to 6 months, mainly concentrated from 2007 to 2014 and from 2017 to 2019 (Figure 3C). In Liaoning province, the periodicity was primarily 4 to 6 months from 2010 to 2015 (Figure 3D).

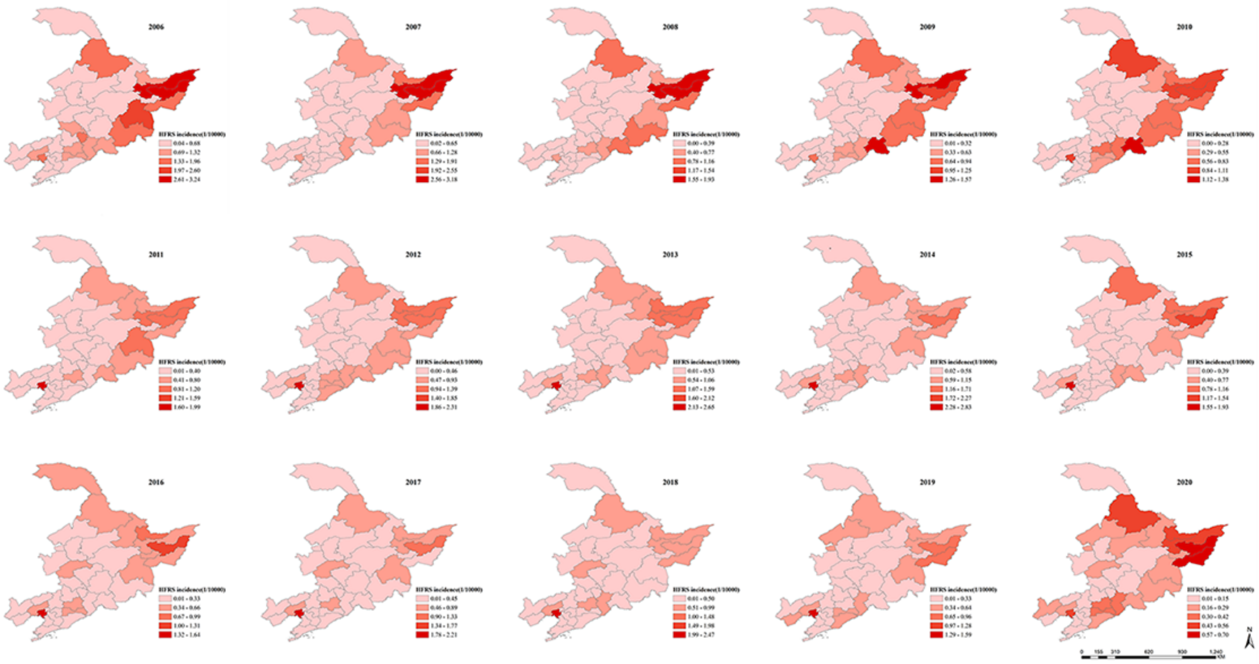
Figure 3. Wavelet power spectrum of monthly HFRS cases in Northeastern China. A mean wavelet coherency spectrum and variance of HFRS wavelet coherency spectrum in the northeast region of China between January 2006 and December 2020. The black line represents the influence cone that delimits the region that is totally not influenced by edge effects. (A) Northeastern China, (B) Heilongjiang province, (C) Jilin province, and (D) Liaoning province. Blue line: global average value of the wavelet spectrum; red dot line: global average value of the spectrum with 95% confidence. HFRS: hemorrhagic fever with renal syndrome.



Spatial Distribution and Cluster

The annual incidence rate of HFRS was unequally distributed among prefecture-level cities in Northeastern China during the 15 years and displayed a decreasing trend by year (Figure 4). Between 2006 and 2010, the areas with high incidence rates of HFRS were mainly concentrated in the eastern and northern parts of Jiamusi City and Shuangyashan City in Heilongjiang province. In contrast, between 2011 and 2015, the number of HFRS cases in cities in Northeastern China gradually decreased. The city with the highest incidence shifted from Shuangyashan City, Jiamusi City, and Hegang City in Heilongjiang province to Panjin City in Liaoning province. Although HFRS cases were still mainly concentrated in eastern Heilongjiang province, the number of patients in the 2 main endemic areas of Jiamusi City and Shuangyashan City decreased substantially, and the distribution range did not expand. Between 2016 and 2020, the trend of HFRS occurrence decreased from east to west. The cities with the annual highest HFRS incidence were Shuangyashan City and Jixi City in the east. The east coastal cities in Northeastern China’s southern part had also become one of the high-risk regions of HFRS, similar to the distribution between 2006 and 2010. Spatial autocorrelation analysis of HFRS incidence for 15 years showed a spatially positive correlation. Moran index fluctuated from 0.02 to 0.43 from 2006 to 2020. For most of the time, HFRS showed a spatially positive correlation with significant spatial aggregation rather than random distribution (Table S1 in Multimedia Appendix 1). The largest Moran index and the strongest aggregation effect was in 2007 during the 15 years (*I*=0.43; *Z*=4.56). There was a positive spatial correlation and aggregated distribution of HFRS prevalence from 2005 to 2010 and in 2016 and 2020, and its overall trend was weakening.

Figure 4. Yearly distribution of hemorrhagic fever with renal syndrome (HFRS) incidence in Northeastern China, from 2006 to 2020. For a higher-resolution version of this figure, see Multimedia Appendix 1.



Geodetector Modeling

HFRS in Northeastern China was characterized by significant SSH. The results of this study showed that the dominant meteorological factors influencing the incidence of HFRS in different regions in Northeastern China had significant differences (all *P*<.001; Table S1 in Multimedia Appendix 2). In Heilongjiang province, the meteorological factors with the greatest influence on HFRS were TEM with a 4-month lag (15.49%), GST with a 4-month lag (14.65%), and PRS with a 5-month lag (7.35%). In Liaoning province, TEM with a 1-month lag (5.79%), GST with a 1-month lag (5.35%), and mean wind speed with a 4-month lag (4.94%) were the top 3 determinants. However, the meteorological factors with the greatest influence on HFRS in Jilin province were precipitation with a 6-month lag (5.52%), maximum evaporation with a 5-month lag (5.36%), and PRS (5.26%). GST with a 1-month lag (4.18%), precipitation with a 3-month lag (3.54%), and TEM with a 4-month lag (3.31%) also had a significant effect on the prevalence of HFRS in Northeastern China (all *P*<.001), as shown in Table S1 in Multimedia Appendix 2.

The effect of any 2 combination of different meteorological factors on HFRS was also quantified simultaneously (Tables S2-5 in Multimedia Appendix 2). The effect of individual factors on the onset of HFRS was enhanced after combining 2 factors compared to the effect of the factors alone. In Heilongjiang province, the q value of PRS with a 5-month lag and GST with a 4-month lag was 0.2680. This indicates the strongest interaction of PRS with a 5-month lag and GST with a 4-month lag could explain 26.80% of the HFRS incidence, which also means that the 2 factors together had a higher interaction than their respective independent effects (Table S5 Multimedia Appendix 2).

In Jilin province, the q value of precipitation with a 6-month lag and PRS was 0.1289, and the explanatory power of the 2 interactions was greater than the sum of the explanatory power of their independent effects, indicating a nonlinear enhancement effect (Table S3 in Multimedia Appendix 2). In Liaoning province, mean relative humidity with a 1-month lag and TEM with a 1-month lag have an explanatory power of 10.90% for the HFRS epidemic (Table S4 in Multimedia Appendix 2).

Time Series Forecasting of HFRS

By constructing the SARIMA model (4, 1, 3)(1, 0, 1)[4], a global trend of fluctuating decrease in HFRS cases was predicted in Northeastern China over the next 5 years. The results showed the training accuracy of the SARIMA model (MAPE=0.17) and the summary performance with the HFRS data set (MAPE=0.23; Table S6 in Multimedia Appendix 2). The model predicted that 8343 cases of HFRS are expected to occur in Northeastern China, with a fluctuation range of 4 to 250 cases per month.

Discussion

Principal Findings

This long-term analysis from Northeastern China is the first comprehensive work to interpret the spatiotemporal distribution and different explanatory power of various meteorological variables for HFRS from 2006 to 2020, providing essential clues to understanding how HFRS developed.

Comparison With Prior Work

From 2006 to 2020, the epidemic situation of HFRS in Northeastern China was complex and severe, with an overall fluctuating downward trend. The epidemic was widespread in Northeastern China, with reports from 36 cities within the range. However, the risk of HFRS incidence varied widely among prefectures. The distribution of prefecture-level cities in

Northeastern China is shown in Figure 1. This decreasing trend is consistent with previous studies [34,35] and primarily due to the implementation of infectious disease prevention and control measures of the Chinese government. With the development of the economy, improvement of medical conditions, and promotion and application of vaccines, the incidence of HFRS in China has been effectively controlled, and the overall incidence rate has decreased significantly [36,37]. Since the Chinese government's implementation of the HFRS immunization program and approximately 2 million doses of HFRS vaccine having been administered annually in China, the distribution and incidence of HFRS infection have changed significantly [38,39]. The evidence revealed a negative correlation between HFRS incidence and vaccination adherence and that vaccination-induced progressive reduction of viral infection in susceptible populations prevents further epidemics and has a notable effect on HFRS prevention and control.

High prevalence of HFRS is found in male individuals in the 30-59 years age group. There are significant differences in the spatial and temporal distribution of HFRS cases in Northeastern China. The eastern and southern prefecture-level cities were the high-risk clusters of HFRS epidemic in Northeastern China, primarily in Heilongjiang province, such as Qitaihe city, Heihe city, and Harbin city. This may be due to the fact that middle-aged male individuals have more exposure to host animals than other populations when they work in endemic areas [40,41]; it also suggests that middle-aged male individuals are a high-risk group for HFRS exposure. In addition, the distribution and number of HFRS host animal populations are influenced by environmental conditions such as temperature and geography, resulting in different distribution and prevalence intensities of HFRS in the different prefecture-level cities [42-44].

The study by He et al [45] suggest that the influence of climate on HFRS becomes increasingly sensitive over time, and it revealed an important clue that the geographic factors with the most significant impact on climate-HFRS association include latitude, distance, and longitude from the coastline, grassland, and woodland. In this study, Heilongjiang province is a prime HFRS epidemic area with a higher incidence of HFRS than the other 2 provinces. The dynamics of HFRS incidence may be due to unique local natural climatic, geographic, and socioeconomic factors. Heilongjiang province is a low mountainous region with high southeast and low northwest topography, located in the midlatitude zone with a cold-temperate continental monsoon climate. Since high-risk HFRS cities are located at the center of the eastern part of Heilongjiang province, their convenient transportation, frequent trade, and tourism with the surrounding areas may, to some extent, contribute to the spread of HFRS. In addition, the land types in the eastern cities are mainly forested and arable land; for example, the arable land area in Qitaihe City accounts for about 31.3% of the total land area of the city and forested land accounts for about 52.8% [46]. Land types are related to the risk of HFRS, which implies that with the development of woodland and farmland, new habitats for hantavirus hosts may also be more available, increasing host density in areas of human activity and further enhancing the risk of HFRS transmission.

The re-emergence of Jiamusi City and Shuangyashan City as hot spots of HFRS in 2020 also suggests that we should focus on strengthening the prevention and control of the epidemic in this region. In addition, the predictions of the SARIMA model suggest that we are still at risk of an epidemic of HFRS in the next 5 years. This information helps health officials better understand the epidemiological patterns of HFRS and encourages them to adapt monitoring, prevention, and control strategies to the HFRS epidemic's periodicity and seasonal fluctuations in Northeastern China.

Additionally, in this study, the seasonality of HFRS was related to the climate condition, with 2 peak incidences per year—summer and winter—and the peak incidence in winter is higher than that in summer in most years. This finding suggests that the influences of meteorological factors could partly explain seasonal variations of HFRS. This characteristic is consistent with previous results [47], which showed that the spatiotemporal dynamics of HFRS cases in Jiangxi province, that is, the HFRS epidemic, has a bimodal seasonal pattern with the primary peak occurring in winter and the secondary peak in early summer. This is because the difference in HFRS incidence between winter and summer is accompanied by the difference in climate between seasons, indicating that changes in meteorological factors play an important role in the dynamics of HFRS [43,48]. Studies have shown seasonal peaks in the HFRS outbreak in China [17,49], and in this study, HFRS risk was much higher in spring in Liaoning and Jilin provinces, and there was a higher peak in Heilongjiang province in winter. The changes in climatic factors are the significant driving force of HFRS and are an important sensitive reason for the occurrence of HFRS [20,45]. In summer and winter, the obvious changes in temperature and precipitation could affect aerosols and crop growth, which provide the conditions for rodent activities [36]; for example, rats may congregate in residential areas during a cold winter to live, increasing the density of rats and their contact with people, which may accelerate the spread of hantavirus [19,43,50].

Possible Implications for Public Health

This study highlights the role of GST, whose separate factor effects were detected to be larger than the effect of temperature and rainfall (Heilongjiang province). This is an outstanding finding of this paper, as few studies have included this indicator in the survey of HFRS. The level of GST is closely related to the growth and development of crops, decomposition of fertilizers, and accumulation of organic matter and is an important environmental factor in plant growth [51]. GST also plays a highly crucial role in microclimate formation [52,53], so it can be hypothesized that soil temperature further influences the survival of host animals by affecting the growth of vegetation and soil microbes, which ultimately leads to changes in host-human transmission of the hantavirus. Furthermore, the interaction of meteorological factors on HFRS was analyzed by the Geodetector, and the interaction of the 2 factors was found to be significantly higher than that of the factors alone. Most of the other meteorological factors interacting with the GST factor also showed enhanced effects, further illustrating the importance of the GST factor.

Limitations

The results of this study also have some limitations. The spatial scale of this study was based on the monthly HFRS case characteristics at the prefecture-level cities in Northeastern China, which may miss some smaller-scale characteristics of HFRS incidence at the county level. Moreover, the HFRS data were obtained from the National Passive Surveillance System, and some unreported data could not be collected. In addition, the current research focused on the temporal and spatial dynamics of HFRS in Northeastern China and has yet to include the factors influencing the disease, such as rodent density, socioeconomic level, and so on. In future studies, more factors will be collected to analyze the relationship between HFRS and environmental-social factors.

Conclusion

Climate factor fluctuations significantly shape the spatiotemporal pattern of HFRS, especially the ground temperature and precipitation factors. Prevention and management should be further focused on high-risk provinces and cities, such as the eastern part of Northeastern China. The health department should make comprehensive prevention and control policies and take measures in advance to prevent HFRS outbreaks in summer and winter. This study will enhance the understanding of meteorological effects of HFRS transmission in Northeastern China over the last 15 years and is helpful to the high-risk region for the scientific adoption of prevention and control strategies of zoonotic disease.

Acknowledgments

We would like to acknowledge the help of the help of Chunyu Zhao from China Medical University (School of Public Health) and Fan Yang from Nanjing University (School of Atmospheric Sciences) for this work.

This study was funded by grants from the National Natural Science Foundation of China (12031010), the Special Grant for the Prevention and Control of Infectious Diseases (2018ZX10713003), and projects from the 13th Five-Year Plan (17SAZ01 and 18QNP063). The funders had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.

Authors' Contributions

Yong Wang, WZ, and Yanding Wang contributed to conceptualization. RJ, XZ, XP, MY, ZL, JG, WY, and YC contributed to data curation. WZ and Yong Wang contributed to funding acquisition. Yanding Wang and XW contributed to materials and analysis tools. WZ and Yong Wang supervised the study. Yanding Wang drafted and revised the manuscript. All authors read and approved the final paper. WZ (zwy0419@126.com) and Yong Wang (ywang7508@sina.com) contributed equally as co-corresponding authors for this paper. WZ and Yong Wang supervised the study and contributed as senior coauthors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Spatial autocorrelation results, mechanism route, and yearly distribution of hemorrhagic fever with renal syndrome (HFRS) incidence.
[DOCX File , 50097 KB - publichealth_v9i1e42673_app1.docx]

Multimedia Appendix 2
Factor detector, interaction detector, and seasonal autoregressive integrated moving average (SARIMA) model results.
[DOC File , 140 KB - publichealth_v9i1e42673_app2.doc]

References

1. Schmaljohn CS, Hasty SE, Harrison SA, Dalrymple JM. Characterization of Hantaan virions, the prototype virus of hemorrhagic fever with renal syndrome. J Infect Dis 1983 Dec 01;148(6):1005-1012. [doi: 10.1093/infdis/148.6.1005] [Medline: 6140290]
2. Jiang H, Zheng X, Wang L, Du H, Wang P, Bai X. Hantavirus infection: a global zoonotic challenge. Virol Sin 2017 Feb 23;32(1):32-43 [FREE Full text] [doi: 10.1007/s12250-016-3899-x] [Medline: 28120221]
3. Sundberg E, Hultdin J, Nilsson S, Ahlm C. Evidence of disseminated intravascular coagulation in a hemorrhagic fever with renal syndrome-scoring models and severe illness. PLoS One 2011 Jun 23;6(6):e21134 [FREE Full text] [doi: 10.1371/journal.pone.0021134] [Medline: 21731657]
4. Jiang H, Du H, Wang LM, Wang PZ, Bai XF. Hemorrhagic fever with renal syndrome: pathogenesis and clinical picture. Front Cell Infect Microbiol 2016 Feb 03;6:1 [FREE Full text] [doi: 10.3389/fcimb.2016.00001] [Medline: 26870699]
5. Zhu Y, Chen Y, Zhu Y, Liu P, Zeng H, Lu N. A retrospective study of acute pancreatitis in patients with hemorrhagic fever with renal syndrome. BMC Gastroenterol 2013 Dec 17;13:171. [doi: 10.1186/1471-230X-13-171] [Medline: 24345089]

6. Hong YM, Moon JC, Yang HC, Kang KP, Kim W, Park SK, et al. Hemorrhagic fever with renal syndrome and coexisting hantavirus pulmonary syndrome. Kidney Res Clin Pract 2012 Jun;31(2):118-120 [FREE Full text] [doi: 10.1016/j.krcp.2012.04.318] [Medline: 26889418]
7. Huong VTQ, Yoshimatsu K, Luan VD, Tuan LV, Nhi L, Arikawa J, et al. Hemorrhagic fever with renal syndrome, Vietnam. Emerg Infect Dis 2010 Feb;16(2):363-365. [doi: 10.3201/eid1602.091204] [Medline: 20113591]
8. Yi YJ, Kang M, Kim WK, Huh K, Klingström J, Song JW, et al. Association between haemorrhagic fever with renal syndrome and cancers. Int J Infect Dis 2021 Dec;113:127-135 [FREE Full text] [doi: 10.1016/j.ijid.2021.10.014] [Medline: 34653656]
9. Klingström J, Granath F, Ekbom A, Björkström NK, Ljunggren H. Increased risk for lymphoma following hemorrhagic fever with renal syndrome. Clin Infect Dis 2014 Oct 15;59(8):1130-1132. [doi: 10.1093/cid/ciu488] [Medline: 24965350]
10. Zhang Y, Zhang F, Wang JB, Zhao ZW, Li MH, Chen HX, et al. Hantaviruses in rodents and humans, Inner Mongolia Autonomous Region, China. Emerg Infect Dis 2009 Jun;15(6):885-891 [FREE Full text] [doi: 10.3201/eid1506.081126] [Medline: 19523286]
11. Tkachenko EA, Ishmukhametov AA, Dzagurova TK, Bernshtein AD, Morozov VG, Siniugina AA, et al. Hemorrhagic fever with renal syndrome, Russia. Emerg Infect Dis 2019 Dec;25(12):2325-2328 [FREE Full text] [doi: 10.3201/eid2512.181649] [Medline: 31742540]
12. Schneider F, Vidal L, Auvray C, Khider Y, Graesslin O. The first French hemorrhagic fever with renal syndrome in pregnant woman. Article in French. J Gynecol Obstet Biol Reprod (Paris) 2009 Sep;38(5):440-442 [FREE Full text] [doi: 10.1016/j.jgyn.2009.04.011] [Medline: 19497685]
13. Morita C, Sugiyama K, Matsuura Y, Kitamura T, Komatsu T, Akao Y, et al. Detection of antibody against hemorrhagic fever with renal syndrome (HFRS) virus in sera of house rats captured in port areas of Japan. Jpn J Med Sci Biol 1983 Feb;36(1):55-57. [doi: 10.7883/yoken1952.36.55] [Medline: 6408289]
14. Zhang YZ, Zou Y, Fu ZF, Plyusnin A. Hantavirus infections in humans and animals, China. Emerg Infect Dis 2010 Aug;16(8):1195-1203 [FREE Full text] [doi: 10.3201/eid1608.090470] [Medline: 20678311]
15. Zou LX, Sun L. Analysis of hemorrhagic fever with renal syndrome using wavelet tools in mainland China, 2004-2019. Front Public Health 2020 Dec 1;8:571984 [FREE Full text] [doi: 10.3389/fpubh.2020.571984] [Medline: 33335877]
16. Zhang WY, Wang LY, Liu YX, Yin WW, Hu WB, Magalhaes RJS, et al. Spatiotemporal transmission dynamics of hemorrhagic fever with renal syndrome in China, 2005-2012. PLoS Negl Trop Dis 2014 Nov 20;8(11):e3344 [FREE Full text] [doi: 10.1371/journal.pntd.0003344] [Medline: 25412324]
17. Shang C, Sun Y, Yin Q, Huang X, Liu X, Zhang Q, et al. Hemorrhagic fever with renal syndrome - Liaoning Province, China, 1999-2018. China CDC Wkly 2020 May 15;2(20):350-354 [FREE Full text] [doi: 10.46234/ccdcw2020.091] [Medline: 34594659]
18. Li CP, Cui Z, Li SL, Magalhaes RJS, Wang B, Zhang C, et al. Association between hemorrhagic fever with renal syndrome epidemic and climate factors in Heilongjiang Province, China. Am J Trop Med Hyg 2013 Nov;89(5):1006-1012 [FREE Full text] [doi: 10.4269/ajtmh.12-0473] [Medline: 24019443]
19. He J, Christakos G, Wu J, Jankowski P, Langousis A, Wang Y, et al. Probabilistic logic analysis of the highly heterogeneous spatiotemporal HFRS incidence distribution in Heilongjiang province (China) during 2005-2013. PLoS Negl Trop Dis 2019 Jan 31;13(1):e0007091 [FREE Full text] [doi: 10.1371/journal.pntd.0007091] [Medline: 30703095]
20. Wu W, Guo JQ, Yin Z, Wang PH, Zhou BS. GIS-based spatial, temporal, and space-time analysis of haemorrhagic fever with renal syndrome. Epidemiol Infect 2009 Dec;137(12):1766-1775. [doi: 10.1017/S0950268809002659] [Medline: 19393118]
21. Li S, Ren H, Hu W, Lu L, Xu X, Zhuang D, et al. Spatiotemporal heterogeneity analysis of hemorrhagic fever with renal syndrome in China using geographically weighted regression models. Int J Environ Res Public Health 2014 Nov 25;11(12):12129-12147 [FREE Full text] [doi: 10.3390/ijerph111212129] [Medline: 25429681]
22. Wang JF, Zhang TL, Fu BJ. A measure of spatial stratified heterogeneity. Ecological Indicators 2016 Aug;67:250-256. [doi: 10.1016/j.ecolind.2016.02.052]
23. Zeng W, Wan X, Lei M, Gu G, Chen T. Influencing factors and prediction of arsenic concentration in Pteris vittata: a combination of geodetector and empirical models. Environ Pollut 2022 Jan 01;292(Pt A):118240. [doi: 10.1016/j.envpol.2021.118240] [Medline: 34619180]
24. Chen Y, Zhou Y, NixiaCiren, Zhang H, Wang C, GesangDeji, et al. Spatiotemporal variations of surface ozone and its influencing factors across Tibet: a Geodetector-based study. Sci Total Environ 2022 Mar 20;813:152651. [doi: 10.1016/j.scitotenv.2021.152651] [Medline: 34954172]
25. Wang Z, Li X, Zhao H. Identifying spatial influence of urban elements on road-deposited sediment and the associated phosphorus by coupling Geodetector and Bayesian Networks. J Environ Manage 2022 Aug 01;315:115170. [doi: 10.1016/j.jenvman.2022.115170] [Medline: 35500491]
26. Song Y, Wang J, Ge Y, Xu C. An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: cases with different types of spatial data. GIScience & Remote Sensing 2020 May 12;57(5):593-610. [doi: 10.1080/15481603.2020.1760434]

27. Ministry of Health. Handbook of Epidemic Hemorrhagic Fever Prevention and Control. Beijing, China: China People's Health Publishing House; 1998.
28. Luo Y, Lv H, Yan H, Zhu C, Ai L, Li W, et al. Meteorological change and hemorrhagic fever with renal syndrome epidemic in China, 2004-2018. Sci Rep 2022 Nov 21;12(1):20037 [FREE Full text] [doi: 10.1038/s41598-022-23945-9] [Medline: 36414682]
29. Xu Q, Li R, Rutherford S, Luo C, Liu Y, Wang Z, et al. Using a distributed lag non-linear model to identify impact of temperature variables on haemorrhagic fever with renal syndrome in Shandong Province. Epidemiol Infect 2018 Oct;146(13):1671-1679. [doi: 10.1017/S095026881800184X] [Medline: 29976265]
30. Wang Y, Wei X, Xiao X, Yin W, He J, Ren Z, et al. Climate and socio-economic factors drive the spatio-temporal dynamics of HFRS in Northeastern China. One Health 2022 Dec;15:100466 [FREE Full text] [doi: 10.1016/j.onehlt.2022.100466] [Medline: 36532667]
31. Moran PAP. The interpretation of statistical maps. J R Stat Soc Series B Stat Methodol 2018 Dec 05;10(2):243-251. [doi: 10.1111/j.2517-6161.1948.tb00012.x]
32. Cazelles B, Chavez M, Magny GCD, Guégan JF, Hales S. Time-dependent spectral analysis of epidemiological time-series with wavelets. J R Soc Interface 2007 Aug 22;4(15):625-636 [FREE Full text] [doi: 10.1098/rsif.2007.0212] [Medline: 17301013]
33. National Catalogue Service For Geographic Information of China. URL: https://www.webmap.cn/mapDataAction.do?method=forw&resType=5&storeId=2 [accessed 2023-04-18]
34. Sun L, Zou LX. Spatiotemporal analysis and forecasting model of hemorrhagic fever with renal syndrome in mainland China. Epidemiol Infect 2018 Oct;146(13):1680-1688. [doi: 10.1017/S0950268818002030] [Medline: 30078384]
35. Liu Q, Liu X, Jiang B, Yang W. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. BMC Infect Dis 2011 Aug 15;11(1):218 [FREE Full text] [doi: 10.1186/1471-2334-11-218] [Medline: 21838933]
36. Tian H, Hu S, Cazelles B, Chowell G, Gao L, Laine M, et al. Urbanization prolongs hantavirus epidemics in cities. Proc Natl Acad Sci U S A 2018 May 01;115(18):4707-4712 [FREE Full text] [doi: 10.1073/pnas.1712767115] [Medline: 29666240]
37. Brocato RL, Hooper JW. Progress on the prevention and treatment of hantavirus disease. Viruses 2019 Jul 04;11(7):610 [FREE Full text] [doi: 10.3390/v11070610] [Medline: 31277410]
38. Wu W, Guo J, An S, Guan P, Ren Y, Xia L, et al. Comparison of two hybrid models for forecasting the incidence of hemorrhagic fever with renal syndrome in Jiangsu Province, China. PLoS One 2015 Aug 13;10(8):e0135492 [FREE Full text] [doi: 10.1371/journal.pone.0135492] [Medline: 26270814]
39. Li Z, Zeng H, Wang Y, Zhang Y, Cheng L, Zhang F, et al. The assessment of Hantaan virus-specific antibody responses after the immunization program for hemorrhagic fever with renal syndrome in northwest China. Hum Vaccin Immunother 2017 Apr 03;13(4):802-807 [FREE Full text] [doi: 10.1080/21645515.2016.1253645] [Medline: 27824286]
40. Turčinov D, Puljiz I, Markotić A, Kuzman I, Begovac J. Clinical and laboratory findings in patients with oliguric and non-oliguric hantavirus haemorrhagic fever with renal syndrome: an analysis of 128 patients. Clin Microbiol Infect 2013 Jul;19(7):674-679 [FREE Full text] [doi: 10.1111/j.1469-0691.2012.03994.x] [Medline: 22963396]
41. Koch J, Brockmann S, Winter C, Kimmig P, Stark K. Significant increase of hantavirus infections in Germany since the beginning of 2007. Euro Surveill 2007 May 03;12(5):E070503.1 [FREE Full text] [doi: 10.2807/esw.12.18.03185-en] [Medline: 17868605]
42. Bai XH, Peng C, Jiang T, Hu ZM, Huang DS, Guan P. Distribution of geographical scale, data aggregation unit and period in the correlation analysis between temperature and incidence of HFRS in mainland China: a systematic review of 27 ecological studies. PLoS Negl Trop Dis 2019 Aug 19;13(8):e0007688 [FREE Full text] [doi: 10.1371/journal.pntd.0007688] [Medline: 31425512]
43. Tian H, Stenseth NC. The ecological dynamics of hantavirus diseases: from environmental variability to disease prevention largely based on data from China. PLoS Negl Trop Dis 2019 Feb 21;13(2):e0006901 [FREE Full text] [doi: 10.1371/journal.pntd.0006901] [Medline: 30789905]
44. Lin H, Liu Q, Guo J, Zhang J, Wang J, Chen H. Analysis of the geographic distribution of HFRS in Liaoning Province between 2000 and 2005. BMC Public Health 2007 Aug 15;7(1):207. [doi: 10.1186/1471-2458-7-207] [Medline: 17697362]
45. He J, Christakos G, Wu J, Cazelles B, Qian Q, Mu D, et al. Spatiotemporal variation of the association between climate dynamics and HFRS outbreaks in Eastern China during 2005-2016 and its geographic determinants. PLoS Negl Trop Dis 2018 Jun 6;12(6):e0006554 [FREE Full text] [doi: 10.1371/journal.pntd.0006554] [Medline: 29874263]
46. Government of Qitaihe City, China. URL: http://www.qth.gov.cn/zjqth/csgk/dlzy/ [accessed 2023-04-21]
47. Yang S, Gao Y, Liu X, Liu X, Liu Y, Metelmann S, et al. Spatiotemporal dynamics of hemorrhagic fever with renal syndrome in Jiangxi province, China. Sci Rep 2020 Aug 31;10(1):14291 [FREE Full text] [doi: 10.1038/s41598-020-70761-0] [Medline: 32868784]
48. Zhang W, Guo W, Fang L, Li C, Bi P, Glass GE, et al. Climate variability and hemorrhagic fever with renal syndrome transmission in Northeastern China. Environ Health Perspect 2010 Jul;118(7):915-920 [FREE Full text] [doi: 10.1289/ehp.0901504] [Medline: 20142167]

49. Zhang S, Wang S, Yin W, Liang M, Li J, Zhang Q, et al. Epidemic characteristics of hemorrhagic fever with renal syndrome in China, 2006-2012. BMC Infect Dis 2014 Jul 11;14(1):384 [FREE Full text] [doi: 10.1186/1471-2334-14-384] [Medline: 25012160]
50. Wu G, Xia Z, Wang F, Wu J, Cheng D, Chen X, et al. Investigation on risk factors of haemorrhagic fever with renal syndrome (HFRS) in Xuancheng City in Anhui Province, mainland China. Epidemiol Infect 2020 Oct 02;148:e248 [FREE Full text] [doi: 10.1017/S0950268820002344] [Medline: 33004084]
51. Olivera Viciedo D, de Mello Prado R, Martinez CA, Habermann E, de Cássia Piccolo M, Calero Hurtado A, et al. Changes in soil water availability and air-temperature impact biomass allocation and C:N:P stoichiometry in different organs of Stylosanthes capitata Vogel. J Environ Manage 2021 Jan 15;278(Pt 1):111540. [doi: 10.1016/j.jenvman.2020.111540] [Medline: 33126195]
52. Dacal M, Bradford MA, Plaza C, Maestre FT, García-Palacios P. Soil microbial respiration adapts to ambient temperature in global drylands. Nat Ecol Evol 2019 Feb 14;3(2):232-238 [FREE Full text] [doi: 10.1038/s41559-018-0770-5] [Medline: 30643242]
53. Moinet GYK, Dhami MK, Hunt JE, Podolyan A, Liáng LL, Schipper LA, et al. Soil microbial sensitivity to temperature remains unchanged despite community compositional shifts along geothermal gradients. Glob Chang Biol 2021 Dec 28;27(23):6217-6231 [FREE Full text] [doi: 10.1111/gcb.15878] [Medline: 34585498]

Abbreviations

CDC: Center for Disease Control and Prevention
GST: mean ground temperature
HFRS: hemorrhagic fever with renal syndrome
MAPE: mean absolute percentage error
PRS: mean pressure
SARIMA: seasonal autoregressive integrated moving average
SSH: spatially stratified heterogeneity
TEM: mean temperature

Edited by Y Khader; submitted 13.09.22; peer-reviewed by J Wang, YL Cheong, P Wang, S Ma; comments to author 02.12.22; revised version received 29.03.23; accepted 11.04.23; published 18.05.23.

Please cite as:
Wang Y, Wei X, Jia R, Peng X, Zhang X, Yang M, Li Z, Guo J, Chen Y, Yin W, Zhang W, Wang Y
The Spatiotemporal Pattern and Its Determinants of Hemorrhagic Fever With Renal Syndrome in Northeastern China: Spatiotemporal Analysis
JMIR Public Health Surveill 2023;9:e42673
URL: https://publichealth.jmir.org/2023/1/e42673
doi:10.2196/42673
PMID:37200083

Original Paper

The Role of Telehealth in Promoting Equitable Abortion Access in the United States: Spatial Analysis

Leah R Koenig[1,2,3], MSPH; Andréa Becker[1], MA, PhD; Jennifer Ko[1,3], MLiS; Ushma D Upadhyay[1,3], MPH, PhD

[1]Advancing New Standards in Reproductive Health, Department of Obstetrics, Gynecology & Reproductive Sciences, University of California, San Francisco, San Francisco, CA, United States

[2]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

[3]Center for Gender and Health Justice, University of California Global Health Institute, Oakland, CA, United States

Corresponding Author:
Ushma D Upadhyay, MPH, PhD
Advancing New Standards in Reproductive Health
Department of Obstetrics, Gynecology & Reproductive Sciences
University of California, San Francisco
Suite 1100
1330 Broadway
San Francisco, CA, 94612
United States
Phone: 1 415 353 4626
Email: ushma.upadhyay@ucsf.edu

Abstract

Background: Even preceding the Supreme Court's 2022 *Dobbs v. Jackson Women's Health Organization* decision, patients in the United States faced exceptional barriers to reach abortion providers. Abortion restrictions disproportionately limited abortion access among people of color, young people, and those living on low incomes. Presently, clinics in states where abortion remains legal are experiencing an influx of out-of-state patients and wait times for in-person appointments are increasing. Direct-to-patient telehealth for abortion care has expanded since its introduction in the United States in 2020. However, the role of this telehealth model in addressing geographic barriers to and inequities in abortion access remains unclear.

Objective: We sought to examine the amount of travel that patients averted by using telehealth for abortion care, and the role of telehealth in mitigating inequities in abortion access by race or ethnicity, age, pregnancy duration, socioeconomic status, rural residence, and distance to a facility.

Methods: We used geospatial analyses and data from patients in the California Home Abortion by Telehealth Study, residing in 31 states and Washington DC, who obtained telehealth abortion care at 1 of 3 virtual abortion clinics. We used patients' residential ZIP code data and data from US abortion facility locations to document the round-trip driving distance in miles, driving time, and public transit time to the nearest abortion facility that patients averted by using telehealth abortion services from April 2021 to January 2022, before the *Dobbs* decision. We used binomial regression to assess whether patients reported that telehealth was more likely to make it possible to access a timely abortion among patients of color, those experiencing food insecurity, younger patients, those with longer pregnancy durations, rural patients, and those residing further from their closest abortion facility.

Results: The 6027 patients averted a median of 10 (IQR 5-26) miles and 25 (IQR 14-46) minutes of round-trip driving, and 1 hour 25 minutes (IQR 46 minutes to 2 hours 30 minutes) of round-trip public transit time. Among a subsample of 1586 patients surveyed, 43% (n=683) reported that telehealth made it possible to obtain timely abortion care. Telehealth was most likely to make it possible to have a timely abortion for younger patients (prevalence ratio [PR] 1.4, 95% CI 1.2-1.6) for patients younger than 25 years of age compared to those 35 years of age or older), rural patients (PR 1.4, 95% CI 1.2-1.6), those experiencing food insecurity (PR 1.3, 95% CI 1.1-1.4), and those who averted over 100 miles of driving to their closest abortion facility (PR 1.6, 95% CI 1.3-1.9).

Conclusions: These findings support the role of telehealth in reducing abortion-related travel barriers in states where abortion remains legal, especially among patient populations who already face structural barriers to abortion care. Restrictions on telehealth abortion threaten health equity.

(JMIR Public Health Surveill 2023;9:e45671) doi:10.2196/45671

KEYWORDS

telehealth; abortion; spatial analysis; health equity; barriers; abortion access; legal; health equity; young people; remote; rural

Introduction

Distance has long been understood to be a key barrier to abortion access in the United States. The distance that patients must travel to the nearest abortion provider has increased substantially in the wake of the July 2022 US Supreme Court's *Dobbs v. Jackson Women's Health Services* decision, which has led to abortion being banned in at least 14 states [1,2]. As a result, providers in protected access states—states where abortion care remains legally available—have experienced an influx of out-of-state patients, which is increasing wait times to care, and may force abortion seekers to travel farther than they would if they could obtain care at the facility closest to their home [3].

Prior studies have found that the further a patient must travel to a provider, the less likely they are to obtain a desired abortion [4,5]. National data indicate that longer travel distances to care due to a spike in antiabortion restrictions in the United States in recent years have contributed to the decreased abortion rate [6]. As of 2019, 20% of people who can get pregnant in the United States lived 43 or more miles from an abortion provider [7] and as of 2018, the United States had 27 "abortion deserts," that is, major cities with no abortion provider within 100 miles [8]. These geographic barriers to abortion facilities exist even within states where abortion is legal [9].

Research has also demonstrated that the impacts of state abortion bans are experienced unequally. Increased distance poses particular challenges for low-income patients due to the lower likelihood of car ownership, loss of wages from the time needed off work, transportation costs for gas or transit fare, as well as the cost of lodging and childcare [10]. People living in rural areas, the South, Midwest, and Mountain West; Black and Hispanic individuals; and minors face greater distances to reach an abortion provider [11,12]. Since the *Dobbs* decision, travel times to abortion facilities have increased disproportionately for Black and Indigenous populations [2].

However, direct-to-patient telehealth abortion may greatly mitigate existing geospatial inequities in abortion access [10,13], despite only being legal in 24 states and Washington, DC as of September 2023 (although more states permit abortion care but prohibit telehealth) [14,15]. In direct-to-patient telehealth abortion, a clinician can interact with patients remotely through videoconferencing or secure messaging and dispense abortion medications via mail-order pharmacies. Prior to 2020, telehealth for abortion was limited to clinic-to-clinic models with remote patient-provider interactions that took place at medical facilities with ultrasonography and other in-person tests [16-20]. Early in the COVID-19 pandemic, professional organizations including The American College of Obstetricians and Gynecologists endorsed a telehealth abortion model [21]. Recent studies have demonstrated the safety and effectiveness of this model of abortion care that can take place without any visits to a clinic or medical facility [22-25]. However, existing research has not captured the travel averted due to telehealth access.

In addition to the potential benefits of direct-to-patient telehealth (subsequently referred to as telehealth) in reducing abortion-related travel, the role of these models in enabling people to obtain a wanted, timely abortion is not well understood. A study conducted in the Pacific Northwest found that patients who used telehealth abortion services resided further from an abortion facility than those who accessed in-clinic care and that some patients chose telehealth even when a clinic is geographically convenient, indicating a preference for telehealth among some patients who obtained an abortion [26]. It is critical to understand who benefits most from abortion models that allow patients to carry out the entire abortion process from home. Few studies have examined the extent to which patients choose telehealth abortion when they have access to in-clinic abortion or whether telehealth is making an otherwise inaccessible abortion possible, therefore reducing barriers to access.

Therefore, we aimed to estimate the amount of travel patients avert by using a telehealth service for medication abortion among a sample of patients who obtained abortions who overwhelmingly resided in states that allow telehealth for abortion care. We also aimed to understand the role that telehealth plays in enabling people to obtain a timely abortion, especially for those who face the greatest barriers to in-clinic care.

Methods

Overview

We used data from the California Home Abortion by Telehealth (CHAT) Study, a study that examines medication abortions provided via telehealth from 3 US virtual clinics, defined as telehealth abortion clinics without brick-and-mortar facilities: Choix, Hey Jane, and Abortion on Demand. While the study was initiated in California, it expanded as the virtual clinics expanded their services to 20 states and Washington, DC. We obtained anonymized clinical chart data from patients who obtained abortion care from these clinics for a defined period between April 2021 and January 2022. At the time of consent to care with the virtual clinic, patients provided permission for the telehealth provider to share their anonymized clinical records with researchers in accordance with each clinic's privacy policy. All clinical data received were deidentified apart from ZIP code (United States postal codes) and dates of service.

In addition, each virtual clinic invited patients who were approved for abortion care during a defined period between June 2021 and January 2022 to participate in a series of 3 surveys on the abortion provider's telehealth platform. Directly following the abortion intake, patients were directed to a page containing detailed information about the study and, if interested, patients provided electronic informed consent prior to their participation. The first survey was administered directly following abortion intake, after obtaining informed consent. Two follow-up surveys were administered within 4 weeks after abortion intake. All surveys were completed on the virtual clinics' telehealth platforms to maximize continuity and survey completion and minimize disruption for participants. Virtual clinics continued to invite patients to complete surveys until approximately 400 patients from each completed all 3 surveys. These participant surveys were linked with their clinical records, and all data were standardized and stored on a secure Research Electronic Data Capture (REDCap; Vanderbilt University) server [27]. We excluded records for abortions where the patient did not take the abortion medications, surveys that were not matched to a clinical chart, and patients with invalid ZIP codes.

Measures

Our key independent variables of interest reflected the amount of travel to the nearest abortion provider averted by using telehealth abortion services. Patient's ZIP codes were recorded in both the clinical chart data and in the surveys. For cases where the clinical chart ZIP code (used for mailing the medications) differed from the ZIP code reported in the survey, we retained the survey ZIP codes with the rationale that some patients may use addresses besides their home addresses for medication delivery. Using ArcGIS (Esri), we first calculated the centroid—the center point of all coordinate points inside the ZIP code polygon—of each unique ZIP code of patients' residences. Next, we used the Advancing New Standards in Reproductive Health 2021 Abortion Facility Database to establish the coordinates of publicly advertising US abortion facilities [28]. We used the ArcGIS Find Nearest feature to calculate the driving distance between the coordinates of the centroid of each patient's ZIP code and the coordinates of the closest abortion facility in miles. We doubled the calculation of 1-way driving distance in miles to calculate the total round-trip driving distance averted in miles. This ArcGIS calculation also generated 1-way travel time in minutes without traffic, which we doubled to develop estimates of round-trip travel time averted and converted to hours. We calculated public transit travel time in hours between patients' ZIP codes and abortion providers' geographic coordinates using the gmapsdistance package in R (R Core Team) [29].

Our outcome of interest was a measure that reflected whether telehealth made it possible to access an abortion in a timely manner (Multimedia Appendix 1). The original item was phrased: "If you didn't have an abortion through telehealth, what would have happened?" Response options were: "I would have gotten an abortion at a clinic soon;" "I would have gotten an abortion at a clinic, but it would have been a while;" "I would have continued the pregnancy;" "I don't know what would have happened;" or "something else." If the patient responded with "something else," they could provide free-text responses. These free-text responses included the following themes: patients would have attempted to self-manage their abortion (primarily seeking abortion pills through other channels), traveled out of state, or sought abortion care at a clinic without a specified time frame. We dichotomized this variable to reflect whether the patient perceived that telehealth made it possible to have an abortion in a timely manner ("I would have gotten an abortion at a clinic, but it would have been a while;" "I would have continued the pregnancy;" "I don't know what would have happened;" or "something else"), versus not ("I would have gotten an abortion at a clinic soon").

We also examined patient characteristics including: patient age at the time of abortion screening (younger than 18, 18-24, 25-29, 30-34, and ≥35 years); pregnancy duration (<35, 35-49, 50-62, and ≥63 days); self-reported race or ethnicity (Asian, Native Hawaiian, or Pacific Islander; Black; Hispanic or Latinx; White; multiracial, Native American, American Indian, or Alaska Native; Middle Eastern or North African; or unknown); urban versus suburban or rural residence based on the rural-urban commuting area codes corresponding to the patient's ZIP code [30]; and the round-trip driving distance to the nearest abortion facility (<5, 5-24, 25-49, 50-99, and ≥100 miles). In regression analyses, we collapsed patient age (younger than 24, 25-29, 30-34, ≥35 years) due to small cell sizes.

Data Analysis

We first described the median, IQR, and total amounts of round-trip driving distance, driving time, and public transit time averted using telehealth for abortion across the clinical chart sample. Next, we calculated median driving distances by the patient characteristic examined and tested for differences in driving distances by each characteristic using Kruskal-Wallis tests.

We used bivariate binomial regression to estimate prevalence ratios (PRs). We then calculated marginal estimates from these binomial regression models to estimate the proportion of participants within each category for whom telehealth made it possible to have an abortion in a timely manner as prevalence percentages (PPs). Analyses were conducted using ArcGIS Online, Stata (version 17.0; StataCorp), and RStudio (version 2022.10.0; Posit, PBC).

Ethical Considerations

This study was approved by the University of California, San Francisco institutional review board (20-32951). Clinical chart data that were deidentified except for patients' ZIP codes and dates of service were obtained from participating virtual clinics in accordance with their privacy policies. The subsample of patients who participated in detailed surveys provided electronic informed consent for their participation and were remunerated with a US $50 electronic debit card after the completion of the final survey. Clinical chart and survey data from participating virtual clinics were standardized and stored on a secure REDCap server [27].

Results

Overview

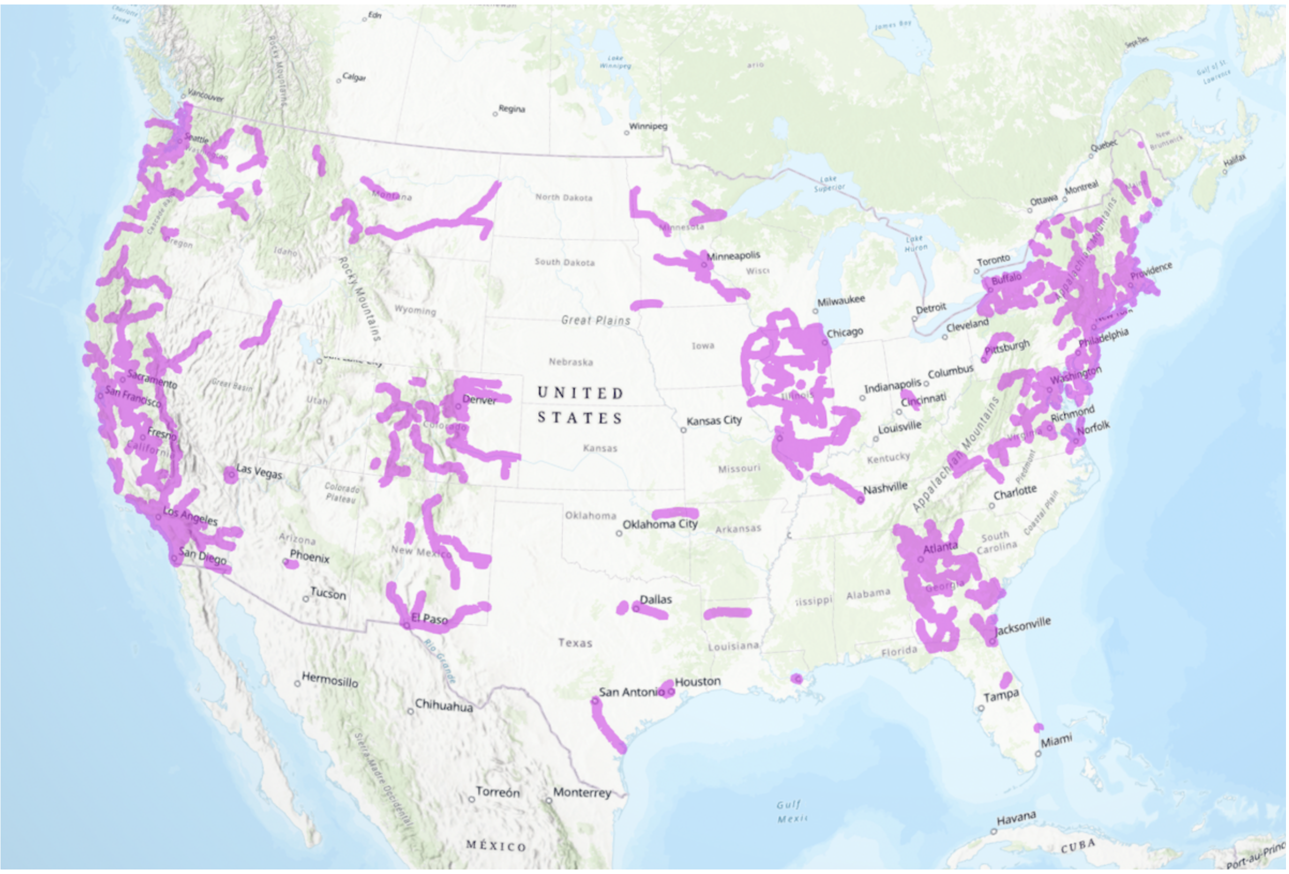
We obtained records for 6154 abortions provided by the virtual clinics between April 2021 and January 2022. Of those, 120 did not take medications, and 7 patients listed ZIP codes from which driving and transit distances could not be calculated, leaving 6027 records included overall. Among these, 1600 patients also participated in CHAT Study baseline surveys. The

ZIP codes in the dataset corresponded to 31 states and Washington, DC.

Travel Averted

The 6027 patients in this analysis averted a median of 10 (IQR 4.5-26.0, range 0.13-566) miles of round-trip driving travel and 25 minutes (IQR 14.1-46.1 minutes, range <1 minute-9 hours) of driving time each. Driving routes averted are depicted in Figure 1 and illustrate the vast distances averted and the distribution of these routes by region. This corresponded to a total of 162,663 miles and 4195 hours of driving averted across the sample (Multimedia Appendix 2). Next, we examined the total public transit time averted for the 5116 (85%) patients for whom public transit routes were available between patients' home ZIP codes and the nearest abortion facility. Patients in the sample averted a median of 1 hour 25 minutes (IQR 45.6 minutes-2 hours 30 minutes, range 2 minutes-48 hours) and a total of 11,720 hours of public transit time. Among the subsample of 1600 patients who completed surveys, patients averted a median of 10 (IQR 4.5-24.6) miles and 25 (IQR 14.3-44.1) minutes of driving, and 1 hour 26 minutes (IQR 46.5 minutes-2 hours 26 minutes) of public transit time. In total, this group of survey participants averted 41,746 miles and 1096 hours of driving, and 3070 hours of public transit time.

Figure 1. Driving routes to the closest abortion facility averted through telehealth (N=6027).



Sample Description

Characteristics of the clinical chart sample and of the survey participants are presented in Table 1. Among the survey subsample, most patients were 25-34 (51%, n=823) years of age and had pregnancy durations between 5 and 7 weeks (55%, n=885). Over half (53%, n=841) identified as White and in the remaining sample 13% (n=203) identified as Hispanic or Latinx; 9% (n=148) identified as Black; 7% (n=105) as Asian, Native Hawaiian, or Pacific Islander; and 19% (n=303) reported more than 1 race or ethnicity, another race or ethnicity, or their race or ethnicity was unknown. In total, 91% (n=1454) of the sample resided in an urban area.

Next, we examined whether patients perceived that telehealth made it possible to have an abortion in a timely manner. Approximately one-fourth (23%, n=371) thought they would have had an abortion at a clinic, but it would have taken some time. Meanwhile, 17% (n=268) did not know what would have happened, and 2% (n=29) said they would have continued the pregnancy. The other 57% (n=903) of the sample thought they would have gotten an abortion at a clinic soon.

We then evaluated differences in travel averted by patient characteristics (Table 2). White patients averted the most driving to the nearest abortion facility (median 12, IQR 5-30 miles), while Asian, Native Hawaiian, or Pacific Islander patients averted the least (median 8, IQR 3-16 miles). Urban residents averted a median of 9 (IQR 4-19) miles of driving time, while those living in rural areas averted a median of 83 (IQR 47-141) miles. The driving miles patients averted did not differ by age, pregnancy duration, or food insecurity.

Table 1. Description of the clinical chart sample and survey subsample.

Factor	Clinical chart sample (n=6027), n (%)	Survey subsample (n=1600), n (%)
Age (years)		
<18	30 (0.5)	8 (0.5)
18-24	1458 (24.2)	448 (28.0)
25-34	3088 (41.2)	823 (51.4)
≥35	1451 (24.1)	321 (20.1)
Pregnancy duration at abortion intake (days)		
<35	1736 (28.8)	448 (28.0)
35-49	3333 (55.3)	885 (55.3)
50-62	825 (13.7)	229 (14.3)
≥63	128 (2.1)	38 (2.4)
Race or ethnicity		
Asian, Native Hawaiian, or Pacific Islander	270 (4.5)	105 (6.6)
Black	413 (6.9)	148 (9.2)
Hispanic or Latinx	339 (5.6)	203 (12.7)
White	2488 (41.3)	841 (52.6)
Multiracial, Native American, American Indian, or Alaska Native; Middle Eastern or North African; or unknown	2517 (41.8)	303 (18.9)
Food insecurity in past month		
No	—[a]	1148 (72.8)
Yes	—	428 (27.2)
Residence		
Suburban or rural	593 (9.8)	146 (9.1)
Urban	5434 (90.2)	1454 (90.9)
Round-trip driving miles averted (miles)		
<5	1705 (28.3)	437 (27.3)
5-24	2765 (45.9)	775 (48.4)
25-49	698 (11.6)	165 (10.3)
50-99	467 (7.8)	120 (7.5)
≥100	392 (6.5)	103 (6.4)
If you didn’t have an abortion through telehealth, what would have happened?		
I would have gotten an abortion at a clinic soon	—	903 (56.9)
I would have gotten an abortion at a clinic, but it would have been a while	—	371 (23.4)
I would have continued the pregnancy	—	29 (1.8)
Something else	—	15 (0.9)
I don’t know what would have happened	—	268 (16.9)

[a] —: not available.

Table 2. Round-trip driving distance averted, by patient characteristics (N=1600).

	Driving miles averted	
	Median (IQR)	*P* value
Age (years)		.54
<18	14.5 (5.1-46.2)	
18-24	9.5 (4.0-24.3)	
25-34	10.2 (4.4-24.0)	
≥35	11.4 (5.1-25.6)	
Pregnancy duration at abortion intake (day)		.12
<35	11.1 (4.7-25.0)	
35-49	10.3 (4.6-23.4)	
50-62	8.8 (3.7-24.5)	
≥63	16.0 (6.3-57.5)	
Race or ethnicity		.001
Asian, Native Hawaiian, or Pacific Islander	7.7 (3.4-15.6)	
Black	9.8 (4.9-21.5)	
Hispanic or Latinx	8.3 (5.0-19.3)	
White	12.0 (4.7-30.2)	
Multiracial, Native American, American Indian, Alaska Native, Middle Eastern or North African, or unknown	9.2 (4.4-19.4)	
Food insecurity in past month		.96
No	10.3 (4.6-23.6)	
Yes	10.2 (4.5-25.5)	
Residence		<.001
Urban	9.3 (4.3-19.4)	
Suburban or rural	82.6 (47.2-140.7)	

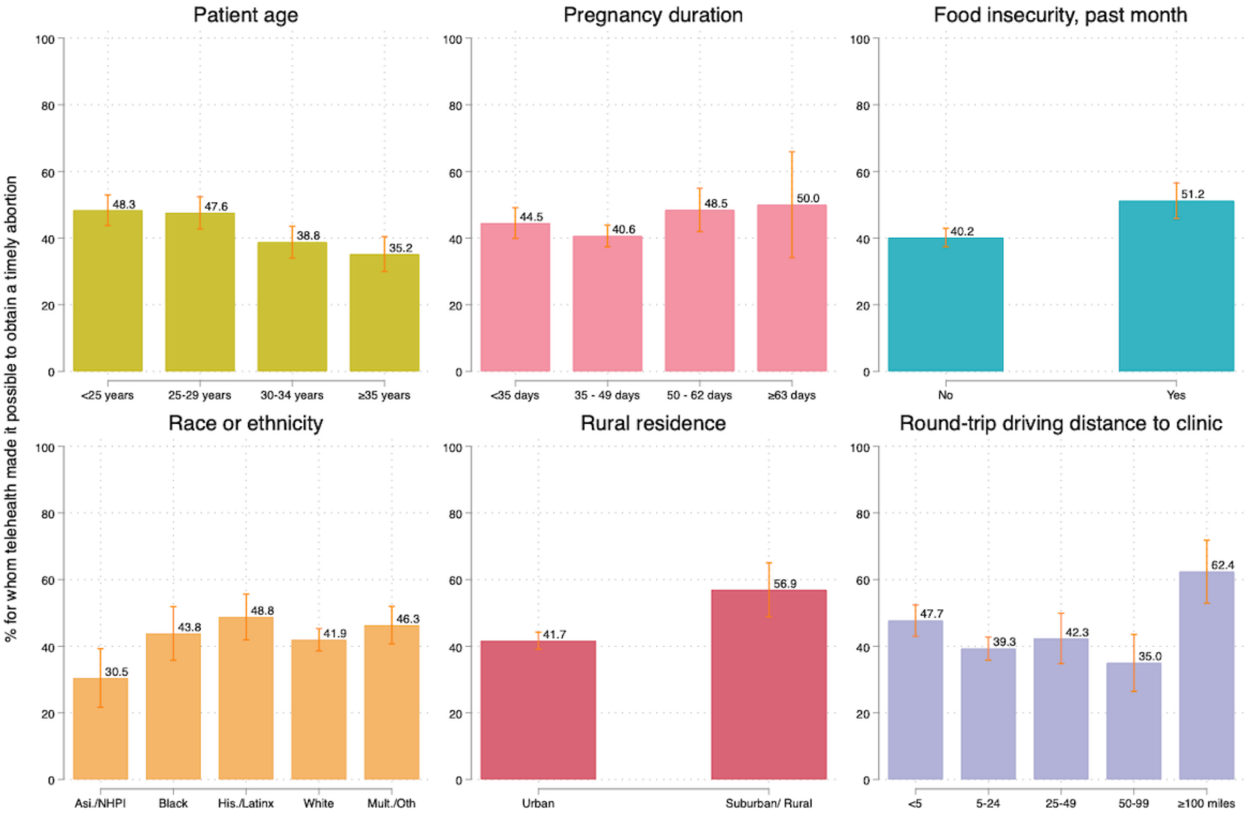
The Role of Telehealth in Abortion Access

Next, we examined the differences in whether telehealth made it possible to obtain an abortion in a timely manner by patient characteristics and by the amount of travel averted (Table 3 and Figure 2). Compared to patients who were 35 years of age or older (PP 35.2%), telehealth was more likely to make abortion possible for patients 24 years of age or younger (PP 48.3%, PR 1.4, 95% CI 1.2-1.6) and for those of 25-29 years of age (PP 47.6%, PR 1.4, 95% CI 1.1-1.6). Patients who experienced food insecurity in the past month were more likely than those who had not experienced food insecurity to have reported that telehealth made a timely abortion possible (PP 51.2% vs 40.2%, PR 1.3, 95% CI 1.1-1.4). When we examined differences by race or ethnicity, patients who were Asian, Native Hawaiian, or Pacific Islander were less likely to report that telehealth made it possible to obtain a timely abortion than White patients (PP 30.5% vs 41.5%, PR 0.7, 95% CI 0.5-1.0). For patients who lived in rural areas, telehealth was more likely to make timely abortion care possible than those who resided in urban areas (PP 56.9% vs 41.7%, PR 1.4, 95% CI 1.2-1.6). Compared to those who averted 5-25 miles of driving to reach an abortion provider (PP 39.3%), those who averted less than 5 miles of round-trip driving (PP 47.7%, PR 1.2, 95% CI 1.1-1.4) or more than 100 miles of round-trip driving (PP 62.4%, PR 1.6, 95% CI 1.3-1.9) were more likely to perceive that telehealth made a timely abortion possible than those that averted less travel. The proportion of patients for whom telehealth made it possible to have an abortion soon was similar across pregnancy duration categories.

Table 3. Associations between patient characteristics and whether telehealth made it possible to obtain a timely abortion (N=1586).

Overall	Prevalence ratio (95% CI)	Marginal prevalence estimates (95% CI)
Patient age (years) at abortion intake		
<25	1.4 (1.2-1.6)	48.3 (43.7-52.9)
25-29	1.4 (1.1-1.6)	47.6 (42.8-52.4)
30-34	1.1 (0.9-1.3)	38.8 (34.0-43.5)
>34	Reference category	35.2 (30.0-40.4)
Pregnancy duration at abortion intake (days)		
<35	Reference category	44.5 (39.9-49.1)
35-49	0.9 (0.8-1.0)	40.6 (37.4-43.9)
50-62	1.1 (0.9-1.3)	48.5 (42.0-55.0)
≥63	1.1 (0.8-1.6)	50.0 (34.1-65.9)
Food did not last in the last month		
No	Reference category	40.2 (37.4-42.9)
Yes	1.3 (1.1-1.4)	51.2 (45.9-56.5)
Race or ethnicity		
Asian, Native Hawaiian, or Pacific Islander	0.7 (0.5-1.0)	30.5 (21.7-39.3)
Black	1.0 (0.9-1.3)	43.8 (35.8-51.9)
Hispanic or Latinx	1.2 (1.0-1.4)	48.8 (41.9-55.6)
White	Reference category	41.9 (38.6-45.3)
Multiracial, other, or unknown	1.1 (1.0-1.3)	46.3 (40.7-52.0)
Rural residence		
Urban	Reference category	41.7 (39.1-44.2)
Suburban or rural	1.4 (1.2-1.6)	56.9 (48.9-65.0)
Round-trip driving miles averted (miles)		
<5	1.2 (1.1-1.4)	47.7 (43.0-52.4)
5-24	Reference category	39.3 (35.8-42.8)
25-49	1.1 (0.9-1.3)	42.3 (34.7-49.9)
50-99	0.9 (0.7-1.2)	35.0 (26.5-43.5)
≥100	1.6 (1.3-1.9)	62.4 (52.9-71.8)

Figure 2. Marginal estimates of the proportion for whom telehealth made it possible to obtain a timely abortion, by patient characteristics, calculated from multivariable binomial regression (N=1586). NHPI: Native Hawaiian or Pacific Islander.



Discussion

Principal Findings

In this study, we found that direct-to-patient telehealth can be an important tool to overcome disparities in abortion access. While the expansion of telehealth is key to improving health equity broadly, abortion care is largely siloed to abortion facilities, which are few and far between and are now closing in record numbers [31]. Therefore, the impact of telehealth on improving equitable abortion access is even greater than for other health care services.

We found that using a telehealth abortion model can avert substantial driving distance and time for those with access to a car and travel time for those who would have used public transit. Our sample drew overwhelmingly from urban areas and from protected access states. Even so, while most of our sample resided close to an abortion facility, 15% lived 25 miles or further, and 7% lived 50 miles or further. Therefore, telehealth can expand the geographic reach of abortion services in states where abortion remains legal.

Telehealth abortion fits within an environmental justice framework, which views telehealth as a means of reducing the carbon footprint of health care services and space as a key component of health disparities [32,33]. Marginalized communities disproportionately bear the brunt of both unequal health policies and the unequal impacts of climate change. This analysis elucidates the connection between environmental justice and reproductive justice by examining abortion access within this framework of geospatial and environmental inequality.

These results suggest that telehealth abortion is poised to address health equity concerns across age, socioeconomic status, and geospatial location. Young people face unique logistical and privacy challenges, many of which can be alleviated by using telehealth, which allows patients to maintain privacy and avoid travel [34]. For low-income patients, the cost of abortion is often a substantial barrier to care, which is only exacerbated when coupled with the costs of associated travel [35]. Telehealth abortion has the potential to address geographic inequities in access to care for those living in abortion deserts or rural areas within protected access states. These findings on telehealth abortion fit within a broader literature that views telehealth as beneficial for reducing health care costs [36], serving underserved and rural patients [37,38], and easing the convenience and comfort of health care provision [39,40].

Compared to the national population of patients who obtained abortions during the same time period, patients in our sample who obtained abortions by telehealth were more likely to be older, White, and have higher socioeconomic statuses [41]. Thus, telehealth abortion may be reaching a subset of patients who are more resourced or have higher technology literacy. However, when patients were asked what would have happened had they not obtained the telehealth abortion, we found that for nearly half, telehealth made it possible to obtain timely abortion care. Telehealth was more likely to play this instrumental role in obtaining an abortion among patient populations who are known to face the most structural barriers to abortion care, such as younger people, those experiencing food insecurity, those residing in rural areas, and those who resided far from an abortion facility [2,42]. Telehealth was less likely to make timely abortion care possible for Asian patients than White patients, perhaps due to urban residence and proximity to abortion providers. Given that telehealth abortion care is situated within a highly unequal health care landscape, there is interest in its role in mitigating inequities in abortion access [43,44]. Our results suggest that telehealth is playing different roles among different demographic groups within our sample. More specifically, while currently, White, urban patients might prefer telehealth due to its convenience or privacy, many patients from groups facing structural barriers to abortion care, such as people of color, those living on lower incomes, those who are younger, and those residing in rural areas, face greater distances to abortion facilities and thus may rely on telehealth for abortion access. These findings suggest that telehealth is key to promoting equitable access to abortion care in the United States, and expanding access to telehealth would be key for these groups.

We found that for patients who resided more than 100 miles from an abortion facility, telehealth was more likely to make it possible to access timely abortion care. This finding echoes prior research that has documented that geospatial barriers can determine whether patients obtain wanted abortions [4,5]. These results bolster the need for alternatives to in-clinic abortion care, especially as geographic access to abortion facilities grows increasingly unequal.

Limitations

This paper has several limitations. Our use of ZIP codes as the location of patient residences reduces the accuracy of the travel distances and travel times we calculated. This likely resulted in overestimated distances and times for some patients and underestimated distances and times for others; however, the net misclassification remains unknown [45]. Only one-fourth of patients elected to participate in the surveys, which may have introduced selection bias and limited the generalizability of these findings to the broader population of virtual clinic patients. Some patients from restricted access states who traveled for abortion care or used various mail-forwarding techniques may have chosen not to disclose their true home ZIP codes. Further, this analysis was conducted on data collected prior to the *Dobbs* decision in June 2022, after which 14 states have banned abortion care. Given the increased wait times for abortion care that have resulted from widespread abortion restrictions, many patients who seek in-person abortion care may need to travel further than their closest abortion facility. Therefore, our analysis may have underestimated the travel averted that we documented in this study. The overwhelming majority of patients in our sample resided in states that allow telehealth for abortion care. Most other states have laws that prohibit telehealth for abortion [15]. Our analysis, therefore, does not capture the full range of travel that could be averted if telehealth abortion care were available across the country. While we did not find that telehealth was more likely to make a critical difference in obtaining an abortion for people of color, this finding may be limited by the racial composition of our sample, in which Black and Hispanic patients, who face the greatest barriers to abortion care, were greatly underrepresented compared to the population of US patients who obtained abortions [41]. Despite these limitations, this study highlights the benefits of telehealth services by being among the first to use geospatial analysis to examine the travel averted, across both driving and public transit transportation, from using a telehealth model for abortion care.

Conclusions

This paper makes important contributions to a growing body of work on telehealth abortion and the broader literature on telehealth. Our findings support the role of telehealth abortion in reducing travel distance, time, and costs, which in turn mitigates the inequities embedded in abortion provision based on geospatial location, socioeconomic status, and local abortion policy. We found that telehealth can play a key role in accessing an otherwise unobtainable or delayed abortion, especially for marginalized patient populations. Future research should examine how telehealth abortion services can be tailored to improve health equity and digital inclusion. These findings may help direct future policy on expanding access to telehealth abortion, especially in the wake of the *Dobbs* decision, which is continuing to exacerbate longstanding inequities in abortion access. While states that ban abortion are unlikely to support telehealth for abortion, there are 6 states that permit abortion but restrict telehealth abortion care [14]. States invested in health equity that want to safeguard access in an increasingly restricted abortion landscape should legalize telehealth abortion. There is a need for federal actions to protect patients who seek and providers who dispense telehealth abortion care across state lines.

Acknowledgments

The authors thank Linda Shin, Kelly Song, and Lisa Peters for their contributions to data collection and management for the CHAT Study. We appreciate input on the design of the CHAT Study from Francine Coeytaux, Karen Meckstroth, and Elisa Wells. We thank Cindy Adam, Mark Adam, Kate Baron, Stephanie Bussmann, Leah Coplon, Lauren Dubey, Lindsay DuBois, Kiki Freedman, Gaby Izarra, Jamie Phifer, and Aisha Wagner for supporting data acquisition. We also thank the Abortion Facility Database team at Advancing New Standards in Reproductive Health for facilitating access to the database, including Rosalyn Schroeder and Shelly Kaller. This research was supported by the BaSe Family Fund, Erik E. and Edith H. Bergstrom Foundation, the Lisa and Douglas Goldman Fund, the Kahle/Austin Foundation, the Isabel Allende Foundation, Jess Jacobs, and Preston-Werner Ventures. LRK was funded in part by training grants from the National Institute on Minority Health and Health Disparities of the National Institutes of Health (award T32MD015070) and the National Institute on Child Health and Human Development (award F31HD111277) for the duration of the study. The funders had no role in the study design, data collection, analysis, the writing of the report, or the decision to submit the paper for publication.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to privacy concerns of patients who obtained abortions and the changing legal status of abortion care but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definition of primary outcome measure.
[PNG File , 165 KB - publichealth_v9i1e45671_app1.png]

Multimedia Appendix 2
Density plot of travel averted by using telehealth.
[PNG File , 230 KB - publichealth_v9i1e45671_app2.png]

References

1. Myers C, Jones R, Upadhyay U. Predicted changes in abortion access and incidence in a post-Roe world. Contraception 2019;100(5):367-373. [doi: 10.1016/j.contraception.2019.07.139] [Medline: 31376381]
2. Rader B, Upadhyay UD, Sehgal NKR, Reis BY, Brownstein JS, Hswen Y. Estimated travel time and spatial access to abortion facilities in the US before and after the Dobbs v Jackson women's health decision. JAMA 2022;328(20):2041-2047 [FREE Full text] [doi: 10.1001/jama.2022.20424] [Medline: 36318194]
3. Katz J, Miller CC, Sanger-Katz M. Interstate abortion travel is already straining parts of the system. The New York Times. 2022. URL: https://www.nytimes.com/2022/07/23/upshot/abortion-interstate-travel-appointments.html [accessed 2023-10-19]
4. Bearak JM, Burke KL, Jones RK. Disparities and change over time in distance women would need to travel to have an abortion in the USA: a spatial analysis. Lancet Public Health 2017;2(11):e493-e500 [FREE Full text] [doi: 10.1016/S2468-2667(17)30158-5] [Medline: 29253373]
5. Pleasants EA, Cartwright AF, Upadhyay UD. Association between distance to an abortion facility and abortion or pregnancy outcome among a prospective cohort of people seeking abortion online. JAMA Netw Open 2022;5(5):e2212065 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.12065] [Medline: 35560050]
6. Brown BP, Hebert LE, Gilliam M, Kaestner R. Distance to an abortion provider and its association with the abortion rate: a multistate longitudinal analysis. Perspect Sex Reprod Health 2020;52(4):227-234. [doi: 10.1363/psrh.12164] [Medline: 33332717]
7. Austin N, Harper S. Quantifying the impact of targeted regulation of abortion provider laws on US abortion rates: a multi-state assessment. Contraception 2019;100(5):374-379. [doi: 10.1016/j.contraception.2019.06.003] [Medline: 31202816]
8. Cartwright AF, Karunaratne M, Barr-Walker J, Johns NE, Upadhyay UD. Identifying national availability of abortion care and distance from major US cities: systematic online search. J Med Internet Res 2018;20(5):e186 [FREE Full text] [doi: 10.2196/jmir.9717] [Medline: 29759954]
9. Johns NE, Foster DG, Upadhyay UD. Distance traveled for Medicaid-covered abortion care in California. BMC Health Serv Res 2017;17(1):287 [FREE Full text] [doi: 10.1186/s12913-017-2241-0] [Medline: 28420438]
10. Upadhyay UD, Grossman D. Telemedicine for medication abortion. Contraception 2019;100(5):351-353. [doi: 10.1016/j.contraception.2019.07.005] [Medline: 31356771]
11. Kelly BC, Brewer SC, Hanson HA. Disparities in distance to abortion care under reversal of Roe V. Wade. Utah Women's Health Review. 2022. URL: https://uwhr.utah.edu/disparities-in-distance-to-abortion-care-under-reversal-of-roe-v-wade%ef%bf%bc/ [accessed 2023-10-19]
12. Myers CK, Ladd D. Did parental involvement laws grow teeth? The effects of state restrictions on minors' access to abortion. J Health Econ 2020;71:102302. [doi: 10.1016/j.jhealeco.2020.102302] [Medline: 32135395]
13. Thompson KMJ, Sturrock HJW, Foster DG, Upadhyay UD. Association of travel distance to nearest abortion facility with rates of abortion. JAMA Netw Open 2021;4(7):e2115530 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.15530] [Medline: 34228128]
14. The availability and use of medication abortion. Kaiser Family Foundation. 2023. URL: https://www.kff.org/womens-health-policy/fact-sheet/the-availability-and-use-of-medication-abortion/ [accessed 2023-10-19]
15. Sobel L, Ramaswamy A, Salganicoff A. The intersection of state and federal policies on access to medication abortion via telehealth. Kaiser Family Foundation. 2022. URL: https://www.kff.org/womens-health-policy/issue-brief/the-intersection-of-state-and-federal-policies-on-access-to-medication-abortion-via-telehealth/ [accessed 2023-10-19]
16. Grindlay K, Lane K, Grossman D. Women's and providers' experiences with medical abortion provided through telemedicine: a qualitative study. Womens Health Issues 2013;23(2):e117-e122. [doi: 10.1016/j.whi.2012.12.002] [Medline: 23410620]

17. Grossman D, Grindlay K. Safety of medical abortion provided through telemedicine compared with in person. Obstet Gynecol 2017;130(4):778-782. [doi: 10.1097/AOG.0000000000002212] [Medline: 28885427]
18. Grossman DA, Grindlay K, Buchacker T, Potter JE, Schmertmann CP. Changes in service delivery patterns after introduction of telemedicine provision of medical abortion in Iowa. Am J Public Health 2013;103(1):73-78 [FREE Full text] [doi: 10.2105/AJPH.2012.301097] [Medline: 23153158]
19. Seymour JW, Thompson TA, Milechin D, Wise LA, Rudolph AE. Potential impact of telemedicine for medication abortion policy and programming changes on abortion accessibility in the United States. Am J Public Health 2022;112(8):1202-1211. [doi: 10.2105/AJPH.2022.306876] [Medline: 35830676]
20. Kohn JE, Snow JL, Grossman D, Thompson TA, Seymour JW, Simons HR. Introduction of telemedicine for medication abortion: changes in service delivery patterns in two U.S. states. Contraception 2021;103(3):151-156. [doi: 10.1016/j.contraception.2020.12.005] [Medline: 33359507]
21. Joint statement on abortion access during the COVID-19 outbreak. The American College of Obstetricians and Gynecologists. 2020. URL: https://www.acog.org/en/news/news-releases/2020/03/joint-statement-on-abortion-access-during-the-covid-19-outbreak [accessed 2022-06-13]
22. Chong E, Shochet T, Raymond E, Platais I, Anger HA, Raidoo S, et al. Expansion of a direct-to-patient telemedicine abortion service in the United States and experience during the COVID-19 pandemic. Contraception 2021;104(1):43-48 [FREE Full text] [doi: 10.1016/j.contraception.2021.03.019] [Medline: 33781762]
23. Upadhyay UD, Koenig LR, Meckstroth KR. Safety and efficacy of telehealth medication abortions in the US during the COVID-19 pandemic. JAMA Netw Open 2021;4(8):e2122320 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.22320] [Medline: 34427682]
24. Upadhyay UD, Raymond EG, Koenig LR, Coplon L, Gold M, Kaneshiro B, et al. Outcomes and safety of history-based screening for medication abortion: a retrospective multicenter cohort study. JAMA Intern Med 2022;182(5):482-491 [FREE Full text] [doi: 10.1001/jamainternmed.2022.0217] [Medline: 35311911]
25. Anger HA, Raymond EG, Grant M, Haskell S, Boraas C, Tocce K, et al. Clinical and service delivery implications of omitting ultrasound before medication abortion provided via direct-to-patient telemedicine and mail in the U.S. Contraception 2021;104(6):659-665. [doi: 10.1016/j.contraception.2021.07.108] [Medline: 34329607]
26. Beardsworth KM, Doshi U, Raymond E, Baldwin MK. Miles and days until medical abortion via TelAbortion versus clinic in Oregon and Washington, USA. BMJ Sex Reprod Health 2022;48(e1):e38-e43 [FREE Full text] [doi: 10.1136/bmjsrh-2020-200972] [Medline: 33789954]
27. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
28. Abortion facility database. ANSIRH. URL: https://www.ansirh.org/abortion-facility-database [accessed 2022-09-27]
29. Melo RA, Zarruk D. Package 'gmapsdistance'. gmapsdistance. 2022. URL: https://cran.r-project.org/web/packages/gmapsdistance/gmapsdistance.pdf [accessed 2023-10-19]
30. Rural-urban commuting area codes. USDA Economic Research Service. 2020. URL: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx [accessed 2023-10-19]
31. Kirstein M, Dreweke J, Jones RK, Philbin J. 100 days post-Roe: at least 66 clinics across 15 US states have stopped offering abortion care. Guttmacher Institute. 2022. URL: https://www.guttmacher.org/node/300216/printable/print [accessed 2023-10-16]
32. McKane RG, Satcher LA, Houston SL, Hess DJ. Race, class, and space: an intersectional approach to environmental justice in New York City. Environ Sociol 2018;4(1):79-92. [doi: 10.1080/23251042.2018.1429177]
33. Tsagkaris C, Hoian AV, Ahmad S, Essar MY, Campbell LW, Grobusch L, et al. Using telemedicine for a lower carbon footprint in healthcare: a twofold tale of healing. J Clim Change Health 2021;1:100006 [FREE Full text] [doi: 10.1016/j.joclim.2021.100006]
34. Ralph L, Hasselbacher L. Adolescents and abortion restrictions: disproportionate burdens and critical warnings. J Adolesc Health 2023;73(2):221-223 [FREE Full text] [doi: 10.1016/j.jadohealth.2023.05.002] [Medline: 37455044]
35. Zuniga C, Thompson T, Blanchard K. Abortion as a catastrophic health expenditure in the United States. Womens Health Issues 2020;30(6):416-425. [doi: 10.1016/j.whi.2020.07.001] [Medline: 32798085]
36. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021;17(2):218-221 [FREE Full text] [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
37. Gilman M, Stensland J. Telehealth and Medicare: payment policy, current use, and prospects for growth. Medicare Medicaid Res Rev 2013;3(4):mmrr.003.04.a04 [FREE Full text] [doi: 10.5600/mmrr.003.04.a04] [Medline: 24834368]
38. Rutledge CM, Kott K, Schweickert PA, Poston R, Fowler C, Haney T. Telehealth and eHealth in nurse practitioner training: current perspectives. AMEP 2017;8:399-409 [FREE Full text] [doi: 10.2147/amep.s116071]
39. Powell RE, Henstenburg JM, Cooper G, Hollander JE, Rising KL. Patient perceptions of telehealth primary care video visits. Ann Fam Med 2017;15(3):225-229 [FREE Full text] [doi: 10.1370/afm.2095] [Medline: 28483887]
40. Koenig L, Ko J, Valladares E, Coeytaux F, Wells E, Lyles C, et al. Patient acceptability of telehealth medication abortion care in the United States, 2021-2022: a cohort study. Am J Public Health 2023.

41. Jones RK, Chiu DW. Characteristics of abortion patients in protected and restricted states accessing clinic-based care 12 months prior to the elimination of the federal constitutional right to abortion in the United States. Perspect Sex Reprod Health 2023;55(2):80-85 [FREE Full text] [doi: 10.1363/psrh.12224] [Medline: 37038835]
42. Kaiser FF. What are the implications of the overturning of Roe v. Wade for racial disparities? Kaiser Family Foundation. 2022. URL: https://www.kff.org/racial-equity-and-health-policy/issue-brief/what-are-the-implications-of-the-overturning-of-roe-v-wade-for-racial-disparities/ [accessed 2023-10-19]
43. Thompson TA, Northcraft D, Carrión F. Addressing structural inequities, a necessary step toward ensuring equitable access to telehealth for medication abortion care during and post COVID-19. Front Glob Womens Health 2022;3:805767 [FREE Full text] [doi: 10.3389/fgwh.2022.805767] [Medline: 35368993]
44. Lyles CR, Sharma AE, Fields JD, Getachew Y, Sarkar U, Zephyrin L. Centering health equity in telemedicine. Ann Fam Med 2022;20(4):362-367 [FREE Full text] [doi: 10.1370/afm.2823] [Medline: 35879077]
45. Seymour JW, Milechin D, Upadhyay UD, Wise LA, Rudolph AE. Improving our estimates: assessing misclassification of abortion accessibility in the United States. Ann Epidemiol 2022;76:98-107. [doi: 10.1016/j.annepidem.2022.10.012] [Medline: 36332754]

Abbreviations

CHAT: California Home Abortion by Telehealth Study
PP: prevalence percentage
PR: prevalence ratio
REDCap: Research Electronic Data Capture

Edited by A Mavragani, T Sanchez; submitted 03.03.23; peer-reviewed by J Seymour, Y Mejova; comments to author 12.04.23; revised version received 05.05.23; accepted 26.09.23; published 07.11.23.

Please cite as:
Koenig LR, Becker A, Ko J, Upadhyay UD
The Role of Telehealth in Promoting Equitable Abortion Access in the United States: Spatial Analysis
JMIR Public Health Surveill 2023;9:e45671
URL: https://publichealth.jmir.org/2023/1/e45671
doi:10.2196/45671
PMID:37934583

Original Paper

Epidemic Characteristics, Spatiotemporal Pattern, and Risk Factors of Other Infectious Diarrhea in Fujian Province From 2005 to 2021: Retrospective Analysis

Yixiao Lu[1*], PhD; Hansong Zhu[2*], MMed; Zhijian Hu[3], PhD; Fei He[3], PhD; Guangmin Chen[2], MMed

[1]School of Public Health (Shenzhen), Shenzhen Campus of Sun Yat-sen University, Shenzhen, China

[2]Fujian Provincial Center for Disease Control and Prevention, The Practice Base on the School of Public Health, Fujian Medical University, Fuzhou, China

[3]Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, Fuzhou, China

[*]these authors contributed equally

Corresponding Author:
Guangmin Chen, MMed
Fujian Provincial Center for Disease Control and Prevention
The Practice Base on the School of Public Health
Fujian Medical University
No.386 Chong'an Road
Fuzhou, 350001
China
Phone: 86 591 87552041
Fax: 86 591 87670235
Email: cgmyf34567@163.com

Abstract

Background: Other infectious diarrhea (OID) continues to pose a significant public health threat to all age groups in Fujian Province. There is a need for an in-depth analysis to understand the epidemiological pattern of OID and its associated risk factors in the region.

Objective: In this study, we aimed to describe the overall epidemic characteristics and spatiotemporal pattern of OID in Fujian Province from 2005 to 2021 and explore the linkage between sociodemographic and environmental factors and the occurrence of OID within the study area.

Methods: Notification data for OID in Fujian were extracted from the China Information System for Disease Control and Prevention. The spatiotemporal pattern of OID was analyzed using Moran index and Kulldorff scan statistics. The seasonality of and short-term impact of meteorological factors on OID were examined using an additive decomposition model and a generalized additive model. Geographical weighted regression and generalized linear mixed model were used to identify potential risk factors.

Results: A total of 388,636 OID cases were recorded in Fujian Province from January 2005 to December 2021, with an average annual incidence of 60.3 (SD 16.7) per 100,000 population. Children aged <2 years accounted for 50.7% (196,905/388,636) of all cases. There was a steady increase in OID from 2005 to 2017 and a clear seasonal shift in OID cases from autumn to winter and spring between 2005 and 2020. Higher maximum temperature, atmospheric pressure, humidity, and precipitation were linked to a higher number of deseasonalized OID cases. The spatial and temporal aggregations were concentrated in Zhangzhou City and Xiamen City for 17 study years. Furthermore, the clustered areas exhibited a dynamic spreading trend, expanding from the southernmost Fujian to the southeast and then southward over time. Factors such as densely populated areas with a large <1-year-old population, less economically developed areas, and higher pollution levels contributed to OID cases in Fujian Province.

Conclusions: This study revealed a distinct distribution of OID incidence across different population groups, seasons, and regions in Fujian Province. Zhangzhou City and Xiamen City were identified as the major hot spots for OID. Therefore, prevention and control efforts should prioritize these specific hot spots and highly susceptible groups.

(JMIR Public Health Surveill 2023;9:e45870) doi:10.2196/45870

KEYWORDS

other infectious diarrhea; spatiotemporal pattern; disease cluster; epidemiological trends; spatial autocorrelation; meteorological factors; environmental factors

Introduction

Background

Diarrheal disease is the second leading cause of malnutrition and death in children aged <5 years, with an estimated 1.7 billion yearly cases of childhood diarrheal disease worldwide. China ranked 11th among countries with the highest number of diarrheal deaths among children aged <5 years [1]. Infectious diarrhea is caused by a host of bacterial, viral, and parasitic organisms [2], and other infectious diarrhea (OID) was defined as infectious diarrhea other than cholera, dysentery, and typhoid or paratyphoid fever [3]. In China, bacterial and viral pathogens, such as *Shigella*, rotavirus, enterocathartic *Escherichia coli*, *Campylobacter jejuni*, and *Salmonella*, are the most commonly identified pathogens associated with OID [4].

OID incidence varies by region in China because of the broad spectrum of pathogens; sociodemographic, environmental, and meteorological factors; hygiene; and living habits [5-7]. Moreover, the seasonal pattern and spatial distribution of OID also exhibit regional variations [8,9]. Understanding the role of environmental and meteorological exposures in these variations is critical for developing targeted public health interventions. Therefore, our study sought to investigate the association between county-level environmental and meteorological factors and the occurrence of OID in Fujian Province.

Analyzing a substantial, 17-year time frame allowed us to capture temporal variations, overall trends, and seasonal fluctuations in OID incidence, facilitating the identification of disease patterns over time and potential seasonality drivers, such as specific meteorological factors. For instance, temperature has been linked to increased diarrhea risk, particularly for diarrhea caused by bacterial pathogens [10,11], whereas heavy rainfall following dry periods may exacerbate incidence [12]. In addition, long-term data analysis provides a comprehensive understanding of epidemiological dynamics and aids in the identification of preventive strategies and interventions.

Spatial clustering methods, such as spatial scan statistics, local indicators of spatial association (LISA), and the Getis-Ord local Gi*(d) statistic, are commonly used for predicting disease risk and examining cluster characteristics based on location, size, and disease incidence [13]. In our study, we aimed to address gaps in OID epidemiology in Fujian Province by comprehensively examining its spatiotemporal distribution and identifying local clusters. We used a combined approach that integrated LISA and spatial scan statistics. This combination ensures a comprehensive identification of the different aspects of spatial patterns and logically consistent outcomes [14-16].

Moreover, recognizing the potential spatial clustering issues of OID, our study adopted a multilevel analysis to investigate the influence of sociodemographic and environmental factors. Specifically, we explored the emerging concern in China regarding potential human exposure to heavy metals through the consumption of contaminated food. Studies conducted in coastal areas have highlighted the primary exposure of soils and foods to highly toxic pollutants, such as arsenic, cadmium, and lead [17,18]. Soil and food pollution resulting from mining, industrialization, and agricultural activities can exacerbate this problem, leading to weakened immune systems [19] and increased susceptibility to diarrheal pathogens.

Southern China, including Fujian Province on the southeastern coast, experiences a typical monsoon climate characterized by high humidity, temperature, and rainfall. This is the region where the majority of OID-related public health emergencies have been documented [20]. In addition, factors such as rapid urbanization, the development of forestry and mining, and prevalent raw seafood eating habits have emerged as potential contributors to the recent rise in OID incidence. Despite the significant burden of OID and the consistent ranking of OID among the top 5 class C notifiable diseases in Fujian, a critical gap persists in the understanding of its epidemiological pattern in this region [21]. Although several studies have investigated OID distribution in the southeastern coastal areas of China [22,23], to the best of our knowledge, no comprehensive and systematic research has been conducted on the spatial distribution and clustering of OID at the county level in Fujian. Therefore, there is a need for an in-depth analysis to understand the epidemiological pattern of OID and its associated risk factors in the region.

Objectives

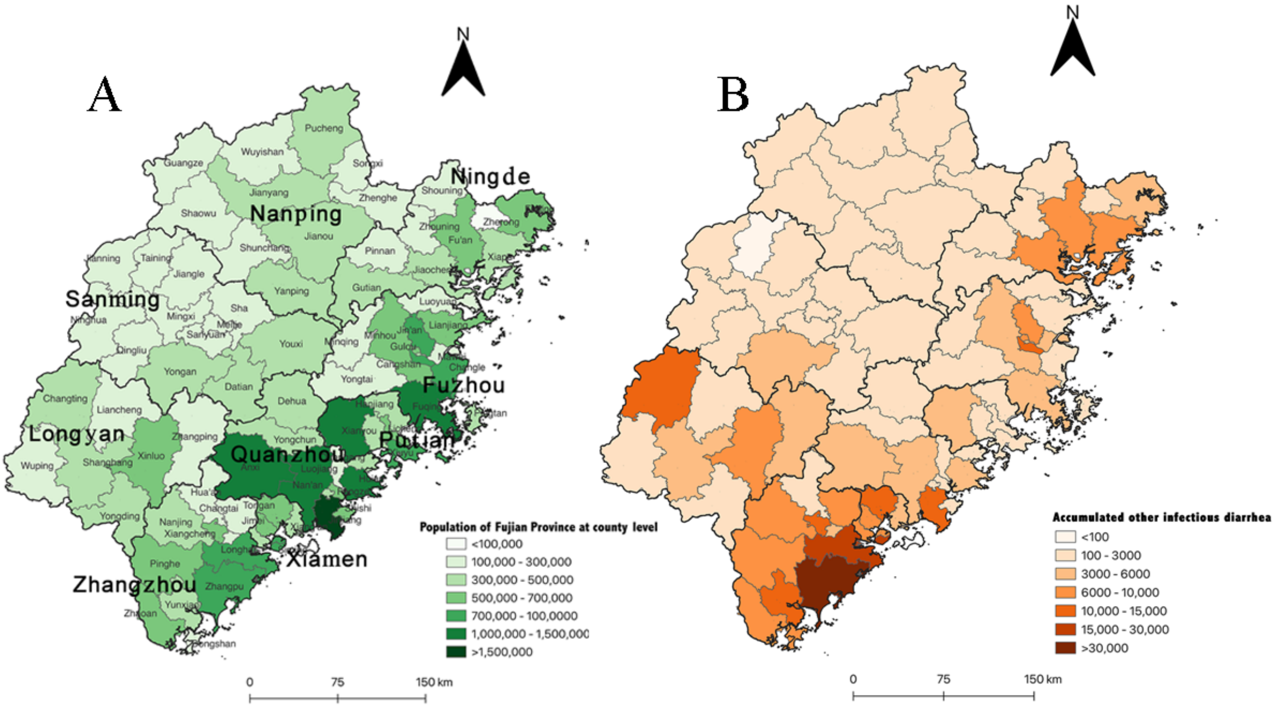
The objective of this study was 2-fold. First, we aimed to provide an overall description of the epidemic distribution patterns and spatiotemporal distribution of OID in Fujian Province from 2005 to 2021. Second, we aimed to explore the linkage between sociodemographic and environmental factors and the occurrence of OID within the study area. By investigating these factors, we sought to enhance our understanding of the epidemiological pattern of OID and inform targeted interventions and preventive strategies to mitigate the impact of OID on public health outcomes in Fujian Province.

Methods

Study Setting

Fujian Province is located on the southeastern coast of China (23°33′–28°20′N, 115°50′–120°40′E) and consists of 6 coastal prefecture-level cities (Nindea, Fuzhou, Putian, Quanzhou, Xiamen, and Zhangzhou) and 6 mountainous cities (Nanping, Sanming, and Longyan), including 85 county-level divisions (Figure 1).

Figure 1. Reported other infectious diarrhea cases at the county level in Fujian Province from January 1, 2005, to December 31, 2010. (A) Population distribution. (B) Accumulated other infectious diarrhea distribution.



Data Sources

The original data on OID in Fujian Province from 2005 to 2021 were obtained from the reporting subsystem of the China Information System for Disease Control and Prevention. OID cases were reported by medical institutions at all levels using the standard report card for infectious diseases, which includes information such as age, sex, occupation, date of onset, classification of disease, registration of residence, and residential address. Residential type is determined by the registration of residence and residential address. Permanent residents are individuals whose registration of residence and residential address are from the same county. The floating population inside city consists of individuals from different counties but the same city. The floating population inside province comprises individuals from different counties and cities within the same province, whereas the floating population outside province includes individuals whose registration of residence is outside Fujian Province. Cases of OID were diagnosed by physicians based on the history of exposure, clinical manifestation, and laboratory testing following the criteria issued by the National Health Commission of the People's Republic of China (WS 271-2007) [24]. Clinically diagnosed cases refer to those with a clinical diagnosis alone, whereas confirmed cases refer to those with both a clinical diagnosis and a pathogen diagnosis.

A total of 402,727 cases were registered in the China Information System for Disease Control and Prevention between January 1, 2005, and December 31, 2021. Suspected cases without a confirmed diagnosis were excluded from the analysis data set, leaving 388,636 confirmed OID cases for inclusion in this study.

The population data used for calculating the incidence rate were sourced from the Fujian Statistical Yearbook, which is officially published on a yearly basis by the Fujian Provincial Bureau of Statistics. County-level data were additionally collected, including demographic factors (permanent residents, urban/rural population, <1-year-old population, and <5-year-old population), economic factors (gross domestic product [GDP] per capita; primary, secondary, and tertiary sector GDP; urbanization level; and highway density), and environmental factors (arsenic, cadmium, chromium, copper, lead, and zinc contents of the soil and fine atmospheric particles with a diameter of ≤2.5 μm [$PM_{2.5}$]).

Statistical Analysis

Time-Series Analysis

The characteristics of OID were summarized by frequency and proportion, and the chi-square test was conducted and odds ratio was calculated to evaluate the differences between subgroups, with a significance level of $P<.05$. A time-series analysis based on the additive decomposition model was performed to estimate the seasonal effects on the reported OID cases in Fujian Province from 2005 to 2021 using the following formula: Xt = seasonal + trend + random. Xt is the number of OID cases, and trend is time expressed in weeks.

To investigate the short-term effects of meteorological factors on OID, we collected daily data on average temperature (°C), maximum temperature (°C), minimum temperature (°C), atmospheric pressure at sea level (mm Hg), average relative humidity (%), and precipitation (mm) from February 2005 to December 2021. The meteorological data were obtained from the daily meteorological data set publicly released by the website Reliable Prognosis [25]. The website provides weather data collected from ground weather stations through the International Free Exchange System for Meteorological Data. The missing data points were imputed using the mean of neighboring values.

To capture the nonlinear relationships between meteorological factors and OID notification cases, we used the generalized additive model (GAM) [26]. We first aggregated weekly values of meteorological factors from the daily data and then averaged the weekly mean values in Fujian Province to derive weekly variables. Then, we detrended the underlying patterns and deseasonalized the weekly OID data. We used a nonparametric spline fitting response model based on a GAM to estimate the effect of meteorological factors on OID, using a Poisson model with the following formula: $y = \alpha + s(x1) + s(x2) + ... + s(xm) + \varepsilon$, incorporates detrended OID data (y); the model intercept (α); and smooth terms (s(x1), s(x2), and s(xm)) for each predictor variable; in addition, a cyclic cubic regression spline basis function (ε) is specified for the time variable. To select smoothing parameters in the GAM, we used the restricted (or residual) maximum likelihood method because the sample size was small.

Spatial Autocorrelation Analysis

The global Moran index was used to detect the spatial autocorrelation of OID cases in Fujian Province [27]. Significant positive spatial autocorrelation of OID cases implies that the distribution of OID cases is more spatially aggregated than a random underlying spatial process. The Moran index is calculated using the following equation: $I = (\mathrm{n}/S_0)(\Sigma^{\mathrm{n}}_{i=1}\Sigma^{\mathrm{n}}_{j=1} W_{i,j}Z_iZ_j)/(\Sigma^{\mathrm{n}}_{i=1}Z_i^2)$, where Z_i is the deviation of an attribute for feature i from its mean (x_i –`X), $W_{i,j}$ is the spatial weight between features i and j, n is the total number of features, and S_o is the aggregate of all spatial weights.

The global Moran I ranges from −1 to 1, with 0 indicating the null hypothesis of no clustering. A higher positive value of Moran I indicates the clustering of OID cases, whereas a lower negative value implies that neighboring areas are characterized by dissimilar OID cases [28]. The queen contiguity method was applied to define a weight matrix specifying the spatial relationships in consideration of the irregularity in the shapes and sizes of county-level divisions in Fujian [29].

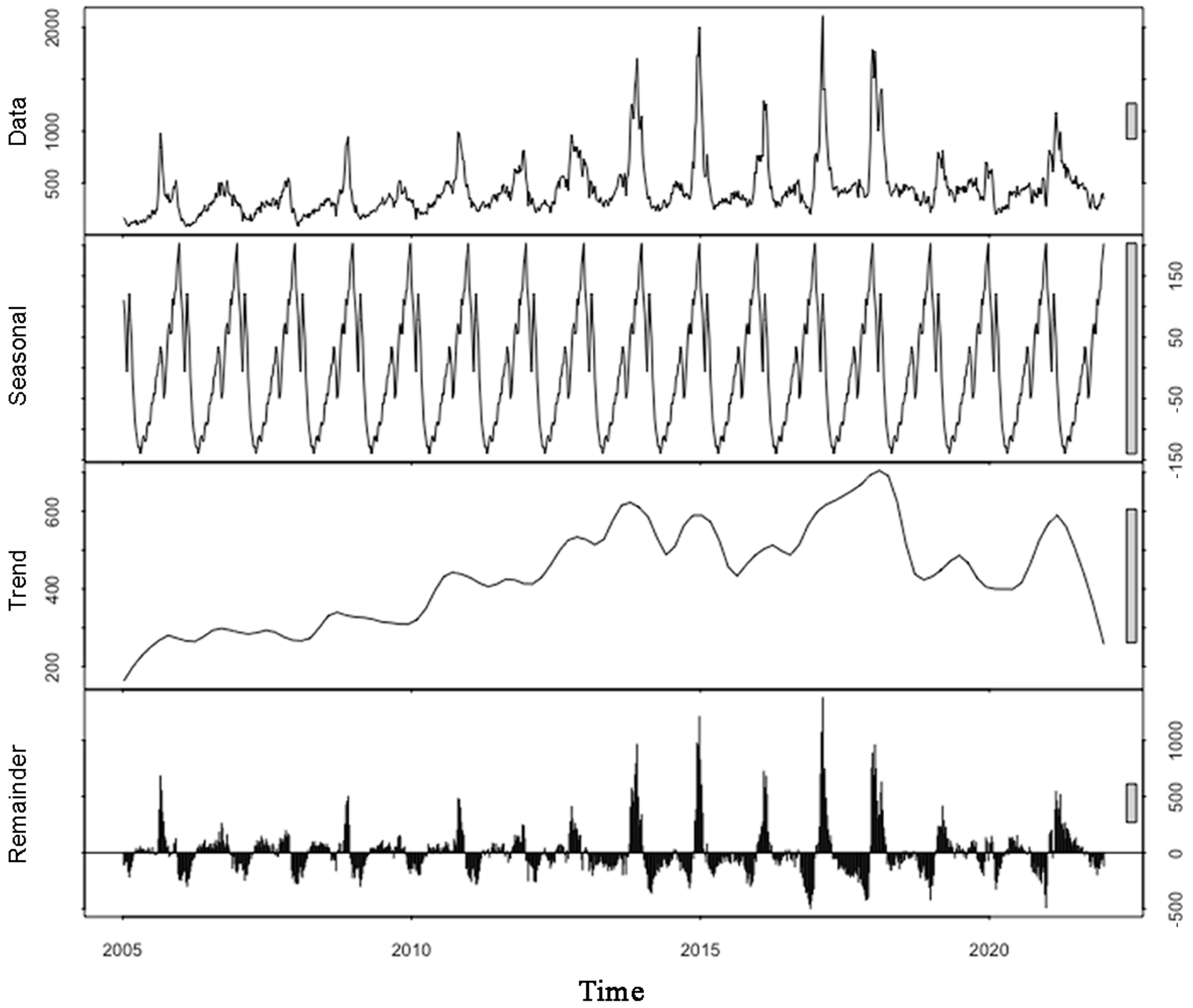
In addition to global measures, local variations of OID cases' spatial patterns were analyzed using a version of local Moran I, and LISA was computed for each location [30]. LISA was performed with 999 permutations to find core clusters and outliers of counties with extreme OID cases unexplained by random variation. As a result, LISA detects clusters as "hot spots" (high-high), "cold spots" (low-low), and "spatial outliers" (high-low or low-high) [31]. A P value <.05 for the cluster or outlier is considered statistically significant.

Retrospective Space-Time Scan Statistics

The SatScan software (version 9.6) applying the Kulldorff method of retrospective space-time analysis was used to identify potential clusters with high rates that significantly exceeded the OID incidence of nearby regions (P<.05). The space-time scan statistics use geographical information and time-periodic variable to explore temporal persistence and possible spatial accumulation. An infinite number of discrete, cylindrical windows are created by the space-time scan, with a circular geographic base and height corresponding to time [32]. Each cylindrical window was evaluated as a potential OID space-time cluster. Assuming that the data followed a Poisson distribution, the distribution and statistical significance of the space-time clusters were analyzed through Monte Carlo replication under the null hypothesis with the default 999 replications to ensure adequate power for defining clustering. The most likely cluster and a number of secondary potential clusters were reported. The relative risk of OID in each cluster was calculated to evaluate the risk of OID in the clustered areas using the following formula [33]: $RR= (c/e)/((C-c)/(C-e))$, where c is the total number of cases in a cluster, e is the number of expected cases in a cluster, and C is the total number of observed cases in Fujian Province.

The OID data from 2005 were used to determine the spatial scanning windows by increasing from 1% to 50% of the population at risk, and the value was limited to 17%, considering no overlap of candidate clusters and covering the largest high-incidence geographic areas [34]. Given that the periodicity of the reported OID cases was approximately 2 to 3 months (Figure 2), the temporal scanning window was set to 90 days. Moreover, each candidate cluster had to include no less than 2 cases and have a minimum of 2 days. Age and sex were used as 2 covariate variables for covariate adjustment in the scan statistics [35]. The retrospective space-time scan statistics were applied for each year of OID in Fujian Province from 2005 to 2021.

Figure 2. The trend and seasonal decomposition analysis of other infectious diarrhea in Fujian Province (from January 1, 2005, to December 31, 2021).



Spatial Regression Analysis

Ordinary least regression (OLS) was used to determine the relationship between OID incidence (of each county) and each of the potential influencing factors (including demographic, economic, and environmental factors). Considering the spatial variation of the regression coefficient, geographical coordinates were included as parts of the regression parameters. Geographical weighted regression (GWR) is an effective method for determining spatial variables by considering local-scale features of geographical elements (county level in this study). The GWR is computed as follows [36]:

$$y_i = \sum_k \beta_k(u_i, v_i)x_{k,i} + \varepsilon_i$$

. Corrected Akaike information criterion (AICc) [37] and R^2 values, which consider model complexity and goodness of fit, were used to evaluate the OLS and GWR models [38].

Generalized linear mixed model (GLMM) [39] was used to investigate the factors influencing the occurrence of clustered types of OID cases, distinguishing whether the OID cases fall within the clustered areas (coded as “1”) or outside of them (coded as “0”) at both individual and county levels. The individual-level variables considered in the analysis were sex (male or female), age (<1 year, 1-5 years, or ≥5 years), and residential type. For the county-level variables (as mentioned earlier), we selected predictors based on the variance inflation factor to mitigate multicollinearity among the model terms. To evaluate the goodness of fit, we used Akaike information criterion (AIC) and Schwarz Bayesian information criterion [40]. Among the different models tested, the model with the lowest AIC and Bayesian information criterion values was deemed to be the most suitable and informative.

All statistical analyses were performed using RStudio (version 1.4.1717) [41]; time-series regression analysis was performed using the *mgcv* R package, spatial autocorrelation analysis was performed using the *spdep* R package, geographically weighted regression was performed using the *GWmodel* R package, GLMM was performed using the *lme4* R package, and the choropleth maps representing the detected disease clusters were generated using QGIS 3.20.2 Odense.

Ethical Considerations

The study was approved by the ethical review board of the Fujian Center for Disease Control and Prevention (Fuzhou, China; number 2020-029), and it was performed in accordance with the principles of the Declaration of Helsinki. Consent to participate was not applicable because this study used OID surveillance data. All data were anonymized and kept confidential to protect the privacy of participants.

Results

Assessment of Case-Reporting Quality

The data set under analysis comprised 388,066 (99.9%) out of 388,636 cases with complete information on the following epidemiological variables: sex, date of birth, occupation, residential type, notification type, symptom onset date, date of diagnosis, and date of data entry. A total of 570 cases without detailed residential addresses were excluded for spatial and spatiotemporal analysis purposes. The reporting delay (between the date of diagnosis and date of data entry) was calculated for all 388,636 cases, and the median (IQR) delay was 0 (−375 to 1) days. We further identified 1879 cases with a negative reporting delay. However, for our subsequent analysis, we used the diagnosis date, rendering the impact of the quality of reporting delay insignificant.

Characteristics of the Reported OID Cases

From January 1, 2005, to December 31, 2021, a total of 388,636 OID cases were reported from 85 county-level divisions (12 cities, 29 districts, and 44 counties) in Fujian Province; the reported OID cases were mainly from coastal areas (Figure 1). Table 1 summarizes the characteristics of the reported OID cases. Of the 388,636 patients with OID, 228,777 (58.9%) were male. The median (IQR) age of the patients was 1 (0-24) year, and 70.5% (273,988/388,636) of the patients were children aged <5 years; patients aged <1 year (108,818/388,636, 28%) accounted for the highest proportion of patients, followed by patients aged 1 to 2 years (88,220/388,636, 22.7%).

Table 2 presents a comparison of the characteristics of patients with OID aged <5 years and patients with OID aged >5 years. The differences in sex, notification type, and residential type between the 2 age groups were statistically significant (*P*<.001). Males aged <5 years had 1.56 times higher odds of being diagnosed with OID than those aged >5 years. Moreover, patients aged <5 years had 6.03 times higher odds of confirming OID through a clinical and pathogen diagnosis than those aged >5 years. In addition, within the floating population (both within city and province), children aged <5 years had 3.18 and 1.44 times higher odds, respectively, of having OID than those aged >5 years.

Table 1. Characteristics of other infectious diarrhea in Fujian Province from 2005 to 2021.

Characteristics	Patients with OID (N=388,636), n (%)
Sex	
Male	228,779 (58.9)
Female	159,857 (41.1)
Age (years)[a]	
0	108,675 (28)
1	88,230 (22.7)
2	30,328 (7.8)
3	15,736 (4)
4	8713 (2.2)
5	22,440 (5.8)
15	49,080 (12.6)
40	37,403 (9.6)
≥60	28,031 (7.2)
Notification type	
Confirmed case	140,140 (36.1)
Clinically diagnosed case	248,496 (63.9)
Occupation	
Diaspora children	245,863 (63.3)
Farmer	59,786 (15.4)
Student	19,520 (5)
Unemployed or housework or retired	18,893 (4.9)
Childcare	14,648 (3.8)
Migrant worker	1691 (0.4)
Medical staff	709 (0.2)
Others	20,755 (5.3)
Unknown	6771 (1.7)
Residential type	
Permanent residents	275,768 (71)
Floating population	112,868 (29)
Inside city	94,159 (24.2)
Inside province	11,790 (3)
Outside province	6919 (1.8)

[a] Age was grouped under 1 year, 2 years, 3 years, etc.

Table 2. Comparison of patients aged <5 years and patients aged >5 years in Fujian Province from 2005 to 2021[a].

Characteristics	Patients aged <5 years (n=251,682), n (%)	Patients aged >5 years (n=136,954), n (%)	OR[b] (95% CI)	*P* value
Sex				<.001
Male	158,439 (63)	70,340 (51.4)	1.56 (1.53-1.59)	
Female	93,243 (37)	66,614 (48.6)	0.64 (0.63-0.66)	
Notification type				<.001
Confirmed case	122,567 (48.7)	17,573 (12.8)	6.03 (5.92-6.15)	
Clinically diagnosed case	129,115 (51.3)	119,381 (87.2)	0.11 (0.11-0.11)	
Residential type				<.001
Permanent residents	161,809 (64.3)	113,959 (83.2)	0.3 (0.29-0.31)	
Floating population				
Inside city	77,612 (30.8)	16,547 (12.1)	3.18 (3.13-3.23)	
Inside province	8436 (3.4)	3354 (2.4)	1.44 (1.36-1.53)	
Outside province	3825 (1.5)	3094 (2.3)	0.65 (0.61-0.69)	

[a] Statistical significance (chi-square test) between patients aged <5 years and patients aged >5 years was set at *P*<.05.

[b] OR: odds ratio.

Incidence and Trend Analysis of the Reported OID Cases

The average annual incidence of OID was 60.3 (SD 16.7) per 100,000 people; the annual incidence of OID peaked in 2017 (at 88.3) and was lowest in 2005 (at 39.5). Table 3 shows the incidence of OID by sex and age. The highest OID incidence was reported in individuals aged <1 year (1362.1 per 100,000 people), followed by individuals aged <2 years (1055.5 per 100,000 people). Low incidence was observed in the individuals aged 15 to 40 years and individuals aged 40 to 60 years (19.2 and 21.2 per 100,000 people, respectively). The average annual incidence among males (71.4 per 100,000) was significantly higher than that among females (52.5 per 100,000; *P*<.05). Specifically, in the age groups >5 and ≥60 years, the incidence among males was higher than that among females, whereas in the age groups 15 and 40 years, the incidence among females was significantly higher than that among males.

Table 3. The incidence of other infectious diarrhea in Fujian Province, stratified by age and sex, from January 1, 2005, to December 31, 2021.

Age (years)[a]	Total		Male		Female	
	Reported cases, n (%)	Average annual incidence (per 100,000)	Reported cases, n (%)	Average annual incidence (per 100,000)	Reported cases, n (%)	Average annual incidence (per 100,000)
0	108,673 (28)	1265.6	70,553 (30.8)	1589.4	38,122 (23.8)	992.4
1	88,229 (22.7)	1042.5	54,575 (23.9)	1202.5	33,655 (21.1)	857.4
2	30,328 (7.8)	354.1	18,514 (8.1)	406.9	11,814 (7.4)	304.2
3	15,736 (4)	197.9	9489 (4.1)	224.1	6247 (3.9)	171.8
4	8713 (2.2)	114.0	5308 (2.3)	131.6	3405 (2.1)	98.6
5	22,440 (5.8)	28.8	14,233 (6.2)	36.2	8207 (5.1)	24.3
15	49,080 (12.6)	18.5	24,276 (10.6)	17.8	24,804 (15.5)	19.1
40	37,403 (9.6)	20.3	17,912 (7.8)	19.4	19,491 (12.2)	21.7
≥60	28,031 (7.2)	35.1	13,919 (6.1)	37.3	14,112 (8.8)	35.5
All	388,633 (100)	60.3	228,777 (100)	70.0	159,856 (100)	51.3

[a] Age was grouped under 1 year, 2 years, 3 years, etc.

The weekly accumulated number of reported OID cases in Fujian Province presented a gyrating upward trend from 2005 to 2018, reaching a peak around July to August in 2017; showed a downward trend until the beginning of 2020; and then rose again (Figure 2, “trend”). The seasonal decomposition analysis revealed an annual seasonal pattern (Figure 2, “seasonal”). However, the peak months varied throughout the study period. In 2005, 2006, 2010, 2012, and 2014, most OID cases were reported in autumn (September-November); conversely, in 2007, 2008, 2009, 2013, 2015, 2016, 2017, 2018, and 2020, the peak period was in winter (December-February). In 2019 and 2021, the peak month was March (Figure 2, “data”). There was a clear seasonal shift in the peak month for OID cases from autumn to winter or spring between 2005 and 2020.

Using meteorological data collected from February 2005 to December 2021 (encompassing 882 weeks), the association between the number of OID cases and meteorological factors was examined. The GAM revealed that atmospheric pressure at sea level (*P*<.001), humidity (*P*<.001), precipitation (*P*=.01), and maximum temperature (*P*=.02) were significantly linked to detrended OID data, whereas average temperature, minimum temperature, and time (weeks) were not significantly related. The results indicated that a higher maximum temperature, atmospheric pressure, humidity, and precipitation were associated with higher detrended OID data. The model explained 12.6% of the deviance in the data, and the adjusted R^2 value was 0.112. A summary of the GAM results, along with its diagnostics, is presented in Table 4.

Table 4. GAM[a] of meteorological factors and other infectious diarrhea in Fujian Province (from February 4, 2005, to December 31, 2021).[b]

Variable (smoothed)	*df*	Reference *df*	*P* value
Time	0.005	8	.59
Average temperature	1.024	1.038	.22
Maximum temperature	2.458	3.109	.02[c]
Minimum temperature	1	1.001	.99
Atmospheric pressure	6.052	7.294	<.001[d]
Relative humidity	2.62	3.349	<.001[d]
Precipitation	1.005	1.01	.01[c]
Adjusted R^2	0.112	N/A[e]	N/A
Deviance explained, %	12.60	N/A	N/A
REML[f]	5216.9	N/A	N/A
Number of weeks	882	N/A	N/A

[a] GAM: generalized additive model.

[b] Intercept: estimate=0.095, SE 3.065; *P*=.98.

[c] *P*<.05.

[d] *P*<.001.

[e] N/A: not applicable.

[f] REML: restricted maximum likelihood.

Spatial and Spatiotemporal Analysis of the Reported OID Cases

The spatial autocorrelation analysis using the global Moran I demonstrated a positive correlation (0.20-0.32, *P*<.001) of OID cases from 2005 to 2020 (Table 5), indicating a significant overall spatial autocorrelation. The localities with clusters of OID cases were identified using local Moran I (LISA). Figure 3 displays the Moran scatterplot for the 17 study years (2005-2021), which reveals that most event points were gathered in the upper-right (high-high) quadrant, confirming the existence of positive spatial autocorrelation among the counties based on OID cases. The LISA cluster map depicted the locations of the high clusters (high surrounded by high), low clusters (low surrounded by low), and spatial outliers (mixture of high and low clusters) in the OID cases.

In Figure 3, the LISA map for all study years (2005-2021) shows that some counties on the south coast of Fujian Province exhibited high-high clusters, whereas the northwest and middle parts were found to be covered by low-low clusters. The yearly LISA map revealed outstanding spatial clusters of OID cases covering specific areas each year. From 2005 to 2007, the clustered areas with the highest density of reported OID cases (high-high) were primarily concentrated in Zhangzhou City, and in 2008, these areas drifted northeastward to Xiamen City and Quanzhou City. Other clustered areas with high OID rates appeared in Ningde City from 2005 to 2011 and in Fuzhou City from 2012 to 2019. Most of the high-low outliers were dispersed in the middle part of Fujian Province (Sanming City and Nanping City), mainly in 2016, 2020, and 2021; these high-low areas were generally surrounded by low-low clusters, which were mountainous areas. The low-high outliers appeared in the same location in Ningde City (northern part) from 2006 to 2018. The clusters and outliers were statistically significant based on *z* test (*P*<.05).

Table 5. The global Moran index of other infectious diarrhea in Fujian Province (from January 1, 2005, to December 31, 2021).

Year	Moran I	*z* score	*P* value
2005	0.22	12.20	<.001
2006	0.21	11.66	<.001
2007	0.24	12.88	<.001
2008	0.23	11.90	<.001
2009	0.30	15.51	<.001
2010	0.25	12.82	<.001
2011	0.23	12.00	<.001
2012	0.21	10.84	<.001
2013	0.22	11.24	<.001
2014	0.23	12.32	<.001
2015	0.28	14.49	<.001
2016	0.28	14.58	<.001
2017	0.32	16.55	<.001
2018	0.29	15.40	<.001
2019	0.26	13.87	<.001
2020	0.20	11.31	<.001
2021	0.31	16.53	<.001

Figure 3. Moran scatterplot and local indices of spatial association cluster map of other infectious diarrhea at county level in Fujian Province (from January 1, 2005, to December 31, 2021). A higher resolution version of the figure is available in Multimedia Appendix 1.

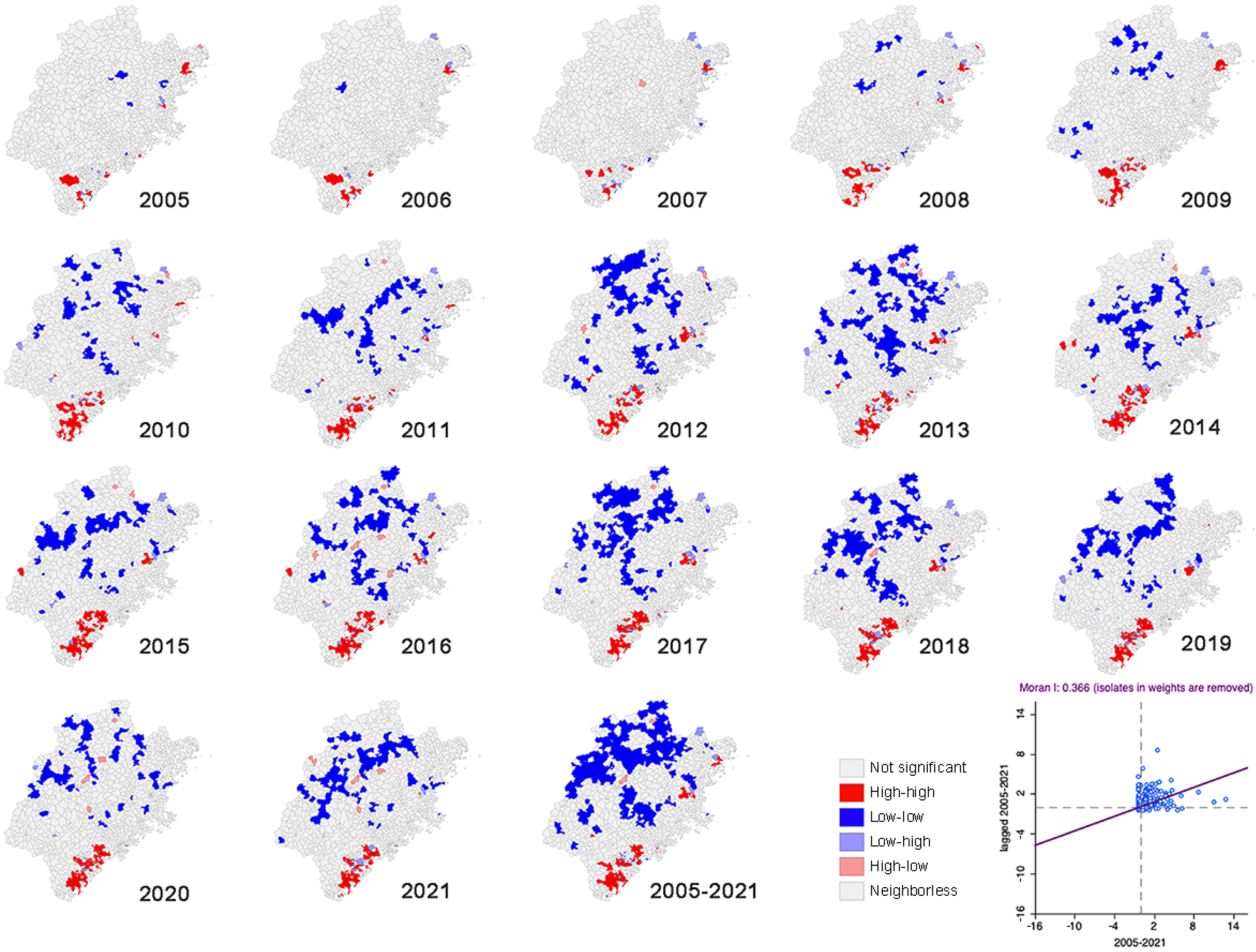


Table 6 shows the most likely cluster of OID cases detected by retrospective space-time scan statistics from 2005 to 2020 in Fujian Province. The time frame was from September to November (2005, 2006, and 2010), mid-June to mid-September (2007), October to December (2008, 2009, and 2012-2014), June to August (2011), January to March (2015-2017), August to October (2018), July to September (2019), and mid-January to mid-April (2020). The average radius of the most likely clusters over the 17-year study period was 56.35 km. The average number of counties included in the most likely cluster was 9.76, with an average relative risk of 8.79.

The most likely clusters were assembled in the southern part of Fujian Province, mainly in Zhangzhou City, Xiamen City, and Quanzhou City (Figure 4). The most likely cluster was detected in the north of Zhangzhou City and south of Xiamen in 2005 and 2006, respectively, and then moved southward to cover the entire region of Zhangzhou City between 2007 and 2010. Again, it moved northward to Xiamen City, covering the northern part of Zhangzhou City and the southern part of Quanzhou City.

Figure 5 illustrates the frequency of OID most likely clusters within districts based on retrospective space-time scan statistics. Two counties in Zhangzhou City (Zhangpu and Longhai) were detected 17 times (each year) using this method, and the median (IQR) frequency for the occurrence of most likely clusters per county was 4.5 (2.25-13.75) times. The southern regions of the province exhibited the highest frequency of OID clusters, whereas no most likely clusters were detected in the central, western, or northeast districts during the study period.

Table 6. The most likely cluster of other infectious diarrhea detected using space-time scan statistics in Fujian Province (from January 1, 2005, to December 31, 2021).

Year	Period	Radius (km)	Observed cases, n	Expected cases	Relative risk	*P* value	Total counties, n	Population, n
2005	September 5 to December 2	53.88	2716	531.0	6.12	<.001	13	5,550,594
2006	August 30 to November 27	53.88	2808	554.1	6.09	<.001	13	5,624,309
2007	June 17 to September 14	58.51	2218	492.2	5.11	<.001	9	4,409,866
2008	October 6 to December 31	97.44	3010	648.4	5.47	<.001	12	5,622,875
2009	October 3 to December 31	97.44	2319	648.4	4.02	<.001	12	5,680,749
2010	August 26 to November 23	72.07	3628	550.6	7.70	<.001	14	5,739,619
2011	May 29 to August 26	39.75	1327	91.0	15.47	<.001	3	1,202,354
2012	October 4 to December 31	45.67	2532	372.1	7.45	<.001	9	3,497,184
2013	October 6 to December 31	45.67	3158	455.5	7.58	<.001	9	3,521,250
2014	October 3 to December 31	45.67	2678	379.6	7.77	<.001	9	3,521,250
2015	January 1 to March 25	45.67	3356	322.8	11.87	<.001	9	3,443,638
2016	January 2 to March 31	45.67	3394	350.2	11.04	<.001	9	3,474,991
2017	January 2 to April 1	45.67	3859	468.4	9.16	<.001	9	3,505,793
2018	January 1 to March 29	63.89	6226	1228.5	6.07	<.001	14	9,479,996
2019	July 29 to October 26	39.75	1893	128.6	15.82	<.001	3	1,341,011
2020	June 28 to September 25	39.75	1630	104.4	16.85	<.001	3	1,341,996
2021	January 16 to April 15	67.54	5851	1186.8	5.92	<.001	16	11,084,758

Figure 4. Annual spatial distribution of space-time clusters of other infectious diarrhea at county level in Fujian Province (from January 1, 2005, to December 31, 2021). A higher resolution version is available in Multimedia Appendix 2.

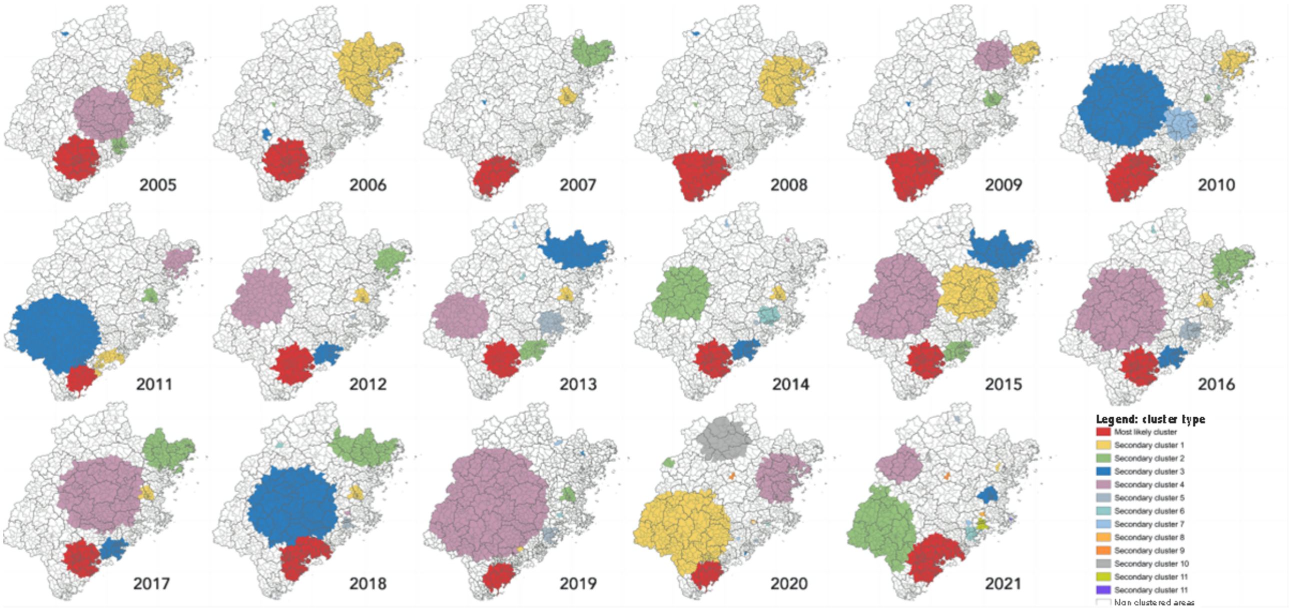
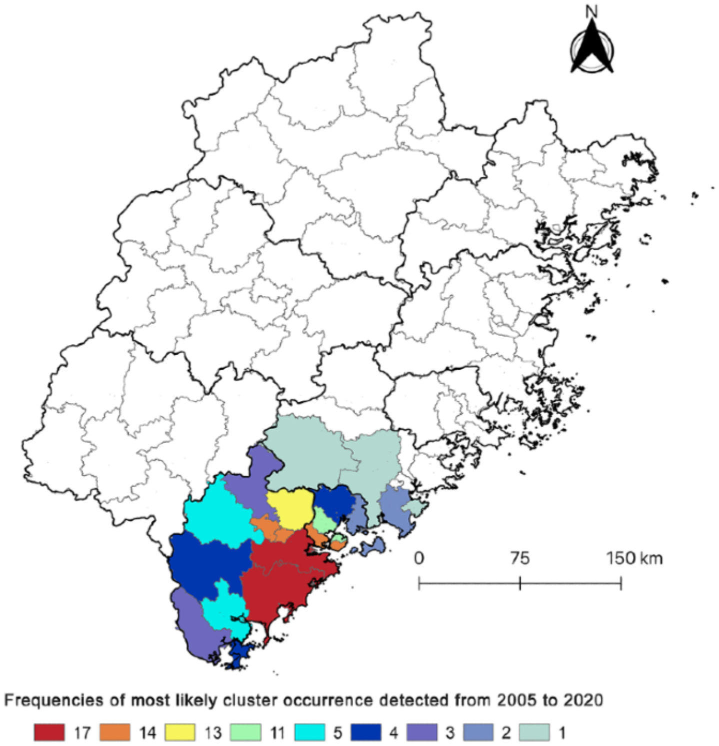


Figure 5. Frequency of most likely cluster occurrence at county level in Fujian Province (from January 1, 2005, to December 31, 2021).



Risk Factors Related to OID Incidence

The relationship between OID incidence and the collected county-level sociodemographic factors was analyzed using a multivariate OLS regression model. The permanent resident population showed a significantly positive association. By contrast, the urban population, rural population, and <5-year-old population showed a significantly negative association with OID cases, and <1-year-old population was positively associated with OID cases, although not statistically significantly. Regarding economic factors, GDP per capita and primary sector's GDP were positively associated with OID cases. By contrast, the secondary and tertiary sectors' GDP and urbanization levels were negatively associated with OID cases. For environmental factors, the contents of arsenic, cadmium, chromium, copper, lead, and zinc in the soil were included in the OLS model, of which arsenic, chromium, and zinc were negatively associated, and the others were positively associated. Table 7 presents a summary of the multivariate OLS with model diagnostics.

A locally linear, nonparametric estimation using the GWR was used to cope with the OLS's nonstationary nature. The adjusted R^2 values rose from 0.389 to 0.405 (Table 7), even if the GWR did not improve the model fit from the OLS model according to the AICc values.

Table 8 and Table 9 summarized the results of GLMM. The response variable “clustered type” in the GLMM indicated whether the OID case is within the clustered areas (“1”) or outside of them (“0”). The clustered areas corresponded to counties that were detected at least once within the most likely clusters from the previous analysis. The GLMM adopts a binomial family with a logit link function, which is appropriate for analyzing binary outcomes. The maximum likelihood method (Laplace approximate) was used for estimation. The model demonstrated good fit, as indicated by AIC=24,672.6 and Bayesian information criterion=24,781.3. The results showed that population density, urbanization level, and copper content had positive estimates, suggesting that an increase in these variables was associated with a higher likelihood of OID cases being within the clustered areas. Conversely, arsenic, cadmium, and chromium contents had negative estimates, implying that a decrease in these variables was associated with a higher likelihood of OID cases being outside the clustered areas.

Table 7. Multivariate ordinary squares regression and geographic weighted regression model results of other infectious diarrhea in Fujian Province (from January 1, 2005, to December 31, 2021).[a]

Variable	Ordinary least regression (OLS) model		
	B	SE	*P* value
Population			
Permanent resident	0.050	0.022	.03[b]
Urban	−0.046	0.021	.03[b]
Rural	−0.049	0.022	.03[b]
<1 year old	0.077	0.044	.08
<5 year old	−0.024	0.011	.03[b]
Economic factor			
GDP[c] per capita	0.009	0.003	.004[d]
Primary sector GDP	0.231	2.653	.93
Secondary sector GDP	−3.582	0.995	<.001[d]
Tertiary sector GDP	−2.292	0.881	.01[b]
Urbanization (%)	−9.438	5.492	.090
Environmental factor			
Arsenic content	−55.301	16.157	<.001[d]
Cadmium content	0.363	1.984	.86
Chromium content	−1.209	0.557	.03[b]
Copper content	0.442	0.759	.56
Lead content	0.019	0.025	.47
Zinc content	−0.051	0.035	.15
AICc[e]	1208.861	N/A[f]	N/A
Adjusted R^2	0.389	N/A	N/A
AICc (GWR[g])	1208.100	N/A	N/A
Adjusted R^2 (GWR)	0.405	N/A	N/A

[a] Intercept: estimate=495.003, SE 317.773; *P*=.12.

[b] *P*<.05.

[c] GDP: gross domestic product.

[d] *P*<.01.

[e] AICc: corrected Akaike information criterion.

[f] N/A: not applicable.

[g] GWR: geographical weighted regression.

Table 8. Fixed effects of generalized linear mixed model results of other infectious diarrhea in Fujian Province (from January 1, 2005, to December 31, 2021).

Coefficients	Estimate	SE	*z* value	*P* value
Intercept	−11.205	0.716	−15.64	<.001
Population density	2.342	0.101	23.21	<.001
Urbanization (%)	14.706	0.187	78.54	<.001
Arsenic content	−12.556	0.387	−32.47	<.001
Cadmium content	1.044	0.032	32.52	<.001
Chrome content	−7.059	0.122	−58.07	<.001
Copper content	1.944	0.169	11.49	<.001

Table 9. Random effects of generalized linear mixed model results of other infectious diarrhea in Fujian Province (from January 1, 2005, to December 31, 2021).

Random effects	Variances	SD
Residential types	0.163	0.404
Age group	0.002	0.049
Sex	0.003	0.056

Discussion

Principal Findings

In this study, we investigated the overall epidemic characteristics, spatiotemporal pattern, and risk factors associated with the occurrence of OID in Fujian Province. By analyzing 17 years of surveillance data, we identified the clustering of OID cases in Zhangzhou City and Xiamen City, with the clustering areas remaining relatively stable over time. Our findings also revealed that densely populated areas with a large <1-year-old population, less economically developed areas, and higher pollution levels were associated with higher OID incidence in the province. Moreover, high population density, urbanization level, and cadmium and copper contents in the soil could contribute to the clustering of OID.

The results highlighted the susceptibility of children aged <5 years; particularly, those aged 0 to 2 years accounted for a significant proportion (196,905/388,636, 50.7%) of the OID cases in Fujian Province. This finding aligns with previous research indicating their susceptibility to OID [42-44]. Viral pathogens such as rotavirus, adenovirus, norovirus, and Salmonella were the 4 most common pathogens in China [45]. Specifically, rotavirus infection increased rapidly in children aged <5 years [7,46]. Previous studies in Fujian Province have also emphasized the significance of rotavirus as a causative pathogen for acute gastroenteritis among children who are hospitalized [47].

Temporal analysis revealed a steady increase in OID cases from 2005 to 2017, with a clear seasonal shift in OID cases from autumn to winter/spring between 2005 and 2020. However, since 2020, there has been subsequent stagnation in case numbers, potentially linked to the impact of the COVID-19 pandemic. Behavioral interventions and policies implemented during the pandemic to suppress the transmission of respiratory infections [48] may have inadvertently reduced the transmission of other infectious diseases, including OID.

Our study found that maximum temperature, atmospheric pressure at sea level, average relative humidity, and precipitation are positively associated with the reported OID cases in Fujian, consistent with previous research in the coastal area of China [22,49,50]. Infectious diarrhea is often caused by microorganisms that are sensitive to environmental conditions [51], and higher temperatures, atmospheric pressure, and humidity may favor the survival and proliferation of microorganisms causing diarrhea [52]. Moreover, increased atmospheric pressure may reduce the dilution of airborne pathogens, leading to an increased concentration of pathogens and further increasing the likelihood of diarrhea transmission [44]. In addition, variations in rainfall and temperature can increase the risk of fecal contamination and diarrhea, particularly during heavy rainfalls and floods. The observed seasonal shift in OID cases from autumn to winter suggests a possible increase in viral infections as the main cause of OID during the winter, possibly owing to a more favorable environment for viral pathogen transmission [53]. However, this hypothesis needs further investigation.

Spatial analysis identified the presence of persistent "hot spots" of OID cases, mainly concentrated in Zhangzhou City and Xiamen City within the study period. These regions have a subtropical monsoon climate, with factors such as high humidity, temperature variation, heavy rainfall, and wind speed potentially facilitating OID transmission [8,22,42]. The predominant pathogen in these areas is rotavirus, which exhibits apparent seasonality during the autumn-winter season [47]. Other risk factors such as the consumption of raw seafood, which often contains pathogens such as rotavirus, norovirus, and *Vibrio parahaemolyticus*, have also been identified [54]. A study conducted in Taiwan on *Campylobacter* gastroenteritis revealed an increased risk of campylobacteriosis in children associated with seafood consumption [55].

In our study, almost 90% (15/17) of the most likely clusters (except for those in 2007 and 2011) were detected in the autumn and winter seasons, consistent with reported gastroenteritis outbreaks caused by norovirus [56]. These results suggest that viral pathogens and foodborne clustered infections are potential contributors to the identified OID clusters in the spatiotemporal analysis. Strengthening outbreak management capacities and implementing robust prevention and control measures in high-incidence areas are necessary. In addition, the high prevalence of human intestinal protozoa [57] and significant environmental pollution in both the land and adjacent sea areas [58] may be associated with the 2 most clustered counties (Zahngpu and Longhai) in Fujian Province found in this study. Furthermore, the clustered areas exhibited a dynamic spreading trend, expanding from the southernmost region to southeast and gradually southward over time. Therefore, close monitoring of the epidemic situation in the surrounding areas is essential.

Regression analysis indicated a positive association between OID cases and factors such as densely populated areas with a large >1-year-old population, less economically developed areas, and higher pollution levels. Children aged <5 years are susceptible to viral infectious diarrhea and are likely to further damage their health [59]. Countermeasures, including rotavirus vaccine [60], improvement of zinc nutriture [61], and dietary counseling [62], have been proven to reduce the high burden of childhood diarrhea. In coastal provinces, such as Fujian, living habits that involve frequent contact with seafood could be related to OID. Hygiene and cooked food education should be enhanced in high-risk areas, such as Zhangzhou City. Poor public health, poor water safety, fecal contamination, and environmental pollution pose obstacles to diarrhea prevention in economically underdeveloped areas [42]. Nevertheless, the impact of social factors on diarrhea remains uncertain [63]. Therefore, further research is needed to explore the major mechanisms underlying the association between infectious diarrhea and certain social factors.

The findings from our study underscore the need for implementing targeted prevention and control measures for OID in Fujian, with a particular focus on susceptible populations, such as young children. Furthermore, the health department should prioritize improving the living environments in Zhangzhou and Xiamen. Regular monitoring of the epidemic characteristics of infectious diarrheal diseases, incorporating meteorological factors into surveillance systems, and developing early warning models can enhance preparedness and response to outbreaks. This study's findings contribute to the understanding of the epidemic characteristics of OID in Fujian Province and provide a vital foundation for informing future risk assessments and facilitating early warning predictions of this disease. Examining the impact of OID in Fujian Province can provide valuable epidemiological insights and effective prevention and control strategies that can be transferred to other regions and countries facing similar challenges.

Limitations

Our study has several limitations. First, the lack of detailed information on the etiology of OID cases hinders further investigation into the pathogenesis of OID in infants and children as well as the identification of specific pathogen-related epidemic patterns. Hence, a province-wide pathogen monitoring and surveillance system is needed, which would provide valuable information for in-depth pathogen characterization. Second, the power of cylindrical scan statistics could be limited to the complex geographic landscape of Fujian Province. More flexible shapes of scan statistics or the use of higher-resolution geographical information can enhance the accuracy and precision of the analysis. Third, although the GAM used in our study captured some of the short-term effects of meteorological factors on OID cases, it explained only a small proportion of the overall variation. Therefore, more suitable forecasting models should be considered to enhance the predictive ability. Fourth, the regression models used to assess the association between OID and risk factors had inherent limitations. Furthermore, the lack of county-level meteorological and socioeconomic data, such as public health resources, further weakened the explanatory power of the regression model. Future research should incorporate predictive models that forecast the incidence of OID in the near and long term, considering the impact of preventive measures. This approach would highlight the importance of ongoing targeted interventions to control OID and improve the health and well-being of susceptible populations. Finally, owing to data limitations, we could not directly assess the influence of the COVID-19 pandemic and its subsequent nonpharmaceutical interventions on OID transmission. Therefore, future research should explore the intersection between the COVID-19 pandemic and OID to gain a comprehensive understanding of the broader impact on public health.

Conclusions

Our study revealed that the incidence of OID in Fujian Province exhibits a distinct distribution across populations, seasons, and regions. Notably, males and children aged >2 years are the most susceptible to OID. Moreover, our analysis demonstrated a shift in the seasonal peak of OID from autumn to winter or spring between 2005 and 2020. Our findings indicated that maximum temperature, atmospheric pressure, relative humidity, and precipitation positively influence the short-term incidence of OID. Zhangzhou City and Xiamen City were identified as the major hot spots for OID in Fujian Province. Therefore, it is essential to prioritize prevention and control efforts in these identified hot spots and among highly susceptible groups. Our findings provide valuable scientific evidence for policy makers to develop targeted intervention measures for OID, considering the specific disease pattern and geographical context.

Acknowledgments

This work was supported by the Major Health Research Project of Fujian Province (grant 2021ZD01001), Natural Science Foundation of Fujian Province (grant 2021J01350), Fujian Provincial Natural Science Foundation Project (grant 2021R0111), Health Science and Technology Project of Fujian Province (grants 2020GGB019 and 2021GGA037), and Fujian Science and Technology Innovation Platform Construction Project (grant 2019Y2001).

Data Availability

The data sets analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

YL, GC, and FH contributed to the study design, methodology, and data analysis. YL wrote the original draft. GC, HZ, and ZH supervised the project and the writing, review, and editing of the manuscript. FH and GC equally contributed as corresponding authors of this manuscript. All the authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Higher resolution version of Figure 3.
[PNG File , 2303 KB - publichealth_v9i1e45870_app1.png]

Multimedia Appendix 2
Higher resolution version of Figure 4.
[PNG File , 2854 KB - publichealth_v9i1e45870_app2.png]

References

1. 2021 pneumonia and diarrhea progress report finds key gains despite toll of the COVID-19 pandemic. Johns Hopkins University. URL: https://www.jhsph.edu/ivac/2021/11/09/2021-pdpr/ [accessed 2022-11-04]
2. Diarrhoeal disease. World Health Organization. 2017. URL: https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease [accessed 2022-11-04]
3. Liu H, Zhang J. Analysis of the pathogens in infectious diarrhea (other than cholera, dysentery, typhoid and paratyphoid) cases reported in China in 2008. Zhonghua Yu Fang Yi Xue Za Zhi 2013;47:328-332 [FREE Full text]
4. Pathogen. Chinese Center for Disease Control and Prevention. URL: https://www.chinacdc.cn/jkzt/crb/bl/grxfxb/zstd/200510/t20051012_24768.html [accessed 2023-04-21]
5. Levy K, Woster AP, Goldstein RS, Carlton EJ. Untangling the impacts of climate change on waterborne diseases: a systematic review of relationships between diarrheal diseases and temperature, rainfall, flooding, and drought. Environ Sci Technol 2016 May 17;50(10):4905-4922 [FREE Full text] [doi: 10.1021/acs.est.5b06186] [Medline: 27058059]
6. Chowdhury FR, Ibrahim QS, Bari MS, Alam MM, Dunachie SJ, Rodriguez-Morales AJ, et al. The association between temperature, rainfall and humidity with common climate-sensitive infectious diseases in Bangladesh. PLoS One 2018 Jun 21;13(6):e0199579 [FREE Full text] [doi: 10.1371/journal.pone.0199579] [Medline: 29928056]
7. Wang LP, Zhou SX, Wang X, Lu QB, Shi LS, Ren X, Chinese Centers for Disease Control and Prevention (CDC) Etiology of Diarrhea Surveillance Study Team. Etiological, epidemiological, and clinical features of acute diarrhea in China. Nat Commun 2021 Apr 29;12(1):2464 [FREE Full text] [doi: 10.1038/s41467-021-22551-z] [Medline: 33927201]
8. Mao Y, Zhang N, Zhu B, Liu J, He R. A descriptive analysis of the Spatio-temporal distribution of intestinal infectious diseases in China. BMC Infect Dis 2019 Sep 02;19(1):766 [FREE Full text] [doi: 10.1186/s12879-019-4400-x] [Medline: 31477044]
9. Phung D, Huang C, Rutherford S, Chu C, Wang X, Nguyen M, et al. Temporal and spatial patterns of diarrhoea in the Mekong Delta area, Vietnam. Epidemiol Infect 2015 Dec;143(16):3488-3497 [FREE Full text] [doi: 10.1017/S0950268815000709] [Medline: 25876699]
10. Naumova EN, Jagai JS, Matyas B, DeMaria A, MacNeill IB, Griffiths JK. Seasonality in six enterically transmitted diseases and ambient temperature. Epidemiol Infect 2007 Feb;135(2):281-292. [doi: 10.1017/S0950268806006698] [Medline: 17291363]
11. Kovats RS, Edwards SJ, Charron D, Cowden J, D'Souza RM, Ebi KL, et al. Climate variability and campylobacter infection: an international study. Int J Biometeorol 2005 Mar 23;49(4):207-214. [doi: 10.1007/s00484-004-0241-3] [Medline: 15565278]

12. Carlton EJ, Eisenberg JN, Goldstick J, Cevallos W, Trostle J, Levy K. Heavy rainfall events and diarrhea incidence: the role of social and environmental factors. Am J Epidemiol 2014 Feb 01;179(3):344-352 [FREE Full text] [doi: 10.1093/aje/kwt279] [Medline: 24256618]
13. Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, Clements AC. Spatial Analysis in Epidemiology. Oxford, UK: Oxford University Press; 2008.
14. Sasson C, Cudnik M, Nassel A, Semple H, Magid D, Sayre M, Columbus Study Group. Identifying high-risk geographic areas for cardiac arrest using three methods for cluster analysis. Acad Emerg Med 2012 Feb;19(2):139-146 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01284.x] [Medline: 22320364]
15. Jacquez GM, Greiling DA. Local clustering in breast, lung and colorectal cancer in Long Island, New York. Int J Health Geogr 2003 Feb 17;2(1):3 [FREE Full text] [doi: 10.1186/1476-072x-2-3] [Medline: 12633503]
16. Laohasiriwong W, Puttanapong N, Luenam A. A comparison of spatial heterogeneity with local cluster detection methods for chronic respiratory diseases in Thailand. F1000Res 2017 Oct;6:1819 [FREE Full text] [doi: 10.12688/f1000research.12128.2] [Medline: 29657710]
17. Liu CP, Luo CL, Gao Y, Li FB, Lin LW, Wu CA, et al. Arsenic contamination and potential health risk implications at an abandoned tungsten mine, southern China. Environ Pollut 2010 Mar;158(3):820-826. [doi: 10.1016/j.envpol.2009.09.029] [Medline: 19910093]
18. Zhuang P, Zou B, Li NY, Li ZA. Heavy metal contamination in soils and food crops around Dabaoshan mine in Guangdong, China: implication for human health. Environ Geochem Health 2009 Dec 13;31(6):707-715. [doi: 10.1007/s10653-009-9248-3] [Medline: 19214759]
19. Alengebawy A, Abdelkhalek ST, Qureshi SR, Wang M. Heavy metals and pesticides toxicity in agricultural soil and plants: ecological risks and human health implications. Toxics 2021 Feb 25;9(3):42 [FREE Full text] [doi: 10.3390/toxics9030042] [Medline: 33668829]
20. Weng X, Wang Z, Ren J, Zhang Y, Yu L, Wang R. Surveillance for public health emergencies caused by infectious diarrhea other than cholera, dysentery, typhoid and paratyphoid in China, 2014–2016. Dis Surveill 2019;34(6):565-570 [FREE Full text] [doi: 10.3784/j.issn.1003-9961.2019.06.020]
21. Fujian Province’s statutory reporting of infectious disease epidemics in September 2022. Fujian Provincial Health Commission. URL: https://wjw.fujian.gov.cn/xxgk/gsgg/yqgg/202210/t20221029_6030448.htm [accessed 2022-11-05]
22. Hu WQ, Li YY, Ma W. [Short-term impact of temperature on infectious diarrhea in southeast coastal area of China, 2005-2013]. Zhonghua Yu Fang Yi Xue Za Zhi 2019 Jan 06;53(1):103-106. [doi: 10.3760/cma.j.issn.0253-9624.2019.01.015] [Medline: 30605971]
23. Chen KC, Lin CH, Qiao QX, Zen NM, Zhen GK, Chen GL, et al. The epidemiology of diarrhoeal diseases in southeastern China. J Diarrhoeal Dis Res 1991 Jun;9(2):94-99. [Medline: 1918839]
24. Diagnostic criteria for infectious diarrhea. National Health Commission of the People's Republic of China. URL: http://www.nhc.gov.cn/wjw/s9491/200704/38817.shtml [accessed 2022-05-10]
25. Weather forecast page view statistics. RU: Reliable Prognosis. URL: https://rp5.ru/stat/508/en [accessed 2023-04-28]
26. Wood SN. Generalized Additive Models: An Introduction with R. 2nd edition. New York City, NY: Chapman and Hall/CRC; 2017.
27. Thompson ES, Saveyn P, Declercq M, Meert J, Guida V, Eads CD, et al. Characterisation of heterogeneity and spatial autocorrelation in phase separating mixtures using Moran's I. J Colloid Interface Sci 2018 Mar 01;513:180-187 [FREE Full text] [doi: 10.1016/j.jcis.2017.10.115] [Medline: 29153711]
28. Moran PA. Notes on continuous stochastic phenomena. Biometrika 1950 Jun;37(1-2):17-23. [Medline: 15420245]
29. Chowdhury AI, Abdullah AY, Haider R, Alam A, Billah SM, Bari S, et al. Analyzing spatial and space-time clustering of facility-based deliveries in Bangladesh. Trop Med Health 2019 Jul 16;47(1):44 [FREE Full text] [doi: 10.1186/s41182-019-0170-9] [Medline: 31346313]
30. Anselin L. Local indicators of spatial association—LISA. Geogr Anal 1995 Apr;27:93-115 [FREE Full text] [doi: 10.1111/j.1538-4632.1995.tb00338.x]
31. Waller LA, Gotway CA. Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley-Interscience; 2004.
32. Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. Am J Public Health 1998 Sep;88(9):1377-1380. [doi: 10.2105/ajph.88.9.1377] [Medline: 9736881]
33. Huang L, Kulldorff M, Gregorio D. A spatial scan statistic for survival data. Biometrics 2007 Mar 16;63(1):109-118. [doi: 10.1111/j.1541-0420.2006.00661.x] [Medline: 17447935]
34. Rao H, Shi X, Zhang X. Using the Kulldorff's scan statistical analysis to detect spatio-temporal clusters of tuberculosis in Qinghai Province, China, 2009-2016. BMC Infect Dis 2017 Aug 21;17(1):578. [doi: 10.1186/s12879-017-2643-y] [Medline: 28826399]
35. Robertson C, Nelson TA, MacNab YC, Lawson AB. Review of methods for space-time disease surveillance. Spat Spatiotemporal Epidemiol 2010 Jul;1(2-3):105-116 [FREE Full text] [doi: 10.1016/j.sste.2009.12.001] [Medline: 22749467]
36. Brunsdon C, Fotheringham AS, Charlton ME. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr Anal 2010 Sep 03;28(4):281-298 [FREE Full text] [doi: 10.1111/j.1538-4632.1996.tb00936.x]

37. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr 1974 Dec;19(6):716-723 [FREE Full text] [doi: 10.1109/tac.1974.1100705]
38. Ordinary least Squares regression (OLS). XLSTAT: Your Data Analysis Solution. URL: https://www.xlstat.com/en/solutions/features/ordinary-least-squares-regression-ols [accessed 2023-11-03]
39. GILMOUR AR, ANDERSON RD, RAE AL. The analysis of binomial data by a generalized linear mixed model. Biometrika 1985 Dec;72(3):593-599 [FREE Full text] [doi: 10.1093/biomet/72.3.593]
40. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods 2012 Jun;17(2):228-243 [FREE Full text] [doi: 10.1037/a0027127] [Medline: 22309957]
41. RStudio: integrated development environment for R. KDnuggets. 2019. URL: https://www.kdnuggets.com/2011/03/rstudio-ide-for-r.html [accessed 2023-11-03]
42. Yang X, Xiong W, Huang T, He J. Meteorological and social conditions contribute to infectious diarrhea in China. Sci Rep 2021 Dec 03;11(1):23374 [FREE Full text] [doi: 10.1038/s41598-021-00932-0] [Medline: 34862400]
43. Chen C, Guan Z, Huang C, Jiang D, Liu X, Zhou Y, et al. Epidemiological trends and hotspots of Other Infectious Diarrhea (OID) in Mainland China: a population-based surveillance study from 2004 to 2017. Front Public Health 2021 Jul 22;9:679853 [FREE Full text] [doi: 10.3389/fpubh.2021.679853] [Medline: 34368054]
44. Wang H, Di B, Zhang T, Lu Y, Chen C, Wang D, et al. Association of meteorological factors with infectious diarrhea incidence in Guangzhou, southern China: a time-series study (2006-2017). Sci Total Environ 2019 Jul 01;672:7-15. [doi: 10.1016/j.scitotenv.2019.03.330] [Medline: 30954825]
45. Hongmei L. Research on the epidemic characteristics and changing trends of other infectious diarrhea diseases in my country from 2005 to 2019. Chinese Center for Disease Control and Prevention. 2020. URL: https://kns.cnki.net/KCMS/detail/detail.aspx?filename=1021001876.nh&dbname=CMFDTEMP [accessed 2023-11-18]
46. Luo HM, Ran L, Meng L, Lian YY, Wang LP. [Analysis of epidemiological characteristics of report cases of rotavirus diarrhea in children under 5 years old in China, 2005-2018]. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Feb 06;54(2):181-186. [doi: 10.3760/cma.j.issn.0253-9624.2020.02.013] [Medline: 32074707]
47. Wu BS, Huang ZM, Weng YW, Chen FQ, Zhang YL, Lin WD, et al. Prevalence and genotypes of rotavirus a and human adenovirus among hospitalized children with acute gastroenteritis in Fujian, China, 2009-2017. Biomed Environ Sci 2019 Mar;32(3):210-214 [FREE Full text] [doi: 10.3967/bes2019.028] [Medline: 30987695]
48. Zipfel CM, Colizza V, Bansal S. The missing season: the impacts of the COVID-19 pandemic on influenza. Vaccine 2021 Jun 23;39(28):3645-3648 [FREE Full text] [doi: 10.1016/j.vaccine.2021.05.049] [Medline: 34078554]
49. Zhou X, Zhou Y, Chen R, Ma W, Deng H, Kan H. High temperature as a risk factor for infectious diarrhea in Shanghai, China. J Epidemiol 2013;23(6):418-423 [FREE Full text] [doi: 10.2188/jea.je20130012] [Medline: 23994865]
50. Fang X, Liu W, Ai J, He M, Wu Y, Shi Y, et al. Forecasting incidence of infectious diarrhea using random forest in Jiangsu Province, China. BMC Infect Dis 2020 Mar 14;20(1):222 [FREE Full text] [doi: 10.1186/s12879-020-4930-2] [Medline: 32171261]
51. Shope R. Global climate change and infectious diseases. Environ Health Perspect 1991 Dec;96:171-174 [FREE Full text] [doi: 10.1289/ehp.9196171] [Medline: 1820262]
52. Chan MC, Mok HY, Lee TC, Nelson EA, Leung TF, Tam WW, et al. Rotavirus activity and meteorological variations in an Asian subtropical city, Hong Kong, 1995-2009. J Med Virol 2013 Nov 12;85(11):2026-2033. [doi: 10.1002/jmv.23686] [Medline: 23852875]
53. Li W, Xiang W, Li C, Xu J, Zhou D, Shang S. Molecular epidemiology of rotavirus A and adenovirus among children with acute diarrhea in Hangzhou, China. Gut Pathog 2020 Apr 13;12(1):19 [FREE Full text] [doi: 10.1186/s13099-020-00359-4] [Medline: 32313556]
54. Chen C, Wu B, Zhang H, Li KF, Liu R, Wang HL, et al. Molecular evolution of GII.P17-GII.17 norovirus associated with sporadic acute gastroenteritis cases during 2013-2018 in Zhoushan Islands, China. Virus Genes 2020 Jun 17;56(3):279-287. [doi: 10.1007/s11262-020-01744-6] [Medline: 32065329]
55. Guo YT, Hsiung CA, Wu FT, Chi H, Huang YC, Liu CC, Taiwan Pediatric Infectious Disease Alliance. Clinical Manifestations and Risk factors of Campylobacter gastroenteritis in children in Taiwan. Biomed J (Forthcoming) 2023 Mar 29 [FREE Full text] [doi: 10.1016/j.bj.2023.03.003] [Medline: 37001586]
56. Guo Z, Huang J, Shi G, Su C, Niu JJ. A food-borne outbreak of gastroenteritis caused by norovirus GII in a university located in Xiamen City, China. Int J Infect Dis 2014 Nov;28:101-106 [FREE Full text] [doi: 10.1016/j.ijid.2014.06.022] [Medline: 25263502]
57. Bao-jian C, Han-guo X, Rong-yan Z, Yan-rong L, Chen-xin L, Xian-liang X, et al. Investigation and analysis of human intestinal protozoal diseases in Fujian province. Chin J Zoonoses 2018:542-545 [FREE Full text] [doi: 10.1142/s1793724816000158]
58. Environmental conditions of land-based sewage outlets into the sea and adjacent sea areas - 2009 Zhangzhou City. Marine Environmental Status Bulletin - China Marine Information Network. URL: https://www.nmdis.org.cn/c/2011-05-05/58745.shtml [accessed 2022-11-08]

59. Thiagarajah JR, Kamin DS, Acra S, Goldsmith JD, Roland JT, Lencer WI, PediCODE Consortium. Advances in evaluation of chronic diarrhea in infants. Gastroenterology 2018 Jun;154(8):2045-2059 [FREE Full text] [doi: 10.1053/j.gastro.2018.03.067] [Medline: 29654747]
60. Hoshino Y, Kapikian AZ. Rotavirus vaccine development for the prevention of severe diarrhea in infants and young children. Trends Microbiol 1994 Jul;2(7):242-249. [doi: 10.1016/0966-842x(94)90629-7] [Medline: 8081651]
61. Black RE. Therapeutic and preventive effects of zinc on serious childhood infectious diseases in developing countries. Am J Clin Nutr 1998 Aug;68(2 Suppl):476S-479S. [doi: 10.1093/ajcn/68.2.476S] [Medline: 9701163]
62. Stauffer WM, Konop RJ, Kamat D. Traveling with infants and young children. Part I: anticipatory guidance: travel preparation and preventive health advice. J Travel Med 2001 Sep;8(5):254-259. [doi: 10.2310/7060.2001.22263] [Medline: 11703908]
63. Alebel A, Tesema C, Temesgen B, Gebrie A, Petrucka P, Kibret GD. Prevalence and determinants of diarrhea among under-five children in Ethiopia: a systematic review and meta-analysis. PLoS One 2018 Jun 28;13(6):e0199684 [FREE Full text] [doi: 10.1371/journal.pone.0199684] [Medline: 29953555]

Abbreviations

AIC: Akaike information criterion
AICc: corrected Akaike information criterion
GAM: generalized additive model
GDP: gross domestic product
GLMM: generalized linear mixed model
GWR: geographic weighted regression
LISA: local indices of spatial association
OID: other infectious diarrhea
OLS: ordinary least regression

Edited by A Mavragani; submitted 20.01.23; peer-reviewed by L Lafrado, T Liu; comments to author 12.04.23; revised version received 03.05.23; accepted 05.09.23; published 30.11.23.

Please cite as:
Lu Y, Zhu H, Hu Z, He F, Chen G
Epidemic Characteristics, Spatiotemporal Pattern, and Risk Factors of Other Infectious Diarrhea in Fujian Province From 2005 to 2021: Retrospective Analysis
JMIR Public Health Surveill 2023;9:e45870
URL: https://publichealth.jmir.org/2023/1/e45870
doi:10.2196/45870
PMID:38032713

Original Paper

Evaluation of a Targeted COVID-19 Community Outreach Intervention: Case Report for Precision Public Health

Isela De La Cerda[1], MS, PhD; Cici X Bauer[2], PhD; Kehe Zhang[2], MPH; Miryoung Lee[1], PhD; Michelle Jones[3], MPH; Arturo Rodriguez[3], DNP; Joseph B McCormick[1], MD; Susan P Fisher-Hoch[1], MD

[1]Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health Brownsville Campus, University of Texas Health Science Center at Houston, Brownsville, TX, United States

[2]Department of Biostatistics and Data Science, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, United States

[3]Public Health Department, City of Brownsville, Brownsville, TX, United States

Corresponding Author:
Susan P Fisher-Hoch, MD
Department of Epidemiology, Human Genetics and Environmental Sciences
School of Public Health Brownsville Campus
University of Texas Health Science Center at Houston
780 Ringgold Road
Brownsville, TX, 78520
United States
Phone: 1 956 755 0600
Email: Susan.P.Fisher-Hoch@uth.tmc.edu

Abstract

Background: Cameron County, a low-income south Texas-Mexico border county marked by severe health disparities, was consistently among the top counties with the highest COVID-19 mortality in Texas at the onset of the pandemic. The disparity in COVID-19 burden within Texas counties revealed the need for effective interventions to address the specific needs of local health departments and their communities. Publicly available COVID-19 surveillance data were not sufficiently timely or granular to deliver such targeted interventions. An agency-academic collaboration in Cameron used novel geographic information science methods to produce granular COVID-19 surveillance data. These data were used to strategically target an educational outreach intervention named “Boots on the Ground” (BOG) in the City of Brownsville (COB).

Objective: This study aimed to evaluate the impact of a spatially targeted community intervention on daily COVID-19 test counts.

Methods: The agency-academic collaboration between the COB and UTHealth Houston led to the creation of weekly COVID-19 epidemiological reports at the census tract level. These reports guided the selection of census tracts to deliver targeted BOG between April 21 and June 8, 2020. Recordkeeping of the targeted BOG tracts and the intervention dates, along with COVID-19 daily testing counts per census tract, provided data for intervention evaluation. An interrupted time series design was used to evaluate the impact on COVID-19 test counts 2 weeks before and after targeted BOG. A piecewise Poisson regression analysis was used to quantify the slope (sustained) and intercept (immediate) change between pre- and post-BOG COVID-19 daily test count trends. Additional analysis of COB tracts that did not receive targeted BOG was conducted for comparison purposes.

Results: During the intervention period, 18 of the 48 COB census tracts received targeted BOG. Among these, a significant change in the slope between pre- and post-BOG daily test counts was observed in 5 tracts, 80% (n=4) of which had a positive slope change. A positive slope change implied a significant increase in daily COVID-19 test counts 2 weeks after targeted BOG compared to the testing trend observed 2 weeks before intervention. In an additional analysis of the 30 census tracts that did not receive targeted BOG, significant slope changes were observed in 10 tracts, of which positive slope changes were only observed in 20% (n=2). In summary, we found that BOG-targeted tracts had mostly positive daily COVID-19 test count slope changes, whereas untargeted tracts had mostly negative daily COVID-19 test count slope changes.

Conclusions: Evaluation of spatially targeted community interventions is necessary to strengthen the evidence base of this important approach for local emergency preparedness. This report highlights how an academic-agency collaboration established and evaluated the impact of a real-time, targeted intervention delivering precision public health to a small community.

(JMIR Public Health Surveill 2023;9:e47981) doi:10.2196/47981

KEYWORDS

community interventions; emergency preparedness; health disparities; intervention evaluation; precision public health; public health informatics; public health intervention; public health; spatial epidemiology; surveillance

Introduction

The SARS-CoV-2 pandemic made the dearth of resources and experience to address the situation in local health departments (LHDs) alarmingly apparent. Aggregated data at the state and national levels are insufficiently granular for real-time strategic local interventions. Small communities in particular suffer from limited support and a long turnaround time for obtaining key information needed to conduct effective surveillance and intervention [1]. Texas-Mexico border counties, such as Cameron County, with high burdens of underlying chronic conditions, were heavily impacted by COVID-19 hospitalizations and deaths [2]. By the end of April 2020, Cameron County reported 2.9% case fatalities, exceeding that of its neighboring county, Hidalgo, reporting 0.8%, assuming similarly accurate data for the 2 counties [3]. The LHD of the City of Brownsville (COB), the largest city within Cameron County, accounting for 43% of its total population, recognized specific needs for COVID-19 information dissemination and the resources for assessing community response. The large percentage of the population without insurance, high poverty rates, and elevated rates of type 2 diabetes and obesity (27% and 50%, respectively) contributed substantially to high morbidity and mortality [4,5]. An added obstacle was the cultural and language barriers, given that 85.7% of households speak a language other than English (usually Spanish) and 93.8% are of Hispanic heritage [4]. The COB public health department recognized the limited resources and approached UTHealth Houston School of Public Health, Brownsville (UTHealth), with whom they had a long relationship, to form a partnership to leverage each entity’s strengths to better serve the community [6]. The collaboration began in early April 2020 and had UTHealth faculty and staff helping COB staff conduct COVID-19 case reports and contact tracing tasks. Additionally, weekly sharing of COB in-house COVID-19 case and testing data were accessed by a UTHealth research team to provide census tract-level data and reports to the COB. This team’s innovative acquisition and use of local surveillance data provided the opportunity to strategically target the efforts of the community intervention named “Boots on the Ground” (BOG).

LHDs are the first line of public health action in emergency situations [7]. Their community interventions and actions result in the most immediate response to public health crises. However, resources to effectively implement real-time interventions and assess their impact are rarely available. The “natural experiment” nature of these interventions presents challenges for their evaluation [8]. Experimental methods, such as randomized controlled trials, are considered the gold standard of evaluation. Randomized controlled trials require the random selection of intervention and control groups, which would be costly, impractical, and time-consuming, and raise ethical concerns in regard to respecting community autonomy in crisis situations [9]. Other valuable methods, such as interrupted time series (ITS) studies have been pivotal in health and public policy for evaluating the effects of community-based interventions [10]. ITS studies are used to establish the time series trend of a particular outcome that has been interrupted by an intervention at a specific point in time resulting in self-controlled study units [11]. To our knowledge, there are currently no studies in relation to spatially targeted community interventions for COVID-19 and very few for infectious diseases in general. A 2021 systematic review of spatially targeted community interventions for infectious diseases identified only 10 studies conducted since 1993, none of which properly addressed the evaluation of the intervention [12].

We conducted a study to evaluate the impact of a targeted community COVID-19 education outreach intervention on the SARS-CoV-2 daily test counts by means of an ITS. These findings could help public health professionals and policy makers make authoritative, evidence-based decisions when responding to similar public health crises.

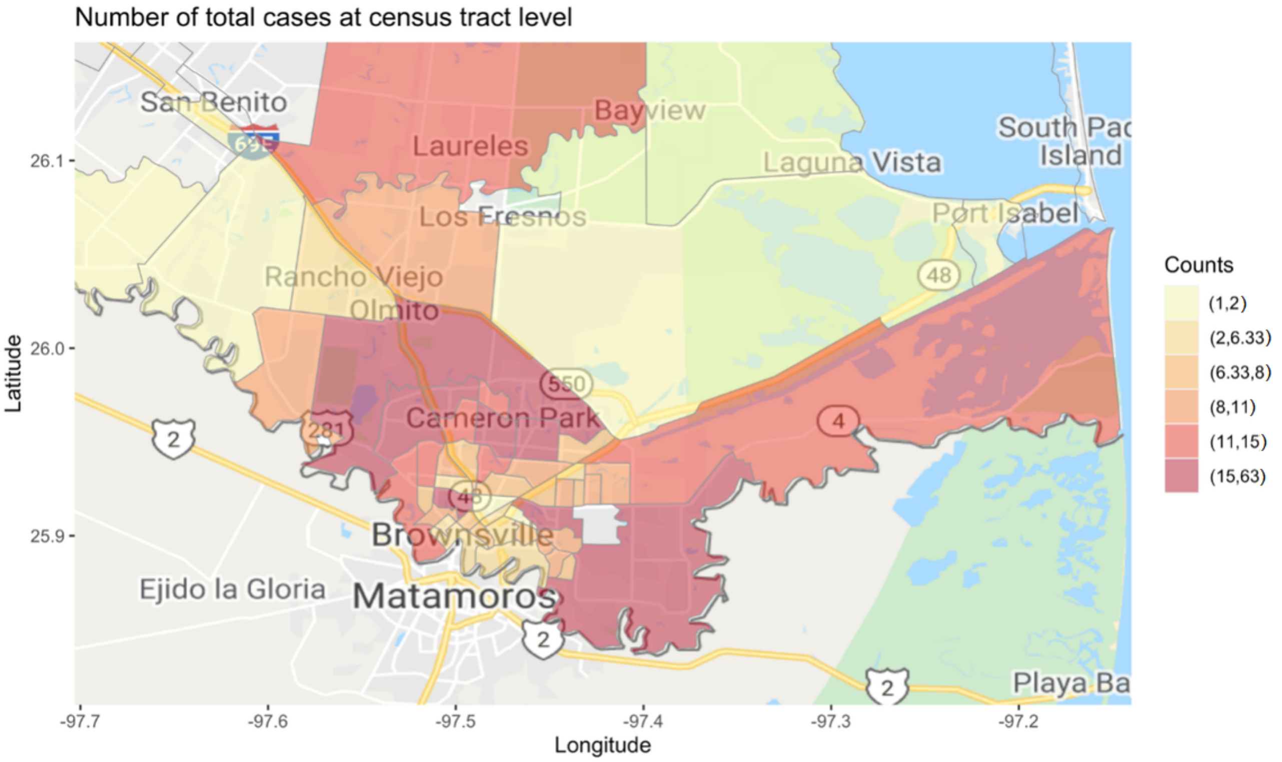
Methods

Boots on the Ground

The first official confirmed COVID-19 case report in the COB was on March 21, 2020. The city opened a municipal SARS-CoV-2 testing site and began its BOG efforts to deliver COVID-19 information to COB residents 2 days later [13]. The initial delivery method of BOG was through blanket events opened to all Brownsville residents at the city public libraries. In early April, the COB joined efforts with epidemiological and advanced statistical expertise at UTHealth, which gave rise to the COVID-19 collaborative working group. Locally collected COB COVID-19 case and testing data included geo-codable but otherwise deidentified data and were securely shared among these 2 working groups on a weekly basis. Applying geospatial science methods, the UTHealth partners used these data to deliver granular COVID-19 maps to the COB (48 census tracts). By April 17, 2020, the first COB census tract COVID-19 case maps (Figure 1) were uploaded by the UTHealth team, and on a weekly basis thereafter.

The COB public health department then used these maps to strategically target their BOG efforts at specific census tracts. The COB epidemiologist selected the census tracts with the highest case counts each week for BOG delivery. If census tracts had similar case counts, the tract with higher obesity prevalence was used to determine which tract was selected. The obesity rates were obtained from the 2019 American Community Survey 5-year estimates [14]. A total of 8 bilingual (English and Spanish) COB public health staff members went door-to-door in the streets of selected census tracts, speaking to residents and leaving packets of COVID-19 information at each house. The information packets were provided in English and Spanish and contained educational information on COVID-19 signs and symptoms as well as details regarding free local SARS-CoV-2 polymerase chain reaction testing sites and available testing opportunities. This delivery method reached a total of 17,170 houses in 18 census tracts. Subsequently, on June 23, delivery efforts shifted to broadcasting methods such as radio, social media, vans with loudspeakers, and news outlets.

Figure 1. COVID-19 case count maps for City of Brownsville census tracts.



Data Sources

The outcome measures chosen to evaluate the impact of BOG were the daily SARS-CoV-2 testing counts. COB COVID-19 testing sites immediately shared all their positive and negative SARS-CoV-2 polymerase chain reaction testing data, which included address information, with the COB. UTHealth experts geocoded addresses to obtain census tract information for each SARS-CoV-2 test, which was then aggregated each day for each of the 18 census tracts. Given the duration of the targeted BOG, from April 21 to June 8, we limited the ITS to only 14 days before and after BOG delivery for each census tract. To avoid a potential spillover effect in our results given the change in BOG delivery method in the last census tract receiving targeted BOG, we limit our time series to only 14 days before and after targeted BOG. For the exposure information, we used the electronic log of BOG efforts maintained by the COB epidemiologist. This log contained the date of targeted BOG delivery for each of the 18 selected census tracts. Based on the log, we created indicator variables representing the pre- and post-BOG day (0 and 1, respectively) status for each tract. An additional variable was created for the time elapsed since BOG delivery for each census tract for the 14 days before and after. We also obtained census tract population size in R Studio's (R Foundation for Statistical Computing) *tidycensus* package using 2015-2019 American Community Survey data [14,15]. This completed the data necessary to conduct the ITS analysis on each of the census tracts.

Ethical Considerations

Data user agreements between the COB and UTHealth were approved by city officials. Data privacy protocols were outlined in the agreement. This study was approved by the UTHealth Committee for the Protection of Human Subjects (HSC-SPH-21-1089). This study was exempt from informed consent given that aggregate level and not subject information was being used.

Statistical Analysis

We hypothesized that the pre- and post-BOG COVID-19 daily test count trend line slope and intercept would significantly increase in BOG-targeted census tracts. For each census tract, we have a series of count data with a known interruption point, which makes piecewise Poisson regression the most suitable model to test our hypothesis [16]. We fit a least squares regression line to each time segment (ie, pre and post) of COVID-19 daily counts, where the change point of the regression line was the day the census tract received targeted BOG, including an offset term of the log of the total population for the census tract. The coefficients of interest in the model were β_2, which estimates the level or intercept (immediate) change in the number of tests per day after the intervention, and β_3, which estimates the change in slope (sustained) trend post BOG compared to the pre-BOG trend along with the P values of the coefficients. Significance for the coefficients was set at a P value of <.05. All statistical analyses were performed in R Studio using the *segmented* package [17,18].

Additional Analysis

To supplement the interpretation of the 18 census tract results, we conducted an additional analysis. The analysis mentioned above was applied to the 30 census tracts in the COB that did not receive BOG during the targeted BOG phase. For these census tracts, we set a date in the middle of the targeted BOG phase as time 0 (May 18, 2020). This additional analysis provided a comparison group to which we compared our targeted BOG subset analysis results.

Results

During the targeted phase of BOG, which lasted from April 21 to June 8, 2020, a total of 18 of the 48 census tracts in the COB received BOG outreach, as shown in Figure S1 in Multimedia Appendix 1. To understand the distribution of the daily test counts for each targeted BOG tract, we included Table S1 in Multimedia Appendix 2 for reference. We obtained the 2 parameters of interest (β_2 and β_3) and their *P* values for each of the piecewise models of the 18 BOG-targeted census tracts. Using a significance *P* value threshold of <.05, significant intercept change (β_2) in testing was observed in 2 census tracts, and significant slope (β_3) change was observed in 5 census tracts. A total of 4 of the 5 tracts with significant slope change had an increase in the slope of the 2-week post-BOG COVID-19 daily test count compared to the pre-BOG slope. A summary of the results is found in Tables 1 and 2.

In the additional analysis of the remaining 30 census tracts in the COB that did not receive BOG outreach during the targeted intervention period, 10 tracts had significant slope changes. A total of 8 of the 10 significant slope changes seen in these census tracts decreased, opposite to what was seen in the slope changes in BOG-targeted census tracts. Table 3 shows the summary of the significant findings and the direction of the coefficients for both groups.

Table 1. Piecewise regression coefficients and *P* values for the change of immediate and sustained COVID-19 daily test count trends for 18 census tracts in the City of Brownsville receiving Boots on the Ground outreach.

Census tract	Immediate		Sustained	
	β_2 coefficient	*P* value	β_3 coefficient	*P* value
1	0.72	.52	–18.15	.99
2	–0.76	.46	1.55	.61
3	0.005	.96	–0.02	.99
4	4.89	.03	0.45	.28
5	–0.76	.48	–21.44	.99
6	0.04	.96	0.06	.99
7	–0.05	.95	0.43	.21
8	1.14	.11	–0.28	.09
9	0.14	.69	–0.45	<.001
10	–0.45	.62	0.27	.54
11	–0.69	.43	–19.98	.24
12	–2.53	.06	0.77	.83
13	–0.27	.73	0.09	.43
14	1.48	.25	0.48	.11
15	0.07	.92	0.34	.02
16	–1.52	.001	0.18	<.001
17	0.44	.15	0.25	.005
18	0.45	.99	0.10	<.001

Table 2. Piecewise regression coefficients and *P* values for the change of immediate and sustained COVID-19 daily test count trends for 30 census tracts in the City of Brownsville not receiving “Boots on the Ground” outreach.

Census tract	Immediate		Sustained	
	β_2 coefficient	*P* value	β_3 coefficient	*P* value
1	1.26	.06	0.11	.12
2	0.59	.03	–0.38	.004
3	–1.02	.08	0.18	.54
4	1.32	.04	0.07	.72
5	0.06	.90	–0.43	.57
6	1.28	.01	–0.21	<.001
7	0.17	.80	–18.79	.93
8	0.79	.19	0.37	.80
9	–0.73	.20	0.66	.08
10	0.20	.76	0.15	.16
11	–0.05	.92	–0.17	.99
12	2.50	<.001	–1.36	<.001
13	–0.16	.89	–0.07	.11
14	–0.33	.55	–1.41	<.001
15	1.28	.05	–0.42	.01
16	–0.09	.88	–0.18	.002
17	0.09	.85	0.59	<.001
18	–0.03	.95	–0.95	.47
19	–0.33	.75	0.23	.40
20	1.35	.04	–0.06	.94
21	0.51	.53	1.02	.90
22	2.07	.02	1.17	.08
23	0.13	.86	0.93	.29
24	0.44	.64	0.52	.99
25	–2.83	<.001	–1.79	<.001
26	–0.91	.14	–0.48	.78
27	–0.62	.50	–20.65	.37
28	0.61	.54	0.12	.83
29	–0.72	.25	–1.10	.37
30	0.23	.58	1.94	.002

Table 3. Direction of coefficients in the significant trend changes observed.

	Received targeted BOG[a]		Did not receive targeted BOG	
	Immediate (n=2), n	Sustained (n=5), n	Immediate (n=8), n	Sustained (n=10), n
Positive impact	1	4	7	2
Negative impact	1	1	1	8

[a] BOG: Boots on the Ground.

Discussion

Principal Findings

The use of local COVID-19 surveillance data, along with collaboration between local public health departments and academia to target a COVID-19 educational outreach intervention, resulted in significant impacts on daily COVID-19 test counts. Most notable was the increase in the sustained trend of testing observed after BOG in census tracts that received targeted BOG. A total of 4 out of 5 census tracts showed significant increased slope changes in testing trends, meaning that the 2-week post-BOG COVID-19 daily test count slope improved significantly from the 2-week pre-BOG slope. Interestingly, the additional analysis of nontargeted BOG COB census tracts was contrary to what we observed in the BOG-targeted tracts. Here, only 20% (2/10) of the census tracts showed a significant increase in sustained testing. This suggests that daily COVID-19 test-seeking in BOG-targeted census tracts saw a sustained improvement (2 weeks after intervention), whereas test-seeking mostly declined in census tracts that did not receive BOG. The perspective of a COB resident documented in an autoethnographic analysis of BOG provides some insight into our interesting findings [6]. The personal method by which targeted BOG census tracts received tailored information potentially reflected a sense of trust among Brownsville residents and made them more receptive to act on the information they were receiving. The long-term benefits of this intervention were observed during the COVID-19 vaccine rollout in early 2021. As BOG delivery shifted later to social media and broadcasting delivery, a city-specific website called BTX Cares was created and served as a trusted source of SARS-CoV2 testing and vaccine information for COB residents [19]. The trust built through the door-to-door delivery of BOG and the use of BTX Cares for COVID-19 information resulted in successful COVID-19 mass vaccination clinic events hosted by the city.

Limitations

There are some limitations to consider in this study. Although we did account for census tract population size, the short time series evaluated prevented us from establishing or adjusting for temporal or seasonal confounders that might arise in time series analysis. The lack of a control group for comparison is an issue to be considered in this case report. We addressed this limitation by providing a comparison group—those census tracts in the COB that did not receive targeted BOG outreach [20]. It is also possible that selecting census tracts with higher rates of infection also drove testing to some extent, but since the infection levels were not dramatically different between tracts in this small city, this effect is likely to have been limited. Our findings provide grounds for further research in the delivery of translational science through targeted public health.

Conclusions

The impact on LHDs of COVID-19 in 2020 was extraordinary in its scale, gravity, and range of challenges it presented. Many years of limited public health funding meant that many counties and cities were underresourced, undertrained, and therefore ill-prepared, for the crisis presented by the pandemic. Addressing immediate needs requires collaboration and new technology. Overall, this program was based on the cooperation of the community with local authorities and with the academic center and ultimately resulted in a trusted channel of communication between city residents and public health professionals that persisted during the COVID-19 vaccine rollout. Combining academia and public health institutions was one way to address the many issues. The actual process was in fact inexpensive and relatively simple in overall execution, even though analytic methods were relatively new and sophisticated. This experience was based on a long-term relationship dating back to the H1N1 influenza outbreak, when similarly, a lack of local resources and data were considerable barriers to effective interventions [1]. What we present here is a case for precision public health that uses new technologies to improve local public health efforts by generating tailored and spatially targeted interventions. Our experiences and findings advocate not only for strong agency-academic collaborations but also for improved, granular population surveillance data sources on a national scale that can be leveraged to target interventions and deliver the correct intervention to the high-risk population in a timely manner [21,22].

Acknowledgments

We would like to acknowledge the efforts of the City of Brownsville public health staff that made this spatially targeted community intervention possible by delivering this intervention to their community. We would also like to thank the endless support from city officials in all our research endeavors, in particular Art Rodriguez, Michelle Jones, and Alvaro Silva from the City of Brownsville.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to data user agreements between the City of Brownsville and UTHealth Houston School of Public Health, Brownsville, but are available from the corresponding author on reasonable request.

Authors' Contributions

IDLC conceived and designed the analysis. AR, MJ, and ML contributed to data acquisition. IDLC and KZ contributed to data curation and processing. IDLC and CXB conducted the data analysis. IDLC, CXB, and ML contributed to the interpretation of the results. IDLC and SPFH drafted the initial manuscript. CXB, SPFH, ML, MJ, JBM, and IDLC contributed critical revisions through the development of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Map of all Cameron County census tracts. Orange colored census tracts are City of Brownsville (COB) tracts that received targeted Boots on the Ground (BOG). Yellow colored census tracts are COB tracts that did not receive targeted BOG.

[DOCX File , 1575 KB - publichealth_v9i1e47981_app1.docx]

Multimedia Appendix 2

24-day SARS-CoV-2 polymerase chain reaction daily test count descriptive statistics for each targeted Boots on the Ground census tract.

[DOCX File , 27 KB - publichealth_v9i1e47981_app2.docx]

References

1. McCormick JB, Yan C, Ballou J, Salinas Y, Reininger B, Gay J, et al. Response to H1N1 in a U.S.-Mexico border community. Biosecur Bioterror 2010;8(3):233-242 [FREE Full text] [doi: 10.1089/bsp.2010.0014] [Medline: 20825334]
2. Cargo K. TAMUCC: COVID-19 transmission rate is increasing in South Texas. Corpus Christi Caller-Times. URL: https://www.caller.com/story/news/local/2020/05/22/tamucc-covid-19-transmission-rate-increasing-south-texas/5244508002/ [accessed 2023-02-14]
3. Historical Texas COVID-19 data. Texas Department of State Health Services. URL: https://dshs.texas.gov/coronavirus/AdditionalData.aspx [accessed 2022-01-26]
4. QuickFacts: Brownsville city, Texas. U.S. Census Bureau. URL: https://www.census.gov/quickfacts/fact/table/brownsvillecitytexas/PST045221 [accessed 2023-02-14]
5. Fisher-Hoch SP, Rentfro AR, Salinas JJ, Pérez A, Brown HS, Reininger BM, et al. Socioeconomic status and prevalence of obesity and diabetes in a Mexican American community, Cameron County, Texas, 2004-2007. Prev Chronic Dis 2010;7(3):A53 [FREE Full text] [Medline: 20394692]
6. Blackburn CC, Jones M, Lee M, Bauer C, Rodriguez A, Garcia R, et al. Autoethnographic examination of data-driven, community-tailored COVID-19 response in Brownsville, Texas. J Humanist Psychol 2022:00221678221118708. [doi: 10.1177/00221678221118708]
7. Institute of Medicine, Board on Health Promotion and Disease Prevention. Public health agencies: their roles in educating public health professionals. In: Gebbie K, Rosenstock L, Hernandez L, editors. Who Will Keep the Public Healthy? Educating Public Health Professionals for the 21st Century. Washington (DC): National Academies Press; 2003.
8. Khullar D, Jena AB. "Natural experiments" in health care research. JAMA Health Forum 2021;2(6):e210290 [FREE Full text] [doi: 10.1001/jamahealthforum.2021.0290] [Medline: 36218753]
9. Lindholm L, Rosén M. What is the "golden standard" for assessing population-based interventions?—Problems of dilution bias. J Epidemiol Community Health 2000;54(8):617-622 [FREE Full text] [doi: 10.1136/jech.54.8.617] [Medline: 10890874]
10. Biglan A, Ary D, Wagenaar AC. The value of interrupted time-series experiments for community intervention research. Prev Sci 2000;1(1):31-49 [FREE Full text] [doi: 10.1023/a:1010024016308] [Medline: 11507793]
11. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017;46(1):348-355 [FREE Full text] [doi: 10.1093/ije/dyw098] [Medline: 27283160]
12. Khundi M, Carpenter JR, Nliwasa M, Cohen T, Corbett EL, MacPherson P. Effectiveness of spatially targeted interventions for control of HIV, tuberculosis, leprosy and malaria: a systematic review. BMJ Open 2021;11(7):e044715 [FREE Full text] [doi: 10.1136/bmjopen-2020-044715] [Medline: 34257091]
13. City of Brownsville, TX-Municipal Government. Facebook. 2022. URL: https://www.facebook.com/watch/?v=536562520555903 [accessed 2022-01-26]
14. American Community Survey (ACS). The United States Census Bureau. 2018. URL: https://www.census.gov/programs-surveys/acs [accessed 2020-09-23]
15. Walker K, Herman M, Eberwein K. tidycensus: load US census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames. The Comprehensive R Archive Network. 2022. URL: https://CRAN.R-project.org/package=tidycensus [accessed 2022-11-22]
16. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther 2002;27(4):299-309 [FREE Full text] [doi: 10.1046/j.1365-2710.2002.00430.x] [Medline: 12174032]
17. Muggeo VM. Segmented: an R package to fit regression models with broken-line relationships. R News 2008;8(1):20-25.
18. Posit. 2022. URL: https://www.posit.co/ [accessed 2023-11-30]

19. Brownsville Cares. 2021. URL: https://btxcares.com/ [accessed 2023-01-19]
20. Bottomley C, Scott JAG, Isham V. Analysing interrupted time series with a control. Epidemiol Methods 2019;8(1):20180010 [FREE Full text] [doi: 10.1515/em-2018-0010]
21. Buckeridge DL. Precision, equity, and public health and epidemiology informatics - a scoping review. Yearb Med Inform 2020;29(1):226-230 [FREE Full text] [doi: 10.1055/s-0040-1701989] [Medline: 32823320]
22. Kee F, Taylor-Robinson D. Scientific challenges for precision public health. J Epidemiol Community Health 2020;74(4):311-314 [FREE Full text] [doi: 10.1136/jech-2019-213311] [Medline: 31974295]

Abbreviations

BOG: Boots on the Ground
COB: City of Brownsville
ITS: interrupted time series
LHD: local health department
UTHealth: UTHealth Houston School of Public Health, Brownsville

Edited by A Mavragani, T Sanchez; submitted 06.04.23; peer-reviewed by H Zhu, A Ramaprasad, C Hudak; comments to author 03.08.23; revised version received 18.08.23; accepted 17.10.23; published 20.12.23.

Please cite as:
De La Cerda I, Bauer CX, Zhang K, Lee M, Jones M, Rodriguez A, McCormick JB, Fisher-Hoch SP
Evaluation of a Targeted COVID-19 Community Outreach Intervention: Case Report for Precision Public Health
JMIR Public Health Surveill 2023;9:e47981
URL: https://publichealth.jmir.org/2023/1/e47981
doi:10.2196/47981
PMID:38117549

Original Paper

Effects of Greenness on Myopia Risk and School-Level Myopia Prevalence Among High School–Aged Adolescents: Cross-sectional Study

Chang Zhang[1], PhD; Cheng Wang[1], PhD; Xin Guo[2,3], PhD; Huiyu Xu[2,3], MSc; Zihao Qin[4], BA; Liyuan Tao[5,6], PhD

[1]Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China
[2]Institute of School Health, Beijing Center for Disease Prevention and Control, Beijing, China
[3]School of Public Health, Capital Medical University, Beijing, China
[4]Faculty of Medicine, University of Queensland, Brisbane, Queensland, Australia
[5]Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing, China
[6]Medical Examination Centre, Peking University Third Hospital, Beijing, China

Corresponding Author:
Liyuan Tao, PhD
Research Center of Clinical Epidemiology
Peking University Third Hospital
49 North Garden Rd
Haidian District Beijing
Beijing, 100191
China
Phone: 86 1082265732
Email: tendytly@163.com

Abstract

Background: Myopia is a serious public health issue. High school–aged adolescents in Beijing have an alarming prevalence of myopia. Therefore, determining myopia protective factors is essential. Green space has a certain association with myopia protective factors that can protect against myopia.

Objective: This study aims to examine the effects of green space around schools on individual myopia risk in high school–aged adolescents and the school-level myopia prevalence.

Methods: Green space was measured using the normalized difference vegetation index (NDVI). A total of 13,380 samples of 51 high schools were selected from a 2021 Beijing Municipal Health Commission survey. Adolescent myopia was defined as a spherical equivalent of ≤–1.00 diopters in the worse eye. Generalized linear mixed models with a binomial error structure were used to analyze the effects of the NDVI on personal myopia risk and adjust them by other factors, such as demographics, exposure time, and outdoor exercise. The effects of the NDVI on school-level myopia prevalence with adjusted demographics and the relative position factors of trees were analyzed through quasibinomial regression.

Results: The overall prevalence of myopia was 80.61% (10,785/13,380, 95% CI 79.93%-81.27%). Although with a 0.1 increase in the 500 and 1000 m buffer NDVIs adjusted by demographic and other factors, the high school–aged personal myopia risk significantly dropped by 16% (odds ratio [OR] 0.84, 95% CI 0.73-0.97) and 12% (OR 0.88, 95% CI 0.79-0.99), respectively. However, only the adjusted 500 m buffer NDVI (by demographics and the position of trees) with a 0.1 increase significantly reduced the school-level myopia prevalence by 15% (OR 0.85, 95% CI 0.74-0.98). Subgroup analysis showed that the adjusted effects of the 500 m buffer NDVI are significant in schoolgirls (OR 0.82, 95% CI 0.72-0.93), juniors (OR 0.82, 95% CI 0.72-0.94), the Han nationality (OR 0.84, 95% CI 0.72-0.97), 1-year exposure (OR 0.84, 95% CI 0.71-0.99) and 3-year exposure (OR 0.78, 95% CI 0.65-0.94).

Conclusions: The greenness of a 500 m buffer around schools is associated with a lower personal myopia risk among adolescents and a lower prevalence of myopia in schools. With regard to prevention and control activities, green space within a 500 m buffer around schools is suggested as an independent protective factor for adolescent myopia.

(JMIR Public Health Surveill 2023;9:e42694) doi:10.2196/42694

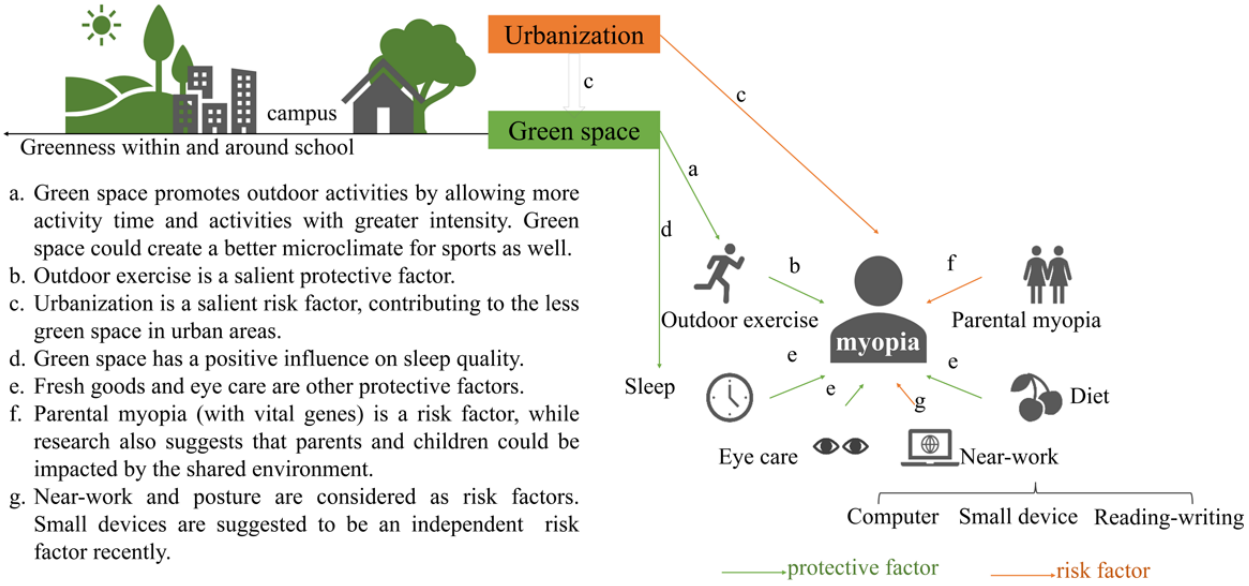
KEYWORDS

high school–aged adolescent; personal myopia risk; school-level myopia prevalence; green space; adjusted effects

Introduction

Myopia, the most common refractive error leading to short-sightedness or nearsightedness [1,2], affects daily life and increases the risk of blindness [3,4]. This serious public health issue is likely to affect approximately half of the world's population by 2050 [5]. Therefore, knowledge of the factors affecting myopia is essential. The genetic and nongenetic factors underlying myopia have been examined [6,7], and nongenetic factors (multiple types) may be more important [8,9]. Research has revealed that urbanization [10,11], outdoor exercise [12-14], parental myopia [15,16], near-work (sustaining a close gaze for reading and writing, using computers, or using small electronic devices) time, body posture [17,18], sleep [19,20], eye care, and diet [21] all impact myopia. These influences vary with demographics [9,22,23] and geographical location [1]. Green space is associated with these factors (Figure 1). Urbanization is considered a risk factor for myopia, possibly due to less green space in urban areas than in suburban or rural areas. Outdoor exercise is a protective factor. More green space promotes more outdoor exercise [24,25] and may reduce near-work time [14]. Moreover, people with higher levels of green space have a lower prevalence of insufficient sleep [26,27]. Hence, greenness may contribute to protection against myopia.

Figure 1. Based on previous studies, greenness may play a role in vision protection against myopia. The green and orange arrow lines indicate protective and risk factors according to previous studies, respectively. Points b, c, e, f, and g represent the influencing factors on myopia, while points a and d explain the associations between green space and influencing factors.



a. Green space promotes outdoor activities by allowing more activity time and activities with greater intensity. Green space could create a better microclimate for sports as well.
b. Outdoor exercise is a salient protective factor.
c. Urbanization is a salient risk factor, contributing to the less green space in urban areas.
d. Green space has a positive influence on sleep quality.
e. Fresh goods and eye care are other protective factors.
f. Parental myopia (with vital genes) is a risk factor, while research also suggests that parents and children could be impacted by the shared environment.
g. Near-work and posture are considered as risk factors. Small devices are suggested to be an independent risk factor recently.

The school is the principal space for school-aged children and adolescents. Previous studies have shown that the greenness level of a campus and the surrounding environment significantly impact students' physical and mental health [28-30]. Recently, researchers have investigated the relationship between school green space and students' myopia [31-33]. Studies focusing on the outdoor surrounding greenness within or around school boundaries have found a negative association between green space exposure and myopia among the school-aged population [34]. However, the evidence for this effect on junior and senior high school students is weak (eg, insufficient study quantity, lacking consideration of trees' distribution in green space). Therefore, there are still gaps in the controlling strategies for high school–aged students' myopia with regard to green space as a protective factor, and more research is warranted.

Our study focused on high school–aged adolescents in Beijing, a population with an alarming prevalence of myopia (once reaching 98.27% in a cross-sectional investigation of senior grade 3 students) [35-37]. We hypothesized that the greenness within or around the school can protect against myopia in the junior and senior high school–aged population. We aimed to determine (1) whether greenness significantly varies between myopia and nonmyopia groups, (2) how greenness acts on individual myopia risk and school myopia prevalence after adjusting for other known influential factors, and (3) whether the greenness protective effect is similar among demographic subgroups. A clarification of these points could elucidate how green space affects myopia in the high school–aged population and potentially contribute to targeted myopia prevention and intervention measures.

Methods

Study Population

The study sample was taken from the monitoring project for common diseases (eg, myopia, obesity) in students and health influence factors conducted by the Beijing Municipal Health Commission in 2021. Various schools involved in this monitoring project were randomly selected from all districts of Beijing. Our study only focused on high school adolescents (junior and senior high schools). We obtained myopia and questionnaire data from 51 high schools and 13,380 students.

Among the 51 schools, 28 (55%) schools were junior high schools, 20 (39%) were senior high schools, and 3 (6%) contained both junior and senior students. Since these 3 school campuses were shared by junior and senior students, there were 48 school plots in total (Figure S1 in Multimedia Appendix 1).

Ethical Considerations

The study was approved by the Ethical Committee of the Peking University Third Hospital (IRB00006761-M2021281) and conducted according to the guidelines of the Declaration of Helsinki.

Myopia Determination

The students included in this study underwent an ophthalmological examination at the school infirmary. The examinations were conducted in accordance with the standards of myopia screening for children and adolescents issued by the Beijing Center for Disease Prevention and Control. The examinations included far-vision examination and refractive examination and were performed by an ophthalmologist having a Chinese national medical practitioner certificate in ophthalmology. Refraction was detected using a desktop noncycloplegic autorefractometer, and all refractometers were accredited to the ISO 10342:2010 standard for ophthalmic instruments and eye refractometers. Adolescent myopia was defined as a spherical equivalent of ≤–1.00 diopters in the worse eye. The worse eye was defined as the eye with the greatest absolute value of the spherical equivalent of the refractive error.

School Green Space

We used the normalized difference vegetation index (NDVI) to measure the schools' green space. The NDVI value ranges from –1 to 1; the higher the NDVI, the higher the level of school green space. GF-2 satellite imagery (multispectral with 3.2 m resolution) was used in our study to calculate the NDVI values. In total, 25 single GF-2 cloud-free images (entirely covering 48 school plots and within 2 years around the myopia data collection time) were selected (data searched from the China Center for Resources Satellite Data and Application [38]). Through radiometric calibration, atmospheric correction and orthorectified (PRC parameters), these images were used to determine the NDVI by brand calculation, according to the following formula: NDVI = (NIR –RED)/(NIR + RED), where RED is the red band and NIR is the near-infrared band. Next, we extracted each plot's average NDVI of buffers (ie, 500 and 1000 m) around the school (Figure S1 in Multimedia Appendix 1). In addition, we analyzed the positions of trees relative to school (teaching) buildings by distinguishing the distribution of trees separated from the buildings with a small space from that with a large space.

Individual Factors

Demographic data were obtained, including birth data (used to convert the sample's age by the current month and year plus the birth month and year), gender (male and female categories), grade (representing exposure time to the school environment: senior/junior grade 1, indicating 1 year of exposure time; senior/junior grades 2 and 3, representing 2 and 3 years of exposure time, respectively), nationality (Han nationality and other categories), place of origin (urban, ie, 6 core urban districts: Dongchen, Xicheng, Haidian, Shijingshan, Fengtai, and Chaoyang [39]; suburban, ie, other 10 districts). In addition, students' heights and weights were collected to calculate their BMI according to the "Screening for Overweight and Obesity Among School-Age Children and Adolescents" protocol (WS/*t* 586—2018).

Information about students' outdoor exercise, including outdoor time on general days (Q1), outdoor time during holidays (Q2), preference for outdoor places at school (Q3), outdoor exercise intensity (Q4), and the number of physical education (PE) classes (Q5); near-work and body gestures, including posture-related distance from the table (Q6) and distance from the screen (Q7), adjusted chair height (Q8), parental control (Q9), and use of electronic devices (Q10); eye care, including the number of daily eye exercises at school (Q11) and the eye rest interval (Q12); dietary and sleep factors, describing sugar (Q13), fresh fruit (Q14), and fresh vegetables (Q15) consumed in the past 7 days; and the average daily sleeping hours (Q16) were acquired from the questionnaire (for details of questions Q1-Q16, see Multimedia Appendix 2).

Statistical Analysis

We analyzed differences between myopia and nonmyopia groups at the single-factor level. Categorical variables (ie, gender, nationality, place of origin, exposure time, position of trees, outdoor exercise factors [Q1-Q3], near-work and body gesture factors [Q6-Q10], eye-care factors [Q11-Q12], and dietary factors [Q13-Q15]) were analyzed using chi-square tests to determine statistical significance, with a 2-sided P value of <.05 as the threshold. Numeric variables (ie, NDVI values, outdoor exercise and sleep factors (Q4, Q5, Q16), age, and BMI) were analyzed using the median Wilcoxon test (P<.05 was considered significant).

Generalized linear mixed models with a binomial error structure were used to test the effects of the NDVI on personal myopia risk. The NDVI and all adjusted factors (ie, demographic factors, outdoor exercise, near-work and body gestures, eye care, diet, sleep, exposure time, and the position of trees) were considered fixed effects. To control for the influence of school variation on adolescents at the personal level, we set "school" as a random factor. Moreover, to assess the influence of different NDVI ranges and their adjusted effects, we developed 6 models. Models 1a, 2a, and 3a represented the regressions of the NDVI (within the school, 500 m buffer around the school, 1000 m buffer around the school, respectively) to myopia. Based on this, adjusted factors were added to Models 1b (ie, Model 1a with the addition of adjusted factors), 2b (ie, Model 2a with the addition of adjusted factors), and 3b (ie, Model 3a with the addition of adjusted factors). For the random and fixed effects for the adjusted NDVI of Models 1b, 2b, and 3b, see Figures S1 and S2 in Multimedia Appendix 3. We used the tolerance and variance inflation factor (VIF) to test collinearity among variables, with tolerance>0.1 and VIF<5 considered as no collinearity. We used the Durbin-Watson (DW) value to test the independence of residuals (with a DW value not close to 0 or 4 meaning that the variables were independent). To analyze the distinct performance of these effects in subgroups, we performed the same generalized linear mixed models under different subgroups divided by school type, gender, nationality, place of origin, and grade (converted to exposure time).

When considering the effects of the NDVI on myopia prevalence at the school level, we first calculated each school's myopia prevalence (myopia students/total students in that school). Second, 6 generalized linear models with a quasibinomial error structure (the quasibinomial was used to replace the binomial for existing overdispersion) were developed in the same way as the regressions earlier. Models 4a, 5a, and 6a only considered the NDVI of 3 buffer scales separately, and Models 4b, 5b, and 6b were the quasibinomial regressions (weighted with the number of students in each school) of the NDVI (within the school, 500 m buffer around the school, 1000 m buffer around the school, respectively) to the school-level myopia prevalence with adjusted factors (ie, the school's average age, gender [proportion of female students], place of origin [urban or suburban], BMI, and position of trees).

Data analyses and visualization were performed using R version 4.1.3 (R Core Team), mainly using the sjplot [40], lme4 [41], and ggplot2 [42] packages.

Results

Myopia Prevalence

The 13,380 students included in this study were aged 11.9-20.9 years (mean 14.86, SD 1.71), and among them, 6803 (50.84%) were male and 6577 (49.16%) female. There were 7850 (58.67%) junior high school-aged students (n=2663, 33.92%, grade 1, n=2610, 33.25%, grade 2, and n=2577, 32.83%, grade 3) and 5530 (41.33%) seniors (n=1908, 34.50%, grade 1, n=1841, 33.29%, grade 2, and n=1781, 32.21%, grade 3). Of the 13,380 adolescents, 5245 (39.20%) were from urban areas and 8135 (60.80%) from suburban areas. In addition, 12,004 (89.72%) of these students were of Han nationality, and 1376 (10.28%) were of other nationalities. According to grade, the exposure times in school were divided into 1, 2, and 3 years, accounting for 4571 (34.16%), 4451 (33.27%), and 4358 (32.57%) of the students, respectively.

The overall myopia prevalence of the high school–aged students was 80.61% (10,785/13,380, 95% CI 79.93%-81.27%). The prevalence of myopia in seniors (4754/5530, 85.97%, 95% CI 84.03%-86.86%) was higher than that in juniors (6031/7850, 76.83%, 95% CI 75.88%-77.75%); see Table 1. With the exception of nationality (*P*=.63) and BMI (*P*=.47), the prevalence of myopia varied significantly with age, gender, and place of origin (*P*<.001); see Table 1. Adolescents with myopia were more likely to be (1) older (median age 14.78 years, IQR 13.75-16.46) compared to those without myopia (median age 14.05 years, IQR 13.01-15.60), (2) schoolgirls (myopia prevalence 5548/6577, 84.4%, 95% CI 83.45%-85.21%) compared to schoolboys (myopia prevalence 5237/6803, 77%, 95% CI 75.96%-77.96%), and (3) urban students (myopia prevalence 4406/5245, 84%, 95% CI 82.98%-84.97%) compared to suburban students (myopia prevalence 6379/8135, 78.4%, 95% CI 77.5%-79.29%).

Table 1. Description of myopia and nonmyopia groups.

Characteristics	Total (N=13,380)	Myopia group (n=10,785, 80.61%)	Nonmyopia group (n=2595, 19.39%)	*P* value[a]
Age (years), median (IQR)	14.63 (13.42-16.31)	14.78 (13.75-16.46)	14.05 (13.01-15.60)	<.001
BMI (kg/m^3), median (IQR)	21.5 (18.9-25.3)	21.4 (18.9-25.2)	21.6 (18.8-25.7)	.47
Gender, n (%)				<.001
Male	6803 (50.84)	5237 (48.56)	1566 (60.35)	N/A[b]
Female	6577 (49.16)	5548 (51.44)	1029 (39.65)	N/A
Nationality, n (%)				.63
Han nationality	12,004 (89.72)	9683 (89.78)	2321 (89.44)	N/A
Other nationalities	1376 (10.28)	1102 (10.22)	274 (10.56)	N/A
School type, n (%)				<.001
Junior	7850 (58.67)	6031 (55.92)	1819 (70.10)	N/A
Senior	5530 (41.33)	4754 (44.08)	776 (29.90)	N/A
Place of origin, n (%)				<.001
Urban area	5245 (39.20)	4406 (40.85)	839 (32.33)	N/A
Suburban area	8135 (60.80)	6379 (59.15)	1756 (67.67)	N/A
Exposure time (years), n (%)				<.001
1	4571 (34.16)	3507 (32.52)	1064 (41.00)	N/A
2	4451 (33.27)	3581 (33.20)	870 (33.53)	N/A
3	4358 (32.57)	3697 (34.28)	661 (25.47)	N/A
Position of trees, n (%)				<.001
From buildings with a small space	4857 (36.30)	4018 (37.26)	839 (32.33)	N/A
From buildings with a large space	8523 (63.70)	6767 (62.74)	1756 (67.67)	N/A

[a] Age and the BMI were reported with *P* values in the Wilcoxon test. Other variables were reported with *P* values in the chi-square test.

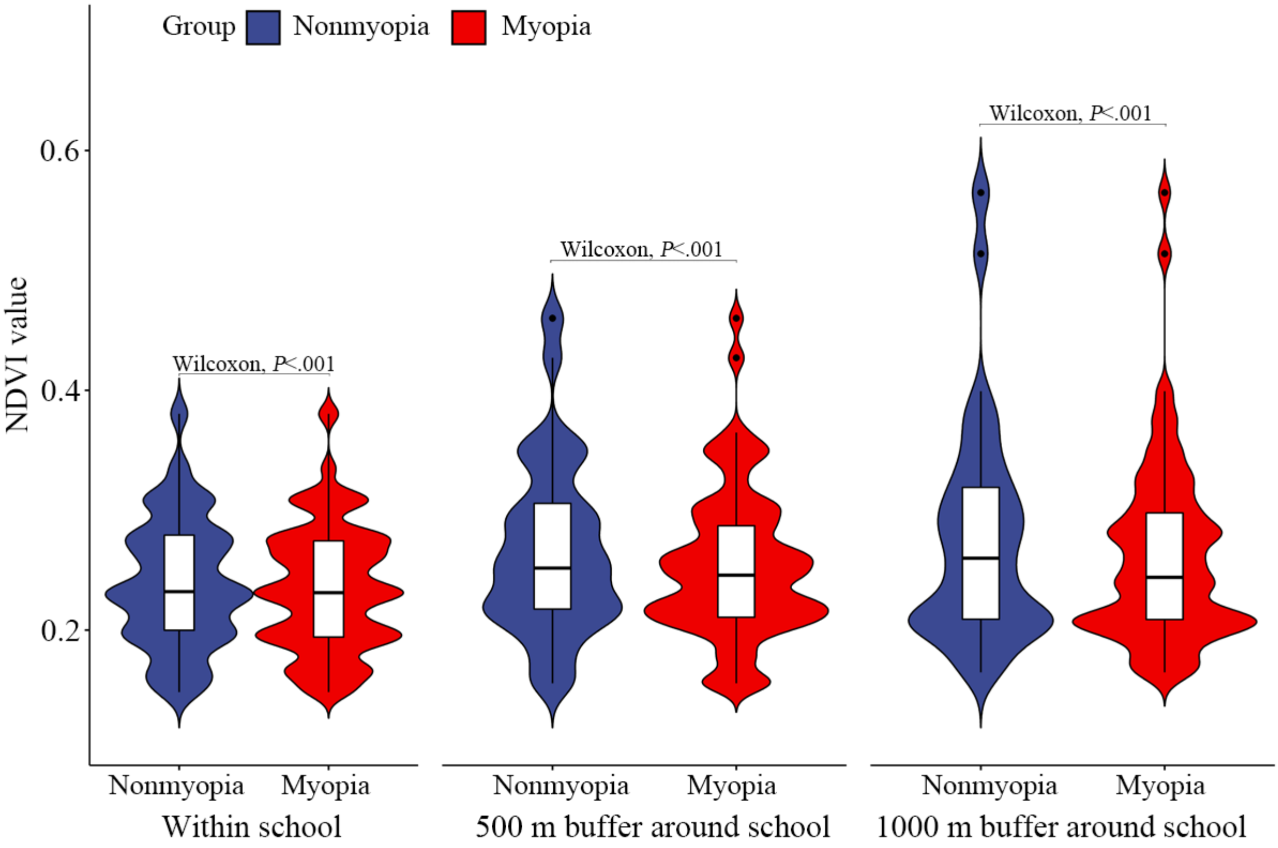
[b] N/A: not applicable.

Factors Influencing the Prevalence of Myopia

We extracted the NDVI within each school boundary and of the 500 and 1000 m buffers (Figure S2 in Multimedia Appendix 1). Significant differences were found in school greenness between the myopia group and the nonmyopia group (*P*<.001); see Figure 2. In the myopia group, the NDVI (1) within the school (median 0.231, IQR 0.194-0.274), (2) within the 500 m buffer (median 0.279, IQR 0.247-0.318), and (3) within the 1000 m buffer (median 0.281, IQR 0.248-0.331) was lower compared to the nonmyopia group (median 0.232, IQR 0.2-0.279; median 0.285, IQR 0.253-0.335; and median 0.296, IQR 0.249-0.351, respectively).

Furthermore, trees located too close to teaching buildings were significantly associated with a higher myopia prevalence (4018/4857, 82.73%, 95% CI 81.64%-83.77%) than trees at a distance with enough space from the buildings (6767/8523, 79.4%, 95% CI 78.53%-80.25%); see Table 1. Moreover, there were significant differences concerning other factors between students with and without myopia (*P*<.001). Students without myopia were likely associated with a higher outdoor level (Q2-Q5), less near-work time with appropriate body posture (Q6-Q8, Q10), better eye care (Q11-Q12), sufficient sleep time (Q16), and high fresh-good intake (Q14); see Multimedia Appendix 2.

Figure 2. Differences in the NDVI between the myopia and nonmyopia groups. The probability density with the minimum, maximum, median, and quartiles are shown in this figure. The red color represents the myopia group, and the blue color represents the nonmyopia group. “Wilcoxon, *P*<.001” means a significant difference in the median by the Wilcoxon test. NDVI: normalized difference vegetation index.



Effects of Influencing Factors on Adolescent Myopia Risk

According to the fixed effects of generalized linear mixed effect models (Table 2), the NDVI of buffers around the school significantly influenced the risk of adolescent myopia without any adjusted factors. With a 0.1 increase in the NDVI of a 500 m buffer (odds ratio [OR] 0.74, 95% CI 0.61-0.89, Model 2a) and a 1000 m buffer (OR 0.79, 95% CI 0.68-0.92, Model 3a) surrounding the school, the adolescent myopia prevalence decreased by 26% and 21%, respectively. However, the decrease in myopia prevalence caused by the 500 m (OR 0.84, 95% CI 0.73-0.97, Model 2b) or the 1000 m (OR 0.88, 95% CI 0.79-0.99, Model 3b) buffer NDVI dropped by 16% and 12%, respectively, due to the adjustment. In addition, neither the NDVI (OR 0.84, 95% CI 0.66-1.06, Model 1a) nor the adjusted NDVI (OR 0.88, 95% CI 0.75-1.04, Model 1b) within the school showed a significant association with the adolescent myopia risk.

Considering all adjusted factors (Table S1 in Multimedia Appendix 4), similar effects of age, gender, exposure time, place of origin, outdoor time during holidays (Q2), preference for outdoor place (Q3), posture-related distance from the table while doing near-work (Q6), and eye-rest interval (Q12) were found to be significantly associated with myopia when adjusting the 500 and 1000 m buffer NDVI. Of these factors, age (OR 1.13, 95% CI 1.06-1.2, Models 2b and 3b) and female gender (OR 1.52, 95% CI 1.38-1.66, Models 2b and 3b) had positive associations with myopia. A close range for reading (Q6; OR 1.22, 95% CI 1.03-1.44, Models 2b and 3b) and an eye-rest interval of 30-60 minutes compared to 15 minutes or less (Q12; OR 1.21, 95% CI 1.05-1.39, Models 2b and 3b) were risk factors for myopia. The outdoor time on holidays was negatively associated with the personal risk of myopia, and persisting in maintaining the outdoor time on holidays reduced the personal risk of myopia by 33% (OR 0.67, 95% CI 0.55-0.82, Models 2b and 3b) compared to hardly doing so. Spending break time outdoors decreased the personal risk of myopia by 10% (OR 0.90, 95% CI 0.8-0.99, Models 2b and 3b) compared to spending break time in the teaching building. In addition, suburbanization was found to lower the personal risk of myopia by approximately 20% when adjusted for NDVI effects (OR 0.81, 95% CI 0.68-0.96 for Model 2b and OR 0.8, 95% CI 0.67-0.95 for Model 3b). However, a 3-year exposure time caused a higher myopia risk than a 1-year exposure time when considering the 500 or the 1000 m buffer NDVI (OR 1.38, 95% CI 1.18-1.63, Model 2b and 3b), but no such significant difference was seen between the 2- and 1-year exposures (OR 1.12, 95% CI 0.99-1.26 for Model 2b and OR 1.12, 95% CI 0.99-1.27 for Model 3b).

Table 2. Summary of the NDVI[a] effect at the personal level and the school level.

	Crude[b]		Adjusted[c]	
	OR[d] (95% CI)[e]	*P* value	OR (95% CI)[e]	*P* value
Personal myopia risk				
Within the school	0.84 (0.66-1.06)	.14	0.88 (0.75-1.04)	.15
500 m buffer around the school	0.74 (0.61-0.89)	.001	0.84 (0.73-0.97)	.02
1000 m buffer around the school	0.79 (0.68-0.92)	.002	0.88 (0.79-0.99)	.04
School myopia prevalence				
Within the school	0.85 (0.67-1.06)	.16	0.94 (0.79-1.11)	.44
500 m buffer around the school	0.73 (0.61-0.87)	<.001	0.85 (0.74-0.98)	.02
1000 m buffer around the school	0.79 (0.69-0.92)	.001	0.89 (0.80-1.01)	.05

[a] NDVI: normalized difference vegetation index.

[b] Odds ratios (ORs) of the NDVI (within the school, 500 m buffer around the school, 1000 m buffer around the school) regressed to the personal myopia risk or the school-level myopia prevalence without multivariable adjustment.

[c] Adjusted ORs with multivariable adjustment.

[d] OR: odds ratio.

[e] OR (95% CI) of the NDVI within the school and of the 500 or 1000 m buffers around the school indicated a 0.1 change in value.

Furthermore, the 500 and 1000 m buffer NDVIs performed differently in various subgroups (Figure 3). The 500 m or 1000 m buffer NDVI significantly influenced juniors (OR 0.82, 95% CI 0.72-0.94 for the 500 m buffer and OR 0.87, 95% CI 0.77-0.97 for the 1000 m buffer), schoolgirls (OR 0.82, 95% CI 0.72-0.93 for the 500 m buffer and OR 0.85, 95% CI 0.76-0.94 for the 1000 m buffer), and the the Han nationality (OR 0.84, 95% CI 0.72-0.97 for the 500 m buffer and OR 0.88, 95% CI 0.78-0.99 for the 1000 m buffer). However, only the 500 m buffer NDVI had a significant association with the suburban group (OR 0.81, 95% CI 0.7-0.94). Moreover, considering the exposure time, the NDVI of buffers around the school (OR 0.78, 95% CI 0.65-0.94 for the 500 m buffer and OR 0.81, 95% CI 0.70-0.94 for the 1000 m buffer) significantly influenced the 3-year exposure group, while only the NDVI of the 500 m buffer around the school (OR 0.84, 95% CI 0.71-0.99) significantly affect the 1-year exposure group.

Figure 3. Effects of the NDVI among subgroups. The effect of the NDVI of buffers around the school were adjusted by demographic factors, outdoor exercise, near-work time, body gestures, eye care, diet, sleep factors, and the position of trees. However, this figure only shows the adjusted effects of the 500 and 1000 m buffer NDVIs surrounding the school. The lines describe the 95% CI range and the points show the OR values, while the yellow color represents the NDVI of the 500 m buffer around the school and the deep-purple color represents the NDVI of the 1000 m buffer around the school. NDVI: normalized difference vegetation index; OR: odds ratio.

Subgroup	OR (95% CI)	*P* value
School type		
Junior		
500 m buffer	0.82 (0.72 to 0.94)	.01
1000 m buffer	0.87 (0.77 to 0.97)	.01
Senior		
500 m buffer	1.28 (0.81 to 2.02)	.29
1000 m buffer	1.16 (0.84 to 1.61)	.36
Gender		
Male		
500 m buffer	0.86 (0.73 to 1.02)	.08
1000 m buffer	0.92 (0.80 to 1.05)	.22
Female		
500 m buffer	0.82 (0.72 to 0.93)	.003
1000 m buffer	0.85 (0.76 to 0.94)	.002
Nationality		
Han nationality		
500 m buffer	0.84 (0.72 to 0.97)	.02
1000 m buffer	0.88 (0.78 to 0.99)	.04
Other nationalities		
500 m buffer	0.86 (0.70 to 1.05)	.14
1000 m buffer	0.89 (0.77 to 1.04)	.15
Exposure time		
1 year		
500 m buffer	0.84 (0.71 to 0.99)	.04
1000 m buffer	0.90 (0.79 to 1.04)	.15
2 years		
500 m buffer	0.88 (0.72 to 1.08)	.22
1000 m buffer	0.94 (0.80 to 1.11)	.44
3 years		
500 m buffer	0.78 (0.65 to 0.94)	.01
1000 m buffer	0.81 (0.70 to 0.94)	.004
Place of origin		
Urban area		
500 m buffer	1.04 (0.74 to 1.46)	.82
1000 m buffer	0.94 (0.72 to 1.23)	.66
Suburban area		
500 m buffer	0.81 (0.70 to 0.94)	.01
1000 m buffer	0.88 (0.78 to 1.01)	.06

Influence of the NDVI on School-Level Myopia Prevalence

The school-level myopia prevalence was not significantly affected by the NDVI within the school (OR 0.85, 95% CI 0.67-1.06, Model 4a); see Table 2, while it was distinctly influenced by the 500 and 1000 m buffer NDVIs without any adjusted factors (Table 2). With a 0.1 increase in the 500 and 1000 m buffer NDVIs, the prevalence of myopia at the school level dropped by 27% (OR 0.73, 95% CI 0.61-0.87, Model 5a) and 21% (OR 0.79, 95% CI 0.69-0.92, Model 6a), respectively. However, after adjusting for age, gender, the BMI, place of origin, and the relative position of trees (Table S2 in Multimedia Appendix 4), only the effects of the 500 m buffer NDVI were still influential (Table 2), with its 0.1 increase significantly reducing the school-level myopia prevalence by 15% (OR 0.85, 95% CI 0.74-0.98, Model 5b). Furthermore, in this situation, the suburban (OR 0.82, 95% CI 0.65-1.02) area no longer showed a significant adjusting effect, while age still had a positive association with the school-level myopia prevalence (OR 1.22, 95% CI 1.13-1.33); see Table S2 in Multimedia Appendix 4.

Discussion

Principal Findings

This research shows the alarming overall prevalence of myopia in senior and junior high school students in Beijing of 80.61%(10,785/13,380, 95% CI 79.93%-81.27%). However, the protective effects of the 500 m buffer NDVI were found on both the adolescent myopia risk (OR 0.84, 95% CI 0.73-0.97) and the school-level myopia prevalence (OR 0.85, 95% CI 0.74-0.98), and it was significantly associated with schoolgirls (OR 0.82, 95% CI 0.71-0.93), juniors (OR 0.82, 95% CI 0.72-0.94), the Han nationality (OR 0.84, 95% CI 0.72-0.97), 1-year exposure (OR 0.84, 95% CI 0.71-0.99), and 3-year exposure (OR 0.78, 95% CI 0.65-0.94) myopia risk.

The prevalence of myopia in Beijing is reported to exceed the average level of other Chinese mainland regions [43,44]. However, in the past decades, a series of actions have been taken by the government to prevent myopia among school-aged children and adolescents [45]. Some studies have noted that the myopia prevalence in junior high school students and senior students dropped from 2018 to 2020 according to cohort investigations. Correspondingly, the green space area ratios of Beijing from 2018 to 2020 were 48.44, 48.46, and 48.96, respectively (Beijing Municipal Forestry and Parks Bureau [46]). Despite the lack of direct evidence to determine the associations between the decrease in the prevalence of myopia and the yearly increasing green space ratio, the benefits of a greener city for children’s and adolescents’ vision development might be meaningful because this type of relationship has been demonstrated at the national or regional level by meta-analysis [47].

Our research found important associations between the green space surrounding schools and students’ myopia, in line with our hypothesis. Thus, more attention should be paid to myopia among students in schools with less greenness around the campus environment. The importance of green space within the 500 m buffer around schools for adolescents’ myopia was emphasized in our study. Similarly, the greenness of the 500 m buffer surrounding the school has been underlined for students’ social and health benefits in other studies [48,49]. We considered that students probably tended to live close to their school to minimize time costs or benefit from opportunities offered by the school (school catchment areas probably encompass the properties nearby) [50]. Thus, the 500 m buffer might also cover the students’ outside-school environments. This situation might provide day-long exposure to greener space and reduce the risk of myopia. Moreover, a greener environment could encourage more outside activities after school, encouraging students to enhance social connections with neighbors of the same age instead of playing alone on electronic devices [51]. In addition, with more green space outside schools, students may have greener routes between school and home, and walking home along an avenue with a tree canopy could be highly inspirational [52]. Indeed, “commuting greenness” has been shown to help reduce the need for spectacles [34]. Furthermore, the marked variation in the prevalence of myopia between suburban and urban schools did not occur when the effects of the NDVI of buffers were considered. It is suggested that the greenness level around schools rather than the urbanization of the location could better determine the school prevalence of myopia. The effects of green space around schools, especially the 500 m buffer, need to be further stressed.

In addition, the effect of green space within the 500 m buffer on myopia prevalence varied in different demographic subgroups. Schoolgirls had a higher myopia prevalence and were slightly more sensitive to the 500 m buffer NDVI than schoolboys. This result showed marginalization, which is usually reported in gender variation; that is, females are marginalized with poor health outcomes but gain more than males from green space exposure [53,54]. However, some studies note that nonenvironmental factors affect myopia in schoolgirls and that gender differences are the effect of puberty hormones, which influence schoolgirls’ eye development progress and increase the tendency to develop myopia [55]. Moreover, the NDVI of the 500 m buffer around the school showed significant effects on reducing personal myopia risk both in 1- and 3-year-exposure-time subgroups, but a 3-year exposure time caused a higher myopia risk than a 1-year exposure time. Although we lack evidence about the effects of exposure time combined with the effects of green space on myopia, it could suggest that students are more likely to benefit from a school environment with adequate green surroundings. In addition, the benefit level could rapidly rise initially when they first enter, and then drop to a certain and steady level afterward. The hypothesis about this trend is also mentioned in other studies for the effects of green space on other health outcomes [56].

Our study also showed important results for other influencing factors that were similar to previous studies. Time spent outdoors has been shown to protect against myopia [14,57], and previous studies have revealed the driving mechanisms. It appears that light quality is a crucial factor in avoiding myopia development, and people with myopia have significantly lower daily light exposure [58,59]. Because the intensity of artificial light in indoor environments is much lower than that in outdoor environments [60], the shade from trees placed close to teaching buildings might make the inner classroom darker, with poor distribution of light, which may aggravate students’ refractive error problems. However, some scholars believe that prolonged exposure to a middle-level-light environment could achieve the same protection against myopia as exposure to a strong-light environment [61]. It is conceivable that the total light exposure accumulated via outdoor activities in green space with medium light intensity could effectively prevent and control myopia development, while providing a visually comfortable environment that is not harmful to the skin [62]. In addition, unlike daytime outdoor light exposure, some researchers suggest that excessive artificial light at night (ALAN) exposure results in myopia [63]. Green space has been identified as having a negative correlation with ALAN [64] and may be a potential protective factor for myopia that mitigates harmful night-light doses for children and adolescents. Furthermore, previous studies have investigated seasonal variation and myopia development [65] and have found that myopia progression is slower in the summer than in the winter. They explain this by the association between outdoor light exposure and myopia. However, these studies may have partly neglected the role of seasonal plant landscape changes with more greenery and its function in promoting more physical activity in the summer than in the winter [66]. Thus, outdoor activities in greener spaces might benefit eye development and protect against ametropia.

Importantly, this study highlighted the same beneficial factors related to good eye care, such as reduction in near-work time and better body posture, as other researchers [1]. These should be key points for myopia prevention and control in Beijing. Furthermore, some studies have summarized the control activities that should be included in government policies, school-based prevention, health sector screening and treatment, and family support for prevention [45]. In view of our findings, since adolescents lack outdoor time [67], we recommend that adequate outdoor time in green space for adolescents should be ensured by policy makers, educators, and parents. In addition, the greenness around schools (particularly within the 500 m buffer) should be improved by tree planting and canopy provision, and the government should support these measures. A higher level of greenness for schools is required because campus green space provides further health benefits (eg, improving mental well-being [28,29,68], reducing heat-related illness [69]) beyond myopia reduction. However, the vegetation distribution should be carefully considered. Creating a vision-friendly campus could include a tree cover around the playground or along inner paths, while leaving appropriate spaces between trees and building edges.

Limitations

However, some limitations exist. First, we used cross-sectional myopia data from 2021 as many online classes were conducted during home quarantine due to the COVID-19 pandemic in 2020, which caused long periods of near-work on electronic devices and less outdoor activity time, as well as an increased prevalence of myopia [70-75]. Thus, the myopia prevalence might be overvalued with fewer environmental influences. Second, we lacked information about the green space exposure level of the samples' previous schools (ie, the NDVI and the exposure duration of their previous schools). Therefore, the exposure time in our study might be estimated inaccurately. In addition, we did not include the myopia duration due to the lack of access to students' previous myopia data. Third, noncycloplegic refractometry may overestimate the prevalence of myopia in teenagers with active accommodation [76]. Fourth, we did not consider other types of green space exposure, such as the accessibility and visibility of greenness.

The NDVI (2D perspective) does not precisely represent human vision greenness (3D perspective) [48], although it is an effective indicator to measure the green space level and can explain its effects on general health outcomes. Since a lack of green view is significantly related to eye fatigue, a greenness vision indicator should be considered. Future studies should address these points.

Conclusion

To the best of our knowledge, this is 1 of the few studies to clarify the effects of school green space with buffers and other relevant factors on junior and senior high school students' myopia prevalence. We emphasize that green space around schools is an independent protective factor for adolescents' myopia, and we suggest the importance of the appropriate distribution of within-campus trees for myopia prevention.

Acknowledgments

We thank anonymous reviewers for comments and suggestions that greatly improved the manuscript. We thank all the on-site investigators of the Beijing Municipal Health Commission for the data collection.

This study was funded by the National Natural Science Foundation of China (82004210), the National Key R&D Program of China (2018YFC1704400), the European Union's Horizon 2020 research and innovation program (grant agreement 821242), and the Central Nonprofit Research Institutions of the Chinese Academy of Forestry (CAFYBB2020ZB008). No payment was received by any of the coauthors for the preparation of this paper.

Authors' Contributions

Conceptualization was performed by CW and XG; data curation by CZ and LT; funding acquisition by CW and LT; investigation by XG and HX; methodology by CZ and LT; project administration by CW and XG; validation by LT; visualization and writing—original draft by CZ; and writing—review and editing by ZQ and LT. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

NDVI values of study plots. NDVI: normalized difference vegetation index.
[PDF File (Adobe PDF File), 31009 KB - publichealth_v9i1e42694_app1.pdf]

Multimedia Appendix 2
Summary of influencing factors.
[PDF File (Adobe PDF File), 341 KB - publichealth_v9i1e42694_app2.pdf]

Multimedia Appendix 3
Random effects and fixed effects of adjusted NDVI models. NDVI: normalized difference vegetation index.
[PDF File (Adobe PDF File), 12424 KB - publichealth_v9i1e42694_app3.pdf]

Multimedia Appendix 4
Summary of regression resluts.
[PDF File (Adobe PDF File), 406 KB - publichealth_v9i1e42694_app4.pdf]

References

1. Baird PN, Saw S, Lanca C, Guggenheim JA, Smith Iii EL, Zhou X, et al. Myopia. Nat Rev Dis Primers 2020 Dec 17;6(1):99. [doi: 10.1038/s41572-020-00231-4] [Medline: 33328468]
2. Morgan IG, Ohno-Matsui K, Saw S. Myopia. Lancet 2012 May;379(9827):1739-1748. [doi: 10.1016/s0140-6736(12)60272-4]
3. Bullimore MA, Ritchey ER, Shah S, Leveziel N, Bourne RRA, Flitcroft DI. The risks and benefits of myopia control. Ophthalmology 2021 Nov 04;128(11):1561-1579 [FREE Full text] [doi: 10.1016/j.ophtha.2021.04.032] [Medline: 33961969]
4. Holden B, Sankaridurg P, Smith E, Aller T, Jong M, He M. Myopia, an underrated global challenge to vision: where the current data takes us on myopia control. Eye (Lond) 2014 Feb;28(2):142-146 [FREE Full text] [doi: 10.1038/eye.2013.256] [Medline: 24357836]
5. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology 2016 May;123(5):1036-1042 [FREE Full text] [doi: 10.1016/j.ophtha.2016.01.006] [Medline: 26875007]
6. Goldschmidt E, Jacobsen N. Genetic and environmental effects on myopia development and progression. Eye (Lond) 2014 Feb;28(2):126-133 [FREE Full text] [doi: 10.1038/eye.2013.254] [Medline: 24357837]
7. Harb EN, Wildsoet CF. Origins of refractive errors: environmental and genetic factors. Annu Rev Vis Sci 2019 Sep 15;5(1):47-72. [doi: 10.1146/annurev-vision-091718-015027] [Medline: 31525141]
8. Morgan IG, French AN, Ashby RS, Guo X, Ding X, He M, et al. The epidemics of myopia: aetiology and prevention. Prog Retin Eye Res 2018 Jan;62:134-149. [doi: 10.1016/j.preteyeres.2017.09.004] [Medline: 28951126]
9. Rudnicka AR, Kapetanakis V, Wathern AK, Logan NS, Gilmartin B, Whincup PH, et al. Global variations and time trends in the prevalence of childhood myopia: a systematic review and meta-analysis. Lancet 2015 Nov;386:S69. [doi: 10.1016/s0140-6736(15)00907-1]
10. Flitcroft DI, Harb EN, Wildsoet CF. The spatial frequency content of urban and indoor environments as a potential risk factor for myopia development. Invest Ophthalmol Vis Sci 2020 Sep 01;61(11):42 [FREE Full text] [doi: 10.1167/iovs.61.11.42] [Medline: 32986814]
11. Pan C, Wu R, Li J, Zhong H. Low prevalence of myopia among school children in rural China. BMC Ophthalmol 2018 Jun 11;18(1):140 [FREE Full text] [doi: 10.1186/s12886-018-0808-0] [Medline: 29890943]
12. Ramamurthy D, Lin Chua SY, Saw S. A review of environmental risk factors for myopia during early life, childhood and adolescence. Clin Exp Optom 2015 Nov;98(6):497-506 [FREE Full text] [doi: 10.1111/cxo.12346] [Medline: 26497977]
13. Wu P, Chen C, Chang L, Niu Y, Chen M, Liao L, et al. Increased time outdoors is followed by reversal of the long-term trend to reduced visual acuity in Taiwan primary school students. Ophthalmology 2020 Nov;127(11):1462-1469 [FREE Full text] [doi: 10.1016/j.ophtha.2020.01.054] [Medline: 32197911]
14. Xiong S, Sankaridurg P, Naduvilath T, Zang J, Zou H, Zhu J, et al. Time spent in outdoor activities in relation to myopia prevention and control: a meta-analysis and systematic review. Acta Ophthalmol 2017 Sep;95(6):551-566 [FREE Full text] [doi: 10.1111/aos.13403] [Medline: 28251836]
15. Low W, Dirani M, Gazzard G, Chan Y, Zhou H, Selvaraj P, et al. Family history, near work, outdoor activity, and myopia in Singapore Chinese preschool children. Br J Ophthalmol 2010 Aug;94(8):1012-1016 [FREE Full text] [doi: 10.1136/bjo.2009.173187] [Medline: 20472747]
16. Giloyan A, Harutyunyan T, Petrosyan V. Risk factors for developing myopia among schoolchildren in Yerevan and Gegharkunik Province, Armenia. Ophthalmic Epidemiol 2017 Apr;24(2):97-103. [doi: 10.1080/09286586.2016.1257028] [Medline: 28032802]
17. Foreman J, Salim AT, Praveen A, Fonseka D, Ting DSW, Guang He M, et al. Association between digital smart device use and myopia: a systematic review and meta-analysis. Lancet Digital Health 2021 Dec;3(12):e806-e818. [doi: 10.1016/s2589-7500(21)00135-7]

18. Huang H, Chang DS, Wu P. The association between near work activities and myopia in children: a systematic review and meta-analysis. PLoS One 2015;10(10):e0140419 [FREE Full text] [doi: 10.1371/journal.pone.0140419] [Medline: 26485393]
19. Chakraborty R, Micic G, Thorley L, Nissen TR, Lovato N, Collins MJ, et al. Myopia, or near-sightedness, is associated with delayed melatonin circadian timing and lower melatonin output in young adult humans. Sleep 2021 Mar 12;44(3):zsaa208. [doi: 10.1093/sleep/zsaa208] [Medline: 33030546]
20. Jee D, Morgan IG, Kim EC. Inverse relationship between sleep duration and myopia. Acta Ophthalmol 2016 May;94(3):e204-e210 [FREE Full text] [doi: 10.1111/aos.12776] [Medline: 26031352]
21. Wang S, Wang Z, Zhang Z, Wu P. Analysis of influencing factors of myopia among children and adolescents in Yanqing District, Beijing in 2018. Capital J Public Health 2021;15(3):143-146. [doi: 10.16760/j.cnki.sdggws.2021.03.008]
22. Pan CW, Dirani M, Cheng CY, Wong TY, Saw SM. The age-specific prevalence of myopia in Asia: a meta-analysis. OVS 2015;92(3):258-266. [doi: 10.1097/opx.0000000000000516]
23. Xu R, Jan C, Song Y, Dong Y, Hu P, Ma J, et al. The association between menarche and myopia and its interaction with related risk behaviors among Chinese school-aged girls: a nationwide cross-sectional study. J Dev Orig Health Dis 2020 Dec;11(6):573-579. [doi: 10.1017/S204017442000077X] [Medline: 32799955]
24. Amoly E, Dadvand P, Forns J, López-Vicente M, Basagaña X, Julvez J, et al. Green and blue spaces and behavioral development in Barcelona schoolchildren: the BREATHE project. Environ Health Perspect 2014 Dec;122(12):1351-1358 [FREE Full text] [doi: 10.1289/ehp.1408215] [Medline: 25204008]
25. Wang M, Qiu M, Chen M, Zhang Y, Zhang S, Wang L. How does urban green space feature influence physical activity diversity in high-density built environment? An on-site observational study. Urban For Urban Green 2021 Jul;62:127129. [doi: 10.1016/j.ufug.2021.127129]
26. Astell-Burt T, Feng X. Does sleep grow on trees? A longitudinal study to investigate potential prevention of insufficient sleep with different types of urban green space. SSM Popul Health 2020 Apr;10:100497 [FREE Full text] [doi: 10.1016/j.ssmph.2019.100497] [Medline: 32025565]
27. Shin JC, Parab KV, An R, Grigsby-Toussaint DS. Greenspace exposure and sleep: a systematic review. Environ Res 2020 Mar;182:109081. [doi: 10.1016/j.envres.2019.109081] [Medline: 31891829]
28. Almeida DQ, Barros H, Ribeiro AI. Residential and school green and blue spaces and intelligence in children: the Generation XXI birth cohort. Sci Total Environ 2022 Mar 20;813:151859. [doi: 10.1016/j.scitotenv.2021.151859] [Medline: 34826494]
29. Bijnens EM, Vos S, Verheyen VV, Bruckers L, Covaci A, De Henauw S, et al. Higher surrounding green space is associated with better attention in Flemish adolescents. Environ Int 2022 Jan 15;159:107016 [FREE Full text] [doi: 10.1016/j.envint.2021.107016] [Medline: 34896669]
30. Liu Q, Luo S, Shen Y, Zhu Z, Yao X, Li Q, et al. Relationships between students' demographic characteristics, perceived naturalness and patterns of use associated with campus green space, and self-rated restoration and health. Urban For Urban Green 2022 Feb;68:127474. [doi: 10.1016/j.ufug.2022.127474]
31. Yang B, Li S, Zou Z, Markevych I, Heinrich J, Bloom MS, et al. Greenness surrounding schools and visual impairment in Chinese children and adolescents. Environ Health Perspect 2021 Oct;129(10):107006. [doi: 10.1289/ehp8429]
32. Yang B, Guo Y, Zou Z, Gui Z, Bao W, Hu L, et al. Exposure to ambient air pollution and visual impairment in children: a nationwide cross-sectional study in China. J Hazard Mater 2021 Apr 05;407:124750. [doi: 10.1016/j.jhazmat.2020.124750] [Medline: 33341569]
33. Yang Y, Chen W, Xu A, Zhao L, Ding X, Li J, et al. Spatial technology assessment of green space exposure and myopia. Ophthalmology 2022 Jan;129(1):113-117 [FREE Full text] [doi: 10.1016/j.ophtha.2021.07.031] [Medline: 34352303]
34. Dadvand P, Sunyer J, Alvarez-Pedrerol M, Dalmau-Bueno A, Esnaola M, Gascon M, et al. Green spaces and spectacles use in schoolchildren in Barcelona. Environ Res 2017 Jan;152:256-262. [doi: 10.1016/j.envres.2016.10.026] [Medline: 27816006]
35. Wu LJ, You QS, Duan JL, Luo YX, Liu LJ, Li X, et al. Prevalence and associated factors of myopia in high-school students in Beijing. PLoS One 2015;10(3):e0120764 [FREE Full text] [doi: 10.1371/journal.pone.0120764] [Medline: 25803875]
36. Chen L, Wu Y, Shi R, Lu R, Zhou M, Xin Z. Analysis of myopia screening results in children and adolescents in Fengtai District, Beijing from 2018 to 2020. Chin J Health Educ 2021;37(6):507-511. [doi: 10.16168/j.cnki.issn.1002-9982.2021.06.006]
37. Guo Y, Duan JL, Liu LJ, Sun Y, Tang P, Lv YY, et al. High myopia in Greater Beijing School Children in 2016. PLoS One 2017;12(11):e0187396 [FREE Full text] [doi: 10.1371/journal.pone.0187396] [Medline: 29121045]
38. China Center for Resources Satellite Data and Application. About CRESDA. URL: https://www.cresda.com/zgzywxyyzxeng/20220726141812992629/sid/list/sid_1.html [accessed 2022-12-20]
39. The People's Government of Beijing Municipality. URL: http://www.beijing.gov.cn [accessed 2022-12-20]
40. Lüdecke D. sjPlot: Data Visualization for Statistics in Social Science. URL: https://cran.r-project.org/web/packages/sjPlot/index.html [accessed 2022-12-20]
41. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
42. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag; 2016.

43. Dong L, Kang YK, Li Y, Wei WB, Jonas JB. Prevalence and time trends of myopia in children and adolescents in China: a systemic review and meta-analysis. Retina (Pa) 2020 Mar;40(3):399-411. [doi: 10.1097/IAE.0000000000002590] [Medline: 31259808]
44. You QS, Wu LJ, Duan JL, Luo YX, Liu LJ, Li X, et al. Prevalence of myopia in school children in greater Beijing: the Beijing Childhood Eye Study. Acta Ophthalmol 2014 Aug;92(5):e398-e406 [FREE Full text] [doi: 10.1111/aos.12299] [Medline: 25165786]
45. Jan CL, Congdon N. Chinese national policy initiative for the management of childhood myopia. Lancet Child Adolesc Health 2018 Dec;2(12):845-846. [doi: 10.1016/s2352-4642(18)30318-3]
46. Beijing Municipal Forestry and Parks Bureau. Statistics. URL: http://yllhj.beijing.gov.cn/zwgk/tjxx/ [accessed 2022-12-20]
47. Peng BA, Naduvilath TJ, Flitcroft DI, Jong M. Is myopia prevalence related to outdoor green space? Ophthalmic Physiol Opt 2021 Nov 05;41(6):1371-1381. [doi: 10.1111/opo.12896] [Medline: 34609002]
48. Labib SM, Lindley S, Huck JJ. Spatial dimensions of the influence of urban green-blue spaces on human health: a systematic review. Environ Res 2020 Jan;180:108869. [doi: 10.1016/j.envres.2019.108869] [Medline: 31722804]
49. Tuen Veronica Leung W, Yee Tiffany Tam T, Pan W, Wu C, Candice Lung S, Spengler JD. How is environmental greenness related to students' academic performance in English and mathematics? Landsc Urban Plan 2019 Jan;181:118-124. [doi: 10.1016/j.landurbplan.2018.09.021]
50. Han J, Cui L, Yu H. Pricing the value of the chance to gain admission to an elite senior high school in Beijing: the effect of the LDHSE policy on resale housing prices. Cities 2021 Aug;115:103238. [doi: 10.1016/j.cities.2021.103238]
51. Putra IGNE, Astell-Burt T, Cliff DP, Vella SA, Feng X. Association between caregiver perceived green space quality and the development of prosocial behaviour from childhood to adolescence: Latent class trajectory and multilevel longitudinal analyses of Australian children over 10 years. J Environ Psychol 2021 Apr;74:101579. [doi: 10.1016/j.jenvp.2021.101579]
52. Marquet O, Hirsch JA, Kerr J, Jankowska MM, Mitchell J, Hart JE, et al. GPS-based activity space exposure to greenness and walkability is associated with increased accelerometer-based physical activity. Environ Int 2022 Jul;165:107317 [FREE Full text] [doi: 10.1016/j.envint.2022.107317] [Medline: 35660954]
53. Sillman D, Rigolon A, Browning MHEM, Yoon HV, McAnirlin O. Do sex and gender modify the association between green space and physical health? A systematic review. Environ Res 2022 Jun;209:112869. [doi: 10.1016/j.envres.2022.112869] [Medline: 35123971]
54. Abdullah AYM, Law J, Perlman CM, Butt ZA. Age- and sex-specific association between vegetation cover and mental health disorders: Bayesian spatial study. JMIR Public Health Surveill 2022 Jul 28;8(7):e34782 [FREE Full text] [doi: 10.2196/34782] [Medline: 35900816]
55. Zoe TC, I-Jong W, Ya-Tang L, Yung-Feng S, Lin L. Polymorphisms in steroidogenesis genes, sex steroid levels, and high myopia in the Taiwanese population. Mol Vis 2011;17:2297-2310.
56. Alcock I, White MP, Wheeler BW, Fleming LE, Depledge MH. Longitudinal effects on mental health of moving to greener and less green urban areas. Environ Sci Technol 2014 Jan 21;48(2):1247-1255 [FREE Full text] [doi: 10.1021/es403688w] [Medline: 24320055]
57. He M, Xiang F, Zeng Y, Mai J, Chen Q, Zhang J, et al. Effect of time spent outdoors at school on the development of myopia among children in China: a randomized clinical trial. JAMA 2015 Sep 15;314(11):1142-1148. [doi: 10.1001/jama.2015.10803] [Medline: 26372583]
58. Landis EG, Yang V, Brown DM, Pardue MT, Read SA. Dim light exposure and myopia in children. Invest Ophthalmol Vis Sci 2018 Oct 01;59(12):4804-4811 [FREE Full text] [doi: 10.1167/iovs.18-24415] [Medline: 30347074]
59. Wen L, Cao Y, Cheng Q, Li X, Pan L, Li L, et al. Objectively measured near work, outdoor exposure and myopia in children. Br J Ophthalmol 2020 Nov 19;104(11):1542-1547 [FREE Full text] [doi: 10.1136/bjophthalmol-2019-315258] [Medline: 32075819]
60. Bhandary SK, Dhakal R, Sanghavi V, Verkicharla PK. Ambient light level varies with different locations and environmental conditions: potential to impact myopia. PLoS One 2021;16(7):e0254027 [FREE Full text] [doi: 10.1371/journal.pone.0254027] [Medline: 34234353]
61. Read SA, Collins MJ. Light exposure and physical activity in myopic and emmetropic children. Optom Vis Sci 2014;9(3):330-341. [doi: 10.1097/opx.0000000000000160]
62. Wu P, Chen C, Lin K, Sun C, Kuo C, Huang H, et al. Myopia prevention and outdoor light intensity in a school-based cluster randomized trial. Ophthalmology 2018 Aug;125(8):1239-1250. [doi: 10.1016/j.ophtha.2017.12.011] [Medline: 29371008]
63. Loman J, Quinn GE, Kamoun L, Ying G, Maguire MG, Hudesman D, et al. Darkness and near work. Ophthalmology 2002 May;109(5):1032-1038. [doi: 10.1016/s0161-6420(02)01012-6]
64. Stanhope J, Liddicoat C, Weinstein P. Outdoor artificial light at night: a forgotten factor in green space and health research. Environ Res 2021 Jun;197:111012. [doi: 10.1016/j.envres.2021.111012] [Medline: 33716029]
65. Gwiazda J, Deng L, Manny R, Norton TT, COMET Study Group. Seasonal variations in the progression of myopia in children enrolled in the correction of myopia evaluation trial. Invest Ophthalmol Vis Sci 2014 Feb 04;55(2):752-758 [FREE Full text] [doi: 10.1167/iovs.13-13029] [Medline: 24408976]

66. Zhou Y, Yang L, Yu J, Guo S. Do seasons matter? Exploring the dynamic link between blue-green space and mental restoration. Urban For Urban Green 2022 Jul;73:127612. [doi: 10.1016/j.ufug.2022.127612]
67. Baró F, Camacho DA, Pérez Del Pulgar C, Triguero-Mas M, Anguelovski I. School greening: right or privilege? Examining urban nature within and around primary schools through an equity lens. Landsc Urban Plan 2021 Apr;208:104019. [doi: 10.1016/j.landurbplan.2020.104019]
68. Chawla L, Keena K, Pevec I, Stanley E. Green schoolyards as havens from stress and resources for resilience in childhood and adolescence. Health Place 2014 Jul;28:1-13. [doi: 10.1016/j.healthplace.2014.03.001] [Medline: 24691122]
69. Lanza K, Alcazar M, Hoelscher DM, Kohl HW. Effects of trees, gardens, and nature trails on heat index and child health: design and methods of the Green Schoolyards Project. BMC Public Health 2021 Jan 07;21(1):98 [FREE Full text] [doi: 10.1186/s12889-020-10128-2] [Medline: 33413276]
70. Moore SA, Faulkner G, Rhodes RE, Brussoni M, Chulak-Bozzer T, Ferguson LJ, et al. Impact of the COVID-19 virus outbreak on movement and play behaviours of Canadian children and youth: a national survey. Int J Behav Nutr Phys Act 2020 Jul 06;17(1):85 [FREE Full text] [doi: 10.1186/s12966-020-00987-8] [Medline: 32631350]
71. Guan H, Okely AD, Aguilar-Farias N, del Pozo Cruz B, Draper CE, El Hamdouchi A, et al. Promoting healthy movement behaviours among children during the COVID-19 pandemic. Lancet Child Adolesc Health 2020 Jun;4(6):416-418. [doi: 10.1016/s2352-4642(20)30131-0]
72. Wong CW, Tsai A, Jonas JB, Ohno-Matsui K, Chen J, Ang M, et al. Digital screen time during the COVID-19 pandemic: risk for a further myopia boom? Am J Ophthalmol 2021 Mar 29;223:333-337 [FREE Full text] [doi: 10.1016/j.ajo.2020.07.034] [Medline: 32738229]
73. Rolland B, Haesebaert F, Zante E, Benyamina A, Haesebaert J, Franck N. Global changes and factors of increase in caloric/salty food intake, screen use, and substance use during the early COVID-19 containment phase in the general population in France: survey study. JMIR Public Health Surveill 2020 Sep 18;6(3):e19630 [FREE Full text] [doi: 10.2196/19630] [Medline: 32589149]
74. Parker K, Uddin R, Ridgers ND, Brown H, Veitch J, Salmon J, et al. The use of digital platforms for adults' and adolescents' physical activity during the COVID-19 pandemic (our life at home): survey study. J Med Internet Res 2021 Feb 01;23(2):e23389 [FREE Full text] [doi: 10.2196/23389] [Medline: 33481759]
75. Zheng Y, Wang W, Zhong Y, Wu F, Zhu Z, Tham Y, et al. A peer-to-peer live-streaming intervention for children during COVID-19 homeschooling to promote physical activity and reduce anxiety and eye strain: cluster randomized controlled trial. J Med Internet Res 2021 Apr 30;23(4):e24316 [FREE Full text] [doi: 10.2196/24316] [Medline: 33882021]
76. Hashemi H, Khabazkhoob M, Asharlous A, Soroush S, Yekta A, Dadbin N, et al. Cycloplegic autorefraction versus subjective refraction: the Tehran Eye Study. Br J Ophthalmol 2016 Aug;100(8):1122-1127. [doi: 10.1136/bjophthalmol-2015-307871] [Medline: 26541436]

Abbreviations

ALAN: artificial light at night
DW: Durbin-Watson
NDVI: normalized difference vegetation index
OR: odds ratio
VIF: variance inflation factor

Edited by Y Khader; submitted 15.09.22; peer-reviewed by Z Huang, J Liu; comments to author 18.11.22; revised version received 26.11.22; accepted 28.11.22; published 09.01.23.

Please cite as:
Zhang C, Wang C, Guo X, Xu H, Qin Z, Tao L
Effects of Greenness on Myopia Risk and School-Level Myopia Prevalence Among High School–Aged Adolescents: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e42694
URL: https://publichealth.jmir.org/2023/1/e42694
doi:10.2196/42694
PMID:36622746

Original Paper

Risk of De Novo Hypertensive Disorders of Pregnancy After Exposure to PM1 and PM2.5 During the Period From Preconception to Delivery: Birth Cohort Study

Zhichao Yuan[1*], MPH; Hai-Jun Wang[1,2*], PhD; Qin Li[1], PhD; Tao Su[3], PhD; Jie Yang[3], MSc; Junjun Chen[4], BSc; Yuanzhou Peng[1], MSc; Shuang Zhou[1], BSc; Heling Bao[1], PhD; Shusheng Luo[1], BSc; Hui Wang[1], PhD; Jue Liu[5], PhD; Na Han[3], MSc; Yuming Guo[6], PhD; Yuelong Ji[1], PhD

[1]Department of Maternal and Child Health, School of Public Health, Peking University, Beijing, China
[2]National Health Commission Key Laboratory of Reproductive Health, Beijing, China
[3]Tongzhou Maternal and Child Health Care Hospital of Beijing, Beijing, China
[4]Department of Electrical and Computer Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States
[5]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[6]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
[*]these authors contributed equally

Corresponding Author:
Yuelong Ji, PhD
Department of Maternal and Child Health
School of Public Health
Peking University
38 Huayuan Road, Haidian District
Beijing, 100089
China
Phone: 86 13161989008
Email: yuelong.ji@pku.edu.cn

Abstract

Background: Particulate matter (PM) is detrimental to the respiratory and circulatory systems. However, no study has evaluated the lag effects of weekly exposure to fine PM during the period from preconception to delivery on the risk of hypertensive disorders of pregnancy (HDPs).

Objective: We set out to investigate the lag effect windows of PM on the risk of HDPs on a weekly scale.

Methods: Data from women with de novo HDPs and normotensive pregnant women who were part of the Peking University Retrospective Birth Cohort, based on the hospital information system of Tongzhou district, were obtained for this study. Meteorological data and data on exposure to fine PM were predicted by satellite remote sensing data based on maternal residential address. The de novo HDP group consisted of pregnant women who were diagnosed with gestational hypertension or preeclampsia. Fine PM was defined as $PM_{2.5}$ and PM_1. The gestational stage of participants was from preconception (starting 12 weeks before gestation) to delivery (before the 42nd gestational week). A distributed-lag nonlinear model (DLNM) was nested in a Cox regression model to evaluate the lag effects of weekly PM exposure on de novo HDP hazard by controlling the nonlinear relationship of exposure–reaction. Stratified analyses by employment status (employed or unemployed), education level (higher or lower), and parity (primiparity or multiparity) were performed.

Results: A total of 22,570 pregnant women (mean age 29.1 years) for whom data were available between 2013 and 2017 were included in this study. The prevalence of de novo HDPs was 6.7% (1520/22,570). Our findings showed that PM_1 and $PM_{2.5}$ were significantly associated with an elevated hazard of HDPs. Exposure to PM_1 during the 5th week before gestation to the 6th gestational week increased the hazard of HDPs. A significant lag effect of $PM_{2.5}$ was observed from the 1st week before gestation to the 6th gestational week. The strongest lag effects of PM_1 and $PM_{2.5}$ on de novo HDPs were observed at week 2 and week 6 (hazard ratio [HR] 1.024, 95% CI 1.007-1.042; HR 1.007, 95% CI 1.000-1.015, respectively, per 10 μg/m^3 increase). The stratified analyses indicated that pregnant women who were employed, had low education, and were primiparous were more vulnerable to PM exposure for de novo HDPs.

Conclusions: Exposure to PM_1 and $PM_{2.5}$ was associated with the risk of de novo HDPs. There were significant lag windows between the preconception period and the first trimester. Women who were employed, had low education, and were primiparous were more vulnerable to the effects of PM exposure; more attention should be paid to these groups for early prevention of de novo HDPs.

(JMIR Public Health Surveill 2023;9:e41442) doi:10.2196/41442

KEYWORDS

air pollution; PM2.5; PM1; hypertensive disorders of pregnancy; preconceptional period; lag effect; pregnancy; hypertension; hypertensive disorders; risk; pollutants; exposure; maternal health; perinatal health; pollution

Introduction

A risk of multiple adverse health outcomes is attributable to increased exposure to the ambient particulate matter (PM). Fine PM is defined as ambient particles with an aerodynamic diameter ≤2.5 µm or ≤1 µm, that is, $PM_{2.5}$ and PM_1, respectively. PM is derived from a wide range of sources, such as forest fires, coal combustion, industrial processes, and traffic emissions [1]. It is composed of various toxic pollutants, including organic carbon, alkanes, metals, sulfates, and nitrates [2,3]. When people inhale the air, fine PM can easily translocate the abovementioned toxic pollutants across biological membranes from the pulmonary alveoli to the blood circulation, further resulting in systemic and tissue injury due to its physical characteristics of tiny size and large surface area [4]. According to the World Health Organization (WHO) air quality guidelines, about 92% of people in the world are exposed to excessive fine PM. Not long ago, the WHO updated the limit on exposure to fine PM, decreasing it from 10 $\mu g/m^3$ to 5 $\mu g/m^3$, based on research findings from the past 15 years. PM exposure is regarded as a risk factor for cardiovascular diseases and adverse perinatal outcomes [5,6]. An accumulating number of studies have used in vitro and in vivo experiments to show that fine PM can induce cardiovascular toxicity [7]. A recent global disease-burden study also indicated that fine PM was associated with risks of low birth weight, preterm birth, and neonatal and infant mortality, particularly in low- and middle-income countries [8].

Emerging evidence indicates a link between PM exposure and hypertensive disorders of pregnancy (HDPs), which are among the leading risk factors for maternal and fetal morbidity and mortality. Gestational hypertension (GH) and preeclampsia (PE) belong to the de novo type of HDP, which accounts for over 85% of HDP diagnoses [9]. HDPs can not only become complicated with proteinuria, edema, seizure, and liver and kidney injury, but also result in a series of adverse pregnancy outcomes, such as intrauterine growth restriction, preterm birth, stillbirth placental abruption, and postpartum hemorrhage [10,11]. In addition, HDPs have grown to become the second leading cause of maternal death in China. Therefore, HDPs are a significant public health problem in China in terms of maternal health [12]. While the etiology of HDPs is still elusive, in recent years an increasing number of researchers have focused on the effects of the environment on cardiovascular health. They have hypothesized that fine PM could result in HDPs via vascular constriction, inflammation, and oxidative stress. However, findings for an association between fine PM and HDPs are inconsistent. A recent Chinese study showed that fine PM exposure during the first trimester was associated with the risk of HDPs [13]. On the other hand, an American study indicated that fine PM exposure increased the development of HDPs during the second trimester rather than the first trimester [14]. This inconsistency may originate from differing study populations and exposure concentrations. Moreover, it has become imperative to draw attention to fine PM exposure before pregnancy. Based on the Developmental Origins of Health and Disease paradigm, the window of sensitivity to fine PM exposure could begin as far back as the preconception period [15]. There is an increasing consensus that preconception exposure to environmental contaminants can affect well-being in pregnancy. Several studies have indicated that air pollution exposure during the preconception period is associated with a risk of termination of pregnancy [16] and gestational diabetes mellitus [17]. The majority of previous studies only focused on the association between air pollution and maternal health in either specific trimesters or the total pregnancy [18,19]. The dynamic association between fine PM exposure and HDPs at different gestational weeks remains to be investigated in Chinese populations. To our knowledge, no study has assessed the association between weekly fine PM exposure, starting during the preconception period, and HDPs, especially de novo HDPs. If an association is shown, preconception precautions for fine PM exposure will become essential practice for lowering the risk of HDPs.

In this retrospective birth cohort study, we collected ambient data on PM exposure and applied a distributed-lag nonlinear model (DLNM) to examine the association between weekly PM exposure during the period from preconception to delivery and de novo HDPs. Additionally, we evaluated sensitivity lag windows for fine PM for HDPs across different sociodemographic strata.

Methods

Study Population

The study participants were pregnant women who were part of the Peking University Retrospective Birth Cohort for whom data were available in the hospital information system of Tongzhou for the period between January 2013 and December 2017. All the participants received antenatal examinations and gave birth at the Tongzhou Maternal and Child Health Care

Hospital, which is the maternal health center for Tongzhou district, Beijing. The sociodemographic and health details of the participants were recorded at the antenatal examinations, including maternal age (<35 years vs ≥35 years), employment status (employed vs unemployed), parity (nulliparous vs multiparous), prepregnancy weight, height, maternal education level (high school or below, junior college, university or above), maternal ethnicity (Han vs non-Han), first day of last menstrual period, and delivery date. Prepregnancy BMI was calculated by dividing weight (in kilograms) by height (in meters squared) and categorized according to the WHO standards for Asian women: underweight (BMI <18.5 kg/m^2), normal weight (BMI 18.5-23 kg/m^2), overweight (BMI 23-27.5 kg/m^2), or obese (BMI >27.5 kg/m^2) [20]. The study data were extracted from the electronic medical records of Tongzhou Maternal and Child Health Care Hospital and were routinely reviewed by in-house professional data engineers every week.

The inclusion criteria of this study were (1) a single gestation, (2) an available residential address, (3) no history of previous HDPs, and (4) a delivery date before the 42nd gestational week. A total of 43,894 pregnant women met the inclusion criteria. We excluded 21,324 pregnant women who met the following exclusion criteria: (1) preexisting conditions, including metabolic syndrome; hemolysis, elevated liver enzymes, slow platelets; or hypertension before the HDP diagnosis; (2) missing data for delivery date; (3) an age at delivery less than 18 years; and (4) missing values for maternal age, ethnicity, education level, employment status, or prepregnancy BMI. Finally, 22,570 women were included in this study. The flowchart in Multimedia Appendix 1 shows the details of participant selection.

Outcome Measurements

The de novo HDP group was defined as patients who experienced GH or PE during the current pregnancy. The normotensive group was defined as those free of any diagnosis of GH or PE and without a history of hypertensive disorders. Information on the history of hypertensive disorders was obtained from initial antenatal examination records. GH or PE diagnoses were made by obstetricians in accord with the Chinese Clinical Practice Guidelines, which are consistent with the guidelines of developed countries [21]. International Classification of Diseases–10 (ICD-10) codes were used to define GH and PE (the ICD codes for GH and PE are described in our previous study [22]). Medical records containing information related to GH and PE were extracted for double confirmation of the diagnosis and to obtain the week of disease diagnosis.

Assessment of Ambient PM Exposure, Temperature, and Relative Humidity

The pregnant women's residential addresses were collected at the initial antenatal examination. Each residential address was transformed into longitude and latitude values to determine the concentration of $PM_{2.5}$ or PM_1 in the center of the nearest 1 km × 1 km area on a grid. Daily PM concentrations were calculated at a spatial resolution of 1 km × 1 km in Beijing during the study period using satellite remote sensing, meteorology, and land-use information. Details of the calculations and validation of PM daily concentration measurements have been reported previously [23]. Data for daily temperature and relative humidity during the study period were obtained from the National Oceanic and Atmospheric Administration [24]. Average daily values for PM, temperature, and relative humidity from preconception (12 before gestation) to delivery (42 gestational weeks) were used to assess atmospheric exposure in our analysis.

Ethical Considerations

The study was approved by the Institutional Review Board of Peking University Health Science Center (IRB00001052-21023). All the study participants provided written consent. The study data were anonymized to protect the privacy of the study participants.

Statistical Analyses

The Student *t* test (2-tailed) and chi-square test were applied to assess differences in the distribution of characteristics between the subjects with HDPs and those who were normotensive. The relationships among $PM_{2.5}$, PM_1, temperature and relative humidity were assessed with the Pearson correlation test.

We followed the weekly exposure to PM_1 and $PM_{2.5}$ from the 12th week before gestation to the 42nd gestational week [25]. Cox proportional hazard models with a DLNM were used to estimate the association between the weekly $PM_{2.5}$ and PM_1 exposures and the hazard of de novo HDPs. The basis of the Cox proportional hazard model is the proportional hazards assumption, which means that all the study participants had the same hazard function. Our study complied with this assumption, which means that the ratio of the hazards for any 2 individuals was constant over time [26]. A DLNM was adopted to assess the effects of weekly PM exposures on hazards for HDPs by the "cross basic" function (cb) by constructing a 2D cross matrix that included exposure doses and time to control for the lag effect of PM exposure and the nonlinear relationship of exposure–reaction. A natural spline with 3 degrees of freedom was used to further accommodate the nonlinear effects of PM exposure, temperature, and relative humidity. The maximum number of lag weeks was set at 54 (from the 12th week before gestation to the 42nd gestational week). The lag effect of PM exposure was evaluated by the model as follows: $\ln(h(t,X)/h0(t)) = \beta1*Zt + \beta2*ns(temperature,3) + \beta3*ns(relative\ humidity,3) + \beta4*covariables$. "$\ln(h(t,X)/h0(t))$" indicates the hazard of de novo HDPs at the specific lag week (t), and the filed event is de novo HDP, noted as X. Zt represented the cross matrix of PM at specific lag weeks. β2 and β3 are the regression coefficients for temperature and relative humidity with 3 degrees of freedom in the natural spline function. β4 is the coefficient for the covariables, including maternal age, employment status, prepregnancy BMI, maternal education level, parity, maternal ethnicity, and conceptional year.

We also performed several sensitivity analyses to test the consistency of the main results. First, we further used natural splines with 4 and 5 degrees of freedom in the cross matrix to assess the hazards of de novo HDPs. Second, considering the clinical differences between GH and PE, we repeated our analyses excluding GH patients to explore the independent association between PM and PE. Meanwhile, stratified analyses were conducted across strata for differences in employment status, parity, and maternal education level.

All statistical analyses were performed with R (version 4.0.0; R Foundation for Statistical Computing). A 2-sided *P* value <.05 was considered a significant difference.

Results

Characteristics of the Participants

There were 1520 participants with de novo HDPs and 21,050 normotensive participants in this study. Table 1 shows the characteristics of the participants. The mean maternal age of the pregnancies was 29.1 (SD 4.1) years. There were 3802/22,570 (16.8%) unemployed women, and 21,222/22,570 (94%) women of Chinese Han ethnicity. A higher percentage of de novo HDPs was observed in the follow sociodemographic strata: low education level, primiparity, and prepregnancy overweight or obesity.

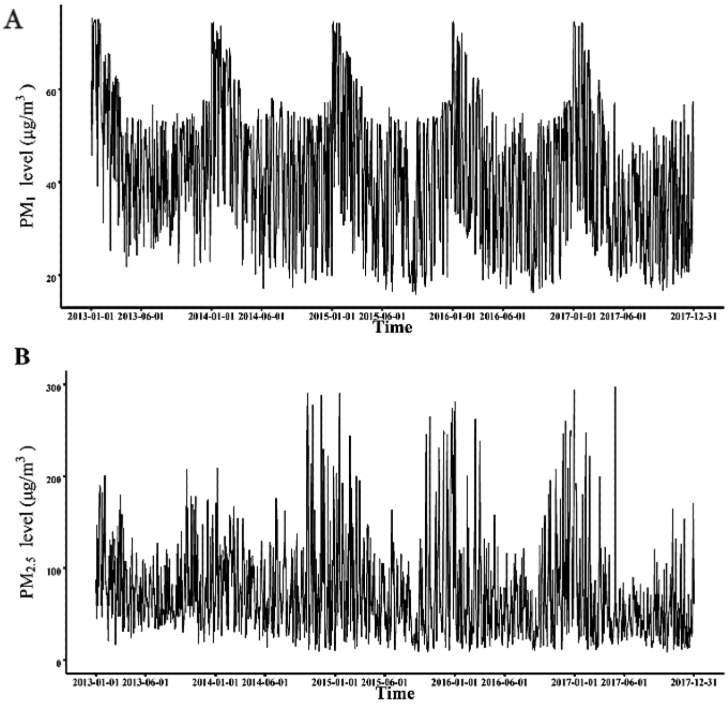
Table 1. Characteristics of mothers in the birth cohorts.

Characteristics	Total (n=22,570)	Normotensive (n=21,050)	Hypertensive disorders of pregnancy (n=1520)	*P* value
Maternal age (years), mean (SD)	29.1 (4.1)	29.1 (4.1)	29.4 (4.4)	.13
Maternal age groups (years), n (%)				<.001
Younger than 35	20,126 (89.2)	18,812 (89.4)	1314 (86.4)	
35 or older	2444 (10.8)	2238 (10.6)	206 (13.6)	
Maternal ethnicity, n (%)				.80
Han	21,222 (94)	19,790 (94)	1432 (94.2)	
Non-Han	1348 (6)	1260 (6)	88 (5.8)	
Parity, n (%)				<.001
Primiparous	15,320 (67.9)	14,222 (67.6)	1098 (72.2)	
Multiparous	7250 (32.1)	6828 (32.4)	422 (27.8)	
Maternal education level, n (%)				<.001
Low	6890 (30.5)	6357 (30.2)	533 (35.1)	
High	15,680 (69.5)	14,693 (68.8)	987 (64.9)	
Employment status, n (%)				.80
Employed	18,768 (83.2)	17,501 (83.1)	1267 (83.4)	
Unemployed	3802 (16.8)	3549 (16.9)	253 (16.6)	
Prepregnancy BMI, n (%)				<.001
Underweight	2372 (10.5)	2293 (10.9)	79 (5.2)	
Normal weight	12,032 (53.3)	11,455 (54.4)	577 (38)	
Overweight	6490 (28.8)	5901 (28)	589 (38.8)	
Obesity	1676 (7.4)	1401 (6.7)	275 (18.1)	

Characteristics of PM Exposure

The distribution of $PM_{2.5}$, PM_1, temperature, and relative humidity is summarized in Multimedia Appendix 2. The average concentration of $PM_{2.5}$ was 74.2 (SD 53.2) μg/m^3, ranging from 8.5 to 398 μg/m^3. The concentration of PM_1 ranged from 16.0 to 75.1 μg/m^3. Mean temperature and relative humidity were 13.5 °C (SD 11.0 °C) and 54.7% (SD 18.2%). Figure 1 presents the daily fluctuation in $PM_{2.5}$ and PM_1 concentration from January 2013 to December 2017, revealing seasonal fluctuations in $PM_{2.5}$ and PM_1 exposure. Meanwhile, $PM_{2.5}$ and PM_1 were strongly correlated (*r*=0.80). The Pearson correlations among $PM_{2.5}$, PM_1, temperature, and relative humidity are depicted in Multimedia Appendix 3.

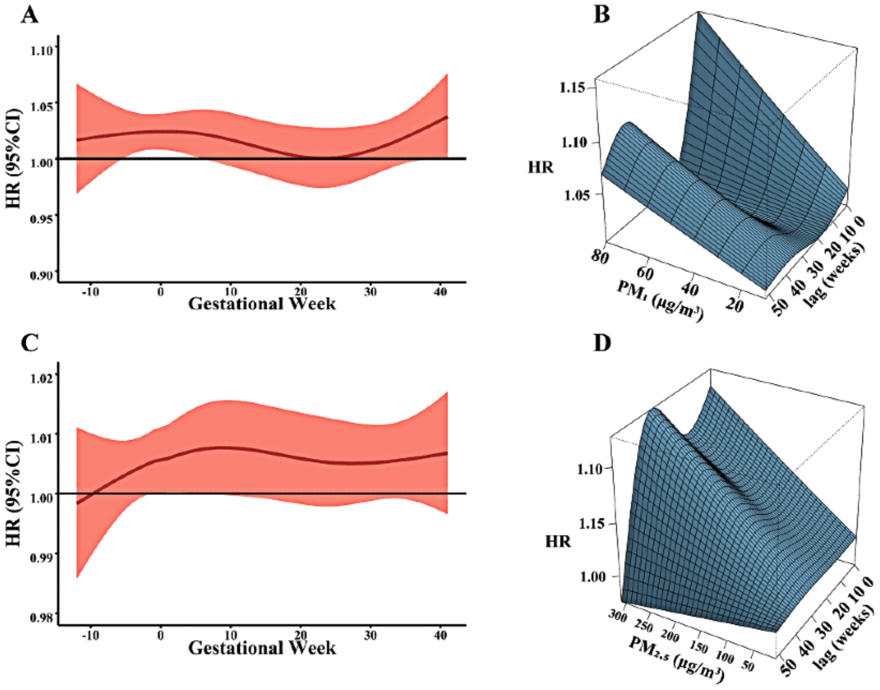
Figure 1. The level of particulate concentration by week. A and B show the daily level of PM_1 and $PM_{2.5}$, respectively, from 2013 to 2017. PM: particulate matter.



Association of Weekly Exposure to PM With HDPs

Figure 2 shows that a 10 $\mu g/m^3$ increase in PM_1 and $PM_{2.5}$ was associated with de novo HDPs in 2D and 3D plots. The hazard of de novo HDPs was associated with PM_1 exposure from the 5th week before gestation to the 6th gestational week, and the strongest effect of PM_1 exposure was observed in the 2nd gestational week (HR 1.024, 95% CI 1.007-1.042; Figure 2A). Figure 2B depicts the effects of PM_1 exposure on de novo HDPs with different exposure levels and different lag weeks, with a significant single peak by lag week from preconception to the first trimester (from lag 36 weeks to lag 47 weeks). For $PM_{2.5}$, the significant lag window was from the 1st week before gestation to the 6th gestational week and the maximum lag effect of $PM_{2.5}$ was found in the 6th gestational week (HR 1.007, 95% CI 1.000-1.015, Figure 2C). Similarly, we also observed a significant single peak from lag 36 weeks to lag 43 weeks when we assessed the association between $PM_{2.5}$ exposure and de novo HDPs for different exposure doses and lag weeks (Figure 2D).

Figure 2. The lag effect of PM on hypertensive disorders of pregnancy in 2D and 3D plots. A and B show the association between PM_1 and de novo hypertensive disorders of pregnancy in 2D and 3D plots, respectively. C and D show the association between $PM_{2.5}$ and de novo hypertensive disorders of pregnancy in 2D and 3D plots, respectively. HR: hazard ratio; PM: particulate matter.



Sensitivity Analyses

We conducted two sensitivity analyses to assess the consistency of the sensitivity windows for PM_1 and $PM_{2.5}$ for de novo HDP hazard. First, using natural splines with 4 and 5 degrees of freedom for the lag constraint, we observed similar results, that is, the significant lag windows for PM_1 and $PM_{2.5}$ spanned the preconception period and the first trimester (Multimedia Appendix 4). Furthermore, we excluded patients with GH to explore the independent association between PM and PE. Multimedia Appendix 5 shows a result consistent with our primary analyses, that is, that the lag windows for PM_1 and $PM_{2.5}$ spanned the preconception period and the first trimester.

Stratified Analyses

In the stratified analyses, we estimated the effects of PM on de novo HDPs across differences in employment status, parity, and maternal education level. In terms of PM_1, a significant sensitivity window for de novo HDPs was observed for employed, primiparous, and low educational status women (Figure 3). Similar findings are shown in Figure 4: women who were employed, primiparous, or had a low education level were more sensitive to $PM_{2.5}$ exposure across the preconception period and the first trimester.

Figure 3. The stratified associations between $PM_{2.5}$ and de novo hypertensive disorders of pregnancy for different maternal characteristics. A and B show the association in unemployed and employed women, respectively. C and D show the association in multiparous and primiparous women, respectively. E and F show the association in women with higher and lower education status, respectively. HR: hazard ratio; PM: particulate matter.

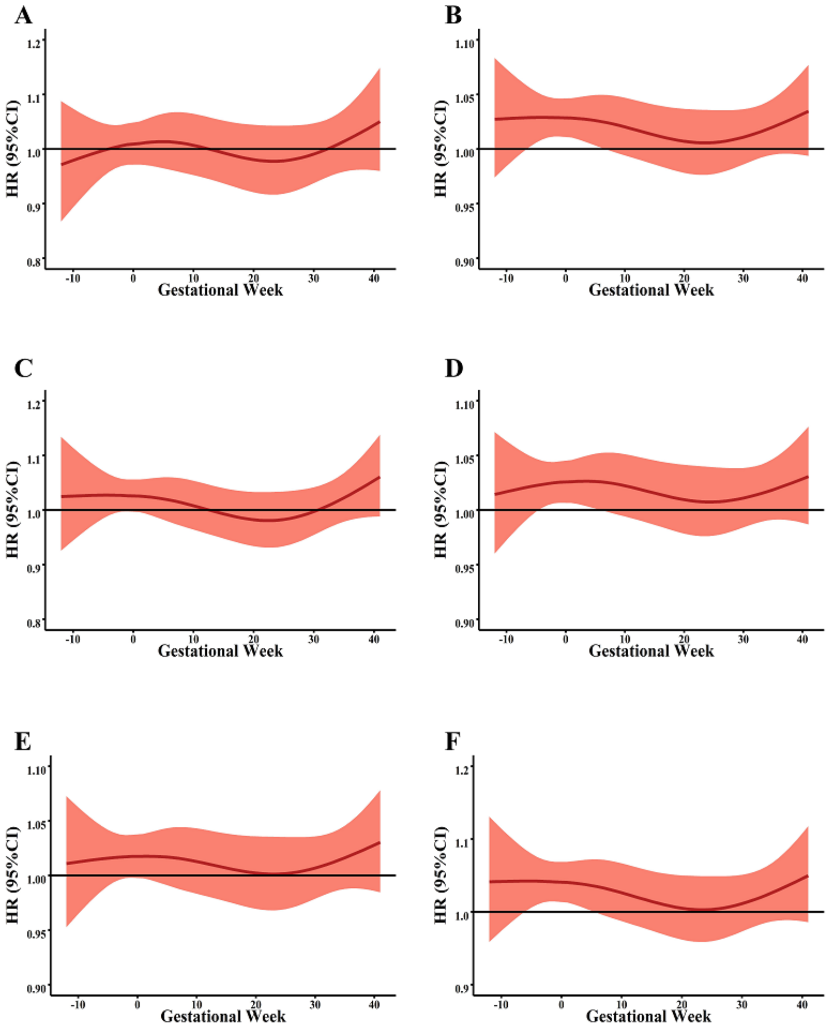
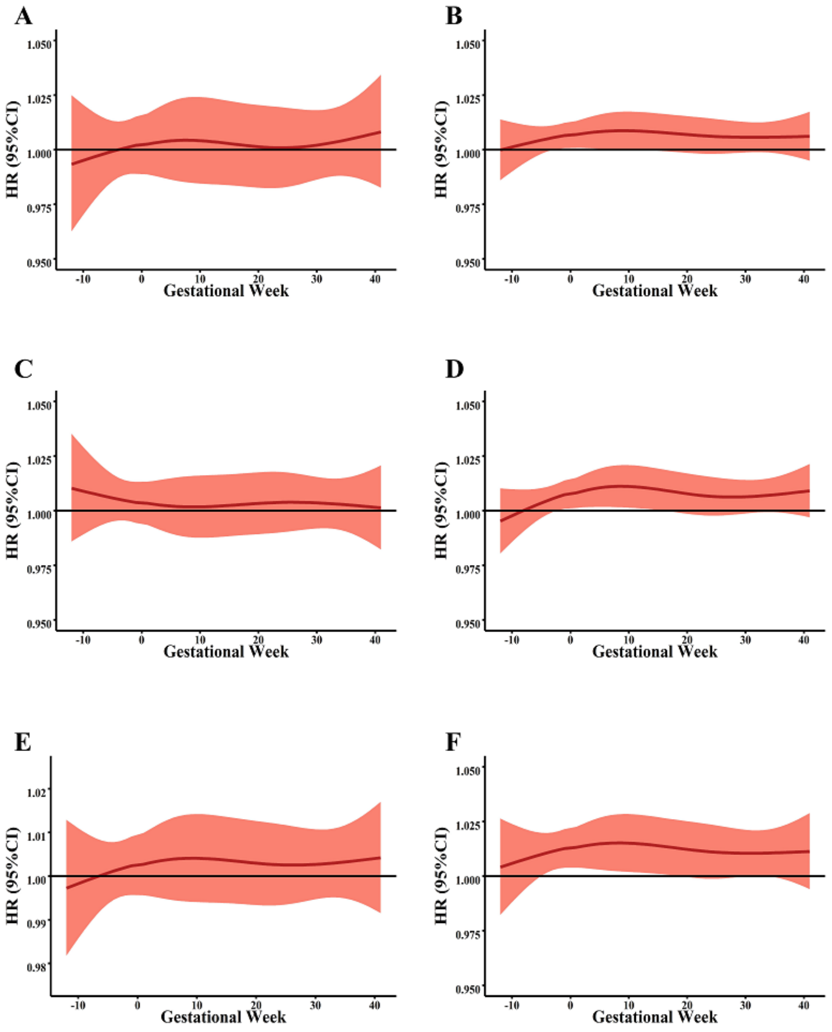


Figure 4. The stratified associations between PM_1 exposure and de novo hypertensive disorders of pregnancy for different maternal characteristics. A and B show the association in unemployed and employed women, respectively. C and D show the association in multiparous and primiparous women, respectively. E and F show the association in women with higher and lower education status, respectively. HR: hazard ratio; PM: particulate matter.



Discussion

Principal Results

This study found that the significant lag windows for PM_1 and $PM_{2.5}$ were from the 5th week before gestation to the 6th gestational week and the 1st week before gestation to the 6th gestational week, respectively, in terms of the de novo HDP hazard. The findings remained robust in sensitivity analyses, which reexamined the associations using different degrees of freedom and after excluding patients with GH. Stratified analyses indicated that pregnant women who were employed, primiparous, or had a lower education level were more sensitive to PM_1 and $PM_{2.5}$ exposure in the prepregnancy period and the first trimester.

In recent years, numerous studies have evaluated the association between air pollution and cardiovascular health during pregnancy. However, there has been no study to assess the adverse effect of preconception air pollution on gestational health in terms of de novo HDPs. We adopted a time-serial study design to explore the weekly effects of PM_1 and $PM_{2.5}$ from the preconception period to delivery. To our knowledge, this is the first study to focus on the detrimental association between PM during both the preconception and gestational periods and de novo HDPs. During the study period, the average exposure to $PM_{2.5}$ was 74.2 $\mu g/m^3$, which is more than 2-fold higher than the recommended maximum level of 35 $\mu g/m^3$ set by the National Ambient Air Quality Standards organization [27]. Moreover, exposure to PM showed seasonal fluctuations, peaking in winter and spring. Considering that our participants were mainly northern Chinese living in Beijing, the capital of China, the high concentration of PM can be attributed to central heating and a dense population.

In this study, the prevalence of de novo HDPs was 6.7% (1520/22,570). However, a lower prevalence (5%) of HDPs was reported in a recent study conducted in Shanghai [13]. In comparison with the Shanghai study, the pregnant subjects in our study were exposed to a higher concentration of PM pollution. Since both Beijing and Shanghai have populations of over 24 million and similar demographic compositions, the prevalence of HDPs is unlikely to have been influenced by differences in industrialization or urbanization between these two megacities. The difference in HDP prevalence might be partially explained by the higher PM exposure in the north due to the climate difference and use of central heating.

In our time-serial study, we observed that the women were sensitive to PM_1 and $PM_{2.5}$ during the preconception period and the first trimester. Our PM exposure measurements were at the week scale, which is more precise than previous studies that have used the trimester scale. Previous studies have made inconsistent findings on the association between fine PM and HDPs. Among 4 US studies, one in California [28] and one in Florida [14] observed that $PM_{2.5}$ exposure was associated with HDPs during pregnancy. However, studies conducted in Rhode Island [29] and New York [30] failed to observe significant associations. There are some potential explanations for these inconsistent results. First, the latitudinal difference may have influenced the temperature and humidity in the different regions. A recent national cohort study indicated that temperature was positively associated with the risk of HDPs during pregnancy [31]. Second, exposure measurement bias may have resulted in a false-negative association. Third, ethnicity differences can affect the association between PM and HDPs. A national cross-sectional study from 2007 to 2018 showed that there was a racial difference in the prevalence of HDPs in the United States [32]. Our study participants mainly lived in Beijing, and 94% were members of the Chinese Han population. Meanwhile, we described our method for high-precision air pollution measurement in a previous study [23]. Our results are consistent with those of the Shanghai study, which had a similar ethnicity and urbanization profile. Thus, climate, measurement, and ethnicity differences were well controlled in our study.

Furthermore, we used a DLNM model, which allowed an examination of the lag effect of PM on HDPs at the week scale, rather than the previously used trimester scale, to improve the precision of the search for the most sensitive gestational week. We also considered potential bias in the model parameters by testing the consistency of our results with different degrees of freedom. Based on the robust results from our sensitivity analyses, we consider that bias from model parameters was not likely. Additionally, our second sensitivity analysis excluding the patients with GE further showed that exposure to fine PM during the preconception period and first trimester was independently associated with the risk of PE. PE is one of the leading causes of adverse pregnancy outcomes, accounting for 10% to 15% of maternal deaths and 15% to 20% of preterm births [33,34]. Early identification of PE risk in clinical practice remains a difficult problem [35]. Our environmental evidence for an association between fine PM exposure and PE might indicate the need for preconception PE prevention to consider environmental hazards.

Considering the strong link with maternal characteristics and gestational blood pressure [36], we further investigated the effects of PM_1 and $PM_{2.5}$ on HDPs across different maternal characteristics, including employment status, parity, and education level. Among employed women, exposure to PM was significantly associated with the risk of HDPs. PM exposure also elevated the risk of HDPs in women who were primiparous and had a low education level, which indicates that women with these maternal characteristics have a higher risk of HDPs with exposure to PM.

Strengths and Limitations

There are several strengths of this study. First, to our knowledge, this was the first to adopt the Cox proportion model with a DLNM to estimate the week-scale sensitivity window to PM for HDPs in a Chinese population. HDPs are a major cause of adverse maternal and prenatal outcomes [37]. However, the identification of risk factors for HDPs has been limited due to complex confounders. Previous studies primarily explored the association between PM and HDPs in a specific trimester [38,39]. Our study analyzed the specific gestational weeks and improved the precision of the sensitivity window. Second, this was the first study to cover the effect of air pollution, starting from the preconception period, on de novo HDPs. Our findings reconfirmed the necessity of focusing on air pollution before pregnancy in terms of de novo HDPs. Third, the large sample size of this study made it possible to determine the effects of PM on subtypes of HDP in stratified analyses.

Some limitations must be noted. First, all the participants were residents of Beijing. The lag windows for PM_1 and $PM_{2.5}$ should be tested in different regions to increase generalizability. Second, PM is a mixture of types of air pollution. The effects of specific PM components should be further explored in future studies.

Conclusions

Exposure to PM_1 and $PM_{2.5}$ was associated with the risk of de novo HDPs. The significant lag windows for PM exposure were between the preconception period and the first trimester. Women who were employed, had a lower education level, or were primiparous were more vulnerable to the effects of PM exposure on the risk of HDPs; more attention should be paid to these groups for the early prevention of de novo HDPs.

Acknowledgments

We sincerely thank the research group of the Peking University Retrospective Birth Cohort, in Tongzhou, for their help with the hospital information system. We appreciate the health professionals at Tongzhou Maternal and Child Health Care Hospital of Beijing for the data collection and management. This study was sponsored by the Beijing Natural Science Foundation (7212144) and National Natural Science Foundation of China (92046019 and 82204055).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study flow chart.
[PNG File , 86 KB - publichealth_v9i1e41442_app1.png]

Multimedia Appendix 2
Distribution of $PM_{2.5}$, PM_1, temperature and relative humidity.
[DOCX File , 17 KB - publichealth_v9i1e41442_app2.docx]

Multimedia Appendix 3
The correlation between ambient exposure.
[PNG File , 74 KB - publichealth_v9i1e41442_app3.png]

Multimedia Appendix 4
The association between PM and de novo HDP with 4 and 5 degrees of freedom for the lag constraint. HDP: hypertensive disorders of pregnancy; PM: particulate matter.
[PNG File , 162 KB - publichealth_v9i1e41442_app4.png]

Multimedia Appendix 5
The association between PM_1 and $PM_{2.5}$, and preeclampsia.
[PNG File , 83 KB - publichealth_v9i1e41442_app5.png]

References

1. Miller M, Newby DE. Air pollution and cardiovascular disease: car sick. Cardiovasc Res 2020 Feb 01;116(2):279-294. [doi: 10.1093/cvr/cvz228] [Medline: 31583404]
2. Tian Y, Liu X, Huo R, Shi Z, Sun Y, Feng Y, et al. Organic compound source profiles of PM2.5 from traffic emissions, coal combustion, industrial processes and dust. Chemosphere 2021 Sep;278:130429. [doi: 10.1016/j.chemosphere.2021.130429] [Medline: 34126680]
3. Curtis L. PM2.5, NO2, wildfires, and other environmental exposures are linked to higher Covid 19 incidence, severity, and death rates. Environ Sci Pollut Res Int 2021 Oct;28(39):54429-54447. [doi: 10.1007/s11356-021-15556-0] [Medline: 34410599]
4. Chang L, Chong WT, Wang X, Pei F, Zhang X, Wang T, et al. Recent progress in research on PM2.5 in subways. Environ Sci Process Impacts 2021 May 26;23(5):642-663. [doi: 10.1039/d1em00002k] [Medline: 33889885]
5. Lin C, Wan C, Liu W, Wang H. PM2.5 induces early epithelial mesenchymal transition in human proximal tubular epithelial cells through activation of IL-6/STAT3 pathway. Int J Mol Sci 2021 Nov 25;22(23):12734 [FREE Full text] [doi: 10.3390/ijms222312734] [Medline: 34884542]
6. Michikawa T, Morokuma S, Takeda Y, Yamazaki S, Nakahara K, Takami A, et al. Maternal exposure to fine particulate matter over the first trimester and umbilical cord insertion abnormalities. Int J Epidemiol 2022 Feb 18;51(1):191-201. [doi: 10.1093/ije/dyab192] [Medline: 34524459]
7. Zhao T, Qi W, Yang P, Yang L, Shi Y, Zhou L, et al. Mechanisms of cardiovascular toxicity induced by PM2.5: a review. Environ Sci Pollut Res Int 2021 Dec;28(46):65033-65051. [doi: 10.1007/s11356-021-16735-9] [Medline: 34617228]
8. Ghosh R, Causey K, Burkart K, Wozniak S, Cohen A, Brauer M. Ambient and household PM2.5 pollution and adverse perinatal outcomes: A meta-regression and analysis of attributable global burden for 204 countries and territories. PLoS Med 2021 Sep;18(9):e1003718 [FREE Full text] [doi: 10.1371/journal.pmed.1003718] [Medline: 34582444]
9. Garovic VD, White WM, Vaughan L, Saiki M, Parashuram S, Garcia-Valencia O, et al. Incidence and long-term outcomes of hypertensive disorders of pregnancy. J Am Coll Cardiol 2020 May 12;75(18):2323-2334 [FREE Full text] [doi: 10.1016/j.jacc.2020.03.028] [Medline: 32381164]
10. Shah NS, Harrington KA, Huang X, Cameron NA, Yee LM, Khan SS. Trends in de novo hypertensive disorders of pregnancy among Asian and Hispanic population subgroups in the United States, 2011 to 2019. JAMA Cardiol 2022 Jul 01;7(7):742-746. [doi: 10.1001/jamacardio.2022.1378] [Medline: 35675084]
11. Garovic VD, Dechend R, Easterling T, Karumanchi SA, McMurtry Baird S, Magee LA, American Heart Association Council on Hypertension; Council on the Kidney in Cardiovascular Disease, Kidney in Heart Disease Science Committee; Council on Arteriosclerosis, ThrombosisVascular Biology; Council on Lifestyle and Cardiometabolic Health; Council on Peripheral Vascular Disease;Stroke Council. Hypertension in pregnancy: diagnosis, blood pressure goals, and pharmacotherapy: a scientific statement from the American Heart Association. Hypertension 2022 Feb;79(2):e21-e41 [FREE Full text] [doi: 10.1161/HYP.0000000000000208] [Medline: 34905954]
12. Feng X, Zhu J, Zhang L, Song L, Hipgrave D, Guo S, et al. Socio-economic disparities in maternal mortality in China between 1996 and 2006. BJOG 2010 Nov;117(12):1527-1536. [doi: 10.1111/j.1471-0528.2010.02707.x] [Medline: 20937073]

13. Su X, Zhao Y, Yang Y, Hua J. Correlation between exposure to fine particulate matter and hypertensive disorders of pregnancy in Shanghai, China. Environ Health 2020 Sep 17;19(1):101 [FREE Full text] [doi: 10.1186/s12940-020-00655-1] [Medline: 32943053]
14. Xu X, Hu H, Ha S, Roth J. Ambient air pollution and hypertensive disorder of pregnancy. J Epidemiol Community Health 2014 Jan;68(1):13-20 [FREE Full text] [doi: 10.1136/jech-2013-202902] [Medline: 24022815]
15. Grandjean P, Barouki R, Bellinger DC, Casteleyn L, Chadwick LH, Cordier S, et al. Life-long implications of developmental exposure to environmental stressors: new perspectives. Endocrinology 2015 Oct;156(10):3408-3415 [FREE Full text] [doi: 10.1210/EN.2015-1350] [Medline: 26241067]
16. Sun S, Wang X, Ding L, Zhang Q, Li N, Sui X, et al. Association between preconceptional air pollution exposure and medical purposes for selective termination of pregnancy. Environ Res 2021 Nov;202:111743. [doi: 10.1016/j.envres.2021.111743] [Medline: 34331927]
17. Zhang M, Wang X, Yang X, Dong T, Hu W, Guan Q, et al. Increased risk of gestational diabetes mellitus in women with higher prepregnancy ambient PM2.5 exposure. Sci Total Environ 2020 Aug 15;730:138982. [doi: 10.1016/j.scitotenv.2020.138982] [Medline: 32388108]
18. Leung M, Weisskopf MG, Laden F, Coull BA, Modest AM, Hacker MR, et al. Exposure to PM2.5 during pregnancy and fetal growth in Eastern Massachusetts, USA. Environ Health Perspect 2022 Jan;130(1):17004 [FREE Full text] [doi: 10.1289/EHP9824] [Medline: 34989624]
19. Wang K, Tian Y, Zheng H, Shan S, Zhao X, Liu C. Maternal exposure to ambient fine particulate matter and risk of premature rupture of membranes in Wuhan, Central China: a cohort study. Environ Health 2019 Nov 14;18(1):96 [FREE Full text] [doi: 10.1186/s12940-019-0534-y] [Medline: 31727105]
20. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
21. Hypertensive Disorders in Pregnancy Subgroup, Chinese Society of Obstetrics and Gynecology, Chinese Medical Association. [Diagnosis and treatment of hypertension and pre-eclampsia in pregnancy: a clinical practice guideline in China（2020）]. Zhonghua Fu Chan Ke Za Zhi 2020 Apr 25;55(4):227-238. [doi: 10.3760/cma.j.cn112141-20200114-00039] [Medline: 32375429]
22. Yuan Z, Wang H, Su T, Yang J, Chen J, Peng Y, et al. The first-trimester gestational weight gain associated with de novo hypertensive disorders during pregnancy: mediated by mean arterial pressure. Front Nutr 2022;9:862323 [FREE Full text] [doi: 10.3389/fnut.2022.862323] [Medline: 35495902]
23. Zhou S, Lin L, Bao Z, Meng T, Wang S, Chen G, et al. The association of prenatal exposure to particulate matter with infant growth: A birth cohort study in Beijing, China. Environ Pollut 2021 May 15;277:116792. [doi: 10.1016/j.envpol.2021.116792] [Medline: 33721799]
24. Climate Data Online. National Oceanic and Atmospheric Administration. URL: https://www.ncdc.noaa.gov/cdo-web/ [accessed 2023-01-11]
25. Lawn JE, Blencowe H, Waiswa P, Amouzou A, Mathers C, Hogan D, Lancet Ending Preventable Stillbirths Series study group, Lancet Stillbirth Epidemiology investigator group. Stillbirths: rates, risk factors, and acceleration towards 2030. Lancet 2016 Feb 06;387(10018):587-603. [doi: 10.1016/S0140-6736(15)00837-5] [Medline: 26794078]
26. Moolgavkar SH, Chang ET, Watson HN, Lau EC. An assessment of the Cox proportional hazards regression model for epidemiologic studies. Risk Anal 2018 Apr;38(4):777-794. [doi: 10.1111/risa.12865] [Medline: 29168991]
27. Ling H, Qing L, Jian X, Lishu S, Liang L, Qian W, et al. Strategies towards PM2.5 attainment for non-compliant cities in China: A case study. J Environ Manage 2021 Nov 15;298:113529. [doi: 10.1016/j.jenvman.2021.113529] [Medline: 34426226]
28. Mobasher Z, Salam MT, Goodwin T, Lurmann F, Ingles SA, Wilson ML. Associations between ambient air pollution and hypertensive disorders of pregnancy. Environ Res 2013 May;123:9-16 [FREE Full text] [doi: 10.1016/j.envres.2013.01.006] [Medline: 23522615]
29. Choe S, Kauderer S, Eliot MN, Glazer KB, Kingsley SL, Carlson L, et al. Air pollution, land use, and complications of pregnancy. Sci Total Environ 2018 Dec 15;645:1057-1064. [doi: 10.1016/j.scitotenv.2018.07.237] [Medline: 30248831]
30. Savitz DA, Elston B, Bobb JF, Clougherty JE, Dominici F, Ito K, et al. Ambient fine particulate matter, nitrogen dioxide, and hypertensive disorders of pregnancy in New York City. Epidemiology 2015 Sep;26(5):748-757 [FREE Full text] [doi: 10.1097/EDE.0000000000000349] [Medline: 26237745]
31. Xiong T, Chen P, Mu Y, Li X, Di B, Li J, et al. Association between ambient temperature and hypertensive disorders in pregnancy in China. Nat Commun 2020 Jun 10;11(1):2925 [FREE Full text] [doi: 10.1038/s41467-020-16775-8] [Medline: 32522990]
32. Bornstein E, Eliner Y, Chervenak FA, Grünebaum A. Racial disparity in pregnancy risks and complications in the US: temporal changes during 2007-2018. J Clin Med 2020 May 10;9(5):1414 [FREE Full text] [doi: 10.3390/jcm9051414] [Medline: 32397663]
33. Duley L. The global impact of pre-eclampsia and eclampsia. Semin Perinatol 2009 Jun;33(3):130-137. [doi: 10.1053/j.semperi.2009.02.010] [Medline: 19464502]

34. Jeyabalan A. Epidemiology of preeclampsia: impact of obesity. Nutr Rev 2013 Oct;71 Suppl 1(0 1):S18-S25 [FREE Full text] [doi: 10.1111/nure.12055] [Medline: 24147919]
35. Marić I, Tsur A, Aghaeepour N, Montanari A, Stevenson DK, Shaw GM, et al. Early prediction of preeclampsia via machine learning. Am J Obstet Gynecol MFM 2020 May;2(2):100100. [doi: 10.1016/j.ajogmf.2020.100100] [Medline: 33345966]
36. Jwa SC, Fujiwara T, Hata A, Arata N, Sago H, Ohya Y. BMI mediates the association between low educational level and higher blood pressure during pregnancy in Japan. BMC Public Health 2013 Apr 25;13(1):389 [FREE Full text] [doi: 10.1186/1471-2458-13-389] [Medline: 23617809]
37. Sinkey RG, Battarbee AN, Bello NA, Ives CW, Oparil S, Tita ATN. Prevention, diagnosis, and management of hypertensive disorders of pregnancy: a comparison of international guidelines. Curr Hypertens Rep 2020 Aug 27;22(9):66 [FREE Full text] [doi: 10.1007/s11906-020-01082-w] [Medline: 32852691]
38. Sun M, Yan W, Fang K, Chen D, Liu J, Chen Y, et al. The correlation between PM2.5 exposure and hypertensive disorders in pregnancy: A meta-analysis. Sci Total Environ 2020 Feb 10;703:134985. [doi: 10.1016/j.scitotenv.2019.134985] [Medline: 31731170]
39. Bai W, Li Y, Niu Y, Ding Y, Yu X, Zhu B, et al. Association between ambient air pollution and pregnancy complications: A systematic review and meta-analysis of cohort studies. Environ Res 2020 Jun;185:109471. [doi: 10.1016/j.envres.2020.109471] [Medline: 32276169]

Abbreviations

DLNM: distributed lag nonlinear model
GE: gestational hypertension
HDP: hypertensive disorder of pregnancy
HR: hazard ratio
ICD-10: International Classification of Diseases–10
PE: preeclampsia
PM: particulate matter
WHO: World Health Organization

Edited by A Mavragani, G Eysenbach; submitted 26.07.22; peer-reviewed by B Padhi, A Finny; comments to author 12.09.22; revised version received 20.09.22; accepted 30.11.22; published 23.01.23.

Please cite as:
Yuan Z, Wang HJ, Li Q, Su T, Yang J, Chen J, Peng Y, Zhou S, Bao H, Luo S, Wang H, Liu J, Han N, Guo Y, Ji Y
Risk of De Novo Hypertensive Disorders of Pregnancy After Exposure to PM1 and PM2.5 During the Period From Preconception to Delivery: Birth Cohort Study
JMIR Public Health Surveill 2023;9:e41442
URL: https://publichealth.jmir.org/2023/1/e41442
doi:10.2196/41442
PMID:36689262

Original Paper

Trends in Exposure to Secondhand Smoke Among Adolescents in China From 2013-2014 to 2019: Two Repeated National Cross-sectional Surveys

Chuanwei Ma[1*], MSc; Yayang Huang[2,3*], MD; Sixuan Li[4,5*], MPH; Min Zhao[6], MD; Xinying Zeng[5], BSc; Xinbo Di[5], BA; Costan G Magnussen[7,8,9], PhD; Bo Xi[1], MD; Shiwei Liu[5], PhD

[1]Department of Epidemiology, School of Public Health, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

[2]Health Education Institute, Beijing Center for Disease Prevention and Control, Beijing, China

[3]Institute of Child and Adolescent Health, School of Public Health, Peking University, Beijing, China

[4]Department of Chronic and Noncommunicable Disease Control and Prevention, Ningbo Municipal Center for Disease Control and Prevention, Ningbo, China

[5]Tobacco Control Office, Chinese Center for Disease Control and Prevention, Beijing, China

[6]Department of Nutrition and Food Hygiene, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

[7]Baker Heart and Diabetes Institute, Victoria, Australia

[8]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland

[9]Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland

[*]these authors contributed equally

Corresponding Author:
Shiwei Liu, PhD
Tobacco Control Office
Chinese Center for Disease Control and Prevention
27 Nanwei Road
Beijing, 100050
China
Phone: 86 1063185150
Email: shiwei_liu@aliyun.com

Abstract

Background: It is well-known that secondhand smoke exposure in childhood or adolescence is positively associated with morbidity and mortality. However, less is known about the current status of and most recent trends in secondhand smoke exposure among adolescents in China.

Objective: We aimed to assess recent changes in the prevalence of secondhand smoke exposure among adolescents in China using nationally representative data.

Methods: We used data from 2 repeated national cross-sectional surveys conducted in 2013-2014 and 2019. A total of 155,117 students (median age 13.5 years) in 2013-2014 and 147,270 students (median age 13.1 years) in 2019 were included in this study. Sociodemographic factors and secondhand smoke exposure information were collected via a standardized questionnaire. Exposure was defined as secondhand smoke exposure ≥1 day during the past 7 days at home or in public places. Other frequencies of secondhand smoke exposure (ie, ≥3 days, ≥5 days, and every day) during the past 7 days were also assessed. The weighted prevalence of secondhand smoke exposure was calculated according to the complex sample design for surveys.

Results: The prevalence of secondhand smoke exposure in any place (home or public places ≥1 day during the past 7 days) decreased from 2013-2014 (72.9%, 95% CI 71.5%-74.3%) to 2019 (63.2%, 95% CI 62%-64.5%), as did exposure at home (2013-2014: 44.4%, 95% CI 43.1%-45.7%; 2019: 34.1%, 95% CI 33.1%-35.2%) and in public places (2013-2014: 68.3%, 95% CI 66.9%-69.6%; 2019: 57.3%, 95% CI 56%-58.6%). The prevalence of secondhand smoke exposure decreased with increased gross domestic product per capita in each of the 2 survey years irrespective of exposure frequency or location. The prevalence of exposure at other frequencies (ie, ≥3 days, ≥5 days, or every day during the past 7 days) also decreased in any place, at home, and in public places. Secondhand smoke exposure was associated with higher school grade level (ninth vs seventh grade: odds ratio [OR] 1.76, 95% CI 1.68-1.84), gender (boys vs girls: OR 1.18, 95% CI 1.15-1.22), urban status (urban vs rural: OR 1.10, 95% CI 1.01-1.19), and cigarette smoking (yes vs no: OR 6.67, 95% CI 5.83-7.62).

Conclusions: Although the prevalence of secondhand smoke exposure among Chinese adolescents declined from 2013-2014 to 2019, it remains unacceptably high. More effective strategies and stronger action are needed in China to further, and dramatically, curb secondhand smoke exposure among adolescents.

(JMIR Public Health Surveill 2023;9:e40782) doi:10.2196/40782

KEYWORDS

secondhand smoke exposure; trends; adolescents; China; secondhand smoke; youth

Introduction

Background

Tobacco products, especially combustible ones, are harmful not only to users' health but also the health of nonsmokers who are exposed to secondhand smoke [1]. The World Health Organization (WHO) reports that exposure to secondhand smoke leads to 1.2 million deaths annually [2]. There is no safe or risk-free level of secondhand smoke exposure for children and adolescents [3], and more than 65,000 children die each year from diseases related to secondhand smoke exposure [2]. Secondhand smoke exposure occurs in various environments (at home or in any public place) and causes short-term damage in childhood (eg, slowed lung growth, worsened asthma, respiratory infections, and sudden infant death syndrome) [1], a long-term burden of cardiovascular disease [4], and mortality in adulthood [5].

China has the largest number of tobacco consumers in the world. Although the prevalence of current tobacco smoking decreased from 1990 to 2019 (males: –18.2%; females: –20.9%) among individuals aged 15 years or older in China, the prevalence of current tobacco smoking remained high in 2019 (males: 49.7%; females: 3.54%) [6]. Secondhand smoke exposure remains a serious public health concern in China despite evidence that has suggested the prevalence of secondhand smoke exposure in the workplace declined from 55.2% in 2010 to 45.3% in 2015 among Chinese nonsmokers aged 16 to 60 years [7]. In addition, national data from 2013 to 2016 showed that 41.7% (95% CI 40.9%-42.3%) of Chinese youth aged 6 to 17 years were exposed to secondhand smoke at home ≥1 day during the past 7 days [8]. This prevalence of secondhand smoke exposure among Chinese adolescents was higher than the global average level (72.9% vs 62.9%, respectively) [9].

Although considerable work on tobacco control has been undertaken in China in recent years, including the protection of teenagers from secondhand smoke exposure [10,11], there are no contemporary data on the prevalence of secondhand smoke exposure among adolescents in this country. Moreover, most previous studies only reported the prevalence of secondhand smoke exposure at home and public places in general [9,12,13]; no studies were able to precisely categorize the types of public places where exposure occurs. Such data are needed to monitor and guide effective interventions to protect adolescents from secondhand smoke exposure.

Objectives

Using repeated nationally representative data, we aimed to assess trends in the prevalence of secondhand smoke exposure between 2013-2014 and 2019 according to different locations (at home, in closed public places, in open public places, and on public transportation).

Methods

Study Participants

We used repeated cross-sectional data on secondhand smoke exposure collected from the Global Youth Tobacco Survey (GYTS)-China, which was a part of the GYTS conducted in China from October 1, 2013, to April 1, 2014, and the China National Youth Tobacco Survey (NYTS), conducted from May 1, 2019, to December 31, 2019. Each of these 2 national surveys was a national, representative, school-based, cross-sectional survey administered to middle school students. Both surveys used similar sampling, questionnaires, and survey methodologies and covered 31 provinces, autonomous regions, and municipalities (hereafter referred to as “provinces”) in China. Both surveys used the same 3-stage stratified cluster random sampling strategy. In the first stage, surveillance points were randomly selected from each province, stratified by urban-rural area by probability proportional to size sampling proportional to population size, to be nationally representative, with a total of 336 surveillance points in 2013-2014 and 347 surveillance points in 2019. In the second stage, 3 middle schools were randomly selected at each surveillance point. In the third stage, 1 class was randomly selected from each grade of the selected school. All students in the selected class were eligible to participate. The median age of the study population was 13.5 years in 2013-2014 and 13.1 years in 2019.

Definition of Secondhand Smoke Exposure

Exposure to secondhand smoke at home was assessed with the question “During the past 7 days, on how many days have people smoked in your presence, in your home?”; exposure in closed public places was assessed with the question “During the past 7 days, on how many days have people smoked in your presence in closed public places other than in your home (such as school buildings, conference halls, gymnasiums, bars, shops, restaurants, shopping malls, movie theaters, etc.)?”; exposure in open public places was assessed with the question “During the past 7 days, on how many days have people smoked in your presence in open public places (such as playgrounds, sidewalks, entrances to buildings, parks, bus stops, parks, etc.)?”; exposure on public transportation was assessed with the question “During the past 7 days, on how many days have people smoked in your presence in any public transportation vehicles (such as trains, buses, or taxicabs, etc.)?" In addition to reporting on secondhand smoke exposure in the individual public places mentioned above, we also unified heterogeneous information by combining exposure that was reported in closed public places, in open public places, and on public transportation together as one variable: "public places." Moreover, we also combined exposure at home and public places as "any place." Exposure to secondhand smoke at any location (ie, at home, in closed public places, in open public places, on public transportation, in any public place, or in any place) was defined as exposure to secondhand smoke on ≥1 day during the past 7 days. We also assessed different frequencies (≥3 days, ≥5 days, and every day during the past 7 days) of secondhand smoke exposure based on the students' responses. To ensure each participant understood that the question was asked about secondhand smoke from others, rather than from themselves, each question was explained by the trained investigators beforehand.

Cigarette smoking was defined as smoking on ≥1 day during the past 30 days with the question "During the past 30 days, on how many days did you smoke cigarettes?" [14]. Exposure to secondhand smoke in schools was assessed with the question "During the past 30 days, did you see anyone smoke inside the school building or outside on school property?" Attitude on the hazards of secondhand smoke exposure was assessed with the question "Do you think the smoke from other people's cigarettes smoking is harmful to you?" with the corresponding answers "Definitely not, probably not, probably yes, and definitely yes." Participants were classified into groups based on whether the gross domestic product (GDP) per capita of their province was low, middle, or high. For 2013-2014 (when the exchange rate was US $1=RMB 6.08), the GDP categories were RMB <35,000, RMB 35,000 to 50,000, and RMB >50,000, respectively. For 2019 (when the exchange rate was US $1=RMB 6.98), the categories were RMB <50,000, RMB 50,000 to 70,000, and RMB >70,000, respectively. Further details of the GDP per capita in 2013 and 2019 in China are available online [15,16].

Ethical Considerations

Both surveys (2013-2014 and 2019) were approved by the National Health Commission, and the proposal and protocols were reviewed by the Institutional Review Board of the Chinese Center for Disease Control and Prevention (202008). All participating students gave oral informed consent. Our study did not contain any identifiable information from individual participants.

Statistical Analysis

The national, regional, and subgroup (by sex, residence, GDP per capita category, and cigarette smoking status) prevalence with the 95% CI of secondhand smoke exposure in 2013-2014 and 2019 among adolescents was calculated with consideration of sampling weights, strata, and primary sampling units from each survey. The absolute change rate with 95% CI in the prevalence of secondhand smoke exposure was calculated as the prevalence in 2013-2014 subtracted from the prevalence in 2019. The chi-square trend test was used to compare differences in prevalence by school grade, sex, residence, region, GDP per capita category, and cigarette smoking status between 2013-2014 and 2019. Multivariable logistic regression models were used to assess the association between secondhand smoke exposure and its potential influencing factors (school grade, sex, residence, region, GDP per capita category, and cigarette smoking status). All analyses were conducted using SAS (SAS Institute). Two-sided P values <.05 were considered statistically significant.

Results

Participant Characteristics

Table 1 presents the demographic data from the 2 national surveys conducted in 2013-2014 and 2019. In 2019, the proportions of adolescents who were in a higher school grade (ninth grade), resided in a rural region, and were current cigarette users were somewhat lower than in 2013-2014. Data from 155,117 adolescents (boys: n=82,057, 52.9%) in the 2013-2014 survey and 147,270 adolescents (boys: n=78,789, 53.5%) in the 2019 survey were included in this study.

Table 1. Demographics of 2 national surveys conducted in 2013-2014 (n=155,117) and 2019 (n=147,270).

	Prevalence in 2013-2014, %	95% CI	Prevalence in 2019, %	95% CI
Grade				
Seventh	33.1	32.5-33.7	34.8	34.2-35.3
Eighth	33.3	32.9-33.7	33.5	33.1-33.9
Ninth	33.6	32.9-34.4	31.7	31.2-32.3
Sex				
Boys	52.9	52.3-53.5	53.5	53.1-54
Girls	47.1	46.5-47.7	46.5	46-46.9
Residence				
Urban	27.9	26-29.8	36.3	34.1-38.5
Rural	72.1	70.2-74	63.7	61.5-65.9
Region				
North	10.6	9.4-11.8	11.1	9.8-12.4
East	26.3	24.7-27.9	28.4	26.3-30.6
Central	17.1	15.8-18.4	18.3	15.9-20.6
South	14.6	12.5-16.8	13.3	11.9-14.8
Southwest	16.9	15.3-18.6	16	14.2-17.9
Northwest	8.3	7.3-9.4	7.3	6.5-8.1
Northeast	6.1	5.1-7.2	5.5	5-6.1
Gross domestic product per capita category				
Low	27.5	25.5-29.5	27	25.1-28.9
Middle	38.4	36.4-40.3	38.6	36.2-41.1
High	34.2	32-36.3	34.4	32.1-36.6
Cigarette smoking				
No	94.1	93.5-94.6	96.1	95.6-96.6
Yes	5.9	5.4-6.5	3.9	3.4-4.4

Trends in Prevalence of Secondhand Smoke Exposure

The overall prevalence of secondhand smoke exposure in any place (on ≥1 day during the past 7 days) decreased from 2013-2014 (72.9%, 95% CI 71.5%-74.3%) to 2019 (63.2%, 95% CI 62%-64.5%), while the prevalence of secondhand smoke exposure at home decreased from 44.4% (95% CI 43.1%-45.7%) in 2013-2014 to 34.1% (95% CI 33.1%-35.2%) in 2019, as did secondhand smoke exposure in public places (2013-2014: 68.3%, 95% CI 66.9%-69.6%; 2019: 57.3%, 95% CI 56%-58.6%), as shown in Tables 2, 3, and 4. The prevalence was high in any place, in public places, and at home across all 31 provinces, with the prevalence of secondhand smoke exposure in public places higher compared to exposure at home in each province in both 2013-2014 and 2019 (Figure 1 and Multimedia Appendix 1, Tables S1-S4). The prevalence exceeded 60% in 2019 in 25 of 31 provinces. The 5 provinces where the prevalence was much higher than other provinces were Yunnan (75.1%, 95% CI 66.2%-84%), Hunan (73.6%, 95% CI 68.9%-78.3%), Beijing (73.2%, 95% CI 69.5%-76.9%), Shanghai (72.5%, 95% CI 62.5%-82.5%), and Gansu (70.2%, 95% CI 62.5%-78%). The prevalence of secondhand smoke exposure also decreased according to different exposure frequencies and locations. For example, there was a downward trend in daily exposure in any place (from 31%, 95% CI 29.9%-32.1% in 2013-2014 to 24.7%, 95% CI 23.8%-25.7% in 2019), at home (from 14.2%, 95% CI 13.6%-14.8% in 2013-2014 to 12.4%, 95% CI 11.8%-13% in 2019), and in public places (from 25.1%, 95% CI 24.2%-26% in 2013-2014 to 18.7%, 95% CI 17.8%-19.5% in 2019), as shown in Tables 2, 3, and 4 and Multimedia Appendix 1, Table S5.

The prevalence of secondhand smoke exposure increased as school grade increased regardless of exposure frequency (≥1 day, ≥3 days, ≥5 days, or 7 days during the past 7 days), location (any place, home, or public place), and survey year (2013-2014 or 2019). The prevalence of secondhand smoke exposure decreased as GDP per capita category increased irrespective of exposure frequency, location, or survey year (Tables 2, 3, and 4 and Multimedia Appendix 1, Table S5). The prevalence of exposure was higher among boys (vs girls), current cigarette smokers (vs nonsmokers), and adolescents who lived in an urban region (vs rural region, except for secondhand smoke exposure at home). A higher prevalence of secondhand smoke exposure was observed in the northwest, south, and southwest regions, with lower exposure in the east and northeast regions (Tables 2, 3, and 4). A downward trend in the prevalence of secondhand smoke exposure in any type of public place (closed public places, open public places, and public transportation) and in schools on ≥1 day during the past 7 days was also observed (Multimedia Appendix 1, Tables S6-S7 and Figure S1).

Table 2. Trends in the prevalence of secondhand smoke exposure in any place (on ≥1 day during the past 7 days) by age, sex, residence, region, gross domestic product per capita category, and status of cigarette smoking among Chinese adolescents from 2013-2014 to 2019.

	Prevalence in 2013-2014, %	95% CI	Prevalence in 2019, %	95% CI	Absolute change in prevalence, %	95% CI
Overall	72.9	71.5 to 74.3	63.2	62 to 64.5	–9.8	–11.8 to –8.1
Grade						
Seventh	65.7	64 to 67.3	55.8	54 to 57.6	–10	–12.6 to –7.7
Eighth	74.3	73.5 to 76.2	66.1	64.5 to 67.6	–8.3	–10.9 to –6.2
Ninth	78.7	77.5 to 79.9	68.5	67.2 to 69.7	–10.3	–12.1 to –8.7
Sex						
Boys	75.5	74.4 to 77	65.4	64 to 66.7	–10.4	–12.3 to –8.7
Girls	69.9	68.3 to 71.4	60.8	59.4 to 62.2	–9.2	–11.3 to –7.2
Residence						
Urban	74.4	73.3 to 75.6	63.5	61.7 to 65.3	–11	–13.3 to –8.9
Rural	72.4	70.5 to 74.2	63.1	61.4 to 64.8	–9.4	–12 to –7
Region						
North	73.1	70.3 to 75.9	63	58.8 to 67.3	–10.1	–15.2 to –4.9
East	68.9	66.8 to 71	59.9	57 to 62.9	–9	–12.6 to –5.4
Central	75.9	73.9 to 78	64	61.5 to 66.4	–12	–15.2 to –8.8
South	67.8	61.8 to 73.8	66.3	63.1 to 69.4	–1.7	–8.6 to 5.3
Southwest	77.5	75.1 to 79.8	65.2	62.1 to 68.2	–12.5	–16.4 to –8.5
Northwest	78.9	75.4 to 82.3	66.3	63.1 to 69.5	–12.6	–17.3 to –7.8
Northeast	75.7	73.8 to 77.6	61.2	57.8 to 64.7	–14.5	–18.5 to –10.5
Gross domestic product per capita category						
Low	75.4	73.4 to 77.4	65.4	62.6 to 68.1	–10.1	–13.6 to –6.7
Middle	75	73.4 to 76.6	64	62.3 to 65.7	–11.1	–13.5 to –8.8
High	69	66.2 to 71.9	60.7	58.4 to 63.1	–8.3	–12 to –4.6
Current cigarette smoking						
No	71.5	70.1 to 72.9	61.9	60.6 to 63.3	–9.6	–11.1 to –8.1
Yes	94.1	93.3 to 95	93.8	92.6 to 95	–0.5	–2 to 1.1

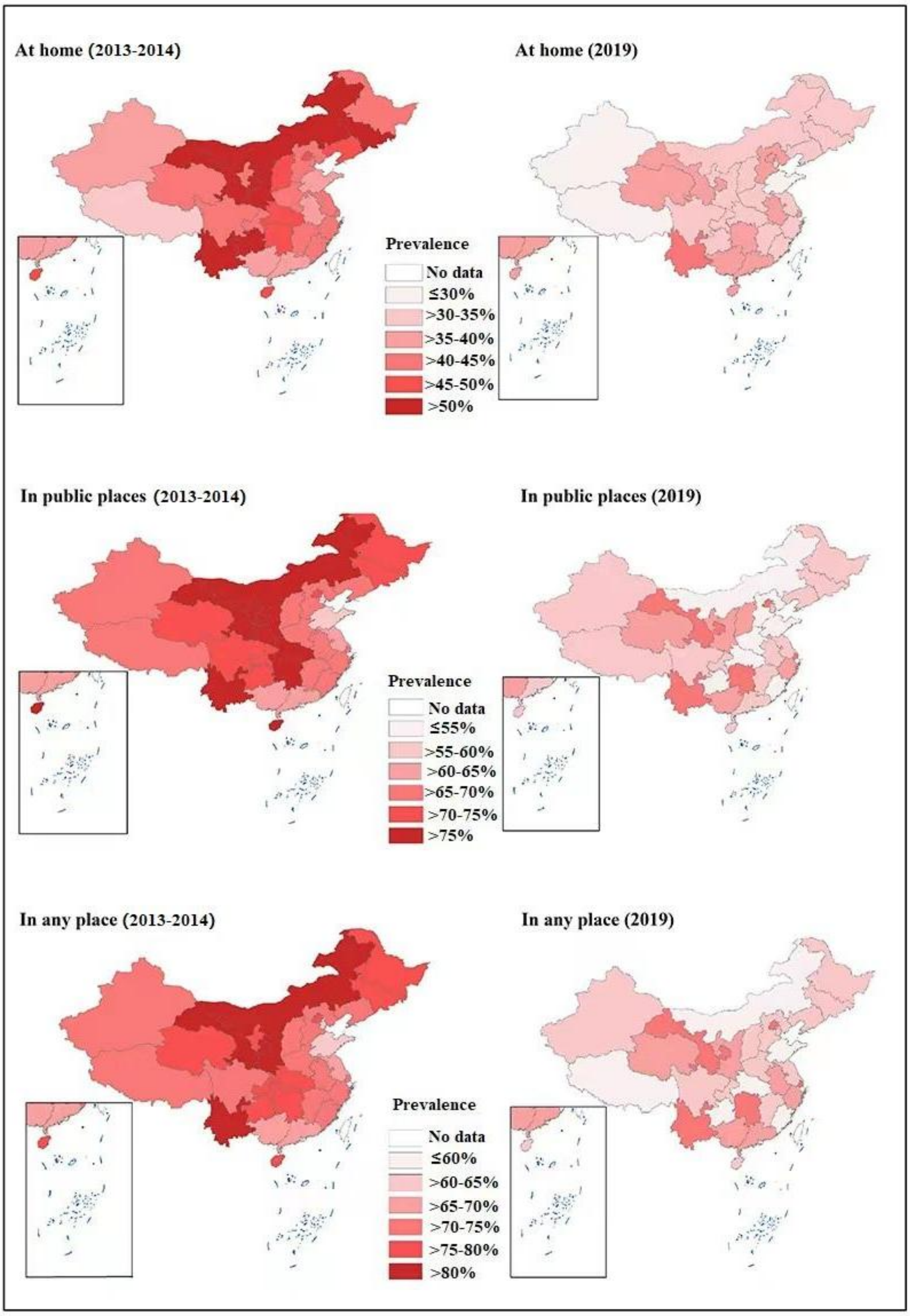
Table 3. Trends in the prevalence of secondhand smoke exposure at home (on ≥1 day during the past 7 days) by age, sex, residence, region, gross domestic product per capita category, and status of cigarette smoking in Chinese adolescents from 2013-2014 to 2019.

	Prevalence in 2013-2014, %	95% CI	Prevalence in 2019, %	95% CI	Absolute change in prevalence, %	95% CI
Overall	44.4	43.1 to 45.7	34.1	33.1 to 35.2	–10.4	–12 to –8.7
Grade						
Seventh	39.6	38.2 to 40.9	30.5	29.3 to 31.8	–9.1	–11 to –7.3
Eighth	46.1	44.4 to 47.7	36	34.7 to 37.3	–10.1	–12.1 to –8.1
Ninth	47.5	46.2 to 48.9	36.1	34.9 to 37.3	–11.5	–13.4 to –9.8
Sex						
Boys	46.4	45.1 to 47.8	36	34.8 to 37.1	–10.5	–12.3 to –8.9
Girls	42.1	40.8 to 43.5	32	30.9 to 33.2	–10.2	–12 to –8.4
Residence						
Urban	43.5	42.1 to 44.8	33.2	31.8 to 34.6	–10.3	–12.3 to –8.3
Rural	44.8	43.1 to 46.5	34.6	33.2 to 36.1	–10.2	–12.5 to –8
Region						
North	46	42.6 to 49.4	35.6	32.1 to 39.2	–10.4	–15.3 to –5.4
East	39.6	37.7 to 41.6	31.6	30 to 33.2	–8	–10.6 to –5.5
Central	45.8	43.7 to 48	33.7	31.5 to 35.9	–12.1	–15.3 to –9
South	40.2	34.9 to 45.5	38	34.1 to 42	–2.2	–8.9 to –4.6
Southwest	50.5	47.7 to 53.2	35.7	32.7 to 38.7	–14.8	–19 to –10.7
Northwest	48.9	45.7 to 52.1	33.1	29.8 to 36.4	–15.7	–20.4 to –11.1
Northeast	46.9	44.6 to 49.2	32.9	30.8 to 35	–14	–17.2 to –10.8
Gross domestic product per capita category						
Low	48.9	46.9 to 50.9	37	34.4 to 39.5	–11.9	–15.2 to –8.6
Middle	44.7	42.9 to 46.4	33.7	32.2 to 35.2	–11	–13.3 to –8.7
High	40.8	38.2 to 43.3	32.4	30.8 to 34	–8.4	–11.4 to –5.4
Current cigarette smoking						
No	42.8	41.5 to 44	32.7	31.7 to 33.7	–10.1	–11.7 to –8.5
Yes	68.2	66.6 to 69.7	66.3	64.1 to 68.4	–1.9	–4.6 to 0.7

Table 4. Trends in the prevalence of secondhand smoke exposure in public places (on ≥1 day during the past 7 days) by age, sex, residence, region, gross domestic product per capita category, and status of cigarette smoking among Chinese adolescents from 2013-2014 to 2019.

	Prevalence in 2013-2014, %	95% CI	Prevalence in 2019, %	95% CI	Absolute change in prevalence, %	95% CI
Overall	68.3	66.9 to 69.6	57.3	56 to 58.6	–11	–12.9 to –9.1
Grade						
Seventh	60.2	58.6 to 61.8	48.9	47.1 to 50.6	–11.4	–13.8 to –9
Eighth	69.8	68 to 71.5	60.1	58.4 to 61.7	–9.7	–12.2 to –7.3
Ninth	74.7	73.5 to 75.9	63.6	62.3 to 65	–11.1	–13 to –9.3
Sex						
Boys	71.3	70 to 72.5	59.5	58.1 to 60.9	–11.9	–13.7 to –10
Girls	64.9	63.4 to 66.4	54.8	53.4 to 56.2	–10.1	–12.2 to –8.1
Residence						
Urban	69.8	68.6 to 71	57.6	55.8 to 59.5	–12.2	–14.5 to –9.9
Rural	67.6	65.9 to 69.4	57.1	55.3 to 58.9	–10.6	–13.1 to –8.1
Region						
North	67.6	64.7 to 70.5	56.6	52.4 to 60.9	–11	–16.2 to –5.9
East	63.5	61.5 to 65.6	53.7	50.7 to 56.6	–9.9	–13.6 to –6.3
Central	71.5	69.2 to 73.7	58.5	56 to 61	–13	–16.4 to –9.6
South	62.8	57 to 68.6	58.7	55.8 to 61.6	–4.1	–10.7 to 2.5
Southwest	73.4	70.8 to 75.9	60.2	57 to 63.4	–13.3	–17.4 to –9.1
Northwest	75.1	71.8 to 78.3	61.9	58.1 to 65.6	–13.2	–18.3 to –8.2
Northeast	70.3	68.2 to 72.5	55.4	52 to 58.8	–14.9	–19.1 to –10.8
Gross domestic product per capita category						
Low	70.5	68.3 to 72.7	59.5	56.8 to 62.3	–11.1	–14.7 to –7.5
Middle	70.7	69.1 to 72.3	58.4	56.6 to 60.2	–12.4	–14.8 to –9.9
High	63.7	61 to 66.4	54.3	51.9 to 56.7	–9.4	–13.1 to –5.8
Current cigarette smoking						
No	66.5	65.2 to 67.9	55.9	54.6 to 57.2	–10.7	–12.6 to –8.8
Yes	91.6	90.4 to 92.8	89.9	88.2 to 91.5	–1.9	–3.9 to 0.2

Figure 1. Trends in the geographical distribution of secondhand smoke exposure (on ≥1 day during the past 7 days) at home, in public places, and in any place among Chinese adolescents from 2013-2014 to 2019.



Association Between Secondhand Smoke Exposure and Potential Factors

In the multivariable logistic regression analysis, adolescents were more likely to be exposed to secondhand smoke in the combined data set (2013-2014 and 2019) if they were in a higher school grade (ninth vs seventh grade: odds ratio [OR] 1.76, 95% CI 1.68-1.84), were boys (boys vs girls: OR 1.18, 95% CI 1.15-1.22), lived in an urban setting (urban vs rural: OR 1.10, 95% CI 1.01-1.19), and were a current cigarette smoker (smoker vs nonsmoker: OR 6.67, 95% CI 5.83-7.62). These results were largely independent of survey year (2013-2014 or 2019; Table 5).

Table 5. Multivariable odds ratios and 95% CIs for factors associated with secondhand smoke exposure (on ≥1 day during the past 7 days) among Chinese adolescents. All variables listed in the table were introduced into logistic regression models.

	In any place, odds ratio (95% CI)			At home, odds ratio (95% CI)			In public places, odds ratio (95% CI)		
	2013-2014	2019	Total	2013-2014	2019	Total	2013-2014	2019	Total
Grade									
Seventh	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Eighth	1.49 (1.40-1.60)	1.52 (1.42-1.62)	1.50 (1.44-1.58)	1.28 (1.23-1.34)	1.26 (1.18-1.33)	1.27 (1.22-1.32)	1.50 (1.40-1.59)	1.55 (1.45-1.65)	1.52 (1.45-1.59)
Ninth	1.86 (1.75-1.99)	1.66 (1.56-1.78)	1.76 (1.68-1.84)	1.34 (1.28-1.40)	1.23 (1.17-1.30)	1.29 (1.25-1.34)	1.89 (1.77-2.01)	1.77 (1.65-1.90)	1.83 (1.74-1.91)
Sex									
Girls	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Boys	1.23 (1.19-1.30)	1.16 (1.12-1.20)	1.18 (1.15-1.22)	1.10 (1.06-1.14)	1.12 (1.09-1.16)	1.10 (1.08-1.12)	1.25 (1.20-1.30)	1.16 (1.11-1.20)	1.18 (1.11-1.20)
Residence									
Rural	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Urban	1.22 (1.09-1.37)	1.09 (0.98-1.22)	1.10 (1.01-1.19)	1.01 (0.92-1.10)	0.99 (0.91-1.08)	0.96 (0.89-1.03)	1.22 (1.10-1.35)	1.11 (0.99-1.23)	1.10 (1.02-1.19)
Region									
North	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
East	0.91 (0.73-1.14)	0.96 (0.77-1.21)	0.93 (0.79-1.09)	0.86 (0.71-1.03)	0.93 (0.77-1.12)	0.89 (0.77-1.01)	0.92 (0.75-1.13)	0.97 (0.78-1.21)	0.94 (0.80-1.10)
Central	1.18 (0.95-1.46)	1.06 (0.85-1.32)	1.09 (0.93-1.28)	1.07 (0.9-1.28)	0.95 (0.78-1.16)	1.00 (0.87-1.15)	1.20 (0.98-1.46)	1.11 (0.89-1.37)	1.12 (0.96-1.31)
South	0.80 (0.61-1.05)	1.20 (0.95-1.52)	0.99 (0.82-1.19)	0.81 (0.65-1.01)	1.16 (0.92-1.47)	0.97 (0.82-1.14)	0.83 (0.64-1.07)	1.14 (0.92-1.40)	0.98 (0.82-1.16)
Southwest	1.24 (1.01-1.51)	1.06 (0.84-1.34)	1.13 (0.96-1.33)	1.17 (0.98-1.39)	0.97 (0.79-1.18)	1.07 (0.94-1.23)	1.28 (1.06-1.54)	1.12 (0.90-1.41)	1.19 (1.02-1.39)
Northwest	1.37 (1.07-1.77)	1.15 (0.90-1.45)	1.26 (1.05-1.51)	1.18 (0.98-1.42)	0.90 (0.73-1.12)	1.06 (0.92-1.23)	1.42 (1.14-1.77)	1.24 (0.97-1.58)	1.34 (1.12-1.59)
Northeast	1.13 (0.94-1.34)	0.91 (0.72-1.15)	1.02 (0.87-1.20)	1.05 (0.89-1.25)	0.90 (0.75-1.07)	1.00 (0.88-1.13)	1.10 (0.93-1.30)	0.94 (0.75-1.17)	1.03 (0.89-1.20)
Gross domestic product per capita category									
Low	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Middle	0.96 (0.81-1.14)	0.99 (0.85-1.15)	0.98 (0.87-1.10)	0.87 (0.76-0.99)	0.95 (0.82-1.09)	0.91 (0.82-1.00)	0.99 (0.85-1.16)	0.99 (0.85-1.14)	0.99 (0.89-1.11)
High	0.86 (0.69-1.07)	0.87 (0.73-1.03)	0.88 (0.76-1.01)	0.88 (0.75-1.04)	0.88 (0.75-1.03)	0.89 (0.79-1.00)	0.86 (0.70-1.06)	0.87 (0.74-1.02)	0.88 (0.77-1.00)
Current cigarette smoking									
No	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Yes	5.31 (4.45-6.35)	8.07 (6.60-9.87)	6.67 (5.83-7.62)	2.51 (2.31-2.73)	3.65 (3.36-3.96)	3.04 (2.87-3.22)	4.53 (3.89-5.28)	6.10 (5.08-7.32)	5.39 (4.80-6.05)

Trends in Attitude Toward Secondhand Smoke Exposure

We observed no change in the proportion of adolescents who thought secondhand smoke exposure was harmful to them between the surveys (2013-2014: 73.9%, 95% CI 73.3%-74.5%; 2019: 73.7%, 95% CI 72.7%-74.8%). This observation was consistent across subgroups by grade, residence, region, and GDP per capita category. However, we noted fewer girls and cigarette smokers and more boys, proportionally, considered smoking to be harmful in 2019 compared to 2013-2014 (Multimedia Appendix 1, Table S8).

Discussion

Principal Findings

A decrease in the prevalence of secondhand smoke exposure occurred from 2013-2014 to 2019 among middle school–aged students in China. This decline was observed irrespective of sex, residence, region, provincial economic level, and location of exposure (in any place, at home, and in public places). However, the prevalence remained high in 2019 (63.2% in any place, 34.1% at home, and 57.3% in public places), highlighting the need for more effective strategies to control tobacco use and further curtail secondhand smoke exposure in China.

To the best of our knowledge, this is the first study to assess in detail the latest information on the prevalence of secondhand smoke exposure and trends in this exposure among adolescents in China using nationally representative data. Overall, in 2019 the prevalence (on ≥1 day during the past 7 days) was 34.1% (95% CI 33.1%-35.2%) at home, which is higher than in the United States (25.3%, 95% CI 23.4%-27.3%), while the prevalence of exposure was similar on public transportation between the 2 countries (China: 23.4%, 95% CI 22.3%-24.4%; United States: 23.3%, 95% CI 21.4%-25.4%) [17]. The prevalence of exposure to secondhand smoke was 57.3% (95% CI 56%-58.6%) in public places and 63.2% (95% CI 62%-64.5%) in any place in 2019, which was higher than in many other countries [9,18,19]. In addition, nearly a quarter (24.7%, 95% CI 23.8%-25.7%) of adolescents were exposed to secondhand smoke in any place daily, one-tenth (12.4%, 95% CI 11.8%-13%) at home and one-fifth (18.7%, 95% CI 17.8%-19.5%) in public places. Data from the 2010-2018 GYTS among 711,366 adolescents aged 12 to 16 years in 142 countries showed that the global prevalence of secondhand smoke exposure (on ≥1 day) was 33.1% (95% CI 32.1%-34.1%) at home, 57.6% (95% CI 56.4%-58.8%) in public places, and 62.9% (95% CI 61.7%-64.1%) in any place; the daily exposure prevalence was 12.3% (95% CI 11.7%-13%) at home, 23.5% (95% CI 22.5%-24.5%) in public places, and 32.5% (95% CI 31.5%-33.6%) in any place [9]. While these data highlight the fact that secondhand smoke exposure remains a serious public health issue worldwide, our data, using the same instruments of exposure measurement, show that exposure in China is comparatively much higher.

One advantage of our data over previous studies is that our exposure assessment was able to further categorize the types of public places where exposure occurred. It is obvious that the health risk should be higher for exposure in enclosed public places than open public places. Our findings showed that the prevalence in enclosed public places (45.8%) was similar to that in open public places (48.6%) in 2019, although the prevalence was lower on public transportation (23.4%). Another important finding is that in 2019 nearly 50% of adolescents were exposed to secondhand smoke in schools, suggesting that many school staff are cigarette smokers who not only set a bad example for students, but also expose students to their secondhand smoke. These data reinforce the fact that secondhand smoke exposure among adolescents remains a serious public health issue, especially in public places, and identifies key exposure sites outside of the home that could be targeted for intervention.

We observed that the prevalence of secondhand smoke exposure increased as school grade increased, which can be explained as older adolescents from higher grades being more accepting of tobacco use among their peers and being more likely to smoke themselves; both factors increase the probability of secondhand smoke exposure [14,20]. We also found that the prevalence of exposure decreased when moving from low to high provincial GDP per capita categories, as others have shown [21,22]. This might be partly explained by lower awareness of the harm of secondhand smoke exposure among adolescents in lower GDP per capita regions than higher GDP per capita regions. Moreover, many economically developed municipalities, such as Beijing and Shanghai, have implemented stricter smoke-free policies than other municipalities or provinces. Notably, however, there was little difference observed in the prevalence of secondhand smoke exposure among adolescents across the 7 geographical regions. These findings further emphasize the fact that secondhand smoke exposure among adolescents remains a serious concern across mainland China. We found that almost all adolescent cigarette smokers (89.9%) were exposed to secondhand smoke (≥1 day during the past 7 days), although exposure was also high in adolescents who did not smoke cigarettes (55.9%). These data are consistent with those from other studies showing a positive relationship between secondhand smoke exposure and cigarette smoking [20,23]. This finding further suggests that preventing smoking among adolescents would help lower the prevalence of secondhand smoke exposure among this population.

Of note, the prevalence of secondhand smoke exposure among adolescents decreased by approximately 10% from 2013-2014 to 2019 regardless of location (in any place, at home, and in public places). In 2005, China ratified the WHO Framework Convention on Tobacco Control (FCTC), which came into force on January 9, 2006 [24]. Since then, a series of tobacco control measures have been drafted in China. In 2007, the Inter-Ministry Coordination and Steering Committee for Implementation of WHO FCTC was established under the State Council [25]. In 2009, the Ministry of Finance and State Taxation Administration decided to adjust the taxation of tobacco products. In 2000 (when the exchange rate was US $1=RMB 8.28), the tax rate was 45% for tobacco products priced RMB ≥50/carton and 30% for products priced RMB <50/carton; in 2009 (when the exchange rate was US $1=RMB 6.81), the rates were 56% for products priced RMB ≥70/carton and 36% for products priced RMB <70/carton) [26]. To protect youth from secondhand smoke exposure, the General Office of the Ministry of Education further strengthened tobacco control in schools (ie, strengthening tobacco control propaganda and education, ensuring that teachers set a good example for students, and establishing a perfect system for tobacco control) in 2010 [27]. Based on the actions above, China has made some progress in tobacco control. It was reported that the prevalence of secondhand smoke exposure among mainly adult people aged ≥15 years declined from 55.2% (95% CI 50.4%-59.9%) in 2010 to 45.3% (95% CI 41.4%-49.2%) in 2015 [7]. Despite these efforts, tobacco control is only slowly improving in China. We found that the prevalence of secondhand smoke exposure at home and in public places is still unacceptably high. To investigate whether a hygiene intervention for Chinese smokers in households with young children would lead to a reduction of secondhand smoke exposure among children, 180 families were randomly divided into an intervention group (n=98 families) and a control group (n=82 families). The results suggested that a smoking hygiene intervention might be an effective measure in reducing the prevalence of secondhand smoke exposure among children [28]. Smoke-free car laws have also been tested as rational and effective measures in reducing secondhand smoke exposure among youth [29,30]. Furthermore, we found that nearly three-fourths of adolescents are aware of the harms to them of secondhand smoke. Feng et al [31] reported that 68.5% of adult nonsmokers and 60.6% of adult smokers believed that secondhand smoke exposure could cause lung cancer. This provides a solid foundation for national smoke-free legislation. These findings call for stronger action to promote the process of achieving a smoke-free world.

Limitations

Although our study has several strengths (2 surveys were conducted in a nationally representative sample of Chinese adolescents and followed strict quality assurance and control procedures to ensure the validity and reliability of the data; to our knowledge, this study reports the most recent prevalence and trend estimates of secondhand smoke exposure among adolescents in mainland China), several limitations should be noted. First, the prevalence of secondhand smoke exposure was estimated by self-report, which is subject to recall bias. However, the questions used to collect the exposure data have been tested with good reliability and adequate validity [32,33]. Further studies using objective biomarkers of secondhand smoke exposure (eg, plasma, saliva, or urinary cotinine or other nicotine biomarkers) could improve the reliability of our results. Second, each of the 2 surveys was a school-based survey, and our estimates of prevalence are conservative given that risk behaviors among those who leave school may be higher than among those who remain in school [34]. Third, only middle school students were included in our analyses; thus, our findings could not be directly generalized to youth in other age groups. Fourth, our study did not provide enough information on secondhand smoke exposure (eg, exposure to different types of tobacco and exposure duration per day). Fifth, data from Taiwan, Hong Kong, and Macao were missing, thus limiting the generalizability of our findings to adolescents throughout China. Sixth, a causal association should not necessarily be inferred, because this study had a cross-sectional design. Seventh, many other factors influencing secondhand smoke exposure, such as social, economic, and individual factors, were not considered in this study.

Conclusions

Although the prevalence of secondhand smoke exposure among Chinese adolescents has declined from 2013-2014 to 2019, exposure remains unacceptably high. These findings suggest more effective strategies and stronger action, especially in public places, but also at home, are needed in China to further curtail secondhand smoke exposure among adolescents.

Acknowledgments

This study was supported by the Youth Team of Humanistic and Social Science of Shandong University (20820IFYT1902), the National Key Research and Development Program of China (2017YFC1310902), the National Natural Science Foundation of China (81872721), and the Chinese Central Government Key Project of Public Health Program (Z195110010005).

Data Availability

The data sets generated during and/or analyzed during the current study are available from the corresponding author S Liu on reasonable request.

Authors' Contributions

BX (xibo2007@126.com) and S Liu are co–corresponding authors. CM, YH, and S Li are cofirst authors. BX and S Liu designed the study and were the principal investigators. CM and YH drafted the first version of the manuscript. S Li analyzed the data. XZ and XD validated the data. S Liu and S Li accessed and verified the data. BX, CGM, MZ, and S Liu critically revised the manuscript. All authors approved the final version of the manuscript and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary files.
[DOC File , 2305 KB - publichealth_v9i1e40782_app1.doc]

References

1. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. U.S. Department of Health and Human Services. 2014. URL: https://www.hhs.gov/sites/default/files/consequences-smoking-exec-summary.pdf [accessed 2023-02-24]
2. Tobacco. World Health Organization. URL: https://www.who.int/health-topics/tobacco#tab=tab_1 [accessed 2022-07-03]
3. Makadia LD, Roper PJ, Andrews JO, Tingen MS. Tobacco use and smoke exposure in children: new trends, harm, and strategies to improve health outcomes. Curr Allergy Asthma Rep 2017 Aug;17(8):55. [doi: 10.1007/s11882-017-0723-0] [Medline: 28741144]
4. Pistilli M, Howard VJ, Safford MM, Lee BK, Lovasi GS, Cushman M, REGARDS Investigators. Association of secondhand tobacco smoke exposure during childhood on adult cardiovascular disease risk among never-smokers. Ann Epidemiol 2019 Apr;32:28-34.e1 [FREE Full text] [doi: 10.1016/j.annepidem.2019.01.012] [Medline: 30799203]
5. Diver WR, Jacobs EJ, Gapstur SM. Secondhand smoke exposure in childhood and adulthood in relation to adult mortality among never smokers. Am J Prev Med 2018 Sep;55(3):345-352. [doi: 10.1016/j.amepre.2018.05.005] [Medline: 30122215]
6. GBD 2019 Tobacco Collaborators. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990-2019: a systematic analysis from the Global Burden of Disease Study 2019. Lancet 2021 Jun 19;397(10292):2337-2360 [FREE Full text] [doi: 10.1016/S0140-6736(21)01169-7] [Medline: 34051883]
7. Xiao L, Jiang Y, Zhang J, Parascandola M. Secondhand smoke exposure among nonsmokers in China. Asian Pac J Cancer Prev 2020 Jul 01;21(S1):17-22 [FREE Full text] [doi: 10.31557/APJCP.2020.21.S1.17] [Medline: 32649166]
8. Xie M, Jia C, Zhang Y, Wang B, Qin N, Cao S, et al. Household exposure to secondhand smoke among Chinese children: status, determinants, and co-exposures. Int J Environ Res Public Health 2020 Jul 30;17(15):5524 [FREE Full text] [doi: 10.3390/ijerph17155524] [Medline: 32751695]
9. Ma C, Heiland EG, Li Z, Zhao M, Liang Y, Xi B. Global trends in the prevalence of secondhand smoke exposure among adolescents aged 12-16 years from 1999 to 2018: an analysis of repeated cross-sectional surveys. Lancet Glob Health 2021 Dec;9(12):e1667-e1678 [FREE Full text] [doi: 10.1016/S2214-109X(21)00365-X] [Medline: 34571047]
10. Yang G, Wang Y, Wu Y, Yang J, Wan X. The road to effective tobacco control in China. Lancet 2015 Mar 14;385(9972):1019-1028. [doi: 10.1016/S0140-6736(15)60174-X] [Medline: 25784349]
11. Fang J, Yang G, Wan X. 'Pro-tobacco propaganda': a case study of tobacco industry-sponsored elementary schools in China. Tob Control 2020 Jul;29(4):447-451 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054646] [Medline: 31302606]
12. Henderson E, Continente X, Fernández E, Tigova O, Cortés-Francisco N, Gallus S, TackSHS project investigators. Secondhand smoke exposure in outdoor children's playgrounds in 11 European countries. Environ Int 2021 Apr;149:105775 [FREE Full text] [doi: 10.1016/j.envint.2020.105775] [Medline: 33228970]
13. Yao T, Sung H, Wang Y, Lightwood J, Max W. Healthcare costs of secondhand smoke exposure at home for U.S. children. Am J Prev Med 2019 Feb;56(2):281-287 [FREE Full text] [doi: 10.1016/j.amepre.2018.08.013] [Medline: 30553690]
14. Ma C, Xi B, Li Z, Wu H, Zhao M, Liang Y, et al. Prevalence and trends in tobacco use among adolescents aged 13-15 years in 143 countries, 1999-2018: findings from the Global Youth Tobacco Surveys. Lancet Child Adolesc Health 2021 Apr;5(4):245-255. [doi: 10.1016/S2352-4642(20)30390-4] [Medline: 33545071]
15. Per Capita Gross Regional Product and Indices. National Bureau of Statistics of China. URL: http://www.stats.gov.cn/tjsj/ndsj/2014/zk/html/Z0315e.htm [accessed 2023-02-24]
16. Gross Regional Product (2019). National Bureau of Statistics of China. URL: http://www.stats.gov.cn/tjsj/ndsj/2020/indexeh.htm [accessed 2023-02-24]
17. Walton K, Gentzke AS, Murphy-Hoefer R, Kenemer B, Neff LJ. Exposure to secondhand smoke in homes and vehicles among US youths, United States, 2011-2019. Prev Chronic Dis 2020 Sep 10;17:E103 [FREE Full text] [doi: 10.5888/pcd17.200107] [Medline: 32915130]
18. Ling MYJ, Lim KH, Hasani WSR, Rifin HM, Majid NLA, Lourdes TGR, et al. Exposure to secondhand smoke among school-going adolescents in Malaysia: Findings from the tobacco and e-cigarettes survey among Malaysian adolescents (TECMA). Tob Induc Dis 2020;18:96 [FREE Full text] [doi: 10.18332/tid/128622] [Medline: 33262682]
19. Kim B, Kim HR. Associations between secondhand smoke and mental health in Korean adolescents. Asia Pac J Public Health 2020 Nov;32(8):406-413. [doi: 10.1177/1010539520957846] [Medline: 32929992]
20. Xi B, Liang Y, Liu Y, Yan Y, Zhao M, Ma C, et al. Tobacco use and second-hand smoke exposure in young adolescents aged 12-15 years: data from 68 low-income and middle-income countries. Lancet Glob Health 2016 Nov;4(11):e795-e805 [FREE Full text] [doi: 10.1016/S2214-109X(16)30187-5] [Medline: 27697484]
21. Rahman M, Haque MN, Rahman MM, Mostofa MG, Zahan MS, Mahmudul Hasan SM, et al. Socio-economic inequalities in secondhand smoke exposure at home in the context of mother-child pairs in Bangladesh. Public Health 2018 Aug;161:106-116. [doi: 10.1016/j.puhe.2018.04.018] [Medline: 30015081]
22. Bird Y, Staines-Orozco H, Moraros J. Adolescents' smoking experiences, family structure, parental smoking and socio-economic status in Ciudad Juárez, Mexico. Int J Equity Health 2016 Feb 20;15:29 [FREE Full text] [doi: 10.1186/s12939-016-0323-y] [Medline: 26897609]

23. Okoli CTC, Kodet J. A systematic review of secondhand tobacco smoke exposure and smoking behaviors: Smoking status, susceptibility, initiation, dependence, and cessation. Addict Behav 2015 Aug;47:22-32. [doi: 10.1016/j.addbeh.2015.03.018] [Medline: 25863004]
24. Tobacco control for 5 years. Central People's Government of the People's Republic of China. URL: http://www.gov.cn/jrzg/2011-01/08/content_1780569.htm [accessed 2023-01-20]
25. Reply from State Council on agreement to set up the Interministerial leading group to be in charge of FCTC implementation. State letter No 41. Central People's Government of the People's Republic of China. 2007. URL: http://www.gov.cn/zwgk/2007-04/30/content_603299.htm [accessed 2023-01-20]
26. Notice of the Ministry of Finance and the State Administration of Taxation on Adjusting the Consumption Tax Policy of Tobacco Products (No.84.2009). The Ministry of Finance and the State Administration of Taxation. URL: https://www.shui5.cn/article/cf/24017.html [accessed 2023-02-24]
27. Opinions on Further Strengthening Smoking Control in Schools. General Office of Ministry of Education. URL: http://www.law-lib.com/law/law_view.asp?id=317247 [accessed 2023-02-24]
28. Abdullah AS, Hua F, Khan H, Xia X, Bing Q, Tarang K, et al. Secondhand smoke exposure reduction intervention in Chinese households of young children: a randomized controlled trial. Acad Pediatr 2015;15(6):588-598. [doi: 10.1016/j.acap.2015.06.008] [Medline: 26300367]
29. Patel M, Thai CL, Meng Y, Kuo T, Zheng H, Dietsch B, et al. Smoke-free car legislation and student exposure to smoking. Pediatrics 2018 Jan;141(Suppl 1):S40-S50 [FREE Full text] [doi: 10.1542/peds.2017-1026H] [Medline: 29292305]
30. Azagba S, Latham K, Shan L. Exposure to secondhand smoke in vehicles among Canadian adolescents: Years after the adoption of smoke-free car laws. Addict Behav Rep 2019 Dec;10:100215 [FREE Full text] [doi: 10.1016/j.abrep.2019.100215] [Medline: 31508478]
31. Feng G, Jiang Y, Zhao L, Meng G, Wu C, Quah AC, et al. Degree of exposure to secondhand smoking and related knowledge, attitude among adults in urban China. Zhonghua Liu Xing Bing Xue Za Zhi 2014 Sep;35(9):998-1001. [Medline: 25492139]
32. Brener ND, Collins JL, Kann L, Warren CW, Williams BI. Reliability of the Youth Risk Behavior Survey Questionnaire. Am J Epidemiol 1995 Mar 15;141(6):575-580. [doi: 10.1093/oxfordjournals.aje.a117473] [Medline: 7900725]
33. Johnson-Kozlow M, Wahlgren DR, Hovell MF, Flores DM, Liles S, Hofstetter CR, et al. Adolescents validly report their exposure to secondhand smoke. J Clin Epidemiol 2010 Aug;63(8):914-919 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.11.015] [Medline: 20346628]
34. Bovet P, Viswanathan B, Faeh D, Warren W. Comparison of smoking, drinking, and marijuana use between students present or absent on the day of a school-based survey. J Sch Health 2006 Apr;76(4):133-137. [doi: 10.1111/j.1746-1561.2006.00081.x] [Medline: 16536852]

Abbreviations

FCTC: Framework Convention on Tobacco Control
GDP: gross domestic product
GYTS: Global Youth Tobacco Survey
NYTS: China National Youth Tobacco Survey
OR: odds ratio
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 05.07.22; peer-reviewed by D Ashley, D KaradoĞan, X Wu, Y Chen; comments to author 09.01.23; revised version received 11.02.23; accepted 14.02.23; published 24.03.23.

Please cite as:
Ma C, Huang Y, Li S, Zhao M, Zeng X, Di X, Magnussen CG, Xi B, Liu S
Trends in Exposure to Secondhand Smoke Among Adolescents in China From 2013-2014 to 2019: Two Repeated National Cross-sectional Surveys
JMIR Public Health Surveill 2023;9:e40782
URL: https://publichealth.jmir.org/2023/1/e40782
doi:10.2196/40782
PMID:36961497

Original Paper

Periodic Characteristics of Hepatitis Virus Infections From 2013 to 2020 and Their Association With Meteorological Factors in Guangdong, China: Surveillance Study

Xixi Zhao[1,2*], PhD; Meijia Li[3*], MPhil; Naem Haihambo[3], MSc; Xinni Wang[4], BSc; Bin Wang[1,2], PhD; Meirong Sun[5], PhD; Mingrou Guo[6], MPhil; Chuanliang Han[6,7,8,9], PhD

[1]The National Clinical Research Center for Mental Disorders and Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China

[2]Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China

[3]Faculty of Psychology and Center for Neuroscience, Vrije Universiteit Brussel, Brussel, Belgium

[4]Faculty of Psychology, Beijing Normal University, Beijing, China

[5]School of Psychology, Beijing Sport University, Beijing, China

[6]The Brain Cognition and Brain Disease Institute, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[7]The Guangdong Provincial Key Laboratory of Brain Connectome and Behavior, Shenzhen, China

[8]Chinese Academy of Sciences Key Laboratory of Brain Connectome and Manipulation, Shenzhen, China

[9]Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Chuanliang Han, PhD
The Brain Cognition and Brain Disease Institute
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
1068 Xueyuan Avenue Shenzhen University Town, Nanshan District
Shenzhen, 518055
China
Phone: 86 18800129802
Email: hanchuanliang2014@163.com

Abstract

Background: In the past few decades, liver disease has gradually become one of the major causes of death and illness worldwide. Hepatitis is one of the most common liver diseases in China. There have been intermittent and epidemic outbreaks of hepatitis worldwide, with a tendency toward cyclical recurrences. This periodicity poses challenges to epidemic prevention and control.

Objective: In this study, we aimed to investigate the relationship between the periodic characteristics of the hepatitis epidemic and local meteorological elements in Guangdong, China, which is a representative province with the largest population and gross domestic product in China.

Methods: Time series data sets from January 2013 to December 2020 for 4 notifiable infectious diseases caused by hepatitis viruses (ie, hepatitis A, B, C, and E viruses) and monthly data of meteorological elements (ie, temperature, precipitation, and humidity) were used in this study. Power spectrum analysis was conducted on time series data, and correlation and regression analyses were performed to assess the relationship between the epidemics and meteorological elements.

Results: The 4 hepatitis epidemics showed clear periodic phenomena in the 8-year data set in connection with meteorological elements. Based on the correlation analysis, temperature demonstrated the strongest correlation with hepatitis A, B, and C epidemics, while humidity was most significantly associated with the hepatitis E epidemic. Regression analysis revealed a positive and significant coefficient between temperature and hepatitis A, B, and C epidemics in Guangdong, while humidity had a strong and significant association with the hepatitis E epidemic, and its relationship with temperature was relatively weak.

Conclusions: These findings provide a better understanding of the mechanisms underlying different hepatitis epidemics and their connection to meteorological factors. This understanding can help guide local governments in predicting and preparing for future epidemics based on weather patterns and potentially aid in the development of effective prevention measures and policies.

(JMIR Public Health Surveill 2023;9:e45199) doi:10.2196/45199

KEYWORDS
hepatitis virus; meteorological factors; epidemics; recrudescence; public health

Introduction

Hepatitis, which describes an inflammation of the liver, is one of the most common diseases in the world. Hepatitis can be caused by a variety of factors. However, viruses are the most prevalent cause. Hepatitis A [1-4], B [5-9], C [10-14], and E [1,15-18] are the 4 most common types of hepatitis in China, and they are caused by the hepatitis A [19,20], B [21], C [22,23], and E viruses (HAV, HBV, HCV, and HEV), respectively. The hepatitis viruses can survive for a long time in a variety of harsh environments and have a strong resistance to external factors [24]. Hepatitis A and E are transmitted through the digestive tract via fecal-oral transmission and contaminated food or water [25,26]. Hepatitis B and C are spread through the blood or bodily fluids of an infected person, mother-to-child transmission, and sexual transmission [8,27-29]. In 2020, the number of confirmed cases of hepatitis A, B, C, and E in Guangdong Province was 1594, 18,4524, 28,068, and 2070, respectively. This represents an increase of 1.10 times, 1.15 times, 1.23 times, and a decrease of 0.75 times, respectively, compared to statistics from 2013. Hepatitis A, B, and C showed an upward trend, whereas hepatitis E showed a downward trend.

Viruses can be transmitted in 2 ways: horizontal and vertical transmission. Horizontal transmission involves the transmission of a virus from one person to another within the same generation, while vertical transmission refers to the transmission of viruses from mother to child. Hepatitis infections often show an age-dependent property, with more severe symptoms observed in adults [19]. People exposed to HBV through horizontal transmission often develop self-limited acute infections, while people who contract the virus from their mothers through vertical transmission tend to become chronic HBV carriers [21]. This difference in infection severity may be due to the difference in immune response between adults and young children [21]. The infection of hepatitis also showed genotype-dependent properties. For example, in HEV, there are 4 genotypes (HEV 1-4) [30]. The prevalence of HEV1 and HEV2 in developing regions have led to large-scale outbreaks. If infection occurs during pregnancy, it leads to severe hepatitis. In contrast, HEV3 and HEV4 are mainly observed in developed regions and are common in both humans and animals, with pigs as the main host [15].

Recrudescence at a fixed frequency is a common feature of infectious diseases throughout the world [31-34]. Previous research has shown that hepatitis epidemics in China exhibit oscillatory properties at a national level [35,36]. Natural forces (eg, temperature [37] and natural disasters [38]) might drive this oscillatory infection. Several examples have shown that epidemics, such as measles [39], pertussis [40-44], influenza [45], and rabies [46-48], can have oscillations. However, whether the oscillatory phenomena of hepatitis epidemics were influenced by environmental factors remains unclear.

In light of the above, in this study, we aimed to explore the oscillatory properties of the hepatitis epidemic in Guangdong (20°13'-25°31' N and 109°39'-117°19' E) as well as the potential natural contributors to the oscillatory outbreaks. We obtained data for the period from January 2013 to December 2020 on 4 notifiable infectious diseases caused by hepatitis viruses (including HAV, HBV, HCV, and HEV) in Guangdong and the meteorological elements (eg, temperature, precipitation, and humidity) in the same time frame. Power spectrum analysis was conducted on these data to capture the oscillatory strength of the outbreaks. We then explored the relationship between meteorological elements and oscillatory properties based on the regression and correlation analyses.

Methods

Data and Sources

Time series data on available monthly reported and confirmed cases of 4 hepatitis diseases (A, B, C, and E) were obtained for Guangdong province in China's mainland, from January 2013 to December 2020, from the Health Commission of Guangdong. The data set is available to the public around the world and is reported monthly. Monthly reported data of meteorological elements (eg, temperature, precipitation, and humidity) of Guangdong province from January 2013 to December 2020 were obtained from the China Statistical Yearbook 2014-2021. The meteorological factors included in this study are continuous values and vary with time, which would directly reflect the actual natural conditions each month. This data set is also available to the public around the world and is reported annually.

Ethical Considerations

For this study, the data we used are open to the public. Our study did not involve any interventions in human participants. This study was approved by the ethics committee of Beijing Sport University, China (2022142H).

Power Spectrum Analysis

We used spectrum analysis to quantify fluctuations and the recurrence of epidemics. Similar methods have been used in classic and modern studies in the field of public health [35,36,49], medicine [50-53], and biological research [54,55]. The power spectral density for each infectious disease was computed via the multitaper method using the Chronux toolbox [56], an open-source data analysis toolbox [57]. The definition of the peak in the power spectrum is the ratio of the power in a specific frequency (eg, once or twice a year) and the power in the surrounding frequencies (0.75-1.25 and 1.75-2.25). In this study, a ratio larger than 3 was considered an obvious oscillatory peak.

Tuning Curves for Monthly Infected Cases

The tuning curve of monthly infected cases depicts the basic character of disease outbreaks, providing a direct view of the situation each month based on the historical data. We took the monthly average number of infected cases of each hepatitis epidemic and computed them into a tuning curve (equation 1). Each type of hepatitis epidemic in this study has a tuning curve, and the periodic pattern within a year would be obvious based on it.

In this equation, N represents the number of the year.

Tuning Curves for Meteorological Factors

The tuning curve of meteorological factors (eg, temperature, precipitation, and humidity) depicts the basic character of natural conditions, providing a direct view of the situation each month based on historical data. We computed the monthly average of meteorological factors into a tuning curve (equation 2). Each meteorological factor in this study has a tuning curve, and the periodic pattern within a year is obvious based on it.

In this equation, N represents the number of the year.

Regression Analysis

The regression model was shown as equation 3.

In this equation, β represents the regression coefficients of the model, and *y(t)* represents the number of cases with hepatitis infection. The regression analysis was conducted using the regress function in MATLAB (2020a) [58].

Correlation Analysis

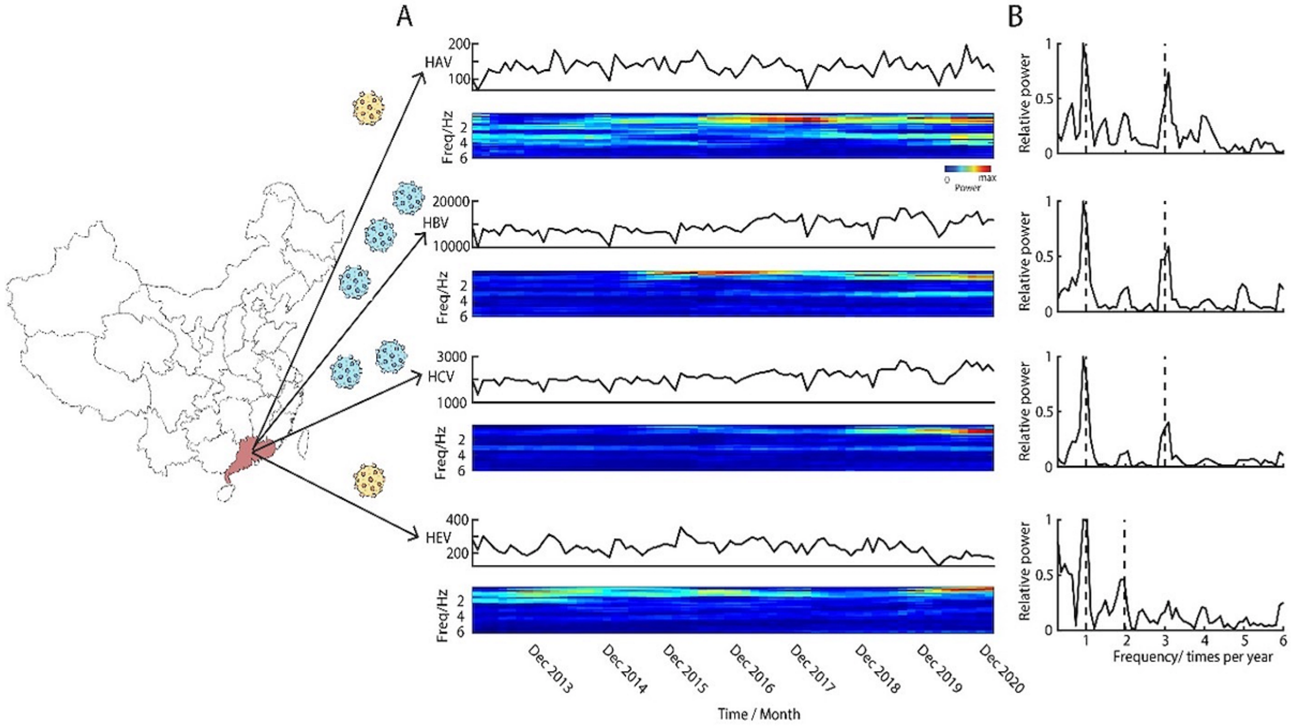
We used the Pearson correlation to measure the relationship between infected cases and meteorological elements. The correlation analysis was conducted using the corr function in MATLAB (2020a).

Results

Periodic Phenomena of Hepatitis Epidemics in Guangdong, China

This study analyzed the monthly data of confirmed cases of 4 hepatitis diseases in Guangdong from January 2013 to December 2020 (Figure 1). Time series data of the 4 hepatitis epidemics (A, B, C, and E) in Guangdong is shown in Figure 1A. It is clear that all hepatitis epidemics in Guangdong have obvious oscillatory patterns based on their power spectrum (Figure 1B), which shows different peaks (eg, peaking once, twice, or 3 times a year). However, for the oscillatory patterns, the hepatitis E epidemic (the fourth row in Figure 1B) differs from the other 3 hepatitis epidemics (first 3 rows in Figure 1B). Hepatitis A, B, and C epidemics show 2 main peaks (occurring once and 3 times per year) in the power spectrums, whereas the hepatitis E epidemic, in addition to the cycle of outbreaks once a year, will have another cycle of outbreaks twice a year.

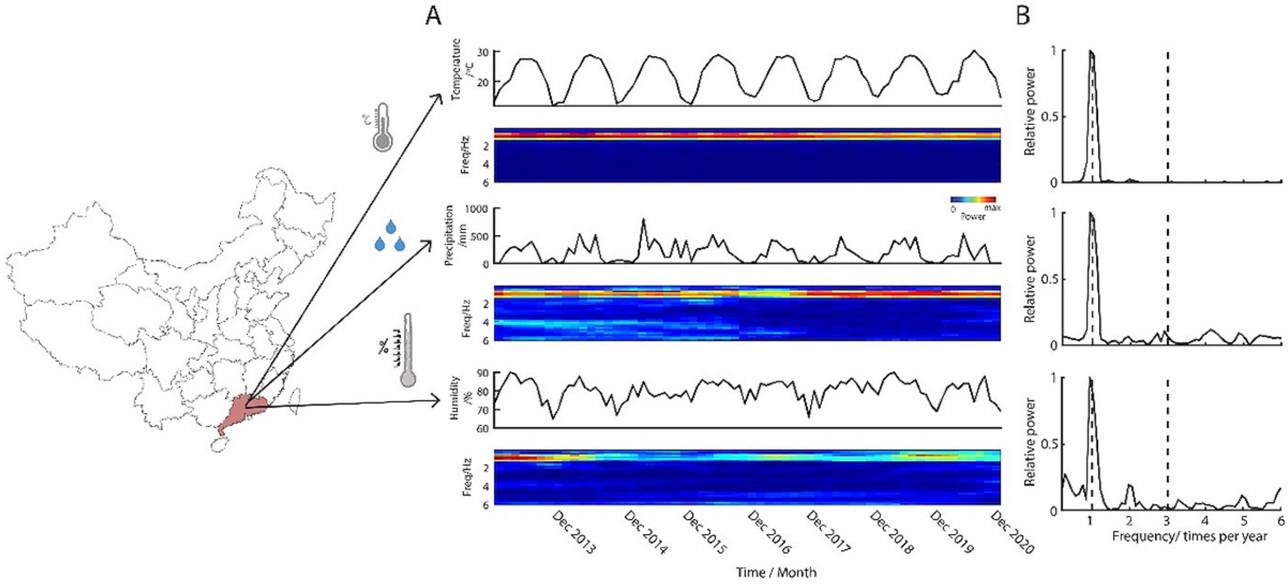
Figure 1. Periodic phenomena of hepatitis epidemics with power spectrum. Left map shows the geological location of Guangdong. (A) shows the monthly incidences of hepatitis A, B, C, and E viruses (HAV, HBV, HCV, and HEV) and its spectrogram from January 2013 to December 2020 in Guangdong. (B) The power spectrum of time series data shown in panel A. The y-axis is the relative power, which is defined as power(f)/max(power(f)).



Periodic Phenomena of Meteorological Elements in Guangdong, China

Meteorological elements might be a potential contributor to the periodic phenomenon of the epidemic; therefore, we also conducted the power spectrum analysis of the time series data of temperature, precipitation, and humidity in Guangdong province (Figure 2A and Figure 2B). All 3 meteorological elements in Guangdong exhibit evident oscillatory patterns throughout time, indicating that the epidemics and meteorological elements may be correlated.

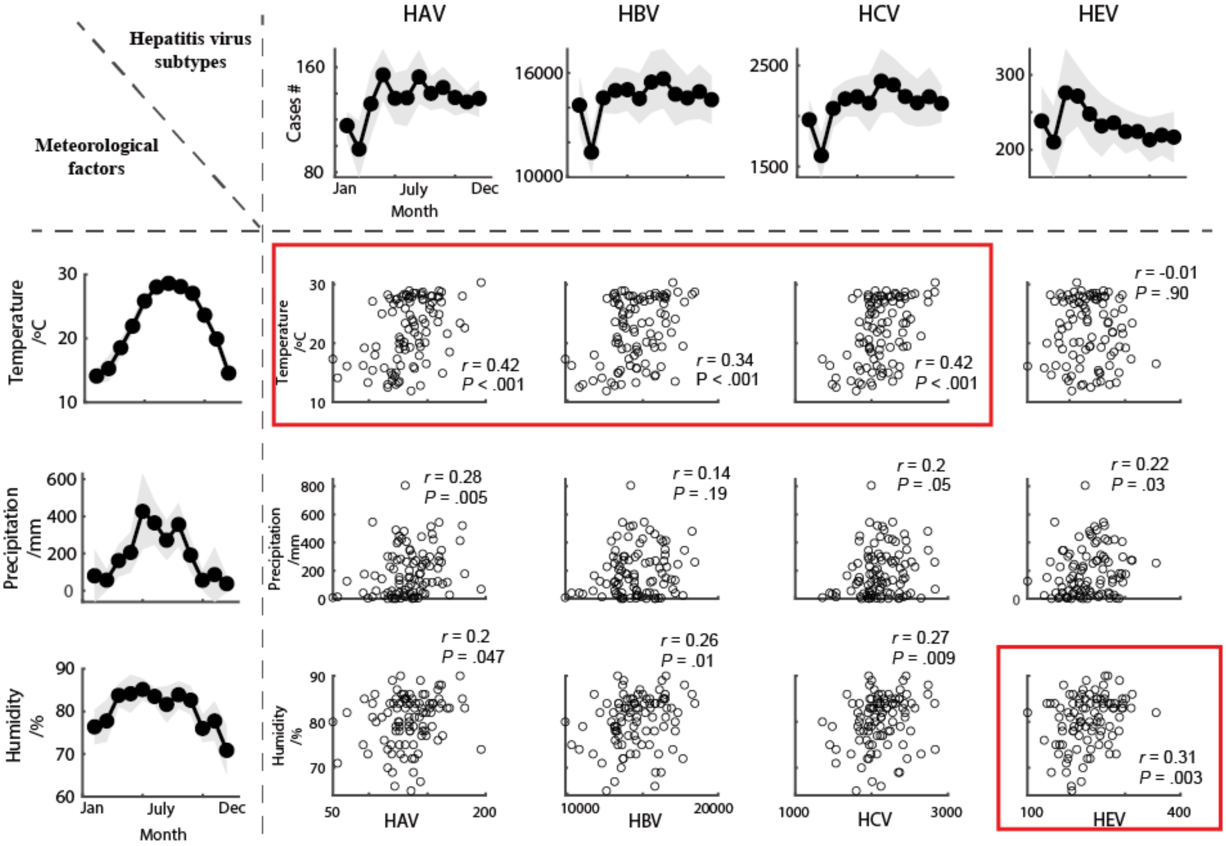
Figure 2. Meteorological elements of hepatitis epidemics with power spectrum. Left map shows the geological location of Guangdong. (A) The monthly time series meteorological elements (temperature, precipitation, and humidity) and its spectrogram from January 2013 to December 2020 in Guangdong. (B) The power spectrum of time series data shown in panel A. The y-axis is the relative power, which is defined as power(f)/max(power(f)).



Relationship Between the Periodic Characteristics of Hepatitis Infection and Meteorological Elements in Guangdong, China

From the observation of the average infected cases of the hepatitis epidemic (first row in Figure 3) and the average meteorological elements (first column in Figure 3), hepatitis infections may be influenced by several natural factors. We then investigated the relationship between hepatitis epidemics and meteorological elements in Guangdong in detail based on multiple analysis approaches.

Figure 3. Relationship between infected cases of hepatitis epidemic and meteorological elements. Top row: the tunning curves for the number of infected cases of 4 hepatitis epidemics over a year, with the grey area denoting the standard error of mean. Left column: the tunning curves for an index of natural factors for each month of each year. Rows 2-4 show scatter plots between the number of cases of each hepatitis type and temperature, precipitation, and humidity, respectively. Each column represents a different type of hepatitis. The red square in each plot indicates conditions that showed significant results in both correlation and regression analyses. HAV: hepatitis A virus; HBV: hepatitis B virus; HCV: hepatitis C virus; HEV: hepatitis E virus.



From the correlation analysis, we discovered a substantial positive association between temperature and the incidence of HAV (*P*<.001), HBV (*P*<.001), and HCV (*P*<.001), showing that the warmer the temperature, the more infected patients there are. Precipitation was positively correlated with HAV (*P*=.005) and HEV (*P*=.03), indicating that the more rainfall there is, the higher the incidence of HAV and HEV. Humidity was positively correlated with HAV (*P*=.047), HBV (*P*=.01), HCV (*P*=.009), and HEV (*P*=.003), showing that the more humid the environment is, the greater the prevalence of hepatitis A, B, C, and E (Figure 3). No additional significant findings were identified. The correlation analysis revealed that temperature had the most significant correlation with HAV, HBV, and HCV, while humidity had the most significant correlation with HEV (depicted by the red square in Figure 3). Since humidity is often affected by precipitation, to further validate their relationship and exclude the influence of other factors (eg, precipitation), we built 2 regression models to investigate how meteorological factors predict the number of hepatitis infections, with and without precipitation as a variable.

From the regression analysis between the number of hepatitis infections and meteorological factors (Table 1), we first checked collinearity by variance inflation factor.

All the variance inflation factors are in the range from 1.5 to 2 (ie, 1.64, 1.93, and 1.78). The correlation coefficients among the 3 meteorological factors (ie, temperature, precipitation, and humidity) were lower than 0.8 (temperature and precipitation: 0.59; temperature and humidity: 0.54; and precipitation and humidity: 0.63, respectively), which would not be affected by collinearity. In the first regression model, using 3 meteorological factors to predict the number of hepatitis infections, the coefficients of temperature in the prediction of hepatitis A, B, and C were significantly positive (hepatitis A: *F*=6.995; *P*<.001, hepatitis B: *F*=5.428; *P*=.001, and hepatitis C: *F*=8.66; *P*<.001) but not significant for the coefficients of precipitation and humidity, suggesting that temperature can positively predict the incidence of hepatitis A, B, and C and is the main contributor to these 3 epidemics. In terms of the hepatitis E epidemic (*F*=4.74; *P*=.004), we found a significant positive coefficient for humidity (*P*=.004) and a weak but significant negative coefficient for temperature (*P*=.03), indicating that a more humid and less warm environment can predict higher incidences of hepatitis E. The second regression model that excluded precipitation as a predictor showed a result that was consistent with the model with 3 predictors, except that the significance level was more robust for humidity to predict HEV infections (Table 2).

In sum, we found that warmer temperatures can predict a higher prevalence of hepatitis A, B, and C in Guangdong. Meanwhile, higher humidity and lower temperature can both predict a higher prevalence of hepatitis E.

Table 1. Statistics for the regression analysis (temperature, precipitation, and humidity).

Regression anlysis	Hepatitis A virus		Hepatitis B virus		Hepatitis C virus		Hepatitis E virus	
	Coefficient	*t* value	Coefficient	*t* value	Coefficient	*t* value	Coefficient	*t* value
β_0	106.94	—[a]	10320.51	—	1425.52	—	49.96	—
$\beta_{1(\text{temperature})}$	1.809	3.71[b]	127.78	3.29[c]	27.59	4.41[b]	–2.05	2.2[d]
$\beta_{2(\text{precipitation})}$	0.001	0.04	–1.04	0.77	–0.22	1.02	0.025	0.76
$\beta_{3(\text{humidity})}$	–0.155	0.31	20.01	0.51	1.57	0.24	2.80	2.93[c]

[a] Not applicable.

[b] *P*<.001.

[c] *P*<.01 ($\beta_{1(\text{temperature})}$ for HBV: *P*=.001; $\beta_{3(\text{humidity})}$ for HEV: *P*=.004).

[d] *P*<.05 (*P*=.03)

Table 2. Statistics for the regression analysis (temperature and humidity).

Regression anlysis	Hepatitis A virus		Hepatitis B virus		Hepatitis C virus		Hepatitis E virus	
	Coefficient	*t* value	Coefficient	*t* value	Coefficient	*t* value	Coefficient	*t* value
β_0	106.23	—[a]	11489.05	—	1675.11	—	21.88	—
$\beta_{1(\text{temperature})}$	1.816	4.06[b]	116.365	3.25[c]	25.15	4.34[b]	–1.78	–2.03[d]
$\beta_{2(\text{humidity})}$	–0.146	–0.34	6.113	0.18	–1.39	–0.25	3.13	3.70[b]

[a] Not applicable.

[b] *P*<.001.

[c] *P*<.01 (*P*=.001).

[d] *P*<.05 (*P*=.04).

Discussion

Principal Findings

In this study, we found that hepatitis epidemics (A, B, C, and E) have different oscillatory properties, and hepatitis A, B, and C in Guangdong have a stronger association with temperature, while hepatitis E showed a stronger association with humidity. Different types of hepatitis will have different periodic characteristics and different associations with different natural environments, which is crucial for epidemic prevention and management. Understanding the temporal characteristics of infectious diseases is essential for effective epidemic prevention, as it may inform the development and implementation of appropriate policies and strategies.

Comparison With Prior Work

To our knowledge, this is the first study to investigate the periodic characteristics of the hepatitis epidemics and their relationship with natural factors, which fills a gap in this field. Previous works have mainly focused on descriptive statistics, detailing the number of infections without describing the precise time characteristics [1,3,18], such as oscillations. Some studies showed that hepatitis epidemics have strong oscillations at a national level [36], but they have not been refined to a more detailed geographical scale (eg, at province levels).

Another issue is the influence of meteorological factors on hepatitis epidemics. HAV and HEV are 2 distinct viruses. Although HAV and HEV share similarities in their fecal-oral transmission route [25,26], our results demonstrate that hepatitis A and E epidemics have different relationships with meteorological factors (hepatitis A epidemic being associated with temperature and hepatitis E epidemic being associated with humidity). This result indicates that even though the 2 viruses share similar transmission modes, the relationship with natural indicators is not necessarily the same, which further indicates that they have unique transmission mechanisms that are not shared, such as host ranges and genotypes.

First, the host range of HAV is limited to humans and several nonhuman primates [59]. There is a tendency for people to engage in more social activities during warmer weather, which increases the likelihood of HAV transmission. In contrast, HEV infects a variety of animal species, including deer, rabbits, and mollusks, but it primarily infects pigs [15,30,60]. Higher humidity is harmful to the health of pigs [61], which will increase contact between pigs and people employed in occupations related to pigs, leading to an increase in HEV infections among the animals and humans who come into contact with an infected animal. Additionally, runoff from outdoor pig farms can contaminate surface water as well as crops that receive this surface water [60]. Higher humidity can increase the risk of HEV contamination of surface water from outdoor pig farms and other sources, and heavy rainfall can lead to the spread of contaminated water to crops and other areas, potentially increasing the incidence of HEV infections. For these reasons, it might be of general value to further investigate water contamination during periods of high rainfall to better understand this phenomenon. Nonetheless, it remains important to implement measures to prevent contamination of water sources and monitor water quality, especially in areas with high levels of outdoor pig farming or other potential sources of contamination.

Second, there are 4 main genotypes of HEV. Genotypes 1 and 2 cause outbreaks or epidemics in humans [62], while genotypes 3 and 4 primarily infect various mammalian species, such as pigs and sheep, but can also infect humans and lead to sporadic cases of hepatitis E. Climate change can affect the environment [63], potentially impacting the quality of water and food sources. For example, the outbreak of HEV genotype 4 infection is more likely to be caused by contaminated tap water rather than contaminated food and contaminated water network on days with heavy rainfall [64]. However, this is not the case for HAV.

Further research is needed to confirm these hypotheses using detailed biological experiments. Previous studies on how temperature affects viral transmission have been equivocal. For example, norovirus prevalence was associated with low water temperatures [65,66], but enterovirus 71 [67] infections began to rise at temperatures above 13°C and declined at temperatures higher than approximately 26°C, exhibiting a V-shaped relationship with the temperature. This phenomenon indicates that transmission patterns are likely influenced by a combination of biological, environmental, and behavioral factors.

Limitations

One limitation of our study is the limited data. Despite being a representative province, our study only includes a single province in China. In the future, we will further obtain more data and apply similar methods to assess infectious diseases in more regions with varying climates. Even though we have made explorations, several factors remain unclear and cannot be addressed at present. For example, hepatitis A and E have similar propagation principles, but their oscillatory strength is very different in Guangdong Province. In the future, higher-dimensional data will be required to provide more clarity. Another potential limitation is that we did not explore the interaction of the variables and the potential nonlinear relationships in the regression model, which should be considered in other studies in the future, where necessary.

Conclusions

Our study makes a link between climate change and the recurrence of infectious disease epidemics, enabling us to predict the magnitude of the epidemics based on various weather conditions. It is essential for local meteorological and medical departments to collaborate closely to prepare for the prevention and control of various epidemics. Additionally, the general public should be informed about the impact of different weather conditions on the spread of disease so that they can take appropriate measures to reduce the spread of viruses.

Acknowledgments

This work was funded by the China Postdoctoral Science Foundation (2022M723299), the Beijing Municipal Hospital Clinical Technology Innovation and Research Plan (XMLX201805), the Beijing Municipal Hospital Research and Development Project (PX2021068), the Advanced Innovation Center for Human Brain Protection Project (3500-12020137), and the Guangdong Province Basic Research Grant (2019A1515110870).

Data Availability

The data set is available to the public, which can be found on the official website of the Guangdong Provincial Health Commission [68].

Authors' Contributions

CH, ML, and XZ conceived and designed the study. CH, ML, NH, BW, MS, and XW contributed to the literature search. CH and ML contributed to data collection. CH and ML contributed to the data analysis and the interpretation of results. All authors contributed to writing the paper.

Conflicts of Interest

None declared.

References

1. Ren X, Wu P, Wang L, Geng M, Zeng L, Zhang J, et al. Changing Epidemiology of Hepatitis A and Hepatitis E Viruses in China, 1990-2014. Emerg Infect Dis 2017 Feb;23(2):276-279 [FREE Full text] [doi: 10.3201/2302.161095] [Medline: 28098527]
2. Li Y, Zhang B, Zhang X, Fan S, Fei L, Yang C, et al. Trend in the incidence of hepatitis A in mainland China from 2004 to 2017: a joinpoint regression analysis. BMC Infect Dis 2022 Aug 01;22(1):663 [FREE Full text] [doi: 10.1186/s12879-022-07651-5] [Medline: 35915398]
3. Sun X, Zhang G, Zhou R, Zheng H, Miao N, Yin Z, et al. Changes in the epidemiology of hepatitis A in three socio-economic regions of China, 1990-2017. Infect Dis Poverty 2019 Oct 03;8(1):80 [FREE Full text] [doi: 10.1186/s40249-019-0591-z] [Medline: 31578150]

4. Wang F, Sun X, Wang F, Zheng H, Jia Z, Zhang G, et al. Changing epidemiology of hepatitis A in China: evidence from three national serological surveys and the National Notifiable Disease Reporting System. Hepatology 2021 Apr 04;73(4):1251-1260. [doi: 10.1002/hep.31429] [Medline: 32592242]
5. Wang H, Men P, Xiao Y, Gao P, Lv M, Yuan Q, et al. Hepatitis B infection in the general population of China: a systematic review and meta-analysis. BMC Infect Dis 2019 Sep 18;19(1):811 [FREE Full text] [doi: 10.1186/s12879-019-4428-y] [Medline: 31533643]
6. Yan Y, Su H, Ji Z, Shao Z, Pu Z. Epidemiology of hepatitis B virus infection in China: current status and challenges. J Clin Transl Hepatol 2014 Mar;2(1):15-22 [FREE Full text] [doi: 10.14218/JCTH.2013.00030] [Medline: 26356070]
7. Lu FM, Li T, Liu S, Zhuang H. Epidemiology and prevention of hepatitis B virus infection in China. J Viral Hepat 2010 Mar;17 Suppl 1:4-9. [doi: 10.1111/j.1365-2893.2010.01266.x] [Medline: 20586928]
8. Zhao H. The epidemiology of hepatitis B virus infection in China. J Viral Hepat 2018 Dec;25(12):1624. [doi: 10.1111/jvh.13004] [Medline: 30199589]
9. Chen S, Li J, Wang D, Fung H, Wong L, Zhao L. The hepatitis B epidemic in China should receive more attention. Lancet 2018 Apr 21;391(10130):1572. [doi: 10.1016/S0140-6736(18)30499-9] [Medline: 29695339]
10. Shen M, Xiao Y, Zhou W, Li Z. Global dynamics and applications of an epidemiological model for hepatitis C virus transmission in China. Discrete Dyn Nat Soc 2015;2015:1-13. [doi: 10.1155/2015/543029]
11. Qu L, Shi Y, Chen K, Lu Y, Ren H. The distribution of hepatitis C virus infection in Shanghai, China: a time-spatial study. BMC Infect Dis 2021 Sep 18;21(1):974 [FREE Full text] [doi: 10.1186/s12879-021-06577-8] [Medline: 34536999]
12. Qi Y, Chen Q, Hao F, Wan Z, Guo H, Lu R, et al. Subtype distribution of hepatitis C virus in Jiangsu, China. J Med Virol 2016 Mar;88(3):498-505. [doi: 10.1002/jmv.24356] [Medline: 26288243]
13. Qin Q, Smith MK, Wang L, Su Y, Wang L, Guo W, et al. Hepatitis C virus infection in China: an emerging public health issue. J Viral Hepat 2015 Mar;22(3):238-244. [doi: 10.1111/jvh.12295] [Medline: 25131856]
14. Mei X, Lu H. Prevalence, diagnosis, and treatment of hepatitis C in Mainland China. Glob Health Med 2021 Oct 31;3(5):270-275 [FREE Full text] [doi: 10.35772/ghm.2021.01080] [Medline: 34782868]
15. Chen Y, Gong Q, Wang Q, Wang W, Wei X, Jiang J, et al. Prevalence of hepatitis E virus among swine in China from 2010 to 2019: A systematic review and meta-analysis. Microb Pathog 2021 Jan;150:104687. [doi: 10.1016/j.micpath.2020.104687] [Medline: 33301857]
16. Tang W, Kong D, Wang Y, Liu M, Hu Q. Hepatitis E virus infection in Wuhan, Central China. Arch Virol 2019 Jan;164(1):27-32. [doi: 10.1007/s00705-018-4036-y] [Medline: 30229302]
17. Zhu Y, Si F, Cao D, Yu X, Yu R, Dong S, et al. Molecular epidemiology of hepatitis E virus infections in Shanghai, China. Virol J 2011 Dec 15;8:541 [FREE Full text] [doi: 10.1186/1743-422X-8-541] [Medline: 22168371]
18. Jia Z, Yi Y, Liu J, Cao J, Zhang Y, Tian R, et al. Epidemiology of hepatitis E virus in China: results from the Third National Viral Hepatitis Prevalence Survey, 2005-2006. PLoS One 2014;9(10):e110837 [FREE Full text] [doi: 10.1371/journal.pone.0110837] [Medline: 25360522]
19. Wang M, Feng Z. Mechanisms of hepatocellular injury in hepatitis A. Viruses 2021 May 08;13(5):861 [FREE Full text] [doi: 10.3390/v13050861] [Medline: 34066709]
20. Debing Y, Neyts J, Thibaut HJ. Molecular biology and inhibitors of hepatitis A virus. Med Res Rev 2014 Sep;34(5):895-917 [FREE Full text] [doi: 10.1002/med.21292] [Medline: 23722879]
21. Tsai K, Kuo C, Ou JJ. Mechanisms of hepatitis B virus persistence. Trends Microbiol 2018 Jan;26(1):33-42 [FREE Full text] [doi: 10.1016/j.tim.2017.07.006] [Medline: 28823759]
22. Moriishi K, Matsuura Y. Mechanisms of hepatitis C virus infection. Antivir Chem Chemother 2003 Nov;14(6):285-297 [FREE Full text] [doi: 10.1177/095632020301400601] [Medline: 14968935]
23. Khatun M, Ray RB. Mechanisms underlying hepatitis C virus-associated hepatic fibrosis. Cells 2019 Oct 14;8(10):1249 [FREE Full text] [doi: 10.3390/cells8101249] [Medline: 31615075]
24. Pirtle EC, Beran GW. Virus survival in the environment. Rev Sci Tech 1991 Sep;10(3):733-748 [FREE Full text] [doi: 10.20506/rst.10.3.570] [Medline: 1782426]
25. Cuthbert JA. Hepatitis A: old and new. Clin Microbiol Rev 2001 Jan;14(1):38-58 [FREE Full text] [doi: 10.1128/CMR.14.1.38-58.2001] [Medline: 11148002]
26. Kamar N, Bendall R, Legrand-Abravanel F, Xia N, Ijaz S, Izopet J, et al. Hepatitis E. The Lancet 2012 Jun;379(9835):2477-2488. [doi: 10.1016/s0140-6736(11)61849-7]
27. Brown JL, Carman WF, Thomas HC. The hepatitis B virus. Baillieres Clin Gastroenterol 1990 Sep;4(3):721-747. [doi: 10.1016/0950-3528(90)90059-p] [Medline: 2149522]
28. Rehermann B, Nascimbeni M. Immunology of hepatitis B virus and hepatitis C virus infection. Nat Rev Immunol 2005 Mar;5(3):215-229. [doi: 10.1038/nri1573] [Medline: 15738952]
29. Dusheiko G. Hepatitis C. Medicine 2007 Jan;35(1):43-48. [doi: 10.1053/j.mpmed.2006.10.011]
30. Nimgaonkar I, Ding Q, Schwartz RE, Ploss A. Hepatitis E virus: advances and challenges. Nat Rev Gastroenterol Hepatol 2018 Feb;15(2):96-110. [doi: 10.1038/nrgastro.2017.150] [Medline: 29162935]
31. Gunning CE, Ferrari MJ, Erhardt EB, Wearing HJ. Evidence of cryptic incidence in childhood diseases. Proc Biol Sci 2017 Aug 30;284(1861):284 [FREE Full text] [doi: 10.1098/rspb.2017.1268] [Medline: 28855364]

32. Anderson RM, Grenfell BT, May RM. Oscillatory fluctuations in the incidence of infectious disease and the impact of vaccination: time series analysis. J Hyg (Lond) 1984 Dec;93(3):587-608 [FREE Full text] [doi: 10.1017/s0022172400065177] [Medline: 6512259]
33. Bhattacharyya S, Ferrari MJ, Bjørnstad ON. Species interactions may help explain the erratic periodicity of whooping cough dynamics. Epidemics 2018 Jun;23:64-70 [FREE Full text] [doi: 10.1016/j.epidem.2017.12.005] [Medline: 29306640]
34. Cao Y, Li M, Haihambo N, Wang X, Zhao X, Wang B, et al. Temporal dynamic characteristics of human monkeypox epidemic in 2022 around the world under the COVID-19 pandemic background. Front Public Health 2023;11:1120470 [FREE Full text] [doi: 10.3389/fpubh.2023.1120470] [Medline: 36778555]
35. Zhao X, Li M, Haihambo N, Jin J, Zeng Y, Qiu J, et al. Changes in temporal properties of notifiable infectious disease epidemics in China during the COVID-19 pandemic: population-based surveillance study. JMIR Public Health Surveill 2022 Jun 23;8(6):e35343 [FREE Full text] [doi: 10.2196/35343] [Medline: 35649394]
36. Han C, Li M, Haihambo N, Cao Y, Zhao X. Enlightenment on oscillatory properties of 23 class B notifiable infectious diseases in the mainland of China from 2004 to 2020. PLoS One 2021;16(6):e0252803 [FREE Full text] [doi: 10.1371/journal.pone.0252803] [Medline: 34106977]
37. Tian H, Yu P, Cazelles B, Xu L, Tan H, Yang J, et al. Interannual cycles of Hantaan virus outbreaks at the human-animal interface in Central China are controlled by temperature and rainfall. Proc Natl Acad Sci U S A 2017 Jul 25;114(30):8041-8046 [FREE Full text] [doi: 10.1073/pnas.1701777114] [Medline: 28696305]
38. Suk JE, Vaughan EC, Cook RG, Semenza JC. Natural disasters and infectious disease in Europe: a literature review to identify cascading risk pathways. Eur J Public Health 2020 Oct 01;30(5):928-935 [FREE Full text] [doi: 10.1093/eurpub/ckz111] [Medline: 31169886]
39. Sumi A, Kamo K, Ohtomo N, Kobayashi N. Study of the effect of vaccination on periodic structures of measles epidemics in Japan. Microbiol Immunol 2007;51(9):805-814 [FREE Full text] [doi: 10.1111/j.1348-0421.2007.tb03976.x] [Medline: 17895597]
40. Fine PE, Clarkson JA. Seasonal influences on pertussis. Int J Epidemiol 1986 Jun;15(2):237-247. [doi: 10.1093/ije/15.2.237] [Medline: 3721687]
41. Broutin H, Guégan J, Elguero E, Simondon F, Cazelles B. Large-scale comparative analysis of pertussis population dynamics: periodicity, synchrony, and impact of vaccination. Am J Epidemiol 2005 Jun 15;161(12):1159-1167. [doi: 10.1093/aje/kwi141] [Medline: 15937025]
42. De Greeff SC, Dekkers ALM, Teunis P, Rahamat-Langendoen JC, Mooi FR, De Melker HE. Seasonal patterns in time series of pertussis. Epidemiol Infect 2009 Oct;137(10):1388-1395. [doi: 10.1017/S0950268809002489] [Medline: 19327200]
43. Wearing HJ, Rohani P. Estimating the duration of pertussis immunity using epidemiological signatures. PLoS Pathog 2009 Oct;5(10):e1000647 [FREE Full text] [doi: 10.1371/journal.ppat.1000647] [Medline: 19876392]
44. Hitz DA, Tewald F, Eggers M. Seasonal Bordetella pertussis pattern in the period from 2008 to 2018 in Germany. BMC Infect Dis 2020 Jul 03;20(1):474 [FREE Full text] [doi: 10.1186/s12879-020-05199-w] [Medline: 32620085]
45. Luo T, Sumi A, Zhou D, Kamo K, Yu B, Zhao D, et al. Study on the effect of measles control programmes on periodic structures of disease epidemics in a large Chinese city. Epidemiol Infect 2011 Feb;139(2):257-264. [doi: 10.1017/S0950268810001056] [Medline: 20478086]
46. Ma C, Hao X, Deng H, Wu R, Liu J, Yang Y, et al. Re-emerging of rabies in Shaanxi province, China, from 2009 to 2015. J Med Virol 2017 Sep;89(9):1511-1519. [doi: 10.1002/jmv.24769] [Medline: 28112421]
47. Zhang Y, Vrancken B, Feng Y, Dellicour S, Yang Q, Yang W, et al. Cross-border spread, lineage displacement and evolutionary rate estimation of rabies virus in Yunnan Province, China. Virol J 2017 Jun 03;14(1):102 [FREE Full text] [doi: 10.1186/s12985-017-0769-6] [Medline: 28578663]
48. Tian H, Feng Y, Vrancken B, Cazelles B, Tan H, Gill MS, et al. Transmission dynamics of re-emerging rabies in domestic dogs of rural China. PLoS Pathog 2018 Dec;14(12):e1007392 [FREE Full text] [doi: 10.1371/journal.ppat.1007392] [Medline: 30521641]
49. Cao Y, Li M, Haihambo N, Zhu Y, Zeng Y, Jin J, et al. Oscillatory properties of class C notifiable infectious diseases in China from 2009 to 2021. Front. Public Health 2022 Aug 11;10:1-9. [doi: 10.3389/fpubh.2022.903025]
50. Cao Y, Han C, Peng X, Su Z, Liu G, Xie Y, et al. Correlation between resting theta power and cognitive performance in patients with schizophrenia. Front Hum Neurosci 2022 Apr 22;16:853994 [FREE Full text] [doi: 10.3389/fnhum.2022.853994] [Medline: 35529780]
51. Han C, Wang T, Wu Y, Li H, Wang E, Zhao X, et al. Compensatory mechanism of attention-deficit/hyperactivity disorder recovery in resting state alpha rhythms. Front Comput Neurosci 2022 Sep 7;16:883065 [FREE Full text] [doi: 10.3389/fncom.2022.883065] [Medline: 36157841]
52. Han C, Guo M, Ke X, Zeng L, Li M, Haihambo N, et al. Oscillatory biomarkers of autism: evidence from the innate visual fear evoking paradigm. Cogn Neurodyn 2023 Apr 10;17(2):459-466. [doi: 10.1007/s11571-022-09839-6] [Medline: 37007195]
53. Han C, Zhao X, Li M, Haihambo N, Teng J, Li S, et al. Enhancement of the neural response during 40 Hz auditory entrainment in closed-eye state in human prefrontal region. Cogn Neurodyn 2023 Apr 07;17(2):399-410. [doi: 10.1007/s11571-022-09834-x] [Medline: 37007205]

54. Han C, Wang T, Wu Y, Li Y, Yang Y, Li L, et al. The generation and modulation of distinct gamma oscillations with local, horizontal, and feedback connections in the primary visual cortex: a model study on large-scale networks. Neural Plast 2021;2021:8874516 [FREE Full text] [doi: 10.1155/2021/8874516] [Medline: 33531893]
55. Han C, Wang T, Yang Y, Wu Y, Li Y, Dai W, et al. Multiple gamma rhythms carry distinct spatial frequency information in primary visual cortex. PLoS Biol 2021 Dec 21;19(12):e3001466 [FREE Full text] [doi: 10.1371/journal.pbio.3001466] [Medline: 34932558]
56. Bell B, Percival DB, Walden AT. Calculating Thomson's spectral multitapers by inverse iteration. GCGS 1993 Mar;2(1):119. [doi: 10.2307/1390958]
57. Chronux. URL: http://chronux.org [accessed 2023-05-31]
58. Sobie EA. An introduction to MATLAB. Sci Signal 2011 Sep 13;4(191):tr7 [FREE Full text] [doi: 10.1126/scisignal.2001984] [Medline: 21934110]
59. Balayan MS. Natural hosts of hepatitis A virus. Vaccine 1992;10 Suppl 1:S27-S31. [doi: 10.1016/0264-410x(92)90537-t] [Medline: 1335654]
60. Khuroo MS, Khuroo MS, Khuroo NS. Transmission of hepatitis E virus in developing countries. Viruses 2016 Sep 20;8(9):253 [FREE Full text] [doi: 10.3390/v8090253] [Medline: 27657112]
61. Xiong Y, Meng Q, Gao J, Tang X, Zhang H. Effects of relative humidity on animal health and welfare. Journal of Integrative Agriculture 2017 Aug;16(8):1653-1658. [doi: 10.1016/s2095-3119(16)61532-0]
62. Nelson KE, Labrique AB, Kmush BL. Epidemiology of genotype 1 and 2 gepatitis E virus infections. Cold Spring Harb Perspect Med 2019 Jun 03;9(6):a031732 [FREE Full text] [doi: 10.1101/cshperspect.a031732] [Medline: 29735579]
63. Cheng X, Liu W, Zhang X, Wang M, Bao C, Wu T. Predicting incidence of hepatitis E using machine learning in Jiangsu Province, China. Epidemiol Infect 2022 Jul 28;150:e149 [FREE Full text] [doi: 10.1017/S0950268822001303] [Medline: 35899849]
64. Chen YJ, Cao NX, Xie RH, Ding CX, Chen EF, Zhu HP, et al. Epidemiological investigation of a tap water-mediated hepatitis E virus genotype 4 outbreak in Zhejiang Province, China. Epidemiol Infect 2016 Dec;144(16):3387-3399 [FREE Full text] [doi: 10.1017/S0950268816001898] [Medline: 27546066]
65. Lopman B, Armstrong B, Atchison C, Gray JJ. Host, weather and virological factors drive norovirus epidemiology: time-series analysis of laboratory surveillance data in England and Wales. PLoS One 2009 Aug 24;4(8):e6671 [FREE Full text] [doi: 10.1371/journal.pone.0006671] [Medline: 19701458]
66. Rohayem J. Norovirus seasonality and the potential impact of climate change. Clinical Microbiology and Infection 2009 Jun;15(6):524-527. [doi: 10.1111/j.1469-0691.2009.02846.x]
67. Chang H, Chio C, Su H, Liao C, Lin C, Shau W, et al. The association between enterovirus 71 infections and meteorological parameters in Taiwan. PLoS ONE 2012 Oct 5;7(10):e46845. [doi: 10.1371/journal.pone.0046845]
68. Guangdong Provincial Health Commission. URL: http://wsjkw.gd.gov.cn/ [accessed 2023-05-31]

Abbreviations

HAV: hepatitis A virus
HBV: hepatitis B virus
HCV: hepatitis C virus
HEV: hepatitis E virus

Edited by A Mavragani, T Sanchez; submitted 20.12.22; peer-reviewed by H Yan, M Peng, W Liu; comments to author 01.02.23; revised version received 18.03.23; accepted 28.04.23; published 15.06.23.

Please cite as:
Zhao X, Li M, Haihambo N, Wang X, Wang B, Sun M, Guo M, Han C
Periodic Characteristics of Hepatitis Virus Infections From 2013 to 2020 and Their Association With Meteorological Factors in Guangdong, China: Surveillance Study
JMIR Public Health Surveill 2023;9:e45199
URL: https://publichealth.jmir.org/2023/1/e45199
doi:10.2196/45199
PMID:37318858

Original Paper

Associations Between Hourly Ambient Particulate Matter Air Pollution and Ambulance Emergency Calls: Time-Stratified Case-Crossover Study

Qiang Zhou[1*], MSc; Hanxu Shi[2*], MPhil; Rengyu Wu[1], BSc; Hong Zhu[1], MSc; Chongzhen Qin[1], MSc; Zhisheng Liang[2], MPhil; Shengzhi Sun[3], PhD; Junfeng Zhao[4], PhD; Yasha Wang[5], PhD; Jie Huang[6], PhD; Yinzi Jin[2,7], PhD; Zhijie Zheng[2], PhD; Jingyan Li[8], PhD; Zhenyu Zhang[2,7], PhD

[1]Shenzhen Center for Prehospital Care, Shenzhen, China
[2]Department of Global Health, Peking University School of Public Health, Beijing, China
[3]School of Public Health, Capital Medical University, Beijing, China
[4]School of Computer Science, Peking University, Beijing, China
[5]National Engineering Research Center of Software Engineering, Peking University, Beijing, China
[6]School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen, China
[7]Institute for Global Health and Development, Peking University, Beijing, China
[8]China National Environmental Monitoring Centre, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Zhenyu Zhang, PhD
Department of Global Health
Peking University School of Public Health
Southern University of Science and Technology
Beijing, 100191
China
Phone: 86 13681165303
Email: zzy@pku.edu.cn

Abstract

Background: Associations between short-term exposure to ambient particulate matter (PM) air pollutants and mortality or hospital admissions have been well-documented in previous studies. Less is known about the associations of hourly exposure to PM air pollutants with ambulance emergency calls (AECs) for all causes and specific causes by conducting a case-crossover study. In addition, different patterns of AECs may be attributed to different seasons and daytime or nighttime periods.

Objective: In this study, we quantified the risk of all-cause and cause-specific AECs associated with hourly PM air pollutants between January 1, 2013, and December 31, 2019, in Shenzhen, China. We also examined whether the observed associations of PM air pollutants with AECs for all causes differed across strata defined by sex, age, season, and the time of day.

Methods: We used ambulance emergency dispatch data and environmental data between January 1, 2013, and December 31, 2019, from the Shenzhen Ambulance Emergency Centre and the National Environmental Monitor Station to conduct a time-stratified case-crossover study to estimate the associations of air pollutants (ie, PM with an aerodynamic diameter less than 2.5 µm [$PM_{2.5}$] or 10 µm [PM_{10}]) with all-cause and cause-specific AECs. We generated a well-established, distributed lag nonlinear model for nonlinear concentration response and nonlinear lag-response functions. We used conditional logistic regression to estimate odds ratios with 95% CIs, adjusted for public holidays, season, the time of day, the day of the week, hourly temperature, and hourly humidity, to examine the association of all-cause and cause-specific AECs with hourly air pollutant concentrations.

Results: A total of 3,022,164 patients were identified during the study period in Shenzhen. Each IQR increase in $PM_{2.5}$ (24.0 µg/m^3) and PM_{10} (34.0 µg/m^3) concentrations over 24 hours was associated with an increased risk of AECs ($PM_{2.5}$: all-cause, 1.8%, 95% CI 0.8%-2.4%; PM_{10}: all-cause, 2.0%, 95% CI 1.1%-2.9%). We observed a stronger association of all-cause AECs with $PM_{2.5}$ and PM_{10} in the daytime than in the nighttime ($PM_{2.5}$: daytime, 1.7%, 95% CI 0.5%-3.0%; nighttime, 1.4%, 95% CI 0.3%-2.6%; PM_{10}: daytime, 2.1%, 95% CI 0.9%-3.4%; nighttime, 1.7%, 95% CI 0.6%-2.8%) and in the older group than in the younger group ($PM_{2.5}$: 18-64 years, 1.4%, 95% CI 0.6%-2.1%; ≥65 years, 1.6%, 95% CI 0.6%-2.6%; PM_{10}: 18-64 years, 1.8%, 95% CI 0.9%-2.6%; ≥65 years, 2.0%, 95% CI 1.1%-3.0%).

Conclusions: The risk of all-cause AECs increased consistently with increasing concentrations of PM air pollutants, showing a nearly linear relationship with no apparent thresholds. PM air pollution increase was associated with a higher risk of all-cause AECs and cardiovascular diseases–, respiratory diseases–, and reproductive illnesses–related AECs. The results of this study may be valuable to air pollution attributable to the distribution of emergency resources and consistent air pollution control.

(JMIR Public Health Surveill 2023;9:e47022) doi:10.2196/47022

KEYWORDS

particulate matter air pollution; ambulance emergency calls; AECs; environmental epidemiology; public health; air pollution; environmental data; patient data

Introduction

It is well-known that exposure to ambient particulate matter (PM) causes huge public health burdens due to substantial excess morbidity and mortality [1]. Over the past decade, a growing number of epidemiological studies have examined associations between short-term exposure to air pollutants and the incidence of cardiovascular disease (ie, acute coronary syndrome and myocardial infarction) [2-9]. China's government has established several policies to control air pollution with great effort, and the severity of air pollution has been eliminated to some extent. However, with continued urbanization and climate change, air pollution still requires focused efforts [10,11]. Therefore, the burden of disease associated with air pollutants is substantial and should be the focus of further research.

Although previous studies have documented the effects of short-term daily air pollution exposure on cardiovascular diseases, less is known about the effects of hourly air pollution exposure using a case-crossover study design. Moreover, most comprehensive studies on air pollutants' health effects have focused on cardiovascular or respiratory diseases–related mortality or hospital admissions, with limited evidence on assessing all-cause health impacts on ambulance emergency calls (AECs). Di et al [12] conducted a study that included 22,433,862 case-days and 76,143,209 control-days and found that each short-term increase of 10 $\mu g/m^3$ was statistically significantly associated with a relative increase of 1.05% (95% CI 0.95%-1.15%) in the daily mortality rate from 2000 to 2012. Ito et al [13] conducted a time-series study that observed a positive association between PM with an aerodynamic diameter less than 2.5 µm [$PM_{2.5}$], daily deaths, and emergency hospitalizations for cardiovascular diseases in New York City. AECs may serve as a more sensitive indicator of the health effects of air pollutants and a more appropriate signal for syndromic surveillance [14,15]. In addition, different patterns of AECs may be attributed to different seasons and daytime or nighttime periods. Identifying these potential patterns can be useful in helping people avoid exposure, but more evidence is needed. Although a few previous studies investigated hourly PM air pollutants' association with disease onset, limitations such as a small sample size with a short period, the lack of all-cause disease as a variable, and only focusing on $PM_{2.5}$ should not be ignored. For example, one Canadian study that only included 500,302 individuals found that exposure to elevated hourly $PM_{2.5}$ during wildfire seasons was associated with increased odds of dispatches related to respiratory and cardiovascular conditions over 5 years [16]. Therefore, it is imperative to conduct a time-stratified case-crossover study that includes a relatively large sample size with a long study period to investigate the associations of PM air pollutants with all-cause and cause-specific AECs. In this study, we quantified the risk of all-cause and cause-specific AECs associated with hourly air pollutants (ie, $PM_{2.5}$ and PM with an aerodynamic diameter less than 10 µm [PM_{10}]) between January 1, 2013, and December 31, 2019, in Shenzhen, China. We also examined whether the observed associations of PM air pollutants with AECs for all causes differed across strata defined by sex, age, season, and the time of day.

Methods

Study Population

Shenzhen is a major city located in the Pearl River Delta region of China, adjacent to Hong Kong. As one of the country's 4 province-level municipalities, Shenzhen has a population of 17.68 million and a gross domestic product per capita of CNY ¥173,700 (US $24,446) as of 2021. The Shenzhen Center for Prehospital Care developed the 120 Emergency Medical Services system in 1994, which includes 73 emergency networks and 103 emergency stations. In 2022, the center received 2,278,232 all-cause emergency calls, an increase of 41.96% from the previous year, and had 295,295 emergency dispatches, an increase of 14.23% from the previous year. Shenzhen has a subtropical monsoon climate that is typical of southern areas of China. We recruited individuals aged >18 years who called the emergency ambulance to reach the hospital, and those with time records linked to transfer were eligible to be included during the period from 2013 to 2019. We used ambulance emergency dispatch data and environmental data between January 1, 2013, and December 31, 2019, from the Shenzhen Ambulance Emergency Centre and the National Environmental Monitor Station, which included patients' self-reported medical claims, the time of AECs, prehospital diagnosis, demographic characteristics, and hourly air pollutant concentrations. Each call in the emergency data set was linked to a patient care report that was completed by the attending paramedics.

Study Design

We used a time-stratified case-crossover study design to estimate the associations of PM pollutants (ie, $PM_{2.5}$ and PM_{10}) with all-cause and cause-specific AECs. Each patient was exposed to PM pollutants before the hour in which the AEC occurred, compared to control periods of the same individual when the AEC did not occur [17,18]. This study design allows for the elimination of potential confounding from all known and unknown time-invariant factors (ie, age, sex, ethnicity, behavioral factors, and socioeconomic status) and covariant factors that vary slowly, such as seasonality [19]. We determined that each case index hour was the hour in which the AEC occurred (ie, case period), and we used 3 or 4 control index hours to match this hour by the same hour of the day, day of the week, month, and year with the case index hour to control for long-term time trends [12]. For example, if the first AEC occurred at 6 AM on Monday, May 20, 2013, we would define 6 AM on Monday, May 20, 2013, as the case index hour and 6 AM on all other Mondays in May 2013 (May 6, 13, and 27) as the control index hours.

Environmental Data Ascertainment

We obtained hourly concentrations of air pollutants (ie, $PM_{2.5}$ and PM_{10}) from the National Environmental Monitor Station during the study period. The station provides real-time data on criteria air pollutants across all nationally controlled monitoring stations operated by the China National Environmental Monitoring Centre with strict standard data quality control procedures. To facilitate communication, we first defined each IQR increase as the difference between the 25th and 75th percentile of air pollutant concentrations [20]. We also defined the extreme concentration (ie, $PM_{2.5}$ and PM_{10}) as the maximum concentration equal to the 97.5th percentile of concentrations over the study period. We calculated a population-weighted average of each IQR concentration increase and the maximum hourly concentration for each hour [21].

AECs Ascertainment

We combined patients' self-reported medical claims and prehospital diagnosis as the final diagnosis. Our primary outcome was all-cause AECs, which we further classified into 3 cause-specific categories: AECs related to cardiovascular diseases, respiratory diseases, and reproductive illnesses. For instance, we classified hypertension, acute ischemic stroke, myocardial infarction, ischemic heart disease, acute coronary syndrome, and unstable angina as cardiovascular diseases–related AECs, whereas calls due to lung diseases, bronchitis, respiratory difficulties, and asthma were categorized as respiratory diseases–related AECs. Similarly, calls related to pregnancy, abortion, or other reproductive disorders were classified as reproductive illnesses–related AECs. Additionally, we created a separate category for AECs due to injury, poisoning, mental health disorders, endocrine disorders, and digestive disorders. We were careful to exclude any vague claims such as dizziness, abdominal pain, discomfort, unconsciousness, headache, and allergy while generating disease stratification. We excluded patients who had both missing self-reported claims and missing prehospital diagnoses. We extracted the relevant demographic information (ie, age, sex, and year) and admission date for each diagnosis. We calculated cause-specific AECs by 24 hours and all-cause AECs by age to reflect the hourly distribution of AECs.

Statistical Analysis

We generated a well-established, distributed lag nonlinear model for nonlinear concentration-response and nonlinear lag-response functions [16]. We modeled the concentration-response function using a natural cubic B-spline with 1 or 2 knots to account for potential nonlinear relationships. We also modeled the lag-response function using a linear function with 1 or 2 knots placed on the log scale of lags up to 48 hours. We used conditional logistic regression to estimate odds ratios (ORs) with 95% CIs, examining the association of all-cause and cause-specific AECs with hourly air pollution concentration. The following covariates were adjusted in our main models: natural spline functions with 3 degrees of freedom for temperature and humidity, public holidays, and the day of the week. We converted ORs with 95% CIs to percentage changes in the risk of all-cause and cause-specific AECs associated with each IQR increase in air pollutant concentrations [20,22,23]. The following equations were used:

$$\text{Percentage change IQR} = (e^{\beta \times \text{IQR}} - 1) \times 100\% \quad \textbf{(1)}$$

$$\text{Lower 95\% CI} = (e^{[\beta - 1.96 \times \text{SE}] \times \text{IQR}} - 1) \times 100\% \quad \textbf{(2)}$$

$$\text{Upper 95\% CI} = (e^{[\beta + 1.96 \times \text{SE}] \times \text{IQR}} - 1) \times 100\% \quad \textbf{(3)}$$

where β is the regression coefficient.

We conducted several stratified analyses by age (18-64 vs ≥65 years), sex (male vs female), season (warm vs cool), and the time of day (daytime vs nighttime) to examine potential effect modifications. The warm season ranged from March to October and the cool season ranged from November to February in Shenzhen. Nighttime is defined as being from 8:00 PM to 7:00 AM in the next day, and daytime is defined as being from 8:00 AM to 7:00 PM within one day [20]. Missing rates of PM air pollutants were less than 0.51%, and we filled in missing data by using the next or previous entry.

We conducted a number of sensitivity analyses to assess the robustness of our results. First, we examined the correlations between several air pollutants to estimate the multicollinearity. Second, we repeated the main analyses based on exposure to the extreme concentration, rather than using only each IQR increase, to observe the robustness of the exposure metric. Third, we used variable key modeling parameters, including modelling the concentration-response functions using a natural B-spline with 3 internal knots for all-cause AECs and 2 or 3 knots for cause-specific AECs and the lag-response function using a natural cubic B-spline with 4 knots for all-cause AECs and 2 or 3 knots for cause-specific AECs placed on the log scale of lags up to 48 hours.

All analyses were performed in R (version 4.2.1; R Foundation for Statistical Computing). The *survival* package was used for conditional logistic regression, and the *dlnm* package was used for the distributed lag nonlinear model.

Ethics Approval

Ethics approval and consent to participate in this project was approved by the Peking University Health Science Center Institutional Review Board (PUIRB-YS2023123). Informed consent was obtained from all participants prior to questionnaire administration.

Results

A total of 3,022,164 patients were identified during the study period in Shenzhen: 64.3% (n=1,942,832) were male, 81.5% (n=2,462,968) were aged 20-64 years, 21.1% (n=636,288) made AECs due to cardiovascular diseases, and 51.4% (n=1,553,480) called the ambulance in the daytime (Table 1). The average hourly concentrations of $PM_{2.5}$ (32.2 $\mu g/m^3$) and PM_{10} (49.6 $\mu g/m^3$) before the index hour were well above the recently updated World Health Organization Global Air Quality Guidelines 2021 [24] (annual average: $PM_{2.5}$, 5.0 $\mu g/m^3$; PM_{10}, 10.0 $\mu g/m^3$; Table S1 in Multimedia Appendix 1). AECs due to cardiovascular diseases rode 2 crests at 9:00 AM and 8:00 PM over 24 hours, the fluctuations of AECs due to respiratory diseases were relatively gentle, and AECs due to reproductive illnesses peaked at 3:00 AM and 11:00 PM (Figure 1). We found significant differences in all-cause and cause-specific AECs between daytime and nighttime, stratified by age (Table 2). For example, the proportion of all-cause AECs was 34.5% (n=128,928) in the ≥65 years group and 49% (n=90,792; <20 years) and 54.1% (n=1,333,684; 20-64 years) in the younger groups at nighttime. We found that all-cause AECs reached their zeniths at 7:00 AM and 9:00 AM in the ≥65 years group, and the 20-64 years group reached their peaks at 6:00 AM and 10:00 PM over 24 hours (Figure S1 in Multimedia Appendix 1).

Table 1. Baseline characteristics of populations in Shenzhen from 2013 to 2019.

Baseline characteristics	Patient (N=3,022,164), n (%)
Age (years)	
<20	185,152 (6.1)
20-64	2,462,968 (81.5)
≥65	373,884 (12.4)
Sex	
Male	1,942,832 (64.3)
Female	1,079,256 (35.7)
Cases by year	
2013	355,816 (11.8)
2014	385,232 (12.7)
2015	399,756 (13.2)
2016	429,948 (14.2)
2017	465,060 (15.4)
2018	475,104 (15.7)
2019	511,248 (17)
Cases by types of AECs[a]	
Cardiovascular	636,288 (21.1)
Respiratory	137,960 (4.6)
Reproductive	208,408 (6.9)
Others[b]	2,039,508 (67.4)
Time of day[c]	
Daytime	1,553,480 (51.4)
Nighttime	1,468,684 (48.6)

[a] AEC: ambulance emergency call.

[b] Others: ambulance emergency calls due to injury, poisoning, mental health disorders, endocrine disorders, and digestive disorders.

[c] Time of day: daytime is from 8:00 AM to 7:00 PM within one day; nighttime is from 8:00 PM to 7:00 AM in the next day.

Figure 1. The distribution of different types of ambulance emergency calls among 3,022,164 patients over 24 hours in Shenzhen from 2013 to 2019.

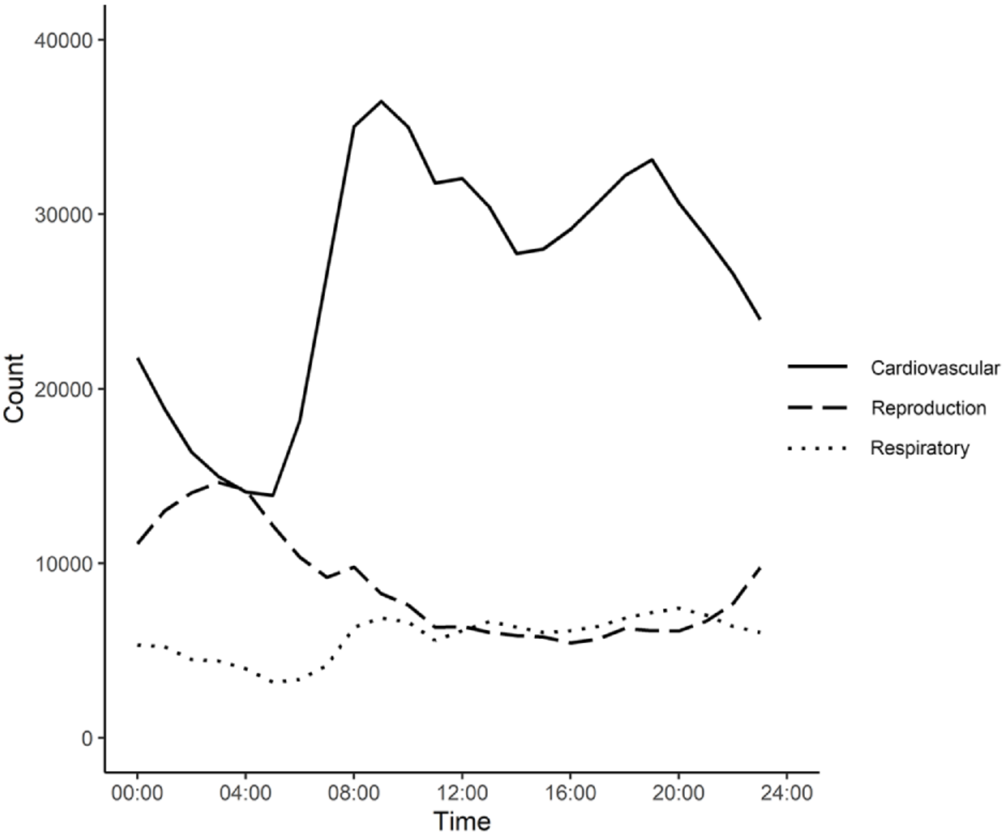


Table 2. Percentage of all types of ambulance emergency calls (AECs) by the time of day and age group in Shenzhen from 2013 to 2019 (N=3,022,164).

AECs	Daytime[a], n (%)	Nighttime[a], n (%)	*P* value
Total population			
All-cause	1,468,684 (48.6)	1,553,480 (51.4)	<.001
Cardiovascular	381,668 (60)	254,620 (40)	<.001
Respiratory	77,072 (55.9)	60,868 (44.1)	<.001
Reproductive	79,484 (38.1)	128,924 (61.9)	<.001
<20 years group			
All-cause	94,360 (51)	90,792 (49)	<.001
Cardiovascular	22,208 (62.7)	13,204 (37.3)	<.001
Respiratory	1,0916 (55)	8920 (45)	<.001
Reproductive	3620 (44.3)	4552 (55.7)	<.001
20-64 years group			
All-cause	1,129,284 (45.9)	1,333,684 (54.1)	<.001
Cardiovascular	257,016 (58.8)	180,048 (41.2)	<.001
Respiratory	35,624 (52.2)	32,628 (47.8)	<.001
Reproductive	72,188 (37.2)	121,716 (62.8)	<.001
≥65 years group			
All-cause	244,956 (65.5)	128,928 (34.5)	<.001
Cardiovascular	102,436 (62.5)	61,360 (37.5)	<.001
Respiratory	30,532 (61.2)	19,320 (38.8)	<.001
Reproductive	—[b]	—	—

[a] Daytime is from 8:00 AM to 7:00 PM within one day; nighttime is from 8:00 PM to 7:00 AM in the next day.

[b] Not applicable.

There were moderate to high correlations among air pollutants (Table S2 in Multimedia Appendix 1). $PM_{2.5}$ was positively correlated with PM_{10}, nitrogen dioxide, ozone, and sulfur dioxide (Spearman *r*=0.93, 0.61, 0.56, and 0.62, respectively), and PM_{10} was positively correlated with $PM_{2.5}$, nitrogen dioxide, ozone, and sulfur dioxide (Spearman *r*=0.93, 0.64, 0.52, and 0.66, respectively).

We found associations of exposure to $PM_{2.5}$ and PM_{10} with a higher incidence of all-cause AECs in the concurrent hour. After that, the associations gradually eliminated and became not statistically significant after approximately 15 to 48 hours (Table 3), so we decided to use the duration of 0 to 24 hours to derive the risk estimates (Figures 2 and 3). Each IQR increase in the concentrations of $PM_{2.5}$ (24.0 μg/m^3) and PM_{10} (34.0 μg/m^3) in 24 hours was associated with a higher risk of AECs ($PM_{2.5}$: all-cause, 1.8%, 95% CI 0.8%-2.4%; cardiovascular diseases, 1.9%, 95% CI 0.2%-3.1%; respiratory diseases, 2.5%, 95%CI –0.9% to 6.1%; reproductive illnesses, 1.9%, 95% CI 1.0%-4.8%; PM_{10}: all-cause, 2.0%, 95% CI 1.1%-2.9%; cardiovascular diseases, 2.0%, 95% CI 0.6%-3.4%; respiratory illnesses, 2.9%, 95% CI –0.6% to 6.6%; reproductive illnesses, 2.1%, 95% CI 0.8%-5.1%; Table 3 and Figures 2 and 3).

Figures 4 and 5 show the concentration-response curves of all-cause and cause-specific AECs for $PM_{2.5}$ and PM_{10}. The risk of all-cause AECs increased consistently with increasing concentrations of PM air pollutants, showing a nearly linear relationship with no apparent thresholds. In general, the risk of all-cause and cause-specific AECs increased dramatically with the accumulated concentration of $PM_{2.5}$ and PM_{10}. The magnitude of the associations varied slightly for the cardiovascular diseases–related AECs, whereas the association had a difference in respiratory diseases– and reproductive illness–related AECs.

In stratified analyses, we observed a stronger association of all-cause AECs with $PM_{2.5}$ and PM_{10} in the daytime than in the nighttime ($PM_{2.5}$: daytime, 1.7%, 95% CI 0.5%-3.0%; nighttime, 1.4%, 95% CI 0.3%-2.6%; PM_{10}: daytime, 2.1%, 95% CI 0.9%-3.4%; nighttime, 1.7%, 95% CI 0.6%-2.8%) and in the older group than the younger group ($PM_{2.5}$: 18-64 years, 1.4%, 95% CI 0.6%-2.1%; ≥65 years, 1.6%, 95% CI 0.6%-2.6%; PM_{10}: 18-64 years, 1.8%, 95% CI 0.9%-2.6%; ≥65 years, 2.0%, 95% CI 1.1%-3.0%; Table 4). For subtypes of AECs, we found stronger associations of respiratory diseases– and reproductive illnesses–related AECs with $PM_{2.5}$ and PM_{10} in the daytime than in the nighttime and weaker associations of cardiovascular diseases–related AECs with $PM_{2.5}$ and PM_{10} in the daytime than in the nighttime (Table 4). Cardiovascular diseases–related

AECs with $PM_{2.5}$ and PM_{10} had stronger association in the older group than in the younger group (Table 4).

A series of sensitivity analyses showed that the results were consistent with each IQR increase when the extreme concentration was considered as the exposure metric (Tables S3 and S4 and Figures S2-S7 in Multimedia Appendix 1). The results were consistent when we adopted an alternative number of knots or degrees of freedom for the distribution of air pollution concentration.

Table 3. Risk of ambulance emergency calls (AECs) associated with each IQR[a] increase in PM2.5[b] and PM10[c] concentration over different lags in Shenzhen from 2013 to 2019. The models were adjusted for public holidays, days of the week, hourly temperature, and hourly humidity.

AECs	0-12 hours, percentage change (95% CI)		0-24 hours, percentage change (95% CI)		0-36 hours, percentage change (95% CI)		0-48 hours, percentage change (95% CI)	
	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}
All-cause	1.7 (0.9 to 2.6)	1.8 (1.0 to 2.6)	1.8 (0.8 to 2.4)	2.0 (1.1 to 2.9)	1.4 (0.5 to 2.4)	1.8 (0.9 to 2.8)	1.3 (0.5 to 2.2)	1.7 (0.9 to 2.6)
Cardiovascular	1.8 (0.3 to 3.2)	1.8 (0.4 to 3.2)	1.9 (0.2 to 3.1)	2.0 (0.6 to 3.4)	1.4 (–0.2 to 3.1)	1.9 (0.3 to 3.6)	1.4 (–0.1 to 2.9)	1.8 (0.3 to 3.3)
Respiratory	2.0 (–1.4 to 5.5)	2.0 (–1.4 to 5.5)	2.5 (–0.9 to 6.1)	2.9 (–0.6 to 6.6)	2.4 (–1.3 to 6.3)	2.2 (–0.6 to 6.7)	1.1 (–1.5 to 5.9)	1.8 (–0.9 to 6.7)
Reproductive	2.3 (0.3 to 5.3)	2.3 (0.1 to 5.3)	1.9 (1.0 to 4.8)	2.1 (0.8 to 5.1)	1.4 (–1.6 to 4.6)	1.9 (–1.2 to 5.1)	2.5 (–0.5 to 5.5)	2.8 (–0.2 to 5.9)

[a] IQR: IQR of $PM_{2.5}$ was defined as the 25th to 75th percentile (24.0 μg/m^3); IQR of PM_{10} was defined as the 25th to 75th percentile (34.0 μg/m^3).

[b] $PM_{2.5}$: particulate matter less than 2.5 μm in diameter.

[c] PM_{10}: particulate matter less than 10 μm in diameter.

Figure 2. Lag structures for the associations of all-cause ambulance emergency calls with each IQR increase in $PM_{2.5}$ and PM_{10} concentration over lags of up to 24 hours in Shenzhen from 2013 to 2019. The overall lag structure curves are calculated using a linear with 2 knots placed on the log scale of lags to model the lag-response association. The solid black lines are the average percentage change in the risk of all-cause ambulance emergency calls with each IQR increase in $PM_{2.5}$ and PM_{10} concentration (24.0 μg/m^3 and 34.0 μg/m^3, respectively), and the dotted lines are the 95% CIs. $PM_{2.5}$ and PM_{10}: particulate matter less than 2.5 and 10 μm in diameter, respectively.

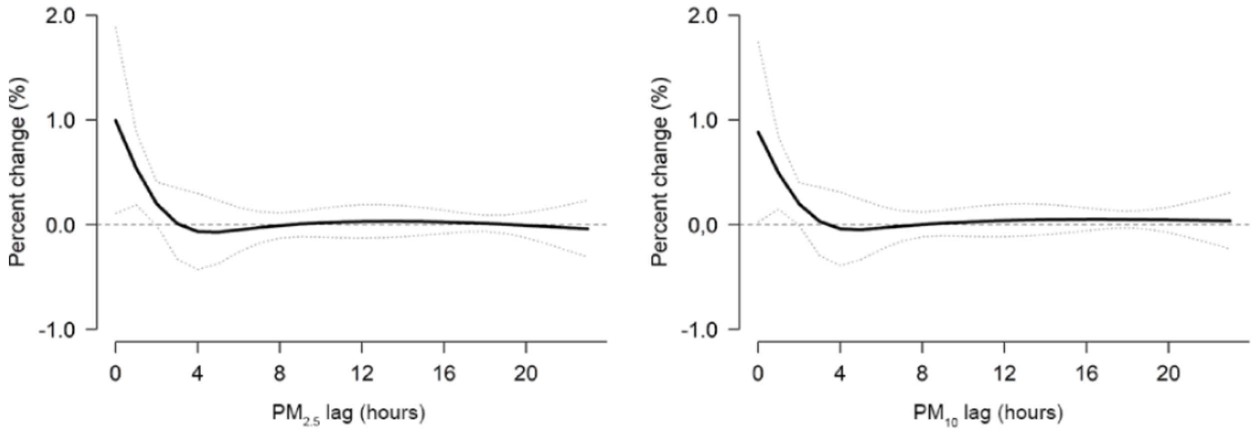


Figure 3. Lag structures for the associations of ambulance emergency calls due to cardiovascular, respiratory, and reproductive diseases with each IQR increase in $PM_{2.5}$ and PM_{10} concentration over lags of up to 24 hours in Shenzhen from 2013 to 2019. A and D represent cardiovascular diseases, B and E represent respiratory diseases, and C and F represent reproductive illnesses. The overall lag structure curves are calculated using a linear with 1 or 2 knots placed on the log scale of lags to model the lag-response association. The black solid lines are the average percentage change in the risk of ambulance emergency calls due to cardiovascular, respiratory, and reproductive diseases with each IQR increase in $PM_{2.5}$ and PM_{10} concentration (24.0 μg/m^3 and 34.0 μg/m^3, respectively), and the dotted lines are the 95% CIs. $PM_{2.5}$ and PM_{10}: particulate matter less than 2.5 and 10 μm in diameter, respectively.

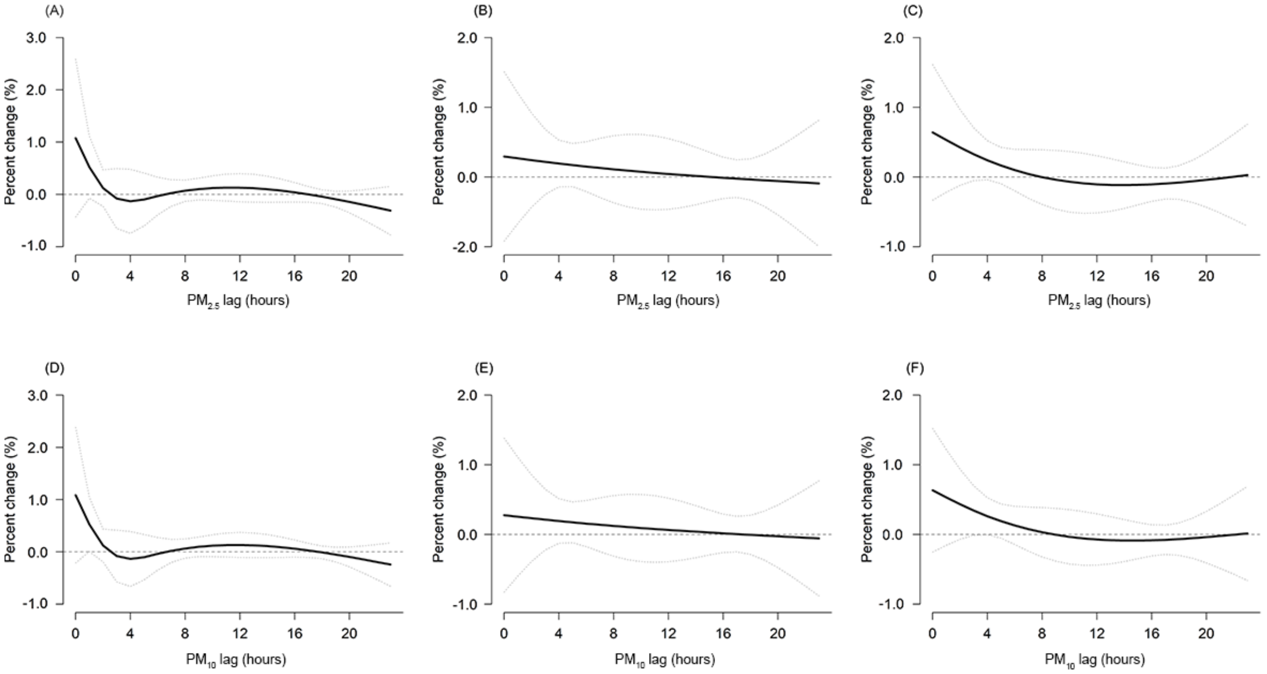


Figure 4. Cumulative concentration-response curves for the association of all-cause ambulance emergency calls with $PM_{2.5}$ and PM_{10} over lags of up to 24 hours in Shenzhen from 2013 to 2019. The cumulative exposure-response curves are calculated using a natural B-spline with 2 knots to model the exposure-response association. The black solid lines are the average percentage change in the risk of all-cause ambulance emergency calls, and the dotted lines are the 95% CIs. $PM_{2.5}$ and PM_{10}: particulate matter less than 2.5 and 10 μm in diameter, respectively.

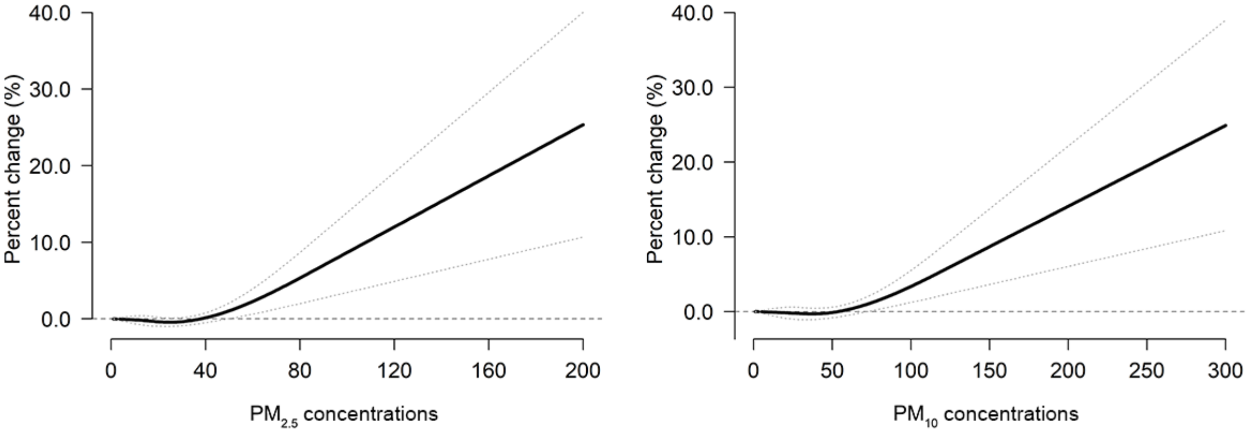


Figure 5. Cumulative concentration-response curves for the associations of ambulance emergency calls due to cardiovascular, respiratory, and reproductive diseases with $PM_{2.5}$ and PM_{10} over lags of up to 24 hours in Shenzhen from 2013 to 2019. A and D represent cardiovascular diseases, B and E represent respiratory diseases, and C and F represent reproductive illnesses. The cumulative exposure-response curves are calculated using a natural B-spline with 1 or 2 knots to model the exposure-response association. The black solid lines are the average percentage change in the risk of ambulance emergency calls due to cardiovascular, respiratory, and reproductive diseases with PM2.5 and PM_{10}, and the dotted lines are the 95% CIs. $PM_{2.5}$ and PM_{10}: particulate matter less than 2.5 and 10 µm in diameter, respectively.

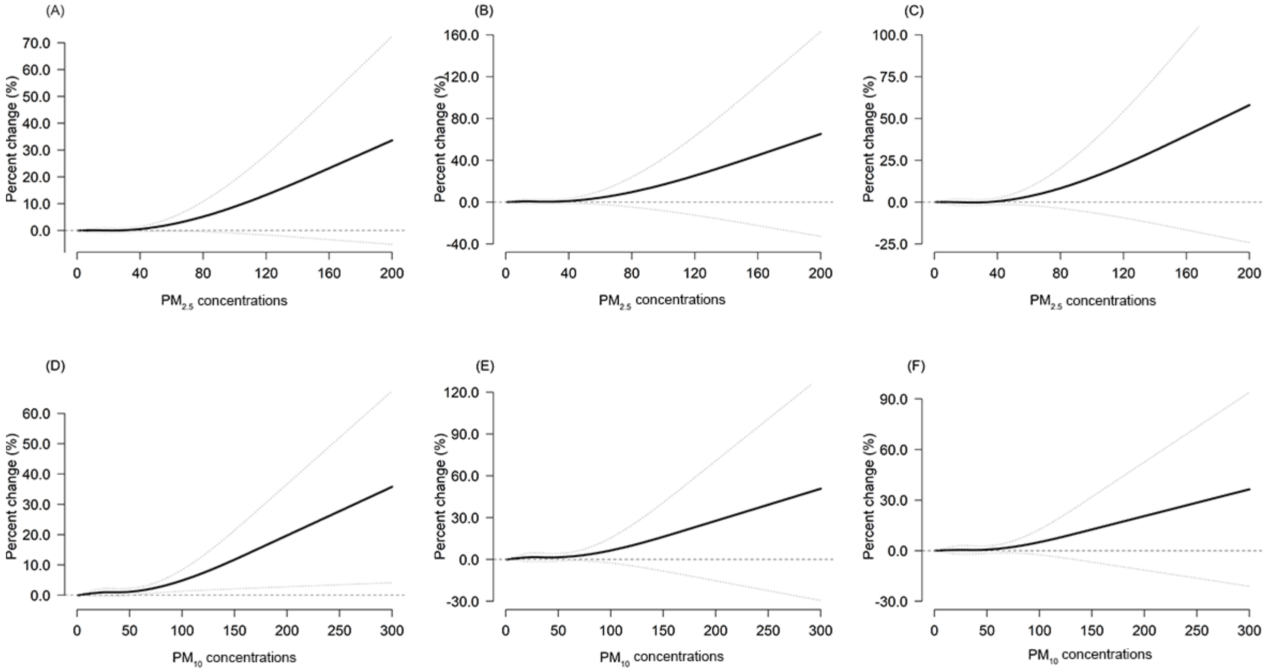


Table 4. Risk of ambulance emergency calls associated with each IQR[a] increase in PM2.5[b] and PM10[c] concentrations over lags of 0-24 hours, stratified by sex, age, season, and the time of day, in Shenzhen from 2013 to 2019. The models were adjusted public holidays, days of the week, hourly temperature, and hourly humidity.

Subgroups	All-cause, percentage change (95% CI)		Cardiovascular diseases, percentage change (95% CI)		Respiratory diseases, percentage change (95% CI)		Reproductive illnesses, percentage change (95% CI)	
	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}
Sex								
Male	1.5 (0.7 to 2.3)	1.8 (0.9 to 2.6)	1.6 (0.1 to 3.0)	1.9 (0.5 to 3.4)	2.5 (–1.2 to 6.4)	2.8 (–0.8 to 6.5)	2.0 (–0.8 to 4.9)	2.3 (–0.6 to 5.2)
Female	1.5 (0.7 to 2.4)	1.8 (1.0 to 2.7)	1.6 (0.2 to 3.1)	1.9 (0.5 to 3.4)	2.6 (–1.2 to 6.5)	2.8 (–0.8 to 6.5)	2.1 (–0.7 to 5.0)	2.3 (–0.6 to 5.3)
Age group (years)								
18-64	1.4 (0.6 to 2.1)	1.8 (0.9 to 2.6)	1.4 (0.1 to 2.8)	1.9 (0.5 to 3.4)	2.3 (–1.2 to 5.9)	2.7 (–0.8 to 6.4)	2.0 (–0.8 to 4.9)	3.7 (–2.1 to 9.9)
≥65	1.6 (0.6 to 2.6)	2.0 (1.1 to 3.0)	1.7 (0.2 to 3.3)	2.2 (0.6 to 3.8)	2.4 (–1.6 to 6.6)	2.8 (–1.1 to 6.8)	1.9 (–1.5 to 5.6)	2.5 (–1.2 to 6.3)
Season[d]								
Warm	1.5 (0.7 to 2.3)	1.8 (0.9 to 2.7)	0.6 (0.2 to 2.8)	1.2 (–0.4 to 2.8)	2.6 (–0.9 to 6.2)	2.8 (–1.3 to 7.0)	3.6 (–2.3 to 9.8)	3.7 (–2.6 to 10.4)
Cool	1.4 (–2.7 to 3.6)	–6.4 (–21.2 to 6.4)	1.6 (–0.2 to 3.4)	–12.0 (–34.0 to 9.4)	1.3 (–2.6 to 5.4)	1.6 (–3.3 to 6.7)	1.0 (–2.8 to 5.1)	2.6 (–6.5 to 12.5)
Time of day[e]								
Daytime	1.7 (0.5 to 3.0)	2.1 (0.9 to 3.4)	1.4 (–0.1 to 3.2)	1.5 (0.2 to 3.9)	2.7 (–2.1 to 7.7)	3.1 (–1.4 to 7.9)	1.5 (–3.0 to 6.3)	2.2 (–2.7 to 7.3)
Nighttime	1.4 (0.3 to 2.6)	1.7 (0.6 to 2.8)	1.6 (–1.7 to 3.8)	1.9 (–0.4 to 4.2)	1.6 (–3.0 to 8.4)	2.9 (–3.0 to 9.2)	5.7 (–2.9 to 15.2)	5.3 (–3.4 to 14.7)

[a] IQR: IQR of $PM_{2.5}$ was defined as the 25th to 75th percentile (24.0 μg/m^3); IQR of PM_{10} was defined as the 25th to 75th percentile (34.0 μg/m^3).

[b] $PM_{2.5}$: particulate matter less than 2.5 μm in diameter.

[c] PM_{10}: particulate matter less than 10 μm in diameter.

[d] Season: warm season is from 1st April 1 to October 31 of each year; cool season is from November 1 to 31 March 31 of each year.

[e] Time of the day: daytime is from 8:00 AM to 7:00 PM within one day; nighttime is from 8:00 PM to 7:00 AM in the next day.

Discussion

Principal Findings

We observed the associations of transient exposure to air pollutants with an increased risk of all-cause AECs, without any discernible threshold effects. The associations were stronger for $PM_{2.5}$ than for PM_{10}. In addition, the risks were apparent and statistically significant within 0-4 hours after exposure to the air pollutants. The excess relative risks showed substantial differences between the time-of-day strata.

Our findings were consistent with a robust body of previous studies demonstrating the positive associations of hourly air pollutants with all-cause and cause-specific AECs [15,16,23,25-30]. However, much of the existing evidence has focused on assessing the associations between cardiovascular diseases–related hospital admissions and exposures at the daily timescale through time-series analyses [31,32]. We have filled these knowledge gaps and generalized these results to the emergency medical service setting, as this study demonstrated that AECs may serve as a sensitive and timely marker of the adverse health effects of air pollution.

Given the use of different exposure metrics, time periods, study designs, and analytical strategies, our results cannot be compared directly across studies. For example, an analysis of Shenzhen AECs between January 18, 2013, and December 31, 2016, found that each 10 μg/m^3 increase in $PM_{2.5}$ and PM_{10} was associated with a 1.44% (95% CI, 0.70-2.19%) and 0.95% (95% CI, 0.39-1.51%) increase in hourly AECs over 5 hours, respectively [23]. In another study in Shenzhen [25], the authors observed positive associations of exposure to $PM_{2.5}$ with an increased risk of emergency department visits across different lag days using daily excessive concentration hours and the daily mean metric. In contrast, relatively weak effects were found for hourly peak $PM_{2.5}$. Rao et al [33] conducted a case-crossover study by using British Columbia Emergency Health Service data from 2010 to 2015 in Canada and showed a positive association between cardiovascular disease–related AECs and $PM_{2.5}$ exposure (OR 1.007, 95% CI 0.997-1.031), and there was a slight increase in the odds of respiratory disease–related AECs at 1.005 (95% CI 0.998, 1.013) over 48 hours. Another study conducted by Ai et al [15] between 2014 and 2016 found that each 10 μg/m^3 increase in $PM_{2.5}$ and $PM1_0$ was associated with a 0.19% (95% CI 0.03-0.35%) and 0.13% (95% CI, 0.02-0.24%) increase of all-cause AECs, respectively, whereas no significant effects of $PM_{2.5}$ and PM_{10} on cardiovascular morbidity were found.

We observed significant effects during the first 24 hours after exposure to the PM pollutants. Several human and animal studies have shown that acute exposure to PM may enhance thrombogenicity through various pathways, including platelet activation, oxidative stress, and the interplay between interleukin-6 and tissue factors [30,34,35]. A few studies have also identified the possible pathways, including endothelial dysfunction, inflammation, dyslipidemia, and autonomic and vascular dysfunction [33,36].

The stratified analyses showed stronger associations of air pollutants with cardiovascular diseases–related AECs in the nighttime than in the daytime, which is in line with existing evidence [22,37]. This may be due to low atmospheric pressure that always occurs during the night (from 1:00 AM to 5:00 AM). Moreover, we found several significant season differences in associations of PM air pollutants with respiratory diseases–related AECs, which could be explained by the virus or allergen being more active in the warm season. Our analysis observed stronger associations of all-cause AECs in the daytime, this could be explained by people often spending more time outdoors during the day and that the associations were estimated based on fixed-site monitors [20]. In addition, there were no associations of air pollutants with AECs stratified by sex in our study. These results were in line with previous literature [7,17], indicating that all patients, regardless of sex, seem to be at higher AEC risks after transient exposure to air pollution.

This study has several limitations. First, we used the individual weighted average of hourly air pollution concentrations as a proxy for personal exposure. Although there is some degree of exposure misclassification that may lead to the underestimation of associations, this limitation may not be avoidable in most time-series studies [22,38-40]. Second, we used prehospital diagnosis rather than medical diagnosis with specific International Classification of Disease codes because data on the final clinical diagnosis were not available. However, we defined our subtypes of AECs using the Medical Priority Dispatch System [16,41], which should not substantially bias our results. Third, we did not have detailed information on the locations of environmental monitoring stations and the proximity of roads, making it challenging to explore the possible effect modification of the associations. Fourth, we were unable to match the monitor station measurements to the closest patient because we were unable to acquire patients' location information. Therefore, we used the average city-level PM concentration as the exposure for patients in this study. Although the impact of this limitation is likely minimal, we plan to gather and analyze this association in future studies [20].

This study has some strengths. First, the Shenzhen Ambulance Emergency Centre provided high-quality data covering most hospitals in Shenzhen. The large sample size and individual-level data allowed us to conduct comprehensive statistical analyses to maximize the validity of the study results. Second, the time-stratified case-crossover study design facilitated causal inference of the study results by adjusting for time-invariant confounding [18,40]. Third, we collected hourly data on air pollutants when all-cause or cause-specific AECs occurred, which characterized the subdaily time effects of exposure to AECs.

Air pollution–related AECs are considered as a public health problem. This study provides the distributions of all-cause and cause-specific AECs over 24 hours and the associations of each IQR increase with AEC risks. The adverse health effects of air pollution are preventable through a combination of reducing exposure, reducing susceptibility, and improving adaptive capacity. In the context of public health preparedness for air pollution, local government response systems should typically provide air pollution–related information to the public, develop strategies to reduce air pollution–related risks, and allocate emergency ambulance resources equitably. It is also important for hospitals and emergency centers to adapt their procedures to meet the increased demands associated with air pollution. For example, local health systems need not only to allocate more resources (ie, health care providers and medical facilities) to general emergency care but also to increase the capacity of health care providers to provide specialized emergency care (ie, cardiovascular and respiratory diseases) at night because of the stronger associations of PM air pollutants found during nighttime hours.

Conclusion

PM air pollutants were associated with a higher relative risk of all-cause AECs and cardiovascular diseases–, respiratory diseases–, and reproductive illnesses–related AECs. The risk of all-cause AECs increased consistently with increasing concentrations of PM air pollutants, showing a nearly linear relationship with no apparent thresholds. The results of this study may be valuable to air pollution attributable to the distribution of emergency resources and consistent air pollution control.

Acknowledgments

This research was supported by the Sanming Project of Medicine in Shenzhen (grant SZSM201911005) and National Key Research and Development Program of China (grant 2022YFC3703003). The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors accept responsibility for the decision to submit for publication.

Data Availability

The data sets generated and/or analyzed during the current study are not publicly available owing to security protocols and privacy regulations but are available from the corresponding author on reasonable request.

Authors' Contributions

QZ, HS, and Z Zhang contributed to conceptualization, methodology, formal analysis, writing—original draft, visualization, and funding acquisition. Z Zhang contributed to writing—review and editing and supervision. RW contributed to resources and writing—review and editing. HZ, CQ, and ZL contributed to writing—review and editing. SS, JZ, YW, JH, YJ, Z Zheng, and JL contributed to writing—review and editing, supervision, and funding acquisition. All authors have read and agreed to the published version of the manuscript. JL (lijy@cnemc.cn) and Z Zhang (zzy@pku.edu.cn) are co-corresponding authors for this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Hourly particulate matter concentrations, ambulance emergency call distribution, air pollutant correlations, and sensitivity analyses.
[DOCX File , 770 KB - publichealth_v9i1e47022_app1.docx]

References

1. Sun S, Weinberger KR, Nori-Sarma A, Spangler KR, Sun Y, Dominici F, et al. Ambient heat and risks of emergency department visits among adults in the United States: time stratified case crossover study. BMJ 2021 Nov 24;375:e065653 [FREE Full text] [doi: 10.1136/bmj-2021-065653] [Medline: 34819309]
2. Nirel R, Adar SD, Dayan U, Vakulenko-Lagun B, Golovner M, Levy I, et al. Fine and coarse particulate matter exposures and associations with acute cardiac events among participants in a telemedicine service: a case-crossover study. Environ Health Perspect 2018 Sep;126(9):97003 [FREE Full text] [doi: 10.1289/EHP2596] [Medline: 30203992]
3. Du Y, Xu X, Chu M, Guo Y, Wang J. Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence. J Thorac Dis 2016 Jan;8(1):E8-E19 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2015.11.37] [Medline: 26904258]
4. Münzel T, Sørensen M, Gori T, Schmidt FP, Rao X, Brook J, et al. Environmental stressors and cardio-metabolic disease: part I-epidemiologic evidence supporting a role for noise and air pollution and effects of mitigation strategies. Eur Heart J 2017 Feb 21;38(8):550-556. [doi: 10.1093/eurheartj/ehw269] [Medline: 27460892]
5. Rasche M, Walther M, Schiffner R, Kroegel N, Rupprecht S, Schlattmann P, et al. Rapid increases in nitrogen oxides are associated with acute myocardial infarction: a case-crossover study. Eur J Prev Cardiol 2018 Nov;25(16):1707-1716. [doi: 10.1177/2047487318755804] [Medline: 29446990]
6. Rich DQ, Utell MJ, Croft DP, Thurston SW, Thevenet-Morrison K, Evans KA, et al. Daily land use regression estimated woodsmoke and traffic pollution concentrations and the triggering of ST-elevation myocardial infarction: a case-crossover study. Air Qual Atmos Health 2018;11(2):239-244 [FREE Full text] [doi: 10.1007/s11869-017-0537-1] [Medline: 29568338]
7. Li J, Liu C, Cheng Y, Guo S, Sun Q, Kan L, et al. Association between ambient particulate matter air pollution and ST-elevation myocardial infarction: a case-crossover study in a Chinese city. Chemosphere 2019 Mar;219:724-729. [doi: 10.1016/j.chemosphere.2018.12.094] [Medline: 30557729]
8. Chen K, Schneider A, Cyrys J, Wolf K, Meisinger C, Heier M, KORA Study Group. Hourly exposure to ultrafine particle metrics and the onset of myocardial infarction in Augsburg, Germany. Environ Health Perspect 2020 Jan;128(1):17003 [FREE Full text] [doi: 10.1289/EHP5478] [Medline: 31939685]
9. Lin H, Wang X, Qian ZM, Guo S, Yao Z, Vaughn MG, et al. Daily exceedance concentration hours: a novel indicator to measure acute cardiovascular effects of PM in six Chinese subtropical cities. Environ Int 2018 Feb;111:117-123. [doi: 10.1016/j.envint.2017.11.022] [Medline: 29190528]
10. Deng Q, Feng J, Gao P, Ni H. Combined effects of vehicles and waste incineration on urban air halogenated and parent polycyclic aromatic hydrocarbons. Environ Int 2023 Jan;171:107720 [FREE Full text] [doi: 10.1016/j.envint.2022.107720] [Medline: 36580736]
11. Air pollution from wildfires, rising heat affected two-thirds of U.S. West. Climate.gov. 2022 Jan 15. URL: https://www.climate.gov/news-features/feed/air-pollution-wildfires-rising-heat-affected-two-thirds-us-west [accessed 2023-01-07]
12. Di Q, Dai L, Wang Y, Zanobetti A, Choirat C, Schwartz JD, et al. Association of short-term exposure to air pollution with mortality in older adults. JAMA 2017 Dec 26;318(24):2446-2456 [FREE Full text] [doi: 10.1001/jama.2017.17923] [Medline: 29279932]
13. Ito K, Mathes R, Ross Z, Nádas A, Thurston G, Matte T. Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. Environ Health Perspect 2011 Apr;119(4):467-473 [FREE Full text] [doi: 10.1289/ehp.1002667] [Medline: 21463978]
14. Chen TT, Zhan ZY, Yu YM, Xu LJ, Guan Y, Ou CQ. Effects of hourly levels of ambient air pollution on ambulance emergency call-outs in Shenzhen, China. Environ Sci Pollut Res Int 2020 Jul;27(20):24880-24888. [doi: 10.1007/s11356-020-08416-w] [Medline: 32337675]

15. Ai S, Wang C, Qian ZM, Cui Y, Liu Y, Acharya BK, et al. Hourly associations between ambient air pollution and emergency ambulance calls in one central Chinese city: implications for hourly air quality standards. Sci Total Environ 2019 Dec 15;696:133956. [doi: 10.1016/j.scitotenv.2019.133956] [Medline: 31450053]
16. Yao J, Brauer M, Wei J, McGrail KM, Johnston FH, Henderson SB. Sub-daily exposure to fine particulate matter and ambulance dispatches during wildfire seasons: a case-crossover study in British Columbia, Canada. Environ Health Perspect 2020 Jun;128(6):67006. [doi: 10.1289/EHP5792] [Medline: 32579089]
17. Bhaskaran K, Hajat S, Armstrong B, Haines A, Herrett E, Wilkinson P, et al. The effects of hourly differences in air pollution on the risk of myocardial infarction: case crossover analysis of the MINAP database. BMJ 2011 Sep 20;343:d5531 [FREE Full text] [doi: 10.1136/bmj.d5531] [Medline: 21933824]
18. Peters A, Dockery DW, Muller JE, Mittleman MA. Increased particulate air pollution and the triggering of myocardial infarction. Circulation 2001 Jun 12;103(23):2810-2815. [doi: 10.1161/01.cir.103.23.2810] [Medline: 11401937]
19. Carracedo-Martínez E, Taracido M, Tobias A, Saez M, Figueiras A. Case-crossover analysis of air pollution health effects: a systematic review of methodology and application. Environ Health Perspect 2010 Aug;118(8):1173-1182 [FREE Full text] [doi: 10.1289/ehp.0901485] [Medline: 20356818]
20. Chen R, Jiang Y, Hu J, Chen H, Li H, Meng X, et al. Hourly air pollutants and acute coronary syndrome onset in 1.29 million patients. Circulation 2022 Jun 14;145(24):1749-1760. [doi: 10.1161/CIRCULATIONAHA.121.057179] [Medline: 35450432]
21. Spangler KR, Weinberger KR, Wellenius GA. Suitability of gridded climate datasets for use in environmental epidemiology. J Expo Sci Environ Epidemiol 2019 Oct;29(6):777-789 [FREE Full text] [doi: 10.1038/s41370-018-0105-2] [Medline: 30538298]
22. Cheng J, Tong S, Su H, Xu Z. Hourly air pollution exposure and emergency department visit for acute myocardial infarction: vulnerable populations and susceptible time window. Environ Pollut 2021 Nov 01;288:117806. [doi: 10.1016/j.envpol.2021.117806] [Medline: 34329072]
23. Wei Y, Wang Y, Di Q, Choirat C, Wang Y, Koutrakis P, et al. Short term exposure to fine particulate matter and hospital admission risks and costs in the Medicare population: time stratified, case crossover study. BMJ 2019 Nov 27;367:l6258 [FREE Full text] [doi: 10.1136/bmj.l6258] [Medline: 31776122]
24. World Health Organization. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva; Switzerland: World Health Organization; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK574594/ [accessed 2023-06-06]
25. Jiao A, Xiang Q, Ding Z, Cao J, Ho HC, Chen D, et al. Short-term impacts of ambient fine particulate matter on emergency department visits: comparative analysis of three exposure metrics. Chemosphere 2020 Feb;241:125012. [doi: 10.1016/j.chemosphere.2019.125012] [Medline: 31606575]
26. Nirel R, Levy I, Adar SD, Vakulenko-Lagun B, Peretz A, Golovner M, et al. Concentration-response relationships between hourly particulate matter and ischemic events: a case-crossover analysis of effect modification by season and air-mass origin. Sci Total Environ 2021 Mar 15;760:143407. [doi: 10.1016/j.scitotenv.2020.143407] [Medline: 33199016]
27. Liang Z, You C, Zhang X, Wang X, Xiao D, He S, et al. Three exposure metrics of size-specific particulate matter associated with acute lower respiratory infection hospitalization in children: a multi-city time-series analysis in China. Sci Total Environ 2022 Apr 10;816:151636. [doi: 10.1016/j.scitotenv.2021.151636] [Medline: 34774633]
28. Lin X, Du Z, Liu Y, Hao Y. The short-term association of ambient fine particulate air pollution with hypertension clinic visits: a multi-community study in Guangzhou, China. Sci Total Environ 2021 Jun 20;774:145707. [doi: 10.1016/j.scitotenv.2021.145707] [Medline: 33611009]
29. Zhang R, Jiang Y, Zhang G, Yu M, Wang Y, Liu G. Association between short-term exposure to ambient air pollution and hospital admissions for transient ischemic attacks in Beijing, China. Environ Sci Pollut Res Int 2021 Feb;28(6):6877-6885. [doi: 10.1007/s11356-020-10963-1] [Medline: 33010014]
30. Lucking AJ, Lundbäck M, Barath SL, Mills NL, Sidhu MK, Langrish JP, et al. Particle traps prevent adverse vascular and prothrombotic effects of diesel engine exhaust inhalation in men. Circulation 2011 Apr 26;123(16):1721-1728. [doi: 10.1161/CIRCULATIONAHA.110.987263] [Medline: 21482966]
31. Konduracka E, Niewiara, Guzik B, Kotynia M, Szolc P, Gajos G, et al. Effect of short-term fluctuations in outdoor air pollution on the number of hospital admissions due to acute myocardial infarction among inhabitants of Kraków, Poland. Pol Arch Intern Med 2019 Feb 28;129(2):88-96 [FREE Full text] [doi: 10.20452/pamw.4424] [Medline: 30688288]
32. Wu K, Ho HC, Su H, Huang C, Zheng H, Zhang W, et al. A systematic review and meta-analysis of intraday effects of ambient air pollution and temperature on cardiorespiratory morbidities: first few hours of exposure matters to life. EBioMedicine 2022 Dec;86:104327 [FREE Full text] [doi: 10.1016/j.ebiom.2022.104327] [Medline: 36323182]
33. Rao X, Zhong J, Brook RD, Rajagopalan S. Effect of particulate matter air pollution on cardiovascular oxidative stress pathways. Antioxid Redox Signal 2018 Mar 20;28(9):797-818 [FREE Full text] [doi: 10.1089/ars.2017.7394] [Medline: 29084451]
34. Robertson S, Miller MR. Ambient air pollution and thrombosis. Part Fibre Toxicol 2018 Jan 03;15(1):1 [FREE Full text] [doi: 10.1186/s12989-017-0237-x] [Medline: 29298690]

35. Smyth E, Solomon A, Birrell MA, Smallwood MJ, Winyard PG, Tetley TD, et al. Influence of inflammation and nitric oxide upon platelet aggregation following deposition of diesel exhaust particles in the airways. Br J Pharmacol 2017 Jul;174(13):2130-2139 [FREE Full text] [doi: 10.1111/bph.13831] [Medline: 28437857]
36. Münzel T, Gori T, Al-Kindi S, Deanfield J, Lelieveld J, Daiber A, et al. Effects of gaseous and solid constituents of air pollution on endothelial function. Eur Heart J 2018 Oct 07;39(38):3543-3550 [FREE Full text] [doi: 10.1093/eurheartj/ehy481] [Medline: 30124840]
37. Lin Z, Chen R, Jiang Y, Xia Y, Niu Y, Wang C, et al. Cardiovascular benefits of fish-oil supplementation against fine particulate air pollution in China. J Am Coll Cardiol 2019 Apr 30;73(16):2076-2085 [FREE Full text] [doi: 10.1016/j.jacc.2018.12.093] [Medline: 31023432]
38. Shin HH, Maquiling A, Thomson EM, Park I, Stieb DM, Dehghani P. Sex-difference in air pollution-related acute circulatory and respiratory mortality and hospitalization. Sci Total Environ 2022 Feb 01;806(Pt 3):150515 [FREE Full text] [doi: 10.1016/j.scitotenv.2021.150515] [Medline: 34627116]
39. Conti S, Harari S, Caminati A, Zanobetti A, Schwartz JD, Bertazzi PA, et al. The association between air pollution and the incidence of idiopathic pulmonary fibrosis in Northern Italy. Eur Respir J 2018 Jan 25;51(1):1700397 [FREE Full text] [doi: 10.1183/13993003.00397-2017] [Medline: 29371377]
40. Chen J, Jiao Z, Liang Z, Ma J, Xu M, Biswal S, et al. Association between temperature variability and global meningitis incidence. Environ Int 2023 Jan;171:107649 [FREE Full text] [doi: 10.1016/j.envint.2022.107649] [Medline: 36470121]
41. Clawson JJ, Dernocoeur K, Murray C. Principles of Emergency Medical Dispatch. 6th ed. Salt Lake City, UT: National Academy of EMD; 2015.

Abbreviations

AEC: ambulance emergency call
OR: odds ratio
PM: particulate matter
PM10: particulate matter less than 10 µm in diameter
PM2.5: particulate matter less than 2.5 µm in diameter

Edited by A Mavragani, T Sanchez; submitted 06.03.23; peer-reviewed by M Wenjun, A Jiao; comments to author 28.03.23; revised version received 12.04.23; accepted 26.05.23; published 20.06.23.

Please cite as:
Zhou Q, Shi H, Wu R, Zhu H, Qin C, Liang Z, Sun S, Zhao J, Wang Y, Huang J, Jin Y, Zheng Z, Li J, Zhang Z
Associations Between Hourly Ambient Particulate Matter Air Pollution and Ambulance Emergency Calls: Time-Stratified Case-Crossover Study
JMIR Public Health Surveill 2023;9:e47022
URL: https://publichealth.jmir.org/2023/1/e47022
doi:10.2196/47022
PMID:37243735

Original Paper

The Combined Effects of Short-Term Exposure to Multiple Meteorological Factors on Unintentional Drowning Mortality: Large Case-Crossover Study

Yingyin Liu[1*], MPH; Xiaomei Dong[1*], PhD; Zhixing Li[1,2], MPH; Sui Zhu[1], PhD; Ziqiang Lin[1], PhD; Guanhao He[1], PhD; Weiwei Gong[3], MPH; Jianxiong Hu[4], MS; Zhulin Hou[6], MS; Ruilin Meng[7], PhD; Chunliang Zhou[8], BS; Min Yu[3], MS; Biao Huang[6], MS; Lifeng Lin[7], PhD; Jianpeng Xiao[4], PhD; Jieming Zhong[3], PhD; Donghui Jin[8], BS; Yiqing Xu[8], MS; Lingshuang Lv[8], MS; Cunrui Huang[9], PhD; Tao Liu[10], PhD; Wenjun Ma[1], PhD

[1]Department of Public Health and Preventive Medicine, School of Medicine, Jinan University, Guangzhou, China
[2]Department of Nosocomial Infection Management, Affiliated Nanfang Hospital of Southern Medical University, Guangzhou, China
[3]Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China
[4]Guangdong Provincial Institute of Public Health, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China
[5]Yunnan Provincial Center for Disease Control and Prevention, Kunming, China
[6]Jilin Provincial Center for Disease Control and Prevention, Changchun, China
[7]Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China
[8]Hunan Provincial Center for Disease Control and Prevention, Changsha, China
[9]Vanke School of Public Health, Tsinghua University, Beijing, China
[10]Disease Control and Prevention Institute of Jinan University, School of Medicine, Jinan University, Guangzhou, China
[*]these authors contributed equally

Corresponding Author:
Wenjun Ma, PhD
Department of Public Health and Preventive Medicine
School of Medicine
Jinan University
601 West Huangpu Avenue
Tianhe District
Guangzhou, 510632
China
Phone: 86 13622280436
Email: mawj@gdiph.org.cn

Abstract

Background: Drowning is a serious public health problem worldwide. Previous epidemiological studies on the association between meteorological factors and drowning mainly focused on individual weather factors, and the combined effect of mixed exposure to multiple meteorological factors on drowning is unclear.

Objective: We aimed to investigate the combined effects of multiple meteorological factors on unintentional drowning mortality in China and to identify the important meteorological factors contributing to drowning mortality.

Methods: Unintentional drowning death data (based on International Classification of Diseases, 10th Edition, codes W65-74) from January 1, 2013, to December 31, 2018, were collected from the Disease Surveillance Points System for Guangdong, Hunan, Zhejiang, Yunnan, and Jilin Provinces, China. Daily meteorological data, including daily mean temperature, relative humidity, sunlight duration, and rainfall in the same period were obtained from the Chinese Academy of Meteorological Science Data Center. We constructed a time-stratified case-crossover design and applied a generalized additive model to examine the effect of individual weather factors on drowning mortality, and then used quantile g-computation to estimate the joint effect of the mixed exposure to meteorological factors.

Results: A total of 46,179 drowning deaths were reported in the 5 provinces in China from 2013 to 2018. In an effect analysis of individual exposure, we observed a positive effect for sunlight duration, a negative effect for relative humidity, and U-shaped associations for temperature and rainfall with drowning mortality. In a joint effect analysis of the above 4 meteorological factors, a 2.99% (95% CI 0.26%-5.80%) increase in drowning mortality was observed per quartile rise in exposure mixture. For the total population, sunlight duration was the most important weather factor for drowning mortality, with a 93.1% positive contribution to the overall effects, while rainfall was mainly a negative factor for drowning deaths (90.5%) and temperature and relative humidity contributed 6.9% and –9.5% to the overall effects, respectively.

Conclusions: This study found that mixed exposure to temperature, relative humidity, sunlight duration, and rainfall was positively associated with drowning mortality and that sunlight duration, rather than temperature, may be the most important meteorological factor for drowning mortality. These findings imply that it is necessary to incorporate sunshine hours and temperature into early warning systems for drowning prevention in the future.

(JMIR Public Health Surveill 2023;9:e46792) doi:10.2196/46792

KEYWORDS

drowning; exposure mixture; quantile g-computation; environmental epidemiology; meteorological factor

Introduction

Drowning is the process of experiencing respiratory impairment from submersion or immersion in liquid [1]; it is a serious public health problem worldwide. The World Health Organization (WHO) reported that an estimated 236,000 people died from drowning in 2019. Globally, drowning is the third leading cause of death by unintentional injury, accounting for 7% of all injury-related deaths [2]. Despite the downward trend in drowning mortality rates [3], there are still considerable disparities across countries. Developed countries tend to have lower drowning mortality; for example, in the United States, the drowning mortality rate was 1.23 per 100,000 from 2010 to 2019 [4]. However, developing regions, like Southeast Asia, are the regions with the highest drowning mortality rates in the world. In the last decades, China has made great progress in drowning prevention and control, and the drowning mortality rate has decreased from 15.09 per 100,000 in 1990 to 3.97 per 100,000 in 2019 [5]; however, this is still a high drowning burden compared with most other countries.

Prior studies on drowning mainly focused on epidemiological characteristics or risk factor analysis [6-8], clinical first aid [9,10], and intervention policy assessment [11,12]. Few studies have investigated the association of meteorological factors with drowning. Among these few studies, several reports indicate that there is a clear seasonal distribution of drowning deaths, with the highest drowning mortality in the summer [13]. The summer heat itself may not increase the risk of drowning, but it may lead to changes in behavior that increase susceptibility to drowning, such as increased water-related activities, increased alcohol consumption, or decreased use of personal flotation devices [14]. In terms of humidity, several studies report that high relative humidity (RH) at high temperatures significantly increases body temperature and increases the chance that people engage in water-related activities such as swimming [15], which in turn leads to an increased risk of drowning. Flooding due to extremely heavy rainfall is also an important cause of drowning mortality. For example, in July 2021, a flood occurred in Henan Province, China, due to persistent extremely heavy rainfall, and this incident resulted in hundreds of drowning deaths [16]. Even though there is no research on the association between duration of sunlight (DS) and drowning, there is research illustrating the importance of daylight hours for those who plan to enjoy swimming in a lake or outdoor pool, thereby potentially increasing the chance of drowning [17]. These studies focused on the effects of single weather factors on drowning. However, in the real world, people are exposed to multiple weather factors simultaneously, not only one at a time. Regrettably, no prior studies have examined the combined effects of exposure to multiple meteorological factors on drowning.

Indeed, it is a challenge to quantify the combined effects of multiple meteorological exposure factors on drowning using traditional statistical models. In order to reduce collinearity in the model, previous multivariate regression models, such as linear regression, generalized linear regression, and generalized additive models (GAMs) have usually explored the association between a single or limited environmental factors and health outcomes. However, this approach could not accurately reveal the true impact of multiple environmental exposures [18]. In recent years, several advanced statistical methods, such as quantile g-computation, have been developed to examine the joint health effects of exposure mixture [19], which provides an opportunity for combined-effect analysis of exposure mixture in environmental epidemiology. Quantile g-computation not only allows for nonlinear correlation analysis, but also calculates the positive and negative weights contributed by each exposure.

This study aimed to investigate the combined effects of multiple meteorological factors, including mean temperature (MT), RH, DS, and rainfall, on drowning mortality based on a data set from 5 provinces in China using a quantile g-computation model [19] and to further quantify the relative importance of each meteorological factor. Our findings may provide scientific information for developing an early warning system for drowning based on multiple meteorological factors to prevent drowning in the context of global warming.

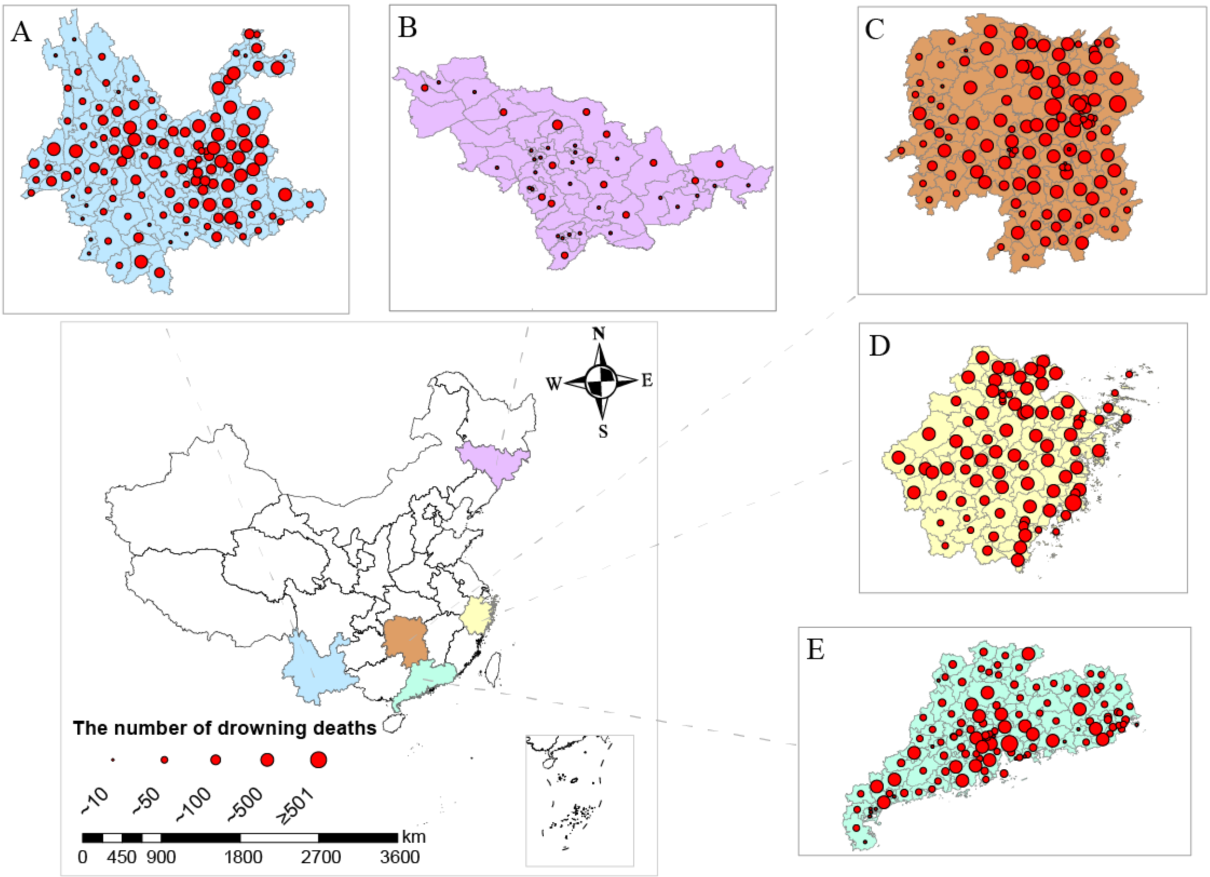
Methods

Data Collection

In this study, the study sites in China were Guangdong Province in south China, Hunan Province in the middle part of the Yangtze River, Zhejiang Province on the southeast coast of China, Yunnan Province in southwest China, and Jilin Province in northeast China. Guangdong and Hunan are both regions with rich sunlight, heat, and water resources. Zhejiang Province is rich in marine resources. Both Jilin and Yunnan provinces have many rivers and lakes.

Drowning mortality data from January 1, 2013, to December 31, 2018, in these 5 Chinese provinces were collected from the Provincial Disease Surveillance Points System. The information on each death included the date of death, cause of death, address, gender, age, education level, and cause of death code. In the International Classification of Diseases, 10th Edition (ICD-10), the codes W65 to 74 are unintentional drowning deaths, excluding drownings due to cataclysms or transportation accidents. The spatial distribution of drowning deaths in the 5 provinces is shown in Figure 1.

Figure 1. Spatial distribution of drowning deaths at the district and county level in 5 provinces. (A) Yunnan Province; (B) Jilin Province; (C) Hunan Province; (D) Zhejiang Province; (E) Guangdong Province.



Daily meteorological data, including daily MT, RH, DS, and rainfall from 2013 to 2018 were collected from the China Meteorological Science Data Center, which collects data from 698 meteorological surveillance points across the country [20]. In order to obtain accurate daily MT and RH at a county level, we used Australian National University Splines (ANUSPLIN) thin plate smoothing [21] to interpolate station meteorological data to form raster data for daily MT and RH with a resolution of 0.01° × 0.01°. In the spatial interpolation model, we introduced latitude, longitude, and altitude as independent variables. The interpolation was validated using the 10-fold crossover method with good prediction accuracy, as previously described [22]. Daily mean DS and rainfall amount were calculated by averaging the values of all stations in each city because these 2 weather factors returned inaccurate data from interpolation in our preliminary analysis.

Air pollution might be a confounding factor, so we also collected data on fine particulate matter (PM2.5) as a representative of air pollutants to control for possible confounding effects. Daily air pollution data (ie, PM2.5) were derived from the National Urban Air Quality Real-Time Release Platform [23]. Because the air pollution surveillance points did not completely cover our study provinces, the land use regression method, based on the random forest model, was used to predict ambient air pollutants; the model showed good prediction capability [24].

Ethics Approval

Ethical approval was obtained from the Medical Research Ethics Committee of Guangdong Provincial Center for Disease Control and Prevention (2019025), and all procedures followed the ethical standards specified by the institution.

Statistical Analysis

This study used a case-crossover design based on temporal stratification, which was further adapted from the case-crossover design of Levy et al [25]. The basic principle is to stratify time so that each case with an outcome is treated as a case group, while exposures in the same year, month, and day of the week are treated as controls, meaning that each case has 3 to 4 controls, thereby effectively controlling for confounding factors, including long-term trends, gender, and age, and effectively avoiding control selection bias in the original case-crossover design and enabling unbiased estimates to be obtained. This research design has been widely used to investigate the effects of short-term environmental exposure on population health, especially in the field of health-effect assessment of meteorological factors [26] and air pollution [27].

Descriptive analyses were represented by numbers, means, maximum values, minimum values, and medians (quartiles), respectively. Spearman correlation was used to examine the correlations between meteorological factors and PM2.5.

Few previous studies have reported the effects of wind speed and air pressure on drowning. Moreover, after preliminary analysis, we observed that wind speed and air pressure had little effect on drowning mortality (Multimedia Appendix 1). Therefore, these 2 weather factors were not included in the joint-effect analysis. In our study, we first conducted individual meteorological factor analysis using a GAM [28] after adjusting for confounding factors (ie, PM2.5) to assess the individual effect of each meteorological factor (MT, RH, DS, and rainfall) on drowning mortality. The GAM equation is as follows:

$$\mathrm{Logit(case)=s(Variable,df)+s(PM_{2.5},df)+}\alpha,$$

where *Logit()* is the connection function, *case* indicates the case type with 1 for cases and 0 for controls; *Variable* represents the meteorological factors to be analyzed, that is, MT, RH, DS and rainfall; *PM2.5* is the concentration of PM2.5; *s()* is the nonparametric smoothing function, which is used to fit the nonlinear relationship between meteorological factors (ie, PM2.5 and drowning, respectively), with 3 degrees of freedom selected by generalized cross-validation [29]; and α denotes the intercept term.

Second, in order to further estimate the joint effect of multiple meteorological factor exposure on drowning mortality, we used a quantile g-computation model [30]. Quantile g-computation is a new approach to estimate the joint effects of exposure mixture [19]. It has been used to evaluate the change in health outcome for each quantile of mixed exposure to multiple environmental factors [31-33], and it does not require the same effect directions between exposure variables and outcome [18]. In this study, the model was implemented through categorizing MT, RH, DS, and rainfall in quartiles, , coded as 0, 1, 2, and 3. The mixture is first fitted nonlinearly to drowning and then linearized according to the results. The quantile g-computation model equation is as follows (confounders *Z* could also be included):

where (the change in drowning for a 1-unit change in all exposures) and each meteorological factor is given a negative or positive weight. The quantile g-computation estimator of the exposure response, ψ, is the sum of the regression coefficients across the included exposures. If all β_j are in the same direction, then the weight for each component (indexed by *k*) is defined as , which is the proportion of the effect due to that component and sums to 1.0. If factors have different directions of effect, then the weights are interpreted as the proportion of the positive or negative partial effect, and the positive and negative weights together sum to 2.0 [32]. After obtaining regression coefficients for 71 cities, values from different cities were combined using meta-analysis to obtain the exposure-response relationship between exposure mixture and drowning risk. The results of the nonlinear fit were expressed as exposure response curves.

The linearized results were expressed in terms of excess risk (ER) with the 95% CI. The formula for ER is as follows:

$$\mathrm{ER} = [\exp(\beta) - 1] \times 100\%$$

where β is the effect of exposure mixture in meta-analysis.

Moreover, we used the quantile g-computation model to conduct subgroup analysis by sex (male, female), age group (<25 years, 25-64 years, ≥65 years) and season (the cold season, from October to March, and the warm season, from April to September). The following formula was used for the between-group variability test:

where b_1 and b_2 are the effect estimates of the 2 groups, respectively, and SE_1 and SE_2 are the SEs of the estimated effects of the 2 groups, respectively [34].

Sensitivity analysis was performed by changing the parameters of the model as well as by adjusting the confounding factors in order to verify the stability and reliability of the model. In the GAM, we conducted a sensitivity analysis by changing the degrees of freedom (*df*=2,4,5) for MT, RH, DS, and rainfall. In the quantile g-computation model, model 1 removed the confounding factor PM2.5 for sensitivity analysis. Model 2, model 3, and model 4 used 2-day, 3-day, and 4-day moving averages for MT, RH, and PM2.5 for sensitivity analysis.

All tests were conducted 2-sided, and effects with $P<.05$ were considered statistically significant. All analyses were conducted with R (version 4.1.0; R Foundation for Statistical Computing) and the *mgcv*, *qgcomp*, *mvmeta*, and *metafor* R packages.

Results

Descriptive Analysis

There were 46,179 drowning deaths in the 5 Chinese provinces from 2013 to 2018, with an average of 21 and a maximum of 68 drowning deaths per day. Each city had an average of 1 drowning death per day, with a maximum of 10 deaths (Table 1). A total of 17,408 people died from drowning during the summer seasons. The number of drowning deaths was the highest in Hunan Province and the lowest in Jilin Province. The number of drowning deaths was much higher for males (n=29,805, 64.5%) than females (n=16,374, 35.5%). The daily MT, RH, rainfall and PM2.5 during the study period were 20.4 °C, 76.7%, 4.5 mm, and 37.6 μg/m^3, respectively (Table 2).

Table 1. Descriptive statistics of meteorological factors and daily drowning deaths by season, province, sex, and age group from 2013 to 2018.

Characteristics	Drowning deaths overall (n=46,179), n (%)	Drowning deaths per day, n					
		Mean (total=21.1)	Minimum (total=3)	P_{25}[a] (total=14)	P_{50}[b] (total=19)	P_{75}[c] (total=26)	Maximum (total=68)
In each city	N/A[d]	1.3	1	1	1	2	10
Season							
Warm	29,145 (63.1)	26.5	6	20	26	32	68
Cold	17,034 (36.9)	15.6	3	12	15	19	40
Sex							
Male	29,805 (64.5)	13.6	1	8	12	18	49
Female	16,374 (35.5)	7.5	1	5	7	10	25
Age (years)							
≤24	13,827 (30)	6.4	1	3	5	9	26
25-64	16,918 (36.6)	7.8	1	5	7	10	27
≥65	15,434 (33.4)	7.1	1	5	7	9	23

[a] P_{25}: 25th percentile.

[b] P_{50}: 50th percentile.

[c] P_{75}: 75th percentile.

[d] N/A: not applicable.

Table 2. Descriptive statistics of meteorological factors from 2013 to 2018.

Meteorological factors	Mean	Minimum	P_{25}[a]	P_{50}[b]	P_{75}[c]	Maximum
Average temperature (°C)	20.6	0.1	15.4	21.8	27.1	33.0
Relative humidity (%)	76.9	50.6	70.6	77.9	84.2	95.2
Sunlight duration (hours)	5.3	0.0	1.3	5.6	8.8	13.9
Rainfall (mm)	2.9	0.0	0.0	0.2	2.7	35.0
PM2.5 (μg/m^3)	38.1	3.6	20.7	31.5	47.6	302.9

[a] P_{25}: 25th percentile.

[b] P_{50}: 50th percentile.

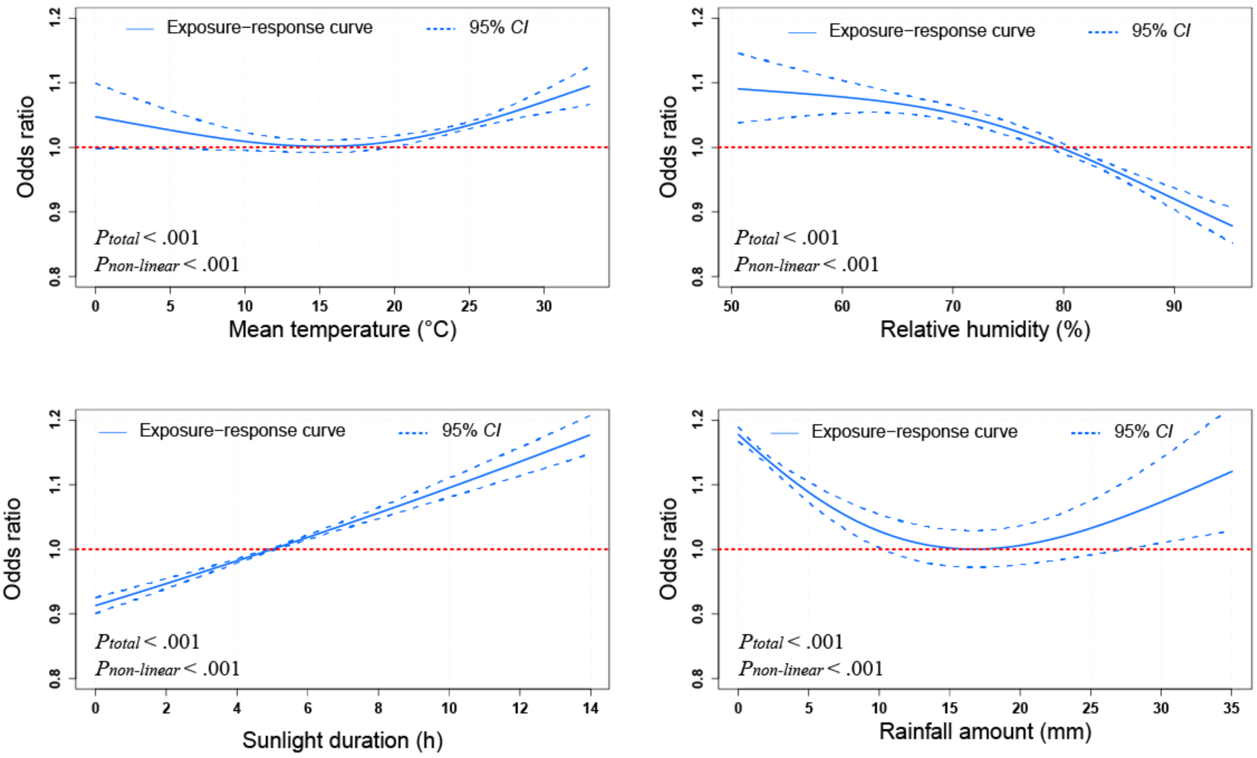
[c] P_{75}: 75th percentile.

Individual Effect of Meteorological Factors on Drowning Mortality

Among the 4 meteorological factors, the strongest correlation was observed between RH and DS (r=–0.61; *P*<.05; Multimedia Appendix 2). As shown in Figure 2, in an individual-effect analysis, there were U-shaped relationships between MT, rainfall and drowning mortality, with the lowest risk at 15.4 °C for MT and 16.6 mm for rainfall, respectively. RH was negatively associated with drowning mortality, while DS was positively associated with drowning mortality.

In the sensitivity analysis of the GAM, the relationship curves for MT, RH, DS, and rainfall with drowning deaths changed little after changing the degrees of freedom in the spline function. (Multimedia Appendix 3)

Figure 2. The associations of daily mean temperature, relative humidity, sunlight duration, and rainfall with drowning mortality in generalized additive models. Solid lines represent the exposure-response curve; dotted lines represent the 95% CI.



Combined Effect of Multiple Meteorological Factors on Drowning Mortality

In the exposure mixture analysis using quantile g-computation, mixed exposure to the 4 meteorological factors was approximately linearly associated with the risk of drowning. The risk tended to increase with each percentile (Figure 3). Meanwhile, we conducted an effect modification analysis of city-level characteristics and found that the risk of drowning caused by meteorological exposure mixture was higher in cities with higher population density, while urban gross domestic product, latitude, longitude, and altitude were found to have only weak modification effects (Multimedia Appendix 4).

After linearization, we observed a 2.99% (95% CI 0.26%-5.80%) drowning mortality increase per quartile rise in the 4 meteorological factors (Figure 4). Subgroup analyses showed that the effect for males was higher than for females with a statistically significant difference (Z=2.21; P=.03). The risk of drowning was heterogeneous across age groups in combined effects. Although those aged 25 to 64 years had the highest risk (ER 4.56%, 95% CI 0.07%-9.25% per quartile rise in the 4 meteorological factors), the difference among all age groups was not statistically significant. The drowning risk attributed to mixed exposure to the 4 weather factors in the warm season seemed higher than in the cold season, but the difference was also not statistically significant (Z=1.61; P=.11).

Figure 5 further shows the relative contribution of each meteorological factor to the combined effect. Sunlight duration had a positive weight with a much higher proportion among all subgroups: 93.1%, 94.1%, 100%, 88.9%, 83.1%, and 83.9% for the total population, males, females, and people aged 0 to 24 years, 25 to 64 years, and ≥65 years, respectively. RH negatively contributed to the combined effect in the total population (9.5%), females (26%), males (9.9%), and young people aged 0 to 24 years (43.2%) and positively contributed in people aged 25 to 64 years (12.4%) and ≥65 years (16.1%). Temperature contributed positively in the total population (6.9%), males (5.9%), the population aged 0 to 24 years (11.1%), and the population aged 25 to 64 years (4.5%), while it made a negative contribution in females (17.8%) and people aged ≥65 years (37.2%). Rainfall always contributed negatively to drowning mortality among all populations (90.5%, 90.2%, 56.2%, 56.8%, 100%, and 62.8% for the total population, males, females, people aged 0 to 24 years, 25 to 64 years, and ≥65 years, respectively; Figure 5).

In a quantile g-computation model sensitivity analysis, the results were basically stable whether adjusting or not adjusting for PM2.5 in the model. The excess risk of drowning increased when moving averages of each weather factor were used in the model (Multimedia Appendix 5).

Figure 3. Exposure-response relationships between meteorological exposure mixture and drowning risk in 71 cities from a quantile g-computation analysis. The line is the exposure response curve, and the blue shaded area is the 95% CI at each percentile.

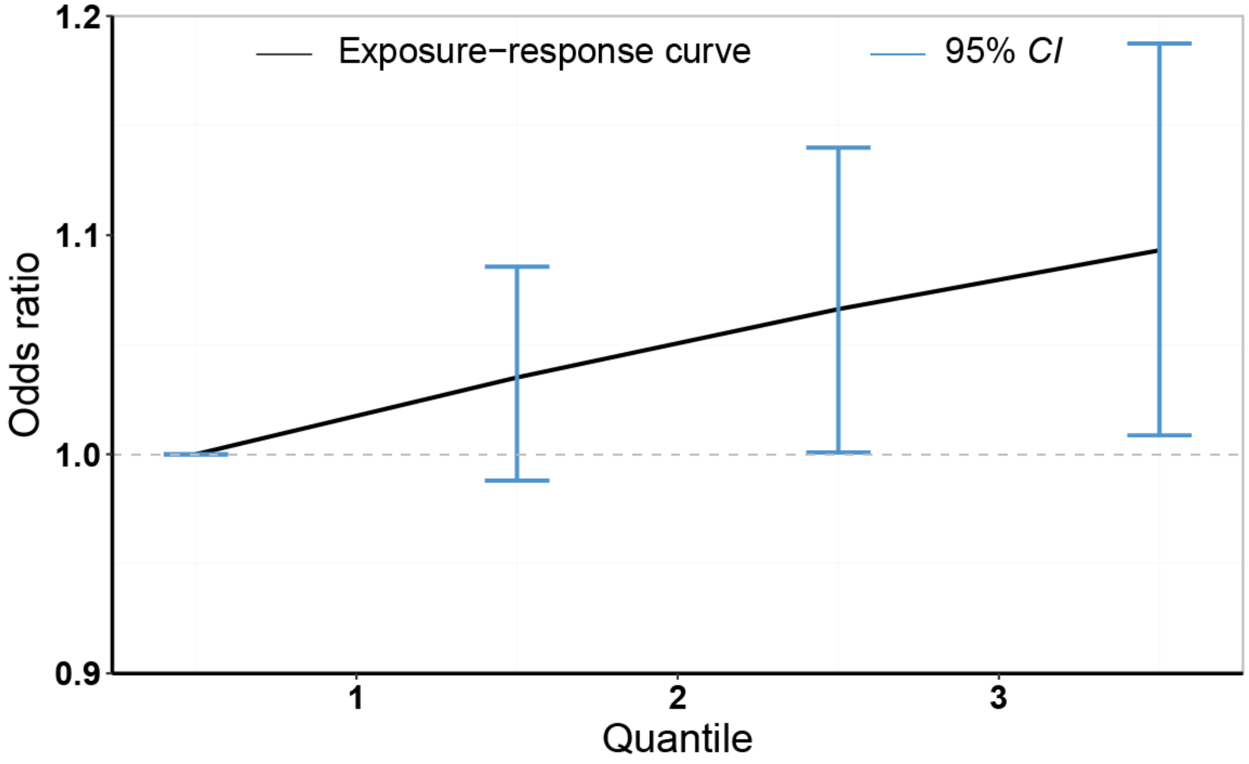


Figure 4. Excess risk (as a percentage) of drowning mortality per quartile increase in mixed daily mean temperature, relative humidity, sunlight duration, and rainfall in 71 cities.

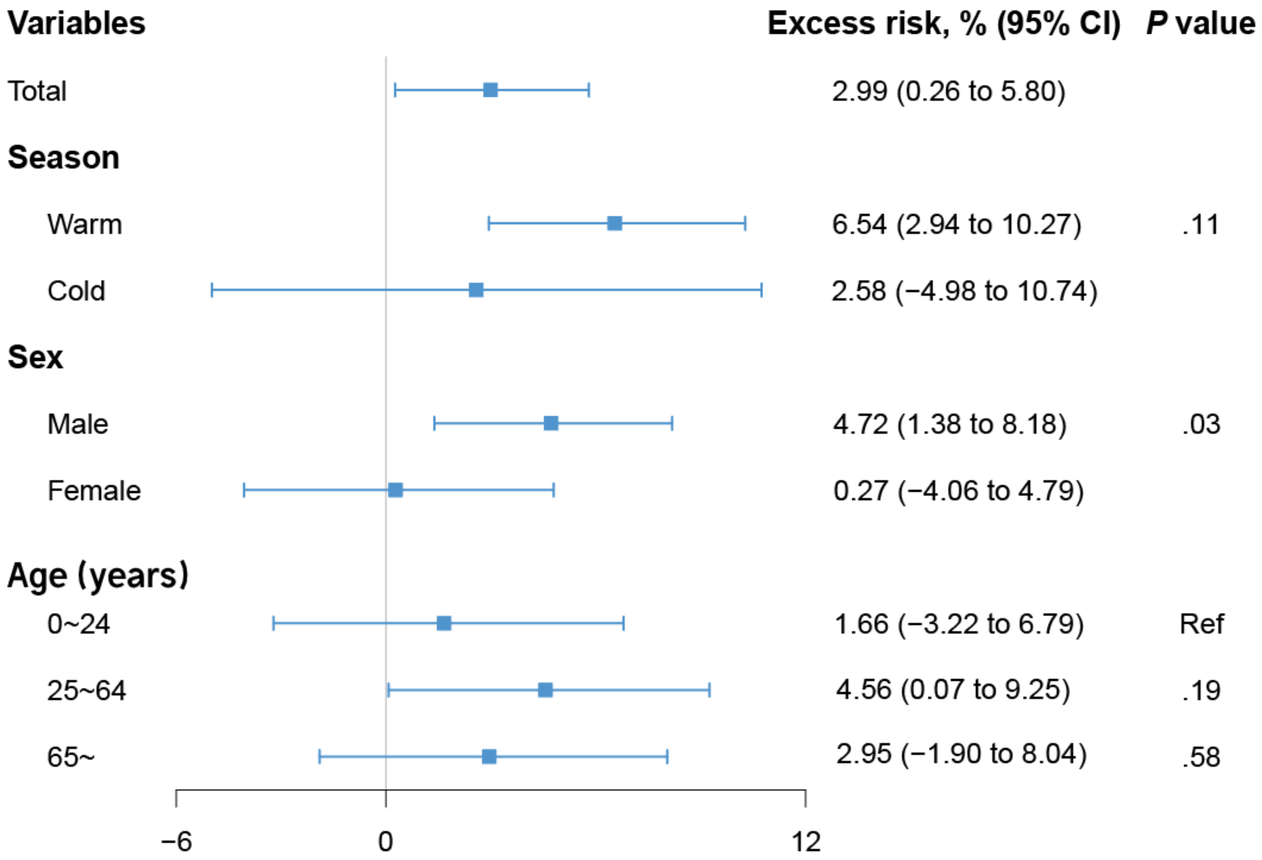
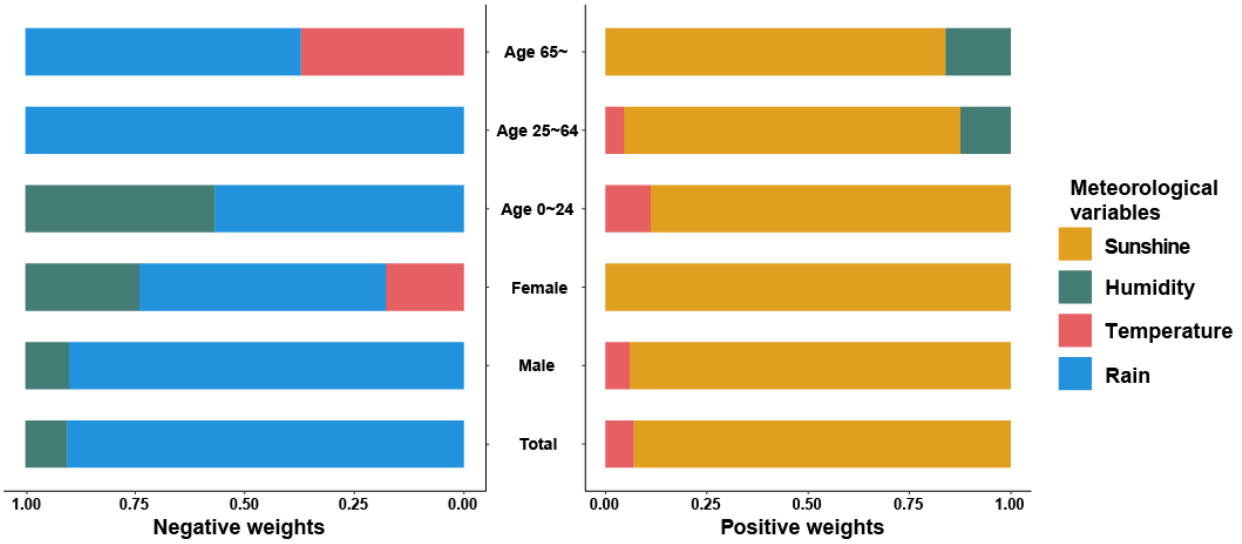


Figure 5. Weights representing the proportion of the positive or negative partial effect of each meteorological factor in quantile g-computation model by sex and age group.



Discussion

Principal Findings

In this study, we examined the combined effect of daily MT, RH, DS, and rainfall on drowning mortality in China. We found that for every quartile increase in meteorological factors, the risk of drowning death in the population increased by 2.99%. We also observed that the 4 meteorological factors had a positively combined effect on drowning mortality in all subgroups except those aged 0 to 24 years; adults aged 25 to 64 years may be more susceptible to mixed exposure to the 4 meteorological factors. Moreover, we showed that DS, rather than MT, was the most important meteorological factor contributing to drowning mortality. These findings have important public health implications for developing an early warning system to reduce drowning mortality in the context of global warming.

Results Interpretation and Implications

An increasing number of epidemiological studies have examined the relationship between exposure to single meteorological factors and the risk of drowning. Consistent with our findings, a study conducted by Chauvin et al [35] reported an increase in the risk of drowning death as ambient temperature rose. When the human body is exposed to high temperatures, the blood vessels in the skin dilate rapidly and sweating increases to accelerate the discharge of excess heat; when entering the water, heavy sweating could lead to electrolyte deficiencies, and swimming in water at this time may lead to dangerous reactions, such as convulsions [36]. Therefore, high temperatures increase the risk of drowning. Additionally, alcohol consumption at high temperatures is an important risk factor for drowning [37,38], as this could affect physiological reactions and behavioral patterns, thereby increasing the risk of drowning. A prior study reported that alcohol consumption was positively correlated with MT [39]. Individuals that consume alcohol have an increased risk of drowning due to greater skin blood flow and sweating. Also, alcohol can cross the human cerebrovascular barrier directly and affect mental and cognitive performance, and the heat sensation associated with alcohol consumption may increase exposure to water [40]. For instance, an Australian study reported that there was a higher tendency to see people consuming alcohol alongside rivers on hot days [41].

To date, we have not found any study on the association of RH with drowning mortality. Among prior studies, RH is often treated as a confounding variable in the model fitting [42,43]. In our study, we observed that RH was negatively associated with the risk of drowning in an individual analysis of meteorological factors and negatively contributed to drowning mortality in most subgroups in a combined-effect analysis. The underlying mechanism of this finding is unclear and requires further in-depth studies in the future.

In terms of DS, existing studies mainly focus on sunlight exposure and skin cancer or suicide [44-46]. The combined effect analysis in our study shows that DS was the most important weather factor in increasing drowning mortality in all subgroups. Possible explanations for this are that (1) longer daylight hours may be associated with higher temperatures, thereby increasing opportunities for water activities, and (2) light has the neurobiological effect of interacting with the brain serotonin system and may influence serotonin-related behavior [46]. Finally, when daylight hours are short, people are less likely to be active outside and are likely to reduce their contact with bodies of water, which in turn may lead to a decreased incidence of drowning.

We further observed that rainfall was a negative factor for unintentional drowning risk in the combined-effect analysis. Previous studies have shown a negative correlation between rainfall and physical activity and a positive association with sedentary time during weekdays [47]. Therefore, rainfall might make it more difficult for people to work outdoors, which in turn may reduce the chance of contact with bodies of water. In contrast, Murray and Carter [48] showed that high rainfall (180.55 mm to 202.25 mm) was associated with increased drowning deaths in Fiji. The rainfall amount in this study was less than 35 mm, which is more likely to reduce people’s willingness to go outside; thus, we observed a reduced risk of drowning. In addition, the deaths in this study did not include flood-related drownings or disaster drownings.

In the stratified analysis, we further observed that males exhibited higher risk than females when exposed to meteorological mixed factors, which is consistent with previous studies. For example, the risk of drowning death is 1.67 times higher in men than women on hot days [14]. Compared to females, males are more likely to work in hot outdoor environments, resulting in potentially greater access to bodies of water and higher risk of drowning. We also observed that for people aged 0 to 24 years, there was no statistically significant association between mixed exposure to the 4 meteorological factors and drowning. This is contrary to previous studies. For instance, a study in Shanghai reported that drowning deaths among children aged 0 to 14 years mainly occurred in the hot season, while other age groups showed less seasonality in drowning [49]. Possible reasons for this are that the young population is more likely to be influenced by individual personality traits (eg, an active and strong spirit for exploring outside) and behavioral patterns and less influenced by external weather factors and the environment; on the other hand, children also end up in bodies of water due to unintentional falls and fights. Meteorological exposure mixture had the greatest effect on people aged 25 to 64 years; one possible explanation for this finding was that adults were more likely to consume alcohol in hot weather, and drinking alcohol increases exposure to water [50]. Moreover, this age group is the main workforce, with a high chance of occupational exposure to bodies of water, which may increase drowning deaths [51].

Several previous studies have estimated the risk of drowning associated with exposure to a single meteorological factor. However, people are exposed to a variety of meteorological factors simultaneously, so it is important to use an exposure mixture approach to measure the combined effects of weather factors on human health. In this study, we estimated the combined effects of 4 weather factors using a quantile g-computation model. This model helps design potential public health interventions for specific sources of exposure [19].

Multiple sectors in the government, such as education departments, health departments, and meteorological departments, should collaborate to develop an early warning system taking into account DS as a potentially critical weather factor to reduce the risk of drowning in the high-risk season. In addition, in areas with a high risk of drowning, governments should provide adequate and safe swimming places to reduce the possibility of people going into dangerous waters for recreation.

Limitations

This study has several limitations that must be noted. First, due to the lack of individual exposure data, we used meteorological data obtained by spatial interpolation as a proxy in each city, and there may be some exposure misclassification bias. Second, although this study is based on 5 provinces, which are to some extent nationally representative, extrapolation of the results needs to be done with caution. Third, drowning is influenced by family factors and socioeconomic factors, which were not included in this study. Finally, we only included unintentional drowning (ICD-10 codes W65-74) in this study, which excludes drowning due to cataclysm (ICD-10 codes X34-X39) and transportation accidents (ICD-10 codes V01-V99). This may not adequately reflect the true number of drowning deaths [52].

Conclusion

This study shows that mixed exposure to the 4 weather factors was positively associated with drowning mortality, which was primarily driven by DS and MT. This finding implies that it is necessary to incorporate sunshine hours and temperature into early warning systems for drowning in the future.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFA0606200), the National Natural Science Foundation of China (42275187), and the Foreign Expert Program of the Ministry of Science and Technology (G2022199006L).

Data Availability

The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

YL and XD are joint first authors. WM had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. WM and TL conceived the original idea and supervised the study. YL and Z Li conducted the data analysis and drafted the manuscript. YL, XD, and Z Li participated in manuscript writing. WG, ZH, RM, CZ, MY, BH, L Lin, JX, JZ, DJ, YX, and L Lv conducted the investigation and contributed to resource organization and data curation. CH was responsible for project administration, resource organization, and data curation. JH, SZ, and Z Lin participated in results review. WM, TL, and GH participated in revision. All authors reviewed the final manuscript. WM and TL are joint corresponding authors and have contributed equally to this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preliminary analysis of air pressure and wind speed in a generalized additive model.
[PNG File , 252 KB - publichealth_v9i1e46792_app1.png]

Multimedia Appendix 2
Spearman correlations between meteorological factors and PM2.5.
[PNG File , 101 KB - publichealth_v9i1e46792_app2.png]

Multimedia Appendix 3
Sensitivity analysis, changing the degrees of freedom in a spline function with mean temperature, relative humidity, sunlight duration, and rainfall in a generalized additive model.
[PNG File , 461 KB - publichealth_v9i1e46792_app3.png]

Multimedia Appendix 4
The modification effects of city-level characteristics.
[PNG File , 346 KB - publichealth_v9i1e46792_app4.png]

Multimedia Appendix 5
Sensitivity analysis with a quantile g-computation model.
[DOCX File , 20 KB - publichealth_v9i1e46792_app5.docx]

References

1. van Beeck EF, Branche C, Szpilman D, Modell J, Bierens JJLM. A new definition of drowning: towards documentation and prevention of a global public health problem. Bull World Health Organ 2005 Nov;83(11):853-856 [FREE Full text] [Medline: 16302042]
2. Drowning. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/drowning [accessed 2022-09-15]
3. Franklin RC, Peden AE, Hamilton EB, Bisignano C, Castle CD, Dingels ZV, et al. The burden of unintentional drowning: global, regional and national estimates of mortality from the Global Burden of Disease 2017 Study. Inj Prev 2020 Oct 28;26(Supp 1):i83-i95 [FREE Full text] [doi: 10.1136/injuryprev-2019-043484] [Medline: 32079663]
4. Drowning facts. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drowning/facts/index.html [accessed 2022-09-15]
5. Global Burden of Disease Study 2019 (GBD 2019) data resources. Institute for Health Metrics and Evaluation. URL: https://ghdx.healthdata.org/gbd-2019 [accessed 2022-09-15]
6. Willcox-Pidgeon S, Franklin RC, Leggat PA, Devine S. Epidemiology of unintentional fatal drowning among migrants in Australia. Aust N Z J Public Health 2021 Jun;45(3):255-262 [FREE Full text] [doi: 10.1111/1753-6405.13102] [Medline: 33900674]
7. Yang L, Nong Q, Li C, Feng Q, Lo SK. Risk factors for childhood drowning in rural regions of a developing country: a case-control study. Inj Prev 2007 Jun;13(3):178-182 [FREE Full text] [doi: 10.1136/ip.2006.013409] [Medline: 17567973]
8. Zhu Y, Jiang X, Li H, Li F, Chen J. Mortality among drowning rescuers in China, 2013: a review of 225 rescue incidents from the press. BMC Public Health 2015 Jul 10;15(1):631 [FREE Full text] [doi: 10.1186/s12889-015-2010-0] [Medline: 26156246]
9. Culp W, Brewster B, Wernicki P. Drowning: lifeguard rescue and resuscitation. Anesthesiology 2010 Jan;112(1):245; author reply 246 [FREE Full text] [doi: 10.1097/ALN.0b013e3181c2951b] [Medline: 20032707]
10. Szpilman D, de Barros Oliveira R, Mocellin O, Webber J. Is drowning a mere matter of resuscitation? Resuscitation 2018 Aug;129:103-106. [doi: 10.1016/j.resuscitation.2018.06.018] [Medline: 29928958]
11. Brenner R. Prevention of drowning in infants, children, and adolescents. Pediatrics 2003 Aug;112(2):440-445. [doi: 10.1542/peds.112.2.440] [Medline: 12897306]
12. Denny S, Quan L, Gilchrist J, McCallin T, Shenoi R, Yusuf S, Council on Injury, Violence, and Poison Prevention. Prevention of drowning. Pediatrics 2021 Aug;148(2):e20190850. [doi: 10.1542/peds.2021-052227] [Medline: 34253571]
13. Peden AE, Franklin RC, Willcox-Pidgeon S. Media reporting of summer drowning: A partial picture, useful for advocacy. Health Promot J Austr 2020 Sep 28;31(3):491-496. [doi: 10.1002/hpja.297] [Medline: 31562788]
14. Fralick M, Denny CJ, Redelmeier DA. Drowning and the influence of hot weather. PLoS One 2013;8(8):e71689 [FREE Full text] [doi: 10.1371/journal.pone.0071689] [Medline: 23977112]
15. Mora C, Dousset B, Caldwell IR, Powell FE, Geronimo RC, Bielecki C, et al. Global risk of deadly heat. Nature Clim Change 2017 Jun 19;7(7):501-506. [doi: 10.1038/nclimate3322]
16. Lancet. Flooding, drowning, and health equity: The Lancet Regional Health - Western Pacific August editorial. Lancet Reg Health West Pac 2021 Aug;13:100264 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100264] [Medline: 34527998]

17. Falk M. Impact of weather conditions on tourism demand in the peak summer season over the last 50years. Tour Manag Perspect 2014 Jan;9:24-35. [doi: 10.1016/j.tmp.2013.11.001]
18. Rosato I, Zare Jeddi M, Ledda C, Gallo E, Fletcher T, Pitter G, et al. How to investigate human health effects related to exposure to mixtures of per- and polyfluoroalkyl substances: A systematic review of statistical methods. Environ Res 2022 Apr 01;205:112565. [doi: 10.1016/j.envres.2021.112565] [Medline: 34915031]
19. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 2020 Apr;128(4):47004 [FREE Full text] [doi: 10.1289/EHP5838] [Medline: 32255670]
20. China Meteorological Science Data Center. URL: http://data.cma.cn/ [accessed 2022-09-15]
21. ANUSPLIN version 4.4 user guide. Australian National University. URL: https://fennerschool.anu.edu.au/research/products/anusplin [accessed 2023-06-30]
22. Liu T, Zhou C, Zhang H, Huang B, Xu Y, Lin L, et al. Ambient temperature and years of life lost: a national study in China. Innovation (Camb) 2021 Feb 28;2(1):100072 [FREE Full text] [doi: 10.1016/j.xinn.2020.100072] [Medline: 34557729]
23. National urban air quality real-time release platform. URL: https://air.cnemc.cn:18007/ [accessed 2022-09-15]
24. Liu T, Meng H, Yu M, Xiao Y, Huang B, Lin L, et al. Urban-rural disparity of the short-term association of PM with mortality and its attributable burden. Innovation (Camb) 2021 Nov 28;2(4):100171 [FREE Full text] [doi: 10.1016/j.xinn.2021.100171] [Medline: 34778857]
25. Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H. Referent selection in case-crossover analyses of acute health effects of air pollution. Epidemiology 2001 Mar;12(2):186-192. [doi: 10.1097/00001648-200103000-00010] [Medline: 11246579]
26. Lee H, Myung W, Kim H, Lee E, Kim H. Association between ambient temperature and injury by intentions and mechanisms: A case-crossover design with a distributed lag nonlinear model. Sci Total Environ 2020 Dec 01;746:141261. [doi: 10.1016/j.scitotenv.2020.141261] [Medline: 32745866]
27. Kim Y, Ng CFS, Chung Y, Kim H, Honda Y, Guo YL, et al. Air pollution and suicide in 10 cities in northeast Asia: A time-stratified case-crossover analysis. Environ Health Perspect 2018 Mar 06;126(3):037002 [FREE Full text] [doi: 10.1289/EHP2223] [Medline: 29529596]
28. Ravindra K, Rattan P, Mor S, Aggarwal AN. Generalized additive models: Building evidence of air pollution, climate change and human health. Environ Int 2019 Nov;132:104987 [FREE Full text] [doi: 10.1016/j.envint.2019.104987] [Medline: 31398655]
29. Wood S. Generalized Additive Models: An Introduction with R. 2nd ed. Boca Raton, FL: Chapman and Hall/CRC; 2017.
30. Keil A. qgcomp: Quantile G-Computation. Cran R Project. URL: https://CRAN.R-project.org/package=qgcomp [accessed 2022-09-15]
31. Li G, Liu J, Lu H, Hu W, Hu M, He J, et al. Multiple environmental exposures and obesity in eastern China: An individual exposure evaluation model. Chemosphere 2022 Jul;298:134316. [doi: 10.1016/j.chemosphere.2022.134316] [Medline: 35302002]
32. Niehoff NM, Keil AP, O'Brien KM, Jackson BP, Karagas MR, Weinberg CR, et al. Metals and trace elements in relation to body mass index in a prospective study of US women. Environ Res 2020 May;184:109396 [FREE Full text] [doi: 10.1016/j.envres.2020.109396] [Medline: 32209500]
33. Sun Y, Li X, Benmarhnia T, Chen J, Avila C, Sacks DA, et al. Exposure to air pollutant mixture and gestational diabetes mellitus in Southern California: Results from electronic health record data of a large pregnancy cohort. Environ Int 2022 Jan;158:106888 [FREE Full text] [doi: 10.1016/j.envint.2021.106888] [Medline: 34563749]
34. Schenker N, Gentleman JF. On judging the significance of differences by examining the overlap between confidence intervals. Am Stat 2001 Aug;55(3):182-186. [doi: 10.1198/000313001317097960]
35. Chauvin M, Kosatsky T, Bilodeau-Bertrand M, Gamache P, Smargiassi A, Auger N. Hot weather and risk of drowning in children: Opportunity for prevention. Prev Med 2020 Jan;130:105885. [doi: 10.1016/j.ypmed.2019.105885] [Medline: 31705939]
36. Maughan RJ, Shirreffs SM. Muscle cramping during exercise: Causes, solutions, and questions remaining. Sports Med 2019 Dec;49(Suppl 2):115-124 [FREE Full text] [doi: 10.1007/s40279-019-01162-1] [Medline: 31696455]
37. Driscoll TR, Harrison JA, Steenkamp M. Review of the role of alcohol in drowning associated with recreational aquatic activity. Inj Prev 2004 Apr;10(2):107-113 [FREE Full text] [doi: 10.1136/ip.2003.004390] [Medline: 15066977]
38. Pajunen T, Vuori E, Vincenzi FF, Lillsunde P, Smith G, Lunetta P. Unintentional drowning: Role of medicinal drugs and alcohol. BMC Public Health 2017 May 19;17(1):388 [FREE Full text] [doi: 10.1186/s12889-017-4306-8] [Medline: 28521790]
39. Hagström H, Widman L, von Seth E. Association between temperature, sunlight hours and alcohol consumption. PLoS One 2019;14(9):e0223312 [FREE Full text] [doi: 10.1371/journal.pone.0223312] [Medline: 31568515]
40. Wasielewski JA, Holloway FA. Alcohol's interactions with circadian rhythms. A focus on body temperature. Alcohol Res Health 2001;25(2):94-100 [FREE Full text] [Medline: 11584555]

41. Peden AE, Franklin RC, Leggat PA. Breathalysing and surveying river users in Australia to understand alcohol consumption and attitudes toward drowning risk. BMC Public Health 2018 Dec 19;18(1):1393 [FREE Full text] [doi: 10.1186/s12889-018-6256-1] [Medline: 30567588]
42. Barnett A, Tong S, Clements A. What measure of temperature is the best predictor of mortality? Environ Res 2010 Aug;110(6):604-611 [FREE Full text] [doi: 10.1016/j.envres.2010.05.006] [Medline: 20519131]
43. Zhang K, Li Y, Schwartz JD, O'Neill MS. What weather variables are important in predicting heat-related mortality? A new application of statistical learning methods. Environ Res 2014 Jul;132:350-359 [FREE Full text] [doi: 10.1016/j.envres.2014.04.004] [Medline: 24834832]
44. Lindqvist PG. The winding path towards an inverse relationship between sun exposure and all-cause mortality. Anticancer Res 2018 Feb;38(2):1173-1178. [doi: 10.21873/anticanres.12337] [Medline: 29374755]
45. Lindqvist P. On the relationship between sun exposure and all-cause mortality. Adv Exp Med Biol 2020;1268:115-122. [doi: 10.1007/978-3-030-46227-7_5] [Medline: 32918215]
46. Vyssoki B, Kapusta ND, Praschak-Rieder N, Dorffner G, Willeit M. Direct effect of sunshine on suicide. JAMA Psychiatry 2014 Nov;71(11):1231-1237. [doi: 10.1001/jamapsychiatry.2014.1198] [Medline: 25208208]
47. Harrison F, Jones AP, Bentham G, van Sluijs EM, Cassidy A, Griffin SJ. The impact of rainfall and school break time policies on physical activity in 9-10 year old British children: a repeated measures study. Int J Behav Nutr Phys Act 2011 May 24;8(1):47 [FREE Full text] [doi: 10.1186/1479-5868-8-47] [Medline: 21609441]
48. Murray K, Carter P. Fatal drownings in Fiji. Asia Pac J Public Health 2017 Jan;29(1):28-34. [doi: 10.1177/1010539516685610] [Medline: 28198649]
49. Yi-chen C, Xiao-pan L, Chen Y, Li-peng HA. Analysis on the epidemiological characteristics and trend of traffic fatality data among residents in Pudong New Area of Shanghai, 2002-2015. Chin J Dis Control Prev 2017;21(6):602-606. [doi: 10.16462/j.cnki.zhjbkz.2017.06.015]
50. Clemens T, Tamim H, Rotondi M, Macpherson AK. A population based study of drowning in Canada. BMC Public Health 2016 Jul 13;16:559 [FREE Full text] [doi: 10.1186/s12889-016-3221-8] [Medline: 27411984]
51. Yang L, Lin Q, Wang Y, Tong Z. Mortality and potential life lost of drowning in Zhoushan residents. J Prev Med. Published online 2019:1116-1119. [doi: 10.19485/j.cnki.issn2096-5087.2019.11.009]
52. Peden AE, Franklin RC, Mahony AJ, Scarr J, Barnsley PD. Using a retrospective cross-sectional study to analyse unintentional fatal drowning in Australia: ICD-10 coding-based methodologies verses actual deaths. BMJ Open 2017 Dec 21;7(12):e019407 [FREE Full text] [doi: 10.1136/bmjopen-2017-019407] [Medline: 29273670]

Abbreviations

ANUSPLIN: Australian National University Splines
DS: duration of sunlight
ER: excess risk
GAM: generalized additive model
ICD-10: International Classification of Diseases, 10th Edition
MT: mean temperature
OR: odds ratio
PM: particulate matter
RH: relative humidity
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 25.02.23; peer-reviewed by F Huang, A Peden, Z Zhang; comments to author 18.04.23; revised version received 05.05.23; accepted 15.06.23; published 20.07.23.

Please cite as:
Liu Y, Dong X, Li Z, Zhu S, Lin Z, He G, Gong W, Hu J, Hou Z, Meng R, Zhou C, Yu M, Huang B, Lin L, Xiao J, Zhong J, Jin D, Xu Y, Lv L, Huang C, Liu T, Ma W
The Combined Effects of Short-Term Exposure to Multiple Meteorological Factors on Unintentional Drowning Mortality: Large Case-Crossover Study
JMIR Public Health Surveill 2023;9:e46792
URL: https://publichealth.jmir.org/2023/1/e46792
doi:10.2196/46792
PMID:37471118

Original Paper

The Association Between Solid Fuel Use and Visual Impairment Among Middle-Aged and Older Chinese Adults: Nationwide Population-Based Cohort Study

Qingqing Jiang[1*], PhD; Shiqi Wang[1*], MSc; Hao Zhang[1], MSc; Yan Guo[1], MSc; Yiling Lou[1], PhD; Shen Huang[1], PhD; Qiqi You[1], MSc; Shiyi Cao[1], PhD

School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Shiyi Cao, PhD
School of Public Health
Tongji Medical College
Huazhong University of Science and Technology
No 13 Hangkong Road
Wuhan, 430030
China
Phone: 86 027 83657984
Email: caoshiyi@hust.edu.cn

Abstract

Background: Indoor air pollution has been reported to have adverse effects on the eye; however, the health effects of exposure to cooking with solid fuels on visual impairment remain unclear in China.

Objective: We aimed to examine the association between cooking with solid fuels and visual impairment, including distance visual impairment (DVI) and near visual impairment (NVI).

Methods: Data were obtained from the China Health and Retirement Longitudinal Study, a nationwide survey of adults aged over 45 years who were enrolled in 2011 (Wave 1) and followed up in Wave 2 (2013), Wave 3 (2015), and Wave 4 (2018). We used Cox proportional hazards models to determine the association between solid fuels use and visual impairment. Additionally, the impact of switching cooking fuel types on vision function were examined through wave-specific data analysis (Wave 1 and Wave 4). Interaction and subgroup analyses were performed to explore the potential effect modifiers. Data were collected using the stratified multistage random sampling method and further analyzed using SPSS 27.0 and R 4.2.1 statistical software packages.

Results: A total of 9559 middle-aged and older Chinese adults without visual impairment at baseline were included in the study, with 51.2% (n=4914) of the participants reporting that they cooked with solid fuels. During the follow-up period, 2644 (27.5%) and 3130 (32.6%) participants developed DVI and NVI, respectively. Compared with the clean fuel users, participants who cooked with solid fuels had a higher risk of DVI (hazards ratio [HR] 1.38, 95% CI 1.28-1.50) and NVI (HR 1.18, 95% CI 1.10-1.27). In addition, switching the cooking fuel type from clean to solid fuels was associated with an elevated risk of DVI (HR 1.51, 95% CI 1.15-1.98) and NVI (HR 1.39, 95% CI 1.06-1.82) compared to persistently using clean fuels during the follow-up period, although no protective effect of switching from solid to clean fuels on NVI was found (P=.52). In subgroup analysis, we found that cooking with solid fuels increased the risk of DVI in participants younger than 65 years (HR 1.41, 95% CI 1.28-1.55), men (HR 1.45, 95% CI 1.28-1.65), urban residents (HR 1.41, 95% CI 1.08-1.75), and smokers (HR 1.43, 95% CI 1.25-1.64). By contrast, negative effects of cooking with solid fuels on NVI were found in nonsmokers (HR 1.21, 95% CI 1.11-1.33) and urban residents (HR 1.20, 95% CI 1.10-1.37).

Conclusions: Cooking with solid fuels was associated with an increased risk of visual impairment among middle-aged and older Chinese adults. These findings indicate that promoting the utilization of clean fuels is conducive to reducing the burden of visual impairment for the public.

(JMIR Public Health Surveill 2023;9:e43914) doi:10.2196/43914

KEYWORDS

indoor air pollution; solid fuel; visual impairment; cohort study; Chinese

Introduction

Incomplete and inefficient combustion of solid fuels (including biomass and coal) is the primary cause of indoor air pollution, affecting approximately half the global population [1]. Prolonged exposure to indoor air pollution has many negative impacts on individuals' health, such as causing respiratory disease, cancer, and eye problems, and is the foremost cause of premature death in developing countries [2]. Among the many human organs that could be affected by indoor air pollution, the eyes are directly exposed to the emissions from solid fuel combustion, including high levels of fine particulate matter ($PM_{2.5}$) and carbon monoxide. These emissions would stimulate the production of reactive oxygen species, which can seriously impair the viability of eye cells and accelerate oxidation of the lens [3]. Moreover, the hazardous materials resulting from solid fuel combustion can also affect the body's cardiovascular circulation system, leading to increased intraocular pressure through the aqueous humor.

Eye problems such as tearing, redness, itching, and stinging are widely reported to be linked to indoor air pollution [3,4]. A recent large cohort study provided evidence that prolonged exposure to indoor air pollution, which was mainly attributed to cooking with solid fuels, would increase the risk of major eye diseases (eg, conjunctiva disorders, cataracts) [5]. Moreover, long-term exposure to eye problems and eye diseases could result in severe visual impairment and even blindness [6-8]. Based on these findings [3-8], it is hypothesized that indoor air pollution exposure may be an environmental risk factor associated with visual impairment.

An estimated 2.2 billion people globally have visual impairments, with nearly half of these cases being preventable or untreated [9]. Untreated vision loss could have significant and lasting effects on a person's ability to perform daily activities and access public services [10]. In China, the most populous country in the word, 30%-50% of adults over the age of 80 years and 7%-20% of adults over the age of 50 years were estimated to have visual impairments [11], ranking second in years lived with disability among all health impairments. As the population is aging rapidly, the burden of visual impairment is predicted to continue to rise [10,12].

Given the widespread use of solid fuels in China [13], evidence on the association of solid fuel use and visual impairment is urgently needed for policy makers to take effective actions to relieve the burden of visual impairment by promoting cleaner household energy.

To fill the above knowledge gap on the link between solid fuel use and visual impairment among Chinese adults, we performed this study to assess the longitudinal association between self-reported cooking with solid fuels and visual impairment in a nationally representative sample obtained from the China Health and Retirement Longitudinal Study (CHARLS). This prospective cohort study assessed the impact of indoor air pollution on visual impairment among middle-aged and older Chinese adults from 2011 to 2018 for the first time and also examined the effects of switching cooking fuel types on vision function over time.

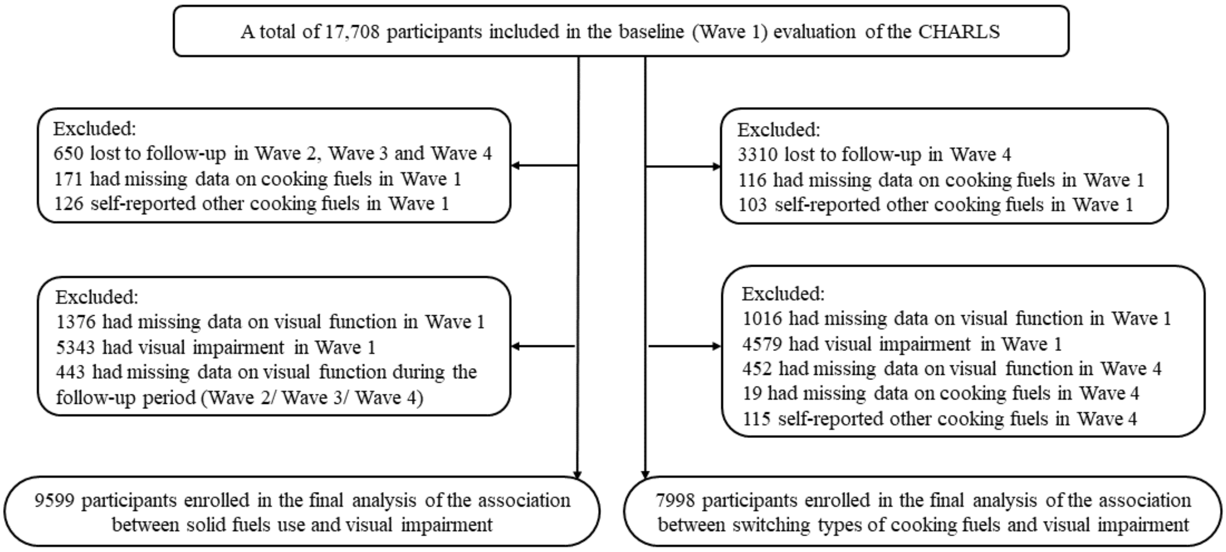
Methods

Study Population

This nationwide population-based cohort study utilized data from the CHARLS, which was implemented by the National School of Development of Peking University. An unbiased and representative sample was obtained by proportional and multistage stratified sampling [14]. In the first stage, all counties in China were stratified by region. Regions were divided by urban districts or rural counties and per capita statistics on gross domestic product. Using the probability-proportional-to-size sampling technique, 150 county-level units were randomly chosen from the sampling frame. In the second stage, three primary sampling units (PSUs) were selected in each county based on a probability proportional to the population size. In the third stage, a random sample of 24 households was selected among all the households mapped in each selected PSU. Finally, for each selected household, one resident aged ≥45 years and their spouse were randomly selected as the participants in the survey. Briefly, the CHARLS survey included Chinese residents from 450 communities covering 150 county-level units in 28 provinces. From May 2011 to March 2012, a total of 17,708 residents with a response rate above 80% participated in the baseline survey (Wave 1). CHARLS respondents are followed every 2 years through a face-to-face computer-assisted personal interview to obtain individual information on sociodemographic characteristics, lifestyle behaviors, indoor air pollution, health status, and other factors of interest. Interviewers helped respondents who had sight problems by reading the questions in the standardized questionnaire during the interview process. To date, follow-up survey data are available for Wave 2 (2013), Wave 3 (2015-2016), and Wave 4 (2018) of CHARLS.

We used the data from all four waves to examine the relationship between solid fuel use and visual impairment; a total of 9559 individuals were selected based on the following exclusion criteria: (1) 650 lost to follow-up in Wave 2, Wave 3, and Wave 4; (2) 171 had missing data on cooking fuels in Wave 1; (3) 126 self-reported other cooking fuels in Wave 1; (4) 1376 had missing data on visual function in Wave 1; (5) 5343 had visual impairment in Wave 1; and (6) 443 had missing data on visual function during the follow-up period (Wave 2/Wave 3/Wave 4) (Figure 1).

Figure 1. Flowchart of the selection process of participants. CHARLS: China Health and Retirement Longitudinal Study.



In addition, we used the wave-specific data of Wave 1 and Wave 4 to examine the association between switching the fuel type and visual impairment; 7998 individuals were selected based on the following exclusion criteria: (1) 3310 lost to follow-up in Wave 4; (2) 116 had missing data on cooking fuels in Wave 1; (3) 103 self-reported other cooking fuels in Wave 1; (4) 1016 had missing data on visual function in Wave 1; (5) 4579 had visual impairment in Wave 1; (6) 452 had missing data on visual function in Wave 4; (7) 19 had missing data on cooking fuels in Wave 4; and (8) 115 self-reported other cooking fuels in Wave 4.

Vision Assessment

Individuals' distance vision was assessed by self-reported visual status when recognizing friends across the street (while wearing corrective lenses or glasses, if applicable) and personal near vision was assessed by self-reported visual status when reading regular newspapers (while wearing glasses or corrective lenses, if applicable). Possible responses to these questions were "poor," "fair," "good," "very good,", and "excellent." Those reporting "poor" for near and distance vision were defined as having near visual impairment (NVI) and distance visual impairment (DVI), respectively. Those reporting "excellent," "very good," "good," or "fair" were defined as having no visual impairment (reference group). Visual impairment assessment and categorization were performed in accordance with previous studies [15-17].

Assessment of Cooking With Solid Fuels

The question "What are your main sources of cooking fuels?" was used to assess indoor cooking fuel types. At baseline, reporting "coal," "crop residue," and "wood" was defined as cooking with solid fuels. The main types of solid fuels were classified into coal and biomass ("crop residue" and "wood"). Reporting "liquefied petroleum," "natural gas," "electricity," and "marsh gas" was defined as cooking with clean fuels. We compared data from respondents who reported cooking with clean fuels to those of the solid fuel users. In addition, we collected information on self-reported cooking fuel types in Wave 4, which had an additional choice of "do not cook" compared to the answers regarding the baseline question. We categorized "do not cook" as clean fuel, and assessed the association of switching cooking fuel types with visual impairment. Therefore, participants could be classified as switching from solid fuels to clean fuels or from clean fuels to solid fuels.

Covariates

Baseline demographics (including age, gender, marital status, education level, residence) and health-related status (including smoking status, alcohol drinking, sleep duration, BMI, and chronic diseases) were included as covariates in our analyses. We classified cohabitation and marriage as married. Education level was divided into illiteracy (0 years) and others (≥1 year). Residence was divided into rural and urban areas [18]. BMI was calculated as weight divided by the square of height (kg/m^2) [19]. Alcohol drinking was classified into drinker and nondrinker. Smoking status was classified into smoker and nonsmoker. Dyslipidemia, chronic lung disease, hypertension, heart disease, diabetes, and cancer were identified by self-reported physician diagnoses.

Statistical Analysis

The baseline characteristics are described according the type of cooking fuels. Continuous variables were summarized as mean (SD) and were compared with independent *t* tests. Categorical variables were summarized as frequencies (percentages) and were compared with the χ^2 test. A Cox proportional hazard model was performed to assess the effects of solid fuel use on visual impairment. We also compared data of participants who switched from clean fuels to solid fuels or from solid fuels to clean fuels with those who persistently used clean or solid fuels through wave-specific analysis. The proportional hazard assumption was examined using the Schoenfeld residual test. Model 1 was an unadjusted model. Model 2 was adjusted for age, gender, BMI, marital status, education level, residence, alcohol drinking, smoking status, and sleep duration. Model 3 was further adjusted for hypertension, dyslipidemia, diabetes, cancer, chronic lung disease, and heart disease based on Model 2. The percentage of missing covariates was less than 5% for all covariates, which conformed to an arbitrary missingness pattern. We used multiple imputation to account for these missing values. The Markov chain Monte Carlo method was used to analyze the complete data set and interpolate each missing value five times to obtain five complete data sets. Statistical analysis was then performed on each filled data set separately and the combined results of the five analyses were used to select the data set with the largest relative efficiency value. Additionally, we performed subgroup analyses and interactions analysis stratified by age, sex, residence area, and smoking status to assess potential effect modifications. The sensitivity analysis was conducted by excluding participants who reported not cooking in Wave 4 to ensure that our findings were robust. Cox proportional hazards model results are presented as hazard ratios with 95% CIs. Statistical analyses were performed using SPSS 27.0 and R 4.2.1 software. Two-tailed *P* values less than .05 were considered statistically significant.

Ethical Considerations

The CHARLS was approved by the Research Ethics Committee of Peking University (IRB00001052-11015). This survey was anonymous and the answers are protected by privacy law. Written informed consent clarifying the study purposes was obtained from each participant before completing the interview.

Results

Characteristics of the Study Participants

A total of 9559 participants without visual impairment at baseline were included. The average age of included participants was 57.7 (SD 9.5) years and 50.3% of them were women. Overall, 51.2% (n=4914) of the participants primarily cooked with solid fuels (Table 1), and 15.3% (n=1469), 13.6% (n=1309), 18.4% (n=1770), 1.4% (n=137), 10.8% (n=1035), and 40.5% (n=3879) of the participants reported cooking with liquefied petroleum, natural gas, electricity, marsh gas, coal, and crop residue/wood burning at baseline, respectively. However, only 32.1% of the participants reported cooking with solid fuels in Wave 4, including 5.6% cooking with coal and 26.5% cooking with crop residue/wood burning (see Multimedia Appendix 1). In general, participants who reported cooking with solid fuels at baseline were more likely to be current smokers and reported less sleep time, lower educational levels, and lower BMI compared with clean fuel users. Significantly more participants living in rural areas and those suffering from chronic lung diseases used solid fuels (both *P*<.001). Additionally, significantly more individuals with hypertension, diabetes, and dyslipidemia used clean fuels for cooking rather than solid fuels (*P*=.04, *P*<.001, and *P*<.001, respectively).

Table 1. Baseline characteristics of the participants according to cooking fuel type.

Characteristic	Total (N=9599)	Solid fuels (n=4914)	Clean fuels (n=4685)	*P* value for difference
Age (years), mean (SD)	57.7 (9.5)	58.3 (9.6)	56.9 (9.5)	.18
Sex, n (%)				.57
Male	4768 (49.7)	2455 (50.0)	2313 (49.4)	
Female	4831 (50.3)	2459 (50.0)	2372 (50.6)	
Education (years), n (%)				<.001
0	2250 (23.4)	1508 (30.7)	742 (15.8)	
≥1	7349 (76.6)	3406 (69.3)	3943 (84.2)	
Marital status, n (%)				.37
Married	8596 (89.6)	4387 (89.3)	4209 (89.8)	
Others	1003 (10.4)	527 (10.7)	476 (10.2)	
Residence, n (%)				<.001
Rural	7186 (74.9)	4570 (93.0)	2616 (55.8)	
Urban	2413 (25.1)	344 (7.0)	2069 (44.2)	
BMI (kg/m^2), mean (SD)[a]	23.4 (3.3)	23.1 (3.4)	23.7 (3.2)	<.001
Smoking status, n (%)				<.001
Smoker	3794 (39.5)	2038 (41.5)	1756 (37.5)	
Nonsmoker	5805 (60.5)	2876 (58.5)	2929 (62.5)	
Alcohol consumption, n (%)				.18
Drinker	3297 (34.3)	1656 (33.7)	1641 (35.0)	
Nondrinker	6302 (65.7)	3258 (66.3)	3044 (65.0)	
Sleep duration (hours), n (%)[b]				.001
<7	4470 (46.9)	2367 (48.5)	2103 (45.2)	
≥7	5068 (53.1)	2515 (51.5)	2553 (54.8)	
Hypertension, n (%)[c]				.04
Yes	2099 (22.0)	1030 (21.1)	1069 (22.9)	
No	7453 (78.0)	3851 (78.9)	3602 (77.1)	
Diabetes, n (%)[d]				<.001
Yes	439 (4.6)	174 (3.6)	265 (5.7)	
No	9077 (95.4)	4689 (96.4)	4388 (94.3)	
Dyslipidemia, n (%)[e]				<.001
Yes	826 (8.7)	327 (6.8)	499 (10.8)	
No	8616 (91.3)	4488 (93.2)	4128 (89.2)	
Cancer, n (%)[f]				.50
Yes	79 (0.8)	37 (0.8)	42 (0.9)	
No	9480 (99.2)	4849 (99.2)	4631 (99.1)	
Chronic lung disease, n (%)[g]				<.001
Yes	844 (8.8)	512 (10.5)	332 (7.1)	
No	8718 (91.2)	4378 (89.5)	4340 (92.9)	
Heart disease, n (%)[h]				.23
Yes	986 (10.3)	486 (10.0)	500 (10.7)	

Characteristic	Total (N=9599)	Solid fuels (n=4914)	Clean fuels (n=4685)	*P* value for difference
No	8564 (89.7)	4398 (90.0)	4166 (89.3)	

[a] Data missing from 425 participants.

[b] Data missing from 61 participants.

[c] Data missing from 47 participants.

[d] Data missing from 83 participants.

[e] Data missing from 157 participants.

[f] Data missing from 40 participants.

[g] Data missing from 37 participants.

[h] Data missing from 49 participants.

Association Between Cooking With Solid Fuels and Visual Impairment

During the follow-up period, 2644 (27.5%) and 3130 (32.6%) of the 9599 total participants reported DVI and NVI, respectively (Table 2). The effects of cooking with solid fuels on visual impairment are presented in Table 2. In the crude model, using solid fuels for cooking elevated the risk of DVI and NVI. The findings from Model 1 and Model 2 were consistent with those of the crude model. After adjusting for extensive confounders (Model 3), we found that participants who cooked with solid fuels had a significantly increased risk of DVI and NVI. In addition, compared with the clean fuel users, the participants who cooked with coal had a significantly increased risk of DVI and NVI. Moreover, cooking with biomass significantly increased the risk of DVI and NVI (see Multimedia Appendix 2).

Through the wave-specific data, we found that participants who switched their cooking type from clean to solid fuels had a 51% elevated risk of DVI and a 39% elevated risk of NVI compared with the persistent clean fuel users. However, there was no evidence of an impact of switching from solid to clean fuels on NVI (*P*=.52). After excluding participants who reported not cooking in Wave 4, our findings remained robust (see Multimedia Appendix 3).

Table 2. Longitudinal association of cooking fuel type, switching fuel type, and visual impairment among middle-aged and older Chinese adults.

Variable	Participants, n	Events/incidence rate (per 1000 person-years)	Model 1[a], HR[b] (95% CI)	Model 2[c], HR (95% CI)	Model 3[d], HR (95% CI)
Distance visual impairment (DVI)					
Cooking fuel type (N=9559)					
Clean fuel (reference)	4685	1064 (39.8)	N/A[e]	N/A	N/A
Solid fuel	4914	1580 (58.1)	1.42 (1.31-1.53)	1.36 (1.26-1.47)	1.38 (1.28-1.50)
Switching fuel type (N=7998)					
From solid to clean fuels					
Persistent use of solid fuels (reference)	2237	462 (29.6)	N/A	N/A	N/A
Solid to clean fuels use	1931	338 (25.0)	0.84 (0.73-0.96)	0.85 (0.74-0.98)	0.85 (0.74-0.98)
From clean to solid fuels					
Persistent use of clean fuels (reference)	3496	419 (17.2)	N/A	N/A	N/A
Clean to solid fuels use	334	62 (26.6)	1.60 (1.22-2.09)	1.48 (1.13-1.93)	1.51 (1.15-1.98)
Near visual impairment (NVI)					
Cooking fuel type (N=9559)					
Clean fuel (reference)	4685	1371 (53.4)	N/A	N/A	N/A
Solid fuel	4914	1759 (65.8)	1.20 (1.11-1.28)	1.18 (1.10-1.26)	1.18 (1.10-1.27)
Switching fuel type (N=7998)					
From solid to clean fuels					
Persistent use of solid fuels (reference)	2237	414 (26.5)	N/A	N/A	N/A
Solid to clean fuels use	1931	346 (25.6)	0.96 (0.83-1.10)	0.96 (0.83-1.10)	0.96 (0.83-1.11)
From clean to solid fuels					
Persistent use of clean fuels (reference)	3496	456 (18.7)	N/A	N/A	N/A
Clean to solid fuels use	334	61 (26.2)	1.44 (1.10-1.88)	1.38 (1.06-1.81)	1.39 (1.06-1.82)

[a] Model 1: Unadjusted.

[b] HR: hazard ratio.

[c] Model 2: Adjusted for age, gender, BMI, marital status, years of education, residence, smoking status, alcohol consumption, and sleep duration.

[d] Model 3: Further adjusted for hypertension, dyslipidemia, diabetes, cancer, chronic lung disease, and heart disease based on Model 2.

[e] N/A: not applicable.

Subgroup Analyses on the Association Between Cooking With Solid Fuels and Visual Impairment

We performed analyses of interactions and subgroups on the relationship between using solid fuels for cooking and visual impairment in Model 3 (Table 3). We found that cooking with solid fuels had an increased risk of DVI among participants younger than 65 years, men, urban residents, and smokers. However, cooking with solid fuels resulted in a higher risk of NVI for nonsmokers and urban residents.

Table 3. Subgroup analysis of the association between solid fuel use and visual impairment among middle-aged and older Chinese adults.

Subgroups	Cooking with solid fuel, n	DVI[a]			NVI[b]		
		Events, n	HR[c] (95% CI)[d]	*P* value for interaction	Events, n	HR (95% CI)[d]	*P* value for interaction
Age (years)				<.001			.06
<65	3657	1082	1.41 (1.28-1.55)		1313	1.17 (1.08-1.26)	
≥65	1257	498	1.27 (1.10-1.48)		446	1.21 (1.04-1.42)	
Sex				.01			.17
Male	2455	667	1.45 (1.28-1.65)		806	1.19 (1.07-1.33)	
Female	2459	913	1.34 (1.21-1.48)		953	1.16 (1.06-1.28)	
Residence				<.001			<.001
Urban	344	83	1.41 (1.08-1.75)		93	1.20 (1.10-1.37)	
Rural	4570	1497	1.16 (1.06-1.27)		1666	1.05 (0.97-1.14)	
Smoking status				<.001			<.001
Smoker	2038	585	1.43 (1.25-1.64)		678	1.12 (0.99-1.26)	
Nonsmoker	2876	995	1.36 (1.23-1.49)		1081	1.21 (1.11-1.33)	

[a] DVI: distance visual impairment.

[b] NVI: near visual impairment.

[c] HR: hazard ratio.

[d] Adjusted for age, gender, BMI, marital status, years of education, residence, smoking status, alcohol consumption, sleep duration, hypertension, dyslipidemia, diabetes, cancer, chronic lung disease, and heart disease.

Discussion

Summary of Findings

This study demonstrated that cooking with solid fuels increased the risks of visual impairment, with an adverse effect of solid fuel use found for both DVI and NVI, especially for participants living in urban areas. Moreover, switching from clean to solid fuels was demonstrated to increase the risk of DVI and NVI compared to the persistent use of clean fuel. These main findings are consistent with evidence from previous studies conducted in countries neighboring China. In India, the older population, who are highly reliant on using unclean cooking fuel, had a higher prevalence of visual impairment [20]. In Nepal, the incidence of cataracts was two-fold higher in women using biomass for cooking compared to that of women using clean fuel [21]. This study thus adds to the accumulating evidence of adverse vision health consequences potentially arising from the use of unclean cooking fuels in low- and middle-income countries [22].

The essential mechanisms of the impacts of solid fuel utilization on visual impairment are undefined [23]. One possible reason is that burning solid fuels produces high levels of gaseous pollutants that increase the risk of eye diseases and may lead to vision impairment. On the one hand, $PM_{2.5}$ is the primary pollutant in solid fuel smog, which represents a mixture of thousands of harmful chemicals such as heavy metals and polycyclic aromatic hydrocarbons [24]. A previous study showed that for each 1 $\mu g/m^3$ rise in ambient $PM_{2.5}$ exposure, the risk of glaucoma increased by 6% [25]. Free radicals accelerate the oxidation of the lens when the eye organ is directly exposed to the fumes of solid fuel, which results in an increased risk of cataracts [26] as the leading cause of vision impairment or loss worldwide [27,28]. On the other hand, incomplete combustion of solid fuel produces carbon monoxide, which is harmful to the eyes due to hypoxia [29]. Another possibility is that indoor air pollutants resulting from solid fuel combustion can affect the body's cardiovascular system and the effect on the aqueous humor can lead to increased intraocular pressure. Moreover, when burning solid fuels, the anterior eyes are more susceptible to damage from sparks, wood dust, or sharp wood and the consequent ocular trauma would increase the risk of vision impairment [30,31]. Heat exposure associated with burning solid fuels may be another potential risk factor for eye diseases [32,33], which would contribute to the development of age-related cataracts and the early onset of presbyopia if the transient temperature elevation in the lens recurs over many years. In addition, previous studies have linked long-term solid fuel use to diabetes [34], depression [35], and cognitive impairment [36], which were reported as potential risk factors affecting vision health.

Previous studies observed that women were more likely to develop eye diseases linked to solid fuel use [37-39], suggesting that as women traditionally played a vital role in cooking, they had observably higher indoor air pollution exposure. However, our study showed that men cooking with solid fuels appeared to be at a higher risk of DVI than women. This result is consistent with the cross-sectional study from India, which failed to find that the visual impairment risk associated with use of unclean cooking fuels was higher for women and in households without a separate kitchen or ventilation [20]. One possible reason for this inconsistency might be differences in the study populations. In our study, most of the individuals were retirees aged over 60 years, who tend to spend most of their time at home, which would increase the frequency of couples cooking together. This situation thereby increases the exposure to indoor air pollution caused by the use of solid fuels for men. Moreover, there is no doubt that the smoking rate is much higher among men than among women and there may be an interaction between vision impairment and solid fuel use due to the apparent gender difference in smoking habits.

More harmful impacts of cooking with solid fuels on NVI were noted among urban residents. Rural areas are usually less economically developed and have a lower population density than urban areas. Higher industrialization in urban areas may increase the risk of visual impairment since rapid industrialization would result in worse indoor air pollution and associated health outcomes [40]. Additionally, the higher population density in urban areas leads to a lower living space per capita, which may elevate the indoor pollutant concentration [41]. Despite some households having installed ventilation devices to prevent severe indoor air pollution, the level of air pollutants in domestic kitchens using solid fuels remains high [42]. In addition, switching cooking fuels from solid to clean fuels failed to decrease the risks of cataracts and conjunctiva disorders compared to the persistent use of solid fuels [5]. Taking this into consideration, it is necessary to call on the general public to use clean energy to reduce indoor air pollution and promote healthy vision, especially in rural areas.

Simultaneously, an interesting result emerging from our study was that switching from clean fuels to solid fuels could increase the risk of visual impairment compared with the continued use of clean fuels. The switch of cooking fuels from clean fuels to biomass fuels might be related to the migration from urban to rural areas among older people in China, which has been a prevalent trend in recent years [43,44]. Older adults generally prefer solid fuels more than younger adults and they are consequently more prone to the adverse effects of indoor air pollution. A cohort study with a 3-year follow-up period found that cooking with biomass fuels was associated with a higher risk of visual impairment among older Chinese adults with a mean baseline age of 82.56 years [43]. The persistent formation of lens fibers and the relative thickening of the lens cortex are important features during the aging process, resulting in changes in the refractive index of the lens [45]. This would in turn result in a higher risk of visual impairment when older adults change their cooking fuels from clean to solid fuels as their eyes may not be able to adapt to the sudden deterioration of air quality caused by solid fuels. Therefore, public health policies and investments that support cleaner household energy can effectively reduce the key sources of ambient air pollution, especially by avoiding the persistent use of solid fuels or switching from clean fuels to solid fuels, to further relieve the burden on visual impairment.

Limitations

This study has several limitations. First, with the rapid development of modernization, urbanization, and industrialization in China over the past three decades, more and more rural residents have gradually begun to use clean energy. Some of the participants in this study may have switched from solid to clean fuels prior to the baseline assessment, which would lead to an underestimate of the impact of solid fuels on visual function. Second, individual visual acuity and the exposure level to household pollutants cannot be accurately estimated due to data unavailability; thus, further research should use an objective assessment to verify the data and adjust for lead-time bias [46]. Third, there remain some potential confounders that we were unable to measure that could also have an impact, including occupational dust [31], heat exposure [33], or sunlight [47].

Conclusion

Overall, our study demonstrated that using solid fuels for cooking was significantly associated with elevated risks of visual impairment among Chinese adults aged over 45 years, especially among participants living in urban areas. Furthermore, switching the cooking fuel type from clean to solid fuels may also significantly increase the risk of visual impairment. Therefore, this study highlights the value of promoting the availability of household clean fuels and encouraging the consistent use of clean fuels to decrease the burden of visual impairment.

Acknowledgments

We thank the China Health and Retirement Longitudinal Study research team for providing the data as well as all of the study participants for their contribution. This study was funded by the Fundamental Research Funds for the Central Universities, Huazhong University of Science and Technology (2020WKZDJC015).

Data Availability

The original data set is publicly available at the China Health and Retirement Longitudinal Study (CHARLS) website [48]. The data analyzed in our study can be obtained from the corresponding author upon reasonable request.

Authors' Contributions

QJ and SW contributed to the literature search, figures, study design, data analyses, data curation, data interpretation, and writing–original draft. HZ and YG contributed to methodology and writing–review & editing. YL, SH, and QY contributed to resources and writing–review & editing. SC contributed to conceptualization, validation, and writing–review & editing. All authors had the final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The frequency of cooking fuel types of the study participants.
[DOCX File , 17 KB - publichealth_v9i1e43914_app1.docx]

Multimedia Appendix 2
Longitudinal association between cooking fuel type and visual impairment among middle-aged and older Chinese adults.
[DOCX File , 17 KB - publichealth_v9i1e43914_app2.docx]

Multimedia Appendix 3
The association of switching cooking fuels type with visual impairment excluding participants who reported not cooking in Wave 4.
[DOCX File , 16 KB - publichealth_v9i1e43914_app3.docx]

References

1. Health Effect Institute. State of global air 2019. 2019. URL: https://www.stateofglobalair.org/sites/default/files/soga_2019_report.pdf [accessed 2023-07-10]
2. Household air pollution. World Health Organization. 2023. URL: https://www.who.int/health-topics/air-pollution#tab=tab_3 [accessed 2023-07-10]
3. Karakoçak BB, Patel S, Ravi N, Biswas P. Investigating the effects of stove emissions on ocular and cancer cells. Sci Rep 2019 Feb 12;9(1):1870. [doi: 10.1038/s41598-019-38803-4] [Medline: 30755694]
4. West S, Bates M, Lee J, Schaumberg D, Lee D, Adair-Rohani H, et al. Is household air pollution a risk factor for eye disease? Int J Environ Res Public Health 2013 Oct 25;10(11):5378-5398 [FREE Full text] [doi: 10.3390/ijerph10115378] [Medline: 24284355]
5. Chan KH, Yan M, Bennett DA, Guo Y, Chen Y, Yang L, China Kadoorie Biobank Study group. Long-term solid fuel use and risks of major eye diseases in China: a population-based cohort study of 486,532 adults. PLoS Med 2021 Jul 29;18(7):e1003716 [FREE Full text] [doi: 10.1371/journal.pmed.1003716] [Medline: 34324491]
6. GBD 2019 Blindness Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. Lancet Glob Health 2021 Feb;9(2):e144-e160 [FREE Full text] [doi: 10.1016/S2214-109X(20)30489-7] [Medline: 33275949]
7. Tham Y, Li X, Wong TY, Quigley HA, Aung T, Cheng C. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology 2014 Nov;121(11):2081-2090. [doi: 10.1016/j.ophtha.2014.05.013] [Medline: 24974815]
8. Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng C, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. Lancet Glob Health 2014 Feb;2(2):e106-e116 [FREE Full text] [doi: 10.1016/S2214-109X(13)70145-1] [Medline: 25104651]
9. Eye care, vision care, vision impairment and blindness. World Health Organization. URL: https://www.who.int/health-topics/blindness-and-vision-loss#tab=tab_1 [accessed 2023-07-10]
10. Xu T, Wang B, Liu H, Wang H, Yin P, Dong W, et al. Prevalence and causes of vision loss in China from 1990 to 2019: findings from the Global Burden of Disease Study 2019. Lancet Public Health 2020 Dec;5(12):e682-e691 [FREE Full text] [doi: 10.1016/S2468-2667(20)30254-1] [Medline: 33271081]
11. Zhao J, Xu X, Ellwein LB, Cai N, Guan H, He M, et al. Prevalence of vision impairment in older adults in rural China in 2014 and comparisons with the 2006 China Nine-Province Survey. Am J Ophthalmol 2018 Jan;185:81-93 [FREE Full text] [doi: 10.1016/j.ajo.2017.10.016] [Medline: 29102607]
12. Li C, Zhu B, Zhang J, Guan P, Zhang G, Yu H, et al. Epidemiology, health policy and public health implications of visual impairment and age-related eye diseases in mainland China. Front Public Health 2022;10:966006 [FREE Full text] [doi: 10.3389/fpubh.2022.966006] [Medline: 36438305]

13. Yu K, Lv J, Qiu G, Yu C, Guo Y, Bian Z, China Kadoorie Biobank Study. Cooking fuels and risk of all-cause and cardiopulmonary mortality in urban China: a prospective cohort study. Lancet Glob Health 2020 Mar;8(3):e430-e439 [FREE Full text] [doi: 10.1016/S2214-109X(19)30525-X] [Medline: 31972151]
14. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
15. Cao G, Yao S, Wang K, Han L, Chen Z, Huang Z, et al. Vision impairment, vision correction, and cognitive decline among middle-aged and older Chinese: results from the China health and retirement longitudinal study. Aging Ment Health 2021 Nov;25(11):2028-2035. [doi: 10.1080/13607863.2020.1828271] [Medline: 33063527]
16. Zhang Q, Cao G, Yao S, Wang C, Chen Z, Hu Y, et al. Self-reported vision impairment, vision correction, and depressive symptoms among middle-aged and older Chinese: Findings from the China health and retirement longitudinal study. Int J Geriatr Psychiatry 2021 Jan;36(1):86-95. [doi: 10.1002/gps.5398] [Medline: 32783270]
17. Shang X, Wu G, Wang W, Zhu Z, Zhang X, Huang Y, et al. Associations of vision impairment and eye diseases with frailty in community-dwelling older adults: a nationwide longitudinal study in China. Br J Ophthalmol 2022 Dec 19:bjo-2022-322048. [doi: 10.1136/bjo-2022-322048] [Medline: 36535748]
18. Administrative area codes and urban-rural codes for statistics in 2022. National Bureau of Statistics of China. URL: http://www.stats.gov.cn/sj/tjbz/tjyqhdmhcxhfdm/2022/index.html [accessed 2022-10-31]
19. Morris ME, Aguilera A. Mobile, social, and wearable computing and the evolution of psychological practice. Prof Psychol Res Pr 2012 Dec;43(6):622-626 [FREE Full text] [doi: 10.1037/a0029041] [Medline: 25587207]
20. Islam S, Upadhyay AK, Mohanty SK, Pedgaonkar SP, Maurer J, O'Donnell O. Use of unclean cooking fuels and visual impairment of older adults in India: a nationally representative population-based study. Environ Int 2022 Jul;165:107302 [FREE Full text] [doi: 10.1016/j.envint.2022.107302] [Medline: 35617815]
21. Patel M, Shrestha MK, Manandhar A, Gurung R, Sadhra S, Cusack R, et al. Effect of exposure to biomass smoke from cooking fuel types and eye disorders in women from hilly and plain regions of Nepal. Br J Ophthalmol 2022 Jan 12;106(1):141-148. [doi: 10.1136/bjophthalmol-2020-316766] [Medline: 32920527]
22. Li X, Guo Y, Liu T, Xiao J, Zeng W, Hu J, et al. The association of cooking fuels with cataract among adults aged 50 years and older in low- and middle-income countries: Results from the WHO Study on global AGEing and adult health (SAGE). Sci Total Environ 2021 Oct 10;790:148093. [doi: 10.1016/j.scitotenv.2021.148093] [Medline: 34102447]
23. Jung SJ, Mehta JS, Tong L. Effects of environment pollution on the ocular surface. Ocul Surf 2018 Apr;16(2):198-205. [doi: 10.1016/j.jtos.2018.03.001] [Medline: 29510225]
24. Gordon SB, Bruce NG, Grigg J, Hibberd PL, Kurmi OP, Lam KH, et al. Respiratory risks from household air pollution in low and middle income countries. Lancet Respir Med 2014 Oct;2(10):823-860 [FREE Full text] [doi: 10.1016/S2213-2600(14)70168-7] [Medline: 25193349]
25. Chua SYL, Khawaja AP, Morgan J, Strouthidis N, Reisman C, Dick AD, UK Biobank Eye Vision Consortium. The relationship between ambient atmospheric fine particulate matter (PM2.5) and glaucoma in a large community cohort. Invest Ophthalmol Vis Sci 2019 Nov 01;60(14):4915-4923. [doi: 10.1167/iovs.19-28346] [Medline: 31764948]
26. Vashist P, Tandon R, Murthy GVS, Barua CK, Deka D, Singh S, ICMR-EYE SEE Study Group. Association of cataract and sun exposure in geographically diverse populations of India: The CASE study. First Report of the ICMR-EYE SEE Study Group. PLoS One 2020;15(1):e0227868 [FREE Full text] [doi: 10.1371/journal.pone.0227868] [Medline: 31971985]
27. Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, Vision Loss Expert Group of the Global Burden of Disease Study. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. Lancet Glob Health 2017 Dec;5(12):e1221-e1234 [FREE Full text] [doi: 10.1016/S2214-109X(17)30393-5] [Medline: 29032195]
28. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1923-1994 [FREE Full text] [doi: 10.1016/S0140-6736(18)32225-6] [Medline: 30496105]
29. Wu PE, Juurlink DN. Carbon monoxide poisoning. CMAJ 2014 May 13;186(8):611-611 [FREE Full text] [doi: 10.1503/cmaj.130972] [Medline: 24396094]
30. Cao Y, Bindslev DA, Kjærgaard SK. Estimation of the in vitro eye irritating and inflammatory potential of lipopolysaccharide (LPS) and dust by using reconstituted human corneal epithelium tissue cultures. Toxicol Mech Methods 2015 Sep 11;25(5):402-409. [doi: 10.3109/15376516.2015.1045660] [Medline: 26361134]
31. Xiang P, Liu R, Sun H, Han Y, He R, Cui X, et al. Molecular mechanisms of dust-induced toxicity in human corneal epithelial cells: water and organic extract of office and house dust. Environ Int 2016;92-93:348-356. [doi: 10.1016/j.envint.2016.04.013] [Medline: 27131017]
32. Sharon N, Bar-Yoseph PZ, Bormusov E, Dovrat A. Simulation of heat exposure and damage to the eye lens in a neighborhood bakery. Exp Eye Res 2008 Jul;87(1):49-55. [doi: 10.1016/j.exer.2008.04.007] [Medline: 18547562]
33. Nandi SK, Rankenberg J, Glomb MA, Nagaraj RH. Transient elevation of temperature promotes cross-linking of α-crystallin-client proteins through formation of advanced glycation endproducts: A potential role in presbyopia and

cataracts. Biochem Biophys Res Commun 2020 Dec 17;533(4):1352-1358 [FREE Full text] [doi: 10.1016/j.bbrc.2020.10.018] [Medline: 33081971]

34. Liu C, Yang C, Zhao Y, Ma Z, Bi J, Liu Y, et al. Associations between long-term exposure to ambient particulate air pollution and type 2 diabetes prevalence, blood glucose and glycosylated hemoglobin levels in China. Environ Int 2016;92-93:416-421 [FREE Full text] [doi: 10.1016/j.envint.2016.03.028] [Medline: 27148900]
35. Li C, Zhou Y, Ding L. Effects of long-term household air pollution exposure from solid fuel use on depression: evidence from national longitudinal surveys from 2011 to 2018. Environ Pollut 2021 Aug 15;283:117350. [doi: 10.1016/j.envpol.2021.117350] [Medline: 34034020]
36. Cao L, Zhao Z, Ji C, Xia Y. Association between solid fuel use and cognitive impairment: a cross-sectional and follow-up study in a middle-aged and older Chinese population. Environ Int 2021 Jan;146:106251 [FREE Full text] [doi: 10.1016/j.envint.2020.106251] [Medline: 33248346]
37. Zetterberg M, Celojevic D. Gender and cataract--the role of estrogen. Curr Eye Res 2015 Feb 02;40(2):176-190. [doi: 10.3109/02713683.2014.898774] [Medline: 24987869]
38. Chan KH, Xia X, Ho K, Guo Y, Kurmi OP, Du H, CKB-Air Collaborative Group. Regional and seasonal variations in household and personal exposures to air pollution in one urban and two rural Chinese communities: A pilot study to collect time-resolved data using static and wearable devices. Environ Int 2021 Jan;146:106217 [FREE Full text] [doi: 10.1016/j.envint.2020.106217] [Medline: 33129001]
39. Ravilla TD, Gupta S, Ravindran RD, Vashist P, Krishnan T, Maraini G, et al. Use of cooking fuels and cataract in a population-based study: The India Eye Disease Study. Environ Health Perspect 2016 Dec;124(12):1857-1862. [doi: 10.1289/EHP193] [Medline: 27227523]
40. Bai R, Lam JCK, Li VOK. A review on health cost accounting of air pollution in China. Environ Int 2018 Nov;120:279-294. [doi: 10.1016/j.envint.2018.08.001] [Medline: 30103126]
41. Vardoulakis S, Giagloglou E, Steinle S, Davis A, Sleeuwenhoek A, Galea KS, et al. Indoor exposure to selected air pollutants in the home environment: a systematic review. Int J Environ Res Public Health 2020 Dec 02;17(23):8972 [FREE Full text] [doi: 10.3390/ijerph17238972] [Medline: 33276576]
42. Pope D, Bruce N, Dherani M, Jagoe K, Rehfuess E. Real-life effectiveness of 'improved' stoves and clean fuels in reducing PM and CO: systematic review and meta-analysis. Environ Int 2017 Apr;101:7-18. [doi: 10.1016/j.envint.2017.01.012] [Medline: 28285622]
43. Zhou Y, Xu M, Ke P, Di H, Gan Y, Feng J, et al. Association of biomass fuel use with the risk of vision impairment among Chinese older adults: a cohort study. Environ Sci Pollut Res Int 2023 Apr;30(19):56273-56283. [doi: 10.1007/s11356-023-26325-6] [Medline: 36917388]
44. Deng Y, Gao Q, Yang T, Wu B, Liu Y, Liu R. Indoor solid fuel use and incident arthritis among middle-aged and older adults in rural China: a nationwide population-based cohort study. Sci Total Environ 2021 Jun 10;772:145395. [doi: 10.1016/j.scitotenv.2021.145395] [Medline: 33578144]
45. Augusteyn RC. On the growth and internal structure of the human lens. Exp Eye Res 2010 Jun;90(6):643-654. [doi: 10.1016/j.exer.2010.01.013] [Medline: 20171212]
46. Duffy SW, Nagtegaal ID, Wallis M, Cafferty FH, Houssami N, Warwick J, et al. Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival. Am J Epidemiol 2008 Jul 01;168(1):98-104. [doi: 10.1093/aje/kwn120] [Medline: 18504245]
47. Delavar A, Freedman DM, Velazquez-Kronen R, Little MP, Kitahara CM, Alexander BH, et al. Ultraviolet radiation and incidence of cataracts in a nationwide US cohort. Ophthalmic Epidemiol 2018 Aug 10;25(5-6):403-411. [doi: 10.1080/09286586.2018.1501077] [Medline: 30095320]
48. Harmonized CHARLS datasets. CHARLS. URL: https://charls.charlsdata.com/pages/data/111/zh-cn.html [accessed 2023-07-10]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
DVI: distance visual impairment
NVI: near visual impairment
PM2.5: fine particulate matter
PSU: primary sampling unit

Edited by A Mavragani; submitted 29.10.22; peer-reviewed by J Lyu, M Raimi, M Mbwogge; comments to author 25.04.23; revised version received 13.05.23; accepted 21.06.23; published 26.07.23.

Please cite as:
Jiang Q, Wang S, Zhang H, Guo Y, Lou Y, Huang S, You Q, Cao S
The Association Between Solid Fuel Use and Visual Impairment Among Middle-Aged and Older Chinese Adults: Nationwide Population-Based Cohort Study
JMIR Public Health Surveill 2023;9:e43914
URL: https://publichealth.jmir.org/2023/1/e43914
doi:10.2196/43914
PMID:37494091

Original Paper

Mortality Risk and Burden From a Spectrum of Causes in Relation to Size-Fractionated Particulate Matters: Time Series Analysis

Jun Yang[1,2*], PhD; Hang Dong[3,4*], MM; Chao Yu[3,4*], MM; Bixia Li[2,5*], MSc; Guozhen Lin[3,4], PhD; Sujuan Chen[2], MSc; Dongjie Cai[2], MSc; Lin Huang[2], MSc; Boguang Wang[2], PhD; Mengmeng Li[6], PhD

[1]School of Public Health, Guangzhou Medical University, Guangzhou, China

[2]Institute for Environmental and Climate Research, Jinan University, Guangzhou, China

[3]Guangzhou Center for Disease Control and Prevention, Guangzhou, China

[4]Institute of Public Health, Guangzhou Medical University and Guangzhou Center for Disease Control and Prevention, Guangzhou, China

[5]Guangdong University of Science and Technology, Dongguan, China

[6]State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Mengmeng Li, PhD
State Key Laboratory of Oncology in South China
Guangdong Provincial Clinical Research Center for Cancer
Sun Yat-sen University Cancer Center
No 651 Dongfeng East Road, Guangzhou 510060
Guangzhou, 510060
China
Phone: 86 020 87345679
Email: limm@sysucc.org.cn

Abstract

Background: There is limited evidence regarding the adverse impact of particulate matters (PMs) on multiple body systems from both epidemiological and mechanistic studies. The association between size-fractionated PMs and mortality risk, as well as the burden of a whole spectrum of causes of death, remains poorly characterized.

Objective: We aimed to examine the wide range of susceptible diseases affected by different sizes of PMs. We also assessed the association between PMs with an aerodynamic diameter less than 1 μm (PM_1), 2.5 μm ($PM_{2.5}$), and 10 μm (PM_{10}) and deaths from 36 causes in Guangzhou, China.

Methods: Daily data were obtained on cause-specific mortality, PMs, and meteorology from 2014 to 2016. A time-stratified case-crossover approach was applied to estimate the risk and burden of cause-specific mortality attributable to PMs after adjusting for potential confounding variables, such as long-term trend and seasonality, relative humidity, temperature, air pressure, and public holidays. Stratification analyses were further conducted to explore the potential modification effects of season and demographic characteristics (eg, gender and age). We also assessed the reduction in mortality achieved by meeting the new air quality guidelines set by the World Health Organization (WHO).

Results: Positive and monotonic associations were generally observed between PMs and mortality. For every 10 $\mu g/m^3$ increase in 4-day moving average concentrations of PM_1, $PM_{2.5}$, and PM_{10}, the risk of all-cause mortality increased by 2.00% (95% CI 1.08%-2.92%), 1.54% (95% CI 0.93%-2.16%), and 1.38% (95% CI 0.95%-1.82%), respectively. Significant effects of size-fractionated PMs were observed for deaths attributed to nonaccidental causes, cardiovascular disease, respiratory disease, neoplasms, chronic rheumatic heart diseases, hypertensive diseases, cerebrovascular diseases, stroke, influenza, and pneumonia. If daily concentrations of PM_1, $PM_{2.5}$, and PM_{10} reached the WHO target levels of 10, 15, and 45 $\mu g/m^3$, 7921 (95% empirical CI [eCI] 4454-11,206), 8303 (95% eCI 5063-11,248), and 8326 (95% eCI 5980-10690) deaths could be prevented, respectively. The effect estimates of PMs were relatively higher during hot months, among female individuals, and among those aged 85 years and older, although the differences between subgroups were not statistically significant.

Conclusions: We observed positive and monotonical exposure-response curves between PMs and deaths from several diseases. The effect of PM_1 was stronger on mortality than that of $PM_{2.5}$ and PM_{10}. A substantial number of premature deaths could be preventable by adhering to the WHO's new guidelines for PMs. Our findings highlight the importance of a size-based strategy in controlling PMs and managing their health impact.

(JMIR Public Health Surveill 2023;9:e41862) doi:10.2196/41862

KEYWORDS

size-fractionated particulate matter; cause-specific mortality; cardiovascular disease; respiratory disease; neoplasm; attributable burden

Introduction

With the rapid socioeconomic development and fast urbanization, air pollution—particularly the particulate matters (PMs)—has become the biggest environmental challenge to human health globally [1]. According to the recent assessment of the global burden of disease, PMs accounted for over 4.1 million deaths [2]. As the largest developing country, China faces a disproportionately high health burden due to PM pollution [3]. The improvement of air pollution control policies warrants an in-depth knowledge and quantification of the health impact of PMs.

PMs consist of discrete particles that vary in size, which is an important characteristic influencing their hazardous effects. PMs with aerodynamic diameters less than 1 μm (PM_1), 2.5 μm ($PM_{2.5}$), and 10 μm (PM_{10}) have been studied extensively due to their ability to enter and deposit in the respiratory tract. Previous epidemiological investigations have identified the harmful impact of PM_{10} and $PM_{2.5}$ on human health [4,5]. In recent years, PM_1 has also raised increasing concerns due to emerging evidence indicating higher health risks associated with PMs of smaller sizes [6-9].

Prior studies mainly focused on the impact of size-fractionated PMs on mortality from common chronic diseases (ie, circulatory and respiratory diseases) [1,3,4,6]. Although less explored, PMs have been linked to an increased risk of developing and dying from diabetes and pancreatic cancer [5,10]. The risk of dying from external causes, including intentional self-harm, was also associated with increased levels of PMs [11]. There is limited evidence to suggest that PMs might also affect other systems, such as digestive, nervous, and genitourinary systems [12,13]. Mechanistic studies have revealed that environmental exposures can cause oxidative stress and inflammation, genomic and epigenetic alterations, mitochondrial dysfunction, endocrine disruption, altered intercellular communication, altered microbiome communities, and impaired nervous system function [14]. We, therefore, hypothesize that size-fractionated PMs can increase the risk of death from diseases involving multiple human body systems, and this excess risk might vary depending on the size of PMs.

We performed a comprehensive evaluation of the relationship between size-fractionated PMs and cause-specific mortality using granular data collected from Guangzhou, a city in China highly polluted by PMs, with PM_1 and $PM_{2.5}$ concentrations reaching as high as 122 $μg/m^3$ and 150 $μg/m^3$ during our study period. Our aim was to examine the wide range of susceptible diseases affected by different sizes of PMs. Specifically, we examined deaths from 10 broad categories of causes (ie, all cause, nonaccidental, cardiovascular, respiratory, digestive, genitourinary, nervous system, endocrine system, external causes, and neoplasms), along with their 26 subcategories. These categories were assumed to be affected by PMs based on existing evidence, as mentioned earlier.

Methods

Health Data

Daily mortality data and weekly counts of influenza-like illness (ILI) for the years 2014-2016 were collected from the Guangzhou Center for Disease Control and Prevention. The causes of death were coded according to the International Classification of Diseases, Tenth Revision. All-cause and nonaccidental mortality were defined using the codes A00-Z99 and A00-R99 (from the International Classification of Diseases, Tenth Revision), respectively. We also extracted data regarding deaths due to cardiovascular (I00-I99), respiratory (J00-J99), digestive (K00-K93), and genitourinary diseases (N00-N99), together with diseases of the nervous system (G00-G99) and endocrine system (D50-D89 and E00-E90). Data on deaths from external causes (V01-Y89) and neoplasms (C00-D48) were also collected. In addition, 26 subcategories within the previously mentioned 10 broad disease categories were also considered (Table S1 in Multimedia Appendix 1). In addition, daily counts of deaths were further categorized by age group (≤64, 65-74, 75-84, and ≥85 years) and gender.

Environmental Data

Daily concentrations of PM_1 during the 2014-2016 period were collected from the monitoring stations of the Chinese Atmosphere Watch Network in Guangzhou [15,16]. Data for $PM_{2.5}$ and PM_{10} were obtained from Guangzhou Bureau of Environmental Protection, together with data for other air pollutants, including ozone, sulfur dioxide, nitrogen dioxide, and carbon monoxide. We averaged the daily concentrations for each pollutant from 11 fixed-site monitoring stations in Guangzhou (Figure S1 in Multimedia Appendix 1). China Meteorological Data Service Center provided the daily meteorological data in Guangzhou, containing daily relative humidity (%); minimum, mean, and maximum temperatures (^oC); and air pressure (hPa).

Ethical Considerations

This study involved only a secondary analysis of daily aggregated and deidentified data, and it is classified as exempt from institutional review board approval according to the Chinese legal documents on ethics review issued by the National Health Commission of the People’s Republic of China (document number: 4; 2023) [17].

Statistical Analyses

We conducted a time-stratified case-crossover method to evaluate the impact of PMs (PM_1, $PM_{2.5}$, and PM_{10}) on mortality [11,18]. The following quasi-Poisson function accounting for overdispersion was used to assess the relationship between size-fractionated PMs and cause-specific mortality:

where Y_t is the observed number of daily deaths on day t; α is the model intercept; $Strata_t$ denotes the time stratum variable used to control for seasonality and long-term trend; NS represents the natural cubic spline function, with 3 dfs for relative humidity (RH_t) and air pressure (PRE_t) and 6 dfs for mean temperature ($TEMP_t$); λ , γ, and η are the vectors of coefficients. β denotes the changes in mortality risk per every 10 μg/m^3 increase in the concentrations of PM pollutants [3,5,6]. Given that influenza could potentially confound the association between air pollution and health [19], the daily average of ILI_t occurrence was also included in the model. Relative risk [$RR=exp(\beta)$] was estimated from the model and the impact of PMs was expressed as the percentage change [$(RR\text{-}1) \times 100\%$] in the daily number of mortality associated with a 10 μg/m^3 increment in PMs.

Furthermore, a smoothing spline function with 3 dfs was used to evaluate the exposure-response relationship between PMs (PM_1, $PM_{2.5}$, and PM_{10}) and mortality risk. We also conducted a sensitivity analysis by changing the dfs (5-9) to test the robustness of the association. To investigate the lag patterns of PMs, we fitted the models using different single lags. Moving average approach was further applied to capture the cumulative lag effects of PMs.

To estimate the excess mortality burden caused by size-fractionated PMs with daily concentrations higher than the recommended target, we considered 5 target levels of $PM_{2.5}$ and PM_{10} as per the World Health Organization’s (WHO) new air quality guidelines. The air quality guideline level and 4 interim targets for $PM_{2.5}$ are 15, 25, 37.5, 50, and 75 μg/m^3, respectively, while for $PM_{10,}$ they are 45, 50, 75, 100, and 150 μg/m^3, respectively [20]. There are no officially announced target levels for PM_1 yet; however, given the high correlation between PM_1 and $PM_{2.5}$ and the fact that they are both secondary pollutants emitted from some common sources [21], we assumed the percentiles of the PM_1 target levels to be the same as those of the $PM_{2.5}$ target levels. We were able to identify the PM_1 target levels (10, 20, 30, 40, and 55 μg/m^3) by locating these percentiles within the PM_1 distribution (Table S2 in Multimedia Appendix 1). The death burden due to PMs was calculated by combining the relative risk of PMs associated with their daily concentrations and the corresponding observed daily number of deaths. The total number of deaths attributable to each specific PM target was computed by summing the excess deaths when daily PM concentrations exceeded that target. The empirical CI (eCI) for the attributable deaths was estimated using Monte Carlo simulation with 1000 replications [5,22].

Stratification analyses were conducted to explore the potential modification effects of season (cold period: November-April; warm period: May-October) and demographical characteristics (gender and age) on the relationship between PMs and mortality. The z statistic was performed to test the statistical difference between the 2 relative risks obtained from the subgroup analyses [23,24].

Sensitivity Analyses

The robustness of our main findings was examined by several analytical strategies. First, we changed the dfs for daily meteorological variables from 3 to 6. Second, we changed the dfs for the time variable from 3 to 9 per year. Finally, to test the confounding influence of other air pollutants, we performed the two-pollutant models by separately introducing a pair of air pollutants that are not highly correlated, thus avoiding collinearity (Spearman correlation coefficient <0.7) [3,25]. All the data analyses were conducted using the R software (version 4.0.1; R Core Team), and a 2-sided P value less than .05 was considered statistically significant.

Results

Summary of Descriptive Statistics

In total, there were 146,459 all-cause deaths from 2014 to 2016 in Guangzhou. The number of deaths was 138,396 for nonaccidental causes, including 56,587 circulatory deaths, 21,395 respiratory deaths, 4587 deaths from digestive diseases, 1303 from nervous system diseases, 1930 from genitourinary diseases, 8052 from external causes, 5469 from endocrine diseases, and 41,709 from neoplasms. The average concentrations of daily PM_1, $PM_{2.5}$, and PM_{10} were 28 (SD 13) μg/m^3, 38 (SD 21) μg/m^3, and 56 (SD 28) μg/m^3, respectively (Table 1 and Table S1 In Multimedia Appendix 1). The correlation coefficients between PM_1, $PM_{2.5}$, PM_{10}, and other air pollutants were less than 0.7, except for nitrogen dioxide. Temperature and relative humidity were negatively correlated with PMs, while air pressure was positively correlated with PMs (Table S3 in Multimedia Appendix 1).

Table 1. Summary statistics of environment and mortality data from 2014 to 2016 in Guangzhou, China.

Variables	Mean (SD)	Minimum, n	Percentile[a]					Maximum, n
			P5	P25	P50	P75	P95	
Air pollutants								
PM_1[b] (µg/m^3)	28 (16)	3	9	16	25	36	57	122
$PM_{2.5}$ (µg/m^3)	38 (21)	6	15	22	33	48	77	150
PM_{10} (µg/m^3)	56 (28)	10	24	35	48	70	108	190
Ozone (µg/m^3)	78 (47)	4	16	42	72	107	163	254
Sulfur dioxide (µg/m^3)	13 (6)	3	5	9	12	16	22	38
Nitrogen dioxide (µg/m^3)	43 (17)	13	22	31	39	52	77	146
Carbon monoxide (mg/m^3)	0.9 (0.2)	0.4	0.6	0.7	0.8	1	1.3	2.5
Meteorological variables								
Temperature (°C)	22 (6)	3	11	17	24	27	30	31
Relative humidity (%)	79 (10)	31	62	74	80	87	93	98
Air pressure (hPa)	1005 (7)	986	994	1000	1004	1010	1016	1028
Daily deaths								
All cause	134 (23)	86	102	118	130	148	174	251
Nonaccidental	126 (22)	80	96	110	124	140	166	238
Circulatory disease	52 (13)	21	35	42	50	59	75	115
Respiratory disease	20 (6)	6	11	15	19	24	32	47
Digestive disease	4 (2)	0	1	3	4	5	8	12
Nervous disease	1 (1)	0	0	0	1	2	3	7
Genitourinary disease	2 (1)	0	0	1	2	3	4	8
External causes	7 (3)	0	3	5	7	9	12	17
Endocrine diseases	5 (2)	0	1	3	5	6	9	15
Neoplasms	38 (6)	17	27	33	38	42	49	58

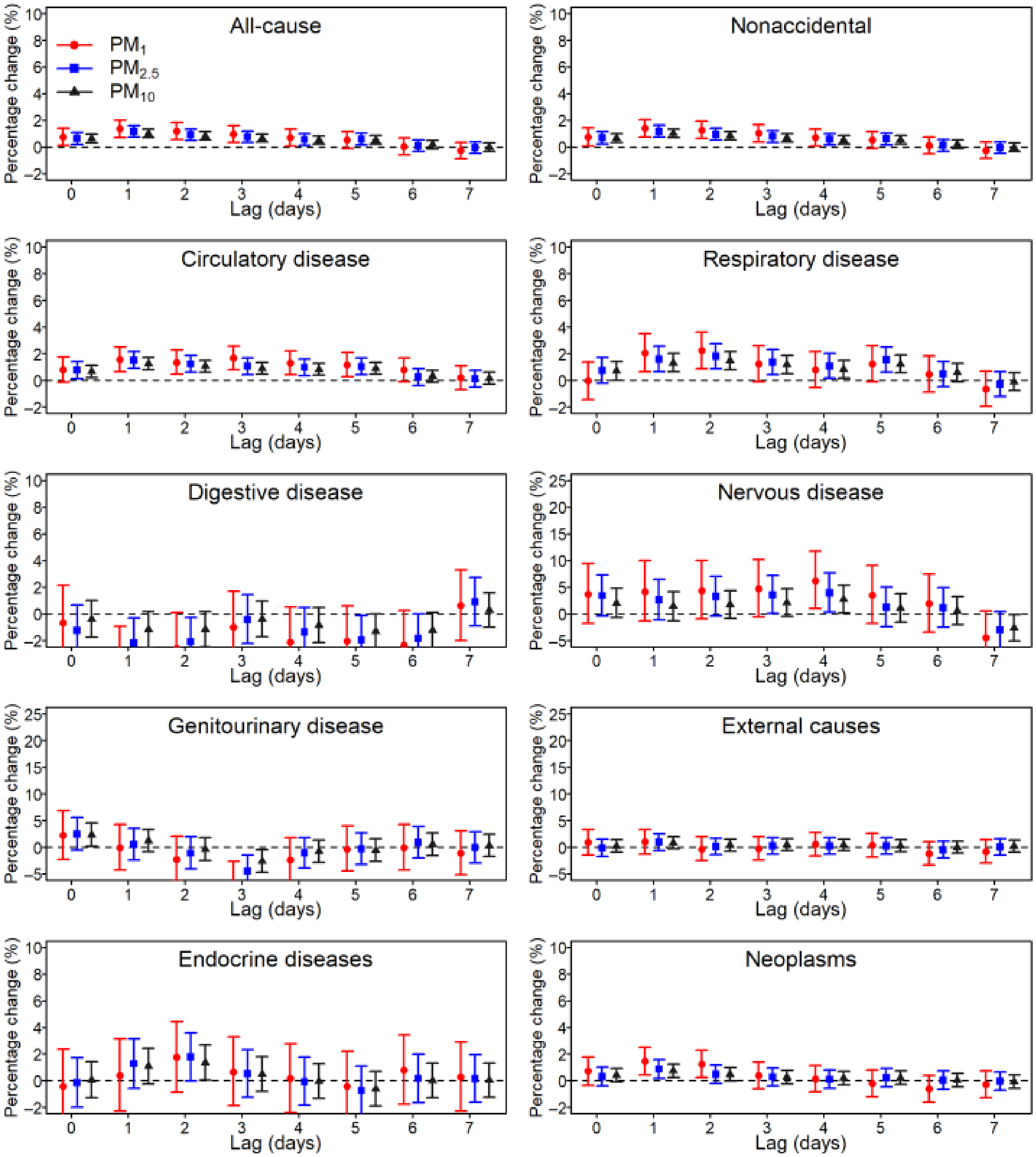
[a] P5, P25, P50, P75, and P95 denote the 5th, 25th, 50th, 75th, and 95th percentiles.

[b] PM_1: particulate matters with an aerodynamic diameter less than 1 µm.

Lag Pattern Effects of PMs on Mortality

The lag pattern effects of PM_1, $PM_{2.5}$, and PM_{10} on mortality risk manifested similar patterns, with estimates peaking at lag 1 or lag 2 for all-cause and nonaccidental mortality as well as for deaths from circulatory diseases, respiratory diseases, and neoplasms. The effects generally lasted for 4 days (Figure 1); however, the effects were statistically nonsignificant at different lag days for other diseases.

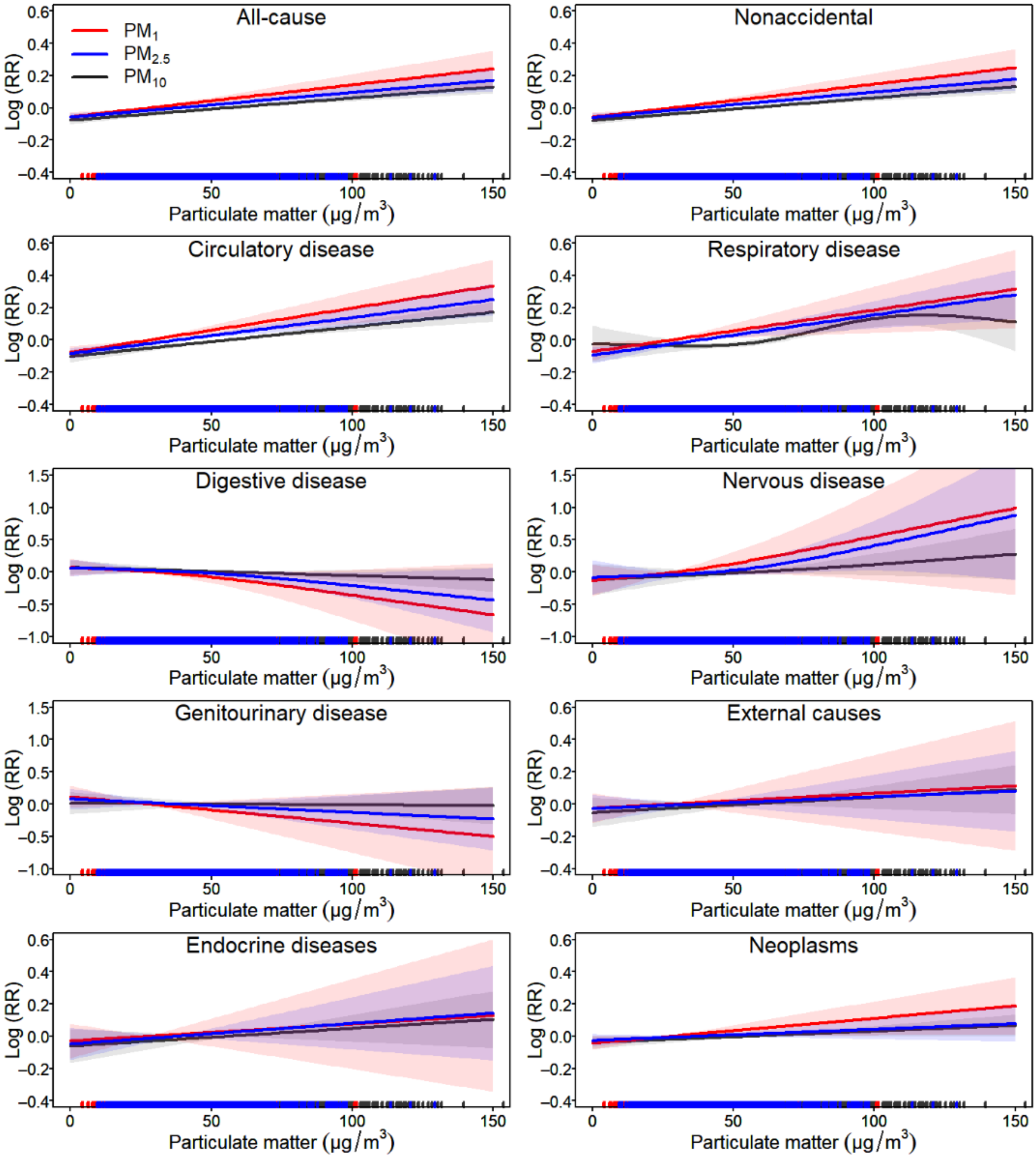
Figure 1. The percentage change (%) of cause-specific mortality associated with a 10 μg/m3 increase in size-fractionated particulate matters (PMs) on lag days 0-10. The vertical red, blue, and black lines denote the effect estimates of PM_1, $PM_{2.5}$, and PM_{10}, respectively, on mortality across lag days.



Dose-Response Relationships Between Size-Fractionated PMs and Mortality

The association between size-fractionated PMs and mortality at lag days 0-3 were generally positive and monotonically increasing (Figure 2). Similar associations were obtained when using 5-9 dfs for the spline function (Figures S2-S4 in Multimedia Appendix 1).

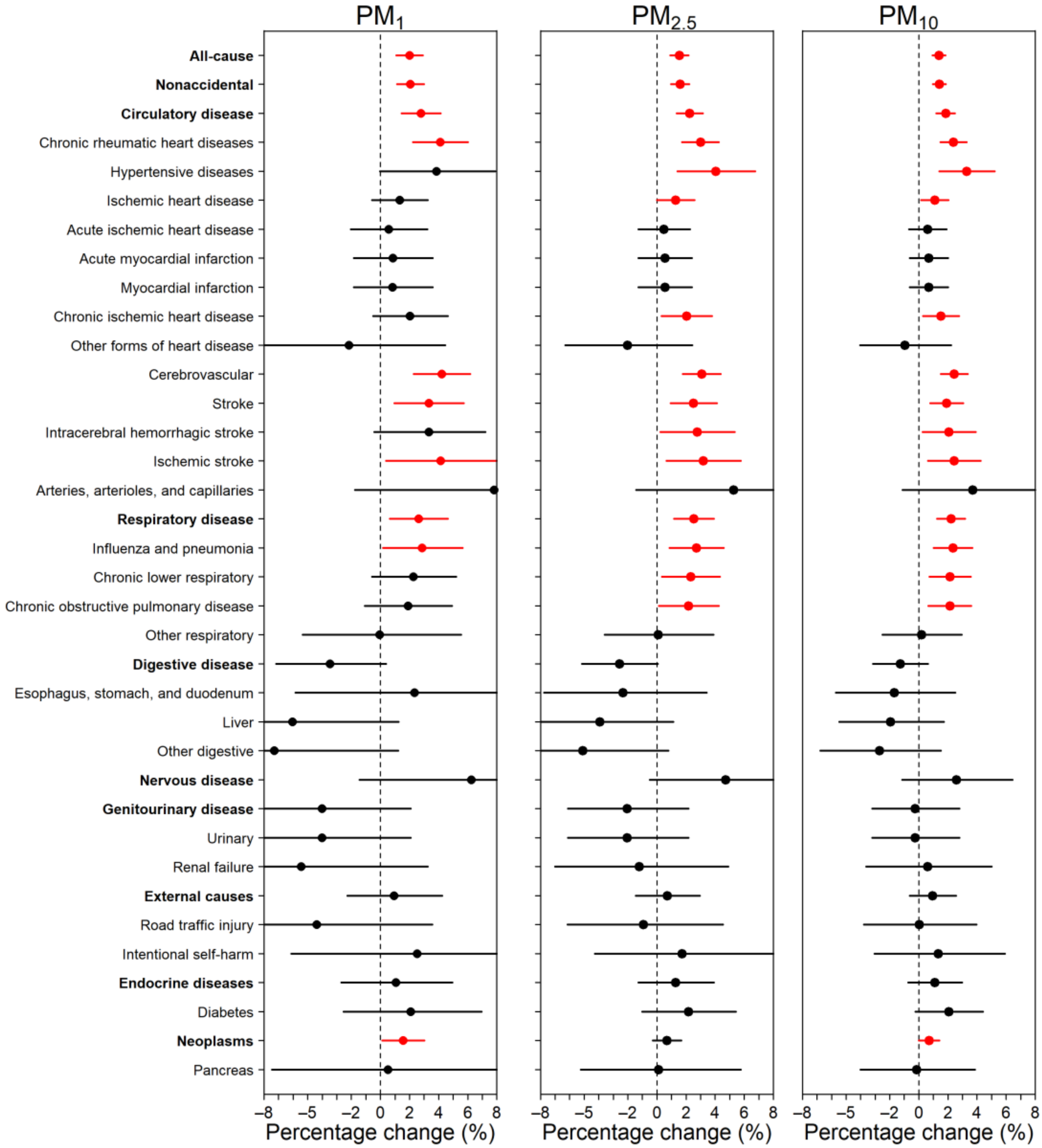
Figure 2. The concentration-response curves of size-fractionated particulate matters (PMs) and cause-specific mortality on lag days 0-3, using a 3-df smoothness for particulate matters. Red, blue, and black lines denote the PM_1-mortality, $PM_{2.5}$-mortality, and PM_{10}-mortality associations, respectively; the shaded areas represent the 95% CIs. RR: relative risk.



Diseases Sensitive to Size-Fractionated PMs

For every 10 μg/m^3 increment in 4-day average concentrations of PM_1, $PM_{2.5}$, and PM_{10}, the risk of all-cause mortality increased by 2.00% (95% CI 1.08%-2.92%), 1.54% (95% CI 0.93%-2.16%), and 1.38% (95% CI 0.95%-1.82%), respectively. The increments in nonaccidental mortality were similar across the 3 PMs (2.06%, 1.59%, and 1.41%). The largest effect estimates were observed for cardiovascular diseases (2.78%, 2.25%, and 1.85%) and respiratory diseases (2.61%, 2.53%, and 2.21%). The association between size-fractionated PMs and deaths from neoplasms were marginally significant, with effect estimates of 1.55% (95% CI 0.11%-3.01%), 0.69% (95% CI –0.28% to 1.67%), and 0.71% (95% CI 0.01%-1.40%; Figure 3 and Table S4 in Multimedia Appendix 1). The effect estimates of all size-fractionated PMs were statistically nonsignificant for deaths from nervous system diseases, digestive system diseases, genitourinary diseases, external causes, and endocrine diseases. For specific subcategories, significant effect estimates of PMs were detected among deaths due to chronic rheumatic heart diseases, hypertensive and cerebrovascular diseases, stroke (notably ischemic stroke), influenza, and pneumonia.

Figure 3. Percentage change (%) of cause-specific mortality per 10 μg/m3 increase in PM_1, $PM_{2.5}$, and PM_{10} on lag days 0-3. For each cause, point estimates (dots) and 95% CIs (horizontal lines) are shown. Red lines indicate statistically significant estimates. PM: particulate matter.



PMs' Effects on Mortality by Season and Individual Characteristics

In the subgroup analyses by season, age, and gender, we found a slightly greater impact of PM_1 on mortality in hot months compared to cold ones. The mortality risk of PM_1 was higher among older people, particularly for those aged ≥85 years. The effect estimates of PM_1 were slightly stronger in female individuals compared to their male counterparts for overall and nonaccidental mortality and deaths from neoplasms. However, the opposite trend was observed for cardiovascular diseases and respiratory diseases (Table 2). A similar trend by seasons and individual characteristics was observed for $PM_{2.5}$ and PM_{10} (Tables S5 and S6 in Multimedia Appendix 1). However, the differences in mortality risks of PM between subgroups (ie, season, age, and gender) were not statistically significant ($P>.05$).

Table 2. Percentage change in mortality per 10 μg/m3 increase in PM_1 on lag days 0-3, stratified by season, age group, and gender.

Variables	Causes of death (%, 95% CI)				
	All cause	Nonaccidental	Cardiovascular	Respiratory	Neoplasms
Total	2.00 (1.08 to 2.92)	2.06 (1.13 to 3.00)	2.78 (1.46 to 4.13)	2.61 (0.63 to 4.63)	1.55 (0.11 to 3.01)
Season					
Cold	1.66 (0.64 to 2.69)	1.74 (0.70 to 2.79)	2.14 (0.69 to 3.63)	2.56 (0.36 to 4.81)	1.29 (–0.35 to 2.96)
Hot	3.11 (1.31 to 4.94)	3.13 (1.29 to 5.00)	5.18 (2.44 to 8.00)	2.79 (–1.20 to 6.95)	2.27 (–0.38 to 5.00)
Age (years)					
0-64	1.22 (–0.24 to 2.70)	1.49 (–0.07 to 3.07)	2.21 (–0.83 to 5.34)	2.70 (–3.94 to 9.79)	0.61 (–1.46 to 2.72)
65-74	1.50 (–0.32 to 3.35)	1.47 (–0.39 to 3.36)	1.74 (–1.21 to 4.78)	–0.87 (–6.20 to 4.75)	2.55 (–0.30 to 5.48)
75-84	2.17 (0.73 to 3.63)	2.21 (0.75 to 3.69)	2.85 (0.84 to 4.90)	2.91 (–0.15 to 6.07)	1.04 (–1.58 to 3.74)
≥85	2.83 (1.22 to 4.47)	2.77 (1.14 to 4.42)	3.45 (1.23 to 5.72)	3.28 (0.33 to 6.31)	4.87 (0.17 to 9.78)
Gender					
Male	1.88 (0.81 to 2.97)	1.96 (0.86 to 3.08)	3.49 (1.81 to 5.20)	2.96 (0.41 to 5.57)	1.05 (–0.71 to 2.83)
Female	2.16 (0.90 to 3.44)	2.20 (0.92 to 3.50)	2.07 (0.26 to 3.90)	2.15 (–0.73 to 5.11)	2.43 (0.04 to 4.89)

As presented in Table 3, the excess all-cause deaths advanced by PM_1 were 7921 (95% eCI 4454-11,206), 6790 (95% eCI 3704-9611), 4899 (95% eCI 2696-6971), 3053 (95% eCI 1790-4393), and 1278 (95% eCI 711-1831), respectively, when the daily concentrations exceeded the target levels of 10, 20, 30, 40, and 55 μg/m^3. For $PM_{2.5}$, the excess death burden was 8303 (95% eCI 5063-11,248), 7194 (95% eCI 4183-9833), 5504 (95% eCI 3245-7620), 3717 (95% eCI 2242-5171), and 1352 (95% eCI 840-1861), with daily concentrations exceeding the WHO targets (15, 25, 37.5, 50, and 75 μg/m^3), respectively (Table S7 in Multimedia Appendix 1). For PM_{10}, 8326 (95% eCI 5980-10,690), 7565 (95% eCI 5130-9602), 4487 (95% eCI 3068-5843), 1991 (95% eCI 1360-2582), and 299 (95% eCI 209-384) deaths were caused, with daily concentrations exceeding the WHO targets (45, 50, 75, 100, and 150 μg/m^3), respectively (Table S8 in Multimedia Appendix 1). Therefore, 5.4%, 5.7%, and 5.7% of all deaths in Guangzhou were preventable by reaching the lowest target levels of PM_1, $PM_{2.5}$, and PM_{10}, respectively.

Table 3. The number of deaths from cause-specific disease (95% empirical CI) advanced by particulate matters (PM) with an aerodynamic diameter less than 1 µm (PM_1) with the concentrations exceeding the target levels from 2014 to 2016.

Cause of death	PM_1 targets (µg/m^3)				
	10	20	30	40	55
All cause	7921 (4454 to 11,206)	6790 (3704 to 9611)	4899 (2696 to 6971)	3053 (1790 to 4393)	1278 (711 to 1831)
Nonaccidental	7717 (4206 to 11,288)	6616 (3672 to 9204)	4776 (2592 to 6768)	2977 (1617 to 4230)	1246 (714 to 1719)
Circulatory disease	4263 (2374 to 6097)	3682 (1998 to 5241)	2688 (1455 to 3861)	1704 (970 to 2372)	727 (428 to 1057)
Chronic rheumatic heart diseases	2659 (1481 to 3744)	2292 (1300 to 3199)	1669 (1014 to 2367)	1060 (588 to 1497)	444 (245 to 609)
Hypertensive diseases	547 (–43 to 1020)	474 (4 to 858)	346 (11 to 625)	221 (17 to 412)	97 (–2 to 178)
Ischemic heart disease	886 (–402 to 2048)	767 (–284 to 1836)	561 (–166 to 1271)	355 (–182 to 879)	152 (–51 to 375)
Acute ischemic heart disease	185 (–688 to 962)	160 (–612 to 887)	118 (–457 to 619)	76 (–308 to 400)	33 (–135 to 184)
Acute myocardial infarction	257 (–595 to 1033)	223 (–550 to 896)	164 (–415 to 666)	104 (–249 to 429)	45 (–116 to 178)
Myocardial infarction	255 (–549 to 1051)	221 (–439 to 894)	163 (–410 to 661)	104 (–218 to 406)	45 (–104 to 180)
Chronic ischemic heart disease	691 (–218 to 1501)	596 (–147 to 1262)	433 (–111 to 944)	271 (–56 to 586)	115 (–43 to 247)
Other forms of heart disease	–125 (–581 to 229)	–107 (–473 to 178)	–76 (–363 to 137)	–46 (–224 to 79)	–19 (–89 to 35)
Cerebrovascular	2592 (1426 to 3681)	2234 (1287 to 3179)	1628 (862 to 2271)	1035 (553 to 1422)	434 (246 to 597)
Stroke	1320 (359 to 2125)	1140 (369 to 1858)	836 (318 to 1333)	532 (150 to 848)	222 (63 to 367)
Intracerebral hemorrhagic stroke	474 (–44 to 948)	412 (–53 to 809)	305 (–60 to 597)	192 (–27 to 382)	81 (–10 to 161)
Ischemic stroke	651 (55 to 1148)	558 (39 to 1003)	405 (31 to 722)	256 (32 to 446)	104 (11 to 187)
Arteries, arterioles, and capillaries	157 (–49 to 293)	135 (–31 to 254)	96 (–27 to 182)	61 (–14 to 110)	26 (–7 to 49)
Respiratory disease	1506 (366 to 2603)	1295 (316 to 2231)	942 (261 to 1585)	600 (161 to 1014)	261 (68 to 438)
Influenza and pneumonia	840 (73 to 1572)	724 (77 to 1344)	527 (15 to 974)	335 (38 to 616)	147 (9 to 262)
Chronic lower respiratory disease	618 (–199 to 1370)	531 (–105 to 1137)	386 (–121 to 815)	246 (–90 to 534)	105 (–40 to 225)
Chronic obstructive pulmonary disease	487 (–258 to 1168)	419 (–257 to 991)	305 (–209 to 729)	195 (–98 to 466)	83 (–48 to 195)
Other respiratory diseases	–4 (–417 to 346)	–3 (–390 to 315)	–3 (–293 to 232)	–2 (–201 to 156)	–1 (–90 to 73)
Digestive disease	–482(–1132 to 46)	–415 (–923 to 8)	–299 (–704 to 15)	–183 (–434 to 27)	–75 (–176 to 8)
Esophagus, stomach, and duodenum	68 (–214 to 270)	59 (–173 to 238)	44 (–156 to 167)	28 (–76 to 109)	12 (–35 to 45)
Liver	–274 (–706 to 58)	–234 (–598 to 59)	–166 (–430 to 34)	–101 (–262 to 17)	–46 (–118 to 8)
Other digestive diseases	–231 (–613 to 25)	–202 (–538 to 29)	–147 (–376 to 6)	–90 (–226 to 13)	–33 (–88 to 4)
Nervous disease	201 (–52 to 398)	172 (–48 to 332)	124 (–45 to 247)	78 (–29 to 154)	33 (–11 to 63)
Genitourinary disease	–241 (–650 to 102)	–208 (–569 to 115)	–152 (–430 to 72)	–95 (–269 to 46)	–41(–113 to 19)
Urinary disease	–241 (–721 to 94)	–208 (–591 to 74)	–152 (–423 to 72)	–95 (–263 to 47)	–41 (–112 to 14)
Renal failure	–171 (–509 to 67)	–146 (–453 to 74)	–106 (–359 to 47)	–67 (–217 to 24)	–29 (–96 to 16)
External causes	206 (–560 to 865)	177 (–424 to 736)	127 (–282 to 546)	78 (–225 to 314)	33 (–81 to 136)
Road traffic injury	–176 (–616 to 118)	–151 (–490 to 93)	–109 (–348 to 65)	–67 (–226 to 46)	–28 (–98 to 18)
Intentional self-harm	67 (–202 to 252)	57 (–153 to 221)	40 (–116 to 160)	25 (–84 to 96)	10 (–33 to 39)
Endocrine diseases	162 (–420 to 707)	140 (–415 to 620)	102 (–278 to 438)	64 (–186 to 273)	28 (–78 to 111)
Diabetes	201 (–312 to 608)	174 (–264 to 526)	128 (–163 to 353)	82 (–118 to 253)	37 (–46 to 105)
Neoplasms	1724 (128 to 3348)	1458 (223 to 2676)	1029 (44 to 1947)	620 (86 to 1144)	247 (13 to 462)

Cause of death	PM_1 targets (μg/m^3)				
	10	20	30	40	55
Pancreas	17 (–275 to 254)	14 (–289 to 212)	10 (–186 to 153)	6 (–102 to 87)	2 (–37 to 31)

Sensitivity Analyses

In the sensitivity analyses, when we used 3-6 dfs for relative humidity, mean temperature, and air pressure and spline function with 5-9 dfs per year for calendar days, the effect sizes remained stable. In the two-pollutant models, after separately including ozone, sulfur dioxide, and carbon monoxide in the main model, the effect estimations of PMs on mortality remained similar and statistically significant (Figures S5-S7 in Multimedia Appendix 1).

Discussion

Principal Findings

To the best of our knowledge, this is the first investigation to assess the associations between size-fractionated PM exposure and deaths from a wide range of causes within the same population. In comparison to PM_{10} and $PM_{2.5}$, we observed a stronger association between PM_1 and mortality risk from cardiorespiratory diseases and neoplasms. Among more specific diseases, significant effect estimates of PM_1 were found among deaths due to hypertensive diseases, chronic rheumatic heart diseases, stroke (notably ischemic stroke), influenza, and pneumonia. Over 5% of all deaths were caused by PMs, with the daily concentrations exceeding the target levels in the WHO's air quality guidelines.

Comparison With Prior Work

Consistent with previous research [8,26-28], we found that the health risk increased with the shrinkage in PM size, with PM_1 ranking the highest in effect estimates. For instance, Zhang et al [8] revealed that for every 10 μg/m^3 increment in PM_1, $PM_{2.5}$, and PM_{10}, the hospital admissions due to cardiovascular diseases increased by 6.7%, 4.5%, and 3.4%, respectively. This gradient in health risk might be mainly attributed to the different particle sizes of PMs, which affect their deposition and absorption in the lungs [7,29]. As the particle size decreases, the surface area increases, leading to a greater catalytic effect on the generation of active oxygen. Smaller molecular size is also conducive to cell absorption, which may further cause the production of high reactive oxygen species, DNA damage, and an increase in interleukin 8 [30]. It is important to note that the relationships between size-fractionated PMs and cause-specific mortality were mostly monotonical in our study, suggesting no safe threshold for these PM pollutants and that any efforts to reduce the PM levels could achieve appreciable health benefits.

A large number of diseases were susceptible to size-fractionated PMs in our study, most of them originating from cardiorespiratory systems. Previous studies mainly focused on the broad categories of diseases and found similar results—the mortality risk was highest for circulatory and respiratory diseases in general [9]. For the subcategories of causes of deaths, significant impact was detected among deaths due to hypertensive diseases, chronic rheumatic heart diseases, stroke (notably ischemic stroke), influenza, and pneumonia for all size-fractionated PMs. This highlights the importance of prioritizing resource allocations during exposure periods and raising awareness of self-protection among patients with these diseases.

It is interesting to note the positive association between PMs and cancer mortality. Attribution bias might partially explain this finding, that is, the incorrect recording of cause of death as cancer, rather than recording the actual causes, if there is a diagnosis of cancer in a person's medical history. In other words, the risk of dying from comorbidities (eg, cardiorespiratory diseases and infection) in patients with cancer increases during highly polluted days, but the cause of death may be falsely assigned to cancer. This bias is almost inevitable, even in developed countries, where death registration data are fairly good [31]. Meanwhile, we cannot exclude the possibility that there is a true relationship between short-term exposure to PMs and cancer mortality. Indeed, there are studies showing the acute impact of PMs on cancer hospitalization and mortality [32,33]. However, we acknowledge that the underlying mechanism remains poorly characterized. Although it is unknown whether this elevated risk of death among patients with cancer is due to competing mortality or cancer-specific mortality, the message is clear—patients with cancer could be vulnerable when exposed to air pollution, and significant efforts are needed to raise awareness among these populations and provide the much-needed protective measures.

Furthermore, the association of size-fractionated PMs with mortality varied by season, age group, and gender. The mortality risks of PMs were stronger during warm seasons, which is in line with previous findings [6,9]. The sources of PMs varied at different temperature levels, and PM_1 may be present in more toxic forms in warmer seasons, potentially transmiting faster and staying longer in the air [34]. During warm seasons, people are prone to open windows and go outdoors more frequently [35], and the absorption of air pollutants may also be enhanced through the temperature regulation system (eg, by increasing sweating, minute ventilation, and cardiac output) [34]. People aged 85 years and older were the most vulnerable to PMs. With the rapid aging in Guangzhou, the death burden is set to increase if stricter controls on air pollution are unavailable in the coming years. The gender effect varies by different health outcomes. For instance, Hu et al [6] (2018) and Yin et al [9] (2020) found that the effect of PM_1 on all-cause mortality among female individuals was larger than that in their male counterparts. Yin et al [27] (2020) reported a stronger impact of PM_1 on male individuals for circulatory diseases. The varying impact of PMs between genders may be partly explained by the biological differences between male and female individuals, such as different gas-blood barrier permeability, particle deposition efficiency, and hormonal status [27]. Another reason may be due to between-gender demographic and behavioral differences, including occupational type, smoking, and lifestyle, which may modify the PM-mortality relationship [6].

Limitations

Some limitations of this study need to be noted. First, there might be diagnostic errors, and therefore, misclassification in the specific causes of death. However, the influence is likely to be small in our study because several quality control measures are available in the Guangzhou death registration system to ensure the accuracy of death classifications [36]. Second, we estimated the death burden of size-fractionated PMs by using several recommended target levels as the reference. However, since no safe threshold was detected for all PMs, our calculations are largely conservative in terms of the health benefits that can be achieved by reducing PMs levels. Finally, as an ecological study, we were unable to explore the individual-level PM-mortality association or control for potential covariates at an individual level, such as smoking habits and occupational exposures. Therefore, the causal relationship cannot be tested in our study. Our study was mainly intended to generate hypotheses on their association, and further research is warranted to test and validate these findings.

Conclusions

Our study indicates that smaller particles are more hazardous to human health. Among specific diseases, significant effect estimates of PM_1 were found for deaths due to hypertensive diseases, chronic rheumatic heart diseases, stroke, influenza, and pneumonia. Our findings highlight the importance of size-based strategies in the control of PMs and management of their health impact.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grants 82003552 and 82204131), Guangdong Basic and Applied Basic Research Foundation (grant 2021A1515110625), and Guangzhou Science and Technology Planning Project (grants 202201011243 and 202201011617).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request and with the permission of Guangzhou Center for Disease Control and Prevention.

Authors' Contributions

JY was in charge of conceptualizing the study, securing resources, writing the original draft of the manuscript, conducting formal analysis, acquiring funding, and providing supervision. HD and CY were responsible for providing resources as well as writing, reviewing, and editing the manuscript. BL contributed to the investigation process as well as writing the original draft, reviewing, and editing the manuscript. GL and BW secured resources and helped in the writing, reviewing, and editing the manuscript. SC, DC, and LH contributed to the investigation process as well as writing, reviewing, and editing the manuscript. ML contributed to the study conceptualization, investigation process, methodology, formal analysis, writing the original draft, funding acquisition, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional statistics.
[DOCX File , 1415 KB - publichealth_v9i1e41862_app1.docx]

References

1. Maji KJ, Ye W, Arora M, Shiva Nagendra SM. PM-related health and economic loss assessment for 338 Chinese cities. Environ Int 2018 Dec;121(Pt 1):392-403 [FREE Full text] [doi: 10.1016/j.envint.2018.09.024] [Medline: 30245362]
2. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
3. Yang J, Zhou M, Li M, Yin P, Hu J, Zhang C, et al. Fine particulate matter constituents and cause-specific mortality in China: a nationwide modelling study. Environ Int 2020 Oct;143:105927 [FREE Full text] [doi: 10.1016/j.envint.2020.105927] [Medline: 32619910]
4. Liu T, Meng H, Yu M, Xiao Y, Huang B, Lin L, et al. Urban-rural disparity of the short-term association of PM with mortality and its attributable burden. Innovation (Camb) 2021 Nov 28;2(4):100171 [FREE Full text] [doi: 10.1016/j.xinn.2021.100171] [Medline: 34778857]

5. Yang J, Zhou M, Zhang F, Yin P, Wang B, Guo Y, et al. Diabetes mortality burden attributable to short-term effect of PM in China. Environ Sci Pollut Res Int 2020 May 23;27(15):18784-18792. [doi: 10.1007/s11356-020-08376-1] [Medline: 32207004]
6. Hu K, Guo Y, Hu D, Du R, Yang X, Zhong J, et al. Mortality burden attributable to PM in Zhejiang province, China. Environ Int 2018 Dec;121(Pt 1):515-522 [FREE Full text] [doi: 10.1016/j.envint.2018.09.033] [Medline: 30292144]
7. Polichetti G, Cocco S, Spinali A, Trimarco V, Nunziata A. Effects of particulate matter (PM(10), PM(2.5) and PM(1)) on the cardiovascular system. Toxicology 2009 Jun 30;261(1-2):1-8. [doi: 10.1016/j.tox.2009.04.035] [Medline: 19379789]
8. Zhang Y, Ding Z, Xiang Q, Wang W, Huang L, Mao F. Short-term effects of ambient PM and PM air pollution on hospital admission for respiratory diseases: case-crossover evidence from Shenzhen, China. Int J Hyg Environ Health 2020 Mar;224:113418. [doi: 10.1016/j.ijheh.2019.11.001] [Medline: 31753527]
9. Yin P, Guo J, Wang L, Fan W, Lu F, Guo M, et al. Higher risk of cardiovascular disease associated with smaller size-fractioned particulate matter. Environ Sci Technol Lett 2020 Jan 29;7(2):95-101. [doi: 10.1021/acs.estlett.9b00735]
10. Wang Y, Li M, Wan X, Sun Y, Cheng K, Zhao X, et al. Spatiotemporal analysis of PM and pancreatic cancer mortality in China. Environ Res 2018 Jul;164:132-139 [FREE Full text] [doi: 10.1016/j.envres.2018.02.026] [Medline: 29486344]
11. Li T, Yan M, Sun Q, Anderson GB. Mortality risks from a spectrum of causes associated with wide-ranging exposure to fine particulate matter: a case-crossover study in Beijing, China. Environ Int 2018 Feb;111:52-59. [doi: 10.1016/j.envint.2017.10.023] [Medline: 29174689]
12. Gariazzo C, Renzi M, Marinaccio A, Michelozzi P, Massari S, Silibello C, BIGEPI Collaborative Group. Association between short-term exposure to air pollutants and cause-specific daily mortality in Italy. A nationwide analysis. Environ Res 2023 Jan 01;216(Pt 3):114676 [FREE Full text] [doi: 10.1016/j.envres.2022.114676] [Medline: 36328229]
13. Guo C, Chan T, Teng Y, Lin C, Bo Y, Chang L, et al. Long-term exposure to ambient fine particles and gastrointestinal cancer mortality in Taiwan: a cohort study. Environ Int 2020 May;138:105640 [FREE Full text] [doi: 10.1016/j.envint.2020.105640] [Medline: 32179321]
14. Peters A, Nawrot TS, Baccarelli AA. Hallmarks of environmental insults. Cell 2021 Mar 18;184(6):1455-1468 [FREE Full text] [doi: 10.1016/j.cell.2021.01.043] [Medline: 33657411]
15. Wei J, Li Z, Guo J, Sun L, Huang W, Xue W, et al. Satellite-derived 1-km-resolution PM concentrations from 2014 to 2018 across China. Environ Sci Technol 2019 Nov 19;53(22):13265-13274. [doi: 10.1021/acs.est.9b03258] [Medline: 31607119]
16. Wang YQ, Zhang XY, Sun JY, Zhang XC, Che HZ, Li Y. Spatial and temporal variations of the concentrations of PM10, PM2.5 and PM1 in China. Atmos Chem Phys 2015 Dec 09;15(23):13585-13598. [doi: 10.5194/acp-15-13585-2015]
17. Chinese Legal Documents on Ethics Review. URL: https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm [accessed 2023-09-26]
18. Guo Y, Barnett AG, Pan X, Yu W, Tong S. The impact of temperature on mortality in Tianjin, China: a case-crossover design with a distributed lag nonlinear model. Environ Health Perspect 2011 Dec;119(12):1719-1725 [FREE Full text] [doi: 10.1289/ehp.1103598] [Medline: 21827978]
19. Yang J, Yang Z, Qi L, Li M, Liu D, Liu X, et al. Influence of air pollution on influenza-like illness in China: a nationwide time-series analysis. EBioMedicine 2023 Jan;87:104421 [FREE Full text] [doi: 10.1016/j.ebiom.2022.104421] [Medline: 36563486]
20. WHO global air quality guidelines: particulate matter (PM25 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240034228 [accessed 2022-06-01]
21. Zhang Y, Lang J, Cheng S, Li S, Zhou Y, Chen D, et al. Chemical composition and sources of PM and PM in Beijing in autumn. Sci Total Environ 2018 Jul 15;630:72-82. [doi: 10.1016/j.scitotenv.2018.02.151] [Medline: 29475115]
22. Gasparrini A, Leone M. Attributable risk from distributed lag models. BMC Med Res Methodol 2014 Apr 23;14(1):55 [FREE Full text] [doi: 10.1186/1471-2288-14-55] [Medline: 24758509]
23. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003 Jan 25;326(7382):219 [FREE Full text] [doi: 10.1136/bmj.326.7382.219] [Medline: 12543843]
24. Chen J, Yang J, Zhou M, Yin P, Wang B, Liu J, et al. Cold spell and mortality in 31 Chinese capital cities: definitions, vulnerability and implications. Environ Int 2019 Jul;128:271-278 [FREE Full text] [doi: 10.1016/j.envint.2019.04.049] [Medline: 31071590]
25. Mela CF, Kopalle PK. The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations. Applied Economics 2002 Mar;34(6):667-677. [doi: 10.1080/00036840110058482]
26. Lin H, Tao J, Du Y, Liu T, Qian Z, Tian L, et al. Particle size and chemical constituents of ambient particulate pollution associated with cardiovascular mortality in Guangzhou, China. Environ Pollut 2016 Jan;208(Pt B):758-766. [doi: 10.1016/j.envpol.2015.10.056] [Medline: 26561449]
27. Yang M, Chu C, Bloom MS, Li S, Chen G, Heinrich J, et al. Is smaller worse? New insights about associations of PM and respiratory health in children and adolescents. Environ Int 2018 Nov;120:516-524. [doi: 10.1016/j.envint.2018.08.027] [Medline: 30153645]

28. Xing X, Hu L, Guo Y, Bloom MS, Li S, Chen G, et al. Interactions between ambient air pollution and obesity on lung function in children: The Seven Northeastern Chinese Cities (SNEC) Study. Sci Total Environ 2020 Jan 10;699:134397. [doi: 10.1016/j.scitotenv.2019.134397] [Medline: 31677469]
29. Samek L, Furman L, Mikrut M, Regiel-Futyra A, Macyk W, Stochel G, et al. Chemical composition of submicron and fine particulate matter collected in Krakow, Poland. Consequences for the APARIC project. Chemosphere 2017 Nov;187:430-439. [doi: 10.1016/j.chemosphere.2017.08.090] [Medline: 28865356]
30. Zou Y, Wu Y, Wang Y, Li Y, Jin C. Physicochemical properties, in vitro cytotoxic and genotoxic effects of PM and PM from Shanghai, China. Environ Sci Pollut Res Int 2017 Aug 5;24(24):19508-19516. [doi: 10.1007/s11356-017-9626-9] [Medline: 28681290]
31. Davies L, Morris L, Hankey B. Increases in thyroid cancer incidence and mortality. JAMA 2017 Jul 25;318(4):389-390. [doi: 10.1001/jama.2017.7906] [Medline: 28742901]
32. Yang Z, Liang H, Deng W, Ji Y. Short-term exposures to PM10 and PM2.5 and lung cancer hospitalization risk in Shenzhen, China: a double negative controls study. Res Sq 2022 Jun 17:1-21 [FREE Full text] [doi: 10.21203/rs.3.rs-1660474/v1]
33. Liu Y, Liu C, Song Y, Zhao J, Gao Y, Zhao L, et al. Short-term association of PM 2.5/PM 10 on lung cancer mortality in Wuhai city, China (2015-2019): a time series analysis. Eur J Cancer Prev 2022 Nov 01;31(6):530-539. [doi: 10.1097/CEJ.0000000000000764] [Medline: 35671253]
34. Li J, Woodward A, Hou X, Zhu T, Zhang J, Brown H, et al. Modification of the effects of air pollutants on mortality by temperature: a systematic review and meta-analysis. Sci Total Environ 2017 Jan 01;575:1556-1570. [doi: 10.1016/j.scitotenv.2016.10.070] [Medline: 27780592]
35. Bergmann S, Li B, Pilot E, Chen R, Wang B, Yang J. Effect modification of the short-term effects of air pollution on morbidity by season: a systematic review and meta-analysis. Sci Total Environ 2020 May 10;716:136985. [doi: 10.1016/j.scitotenv.2020.136985] [Medline: 32044481]
36. Xu X, Dong H, Li L, Liu W, Lin G, Ou C. Trends and seasonality in cause-specific mortality among children under 15 years in Guangzhou, China, 2008-2018. BMC Public Health 2020 Jul 16;20(1):1117 [FREE Full text] [doi: 10.1186/s12889-020-09189-0] [Medline: 32678015]

Abbreviations

eCI: empirical confidence interval
ILI: influenza-like illness
PM: particulate matter
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 12.08.22; peer-reviewed by A Domnich, C Yang; comments to author 15.06.23; revised version received 07.07.23; accepted 29.08.23; published 09.10.23.

Please cite as:
Yang J, Dong H, Yu C, Li B, Lin G, Chen S, Cai D, Huang L, Wang B, Li M
Mortality Risk and Burden From a Spectrum of Causes in Relation to Size-Fractionated Particulate Matters: Time Series Analysis
JMIR Public Health Surveill 2023;9:e41862
URL: https://publichealth.jmir.org/2023/1/e41862
doi:10.2196/41862
PMID:37812487

Original Paper

The Relationship Between Soil-Transmitted Helminth Infections and Environmental Factors in Puerto Iguazú, Argentina: Cross-Sectional Study

Ernesto Candela[1], MSc; Carolina Goizueta[2], MSc; Leonardo Sandon[2], MSc; Carla Muñoz-Antoli[1], PhD; Maria Victoria Periago[2,3], PhD

[1]Department of Pharmacy Pharmaceutical Technology and Parasitology, Universitat de València, Burjassot, Spain
[2]Mundo Sano Foundation, Buenos Aires, Argentina
[3]National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

Corresponding Author:
Ernesto Candela, MSc
Department of Pharmacy Pharmaceutical Technology and Parasitology
Universitat de València
C/ Vicent Andres Estelles s/n
Burjassot, 46100
Spain
Phone: 34 653313156
Email: ercansen@alumni.uv.es

Abstract

Background: Soil-transmitted helminths (STHs) are widely distributed throughout the world. Various factors, including the environment, socioeconomic characteristics, and access to water and sanitation, play an important role in the spread and persistence of these parasites within communities. They, in turn, affect the growth and development of members of the community, especially children. Studies in the northern provinces of Argentina have shown variable prevalence of STHs, but the factors associated with their presence have not been completely elucidated.

Objective: This cross-sectional study aimed to identify the socioeconomic and environmental factors related to STH infection in indigenous villages located in Puerto Iguazú (Misiones), Argentina.

Methods: Between 2018 and 2019, stool samples were collected from individuals ≥1 year residing in 3 villages: Mini-Marangatú, Yriapú, and Fortín Mbororé. Standard parasitological methods were used to determine STH prevalence. Standardized questionnaires were used to assess participants' habits, customs, and household characteristics, and environmental data were obtained through satellite imagery. Multilinear regression with Akaike information criterion stepwise variables was used to explore relevant associations.

Results: A total of 342 individuals from the 3 villages participated in this study. The prevalence of STHs varied across villages: 89.6% (43/48), in Mini-Marangatú, 80.8% (101/125) in Yriapú, and 68.5% (115/169) in Fortín Mbororé. Notably, there was a significant difference in hookworm infection among the villages (*P*=.02). The analysis highlighted the significant influence of specific environmental factors on STH presence and spatial distribution, particularly in relation to hookworm infection. Vegetation patterns represented by the Vegetation Heterogeneity Index, created ad hoc for this study, emerged as a critical factor, with 2 significant predictors related to it (*P*=.002 and *P*=.004) alongside impervious surface density with a significant predictor (*P*<.001). The multilinear regression model yielded a high *F* test score (F_{108}=4.75, *P*<.001), indicating a strong fit (R^2=0.5465). Furthermore, socioeconomic factors, including walking barefoot in houses with dirt floors and overcrowding, were significantly correlated with hookworm infection intensity (*P*<.001 and *P*=.001, respectively). We also used the multilinear regression model to calculate hookworm infection intensity (F_{110}=21.15, *P*<.001; R^2=0.4971).

Conclusions: Our study underscores the complexity of STH transmission, as villages with similar living conditions and environmental characteristics displayed varied STH prevalence and spatial distribution. Specific environmental factors, such as vegetation pattern and impervious surface density, played major roles in STH presence, demonstrating the crucial relationship between environmental factors and hookworm infection distribution. Moreover, our findings emphasize the significant influence of socioeconomic factors on hookworm infection intensity. By gaining insights into this complex interplay, our research contributes to a better understanding of STH transmission characteristics, thereby informing targeted public health interventions for effective control.

(JMIR Public Health Surveill 2023;9:e41568) doi:10.2196/41568

KEYWORDS

soil-transmitted helminths; hookworm; prevalence; intensity: distribution: Iguazú; Argentina

Introduction

Soil-transmitted helminth (STH) infections are the most prevalent among the neglected tropical diseases (NTDs) worldwide, affecting over 1.5 billion people as of 2023 [1]. These diverse diseases have an enormous impact on individuals, families, and communities in low- and middle-income countries, and they constitute a serious obstacle to socioeconomic development and quality of life, leading to loss of productivity and exacerbating poverty [2]. NTDs are mainly prevalent in tropical and subtropical areas, where they mostly affect impoverished communities and disproportionately affect women and children [2,3]. STHs are intestinal parasites (IPs), with the most common species affecting humans being *Ascaris lumbricoides*, *Trichuris trichiura*, and the hookworms *Necator americanus* and *Ancylostoma duodenale* [4]. Although *Strongyloides stercoralis* is a common STH in Latin America and the Caribbean, it is not included in this group due to its complex life cycle and specific characteristics for diagnosis, quantification, and treatment [5,6].

Previous studies have shown the involvement of different factors in the transmission of STHs, including socioeconomic factors, such as the Human Development Index; nutritional and immunological factors; and environmental factors, such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) [7-9]. All these factors exert an economic impact on the population and thus also play a role in the perpetuation of poverty [10,11].

Argentina has a heterogeneous prevalence of STHs (between 0% and 88.9%) throughout the country, and the northeast and northwest provinces of Misiones, Chaco, or Salta are identified as endemic [12-16], with varying rates of infection depending on socioeconomic status, sanitary and environmental conditions, and access to water [16-19]. Despite its high prevalence of STHs, Argentina does not currently have a deworming program as approved by the World Health Assembly (WHA) through resolution WHA54.19 [20].

The Misiones province is composed of 10,218 indigenous people from the Mbyá-Guaraní ethnic group distributed among 116 communities that live mostly in rural areas [21,22] under precarious and poor sanitary conditions, with high rates of malnutrition among children [17]. IPs are highly endemic in this area, especially among indigenous communities [23], with varying prevalence rates depending on age, hygiene habits, access to water and basic sanitation, and nutritional conditions, among others [16,17,24]. These factors can cause or aggravate malnutrition, leading to anemia and growth delays in children [18,25].

Geographic information system (GIS), remote sensing (RS), and digital elevation model (DEM) technologies provide information to identify and analyze determinants of disease distribution, thereby facilitating the development of risk prediction models in relation to environmental variables. This, in turn, aids the design of strategies to identify and prevent these infections. In this study, we collected parasitological and socioeconomic data from different Mbyá-Guaraní villages of Puerto Iguazú, Misiones, Argentina, as well as environmental variables, to determine the main factors associated with STH infection in this region.

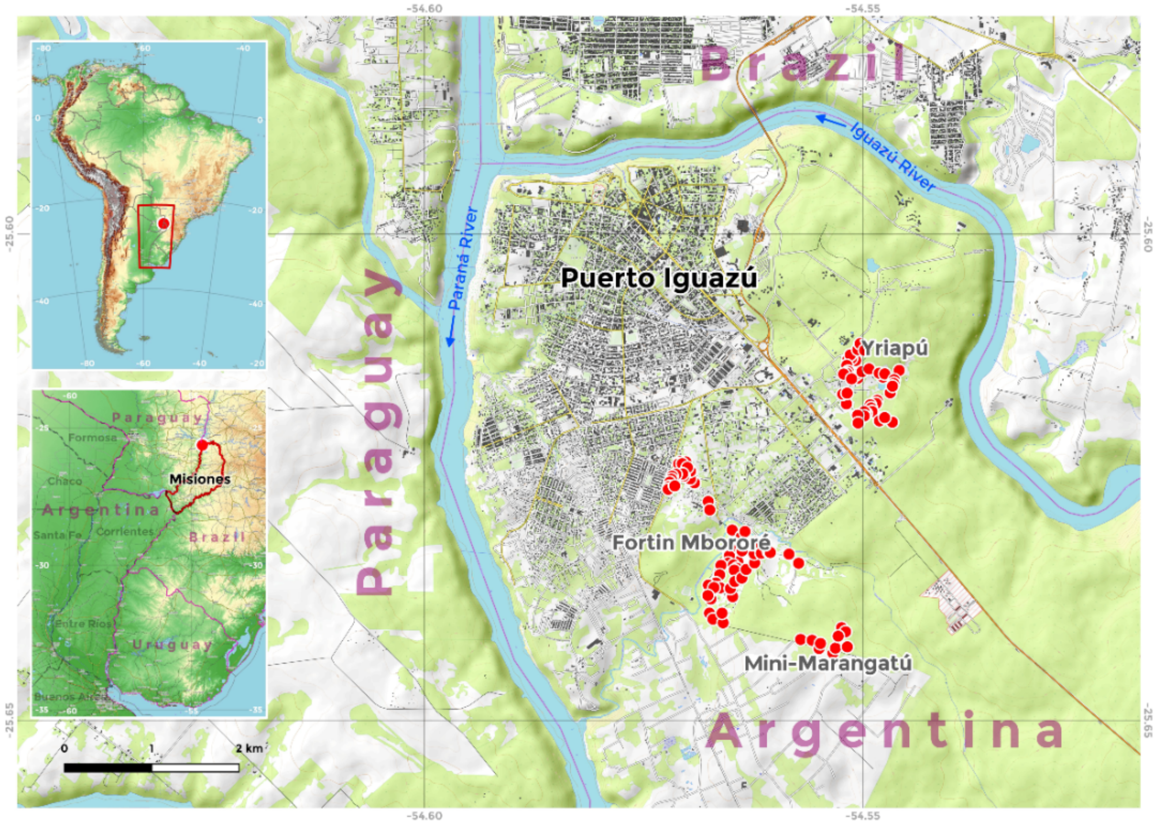
Methods

Study Area

Puerto Iguazu is a city located in the province of Misiones in northeastern Argentina. It is naturally divided by the Paraná and Iguazú Rivers, which act as the physical borders between Argentina, Brazil, and Paraguay. According to the most recent national census data, the city has over 42,800 inhabitants [26].

This study was carried out in the rural area of Puerto Iguazú, around the city´s periphery, where Mbyá Guaraní indigenous communities have settled into 3 villages: Fortín Mbororé, Mini-Marangatú, and Yriapú. Previous studies in this area have shown a high prevalence of IPs, specifically STHs [12,16,27]. Moreover, these villages are adequately sized to enable the enrollment of enough participants to explore the relationship between STH infection and its determinants, with 200 families in Fortín Mbororé, 35 in Mini-Marangatú, and 100 in Yriapú (Figure 1).

Figure 1. Map of the study area. Points represent each of the georeferenced households from the 3 villages included in the study, Fortín Mbororé, Mini-Marangatú, and Yriapú, which are adjacent to Puerto Iguazú, Misiones, Argentina. Map created using QGIS with overlay imagery from OpenTopoMap (CC-BY-SA) via the QuickMapServices plugin (version 0.19.29). Copyright belongs to OpenStreetMap contributors [28]. Map data are from August 27, 2021.



The family subsistence economy of the communities that participated in this study is from guided tours organized to visit the village, handcrafts, and social plans [12]. These communities share similar water and sanitation conditions and are homogeneous in their economic status, with low monthly incomes. Living conditions are characterized by a lack of water and sanitation [19]. Houses are scattered, surrounded by vegetation, and made of adobe bricks with unimproved roofs and dirt floors.

Study Design

Stool samples were collected from participants from all 3 villages. The households were visited, georeferenced, and characterized using a questionnaire. Stool containers were provided, along with verbal instructions on how to collect the samples, and retrieved on the following day. The fresh samples were transported without a fixative in a refrigerated icebox and kept at 4 °C in the lab until analysis within 24 hours of collection. The inclusion criteria were based on age (participants had to be at least 1 year of age) and willingness to participate, evidenced through the informed consent process. Individuals who lived or worked for long periods of time outside the study area and those with conditions that impaired an understanding of the consent process were excluded from the study. Considering that the area is in the tropical forest biome, which involves deforestation of tropical and subtropical forests for agriculture, the extension of grasslands may affect the viability of STH eggs and larvae. Nevertheless, several publications have suggested that the forest mass is in the process of forest transition, where the recovery of natural systems such as forests is taking place [29].

Ethics Approval

The institutional review board of the Ministry of Health of the Province of Misiones approved this study protocol and consent forms (171403/2018). Written consent was obtained from all the parents/guardians for children under 16 years of age. All individuals aged 16 years and older are considered adults in Argentina; therefore, written consent was obtained from them directly.

Parasitological Analysis

The collected stool samples were analyzed using the Ritchie concentration technique, Baermann concentration, and Kato-Katz technique to measure the infection intensity of STH, as previously described [12]. The parasitological parameters used were the prevalence of IPs and the intensity of STH infections in eggs per gram of feces.

Living Conditions and Environmental Characterization

Habits, customs, and household characteristics were analyzed at the household level using GIS. The variables, which were collected using a standardized questionnaire, are summarized in Table 1.

Table 1. Variables analyzed at individual, household, and village levels.

Variables	Description	Results
Education	Education level of household inhabitants	From none to university level
Income	Source of income	Formal employment, public sector, tourism, animal farming, crafts, and social plans, among others
Animal farming	Presence of pets or animal farming	Dogs, cats, chickens, pigs, ducks, and other
Health coverage	Type of health insurance	Public system, private health insurance, and prepaid health insurance, among others
Household characteristics		
Roof	Type of material	Wood, branches, adobe, and metal sheets
Wall	Type of material	Wood, cement, dirt, bricks, and adobe
Floor	Type of material	Wood, dirt floor, cement, and smoothed floor
Electricity	Source of electricity	None, public network, generator, and other
Water source	For human consumption, cooking, or washing	Borehole, tap water, well, and other
Water treatment	Any type of treatment	Boiling, chemical treatment, and other
Excreta disposal	Excretal disposal	Open defecation or latrine
Cooking	Source of heat for cooking	Gas stove or oven, electric stove, wood stove, and other
Garbage disposal	Type of disposal	Municipal system, burning, burying, and other
Behavioral aspects		
Barefoot	Use of footwear	Yes or no
Hand washing	Practice of handwashing before eating and after defecation	Yes or no

Environmental and geographical data were collected using RS together with DEM. From the RS, the Soil-Adjusted Vegetation Index (SAVI), Vegetation Heterogeneity Index (VHI), Enhanced Normalized Difference Impervious Surfaces Index (ENDISI), and Bare Soil Index (BSI) were obtained. From the DEM, the Topographic Position Index (TPI) and Topographic Wetness Index (TWI) were calculated. The characteristics and sources of these indices are detailed in Table 2 and Multimedia Appendix 1.

Table 2. Environmental and geographical indices used to determine their association with the presence of soil-transmitted helminths (STH) in villages from Puerto Iguazú, Misiones, Argentina.

Index	Source	Characteristics
Topographic Position Index [30]	30-m resolution DEM[a] from the IGN[b] [31]	• This index measures the altitude of a point with respect to its surrounding area. • Positive values indicate that the central point is located higher than its average surroundings (such as ridges and hilltops), negative values indicate a position lower than the average (valley and sinkholes), and a near-zero value indicates a flat or continuous slope. • The index used is a multiscale mean of 100-, 200- and 500-meter radii.
Topographic Wetness Index [32]	30-m resolution DEM from IGN [31]	• It identifies potential points of water accumulation (humidity) based on topographic elements. • This index is used as a proxy for the measurement of soil humidity.
Soil-Adjusted Vegetation Index [33]	2019 mean annual surface reflectance from Sentinel 2 imagery [34], retrieved and processed with GEE[c] [35]	• This is a modification of the Normal Density Vegetation Index, which corrects for the brightness of the soil when vegetation is scarce. • It is used to estimate the quantity, quality, and development of vegetation through RSd.
Vegetation Heterogeneity Index	2019 mean annual surface reflectance from Sentinel 2 imagery [34], retrieved and processed with GEE [35]	• The index is used to identify areas of more or less vegetation with respect to their surroundings (80 m radius). • The SD of these values was also mapped to show the heterogeneity of the vegetation in the study area. Low values indicate a subtle variation, while high levels show abrupt changes in the vegetation around each household.
Enhanced Normalized Difference Impervious Surfaces Index [36]	2019 mean annual surface reflectance from Sentinel 2 imagery [34], retrieved and processed with GEE [35]	• It detects impervious surfaces (buildings, asphalted roads, etc). • A threshold of 0.15 was applied to set a value of 1 for impervious areas and 0 for other surfaces. • This binary map was used to get a map with the imperviousness index within a radius of 500 m for each household, with continuous values from 0 to 1.
Bare Soil Index [37]	2019 mean annual surface reflectance from Sentinel 2 imagery [34], retrieved and processed with GEE [35]	• This index combines blue, red, and infrared bands to capture variations in the soil. • Short infrared and red bands are used to quantify the mineral composition of the soil, while blue bands and infrared bands are used for vegetation cover. • Values range between −1 and 1, with higher values indicating bare soil.

[a] DEM: digital elevation model.

[b] IGN: Instituto Geográfico Nacional (National Geographical Institute).

[c] GEE: Google Earth Engine.

[d] RS: remote sensing.

Statistical Analyses

Data were analyzed using Stata 12 software (StataCorp) and RStudio (R Foundation for Statistical Computing). Measures were evaluated using proportion with 95% CIs and means with SDs. The chi-square test was used to compare significant associations between different variables. The age variable was categorized into 5 groups: group 1 (0-5 years), group 2 (6-10 years), group 3 (11-20 years), group 4 (21-40 years), and group 5 (>40 years).

To determine the spatial distribution of STH infection in the study area, an algorithm based on the kernel density estimation (KDE) [38-40] technique was used to identify areas where infection was more prevalent than expected, assuming a homogenous distribution (null hypothesis), by calculating the difference from it using the SD as the unit of measurement. Therefore, positive or negative values indicated higher or lower values of infection than expected under the null hypothesis. With the main results and this same method (a quartic kernel shape and bandwidth of 200 m), the distribution of other variables was calculated: intensity of hookworm infection, households with dirt floors where inhabitants walked barefoot, and households with overcrowding. Variables that showed a significant correlation with the presence and intensity of hookworm infection were then used in a multiple linear regression analysis [41]. Predictors of infection were selected using a stepwise method that selected the best predictors using the Akaike information criterion. Values were considered significant at *P*<.05, with a 95% CI. The full statistical report is available in Multimedia Appendix 2.

Results

Study Population

A total of 342 individuals from the 3 villages participated in this study and provided stool samples: 169 (49.4%) individuals from Fortin Mbororé, 125 (44.5%) from Yriapú, and 48 (14%) from Mini-Marangatú. The population distribution in the 3 communities was 53.8% (91/168) men and 46.2% (78/169) women in Fortin Mbororé, 56.8% (71/125) men and 43.2% (54/125) women in Yriapú, and 50% (24/48) for both sexes in Mini-Marangatú. The mean age of participants was 21 (SD 17.9) years in Fortin Mbororé, 10.4 (SD 11.43) years in Yriapú, and 15 (SD 12.21) years in Mini-Marangatú.

Prevalence of IPs

The overall prevalence of IPs in the 3 villages was 95.8% (46/48) in Mini-Marangatú, 95.2% (119/125) in Yriapú, and 91.1% (154/169) in Fortín Mbororé. Protozoan infection ranged from 87.5% (42/48) in Mini-Marangatú to 81.6% (102/125) in Yriapú, while helminth infections were highest in Mini-Marangatú (44/448, 91.7%), followed by Yriapú (104/125, 83.2%) and lower in Fortin Mbororé (126/169, 74.6%) (Table 3). The STH prevalence was 68.1% (115/169) in Fortín Mbororé, 80.8% (101/125) in Yriapú, and 89.6% (43/48) in Mini-Marangatú. Infection caused by *T. trichiura* was only detected in Mini-Marangatú village, with only 1 (2.1%) case, and no *A. lumbricoides* infections were detected in this village. The most prevalent STH was hookworm, reaching statistically different infection rates ($\chi^2_2=7.6$, $P=.02$) between the 3 villages: Mini-Marangatú (42/48, 87.5%), Yriapú (92/125, 73.6%), and Fortín Mbororé (114/169, 67.5%). The descriptive characteristics of STH infections in the 3 villages are provided in Table 3.

Table 3. Descriptive characteristics and prevalence of intestinal parasites in individuals from Fortin Mbororé, Yriapú, and Mini-Marangatú.

Characteristics	Fortin Mbororé	Yriapú	Mini-Marangatú
Age (years), mean (SD)	21 (17.9)	10.4 (11.43)	15 (12.21)
Age range	1-87	1-54	1-49
Gender, n (%)			
Female	78 (46.2)	54 (43.2)	24 (50)
Male	91 (53.8)	71 (56.8)	24 (50)
Prevalence of protozoans, n (%); range (95% CI)	138 (81.7); 75-86.8	102 (81.6); 73.7-87.5	42 (87.5); 74.2-94.4
Entamoeba coli	69 (40.8); 33.6-48.5	73 (58.4); 49.5-66.8	22 (45.8);32-60.4
Entamoeba complex	18 (10.7); 6.8-16.3	12 (9.6);5.5-16.3	3 (6.3);1.9-18.3
Entamoeba hartmanni	19 (11.2); 7.3-17	24 (19.2); 13.1-27.2	12 (25); 14.5-39.6
Endolimax nana	34 (20.1); 14.7-26.9	13 (10.4); 6.1-17.2	10 (20.8); 11.3-35.2
Iodamoeba butschlii	10 (5.9); 3.2-10.7	2 (1.6);0.4-6.3	1 (2.1); 0.3-14.2
Chilomastix mesnili	6 (3.6); 1.6-7.7	16 (12.8); 7.9-20	6 (12.5); 5.6-25.8
Giardia intestinalis	40 (23.7); 17.8-30.7	38 (30.4); 22.9-39.1	14 (29.2); 17.8-44
Blastocystis spp.	93 (55); 47.4-62.4	56 (44.8); 36.2-53.7	32 (66.7); 51.8-78.8
Prevalence of helminths, n (%); range (95% CI)	126 (74.6); 67.4-80.6	104 (83.2); 75.5-88.8	44 (91.7); 79.2-96.9
Enterobius vermicularis	3 (1.8); 0.6-5.4	5 (4); 1.7-9.4	—[a]
Hymenolepis nana	28 (16.6); 11.6-23	36 (28.8); 21.5-37.5	6 (12.5); 5.6-25.8
Trichuris trichiura	—	—	1 (2.1); 0.3-14.2
Ascaris lumbricoides	3 (1.8); 0.6-5.4	21 (17.2); 11.4-25.1	—
Hookworm	114 (67.5); 60-74.2	92 (73.6); 65.1-80.7	42 (87.5); 74.2-94.4
Strongyloides stercoralis	11 (6.5); 3.6-11.4	44 (35.48); 27.5-44.4	14 (29.2); 17.8-44

[a] —: not available.

Hookworm infection was higher in the age groups ranging from 0 to 5 years and from 6 to 10 years in Yriapú and lower in Fortin Mbororé, especially in the western area located near the urban area of Puerto Iguazú. Moreover, statistical differences between age groups were observed within Fortin Mbororé ($\chi^2_4=27.9$, $P<.001$) and Mini-Marangatú ($\chi^2_4=17.5$, $P=.002$). Mixed infections with different STH species were also observed (Tables 4 and 5), together with the intensity of infection. *Trichuris trichiura* and *A. lumbricoides* were present mainly as light-intensity infections. On the other hand, heavy-intensity hookworm infections were detected in all 3 villages. The highest rate of individuals with heavy infections (22/42, 52.4%) were found in Mini-Marangatú, but no statistical differences were observed between the types of intensity.

Table 4. Intensity of soil-transmitted helminth (STH) infections in individuals from Fortin Mbororé, Yriapú, and Mini-Marangatú.

Infections	Fortin Mbororé, n (%)			Yriapú, n (%)			Mini-Marangatú, n (%)		
	Light	Moderate	Heavy	Light	Moderate	Heavy	Light	Moderate	Heavy
Hookworm	67 (58.8)	12 (10.5)	35 (30.7)	61 (66.3)	11 (12)	20 (21.7)	15 (35.7)	5 (11.9)	22 (52.4)
A. lumbricoides	2 (66.7)	—[a]	1 (33.3)	12 (57.1)	6 (28.6)	3 (14.3)	—	—	—
T. trichiura	—	—	—	—	—	—	1 (100)	—	—

[a] —: not available.

Table 5. Number of mixed soil-transmitted helminth (STH) infections in individuals from Fortin Mbororé, Yriapú, and Mini-Marangatú.

Mixed STH infections	Fortin Mbororé, n (%)	Yriapú, n (%)	Mini-Marangatú, n (%)
Hookworm/*S. stercoralis*	10 (76.9)	37 (56.9)	13 (92.9)
Hookworm/*A. lumbricoides*	3 (23.1)	19 (29.2)	—[a]
Hookworm/*S. stercoralis*/*A. lumbricoides*	—	9 (13.9)	—
Hookworm/*S. stercoralis*/*T. trichiura*	—	—	1 (7.1)

[a] —: not available.

Living Conditions

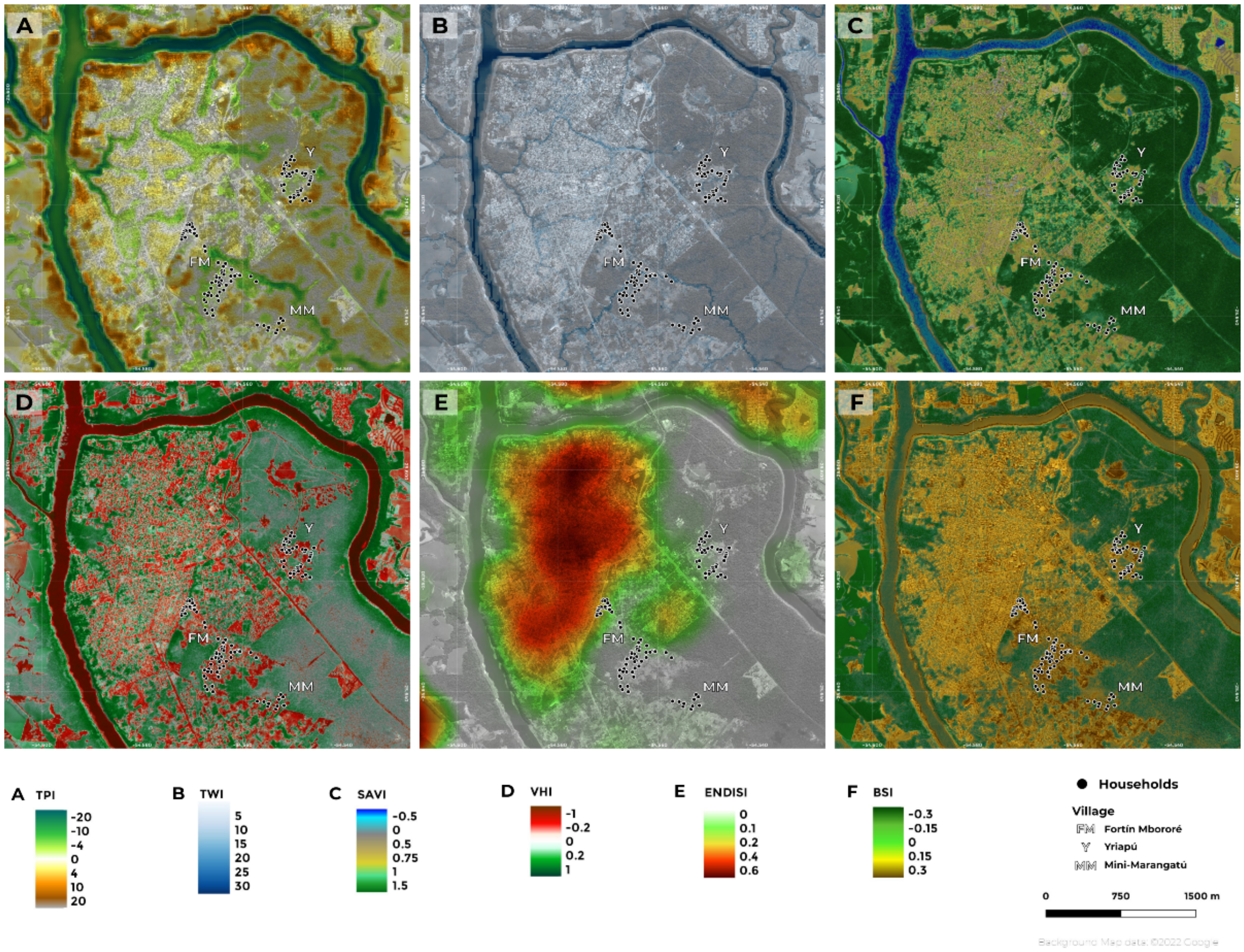
Living conditions between the villages were usually similar. The average number of inhabitants per household was 5.3 for Fortin Mbororé, 5.1 for Yriapú, and 5.8 for Mini-Marangatú. Most households had a single room for sleeping; therefore, overcrowding was common. Generally, houses were made of wooden walls and dirt floors—90% (43/48) in Mini-Marangatú and 50% (61/121) in Yriapú. In the case of Fortin Mbororé, this figure was reduced to 35% (58/166) since 40% (66/166) of the households had cement floors. Practically the entire population, both children and adults from all 3 villages, walked barefoot. Although 22.9% (38/166) of households practiced open defecation, most had a latrine that consisted of a simple ground excavation. With respect to the source of drinking water, all the families in Mini-Marangatú obtained their water from boreholes, along with 75% (92/121) in Yriapú and 49% (81/166) in Fortin Mbororé. Family incomes were low and precarious, mostly coming from animal farming, crafts, or social plans. In the newer village of Mini-Marangatú, which branched off from Fortin Mbororé, 60% (99/166) of the families obtained their income from crafts.

Only the type of floor was observed to be associated with hookworm transmission, with a significantly higher prevalence of hookworm found in individuals from Yriapú village living in dirt floor houses (χ^2_3=8.8, P=.03). Lack of sanitation and hygiene, water source, the use of a latrine with simple ground excavation, and source of income were not related to a higher prevalence of hookworms. Walking barefoot and living in overcrowding conditions were significantly related to the intensity of hookworm infection (P<.001 and P=.003, respectively; F_{110}=46.2, P<.001), indicating the significance of the model (R^2=0.46).

Environmental Characterization

Figure 2 shows the distribution of the different environmental indices in the study area. TPI helps discriminate between areas that are depressed and those that have some prominence and thus are less prone to accumulating water, and TWI detects hydrological flow paths and thus proximity to streaming water and other bodies of water. The values of these 2 indices indicate that the study area was irregular with depressed areas and small hills (Figure 2A), with many water courses running through it, including the Mbocay stream (Figure 2B). Yriapú is located at the edge of a slightly depressed area, while Fortin Mbororé is divided into 2 parts by the Mbocay stream. Both sides of the village are on flat ground. Mini-Marangatú is located close to the crest of a small hill.

Figure 2. Distribution of the different indexes used in the study area from Puerto Iguazú, Misiones, Argentina. (A) Topographic Position Index (TPI). (B) Topographic Wetness Index (TWI). (C). Soil-Adjusted Vegetation Index (SAVI). (D) Vegetation Heterogeneity Index (VHI). (E) Enhanced Normalized Difference Impervious Surface Index (ENDISI). (F) Bare Soil Index (BSI). Map created using QGIS with background imagery from Google Maps via the QuickMapServices plugin (version 0.19.29). Copyright 2021 Google. Map data 2021 Google.

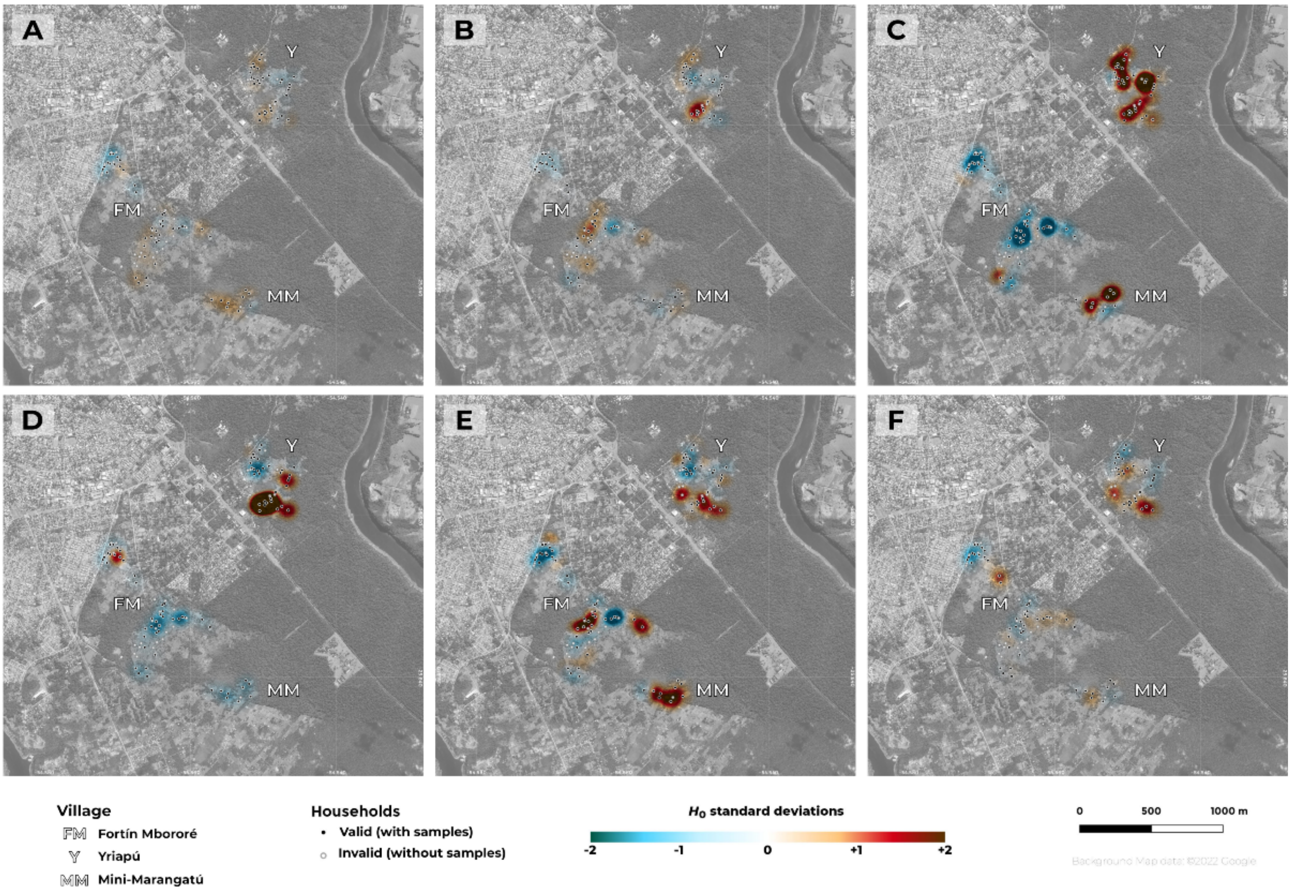


For the presence of vegetation, SAVI and VHI were used, with SAVI measuring vigor and VHI estimating the heterogeneity of the vegetation landscape. As depicted in Figures 2C and D, some differences in the distribution of the SAVI were observed between the villages. Western Fortin Mbororé had lower values than the central area of the village, while both Mini-Marangatú and Yriapú were surrounded by more vigorous vegetation (Figure 2C). Through the VHI, the difference between bare soil and the presence of vegetation was greatest in Mini-Marangatú and very small in Fortin Mbororé. The other 2 indices, ENDISI and BSI, were used to indicate the presence of bare soil (Figure 2E and 2F) but with different focuses: ENDISI on urbanized areas and BSI on natural soil. Again, Fortin Mbororé had a greater presence of bare soil around the houses, especially in the western area where households were located close to the urban city of Puerto Iguazú, whereas both Yriapú and Mini-Marangatú had patches of bare soil and vegetation.

Spatial Distribution of STH

The KDE technique was used to analyze the spatial distribution of STH in the study area and observe differences between the villages. As shown in Figure 3A, the distribution of hookworm-positive individuals was not entirely homogenous, and the differences in the SD were not pronounced and were present throughout the entire study area. There was a slight concentration of cases in Mini-Marangatú, followed by Yriapú. Some differences were observed within Fortin Mbororé, where households closer to the urbanized area of Puerto Iguazú (northwest) had lower values than those farther away (southwest).

Figure 3. Spatial distribution obtained using the kernel density estimation (KDE) technique for (A) hookworm infection, (B) hookworm intensity, (C) Strongyloides stercoralis infection, (D) Ascaris lumbricoides infection, (E) households with dirt floors where inhabitants walk barefoot, and (F) households with overcrowding. Map created using QGIS with background imagery from Google Maps via the QuickMapServices plugin (version 0.19.29). Copyright 2021 Google. Map data 2021 Google.



With respect to the intensity of hookworm infection (Figure 3B), positive cases from Yriapú were mostly of light intensity, while higher-intensity cases were clustered only in the southern area of Yriapú. Although the prevalence of other STH was lower than that of hookworm, the KDE analysis also showed a heterogeneous distribution, with a marked difference in the SD of *S. stercoralis*, showing a high concentration of positive cases in Yriapú and Mini-Marangatú compared to Fortin Mbororé (Figure 3C). *Ascaris lumbricoides* infections were detected mostly in Yriapú village (Figure 3D).

Figure 3E shows the distribution of households with dirt floors and individuals who usually walk barefoot. Given that many houses in Fortin Mbororé had cement floors, the difference in the SD was more evident in Mini-Marangatú and southwestern Fortin Mbororé compared to Yriapú. The distribution of households with overcrowding (Figure 3F) was homogeneous in the study area, although there was a clustering in Yriapú.

The multivariate model used to identify the factors most related to the distribution of hookworm in the study area selected the following predictors: mean VHI, SD of VHI, and ENDISI and households with dirt floors and individuals who walk barefoot. In this model, only the environmental variables were significantly associated (P<.001) with the presence of hookworm infection, which explained 56% (248/342) of the variability observed in the distribution of cases in the study area (F_{108}=34.75, P<.001). The predictors for hookworm intensity selected by the multivariate model were the distribution of dirt floor households with individuals that walked barefoot and households with overcrowding. Both variables showed a significant association (P<.001) with the intensity of hookworm infection and predicted 45% (153/342) of the variability observed in the distribution of the cases.

Discussion

Principal Findings

This study uncovers factors associated with the presence and intensity of STH infection, especially hookworm, in rural areas of Puerto Iguazú. The high prevalence observed coincides with previous studies in other rural areas of the region [5,15,42,43], as well as a higher prevalence in the infant population [3,5]. Hookworm prevalence was above 70% in all villages (114/169 in Fortín Mbororé, 92/125 in Yriapú, and 42/48 in Mini-Marangatú), with the highest prevalence observed in Mini-Marangatú, located in the easternmost area.

In the multivariate analysis for hookworm infection, barefoot walking and overcrowding were implicated in an increase in the intensity of infection, while the type of household floor material was associated with an increase in transmission. These results coincide with those already observed in other studies [12,44,45]. Although other studies indicate that hygienic conditions, water, type of latrine, and level of education are related to a higher hookworm prevalence [16,19,46-48], in this study, these factors did not show statistical significance. These different results may be due to the uniformity of these variables throughout the different villages since most of the households shared the same characteristics. Moreover, although the villages were independent population centers managed by their own community leader or cacique, they were relatively close to each other, thus sharing similar environmental and economic conditions.

Since hygienic conditions have been shown to greatly influence transmission through the fecal-oral route [19], as in the case of *A. lumbricoides*, the proximity to the Mbocay stream in the villages of Fortin Mbororé (eastern zone) and Mini-Marangatú could have generated different *A. lumbricoides* infection rates between the villages. In addition, the high density of construction observed in the western area of Fortin Mbororé, through the ENDISI, may have affected the transmission of this parasite, given that low humidity can hinder egg embryonation [9].

With respect to environmental factors and their predictors, previous studies have highlighted the role of humidity, temperature, and soil type in the transmission of STH [16,45,48,49]. Higher humidity and temperature have been shown to be favorable conditions for the survival of heterogonic stages of *S. stercoralis* [50,51]. A joint analysis of the TPI and TWI allowed us to determine that the village of Yriapú is situated in a depressed area with higher humidity accumulation, which could be a factor behind the high prevalence of *S. stercoralis* observed in this village. Available studies also show the association of humidity and vegetation indices in the prevalence of *S. stercoralis* [52].

To analyze the effect of vegetation on the transmission of STH, the VHI, which was developed specifically for this study, helped identify areas with more or less vegetation with respect to each surrounding household. The algorithm is similar to the TPI; however, instead of elevation, the data source is a vegetational index (in this case, SAVI). High values of this index around households were associated with increased hookworm infection, which indicates the importance not only of the vigorousness of vegetation but also its distribution pattern in the landscape of the study area. High values of this index around the Mini-Marangatú households coincided with the highest prevalence of hookworm in this area. These vegetation and humidity indices, which demonstrate the influence of vegetation on hookworm distribution, coincided with the risk of hookworm infection in areas of VHI values reported in other studies [44,49,53].

Previous studies suggested that infection is more probable in households surrounded by bare soil, among other factors [52,54-56]. Fortin Mbororé has bare soil surrounding its houses, especially in the western area located close to the urbanized area of the city. Although bare soil is an environmental factor that has been shown to be associated with STH infections, this area had lower hookworm infection rates.

The ENDISI aids in detecting impervious surfaces, particularly the amount of water stored. The survival of hookworm larvae depends on the soil's water-retaining properties. When the soil dries out, the water is restricted to the thin film around individual soil particles, and the infective larval stage remains quiescent in the moisture film until it makes contact with its host [57]. The lower hookworm infection rates observed in Fortin Mbororé may be because this area is highly influenced by the ENDISI since urban development and infrastructure density modify soil permeability and moisture. This can have a negative effect on the persistence and survival of this type of parasite, which requires humidity to survive and develop its larval stages [51,58].

In general, villages sharing similar living conditions and environmental characteristics contributed to the high infection rates observed. This study demonstrated that living conditions play a role in the intensity of hookworm infection, and environmental variables are significantly associated with its presence. However, the specific differences observed in certain areas aid in elucidating how human development and social and sanitary conditions may influence lower infection rates among individuals and villages located in endemic areas. Evidence from previous studies and the results obtained herein show that environmental factors such as temperature, vegetation, and humidity play a role in the presence and maintenance of STHs in the soil. Therefore, environmental changes caused by climate change could modify the distribution of these parasites, although deforestation, bare soil, high temperature, and lack of humidity could restrict their presence.

The limitations of this study include the sensitivity of the techniques used to detect STHs since low burdens of infection could be missed. Additionally, in this area, hookworm was the most prevalent STH; therefore, studies conducted in areas with a greater presence of other species of the group would be beneficial. Further studies in more heterogeneous communities with similar environmental characteristics could help aid our understanding of the socioeconomic and building characteristics that determine the presence and intensity of STH infections.

Conclusion

This study, conducted in an endemic area for STHs, especially hookworms, reinforces the importance of the environment in the establishment of this group of parasites, which require passage through the soil for their development. Additionally, we observed that living conditions, like walking barefoot, having dirt floors, or overcrowding, are associated with the intensity of hookworm infection. Given that environmental variables cannot be changed, it is important to work on those aspects that can be modified, such as the characteristics of the house, the availability of water and sanitation, and periodic deworming as suggested by the World Health Organization deworming guidelines [59].

Acknowledgments

We appreciate the participation of the community from Fortín Mbororé, Yriapú, and Mini-Marangatú and permission to work in the village from their leaders (caciques). We would like to thank all our colleagues from Mundo Sano's office in Iguazú for collaborating with us on the study logistics and data collection. Finally, we would like to acknowledge the institutional support from Marcelo Abril (Mundo Sano) and Marcelo Scavuzzo and his team from Instituto de Altos Estudios Espaciales "Mario Gulich," Comisión Nacional de Actividades Espaciales (CONAE).

Data Availability

Data supporting the conclusions of this article are included within the article. The datasets used and/or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

EC carried out laboratory analyses, analyzed and interpreted the data, and wrote the draft of the manuscript. CG was involved in the acquisition of data and community relations with the cacique. LS developed, analyzed, and interpreted the mathematical models as well as the figures. CMA was involved in the conception and design of the study, analysis, and interpretation, and revising the final draft of the manuscript. MVP was involved in the conception and design of the study, acquisition of data, analysis and interpretation, and revision of the final draft of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Environmental index formulas.
[DOCX File , 16 KB - publichealth_v9i1e41568_app1.docx]

Multimedia Appendix 2
Report of the statistical models used.
[DOCX File , 20 KB - publichealth_v9i1e41568_app2.docx]

References

1. Soil-transmitted helminth infections. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/soil-transmitted-helminth-infections [accessed 2023-07-20]
2. Working to overcome the global impact of neglected tropical diseases: first WHO report on neglected tropical diseases. World Health Organization. URL: https://apps.who.int/iris/handle/10665/44440 [accessed 2022-04-30]
3. Harhay MO, Horton J, Olliaro PL. Epidemiology and control of human gastrointestinal parasites in children. Expert Rev Anti Infect Ther 2010 Feb 10;8(2):219-234 [FREE Full text] [doi: 10.1586/eri.09.119] [Medline: 20109051]
4. Eliminating soil-transmitted helminthiases as a public health problem in children. Progress report 2001-2010 and strategic plan 2011-2020. World Health Organization. Geneva, Switzerland; 2012. URL: https://iris.who.int/bitstream/handle/10665/44804/9789241503129_eng.pdf?sequence=1 [accessed 2022-04-30]
5. Saboyá M, Catalá L, Ault S, Nicholls R. Prevalence and intensity of infection of soil-transmitted helminths in Latin America and the Caribbean countries: Mapping at second administrative level 2000-2010. Pan American Health Organization. 2011. URL: https://www3.paho.org/hq/dmdocuments/2011/STH-segundo-nivel-prevalence-LAC-MARCH-2011.pdf [accessed 2022-04-30]
6. Siddiqui AA, Berk SL. Diagnosis of Strongyloides stercoralis infection. Clin Infect Dis 2001 Oct 01;33(7):1040-1047. [doi: 10.1086/322707] [Medline: 11528578]
7. Chammartin F, Guimarães LH, Scholte RG, Bavia ME, Utzinger J, Vounatsou P. Spatio-temporal distribution of soil-transmitted helminth infections in Brazil. Parasit Vectors 2014 Sep 18;7(1):440 [FREE Full text] [doi: 10.1186/1756-3305-7-440] [Medline: 25230810]
8. Pullan RL, Kabatereine NB, Quinnell RJ, Brooker S. Spatial and genetic epidemiology of hookworm in a rural community in Uganda. PLoS Negl Trop Dis 2010 Jun 15;4(6):e713 [FREE Full text] [doi: 10.1371/journal.pntd.0000713] [Medline: 20559556]
9. Chammartin F, Scholte RG, Malone JB, Bavia ME, Nieto P, Utzinger J, et al. Modelling the geographical distribution of soil-transmitted helminth infections in Bolivia. Parasit Vectors 2013 May 25;6(1):152 [FREE Full text] [doi: 10.1186/1756-3305-6-152] [Medline: 23705798]

10. Hotez P, Bundy DAP, Beegle K, Brooker S, Drake L, De SN. Helminth infections: soil-transmitted helminth infections and schistosomiasis. In: Disease Control Priorities in Developing Countries. 2nd edition. Washington,: Oxford University Press; 2006:467-482.
11. Alvarez Di Fino EM, Rubio J, Abril MC, Porcasi X, Periago MV. Risk map development for soil-transmitted helminth infections in Argentina. PLoS Negl Trop Dis 2020 Feb 10;14(2):e0008000. [doi: 10.1371/journal.pntd.0008000]
12. Candela E, Goizueta C, Periago MV, Muñoz-Antoli C. Prevalence of intestinal parasites and molecular characterization of Giardia intestinalis, Blastocystis spp. and Entamoeba histolytica in the village of Fortín Mbororé (Puerto Iguazú, Misiones, Argentina). Parasit Vectors 2021 Oct 01;14(1):510 [FREE Full text] [doi: 10.1186/s13071-021-04968-z] [Medline: 34598722]
13. Fleitas PE, Floridia-Yapur N, Nieves EE, Echazu A, Vargas PA, Caro NR, et al. Strongyloides stercoralis and Trypanosoma cruzi coinfections in a highly endemic area in Argentina. PLoS Negl Trop Dis 2022 Feb 4;16(2):e0010179. [doi: 10.1371/journal.pntd.0010179]
14. Cimino RO, Fleitas P, Fernández M, Echazú A, Juarez M, Floridia-Yapur N, et al. Seroprevalence of the strongyloides stercoralis infection in humans from Yungas Rainforest and Gran Chaco Region from Argentina and Bolivia. Pathogens 2020 May 20;9(5):394. [doi: 10.3390/pathogens9050394]
15. Socías ME, Fernández A, Gil JF, Krolewiecki AJ. Soil transmitted helminthiasis in Argentina: A systematic review. Medicina (Buenos Aires) 2014;74(1):29-36 [FREE Full text]
16. Rivero MR, De Angelo C, Nuñez P, Salas M, Motta CE, Chiaretta A, et al. Environmental and socio-demographic individual, family and neighborhood factors associated with children intestinal parasitoses at Iguazú, in the subtropical northern border of Argentina. PLoS Negl Trop Dis 2017 Nov 20;11(11):e0006098 [FREE Full text] [doi: 10.1371/journal.pntd.0006098] [Medline: 29155829]
17. Zonta M, Oyhenart E, Navone G. Nutritional status, body composition, and intestinal parasitism among the Mbyá-Guaraní communities of Misiones, Argentina. Am J Hum Biol 2010;22(2):193-200. [doi: 10.1002/ajhb.20977] [Medline: 19642206]
18. Gamboa MI, Navone GT, Orden AB, Torres MF, Castro LE, Oyhenart EE. Socio-environmental conditions, intestinal parasitic infections and nutritional status in children from a suburban neighborhood of La Plata, Argentina. Acta Trop 2011 Jun;118(3):184-189. [doi: 10.1016/j.actatropica.2009.06.015] [Medline: 19577532]
19. Echazú A, Bonanno D, Juarez M, Cajal SP, Heredia V, Caropresi S, et al. Effect of poor access to water and sanitation as risk factors for soil-transmitted helminth infection: selectiveness by the infective route. PLoS Negl Trop Dis 2015 Sep 30;9(9):e0004111 [FREE Full text] [doi: 10.1371/journal.pntd.0004111] [Medline: 26421865]
20. Schistosomiasis and soil-transmitted helminth infections. World Health Assembly. 2001 May 22. URL: https://www.who.int/publications/i/item/wha54.19 [accessed 2022-04-20]
21. Portal de Datos Justicia Argentina. Instituto Nacional de Asuntos Indígenas (INAI). URL: http://datos.jus.gob.ar/dataset/listado-de-comunidades-indigenas/archivo/8f9af332-83ff-4ea6-a6a5-7a90371a41fb [accessed 2022-04-20]
22. Estructura demográfica de la población Mbya Guaraní Provincia de Misiones. Ministerio de Salud de la Provincia de Misiones. 2019. URL: https://salud.misiones.gob.ar/wp-content/uploads/2019/10/ESTRUCTURA-DEMOGRAFICA-DE-POBLACION-MBYA-GUARANI.pdf [accessed 2022-04-22]
23. Chin YT, Lim YAL, Chong CW, Teh CSJ, Yap IKS, Lee SC, et al. Prevalence and risk factors of intestinal parasitism among two indigenous sub-ethnic groups in Peninsular Malaysia. Infect Dis Poverty 2016 Jul 18;5(1):77 [FREE Full text] [doi: 10.1186/s40249-016-0168-z] [Medline: 27430215]
24. Navone GT, Gamboa MI, Oyhenart EE, Orden AB. Intestinal parasitosis in Mbyá-Guaraní populations from Misiones Province, Argentina: epidemiological and nutritional aspects. Cad Saude Publica 2006 May;22(5):1089-1100 [FREE Full text] [doi: 10.1590/s0102-311x2006000500022] [Medline: 16680362]
25. Ellwanger JH, Ziliotto M, Kulmann-Leal B, Chies JAB. Iron deficiency and soil-transmitted helminth infection: classic and neglected connections. Parasitol Res 2022 Dec 19;121(12):3381-3392. [doi: 10.1007/s00436-022-07697-z] [Medline: 36258094]
26. El municipio de Puerto Iguazú. Municipalidades en Argentina. URL: https://www.municipalidad-argentina.com.ar/municipalidad-puerto-iguazu.html [accessed 2022-04-22]
27. Rivero M, De Angelo C, Nuñez P, Salas M, Liang S. Intestinal parasitism and nutritional status among indigenous children from the Argentinian Atlantic Forest: Determinants of enteroparasites infections in minority populations. Acta Trop 2018 Nov;187:248-256. [doi: 10.1016/j.actatropica.2018.08.015] [Medline: 30125528]
28. OpenStreetMap Foundation. Copyright and License.: OpenStreetMap URL: https://www.openstreetmap.org/copyright [accessed 2023-10-25]
29. Izquierdo A, De Angelo CD, Aide T. Thirty Years of human demography and land-use change in the Atlantic forest of Misiones, Argentina: an evaluation of the forest transition model. Ecol Soc 2008;13(2) [FREE Full text] [doi: 10.5751/es-02377-130203]
30. Weiss AD. Topographic position and landforms analysis. 1999 Presented at: 4th Symposium on the Environmental Monitoring and Assessment Program (EMAP); April 6-8; Seattle, WA.
31. Modelo digital de elevaciones de la República de Argentina versión 2. Instituto Geográfico Nacional República Argentina. 2019. URL: https://www.ign.gob.ar/archivos/Informe_MDE-Ar_v2.0_30m.pdf [accessed 2022-01-30]

32. Mitchell-Box K, Braun KL. Fathers' thoughts on breastfeeding and implications for a theory-based intervention. J Obstet Gynecol Neonatal Nurs 2012;41(6):E41-E50. [doi: 10.1111/j.1552-6909.2012.01399.x] [Medline: 22861175]
33. da Silva VS, Salami G, da Silva MIO, Silva EA, Monteiro Junior JJ, Alba E. Methodological evaluation of vegetation indexes in land use and land cover (LULC) classification. Geol Ecol Landsc 2019 Apr 28;4(2):159-169. [doi: 10.1080/24749508.2019.1608409]
34. Sentinel-2 MSI: MultiSpectral Instrument, Level-2A. Google Developer. URL: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR [accessed 2022-01-31]
35. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ 2017 Dec;202:18-27 [FREE Full text] [doi: 10.1016/j.rse.2017.06.031]
36. Chen J, Yang K, Chen S, Yang C, Zhang S, He L. Enhanced normalized difference index for impervious surface area estimation at the plateau basin scale. J Appl Rem Sens 2019 Jan 1;13(01):1. [doi: 10.1117/1.jrs.13.016502]
37. Wanhui C, Liangyun L, Chao Z, Jihua W, Jindi W. Monitoring the seasonal bare soil areas in Beijing using multitemporal TM images. 2004 Presented at: IEEE International Geoscience and Remote Sensing Symposium; Sept 20-24; Anchorage, AK p. 3379-3382.
38. Chen Y. A tutorial on kernel density estimation and recent advances. Biostat Epidemiol 2017 Dec 01;1(1):161-187. [doi: 10.1080/24709360.2017.1396742]
39. Mitchell-Box K, Braun KL. Fathers' thoughts on breastfeeding and implications for a theory-based intervention. J Obstet Gynecol Neonatal Nurs 2012;41(6):E41-E50. [doi: 10.1111/j.1552-6909.2012.01399.x] [Medline: 22861175]
40. Michael Kramer. Spatial Epidemiology.: GitHub; 2021. URL: https://mkram01.github.io/EPI563-SpatialEPI/EPI563-SpatialEPI-eBook.pdf [accessed 2022-01-31]
41. Rodrigo JA. Introducción a la regresión lineal simple. Condiciones. 2016. URL: https://www.cienciadedatos.net/documentos/25_regresion_lineal_multiple.html [accessed 2022-01-31]
42. Echazú A, Juarez M, Vargas PA, Cajal SP, Cimino RO, Heredia V, et al. Albendazole and ivermectin for the control of soil-transmitted helminths in an area with high prevalence of Strongyloides stercoralis and hookworm in northwestern Argentina: A community-based pragmatic study. PLoS Negl Trop Dis 2017 Oct 9;11(10):e0006003 [FREE Full text] [doi: 10.1371/journal.pntd.0006003] [Medline: 28991899]
43. Periago MV, García R, Astudillo OG, Cabrera M, Abril MC. Prevalence of intestinal parasites and the absence of soil-transmitted helminths in Añatuya, Santiago del Estero, Argentina. Parasit Vectors 2018 Dec 14;11(1):638 [FREE Full text] [doi: 10.1186/s13071-018-3232-7] [Medline: 30547815]
44. Pullan RL, Bethony JM, Geiger SM, Cundill B, Correa-Oliveira R, Quinnell RJ, et al. Human helminth co-infection: analysis of spatial patterns and risk factors in a Brazilian community. PLoS Negl Trop Dis 2008 Dec 23;2(12):e352 [FREE Full text] [doi: 10.1371/journal.pntd.0000352] [Medline: 19104658]
45. Anegagrie M, Lanfri S, Aramendia AA, Scavuzzo CM, Herrador Z, Benito A, et al. Environmental characteristics around the household and their association with hookworm infection in rural communities from Bahir Dar, Amhara Region, Ethiopia. PLoS Negl Trop Dis 2021 Jun 22;15(6):e0009466 [FREE Full text] [doi: 10.1371/journal.pntd.0009466] [Medline: 34157019]
46. de Silva NR, Brooker S, Hotez PJ, Montresor A, Engels D, Savioli L. Soil-transmitted helminth infections: updating the global picture. Trends Parasitol 2003 Dec;19(12):547-551. [doi: 10.1016/j.pt.2003.10.002] [Medline: 14642761]
47. Strunz EC, Addiss DG, Stocks ME, Ogden S, Utzinger J, Freeman MC. Water, sanitation, hygiene, and soil-transmitted helminth infection: a systematic review and meta-analysis. PLoS Med 2014 Mar;11(3):e1001620 [FREE Full text] [doi: 10.1371/journal.pmed.1001620] [Medline: 24667810]
48. Aw JYH, Clarke NE, Mayfield HJ, Lau CL, Richardson A, Vaz Nery S. Novel statistical approaches to identify risk factors for soil-transmitted helminth infection in Timor-Leste. Int J Parasitol 2021 Aug;51(9):729-739 [FREE Full text] [doi: 10.1016/j.ijpara.2021.01.005] [Medline: 33798561]
49. Campbell SJ, Nery SV, Wardell R, D'Este CA, Gray DJ, McCarthy JS, et al. Water, Sanitation and Hygiene (WASH) and environmental risk factors for soil-transmitted helminth intensity of infection in Timor-Leste, using real time PCR. PLoS Negl Trop Dis 2017 Mar 27;11(3):e0005393 [FREE Full text] [doi: 10.1371/journal.pntd.0005393] [Medline: 28346536]
50. Viney ME, Lok JB. The biology of Strongyloides spp. WormBook 2015 Jul 16:1-17 [FREE Full text] [doi: 10.1895/wormbook.1.141.2] [Medline: 26183912]
51. PAULA FM, COSTA-CRUZ JM. Epidemiological aspects of strongyloidiasis in Brazil. Parasitology 2011 Aug 03;138(11):1331-1340. [doi: 10.1017/s003118201100120x]
52. Chaiyos J, Suwannatrai K, Thinkhamrop K, Pratumchart K, Sereewong C, Tesana S, et al. MaxEnt modeling of soil-transmitted helminth infection distributions in Thailand. Parasitol Res 2018 Nov 17;117(11):3507-3517. [doi: 10.1007/s00436-018-6048-7] [Medline: 30120589]
53. Ruberanziza E, Owada K, Clark NJ, Umulisa I, Ortu G, Lancaster W, et al. Mapping soil-transmitted helminth parasite infection in Rwanda: estimating endemicity and identifying at-risk populations. Trop Med Infect Dis 2019 Jun 14;4(2):93 [FREE Full text] [doi: 10.3390/tropicalmed4020093] [Medline: 31207897]

54. Knopp S, Mohammed KA, Simba Khamis I, Mgeni AF, Stothard JR, Rollinson D, et al. Spatial distribution of soil-transmitted helminths, including Strongyloides stercoralis, among children in Zanzibar. Geospat Health 2008 Nov 01;3(1):47-56. [doi: 10.4081/gh.2008.231] [Medline: 19021108]
55. Riess H, Clowes P, Kroidl I, Kowuor DO, Nsojo A, Mangu C, et al. Hookworm infection and environmental factors in mbeya region, Tanzania: a cross-sectional, population-based study. PLoS Negl Trop Dis 2013 Sep 5;7(9):e2408 [FREE Full text] [doi: 10.1371/journal.pntd.0002408] [Medline: 24040430]
56. Mudenda NB, Malone JB, Kearney MT, Mischler PD, Nieto PDM, McCarroll JC, et al. Modelling the ecological niche of hookworm in Brazil based on climate. Geospat Health 2012 Sep 01;6(3):S111-S123. [doi: 10.4081/gh.2012.129] [Medline: 23032277]
57. Mabaso MLH, Appleton CC, Hughes JC, Gouws E. The effect of soil type and climate on hookworm (Necator americanus) distribution in KwaZulu-Natal, South Africa. Trop Med Int Health 2003 Aug 18;8(8):722-727 [FREE Full text] [doi: 10.1046/j.1365-3156.2003.01086.x] [Medline: 12869093]
58. Eslahi AV, Hashemipour S, Olfatifar M, Houshmand E, Hajialilo E, Mahmoudi R, et al. Global prevalence and epidemiology of Strongyloides stercoralis in dogs: a systematic review and meta-analysis. Parasit Vectors 2022 Jan 10;15(1):21 [FREE Full text] [doi: 10.1186/s13071-021-05135-0] [Medline: 35012614]
59. Preventive chemotherapy in human helminthiasis: coordinated use of anthelminthic drugs in control interventions: a manual for health professionals and programme managers. World Health Organization. 2006. URL: https://apps.who.int/iris/handle/10665/43545 [accessed 2022-04-26]

Abbreviations

BSI: Bare Soil Index
DEM: digital elevation model
ENDISI: Enhanced Normalized Difference Impervious Surfaces Index
EVI: Enhanced Vegetation Index
GIS: geographic information system
IP: intestinal parasite
KDE: kernel density estimation
NDVI: Normalized Difference Vegetation Index
NTD: neglected tropical disease
RS: remote sensing
SAVI: Soil-Adjusted Vegetation Index
STH: soil-transmitted helminth
TPI: Topographic Position Index
TWI: Topographic Wetness Index
VHI: Vegetation Heterogeneity Index
WHA: World Health Assembly

Edited by A Mavragani, T Sanchez; submitted 31.07.22; peer-reviewed by MM Patwary, MF Mohd Hisham, J Ellwanger; comments to author 06.07.23; revised version received 26.07.23; accepted 19.09.23; published 07.11.23.

Please cite as:
Candela E, Goizueta C, Sandon L, Muñoz-Antoli C, Periago MV
The Relationship Between Soil-Transmitted Helminth Infections and Environmental Factors in Puerto Iguazú, Argentina: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e41568
URL: https://publichealth.jmir.org/2023/1/e41568
doi:10.2196/41568
PMID:37934580

Original Paper

The Associations of Weekend Warrior Activity Patterns With the Visceral Adiposity Index in US Adults: Repeated Cross-sectional Study

Kai Wang[1], MSc; Fang Xia[1], MSc; Qingwen Li[1], MSc; Xin Luo[1], MSc; Jinyi Wu[1], MSc

[1]Department of Public Health, Wuhan Fourth Hospital, Wuhan, China

Corresponding Author:
Jinyi Wu, MSc
Department of Public Health
Wuhan Fourth Hospital
Hanzheng street #473
Wuhan, 430000
China
Phone: 86 18140520597
Email: wjypuai@outlook.com

Abstract

Background: According to previous reports, obesity especially visceral fat has become an important public health problem, causing an estimation of 20.5 disability-adjusted life years per 1000 inhabitants. Those who exercised for 1 or 2 days per week and reached the recommended 150 minutes of moderate physical activity (PA) per week have been defined as "weekend warriors" (WWs). Although the benefits of PA in suppressing obesity have been widely studied, the association of WWs with the Visceral Adiposity Index (VAI) and the difference between WW activity and regular PA are yet to be explored.

Objective: This study aims to explore the association between WW activity and other PA patterns with VAI in US adults.

Methods: The National Health and Nutrition Examination Survey 2007-2016 data set was used, and the analytic sample was limited to adults 20 years and older who had complete information about VAI, PA patterns, and other covariates, including demographic characteristics, behavioral factors, and disease conditions. Participants' characteristics in different PA pattern groups were tested using the Rao and Scott adjusted χ^2 test and ANOVA. Univariate and multivariate stepped linear regression models were then used to explore the association between the PA pattern and VAI. Finally, stratified analyses and interaction effects were conducted to investigate whether the association was stable among subgroups.

Results: The final sample included 9642 adults 20 years or older, which is representative of 158.1 million noninstitutionalized US adults, with 52.15% (n=5169) being male and 70.8% (n=4443) being non-Hispanic White. Gender, age group, race, education level, income level, marital status, smoking status, alcoholism, VAI, cardiovascular disease, and diabetes were all correlated with the PA pattern, but no relationship between hypertension and PA pattern was observed. After adjusting for demographic covariates, smoking status, alcoholism, cardiovascular disease, diabetes, and hypertension, WW and regularly active adults had a β of .307 (95% CI –0.611 to –0.003) and .354 (95% CI –0.467 to –0.241), respectively, for reduced VAI when compared with inactive adults, but no significant effect of lowering VAI (β=–.132, 95% CI –0.282 to 0.018) was observed in insufficiently active adults when compared with inactive adults. Besides, no significant difference was exhibited between WW adults and regularly active adults (β=.047, 95% CI –0.258 to 0.352), suggesting WW adults had the same benefit of decreasing VAI as regularly active adults. Stratified analyses results exhibited that WW activity was related to reduced VAI in female adults aged 20-44 years who were non-Hispanic Black, other, or multiracial; high school or General Educational Development education; and never married, and the association between PA pattern and VAI remained stable in all demographic subgroups.

Conclusions: Compared with inactive adults, WWs could reduce VAI, and there was no significant difference between WWs and regular active adults in decreasing VAI. Our study provides compelling evidence of the beneficial effect of WW activity on visceral obesity.

(JMIR Public Health Surveill 2023;9:e41973) doi:10.2196/41973

KEYWORDS

weekend warrior; Visceral Adiposity Index; NHANES; physical activity; obesity

Introduction

Physical activity (PA) has been proven to be beneficial by a number of studies, including for reducing chronic disease risk, improving mental state, and prolonging life [1-3]. The World Health Organization recommended that people aged 18 to 64 years should perform at least 150 minutes of moderate-intensity aerobic exercise per week, 75 minutes of high-intensity aerobic exercise per week, or an equivalent combination [4]. Randomized clinical trials have also shown that short-term periodic intermittent PA could improve cardiopulmonary health and other health outcomes, such as blood lipid levels and obesity [5].

In contrast to those who did 30 minutes of moderate-intensity PA for 5 days per week, those who did all exercise on 1 or 2 days per week were usually defined as “weekend warriors” (WWs) [6,7]. As society has become more fast paced worldwide, participating in sports frequently may be less suitable for a busy lifestyle, which has caused a gradual increase in the proportion of WWs. However, at present, the scientific community still knows little about WWs and whether their exercise schedule is more beneficial than inactivity or if WW activity had the same benefits as regularly active (RA) adults.

According to previous reports, obesity, especially visceral fat, has become an important public health problem, causing an estimation of 20.5 disability-adjusted life years per 1000 inhabitants. The economic impact of care for comorbidities associated with obesity could amount to US $2.158 billion. [8]. Obesity is related with various diseases, including cardiovascular disease (CVD), prodrome diabetes, type 2 diabetes mellitus (T2DM), hypertension, hyperlipidemia, sleep apnea, and some malignant tumors [9]. BMI as a routine weight measurement index has been widely used in scientific research [10,11]. However, BMI focuses on measuring overweight, which is not a reliable index to estimate body fat distribution. In contrast, the Visceral Adiposity Index (VAI) is a simple gender-specific indicator of visceral fat dysfunction, which can estimate body fat distribution. Moreover, VAI could be used as a risk predictor of functional impairment and disease, including for CVD, T2DM, nonalcoholic cirrhosis, or erectile dysfunction [12-18]. According to previous research, it is widely accepted that PA could reduce BMI and prevent obesity [10,19,20]. However, the association of WW activity with VAI and the difference between WW activity and regular PA remain to be explored.

Therefore, we obtained data from the National Health and Nutrition Examination Survey (NHANES) to conduct an analysis concerning WW activity and VAI. Through this research, we would like to make the public aware of the benefits of PA on VAI and make suggestions on how to reduce VAI to further reduce various metabolic disease risks. At the same time, the impact of different modes of PA on VAI is an interesting research direction. We also used large-scale data to analyze the different PA modes and VAI so as to provide guidance on the choice of PA mode.

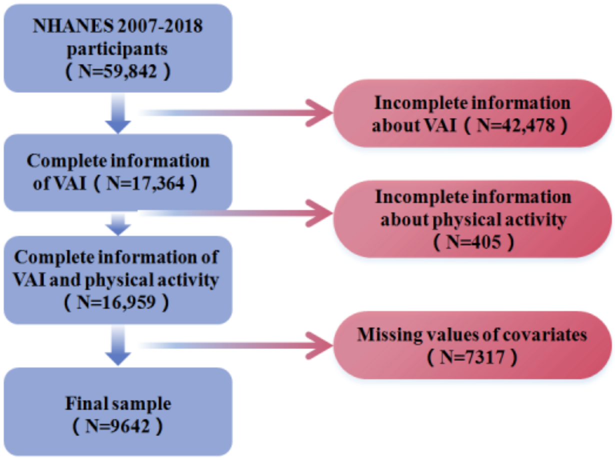
Methods

Study Population

Data was derived from the six continuous NHANES cycles from 2007-2008 to 2017-2018, which is a nationally representative population-based survey for assessing adult and child health and nutritional status in the United States [21]. The examination components consisted of medical, dental, and physiological measurements, and laboratory tests, which were supervised by trained medical personnel. Furthermore, the adoption of various modern equipment and compensation for the participants enabled the NHANES to collect reliable and high-quality data.

The total sample size of adults from the 2007 to 2018 cycle was 9642. Additional details on the study design, sampling, and exclusion criteria are described in Figure 1.

Figure 1. Flowchart of study design. NHANES: National Health and Nutrition Examination Survey; VAI: Visceral Adiposity Index.



Ethics Approval

All participants in NHANES provide informed consent and NHANES is approved by the research ethics review board of the Centers for Disease Control and Prevention [22].

Outcome Ascertainment

The VAI was measured with the following sex-specific formula: VAI = [waist circumference (WC) / 39.68 + (1.88 * BMI)] * (triglycerides / 1.03) * (1.31 / high-density lipoprotein cholesterol [HDL-C]) for males and VAI = [WC / 36.58 + (1.89

* BMI)] * (triglycerides / 0.81) * (1.52 / HDL-C) for females, in which triglycerides and HDL-C were expressed in mmol/L, WC expressed in cm, and BMI expressed in kg/m^2 [23]. The Cobas 6000 Chemistry Analyzer is optimized for workloads using a combination of photometric and ion-selective electrode determinations (c501 module), and electrochemiluminescence technology in the detection of triglyceride concentration [24]. HDL-C was measured with a magnesium/dextran sulfate solution that was first added to the specimen to form water-soluble complexes with non–HDL-C fractions [25].

PA and Other Covariates

PA in this study was assessed with a PA questionnaire in which participants were asked about the frequency and duration of vigorous and moderate sports, fitness, and recreational activities for at least 10 continuous minutes in a typical week. PA was calculated with the combination of frequency (times per week) and duration (duration per time), and total PA was calculated with the formula 2 * vigorous PA + moderate PA, since 1 minute of vigorous-intensity activity was defined as equivalent to 2 minutes of moderate-intensity activity according to PA guidelines [4]. Therefore, PA patterns were categorized into the following: inactive (no vigorous or moderate PA), insufficiently active (<150 minutes per week of total PA), WW (at least 150 minutes per week of total PA in 1 or 2 sessions), and RA (at least 150 minutes per week of total PA in more than 2 sessions).

Based on previous studies, covariates in this study included demographic data such as gender (male, female), age (20-44, 45-64, ≥65 years), race (non-Hispanic White, non-Hispanic Black, Mexican Americans, and other or multiracial races), education level (less than high school graduate, high school degree or equivalent, more than high school degree), income level (measured with as the ratio of family income to poverty [PIR]: low income PIR≤1.3, middle income 1.3<PIR<3.5, or high income PIR≥3.5), marital status (married or living with partner; divorced, separated, or widowed; never married). Behavioral risk factors included smoking status (never smoked, former smoker, current smoker) and alcoholism (yes, no). CVD was defined as a self-reported congestive heart failure, coronary heart disease, angina, myocardial infarction, or stroke diagnosed by a professional doctor. Diabetes was defined as a fasting plasma glucose ≥126 mg/dL, 2-hour plasma glucose ≥200 mg/dL, hemoglobin A_{1c} ≥6.5%, or self-reported diabetes diagnosed by a professional doctor. Hypertension was defined as an average systolic pressure ≥140 mm Hg and diastolic pressure ≥90 mm Hg in 3 tests [26-29].

Statistical Analysis

According to analytic guidelines published by the National Center for Health Statistics, stratum and primary sampling units were taken into account for the complex, multistage, probability sampling design. Since 6 consecutive cycles were derived from NHANES, the 2-year mobile examination center weight divided by 6 was adopted to be representative of the general population.

Initially, participants' characteristics in different groups were tested using the Rao and Scott [30] adjusted χ^2 test and ANOVA. Univariate and multivariate binary stepped logistic regression models were then used to explore the association between PA pattern and depression risk. Model 1 was nonadjusted, while model 2 was adjusted for demographic data. Model 3 was adjusted for demographic data, behavioral factors, and disease condition. Finally, a stratified analyses and interaction effect were conducted to investigate whether the association was stable among subgroups. Statistical analyses were performed using the Stata software (version 16.0, StataCorp LLC). All statistical tests were 2-sided, and significance was considered at α=.05.

Results

Characteristics of Study Participants

As described in Figure 1, 16,059 participants with complete information about VAI and PA pattern were enrolled in the study, and 1542, 2417, 10, 2464, and 884 participants were excluded because of incomplete information about income level, marital status, smoking status, alcoholism, and hypertension, respectively. The final sample included 9642 adults 20 years and older, which is representative for 158.1 million noninstitutionalized US adults, with 52.15% (n=5169) being male and 70.8% (n=4443) being non-Hispanic White.

The characteristics of participants in different PA pattern groups are presented in Table 1. Gender, age group, race, education level, income level, marital status, smoking status, alcoholism, VAI, CVD, and diabetes were all correlated with the PA pattern, but no relationship between hypertension and the PA pattern was observed. To be more specific, WW adults were more likely to be male, be aged 20-44 years, be non-Hispanic White, have some college or above, have a PIR≥3.5, be married or living with a partner, have never smoked, have lower VAI, be nonalcoholic, have CVD, and have diabetes.

Table 1. Participants characteristics according to physical activity pattern.

Characteristics	Overall (n=9642)	Inactive (n=4854)	Insufficiently active (n=1474)	Weekend warrior (n=190)	Regularly active (n=3124)	*P* value
Gender, n (%)						<.001
Male	5169 (52.15)	2543 (50.70)	704 (47.96)	151 (73.45)	1771 (54.69)	
Female	4473 (47.85)	2311 (49.30)	770 (52.04)	39 (26.55)	1353 (45.31)	
Age group (years), n (%)						<.001
20-44	4074 (44.82)	1694 (38.09)	625 (44.00)	146 (51.70)	4609 (51.70)	
45-64	3424 (37.78)	1877 (41.42)	536 (38.23)	40 (33.98)	971 (33.98)	
≥65	2144 (17.40)	1283 (20.49)	313 (17.77)	4 (0.02)	544 (14.32)	
Race, n (%)						<.001
Non-Hispanic White	4443 (70.80)	2192 (68.37)	704 (74.21)	56 (56.44)	1491 (73.00)	
Non-Hispanic Black	1877 (9.43)	979 (10.53)	278 (8.30)	49 (14.54)	571 (8.34)	
Mexican American	1398 (8.08)	786 (9.50)	180 (6.42)	38 (12.87)	394 (6.84)	
Other or multiracial	1924 (11.69)	897 (11.61)	312 (11.07)	47 (16.16)	668 (11.83)	
Education level, n (%)						<.001
Less than high school graduate	2119 (14.26)	1508 (22.15)	209 (9.41)	41 (17.34)	361 (6.65)	
High school graduate	2226 (22.49)	1248 (27.14)	329 (21.90)	45 (19.08)	604 (17.26)	
Some college or above	5297 (63.25)	2098 (50.71)	936 (68.69)	104 (63.58)	2159 (76.08)	
Income level, n (%)						<.001
PIR[a]≤1.3	2916 (20.30)	1818 (27.11)	345 (15.78)	60 (23.01)	693 (13.88)	
1.3<PIR<3.5	3660 (35.41)	1933 (38.94)	568 (35.15)	76 (40.50)	1083 (30.92)	
PIR≥3.5	3066 (44.29)	1103 (33.95)	561 (49.07)	54 (36.48)	1348 (55.20)	
Marital status, n (%)						<.001
Married or living with partner	5845 (64.72)	2949 (64.26)	912 (66.87)	104 (55.50)	1880 (64.80)	
Divorced, separated, or widowed	2068 (17.80)	1201 (21.02)	312 (16.70)	23 (12.00)	532 (14.69)	
Never married	1729 (17.48)	704 (14.73)	250 (16.43)	63 (32.51)	712 (20.51)	
Smoking status, n (%)						<.001
Never	4800 (51.12)	2168 (45.36)	779 (53.79)	113 (63.09)	1740 (56.30)	
Former	2649 (27.82)	1343 (27.18)	425 (29.20)	28 (14.75)	853 (28.67)	
Current	2193 (21.07)	1343 (27.46)	270 (17.01)	49 (22.16)	531 (15.03)	
Alcoholism, n (%)						<.001
No	7982 (83.85)	3858 (79.79)	1256 (85.23)	164 (87.88)	2704 (87.98)	
Yes	1660 (16.15)	996 (20.21)	218 (14.77)	26 (12.12)	420 (12.02)	
Cardiovascular disease, n (%)						<.001
No	8637 (91.85)	4199 (89.12)	1348 (93.72)	184 (98.70)	2906 (93.97)	
Yes	1005 (8.15)	655 (10.88)	126 (6.28)	6 (1.30)	218 (6.03)	
Diabetes, n (%)						<.001
No	7689 (85.04)	3615 (79.69)	1213 (87.48)	178 (94.75)	2683 (89.95)	
Yes	1953 (14.96)	1239 (20.31)	261 (12.52)	12 (5.25)	441 (10.05)	
Hypertension, n (%)						.17
No	9355 (97.44)	4697 (97.23)	1425 (96.72)	184 (98.12)	3049 (97.99)	

Characteristics	Overall (n=9642)	Inactive (n=4854)	Insufficiently active (n=1474)	Weekend warrior (n=190)	Regularly active (n=3124)	*P* value
Yes	287 (2.56)	157 (2.77)	49 (3.28)	6 (1.88)	75 (2.01)	
Visceral Adiposity Index, (SD)	1.98 (2.61)	2.24 (3.04)	1.95 (1.84)	1.71 (1.72)	1.68 (2.33)	<.001

[a] PIR: ratio of family income to poverty.

Relationship Between PA Pattern and VAI

The results of the binary univariate and multivariate logistic regression models of PA pattern on VAI were presented in Table 2 and Figure 2, and all models revealed a significantly lower VAI among WW and RA adults when compared with inactive adults. After adjusting for demographic covariates, smoking status, alcoholism, CVD, diabetes, and hypertension, WW and RA adults' VAI reduced by 0.307 (95% CI –0.611 to –0.003) and 0.354 (95% CI –0.467 to –0.241), respectively, but no significant effect was observed for VAI reduction (β=–.132, 95% CI –0.282 to 0.018) in insufficiently active adults. Moreover, no significant difference was exhibited between WW adults and RA adults, suggesting WW adults had the same benefit of decreasing VAI as RA adults.

Table 2. The relationship between physical activity pattern and Visceral Adiposity Index in adults.

Characteristics	Model 1[a]		Model 2[b]		Model 3[c]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Inactive	Reference	N/A[d]	Reference	N/A	Reference	N/A
Insufficiently active	–.291(–0.463 to –0.118)	.001	–.212 (–0.374 to –0.049)	.01	–.132 (–0.282 to 0.018)	.08
Weekend warrior	–.535 (–0.830 to –0.239)	.001	–.434 (–0.732 to –0.135)	.005	–.307 (–0.611 to –0.003)	.04
Regularly active	–.557 (–0.671 to –0.443)	<.001	–.441 (–0.556 to –0.326)	<.001	–.354 (–0.467 to –0.241)	<.001
Regularly active	Reference	N/A	Reference	N/A	Reference	N/A
Inactive	.557 (0.443 to 0.671)	<.001	.440 (0.326 to 0.556)	<.001	.354 (0.241 to 0.468)	<.001
Insufficiently active	.266 (0.106 to 0.426)	.001	.229 (0.063 to 0.395)	.007	.222 (0.060 to 0.385)	.008
Weekend warrior	.022 (–0.287 to 0.332)	.89	.007 (–0.294 to 0.309)	.96	.047 (–0.258 to 0.352)	.76

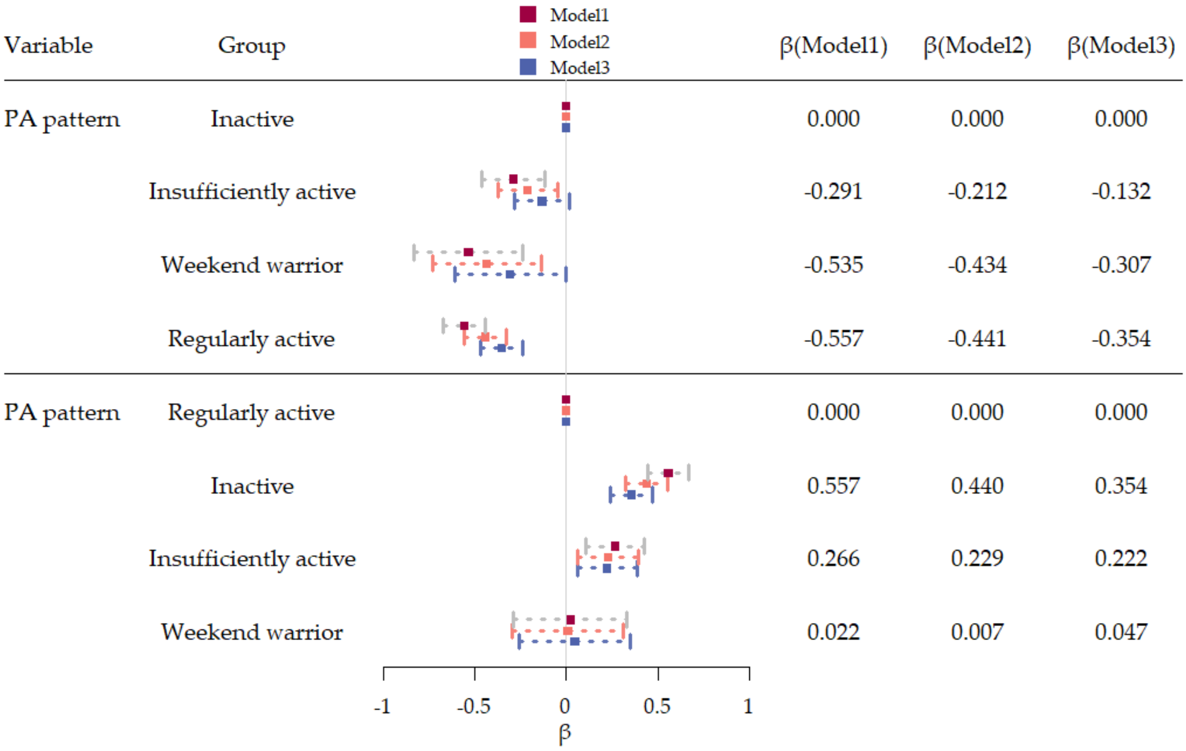
[a] Model 1 was the univariate model in which no covariates were adjusted.

[b] Model 2 was adjusted for demographic covariates, including gender, age group, race, education level, income level, and marital status.

[c] Model 3 was additionally adjusted for smoking status, alcoholism, cardiovascular disease, diabetes, and hypertension.

[d] N/A: not applicable.

Figure 2. The forest plot of PA pattern and the Visceral Adiposity Index in adults. PA: physical activity.



Variable	Group	β(Model1)	β(Model2)	β(Model3)
PA pattern	Inactive	0.000	0.000	0.000
	Insufficiently active	-0.291	-0.212	-0.132
	Weekend warrior	-0.535	-0.434	-0.307
	Regularly active	-0.557	-0.441	-0.354
PA pattern	Regularly active	0.000	0.000	0.000
	Inactive	0.557	0.440	0.354
	Insufficiently active	0.266	0.229	0.222
	Weekend warrior	0.022	0.007	0.047

Stratified Analyses According to Demographic Characteristics

To better understand the association between PA pattern and VAI in various subgroups, we conducted stratified analyses according to demographic characteristics. As described in Table 3, the WW activity pattern was related to a reduced VAI in female participants aged 20-44 years who were non-Hispanic Black, other, or multiracial, had a high school degree or General Educational Development, and were never married. The results of the interaction analyses revealed that all *P* values did not reach the significant level, indicating that the association between PA pattern and VAI remained stable in all demographic subgroups.

Table 3. The relationship between physically activity pattern and Visceral Adiposity Index in subgroups[a].

Characteristics	Inactive	Insufficiently active, β (95% CI)	Weekend warrior, β (95% CI)	Regularly active, β (95% CI)	*P* value for interaction
Gender					.17
Male	Reference	–.123 (–0.363 to 0.118)	–.229 (–0.636 to 0.177)	–.373 (–0.547 to –0.199)	
Female	Reference	–.152 (–0.334 to 0.030)	–0.621 (–0.902 to –0.339)	–.327 (–0.502 to –0.151)	
Age group (years)					.61
20-44	Reference	–.262 (–0.483 to –0.041)	–.388 (–0.776 to 0.000)	–.425 (–0.603 to –0.247)	
45-64	Reference	–.055 (–0.361 to 0.250)	–.314(–0.883 to 0.259)	–.269 (–0.558 to 0.021)	
≥65	Reference	–.053 (–0.345 to 0.240)	–.251 (–0.608 to 0.107)	–.375 (–0.554 to –0.195)	
Race					.28
Non-Hispanic White	Reference	–.148(–0.345 to 0.049)	–.270 (–0.744 to 0.204)	–.390 (–0.548 to –0.233)	
Non-Hispanic Black	Reference	–.136 (–0.316 to 0.045)	–.345 (–0.600 to –0.089)	–.236 (–0.447 to –0.026)	
Mexican American	Reference	–.414 (–0.744 to –0.083)	–.384 (–1.430 to 0.661)	–.198 (–0.578 to 0.183)	
Other or multiracial	Reference	.093 (–0.243 to 0.429)	–.493 (–0.928 to –0.059)	–.370 (–0.615 to –0.125)	
Education level					.99
Less than high school graduate	Reference	–.243 (–0.539 to 0.053)	–.154 (–0.722 to 0.414)	–.369 (–0.684 to –0.053)	
High school graduate or GED[b]	Reference	–.241 (–0.533 to 0.052)	–.446 (–0.805 to –0.086)	–.361 (–0.596 to –0.125)	
Some college or above	Reference	–.098 (–0.300 to 0.105)	–.271 (–0.727 to 0.184)	–.351 (–0.499 to –0.202)	
Income level					.58
PIR[c]≤1.3	Reference	–.293 (–0.541 to –0.044)	–.225 (–0.660 to 0.211)	–.414 (–0.651 to –0.177)	
1.3<PIR<3.5	Reference	–.111 (–0.388 to 0.166)	–.166 (–0.781 to 0.450)	–.220 (–0.474 to 0.034)	
PIR≥3.5	Reference	–.108 (–0.323 to 0.107)	–.392 (–0.822 to 0.037)	–.429 (–0.587 to –0.272)	
Marital status					.16
Married or living with partner	Reference	–.099 (–0.295 to 0.097)	–.188 (–0.722 to 0.347)	–.448 (–0.603 to –0.295)	
Divorced, separated, or widowed	Reference	–.255 (–0.541 to 0.032)	–.287 (–0.767 to 0.194)	–.145 (–0.478 to 0.188)	
Never married	Reference	–.121 (–0.474 to 0.231)	–.412 (–0.657 to –0.167)	–.168 (–0.437 to 0.101)	

[a] The models were adjusted for demographic characteristics, behavioral factors, and disease conditions.

[b] GED: General Educational Development.

[c] PIR: ratio of family income to poverty.

Discussion

With NHANES, we found that the WW exercise pattern in US adults is helpful to reduce VAI compared with no exercise. In particular, there was no difference between WWs and RA adults, suggesting that WW activity had the same VAI reduction effect as regular activity. Moreover, stratified analysis and interaction analysis confirmed the stability of the results. This might suggest for people who cannot exercise regularly that WW activity could reduce visceral obesity and improve health.

To the best of our knowledge, this is the first analysis to explore the impact of PA patterns, especially of WWs, on VAI. At present, most studies focus on the influence of the sedentary behaviors and leisure time PA (LTPA) ratio on disease or health status [31,32]. Through a cohort study of a nationally representative sample of US cancer survivors, it was found that the combination of sedentary and physical inactivity was associated with the high risk of death [2]. Including the mode of PA adds to the analysis of the LTPA ratio and concerns not only the duration but also the frequency of exercise. Therefore, the NHANES database was used in this study to determine the effects of WW and other PA patterns on visceral fat distribution with VAI.

We took inactivity as the reference. In model 1, no covariates were considered, and insufficiently active, WW, and RA adults had reduced VAI. In model 2, after adjusting for demographic data, the results were similar to the univariate analysis. In model 3, WW and RA adults had reduced VAI after adjusting for demographic data, behavioral factors, and disease conditions, while no significant difference was observed in insufficiently active adults and inactive adults. In detail, RA adults had the largest negative correlation with VAI (β=–.35; *P*<.001), and

WW adults also had an obvious negative correlation (β=–.31; *P*=.047). Eekelen et al [33] found that moderate-to-vigorous PA (MVPA) was associated with less body fat, visceral fat, and liver fat. Mild PA seems to be associated with less body fat but not visceral fat or liver fat. This suggests that exercise would preferentially reduce visceral fat compared to caloric restriction, which might be because visceral fat is more metabolically active and sensitive to lipolytic activation in the adrenal system. Similarly, inactive and insufficiently active adults had elevated VAI when compared with RA adults. Nevertheless, no significant difference was observed in WW and RA adults, indicating the importance of PA duration rather than frequency.

Moreover, the linear regression showed other influencing factors of VAI (see Multimedia Appendix 1 for details). Using no exercise or regular exercise as a reference, age, race, income, marriage, smoking, and diabetes all had an impact on VAI. In the age group, there was no difference in VAI between those aged 20-44 years and those aged 45-64 years, but adults 65 years or older were significantly negatively correlated with VAI, which might be caused by emaciation in older adults. Cameron et al [11] reported the interaction between age and PA. They found that a decrease in BMI and body fat percentage was greater in the older adult group than in the younger group, which was consistent with our results.

Our results showed that non-Hispanic Black participants were more likely to reduce VAI than non-Hispanic White participants. Some studies found that, after correcting for total fat mass, Black women had less visceral adipose tissue than White women, while White people had more visceral adipose tissue than African American people [10]. African American people were more likely to reduce visceral adipose tissue, which supported our results.

There was no difference between middle-income and low-income groups on VAI, while the VAI of the high-income group was significantly lower than that of the low-income group. Ameye and Swinnen [34] reported that obesity varies with income but in a nonlinear way. Overall, obesity in low-income countries increased with greater income but was not related with income in middle-income countries, while it decreases with greater income in high-income countries. Since our data set was from the United States, it was in line with Ameye and Swinnen [34], and it may be related to social stigma around overweight and good medical resources in high-income countries.

Unmarried people were more likely to reduce VAI than married people. According to Lee et al [35], the prevalence of abdominal obesity in married participants was higher than those with other marriage statuses, which was consistent with our results. This marriage-related difference might be due to the fact that marriage increased the frequency of meals and snacks, thus increasing total energy consumption. Compared with married people, unmarried people were more eager to lose weight and maintain weight to make themselves more attractive for future marriage.

Smokers might have higher VAI than nonsmokers. Wehby et al [36] used a sample of 1057 mothers from Norway and found a heterogeneous effect of smoking on BMI. With an increase in smoking, BMI would increase at low/medium BMI levels, while BMI would decrease at high BMI levels. They observed genetic effects on the relationship between smoking and increased BMI, and gene analysis (eg, *CHRNA3*) had consistent evidence that increased BMI was related with smoking, which may explain our results.

Participants who were diabetic could increase VAI more than those without diabetes. According to Haslam's [37] review, experimental results showed that after 6 months of overfeeding, the BMI of young men without a family history of diabetes would increase to 28.0 kg/m^2, and the levels of fasting plasma insulin, glucose, and triglyceride would also increase reversibly, thereby impairing glucose tolerance. About 90% of patients with type 2 diabetes had a BMI greater than 23.0 kg/m^2. Patients who are diabetic might take drugs that cause weight gain, while vulnerable individuals who are already obese may take drugs that cause hypoglycemia.

VAI has been proven by many studies to be a predictor of cardiovascular and metabolic diseases such as hypertension, T2DM, CVD, and nonalcoholic fatty liver disease. Therefore, we planned to conduct a structural equation model in the next study to analyze the mediation effect of VAI: whether the activity pattern might affect VAI and related diseases. This future research direction would improve the relationship between exercise, diet, and visceral obesity, and further specify the analysis of activity and diet patterns.

There were several advantages in this study. First, we adopted a large-scale data set including 9642 participants, thus ensuring representativeness. Second, the topic of activity and obesity was specified in this study, thus the association between visceral obesity and different exercise modes. Finally, this study used a hierarchical linear regression method to analyze the covariates of VAI in detail. However, there were still some limitations. First, this study was based on a cross-sectional investigation, which could not infer the causal relationship between activity patterns and VAI. Second, there are many factors related to VAI, and the variables included in this study are limited. Third, since this study focused on the activity patterns, it did not further explore the relationship between exercise intensity, exercise/sedentary ratio, and VAI, which could be explored in the next study.

In conclusion, although the benefits of physical activities were well known, many people had limited time to engage in activities, and MVPA only accounted for 5% of the total time for daytime activities. Therefore, the role of WW activity is worth researching. According to our results, we found that, compared with no activity, WWs could reduce VAI, and there was no difference between WWs and RA adults, indicating that WW activity is worth promoting for people who cannot exercise regularly. This research also suggested that the duration of PA was more important than the frequency of activity. With this study, we want to make the public aware of the benefits about WW PA and encourage those who are tired from work but want to keep healthy to exercise during the weekend to reduce VAI.

Data Availability

The data sets were derived from NHANES 2007-2008 to 2017-2018 [38], and only publicly available data was used in this study. The data sets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Other risk factors of Visceral Adiposity Index.
[DOCX File , 15 KB - publichealth_v9i1e41973_app1.docx]

References

1. Ando S, Koyama T, Kuriyama N, Ozaki E, Uehara R. The association of daily physical activity behaviors with visceral fat. Obes Res Clin Pract 2020;14(6):531-535. [doi: 10.1016/j.orcp.2020.10.004] [Medline: 33168484]
2. Whitaker KM, Pereira MA, Jacobs DR, Sidney S, Odegaard AO. Sedentary behavior, physical activity, and abdominal adipose tissue deposition. Med Sci Sports Exerc 2017 Mar;49(3):450-458 [FREE Full text] [doi: 10.1249/MSS.0000000000001112] [Medline: 27749387]
3. Prince S, Kristjansson E, Russell K, Billette J, Sawada M, Ali A, et al. Relationships between neighborhoods, physical activity, and obesity: a multilevel analysis of a large Canadian city. Obesity (Silver Spring) 2012 Oct;20(10):2093-2100 [FREE Full text] [doi: 10.1038/oby.2011.392] [Medline: 22262164]
4. Piercy K, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. JAMA 2018 Nov 20;320(19):2020-2028 [FREE Full text] [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
5. Courtemanche C, Tchernis R, Ukert B. The effect of smoking on obesity: evidence from a randomized trial. J Health Econ 2018 Jan;57:31-44. [doi: 10.1016/j.jhealeco.2017.10.006] [Medline: 29179027]
6. O'Donovan G, Sarmiento OL, Hamer M. The rise of the "Weekend Warrior". J Orthop Sports Phys Ther 2018 Aug;48(8):604-606. [doi: 10.2519/jospt.2018.0611] [Medline: 30064334]
7. Gubelmann C, Vollenweider P, Marques-Vidal P. Of weekend warriors and couch potatoes: socio-economic determinants of physical activity in Swiss middle-aged adults. Prev Med 2017 Dec;105:350-355 [FREE Full text] [doi: 10.1016/j.ypmed.2017.10.016] [Medline: 28988999]
8. Gil-Rojas Y, Garzón A, Hernández F, Pacheco B, González D, Campos J, et al. Burden of disease attributable to obesity and overweight in colombia. Value Health Reg Issues 2019 Dec;20:66-72. [doi: 10.1016/j.vhri.2019.02.001] [Medline: 31035116]
9. Conway B, Rene A. Obesity as a disease: no lightweight matter. Obes Rev 2004 Aug;5(3):145-151. [doi: 10.1111/j.1467-789X.2004.00144.x] [Medline: 15245383]
10. Camhi S, Bray G, Bouchard C, Greenway F, Johnson W, Newton R, et al. The relationship of waist circumference and BMI to visceral, subcutaneous, and total body fat: sex and race differences. Obesity (Silver Spring) 2011 Feb;19(2):402-408 [FREE Full text] [doi: 10.1038/oby.2010.248] [Medline: 20948514]
11. Cameron N, Godino J, Nichols JF, Wing D, Hill L, Patrick K. Associations between physical activity and BMI, body fatness, and visceral adiposity in overweight or obese Latino and non-Latino adults. Int J Obes (Lond) 2017 Jun;41(6):873-877 [FREE Full text] [doi: 10.1038/ijo.2017.49] [Medline: 28220040]
12. Xu C, Ma Z, Wang Y, Liu X, Tao L, Zheng D, et al. Visceral adiposity index as a predictor of NAFLD: a prospective study with 4-year follow-up. Liver Int 2018 Dec;38(12):2294-2300. [doi: 10.1111/liv.13941] [Medline: 30099825]
13. Nusrianto R, Tahapary DL, Soewondo P. Visceral adiposity index as a predictor for type 2 diabetes mellitus in Asian population: a systematic review. Diabetes Metab Syndr 2019;13(2):1231-1235. [doi: 10.1016/j.dsx.2019.01.056] [Medline: 31336469]
14. Hudzik B, Nowak J, Szkodziński J, Zubelewicz-Szkodzińska B. Visceral adiposity in relation to body adiposity and nutritional status in elderly patients with stable coronary artery disease. Nutrients 2021 Jul 09;13(7):2351 [FREE Full text] [doi: 10.3390/nu13072351] [Medline: 34371863]
15. Keum N, Lee D, Kim R, Greenwood D, Giovannucci E. Visceral adiposity and colorectal adenomas: dose-response meta-analysis of observational studies. Ann Oncol 2015 Jun;26(6):1101-1109 [FREE Full text] [doi: 10.1093/annonc/mdu563] [Medline: 25480876]
16. Fujiwara N, Nakagawa H, Kudo Y, Tateishi R, Taguri M, Watadani T, et al. Sarcopenia, intramuscular fat deposition, and visceral adiposity independently predict the outcomes of hepatocellular carcinoma. J Hepatol 2015 Jul;63(1):131-140 [FREE Full text] [doi: 10.1016/j.jhep.2015.02.031] [Medline: 25724366]
17. Bolat MS, Kocamanoglu F, Ozbek ML, Buyukalpelli R, Asci R. Can high Visceral Adiposity Index be a risk factor for sexual dysfunction in sexually active men? J Sex Med 2020 Oct;17(10):1926-1933. [doi: 10.1016/j.jsxm.2020.06.014] [Medline: 32712095]

18. Bagyura Z, Kiss L, Lux Á, Csobay-Novák C, Jermendy Á, Polgár L, et al. Association between coronary atherosclerosis and visceral adiposity index. Nutr Metab Cardiovasc Dis 2020 May 07;30(5):796-803 [FREE Full text] [doi: 10.1016/j.numecd.2020.01.013] [Medline: 32127334]
19. Ross R, Soni S, Houle S. Negative energy balance induced by exercise or diet: effects on visceral adipose tissue and liver fat. Nutrients 2020 Mar 25;12(4):891 [FREE Full text] [doi: 10.3390/nu12040891] [Medline: 32218121]
20. Nigatu Haregu T, Khayeka-Wandabwa C, Ngomi N, Oti S, Egondi T, Kyobutungi C. Analysis of patterns of physical activity and sedentary behavior in an urban slum setting in Nairobi, Kenya. J Phys Act Health 2016 Aug;13(8):830-837. [doi: 10.1123/jpah.2015-0510] [Medline: 26998581]
21. Xia F, Li Q, Luo X, Wu J. Association between urinary metals and leukocyte telomere length involving an artificial neural network prediction: Findings based on NHANES 1999-2002. Front Public Health 2022;10:963138 [FREE Full text] [doi: 10.3389/fpubh.2022.963138] [Medline: 36172207]
22. NCHS Ethics Review Board (ERB) approval. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhanes/irba98.htm#print [accessed 2022-12-28]
23. Amato M, Giordano C, Galia M, Criscimanna A, Vitabile S, Midiri M, AlkaMeSy Study Group. Visceral Adiposity Index: a reliable indicator of visceral fat function associated with cardiometabolic risk. Diabetes Care 2010 Apr;33(4):920-922 [FREE Full text] [doi: 10.2337/dc09-1825] [Medline: 20067971]
24. Triglycerides laboratory procedure manual. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/data/nhanes/2017-2018/labmethods/TRIGLY-J-MET-508.pdf [accessed 2022-12-28]
25. Direct HDL laboratory procedure manual. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/data/nhanes/2017-2018/labmethods/HDL-J-MET-508.pdf [accessed 2022-12-28]
26. Montgomery J, Lu J, Ratliff S, Mezuk B. Food insecurity and depression among adults with diabetes: results from the National Health and Nutrition Examination Survey (NHANES). Diabetes Educ 2017 Jun;43(3):260-271 [FREE Full text] [doi: 10.1177/0145721717699890] [Medline: 28436293]
27. Gay IC, Tran DT, Paquette DW. Alcohol intake and periodontitis in adults aged ≥30 years: NHANES 2009-2012. J Periodontol 2018 Jun;89(6):625-634. [doi: 10.1002/JPER.17-0276] [Medline: 29572839]
28. ALHarthi SS, Natto ZS, Midle JB, Gyurko R, O'Neill R, Steffensen B. Association between time since quitting smoking and periodontitis in former smokers in the National Health and Nutrition Examination Surveys (NHANES) 2009 to 2012. J Periodontol 2019 Jan;90(1):16-25. [doi: 10.1002/JPER.18-0183] [Medline: 30102767]
29. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988-2012. JAMA 2015 Sep 08;314(10):1021-1029. [doi: 10.1001/jama.2015.10029] [Medline: 26348752]
30. Rao JNK, Scott AJ. On simple adjustments to chi-square tests with sample survey data. Ann Stat 1987 Mar 1;15(1):260-271. [doi: 10.1214/aos/1176350273]
31. Hobbs M, Griffiths C, Green M, Jordan H, McKenna J. P56 Associations between the physical activity and food environment and obesity: a cross sectional study of UK adults. J Epidemiol Community Health 2016 Sep 13;70(Suppl 1):A78.2-A7A78. [doi: 10.1136/jech-2016-208064.155]
32. Edimo Dikobo SJ, Lemieux I, Poirier P, Després JP, Alméras N. Leisure-time physical activity is more strongly associated with cardiometabolic risk than occupational physical activity: Results from a workplace lifestyle modification program. Prog Cardiovasc Dis 2022 Dec 21 [FREE Full text] [doi: 10.1016/j.pcad.2022.12.005] [Medline: 36565734]
33. Winters-VAN Eekelen E, VAN DER Velde JHPM, Boone SC, Westgate K, Brage S, Lamb HJ, et al. Objectively measured physical activity and body fatness: associations with total body fat, visceral fat, and liver fat. Med Sci Sports Exerc 2021 Nov 01;53(11):2309-2317 [FREE Full text] [doi: 10.1249/MSS.0000000000002712] [Medline: 34081059]
34. Ameye H, Swinnen J. Obesity, income and gender: The changing global relationship. Global Food Secur 2019 Dec;23:267-281. [doi: 10.1016/j.gfs.2019.09.003]
35. Lee J, Shin A, Cho S, Choi J, Kang D, Lee J. Marital status and the prevalence of obesity in a Korean population. Obes Res Clin Pract 2020;14(3):217-224. [doi: 10.1016/j.orcp.2020.04.003] [Medline: 32418738]
36. Wehby GL, Murray JC, Wilcox A, Lie RT. Smoking and body weight: evidence using genetic instruments. Econ Hum Biol 2012 Mar;10(2):113-126 [FREE Full text] [doi: 10.1016/j.ehb.2011.09.002] [Medline: 22024417]
37. Haslam D. Obesity and diabetes: the links and common approaches. Prim Care Diabetes 2010 Jul;4(2):105-112. [doi: 10.1016/j.pcd.2010.04.002] [Medline: 20447884]
38. NHANES 2017-2018. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 [accessed 2022-12-28]

Abbreviations

CVD: cardiovascular disease
HDL-C: high-density lipoprotein cholesterol
LTPA: leisure time physical activity
MVPA: moderate-to-vigorous physical activity
NHANES: National Health and Nutrition Examination Survey

PA: physical activity
PIR: ratio of family income to poverty
RA: regularly active
T2DM: type 2 diabetes mellitus
VAI: Visceral Adiposity Index
WC: waist circumference
WW: weekend warrior

Edited by G Eysenbach; submitted 16.08.22; peer-reviewed by Q Li, D Wang; comments to author 26.10.22; revised version received 13.11.22; accepted 04.12.22; published 11.01.23.

Please cite as:
Wang K, Xia F, Li Q, Luo X, Wu J
The Associations of Weekend Warrior Activity Patterns With the Visceral Adiposity Index in US Adults: Repeated Cross-sectional Study
JMIR Public Health Surveill 2023;9:e41973
URL: https://publichealth.jmir.org/2023/1/e41973
doi:10.2196/41973
PMID:36630179

Original Paper

Changes in Online Food Access During the COVID-19 Pandemic and Associations With Deprivation: Longitudinal Analysis

Matthew Keeble[1], MSc; Jean Adams[1], PhD; Thomas Burgoine[1], PhD

[1]MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom

Corresponding Author:
Matthew Keeble, MSc
MRC Epidemiology Unit, Institute of Metabolic Science
University of Cambridge School of Clinical Medicine
Level 3, Institute of Metabolic Science
Cambridge, CB2 0SL
United Kingdom
Phone: 44 1223 330315
Email: matthew.keeble@mrc-epid.cam.ac.uk

Abstract

Background: Food prepared out of home is typically energy-dense and nutrient-poor. Online food delivery services have become a popular way to purchase such food. The number of accessible food outlets through these services can influence how frequently they are used. Anecdotally, food outlet access through online food delivery services increased in England between 2020 and 2022, in the context of the COVID-19 pandemic. However, the extent to which this access changed is poorly understood.

Objective: We aimed to investigate monthly changes in online access to food prepared out of home in England in the context of the first 2 years of the COVID-19 pandemic compared with November 2019 and the extent to which any changes were associated with deprivation.

Methods: In November 2019 and monthly between June 2020 and March 2022, we used automated data collection to construct a data set containing information about all food outlets in England registered to accept orders through the leading online food delivery service. Across postcode districts, we identified the number and percentage of food outlets registered to accept orders and the number that was accessible. We used generalized estimating equations (adjusted for population density, the number of food outlets in the physical food environment, and rural/urban classification) to investigate the change in outcomes compared with prepandemic levels (November 2019). We stratified analyses by deprivation quintile (Q).

Results: Across England, the summed number of food outlets registered to accept orders online increased from 29,232 in November 2019 to 49,752 in March 2022. Across postcode districts, the median percentage of food outlets registered to accept orders online increased from 14.3 (IQR 3.8-26.0) in November 2019 to 24.0 (IQR 6.2-43.5) in March 2022. The median number of food outlets accessible online decreased from 63.5 (IQR 16.0-156.0) in November 2019 to 57.0 (IQR 11.0-163.0) in March 2022. However, we observed variation by deprivation. In March 2022, the median number of outlets accessible online was 175.0 (IQR 104.0-292.0) in the most deprived areas (Q5) compared with 27.0 (IQR 8.5-60.5) in the least deprived (Q1). In adjusted analyses, we estimated that the number of outlets accessible online in the most deprived areas was 10% higher in March 2022 compared with November 2019 (incidence rate ratios: 1.10, 95% CI 1.07-1.13). In the least deprived areas, we estimated a 19% decrease (incidence rate ratios: 0.81, 95% CI 0.79-0.83).

Conclusions: The number of food outlets accessible online increased only in the most deprived areas in England. Future research might attempt to understand the extent to which changes in online food access were associated with changes in online food delivery service use and the possible implications on diet quality and health.

(JMIR Public Health Surveill 2023;9:e41822) doi:10.2196/41822

KEYWORDS

COVID-19; digital food environment; fast foods; online food delivery services; public health

Introduction

Purchasing and consuming food prepared out of home has become increasingly popular in many countries [1]. This food accounted for over 50% of total food expenditure in the United States in 2018 [2], and more than one-quarter of total food and beverage expenditure in the United Kingdom between 2015 and 2017 [3]. Food prepared out of home is often energy-dense [4,5], with the majority of items that are served by large chain restaurants exceeding recommended levels for sodium, fat, saturated fat, or sugars [6]. In turn, more frequent consumption of food prepared out of home has been positively associated with bodyweight [7]. Online food delivery services are now an established way of purchasing such food [8]. It is plausible, therefore, that the use of these services has implications for dietary practices, obesity, and health [9].

When using online food delivery services (such as UberEats), customers receive aggregated information about all food outlets that will deliver to them based on their location. Customers then select a food outlet and place and pay for their order on a single platform. Orders are forwarded to food outlets where meals are prepared before being delivered by couriers working for them or the online food delivery service [9]. As in the physical food environment [10], there is evidence that the number of food outlets accessible to individuals online (ie, the number that will deliver to them) can influence online food delivery service use. Among adults living in the United Kingdom, only those with the highest number of food outlets accessible online (between 182 and 879 outlets) self-reported any online food delivery service use in the previous week [11]. Even before the COVID-19 pandemic, the potential for this food-purchasing practice was not equally distributed across England, where the adjusted mean number of food outlets accessible online in 2019 was over 100 in the most deprived areas, compared with 70 in the least deprived areas [12]. This difference could contribute to known socioeconomic inequalities in diet-related health [13].

Government responses to the COVID-19 pandemic with respect to out of home food retail varied across countries worldwide. The Supplemental Nutrition Assistance Program was extended to include digital food retail in the United States [14], and the rules regarding the delivery of food prepared out of home and alcohol were relaxed in Australia [15]. In the United Kingdom, emergency regulations that allowed bars, cafés, pubs, and restaurants to offer a takeaway food service without specific permission were in place between March 2020 and March 2022 [16-18]. Further details are shown in Table 1.

Table 1. Details of selected emergency regulations introduced, action taken, and broader events in England in relation to the COVID-19 pandemic and food outlets selling food prepared out of home.

Time	Emergency regulations introduced, action taken, and broader events
Before March 2020	• Food outlets in the physical food environment operated with a *primary use* based on their core business operations. The *primary use* of hot food takeaway outlets would be to serve food prepared out of home for off-premises consumption after previously receiving Local Authority[a] planning permission to operate. • Other establishments (bars, cafés, pubs, and restaurants) could offer a hot food takeaway service *in addition* to their primary use but would need Local Authority planning permission to do so. • All food outlets with a customer-facing premises in the physical food environment could feasibly register to accept orders online regardless of their primary use without Local Authority planning permission[b].
March 2020	• Bars, cafés, pubs, and restaurants were forced to close for on-premises food consumption as part of *first* national lockdown[c]. • Emergency regulations allowed bars, cafés, pubs, and restaurants to offer a hot food takeaway service in addition to their primary use introduced.
July 2020	• Bars, cafés, pubs, and restaurants were allowed to reopen for on-premises food consumption but only with table service and with restricted capacity.
August 2020	• The *Eat Out to Help Out* scheme that offered a 50% discount on meals, up to £10 (US $12) per person, every Monday, Tuesday, and Wednesday between August 3 and 31 was introduced for on-premises food consumption.
November 2020	• Bars, cafés, pubs, and restaurants were forced to close for on-premises food consumption as part of *second* national lockdownc. • Emergency regulations introduced in March 2020 were extended until March 2022.
December 2020	• Bars, cafés, pubs, and restaurants were allowed to reopen for on-premises food consumption but only with table service and with restricted capacity.
January 2021	• Bars, cafés, pubs, and restaurants were forced to close for on-premises food consumption as part of *third* national lockdownc.
April 2021	• Bars, cafés, pubs, and restaurants were allowed to reopen for on-premises food consumption but only with table service and with restricted capacity.
July 2021	• Restrictions necessitating table service only and capacity limits for on-premises food consumption inside bars, cafés, pubs, and restaurants ended.
March 2022	• Emergency regulations introduced in March 2020 ended.
After March 2022	• Bars, cafés, pubs, and restaurants *should* revert to their primary use and *should* stop offering a hot food takeaway service if adopted as part of emergency regulations. To continue offering an additional hot food takeaway service, Local Authority planning permission would be required. • All food outlets with customer-facing premises in the physical food environment can feasibly register to accept orders online regardless of their primary use without Local Authority planning permission.

[a] Local Authorities are administrative bodies that operate at a subnational and subregional level.

[b] Registration to accept orders through online food delivery services remained viable regardless of any subsequent action or change.

[c] As part of lockdown orders in England, individuals were instructed to remain *at home*.

With regard to the link between the information outlined in Table 1 and online food delivery services, food sold through these services is typically prepared in the kitchens of food outlets that are customer facing in the physical food environment [9]. The location of these premises is socioeconomically distributed in England, with higher numbers and concentrations in more deprived areas [19]. As such, in more deprived areas, it is possible that a greater number of existing food outlets chose to operate within emergency regulations introduced during the COVID-19 pandemic and registered to accept orders online. In turn, this could influence absolute and relative levels of online food accessibility and aforementioned inequalities.

First, we aimed to investigate changes in levels of online access to food prepared out of home in England during the first 2 years of the COVID-19 pandemic. Second, we aimed to identify the extent to which any changes were associated with area-level deprivation.

Methods

In this study, we built on our previous research that investigated cross-sectional associations between deprivation and online food outlet access in England [12]. Given that we have previously reported our methods in full, in this study we provide an overview.

Study Setting, Period, and Analytic Scale

The setting for our study was England, and the study period was from November 2019 to March 2022, which coincided with the end of emergency regulations introduced during the

COVID-19 pandemic (Table 1). Data were not collected between December 2019 and May 2020, meaning that first available data after November 2019 were from June 2020.

We completed analyses at the postcode district level because this reflects how food outlets registered to accept orders online delineated where they would deliver to in November 2019. Postcode districts are contained in the first half of full postcodes. For example, for the *postcode* CB2 0QQ, the *postcode district* is CB2. Postcode districts have an average size of 33 square miles [20] and a median population of 23,610 (IQR 13,320-34,560) [21].

We used postcode district boundary data from 2012, sourced from the United Kingdom Data Service [22], to map postcode districts in a Geographic Information System (GIS; ArcGIS version 10.7.1; ESRI Inc). We considered 2118 postcode districts eligible for inclusion in our analyses, reflecting those with boundaries entirely within or intersecting the border of England.

Exposure Measure

We modeled our exposure measure (time) based on the frequency of data collection for our outcomes (monthly). We investigated changes in each of our outcomes over time compared with baseline (November 2019).

Outcome Measures

Data Source and Collection

For each outcome, we collected data from an online food delivery service that was considered to be the UK market leader, Just Eat. In 2019, approximately 30,000 food outlets in the United Kingdom were registered to accept orders through this service, and there were approximately 170 million orders placed by customers [23,24]. In pilot analyses conducted in 2020, only a minority of food outlets registered to accept orders through Deliveroo (the next largest online food delivery service platform in the United Kingdom) were not also registered to accept orders through Just Eat [12]. Therefore, we collected data from Just Eat as a representation of online food delivery services. We refer to our data source as “the online food delivery service” hereafter.

In November 2019 and then monthly between June 2020 and March 2022, we used a web-browser extension to identify and collect the postcode of all food outlets in England that were registered to accept orders through the online food delivery service [25].

Number of Food Outlets Registered to Accept Orders Online

To accept orders through the online food delivery service, food outlets must register with them. In doing so, food outlets will have information about their business displayed on the online food delivery service platform irrespective of the location where they would deliver. We used Doogal (a free-to-use web service) to geocode the postcode of each food outlet identified during data collection [26] and excluded those that could not be geocoded (monthly range 0.08%-1.40%). We then mapped the locations of food outlets in our GIS based on supplied coordinates and counted the number located in each postcode district boundary.

Number of Food Outlets Accessible Online

In November 2019, June 2020, and July 2020, food outlets registered to accept orders online published the postcode districts that they would deliver to as their delivery areas. After July 2020, the online food delivery service no longer published this information. Therefore, between August 2020 and March 2022, after we searched the online food delivery service website by a given postcode district, we identified unique food outlets that appeared in our search results. We counted this number to determine those that would be accessible for a given population living in each postcode district (this is the number of food outlets that could be ordered from). Although we used 2 approaches, resultant data used in the outcome construction were the same.

Percentage of Food Outlets Registered to Accept Orders Online

We compared the number of food outlets registered to accept orders online with the number of outlets located in the physical food environment of the same postcode district. In doing so, we estimated the percentage of food outlets registered to accept orders online.

For the denominator (the number of outlets located in the physical food environment of a postcode district), we used the Ordnance Survey Points of Interest (OS POI) data set. These are commercial data containing information about retailers across multiple sectors, collated from more than 170 suppliers [27]. We extracted information for the following food outlet classifications: “Fast food and takeaway outlets” (food outlets selling food for off-premises consumption); “Fast food delivery services” (food outlets selling food for delivery, not explicitly online); “Fish and Chip shops” (food outlets selling a traditional British cuisine, typically for off-premises consumption); “Restaurants” (food outlets selling food for on-premises consumption); “Pubs, Bars, Inns” (establishments that primarily serve alcohol, that can also sell food for on-premises consumption); and “Cafe, Snack Bars & Tea Rooms” (food outlets selling food with no distinguishable consumption location). We selected these classifications based on a priori knowledge that they included food outlets typically registered to accept orders online and to reflect emergency regulations introduced during the COVID-19 pandemic (Table 1). We used coordinates supplied with OS POI data to map the locations of food outlets in our GIS. These coordinates have a stated accuracy of 1 meter [27].

We matched monthly data from the online food delivery service with OS POI data collected quarterly (Multimedia Appendix 1). We did not match individual food outlets listed in each data set, meaning that this outcome is the number of food outlets registered to accept orders online (based on data from the online food delivery service) as a percentage of the number of food outlets in the physical food environment (based on OS POI data set). Although we report a percentage, we acknowledge that in the strictest sense, we did not calculate it as such. We bounded this measure between 0% and 100% because the number of food outlets registered to accept orders online should not exceed the number of food outlets in the physical food environment. When the percentage exceeded 100%, it represented that a retailer not classified as a food outlet in OS POI data (according to our included classifications) was registered to accept orders online. We excluded postcode districts when this occurred (n=3).

Covariates

Food sold through online food delivery services is typically prepared in the kitchens of food outlets located in the physical food environment. Therefore, online food outlet access might be a function of physical food outlet access. We used OS POI data to account for the number of food outlets in the physical food environment when we did not use it as the denominator (ie, for the percentage of food outlets registered to accept orders online).

We used data from the 2019 Index of Multiple Deprivation to measure relative deprivation. This is a compound measure that includes metrics across the following domains: income deprivation, employment deprivation, crime, health deprivation and disability, education, skills and training deprivation, barriers to housing and services, and living environment deprivation [28]. Deprivation values are available for lower super output areas (LSOAs) in England, which are administrative boundaries with a mean residential population of 1500 people [21]. We calculated the mean Index of Multiple Deprivation value of LSOAs within, and intersecting, the boundary of each postcode district. For analyses, we split postcode districts into quintiles (Qs) of deprivation, with Q5 being the most deprived.

We used data from the 2011 rural/urban classification to categorize postcode districts as “rural,” when LSOAs within or intersecting their boundary were most frequently rural (populations <10,000 people within combined settlements, most of whom lived in rural-related areas), or “urban,” when intersecting LSOAs were most frequently urban (populations >10,000 people within combined settlements, most of whom lived in urban-related areas) [29]. We used data from the 2011 UK census [30] to identify the number of individuals who usually lived in a postcode district. Of the 2118 postcode districts, data for rural/urban classification and population density were available for 2097 (99%) and 2088 (95.4%) postcode districts, respectively, and did not change over the study period.

Statistical Analysis

Overview

We used the longitudinal analysis (“xt”) suite of tools in Stata (version 16.1; StataCorp) to complete statistical analysis [31]. We report findings from the start (November 2019) to the end (March 2022) of the study period in the *Results* section. Findings for all time points (November 2019 and then monthly between June 2020 and March 2022) are available in Multimedia Appendix 1.

Descriptive Statistics

For each measure, we calculated the median (IQR) or the mean (SD) and the percentage change from baseline at each time point.

Inferential Statistics

We used generalized estimating equations (GEE) to investigate changes in each outcome over time (the exposure measure).

We completed a complete case analysis, whereby included postcode districts had complete data on all relevant measures. Data for count-based outcomes (the number of food outlets registered to accept orders online and the number of food outlets accessible online) were not normally distributed and were overdispersed. We used negative binomial GEE to account for this. Negative binomial GEE report incidence rate ratios (IRRs) and 95% CIs. In the context of this study, IRRs are the expected change in the outcome value at each time point compared with the baseline value (November 2019). For the percentage of food outlets registered to accept orders online, we rescaled the data to be between 0 and 1 and specified a binomial distribution [32]. Model coefficients for this outcome are the change at each time point compared with baseline values. We first completed unadjusted analyses and then analyses adjusted for covariates. For the number of food outlets registered to accept orders online and the number of food outlets accessible online, we included population density, rural/urban classification, and the number of food outlets in the physical food environment as covariates in our adjusted model. For the percentage of food outlets registered to accept orders online, we only included population density and rural/urban classification as covariates.

We report the findings from our adjusted models in the *Results* section, and our unadjusted models in Multimedia Appendix 1. From our adjusted models, we also estimated the mean count from IRRs and the mean percentage from coefficients and report these in the *Results* section. We report the respective IRRs and coefficients in Multimedia Appendix 1.

Associations With Deprivation

For each outcome, we included an interaction term between time and deprivation in our adjusted GEE and completed a post hoc test for significance (with statistical significance set at $P<.01$ to account for multiple testing). When interaction terms were significant, we completed analyses stratified by deprivation. In November 2019, there were inequalities in access to food outlets selling food prepared out of home online [12]. For each outcome, we calculated a slope index of inequality measure at baseline (November 2019) and at the end of the study period (March 2022) to investigate how inequalities changed over time. This measure of inequality is the difference in the respective outcome between the least and most deprived areas, estimated using linear regression [33,34].

Ethical Considerations

Our study relied on publicly available data. Research ethics committee approval was not required.

Results

Overview

A descriptive summary of online food access in England is shown in Table 2 and Tables S1-S8 in Multimedia Appendix 1.

Table 2. Descriptive summary of online food accessibility at the postcode district level in England, stratified by deprivation quintile[a].

	Deprivation quintile, median (IQR)					England, median (IQR)
	1 (least deprived)	2	3	4	5 (most deprived)	
Number of food outlets registered to accept orders online[b]						
Count						
November 2019	3.0 (1.0 to 8.0)	4.0 (1.0 to 12.0)	6.0 (1.0 to 18.0)	13.0 (3.0 to 25.0)	24.0 (12.0 to 39.0)	7.0 (1.0 to 21.0)
March 2022	5.0 (1.0 to 15.0)	8.0 (1.0 to 21.0)	10.0 (2.0 to 29.0)	21.0 (4.0 to 41.0)	35.0 (20.0 to 59.0)	13.0 (3.0 to 34.0)
Change from baseline (%)[c]						
March 2022	80.0 (22.2 to 120.0)	69.0 (33.3 to 120.0)	66.7 (30.0 to 106.3)	62.8 (33.3 to 100.0)	57.9 (34.0 to 87.5)	65.4 (33.3 to 100.0)
Number of food outlets accessible online[d]						
Count						
November 2019	37.0 (14.0 to 70.5)	38.0 (10.0 to 96.0)	62.0 (8.5 to 134.5)	86.0 (12.0 to 190.0)	164.0 (87.0 to 273.0)	63.5 (16.0 to 156.0)
March 2022	27.0 (8.5 to 60.5)	27.0 (6.0 to 103.0)	50.0 (6.5 to 133.5)	95.0 (13.0 to 217.0)	175.0 (104.0 to 292.0)	57.0 (11.0 to 163.0)
Change from baseline (%)[c]						
March 2022	−12.7 (−48.6 to 20.0)	−7.5 (−48.8 to 28.6)	−1.1 (−35.0 to 33.2)	13.8 (−18.2 to 52.5)	13.1 (−7.9 to 33.8)	0.0 (−32.0 to 33.3)
Percentage of food outlets registered to accept orders online[e]						
Percent (%)						
November 2019	7.9 (2.2 to 14.9)	8.7 (1.5 to 19.0)	12.5 (1.9 to 23.5)	20.4 (6.5 to 30.8)	27.8 (19.7 to 37.4)	14.3 (3.8 to 26.0)
March 2022	13.2 (4.7 to 25.9)	14.9 (4.1 to 33.3)	20.5 (5.3 to 36.6)	30.8 (11.3 to 46.6)	41.9 (30.7 to 52.4)	24.0 (7.7 to 41.0)
Change from baseline (%)[c]						
March 2022	70.7 (16.7 to 116)	62.0 (25.0 to 106.2)	55.1 (24.0 to 98.9)	53.0 (24.8 to 88.9)	44.8 (22.4 to 75.8)	55.3 (23.0 to 96.1)

[a] Data are reported as median (IQR); postcode districts are small geographical units used for mail routing in England.

[b] Food outlets in the physical food environment registered to accept orders through the UK market leading online food delivery service.

[c] Baseline=November 2019.

[d] Food outlets registered to accept orders through the UK marketing leading online food delivery service that would deliver to a given postcode district.

[e] Calculated as the number of food outlets registered to accept orders through the UK market leading online food delivery service compared with the number of the physical food environment.

Number of Food Outlets Registered to Accept Orders Online

The summed number of food outlets registered to accept orders online in England increased from 29,232 in November 2019 to 49,752 in March 2022, equating to 70.2% growth (Figure 1).

The median number of food outlets registered to accept orders online per postcode district was 7.0 (IQR 1.0-21.0) in November 2019 and 13.0 (IQR 3.0-34.0) in March 2022. The median percent change from baseline (November 2019) per postcode district was 65.4 (IQR 33.3-100.0) in March 2022. The overall increase in the number of food outlets registered to accept orders was significant at each time point in our adjusted model (Table S10 in Multimedia Appendix 1 provides the IRRs), and there was significant effect modification by deprivation (*P*<.001). Estimated means calculated from IRRs of our adjusted negative binomial GEE are shown in Figure 2. At each level of deprivation, we observed that the estimated number of outlets registered to accept orders online had initially increased from baseline levels (November 2019); this decreased immediately after June 2020 and was then followed by a more consistent upward trend. The estimated number was consistently highest in the most deprived areas (Q5) and lowest in the least deprived (Q1). Absolute growth over time was also highest in the most deprived areas. For these areas, the estimated number was 39.1 outlets at the end of the study period, compared with 24.6 outlets at baseline, whereas for the least deprived areas, this was 6.6 and 3.5 outlets, respectively. The slope index of inequality between the least and most deprived areas was 5.0 (95% CI 4.5-4.5) outlets at baseline, and 7.8 (95% CI 7.0-8.6) outlets in March 2022.

Figure 1. Summed number (count) of food outlets registered to accept orders online in England between November 2019 and March 2022. Shaded bars represent time points reported in the Results section: baseline (November 2019) and end (March 2022). No data were available from December 2019 to May 2020.

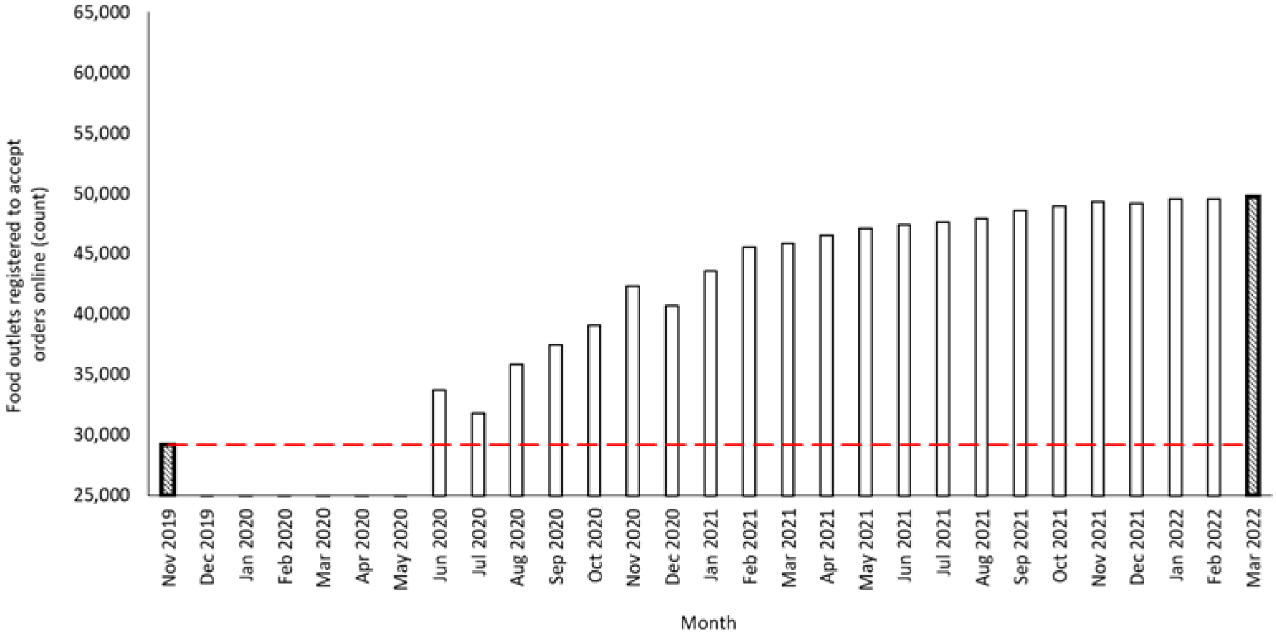
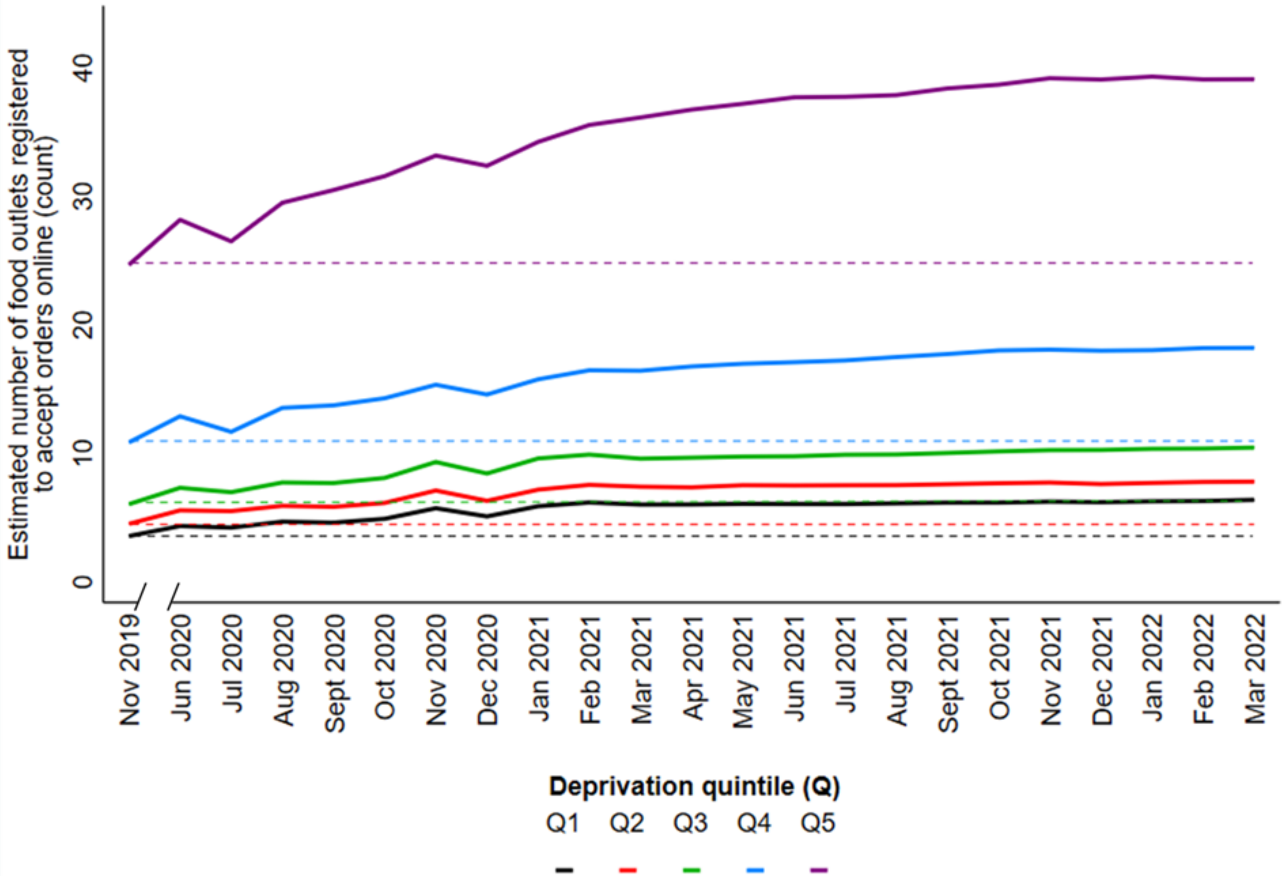


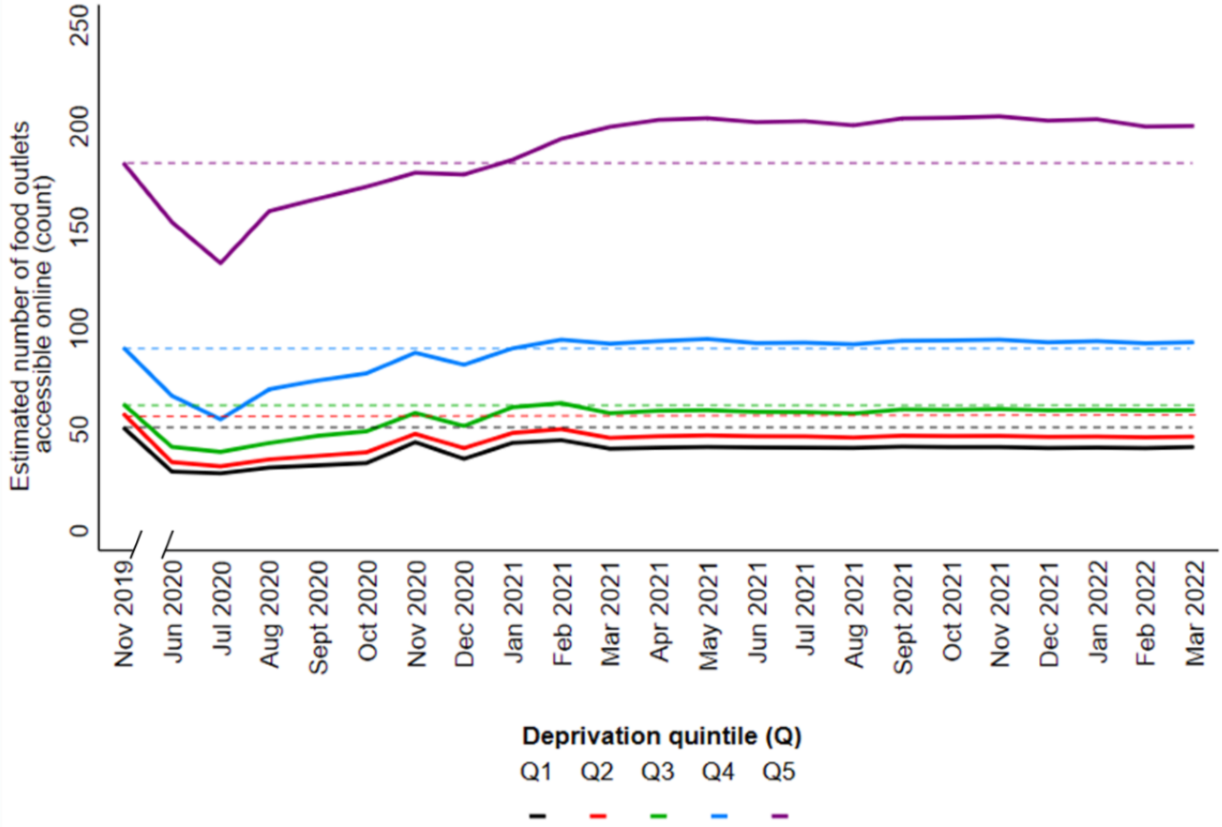
Figure 2. Estimated mean number (count) of food outlets registered to accept orders online in England, stratified by deprivation Q, where Q5 represents the most deprived postcode districts. Estimated number calculated from IRRs of a negative binomial generalized estimated equation adjusted for population density, rural/urban classification, and the number of food outlets in the physical food environment. In all, 2067 postcode districts were included. IRR: incidence rate ratio; Q: quintile.



Number of Food Outlets Accessible Online

The median number of food outlets accessible online per postcode district was 63.5 (IQR 16.0-156.0) in November 2019 and 57.0 (IQR 11.0-163.0) in March 2022. The median percent change from baseline per postcode district was 0.0 (IQR −32.0 to 33.3) in March 2022. The overall decrease in the number of food outlets accessible online was significant at each time point in our adjusted model (Table S12 in Multimedia Appendix 1 provides the IRRs), and there was a significant effect modification by deprivation (*P*<.001). Estimated means calculated from the IRRs of our adjusted negative binomial GEE are shown in Figure 3. At each level of deprivation, we observed that the estimated number of food outlets accessible online had decreased from baseline in June 2020. Although this was followed by an upward trajectory, the estimated number remained lower than baseline in less deprived areas (Q1-Q3) but surpassed baseline in those in the 2 upper Qs of deprivation (Q4 and Q5). For the most deprived areas (Q5), the estimated number increased from 181.9 outlets in November 2019 to 200.0 outlets in March 2022, and this contributed to an increasing gap in online food outlet access between the least and most deprived areas. The slope index of inequality between the least and most deprived areas was 32.0 (95% CI 28.1-35.9) outlets at baseline and 37.3 (95% CI 31.8-42.9) outlets in March 2022.

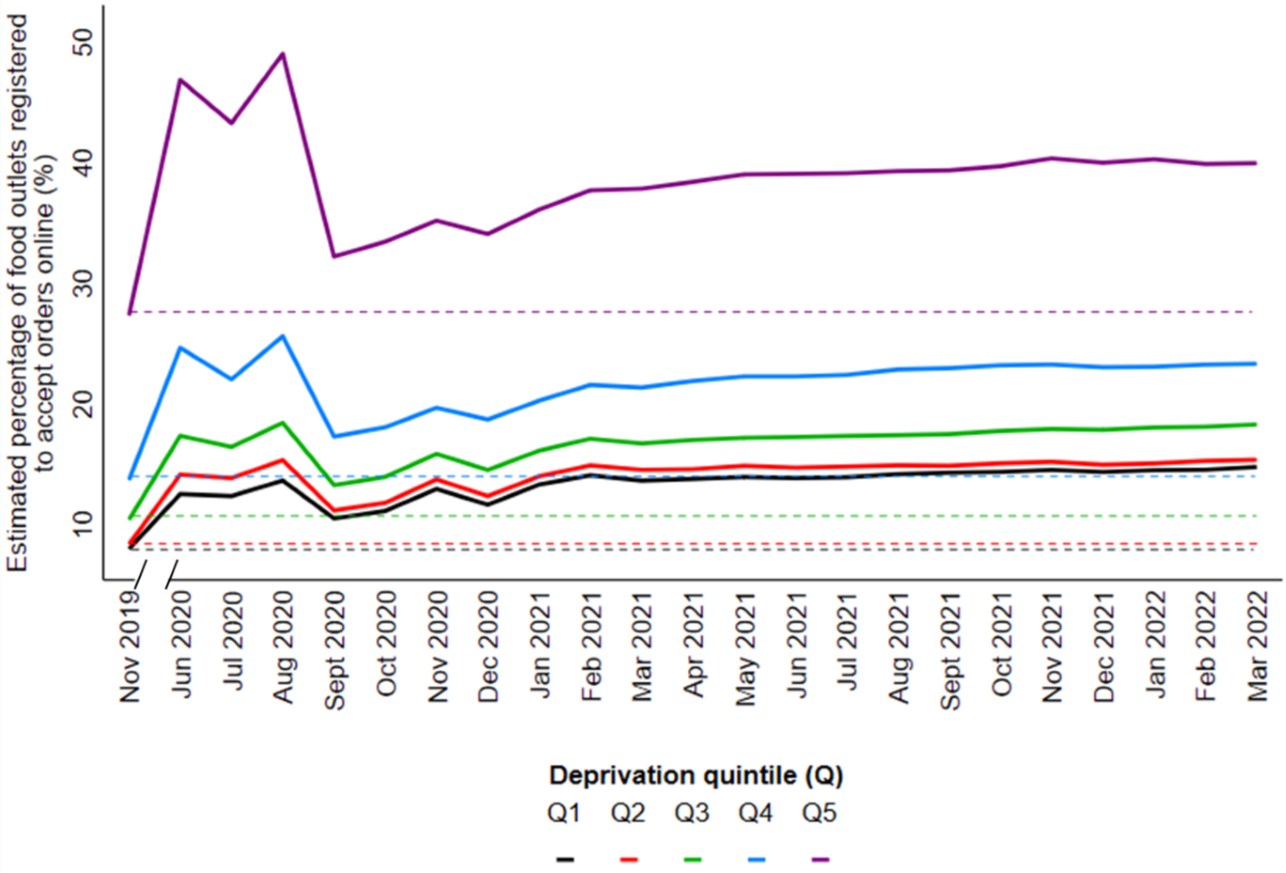
Figure 3. Estimated mean number (count) of food outlets accessible online in England, stratified by deprivation Q, where Q5 represents the most deprived postcode districts. Estimated number calculated from IRRs of a negative binomial generalized estimated equation adjusted for population density, rural/urban classification, and the number of food outlets in the physical food environment. In all, 2067 postcode districts were included. IRR: incidence rate ratio; Q: quintile.



Percentage of Food Outlets Registered to Accept Orders Online

The median number of food outlets registered to accept orders online as a percentage of the number of food outlets in the physical food environment per postcode district was 14.3 (IQR 3.8-26.0) in November 2019 and 24.0 (IQR 6.2-43.5) in March 2022. The median percent change from baseline per postcode district was 55.3 (IQR 23.0-96.1) in March 2022. The overall increase in the percentage of food outlets registered to accept orders online was significant at each time point in our adjusted model (Table S14 in Multimedia Appendix 1 provides the coefficients), and there was a significant effect modification by deprivation (*P*<.001). Estimated means calculated from the coefficients of our adjusted GEE are shown in Figure 4. We observed an initial increase from baseline in June 2020 that was followed by a slight decline, a second increase that equaled or surpassed previous levels, and then another decline before a more stable increase. Although this trend was evident across all levels of deprivation and the estimated percentage was significantly increased at each level of deprivation by the end of the study period, the magnitude varied. Nevertheless, the estimated mean percentage was the highest in the most deprived areas (40.0% in March 2022 compared with 27.4% in November 2019), with these areas also having a higher growth over time in absolute terms, compared with other areas. The slope index of inequality between the least and most deprived areas was 4.5% (95% CI 4.1-4.9) at baseline and 5.9% (95% CI 5.4-6.5) in March 2022.

Figure 4. Estimated mean percentage (%) of food outlets registered to accept orders online in England, stratified by derivation Q, where Q5 represents the most deprived postcode districts. Estimated percentage calculated from coefficients of a generalized estimating equation adjusted for population density and rural/urban classification. In all, 2065 postcode districts were included. Q: quintile.



Discussion

Summary of Findings

To the best of our knowledge, our study was the first to examine changes in levels of online food access across an entire country during the COVID-19 pandemic. We identified an increase in the number of food outlets registered to accept orders online, reaching approximately 50,000 in England by March 2022. There was a parallel increase in the percentage of food outlets in the physical food environment registered to accept orders online. In contrast, the number of food outlets accessible online, that is, those that would deliver to a given population living in a postcode district, was on average lower in March 2022 than before the COVID-19 pandemic. The magnitude of change for all measures of online food access varied by level of deprivation. The most deprived postcode districts consistently had the highest number and percentage of food outlets registered to accept orders online. Importantly, contrary to the trend for England as a whole, only in the most deprived postcode districts was the number of food outlets accessible online higher in March 2022 than at baseline. We did not observe this for postcode districts at any other level of deprivation, with the number *decreasing* in less deprived areas. As a result, there is some evidence that existing socioeconomic inequalities in the opportunity to use online food delivery services to purchase food prepared out of home widened over time.

Comparison With Prior Work

Our finding that an increased number of food outlets had registered to accept orders online in the first 2 years of the COVID-19 pandemic aligns with the reports presented in business and news media from major online food delivery services operating in England [35,36]. Moreover, the decrease in July 2020 accords with contemporaneous reports of food outlet closures and decreased order volume through online food delivery services [37]. For the most part, after July 2020, we observed an increase in the number of food outlets registered to accept orders online. As opportunities for on-premises food consumption were limited by national stay-at-home orders imposed in March 2020, followed by periods of time with restrictions on the capacity for on-premises dining, food outlet owners perhaps made a strategic business decision to register to accept orders online as a way to generate revenue.

For England as a whole, we identified an overall decline in the number of food outlets *accessible* online (ie, the number that will deliver to a given population living in a postcode district) between November 2019 and March 2022. This is in contrast to the increased number of food outlets *registered* to accept orders over the same period (irrespective of whom they would deliver to). The decline we observed was particularly pronounced in the early stages of the COVID-19 pandemic, which might have reflected a period of transition among those that had only recently registered to accept orders online. Food outlets unaccustomed to food delivery plausibly operated with a more limited delivery radius to ensure they could fulfill customer orders. Moreover, at this time, there was broader workforce capacity instability owing to self-isolation rules, unclear restrictions on maximum travel distances legally allowed, and concerns for online food delivery service courier safety [38]. All of which might have contributed to the implementation of limited delivery areas.

Although the initial decline in the number of food outlets accessible online was followed by an increase at all levels of deprivation, it was only among more deprived areas, and particularly the most deprived, that the number eventually surpassed prepandemic levels over the period of our study. The number of food outlets accessible online was already the highest in the most deprived areas in England before the pandemic [12]. However, we found evidence of an increasing divergence between the most and least deprived areas, suggesting that inequality in online food access widened during the period of our study. A perceived lack of demand in less deprived areas might have meant that registered food outlets chose not to include them in their delivery areas. Moreover, the delivery of food sold through online food delivery services is mostly completed by couriers on bicycles. As such, there is a natural limit in the distance that can be traveled while maintaining food quality. In rural, less deprived areas, this challenge might be insurmountable. Regardless, given that the number of food outlets accessible online is positively associated with online food delivery service use [11], it is possible that food-purchasing practices, diet, and health were negatively influenced during the COVID-19 pandemic, especially among populations living in more deprived areas. Although adults living in the United Kingdom self-reported a decrease in their diet quality during the COVID-19 pandemic [39], evidence on changes in the frequency of having food prepared out of home delivered is inconsistent [40]. Further research is required to understand the extent to which the changes in online food outlet access we observed were associated with changes in online food delivery service use during the same period and subsequent implications for diet and health. This future research could incorporate further measures of food access, for example, accounting for the affordability of food sold, which is recognized by online food delivery service customers as being an important consideration preceding use [41]. It is plausible that it is not financially viable for populations living in the most deprived areas to use online food delivery services despite the increase in the number of food outlets accessible, especially owing to the increase in the cost of living in the United Kingdom and elsewhere [42].

The number of food outlets registered to accept orders online as a percentage of the number of food outlets in the physical food environment had increased by the end of the study period. However, this increase was only apparent after a period of instability. Owners of food outlets with customer-facing premises in the physical food environment reported that although being registered to accept orders online was a way for customers to access their food when on-premises food consumption and travel was restricted, simultaneously managing orders placed in-person and online was difficult [38]. If food outlets that did not previously accept orders online only did so out of necessity when on-premises food consumption was restricted, it is plausible that they subsequently deregistered when restrictions ended. This scenario would partly explain the successive increases and decreases that coincided with the start and end of restrictions introduced during the COVID-19 pandemic (Figure 4).

Although the relative change over time with respect to the percentage of food outlets registered to accept orders online was consistent across all levels of deprivation, the absolute change was highest in the most deprived areas, which led to a widening of absolute inequality in online access to food prepared out of home. Food sold through online food delivery services is typically prepared in the kitchen facilities of food outlets located in the physical food environment [9]. Our finding is likely a reflection of the existing urban form in the most deprived areas in England, which is characterized by a high density and concentration of food outlets [43,44]. Although we cannot be certain, if more food outlets were operating within emergency regulations introduced during the COVID-19 pandemic, more outlets may have also registered to accept orders online. Our findings therefore provide evidence to suggest that the introduction of emergency regulations at least partly contributed to widening inequality in this exposure.

There remains considerable scope for the percentage of food outlets in the physical food environment that are registered to accept orders online to increase. This emphasizes the coexisting and overlapping nature of digital and physical food environments, which together provide multiple opportunities to purchase energy-dense, nutrient-poor food [45]. However, there are likely to be natural limits to growth because all food outlets do not necessarily need or want to register to accept orders online. In fact, the relative stability from June 2021 onward suggests that a plateau might have already occurred. An important limitation of this paper is that we do not have additional prepandemic data that would allow us to account for any existing trends in analyses. Nevertheless, our findings and associated data can contribute to future surveillance of longer-term trends.

Possible Implications for Public Health and Policy

Urban planning has been used by over half of local authorities in England to promote the creation of healthier physical food environments, specifically by preventing new takeaway food outlets from opening [46]. To our knowledge, restrictions targeted specifically at online food delivery services are not in place. Although not yet fully clear, similar to the physical food environment, interventions to restrict access to food prepared out of home through these services might be increasingly necessary in the future [47,48]. At this point, public health interventions that do not necessarily attempt to restrict online food access but instead attempt to mitigate the potential public health burden posed by online food delivery service use possibly represent the most appropriate route to regulation. For example, changing the nutritional composition of food sold inside outlets would be expected to have a crossover effect and also change food sold online. Additionally, in England, in 2022, mandatory calorie labeling of non–prepacked food and soft drinks was introduced for retailers that operate with more than 250 employees, which meant that this information had to be displayed both inside food outlets and on the platforms of online food delivery services [49,50]. The introduction of this regulation recognizes the role of the digital food environment in the purchase of food prepared out of home. Research in the future could seek to further understand if and how regulations in the physical food environment can be extended to the digital food environment.

Furthermore, there has been evolution in the preparation location for the food sold through online food delivery services. This is best demonstrated by the development of facilities known as dark kitchens [51]. These facilities allow food businesses to register to accept orders through online food delivery services and prepare meals for delivery without the financial costs of having a customer-facing premises. To date, there is only limited evidence on the dark kitchen business model, which may not be generalizable beyond the data collection context of a single London borough [52]. Nevertheless, the development of these facilities plausibly influenced our findings related to the percentage of food outlets registered to accept orders online. Future research might seek to understand the public health implications of the dark kitchen business model, including how it uniquely contributes to the number of food outlets accessible online.

Methodological Considerations and Limitations

Our data collection spanned 2 years and allowed us to closely monitor trends in important metrics of online food outlet access. In doing so, we present new baseline levels for future assessment. Nevertheless, our study is not without limitations. Our analysis was observational, and we cannot definitively conclude that the changes we observed were due to the COVID-19 pandemic. We also do not have sufficient prepandemic data that would allow underlying or existing trends to be accounted for in analyses. This is particularly the case for changes between November 2019 and June 2020 when we had no data, and we could not determine whether changes coincided with the COVID-19 pandemic.

We completed data collection at a single time point on a monthly basis. If food outlets were only registered to accept orders or accessible in the intervening period, they might not have been returned in our searches. However, the changes in the number of food outlets accessible online that we observed, especially the initial decrease in June 2020, is similar to reports of food outlet closures and decreased order volume at this time [37].

We did not track individual food outlets over the study period. Instead, we studied the total number of food outlets registered to accept orders at any given time point. This means that we are unable to comment on the number that were newly registered during the COVID-19 pandemic. It would be interesting for future work to investigate the extent to which food outlets that registered only during this period remained registered.

Our findings demonstrate *potential* online food outlet access. For this to be realized, any given individual must be an online food delivery service customer. As such, our findings do not necessarily translate into realized individual-level online food outlet access [53].

We used postcode districts as our unit of analysis and acknowledge the possibility that our findings are subject to the modifiable areal unit problem. Although the spatial unit that we adopted for analyses has the potential to introduce bias [54], our approach was data driven to allow consistency with our previous research [12]. Relatedly, using postcode districts as our unit of analysis also meant that we were limited to using boundary data from 2012. These data were not temporally aligned with further covariate data that were the most recently published or collected.

Conclusions

We investigated changes in multiple measures of food access through online food delivery services during the first 2 years of the COVID-19 pandemic. We identified that the number of food outlets in England that were registered to accept orders online increased. In parallel, the number of food outlets registered to accept orders online as a percentage of the number of food outlets in the physical food environment increased. Although the number of food outlets that could be accessed decreased for the whole of England, trends differed by the level of deprivation. The number of outlets decreased compared with prepandemic levels in the least deprived areas and increased to surpass prepandemic levels *only* in the most deprived.

Overall, then, we identified considerable changes in measures of online food accessibility during a period when on-premises food consumption was often restricted. Our data from March 2022 represent a new baseline to which future changes in measures of online food accessibility can be investigated. Future research might attempt to understand the extent to which the changes we identified were associated with the changes in online food delivery service use and, in turn, the implications for diet quality and health.

Acknowledgments

MK was funded by the National Institute for Health and Care Research (NIHR) School for Public Health Research (SPHR; grant number PD_SPH_2015). This work was supported by the Medical Research Council (grant number MC_UU_00006_7). The views expressed are those of the authors and not necessarily those of any of the above named funders. The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; or in writing the manuscript.

Data Availability

The data that support the findings of this study were collected from sources available in the public domain: Just Eat Takeaway, the UK Census, the Indices of Multiple Deprivation, and UK boundary data. Data availability is subject to request and license restrictions. Further data were sourced from Digimap Edina, and we included data from PointX (c) DatabaseRight/Copyright (2022) and OS (c) Crown Copyright (2022). Data collected from Just Eat are accessible upon reasonable request; however, their use is for noncommercial purposes only.

Authors' Contributions

MK was responsible for conceptualization, data curation, formal analysis, methodology, and writing—original draft. JA and TB contributed to conceptualization, methodology, supervision, and writing—review and editing. All authors read and agreed to the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 120 KB - publichealth_v9i1e41822_app1.docx]

References

1. Wellard-Cole L, Davies A, Allman-Farinelli M. Contribution of foods prepared away from home to intakes of energy and nutrients of public health concern in adults: a systematic review. Crit Rev Food Sci Nutr 2022;62(20):5511-5522. [doi: 10.1080/10408398.2021.1887075] [Medline: 33596740]
2. Food prices and spending. United States Department of Agriculture Economic Research Service. 2020. URL: https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/food-prices-and-spending/ [accessed 2020-09-16]
3. Cornelsen L, Berger N, Cummins S, Smith RD. Socio-economic patterning of expenditures on 'out-of-home' food and non-alcoholic beverages by product and place of purchase in Britain. Soc Sci Med 2019 Aug;235:112361. [doi: 10.1016/j.socscimed.2019.112361] [Medline: 31262504]
4. Jaworowska AM, Blackham TM, Long R, Taylor C, Ashton M, Stevenson L, et al. Nutritional composition of takeaway food in the UK. Nutr Food Sci 2014;44(5):414-430 [FREE Full text] [doi: 10.1108/nfs-08-2013-0093]
5. Robinson E, Jones A, Whitelock V, Mead BR, Haynes A. (Over)eating out at major UK restaurant chains: observational study of energy content of main meals. BMJ 2018 Dec 12;363:k4982 [FREE Full text] [doi: 10.1136/bmj.k4982] [Medline: 30541906]
6. Huang Y, Burgoine T, Theis DR, Adams J. Differences in energy and nutrient content of menu items served by large chain restaurants in the USA and the UK in 2018. Public Health Nutr 2022 Jun 01;25(10):2671-2679. [doi: 10.1017/S1368980022001379] [Medline: 35642073]
7. Pereira MA, Kartashov AI, Ebbeling CB, Van Horn L, Slattery ML, Jacobs Jr DR, et al. Fast-food habits, weight gain, and insulin resistance (the CARDIA study): 15-year prospective analysis. Lancet 2005 Jan;365(9453):36-42. [doi: 10.1016/S0140-6736(04)17663-0] [Medline: 15639678]
8. Keeble M, Adams J, Sacks G, Vanderlee L, White CM, Hammond D, et al. Use of online food delivery services to order food prepared away-from-home and associated sociodemographic characteristics: a cross-sectional, multi-country analysis. Int J Environ Res Public Health 2020 Jul 17;17(14):5190 [FREE Full text] [doi: 10.3390/ijerph17145190] [Medline: 32709148]
9. Granheim SI, Løvhaug AL, Terragni L, Torheim LE, Thurston M. Mapping the digital food environment: a systematic scoping review. Obes Rev 2022 Jan;23(1):e13356. [doi: 10.1111/obr.13356] [Medline: 34519396]
10. Caspi CE, Sorensen G, Subramanian SV, Kawachi I. The local food environment and diet: a systematic review. Health Place 2012 Sep;18(5):1172-1187 [FREE Full text] [doi: 10.1016/j.healthplace.2012.05.006] [Medline: 22717379]
11. Keeble M, Adams J, Vanderlee L, Hammond D, Burgoine T. Associations between online food outlet access and online food delivery service use amongst adults in the UK: a cross-sectional analysis of linked data. BMC Public Health 2021 Oct 31;21(1):1968 [FREE Full text] [doi: 10.1186/s12889-021-11953-9] [Medline: 34719382]
12. Keeble M, Adams J, Bishop TR, Burgoine T. Socioeconomic inequalities in food outlet access through an online food delivery service in England: a cross-sectional descriptive analysis. Appl Geogr 2021 Aug;133:102498 [FREE Full text] [doi: 10.1016/j.apgeog.2021.102498] [Medline: 34345056]
13. Peeters A, Blake MR. Socioeconomic inequalities in diet quality: from identifying the problem to implementing solutions. Curr Nutr Rep 2016 Jul 1;5(3):150-159 [FREE Full text] [doi: 10.1007/s13668-016-0167-5]
14. Foster IS, Liu SY, Hoffs CT, LeBoa C, Chen AS, Rummo PE. Disparities in SNAP online grocery delivery and implementation: lessons learned from California during the 2020-21 COVID pandemic. Health Place 2022 Jul;76:102811 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102811] [Medline: 35605572]
15. Colbert S, Wilkinson C, Thornton L, Richmond R. COVID-19 and alcohol in Australia: industry changes and public health impacts. Drug Alcohol Rev 2020 Jul;39(5):435-440 [FREE Full text] [doi: 10.1111/dar.13092] [Medline: 32483882]

16. New regulations created by Secretary of State for business closure (COVID-19). United Kingdom Government. 2020 Mar 22. URL: https://www.gov.uk/government/news/new-regulations-created-by-secretary-of-state-for-business-closure-covid-19 [accessed 2022-05-31]
17. Government to grant permission for pubs and restaurants to operate as takeaways as part of coronavirus response. United Kingdom Government. 2020 Mar 17. URL: https://www.gov.uk/government/news/government-to-grant-permission-for-pubs-and-restaurants-to-operate-as-takeaways-as-part-of-coronavirus-response [accessed 2022-05-31]
18. Town and Country Planning. 2020 No. 1243. United Kingdom Legislation. 2020. URL: https://www.legislation.gov.uk/uksi/2020/1243/made/data.pdf [accessed 2022-05-31]
19. Maguire ER, Burgoine T, Monsivais P. Area deprivation and the food environment over time: a repeated cross-sectional study on takeaway outlet density and supermarket presence in Norfolk, UK, 1990-2008. Health Place 2015 May;33:142-147 [FREE Full text] [doi: 10.1016/j.healthplace.2015.02.012] [Medline: 25841285]
20. All you need to know about postcodes but were afraid to ask. Beacon Dodsworth. 2020. URL: https://beacon-dodsworth.co.uk/blog/all-you-need-to-know-about-postcodes-but-were-afraid-to-ask/ [accessed 2022-07-11]
21. Census geography. Office for National Statistics. 2016. URL: https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography [accessed 2022-05-10]
22. Census boundary data. UK Data Service. 2020. URL: https://census.ukdataservice.ac.uk/get-data/boundary-data.aspx [accessed 2020-09-15]
23. Restaurant delivery and takeaway in the United Kingdom - statistics and facts. Statista. 2020. URL: https://www.statista.com/topics/4679/food-delivery-and-takeaway-market-in-the-united-kingdom-uk/ [accessed 2021-02-20]
24. Just Eat Takeaway.com Q4 2020 Trading Update. Just Eat. 2021. URL: https://www.justeattakeaway.com/media/press-releases/ [accessed 2021-01-15]
25. Webscraper IO. 2022. URL: https://www.webscraper.io/ [accessed 2022-05-24]
26. Doogal. 2022. URL: https://www.doogal.co.uk/ [accessed 2022-05-24]
27. Points of Interest - Technical Specification. Ordnance Survey. 2019. URL: https://www.ordnancesurvey.co.uk/documents/product-support/tech-spec/points-of-interest-technical-specification-v4.3.pdf [accessed 2022-07-11]
28. English indices of deprivation 2019. Ministry of Housing Communities and Local Government. 2019. URL: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [accessed 2022-07-11]
29. 2011 rural/urban classification. Office for National Statistics. 2016. URL: https://www.ons.gov.uk/methodology/geography/geographicalproducts/ruralurbanclassifications/2011ruralurbanclassification [accessed 2022-07-11]
30. Usual resident population. Office for National Statistics. 2013. URL: https://www.nomisweb.co.uk/census/2011/ks101ew [accessed 2022-07-11]
31. Longitudinal-Data / Panel-Data reference manual Release 17. StataCorp. 2021. URL: https://www.stata.com/manuals/xt.pdf [accessed 2022-02-11]
32. How does one do regression when the dependent variable is a proportion? UCLA: Statistical Consulting Group. 2021. URL: https://stats.oarc.ucla.edu/stata/faq/how-does-one-do-regression-when-the-dependent-variable-is-a-proportion/ [accessed 2022-05-10]
33. Khang YH, Lim D, Bahk J, Kim I, Kang HY, Chang Y, et al. A publicly well-accepted measure versus an academically desirable measure of health inequality: cross-sectional comparison of the difference between income quintiles with the slope index of inequality. BMJ Open 2019 Jun 27;9(6):e028687 [FREE Full text] [doi: 10.1136/bmjopen-2018-028687] [Medline: 31248930]
34. Renard F, Devleesschauwer B, Speybroeck N, Deboosere P. Monitoring health inequalities when the socio-economic composition changes: are the slope and relative indices of inequality appropriate? Results of a simulation study. BMC Public Health 2019 May 30;19(1):662 [FREE Full text] [doi: 10.1186/s12889-019-6980-1] [Medline: 31146708]
35. Nicolle E. Uber Eats orders skyrocket 160 per cent in lockdown. City A.M. 2020 Jun 29. URL: https://www.cityam.com/uber-eats-orders-in-the-uk-skyrocket-as-lockdown-powers-up-delivery-apps/ [accessed 2022-05-31]
36. Q1 Trading Update. Deliveroo. 2022. URL: https://corporate.deliveroo.co.uk/investors/results-reports-presentations/ [accessed 2022-05-31]
37. How has coronavirus affected pubs, cafes and restaurants? Economics Observatory. 2020 Jul 2. URL: https://www.economicsobservatory.com/how-has-coronavirus-affected-pubs-cafes-and-restaurants [accessed 2022-06-06]
38. Ecker Y, Strüver A. Towards alternative platform futures in post-pandemic cities? A case study on platformization and changing socio-spatial relations in on-demand food delivery. Digit Geogr Soc 2022;3:100032 [FREE Full text] [doi: 10.1016/j.diggeo.2022.100032]
39. Robinson E, Boyland E, Chisholm A, Harrold J, Maloney NG, Marty L, et al. Obesity, eating behavior and physical activity during COVID-19 lockdown: a study of UK adults. Appetite 2021 Jan 01;156:104853 [FREE Full text] [doi: 10.1016/j.appet.2020.104853] [Medline: 33038479]
40. Acton RB, Vanderlee L, Cameron AJ, Goodman S, Jáuregui A, Sacks G, et al. Self-reported impacts of the COVID-19 pandemic on diet-related behaviors and food security in 5 countries: results from the International Food Policy Study 2020. J Nutr 2022 Jun 13;152(Suppl 1):35S-46S [FREE Full text] [doi: 10.1093/jn/nxac025] [Medline: 35274699]

41. Keeble M, Adams J, Burgoine T. Investigating experiences of frequent online food delivery service use: a qualitative study in UK adults. BMC Public Health 2022 Jul 16;22(1):1365 [FREE Full text] [doi: 10.1186/s12889-022-13721-9] [Medline: 35842625]
42. Robinson E. Obesity and the cost of living crisis. Int J Obes (Lond) 2023 Feb;47(2):93-94 [FREE Full text] [doi: 10.1038/s41366-022-01242-9] [Medline: 36456646]
43. Townshend T, Lake AA. Obesogenic urban form: theory, policy and practice. Health Place 2009 Dec;15(4):909-916. [doi: 10.1016/j.healthplace.2008.12.002] [Medline: 19201641]
44. Townshend TG. Toxic high streets. J Urban Des 2017;22(2):167-186 [FREE Full text] [doi: 10.1080/13574809.2015.1106916]
45. Lord C, Bates O, Friday A, McLeod F, Cherrett T, Martinez-Sykora A, et al. The sustainability of the gig economy food delivery system (Deliveroo, UberEATS and Just-Eat): histories and futures of rebound, lock-in and path dependency. Int J Sustain Transp (forthcoming) 2022 May 02:1-13 [FREE Full text] [doi: 10.1080/15568318.2022.2066583]
46. Keeble M, Burgoine T, White M, Summerbell C, Cummins S, Adams J. How does local government use the planning system to regulate hot food takeaway outlets? A census of current practice in England using document review. Health Place 2019 May;57:171-178 [FREE Full text] [doi: 10.1016/j.healthplace.2019.03.010] [Medline: 31055107]
47. Digital food environments: factsheet. World Health Organization, Regional Office for Europe. 2021. URL: https://apps.who.int/iris/bitstream/handle/10665/342072/WHO-EURO-2021-2755-42513-59052-eng.pdf?sequence=1&isAllowed=y [accessed 2022-03-07]
48. Slide to order: a food systems approach to meals delivery apps: WHO European Office for the Prevention and Control of Noncommunicable diseases. World Health Organization, Regional Office for Europe. 2021. URL: https://apps.who.int/iris/handle/10665/350121 [accessed 2022-03-08]
49. Robinson E, Marty L, Jones A, White M, Smith R, Adams J. Will calorie labels for food and drink served outside the home improve public health? BMJ 2021 Jan 20;372:n40. [doi: 10.1136/bmj.n40] [Medline: 33472836]
50. Yeo GS. Is calorie labelling on menus the solution to obesity? Nat Rev Endocrinol 2022 Aug;18(8):453-454 [FREE Full text] [doi: 10.1038/s41574-022-00705-3] [Medline: 35680986]
51. Shapiro A. Platform urbanism in a pandemic: dark stores, ghost kitchens, and the logistical-urban frontier. J Consum Cult 2023 Feb;23(1):168-187 [FREE Full text] [doi: 10.1177/14695405211069983]
52. Rinaldi C, D'Aguilar M, Egan M. Understanding the online environment for the delivery of food, alcohol and tobacco: an exploratory analysis of 'dark kitchens' and rapid grocery delivery services. Int J Environ Res Public Health 2022 May 02;19(9):5523 [FREE Full text] [doi: 10.3390/ijerph19095523] [Medline: 35564918]
53. Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. Am J Public Health 1994 May;84(5):819-824. [doi: 10.2105/ajph.84.5.819] [Medline: 8179055]
54. Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. Environ Plan A 1991 Jul;23(7):1025-1044 [FREE Full text] [doi: 10.1068/a231025]

Abbreviations

GEE: generalized estimating equations
GIS: Geographic Information System
IRR: incidence rate ratio
LSOA: lower super output area
OS POI: Ordnance Survey Points of Interest
Q: quintile

Edited by T Sanchez, A Mavragani; submitted 10.08.22; peer-reviewed by S Jia, A Kalbus; comments to author 10.01.23; revised version received 08.02.23; accepted 23.02.23; published 17.04.23.

Please cite as:
Keeble M, Adams J, Burgoine T
Changes in Online Food Access During the COVID-19 Pandemic and Associations With Deprivation: Longitudinal Analysis
JMIR Public Health Surveill 2023;9:e41822
URL: https://publichealth.jmir.org/2023/1/e41822
doi:10.2196/41822
PMID:36848236

Original Paper

Temporal Trends in Cardiovascular Health Status Among Chinese School-Aged Children From 1989 to 2018: Multiwave Cross-Sectional Analysis

Xijie Wang[1*], MD, PhD; Bin Dong[2*], MD, PhD; Feifei Huang[3*], MD, PhD; Ji Zhang[4], MD, PhD; Rongxin He[1], PhD; Shufa Du[5], MD; Jiguo Zhang[3], PhD; Jun Ma[2], MD, PhD; Huijun Wang[3], MD, PhD; Bing Zhang[3], PhD; Wannian Liang[1], MD

[1]Vanke School of Public Health, Tsinghua University, Beijing, China
[2]Institute of Child and Adolescent Health, School of Public Health, Peking University, Beijing, China
[3]National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention, Beijing, China
[4]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[5]Department of Nutrition and Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[*]these authors contributed equally

Corresponding Author:
Bing Zhang, PhD
National Institute for Nutrition and Health
Chinese Center for Disease Control and Prevention
No.27, Nanwei Road
Beijing, 100050
China
Phone: 86 10 66237298
Email: zhangbing@chinacdc.cn

Abstract

Background: Despite the release of updated metrics for Life’s Essential 8 (LE8), key indicators for assessing cardiovascular health (CVH) status, there is currently no report on their distribution among Chinese children.

Objective: This study aimed to assess the nationwide distribution of CVH in Chinese school-aged children using LE8 scores and analyze temporal changes in these scores over time.

Methods: Participants aged 7 to 19 years from 11 waves (between 1989 and 2018) of the China Health and Nutrition Survey were included in this study. LE8 components were grouped into 2 domains of health behaviors (diet, physical activity, nicotine exposure, sleep) and health factors (BMI, blood lipids, blood glucose, blood pressure). Scores of overall CVH and each LE8 metric were calculated individually. Temporal changes were assessed with joint point regression models by rural and urban living residence. The causal relationships between health behaviors and health factors that changed the most over time were built with cross-lagged panel models.

Results: A total of 21,921 participants, 52.6% (n=11,537) of whom were male, who had data for at least 4 CVH components were included in the analysis. The mean age was 13 (SD 3.6) years. The overall CVH score remained stable in most regions, with the lowest found in Shandong from East China, which had a mean between 67 (SD 10.9) and 67.2 (SD 12.4). In contrast, the highest score was found in Guizhou from Southwest China, with a mean between 71.4 (SD 10.8) and 74.3 (SD 10.3). In rural areas, the diet score decreased significantly from 1997 onward with a speed of 0.18 (95% CI: 0.15-0.21; *P*<.001) per year, and the BMI score decreased significantly from 2005 onward with a speed of 0.56 (95% CI 0.44-0.68; *P*<.001) per year. In urban areas, the diet score decreased from 1994 onward with a speed of 0.03 (95% CI: 0.001-0.07; *P*=.04) per year, and the BMI score decreased from 2002 onward with a speed of 0.63 (95% CI 0.47-0.79; *P*<.001) per year. The sleep score dropped constantly in both urban and rural areas, with a speed of 0.69 (95% CI 0.58-0.80; *P*<.001) and 0.69 (95% CI: 0.52-0.86; *P*<.001) per year, respectively. A decline in the diet score led to a decline in the BMI score with a coefficient of 0.190 (95% CI 0.030-0.351; *P*=.02), while a decline in the BMI score led to a decline in sleep health with a coefficient of 0.089 (95% CI 0.010-0.168; *P*=.03).

Conclusions: Chinese school-aged children and adolescents were generally of moderate CVH status, but mutual influences existed between CVH metrics. Dietary interventions should be prioritized for promoting overall CVH in the future.

(JMIR Public Health Surveill 2023;9:e45564) doi:10.2196/45564

KEYWORDS

cardiovascular health; school-aged children; temporal change; China Health and Nutrition Survey

Introduction

Cardiovascular disease (CVD) is the leading cause of death and disability in the world, in large part because of risk factors modifiable by changes in behavior [1,2]. Individuals with an unfavorable cardiovascular health (CVH) status during childhood are more likely to develop subclinical atherosclerosis and subsequent CVD later in life [3]. Controlling and reducing CVD-related mortality and morbidity are not only one of the key indicators of the United Nations' Sustainable Development Goals [2] but also one of the top priorities of the 2030 Healthy China Plan [4].

Although the overall CVD mortality has stalled in recent years, especially in middle- and low-income countries [5], efficient prevention strategies could further effectively reduce age-adjusted death rates due to CVD [6,7]. This points to the importance of recognizing, evaluating, and maintaining overall CVH.

The newly released Life's Essential 8 (LE8), key measures for improving heart health released by the American Heart Association, provide a comprehensive tool to evaluate overall CVH status and specific metrics of CVH behaviors or factors. Compared with the previous version, Life's Simple 7 [8], the new version provides not only classifications of CVH status but also continuous scores that are helpful for both cross-sectional and longitudinal analyses on differences and trends in CVH status distribution. The revised version incorporates newly identified risk factors, such as sleep health, alongside metrics that have demonstrated stronger associations with long-term CVD, such as the replacement of total cholesterol with non–high-density lipoprotein (HDL) cholesterol [9,10].

The newly released metrics have garnered significant attention in the field of CVH research. However, prior studies have not addressed either the national distribution of CVH status among Chinese school-aged children or the temporal changes in CVH metrics over time. Accordingly, in this study, we collected data pertaining to children and adolescents aged 7 to 19 years from waves spanning 1989 to 2018 in the China Health and Nutrition Survey (CHNS). We aimed to analyze temporal trends and geographical variations in CVH in Chinese school-aged children. Additionally, we aimed to investigate whether changes in health behaviors over the past 3 decades have had an impact on health factors.

Methods

Study Participants

The study data were obtained from 11 waves of the CHNS, an open prospective cohort study carried out by the University of North Carolina at Chapel Hill in partnership with the National Institute for Nutrition and Health of the Chinese Center for Disease Control and Prevention [11]. The baseline survey was conducted in 1989 and followed up in 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, 2015, and 2018, respectively.

This project covered 15 provinces (autonomous regions and municipalities) including Heilongjiang, Liaoning (did not participate in 1997), Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, Guizhou, Beijing, Shanghai, Chongqing, Shaanxi, Zhejiang, and Yunnan. Notably, Heilongjiang only participated from 1997 onward; Beijing, Chongqing, and Shanghai participated from 2011 onward; and Zhejiang, Yunnan, and Shaanxi participated from 2015 onward. A multistage, stratified, clustered random method was used for sampling. Based on the national administrative regions, a weighted sampling scheme was used to randomly select 2 cities and 4 counties from each province. Next, 2 urban and 2 suburban neighborhoods were randomly selected in each city, and 1 township and 3 villages were randomly selected in each county. Urban neighborhoods in the cities and townships in the counties were grouped into urban areas, and suburban neighborhoods in the cities and villages in the counties were grouped into rural areas. More design details are described elsewhere [11,12].

Study Measurements

Dietary Survey

A 3-day consecutive 24-hour-dietary recall method was employed to gather information on the consumption of all foods and drinks. Due to notable variations in the Chinese diet between weekdays and weekends, we recorded the consumption for 2 workdays and 1 weekend day. After training, investigators conducted face-to-face interviews, asking participants about their food and beverage intake over the past 24 hours for 3 consecutive days. Simultaneously, a household weighing method was used to document the consumption of cooking oil and condiments over the corresponding 3 days. The allocation of cooking oil and condiment consumption to individuals was based on how often they cooked at home, meal proportions, and the ratio of individual energy intake to household energy intake. To minimize the recall bias, respondents were asked to write down what they ate and drank every day.

Food items were categorized according to the categories outlined in the China Food Composition Tables. In this study, the term "milk" referred to the equivalent of liquid milk, encompassing various dairy and dairy products converted based on the protein content of every 100 grams of the edible portion. The category "beans" included soybeans, other legumes, and their products. The intakes of soybean flour and soybean milk were converted into the equivalent of soybeans, considering the protein content of every 100 grams of the edible portion. Salt intake consisted of sodium chloride (NaCl) from cooking salt, paste, and sauce.

Anthropometric Measurements

The participants' height, weight, systolic blood pressure (SBP), and diastolic blood pressure (DBP) were measured by trained investigators. Height was recorded to the nearest 0.1 centimeters.

Weight was recorded to the nearest 0.1 kilograms. BMI was calculated as weight (kg)/height^2 (m^2).

Blood Sample Collection

In the years 2009, 2015, and 2018, the CHNS project collected fasting venous blood samples from participants aged 7 years and above. The fasting plasma glucose (FPG) concentration was measured using the GOD-PAP (glucose oxidase-phenol-4-aminoantipyrineenzymatic) method (Randox Laboratories Ltd). Glycated hemoglobin (HbA1c) levels were measured using a high-performance liquid chromatography system (model HLC-723 G7; Tosoh Corp). Participants' diagnosis of diabetes mellitus was confirmed based on doctor diagnoses, as reported by participants themselves or by their parents in cases where adolescents could not answer. Total cholesterol and HDL cholesterol were measured using the CHOD-PAP (cholesterol oxidase: P-aminophenazone) method (Kyowa Medex Co Ltd). Non-HDL cholesterol was calculated as the difference between total cholesterol and HDL cholesterol.

Nicotine Exposure

Nicotine exposure was assessed through active smoking and secondhand smoke exposure. Active smoking was categorized into three groups: (1) individuals who had never smoked, (2) those who had smoked at any point or had quit smoking for more than 30 days, and (3) those who were currently smoking or had quit smoking within the last 30 days. Secondhand smoke exposure was defined as living with people who smoke.

Physical Activity

Physical activity data, comprehensively collected in this project since 2000, were obtained through questionnaires. Physical activity encompassed occupational, domestic, travel, and leisure activities.

Sleep Duration Data

Sleep duration data, including both night sleep time and daytime sleep time, were collected from questionnaires as part of the study.

Quantification of CVH Components

The methods for measuring, evaluating, and quantifying the 8 CVH components are displayed in Tables S1 and S2 of Multimedia Appendix 1. These methods closely aligned with the original LE8 metrics tailored for children and adolescents under the age of 19 years. The 8 components comprised of 4 health behaviors (diet, physical activity, nicotine exposure, and sleep health) and 4 health factors (BMI, non-HDL cholesterol, FPG, and blood pressure), as defined by the LE8 reference. Furthermore, sleep health was evaluated in accordance with the Healthy China Initiative (2019-2030), which recommends a sleep duration of 10 hours for primary school students, 9 hours for secondary school students, and 8 hours for high school students. Age and sex-specific BMI percentiles were generated based on growth references from the World Health Organization [13]. SBP and DBP were transformed into age, sex, and height–specific percentiles, and evaluated based on the American Academy of Pediatrics' Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents, released in 2017 [14].

Statistical Analyses

Each of the LE8 components was scored on a scale ranging from 0 to 100, with the total CVH score calculated as the average of the 8 component scores. Overall CVH scores between 80 and 100 were categorized as high CVH, scores between 50 and 79 were categorized as moderate CVH, and scores between 0 and 49 were classified as low CVH. The mean and SD were calculated for continuous variables, and numbers and percentages were calculated for categorical variables. The median and IQR were reported for non-HDL cholesterol. Geographical regions were categorized as North China, Northeast China, East China, Central China, Southwest China, and Northwest China according to codes for the administrative divisions of the People's Republic of China (GB/*t* 2260), as outlined in Table S3 of Multimedia Appendix 1. Joint point regression models were used to analyze the differences in CVH scores over time, considering rural or urban living residence. Adjustments were made for the children's age and sex as appropriate. Since scores for both health behaviors and health factors decreased sharply after 2004, cross-lagged panel models were applied to explore causal relationships between changes in health behaviors and health factors. Model fit was evaluated based on the Comparative Fit Index and Tucker-Lewis Index scores greater than 0.95, with a root mean squared error of approximation less than 0.05 being indicative of good fit [15]. All statistical analyses were performed using Stata software (version 14.0; StataCorp). The associations were considered significant at 2-sided $P<.05$.

Ethics Approval

This study was approved by the institutional review board of the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health of the Chinese Center for Disease Control and Prevention (2018-004). Participants aged 7 to 17 years provided signed informed consent from themselves and their guardians, while participants aged 18 years and above provided their own signed informed consent forms to participate in the original surveys. The data for our study were completely deidentified and did not contain any information that could be traced back to the individual participants.

Results

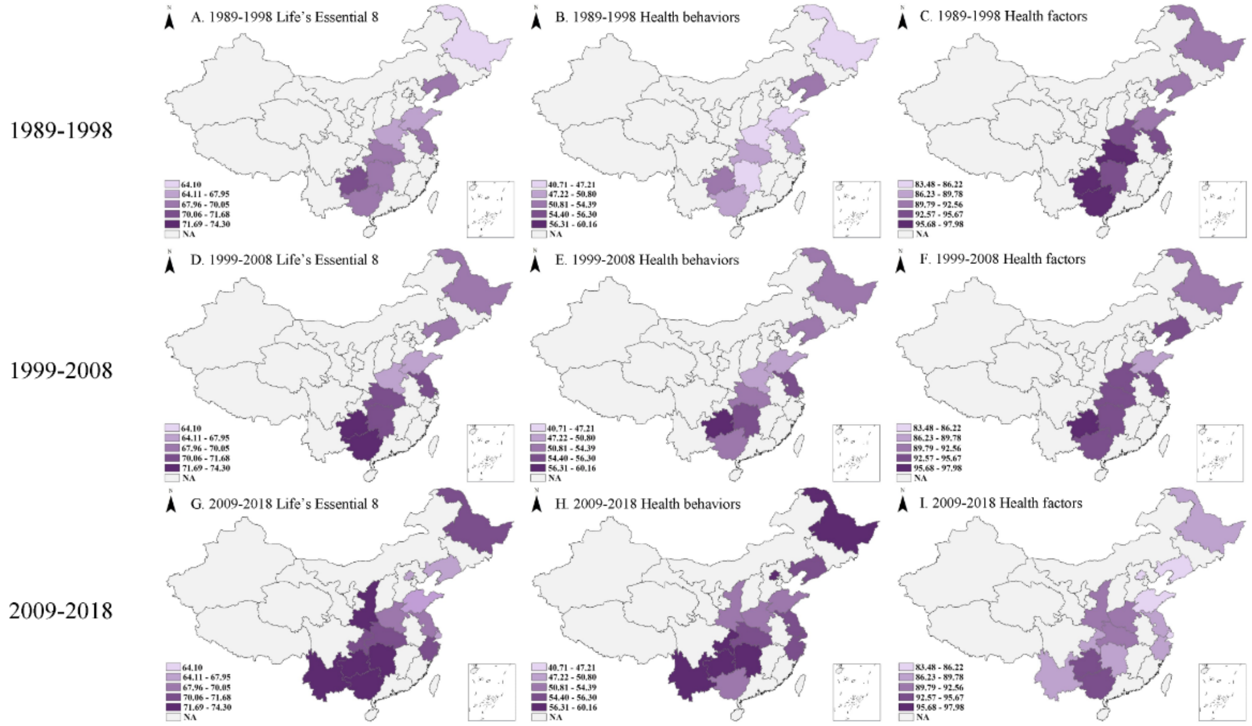
In total, 21,921 (77%) out of 28,477 participants, 52.6% (n=11,537) of them male, aged between 7 and 19 (mean 13, SD 3.6) years, with data for at least 4 CVH components, were included in the study analysis. Among the participants, 6071 (27.7%) were from urban areas. Demographic characteristics, as well as CVH status, are detailed in Table 1 by geographic regions, and comparisons of demographic characteristics between the excluded and included children are displayed in Table S4 of Multimedia Appendix 1. The sample distribution at each research wave is displayed in Table S5 of Multimedia Appendix 1.

Table 1. Demographic characteristics of participants by geographic regions.

Characteristics	Overall	North China (n=403)	Northeast China (n=2746)	East China (n=4271)	Central China (n=10,208)	Southwest China (n=3830)	Northwest China (n=463)	Difference (*P* value)
Urban area, n (%)	6071 (27.7)	321 (79.7)	643 (23.4)	1418 (33.2)	2561 (25.1)	992 (25.9)	136 (29.4)	<.001
Male sex, n (%)	11,537 (52.6)	213 (52.9)	1418 (51.6)	2242 (52.5)	5420 (53.1)	2005 (52.4)	239 (51.6)	.803
Age (years), mean (SD)	13 (3.6)	12.6 (3.5)	13.1 (3.5)	13.1 (3.6)	13 (3.6)	12.9 (3.6)	12.2 (3.2)	<.001
Diet score, mean (SD)	28.4 (9)	28.8 (10.8)	30.1 (10.4)	28.9 (9.4)	28.2 (8.7)	27.2 (7.7)	26.7 (7.4)	<.001
Physical activity, n (%)								<.001
≥420 min	5477 (30.8)	163 (40.7)	615 (26.9)	907 (26.7)	2343 (29.2)	1292 (40.3)	157 (34.7)	
360 to 419 min	701 (4)	15 (3.7)	86 (3.8)	132 (3.9)	318 (4)	133 (4.2)	17 (3.8)	
300 to 359 min	940 (5.3)	19 (4.7)	142 (6.2)	156 (4.6)	395 (4.9)	207 (6.5)	21 (4.6)	
240 to 299 min	1041 (5.9)	25 (6.2)	150 (6.6)	182 (5.4)	443 (5.5)	219 (6.8)	22 (4.9)	
120 to 239 min	3259 (18.3)	73 (18.2)	499 (21.8)	534 (15.7)	1482 (18.5)	596 (18.6)	75 (16.6)	
1 to 119 min	3300 (18.6)	55 (13.7)	446 (19.5)	728 (21.5)	1575 (19.6)	413 (12.9)	83 (18.3)	
Never	3049 (17.2)	51 (12.7)	350 (15.3)	755 (22.3)	1469 (18.3)	346 (10.8)	78 (17.2)	
Cigarette use, n (%)								<.001
Never tried	10356 (94.8)	205 (99.5)	1348 (94.4)	2203 (98.2)	4629 (94.4)	1763 (91.4)	208 (99.5)	
Tried but >30 days ago	13 (0.1)	—[a]	3 (0.2)	1 (0)	5 (0.1)	4 (0.2)	—	
Used within 30 days	552 (5.1)	1 (0.5)	77 (5.4)	39 (1.7)	271 (5.5)	163 (8.5)	1 (0.5)	
Sleep per night (hours), mean (SD)	8.7 (1.2)	8.4 (1.1)	8.5 (1.2)	8.6 (1.2)	8.8 (1.2)	8.9 (1.1)	8.4 (1.1)	<.001
Height (cm), mean (SD)	146.6 (16.8)	153 (17.6)	150.9 (16.5)	149.8 (16.7)	145 (16.5)	143.1 (16.3)	149.5 (16.2)	<.001
Weight (kg), mean (SD)	40.1 (14.4)	48.8 (17.1)	43.5 (14.6)	43.8 (14.7)	38.1 (13.2)	37.6 (15.1)	42.8 (14.1)	<.001
BMI (kg/m^2), mean (SD)	18 (3.2)	20.2 (4.3)	18.6 (3.4)	19 (3.4)	17.5 (3)	17.7 (3)	18.7 (3.6)	<.001
Non-HDL[b] cholesterol (mg/dL), median (IQR)	94.9 (33.3)	104.6 (45.8)	96.2 (30.9)	99.8 (38.3)	95.5 (32.5)	94 (30.9)	84.3 (32.1)	<.001
Fasting plasma glucose (mg/dL), mean (SD)	89.9 (13.9)	91.2 (10.5)	86.9 (17.1)	88.6 (9.9)	91.3 (17.5)	90 (11.5)	90.1 (8.4)	<.001
Systolic blood pressure (mmHg), mean (SD)	99.9 (13.5)	106.4 (11.7)	102.6 (13.9)	102.6 (13.7)	98.9 (13)	96.8 (13.1)	100.4 (13.5)	<.001
Diastolic blood pressure (mmHg), mean (SD)	65.4 (9.8)	68.5 (7.6)	68 (10.4)	66.6 (9.8)	64.7 (9.5)	63.7 (9.6)	66.7 (8.7)	<.001

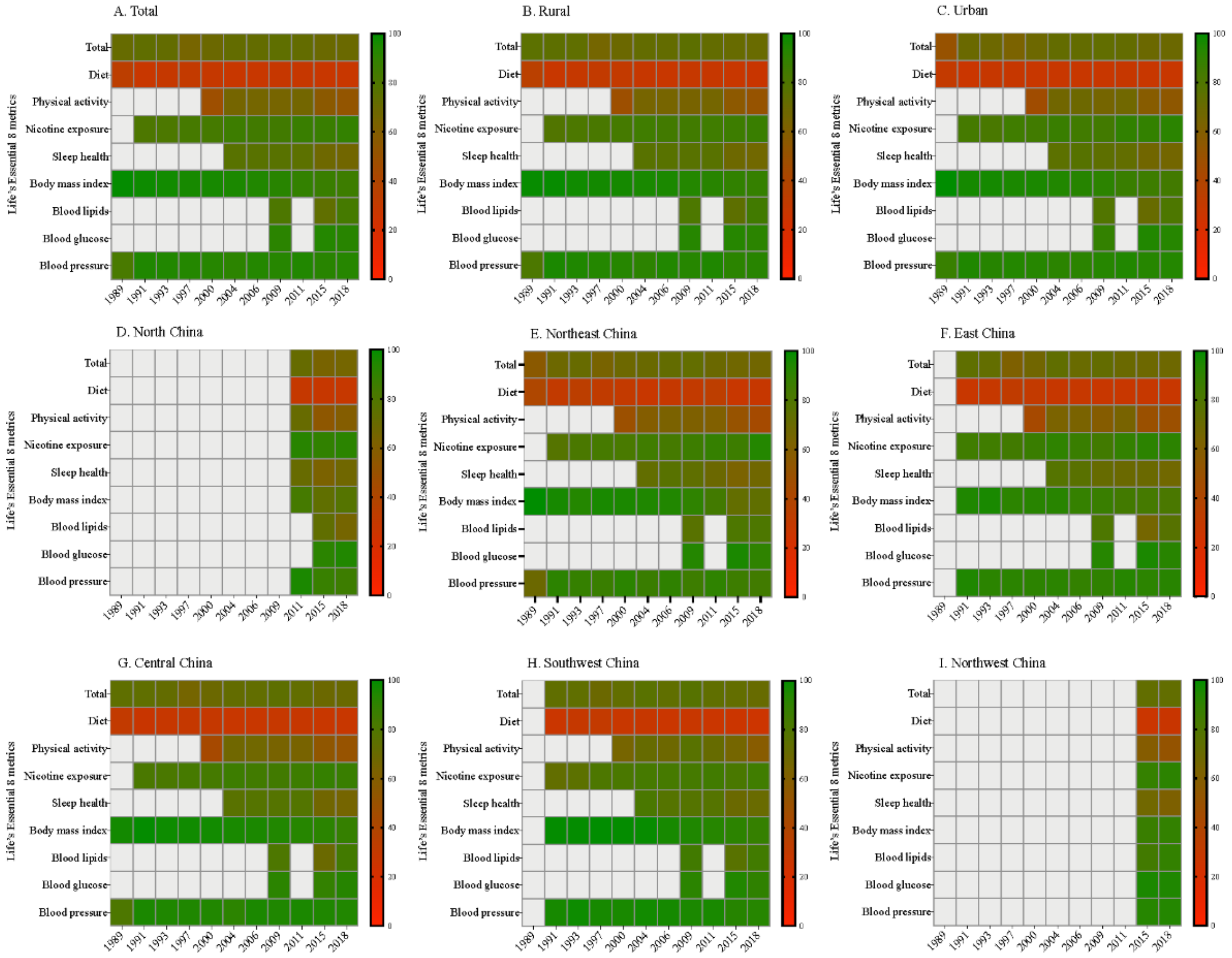
During the past 3 decades, the overall CVH score remained stable in most of the involved provinces, with the lowest score observed in Shandong from East China, with a mean between 67 (SD 10.9) and 67.2 (SD 12.4). The highest score was observed in Guizhou from Southwest China, with a mean between 71.4 (SD 10.8) and 74.3 (SD 10.3). The most significant change in overall CVH status was observed in Heilongjiang from Northeast China, with scores increasing from a mean of 64.1 (SD 11.1) to 71.1 (SD 10.3). Health behavior scores, though much lower than overall scores, showed a gradual increase during the decades. The greatest increase was found in Guizhou from Southwest China and Hunan from Central China, with health behavior scores increasing from a mean of 50.8 (SD 18.1) to 60.2 (SD 14.3) and from 46.8 (SD 17.5) to 56.7 (SD 15.7), respectively. Conversely, health factor scores decreased across all areas, with the biggest decline found in Liaoning from Northeast China (mean 92.3, SD 13.8 to mean 84, SD 17.8) and Shandong from East China (mean 91.6, SD 14.5 to mean 83.6, SD 16.7). Temporal changes in CVH scores by province are shown in Figure 1.

Figure 1. Distribution and temporal change of overall cardiovascular health (CVH) status in Chinese school-aged children and adolescents from 1989 to 2018 based on natural breaks (Jenks) classification. A. Life’s Essential 8 score (1989-1998); B. Health behavior score (1989-1998); C. Health factor score (1989-1998); D. Life’s Essential 8 score (1999-2008); E. Health behavior score (1999-2008); F. Health factor score (1999-2008); G. Life’s Essential 8 score (2009-2018); H. Health behavior score (2009-2018); I. Health factor score (2009-2018).



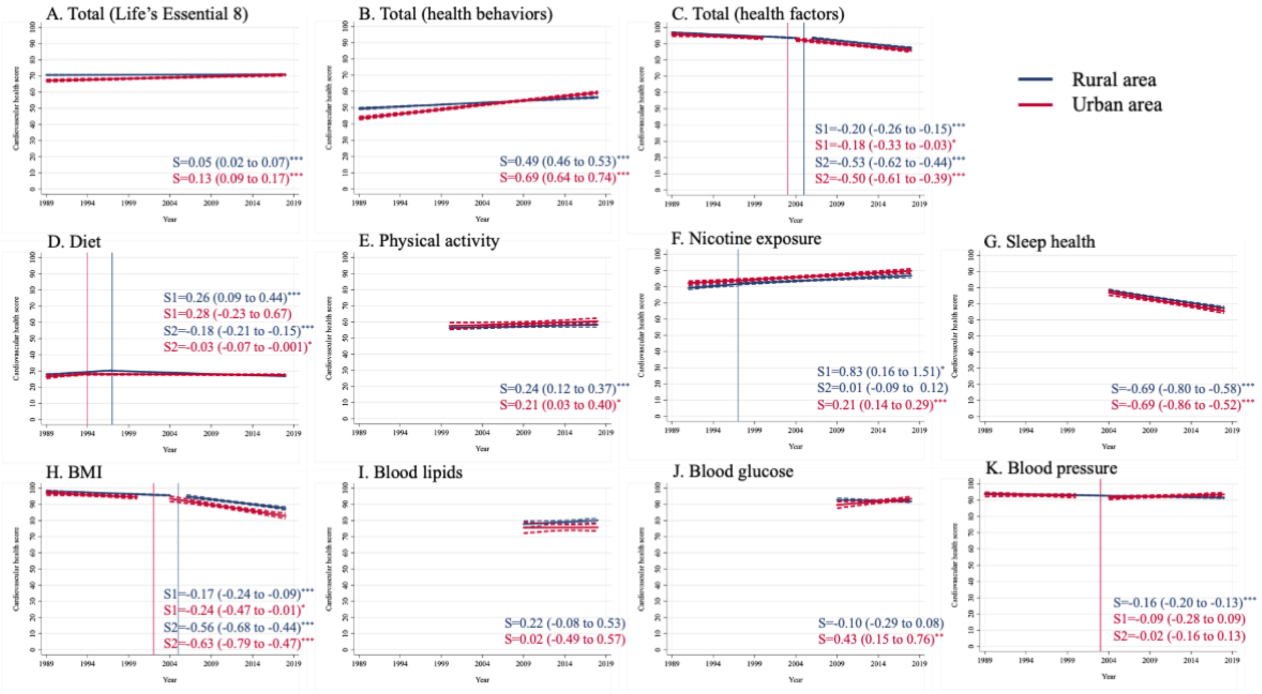
Regarding specific CVH metrics, the healthy diet score consistently remained the lowest for all research waves and in all 6 geographical regions. Worsened sleep health was observed in all geographical regions, while worsened BMI was particularly prominent in North and Northeast China (Figure 2 and Figure S1 in Multimedia Appendix 1).

Figure 2. Changes in cardiovascular health metrics for Chinese school-aged children and adolescents from 1989 to 2018 by geographic region.



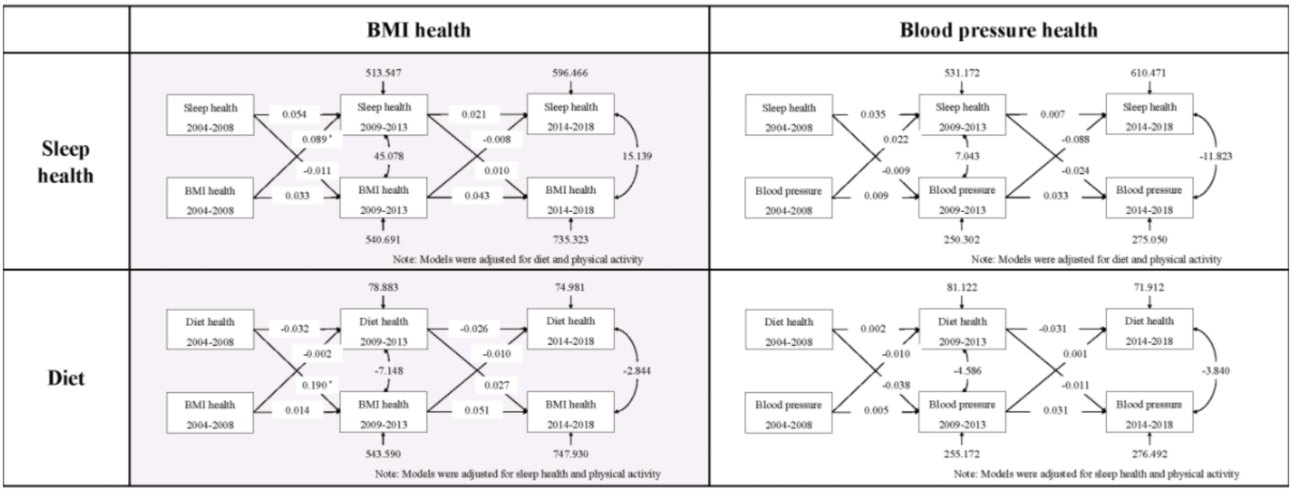
Joint point regression models were then applied to analyze the temporal changes in CVH metrics. Although similar trends were observed in both rural and urban areas, time lags in break points were identified for several metrics (Figure 3). Generally, overall CVH scores and health behavior scores displayed linear growth trends in both rural and urban areas. Total health factor scores decreased faster in rural areas since 2005, with a speed of 0.53 (95% CI 0.44-0.62; *P*<.001) per year, and in urban areas since 2003, with a speed of 0.50 (95% CI 0.39-0.61; *P*<.001). Decreases in diet scores became more significant in rural areas since 1997, with a speed of 0.18 (95% CI 0.15-0.21; *P*<.001) and in urban areas since 1994, with a speed of 0.03 (95% CI 0.001-0.07; *P*=.04). The decrease in BMI scores became more pronounced in rural areas since 2005, with a speed of 0.56 (95% CI 0.44-0.68; *P*<.001) and in urban areas since 2002, with a speed of 0.63 (95% CI 0.47-0.79; *P*<.001). Constant decreases were observed in sleep health scores, with a speed of 0.69 (95% CI 0.58-0.80; *P*<.001) in rural areas and 0.69 (95% CI 0.52-0.86; *P*<.001) in urban areas.

Figure 3. Joint point regression analysis of time trends in cardiovascular health scores in Chinese school-aged children and adolescents from 1989 to 2018. Where a joint point was recognized, S1 referred to the slope before the joint point, and S2 referred to the slope after the joint point. S: slope.



Cross-lagged panel models were constructed between health behaviors (including diet, physical activity, and sleep health) and health factors (including BMI and blood pressure health) during the study waves spanning 2004 to 2008, 2009 to 2013, and 2014 to 2018 (Figure 4). The results indicated that lowered BMI health between 2004 and 2008 resulted in lowered sleep health between 2009 and 2013, with a coefficient of 0.089 (95% CI 0.010-0.168; *P*=.03). Concurrently, lowered diet scores from 2004 to 2008 resulted in lowered BMI health from 2009 to 2013, with a coefficient of 0.190 (95% CI 0.030-0.351; *P*=.02).

Figure 4. Cross-lagged standard regression coefficient of sleep health, diet health, BMI, and blood pressure health between 2009 and 2013 and 2014 and 2018. BP: blood pressure.*Indicated to *P*<.05.



Discussion

Principal Findings

To our knowledge, this is the first study examining CVH status and its temporal changes over the past decades in Chinese school-aged children and adolescents using the new LE8 metrics. From 1989 to 2018, the overall LE8 score remained relatively stable at a moderate level. Simultaneously, the overall score for health behaviors moderately increased, while that of health factors declined significantly over time. Therefore, we inferred that Chinese children and adolescents may face a prolonged downward trend in overall CVH status in the future.

The overall level of CVH scores in Chinese children in the recent decade was compatible with those of children in the United States, ranging between 64.6 and 74.1, depending on the metrics used for calculation [16]. Both groups shared suboptimal performance in diet, but Chinese children and adolescents had relatively lower scores in health behaviors such as physical activity and sleep health. The initiation of CVD and related diseases in childhood underscores the critical importance of LE8 metrics, turning childhood and youth into new frontiers for CVD prevention [17]. As the behavioral change theory has emphasized, changes in human behavior should be based on education, enhanced self-motivation, and adequate external support, eventually resulting in changes in health outcomes [18,19]. Early and sustained multicomponent educational programs that focus on health promotion in children and adolescents could represent a critical window of opportunity to potentially prevent disease in later life [20].

The results of our analysis indicated that the overall CVH level was lower in the eastern part of China compared to the western part, similar to findings reported for Chinese adults from the 2021 Report on Cardiovascular Health and Disease in China [21,22]. In contrast with reports from countries in North America, where lower CVH status is typically associated with less favorable economic conditions [23], we found poorer CVH status in areas with better economic conditions in China. This may be attributed to wider socioeconomic gaps and the prevalence of unhealthy lifestyles in urban and higher-income areas [24]. However, further research is needed for a more detailed exploration of the social determinants of CVH.

Several metrics in the LE8 were particularly noteworthy in this study, highlighting persistently low diet quality, fluctuating levels of physical activity, and persistent declines in sleep and BMI health. Of more concern is that although there was a time lag between urban and rural areas, the trend and speed of the simultaneous decline in rural areas for these indicators were comparable—or even greater than—that of urban areas.

Enhancing physical fitness in school-aged children and adolescents has been a top priority for the Chinese government. The State Council of the People's Republic of China's initiatives "Suggestions on Strengthening Youth Sports and Enhancing Youth Physical Fitness," released in 2007 [25], and the "Sunshine Sports Program," which encourages all school-aged children and adolescents to have at least 1 hour of outdoor physical activity, have been implemented. Nationwide survey findings from both our analysis and the Chinese National Survey on Students' Constitution and Health showed a significant improvement in weekly exercise hours and physical fitness during the following years [26,27]. However, our analysis reveals that physical activity scores have continued to decline since 2011, primarily due to lower scores in Northwest China. The geographical imbalance in policy implementation is evident and should be addressed in future efforts.

Sleep health has drawn attention in some high-income countries, and public health initiatives have been implemented to promote better sleep health [28,29]. In the United States, national surveys indicate an improvement in sleep duration due to earlier bedtimes and longer sleep durations in various segments of the population [30]. Insufficient sleep can directly impact growth and development in children and adolescents, acting as a medium to spread the adverse effects of modern social development on CVH through its impact on the endocrine system [31-34].

A population-based analysis found a causal cycle between poor diet, worsened BMI, and shortened sleep time, as all 3 metrics have experienced significant decreases during the latest decade. This suggests that overall CVH in Chinese children and adolescents could soon worsen. The LE8 guideline emphasizes that overall CVH status depends not on a single metric but on a combination of all 8 metrics. Emphasizing certain aspects when policymaking could lead to joint improvements in other metrics, making it crucial to avoid geographical biases in policy implementations.

Our study results indicate that it might be cost-effective and quicker to prioritize diet health interventions for CVH promotion, as diet could impact other CVH metrics both directly and indirectly. Moreover, comprehensive intervention strategies tailored to specific regional conditions are needed for further CVH promotion, emphasizing the need for developing evidence-based behavior change techniques [35].

It is important to note that our study only reflects the temporal change and distribution of CVH status in school-aged children and adolescents until 2018. Since then, a series of government policies, including the "Healthy China Initiative (2019-2030)," "Implementation Plan for Obesity Prevention and Control in Children and Adolescents," and "Implementation Plan for Comprehensive Prevention and Control of Myopia in Children and Adolescents," have been released to guide health promotion programs targeting Chinese children and adolescents. Key factors such as the prevalence of obesity and myopia were listed as part of assessment indicators to ensure that local governments attach enough importance to them. Dietary health was also addressed in nearly all the aforementioned policies, with related health education packages to aid in implementation. With such a strong driving force, we anticipate significant improvements in CVH and overall health status for Chinese school-aged children and adolescents by 2030, setting a valuable example for other countries.

Limitations

This study has several limitations. First, although the study data originated from one of China's largest and longest-standing nationwide studies of child and adolescent health status, there is an inherent bias in sample selection. Approximately 23% (n=6556) of the participants from the original data set were excluded due to missing data (mainly from blood samples or surveys before 2000). Although the excluded participants were not likely to have influenced the findings of this study, which focuses on CVH changes in the past 20 years, we are unable to ascertain if they would influence the overall CVH scores and their trends before 2000. Nevertheless, the original CHNS sample was not designed to be representative of China but rather to capture a diverse range of economic and demographic circumstances [12] Therefore, caution should be taken when generalizing these findings. Despite these limitations, based on the setting of the original survey, our study results still provide valuable insights into the changing trends of CVH status among children and adolescents from different socioeconomic development levels in China.

Second, although the original CHNS was designed to be a prospective cohort study, our study adopted a multiwave cross-sectional study design due to the lack of personal information. Therefore, the analyses were conducted at the population level, and further studies with individual-level data could provide confirmation of the reported population-level data. Third, concerns about the reliability of the self-reported dietary data have been raised in the scientific community, and we observed lower reported energy intakes in many Chinese children in our previous studies. However, given the complexity of the Chinese diet, objective dietary survey methods are still being developed.

Conclusions

Based on the latest LE8 metrics, Chinese school-aged children and adolescents generally exhibited a moderate level of CVH over the past decades. Poor performance in health behaviors, including diet, sleep, and physical activity may contribute to worsening overall CVH in the future. Therefore, policies and interventions addressing diet health may be the breakthrough point for promoting overall CVH status in Chinese children and adolescents.

Acknowledgments

The China Health and Nutrition Survey (CHNS) survey was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01-HD30880 and R01-HD38700); National Institute on Aging (R01-AG065357); National Heart, Lung, and Blood Institute (R01- HL108427I); National Institutes of Health (NIH) Fogarty International Center (D43-TW009077); Carolina Population Center (P2C-HD050924 and P30-AG066615) from 1989 to 2019; and the Ministry of Finance of the People’s Republic of China. This work was supported by the National Natural Science Foundation (81673192 to JM, 81903344 to BD and 82204067 to XW) and Postdoctoral Research Specialization of Vanke School of Public Health, Tsinghua University (2022BH011 to XW).

The authors greatly appreciated the colleagues and primary health care staff from the CHNS investigation groups for their hard work on data collection. We would also like to express our gratitude to Professor Zhiqiang Wang, who provided generous support when drafting this manuscript.

Data Availability

Partial data from 1989 to 2011 are available upon request to the corresponding author. The remaining data are not publicly available.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File , 119 KB - publichealth_v9i1e45564_app1.docx]

References

1. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, American Heart Association Council on Epidemiology Prevention Statistics Committee Stroke Statistics Subcommittee. Heart disease and stroke statistics-2018 update: A report from the American Heart Association. Circulation 2018 Mar 20;137(12):e67-e492 [FREE Full text] [doi: 10.1161/CIR.0000000000000558] [Medline: 29386200]
2. Reddy KS, Mathur MR. Global burden of CVD. In: Handbook of Global Health. Cham, Switzerland: Springer; May 12, 2021.
3. Pool LR, Krefman AE, Labarthe DR, Greenland P, Juonala M, Kähönen M, et al. The timing and sequence of cardiovascular health decline. Am J Prev Med 2021 Oct;61(4):545-553. [doi: 10.1016/j.amepre.2021.04.010] [Medline: 34238623]
4. Bei Y, Yang T, Xiao J. Cardiovascular medicine in China: what can we do to achieve the Healthy China 2030 plan? BMC Med 2018 Aug 24;16(1):132 [FREE Full text] [doi: 10.1186/s12916-018-1133-4] [Medline: 30139358]
5. Ma L, Chen W, Gao R, Liu LS, Zhu ML, Wang YJ, et al. China cardiovascular diseases report 2018: an updated summary. J Geriatr Cardiol 2020 Jan;17(1):1-8 [FREE Full text] [doi: 10.11909/j.issn.1671-5411.2020.01.001] [Medline: 32133031]
6. Hyattsville (MD). Health, United States, 2017: With Special Feature on Mortality. In: National Center for Health Statistics. US: National Center for Health Statistics; 2018.
7. Jagannathan R, Patel SA, Ali MK, Narayan KMV. Global updates on cardiovascular disease mortality trends and attribution of traditional risk factors. Curr Diab Rep 2019 Jun 20;19(7):44. [doi: 10.1007/s11892-019-1161-2] [Medline: 31222515]
8. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van HL, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic impact goal through 2020 and beyond. Circulation 2010 Feb 2;121(4):586-613 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192703] [Medline: 20089546]

9. Brunner FJ, Waldeyer C, Ojeda F, Salomaa V, Kee F, Sans S, et al. Application of non-HDL cholesterol for population-based cardiovascular risk stratification: results from the Multinational Cardiovascular Risk Consortium. Lancet 2019 Dec;394(10215):2173-2183. [doi: 10.1016/s0140-6736(19)32519-x]
10. Juonala M, Wu F, Sinaiko A, Woo JG, Urbina EM, Jacobs D, et al. Non-HDL cholesterol levels in childhood and carotid intima-media thickness in adulthood. Pediatrics 2020 Apr;145(4):1-9 [FREE Full text] [doi: 10.1542/peds.2019-2114] [Medline: 32209701]
11. Zhang B, Zhai FY, Du SF, Popkin BM. The China Health and Nutrition Survey, 1989-2011. Obes Rev 2014 Jan;15 Suppl 1(0 1):2-7 [FREE Full text] [doi: 10.1111/obr.12119] [Medline: 24341753]
12. Popkin BM, Du S, Zhai F, Zhang B. Cohort Profile: The China Health and Nutrition Survey--monitoring and understanding socio-economic and health change in China, 1989-2011. Int J Epidemiol 2010 Dec;39(6):1435-1440 [FREE Full text] [doi: 10.1093/ije/dyp322] [Medline: 19887509]
13. de Onis M. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ 2007 Sep 01;85(09):660-667. [doi: 10.2471/blt.07.043497]
14. Flynn JT, Kaelber DC, Baker-Smith CM, Blowey D, Carroll AE, Daniels SR, Subcommittee on Screening and Management of High Blood Pressure in Children. Clinical practice guideline for screening and management of high blood pressure in children and adolescents. Pediatrics 2017 Sep;140(3). [doi: 10.1542/peds.2017-1904] [Medline: 28827377]
15. Ehrminger M, Brunet-Gouet E, Cannavo A, Aouizerate B, Cussac I, Azorin J, FondaMental Advanced Centers of Expertise in Bipolar Disorders (FACE-BD) Collaborators, et al. Longitudinal relationships between cognition and functioning over 2 years in euthymic patients with bipolar disorder: a cross-lagged panel model approach with the FACE-BD cohort. Br J Psychiatry 2021 Feb;218(2):80-87. [doi: 10.1192/bjp.2019.180] [Medline: 31407639]
16. Lloyd-Jones DM, Ning H, Labarthe D, Brewer L, Sharma G, Rosamond W, et al. Status of cardiovascular health in US adults and children using the American Heart Association’s new “Life’s Essential 8” Metrics: prevalence estimates from the National Health and Nutrition Examination Survey (NHANES), 2013 through 2018. Circulation 2022 Sep 13;146(11):822-835. [doi: 10.1161/circulationaha.122.060911]
17. Abrignani MG, Lucà F, Favilli S, Benvenuto M, Rao CM, Di Fusco SA, Cardiovascular Prevention Area, Young Cardiologists Area, Paediatric Cardiology Task Force of the Associazione Nazionale Medici Cardiologi Ospedalieri (ANMCO), Heart Care Foundation. Lifestyles and cardiovascular prevention in childhood and adolescence. Pediatr Cardiol 2019 Aug 24;40(6):1113-1125. [doi: 10.1007/s00246-019-02152-w] [Medline: 31342115]
18. Gardner B, Rebar A. Habit formation and behavior change. Oxford Res Encycl Psychol 2019:26. [doi: 10.1093/acrefore/9780190236557.013.129]
19. Beauchamp MR, Crawford KL, Jackson B. Social cognitive theory and physical activity: mechanisms of behavior change, critique, and legacy. Psychol Sport Exerc 2019 May;42:110-117. [doi: 10.1016/j.psychsport.2018.11.009] [Medline: 30872122]
20. Fernandez-Jimenez R, Al-Kazaz M, Jaslow R, Carvajal I, Fuster V. Children present a window of opportunity for promoting health: JACC review topic of the week. J Am Coll Cardiol 2018 Dec 25;72(25):3310-3319 [FREE Full text] [doi: 10.1016/j.jacc.2018.10.031] [Medline: 30527619]
21. Ma L, Wang Z, Fan J, Hu S. Epidemiology and management of hypertension in China: an analysis using data from the Annual Report on Cardiovascular Health and Diseases in China (2021). Chin Gen Pract 2022;25(30):3715-3720. [doi: 10.12114/j.issn.1007-9572.2022.0502]
22. Hu SS, The Writing Committee Of The Report On Cardiovascular Health And Diseases in China. Report on cardiovascular health and diseases in China 2021: an updated summary. J Geriatr Cardiol 2023 Jun 28;20(6):399-430 [FREE Full text] [doi: 10.26599/1671-5411.2023.06.001] [Medline: 37416519]
23. Movsisyan N, Vinciguerra M, Medina-Inojosa J, Lopez-Jimenez F. Cardiovascular diseases in Central and Eastern Europe: a call for more surveillance and evidence-based health promotion. Ann Glob Health 2020 Feb 26;86(1):21 [FREE Full text] [doi: 10.5334/aogh.2713] [Medline: 32166066]
24. Patel N, Kalra R, Bhargava A, Arora G, Arora P. Ideal cardiovascular health among American adults after the economic recession of 2008-2009: insights from NHANES. Am J Med 2019 Oct;132(10):1182-1190.e5 [FREE Full text] [doi: 10.1016/j.amjmed.2019.06.004] [Medline: 31278932]
25. Suggestions on strengthening youth sports and enhancing youth physical fitness. State Council of the People's Republic of China. 2007. URL: http://www.gov.cn/gongbao/content/2007/content_663655.htm [accessed 2022-08-18]
26. Dong Y, Lau PWC, Dong B, Zou Z, Yang Y, Wen B, et al. Trends in physical fitness, growth, and nutritional status of Chinese children and adolescents: a retrospective analysis of 1·5 million students from six successive national surveys between 1985 and 2014. Lancet Child Adolesc Health 2019 Dec;3(12):871-880. [doi: 10.1016/s2352-4642(19)30302-5]
27. Department of Physical Health and Arts Education, Ministry of Education. Release report of the Eighth National Survey on Student Physical Fitness and Health. Chin J Sch Health 2021;42(9):1281-1282. [doi: 10.16835/j.cnki.1000-9817.2021.09.001]
28. Chaput J. The integration of pediatric sleep health into public health in Canada. Sleep Med 2019 Apr;56:4-8. [doi: 10.1016/j.sleep.2018.06.009] [Medline: 30076116]

29. Hale L, Troxel W, Buysse DJ. Sleep health: an opportunity for public health to address health equity. Annu Rev Public Health 2020 Apr 02;41(1):81-99 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094412] [Medline: 31900098]
30. Basner M, Dinges D. Sleep duration in the United States 2003-2016: first signs of success in the fight against sleep deficiency? Sleep 2018 Apr 01;41(4):1-16. [doi: 10.1093/sleep/zsy012] [Medline: 29325164]
31. St-Onge M, Grandner MA, Brown D, Conroy MB, Jean-Louis G, Coons M, et al. Sleep duration and quality: impact on lifestyle behaviors and cardiometabolic health: a scientific statement from the American Heart Association. Circulation 2016 Nov;134(18). [doi: 10.1161/cir.0000000000000444]
32. Serin Y, Acar Tek N. Effect of circadian rhythm on metabolic processes and the regulation of energy balance. Ann Nutr Metab 2019 Apr 23;74(4):322-330 [FREE Full text] [doi: 10.1159/000500071] [Medline: 31013492]
33. Al-Rawi N, Madkour M, Jahrami H, Salahat D, Alhasan F, BaHammam A, et al. Effect of diurnal intermittent fasting during Ramadan on ghrelin, leptin, melatonin, and cortisol levels among overweight and obese subjects: A prospective observational study. PLoS One 2020 Aug 26;15(8):e0237922 [FREE Full text] [doi: 10.1371/journal.pone.0237922] [Medline: 32845924]
34. Xu Y, Yu Y, Huang Y, Wan Y, Su P, Tao F, et al. Exposure to bedroom light pollution and cardiometabolic risk: A cohort study from Chinese young adults. Environ Pollut 2022 Feb 01;294:118628. [doi: 10.1016/j.envpol.2021.118628] [Medline: 34883146]
35. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Transl Behav Med 2018 Dec 01;8(2):212-224 [FREE Full text] [doi: 10.1093/tbm/ibx019] [Medline: 29381786]

Abbreviations

CHNS: China Health and Nutrition Survey
CHOD-PAP: cholesterol oxidase: P-aminophenazone
CVD: cardiovascular disease
CVH: cardiovascular health
DBP: diastolic blood pressure
FPG: fasting plasma glucose
GOD-PAP: glucose oxidase-phenol-4-aminoantipyrineenzymatic
HbA1c: glycated hemoglobin
HDL: high-density lipoprotein
LE8: Life’s Essential 8
NaCl: sodium chloride
SBP: systolic blood pressure

Edited by A Mavragani, T Sanchez; submitted 12.01.23; peer-reviewed by TC Chan, Y Jianrong; comments to author 18.05.23; revised version received 02.06.23; accepted 24.08.23; published 23.10.23.

Please cite as:
Wang X, Dong B, Huang F, Zhang J, He R, Du S, Zhang J, Ma J, Wang H, Zhang B, Liang W
Temporal Trends in Cardiovascular Health Status Among Chinese School-Aged Children From 1989 to 2018: Multiwave Cross-Sectional Analysis
JMIR Public Health Surveill 2023;9:e45564
URL: https://publichealth.jmir.org/2023/1/e45564
doi:10.2196/45564
PMID:37870895

Original Paper

Associations of Country-Specific and Sociodemographic Factors With Self-Reported COVID-19–Related Symptoms: Multivariable Analysis of Data From the CoronaCheck Mobile Health Platform

Elke Humer[1], PD, PhD; Thomas Keil[2,3,4], Prof Dr; Carolin Stupp[2,3]; Winfried Schlee[5,6], Prof Dr; Manfred Wildner[3,7], Prof Dr; Peter Heuschmann[2,8], Prof Dr; Michael Winter[2], Dr rer nat; Thomas Probst[1*], Prof Dr; Rüdiger Pryss[2*], Prof Dr

[1]Department for Psychosomatic Medicine and Psychotherapy, University for Continuing Education Krems, Krems, Austria
[2]Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany
[3]State Institute of Health, Bavarian Health and Food Safety Authority, Erlangen, Germany
[4]Institute of Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, Berlin, Germany
[5]Department of Psychiatry and Psychotherapy, University of Regensburg, Regensburg, Germany
[6]Eastern Switzerland University of Applied Sciences, St Gallen, Switzerland
[7]Pettenkofer School of Public Health, University of Munich, Munich, Germany
[8]Clinical Trial Center Würzburg, University Hospital Würzburg, Würzburg, Germany
[*]these authors contributed equally

Corresponding Author:
Rüdiger Pryss, Prof Dr
Institute of Clinical Epidemiology and Biometry
University of Würzburg
Am Schwarzenberg Haus A15
Würzburg, 97078
Germany
Phone: 49 931201 ext 46471
Email: pryss_r@ukw.de

Abstract

Background: The COVID-19 symptom-monitoring apps provide direct feedback to users about the suspected risk of infection with SARS-CoV-2 and advice on how to proceed to prevent the spread of the virus. We have developed the CoronaCheck mobile health (mHealth) platform, the first free app that provides easy access to valid information about the risk of infection with SARS-CoV-2 in English and German. Previous studies have suggested that the clinical characteristics of individuals infected with SARS-CoV-2 vary by age, gender, and viral variant; however, potential differences between countries have not been adequately studied.

Objective: The aim of this study is to describe the characteristics of the users of the CoronaCheck mHealth platform and to determine country-specific and sociodemographic associations of COVID-19–related symptoms and previous contacts with individuals infected with COVID-19.

Methods: Between April 8, 2020, and February 3, 2022, data on sociodemographic characteristics, symptoms, and reports of previous close contacts with individuals infected with COVID-19 were collected from CoronaCheck users in different countries. Multivariable logistic regression analyses were performed to examine whether self-reports of COVID-19–related symptoms and recent contact with a person infected with COVID-19 differed between countries (Germany, India, South Africa), gender identities, age groups, education, and calendar year.

Results: Most app users (N=23,179) were from Germany (n=8116, 35.0%), India (n=6622, 28.6%), and South Africa (n=3705, 16.0%). Most data were collected in 2020 (n=19,723, 85.1%). In addition, 64% (n=14,842) of the users were male, 52.1% (n=12,077) were ≥30 years old, and 38.6% (n=8953) had an education level of more than 11 years of schooling. Headache, muscle pain, fever, loss of smell, loss of taste, and previous contacts with individuals infected with COVID-19 were reported more frequently by users in India (adjusted odds ratios [aORs] 1.3-8.3, 95% CI 1.2-9.2) and South Africa (aORs 1.1-2.6, 95% CI 1.0-3.0) than those in Germany. Cough, general weakness, sore throat, and shortness of breath were more frequently reported in

India (aORs 1.3-2.6, 95% CI 1.2-2.9) compared to Germany. Gender-diverse users reported symptoms and contacts with confirmed COVID-19 cases more often compared to male users.

Conclusions: Patterns of self-reported COVID-19–related symptoms and awareness of a previous contact with individuals infected with COVID-19 seemed to differ between India, South Africa, and Germany, as well as by gender identity in these countries. Viral symptom–collecting apps, such as the CoronaCheck mHealth platform, may be promising tools for pandemics to support appropriate assessments. Future mHealth research on country-specific differences during a pandemic should aim to recruit representative samples.

(JMIR Public Health Surveill 2023;9:e40958) doi:10.2196/40958

KEYWORDS

COVID-19; COVID-19 symptoms; gender; India; South Africa; Germany; symptoms; app; information; English; sociodemographic; weakness; muscle pain; pain; age; education

Introduction

The emergence of COVID-19, caused by SARS-CoV-2 infection, has led to a major ongoing public health crisis worldwide [1]. By April 2022, the COVID-19 pandemic had been associated with more than 6 million deaths out of more than 500 million confirmed cases worldwide [2]. To contain the spread of the virus, various prevention strategies, such as lockdowns, quarantine, and stringent hygiene practices, have been adopted in countries worldwide [3]. The transmission of SARS-CoV-2 occurs primarily through contact with infected individuals via respiratory droplets and aerosols [4,5]. Symptoms vary widely, with fever, cough, and fatigue among the most common [6,7], while loss of smell and loss of taste were described as the most prominent COVID-19 symptoms until 2021, before the Omicron variant became prevalent [7-9]. Symptoms may appear after an incubation period of 2-18 days after contact with the virus [3,10]. Viral shedding may occur 3 days before the onset of symptoms, and more than 50% of SARS-CoV-2 transmissions are reported to occur in asymptomatic infected individuals [11]. Therefore, early identification and isolation of individuals infected with COVID-19 is particularly important to reduce transmission [12,13]. To effectively prevent and manage COVID-19 cases (eg, by detecting clusters) at the population level, various mobile health (mHealth) apps have been developed [3,14]. Symptom-monitoring apps aim to determine whether the user may be infected with SARS-CoV-2 by asking a series of screening questions that include symptoms, such as fever, cough, and pain, and by documenting potential contacts with infected individuals. These apps aim to provide the user with information about the potential risk of being infected and to give advice about further behavior, such as avoiding physical contact with others or wearing a face mask to reduce the spread [3]. Some evidence exists suggesting that the reported clinical characteristics of individuals infected with SARS-CoV-2 vary by age, gender, variant of the virus, and type of reporting system (eg, self-reports vs clinician reported) [7,11,15-17]; however, possible differences between countries have not yet been sufficiently investigated. Since health systems operate differently worldwide (eg, centralized vs decentralized) and have responded differently to the pandemic itself, such country-specific comparisons are important [18]. Official COVID-19 data often do not reflect differences between countries, because they are subject to various uncertainties, such as small numbers of COVID-19 tests and shortcomings in the monitoring systems of countries with limited resources. Therefore, studies comparing cross-national cases run the risk of arriving at incorrect conclusions [19].

In the beginning, the CoronaCheck app was intended to enable symptom monitoring to offer quick advice to those affected. This advice needed to follow the official recommendations of the Robert Koch Institute as the National German Public Health Institute so that those affected could assess their symptoms to the best of the current knowledge and to complement the official support hotlines. Another goal was to provide certain segments of the population with easier access to information than through a telephone hotline. These included adolescents, people with hearing problems who may not fully understand the telephone voice, and migrants or tourists without sufficient language skills for a hotline in the local language.

CoronaCheck is an open-science mHealth platform with 109,603 installations and 88,537 completed questionnaires as of February 3, 2022. The app was developed in the first months of the COVID-19 pandemic in Germany in collaboration with university partners from Bavaria and software companies. The app is based on the TrackYourHealth platform [20,21] and is guided by the information and official recommendations of health authorities. The goal is to provide users with quick and easy-to-perform symptom screening and documentation of contact information to identify their risk of infection with SARS-CoV-2 [22]. CoronaCheck adheres to the Medical Device Regulations (MDR) and has been released on the official app stores of Google (April 30, 2020, 8:25 p.m.) and Apple (April 24, 2020, 12:47 p.m.). The app collects anonymous information about users' sociodemographic characteristics, COVID-19–related symptoms, and recent close contact with an infected individual to provide direct feedback to users. This includes recommendations on how individuals should act, such as seeking medical advice, avoiding contact with others, and taking other protective measures. This study aims to describe the CoronaCheck user characteristics based on data from Germany, India, and South Africa and to determine associations of self-reports of COVID-19–related symptoms and previous contacts with individuals infected with COVID-19 with country, sociodemographic characteristics (gender identity, age group, educational level), calendar year, and user status (reported for oneself vs reported for another person).

Methods

Ethical Considerations

The study was approved by the Ethics Committee of the University of Würzburg (ethical approval no. 71/20-me) and the university's data protection officer and was carried out in accordance with the General Data Protection Regulations of the European Union [23]. To have their anonymous data included in this study, all app users had to provide informed consent and agree that the data can be used for research purposes. With regard to data protection, users were also given comprehensive information at the beginning, and this information can also be viewed at any time [24].

Experimental Design

CoronaCheck was developed in a scientific collaboration between 2 German university partners from Bavaria and software companies and is based on the openly available information about SARS-CoV-2 and corresponding recommendations of the Robert Koch Institute, the German national health authority [22]. CoronaCheck combines the ideas of patient-reported outcomes and mobile sensing with direct feedback from the app to users based on their answers to the questions. The overall goal is to provide a quick test that can be easily performed at any time and with any change in symptoms. In developing the app, it was important to comply with the MDR. The app was released in the official app stores of Google and Apple and complies with the MDR.

On April 8, 2020 (10 weeks after the first COVID-19 cases were reported in Germany), data collection from this app began in Germany, followed by the first data from South Africa and India on April 24 and 28, 2020, respectively.

Ascertainment of Demographic Information

We collected data on age (years, categorized in 10-year intervals up to 79 years, ie, ≤9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, ≥80 years), gender identity (male, female, diverse, not specified), educational level (defined by years of schooling in 4 categories: ≤9 years, 10-11 years, ≥12 years, and "not reported" if users did not answer this question), and the information for whom the questionnaire was filled out (oneself, another person, not specified).

Ascertainment of COVID-19–Associated Symptoms

Participants were asked about the presence of 11 different COVID-19–related symptoms in the past 24 hours (yes/no): fever (defined as a temperature of 37.5 °C or more), sore throat, stuffed or runny nose, cough, loss of smell, loss of taste, shortness of breath, headache, muscle or joint pain, diarrhea, and general weakness. In the case of diarrhea, the symptom was assessed only in the first 2 months (April-June 2020), before it was deleted in accordance with the recommendations of the Robert Koch Institute.

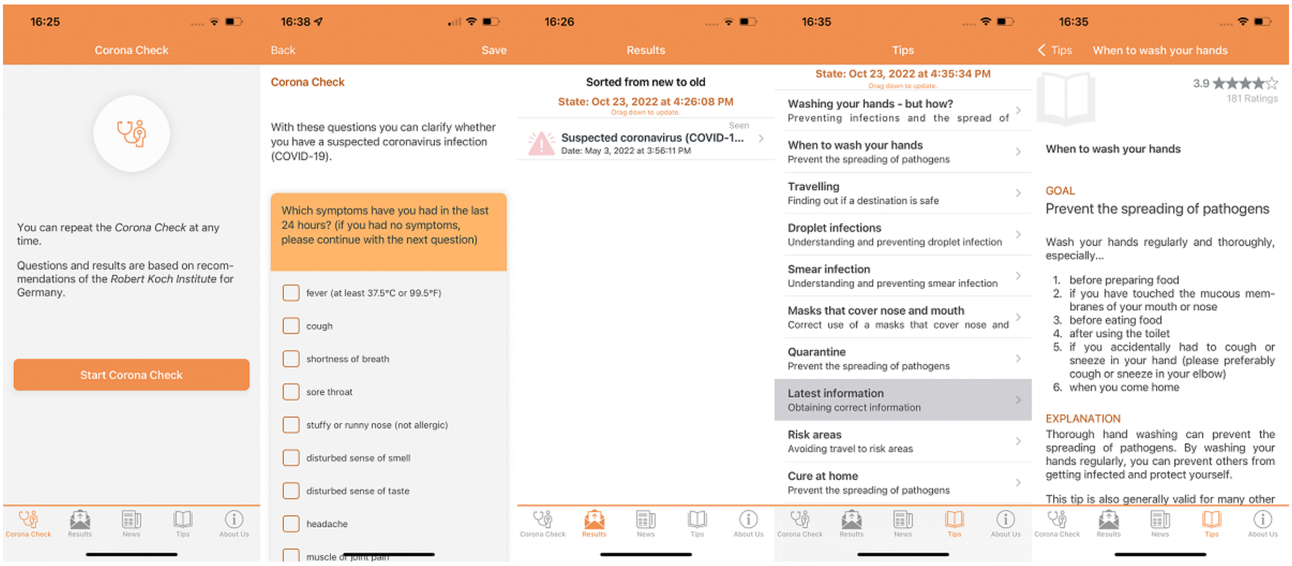
Assessment of Close Contact With a Confirmed Case of COVID-19

We asked participants whether they or the person for whom they completed the questionnaire had had close contact with a person with confirmed COVID-19 in the past 14 days (yes, no, or not specified). We did not ask whether this confirmed case was based on a polymerase chain reaction (PCR) test, a medical diagnosis (including clinical radiology tests), or other information.

App Development and Illustrations

The app was developed based on the TrackYourHealth platform [20,21]. This always consists of a native iOS (Objective C) and a native Android (Java) app as well as a server backend. The latter includes a relational database and interface for mobile apps. The interface is based on the representational state transfer (REST) architectural style, and the connection to the apps is carried out via an encrypted secure sockets layer (SSL) channel [25]. The apps include a questionnaire for the COVID-19 self-check and a module that calculates the feedback for the symptoms. The module consists of the server-side calculation and a component in the apps that displays the feedback. Furthermore, 2 other modules are enabled in the app, one that displays general news located on the server and one that provides tips about COVID-19. The tips are intended to make everyday life easier and can also be evaluated by users in terms of their usefulness. Finally, the app provides information about the team, privacy, and declarations of conformity. The server for the data acquisition is located at the Service Center Medical Informatics at the University Hospital in Würzburg (SMI), and the app has been published in the official stores of Apple and Google. The last thing to mention is that all information (ie, the questionnaire, feedbacks, tips, news) is cached on the mobile devices, but server-side changes appear as an update in the app. Figure 1 shows screenshots of the app. The first screenshot shows the screen to start the test, the second the test in action, the third the feedback screen, the fourth the list of tips, and the fifth a detail page for a tip.

Figure 1. App illustration: screenshots of (1) Start Check, (2) Questionnaire, (3) Feedback Module, (4) Tips, and (5) Detailed Tip.



You can repeat the *Corona Check* at any time.

Questions and results are based on recommendations of the *Robert Koch Institute* for Germany.

With these questions you can clarify whether you have a suspected coronavirus infection (COVID-19).

Which symptoms have you had in the last 24 hours? (if you had no symptoms, please continue with the next question)

- ☐ fever (at least 37.5°C or 99.5°F)
- ☐ cough
- ☐ shortness of breath
- ☐ sore throat
- ☐ stuffy or runny nose (not allergic)
- ☐ disturbed sense of smell
- ☐ disturbed sense of taste
- ☐ headache
- ☐ muscle or joint pain

Sorted from new to old

State: Oct 23, 2022 at 4:26:08 PM

Drag down to update.

Suspected coronavirus (COVID-1...

Date: May 3, 2022 at 3:56:11 PM

State: Oct 23, 2022 at 4:35:34 PM

Drag down to update.

Washing your hands - but how?
Preventing infections and the spread of

When to wash your hands
Prevent the spreading of pathogens

Travelling
Finding out if a destination is safe

Droplet infections
Understanding and preventing droplet infection

Smear infection
Understanding and preventing smear infection

Masks that cover nose and mouth
Correct use of a masks that cover nose and

Quarantine
Prevent the spreading of pathogens

Latest information
Obtaining correct information

Risk areas
Avoiding travel to risk areas

Cure at home
Prevent the spreading of pathogens

When to wash your hands

GOAL

Prevent the spreading of pathogens

Wash your hands regularly and thoroughly, especially...

1. before preparing food
2. if you have touched the mucous membranes of your mouth or nose
3. before eating food
4. after using the toilet
5. if you accidentally had to cough or sneeze in your hand (please preferably cough or sneeze in your elbow)
6. when you come home

EXPLANATION

Thorough hand washing can prevent the spreading of pathogens. By washing your hands regularly, you can prevent others from getting infected and protect yourself.

This tip is also generally valid for many other

Data Management

Only participants who met the following criteria were included in the final analysis: provided consent to use data for research purposes; answered questions about age, user status, and gender identity; provided GPS data; and answered questions about previous contacts (in the past 14 days) with a person infected with COVID-19 and the presence or absence of the 11 symptoms listed before. In case the app was used multiple times by a single person, only the first recording was used. It should be noted that every time a questionnaire is filled out, the GPS location is saved, provided that the user gives their consent. The location permission is regulated by the operating system (Android or iOS). We do not store the raw location, but we process the data we receive from the operating system to make it more coarse-grained, and only store a value with an accuracy of 11.1 km. The level of detail with which we store the GPS sensor data was determined in discussions with all stakeholders, the institutional review board, and relevant laws and regulations. Further, only data from countries with more than 3000 users (Germany, India, and South Africa) were included in the statistical analyses. Data from other countries were comparatively small (each <600 users). The educational level was categorized into 3 groups (<12 years, ≥12 years, missing). Since most reports were collected in 2020 and the first variants of concern emerged around the turn of 2020/2021 [26], we defined 2 categories (2020 and 2021/2022) for the calendar year variable.

Statistical Analysis

Descriptive statistics were conducted to describe the CoronaCheck user characteristics. To compare sociodemographic characteristics (gender identity, age, education), calendar year, and user status among app users from Germany, India, South Africa, and other countries, chi-square tests were applied. Chi-square tests and Fisher exact tests were performed to compare the proportion of users who reported COVID-19–related symptoms and contact with a person infected with COVID-19 between 3 countries (Germany, India, and South Africa), gender identity (3 categories), educational level (3 categories), calendar years (2020 vs 2021/2022), and user status (symptoms reported for oneself vs another person). Univariate logistic regressions were calculated to assess the associations between reporting COVID-19–related symptoms and reporting contact with a person who tested positive for COVID-19 with country, gender, age (in 10-year categories), calendar year, educational level, and user status. Multivariable logistic regression was conducted to adjust the data for country, gender identity, age, calendar year, education level, and user status. Odds ratios (ORs) and their 95% CIs were estimated to assess the statistical uncertainty. All statistical analyses were performed using SPSS version 26 (IBM Corp), and all tests were 2-tailed.

Results

Study Population

From April 8, 2020, to February 3, 2022, we received a total of 88,537 recordings. Recordings with missing GPS data (n=29,449, 33.3%), missing consent that data can be used for research purposes (n=30,836, 34.8%), or missing information about age (n=2001, 2.3%), gender identity (n=1927, 2.2%), or user status (n=10,039, 11.3%) were excluded. Only the first recording per person was included in the case of multiple use. After applying these exclusion criteria, 23,179 recordings remained. All recordings included answers on the questions about previous contact (in the past 14 days) with a person infected with COVID-19 and the presence or absence of the 11 symptoms listed before. Most recordings were from Germany (n=8116, 35.0%), India (n=6622, 28.6%), and South Africa (n=3705, 16.0%). Recordings from 131 other countries ranged from 1 (0.004%) to 595 (2.6%) per country and were summarized into 1 group, "other countries" (n=4736, 20.4%).

The characteristics of the sample stratified by country are summarized in Table 1. In short, about half of the users were younger than 30 years old, with a higher proportion of young people using the app in India than in Germany. On average, 35.5% (n=8218) of the users were women, with the lowest proportion of women using the app in India.

The 4736 (20.4%) app users from “other countries” were excluded from further analysis, yielding a final analytic cohort of 18,443 (79.6%) individuals.

Table 1. Characteristics of CoronaCheck app users stratified by country.

Characteristics	Country				Total (N=23,179), n (%)
	Germany (n=8116, 35.0%), n (%)	India (n=6622, 28.6%), n (%)	South Africa (n=3705, 16.0%), n (%)	Other countries (n=4736, 20.4%), n (%)	
Gender identity (*P*<.001)					
Man	4769 (58.8)	4686 (70.8)	2082 (56.2)	3305 (69.8)	14,842 (64.0)
Woman	3308 (40.8)	1893 (28.6)	1611 (43.5)	1406 (29.7)	8218 (35.5)
Diverse	39 (0.5)	43 (0.6)	12 (0.3)	25 (0.5)	119 (0.5)
Age (years; *P*<.001)					
< 30	1964 (24.2)	4649 (70.2)	1982 (53.5)	2507 (52.9)	11,102 (47.9)
30-49	2722 (33.5)	1556 (23.5)	1469 (39.6)	1480 (31.4)	7227 (31.2)
≥50	3430 (42.3)	417 (6.3)	254 (6.9)	749 (15.8)	4850 (20.9)
Education (years; *P*<.001)					
<12	3827 (47.2)	1849 (27.9)	891 (24.0)	1372 (29.0)	7939 (34.3)
≥12	3030 (37.3)	2047 (30.9)	1906 (51.4)	1970 (41.6)	8953 (38.6)
Missing	1259 (15.5)	2726 (41.2)	908 (24.5)	1394 (29.4)	6287 (27.1)
Calendar year (*P*<.001)					
2020	7084 (87.3)	5805 (87.7)	3094 (83.5)	3740 (79.0)	19,723 (85.1)
2021/2022	1032 (12.7)	817 (12.3)	611 (16.5)	996 (21.0)	3456 (14.9)
User status (*P*<.001)					
Reported for oneself	7499 (92.4)	5222 (78.9)	3314 (89.4)	4032 (85.1)	20,067 (86.6)
Reported for another person	617 (7.6)	1400 (21.1)	391 (10.6)	704 (14.9)	3112 (13.4)

Occurrence of COVID-19–Related Symptoms and Close Contact With an Infected Person

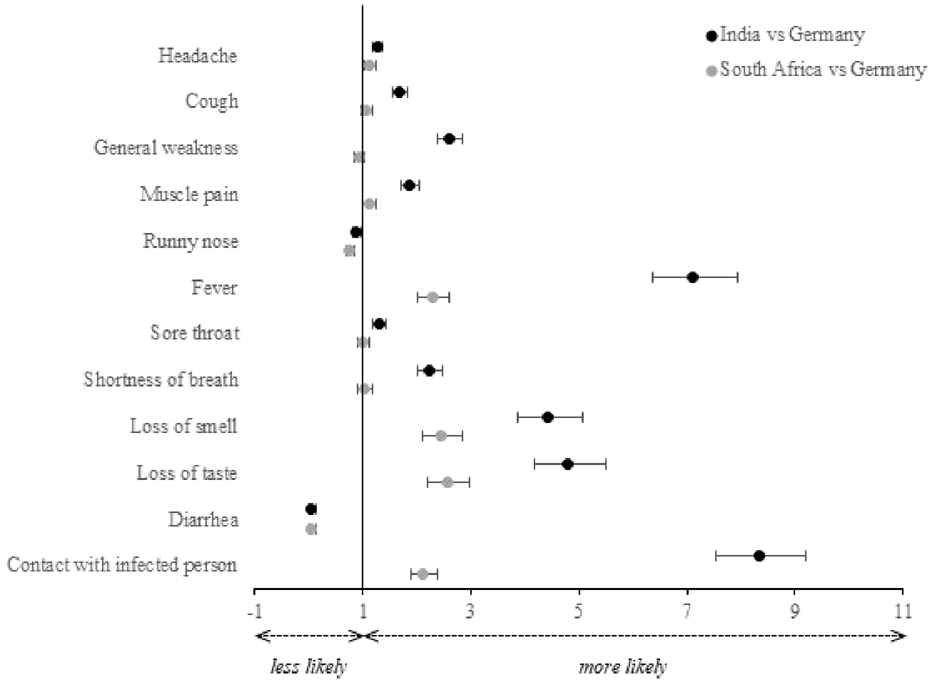
Of the 18,443 recordings collected in Germany, India, and South Africa, the 3 most frequent self-reported symptoms were headache (n=5351, 29.0%), cough (n=5219, 28.3%), and general weakness (n=4592, 24.9%), followed by muscle pain (n=4131, 22.4%), runny nose (n=3984, 21.6%), fever (n=3815, 20.7%), and sore throat (n=3646, 19.8%). Shortness of breath (n=2852, 15.5%), loss of smell (n=2224, 12.1%), loss of taste (n=2161, 11.7%), and diarrhea (n=136, 0.7%) were reported less frequently.

In 5105 (27.7%) of the 18,443 recordings collected in the 3 countries, the app users reported close contact with someone infected with SARS-CoV-2 in the past 14 days.

Exploration of Country Differences

We found differences between the 3 countries in all symptoms (Figure 2 and Multimedia Appendix 1, Tables S1 and S2). We observed higher odds of reporting headache, cough, general weakness, muscle pain, fever, sore throat, shortness of breath, and loss of smell and taste: the adjusted odds ratios (aORs) were 1.27-7.10 (95% CI 1.17-7.95) for app users from India compared to Germany. In contrast, app users from India reported diarrhea less frequently than those from Germany (aOR 0.05, 95% CI 0.02-0.12). App users from South Africa were more likely to report headache, muscle pain, fever, and loss of smell and taste (aORs 1.11-2.56, 95% CI 1.00-2.98), while they were less likely to report runny nose (aOR 0.76, 95% CI 0.68-0.84) and diarrhea (aOR 0.03, 95% CI 0.01-0.12) than users from Germany. App users from India were significantly more likely to have knowledge of close contact with a person positive for SARS-CoV-2 in the past 14 days compared to Germany (aOR 8.33, 95% CI 7.53-9.21). To a lesser extent, this was also the case for app users from South Africa compared to Germany (aOR 2.12, 95% CI 1.89-2.38).

Figure 2. aORs and their 95% CIs stratified by the country of the CoronaCheck app users. aOR: adjusted odds ratio.

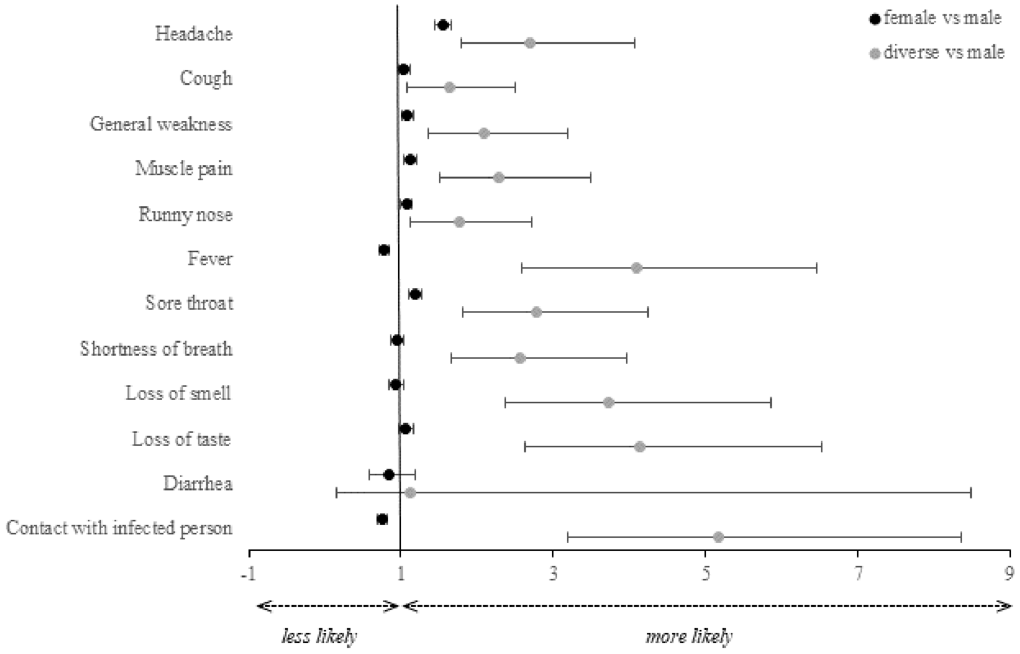


Exploration of Sociodemographic Differences

Gender Identity Differences

Our results showed gender identity differences for all symptoms except diarrhea (Figure 3 and Multimedia Appendix 1, Tables S3 and S4). Women app users reported fever less often (aOR 0.78, 95% CI 0.72-0.85) and headache, general weakness, muscle pain, runny nose, and sore throat more often (aORs 1.07-1.55, 95% CI 1.00-1.65) compared to men. Except for diarrhea, gender-diverse users reported COVID-19–related symptoms more often compared to men (aORs 1.64-4.14, 95% CI 1.08-6.53). Women were less likely than men to know of recent contact with a person infected with SARS-CoV-2 (aOR 0.76, 95% CI 0.70-0.82). Gender-diverse users had 5.16 times higher odds (95% CI 3.18-8.36) of reporting close contact with a person infected with SARS-CoV-2 in the past 14 days compared to app users who were men.

Figure 3. aORs and their 95% CIs stratified by the gender identity of the CoronaCheck app users. aOR: adjusted odds ratio.



Age Effects

Older persons were less likely to report headache, cough, general weakness, runny nose, fever, sore throat, and loss of smell (aORs 0.86-0.96, 95% CI 0.84-0.99, per 10-year category), whereas they were more likely to report muscle pain (aOR 1.03, 95% CI 1.01-1.05, per 10-year category; Table 2). Older persons were less likely to report close contact with a person positive for SARS-CoV-2 in the past 14 days (aOR 0.87, 95% CI 0.85-0.89, per 10-year category).

Table 2. Unadjusted ORs[a] and aORs[b] by 10-year age category (N=18,445).[c]

Variable	Unadjusted OR (95% CI)	aOR (95% CI)
Headache	0.84 (0.82-0.86)	0.87 (0.85-0.89)
Cough	0.85 (0.84-0.87)	0.90 (0.89-0.92)
General weakness	0.87 (0.86-0.89)	0.96 (0.94-0.99)
Muscle pain	0.96 (0.94-0.98)	1.03 (1.01-1.05)
Runny nose	0.87 (0.85-0.89)	0.86 (0.84-0.88)
Fever	0.76 (0.74-0.77)	0.92 (0.90-0.94)
Sore throat	0.84 (0.82-0.86)	0.87 (0.85-0.89)
Shortness of breath	0.91 (0.89-0.93)	0.99 (0.97-1.02)
Loss of smell	0.81 (0.79-0.83)	0.95 (0.92-0.98)
Loss of taste	0.83 (0.81-0.86)	0.98 (0.95-1.02)
Diarrhea	1.17 (1.08-1.28)	0.92 (0.84-1.01)
Contact with infected person	0.71 (0.70-0.73)	0.87 (0.85-0.89)

[a] OR: odds ratio.

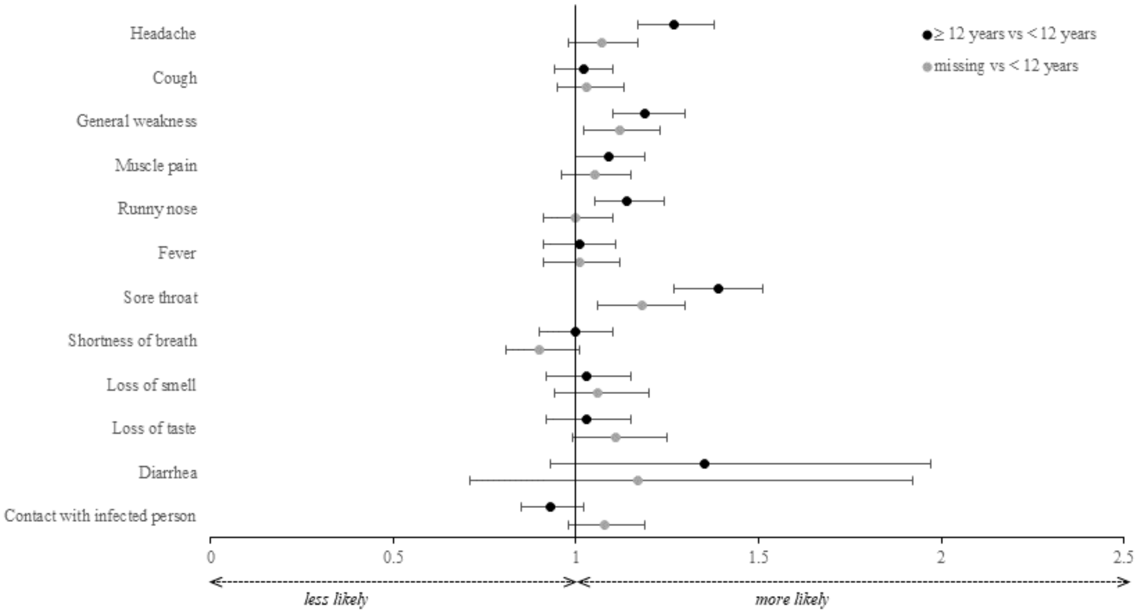
[b] aOR: adjusted odds ratio.

[c] The multivariable regression model was adjusted for country, gender identity, calendar year, education, and user status.

Differences Concerning the Educational Level

App users with at least 12 years of school education reported headache, general weakness, muscle pain, runny nose, and sore throat more often compared to users with 11 or fewer years of education (aORs 1.09-1.39, 95% CI 1.00-1.51; Figure 4 and Multimedia Appendix 1, Tables S5 and S6).

Figure 4. aORs and their 95% CIs stratified by educational level (years of schooling). aOR: adjusted odds ratio.

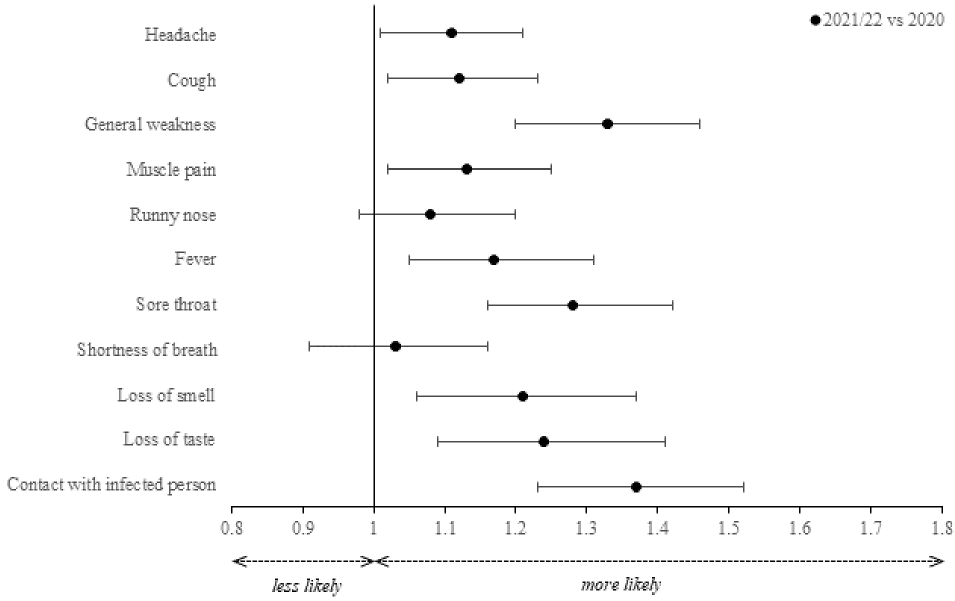


Exploration of Differences Between Calendar Years

Most COVID-19–associated symptoms were more likely to be reported by app users in 2021/2022 compared to 2020 (aORs 1.11-1.33, 95% CI 1.01-1.46; Figure 5 and Multimedia Appendix 1, Tables S7 and S8). The occurrence of runny nose and shortness of breath was not associated with the calendar year.

The reporting of close contact with someone infected with SARS-CoV-2 in the past 14 days was more likely in 2021/2022 compared to 2020 (aOR 1.37, 95% CI 1.23-1.52).

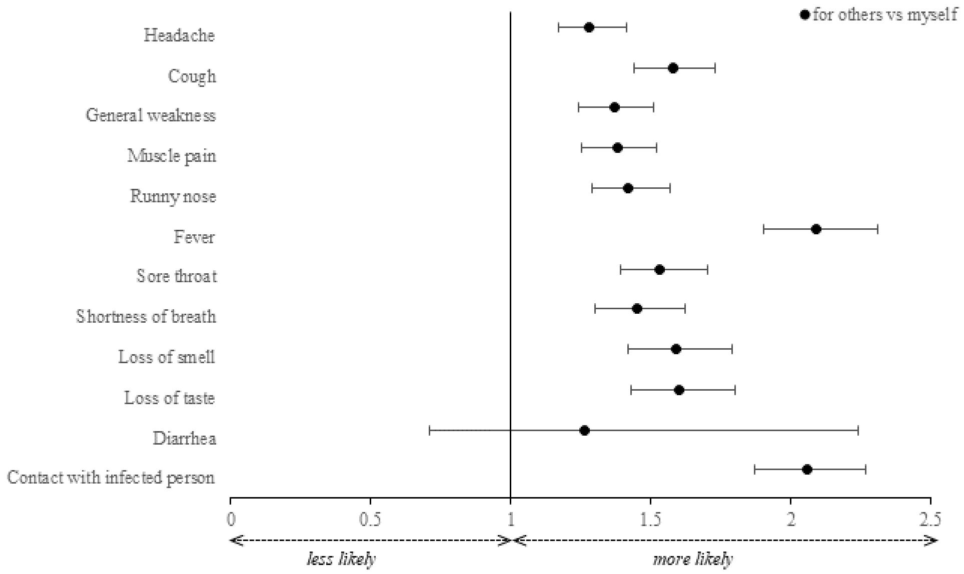
Figure 5. aORs and their 95% CIs stratified by the year of using the CoronaCheck app. aOR: adjusted odds ratio.



Exploration of Differences Concerning the Status of the Person (Reported for Oneself vs for Others)

We found differences between persons for whom symptoms were reported for all symptoms except diarrhea (Figure 6 and Multimedia Appendix 1, Tables S9 and S10). Symptoms as well as close contact with a person positive for SARS-CoV-2 in the past 14 days were more often reported when the data were entered for another person compared to for oneself (aORs 1.28-2.09, 95% CI 1.17-2.31).

Figure 6. aORs and their 95% CIs stratified by user status (ie, user-reported symptoms for another person or for oneself). aOR: adjusted odds ratio.



Discussion

Principal Findings

Our study showed that self-reported COVID-19–related symptoms and reports of previous contact with individuals infected with COVID-19 vary by country, gender identity, age, calendar year, education, and user status.

Overall, the most frequently reported symptoms were headache, cough, and general weakness. Symptoms commonly included in COVID-19 case definitions, such as shortness of breath and fever [27,28], were less frequently reported by app users. Especially app users from India (compared to Germany), as well as gender-diverse (compared to men) users were more likely to report COVID-19–related symptoms and previous close contact with infected individuals.

Comparison With Prior Work

Our findings regarding the occurrence of common symptoms, such as headache, general weakness, and runny nose, are consistent with previous studies. Specifically, previous studies in individuals infected with PCR-confirmed SARS-CoV-2 have shown that fatigue, headache, malaise, myalgia, and upper respiratory symptoms (sore throat, cough, sneezing, rhinitis) occur early after symptom onset, whereas symptoms considered more characteristic of COVID-19, such as lower respiratory and chemosensory symptoms, occur later [17,29,30].

Our study showed differences in self-reported COVID-19 symptoms primarily between Germany, India, and South Africa. App use from India and South Africa was associated with a disproportionately high likelihood of experiencing COVID-19–related symptoms and knowledge of close contact with an infected person. Specifically, the likelihood of reporting contact with a person infected with COVID-19 was 8.3-fold higher among users from India compared to Germany. Most symptoms were also more likely to be reported by users from India than by German users. World Health Organization (WHO) official data on confirmed COVID-19 cases show conflicting results, with higher numbers of confirmed cases and deaths per capita in Germany than in India and South Africa [2]. In more detail, by the beginning of February 2022, the cumulative confirmed COVID-19 cases amounted to 15% of the German population, higher than in South Africa (6%) and India (3%). The discrepancy is likely due to several factors of uncertainty in the COVID-19 data, such as the low number of COVID-19 tests and deficiencies in surveillance systems in India that result in underreporting of infections and deaths [31]. Studies that compare cross-national cases run the risk of reaching incorrect conclusions [19]. Thus, a higher likelihood of reporting a previous contact with a person testing positive for COVID-19 in our study and the reporting of symptoms associated with COVID-19 suggest that there is substantial underreporting bias in the existing data from India. In addition, higher odds of reporting previous contact with a person positive for COVID-19 (OR 2.12) as well as higher odds of reporting several symptoms were observed for app users from South Africa compared to Germany. Although South Africa is the country that performs the most COVID-19 testing among African countries and is one of the few countries that produces vital statistics, it is likely that the number of COVID-19 cases and deaths is underestimated [32]. The findings suggest that effective management of the pandemic is particularly difficult in countries with limited resources, inadequate testing, and weak health systems.

Gender identity differences were found in this study. Several symptoms were more likely to be reported by women and gender-diverse users than by men. Exceptions to the general assumption that women are more likely to report symptoms than men were observed for fever. Previous studies of individuals testing positive for COVID-19 support the assumption that women are more likely than men to be affected by headache and sore throat, as well as rhinitis [15,17,33,34]. Results regarding muscle pain [15,17] are equivocal, and fever seems to be less prevalent in women [33]. Female users less frequently reported contact with a person positive for COVID-19 compared to male users. Previous studies highlight that women are more willing to adhere to protective measures [35,36], which might be 1 explanation for the lower odds of self-reported contact with an infected person in women individuals and possibly also to the overall lower proportion of women app users.

The impact of COVID-19 on COVID-19 symptoms in gender-diverse individuals has not been well captured; however, the more than 5-fold increase in the likelihood of gender-diverse users reporting prior contact with an individual who tested positive for COVID-19, as well as the higher odds of reporting symptoms observed in this study, is supported by the notion that gender-diverse individuals are more vulnerable to COVID-19, which extends to the areas of increased barriers to COVID-19 transmission prevention strategies, decreased physical and mental health, and barriers to health services [37]. As such, the higher odds of reporting contact with persons infected with COVID-19 corroborates the reported reduced options of gender-diverse individuals to practice physical distancing, as they are more likely to be essential workers, requiring in-person work and often commuting by public transport [37]. Other discussed aspects are the higher behavioral risks, such as substance use and smoking, which are common among gender-diverse people [38]. Studies have also reported a higher risk for chronic health conditions, such as diabetes, in gender-diverse individuals [39], and they often face barriers to health care services [40]. Greater vaccination hesitancy has also been reported, likely associated with medical mistrust due to the historical mistreatment of gender-diverse individuals [37]. Thus, all these factors might place this population at greater risk for COVID-19–related illness. Considering these previous reports as well as the findings of the study at hand, gender-diverse individuals seem to be among those minorities most strongly affected by the COVID-19 pandemic.

Older age was found to be associated with a lower likelihood of reporting contact with a person infected with COVID-19. In support of this finding, previous studies have found an association between older age and self-reported adherence to COVID-19 public health measures [41]). Reports of age-related differences in the clinical presentation of patients with COVID-19 suggest that older patients with COVID-19 are more likely to have general symptoms (fever, fatigue, gastrointestinal symptoms) than specific symptoms (loss of smell and taste) [17,33,34]. In this study, the odds of reporting almost all COVID-19–related symptoms decreased with increasing age. Given the lower likelihood of being a contact in older individuals, it can be surmised that the proportion of app users infected with COVID-19 was lower among older users, which likely contributed to their lower frequency of reported symptoms.

The increased odds of reporting contact with someone who tested positive for SARS-CoV-2 in 2021/2022 compared to 2020 is supported by the increased transmission of the variants of concern, which predominated from 2021 onward. The Alpha variant (B.1.1.7), first identified in the United Kingdom at the end of 2020; the Beta variant (B.1.351), first identified in South Africa at the end of 2020; the Delta variant (B.1.617.2), first identified in India at the end of 2020; and the Omicron variant (B.1.1.529), first identified in South Africa at the end of 2021, have roughly 50% increased transmission [11]. In addition, the higher odds of reporting symptoms after 2020 compared to the year in which the SARS-CoV-2 wild-type virus was dominant might be due to the increased disease severity of the Alpha, Beta, and Delta variants [26].

In this study, the person for whom the symptom check was used affected the outcome: the likelihood of reporting contact with a person positive for COVID-19 and various COVID-19–related symptoms was higher when the app was used to check symptoms for another person. Thus, it appears that the app was used more frequently for other individuals when there was a specific suspicion of infection. Differences in reporting systems have been observed previously, with a higher percentage of symptoms reported in self-reports than in physician reports [7]. Therefore, the use of a self-report app seems to be an appropriate method for recording COVID-19–related symptoms.

Strengths

Our study had several strengths. First, data were collected over a long time interval that began early in the pandemic. Data were collected across different countries, with considerable differences in health care systems, economic status, and governmental regulations to combat the spread of the virus, therefore increasing the generalizability of the findings. The number of gender-diverse individuals among app users also allowed us to identify the specifics of symptoms reported by COVID-19 among gender minorities. In addition, it was possible to develop the app in a relatively short period and adapt it to the MDR, which meant that, among other things, special consideration was given to the risks involved in using the app. The app made use of GPS measurements when users allowed it, which shows that such measurements are also welcomed by many users. Finally, the content of the app was based on the recommendations of the Robert Koch Institute, the national public health institute in Germany. We therefore regularly adapted the app to the changing official recommendations.

Limitations

Our study also had some limitations that need to be noted. First, objective data on whether app users tested positive for COVID-19 or not were not available. Thus, we were not able to validate the subjectively reported information about COVID-19 test positivity. Moreover, no objective data were available if the infected contact person of the user had a positive COVID-19 test; however, the corresponding question asked specifically about contact with a person who was a confirmed case.

Second, all information about specific symptoms was self-reported by the users, which was not validated by, for example, doctors' diagnoses. In those cases where the questionnaire was filled out for another person, the validity of the data is further limited. This especially applied to the youngest and oldest app users, where a higher proportion of reported information was not entered by themselves (51.5% among children ≤9 years old and 21.1% among the elderly ≥80 years old vs 12.1% among 10-79-year-old persons).

Third, participants' socioeconomic status was limited to the information about the years of education in school. This information was further limited by the high number of missing entries as well as by differences in the educational systems between countries.

Fourth, we had to adjust the app several times to meet the changing official recommendations of the German public health authority, the Robert Koch Institute, over the course of the pandemic. Therefore, during the overall time of app usage, especially in the early phase of the pandemic, some information was slightly differently presented to users. For example, the question regarding diarrhea was only administered for the first 2 months before it was removed.

Furthermore, as data collection was based on a mobile app, results might not be generalized to individuals with lower affinity for digital health apps. In concrete terms, this also means that information and selection bias must be taken into account. However, observer and recall bias can be reduced to some extent by an app in terms of patient-reported outcomes. Therefore, app-based data have strengths as well as weaknesses that need to be considered. Moreover, mHealth apps are always not just tools but also social phenomena that create a feedback loop between app and user, which means that users continuously adjust their behavior as they use the app. However, since the CoronaCheck app has collected a comparatively large amount of data in different cultures compared to other mHealth apps, certain differences may have been balanced out in the sense of generalization. Finally, it should be mentioned that apps that are made available on multiple mobile operating systems will also always generate differences in the resulting data, as the user interface is never completely identical.

Conclusion

Our app-based results indicated country-specific differences in COVID-19–related symptom patterns. It seems that self-reported COVID-19–related symptoms and awareness about previous contacts with individuals infected with COVID-19 are more likely among app users from India and South Africa (compared to users from Germany), as well as in gender-diverse (compared to male) users. Symptom assessment tools could play a positive role in the context of widespread infectious diseases. In general, with COVID-19, we saw that the way symptoms were reported and the way infection and COVID-19 deaths were defined were different [42]. For future pandemics, app-based global data collection, such as that of CoronaCheck, is another tool to better equip us and support appropriate assessments.

Future mHealth research on country-specific differences during a viral pandemic should aim to recruit representative samples. If pseudonymous assessments are possible, studies should collect information about symptoms, confirmed testing, and adherence to public health recommendations longitudinally.

Acknowledgments

We are grateful to all app users for providing data for research purposes. We thank Rachel Dale for proofreading the manuscript and Marc Schickler for his help in visualizing the results.

Data Availability

The data presented in this study are only available on request from the corresponding author. The data are not publicly available, because participants' informed consent did not cover public deposition of data.

Authors' Contributions

EH analyzed the data and drafted the manuscript. TK was the coprincipal investigator, supervised the development of the app regarding the medical content, planned and developed the questionnaire, and participated in the planning of the statistical analyses and writing of the manuscript. CS participated in the development of the study design and questionnaire and updated the medical content and corresponding recommendations by state and nationwide health authorities. WS and MW participated in the development of the study design. PH spent resources and provided advice. MW participated in the statistical analyses. TP analyzed the data and drafted the manuscript. RP was the principal investigator, was responsible for the study design and technical development, and drafted the manuscript. All coauthors have reviewed and approved the manuscript for submission. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 73 KB - publichealth_v9i1e40958_app1.docx]

References

1. Li J, Lai S, Gao GF, Shi W. The emergence, genomic diversity and global spread of SARS-CoV-2. Nature 2021 Dec 08;600(7889):408-418. [doi: 10.1038/s41586-021-04188-6] [Medline: 34880490]
2. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. URL: https://covid19.who.int [accessed 2022-03-14]
3. John Leon Singh H, Couch D, Yap K. Mobile health apps that help with COVID-19 management: scoping review. JMIR Nurs 2020 Aug 6;3(1):e20596 [FREE Full text] [doi: 10.2196/20596] [Medline: 32897271]
4. Mary L, Raj S. A survey on SARS-CoV-2 (COVID-19) using machine learning techniques. 2021 Presented at: 2nd International Conference on Smart Electronics and Communication (ICOSEC); October 7-9, 2021; Trichy, India p. 1612-1617. [doi: 10.1109/icosec51865.2021.9591841]
5. Spinato G, Gaudioso P, Boscolo Rizzo P, Fabbris C, Menegaldo A, Mularoni F, et al. Risk management during COVID-19: safety procedures for otolaryngologists. Acta Biomed 2021 Feb 16;92(1):e2021105 [FREE Full text] [doi: 10.23750/abm.v92i1.11281] [Medline: 33682810]
6. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): a systematic review and meta-analysis of 148 studies from 9 countries. PLoS One 2020 Jun 23;15(6):e0234765 [FREE Full text] [doi: 10.1371/journal.pone.0234765] [Medline: 32574165]
7. Mizrahi B, Shilo S, Rossman H, Kalkstein N, Marcus K, Barer Y, et al. Longitudinal symptom dynamics of COVID-19 infection. Nat Commun 2020 Dec 04;11(1):6208 [FREE Full text] [doi: 10.1038/s41467-020-20053-y] [Medline: 33277494]
8. Callejon-Leblic MA, Moreno-Luna R, Del Cuvillo A, Reyes-Tejero IM, Garcia-Villaran MA, Santos-Peña M, et al. Loss of smell and taste can accurately predict COVID-19 infection: a machine-learning approach. J Clin Med 2021 Feb 03;10(4):570 [FREE Full text] [doi: 10.3390/jcm10040570] [Medline: 33546319]
9. Sudre CH, Keshet A, Graham MS, Joshi AD, Shilo S, Rossman H, et al. Anosmia, ageusia, and other COVID-19-like symptoms in association with a positive SARS-CoV-2 test, across six national digital surveillance platforms: an observational study. Lancet Digital Health 2021 Jul;3(9):e577-e586. [doi: 10.1016/s2589-7500(21)00115-1]
10. Elias C, Sekri A, Leblanc P, Cucherat M, Vanhems P. The incubation period of COVID-19: a meta-analysis. Int J Infect Dis 2021 Mar;104:708-710 [FREE Full text] [doi: 10.1016/j.ijid.2021.01.069] [Medline: 33548553]
11. Long B, Carius BM, Chavez S, Liang SY, Brady WJ, Koyfman A, et al. Clinical update on COVID-19 for the emergency clinician: presentation and evaluation. Am J Emerg Med 2022 Apr;54:46-57 [FREE Full text] [doi: 10.1016/j.ajem.2022.01.028] [Medline: 35121478]
12. Quer G, Radin JM, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. Nat Med 2021 Jan 29;27(1):73-77. [doi: 10.1038/s41591-020-1123-x] [Medline: 33122860]

13. Rai P, Kumar BK, Deekshit VK, Karunasagar I, Karunasagar I. Detection technologies and recent developments in the diagnosis of COVID-19 infection. Appl Microbiol Biotechnol 2021 Jan 04;105(2):441-455 [FREE Full text] [doi: 10.1007/s00253-020-11061-5] [Medline: 33394144]
14. Almalki M, Giannicchi A. Health apps for combating COVID-19: descriptive review and taxonomy. JMIR Mhealth Uhealth 2021 Mar 02;9(3):e24322 [FREE Full text] [doi: 10.2196/24322] [Medline: 33626017]
15. Sebo P, Maisonneuve H, Lourdaux J, Cuzin C, Floquet M, Tudrej B, et al. Self-reported symptoms in French primary care SARS-CoV-2 patients: association with gender and age group. Fam Pract 2021 Aug 27;38(Suppl 1):i37-i44 [FREE Full text] [doi: 10.1093/fampra/cmab009] [Medline: 34448483]
16. Ahmed MM, Sayed AM, El Abd D, Fares S, Said MSM, Elsayed Sedik Ebrahim E. Gender difference in perceived symptoms and laboratory investigations in suspected and confirmed COVID-19 cases: a retrospective study. J Prim Care Community Health 2021 Aug 19;12:21501327211039718 [FREE Full text] [doi: 10.1177/21501327211039718] [Medline: 34407661]
17. Wiegele PN, Kabar I, Kerschke L, Froemmel C, Hüsing-Kabar A, Schmidt H, et al. Symptom diary-based analysis of disease course among patients with mild coronavirus disease, Germany, 2020. Emerg Infect Dis 2021 May;27(5):1353-1361 [FREE Full text] [doi: 10.3201/eid2705.204507] [Medline: 33900166]
18. Buyukkececi Z. Cross-country differences in anxiety and behavioral response to the Covid-19 pandemic. Eur Soc 2020 Oct 08;23(sup1):S417-S447. [doi: 10.1080/14616696.2020.1828975]
19. Biswas RK, Afiaz A, Huq S. Underreporting COVID-19: the curious case of the Indian subcontinent. Epidemiol Infect 2020 Sep 11;148:e207. [doi: 10.1017/s0950268820002095]
20. Vogel C, Pryss R, Schobel J, Schlee W, Beierle F. Developing apps for researching the COVID-19 pandemic with the TrackYourHealth platform. 2021 Presented at: IEEE/ACM 8th International Conference on Mobile Software Engineering and Systems (MobileSoft); May 17-19, 2021; Madrid, Spain p. 65-68. [doi: 10.1109/mobilesoft52590.2021.00015]
21. Beierle F, Schobel J, Vogel C, Allgaier J, Mulansky L, Haug F, et al. Corona Health—a study- and sensor-based mobile app platform exploring aspects of the COVID-19 pandemic. IJERPH 2021 Jul 10;18(14):7395. [doi: 10.3390/IJERPH18147395]
22. Holfelder M, Mulansky L, Schlee W, Baumeister H, Schobel J, Greger H, et al. Medical device regulation efforts for mhealth apps during the COVID-19 pandemic—an experience report of Corona Check and Corona Health. J 2021 Jun 16;4(2):206-222. [doi: 10.3390/j4020017]
23. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). OJEU 2016 May 4;L119:1-88. [doi: 10.5593/sgemsocial2019v/1.1/s02.022]
24. Corona Health. Privacy Policy. URL: https://www.corona-health.net/en/privacy/ [accessed 2023-01-20]
25. Pryss R, Schobel J, Reichert M. Requirements for a flexible and generic API enabling mobile crowdsensing mHealth applications. 2018 Presented at: 4th International Workshop on Requirements Engineering for Self-Adaptive, Collaborative, and Cyber Physical Systems (RESACS); August 20, 2018; Banff, AB, Canada p. 24-31. [doi: 10.1109/resacs.2018.00010]
26. Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, et al. The biological and clinical significance of emerging SARS-CoV-2 variants. Nat Rev Genet 2021 Dec 17;22(12):757-773 [FREE Full text] [doi: 10.1038/s41576-021-00408-x] [Medline: 34535792]
27. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
28. Shen Y, Zheng F, Sun D, Ling Y, Chen J, Li F, et al. Epidemiology and clinical course of COVID-19 in Shanghai, China. Emerg Microbes Infect 2020 Dec 07;9(1):1537-1545 [FREE Full text] [doi: 10.1080/22221751.2020.1787103] [Medline: 32573353]
29. Saad N, Moek F, Steitz F, Murajda L, Bärnighausen T, Zoller T, et al. A longitudinal study on symptom duration and 60-day clinical course in non-hospitalised COVID-19 cases in Berlin, Germany, March to May, 2020. Eurosurveillance 2021 Oct 28;26(43):1-9. [doi: 10.2807/1560-7917.es.2021.26.43.2001757]
30. Bliddal S, Banasik K, Pedersen OB, Nissen J, Cantwell L, Schwinn M, et al. Acute and persistent symptoms in non-hospitalized PCR-confirmed COVID-19 patients. Sci Rep 2021 Jun 23;11(1):13153 [FREE Full text] [doi: 10.1038/s41598-021-92045-x] [Medline: 34162913]
31. Chatterjee P. Is India missing COVID-19 deaths? Lancet 2020 Sep;396(10252):657. [doi: 10.1016/s0140-6736(20)31857-2]
32. Okonji EF, Okonji OC, Mukumbang FC, Van Wyk B. Understanding varying COVID-19 mortality rates reported in Africa compared to Europe, Americas and Asia. Trop Med Int Health 2021 Jul;26(7):716-719 [FREE Full text] [doi: 10.1111/tmi.13575] [Medline: 33733568]
33. Lechien JR, Chiesa-Estomba CM, Place S, Van Laethem Y, Cabaraux P, Mat Q, COVID-19 Task Force of YO-IFOS. Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019. J Intern Med 2020 Sep 17;288(3):335-344 [FREE Full text] [doi: 10.1111/joim.13089] [Medline: 32352202]
34. Lee Y, Min P, Lee S, Kim S. Prevalence and duration of acute loss of smell or taste in COVID-19 patients. J Korean Med Sci 2020 May 11;35(18):e174 [FREE Full text] [doi: 10.3346/jkms.2020.35.e174] [Medline: 32383370]

35. Coroiu A, Moran C, Campbell T, Geller AC. Barriers and facilitators of adherence to social distancing recommendations during COVID-19 among a large international sample of adults. PLoS One 2020 Oct 7;15(10):e0239795 [FREE Full text] [doi: 10.1371/journal.pone.0239795] [Medline: 33027281]
36. Lin T, Harris EA, Heemskerk A, Van Bavel JJ, Ebner NC. A multi-national test on self-reported compliance with COVID-19 public health measures: the role of individual age and gender demographics and countries' developmental status. Soc Sci Med 2021 Oct;286:114335 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114335] [Medline: 34450390]
37. Zubizarreta D, Trinh M, Reisner SL. COVID-19 risk and resilience among U.S. transgender and gender diverse populations. Am J Prev Med 2022 Feb;62(2):299-303 [FREE Full text] [doi: 10.1016/j.amepre.2021.07.017] [Medline: 34716015]
38. Buchting FO, Emory KT, Scout, Kim Y, Fagan P, Vera LE, et al. Transgender use of cigarettes, cigars, and e-cigarettes in a national study. Am J Prev Med 2017 Jul;53(1):e1-e7 [FREE Full text] [doi: 10.1016/j.amepre.2016.11.022] [Medline: 28094133]
39. Dragon CN, Guerino P, Ewald E, Laffan AM. Transgender Medicare beneficiaries and chronic conditions: exploring fee-for-service claims data. LGBT Health 2017 Dec;4(6):404-411 [FREE Full text] [doi: 10.1089/lgbt.2016.0208] [Medline: 29125908]
40. Jones BA, Bowe M, McNamara N, Guerin E, Carter T. Exploring the mental health experiences of young trans and gender diverse people during the Covid-19 pandemic. Int J Transgend Health 2021 Mar 01:1-13. [doi: 10.1080/26895269.2021.1890301]
41. Wright L, Fancourt D. Do predictors of adherence to pandemic guidelines change over time? A panel study of 22,000 UK adults during the COVID-19 pandemic. Prev Med 2021 Dec;153:106713 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106713] [Medline: 34242662]
42. HSC Public Health Agency. Why Are There Differences in the Reporting of COVID-19 Related Deaths?. URL: https://www.publichealth.hscni.net/node/5174 [accessed 2022-06-07]

Abbreviations

aOR: adjusted odds ratio
MDR: Medical Device Regulations
mHealth: mobile health
OR: odds ratio
PCR: polymerase chain reaction

Edited by T Sanchez, A Mavragani; submitted 11.07.22; peer-reviewed by S Pesälä, E Grasaas; comments to author 18.10.22; revised version received 07.11.22; accepted 13.12.22; published 03.02.23.

Please cite as:
Humer E, Keil T, Stupp C, Schlee W, Wildner M, Heuschmann P, Winter M, Probst T, Pryss R
Associations of Country-Specific and Sociodemographic Factors With Self-Reported COVID-19–Related Symptoms: Multivariable Analysis of Data From the CoronaCheck Mobile Health Platform
JMIR Public Health Surveill 2023;9:e40958
URL: https://publichealth.jmir.org/2023/1/e40958
doi:10.2196/40958
PMID:36515987

Original Paper

Epidemiology, Symptomatology, and Risk Factors for Long COVID Symptoms: Population-Based, Multicenter Study

Martin Chi-Sang Wong[1,2,3,4*], MD, MPH; Junjie Huang[1,2*], PhD; Yuet-Yan Wong[1], MA; Grace Lai-Hung Wong[5,6,7], MBChB, MD; Terry Cheuk-Fung Yip[5,6,7], PhD; Rachel Ngan-Yin Chan[8], MPhil, PhD; Steven Wai-Ho Chau[8], MBBS, MSc; Siew-Chien Ng[5,7], MBBS, PhD; Yun-Kwok Wing[8], MBChB; Francis Ka-Leung Chan[5,7], MBChB, MD

[1]Jockey Club School of Public Health & Primary Care, The Chinese University of Hong Kong, Sha Tin, China
[2]Centre for Health Education and Health Promotion, Faculty of Medicine, The Chinese University of Hong Kong, Sha Tin, China
[3]School of Public Health, Peking University, Beijing, China
[4]School of Public Health, The Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[5]Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Sha Tin, China
[6]Medical Data Analytics Centre, The Chinese University of Hong Kong, Sha Tin, China
[7]Institute of Digestive Disease, Faculty of Medicine, The Chinese University of Hong Kong, Sha Tin, China
[8]Department of Psychiatry, Faculty of Medicine, The Chinese University of Hong Kong, Tai Po, China
[*]these authors contributed equally

Corresponding Author:
Francis Ka-Leung Chan, MBChB, MD
Department of Medicine and Therapeutics
Faculty of Medicine
The Chinese University of Hong Kong
Room 41028, 2/F, Special Block E
Sha Tin
China
Phone: 852 35051339
Fax: 852 26471557
Email: fklchan@cuhk.edu.hk

Abstract

Background: Long COVID induces a substantial global burden of disease. The pathogenesis, complications, and epidemiological and clinical characteristics of patients with COVID-19 in the acute phase have been evaluated, while few studies have characterized the epidemiology, symptomatology, and risk factors of long COVID symptoms. Its characteristics among patients with COVID-19 in the general population remain unaddressed.

Objective: We examined the prevalence of long COVID symptoms, its symptom patterns, and its risk factors in 4 major Chinese cities in order to fill the knowledge gap.

Methods: We performed a population-based, multicenter survey using a representative sampling strategy via the Qualtrics platform in Beijing, Shanghai, Guangzhou, and Hong Kong in June 2022. We included 2712 community-dwelling patients with COVID-19 and measured the prevalence of long COVID symptoms defined by the World Health Organization (WHO), and their risk factors. The primary outcomes were the symptoms of long COVID, with various levels of impact. A descriptive analysis of the prevalence and distribution of long COVID symptoms according to disease severity was conducted. A sensitivity analysis of increasing the number of long COVID symptoms was also conducted. Univariate and multivariate regression analyses were performed to examine the risk factors of severe long COVID symptoms, including age, gender, marital status, current occupation, educational level, living status, smoking habits, monthly household income, self-perceived health status, the presence of chronic diseases, the use of chronic medication, COVID-19 vaccination status, and the severity of COVID-19.

Results: The response rate was 63.6% (n=2712). The prevalence of long COVID, moderate or severe long COVID, and severe long COVID was 90.4% (n=2452), 62.4% (n=1692), and 31.0% (n=841), respectively. Fatigue (n=914, 33.7%), cough (n=865, 31.9%), sore throat (n=841, 31.0%), difficulty in concentrating (n=828, 30.5%), feeling of anxiety (n=817, 30.2%), myalgia (n=811, 29.9%), and arthralgia (n=811, 29.9%) were the most common severe long COVID symptoms. From multivariate regression analysis, female gender (adjusted odds ratio [aOR]=1.49, 95% CI 1.13-1.95); engagement in transportation, logistics, or the discipline workforce (aOR=2.52, 95% CI 1.58-4.03); living with domestic workers (aOR=2.37, 95% CI 1.39-4.03); smoking (aOR=1.55, 95% CI 1.17-2.05); poor or very poor self-perceived health status (aOR=15.4, 95% CI 7.88-30.00); ≥3 chronic diseases (aOR=2.71, 95% CI 1.54-4.79); chronic medication use (aOR=4.38, 95% CI 1.66-11.53); and critical severity of COVID-19 (aOR=1.52, 95% CI 1.07-2.15) were associated with severe long COVID. Prior vaccination with ≥2 doses of COVID-19 vaccines was a protective factor (aOR=0.35-0.22, 95% CI 0.08-0.90).

Conclusions: We examined the prevalence of long COVID symptoms in 4 Chinese cities according to the severity of COVID-19. We also evaluated the pattern of long COVID symptoms and their risk factors. These findings may inform early identification of patients with COVID-19 at risk of long COVID and planning of rehabilitative services.

(JMIR Public Health Surveill 2023;9:e42315) doi:10.2196/42315

KEYWORDS

COVID-19; epidemiology; symptom; risk factor; long COVID; multicenter survey; general population

Introduction

COVID-19 has posed an unprecedented challenge to public health worldwide. As of July 9, 2022, more than 551 million cases and 6.3 million deaths have been reported [1]. Apart from the acute phase of the infection, the disease may also manifest as persistent, lingering symptoms known as long COVID [2-5]. The World Health Organization (WHO) has defined long COVID as a condition that occurs in patients with COVID-19 usually 3 months from the onset of the disease, with symptoms that last for at least 2 months and that the symptoms could not be attributed to an alternative diagnosis [6]. These long-term symptoms could exert a devastating effect [7], as they could involve multiple organ systems, and have been linked to psychosocial consequences [8]. A study has found that almost 10% of people with COVID-19 reported long COVID stmptoms 4-12 weeks after the infection [9,10].

The impact of long COVID symptoms has been observed even in patients with mild COVID-19, patients who do not require respiratory support or intensive care, moderate-to-severe cases among hospitalized patients who turn out to have cleared SARS-CoV-2 and have been discharged from hospitals, mild-to-moderate cases in outpatient clinics, and even children [11]. It has been reported that a substantial proportion of patients with COVID-19 continue to suffer from long-term adverse effects involving almost all bodily systems, including respiratory, gastrointestinal, neuropsychiatric, cardiovascular, and dermatological systems [12]. A recent study performed by our team showed that approximately 76% of patients with COVID-19 suffered from long COVID symptoms, including fatigue, poor memory, and hair loss, within 6 months of hospital admission [13].

The prevalence of long COVID symptoms appeared not only in Hong Kong and other Chinese cities but also across different countries [14]. For example, a study in Germany described different symptoms of long COVID, including headache, cough, shortness of breath, fatigue, dysosmia or anosmia, brain fog, or cognitive impairment [15]. Similar to these studies, the pattern of long COVID symptoms and their impacts have been investigated. For example, fatigue was the most common symptom in patients who had long COVID irrespective of the severity of their initial infection. Patients who had severe fatigue between 3 and 6 months might not encounter symptom improvement or resolution in the long term. Even those patients who had mild COVID-19 could develop long-term symptoms, including cough, fatigue, low-grade fever, shortness of breath, headaches, chest pain, muscle pains and weakness, metabolic disruption, neurocognitive difficulties, and gastrointestinal upset [3]. Long COVID symptoms could largely affect health-related quality of life and activities of daily living [16,17].

There could be many contributing factors that lead to prolonged patient recovery, such as deconditioning, relapse or reinfection, weak or absent antibody response, inflammatory and other immune reactions, and posttraumatic stress [3]. Risk factors that are associated with sociodemographic characteristics are important. These include gender, employment status, marital status, religion, multimorbidity, and living status [18-21].

Although the pathogenesis, complications, and epidemiological and clinical characteristics of patients with COVID-19 in the acute phase have been evaluated [22,23], few studies have characterized the epidemiology, symptomatology, and risk factors of long COVID. Previous studies examining these attributes are relatively few, mainly focused on hospitalized patients, or used a convenience sampling strategy [8]. In addition, the timing of assessment varied from 14 days to 6 months—or was not even reported—which is not consistent with the time frame of long COVID proposed by WHO, which is 3 months after COVID-19 infection [6]. In a recent review by Crook et al [8], 1 of the top priorities of research includes the precision epidemiology and risk factors of long COVID.

Hence, we aimed to fill these knowledge gaps by evaluating the prevalence of long COVID symptoms in 4 major Chinese cities using a representative sampling methodology. We also examined the distribution of COVID-19 symptoms and explored the risk factors of long COVID symptoms. We hypothesized that age, gender, the presence of chronic diseases, the use of chronic medications, the severity of COVID-19, and not receiving COVID-19 vaccines are associated with the occurrence of long COVID symptoms based on findings from recent studies [18,24,25].

Methods

Study Design

This was a cross-sectional survey of patients with COVID-19, aged ≥18 years, and residing in 4 cities of China: Beijing, Shanghai, Guangzhou, and Hong Kong. The study was launched on June 2, 2022, and closed on June 28, 2022. We used Qualtrics [26], an online survey platform. Inclusion criteria included age ≥18 years, with a diagnosis of COVID-19 from January 2020 to June 2022 in Beijing, Shanghai, Guangzhou, or Hong Kong. We excluded subjects without a history of COVID-19 infection. The survey collected the dates of COVID-19 vaccination, COVID-19 diagnosis, and recovery.

Ethical Considerations

The Survey and Behavioural Research Ethics Committee of the Chinese University of Hong Kong approved the study (approval no. SBRE-21-0730). All eligible participants provided digital informed consent via electronic means at the beginning of the survey. The survey did not collect personally identifiable information, and we encrypted email addresses collected for survey distribution as anonymized respondent identities. All respondents' data were recorded in an anonymous manner with a unique internally generated reference key. We assured the respondents of the anonymous, confidential nature of the study, where only aggregate data would be reported. We strictly adhered to the data privacy policy in the approved study protocol.

Survey Instrument

An expert panel consisting of epidemiologists, physicians, public health professionals, professors, and biostatisticians composed and validated the survey. The survey was pilot-tested to optimize coherence and clarity and was subsequently revised according to the pilot findings. During the curation of survey items, we referred to published studies that used validated survey instruments for long COVID symptoms [13,27]. The survey comprised 31 questions and required a median of 13 minutes to complete. The information sheet of the survey provided background information about COVID-19. The survey was delivered in traditional Chinese for participants in Hong Kong and simplified Chinese for respondents in the other 3 cities of mainland China. Before commencement of the survey, prospective participants confirmed that they were aged 18 years or older and had previously been diagnosed with COVID-19.

The survey collected respondents' sociodemographic information, past medical history, use of chronic medications, history of COVID-19 vaccines received, and severity of COVID-19. The survey contained several parts: First, we collected the respondents' sociodemographic information, including gender, marital status, employment status, types of jobs, educational level, living status, smoking habits, income level, self-reported health status, and long-term medication. In addition, we asked for the dates of positive and negative COVID-19 tests, the severity of COVID-19, the number of COVID-19 vaccines that were received, and the dates of taking vaccination. To identify the severity of COVID-19, we used the COVID-19 WHO severity classification [28] to assign respondents to 4 categories (mild, moderate, severe, and critical) based on the presence of pneumonia (eg, adults with clinical signs of pneumonia, respiratory rate>30 breaths/minute, severe respiratory distress, or SpO_2<90% on room air; WHO mentioned that a diagnosis can be made on clinical grounds with the assistance of chest imaging, such as radiograph, computed tomography (CT) scan, or ultrasound), hypoxia requiring oxygen, the need for mechanical ventilation, and a history of admission to the intensive care unit (ICU) due to COVID-19.

The respondents' medical history before and after the diagnosis of COVID-19 was also collected. Respondents were asked whether they had certain types of diseases before diagnosis (yes or no), the year of disease diagnosis, whether they took medications for the diseases (yes or no), and an open-ended field allowing them to document their medications. Medical conditions on the survey were infection, tumor, metabolic diseases, hematological disorders, mental illness, nervous system disease, circulatory system disease, respiratory disease, digestive disease, reproductive and urinary diseases, pregnancy complications, immune system disease, skin and subcutaneous tissue diseases, and musculoskeletal and connective tissue disease, which have been used in previous studies on the association between comorbidities and COVID-19 infection [29].

We also inquired about the presence of COVID-19 symptoms encountered by the respondents that were persistent for at least 3 months after COVID-19 infection. Referring to previous surveys [13,27], we listed 30 common symptoms: fatigue, fever, chills, inability to perform exercise, night sweats, hair loss, headache, dizziness, chest pain, rapid heartbeat, cough, sputum, sore throat, runny nose, dyspnea, arthralgia, myalgia, nausea, vomiting, diarrhea, abdominal pain, stomachache, anosmia, loss of taste, blurred vision, difficulty in concentration, difficulty in fall into asleep, feeling anxious, feeling sad, and memory problems. An open-ended field allowed the respondents to document other symptoms. These symptoms also appeared in "Symptoms of Infection with Coronavirus-19 (SIC)," which is a comprehensive patient-reported outcome measure developed to evaluate vaccines and treatment for COVID-19 [30]. To capture clinically meaningful symptoms, we assigned a 4-point Likert scale to assess the impact of the symptoms on the respondents' activities of daily living: 1 (no impact), 2 (little impact), 3 (moderate impact), and 4 (high impact). In addition, we classified the symptoms according to the bodily systems involved.

Sampling Strategy

The sample partners of Qualtrics randomly selected eligible respondents for the study, aiming for national representation through routers and a sophisticated application programming interface (API). To enhance the representativeness of the survey, we used a simple random sampling strategy to recruit potential participants by trying to match population demographics to our survey. The study collected data on the population of each city, and eligible participants were randomly selected and invited to complete the survey. The survey was accessed through the website, a smartphone, or other e-devices. Data scrubbing was subsequently performed after receiving all survey responses to remove unfavorable data, optimizing data accuracy and reliability. A more detailed survey methodology has been described elsewhere [26].

Outcome Variables and Sample Size

The primary outcome variable was the prevalence of long COVID symptoms, moderate or severe long COVID symptoms, and severe long COVID symptoms. We defined these 3 outcome variables as having at least 1 symptom with any degree of impact, at least 1 symptom with moderate or severe impact, and at least 1 symptom with severe impact, respectively. The secondary outcomes included the distribution of long COVID symptoms and their patterns. We assumed the proportion of the primary outcome as 50%, which provided a maximum sample size for each group. From the formula precision=1.96 × √[(p) × (1 – p)/N], where "p" refers to proportion of the primary outcome, a sample size of approximately 550 respondents would achieve a precision level of 0.04 for each city, so a total of 2200 respondents were required.

Statistical Analysis

We conducted a descriptive analysis of the prevalence and distribution of long COVID symptoms according to their severity. As the prevalence of self-reported long COVID symptoms was high, we also conducted a sensitivity analysis of increasing the number of long COVID symptoms. The proportion of long COVID symptoms, moderate or severe long COVID symptoms, and severe long COVID symptoms was computed for each city. To examine the risk factors of severe long COVID symptoms, we performed univariate and multivariate regression analyses with the primary outcome as a binary variable. The covariates included age, gender, marital status, current occupation, educational level, living status, smoking habits, monthly household income, self-perceived health status, the presence of chronic diseases, the use of chronic medication, COVID-19 vaccination status, and the severity of COVID-19. We classified the respondents’ occupations based on the recommendation by the Environmental Modelling Group (EMG) [31]. We entered all covariates with $P<.25$ in univariate analysis into the regression model and evaluated their statistical significance defined as $P<.05$. We tested for variable interaction and multicollinearity in the regression analysis.

Results

Participant Characteristics

We received 7161 total entries and screened out 3862 (53.9%) surveys due to ineligibility. A total of 830 (11.6%) surveys were overquota, and quality control was performed on 248 (3.5%) surveys. There were 4024 (56.2%) complete surveys, and 1312 (32.6%) were removed by data scrubbing. The total sample size achieved was 2712 (37.9%), with a response rate of 63.6%. The number of respondents from Beijing, Shanghai, Guangzhou, and Hong Kong was 725 (26.7%), 718 (26.5%), 715 (26.4%), and 554 (20.4%), respectively. Among them, the majority (n=2182, 80.5%) were aged 25-44 years, female (n=1626, 60.0%), married (n=2422, 89.3%), engaged in full-time jobs (n=2612, 96.3%), and with a tertiary or higher educational level (n=2547, 93.9%). Most respondents engaged in occupations at high risk for COVID-19, including the catering industry (n=533, 19.7%), followed by transportation, logistics, and the discipline workforce (n=344, 12.7%) and the health care sector (n=201, 7.4%) [18]. Most were living with their spouse (n=2400, 88.5%) or children (n=2230, 82.2%), were nonsmokers (n=1817, 67.0%), and were approximately equally distributed according to their monthly household income in the strata HK $10,000 to ≥$60,000 (US $1275.04-$7650.22). The majority perceived their health as good (n=994, 36.7%) or average (n=952, 35.1%), followed by very good (n=466, 17.2%), poor (n=256, 9.4%), and very poor (n=44, 1.6%). Around 2515 (92.7%) respondents did not use chronic medications, and 335 (12.3%) reported at least 2 chronic conditions before COVID-19 diagnosis (Table 1).

Table 1. Respondent characteristics (N=2712).

Characteristics	Respondents, n (%)
Age (years)	
18-24	139 (5.1)
25-34	1150 (42.4)
35-44	1032 (38.1)
45-54	302 (11.1)
55-64	71 (2.6)
≥65	18 (0.7)
Gender	
Male	1086 (10.2)
Female	1626 (60.0)
Marital status	
Single	277 (10.2)
Married	2422 (89.3)
Divorced	10 (0.4)
Widowed	3 (0.1)
Job status	
Full-time	2612 (96.3)
Part-time	52 (1.9)
Retired	19 (0.7)
Housewife	16 (0.6)
Others	13 (0.5)
Current occupation: higher-exposure group	
Catering industry	533 (19.7)
Health care sector	201 (7.4)
Transportation, logistics, discipline workforce	344 (12.7)
Current occupation: lower-exposure group	
Government or community	6 (0.2)
Finance/insurance/banking	53 (2.0)
IT/construction/manufacturing	189 (7.0)
Student	94 (3.5)
Unemployed/retired	177 (6.5)
Housewife	74 (2.7)
Educational level	
Primary or below	11 (0.4)
Secondary	154 (5.7)
Tertiary or above	2547 (93.9)
Living with[a]	
Children	2230 (82.2)
Spouse	2400 (88.5)
Domestic worker	44 (1.6)
Single	111 (4.1)
Others	180 (6.6)

Characteristics	Respondents, n (%)
Smoking status	
Nonsmoker	1817 (67.0)
Ex-smoker quitted for ≥1 year	709 (26.1)
Smoker	186 (6.9)
Monthly household income (HK $/US $[b])	
<10,000/<1275.04	128 (4.7)
10,000-19,999/1275.04-2549.95	584 (21.5)
20,000-29,999/2550.07-3824.98	747 (27.5)
30,000-59,999/3825.11-7650.09	697 (25.7)
≥60,000/≥7650.22	542 (20.0)
Receiving a living allowance	6 (0.2)
Refused to answer	8 (0.3)
Self-perceived health status	
Very good	466 (17.2)
Good	994 (36.7)
Average	952 (35.1)
Poor	256 (9.4)
Very poor	44 (1.6)
Chronic medication use	
No	2515 (92.7)
Yes	197 (7.3)
Number of chronic conditions before COVID-19 confirmed	
0	1988 (73.3)
1	389 (14.3)
2	90 (3.3)
≥3	245 (9.0)

[a] The total proportion was more than 100% as we allowed multiple responses.

[b] HK $1=US $0.13.

Severity of COVID-19

Most respondents suffered from pneumonia at the time of COVID-19 diagnosis (n=2141, 78.9%), required hospital admission for COVID-19 management (n=1944, 71.7%), needed oxygen due to COVID-19 (n=1427, 52.6%), and received antiviral agents (n=1747, 64.4%); see Multimedia Appendix 1, Table S1. A significant proportion necessitated artificial ventilation (n=986, 36.4%) and admission to the ICU (n=940, 34.7%). Up to 1552 (57.2%) of them had severe or critical COVID-19, and 1160 (42.8%) had mild or moderate severity of COVID-19. Most of the respondents had received vaccination before COVID-19 diagnosis, with the majority having taken 3 or more doses of Coronavac-Sinovac (n=1525, 56.2%), following by those who had taken 2 doses of Coronavac-Sinovac (n=786, 29.0%). There were about 169 (6.2%) respondents who had taken 2 or more doses of Pfizer-BioNTech.

Long COVID Symptoms

Multimedia Appendix 1, Table S2, shows the distribution of long COVID symptoms. The highest proportion of respondents suffered from fatigue (n=914, 33.7%), cough (n=865, 31.9%), sore throat (n=841, 31.0%), difficulty in concentrating (n=828, 30.5%), feeling of anxiety (n=817, 30.2%), myalgia (n=811, 29.9%), arthralgia (n=811, 29.9%), sputum production (n=789, 29.4%), and difficulty in falling asleep (n=792, 29.2%); see Figure 1. The prevalence of having long COVID symptoms of any severity, moderate or high severity, and high severity was 90.4% (n=2452), 62.4% (n=1692), and 31.0% (n=841), respectively. The symptoms were qualified in terms of self-perceived severity, and sensitivity analysis of using 2 or more symptoms showed a better estimate of long COVID syndrome/disorder (Figure 2). These prevalence figures were the highest in Hong Kong (n=541, 97.7%, n=385, 69.5%, and n=194, 35.0%, respectively) and Shanghai (n=662, 92.2%, n=461, 64.2%, and n=287, 40.0%, respectively), followed by Guangzhou (n=630, 88.1%, n=451, 63.1%, and n=212, 29.7%, respectively) and Beijing (n=619, 85.4%, n=395, 54.5%, and n=148, 20.4%, respectively); see Figure 3. The prevalence figures were the highest in the second wave (from July to December 2020; n=67, 91.8%, n=48, 65.8%, and n=16, 21.9%, respectively) and the third wave (from January to June 2021; n=453, 91.7%, n=329, 66.6%, and n=175, 35.4%, respectively), followed by the first wave (from January to June 2020; n=122, 90.4%, n=64, 47.4%, and n=36, 26.7%, respectively), the fifth wave (from January to June 2022; n=1446, 90.1%, n=998, 62.2%, and n=503, 31.3%, respectively), and the fourth wave (from July to December 2021; n=364, 89.9%, n=253, 62.5%, and n=111, 27.4%, respectively); see Multimedia Appendix 1, Figure S1. A slightly negative correlation was found between the timing of COVID-19 and the number of long COVID symptoms (β=–.046, *P*=.016), supporting that long COVID symptom numbers decline with a longer duration of follow-up.

Figure 1. Distribution of long COVID symptoms. The highest proportion of respondents suffered from fatigue (n=914, 33.7%), cough (n=865, 31.9%), sore throat (n=841, 31.0%), difficulty in concentrating (n=828, 30.5%), feeling of anxiety (n=817, 30.2%), myalgia (n=811, 29.9%), arthralgia (n=811, 29.9%), sputum production (n=789, 29.4%), and difficulty in falling asleep (n=792, 29.2%).

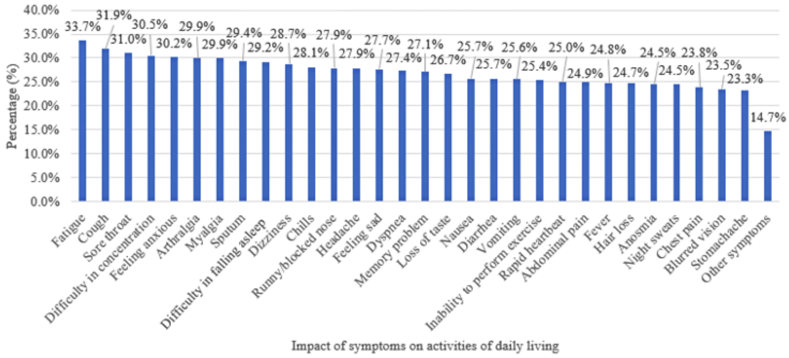


Figure 2. The prevalence of having long COVID symptoms of any severity, moderate or high severity, and high severity was 90.4% (n=2452), 62.4% (n=1692), and 31.0% (n=841), respectively. The symptoms were qualified in terms of self-perceived severity, and sensitivity analysis of using more than 1 symptom showed a better estimate of long COVID syndrome/disorder.

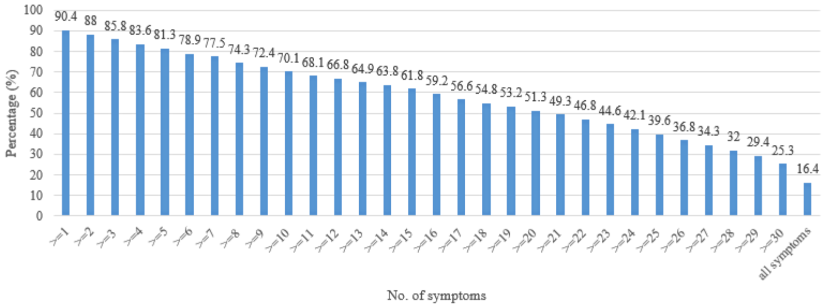
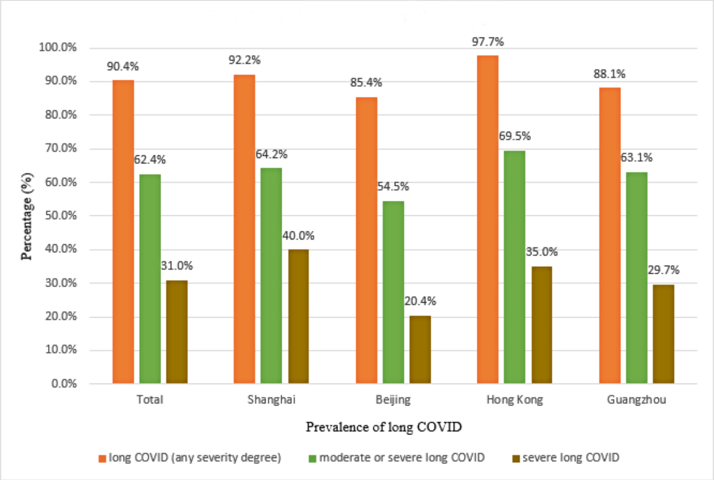


Figure 3. Prevalence of long COVID with severity. These prevalence figures were the highest in Hong Kong (n=541, 97.7%, n=385, 69.5%, and n=194, 35.0%, respectively) and Shanghai (n=662, 92.2%, n=461, 64.2%, and n=287, 40.0%, respectively), followed by Guangzhou (n=630, 88.1%, n=451, 63.1%, and n=212, 29.7%, respectively) and Beijing (n=619, 85.4%, n=395, 54.5%, and n=148, 20.4%, respectively).



Factors Associated With Long COVID Symptoms of Moderate-to-High Severity

From multivariate regression analysis (Table 2), female subjects (adjusted odds ratio [aOR]=1.49, 95% CI 1.13-1.95); patients engaged in transportation, logistics, or the discipline workforce (aOR=2.52, 95% CI 1.58-4.03); patients living with domestic workers (aOR=2.37, 95% CI 1.39-4.03); smokers/ex-smokers (aOR=1.55, 95% CI 1.17-2.05); respondents with poor or very poor self-perceived health status (aOR=15.4, 95% CI 7.88-30.00); those with more chronic diseases (1 chronic disease: aOR=1.92, 95% CI 1.24-2.97; 2 chronic diseases: aOR=2.71, 95% CI 1.21-6.05; ≥3 chronic diseases: aOR=2.71, 95% CI 1.54-4.79); chronic medication use (aOR=4.38, 95% CI 1.66-11.53); and critical severity of COVID-19 (aOR=1.52, 95% CI 1.07-2.15) were associated with the development of severe long COVID. Prior COVID-19 vaccination with Coronavac-Sinovac (aOR=0.35, 95% CI 0.14-0.90, *P*=.03) or Pfizer-BioNTech (aOR=0.22, 95% CI 0.08-0.63, *P*=.005) for at least 2 doses was a protective factor.

Table 2. Risk factors of severe long COVID (N=1030).

Characteristics	Respondents, n (%)	cOR[a] (95% CI)	*P* value	aOR[b] (95% CI)	*P* value
Age (years)					
35-44	417 (60.5)	1 (reference)	.167	1 (reference)	.337
18-24	57 (71.3)	1.62 (0.97-2.69)	.064	1.15 (0.59-2.22)	.686
25-34	372 (58.5)	0.92 (0.74-1.14)	.452	1.09 (0.84-1.42)	.508
45-54	138 (63.6)	1.14 (0.83-1.56)	.418	0.71 (0.47-1.08)	.108
≥55	46 (65.7)	1.25 (0.75-2.10)	.397	1.20 (0.65-2.20)	.567
Gender					
Male	357 (57.3)	1 (reference)	N/A[c]	1 (reference)	N/A
Female	673 (63.0)	1.27 (1.04-1.55)	.022	1.49 (1.13-1.95)	.004
Marital status					
Married	925 (60.8)	1 (reference)	N/A	N/A	N/A
Single/divorced/widowed	105 (61.8)	1.04 (0.75-1.44)	.802	N/A	N/A
Current occupation: higher-exposure group					
Catering industry	206 (63.4)	1 (reference)	<.001	1 (reference)	<.001
Health care sector	63 (59.4)	0.85 (0.54-1.33)	.466	0.98 (0.58-1.65)	.941
Transportation, logistics, discipline workforce	205 (84.0)	3.04 (2.01-4.58)	<.001	2.52 (1.58-4.03)	<.001
Current occupation: lower-exposure group					
Government or community	3 (60.0)	0.87 (0.14-5.26)	.876	0.39 (0.04-3.52)	.405
Finance/insurance/banking	18 (64.3)	1.04 (0.46-2.33)	.924	1.32 (0.54-3.23)	.542
IT/construction/manufacturing	44 (50.0)	0.58 (0.36-0.93)	.024	0.36 (0.20-0.64)	.001
Student	27 (67.5)	1.20 (0.60-2.41)	.610	1.10 (0.45-2.70)	.831
Unemployed/retired	56 (61.5)	0.92 (0.57-1.49)	.747	0.60 (0.34-1.06)	.078
Housewife	30 (63.8)	1.02 (0.54-1.93)	.953	1.07 (0.52-2.19)	.851
Educational level					
Secondary or below	73 (70.2)	1 (reference)	N/A	1 (reference)	N/A
Tertiary or above	957 (60.3)	0.64 (0.42-0.99)	.046	0.62 (0.37-1.05)	.073
Living status					
Children	856 (61.2)	1.09 (0.84-1.41)	.513	1 (reference)	N/A
Spouse	916 (60.9)	1.00 (0.73-1.36)	.979	N/A	N/A
Single	63 (56.8)	0.83 (0.56-1.23)	.358	N/A	N/A
Domestic workers/others	70 (70.0)	1.54 (0.99-2.38)	.055	2.37 (1.39-4.03)	.001
Smoking status					
Nonsmoker	618 (55.0)	1 (reference)	N/A	1 (reference)	N/A
Smoker	412 (72.5)	2.16 (1.74-2.69)	<.001	1.55 (1.17-2.05)	.002
Monthly household income (HK $/US $[d])					
<10,000/<1275.04 or Comprehensive Social Security Assistance (CSSA)/Disability Allowance (DA)	44 (69.8)	1 (reference)	.011	1 (reference)	.060
10,000-19,999/1275.04-2549.95	218 (56.6)	0.56 (0.32-1.00)	.051	0.92 (0.47-1.82)	.811
20,000-29,999/2550.07-3824.98	264 (56.8)	0.57 (0.32-1.00)	.051	1.10 (0.55-2.17)	.790
30,000-59,999/3825.11-7650.09	244 (64.7)	0.79 (0.44-1.41)	.430	1.55 (0.78-3.11)	.214
≥60,000/≥7650.22	260 (64.7)	0.79 (0.44-1.41)	.424	1.25 (0.62-2.53)	.526
Self-perceived health status					

Characteristics	Respondents, n (%)	cOR[a] (95% CI)	*P* value	aOR[b] (95% CI)	*P* value
Very good	109 (39.6)	1 (reference)	<.001	1 (reference)	<.001
Good	279 (46.3)	1.32 (0.98-1.76)	.064	1.51 (1.09-2.09)	.013
Average	467 (74.5)	4.45 (3.29-6.01)	<.001	5.06 (3.58-7.13)	<.001
Poor/very poor	175 (93.1)	20.5 (11.1-37.9)	<.001	15.4 (7.88-30.0)	<.001
Presence of chronic diseases					
0	711 (54.5)	1 (reference)	<.001	1 (reference)	<.001
1	117 (74.5)	2.44 (1.68-3.55)	<.001	1.92 (1.24-2.97)	.003
2	37 (80.4)	3.43 (1.64-7.16)	.001	2.71 (1.21-6.05)	.015
≥3	165 (89.2)	6.88 (4.27-11.08)	<.001	2.71 (1.54-4.79)	.001
Chronic medication use					
No	955 (59.2)	1 (reference)	N/A	1 (reference)	N/A
Yes	75 (93.8)	10.3 (4.15-25.7)	<.001	4.38 (1.66-11.5)	.003
Vaccination status					
No vaccine taken	30 (83.3)	1 (reference)	.195	1 (reference)	.073
Coronavac-Sinovac, 1 dose	40 (63.5)	0.35 (0.13-0.96)	.042	0.36 (0.12-1.11)	.075
Coronavac-Sinovac, ≥2 doses	877 (60.4)	0.31 (0.13-0.74)	.008	0.35 (0.14-0.90)	.030
Pfizer-BioNTech, 1 dose	3 (60.0)	0.30 (0.04-2.20)	.236	0.17 (0.02-1.77)	.139
Pfizer-BioNTech, ≥2 doses	59 (59.0)	0.29 (0.11-0.75)	.011	0.22 (0.08-0.63)	.005
Others	21 (58.3)	0.28 (0.09-0.84)	.023	0.22 (0.07-0.76)	.016
Severity of COVID-19					
Mild	134 (47.9)	1 (reference)	<.001	1 (reference)	.009
Moderate	286 (54.1)	1.28 (0.96-1.71)	.093	0.96 (0.69-1.34)	.811
Severe	191 (66.3)	2.15 (1.53-3.01)	<.001	1.34 (0.90-1.98)	.150
Critical	419 (70.4)	2.59 (1.93-3.48)	<.001	1.52 (1.07-2.15)	.018

[a] cOR: crude odds ratio.

[b] aOR: adjusted odds ratio.

[c] N/A: not applicable.

[d] HK $1=US $0.13.

Discussion

Principal Findings

This multicenter survey involving 4 Chinese cities showed that the prevalence of long COVID symptoms was high for symptoms of any severity, moderate or high severity, and high severity. We reported the most common long COVID symptoms and their distribution, as well as the risk factors of long COVID symptoms. Notably, COVID-19 vaccination was associated with a protective effect on the development of long COVID symptoms.

The prevalence of long COVID symptoms was high across different cities. Several studies have reported different incidence rates of long COVID symptoms at different observation periods, including 76% of patients at 6 months [5], 32.6%-87.0% of patients at 60 days [32,33], and 96% of patients at 90 days [34]. A study conducted in the Chinese city of Wuhan included 2469 patients who had confirmed COVID-19 and were discharged from the hospital between January and May 2020. Those who required high-flow nasal cannula (HFNC), noninvasive mechanical ventilation (NIV), or invasive mechanical ventilation (IMV) were more likely to suffer from long COVID symptoms and complications that affected their quality of life, including limitations in their mobility, usual activities, and mental health [14]. These findings were compatible with our results in which patients with critical severity of COVID-19 were more likely to have severe long COVID symptoms. The prevalence of long COVID symptoms was also common in other countries. For instance, a study found that 62% of 89 interviewed patients in the United Kingdom had long COVID symptoms for over 3 months, 52% for 6 months, and 49% for 9 months after their hospital admission [35]. In France, 51% of patients presented with at least 1 symptom after their COVID-19 diagnosis. Regarding severity, 20% of the patients required ICU care, including IMV, vasopressors, and extracorporeal membrane oxygenation during COVID-19 [31]. By using the same WHO definition of COVID-19 severity, a study in Turkey reported 47.5% of the participants suffering from 1 or more persistent symptoms, including outpatient and inpatient clinics, irrespective of disease severity. Most studies have reported high prevalence rates of long COVID symptoms, including Spain, Bangladesh, the United Kingdom, the United States, Nigeria, and Denmark [5].

In addition, research on the prevalence of long COVID symptoms has begun. Apart from respiratory symptoms, fatigue and neuropsychiatric symptoms have been the most frequently reported manifestations of long COVID. The UK Office for National Statistics (ONS) estimated the 5-week prevalence of fatigue to be 11.9% among patients with COVID-19 [36]. One cross-sectional study concluded that 92.9% and 93.5% of hospitalized and nonhospitalized patients with COVID-19, respectively, suffered from ongoing fatigue for 79 days following illness onset [17]. The exact mechanism of its appearance after COVID-19 is currently speculative. This has been attributed to dysfunctional inflammatory response pathways [37]. In addition, a repertoire of central, peripheral, and psychological factors might play a role [8], including congestion of the glymphatic system, hypometabolism in the frontal lobe and cerebellum, and direct SARS-CoV-2 infection of skeletal muscle fibres and neuromuscular junctions [8]. With regard to neuropsychiatric symptoms, the underlying mechanism may be related to glial cell activation, which damages neurons, in addition to hyperinflammatory and hypercoagulable states leading to increased risk of thrombotic events. These could lead to infiltration of blood-derived substances and leukocytes to the brain parenchyma [8].

The risk factors of long COVID symptoms have not been extensively studied. In some studies, certain factors that increase the risk of COVID-19 do not seem to increase the risk of long COVID [8]. For instance, although the male sex has been found to be a risk factor for contracting COVID-19, the ONS reported that the prevalence of any long COVID symptoms is higher in women than that in men (23.6% vs 20.7%) [36]. This observation is consistent with our findings where the female gender was associated with a significantly higher risk of severe long COVID symptoms.

Across different regions, studies have shown that particular occupational groups may be at higher risk of contracting COVID-19, which in turn may affect the risk of long COVID symptoms. High-risk occupations included those engaged in health care, public service, public transportation, material moving, elementary services, and other essential sectors with workers reporting higher rates of mortality. A study conducted across 46 states in the United States found that the per capita age-standardized mortality rate among essential workers was 30.4 per 100,000 individuals compared to workers in nonessential industries (15.5 per 100,000 individuals) [38]. As the risk of long COVID symptoms was associated with employment in transportation, logistics, and the discipline workforce in this study, this could be attributed to increased exposure to SARS-CoV-2, leading to a higher likelihood of experiencing persisting symptoms.

Living status, which was also regarded as an environmental risk factor, has also been examined by other studies on COVID-19 infection and long COVID symptoms. Despite a poor living environment, such as low air quality and transportation insecurity, environmental exposure to social contact that was affected by sociodemographic characteristics is also associated with the risk of COVID-19 infection. This might increase the risk of suffering from long COVID symptoms. These exposures include cohabitation and living in overcrowded housing [18]. There is a lack of studies that investigate the relationship between COVID-19 infection or long COVID symptoms and the living status of domestic workers, while our study showed that it is a significant risk factor. This may be due to environmental exposure to social contact.

Previous meta-analyses have demonstrated an increased risk of severe COVID-19 and death in current and former smokers compared to nonsmokers [39,40]. However, smoking is not a confirmed risk factor for long COVID symptoms and has only been shown to be a predictor (P>.001) of longer symptom duration [41]. Our findings also demonstrated that smoking prevalence is associated with a higher risk of long COVID symptoms. Smoking may potentially be a risk factor for some long COVID symptoms as smokers tend to have a greatly weakened immune and cardiovascular system, thus increasing their susceptibility to various health complications.

Multimorbidity with pre-existing health conditions [42], particularly asthma [43], has displayed a heightened risk of developing persistent symptoms and increased risk of COVID-19 fatality [44]. Multimorbidity has been positively associated with medication use, with 50% of adults in high-income countries taking up to 5 medications or more [45,46]. Poorer perceived health status may be considered a risk factor for long COVID symptoms, as observed in a study wherein 87% of the participants had a good self-reported health status before the pandemic but 83.3% reported moderate-to-poor self-reported health 6 months after initial onset [24]. This is further supported by a cohort study conducted in the Netherlands 3 months following recovery, where the health status of patients with COVID-19 was generally reported to be poor, with significant impairment in the domains of functional impairment (64%), fatigue (69%), and quality of life (72%) [47].

COVID-19 vaccination, including Coronavac-Sinovac and Pfizer-BioNTech, had protective effect on the development of long COVID symptoms. A community-based study in the United Kingdom that included 6729 patients with any severity of long COVID symptoms found that vaccine doses are associated with significant reduction in experiencing long COVID symptoms (first vaccine dose: 12.8% decrease, P<.001; second vaccine dose: 8.8% decrease, P=.003) [48]. Vaccination is able to prevent reinfection among patients with long COVID symptoms, as persons who take a single vaccine dose have equal or higher antibody titers than those who do not take any vaccination, which is consistent with our study findings [25]. The impact of vaccination on long COVID symptoms could vary among respondents or between different vaccine types; however, the vaccinated population does have a lower chance of suffering from long COVID symptoms compared with those who have not been vaccinated [25,48,49].

Strengths and Limitations

This study was relatively large scale as compared to other surveys, and it adopted a representative sampling strategy involving 4 major cities in China. The response rate was high relative to other similar studies, thus enhancing its generalizability to other settings. Nevertheless, several limitations should be addressed. First, its cross-sectional nature did not allow a cause-and-effect relationship to be established due to the possibility of reverse causality. However, the objective of this study was to identify independent risk factors instead of causes of long COVID. In addition, there existed potential recall bias as some study participants self-reported the presence of long COVID symptoms, especially those with an earlier diagnosis of COVID-19. It should also be noted that the definition of long COVID is different across various authorities, with the National Institute for Health and Care Excellence (NICE) and the US Centers for Disease Control and Prevention (CDC) using 4-12 weeks and 4 weeks, respectively, as the time frame of its persistent symptoms. Furthermore, there could be residual confounders in our multivariate regression analysis, and these could vary according to different cities where the management strategies for COVID-19 might be different. In particular, it is unknown whether the study participants received certain treatments to relieve their long COVID symptoms at the initial stage of the development, thus reducing their risk of long COVID symptoms. Lastly, as in most surveys, not all potential study participants in the sampling frame were eligible before the recruitment process. In addition, the older population who may have a substantially higher burden of both COVID-19 and long COVID symptoms were underrepresented in this study. As the survey had to be completed through e-devices, relatively fewer older people were included as a higher proportion of them did not know how to use e-devices. Future study on the prevalence of long COVID in the older population is warranted.

Conclusion

We examined the prevalence of long COVID symptoms in 4 Chinese cities and the effect of the severity of COVID-19. We also examined the pattern of long COVID symptoms, as well as the risk factors, including gender, occupational groups, living status, lifestyle, multimorbidity, and vaccination status. Fatigue, cough, sore throat, difficulty in concentration, feeling of anxiety, myalgia, and arthralgia are the most common severe long COVID symptoms. The female gender; engagement in transportation, logistics, or the discipline workforce; living with domestic workers; smoking; poor self-perceived health status; chronic diseases; chronic medication use; and critical severity of COVID-19 are associated with severe long COVID. Prior vaccination with ≥2 doses of COVID-19 vaccines is a protective factor. Our findings may inform early identification of patients with COVID-19 at risk of long COVID and planning of rehabilitative services.

Acknowledgments

All authors were responsible for the study concept and design. MC-SW, JH, and Y-YW were responsible for data acquisition and analysis, had full access to all of the data in the study, and take responsibility for data integrity and the accuracy of the data analysis. All authors were responsible for the interpretation of data, drafting, and critical revision of the manuscript for important intellectual content.

This work was supported in part by the Health and Medical Research Fund (HMRF) – Food and Health Bureau Commissioned Research on COVID-19 (project reference number: COVID1903002) to FK-LC and MC-SW. The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

S-CN and FK-LC are scientific cofounders of Geniebiome Ltd.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File , 27 KB - publichealth_v9i1e42315_app1.docx]

References

1. WHO coronavirus (COVID-19) dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2022-07-09]
2. How long does COVID-19 last? Kings College, London, 2020. ZOE Limited. URL: https://covid19.joinzoe.com/post/covid-long-term?fbclid=IwAR1RxIcmmdL-EFjh_aI- [accessed 2023-02-23]
3. Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. BMJ 2020 Aug 11;370:m3026. [doi: 10.1136/bmj.m3026] [Medline: 32784198]
4. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr;27(4):601-615 [FREE Full text] [doi: 10.1038/s41591-021-01283-z] [Medline: 33753937]

5. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet 2021 Jan 16;397(10270):220-232 [FREE Full text] [doi: 10.1016/S0140-6736(20)32656-8] [Medline: 33428867]
6. Post COVID-19 condition (long COVID). World Health Organization. URL: https://www.who.int/srilanka/news/detail/16-10-2021-post-covid-19-condition [accessed 2023-02-23]
7. Ladds E, Rushforth A, Wieringa S, Taylor S, Rayner C, Husain L, et al. Persistent symptoms after Covid-19: qualitative study of 114 "long Covid" patients and draft quality principles for services. BMC Health Serv Res 2020 Dec 20;20(1):1144 [FREE Full text] [doi: 10.1186/s12913-020-06001-y] [Medline: 33342437]
8. Crook H, Raza S, Nowell J, Young M, Edison P. Long covid-mechanisms, risk factors, and management. BMJ 2021 Jul 26;374:n1648. [doi: 10.1136/bmj.n1648] [Medline: 34312178]
9. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV, WHO Clinical Case Definition Working Group on Post-COVID-19 Condition. A clinical case definition of post-COVID-19 condition by a Delphi consensus. Lancet Infect Dis 2022 Apr;22(4):e102-e107. [doi: 10.1016/s1473-3099(21)00703-9]
10. Mikuteit M, Heinemann S, Roder S, Niewolik J, Schröder D, Vahldiek K, et al. Long-term consequences of COVID-19 and the pandemic: protocol for a web-based, longitudinal observational study (defeat). JMIR Res Protoc 2022 Oct 26;11(10):e38718 [FREE Full text] [doi: 10.2196/38718] [Medline: 36108134]
11. Yong SJ. Long COVID or post-COVID-19 syndrome: putative pathophysiology, risk factors, and treatments. Infect Dis (Lond) 2021 Oct 22;53(10):737-754 [FREE Full text] [doi: 10.1080/23744235.2021.1924397] [Medline: 34024217]
12. Baig AM. Chronic COVID syndrome: need for an appropriate medical terminology for long-COVID and COVID long-haulers. J Med Virol 2021 May;93(5):2555-2556. [doi: 10.1002/jmv.26624] [Medline: 33095459]
13. Liu Q, Mak JWY, Su Q, Yeoh YK, Lui GC, Ng SSS, et al. Gut microbiota dynamics in a prospective cohort of patients with post-acute COVID-19 syndrome. Gut 2022 Mar 26;71(3):544-552 [FREE Full text] [doi: 10.1136/gutjnl-2021-325989] [Medline: 35082169]
14. COVID-19 rapid guideline: managing the long-term effects of COVID-19 NICE guideline. National Institute for HealthCare Excellence. URL: https://www.nice.org.uk/guidance/ng188 [accessed 2022-08-13]
15. Koczulla AR, Ankermann T, Behrends U, Berlit P, Böing S, Brinkmann F, et al. [S1 guideline post-COVID/long-COVID]. Pneumologie 2021 Nov 02;75(11):869-900 [FREE Full text] [doi: 10.1055/a-1551-9734] [Medline: 34474488]
16. Van Herck M, Goërtz YMJ, Houben-Wilke S, Machado FVC, Meys R, Delbressine JM, et al. Severe fatigue in long COVID: web-based quantitative follow-up study in members of online long COVID support groups. J Med Internet Res 2021 Sep 21;23(9):e30274 [FREE Full text] [doi: 10.2196/30274] [Medline: 34494964]
17. Goërtz YMJ, Van Herck M, Delbressine JM, Vaes AW, Meys R, Machado FV, et al. Persistent symptoms 3 months after a SARS-CoV-2 infection: the post-COVID-19 syndrome? ERJ Open Res 2020 Oct 10;6(4):00542-2020 [FREE Full text] [doi: 10.1183/23120541.00542-2020] [Medline: 33257910]
18. Rozenfeld Y, Beam J, Maier H, Haggerson W, Boudreau K, Carlson J, et al. A model of disparities: risk factors associated with COVID-19 infection. Int J Equity Health 2020 Jul 29;19(1):126 [FREE Full text] [doi: 10.1186/s12939-020-01242-z] [Medline: 32727486]
19. Meza-Torres B, Delanerolle G, Okusi C, Mayor N, Anand S, Macartney J, et al. Differences in clinical presentation with long COVID after community and hospital infection and associations with all-cause mortality: English Sentinel Network Database Study. JMIR Public Health Surveill 2022 Aug 16;8(8):e37668 [FREE Full text] [doi: 10.2196/37668] [Medline: 35605170]
20. Laires PA, Dias S, Gama A, Moniz M, Pedro AR, Soares P, et al. The association between chronic disease and serious COVID-19 outcomes and its influence on risk perception: survey study and database analysis. JMIR Public Health Surveill 2021 Jan 12;7(1):e22794 [FREE Full text] [doi: 10.2196/22794] [Medline: 33433397]
21. Yanover C, Mizrahi B, Kalkstein N, Marcus K, Akiva P, Barer Y, et al. What factors increase the risk of complications in SARS-CoV-2-infected patients? A cohort study in a nationwide Israeli health organization. JMIR Public Health Surveill 2020 Aug 25;6(3):e20872 [FREE Full text] [doi: 10.2196/20872] [Medline: 32750009]
22. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. JAMA 2020 Aug 25;324(8):782-793. [doi: 10.1001/jama.2020.12839] [Medline: 32648899]
23. Cevik M, Kuppalli K, Kindrachuk J, Peiris M. Virology, transmission, and pathogenesis of SARS-CoV-2. BMJ 2020 Oct 23;371:m3862. [doi: 10.1136/bmj.m3862] [Medline: 33097561]
24. Vaes AW, Goërtz YMJ, Van Herck M, Machado FVC, Meys R, Delbressine JM, et al. Recovery from COVID-19: a sprint or marathon? 6-month follow-up data from online long COVID-19 support group members. ERJ Open Res 2021 Apr 26;7(2):00141-2021 [FREE Full text] [doi: 10.1183/23120541.00141-2021] [Medline: 34041295]
25. Tsuchida T, Hirose M, Inoue Y, Kunishima H, Otsubo T, Matsuda T. Relationship between changes in symptoms and antibody titers after a single vaccination in patients with Long COVID. J Med Virol 2022 Jul 08;94(7):3416-3420 [FREE Full text] [doi: 10.1002/jmv.27689] [Medline: 35238053]
26. Survey methodology and compliance best practices. Qualtrics. URL: https://www.qualtrics.com/support/survey-platform/survey-module/survey-checker/survey-methodology-compliance-best-practices/ [accessed 2022-08-13]

27. Hughes SE, Haroon S, Subramanian A, McMullan C, Aiyegbusi OL, Turner GM, et al. Development and validation of the symptom burden questionnaire for long covid (SBQ-LC): Rasch analysis. BMJ 2022 Apr 27;377:e070230 [FREE Full text] [doi: 10.1136/bmj-2022-070230] [Medline: 35477524]
28. Clinical management of COVID-19: interim guidance, 27 May 2020. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/332196 [accessed 2022-08-13]
29. Lau LH, Wong SH, Yip TCF, Wong GLH, Wong VWS, Sung JJY. Collateral effect of coronavirus disease 2019 pandemic on hospitalizations and clinical outcomes in gastrointestinal and liver diseases: a territory-wide observational study in Hong Kong. Gastroenterology 2020 Nov;159(5):1979-1981.e3 [FREE Full text] [doi: 10.1053/j.gastro.2020.07.042] [Medline: 32721440]
30. Romano C, Fehnel S, Stoddard J, Sadoff J, Lewis S, McNulty P, et al. Development of a novel patient-reported outcome measure to assess signs and symptoms of COVID-19. J Patient Rep Outcomes 2022 Jul 29;6(1):85 [FREE Full text] [doi: 10.1186/s41687-022-00471-w] [Medline: 35904710]
31. COVID-19 risk by occupation and workplace. EMG – Transmission Group. 2021 Feb 11. URL: https://tinyurl.com/ytvtardd [accessed 2023-02-25]
32. Chopra V, Flanders SA, O'Malley M, Malani AN, Prescott HC. Sixty-day outcomes among patients hospitalized with COVID-19. Ann Intern Med 2021 Apr;174(4):576-578. [doi: 10.7326/m20-5661]
33. Carfì A, Bernabei R, Landi F, Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent symptoms in patients after acute COVID-19. JAMA 2020 Aug 11;324(6):603-605 [FREE Full text] [doi: 10.1001/jama.2020.12603] [Medline: 32644129]
34. Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine 2021 Aug;38:101019 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101019] [Medline: 34308300]
35. Datta SD, Talwar A, Lee JT. A proposed framework and timeline of the spectrum of disease due to SARS-CoV-2 infection: illness beyond acute infection and public health implications. JAMA 2020 Dec 08;324(22):2251-2252. [doi: 10.1001/jama.2020.22717] [Medline: 33206133]
36. Prevalence of long COVID symptoms and COVID-19 complications. UK Office for National Statistics. 2020. URL: https://tinyurl.com/ymths8aj [accessed 2022-08-13]
37. Islam MF, Cotler J, Jason LA. Post-viral fatigue and COVID-19: lessons from past epidemics. Fatigue: Biomed Health Behav 2020 Jun 25;8(2):61-69. [doi: 10.1080/21641846.2020.1778227]
38. Chen Y, Chen R, Charpignon M, Kiang M, Riley A, Glymour M. COVID-19 mortality among working-age Americans in 46 states, by industry and occupation. medRxiv 2022.03.29.2227 2022. [doi: 10.1101/2022.03.29.22273085]
39. Reddy RK, Charles WN, Sklavounos A, Dutt A, Seed PT, Khajuria A. The effect of smoking on COVID-19 severity: a systematic review and meta-analysis. J Med Virol 2021 Feb 13;93(2):1045-1056 [FREE Full text] [doi: 10.1002/jmv.26389] [Medline: 32749705]
40. Umnuaypornlert A, Kanchanasurakit S, Lucero-Prisno DE, Saokaew S. Smoking and risk of negative outcomes among COVID-19 patients: a systematic review and meta-analysis. Tob Induc Dis 2021 Feb 4;19(February):09-13 [FREE Full text] [doi: 10.18332/tid/132411] [Medline: 33551713]
41. Hossain MA, Hossain KMA, Saunders K, Uddin Z, Walton LM, Raigangar V, et al. Prevalence of long COVID symptoms in Bangladesh: a prospective inception cohort study of COVID-19 survivors. BMJ Glob Health 2021 Dec 14;6(12):e006838 [FREE Full text] [doi: 10.1136/bmjgh-2021-006838] [Medline: 34906986]
42. Stavem K, Ghanima W, Olsen MK, Gilboe HM, Einvik G. Persistent symptoms 1.5-6 months after COVID-19 in non-hospitalised subjects: a population-based cohort study. Thorax 2021 Apr 03;76(4):405-407 [FREE Full text] [doi: 10.1136/thoraxjnl-2020-216377] [Medline: 33273028]
43. Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. Nat Med 2021 Apr;27(4):626-631 [FREE Full text] [doi: 10.1038/s41591-021-01292-y] [Medline: 33692530]
44. Treskova-Schwarzbach M, Haas L, Reda S, Pilic A, Borodova A, Karimi K, et al. Pre-existing health conditions and severe COVID-19 outcomes: an umbrella review approach and meta-analysis of global evidence. BMC Med 2021 Aug 27;19(1):212 [FREE Full text] [doi: 10.1186/s12916-021-02058-6] [Medline: 34446016]
45. Ecks S. Multimorbidity, polyiatrogenesis, and COVID-19. Med Anthropol Q 2020 Dec 04;34(4):488-503 [FREE Full text] [doi: 10.1111/maq.12626] [Medline: 33274530]
46. Bustos-Vázquez E, Padilla-González E, Reyes-Gómez D, Carmona-Ramos MC, Monroy-Vargas JA, Benítez-Herrera AE, et al. Survival of COVID-19 with multimorbidity patients. Healthcare (Basel) 2021 Oct 22;9(11):1423 [FREE Full text] [doi: 10.3390/healthcare9111423] [Medline: 34828470]
47. van den Borst B, Peters J, Brink M, Schoon Y, Bleeker-Rovers C, Schers H, et al. Comprehensive health assessment 3 months after recovery from acute coronavirus disease 2019 (COVID-19). Clin Infect Dis 2021 Sep 07;73(5):e1089-e1098 [FREE Full text] [doi: 10.1093/cid/ciaa1750] [Medline: 33220049]
48. Ayoubkhani D, Bermingham C, Pouwels KB, Glickman M, Nafilyan V, Zaccardi F, et al. Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study. BMJ 2022 May 18;377:e069676 [FREE Full text] [doi: 10.1136/bmj-2021-069676] [Medline: 35584816]

49. Strain WD, Sherwood O, Banerjee A, Van der Togt V, Hishmeh L, Rossman J. The impact of COVID vaccination on symptoms of long COVID: an international survey of people with lived experience of long COVID. Vaccines (Basel) 2022 Apr 21;10(5):652 [FREE Full text] [doi: 10.3390/vaccines10050652] [Medline: 35632408]

Abbreviations

aOR: adjusted odds ratio
ICU: intensive care unit
IMV: invasive mechanical ventilation
ONS: Office for National Statistics
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 31.08.22; peer-reviewed by Y Wang, P Wang; comments to author 22.12.22; revised version received 09.01.23; accepted 12.01.23; published 07.03.23.

Please cite as:
Wong MCS, Huang J, Wong YY, Wong GLH, Yip TCF, Chan RNY, Chau SWH, Ng SC, Wing YK, Chan FKL
Epidemiology, Symptomatology, and Risk Factors for Long COVID Symptoms: Population-Based, Multicenter Study
JMIR Public Health Surveill 2023;9:e42315
URL: https://publichealth.jmir.org/2023/1/e42315
doi:10.2196/42315
PMID:36645453

Original Paper

Population Size Estimation of Men Who Have Sex With Men in Rwanda: Three-Source Capture-Recapture Method

Elysee Tuyishime[1,2], BSc, MSc; Catherine Kayitesi[3], BSc, MPH; Gentille Musengimana[2], BSc, MSc; Samuel Malamba[2], PhD; Hailegiorgis Moges[2], BSc, MSc; Ida Kankindi[2], MPH, MD; Horacio Ruisenor Escudero[4], MD, PhD; Ignace Habimana Kabano[1], PhD; Tom Oluoch[2], PhD; Eric Remera[3], PhD; Angela Chukwu[1], Prof Dr

[1]African Center of Excellence in Data Science, University of Rwanda, Kigali, Rwanda

[2]Division of Global HIV and Tuberculosis, Center for Global Health, US Centers for Disease Control and Prevention, Kigali, Rwanda

[3]HIV/AIDS, Sexually Transmitted Infections and Viral Hepatitis Division, Rwanda Biomedical Center, Kigali, Rwanda

[4]Key Population Surveillance Team, Epidemiology and Surveillance Branch, Division of Global HIV and Tuberculosis, Center of Global Health, US Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Elysee Tuyishime, BSc, MSc
Division of Global HIV and Tuberculosis
Center for Global Health
US Centers for Disease Control and Prevention
337H+F58, Kigali
Kigali, 28
Rwanda
Phone: 250 782723453 ext 2580
Email: obx5@cdc.gov

Abstract

Background: Globally, men who have sex with men (MSM) continue to bear a disproportionately high burden of HIV infection. Rwanda experiences a mixed HIV epidemic, which is generalized in the adult population, with aspects of a concentrated epidemic among certain key populations at higher risk of HIV infection, including MSM. Limited data exist to estimate the population size of MSM at a national scale; hence, an important piece is missing in determining the denominators to use in estimates for policy makers, program managers, and planners to effectively monitor HIV epidemic control.

Objective: The aims of this study were to provide the first national population size estimate (PSE) and geographic distribution of MSM in Rwanda.

Methods: Between October and December 2021, a three-source capture-recapture method was used to estimate the MSM population size in Rwanda. Unique objects were distributed to MSM through their networks (first capture), who were then tagged according to MSM-friendly service provision (second capture), and a respondent-driven sampling survey was used as the third capture. Capture histories were aggregated in a 2k–1 contingency table, where k indicates the number of capture occasions and “1” and “0” indicate captured and not captured, respectively. Statistical analysis was performed in R (version 4.0.5) and the Bayesian nonparametric latent-class capture-recapture package was used to produce the final PSE with 95% credibility sets (CS).

Results: We sampled 2465, 1314, and 2211 MSM in capture one, two, and three, respectively. There were 721 recaptures between captures one and two, 415 recaptures between captures two and three, and 422 recaptures between captures one and three. There were 210 MSM captured in all three captures. The total estimated population size of MSM above 18 years old in Rwanda was 18,100 (95% CS 11,300-29,700), corresponding to 0.70% (95% CI 0.4%-1.1%) of total adult males. Most MSM reside in the city of Kigali (7842, 95% CS 4587-13,153), followed by the Western province (2469, 95% CS 1994-3518), Northern province (2375, 95% CS 842-4239), Eastern province (2287, 95% CS 1927-3014), and Southern province (2109, 95% CS 1681-3418).

Conclusions: Our study provides, for the first time, a PSE of MSM aged 18 years or older in Rwanda. MSM are concentrated in the city of Kigali and are almost evenly distributed across the other 4 provinces. The national proportion estimate bounds of MSM out of the total adult males includes the World Health Organization’s minimum recommended proportion (at least 1.0%) based on 2012 census population projections for 2021. These results will inform denominators to be used for estimating service coverage and fill existing information gaps to enable policy makers and planners to monitor the HIV epidemic among MSM nationally. There is an opportunity for conducting small-area MSM PSEs for subnational-level HIV treatment and prevention interventions.

(JMIR Public Health Surveill 2023;9:e43114) doi:10.2196/43114

KEYWORDS

population size; men who have sex with men; nationwide; capture-recapture; RDS; three-source; Rwanda; HIV

Introduction

Globally, men who have sex with men (MSM) continue to bear a disproportionately high burden of HIV infection [1]. In sub-Saharan Africa, same-sex behaviors have become an area of interest for HIV research. The results from recent studies indicate the widespread existence of MSM groups across sub-Saharan Africa and high rates of HIV infection among this population [2]. As countries approach HIV epidemic control in sub-Saharan Africa, attention has shifted to key affected populations, including female sex workers (FSW), people who inject drugs (PWID), transgender women, and MSM who are at a higher risk of HIV acquisition [3].

These key populations account for less than 5% of the global population, but they and their sexual partners comprised 70% of new HIV infections in 2021 [4]. MSM living in sub-Saharan African countries, where homosexual activity has often been severely criminalized, are hard to reach with HIV prevention programs. Currently, the risk of getting an HIV infection is nearly five times higher among MSM in countries where there is criminalization for homosexuality compared to that for MSM in countries where no criminalization for homosexuality occurs [5].

Biologically, unprotected receptive anal sex poses a much higher risk than unprotected receptive vaginal sex, whether the receptive partner is male or female; the risk of HIV transmission during anal intercourse may be approximately 18 times greater than that during vaginal intercourse [1]. By this fact, MSM are at an extremely high risk of becoming infected with HIV. The global estimated HIV prevalence among MSM is 21%, with 6% in sub-Saharan countries, ranging from 3.8% up to 31% [6,7]. In addition, people with marginalized sexual or gender identities or behaviors sometimes lack the ability to protect themselves from HIV infection due to structural factors, including self-stigmatization, discrimination, and lack of access to information and services [8]. The prevalence ratios are particularly elevated in West and Central Africa as well as in low-prevalence countries [9].

Rwanda is in the East African region, bordering four countries: Tanzania, Uganda, the Democratic Republic of the Congo, and Burundi. The country is divided into five administrative regions, including four provinces, the city of Kigali, and 30 districts as another subnational unit level. Rwanda experiences a mixed HIV epidemic, which is generalized in the adult population, with an adult (15-49 years) HIV prevalence stabilized at around 2.6%, and aspects of a concentrated epidemic among specific key populations at higher risk of HIV infection, with a rate of 45% among FSWs [10] and 6.5% among MSM [11]. In Rwanda, MSM are considered one of the key populations for HIV prevention and care in the HIV and AIDS national strategic plan [12].

Although the national strategic plan suggests a special emphasis on key populations and that the epidemic is likely concentrated among these populations, limited data exist to estimate the actual population size at the national scale [12]. To date, no nationwide study has been performed to estimate the number of MSM in Rwanda. However, the Project San Francisco study performed in 2018 using a service multiplier and unique object identifier approach estimated the size of MSM in the city of Kigali and surrounding areas to be 8411, ranging between 6760 and 11,151 [11]. Lack of reliable estimates of the MSM national population size misses important information to determine the denominators to estimate need, coverage, and the spread of new infections that policy makers and planners use to monitor HIV epidemic control.

In response to this information gap, we performed a national population size estimation of MSM in Rwanda from October to December 2021, aimed at providing national-level and provincial-level estimates.

Methods

Study Population

The study population included adult males aged 18 years and older who self-reported as gay/bisexual or having had anal sex with a male in the last 12 months and lived primarily in Rwanda during the past 12 months. By living primarily in Rwanda, we refer to having spent most of their time in Rwanda over the last 12 months, regardless of any possibility of traveling out of the country. The study excluded any MSM who were not willing to voluntarily participate.

Study Design and Setting

Estimating the size of hard-to-reach populations has long been a challenging task due to the fact that sampling frames often do not exist for these populations, some populations are dynamic both in time and space, and members may be afraid to identify themselves in the community. There are multiple methods suggested to estimate the size of populations without a sampling frame, including venue-based sampling [13], time-location sampling [14], respondent-driven sampling (RDS) [15], multiplier methods [16], the network scale-up method [17], successive sampling-population size estimation [18], and capture-recapture (CRC) [19]. No gold-standard method currently exists, and many methods have been used, each with different strengths and weaknesses.

The CRC method has been shown to be useful in estimating the size of hard-to-reach populations [20-22]. Based on its statistical foundation, CRC, as an empirical population size estimate (PSE) method, produces PSEs with higher precision. Research suggests that, compared with two-source CRC, conducting CRC with multiple (three or more) sampling/capture rounds strengthens the design, produces more robust estimates, and relaxes the CRC requirement (assumption) of sample independence [23]. The CRC method is increasingly being used by several researchers in epidemiology and public health [24-26], and this method was previously used to estimate the national size estimate of FSWs in Rwanda, demonstrating that acquired implementation experiences are crucial for such estimates [27].

We used the three-source capture-recapture (3S-CRC) method to estimate the size of MSM in Rwanda. CRC consists of marking a random portion of the population of interest at the first encounter. Subsequently, another portion of the population is drawn to observe how many were marked initially. The greater the number of those tagged, the smaller the PSE. Tagging can be replicated multiple times as desired. The 3S-CRC approach has been shown to be a robust method in estimating a population without a sampling frame, such as FSWs, MSM, and PWID [28], and is described elsewhere in more detail [29]. The MSM 3S-CRC began with a capture stage. During this stage, members of the MSM population were “encountered” and then “marked” by providing a memorable and easily identifiable gift with a specific design that is hard to find on the local market. After a 1-week interval, a second capture (recapture) was initiated by offering MSM-friendly services to MSM countrywide. During the third capture (recapture), an RDS approach was used [30,31] with specific questions incorporated to identify those MSM encountered during the previous capture occasions.

During capture one, MSM key informants were selected by MSM community-based organizations (CBOs) to facilitate the distribution of unique objects within MSM associations and groups and those who are neither members of any association nor MSM group. A list of MSM associations and groups was developed with their corresponding numbers of members in all provinces of the country: 14 from Eastern province, 12 from the city of Kigali, 10 from Southern province, and 8 from each of the Northern and Western provinces. Each province was assigned a different unique color of the unique objects to be distributed. The number of objects to be distributed within each association or group was determined by the probability proportional to the association or group size; within an association or a group, systematic sampling was found to be appropriate to determine who receives the unique object according to the existing list of MSM association or group members. The approach comprises determining a sampling interval as well as a random start on the list of MSM association or group members.

Each selected MSM member within an association or a group was given three unique objects: one for himself and another two to hand out to other MSM he knows but who do not belong to any association or group. The distribution process for unique objects within an MSM association or group was facilitated by a well-trained team of distributors with MSM key informants’ support. Branded keychains valued at no more than US $3 were used as unique objects. MSM were provided with the study objectives and advised to keep the received unique objects in a safe place as they may be asked to show them later for verification. This activity took 1 week to be completed.

There were possibilities that some objects could not be successfully distributed and not returned as well as multiple objects given to the same person. To minimize any related bias on the PSEs, a debrief was given to the MSM key informant as well as the MSM association/group members focusing on study objectives; eligibility criteria to receive the object; object distribution process; and key considerations, including prompting the object receiver to check if he had not been approached by someone else in the same study context to avoid duplication. In addition, the monetary value of the unique object was set to US $3 to minimize the intention for the object distributor to keep the objects for themselves or the object receiver being willing to receive more than one object. The object distribution process was monitored daily by the object distributor and MSM key informant to report the number of objects distributed successfully as well as those not successfully distributed, which were returned physically.

The following week, capture two was initiated, where MSM were tagged using selected specific MSM-friendly services. The selection process for services was facilitated by MSM key informants and their CBOs. Given that MSM are given health services through usual health facilities but equipped with MSM-friendly settings and packages, the latter were used to offer specific MSM-friendly services, including distribution of condoms and lubricants. For this purpose, 23 health facilities were selected across the country where an MSM key informant and a health worker who usually serve MSM at the same facility were assigned to provide the selected service to MSM and to record corresponding information in the study purposes. The key informant role was to serve as the receptionist and to facilitate screening to create a welcoming and friendly environment for MSM, while the health worker served to provide services and record necessary information. Prior to the second capture, community mobilizers served on the mobilizing of upcoming service provision to MSM to increase awareness and to break the fear and stigma of participation. Those who accepted the offered services were counted as captured and were asked whether they had received the distributed unique object during the previous week. Responses (yes/no) were recorded depending on whether the MSM had the gift with him or was able to correctly identify the unique object received.

Capture two was followed by the Integrated Behavioral and Biological Surveillance Survey (IBBSS) using RDS [32], which served as the opportunity for the third capture. RDS is a variant of chain-referral sampling that employs Markov-chain theory and the theory of biased networks to reduce biases generally associated with chain-referral methods [31]. It was proven that even though sampling begins with a purposively chosen group of initial subjects, as is the case for most chain-referral samples, the composition of the ultimate sample is wholly independent of those initial subjects [33]. During capture three, all MSM recruited by their peers using the RDS approach were counted as captured. There were eight study sites distributed according to the administrative provinces: one in each of the Northern, Southern, and Eastern provinces; two in the Western province; and three in the city of Kigali. MSM who met the study eligibility criteria and were well-connected and respected among their MSM peers were selected as seeds and marked the beginning of each referral chain. During this study, the investigator contacted the seed via the implementing partners, including nongovernmental organizations and CBOs working with the MSM population. Three seeds were selected at each study site for a total of 24 seeds.

During the RDS survey, participants were asked whether they had received unique objects during capture one and/or had received provided services during capture two during the second week of study implementation. Several prompts were used to confirm prior participation, including having received the object physically or being able to correctly identify the received object on a laminated card with images of multiple objects. MSM were not allowed to participate more than once in the same round; an interoperable fingerprint system was used to verify that each participant was recorded only once during capture three and to detect any duplication within the same site and across all study sites.

Fingerprint machines were installed at each study site and connected to the internet for easy and real-time data synchronization. Once a fingerprint was recorded, it was automatically converted into alpha-numerical codes and transferred to the central server to synchronize all study site–level data. With this installed system, we could identify any MSM trying to reregister his fingerprint at the same site or at a different site. The fingerprint machines converted the recorded fingerprints into alpha-numerical codes with no reverse backway possibility, and then the alpha-numerical codes were stored as participant identifiers to ensure participants' confidentiality.

Sample Size and Sampling

For the first two initial captures, the sample size was based on Rwanda Population-Based HIV Impact Assessment 2019 [34], where the proportion of all males aged 18+ years with at least one male sexual partner in the last 12 months by province (0.30% for the city of Kigali, and 0.16%, 0.08%, 0.24%, and 0.24% for the Eastern, Northern, Southern, and Western provinces, respectively) was obtained to estimate the pooled and provincial-level stratified sample size. Assuming a design effect of 1.5, a precision of 0.5%, and adjusting for a 15% loss of coupons estimated from a previous CRC-related study [27], we calculated the required minimum sample size. We estimated a total of 2705 objects needed to be distributed across provinces as follows: 803 in the city of Kigali, 586 in Western province, 658 in Southern province, 219 in Northern province, and 439 in Eastern province during each of the first two captures.

For the RDS survey, the sample size calculation was based on the results from the previous IBBSS conducted in Rwanda in 2020. The formula for the sample size calculation for an RDS study was used according to the method of Salganik [35], assuming an MSM HIV prevalence of 11.3% in the city of Kigali, 6.4% in Western province, 1.4% in Southern province, 3.1% in Northern province, and 1.2% in Eastern province; a design effect of 1.5; precision *w* 0.025; and a nonresponse rate of 10%. We estimated the required minimum total sample size of 2210 across the country distributed by province as follows: 1027 in the city of Kigali, 613 in Western province, 141 in Southern province, 308 in Northern province, and 121 in Eastern province.

All of the estimated sample sizes for each capture and the statistical power were validated using MS-CRC Power Analysis of the shinyrecap application [36].

Data Management

Data were captured using 3 electronic forms (capture one, capture two, and capture three forms) using tablets with data-entry forms designed using Open Data Kit (ODK) [37]. The tablets were password-protected to ensure that only the study team had access to the data. Other data were captured using Microsoft Excel sheets, which were also password-protected. Data cleaning, entailing deduplication, review of data consistency, and completion of any missing data, was performed daily.

Unique identifiers were used to ensure that study participants were only enrolled once. Once data were deduplicated, all identifiers were removed prior to creating the final data sets for analysis.

Statistical Analysis

In preparation for analysis, participant-level data from ODK were exported into R v.4.0.5 for Windows, and cleaning steps were performed based on preset exclusion criteria and logical flow. Data were analyzed by province. Aggregated data sets detailing counts of each capture/recapture combination were produced for each subset. The Bayesian nonparametric latent-class model, which is flexible and able to accommodate various forms of heterogeneity in capture probabilities, was used to produce the final PSE with 95% credible sets (CS) from aggregate data sets.

All analyses were performed using the latent-class model for capture-recapture (LCMCR) package in R v.4.0.5 for Windows [38,39]. LCMCR is a novel Bayesian nonparametric method for estimating the size of hard-to-reach populations from CRC data. This method is based on a Dirichlet process mixture, capable of accommodating complex patterns of heterogeneity of captures, and can transparently modulate its complexity without a separate model selection step [40,41]. The Bayesian nonparametric latent-class approach posits that the population is divided into several groups with members in each group having the same homogeneous capture probability. The number of homogeneous strata in a population is uncertain and covariates that identify those classes may not be available. Thus, the strata are said to be latent and strata identities are treated as missing data. Estimation is naturally accomplished using mixtures of distributions [42].

Uninformative priors (ie, those with minimal influence on the inference and dominated by the likelihood function) were specified for the Dirichlet process parameters α and γ (0.25, 0.25). We used K=5 latent classes; 10,000 samples from the posterior distribution were drawn with a burn-in of 10,000 iterations and a thinning interval of 1000 iterations to specify the Markov Chain Monte Carlo (MCMC) sampling. Convergence of the MCMC sampling was assessed using trace plots and a histogram of the posterior probability distribution for population size.

Median PSEs with 95% CS for 3S-CRC were produced overall and by province. To facilitate the interpretation of results and application of estimates for programs, the highest density intervals (HD Interval package in R) are presented. Analysis outputs include the median PSEs with 95% CS.

Ethical Considerations

The survey received ethical approval from the local Institutional Review Board, Rwanda National Ethics Committee. It was also reviewed in accordance with the US Centers for Disease Control and Prevention (CDC) human research protection procedures and was determined to be research with CDC nonengaged. For the PSE study component, we received a waiver of the informed consent requirement for all MSM participants because no personally identifiable information was collected; however, verbal consent was sought since the data collected were restricted to only recording whether the MSM accepted the gift during subsequent capture.

Since same-sex relationships remain stigmatized in Rwanda, we anticipated a small risk of physical and/or verbal violence in cases where a study participant was identified in the community. Therefore, the study investigator ensured that the study was conducted anonymously to protect the identity of participants and ensure confidentiality of the data collected; no names or any other personally identifiable information were recorded anywhere. Completed questionnaires (identifying individuals by only identification numbers) were kept with the study coordinator during the fieldwork. All forms containing any study information were kept in a locked cabinet accessible by only authorized study personnel. Additionally, the electronic data set identified by the unique code was password-protected and accessed by authorized personnel using a computer backed up to a server located at Rwanda Biomedical Center. Furthermore, for all study members participating in human subjects research and data collection, it was a requirement for the study team to have participated in trainings on human subjects research, confidentiality, and interviewing techniques before commencing study activities. A confidentiality agreement was signed by all study investigators, coinvestigators, and data collectors.

Results

Capture One

Keychains with unique designs were distributed to the MSM through their corresponding associations, groups, and key informants (Figure 1). Table 1 summarizes the results from capture one; a total of 2465 out of the 2723 objects (90.53%) were successfully distributed.

Figure 1. Example of a unique object (keychain) distributed in capture one.

Table 1. Provincial-level object distribution breakdown.

Location	Objects assigned, n	Objects successfully distributed, n	Objects not distributed successfully and returned, n
Eastern province	636	558	78
City of Kigali	894	885	9
Northern province	166	150	16
Southern province	585	515	70
Western province	442	357	85
Total	2723	2465	258

Capture Two

MSM-friendly services were provided at 23 health facilities that typically have a key population service package countrywide. Lubricants and condoms were distributed during the capture-two period, during which a total of 1340 of the anticipated 2705 MSM (49.54%) came for health services at health facilities. Out of the 1340 MSM who came for the services, 1314 (98.06%) met the inclusion criteria and were offered the services, and among them, 721 (54.87%) were identified as having received the distributed unique object during the previous week. Table 2 provides the province-level breakdown for capture two.

Table 2. Provincial-level service provision among men who have sex with men (MSM) in capture two.

Location	Health facilities, n	Anticipated MSM to be offered services, n	MSM received at health facility, n	MSM received offered services, n	MSM received distributed unique object during the previous week, n
Eastern province	5	439	343	337	211
City of Kigali	5	803	510	497	185
Northern province	1	219	28	25	4
Southern province	7	658	291	291	195
Western province	5	586	168	164	126
Total	23	2705	1340	1314	721

Capture Three

During capture three, RDS was used for participant recruitment. Every MSM recruited during RDS was screened for eligibility and considered as captured during capture three once he consented to participate. A total of 2211 MSM were captured during this capture occasion. Among those captured during capture three, 422 (19.09%) were identified as having received distributed unique objects during capture one, whereas 415 (18.77%) were identified as having received MSM-friendly services during capture two. Table 3 provides the provincial-level breakdown for capture three.

Table 3. Provincial-level data among men who have sex with men (MSM) in capture three.

Location	Anticipated MSM to be captured, n	MSM captured, n	MSM received distributed unique object, n	MSM received provided MSM-friendly services, n	MSM received both unique object and MSM-friendly services, n
Eastern province	121	126	50	64	36
City of Kigali	1027	1,021	128	124	42
Northern province	308	303	20	18	4
Southern province	141	152	42	54	28
Western province	613	609	182	155	100
Total	2210	2211	422	415	210

Population Size Estimation

Overall, we sampled 2465, 1314, and 2211 MSM in capture one, two, and three, respectively. There were 721 recaptures between captures one and two, 415 recaptures between captures two and three, and 422 recaptures between captures one and three. There were 210 MSM captured in all three captures. The Venn diagram in Figure 2 provides the summary results for all three capture occasions and overlaps.

Before conducting the CRC analysis, we explored dependency between captures by testing for homophily in the RDS recruitment chain based on the capture history variable, finding a homophily value of 1.016596, which indicates nondependency.

The trace plot in Figure 3 presents the simulation results over 10,000 samples, demonstrating the population size distribution for a converging simulation result based on the random noise shape observed between the values 10,000 and 20,000 on the Y-axis. Figure 4 presents the posterior distribution (ie, population size distribution) provided by the model.

The final MSM PSEs for the overall and provincial levels are presented in Table 4. The median from the posterior probability distribution is used as the point estimate and 95% CS are used to describe uncertainty.

We estimated the overall population of MSM in Rwanda to be 18,100 (95% CS 11,300-29,700), with the majority living in the city of Kigali (7842, 95% CS 4587-13,153). The MSM PSEs were similar for the remaining 4 provinces (Northern, Southern, Eastern, and Western provinces).

Figure 2. Venn diagram representing individual capture results and overlaps between capture occasions. Numbers represent the numbers of men who have sex with men captured at each round.

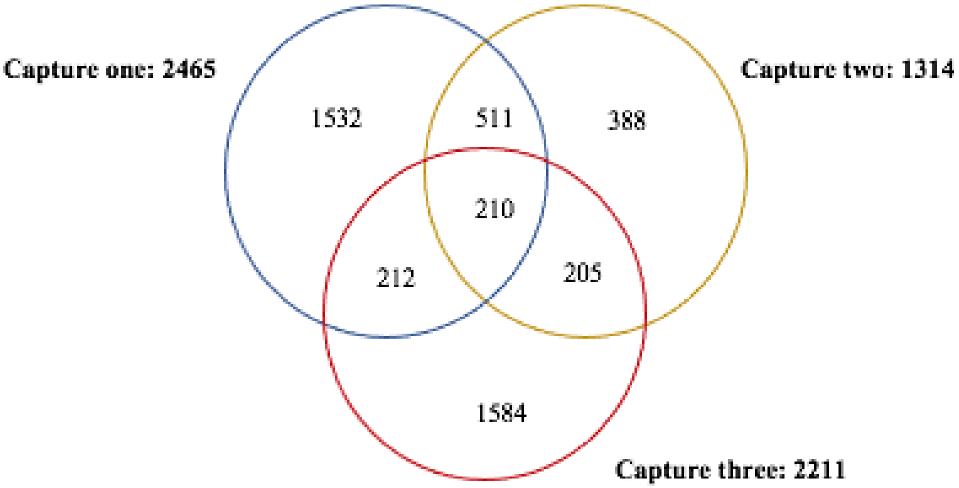


Figure 3. Trace plot for the population size estimate.

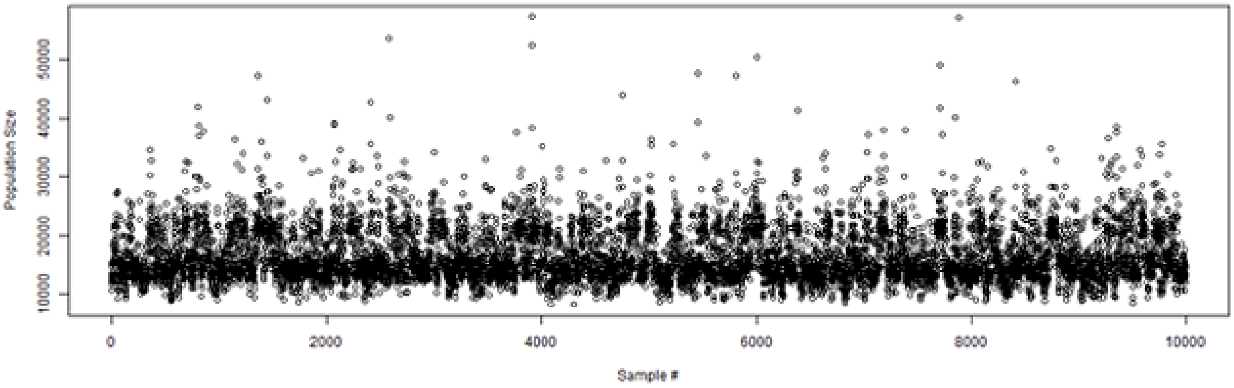


Figure 4. Histogram of the posterior probability distribution for population size.

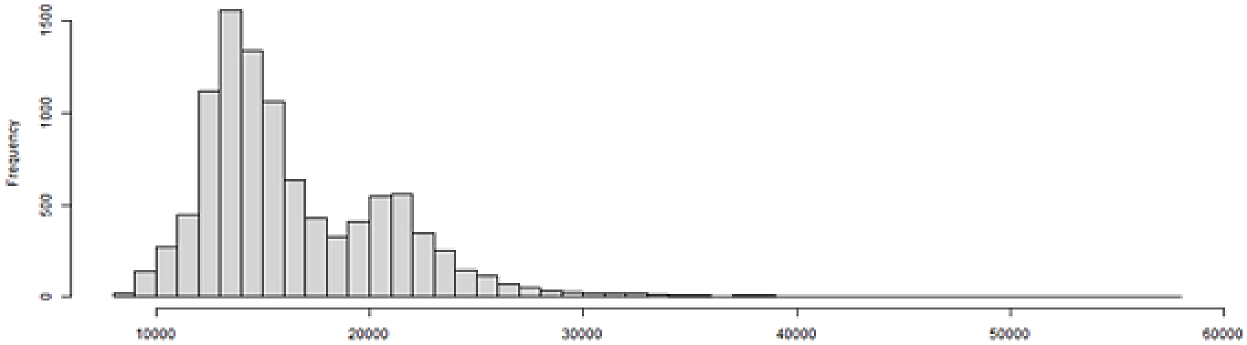


Table 4. Population size estimate (PSE) of men who have sex with men, Rwanda 2022.

Location	Proportion estimate[a], % (95% CI)	Median PSE (95% CS[b])
Eastern province	0.3 (0.3-0.5)	2287 (1927-3014)
City of Kigali	2.7 (1.6-4.6)	7842 (4587-13,153)
Northern province	0.5 (0.2-1.0)	2375 (842-4239)
Southern province	0.4 (0.3-0.6)	2109 (1681-3418)
Western province	0.3 (0.3-0.5)	2469 (1994-3518)
Overall	0.7 (0.4-1.1)	18,100 (11,300-29,700)

[a] Proportion estimates of men were based on the 2012 Census data by National Institute of Statistics of Rwanda 2021 population size projections.

[b] CS: credibility set.

Discussion

This MSM PSE presents the first use of the 3S-CRC approach to estimate MSM population size on a nationwide scale in Rwanda. The 3S-CRC method provided an estimate for MSM in the city of Kigali of 7842 (95% CS 4587-13,153), which was similar to the previous estimate obtained by “Project San Francisco” of 8411 (95% CS 6760- 11,151) [11]. The slight differences in the 2018 and 2021 Kigali MSM PSEs can be explained by considering the geographical coverage and difference in methodologies used. The distribution of MSM in each of the remaining 4 provinces was fairly uniform and lower than the estimates in Kigali. Differences in estimates of the MSM population size distribution across the country may also reflect long-term movement patterns among MSM, from rural to urban as well as from smaller to larger urbanized contexts [43].

To some degree, MSM size estimates are influenced by the proportion of MSM who may have decided to not participate in the study due to potential privacy concerns, a potentially significant element given the burden of stigma and heteronormative behavioral expectations (eg, marriage and parenting) for MSM in Rwanda. Our overall MSM PSE represents 0.7% (95% CI 0.4%-1.1%) of the total adult male population in Rwanda based on 2012 Census data collected by National Institute of Statistics of Rwanda 2021 population size projections. The 2020 World Health Organization (WHO) and Joint United Nations Program on HIV/AIDS (UNAIDS) technical brief recommends the revision of the MSM PSE for those countries, using an MSM PSE less than 1% of total adult males based on the region [28]. The UNAIDS monitoring system through 2019 estimated a global median proportion of adult men who had sex with another man in the previous year of 1.9% across 38 low- or middle-income countries, and this proportion was estimated at 1.45% in Eastern and Southern Africa, where Rwanda is located [44,45]. Accordingly, the estimate from the current study aligns with the WHO recommendation regarding MSM PSEs [7].

Different approaches are being used for size estimation of hidden populations, each with various strengths and limitations, and there is currently no consensus on a gold-standard method [46]. The traditional CRC method employs two captures; however, extra captures can be added to increase the number of data points from which estimates are created, resulting in more stable and robust PSEs [47]. The 3S-CRC method has been commonly used in epidemiology to estimate the size of key populations targeted by health interventions for certain health conditions owing to its mathematically grounded and defensible results [23,48,49]. Several studies have used the 3S-CRC method to estimate the size of certain population groups without a sampling frame, including FSW, MSM, and PWID [28,50-52]. The final estimate for this study was based on the 3S-CRC data set. In summary, four major assumptions must be met for CRC to give reliable population estimates: individual captures are independent, the population is closed, the capture history is correct for all members of the target population, and the chance of getting caught is homogeneous [53].

To minimize dependencies between captures, we used different distribution settings for each capture occasion. During the first capture, members of the MSM population were tagged by the keychain provided through their corresponding associations, groups, and key informants. In the second capture, MSM were tagged by being offered MSM-friendly services through health facilities that usually serve MSM nationwide. In the third capture, we used the RDS method, in which all recruited MSM were considered as captured. For all three captures, a 1-week time interval was used between two consecutive capture occasions to minimize recall bias and to fulfill the population closeness assumption. At each capture of the first two capture rounds, unique object distribution and MSM-friendly services provision procedures included a random aspect to ensure that the chance of getting caught was homogeneous. However, our estimates might be limited with missing a random sampling aspect during the third capture round where RDS was used.

There are several plausible constraints to the design of our estimation activity. A possible limitation is the underlying 3S-CRC assumptions that might have influenced the validity of our findings, leading to reduced accuracy of population sizes and wide confidence ranges. In particular, we employed unique objects as a tagging strategy to protect the anonymity of sampled populations. However, not all individuals were carrying the received object at the subsequent capture occasion, complicating the identification of recaptures. Furthermore, we had to assume that the person presenting the object is the same person who received the object (an essential limitation present in anonymous sampling–based CRC). We tried to mitigate these limitations by offering the opportunity to identify the correct object from a laminated card with several pictures including the correct unique object for those presenting without unique objects. There was a possibility of guessing or having seen the object and therefore biasing the PSE downward. To overcome possible participation duplicates at enrollment during the third capture, a biometric system using fingerprint identification was installed and employed across all study sites. We also acknowledge any possible selection bias that might have been influenced by the established study inclusion criteria.

The key strength of our study is that it is powered to provide national- and provincial-level PSEs for MSM in Rwanda for the very first time. Sampling considered administrative provinces as strata and targeted 28 (out of 30) districts in Rwanda, which are nationally representative and reflective of the demography. The selected districts included key urban areas with a high likelihood of expanding the catchment to include participation by MSM based in rural areas. Furthermore, during the third capture occasion, RDS was used giving more confidence in reaching MSM with lower social visibility.

The final estimate of the MSM population size in Rwanda is based on a Bayesian approach to account for the complex patterns of heterogeneity between captures and the aggregation of homogeneous strata into latent classes. While other statistical techniques make reasonably strong assumptions about the structure of the joint distribution of capture patterns, the latent-class Bayesian method is a model-averaging strategy that seeks to estimate the joint distribution as directly as feasible from the data [40].

In conclusion, this study provides, for the first time, an estimated population size of MSM aged 18 years and above in Rwanda. The results will allow national programs and implementation partners to invest in HIV services at a level that is commensurate with need, coverage, and new infections. These data enable policy makers and planners to monitor HIV epidemic control nationwide, specifically among the MSM population, and to plan for other health services such as the prevention and treatment of sexually transmitted infections, among others. While these estimates are usable at the national and provincial levels, further work is needed on small-area estimation to align the PSE results with the intended HIV treatment and prevention interventions at subnational levels among MSM. Furthermore, we acknowledge that there are still limitations of estimating some hard-to-reach MSM groups, which is another potential area for further research.

Acknowledgments

The authors would like to thank all those that made valuable contributions to the success of this project. This population size estimation of men who have sex with men has been supported by the President’s Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention (CDC) under the terms of GH002203. The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the funding agencies.

Data Availability

Data are owned by Rwanda Biomedical Center (RBC); all data access enquiries should be directed to RBC through the following email address: info@rbc.gov.rw.

Authors' Contributions

All authors made significant contributions to the study’s design, conduct, or data analysis and interpretation, as well as writing or editing portions of the work and approving the final version for publication.

Conflicts of Interest

None declared.

References

1. Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. Lancet 2012 Jul 28;380(9839):367-377 [FREE Full text] [doi: 10.1016/S0140-6736(12)60821-6] [Medline: 22819660]
2. Sandfort TGM, Dominguez K, Kayange N, Ogendo A, Panchia R, Chen YQ, et al. HIV testing and the HIV care continuum among sub-Saharan African men who have sex with men and transgender women screened for participation in HPTN 075. PLoS One 2019 May 31;14(5):e0217501 [FREE Full text] [doi: 10.1371/journal.pone.0217501] [Medline: 31150447]
3. Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000-2006: a systematic review. PLoS Med 2007 Dec;4(12):e339 [FREE Full text] [doi: 10.1371/journal.pmed.0040339] [Medline: 18052602]
4. Frescura L, Godfrey-Faussett P, Feizzadeh AA, El-Sadr W, Syarif O, Ghys PD, on behalf of the 2025 testing treatment target Working Group. Achieving the 95 95 95 targets for all: a pathway to ending AIDS. PLoS One 2022;17(8):e0272405 [FREE Full text] [doi: 10.1371/journal.pone.0272405] [Medline: 35925943]
5. Beyrer C, Baral SD, Collins C, Richardson ET, Sullivan PS, Sanchez J, et al. The global response to HIV in men who have sex with men. Lancet 2016 Jul 09;388(10040):198-206. [doi: 10.1016/S0140-6736(16)30781-4] [Medline: 27411880]

6. UNAIDS Global AIDS Update 2022. United Nations Program on HIV/AIDS. URL: https://www.unaids.org/en/resources/documents/2022/in-danger-global-aids-update [accessed 2023-03-13]
7. World Health Organization, UNAIDS. Recommended population size estimates of men who have sex with men. Joint United Nations Program on HIV/AIDS. 2020 Nov. URL: https://www.unaids.org/sites/default/files/media_asset/2020-recommended-population-size-estimates-of-men-who-have-sex-with-men_en.pdf [accessed 2023-03-13]
8. Ayala G, Beck J, Lauer K, Reynolds R, Sundararaj M, Network Contact: The Global Forum on MSM and HIV (MSMGF). Social discrimination against men who have sex with men (MSM) implications for HIV policy and programs. The Communication Initiative Network. 2010 Sep. URL: https://www.comminit.com/global/content/social-discrimination-against-men-who-have-sex-men-msm-implications-hiv-policy-and-progr [accessed 2023-03-13]
9. Case K, Johnson LF, Mahy M, Marsh K, Supervie V, Eaton JW. Summarizing the results and methods of the 2019 Joint United Nations Programme on HIV/AIDS HIV estimates. AIDS 2019 Dec 15;33(Suppl 3):S197-S201 [FREE Full text] [doi: 10.1097/QAD.0000000000002440] [Medline: 31725433]
10. Nsanzimana S, Mills EJ, Harari O, Mugwaneza P, Karita E, Uwizihiwe JP, et al. Prevalence and incidence of HIV among female sex workers and their clients: modelling the potential effects of intervention in Rwanda. BMJ Glob Health 2020 Aug;5(8):e002300 [FREE Full text] [doi: 10.1136/bmjgh-2020-002300] [Medline: 32764126]
11. Twahirwa Rwema JO, Lyons CE, Herbst S, Liestman B, Nyombayire J, Ketende S, et al. HIV infection and engagement in HIV care cascade among men who have sex with men and transgender women in Kigali, Rwanda: a cross-sectional study. J Int AIDS Soc 2020 Oct;23(Suppl 6):e25604 [FREE Full text] [doi: 10.1002/jia2.25604] [Medline: 33000912]
12. Fourth Health Sector Strategic Plan July 2018-June 2024. Republic of Rwanda Ministry of Health. 2018. URL: https://www.childrenandaids.org/sites/default/files/2018-05/Rwanda_Nat%20Health%20Sector%20Plan_2018-2024.pdf [accessed 2023-03-14]
13. Muhib FB, Lin LS, Stueve A, Miller RL, Ford WL, Johnson WD, Community Intervention Trial for Youth Study Team. A venue-based method for sampling hard-to-reach populations. Public Health Rep 2001;116(Suppl 1):216-222 [FREE Full text] [doi: 10.1093/phr/116.S1.216] [Medline: 11889287]
14. Karon JM, Wejnert C. Statistical methods for the analysis of time-location sampling data. J Urban Health 2012 Jun 16;89(3):565-586 [FREE Full text] [doi: 10.1007/s11524-012-9676-8] [Medline: 22421885]
15. McCreesh N, Frost SDW, Seeley J, Katongole J, Tarsh MN, Ndunguse R, et al. Evaluation of respondent-driven sampling. Epidemiology 2012 Jan;23(1):138-147 [FREE Full text] [doi: 10.1097/EDE.0b013e31823ac17c] [Medline: 22157309]
16. Medhi G, Mahanta J, Akoijam BS, Adhikary R. Size estimation of injecting drug users (IDU) using multiplier method in five districts of India. Subst Abuse Treat Prev Policy 2012 Feb 21;7:9 [FREE Full text] [doi: 10.1186/1747-597X-7-9] [Medline: 22348548]
17. Bernard H, Hallett T, Iovita A, Johnsen EC, Lyerla R, McCarty C, et al. Counting hard-to-count populations: the network scale-up method for public health. Sex Transm Infect 2010 Dec;86(Suppl_2):ii11-ii15 [FREE Full text] [doi: 10.1136/sti.2010.044446] [Medline: 21106509]
18. Johnston LG, McLaughlin KR, El Rhilani H, Latifi A, Toufik A, Bennani A, et al. Estimating the size of hidden populations using respondent-driven sampling data: case examples from Morocco. Epidemiology 2015 Nov;26(6):846-852 [FREE Full text] [doi: 10.1097/EDE.0000000000000362] [Medline: 26258908]
19. Wesson PD. If you are not counted, you don't count: estimating the size of hidden populations. PhD thesis, UC Berkeley. eScholarship. 2016. URL: https://escholarship.org/uc/item/7gj5h8cp [accessed 2023-03-13]
20. Pollock KH. A capture-recapture design robust to unequal probability of capture. J Wildlife Manag 1982;46(3):752-757. [doi: 10.2307/3808568]
21. Tilling K. Capture-recapture methods--useful or misleading? Int J Epidemiol 2001 Feb;30(1):12-14. [doi: 10.1093/ije/30.1.12] [Medline: 11171841]
22. Böhning D, van der Heijden PGM, Bunge J. Capture-recapture methods for the social and medical sciences. Boca Raton: CRC Press; 2018.
23. van Hest R. Capture recapture methods in surveillance of tuberculosis and other infectious diseases. PhD thesis. Rotterdam: Erasmus University; 2007.
24. Guure C, Dery S, Afagbedzi S, Tun W, Weir SS, Quaye S, et al. National and subnational size estimation of female sex workers in Ghana 2020: comparing 3-source capture-recapture with other approaches. PLoS One 2021;16(9):e0256949 [FREE Full text] [doi: 10.1371/journal.pone.0256949] [Medline: 34550978]
25. McIntyre AF, Mitchell A, Stafford KA, Nwafor SU, Lo J, Sebastian V, et al. Key population size estimation to guide HIV epidemic responses in Nigeria: Bayesian analysis of 3-source capture-recapture data. JMIR Public Health Surveill 2022 Oct 26;8(10):e34555 [FREE Full text] [doi: 10.2196/34555] [Medline: 36287587]
26. Coumans A, Cruyff M, Van der Heijden PGM, Wolf J, Schmeets H. Estimating homelessness in the Netherlands using a capture-recapture approach. Soc Indic Res 2015 Nov 16;130(1):189-212. [doi: 10.1007/s11205-015-1171-7]
27. Musengimana G, Tuyishime E, Remera E, Dong M, Sebuhoro D, Mulindabigwi A, et al. Female sex workers population size estimation in Rwanda using a three-source capture-recapture method. Epidemiol Infect 2021 Mar 18;149:e84 [FREE Full text] [doi: 10.1017/S0950268821000595] [Medline: 33734058]

28. Doshi R, Apodaca K, Ogwal M, Bain R, Amene E, Kiyingi H, et al. Estimating the size of key populations in Kampala, Uganda: 3-source capture-recapture study. JMIR Public Health Surveill 2019 Aug 12;5(3):e12118 [FREE Full text] [doi: 10.2196/12118] [Medline: 31407673]
29. Becker N, Hyde C. Estimating population size from multiple recapture experiments. Stoch Process Their Appl 1990;36(1):77-83. [doi: 10.1016/0304-4149(90)90043-R]
30. Gile KJ, Handcock MS. Respondent-driven sampling: an assessment of current methodology. Sociol Methodol 2010 Aug 03;40(1):285-327 [FREE Full text] [doi: 10.1111/j.1467-9531.2010.01223.x] [Medline: 22969167]
31. Heckathorn D. Respondent-driven sampling: a new approach to the study of hidden populations. Social Prob 2014;44(2):174-199. [doi: 10.2307/3096941]
32. Abdul-Quader A, Heckathorn DD, Sabin K, Saidel T. Implementation and analysis of respondent driven sampling: lessons learned from the field. J Urban Health 2006 Nov;83(6 Suppl):i1-i5 [FREE Full text] [doi: 10.1007/s11524-006-9108-8] [Medline: 17058119]
33. McCreesh N, Frost SDW, Seeley J, Katongole J, Tarsh MN, Ndunguse R, et al. Evaluation of respondent-driven sampling. Epidemiology 2012 Jan;23(1):138-147 [FREE Full text] [doi: 10.1097/EDE.0b013e31823ac17c] [Medline: 22157309]
34. The Rwanda Population-based HIV Impact Assessment. Rwanda Final Report 2018-2019. PHIA Project. URL: https://phia.icap.columbia.edu/rwanda-final-report/ [accessed 2023-03-15]
35. Salganik M. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. J Urban Health 2006 Nov;83(6 Suppl):i98-112 [FREE Full text] [doi: 10.1007/s11524-006-9106-x] [Medline: 16937083]
36. McIntyre AF, Fellows IE, Gutreuter S, Hladik W. Population size estimation from capture-recapture studies using shinyrecap: design and implementation of a web-based graphical user interface. JMIR Public Health Surveill 2022 Apr 26;8(4):e32645 [FREE Full text] [doi: 10.2196/32645] [Medline: 35471234]
37. Hartung C, Lerer A, Anokwa Y, Tseng C, Brunette W, Borriello G. Open data kit: tools to build information services for developing regions. 2010 Presented at: the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development; December 13-15, 2010; London, UK.
38. Manrique-Vallier D. LCMCR: Bayesian Non-Parametric Latent-Class Capture-Recapture. CRAN R project. URL: https://CRAN.R-project.org/package=LCMCR [accessed 2023-03-13]
39. Biggeri A. Latent class models for varying catchability and correlation among sources in capture-recapture estimation of the size of a human population. Statistica Applicata 1999;11(3):1-14 [FREE Full text]
40. Manrique-Vallier D. Bayesian population size estimation using Dirichlet process mixtures. Biometrics 2016 Dec;72(4):1246-1254. [doi: 10.1111/biom.12502] [Medline: 26954906]
41. Teh YW. Dirichlet process. In: Sammut C, Webb GI, editors. Encyclopedia of machine learning. Boston, MA: Springer; 2010:280-287.
42. Link WA. Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. Biometrics 2003 Dec;59(4):1123-1130. [doi: 10.1111/j.0006-341x.2003.00129.x] [Medline: 14969493]
43. Lieb S, Fallon SJ, Friedman SR, Thompson DR, Gates GJ, Liberti TM, et al. Statewide estimation of racial/ethnic populations of men who have sex with men in the U.S. Public Health Rep 2011;126(1):60-72 [FREE Full text] [doi: 10.1177/003335491112600110] [Medline: 21351396]
44. World Health Organization. Key population strategic information: recommended population size estimates of men who have sex with men: technical brief. UNAIDS. 2020. URL: https://www.unaids.org/sites/default/files/media_asset/2020-recommended-population-size-estimates-of-men-who-have-sex-with-men_en.pdf [accessed 2023-03-13]
45. Baral S, Turner RM, Lyons CE, Howell S, Honermann B, Garner A, et al. Population size estimation of gay and bisexual men and other men who have sex with men using social media-based platforms. JMIR Public Health Surveill 2018 Feb 08;4(1):e15 [FREE Full text] [doi: 10.2196/publichealth.9321] [Medline: 29422452]
46. Quaye S, Fisher Raymond H, Atuahene K, Amenyah R, Aberle-Grasse J, McFarland W, Ghana Men Study Group. Critique and lessons learned from using multiple methods to estimate population size of men who have sex with men in Ghana. AIDS Behav 2015 Feb;19(Suppl 1):S16-S23. [doi: 10.1007/s10461-014-0943-4] [Medline: 25704987]
47. Abdul-Quader AS, Baughman AL, Hladik W. Estimating the size of key populations: current status and future possibilities. Curr Opin HIV AIDS 2014 Mar;9(2):107-114 [FREE Full text] [doi: 10.1097/COH.0000000000000041] [Medline: 24393694]
48. Des Jarlais D, Khue PM, Feelemyer J, Arasteh K, Thi Huong D, Thi Hai Oanh K, et al. Using dual capture/recapture studies to estimate the population size of persons who inject drugs (PWID) in the city of Hai Phong, Vietnam. Drug Alcohol Depend 2018 Apr 01;185:106-111 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.11.033] [Medline: 29432973]
49. Vuylsteke B, Sika L, Semdé G, Anoma C, Kacou E, Laga M. Estimating the number of female sex workers in Côte d'Ivoire: results and lessons learned. Trop Med Int Health 2017 Sep;22(9):1112-1118 [FREE Full text] [doi: 10.1111/tmi.12911] [Medline: 28627038]
50. Baluku M, Wamala T. When and how do individuals transition from regular drug use to injection drug use in Uganda? Findings from a rapid assessment. Harm Reduct J 2019 Dec 23;16(1):73 [FREE Full text] [doi: 10.1186/s12954-019-0350-2] [Medline: 31870396]

51. Apodaca K, Doshi RH, Ogwal M, Kiyingi H, Aluzimbi G, Musinguzi G, et al. Capture-recapture among men who have sex with men and among female sex workers in 11 towns in Uganda. JMIR Public Health Surveill 2019 Apr 03;5(2):e12316 [FREE Full text] [doi: 10.2196/12316] [Medline: 30942697]
52. Okiria AG, Bolo A, Achut V, Arkangelo GC, Michael ATI, Katoro JS, et al. Novel approaches for estimating female sex worker population size in conflict-affected South Sudan. JMIR Public Health Surveill 2019 Mar 18;5(1):e11576 [FREE Full text] [doi: 10.2196/11576] [Medline: 30882356]
53. Neugebauer R, Wittes J. Voluntary and involuntary capture-recapture samples--problems in the estimation of hidden and elusive populations. Am J Public Health 1994 Jul;84(7):1068-1069. [doi: 10.2105/ajph.84.7.1068] [Medline: 8017525]

Abbreviations

3S-CRC: three-source capture-recapture
CBO: community-based organization
CDC: Centers for Disease Control and Prevention
CRC: capture-recapture
CS: credibility set
FSW: female sex worker
IBBSS: Integrated Behavioral and Biological Surveillance Survey
LCMCR: latent-class capture-recapture
MCMC: Markov Chain Monte Carlo
MSM: men who have sex with men
ODK: Open Data Kit
PSE: population size estimate
PWID: people who inject drugs
RDS: respondent-driven sampling
UNAIDS: Joint United Nations Program on HIV/AIDS
WHO: World Health Organization

Edited by Y Khader; submitted 30.09.22; peer-reviewed by K McLaughlin, C Garcia; comments to author 09.11.22; revised version received 24.12.22; accepted 03.02.23; published 27.03.23.

Please cite as:
Tuyishime E, Kayitesi C, Musengimana G, Malamba S, Moges H, Kankindi I, Escudero HR, Habimana Kabano I, Oluoch T, Remera E, Chukwu A
Population Size Estimation of Men Who Have Sex With Men in Rwanda: Three-Source Capture-Recapture Method
JMIR Public Health Surveill 2023;9:e43114
URL: https://publichealth.jmir.org/2023/1/e43114
doi:10.2196/43114
PMID:36972131

Corrigenda and Addenda

Correction: Racial/Ethnic Disparity in Association Between Fetal Alcohol Syndrome and Alcohol Intake During Pregnancy: Multisite Retrospective Cohort Study

Sarah Soyeon Oh[1,2], PhD; Bada Kang[3], PhD; Jewel Park[4], MD, MPH; SangMin Kim[5], MD; Eun-Cheol Park[2,6], MD, PhD; Seung Hee Lee[7], MSN; Ichiro Kawachi[1], MBChB, PhD

[1]Department of Social and Behavioral Sciences, Harvard TH Chan School of Public Health, Boston, MA, United States
[2]Institute of Health Services Research, Yonsei University College of Medicine, Seoul, Republic of Korea
[3]Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, Seoul, Republic of Korea
[4]Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[5]Harvard Medical School, Boston, MA, United States
[6]Department of Preventive Medicine and Public Health, Yonsei University College of Medicine, Seoul, Republic of Korea
[7]College of Nursing and Brain Korea 21 Four Project, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Bada Kang, PhD
Mo-Im Kim Nursing Research Institute
Yonsei University College of Nursing
50-1 Yonsei-Ro
Seodaemun-Gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 3283
Email: bdkang@yuhs.ac

Related Article:

Correction of: https://publichealth.jmir.org/2023/1/e45358

(JMIR Public Health Surveill 2023;9:e48460) doi:10.2196/48460

In "Racial/Ethnic Disparity in Association Between Fetal Alcohol Syndrome and Alcohol Intake During Pregnancy: Multisite Retrospective Cohort Study" (JMIR Public Health Surveill 2023;9:e45358) the authors noted one error.

The authorship was previously published as:

> *Sarah Soyeon Oh, SangMin Kim, Bada Kang, Jewel Park, Seung Hee Lee, Eun-Cheol Park, Ichiro Kawachi*

And will be changed to read as follows:

> *Sarah Soyeon Oh, Bada Kang, Jewel Park, SangMin Kim, Eun-Cheol Park, Seung Hee Lee, Ichiro Kawachi*

The correction will appear in the online version of the paper on the JMIR Publications website on April 28, 2023 together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 24.04.23; this is a non–peer-reviewed article;accepted 24.04.23; published 28.04.23.

Please cite as:
Oh SS, Kang B, Park J, Kim S, Park EC, Lee SH, Kawachi I
Correction: Racial/Ethnic Disparity in Association Between Fetal Alcohol Syndrome and Alcohol Intake During Pregnancy: Multisite Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e48460
URL: https://publichealth.jmir.org/2023/1/e48460
doi:10.2196/48460
PMID:37116189

Original Paper

Global, Regional, and National Prevalence of Gout From 1990 to 2019: Age-Period-Cohort Analysis With Future Burden Prediction

Qiyu He[1*], MD, PhD; Tsz-Ngai Mok[2,3*], MD, PhD; Tat-Hang Sin[4*], MD; Jiaying Yin[5], MPH; Sicun Li[2,6], BA; Yiyue Yin[7], MSc; Wai-Kit Ming[2,8*], MPH, MMSc, MD, PhD; Bin Feng[9*], MD

[1] State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences, Peking Union Medical College, Bejing, China

[2] Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong, Hong Kong

[3] Department of Orthopedic Surgery and Sports Medicine Center, The First Affiliated Hospital and The First Clinical College, Jinan University, Guangzhou, China

[4] Department of Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[5] School of Public Health, University of Hong Kong, Hong Kong, Hong Kong

[6] School of Humanities and Communication, Guangdong University of Finance and Economics, Guangzhou, China

[7] Department of Biostatistics, College of Science, City University of Hong Kong, Hong Kong, Hong Kong

[8] School of Public Policy and Management, Tsinghua University, Beijing, China

[9] Department of Orthopedic Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[*] these authors contributed equally

Corresponding Author:
Bin Feng, MD
Department of Orthopedic Surgery
Peking Union Medical College Hospital
Chinese Academy of Medical Sciences and Peking Union Medical College
No 1, Shuaifuyuan Street
Beijing
China
Phone: 86 69155200
Email: fengbin@pumch.cn

Abstract

Background: Gout is a common and debilitating condition that is associated with significant morbidity and mortality. Despite advances in medical treatment, the global burden of gout continues to increase, particularly in high–sociodemographic index (SDI) regions.

Objective: To address the aforementioned issue, we used age-period-cohort (APC) modeling to analyze global trends in gout incidence and prevalence from 1990 to 2019.

Methods: Data were extracted from the Global Burden of Disease Study 2019 to assess all-age prevalence and age-standardized prevalence rates, as well as years lived with disability rates, for 204 countries and territories. APC effects were also examined in relation to gout prevalence. Future burden prediction was carried out using the *Nordpred* APC prediction of future incidence cases and the Bayesian APC model.

Results: The global gout incidence has increased by 63.44% over the past 2 decades, with a corresponding increase of 51.12% in global years lived with disability. The sex ratio remained consistent at 3:1 (male to female), but the global gout incidence increased in both sexes over time. Notably, the prevalence and incidence of gout were the highest in high-SDI regions (95% uncertainty interval 14.19-20.62), with a growth rate of 94.3%. Gout prevalence increases steadily with age, and the prevalence increases rapidly in high-SDI quantiles for the period effect. Finally, the cohort effect showed that gout prevalence increases steadily, with the risk of morbidity increasing in younger birth cohorts. The prediction model suggests that the gout incidence rate will continue to increase globally.

Conclusions: Our study provides important insights into the global burden of gout and highlights the need for effective management and prophylaxis of this condition. The APC model used in our analysis provides a novel approach to understanding the complex trends in gout prevalence and incidence, and our findings can inform the development of targeted interventions to address this growing health issue.

(JMIR Public Health Surveill 2023;9:e45943) doi:10.2196/45943

KEYWORDS

gout; prevalence; age-period-cohort analysis; Global Burden of Disease Study 2019; prediction; Bayesian age-period-cohort analysis; Norped age-period-cohort analysis

Introduction

Background

Gout is a ubiquitous chronic disease caused by an elevation of urate concentrations in articular and nonarticular structures. It can lead to highly intensive pain, lingering discomfort, and inflammation, as well as limited range of motion at the joints, with a high population burden [1]. With a high recurrence rate, many patients cannot undergo standard therapy, which leads to poor health-related quality of life [2]. Besides the increased personal financial burden, direct and indirect annual financial burdens exist for those who cannot control gout well [3]. Patients with cardiovascular disease who have also been diagnosed with gout have higher mortality rates, regardless of age, sex, and comorbidities factors [4]. The mortality gap of gout, which is the association with premature death, has stayed consistent over 2 decades [5,6].

Despite the heavy population burden, especially in the high–sociodemographic index (SDI) countries, the latest studies only focused on the regional or short-period gout burden [7-10]. Because of technological advancements, people's lifestyles have changed dramatically over the past 2 decades. The tendency and pattern of gout have changed owing to the lifestyle changes worldwide, even with treatment recommendations by many rheumatological societies. However, health policies have not provided any effective action strategies to cope with the steadily increasing trend of gout incidence [11-13]. To break through the barrier, it is essential to provide different approaches regarding the epidemiology of gout to evaluate public health development over time. In this regard, an analysis of prevalence trends, with a particular focus on their associations with age-period-cohort (APC) effects, has the potential to delineate the success of different aspects of health care delivery and identify the remaining treatment gaps [14,15].

Using the APC model, we can potentially mine ethnic disparities and geographic differences, as well as the effectiveness of the monitoring system. As the comparison of collinearity factors should be determined through the additional constraints [16]. This is a valid model for providing valuable insight into various public health issues. It could compare the efficiency of health care provision of different regions and identify the emblematic countries or regions that have acted efficiently on gout prevalence [17-20].

Objectives

This study aims to use the APC model to analyze the global-, regional-, and national-level prevalence of gout. The *Nordpred* and Bayesian APC (BAPC) prediction models were used to estimate future gout incidence. Although Jeong et al [21] have analyzed the global burden of gout and its associations with various factors such as sex, age, and region, as well as risk factors such as obesity and potential genetic factors, using Global Burden of Disease (GBD) Study 2019 data, this study aims to analyze global trends in gout incidence and prevalence and discuss local health care delivery to learn what may lead to novel action for better management in the future [21]. Therefore, by analyzing the GBD Study 2019 data from 1990 to 2019 by age, sex, and SDI, we have provided an in-depth gout report.

Methods

Overview

All data were extracted from the GBD Study 2019 [22], which provides descriptive epidemiological data on 369 diseases and injuries in 204 countries and territories from 1990 to 2019, including incidence, prevalence, mortality, years of life lost, years lived with disability (YLD), and disability-adjusted life years [23]. To identify the data source from a systematic review of published studies, searches of government and international organization websites, published reports, and primary data sources were undertaken, and librarians reviewed each data source. Diagnosis codes from both the International Classification of Diseases, Ninth Revision (ICD-9; 274.x), and the International Classification of Diseases, Tenth Revision (ICD-10; M10.x), were used to identify cases of gout in the GBD Study 2019 data set. The search terms used included *acute gout*, *gouty arthropathy*, *gouty neuritis*, and *primary gout*. According to the American College of Rheumatology diagnostic criteria for gout, the gold standard consists of checking for the presence of monosodium urate crystals in the synovial fluid of the affected joint.

Ethical Considerations

This study used data from the GBD Study 2019, which was approved by the institutional review board of the University of Washington School of Medicine. As this is a secondary analysis of existing data, no additional human participant research ethics review or informed consent was required. The original data collection obtained informed consent from study participants or was granted exemptions by the institutional review board. Study data were anonymized and deidentified to protect the privacy and confidentiality of study participants.

Analysis of Overall Temporal Trends

Using all-age prevalence (number) and age-standardized prevalence rates, we assessed the temporal trend in gout prevalence over the study period from 1990 to 2019. Age-standardized population prevalence was calculated using the global data from the GBD Study 2019. Measuring age-standardized population prevalence can provide a method that compares prevalence among countries and SDI quantiles with different populations. As they provide an indirect insight into epidemiological severity, YLD data were examined using an age-standardized YLD rate for the study period from 1990 to 2019. YLD refers to the number of years that an individual with an illness that affects their quality of life, in this case, a patient with a confirmed diagnosis of gout, survives until death [24].

APC Analysis

The APC model was used mainly to analyze gout morbidity trends and predict future gout burden. The model considers 3 factors: age, period, and cohort. Time trends of diseases are usually explained by age effects, period effects, and cohort effects, as well as random variation. The critical problem of descriptive epidemiology in this case was to describe the APC results of disease occurrence under the condition of gout rates. In the APC model, the period effect refers to the change of human factors affecting the gout rate in the population, such as the development of disease diagnosis technology, screening and early detection, changes in disease definition and registration, and treatment improvement. These human factors may affect disease rates in different periods, resulting in period effects. The age effect is one of the most critical determinants of disease occurrence. The cohort effect refers to changes in disease rates caused by different levels of exposure to risk factors in different generations. The APC model estimates the APC effects by SDI quintiles. The age effect uses longitudinal and cross-sectional age curves, representing age-related gout prevalence; fitted temporal trends display the period effect, period risk ratio, and period deviations; and cohort effects are displayed using cohort RR, local drifts, and cohort deviations. In the APC model, net drifts represent the overall trend in prevalence or incidence over time, whereas local drifts reflect the trends within a specific time period or age group. Age effects measure the changes in prevalence or incidence associated with aging, period effects refer to changes that affect all age groups at a particular time point, and cohort effects reflect the changes in prevalence or incidence for individuals born in the same year. The model also included the issue of covariance among age, period, and cohort, which was addressed by the identifiability constraint that the sum of the APC coefficients must equal 0. However, this constraint can lead to difficulties in estimating the separate APC effects. To address this issue, the estimable functions approach is used, which involves combining linear constraints on the APC coefficients to estimate these effects. Specifically, the sum-to-0 constraint is used to estimate the net drifts (long-term trends) and local drifts (deviations from the long-term trends) for each effect. The net drifts correspond to the overall change in the outcome over time, whereas the local drifts correspond to the deviations from the overall change. The APC model can thus provide insights into the unique APC effects on gout prevalence and incidence over time. In the case of gout, the model grouped age effect into 10-year intervals, period effect into 5-year intervals, and cohort effect into 5-year intervals to simplify the analysis. The x intercept for the cohort effect is set at the year 1945, and the reference choice (the baseline for comparison) is indicated with a dashed line.

Prediction Model

To reflect the trends in gout burden, the number of new cases of gout from 2019 to 2042 was predicted using a *Nordpred* APC analysis by sex. The R software (R Foundation for Statistical Computing) package *Nordpred* was used for this analysis, which takes into account changing rates and changing population structures, as demonstrated in previous studies [25]. In addition, to facilitate comparison with the predicted results, we calculated the absolute number of events that would occur if the rates remained stable (baseline reference), decreased by 1% per year (optimistic reference), or increased by 1% per year (pessimistic reference) based on actual observed rates in 2019. To validate the stability of the prediction results, we performed a sensitivity analysis using the BAPC model integrated nested Laplace approximations (INLA) with the *BAPC* and *INLA* packages in R.

Statistical Analysis

Our study used the APC model to analyze gout morbidity and mortality trends and predict future gout burden. The APC model estimates the time trend of prevalence within each age group, expressed as the annual percentage change of age-specific prevalence, reflecting the birth cohort effect trend, whereas Jeong et al [21] used counts, age-standardized rates, age-standardized percentage changes of prevalence, incidence, and YLD to quantify global trends in the burden of gout using the GBD Study 2019 standard population from 1990 to 2019. The Wald chi-square test results indicate whether there is a birth cohort effect and whether the test results are statistically different. An important implication is that a single age-standardized rate curve and estimated annual percentage change value do not adequately describe the time trend in each age group. A drift absolute value of >1% is considered a material change, which is tested by the Wald chi-square test. All analyses were 2-sided, and statistical significance was considered at *P*<.05. Statistical analysis was conducted using R software (version 3.6.3), and the result was presented using 95% uncertainty intervals (UIs).

Results

Global and Regional Trends in Gout Incidence From 1990 to 2019

The global population of individuals with gout increased in the past 30 years from 22 million (95% UI 17.52-27.37) to 53 million (95% UI 43.38-66.34). The growth rate of gout *incidence* was 63.44% (95% UI 57.33%-68.78%). YLD increased from 0.69 million (95% UI 0.44-99.37) to 1.67 million (95% UI 1.07-2.39). The growth rate of global YLD was 51.12% (95% UI 45.91%-56.51%). The sex ratio stayed consistent at 3:1 (male to female), but the gout *incidence* kept increasing over time. The number of male individuals with gout increased from 16 million (95% UI 13.22-20.68) to 40 million (95% UI 32.75-50.10). The growth rate of gout incidence in male individuals was 70.15% (95% UI 64.55%-75.74%). The number of female individuals with gout increased from 5.3 million (95% UI 4.25-6.75) to 13.2 million (95% UI 106.03-163.79), and the growth rate of gout incidence in female individuals was 68.70% (95% UI 63.70%-73.43%). In the high- and middle-SDI regions, the number of individuals with gout skyrocketed from 7 million (95% UI 5.67-8.84) to 17 million (95% UI 14.19-20.62) and from 6 million (95% UI 4.64-7.26) to 15 million (95% UI 12.28-19.40), respectively, with growth rates of 94.3% (95% UI 84.33%-107.15%) and 90.46% (81.34%-98.45%), respectively (Figure 1; Table 1).

Figure 1. The age effect for gout is shown with all sociodemographic index (SDI) quantiles on total prevalence cases (number) and prevalence rate per 100,000 population from (A) 1990 to (B) 2019. The age effect is shown with all SDI quantiles on years lived with disability (YLD; number) and the YLD rate per 100,000 population from (C) 1990 to (D) 2019. The global tendency under the age effect on both gender.

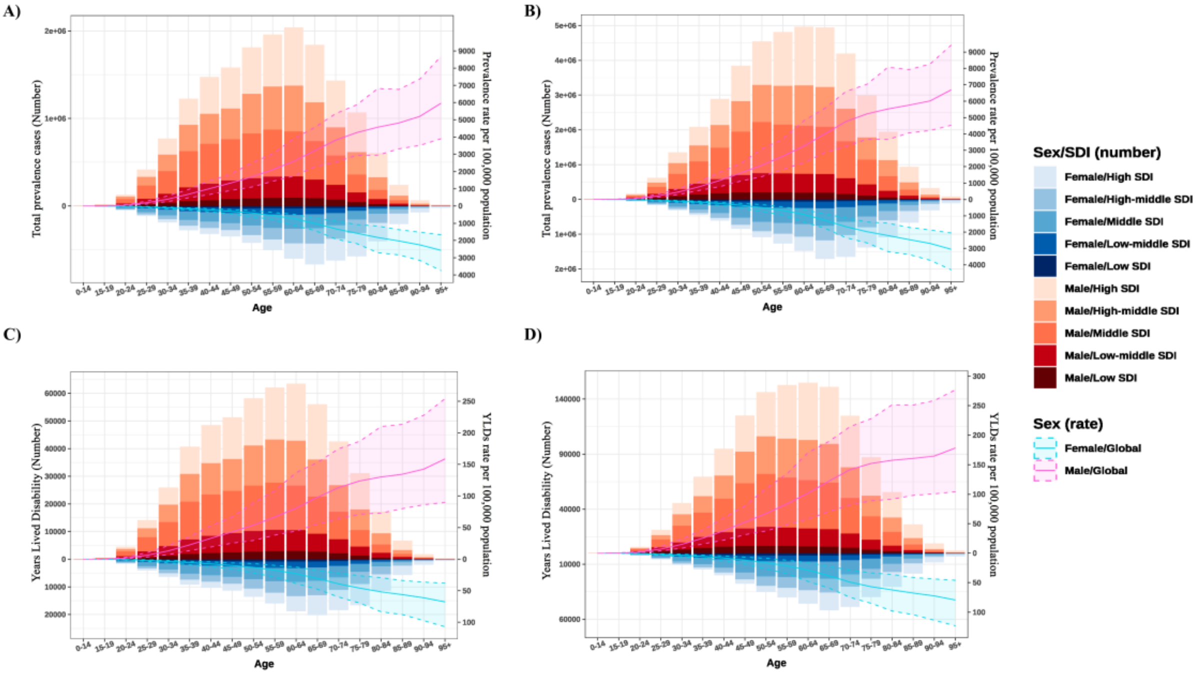


Table 1. Trends in gout prevalence and years lived with disability (YLD) across sociodemographic index (SDI) quintiles from 1990 to 2019.[a]

	Global, 1990-2019	High SDI, 1990-2019	High-middle SDI, 1990-2019	Middle SDI, 1990-2019	Low-middle SDI, 1990-2019	Low SDI, 1990-2019
All-age incidence						
Cell number×10,000, n (95% UI[b])	2206.37 (1751.87-2737.37)-5387.18 (4338.32-6634.23)	711.33 (566.90-884.13)-1713.24 (1419.06-2061.89)	538.4 (428.55-671.04)-1239.95 (981.62-1549.99)	585.1 (464.74-726.11)-1549.82 (1228.03-1940.43)	265.61 (210.77-331.26)-635.63 (505.79-796.65)	105.05 (83.22-130.97)-246.51 (196.23-307.16)
Annual change, % (95% UI)	63.44 (57.33-68.78)	71.46 (62.62-81.17)	82.60 (75.02-89.65)	74.35 (65.47-82.43)	60.86 (55.38-65.70)	16.23 (13.61-18.94)
All-age prevalence rate						
Rate per 100,000, n (95% UI)	412.42 (327.46-511.67)-696.25 (857.42-6634.23)	865.35 (689.65-1075.57)-1690.61 (1400.31-2034.66)	468.00 (372.51-583.30)-866.85 (686.26-1083.60)	340.82 (270.71-422.96)-646.68 (512.41-809.67)	235.13 (186.58-293.25)-360.34 (286.73-451.62)	198.91 (157.58-247.98)-246.51 (196.23-307.16)
Annual change, % (95% UI)	69.24 (63.79-74.51)	94.32 (84.33-107.15)	84.66 (78.25-89.98)	90.46 (81.34-98.45)	53.25 (50.48-58.90)	11.47 (9.38-13.71)
YLD						
Cell number×10,000, n (95% UI)	69.12 (43.74-99.37)-167.4 (106.80-239.34)	21.92 (13.93-31.51)-52.03 (33.75-73.50)	16.93 (10.60-24.45)-38.83 (24.16-55.90)	18.61 (11.70-26.86)-48.85 (30.57-70.67)	8.34 (5.24-12.01)-19.88 (12.40-28.57)	3.29 (2.07-4.73)-7.75 (4.84-11.22)
Annual change, % (95% UI)	51.12 (45.91-56.51)	65.90 (56.91-77.24)	60.64 (53.55-66.88)	57.33 (48.88-65.49)	43.12 (37.42-49.04)	17.96 (13.66-22.58)
APC[c] model estimates						
Annual net drift[d] of prevalence, % (95% CI)	0.85 (0.80-0.89)	1.596 (1.5332-1.6588)	0.9507 (0.9014-1)	0.7075 (0.6497-0.7654)	0.3088 (0.2879-0.3297)	0.1497 (0.1382-0.1613)
Age-standardized prevalence rate						
Rate per 100,000, n (95% UI)	532.99 (425.35-657.86)-652.24 (528.56-798.60)	711.83 (565.01-887.95)-1041.01 (862.79-1244.91)	492.64 (393.93-611.45)-624.81 (495.85-778.15)	507.06 (405.34-632.91)-597.19 (476.09-740.10)	403.01 (324.18-502.21)-439.20 (351.05-550.29)	406.42 (324.02-512.66)-424.68 (339.15-532.13)
Annual change, % (95% UI)	23.69 (21.65-26.25)	47.95 (40.56-57.67)	28.97 (26.99-30.76)	19.79 (17.53-21.97)	10.56 (9.15-12.03)	5.74 (4.43-7.15)
Age-standardized YLD rate						
Rate per 100,000, n (95% UI)	16.55 (10.44-23.79)-20.23 (12.91-28.88)	22.03 (13.87-31.73)-32.05 (20.59-44.73)	15.41 (9.67-22.17)-19.62 (12.31-28.20)	15.89 (9.95-22.93)-18.70 (11.89-26.99)	12.44 (7.88-18.03)-13.59 (8.61-19.60)	12.48 (7.93-18.03)-13.11 (8.27-19.07)
Annual change, % (95% UI)	25.94 (23.40-28.94)	43.04 (35.31-52.93)	33.70 (29.86-37.29)	20.72 (17.16-24.47)	15.99 (12.58-19.37)	12.68 (9.27-16.21)

[a] Age-standardized mortality rate is computed by direct standardization with the global standard population in the Global Burden of Disease Study 2019.

[b] UI: uncertainty interval.

[c] APC: age-period-cohort.

[d] Net drifts are estimates derived from the age-period-cohort model and denote overall annual percentage change in mortality, which captures the contribution of the effects from calendar time and successive birth cohorts.

Global and Regional Trends in Gout Prevalence From 1990 to 2019

The gout prevalence trends are similar to the incidence trends. The global all-age prevalence rate increased from 412.42 (95% UI 327.46-511.67) to 696.25 (95% UI 857.42-6634.23) per 10,000 population, and the age-standardized prevalence rate increased from 532.99 (95% UI 425.35-657.86) to 652.24 (95% UI 528.56-798.60) per 10,000 population (Figure 2; Table 1). The high-SDI quantiles have the highest increase in prevalence from 711.83 (95% UI 565.01-887.95) to 1041.01 (95% UI 862.79-1244.91) per 10,000 population, an increase of 47.95%, followed by the high-middle–SDI quantiles, where the age-standardized prevalence increased from 492.64 (95% UI 393.93-611.45) to 624.81 (95% UI 495.85-778.15) per 10,000 population. The prevalence depends on the regions' SDI development (Figure 3; Table 1).

Figure 2. The period effect for gout is shown with all sociodemographic index (SDI) quantiles on (A) total prevalence cases (number) and age-standardized prevalence rate per 100,000 population and (B) years lived with disability (YLD; number) and standardized YLD rate per 100,000 population. The global tendency under the age effect on both sexes is shown as line charts.

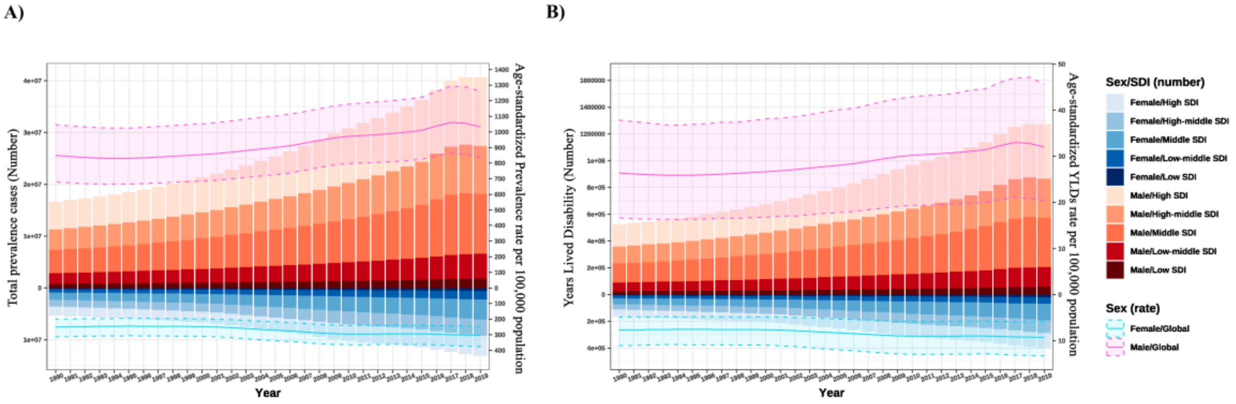
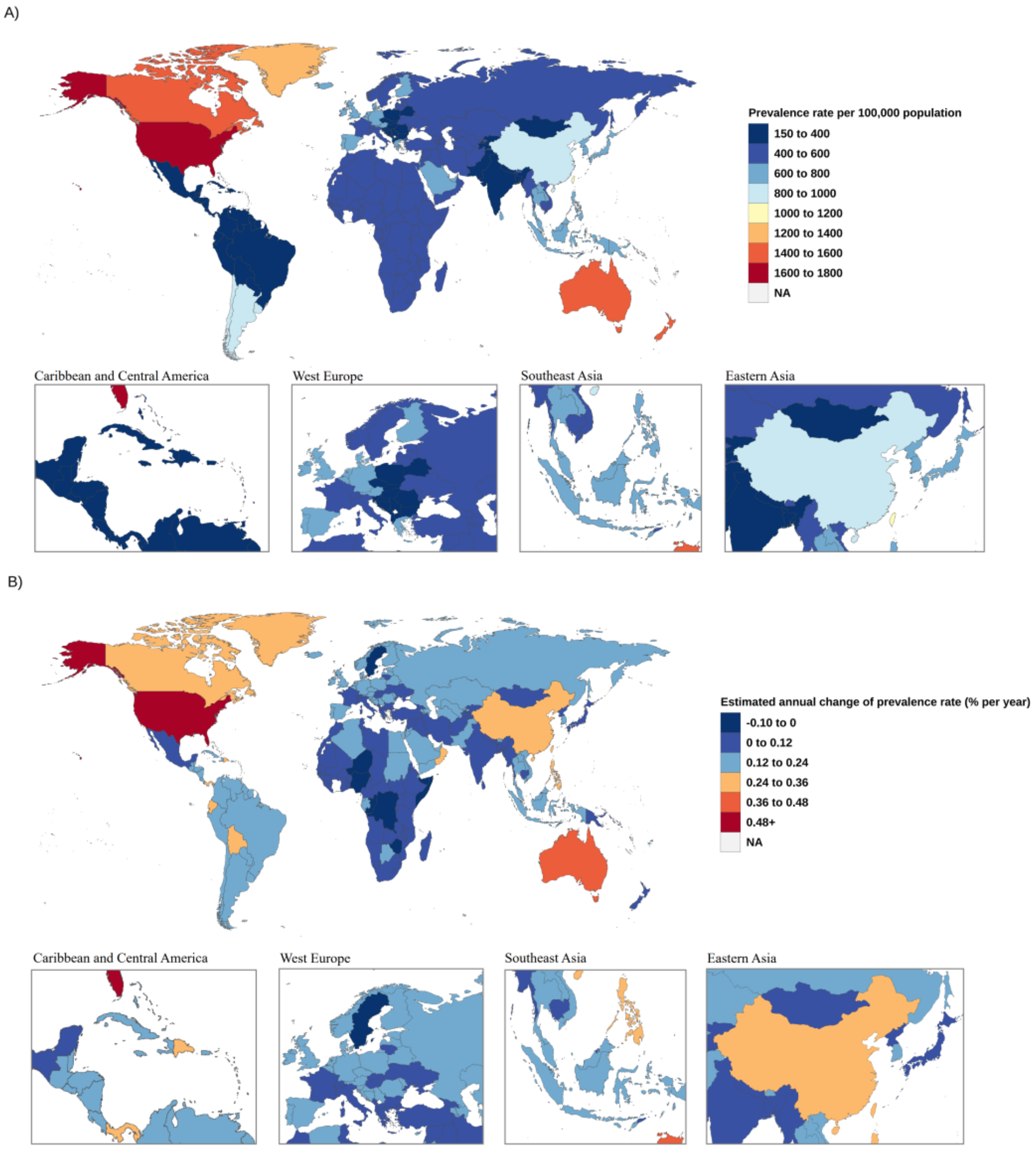


Figure 3. (A) The all-age prevalence for gout in 2019 in 204 countries and territories. (B) Net drift of gout prevalence from 1990 to 2019 in 204 countries and territories.



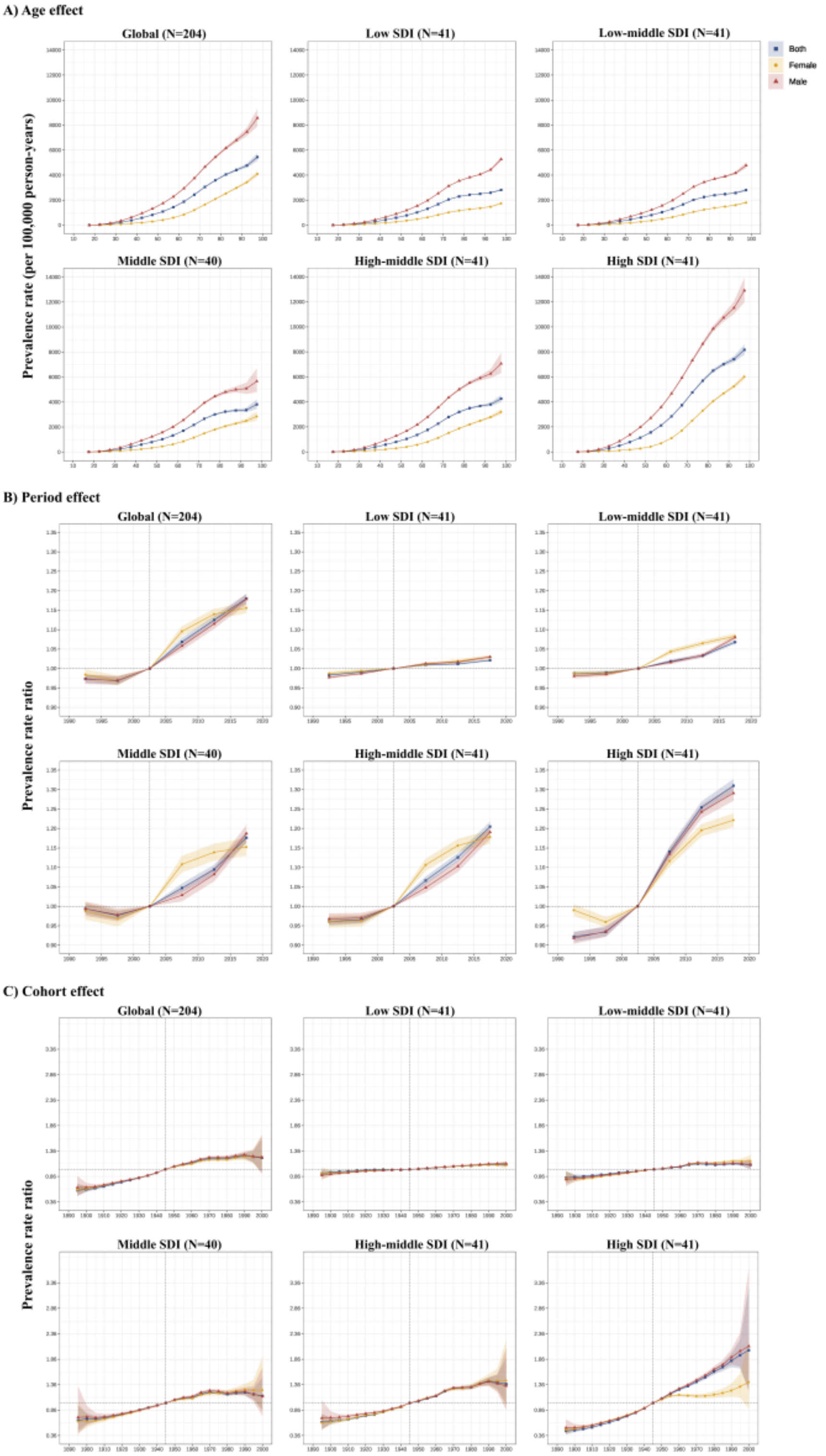
APC Effects on Gout Prevalence

To extend the previous studies, we used the APC model to analyze gout burden [21]. Globally, the prevalence of the age effect increases steadily from the age of 30 years in both sexes. The prevalence rises the most readily in high-SDI quintiles across all age groups compared with other quintiles.

Regarding the period effect, gout prevalence witnessed a growth tendency across most SDI quintiles over the study period, with only low-SDI countries remaining nearly constant over the past 2 decades. High-SDI, high-middle–SDI, and middle-SDI countries had the most notable increase in period prevalence rates across the period from 1990 to 2019. By contrast, low-SDI and low-middle–SDI countries stayed relatively stable in terms of prevalence rates. Surprisingly, in terms of the period effect, the gout prevalence rate in female individuals was higher than that in male individuals across the period from 2003 to 2015 in low-middle–SDI, middle-SDI, and high-middle–SDI countries. Still, the data showed that the gout prevalence rate in male individuals was higher than that in female individuals (Figure 1).

The morbidity is increased and the younger age population is readily exposed to the risk factor. The growth of cohort effects was remarkable in high-SDI countries. High-SDI countries had progressively increasing morbidity rates among those born after 1900. The sex ratio has shifted in middle-SDI and high-middle–SDI countries, where female individuals have a higher morbidity rate than male individuals in terms of the cohort effect (Figure 4).

Figure 4. Age-period-cohort effects on gout prevalence by sociodemographic index (SDI) quintiles. (A) Age effects are shown by the fitted longitudinal age curves of prevalence (per 100,000 person-years) adjusted for period deviations. (B) Period effects are shown by the relative risk of prevalence (prevalence rate ratio) and computed as the ratio of age-specific rates from 1990 to 1994 (the referent period) to 2015-2019. (C) Cohort effects are shown by the relative risk of prevalence and computed as the ratio of age-specific rates from the 1925 cohort to the 2015 cohort, with the referent cohort set at 1960. The dots and shaded areas denote prevalence rates and prevalence rate ratios, respectively, and their corresponding 95% CIs.

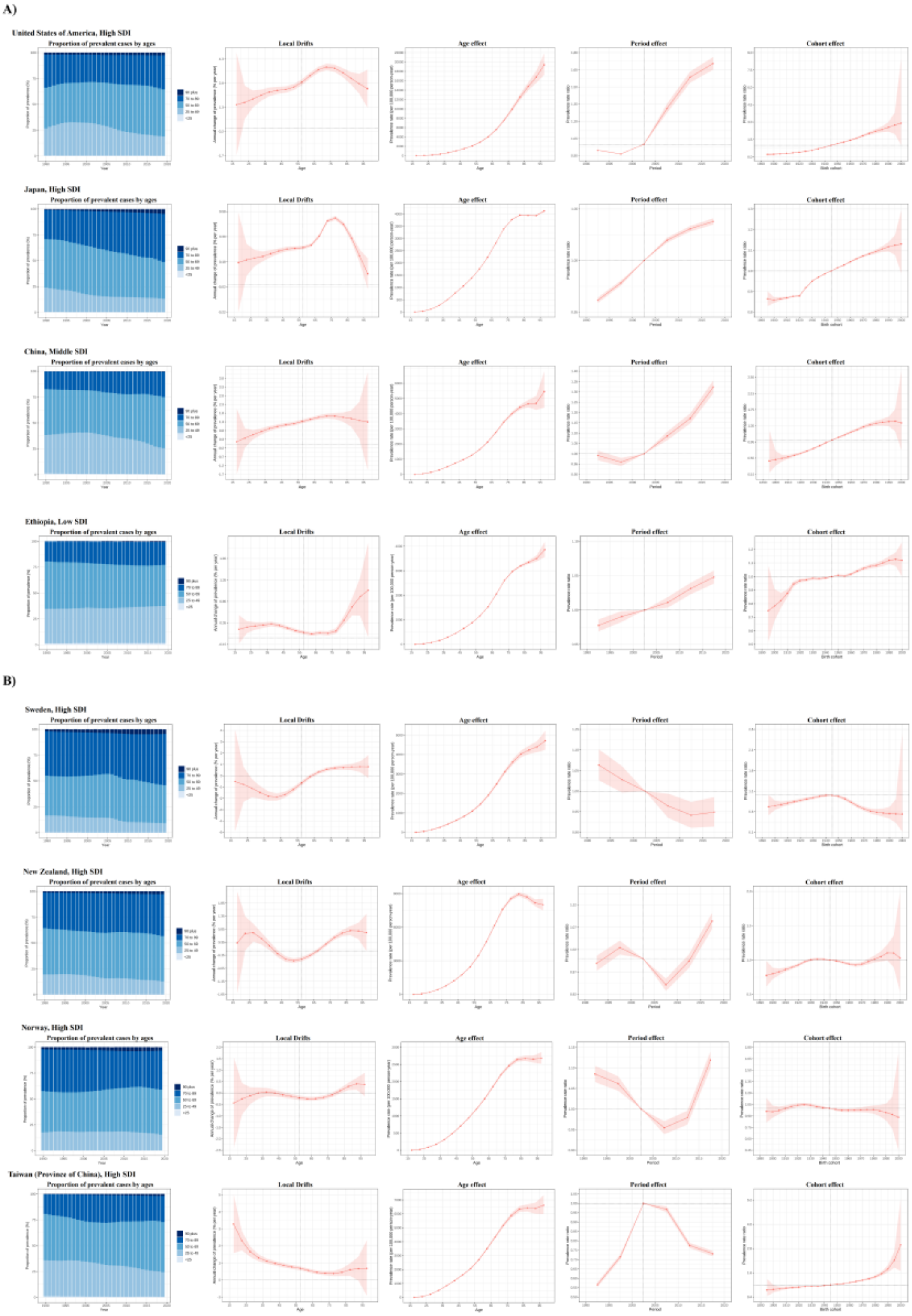


APC Effects in Typical Countries or Regions

Several countries or regions across SDI quintiles displayed significant developments in gout morbidity by APC effects around the world. The United States showed a trend typical of high-SDI countries, where growing gout prevalence was measured across all age groups. In Japan, another example of a high-SDI country, increased morbidity stood out for its notable net drift and demonstrated an emerging transition in the age distribution of morbidity. Moreover, China had a shift in the age distribution of morbidity, with significantly increased risk in those born after 1970, and the gout prevalence increased dramatically after 2003 in this middle-SDI country. Regarding low-SDI countries, Ethiopia showed a similar phenomenon to that in high-SDI countries because the prevalence kept growing over time. People aged <75 years had significant local drift in terms of being diagnosed with gout. People born after 2000 had a significantly higher risk of developing gout (Figure 5A).

Countries displaying relatively unfavorable APC effects on prevalence are shown in Figure 5B. Sweden, a high-SDI country, showed increasing risk throughout the age effect, but the period effect reduced the prevalence. In addition, people born after 2000 in Sweden have a lower risk of developing gout. New Zealand and Norway showed a descending trend in the period effect but rebooted the prevalence rate after 2008. By contrast, Taiwan showed the opposite result, with the prevalence rate consistently decreasing after 2003.

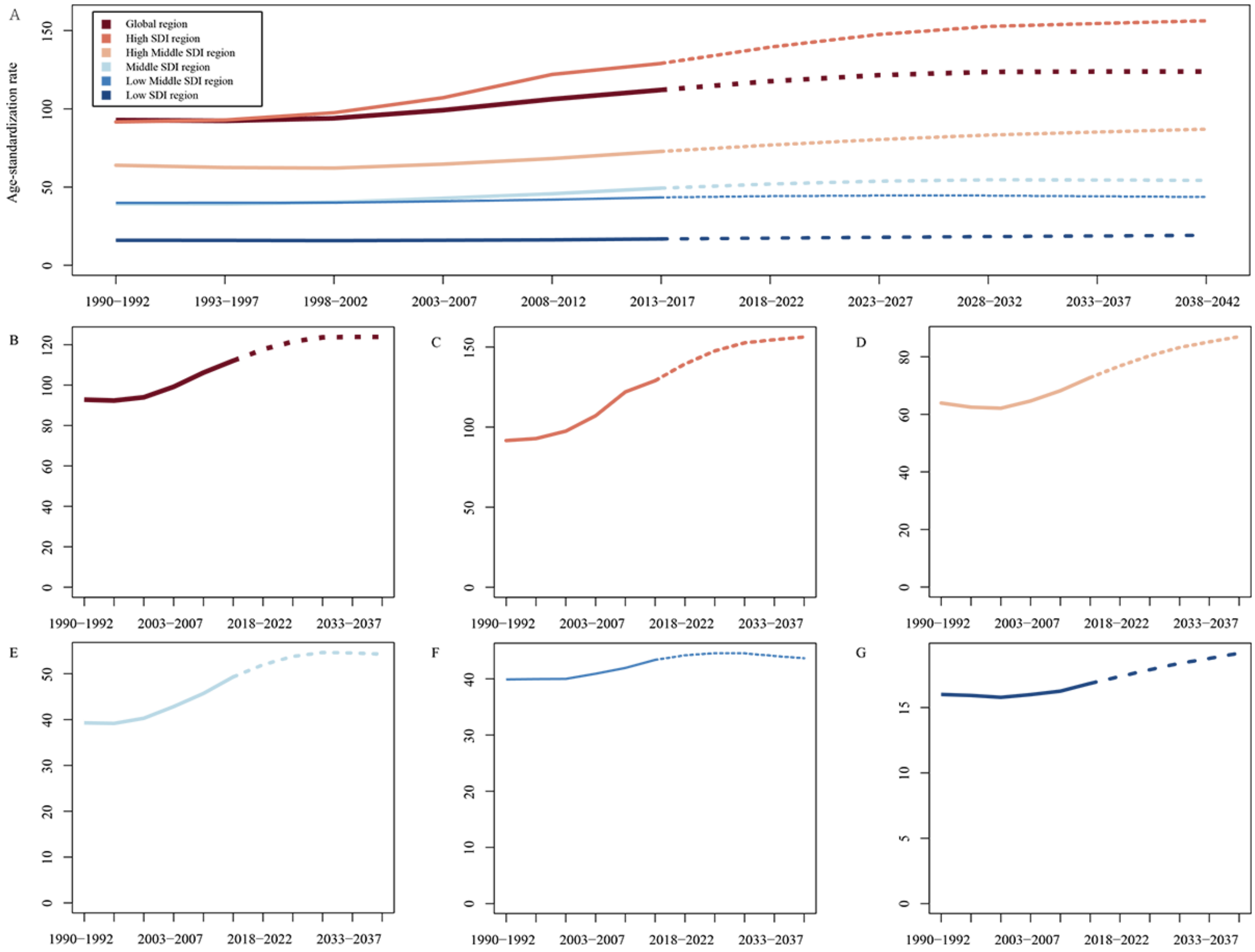
Figure 5. Favorable (A) and unfavorable (B) age-period-cohort effects on exemplar countries across sociodemographic index (SDI) quintiles. Age distribution of deaths shows the relative proportion of morbidity from each age group from 1990 to 2019. Local drifts indicate the annual percentage change in prevalence rate (%) across 5-year age groups (from 0-4 to 65-69 years). Age effects are represented by the fitted longitudinal age curves of prevalence (per 100,000 person-years) adjusted for period deviations. Period effects are represented by the relative risk of prevalence (prevalence rate ratio) and computed as the ratio of age-specific rates in each period compared with the referent period from 1990 to 1994. Cohort effects are represented by the relative risk of prevalence (prevalence rate ratio) and computed as the ratio of age-specific rates in each cohort compared with the referent 1960 cohort. The shaded areas indicate the corresponding 95% CIs of each point estimate. See high-resolution image in Multimedia Appendix 1.



Trends in Gout Incidence Over Time

The predicted incidence rates of gout per 100,000 population were calculated using the *Nordpred* APC prediction for different SDI regions (Figure 6). Overall, the predicted incidence rates of gout increased over time in all regions. The global predicted incidence rate increased from 92.79 per 100,000 population in the period from 1990 to 1992 to 123.83 per 100,000 population in the period from 2038 to 2042. The highest predicted incidence rate was observed in the high-SDI quantiles region, with a predicted incidence rate of 156.28 per 100,000 population in the period from 2038 to 2042. The lowest predicted incidence rate was observed in the low-SDI quantiles region, with a predicted incidence rate of 19.16 per 100,000 population in the period from 2038 to 2042. There were also notable differences in the predicted incidence rates among regions; for instance, the predicted incidence rates in the high-middle–SDI quantiles region were consistently lower than those in the high-SDI quantiles region but higher than those in the middle-SDI quantiles and low-middle–SDI quantiles regions. Similarly, the predicted incidence rates in the low-middle–SDI quantiles region were consistently lower than those in the middle-SDI quantiles region but higher than those in the low-SDI quantiles region.

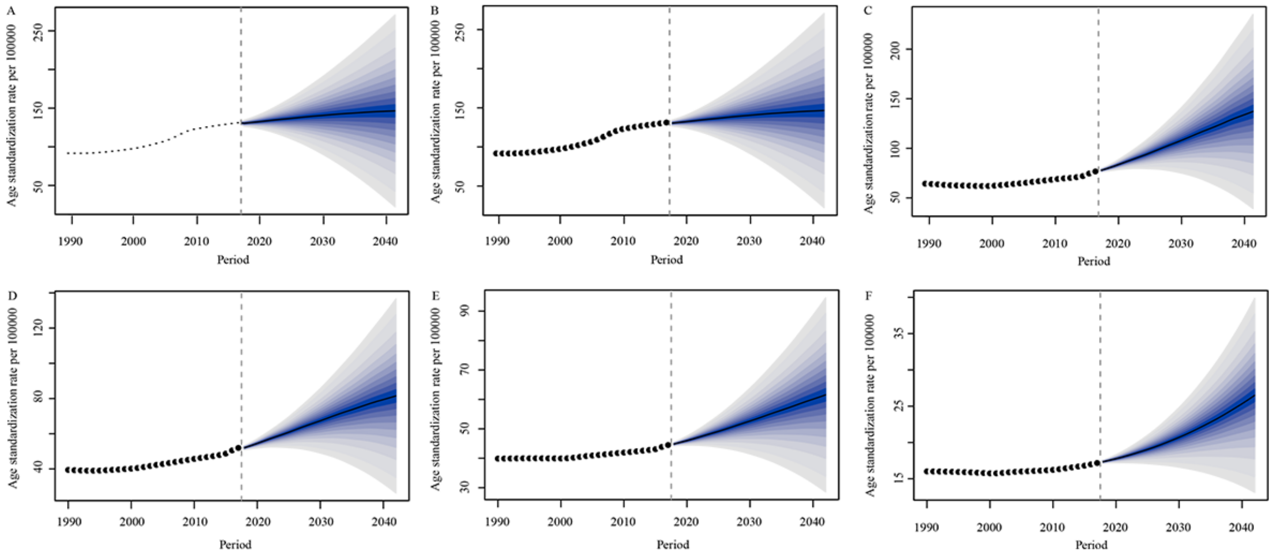
Figure 6. Trends in gout by Nordpred age-period-cohort prediction model in (A) global scale. To minimize the scale bias, the plots were separated as follows: (B) global scale, (C) high–sociodemographic index (SDI) quantiles, (D) high-middle–SDI quantiles, (E) middle-SDI quantiles, (F) low-middle–SDI quantiles, and (G) low-SDI quantiles.



Overall, the number of cases of gout increased over time in all regions. Globally, in 2023, a total of 10,016,336 cases were reported, which is projected to increase to 12,082,807 cases in 2035. The low-SDI quantiles region had the lowest number of cases, with 490,786.4 cases reported in 2018, which is projected to increase to 788,807.5 cases in 2042. There were also notable differences in the number of cases among regions; for instance, the high-SDI quantiles region had the second highest number of cases, with 2,242,423 cases reported in 2018, which is projected to increase to 2,395,031 cases in 2035. The low-middle–SDI quantiles region had a slightly higher number of cases than the middle-SDI quantiles region, but the trend is expected to reverse by 2035, with the middle-SDI quantiles region projected to have more cases (Multimedia Appendix 2).

The BAPC prediction model was used as a sensitivity analysis to validate the findings from the APC analysis. Consistent with the APC analysis, the BAPC prediction model suggested that there would be an overall increase in the age-standardized incidence of the disease from 1990 to 2042 across all 6 regions. The highest increase in age-standardized incidence is expected to occur in the high-middle–SDI quantiles, with the incidence increasing from 64.30 per 100,000 population to 137.14 per 100,000 population. The high-SDI quantiles are also expected to experience a significant increase in age-standardized incidence, from 91.68 per 100,000 population to 146.51 per 100,000 population. The middle-SDI quantiles are expected to see an increase from 39.44 per 100,000 population to 81.48 per 100,000 population. The low-middle–SDI quantiles are expected to experience an increase from 39.86 per 100,000 population to 61.54 per 100,000 population. The low-SDI quantiles are expected to experience the smallest increase in age-standardized incidence, from 15.99 per 100,000 population to 26.49 per 100,000 population. Overall, the BAPC prediction model suggests that there will be a significant increase in the incidence of the disease across all regions, with the highest increases occurring in regions with higher-SDI quantiles (Figure 7).

Figure 7. Trends in gout by Bayesian age-period-cohort prediction model in (A) global scale, (B) high–sociodemographic index (SDI) quantiles, (C) high-middle–SDI quantiles, (D) middle-SDI quantiles, (E) low-middle–SDI quantiles, and (F) low-SDI quantiles.



Discussion

Principal Findings

Gout is among the most common chronic rheumatic and inflammatory arthritis diseases worldwide. Understanding trends in gout prevalence is key for governments to distribute health care resources better [26]. Despite gout having a low mortality rate and high *curable* rates without comorbidities, gout prevalence has steadily increased globally owing to inadequate management by both medical professionals and patients in many countries [27,28]. The net drift in the majority of high-SDI quantiles increased most readily in terms of both period and cohort effects. To our knowledge, this is the first study using the APC model to analyze global trends in gout prevalence, enabling comparisons among different countries and predictions of future incidence. Compared with previous publications [7,21], our study advances the field by providing a more comprehensive and nuanced understanding of gout prevalence trends and contributing to public health knowledge through future burden estimations. To be precise, the analysis of age effects, period effects, and cohort effects allowed us to determine the prevalence trends by period and birth cohort in individual regions, providing an in-depth evaluation of the quality of the health care system with regard to gout management. The model allowed a comparison among countries and the discovery of some countries or regions with effective interventions. By highlighting these responsive quantiles, we hope that inspiration will strike for a novel or potential idea to better manage gout.

This study illustrates the increase in gout prevalence over time from 1990 to 2019. This tendency would not change, according to our APC model prediction. The major challenge was that the highest gout prevalence occurred in the high-SDI quantiles where the quality of medical care is outstanding. This is an indication, perhaps, that advanced interventions and techniques are not sufficient to solve the problem; rather, this points to the need for the entire medical system to work together. With populations increasing worldwide and with an increase in the aging population, the situation regarding gout prevalence is delicately poised. This analysis of the gout burden in different regions, therefore, is important from the point of view of learning how the medical system works in countries that have demonstrated superior management of gout.

A previous study reported the global gout burden during the period from 1990 to 2010 [8]. With a 9-year data gap, it is necessary to perform updated research to learn trends in gout prevalence and identify the challenges our health system may face in the future. In addition, it is essential to evaluate the effectiveness of the gout management strategies used during the period from 2010 to 2019. Compared with the Disease Modelling Meta-Regression model, the APC model could divide data into 3 spatially varying parameters with conditioning. The conditioning order-dependent model was validated on the spatially varying coefficients; therefore, the model was available for the individual country calculations [16]. Consequently, the APC model was more favorable with regard to the Watanabe-Akaike information criterion than the model with independent random effects [29].

Our analysis of gout prevalence estimates suggests that many low- and middle-income countries, especially those with a high SDI, have had an increasing gout prevalence tendency. Still, the low-SDI countries have relatively stable conditions regarding gout prevalence. We also found that people born after 2000 were more likely to develop gout than those born after 1900, suggesting that awareness of health care resources and education for gout are insufficient [7]. It is possible that patients who experience gout flares may be ashamed of admitting their lifestyles, including the excessive intake of purine-rich foods and alcohol, and this may be combined with low health literacy, because of which they do not receive sufficient information on gout [30]. Most patients with gout visit inpatient clinics only for relief of gout flares, and it has been suggested that the weak relationship between the physician and the patient with gout is leading to treatment nonadherence [31,32]. Another possible explanation for the higher likelihood of gout development among individuals born after 2000 compared with those born after 1900 is the improvement in people's standard of living and dietary habits, which may have led to an earlier onset of gout [33]. This highlights the importance of considering multiple factors when analyzing the trends in gout prevalence and the burden of gout.

Age-standardized prevalence and YLD can provide a clear picture for comparisons among different economic countries or regions. Because of the varying population structure in different SDI countries and regions, we should interpret the results carefully. In lower-SDI countries, citizens tend to have shorter life expectancy and higher birth rates; therefore, the all-age prevalence is relatively lower than the standardized-age prevalence [34,35]. To cope with the actual situation, health care systems may consider investing more resources in *gout flare* services, especially effective treatment. However, the aging population structure is one of the common social issues in higher-SDI countries. Although the standardized-age prevalence is lower than the all-aged prevalence, because older adults have become a prominent part of the population and also because of rising obesity rates, the health burden would keep increasing. Prevention and long-term management are key to coping with the situation [36,37].

Although the standardized-age prevalence of chronic diseases is lower among younger age groups, the increasing prevalence of obesity and the aging population in the United States has contributed to a considerable health burden. As one of the high-SDI countries, the United States has experienced the most substantial incidence and prevalence rates in the world, with an estimated 1 million cases [38]. Although the United States provides one of the best medical services in the world, gout is the most common inflammatory arthritis disease with significant morbidity and mortality [39]. Gout management strategies, including monitoring and treatment options, are far from enough. According to our results, gout prevalence in the United States has yet to plateau. The rates of obesity, hypertension, and chronic kidney disease keep rising because of the aging population and longer life expectancy [40-42]. As a result, the quality of life may not be improved even with advanced medical services.

In contrast to the United States, Japan, a high-SDI country, has a lower prevalence rate, despite a steadily increasing trend. Japanese traditionally favor purine-rich foods, including seafood, miso, soy sauce, and umami broth [43]. A possible reason for Japan having a lower prevalence rate than the United States is the country's treat-to-target approach to gout [44,45]. The Japanese medical profession has suggested that patients who present with hyperuricemia and those without symptoms should undergo urate-lowering therapy because hyperuricemia is the etiology causing gout flares [46]. This approach is one of the breakthrough ideas for gout prevention. Checking for the presence of urate in the blood during routine examination may be a possible way to control and prevent gout.

China is the world's second most populous country, with a population of 1.4 billion. The prevalence of gout in China was lower than that of the United States and European countries in 1990, the reasons for which are differences in genetic factors and exposure [7]. In addition, in 1990, China's public health institutions were undeveloped, and most public health workers had a poor opinion of the whole system [47]. Therefore, it became a challenge to collect national-level data; as a result, detection accuracy was unsatisfactory, and it remained so until the public health system was reformed. Subsequently, because of rapid westernization and urbanization, the Chinese changed their lifestyles; thus, gout prevalence kept rising in this large-population country with higher diagnosis accuracy. However, various degrees of urbanization and levels of quality of hospitals in different regions or provinces could lead to an extensive range of prevalence [48-50]. It has been reported that the obesity rate could reflect gout prevalence indirectly [51]; thus, it may be an index for the public health professional for the surveillance and prediction of gout prevalence. In low and middle tropical countries, the focus is mainly on infectious diseases and parasites, with insufficient epidemiological surveillance. Because of the younger population, the noncommunicable disease burden is underestimated in these regions. After age standardization, the mortality rate in low and middle tropical countries is more severe than that in high-income countries owing to lack of proper management and health education [52]. To better distribute resources, governments should financially support public health institutions to screen for a *curable* chronic disease, which is gout in this case [53].

A few countries such as Sweden and Taiwan showed a descendant trend in gout prevalence in terms of the period effect. However, several studies have reported that gout management in these countries was not satisfactory owing to the low rate of urate-lowering therapy use, long-term management issues, and poor gout management education [54-56]. Kuo et al [56] explained that the phenomenon could be due to statistical bias, including the delay in prevalence reaction time, short follow-up time, and overestimation of incidence rates from previous studies. However, our APC model presented the period effect from 1990 to 2019; therefore, the length of the follow-up period is sufficient. Besides the education level, ethnic differences, income, and occupation, the possible reason for the aforementioned phenomenon is the different levels of diagnostic certainty [26,57]. In addition, these countries have improved health care systems with regard to gout management because the government provides public data for researchers to study the disease. In Sweden, the Skåne Healthcare Register provides nationwide data for rheumatologists to study gout. Therefore, the rheumatologists could analyze the data with different dimensions, including by region and population [54,58-62]. In Taiwan, the researchers focused on the comorbidities caused by gout [63-68]. As mentioned previously, the mortality gap has stayed consistent over 2 decades. Gout complications could also affect the quality of life catastrophically. Research is necessary to understand and analyze the relationship between gout and its complications. It is also essential to explore the correlations between gout and other diseases. These studies are not only for academic use but also for practical use in the future. Although gout prevalence showed a descendant trend in Sweden, the Swedish public health professional is not satisfied regarding the prevalence and incidence of gout in the country [69]. Therefore, awareness should be created of chronic inflammatory arthritis prevalence on a worldwide level, and immediate action should be taken, including gout management education for both patients and healthy populations.

Although gout is a chronic disease, the primary management involves short-term symptomatic treatment [70,71]. Chronic diseases are often established integrated health care systems with multidisciplinary input [72]. It could prevent comorbidity effectively with the advanced policy science [73]. Primary health care for chronic diseases involves different approaches. At outpatient clinics, patients receive comprehensive advice regarding, for example, nutrition intake, individual susceptibility, body weight, and behavior changes. As gout is caused by multiple factors, the individual approach, such as nursing case management, shows significantly positive results [74]. Regarding health care system approaches, a medical record network set up among hospitals can play an important role. A patient's medical history could track gout flare frequency, which could enable evaluation of the effectiveness of the management strategy. Of note, a patient's medical history is a record of their lifestyle in the past (eg, alcohol use and body weight). Hence, the establishment of a medical record network may help to provide effective support for, and monitoring of, patients with gout [75]. It has been reported that low medication adherence among patients with gout is due to time constraints, cost, and scarcity of incentives [76]. Technology could help to overcome these barriers; for example, mobile phone apps and personal activity tracker tools could provide a novel approach to helping patients to change health behaviors [77]. Monitoring patients' health state remotely, improving patient convenience, and protecting patient privacy could enhance patient medication adherence. It may also be meaningful for health care departments to refer to the policy on, and the management of, other common chronic disease (eg, cardiovascular disease, diabetes mellitus, and asthma).

Limitations

This study includes limitations. First, limited primary data could be obtained for analysis because the data was modeled in a higher setting by the Institute for Health Metrics and Evaluation, an independent global health research center at the University of Washington, which coordinated the GBD Study 2019. Therefore, the results are heavily dependent on the GBD Study 2019 results data. It is essential to ensure the quality of country-level data collection for a health system because the public health professional can evaluate the effectiveness of the local guideline and policy with in-depth studies. Cohort studies and prospective studies are essential for studying health issues in various populations. An epidemiological study could be combined with pharmacology, mental health, etiology, and genetic studies. Second, the APC model was the mathematical approach used. However, it does not fully reflect the realistic situation. Therefore, the results should be interpreted with caution.

Conclusions

Our study provides important insights into the global burden of gout and highlights the need for effective management and prophylaxis of this condition. The APC model used in our analysis provides a novel approach to understanding the complex trends in gout prevalence and incidence. Gout prevalence in the global population has increased significantly over the past 3 decades, particularly in high-SDI regions. Our study also identified an increasing trend in gout incidence, which is predicted to continue in the coming years. These findings emphasize the urgency of addressing the inadequate strategies for management and prevention of gout in many countries. To cope with the increasing gout prevalence, we recommend the integration of multidisciplinary input and the establishment of medical record networks to improve gout management.

Acknowledgments

This study was supported by Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (CIFMS; No. 2022-I2M-C&T-B-031), the National High Level Hospital Clinical Research Funding (No. 2022-PUMCH-A-124), the National Natural Science Foundation in China (No. 81871740), and the City University of Hong Kong New Research Initiatives/Infrastructure Support from Central (APRC; 9610589).

Data Availability

All data were obtained via the Global Burden of Disease 2019 study.

Authors' Contributions

QH, T-NM, and BF contributed to the conception and design of the study. SL organized the database. YY performed the statistical analysis. T-NM and QH wrote the first draft of the manuscript. T-HS wrote sections of the manuscript. W-KM and BF supervised the research. W-KM and BF were the cocorresponding authors. All authors contributed to manuscript revision and read and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
High-resolution image for Figure 5.
[PDF File (Adobe PDF File), 8676 KB - publichealth_v9i1e45943_app1.pdf]

Multimedia Appendix 2
Trends in observed (dashed lines) and predicted (solid lines) gout in number of incidence cases in the (A) global scale, (B) high–sociodemographic index (SDI) quantiles, (C) high-middle–SDI quantiles, (D) middle-SDI quantiles, (E) low-middle–SDI quantiles, and (F) low-SDI quantiles. Shading indicates whether the rate remained stable (baseline reference), decreased by 1% per year (optimistic reference, lower limit), or increased by 1% per year (pessimistic reference, upper limit) based on the observed rate in 2019.
[PNG File , 536 KB - publichealth_v9i1e45943_app2.png]

References

1. Clebak KT, Morrison A, Croad JR. Gout: rapid evidence review. Am Fam Physician 2020 Nov 01;102(9):533-538 [FREE Full text] [Medline: 33118789]
2. Chandratre P, Roddy E, Clarson L, Richardson J, Hider S, Mallen CD. Health-related quality of life in gout: a systematic review. Rheumatology (Oxford) 2013 Nov;52(11):2031-2040 [FREE Full text] [doi: 10.1093/rheumatology/ket265] [Medline: 23934311]
3. Mody GM. Rheumatology in Africa-challenges and opportunities. Arthritis Res Ther 2017 Mar 07;19(1):49 [FREE Full text] [doi: 10.1186/s13075-017-1259-3] [Medline: 28270202]
4. Huang C, Wang T, Chen Y, Hu J, Ku P, Hsieh C, et al. Gout as a risk factor for acute myocardial infarction: evidence from competing risk model analysis. J Investig Med 2021 Aug 05;69(6):1161-1167. [doi: 10.1136/jim-2020-001714] [Medline: 34001632]
5. Fisher MC, Rai SK, Lu N, Zhang Y, Choi HK. The unclosing premature mortality gap in gout: a general population-based study. Ann Rheum Dis 2017 Jul 25;76(7):1289-1294. [doi: 10.1136/annrheumdis-2016-210588] [Medline: 28122760]
6. Kuo C, Grainge MJ, Mallen C, Zhang W, Doherty M. Comorbidities in patients with gout prior to and following diagnosis: case-control study. Ann Rheum Dis 2016 Jan 14;75(1):210-217 [FREE Full text] [doi: 10.1136/annrheumdis-2014-206410] [Medline: 25398375]
7. Tang Y, Zhang L, Zhu S, Pan J, Zhou S, He T, et al. Gout in China, 1990-2017: the Global Burden of Disease Study 2017. Public Health 2021 Feb;191:33-38. [doi: 10.1016/j.puhe.2020.06.029] [Medline: 33482625]
8. Smith E, Hoy D, Cross M, Merriman TR, Vos T, Buchbinder R, et al. The global burden of gout: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014 Aug 03;73(8):1470-1476. [doi: 10.1136/annrheumdis-2013-204647] [Medline: 24590182]
9. Liu K, Yao Y, Chen W, Mao Y, Ye D, Wen C. Modifiable risk factors and incidence of gout: estimation of population attributable fraction in the US. Semin Arthritis Rheum 2022 Aug;55:152040. [doi: 10.1016/j.semarthrit.2022.152040] [Medline: 35679791]
10. Zhu B, Wang Y, Zhou W, Jin S, Shen Z, Zhang H, et al. Trend dynamics of gout prevalence among the Chinese population, 1990-2019: a joinpoint and age-period-cohort analysis. Front Public Health 2022 Oct 12;10:1008598 [FREE Full text] [doi: 10.3389/fpubh.2022.1008598] [Medline: 36311630]
11. Zhou J, Zhou F, Wang W, Zhang X, Ji Y, Zhang P, et al. Epidemiological features of NAFLD from 1999 to 2018 in China. Hepatology 2020 May 29;71(5):1851-1864. [doi: 10.1002/hep.31150] [Medline: 32012320]
12. Platts-Mills TA. The allergy epidemics: 1870-2010. J Allergy Clin Immunol 2015 Jul;136(1):3-13 [FREE Full text] [doi: 10.1016/j.jaci.2015.03.048] [Medline: 26145982]
13. Danve A, Neogi T. Rising Global Burden of Gout: time to act. Arthritis Rheumatol 2020 Nov 17;72(11):1786-1788 [FREE Full text] [doi: 10.1002/art.41453] [Medline: 33150696]
14. Knowles RL, Bull C, Wren C, Dezateux C. Mortality with congenital heart defects in England and Wales, 1959-2009: exploring technological change through period and birth cohort analysis. Arch Dis Child 2012 Oct 29;97(10):861-865. [doi: 10.1136/archdischild-2012-301662] [Medline: 22753769]
15. Su Z, Zou Z, Hay SI, Liu Y, Li S, Chen H, et al. Global, regional, and national time trends in mortality for congenital heart disease, 1990-2019: an age-period-cohort analysis for the Global Burden of Disease 2019 study. EClinicalMedicine 2022 Jan;43:101249 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101249] [Medline: 35059612]
16. Chernyavskiy P, Little M, Rosenberg PS. Spatially varying age-period-cohort analysis with application to US mortality, 2002-2016. Biostatistics 2020 Oct 01;21(4):845-859 [FREE Full text] [doi: 10.1093/biostatistics/kxz009] [Medline: 31030216]
17. Okui T. An age-period-cohort analysis for prevalence of common psychiatric disorders in Japan, 1999-2017. Soc Psychiatry Psychiatr Epidemiol 2021 Apr 22;56(4):639-648 [FREE Full text] [doi: 10.1007/s00127-020-01988-5] [Medline: 33221927]
18. Kim HS, Eun SJ. Age-period-cohort analysis of trends in infectious disease mortality in South Korea from 1983 to 2017. Int J Environ Res Public Health 2021 Jan 21;18(3):906 [FREE Full text] [doi: 10.3390/ijerph18030906] [Medline: 33494300]

19. Kim N, Kawachi I. Age period cohort analysis of chewing ability in Korea from 2007 to 2018. Sci Rep 2021 Jul 19;11(1):14660 [FREE Full text] [doi: 10.1038/s41598-021-94086-8] [Medline: 34282165]
20. Ma Y, Cui Y, Hu Q, Mubarik S, Yang D, Jiang Y, et al. Long-term changes of HIV/AIDS incidence rate in China and the U.S. Population from 1994 to 2019: a join-point and age-period-cohort analysis. Front Public Health 2021;9:652868 [FREE Full text] [doi: 10.3389/fpubh.2021.652868] [Medline: 34869132]
21. Jeong YJ, Park S, Yon DK, Lee SW, Tizaoui K, Koyanagi A, et al. Global burden of gout in 1990-2019: a systematic analysis of the Global Burden of Disease study 2019. Eur J Clin Invest 2023 Apr 07;53(4):e13937. [doi: 10.1111/eci.13937] [Medline: 36511834]
22. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
23. GBD 2019 DiseasesInjuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
24. Preedy VR, Watson RR, editors. Years lived with disability. In: Handbook of Disease Burdens and Quality of Life Measures. New York City, NY: Springer; 2010.
25. Simões TC, Borges LF, Parreira de Assis AC, Silva MV, Dos Santos J, Meira KC. Chagas disease mortality in Brazil: a Bayesian analysis of age-period-cohort effects and forecasts for two decades. PLoS Negl Trop Dis 2018 Sep 28;12(9):e0006798 [FREE Full text] [doi: 10.1371/journal.pntd.0006798] [Medline: 30265661]
26. Dehlin M, Jacobsson L, Roddy E. Global epidemiology of gout: prevalence, incidence, treatment patterns and risk factors. Nat Rev Rheumatol 2020 Jul 15;16(7):380-390. [doi: 10.1038/s41584-020-0441-1] [Medline: 32541923]
27. Rees F, Jenkins W, Doherty M. Patients with gout adhere to curative treatment if informed appropriately: proof-of-concept observational study. Ann Rheum Dis 2013 Jun 07;72(6):826-830. [doi: 10.1136/annrheumdis-2012-201676] [Medline: 22679303]
28. Goldfien RD, Ng MS, Yip G, Hwe A, Jacobson A, Pressman A, et al. Effectiveness of a pharmacist-based gout care management programme in a large integrated health plan: results from a pilot study. BMJ Open 2014 Jan 10;4(1):e003627 [FREE Full text] [doi: 10.1136/bmjopen-2013-003627] [Medline: 24413343]
29. Watanabe S. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. J Mach Learn Res 2010;10:3571-3594.
30. Murdoch R, Barry MJ, Choi HK, Hernandez D, Johnsen B, Labrador M, et al. Gout, Hyperuricaemia and Crystal-Associated Disease Network (G-CAN) common language definition of gout. RMD Open 2021 Apr 26;7(2):e001623 [FREE Full text] [doi: 10.1136/rmdopen-2021-001623] [Medline: 33903281]
31. van Onna M, Hinsenveld E, de Vries H, Boonen A. Health literacy in patients dealing with gout: a qualitative study. Clin Rheumatol 2015 Sep 12;34(9):1599-1603. [doi: 10.1007/s10067-014-2838-1] [Medline: 25501463]
32. Fields TR, Batterman A. How can we improve disease education in people with gout? Curr Rheumatol Rep 2018 Mar 08;20(3):12 [FREE Full text] [doi: 10.1007/s11926-018-0720-x] [Medline: 29516200]
33. Li Q, Zou Y, Lian S, Liang J, Bi Y, Deng C, et al. Sugar-sweeten beverage consumption is associated with more obesity and higher serum uric acid in Chinese male gout patients with early onset. Front Nutr 2022 Jul 12;9:916811 [FREE Full text] [doi: 10.3389/fnut.2022.916811] [Medline: 35903455]
34. Roth G, Mensah G, Johnson C, Addolorato G, Ammirati E, Baddour L, GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
35. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2095-2128. [doi: 10.1016/S0140-6736(12)61728-0] [Medline: 23245604]
36. Zhang W, Doherty M, Bardin T, Pascual E, Barskova V, Conaghan P, EULAR Standing Committee for International Clinical Studies Including Therapeutics. EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2006 Oct 28;65(10):1312-1324 [FREE Full text] [doi: 10.1136/ard.2006.055269] [Medline: 16707532]
37. Jordan K, Cameron J, Snaith M, Zhang W, Doherty M, Seckl J, British Society for RheumatologyBritish Health Professionals in Rheumatology Standards, GuidelinesAudit Working Group (SGAWG). British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of gout. Rheumatology (Oxford) 2007 Aug;46(8):1372-1374. [doi: 10.1093/rheumatology/kem056a] [Medline: 17522099]
38. Chen-Xu M, Yokose C, Rai SK, Pillinger MH, Choi HK. Contemporary prevalence of gout and hyperuricemia in the United States and decadal trends: the national health and nutrition examination survey, 2007-2016. Arthritis Rheumatol 2019 Jun 15;71(6):991-999 [FREE Full text] [doi: 10.1002/art.40807] [Medline: 30618180]
39. Keller SF, Mandell BF. Management and cure of gouty arthritis. Med Clin North Am 2021 Mar;105(2):297-310. [doi: 10.1016/j.mcna.2020.09.013] [Medline: 33589104]

40. Fryar CD, Ostchega Y, Hales CM, Zhang G, Kruszon-Moran D. Hypertension prevalence and control among adults: United States, 2015-2016. NCHS Data Brief 2017 Oct(289):1-8 [FREE Full text] [Medline: 29155682]
41. Murphy D, McCulloch CE, Lin F, Banerjee T, Bragg-Gresham JL, Eberhardt MS, et al. Trends in prevalence of chronic kidney disease in the United States. Ann Intern Med 2016 Aug 02;165(7):473. [doi: 10.7326/m16-0273]
42. Hales CM, Fryar CD, Carroll MD, Freedman DS, Ogden CL. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007-2008 to 2015-2016. JAMA 2018 Apr 24;319(16):1723-1725 [FREE Full text] [doi: 10.1001/jama.2018.3060] [Medline: 29570750]
43. Kaneko K, Aoyagi Y, Fukuuchi T, Inazawa K, Yamaoka N. Total purine and purine base content of common foodstuffs for facilitating nutritional therapy for gout and hyperuricemia. Biol Pharm Bull 2014;37(5):709-721 [FREE Full text] [doi: 10.1248/bpb.b13-00967] [Medline: 24553148]
44. Kiltz U, Smolen J, Bardin T, Cohen Solal A, Dalbeth N, Doherty M, et al. Treat-to-target (T2T) recommendations for gout. Ann Rheum Dis 2017 Apr 22;76(4):632-638. [doi: 10.1136/annrheumdis-2016-209467] [Medline: 27658678]
45. FitzGerald JD, Dalbeth N, Mikuls T, Brignardello-Petersen R, Guyatt G, Abeles AM, et al. 2020 American College of rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020 Jun 11;72(6):744-760 [FREE Full text] [doi: 10.1002/acr.24180] [Medline: 32391934]
46. Yamanaka H. [Revised version of Guideline for the Management of Hyperuricemia and Gout]. Nihon Rinsho 2008 Apr;66(4):643-646. [Medline: 18409508]
47. Wang L, Wang Z, Ma Q, Fang G, Yang J. The development and reform of public health in China from 1949 to 2019. Global Health 2019 Jul 02;15(1):45 [FREE Full text] [doi: 10.1186/s12992-019-0486-6] [Medline: 31266514]
48. Nan H, Qiao Q, Dong Y, Gao W, Tang B, Qian R, et al. The prevalence of hyperuricemia in a population of the coastal city of Qingdao, China. J Rheumatol 2006 Jul;33(7):1346-1350. [Medline: 16821269]
49. Miao Z, Li C, Chen Y, Zhao S, Wang Y, Wang Z, et al. Dietary and lifestyle changes associated with high prevalence of hyperuricemia and gout in the Shandong coastal cities of Eastern China. J Rheumatol 2008 Sep;35(9):1859-1864. [Medline: 18634142]
50. Li R, Sun J, Ren L, Wang H, Liu W, Zhang X, et al. Epidemiology of eight common rheumatic diseases in China: a large-scale cross-sectional survey in Beijing. Rheumatology (Oxford) 2012 Apr 16;51(4):721-729. [doi: 10.1093/rheumatology/ker370] [Medline: 22179737]
51. Zhai F, Wang H, Du S, He Y, Wang Z, Ge K, et al. Prospective study on nutrition transition in China. Nutr Rev 2009 May;67 Suppl 1:S56-S61. [doi: 10.1111/j.1753-4887.2009.00160.x] [Medline: 19453679]
52. Ezzati M, Pearson-Stuttard J, Bennett JE, Mathers CD. Acting on non-communicable diseases in low- and middle-income tropical countries. Nature 2018 Jul 25;559(7715):507-516 [FREE Full text] [doi: 10.1038/s41586-018-0306-9] [Medline: 30046068]
53. World Health Organization. Regional Office for the Western Pacific. People's Republic of China Health System Review. Manila: WHO Regional Office for the Western Pacific; 2015.
54. Dehlin M, Drivelegka P, Sigurdardottir V, Svärd A, Jacobsson LT. Incidence and prevalence of gout in Western Sweden. Arthritis Res Ther 2016 Jul 13;18(1):164 [FREE Full text] [doi: 10.1186/s13075-016-1062-6] [Medline: 27412614]
55. Kapetanovic MC, Hameed M, Turkiewicz A, Neogi T, Saxne T, Jacobsson L, et al. Prevalence and incidence of gout in southern Sweden from the socioeconomic perspective. RMD Open 2016 Nov 01;2(2):e000326 [FREE Full text] [doi: 10.1136/rmdopen-2016-000326] [Medline: 27933209]
56. Kuo C, Grainge MJ, See L, Yu K, Luo S, Zhang W, et al. Epidemiology and management of gout in Taiwan: a nationwide population study. Arthritis Res Ther 2015 Jan 23;17(1):13 [FREE Full text] [doi: 10.1186/s13075-015-0522-8] [Medline: 25612613]
57. Gardner MJ, Power C, Barker DJ, Padday R. The prevalence of gout in three English towns. Int J Epidemiol 1982 Mar;11(1):71-75. [doi: 10.1093/ije/11.1.71] [Medline: 7085182]
58. Dehlin M, Jacobsson L. Trends in gout hospitalization in Sweden. J Rheumatol 2018 Jan 01;45(1):145-146. [doi: 10.3899/jrheum.170381] [Medline: 29292318]
59. Wändell P, Carlsson AC, Li X, Gasevic D, Ärnlöv J, Holzmann MJ, et al. Gout in immigrant groups: a cohort study in Sweden. Clin Rheumatol 2017 May 13;36(5):1091-1102 [FREE Full text] [doi: 10.1007/s10067-016-3525-1] [Medline: 28091806]
60. Dehlin M, Sandström TZ, Jacobsson LT. Incident gout: risk of death and cause-specific mortality in Western Sweden: a prospective, controlled inception cohort study. Front Med (Lausanne) 2022 Feb 24;9:802856 [FREE Full text] [doi: 10.3389/fmed.2022.802856] [Medline: 35280894]
61. Wändell P, Carlsson AC, Ljunggren G. Gout and its comorbidities in the total population of Stockholm. Prev Med 2015 Dec;81:387-391. [doi: 10.1016/j.ypmed.2015.10.003] [Medline: 26500085]
62. Dehlin M, Scheepers L, Landgren A, Josefsson L, Svensson K, Jacobsson L. Lifestyle factors and comorbidities in gout patients compared to the general population in Western Sweden: results from a questionnaire study. Scand J Rheumatol 2022 Sep 10;51(5):390-393. [doi: 10.1080/03009742.2022.2035952] [Medline: 35266438]

63. Chuang J, Lee J, Leu T, Hidajah AC, Chang Y, Li C. Association of gout and colorectal cancer in Taiwan: a nationwide population-based cohort study. BMJ Open 2019 Oct 09;9(10):e028892 [FREE Full text] [doi: 10.1136/bmjopen-2019-028892] [Medline: 31601586]
64. Chen H, Hsu Y, Chen C, Wang P, Lin C, Cheng S, et al. Gout and the risk of epilepsy: a population-based cohort study. Medicine (Baltimore) 2020 Jun 26;99(26):e20823 [FREE Full text] [doi: 10.1097/MD.0000000000020823] [Medline: 32590769]
65. Perng W, Hung Y, Lai W, Chu K, Wang Y, Wei JC. Increased risk of glaucoma amongst new-onset gout patients aged 20-39 years: a nationwide population-based cohort study in Taiwan. Int J Clin Pract 2021 Jul 02;75(7):e14169. [doi: 10.1111/ijcp.14169] [Medline: 33788372]
66. Chung M, Hung P, Hsiao P, Wu L, Chang C, Wu M, et al. Association of sodium-glucose transport protein 2 inhibitor use for type 2 diabetes and incidence of gout in Taiwan. JAMA Netw Open 2021 Nov 01;4(11):e2135353 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.35353] [Medline: 34797368]
67. Hu L, Yang AC, Lee S, You Z, Tsai S, Hu C, et al. Risk of Parkinson's disease following gout: a population-based retrospective cohort study in Taiwan. BMC Neurol 2020 Sep 08;20(1):338 [FREE Full text] [doi: 10.1186/s12883-020-01916-9] [Medline: 32900384]
68. Huang K, Tai C, Tsai Y, Kuan Y, Lee CY. Correlation between gout and coronary heart disease in Taiwan: a nationwide population-based cohort study. Acta Cardiol Sin 2019 Nov;35(6):634-640 [FREE Full text] [doi: 10.6515/ACS.201911_35(6).20190403B] [Medline: 31879516]
69. Sigurdardottir V, Svärd A, Jacobsson L, Dehlin M. Gout in Dalarna, Sweden - a population-based study of gout occurrence and compliance to treatment guidelines. Scand J Rheumatol 2022 Oct 27:1-8. [doi: 10.1080/03009742.2022.2132055] [Medline: 36300710]
70. Roddy E, Clarkson K, Blagojevic-Bucknall M, Mehta R, Oppong R, Avery A, et al. Open-label randomised pragmatic trial (CONTACT) comparing naproxen and low-dose colchicine for the treatment of gout flares in primary care. Ann Rheum Dis 2020 Feb 30;79(2):276-284 [FREE Full text] [doi: 10.1136/annrheumdis-2019-216154] [Medline: 31666237]
71. Rainer TH, Cheng CH, Janssens HJ, Man CY, Tam LS, Choi YF, et al. Oral prednisolone in the treatment of acute gout. Ann Intern Med 2016 Feb 23;164(7):464. [doi: 10.7326/m14-2070]
72. Coffey L, Mahon C, Gallagher P. Perceptions and experiences of diabetic foot ulceration and foot care in people with diabetes: a qualitative meta-synthesis. Int Wound J 2019 Feb 04;16(1):183-210 [FREE Full text] [doi: 10.1111/iwj.13010] [Medline: 30393976]
73. Mozaffarian D. Dietary and policy priorities for cardiovascular disease, diabetes, and obesity. Circulation 2016 Jan 12;133(2):187-225. [doi: 10.1161/circulationaha.115.018585]
74. da Silva M, de Fátima Mantovani M, Castanho Moreira R, Perez Arthur J, Molina de Souza R. Nursing case management for people with hypertension in primary health care: a randomized controlled trial. Res Nurs Health 2020 Jan 11;43(1):68-78. [doi: 10.1002/nur.21994] [Medline: 31710134]
75. Bodenheimer T. Helping patients improve their health-related behaviors: what system changes do we need? Dis Manag 2005 Oct;8(5):319-330. [doi: 10.1089/dis.2005.8.319] [Medline: 16212517]
76. Rai S, Choi H, Choi S, Townsend A, Shojania K, De Vera MA. Key barriers to gout care: a systematic review and thematic synthesis of qualitative studies. Rheumatology (Oxford) 2018 Jul 01;57(7):1282-1292 [FREE Full text] [doi: 10.1093/rheumatology/kex530] [Medline: 29672772]
77. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. J Telemed Telecare 2016 Oct 18;24(1):22-30. [doi: 10.1177/1357633x16673538]

Abbreviations

APC: age-period-cohort
BAPC: Bayesian age-period-cohort
GBD: Global Burden of Disease
ICD-10: International Classification of Diseases, Tenth Revision
ICD-9: International Classification of Diseases, Ninth Revision
INLA: integrated nested Laplace approximations
SDI: sociodemographic index
UI: uncertainty interval
YLD: years lived with disability

Edited by A Mavragani, T Sanchez; submitted 23.01.23; peer-reviewed by C Cao, M Shapiro; comments to author 06.03.23; revised version received 26.03.23; accepted 20.04.23; published 07.06.23.

Please cite as:
He Q, Mok TN, Sin TH, Yin J, Li S, Yin Y, Ming WK, Feng B
Global, Regional, and National Prevalence of Gout From 1990 to 2019: Age-Period-Cohort Analysis With Future Burden Prediction
JMIR Public Health Surveill 2023;9:e45943
URL: https://publichealth.jmir.org/2023/1/e45943
doi:10.2196/45943
PMID:37285198

Original Paper

Web-Based Harm Reduction Intervention for Chemsex in Men Who Have Sex With Men: Randomized Controlled Trial

Edmond Pui Hang Choi[1], PhD; Kitty Wai Ying Choi[1,2], MSocSc; Chanchan Wu[1], MSc; Pui Hing Chau[1], PhD; Jojo Yan Yan Kwok[1], PhD; William Chi Wai Wong[3], MD; Eric Pui Fung Chow[4,5,6], PhD

[1]School of Nursing, The University of Hong Kong, Hong Kong, China (Hong Kong)
[2]Sticky Rice Love, Hong Kong, China (Hong Kong)
[3]Department of Family Medicine and Primary Care, School of Clinical Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)
[4]Melbourne Sexual Health Centre, Alfred Health, Melbourne, Australia
[5]Central Clinical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia
[6]Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Edmond Pui Hang Choi, PhD
School of Nursing
The University of Hong Kong
5/F, Academic Building, 3 Sassoon Road
Pok Fu Lam
Hong Kong,
China (Hong Kong)
Phone: 852 3917 6972
Email: h0714919@connect.hku.hk

Abstract

Background: Men who have sex with men (MSM) who practice chemsex have a higher likelihood of engaging in risky sexual behaviors and higher rates of HIV infection and other sexually transmitted infections (STIs) than those who do not.

Objective: This trial aimed to evaluate the effectiveness of a web-based intervention in reducing the sexual harms of chemsex among MSM.

Methods: The study was a 2-arm, assessor-blinded, randomized, parallel-group trial with a 3-month follow-up period. The study was conducted in the year 2021 in Hong Kong. Underpinned by the theory of planned behaviors and a harm reduction approach, the intervention consisted of interactive components and knowledge-based information about chemsex. Participants in the control group received brief information and content about sexual violence. The primary outcome was self-efficacy in refusing risky sexual behaviors and chemsex, as measured by the Condom Self-Efficacy Scale (CSES), Self-Efficacy for Sexual Safety (SESS) instrument, and Drug Avoidance Self-Efficacy Scale (DASES). The secondary outcomes included intentions to have chemsex, actual engagement in chemsex, HIV and other STI testing, and condom use in the last 3 months. All outcomes were self-reported. An online structured questionnaire was used to collect data.

Results: In total, 316 MSM enrolled in the study. The intervention group demonstrated a significantly larger improvement in condom-use self-efficacy (as measured by CSES scores; time-by-group interaction: β=4.52, 95% CI 2.03-7.02; *P*<.001), self-efficacy for sexual safety (as measured by SESS scores; time-by-group interaction: β=2.11, 95% CI 0.66-3.56; *P*=.004), and drug avoidance self-efficacy (as measured by DASES scores; time-by-group interaction: β=6.98, 95% CI 1.75-12.22; *P*=.009). Regarding the secondary outcomes, participants in the intervention group demonstrated a significantly larger reduction in the likelihood of having engaged in chemsex in the last 3 months (time-by-group interaction: odds ratio [OR]=0.23, 95% CI 0.10-0.53; *P*=.001) and likelihood of having had the intention to engage in chemsex in the last 3 months (time-by-group interaction: OR=0.37, 95% CI 0.18-0.78; *P*=.009). Participants in the intervention group also showed a significantly larger increase in the likelihood of having undergone HIV testing in the last 3 months (time-by-group interaction: OR=3.08, 95% CI 1.72-5.54; *P*<.001).

Conclusions: This study suggests that a web-based intervention with a harm reduction approach can enhance the self-efficacy of MSM in refusing risky sexual behaviors and chemsex and improve the uptake of HIV testing. We also provide initial evidence that such interventions can reduce both the intention of MSM to engage in chemsex and their actual engagement in chemsex.

Trial Registration: ISRCTN Registry ISRCTN20134522; https://www.isrctn.com/ISRCTN20134522.

International Registered Report Identifier (IRRID): RR2-10.1186/s12889-021-10742-8

(JMIR Public Health Surveill 2023;9:e42902) doi:10.2196/42902

KEYWORDS

chemsex; drug abuse; eHealth; men who have sex with men; trial; prevention; Chinese

Introduction

"Chemsex" is defined as the use of psychoactive substances before or during planned sex to facilitate, initiate, prolong, sustain, or intensify the sexual encounter [1,2]. There is no universal and systematic definition of types of chemsex drugs because the drugs used might vary based on place and time [3]. Nonetheless, there are 4 substances typically associated with chemsex: methamphetamine, mephedrone, γ-butyrolactone/γ-hydroxybutyric acid, and ketamine [4]. Chemsex occurs across sexual orientations and genders but is considerably more common among men who have sex with men (MSM) [5-7]. A systematic review reported that the prevalence of recent (within the last 6 months) engagement in chemsex among MSM populations is high, with estimates ranging from 9.9% to 93.7% (according to 12 studies) [8]. Motivations for engagement in chemsex vary. A systematic review reported that MSM engage in chemsex because it can increase their stamina and arousal levels, allowing them to engage in sex for sustained periods; the lowering of inhibitions induced by chemsex can also provide a more immediate and sustained interaction with sex partners [2].

Empirical studies have reported that the practice of risky sexual behaviors is prevalent among MSM who engage in chemsex. A systematic review reported that the prevalence of condomless sex ranged from 17% to 100% among MSM who engaged in chemsex [8]. A study in the Netherlands also reported that the participating MSM who engaged in chemsex were significantly more likely to practice condomless anal sex than the MSM who did not (84.3% vs 61.1%, respectively; $P<.001$) [9]. According to a qualitative study conducted in the United Kingdom, MSM perceive condomless sex as the norm during chemsex [10]. Chemsex is also associated with group sex, fisting, and a higher number of sexual partners [8,11,12]. Robust evidence suggests that chemsex is strongly associated with high-risk sexual behaviors linked to the risk of acquisition of HIV and other sexually transmitted infections (STIs).

MSM who engage in chemsex experience higher rates of HIV infection, other STIs, and hepatitis C infection than those who do not [4]. A cohort study in Canada reported that practicing chemsex is linked to an increased incidence of gonorrhea and chlamydia and that this effect is stronger for people who use multiple chemsex substances [13]. A study in the United Kingdom reported that the rate of new HIV diagnoses is significantly higher in MSM who practice chemsex than in those who do not (8.6% vs 1.8%, respectively) [12]. The study also reported that MSM who practice chemsex have higher odds of having a serodiscordant HIV-positive sex partner (adjusted odds ratio [OR]=6.83) [12]. The high frequency of STIs among chemsex users highlights the importance of primary, secondary, and tertiary prevention of STIs and chemsex among MSM.

Globally, "undetectable equals untransmittable (U=U)" and pre-exposure prophylaxis (PrEP) have become important elements in HIV prevention programs [14]. However, the uptake of U=U and PrEP has been limited and slow in the Asia-Pacific region [14]. It was reported that U=U has not yet been widely applied in clinical settings in the Asia-Pacific region to empower people living with HIV to use antiretroviral treatments to achieve and maintain their untransmittable status and to live normal sexual and social lives [14]. However, there has been no systematic study evaluating the use of U=U in Hong Kong. According to the most recent review of the HIV/AIDS situation in Hong Kong, sexual transmission remained the major mode of HIV transmission [15].

In addition, even though PrEP is a very effective HIV prevention method [16], unequal access to PrEP services continues to negatively affect many people who could benefit from it worldwide. It was estimated that just under 1 million people had initiated oral PrEP by 2020, which is far less than the 2020 Joint United Nations Programme on HIV/AIDS target of 3 million [17]. Furthermore, access to PrEP is still highly concentrated in a small number of countries. By the end of 2020, fewer than 20 countries recorded more than 10,000 PrEP initiations [17]. In Asia, which was the source of nearly 20% of the world's HIV infections in 2019, only a few countries (eg, Thailand and Vietnam) have implemented national guidelines outlining PrEP as an HIV prevention strategy [17,18]. Moreover, another study estimated that Southeast Asia contributes less than 5% of all the PrEP initiations recorded worldwide [19]. Despite increasing scientific evidence on its effectiveness and safety, international recommendations, and the rising global adoption of PrEP in HIV prevention, PrEP is currently not available as a part of public health care services in Hong Kong. People in Hong Kong can only obtain PrEP in the private sector at a very high cost (about HK $8000/month [US $1032]) or purchase it from other countries [20]. In areas where the availability of PrEP is still limited, behavioral interventions to promote safer sexual practices, such as consistent condom use and regular HIV testing, still play an important role in HIV prevention.

eHealth, health services and information delivered electronically through the internet (eg, through a website) [21], has become a commonly used modality for health promotion. A systematic review of the process evaluation of 8 eHealth interventions reported that eHealth interventions targeting sexual risk and substance use are acceptable for MSM across different sociodemographic groups [22]. Another systematic review and meta-analysis of 46 studies supported the effectiveness of eHealth privacy interventions in promoting HIV-preventive behaviors, such as condom use and STI testing, among MSM [23]. Thanks to the well-established internet access that most people have worldwide, one of the most important advantages of eHealth is the high degree of accessibility it offers by granting access to remotely located people [24]. Moreover, compared with face-to-face interventions, eHealth interventions provide greater anonymity, privacy, and accessibility, which are particularly relevant and important for promoting sexual health among MSM.

Although several systematic reviews have provided robust evidence about the effectiveness of eHealth interventions in improving sexual health among MSM [23,25,26], there has been no study related to chemsex and its prevention. To address this knowledge gap, this study aimed to evaluate the effectiveness of a web-based intervention in reducing the sexual harms of chemsex among MSM, with the objectives of strengthening the self-efficacy of MSM in refusing risky sexual behaviors and chemsex, reducing both their intention to engage in chemsex and their actual chemsex behaviors, enhancing consistent condom use during both sober and drug-influenced sex, and increasing motivation toward the practice of regular HIV and STI testing.

We hypothesized that the web-based intervention could increase participants' self-efficacy in refusing risky sexual behaviors and chemsex, reduce both the intention to engage in chemsex and actual chemsex behaviors, enhance consistent condom use, and increase the uptake of HIV and STI testing.

Methods

Study Design

This study was a 2-arm, assessor-blinded, randomized, parallel-group trial with a 3-month follow-up period. The trial protocol has been published [27] and registered. There were no important changes to the methods and study content after trial commencement.

Study Participants and Setting

The study participants were enrolled using convenience sampling from June 15, 2021, to November 5, 2021. To ensure that we could recruit the required number of study participants from diverse backgrounds, potential participants were approached using social media (eg, Instagram). Local nongovernmental organizations that target MSM populations in Hong Kong also helped recruit potential participants based on their existing networks. Participants were eligible for inclusion in the trial if they self-identified as cis-male MSM aged ≥18 years with internet access and the ability to read and understand Chinese.

Screening, Baseline Assessment, and Randomization

An online screening questionnaire was used to screen the study participants for eligibility. Eligible participants were asked to sign an electronic consent form and use their email addresses to register for the study. They also needed to set a personal password to gain access to the study intervention. After completing these steps, the participants were asked to complete an online baseline questionnaire. After completing the baseline questionnaire, participants were randomly assigned to either the intervention group or the control group via a computer-generated block randomization procedure (with a block size of 4 with no stratification) with a 1:1 randomization ratio. The computer-generated sequence was created by the independent programmer who developed the online platform. The online platform conducted masking and allocation concealment. All participants and research team members were blinded to the allocation sequence before the allocation. After randomization, the study participants were automatically guided to the web content associated with their allocation.

Interventions

The intervention was divided into 2 parts. The first part involved an interactive component. Enrolled participants were first invited to complete 2 quizzes to review their level of understanding of chemsex. The quizzes assessed how much the participants knew about chemsex. Each quiz contained 10 multiple-choice questions. The 2 sets of quizzes differed in terms of the level of difficulty. Participants received their score after completing the quiz. In the second part, participants were given knowledge-based information about chemsex and its potential risks and legal consequences. The side effects of different chemsex substances were also covered. Additionally, the participants were presented with information about how they could protect themselves from contracting HIV and other STIs and regarding local resources for emotional support and HIV and STI testing.

The contents of the intervention were developed based on the theory of planned behaviors [28,29]. For example, the interventions aimed to (1) lower participants' desire to engage in chemsex by enhancing their knowledge of chemsex (eg, regarding its side effects, risks, and legal consequences) and (2) improve the consistency of participants' condom use by enhancing their attitude toward and knowledge about condom use. Also, the contents of the intervention aimed to intervene from the perspective of harm reduction, primarily with regard to sexual harm. Because it is difficult for current chemsex users to withdraw from drug use over a short period, being empowered to carry out safer sex practices, even in drug-influenced situations, may be a more realistic way to lower the risks associated with HIV and STI transmission. For example, checklists of what to pay attention to before, during, and after chemsex were provided to foster responsible attitudes toward and practices of chemsex. The safety precautions an individual can take to reduce harm related to chemsex practice were also emphasized. Participants in the intervention group were expected to spend only 30 minutes to 45 minutes to complete the overall intervention.

The control group was offered brief information and educational content about sexual violence, which was not relevant to the chemsex intervention component. There was no interactive component in the control group. The details of the intervention and control groups are shown in Multimedia Appendix 1.

Study participants had unlimited access to their allocated content during the study period. However, the content was only available to participants with a registered email address and password to minimize contamination between the intervention and control groups.

Outcomes

The primary outcome was self-efficacy in refusing risky sexual behaviors and chemsex. As there was no well-established study instrument to specifically assess this primary outcome, the traditional Chinese version of the Condom Self-Efficacy Scale (CSES) [30], Self-Efficacy for Sexual Safety (SESS) instrument [31], and Drug Avoidance Self-Efficacy Scale (DASES) [32] were used to measure the primary outcome.

The CSES is a 14-item instrument covering 3 domains: (1) consistent condom use, (2) correct condom use, and (3) condom use communication. The total score ranges from 14 to 70, with a higher score indicating a higher level of condom use efficacy [30]. The Cronbach alpha coefficient of the CSES was 0.94 in the current sample. The SESS is a 7-item instrument for assessing participants' confidence in practicing safer sex. The total score ranges from 7 to 35, with a higher score indicating a higher level of self-efficacy for safe sex [31]. The Cronbach alpha coefficient of the SESS was 0.89 in the current sample. The DASES was used to assess drug avoidance self-efficacy. It includes 16 items assessing abstinence self-efficacy across different high-risk situations. The total score ranges from 16 to 112, with a higher score indicating a higher level of self-efficacy for resisting drug use [32]. The Cronbach alpha coefficient of the DASES was 0.89 in the current sample.

The secondary outcomes included (1) an intention to have chemsex in the last 3 months, (2) actual engagement in chemsex in the last 3 months, and (3) HIV and other STI testing in the last 3 months. We assessed and evaluated participants' intentions to have chemsex in the last 3 months because we wanted to supplement the outcomes relating to their actual engagement in chemsex. It is known that people who intend to have chemsex do not necessarily eventually engage in chemsex. Therefore, their intentions to have chemsex in the past 3 months were assessed. For participants who engaged in chemsex in the last 3 months, data were collected on the (4) practice of condomless sex during non-chemsex in the last 3 months and (5) practice of condomless sex during chemsex in the last 3 months.

All study outcomes were self-assessed at baseline and 3-month follow-up interviews through an online structured questionnaire. The online questionnaire was pilot tested. A trained research assistant would send reminders (via email or instant messaging) to the study participants to ask them to complete the follow-up interview. There was no change to trial outcomes after the study commenced.

Sample Size

To detect a small-to-moderate between-group difference (Cohen d=0.4) in the primary outcome through an independent samples *t* test and to achieve a power of 80% at a .05 level of 2-sided significance, at least 200 participants (100 participants in each group) were required. Based on a systematic review of eHealth interventions about HIV and STI prevention among MSM [26], we estimated that the dropout rate would be 20%. Therefore, at least 250 participants would need to be recruited. G*Power was used to calculate the sample size.

Statistical Analysis

Descriptive statistics were used to summarize the sociodemographic characteristics and study outcomes of the participants at each time point. Baseline characteristics and study outcomes (1) between the intervention and control groups and (2) between participants who completed the study and those who did not (dropouts) were compared using Fisher exact tests for categorical variables or an independent samples *t* test for continuous variables.

The intention-to-treat principle was applied. Linear mixed effects models were used to assess the differential change in continuous outcomes (ie, scores on the CSES, SESS, and DASES). Similarly, generalized linear mixed effects models with logit links were used to analyze the binary outcomes (ie, intention to have chemsex, actual engagement in chemsex, undergone HIV testing, undergone other STI testing, practice of condomless sex during non-chemsex, and practice of condomless sex during chemsex in the last 3 months). Time, group, and interaction between group and time were included as independent variables. Mixed effects models can accommodate missing data and do not require imputation of missing observations, providing a natural way to deal with missing values or dropouts. For the main analysis, no covariates were put in the models. To supplement the main analysis, complete case analysis was conducted. Statistical analysis was performed using SPSS version 25.0 (IBM Corp, Armonk, NY). All statistical tests were 2-tailed, with a 5% level of statistical significance. The data set and analyses were checked by 2 researchers.

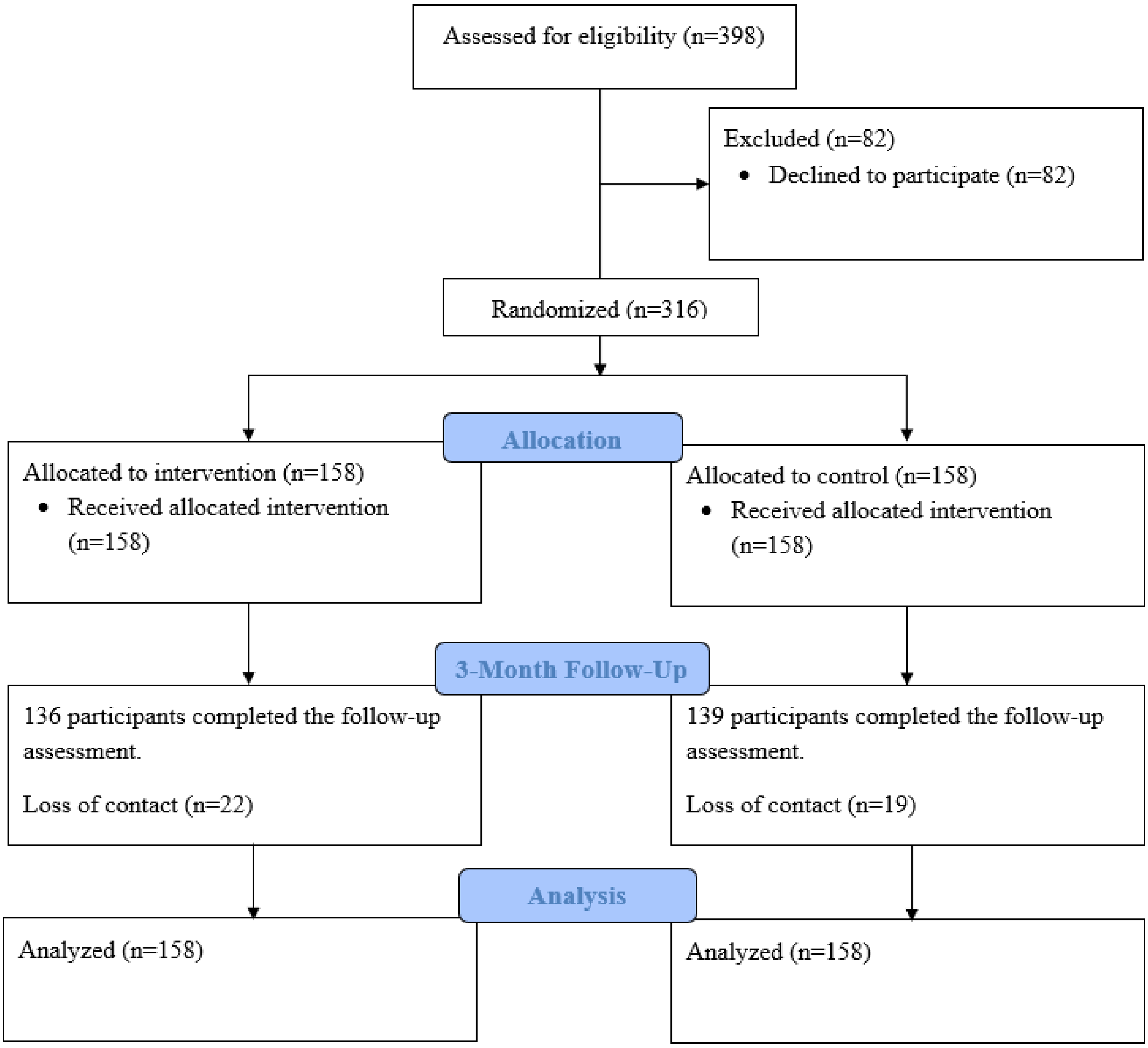
Ethics

The study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (HKU/HA HKW IRB; reference number: UW 20-650). Electronic informed consent was obtained from each study participant. Study information such as aims, nature of the study, and brief overview of the content of the intervention were provided in the consent form. This trial followed the CONSORT-EHEALTH statement (Multimedia Appendix 2) and the International Conference of Harmonisation Guidelines for Good Clinical Practice. Study data were de-identified. Only the research team had access to the study data. We established a hotline and email account for enquiries and support.

Results

In total, 316 participants were enrolled and randomized into intervention (n=158) and control (n=158) groups. However, 41 participants (22 participants in the intervention group and 19 participants in the control group) did not complete the follow-up survey. The overall dropout rate was 13.0%, (41/316), and there was no significant difference in the dropout rate between the intervention group and control group. Figure 1 shows the CONSORT flow diagram.

Figure 1. CONSORT flow diagram.



Participant Characteristics

Table 1 shows the study participants' baseline characteristics. The mean age was 27.34 (SD 6.77) years. Of the 316 participants, 265 (83.9%) were homosexual, and 51 (16.1%) were bisexual. Regarding relationship status, 172 (54.4%) of the 316 participants were in a relationship or married, and 144 (45.6%) participants were single. Of the 316 participants, 199 (63.0%) were employed full-time, and 117 (37.0%) were not employed full-time. Regarding income, 167 (52.8%) of the 316 participants had a monthly personal income less than HK $20,000, and 149 (47.2%) had a monthly personal income of at least HK $20,000.

In total, 104 (32.9%) of the 316 participants had undergone HIV testing in the last 3 months, while 64 (20.3%) of the 316 participants had undergone other STI testing in the last 3 months. Regarding chemsex, 57 (18.0%) of the 316 participants had intended to have chemsex in the last 3 months, while 84 (26.6%) of the 316 participants had engaged in chemsex during their lifetime and 51 (16.1%) of the 316 participants had engaged in chemsex in the last 3 months.

No significant heterogeneity of the demographic data and study outcomes was found between the intervention group and control group, except for relationship status and HIV testing in the last 3 months. No significant heterogeneity of the demographic data and study outcomes was found among those who completed the study (non-dropout) versus those who did not complete the study (dropouts), except for sexual orientation and chemsex in the last 3 months (Multimedia Appendix 3).

Table 1. Baseline characteristics and study outcomes of the study participants.

Characteristics and outcomes	Total (n=316)	Intervention group (n=158)	Control group (n=158)	*P* value[a]
Baseline sociodemographic characteristics				
Age (years), mean (SD)	27.34 (6.77)	27.78 (6.98)	26.90 (6.55)	.25
Sexual orientation, n (%)				.76
Homosexual	265 (83.9)	134 (84.8)	131 (82.9)	
Bisexual	51 (16.1)	24 (15.2)	27 (17.1)	
Relationship status, n (%)				.009
In a relationship/married	172 (54.4)	74 (46.8)	98 (62.0)	
Single	144 (45.6)	84 (53.2)	60 (38.0)	
Educational level, n (%)				.77
Bachelor degree or above	261 (82.6)	129 (81.6)	132 (83.5)	
Less than a bachelor degree	55 (17.4)	29 (18.4)	26 (16.5)	
Employment status, n (%)				.99
Employed full-time	199 (63.0)	99 (62.7)	100 (63.3)	
Not employed full-time	117 (37.0)	59 (37.3)	58 (36.7)	
Monthly personal income (HK $), n (%)				.65
≥20,000	149 (47.2)	77 (48.7)	72 (45.6)	
<20,000	167 (52.8)	81 (51.3)	86 (54.4)	
Chemsex (lifetime), n (%)				.90
Yes	84 (26.6)	41 (25.9)	43 (27.2)	
No	232 (73.4)	117 (74.1)	115 (72.8)	
Primary outcomes				
Self-Efficacy for Sexual Safety Scale, mean (SD)	25.42 (6.07)	25.13 (6.08)	25.70 (6.07)	.41
Condom Self-Efficacy Scale, mean (SD)				
Consistent Use subscale	11.65 (2.82)	11.54 (2.82)	11.75 (2.82)	.51
Correct Use subscale	23.87 (4.86)	23.87 (4.74)	23.87 (4.98)	.99
Communication subscale	19.15 (4.38)	19.01 (4.45)	19.28 (4.32)	.59
Total score	54.66 (11.12)	54.43 (11.04)	54.90 (11.24)	.71
Drug Avoidance Self-Efficacy Scale, mean (SD)	83.97 (20.81)	81.91 (21.27)	86.03 (20.20)	.08
Secondary outcomes				
Had chemsex in the last 3 months, n (%)				.76
Yes	51 (16.1)	27 (17.1)	24 (15.2)	
No	265 (83.9)	131 (82.9)	134 (84.8)	
Intended to have chemsex in the last 3 months, n (%)				.99
Yes	57 (18.0)	28 (17.7)	29 (18.4)	
No	259 (82.0)	130 (82.3)	129 (81.6)	
Underwent HIV testing in the last 3 months, n (%)				.006
Yes	104 (32.9)	40 (25.3)	64 (40.5)	
No	212 (67.1)	118 (74.7)	94 (59.5)	
Underwent other STI[b] testing in the last 3 months, n (%)				.21
Yes	64 (20.3)	27 (17.1)	37 (23.4)	
No	252 (79.7)	131 (82.9)	121 (76.6)	

Characteristics and outcomes	Total (n=316)	Intervention group (n=158)	Control group (n=158)	*P* value[a]
Had condomless sex during non-chemsex in the last 3 months (n=51)[c], n (%)				.49
Yes	49 (96.1)	25 (92.6)	24 (100)	
No	2 (3.9)	2 (7.4)	0 (0)	
Had condomless sex during chemsex in the last 3 months (n=51)[c], n (%)				.11
Yes	47 (92.2)	23 (85.2)	24 (100)	
No	4 (7.8)	4 (14.8)	0 (0)	
Study dropouts				
Participants who completed the follow-up interview, n (%)				.74
Yes	275 (87.0)	136 (86.1)	139 (88.0)	
No	41 (13.0)	22 (13.9)	19 (12.0)	

[a] Fisher exact test for categorical variables or independent *t* test for continuous variables.

[b] STI: sexually transmitted infection.

[c] Only participants who engaged in chemsex in the last 3 months answered the questions about condom use. At the baseline assessment, 51 participants had chemsex in the last 3 months.

Primary Outcomes

The effects of the intervention on the study outcomes are summarized in Table 2. Regarding the SESS, the intervention group demonstrated significantly larger improvement in self-efficacy for sexual safety (time-by-group interaction: β=2.11, 95% CI 0.66-3.56; *P*=.004). The intervention group also demonstrated significantly larger improvement in condom use self-efficacy (as measured by their total CSES score; time-by-group interaction: β=4.52, 95% CI 2.03-7.02; *P*<.001) and drug avoidance self-efficacy (as measured by the DASES; time-by-group interaction: β=6.98, 95% CI 1.75-12.22; *P*=.009). The results of the CSES subscale scores are shown in Multimedia Appendix 4.

Table 2. Mixed effects models for comparison of study outcomes.

Outcomes	Intervention group			Control group			Group × time effect, β[a] or OR[b,c] (95% CI)	*P* value[d,e]
	Results, mean (95% CI)[a] or n (%)[b]	Within-group change from baseline, β[a] or OR[b] (95% CI)	*P* value	Results, mean (95% CI)[a] or n (%)[b]	Within-group change from baseline, β[a] or OR[b] (95% CI)	*P* value		
Primary outcomes								
Self-Efficacy for Sexual Safety Scale[f]								
Baseline	25.13 (24.18 to 26.08)	N/A[g]	N/A	25.70 (24.75 to 26.65)	N/A	N/A	N/A	N/A
Follow-up	28.59 (27.69 to 29.49)	3.46 (2.43 to 4.48)	<.001	27.05 (26.16 to 27.94)	1.35 (0.33 to 2.37)	.01	2.11 (0.66 to 3.56)	.004
Condom Self-Efficacy Scale: total score[h]								
Baseline	54.43 (52.69 to 56.17)	N/A	N/A	54.90 (53.15 to 56.64)	N/A	N/A	N/A	N/A
Follow-up	60.48 (58.87 to 62.09)	6.05 (4.28 to 7.82)	<.001	56.43 (54.83 to 58.03)	1.53 (–0.23 to 3.29)	.09	4.52 (2.03 to 7.02)	<.001
Drug Avoidance Self-Efficacy Scale[i]								
Baseline	81.91 (78.67 to 85.16)	N/A	N/A	86.03 (82.79 to 89.28)	N/A	N/A	N/A	N/A
Follow-up	91.06 (87.94 to 94.17)	9.15 (5.43 to 12.86)	<.001	88.19 (85.11 to 91.28)	2.16 (–1.53 to 5.85)	.25	6.98 (1.75 to 12.22)	.009
Secondary outcomes								
Had chemsex in the last 3 months								
Baseline	27 (17.1)	N/A	N/A	24 (15.2)	N/A	N/A	N/A	N/A
Follow-up	7 (5.1)	0.26 (0.12 to 0.58)	.001	24 (17.3)	1.17 (0.85 to 1.59)	.34	0.23 (0.10 to 0.53)	.001
Intended to have chemsex in the last 3 months								
Baseline	28 (17.7)	N/A	N/A	29 (18.4)	N/A	N/A	N/A	N/A
Follow-up	9 (6.6)	0.33 (0.17 to 0.65)	.001	23 (16.5)	0.88 (0.65 to 1.21)	.43	0.37 (0.18 to 0.78)	.009
Underwent HIV testing in the last 3 months[j]								
Baseline	37 (24.2)	N/A	N/A	57 (38.5)	N/A	N/A	N/A	N/A
Follow-up	48 (36.6)	1.81 (1.16 to 2.84)	.01	35 (26.9)	0.59 (0.40 to 0.86)	.006	3.08 (1.72 to 5.54)	<.001
Underwent other STI[k] testing in the last 3 months								
Baseline	27 (17.1)	N/A	N/A	37 (23.4)	N/A	N/A	N/A	N/A
Follow-up	32 (23.5)	1.49 (0.91 to 2.44)	.11	33 (23.7)	1.02 (0.69 to 1.51)	.93	1.46 (0.78 to 2.76)	.23
Had condomless sex during non-chemsex in the last 3 months (n=31)[l]								
Baseline	4 (57.1)	N/A	N/A	22 (91.7)	N/A	N/A	N/A	N/A
Follow-up	5 (71.4)	1.88 (0.57 to 6.13)	.29	24 (100)	175.39 (40.00 to 769.70)	<.001	0.01 (0.002 to 0.07)	<.001
Had condomless sex during chemsex in the last 3 months (n=31)[l]								
Baseline	3 (42.9)	N/A	N/A	19 (79.2)	N/A	N/A	N/A	N/A

Outcomes	Intervention group			Control group			Group × time effect, β[a] or OR[b,c] (95% CI)	*P* value[d,e]
	Results, mean (95% CI)[a] or n (%)[b]	Within-group change from baseline, β[a] or OR[b] (95% CI)	*P* value	Results, mean (95% CI)[a] or n (%)[b]	Within-group change from baseline, β[a] or OR[b] (95% CI)	*P* value		
Follow-up	6 (85.7)	8.00 (0.92 to 69.55)	.06	22 (91.7)	2.90 (0.87 to 9.65)	.08	2.76 (0.23 to 32.83)	.41

[a] Primary outcomes.

[b] Secondary outcomes, with “no” as the reference category.

[c] OR: odds ratio.

[d] Linear mixed effects models for primary outcomes, with the control group as the reference category. Baseline characteristics were not adjusted in the models.

[e] Generalized linear mixed-effects models for secondary outcomes, with logit links that were used to analyze the binary outcomes and the control group as the reference category. Baseline characteristics were not adjusted in the models.

[f] The total score ranges from 7 to 35, with a higher score indicating a higher level of self-efficacy for safe sex.

[g] N/A: not applicable.

[h] The total score ranges from 14 to 70, with a higher score indicating a higher level of condom use efficacy.

[i] The total score ranges from 16 to 112, with a higher score indicating a higher level of self-efficacy to resist drug use.

[j] Participants who reported being HIV-positive at the baseline assessment were excluded from the analysis.

[k] STI: sexually transmitted infection.

[l] Only participants who engaged in chemsex in the last 3 months answered the questions about condom use. At the follow-up assessment, only 31 participants had chemsex in the last 3 months.

Secondary Outcomes

Participants in the intervention group demonstrated significantly larger reductions in the likelihood to have engaged in chemsex in the last 3 months (time-by-group interaction: OR=0.23, 95% CI 0.10-0.53; *P*=.001) and intention to engage in chemsex in the last 3 months (time-by-group interaction: OR=0.37, 95% CI 0.18-0.78; *P*=.009). Participants in the intervention group also showed a significantly larger increase in the likelihood to have undergone HIV testing in the last 3 months (time-by-group interaction: OR=3.08, 95% CI 1.72-5.54; *P*<.001). However, no significant group-by-time interaction effect was observed on the outcome of other STI testing in the last 3 months (*P*=.23). During the follow-up interviews, only 31 participants reported engaging in chemsex in the last 3 months. A statistically significant group-by-time interaction effect was observed for the outcome of condomless sex during non-chemsex in the last 3 months (*P*<.001). No significant group-by-time interaction effect was observed on the outcome of condomless sex during chemsex in the last 3 months (*P*=.41). The results of the between-group differences in all study outcomes at each time point are shown in Multimedia Appendix 5. Given that there was a statistically significant difference in relationship status between the intervention group and control group at the baseline assessment, an adjusted analysis was conducted (Multimedia Appendix 6). The results of the complete case analysis (n=275) are shown in Multimedia Appendix 7. The results are consistent with those of the main analysis.

Discussion

Principal Findings

In this randomized controlled trial, we found that a web-based intervention with a harm reduction approach could enhance the self-efficacy of MSM in refusing risky sexual behaviors and chemsex. The intervention enhanced HIV testing but had no effects on other STI testing. Importantly, we found that the intention of MSM to engage in chemsex and their actual engagement in chemsex can also be reduced by the intervention. However, it should be noted that, given the small sample size of participants who reported engaging in chemsex in the last 3 months during the follow-up interviews, the findings related to condom use might not be reliable and should be interpreted with caution. Nonetheless, the clinical significance and academic merits of the findings are substantial because this study has provided one of the very first pieces of evidence on the effectiveness of eHealth interventions in reducing the sexual harm of chemsex among MSM, which echoes the recent review by Strong and colleagues [33] that called for more action and identified a need for more harm reduction interventions related to chemsex. In addition, one important feature of the trial was that the contents of the intervention were relatively brief. Participants in the intervention group were expected to spend only 30 minutes to 45 minutes to complete the overall intervention. Such digitally delivered brief interventions can be easily disseminated to hard-to-reach populations such as MSM who practice chemsex, in a cost-effective manner.

The intervention in the current trial was developed through a harm reduction approach because most qualitative evidence suggests that many MSM prefer reducing the harm associated with drug use (eg, by learning how to better manage their use) instead of abstaining altogether [33,34]. Furthermore, while highlighting what kinds of tailored interventions can assist in the reduction of harm, Herrijgers and colleagues [35] showed in their schematic overview that MSM were exposed to different degrees of harm before, during, and after chemsex events. To meet the actual needs of MSM, the intervention contents of the current trial were developed in a comprehensive manner, addressing the different aspects of harm related to chemsex [33,35]. We also added interactive content (quizzes and instant feedback) to the intervention because a meta-analysis of 12 studies reported that a new media intervention using interactive components yielded significant effects on condom use, while interventions using only static content did not yield positive effects on enhancing condom use [36].

Compared with the control group, the intervention group demonstrated statistically significant effects on the primary outcomes with a small-to-moderate effect size (in Multimedia Appendix 7) and on most secondary outcomes (ie, HIV testing, intention to have chemsex, and actual engagement in chemsex). Our findings are in line with those reported by a systematic review and meta-analysis of 46 studies, which supported the effectiveness of eHealth interventions in promoting HIV-preventive behaviors among MSM with a small effect size [23]. Coupled with previous studies [23], our study further supports the benefits of eHealth interventions to reduce risky sexual behaviors among MSM, which ultimately decreases the burdens of HIV and other STIs in the MSM population.

Even though PrEP has provided a new means of harm reduction (ie, primary HIV prevention) in the context of chemsex [33], the uptake of PrEP has remained low in some regions, such as Southeast Asia. A study reported that only 1% had ever used PrEP among MSM in Hong Kong [20]. In geographic areas where the availability of PrEP and access to it are still limited, eHealth interventions to promote safer sexual practice and regular HIV testing still play an important role in HIV prevention. Therefore, this trial still has significant public health implications for helping to control the burgeoning HIV epidemic among MSM.

Another important benefit of the current study is that the intervention was found to reduce the participants' intention to have chemsex and their actual engagement in chemsex. Given that chemsex tends to be hidden and private and that those who practice chemsex are difficult to reach [37], eHealth interventions can provide MSM with a private, anonymous, and easily accessible platform offering supportive care and useful information. The current study provides empirical evidence for governments, health authorities, and other stakeholders to put more resources into delivering chemsex-preventive interventions and care using online platforms. Also, more implementation research is needed to plan how to position and scale up eHealth interventions regarding chemsex alongside other community-based HIV-preventive interventions.

This study found no significant effect on other STI testing, implying that future interventions should strengthen the components intended to encourage MSM to undergo regular STI testing. Nonetheless, the negative result could be explained by the substantial reduction in the availability of sexual health services during the COVID-19 pandemic [38]. A study in Australia reported that chlamydia and gonorrhea testing dropped significantly during the COVID-19 pandemic [39]. Unlike HIV self-testing, which is available through the government, community-based organizations, and registered pharmaceutical stores [40], MSM in Hong Kong might have experienced difficulty undergoing other forms of STI testing during the pandemic. Also, the attitudes and perceptions toward the consequences of other STIs among MSM could have led to the insignificant finding [41]. Specifically, according to a qualitative study among MSM in Hong Kong, the MSM community is less cautious about other STIs because they are curable and treatable [42]. However, given the rising number of other STIs in MSM, there is an urgent need to strengthen STI screening and care for MSM [43].

Limitations

There are several limitations and implications for further studies. First, social media, such as Instagram, was used to recruit participants. An online screening survey and electronic consent forms were also used in this study. Participants who were not technology-savvy may thus have been excluded in the recruitment phase. Therefore, the study findings may not be applicable to those who are not technology-savvy and those with low eHealth literacy. Second, study participants were not blinded to group allocation. The study might be subject to performance bias, as participant knowledge of group allocation may have affected their behaviors. Third, we only evaluated participants' intention to have chemsex in the last 3 months. Further studies should assess their intention to have chemsex in the future such as in the next 3 months. Fourth, the study outcomes were measured 3 months after the baseline, so the long-term sustainability of the intervention remains unknown. Trials with a longer follow-up period are needed to evaluate the sustained effect of eHealth interventions in enhancing the self-efficacy of MSM in refusing risky sexual behaviors and chemsex. Also, an analysis of the cost-effectiveness of eHealth interventions for chemsex should be considered. Fifth, we did not record information about the study participants' engagement, such as the duration or frequency of their visits. These parameters should be recorded in the future to further investigate how study engagement impacts the effectiveness of interventions. Sixth, during the follow-up interviews, the number of people who had engaged in chemsex in the last 3 months was small (n=31), which might have affected the statistical power and precision of some of the secondary outcomes. Seventh, qualitative interviews should be conducted to obtain qualitative feedback from study participants in the future. Finally, similar trials should be developed and tested in other geographic areas to provide more empirical evidence to support the effectiveness of such interventions in enhancing the self-efficacy of MSM in refusing risky sexual behaviors and chemsex.

Conclusion

This study suggests that a web-based intervention with a harm reduction approach can enhance the self-efficacy of MSM in refusing risky sexual behaviors and chemsex and improve HIV testing in the 3-month follow-up interview. We also provide initial evidence that such interventions can reduce both the intention of MSM to engage in chemsex and their actual engagement in chemsex.

Acknowledgments

The study was sponsored by the Council for the AIDS Trust Fund (reference number: MSS 338 R).

This HIV/AIDS prevention part of the project/program is sponsored by the AIDS Trust Fund. The content of this article represents the opinion of our organization only. It does not represent the position of the AIDS Trust Fund. The AIDS Trust Fund is not responsible for any claims, demands, or liabilities whatsoever arising from or in connection with the use of any information contained in this article or the participation of the sponsored project/program.

EPFC is supported by an Australian National Health and Medical Research Council (NHMRC) Emerging Leadership Investigator Grants (GNT1172873) outside of the submitted work.

Data Availability

The content of the web-based intervention, data sets generated during or analyzed during the current study, and related syntax are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Content of the intervention and control groups.
[PDF File (Adobe PDF File), 90 KB - publichealth_v9i1e42902_app1.pdf]

Multimedia Appendix 2
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1231 KB - publichealth_v9i1e42902_app2.pdf]

Multimedia Appendix 3
Comparisons of baseline characteristics and study outcomes between non-dropout and dropout participants.
[PDF File (Adobe PDF File), 101 KB - publichealth_v9i1e42902_app3.pdf]

Multimedia Appendix 4
Mixed effects models for comparison of subscale scores of the Condom Self-Efficacy Scale.
[PDF File (Adobe PDF File), 62 KB - publichealth_v9i1e42902_app4.pdf]

Multimedia Appendix 5
Mixed effects models for comparison of study outcomes with the results of between-group difference at each time point.
[PDF File (Adobe PDF File), 142 KB - publichealth_v9i1e42902_app5.pdf]

Multimedia Appendix 6
Mixed effects models for comparison of study outcomes with an adjustment of baseline characteristics.
[PDF File (Adobe PDF File), 56 KB - publichealth_v9i1e42902_app6.pdf]

Multimedia Appendix 7
Complete case analysis for comparison of study outcomes (n=275).
[PDF File (Adobe PDF File), 20 KB - publichealth_v9i1e42902_app7.pdf]

References

1. McCall H, Adams N, Mason D, Willis J. What is chemsex and why does it matter? BMJ 2015 Nov 03;351:h5790. [doi: 10.1136/bmj.h5790] [Medline: 26537832]
2. Maxwell S, Shahmanesh M, Gafos M. Chemsex behaviours among men who have sex with men: A systematic review of the literature. Int J Drug Policy 2019 Jan;63:74-89. [doi: 10.1016/j.drugpo.2018.11.014] [Medline: 30513473]

3. Evers Y, Hoebe C, Dukers-Muijrers N, Kampman C, Kuizenga-Wessel S, Shilue D, et al. Sexual, addiction and mental health care needs among men who have sex with men practicing chemsex - a cross-sectional study in the Netherlands. Prev Med Rep 2020 Jun;18:101074 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101074] [Medline: 32226730]
4. Bohn A, Sander D, Köhler T, Hees N, Oswald F, Scherbaum N, et al. Chemsex and mental health of men who have sex with men in Germany. Front Psychiatry 2020 Nov 4;11:542301 [FREE Full text] [doi: 10.3389/fpsyt.2020.542301] [Medline: 33329083]
5. Berg RC, Amundsen E, Haugstvedt Å. Links between chemsex and reduced mental health among Norwegian MSM and other men: results from a cross-sectional clinic survey. BMC Public Health 2020 Nov 25;20(1):1785 [FREE Full text] [doi: 10.1186/s12889-020-09916-7] [Medline: 33238948]
6. Malandain L, Mosser S, Mouchabac S, Blanc J, Alexandre C, Thibaut F. Chemical sex (chemsex) in a population of French university students. Dialogues Clin Neurosci 2021 Jun 01;23(1):39-43 [FREE Full text] [doi: 10.1080/19585969.2022.2042163] [Medline: 35860173]
7. Lawn W, Aldridge A, Xia R, Winstock AR. Substance-linked sex in heterosexual, homosexual, and bisexual men and women: an online, cross-sectional "Global Drug Survey" report. J Sex Med 2019 May;16(5):721-732. [doi: 10.1016/j.jsxm.2019.02.018] [Medline: 30952548]
8. Tomkins A, George R, Kliner M. Sexualised drug taking among men who have sex with men: a systematic review. Perspect Public Health 2019 Jan 30;139(1):23-33. [doi: 10.1177/1757913918778872] [Medline: 29846139]
9. Drückler S, van Rooijen MS, de Vries HJ. Chemsex among men who have sex with men: a sexualized drug use survey among clients of the sexually transmitted infection outpatient clinic and users of a gay dating app in Amsterdam, the Netherlands. Sexual Trans Dis 2018;45(5):325-331. [doi: 10.1097/olq.0000000000000753]
10. Bourne A, Reid D, Hickson F, Torres-Rueda S, Weatherburn P. Illicit drug use in sexual settings ('chemsex') and HIV/STI transmission risk behaviour among gay men in South London: findings from a qualitative study. Sex Transm Infect 2015 Dec 09;91(8):564-568. [doi: 10.1136/sextrans-2015-052052] [Medline: 26163510]
11. Sewell J, Miltz A, Lampe FC, Cambiano V, Speakman A, Phillips AN, Attitudes to and Understanding of Risk of Acquisition of HIV (AURAH) Study Group. Poly drug use, chemsex drug use, and associations with sexual risk behaviour in HIV-negative men who have sex with men attending sexual health clinics. Int J Drug Policy 2017 May;43:33-43. [doi: 10.1016/j.drugpo.2017.01.001] [Medline: 28189979]
12. Pakianathan M, Whittaker W, Lee M, Avery J, Green S, Nathan B, et al. Chemsex and new HIV diagnosis in gay, bisexual and other men who have sex with men attending sexual health clinics. HIV Med 2018 May 22;19(7):485-490 [FREE Full text] [doi: 10.1111/hiv.12629] [Medline: 29790254]
13. Flores Anato JL, Panagiotoglou D, Greenwald ZR, Blanchette M, Trottier C, Vaziri M, et al. Chemsex and incidence of sexually transmitted infections among Canadian pre-exposure prophylaxis (PrEP) users in the l'Actuel PrEP Cohort (2013-2020). Sex Transm Infect 2022 Dec;98(8):549-556 [FREE Full text] [doi: 10.1136/sextrans-2021-055215] [Medline: 35039437]
14. Phanuphak N, Ramautarsing R, Chinbunchorn T, Janamnuaysook R, Pengnonyang S, Termvanich K, et al. Implementing a status-neutral approach to HIV in the Asia-Pacific. Curr HIV/AIDS Rep 2020 Oct 29;17(5):422-430 [FREE Full text] [doi: 10.1007/s11904-020-00516-z] [Medline: 32725317]
15. CHP reviews local HIV/AIDS situation in second quarter of 2022. The Government of the Hong Kong Special Administrative Region. 2022. URL: https://www.info.gov.hk/gia/general/202208/30/P2022083000386.htm [accessed 2022-12-24]
16. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
17. Bavinton BR, Grulich AE. HIV pre-exposure prophylaxis: scaling up for impact now and in the future. The Lancet Public Health 2021 Jul;6(7):e528-e533. [doi: 10.1016/s2468-2667(21)00112-2]
18. Data by Country. PrEPWatch. URL: https://www.prepwatch.org/data-by-country/ [accessed 2022-12-24]
19. To KW, Lee SS. HIV pre-exposure prophylaxis in South East Asia: A focused review on present situation. Int J Infect Dis 2018 Dec;77:113-117 [FREE Full text] [doi: 10.1016/j.ijid.2018.10.027] [Medline: 30395980]
20. Wang Z, Lau JTF, Fang Y, Ip M, Gross DL. Prevalence of actual uptake and willingness to use pre-exposure prophylaxis to prevent HIV acquisition among men who have sex with men in Hong Kong, China. PLoS One 2018 Feb 12;13(2):e0191671 [FREE Full text] [doi: 10.1371/journal.pone.0191671] [Medline: 29432470]
21. Øyeflaten I, Johansen T, Nielsen CV, Johnsen TL, Tveito TH, Momsen AH. eHealth interventions to facilitate work participation. JBI Database of Systematic Reviews and Implementation Reports 2019;17(6):1026-1033. [doi: 10.11124/jbisrir-2017-003953]
22. Meiksin R, Melendez-Torres GJ, Falconer J, Witzel TC, Weatherburn P, Bonell C. eHealth interventions to address sexual health, substance use, and mental health among men who have sex with men: systematic review and synthesis of process evaluations. J Med Internet Res 2021 Apr 23;23(4):e22477 [FREE Full text] [doi: 10.2196/22477] [Medline: 33890855]
23. Xin M, Viswanath K, Li AY, Cao W, Hu Y, Lau JT, et al. The effectiveness of electronic health interventions for promoting HIV-preventive behaviors among men who have sex with men: meta-analysis based on an integrative framework of design

and implementation features. J Med Internet Res 2020 May 25;22(5):e15977 [FREE Full text] [doi: 10.2196/15977] [Medline: 32449685]

24. Meier CA, Fitzgerald MC, Smith JM. eHealth: extending, enhancing, and evolving health care. Annu Rev Biomed Eng 2013 Jul 11;15(1):359-382. [doi: 10.1146/annurev-bioeng-071812-152350] [Medline: 23683088]
25. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. J Med Internet Res 2014 May 26;16(5):e134 [FREE Full text] [doi: 10.2196/jmir.3393] [Medline: 24862459]
26. Nguyen LH, Tran BX, Rocha LEC, Nguyen HLT, Yang C, Latkin CA, et al. A systematic review of eHealth interventions addressing HIV/STI prevention among men who have sex with men. AIDS Behav 2019 Sep 10;23(9):2253-2272 [FREE Full text] [doi: 10.1007/s10461-019-02626-1] [Medline: 31401741]
27. Choi EPH, Chau PH, Wong WCW, Kowk JYY, Choi KWY, Chow EPF. Developing and testing of an interactive internet-based intervention to reduce sexual harm of sexualised drug use ('chemsex') among men who have sex with men in Hong Kong: a study protocol for a randomised controlled trial. BMC Public Health 2021 Apr 13;21(1):713 [FREE Full text] [doi: 10.1186/s12889-021-10742-8] [Medline: 33849517]
28. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
29. Wang Z, Yang X, Mo PKH, Fang Y, Ip TKM, Lau JTF. Influence of social media on sexualized drug use and chemsex among Chinese men who have sex with men: observational prospective cohort study. J Med Internet Res 2020 Jul 24;22(7):e17894 [FREE Full text] [doi: 10.2196/17894] [Medline: 32706705]
30. Zhao Y, Wong CKH, Miu HYH, Yuen WWY, Chin WY, Luo T, et al. Translation and validation of a Condom Self-Efficacy Scale (CSES) Chinese version. AIDS Educ Prev 2016 Dec;28(6):499-510. [doi: 10.1521/aeap.2016.28.6.499] [Medline: 27925485]
31. Alvy LM, McKirnan DJ, Mansergh G, Koblin B, Colfax GN, Flores SA, Project MIX Study Group. Depression is associated with sexual risk among men who have sex with men, but is mediated by cognitive escape and self-efficacy. AIDS Behav 2011 Aug 9;15(6):1171-1179. [doi: 10.1007/s10461-010-9678-z] [Medline: 20217471]
32. Martin GW, Wilkinson DA, Poulos CX. The drug avoidance self-efficacy scale. Journal of Substance Abuse 1995 Jan;7(2):151-163. [doi: 10.1016/0899-3289(95)90001-2]
33. Strong C, Huang P, Li C, Ku SW, Wu H, Bourne A. HIV, chemsex, and the need for harm-reduction interventions to support gay, bisexual, and other men who have sex with men. The Lancet HIV 2022 Oct;9(10):e717-e725. [doi: 10.1016/s2352-3018(22)00124-2]
34. Graf N, Dichtl A, Deimel D, Sander D, Stöver H. Chemsex among men who have sex with men in Germany: motives, consequences and the response of the support system. Sex. Health 2018;15(2):151. [doi: 10.1071/sh17142]
35. Herrijgers C, Poels K, Vandebosch H, Platteau T, van Lankveld J, Florence E. Harm reduction practices and needs in a Belgian chemsex context: findings from a qualitative study. Int J Environ Res Public Health 2020 Dec 04;17(23):9081 [FREE Full text] [doi: 10.3390/ijerph17239081] [Medline: 33291855]
36. Swanton R, Allom V, Mullan B. A meta-analysis of the effect of new-media interventions on sexual-health behaviours. Sex Transm Infect 2015 Feb 28;91(1):14-20. [doi: 10.1136/sextrans-2014-051743] [Medline: 25433051]
37. Pires CV, Gomes FC, Caldas J, Cunha M. Chemsex in Lisbon? Self-reflexivity to uncover the scene and discuss the creation of community-led harm reduction responses targeting chemsex practitioners. Contemporary Drug Problems 2022 Apr 21;49(4):434-452. [doi: 10.1177/00914509221094893]
38. Phillips T, Fairley C, Donovan B, Ong J, McNulty A, Marshall L, et al. Sexual health service adaptations to the coronavirus disease 2019 (COVID-19) pandemic in Australia: a nationwide online survey. Aust N Z J Public Health 2021 Dec;45(6):622-627 [FREE Full text] [doi: 10.1111/1753-6405.13158] [Medline: 34473388]
39. Thng C, Hughes I, Poulton G, O'Sullivan M. 18 months on: an interrupted time series analysis investigating the effect of COVID-19 on chlamydia and gonorrhoea testing and test positivity at the Gold Coast, Australia. Sex. Health 2022 Apr 19;19(2):127-131. [doi: 10.1071/sh21231]
40. Chan PS, Chidgey A, Lau J, Ip M, Lau JT, Wang Z. Effectiveness of a novel HIV Self-Testing Service with Online Real-Time Counseling Support (HIVST-Online) in Increasing HIV Testing Rate and Repeated HIV testing among men who have sex with men in Hong Kong: results of a pilot implementation project. Int J Environ Res Public Health 2021 Jan 15;18(2):729 [FREE Full text] [doi: 10.3390/ijerph18020729] [Medline: 33467770]
41. Sarno EL, Macapagal K, Newcomb ME. "The Main Concern is HIV, Everything Else is Fixable": indifference toward sexually transmitted infections in the era of biomedical HIV prevention. AIDS Behav 2021 Aug 20;25(8):2657-2660 [FREE Full text] [doi: 10.1007/s10461-021-03226-8] [Medline: 33745099]
42. Choi KWY, Choi EPH, Chow EPF, Wan EYF, Wong WCW, Wong JYH, et al. The experience of using dating applications for sexual hook-ups: a qualitative exploration among HIV-negative men who have sex with men in Hong Kong. J Sex Res 2021 Jul 03;58(6):785-794. [doi: 10.1080/00224499.2021.1886227] [Medline: 33656408]
43. Ramchandani MS, Golden MR. Confronting rising STIs in the era of PrEP and treatment as prevention. Curr HIV/AIDS Rep 2019 Jun 10;16(3):244-256 [FREE Full text] [doi: 10.1007/s11904-019-00446-5] [Medline: 31183609]

Abbreviations

CSES: Condom Self-Efficacy Scale
DASES: Drug Avoidance Self-Efficacy Scale
MSM: men who have sex with men
NHMRC: National Health and Medical Research Council
PrEP: pre-exposure prophylaxis
SESS: Self-Efficacy for Sexual Safety
STI: sexually transmitted infection
U=U: undetectable equals untransmittable

Edited by A Mavragani; submitted 23.09.22; peer-reviewed by S John, Z Ni; comments to author 19.10.22; revised version received 10.11.22; accepted 29.11.22; published 05.01.23.

Please cite as:
Choi EPH, Choi KWY, Wu C, Chau PH, Kwok JYY, Wong WCW, Chow EPF
Web-Based Harm Reduction Intervention for Chemsex in Men Who Have Sex With Men: Randomized Controlled Trial
JMIR Public Health Surveill 2023;9:e42902
URL: https://publichealth.jmir.org/2023/1/e42902
doi:10.2196/42902
PMID:36602853

Original Paper

Quality of Life and Mental Health of Chinese Sexual and Gender Minority Women and Cisgender Heterosexual Women: Cross-sectional Survey and Mediation Analysis

Chanchan Wu[1], MSc; Pui Hing Chau[1], PhD; Edmond Pui Hang Choi[1], PhD

[1]School of Nursing, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Edmond Pui Hang Choi, PhD
School of Nursing
The University of Hong Kong
5/F, Academic Building, 3 Sassoon Road
Pok Fu Lam
Hong Kong, 999077
China (Hong Kong)
Phone: 852 3917 6972
Email: h0714919@connect.hku.hk

Abstract

Background: Health-related research on sexual minority populations in China is lacking, and research on sexual and gender minority women (SGMW, including transgender women and persons of other gender identities assigned female at birth of all sexual orientations, and cisgender women with nonheterosexual orientations) is even less. Currently, there are limited surveys related to mental health in Chinese SGMW, but there are no studies on their quality of life (QOL), no studies comparing the QOL of SGMW with that of cisgender heterosexual women (CHW), and no studies on the relationship between sexual identity and the QOL as well as associated mental health variables.

Objective: This study aims to evaluate the QOL and mental health in a diverse sample of Chinese women and make comparisons between SGMW and CHW and then investigate the relationship between sexual identity and the QOL through the role of mental health.

Methods: A cross-sectional online survey was conducted from July to September 2021. All participants completed a structured questionnaire containing the World Health Organization Quality of Life–abbreviated short version (WHOQOL-BREF), the 9-item Patient Health Questionnaire (PHQ-9), the 7-item Generalized Anxiety Disorder scale (GAD-7), and the Rosenberg Self-Esteem Scale (RSES).

Results: In total, 509 women aged 18-56 years were recruited, including 250 (49.1%) CHW and 259 (50.9%) SGMW. Independent *t* tests showed that the SGMW reported significantly lower levels of QOL, higher levels of depression and anxiety symptoms, and lower self-esteem than the CHW. Pearson correlations showed that every domain and the overall QOL were positively associated with mental health variables, with moderate-to-strong correlations (r range 0.42-0.75, *P*<.001). Multiple linear regressions found that participants belonging to the SGMW group, current smokers, and women with no steady partner were associated with a worse overall QOL. The mediation analysis found that depression, anxiety, and self-esteem significantly completely mediated the relationship between sexual identity and physical, social, and environment domains of the QOL, while the relationship between sexual identity and the overall QOL and psychological QOL was partially mediated by depression and self-esteem.

Conclusions: The SGMW had poorer levels of QOL and a worse mental health status than the CHW. The study findings affirm the importance of assessing mental health and highlight the need to design targeted health improvement programs for the SGMW population, who may be at higher risk of a poor QOL and mental health.

(JMIR Public Health Surveill 2023;9:e42203) doi:10.2196/42203

KEYWORDS

quality of life; depression; anxiety; self-esteem; sexual and gender minority women; mental health; mediation analysis; China; minority; women; survey; social

Introduction

Public health surveillance systems, to date, have focused on cisgender heterosexual (CH) individuals, whose gender identity is consistent with their sex assigned at birth and who are sexually or romantically attracted to people of the opposite sex, while people with other gender and sexual orientation identities have historically been less visible [1]. In general, all non-CH individuals could be characterized as sexual and gender minority (SGM) populations. Specifically, gender minorities include but are not limited to people who are transgender (whose gender identity and sex assigned at birth do not correspond), gender queer (whose gender identity falls outside the traditional gender), gender fluid (whose gender identity is not fixed), gender questioning (who are unsure about or are exploring their gender identity), and other nonbinary genders outside of the traditional gender binary of male and female. Similarly, sexual orientation minorities include but are not limited to homosexual (gay and lesbian, people who are primarily sexually or romantically attracted to people of the same sex or gender as themselves), bisexual (attraction to multiple genders), pansexual (attraction to all genders/regardless of gender), and asexual (who experience little or no sexual attraction to others) or other nonheterosexual orientations.

A growing body of research has explored health disparities among people of different sexual orientations, and sexual minority populations have consistently been found to report worse health outcomes than their CH peers, especially in the area of mental health [2-9]. Specifically, studies investigating mental health and its relationship with sexual orientations have shown that sexual minority populations are more likely to experience mental health problems than CH individuals, with depression and anxiety being the most commonly diagnosed mental health disorders [3,10-16]. In a comprehensive meta-analysis of large samples, sexual minority populations were twice as likely to suffer from depression and anxiety as CH individuals [17]. It is worth noting that from a gender perspective, mental disorders are more commonly diagnosed in women than in men globally [18,19], so relevant research on women is crucial. Even in the field of research on SGM populations, more attention has been paid to males than females [20]; therefore, sexual and gender minority women (SGMW, including transgender women and persons of other gender identities assigned female at birth of all sexual orientations, and cisgender women with nonheterosexual orientations) need more public health attention.

As emphasized by the World Health Organization (WHO), mental health is “not just the absence of mental disorder” but “a state of well-being in which individuals realize their own potential and can cope with normal stresses of life” [19]. In other words, a healthy mental state is not only the absence of depression, anxiety, or other symptoms but also the presence of positive components of well-being. Self-esteem, literally defined as the degree to which individuals believe themselves to be valuable or adequate [21], is one such essential component and strongly related to happiness [22]. Similar to gender differences in mental disorders, a meta-analysis of self-esteem also revealed that women report lower levels of self-esteem than men [23]. Although high levels of self-esteem are associated with positive self-attitudes and are protective against adverse mental health outcomes [24], low levels of self-esteem, more commonly reported by people belonging to sexual minority populations [25], increase the likelihood of mental health difficulties.

In addition to mental health, a poor quality of life (QOL) has been shown to be positively associated with adverse mental health outcomes [26]. The QOL is widely used to present individuals’ perceptions of their physical and psychological health, social relationships, and environment [27]. However, there is compelling evidence that sexual minority populations are more likely to report a poorer QOL than CH individuals [28], especially psychological QOL [29]. To promote the overall health of SGM populations, it is not sufficient to only examine differences in the incidence of mental disorders and levels of the QOL without understanding the in-depth relationships and mechanisms among them. A recent review examining mediators of the depression disparities between sexual minority and CH populations suggested that the psychological process may also play a role, but this has been understudied to date [30], and more mediation analyses are needed.

Health-related research on SGMW is growing slowly worldwide [31-33], but there is less research on Chinese SGMW. According to published data from Western countries, sexual minority women experience a higher incidence of mental health disorders, including anxiety and depression, and also generally report a worse QOL than cisgender heterosexual women (CHW) [34,35]. A recent scoping review summarizing the holistic health of homosexual and bisexual Chinese over the past 20 years showed that the academic attention given to the male population is much higher than that given to female groups (96.46% vs 1.32%, respectively) [20]. Although there are limited surveys investigating the mental health of SGMW in China, results regarding depression and anxiety are less consistent among women from different countries with varied sexual identities [13,36-38]. Furthermore, there is currently no research on the QOL of Chinese SGMW and the relationship between sexual identity and the QOL, highlighting the need to research the health of women, especially SGMW in China.

The associations between sexual identity and the QOL as well as mental health vary across studies conducted in different countries, and there are currently no studies aimed at explaining these relationships among Chinese SGM populations, let alone SGMW. Therefore, this study aims to first examine the QOL and mental health of Chinese adult women of different sexual identities, with mental health variables including depression and anxiety as well as self-esteem, and then investigate the relationships between sexual identity and the QOL through the role of mental health variables in the Chinese context. Our main hypotheses were as follows: (1) SGMW in China may report lower levels of the QOL and more mental health disorders compared to CHW, (2) a better mental health status may positively be associated with higher levels of the QOL, and (3) mental health variables mediate the effect of sexual identity on the QOL.

Methods

Participants and Procedures

This was a cross-sectional study conducted online in China. Given the relative sensitivity and invisibility of our study population under the influence of traditional Chinese culture, multiple recruitment methods were used, including (1) convenient recruitment via 4 popular nongovernmental organizations (NGOs) serving female or SGM populations in mainland China (TrueSelf, r&B bisexual community, Period Pride China, and the Wuhan LGBT Center); (2) respondent-driven recruitment, where respondents were encouraged to help recruit potential peers through their network of connections [39,40]; and (3) social platform recruitment by releasing the study poster to some popular online communities.

Eligible participants included Chinese women who were (1) at least 18 years old; (2) self-identified as "female," either cisgender or transgender women or gender nonbinary individuals whose sex assigned at birth was female; and (3) able to read and understand Mandarin. Data were collected using an online survey platform (Wenjuanxing), and data collection was initiated after informed consent was obtained.

Ethical Considerations

Ethical approval was obtained from the Human Research Ethics Committee (HREC) of the University of Hong Kong (reference no. EA210325). All participants provided informed consent.

Study Instruments

Participants who consented to join the study were asked to fill out all the following questionnaires in Mandarin.

The 9-Item Patient Health Questionnaire

The 9-item Patient Health Questionnaire (PHQ-9) is a depression module that measures the presence and severity of depression symptoms over the past 2 weeks according to the *Diagnostic and Statistical Manual of Mental Disorders IV* (DSM-IV) criteria for assessing symptoms of depression [41]. The PHQ-9 is a self-report questionnaire, and each item ranges from 0 (not at all) to 3 (nearly every day), with a summed score ranging from 0 to 27. The Chinese version of the PHQ-9 has been previously evaluated for reliability and validity, showing that it is a valid and efficient tool for screening depression [42]. This study used this Chinese version of the PHQ-9, and the Cronbach α of the scale among the study sample was .91.

The 7-Item Generalized Anxiety Disorder Scale

The 7-item Generalized Anxiety Disorder scale (GAD-7) is a self-reported screening tool to identify anxiety symptoms in the past 2 weeks. Each of the 7 items is scored from 0 (not at all) to 3 (nearly every day), and the GAD-7 scale score ranges from 0 to 21 [43]. The Chinese version of GAD-7 has been previously validated and is commonly used [44,45]. This study used this Chinese version, and the Cronbach α of the scale among the study sample was .94.

The Rosenberg Self-Esteem Scale

The Rosenberg Self-Esteem Scale (RSES) is a 10-item self-report scale evaluating an individual's self-esteem [21]. Each item is answered using a 4-point Likert scale ranging from 1 (strongly agree) to 4 (strongly disagree), while half items are reverse-scored ranging from 1 (strongly disagree) to 4 (strongly agree). The Chinese version of the questionnaire has been previously validated [46-49], and this study used this Chinese version to measure participants' self-esteem. All scores of the 10 items are summed as a global score on a continuous scale, and higher scores indicate lower overall self-esteem. In this study, the Cronbach α of the scale was .77.

The World Health Organization Quality of Life–Abbreviated Short Version

The World Health Organization Quality of Life–abbreviated short version (WHOQOL) is a QOL assessment tool developed by the WHOQOL Group and includes 100 items on a 5-point scale [27]. Based on that, an abbreviated short version of the WHOQOL instrument (WHOQOL-BREF) has been developed and is widely used in assessing the QOL of the public [50,51]. The WHOQOL-BREF contains 26 items in total, including 2 (8%) overall items and 24 (92%) facet items on 4 domains: physical (7, 29%, items), psychological (6, 25%, items), social relationships (3, 13%, items), and environment (8, 33%, items). Regarding the scoring, there are 3 negatively phrased items (items 3, 4, and 26) that should be reversed before calculating the scores. The mean score of items within each domain is used to calculate the domain score, and each domain score can be transformed into a score ranging from 4 to 20, with higher scores representing a better QOL [27,51,52]. The Chinese version of the WHOQOL-BREF has been previously validated [51] and was used in this study. In this study, the Cronbach α coefficient of the total scale was .93 and of the 4 subscales was .79, .88, .68, and .83, respectively.

Sociodemographic Information

All participants completed a set of sociodemographic items, including their basic information (age, employment, race/ethnicity, education level, income, relationship status) and their sexual identity (gender identity and sexual orientation identity).

Sample Size Justification

The sample size calculation was performed using G*Power 3.1 [53]. To detect a statistically significant difference between CHW and SGMW using the independent samples *t* test with 80% power, 5% level of significance (2-tailed), and a Cohen *d* effect size of 0.3, a minimum sample size of 352 participants was needed.

Statistical Analysis

Descriptive statistics (counts and percentages, means, and SDs) were used to illustrate participants' demographic characteristics, QOL, and mental health outcomes.

Comparisons Between SGMW and CHW

To compare the scores of mental health variables (PHQ-9, GAD-7, RSES) and the QOL (WHOQOL-BREF) between SGMW and CHW, independent *t* tests were conducted and the Cohen *d* effect size was also calculated, with cut-offs of 0.2, 0.5, and 0.8 for small, medium, and large effect sizes, respectively [54]. To compare the percentages of participants' demographic characteristics, chi-square tests were conducted. Pearson correlations were used to identify the correlations between mental health variables and the QOL, and the correlations were defined as strong (≥0.5), moderate (≥0.3 and <0.5), or weak (<0.3). One-way ANOVA was performed to compare the mean scores in 5 sexual orientation groups. Next, multiple linear regression analyses were performed to identify factors associating mental health and the QOL separately.

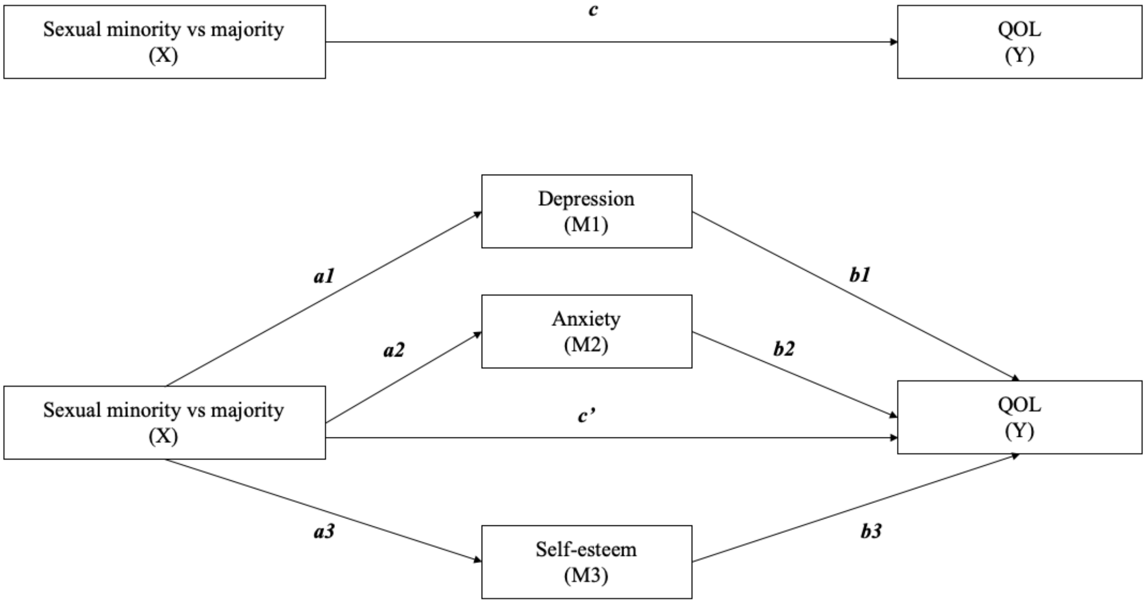
Parallel Mediation Analysis

To examine mediating effects, we used a bootstrapping technique using the PROCESS macro [55], and parallel mediation models were assessed. We implemented 5000 bootstrap samples for the percentile bootstrap CI, and a 95% confidence level was used. We proposed the effects of sexual identity (independent X: sexual minority vs majority) through 3 mediating mental health variables (parallel mediators: M1, depression; M2, anxiety; and M3, self-esteem) on the QOL (dependent Y: physical QOL, psychological QOL, social relationship QOL, environment QOL, and overall QOL). The existing literature has suggested that many sociodemographic factors are associated with the QOL [55-57], so age, income, alcohol use, smoking status, and relationship status were included as covariates to control for their potential effects in the mediation analysis.

As shown in Figure 1, “a” is the effect of the independent variable (X) on mental health (mediators M1-M3), “b” is the effect of mental health (M1-M3) on the QOL (Y), “c” is the total effect of X on Y, and c’ is the direct effect of X on Y. An indirect effect was considered significant if the CI did not contain 0. If the indirect effect is significant, it means that there is a mediation effect. Meanwhile, if the direct effect disappears, then there is complete mediation, whereas if the direct effect remains, then there is partial mediation [58]. Thus, we expected that sexual minority status would be positively associated with mental disorders (higher levels of depression and anxiety but lower levels of self-esteem) and, in turn, be negatively associated with the QOL.

Data were analyzed using IBM Statistical Package for the Social Sciences (SPSS) version 27.0, and the PROCESS macro version 4.1 for SPSS was used to conduct the mediation analysis [55]. All significance tests were 2-tailed, and findings with *P*<.05 were considered statistically significant.

Figure 1. Proposed parallel mediation model; a: effect of the independent variable (X) on mental health (M1-M3), b: effect of mental health on QOL (Y), c: total effect of X on Y, c’: direct effect of X on Y. M: mediator; QOL: quality of life.



Results

Demographic Characteristics of Study Participants

Of the 524 questionnaires collected between July and September 2021, 15 (2.9%) were unqualified since they were cisgender male (9/15, 60%) or less than 18 years old (4/15, 27%) or invalid data with all same answers (2/15, 13%), resulting in a total of 509 (97.1%) adult women enrolled in the study with no missing data. Specifically, 250 (49.1%) were CHW and categorized as the CHW group, while 259 (50.9%) were either sexual orientation or gender minority women and categorized as the SGMW group. The sexual orientation and gender identity of participants are shown in Table 1.

The mean age of the overall sample was 25.57 (SD 5.77) years, ranging from 18 to 56 years (Table 2). The mean age of the CHW was 26.75 (SD 6.31) years and was 24.43 (SD 4.94) years for the SGMW. Most participants were Han people (n=466, 91.6%) in the age group of 21-30 years (n=347, 68.2%) and had college or bachelor’s degrees (n=284, 55.8%), and around two-fifths were full-time students (n=222, 43.6%). About half of the women were single, and half had a steady partner (n=258, 50.7%, vs n=251, 49.3%), and a total of 248 (48.7%) women had full-time jobs. One-third of the women (n=178, 35.0%) never had a drink before, and the majority (n=428, 84.1%) had never smoked or had already quit smoking. Notably, 9 (1.8%) participants had a history of drug use.

Table 1. Frequency distribution of participants by gender and sexual orientation identity (N=509).

Gender identity	Sexual orientation identity, n (%)					
	Homosexual	Heterosexual	Bisexual	Pansexual	Others[a]	Total
Cisgender women	51 (10.0)	250 (49.1)[b]	74 (14.5)	33 (6.5)	28 (5.5)	436 (85.7)
Transgender women	8 (1.6)	1 (0.2)	1 (0.2)	1 (0.2)	2 (0.4)	13 (2.6)
Gender queer	10 (2.0)	1 (0.2)	4 (0.8)	12 (2.4)	1 (0.2)	28 (5.5)
Gender fluid	0	0	4 (0.8)	2 (0.4)	1 (0.2)	7 (1.4)
Other gender nonbinary[c]	2 (0.4)	7 (1.4)	5 (1.0)	4 (0.8)	7 (1.4)	25 (4.9)
Total	71 (13.9)	259 (50.9)	88 (17.3)	52 (10.2)	39 (7.7)	509 (100.0)

[a] Other minority sexual orientation identities, including asexual, gynephilia/femininity, and all other nonheterosexual orientations.

[b] The cisgender heterosexual women (CHW) group in this study comprised 250 (49.1%) women. Correspondingly, the other 259 (50.9%) sexual orientation and gender minority women were categorized as the sexual and gender minority women (SGMW) group.

[c] This group included gender questioning (someone who is unsure about or is exploring their gender identity), no gender, and other nonbinary genders that are noncisgender and beyond social norms.

Table 2. Characteristics of the study sample with comparisons.

Characteristics	Overall (N=509)	SGMW[a] (n=259, 50.1%)	CHW[b] (n=250, 49.1%)
Age (years), n (%)			
18-20	94 (18.5)	61 (23.6)	33 (13.2)
21-30	347 (68.2)	171 (66.0)	176 (70.4)
31-40	56 (11.0)	25 (9.7)	31 (12.4)
41-56	12 (2.4)	2 (0.8)	10 (4.0)
Mean age, mean (SD); *P*<.001[c]	25.57 (5.77)	24.43 (4.94)	26.75 (6.31)
Number of friends, n (%)			
0	144 (28.3)	48 (33.3)	96 (66.7)
1-5	154 (30.3)	96 (62.3)	58 (37.7)
6-10	74 (14.5)	50 (67.6)	24 (32.4)
11-20	43 (8.4)	28 (65.1)	15 (34.9)
21-50	40 (7.9)	20 (50.0)	20 (50.0)
51-100	27 (5.3)	12 (44.4)	15 (55.6)
>100	27 (5.3)	5 (18.5)	22 (81.5)
Mean numbers of friends around, mean (SD); *P*=.005	33.44 (108.28)	20.14 (71.40)	47.22 (135.14)
Education, n (%); *P*<.001			
High school and below	33 (6.5)	21 (8.1)	12 (4.8)
College/bachelor	284 (55.8)	167 (64.5)	117 (46.8)
Graduate degree and above	192 (37.7)	71 (27.4)	121 (48.4)
Ethnic, n (%); *P*=.12			
Han people	466 (91.6)	242 (93.4)	224 (89.6)
Others (Muslim, Zhuang people, etc)	43 (8.4)	17 (6.6)	26 (10.4)
Employment, n (%); *P*<.001			
Unemployed	39 (7.7)	29 (11.2)	10 (4.0)
Full-time student	222 (43.6)	118 (45.6)	104 (41.6)
Organization officer	42 (8.3)	23 (8.9)	19 (7.6)
Professional/technical personnel	103 (20.2)	29 (11.2)	74 (29.6)
Administrative personnel	17 (3.3)	7 (2.7)	10 (4.0)
Service personnel	33 (6.5)	19 (7.3)	14 (5.6)
Others (famers/freelancers)	53 (10.4)	34 (13.1)	19 (7.6)
Monthly income (CNY/US $[d]), n (%);*P*<.001			
≤CNY 5000.00/US $743.89	282 (55.4)	162 (62.5)	120 (48.0)
>CNY 5000.00/US $743.89	227 (44.6)	97 (37.5)	130 (52.0)
Relationship, n (%); *P*<.001			
Have a steady partner	251 (49.3)	115 (44.4)	136 (54.4)
Have no steady partner(s)	258 (50.7)	144 (55.6)	114 (45.6)
Alcohol use, n (%); *P*<.001			
Never drank before	178 (35.0)	68 (26.3)	110 (44.0)
Current alcohol user	331 (65.0)	191 (73.7)	140 (56.0)
Drug use, n (%); *P*=.10			
Never used before	500 (98.2)	252 (97.3)	248 (99.2)

Characteristics	Overall (N=509)	SGMW[a] (n=259, 50.1%)	CHW[b] (n=250, 49.1%)
Have used drug before	9 (1.8)	7 (2.7)	2 (0.8)
Smoking, n (%); *P*<.001			
Never smoked or have quit smoking	428 (84.1)	188 (72.6)	240 (96.0)
Current smoker	81 (15.9)	71 (27.4)	10 (4.0)

[a] SGMW: sexual and gender minority women. The percentages for this column were all calculated with 259 SGMW as the denominator.

[b] CHW: cisgender heterosexual women. The percentages for this column were all calculated with 250 CHW as the denominator.

[c] All *P* values were reported by conducting comparisons performing independent *t* tests or chi-square tests.

[d] CNY 1=US $0.15.

QOL and Mental Health Among Study Participants With Comparisons

Table 3 shows the mean scores of the QOL (as measured by the WHOQOL-BREF) and mental health outcomes (as measured by the PHQ-9, GAD-7, and RSES) for our study participants. The mean QOL scores of each domain in the overall sample were ranked as follows: physical QOL (mean 14.05, SD 2.71), psychological QOL (mean 13.74, SD 3.13), environment QOL (mean 13.74, SD 2.68), and social relationship QOL (mean 13.62, SD 3.07). The mean scores of mental health variables in the overall sample were 7.85 (SD 5.92) for depression, 6.00 (SD 5.23) for anxiety, and 28.86 (SD 4.66) for self-esteem. All comparisons indicated that compared to the CHW group, the SGMW group reported significantly lower levels of overall QOL (52.77 vs 57.61) and QOL in all domains, lower levels of self-esteem (28.04 vs 29.72), but higher levels of depression (9.16 vs 6.50) and anxiety (5.49 vs 4.77) symptoms (all *P*<.01). The comparisons among 5 different sexual orientation groups were also conducted (Multimedia Appendix 1).

Table 3. QOL[a] and mental health of the study participants with comparisons.

QOL and mental health	Overall (N=509)	SGMW[b] (n=259, 50.9%)	CHW[c] (n=250, 49.1%)	*P* value	Cohen *d* (SGMW-CHW)
QOL (WHOQOL-BREF[d]), mean (SD)					
Physical	14.05 (2.71)	13.49 (2.87)	14.63 (2.39)	<.001	–0.43
Psychological	13.74 (3.13)	12.94 (3.35)	14.57 (2.65)	<.001	–0.54
Social relationship	13.62 (3.07)	13.02 (3.35)	14.23 (2.61)	<.001	–0.40
Environment	13.74 (2.68)	13.32 (2.90)	14.18 (2.35)	<.001	–0.33
Total score	55.15 (9.90)	52.77 (10.69)	57.61 (8.34)	<.001	–0.50
Mental health, mean (SD)					
Depression (PHQ-9[e])	7.85 (5.92)	9.16 (6.36)	6.50 (5.11)	<.001	0.46
Anxiety (GAD-7[f])	6.00 (5.23)	6.93 (5.49)	5.04 (4.77)	<.001	0.37
Self-esteem (RSES[g])	28.86 (4.66)	28.04 (4.95)	29.72 (4.18)	<.001	–0.37

[a] QOL: quality of life.

[b] SGMW: sexual and gender minority women.

[c] CHW: cisgender heterosexual women.

[d] WHOQOL-BREF: World Health Organization Quality of Life–abbreviated short version.

[e] PHQ-9: 9-item Patient Health Questionnaire.

[f] GAD-7: 7-item Generalized Anxiety Disorder scale.

[g] RSES: Rosenberg Self-Esteem Scale.

Correlations Between the QOL and Mental Health

Pearson correlations between the QOL and depression, anxiety, and self-esteem are displayed in Table 4. As hypothesized, both depression and anxiety symptoms were significantly negatively associated with all domains of the QOL as well as the overall QOL, with moderate-to-strong correlations (*r* range –0.42 to –0.68, all *P*<.001). Self-esteem was significantly positively associated with all domains of the QOL as well as the overall QOL, with all strong correlations (*r* range 0.52-0.75, all *P*<.001).

Table 4. Correlations between the QOL[a] and mental health variables.

Mental health	QOL (WHOQOL-BREF[b])				
	Physical	Psychological	Social relationship	Environment	Total score
Depression (PHQ-9[c])					
r value	–0.66	–0.67	–0.52	–0.47	–0.68
P value	<.001	<.001	<.001	<.001	<.001
Anxiety (GAD-7[d])					
r value	–0.56	–0.55	–0.42	–0.45	–0.58
P value	<.001	<.001	<.001	<.001	<.001
Self-esteem (RSES[e])					
r value	0.60	0.75	0.53	0.52	0.71
P value	<.001	<.001	<.001	<.001	<.001

[a] QOL: quality of life.

[b] WHOQOL-BREF: World Health Organization Quality of Life–abbreviated short version.

[c] PHQ-9: 9-item Patient Health Questionnaire.

[d] GAD-7: 7-item Generalized Anxiety Disorder scale.

[e] RSES: Rosenberg Self-Esteem Scale.

Regressions of the QOL and Mental Health

Tables 5-12 display the results of multiple linear regressions. The identity of being a sexual minority woman had a significant negative association with the QOL in all domains as well as the overall QOL (all *P*<.01). In the regression models of mental health variables, belonging to the SGMW group was associated with more severe depression (*P*=.002) and anxiety (*P*=.02) symptoms as well as lower self-esteem levels (*P*=.02). The current smoking status was significantly negatively associated with the overall QOL and all other domains of the QOL except social relationship QOL. Regarding the regressions of mental health variables, the current smoking status was significantly associated with higher levels of depression (*P*<.001) and anxiety (*P*=.006) and lower levels of self-esteem (*P*=.002). Furthermore, having a steady partner was found to be significantly positively associated with the overall QOL (*P*=.006), psychological QOL (*P*=.02), and social relationship QOL (*P*<.001).

Table 5. Regressions of sexual identity on the physical QOL[a].

Characteristics	Physical QOL (R-square=7.9%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[b] vs SGMW[c])	0.77	.003	0.27 to 1.26
Age	0.00	.93	–0.04 to 0.05
Alcohol user (yes vs no)	–0.40	.12	–0.90 to 0.10
Drug user (yes vs no)	1.00	.27	–0.79 to 2.79
Current smoker (yes vs no)	–0.84	.02	–1.53 to –0.15
Income (≥CNY 5000/US $743.89[d] vs below)	0.52	.052	0.00 to 1.05
Have a steady partner (vs no)	0.41	.09	–0.07 to 0.89

[a] QOL: quality of life.

[b] CHW: cisgender heterosexual women.

[c] SGMW: sexual and gender minority women.

[d] CNY 1=US $0.15.

Table 6. Regressions of sexual identity on the psychological QOL[a].

Characteristics	Psychological QOL (R-square=11.3%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[b] vs SGMW[c])	1.13	<.001	0.56 to 1.69
Age	0.02	.48	–0.03 to 0.07
Alcohol user (yes vs no)	–0.11	.70	–0.68 to 0.46
Drug user (yes vs no)	0.63	.54	–1.40 to 2.66
Current smoker (yes vs no)	–1.38	.001	–2.16 to –0.60
Income (≥CNY 5000/US $743.89[d] vs below)	0.43	.15	–0.16 to 1.03
Have a steady partner (vs no)	0.64	.02	0.10 to 1.19

[a] QOL: quality of life.

[b] CHW: cisgender heterosexual women.

[c] SGMW: sexual and gender minority women.

[d] CNY 1=US $0.15.

Table 7. Regressions of sexual identity on the social relationship QOL[a].

Characteristics	Social relationship QOL (R-square=9.7%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[b] vs SGMW[c])	0.90	.002	0.35 to 1.46
Age	–0.07	.01	–0.12 to –0.01
Alcohol user (yes vs no)	–0.60	.04	–1.16 to –0.04
Drug user (yes vs no)	–1.45	.16	–3.45 to 0.56
Current smoker (yes vs no)	–0.74	.06	–1.51 to 0.04
Income (≥CNY 5000/US $743.89[d] vs below)	0.27	.37	–0.32 to 0.86
Have a steady partner (vs no)	1.10	<.001	0.57 to 1.64

[a] QOL: quality of life.

[b] CHW: cisgender heterosexual women.

[c] SGMW: sexual and gender minority women.

[d] CNY 1=US $0.15.

Table 8. Regressions of sexual identity on the environment QOL[a].

Characteristics	Environment QOL (R-square=4.4%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[b] vs SGMW[c])	0.67	.008	0.17 to 1.17
Age	–0.03	.20	–0.08 to 0.02
Alcohol user (yes vs no)	0.01	.97	–0.50 to 0.51
Drug user (yes vs no)	–0.21	.82	–2.01 to 1.59
Current smoker (yes vs no)	–0.74	.04	–1.44 to –0.05
Income (≥CNY 5000/US $743.89[d] vs below)	0.41	.13	–0.12 to 0.94
Have a steady partner (vs no)	0.30	.23	–0.19 to 0.78

[a] QOL: quality of life.

[b] CHW: cisgender heterosexual women.

[c] SGMW: sexual and gender minority women.

[d] CNY 1=US $0.15.

Table 9. Regressions of sexual identity on the overall QOL[a].

Characteristics	Overall QOL (R-square=10.3%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[b] vs SGMW[c])	3.47	<.001	1.68 to 5.26
Age	–0.08	.37	–0.24 to 0.09
Alcohol user (yes vs no)	–1.10	.23	–2.91 to 0.71
Drug user (yes vs no)	–0.02	.99	–6.47 to 6.43
Current smoker (yes vs no)	–3.70	.004	–6.19 to –1.21
Income (≥CNY 5000/US $743.89[d] vs below)	1.64	.09	–0.26 to 3.54
Have a steady partner (vs no)	2.45	.006	0.72 to 4.18

[a] QOL: quality of life.

[b] CHW: cisgender heterosexual women.

[c] SGMW: sexual and gender minority women.

[d] CNY 1=US $0.15.

Table 10. Regressions of sexual identity on mental health (depression).

Characteristics	Depression (R-square=8.5%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[a] vs SGMW[b])	-1.74	.002	–2.82 to –0.66
Age	-0.07	.16	–0.17 to 0.03
Alcohol user (yes vs no)	0.86	.12	–0.24 to 1.95
Drug user (yes vs no)	-2.01	.31	–5.91 to 1.89
Current smoker (yes vs no)	2.50	.001	0.99 to 4.00
Income (≥CNY 5000/US $743.89[c] vs below)	-0.03	.95	–1.18 to 1.12
Have a steady partner (vs no)	-0.42	.43	–1.47 to 0.63

[a] CHW: cisgender heterosexual women.

[b] SGMW: sexual and gender minority women.

[c] CNY 1=US $0.15.

Table 11. Regressions of sexual identity on mental health (anxiety).

Characteristics	Anxiety (R-square=6.9%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[a] vs SGMW[b])	-1.11	.02	–2.07 to –0.15
Age	-0.03	.46	–0.12 to 0.06
Alcohol user (yes vs no)	1.06	.03	0.09 to 2.04
Drug user (yes vs no)	-2.34	.19	–5.81 to 1.13
Current smoker (yes vs no)	1.87	.006	0.53 to 3.21
Income (≥CNY 5000/US $743.89[c] vs below)	-0.33	.53	–1.35 to 0.70
Have a steady partner (vs no)	-0.68	.15	–1.61 to 0.25

[a] CHW: cisgender heterosexual women.

[b] SGMW: sexual and gender minority women.

[c] CNY 1=US $0.15.

Table 12. Regressions of sexual identity on mental health (self-esteem).

Characteristics	Self-esteem (R-square=7.8%)		
	Coefficient	*P* value	95% CI
Sexual identity (CHW[a] vs SGMW[b])	1.01	.02	0.15 to 1.86
Age	0.00	.97	–0.08 to 0.08
Alcohol user (yes vs no)	0.04	.92	–0.82 to 0.91
Drug user (yes vs no)	1.58	.31	–1.50 to 4.66
Current smoker (yes vs no)	-1.91	.002	–3.10 to –0.72
Income (≥CNY 5000/US $743.89[c] vs below)	1.33	.004	0.42 to 2.23
Have a steady partner (vs no)	0.60	.16	–0.23 to 1.43

[a] CHW: cisgender heterosexual women.

[b] SGMW: sexual and gender minority women.

[c] CNY 1=US $0.15.

Mediation Effects Between Sexual Identity and the QOL via Mental Health

Figures 2-6 and Table 13 show the effects of parallel mediation analyses and the corresponding unstandardized effect coefficients of mental health on the relationship between sexual identity status and the QOL. As demonstrated, individuals' sexual identity was significantly associated with depression (β=–1.72, *P*=.002), anxiety (β=–1.12, *P*=.02), and self-esteem (β=.93, *P*=.03). As mediators, depression and self-esteem were significantly associated with the QOL in all domains as well as the overall QOL (β ranging from –.07 to –.68 and .19 to .98, respectively; all *P*<.05), while anxiety was only significantly associated with the QOL in the environment domain (β=–.07, *P*=.04). Likewise, the indirect effects of depression, anxiety, and self-esteem were only significant (95% CI did not contain 0) in the environment domain of the QOL, while the indirect effects of depression and self-esteem were significant in all other domains of the QOL as well as the overall QOL, indicating that anxiety is only a mediator in the environment domain but not in other aspects. After considering the mediation effects of depression, anxiety, and self-esteem, the direct effects (c') between sexual identity and the physical QOL (β=.20, *P*=.28), social relationship QOL (β=.35, *P*=.14), and environment QOL (β=.21, *P*=.32) were no longer statistically significant, supporting complete mediation effects. In contrast, the direct effects between sexual identity and the psychological QOL (β=.44, *P*=.02) and the overall QOL (β=1.20, *P*=.04) remained statistically significant, supporting partial mediation effects.

Figure 2. Parallel mediation model for the physical QOL. M: mediator; QOL: quality of life.

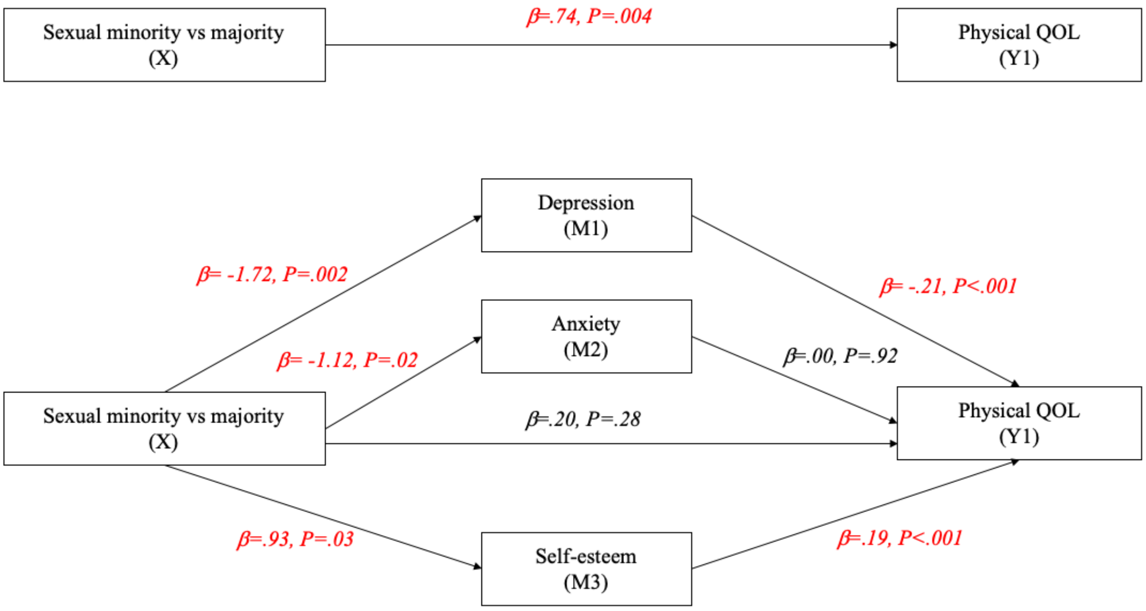


Figure 3. Parallel mediation model for the psychological QOL. M: mediator; QOL: quality of life.

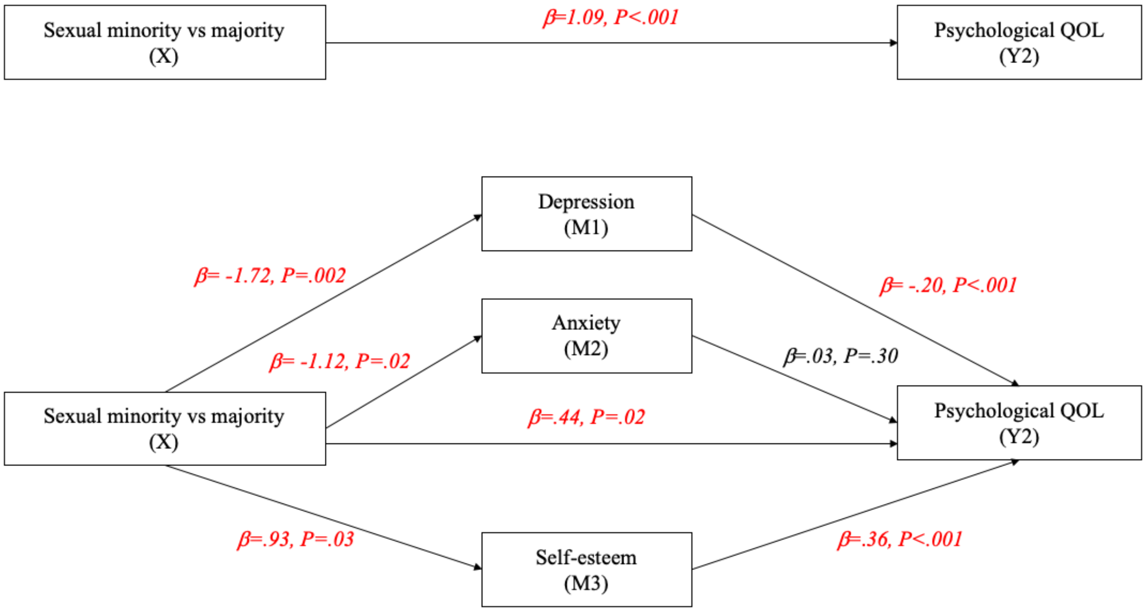


Figure 4. Parallel mediation model for the social relationship QOL. M: mediator; QOL: quality of life.

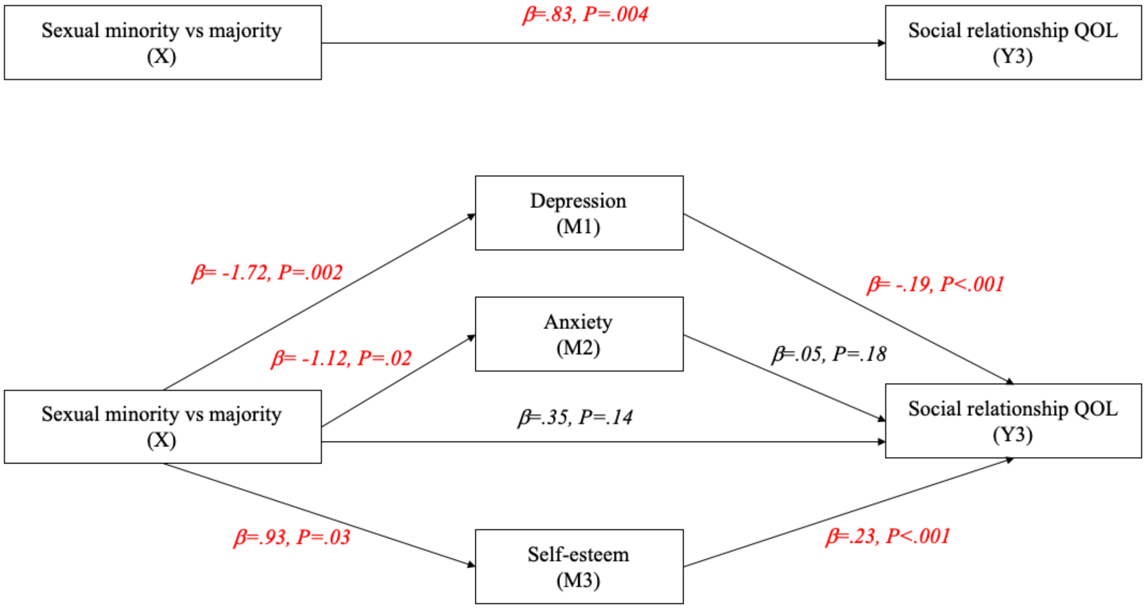


Figure 5. Parallel mediation model for the environment QOL. M: mediator; QOL: quality of life.

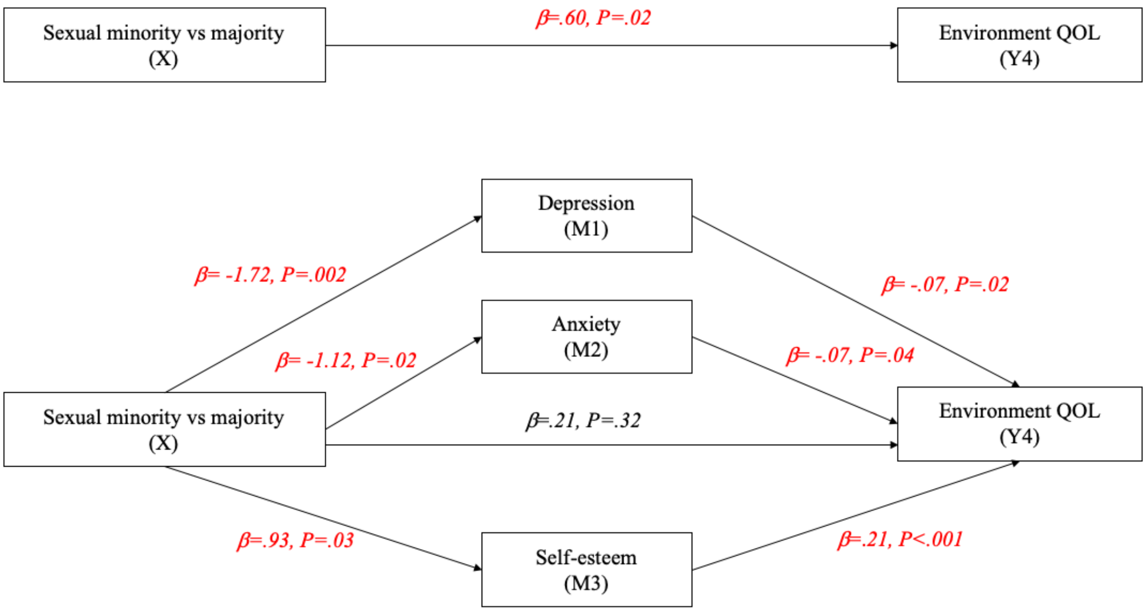


Figure 6. Parallel mediation model on the overall QOL. M: mediator; QOL: quality of life.

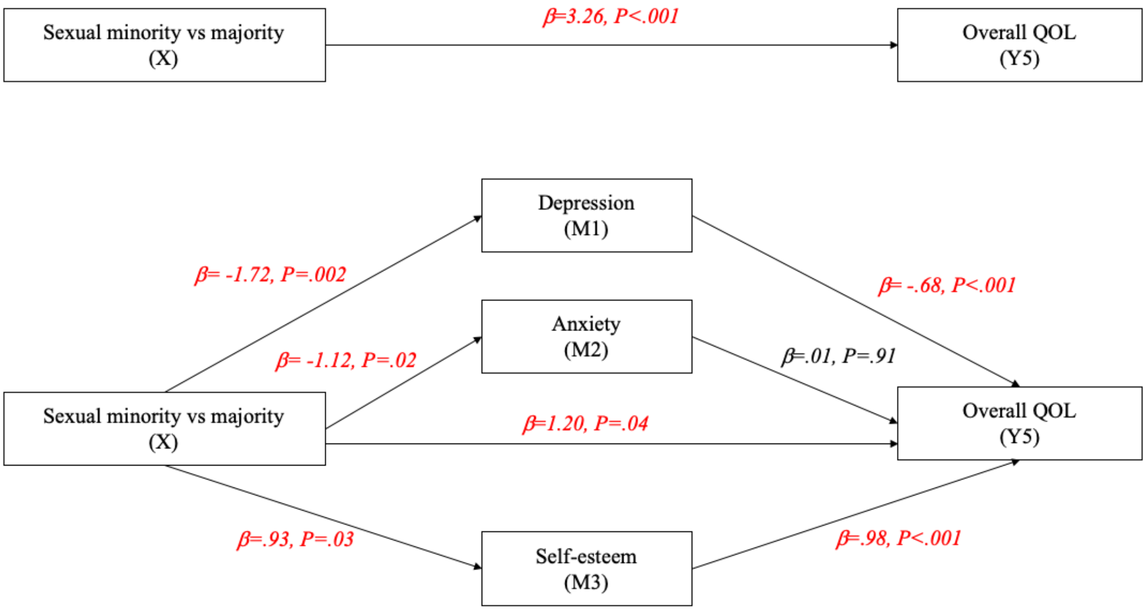


Table 13. Mediation effects of mental health on the relationship between sexual identity and the QOL[a].

Dependent variable (Y)	Total effect (c)			Direct effect (c')			X→M[b] (a1, a2, a3)[c]			M→Y (b1, b2, b3)[d]			Indirect effect (a×b)		
	β[e]	*P* value	SE	β	*P* value	SE	β	*P* value	SE	β	*P* value	SE	β	SE	95% CI
Sexual identity → physical QOL (X→Y1)															
X→M1→Y1	.74	.004	.25	.20	.28	.18	–1.72	.002	.55	–.21	<.001	.03	.37	.13	0.12 to 0.65
X→M2→Y1	—[f]	—	—	—	—	—	–1.12	.02	.49	.00	.92	.03	.00	.04	–0.08 to 0.08
X→M3→Y1	—	—	—	—	—	—	.93	.03	.43	.19	<.001	.02	.17	.08	0.02 to 0.35
Sexual identity → psychological QOL (X→Y2)															
X→M1→Y2	1.09	<.001	.29	.44	.02	.18	–1.72	.002	.55	–.20	<.001	.03	.35	.13	0.11 to 0.61
X→M2→Y2	—	—	—	—	—	—	–1.12	.02	.49	.03	.30	.03	–.03	.04	–0.12 to 0.03
X→M3→Y2	—	—	—	—	—	—	.93	.03	.43	.36	<.001	.02	.33	.15	0.03 to 0.63
Sexual identity → social relationship QOL (X→Y3)															
X→M1→Y3	.83	.004	.28	.35	.14	.24	–1.72	.002	.55	–.19	<.001	.03	.33	.12	0.11 to 0.59
X→M2→Y3	—	—	—	—	—	—	–1.12	.02	.49	.05	.18	.04	–.05	.05	–0.18 to 0.03
X→M3→Y3	—	—	—	—	—	—	.93	.03	.43	.23	<.001	.03	.21	.10	0.02 to 0.41
Sexual identity → environment QOL (X→Y4)															
X→M1→Y4	.60	.02	.25	.21	.32	.22	–1.72	.002	.55	–.07	.02	.03	.12	.07	0.01 to 0.28
X→M2→Y4	—	—	—	—	—	—	–1.12	.02	.49	–.07	.04	.03	.08	.06	0.00 to 0.20
X→M3→Y4	—	—	—	—	—	—	.93	.03	.43	.21	<.001	.03	.19	.09	0.02 to 0.37
Sexual identity → overall QOL (X→Y5)															
X→M1→Y5	3.26	<.001	.91	1.20	.04	.59	–1.72	.002	.55	–.68	<.001	.09	1.17	.43	0.37 to 2.06
X→M2→Y5	—	—	—	—	—	—	–1.12	.02	.49	.01	.91	.09	–.01	.11	–0.26 to 0.21
X→M3→Y5	—	—	—	—	—	—	.93	.03	.43	.98	<.001	.07	.91	.42	0.09 to 1.75

[a] QOL: quality of life.

[b] M: mediator.

[c] a1, a2, and a3 are the effects of the sexual identity (X) on the mental health mediators depression (M1), anxiety (M2), and self-esteem (M3), respectively.

[d] Correspondingly, b1, b2, and b3 are the effects of M1, M2, and M3, respectively, on the QOL (Y).

[e] Unstandardized coefficients (β) are reported. Bootstrap sample size=5000.

[f] Not applicable.

Discussion

Principal Findings

The QOL is widely used to describe individuals' perceptions of general and holistic health [27,52]. SGMW in this study reported significantly lower levels of overall and domain-specific QOL, higher levels of depression and anxiety, and lower levels of self-esteem compared to CHW. Our findings are consistent with several large-scale reviews of the mental health of sexual minority populations in Western countries [3,10], all suggesting that people belonging to sexual minorities are at higher risk of mental health disorders. Meanwhile, our study findings are consistent with the findings in Hong Kong, both showing that Chinese sexual minority populations report poorer levels of the QOL than CH people [28]. In terms of positive mental health, the CHW in our study reported significantly higher levels of self-esteem than the SGMW, which may be related to the fact that the CH identity is mainstreamed and encouraged in Chinese society, suggesting the need to consider specific social contexts when exploring the health of SGM populations.

In terms of cultural issues affecting the QOL and mental health of SGMW in China, it is critical to consider the deep-rooted traditional Confucian Chinese culture that highly values family and inheritance. Specifically, being a member of SGM populations in China is associated with internalized homophobia and stressful feelings in the face of social expectations [25,59,60], which is found to be negatively correlated with lower self-evaluation [61] and social well-being [62]. Compared to the twentieth century, the living conditions of sexual minority populations in China do not seem to have changed much and lesbians have always been more hidden and invisible than gays [63]. According to a recent national survey on sexuality in China, for women born in the 1980s and 1990s (similar to the age group of the population surveyed in this study), 30.3% and 18.1%, respectively, believe homosexual sex is always wrong, and both rates are higher among men in the same age groups [64].

Furthermore, despite a similar cultural background, the overall participants in our study (mainly from mainland China) scored lower than SGMW in Taiwan using the same RSES [25], which may be related to the differences in policy—same-sex marriage has been legalized in Taiwan since 2019 [65]. However, governments in mainland China, Hong Kong, and Macau do not recognize same-sex marriage, which may further lead the sexual minority populations to feel isolated and fear social condemnation [60]. Despite the fact that homosexuality is no longer classified as a pathology under the *Chinese Classification of Mental Disorders* since 2001 [66], the SGMW in our study reported a more worrisome QOL and mental health disorders than the CHW. Thus, future researchers should not only pay attention to SGM populations' current health status but also fully consider social contexts and living environments before designing adaptive health improvement interventions.

Both correlation and regression analyses in this study showed that the overall QOL as well as the QOL in each domain are significantly associated with depression, anxiety, and self-esteem, which is in line with previous studies. Specifically, depression was shown to negatively impact the QOL in our study, as has been seen in similar pathway analysis studies in other populations [67]. Previous studies have also suggested that people with anxiety report an impaired QOL compared to the general population [68], and both depression and anxiety play an important role in the QOL [69], supporting our study results. On the hypothesized mediation effects of mental health, our results indicated that sexual identity not only directly affects the QOL but also indirectly impacts the QOL through the complete or partial mediating effects of mental health variables, which is consistent with previous studies supporting the mediation effects of mental health [56,70]. The pathways from sexual identity to a reduced QOL shown in our study can be attributed to depression, anxiety, and self-esteem. SGMW reported higher levels of depression and anxiety but lower levels of self-esteem than CHW, which was subsequently related to lower levels of the QOL. To minimize the negative effects of sexual identity on the QOL, mental health may be promoted through interventions targeting SGMW.

Given the prior neglect of the health status of SGM populations and the worrisome levels of the QOL and mental health among SGMW, the participants in this study, our results also have significant implications for professional practice. It is recommended that health and social care professionals proactively attend to the mental health of people belonging to sexual minorities when designing or providing interventions. Notably, we found that the proportions of current smokers and alcohol users were significantly higher in SGMW than in CHW, which is supported by previous studies [71-74]. As shown in the multiple linear regressions, current smokers had poorer levels of the QOL in each domain as well as the overall QOL, which is also consistent with the review conclusions that smoking is negatively associated with the QOL [75]. In addition, the current smoking status was significantly associated with all mental health outcomes in this study, as indicated by higher levels of depression and anxiety and lower levels of self-esteem in current smokers compared to nonsmokers, so the substance use among SGMW in China requires more attention and corresponding improvement programs are needed. Existing evidence shows that smoking cessation is associated with reduced depression and anxiety and could significantly improve the QOL as well [76]. With regard to specific smoking cessation programs for SGM populations, a review summarized that interventions targeting SGM populations are more effective than those targeting the general population [74] and a connection to specific sexual minority communities could protect young sexual minority women in the United States from the risk of smoking [71]; therefore, further research to examine whether sexual minority community engagement can reduce the likelihood of smoking among Chinese SGM populations is warranted.

Although there were no significant associations between having a steady partner and mental health variables in our study, being in a steady relationship was found to be significantly positively associated with the overall QOL, psychological QOL, and social relationship QOL. Such findings that the relationship status could positively affect the QOL are consistent with a published review showing that being in a relationship with a regular partner is a protective factor for health [77], so future interventions that take relationship factors into account may be promising. There is also ample evidence that social support has protective effects on the mental health and QOL of sexual minority populations [9,29,78,79], and perceived support from participants' partners is related to lower risk of depression among SGMW [25]. In contrast, stressors, such as victimization and lower social or family support, may contribute to higher levels of depression among sexual minority populations compared to CH individuals [30]. However, there are few studies examining the comprehensive factors influencing the holistic health of Chinese SGM individuals, and studies only on Chinese SGMW are basically nonexistent; thus, it is strongly recommended that future researchers conduct specific explorations targeting female minority populations.

Sexual minority populations have been found to experience higher rates of physical and mental health disorders and poorer levels of the QOL [2-9,80]; thus, interventions should be designed to help improve their holistic health, especially the QOL. The finding of this study that mental health could mediate the association between sexual identity and the QOL highlights the need for all relevant stakeholders to assess SGM populations' mental health when providing health care or social service. Based on our findings, we also recommend that future interventions aimed at improving the QOL of sexual and gender minorities also emphasize positive mental health components, such as self-esteem, thereby improving their QOL.

To sum up, this is the first study to examine both the QOL and mental health among Chinese adult women with diverse sexual identities and also the first to conduct comparisons between SGMW and CHW, showing that SGMW experience worse levels of the QOL and mental health compared to CHW. Although the associations between the QOL and mental health are well established in general populations, there has been little research on these relationships in people of different gender and sexual orientation identities, so this study is also the first to report the interrelationships between sexual identity and the QOL as well as mental health (depression, anxiety, and self-esteem). Moreover, our study is also the first to explore and confirm the mediating role of mental health in the impact of sexual identity on the QOL.

Limitations

This study had several limitations. First, although the QOL and mental health variables explored in this study are generally investigated as separate constructs, this has certain limitations as the meanings of these concepts are intertwined and may influence each other. Although due to the cross-sectional design of this study, the comprehensive relationships within these concepts could not be adequately explored, further studies using more diverse research methodologies, such as qualitative approaches or longitudinal designs, are needed. Second, most of our study participants were young, and nearly half had no paid employment or were still full-time students; therefore, the results might not be generalizable to the broader Chinese female population. Third, although the sample size was sufficient for comparisons between SGMW as a whole and CHW, subgroup analyses by specific sexual orientation may not be powerful enough. Future studies with larger, more diverse, and representative samples are needed.

Conclusion

This study demonstrated that SGMW have lower levels of the QOL, higher levels of depression and anxiety symptoms, and lower levels of self-esteem than heterosexual women. The overall QOL and QOL in each domain were found to be positively associated with good mental health outcomes, with moderate-to-strong correlations. Multiple linear regressions found that being a sexual or gender minority, current smoker, and woman with no steady partner are associated with a worse overall QOL. Mediation analysis showed that depression, anxiety, and self-esteem play significant mediating roles in the relationship between sexual identity and the QOL in Chinese adult women, affirming the importance of assessing mental health when investigating or improving the QOL. Longitudinal studies and evidence-based intervention programs are further needed, with a particular focus on sexual minority women.

Acknowledgments

We are grateful to all the participants for sharing their experiences with us. We also thank TrueSelf Shenzhen, Period Pride China, Wuhan LGBT Center, the r&B bisexual community, and all those who helped transfer this survey to their friends for their contribution to the questionnaire distribution.

Publication made possible in part by support from the University of Hong Kong (HKU) Libraries Open Access Author Fund sponsored by the HKU Libraries.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparisons of quality of life and mental health among 5 sexual orientation groups.
[DOCX File , 70 KB - publichealth_v9i1e42203_app1.docx]

References

1. Lee JGL, Ylioja T, Lackey M. Identifying Lesbian, Gay, Bisexual, and Transgender Search Terminology: A Systematic Review of Health Systematic Reviews. PLoS One 2016;11(5):e0156210 [FREE Full text] [doi: 10.1371/journal.pone.0156210] [Medline: 27219460]
2. Zheng L, Hart TA, Noor SW, Wen G. Stressors based on sexual orientation and mental health among lesbian, gay, and bisexual individuals in China: minority stress and perceived pressure to get married. Arch Sex Behav 2020 Jul 13;49(5):1769-1782. [doi: 10.1007/s10508-020-01693-z] [Medline: 32285312]
3. Plöderl M, Tremblay P. Mental health of sexual minorities. A systematic review. Int Rev Psychiatry 2015 Nov 09;27(5):367-385. [doi: 10.3109/09540261.2015.1083949] [Medline: 26552495]
4. Mundle G, Mahler L, Bhugra D. Homosexuality and mental health. Int Rev Psychiatry 2015;27(5):355-356. [doi: 10.3109/09540261.2015.1109790] [Medline: 26569633]
5. Frost DM, Lehavot K, Meyer IH. Minority stress and physical health among sexual minority individuals. J Behav Med 2015 Feb;38(1):1-8 [FREE Full text] [doi: 10.1007/s10865-013-9523-8] [Medline: 23864353]
6. Su X, Zhou AN, Li J, Shi L, Huan X, Yan H, et al. Depression, loneliness, and sexual risk-taking among HIV-negative/unknown men who have sex with men in China. Arch Sex Behav 2018 Oct;47(7):1959-1968 [FREE Full text] [doi: 10.1007/s10508-017-1061-y] [Medline: 29147806]
7. Amos R, Manalastas EJ, White R, Bos H, Patalay P. Mental health, social adversity, and health-related outcomes in sexual minority adolescents: a contemporary national cohort study. Lancet Child Adolesc Health 2020 Jan;4(1):36-45. [doi: 10.1016/s2352-4642(19)30339-6]

8. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003 Sep;129(5):674-697 [FREE Full text] [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
9. Fredriksen-Goldsen KI, Emlet CA, Kim H, Muraco A, Erosheva EA, Goldsen J, et al. The physical and mental health of lesbian, gay male, and bisexual (LGB) older adults: the role of key health indicators and risk and protective factors. Gerontologist 2013 Aug;53(4):664-675 [FREE Full text] [doi: 10.1093/geront/gns123] [Medline: 23034470]
10. Ross LE, Salway T, Tarasoff LA, MacKay JM, Hawkins BW, Fehr CP. Prevalence of depression and anxiety among bisexual people compared to gay, lesbian, and heterosexual individuals: a systematic review and meta-analysis. J Sex Res 2018;55(4-5):435-456. [doi: 10.1080/00224499.2017.1387755] [Medline: 29099625]
11. Eres R, Postolovski N, Thielking M, Lim MH. Loneliness, mental health, and social health indicators in LGBTQIA+ Australians. Am J Orthopsychiatry 2021;91(3):358-366. [doi: 10.1037/ort0000531] [Medline: 33315419]
12. Akré E, Anderson A, Stojanovski K, Chung KW, VanKim NA, Chae DH. Depression, anxiety, and alcohol use among LGBTQ+ people during the covid-19 pandemic. Am J Public Health 2021 Sep;111(9):1610-1619. [doi: 10.2105/ajph.2021.306394]
13. Steele LS, Ross LE, Dobinson C, Veldhuizen S, Tinmouth JM. Women's sexual orientation and health: results from a Canadian population-based survey. Women Health 2009;49(5):353-367. [doi: 10.1080/03630240903238685] [Medline: 19851942]
14. Gonzales G, Green J. Medication use among sexual-minority populations for self-reported feelings of depression and anxiety. Psychiatr Serv 2020 Apr 01;71(4):343-354. [doi: 10.1176/appi.ps.201900219] [Medline: 31847734]
15. Gilbey D, Mahfouda S, Ohan J, Lin A, Perry Y. Trajectories of mental health difficulties in young people who are attracted to the same gender: a systematic review. Adolescent Res Rev 2019 Aug 09;5(3):281-293. [doi: 10.1007/s40894-019-00128-8]
16. Chan RC, Operario D, Mak WW. Bisexual individuals are at greater risk of poor mental health than lesbians and gay men: the mediating role of sexual identity stress at multiple levels. J Affect Disord 2020 Jan 01;260:292-301. [doi: 10.1016/j.jad.2019.09.020] [Medline: 31521866]
17. Semlyen J, King M, Varney J, Hagger-Johnson G. Sexual orientation and symptoms of common mental disorder or low wellbeing: combined meta-analysis of 12 UK population health surveys. BMC Psychiatry 2016 Mar 24;16:67 [FREE Full text] [doi: 10.1186/s12888-016-0767-z] [Medline: 27009565]
18. World Health Organization. Depression. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2023-02-02]
19. World Health Organization. Social Determinants of Mental Health. URL: https://www.who.int/publications/i/item/9789241506809 [accessed 2023-02-02]
20. Wu C, Choi EPH, Chau PH. The holistic health status of Chinese homosexual and bisexual adults: a scoping review. Front Public Health 2021 Aug 24;9:710575 [FREE Full text] [doi: 10.3389/fpubh.2021.710575] [Medline: 34504828]
21. Rosenberg M. Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press; 1965:0037-7732.
22. Baumeister RF, Campbell JD, Krueger JI, Vohs KD. Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? Psychol Sci Public Interest 2003 May 24;4(1):1-44. [doi: 10.1111/1529-1006.01431] [Medline: 26151640]
23. Gentile B, Grabe S, Dolan-Pascoe B, Twenge JM, Wells BE, Maitino A. Gender differences in domain-specific self-esteem: a meta-analysis. Rev Gen Psychol 2009 Mar 01;13(1):34-45. [doi: 10.1037/a0013689]
24. Kogler L, Seidel E, Metzler H, Thaler H, Boubela RN, Pruessner JC, et al. Impact of self-esteem and sex on stress reactions. Sci Rep 2017 Dec 08;7(1):17210 [FREE Full text] [doi: 10.1038/s41598-017-17485-w] [Medline: 29222516]
25. Wang Y, Miao N, Chang S. Internalized homophobia, self-esteem, social support and depressive symptoms among sexual and gender minority women in Taiwan: an online survey. J Psychiatr Ment Health Nurs 2021 Aug;28(4):601-610. [doi: 10.1111/jpm.12705] [Medline: 33085822]
26. Wang Y, Chang S, Miao N. Health status and quality of life of middle-aged and older Taiwanese sexual and gender minorities. J Nurs Scholarsh 2021 May;53(3):369-377. [doi: 10.1111/jnu.12640] [Medline: 33675160]
27. World Health Organization. WHOQOL: Measuring Quality of Life. URL: https://www.who.int/tools/whoqol [accessed 2023-02-02]
28. Wong JY, Choi EP, Lo HH, Wong W, Chio JH, Choi AW, et al. Dating violence, quality of life and mental health in sexual minority populations: a path analysis. Qual Life Res 2017 Apr;26(4):959-968. [doi: 10.1007/s11136-016-1415-2] [Medline: 27679496]
29. Kim S, Choi-Kwon S. Physical and mental health related quality of life and their influencing factors on sexual minority women in Korea. Int J Environ Res Public Health 2021 Feb 22;18(4):2115 [FREE Full text] [doi: 10.3390/ijerph18042115] [Medline: 33671565]
30. Argyriou A, Goldsmith KA, Rimes KA. Mediators of the disparities in depression between sexual minority and heterosexual individuals: a systematic review. Arch Sex Behav 2021 Apr;50(3):925-959 [FREE Full text] [doi: 10.1007/s10508-020-01862-0] [Medline: 33689086]
31. Schulman JK, Erickson-Schroth L. Mental health in sexual minority and transgender women. Med Clin North Am 2019 Jul;103(4):723-733. [doi: 10.1016/j.mcna.2019.02.005] [Medline: 31078203]

32. Roberts SJ. Health care of sexual minority women. Nurs Clin North Am 2018 Jun;53(2):227-239. [doi: 10.1016/j.cnur.2018.01.009] [Medline: 29779515]
33. Hughes TL, Johnson TP, Steffen AD, Wilsnack SC, Everett B. Lifetime victimization, hazardous drinking, and depression among heterosexual and sexual minority women. LGBT Health 2014 Sep;1(3):192-203 [FREE Full text] [doi: 10.1089/lgbt.2014.0014] [Medline: 26789712]
34. Yette EM, Ahern J. Health-related quality of life among black sexual minority women. Am J Prev Med 2018 Sep;55(3):281-289 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.037] [Medline: 30122211]
35. Lewis RJ, Ehlke SJ, Shappie AT, Braitman AL, Heron KE. Health disparities among exclusively lesbian, mostly lesbian, and bisexual young women. LGBT Health 2019;6(8):400-408 [FREE Full text] [doi: 10.1089/lgbt.2019.0055] [Medline: 31738644]
36. Petterson LJ, VanderLaan DP, Persson TJ, Vasey PL. The relationship between indicators of depression and anxiety and sexual orientation in Canadian women. Arch Sex Behav 2018 May;47(4):1173-1182. [doi: 10.1007/s10508-017-1099-x] [Medline: 29075928]
37. Hu J, Hu J, Huang G, Zheng X. Life satisfaction, self-esteem, and loneliness among LGB adults and heterosexual adults in China. J Homosex 2016 Aug 05;63(1):72-86. [doi: 10.1080/00918369.2015.1078651] [Medline: 26244408]
38. Beaber TE, Werner PD. The relationship between anxiety and sexual functioning in lesbians and heterosexual women. J Homosex 2009;56(5):639-654. [doi: 10.1080/00918360903005303] [Medline: 19591037]
39. Crawford FW, Wu J, Heimer R. Hidden population size estimation from respondent-driven sampling: a network approach. J Am Stat Assoc 2018;113(522):755-766. [doi: 10.1080/01621459.2017.1285775] [Medline: 30828120]
40. Hequembourg AL, Panagakis C. Maximizing respondent-driven sampling field procedures in the recruitment of sexual minorities for health research. SAGE Open Med 2019;7:2050312119829983 [FREE Full text] [doi: 10.1177/2050312119829983] [Medline: 30800299]
41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
42. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2014 Sep;36(5):539-544 [FREE Full text] [doi: 10.1016/j.genhosppsych.2014.05.021] [Medline: 25023953]
43. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
44. Zhang C, Wang T, Zeng P, Zhao M, Zhang G, Zhai S, et al. Reliability, validity, and measurement invariance of the General Anxiety Disorder scale among Chinese medical university students. Front Psychiatry 2021 May 19;12:648755 [FREE Full text] [doi: 10.3389/fpsyt.2021.648755] [Medline: 34093269]
45. Gong Y, Zhou H, Zhang Y, Zhu X, Wang X, Shen B, et al. Validation of the 7-item Generalized Anxiety Disorder scale (GAD-7) as a screening tool for anxiety among pregnant Chinese women. J Affect Disord 2021 Mar 01;282:98-103. [doi: 10.1016/j.jad.2020.12.129] [Medline: 33401129]
46. Shen Z, Cai T. Disposal to the 8th item of Rosenberg Self Esteem Scale (Chinese version). Chin Ment Health J 2008;22(9):661-663.
47. Martin CR, Thompson DR, Chan DS. An examination of the psychometric properties of the Rosenberg Self-Esteem Scale (RSES) in Chinese acute coronary syndrome (ACS) patients. Psychol Health Med 2006 Nov;11(4):507-521. [doi: 10.1080/13548500500407367] [Medline: 17129926]
48. Xu M, Leung S. Bifactor structure for the categorical Chinese Rosenberg Self-Esteem Scale. Span J Psychol 2016 Oct 11;19:E67. [doi: 10.1017/sjp.2016.66] [Medline: 27724991]
49. Wu C. An examination of the wording effect in the Rosenberg Self-Esteem Scale among culturally Chinese people. J Soc Psychol 2008 Oct;148(5):535-552. [doi: 10.3200/socp.148.5.535-552]
50. World Health Organization. The World Health Organization quality of life (WHOQOL) - BREF, 2012 Revision. URL: https://apps.who.int/iris/handle/10665/77773 [accessed 2023-02-02]
51. Xia P, Li N, Hau K, Liu C, Lu Y. Quality of life of Chinese urban community residents: a psychometric study of the mainland Chinese version of the WHOQOL-BREF. BMC Med Res Methodol 2012 Mar 27;12:37 [FREE Full text] [doi: 10.1186/1471-2288-12-37] [Medline: 22452994]
52. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
53. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/brm.41.4.1149]
54. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge; 1988.
55. Hayes A. In: Little TD, editor. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. Third ed. New York, NY: Guilford Press; 2022.

56. Choi EPH, Wong JYH, Fong DYT. Mental health and health-related quality of life of Chinese college students who were the victims of dating violence. Qual Life Res 2017 Apr;26(4):945-957. [doi: 10.1007/s11136-016-1413-4] [Medline: 27660071]
57. Liu K, Cheng F, Dong H, Dong X, Xu J. Sexual orientation and quality of life of people living with HIV/AIDS in China: evidence from an institutional-based cross-sectional study. Qual Life Res 2022 Jan;31(1):125-134. [doi: 10.1007/s11136-021-02877-3] [Medline: 34028639]
58. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. Annu Rev Psychol 2007 Jan;58:593-614 [FREE Full text] [doi: 10.1146/annurev.psych.58.110405.085542] [Medline: 16968208]
59. Shi X, Xu W, Zheng Y. Heterosexual marital intention: effects of internalized homophobia, homosexual identity, perceived family support, and disclosure among Chinese gay and bisexual men. J Homosex 2020 Dec 03;67(4):452-467. [doi: 10.1080/00918369.2018.1547558] [Medline: 30507292]
60. Xu W, Zheng L, Xu Y, Zheng Y. Internalized homophobia, mental health, sexual behaviors, and outness of gay/bisexual men from Southwest China. Int J Equity Health 2017 Feb 17;16(1):36 [FREE Full text] [doi: 10.1186/s12939-017-0530-1] [Medline: 28222806]
61. Ren Z, Hood RW. Internalized Homophobia Scale for gay Chinese men: conceptualization, factor structure, reliability, and associations with hypothesized correlates. Am J Mens Health 2018 Sep 19;12(5):1297-1306 [FREE Full text] [doi: 10.1177/1557988318768603] [Medline: 29671369]
62. Sun S, Pachankis JE, Li X, Operario D. Addressing minority stress and mental health among men who have sex with men (MSM) in China. Curr HIV/AIDS Rep 2020 Feb;17(1):35-62 [FREE Full text] [doi: 10.1007/s11904-019-00479-w] [Medline: 31950336]
63. Ruan FF. Sex in China: Studies in Sexology in Chinese Culture. New York, NY: Springer-Verlag; 1991.
64. Yu J, Luo W, Xie Y. Sexuality in China: a review and new findings. Chin J Sociol 2022 Aug 03;8(3):293-329. [doi: 10.1177/2057150x221114599]
65. Yen C, Ko N, Huang Y, Chen M, Lin I, Lu W. Preference about laws for the legal recognition of same-sex relationships in Taiwanese people before and after same-sex marriage referenda: a Facebook survey study. Int J Environ Res Public Health 2020 Mar 18;17(6):2000 [FREE Full text] [doi: 10.3390/ijerph17062000] [Medline: 32197386]
66. Chen Y. Chinese Classification of Mental Disorders (CCMD-3): towards integration in international classification. Psychopathology 2002 Jul 17;35(2-3):171-175. [doi: 10.1159/000065140] [Medline: 12145505]
67. Yan F, Tang S, Goldsamt L, Wang H, Chen J, Li X. Interrelationships between intimate partner violence, coping style, depression, and quality of life among the regular female sexual partners of men who have sex with men. J Interpers Violence 2022 Jan;37(1-2):NP651-NP670. [doi: 10.1177/0886260520917519] [Medline: 32390497]
68. Allgulander C, Jørgensen T, Wade A, François C, Despiegel N, Auquier P, et al. Health-related quality of life (HRQOL) among patients with generalised anxiety disorder: evaluation conducted alongside an escitalopram relapse prevention trial. Curr Med Res Opin 2007 Oct;23(10):2543-2549. [doi: 10.1185/030079907X226087] [Medline: 17825130]
69. Lu Y, Jiang Y, Gu L. Using path analysis to investigate the relationships between depression, anxiety, and health-related quality of life among patients with coronary artery disease. Qual Life Res 2019 Oct;28(10):2695-2704. [doi: 10.1007/s11136-019-02207-8] [Medline: 31098799]
70. Choi EPH, Lam CLK, Chin WY. Mental health mediating the relationship between symptom severity and health-related quality of life in patients with lower urinary tract symptoms. Low Urin Tract Symptoms 2016 Sep;8(3):141-149. [doi: 10.1111/luts.12086] [Medline: 27619778]
71. Johns M, Pingel E, Youatt E, Soler J, McClelland S, Bauermeister J. LGBT community, social network characteristics, and smoking behaviors in young sexual minority women. Am J Community Psychol 2013 Sep;52(1-2):141-154 [FREE Full text] [doi: 10.1007/s10464-013-9584-4] [Medline: 23783884]
72. Blosnich JR, Jarrett T, Horn K. Racial and ethnic differences in current use of cigarettes, cigars, and hookahs among lesbian, gay, and bisexual young adults. Nicotine Tob Res 2011 Jun;13(6):487-491. [doi: 10.1093/ntr/ntq261] [Medline: 21330283]
73. Brewster KL, Tillman KH. Sexual orientation and substance use among adolescents and young adults. Am J Public Health 2012 Jun;102(6):1168-1176. [doi: 10.2105/ajph.2011.300261]
74. Berger I, Mooney-Somers J. Smoking cessation programs for lesbian, gay, bisexual, transgender, and intersex people: a content-based systematic review. Nicotine Tob Res 2017 Nov 07;19(12):1408-1417. [doi: 10.1093/ntr/ntw216] [Medline: 27613909]
75. Goldenberg M, Danovitch I, IsHak WW. Quality of life and smoking. Am J Addict 2014;23(6):540-562. [doi: 10.1111/j.1521-0391.2014.12148.x] [Medline: 25255868]
76. Taylor G, McNeill A, Girling A, Farley A, Lindson-Hawley N, Aveyard P. Change in mental health after smoking cessation: systematic review and meta-analysis. BMJ 2014 Feb 13;348:g1151 [FREE Full text] [doi: 10.1136/bmj.g1151] [Medline: 24524926]
77. Johns MM, Liddon N, Jayne PE, Beltran O, Steiner RJ, Morris E. Systematic mapping of relationship-level protective factors and sexual health outcomes among sexual minority youth: the role of peers, parents, partners, and providers. LGBT Health 2018 Jan;5(1):6-32 [FREE Full text] [doi: 10.1089/lgbt.2017.0053] [Medline: 29271692]

78. Wang C, Lin H, Chen M, Ko N, Chang Y, Lin I, et al. Effects of traditional and cyber homophobic bullying in childhood on depression, anxiety, and physical pain in emerging adulthood and the moderating effects of social support among gay and bisexual men in Taiwan. NDT 2018 May;14:1309-1317. [doi: 10.2147/ndt.s164579]
79. Lewis RJ, Milletich RJ, Kelley ML, Woody A. Minority stress, substance use, and intimate partner violence among sexual minority women. Aggress Violent Behav 2012 May;17(3):247-256. [doi: 10.1016/j.avb.2012.02.004]
80. Pakula B. Sexual Identity, Minority Stress, and the Mental Health of Lesbian, Gay, Bisexual, and Heterosexual Canadians. Vancouver: University of British Columbia; 2017.

Abbreviations

CH: cisgender heterosexual
CHW: cisgender heterosexual women
GAD-7: 7-item Generalized Anxiety Disorder scale
PHQ-9: 9-item Patient Health Questionnaire
QOL: quality of life
RSES: Rosenberg Self-Esteem Scale
SGM: sexual and gender minority
SGMW: sexual and gender minority women
WHOQOL-BREF: World Health Organization Quality of Life–abbreviated short version

---

Edited by Y Khader; submitted 26.08.22; peer-reviewed by A AL-Asadi, A King; comments to author 16.12.22; revised version received 06.01.23; accepted 18.01.23; published 22.02.23.

Please cite as:
Wu C, Chau PH, Choi EPH
Quality of Life and Mental Health of Chinese Sexual and Gender Minority Women and Cisgender Heterosexual Women: Cross-sectional Survey and Mediation Analysis
JMIR Public Health Surveill 2023;9:e42203
URL: https://publichealth.jmir.org/2023/1/e42203
doi:10.2196/42203
PMID:36811941

Original Paper

Effects of Bullying on Anxiety, Depression, and Posttraumatic Stress Disorder Among Sexual Minority Youths: Network Analysis

Jiaqi Li[1,2,3,4*], BS; Yu Jin[5*], PhD; Shicun Xu[6,7*], PhD; Amanda Wilson[8], PhD; Chang Chen[1,2,3,4], PhD; Xianyu Luo[5], BSc; Yuhang Liu[5], BSc; Xi Ling[1,2,3,4], LLB; Xi Sun[7], PhD; Yuanyuan Wang[1,2,3,4], PhD, Prof Dr

[1]School of Psychology, South China Normal University, Guangzhou, China
[2]Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, Guangzhou, China
[3]Center for Studies of Psychological Application, South China Normal University, Guangzhou, China
[4]Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, Guangzhou, China
[5]College of Education for the Future, Beijing Normal University, Beijing, China
[6]Northeast Asian Research Center, Jilin University, Changchun, China
[7]Department of Population, Resources and Environment, Jilin University, Changchun, China
[8]Faculty of Health and Life Sciences, De Montfort University, De Montfort, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Yuanyuan Wang, PhD, Prof Dr
School of Psychology
South China Normal University
55 West Zhongshan Avenue
Tianhe District
Guangzhou, 510660
China
Phone: 86 13138157257
Email: angelayuanyuanwang@gmail.com

Abstract

Background: Bullying victimization is highly prevalent among sexual minority youths, particularly in educational settings, negatively affecting their mental health. However, previous studies have scarcely explored the symptomatic relationships among anxiety, depression, and posttraumatic stress disorder (PTSD) among sexual minority youths who experienced bullying on college campuses.

Objective: The objectives of our study were to (1) characterize the anxiety-depression-PTSD network structures of gay or lesbian, bisexuals, and other sexual minority youths previously bullied on college campuses; and (2) compare symptomatic associations in the anxiety-depression-PTSD networks among bullied sexual minority youths and heterosexual youths' groups.

Methods: This cross-sectional study recruited college participants from Jilin Province, China. Data were analyzed using a subset of the data extracted after screening for sexual orientation and history of bullying victimization. Sexual minority youths were then divided into 3 subgroups: gay or lesbian (homosexual), bisexual, and other. Mental health symptom severity was assessed using scales: the 7-item Generalized Anxiety Disorder Scale measuring anxiety, the 9-item Patient Health Questionnaire measuring depression, and the 10-item Trauma Screening Questionnaire measuring PTSD symptoms. Combining the undirected and Bayesian network analyses, the anxiety-depression-PTSD networks were compared among sexual minority youths subgroups, and the difference between heterosexual youths and sexual minority youths was investigated. Chi-square tests were used to compare the difference in categorical variables, while independent-sample *t* tests were run on continuous variables.

Results: In this large-scale sample of 89,342 participants, 12,249 identified as sexual minority youths, of which 1603 (13.1%, 95% CI 12.5%-13.7%) reported being bullied on college campuses in the past year. According to the expected influence (EI) and bridge expected influence (bEI) index, in the global network structure of anxiety, depression, and PTSD, sad mood (EI=1.078, bEI=0.635) and irritability (EI=1.077, bEI=0.954) were identified as central and bridge symptoms; emotional cue reactivity (EI=1.015) was a central symptom of PTSD in this global network. In the anxiety-depression-PTSD Bayesian network, anhedonia had the highest prediction priority for activating other symptoms; and feeling afraid linked symptoms from anxiety to the PTSD community. Compared to their heterosexual counterparts, sexual minority youths exhibited a stronger association between difficulty concentrating and appetite. The “sad mood-appetite” edge was strongest in the gay or lesbian network; the “irritability-exaggerated startle response” edge was strongest in the bisexual network.

Conclusions: For the first time, this study identified the most central and bridge symptoms (sad mood and irritability) within the depression-anxiety-PTSD network of sexual minority youths with past bullying-victim experiences on college campuses. Emotional cue reactivity, anhedonia, and feeling afraid were other vital symptoms in the comorbid network. Symptomatic relationships existed showing heterogeneity in bullied heterosexual youths and sexual minority youth networks, which also was present within the sexual minority youth subgroups. Consequently, refined targeted interventions are required to relieve anxiety, depression, and PTSD symptoms.

(JMIR Public Health Surveill 2023;9:e47233) doi:10.2196/47233

KEYWORDS

sexual minority youths; bullying victimization; anxiety; depression; posttraumatic stress disorder; network analysis; Bayesian network

Introduction

Bullying is defined as the intent or behavior that seeks to harm another person repeatedly, with an imbalance of power [1,2]. It is reported in the literature that the rate of bullying in higher education (on college or university campuses) varies across cultures, ranging from 1.7% to 25.2% [3-5]. Particularly, sexual minority youths, who identify their sexual orientations as nonheterosexual, such as lesbian, gay, or bisexual, are vulnerable to discrimination and bullying [5,6], and have higher odds of experiencing bullying on campus than their heterosexual counterparts [7-10]. A national survey in the United States further reported that sexual minority youths experience a high rate of verbal harassment (76.1%), sexual harassment (53.7%), physical harassment (31.2%), and physical assault (12.5%) [11]. In addition, after being bullied, sexual minority youths may experience adverse physical health and persistent negative effects on their mental health [12].

A considerable amount of literature has reported that sexual minority youths exhibit higher rates of anxiety, depression, and posttraumatic stress disorder (PTSD) than their heterosexual counterparts [13-16]. As a recent meta-analysis noted [17], compared to heterosexual counterparts, lesbian or gay, as well as bisexual individuals yielded a higher risk of anxiety and depression. In addition, sexual minority individuals were found to be vulnerable to adverse experiences, such as being harassed and even bullied in school, increasing the risk for anxiety, depression, and PTSD than those participants who have not experienced adverse life events [18].

Past research has found differences in mental health problems among subgroups of sexual minorities [17,19]. For example, the result from a longitudinal study suggests that bullying victimization is connected to stronger depressive symptoms among bisexual women, and stronger anxiety symptoms among bisexual men [20]. Bisexual individuals are also more likely to report PTSD than their gay or lesbian counterparts [21]. Pansexual women appear to be more open about their sexuality but with a higher consciousness in regard to their sexual orientation-based stigma than bisexual women, which is linked to poorer mental health [22]. Further, asexual men are suggested to have higher scores of depression and psychiatric disturbance than their homosexual and bisexual counterparts [23]. The above studies show a high heterogeneity of mental health outcomes among sexual minority subgroups. However, whether these differences remain consistent among sexual minority youths who are being bullied on college campuses requires further exploration.

Minority stress theory provides a conceptual model that the mental health issues of sexual minorities are partially derived from minority stressors, which are defined as the excess stress to which individuals from stigmatized social categories are exposed to due to their minority position [24], such as higher rates of bullying and discrimination. Furthermore, minority stress can cause biological responses that are related to mental health symptoms [25]. For example, exposure to general minority stress can affect one’s gene expression, which is connected to cardiovascular function [26] and exerts an interactional effect between anxiety, depression, and PTSD [27-29]. Meanwhile, it may increase the dysregulation of the hypothalamus-pituitary-adrenal axis, directly adding to the risk of depression, anxiety, and PTSD [30,31]. Nevertheless, the symptomatic relationship between anxiety, depression, and PTSD among sexual minority youths who are bullied on college campuses lacks investigation.

In recent years, several published studies have adopted network analysis in both clinical psychology or psychiatric fields to explore the relationship between symptoms of various disorders [32-34]. The undirected network analysis depicts the interactions between individual symptoms, with a symptom being termed a “node,” and the partial correlation coefficient between 2 nodes determining their “edge” [35]. According to the expected influence (EI) index and the bridge expected influence (bEI) index, the central and bridge symptoms of the network structure could be identified, respectively [36,37]. Furthermore, in this study, Bayesian network analysis was applied to clarify the directed relationship between symptoms further. Based on the completed partially directed acyclic graph (CPDAG), Bayesian network analysis can unclose the putative directions of the potential causal relationship between 2 nodes [38], and it could help to discover activating paths among symptoms of anxiety, depression, and PTSD network [33].

Given these premises, this study was implemented using undirected and Bayesian network analyses to achieve 2 main objectives: first, to characterize the symptomatic relationships in the anxiety-depression-PTSD network among bullied sexual minority youths and sexual minority youth subgroups; second, to compare the network structures among bullied sexual minority youth subgroups. For aim 1, based on previous results [34,39,40], the researchers hypothesized that depressive and anxious symptoms would be the most central and bridge among all symptoms. For aim 2, considering sexual minority youth subgroups have various prevalence and symptomatic manifestations of anxiety, depression, and PTSD [14,21], the researchers hypothesized that different strengths of symptomatic associations would exist among sexual minority youth subgroups who had disclosed experiences of being bullied. In addition, the researchers aimed to enrich the field, as well as help offer targeted suggestions and interventions for specific symptoms associated with poor mental health.

Methods

Participants and Settings

This study was designed with a cross-sectional approach, recruiting College students from 63 Universities or Colleges in Jilin Province, China, from October to November 2021. Students completed an electronic questionnaire with a Quick Response code, delivered by their teachers with their classmates. The inclusion criteria for recruitment were (1) being 15 years old or older; (2) studying at universities or colleges in Jilin Province, China; (3) understanding the Chinese questionnaire and having the ability to provide informed consent.

Ethical Considerations

Ethical approval for this study was granted by the Ethics Committee of Jilin University (NO20210929 [11 October 2021]), following the 1964 Helsinki Declaration and its amendments in 2013. Electronic informed consent was provided by all participants, with consent that their answers could be applied to any secondary analyses.

Measurements

Participants' sexual orientation was measured by a single-option question with 6 items: "Which of the following better describes your sexual orientation?" Except for heterosexuality, all other sexual orientations are combined to form the sexual minority youth group [41,42], which was then further divided into 3 subgroups, including homosexuality (gay or lesbian), bisexuality, and others (asexuality, pansexuality, and uncertain). Anxiety was measured by the 7-item Generalized Anxiety Disorder Questionnaire [43], depression by the 9-item Patient Health Questionnaire [44], and PTSD by the 10-item Trauma Screening Questionnaire [45]. The experience of being bullied during higher education was measured by a single question: "Have you ever been bullied on campus in the past year?" Details of measurements are presented in the Section S1 in Multimedia Appendix 1, Additional information for networks.

Statistical Analysis

Descriptive Analysis

In total, 117,769 participants were recruited and 89,342 were included according to the criteria and completion of the questionnaire. A descriptive analysis was run on the subdata extracted from screening for sexual orientation and history of bullying victimization. In the data set of those who had been bullied on college campuses in the past, they were divided into groups of heterosexual and sexual minority youth groups. Sociodemographic variables included participants' age, sex, residence, ethnicity, family type, current annual income, and whether they were an only child. The categorical variables in heterosexual or sexual minority youth groups were compared by chi-square tests, while continuous variables were compared by two-tailed independent-sample *t* tests.

Undirected Network Estimation

This study used R programming [46] to structure the undirected network, with the R package "qgraph" (R Foundation for Statistical Computing) to visualize the Graphical Gaussian Model [35]. The R package "mgm" assessed the predictability of a node, referring to the variance of a node that could be explained by all others [47]. The differences between the sexual minority youths and heterosexual network structures and among sexual minority youth subgroups were compared using the R package "Network Comparison Test" with 1000 permutations [48]. For more detail on the method, see Section S2 in Multimedia Appendix 1, Additional information for networks.

Bayesian Network Estimation

In this study, DAG was modelled by the R package "bnlearn," with the hill-climbing algorithm [49]. To further obtain a clear direction among symptoms, the Markov equivalence classes of the DAG were drawn, which can be described uniquely by a CPDAG [50]. In CPDAG, all arrows of undirected edges can be invertible, while all directed edges cannot be converted into undirected edges. Finally, in the determined CPDAG, the direct activating paths in the Bayesian network are more conclusive. More details are shown in Section S2 in Multimedia Appendix 1, Additional information for networks.

Results

Descriptive Statistics

Table 1 presents the sociodemographic characteristics of participants. Within groups, 5063 (6%) heterosexual participants and 1603 (13.1%) sexual minority youth participants reported being bullied on college campuses in the past year. The prevalence of bullying among heterosexual youths was 6% (95% CI 5.9%-6.2%), and 13.1% (95% CI 12.5%-13.7%) among sexual minority youths. Furthermore, among gay participants, the prevalence of bullying was 19.9% (95% CI 17.1%-23.1%), 12.2% (95% CI 9.8%-15.1%) among lesbians, 15.6% (95% CI 13.6%-17.9%) among bisexual males, and 12% (95% CI 11.2%-12.9%) among bisexual females. The odds of being bullied among sexual minority males was 14.4% (95% CI 12.7%-16.2%), and 11.8% (95% CI 10.6%-13%) among sexual minority females.

Table 1. Sociodemographic characteristics of participants who had been bullied on college campuses.

	Heterosexuals (n=5063), n (%)	Sexual minority youths (n=1603), n (%)	Lesbian (n=77), n (%)	Gay (n=142), n (%)	Bisexuals (n=815), n (%)	Others (n=569), n (%)	X^2[a]	Degrees of freedom (*df*)	*P* value
Sex							94.2	1	<.001[b]
Male	2444 (48.3)	552 (34.4)	N/A[c]	142 (100)	176 (21.6)	234 (41.1)			
Female	2619 (51.7)	1051 (65.6)	77 (100)	N/A	639 (78.4)	335 (58.9)			
Residence							44.6	1	<.001[b]
City	2513 (49.6)	949 (59.2)	51 (66.2)	71 (50)	494 (60.6)	333 (58.5)			
Town and county	2550 (50.4)	654 (40.8)	26 (33.8)	71 (50)	321 (39.4)	236 (41.5)			
Ethnicity							1.0	1	.31
Han	4543 (89.7)	1424 (88.8)	74 (96.1)	119 (83.8)	728 (89.3)	503 (88.4)			
Others	520 (10.3)	179 (11.2)	3 (3.9)	23 (16.2)	87 (10.7)	66 (11.6)			
Family type							17.2	2	<.001[b]
Nuclear family	3311 (65.4)	1005 (62.7)	42 (54.5)	86 (60.6)	502 (61.6)	375 (65.9)			
Above 3 generation	951 (18.8)	274 (17.1)	12 (15.6)	26 (18.3)	143 (17.5)	93 (16.3)			
Others	801 (15.8)	324 (20.2)	23 (29.9)	30 (21.1)	170 (20.9)	101 (17.8)			
Current annual income (US $)							20.0	3	<.001[b]
<930	1659 (32.8)	516 (32.2)	27 (35)	47 (33.1)	234 (28.7)	208 (36.6)			
930-2169	1688 (33.3)	463 (28.9)	16 (20.8)	33 (23.2)	247 (30.3)	167 (29.3)			
2170-3565	763 (15.1)	251 (15.6)	14 (18.2)	27 (19)	143 (17.5)	67 (11.8)			
≥3565	953 (18.8)	373 (23.3)	20 (26)	35 (24.6)	191 (23.4)	127 (22.3)			
Only-child status							19.4	1	<.001[b]
Yes	2340 (46.2)	842 (52.5)	42 (54.5)	72 (50.7)	408 (50.1)	320 (56.2)			
No	2723 (53.8)	761 (47.5)	35 (45.5)	70 (49.3)	407 (49.9)	249 (43.8)			
Age (years), mean (SD)	19.54 (1.72)	19.57 (1.67)	19.55 (1.66)	19.74 (1.70)	19.51 (1.63)	19.63 (1.73)	–0.7[d]	6659	.49
GAD-7[e], mean (SD)	6.25 (474)	7.82 (5.44)	8.06 (5.90)	7.61 (5.30)	7.93 (5.15)	7.69 (5.80)	–11.2[d]	6664	<.001[b]
PHQ-9[f], mean (SD)	8.06 (5.30)	10.15 (6.33)	10.74 (7.52)	9.45 (5.42)	10.16 (5.96)	10.23 (6.85)	–13.1[d]	6664	<.001[b]
PTSD-10[g], mean (SD)	4.88 (3.16)	5.75 (3.12)	5.94 (3.38)	5.58 (3.20)	5.91 (2.97)	5.55 (3.27)	–9.7[d]	6664	<.001[b]

[a] X^2: used to compare differences between heterosexual and sexual minority youth groups among specified demographic variables.

[b] There was a significant difference existing with *P*<.001.

[c] N/A: Not applicable.

[d] T: used to compare differences between heterosexual and sexual minority youth groups among specified demographic variables.

[e] GAD-7: the 7-item Generalized Anxiety Disorders Scale.

[f] PHQ-9: the 9-item Patient Health Questionnaire.

[g] PTSD-10: measured by the 10-item Trauma Screening Questionnaire.

Compared to their heterosexual counterparts, more sexual minority females, and more sexual minority youths living in urban residences reported being bullied on the campuses. In addition, experiences of being bullied were more likely to occur among participants from nuclear families, with lower annual family income, or among those who had siblings. Bullied sexual minority youths also reported higher rates of anxiety, depression, and PTSD symptoms than bullied heterosexual youths.

Undirected Network Structure

As seen in Figure 1, according to the centrality index EI, GAD2 (Generalized Anxiety Disorders Scale; “control worry,” EI=1.134) had the highest EI, followed by PHQ4 (Patient Health Questionnaire; “energy,” EI=1.116), GAD3 (“worry a lot,” EI=1.085), PHQ2 (“sad mood,” EI=1.078), GAD6 (“irritability,” EI=1.077), and PTSD4 (“energy,” EI=1.015). These 6 symptoms played the most central role in the global network of anxiety, depression, and PTSD symptoms among all bullied sexual minority youths. The central network structures of the 3 sexual minority youth subgroups and their centrality index EI are shown in Table S1 and Figures S1, S2, S3, and S4 in Multimedia Appendix 1, Additional information for networks.

As seen in Figure 2, according to bEI, GAD6 ("irritability," bEI=0.954), PHQ8 ("motor," bEI=0.883), GAD7 ("feeling afraid," bEI=0.758), GAD5 ("restless," bEI=0.718), GAD1 ("nervous," bEI=0.694), and PHQ2 ("sad mood," bEI=0.635) had crucial bridging roles in connecting to the other symptoms of anxiety, depression, or PTSD, which may possibly transform from 1 community to another. The bridge network structures of the 3 sexual minority youth subgroups and their bridge bEI index are shown in Table S1 and Figures S5, S6, S7, and S8 in Multimedia Appendix 1, Additional information for networks.

Figure 1. The global network structure of central symptoms among all bullied sexual minority youths (N=1603). GAD: the 7-item Generalized Anxiety Disorders Scale; LGB: homosexuality (gay or lesbian), bisexuality and other sexual minority youths; PHQ: the 9-item Patient Health Questionnaire; PTSD: posttraumatic stress disorder measured by the 10-item Trauma Screening Questionnaire.

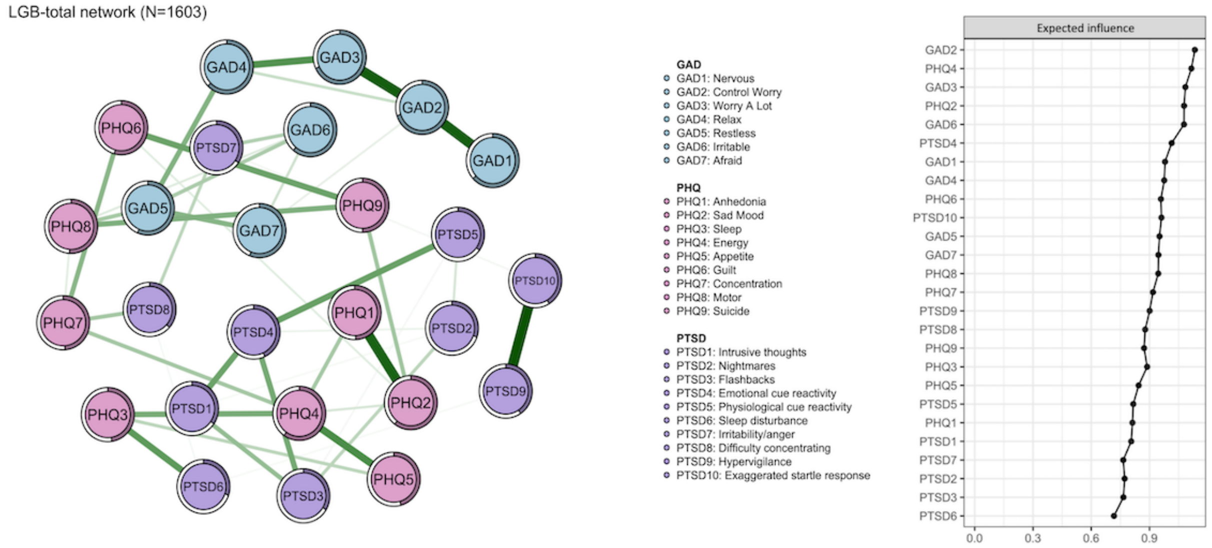
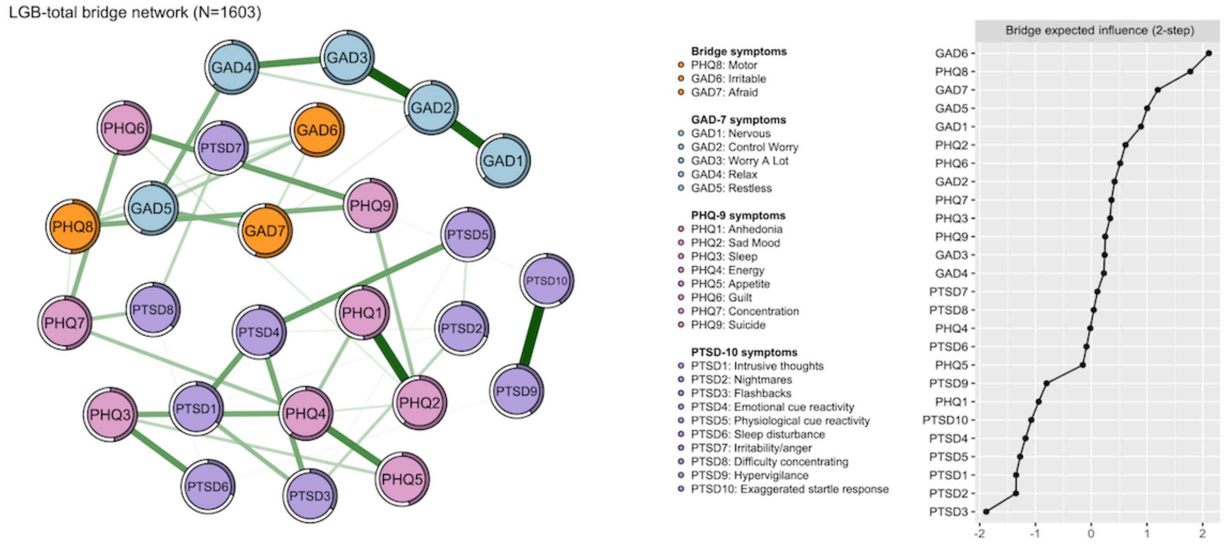


Figure 2. The global network structure of bridge symptoms among all bullied sexual minority youths (N=1603). GAD: the 7-item Generalized Anxiety Disorders Scale; LGB: homosexuality (gay or lesbian), bisexuality and other sexual minority youths; PHQ: the 9-item Patient Health Questionnaire; PTSD: posttraumatic stress disorder measured by the 10-item Trauma Screening Questionnaire.



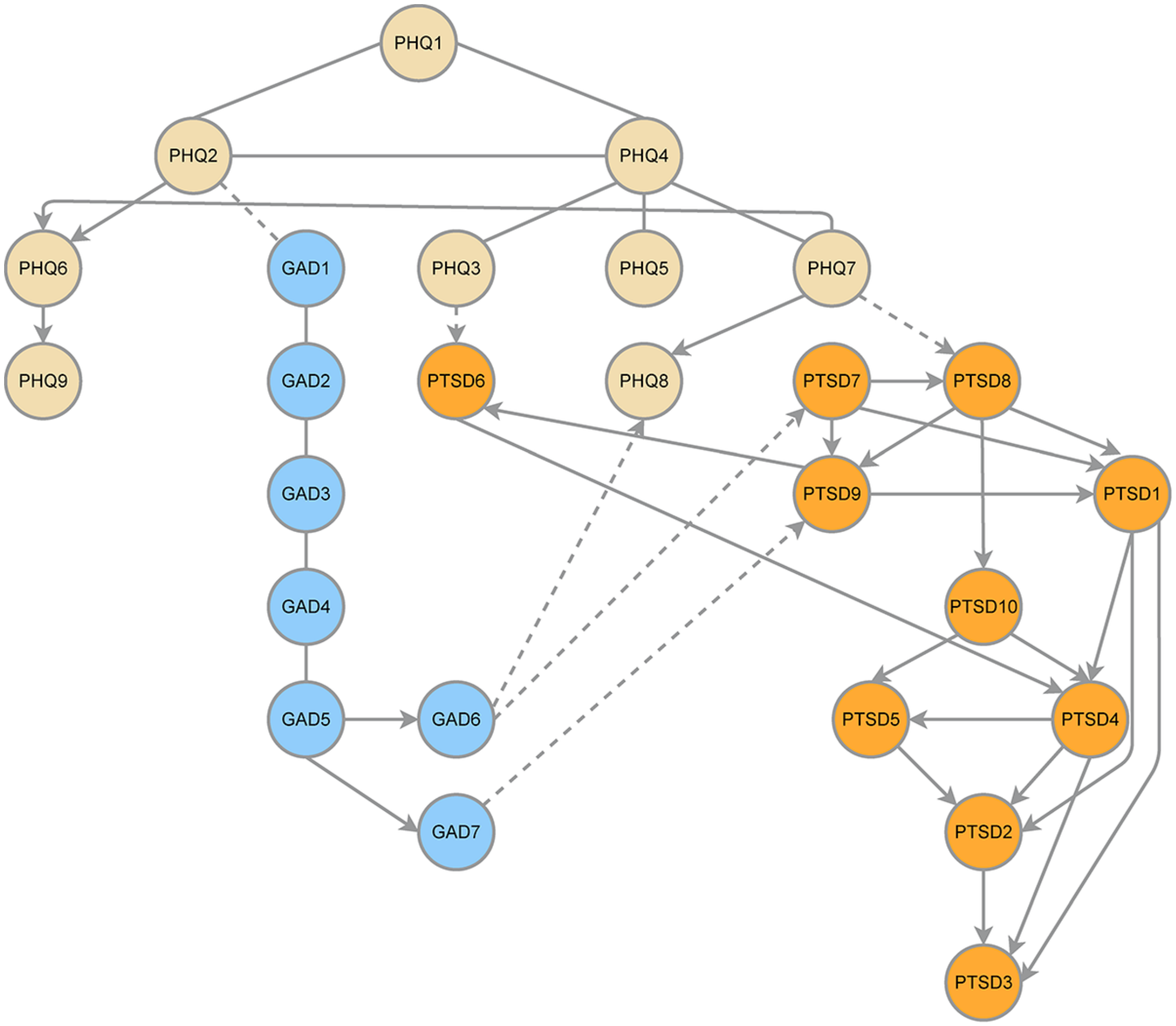
The average predictability of nodes was 48.67%, which means that, on average, the variance of 48.67% per node could be explained by its neighbors. In addition, all networks show good stability in this study (see in Figures S9, S10, S11, and S12 in Multimedia Appendix 1, Additional information for networks).

Bayesian Network Structure

Figure 3 shows the CPDAG results, indicating that PHQ1 ("anhedonia") had the highest predictive priority for other symptoms occurring in the anxiety-depression-PTSD network. PHQ2 ("sad mood"), GAD6 ("irritability"), and GAD7 ("feeling afraid") were identified as 3 crucial bridge symptoms, whose downstream symptoms had more branches and could activate more pathways. In addition, there were reversible connections between PHQ1 ("anhedonia"), PHQ2 ("sad mood"), and PHQ4 ("appetite"). Whereas, the edges between PHQ2 ("sad mood") and PHQ6 ("guilt"), between GAD6 ("irritability"), PHQ8 ("motor"), and PTSD7 ("irritability or anger"), and between

GAD7 (“feeling afraid”) to PTSD9 (“hypervigilance”) were unidirectional and irreversible. These results indicate the most likely direction from both depression and anxiety symptoms to PTSD symptoms, rather than vice versa. More results of CPDAG can be found in Figure S13 in Multimedia Appendix 1, Additional information for networks.

Figure 3. The Bayesian network of anxiety, depression, and posttraumatic stress disorder among bullied sexual minority youths, based on a completed partially directed acyclic graph. GAD: the 7-item Generalized Anxiety Disorders Scale; PHQ; the 9-item Patient Health Questionnaire; PTSD: posttraumatic stress disorder measured by the 10-item Trauma Screening Questionnaire. The edges without directions refer to reversiblity. The solid edges with direction refer to irreversiblity. The dotted edges with direction refer to irreversibility and arrowing to a node in another symptom community.

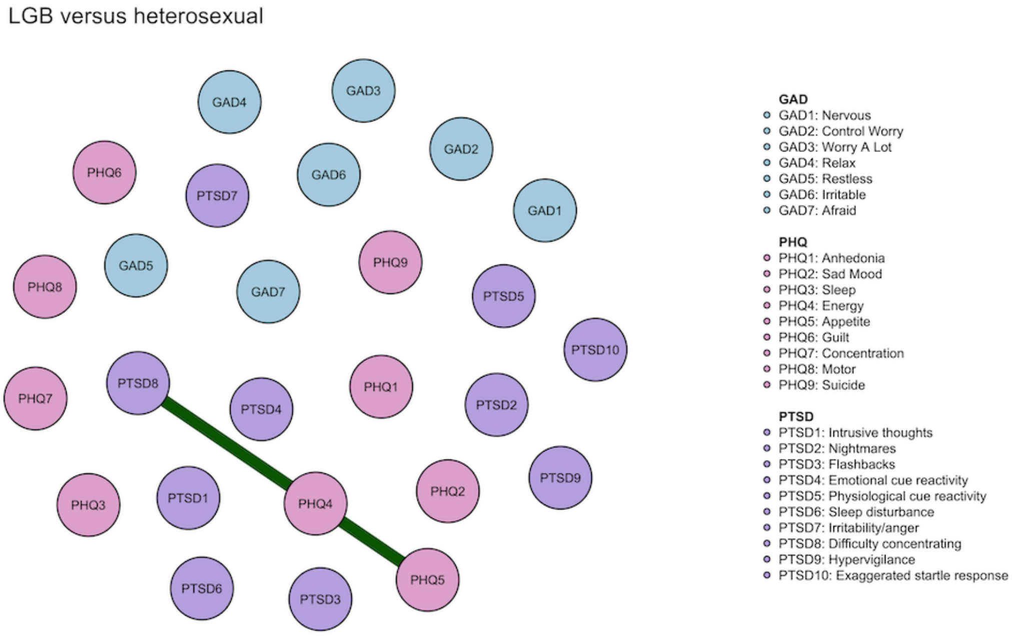


Network Comparison Structures

Figure 4 shows the results when comparing the network structures between the sexual minority youths and heterosexual groups. The edge of PTSD8 (“difficulty concentrating”)-PHQ5 (“appetite”) was significantly stronger in the sexual minority youth network structure than in the heterosexual.

Among the 3 subgroups of sexual minority youths, there was no significant difference in edges between the gay or lesbian network and the other sexual minority youth networks. However, the edge of PHQ2 (“sad mood”)-PHQ5 (“appetite”) was strongest in the gay or lesbian network and the edge of PTSD7 (“irritability”)-PTSD10 (“exaggerated startle response”) was strongest in the bisexual sexual minority youth networks (Figure S14 in Multimedia Appendix 1, Additional information for networks).

Figure 4. The comparison of network structure between sexual minority and heterosexual youths. GAD: the 7-item Generalized Anxiety Disorders Scale; PHQ: the 9-item Patient Health Questionnaire; PTSD: posttraumatic stress disorder measured by the 10-item Trauma Screening Questionnaire.


LGB versus heterosexual

Discussion

Principal Findings

This is the first study based on a large-scale sample to investigate the symptomatic relationship of anxiety, depression, and PTSD among sexual minority youths bullied on college campuses in the past. The results partly supported hypothesis 1: sad mood and irritability were identified as the most frequently crucial activating symptoms of poor mental health, where emotional cue reactivity was the most central symptom of the PTSD community. In the Bayesian network, anhedonia had the highest predictive priority to activate the occurrence of all other symptoms; feeling afraid played a bridging role in the anxiety-depression-PTSD network. The results supported hypothesis 2: compared to heterosexuals, sexual minority youths exhibited a stronger association between difficulty concentrating and appetite. The "sad mood-appetite" edge was strongest in the gay or lesbian network, and the "irritability-exaggerated startle response" edge was strongest in the bisexual network.

Sexual minority youths has been frequently reported to experience discrimination, prejudice, and stigma-bias because of their sexual orientation [6,51], associated with bullying victimization of sexual minority youths [52,53]. Our study further reported a high prevalence of being bullied on college campuses among sexual minority youths, which was 13.3% (95% CI 12.5%-13.7%, n=1603). Consistent with previous findings, exposure to bullying on college campuses was found to be related to the onset of mental health problems among sexual minority youths [53]. As a meta-analysis supported, individuals who experience bullying victimization had increased odds of depression, anxiety, and PTSD [54].

Sad mood and irritability were the most crucial symptoms (both central and bridge) in all anxiety-depression-PTSD network models. Previous network analyses have consistently identified sad mood as the central and bridge symptom in the structure of the anxiety-depression network [36,55]. This result may be explained by the external stressors that sexual minority youths face, making them feel sad [51,56]. According to a US national survey on youths' risk behavior, high rates of in-school bullying victimization for homosexuals (28.22%, n=91) and bisexuals (34.03%, n=318) occurred because of external discrimination [7]. When under stressful conditions, if those bullied students experience insufficient support from teachers, they could feel abandoned and hopeless [57]. In terms of irritability, 2 prior studies also noted that irritability had a high betweenness index (another bridge centrality index) [34], and was identified as one of the central symptoms in the depression-anxiety-PTSD network [58]. Research has indicated that, once bullied, an inability to find a coping strategy or resist bullying behaviors can make individuals feel frustrated, which may trap them in rumination anger [9]. Moreover, traumatic memories of bullying victimization could generate severe distress and irritability as well [59]. Our results showed irritability also had a high predictive priority of the occurrence of PTSD symptoms. Findings indicate that for those who do withstand cumulative bullying victimization, their subjective aggression can be triggered, potentially increasing their irritable mood and PTSD symptoms simultaneously [60]. Notably, emotional cue reactivity (in this study referring to feeling upset by reminders of the events) was the most central symptom in the PTSD network. This may be explained by the influences of autobiographical memories (a mixed memory of complicated life events) [61]. Experiences of being bullied incur negative emotional cues, such as insults and physical aggression [62], and these negative memories would, in turn, cause further distress.

In the Bayesian network, anhedonia had the highest predictive priority of activating other symptoms. It parallels a previous network analysis that identified anhedonia as a central symptom in the anxiety-depression network [36]. Previous findings reported that bullying behaviors could lead to poorer mental health for bullying victims [63]. After being bullied, sexual minority youths may form a negative schema about an institution, which influences their interaction with others, decreasing their interest in participating in activities with their peers [64]. Another possible explanation could be a lack of support from teachers and families after exposure to being bullied, which further hinders sexual minority youths' mental health [65] and also makes it harder for them to perceive pleasure [66]. Additionally, being bullied may negatively influence the release of dopamine, a chemical produced by neuronal receptors that relate to feeling good, in the case of bullying, this causes a lack of pleasure [67]. Feeling afraid also played a significant bridging role in the Bayesian network, pointing to the occurrence of hypervigilance in the PTSD community. On the one hand, past bullying victimization could lead sexual minority youths to be afraid of reoccurring bullying [68]; on the other hand, without efficient support from others, exposing perpetrators of bullying could trigger their fear of retaliation [69]. In addition, feeling afraid is linked to the symptoms of PTSD in this study. As early findings stated, those bullied could incur idiopathic nightmares and sleep problems, leading to hypervigilance [70]. From the neuropsychological perspective, fear-related threats increase the steady-state visual evoked amplitudes, thus heightening individuals' vigilance [71]. Based on this fear of repeated bullying, sexual minority youths' visual attention to possible danger may present with over-heightened reactions.

Compared to the heterosexual network structure, difficulty concentrating was found to be more closely associated with appetite in the sexual minority youth network structure. As previous findings stated, sexual minority youths were more vulnerable to weight-based bullying by peers than heterosexual youths [72,73]. Moreover, a motivation to avoid or prevent bullying may be a key driver for those who seek treatment for obesity [74]. Thus, the ideal body image and weight bias may result in more sexual minority youths choosing to diet in order to lose weight or engage in harmful weight-control strategies, such as purging, to maintain a satisfactory appearance [75]. It should be noted that over-preoccupation with body shape might impair concentration in other areas, supported by a previous study in which those who experienced bullying reported that they perceived their attention was distracted due to the fear associated with the experience, which led to difficulty concentrating when learning and during academic achievement, further hindering their state of mental well-being [76].

Differently, compared to bullied bisexual youths, sad mood had a stronger correlation with appetite among bullied gay or lesbian youths. Researchers consistently report more disordered eating behaviors in gay or lesbian youths than in bisexuals [77]. In addition, gay or lesbian youths were reported to have more body dissatisfaction, driven by an ideal body shape [78,79]. Negative mood also relates to subjective appetite, causing poor appetite or overeating [80]. Since discrimination and bullying incur a sad mood for gay or lesbian youths, they may choose binge eating to elevate their mood momentarily [81]. Furthermore, compared to other sexual minority youths, bisexual youths exhibited a stronger correlation between irritability and exaggerated startle response. This may be due to bisexual youths experiencing dual discrimination from homosexual and heterosexual communities [82]. Since bullying shapes a threatening context for bisexual youths, they may become irritable and exhibit a high startle reactivity [83]. Meanwhile, higher irritability can alter attention bias, with decreased functional connectivity between the left inferior frontal gyrus and a periaqueductal grey area, which may increase the response to frustrating stimuli [84].

Limitations

Several limitations should be noted in this study. First, as a cross-sectional study, it is hard to explore the dynamic process of psychopathological networks [85], thus, further longitudinal research is required to obtain causal conclusions. Second, the replicability and generalizability of network structures are still limited across various samples [86]. In the future, neuroimaging studies are required to explore the organic effects on the brain. Third, measurements were limited in several aspects: (1) the minority stressors, such as lacking peers or family support, were not directly measured, and the results could be biased because of other potential influences. (2) There may be a part of the youths who have hesitancy toward or concealment of being identified as sexual minority youths for fear of being exposed in the context. (3) In addition, the interaction of gender identity and sexual orientation should be considered in future investigations. For example, a transgender man attracted to a cisgender man could also consider himself gay. Further studies should be conducted covering diverse gender identities in detail. (4) Because the experience of being bullied on college campuses in the past was measured by a self-report as well, in which the standard to determine the conceptualization of bullying varies across individuals, it was inevitable to include samples with different severities of bullying victimization. Future research should provide more detailed standards to define bullying and explore the underlying mechanism between various types or levels of bullying behaviors and the impact this has on the mental health of sexual minority populations. Bullying off-campus perpetrated by peers should be measured, and its impacts must be explored. (5) The bias of recall errors or social desirability effects is unavoidable for data from subjective reporting scales. Finally, other psychiatric disorders, such as obsessive-compulsive disorder or eating disorders, were not measured directly, which may affect the results.

Conclusions

This cross-sectional study based on a large-scale sample, combined the undirected network analysis and Bayesian network analysis, for the first time, identifying the most central and bridge symptoms (sad mood and irritability), as well as the central role of emotional cue reactivity within the depression-anxiety-PTSD network of sexual minority youths who were bullied on college campuses in the past. CPDAG results also indicated the vital roles of anhedonia and feeling afraid in the global Bayesian network in this study. In addition, the comparisons of networks supported three findings: (1) compared to heterosexuals, the edge of "difficulty concentrating" was stronger than in the global sexual minority youth network, (2) the edge of "sad mood-appetite" was strongest in the gay or lesbian network, and (3) the edge of "irritability-exaggerated startle response" was strongest in the bisexual network. These findings should be considered when offering targeted support for improving the physical health and mental well-being of sexual minority youths. Consequently, refined targeted interventions could be formulated [6] to relieve the symptoms of anxiety, depression, and PTSD comorbidity among sexual minority youths.

Acknowledgments

The conception and design of this study was done by YW and SX. SX and XS did the data collection. Data quality control was done by SX. Data analysis and all figures were done by YJ, CC, XLuo, YL, and XLing. This study's write-up was performed by JL, YJ, and SX. Critical comments were made by YW and AW. This study was supervised by YW and SX. We thank all the participants and staff involved in this study. This study was supported by the National Natural Science Foundation of China (82201708).

Data Availability

The data set for this specific paper is available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional information for networks.
[DOCX File , 7006 KB - publichealth_v9i1e47233_app1.docx]

References

1. Smith PK. Bullying: definition, types, causes, consequences and intervention. Soc Personal Psychol Compass 2016;10(9):519-532. [doi: 10.1111/spc3.12266]
2. Younan B. A systematic review of bullying definitions: how definition and format affect study outcome. J Aggress Confl Peace Res 2019;11(2):109-115. [doi: 10.1108/jacpr-02-2018-0347]
3. Lund EM, Ross SW. Bullying perpetration, victimization, and demographic differences in college students: a review of the literature. Trauma Violence Abuse 2017 Jul;18(3):348-360. [doi: 10.1177/1524838015620818] [Medline: 26759417]
4. Pörhölä M, Cvancara K, Kaal E, Kunttu K, Tampere K, Torres MB. Bullying in university between peers and by personnel: cultural variation in prevalence, forms, and gender differences in four countries. Soc Psychol Educ 2019;23(1):143-169 [FREE Full text] [doi: 10.1007/s11218-019-09523-4]
5. Sinkkonen HM, Puhakka H, Meriläinen M. Bullying at a university: students' experiences of bullying. Studies in Higher Education 2012;39(1):153-165. [doi: 10.1080/03075079.2011.649726]
6. Smith TE, Bauerband LA, Aguayo D, McCall CS, Huang FL, Reinke WM, et al. School bullying and gender minority youth: victimization experiences and perceived prevalence. School Psych Rev 2022:1-14. [doi: 10.1080/2372966x.2021.2002123]
7. Angoff HD, Barnhart WR. Bullying and cyberbullying among LGBQ and heterosexual youth from an intersectional perspective: findings from the 2017 national youth risk behavior survey. J Sch Violence 2021;20(3):274-286. [doi: 10.1080/15388220.2021.1879099]
8. Wensley K, Campbell M. Heterosexual and nonheterosexual young university students' involvement in traditional and cyber forms of bullying. Cyberpsychol Behav Soc Netw 2012;15(12):649-654. [doi: 10.1089/cyber.2012.0132] [Medline: 23078337]
9. Zsila Á, Urbán R, Demetrovics Z. Anger rumination and unjust world beliefs moderate the association between cyberbullying victimization and psychiatric symptoms. Psychiatry Res 2018;268:432-440 [FREE Full text] [doi: 10.1016/j.psychres.2018.08.001] [Medline: 30130710]
10. Kann L, McManus T, Harris WA, Shanklin SL, Flint KH, Queen B, et al. Youth risk behavior surveillance - United States, 2017. MMWR Surveill Summ 2018;67(8):1-114 [FREE Full text] [doi: 10.15585/mmwr.ss6708a1] [Medline: 29902162]
11. Kosciw JG, Clark CM, Menard L. The 2021 National School Climate Survey: The experiences of LGBTQ+ youth in our nation's schools. New York: GLSEN; 2022. URL: https://healtheducationresources.unesco.org/library/documents/2021-national-school-climate-survey-experiences-lgbtq-youth-our-nations-schools [accessed 2023-09-22]
12. Russell ST, Fish JN. Mental health in lesbian, gay, bisexual, and transgender (LGBT) youth. Annu Rev Clin Psychol 2016;12:465-487 [FREE Full text] [doi: 10.1146/annurev-clinpsy-021815-093153] [Medline: 26772206]
13. Mittleman J. Sexual minority bullying and mental health from early childhood through adolescence. J Adolesc Health 2019;64(2):172-178. [doi: 10.1016/j.jadohealth.2018.08.020] [Medline: 30392862]

14. Ross LE, Salway T, Tarasoff LA, MacKay JM, Hawkins BW, Fehr CP. Prevalence of depression and anxiety among bisexual people compared to gay, lesbian, and heterosexual individuals: a systematic review and meta-analysis. J Sex Res 2018;55(4-5):435-456 [FREE Full text] [doi: 10.1080/00224499.2017.1387755] [Medline: 29099625]
15. Poteat VP, O'Brien MD, Rosenbach SB, Finch EK, Calzo JP. Depression, anxiety, and interest in mental health resources in school-based gender-sexuality alliances: implications for sexual and gender minority youth health promotion. Prev Sci 2021;22(2):237-246 [FREE Full text] [doi: 10.1007/s11121-020-01196-2] [Medline: 33410118]
16. Burton CM, Marshal MP, Chisolm DJ, Sucato GS, Friedman MS. Sexual minority-related victimization as a mediator of mental health disparities in sexual minority youth: a longitudinal analysis. J Youth Adolesc 2013;42(3):394-402 [FREE Full text] [doi: 10.1007/s10964-012-9901-5] [Medline: 23292751]
17. Wittgens C, Fischer MM, Buspavanich P, Theobald S, Schweizer K, Trautmann S. Mental health in people with minority sexual orientations: A meta-analysis of population-based studies. Acta Psychiatr Scand 2022;145(4):357-372 [FREE Full text] [doi: 10.1111/acps.13405] [Medline: 35090051]
18. Schnarrs PW, Stone AL, Bond MA, Salcido R, Dorri AA, Nemeroff CB. Development and psychometric properties of the sexual and gender minority adverse childhood experiences (SGM-ACEs): Effect on sexual and gender minority adult mental health. Child Abuse Negl 2022;127:105570 [FREE Full text] [doi: 10.1016/j.chiabu.2022.105570] [Medline: 35231816]
19. Borgogna NC, McDermott RC, Aita SL, Kridel MM. Anxiety and depression across gender and sexual minorities: implications for transgender, gender nonconforming, pansexual, demisexual, asexual, queer, and questioning individuals. Psychol Sex Orientat Gend Divers 2019;6(1):54-63. [doi: 10.1037/sgd0000306]
20. Katz-Wise SL, Rosario M, Calzo JP, Scherer EA, Sarda V, Austin SB. Associations of timing of sexual orientation developmental milestones and other sexual minority stressors with internalizing mental health symptoms among sexual minority young adults. Arch Sex Behav 2017;46(5):1441-1452 [FREE Full text] [doi: 10.1007/s10508-017-0964-y] [Medline: 28271349]
21. Carey FR, LeardMann CA, Lehavot K, Jacobson IG, Kolaja CA, Stander VA, et al. Health disparities among lesbian, gay, and bisexual service members and veterans. Am J Prev Med 2022;63(4):521-531 [FREE Full text] [doi: 10.1016/j.amepre.2022.04.034] [Medline: 35794031]
22. Morandini J, Strudwick J, Menzies R, Dar-Nimrod I. Differences between Australian bisexual and pansexual women: an assessment of minority stressors and psychological outcomes. Psychol Sex 2022;14(1):233-251. [doi: 10.1080/19419899.2022.2100717]
23. Yule MA, Brotto LA, Gorzalka BB. Mental health and interpersonal functioning in self-identified asexual men and women. Psychol Sex 2013;4(2):136-151. [doi: 10.1080/19419899.2013.774162]
24. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003;129(5):674-697 [FREE Full text] [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
25. Flentje A, Heck NC, Brennan JM, Meyer IH. The relationship between minority stress and biological outcomes: a systematic review. J Behav Med 2020;43(5):673-694 [FREE Full text] [doi: 10.1007/s10865-019-00120-6] [Medline: 31863268]
26. Flentje A, Kober KM, Carrico AW, Neilands TB, Flowers E, Heck NC, et al. Minority stress and leukocyte gene expression in sexual minority men living with treated HIV infection. Brain Behav Immun 2018;70:335-345 [FREE Full text] [doi: 10.1016/j.bbi.2018.03.016] [Medline: 29548994]
27. Dedert EA, Calhoun PS, Watkins LL, Sherwood A, Beckham JC. Posttraumatic stress disorder, cardiovascular, and metabolic disease: a review of the evidence. Ann Behav Med 2010;39(1):61-78 [FREE Full text] [doi: 10.1007/s12160-010-9165-9] [Medline: 20174903]
28. Gustad LT, Bjerkeset O, Strand LB, Janszky I, Salvesen Ø, Dalen H. Cardiac function associated with previous, current and repeated depression and anxiety symptoms in a healthy population: the HUNT study. Open Heart 2016;3(1):e000363 [FREE Full text] [doi: 10.1136/openhrt-2015-000363] [Medline: 26925243]
29. Johnson AK, Hayes SN, Sawchuk C, Johnson MP, Best PJ, Gulati R, et al. Analysis of posttraumatic stress disorder, depression, anxiety, and resiliency within the unique population of spontaneous coronary artery dissection survivors. J Am Heart Assoc 2020;9(9):e014372 [FREE Full text] [doi: 10.1161/JAHA.119.014372] [Medline: 32342736]
30. Almeida FB, Pinna G, Barros HMT. The role of HPA axis and allopregnanolone on the neurobiology of major depressive disorders and PTSD. Int J Mol Sci 2021;22(11):5495 [FREE Full text] [doi: 10.3390/ijms22115495] [Medline: 34071053]
31. Hinds JA, Sanchez ER. The role of the Hypothalamus–Pituitary–Adrenal (HPA) axis in test-induced anxiety: assessments, physiological responses, and molecular details. Stresses 2022;2(1):146-155. [doi: 10.3390/stresses2010011]
32. Amore M, Murri MB, Calcagno P, Rocca P, Rossi A, Aguglia E, et al. The association between insight and depressive symptoms in schizophrenia: Undirected and Bayesian network analyses. Eur Psychiatry 2020;63(1):1-21 [FREE Full text] [doi: 10.1192/j.eurpsy.2020.45] [Medline: 32372731]
33. McNally RJ, Heeren A, Robinaugh DJ. A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. Eur J Psychotraumatol 2017;8(sup3):1341276 [FREE Full text] [doi: 10.1080/20008198.2017.1341276] [Medline: 29038690]

34. Price M, Legrand AC, Brier ZMF, Hébert-Dufresne L. The symptoms at the center: examining the comorbidity of posttraumatic stress disorder, generalized anxiety disorder, and depression with network analysis. J Psychiatr Res 2019;109:52-58 [FREE Full text] [doi: 10.1016/j.jpsychires.2018.11.016] [Medline: 30502492]
35. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: a tutorial paper. Behav Res Methods 2018;50(1):195-212 [FREE Full text] [doi: 10.3758/s13428-017-0862-1] [Medline: 28342071]
36. Beard C, Millner AJ, Forgeard MJC, Fried EI, Hsu KJ, Treadway MT, et al. Network analysis of depression and anxiety symptom relationships in a psychiatric sample. Psychol Med 2016;46(16):3359-3369 [FREE Full text] [doi: 10.1017/S0033291716002300] [Medline: 27623748]
37. Jones PJ, Ma R, McNally RJ. Bridge centrality: a network approach to understanding comorbidity. Multivariate Behav Res 2021;56(2):353-367. [doi: 10.1080/00273171.2019.1614898] [Medline: 31179765]
38. Pearl J, Russell S. Bayesian networks. 2011. URL: https://escholarship.org/uc/item/53n4f34m [accessed 2023-09-22]
39. Contreras A, Nieto I, Valiente C, Espinosa R, Vazquez C. The study of psychopathology from the network analysis perspective: a systematic review. Psychother Psychosom 2019;88(2):71-83 [FREE Full text] [doi: 10.1159/000497425] [Medline: 30889609]
40. Jin Y, Xu S, Hu Z, Li J, Li H, Wang X, et al. Co-occurrence of PTSD and affective symptoms in a large sample with childhood trauma subtypes: a network analysis. Front Public Health 2023;11:1093687 [FREE Full text] [doi: 10.3389/fpubh.2023.1093687] [Medline: 36960378]
41. Rice K. Pansexuality. Int Encyclopedia Human Sexuality 2015:861-1042. [doi: 10.1002/9781118896877.wbiehs328]
42. Hsieh N, Shuster SM. Health and health care of sexual and gender minorities. J Health Soc Behav 2021;62(3):318-333. [doi: 10.1177/00221465211016436] [Medline: 34528481]
43. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097 [FREE Full text] [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
44. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
45. Brewin CR, Rose S, Andrews B, Green J, Tata P, McEvedy C, et al. Brief screening instrument for post-traumatic stress disorder. Br J Psychiatry 2018;181(2):158-162 [FREE Full text] [doi: 10.1192/bjp.181.2.158]
46. Ihaka R, Gentleman R. R: a language for data analysis and graphics. J Comput Graph Stat 1996;5(3):299-314. [doi: 10.1080/10618600.1996.10474713]
47. Haslbeck J, Waldorp LJ. mgm: estimating time-varying mixed graphical models in high-dimensional data. J Stat Soft 2020;93(8):1-8. [doi: 10.18637/jss.v093.i08]
48. van Borkulo CD, van Bork R, Boschloo L, Kossakowski JJ, Tio P, Schoevers RA, et al. Comparing network structures on three aspects: a permutation test. Psychol Methods 2022. [doi: 10.1037/met0000476] [Medline: 35404628]
49. Scutari M. Learning bayesian networks with the bnlearn R package. J Stat Softw 2010;35(3):1-22. [doi: 10.48550/arXiv.0908.3817]
50. Castelletti F, Consonni G, Della Vedova ML, Peluso S. Learning markov equivalence classes of directed acyclic graphs: an objective bayes approach. Bayesian Anal 2018;13(4):1235-1260. [doi: 10.1214/18-ba1101]
51. Ayhan CHB, Bilgin H, Uluman OT, Sukut O, Yilmaz S, Buzlu S. A systematic review of the discrimination against sexual and gender minority in health care settings. Int J Health Serv 2020;50(1):44-61 [FREE Full text] [doi: 10.1177/0020731419885093] [Medline: 31684808]
52. D'Urso G, Pace U. Homophobic bullying among adolescents: the role of insecure-dismissing attachment and peer support. J LGBT Youth 2019;16(2):173-191. [doi: 10.1080/19361653.2018.1552225]
53. Zhao M, Xiao D, Wang W, Wu R, Zhang W, Guo L, et al. Association among maltreatment, bullying and mental health, risk behavior and sexual attraction in chinese students. Acad Pediatr 2021;21(5):849-857. [doi: 10.1016/j.acap.2020.11.024] [Medline: 33279736]
54. Moore SE, Norman RE, Suetani S, Thomas HJ, Sly PD, Scott JG. Consequences of bullying victimization in childhood and adolescence: a systematic review and meta-analysis. World J Psychiatry 2017;7(1):60-76 [FREE Full text] [doi: 10.5498/wjp.v7.i1.60] [Medline: 28401049]
55. Kaiser T, Herzog P, Voderholzer U, Brakemeier EL. Unraveling the comorbidity of depression and anxiety in a large inpatient sample: network analysis to examine bridge symptoms. Depress Anxiety 2021;38(3):307-317 [FREE Full text] [doi: 10.1002/da.23136] [Medline: 33465284]
56. Bettis AH, Liu RT. Population-based analysis of temporal trends in the prevalence of depressed mood among sexual minority and heterosexual youths from 1999 through 2017. JAMA Pediatr 2019;173(12):1201-1203 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.3804] [Medline: 31633768]
57. Tholander M, Lindberg A, Svensson D. “A freak that no one can love”: difficult knowledge in testimonials on school bullying. Res Pap 2019;35(3):359-377 [FREE Full text] [doi: 10.1080/02671522.2019.1568534]
58. Ma S, Yang J, Xu J, Zhang N, Kang L, Wang P, et al. Using network analysis to identify central symptoms of college students' mental health. J Affect Disord 2022;311:47-54 [FREE Full text] [doi: 10.1016/j.jad.2022.05.065] [Medline: 35588909]

59. Källmén H, Hallgren M. Bullying at school and mental health problems among adolescents: a repeated cross-sectional study. Child Adolesc Psychiatry Ment Health 2021;15(1):74 [FREE Full text] [doi: 10.1186/s13034-021-00425-y] [Medline: 34906157]
60. Kircanski K, White LK, Tseng WL, Wiggins JL, Frank HR, Sequeira S, et al. A latent variable approach to differentiating neural mechanisms of irritability and anxiety in youth. JAMA Psychiatry 2018;75(6):631-639 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.0468] [Medline: 29625429]
61. Sheldon S, Williams K, Harrington S, Otto AR. Emotional cue effects on accessing and elaborating upon autobiographical memories. Cognition 2020;198:104217 [FREE Full text] [doi: 10.1016/j.cognition.2020.104217] [Medline: 32044616]
62. Solomon DT, Combs EM, Allen K, Roles S, DiCarlo S, Reed O, et al. The impact of minority stress and gender identity on PTSD outcomes in sexual minority survivors of interpersonal trauma. Psychol Sex 2019;12(1-2):64-78. [doi: 10.1080/19419899.2019.1690033]
63. Evans CBR, Smokowski PR, Rose RA, Mercado MC, Marshall KJ. Cumulative bullying experiences, adolescent behavioral and mental health, and academic achievement: an integrative model of perpetration, victimization, and bystander behavior. J Child Fam Stud 2018;27(9):2415 [FREE Full text] [doi: 10.1007/s10826-018-1078-4] [Medline: 30174382]
64. Braun A, Santesteban-Echarri O, Cadenhead KS, Cornblatt BA, Granholm E, Addington J. Bullying and social functioning, schemas, and beliefs among youth at clinical high risk for psychosis. Early Interv Psychiatry 2022;16(3):281-288. [doi: 10.1111/eip.13157] [Medline: 33938145]
65. Day JK, Fish JN, Grossman AH, Russell ST. Gay-straight alliances, inclusive policy, and school climate: LGBTQ youths' experiences of social support and bullying. J Res Adolesc 2020;30 Suppl 2(Suppl 2):418-430 [FREE Full text] [doi: 10.1111/jora.12487] [Medline: 30861243]
66. Luk JW, Gilman SE, Haynie DL, Simons-Morton BG. Sexual orientation and depressive symptoms in adolescents. Pediatrics 2018;141(5):e20173309 [FREE Full text] [doi: 10.1542/peds.2017-3309] [Medline: 29661939]
67. Ke T, De Simoni S, Barker E, Smith P. The association between peer-victimisation and structural and functional brain outcomes: a systematic review. JCPP Adv 2022;2(2):e12081 [FREE Full text] [doi: 10.1002/jcv2.12081] [Medline: 37431463]
68. Barnett AP, Molock SD, Nieves-Lugo K, Zea MC. Anti-LGBT victimization, fear of violence at school, and suicide risk among adolescents. Psychol Sex Orientat Gend Divers 2019;6(1):88-95 [FREE Full text] [doi: 10.1037/sgd0000309] [Medline: 37383950]
69. Bjereld Y, Daneback K, Mishna F. Adults’ responses to bullying: the victimized youth’s perspectives. Res Pap Educ 2019;36(3):257-274 [FREE Full text] [doi: 10.1080/02671522.2019.1646793]
70. Nielsen T. The stress acceleration hypothesis of nightmares. Front Neurol 2017;8:201 [FREE Full text] [doi: 10.3389/fneur.2017.00201] [Medline: 28620339]
71. Kastner-Dorn AK, Andreatta M, Pauli P, Wieser MJ. Hypervigilance during anxiety and selective attention during fear: using steady-state visual evoked potentials (ssVEPs) to disentangle attention mechanisms during predictable and unpredictable threat. Cortex 2018;106:120-131 [FREE Full text] [doi: 10.1016/j.cortex.2018.05.008] [Medline: 29929061]
72. Puhl RM, Lessard LM. Weight stigma in youth: prevalence, consequences, and considerations for clinical practice. Curr Obes Rep 2020;9(4):402-411. [doi: 10.1007/s13679-020-00408-8] [Medline: 33079337]
73. Bucchianeri MM, Gower AL, McMorris BJ, Eisenberg ME. Youth experiences with multiple types of prejudice-based harassment. J Adolesc 2016;51:68-75. [doi: 10.1016/j.adolescence.2016.05.012] [Medline: 27310725]
74. Reece LJ, Bissell P, Copeland RJ. 'I just don't want to get bullied anymore, then I can lead a normal life'; insights into life as an obese adolescent and their views on obesity treatment. Health Expect 2016;19(4):897-907 [FREE Full text] [doi: 10.1111/hex.12385] [Medline: 27403849]
75. Himmelstein MS, Puhl RM, Watson RJ. Weight-based victimization, eating behaviors, and weight-related health in sexual and gender minority adolescents. Appetite 2019;141:104321 [FREE Full text] [doi: 10.1016/j.appet.2019.104321] [Medline: 31202917]
76. Al-Raqqad HK, Al-Bourini ES, Al Talahin FM, Aranki RME. The impact of school bullying on students' academic achievement from teachers point of view. Int Educ Stud 2017;10(6):44-50. [doi: 10.5539/ies.v10n6p44]
77. Cao Z, Cini E, Pellegrini D, Fragkos KC. The association between sexual orientation and eating disorders-related eating behaviours in adolescents: a systematic review and meta-analysis. Eur Eat Disord Rev 2023;31(1):46-64 [FREE Full text] [doi: 10.1002/erv.2952] [Medline: 36367345]
78. Yean C, Benau EM, Dakanalis A, Hormes JM, Perone J, Timko CA. The relationship of sex and sexual orientation to self-esteem, body shape satisfaction, and eating disorder symptomatology. Front Psychol 2013;4:887 [FREE Full text] [doi: 10.3389/fpsyg.2013.00887] [Medline: 24348441]
79. Cella S, Iannaccone M, Ascione R, Cotrufo P. Body dissatisfaction, abnormal eating behaviours and eating disorder attitude in homo- and heterosexuals. Eat Weight Disord 2010;15(3):e180-e185. [doi: 10.3275/6866] [Medline: 20160471]
80. Hepworth R, Mogg K, Brignell C, Bradley BP. Negative mood increases selective attention to food cues and subjective appetite. Appetite 2010;54(1):134-142. [doi: 10.1016/j.appet.2009.09.019] [Medline: 19815043]
81. Svaldi J, Werle D, Naumann E, Eichler E, Berking M. Prospective associations of negative mood and emotion regulation in the occurrence of binge eating in binge eating disorder. J Psychiatr Res 2019;115:61-68. [doi: 10.1016/j.jpsychires.2019.05.005] [Medline: 31121393]

82. Doan Van EE, Mereish EH, Woulfe JM, Katz-Wise SL. Perceived discrimination, coping mechanisms, and effects on health in bisexual and other non-monosexual adults. Arch Sex Behav 2019;48(1):159-174 [FREE Full text] [doi: 10.1007/s10508-018-1254-z] [Medline: 30105618]
83. Radoman M, Akinbo FD, Rospenda KM, Gorka SM. The impact of startle reactivity to unpredictable threat on the relation between bullying victimization and internalizing psychopathology. J Psychiatr Res 2019;119:7-13 [FREE Full text] [doi: 10.1016/j.jpsychires.2019.09.004] [Medline: 31520836]
84. Tseng WL, Deveney CM, Stoddard J, Kircanski K, Frackman AE, Yi JY, et al. Brain mechanisms of attention orienting following frustration: associations with irritability and age in youths. Am J Psychiatry 2019;176(1):67-76 [FREE Full text] [doi: 10.1176/appi.ajp.2018.18040491] [Medline: 30336704]
85. Ebner-Priemer UW, Eid M, Kleindienst N, Stabenow S, Trull TJ. Analytic strategies for understanding affective (in)stability and other dynamic processes in psychopathology. J Abnorm Psychol 2009;118(1):195-202. [doi: 10.1037/a0014868] [Medline: 19222325]
86. Forbes MK, Wright AGC, Markon KE, Krueger RF. Evidence that psychopathology symptom networks have limited replicability. J Abnorm Psychol 2017;126(7):969-988 [FREE Full text] [doi: 10.1037/abn0000276] [Medline: 29106281]

Abbreviations

bEI: bridge expected influence
CPDAG: completed partially directed acyclic graph
EI: expected influence
GAD: Generalized Anxiety Disorder Scale
PHQ: Patient Health Questionnaire
PTSD: posttraumatic stress disorder

Edited by A Mavragani, T Sanchez; submitted 13.03.23; peer-reviewed by E Fedele, M Scutari; comments to author 17.05.23; revised version received 06.06.23; accepted 05.09.23; published 01.11.23.

Please cite as:
Li J, Jin Y, Xu S, Wilson A, Chen C, Luo X, Liu Y, Ling X, Sun X, Wang Y
Effects of Bullying on Anxiety, Depression, and Posttraumatic Stress Disorder Among Sexual Minority Youths: Network Analysis
JMIR Public Health Surveill 2023;9:e47233
URL: https://publichealth.jmir.org/2023/1/e47233
doi:10.2196/47233
PMID:37910159

Original Paper

Small Area Forecasting of Opioid-Related Mortality: Bayesian Spatiotemporal Dynamic Modeling Approach

Cici Bauer[1], PhD; Kehe Zhang[1], MS; Wenjun Li[2], PhD; Dana Bernson[3], MPH; Olaf Dammann[4,5], MD, PhD; Marc R LaRochelle[6,7], MPH, MD; Thomas J Stopka[4,8,9], MHS, PhD

[1]Department of Biostatistics and Data Science, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, United States

[2]Department of Public Health, University of Massachusetts Lowell, Lowell, MA, United States

[3]Office of Population Health, Department of Public Health, The Commonwealth of Massachusetts, Boston, MA, United States

[4]Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, MA, United States

[5]Department of Gynecology and Obstetrics, Hannover Medical School, Hannover, Germany

[6]Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, United States

[7]Grayken Center for Addiction, Boston Medical Center, Boston, MA, United States

[8]Department of Urban and Environmental Policy and Planning, Tufts University, Medford, MA, United States

[9]Department of Community Health, Tufts University, Medford, MA, United States

Corresponding Author:
Cici Bauer, PhD
Department of Biostatistics and Data Science
School of Public Health
The University of Texas Health Science Center at Houston
1200 Pressler Street, Room E819
Houston, TX, 77030
United States
Phone: 1 713 500 9581
Email: cici.x.bauer@uth.tmc.edu

Abstract

Background: Opioid-related overdose mortality has remained at crisis levels across the United States, increasing 5-fold and worsened during the COVID-19 pandemic. The ability to provide forecasts of opioid-related mortality at granular geographical and temporal scales may help guide preemptive public health responses. Current forecasting models focus on prediction on a large geographical scale, such as states or counties, lacking the spatial granularity that local public health officials desire to guide policy decisions and resource allocation.

Objective: The overarching objective of our study was to develop Bayesian spatiotemporal dynamic models to predict opioid-related mortality counts and rates at temporally and geographically granular scales (ie, ZIP Code Tabulation Areas [ZCTAs]) for Massachusetts.

Methods: We obtained decedent data from the Massachusetts Registry of Vital Records and Statistics for 2005 through 2019. We developed Bayesian spatiotemporal dynamic models to predict opioid-related mortality across Massachusetts' 537 ZCTAs. We evaluated the prediction performance of our models using the one-year ahead approach. We investigated the potential improvement of prediction accuracy by incorporating ZCTA-level demographic and socioeconomic determinants. We identified ZCTAs with the highest predicted opioid-related mortality in terms of rates and counts and stratified them by rural and urban areas.

Results: Bayesian dynamic models with the full spatial and temporal dependency performed best. Inclusion of the ZCTA-level demographic and socioeconomic variables as predictors improved the prediction accuracy, but only in the model that did not account for the neighborhood-level spatial dependency of the ZCTAs. Predictions were better for urban areas than for rural areas, which were more sparsely populated. Using the best performing model and the Massachusetts opioid-related mortality data from 2005 through 2019, our models suggested a stabilizing pattern in opioid-related overdose mortality in 2020 and 2021 if there were no disruptive changes to the trends observed for 2005-2019.

Conclusions: Our Bayesian spatiotemporal models focused on opioid-related overdose mortality data facilitated prediction approaches that can inform preemptive public health decision-making and resource allocation. While sparse data from rural and less populated locales typically pose special challenges in small area predictions, our dynamic Bayesian models, which maximized information borrowing across geographic areas and time points, were used to provide more accurate predictions for small areas. Such approaches can be replicated in other jurisdictions and at varying temporal and geographical levels. We encourage the formation of a modeling consortium for fatal opioid-related overdose predictions, where different modeling techniques could be ensembled to inform public health policy.

(JMIR Public Health Surveill 2023;9:e41450) doi:10.2196/41450

KEYWORDS

opioid-related mortality; small area estimation; spatiotemporal models; Bayesian; forecasting

Introduction

Opioid-related overdoses continue to be at crisis levels in communities across the United States, with more than 75,673 fatal overdoses in the 12-month period ending in April 2021 [1-7], worsening during the COVID-19 pandemic [8]. Opioid-related deaths increased more than 5-fold in Massachusetts between 2000 and 2016, with more than 2000 per year from 2016 to 2021 [9,10]. Fatal opioid-related overdose rates above the national level have been ubiquitous across communities in Massachusetts [11]. Despite this crisis, public health responses to the opioid overdose epidemic have been limited by an inability to rapidly identify current fatal overdose patterns, predict future local clusters, and evaluate the effectiveness of interventions.

Identification and prediction of local fatal opioid overdoses require a comprehensive and high-quality surveillance system that provides data sources to capture granular geographic information of the fatal opioid overdose cases across space and time. Ideally, additional information such as individual demographics, past medical history (particularly mental health history), local drug supply, and other risk factors should be included as well to enhance our understanding of the opioid crisis. Many states have established surveillance systems to monitor opioid-related morbidity and mortality to inform planning and evaluate control efforts. Some surveillance systems include unlinked individual data sources for different opioid-related reporting (eg, vital records and prescription drug monitoring programs), and some aim to provide an individually linked database across various data sources [12]. While the latter provides much enhanced data capacity for a wide range of opioid-related research, creating such linked databases takes substantial time and financial resources.

Accurate identification and prediction of fatal opioid-related overdose trends also requires sophisticated spatial and predictive analytical approaches, as noted in a recent review of methodological approaches for the prediction of opioid use–related epidemics in the United States [13]. Most published literature has focused on the identification of community- or neighborhood-level risk factors for opioid-related overdose. For example, Bozorgi et al [14] explored different machine learning and spatial analytical approaches to identify leading contextual risk factors for drug overdose at the block group level in South Carolina. A similar study by Schell et al [15] also focused on identifying new neighborhood-level predictors of opioid-related overdose deaths in Rhode Island at the census block group level, using least absolute shrinkage and selection operator and random forest algorithms. Abell-Hart et al [16] identified counties with a high number of underserved opioid overdose patients in New York state. Basak et al [17] detected spatiotemporal hot spots (ie, counties) at high risk of prescription opioid misuse and overdose using regression models that relied on Medicare claims data in Virginia, North Carolina, and West Virginia.

Statistical models or machine learning techniques proposed specifically for prediction purposes are limited, with most previous studies performed at the US county or state level [18,19], lacking the spatial granularity needed to guide local public health departments for preemptive actions. Bayesian spatiotemporal models have received substantial attention in the past several years in opioid-related research, given their ability to include temporal and spatial correlations and improved precision in small area estimation [20-22]. Sumetsky et al [20] for instance, developed a Bayesian logistic growth model for opioid overdose mortality predictions for 146 counties in North and South Carolina.

In this study, we developed and validated several Bayesian spatiotemporal dynamic predictive models designed for small area forecasting of opioid-related overdose mortality at the ZIP Code Tabulation Area (ZCTA) level. We investigated the benefits of including various area-level demographic and socioeconomic factors in improving the prediction. The predictive performance was assessed using opioid-related overdose mortality data from Massachusetts between 2005 and 2019. We identified the top ZCTAs with the highest predicted fatal opioid overdose rates or counts and by urban and rural areas. Our prediction results can help inform local public health departments' planning and targeting of resource allocations.

Methods

Opioid-Related Mortality Data

Opioid-related mortality data were obtained from the Massachusetts Department of Public Health's (MDPH's) Registry of Vital Records and Statistics (RVRS) for 2005 through 2019. International Classification of Disease, Tenth Revision, codes for mortality were used to select from the *underlying cause of death* field by RVRS staff and identify poisonings or overdoses: X40-X44, X60-X64, X85, and Y10-Y14. All multiple *underlying cause of death* fields were then used to define opioid-related death with T40.0, T40.1, T40.2, T40.3, T40.4, and T40.6. We excluded opioid-related deaths for individuals younger than 20 years, since fatal opioid-related death was rare in this age group (n=191). This age grouping also allows for comparison of opioid-related overdose rates with other drug overdose studies [23]. We used the ZCTA of the reported residential address for decedents at the time of the fatal overdose and excluded those either with missing addresses or those with addresses outside of Massachusetts' boundaries (n=642). The final analytic sample included 16,377 fatal opioid-related overdoses from 2005 to 2019. The flowchart for deriving the analytic samples used for this analysis, summarized by each year, is presented in Figure S1 in Multimedia Appendix 1. We also developed a web-based dashboard to present the map and summaries of these data [24]. Requests for the opioid-related decedent data should be made directly to the MDPH's RVRS.

ZCTA Demographic and Socioeconomic Factors

ZCTA-level demographic and socioeconomic variables were obtained from 2005 to 2019 from the American Community Survey 5-year estimates (2005-2009, 2010-2014, and 2015-2019, respectively) [25]. We used the total population counts and characteristics for people aged 20 years or older at the ZCTA level to account for differences in population size. The ZCTA-level demographic variables included race or ethnicity (proportions of White and Hispanic individuals), education (proportion of individuals with a bachelor's degree), employment (proportion of unemployed individuals), poverty level (proportion of individuals living under the federal poverty level), income (per capita income in US $), living conditions (proportion of renters), transportation (proportion of individuals without vehicles), and English speaking (proportions of individuals with limited English). A detailed summary of these covariates, for all ZCTAs and those stratified by urban or rural areas, is presented in Table S1 in Multimedia Appendix 1. These covariates were considered from a larger set of community-level covariates and were selected on the basis of the literature and our prior research. Following calculation of descriptive statistics and bivariate analyses, we further narrowed the list to a subset of population-level measures at the ZCTA level. The final set of covariates included in this analysis was chosen after fitting a simple Poisson regression model and selecting those with a significant association with opioid-related overdose mortality (see Figure S5 in Multimedia Appendix 1).

Urban and Rural ZCTAs

The Massachusetts State Office of Rural Health uses a composite scoring system to categorize rurality for each ZCTA. Among the 537 ZCTAs included in this analysis, 390 (73%) were classified as urban areas, and the rest as rural areas. A map of the ZCTAs, color-coded by their urban or rural status, is provided in Figure S2 in Multimedia Appendix 1.

Statistical Models

Our choice of dynamic spatiotemporal model was motivated by a previous study of drug overdose rates across US counties, which showed that the strongest predictor of overdose rates was overdose rates in nearby counties from the previous year. We considered the following Bayesian dynamic spatiotemporal models [26,27]. Y_{it} denotes the number of opioid-related fatal overdose counts in ZCTA i (i=1, …, I=537) and year t (t=2005, …, t=2019). We assumed a Poisson distribution with mortality rate μ_{it}; that is, $Y_{it}|\mu_{it} \sim Poisson(N_{it}\mu_{it})$, with N_{it} representing the population size as an offset. The mortality rate μ_{it} was then decomposed using the following models:

Model 1: $\log(\mu_{it}) = \alpha_i + S(t) + \eta_t$, $\eta_t \sim AR(1)$,

Model 2: $\log(\mu_{it}) = \alpha_i + S(t) + \omega_{it}$, $\omega_{it} \sim GMRF(\tau\Sigma_{AR(1)} \otimes \Sigma_I)$, and

Model 3: $\log(\mu_{it}) = \alpha + S(t) + \omega_{it}$, $\omega_{it} \sim GMRF(\tau\Sigma_{AR(1)} \otimes \Sigma_{CAR})$

In all 3 models, $S(t)$ captured the overall (or marginal) temporal trend and could be modeled via linear or quadratic terms or spline functions. The major differences among the 3 models were how random effects (η or ω) were modeled in the spatiotemporal interactions. In model 1, α_i was the random effect accounting for the difference of the opioid-related mortality at the ZCTA level, and η_t was a first-order autogressive latent effect $AR(1)$ for the temporal trend. Essentially, model 1 regressed the opioid-related mortality rate at time t, on the log scale, over the previous time point t–1; however, instead of running one model for each ZCTA, the Bayesian framework allowed pooling of the ZCTAs to improve prediction. In models 2 and 3, α was the overall intercept, and the ZCTA-level differences of the opioid-related mortality rates were absorbed in the spatiotemporal interaction term ω_{it}, which now accounted for the spatial dependency between ZCTAs. Models 2 and 3 were inseparable spatiotemporal models, where the interaction term ω_{it} had a Gaussian Markov Random Field (GMRF) with mean 0 and covariance matrix $\tau\Sigma_T \otimes \Sigma_S$ [28]. The interaction term was the Kronecker product of the temporal structure Σ_T and the spatial structure Σ_S, and τ was the precision parameter (ie, reciprocal of the variance parameter). We again used the $AR(1)$ temporal structure in models 2 and 3 but considered different structures for the spatial dependency. In model 2, the spatial structure was an $I \times I$ identity matrix. Although the identity matrix may appear to assume independence of the areas, the hierarchical structure of the model imposed borrowing information from the neighboring areas. In model 3, we assumed a conditional autoregressive model (CAR) [29] as the spatial structure, which assumed that ZCTAs that were geographically adjacent were more similar than those that were far away. We note that other spatial or temporal structures (eg, random walk models) can also be considered in the proposed model framework.

Models 1, 2, and 3 were referred to as the respective base models. We then added ZCTA-level demographic and socioeconomic variables x_{it} and the urban or rural indicator variable described above, to each of the base models to assess potential improved prediction performance. All models were fit in a Bayesian framework, with the priors chosen as the noninformative priors commonly used in spatiotemporal models [28]. In our application, it turned out that the interaction term was sufficient to capture the spatial-temporal variation, so we dropped the marginal temporal term. We used the posterior median and 90% credible intervals (CrIs) for inference. The

90% CrI was preferred over the commonly used 95% CrI, as the former was better suited for prediction purposes to avoid overly wide uncertainty. All analyses were performed in RStudio [30] and R package integrated nested Laplace approximation [31].

To assess the predictive performance of the models, we used a one-year ahead approach: assuming we observed data from 2005 to year *t*, we predicted the opioid-related mortality rates and counts for year *t*+1. We started by considering the observed data from 2005 to 2015, and the prediction was carried out for 2016. Then we considered observed data from 2005 to 2016 and predicted counts and rates were calculated for 2017. The one-year ahead prediction procedures were carried out for each year from 2016 to 2019, and the predictions were obtained for each ZCTA, along with the prediction uncertainty (posterior 90% CrI). We used the following metrics to compare the predictive performance across different models: mean absolute error (MAE), root mean square error (RMSE), and the rank difference (RD). MAE was defined as the average absolute difference between the predictive and the observed values; that is, , for count or rate Y_{it}. RMSE was defined as the average squared difference between the predictive and the observed values; that is, . RMSE has the advantage that it was on the same scale as the outcome variable and was thus easy to interpret. For example, if the prediction was for the ZCTA-level mortality count (ie, how many people would die from fatal opioid-related overdose), then an RMSE of 10 would roughly mean that on average the prediction was off by 10 mortality cases. Smaller MAE and RMSE would indicate better performance. The RD was motivated by the correct classification of opioid mortality by group membership, where we divided the ZCTAs into quintiles Q1 to Q5, based on the observed fatal opioid-related overdose rates and counts, and then compared to the quintiles for the predicted rates and counts. The RD was then calculated by the percentage of times when the classification was correct. Higher values of RD indicated better performance.

Ethical Considerations

This study was reviewed by the Health Sciences institutional review board at Tufts University and was designated as exempt from ethics approval (IRB reference number: 13288).

Results

The number of fatal opioid-related overdoses among people aged 20 years or older increased from 557 in 2005 to 1912 in 2019, corresponding to an increase in rate from 7.95 in 2005 to 34.4 in 2019 per 100,000 population. Out of a total of 16,377 opioid-related fatalities between 2005 and 2019, overall, 91% occurred in urban ZCTAs. The observed rates at the ZCTA level were highly variable (Figure S3 in Multimedia Appendix 1), ranging from 0 to 1316 per 100,000 population. The extreme values in the observed rates tended to occur in areas with small populations. The high instability of the observed rates, commonly known as the small area estimation problem, poses a special challenge in developing accurate prediction models, which our Bayesian spatiotemporal models helped address.

Results of the predictive performance among the proposed candidate models, using the one-year ahead approach, are presented in Table 1 (for the ZCTA-level fatal opioid-related overdose count) and Table 2 (for the ZCTA-level fatal opioid-related overdose rate). Prediction errors were summarized for all ZCTAs and stratified by rural and urban ZCTAs. Overall, model 3 with the inseparable spatiotemporal interaction term of $AR(1)$ and CAR structures performed best for fatal opioid-related overdose predictions, as indicated by smaller MAE and RMSE values. The RMSE showed that the predicted opioid-related death counts, on average, were off only by 2 counts per area. As expected, predictions were better for urban areas than for rural areas, since most of the deaths occurred in urban areas. Addition of demographic and socioeconomic variables generally improved the prediction performance, particularly for model 1; however, the improvement was attenuated in models 2 and 3 where we assumed inseparable spatiotemporal models. The smaller improvement in model 3 was likely because the CAR model already captured the spatial patterns tied to the demographic and socioeconomic factors and hence served as a surrogate of the demographic and socioeconomic variables for predictions. The best performing model using the rank difference showed a less clear pattern, as the values were very similar across the different models. This was particularly true when focusing on the rural ZCTAs, likely due to their small predicted rates or counts and hence very narrow ranges for each quintile.

Table 1. Predictive performance assessment of the ZIP Code Tabulation Area (ZCTA) level opioid-related overdose death count, using the root mean and rank difference, and one-year ahead prediction starting with 2016[a]. The smallest root mean square error, mean absolute difference, and highest rank difference for each row are depicted in italics, indicating the best performing models.

Count year	ZCTA type	Model 1		Model 2		Model 3	
		Base Model	Add SDOH[b]	Base Model	Add SDOH	Base Model	Add SDOH
Root mean square error							
2016	All	2.56	2.63	2.98	2.85	*2.35*	2.40
2017	All	2.06	1.98	2.29	2.20	2.04	*1.94*
2018	All	2.53	2.47	2.72	2.56	2.47	*2.44*
2019	All	2.38	2.33	2.57	2.43	2.19	*2.18*
2016	Urban	2.90	2.98	3.38	3.22	*2.65*	2.70
2017	Urban	2.34	2.25	2.62	2.50	2.30	*2.20*
2018	Urban	2.83	*2.75*	3.04	2.84	2.76	2.71
2019	Urban	2.65	2.59	2.87	2.70	2.46	*2.45*
2016	Rural	1.26	1.32	1.42	1.43	*1.20*	1.28
2017	Rural	0.98	*0.97*	1.03	1.03	1.03	0.98
2018	Rural	1.45	1.45	1.62	1.62	*1.43*	1.48
2019	Rural	1.42	1.41	1.46	1.47	*1.15*	1.16
Mean absolute difference							
2016	All	0.92	0.98	1.43	1.32	*0.67*	0.74
2017	All	*0.01*	0.02	0.61	0.50	0.26	0.20
2018	All	0.55	0.56	0.82	0.73	*0.32*	0.37
2019	All	0.17	0.19	0.58	0.51	*0.00*	0.06
2016	Urban	1.15	1.21	1.77	1.60	*0.85*	0.91
2017	Urban	*0.04*	0.06	0.79	0.63	0.29	0.22
2018	Urban	0.61	0.60	0.91	0.78	*0.30*	0.35
2019	Urban	0.13	0.14	0.64	0.52	*0.03*	0.04
2016	Rural	0.30	0.36	0.52	0.55	*0.22*	0.28
2017	Rural	0.14	*0.08*	0.13	0.16	0.19	0.13
2018	Rural	0.40	0.44	0.58	0.60	*0.38*	0.43
2019	Rural	0.29	0.33	0.43	0.48	*0.06*	0.12
Rank difference							
2016	All	0.50	0.51	0.50	0.49	0.52	*0.53*
2017	All	0.50	0.50	0.49	0.48	0.51	*0.52*
2018	All	0.50	0.51	0.50	*0.52*	0.49	*0.52*
2019	All	0.51	0.51	0.49	0.50	0.50	*0.51*
2016	Urban	0.54	0.54	0.55	0.54	*0.57*	0.55
2017	Urban	*0.53*	*0.53*	0.51	0.52	0.52	0.52
2018	Urban	0.54	0.55	0.55	*0.56*	0.54	0.54
2019	Urban	0.58	0.60	0.54	0.57	0.61	*0.62*
2016	Rural	*0.40*	0.39	*0.40*	0.38	0.37	0.35
2017	Rural	*0.43*	0.42	0.39	0.41	0.41	0.40
2018	Rural	*0.38*	0.37	0.34	0.36	0.37	0.37
2019	Rural	0.34	0.37	0.33	0.34	0.37	*0.38*

[a] For example, using data from 2005 to 2015, we predicted the fatal opioid-related overdose count in year 2016 for each of the 537 ZIP Code Tabulation Areas (ZCTAs) in Massachusetts, and compared them to the observed data to assess performance. Similarly, for 2017, we used data from 2006 to 2016, and predicted for 2017. The prediction error was assessed for counts for all ZCTAs and stratified by rural and urban status. Base model refers to Bayesian dynamic spatiotemporal models without any covariates; model results with added covariates of area-level contextual factors are included in the column “Add SDOH.”

[b] SDOH: Social determinants of health.

Since our 1-year prediction assessments suggested that the best performing model was model 3 with the included demographic and socioeconomic variables, we used this model for 2020 and 2021 predictions based on the opioid-related mortality data for 2005-2019. We assumed that ZCTA-level population size and demographic and socioeconomic variables had the same values in 2020 and 2021 as those in 2019. We carried out the prediction for each ZCTA and at the state level. Figure 1 presents the fitted temporal trends of opioid-related mortality in Massachusetts between 2005 and 2019, with predictions carried out for 2020 and 2021 (the gray shaded area), both as the rate per 100,000 population (panel A) and total counts (panel B). Each line represents an individual ZCTA, color-coded by urban or rural status. Figure 2 presents the maps of the predicted opioid-related mortality rates and counts for 2020 and 2021. At the ZCTA-level, the predicted opioid-related mortality rate ranged from 6.11 to 162 per 100,000 population for 2020, and 6.05 to 158 per 100,000 population in 2021. Note that because of the “smoothing” effect in Bayesian models, the estimated or predicted rate could be very close to 0 but would never be exactly 0. Therefore, for ZCTAs that may have zero observed deaths, the predicted rate would not be zero but would be skewed toward the average rate. In addition, any ZCTAs that may have low observed rates but were “surrounded” by ZCTAs with high rates would have higher predicted rates, reflecting how the spatial models borrow information from neighboring ZCTAs. Because of the large variation in population size by ZCTA, the spatial patterns seen in the maps in Figure 2 for predicted rate and counts are quite different, where the latter are generally centered around highly populated areas. The predicted count ranges from close to 0 (after rounding up to integers) to 30 (90% CrI 21-41) for 2020, and 29 (90% CrI 19-44) for 2021. We also identified the top 5 ZCTAs with the highest predicted counts, by urban or rural classifications, and presented the prediction results in Figure 3. The urban ZCTAs were located within cities with high fatal opioid-related overdose risks: Lawrence, Lynn, Quincy, Brockton, and New Bedford (Figure 3C). The rural ZCTAs were located in municipalities that were known to have high risks for fatal opioid-related overdoses, including Pittsfield, North Adams, Greenfield, Westfield, and Billerica (Figure 3D).

At the state level, the predicted opioid-related mortality rate for the population aged 20 years and older was 35.79 per 100,000 population (90% CrI 29.4-43.4) for 2020, and 35.81 (90% CrI 26.9-46.0) for 2021. For urban areas, the predicted rate was 36.0 (90% CrI 29.2-43.7) in 2020, and 35.5 (90% CrI 26.8-46.1) in 2021. For rural areas, the predicted rate was 33.9 (90% CrI 26.6-43.3) in 2020, and 34.6 (90% CrI 25.2-53.4) in 2021. The CrIs for the prediction were wider in 2021 than in 2020 and in rural areas than in urban areas. This was expected as the further ahead we attempted to predict outcomes, the less certainty we would have. Our prediction suggested that the opioid-related overdose death counts for 2020 would be 1887 (90% CrI 1549-2285) for the whole state, with 1682 (90% CrI 1364-2041) in urban ZCTAs, and 203 (90% CrI 160-260) in rural ZCTAs, and those for 2021 would be 1888 (90% CrI 1419-2426) for the whole state with 1656 (90% CrI 1253-2149) in urban ZCTAs, and 208 (90% CrI 151-320) in rural ZCTAs.

Table 2. Predictive performance assessment of the ZIP Code Tabulation Area (ZCTA) level opioid-related overdose death rate, using the root mean and rank difference, and one-year ahead prediction starting with 2016[a]. The smallest root mean square error, mean absolute difference, and highest rank difference for each row are depicted in italics, indicating the best performing models.

Rate year	ZCTA type	Model 1		Model 2		Model 3	
		Base Model	Add SDOH[b]	Base Model	Add SDOH	Base Model	Add SDOH
Root mean square error							
2016	All	41.18	40.27	42.84	41.94	40.28	*38.95*
2017	All	29.21	30.18	29.01	30.24	*28.63*	29.82
2018	All	40.54	*40.22*	42.63	41.83	40.47	40.65
2019	All	74.80	75.64	75.75	76.03	*73.26*	75.39
2016	Urban	44.53	43.39	46.52	45.45	43.57	*41.65*
2017	Urban	30.31	31.76	30.75	32.30	*29.63*	31.13
2018	Urban	38.30	*37.67*	40.82	39.71	38.17	38.24
2019	Urban	83.48	84.42	84.48	84.74	*81.59*	84.13
2016	Rural	30.50	30.45	30.98	30.74	*29.79*	30.62
2017	Rural	26.08	25.53	*23.76*	23.91	25.80	26.02
2018	Rural	*46.00*	46.33	47.10	47.03	46.04	46.47
2019	Rural	44.05	44.49	44.90	45.31	*43.99*	44.38
Mean absolute difference							
2016	All	7.29	7.27	13.73	12.44	*4.91*	4.95
2017	All	3.70	4.01	4.31	*2.81*	5.87	5.83
2018	All	5.89	5.36	10.59	9.26	4.10	*4.05*
2019	All	4.47	4.01	10.07	8.83	1.60	*1.57*
2016	Urban	10.02	9.41	16.37	14.41	7.25	*7.01*
2017	Urban	*2.11*	3.09	5.69	3.56	5.09	5.24
2018	Urban	5.72	4.61	10.05	8.15	3.01	*2.83*
2019	Urban	4.60	3.64	10.02	8.24	2.07	*1.91*
2016	Rural	*0.00*	1.57	6.68	7.20	1.31	0.55
2017	Rural	7.94	6.44	*0.64*	0.84	7.93	7.40
2018	Rural	*6.34*	7.39	12.04	12.23	6.98	7.29
2019	Rural	4.13	5.02	10.22	10.40	*0.35*	0.68
Rank difference							
2016	All	0.30	*0.34*	0.31	0.33	0.31	0.32
2017	All	0.32	0.36	0.32	*0.37*	0.34	0.35
2018	All	0.29	0.31	0.31	0.30	0.31	*0.34*
2019	All	0.33	0.34	0.34	*0.35*	0.34	0.32
2016	Urban	0.32	*0.38*	0.31	0.37	0.33	*0.38*
2017	Urban	0.34	*0.38*	0.35	0.35	0.37	*0.38*
2018	Urban	0.32	0.34	0.30	0.33	0.33	0.34
2019	Urban	0.36	0.36	0.36	0.37	*0.38*	0.36
2016	Rural	0.28	0.24	*0.30*	0.24	0.29	0.26
2017	Rural	0.19	0.21	*0.30*	0.27	0.22	0.23
2018	Rural	0.21	*0.26*	0.24	0.24	*0.26*	*0.26*
2019	Rural	*0.28*	0.25	0.25	0.26	0.27	*0.28*

[a] For example, using data from 2005 to 2015, we predicted the fatal opioid-related overdose rate in 2016 for each of the 537 ZIP Code Tabulation Areas (ZCTAs) in Massachusetts, and compared the predicted rates to the observed rates to assess prediction performance. Similarly, for 2017, we used data from 2006 to 2016 and predicted for 2017. The prediction error was assessed for rates for all ZCTAs and stratified by rural and urban status. Base model refers to Bayesian dynamic spatiotemporal models without any covariates; model results with added covariates of area-level contextual factors are included in the column labeled "Add SDOH."

[b] SDOH: Social determinants of health.

Figure 1. Fitted temporal trends for opioid-related overdose (OD) mortality in Massachusetts between 2005 and 2019, with predictions for 2020 and 2021 (in gray shade). Each line represents a ZIP Code Tabulation Area (ZCTA) color-coded by its urban or rural status. Opioid-related mortality data were obtained from the Massachusetts Registry for Vital Records and Statistics, and predictions were made on the fatal opioid-related overdose rates per 100,000 population (panel A) and count (panel B). These results are from the Bayesian dynamic spatiotemporal Model 3 with ZCTA level demographic and socioeconomic variables.

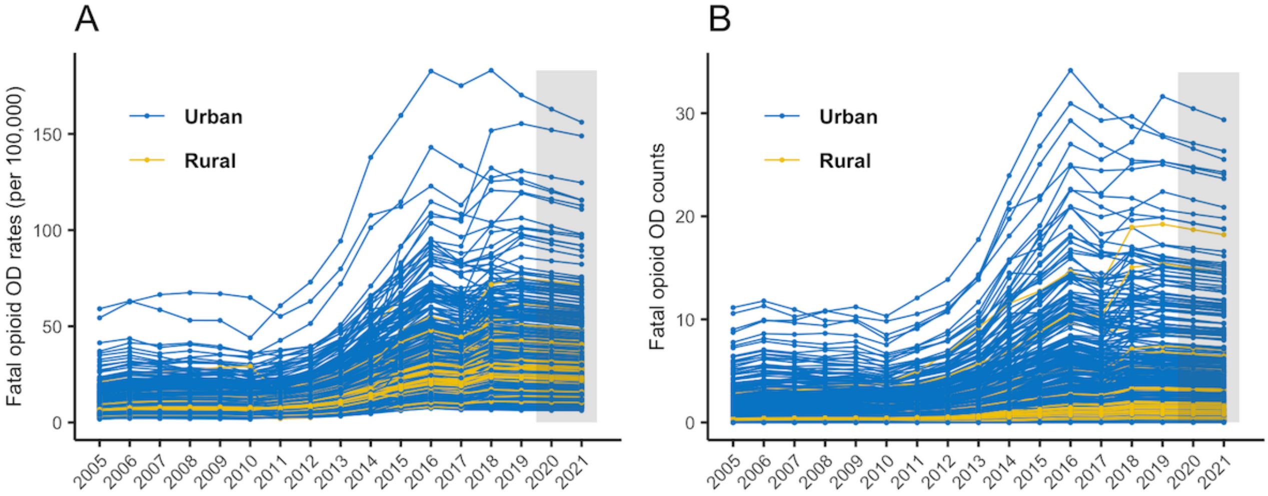


Figure 2. Maps of the predicted ZIP Code Tabulation Area (ZCTA)–level fatal opioid-related overdose (OD) rates (per 100,000 population) and counts for 2020 (A and B) and 2021 (C and D). Data were obtained from Massachusetts Registry for Vital Records and Statistics for 2005 to 2019 and used to predict for 2020 and 2021. These predictions were obtained from the proposed Bayesian dynamic spatiotemporal Model 3 with ZCTA level demographic and socioeconomic variables. In each panel, the embedded histograms present the distribution of the predicted fatal opioid-related overdose rates or counts for that year.

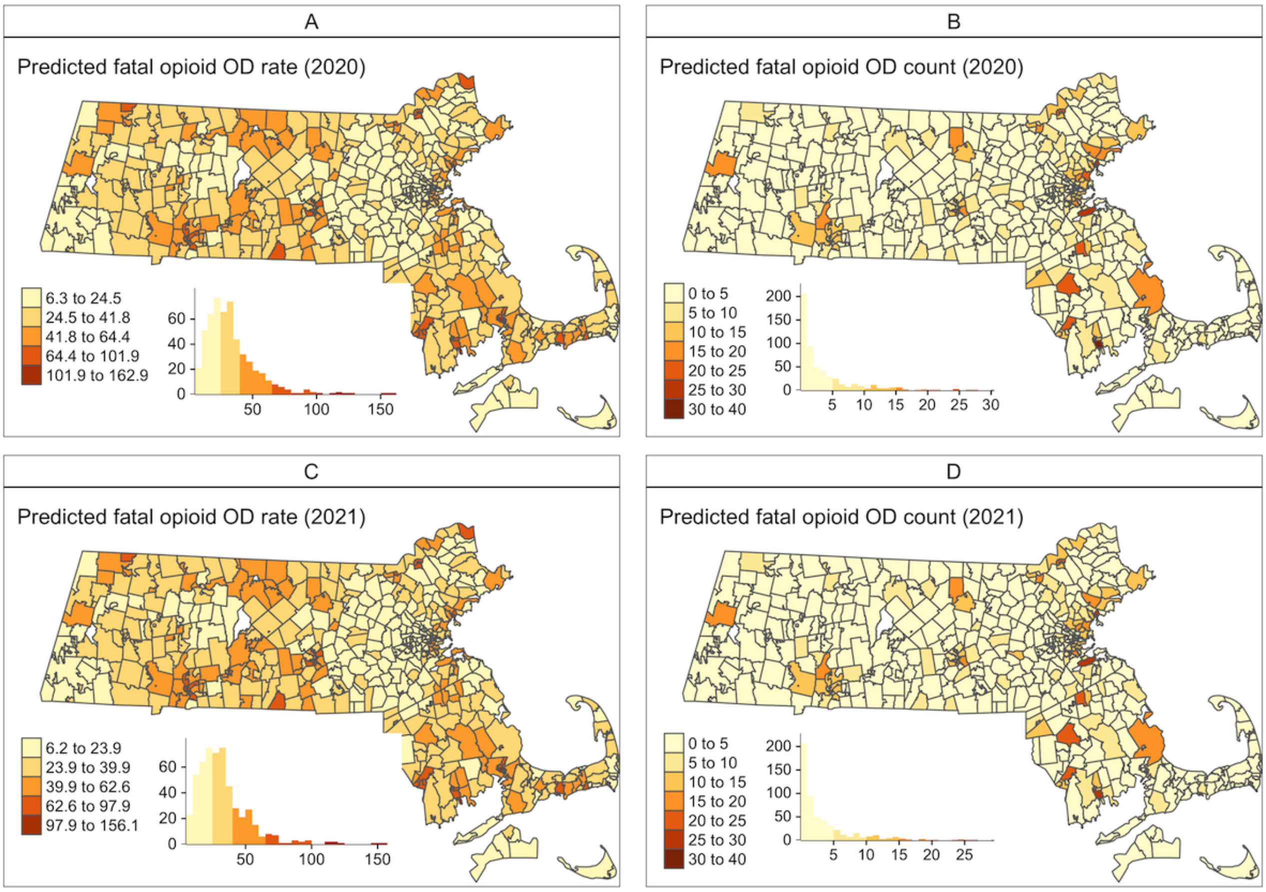
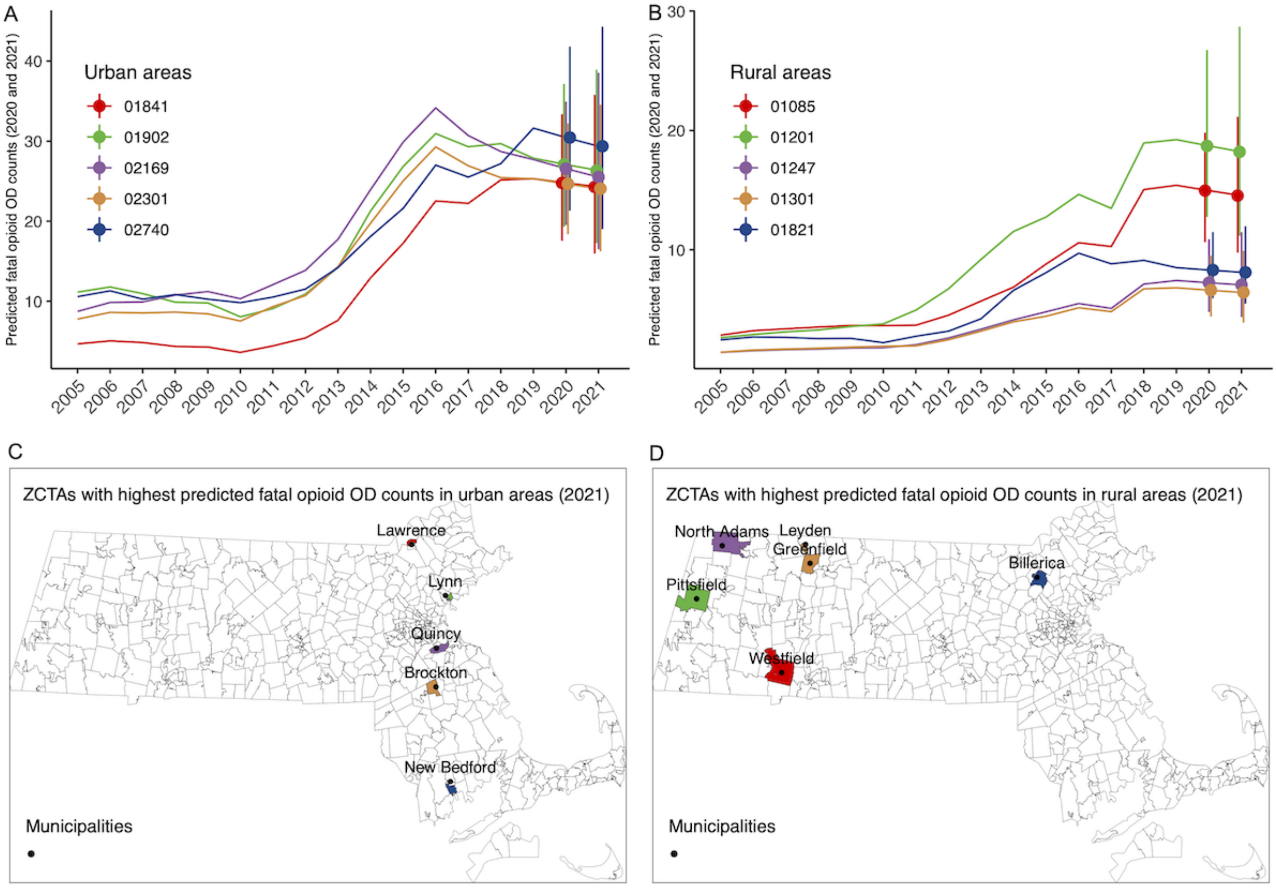


Figure 3. Five selected ZIP Code Tabulation Areas (ZCTAs) with the highest predicted fatal opioid-related overdose (OD) rates or counts for urban (A, C) and rural (B, D) areas. The vertical bars in panels A and B present the 90% posterior credible intervals (CrIs) of the predictions for years 2020 and 2021. Maps (C, D) present the corresponding locations of the identified ZCTAs. Predictions were performed using the proposed Bayesian dynamic spatiotemporal model 3 with ZCTA level demographic and socioeconomic variables and fatal opioid-related overdose data from 2005 to 2019 in Massachusetts.



Discussion

In this analysis, we developed and compared several Bayesian spatiotemporal dynamic models for predicting small area opioid-related mortality. The prediction performance was evaluated with data from the MDPH RVRS for 2005 through 2019 using the one-year ahead approach, with and without ZCTA-level demographic and socioeconomic variables. Using data from 2005 through 2019 and the best performing model, we also predicted the fatal opioid-related death rates and counts for 2020 and 2021, respectively, along with uncertainty assessments. We also identified the ZCTAs—by urban and rural status—that had the highest predicted opioid-related mortality, both by rates and by counts.

Our prediction showed that, if there was no interruptive change to the trends observed for 2005-2019, we would observe a stabling trend of the fatal opioid-related overdose in Massachusetts for 2020 and 2021. The stabling trend would be applicable to the entire state (ie, the state-level total fatal opioid-related overdose) as well as for most ZCTAs in Massachusetts. Our prediction also identified the ZCTAs deviating from this stabilizing trend, with continually increasing rates above the state average. Identifying ZCTAs with the predicted high risks allows for the possibility of preemptive and geo-targeted public health interventions. The prediction models presented here allowed for a more granular depiction of existing and expected trends in opioid-related overdose deaths over spatially granular units, which differed from other existing models developed for predication at larger geographical scales (eg, state or county). Such predictions are instrumental for state and local public health departments' planning to identify and address potential service gaps for deploying harm reduction and treatment interventions. Local data are largely limited both in type and quantity; drug seizure data, for example, are often not available at more granular levels than the state, limiting input information in predictive modeling. However, up-to-date data at the local scale are instrumental to developing the prediction models and engaging communities in designing and implementing data-driven responses to reduce opioid-related harms [32].

At the time of this analysis, we only obtained the fatal opioid-related overdose data for Massachusetts from 2005 to 2019 and partial data in 2020 (incomplete data with only the first 6 months). We designed our analysis and prediction evaluation on the basis of the data available to us, but we lacked the official statistics for 2020 and 2021 to assess the prediction accuracy for those years. The latest Massachusetts opioid-related overdose data brief released in December 2022 [33] reported the confirmed opioid-related overdose deaths among Massachusetts residents for 2020 and 2021. Although the data brief reported the total number and rate including all ages, which was different from the adult population we focused on, it was clear that we underpredicted the total number of fatal opioid-related overdose in Massachusetts for 2020 and 2021. The data brief also noted a 9% increase in the opioid-related overdose death rate in Massachusetts for 2021 over 2020. However, since the brief only reported the total opioid overdose death counts at the state level and not by ZCTAs, we were not able to assess the underprediction at a more granular spatial level. The sharp increase in the fatal opioid-related overdose was likely due to the substantive impact of the COVID-19 pandemic [8,34], which disrupted any established trends prior to 2020 that informed our prediction modeling. The time lag in the drug overdose database had been identified as a main barrier to developing fatal overdose predictive models, as reported by Borquez and Martin [35]. Our future work, in collaboration with the MDPH, plans to use the statewide Public Health Data warehouse [36] to obtain timelier and rich data sources with local information in order to improve predictive performance in small areas.

It is important to consider the limitations inherent in the highlighted analyses, and recommendations that should be considered when developing future opioid-related overdose prediction models. First, the spatial units of ZCTAs used in our study may not be the ideal spatial unit for prediction, as the population size within ZCTAs vary substantially, compared to other spatial units of analysis (eg, census tracts) where, by design, the population sizes are much more homogenous. In addition, ZCTAs may not be sufficiently granular, from a spatial perspective, to identify local hot spots. However, they are extensively used in spatial analysis as they are most readily available in many aggregated data sources (eg, surveillance or insurance claims databases). ZCTAs provide useful geospatial information for analyses while often also satisfying data privacy concerns. Second, a better understanding of the contributing factors to local opioid overdose trends and patterns, with reliable measures representing such factors, would clearly improve prediction power and accuracy. For example, toxicology data would have helped us to include the appearance of fentanyl in the local drug supply, to aid the prediction in the shifting “waves” in opioid-related fatal overdose [37]. Third, we used the ZCTA of the decedents’ residences, rather than the injury addresses or the locations where the fatal opioid overdoses were recorded. Injury data generally have a high level of missingness (~50% in MA), and the recorded death location, if not at the decedent’s place of residence, is often recorded at a hospital, even though the injury (ie, overdose event) typically occurred elsewhere. Finally, and perhaps most challenging, is to incorporate the impact of emerging phenomena such as the COVID-19 pandemic into predictions of opioid-related overdose trends, which was not possible for our analyses given the data availability at the time of analyses. This task requires the real-time data inputs and requires a joint effort among researchers and practitioners from multiple agencies, institutions, and sectors. We have seen a lot of progress made on this front during the COVID-19 pandemic, and we hope to see more progress in the drug overdose research in the future.

Despite the abovementioned data source and methodological limitations, our models showed promise in providing reasonable 1-year forecasts of opioid-related mortality in MA with geographic granularity using existing data, as the short-term point estimates for the number of overdoses tended to be close to the true value. Our Bayesian spatiotemporal models further demonstrated the advantages of incorporating inseparable spatiotemporal dependencies over the simpler regression models without such dependence. Since the assumed spatial dependency structure captured the spatial patterns tied to many demographic and socioeconomic factors, such models do not rely on the knowledge of the future measures of these factors in predicting opioid-related mortality. Prediction is a challenging problem in general, and though many models have been developed in various contexts, it is almost impossible to find one single model or approach that universally performs best [38]. Although our analysis could not investigate all possible predictive models for fatal opioid-related overdoses, we provided a novel approach to forecasting overdose events for small geographic areas. Compared to other predictive approaches, Bayesian models provide a natural framework where the prediction can be conveniently included in the model fitting process, by treating predictions as missing values. Our findings demonstrated the utility of sophisticated Bayesian spatiotemporal dynamic models in supporting state and local opioid surveillance and the ability to provide prediction at a granular geographic level, offering a unique opportunity for preemptive public health and policy interventions, replacing reactionary public health responses. Echoing Borquez and Martin [35], we encourage the field to consider a modeling consortium for opioid-related prediction models, where different modeling techniques could be ensembled.

Acknowledgments

This research was partially supported by funding from the Centers for Disease Control and Prevention (Assessing High Risk Opioid Prescribers and Fatal and Non-Fatal Opioid Overdose grant INTF7311HH2 500224100 awarded to the principal investigator TJS), and the National Institute Drug Abuse (grant R01DA054267 awarded to CB and TJS).

Authors' Contributions

CB and TJS conceived and designed the analysis. TJS contributed to the acquisition of the data. CB and KZ contributed to data processing and curation, conducted the data analysis, and drafted the initial manuscript. WL, MRL, DB, and TJS contributed to critical revision of the manuscript. All authors contributed to the interpretation of the findings, and reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Trends of opioid-related deaths in Massachusetts.
[PDF File (Adobe PDF File), 9073 KB - publichealth_v9i1e41450_app1.pdf]

References

1. Ciccarone D. Fentanyl in the US heroin supply: A rapidly changing risk environment. Int J Drug Policy 2017 Aug;46:107-111 [FREE Full text] [doi: 10.1016/j.drugpo.2017.06.010] [Medline: 28735776]
2. Ciccarone D. The triple wave epidemic: Supply and demand drivers of the US opioid overdose crisis. International Journal of Drug Policy 2019 Sep;71:183-188 [FREE Full text] [doi: 10.1016/j.drugpo.2019.01.010] [Medline: 30718120]
3. Understanding drug overdoses and deaths. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugoverdose/epidemic/index.html [accessed 2023-01-19]
4. Somerville NJ, O'Donnell J, Gladden RM, Zibbell JE, Green TC, Younkin M, et al. Characteristics of Fentanyl Overdose - Massachusetts, 2014-2016. MMWR Morb Mortal Wkly Rep 2017 Apr 14;66(14):382-386 [FREE Full text] [doi: 10.15585/mmwr.mm6614a2] [Medline: 28406883]
5. Scholl L, Seth P, Kariisa M, Wilson N, Baldwin G. Drug and Opioid-Involved Overdose Deaths - United States, 2013-2017. MMWR Morb Mortal Wkly Rep 2018 Jan 04;67(5152):1419-1427 [FREE Full text] [doi: 10.15585/mmwr.mm675152e1] [Medline: 30605448]
6. Ahmad FB, Cisewski JA, Rossen LM, Sutton P. Provisional drug overdose death counts. National Center for Health Statistics. 2023. URL: https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm [accessed 2023-01-19]
7. Watts M. Overdose deaths accelerating COVID-19 pandemic, expanded prevention efforts needed. Florida Department of Health in Paso County. 2021. URL: http://pasco.floridahealth.gov/newsroom/2021/02/overdose-deaths-accelerating-during-covid-19.html [accessed 2023-01-19]
8. AMA issue brief: drug overdose epidemic worsened during COVID pandemic. Alliant Health Solutions. 2021. URL: https://www.alliantquality.org/media_library/ama-issue-brief-drug-overdose-epidemic-worsened-during-covid-pandemic/ [accessed 2023-01-19]
9. Data Brief: Opioid-Related Overdose Deaths among Massachusetts Residents 2022 December. Current opioid statistics | Mass.gov. 2022 Dec. URL: https://www.mass.gov/lists/current-opioid-statistics#updated-data-%E2%80%93-as-of-december-2022- [accessed 2023-01-18]
10. Data brief: opioid-related overdose deaths among Massachusetts residents (2020-11). Massachusetts Department of Public Health. 2020. URL: https://archives.lib.state.ma.us/handle/2452/837612 [accessed 2023-01-19]
11. Number of unintentional opioid-related deaths by county, MA residents: 2000-2015. Massachusetts Department of Public Health. 2016. URL: http://www.mass.gov/eohhs/docs/dph/quality/drugcontrol/county-level-pmp/overdose-deaths-by-county-including-map-may-2016.pdf [accessed 2023-01-19]
12. Bharel M, Bernson D, Averbach A. Using Data to Guide Action in Response to the Public Health Crisis of Opioid Overdoses. NEJM Catalyst 2020 Sep;1(5). [doi: 10.1056/cat.19.1118]
13. Marks C, Carrasco-Escobar G, Carrasco-Hernández R, Johnson D, Ciccarone D, Strathdee SA, et al. Methodological approaches for the prediction of opioid use-related epidemics in the United States: a narrative review and cross-disciplinary call to action. Transl Res 2021 Aug;234:88-113 [FREE Full text] [doi: 10.1016/j.trsl.2021.03.018] [Medline: 33798764]
14. Bozorgi P, Porter DE, Eberth JM, Eidson JP, Karami A. The leading neighborhood-level predictors of drug overdose: A mixed machine learning and spatial approach. Drug Alcohol Depend 2021 Dec 01;229(Pt B):109143. [doi: 10.1016/j.drugalcdep.2021.109143] [Medline: 34794060]
15. Schell RC, Allen B, Goedel WC, Hallowell BD, Scagos R, Li Y, et al. Identifying predictors of opioid overdose death at a neighborhood level with machine learning. Am J Epidemiol 2022 Feb 19;191(3):526-533 [FREE Full text] [doi: 10.1093/aje/kwab279] [Medline: 35020782]
16. Abell-Hart K, Rashidian S, Teng D, Rosenthal RN, Wang F. Where Opioid Overdose Patients Live Far From Treatment: Geospatial Analysis of Underserved Populations in New York State. JMIR Public Health Surveill 2022 Apr 12;8(4):e32133 [FREE Full text] [doi: 10.2196/32133] [Medline: 35412467]
17. Basak A, Cadena J, Marathe A, Vullikanti A. Detection of Spatiotemporal Prescription Opioid Hot Spots With Network Scan Statistics: Multistate Analysis. JMIR Public Health Surveill 2019 Jun 17;5(2):e12110 [FREE Full text] [doi: 10.2196/12110] [Medline: 31210142]
18. Lyle Cooper R, Thompson J, Edgerton R, Watson J, MacMaster SA, Kalliny M, et al. Modeling dynamics of fatal opioid overdose by state and across time. Prev Med Rep 2020 Dec;20:101184 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101184] [Medline: 32995141]

19. Campo DS, Gussler JW, Sue A, Skums P, Khudyakov Y. Accurate spatiotemporal mapping of drug overdose deaths by machine learning of drug-related web-searches. PLoS One 2020 Dec 7;15(12):e0243622 [FREE Full text] [doi: 10.1371/journal.pone.0243622] [Medline: 33284864]
20. Sumetsky N, Mair C, Wheeler-Martin K, Cerda M, Waller LA, Ponicki WR, et al. Predicting the Future Course of Opioid Overdose Mortality: An Example From Two US States. Epidemiology 2021 Jan;32(1):61-69 [FREE Full text] [doi: 10.1097/EDE.0000000000001264] [Medline: 33002963]
21. Bauer C, Champagne-Langabeer T, Bakos-Block C, Zhang K, Persse D, Langabeer JR. Patterns and risk factors of opioid-suspected EMS overdose in Houston metropolitan area, 2015-2019: A Bayesian spatiotemporal analysis. PLoS One 2021 Mar 11;16(3):e0247050 [FREE Full text] [doi: 10.1371/journal.pone.0247050] [Medline: 33705402]
22. Li ZR, Xie E, Crawford FW, Warren JL, McConnell K, Copple JT, et al. Suspected heroin-related overdoses incidents in Cincinnati, Ohio: A spatiotemporal analysis. PLoS Med 2019 Nov 12;16(11):e1002956 [FREE Full text] [doi: 10.1371/journal.pmed.1002956] [Medline: 31714940]
23. Jalal H, Buchanich JM, Roberts MS, Balmert LC, Zhang K, Burke DS. Changing dynamics of the drug overdose epidemic in the United States from 1979 through 2016. Science 2018 Sep 21;361(6408) [FREE Full text] [doi: 10.1126/science.aau1184] [Medline: 30237320]
24. ZCTA level fatal opioid overdose rate in Massachusetts (MA). URL: https://spatiotemporal-data-science.shinyapps.io/MA_Interactive_Map/ [accessed 2023-01-19]
25. American Community Survey (ACS). United States Census Bureau. URL: https://www.census.gov/programs-surveys/acs [accessed 2023-01-19]
26. Knorr-Held L, Besag J. Modelling risk from a disease in time and space. Stat Med 1998 Sep 30;17(18):2045-2060. [doi: 10.1002/(sici)1097-0258(19980930)17:18<2045::aid-sim943>3.0.co;2-p] [Medline: 9789913]
27. Bauer C, Wakefield J, Rue H, Self S, Feng Z, Wang Y. Bayesian penalized spline models for the analysis of spatio-temporal count data. Stat Med 2016 May 20;35(11):1848-1865 [FREE Full text] [doi: 10.1002/sim.6785] [Medline: 26530705]
28. Rue H, Held L. Gaussian Markov Random Fields: Theory and Applications. London: Chapman & Hall; 2005.
29. Besag J. Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical Society: Series B (Methodological) 2018 Dec 05;36(2):192-225. [doi: 10.1111/j.2517-6161.1974.tb00999.x]
30. RStudio is an integrated development environment (IDE) for R. GitHub. URL: https://github.com/rstudio/rstudio [accessed 2023-01-19]
31. Lindgren F, Rue H. Bayesian spatial modelling with R-INLA. J Stat Softw 2015;63(19):1-25. [doi: 10.18637/jss.v063.i19]
32. Sprague Martinez L, Rapkin BD, Young A, Freisthler B, Glasgow L, Hunt T, et al. Community engagement to implement evidence-based practices in the HEALing communities study. Drug Alcohol Depend 2020 Dec 01;217:108326 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108326] [Medline: 33059200]
33. Data Brief: Opioid-Related Overdose Deaths among Massachusetts Residents 2022 December. Massachusetts Department of Public Health. 2022 Dec. URL: https://www.mass.gov/lists/current-opioid-statistics#updated-data-%E2%80%93-as-of-december-2022- [accessed 2023-01-19]
34. Slavova S, Rock P, Bush HM, Quesinberry D, Walsh SL. Signal of increased opioid overdose during COVID-19 from emergency medical services data. Drug Alcohol Depend 2020 Sep 01;214:108176 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108176] [Medline: 32717504]
35. Borquez A, Martin NK. Fatal overdose: Predicting to prevent. Int J Drug Policy 2022 Jun;104:103677 [FREE Full text] [doi: 10.1016/j.drugpo.2022.103677] [Medline: 35550852]
36. Data Brief: Chapter 55 – 2017 Legislative Report. Public Health Data Warehouse (PHD): Publications. 2017. URL: https://www.mass.gov/lists/public-health-data-warehouse-phd-publications [accessed 2022-01-19]
37. U.S. Congress. Senate. Subcommittee on Africa and Global Health Policy of the Committee on Foreign Relations. Zimbabwe After the Elections: Hearing before the Subcommittee on Africa and Global Health Policy of the Committee on Foreign Relations. In: The SHAFR Guide Online. Leiden: BRILL; 2022.
38. COVID-19 Forecasts: Deaths. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/science/forecasting/forecasting-us.html [accessed 2023-01-18]

Abbreviations

CAR: conditional autoregressive model
CrI: credible interval
MAE: mean absolute error
MDPH: Massachusetts Department of Public Health
RD: rank difference
RMSE: root mean square error
RVRS: Registry of Vital Records and Statistics
ZCTA: ZIP Code Tabulation Area

Edited by A Mavragani, H Bradley; submitted 26.07.22; peer-reviewed by Z Li, WHJ Lo-Ciganic, S Linder; comments to author 24.10.22; revised version received 14.12.22; accepted 26.12.22; published 10.02.23.

Please cite as:
Bauer C, Zhang K, Li W, Bernson D, Dammann O, LaRochelle MR, Stopka TJ
Small Area Forecasting of Opioid-Related Mortality: Bayesian Spatiotemporal Dynamic Modeling Approach
JMIR Public Health Surveill 2023;9:e41450
URL: https://publichealth.jmir.org/2023/1/e41450
doi:10.2196/41450
PMID:36763450

Original Paper

Nasopharyngeal Cancer Incidence and Mortality in 185 Countries in 2020 and the Projected Burden in 2040: Population-Based Global Epidemiological Profiling

Yanting Zhang[1], PhD; Harriet Rumgay[2], PhD; Mengmeng Li[3,4], PhD; Sumei Cao[3,4], PhD; Wanqing Chen[5], PhD

[1]Department of Epidemiology and Health Statistics, School of Public Health, Guangdong Medical University, Dongguan, China

[2]Cancer Surveillance Branch, International Agency for Research on Cancer, Lyon, France

[3]Department of Cancer Prevention, Sun Yat-sen University Cancer Center, Guangzhou, China

[4]State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, China

[5]Office of Cancer Screening, National Cancer Center of China/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Corresponding Author:
Yanting Zhang, PhD
Department of Epidemiology and Health Statistics, School of Public Health, Guangdong Medical University
No.1 Xincheng Road
Dongguan, 523808
China
Phone: 86 076922896050
Email: zhangyt@gdmu.edu.cn

Abstract

Background: Nasopharyngeal cancer (NPC) is one of the most common head and neck cancers.

Objective: This study describes the global epidemiological profiles of NPC incidence and mortality in 185 countries in 2020 and the projected burden in 2040.

Methods: The estimated numbers of NPC cases and deaths were retrieved from the GLOBOCAN 2020 data set. Age-standardized incidence rates (ASIRs) and age-standardized mortality rates (ASMRs) were calculated using the world standard. The future number of NPC cases and deaths by 2040 were estimated based on global demographic projections.

Results: Globally, approximately 133,354 cases and 80,008 deaths from NPC were estimated in 2020 corresponding to ASIRs and ASMRs of 1.5 and 0.9 per 100,000 person-years, respectively. The largest numbers of both global cases and deaths from NPC occurred in Eastern Asia (65,866/133,354, 49.39% and 36,453/80,008, 45.56%, respectively), in which China contributed most to this burden (62,444/133,354, 46.82% and 34,810/80,008, 43.50%, respectively). The ASIRs and ASMRs in men were approximately 3-fold higher than those in women. Incidence rates varied across world regions, with the highest ASIRs for both men and women detected in South-Eastern Asia (7.7 and 2.5 per 100,000 person-years, respectively) and Eastern Asia (3.9 and 1.5 per 100,000 person-years, respectively). The highest ASMRs for both men and women were found in South-Eastern Asia (5.4 and 1.5 per 100,000 person-years, respectively). By 2040, the annual number of cases and deaths will increase to 179,476 (46,122/133,354, a 34.58% increase from the year 2020) and 113,851 (33,843/80,008, a 42.29% increase), respectively.

Conclusions: Disparities in NPC incidence and mortality persist worldwide. Our study highlights the urgent need to develop and accelerate NPC control initiatives to tackle the NPC burden in certain regions and countries (eg, South-Eastern Asia, China).

(JMIR Public Health Surveill 2023;9:e49968) doi:10.2196/49968

KEYWORDS

nasopharyngeal cancer; incidence; mortality; epidemiology; worldwide

Introduction

Background

Nasopharyngeal cancer (NPC) is an epithelial carcinoma that arises from the mucosal lining of the nasopharynx. NPC is one of the most common head and neck cancers and is characterized by remarkable geographic variation. Historically, NPC incidence rates were less than 1 per 100,000 person-years in most parts of the world, but they have been higher than 20 per 100,000 person-years in South-Eastern Asia and Southern China in the past decades [1-3]. In contrast, NPC incidence and mortality rates in most countries worldwide have significantly decreased over the past decades [4,5]. Notably, NPC is usually diagnosed in advanced stages and has a very poor prognosis [6]. For example, the 5-year net survival is estimated to be 47% in China [7].

Risk Factors Associated With NPC Occurrence

NPC is a largely preventable disease due to its many modifiable risk factors. The major etiological factors for NPC include Epstein-Barr virus (EBV) infection, tobacco smoking, intake of salted fish and other salt-preserved food, and occupational exposure to wood dust [8]. EBV infection has been consistently identified as an important risk factor, with a dose-response relationship between EBV antibody level and NPC risk [9,10]. The population attributable fraction for NPC incidence due to EBV infection has been estimated to be 85% [11]. Compared with people who never smoke, people who formerly smoked have a 60% greater risk of developing NPC [12]. The relative risk of NPC associated with weekly consumption of Chinese-style salt-preserved fish ranges from 1.1 to 4, whereas that associated with daily consumption ranges from 1.8 to 20 compared with no or rare consumption [13,14]. Occupational exposure to wood dust is estimated to be associated with a 50% increased risk of NPC [15].

Research Significance and Objective

Given the strong association of NPC with its modifiable risk factors and the changing epidemiological profile due to trends in NPC incidence and mortality, understanding the current epidemiological profile of international variations in NPC incidence and mortality is essential. This would allow public health policy makers to make evidence-based decisions for primary prevention and optimize the allocation of resources to reduce the global burden of NPC. Considering the growing and aging global population, predicting the future NPC burden is vital for better planning of future cancer control programs. We therefore examined the geographic variations in NPC incidence and mortality across world regions and countries worldwide based on the GLOBOCAN estimates for the year 2020 and predict the future NPC burden up to 2040 based on demographic projections.

Methods

Data Sources

The numbers of new cases of and deaths from NPC (International Classification of Diseases, tenth revision C11) were extracted from the GLOBOCAN 2020 database for 185 countries or territories by sex and by 5-year age groups (0-4, 5-9, …, 80-84, 85 years and older) [16-18]. Corresponding population data for 2020 were retrieved from the United Nations website [19]. The population projections used in this study are based on the future fertility rates [19]. The data sources and methods used in compiling the global cancer estimates for 2020 have been described in detail elsewhere [17]. Briefly, the GLOBOCAN national estimates are dependent on the availability of recorded high-quality national and subnational incidence (from population-based cancer registries) and national mortality data (from vital registration systems) [17]. Nine methods were employed depending on the best available data on cancer-specific incidence or mortality data [17]. The hierarchical set of methods includes utilization of short-term and long-term prediction models based on historical observed data, estimated incidence, mortality based on modelled survival, approximation using observed data at the subnational level, or data from neighboring countries [17]. In countries where national mortality data were available but national or subnational cancer registries were not, national incidence estimation relied on national mortality estimates and modelling of the mortality to incidence ratio from neighboring countries [17]. In countries where neither mortality nor incidence data were available, incidence and mortality estimates were based on mortality or incidence data of neighboring countries [17]. Thus, the validity of national cancer incidence and mortality estimates is dependent on the degree of representativeness and the quality of the source information [17]. The methods used to derive the 2020 estimates correspond to those used previously for 2018, 2012, and 2008 [20-22].

Statistical Analysis

We present tables and figures of the estimated new cases and deaths as well as 2 summary measures by using direct standardization, namely, the age-standardized incidence rates (ASIRs) and age-standardized mortality rates (ASMRs) per 100,000 person-years based on the 1966 Segi–Doll World standard population [23,24] and the cumulative risk of being diagnosed with or dying from NPC before the age of 75 years, assuming the absence of competing causes of death [25]. We predicted the future number of NPC cases and deaths worldwide by the United Nation’s 4-tier Human Development Index (HDI), where HDI was used to assess the cancer burden at varying levels of development (low, medium, high, and very high HDI) [26] and in China up to the year 2040 based on demographic projections and scenarios of annually increasing (+1%, +2%, +3%, +4%), stable (0%), or decreasing (–1%, –2%, –3%, –4%) rates from the baseline year of 2020. We did not use scenarios of rates changing by SD 5% or more because such changes would be unlikely to occur in real life [4,5]. Predictions were calculated by applying the age-specific rates for the year 2020 (and each of the increasing or decreasing scenarios described) to the corresponding projected population data as estimated by the United Nations Development Program. The results are presented by country and aggregated across 20 United Nations–defined world regions [19] and according to the HDI group in 2020. Data management and analyses were performed in R software (version 4.0.2; R Foundation for Statistical

Computing) [27]. Figures were plotted using SigmaPlot software (version 12.5; Systat Software Inc) [28]. Global maps of NPC incidence and mortality rates by country are shown using R software (version 4.0.2; R Foundation for Statistical Computing) [27].

Ethical Considerations

This study does not involve human participants and animals. Ethics approval was not required for this study, as this study used existing nonidentifiable data that were aggregated at the population level.

Results

Global Burden of NPC Incidence and Mortality

In 2020, an estimated 133,354 people were diagnosed with NPC worldwide, corresponding to an ASIR of 1.5 per 100,000 person-years (Table 1). More men (96,371 cases) than women (36,983 cases) were diagnosed with NPC, and the ASIRs in men were approximately 3-fold higher than those in women (2.2 vs 0.8 per 100,000 person-years, respectively) (Multimedia Appendix 1). Globally, an estimated 80,008 people died from NPC, corresponding to an ASMR of 0.9 per 100,000 person-years (Table 1). Mortality in men was also higher than that in women, with 58,094 and 21,914 deaths, corresponding to ASMRs of 1.3 and 0.5 per 100,000 person-years, respectively (Multimedia Appendix 1). In addition, the cumulative risk of being diagnosed with and dying from NPC before the age of 75 years was 1 in 476 (0.21%) and 1 in 667 (0.15%), respectively (Table 1).

Table 1. Nasopharyngeal cancer incidence and mortality in both sexes combined in 2020 by world region and human development index level.

	Population (N=7,794,799)	Incidence (n=133,354)			Mortality (n=80,008)		
	Values (in thousands), n (%)	Cases, n (%)	ASIR[a]	Cumulative risk[b]	Deaths, n (%)	ASMR[c]	Cumulative risk[b]
Europe							
Northern Europe	106,261 (1.4)	415 (0.3)	0.26	0.04	247 (0.3)	0.12	0.03
Western Europe	196,146 (2.5)	1304 (1)	0.40	0.06	502 (0.6)	0.12	0.03
Southern Europe	153,423 (2)	1584 (1.2)	0.64	0.09	746 (0.9)	0.24	0.05
Central and Eastern Europe	293,013 (3.8)	1901 (1.4)	0.43	0.06	1091 (1.4)	0.22	0.04
America							
Northern America	368,870 (4.7)	2177 (1.6)	0.41	0.06	1071 (1.3)	0.16	0.04
South America	430,760 (5.5)	1423 (1.1)	0.28	0.05	797 (1)	0.15	0.03
Central America	179,670 (2.3)	309 (0.2)	0.17	0.03	181 (0.2)	0.10	0.02
Caribbean	43,532 (0.6)	313 (0.2)	0.56	0.10	198 (0.3)	0.34	0.07
Asia							
Eastern Asia	1,678,090 (21.5)	65,866 (49.4)	2.70	0.35	36,453 (45.6)	1.40	0.23
China	1,447,470 (18.6)	62,444 (46.8)	3.00	0.40	34,810 (43.5)	1.60	0.27
South-Central Asia	2,014,709 (25.8)	8366 (6.3)	0.43	0.08	6117 (7.7)	0.32	0.06
South-Eastern Asia	668,620 (8.6)	36,747 (27.6)	5.00	0.75	24,219 (30.3)	3.30	0.59
Western Asia	278,429 (3.6)	2680 (2)	1.00	0.15	1645 (2.1)	0.63	0.11
Oceania							
Australia and New Zealand	30,322 (0.4)	176 (0.1)	0.42	0.06	101 (0.1)	0.19	0.04
Melanesia	11,123 (0.1)	22 (0.1)	0.25	0.03	14 (0.1)	0.17	0.02
Micronesia/Polynesia	1233 (<0.1)	30 (0.1)	2.20	0.24	26 (0.1)	1.93	0.21
Africa							
Northern Africa	246,233 (3.2)	3525 (2.6)	1.60	0.25	2113 (2.6)	0.98	0.20
Western Africa	401,861 (5.2)	1906 (1.4)	0.70	0.10	1289 (1.6)	0.53	0.09
Southern Africa	67,504 (0.9)	212 (0.2)	0.34	0.07	145 (0.2)	0.25	0.06
Middle Africa	179,595 (2.3)	1212 (0.9)	1.10	0.15	852 (1.1)	0.86	0.13
Eastern Africa	445,406 (5.7)	3186 (2.4)	1.10	0.20	2201 (2.8)	0.88	0.18
Human Development Index							
Very high human development index	1,564,286 (20.1)	14,911 (11.2)	0.66	0.09	7828 (9.8)	0.30	0.05
High human development index	2,909,468 (37.3)	93,153 (69.9)	2.50	0.35	54,850 (68.6)	1.40	0.24
Medium human development index	2,327,556 (29.9)	19,543 (14.7)	0.89	0.15	13,314 (16.6)	0.62	0.12
Low human development index	990,175 (12.7)	5722 (4.3)	0.89	0.15	3996 (5)	0.69	0.13
World	7,794,799 (100)	133,354 (100)	1.50	0.21	80,008 (100)	0.88	0.15

[a] ASIR: age-standardized incidence rate per 100,000 person-years.

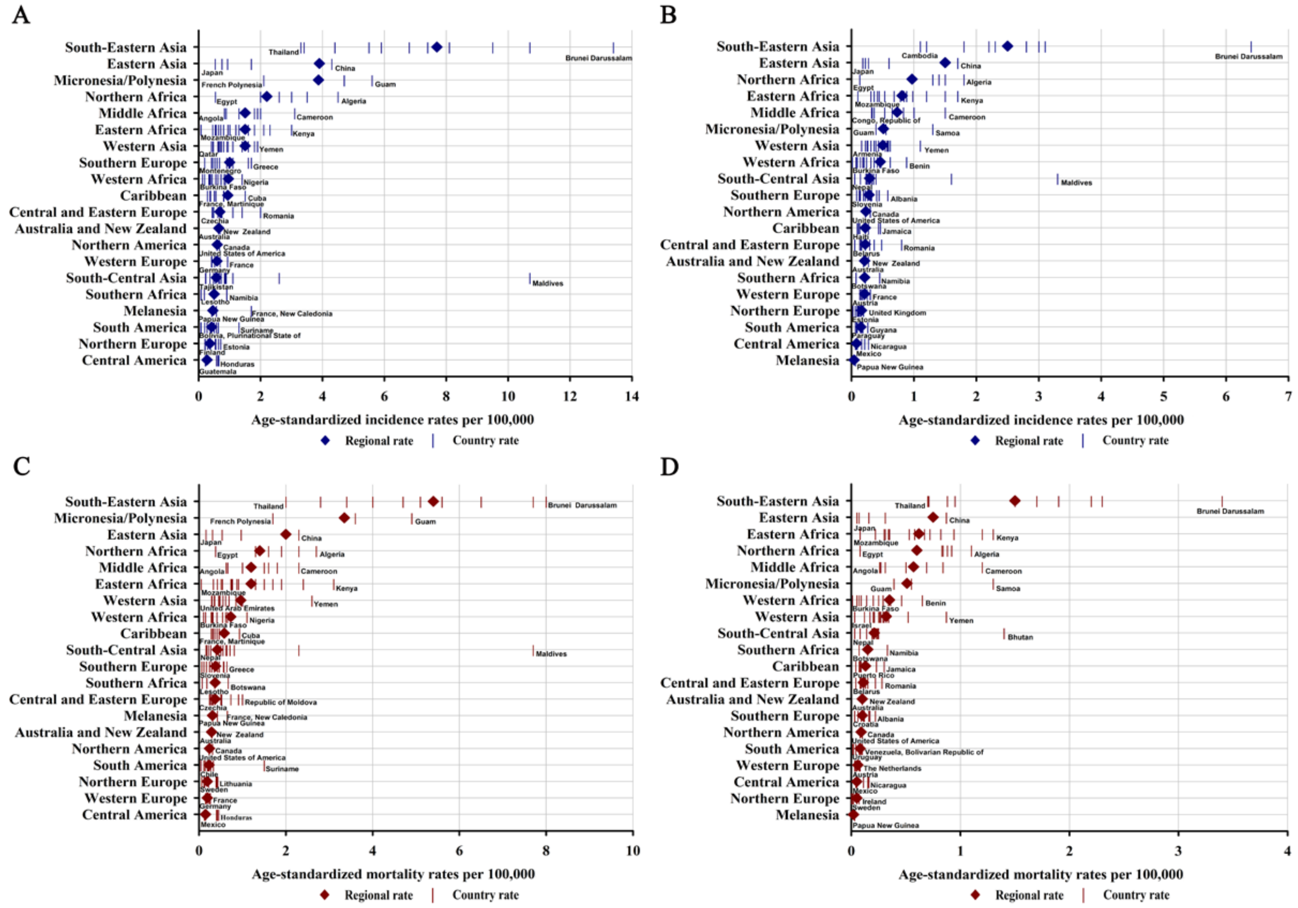
[b] Cumulative risk of being diagnosed with or dying from nasopharyngeal cancer before the age of 75 years in 2020.

[c] ASMR: age-standardized mortality rate per 100,000 person-years.

Geographical Variations in NPC Incidence and Mortality by World Region

The largest numbers of cases and deaths from NPC in both sexes combined were estimated in Eastern Asia (65,866/133,354, 49.39% of total NPC cases and 36,453/80,008, 45.56% of total NPC deaths), followed by South-Eastern Asia (36,747/133,354, 27.55% and 24,219/80,008, 30.15%, respectively) and South-Central Asia (8366/133,354, 6.27% and 6117/80,008, 7.64%, respectively) (Multimedia Appendix 2). A male predominance in NPC cases and deaths was consistent across all world regions (Multimedia Appendix 1). The incidence rates of NPC showed approximately 29-fold variation in men and 63-fold variation in women across world regions (Figure 1A, Figure 1B, and Multimedia Appendix 1). In men, the ASIR per 100,000 person-years was the highest in South-Eastern Asia (7.7), followed by Eastern Asia (3.9) and Micronesia/Polynesia (3.9), but the lowest ASIR was in Central America (0.3). In women, the highest ASIR per 100,000 person-years was detected in South-Eastern Asia (2.5), followed by Eastern Asia (1.5), and Northern Africa (1.0), and the lowest was found in Melanesia (less than 0.1). The mortality rates of NPC varied approximately by 36-fold in men and 75-fold in women among world regions (Figure 1C, Figure 1D, and Multimedia Appendix 1). In men, the highest ASMR per 100,000 person-years was detected in South-Eastern Asia (5.4), followed by Micronesia/Polynesia (3.4) and Eastern Asia (2.0), and the lowest was observed in Central America (0.2). In women, the highest ASMR per 100,000 person-years was found in South-Eastern Asia (1.5), followed by Eastern Asia (0.8) and Eastern Africa (0.6), and the lowest was seen in Melanesia (<0.1). Sex-related disparities were also noted, with NPC ASIRs and ASMRs higher in men than in women across world regions. For example, the male-to-female ASIR and ASMR ratios ranged from 1.9 and 1.9 in Eastern Africa to 11.5 and 15.5 in Melanesia, respectively.

Figure 1. Age-standardized rates per 100,000 person-years of nasopharyngeal cancer incidence in (A) men and (B) women and mortality in (C) men and (D) women, which are ordered according to descending incidence rates by world region. Countries with zero cases are not shown in this figure.

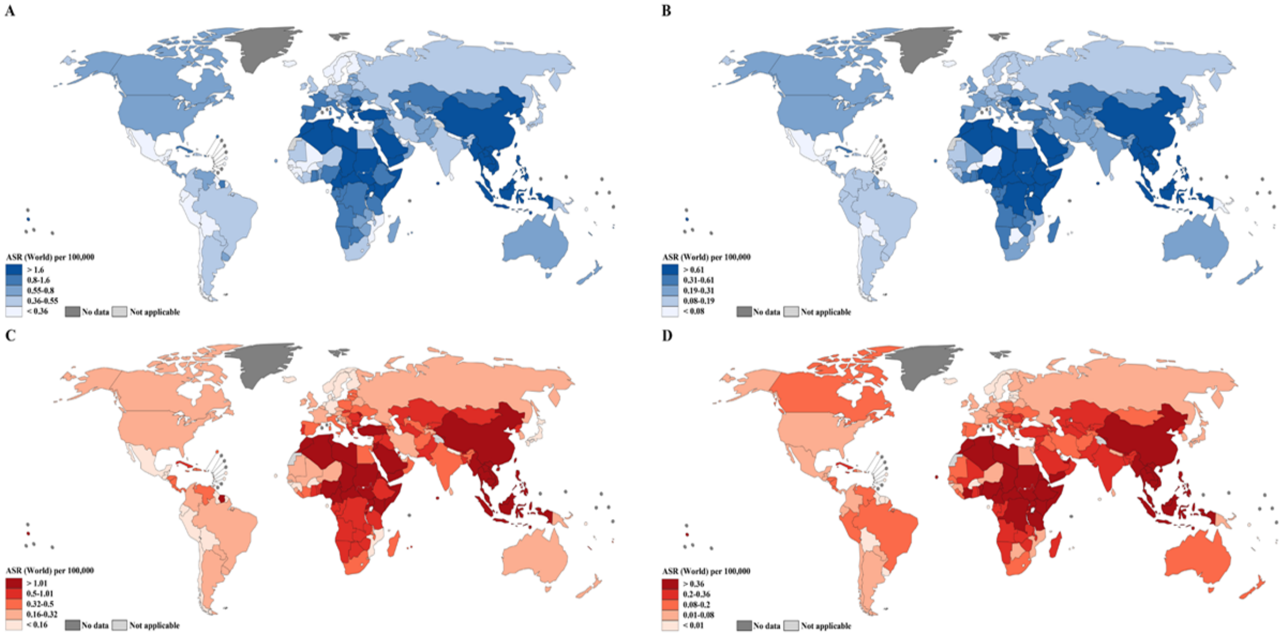


Geographical Variations in NPC Incidence and Mortality by Country

At the national level, China (62,444/133,354, 46.82% of global NPC cases; 34,810/80,008, 43.50% of global NPC deaths) was the greatest contributor to the global NPC burden due to its large population (1,447,470,000/7,794,799,000, 18.56% of the global population) and high incidence and mortality rates (ASIR, 3.0 per 100,000 person-years; ASMR, 1.6 per 100,000 person-years) (Table 1). In terms of incidence rates, the highest ASIR per 100,000 person-years occurred in Brunei Darussalam (13.4), followed by Maldives (10.7) and Indonesia (10.7), in men and in Brunei Darussalam (6.4), followed by Maldives (3.3) and Malaysia (3.1), in women (Figure 1A, Figure 1B, Figure 2A, and Figure 2B). Notably, the incidence rates varied markedly within world regions. For example, in the high-risk region South-Eastern Asia, the ASIRs ranged from 3.3 per 100,000 person-years in Thailand to 13.4 per 100,000 person-years in Brunei Darussalam (high-to-low ASIR ratio = 4.1) in men and from 1.1 in Cambodia to 6.4 in Brunei

Darussalam (high-to-low ASIR ratio = 5.8) in women (Figure 1A and Figure 1B). In terms of mortality rates, the highest ASMR per 100,000 person-years in men was found in Brunei Darussalam (8.0), followed by Maldives (7.7) and Indonesia (7.7), and the highest ASMR per 100,000 person-years in women was found in Brunei Darussalam (3.4), followed by Timor-Leste (2.3) and Lao People's Democratic Republic (2.2) (Figure 1C, Figure 1D, Figure 2C, and Figure 2D). Considerable variations in NPC mortality were also evident within world regions. For example, in South-Eastern Asia, the ASMRs per 100,000 person-years ranged from 2.0 in Thailand to 8.0 in Brunei Darussalam (high-to-low ASMR ratio = 4.0) in men and from 0.7 in Thailand to 3.4 in Brunei Darussalam (high-to-low ASMR ratio = 4.9) in women (Figure 1C and Figure 1D).

Figure 2. Global map of age-standardized rates of nasopharyngeal cancer incidence in (A) men and (B) women and mortality in (C) men and (D) women per 100,000 person-years by country. ASR: age-standardized rate.



NPC Incidence and Mortality by Level of Human Development

By HDI group, the vast majority of NPC cases and deaths occurred among 37.32% (2,909,468,000/7,794,799,000) of the world population living in high HDI countries, representing 69.85% (93,153/133,354) of new cases and 68.55% (54,850/80,008) of deaths globally (Table 1). Both ASIRs and ASMRs were the highest in high HDI countries, with the rates being 2.5 and 1.4 per 100,000 person-years, respectively. Notably, both the lowest ASIRs and ASMRs were observed in very high HDI countries. The ASIRs and ASMRs in low HDI countries were similar to those in medium HDI countries.

Predicted Number and Percentage Increase of Cases and Deaths From NPC

Worldwide, an estimated 179,476 new NPC cases are projected to occur in 2040—an increase of approximately 34.58% (46,122/133,354)—compared with 133,354 cases in 2020, assuming that global incidence rates in 2020 remain stable (Figure 3A). Moreover, a 2% annual increase in incidence rates from 2020 would more than double the total annual NPC cases by 2040 (Figure 3A). Notably, it would require a 2% annual decrease in mortality rates to ensure that there would be fewer NPC cases in 2040 (119,710 cases) than there were in 2020 (133,354 cases). In terms of mortality, NPC deaths were estimated to increase by approximately 42.29% (33,843/80,008), from 80,008 in 2020 to 113,851 in 2040, assuming that mortality rates in 2020 remained unchanged (Figure 3B). An annual decline of 2% in mortality rates would be required to achieve fewer NPC deaths in 2040 compared to the estimated deaths in 2020 (Figure 3B). By HDI group, the largest absolute increase in NPC cases and deaths is estimated to occur in high HDI countries, with 26.7% (24,874/93,153) more cases (24,874 additional cases) and 40.15% (22,024/54,850) more deaths (22,024 additional deaths) per year by 2040 (Multimedia Appendix 3), assuming that incidence and mortality rates in 2020 remain constant, reflecting the already high rates in high HDI countries and its large population, which will continue to grow. However, the greatest relative increases in cases and deaths will occur in low HDI countries (5181/5722, 90.54% and 3808/3996, 95.29% increase, respectively) (Multimedia Appendix 3). Notably, decreases in both incidence and mortality rates would need to be greater than 4%, 3%, 2%, and 1% to retain the predicted NPC burden in 2040 at the level observed in 2020 for low, medium, high, and very high HDI countries, respectively (Figures 4-5).

Figure 3. Predicted number of (A) new cases and (B) deaths from nasopharyngeal cancer, assuming 9 scenarios of annual change in global rates between 2020 and 2040, both sexes combined. Nine scenarios include annually increasing (+1%, +2%, +3%, +4%), stable (0%), and decreasing (–1%, –2%, –3%, –4%) rates from the baseline year of 2020.

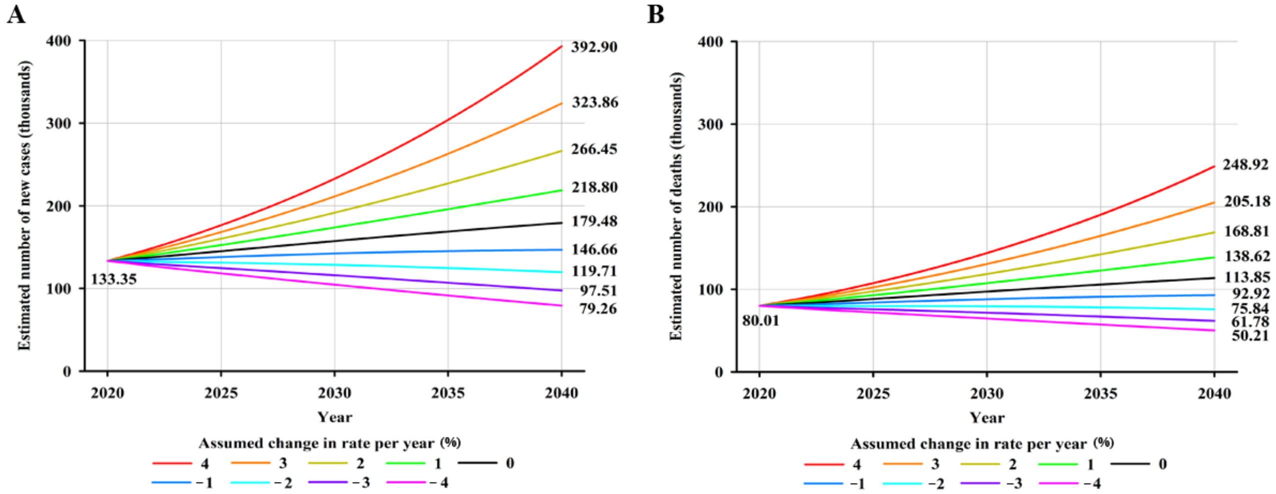


Figure 4. Predicted number of new nasopharyngeal cancer cases, assuming 9 scenarios of annual change in rates between 2020 and 2040, in (A) countries with very high human development index, (B) countries with high human development index, (C) countries with medium human development index, and (D) countries with low human development index, both sexes combined. Nine scenarios consist of annually increasing (+1%, +2%, +3%, +4%), stable (0%), and decreasing (–1%, –2%, –3%, –4%) rates from the baseline year of 2020.

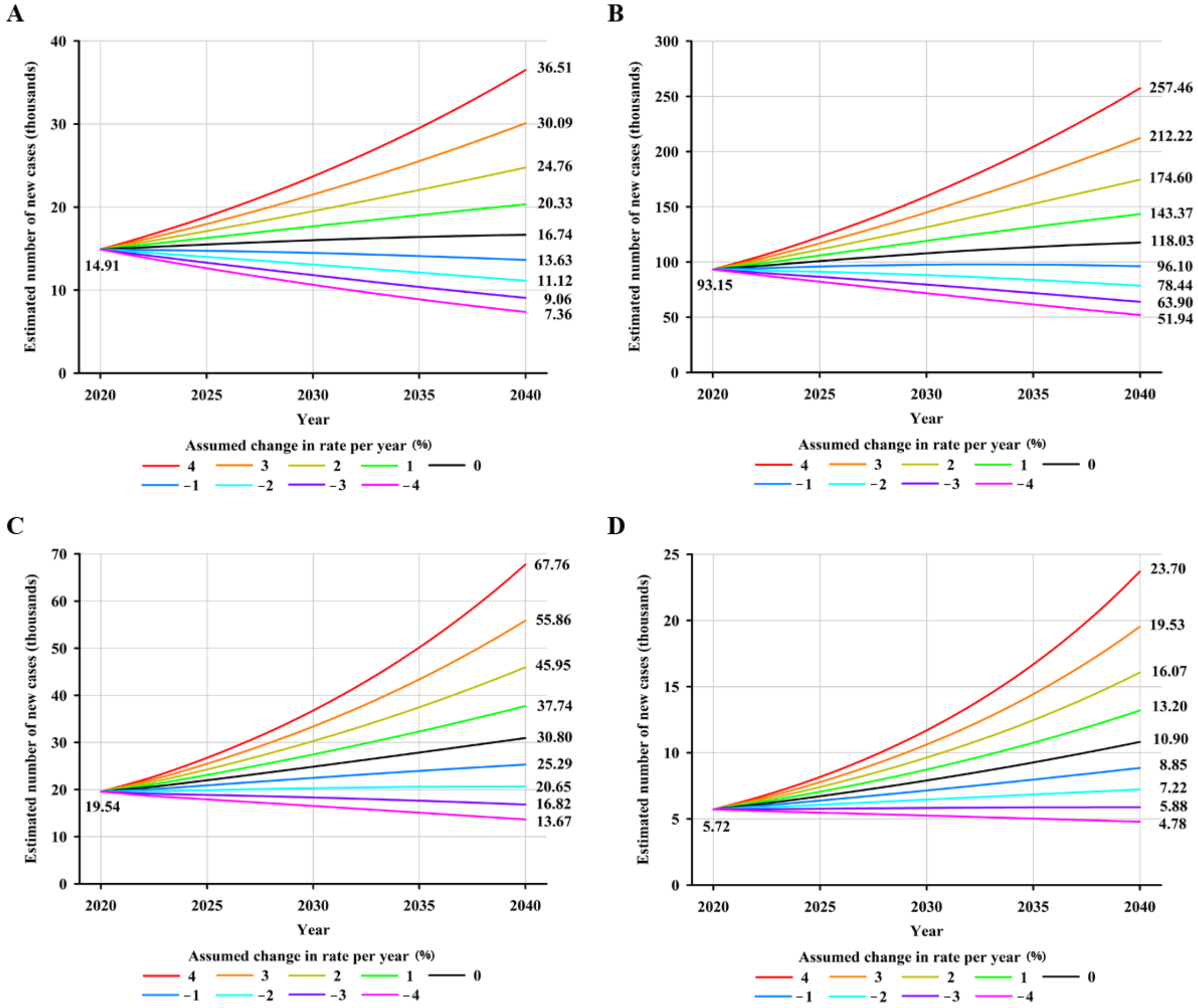
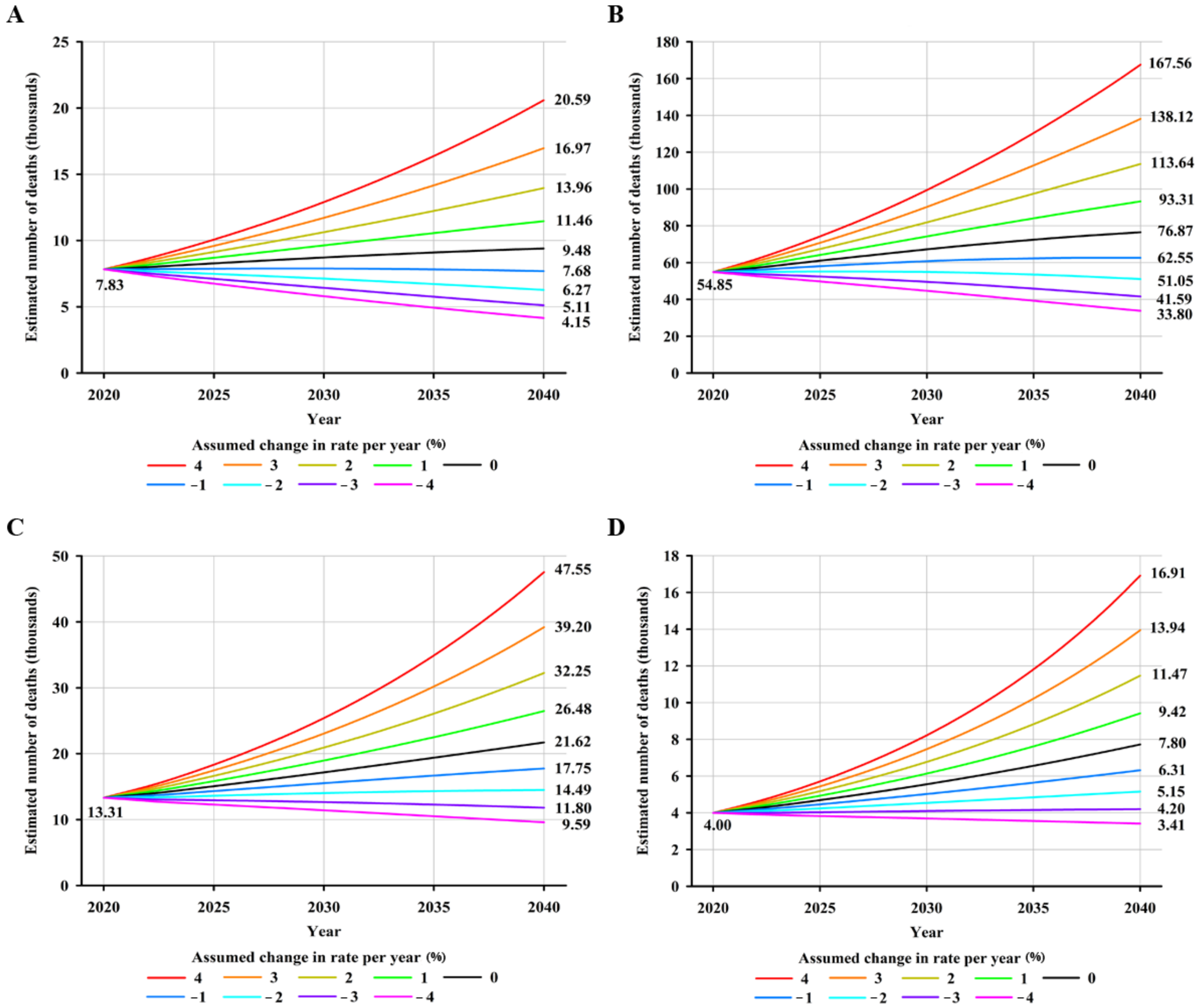
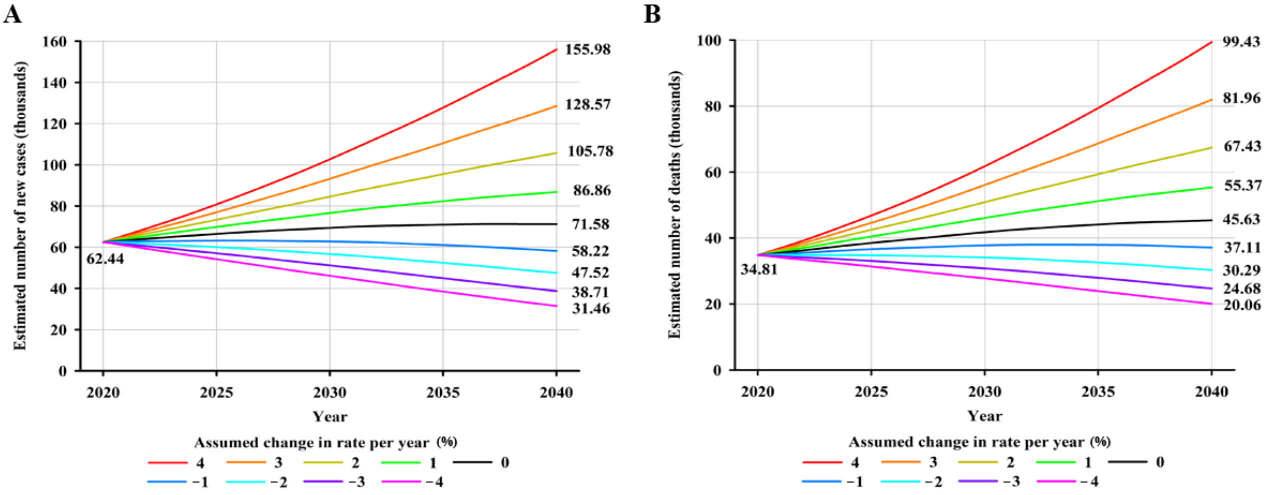


Figure 5. Predicted number of deaths from nasopharyngeal cancer assuming 9 scenarios of annual change in rates between 2020 and 2040 in (A) countries with very high human development index, (B) countries with high human development index, (C) countries with medium human development index, and (D) countries with low human development index, both sexes combined. Nine scenarios consist of annually increasing (+1%, +2%, +3%, +4%), stable (0%), and decreasing (–1%, –2%, –3%, –4%) rates from the baseline year of 2020.



In China, the number of NPC cases was estimated to increase by approximately 13.99% (8742/62,444), from 62,444 in 2020 to 71,186 in 2040, assuming that incidence rates in 2020 remained stable (Figure 6A). Similarly, NPC deaths are set to increase by close to 30.35% (10,568/34,810) based on demographic changes alone, from 34,810 in 2020 to 45,378 in 2040 (Figure 6B). As an illustration, it would take greater than 1% and 2% decline in the annual incidence and mortality rate from 2020 to 2040 to achieve fewer cases and deaths in 2040 compared to the level seen in 2020, respectively.

Figure 6. Predicted number of (A) new cases and (B) deaths from nasopharyngeal cancer assuming 9 scenarios of annual change in rates between 2020 and 2040 in China (both sexes combined). Nine scenarios consist of annually increasing (+1%, +2%, +3%, +4%), stable (0%), and decreasing (–1%, –2%, –3%, –4%) rates from the baseline year of 2020.



Discussion

Principal Findings

Globally, approximately 133,354 new NPC cases were reported and 80,008 deaths occurred in 2020 [16-19]. The largest numbers of cases and deaths of NPC were estimated in Eastern Asia, in which China contributed most to this burden [29]. NPC incidence and mortality rates in men were substantially higher than those in women across all world regions. The highest incidence rates for both men and women were detected in South-Eastern Asia and Eastern Asia, and the mortality rates for both men and women were the highest in South-Eastern Asia. The number of cases and deaths from NPC will increase by more than 35% over the next 20 years worldwide due to population growth and aging alone, with annual NPC cases and deaths expected to almost double in low HDI countries by 2040 [26].

Interpretation of the Disparities of NPC Burden by Geography and by Gender

The changing epidemiological profile and the disparities of NPC burden by geography and by gender might be largely associated with differences in the prevalence of its risk factors and improvement of diagnostic and treatment techniques. Previous studies have shown that NPC incidence and mortality rates have significantly decreased in most European, Northern American, and Eastern and South-Eastern Asian countries over the past decades [4,5]. The declining trends in NPC incidence rates in Europe and Northern America might be related to decreased tobacco smoking prevalence, and the declines in NPC incidence rates in Eastern and South-Eastern Asia might be explained by the decreased intake of salted fish and preserved food [4], while decreases in mortality rates might also be the result of advances in diagnostic (eg, diagnostic imaging accuracy) and radiotherapy techniques (eg, the introduction of intensity-modulated radiation therapy) [5]. In particular, the reduction of NPC burden in China is partly due to the implementation of EBV screening for early detection of NPC since the 1970s [29].

The relatively higher NPC burden in Eastern and South-Eastern Asian countries could be largely related to the high prevalence of EBV infection [30]. In addition, 316 million adults are estimated to smoke in China, who account for nearly one-third of the Chinese who smoke and contribute to 40% of the tobacco consumption worldwide [31]. The higher NPC burden in China could therefore be partly explained by its large tobacco consumption and production. The relatively larger consumption of salt-preserved foods in most Eastern and South-Eastern Asian and Northern African countries might also explain part of the remarkable international variations of NPC burden [2]. The elevated NPC burden observed in men compared with that observed in women across all regions can be partly explained by the higher smoking prevalence and occupational exposure to carcinogens [32,33]. In addition, the decreased NPC burden in women could be related to the potentially protective effect of endogenous estrogens [33,34]. Of note, genetics might also play an important role in the pathogenesis of NPC. Previous studies showed that NPC incidence in Southern China is about 20-50 times higher than that in western countries [1-3]. Despite immigration of second- and third-generation Chinese people to western countries, they are still at a higher risk for NPC, thereby demonstrating a sustained risk for NPC even after changing environments [35,36].

Interpretation of the Projected Number of Cases and Deaths From NPC

Our findings show that the projected number of cases and deaths from NPC are expected to increase to 179,476 and 113,851 by 2040 worldwide, respectively, as a result of population growth and aging alone. A 2% annual decrease in global incidence and mortality rates would be needed to halt the increasing NPC burden by 2040. Notably, our findings show that even larger decreases in incidence and mortality rates would be required to reduce the future NPC burden in low and medium HDI countries. Considering these changes, the reallocation of resources for NPC primary prevention programs aimed at reducing population levels of EBV infection, tobacco smoking, intake of salted fish and other salt-preserved food, and occupational exposure to NPC carcinogens in certain regions and countries and the increased access to early detection modalities (eg, NPC screening) and health care services for high-risk populations identified by our study are crucial for reducing the global NPC burden.

Comparison With Prior Work

To the best of our knowledge, this study is the first to provide a comprehensive picture of the most up-to-date epidemiological profile of NPC incidence and mortality on a global scale based on the GLOBOCAN 2020 data set, which is highly relevant for cancer control and clinical practice. The numbers and rates of NPC presented in this study are estimates based on the best available data (reviewed for their completeness, coverage, and accuracy) from population-based cancer registries.

Limitations

There are several limitations in this study. First, although our findings are based on the best available and high-quality data, caution is warranted when interpreting the findings for countries with limited coverage from population-based cancer registries and where proxy data were used to obtain national estimates [17]. Second, the GLOBOCAN estimates did not account for the impact of the COVID-19 pandemic on cancer diagnoses because the GLOBOCAN estimates were based on extrapolations of previous years of cancer data [17]. Third, the projections of the future burden of NPC in 2040 considered neither the recent changes in NPC incidence and mortality rates nor the heterogeneity in NPC incidence and mortality trends among countries. Thus, the predictions in our study likely represent an overestimate of the future NPC burden, given the declining trends of NPC incidence and mortality rates in recent decades worldwide [4,5] and should be interpreted with caution. Finally, we could not distinguish the required changes in the modifiable risk factors and their corresponding effects on predicted NPC incidence and mortality rates, which suggests the need for further studies to clarify the impact of changes in exposure to modifiable risk factors on the burden of NPC.

Conclusions

NPC remains a considerable public health challenge worldwide. NPC incidence and mortality vary markedly across world regions and countries and between men and women, and this variation is likely related to differences in EBV infection, tobacco smoking, intake of salted fish and other salt-preserved food, and occupational exposure to NPC carcinogens. Given the remarkable geographic disparity of NPC burden across world regions and countries, this study highlights the urgent need to formulate more effective primary NPC prevention strategies and prioritize allocation of NPC prevention and treatment resources for high-risk populations to tackle the NPC burden in certain regions and countries.

Acknowledgments

We gratefully acknowledge the contribution of all cancer registry staff in generating the cancer registry data used to create the cancer incidence and mortality estimates in the GLOBOCAN database. Where authors are identified as personnel of the International Agency for Research on Cancer and the World Health Organization (WHO), the authors alone are responsible for the views expressed in this paper, and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer and WHO. YZ is supported by funds for PhD researchers of Guangdong Medical University (grant 4SG23187G). ML is supported by Science and Technology Projects in Guangzhou (grant 202201011243). All funding parties did not have any role in the design of this study or in the explanation of the data.

Data Availability

Data are available in the Global Cancer Observatory, a public open access repository [37]. Processed data sets used in this analysis are available from the corresponding author upon reasonable request.

Authors' Contributions

YZ contributed to conceptualization, supervision, project administration, data curation, formal analysis, methodology, software visualization, writing the original draft, review, and editing. HR contributed to validation, software visualization, writing, review, and editing. ML, SC, and WC contributed to review and editing of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Nasopharyngeal cancer incidence and mortality in 2020 by sex, world region, and human development index level.

[DOCX File , 23 KB - publichealth_v9i1e49968_app1.docx]

Multimedia Appendix 2

Distribution of (A) nasopharyngeal cancer cases and (B) deaths by world region in 2020, both sexes combined.

[PNG File , 134 KB - publichealth_v9i1e49968_app2.png]

Multimedia Appendix 3

Predicted percentage change (absolute numbers are shown above bars) of (A) new cases and (B) deaths from nasopharyngeal cancer (both sexes combined) between 2020 and 2040 globally and by HDI, assuming that incidence rates and mortality rates in 2020 remained unchanged. HDI: Human Development Index.

[PNG File , 74 KB - publichealth_v9i1e49968_app3.png]

References

1. Parkin D, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55(2):74-108 [FREE Full text] [doi: 10.3322/canjclin.55.2.74] [Medline: 15761078]
2. Chang E, Adami H. The enigmatic epidemiology of nasopharyngeal carcinoma. Cancer Epidemiol Biomarkers Prev 2006 Oct;15(10):1765-1777. [doi: 10.1158/1055-9965.EPI-06-0353] [Medline: 17035381]
3. Wei K, Zheng R, Zhang S, Liang Z, Li Z, Chen W. Nasopharyngeal carcinoma incidence and mortality in China, 2013. Chin J Cancer 2017 Nov 09;36(1):90 [FREE Full text] [doi: 10.1186/s40880-017-0257-9] [Medline: 29122009]
4. Tang L, Chen W, Xue W, He Y, Zheng R, Zeng Y, et al. Global trends in incidence and mortality of nasopharyngeal carcinoma. Cancer Lett 2016 Apr 28;374(1):22-30. [doi: 10.1016/j.canlet.2016.01.040] [Medline: 26828135]
5. Carioli G, Negri E, Kawakita D, Garavello W, La Vecchia C, Malvezzi M. Global trends in nasopharyngeal cancer mortality since 1970 and predictions for 2020: Focus on low-risk areas. Int J Cancer 2017 May 15;140(10):2256-2264 [FREE Full text] [doi: 10.1002/ijc.30660] [Medline: 28224615]
6. Wei WI, Sham JS. Nasopharyngeal carcinoma. The Lancet 2005 Jun;365(9476):2041-2054. [doi: 10.1016/s0140-6736(05)66698-6]
7. Soerjomataram I, Cabasag C, Bardot A, Fidler-Benaoudia MM, Miranda-Filho A, Ferlay J, et al. Cancer survival in Africa, Central and South America, and Asia (SURVCAN-3): a population-based benchmarking study in 32 countries. The Lancet Oncology 2023 Jan;24(1):22-32. [doi: 10.1016/s1470-2045(22)00704-5]
8. United Nations. URL: https://esa.un.org/unpd/wpp/Publications/Files/Key_Findings_WPP_2015.pdf [accessed 2022-12-30]
9. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Biological agents. IARC Monogr Eval Carcinog Risks Hum 2012;100(Pt B):1-441. [Medline: 23189750]
10. Chan CK, Mueller N, Evans A, Harris NL, Comstock GW, Jellum E, et al. Epstein-Barr virus antibody patterns preceding the diagnosis of nasopharyngeal carcinoma. Cancer Causes Control 1991 Mar;2(2):125-131. [doi: 10.1007/bf00053132]
11. de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. The Lancet Global Health 2020 Feb;8(2):e180-e190. [doi: 10.1016/s2214-109x(19)30488-7]
12. Xue W, Qin H, Ruan H, Shugart YY, Jia W. Quantitative association of tobacco smoking with the risk of nasopharyngeal carcinoma: a comprehensive meta-analysis of studies conducted between 1979 and 2011. Am J Epidemiol 2013 Aug 01;178(3):325-338 [FREE Full text] [doi: 10.1093/aje/kws479] [Medline: 23785114]
13. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Personal habits and indoor combustions. IARC Monogr Eval Carcinog Risks Hum 2012;100(Pt E):1-538. [Medline: 23193840]
14. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Some naturally occurring substances: food items and constituents, heterocyclic aromatic amines and mycotoxins. In: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 56. Lyon, France: International Agency for Research on Cancer (IARC); 1993.
15. Beigzadeh Z, Pourhassan B, Kalantary S, Golbabaei F. Occupational exposure to wood dust and risk of nasopharyngeal cancer: A systematic review and meta-analysis. Environ Res 2019 Apr;171:170-176. [doi: 10.1016/j.envres.2018.12.022] [Medline: 30677637]
16. Sung H, Ferlay J, Siegel R, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
17. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Online ahead of print. Int J Cancer 2021 Apr 05:1-12 [FREE Full text] [doi: 10.1002/ijc.33588] [Medline: 33818764]
18. Ferlay J, Ervik M, Lam F. Global cancer observatory: cancer today. International Agency for Research on Cancer. URL: https://gco.iarc.fr/today/home [accessed 2022-12-30]
19. Standard country or area codes for statistical use (M49). United Nations Statistics Division. URL: https://unstats.un.org/unsd/methodology/m49/ [accessed 2022-12-30]
20. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin D, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019 Apr 15;144(8):1941-1953 [FREE Full text] [doi: 10.1002/ijc.31937] [Medline: 30350310]
21. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 01;136(5):E359-E386 [FREE Full text] [doi: 10.1002/ijc.29210] [Medline: 25220842]
22. Ferlay J, Shin H, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010 Dec 15;127(12):2893-2917 [FREE Full text] [doi: 10.1002/ijc.25516] [Medline: 21351269]

23. Segi M, Kurihara M. Cancer mortality for selected sites in 24 countries (1950-1957). 2edn. 1960. URL: https://books.google.ca/books/about/Cancer_Mortality_for_Selected_Sites_in_2.html?id=tH9rAAAAMAAJ&redir_esc=y [accessed 2023-09-11]
24. Doll R, Payne P, Waterhouse J. Cancer Incidence in 5 Continents: A Technical Report. Heidelberg: Springer Berlin; 1966.
25. Day N. Cancer incidence in five continents. cumulative rate and cumulative risk. IARC Sci Publ 1992(120):862-864. [Medline: 1284607]
26. Human development report 2020. United Nations Development Programme. 2020. URL: https://hdr.undp.org/content/human-development-report-2020 [accessed 2023-09-14]
27. The R project for statistical computing. R Foundation. URL: https://www.R-project.org/ [accessed 2022-12-30]
28. SigmaPlot for graphing and data visualization. Systat. URL: https://systatsoftware.com [accessed 2022-12-30]
29. Cao S, Simons MJ, Qian C. The prevalence and prevention of nasopharyngeal carcinoma in China. Chin J Cancer 2011 Feb;30(2):114-119 [FREE Full text] [doi: 10.5732/cjc.010.10377] [Medline: 21272443]
30. Hu C, Wang W, Chu X, Ren Z, Lyu J. Global, regional, and national burden of nasopharyngeal carcinoma from 1990 to 2017-results from the Global Burden of Disease Study 2017. Head Neck 2020 Nov;42(11):3243-3252. [doi: 10.1002/hed.26378] [Medline: 33463843]
31. Parascandola M, Xiao L. Tobacco and the lung cancer epidemic in China. Transl Lung Cancer Res 2019 May;8(Suppl 1):S21-S30 [FREE Full text] [doi: 10.21037/tlcr.2019.03.12] [Medline: 31211103]
32. Jia W, Qin H. Non-viral environmental risk factors for nasopharyngeal carcinoma: a systematic review. Semin Cancer Biol 2012 Apr;22(2):117-126. [doi: 10.1016/j.semcancer.2012.01.009] [Medline: 22311401]
33. Xie S, Yu IT, Tse L, Mang OW, Yue L. Sex difference in the incidence of nasopharyngeal carcinoma in Hong Kong 1983-2008: suggestion of a potential protective role of oestrogen. Eur J Cancer 2013 Jan;49(1):150-155. [doi: 10.1016/j.ejca.2012.07.004] [Medline: 22892061]
34. OuYang P, Zhang L, Lan X, Xie C, Zhang W, Wang Q, et al. The significant survival advantage of female sex in nasopharyngeal carcinoma: a propensity-matched analysis. Br J Cancer 2015 Apr 28;112(9):1554-1561 [FREE Full text] [doi: 10.1038/bjc.2015.70] [Medline: 25742485]
35. Chang E, Ye W, Zeng Y, Adami H. The evolving epidemiology of nasopharyngeal carcinoma. Cancer Epidemiol Biomarkers Prev 2021 Jun;30(6):1035-1047 [FREE Full text] [doi: 10.1158/1055-9965] [Medline: 33849968]
36. Mousavi SM, Sundquist J, Hemminki K. Nasopharyngeal and hypopharyngeal carcinoma risk among immigrants in Sweden. Int J Cancer 2010 Dec 15;127(12):2888-2892 [FREE Full text] [doi: 10.1002/ijc.25287] [Medline: 21351268]
37. Global Cancer Observatory. URL: https://gco.iarc.fr/ [accessed 2023-09-11]

Abbreviations

ASIR: age-standardized incidence rate
ASMR: age-standardized mortality rate
EBV: Epstein-Barr virus
HDI: Human Development Index
NPC: nasopharyngeal cancer

Edited by A Mavragani, T Sanchez; submitted 14.06.23; peer-reviewed by L Fu, X Wu; comments to author 30.06.23; revised version received 02.07.23; accepted 24.08.23; published 20.09.23.

Please cite as:
Zhang Y, Rumgay H, Li M, Cao S, Chen W
Nasopharyngeal Cancer Incidence and Mortality in 185 Countries in 2020 and the Projected Burden in 2040: Population-Based Global Epidemiological Profiling
JMIR Public Health Surveill 2023;9:e49968
URL: https://publichealth.jmir.org/2023/1/e49968
doi:10.2196/49968
PMID:37728964

Original Paper

Optimal Look-Back Period to Identify True Incident Cases of Diabetes in Medical Insurance Data in the Chinese Population: Retrospective Analysis Study

Wenyi Yang[1,2*], BM; Baohua Wang[3*], MM; Shaobo Ma[4], MM; Jingxin Wang[1,2], BM; Limei Ai[1,2], BM; Zhengyu Li[5], BM; Xia Wan[1,2], MD

[1]Institute of Basic Medical Sciences, Chinese Academy of Medical Science, Beijing, China

[2]School of Basic Medicine, Peking Union Medical College, Beijing, China

[3]Chinese Center for Disease Control and Prevention, National Institute for Prevention and Control of Chronic Noncommunicable Diseases, Beijing, China

[4]Weifang Medical Insurance Center, Weifang, China

[5]Department of Clinical Medicine, Qingdao University Medical College, Qingdao, China

[*]these authors contributed equally

Corresponding Author:
Xia Wan, MD
Institute of Basic Medical Sciences
Chinese Academy of Medical Science
Dongdan Street, 5th
Beijing, 100052
China
Phone: 86 01065233870
Email: xiawan@ibms.pumc.edu.cn

Abstract

Background: Accurate estimation of incidence and prevalence is vital for preventing and controlling diabetes. Administrative data (including insurance data) could be a good source to estimate the incidence of diabetes. However, how to determine the look-back period (LP) to remove cases with preceding records remains a problem for administrative data. A short LP will cause overestimation of incidence, whereas a long LP will limit the usefulness of a database. Therefore, it is necessary to determine the optimal LP length for identifying incident cases in administrative data.

Objective: This study aims to offer different methods to identify the optimal LP for diabetes by using medical insurance data from the Chinese population with reference to other diseases in the administrative data.

Methods: Data from the insurance database of the city of Weifang, China from between January 2016 and December 2020 were used. To identify the incident cases in 2020, we removed prevalent patients with preceding records of diabetes between 2016 and 2019 (ie, a 4-year LP). Using this 4-year LP as a reference, consistency examination indexes (CEIs), including positive predictive values, the κ coefficient, and overestimation rate, were calculated to determine the level of agreement between different LPs and an LP of 4 years (the longest LP). Moreover, we constructed a retrograde survival function, in which survival (ie, incident cases) means not having a preceding record at the given time and the survival time is the difference between the date of the last record in 2020 and the most recent previous record in the LP. Based on the survival outcome and survival time, we established the survival function and survival hazard function. When the survival probability, S(t), remains stable, and survival hazard converges to zero, we obtain the optimal LP. Combined with the results of these two methods, we determined the optimal LP for Chinese diabetes patients.

Results: The κ agreement was excellent (0.950), with a high positive predictive value (92.2%) and a low overestimation rate (8.4%) after a 2-year LP. As for the retrograde survival function, S(t) dropped rapidly during the first 1-year LP (from 1.00 to 0.11). At a 417-day LP, the hazard function reached approximately zero (h_t=0.000459), S(t) remained at 0.10, and at 480 days, the frequency of S(t) did not increase. Combining the two methods, we found that the optimal LP is 2 years for Chinese diabetes patients.

Conclusions: The retrograde survival method and CEIs both showed effectiveness. A 2-year LP should be considered when identifying incident cases of diabetes using insurance data in the Chinese population.

(JMIR Public Health Surveill 2023;9:e46708) doi:10.2196/46708

KEYWORDS

diabetes; incident cases; administrative data; look-back period; retrograde survival function

Introduction

Diabetes is a severe, long-term disease that significantly impacts the lives of individuals, families, and societies worldwide [1]. In the past 3 decades, the prevalence of type 2 diabetes has risen dramatically. About 422 million people worldwide have diabetes, and 1.5 million deaths are directly attributable to diabetes each year, as reported by World Health Organization in 2022. In 2013, China had the largest number of patients with diabetes and the second-highest spending on diabetes and its complications worldwide [2,3]. Therefore, to achieve global agreement and halt the worldwide rise in diabetes by 2025, it is crucial to prevent and control diabetes in China.

An accurate estimate of incidence and prevalence plays a crucial role in properly preventing and controlling diabetes. Although large-scale and representative surveys could be used for precise estimation of incidence [4-6], they are expensive, laborious, and time-consuming. Currently, only a few areas have established incident case surveillance systems for chronic diseases in China, and only for some chronic diseases (such as diabetes and hypertension), with short monitoring periods and few sample surveillance points, which limits their usefulness. An alternative method is to use administrative data, such as medical insurance data or disease registration data. Advantages of using these administrative data include easy access, low cost, and large sample size (covering the majority of residents), as well as the ability to follow up with participants longitudinally for many years; this has been confirmed by many researchers in other countries [7-11]. Since medical insurance reforms in 2012 [12], China has vigorously promoted medical policies and expanded coverage of the population [13]. Effective data have gradually accumulated during the past few years, making it possible to conduct research in this population.

However, multiple records of the same patient make it a general challenge to identify true incident cases through medical insurance data, as it is difficult to define the starting point of certain diseases for a patient. Before the index year, a look-back (LP) period is used to deal with this problem. Patients with a diabetes diagnosis in the index year and without a diabetes diagnosis in the LP would be identified as incident cases. Until now, there has been no consensus on the length of the LP. An insufficient LP might cause overestimation of incidence [14], whereas an overlong LP would limit the number of reporting years and the usefulness of the database. Researchers have found that different diseases have different LPs due to specific trajectories [11]. For diabetes, Brameld et al [15] identified an LP of 13 years using data from the Western Australian Data Linkage System, while Asghari et al identified an LP of 5 years using data from the Régie de l’assurance maladie du Québec [7]. We believe that different patterns of medical contact cycles among countries may cause these different findings on LPs. However, there is no study on the optimal LP in the Chinese population for diabetes.

In this study, we will use 2 methods to identify the optimal LP for diabetes in the Chinese population, using data from an insurance database for 2016 to 2020 obtained from the Medical Security Bureau (MSB) of the city of Weifang, China, as an example. Weifang, a prefecture-level city in Shandong province with 9.4 million long-term residents, has a systematic method for ensuring medical service, and the coverage rate for medical insurance reached 96.15% in 2021. The aim of this study is to identify the optimal LP for diabetes for the Chinese population and additionally to compare different results by using different methods to obtain a reference for other diseases in the administrative data.

Methods

Ethical Considerations

The proposal for this study was reviewed by Chinese Academy of Medical Science & School of Basic Medicine (033-2018). The insurance claims data were obtained from the MSB of Weifang. In order to protect the privacy of patients, all personally identifiable information in the database is obfuscated.

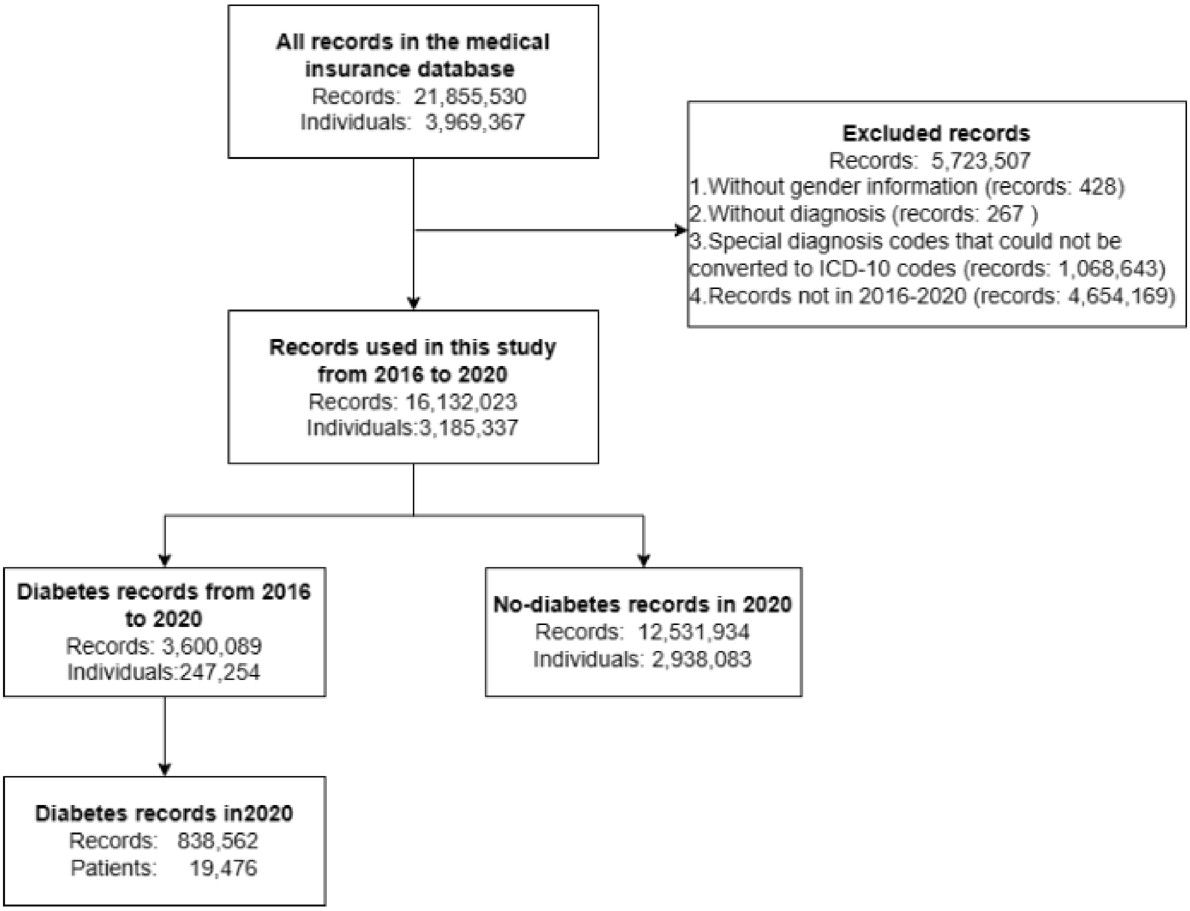
Data Sources

The database used in this study included all in- or outpatient claim records of patients in Weifang for 2016 to 2020 provided by the local MSB. The social health insurance system covered approximately all residents in Weifang (close to 9 million individuals) during the period from 2016 to 2020. Each claim record contains (1) basic information about the patient, including personal identification number, gender, date of birth, and type of medical insurance; (2) medical information, including hospital name and code, admission and discharge time, and diagnosis of disease; and (3) details of medical service expenses, including total expenses, out-of-pocket expenses, and reimbursement expenses.

Data Processing

Firstly, due to the changes in the Weifang disease coding system and the special disease codes in outpatient chronic-disease records, the initial disease codes in the database were partly based on the International Classification of Diseases, 10th Revision (ICD-10). Therefore, the diagnoses were harmonized with the ICD-10 according to the code-disease mapping table provided by the MSB of Weifang. Later, records with critical missing information, such as gender (418/21,855,530, <0.001%) or disease diagnosis (1,068,910/21,855,530, 4.89%), were deleted. The data processing is shown in Figure 1.

Figure 1. Flow chart showing included and excluded records and the process used for data cleaning of the insurance claim data (January 1, 2016, to December 31, 2020). ICD-10: International Classification of Disease, 10th Revision.

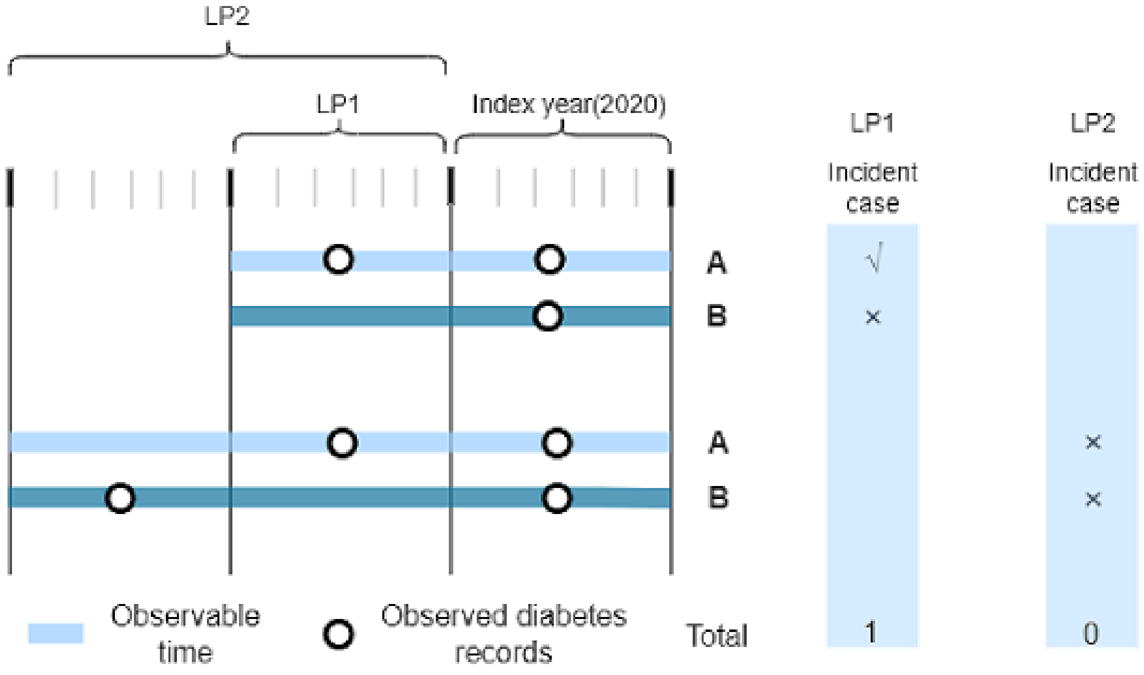


Definition of Incident Diabetes for the Application of Different LPs

First, we defined diabetes (ICD-10 codes E10-E14) cases as those with at least 1 inpatient claim record or 3 outpatient claim records with a diagnosis of diabetes, considering that inpatient records are more reliable than outpatient records [16]. In this database, we found that all patients defined as having diabetes had more than 3 outpatient records, reflecting the regularity of medical contact.

Second, we defined incident diabetes cases. The procedure to identify incident cases was to exclude cases with any in- or outpatient records of diabetes in the given LP, before the index year. Different LPs may classify patients as having different statuses. Two scenarios for assumed different LPs are shown in Figure 2. Scenario 1 (LP1) shows that patient B will be identified as an incident case but patient A will not. Scenario 2 (LP2) shows that with the extension of the LP, more records of patient B might be caught, but patient A and patient B will both be identified as prevalent cases. In order to use as long an LP as possible in this study, we used 2020 as the index year and a 4-year LP (2016-2019) as the longest LP.

Figure 2. Definition of incident cases under 2 scenarios. Scenario 1 (LP1): patient B will be identified as an incident case but patient A will not; scenario 2 (LP2): both patient A and patient B are not identified as incident diabetes cases. LP: look-back period.



Statistical Analyses

Two methods were developed to identify the optimal LP. First, consistency examination indexes (CEIs) were used to determine the level of agreement between different LPs from a quarter of a year to 4 years. Second, a retrograde survival function method was used. Finally, the optimal LP was selected after considering the results from the above approaches.

The CEIs include the κ value, positive predictive values (PPVs), and overestimation value. The relative formulas have been reported in other papers [4,15]. The κ value is categorized as follows: values between –1 and 0 represent “no agreement”; between 0 and 0.20 represent “poor agreement”; between 0.21 and 0.40 represent “slight agreement”; between 0.41 and 0.60 represent “fair agreement”; between 0.61 and 0.80 represent “good agreement”; between 0.81 and 0.90 represent “very good agreement”; and between 0.91 and 1.0 represent “excellent agreement” [17]. The PPVs are used to estimate the probability of identifying a true incident case among those identified as being new cases [18], calculated using the 4-year LP as the reference time. The overestimation of incident cases decreased with extension of the LP, which reflects the impact of the length of the LP impacts on the number of incident cases. All the CEIs mentioned above were calculated at quarter-year intervals during the first year of the LP and 1-year intervals after the first year, since the number of identified incident cases varied greatly within the first year of the LP in previous studies [7,12,16].

With the retrograde survival function method, survival (ie, being an incident case) means not having a preceding record at the given time. In detail, for each record in this study, let *a* represent the date of the last record of diabetes of a patient in 2020, *b* the date of the most recent previous record in the LP, and *c* the first date of the LP. The retrograde survival time is (*a*–*b*) for patients with preceding records. Patients without preceding records during the LP, indicating survival, are defined as censored cases, with the censored time (*a* – *c*). Based on the survival time and censored time, we established the survival function. Survival probability, S(t), the probability of surviving at a specific time, is represented by the equation S(t)=P (T>t), where *T* is the year of the LP, calculated using the Kaplan-Meier estimator. Then, we used the following 2 approaches to determine the optimal LP based on the survival function. For the first approach, we constructed a hazard function. The hazard function represents the instantaneous probability of having the end point event during (t,Δt) conditional on survival up to *t* or later. In actuarial terms, it is calculated by , where $d_{\Delta t}$ is the number of patients with the previous record during (t,Δt) and n_t is the patient at risk during the interval. Since the retrograde hazard of diabetes will decrease and converge to zero and the survival function will converge to the “true” probability of being incident, we assumed that a time point (*tf*) exists where . At this point (*tf*), all cases with preceding records have been found, and the remaining cases are risk-free; this is exactly the optimal LP. To estimate (*tf*), for practical purposes, we defined (*tf*) as the first time point when $h_t \le 0.0005$. For the second approach, we counted the frequency of S(t) corresponding to each day calculated by the retrograde survival function. We assumed that there was a stable S(t), where the frequency was high and increased significantly, representing that S(t) stays at S(t_f) relatively constantly over a long time and that the change in S(t) approaches zero. The t_f will be considered as the point when S(t) first reaches S(t_f).

Statistical analyses were performed using SAS (version 9.4; SAS institute).

Results

After removing records according to the criteria above, 16,132,023 records from medical insurance claims data from 2016 to 2020 were used in this study; these records came from 3,185,337 individuals. There were 3,600,089 records and 247,254 cases of diabetes in total. The insurance record history of 109,476 diabetes prevalent cases in 2020 was screened with the different LPs to identify incident cases. Inclusion and exclusion details are shown in Figure 1, and basic information on the data is shown in Table 1.

Table 1. Basic information on the data used in this study (January 1, 2016, to December 31, 2020). "Insurance records" represents all claims and diabetes records in each year. "Individuals with insurance records" represents the number of persons with claims in each year. "Visits per capita" was calculated as the number of insurance records divided by the number of individuals with insurance records.

	Insurance records, n		Individuals with insurance records, n		Visits per capita, n	
	Inpatient	Outpatient	Inpatient	Outpatient	Inpatient	Outpatient
Entire database						
2016	1,227,374	1,326,862	894,726	128,819	1.37	10.30
2017	1,459,273	1,583,937	1,027,600	151,130	1.42	10.48
2018	1,484,065	1,815,786	1,024,998	168,547	1.45	10.77
2019	1,559,320	2,123,446	1,061,351	196,294	1.47	10.82
2020	1,393,165	2,158,795	947,821	220,560	1.47	9.79
Diabetes cases						
2016	26,597	526,619	23,117	58,703	1.15	8.97
2017	31,053	618,289	26,833	67,542	1.16	9.15
2018	31,668	699,739	27,425	75,361	1.15	9.29
2019	34,131	793,431	29,594	87,197	1.15	9.10
2020	30,604	807,958	27,038	97,396	1.13	8.30

Applying the longest LP identified 19,086 incident cases of diabetes. Only 17% (19,087/109,476) of the prevalent patients in 2020 were identified as incident cases as they did not have preceding records for diabetes in the longest LP before 2020, revealing that around 80% (90,389/109,476) of patients in 2020 were not truly incident cases. The κ coefficient and PPVs for these results rose with extension of the LP. Agreement was good when a quarter-year LP (κ=0.751) was used, changing to very good (κ=0.899) and excellent (κ=0.950) when the 1-year and 2-year LPs were used. As for PPVs, to achieve a PPV of at least 90% (10% of false positives) [19], an LP of 2 years was necessary. The overestimation rate decreased with the extension of the length of the LP and did so especially sharply during the 1-year LP. After a 2-year LP, the overestimation decreased by under 10%. In general, according to the results of these CEIs, the optimal LP is 2 years for the Chinese population. Detailed information on the CEIs is shown in Table 2.

Table 2. Positive predictive values, κ coefficients, and overestimation with different look-back periods, all of which represent the agreement between the incident cases with different look-back periods and the 4-year look-back period as the longest period. Incident cases are defined as those not having previous records in the corresponding look-back period.

LP[a] (years)	Incident cases, n	PPV[b], %	κ	Overestimation, %
0.25	29,407	0.674	0.751	0.485
0.5	26,195	0.756	0.825	0.323
0.75	24,172	0.819	0.876	0.220
1	23,288	0.850	0.899	0.176
2	21,472	0.922	0.950	0.084
3	20,468	0.968	0.980	0.033
4	19,806	1	N/A[c]	0

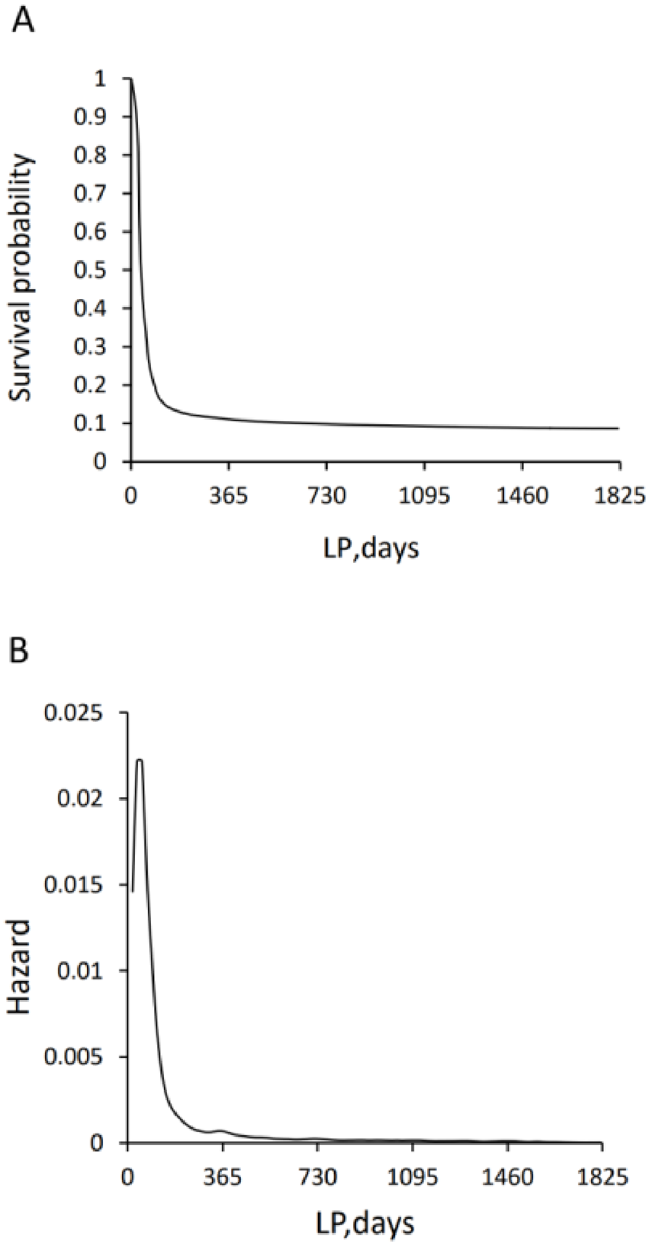
[a] LP: look-back period.

[b] PPV: positive predictive value.

[c] N/A: not applicable.

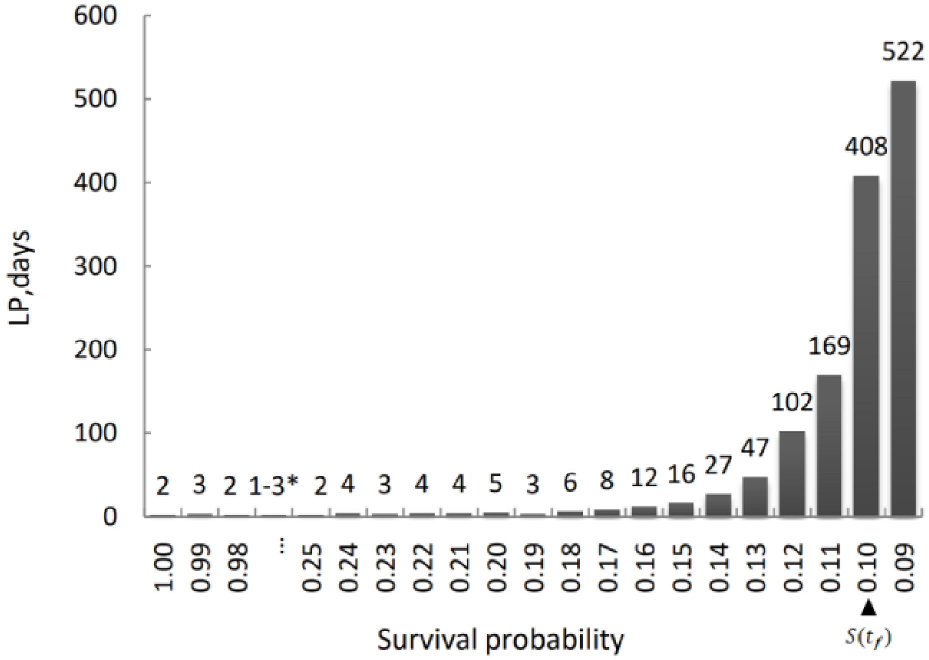
The retrograde survival function graphically represented the exclusion, showing the greatest decrease during the first year of screening (Figure 3). During the first 1-year LP, S(t) dropped from 1.00 to 0.11 (365 days) rapidly. The more days were included in the LP, the more the survival function converged to the “true” probability of being incident [18]. According to the hazard function, the first time point when $h_t \leq 0.0005$ was with a 417-day LP: h_t=0.000459; S(t)=0.11. This meant that the derivative (slope) of the survival function tended toward zero at this time.

Figure 3. (A) Change in survival probability with LPs with different numbers of days. Survival probability represents the probability of being an incident case at a specific day in the LP, which is calculated using the Kaplan-Meier estimator. (B) Change in survival hazard with LPs with different numbers of days. Survival hazard represents the instantaneous probability of having the end point event. LP: look-back period.



The results for frequency showed that the frequency of S(t) was small, from 1.00 to 0.13, indicating that S(t) changed greatly. A change in S(t) from 0.11 to 0.10 means that the frequency has significantly increased, from 169 to 408, which is more than double (Figure 4). This means that the risk for a preceding record at this time remained relatively constant at 0.10 over a very long time, and the first day when S(t) reached 0.10 was 480 days. Combining the results for frequency and hazard function, the retrograde survival function method showed that the optimal LP was more than 400 days, which is closest to 2 years.

Figure 4. The frequency of S(t). *The frequency of S(t) from 0.97 to 0.26 was between 1 and 3. LP: look-back period.



When using 2 years as the optimal LP, there were 21,472 incident cases. The number of female patients (n=10,862) was a little more than that of male patients (n=10,610), and 84.74% (18,192/21,472) of incident cases were between the ages of 45 and 79 years (Table 3).

Table 3. Incident cases applying a 2-year look-back period (ie, using 2018 and 2019 to identify incident cases in 2020) with different ages and genders.

Age group (years)	Male patients, n (n=10,610)	Female patients, n (n=10,862)	Total patients, n (n=21,472)
0-4	4	7	11
5-9	7	16	23
10-14	36	29	65
15-19	34	23	57
20-24	52	45	97
25-29	99	63	162
30-34	308	194	502
35-39	401	178	579
40-44	577	271	848
45-49	961	538	1499
50-54	1432	1073	2505
55-59	1806	1720	3526
60-64	1544	1778	3322
65-69	1504	2061	3565
70-74	1000	1533	2533
75-79	481	764	1245
80-84	269	394	663
≥85	95	175	270

Discussion

Principal Results

In our study, a 2-year LP had good CEIs with excellent κ agreement (0.950>0.9), a high PPV (92.2%) and a lower than 10% overestimation rate. As for the retrograde survival function, it showed that a 2-year LP can reliably distinguish new cases in a prevalent pool. According to the results of the methods above, we confirmed that the optimal LP for identifying incident cases of diabetes using the Chinese insurance database provided by the MSB of Weifang is around 2 years, and all the methods are effective and stable.

Comparison With Prior Work

Optimal Length of LP Compared to Other Work

A 2-year LP for diabetes incident case identification using administrative data is the same length as used by a few previous studies [20,21]. A study in Manitoba, Canada, also showed a probability of 0.96 for patients with diabetes to have subsequent medical contact for diabetes within 2 years [22]. However, our results differ from those of Asghari et al [7], who reported an LP of 5 years for health insurance data from Quebec, Canada. A possible reason might be different patterns and medical contact cycles among Chinese and Canadian populations caused by the different medical insurance policies. The hazard function figure (Figure 3B) reflects this difference. In our study, the hazard function was not monotonic, indicating that a few diabetes patients had medical contact irregularly in this database, unlike the Canadian insurance data. n addition, different judgment criteria on CEIs and h_t in these studies might have caused the differences.

Methodological Comparison

As for retrograde survival function, Brameld et al [15] chose a hazard rate under 0.00001 as the standard when using the retrograde survival method for identifying the optimal LP for diabetes, resulting in a 13-year optimal LP for diabetes among a group of LPs with the longest LP being 15 years, much longer than our study and most other studies [20,23,24]. We think an overly strict standard may have caused an overly long LP. To avoid an overly long estimation of optimal LP, we used $h_t \leq 0.0005$ as the criterion, which is similar to the study by Asghari et al [7], which showed that a 2-year LP was optimal for diabetes. That study defined the optimal LP as the time when S(t) was stable and constant [7]; however, the authors did not provide a detailed definition for stable S(t), that is, $S(t_f)$. In our study, we tried a new approach. We counted the frequency of S(t) and defined $S(t_f)$ by the frequency of S(t) being greatest and increasing significantly.

Regarding CEIs, such as the κ coefficient, PPVs, and overestimation rate, although they are convenient to calculate and used by many researchers, their criteria are not completely consistent between studies. The overestimation rate has been used by a few researchers to evaluate the impact of varying lengths of LP [11,21,25-27], with values such as 10% [11] and 20% [21]. It is the same for PPVs, as some studies have chosen 80% as the evaluation criterion [19], while some have chosen 90% [28]. The criteria of the κ coefficient have generally between consistent (0.8 or 0.9) [7,18], but some researchers have considered that such high agreement may be related to the high number of prevalent cases of diabetes [7]. Therefore, the effectiveness of κ may remain to be further verified.

Because of the inconsistency of criteria, Beaudet et al [18] chose to combine the results of κ and PPV to identify the optimal LP. In our study, we compared the results of the κ coefficient, PPV, the overestimation rate, and the retrograde survival method. We found that a κ coefficient of 0.9, 90% PPV, 10% overestimation, and 0.0005 hazard rate may be the most suitable criteria, returning stable and constant results, with a 2-year LP as the optimal period.

In general, the CEIs were convenient to calculate, while retrograde survival curves showed the probability of "surviving" converging to the plateau, allowing us to obtain the optimal LP visually and quantitively. Moreover, the criteria of these methods are different. Previous studies have mostly used only one method for identifying the optimal LP, whereas 2 different methods, as used in this study, may lead to more robust results, which could then be referred to when analyzing LPs for other diseases in the administrative data.

Number of Incident Diabetes Cases

In our study, there were 109,467 diabetes cases in 2020 before LP adjustment. After using the 2-year LP, the number of incident cases decreased to 21,472, which means this adjustment is important. In addition, our study showed that 84.74% (18,192/21,472) of incident cases were between the ages of 45 to 79 years, indicating that the middle-aged and older population has a high incidence of diabetes. We should make efforts to focus on this population to prevent diabetes.

Innovativeness

Our study is the first to identify 2 years as the optimal LP for diabetes in the Chinese population when using insurance data, which supports efforts to estimate the incidence of diabetes among populations using insurance data. More importantly, we used 2 different methods to identify the optimal LP, a method that could be referred to when analyzing LPs for other diseases in administrative data.

Limitations

Our study has some limitations. In our database, without more detailed records for disease classification, it was not possible to differentiate between type 1 and type 2 diabetes. However, there are far fewer patients with type 1 diabetes than with type 2 diabetes in the overall population. Thus, we considered the study population to be generally homogeneous and that these results may be representative and stable. Furthermore, we only applied up to a 4-year LP. Although the results of the retrograde survival function analysis showed stabilization of risk of being a preceding case after a 2-year LP, we expect to use longer LPs in future studies with more accumulated data.

Conclusion

All CEIs (the κ coefficient, PPV, and overestimation rate) and the retrograde survival function method can effectively identify the optimal LP for diabetes incident cases, with stable and constant results. We found that a 2-year LP was optimal to remove preceding diabetes cases and identify the true incidence of diabetes using medical insurance data in the Chinese population.

Acknowledgments

The authors thank all participants and staff of the Weifang Medical Insurance Center for their contributions to this work. The authors also acknowledge the contribution to this study of the Center for Disease Control and Prevention of Weifang City. The authors appreciate the support of the China Medical Board (CMB), which supported the CMB Collaborating Program in Burden of Diseases in China (15-208).

Data Availability

The insurance database used in this study is not publicly available due to a confidentiality agreement with Medical Security Bureau of Weifang but it is available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9 edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
2. Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. JAMA 2017 Jun 27;317(24):2515-2523 [FREE Full text] [doi: 10.1001/jama.2017.7596] [Medline: 28655017]
3. Wang L, Peng W, Zhao Z, Zhang M, Shi Z, Song Z, et al. Prevalence and treatment of diabetes in China, 2013-2018. JAMA 2021 Dec 28;326(24):2498-2506 [FREE Full text] [doi: 10.1001/jama.2021.22208] [Medline: 34962526]
4. Shibayama T, Noguchi H, Takahashi H, Tamiya N. Relationship between social engagement and diabetes incidence in a middle-aged population: Results from a longitudinal nationwide survey in Japan. J Diabetes Investig 2018 Sep;9(5):1060-1066 [FREE Full text] [doi: 10.1111/jdi.12820] [Medline: 29430865]
5. Hawkins Carranza F, Corbatón-Anchuelo A, Bermejo Pareja F, Martín-Arriscado Arroba C, Vega-Quiroga S, Benito-León J, et al. Incidence of type 2 diabetes in the elderly in Central Spain: Association with socioeconomic status, educational level, and other risk factors. Prim Care Diabetes 2022 Apr;16(2):279-286 [FREE Full text] [doi: 10.1016/j.pcd.2021.12.016] [Medline: 35065897]
6. Boonpor J, Ho FK, Gray SR, Celis-Morales CA. Association of self-reported walking pace with type 2 diabetes incidence in the UK Biobank Prospective Cohort Study. Mayo Clin Proc 2022 Sep;97(9):1631-1640. [doi: 10.1016/j.mayocp.2022.02.028] [Medline: 36058577]
7. Asghari S, Courteau J, Carpentier AC, Vanasse A. Optimal strategy to identify incidence of diagnostic of diabetes using administrative data. BMC Med Res Methodol 2009 Aug 28;9(1):62 [FREE Full text] [doi: 10.1186/1471-2288-9-62] [Medline: 19715586]
8. Cohen A, Mok E, Simard M, Dubé M, Larocque I, Plante C, et al. Increasing incidence of type 1 and type 2 diabetes among Canadian children. Can J Diabetes 2022 Mar;46(2):189-195. [doi: 10.1016/j.jcjd.2021.08.006] [Medline: 35101343]
9. Ortiz-Marrón H, Del Pino Valero V, Esteban-Vasallo M, Zorrilla Torras B, Ordobás Gavín M. Evolution of the incidence of type 1 diabetes mellitus in the Community of Madrid, 1997-2016. An Pediatr (Engl Ed) 2021 Oct;95(4):253-259 [FREE Full text] [doi: 10.1016/j.anpede.2020.08.011] [Medline: 34481795]
10. Purkiss SF, Keegel T, Vally H, Wollersheim D. Estimates of drug treated diabetes incidence and prevalence using Australian administrative pharmaceutical data. Int J Popul Data Sci 2021 May 10;6(1):1398 [FREE Full text] [doi: 10.23889/ijpds.v6i1.1398] [Medline: 34007898]
11. Abbas S, Ihle P, Köster I, Schubert I. Estimation of disease incidence in claims data dependent on the length of follow-up: a methodological approach. Health Serv Res 2012 Apr;47(2):746-755 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01325.x] [Medline: 21958217]
12. Wu R, Li N, Ercia A. The effects of private health insurance on universal health coverage objectives in China: A systematic literature review. Int J Environ Res Public Health 2020 Mar 19;17(6):2049 [FREE Full text] [doi: 10.3390/ijerph17062049] [Medline: 32204527]
13. Meng Q, Xu L, Zhang Y, Qian J, Cai M, Xin Y, et al. Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study. Lancet 2012 Mar 03;379(9818):805-814. [doi: 10.1016/S0140-6736(12)60278-5] [Medline: 22386034]

14. Czwikla J, Jobski K, Schink T. The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data. BMC Med Res Methodol 2017 Aug 14;17(1):122 [FREE Full text] [doi: 10.1186/s12874-017-0407-4] [Medline: 28806932]
15. Brameld K, Holman C, Lawrence D, Hobbs M. Improved methods for estimating incidence from linked hospital morbidity data. Int J Epidemiol 2003 Aug;32(4):617-624. [doi: 10.1093/ije/dyg191] [Medline: 12913039]
16. Lin J, Shau W, Lai M. Sex- and age-specific prevalence and incidence rates of sight-threatening diabetic retinopathy in Taiwan. JAMA Ophthalmol 2014 Aug;132(8):922-928. [doi: 10.1001/jamaophthalmol.2014.859] [Medline: 24809869]
17. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993 May;46(5):423-429. [doi: 10.1016/0895-4356(93)90018-v] [Medline: 8501467]
18. Beaudet N, Courteau J, Sarret P, Vanasse A. Improving the selection of true incident cases of low back pain by screening retrospective administrative data. Eur J Pain 2014 Aug;18(7):923-931. [doi: 10.1002/j.1532-2149.2013.00437.x] [Medline: 24343901]
19. Vanasse A, Courteau J, Fleury M, Grégoire JP, Lesage A, Moisan J. Treatment prevalence and incidence of schizophrenia in Quebec using a population health services perspective: different algorithms, different estimates. Soc Psychiatry Psychiatr Epidemiol 2012 Apr;47(4):533-543. [doi: 10.1007/s00127-011-0371-y] [Medline: 21445625]
20. Koo BK, Lee C, Yang BR, Hwang S, Choi N. The incidence and prevalence of diabetes mellitus and related atherosclerotic complications in Korea: a National Health Insurance Database Study. PLoS One 2014;9(10):e110650 [FREE Full text] [doi: 10.1371/journal.pone.0110650] [Medline: 25329714]
21. Rosenlund M, Ekström N, Törnblom M, Wintzell V, Stark JH, Titievsky L. Impact of variable look-back periods on the incidence rates of chronic diseases using real world data. Pharmacoepidemiol Drug Saf 2020 Sep;29(9):1086-1092. [doi: 10.1002/pds.5066] [Medline: 32648308]
22. Blanchard JF, Ludwig S, Wajda A, Dean H, Anderson K, Kendall O, et al. Incidence and prevalence of diabetes in Manitoba, 1986-1991. Diabetes Care 1996 Aug;19(8):807-811. [doi: 10.2337/diacare.19.8.807] [Medline: 8842595]
23. Andes LJ, Li Y, Srinivasan M, Benoit SR, Gregg E, Rolka DB. Diabetes prevalence and incidence among Medicare beneficiaries - United States, 2001-2015. MMWR Morb Mortal Wkly Rep 2019 Nov 01;68(43):961-966 [FREE Full text] [doi: 10.15585/mmwr.mm6843a2] [Medline: 31671084]
24. Bratu A, McLinden T, Kooij K, Ye M, Li J, Trigg J, et al. Incidence of diabetes mellitus among people living with and without HIV in British Columbia, Canada between 2001 and 2013: a longitudinal population-based cohort study. BMJ Open 2021 May 11;11(5):e048744 [FREE Full text] [doi: 10.1136/bmjopen-2021-048744] [Medline: 33980535]
25. Stein JD, Blachley TS, Musch DC. Identification of persons with incident ocular diseases using health care claims databases. Am J Ophthalmol 2013 Dec;156(6):1169-1175.e3 [FREE Full text] [doi: 10.1016/j.ajo.2013.06.035] [Medline: 23972306]
26. Schmedt N, Khil L, Berger K, Riedel O. Incidence of multiple sclerosis in Germany: A cohort study applying different case definitions based on claims data. Neuroepidemiology 2017;49(3-4):91-98. [doi: 10.1159/000481990] [Medline: 29136628]
27. Camplain R, Kucharska-Newton A, Cuthbertson CC, Wright JD, Alonso A, Heiss G. Misclassification of incident hospitalized and outpatient heart failure in administrative claims data: the Atherosclerosis Risk in Communities (ARIC) study. Pharmacoepidemiol Drug Saf 2017 Apr;26(4):421-428 [FREE Full text] [doi: 10.1002/pds.4162] [Medline: 28120359]
28. Benchimol EI, Guttmann A, Mack DR, Nguyen GC, Marshall JK, Gregor JC, et al. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada. J Clin Epidemiol 2014 Aug;67(8):887-896. [doi: 10.1016/j.jclinepi.2014.02.019] [Medline: 24774473]

Abbreviations

CEI: consistency examination index
ICD-10: International Classification of Diseases, 10th Revision
LP: look-back period
MSB: Medical Security Bureau
PPV: positive predictive value

Edited by A Mavragani, T Sanchez; submitted 22.02.23; peer-reviewed by A Khatami, C Yang; comments to author 04.08.23; revised version received 23.08.23; accepted 26.09.23; published 06.11.23.

Please cite as:
Yang W, Wang B, Ma S, Wang J, Ai L, Li Z, Wan X
Optimal Look-Back Period to Identify True Incident Cases of Diabetes in Medical Insurance Data in the Chinese Population: Retrospective Analysis Study
JMIR Public Health Surveill 2023;9:e46708
URL: https://publichealth.jmir.org/2023/1/e46708
doi:10.2196/46708
PMID:37930785

Original Paper

Application of Machine Learning Prediction of Individual SARS-CoV-2 Vaccination and Infection Status to the French Serosurveillance Survey From March 2020 to 2022: Cross-Sectional Study

Stéphanie Bougeard[1], PhD; Adeline Huneau-Salaun[1]; Mikael Attia[2], PhD; Jean-Baptiste Richard[3], MSc; Caroline Demeret[2], PhD; Johnny Platon[3]; Virginie Allain[1]; Stéphane Le Vu[4], PhD; Sophie Goyard[5], PhD; Véronique Gillon[6], PhD; Sibylle Bernard-Stoecklin[7], PhD; Bernadette Crescenzo-Chaigne[2], PhD; Gabrielle Jones[7]; Nicolas Rose[1], PhD; Sylvie van der Werf[2], PhD; Olivier Lantz[6*], PhD, MSc; Thierry Rose[5*], PhD; Harold Noël[7*], PhD

[1]Epidemiology, Health and Welfare, Laboratory of Ploufragan-Plouzané-Niort, French Agency for Food, Environmental, Occupational Health & Safety, Ploufragan, France

[2]Unit of Molecular Genetics of RNA Viruses, Institut Pasteur, Paris, France

[3]Data Support, Processing and Analysis Department, Santé publique France, Saint-Maurice, France

[4]EPI-PHARE, Saint-Denis, France

[5]Diagnostic Test Innovation and Development Core Facility, Institut Pasteur, Paris, France

[6]Clinical Immunology Laboratory, Institut Curie, Paris, France

[7]Infectious Disease Division, Santé publique France, Saint-Maurice, France

[*]these authors contributed equally

Corresponding Author:
Stéphanie Bougeard, PhD
Epidemiology, Health and Welfare
Laboratory of Ploufragan-Plouzané-Niort
French Agency for Food, Environmental, Occupational Health & Safety
BP 53 - Technopole Saint Brieuc Armor
Ploufragan, 22440
France
Phone: 33 296010150
Email: stephanie.bougeard@anses.fr

Abstract

Background: The seroprevalence of SARS-CoV-2 infection in the French population was estimated with a representative, repeated cross-sectional survey based on residual sera from routine blood testing. These data contained no information on infection or vaccination status, thus limiting the ability to detail changes observed in the immunity level of the population over time.

Objective: Our aim is to predict the infected or vaccinated status of individuals in the French serosurveillance survey based only on the results of serological assays. Reference data on longitudinal serological profiles of seronegative, infected, and vaccinated individuals from another French cohort were used to build the predictive model.

Methods: A model of individual vaccination or infection status with respect to SARS-CoV-2 obtained from a machine learning procedure was proposed based on 3 complementary serological assays. This model was applied to the French nationwide serosurveillance survey from March 2020 to March 2022 to estimate the proportions of the population that were negative, infected, vaccinated, or infected and vaccinated.

Results: From February 2021 to March 2022, the estimated percentage of infected and unvaccinated individuals in France increased from 7.5% to 16.8%. During this period, the estimated percentage increased from 3.6% to 45.2% for vaccinated and uninfected individuals and from 2.1% to 29.1% for vaccinated and infected individuals. The decrease in the seronegative population can be largely attributed to vaccination.

Conclusions: Combining results from the serosurveillance survey with more complete data from another longitudinal cohort completes the information retrieved from serosurveillance while keeping its protocol simple and easy to implement.

(JMIR Public Health Surveill 2023;9:e46898) doi:10.2196/46898

KEYWORDS

SARS-CoV-2; serological surveillance; infection; vaccination; machine learning; seroprevalence; blood testing; immunity; survey; vaccine response; French population; prediction

Introduction

In the ongoing global effort to contain the SARS-CoV-2 pandemic, population-wide serological surveys are recommended for disease surveillance and policymaking [1]. Serological assays directly measure the antibody response to SARS-CoV-2 resulting from viral infection or vaccination. Monitoring seroprevalence is of paramount interest to complement case-based surveillance that does not capture subclinical cases or people using self-tests and to evaluate the effectiveness of the vaccination strategy over time. Therefore, the World Health Organization's UNITY initiative promotes serological surveys by providing guidelines to standardize worldwide serological studies.

Despite the availability of well-known methods as well as guidelines, implementation of serological surveys remains challenging in terms of the resources and logistic means needed to obtain samples [2]. Collecting residual sera from routine clinical blood testing represents an easy and inexpensive solution. This strategy was chosen for the nationwide SARS-CoV-2 serosurvey in France (SERPICO) [3]. This survey, conducted by the national health agency Santé publique France and the National Reference Centre for Respiratory Infections Viruses of the Institut Pasteur, aimed to estimate the seroprevalence of anti–SARS-CoV-2 antibodies in the French population over time according to gender, age, and region. The humoral immune status of individuals was determined by 3 complementary serological assays: 2 luciferase-linked immunosorbent assays (LuLISAs) detecting the nucleocapsid (LN) and the spike (LS) proteins of SARS-CoV-2 and a pseudoneutralization assay (PNT). In total, 8 collection periods between March 2020 and March 2022 monitored the evolution of seroprevalence in the French general population.

Substantial differences were reported in the effectiveness and duration of natural versus vaccine-conferred or hybrid immunity against SARS-CoV-2 reinfection [4,5]. In particular, prior infection after 1 dose of vaccine elicited antispike IgG antibody responses with higher peak levels or longer half-lives than 1 or 2 vaccinations in seronegative individuals [6]. We therefore aimed to reconstruct the proportion of infected versus vaccinated or infected and vaccinated individuals over time as it may carry meaningful lessons and potential applications for future vaccine-preventable disease pandemics.

Although the SERPICO serosurvey lacked data documenting the status of the sampled individuals with respect to previous SARS-CoV-2 infection or vaccination, the detailed serological data presented here could allow novel approaches to monitor the SARS-CoV-2 seroprevalence of the French population with regard to infection or vaccination.

In this study, we propose a model derived from a machine learning procedure to predict individuals' immune status with respect to SARS-CoV-2 infection and vaccination based on results from 3 serological assays. We applied this predictive model to the SERPICO serosurvey to characterize immune status resulting from natural infection and vaccination for the French population between 2020 and 2022.

Methods

Serological Assays

The National Reference Centre for Respiratory Infections Viruses and the Diagnostic Test Innovation and Development core facility of the Institut Pasteur developed 3 serological assays: 2 LuLISAs detecting the LN and LS proteins of SARS-CoV-2 and a PNT [7]. These serological results are considered explanatory variables in the predictive models. They were expressed as $\log_{10}$ of their original value.

Data

This study used 3 data sets (Table 1). The CURIE-O-SA [7] and prepandemic [3,7] data sets were used to build and validate the model, and the SERPICO [3] data set was used for application of the model.

The CURIE-O-SA study is a large cohort of 1917 workers in a hospital and research center specialized in oncology. Of the 4394 individuals included in this data set, 77% (n=3595) were men and 23% (n=899) were women. The median age was 38 (IQR 19-82) years with 94% (n=4130) of participants between 20 and 59 years of age. The CURIE-O-SA serological assay results were available at different sampling times ("individual date") with a mean number of times per participant equal to 2.5 (SD 1.2) and 73% (n=3208) of individuals with 1, 2, or 3 sampling times. Information on SARS-CoV-2 history (symptoms and date of positive reverse transcription polymerase chain reaction [RT-PCR] test, if any) and on SARS-CoV-2 vaccination (number of injections and dates) was collected. The uninfected status (0) was set if the following 3 conditions were met: no positive PCR test, no declaration of ageusia or anosmia, and a $\log_{10}$LN value below 4.60. Individuals with $\log_{10}$LN>4.60 and no positive PCR test were considered potential subclinical cases and were excluded. The infected status (1) was set for individuals with a positive RT-PCR result history only; individuals reporting symptoms for SARS-CoV-2 with no positive RT-PCR result were excluded. The vaccination status (0=unvaccinated, 1=1 injection, 2=2 injections) was defined while taking into account a delay of 15 days after vaccination (immunity onset) and no more than 6 months after the last injection (immunity waning). Of the 4394 results, 80.4% (n=3532) of results were negative (uninfected and unvaccinated), 5.2% (n=231) were infected and unvaccinated, 10.9% (n=477) were vaccinated twice, 1.9% (n=82) were vaccinated once, and 1.6% (n=72) were infected and vaccinated.

The prepandemic results came from healthy donors from a blood bank before 2019 and were expected to be predicted negative (uninfected and unvaccinated) by the model. In this data set, of the 233 individuals, 32.2% (n=75) were men and 67.8% (n=158) were women with a median age of 44 (IQR 18-81) years.

The SERPICO data were considered application data for the model. The SERPICO survey monitored the evolution of seroprevalence of anti–SARS-CoV-2 antibodies in the French population (mainland area) from March 2020 to March 2022 with 8 periods of sampling. Results for the 3 serological assays (LN, LS, and PNT) were available for 23,886 samples.

Table 1. Description of the 3 data sets under study.

Data set	Description	Use	Number of results	Variables
CURIE-O-SA [7] (n=4394)	Cohort study among health workers; April 2020-November 2021	Reference for SARS-CoV-2 status; reference for SARS-CoV-2 vaccination status	4394 individuals × dates	LS[a], LN[b], PNT[c] values; gender; age; SARS-CoV-2 infection status (0=uninfected, 1= infected); SARS-CoV-2 vaccination status (0=unvaccinated, 1=1 injection, 2=2 injections)
Prepandemic [3,7] (n=233)	Blood donors; 2014-2018	Control for uninfected SARS-CoV-2 status; control for unvaccinated SARS-CoV-2 status	233 individuals	LS, LN, PNT values; gender; age
SERPICO [3] (n=23,886)	Transversal nationwide survey, residual sera; March 2020-March 2022	Application	23,886 individuals × dates	LS, LN, PNT values; gender; age; region

[a] LN: nucleocapsid.

[b] LS: spike protein.

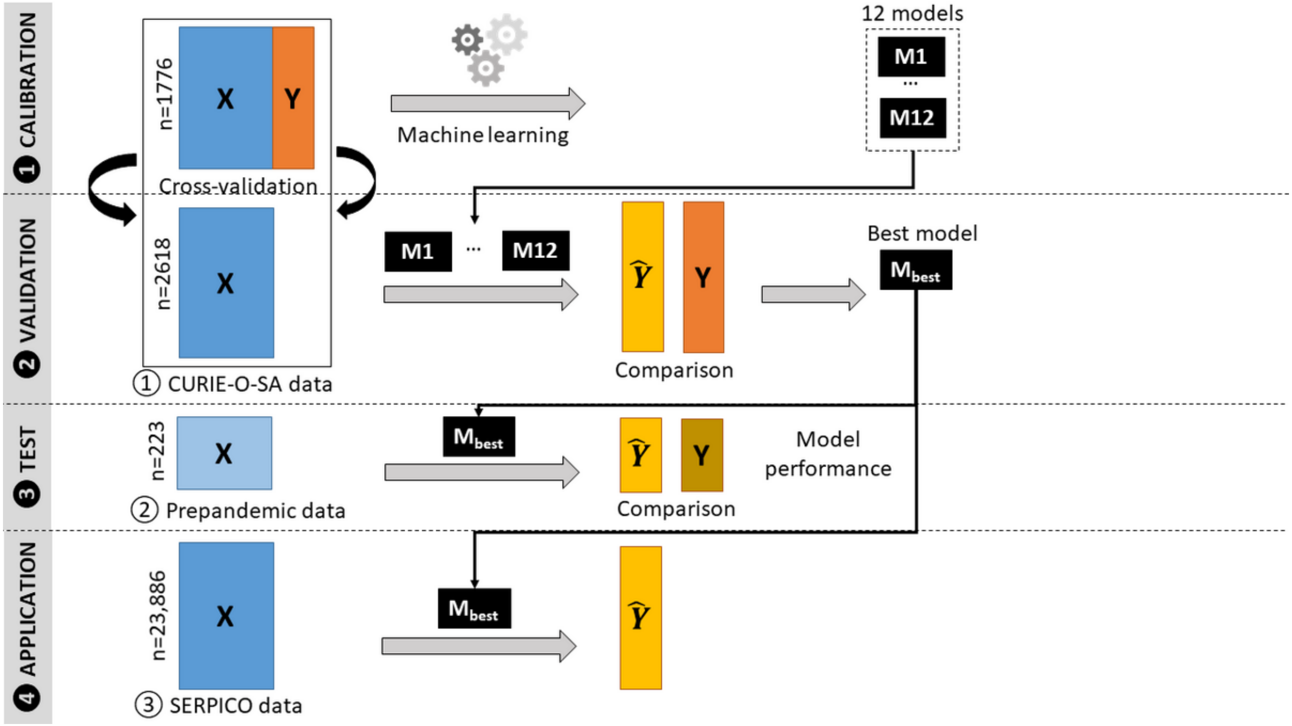
[c] PNT: pseudoneutralization assay.

Machine Learning Procedure

The model was built using a 4-step procedure involving calibration, validation, testing, and application (Figure 1). It aimed to predict the SARS-CoV-2 status at the individual level from 3 serological assays (LS, LN, and PNT) and 2 covariables (age and gender). The SARS-CoV-2 status was defined as a variable with 5 modalities: uninfected-unvaccinated (NEG), vaccinated with 1 dose (VAC1), vaccinated with 2 doses (VAC2), infected-unvaccinated (INF), and infected-vaccinated regardless of the number of doses (INF.VAC). A machine learning procedure was applied to challenge several classification models [8]. The model with the best predictive performance was selected from a cross-validation procedure on the first data set (CURIE-O-SA) then validated on an external data set (prepandemic). Finally, the selected model was applied to the French nationwide serosurveillance survey (SERPICO) to infer the proportions of the 5 SARS-CoV-2 statuses.

More precisely, 12 parametric and nonparametric classification models from 5 main statistical families were assessed within a machine-learning procedure. The models came from the following families: (1) generalized linear models (standard multinomial regression, penalized generalized linear model), (2) Bayesian models (naive Bayesian classification), (3) factor analysis models (discriminant analysis, mixture discriminant analysis, regularized discriminant analysis, kernel partial least squares discriminant analysis), (4) decision tree models (bagged classification and regression tree, Quinlan's C5.0 algorithm), and (5) nonparametric classification models (K nearest neighbors, support vector machine, neural network) [9]. To select the best predictive model and avoid overfitting, a repeated (200 times) 2-fold cross-validation was applied. To ensure the independence of the "individual x date" units, calibration data were randomly selected so that there was only 1 sample per individual involved in the CURIE-O-SA cohort; the remaining data were used for validation. The model parameters were optimized by means of a 10-fold cross-validation procedure on the calibration data. Percentages of well-classified samples for each modality of the status were calculated on the validation data. The best model was the one that best predicted the status modalities on average. Using the mean predictive performance of the 5 statuses to select the best model (rather than the overall prediction) can be considered oversampling of the rare modalities [10]. Repeated bootstrap simulations (50 times) were used to obtain 95% CIs. The overall procedure was implemented in R software (version 4.1.2, R Foundation for Statistical Computing) [11] by means of the "caret" package (version 6.0-90) [12].

Figure 1. Sketch graph of the 3-step model-building procedure (calibration, validation, and testing) and the application step. Regardless of the data set (ie, CURIE-O-SA, prepandemic, or SERPICO), the X explanatory data contained the $\log_{10}$ values of the 3 serological assays ($\log_{10}$LN, $\log_{10}$LS, and $\log_{10}$PNT) and the Y outcome data, with the SARS-CoV-2 status being known (CURIE-O-SA, prepandemic) or unknown (SERPICO).



Ethical Considerations

The study was based on a secondary use of pseudonymized data collected from health professionals and already published [3,7]. According to French law, such studies are not required to receive ethics committee approval.

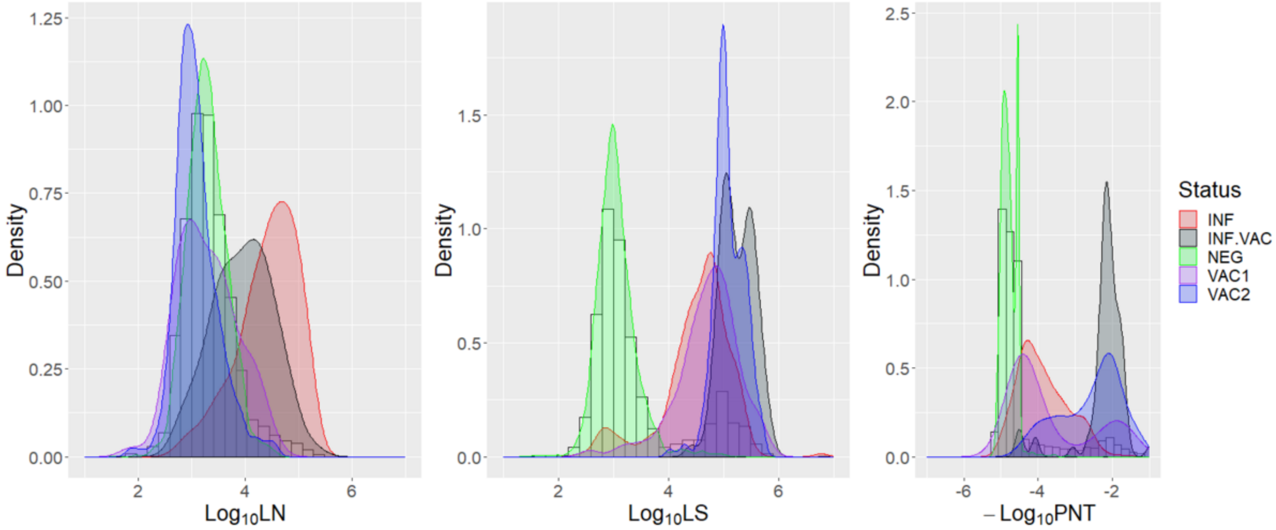
Results

Descriptive Statistics

The CURIE-O-SA data set used to build and validate the predictive model is illustrated in Figure 2. The serological LN results differentiated INF and to a lesser extent INF.VAC individuals from the others. The serological results targeting the SARS-CoV-2 LS separated the NEG individuals from the others. INF individuals developed anti-LS and anti-LN protein immunoglobulins, while VAC1 and VAC2 individuals developed anti-LS immunoglobulins only. This difference in immunization between infected (and unvaccinated) and vaccinated individuals was expected in a cohort where VAC1 and VAC2 individuals were vaccinated by vaccines targeting the LS only. The PNT results separated the INF.VAC individuals from the others as INF.VAC individuals presented a higher response to the PNT assay than did VAC1 and VAC2 individuals.

The prepandemic data set used for model validation consisted of 223 negative sera with the following average values for the serological assays ($\log_{10}$ values): 3.16 (SD 0.30) for LN, 3.16 (SD 0.20) for LS, and 5.07 (SD 0.11) for PNT.

Figure 2. Density estimation of the $\log_{10}$ values of the serological assays (LN, LS, and PNT) according to their known SARS-CoV-2 status in the CURIE-O-SA cohort. INF: infected-unvaccinated; INF.VAC: infected-vaccinated regardless of the number of doses; LN: nucleocapsid; LS: spike protein; NEG: uninfected-unvaccinated; PNT: pseudoneutralization assay; VAC1: vaccinated with 1 dose; VAC2: vaccinated with 2 doses.



Model

The percentages of well-classified results for each status and for all the models under study were calculated (Table 2).

The selected model was the mixture discriminant analysis with the best average prediction performance (mean 69.9%, SD 0.5%). More precisely, 98.4% (SD 0.2%) of the NEG status was well-predicted, as well as 40.6% (SD 6.9%) of the VAC1, 87.5% (SD 3.1%) of the VAC2, 42.6% (SD 11.3%) of the INF.VAC, and 78.9% (SD 3.8%) of the INF statuses. Best performances were obtained for the modalities with the largest numbers of “individual × date” units (NEG, VAC2, and INF). The VAC1 individuals were usually (60%) incorrectly predicted as VAC2; the error in the prediction did not depend on age or gender. The INF.VAC individuals were usually (84.4%) incorrectly predicted as VAC2; the error in the prediction did not depend on age and gender. As expected, this model predicted 99.5% of the negative prepandemic sera (Figure 3).

The model was applied to the prepandemic data (n=223). It predicted 222 results as negative and 1 as VAC1. This confirms the ability of the model to predict infection-free individuals, as 99.5% of the results were correctly predicted as NEG. This performance is in accordance with that calculated by means of cross-validation (98.4%). However, this performance is expected in that the model has many NEG values in the calibration data set (3532/4394; 80.4%) and is tested on expected NEG results. The only result not correctly predicted was predicted to be VAC1 with values $\log_{10}$LN=3.09, $\log_{10}$LS=4.44, and $\log_{10}$PNT=5.04. This individual had higher LS values than expected for NEG individuals ($\log_{10}$LN=3.28, $\log_{10}$LS=3.05, and $\log_{10}$PNT=4.78 for average NEG individuals in the CURIE-O-SA data set).

Table 2. Predictive performances of different models obtained by means of a repeated (200 times) 2-fold cross-validation using the CURIE-O-SA cohort (n=4394 individual × date units from April 2020-November 2021).

Model	Predictive performance (%), mean (SD)					
	NEG[a] (n=3532)	VAC1[b] (n=82)	VAC2[c] (n=477)	INF.VAC[d] (n=72)	INF[e] (n=231)	Average[f]
Multinomial regression	99.1 (0.2)	3.4 (4.7)	95.4 (1.5)	29.1 (9.8)	68.1 (2.9)	59 (41.8)
Penalized generalized linear model	99.1 (0.2)	1.9 (3.6)	95.4 (1.3)	24.8 (12.6)	68 (2.8)	57.8 (43.1)
Naive Bayesian classification	98.5 (0.2)	6.7 (4.7)	91.3 (1.9)	47.8 (8.7)	83.8 (1.8)	65.6 (38.3)
Linear discriminant analysis	98.7 (0.1)	31.1 (5.8)	81.1 (4)	45 (6.1)	77.9 (1.8)	66.8 (27.8)
Regularized discriminant analysis	98.3 (0.2)	22.9 (11.3)	91.5 (3)	46.9 (6.3)	82.5 (2.8)	68.4 (32.2)
Mixture discriminant analysis[g]	98.4 (0.2)	40.6 (6.9)	87.5 (3.1)	42.6 (11.3)	78.9 (3.8)	69.6 (26.5)
Kernel PLS[h] discriminant analysis	99.7 (0.1)	0 (0)	90.3 (1)	0 (0)	41.7 (3.1)	46.3 (47.7)
Bagged CART[i]	98.7 (0.2)	22.7 (7.8)	89.6 (2.6)	31.1 (8.8)	72.3 (4.6)	62.9 (34.3)
Quinlan's C5.0 algorithm	98.8 (0.4)	13.3 (9.6)	91.6 (3.5)	29 (14.5)	72.4 (7.8)	61 (38.1)
K nearest neighbors	98.8 (0.2)	18.9 (7.7)	92 (2.4)	31.4 (9.9)	70.2 (4.3)	62.8 (35.8)
Support vector machine	98.9 (0.2)	1.1 (2.6)	95.6 (1.5)	23.3 (15.8)	76.2 (2.8)	59 (44.3)
Neural network	98.7 (0.3)	14.1 (10.7)	94 (2.4)	30.7 (15.6)	77.1 (5.5)	62.9 (38.3)

[a] NEG: uninfected-unvaccinated.

[b] VAC1: vaccinated with 1 dose.

[c] VAC2: vaccinated with 2 doses.

[d] INF.VAC: infected-vaccinated regardless of the number of doses.

[e] INF: infected-unvaccinated.

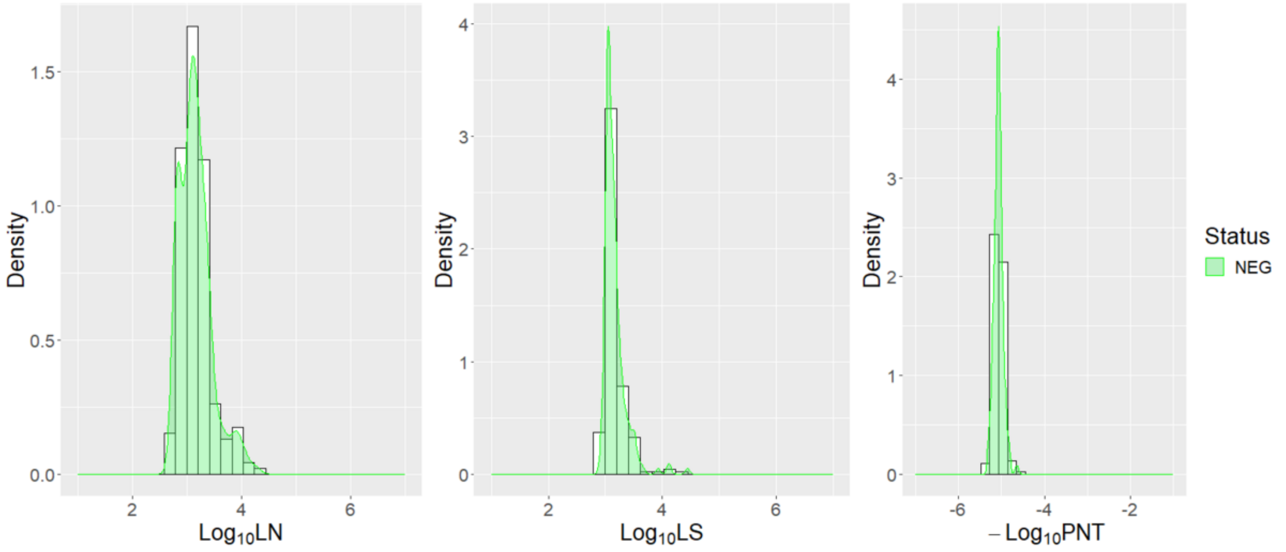
[f] Mean predictive performance across the 5 statuses.

[g] Selected model.

[h] PLS: partial least squares.

[i] CART: classification and regression tree.

Figure 3. Density estimation of the $\log_{10}$ values of the serological assays (LN, LS, and PNT) according to their known SARS-CoV-2 status in the prepandemic data set. LN: nucleocapsid; LS: spike protein; NEG: uninfected-unvaccinated; PNT: pseudoneutralization assay.



Prediction for the SERPICO Serosurvey

The model was applied to the French nationwide serosurveillance survey SERPICO (n=23,886). The consistency of the predictions was confirmed by comparing the predicted results to internal and external reference information (Multimedia Appendix 1). It follows that the estimation of the vaccinated proportion of the French population was correctly predicted, especially between March 2020 and June 2021. Predictions were less accurate for the last 2 periods as the proportion of individuals with a complete vaccination scheme

(VAC2 and INF.VAC) tended to be underestimated in the prediction.

The average percentages of the 5 predicted SARS-CoV-2 immune statuses are shown in Figure 4 for the 8 sampling periods. For the first 4 periods, individuals were mostly predicted as NEG (March 2020: mean 99.1%, SD 0.2%; October 2020: mean 94.7%, SD 0.4%). Between February 2021 and March 2022, the number of INF individuals increased from 7.5% (SD 0.4%) to 16.8% (SD 0.7%). From June 2021, the numbers of vaccinated (VAC1 and VAC2) and INF.VAC individuals increased in relation with the rollouts of vaccination for the older population from the end of December 2020, for any person older than 12 years from June 2021, and for children aged 5 to 11 years from December 2021 in France. The prediction performances of the model were usually better for the most common statuses (NEG and INF). The INF.VAC status was better predicted in the 4 later periods during the vaccination rollout.

No difference in SARS-CoV-2 predicted status was observed between genders, except in October 2021 (Figure 5, left panel). At this time, the percentage of predicted NEG men was higher than that of women. No other significant difference was observed due to large CIs associated with the VAC1, VAC2, and INF.VAC predictions.

The percentage of INF individuals was similar in all age groups over the 7 sampling periods (Figure 5, right panel). In June 2021, the proportions of individuals predicted as VAC1 and VAC2 increased first in older age groups (60-69, 70-79, and older than 80 years) as vaccination was rolled out for these age groups first. Conversely, the proportion of individuals predicted to be NEG remained higher for the younger age groups during the same period. The proportion of NEG individuals was still higher in children aged 0 to 9 years in comparison with other age groups in October 2021 because vaccination was open only to children aged 5 to 11 years.

Predictions for the 5 SARS-CoV-2 statuses can be illustrated according to French administrative regions (Figure 6). The epidemic first progressed in the Eastern part of France between March and May 2021 and then spread to the Paris region and to the northern part of France in October 2021. Spatial variability was observed in immune status, with the predicted fraction of vaccinated individuals being higher in the western part of France than in other regions in June and October 2021.

Figure 4. Percentages (95% CI) of SARS-CoV-2 predicted status for the 8 collection periods. INF: infected-unvaccinated; INF.VAC: infected-vaccinated regardless of the number of doses; NEG: uninfected-unvaccinated; VAC1: vaccinated with 1 dose; VAC2: vaccinated with 2 doses.

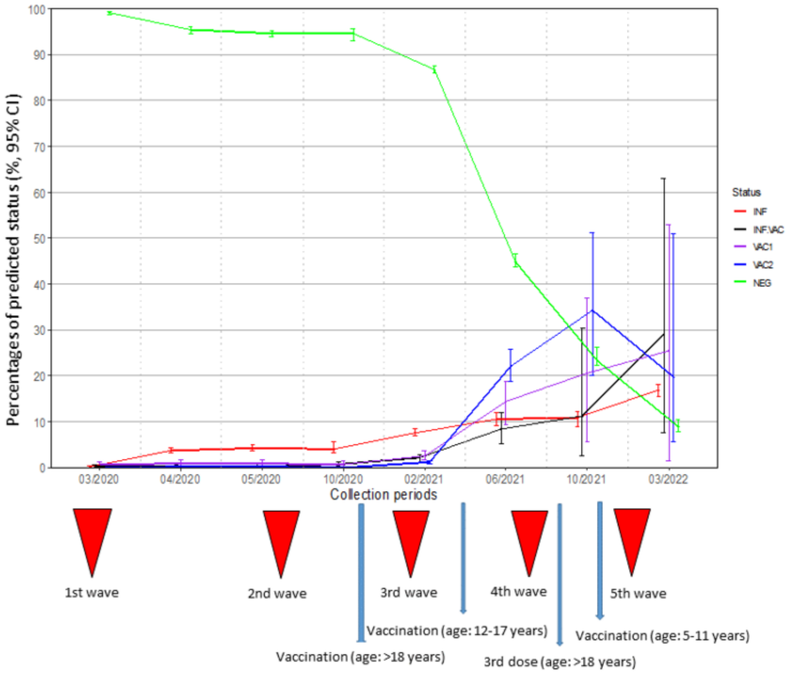


Figure 5. Percentages (95% CI) of SARS-CoV-2 predicted statuses for the 8 collection periods according to gender (left panel) and age group (right panel) in the SERPICO data set. INF: infected-unvaccinated; INF.VAC: infected-vaccinated regardless of the number of doses; NEG: uninfected-unvaccinated; VAC1: vaccinated with 1 dose; VAC2: vaccinated with 2 doses.

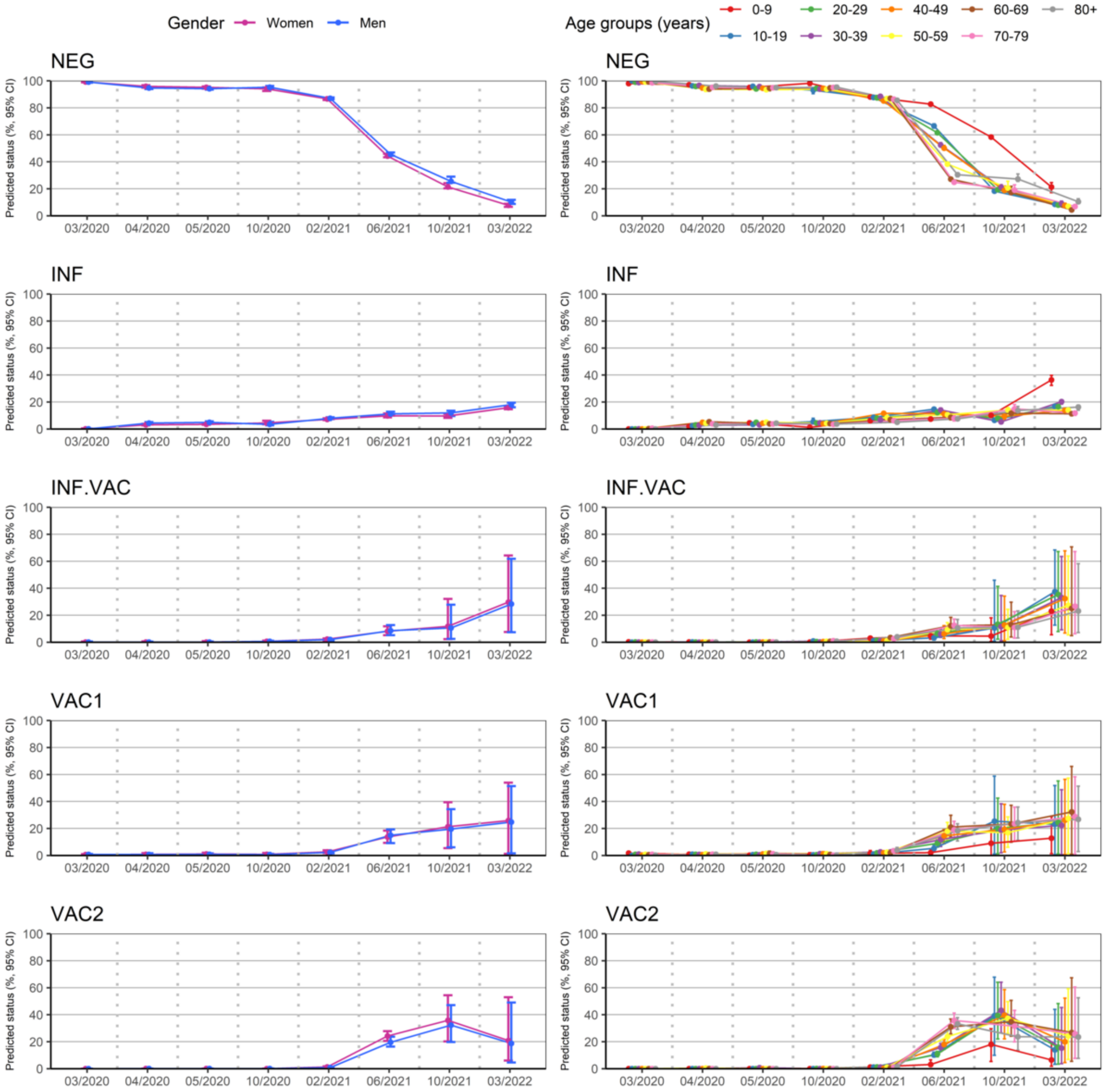
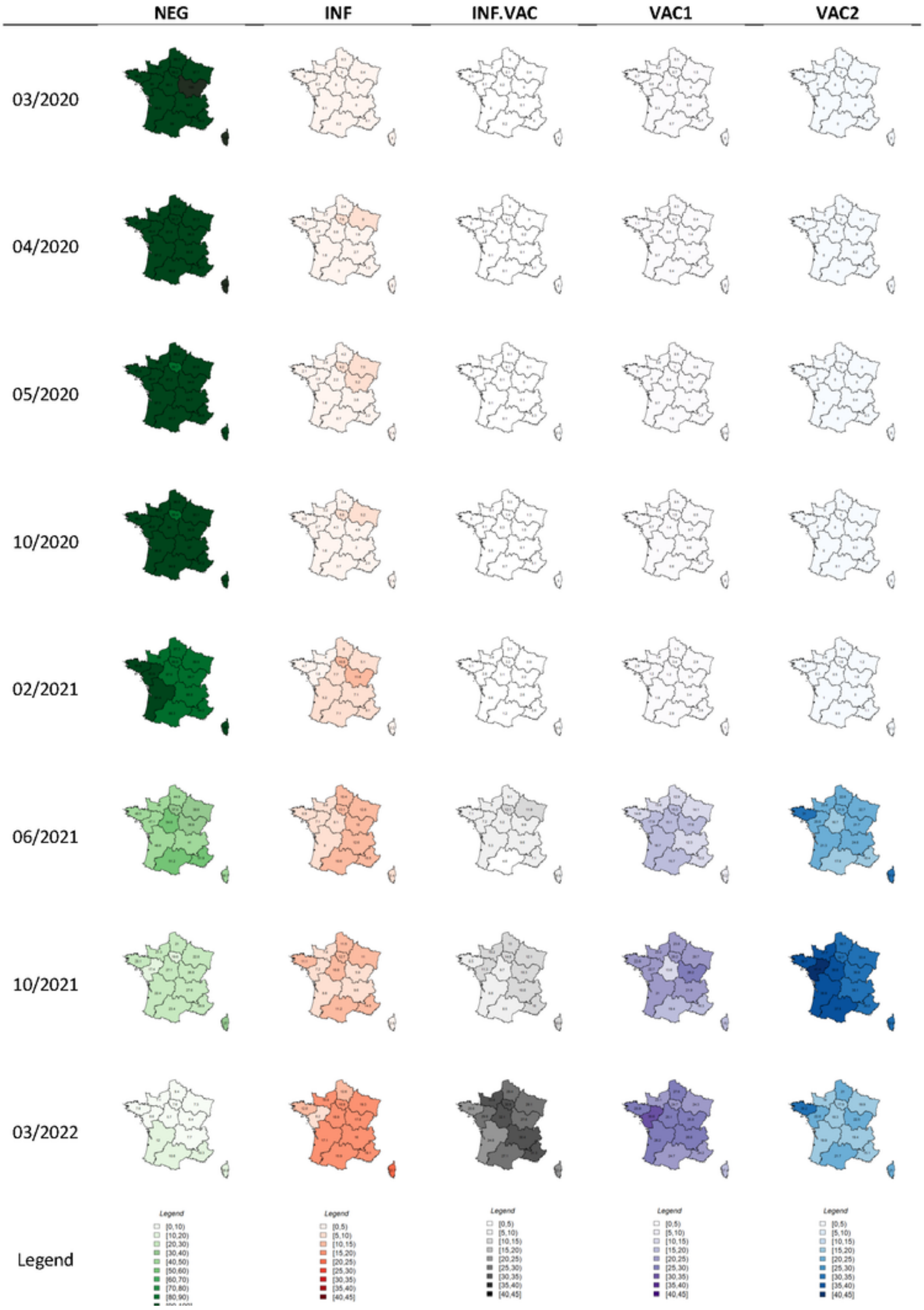


Figure 6. Percentages of SARS-CoV-2 predicted status for the 8 collection periods according to regions in France for the SERPICO data set. INF: infected-unvaccinated; INF.VAC: infected-vaccinated regardless of the number of doses; NEG: uninfected-unvaccinated; VAC1: vaccinated with 1 dose; VAC2: vaccinated with 2 doses. For a higher-resolution version of this figure, see Multimedia Appendix 2.



Discussion

Serological Assays

The objective of our study was to predict the infected or vaccinated status of the individuals enrolled in the SARS-CoV-2 serosurveillance survey without prior information on their SARS-CoV-2 infection or vaccination status. For that purpose, reference data on serological profiles of seronegative, infected, and vaccinated individuals from another French cohort were used to build the predictive model. This approach was possible because serological testing was carried out using the same assays developed by the National Reference Centre for Respiratory Diseases of the Institut Pasteur, ensuring the comparability of the serological results between studies.

LuLISA-S is a sensitive assay for identifying individuals exposed to SARS-CoV-2 infection or vaccination. As the vaccines used in France target the SARS-CoV-2 spike protein only, we were confident in the ability of the LuLISA-N assay to distinguish naturally infected individuals from vaccinated ones in the CURIE-O-SA and SERPICO surveys. The LN and LS assays were studied in the context of infection and vaccination [13]. What was less expected was the determinant role of the PNT assay in differentiating INF.VAC individuals from others. The infected status was characterized by a higher pseudoneutralization capacity than that in vaccinated individuals only, confirming previous observations of the higher protection conferred by infection followed by vaccination [4]. Therefore, the 3 serological assays were complementary and essential to obtain a clear distinction between all SARS-CoV-2 statuses in the population, despite the fact that using 3 assays is costly and time-consuming for serosurveillance purposes.

Machine Learning and Data-Driven Analysis

To build models predicting the SARS-CoV-2 status from serological assays, a machine learning procedure was applied. This procedure made it possible to compare a large number of classification models and select the most predictive one. This kind of procedure was previously used for prediction of SARS-CoV-2 status based on serological test results in a population vaccinated with whole virion vaccines [8].

The proposed procedure was based on raw serological results without setting any thresholds, contrary to previous works [3]. The thresholds used in the SERPICO survey were designed to maximize the specificity of the assays in a context of a low prevalence of SARS-CoV-2 infection in France from March 2020 to April 2021. In this study, adopting a data-driven strategy without any assay threshold enabled us to capture the evolution of the SARS-CoV-2 epidemic without modifying the assay interpretations over time [14]. The proposed model exhibited similar performance to the standard method for distinguishing negative versus positive SARS-CoV-2 immune statuses; however, it provided additional information about the positive status (ie, vaccinated, infected, or both; Multimedia Appendix 1).

To ease the use of the model, an R-shiny application was developed for the National Reference Centre and implemented in a user-friendly environment. Input data from the serological assays can be uploaded as a data frame and the most likely SARS-CoV-2 status is given [15].

Vaccination Impact on Seroprevalence

Application of the predictive model to serosurveillance results gave useful insights for interpreting the evolution of SARS-CoV-2 seroprevalence in France. Between June and October 2020, the proportion of individuals who were seronegative to SARS-CoV-2 infection was still higher than 90% despite the first 2 SARS-CoV-2 waves that greatly impacted health services in March and September 2020. The observed quasistability of seroprevalence over this period could be explained by an increase in seroprevalence due to the waves of infections, tempered by the fairly rapid decrease in anti–SARS-CoV-2 antibodies.

Between October 2020 and March 2022, the proportion of seronegative individuals decreased. As the proposed predictions showed that the proportion of INF individuals did not increase over the same period, the decrease of the seronegative population could be mainly attributed to vaccination. Additionally, the decrease of the NEG population occurred earlier in the older age groups (older than 60 years) who were targeted first by the vaccination program. A part of this older population also benefited from a hybrid immunity due to vaccination and infection, although to a lesser degree than the rest of the adult population. On the contrary, two-thirds of children aged 0 to 9 years were still seronegative to SARS-CoV-2 in October 2021 due to a later rollout of vaccination for children aged 5 to 11 years. The expected percentages infected children aged 0 to9 years increased from 1.3% to 36.4% from October 2021 to March 2021. This clearly shows the impact of the Omicron strain on this poorly vaccinated population. All together, these results confirm that population immunity toward SARS-CoV-2 infection would progress very slowly without vaccination in the French population [16]. Moreover, such a strategy—necessarily combined with continued restrictive measures aimed to avoid overwhelming the health care system—would have had a tremendous impact on the economy and mental health.

Consistency of the Predictions

The proposed predictive model produced SARS-CoV-2 immune status predictions in accordance with the observed SARS-CoV-2 epidemiological situation in France from March 2020 to March 2022 as results were only given in terms of infection prevalence (Multimedia Appendix 1). Predictions by region, gender, and age were consistent with the epidemiological weekly observations of the SARS-CoV-2 epidemic [17].

The consistency of the predictions was largely due to the use of 3 complementary serological assays that enabled us to finely distinguish the 5 SARS-CoV-2 statuses through a machine learning procedure. The main limit of the predictive models is associated with the data from the cohort used for calibration and validation of the models. Indeed, women and middle-aged individuals were overrepresented in the CURIE-O-SA cohort. The lack of profile diversity in the calibration and validation data set may explain why including the covariates (gender and age) did not improve model predictions. In addition, the CURIE-O-SA study took place from April 2020 to November 2021; therefore, it did not capture key evolutions in the SARS-CoV-2 epidemic in France, such as the emergence of the Omicron variant (November 2021) that provoked a different response to the PNT assay used in our study. In addition, infected and vaccinated individuals became more frequent in the overall population over time. This profile of individuals and individuals vaccinated with 1 dose were rare in the CURIE-O-SA cohort, leading to poor accuracy in the prediction of those statuses. The high percentages of individuals predicted to be vaccinated with 1 dose in October 2021 and March 2022 may have been due to the waning immunity of individuals vaccinated with 2 doses for a long time [5,18-20]. Lastly, the CURIE-O-SA cohort did not cover the rollout of the third vaccination dose. The validity limits of the predictive models were thus reached in October 2021, as exemplified by the overlapping CIs of the different percentages of vaccination statuses. The collection of new reference results including both results for the 3 serological assays and SARS-CoV-2 infection and vaccination history of the individuals is needed to update the model. Maintaining serosurveillance distinguishing vaccinated or infected populations is of interest to describe the evolution of SARS-CoV-2 immunity in the overall population and to understand immunity waning over time, but additional samples are needed for that purpose.

Conclusion and Perspectives

A predictive model of individual vaccination or infection status with respect to SARS-CoV-2 was proposed based on 3 complementary serological assays and based on a machine learning procedure. This model was applied to the French nationwide serosurveillance survey from 2020 to 2022 to estimate the proportions of the French population that were seronegative, infected, vaccinated (1 or 2 doses), or infected and vaccinated, as this data set included no prior information on the SARS-CoV-2 infection or vaccination status of the individuals. This allowed us to follow the level of SARS-CoV-2 infection and the vaccine response profile of the French population over time.

Combining the results from the serosurveillance survey with previously acquired results from a cohort studied longitudinally improved the information retrieved from serosurveillance while keeping its protocol simple and easy to implement (no need to collect SARS-CoV-2 information on a large sample of individuals). We think that such a combination strategy is of interest to improve serosurveillance of emerging vaccine-preventable diseases. The results of our predictive model make it possible to measure the crucial contribution of SARS-CoV-2 vaccination to rapidly reach a level of collective immunity that has made it possible to relax sanitary measures without overloading the health care system. Indeed, population immunity toward SARS-CoV-2 infection would have progressed very slowly without vaccination.

Acknowledgments

We thank from the Institut Pasteur, the Recombinant Protein Production and Purification core facility for SARS-CoV-2 protein preparation, the Molecular Biophysics core facility for their quality checking, and Yves L. Janin (Unit of Chemistry and Biocatalysis) for providing the Hikarazine luciferase-linked immunosorbent assay substrate. We also thank the whole Integrated Collections for Adaptive Research in Biomedicine (ICAReB) team and COVID-19 support staff at the Institut Pasteur, the team from the Eurofins Biomnis Sample Library, and the team from CerbaHealthcare for serum sample management.

We also thank the personnel of the Institut Curie who volunteered to participate to the CURIE-O-SA study, which was set up and managed by the staff of the clinical and laboratory departments of the Institut Curie.

The blood and clinical study at the Institut Curie was funded by Fondation de France, Agence Nationale de la Recherche (grant ANR-21-COVR-002) and institutional funding from the Institut Curie.

Authors' Contributions

SB, AHS, HN, and SBS wrote the manuscript. OL and HN designed the study. SB and VA analyzed and visualized the data. TR, SVDW, MA, CD, SG and BCC performed laboratory investigations. MA, JBR, CD, JP, VA, SLV, SG, VG, BCC, GJ, NR, SVDW, OL, and TR revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consistency of the SERPICO predictions.
[DOCX File , 20 KB - publichealth_v9i1e46898_app1.docx]

Multimedia Appendix 2
Percentages of SARS-CoV-2 predicted status for the 8 collection periods according to regions in France for the SERPICO data set. INF: infected-unvaccinated; INF.VAC: infected-vaccinated regardless of the number of doses; NEG: uninfected-unvaccinated; VAC1: vaccinated with 1 dose; VAC2: vaccinated with 2 doses.
[PDF File (Adobe PDF File), 903 KB - publichealth_v9i1e46898_app2.pdf]

References

1. Murhekar MV, Clapham H. COVID-19 serosurveys for public health decision making. Lancet Glob Health 2021 May;9(5):e559-e560 [FREE Full text] [doi: 10.1016/S2214-109X(21)00057-7] [Medline: 33705691]
2. Metcalf J, Paireau J, O'Driscoll M, Pivette M, Hubert B, Pontais I, et al. Comparing the age and sex trajectories of SARS-CoV-2 morbidity and mortality with other respiratory pathogens. R Soc Open Sci 2022 Jun;9(6):211498 [FREE Full text] [doi: 10.1098/rsos.211498] [Medline: 35719888]
3. Le Vu S, Jones G, Anna F, Rose T, Richard J, Bernard-Stoecklin S, et al. Prevalence of SARS-CoV-2 antibodies in France: results from nationwide serological surveillance. Nat Commun 2021 May 21;12(1):3025 [FREE Full text] [doi: 10.1038/s41467-021-23233-6] [Medline: 34021152]
4. Hall V, Foulkes S, Insalata F, Kirwan P, Saei A, Atti A, et al. Protection against SARS-CoV-2 after Covid-19 Vaccination and Previous Infection. N Engl J Med 2022 Mar 31;386(13):1207-1220. [doi: 10.1056/nejmoa2118691]

5. Goldberg Y, Mandel M, Bar-On YM, Bodenheimer O, Freedman LS, Ash N, et al. Protection and waning of natural and hybrid immunity to SARS-CoV-2. N Engl J Med 2022 Jun 09;386(23):2201-2212. [doi: 10.1056/nejmoa2118946]
6. Wei J, Matthews PC, Stoesser N, Diamond I, Studley R, Rourke E, et al. SARS-CoV-2 antibody trajectories after a single COVID-19 vaccination with and without prior infection. Nat Commun 2022 Jun 29;13(1):3748 [FREE Full text] [doi: 10.1038/s41467-022-31495-x] [Medline: 35768431]
7. Anna F, Goyard S, Lalanne AI, Nevo F, Gransagne M, Souque P, et al. High seroprevalence but short-lived immune response to SARS-CoV-2 infection in Paris. Eur J Immunol 2021 Jan;51(1):180-190 [FREE Full text] [doi: 10.1002/eji.202049058] [Medline: 33259646]
8. Singh P, Ujjainiya R, Prakash S, Naushin S, Sardana V, Bhatheja N, Promila, et al. A machine learning-based approach to determine infection status in recipients of BBV152 (Covaxin) whole-virion inactivated SARS-CoV-2 vaccine for serological surveys. Comput Biol Med 2022 Jul;146:105419 [FREE Full text] [doi: 10.1016/j.compbiomed.2022.105419] [Medline: 35483225]
9. Carobene A, Famiglini L, Sabetta E, Naclerio A, Banfi G. Machine learning and laboratory values in the diagnosis, prognosis and vaccination strategy of COVID-19. In: Lidströmer N, Eldar YC, editors. Artificial Intelligence in Covid-19. Cham: Springer; 2022:121-156.
10. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. New York: Springer; 2009.
11. Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H, Bing G. Learning from class-imbalanced data: review of methods and applications. Expert Syst Appl 2017 May;73:220-239. [doi: 10.1016/j.eswa.2016.12.035]
12. The R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2023-10-25]
13. Kuhn M. Caret: classification and regression training. CRAN. URL: https://cran.r-project.org/web/packages/caret/index.html [accessed 2023-10-25]
14. Si Y, Covello L, Wang S, Covello T, Gelman A. Beyond vaccination rates: a synthetic random proxy metric of total SARS-CoV-2 immunity seroprevalence in the community. Epidemiology 2022 Jul 01;33(4):457-464 [FREE Full text] [doi: 10.1097/EDE.0000000000001488] [Medline: 35394966]
15. Léon L, Pillonel J, Jauffret-Roustide M, Barin F, Le Strat Y. Estimating prevalence from dried blood spots without using biological cut-offs: application of a novel approach to hepatitis C virus in drug users in France (ANRS-Coquelicot survey). Epidemiol Infect 2019 Jun 17;147:1-7. [doi: 10.1017/s0950268819001043]
16. Bougeard S. Prediction of Individual Vaccination and/or Infection Status With Respect to SARS-CoV-2 According to 3 Serological Assays (LN, LS, PNT). URL: https://sbougeard.shinyapps.io/applishinysarscovtwo [accessed 2023-10-25]
17. Salje H, Tran Kiem C, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. Science 2020 Jul 10;369(6500):208-211 [FREE Full text] [doi: 10.1126/science.abc3517] [Medline: 32404476]
18. COVID 19 point epidemiologique. Santé Publique France. URL: https://www.santepubliquefrance.fr/recherche/#search=COVID%2019%20%20%20point%20epidemiologique&publications=donn%C3%A9es&regions=National&sort=date [accessed 2023-10-25]
19. Levin EG, Lustig Y, Cohen C, Fluss R, Indenbaum V, Amit S, et al. Waning immune humoral response to BNT162b2 Covid-19 vaccine over 6 months. N Engl J Med 2021 Dec 09;385(24):e84. [doi: 10.1056/nejmoa2114583]
20. Gaebler C, Wang Z, Lorenzi JCC, Muecksch F, Finkin S, Tokuyama M, et al. Evolution of antibody immunity to SARS-CoV-2. Nature 2021 Mar;591(7851):639-644 [FREE Full text] [doi: 10.1038/s41586-021-03207-w] [Medline: 33461210]

Abbreviations

INF: infected-unvaccinated
INF.VAC: infected-vaccinated regardless of the number of doses
LN: nucleocapsid
LS: spike protein
LuLISA: luciferase-linked immunosorbent assay
NEG: uninfected-unvaccinated
PNT: pseudoneutralization assay
RT-PCR: reverse transcription polymerase chain reaction
VAC1: vaccinated with 1 dose
VAC2: vaccinated with 2 doses

Edited by A Mavragani, T Sanchez; submitted 02.03.23; peer-reviewed by L Espinosa, T Rana; comments to author 11.05.23; revised version received 30.05.23; accepted 20.06.23; published 28.11.23.

Please cite as:
Bougeard S, Huneau-Salaun A, Attia M, Richard JB, Demeret C, Platon J, Allain V, Le Vu S, Goyard S, Gillon V, Bernard-Stoecklin S, Crescenzo-Chaigne B, Jones G, Rose N, van der Werf S, Lantz O, Rose T, Noël H
Application of Machine Learning Prediction of Individual SARS-CoV-2 Vaccination and Infection Status to the French Serosurveillance Survey From March 2020 to 2022: Cross-Sectional Study
JMIR Public Health Surveill 2023;9:e46898
URL: https://publichealth.jmir.org/2023/1/e46898
doi:10.2196/46898
PMID:38015594

Original Paper

Comparing Telephone Survey Responses to Best-Corrected Visual Acuity to Estimate the Accuracy of Identifying Vision Loss: Validation Study

John Wittenborn[1], BSc, BA; Aaron Lee[2], MD, MSc; Elizabeth A Lundeen[3], PhD; Phoebe Lamuda[1], SM; Jinan Saaddine[3], MD, MPH; Grace L Su[2], MD; Randy Lu[2], BSc; Aashka Damani[2], BSE; Jonathan S Zawadzki[2], BA; Colin P Froines[2], BA; Jolie Z Shen[2], MD; Timothy-Paul H Kung[2], MBA; Ryan T Yanagihara[2], MD; Morgan Maring[2], MS; Melissa M Takahashi[2], BS; Marian Blazes[2], MD; David B Rein[1], PhD

[1]Public Health Analytics, National Opinion Research Center at the University of Chicago, Chicago, IL, United States
[2]Department of Ophthalmology, University of Washington, Seattle, WA, United States
[3]Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
John Wittenborn, BSc, BA
Public Health Analytics
National Opinion Research Center at the University of Chicago
55 East Monroe Street, 30th Floor
Chicago, IL, 60603
United States
Phone: 1 3125195718
Fax: 1 3127594004
Email: wittenborn-john@norc.org

Abstract

Background: Self-reported questions on blindness and vision problems are collected in many national surveys. Recently released surveillance estimates on the prevalence of vision loss used self-reported data to predict variation in the prevalence of objectively measured acuity loss among population groups for whom examination data are not available. However, the validity of self-reported measures to predict prevalence and disparities in visual acuity has not been established.

Objective: This study aimed to estimate the diagnostic accuracy of self-reported vision loss measures compared to best-corrected visual acuity (BCVA), inform the design and selection of questions for future data collection, and identify the concordance between self-reported vision and measured acuity at the population level to support ongoing surveillance efforts.

Methods: We calculated accuracy and correlation between self-reported visual function versus BCVA at the individual and population level among patients from the University of Washington ophthalmology or optometry clinics with a prior eye examination, randomly oversampled for visual acuity loss or diagnosed eye diseases. Self-reported visual function was collected via telephone survey. BCVA was determined based on retrospective chart review. Diagnostic accuracy of questions at the person level was measured based on the area under the receiver operator curve (AUC), whereas population-level accuracy was determined based on correlation.

Results: The survey question, “Are you blind or do you have serious difficulty seeing, even when wearing glasses?” had the highest accuracy for identifying patients with blindness (BCVA ≤20/200; AUC=0.797). The highest accuracy for detecting any vision loss (BCVA <20/40) was achieved by responses of “fair,” “poor,” or “very poor” to the question, “At the present time, would you say your eyesight, with glasses or contact lenses if you wear them, is excellent, good, fair, poor, or very poor” (AUC=0.716). At the population level, the relative relationship between prevalence based on survey questions and BCVA remained stable for most demographic groups, with the only exceptions being groups with small sample sizes, and these differences were generally not significant.

Conclusions: Although survey questions are not considered to be sufficiently accurate to be used as a diagnostic test at the individual level, we did find relatively high levels of accuracy for some questions. At the population level, we found that the relative prevalence of the 2 most accurate survey questions were highly correlated with the prevalence of measured visual acuity loss among nearly all demographic groups. The results of this study suggest that self-reported vision questions fielded in national surveys are likely to yield an accurate and stable signal of vision loss across different population groups, although the actual measure of prevalence from these questions is not directly analogous to that of BCVA.

(JMIR Public Health Surveill 2023;9:e44552) doi:10.2196/44552

KEYWORDS

vision; blindness; surveillance; survey; acuity; validation; visual health; optometry clinic; eye disease; vision loss

Introduction

Vision loss and blindness affect approximately 7.08 million Americans and cost the US economy US $134 billion per year [1,2]. Improving and targeting public health and medical interventions to reduce this burden depend on measuring changes in vision loss and blindness prevalence over time at the national, state, and local level. The 2016 National Academies of Sciences, Engineering, and Medicine report *Making Eye Health a Population Health Imperative* [3] highlighted the role that epidemiological surveillance of vision loss could play in mitigating the burden of vision loss. This report included recommendations for the Centers for Disease Control and Prevention (CDC) to develop case definitions to measure visual health, evaluate the utility of self-reported vision measures, and establish a national surveillance system to track vision and eye health indicators in available data sources [3]. A particular challenge to meeting these recommendations is reconciling the different types of data and case definitions used to identify vision loss and blindness prevalence. Historically, most estimates of the prevalence of vision loss and blindness available in the published literature are derived from either population-based examination studies that objectively measure visual function in a defined geographic area and apply these rates to the national population or from nationally representative surveys that measure self-reported vision loss through responses to a wide variety of survey questions [4-8].

Direct measurement of best-corrected visual acuity (BCVA) in the better-seeing eye as evaluated by a trained eye-care technician is accepted as the gold standard method to determine vision impairment or blindness at the person level. Several high-quality population-based studies have attempted to measure BCVA among all persons or persons of a specific race/ethnicity group residing within specific localities [5,6,9-11]. The National Health and Nutrition Examination Survey (NHANES), which approximated BCVA using autorefractors from 1999-2008, is the only examination study derived from a nationally representative sample [4,12]. However, the high cost of examination studies limits their sample size and collection frequency, and study recruitment logistics result in the exclusions of important populations including younger persons, persons with risk factors, the oldest age groups, and persons living in institutional settings such as nursing homes. As a result, examination studies have limited capacity to measure and monitor prevalence for the entire population, identify trends or disparities over time among sociodemographic and risk factor groups, or identify variation in prevalence at the state or local level.

In contrast, survey-based measures of self-reported vision loss and blindness offer advantages of timeliness and much larger sample sizes and typically include richer covariate information on demographic, socioeconomic, behavioral, access to care, and health risk factors. At least 16 nationally representative surveys include self-reported vision questions, including several at the state level and one at the county level [13,14]. Given their low cost of administration, self-reported vision questions can be extremely useful for population health surveillance if they yield a consistent signal of vision problems at the population level, even if they are less precise at the individual level. However, the large number of different survey questions and differences in how they are administered in different surveys result in a wide range of estimates, and experts have expressed skepticism about their accuracy [15,16]. Currently, the validity and correlation of self-reported vision for determining measured visual function is unknown.

Two recent surveillance estimates of the prevalence of vision loss by the CDC’s Vision and Eye Health Surveillance System (VEHSS) and the World Health Organization’s (WHO) Global Burden of Disease project used statistical modeling approaches to combine information from both examination- and survey-based data sources [17]. The VEHSS composite prevalence estimates of blindness and visual acuity loss used Bayesian meta-analytic methods to combine prevalence information from multiple published examination studies that measured BCVA, benchmarked to the NHANES examination measures. The analysis then used self-reported information from the American Community Survey and the National Survey of Children’s Health to predict visual acuity loss prevalence among populations that were excluded or insufficiently captured by examination-based data, including children, older age groups, and persons living in institutional settings, as well as to allocate prevalence by state and county based on the 2019 US population [2]. The WHO estimates similarly combine examination- and survey-based data to estimate burden in WHO regions for which insufficient data exist.

However, more research on the relationship between self-reported vision problems and measured BCVA is vital to better understand the utility of these questions for public health surveillance. Determining the extent to which variation in self-reported vision problems correlates to variation in underlying visual function will support the ongoing development and interpretability of novel surveillance estimates that leverage multiple types of data. In this paper, we sought to estimate the accuracy of self-reported vision loss measures fielded in national surveys by comparing their results to a gold standard of evaluated BCVA at both the individual and population level.

Methods

We conducted a retrospective chart review and concurrent telephone interview among a sample of patients from the University of Washington (UW) ophthalmology and optometry clinics and calculated statistics to inform the accuracy and validity of telephone survey self-reported visual function compared to examination-measured BCVA. We evaluated findings both at the person level, where we assessed the diagnostic accuracy of the survey questions for identifying BCVA loss, and at the population level, where we compared sample-level prevalence rates derived from survey responses to those based on BCVA.

Participant Sample

We sampled study participants from current or recent patients from the UW ophthalmology and optometry clinics. First, we restricted a patient-level electronic health record file to only include patients with a visit from May 2018 through April 2020. We then sampled 1200 patients for inclusion such that approximately 20% had diagnoses for age-related macular degeneration, 20% had diagnoses for diabetic retinopathy, and 20% had diagnoses for glaucoma. Among these patients, we attempted to oversample patients with severe or vision-threatening stages of disease. We sampled for 2 additional cohorts of 20% each, which excluded any patients with diagnoses of age-related macular degeneration, diabetic retinopathy, or glaucoma: one cohort in which patients had presenting visual acuity loss of 20/40 or worse in the better-seeing eye and another cohort in which patients had no record of visual acuity loss. The diagnosed prevalence of cataract was so high that we did not need to oversample for patients with this condition. The overall sample size was too low to specifically oversample for additional diseases, so the prevalence of these conditions in the sample was based on the background rate of selected patients. Between the summer of 2020 and the spring of 2021, UW ophthalmology students and residents recruited and obtained informed consent from 669 patients, of whom 438 (65.5%) completed the telephone survey interview.

Chart Review

We abstracted chart information for all 669 recruited patients from October 2020 until March 2021. The project team developed a secure, web-based form that allowed chart abstractors to complete their reviews remotely (due to COVID-19) and automatically aggregate and store the review information. The form was reviewed by an external panel of experts in ophthalmology, and recommended revisions were incorporated. The form captured presenting (or habitual) acuity and BCVA in each eye and contained additional free-text fields designed to capture patient vision information from electronic health record assessment and treatment plan information. The abstractions were conducted by 2 UW ophthalmology residents (GLS and RL) under the supervision of a professor (AL). Study authors manually reviewed the acuity fields and assigned a single *logMar* value to each patient based on BCVA in the better-seeing eye. These values were re-reviewed by the original abstractors. For reporting purposes, acuity is shown in Snellen acuity ratios.

Telephone Survey

A telephone survey was administered to consented patients in 2 rounds (November 2020 and May 2021) to minimize the amount of time between the survey and the patients' recruitment and consent. Nonrespondents from the first round were recontacted in the second round. The survey was administered by professional interviewers with experience conducting federally funded national surveys. The survey included sections with questions related to patient sociodemographic characteristics, eye diseases, visual function, and vision-related difficulties with activities of daily living (ADL). The question order within each section was randomized for each respondent.

To select questions to include in the survey, we first identified vision-related survey questions from 16 recent or ongoing national surveys [14]. We reviewed the questions with the CDC's Vision Health Initiative and an expert advisory panel convened under the VEHSS, as well as experts in telephone survey administration and psychometrics at NORC at the University of Chicago. We selected three visual function questions from current or recent national surveys: (Question [Q] 1) "Are you blind, or do you have serious difficulty seeing, even when wearing glasses?" which has been administered by the American Community Survey since 2008 and the Behavioral Risk Factor Surveillance System core module since 2013. (Q2) "At the present time, would you say your eyesight, with glasses or contact lenses if you wear them, is excellent, good, fair, poor, or very poor" which is a scaled response question included in NHANES from 1999-2008. (Q3) "Are you blind or unable to see at all?" was administered by the National Health Interview Survey (NHIS) since 1999. We considered but ultimately did not select other questions, including the NHIS question, "Do you have any trouble seeing, even when wearing glasses or contact lenses?" because the wording was deemed to be too similar to Q1.

We also added 2 additional questions not included in prior surveys: "Have you ever been told by a doctor or other health professional that you..." (Q4) "have visual impairment?" and (Q5) "are blind?" which are similar in structure to self-reported eye disease questions that were also included in the survey. We included 6 questions on vision-related ADL that were fielded in the NHIS from 2016-2018, which all began with, "Even when wearing glasses or contacts lenses, because of your eyesight, how difficult is it for you to..." followed by descriptions of ADL, including reading newsprint and driving during the day. Full details and results for the ADL questions are included in Multimedia Appendix 1.

Analyses

We calculated descriptive statistics for all survey responses and chart abstraction values. To assess diagnostic accuracy of questions at the person level, we calculated the sensitivity, specificity, and area under the receiver operator curve (AUC), comparing survey responses to a gold standard of chart-reviewed BCVA. We did not report positive and negative predictive values because these varied based on the underlying prevalence of vision loss in the sample, and by design, we oversampled for vision loss. For Q1 and Q2, we calculated box-and-whisker plots including the mean, median, and IQR of *logMar* values for each survey response value. To assess the validity of self-reported survey questions for quantifying population-level prevalence rates of BCVA-assessed vision loss and blindness, we compared the sample-level prevalence rate and 95% CI, overall and by subgroup. To test for response bias in the telephone survey, we calculated Kruskal-Wallis or chi-square *P* values for differences in baseline demographic characteristics for respondents compared to nonrespondents (Multimedia Appendix 2).

Ethics Approval

The study protocol was reviewed and approved by the UW Institutional Review Board (STUDY00008957) and conforms to the Declaration of Helsinki. All patients included in the study provided informed consent for inclusion in the study, including consenting to the telephone interview and secondary analysis of their medical records and claims information. Consent was obtained via telephone due to COVID-19 restrictions on in-person research activities. No person-level data from this study are released. Participants received no compensation for participation in the study.

Results

Sample Characteristics

The study sample included 438 patients who completed the telephone survey, which was 65.5% of the 669 patients who consented for participation (Table 1). On average, the interview was conducted 5 months after the most recent acuity measurement. Nearly half (n=215, 49.1%) of the 438 patients were interviewed within 120 days of their last acuity measurement and 86.5% (n=379) were interviewed within a year of their last exam, where the longest time difference was 1 year 3 months. We found no significant differences in AUCs associated with duration of time between the survey and exam. We also found no significant differences in demographic characteristics between patients who did and did not respond to the telephone survey (age: *P*=.53; sex: *P*=.44; and race/ethnicity: *P*=.12; Multimedia Appendix 2). The sample was 54.1% (n=237) female, 70.5% (n=309) non-Hispanic White, 9.6% (n=42) non-Hispanic Black, 4.8% (n=21) Hispanic, and 11.6% (n=51) other races/ethnicities. The majority (n=251, 57.3%) of patients were aged 65-84 years; 25.8% (n=113) were aged 40-64 years, 8% were aged (n=35) 85+ years, 8% were aged (n=35) 18-39 years, and 0.9% were aged (n=4) 0-17 years. Of all the respondents, 38.4% (n=168) reported education beyond a bachelor’s degree, 28.8% (n=126) had an associate’s or bachelor’s degree, and 32.4% (n=142) had less than a college degree. Nearly two-thirds (n=281, 64.2%) of patients had BCVA better than 20/40, equating to normal vision; 17.8% (n=78) had mild impairment of 20/40 to <20/80; 4.8% (n=21) had moderate impairment of 20/80 to <20/200; and 13.2% (n=58) were blind, including 2.1% (n=9) with 20/200 to <20/400 and 11.2% (n=49) with blindness defined as acuity ≤20/400, difficulty counting fingers or seeing hand motions, or having no light perception.

Table 1. Demographic characteristics of the study population.

Characteristic	Patient (n=438), n (%)
Sex	
Male	199 (45.4)
Female	237 (54.1)
Age (years)	
0-17	4 (0.9)
18-39	35 (8)
40-64	113 (25.8)
65-84	251 (57.3)
85+	35 (8)
Race/ethnicity	
Hispanic	21 (4.8)
Non-Hispanic Black	42 (9.6)
Non-Hispanic White	309 (70.5)
Other	51 (11.6)
Unknown	15 (3.4)
Education	
Less than college degree	142 (32.4)
Associate's or bachelor's degree	126 (28.8)
Graduate degree	168 (38.4)
Unknown	2 (0.5)
Best-corrected visual acuity in the better-seeing eye	
>20/40 (normal vision)	281 (64.2)
20/40 to <20/80 (mild vision impairment)	78 (17.8)
20/80 to <20/200 (moderate vision impairment)	21 (4.8)
20/200 to <20/400 (blindness)	9 (2.1)
≤20/400 (blindness based on acuity, CF[a], HM[b], or NLP[c])	49 (11.2)

[a] CF: counting fingers.

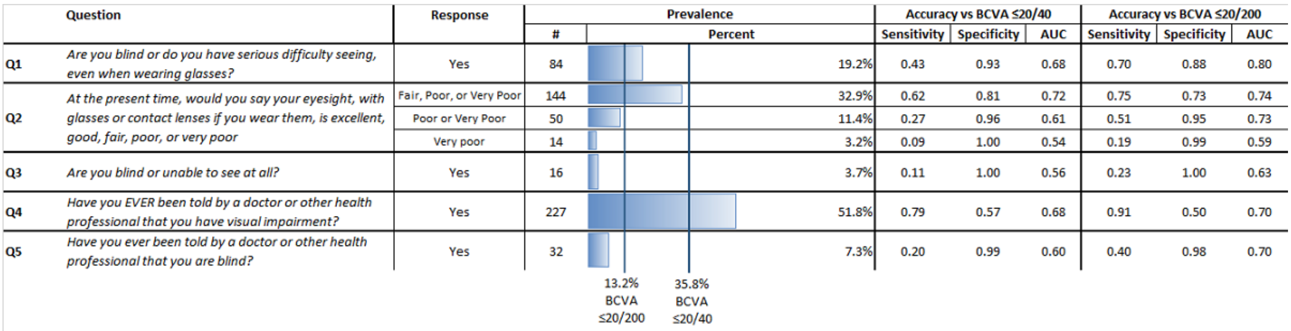
[b] HM: hand motion.

[c] NLP: no light perception.

Survey Responses

Of respondents, 19.2% (n=84) self-reported "yes" to the question Q1, "Are you blind or do you have serious difficulty seeing, even when wearing glasses?"; 3.7% (n=16) responded that they are "blind or unable to see at all" (Q3); and 55.5% (n=227) and 7.3% (n=32) of respondents stated they had ever been told by a doctor that they had visual impairment (Q4) and blindness (Q5), respectively (Figure 1).

Figure 1. Sample prevalence, sensitivity, specificity, and area under the receiver operator curve (AUC) of survey questions versus any vision loss or blindness. The sample prevalence of survey response options compared to the gold standard of best-corrected visual acuity (BCVA) thresholds of ≤20/40 (vision loss) and ≤20/200 (blindness) is shown. The diagnostic accuracy of each question is compared to BCVA-assessed vision loss or blindness. Sensitivity (sens.) refers to the rate at which the question accurately identifies vision loss or blindness in comparison to BCVA (true positive), whereas specificity (spec.) refers to the rate at which the question correctly identified BCVA vision, indicating no vision loss or blindness (true negative). AUC measures both sensitivity and specificity and, therefore, is the primary measure of diagnostic accuracy reported in this analysis. Q: question.



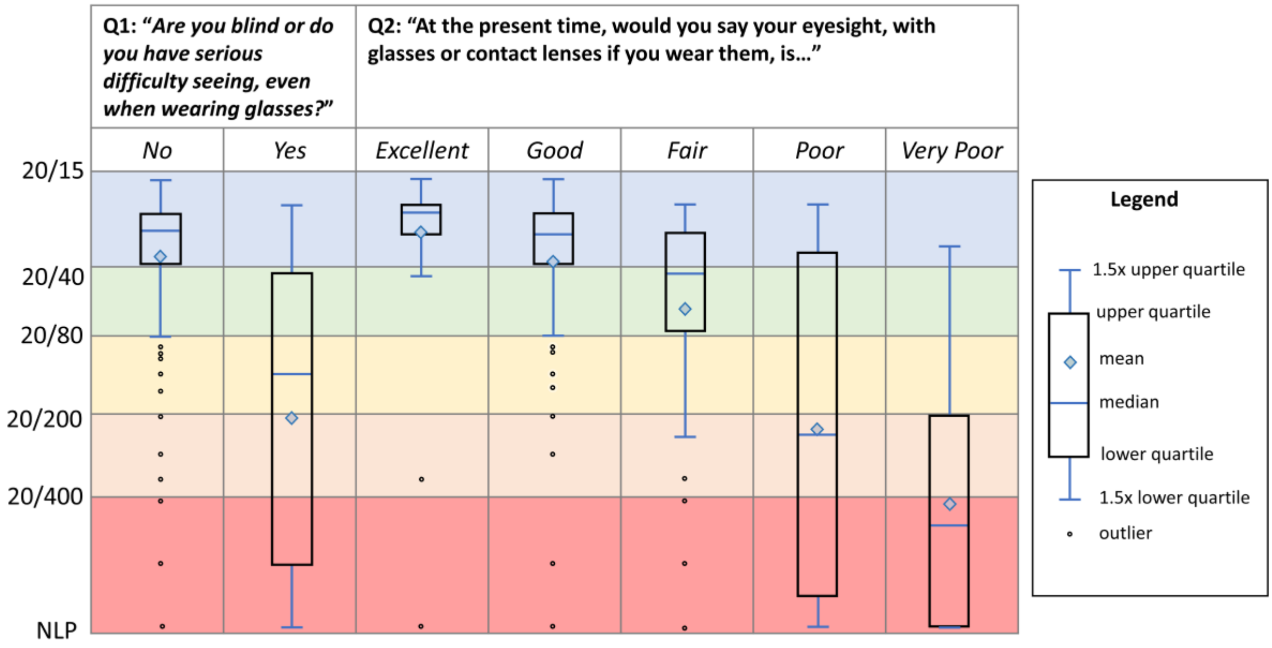
	Question	Response	Prevalence		Accuracy vs BCVA ≤20/40			Accuracy vs BCVA ≤20/200		
			#	Percent	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Q1	*Are you blind or do you have serious difficulty seeing, even when wearing glasses?*	Yes	84	19.2%	0.43	0.93	0.68	0.70	0.88	0.80
Q2	*At the present time, would you say your eyesight, with glasses or contact lenses if you wear them, is excellent, good, fair, poor, or very poor*	Fair, Poor, or Very Poor	144	32.9%	0.62	0.81	0.72	0.75	0.73	0.74
		Poor or Very Poor	50	11.4%	0.27	0.96	0.61	0.51	0.95	0.73
		Very poor	14	3.2%	0.09	1.00	0.54	0.19	0.99	0.59
Q3	*Are you blind or unable to see at all?*	Yes	16	3.7%	0.11	1.00	0.56	0.23	1.00	0.63
Q4	*Have you EVER been told by a doctor or other health professional that you have visual impairment?*	Yes	227	51.8%	0.79	0.57	0.68	0.91	0.50	0.70
Q5	*Have you ever been told by a doctor or other health professional that you are blind?*	Yes	32	7.3%	0.20	0.99	0.60	0.40	0.98	0.70

Person-Level Accuracy of Survey Responses for Indicating BCVA

Figures 1 and 2 show associations between BCVA and survey responses at the person level. Figure 1 reports the sensitivity, specificity, and AUC of each survey question for predicting either vision loss (≤20/40) or blindness (≤20/200). For any vision loss, Q4 had the highest sensitivity (0.79), whereas Q3 had 100% specificity. Q2 achieved the highest AUC (0.72) and the smallest percentage difference when including any response of “fair,” “poor,” or “very poor” to indicate vision loss (data not shown). Q1 had the second highest AUC of 0.68, but this appears to be achieved through high specificity (0.930) despite relatively low sensitivity (0.43). Q1 also underestimated vision loss by 46%. For blindness, the highest sensitivity and specificity were achieved by Q4 (0.91) and Q3 (0.995), respectively. Q1 had the highest AUC at 0.80 but overestimated blindness by 47% relative to BCVA. The second highest AUC was achieved by Q2 (0.74) when including responses of “fair,” “poor,” or “very poor” to indicate blindness. Q2 responses of “poor” or “very poor” resulted in the closest overall prevalence estimate for blindness, underpredicting blindness by 12% relative to BCVA. Multimedia Appendix 3 shows the receiver operating characteristic curves associated with these results.

Figure 2 includes box plots depicting the distribution of visual acuity values by survey responses to Q1 and Q2 (selected because these questions achieved the highest AUC values for blindness and vision loss, respectively). For responses to Q1, the IQR of “no” responses is almost entirely within the range of normal acuity values of <20/40, whereas the lower threshold of 1.5 × IQR for “no” is almost entirely within the range of mild visual impairment of 20/40 to < 20/80. The IQR of “yes” responses ranges from just over 20/40 to <20/400, but the upper threshold of 1.5 × IQR is nearly 20/20. For Q2, the box plots show that the mean and median acuity values of the 5-response options fall within 5 corresponding levels of acuity loss, although the IQR and outer thresholds span multiple acuity categories.

Figure 2. Distribution of best-corrected visual acuity (BCVA) in the better-seeing eye by survey response value. The box-and-whisker plots show the distribution of BCVA among patients who gave a response to both Q1 and Q2. These questions were selected because they were the 2 questions with the highest diagnostic accuracy for BCVA blindness and vision loss, respectively. The box-and-whisker plots show the mean, median, upper and lower quartiles, 1.5× quartiles, and outliers defined as values outside of the range of 1.5× the lower and upper quartiles. The distribution of BCVA generally corresponds with the survey response. However, for Q2, “Excellent” and “Good” responses are both primarily within the normal vision range, whereas “Poor” responses exhibits a very wide distribution. NLP: no light perception; Q: question.

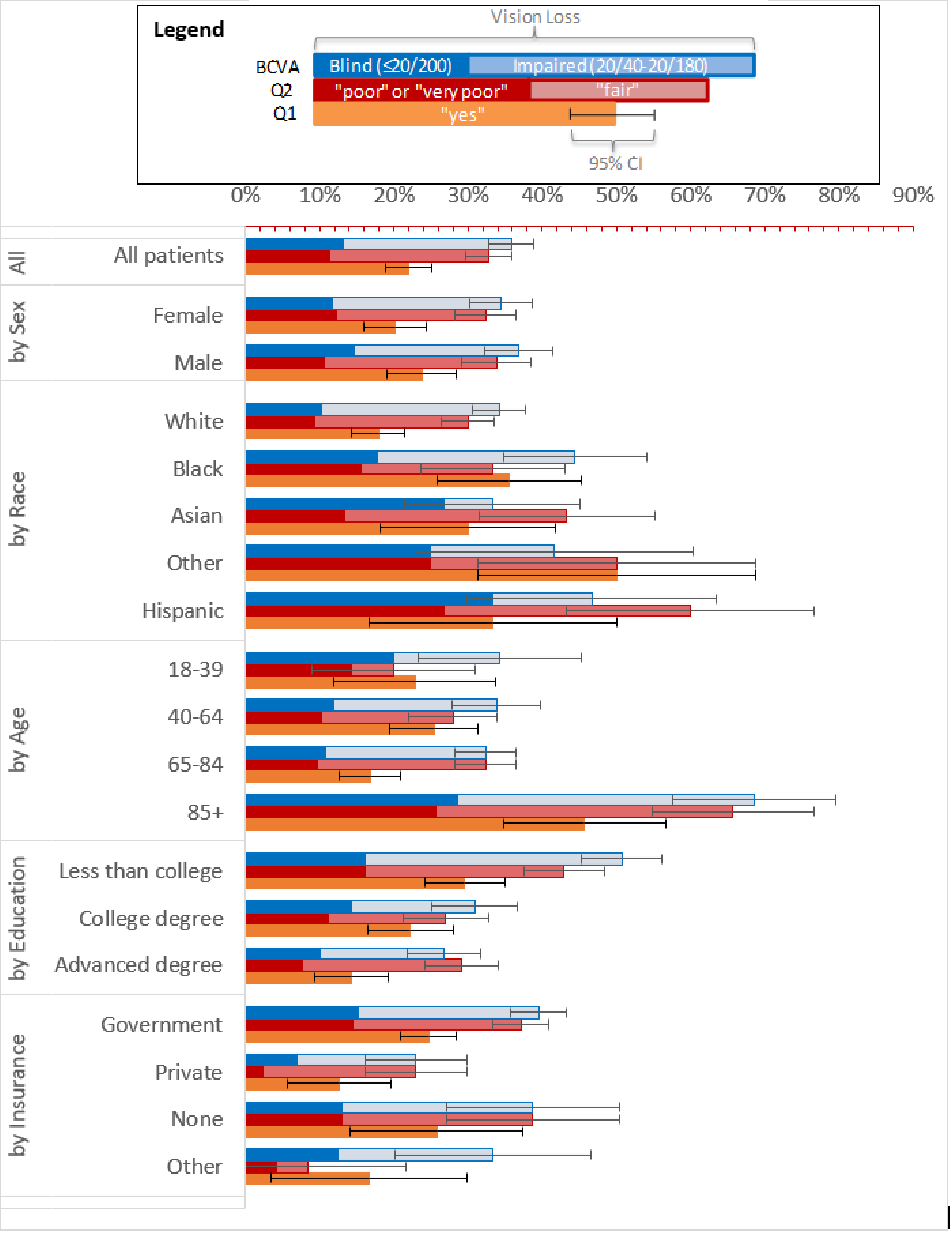


Population-Level Vision Loss and Blindness Prevalence

Figure 3 shows the population-level prevalence rates for BCVA and Q1 and Q2 responses overall and by sex, race/ethnicity, age group, insurer, and education status among the sample. BCVA is shown as a stacked bar with blindness on the left and impairment on the right; the cumulative sum of these bar charts is equivalent to any BCVA loss. The prevalence of vision loss for all patients was 35.8% (157/438). Responses for Q2 are similarly stacked, such that the left bar includes responses of “poor” or “very poor” and the right bar includes “fair”; thus, the stacked cumulative total would equate to the responses of “fair,” “poor,” or “very poor,” which had an overall prevalence of 32.9% (n=144). Q1 “yes” responses had an overall prevalence of 21.9% (96/438).

At the population level, Q1 and Q2 responses exhibit similar trends and disparities in prevalence rates as those of BCVA. The two Q2 response groupings are closely correlated with the 2 BCVA categories. For Q2, the relative difference in prevalence rate of “poor” or “very poor” responses was 14% lower than that of BCVA blindness, whereas the prevalence of Q2 responses of “fair,” “poor,” or “very poor” was 8% lower than BCVA vision loss. Q1 response of “yes” relatively overpredicts BCVA blindness (≤20/200) by 66% and underpredicts BCVA vision loss (≤20/40) by 39%. The relative relationships in prevalence between BCVA, Q1, and Q2 were stable across most sample subgroups. Prevalence based on Q1 falls within the range of BCVA blindness to vision loss for every subgroup except the “Other race” subgroup, although this result is based on a sample of only 12 patients. Similarly, the cumulative responses from Q2 fell within the range of blindness to vision loss for all subgroups except the “Other race,” “Hispanic,” “Advanced degree,” and “Other insurance” subgroups.

Figure 3. Sample-level prevalence rates of best-corrected visual acuity (BCVA)–assessed blindness and visual impairment, Q2 responses of “poor" or "very poor" and “fair,” and Q1 response of “yes.” The relationship between population-level prevalence rates based on BCVA, Q1, and Q2 at the population level is shown. BCVA is based on the better-seeing eye. Q1 is “Are you blind or do you have serious difficulty seeing, even when wearing glasses?” and Q2 is ”At the present time, would you say your eyesight, with glasses or contact lenses if you wear them, is excellent, good, fair, poor, or very poor?” For each population group, there are 3 bars. The top blue bar is BCVA and depicts 2 stacked bars. The darker blue bar on the left is the prevalence of blindness (≤20/200). The lighter blue bar stacked to the right is the prevalence of persons with visual impairment (20/40 to 20/180). Thus, the sum of the upper, blue stacked bars represents any vision loss (≤20/40). The red stacked bar below depicts prevalence results from Q2. The dark red bar on the left includes patients who reported “poor” or “very poor” eyesight. The lighter red bar stacked to the right is the prevalence of “fair” responses. Thus, the value of the dark and light red stacked bars represents patients who reported “very poor,” “poor,” or “fair” eyesight. The lower orange bar is the prevalence of “yes” responses to Q1. Q: question.



Discussion

This study provides evidence on the accuracy and validity of vision and eye health survey questions, including those currently or recently deployed through federally funded national surveys. We measured the concordance between self-reported vision loss or blindness and vision-related ADL limitations versus vision measured with BCVA among a sample of patients from ophthalmology and optometry clinics and reported results both at the person and population level. The highest AUC for predicting blindness (0.80) was achieved by Q1, “Are you blind, or do you have serious difficulty seeing, even when wearing glasses?” The highest AUC (0.72) for predicting any BCVA loss was achieved by Q2, “At the present time, would you say your eyesight, with glasses or contact lenses if you wear them, is…” when the responses of “fair,” “poor,” or “very poor” are used to indicate any vision loss. Prevalence rates based on Q2 or adjusted Q1 responses were highly correlated to population-level prevalence of BCVA blindness and vision loss and suggest that at the population level, variation in self-reported prevalence based on Q1 and Q2 is likely to accurately reflect patterns and variations in the underlying prevalence of visual acuity loss and blindness.

This study has several limitations. Our study sample consisted of current or recent patients from optometry and ophthalmology clinics, which are not representative of the general population and may have higher awareness of their visual health. This sample also exhibited higher levels of education and, due to our intentional oversampling, a much higher prevalence of evaluated BCVA impairment than would be seen in a nationally representative population. Due to COVID-19–related restrictions on in-person research, patient acuity measures were collected through retrospective chart review. Among the 669 recruited patients, 231 (34.5%) did not complete a telephone interview. In addition, we define vision loss based on BCVA only. We did not consider other measures of vision loss such as reduced visual field, contrast sensitivity, or near-distance acuity because this information was not routinely recorded in the medical charts. Due to this limitation, we likely misclassified some patients with better vision. We additionally compared self-reported values to BCVA as opposed to presenting visual acuity because of data quality concerns regarding how consistently presenting or habitual acuity may have been recorded in patient charts and because historically, vision loss and blindness prevalence estimates are reported based on BCVA. Finally, time differences between the patients’ survey interview and the date of their nearest examination may bias results. Presumably, longer time periods would reduce the apparent accuracy of the self-reported responses, although we did not find a significant association between this time difference and AUC.

Identifying and documenting the concordance and relationship between self-reported vision problems and BCVA is important to enhance national surveillance of vision loss and eye problems. Historically, visual health epidemiological estimates were derived from either national surveys or examination studies, each of which have inherent strengths and limitations. Examination studies provide estimates of objectively measured vision loss, but due to their high cost and complexity, they are all based on small, localized samples, many of which are arguably out of date. National self-reported surveys provide large, representative, and ongoing samples that examination studies lack, but self-reported instruments previously lacked evidence of validity as proxy measures of actual vision loss.

The purpose of this study was to evaluate the validity of self-reported vision measures to assess population health. The results of this study imply that self-reported questions should not be used as a substitute for clinical vision evaluation but do support the interpretation of self-reported survey questions as proxy measures of variation in population visual health for use in surveillance. The recent surveillance estimates of the prevalence of visual acuity loss produced by the CDC’s VEHSS and the WHO both leverage the strengths of examination studies (their robust measurement of vision loss) as well as that of national surveys (their large, ongoing representative samples) to provide more complete and detailed epidemiological estimates than can be supported by either type of data alone [2,17]. Our findings lend support to the methodologies and assumptions of these ongoing national and international visual health surveillance efforts by measuring the accuracy of self-reported vision at the individual and population level.

Acknowledgments

The authors would like to thank the Vision and Eye Health Surveillance System (VEHSS) expert panel for sharing their expertise and reviewing the study protocol and abstraction forms. Their input greatly improved the quality of this study: Bonnielin Swenor, PhD, MPH, Johns Hopkins Bloomberg School of Public Health; Cynthia Owsley, PhD, MSPH, University of Alabama at Birmingham; David Friedman, MD, PhD, MPH, Harvard Medical School, Massachusetts Eye and Ear; Sandra Block, OD, MEd, MPH, Illinois College of Optometry; and Yao Liu, MD, MS, University of Wisconsin.

This manuscript was supported by funding from the Centers for Disease Control and Prevention (CDC) Vision Health Initiative via cooperative agreement U01DP006444, “Research to Enhance the U.S. Vision and Eye Health Surveillance System.” The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC.

Data Availability

Data collected for this study are restricted per the institution review board approval at the University of Washington and are not publicly available.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Activities of daily living (ADL) questions.
[DOCX File , 22 KB - publichealth_v9i1e44552_app1.docx]

Multimedia Appendix 2
Baseline characteristics for all consented patients.
[DOCX File , 17 KB - publichealth_v9i1e44552_app2.docx]

Multimedia Appendix 3
Receiver operator curves (ROC).
[DOCX File , 142 KB - publichealth_v9i1e44552_app3.docx]

References

1. Rein DB, Wittenborn JS, Zhang P, Sublett F, Lamuda PA, Lundeen EA, et al. The economic burden of vision loss and blindness in the United States. Ophthalmology 2022 Apr;129(4):369-378. [doi: 10.1016/j.ophtha.2021.09.010] [Medline: 34560128]
2. Flaxman AD, Wittenborn JS, Robalik T, Gulia R, Gerzoff RB, Lundeen EA, Vision and Eye Health Surveillance System study group. Prevalence of visual acuity loss or blindness in the US: a Bayesian meta-analysis. JAMA Ophthalmol 2021 Jul 01;139(7):717-723 [FREE Full text] [doi: 10.1001/jamaophthalmol.2021.0527] [Medline: 33983373]
3. National Academies of Sciences, Engineering, and Medicine, Medicine Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on Public Health Approaches to Reduce Vision Impairment and Promote Eye Health. In: Welp A, Woodbury RB, McCoy MA, Teutsch SM, editors. Making Eye Health a Population Health Imperative: Vision for Tomorrow. Washington, DC: National Academies Press (US); 2016.
4. Zhang X, Cotch MF, Ryskulova A, Primo SA, Nair P, Chou C, et al. Vision health disparities in the United States by race/ethnicity, education, and economic status: findings from two nationally representative surveys. Am J Ophthalmol 2012 Dec;154(6 Suppl):S53-62.e1 [FREE Full text] [doi: 10.1016/j.ajo.2011.08.045] [Medline: 23158224]
5. Congdon N, O'Colmain B, Klaver CCW, Klein R, Muñoz B, Friedman DS, Eye Diseases Prevalence Research Group. Causes and prevalence of visual impairment among adults in the United States. Arch Ophthalmol 2004 Apr 01;122(4):477-485. [doi: 10.1001/archopht.122.4.477] [Medline: 15078664]
6. Varma R, Kim JS, Burkemper BS, Wen G, Torres M, Hsu C, Chinese American Eye Study Group. Prevalence and causes of visual impairment and blindness in Chinese American sdults: the Chinese American Eye Study. JAMA Ophthalmol 2016 Jul 01;134(7):785-793. [doi: 10.1001/jamaophthalmol.2016.1261] [Medline: 27196952]
7. Varma R, Ying-Lai M, Klein R, Azen SP, Los Angeles Latino Eye Study Group. Prevalence and risk indicators of visual impairment and blindness in Latinos: the Los Angeles Latino Eye Study. Ophthalmology 2004 Jun;111(6):1132-1140. [doi: 10.1016/j.ophtha.2004.02.002] [Medline: 15177963]
8. Vitale S, Cotch MF, Sperduto RD. Prevalence of visual impairment in the United States. JAMA 2006 May 10;295(18):2158-2163. [doi: 10.1001/jama.295.18.2158] [Medline: 16684986]
9. Muñoz B, West SK, Rodriguez J, Sanchez R, Broman AT, Snyder R, et al. Blindness, visual impairment and the problem of uncorrected refractive error in a Mexican-American population: Proyecto VER. Invest Ophthalmol Vis Sci 2002 Mar;43(3):608-614. [Medline: 11867574]
10. Fisher DE, Shrager S, Shea SJ, Burke GL, Klein R, Wong TY, et al. Visual impairment in White, Chinese, Black, and Hispanic participants from the Multi-Ethnic Study of Atherosclerosis cohort. Ophthalmic Epidemiol 2015 Sep 23;22(5):321-332 [FREE Full text] [doi: 10.3109/09286586.2015.1066395] [Medline: 26395659]
11. Mansberger SL, Romero FC, Smith NH, Johnson CA, Cioffi GA, Edmunds B, et al. Causes of visual impairment and common eye problems in Northwest American Indians and Alaska Natives. Am J Public Health 2005 May;95(5):881-886. [doi: 10.2105/ajph.2004.054221]
12. Zebardast N, Friedman DS, Vitale S. The prevalence and demographic associations of presenting near-vision impairment among adults living in the United States. Am J Ophthalmol 2017 Feb;174:134-144 [FREE Full text] [doi: 10.1016/j.ajo.2016.11.004] [Medline: 27865728]
13. Crews JE, Lollar DJ, Kemper AR, Lee LM, Owsley C, Zhang X, et al. The variability of vision loss assessment in federally sponsored surveys: seeking conceptual clarity and comparability. Am J Ophthalmol 2012 Dec;154(6 Suppl):S31-44.e1. [doi: 10.1016/j.ajo.2011.10.029] [Medline: 23158222]
14. Davidson C, Wittenborn J, Rein D. NORC at the University of Chicago. 2018 May 1. URL: https://www.norc.org/PDFs/VEHSS/ReviewNationalSurveyDataVEHSS.pdf [accessed 2023-02-23]

15. Rein DB, Lamuda PA, Wittenborn JS, Okeke N, Davidson CE, Swenor BK, et al. Vision impairment and blindness prevalence in the United States: variability of vision health responses across multiple national surveys. Ophthalmology 2021 Jan;128(1):15-27 [FREE Full text] [doi: 10.1016/j.ophtha.2020.06.064] [Medline: 32663529]
16. Sommer A. Prevalence of visual impairment in the United States: ease versus accuracy. Ophthalmology 2021 Jan;128(1):28-29. [doi: 10.1016/j.ophtha.2020.08.036] [Medline: 33349342]
17. GBD 2019 Blindness and Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. Lancet Glob Health 2021 Feb;9(2):e130-e143 [FREE Full text] [doi: 10.1016/S2214-109X(20)30425-3] [Medline: 33275950]

Abbreviations

ADL: activities of daily living
AUC: area under the receiver operator curve
BCVA: best-corrected visual acuity
CDC: Centers for Disease Control and Prevention
NHANES: The National Health and Nutrition Examination Survey
NHIS: National Health Interview Survey
Q: question
UW: University of Washington
VEHSS: Vision and Eye Health Surveillance System
WHO: World Health Organization

Edited by A Mavragani; submitted 23.11.22; peer-reviewed by D VanNasdale, J Crews; comments to author 07.12.22; revised version received 13.12.22; accepted 06.01.23; published 07.03.23.

Please cite as:
Wittenborn J, Lee A, Lundeen EA, Lamuda P, Saaddine J, Su GL, Lu R, Damani A, Zawadzki JS, Froines CP, Shen JZ, Kung TPH, Yanagihara RT, Maring M, Takahashi MM, Blazes M, Rein DB
Comparing Telephone Survey Responses to Best-Corrected Visual Acuity to Estimate the Accuracy of Identifying Vision Loss: Validation Study
JMIR Public Health Surveill 2023;9:e44552
URL: https://publichealth.jmir.org/2023/1/e44552
doi:10.2196/44552
PMID:36881468

Original Paper

Defining County-Level Terrestrial Rabies Freedom Using the US National Rabies Surveillance System: Surveillance Data Analysis

Amber Kunkel[1,2], ScD; Gabriella Veytsel[1], MPH; Sarah Bonaparte[1], MPH; Haillie Meek[3], DVM; Xiaoyue Ma[1], MPH, MEd; Amy J Davis[4], PhD; Jesse Bonwitt[1], BVSc; Ryan M Wallace[1], DVM

[1]Division of High-Consequence Pathogens and Pathology, National Center for Emerging & Zoonotic Infectious Diseases, US Centers for Disease Control and Prevention, Atlanta, GA, United States

[2]Epidemic Intelligence Service, US Centers for Disease Control and Prevention, Atlanta, GA, United States

[3]Epidemiology Elective Program, US Centers for Disease Control and Prevention, Atlanta, GA, United States

[4]Wildlife Services, Animal and Plant Health Inspection Service, United States Department of Agriculture, Fort Collins, CO, United States

Corresponding Author:
Sarah Bonaparte, MPH
Division of High-Consequence Pathogens and Pathology
National Center for Emerging & Zoonotic Infectious Diseases
US Centers for Disease Control and Prevention
1600 Clifton Road
Atlanta, GA, 30329
United States
Phone: 1 4046393538
Email: sarahcbonaparte@gmail.com

Abstract

Background: Rabies is a deadly zoonotic disease with nearly 100% fatality rate. In the United States, rabies virus persists in wildlife reservoirs, with occasional spillover into humans and domestic animals. The distribution of reservoir hosts in US counties plays an important role in public health decision-making, including the recommendation of lifesaving postexposure prophylaxis upon suspected rabies exposures. Furthermore, in surveillance data, it is difficult to discern whether counties have no cases reported because rabies was not present or because counties have an unreported rabies presence. These epizootics are monitored by the National Rabies Surveillance System (NRSS), to which approximately 130 state public health, agriculture, and academic laboratories report animal rabies testing statistics. Historically, the NRSS classifies US counties as free from terrestrial rabies if, over the previous 5 years, they and any adjacent counties did not report any rabies cases and they tested ≥15 reservoir animals or 30 domestic animals.

Objective: This study aimed to describe and evaluate the historical NRSS rabies-free county definition, review possibilities for improving this definition, and develop a model to achieve more precise estimates of the probability of terrestrial rabies freedom and the number of reported county-level terrestrial rabies cases.

Methods: Data submitted to the NRSS by state and territorial public health departments and the US Department of Agriculture Wildlife Services were analyzed to evaluate the historical rabies-free definition. A zero-inflated negative binomial model created county-level predictions of the probability of rabies freedom and the expected number of rabies cases reported. Data analyzed were from all animals submitted for laboratory diagnosis of rabies in the United States from 1995 to 2020 in skunk and raccoon reservoir territories, excluding bats and bat variants.

Results: We analyzed data from 14,642 and 30,120 county-years in the raccoon and skunk reservoir territories, respectively. Only 0.85% (9/1065) raccoon county-years and 0.79% (27/3411) skunk county-years that met the historical rabies-free criteria reported a case in the following year (99.2% negative predictive value for each), of which 2 were attributed to unreported bat variants. County-level model predictions displayed excellent discrimination for detecting zero cases and good estimates of reported cases in the following year. Counties classified as rabies free rarely (36/4476, 0.8%) detected cases in the following year.

Conclusions: This study concludes that the historical rabies freedom definition is a reasonable approach for identifying counties that are truly free from terrestrial raccoon and skunk rabies virus transmission. Gradations of risk can be measured using the rabies prediction model presented in this study. However, even counties with a high probability of rabies freedom should maintain rabies testing capacity, as there are numerous examples of translocations of rabies-infected animals that can cause major changes in the epidemiology of rabies.

(JMIR Public Health Surveill 2023;9:e43061) doi:10.2196/43061

KEYWORDS

rabies; surveillance; zoonoses; zoonosis; model; disease spread; infection spread; animal; predict; public health

Introduction

Background

Rabies is a deadly zoonotic disease caused by an RNA virus (rabies virus) in the *Lyssavirus* genus, which causes acute progressive encephalitis in mammals with nearly 100% fatality rate [1,2]. Globally, an estimated 60,000 people die annually from rabies infection, with the highest fatality rates attributable to the dog-mediated rabies virus variant (DMRVV) in African and Asian countries [3]. However, established dog population management and vaccination methods have successfully eliminated DMRVV in most Western countries [4].

With the elimination of DMRVV in the United States, terrestrial mesocarnivores and bats are the remaining rabies reservoirs in the United States [5]. Reported infections in domestic animals have decreased since the implementation of animal control and vaccination programs in the 1940s and 1950s in the United States; however, contact with wildlife and unvaccinated domestic animals can still pose a threat of infection [6,7]. Human rabies deaths in the United States are rare, and rabies pre-exposure prophylaxis and postexposure prophylaxis (PEP) are effective when administered according to the Advisory Committee on Immunization Practices guidance [6,8,9]. Annually, an estimated 55,000 to 60,000 people in the United States receive rabies PEP for suspected rabies exposures at an estimated cost of >US $150 million [6,10]. The distribution of reservoir hosts in a given US county (ie, local administrative unit) can inform rabies risk assessment algorithms to determine if rabies PEP is required following a human exposure, based on the epizootiology of the area where the exposure occurred [11]. Therefore, understanding the geographic and temporal patterns of animal rabies in the United States through routine surveillance is essential for public health interventions such as the recommendation of costly but lifesaving rabies PEP as well as for developing and evaluating wildlife rabies management actions and responding to unexpected rabies occurrences [7].

Rabies reservoirs in the United States include multiple bat species and 5 terrestrial mesocarnivores, including raccoons (*Procyonlotor*), skunks (family *Mephitidae*), foxes (*Vulpes* spp and *Urocyoncinereoargenteus*), and the small Indian mongoose (*Herpestesauropunctatus*) in Puerto Rico [7]. Substantial geographical variation exists in the risk of rabies exposure in the United States owing to the distinct geographic distribution of the different rabies virus variants (RVVs) associated with these terrestrial reservoirs [12]. The eastern raccoon RVV, for example, was first reported in Florida in the 1950s and later spread throughout the east coast following an apparent translocation event [13,14]. Three distinct skunk RVVs exist: south central, north central, and California skunks, named for their geographical locations within the United States. Other regions of the United States, such as the northwestern states of Washington and Oregon, are believed to be free from terrestrial rabies reservoirs. However, bat rabies is present throughout the continental US as it is uniquely unconstrained by the geographic barriers that define the territories of terrestrial reservoirs [5,7,12].

Animal and human rabies are both nationally notifiable conditions in the United States [15,16]. The US Centers for Disease Control and Prevention (CDC) manages the National Rabies Surveillance System (NRSS), which comprises approximately 130 jurisdictional public health, agriculture, and academic laboratories that conduct passive public health surveillance, as well as the US Department of Agriculture National Wildlife Rabies Management Program, which conducts active case surveillance in areas where wildlife vaccinations are distributed [12]. Jurisdictional laboratories conduct diagnostic testing; some perform viral characterization for positive samples; and the US CDC offers laboratory assistance for diagnostics, viral typing, and exposure tracing and assessment. Animal and human rabies case definitions and data elements reportable to the NRSS are defined by the Council of State and Territorial Epidemiologists Animal Rabies and Human Rabies position statements, respectively [17,18].

Objectives

As the NRSS is largely a passive public health surveillance system, it can be difficult to discern through surveillance data if counties report no animal rabies cases because of the absence of rabies or because of the inability to detect cases. Historically, the NRSS defines counties in the United States as *terrestrial rabies free* if there have been no terrestrial rabies cases reported in that county or any adjacent county for the previous 5 years and if the county has achieved a sufficient level of surveillance testing for rabies over those 5 years [12]. However, no previous attempts have been made to validate or improve this definition since its inception in 2005 [19]. Therefore, the objectives of this analysis were to describe and evaluate the performance of the historical NRSS rabies-free county definition, review the possibilities for improving this definition, and develop a model to achieve more precise estimates of the probability of terrestrial rabies freedom and the number of reported terrestrial rabies cases at the county level.

Methods

Domestic Animal Rabies Testing and Reporting in the United States

The NRSS collects data from state public health, agriculture, and academic laboratories that test animal samples for rabies using the direct fluorescent antibody test or other diagnostic methods recommended by the Council for State and Territorial

Epidemiologist's Animal Rabies position statement [17]. Confirmed animal cases are reportable to the state public health officials and notifiable to the CDC [15,17,20]. Recommended data elements included the species, capture location, test date, and RVV (when available) of each animal tested for rabies virus. The NRSS has been relatively unchanged since 1995 in terms of consistent diagnostic approaches, access to testing, and epidemiological circumstances.

The passive public health surveillance system prioritizes the testing of animals involved in human or domestic animal exposures, although specific testing criteria vary by jurisdiction. Passive public health surveillance accounted for 95% of reported rabid animals. Active surveillance is conducted in selected high-priority areas and informs wildlife rabies management such as management of oral rabies vaccination campaigns. Passive public health data and active surveillance data were both included in this analysis.

Historical Rabies-Free Definition

The historical definition used to classify counties as terrestrial rabies free is as follows [12]:

- No terrestrial rabies cases reported in that county for ≥5 years;
- No terrestrial rabies cases reported in any neighboring counties for ≥5 years; and
- ≥15 surveillance points were tested over the past 5 years, where surveillance points=0.5 × (number of domestic animals) + 1 × (number of reservoir animals tested). Companion animals and livestock (eg, dogs, cats, cattle, and horses) are considered domestic animals. Foxes, raccoons, skunks, and mongooses are considered reservoir animals, regardless of the state or county in which they are found.

Data Set

This analysis used county-level data from the NRSS from 1995 to 2020 regarding the number and species of animals tested, and positive, for rabies each year (excluding bats) as well as accompanying viral characterization results when available. Bats and bat variants were excluded to focus the analysis on rabies from terrestrial reservoirs; the entire continental United States is considered to be at risk for rabies transmitted by bats. First, we evaluated the historical surveillance definition using measures of sensitivity, specificity, positive predictive value, and negative predictive value. Next, we applied a zero-inflated negative binomial model to describe the contributions of each individual variable involved in the historical definition, evaluate alternative definitions of rabies freedom, and predict the presence or absence and the number of rabies cases that would be reported in a given county in the following year.

All bats submitted for testing, as well as rabies virus–positive terrestrial mammals recorded as having bat RVVs, were excluded from the analysis. As data files for some states before 2014 included inconsistent reporting of negative test results, we included data from each state starting from the year after which <3 in 4 consecutive years contained incomplete rabies test results (1995 for most states; 1998 for Delaware; 2002 for Florida, Mississippi, New Mexico, and Vermont; 2006 for California, Georgia, Iowa, and South Carolina; and 2014 for Oklahoma). Eight remaining state-years had missing county information for >50% of negative results. In these data files, to avoid loss of these data in the analysis, the negative results were assigned to counties using the average distribution from the 4 closest years of complete data. Nineteen state-years had ≤10 samples tested or <50% of samples testing negative, suggesting that negative results had not been fully reported. In these data files, we set the number of negatives equal to the average of the 4 closest years of the complete data. Overall, 6.99% (104/1486) of the state-years of data were removed for missing negatives and 2% (27/1350) had negative results imputed.

The total number and species of animals tested and positive were aggregated by county and year. All terrestrial species were included unless they were noted as having bat RVVs. The historical rabies freedom definition requires data on the detection of rabies in the preceding 5 years. We used the first 8 continuous years of data from each county (regardless of whether any samples were submitted) as a baseline and began predicting the presence or absence of rabies from the 9th year of the data. Thus, our predictions began in 2003 for counties in the states with complete data.

Models were fitted separately for counties in states that were historically designated as raccoon and skunk reservoir territory based on past detections of the respective RVVs, as reported by the NRSS [12]. Counties in states with other terrestrial rabies reservoirs (Arctic fox variant, gray fox variant, and dog-mongoose variant) were excluded to focus on the most common RVVs found in the United States. For the raccoon territory, we first selected all counties that had at least 1 recorded positive case in the data set and were located in states where raccoons were the primary terrestrial reservoir. We then added all counties within 100 km of this area (county centroid-to-centroid) unless the new county was located in a state with a different terrestrial reservoir. We followed the same procedure to define the counties as skunk territories.

For modeling analyses, data were separated into "training" and "validation" data sets. Data from 2018 were used to train the models, whereas data from 2019 and 2020 were used as separate validation data sets. The results of nonparametric analyses are presented for both the training and validation data sets.

Modeling Approach

We used a zero-inflated negative binomial regression model to describe the presence or absence of terrestrial rabies and the number of reported terrestrial rabies cases per county-year. A zero-inflated negative binomial model is appropriate for count data when there is a greater than expected number of zeros in the data set, and the process generating these zeros is independent of the count process [21,22]. In this case, excess zeros may occur in counties that are truly rabies free, that is, not enzootic for terrestrial RVVs. The counting process represents rabid animal observations in counties where rabies is enzootic. The first process is encoded as a logistic model for rabies freedom (1a) and the second as a negative binomial count model (1b). We refer to the zero-inflated negative binomial model as the rabies prediction model.

Freedom model 1a: $\text{logit}(P(\textit{rabies free})) = \alpha_0 + \alpha_1 A + \alpha_2 B + \alpha_3 C + \alpha_4 D$ **(1)**

Count model 1b: $\log(E(\textit{cases reported}|\textit{not rabies free})) = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 E$ **(2)**

Where:

A = log(number of rabid animals reported in the county of interest in the past year + 0.5) **(3)**

B = log(number of rabid animals reported in any neighboring county in the past year + 0.5) **(4)**

C = log(years with no rabies reported in the county of interest) [1=case reported in past year, 2=last case reported in the year before last, etc] **(5)**

D = log(years with no rabies reported in any neighboring county) [1=case reported in the past year, 2=last case reported in the year before last, etc] **(6)**

E = log(surveillance points in the past 5 years + 0.5) **(7)**

The count portion of the model estimates the number of cases that will be reported if rabies is enzootic. The freedom portion of the model estimates the probability that there truly are no cases within the county. From these model outputs, we can derive three outcomes: (1) the probability that the county is rabies free, (2) the probability that the county would observe rabies cases (if present) through their surveillance efforts, and (3) the probability that the county would have cases that would go unreported by surveillance efforts.

Predictor variables for each component model were selected based on the historical rabies-free definition and a consideration of possible causal pathways. For example, surveillance may have a causal effect on the number of rabies cases reported (count model) but not on the underlying endemicity of terrestrial rabies RVVs (freedom model).

We natural log transformed all the variables to provide a better fit. We defined the number of years since rabies was last reported as ranging from 1 (last year) to 9 (cutoff for any >8). Variables for surveillance points and cases reported in the prior year were shifted by 0.5 to avoid zero values in log-transformed variables.

Model Validation and Comparison

To evaluate the impacts of different surveillance effort definitions, we applied different definitions of surveillance effort to the rabies prediction model: including only domestic or only reservoir animals, adjusting the relative value of surveillance effort by changing the coefficient applied to domestic animals from 0.1 to 1 times that of reservoir animals, applying specific coefficients to different domestic and wild animal species based on their reported test positivity rates, changing the number of years of surveillance effort considered from 1 to 8, and normalizing surveillance effort by county size or population.

Several criteria were used to evaluate different model formulations. First, we calculated the Spearman correlation between the predicted and reported cases each year. Second, we computed the area under the receiver operating characteristic curves (AUCs), treating the observed data in the prediction year as "true" values. We considered two different options as predicted values for AUCs: (1) the model-predicted probability of rabies freedom and (2) the model-predicted probability of reporting 0 cases. As the number of observed cases in the prediction year is not a perfect measure of rabies freedom, we also considered the detection of any rabies cases in the following 5 years as a marker of true rabies presence in a sensitivity analysis.

$P(\textit{rabies free}) = \textit{expit}(\alpha_0 + \alpha_1 A + \alpha_2 B + \alpha_3 C + \alpha_4 D)$ **(8)**

$P(\textit{0 cases reported}) = P(\textit{rabies free}) + (1 - P(\textit{rabies free})) \times P(\textit{0 cases reported}|\textit{not rabies free})$ **(10)**

Where:

r=the negative binomial dispersion parameter

In surveillance data, it is difficult to discern which counties had zero cases reported because rabies was not present versus which counties had unreported rabies. To approximate how well the model captures observing at least 1 case when a county is enzootic, we calculated the AUC by comparing the predicted probability of rabies observing at least 1 positive case if rabies is present in a county (ie, the probability of >0 cases reported from model 1b) with county-level bat observation data in a sensitivity analysis using bat variant data. Rabid bats are considered to be present everywhere in the continental United States, so we hypothesized that counties with a high probability of observing terrestrial rabies would also be more likely to observe rabid bats. However, counties not observing rabid bats would still not be considered "rabies free," hence their exclusion from model fitting and the main analysis.

Calculations

All analyses were performed using R [23]. The zero-inflated models were run using the package *pscl* [24].

Ethical Considerations

These data were reported to the CDC by the State Public Health Departments under the authority of the Council for State and Territorial Epidemiologists, Position Statement 22-ID-06. As this study is a secondary analysis of data, institutional review board approval was not needed.

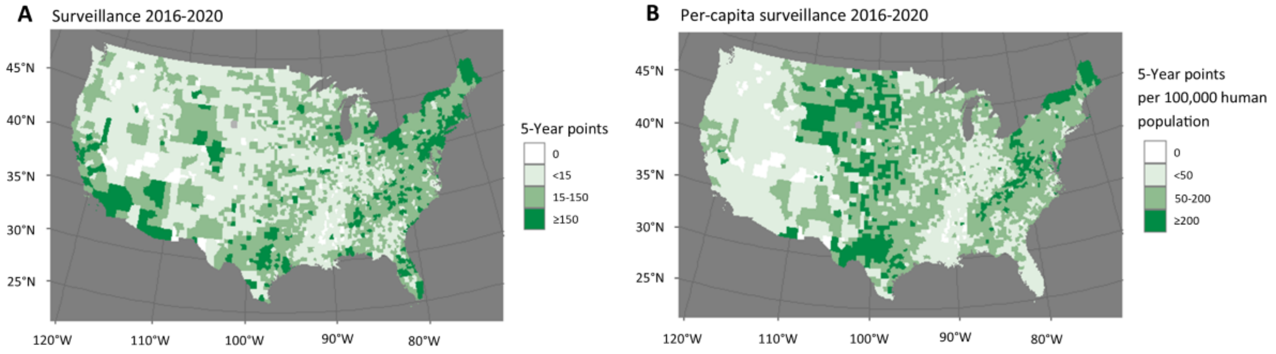
Results

Overview

Our analyses for raccoon territory included data from 1005 counties across 23 states beginning in 2003 for a total of 14,642 county-years. For skunk territory, we included 1799 counties across 25 states beginning in 2003 for a total of 30,120 county-years. Counties from some states were included after 2003, as described in the *Methods* section.

Of the county-years included for raccoon territory analysis, 57% (8346/14,642) had reported a terrestrial rabies case in the past year, 2% (293/14,642) had not reported a case for exactly 5 years, and 23% (3368/14,642) had not reported a case for >8 years. In skunk territory, only 17% (5120/30,120) of county-years had reported a case in the past year, 3% (904/30,120) had not reported a case for exactly 5 years, and 54% (16,265/30,120) had not reported a case for >8 years. The median number of surveillance points, under the existing NRSS definition, over the preceding 5-year period was 63.0 (IQR 22.0-143.9) in raccoon territory counties and 14.5 (IQR 5.5-39.0) in skunk territory counties. This corresponds to a median of 98.2 (IQR 55.7-153.6) surveillance points per 100,000 people in raccoon territory and 72.9 (IQR 36.2-147.2) per 100,000 people in skunk territory. The overall and per capita surveillance points from 2016 to 2020 are shown in Figure 1.

Figure 1. Rabies surveillance points (A) and per-capita rabies surveillance points (B) tested in the contiguous United States during 2016 to 2020 at the county level. Each reservoir animal tested contributes 1 surveillance point, and each domestic animal contributes 0.5 surveillance points.



Nonparametric Evaluation of Historical Definition

Only 0.85% (9/1065) of counties in raccoon territory and 0.79% (27/3411) of counties in skunk territory had rabies reported in a year for which rabies was predicted to be absent using the historical rabies-free definition from 2003 to 2020 (Table 1). According to the historical definition, counties predicted as rabies free for a given year had no cases reported 99.2% of the time in both raccoon and skunk territories (negative predictive value). Counties predicted to be rabies free for a given year also had a high probability of no cases being reported throughout the following 5 years (97.8% in raccoon territory and 94.9% in skunk territory). Only a small proportion of counties with zero rabies cases reported in a given year met the historical rabies-free criteria (16.6% in raccoon territory and 13.5% in skunk territory).

We performed a deeper investigation into the 9 counties in raccoon territory where rabies was reported in a year for which rabies was predicted to be absent. Of these 9 detections, 2 (22%) were attributed to bat variants that had not been reported to the CDC and 1 (11%) was associated with a data entry error in the state submission file. Of the remainder, 3 were associated with a rabies incursion into northeastern Ohio in 2004 and 1 with a translocated rabid cat [25]. No explanation was found for the remaining 2 counties. Therefore, only 56% (5/9) of detections in raccoon territory were falsely predicted to be rabies free.

Table 1. Nonparametric evaluation of historical rabies freedom definition in counties in both skunk and raccoon territories.

Reservoir and county-level rabies prediction (historical rabies-free definition)	Rabies reported in county in the next 1 year			Rabies reported in county in the next 5 years		
	Yes	No	Total	Yes	No	Total
Raccoon						
Present, n	8264	5313	13,577	7756	2090	9846
Free, n	9	1056	1065	17	759	776
P(predicted present|reported)	99.9%	99.9%	99.9%	99.8%	99.8%	99.8%
P(predicted free|not reported)	16.6%	16.6%	16.6%	26.6%	26.6%	26.6%
P(reported|predicted present)	60.9%	60.9%	60.9%	78.8%	78.8%	78.8%
P(not reported|predicted free)	99.2%	99.2%	99.2%	97.8%	97.8%	97.8%
Skunk						
Present, n	4944	21,765	26,709	8298	12,074	20,372
Free, n	27	3384	3411	131	2421	2552
P(predicted present|reported)	99.5%	99.5%	99.5%	98.4%	98.4%	98.4%
P(predicted free|not reported)	13.5%	13.5%	13.5%	16.7%	16.7%	16.7%
P(reported|predicted present)	18.5%	18.5%	18.5%	40.7%	40.7%	40.7%
P(not reported|predicted free)	99.2%	99.2%	99.2%	94.9%	94.9%	94.9%

Rabies Prediction Model Fit and Parameters

The Spearman correlation between the predicted and reported number of cases was 0.77 in raccoon territory and 0.44 in skunk territory and was applied to the test data sets using the original model (Table 2). Varying the surveillance definitions as described in the *Methods* section did not substantially improve model fit (Table 2), with AUCs that generally varied only by 0.01 or 0.02 across all models.

Within both the skunk and raccoon reservoir territories, counties predicted to have a higher probability of observing rabid animals (if present) were more likely to report rabid bats (AUC 0.76 in raccoon territory and 0.78 in skunk territory).

Reviewing the rabies prediction model parameters, the probability that a county is free from rabies decreases with the increasing number of rabid animals reported in the previous year in the county of interest. The probability decreases with an increasing number of rabid animals reported in the previous year in any adjacent counties, although the latter effect is less impactful. The probability of rabies freedom increases with increasing years since the last rabies case was reported in the county of interest and also increases with increasing years since the last rabies case was reported in any adjacent counties, although again the latter effect is less impactful on the model prediction (model 1A in Table 3).

If a county is not free from rabies, the number of reported cases in 1 year is positively associated with the number of positive reported cases in the following year. The number of cases reported also increases with an increasing number of rabid animals reported in any adjacent counties, although the effect is less strong. Finally, the number of cases reported the next year increases with surveillance points tested over the past 5 years in the county of interest (model 1B in Table 3).

Table 2. Performance of original model and models using different surveillance point definitions on the training and validation data sets.

	Spearman correlation predicted: reported cases	AUC[a] predicted: reported zero cases	AUC predicted free: reported zero cases
Raccoon model			
Original model, training data set	0.82	0.93	0.93
Different surveillance definitions[b], training data set	0.81-0.83	0.93-0.93	0.93-0.93
Original model predictions compared with rabies presence or absence in the next 5 years, training data set	0.89	0.97	0.97
Original model applied to validation data set (2019)	0.78	0.90	0.90
Original model applied to validation data set (2020)	0.77	0.90	0.90
Skunk model			
Original model, training data set	0.52	0.89	0.89
Different surveillance definitions[b], training data set	0.51-0.52	0.89-0.89	0.89-0.89
Original model predictions compared with rabies presence or absence in the next 5 years, training data set	0.66	0.88	0.88
Original model applied to test data set (2019)	0.44	0.91	0.91
Original model applied to test data set (2020)	0.44	0.91	0.91

[a] AUC: area under the receiver operating characteristic curve.

[b] Explored definition included the following: including only domestic or only reservoir animals, adjusting the relative value of surveillance effort by changing the coefficient applied to domestic animals from 0.1 to 1 times that of reservoir animals, applying specific coefficients to different species based on their observed test positivity rates, changing the number of years of surveillance effort considered from 1 to 8, and normalizing surveillance effort by county size or population.

Table 3. Rabies prediction zero-inflated negative binomial model parameters (fit to county-level data on the number of rabid animals detected from 2003 to 2018). Models were fit separately in raccoon reservoir territory and skunk reservoir territory.

	Raccoon territory				Skunk territory			
	Rabies freedom model (1A)		Rabies count model (1B)		Rabies freedom model (1A)		Rabies count model (1B)	
	Value (SE[a])	*P* value	Value (SE)	*P* value	Value (SE)	*P* value	Value (SE)	*P* value
Intercept (α_0/β_0)	−1.22 (N/A[b])	N/A	−1.18 (N/A)	N/A	−2.00 (N/A)	N/A	−1.46 (N/A)	N/A
α_1/β_1: slope log(number of rabid animals reported in the county of interest in the past year + 0.5)	−0.59 (0.08)	<.001	0.38 (0.01)	<.001	−0.51 (0.13)	<.001	0.39 (0.02)	<.001
α_2/β_2: slope log(number of rabid animals reported in any neighboring county in the past year + 0.5)	−0.45 (0.06)	<.001	0.22 (0.01)	<.001	−0.40 (0.02)	<.001	0.24 (0.02)	<.001
α_3: slope log(years with no rabies reported in the county of interest)	1.28 (0.11)	<.001	—[c]	—	1.22 (0.10)	<.001	—	—
α_4: log(years with no rabies reported in any neighboring county)	0.62 (0.14)	<.001	—	—	0.91 (0.08)	<.001	—	—
β_3: slope log(surveillance points in the past 5 years + 0.5)	—	—	0.30 (0.01)	<.001	—	—	0.32 (0.01)	<.001
I(r): dispersion parameter	—	—	1.09 (0.03)	<.001	—	—	−0.16 (0.04)	<.001

[a] SE: sensitivity.

[b] N/A: not applicable.

[c] Not available.

Model-Based Evaluation of Historical Definition

The model predicts an approximately 98% probability of observing zero cases in the next year in both raccoon and skunk territories (Table 4) for a hypothetical county meeting the historical rabies-free definition (ie, 15 surveillance points, 5 years with no cases in the county of interest, and 5 years with no cases in any adjacent counties). This probability is equivalent to the probability of rabies in counties predicted to be rabies free from Table 1, except that Table 1 considers all counties meeting or exceeding the historical rabies-free criteria, whereas Table 4 considers a scenario that exactly meets the criteria.

Table 4. Model predictions for a scenario meeting historical rabies freedom criteria. In contrast to this assumes exactly 5 years with no cases reported in the county of interest or any neighboring counties and exactly 15 surveillance points over the past 5 years.

	Model: raccoon territory	Model: skunk territory
NPV[a]: probability of zero cases reported in the next year, historical criteria are met	0.982	0.979
Probability of being rabies free in the next year, historical criteria are met	0.947	0.920
Probability that a case occurs in the next year but goes unreported, historical criteria are met	0.034	0.059

[a] NPV: negative predictive value.

Separating the probability of zeros derived from the rabies freedom model (ie, county is truly nonendemic for rabies) and those derived from the count model (ie, county is endemic for rabies, but no rabies cases are reported), the probability of being truly rabies free is slightly higher for raccoon territory (0.95) than skunk territory (0.92), and the probability that rabies is present but unrecognized is slightly lower in raccoon territory (0.03) than in skunk territory (0.06).

County-Specific Predictions: Rabies Freedom and Number of Reported Cases

In contrast to the binary yes or no cutoff of the historical rabies freedom definition, the rabies prediction model allows each county to have a predicted probability of being free from rabies and the expected number of terrestrial rabies cases reported in the following year. The median predicted probability of rabies freedom in 2020 was 0.10 (IQR 0.04-0.71) for counties in raccoon territory and 0.91 (IQR 0.55-0.96) for counties in skunk territory. The median expected number of rabies cases reported per county in 2020 was 1.4 (IQR 0.2-3.6) in raccoon territory and 0.03 (IQR 0.01-0.2) in skunk territory. For comparison, the median reported number of rabies cases reported in 2020 was 1.0 (IQR 0.0-3.0) in raccoon territory and 0.0 (IQR 0.0-0.0) in skunk territory. Counties with cases reported in 2020 tended to have a low probability of freedom from rabies (Figure 1).

Discussion

Principal Findings

The results of this study show that the historical definition of rabies freedom in the United States is largely successful in identifying counties with the lowest risk of rabies. Applying this definition, 99.2% of counties predicted as rabies free did not have a case reported in the following year. Modeling analyses showed that all 3 components of this definition (in-county presence, neighboring county presence, and surveillance effort) play an important role in making this prediction. Increased number of years since a case was last reported in the county of interest or any contiguous counties is associated with increased probability of rabies freedom, and increased surveillance over the past 5 years is associated with better rabies observation in the subsequent year in endemic counties.

Although the evaluation method used in these analyses supports the use of the historical definition, it also presents a more nuanced assessment of rabies freedom. Although we did not find an advantage of changing the definition of surveillance points from its historical definition (0.5 points per domestic animal tested and 1 point per reservoir animal tested), the model allowed us to estimate an individualized risk for each county rather than a binary designation of rabies presence or absence (Figures 2 and 3). Predictions of the probability of rabies freedom and the expected number of cases reported in 2019 and 2020 correlated with actual cases reported in the same years.

This analysis only considered variables related to historical rabies surveillance and detection, as in the historical definition. The inclusion of other variables such as county population and the size of land use could be considered in the future. In addition, modifications to these methods that consider the point locations of samples could be considered when subcounty designations are needed. Subcounty location data (eg, Geographic Information Systems coordinates) are not routinely reported to the CDC's NRSS, necessitating an aggregated county-level analysis. Trade-offs between distance and time from the last case reported could also be considered (eg, is it possible to consider an area "rabies free" if a recent case was reported but is very far from the nearest positive?).

Figure 2. Raccoon territory model predictions of the probability of rabies freedom (A) and the number of expected rabies cases (B), compared with the number of actual reported rabies cases per county (C) in 2020. Model estimated probability that a rabies case will occur but go unreported (D).

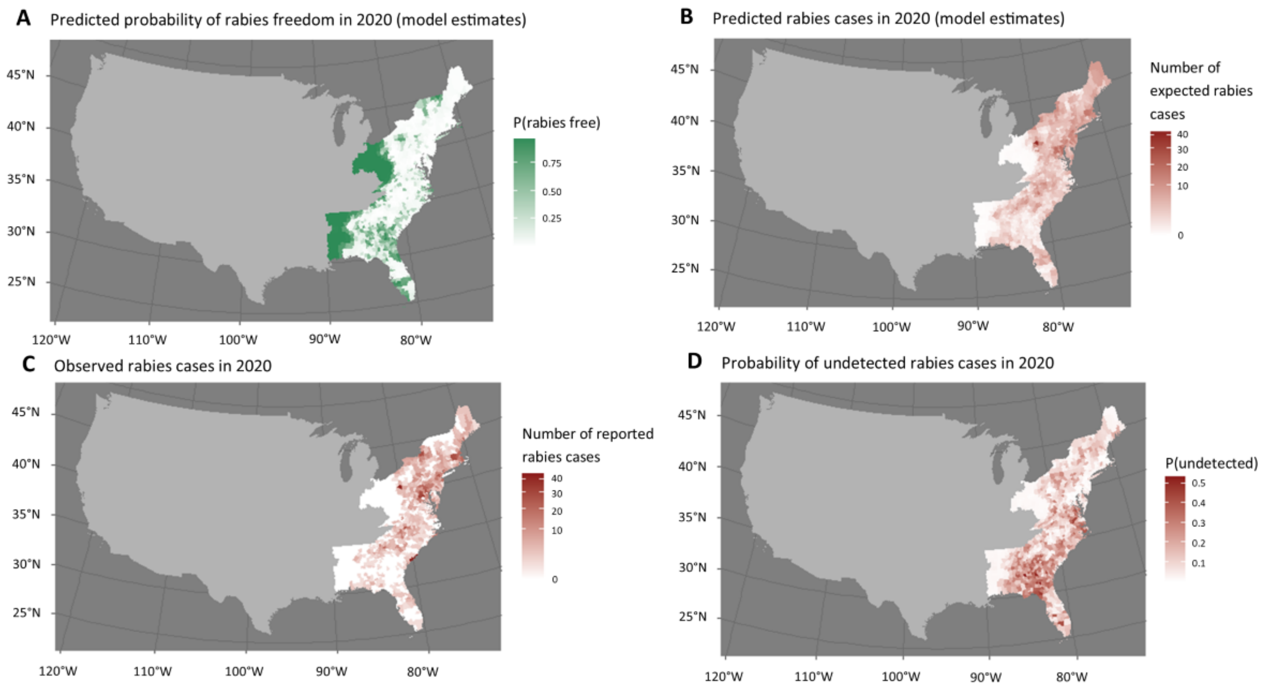
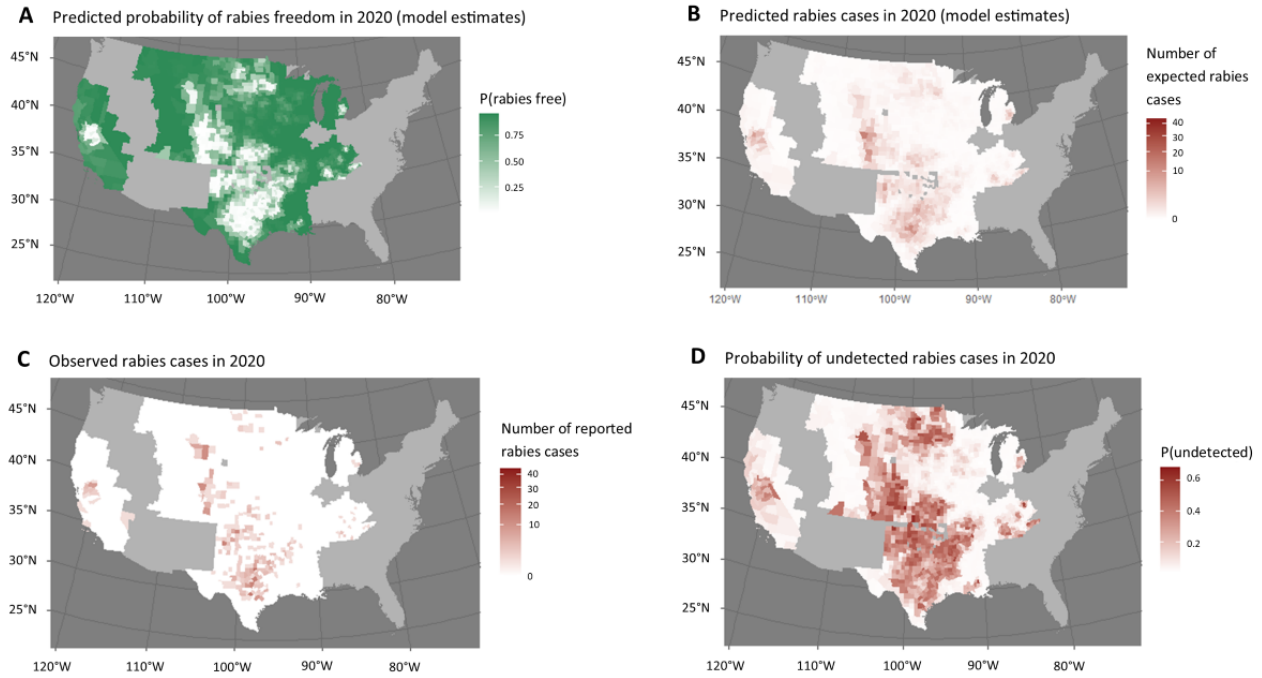


Figure 3. Skunk territory model predictions of the probability of rabies freedom (A) and the number of expected rabies cases (B), compared with the number of actual reported rabies cases per county (C) in 2020. Model estimated probability that a rabies case will occur but go unreported (D).



Designations of terrestrial rabies endemicity are important for clinical, public health, and wildlife management decision-making at a local, state, and national scale. The model presented in this study uses data only from the United States, and model parameters may not be generalizable to other countries, particularly those where DMRVVs may be found or where surveillance systems differ substantially from the US NRSS. When considering the application of these methods to locations outside of the United States, several important considerations must be made. First, the models described here differed according to the reservoir species. This suggests that a modeling process is required for each unique rabies reservoir species, of which >30 species have been described globally. Second, administrative divisions may differ substantially from US counties by size or population, making rabies detection more or less likely.

The United States has one of the most robust rabies surveillance systems in the world, with more laboratories and testing per capita than nearly any other country. From 2016 to 2020, nearly all counties in the United States tested at least 1 terrestrial animal for rabies (Figure 1). Therefore, under the US system, the amount of testing conducted in a county-year may not be as important as having access to laboratories and relatively high rabies awareness among the public and health jurisdictions. According to the US surveillance scheme, the cost of testing and access to laboratories are rarely cited as reasons for not testing animals suspected of having rabies. Although access to rabies testing in the United States is robust, 1 sensitivity analysis presented in this study showed that counties predicted to detect more terrestrial rabies cases (ie, those with greater surveillance efforts for terrestrial animals and more terrestrial animal rabies reported in the past year) are more likely to detect rabid bats in the following year. This could indicate that accessing testing is not equivalent in all counties, and future analyses could consider investigating ways to identify locations where a lack of testing could overshadow rabies virus transmission. In countries where access to testing is less secure and awareness of rabies and appropriate postexposure behaviors is lacking, predictive variables (such as number of samples tested) are likely to be more important in predicting the presence or absence of rabies in a defined area.

Limitations

This study had several limitations. Variant typing was not performed on all rabid animals in the United States, and some rabid animals with bat variants may have been inadvertently included in our data set. However, previous studies have found that >99% of terrestrial mammals in an area with a terrestrial RVV have the expected viral variant [26]. Of the 9 county-years for which rabies was reported despite meeting the rabies freedom definition, at least 2 had bat variants that were not reported to the CDC at the time of the analysis. Some states inconsistently reported negative rabies testing results in the early years of our data, leading us to exclude certain years of data from some states; however, <10% of the data were affected. In addition, we included data from the United States only, so counties bordering Mexico or Canada have incomplete data on rabies cases in the surrounding areas. We only explicitly developed models for raccoon and skunk RVV territories, as the number of counties with other terrestrial rabies variants in the United States is small. Counties in states, such as Arizona and New Mexico, with multiple RVVs were not included when fitting the models, as the rabies dynamics may differ in these regions. It is possible that animal location data submitted to the NRSS may not reflect the location where the animal was exposed to rabies but could reflect other locations noted by case investigators such as the location where the animal was found.

Finally, it is possible that even counties in which no rabies is reported over long periods may not truly be free from rabies.

Conclusions

In conclusion, the historical rabies freedom definition is a reasonable approach for identifying counties that are truly free from terrestrial raccoon and skunk rabies virus transmission. Gradations in risk can also be measured using the rabies prediction model presented in this study, with AUCs >0.9. Nevertheless, bat variant rabies in terrestrial mammals remains a possibility throughout the United States, excluding Hawaii, and translocation events can occur, which can lead to unpredictable shifts in rabies epidemiology [16]. Even counties with a high probability of rabies freedom should maintain vigilance and rabies testing capacity, as there are numerous examples of anthropogenic and natural translocations of rabies-infected animals that can cause major changes in the epidemiology of rabies.

Acknowledgments

The findings and conclusions of this study are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention. The authors thank the state and territorial public health and agriculture departments and laboratories for their annual contribution of rabies surveillance data to the National Rabies Surveillance System.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Fooks AR, Banyard AC, Horton DL, Johnson N, McElhinney LM, Jackson AC. Current status of rabies and prospects for elimination. Lancet 2014 Oct 11;384(9951):1389-1399 [FREE Full text] [doi: 10.1016/S0140-6736(13)62707-5] [Medline: 24828901]
2. Warrell MJ, Warrell DA. Rabies: the clinical features, management and prevention of the classic zoonosis. Clin Med (Lond) 2015 Feb;15(1):78-81 [FREE Full text] [doi: 10.7861/clinmedicine.14-6-78] [Medline: 25650205]
3. Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, Global Alliance for Rabies Control Partners for Rabies Prevention. Estimating the global burden of endemic canine rabies. PLoS Negl Trop Dis 2015 Apr;9(4):e0003709 [FREE Full text] [doi: 10.1371/journal.pntd.0003709] [Medline: 25881058]
4. Wallace RM, Undurraga EA, Blanton JD, Cleaton J, Franka R. Elimination of dog-mediated human rabies deaths by 2030: needs assessment and alternatives for progress based on dog vaccination. Front Vet Sci 2017;4:9 [FREE Full text] [doi: 10.3389/fvets.2017.00009] [Medline: 28239608]
5. Velasco-Villa A, Reeder SA, Orciari LA, Yager PA, Franka R, Blanton JD, et al. Enzootic rabies elimination from dogs and reemergence in wild terrestrial carnivores, United States. Emerg Infect Dis 2008 Dec;14(12):1849-1854 [FREE Full text] [doi: 10.3201/eid1412.080876] [Medline: 19046506]
6. Pieracci EG, Pearson CM, Wallace RM, Blanton JD, Whitehouse ER, Ma X, et al. Vital signs: trends in human rabies deaths and exposures - United States, 1938-2018. MMWR Morb Mortal Wkly Rep 2019 Jun 14;68(23):524-528 [FREE Full text] [doi: 10.15585/mmwr.mm6823e1] [Medline: 31194721]
7. Ma X, Bonaparte S, Toro M, Orciari LA, Gigante CM, Kirby JD, et al. Rabies surveillance in the United States during 2020. J Am Vet Med Assoc 2022 May 05;260(10):1157-1165 [FREE Full text] [doi: 10.2460/javma.22.03.0112] [Medline: 35522584]
8. Kunkel A, Minhaj FS, Whitehill F, Austin C, Hahn C, Kieffer AJ, et al. Notes from the field: three human rabies deaths attributed to bat exposures - United States, August 2021. MMWR Morb Mortal Wkly Rep 2022 Jan 07;71(1):31-32 [FREE Full text] [doi: 10.15585/mmwr.mm7101a5] [Medline: 34990442]
9. Rao AK, Briggs D, Moore SM, Whitehill F, Campos-Outcalt D, Morgan RL, et al. Use of a modified preexposure prophylaxis vaccination schedule to prevent human rabies: recommendations of the advisory committee on immunization practices — United States, 2022. Centers for Disease Control and Prevention. 2022 May 6. URL: https://www.cdc.gov/mmwr/volumes/71/wr/mm7118a2.htm [accessed 2023-03-20]
10. Whitehouse ER, Person MK, Brown CM, Slavinski S, Rao AK, Blanton JD. Evaluating surveillance for and estimating administration of rabies postexposure prophylaxis in the United States, 2012-2018. PLoS Negl Trop Dis 2021 Oct;15(10):e0009878 [FREE Full text] [doi: 10.1371/journal.pntd.0009878] [Medline: 34695115]
11. Manning SE, Rupprecht CE, Fishbein D, Hanlon CA, Lumlertdacha B, Guerra M, Advisory Committee on Immunization Practices Centers for Disease ControlPrevention (CDC). Human rabies prevention--United States, 2008: recommendations of the Advisory Committee on Immunization Practices. MMWR Recomm Rep 2008 May 23;57(RR-3):1-28 [FREE Full text] [Medline: 18496505]

12. Ma X, Monroe BP, Wallace RM, Orciari LA, Gigante CM, Kirby JD, et al. Rabies surveillance in the United States during 2019. J Am Vet Med Assoc 2021 Jun 01;258(11):1205-1220. [doi: 10.2460/javma.258.11.1205] [Medline: 33978439]
13. Scatterday J, Schneider N, Lewis A. Sporadic animal rabies in Florida. Public Health Rep (1896) 1960 Oct;75(10):945-953 [FREE Full text] [Medline: 13747108]
14. Jenkins SR, Winkler WG. Descriptive epidemiology from an epizootic of raccoon rabies in the Middle Atlantic States, 1982-1983. Am J Epidemiol 1987 Sep;126(3):429-437. [doi: 10.1093/oxfordjournals.aje.a114674] [Medline: 3497576]
15. Rabies, animal. National Notifiable Diseases Surveillance System (NNDSS). URL: https://ndc.services.cdc.gov/conditions/rabies-animal/ [accessed 2021-01-26]
16. Rabies, human. National Notifiable Diseases Surveillance System (NNDSS). URL: https://ndc.services.cdc.gov/conditions/rabies-human/ [accessed 2021-02-26]
17. Revision of public health reporting and timeframes for national notification for animal rabies. Council of State and Territorial Epidemiologists. 2022. URL: https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/ps2022/22-ID-06_Animal_Rabies.pdf [accessed 2023-02-20]
18. Revision of the surveillance case definition for human rabies. CSTE Position Statement Archive. 2010. URL: https://cdn.ymaws.com/www.cste.org/resource/resmgr/PS/10-ID-16.pdf [accessed 2023-03-20]
19. Blanton JD, Krebs JW, Hanlon CA, Rupprecht CE. Rabies surveillance in the United States during 2005. J Am Vet Med Assoc 2006 Dec 15;229(12):1897-1911 [FREE Full text] [doi: 10.2460/javma.229.12.1897] [Medline: 17173527]
20. Wharton M, Chorba TL, Vogt RL, Morse DL, Buehler JW. Case definitions for public health surveillance. MMWR Recomm Rep 1990 Oct 19;39(RR-13):1-43 [FREE Full text] [Medline: 2122225]
21. Blasco‐Moreno A, Pérez‐Casany M, Puig P, Morante M, Castells E. What does a zero mean? Understanding false, random and structural zeros in ecology. Methods Ecol Evol 2019 May 09;10(7):949-959. [doi: 10.1111/2041-210x.13185]
22. Wenger SJ, Freeman MC. Estimating species occurrence, abundance, and detection probability using zero-inflated distributions. Ecology 2008 Oct;89(10):2953-2959. [doi: 10.1890/07-1127.1] [Medline: 18959332]
23. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
24. Achim Z, Christian K, Jackman S. Regression models for count data in R. J Stat Softw 2008 Jul;27(8):1-25. [doi: 10.18637/jss.v027.i08]
25. Singh AJ, Chipman RB, de Fijter S, Gary R, Haskell MG, Kirby J, et al. Translocation of a stray cat infected with rabies from North Carolina to a terrestrial rabies-free county in Ohio, 2017. MMWR Morb Mortal Wkly Rep 2018 Oct 26;67(42):1174-1177 [FREE Full text] [doi: 10.15585/mmwr.mm6742a2] [Medline: 30359345]
26. Wallace RM, Gilbert A, Slate D, Chipman R, Singh A, Cassie W, et al. Right place, wrong species: a 20-year review of rabies virus cross species transmission among terrestrial mammals in the United States. PLoS One 2014;9(10):e107539 [FREE Full text] [doi: 10.1371/journal.pone.0107539] [Medline: 25295750]

Abbreviations

AUC: area under the receiver operating characteristic curve
CDC: Centers for Disease Control and Prevention
DMRVV: dog-mediated rabies virus variant
NRSS: National Rabies Surveillance System
PEP: postexposure prophylaxis
RVV: rabies virus variant

Edited by A Mavragani, T Sanchez; submitted 30.09.22; peer-reviewed by K Brown, M Salman, H Schwermer; comments to author 07.12.22; revised version received 27.01.23; accepted 14.02.23; published 07.04.23.

Please cite as:
Kunkel A, Veytsel G, Bonaparte S, Meek H, Ma X, Davis AJ, Bonwitt J, Wallace RM
Defining County-Level Terrestrial Rabies Freedom Using the US National Rabies Surveillance System: Surveillance Data Analysis
JMIR Public Health Surveill 2023;9:e43061
URL: https://publichealth.jmir.org/2023/1/e43061
doi:10.2196/43061
PMID:37027194

Original Paper

Survival and Trends in Annualized Hazard Function by Age at Diagnosis Among Chinese Breast Cancer Patients Aged ≤40 Years: Case Analysis Study

Jun Wang[1*], MD; Ting Luo[2*], MD; Zhong-zheng Xiang[1*], MD; Ming-min He[1], MD; Yuan-yuan Zeng[1], MD; Tian Yang[1], MD; Xiao-yuan Wei[1], MD; Siting Yu[1], MD; Ze-lei Dai[1], MD; Ning-yue Xu[1], MD; Chen-feng Tan[1], MD; Lei Liu[1*], MD

[1]Department of Head and Neck Oncology, West China Hospital, Sichuan University, Chengdu, China

[2]Breast Disease Center, Cancer Center, West China Hospital, Sichuan University, Chengdu, China

[*]these authors contributed equally

Corresponding Author:
Lei Liu, MD
Department of Head and Neck Oncology
West China Hospital
Sichuan University
No. 37 Guoxue Alley
Wuhou District
Chengdu, 610044
China
Phone: 86 18805053037
Email: liuleihx@gmail.com

Abstract

Background: Young breast cancer patients are more likely to develop aggressive tumor characteristics and a worse prognosis than older women, and different races and ethnicities have distinct epidemiologies and prognoses. However, few studies have evaluated the clinical biological features and relapse patterns in different age strata of young women in Asia.

Objective: We aimed to explore survival differences and the hazard function in young Chinese patients with breast cancer (BC) by age.

Methods: The patients were enrolled from West China Hospital, Sichuan University. The chi-squared test, a Kaplan-Meier analysis, a log-rank test, a Cox multivariate hazards regression model, and a hazard function were applied for data analysis. Locoregional recurrence–free survival (LRFS), distant metastasis–free survival (DMFS), breast cancer–specific survival (BCSS), and overall survival (OS) were defined as end points.

Results: We included 1928 young BC patients diagnosed between 2008 and 2019. Patients aged 18 to 25, 26 to 30, 31 to 35, and 36 to 40 years accounted for 2.7% (n=53), 11.8% (n=228), 27.7% (n=535), and 57.7% (n=1112) of the patients, respectively. The diagnosis of young BC significantly increased from 2008 to 2019. Five-year LRFS, DMFS, BCSS, and OS for the entire population were 98.3%, 93.4%, 94.4%, and 94%, respectively. Patients aged 18 to 25 years had significantly poorer 5-year LRFS ($P<.001$), 5-year DMFS ($P<.001$), 5-year BCSS ($P=.04$), and 5-year OS ($P=.04$) than those aged 31 to 35, 26 to 30, and 36 to 40 years. The hazard curves for recurrence and metastasis for the whole cohort continuously increased over the years, while the BC mortality risk peaked at 2 to 3 years and then slowly decreased. When stratified by age, the annualized hazard function for recurrence, metastasis, and BC mortality in different age strata showed significantly different trends, especially for BC mortality.

Conclusions: The annual diagnosis of young BC seemed to increase in Chinese patients, and the distinct age strata of young BC patients did not differ in survival outcome or failure pattern. Our results might provide strategies for personalized management of young BC.

(JMIR Public Health Surveill 2023;9:e47110) doi:10.2196/47110

KEYWORDS

breast cancer; young age; age strata; survival; annual hazard function; China

Introduction

Breast cancer (BC) is the most frequently diagnosed cancer (approximately 31% of all cancer sites) and is the second cause of mortality among female patients based on a 2022 prediction [1]. Young BC, defined as BC occurring in people aged ≤40 years, has always been a hotly discussed issue due to its lower incidence but poorer prognosis than BC in older patients [2]. The annual number of diagnoses of young BC is approximately 14,000 in the United States, with an estimated incidence of 5% to 7%, while a higher proportion is reported in Asia, up to 20% [3-5]. The diagnosis of young BC has sharply increased in several countries over the last years [3,6].

Young BC patients are more likely to develop aggressive tumor characteristics and have a worse prognosis compared with older patients [7-15]. Previous studies that used population-based data have reported that young BC patients have higher rates of advanced, poorly differentiated tumors, estrogen receptor (ER) negativity, progesterone receptor (PR) negativity, human epidermal growth factor receptor 2 (HER-2) positivity, a higher Ki-67 index, and lymphovascular invasion [7-12]. Young BC patients have also been definitively demonstrated to have a higher proportion of invasive molecular subtypes, including triple-negative and HER-2–positive subtypes; higher rates of distant disease at initial diagnosis; and poor long-term survival outcomes compared with older patients [13-15].

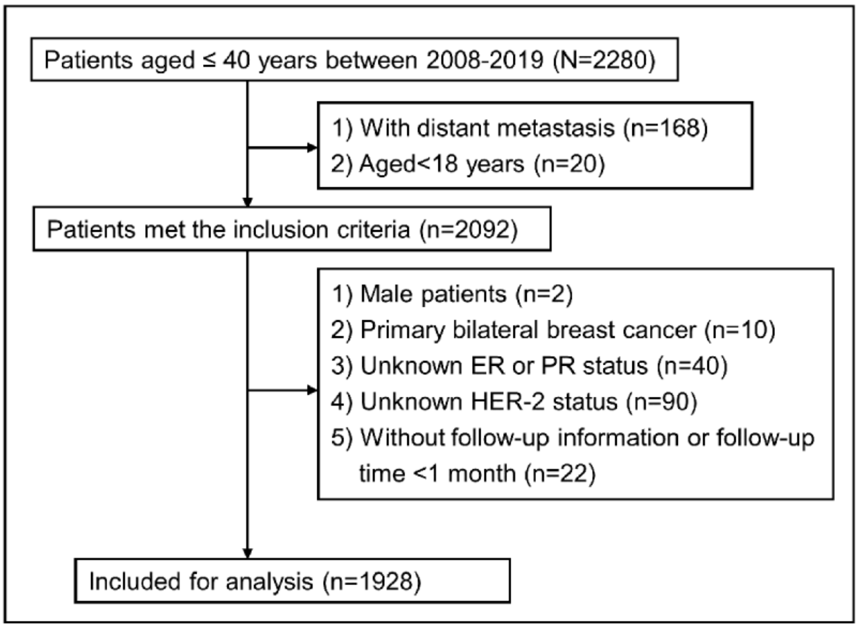
Several studies have reported that the epidemiology and prognosis of young BC also vary in different races and ethnicities [16-18]. The incidence of young BC in the United States is significantly lower than that in Asian countries (approximately 7% vs approximately 20%) [3-6]. Young African-American BC patients have increased risk of BC at a younger age, higher pathological grade, and higher rates of hormone receptor (HR) negativity compared to young White women [16,17]. In addition, young Asian patients have higher proportions of advanced-stage cancers and lower rates of poorly differentiated cancers and invasive BC subtypes (triple-negative and luminal B subtypes) but a better prognosis than young White patients [18]. However, few studies have evaluated the clinical biological features and relapse patterns in different age strata among young women in Asia. Therefore, our study aims to explore the clinicopathological characteristics, survival outcomes, and hazard function of Chinese patients aged ≤40 years by age group (18-25, 26-30, 31-35, and 36-40 years).

Materials

Patients

The patient data were extracted from the database of West China Hospital, Sichuan University. The database, the Breast Cancer Information Management System (BCIMS), prospectively collects patient information from medical records on demographics, tumor characteristics, treatment, and follow-up. We included patients based on the following criteria: (1) diagnosis with BC between 2008 and 2019; (2) age ≤40 years; (3) nonmetastatic disease; and (4) availability of detailed information on age, tumor stage, nodal stage, clinical stage, ER status, PR status, HER-2 status, molecular subtype, surgery, neoadjuvant/adjuvant chemotherapy, radiotherapy, endocrine therapy, anti–HER-2 targeted therapy, and follow-up. We excluded patients with the following characteristics: (1) male, 2) age ≤18 years, and (3) bilateral BC. The inclusion flow sheet of the patients is shown in Figure 1.

Figure 1. Inclusion flow sheet of patients. ER: estrogen receptor; HER-2: human epidermal growth factor receptor 2; PR: progesterone receptor.



Variables

The following patient demographic and clinicopathological variables were included: age (18-25, 26-30, 31-35, and 36-40 years), tumor stage (tumor in situ [TIS], T1, T2, T3, T4), nodal stage (N0, N1, N2, N3), clinical stage (0, I, II, III), pathological grade (well differentiated, moderately differentiated, poorly differentiated/undifferentiated, unknown), ER status, PR status, HER-2 status, molecular subtype (luminal A, luminal B, HER-2 positive, triple-negative), surgery (breast-conserving, mastectomy, unknown), neoadjuvant chemotherapy, adjuvant chemotherapy, radiotherapy, endocrine therapy, and targeted therapy.

Treatment, Follow-Up, and End Points

Surgery, chemotherapy, and radiotherapy regimens were formulated by a cooperative, multidisciplinary group including surgeons, oncologists, radiotherapy physicians, and patients. Neoadjuvant/adjuvant chemotherapy depended on advanced tumor stage, advanced nodal stage, invasive molecular subtype, such as triple-negative BC, and the willingness to undergo breast-conserving surgery. Radiotherapy was administered to patients receiving breast-conserving surgery with a positive margin and positive axillary lymph nodes. Patients who were HR positive received neoadjuvant or adjuvant endocrine therapy, and patients with HER-2 overexpression were treated with an anti–HER-2 targeted therapy, permitting economic conditions. Follow-up information was collected from medical records, office visits, and telephone calls every 3 months in the first 2 years, semiannually during years 2 to 5 years, and once a year after 5 years. The end points in this study were locoregional recurrence–free survival (LRFS), distant metastasis–free survival (DMFS), breast cancer–specific survival (BCSS), and overall survival (OS). The definitions of LRFS, DMFS, BCSS, and OS were stated in our previous study [19].

Statistical Analysis

The chi-squared test was used to compare differences in patient baseline characteristics in different age groups. Excel (2016 version; Microsoft Corp) was used to draw variation trends between 2008 and 2019 in the 4 age groups. A Kaplan-Meier analysis was applied to draw survival curves for LRFS, DMFS, OS, and BCSS. A log-rank test was used to compare differences between the 4 groups. A Cox multivariate hazards regression model was used to identify protective and risk factors for predicting LRFS, DMFS, OS, and BCSS. Annualized hazard rates for the whole group and different age strata, defined as percentage of events occurring within a time interval, were calculated using maximum likelihood estimate of a piece-wise exponential model. SPSS (version 25.0; IBM Corp) and Excel were used to analyze and map data. *P* values less than .05 (2-tailed) were considered statistically significant.

Ethical Considerations

Our study was approved by the Biomedical Ethics Committee of West China Hospital, Sichuan University (2020427). Informed consent was obtained from the participants when they first received treatment in our institution. Patient privacy was well protected due to the deidentification of their information.

Results

Patient Characteristics and Treatment Information

A total of 1928 patients met the inclusion criteria and were included for analysis. Patients aged 18 to 25 years, 26 to 30 years, 31 to 35 years, and 36 to 40 years accounted for 2.7% (n=53), 11.8% (n=228), 27.7% (n=535), and 57.7% (n=1112) of the participants, respectively. The proportions of patients at the TIS, T1, T2, T3, and T4 stages were 1.5% (n=28), 36.2% (n=697), 50.5% (n=976), 6.3% (n=121), and 5.5% (n=106), respectively. In addition, 44.6% (n=860), 32.2% (n=621), 11.8% (n=228), and 11.4% (n=219) of the patients were at the N0, N1, N2, and N3 stages, respectively.

The majority of the patients were at an early clinical stage (n=1405, 72.9%), had moderately or poorly differentiated tumors (n=1342, 69.7%), were ER positive (n=1345, 69.8%), were PR positive (n=1288, 66.8%), and were HER-2 negative (n=1383, 71.7%). Regarding molecular subtype, 17.9% (n=346), 54.9% (n=1059), 10.1% (n=194), and 17.1% (n=329) of the patients had the luminal A, luminal B, HER-2–positive, and triple-negative subtypes, respectively. Detailed patient characteristics are presented in Table 1.

In total, 17.4% (n=335) and 79.5% (n=1534) of the patients received breast-conserving surgery and mastectomy, respectively, while 3.1% (n=59) were not treated with surgery or had no record of surgery. In total, 19.8% (n=382) of the patients received neoadjuvant chemotherapy, 89.8% (n=1731) of the patients received adjuvant chemotherapy, and 45.7% (n=881) of the patients were treated with radiotherapy. In addition, 71% (n=1389) of the patients received endocrine therapy and 17.6% (n=340) were treated with targeted therapy (Table 1).

There were no statistical differences in baseline characteristics between the 4 groups (ages 18-25, 26-30, 31-35, and 36-40 years), including tumor stage (*P*=.53), nodal stage (*P*=.65), clinical stage (*P*=.33), ER status (*P*=.42), PR status (*P*=.16), HER-2 status (*P*=.89), pathological grade (*P*=.07), and molecular subtype (*P*=.43). However, patients aged 18 to 25 years were more likely to receive breast-conserving surgery (*P*<.001), adjuvant chemotherapy (*P*=.006), and radiotherapy (*P*=.02), while patients aged 36 to 40 years were more likely to receive mastectomy (*P*<.001). Patients aged 26 to 30 years were more likely to be treated with neoadjuvant chemotherapy (*P*<.001) and targeted therapy (*P*=.04; Table 1).

Table 1. Baseline characteristics of overall population.

Variables	Participants (n=1928), n (%)	Participants by age group (years), n (%)				*P* value
		18-25 (n=53)	26-30 (n=228)	31-35 (n=535)	36-40 (n=1112)	
Tumor stage						.53
Tumor in situ	28 (1.5)	0 (0)	4 (1.8)	10 (1.9)	14 (1.3)	
T1	697 (36.2)	24 (45.3)	81 (35.5)	187 (35)	405 (36.4)	
T2	976 (50.5)	22 (41.5)	110 (48.3)	265 (49.4)	579 (52.1)	
T3	121 (6.3)	4 (7.5)	19 (8.3)	40 (7.5)	58 (5.2)	
T4	106 (5.5)	3 (5.7)	14 (6.1)	33 (6.2)	56 (5)	
Nodal stage						.65
N0	860 (44.6)	20 (37.7)	101 (44.3)	237 (44.3)	502 (45.1)	
N1	621 (32.2)	15 (28.3)	71 (31.1)	174 (32.5)	361 (32.5)	
N2	228 (11.8)	9 (17)	23 (10.1)	65 (12.1)	131 (11.8)	
N3	219 (11.4)	9 (17)	33 (14.5)	59 (11)	118 (10.6)	
Clinical stage						.33
0	23 (1.2)	0 (0)	4 (1.8)	6 (1.1)	13 (1.2)	
I	445 (23.1)	17 (32.1)	52 (22.8)	119 (22.2)	257 (23.1)	
II	937 (48.6)	17 (32.1)	103 (45.2)	262 (49)	555 (49.9)	
III	523 (27.1)	19 (35.8)	69 (30.2)	148 (27.7)	287 (25.8)	
Pathological grade						.07
Well differentiated	34 (1.8)	0 (0)	7 (3.1)	8 (1.5)	19 (1.7)	
Moderately differentiated	556 (28.8)	11 (20.8)	60 (26.3)	178 (33.3)	307 (27.6)	
Poorly differentiated/undifferentiated	786 (40.8)	21 (39.6)	88 (38.6)	201 (37.6)	476 (42.8)	
Unknown	552 (28.6)	21 (39.6)	73 (32)	148 (27.6)	310 (27.9)	
Estrogen receptor status						.42
Positive	1345 (69.8)	33 (62.3)	162 (71.1)	364 (68)	786 (70.7)	
Negative	583 (30.2)	20 (37.7)	66 (28.9)	171 (32)	326 (29.3)	
Progesterone receptor status						.16
Positive	1288 (66.8)	30 (56.6)	147 (64.5)	349 (65.2)	762 (68.5)	
Negative	640 (33.2)	23 (43.4)	81 (35.5)	186 (34.8)	350 (31.5)	
Human epidermal growth factor receptor 2 status						.89
Positive	545 (28.3)	16 (30.2)	69 (30.3)	149 (27.9)	311 (28)	
Negative	1383 (71.7)	37 (69.8)	159 (69.7)	386 (72.1)	801 (72)	
Molecular subtype						.43
Luminal A	346 (17.9)	3 (5.7)	42 (18.4)	102 (19.1)	199 (17.9)	
Luminal B	1059 (54.9)	32 (60.4)	125 (54.8)	280 (52.3)	622 (55.9)	
HER-2 positive	194 (10.1)	6 (11.3)	20 (8.8)	54 (10.1)	114 (10.3)	
Triple negative	329 (17.1)	12 (22.6)	41 (18)	99 (18.5)	177 (15.9)	
Surgery						<.001
Breast conserving surgery	335 (17.4)	15 (28.3)	58 (25.4)	105 (19.6)	157 (14.1)	
Mastectomy	1534 (79.5)	35 (66)	165 (72.4)	409 (76.5)	925 (83.2)	
Unknown	59 (3.1)	3 (5.7)	5 (2.2)	21 (3.9)	30 (2.7)	
Neoadjuvant chemotherapy						<.001

Variables	Participants (n=1928), n (%)	Participants by age group (years), n (%)				*P* value
		18-25 (n=53)	26-30 (n=228)	31-35 (n=535)	36-40 (n=1112)	
Yes	382 (19.8)	15 (28.3)	69 (30.3)	107 (20)	191 (17.2)	
No	1546 (80.2)	38 (71.7)	159 (69.7)	428 (80)	921 (82.8)	
Adjuvant chemotherapy						.006
Yes	1731 (89.8)	53 (100)	201 (88.2)	466 (87.1)	1011 (90.9)	
No	197 (10.2)	0 (0)	27 (11.8)	69 (12.9)	101 (9.1)	
Radiotherapy						.02
Yes	881 (45.7)	35 (66)	110 (48.2)	242 (45.2)	494 (44.4)	
No	1047 (54.3)	18 (34)	118 (51.8)	293 (54.8)	618 (55.6)	
Endocrine therapy						.22
Yes	1389 (72)	37 (69.8)	175 (76.8)	372 (69.5)	805 (72.4)	
No	539 (28)	16 (30.2)	53 (23.2)	163 (30.5)	307 (27.6)	
Targeted therapy						.04
Yes	340 (17.6)	11 (20.8)	50 (21.9)	106 (19.8)	173 (15.6)	
No	1588 (82.4)	42 (79.2)	178 (78.1)	429 (80.2)	939 (84.4)	

Trends in Annual Diagnosis by Age

The trends in the diagnosis of BC in the whole cohort and patients aged 18 to 25, 26 to 30, 31 to 35, and 36 to 40 years from 2008 to 2019 are presented in Table 2. The proportion of young BC patients significantly increased from 4.7% (n=92) in 2008 to 12.8% (n=247) in 2019. When stratified by age, annual diagnosis of patients aged 36 to 40 years significantly decreased from 59.4% (52/92) in 2008 to 44.6% (86/194) in 2018, while the diagnosis of patients aged 26 to 30 showed a significant upward trend, from 5.9% (6/92) in 2008 to 17.5% (34/194) in 2018. However, there were no significant tendencies in the diagnosis of patients aged 18 to 25 years, with 3% (3/92) in 2008 and 5.1% (10/194) in 2018; for patients aged 31 to 35 years, diagnosis was 31.7% (31/92) in 2008 and 32.8% (64/194) in 2018 (Table 2).

Table 2. Trends of annual diagnosis of patients in different age groups (18-25, 26-30, 31-35, and 36-40 years).

Year	Age group (years), n (%)			
	18-25	26-30	31-35	36-40
2008 (n=92)	3 (3)	6 (5.9)	31 (31.7)	52 (59.4)
2009 (n=140)	5 (4.2)	11 (7.7)	36 (25.9)	88 (62.2)
2010 (n=120)	1 (0.8)	15 (12.5)	26 (21.7)	78 (65)
2011 (n=152)	3 (1.9)	10 (6.5)	32 (20.8)	107 (70.8)
2012 (n=160)	4 (2.5)	13 (8)	41 (25.9)	102 (63.6)
2013 (n=146)	3 (2.7)	9 (6.1)	30 (20.3)	104 (70.9)
2014 (n=144)	3 (2.1)	17 (11.8)	33 (22.9)	91 (63.2)
2015 (n=176)	3 (1.7)	26 (15.3)	50 (28.2)	97 (54.8)
2016 (n=149)	7 (4.7)	23 (15.4)	40 (26.8)	79 (53)
2017 (n=208)	5 (2.4)	40 (19.2)	68 (32.7)	95 (45.7)
2018 (n=194)	10 (5.1)	34 (17.5)	64 (32.8)	86 (44.6)
2019 (n=247)	6 (2.4)	24 (9.6)	84 (33.7)	133 (54.2)

Survival and Prognostic Analysis by Age

With a median follow-up time of 75.7 months (range 1.1-173 months), 75 (3.9%) recurrences, 236 (12.2%) cases of distant metastasis, 132 (6.8%) BC-related deaths, and 138 (7.2%) other deaths occurred among the 1928 participants. In the entire group, 5-year LRFS, DMFS, BCSS, and OS were 98.3%, 93.4%, 94.4%, and 94%, respectively. There were better survival outcomes with age in BC patients aged ≤40 years. Patients aged 18 to 25 years had significantly poorer 5-year LRFS (ages 18-25, 26-30, 31-35, and 36-40 years: 88.5%, 96.7%, 98.6%, and 98.9%, respectively; *P*<.001; Figure 2A), 5-year DMFS (ages

18-25, 26-30, 31-35, and 36-40 years: 82.8%, 89.1%, 93.7%, and 94.6%, respectively; *P*<.001; Figure 2B), 5-year BCSS (ages 18-25, 26-30, 31-35, and 36-40 years: 85%, 91.8%, 95.8%, and 94.5%, respectively; *P*=.04; Figure 2C), and 5-year OS (age 18-25, 26-30, 31-35, and 36-40 years: 85%, 91.3%, 95.8%, and 94%, respectively; *P*=.04; Figure 2D) than those aged 31-35, 26-30, and 36-40 years. Cox multivariate regression model showed that age was a significant predictor for LRFS (hazard ratio [HR] 0.645, 95% CI 0.500-0.831; *P*<.001) and DMFS (HR 0.743, 95% CI 0.641-0.861; *P*<.001), while it was not a predictor for BCSS (HR 0.889, 95% CI 0.729-1.805; *P*=.25) or OS (HR 0.867, 95% CI 0.709-1.061; *P*=.17; Table 3).

Figure 2. Survival curves of locoregional recurrence–free survival, distant metastasis–free survival, breast cancer–specific survival, and overall survival for young patients of different ages.

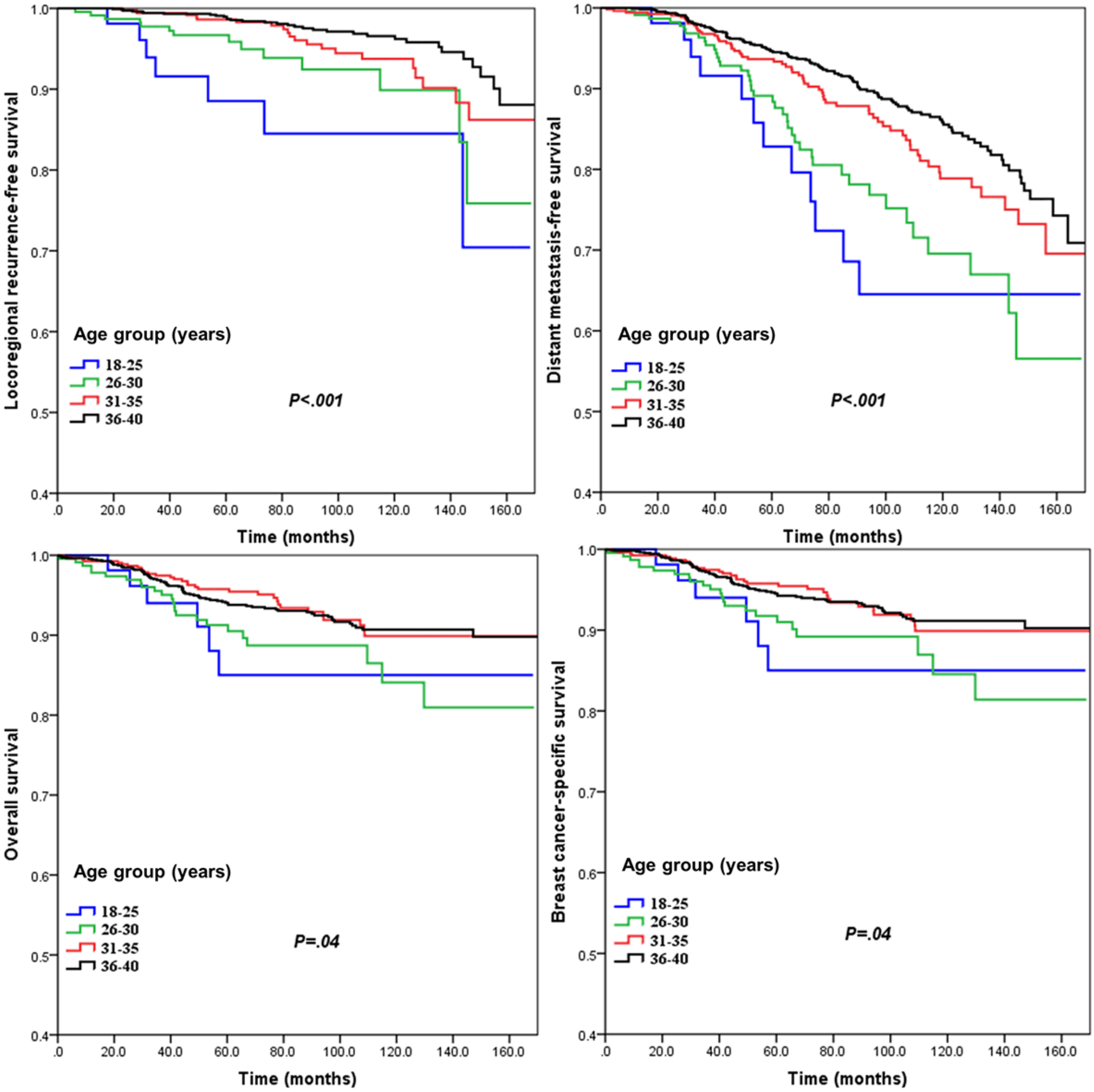


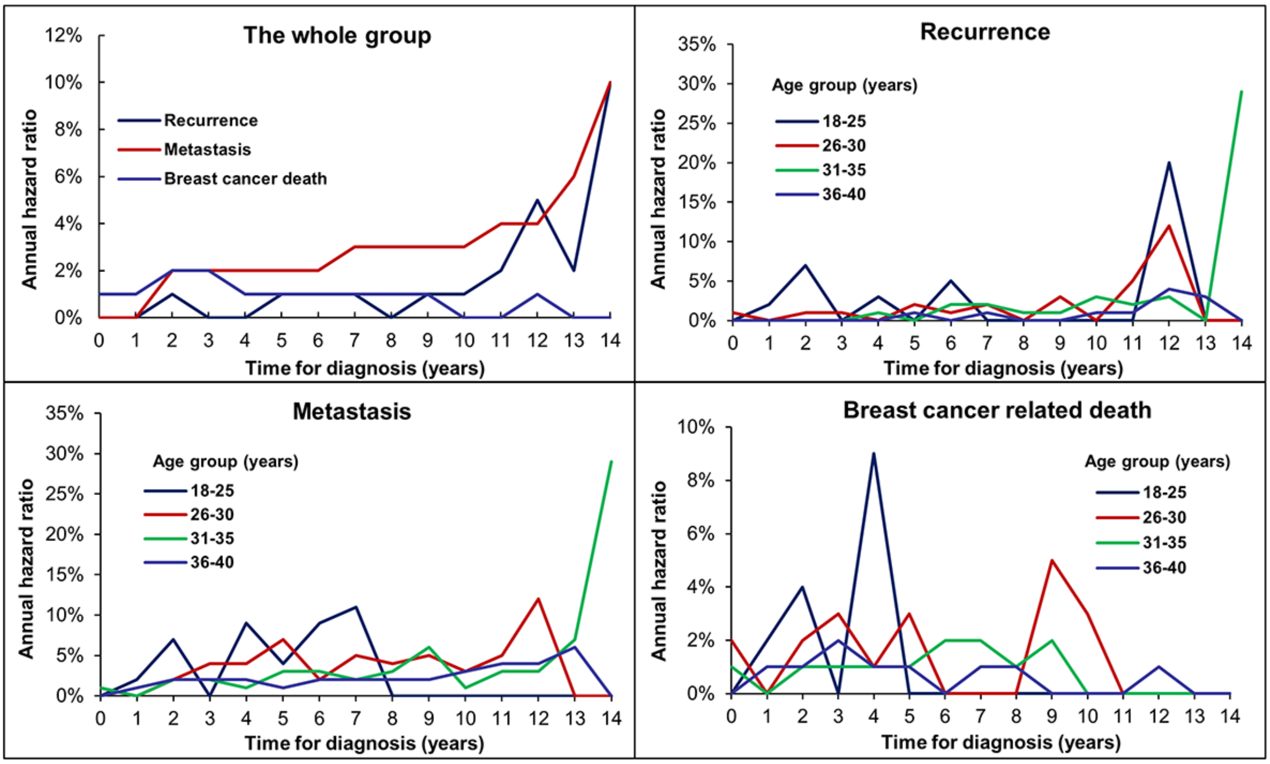
Table 3. Cox multivariate analysis of locoregional recurrence–free survival, distant metastasis–free survival, breast cancer–specific survival, and overall survival in all patients.

Outcome and age group (years)	Hazard ratio (95% CI)	*P* value
Locoregional recurrence–free survival		
18-25	1	
26-30	0.551 (0.193-1.571)	.27
31-35	0.280 (0.105-0.752)	.01
36-40	0.228 (0.088-0.590)	.002
Distant metastasis–free survival		
18-25	1	
26-30	0.967 (0.489-1.910)	.92
31-35	0.539 (0.280-1.035)	.06
36-40	0.448 (0.238-0.845)	.01
Breast cancer–specific survival		
18-25	1	
26-30	0.959 (0.378-2.432)	.93
31-35	0.518 (0.210-1.227)	.15
36-40	0.596 (0.252-1.410)	.24
Overall survival		
18-25	1	
26-30	0.997 (0.395-2.515)	.99
31-35	0.517 (0.210-1.272)	.15
36-40	0.644 (0.273-1.518)	.32

Annualized Hazard Curve of Recurrence, Metastasis, and BC Death

We also explored the annualized hazard trends of recurrence, metastasis, and BC mortality in this population (Figure 3). The hazard curves for recurrence and metastasis in the whole cohort continuously increased over time and did not reach a peak within the follow-up of 14 years. The BC mortality risk curve peaked at 2 to 3 years (at 2%), maintained a steady lower level after 3 years, then changed to a slowly decreasing plateau (Figure 3). After conducting a stratified analysis by age, we found that patients aged 18 to 25 years had peak BC mortality at 2 and 4 years and still had peak metastasis risk before the first 8 years. Patients aged 26 to 30, 31 to 35, and 36 to 40 years all showed a slow upward trend of recurrence and metastasis risk during follow-up. In addition, patients aged 26 to 30 years had high risk of BC mortality during years 0 to 6 and 8 to 11, while patients aged 31 to 35 years had high risk before the first 10 years. The hazard function for BC mortality in patients aged 36 to 40 years maintained a high plateau before the first 8 years then changed to a low death rate (Figure 3).

Figure 3. Annualized hazard curves for recurrence, metastasis, and breast cancer–related mortality for the whole group and distinct age strata of the patients.



Discussion

Principal Findings

We provide updated data on the clinicopathological characteristics, survival outcomes, and hazard function of different age strata of Chinese BC patients aged younger than 40 years. The main finding was that different age strata of young BC patients had different rates of survival: the younger the patient, the worse the prognosis. In addition, we found the risk of recurrence and metastasis continuously increased over time, and the failure patterns of different age strata were significantly different. Our result demonstrating different survival outcomes and failure patterns in different age strata could be used to tailor personalized management of young BC patients.

Previous studies have demonstrated that young BC patients have worse prognoses than older patients; 5-year BCSS in patients aged ≤35 years was approximately 80%, while it was 90% in patients ≥60 years [7,20]. The reasons for this biological difference between young and old patients have not been well elaborated. Several possible reasons might explain the phenomenon. First, younger patients are more likely to develop tumors with aggressive characteristics, such as higher Ki-67 and a triple-negative subtype [7-12]. Second, younger patients have higher expression of RANK-ligand, c-kit, mammary stem cell, and BRCA1 mutation signatures and the deregulation of PI3K and Myc pathways, which are associated with a poor prognosis [21-23]. Third, young BC patients have increased risk of psychosocial stress, and their treatment strategies are more likely to be affected by considerations of sexual function, fertility, beauty, body image, and their careers [24,25]. In this study, the 5-year BCSS and OS in young patients were 94.3% and 93.9%, respectively. These survival rates were higher than those of older Chinese patients in a previous study; the reason might be that more patients had early-stage tumors (6279/7553, 83.1%), as well as nodal-stage (6821/7553, 90.3%) and luminal-subtype (5937/7553, 78.6%) tumors in this study [26]. In addition, survival outcomes among younger patients in our study were also higher than those in younger women in the United States, which might be attributable to racial, dietary, climate, and living differences leading to Asian BC patients having better survival than White patients [7,26,27].

It has not been fully delineated whether survival and prognoses are the same in different age strata among young women. A study conducted by Fredholm et al [28] included 1120 women with stage I to III cancer; their results showed that patients aged <30 years had significantly poorer 5-year BCSS than those aged 35 to 40 years (80% vs 86%; *P*<.001). However, the opposite result was found in another population-based report in Italy: there was no survival difference among patients aged <25, 25 to 29, and 30 to 34 years in locoregional relapse (*P*=.87), distant metastasis (*P*=.40), BCSS (*P*=.58), and OS (*P*=.99) [27]. In our study, there were significant differences in LRFS (*P*<.001), DMFS (*P*<.001), BCSS (*P*=.035), and OS (*P*=.037) in younger patients by age, and patients aged <25 years old had the worst survival rate. The reasons for the difference might be that the sample size in the previous study was small (n=497) and that the included patients were White; these are important differences from our study [27]. Therefore, it is important to establish policies for better managing and improving the survival of very young Chinese BC patients.

There is limited evidence available evaluating the risk of relapse and death in young BC patients. A recent analysis from the

International Breast Cancer Study Group clinical trials included 4105 BC patients and explored the patterns of late BC recurrence. The authors demonstrated that the annualized hazard of recurrence peaked at 2 years and then decreased slowly; however, the study did not stratify the analysis for young BC patients [29]. Another study by de la Rochefordiere and his colleagues [30] assessed failure patterns in young age groups. They included 3371 women aged ≤55 years and divided them into 3 groups (≤33 years, 34-40 years, and ≥40 years). The results showed that the annual hazard rate of relapse peaked at 2 years and patients aged ≤33 years had a higher relapse risk than those aged 34 to 40 years for about 5 years, but a lower risk after 5 years [30]. A similar result was found in our study: BC death risk peaked at 2 and 3 years after diagnosis; however, totally different trends were observed for the annual hazard of recurrence and metastasis, which both showed a slow upward trend in our study that did not reach a peak during follow-up, possibly attributable to differences in race and sample size between the studies (1950 patients in this study vs 456 patients in the past study) [29]. The reasons for the different failure patterns might include differences in the expression of RANK-ligand, c-kit, mammary stem cell, and BRCA1 mutation signatures, as well as in the PI3K and Myc pathways [21-23]. Therefore, it is essential to further explore the optimal management of young BC patients to reduce the risk of recurrence and metastasis.

Limitations

Our study has several limitations that might affect the results. First, the patient data were extracted from a database that was built beforehand; selection bias in retrospective studies is inevitable. Second, our data were collected at a single center, and the result thus cannot represent all patients in China. Third, this study only included hospitalized patients, which probably does not represent all patients with cancer (ie, there may be patients that have cancer but are not hospitalized). Finally, the sample size in our study was relatively small, especially for patients aged 18 to 25 years. Therefore, a multicenter prospective study with a larger sample size should be conducted to further explore the characteristics of young BC patients and their management. Despite the limitations of our study, we have expanded the understanding of young BC in different races and ethnicities.

Conclusions

Our study demonstrates that the annual diagnosis of young BC increased in Chinese patients. The diagnostic rate of patients aged 36 to 40 years decreased, while it remained stable in those aged 18 to 25 and 31 to 35 years. Young BC patients of different age groups did not differ in survival outcomes, but the younger the patient, the worse were the 5-year LRFS, DMFS, BCSS, and OS. In addition, there were significantly different failure patterns in different age strata among the young BC patients, especially for BC mortality. Our results demonstrating different survival outcomes and failure patterns in different age strata may allow tailored, personalized management for young BC patients.

Acknowledgments

We thank all the patients who participated in this study.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JW and ZZX drafted the manuscript. TL and LL conceived the study. YYZ, TY, XYW, SY, NYX, ZLD, and CFT acquired and organized the data sets. JW, YYZ, TY, and XYW conducted the statistical analyses. MMH assisted in revising the manuscript. JW, ZZX, TL, and LL participated in the study design. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022 Jan;72(1):7-33 [FREE Full text] [doi: 10.3322/caac.21708] [Medline: 35020204]
2. Copson E, Eccles B, Maishman T, Gerty S, Stanton L, Cutress R, POSH Study Steering Group. Prospective observational study of breast cancer treatment outcomes for UK women aged 18-40 years at diagnosis: the POSH study. J Natl Cancer Inst 2013 Jul 03;105(13):978-988. [doi: 10.1093/jnci/djt134] [Medline: 23723422]
3. Brinton L, Sherman M, Carreon J, Anderson W. Recent trends in breast cancer among younger women in the United States. J Natl Cancer Inst 2008 Nov 19;100(22):1643-1648 [FREE Full text] [doi: 10.1093/jnci/djn344] [Medline: 19001605]

4. Cherbal F, Gaceb H, Mehemmai C, Saiah I, Bakour R, Rouis AO, et al. Distribution of molecular breast cancer subtypes among Algerian women and correlation with clinical and tumor characteristics: a population-based study. Breast Dis 2015;35(2):95-102. [doi: 10.3233/BD-150398] [Medline: 25736840]
5. Schlichting JA, Soliman AS, Schairer C, Harford JB, Hablas A, Ramadan M, et al. Breast cancer by age at diagnosis in the Gharbiah, Egypt, population-based registry compared to the United States Surveillance, Epidemiology, and End Results Program, 2004-2008. Biomed Res Int 2015;2015:381574 [FREE Full text] [doi: 10.1155/2015/381574] [Medline: 26495294]
6. Pollán M. Epidemiology of breast cancer in young women. Breast Cancer Res Treat 2010 Sep;123 Suppl 1:3-6. [doi: 10.1007/s10549-010-1098-2] [Medline: 20711654]
7. Gnerlich J, Deshpande A, Jeffe D, Sweet A, White N, Margenthaler J. Elevated breast cancer mortality in women younger than age 40 years compared with older women is attributed to poorer survival in early-stage disease. J Am Coll Surg 2009 Mar;208(3):341-347 [FREE Full text] [doi: 10.1016/j.jamcollsurg.2008.12.001] [Medline: 19317994]
8. Sidoni A, Cavaliere A, Bellezza G, Scheibel M, Bucciarelli E. Breast cancer in young women: clinicopathological features and biological specificity. Breast 2003 Aug;12(4):247-250. [doi: 10.1016/s0960-9776(03)00095-x] [Medline: 14659308]
9. Collins LC, Marotti JD, Gelber S, Cole K, Ruddy K, Kereakoglow S, et al. Pathologic features and molecular phenotype by patient age in a large cohort of young women with breast cancer. Breast Cancer Res Treat 2012 Feb;131(3):1061-1066. [doi: 10.1007/s10549-011-1872-9] [Medline: 22080245]
10. Keegan TH, DeRouen MC, Press DJ, Kurian AW, Clarke CA. Occurrence of breast cancer subtypes in adolescent and young adult women. Breast Cancer Res 2012 Mar 27;14(2):R55 [FREE Full text] [doi: 10.1186/bcr3156] [Medline: 22452927]
11. Colleoni M, Rotmensz N, Robertson C, Orlando L, Viale G, Renne G, et al. Very young women (<35 years) with operable breast cancer: features of disease at presentation. Ann Oncol 2002 Feb;13(2):273-279 [FREE Full text] [doi: 10.1093/annonc/mdf039] [Medline: 11886005]
12. Bharat A, Aft RL, Gao F, Margenthaler JA. Patient and tumor characteristics associated with increased mortality in young women (≤40 years) with breast cancer. J Surg Oncol 2009 Sep 01;100(3):248-251. [doi: 10.1002/jso.21268] [Medline: 19330813]
13. Morrison DH, Rahardja D, King E, Peng Y, Sarode VR. Tumour biomarker expression relative to age and molecular subtypes of invasive breast cancer. Br J Cancer 2012 Jul 10;107(2):382-387 [FREE Full text] [doi: 10.1038/bjc.2012.219] [Medline: 22713661]
14. Azim HA, Partridge AH. Biology of breast cancer in young women. Breast Cancer Res 2014 Aug 27;16(4):427 [FREE Full text] [doi: 10.1186/s13058-014-0427-5] [Medline: 25436920]
15. Johnson RH, Chien FL, Bleyer A. Incidence of breast cancer with distant involvement among women in the United States, 1976 to 2009. JAMA 2013 Feb 27;309(8):800-805. [doi: 10.1001/jama.2013.776] [Medline: 23443443]
16. Newman LA. Breast cancer disparities: high-risk breast cancer and African ancestry. Surg Oncol Clin N Am 2014 Jul;23(3):579-592. [doi: 10.1016/j.soc.2014.03.014] [Medline: 24882352]
17. Dunn BK, Agurs-Collins T, Browne D, Lubet R, Johnson KA. Health disparities in breast cancer: biology meets socioeconomic status. Breast Cancer Res Treat 2010 Jun;121(2):281-292. [doi: 10.1007/s10549-010-0827-x] [Medline: 20437200]
18. Shoemaker ML, White MC, Wu M, Weir HK, Romieu I. Differences in breast cancer incidence among young women aged 20-49 years by stage and tumor characteristics, age, race, and ethnicity, 2004-2013. Breast Cancer Res Treat 2018 Jun;169(3):595-606 [FREE Full text] [doi: 10.1007/s10549-018-4699-9] [Medline: 29445940]
19. Wang J, Zhong X, Luo T, Xiang Z, Zeng Y, Yang T, et al. The value of the 8th edition of American Joint Committee on Cancer Pathological Prognostic Staging on the selection of postmastectomy radiotherapy for T1-2N1 breast cancer. J Oncol 2022;2022:7550323 [FREE Full text] [doi: 10.1155/2022/7550323] [Medline: 36330354]
20. Assi H, Khoury K, Dbouk H, Khalil L, Mouhieddine T, El Saghir NS. Epidemiology and prognosis of breast cancer in young women. J Thorac Dis 2013 Jun;5 Suppl 1(Suppl 1):S2-S8 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2013.05.24] [Medline: 23819024]
21. Azim HJ, Michiels S, Bedard P, Singhal S, Criscitiello C, Ignatiadis M, et al. Elucidating prognosis and biology of breast cancer arising in young women using gene expression profiling. Clin Cancer Res 2012 Mar 01;18(5):1341-1351 [FREE Full text] [doi: 10.1158/1078-0432.CCR-11-2599] [Medline: 22261811]
22. Huzarski T, Byrski T, Gronwald J, Górski B, Domagala P, Cybulski C, et al. Ten-year survival in patients with BRCA1-negative and BRCA1-positive breast cancer. J Clin Oncol 2013 Sep 10;31(26):3191-3196. [doi: 10.1200/JCO.2012.45.3571] [Medline: 23940229]
23. Anders CK, Acharya CR, Hsu DS, Broadwater G, Garman K, Foekens JA, et al. Age-specific differences in oncogenic pathway deregulation seen in human breast tumors. PLoS One 2008 Jan 02;3(1):e1373 [FREE Full text] [doi: 10.1371/journal.pone.0001373] [Medline: 18167534]
24. Freedman RA, Partridge AH. Management of breast cancer in very young women. Breast 2013 Aug;22 Suppl 2:S176-S179 [FREE Full text] [doi: 10.1016/j.breast.2013.07.034] [Medline: 24074783]
25. Gabriel CA, Domchek SM. Breast cancer in young women. Breast Cancer Res 2010;12(5):212 [FREE Full text] [doi: 10.1186/bcr2647] [Medline: 21067532]

26. Lim DW, Giannakeas V, Narod SA. Survival differences in Chinese versus White women with breast cancer in the United States: A SEER-based analysis. JCO Glob Oncol 2020 Oct;6:1582-1592 [FREE Full text] [doi: 10.1200/GO.20.00316] [Medline: 33079607]
27. Cancello G, Maisonneuve P, Mazza M, Montagna E, Rotmensz N, Viale G, et al. Pathological features and survival outcomes of very young patients with early breast cancer: how much is "very young"? Breast 2013 Dec;22(6):1046-1051 [FREE Full text] [doi: 10.1016/j.breast.2013.08.006] [Medline: 24091127]
28. Fredholm H, Magnusson K, Lindström LS, Garmo H, Fält SE, Lindman H, et al. Long-term outcome in young women with breast cancer: a population-based study. Breast Cancer Res Treat 2016 Nov;160(1):131-143 [FREE Full text] [doi: 10.1007/s10549-016-3983-9] [Medline: 27624330]
29. Colleoni M, Sun Z, Price KN, Karlsson P, Forbes JF, Thürlimann B, et al. Annual hazard rates of recurrence for breast cancer during 24 years of follow-up: Results from the International Breast Cancer Study Group Trials I to V. J Clin Oncol 2016 Mar 20;34(9):927-935 [FREE Full text] [doi: 10.1200/JCO.2015.62.3504] [Medline: 26786933]
30. de la Rochefordiere A, Asselain B, Campana F, Scholl SM, Fenton J, Vilcoq JR, et al. Age as prognostic factor in premenopausal breast carcinoma. Lancet 1993 Apr 24;341(8852):1039-1043. [doi: 10.1016/0140-6736(93)92407-k] [Medline: 8096955]

Abbreviations

BC: breast cancer
BCSS: breast cancer–specific survival
DMFS: distant metastasis–free survival
ER: estrogen receptor
HER-2: human epidermal growth factor receptor 2
HR: hormone receptor
LRFS: locoregional recurrence–free survival
OS: overall survival
PR: progesterone receptor
TIS: tumor in situ

Edited by A Mavragani, T Sanchez; submitted 10.03.23; peer-reviewed by Z Zhu, Y Pang, CF Cai; comments to author 12.04.23; revised version received 19.04.23; accepted 15.06.23; published 18.07.23.

Please cite as:
Wang J, Luo T, Xiang ZZ, He MM, Zeng YY, Yang T, Wei XY, Yu S, Dai ZL, Xu NY, Tan CF, Liu L
Survival and Trends in Annualized Hazard Function by Age at Diagnosis Among Chinese Breast Cancer Patients Aged ≤40 Years: Case Analysis Study
JMIR Public Health Surveill 2023;9:e47110
URL: https://publichealth.jmir.org/2023/1/e47110
doi:10.2196/47110
PMID:37463020

Original Paper

Improving Surveillance of Human Tick-Borne Disease Risks: Spatial Analysis Using Multimodal Databases

Sarah P Maxwell[1], PhD; Chris Brooks[2], PhD; Dohyeong Kim[1], PhD; Connie L McNeely[3], PhD; Seonga Cho[4], MA; Kevin C Thomas[2], PhD

[1]School of Economic, Political and Policy Sciences, University of Texas at Dallas, Richardson, TX, United States
[2]Laboratory for Human Neurobiology, Boston University School of Medicine, Boston, MA, United States
[3]Schar School of Policy and Government, George Mason University, Fairfax, VA, United States
[4]Department of Geography, University of California, Santa Barbara, Santa Barbara, CA, United States

Corresponding Author:
Sarah P Maxwell, PhD
School of Economic, Political and Policy Sciences
University of Texas at Dallas
800 West Campbell Avenue
Richardson, TX, 75080
United States
Phone: 1 972 883 6469
Email: sarah.maxwell@utdallas.edu

Abstract

Background: The extent of tick-borne disease (TBD) risk in the United States is generally unknown. Active surveillance using entomological measures, such as presence and density of infected nymphal *Ixodes scapularis* ticks, have served as indicators for assessing human risk, but results have been inconsistent and passive surveillance via public health systems suggests TBDs are underreported.

Objective: Research using various data sources and collection methods (eg, Google Trends, apps, and tick bite encounters [TBEs] reports) has shown promise for assessing human TBD risk. In that vein, and engaging a One Health perspective, this study used multimodal databases, geographically overlaying patient survey data on TBEs and concomitant reports of TBDs with data drawn from other sources, such as canine serological reports, to glean insights and to determine and assess the use of various indicators as proxies for human TBD risk.

Methods: This study used a mixed methods research strategy, relying on triangulation techniques and drawing on multiple data sources to provide insights into various aspects of human disease risk from TBEs and TBDs in the United States. A web-based survey was conducted over a 15-month period beginning in December 2020 to collect data on TBEs. To maximize the value of the covariate data, related analyses included TBE reports that occurred in the United States between January 1, 2000, and March 31, 2021. TBEs among patients diagnosed with Lyme disease were analyzed at the county level and compared to *I scapularis* and *I pacificus* tick presence, human cases identified by the Centers for Disease Control and Prevention (CDC), and canine serological data. Spatial analyses employed multilayer thematic mapping and other techniques.

Results: After cleaning, survey results showed a total of 249 (75.7%) TBEs spread across 148 respondents (61.9% of all respondents, 81.7% of TBE-positive respondents); 144 (4.7%) counties in 30 states (60%) remained eligible for analysis, with an average of 1.68 (SD 1.00) and median of 1 (IQR 1) TBEs per respondent. Analysis revealed significant spatial matching at the county level among patient survey reports of TBEs and disease risk indicators from the CDC and other official sources. Thematic mapping results included one-for-one county-level matching of reported TBEs with at least 1 designated source of human disease risk (ie, positive canine serological tests, CDC-reported Lyme disease, or known tick presence).

Conclusions: Use of triangulation methods to integrate patient data on TBE recall with established canine serological reports, tick presence, and official human TBD information offers more granular, county-level information regarding TBD risk to inform clinicians and public health officials. Such data may supplement public health sources to offer improved surveillance and provide bases for developing robust proxies for TBD risk among humans.

(JMIR Public Health Surveill 2023;9:e43790) doi:10.2196/43790

KEYWORDS

tick-borne disease surveillance; Lyme disease; tick bite encounter; One Health model; triangulation; entomology; entomological; tick; thematic mapping; spatial; risk; surveillance; vector

Introduction

Tick-borne illness is steadily increasing across the United States. Lyme disease (LD), in particular, is the fastest growing vector-borne disease in the United States and accounts for the majority of all tick-borne diseases (TBDs) in the country [1,2]. Research using various data sources and collection methods (eg, Google trends, apps, and tick bite encounters' [TBEs'] reports) has shown promise for assessing human TBD risk [3-9]. However, the extent of TBD risk is relatively unknown. Entomological measures, such as presence and density of infected nymphal *Ixodes scapularis* ticks, have served as indicators for assessing human LD risk [10-13], but results have been inconsistent [11,14]. These measures do not consider human behavior [11]. Creating models that incorporate human behavior and adverse events—for example, TBE recall—can produce improved and more nuanced approaches for assessing human TBD risk. Internet and database searches on LD have demonstrated that information-seeking behaviors share similar temporal and spatial trends with known epidemiological reports [8,9]. Similar information seeking behaviors among health care providers also follow similar regional and temporal patterns to know human disease trends [10].

LD is transmitted by ticks in the *Ixodes* genus. *I scapularis*, the blacklegged tick, and *I pacificus*, the western blacklegged tick, are the primary vectors of the spirochete (*Borrelia burgdorferi*) that cause LD [10,15]. The blacklegged tick transmits other pathogens, including *Anaplasma phagocytophilum*, *Borrelia miyamotoi*, *Babesia microti*, *Powassan virus*, and ehrlichiosis associated with *Ehrlichia muris* eauclairensis [10,16]. *I pacificus* is known to transmit *Borrelia burgdorferi* sensu stricto, *Borrelia mayonii*, *Anaplasma phagocytophilum*, and *Borrelia miyamotoi* [16]. In addition, transmission of TBD pathogens to humans can result in related coinfections, such as Babesia, ehrlichiosis, or granulocytic anaplasmosis [16,17]. Recent modeling suggests the expansion of both *I scapularis* and *I pacificus* in areas where surveillance previously did not indicate reported or established tick populations [18]. The spread of *I scapularis* and *I pacificus* ticks with accompanying pathogens points to LD and other TBDs as a major public health risk [10], but, again, the actual prevalence of multiple TBDs or coinfections from related tick encounters is largely unknown, posing a problem for public health planning. Public health systems face significant challenges in identifying the spread of tick populations, as well as the associated risk and burden of TBDs among humans [19]. TBDs in this regard—accounting for 95% of all vector-borne diseases in the United States—are an especially serious and growing public health threat, with LD accounting for more than 70% of cases [20].

Veterinary studies that capture TBDs are shown to be useful in identifying possible coinfections and providing an important basis for human risk comparison [11], especially in areas where the Centers for Disease Control and Prevention (CDC) indicates TBD existence but where the actual number of cases may be unknown [12]. This study incorporates multimodal data triangulation using TBE, canine serological reports, public health data, and tick presence indicators to assess human disease risk among survey respondents reporting a TBE and concomitant diagnosis of a TBD.

Methods

Study Design

Various studies have introduced the use of web-based and other survey data along with other sources of patient information, on matters such as TBEs to demonstrate promising approaches and indicators for investigating human disease risk [3,4]. This study builds upon and extends recent research to include data on human TBEs at the county level and associated self-reports of TBDs in not only officially designated endemic regions but, importantly, also regions perceived as nonendemic in the United States relative to known tick and human and canine disease case reports.

The analytical objective in this study was to assess human TBD risk by employing a mixed methods approach using triangulation techniques and multimodal databases, geographically overlaying TBE reports to official data sources. Specifically, a web-based survey was conducted over a 15-month period beginning in December 2020 to collect data on TBEs. To maximize the value of the covariate data, related analyses included only TBE reports that occurred in the United States between January 1, 2000, and March 31, 2021. Adopting a One Health perspective, TBEs among patients diagnosed with LD were analyzed at the county level and compared to *I scapularis* and *I pacificus* tick presence, human cases identified by the CDC, and canine serological data. Spatial analyses included multilayer thematic mapping and statistical analyses, involving the amalgamated use of county-level patient TBEs, canine serological cases, tick presence, official CDC-reported human cases as an indicator for human TBD risk, and US Census data. Mapping and analyses of data conducted at the county level across the resultant data sets were supplemented by Fisher's exact tests of independence.

Ethical Considerations

The study protocol was approved by the institutional review board at the University of Texas at Dallas (IRB-21-149). Institutional review board approval was obtained prior to the survey, under formal adoption of the Declaration of Helsinki. All answers were completely anonymous and patients consented to voluntarily take the survey. All survey respondents provided informed consent. No personally associated identification data were collected or reported as such. Human subjects were not compensated. Data are available upon request.

Tick Bite Encounters and Respondent Data

The Tick Bite Encounter Survey (henceforth referred to as the "survey") was conducted on the internet via social media and shared through selected national TBD-related nonprofit organizational websites. The survey was available beginning in December 2020 for 15 months via an anonymous link (administered by Qualtrics). Designed to engage individuals diagnosed with a TBD, participation was voluntary, with respondents constituting a convenience sample for the study. Survey respondents were asked to report their diagnoses, tick-bite recall by month and year, and by county or zip code and state where tick bites occurred. Respondents listed county or zip code where and when they recalled receiving up to 4 tick bites or TBEs, and data were assessed to ensure entries matched corresponding states as reported. To maximize the value of the covariate data (eg, CDC-positive LD diagnoses), only TBEs that occurred in the United States between January 1, 2000, and March 31, 2021, were included in the analysis. In addition, TBEs were only analyzed if the location provided by the respondent could be unambiguously localized to a single zip code or county. Selected demographic information was also collected, including those who did not recall a TBE.

Additional study end points derived from official national databases were obtained from two main sources: (1) official by-county databases maintained by the CDC and US Census Bureau: (A) total number of human LD cases that met CDC diagnostic criteria and were recorded by state health officials to the CDC between 2000 and 2019; (B) counties officially established and reported by the CDC to contain *I scapularis* or *I pacificus* ticks as of 2020). The CDC notes that “counties classified as ‘no records’ should not be interpreted as the tick being absent. No records could be a result of a lack of sampling efforts, tick collections, or a lack of reporting or publishing the results of sampling efforts” [21]. Accordingly, for each county, tick presence was coded as present, absent, or unknown (no data); and (C) county populations as reported by the US Census Bureau’s Intercensal Estimates for 2000 to 2010 and 2010 to 2020. (2) Companion Animal Parasite Council county-level databases: total number of serological tests conducted on canines in 2020 (specifically: number of tests positive for ehrlichiosis; number of tests positive for anaplasmosis; number of tests positive for LD). The maps were presented as demonstrating county-level overlap of TBEs by each human disease risk indicator (ie, canine serological reports, CDC LD cases, and tick presence). In addition to thematic and multilevel mapping analyses using the ESRI’s ArcGIS (version 10.7), cross tabulations are provided across all data sources for all counties containing 1 or more survey TBEs, aligning with the thematic maps.

Results

Overview

Statistical analyses were conducted at the county level for all counties which had at least 1 TBE reported on the survey. Of the 239 respondents who completed the web-based survey, 182 (76.2%) reported at least 1 TBE, as shown in Table 1. A combined total of 329 TBEs were reported, with an average of 1.82 (SD 1.11) and a median of 1 (IQR 1) TBE per TBE-positive respondent. TBEs were filtered to restrict the sample to respondents who met the following criteria: the TBE could be definitively localized to a single county or zip code in a US state, and the TBE could be definitively localized to a single year between 2000 and 2021, inclusive. Note that some respondents indicated the location of TBEs using zip codes that occasionally encompassed areas crossing and belonging to different counties; in those cases, the TBE was considered to have occurred in the zip code’s “primary county” [22] and was considered “definitively localized” for analytical purposes. After applying these criteria, a total of 249 (75.7%) TBEs were apparent across 148 respondents (61.9% of all respondents, 81.8% of “TBE-positive” respondents), 144 (4.8%) counties in 30 states (60%) remained eligible for analysis, with an average of 1.68 (SD 1.00) and median of 1 (IQR 1) TBEs per respondent. County-level statistics appear in Table 2, with 144 eligible counties and their county-level data in the final data set. Unless otherwise noted, all further county-level proportions and percentages described herein were derived from this set of 144 counties; likewise, all state-level statistics were derived from the set of 30 states. None of the 144 counties or 30 states analyzed was missing survey TBE data, as only counties with one or more reported TBEs were included in the analysis. Of these, 14 (9.72%) counties across 7 (23.33%) states reported no CDC LD cases. These data include official LD reports from the CDC; patient reports of TBEs via web-based survey; established tick presence, indicating counties with habitats suitable for ticks; and canine serological data for Lyme, ehrlichiosis, and anaplasmosis.

Table 1. Survey respondents and tick bite recall (December 2020 to March 2021).

	Recall tick bite, n (%)	Do not recall tick bite, n (%)	All respondents, n (%)
Age			
9 to 45 years old	65 (66.3)	33 (33.7)	98 (41)
46 or more years old	116 (82.9)	24 (17.1)	140 (58.6)
Age not reported	1 (100)	0 (0)	1 (0.4)
Total	182 (76.2)	57 (23.9)	239 (100)
Number of years sick			
0 to 5 years	113 (77.4)	33 (22.6)	146 (61.1)
6 or more years	67 (75.3)	22 (24.7)	89 (37.2)
Not reported	2 (50)	2 (50)	4 (1.7)
Total	182 (76.2)	57 (23.6)	239 (100)

Table 2. Number of US counties with reported TBDs[a] by frequency of primary end points.

	Counties with disease indicator present, n (%)	Counties with no disease indicator present, n (%)	Data unavailable, n (%)	Total, n (%)
Human cases				
CDC[b] LD[c]	130 (90.3)	14 (9.7)	0 (0)	144 (100)
Survey TBE[d]	144 (100)	0 (0)	0 (0)	144 (100)
Tick presence				
Ixodes scapularis	110 (76.4)	34 (23.6)	0 (0)	144 (100)
I pacificus	2 (1.4)	142 (98.6)	0 (0)	144 (100)
Canine serological tests				
Canine any TBD	126 (87.5)	3 (2.1)	15 (10.4)	144 (100)
Canine LD	113 (78.5)	16 (11.1)	15 (10.4)	144 (100)
Canine anaplasmosis	106 (73.6)	23 (15.9)	15 (10.4)	144 (100)
Canine ehrlichiosis	122 (84.7)	7 (4.9)	15 (10.4)	144 (100)

[a] TBD: tick-borne disease.

[b] CDC: Centers for Disease Control and Prevention.

[c] LD: Lyme disease.

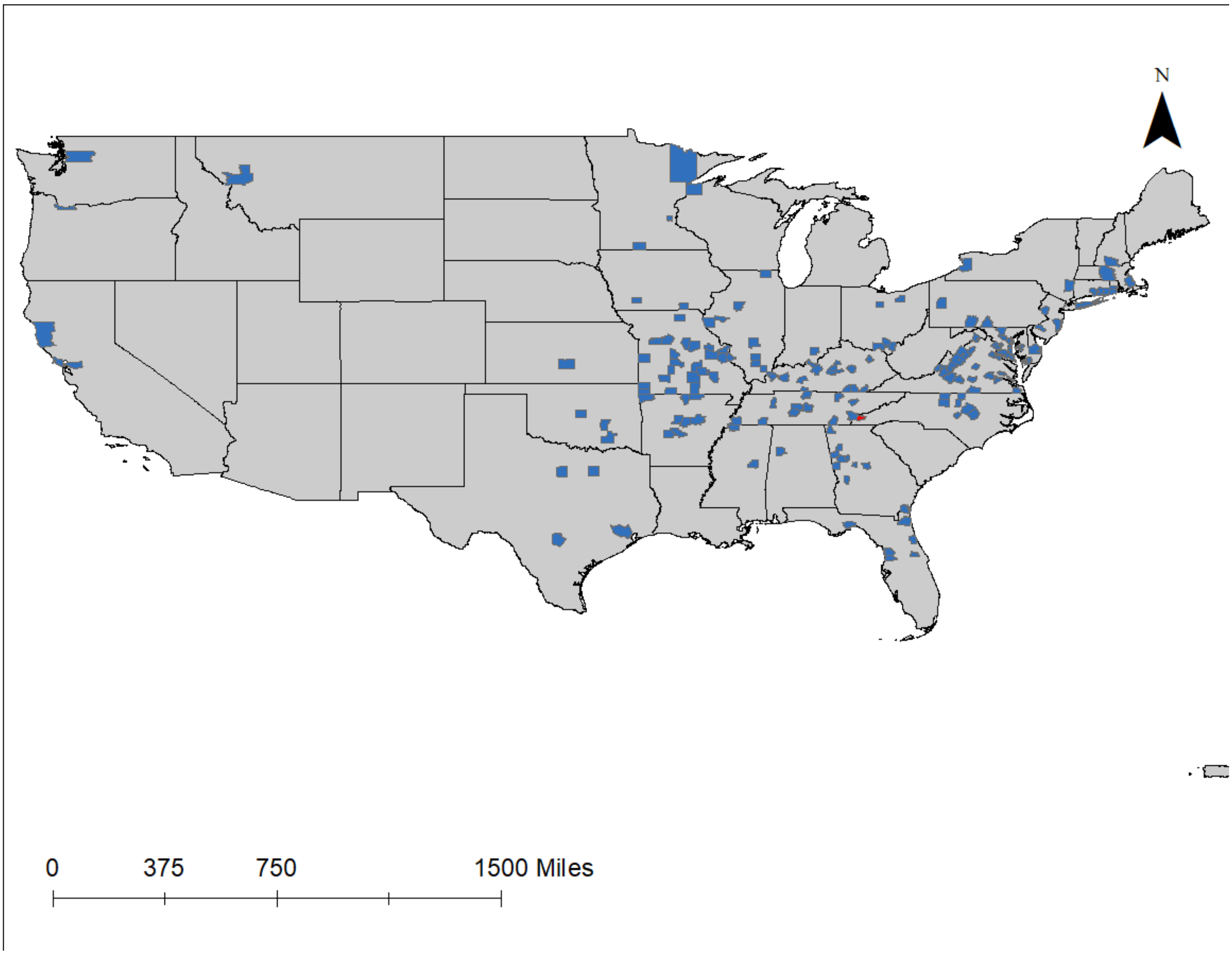
[d] TBE: tick bite encounter.

Self-Reported TBEs

Using triangulation methods, Figures 1-5 show tick presence and the existence of human and canine disease by county in the United States. Multilayered thematic maps of self-reported TBEs among patients diagnosed with LD and coinfections, in addition to statistical alignments and comparisons of data regarding TBE reports and recall, was developed and performed.

Taken together, the maps offer a representation of human TBD risk, based on the more nuanced approach engaging triangulation of human and canine reports in addition to tick presence. Survey-reported TBEs in 144 counties, as shown in Figure 1, are then compared to official disease and tick data. Official CDC data are also compared separately to demonstrate county level overlap via human and canine laboratory testing and tick presence.

Figure 1. Survey respondents with LD and reported tick bite encounters by county. LD: Lyme disease.



Mapping and analyses of subsequent county-level data also included Fisher's Exact tests of independence, as reported in Tables 3-6, to evaluate the likelihood that the co-occurrence or "county overlap" of given pairs of variables was random. A significant outcome ($P<.05$ level of significance) provides evidence that the 2 variables may be used as proxies for one other.

Table 3. Association across US counties by tick presence (significance set at $P<.05$).

Fisher exact test	*Ixodes pacificus* − Counties	*I pacificus* + Counties	Total
I scapularis – Counties	32	2	34
I scapularis + Counties	110	0	110
Total	142	2	144

Table 4. Association of US counties with survey respondent TBEs[a] by canine LD[b] and CDC[c] LD cases (significance set at $P<.05$).

Fisher exact test	County has reports of CDC LD	County has no reports of CDC LD	Total
County has reports of canine LD	5	11	16
County has no reports of canine LD	5	108	113
Data unavailable	4	11	15
Total	14	130	144

[a] TBE: tick bite encounter.

[b] LD: Lyme disease.

[c] CDC: Centers for Disease Control and Prevention.

Table 5. Association of US counties by survey respondent TBEs[a] based on canine anaplasmosis and CDC[b] LD[c] cases (significance set at *P*<.05).

Fisher exact test	County has reports of CDC LD	No county reports of CDC LD	Total
County has reports of canine anaplasmosis	5	18	23
County has no reports of canine anaplasmosis	5	101	106
Data unavailable	4	11	15
Total	14	130	144

[a] TBE: tick bite encounter.

[b] CDC: Centers for Disease Control and Prevention.

[c] LD: Lyme disease.

Table 6. Association of counties by survey respondent TBE[a] with canine ehrlichiosis and CDC[b] LD[c] Cases (significance set at *P*<.05).

Fisher exact test	County has CDC reports of LD	County has CDC reports of no LD	Total
County has reports of canine ehrlichiosis	3	4	7
County has reports of no canine ehrlichiosis	7	115	122
Data unavailable	4	11	15
Total	14	130	144
Fisher exact outcome	<0.05	<0.05	<0.05

[a] TBE: tick bite encounter.

[b] CDC: Centers for Disease Control and Prevention.

[c] LD: Lyme disease.

Human LD, Survey TBEs, and Canine LD

Figures 2-5 present integrated, multilayer thematic maps that show overlap of survey respondent TBEs with canine cases of LD, ehrlichiosis, and anaplasmosis, and human cases of LD (the only available county-level TBD data from the CDC). Figure 2 presents a thematic overlap of CDC and canine cases of LD within counties with at least 1 survey-reported TBE. Yellow counties indicate containing canine LD, CDC-reported LD cases, and where TBE survey respondents reported a TBE and concomitant disease. Red counties represent overlap among canine LD and TBE survey respondents, but without CDC reports of human LD. Green graduated shaded counties are all LD counties in the United States as reported by the CDC. Table 4 provides statistical data representing LD by county for respondent TBE, canine LD, and CDC LD. Some counties do not have canine data, so disease presence among dogs in those counties is unknown.

Figure 2. Respondent TBEs with canine and CDC LD cases by county (CDC LD cases 2000-2019). CDC: Centers for Disease Control and Prevention; LD: Lyme disease; TBE: tick bite encounter.

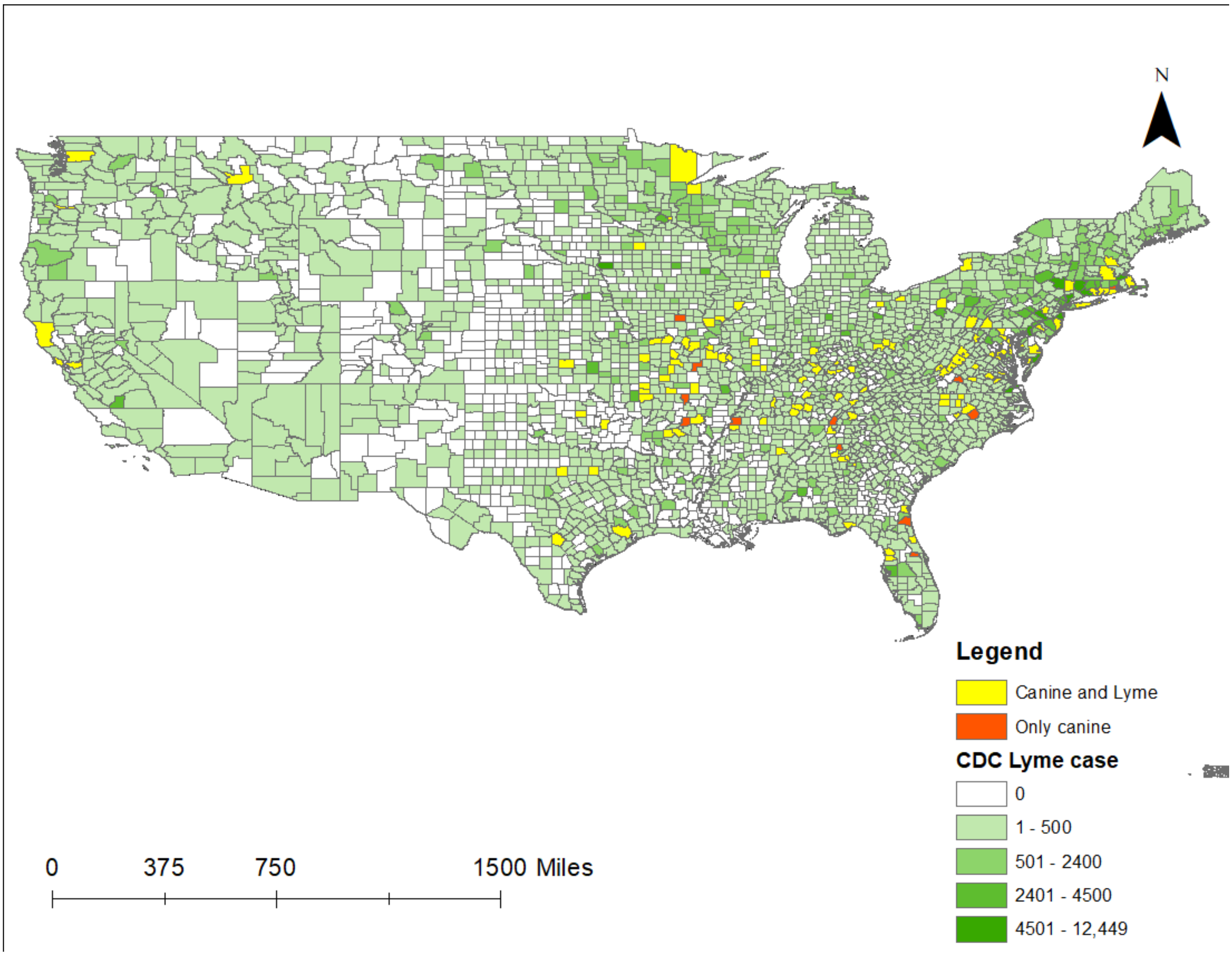


Figure 3 offers a map of CDC data and canine cases of anaplasmosis within TBE survey respondent counties. Blue counties indicate where canine cases overlap with CDC LD cases and survey-reported TBEs. Light blue counties represent overlap among canine LD and TBE survey respondents, but without CDC reports of human LD. Green counties are all LD counties with TBE survey respondent overlap. Table 5 highlights the thematic mapping overlaying, indicating significant association between canine anaplasmosis and CDC cases of LD within counties that had reports of a human TBE.

Figure 3. Respondent TBEs with canine anaplasmosis and CDC LD cases by county (CDC LD cases 2000-2019). CDC: Centers for Disease Control and Prevention; LD: Lyme disease; TBE: tick bite encounter.

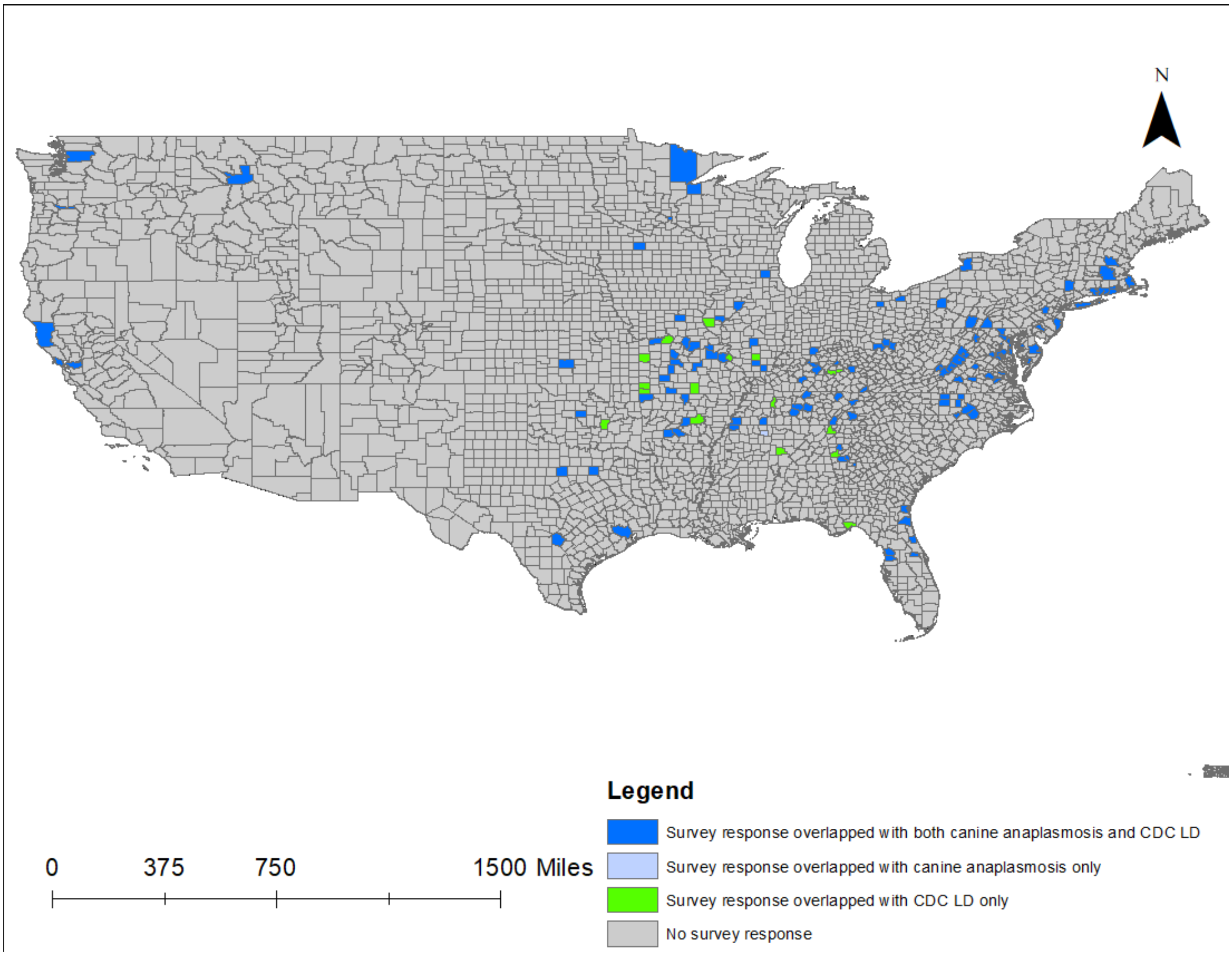
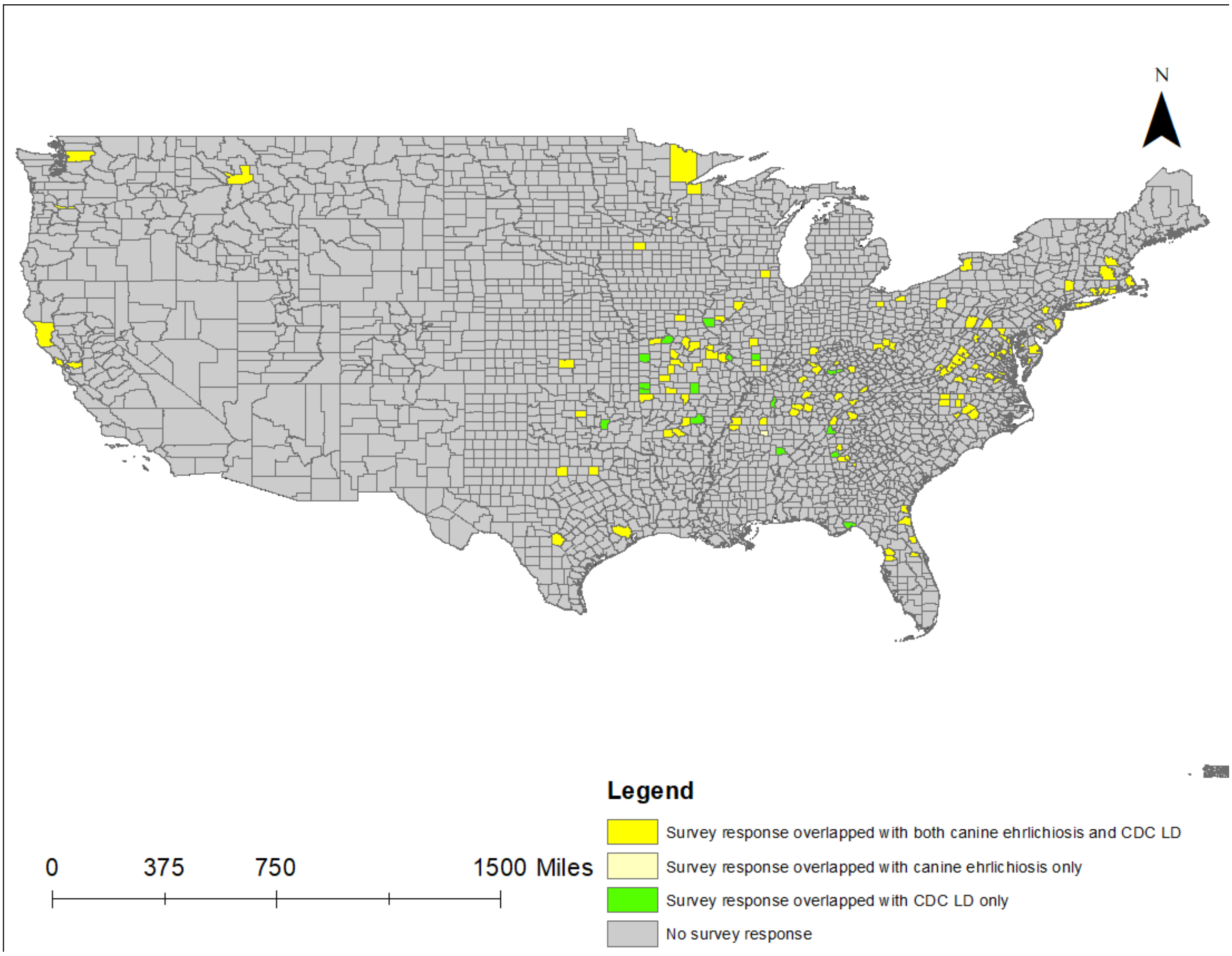


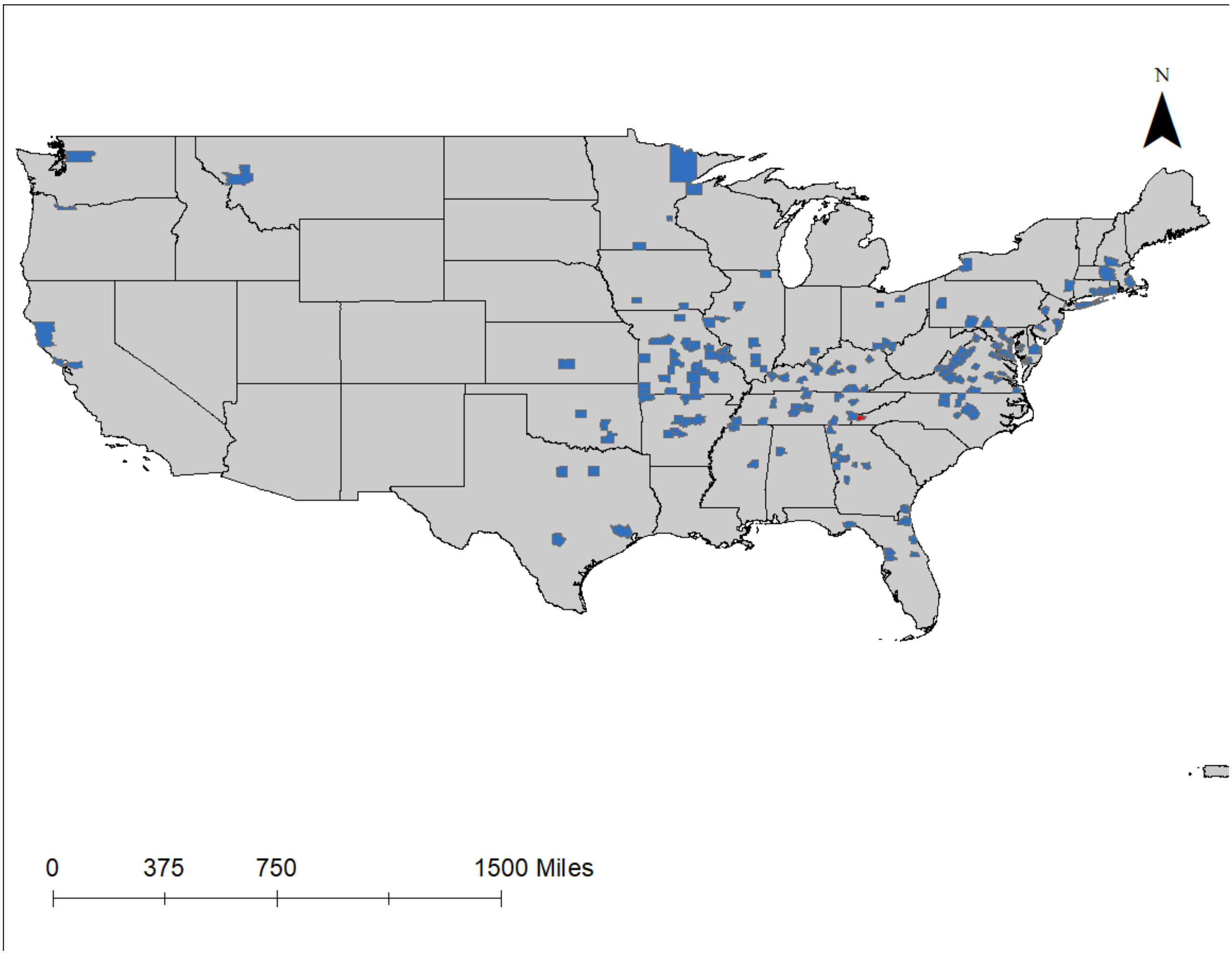
Figure 4 presents a thematic overlay of CDC and canine ehrlichiosis cases within TBE survey respondent counties. Yellow counties indicate where canine LD cases overlap with CDC LD cases and survey reported TBEs. Light yellow shading represents overlap between canine LD and TBE survey respondents, but without CDC reports of human LD. Green shading likewise represents counties with overlap between CDC LD and TBE survey data, but with no canine LD cases. Table 6 highlights the thematic mapping indicating significant associations between canine ehrlichiosis and CDC reported cases of LD within counties with human TBE reports.

Figure 4. Respondent TBEs with canine ehrlichiosis and CDC LD cases by county (CDC LD cases 2000-2019). CDC: Centers for Disease Control and Prevention; LD: Lyme disease; TBE: tick bite encounter.



In sum, the maps offer visual representation of TBE overlap with official data from 3 primary sources: tick presence, CDC-reported LD cases, and positive canine tests for LD, ehrlichiosis, and anaplasmosis. Using thematic overlays to triangulate matching human disease risk, every county overwhelmingly matched TBE survey reports and at least 1 disease risk indicator from official sources. Figure 5 is a summary map, depicting all counties where a survey respondent reported a TBE and a LD diagnosis and where the TBE matched at least one of the triangulated data points of tick presence or human or canine disease. The scope of influence is substantial in this respect; out of a total of 144 counties included in the final analytical data set, only 1 did not overlap with an official source of disease risk—Graham County, North Carolina—which appears as the only county in red on the summary map in Figure 5. In other words, the respondent TBEs and subsequent self-reported LD offer a one-to-one match with human or canine official serological reports. The lack of overlap in Graham County does not mean that disease risk does not exist but, rather, was the result of missing canine data. In fact, neighboring counties showed positive serological tests for LD, ehrlichiosis, and anaplasmosis in canines, and ticks do not adhere to county lines.

Figure 5. US counties with survey respondent reported TBEs and LD diagnoses, with TBEs matched to triangulated data points of tick presence or human or canine disease. LD: Lyme disease; TBE: tick bite encounter.



Discussion

Implications of the Relative Proportions of CDC LD Cases and Survey TBEs

To date, surveillance of TBDs typically is presented in aggregate national and state level databases, hindering analyses at levels that may be more revealing and provide more detailed depictions of relevant indications of TBD risk. The current public health system in the United States does not recognize as endemic areas that may cluster in various states outside of the Northeast. Data for most TBDs are provided by the CDC, aggregated from state public health departments following positive laboratory results. Although county-level data on TBDs are collected by some states, the CDC reports such data for LD only. In addition, while LD reports are available by county, no distinctions are drawn between locally as opposed to travel-acquired disease. Accordingly, TBE likelihood—which is geographically dependent—could not be confidently determined from such data alone.

However, multilayer thematic mapping has demonstrated clear overlapping among TBEs and concomitant human disease reports with canine and CDC LD reports and tick presence data by county. As noted, previous studies have demonstrated the utility of multilayer thematic mapping using canine and self-reported TBEs, particularly as it pertains to ecologically distinct regions [23]. Specifically, the survey sample, although relatively small, was national in scope. However, respondents reported TBEs only in western, mid-Atlantic, and eastern regions where LD and ticks are either established or considered endemic. Note that this could be an artifact and limitation of the data with the use of respondents as a convenience sample. In any case, a one-to-one match was found for TBE and human disease self-reports with at least one of the human disease risk indicators. The critical point here, as revealed through the mapping and the statistical analyses, is the substantial overlap in the TBE counties among canine (especially canine ehrlichiosis), CDC reported LD cases, and presence of ticks carrying associated pathogens.

Triangulating TBD Risk via Proxy Data

Developing robust proxies for human TBD risk to address data and testing limitations is an important analytical undertaking, especially until objective, standardized, and centralized diagnostic monitoring can be implemented for a more detailed picture below the state level as needed. Research has pointed to similar spatial distributions and county-level findings among canines and humans using official public health data, in addition to self-reports of TBEs and subsequent diagnoses (both clinical and CDC+) and tick infectivity [23-26]. The use of other data sources, including patient surveys, also points to other surveillance data and techniques that have the potential to inform public health practitioners when local data are scarce or clinical evaluations are made in geographic areas not generally considered endemic [27,28].

This study has explored paths for improving disease surveillance by triangulating TBD risk via proxy data. Proxies were considered to help establish risk, relying on “indirect data” such as surveys and veterinary assessments to supplement and extend official data when available. The use of canine TBDs as proxies was supported by research indicating, for example, links between companion pets and human disease. In fact, owners of dogs and cats are at increased risk of encountering ticks and of developing TBDs [29], which also reflects the One Health perspective given connections among humans and their pets. In other words, canine TBD data are not fully divested from human TBD data and, so, with proper consideration of their differences and precise characterization of their common factors, they can usefully serve as proxies, as demonstrated in this study.

Limitations and Recommendations for Future Work

The survey was intentionally aimed at individuals likely to have experienced TBEs and to have a high probability of participation and response (eg, web-based support groups focused on TBDs). This approach naturally sat the survey respondent sample apart from the general population and could lead to underestimation of TBEs. In addition to survey limitations and possible bias and simple recollection errors by respondents, some statistical issues can also be noted. In particular, the canine data represented cases only from 2020, rather than across the 2000-2021 study period. Future work will incorporate more nuanced multilevel models that can account for unequal variances, mixed and fixed effects, nesting, and other confounding issues. Additional data too will address broader contextual concerns and data set gaps. For example, other tick species (eg, the Lone Star Tick) can act as vectors for TBDs and are extant in different geographical locations. Inclusion of such entomological data can improve analytical coverage. In addition, expanded historical data will help elucidate long-term trends that may not be detectable over shorter periods and will aid in outcome interpretation.

Conclusions

Controversy arises as a burgeoning population reports TBD diagnoses that do not match official reports. LD and other TBDs can be difficult to diagnose and physicians and health care professionals may be unaware of growing risks and public health threats in locally clustered areas, particularly in areas generally considered nonendemic [23]. Diagnostic ambiguity is exacerbated by symptomatic overlap in TBD presentation, as well as insufficient tracking and characterization of the various vectors and TBEs that cause them.

To date, surveillance of TBDs in the United States is a patchwork of activities that are unable to provide a satisfactory picture of risk or infection rates across TBDs, not to mention overlapping comorbidities in infected humans. Resulting controversies plague the patient-physician-public health relationship and discourse, with LD and other TBD sufferers turning to advocacy in demand for diagnosis and treatment. The importance of identifying opportunities to improve surveillance and data for determining TBD risk cannot be overstated for informing the public and the medical community.

Acknowledgments

The authors wish to thank the hundreds of respondents who completed the web-based survey and whose reporting made this work possible.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

SPM, CB, KCT, CLM, DK, and SC conceptualized the study. CB, SC, DK, KCT, and SPM designed the study methodology. CB, DK, SC, SPM, and KCT validated the findings. CB, SC, and DK carried out formal analysis. SPM, CB, SC, DK, CLM, and KCT carried out the investigation. SPM and CB acquired the resources for the study. CB and SC curated the data. SPM, CB, CLM, DK, SC, and KCT drafted the manuscript. SPM, CB, CLM, SC, and DK reviewed and edited the manuscript. DK, SC, SPM, and CB visualized the data. SPM, KCT, and DK supervised the study. SPM was in charge of the study’s administration. All authors have read and agreed to the published version of this paper.

Conflicts of Interest

None declared.

References

1. Rodino KG, Theel ES, Pritt BS. Tick-borne diseases in the United States. Clin Chem 2020;66(4):537-548 [FREE Full text] [doi: 10.1093/clinchem/hvaa040] [Medline: 32232463]
2. Preventing tick bites. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/lyme/prev/index.html [accessed 2021-09-10]

3. Scheerer C, Rüth M, Tizek L, Köberle M, Biedermann T, Zink A. Googling for ticks and Borreliosis in Germany: nationwide google search analysis from 2015 to 2018. J Med Internet Res 2020;22(10):e18581 [FREE Full text] [doi: 10.2196/18581] [Medline: 33064086]
4. Beaujean D, van Velsen L, van Gemert-Pijnen JE, Maat A, van Steenbergen JE, Crutzen R. Using risk group profiles as a lightweight qualitative approach for intervention development: an example of prevention of tick bites and lyme disease. JMIR Res Protoc 2013;2(2):e45 [FREE Full text] [doi: 10.2196/resprot.2760] [Medline: 24172875]
5. Pesälä S, Virtanen MJ, Sane J, Jousimaa J, Lyytikäinen O, Murtopuro S, et al. Health care professionals' evidence-based medicine internet searches closely Mimic the known seasonal variation of Lyme Borreliosis: a register-based study. JMIR Public Health Surveill 2017;3(2):e19 [FREE Full text] [doi: 10.2196/publichealth.6764] [Medline: 28400357]
6. Fernandez MP, Bron GM, Kache PA, Larson SR, Maus A, Gustafson D, et al. Usability and feasibility of a smartphone app to assess human behavioral factors associated with tick exposure (The Tick App): quantitative and qualitative study. JMIR Mhealth Uhealth 2019;7(10):e14769 [FREE Full text] [doi: 10.2196/14769] [Medline: 31651409]
7. Lami F, Asi W, Khistawi A, Jawad I. Syndromic surveillance of communicable diseases in mobile clinics during the Arbaeenia Mass gathering in Wassit Governorate, Iraq, in 2014: cross-sectional study. JMIR Public Health Surveill 2019;5(4):e10920 [FREE Full text] [doi: 10.2196/10920] [Medline: 31593544]
8. Pesälä S, Virtanen MJ, Sane J, Mustonen P, Kaila M, Helve O. Health information-seeking patterns of the general public and indications for disease surveillance: register-based study using lyme disease. JMIR Public Health Surveill 2017;3(4):e86 [FREE Full text] [doi: 10.2196/publichealth.8306] [Medline: 29109071]
9. Gillis T, Garrison S. Confounding effect of undergraduate semester-driven "Academic" internet searches on the ability to detect true disease seasonality in google trends data: fourier filter method development and demonstration. JMIR Infodemiology 2022;2(2):e34464 [FREE Full text] [doi: 10.2196/34464] [Medline: 37113451]
10. Eisen RJ, Eisen L. The blacklegged tick, Ixodes scapularis: an increasing public health concern. Trends Parasitol 2018;34(4):295-309. [doi: 10.1016/j.pt.2017.12.006] [Medline: 29336985]
11. Hook SA, Nawrocki CC, Meek JI, Feldman KA, White JL, Connally NP, et al. Human-tick encounters as a measure of tickborne disease risk in lyme disease endemic areas. Zoonoses Public Health 2021;68(5):384-392. [doi: 10.1111/zph.12810] [Medline: 33554467]
12. Feldman KA, Connally NP, Hojgaard A, Jones EH, White JL, Hinckley AF. Abundance and infection rates of Ixodes scapularis nymphs collected from residential properties in Lyme disease-endemic areas of Connecticut, Maryland, and New York. J Vector Ecol 2015;40(1):198-201 [FREE Full text] [doi: 10.1111/jvec.12153] [Medline: 26047204]
13. Eisen L, Dolan MC. Evidence for personal protective measures to reduce human contact with blacklegged ticks and for environmentally based control methods to suppress host-seeking blacklegged ticks and reduce infection with Lyme disease spirochetes in tick vectors and rodent reservoirs. J Med Entomol 2016;53(5):1063-1092 [FREE Full text] [doi: 10.1093/jme/tjw103] [Medline: 27439616]
14. Pepin KM, Eisen RJ, Mead PS, Piesman J, Fish D, Hoen AG, et al. Geographic variation in the relationship between human Lyme disease incidence and density of infected host-seeking Ixodes scapularis nymphs in the Eastern United States. Am J Trop Med Hyg 2012;86(6):1062-1071 [FREE Full text] [doi: 10.4269/ajtmh.2012.11-0630] [Medline: 22665620]
15. Eisen RJ, Eisen L, Beard CB. County-scale distribution of Ixodes scapularis and Ixodes pacificus (Acari: Ixodidae) in the Continental United States. J Med Entomol 2016;53(2):349-386 [FREE Full text] [doi: 10.1093/jme/tjv237] [Medline: 26783367]
16. Tickborne diseases of the United States. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ticks/tickbornediseases/index.html [accessed 2023-07-20]
17. Kocan KM, de la Fuente J, Coburn LA. Insights into the development of Ixodes scapularis: a resource for research on a medically important tick species. Parasit Vectors 2015;8:592 [FREE Full text] [doi: 10.1186/s13071-015-1185-7] [Medline: 26576940]
18. Hahn MB, Jarnevich CS, Monaghan AJ, Eisen RJ. Modeling the geographic distribution of Ixodes scapularis and Ixodes pacificus (Acari: Ixodidae) in the contiguous United States. J Med Entomol 2016;53:1176-1191 [FREE Full text] [doi: 10.1186/s13071-015-1185-7] [Medline: 26576940]
19. Madison-Antenucci S, Kramer LD, Gebhardt LL, Kauffman E. Emerging tick-borne diseases. Clin Microbiol Rev 2020;33(2):e00083-18 [FREE Full text] [doi: 10.1128/CMR.00083-18] [Medline: 31896541]
20. Adams DA, Thomas KR, Jajosky RA, Foster L, Sharp P, Onweh DH, Nationally Notifiable Infectious Conditions Group. Summary of notifiable infectious diseases and conditions—United States, 2014. MMWR Morb Mortal Wkly Rep 2016;63(54):1-152 [FREE Full text] [doi: 10.15585/mmwr.mm6354a1] [Medline: 27736829]
21. Tick surveillance data sets. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ticks/surveillance/TickSurveillanceData.html [accessed 2023-05-29]
22. Zip Data Maps. URL: https://www.zipdatamaps.com [accessed 2021-10-15]
23. Maxwell SP, McNeely CL, Thomas K, Brooks C. Tick-borne surveillance patterns in perceived non-endemic geographic areas: human tick encounters and disease outcomes. Healthcare (Basel) 2021;9(6):771 [FREE Full text] [doi: 10.3390/healthcare9060771] [Medline: 34205506]

24. Maxwell SP, Brooks C, Kim P, Kim D, McNeely CL, Thomas K. Understanding habitats and environmental conditions of white-tailed deer population density and public health data to aid in assessing human tick-borne disease risk. Microorganisms 2023;11(4):865 [FREE Full text] [doi: 10.3390/microorganisms11040865] [Medline: 37110288]
25. Brooks C, McNeely CL, Maxwell SP, Thomas KC. Assessing tick-borne disease risk and surveillance: toward a multi-modal approach to diagnostic positioning and prediction. Microorganisms 2022;10(4):832 [FREE Full text] [doi: 10.3390/microorganisms10040832] [Medline: 35456881]
26. Thomas K, Brooks C, McNeely CL, Maxwell SP. Tick-borne disease risk and exposure among vulnerable populations in perceived non-endemic regions. Zoonotic Dis 2022;2(3):111-116 [FREE Full text] [doi: 10.3390/zoonoticdis2030011]
27. Kim D, Maxwell S, Le Q. Spatial and temporal comparison of perceived risks and confirmed cases of Lyme disease: an exploratory study of google trends. Front Public Health 2020;8:395 [FREE Full text] [doi: 10.3389/fpubh.2020.00395] [Medline: 32923420]
28. Maxwell SP. The elusive understanding of Lyme disease in non-endemic geographic areas: an exploratory survey of patients with chronic symptoms in Texas. J Patient Exp 2020;7(6):1621-1626 [FREE Full text] [doi: 10.1177/2374373520926821] [Medline: 33457622]
29. Jones EH, Hinckley AF, Hook SA, Meek JI, Backenson B, Kugeler KJ, et al. Pet ownership increases human risk of encountering ticks. Zoonoses Public Health 2018;65(1):74-79 [FREE Full text] [doi: 10.1111/zph.12369] [Medline: 28631423]

Abbreviations

CDC: Centers for Disease Control and Prevention
LD: Lyme disease
TBD: tick-borne disease
TBE: tick bite encounter

Edited by A Mavragani, T Sanchez; submitted 24.10.22; peer-reviewed by S Pesälä, C Sun; comments to author 11.04.23; revised version received 24.04.23; accepted 27.06.23; published 23.08.23.

Please cite as:
Maxwell SP, Brooks C, Kim D, McNeely CL, Cho S, Thomas KC
Improving Surveillance of Human Tick-Borne Disease Risks: Spatial Analysis Using Multimodal Databases
JMIR Public Health Surveill 2023;9:e43790
URL: https://publichealth.jmir.org/2023/1/e43790
doi:10.2196/43790
PMID:37610812

Original Paper

The Environmental and Socioeconomic Effects and Prediction of Patients With Tuberculosis in Different Age Groups in Southwest China: A Population-Based Study

Wen Wei[1], PhD; Lan Xia[2], MSc; Jianlin Wu[2], BSc; Zonglei Zhou[1], PhD; Wenqiang Zhang[1], PhD; Rongsheng Luan[1], MSc

[1]Department of Epidemiology and Biostatistics, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

[2]Department of Tuberculosis, Center for Disease Control and Prevention of Sichuan Province, Chengdu, China

Corresponding Author:
Rongsheng Luan, MSc
Department of Epidemiology and Biostatistics
West China School of Public Health and West China Fourth Hospital
Sichuan University
16#, Section 3, Renmin Nan Lu
Chengdu, 610041
China
Phone: 86 028 85501604
Fax: 86 028 85501604
Email: luan_rs@scu.edu.cn

Abstract

Background: While the End Tuberculosis (TB) Strategy has been implemented worldwide, the cause of the TB epidemic is multifactorial and not fully understood.

Objective: This study aims to investigate the risk factors of TB and incorporate these factors to forecast the incidence of TB infection across different age groups in Sichuan, China.

Methods: Correlation and linear regression analyses were conducted to assess the relationships between TB cases and ecological factors, including environmental, economic, and social factors, in Sichuan Province from 2006 to 2017. The transfer function-noise model was used to forecast trends, considering both time and multifactor effects.

Results: From 2006 to 2017, Sichuan Province had a reported cumulative incidence rate of 1321.08 cases per 100,000 individuals in male patients and 583.04 cases per 100,000 individuals in female patients. There were significant sex differences in the distribution of cases among age groups (trend χ^2_{25}=12,544.4; *P*<.001). Ganzi Tibetan Autonomous Prefecture had the highest incidence rates of TB in both male and female patients in Sichuan. Correlation and regression analyses showed that the total illiteracy rate and average pressure at each measuring station (for individuals aged 15-24 years) were risk factors for TB. The protective factors were as follows: the number of families with the minimum living standard guarantee in urban areas, the average wind speed, the number of discharged patients with invasive TB, the number of people with the minimum living standard guarantee in rural areas, the total health expenditure as a percentage of regional gross domestic product, and being a single male individual (for those aged 0-14 years); the number of hospitals and number of health workers in infectious disease hospitals (for individuals aged 25-64 years); and the amount of daily morning and evening exercise, the number of people with the urban minimum living standard guarantee, and being married (for female individuals aged ≥65 years). The transfer function-noise model indicated that the incidence of TB in male patients aged 0-14 and 15-24 years will continue to increase, and the incidence of TB in female patients aged 0-14 and ≥65 years will continue to increase rapidly in Sichuan by 2035.

Conclusions: The End TB Strategy in Sichuan should consider environmental, educational, medical, social, personal, and other conditions, and further substantial efforts are needed especially for male patients aged 0-24 years, female patients aged 0-14 years, and female patients older than 64 years.

(JMIR Public Health Surveill 2023;9:e40659) doi:10.2196/40659

KEYWORDS

tuberculosis; risk factors; age; sex; prediction; TB control; tuberculosis control

Introduction

Tuberculosis (TB) has been the leading cause of the global disease burden. Approximately 25% of the population worldwide are infected with *M tuberculosis* [1]. In 2019, according to the World Health Organization, there were approximately 10 million new cases of TB globally and 1.2 million deaths [2]. Of note, almost 90% of individuals who are infected with TB each year originate from low-income countries [2]. Poor health services, malnutrition, and crowded working and living conditions in these countries leads to the increased risk of TB across populations. Fighting poverty has become a major theme for the World Health Organization, which aims to end the global TB epidemic using “the End TB Strategy” [2]. According to the End TB Strategy, the targeted percentage reduction in the absolute number of TB deaths and incidence rate for 2035 are 95% and 90% of the 2015 baseline, respectively [1]. However, the cause of the TB epidemic is multifactorial and not fully understood.

Most people (approximately 90%) develop the disease in adulthood, with men being more susceptible than women [1-4]. Previous ecological studies have demonstrated a significant association between TB cases and ecological factors, including environmental, economic, health, and social conditions [3-11]. Although ecological risk factors for TB at the population and individual levels have raised great concerns, these results are not consistent [1,2,11-16]. The cause of TB varies in different populations and countries, especially in different groups (eg, of different ages and sexes) or in areas with a high TB burden [1], which is a major barrier for the End TB Strategy. China had the third-largest TB burden in the world in 2019, while Sichuan Province, known for its geographical and ethnic diversity, has a top-ranked TB burden, thus providing the opportunity to comprehensively identify specific risk factors in a complex background.

Furthermore, elucidating the trend of TB incidence with identified TB factors helps to assess the effectiveness of containing measurements that may aid policy maker decisions and public health practice. Therefore, our research aimed to identify these potential risk factors for TB in Southwest China using multiple regression and to predict the trend of TB incidence using a transfer function-noise (TFN) model [17-21]. The overarching goal of our study was to gain insight into the secular TB trend, providing implications for advancing TB prevention and control strategies to achieve the targets of the End TB Strategy.

Methods

Data Sources and Objects

Geographic Information

We obtained data from China’s Geographic Information Center on Sichuan Province in 2009. The map of county-level administrative divisions included prefectures, counties, cities, and districts. In total, Sichuan Province has 21 cities or prefectures and 181 districts, counties, and cities, with a total of 1.4 million inhabitants.

Social, Economic, Environmental, Education, and Health Information

Data were extracted from the 2006-2017 Statistical Yearbook of Sichuan Provincial Bureau of Statistics. All TB cases were grouped by sex and age. The age groups were as follows: children (aged 0-14 years), youths (aged 15-24 years), adults (aged 25-64 years), and older individuals (aged >64 years).

From the age-stratified population of the “Epidemic Information Network Direct Reporting System” from 2006 to 2017, we enrolled the subpopulation in Sichuan Province. All population data were for permanent residents, specific to county administrative divisions, and included population data by age and sex. Information on TB and HIV/AIDS was obtained from the TB Information Management System of the Chinese Disease Prevention and Control Information System and the Statistical Yearbook published by the Sichuan Provincial Bureau of Statistics. The incidence of TB was analyzed in different age groups.

Statistical Analysis

The ecological analysis used data such as case reports, registered case data, and ecological information to explore the risk factors related to the prevalence of TB. Data on TB cases and the incidence rate of TB were collected per administrative division, sex, and age group. The variables for each model were chosen by auto-modeling.

The 128 ecological factors were all obtained from 12 years (2006-2017) of data. The transfer function model was fitted by the autoregressive integrated moving average (ARIMA) model by time series [17-21], using expert modeling and adding independent variables for the model fitted to choose the variables through the fitted data and prediction data; we used 3 different models (the Grey model, the ARIMA model, the TFN model) to identify the model that most closely aligned with the real TB data of 2018.

TFN Model

The expert modeler selected the optimal model from multiple fitted models, when the R^2 value reached the ideal state. The linear regression analysis adopts the stepwise regression analysis method and uses multiple models to fit and to achieve a better R^2 value in four different age groups. Finally, only the best model was used for display. Multiple stepwise regression was carried out for multivariate analysis (stepwise regression rules: F-to-enter≥3.840, F-to-remove≤2.710). The models were established by each factor, stepwise in and out of the models. Multimedia Appendix 1 shows the stepwise regression summary; when the R^2 and standard estimated error reached the best value, the model concluded.

Univariate analyses (Pearson correlation analysis) and multivariate analyses (regression analysis) were used to analyze the protective factors and risk factors for TB in Sichuan Province according to sex and age group (trend χ^2_{25}=12,544.4; $P<.001$). Statistical analysis (descriptive analysis and cluster analysis of spatiotemporal scans) and prediction of TB incidence in 2035 were performed according to sex and age.

The autocorrelation test of the residuals uses the Durbin-Watson (DW) test, with the following test statistic:

DW values occur in a range from 0 to 4 as follows: a DW value of 0 indicates complete positive autocorrelation, values between 0 and 1.5 indicate positive autocorrelation, values between 1.5 and 2.5 indicate no autocorrelation, values between 2.5 and 4 indicate negative autocorrelation, and a value of 4 indicates complete negative autocorrelation. In addition, the autocorrelation function and partial autocorrelation function show whether the data sequence reached a stable state. At the same time, the R^2 and Bayesian information criterion values of the TFN model were used before and after data unit standardization to find the best model. The closer the DW value is to 2, the more independent the observations of multiple linear regression are.

Considering the effect of time and multiple factors, we used Panel regression, Poisson regression, and Lasso regression for the analysis; however, the data were not suitable for these regressions; for example, our study has a large number of factors to explore the relationship, so the Panel regression could not include all 128 factors and ID and time to fit. The pilot analyses revealed that the random effect model was better than the pool model and fixed effect model. For these reasons, we did not use other regression methods.

We included all reported cases of TB during 2006-2017; these data possibly contain information and selection bias. The TFN model is a multivariate time series analysis method that can be seen as a combination of the ARIMA model and a multiple regression model. We used SPSS 23.0 (IBM Corp) and ArcGis Map 10.6 (ESRI Inc) to create spatiotemporal scans and to conduct the analyses that predicted TB trends by sex and age. The three main steps were as follows: model identification, parameter estimation, and model testing. For the calculation of the *P* value, the analysis was applied under the assumptions of unequal variances and a statistical significance of $P<.05$.

Ethics Approval

Data collection of TB was required by the Law of the People's Republic of China on Prevention and Treatment of Infectious Diseases. The ethics approval in this study was granted by the Ethics Committee of Sichuan Center for Disease Control and Prevention (SCCDCIRB2022-001).

Results

Annual TB Cases and Incidence Rate

From 2006 to 2017, Sichuan Province reported 548,584 cases of pulmonary TB in male patients, with a reported cumulative incidence rate of 1321.08 cases per 100,000 individuals (average 110.09 cases/100,000 individuals), and 235,149 cases of pulmonary TB in female patients, with a reported cumulative incidence rate of 583.04 cases per 100,000 individuals (average 48.59 cases/100,000 individuals). Thus, there were approximately 2.33 times more cases in male patients than in female patients. The reported cumulative incidence of TB in Sichuan Province from 2006 to 2017 was 961.71 cases per 100,000 individuals (average 80.14 cases/100,000 individuals). These TB cases mainly occurred in individuals aged 15-64 years, which accounted for 82.02% (n=642,808) of the total cases.

As shown in Figure 1, there were sex differences in the distribution of cases among age groups, and these differences were significant (trend χ^2_{25}=12,544.4; $P<.001$). The number of TB cases in Sichuan Province peaked in individuals aged 20-24 years and gradually decreased in individuals older than 64 years (Figure 1). During these 12 years (2006-2017), the incidence rate in those aged 80-85 years was lower than in those aged 60-79 years, while individuals older than 70 years had the highest TB incidence peak (Figure 2).

Figure 1. TB cases in Sichuan Province of China during 2006-2017. TB: tuberculosis.

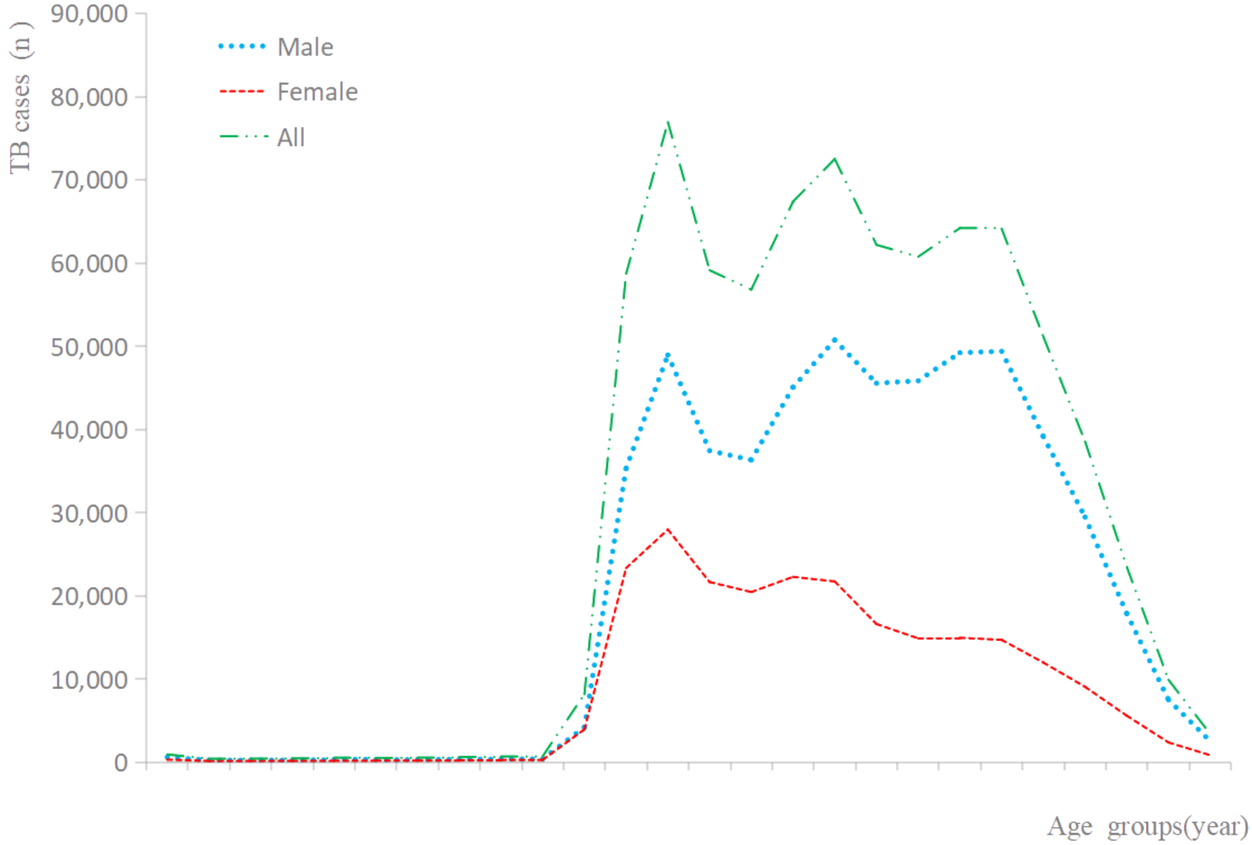
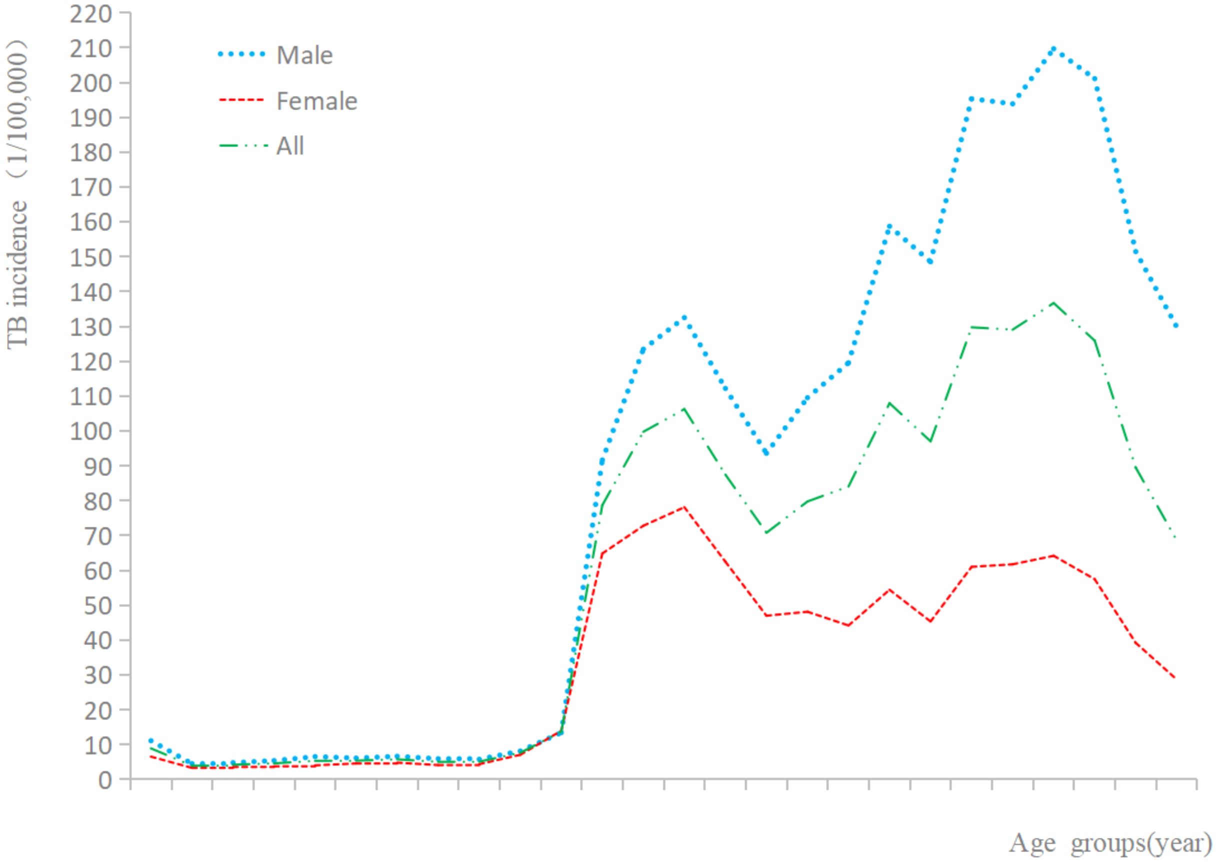


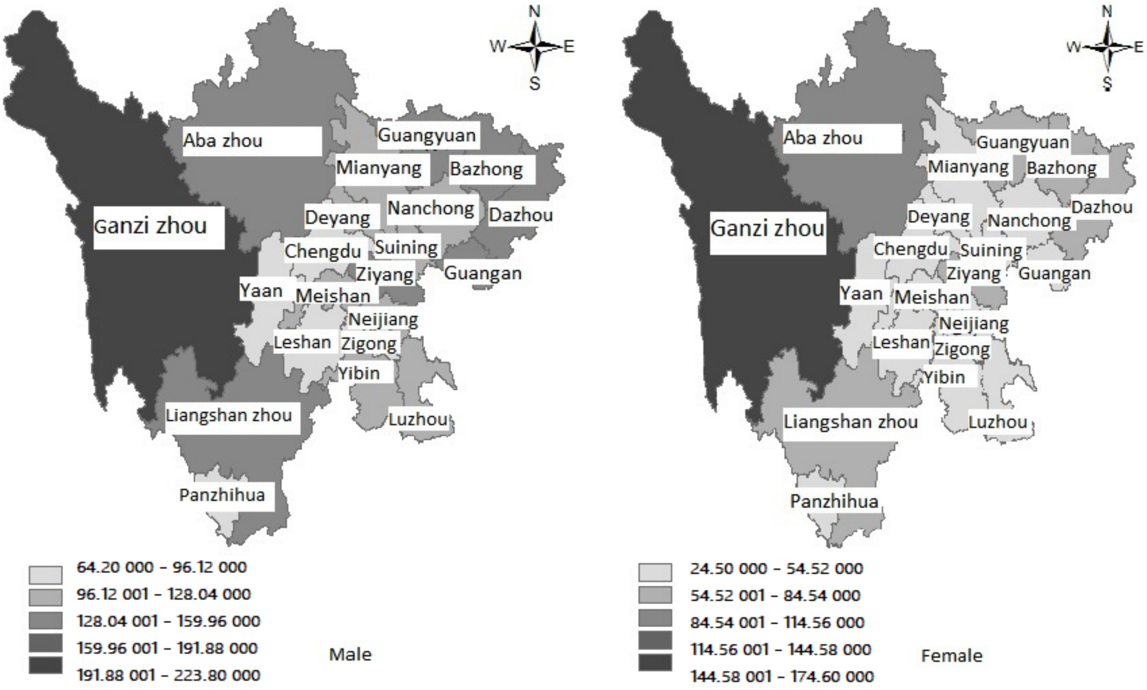
Figure 2. TB incidence rate in Sichuan Province of China during 2006-2017. TB: tuberculosis.



Geographic Distribution

Figure 3 displays the TB cumulative incidence in 21 cities/prefectures during this time period. The 5 cities/prefectures with the highest incidence rates of TB in male patients were Ganzi Tibetan Autonomous Prefecture, Aba Tibetan and Qiang Autonomous Prefecture, Guangyuan City, Liangshan Yi Autonomous Prefecture, and Dazhou. The five cities/prefectures with the lowest incidence rates of TB in male patients were Chengdu, Panzhihua, Ya’an, Leshan, and Zigong. In female patients, the top five cities/prefectures according to TB incidence rates were Ganzi Tibetan Autonomous Prefecture, Aba Tibetan and Qiang Autonomous Prefecture, Liangshan Yi Autonomous Prefecture, Guangyuan City, and Dazhou; the bottom five were Chengdu, Zigong, Ya’an, Leshan, and Panzhihua.

Figure 3. The incidence of male and female TB in Sichuan Province of China during 2006-2017(1/100,000). TB: tuberculosis.



Relationships Among the Variables

Across the four age groups, there were significant associations between TB incidence and the environmental and socioeconomic effects in five aspects, including other diseases as well as health resources; economic and social factors; pollutant, forestry, and meteorology variables; population structure, marriage, and habit variables; and educational investment and education level (all *P*<.05; Multimedia Appendix 1, Tables S1-S5).

Multivariate Analysis

As shown in Table 1, a linear regression analysis was performed on the incidence of TB and the various factors, and a multivariate analysis was performed on the four age groups in the whole population, the male population, and the female population. The regression model exhibited multivariate multicollinearity; the DW values for the four age groups (individuals aged 0-14, 15-24, 25-64, or ≥65 years) were 2.433, 1.340, 2.491, and 3.114, respectively.

In addition, in those aged 0-14 years, the factors affecting the risk of TB were as follows: the number of families with the urban minimum living standard guarantee, average wind speed, number of patients with invasive pulmonary TB discharged, number of people with the rural minimum living standard guarantee, total health expenditure as a percentage of regional gross domestic product (GDP), and being a single man. These factors reduced the risk of TB. The incidence of TB in those aged 15-24 years, sex ratio, number of forest fires, and regular exercise rate of adults increased the risk of TB in children (Table 1).

In those aged 15-24 years, the incidence of TB in those aged 0-14 years, total illiteracy rate, number of infectious disease hospitals, enrollment number (secondary school and above), exposure to secondhand smoke among nonsmokers (%), fraction of married men (%), number of forest fires, and number of junior high school boys (%) increased the risk of TB (Table 1).

In those aged 25-64 years, the number of hospitals, number of health workers in infectious disease hospitals, TB rate in those aged ≥65 years, daylight duration, and male life expectancy decreased the risk of TB (Table 1).

Finally, in the group of people older than 64 years, the risk of TB was increased by the following factors: the rate of TB cases in female patients, cumulative precipitation over 24 hours (8 PM to 8 PM), rate of TB in those aged 25-64 years, total health expenditure as a percentage of regional GDP, number of families with the rural minimum living standard guarantee, rate of TB in patients aged 15-24 years, and civil medical assistance. The risk of TB decreased with the amount of daily morning and evening exercise and the number of people with the urban minimum living standard guarantee and in women who were married (Table 1).

Table 1. Multivariate analysis of tuberculosis (TB) incidence by age group (individuals aged 0-14, 15-24, 25-64, or ≥65 years) in Sichuan Province.

Model and variables	B (95% CI)	*T* test (*df*)	*P* value
Model 1 (dependent variable: TB incidence in those aged 0-14 years; Durbin-Watson value: 2.433)			
(constant)	–8.259 (–8.269 to –8.249)	–10626.577 (11)	<.001
TB rate in those aged 15-24 years	1.01E-01 (0.101 to 0.101)	14772.916 (11)	<.001
Sex ratio	8.387 (8.379 to 8.394)	14578.003 (11)	<.001
Number of forest fires	0.001 (0.001 to 0.001)	3593.154 (11)	<.001
Number of families with the minimum living standard guarantee in urban areas (households)	–7.438E-07 (0 to 0)	–2011.933 (11)	<.001
Average wind speed	–0.005 (–0.005 to –0.005)	–1576.859 (11)	<.001
Number of patients with invasive TB discharged	–1.505E-05 (0 to 0)	–4259.893 (11)	<.001
Number of people with the minimum living standard guarantee in rural areas (person)	–1.41E-07 (0 to 0)	–1813.338 (11)	<.001
Regular exercise rate	1.82 (1.8 to 1.839)	1199.617 (11)	.001
Total health expenditure as a percentage of regional GDP[a] (%)	–0.007 (–0.008 to –0.007)	–377.144 (11)	.002
Unmarried men (%)	–0.001 (–0.001 to –0.001)	–105.253 (11)	.006
Model 2 (dependent variable: TB incidence in those aged 15-24 years; Durbin-Watson value: 1.340)			
(constant)	–2541.142 (–2542.913 to –2539.371)	–18234.07 (11)	<.001
TB rate in those aged 0-14 years	6.227 (6.226 to 6.228)	82051.414 (11)	<.001
Total illiteracy rate (%)	1.789 (1.788 to 1.789)	27895.184 (11)	<.001
Average pressure at each measuring station	0.307 (0.307 to 0.307)	18170.576 (11)	<.001
Number of infectious disease hospitals	0.854 (0.853 to 0.854)	20744.736 (11)	<.001
Enrollment number (secondary school and above)	0.000004396 (0 to 0)	13489.017 (11)	<.001
Exposure to secondhand smoke among nonsmokers (%)	0.094 (0.094 to 0.094)	11336.041 (11)	<.001
Male illiteracy rate (%)	–0.617 (–0.618 to –0.615)	–5377.636 (11)	<.001
Married men (%)	0.083 (0.083 to 0.083)	3963.885 (11)	<.001
Number of forest fires	0.001 (0.001 to 0.001)	895.709 (11)	.001
Junior high school boys (%)	0.004 (0.004 to 0.005)	128.503 (11)	.005
Model 3 (dependent variable: TB incidence in those aged 25-64 years; Durbin-Watson value: 2.491)			
(constant)	10.268 (10.218 to 10.317)	2642.64 (11)	<.001
TB rate	1.123 (1.123 to 1.123)	99793.924 (11)	<.001
Number of hospitals	–0.009 (–0.009 to –0.009)	–54131.67 (11)	<.001
Number of health workers in infectious disease hospitals	–0.006 (–0.006 to –0.006)	–32021.619 (11)	<.001
TB rate in those aged >64 years	–0.048 (–0.048 to –0.048)	–26994.198 (11)	<.001
Daylight duration	–0.178 (–0.178 to –0.178)	–14248.119 (11)	<.001
Life expectancy	0.421 (0.418 to 0.424)	1887.559 (11)	<.001
Diabetes mortality rate (1/100,000)	0.117 (0.116 to 0.117)	4809.176 (11)	<.001
Total expenditure per person in urban residents (¥)	0.00002666 (0 to 0)	915.623 (11)	.001
Number of public health workers per 1000 people	0.062 (0.06 to 0.065)	356.061 (11)	.002
Male life expectancy	–0.032 (–0.034 to –0.029)	–171.099 (11)	.004
Model 4 (dependent variable: TB incidence in those aged >64 years; Durbin-Watson value: 3.114)			
(constant)	–28.128 (–28.41 to –27.846)	–1266.037 (11)	.001
TB rate in women	2.98E+00 (2.979 to 2.98)	70646.596 (11)	<.001
Amount of daily morning and evening exercise	–5.24E-05 (0 to 0)	–29473.402 (11)	<.001

Model and variables	B (95% CI)	*T* test (*df*)	*P* value
Cumulative precipitation in 24 hours (8 PM to 8 PM)	0.985 (0.985 to 0.985)	44529.735 (11)	<.001
TB rate in those aged 25-64 years	1.391 (1.39 to 1.392)	15700.111 (11)	<.001
Number of people with the urban minimum living standard guarantee (person)	–0.00001256 (0 to 0)	–3962.971 (11)	<.001
Total health expenditure as a percentage of regional GDP (%)	1.51 (1.506 to 1.513)	5632.666 (11)	<.001
Number of families with the rural minimum living standard guarantee (households)	0.000001464 (0 to 0)	1182.888 (11)	.001
TB rate in women aged 15-24 years	0.132 (0.13 to 0.133)	1161.821 (11)	.001
Married women (%)	–0.192 (–0.196 to –0.189)	–676.252 (11)	.001
Civil medical assistance (times used)	1.04E-09 (0 to 0)	44.526 (11)	.01

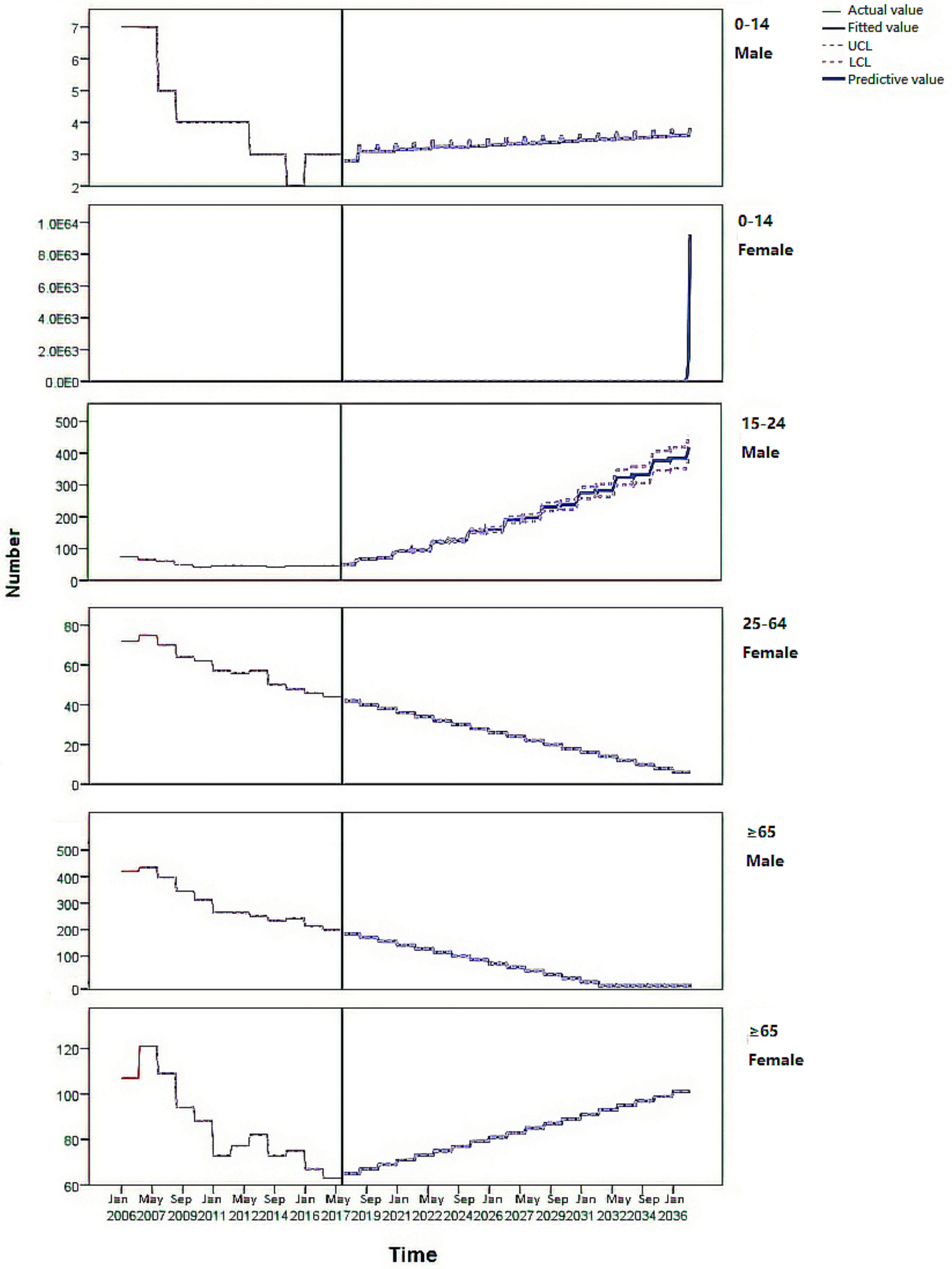
[a] GDP: gross domestic product.

Forecast Trend

In addition, the TFN model predicted that in 2035, the incidence of TB in male patients aged 0-14, 15-24, 25-64, and ≥65 years per 100,000 individuals would be 4.00, 377.08, 0.00, and –54.00, respectively (Figure 4). Our model predicted that the incidence of TB in males aged 0-14, 15-24, 25-64, and ≥65 years would not fall to 0 by 2035 and that the incidence rate of TB in male patients aged 0-14 and 15-24 years would increase.

Moreover, the TFN results showed that the incidence of TB among female patients aged 0-14, 15-24, 25-64, and ≥65 years per 100,000 individuals in 2018 was 3.00, unpredictable, 42.00, and 65.00, respectively. However, in 2035, these rates were predicted to increase, and TB incidence among female patients aged 0-14 years was predicted to be higher than among women aged ≥65 years, which might also exhibit an upward trend (Figure 4).

Figure 4. The transfer function-noise model predicts the trend of age-specific incidence rates (male/female: aged 0-14, 15-24, 25-64, and ≥65 years). LCL: lower confidence limit; UCL: upper confidence limit.



Model Selection

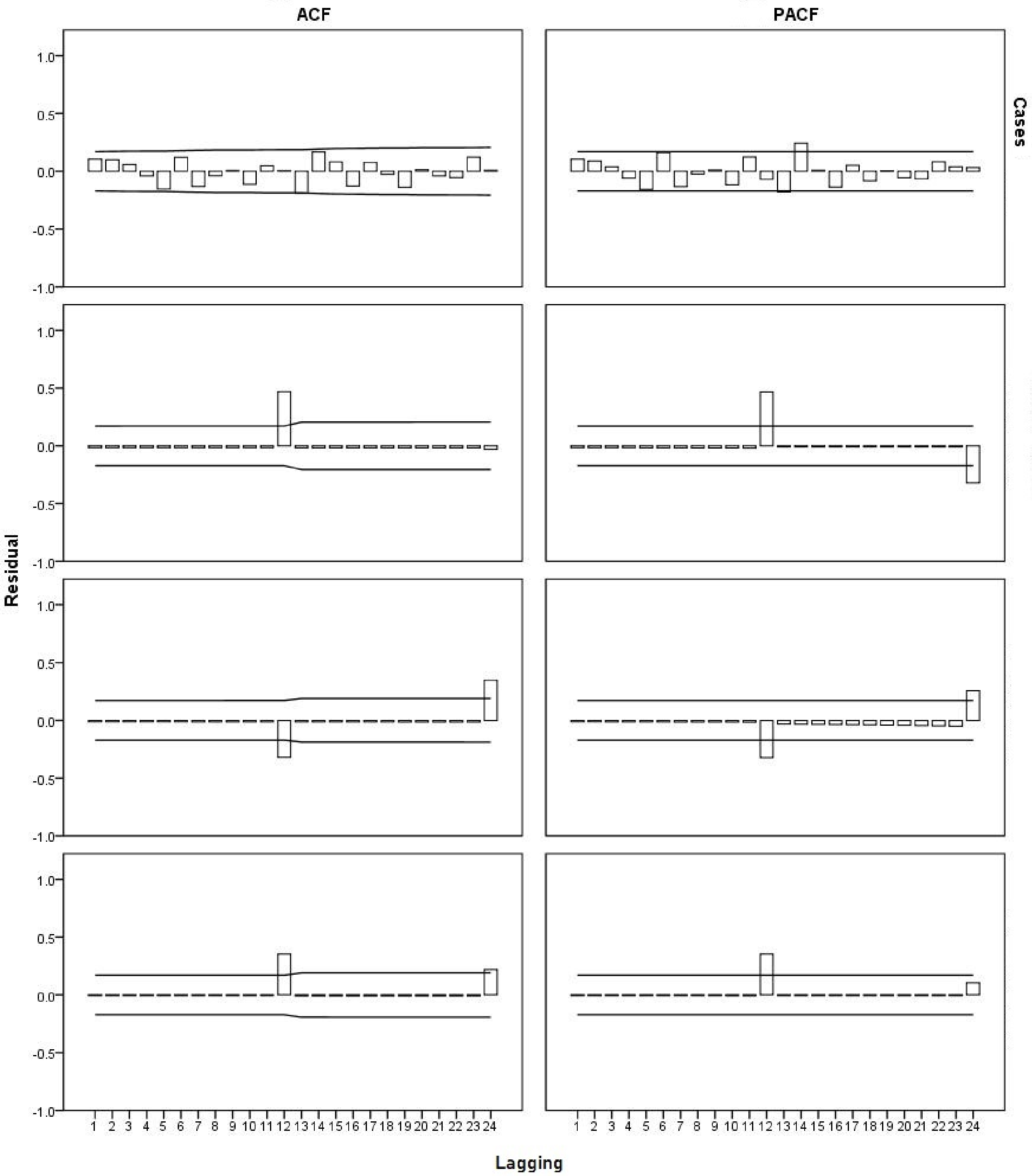
The TFN model was used to fit the number and incidence rate of TB cases reported in Sichuan Province during 2006-2017, establish and judge the models, and predict TB incidence in Sichuan Province during 2018-2035. The TFN model was fitted for different groups to determine the TB incidence. The R^2 and Bayesian information criterion of the models and evaluations were all ideal, as shown in Table S6 in Multimedia Appendix 1.

The TFN models used female and male data separately in the multivariate models, and the total of all cases was 783,735. From 2006 to 2017, Sichuan Province reported 548,585 cases of pulmonary TB in male patients, with a reported incidence rate of 11 cases per 100,000 individuals, and 235,150 cases of pulmonary TB in female patients, with a reported cumulative incidence rate of 4.85 cases per 100,000 individuals. There were 12-year observations per model at the population level. For each model, using stepwise regression methods, 10 well-fitting models were established (a better R^2 value), and the rule was carried out for multivariate analysis (stepwise regression rules: F-to-enter≥3.840, F-to-remove≤2.710).

Additionally, we adjusted the variables of the data to be standardizing variables, and the result did not change. The findings also show that the data trend is relatively stable after multiple data processing, such as splitting or logarithmic processing. Both the destandardized and standardized results and indicators showed that the models were the same.

The TFN model, number and incidence rate of TB cases, autocorrelation function, and partial autocorrelation function showed that the data sequence reached a stable state, and the effect of white noise was eliminated (Figure 5).

Figure 5. Transfer function-noise model: tuberculosis and incidence rate of ACF and PACF in the whole population. ACF: autocorrelation function; PACF: partial autocorrelation function.



We found that the total illiteracy rate and average pressure at each measuring station (in individuals aged 15-24 years) were risk factors for TB incidence overall. Factors that protected against TB incidence were the number of families with the minimum living standard guarantee in urban areas, average wind speed, number of patients with invasive TB discharged, number of people with the minimum living standard guarantee in rural areas, total health expenditure as a percentage of regional GDP, being a single man (aged 0-14), number of hospitals, number of health workers in infectious disease hospitals (for those aged 25-64 years), amount of daily morning and evening exercises, number of people with the urban minimum living standard guarantee, and being a married woman (aged ≥65 years). The noise transfer-function model predicts that until 2035, the incidence of TB in male patients aged 0-14 and 15-24 years will increase, while in female patients aged 0-14 and ≥65 years, the incidence will increase rapidly.

Discussion

Principal Findings

To our knowledge, this is the most comprehensive ecological correlation and time series trend forecast analyses to investigate the associations of ecological factors with TB epidemic and the secular trend of TB incidence. The ecological correlation analyses demonstrated an overall limited influence of the environmental and socioeconomic effects and prediction on the TB epidemic. Incorporating these environmental and socioeconomic conditions into time series trend forecast models found that the TB incidence rate will continue to increase among male patients aged 0-24 years, female patients aged 0-14 years, and female patients older than 64 years, which should be considered as targeted populations to end the TB epidemic before 2035.

Compared with the worldwide rate and those for high-income countries, TB incidence in rural areas of Sichuan Province was very high, which was supported by a significant increase risk of TB infection among farmers compared with individuals with other occupations and in clustering areas of ethnic minorities in western and northeastern Sichuan [12]. Given that most people reside in rural areas worldwide, more effective strategies are needed to contain the TB epidemic.

Additionally, in China, the highest TB burden was found in those aged >70 years, the same as our findings that they had the peak TB incidence [22]; nevertheless, their incidence of this disease will not grow fast. Meanwhile, male patients aged 0-24 years, female patients aged 0-14 years, and female patients older than 64 years will have substantial growth by 2035 and take a heavy TB burden. Because of health problems and risk factors of chronic disease such as hypoimmunity, high systolic blood pressure, and high total cholesterol, with greater longevity, may predispose older adults to TB [22,23]. Apart from social aging and the above reasons, further national surveys to estimate the latent reasons for the increasing incidence of TB in female patients are required.

Of note, the risk factors for TB incidence varied across different ages. For children aged 0-14 years, we found that the proportion reflected by the total health expenditure divided by the regional GDP was associated with a decrease in the incidence of TB, while an inverse association was observed in the population older than 64 years. In-depth investigations need to reveal its role in TB incidence. In addition, a decrease in the number of forest fires led to a decrease in TB incidence. This finding was in line with those of Chen et al [11], who reported an association between decreased per capita living space and TB incidence. Forest fires may lead to air pollution, causing more patients with TB to seek diagnosis and treatment. Chen et al [11] suggested that future studies should consider social factors such as income and education. Furthermore, we found that in individuals aged 15-24 years, the total illiteracy rate can predict the incidence of TB.

For female individuals older than 64 years, daily morning and evening exercise, the number of people with the urban minimum living standard guarantee, and marriage (for women) may reduce the incidence of TB. Married women who do not require assistance for extreme poverty may have better nutrition and better health well-being [1,2], while the urban minimum living standard guarantees and the aid of a spouse provides support to seek TB treatment, especially for older persons (aged ≥65 years). Collectively, social and environmental supports also help control TB incidence in Sichuan.

Moreover, our findings did not support the protective role of neonatal BCG vaccination in TB incidence across all age groups. Notably, delivering the neonatal BCG vaccination displayed a negative correlation with TB incidence across all age groups, principally in line with previous findings [1,2]. Given that BCG vaccines offer the best chance to contain the accelerating spread of multidrug-resistant TB, more data are required to draw definitive conclusions on its role in TB incidence.

Limitations

This study also has several limitations. First, we did not take into account determinations of genetic predisposition, treatments, treatment efficacy, and other ecological factors due to limited data availability. Second, all data in this study was obtained from Sichuan Province, and the results were subject to the inherent limitations of an ecological study; thus, principal findings may not be generalizable. Prospectively, further long-term cohort studies with large sample sizes as well as randomized controlled experimental studies are needed to determine the causal relationships.

Conclusion

In conclusion, our study identifies a series of environmental and socioeconomic conditions associated with TB incidence through the most comprehensive ecological analyses. Incorporating these TB factors into the TFN model clarifies that male individuals aged 0-24 years, female individuals aged 0-14 years, and female individuals older than 64 years are barriers to the End TB Strategy in Sichuan, and provides new insights into TB prevention and control strategies.

Acknowledgments

We appreciated Professor Jianwei Zhao and Professor Chuang Chen for help with the study design and writing, and Professor Shujuan Yang, Professor Fangfang Zeng, and Dr Sui Zhu for revising the manuscript.

Data Availability

All data used in the analysis of this manuscript are freely available by contacting the corresponding author. Data are available in a public open access repository, available on reasonable request.

Authors' Contributions

RL and WW designed the study and searched the publications. WW and LX analyzed the data and wrote the Methods section. JW wrote the Results section and checked all the data. ZZ and WZ wrote the Discussion section. All authors were responsible for data synthesis and have read and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material.
[DOC File , 410 KB - publichealth_v9i1e40659_app1.doc]

References

1. Global tuberculosis report 2021. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240037021 [accessed 2021-10-14]
2. Global tuberculosis report 2020. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240013131 [accessed 2020-10-15]
3. Efsen AMW, Panteleev AM, Grint D, Podlekareva DN, Vassilenko A, Rakhmanova A, HIV/TB Study Group. TB meningitis in HIV-positive patients in Europe and Argentina: clinical outcome and factors associated with mortality. Biomed Res Int 2013;2013:373601. [doi: 10.1155/2013/373601] [Medline: 24699884]
4. Thwaites GE, Duc Bang N, Huy Dung N, Thi Quy H, Thi Tuong Oanh D, Thi Cam Thoa N, et al. The influence of HIV infection on clinical presentation, response to treatment, and outcome in adults with Tuberculous meningitis. J Infect Dis 2005 Dec 15;192(12):2134-2141. [doi: 10.1086/498220] [Medline: 16288379]
5. Vinnard C, Macgregor RR. Tuberculous meningitis in HIV-infected individuals. Curr HIV/AIDS Rep 2009 Aug;6(3):139-145 [FREE Full text] [doi: 10.1007/s11904-009-0019-7] [Medline: 19589299]
6. Thwaites GE, Nguyen DB, Nguyen HD, Hoang TQ, Do TTO, Nguyen TCT, et al. Dexamethasone for the treatment of tuberculous meningitis in adolescents and adults. N Engl J Med 2004 Oct 21;351(17):1741-1751. [doi: 10.1056/NEJMoa040573] [Medline: 15496623]
7. Katrak SM, Shembalkar PK, Bijwe SR, Bhandarkar LD. The clinical, radiological and pathological profile of tuberculous meningitis in patients with and without human immunodeficiency virus infection. J Neurol Sci 2000 Dec 01;181(1-2):118-126. [doi: 10.1016/s0022-510x(00)00440-8] [Medline: 11099721]
8. Berenguer J, Moreno S, Laguna F, Vicente T, Adrados M, Ortega A, et al. Tuberculous meningitis in patients infected with the human immunodeficiency virus. N Engl J Med 1992 Mar 05;326(10):668-672. [doi: 10.1056/NEJM199203053261004] [Medline: 1346547]
9. Thwaites G, Chau T, Stepniewska K, Phu N, Chuong L, Sinh D, et al. Diagnosis of adult tuberculous meningitis by use of clinical and laboratory features. Lancet 2002 Oct 26;360(9342):1287-1292. [doi: 10.1016/s0140-6736(02)11318-3] [Medline: 12414204]
10. Cecchini D, Ambrosioni J, Brezzo C, Corti M, Rybko A, Perez M, et al. Tuberculous meningitis in HIV-infected and non-infected patients: comparison of cerebrospinal fluid findings. Int J Tuberc Lung Dis 2009 Feb;13(2):269-271. [Medline: 19146759]
11. Chen W, Shu W, Wang M, Hou Y, Xia Y, Xu W, et al. Pulmonary tuberculosis incidence and risk factors in rural areas of China: a cohort study. PLoS One 2013;8(3):e58171 [FREE Full text] [doi: 10.1371/journal.pone.0058171] [Medline: 23554875]
12. Xia Y, Li T, Wang DX, He JG, Zhang PR. Tuberculosis epidemic status in Sichuan Province during 2004 and 2014 and trend prediction. J Prev Med Inf 2015;12(1):946-949.
13. Xia L, Zhu S, Chen C, Rao Z, Xia Y, Wang D, et al. Spatio-temporal analysis of socio-economic characteristics for pulmonary tuberculosis in Sichuan province of China, 2006-2015. BMC Infect Dis 2020 Jun 22;20(1):433 [FREE Full text] [doi: 10.1186/s12879-020-05150-z] [Medline: 32571231]

14. Im C, Kim Y. Spatial pattern of tuberculosis (TB) and related socio-environmental factors in South Korea, 2008-2016. PLoS One 2021;16(8):e0255727 [FREE Full text] [doi: 10.1371/journal.pone.0255727] [Medline: 34352032]
15. Pawlowski A, Jansson M, Sköld M, Rottenberg ME, Källenius G. Tuberculosis and HIV co-infection. PLoS Pathog 2012 Feb;8(2):e1002464 [FREE Full text] [doi: 10.1371/journal.ppat.1002464] [Medline: 22363214]
16. Park J, Kwon J, Kim T, Heo J. Future inflow simulation considering the uncertainties of TFN Model and GCMs on Chungju Dam Basin. J Korea Water Resources Assoc 2014 Feb 28;47(2):135-143. [doi: 10.3741/jkwra.2014.47.2.135]
17. Yi M, Lee K. Transfer function-noise modelling of irregularly observed groundwater heads using precipitation data. J Hydrology 2004 Mar;288(3-4):272-287. [doi: 10.1016/j.jhydrol.2003.10.020]
18. Liu Y, Zhu J, Liu X, Li X. Application of transfer function-noise model in the prediction of child mortality rate under age 5 in China. Chin J Obs/Gyne Pediatr 2008;4(2):25-27.
19. Liwei J, Tong W, Weiqi W. The transfer function model and its application in hospital business income. Modern Prev Med 2007;34(20):3858-3564.
20. Bierkens MFP, Knotters M, van Geer FC. Calibration of transfer function-noise models to sparsely or irregularly observed time series. Water Resources Res 2010 Jul 09;35(6):1741-1750. [doi: 10.1029/1999wr900083]
21. Chen X, Dong X, Wang S, Yang Z, Li M, Wang Y. Time series analysis of reported incidence of tuberculosis in Guangzhou from 1997 to 2009. Chin J Health Stat 2011;28(5):548-552.
22. Zhu S, Xia L, Yu S, Chen S, Zhang J. The burden and challenges of tuberculosis in China: findings from the Global Burden of Disease Study 2015. Sci Rep 2017 Nov 06;7(1):14601. [doi: 10.1038/s41598-017-15024-1] [Medline: 29097809]
23. China Statistical Yearbook. National Bureau of Statistics of China. 2016. URL: http://www.stats.gov.cn/tjsj/ndsj/2016/indexch.htm [accessed 2022-12-19]

Abbreviations

ARIMA: autoregressive integrated moving average
DW: Durbin-Watson
GDP: gross domestic product
TB: tuberculosis
TFN: transfer function-noise

Edited by A Mavragani; submitted 01.07.22; peer-reviewed by T Li, S Sarejloo, M Wambura; comments to author 12.09.22; revised version received 04.10.22; accepted 01.12.22; published 13.01.23.

Please cite as:
Wei W, Xia L, Wu J, Zhou Z, Zhang W, Luan R
The Environmental and Socioeconomic Effects and Prediction of Patients With Tuberculosis in Different Age Groups in Southwest China: A Population-Based Study
JMIR Public Health Surveill 2023;9:e40659
URL: https://publichealth.jmir.org/2023/1/e40659
doi:10.2196/40659
PMID:36456535

Original Paper

Decreased Seasonal Influenza Rates Detected in a Crowdsourced Influenza-Like Illness Surveillance System During the COVID-19 Pandemic: Prospective Cohort Study

Autumn Gertz[1], MSc; Benjamin Rader[1,2], MPH; Kara Sewalk[1], MPH; Tanner J Varrelman[1], PhD; Mark Smolinski[3], MPH, MD; John S Brownstein[1,4], PhD

[1]Computational Epidemiology Lab, Boston Children's Hospital, Boston, MA, United States
[2]Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States
[3]Ending Pandemics, San Francisco, CA, United States
[4]Harvard Medical School, Boston, MA, United States

Corresponding Author:
John S Brownstein, PhD
Computational Epidemiology Lab
Boston Children's Hospital
300 Longwood Avenue
Boston, MA, 02115
United States
Phone: 1 8572185188
Email: info@outbreaksnearme.org

Abstract

Background: Seasonal respiratory viruses had lower incidence during their 2019-2020 and 2020-2021 seasons, which overlapped with the COVID-19 pandemic. The widespread implementation of precautionary measures to prevent transmission of SARS-CoV-2 has been seen to also mitigate transmission of seasonal influenza. The COVID-19 pandemic also led to changes in care seeking and access. Participatory surveillance systems have historically captured mild illnesses that are often missed by surveillance systems that rely on encounters with a health care provider for detection.

Objective: This study aimed to assess if a crowdsourced syndromic surveillance system capable of detecting mild influenza-like illness (ILI) also captured the globally observed decrease in ILI in the 2019-2020 and 2020-2021 influenza seasons, concurrent with the COVID-19 pandemic.

Methods: Flu Near You (FNY) is a web-based participatory syndromic surveillance system that allows participants in the United States to report their health information using a brief weekly survey. Reminder emails are sent to registered FNY participants to report on their symptoms and the symptoms of household members. Guest participants may also report. ILI was defined as fever and sore throat or fever and cough. ILI rates were determined as the number of ILI reports over the total number of reports and assessed for the 2016-2017, 2017-2018, 2018-2019, 2019-2020, and 2020-2021 influenza seasons. Baseline season (2016-2017, 2017-2018, and 2018-2019) rates were compared to the 2019-2020 and 2020-2021 influenza seasons. Self-reported influenza diagnosis and vaccination status were captured and assessed as the total number of reported events over the total number of reports submitted. CIs for all proportions were calculated via a 1-sample test of proportions.

Results: ILI was detected in 3.8% (32,239/848,878) of participants in the baseline seasons (2016-2019), 2.58% (7418/287,909) in the 2019-2020 season, and 0.27% (546/201,079) in the 2020-2021 season. Both influenza seasons that overlapped with the COVID-19 pandemic had lower ILI rates than the baseline seasons. ILI decline was observed during the months with widespread implementation of COVID-19 precautions, starting in February 2020. Self-reported influenza diagnoses decreased from early 2020 through the influenza season. Self-reported influenza positivity among ILI cases varied over the observed time period. Self-reported influenza vaccination rates in FNY were high across all observed seasons.

Conclusions: A decrease in ILI was detected in the crowdsourced FNY surveillance system during the 2019-2020 and 2020-2021 influenza seasons, mirroring trends observed in other influenza surveillance systems. Specifically, the months within seasons that overlapped with widespread pandemic precautions showed decreases in ILI and confirmed influenza. Concerns persist regarding respiratory pathogens re-emerging with changes to COVID-19 guidelines. Traditional surveillance is subject to changes in health care behaviors. Systems like FNY are uniquely situated to detect disease across disease severity and care seeking, providing key insights during public health emergencies.

(JMIR Public Health Surveill 2023;9:e40216) doi:10.2196/40216

KEYWORDS

participatory surveillance; influenza; crowdsourced data; disease surveillance; surveillance; COVID-19; respiratory; transmission; detection; survey; sore throat; fever; cough; vaccination; diagnosis; precautions

Introduction

Globally, seasonal respiratory viruses had lower incidence during the 2019-2020 season, which overlapped with the emergence of COVID-19 [1-4]. In the United States, decreases in influenza activity were also observed in 2019-2021 [2,5-8]. A national influenza-like illness (ILI) reference configured using an average of the previous few seasons was established at 2.6%; however, no week within the 2020-2021 season reached an ILI rate above 1.7% [8]. While influenza burden varies seasonally [9], COVID-19 contributed to this widespread decrease [5-7,10].

Many nonpharmaceutical interventions (NPIs) enacted for COVID-19 are also recommended for pandemic influenza [11-13]; therefore, their extensive uptake impacted the cocirculation of influenza. For example, those with mild symptoms for both illnesses are advised to remain home [13]. Traditional disease surveillance may omit these mild illnesses and not capture health care–related changes [14,15]. Crowdsourced estimates can be less susceptible to fluctuations in care seeking resulting from the pandemic [15,16]. Here, we use Flu Near You (FNY), a previously validated crowdsourced influenza surveillance system, to measure seasonal influenza incidence during the COVID-19 pandemic [17].

FNY is uniquely situated to further assess changes in seasonal ILI during the pandemic [14,17,18]. Historically, FNY has provided ILI trend insights that complement traditional surveillance and display compatibility with data from systems like the US Centers for Disease Control and Prevention's Influenza-like Illness Surveillance Network [17,18]. This study aims to assess whether crowdsourced syndromic surveillance capable of capturing mild influenza also revealed the global decrease in ILI during influenza seasons concurrent with the COVID-19 pandemic.

Methods

FNY Participatory Syndromic Surveillance System

Flu Near You (FNY) is a web-based participatory syndromic surveillance system where participants in the United States report their health using a weekly survey. Guest participants may also submit ad hoc reports. Participants are asked if they or household members feel healthy or sick. If sick, participants are prompted with a populated list of symptoms to select (Multimedia Appendix 1). Seasonal influenza vaccination status is also collected. Self-reported influenza diagnosis data were captured starting February 2020.

Population Sample

Data from 2016-2021 were included, with influenza seasons defined as *Morbidity and Mortality Weekly Report* weeks 40 through 20 (Multimedia Appendix 2, Table S1). Data were included for the entire FNY week that captured the influenza season start and end dates. Weekly reports by both registered and guest participants were included. Registered participants had a single identifier, while guest participants were assigned one per report. FNY used token validation between the front and back ends of the website to reduce the risk of spam reporting. Reports by registered participants were deduplicated by week if multiple reports were captured, with the most recent report included. For demographic analysis, each user was counted once per flu season, and those with missing data were excluded. No further restrictions were applied [17].

Ethical Considerations

This study of FNY data received approval from the Boston Children's Hospital Institutional Review Board (IRB-P00023700) and received a waiver of informed consent. FNY participants may report anonymously. Those who opted to provide contact information were deidentified. No compensation was provided to participants.

Analysis

ILI was defined as fever in addition to sore throat, cough, or both [19]. Influenza diagnosis was classified as self-reported influenza diagnosis (hereafter called self-reported diagnosis). Influenza vaccination rate was defined as the proportion of unique reporters who indicated being vaccinated any time during the influenza season. ILI rates were determined as the number of reports meeting the ILI definition over the total reports during that time. Season and monthly ILI rates were assessed with October and May rates truncated to align with the defined season. CIs were calculated via a 1-sample test of proportions.

Total ILI cases and unique reports from the 2016-2017, 2017-2018, and 2018-2019 influenza seasons were averaged by month and are presented as "baseline seasons." Demographics across the baseline seasons were aggregated to retain their specific proportions, accounting for participants reporting across seasons. Analyses were conducted with R (version 4.0.4; R Foundation for Statistical Computing).

Results

Population

During the baseline influenza seasons, 237,309 unique individuals reported on FNY, while 68,800 individuals reported for the 2019-2020 season and 25,996 in 2020-2021. There were more female respondents to FNY across seasons (Table 1). Higher proportions of older individuals reported across seasons (Table 1). In 2019-2020 and 2020-2021, respondents aged 65 years and older comprised the largest proportions (n=9812, 32.6% and n=8236, 39.9%, respectively) among those who reported age (n=30,139, 43.8% and n=20,650, 79.4%). Those younger than 18 years were the smallest proportion across influenza seasons (Table 1).

Table 1. Characteristics of respondents for age and sex.

	Respondents, n (%)		
	Baseline flu seasons (N=274,235[a])	2019-2020 flu season (n=68,800)	2020-2021 flu season (n=25,996)
Respondents who reported age	85,745 (31.3)	30,139 (43.8)	20,650 (79.4)
Age group (years)[b]			
13-17	6625 (7.7)	2419 (8.0)	1448 (7.0)
18-34	10,409 (12.1)	3445 (11.4)	2035 (9.9)
35-44	10,356 (12.1)	3269 (10.9)	1733 (8.4)
45-54	14,291 (16.7)	4638 (15.4)	2697 (13.0)
55-64	19,874 (23.2)	6556 (21.8)	4501 (21.8)
≥65	24,190 (28.2)	9812 (32.6)	8236 (39.9)
Respondents who reported sex	85,741 (31.3)	29,897 (43.5)	20,561 (79.1)
Sex[c]			
Female	52,221 (60.9)	17,980 (56.6)	12,110 (58.9)
Male	33,520 (39.1)	11,917 (43.4)	8451 (41.1)

[a] Combined demographics from the 2016-2017, 2017-2018, and 2018-2019 influenza seasons.

[b] Percentages reflect the proportion among those who reported age.

[c] Percentages reflect the proportion among those who reported sex.

ILI Trends

ILI variability was observed across the baseline seasons (Table 2). The characteristics of ILI captured on FNY in the described seasons followed patterns observed in traditional surveillance [9]. The 2016-2017, 2018-2019, and 2019-2020 influenza seasons were classified as moderate, and the 2017-2018 season was severe [9]. On FNY, the observed rate of ILI was 1.51% in 2016-2017, 4.98% in 2017-2018, and 4% in 2018-2019. Across the baseline seasons, the rate of ILI was 3.8%. Both the 2019-2020 and 2020-2021 seasons had lower ILI rates compared to the baseline seasons (Table 2). An ILI rate of 2.55% was observed in 2019-2020 and a rate of only 0.27% was captured in 2020-2021.

Monthly ILI rates displayed expected fluctuations within influenza seasons. The baseline seasons had ILI rates between 2.1% (95% CI 1.89%-2.34%) and 5.79% (95% CI 5.56%-6.02%), with peak activity seen from December to February (Figure 1). ILI activity peaked in January 2020 (4.5%) before moderately declining in February (3.29%) and further declining over the remainder of the season (Figure 1). For the 2020-2021 season, there was no evident peak in ILI (Figure 1).

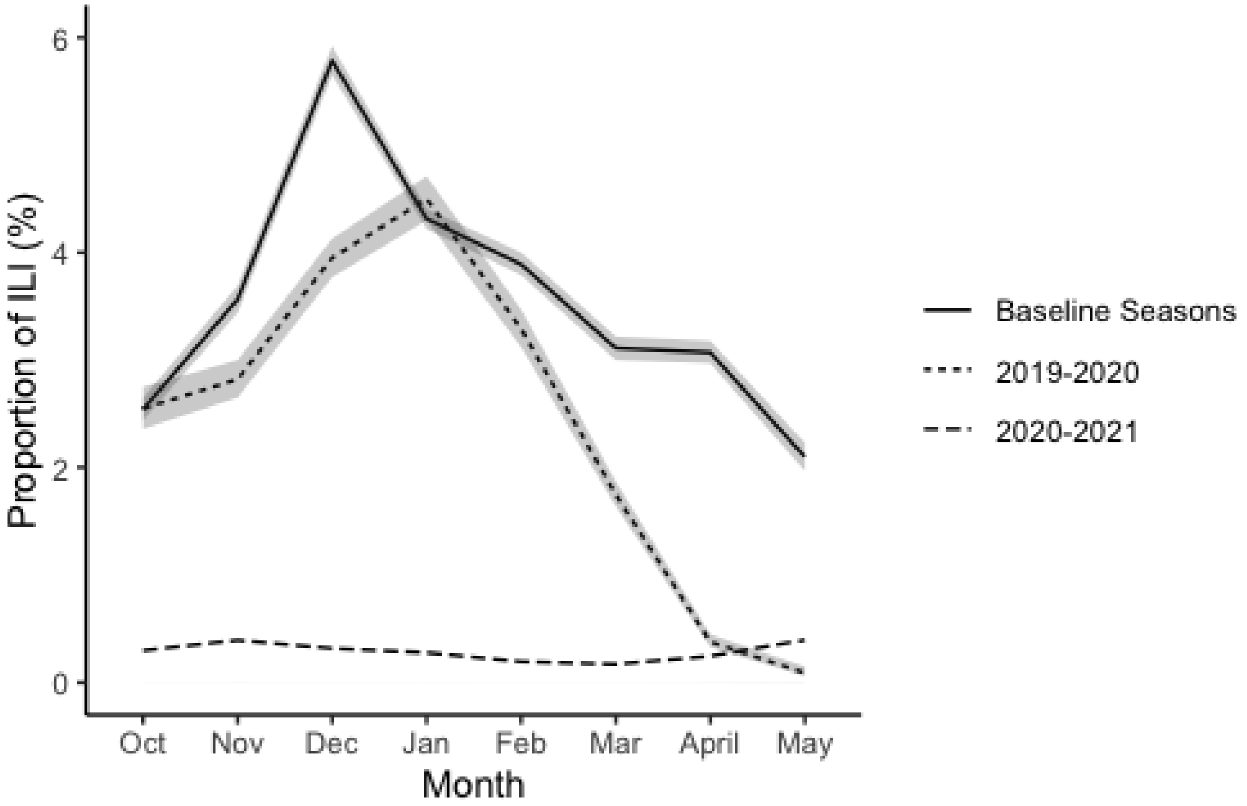
Table 2. Influenza-like-illness (ILI; defined as fever with cough, sore throat, or both) across influenza seasons on Flu Near You.

Influenza season	ILI cases, n	Total reports[a], n	ILI rate, % (95% CI)	Peak month(s)	Rate during peak month(s), %
2016-2017	3146	207,901	1.51 (1.46-1.57)	Feb	3.82
2017-2018	17,616	353,758	4.98 (4.91-5.05)	Dec	10.17
2018-2019	11,477	287,219	4 (3.92-4.07)	Feb	4.94
2019-2020	7418	287,909	2.55 (2.52-2.64)	Jan	4.5
2020-2021	546	201,079	0.27 (0.25-0.3)	Nov, May	0.4
Baseline seasons[b]	32,239	848,878	3.8 (3.76-3.83)	Dec	5.79

[a] Number of unique weekly reports submitted to Flu Near You.

[b] Combined ILI rate from the 2016-2017, 2017-2018, and 2018-2019 influenza seasons.

Figure 1. Monthly rates of influenza-like illness (ILI) detected in Flu Near You by influenza season. The 95% CIs are presented with the rates.



Self-Reported Influenza Diagnosis

Reductions in self-reported influenza diagnosis were detected in FNY (Table 3). Starting at the point of collection (February 2020), the rates of self-reported influenza diagnosis continually decreased (Table 3). February 2020 saw the highest rate of self-reported diagnosed influenza (0.97%) and the highest positivity rate (11.72%) among those tested. The lowest rates of influenza (0.01%) were detected in May 2020 and February and March 2021. From February to May of the 2019-2020 influenza season, the overall influenza rate was 0.21%, and the positivity rate was 7.04%. For the entire 2020-2021 influenza season, there was a lower rate of influenza (0.03%) and a lower positivity rate (1.1%). Reduced reports of influenza-positive ILI were also observed (Table 3).

Table 3. Self-reported influenza diagnoses in Flu Near You population.

Time period	Confirmed diagnoses of influenza[a]/total reports, n/N (%)	Positivity[b] (confirmed diagnoses/total tested for influenza), n/N (%)	Confirmed cases of ILI[c,d]/ILI cases, n/N (%)
2019-2020			
February	389/40,169 (0.97)	389/3318 (11.72)	109/1321 (8.25)
March	243/52,499 (0.46)	243/4135 (5.88)	87/919 (9.47)
April	18/38,096 (0.05)	18/1197 (1.5)	14/146 (9.59)
May	1/16,962 (0.01)	1/517 (0.19)	0/16 (0)
2020-2021			
October	7/24,812 (0.03)	7/606 (1.16)	5/74 (6.76)
November	5/25,023 (0.02)	5/824 (0.61)	3/99 (3.03)
December	10/31,181 (0.03)	10/798 (1.25)	0/99 (0)
January	13/30,157 (0.04)	13/751 (1.73)	7/84 (8.33)
February	4/30,097 (0.01)	4/534 (0.75)	1/59 (1.69)
March	4/36,149 (0.01)	4/610 (0.66)	3/62 (4.84)
April	4/16,629 (0.02)	4/359 (1.11)	1/41 (2.44)
May	6/7031 (0.09)	6/301 (1.99)	2/28 (7.14)

[a] Self-reported influenza diagnoses among entire reporting population by month.

[b] Self-reported influenza diagnoses among those who reported influenza testing by month.

[c] ILI: influenza-like-illness.

[d] Self-reported influenza diagnoses among those who met the definition of ILI by month.

Influenza Vaccination

Influenza vaccination rates varied seasonally, with 58.6% (160,610/274,235) of participants being vaccinated in the baseline seasons. Both the 2016-2017 and 2018-2019 seasons had over 80% vaccinated participants (23,632/27,024, 87.4% and 68,588/84,185, 81.5%, respectively). In the 2017-2018 season, which had more guest participants, 42% (68,390/163,026) of participants reported influenza vaccination. The 2019-2020 season had a high vaccination rate of 84% (57,823/68,800). The 2020-2021 season had a lower vaccination rate (66.5%) (17,299/25,996) compared to the previous season.

Discussion

FNY observed unseasonably low rates of ILI during 2019-2021. Late 2019 had higher rates of ILI than early 2020, mirroring a globally observed reduction in ILI [2,3,5-7,10]. FNY observed ILI decreasing in February 2020, slightly ahead of the nationwide implementation of COVID-19 precautions. This early decrease is likely attributed to high health literacy among FNY participants, as they are probable early adopters of NPIs [17]. For the entirety of the 2020-2021 influenza season, FNY ILI rates were comparable to those outside of the viral season [3,8]. Influenza diagnoses on FNY also decreased from February to May 2020, mirroring ILI trends. This alignment of trends displays the importance of complementing traditional surveillance with participatory systems [20,21], and further suggests that prevention measures for COVID-19 may have secondarily reduced transmission of seasonal influenza.

FNY participants have historically reported influenza vaccination rates around 80% [22], compared to under 50% in the US population [23-25]. Both 2019-2020 and 2020-2021 had higher vaccination rates than the baseline seasons, potentially contributing to low rates of ILI. However, this correlation was impractical to assess due to the high influenza vaccination rates among FNY participants.

This study has limitations worth noting. First, the FNY population is not representative of the US population, limiting the scope of the results [14,15,18]. However, as the FNY data are longitudinal, we are able to compare influenza seasons within this specific population. Second, self-reported data are subject to bias, which can lead to misclassification. Furthermore, ILI is a coarse measurement that may detect other circulating respiratory infections. Therefore, ILI rates may include other seasonal viruses and have inflated baseline levels [1]. Contrarily, COVID-19 may have been captured as ILI. In April and May 2021, there was a delivery issue with FNY reminders, resulting in fewer reports than expected. The minimal increase in detected ILI is likely attributable to less reporting of being healthy, as sick individuals are more likely to report unprompted. Lastly, Outbreaks Near Me [26] was launched in March 2020, contributing to lower reporting on FNY. The low rates of ILI during the 2020-2021 season suggest prevention measures for COVID-19 may have secondarily mitigated transmission of seasonal influenza. However, interpretation of ILI during periods of low circulation should be approached cautiously and may be best studied with the use of zero-inflation models that can appropriately account for zeros in the syndromic surveillance data.

Concerns persist regarding respiratory pathogens re-emerging with COVID-19 guideline changes and declined NPI use [5-7]. The fall 2022 “triple-demic” exemplifies the legitimacy of this concern. Traditional surveillance provides valuable insights but can be subject to delays and changes to the health care system [27]. For example, at-home diagnostics data revealed additional populations traditional surveillance is prone to omit [28]. Crowdsourced surveillance systems like FNY are uniquely situated to detect disease among populations who may be missed by traditional surveillance, thereby providing key insights early in public health emergencies.

Acknowledgments

The authors thank the Flu Near You team, as well as the platform participants. JSB and MS conceptualized the survey and platform. JSB, MS, KS, and AG facilitated project administration. AG and KS conceptualized the study; AG, BR, KS, and JSB designed the analysis; AG conducted the analyses; BR and TJV provided analytic supervision. AG drafted the initial manuscript; all authors contributed to the revised draft and approved the final manuscript. This study and analysis were not funded.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Flu near you questionnaire.
[DOCX File , 17 KB - publichealth_v9i1e40216_app1.docx]

Multimedia Appendix 2
Table S1. Influenza season start and end dates.
[DOCX File , 17 KB - publichealth_v9i1e40216_app2.docx]

References

1. Maharaj AS, Parker J, Hopkins JP, Gournis E, Bogoch II, Rader B, et al. The effect of seasonal respiratory virus transmission on syndromic surveillance for COVID-19 in Ontario, Canada. Lancet Infect Dis 2021 May;21(5):593-594 [FREE Full text] [doi: 10.1016/S1473-3099(21)00151-1] [Medline: 33773620]
2. Chan KH, Lee P, Chan CY, Lam KBH, Ho P. Monitoring respiratory infections in covid-19 epidemics. BMJ 2020 May 04;369:m1628. [doi: 10.1136/bmj.m1628] [Medline: 32366507]
3. Olsen SJ, Azziz-Baumgartner E, Budd AP, Brammer L, Sullivan S, Pineda RF, et al. Decreased influenza activity during the COVID-19 Pandemic - United States, Australia, Chile, and South Africa, 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 18;69(37):1305-1309 [FREE Full text] [doi: 10.15585/mmwr.mm6937a6] [Medline: 32941415]
4. Weekly seasonal influenza updates. World Health Organization. URL: https://www.who.int/teams/global-influenza-programme/surveillance-and-monitoring/influenza-updates [accessed 2021-04-26]
5. Baker RE, Park SW, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. Proc Natl Acad Sci USA 2020 Dec 01;117(48):30547-30553 [FREE Full text] [doi: 10.1073/pnas.2013182117] [Medline: 33168723]
6. Sinha P, Reifler K, Rossi M, Sagar M. Coronavirus disease 2019 mitigation strategies were associated with decreases in other respiratory virus infections. Open Forum Infect Dis 2021 Jun;8(6):ofab105 [FREE Full text] [doi: 10.1093/ofid/ofab105] [Medline: 34514014]
7. Huang QS, Wood T, Jelley L, Jennings T, Jefferies S, Daniells K, et al. Impact of the COVID-19 nonpharmaceutical interventions on influenza and other respiratory viral infections in New Zealand. Nat Commun 2021 Feb 12;12(1):1001 [FREE Full text] [doi: 10.1038/s41467-021-21157-9] [Medline: 33579926]
8. Weekly U.S. influenza surveillance report. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/weekly/index.htm [accessed 2023-12-19]
9. Past Seasons Estimated Influenza Disease Burden. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/about/burden/past-seasons.html [accessed 2023-12-19]
10. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Health 2020 May;5(5):e279-e288 [FREE Full text] [doi: 10.1016/S2468-2667(20)30090-6] [Medline: 32311320]

11. World Health Organization Writing Group, Bell D, Nicoll A, Fukuda K, Horby P, Monto A, et al. Non-pharmaceutical interventions for pandemic influenza, international measures. Emerg Infect Dis 2006 Jan;12(1):81-87 [FREE Full text] [doi: 10.3201/eid1201.051370] [Medline: 16494722]
12. World Health Organization Writing Group, Bell D, Nicoll A, Fukuda K, Horby P, Monto A, et al. Non-pharmaceutical interventions for pandemic influenza, national and community measures. Emerg Infect Dis 2006 Jan;12(1):88-94 [FREE Full text] [doi: 10.3201/eid1201.051371] [Medline: 16494723]
13. Influenza (seasonal). World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal) [accessed 2023-12-19]
14. Baltrusaitis K, Vespignani A, Rosenfeld R, Gray J, Raymond D, Santillana M. Differences in regional patterns of influenza activity across surveillance systems in the United States: comparative evaluation. JMIR Public Health Surveill 2019 Sep 14;5(4):e13403 [FREE Full text] [doi: 10.2196/13403] [Medline: 31579019]
15. Baltrusaitis K, Reed C, Sewalk K, Brownstein JS, Crawley AW, Biggerstaff M. Healthcare-seeking behavior for respiratory illness among Flu Near You participants in the United States during the 2015-2016 through 2018-2019 influenza seasons. J Infect Dis 2022 Aug 24;226(2):270-277 [FREE Full text] [doi: 10.1093/infdis/jiaa465] [Medline: 32761050]
16. Czeisler, Marynak K, Clarke KEN, Salah Z, Shakya I, Thierry JM, et al. Delay or avoidance of medical care because of COVID-19-related concerns - United States, June 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 11;69(36):1250-1257 [FREE Full text] [doi: 10.15585/mmwr.mm6936a4] [Medline: 32915166]
17. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: crowdsourced symptom reporting spanning 2 influenza seasons. Am J Public Health 2015 Oct;105(10):2124-2130. [doi: 10.2105/AJPH.2015.302696] [Medline: 26270299]
18. Baltrusaitis K, Santillana M, Crawley AW, Chunara R, Smolinski M, Brownstein JS. Determinants of participants' follow-up and characterization of representativeness in Flu Near You, a participatory disease surveillance system. JMIR Public Health Surveill 2017 Apr 07;3(2):e18 [FREE Full text] [doi: 10.2196/publichealth.7304] [Medline: 28389417]
19. U.S. influenza surveillance: purpose and methods. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/weekly/overview.htm [accessed 2023-12-19]
20. Wittwer S, Paolotti D, Lichand G, Leal Neto O. Participatory surveillance for COVID-19 trend detection in Brazil: cross-sectional study. JMIR Public Health Surveill 2023 Apr 26;9:e44517 [FREE Full text] [doi: 10.2196/44517] [Medline: 36888908]
21. Ackley SF, Pilewski S, Petrovic VS, Worden L, Murray E, Porco TC. Assessing the utility of a smart thermometer and mobile application as a surveillance tool for influenza and influenza-like illness. Health Informatics J 2020 Sep;26(3):2148-2158 [FREE Full text] [doi: 10.1177/1460458219897152] [Medline: 31969046]
22. Maintaining user engagement in an infectious disease surveillance-related citizen science project. Citizen Science: Theory and Practice. URL: https://theoryandpractice.citizenscienceassociation.org/article/10.5334/cstp.302/ [accessed 2023-12-19]
23. Influenza vaccination coverage for persons 6 months and older | FluVaxView | seasonal influenza (flu). US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/fluvaxview/interactive.htm [accessed 2023-12-19]
24. Flu vaccination coverage, United States, 2019-20 influenza season | FluVaxView | seasonal influenza (flu) |. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/fluvaxview/coverage-1920estimates.htm [accessed 2023-12-19]
25. Early-season influenza vaccination uptake and intent among adults - United States, September 2020. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/fluvaxview/nifs-estimates-sept2020.htm [accessed 2023-12-19]
26. Outbreaks Near Me. URL: https://outbreaksnearme.org [accessed 2023-12-19]
27. Furlow B. Triple-demic overwhelms paediatric units in US hospitals. Lancet Child Adolesc Health 2023 Feb;7(2):86. [doi: 10.1016/S2352-4642(22)00372-8] [Medline: 36529148]
28. Rader B, Gertz A, Iuliano AD, Gilmer M, Wronski L, Astley CM, et al. Use of at-home COVID-19 tests - United States, August 23, 2021-March 12, 2022. MMWR Morb Mortal Wkly Rep 2022 Apr 01;71(13):489-494 [FREE Full text] [doi: 10.15585/mmwr.mm7113e1] [Medline: 35358168]

Abbreviations

FNY: Flu Near You
ILI: influenza-like illness
NPI: nonpharmaceutical intervention

Edited by G Eysenbach; submitted 13.06.22; peer-reviewed by S Ps, C Dalton, Z Li, O Leal Neto, E Lau, J Jiang; comments to author 05.07.23; revised version received 24.07.23; accepted 14.11.23; published 28.12.23.

Please cite as:
Gertz A, Rader B, Sewalk K, Varrelman TJ, Smolinski M, Brownstein JS
Decreased Seasonal Influenza Rates Detected in a Crowdsourced Influenza-Like Illness Surveillance System During the COVID-19 Pandemic: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e40216
URL: https://publichealth.jmir.org/2023/1/e40216
doi:10.2196/40216
PMID:38153782

Original Paper

Demographic Determinants and Geographical Variability of COVID-19 Vaccine Hesitancy in Underserved Communities: Cross-sectional Study

Jennifer L Matas[1*], MPH; Latrice G Landry[2,3,4*], MS, MMSc, PhD; LaTasha Lee[5], MPH, PhD; Shantoy Hansel[1], MPH; Makella S Coudray[1], MPH, PhD; Lina V Mata-McMurry[1], MD; Nishanth Chalasani[6], BA; Liou Xu[1], PhD; Taylor Stair[1], BBA; Christina Edwards[1], MHA; Gary Puckrein[1], PhD; William Meyer[7], PhD; Gary Wiltz[8], MD; Marian Sampson[9], MD; Paul Gregerson[10], MBA, MD; Charles Barron[11], MD; Jeffrey Marable[12], MD; Olakunle Akinboboye[13], MBA, MPH, MD; Dora Il'yasova[1,14], PhD

[1]National Minority Quality Forum, Washington, DC, United States
[2]University of Pennsylvania, Philadelphia, PA, United States
[3]Dana Farber Cancer Institute, Boston, MA, United States
[4]Harvard Medical School, Boston, MA, United States
[5]The George Washington University, Washington, DC, United States
[6]Morsani College of Medicine, University of South Florida, Tampa, FL, United States
[7]Quest Diagnostics, Secaucus, NJ, United States
[8]Teche Action Clinic, Franklin, LA, United States
[9]Osceola Community Health Services, Kissimmee, FL, United States
[10]John Wesley Community Health, Commerce, CA, United States
[11]Aunt Martha’s Health and Wellness, Olympia Fields, IL, United States
[12]PrimaryOne Health, Columbus, OH, United States
[13]Queens Heart Institute, Rosedale, NY, United States
[14]Department of Family Medicine and Community Health, School of Medicine, Duke University, Durham, NC, United States
[*]these authors contributed equally

Corresponding Author:
Dora Il'yasova, PhD
National Minority Quality Forum
1201 15th Street, NW
Suite 340
Washington, DC, 20005
United States
Phone: 1 202 223 7560 ext 175
Email: dilyasova@nmqf.org

Abstract

Background: COVID-19 hospitalizations and deaths disproportionately affect underserved and minority populations, emphasizing that vaccine hesitancy can be an especially important public health risk factor in these populations.

Objective: This study aims to characterize COVID-19 vaccine hesitancy in underserved diverse populations.

Methods: The Minority and Rural Coronavirus Insights Study (MRCIS) recruited a convenience sample of adults (age≥18, N=3735) from federally qualified health centers (FQHCs) in California, the Midwest (Illinois/Ohio), Florida, and Louisiana and collected baseline data in November 2020-April 2021. Vaccine hesitancy status was defined as a response of “no” or “undecided” to the question “Would you get a coronavirus vaccine if it was available?” (“yes” categorized as not hesitant). Cross-sectional descriptive analyses and logistic regression models examined vaccine hesitancy prevalence by age, gender, race/ethnicity, and geography. The expected vaccine hesitancy estimates for the general population were calculated for the study counties using published county-level data. Crude associations with demographic characteristics within each region were assessed using the chi-square test. The main effect model included age, gender, race/ethnicity, and geographical region to estimate adjusted odds ratios (ORs) and 95% CIs. Interactions between geography and each demographic characteristic were evaluated in separate models.

Results: The strongest vaccine hesitancy variability was by geographic region: California, 27.8% (range 25.0%-30.6%); the Midwest, 31.4% (range 27.3%-35.4%); Louisiana, 59.1% (range 56.1%-62.1%); and Florida, 67.3% (range 64.3%-70.2%). The expected estimates for the general population were lower: 9.7% (California), 15.3% (Midwest), 18.2% (Florida), and 27.0% (Louisiana). The demographic patterns also varied by geography. An inverted U-shaped age pattern was found, with the highest prevalence among ages 25-34 years in Florida (n=88, 80.0%,) and Louisiana (n=54, 79.4%; *P*<.05). Females were more hesitant than males in the Midwest (n= 110, 36.4% vs n= 48, 23.5%), Florida (n=458, 71.6% vs n=195, 59.3%), and Louisiana (n= 425, 66.5% vs. n=172, 46.5%; *P*<.05). Racial/ethnic differences were found in California, with the highest prevalence among non-Hispanic Black participants (n=86, 45.5%), and in Florida, with the highest among Hispanic (n=567, 69.3%) participants (*P*<.05), but not in the Midwest and Louisiana. The main effect model confirmed the U-shaped association with age: strongest association with age 25-34 years (OR 2.29, 95% CI 1.74-3.01). Statistical interactions of gender and race/ethnicity with the region were significant, following the pattern found by the crude analysis. Compared to males in California, the associations with the female gender were strongest in Florida (OR=7.88, 95% CI 5.96-10.41) and Louisiana (OR=6.09, 95% CI 4.55-8.14). Compared to non-Hispanic White participants in California, the strongest associations were found with being Hispanic in Florida (OR=11.18, 95% CI 7.01-17.85) and Black in Louisiana (OR=8.94, 95% CI 5.53-14.47). However, the strongest race/ethnicity variability was observed within California and Florida: the ORs varied 4.6- and 2-fold between racial/ethnic groups in these regions, respectively.

Conclusions: These findings highlight the role of local contextual factors in driving vaccine hesitancy and its demographic patterns.

(JMIR Public Health Surveill 2023;9:e34163) doi:10.2196/34163

KEYWORDS

COVID-19 vaccine; vaccine hesitancy; underrepresented in research; minority populations; federally qualified health center (FQHC); public health; COVID-19

Introduction

Vaccine hesitancy has been defined as a “delay in acceptance or refusal of vaccination despite the availability of vaccination services” and is recognized among the 10 major public health threats by the World Health Organization (WHO) [1]. A recent review of the literature showed that vaccine hesitancy has become 1 of the key research topics [2]. The Centers for Disease Control and Prevention (CDC) published a map with the estimates of COVID-19 vaccine hesitancy, demonstrating its wide geographical variability [3]. Such variability emphasizes the importance of cultural, social, and economic contexts in driving COVID-19 vaccine hesitancy. It has been documented that COVID-19 vaccine hesitancy and vaccination rates have been low among racial/ethnic minority populations [4], even though these populations have been disproportionately affected by the pandemic, as shown by the rates of hospitalizations and deaths [5-8]. However, in contrast to the United States, a greater prevalence of COVID-19 vaccine hesitancy was not found among Black participants in the United Kingdom [9], emphasizing the importance of contextual cultural and socioeconomic factors.

The focus of our research is underserved communities in different geographical regions of the United States with vastly different racial/ethnic, cultural, and socioeconomic contexts. People in historically underserved communities have a higher proportion of racial/ethnic minorities, groups that are disproportionally affected by COVID-19 [5-8]. Moreover, people in underserved communities are at higher risk of more severe COVID-19 outcomes because of a greater prevalence of comorbidities associated with severe COVID-19 [10]. Understanding the variability in vaccination interest within and between minority and underserved communities is essential as effective vaccine administration policies and interventions require a nuanced understanding of the obstacles [11]. To date, few studies of COVID-19 vaccine hesitancy have reported estimates from underserved communities; this includes studies conducted in Delaware, North Carolina, and Monterey County, California [12-14]. Two studies were conducted in the year 2020, one study in North Carolina and one in California. The study in North Carolina classified overall 69% of respondents as vaccine hesitant, with the highest prevalence among White (74%), followed by Black (62.7%) and Hispanic (59.5%) participants. The study in Monterey County, California, included Hispanic farmworkers and reported 48.4% “not extremely likely to get vaccinated.” The study in Delaware was conducted in March-May 2021 and reported an overall 42% vaccine hesitancy prevalence, with 45% of Black participants expressing vaccine hesitancy. These previous studies also demonstrate that vaccine hesitancy is likely to be driven by contextual factors specific to geographic regions.

Here, we present the results from a large Minority and Rural Coronavirus Insights Study (MRCIS; N=3462) conducted in November 2020-March 2021 among underserved racially/ethnically diverse communities across 4 geographical US regions. The MRCIS is a multisite, community-based, epidemiologic investigation of the social and structural determinants of health, clinical, environmental, and genetic factors associated with the COVID-19 pandemic in minority and rural communities in the United States. Our objectives are to describe the demographic determinants of COVID-19 vaccine hesitancy within each geographical region and compare the estimates of vaccine hesitancy in these underserved communities with those expected for the general population in their respective local regions.

Methods

Study Design

This cross-sectional analysis of COVID-19 vaccine hesitancy prevalence was conducted using the baseline data collected by the MRCIS. A total of 3735 participants were enrolled, 155 (4.1%) participants were not qualified to participate (under the age of 18 years), and 29 (0.8%) participants were removed from the study due to missing informed consent, resulting in a study population of 3551 (95.1%) participants. The analytical sample included 3491 (98.3%) participants after the exclusion of 60 (1.7%) participants, who responded to the vaccine hesitancy question as “N/A” or left that question blank. As the last step of the analysis, we excluded 29 participants who were erroneously enrolled twice with a separate ID; this exclusion yielded a final analytical sample of 3462.

Study Participants

In November 2020, the National Minority Quality Forum (NMQF) launched the MRCIS, a prospective longitudinal investigation of risk and socioeconomic factors associated with the disproportionate impact of COVID-19 on minority and rural communities. Federally qualified health centers (FQHCs), funded through the Health Resources & Services Administration (HRSA), were invited to partner with the MRCIS as community-based health care providers. The FQHCs were established to operate in communities that are underserved and therefore have been underrepresented in research. Research sites were selected based on the high proportion of deaths in minority populations that they serve compared to the proportion of minorities in the state population. In total, 5 community health centers in 4 geographically diverse HRSA regions (regions 6, 4, 5, and 9) were selected to participate in the MRCIS. In the Southeast region, participants were recruited from FQHCs located in HRSA regions 4 and 6. Participants in HRSA region 4 were recruited from Osceola Community Health Services (OCHS), located in Kissimmee, Florida, and participants in HRSA region 6 were recruited from the Teche Action Clinic (TAC), located in Franklin, Louisiana. In the Southwest region, HRSA region 9, participants were recruited from the John Wesley County Hospital (JWCH), an FQHC located in Los Angeles, California. In HRSA region 5, recruitment occurred at 2 FQHCs: Aunt Martha (AM) in Olympia Fields, Illinois, and Primary One (PO) in Columbus Ohio.

Volunteers were recruited from a convenience sample of adults (age≥18 years). Recruitment strategies varied by site and included 1 or more of the following methods: (1) in-person recruitment at the site; (2) in-person recruitment at study centers set up in the communities, including the recreational center, a low-income housing complex, the fire department, and homeless shelters; and (3) advertising through flyers distributed to the community. Participants did not receive any compensation for participation in this study.

Ethical Considerations

At enrollment, participants completed an informed consent form and a baseline survey. The protocol was reviewed and approved by the WIRB-Copernicus Group Institutional Review Board (WCG IRB; #1292174). Each participant received a unique study ID, and the identifying information was removed from the data set for analysis to maintain patient privacy and confidentiality.

Data Collection and Outcome Definition

The data used in this analysis were collected in the MRCIS using a baseline survey, including self-reported age, gender, race, and ethnicity. The question for reporting race was the following: “Select all that apply for race: American Indian/Alaskan Native, Asian, Black/African American/Native Hawaiian/Other Pacific Islander, White, Other, and Prefer Not to Answer.” The options for the ethnicity question were the following: “Hispanic or Latino,” “Not Hispanic or Latino,” and “Prefer not to answer.” We used the race and ethnicity questions to develop combined race/ethnicity categories. Although participants were allowed to select more than 1 category for the race, no one in our sample reported multiple races. All individuals who self-reported Hispanic/Latino as their ethnicity were characterized as Hispanic/Latino regardless of what they indicated as their race. We developed 4 mutually exclusive racial/ethnic categories as follows: Hispanic/Latino, non-Hispanic Black, non-Hispanic White, and non-Hispanic other.

Participants were asked, “Would you get a coronavirus vaccine if it was available,” with the option to answer “yes,” “no,” or “undecided.” According to the WHO definition [1], vaccine hesitancy was defined as a response of “no” or “undecided.” The responses were recorded from November 2020 to March 2021.

Statistical Analysis

Vaccine Hesitancy Prevalence by Demographic Characteristics and Geographical Region

Self-reported age was categorized into the following age groups: 18-24, 25-34, 35-44, 45-54, 55-64, and ≥65 years. Race and ethnic information were categorized as noted before (Hispanic/Latino, non-Hispanic Black/African American, non-Hispanic White, and non-Hispanic other). Gender was characterized as female, male, and other (n=3). Due to a small sample size, the “Other” gender category was excluded from the regression analysis that adjusted for gender as a confounder. Descriptive statistics, including counts and proportions comparing individual-level factors in the overall study population and by vaccine hesitancy (not hesitant vs hesitant), were calculated. The strength of the adjusted associations of vaccine hesitancy prevalence with the demographic characteristics and geographical regions was assessed in logistic regression models by calculating odds ratios (ORs) and their 95% CIs. The main effect model included age (6 categories), gender (females vs males), race/ethnicity (4 categories), and HRSA site (4 sites). The reference groups for age, gender, race/ethnicity, and HRSA site were individuals aged 65 years and above, males, non-Hispanic White, and California, respectively.

Observed and Expected Vaccine Hesitancy Prevalence

The observed vaccine hesitancy prevalence and corresponding 95% CIs were calculated for each of the study sites. The expected vaccine hesitancy was calculated based on the county-level estimates reported by the CDC [3].

Each MRCIS participant was assigned to a residential county based on the reported residential zip code and the US Department of Housing and Urban Development (HUD) crosswalk files [15]. These crosswalk files are derived directly from the United States Postal Service (USPS) and are updated quarterly reflecting changes in zip code configurations. Converting the MRCIS participant zip code to residential counties enabled us to compare the MRCIS prevalence of vaccine hesitancy to the national estimates reported by the CDC [3]. National COVID-19 vaccine hesitancy data were downloaded directly from the CDC website. The CDC's outcome definition for vaccine hesitancy was derived from the US Census Bureau's Household Pulse Survey (HPS), which asked the following survey question: "Once a vaccine to prevent COVID-19 is available to you, would you…get a vaccine?" Participants were given the following response options: (1) "definitely get a vaccine," (2) "probably get a vaccine," (3) "unsure," (4) "probably not get a vaccine," and (5) "definitely not get a vaccine." The estimate of vaccine hesitancy we used was "hesitant or unsure," which directly compares with our definition of vaccine hesitancy [3].

For each region, the expected number of hesitant subjects was calculated by multiplying the number of MRCIS participants per county by the CDC's reported vaccine hesitancy prevalence. For example, 930 participants resided in Los Angeles County, California, where the vaccine hesitancy was estimated by the CDC as 9.6%, so the expected number of vaccine-hesitant participants was 89.2. With 27.1% vaccine hesitancy estimates for St Mary Parish, Louisiana, among 882 participants, 239 were expected to be hesitant. The expected number of participants was summed across each study site, and their percentage (ie, percentage of the participants at the site) was calculated. In addition to the expected vaccine hesitancy prevalence, we provided a county-based range for estimated vaccine hesitancy in each MRCIS site.

The assessment of COVID-19 vaccine hesitancy was conducted by the CDC and our study in different periods (mid-2021 vs November 2020-March 2021). Based on the previously published time trends of COVID-19 vaccine hesitancy, it is likely that our expected estimates are approximately 1/3 greater compared to the CDC's reported ones [16]. Therefore, we presented time-adjusted expected prevalence in the general population in the counties of our participants' residences.

Demographic Patterns of Vaccine Hesitancy Prevalence by Geographical Region

Age, gender, and the race/ethnicity patterns of vaccine hesitancy were explored in descriptive analyses (Figures 1-3). The Pearson chi-square test was performed to determine the statistical significance of each demographic pattern within each study site. As this crude analysis strongly suggested the modification of the demographic patterns by geographical region, multivariable logistic regression models were performed to evaluate the interactions between each of the demographic covariates and the study site in separate logistic regression models. In the model exploring the interaction between age and geographical region, individuals in California aged 55 years and above served as the common reference group. In this model, we collapsed the 2 oldest categories (55-64 and ≥65 years) because the main effects model detected similar adjusted associations of these age categories with vaccine hesitancy prevalence. In the model exploring the interaction between gender and geographical region, males in California served as the common reference group. Lastly, non-Hispanic White participants in California served as the common reference group in the model exploring the interaction between race/ethnicity and geographical region. A type 1 error rate of 5% was used in the entire analysis.

Akaike information criterion (AIC) statistics were used to assess model fitness, in which lower AIC estimates indicated improved model fitness. We reported the following AIC statistics:

- Main effects multivariable logistic regression model AIC=4218.985
- Multivariable logistic regression model with the interaction between age and region AIC=4227.904
- Multivariable logistic regression with the interaction of gender and region AIC=4213.966
- Multivariable logistic regression with the interaction of race/ethnicity and region AIC=4200.983

As demonstrated by the AIC, the interaction terms between age and region did not improve the model, with the interaction term *P* value of .32.

All analyses were conducted in SAS version 9.4 (SAS Institute) and R version 4.2.1 (Comprehensive R Archive Network).

Figure 1. Vaccine hesitancy by age and geographical location, the MRCIS (2020-2021). CA: California; FL: Florida; HRSA: Health Resources & Services Administration; IL: Illinois; LA: Louisiana; MRCIS: Minority and Rural Coronavirus Insights Study; OH: Ohio.

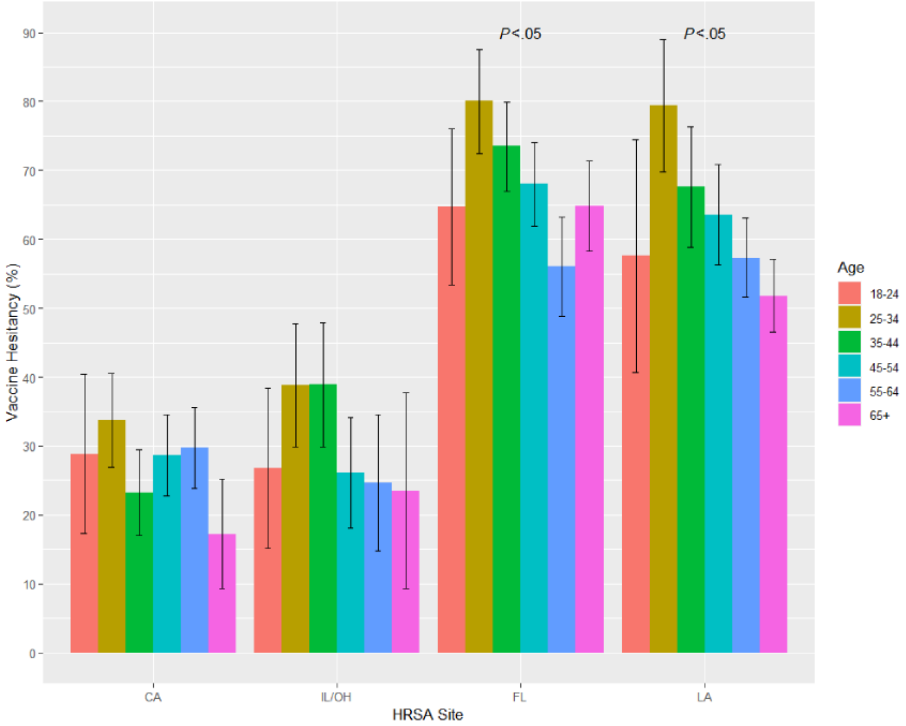


Figure 2. Vaccine hesitancy by gender and geographical location, the MRCIS (2020-2021). CA: California; FL: Florida; HRSA: Health Resources & Services Administration; IL: Illinois; LA: Louisiana; MRCIS: Minority and Rural Coronavirus Insights Study; OH: Ohio.

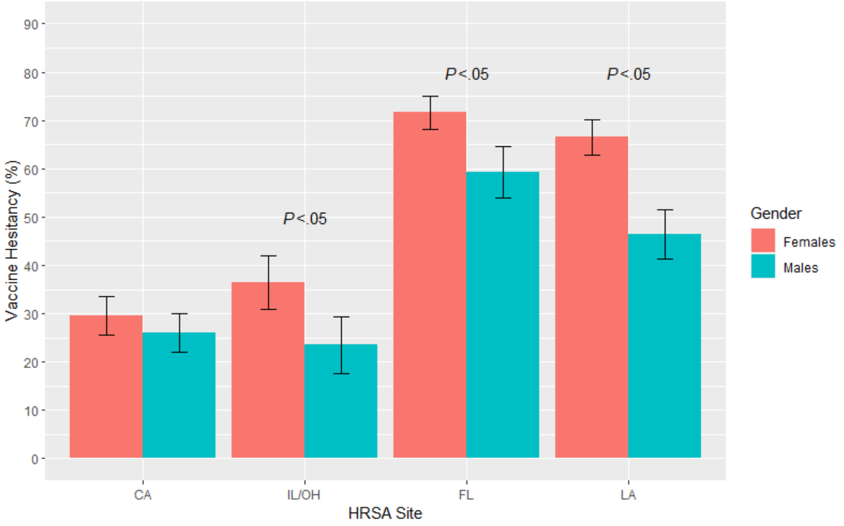
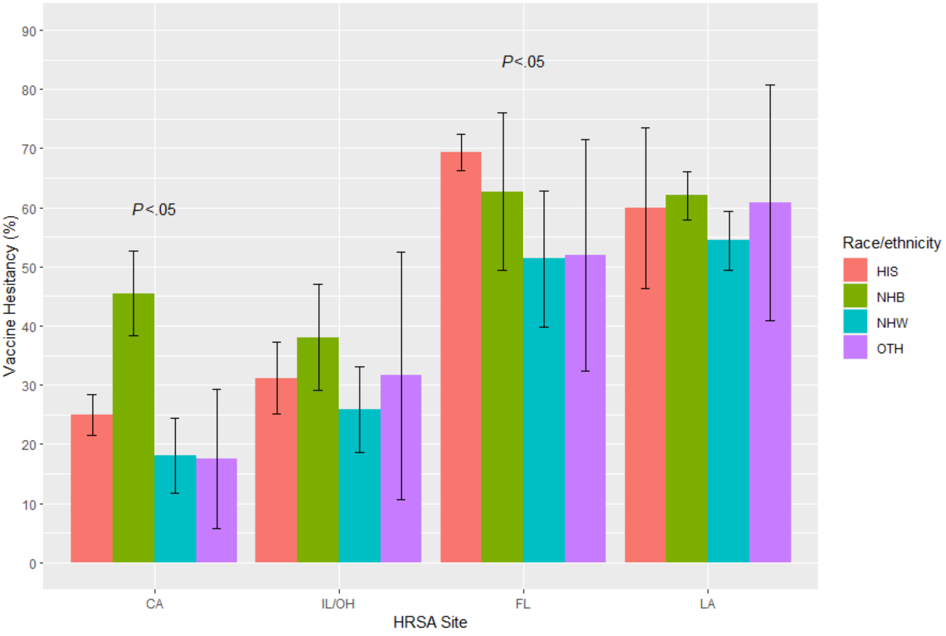


Figure 3. Vaccine hesitancy by race ethnicity and geographical location, the MRCIS (2020-2021). CA: California; FL: Florida; HIS: Hispanic/Latino; HRSA: Health Resources & Services Administration; IL: Illinois; LA: Louisiana; MRCIS: Minority and Rural Coronavirus Insights Study; NHB: non-Hispanic Black/African American; NHW: non-Hispanic White; OH: Ohio; OTH: non-Hispanic other.



Results

Participant Details

Our study population was diverse, including females (n=2076, 60.0%) and males (n=1379, 39.8%) aged 18-94 years. The 3 major racial/ethnic groups were well represented, with the largest group presented by Hispanic participants (n=1684, 48.6%), followed by non-Hispanic Black/African American (n=896, 25.9%) and non-Hispanic White (n=740, 21.4%) participants. Other non-Hispanic racial/ethnic groups were presented by 107 (3.1%) participants; due to the small sample size, this demographic group was not further stratified (Table 1). Among females, 781/2076 (37.6%) were considered as being of reproductive age (18-44 years).

In this study population, vaccine hesitancy exceeded 50% among the following demographic groups: females, participants 25-44 years of age, and non-Hispanic Black/African American participants (Table 1). Overall, no linear trend was apparent in vaccine hesitancy with the increase in age. In relation to geography, a striking variability of hesitancy prevalence was observed: the highest prevalence was found in Florida (653/1680, 38.9%) and Louisiana (597/1680, 35.5%); see Table 1. The differences between the highest (Florida) and the lowest (California) hesitancy prevalence were greater than 2-fold (Table 1). Because of this strong geographic variability, we explored whether the observed geographic differences were similar to what would be expected in the general population within the same geographic areas.

The range of vaccine hesitancy in the general population is presented for the counties where our participants resided (Table 2). The expected vaccine hesitancy in the general population overall resembled the geographical pattern that we observed in the underserved communities, that is, lower in California and the Midwest (Illinois/Ohio) compared to Florida and Louisiana. Considering the expected time trend, that is, approximately 1/3-fold reduction in hesitancy by mid-2021 (period of estimates from the CDC) as opposed to 2020-early 2021, when our study was conducted, we adjusted the expected vaccine hesitancy prevalence as presented in Table 2. Even after such an adjustment, the observed vaccine hesitancy prevalence in the underserved population tended to be greater than expected in the general population (Table 2). These results further emphasized the importance of assessing vaccine hesitancy in the underserved population.

Examination of the crude (unadjusted) prevalence estimates for vaccine hesitancy showed that their demographic patterns vary between geographical regions. There was no clear age pattern among the California and Midwest participants (Figure 1). In other regions, a trend for greater hesitancy among participants aged 25-44 years was evident (*P*<.05); see Figure 1. Greater vaccine hesitancy among females was noticeable in all regions but was more pronounced (*P*<.05) in Florida, Louisiana, and the Midwest (Figure 2). Racial/ethnic differences were detected in California and Florida (*P*<.05) but not in Louisiana and the Midwest (Figure 3).

The adjusted estimates of the associations between vaccine hesitancy prevalence and both demographic and geographical characteristics (main effect model, Table 3) followed a pattern similar to the observations presented in Table 1. An inverted U-shaped correlation was found between vaccine hesitancy and age. Specifically, vaccine hesitancy was less associated with the younger and older age groups, whereas the association with age was 1.4-2.3-fold greater among adults aged 25-54 years compared to older adults aged ≥65 years (Table 3). Within the 25-54-year age group, there was a clear trend of the inverse association with age, with vaccine hesitancy declining among participants older than 25-34 years (Table 3). Females were approximately 1.7-fold more likely to express vaccine hesitancy compared to males. We hypothesized that vaccine hesitancy would be even more strongly associated with the female gender among females of reproductive age and tested the interaction between gender and age categorized as reproductive (18-44 years) and older (≥45 years); this model with age-gender interaction did not reveal differences in the association with gender by age. Non-Hispanic Black/African American and Hispanic participants had 1.8- and 1.2-fold greater odds of vaccine hesitancy compared to non-Hispanic White participants, respectively (Table 3). Overall, the magnitude of the strongest associations between demographic characteristics and vaccine hesitancy clustered around ORs of 2. Much stronger associations were found with geographical regions. Compared to California, even after adjustment for age, gender, and race/ethnicity, the odds of vaccine hesitancy were almost 6-fold greater in Florida and 4-fold greater in Louisiana (Table 3). Thus, the main effect analysis demonstrated that the geographical region has the strongest association with vaccine hesitancy.

Table 1. Population characteristics and vaccine hesitancy status in the MRCIS[a] (2020-2021).

Characteristics	Participants (N=3462), n (%)	Vaccine hesitancy	
		Not hesitant (n=1782, 51.5%), n (%)	Hesitant (n=1680, 48.5%), n (%)
Age (years)			
<25	216 (6.2)	121 (6.8)	95 (5.6)
25-34	478 (13.8)	229 (12.9)	249 (14.8)
35-44	586 (16.9)	292 (16.4)	294 (17.5)
45-54	737 (21.3)	381 (21.4)	356 (21.2)
55-64	775 (22.4)	421 (23.6)	354 (21.1)
≥65	665 (19.2)	334 (18.7)	331 (19.7)
Missing	5 (0.1)	4 (0.2)	1 (0.1)
Gender			
Female	2076 (60.0)	937 (52.6)	1139 (67.8)
Male	1379 (39.8)	840 (47.1)	539 (32.1)
Other	3 (0.1)	3 (0.2)	0
Missing	4 (0.1)	2 (0.1)	2 (0.1)
Race/ethnicity			
Hispanic/Latino	1684 (48.6)	869 (48.8)	815 (48.5)
Non-Hispanic Black/African American	896 (25.9)	398 (22.3)	498 (29.6)
Non-Hispanic White	740 (21.4)	429 (24.1)	311 (18.5)
Non-Hispanic other	107 (3.1)	67 (3.8)	40 (2.4)
Missing	35 (1.0)	19 (1.1)	16 (1.0)
HRSA[b] site			
California	971 (100.0)	701 (39.3)	270 (16.1)
Florida	971 (100.0)	318 (17.8)	653 (38.9)
Midwest (Illinois/Ohio)	510 (100.0)	350 (19.6)	160 (9.5)
Louisiana	1010 (100.0)	413 (23.2)	597 (35.5)

[a] MRCIS: Minority and Rural Coronavirus Insights Study.

[b] HRSA: Health Resources & Services Administration.

Table 2. Vaccine hesitancy estimates in the general population, as reported by the CDC[a] (2021), and in the underserved population, as observed by the MRCIS[b] (2020-2021).

Sites of data collection	Number of counties in each HRSA[c] site	CDC vaccine hesitance prevalence estimates in HRSA site counties	Expected vaccine hesitancy prevalence[d] (adjustment for time trend)[e]	MRCIS observed prevalence of vaccine hesitancy (95% CI)
California	6	8.2%-13.3%	9.7% (12.9%)	27.8% (25%-30.6%)
Midwest (Illinois/Ohio)	23	10.4%-23.7%	15.3% (20.3%)	31.8% (27.7%-35.8%)
Florida	8	16.0%-20.4%	18.2% (24.2%)	67.3% (64.3%-70.2%)
Louisiana	11	21.7%-27.9%	27.0% (36%)	59.1% (56.1%-62.1%)

[a] CDC: Centers for Disease Control and Prevention.

[b] MRCIS: Minority and Rural Coronavirus Insights Study.

[c] HRSA: Health Resources & Services Administration.

[d] Based on the CDC estimates for the counties of the participant’s residence (2021).

[e] Based on the previously published time trend estimates, the expected prevalence of vaccine hesitancy decreased by 1/3 between 2020 and 2021; therefore, we increase the estimate for the expected vaccine hesitancy to reflect what would be expected at the time of our survey.

Table 3. Association of vaccine hesitancy with demographic characteristics and geographical location, as observed by the MRCIS[a] (2020-2021).

Main Effect Model	Adjusted OR[b] (95% CI)
Age (years)	
18-24	1.33 (0.94-1.87)
25-34	2.29 (1.74-3.01)
35-44	1.59 (1.24-2.04)
45-54	1.38 (1.10-1.75)
55-64	1.12 (0.89-1.40)
≥65	Reference
Gender	
Female	1.67 (1.44-1.94)
Male	Reference
Race/ethnicity	
Hispanic/Latino	1.24 (0.99-1.55)
Non-Hispanic Black	1.79 (1.45-2.21)
Non-Hispanic White	Reference
Non-Hispanic other	0.93 (0.59-1.46)
HRSA[c] site	
California	Reference
Midwest (Illinois/Ohio)	1.11 (0.87-1.42)
Florida	5.81 (4.72-7.16)
Louisiana	4.03 (3.20-5.08)

[a] MRCIS: Minority and Rural Coronavirus Insights Study.

[b] OR: odds ratio.

[c] HRSA: Health Resources & Services Administration.

We further explored the adjusted associations between each demographic characteristic and vaccine hesitancy, considering the geographical differences (models with interaction terms, Tables 4 and 5). These models showed significant interactions (*P*<.05 for the interaction term) between the geographical region and 2 demographic characteristics, namely gender and race/ethnicity, not with age. The adjusted association with gender followed the same tendencies revealed by the crude analysis (Figure 2 and Table 4). Within each gender group, the association between the female gender and vaccine hesitancy was strongest in Florida, followed by Louisiana and the Midwest (Table 4). Within each region, the tendency of a stronger association with the female gender was obvious; however, only in Louisiana, the 95% CIs for the gender-specific estimates did not overlap, indicating sufficient precision to detect gender differences in this region.

The racial/ethnic differences in the adjusted estimates of vaccine hesitancy also followed the pattern found in the crude analysis (Figure 3). As compared to non-Hispanic White participants in California, all racial/ethnic groups in Florida and Louisiana had a greater association with vaccine hesitancy (Table 5). In California, non-Hispanic Black/African American participants clearly had greater odds of vaccine hesitancy compared to non-Hispanic White participants. However, in other regions, the likelihood that non-Hispanic Black participants would express vaccine hesitancy was similar to other racial/ethnic groups. The strongest association of vaccine hesitancy with being non-Hispanic White was observed in Louisiana, whereas the strongest association with being Hispanic was observed in Florida. Thus, in this study, the demographic patterns of vaccine hesitancy, especially racial/ethnic differences, strongly depended on local contextual factors defined by geographical region.

Table 4. Association between vaccine hesitancy and gender modified by geographical location, as observed by the MRCIS[a] (2020-2021).

Gender	California, OR[b] (95% CI)[c]	Midwest (Illinois/Ohio), OR (95% CI)[c]	Florida, OR (95% CI)[c]	Louisiana, OR (95% CI)[c]
Female	1.14 (0.86-1.52)	1.58 (1.14-2.17)	7.88 (5.96-10.41)	6.09 (4.55-8.14)
Male	Reference[d]	0.83 (0.56-1.23)	4.66 (3.41-6.35)	2.80 (2.03-3.85)

[a] MRCIS: Minority and Rural Coronavirus Insights Study.

[b] OR: odds ratio.

[c] Multivariable logistic regression models were adjusted for age and race/ethnicity.

[d] Males in California served as the common reference group.

Table 5. Association between vaccine hesitancy and race/ethnicity modified by geographical location, as observed by the MRCIS[a] (2020-2021).

Race/ethnicity	California, OR[b] (95% CI)[c]	Midwest (Illinois/Ohio), OR (95% CI)[c]	Florida, OR (95% CI)[c]	Louisiana, OR (95% CI)[c]
Hispanic	1.49 (0.92-2.41)	1.95 (1.15-3.30)	11.07 (6.93-17.70)	7.25 (3.50-15.02)
Non-Hispanic Black/African American	4.61 (2.71-7.84)	2.60 (1.44-4.69)	8.07 (3.90-16.67)	8.88 (5.48-14.37)
Non-Hispanic Whites	Reference[d]	1.53 (0.85-2.74)	5.38 (2.81-10.29)	6.68 (4.09-10.92)
Non-Hispanic other	0.99 (0.39-2.52)	1.84 (0.62-5.41)	5.63 (2.26-14.02)	8.32 (3.19-21.68)

[a] MRCIS: Minority and Rural Coronavirus Insights Study.

[b] OR: odds ratio.

[c] Multivariable logistic regression models were adjusted for age and gender.

[d] Non-Hispanic White participants in California served as the common reference group.

Discussion

Principal Findings

Our study of COVID-19 vaccine hesitancy in underserved populations revealed several important findings. We found profound differences in vaccine hesitancy between geographical regions (Table 1). These observed differences might be expected based on the estimates published by the CDC (Table 2). However, our estimates for the MRCIS population showed a clear tendency of greater vaccine hesitancy than expected in the general population. This finding is important because the underserved population is at a greater risk of worse outcomes of this infection [5-8]. Thus, our study highlights the need to clearly communicate the benefits and risks of the vaccine to this high-risk population.

Another important finding is the difference in demographic patterns of vaccine hesitancy by geographical region. Although in California, the greatest vaccine hesitancy was among non-Hispanic Black/African American participants, this finding cannot be extrapolated to Louisiana and Florida. Similarly, the highest hesitancy prevalence among Hispanic participants in Florida cannot be extrapolated to other regions. These findings suggest that local contextual factors are major contributors to vaccine hesitancy in contrast to demographic characteristics. Future studies of vaccine hesitancy and other health behaviors should consider the local specifics. Additionally, these findings could prove important to developing effective local public health initiatives as streamlined, catch-all approaches may lack the nuance and specificity needed to impact vulnerable communities.

Results of prior studies support our findings, emphasizing the importance of local contextual factors in driving racial/ethnic patterns of vaccine hesitancy. A prior study conducted in the underserved population of North Carolina found the highest prevalence among White (74%), followed by Black (62.7%) and Hispanic (59.5%) participants, whereas a study in the Delaware underserved community found the greatest hesitancy among Black participants (45%). The online study conducted by the market research firm YouGov [17] found a similar prevalence of vaccine hesitancy among Black and White participants in 2020, once again confirming that racial/ethnic patterns strongly depend on the study population examined and cannot be extrapolated from one population to another.

We found that vaccine hesitancy tended to be lower among participants younger than 25 years and older than 55 years in the Midwest, Florida, and Louisiana (Figure 1 and Table 3). Evaluating the interaction between age and geographical region strongly suggests that in this study, the age patterns adjusted for gender and race/ethnicity did not differ by geography. The highest vaccine prevalence and the strongest association with age were found among ages 25-34 years. Similarly, among US Facebook respondents in January-May 2021, this age group had the highest vaccine hesitancy prevalence [18]. Another study conducted for the New York City Transport Workers Union found that adults aged 50 years and above were less likely to be COVID-19 vaccine hesitant than their younger counterparts [19]. These investigations examined study populations different from this report both socially and geographically. Thus, the age pattern with the high prevalence of vaccine hesitancy among ages 25-34 years appears to be consistent and less sensitive to the local contextual factors as opposed to racial/ethnic patterns.

The tendency of females to be more likely vaccine hesitant was apparent in all geographical regions but especially pronounced in Louisiana (Figure 2 and Table 4). Similar results were found by the US Facebook study, with females being more likely hesitant compared to men [18]. In addition, the analysis of a nationally representative sample from the 2021 Household Pulse Survey, administered by the US Census Bureau, showed greater vaccine hesitancy among US women [20]. Similar to the age patterns, gender differences in vaccine hesitancy may not be as sensitive to the local contextual factors and, therefore, are more generalizable to different population segments in the United States compared to racial/ethnic differences.

Limitations

There are several limitations to our study. Because the participants were drawn from volunteers and represented a convenience sample, the results may be liable to self-selection bias. To discuss whether such a bias potentially increased or decreased participation among the vaccine-hesitant individuals, we considered the major incentive for participation (ie, the availability of COVID-19 testing). Such an incentive suggests that volunteers were concerned about their infection status. However, neither the COVID-19 incidence rate nor mortality correlated with vaccine hesitancy worldwide [21]. Thus, it is unclear whether a potential self-selection bias was likely to increase or decrease participation among the hesitant subgroup, given the interest in the COVID-19 test. Because similar geographical patterns were found by the CDC data (Table 2), likely, self-selection did not distort the observed differences by geographical region. Similar age and gender patterns found in our and other studies suggest that either self-selection bias may have influenced all these studies or a potential self-selection bias did not distort these associations. Finally, self-selection bias could influence racial/ethnic patterns. However, the direction of the bias should have been specific to contextual local factors governing the observed geographical differences. Overall, even considering the likelihood of self-selection bias, our study clearly demonstrates the importance of local contextual factors in driving COVID-19 vaccine hesitancy. Another limitation of our study is the lack of additional data to further investigate the reasons for the revealed differences in vaccine hesitancy. Which local contextual factors are important remains unknown. We will focus our future investigations on answering this question.

Conclusion

In summary, our findings strongly suggest that local contextual factors drive the overall level of vaccine hesitancy in different regions, as demonstrated by the low prevalence in all demographic groups in California compared to Florida and Louisiana. Importantly, in all geographical regions, our estimates of vaccine hesitancy in the underserved population tended to be greater than what was expected in the general population in those same regions. Future investigations on this topic should include recording reasons for vaccine hesitancy to gain further insight into this subject. The same vaccine hesitancy factors will likely play a significant role in curbing future epidemics through vaccinations.

Acknowledgments

The National Minority Quality Forum was responsible for the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, approval, and the decision to submit the manuscript for publication. Quest Diagnostics provided in-kind and staffing support for the diagnostic tests during the enrollment process. The Centene Corporation provided startup support for the study.

Conflicts of Interest

None declared.

References

1. Ten threats to global health in 2019. World Health Organization. 2019. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2022-11-15]
2. Bussink-Voorend D, Hautvast JLA, Vandeberg L, Visser O, Hulscher MEJL. A systematic literature review to clarify the concept of vaccine hesitancy. Nat Hum Behav 2022 Dec 22;6(12):1634-1648. [doi: 10.1038/s41562-022-01431-6] [Medline: 35995837]

3. Estimates of vaccine hesitancy for COVID-19. Centers for Disease Control and Prevention. URL: https://data.cdc.gov/stories/s/Vaccine-Hesitancy-for-COVID-19/cnd2-a6zw/ [accessed 2023-03-20]
4. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. No populations left behind: vaccine hesitancy and equitable diffusion of effective COVID-19 vaccines. J Gen Intern Med 2021 Jul;36(7):2130-2133 [FREE Full text] [doi: 10.1007/s11606-021-06698-5] [Medline: 33754319]
5. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 - COVID-NET, 14 states, March 1-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Apr 17;69(15):458-464 [FREE Full text] [doi: 10.15585/mmwr.mm6915e3] [Medline: 32298251]
6. Risk for COVID-19 infection, hospitalization, and death by race/ethnicity. Centers for Disease Control and Prevention. 2022 Dec 28. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html [accessed 2023-03-20]
7. Abedi V, Olulana O, Avula V, Chaudhary D, Khan A, Shahjouei S, et al. Racial, economic, and health inequality and COVID-19 infection in the United States. J Racial Ethn Health Disparities 2021 Jun 01;8(3):732-742 [FREE Full text] [doi: 10.1007/s40615-020-00833-4] [Medline: 32875535]
8. Gross CP, Essien UR, Pasha S, Gross JR, Wang S, Nunez-Smith M. Racial and ethnic disparities in population-level Covid-19 mortality. J Gen Intern Med 2020 Oct 04;35(10):3097-3099 [FREE Full text] [doi: 10.1007/s11606-020-06081-w] [Medline: 32754782]
9. Nguyen LH, Joshi AD, Drew DA, Merino J, Ma W, Lo C, COPE Consortium. Self-reported COVID-19 vaccine hesitancy and uptake among participants from different racial and ethnic groups in the United States and United Kingdom. Nat Commun 2022 Feb 01;13(1):636 [FREE Full text] [doi: 10.1038/s41467-022-28200-3] [Medline: 35105869]
10. Grasselli G, Greco M, Zanella A, Albano G, Antonelli M, Bellani G, COVID-19 Lombardy ICU Network. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. JAMA Intern Med 2020 Oct 01;180(10):1345-1355 [FREE Full text] [doi: 10.1001/jamainternmed.2020.3539] [Medline: 32667669]
11. Schoch-Spana M, Brunson E, Long R, Ruth A, Ravi S, Trotochaud M, et al. The public's role in COVID-19 vaccination: human-centered recommendations to enhance pandemic vaccine awareness, access, and acceptance in the United States. Vaccine 2021 Sep 24;39(40):6004-6012 [FREE Full text] [doi: 10.1016/j.vaccine.2020.10.059] [Medline: 33160755]
12. Wang SX, Bell-Rogers N, FNP-c, Dillard D, Harrington MA. COVID-19 vaccine hesitancy in Delaware's underserved communities. Dela J Public Health 2021 Sep;7(4):168-175 [FREE Full text] [doi: 10.32481/djph.2021.09.022] [Medline: 34604782]
13. Doherty IA, Pilkington W, Brown L, Billings V, Hoffler U, Paulin L, et al. COVID-19 vaccine hesitancy in underserved communities of North Carolina. PLoS One 2021;16(11):e0248542 [FREE Full text] [doi: 10.1371/journal.pone.0248542] [Medline: 34723973]
14. Mora A, Lewnard J, Kogut K, Rauch S, Morga N, Jewell N. Impact of the COVID-19 pandemic and vaccine hesitancy among farmworkers from Monterey County, California. medRxiv. 2020.12.18.2024 Preprint posted online December 22, 2020 . [doi: 10.1101/2020.12.18.20248518]
15. HUD USPS zip code crosswalk files. Office of Policy Development and Research (PD&R). URL: https://www.huduser.gov/portal/datasets/usps_crosswalk.html [accessed 2023-03-20]
16. King W, Rubinstein M, Reinhart A, Mejia R. Time trends and factors related to COVID-19 vaccine hesitancy from January-May among US adults: findings from a large-scale national survey. medRxiv Preprint posted online October 25, 2021. [doi: 10.1101/2021.07.20.21260795]
17. Padamsee TJ, Bond RM, Dixon GN, Hovick SR, Na K, Nisbet EC, et al. Changes in COVID-19 vaccine hesitancy among black and white individuals in the US. JAMA Netw Open 2022 Jan 04;5(1):e2144470 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.44470] [Medline: 35061038]
18. Nguyen QC, Yardi I, Gutierrez FXM, Mane H, Yue X. Leveraging 13 million responses to the U.S. COVID-19 Trends and Impact Survey to examine vaccine hesitancy, vaccination, and mask wearing, January 2021-February 2022. BMC Public Health 2022 Oct 13;22(1):1911 [FREE Full text] [doi: 10.1186/s12889-022-14286-3] [Medline: 36229804]
19. Meltzer GY, Harris J, Hefner M, Lanternier P, Gershon RRM, Vlahov D, et al. Associations between COVID-19 vaccine hesitancy and socio-spatial factors in NYC transit workers 50 years and older. Int J Aging Hum Dev 2023 Jan;96(1):76-90 [FREE Full text] [doi: 10.1177/00914150221106709] [Medline: 35702009]
20. Morales D, Beltran T, Morales S. Gender, socioeconomic status, and COVID-19 vaccine hesitancy in the US: an intersectionality approach. Sociol Health Illn 2022;44(6):953-971 [FREE Full text] [doi: 10.1111/1467-9566.13474]
21. Lazarus JV, Wyka K, White TM, Picchio CA, Rabin K, Ratzan SC, et al. Revisiting COVID-19 vaccine hesitancy around the world using data from 23 countries in 2021. Nat Commun 2022 Jul 01;13(1):3801 [FREE Full text] [doi: 10.1038/s41467-022-31441-x] [Medline: 35778396]

Abbreviations

AIC: Akaike information criterion

CDC: Centers for Disease Control and Prevention
FQHC: federally qualified health centers
HRSA: Health Resources & Services Administration
MRCIS: Minority and Rural Coronavirus Insights Study
OR: odds ratio
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 10.12.21; peer-reviewed by M Behzadifar, C Campos-Castillo, MP Nowalk; comments to author 09.04.22; revised version received 31.05.22; accepted 22.02.23; published 27.04.23.

Please cite as:
Matas JL, Landry LG, Lee L, Hansel S, Coudray MS, Mata-McMurry LV, Chalasani N, Xu L, Stair T, Edwards C, Puckrein G, Meyer W, Wiltz G, Sampson M, Gregerson P, Barron C, Marable J, Akinboboye O, Il'yasova D
Demographic Determinants and Geographical Variability of COVID-19 Vaccine Hesitancy in Underserved Communities: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e34163
URL: https://publichealth.jmir.org/2023/1/e34163
doi:10.2196/34163
PMID:36811869

Original Paper

The Use of Crisis Services Following the Mass School Shooting in Uvalde, Texas: Quasi-Experimental Event Study

Kirsty J Weitzel[1], MSc; Robert F Chew[1], MSc; Adam Bryant Miller[2,3], PhD; Caroline W Oppenheimer[2], PhD; Ashley Lowe[4], MPH; Anna Yaros[2], PhD

[1]Center for Data Science and AI, RTI International, Research Triangle Park, NC, United States

[2]Center on Social Determinants, Risk Behaviors & Prevention Science, RTI International, Research Triangle Park, NC, United States

[3]Department of Psychology & Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]Transformative Research Unit for Equity, RTI International, Research Triangle Park, NC, United States

Corresponding Author:
Kirsty J Weitzel, MSc
Center for Data Science and AI
RTI International
3040 East Cornwallis Road
PO Box 12194
Research Triangle Park, NC
United States
Phone: 1 9195416000
Email: kweitzel@rti.org

Abstract

Background: Mass shootings result in widespread psychological trauma for survivors and members of the affected community. However, less is known about the broader effects of indirect exposure (eg, media) to mass shootings. Crisis lines offer a unique opportunity to examine real-time data on the widespread psychological effects of mass shootings.

Objective: Crisis Text Line is a not-for-profit company that provides 24/7 confidential SMS text message–based mental health support and crisis intervention service. This study examines changes in the volume and composition of firearm-related conversations at Crisis Text Line before and after the mass school shooting at Robb Elementary School on May 24, 2022, in Uvalde, Texas.

Methods: A quasi-experimental event study design was used to compare the actual volume of firearm-related conversations received by Crisis Text Line post shooting to forecasted firearm conversation volume under the counterfactual scenario that a shooting had not occurred. Conversations related to firearms were identified among all conversations using keyword searches. Firearm conversation volume was predicted using a seasonal autoregressive integrated moving average model trained on the 3 months of data leading up to the shooting. Additionally, proportions of issue tags (topics coded post conversation by volunteer crisis counselors at Crisis Text Line after the exchange) were compared in the 4 days before (n=251) and after (n=417) the shooting to assess changes in conversation characteristics. The 4-day window was chosen to reflect the number of days conversation volume remained above forecasted levels.

Results: There was a significant increase in the number of conversations mentioning firearms following the shooting, with the largest spike (compared to forecasted numbers) occurring the day after the shooting (n=159) on May 25, 2022. By May 28, the volume reverted to within the 95% CI of the forecasted volume (n=77). Within firearm conversations, “grief” issue tags showed a significant increase in proportion in the week following the shooting, while “isolation/loneliness,” “relationships,” and “suicide” issue tags showed a significant decrease in proportions the week following the shooting.

Conclusions: The results suggest that the Uvalde school shooting may have contributed to an increase in demand for crisis services, above what would be expected given historical trends. Additionally, we found that these firearm-related crises conversations immediately post event are more likely to be related to grief and less likely to be related to suicide, loneliness, and relationships. Our findings provide some of the first data showing the real-time repercussions for the broader population exposed to school shooting events. This work adds to a growing evidence base documenting and measuring the rippling effects of mass shootings outside of those directly impacted.

(JMIR Public Health Surveill 2023;9:e42811) doi:10.2196/42811

KEYWORDS

mass shooting; Crisis Text Line; firearms; mental health; suicide; trauma; indirect exposure; public health; crisis management; mental health support; mass casualty; shooting; crisis service

Introduction

Mass shootings result in widespread psychological distress and trauma symptoms for survivors and members of the impacted community [1]. Research demonstrates that following a mass shooting, higher rates of psychological symptoms (eg, depressive symptoms, increased feelings of fear, and decreased perceptions of safety) are found in the immediately affected community [2]. While there is emerging evidence suggesting that mass shootings may have negative mental health consequences that reach beyond local communities to broader populations through indirect exposure (eg, conversations, media exposure) [2], information about immediate effects of mass shootings on broader public well-being is nearly absent in the literature. One study showed that teachers who had increased exposure to school shooting–related media had increased experiences of secondary trauma [3]. Advancing understanding of indirect exposure and resulting mental health–related issues will inform future public health response.

There are several challenges associated with studying the effects of indirect exposure to mass shootings using traditional survey methods. Mass shooting events are unpredictable low-base rate events. Scheduled data collection of longitudinal or repeated cross-sectional studies may not align with when mass shootings occur, leading to excessively inefficient and costly survey operations. Furthermore, prolonged periods between exposure and data collection may lead to recall bias [4], reducing the quality of responses and the ability to make strong inferences. Recently, alternative study designs have been used to address these methodological challenges. For example, studies have analyzed the content of social media posts before and after mass shooting events to show increases in expression of emotions in the general population [5,6].

Data from crisis lines provide a unique opportunity to examine real-time data on the widespread psychological effects of mass shootings. Crisis Text Line (CTL) is a not-for-profit company that provides 24/7 confidential crisis counseling via SMS text messaging and WhatsApp. The majority of texters served by CTL in the United States are younger than 25 years (76% aged <25 years) [7]. CTL provides rich data specifically focused on mental health–related crises in the moments that they occur. Previous studies have used time series research designs with deidentified CTL data to analyze changes in the number and types of crises due to exposure to *13 Reasons Why* on Netflix [8], celebrity suicide deaths [9], and Hurricane Florence [10]. These prior investigations collectively demonstrate that poignant events that have widespread coverage on various media outlets (including social media) have measurable impact on the larger corpus of individuals texting CTL.

In this study, we investigated CTL's anonymized conversations related to firearm violence, including mass shootings, suicide, and other firearm-related issues. To our knowledge, no existing study has examined firearm-related crises in CTL conversations or other crisis lines. We examined changes in the volume and composition of CTL firearm-related texts before and after the Robb Elementary School mass school shooting on May 24, 2022, in Uvalde, Texas. Based on previous investigations with deidentified CTL data and national events [8-10], we hypothesized that conversations mentioning firearms would increase in the days following the mass shooting. Because work has not examined the broader effects of indirect exposure to mass shootings on mental health, we did not have a specific hypothesis about how social and mental health–related stressors may change before and after a mass shooting event. The objective of this study was to provide direct evidence of the effects of mass shootings on the broader public's mental health.

Methods

Overview

Anonymized and deidentified full crisis conversation transcripts and postconversation survey data completed by volunteer crisis counselors (vCCs) was obtained from CTL for the dates surrounding the Robb Elementary School shooting (March 1 to May 30, 2022). We filtered the anonymized and deidentified individual CTL conversations by 123 firearm-related keywords (eg, "gun," "shooting") to create a corpus of firearm-related crisis conversations. The keyword list started with common terms for firearms and was expanded using a word embedding similarity search [11] to discover firearm terms in the conversation text that were not present in the initial list (eg, "gsw," "9 mm"), expanding the number of keyword terms from 23 to 37. Word embeddings were generated using the Word2Vec algorithm [12] trained on all CTL conversations from September 2018 to August 2021 (N=2,539,460 conversations), which is the date range of focus for additional ongoing projects with the CTL data. Keywords were expanded to 123 words by examining conversations where vCCs' postconversation survey notes had a firearm keyword listed as the potential means of suicide but that were not already identified based on the conversation text using our existing keyword list. A full list of keywords is available in Multimedia Appendix 1.

A quasi-experimental event study design [13] was used to compare the actual firearm conversation volume post event to forecasted firearm conversation volume under the counterfactual scenario that a shooting had not occurred. Expected volumes were estimated using a seasonal autoregressive integrated moving average time series model with an order of (0, 1, 1) and seasonal order of (1, 1, 1, 7) on daily counts for a pre-event period (March 1 to May 23, 2022) to forecast the firearm conversation volume post event (May 4 to May 30, 2022). Other mass shooting events in the pre-event period were incorporated into the model using an exogenous dummy variable (1=mass shooting, 0=no shooting). Other mass shooting events were identified using the Gun Violence Archive [14] and the US Congress definition of mass public shooting [15]. A start date of March 1, 2022, was chosen to focus on recent months, though a start date of January 1, 2022, was also tested with similar results. An end date of May 30, 2022, was chosen to exclude the June 1, 2022, mass shooting in Tulsa, OK.

Lastly, we compared proportions of conversation issue tags (characteristics of the text conversation coded by CTL vCCs during/after the exchange) 4 days before (n=251) and after (n=417) the shooting to assess changes in conversation characteristics. The duration of the time before and after the event was chosen empirically to represent days in which the actual volume of firearm-related conversations exceeded the expected. Chi-square tests were used to test the hypothesis that issue tag proportions differed in the pre- and postevent periods.

Data Management

Data from CTL was accessed on a secure remote server hosted by CTL. All data used in this study will be deleted upon completion of the broader Centers for Disease Control and Prevention–funded research project.

Ethical Considerations

This project was reviewed by the RTI Office of Research Protection Institutional Review Board (IRB; STUDY00021510) and deemed to fall under the category of Not Human Subjects Research (public health practices and surveillance activities). Data used within this study are collected by existing infrastructure at CTL as part of the crisis services they provide. All data gathered from texters including text conversations and postconversation survey data provided by CTL are anonymized and deidentified before transfer to the research team for analysis. Texters agree to the use of their anonymized information for research purposes as part of the CTL's terms of service automatically shared before being connected with a vCC [16]. Anonymized texter information is shared for public good in aggregate at CrisisTrends.org and with select researchers following a rigorous screening process and IRB approval, and texters can request that their conversations be removed from CTL's systems at any time following their conversations.

Results

We observed a significant increase in the number of conversations mentioning firearms following the shooting at Robb Elementary School (Figure 1). The largest spike in conversations mentioning firearms occurred the day after the shooting (n=159), May 25, 2022, with volume reverting to within the 95% CI of forecasted volume 4 days after the shooting on May 28 (n=77).

Table 1 reports the frequency and proportions of issue tags for conversations in the immediate 4-day pre- and postevent period. For firearm conversations during the postevent period, the count of all issue tags increased, except for the "COVID-19" issue tag. To control for increases in tags post event, we examined the percentage of total firearm conversations in the pre- and postevent period that each unique tag was used. Comparing pre-event to postevent, the "grief" issue tag increased significantly in the days following the shooting (*d*=0.06; *P*=.03), while the proportions for the "isolation/loneliness" (*d*=–0.07; *P*=.04), "relationships" (*d*=–0.11; *P*<.001), and "suicide" (*d*=–0.14; *P*<.001) issue tags decreased in the days following the shooting.

Figure 1. Count of Crisis Text Line firearm-related conversations following the Robb Elementary School shooting.

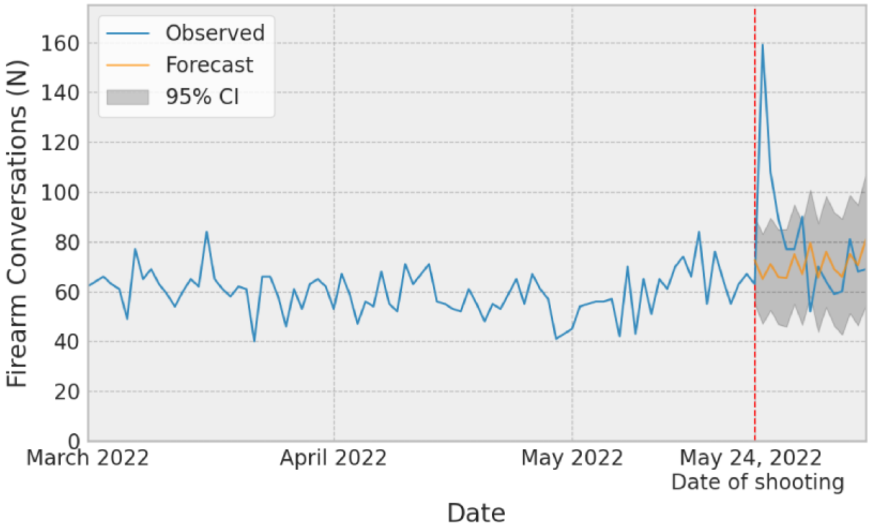


Table 1. Issue tag proportions for conversations prior and in the week following the shooting sorted by pre-post difference in proportions.

Issue tag	Count before, n	Count after, n	Pre-event proportion (n=251)	Postevent proportion (n=417)	Difference	*P* value
Grief	19	55	0.08	0.13	0.06	*.03* [a]
Anxiety	79	154	0.31	0.37	0.05	.15
Abuse (emotional)	8	17	0.03	0.04	0.01	.56
Gender/sexual identity	3	7	0.01	0.02	0.00	.62
Racism	4	8	0.02	0.02	0.00	.76
Depression/sadness	6	11	0.37	0.37	0.00	.99
Eating disorder/body image	92	154	0.02	0.02	0.00	.99
Abuse (physical)	6	8	0.02	0.02	0.00	.68
COVID-19	3	2	0.01	0.00	–0.01	.03
Abuse (unspecified)	7	8	0.03	0.02	–0.01	.46
Abuse (sexual)	10	12	0.04	0.03	–0.01	.44
Bullying	10	11	0.05	0.03	–0.01	.36
Substance use	12	14	0.04	0.02	–0.02	.24
Self-harm	27	34	0.11	0.08	–0.03	.26
Isolation/loneliness	69	87	0.27	0.21	–0.07	*.04*
Relationships	95	112	0.39	0.26	–0.13	*<.001*
Suicide	118	141	0.47	0.34	–0.13	*<.001*

[a] Italics indicate values that are statistically significant.

Discussion

The objective of this investigation was to examine the effect of a mass shooting event on the broader public's mental health. To do so, we examined the immediate effects of a widely publicized mass shooting in the United States on volume of mental health crisis text conversations that mentioned firearms. We also examined the psychosocial context of those text conversations by investigating changes in issue tags associated with firearm conversations as coded by the vCCs. Consistent with our hypothesis, we found that firearm-related crisis conversations following the Uvalde school shooting significantly increased. In addition to this main finding, we observed an increase in the proportion of grief-related conversations that mentioned firearms and a decrease in suicide-related conversations that mentioned firearms. We also observed a decrease in firearm conversations associated with relationships or feelings of isolation/loneliness following the Uvalde school shooting. Overall, we saw evidence that a widely publicized mass shooting influenced the broader public's mental health beyond the specific location of the mass shooting event.

Our innovative methodological approach of analyzing real-time mental health crises through CTL conversations allowed us to address a critical gap in knowledge of the psychological effects of mass shootings on remote populations of individuals who are not directly exposed to the mass shooting event. Only a couple of prior studies have investigated the effects of mass shooting events on youth populations beyond the communities directly involved in these events, and these studies have mainly focused on youth expressions of emotion such as fear. However, our study findings may have the potential to have a significant impact on mental health crises (beyond general expressions of negative emotions) among the public. Interestingly, the increase in crisis conversation volume existed even when accounting for other mass shootings in the study period. It is possible that the Uvalde shooting has a particularly strong impact on users of CTL because school shootings (relative to other forms of mass shootings) appear to significantly affect youth mental health [17], and the CTL population is largely composed of youth [7]. Of course, other confounding factors such as the amount of media coverage or social media mentions may also be influential.

The increase in proportion of grief-related conversations, along with the decrease in suicide-related conversations, further suggest that firearm crisis conversations post event may be focused more on loss rather than suicide as is typically more common in firearm-related mental health crisis conversations. These findings are consistent with prior research showing that grief is a common reaction immediately following school shootings [18]. We also observed a decrease in firearm conversations associated with isolation and relationship issues. Although the data does not illustrate why this shift occurred, we speculate that an increase in social solidarity after the mass shooting event leads to fewer crises due to isolation and relationship factors. This possibility is consistent with prior studies suggesting that community members often reach out to provide more emotional support to one another after traumatic events [19]. Together, our results demonstrate that firearm-related mental health crisis conversations immediately following a mass shooting differ from typical firearm conversations in that they are more likely to be associated with grief issues and less likely to be associated with suicide, isolation, and relationship issues.

Several limitations of the study should be noted. First, we assume that the firearm conversations represent users that were not directly impacted by the Uvalde shooting (survivors, direct family members, etc) due to the relatively smaller number of people expected to incur direct versus indirect effects. Future studies should use geographic indicators to increase the chance of examining those directly related to a mass shooting event. Second, our study design does not address why the text volume associated with the Uvalde school shooting is greater than other mass shootings observed in the study period. Future research studying the magnitude of indirect impacts of mass shootings across different types of events could greatly enhance the literature, for example, the impact of school shootings compared to nonschool shootings. Additionally, this study does not attempt to quantify the impact of increased awareness of or referrals to CTL due to the Uvalde shooting. Lastly, issue tags were designated by the vCCs and thus relatively subjective. Future work could study the reliability of vCCs' tagging to help better understand how mass shootings may be related to increases in firearm-related mental health crises. Overall, our findings support two main conclusions. First, the Uvalde, TX mass school shooting increased firearm-related text conversations at a national mental health crisis line even after factoring in the effects of other mass shootings that occurred proximally in time. Our findings provide some of the first data showing the real-time repercussions for the broader population exposed to school shooting events. Public health officials must recognize that the reverberations of school shootings on mental health extend beyond just the immediate community where the event takes place. Second, these firearm-related crises conversations may be most frequently related to grief. Thus, findings suggest that public health interventions that target feelings of grief may have the potential to reduce mental health crises immediately following a school mass shooting. Overall, the findings lay the foundation for future research on effective public mental health interventions that may be implemented after school mass shootings.

Acknowledgments

This work was funded by grant R01CE003295 from the Centers for Disease Control and Prevention. We thank Crisis Text Line for making anonymized and deidentified data available to our team for research purposes.

Data Availability

Data from Crisis Text Line used in this research is not available publicly due to concerns of privacy and anonymity of texters who use Crisis Text Line's service. Aggregated anonymized data from more than 7 million conversations with people in crisis supported by Crisis Text Line is available at CrisisTrends.org.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Firearm keywords used in conversation identification.
[DOCX File , 23 KB - publichealth_v9i1e42811_app1.docx]

References

1. Hoffman B. Inside Terrorism: Third Edition. New York, NY: Columbia University Press; 2017.
2. Lowe SR, Galea S. The mental health consequences of mass shootings. Trauma Violence Abuse 2017 Jan;18(1):62-82. [doi: 10.1177/1524838015591572] [Medline: 26084284]
3. Stevens T, Barnard‐Brak L, Roberts B, Acosta R, Wilburn S. Aggression toward teachers, interaction with school shooting media, and secondary trauma: lockdown drills as moderator. Psychol Schs 2019 Dec 18;57(4):583-605. [doi: 10.1002/pits.22329]
4. Coughlin SS. Recall bias in epidemiologic studies. J Clin Epidemiol 1990;43(1):87-91. [doi: 10.1016/0895-4356(90)90060-3] [Medline: 2319285]
5. Jones NM, Wojcik SP, Sweeting J, Silver RC. Tweeting negative emotion: an investigation of Twitter data in the aftermath of violence on college campuses. Psychol Methods 2016 Dec;21(4):526-541 [FREE Full text] [doi: 10.1037/met0000099] [Medline: 27918180]
6. Jones NM, Brymer M, Silver RC. Using big data to study the impact of mass violence: opportunities for the traumatic stress field. J Trauma Stress 2019 Oct;32(5):653-663. [doi: 10.1002/jts.22434] [Medline: 31593620]
7. Pisani AR, Gould MS, Gallo C, Ertefaie A, Kelberman C, Harrington D, et al. Individuals who text crisis text line: key characteristics and opportunities for suicide prevention. Suicide Life Threat Behav 2022 Jun;52(3):567-582 [FREE Full text] [doi: 10.1111/sltb.12872] [Medline: 35615898]

8. Thompson LK, Michael KD, Runkle J, Sugg MM. Crisis Text Line use following the release of Netflix series 13 Reasons Why Season 1: time-series analysis of help-seeking behavior in youth. Prev Med Rep 2019 Jun;14:100825 [FREE Full text] [doi: 10.1016/j.pmedr.2019.100825] [Medline: 30828539]
9. Sugg MM, Michael KD, Stevens SE, Filbin R, Weiser J, Runkle JD. Crisis text patterns in youth following the release of Season 2 and celebrity suicides: a case study of summer 2018. Prev Med Rep 2019 Dec;16:100999 [FREE Full text] [doi: 10.1016/j.pmedr.2019.100999] [Medline: 31750076]
10. Runkle JD, Michael KD, Stevens SE, Sugg MM. Quasi-experimental evaluation of text-based crisis patterns in youth following Hurricane Florence in the Carolinas, 2018. Sci Total Environ 2021 Jan 01;750:141702. [doi: 10.1016/j.scitotenv.2020.141702] [Medline: 32861078]
11. Diaz F, Mitra B, Craswell N. Query expansion with locally-trained word embeddings. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016 Presented at: 54th Annual Meeting of the Association for Computational Linguistics; August 2016; Berlin, Germany p. 367-377. [doi: 10.18653/v1/p16-1035]
12. Mikolov T, Chen K, Corado G, Dean J. Efficient estimation of word representations in vector space. arXiv. Preprint posted online on January 16, 2013. [doi: 10.48550/arXiv.1301.3781]
13. Huntington-Klein N. The Effect: An Introduction to Research Design and Causality. URL: https://theeffectbook.net/ [accessed 2022-09-06]
14. Mass shootings. Gun Violence Archive. URL: https://www.gunviolencearchive.org/mass-shooting [accessed 2022-09-07]
15. Public law 112–265—Jan. 14, 2013. Library of Congress. URL: https://www.congress.gov/112/plaws/publ265/PLAW-112publ265.pdf [accessed 2022-09-06]
16. Crisis Text Line. URL: https://www.crisistextline.org/privacy/ [accessed 2023-01-20]
17. Cimolai V, Schmitz J, Sood AB. Effects of mass shootings on the mental health of children and adolescents. Curr Psychiatry Rep 2021 Feb 11;23(3):12. [doi: 10.1007/s11920-021-01222-2] [Medline: 33570688]
18. Wayment HA, Silver RC. Grief and solidarity reactions 1 week after an on-campus shooting. J Interpers Violence 2021 Mar;36(5-6):NP2423-NP2442. [doi: 10.1177/0886260518766431] [Medline: 29589521]
19. Matsubayashi T, Sawada Y, Ueda M. Natural disasters and suicide: evidence from Japan. Soc Sci Med 2013 Apr;82:126-133. [doi: 10.1016/j.socscimed.2012.12.021] [Medline: 23305723]

Abbreviations

CTL: Crisis Text Line
vCC: volunteer crisis counselor

Edited by A Mavragani; submitted 20.09.22; peer-reviewed by A Teles, IH Song, C Baxter; comments to author 25.11.22; revised version received 16.12.22; accepted 22.12.22; published 08.02.23.

Please cite as:
Weitzel KJ, Chew RF, Miller AB, Oppenheimer CW, Lowe A, Yaros A
The Use of Crisis Services Following the Mass School Shooting in Uvalde, Texas: Quasi-Experimental Event Study
JMIR Public Health Surveill 2023;9:e42811
URL: https://publichealth.jmir.org/2023/1/e42811
doi:10.2196/42811
PMID:36753321

Original Paper

Impacts of GlobalConsent, a Web-Based Social Norms Edutainment Program, on Sexually Violent Behavior and Bystander Behavior Among University Men in Vietnam: Randomized Controlled Trial

Kathryn M Yount[1], PhD; Yuk Fai Cheong[1], PhD; Irina Bergenfeld[1], MPH; Quach Thu Trang[2], MA; Jessica M Sales[1], PhD; Yiman Li[1], MScPH; Tran Hung Minh[2], MD

[1]Rollins School of Public Health, Emory University, Atlanta, GA, United States
[2]Center for Creative Initiatives in Health and Population, Hanoi, Vietnam

Corresponding Author:
Kathryn M Yount, PhD
Rollins School of Public Health
Emory University
1518 Clifton Rd NE
Atlanta, GA, 30322
United States
Phone: 1 4047278511
Email: kyount@emory.edu

Abstract

Background: *Sexual violence* against women is prevalent worldwide. Prevention programs that treat men as allies and integrate a bystander framework are emerging in lower income settings, but evidence of their effectiveness is conflicting.

Objective: This study aimed to test the impact of GlobalConsent on sexually violent behavior and prosocial bystander behavior among university men in Vietnam.

Methods: We used a double-blind, parallel intervention versus control group design with 1:1 randomization at 2 universities. A total of 793 consenting heterosexual or bisexual men aged 18-24 years who matriculated in September 2019 were enrolled and assigned randomly to GlobalConsent or an attention-control adolescent health education (AHEAD) program. GlobalConsent is an adapted, theory-based, 6-module web-based intervention with diverse behavior change techniques and a locally produced serial drama. AHEAD is a customized, 6-module attention-control program on adolescent health. Both the programs were delivered to computers and smartphones over 12 weeks. Self-reported sexually violent behaviors toward women in the prior 6 months and prosocial bystander behaviors in the prior year were measured at 0, 6, and 12 months.

Results: More than 92.7% (735/793) of men in both study arms completed at least 1 program module, and >90.2% (715/793) of men completed all 6 modules. At baseline, a notable percentage of men reported any sexually violent behavior (GlobalConsent: 123/396, 31.1%; AHEAD: 103/397, 25.9%) in the prior 6 months. Among men receiving GlobalConsent, the odds of reporting a high level (at least 2 acts) of sexually violent behavior at the endline were 1.3 times the odds at baseline. Among men receiving AHEAD, the corresponding odds ratio was higher at 2.7. The odds of reporting any bystander behavior at endline were 0.7 times the odds at baseline for GlobalConsent, and the corresponding odds ratio for AHEAD was lower at 0.5.

Conclusions: Compared with a health attention-control condition, GlobalConsent has sustained favorable impacts on sexually violent behavior and prosocial bystander behavior among matriculating university men in Vietnam, who would otherwise face increasing risks of sexually violent behavior. GlobalConsent shows promise for national scale-up and regional adaptations.

Trial Registration: ClinicalTrials.gov NCT04147455; https://clinicaltrials.gov/ct2/show/NCT04147455

International Registered Report Identifier (IRRID): RR2-10.1186/s12889-020-09454-2

(JMIR Public Health Surveill 2023;9:e35116) doi:10.2196/35116

KEYWORDS

behavior change communication; bystander behavior; campus sexual assault; educational entertainment (edutainment); sexual violence; social cognitive theory; social norms theory; Vietnam; mobile phone

Introduction

Background

Sexual violence is defined as a sexual act committed against a person when consent is not freely given. Sexual violence ranges from unwanted noncontact sexual experiences to completed forced penetration [1]. Among sexually experienced women aged 15-19 years, forced sexual debut is common worldwide (15%), including in the Asia and Pacific (14%) [2], and women comprise 91% of the victims [3]. The physical, psychological, and economic aftermath of sexual violence often is severe for victims [4,5] and is costly for societies [6].

In Vietnam, sexual violence persists [7,8] despite the legal reforms to address it. Men may discount, excuse, or deny acts of sexual violence [9], and their reported rates of sexually violent behavior (0.2%) are lower than women's reported rates of victimization (12%) [10]. Thus, prevention with men is crucial to create environments where women's freedom from violence is possible [11]. However, young men are often difficult to reach [12] and may resist participation in prevention programs [13]. The behavior of bystanders, or witnesses of sexual violence, may also be gender specific [14]. Thus, prevention with men that integrates a bystander framework and treats men as "allies" [15] may address attitudes and behaviors while decreasing resistance to participation [16,17].

However, evidence of programmatic impacts from rigorous evaluations is limited [18]. In an Ovid database search for "sexual violence," "intervention," and "review," we identified 4 reviews published since 2016 that focused on interventions to prevent sexual violence in adolescents and young adults [19-22]. A 2017 review of reviews identified a few interventions focused on boys [19]. A subsequent review of 44 bystander intervention programs in North America found that, on average, programs involved a single session (75%) of <2 hours (mean 116 minutes; range 10-480 minutes) in college populations (75%) as well as in-person presentations (68%) and discussions (54%) with mixed-gender groups (56%). Relatively few interventions were tailored to men (27%) and involved web-based delivery (11%; [20]). Moreover, study designs had important limitations: the majority were not randomized controlled trials (RCTs; 62%) and had small sample sizes (mean 536, range 1-4311) in White populations (73%), high attrition (36%) [20], and short follow-up periods of ≤6 months (89%) focused on nonbehavioral outcomes (66%). A third systematic review focused on interventions to prevent intimate partner, dating, and sexual violence in men and boys and found heterogeneity across the 9 included studies in program content and delivery strategies, study designs, sample sizes, and outcome measurement [21]. Most studies used cluster-randomized designs, recruited undergraduate college students, and evaluated a multisession program delivered via group sessions; however, only 1 program reduced men's self-reported sexually violent behavior, and most studies were based in the United States. A fourth review that focused on intervention studies to change "hegemonic masculinities" found that 8 of the 10 included quantitative studies were conducted in the United States or Africa, only 1 was web-based, and impacts on sexually violent behavior were mixed [22]. Thus, especially in lower- and middle-income countries (LMICs), gender-specific interventions with men to prevent sexually violent behavior are rare, and theoretically grounded, web-based sexual violence prevention interventions engaging men and involving a bystander framework have not been evaluated in such settings.

This study tested the impact of GlobalConsent on sexually violent behavior and prosocial bystander behavior in university men in Vietnam. The team adapted GlobalConsent from RealConsent, an evidence-based web-delivered intervention tested among university men in the United States [23]. RealConsent is based on formative research [24], social cognitive theory [25], social norms theory [26], and the bystander education model [27]. Sexually violent behavior is theorized to arise from the interplay of sociocontextual factors, personal factors, and behavior. To operationalize this theory of change, programmatic features of RealConsent include gender-specific content [28-31] that resonates with the viewer [32] and diverse behavior change techniques [33], such as providing information and instruction on obtaining effective consent for sex and intervening safely, modeling communication and intervening behaviors, showing positive outcomes for obtaining consent and intervening plus negative outcomes for perpetrating and not intervening, and reinforcing with positive feedback [34]. RealConsent uses didactic presentation of material via integrated audio, video, and infographics; problem-based learning with interactivity and reinforced practice; short videos or animations to model behavior; and educational entertainment, starting and ending each module with a brief episode of a serial drama [23,35].

RealConsent aimed to change 2 primary behaviors by changing 7 cognitive, attitudinal, and affective mediators. The 2 primary behaviors were prosocial intervening behaviors, such as trying to stop a peer from being sexually coercive and sexually violent behavior toward women. The 7 mediators included knowledge of the elements of sexual consent, knowledge and skills to intervene safely, misperceptions of norms about sexual violence and rape, negative attitudes about date rape, positive masculinity, skills in sexual communication, and empathy for the victims of sexual violence. Results from a randomized controlled trial supported the efficacy of RealConsent. A random probability sample of 743 undergraduate men aged 18-24 years attending a large, public, urban university in the Southeastern United States was randomized to RealConsent (376/743, 51%) or a web-based health education attention-control program (367/743, 49%) [23]. Participants were surveyed on the web at baseline, after intervention, and 6 months after intervention. Six 30-minute modules were delivered via a password-protected web portal. At 6 months after intervention, compared with the control group, the RealConsent group intervened more often and engaged in less sexual violence. They reported greater knowledge about legal definitions of sexual assault and elements of effective consent; lower adherence to rape myths, negative date-rape attitudes, hypergender ideology, and hostility to women; greater empathy for rape victims; greater intentions to intervene; less positive outcome expectancies for nonconsensual sex; more positive outcome expectancies for intervening; and less comfort with other men's inappropriate behavior. The trial experienced high loss to follow-up in the intervention (67%) and control (75%) arms; therefore, results in the retained sample should be interpreted with caution.

Objectives

To adapt RealConsent for delivery to a new context, we followed the Centers for Disease Control and Prevention's 5-step process [36]. Step 1, *assess*, involved assessing the target population, the evidence-based intervention (EBI) being considered for implementation, and the implementing agency's capacity to use the intervention. Step 2, *select*, entailed determining whether to adopt the EBI without adaptations, with adaptations, or to choose another EBI for adaptation. Step 3, *prepare*, involved making the necessary adaptations to the EBI while retaining the core elements. Step 4, *pilot*, entailed piloting the adapted intervention and developing a plan for implementation. We conducted qualitative research to implement steps 1 to 4 [37], and those findings are presented elsewhere [38-40].

To complete step 5, *implement*, we undertook an RCT to test, among men attending 2 universities in Vietnam, the impact of GlobalConsent versus a customized adolescent health education (AHEAD) attention-control condition on changes in 2 primary behavioral outcomes—sexually violent behavior and prosocial bystander behavior. The aim of this paper was to present findings from the RCT with regard to 3 a priori hypotheses. First, we expected that GlobalConsent would *mitigate increases in sexually violent behavior* that may occur when men matriculate into university and interact with women with less parental supervision [41-43]. Second, we expected that GlobalConsent would *increase prosocial bystander behavior* relative to attention-control conditions. Third, we expected that GlobalConsent would influence these behavioral outcomes directly and indirectly through changes in knowledge, attitudinal, and affective secondary outcomes. The results presented here focused on the primary, unmediated behavioral outcomes.

Methods

Setting

Universities in Vietnam were suitable contexts for adapting RealConsent, given their similarities to universities in the United States. In Vietnam, undergraduate study is 4-6 years, with 2 foundational years and 2-4 years for specialization. Except for political education and national defense, universities design their curricular and extracurricular activities and maintain networks through the Ministry of Education and Training, professional associations, and the Youth Union, all pathways for the national scale-up of educational programs. The study sites were 2 universities located in Hanoi. One is a 120-year-old state school that trains 1000 students yearly in health professions. The other is a 34-year-old private university that trains 7000 students annually across diverse disciplines. Both universities provided letters of support for this study.

Trial Design

The study design, detailed elsewhere [37], applied a double-blind, parallel intervention versus attention-control group design with balanced (1:1) randomization (Multimedia Appendix 1).

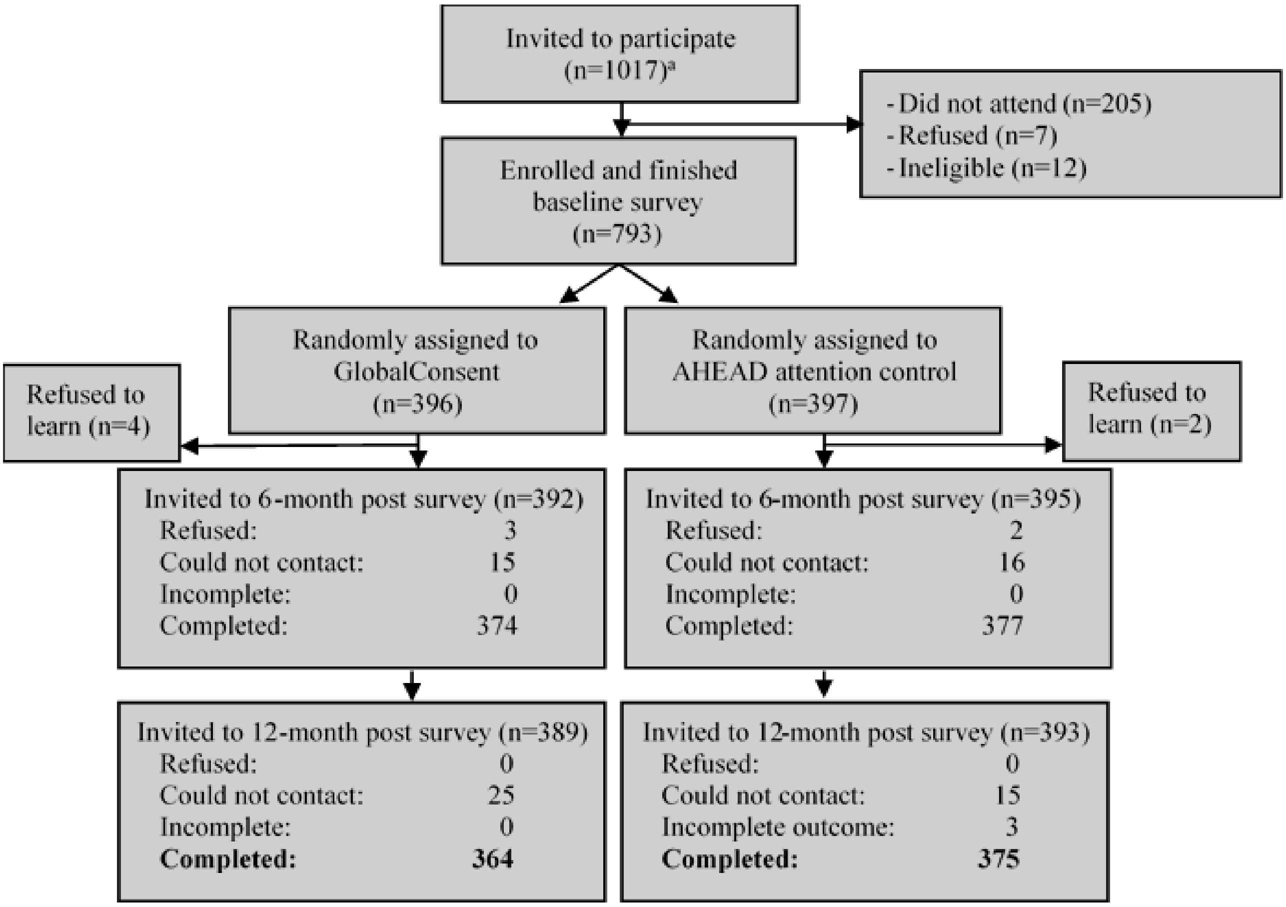
Ethics Approval

Data were collected following established ethical guidelines for research on gender-based violence [37]. The institutional review boards of Emory University (IRB00099860) and Hanoi University of Public Health (017–384/DD-YTCC) reviewed and approved the study protocol, including the web-based consent forms used in the study.

Sample Eligibility, Participation, and Retention

Eligible men for the study were heterosexual or bisexual, aged 18-24 years, and matriculating in September 2019 at either of the 2 study sites. To identify eligible men at university 1, overall, 56% (5/9) of the departments were included in the sample because all departments had an agreed threshold of ≥15 cismale students. For university 2, a total of 59% (13/22) of the departments having at least 20 cismale students were included. All first-year cismale students of the included departments were invited to participate (n=1017; Figure 1; power calculations are presented in the protocol paper by Yount et al [37]). Of these 1017 men, 205 (20.1%) did not attend the orientation, 7 (0.7%) declined, and 12 (1.2%) did not meet the inclusion criteria. Eligible, consenting men who completed an in-person, computer-assisted self-administered baseline survey at their home university in September 2019 (793/1017, 78%) were assigned a random number generated in Microsoft Excel (Microsoft Corp), sorted in ascending order. Of the 793 eligible men, the first 397 (50.1%) were assigned to the AHEAD attention-control group, and the remaining 396 (49.9%) men were assigned to the GlobalConsent group. The participants were blinded to these assignments because of the use of an attention-control condition, and the analysis team was blinded to these assignments throughout the analysis. Furthermore, 6 assigned participants (4 GlobalConsent and 2 AHEAD) declined to login to their assigned learning program. The participants were able to access their assigned programs from November 2019 to January 2020. Owing to the COVID-19 pandemic, posttest 1 (completed in April-May 2020, 3 months after intervention, and 6 months after baseline) and posttest 2 for trial endline (completed in October-November 2020, 9 months after intervention, and 12 months after baseline) were self-administered remotely via the web to ensure completion, regardless of students' location of residence. Some internet disruptions were apparent, as the mean number of survey attempts was >1 in both groups and posttest rounds. However, the agreement across survey attempts between outcome-related responses exceeded 95%. In the GlobalConsent group, 94.4% (374/396) and 91.9% (364/396) students completed posttest 1 and 2, respectively. In the AHEAD group, 95% (377/397) and 94.5% (375/397) of students completed posttest 1 and 2, respectively. Reasons for high retention may have included a cultural tendency to adherence in Vietnam, a progressive compensation schedule for survey and program module completion, and text messages and emails to remind students to complete their studies [37].

Figure 1. Flow diagram of sample enrollment, allocation, and follow-up randomized controlled trial of GlobalConsent. [a] Of the 1017 eligible men, 362 were from University 1 and 655 were from University 2. AHEAD: adolescent health education.



GlobalConsent Program

The adapted GlobalConsent program took into account qualitative findings [38,40] and feedback from focus group discussions with cismale students and university stakeholders on the RealConsent program and the adapted GlobalConsent storyboards before production (Multimedia Appendices 2 and 3). In total, 7 major adaptations have been observed. First, the local cismale characters were grounded in findings from the formative qualitative study, including the family and social contexts in which young men live in Vietnam. The male characters were adapted to evolve in contextualized ways, showing how their gender attitudes and awareness of sexual violence and response were changing. Thus, 5 cismale characters were created to reflect different initial masculinity typologies and different speeds and degrees of "masculinity development" over the program. These characters reflected (1) *positive masculinity* throughout; (2) *somewhat positive masculinity* throughout; (3) *somewhat more traditional masculinity* initially and becoming more aware earlier in the program (modules 2-3), as reflected in verbal tone, language, and marked behavior change; (4) *more traditional masculinity* initially and struggling but becoming more aware, as reflected in a scene in a karaoke bar when he returns to a hotel with drunk young woman and decides not to have sex; and (5) *very traditional masculinity* initially using inappropriate behavior, sanctioned by friends, and becoming more aware of modules 4-5, again as reflected in tone, language, and behavior.

Second, the cisfemale partners of these cismale characters were added because sexual coercion and violence in dating relationships emerged as salient in the formative qualitative findings. Third, "coach-talk" segments for reinforced learning were removed, as they relied on English vernacular, were misunderstood by Vietnamese stakeholders, and caused confusion about key messages. Fourth, "key questions" and "takeaways" were added to specific segments and scenarios, as these forms of reinforced learning were more familiar in Vietnam and helped to reduce misinterpretation of key messages. Fifth, content related to knowledge, attitudinal, and affective pathways in the theory of change was adapted to be relevant for Vietnam, and definitions of sexual violence were expanded to include the full spectrum of sexually violent behaviors. Sixth, anonymized narratives from the qualitative interviews were incorporated into the serial drama script to ensure comprehension and cultural suitability. Finally, all scenarios were refilmed or reanimated to resonate with the learning style of the cismale students in Vietnam.

The final GlobalConsent program included six 30-minute modules, each ranging in the number of segments and types of activities, with diverse actors and languages suitable for students who self-identify as heterosexual or bisexual men in Vietnam. Modules—grounded in social cognitive theory, social norms theory, and a bystander framework [37]—covered 6 topics to address hypothesized knowledge, normative or attitudinal, and affective mediators of the program and behavioral outcomes. These topics included consent for sex, rape-myth beliefs and norms regarding gender roles, effective communication, alcohol and rape, victim empathy, and bystander intervention. Each module included didactic activities, interactivity, and episodes of an adapted serial drama that modeled positive behaviors with cismale peers and cisfemale sex partners. The program was formatted for delivery to the computers and smartphones. Participants were provided with unique login information to access the platform and could work at their own pace, could not skip segments, and were encouraged by email and text messages to complete all modules in 12 weeks. The learning app captured the number of times a student logged in to each module and the total minutes a student spent on each module.

AHEAD Attention-Control Program

The team developed AHEAD, the attention-control program, from open-access content suitable for young people in Vietnam [37]. To approximate the format, intensity, and duration of GlobalConsent, AHEAD was developed as a web-delivered, multimedia program audio-narrated in Vietnamese with 6 learning modules that were 35-45 minutes long on brain development, nutrition, physical activity, substance use, sleep, and agency. Similar to GlobalConsent, participants could work at their own pace, could not skip segments, and were encouraged by email and text messages to complete all modules in 12 weeks. The use of an attention-control condition was justified because the original trial of RealConsent also used an attention-control condition and the team wanted the trial of GlobalConsent to include a similar comparator. In addition, the use of an attention-control condition accounted for the web-based mode of delivery in both study arms, so any effects of the content of GlobalConsent could be compared with those of a general *adolescent health education* program that did not focus on the theoretical knowledge, attitudinal, and affective pathways by which GlobalConsent was theorized to reduce men’s sexually violent behavior and increase their prosocial bystander behavior.

Data

Primary Outcomes

The primary outcomes for this analysis were sexually violent behavior, measured using the Sexual Experiences Survey [44], and prosocial bystander behavior, measured using 4 items from an adapted Bystander Behavior Scale [45] (Table 1). The Sexual Experiences Survey asks about the perpetration of 7 acts of contact sexual violence, ranging from unwanted touching to forced penetration, using any of 7 physical or nonphysical tactics, such as holding someone down or threatening to end the relationship (35 items total). The Sexual Experiences Survey also asked about the perpetration of 10 acts of noncontact sexual violence, such as masturbating in front of someone when they did not agree. We captured *sexually violent behavior* based on the reported frequency (never, once, twice, and ≥3 times) of all 45 items in the prior 6 months. We captured *prosocial bystander behavior* based on the reported frequency (never, once, and ≥2 times) in the prior 12 months of 4 acts, such as “I have talked with someone about sexual or dating violence as an issue for our university.” The frequency of missingness was 0.1%-0.7% across all items, and these responses were recoded as 0, “not reported.” We created kernel density plots of the summative scores for each primary outcome (Multimedia Appendix 4). On the basis of their distributions, we captured 3 measurement scales for each behavioral outcome to assess whether the GlobalConsent program had linear or threshold effects. These measurement scales were any reported act (1 or more vs 0 or none reported), many reported acts (3 or more vs 0-2 or none reported), and the number of reported acts (Table 1).

Table 1. Primary reported behavioral outcomes in the randomized controlled trial of GlobalConsent.

Outcomes	Instrument	Response options	Sample size, n	Example item	“Any act”	“Many acts”	Sum
Sexually violent behavior, past 6 months	Sexual Experiences Survey [44]	- 0, Never - 1, Once - 2, Twice - 3, ≥3 times	45	I took photos or videotapes of someone while they were undressing, were nude, or were having sex, when they did not agree to it.	- 1, Any act - 0, No act[a]	- 1, 3 or more acts - 0, 0-2 acts[a]	0-135
Prosocial bystander behavior, past 12 months	Bystander Behavior Scale [45]	- 0, Never - 1, Once - 2, ≥2 times	4	I have encouraged others to learn more and get involved in preventing sexual or dating violence.	- 1, Any act - 0, No act[a]	- 1, 3 or more acts - 0, 0-2 acts[a]	0-8

[a] Missing responses for the number of acts were coded 0.

Exposures

Participants were assigned randomly to either the GlobalConsent treatment (=1) or the AHEAD attention-control program (=0). Time was coded 0 for baseline and 1 for subsequent time points. A time-by-treatment arm interaction term was created to assess the difference-in-difference (DID) for GlobalConsent versus AHEAD at “endline” (posttest 1 and 2 combined) versus baseline.

Covariates

Covariates included age in years, program of study (health sciences, others), relationship history (ever, never), living situation (parents, other relatives, dormitory, off campus, and other), ethnicity (majority, minority), sexual orientation (heterosexual, bisexual), religion (any, none), and residence in Hanoi for at least 1 year before baseline (yes, no). In addition, past sexual history (ever, never), sexual preference (women only, others), gender identity, and expression (very feminine=1 to very masculine=7) were collected at the posttest 1. At posttest 2, an adapted Vietnam Prevalence Study on Child Maltreatment [46] scale measured maltreatment before the age of 18 years with 28 items capturing any exposure to behaviors by adults in 5 domains: physical, sexual, and emotional abuse as well as physical and emotional neglect. Dichotomous (any, none) variables were created for each domain of child maltreatment. Any exposure (yes, no) to web-based sexually explicit material in the past 6 months was measured for 5 domains (textual, partial nudity, full nudity, nonviolent sexual acts, and violent sexual acts) [47].

Statistical Analysis

First, we conducted descriptive analyses of the sample overall and by study arm to assess qualitative balance across groups. Second, we assessed levels of sexually violent behavior and prosocial intervening behavior at each study wave by study arm. Finally, we conducted DID modeling to estimate the odds ratios (ORs), incidence rate ratios, and 95% CIs for sexually violent behavior and prosocial bystander behavior at endline (posttest 1 and 2 combined) versus baseline. For binary outcomes, we performed logistic regression with the assigned treatment arm, time, and treatment-by-time interaction (DID variable) as predictors. For count outcomes, we performed a negative binomial regression, which accounted for overdispersion in the outcome. To probe significant treatment-by-time interaction or DID effects, we computed and compared the OR and incidence rate ratio of the occurrence of the behavior separately for each group. All analyses were performed using Stata (version 16; StataCorp [48]).

Role of the Funding Source

The funder of the study had no role in study design, data collection, analysis, interpretation, or writing of the report. All authors had full access to all study data and were responsible for the decision to submit this paper for publication.

Results

Sample Characteristics

The characteristics of men across the groups were similar at baseline and follow-up waves (Tables 2-4). At baseline, men were aged 18 years, on average, and a majority were attending university 2, majoring in nonhealth fields, of Kinh ethnicity, heterosexual, nonreligious, and had never been in a dating or sexual relationship. The living arrangements and residential location of men were similarly distributed across the groups (Table 2). After baseline, high percentages of men in both the GlobalConsent and AHEAD groups, respectively, logged into the program (372/396, 93.9% vs 386/397, 97.2%), completed at least 1 module (361/396, 91.2% vs 374/397, 94.2%), and completed all modules (355/396, 89.6% vs 362/397, 91.2%).

At posttest 1, most men in both groups had never had sex and sexually preferred only women (Table 3). In addition, reported gender identities and gender presentations were predominantly masculine, with average scores ≥6.0. About three-fourths (569/751, 75.8%) of men had ever been exposed to textual sexually explicit material or “top nudity.” Almost two-thirds (490/751, 65.2%) had ever been exposed to “full nudity” or nonviolent sexual acts. Approximately 40.9% (302/739; ie, about 4 in 10) had ever been exposed to sexually explicit material showing violent sexual acts against women.

Reported experiences of child maltreatment, by type, were similar in the GlobalConsent and AHEAD groups (Table 4). More than half (401/735, 54.6%) of the men reported experiencing emotional abuse in childhood, and about 1 in 5 reported experiencing emotional neglect in childhood. Almost half (346/736, 47%) of the men reported experiencing physical abuse in childhood and ≤5% reported experiencing physical neglect in childhood. More than 1 in 10 men reported experiencing sexual abuse during childhood. High levels of childhood maltreatment are consistent with 2014 population-based national data from secondary school students. Except for emotional neglect, which is higher, the prevalence estimates for childhood maltreatment in our sample are lower than the national average [46].

Table 2. Baseline Characteristics of men matriculating in 2 universities in Vietnam in September 2019, overall and by groups randomized to GlobalConsent treatment and adolescent health education (AHEAD) attention-control conditions.

Covariates measured at baseline	GlobalConsent (n=396)[a]	AHEAD (n=397)	Total (n=793)[a]
University, n (%)			
University 1	172 (43.4)	173 (43.6)	345 (43.5)
University 2	224 (56.6)	224 (56.4)	448 (56.5)
Major, n (%)			
Health	185 (46.7)	181 (45.6)	366 (46.2)
Others	211 (53.3)	216 (54.4)	427 (53.9)
Age (years), mean (SD)	18.1 (0.4)	18.1 (0.3)	18.1 (0.4)
Ethnicity, n (%)			
Majority (Kinh)	377 (96.2)	379 (95.5)	756 (95.8)
Minority	15 (3.8)	18 (4.5)	33 (4.2)
Sexual orientation, n (%)			
Heterosexual	378 (95.5)	381 (96)	759 (95.7)
Bisexual	18 (4.6)	16 (4)	34 (4.3)
Religion, n (%)			
Any	62 (15.7)	73 (18.4)	135 (17)
None	334 (84.3)	324 (81.6)	658 (83)
Relationship status, n (%)			
Ever in a relationship	179 (45.2)	187 (47.1)	366 (46.2)
Never in a relationship	217 (54.8)	210 (52.9)	427 (53.9)
Living situation, n (%)			
With parents	233 (33.6)	119 (30)	252 (31.8)
With other relatives	54 (13.6)	52 (13)	106 (13.4)
Dormitory or on campus	52 (13.1)	68 (17.1)	120 (15.1)
Off-campus alone or with nonrelatives	145 (36.6)	138 (34.8)	283 (35.7)
Other do not know	12 (3)	20 (5)	32 (4)
Lived in Hanoi at least 1 year, n (%)			
Yes	203 (51.3)	183 (46.1)	386 (48.7)
No	193 (48.7)	214 (53.9)	407 (51.3)

[a] *Missing at baseline*: ethnicity 4 GlobalConsent.

Table 3. Characteristics measured at posttest 1 (6-months after baseline) of men matriculating in 2 universities in Vietnam in September 2019, overall and by groups randomized to GlobalConsent treatment and adolescent health education (AHEAD) attention-control conditions.

Covariates measured at posttest 1 (6 months after baseline)	GlobalConsent (n=374)[a]	AHEAD (n=377)[a]	Total (n=751)[a]
Survey attempts, mean (SD)	1.6 (1.9)	1.4 (0.8)	1.5 (1.5)
Ever had sex, n (%)			
Yes	66 (18.1)	80 (22.3)	146 (20.2)
Never	299 (81.9)	279 (77.7)	578 (79.8)
Sexual preference, n (%)			
Only women	297 (81.4)	303 (84.4)	600 (82.9)
Others or do not know	68 (18.6)	56 (15.6)	124 (17.1)
Gender identity (very feminine=1 to very masculine=7), mean (SD)	6.1 (0.9)	6.2 (0.9)	6.2 (0.9)
Gender presentation (very feminine=1 to very masculine=7), mean (SD)	6.0 (0.9)	6.0 (0.9)	6.0 (0.9)
Exposure to SEM[b]: textual, n (%)			
Yes	286 (76.5)	297 (78.8)	583 (77.6)
No	88 (23.5)	80 (21.2)	168 (22.4)
Exposure to SEM: top nudity, n (%)			
Yes	279 (74.6)	290 (76.9)	569 (75.8)
No	95 (25.4)	87 (23.1)	182 (24.2)
Exposure to SEM: full nudity, n (%)			
Yes	241 (64.4)	249 (66.1)	490 (65.2)
No	133 (35.6)	128 (34)	261 (34.8)
Exposure to SEM: nonviolent sex, n (%)			
Yes	234 (62.8)	244 (64.7)	480 (63.9)
No	139 (37.2)	133 (35.8)	271 (36.1)
Exposure to SEM: violent sex, n (%)			
Yes	143 (38.2)	159 (42.2)	302 (40.2)
No	231 (61.8)	218 (57.8)	449 (59.8)
Lost to follow-up at posttest 1, n (%)	22 (5.6)	20 (5)	42 (5.3)

[a] *Missing at posttest 1*: ever had sex 9 GlobalConent (GC), 18 AHEAD; sexual preference 9 GC, 18 AHEAD; gender identity 11 GC, 18 AHEAD; gender presentation 9 GC, 18 AHEAD.

[b] SEM: sexually explicit material.

Table 4. Characteristics measured at posttest 2 (12-months after baseline) of men matriculating in 2 universities in Vietnam in September 2019, overall and by groups randomized to GlobalConsent treatment and adolescent health education (AHEAD) attention-control conditions.

Covariates measured at posttest 2 (12 months after baseline)	GlobalConsent (n=364)[a]	AHEAD (n=375)[a]	Total (n=739)[a]
Child maltreatment: emotional abuse, n (%)			
Yes	188 (51.9)	213 (56.6)	401 (54.6)
No	173 (48.1)	161 (43.1)	334 (45.4)
Child maltreatment: emotional neglect, n (%)			
Yes	75 (20.7)	65 (17.4)	140 (19)
No	287 (79.3)	309 (82.6)	596 (81)
Child maltreatment: physical abuse, n (%)			
Yes	171 (47.2)	175 (46.8)	346 (47)
No	191 (52.8)	199 (53.2)	390 (53)
Child maltreatment: physical neglect, n (%)			
Yes	9 (2.5)	19 (5.1)	28 (3.8)
No	353 (97.5)	355 (94.9)	708 (96.2)
Child maltreatment: sexual abuse, n (%)			
Yes	55 (15.2)	48 (12.8)	103 (14)
No	307 (84.8)	326 (87.2)	633 (86)
Lost to follow-up at posttest 2, from baseline n (%)	32 (8.1)	22 (5.5)	54 (6.8)

[a] *Missing at posttest 2*: physical abuse, emotional abuse, sexual abuse, physical neglect, and emotional neglect same 2 in GlobalConsent and same 1 in AHEAD.

Sexually Violent Behavior and Prosocial Intervening Behavior

Reported rates of any sexually violent behavior were 31.1% (123/396), 21.4% (80/374), and 31.9% (116/364) across survey waves in the GlobalConsent group and 25.9% (103/397), 26.3% (99/377), and 31.5% (118/375) across waves in the AHEAD group (Table 5). Rates of "high" sexually violent behavior (>2 acts) were 21% (83/396), 13.6% (51/374), and 24.7% (90/364) across waves in the GlobalConsent group and 17.1% (68/397), 18% (68/377), and 23.2% (87/375) across waves in the AHEAD group. The mean count of sexually violent acts was 3.2, 2.2, and 3.1 across waves in the GlobalConsent group and 2.4, 3.6, and 3.2 across waves in the AHEAD group.

Table 5. Sexually violent behavior and prosocial bystander behavior at baseline, posttest 1 (6 months after baseline), and posttest 2 (12 months after baseline) of men matriculating at 2 universities in Vietnam in September 2019 and randomized to receive the GlobalConsent program or the adolescent health education (AHEAD) attention-control program.

Outcomes	Baseline		Posttest 1		Posttest 2	
	AHEAD (n=397)	GlobalConsent (n=396)	AHEAD (n=377)	GlobalConsent (n=374)	AHEAD (n=375)	GlobalConsent (n=364)
Sexually violent behavior						
Yes any, n (%)	103 (25.9)	123 (31.1)	99 (26.3)	80 (21.4)	118 (31.5)	116 (31.9)
Yes, high (>2 acts), n (%)	68 (17.1)	83 (21)	68 (18)	51 (13.6)	87 (23.2)	90 (24.7)
Count, mean (SE); range	2.4 (9.7); 0-111	3.2 (11.6); 0-96	3.55 (12.1); 0-112	2.24 (8.1); 0-70	3.2 (9.9); 0-79	3.1 (11.7); 0-135
Prosocial bystander behavior						
Yes any, n (%)	258 (65)	258 (65.2)	N/A[a]	N/A	170 (45.3)	203 (55.8)
Yes high (>2 acts), n (%)	141 (35.5)	163 (41.2)	N/A	N/A	108 (28.8)	138 (37.9)
Count, mean (SE); range	2.2 (0.1); 0-8	2.4 (0.1); 0-8	N/A	N/A	1.7 (0.1); 0-8	2.3 (0.1); 0-8

[a] N/A: not applicable (as the 12-month window of observation at posttest 1 overlaps with the 12-month window of observation at baseline).

Reported rates of any prosocial bystander behavior were 65.2% (258/396) and 55.8% (203/364) across survey waves in the GlobalConsent group and 65% (258/397) and 45.3% (170/375) across waves in the AHEAD group (Table 5). Rates of "high" prosocial bystander behavior (>2 acts) were 41.2% (163/396) and 37.9% (138/364) across waves in the GlobalConsent group and were 35.5% (141/397) and 28.8% (108/375) across waves in the AHEAD group. The mean counts of prosocial bystander acts were 2.4 and 2.3 across waves in the GlobalConsent group and were 2.2 and 1.7 across waves in the AHEAD group.

Effects of GlobalConsent on Behavioral Outcomes

Significant interaction effects in the DID models suggested favorable impacts of GlobalConsent relative to AHEAD on sexually violent behavior (>2 acts; OR 0.5, 95% CI 0.3-0.7; *P*=.001) and prosocial bystander behavior (any act; OR 1.5, 95% CI 1.0-2.3; *P*=.05; Table 6). Marginally significant interaction effects in DID models for any sexually violent behavior and the count for prosocial bystander acts suggested a consistent, favorable impact of GlobalConsent relative to AHEAD across the measurement scales of these 2 behavioral outcomes.

Table 6. Difference-in-difference odds ratios (ORs) for the logistic model or incidence rate ratios (IRRs) for the negative binomial model, comparing endline (posttest 1 and 2 combined) versus baseline for 730 university men in the GlobalConsent versus adolescent health education groups.

Outcomes	OR IRR (95% CI)	*P* value
Sexually violent behavior		
Yes, any act	0.71 (0.50-1.00)	.05
Yes, >2 acts	0.47 (0.31-0.72)	.001
Count	0.65 (0.37-1.14)	.13
Prosocial bystander behavior		
Yes, any act	1.51 (1.00-2.27)	.05
Yes, >2 acts	1.19 (0.78-1.81)	.42
Count	1.27 (1.00-1.60)	.05

To probe the interaction effects further, we computed the ORs of sexually violent behavior (>2 acts) and prosocial bystander behavior (any act) at endline (posttest 1 and 2 combined) versus baseline separately for the GlobalConsent and AHEAD groups. Among the GlobalConsent participants, the odds of perpetrating >2 acts of sexual violence at the endline were 1.3 times the odds at baseline. For the AHEAD group, the odds of perpetrating >2 acts of sexual violence at the endline were 2.7 times the odds at baseline. Among the GlobalConsent participants, the odds of any prosocial behavior at endline were 0.7 times the odds at baseline. For the AHEAD group, the odds of any prosocial behavior at posttest were 0.5 times the odds at baseline.

Discussion

Principal Findings

This randomized controlled trial recruited 793 university men in Vietnam to test the efficacy of GlobalConsent, a culturally tailored adaptation of an efficacious, theoretically grounded web-based sexual violence prevention program originally designed for US college men [23]. In this study, compared with men in the attention-control group, men in the GlobalConsent group had lower odds of engaging in sexually violent behavior at the endline than at the baseline. In addition, compared with men in the AHEAD group, men in the GlobalConsent group had higher odds of any prosocial bystander behavior at the endline than at baseline. The favorable impacts of GlobalConsent versus AHEAD were consistent across the measurement scales of these 2 behavioral outcomes. Such behavior changes are notable, as this is the first web-delivered sexual violence prevention program for cismale college students implemented in a middle-income setting. Most sexual violence prevention programs globally have measured changes in reported behavioral *intentions* rather than in reported *behavior* [49].

Study Limitations and Strengths

This study had some limitations. First, the study sites were limited to 2 urban universities to accommodate program adaptation, production, and testing. Therefore, the findings cannot be generalized to all university men in Vietnam. However, the 2 study universities represented public and private institutions of higher education, diverse faculties of study, and diverse student bodies from urban and provincial Vietnam. Second, behavioral outcomes were self-reported and may have been subject to social desirability bias. To mitigate this limitation, the team relied on a validated instrument to measure sexual violence and self-administration, which enhances privacy and may enhance honest disclosures. Third, we did not interview women on campus about changes in behavior, as reported by men, owing to resource constraints and the ethical and logistical challenges of linking women’s reports of men’s behavior at the individual level. Fourth, some nonsignificant differences in behavior and attenuation of behavior change at 12 months suggest the need for booster training after the 3-hour program period. Fifth, the 12-month follow-up period allowed for a short-term impact assessment of GlobalConsent. However, the follow-up period in this study exceeded that in most similar studies. Long-term follow-up of participants would be beneficial to assess the impact of GlobalConsent on behavioral outcomes throughout men’s time at university and beyond. Sixth, similar to other web-based interventions, the content of GlobalConsent may become outdated, and updating or future adaptations to GlobalConsent may incur new costs. To mitigate this limitation, the team adapted a program script and content to ensure maximum durability. Seventh, participants in the GlobalConsent group may have become unblinded to their treatment assignment. Although a possibility, having an attention-control condition in which adolescent-focused health content is provided should reduce the likelihood of unblinding because health-related content is provided in both arms, and participants were not informed about the specific theoretical knowledge-, attitudinal-, and empathy-related pathways by which the GlobalConsent programing was expected to operate. Finally, some loss to follow-up was observed in both study arms. However, in absolute terms and relative to similar intervention studies [20-23], attrition from our study was very low and similarly low for the treatment and control arms. In addition, the baseline characteristics of those who were lost to follow-up were similar across the treatment and control arms. These findings reduce concerns that men who were sexually violent after baseline attrited more often from the treatment group than the control group; however, this possibility cannot be ruled out. Finally, as the DID method focused on changes, it accommodated the baseline differences in both main outcomes between the 2 arms.

Recognizing these potential limitations, the strengths of this study were many. They included a randomized controlled design, customized attention-control condition, web-based delivery of GlobalConsent and attention-control programs to computers and smartphones, a large sample of participating university men, outstanding participation and retention, and refined measurement of primary behavioral outcomes and secondary (mediating) outcomes at 3 months and 9 months after intervention. Ours is among the few RCTs in the field, especially in LMICs, to assess impacts on 2 behaviors and to assess the impact on behavior at 2 time points after intervention.

Implications for Research and Public Health Practice

This study fills critical knowledge gaps on preventing sexual violence globally. In LMICs, interventions to prevent sexual violence in young men are rare [19]. This project adds to the limited evidence from LMICs on efficacious strategies to prevent sexual violence and promote prosocial bystander behaviors in university men. As GlobalConsent is an adaptation of RealConsent, this study adds to the limited evidence on the replication of efficacious sexual violence prevention interventions [50]. The review of bystander interventions by Mujal et al [20] identified 2 programs with evidence of replication; both were group-delivered and relied on presentations and discussions—techniques that can be resource- and cost-intensive. This study provides empirical support for the efficacy of a web-based edutainment approach for sexual violence prevention delivered to the computers and smartphones of students on college campuses, which may be more appealing and feasible in resource-constrained LMICs [51,52]. Finally, globally, most sexual violence prevention interventions have used in-person, small-group formats, with limited reach, standardization, and impact [50,53]. The web-based platform and GlobalConsent delivered to computers and smartphones make wide-scale implementation with higher fidelity more feasible than in-person programs. A large-scale implementation study of GlobalConsent in universities across Vietnam would be a robust next step to assess its effectiveness, and information about barriers and facilitators to university delivery of this intervention would be critical for moving GlobalConsent from research to practice.

This study reveals the value of using a systematic process to adapt evidence-based programs before they are delivered in new settings and populations. The process [36] used here allowed for the preservation of core content responsible for the efficacy of RealConsent and flexibility to add and to tailor other elements for university populations in Vietnam. Although the results here cannot confirm that GlobalConsent was an optimal adaptation to RealConsent, findings from the trial still suggest that following a systematic process to adapt an EBI may yield an adapted program that is efficacious cross culturally. This study demonstrates that GlobalConsent was able to reduce the risk of sexual violence and to promote prosocial bystander behavior in this LMIC. The potentially smaller effects of a web-based intervention are counterbalanced by GlobalConsent's highly standardized delivery relative to in-person, small-group formats; therefore, the consistent delivery and scalability of GlobalConsent to the national level is likely to be high. Research to adapt and test GlobalConsent in other LMICs is warranted.

Conclusions

Given the high rates of sexual violence against women, growing numbers of young men and women attaining postsecondary education, and rapid increases in access to the internet and smartphones globally, universities are ideal settings to provide novel sexual violence prevention programs during this critical developmental window, when young men are leaving home and are at increasing risk of sexually violent behavior. Evidence-based sexual violence prevention programs such as GlobalConsent, which are cost-effective, easily implemented by universities, and appealing to a diverse student population, are needed globally. Reviews of sexual violence prevention programs have described the importance of bystander approaches and engaging men as women's allies in preventing sexual violence. Equally important is the ability of such efficacious programs to reach large populations to expand their impact. This study demonstrates that theoretically grounded, web-based edutainment with gender-specific content, customized didactic and interactive behavior change techniques, and an adapted serial drama make GlobalConsent potentially scalable in Vietnam and adaptable to other LMICs, where efficacious sexual violence prevention programs are needed.

Acknowledgments

This study was supported by a grant from an anonymous charitable foundation. The authors would like to thank Dr Laura Salazar for consultation during proposal development and Ms Emma Jackson in the Hubert Department of Global Health, Rollins School of Public Health for assistance with the project implementation. The authors would also like to thank Dang Hong Linh, Nguyen Phuong Thao, and Le Thu Giang for their assistance to the Center for Creative Initiatives in Health and Population in project implementation. Finally, the authors would like to thank anonymous reviewers of prior versions of this paper and all study participants and volunteers without whom this project would not have been possible.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1395 KB - publichealth_v9i1e35116_app1.pdf]

Multimedia Appendix 2
GlobalConsent and adolescent health education (AHEAD) program screenshots.
[PDF File (Adobe PDF File), 1540 KB - publichealth_v9i1e35116_app2.pdf]

Multimedia Appendix 3
Documentation of adaptation of RealConsent to create GlobalConsent.
[DOCX File , 32 KB - publichealth_v9i1e35116_app3.docx]

Multimedia Appendix 4
Kernel density plots of behavioral outcomes measured as counts.
[DOCX File , 6159 KB - publichealth_v9i1e35116_app4.docx]

References

1. Basile K, Saltzman LE. Sexual violence surveillance: uniform definitions and recommended data elements. U.S. Department of Justice Office of Justice Programs. 2002. URL: https://www.ojp.gov/ncjrs/virtual-library/abstracts/sexual-violence-surveillance-uniform-definitions-and-recommended [accessed 2022-12-26]
2. Decker MR, Latimore AD, Yasutake S, Haviland M, Ahmed S, Blum RW, et al. Gender-based violence against adolescent and young adult women in low- and middle-income countries. J Adolesc Health 2015 Feb;56(2):188-196. [doi: 10.1016/j.jadohealth.2014.09.003] [Medline: 25620301]
3. Rennison C. Rape and sexual assault: reporting to police and medical attention, 1992-2000. U.S. Department of Justice Office of Justice Programs. 2002 Aug. URL: https://www.ojp.gov/ncjrs/virtual-library/abstracts/rape-and-sexual-assault-reporting-police-and-medical-attention-1992 [accessed 2022-12-26]
4. Gonzalez F, Benuto L, Casas J. Prevalence of interpersonal violence among Latinas: a systematic review. Trauma Violence Abuse 2020 Dec;21(5):977-990 [FREE Full text] [doi: 10.1177/1524838018806507] [Medline: 30554556]
5. Amar AF, Gennaro S. Dating violence in college women: associated physical injury, healthcare usage, and mental health symptoms. Nurs Res 2005;54(4):235-242. [doi: 10.1097/00006199-200507000-00005] [Medline: 16027566]
6. Wiersema B, Cohen M. Victim Costs and Consequences A New Look - Scholar's Choice Edition. Sacramento, California: Creative Media Partners, LLC; 1996.
7. Waibel G, Glück S. More than 13 million: mass mobilisation and gender politics in the Vietnam Women's Union. Gender Develop 2013 Jul 04;21(2):343-361 [FREE Full text] [doi: 10.1080/13552074.2013.802148]
8. Vu H, Schuler S, Hoang T, Quach T. Divorce in the context of domestic violence against women in Vietnam. Cult Health Sex 2014 Jun;16(6):634-647 [FREE Full text] [doi: 10.1080/13691058.2014.896948] [Medline: 24730640]
9. James-Hawkins L, Salazar K, Hennink M, Ha V, Yount K. Norms of masculinity and the cultural narrative of intimate partner violence among men in Vietnam. J Interpers Violence 2019 Nov;34(21-22):4421-4442 [FREE Full text] [doi: 10.1177/0886260516674941] [Medline: 29294621]
10. Yount K, Higgins E, VanderEnde K, Krause K, Minh T, Schuler S, et al. Men's perpetration of intimate partner violence in Vietnam. Men Masculinities 2015 Feb 24;19(1):64-84 [FREE Full text] [doi: 10.1177/1097184x15572896]
11. Lonsway KA, Klaw EL, Berg DR, Waldo CR, Kothari C, Mazurek CJ, et al. Beyond "no means no". J Interpers Violence 2016 Jul 02;13(1):73-92 [FREE Full text] [doi: 10.1177/088626098013001005]
12. Rich MD, Utley EA, Janke K, Moldoveanu M. "I'd rather be doing something else:" male resistance to rape prevention programs. J Men's Studies 2010 Oct 1;18(3). [doi: 10.3149/jms.1803.268]
13. Payne D, Lonsway K, Fitzgerald L. Rape myth acceptance: exploration of its structure and its measurement using theillinois rape myth acceptance scale. J Res Personality 1999 Mar;33(1):27-68 [FREE Full text] [doi: 10.1006/jrpe.1998.2238]
14. Amar A, Sutherland M, Laughon K. Gender differences in attitudes and beliefs associated with bystander behavior and sexual assault. J Forensic Nurs 2014;10(2):84-91; quiz E1. [doi: 10.1097/JFN.0000000000000024] [Medline: 24762431]
15. Deitz S, Blackwell K, Daley P, Bentley B. Measurement of empathy toward rape victims and rapists. J Personality Social Psychol 1982 Aug;43(2):372-384 [FREE Full text] [doi: 10.1037/0022-3514.43.2.372]
16. Amar AF, Tuccinardi N, Heislein J, Simpson S. Friends Helping Friends: a nonrandomized control trial of a peer-based response to dating violence. Nurs Outlook 2015 Jul;63(4):496-503. [doi: 10.1016/j.outlook.2015.01.004] [Medline: 26187089]

17. Berkowitz A. Fostering men's responsibility for preventing sexual assault. In: Preventing Violence in Relationships: Interventions Across the Life Span. Washington, D.C., United States: American Psychological Association; 2002.
18. Newlands R. A critical review of sexual violence prevention on college campuses. Acta Psychopathologica 2016 Apr 2;2(2):14. [doi: 10.4172/2469-6676.100040]
19. Yount K, Krause K, Miedema S. Preventing gender-based violence victimization in adolescent girls in lower-income countries: systematic review of reviews. Soc Sci Med 2017 Nov;192:1-13 [FREE Full text] [doi: 10.1016/j.socscimed.2017.08.038] [Medline: 28941786]
20. Mujal GN, Taylor ME, Fry JL, Gochez-Kerr TH, Weaver NL. A systematic review of bystander interventions for the prevention of sexual violence. Trauma Violence Abuse 2021 Apr 16;22(2):381-396. [doi: 10.1177/1524838019849587] [Medline: 31204606]
21. Graham LM, Embry V, Young B, Macy RJ, Moracco KE, Reyes HL, et al. Evaluations of prevention programs for sexual, dating, and intimate partner violence for boys and men: a systematic review. Trauma Violence Abuse 2021 Jul;22(3):439-465. [doi: 10.1177/1524838019851158] [Medline: 31262233]
22. Pérez-Martínez V, Marcos-Marcos J, Cerdán-Torregrosa A, Briones-Vozmediano E, Sanz-Barbero B, Davó-Blanes MC, et al. Positive masculinities and gender-based violence educational interventions among young people: a systematic review. Trauma Violence Abuse 2021 Jul 20:15248380211030242 (forthcoming). [doi: 10.1177/15248380211030242] [Medline: 34282677]
23. Salazar LF, Vivolo-Kantor A, Hardin J, Berkowitz A. A web-based sexual violence bystander intervention for male college students: randomized controlled trial. J Med Internet Res 2014 Sep 05;16(9):e203 [FREE Full text] [doi: 10.2196/jmir.3426] [Medline: 25198417]
24. Salazar L, Vivolo-Kantor A, McGroarty-Koon K. Formative research with college men to inform content and messages for a web-based sexual violence prevention program. Health Commun 2017 Sep;32(9):1133-1141 [FREE Full text] [doi: 10.1080/10410236.2016.1214219] [Medline: 27593559]
25. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164 [FREE Full text] [doi: 10.1177/1090198104263660] [Medline: 15090118]
26. Fabiano P, Perkins H, Berkowitz A, Linkenbach J, Stark C. Engaging men as social justice allies in ending violence against women: evidence for a social norms approach. J Am Coll Health 2003;52(3):105-112 [FREE Full text] [doi: 10.1080/07448480309595732] [Medline: 14992295]
27. Banyard V, Moynihan M, Plante E. Sexual violence prevention through bystander education: an experimental evaluation. J Community Psychol 2007 May;35(4):463-481 [FREE Full text] [doi: 10.1002/jcop.20159]
28. Lewis M, Neighbors C. Optimizing personalized normative feedback: the use of gender-specific referents. J Stud Alcohol Drugs 2007 Mar;68(2):228-237 [FREE Full text] [doi: 10.15288/jsad.2007.68.228] [Medline: 17286341]
29. Vladutiu C, Martin S, Macy R. College- or university-based sexual assault prevention programs: a review of program outcomes, characteristics, and recommendations. Trauma Violence Abuse 2011 Apr;12(2):67-86 [FREE Full text] [doi: 10.1177/1524838010390708] [Medline: 21196436]
30. Brecklin LR, Forde DR. A meta-analysis of rape education programs. Violence Vict 2001 Jan 01;16(3):303-321. [doi: 10.1891/0886-6708.16.3.303]
31. Gibbons R, Evans J. The evaluation of campus-based gender violence prevention programming: what we know about program effectiveness and implications for practitioners. National Online Resource Centre on Violence Against Women. 2013 Jan. URL: http://www.vawnet.org [accessed 2023-01-06]
32. Ritterband L, Thorndike F, Cox D, Kovatchev B, Gonder-Frederick L. A behavior change model for internet interventions. Ann Behav Med 2009 Aug;38(1):18-27 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
33. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
34. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387 [FREE Full text] [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
35. Hurling R, Fairley BW, Dias MB. Internet-based exercise intervention systems: are more interactive designs better? Psychol Health 2006 Dec;21(6):757-772. [doi: 10.1080/14768320600603257]
36. McKleroy V, Galbraith J, Cummings B, Jones P, Harshbarger C, Collins C, ADAPT Team. Adapting evidence-based behavioral interventions for new settings and target populations. AIDS Educ Prev 2006 Aug;18(4 Suppl A):59-73 [FREE Full text] [doi: 10.1521/aeap.2006.18.supp.59] [Medline: 16987089]
37. Yount KM, Minh TH, Trang QT, Cheong YF, Bergenfeld I, Sales JM. Preventing sexual violence in college men: a randomized-controlled trial of GlobalConsent. BMC Public Health 2020 Sep 01;20(1):1331 [FREE Full text] [doi: 10.1186/s12889-020-09454-2] [Medline: 32873262]
38. Bergenfeld I, Lanzas G, Trang QT, Sales J, Yount KM. Rape myths among university men and women in Vietnam: a qualitative study. J Interpers Violence 2022 Feb;37(3-4):NP1401-NP1431. [doi: 10.1177/0886260520928644] [Medline: 32529939]

39. Bergenfeld I, Tamler I, Sales J, Trang Q, Minh T, Yount K. Navigating changing norms around sex in dating relationships: a qualitative study of young people in Vietnam. Sexuality Culture 2021 Sep 08;26(2):514-530 [FREE Full text] [doi: 10.1007/s12119-021-09905-x]
40. Lewis P, Bergenfeld I, Thu Trang Q, Minh T, Sales J, Yount K. Gender norms and sexual consent in dating relationships: a qualitative study of university students in Vietnam. Cult Health Sex 2022 Mar;24(3):358-373 [FREE Full text] [doi: 10.1080/13691058.2020.1846078] [Medline: 33512308]
41. Bui T. 'Being a good girl': mother-daughter sexual communication in contemporary Vietnam. Cult Health Sex 2020 Jul;22(7):794-807 [FREE Full text] [doi: 10.1080/13691058.2020.1742383] [Medline: 32463343]
42. Ghuman S, Loi VM, Huy VT, Knodel J. Continuity and change in premarital sex in Vietnam. Int Fam Plan Perspect 2006 Dec;32(4):166-174 [FREE Full text] [doi: 10.1363/3216606] [Medline: 17237013]
43. Nguyen PA. ‘relationships based on love and relationships based on needs’: emerging trends in youth sex culture in contemporary urban Vietnam. Mod Asian Stud 2007 Jan 18;41(2):287-313 [FREE Full text] [doi: 10.1017/s0026749x05002258]
44. Koss M, Oros C. Sexual Experiences Survey: a research instrument investigating sexual aggression and victimization. J Consulting Clin Psychol 1982;50(3):455-457 [FREE Full text] [doi: 10.1037/0022-006x.50.3.455]
45. McMahon S, Allen C, Postmus J, McMahon S, Peterson N, Lowe Hoffman M. Measuring bystander attitudes and behavior to prevent sexual violence. J Am Coll Health 2014;62(1):58-66 [FREE Full text] [doi: 10.1080/07448481.2013.849258] [Medline: 24313697]
46. Tran N, Alink L, Van Berkel SR, Van Ijzendoorn MH. Child maltreatment in Vietnam: prevalence and cross-cultural comparison. J Aggression Maltreatment Trauma 2016 Dec 28;26(3):211-230 [FREE Full text] [doi: 10.1080/10926771.2016.1250851]
47. Livingstone S, Haddon L, Görzig A, Ólafsson K. Technical Report and User Guide: The 2010 EU Kids Online Survey. LSE, London: EU Kids Online; 2011.
48. StataCorp. Stata Statistical Software: Release 16. College Station, TX: StatCorp LP; 2019.
49. Garrity S. Sexual assault prevention programs for college-aged men: a critical evaluation. J Forensic Nurs 2011 Mar;7(1):40-48. [doi: 10.1111/j.1939-3938.2010.01094.x] [Medline: 21348933]
50. DeGue S, Valle L, Holt M, Massetti G, Matjasko J, Tharp A. A systematic review of primary prevention strategies for sexual violence perpetration. Aggress Violent Behav 2014;19(4):346-362 [FREE Full text] [doi: 10.1016/j.avb.2014.05.004] [Medline: 29606897]
51. Martin P, Alberti C, Gottot S, Bourmaud A, de La Rochebrochard E. Expert opinions on web-based peer education interventions for youth sexual health promotion: qualitative study. J Med Internet Res 2020 Nov 24;22(11):e18650 [FREE Full text] [doi: 10.2196/18650] [Medline: 33231552]
52. Duan YP, Wienert J, Hu C, Si GY, Lippke S. Web-based intervention for physical activity and fruit and vegetable intake among Chinese university students: a randomized controlled trial. J Med Internet Res 2017 Apr 10;19(4):e106 [FREE Full text] [doi: 10.2196/jmir.7152] [Medline: 28396306]
53. Tharp AT, DeGue S, Valle LA, Brookmeyer KA, Massetti GM, Matjasko JL. A systematic qualitative review of risk and protective factors for sexual violence perpetration. Trauma Violence Abuse 2013 Apr 27;14(2):133-167. [doi: 10.1177/1524838012470031] [Medline: 23275472]

Abbreviations

AHEAD: adolescent health education
DID: difference-in-difference
EBI: evidence-based intervention
LMIC: lower- and middle-income country
OR: odds ratio
RCT: randomized controlled trial

Edited by A Mavragani, H Bradley; submitted 02.12.21; peer-reviewed by T Lefèvre, H Williamson; comments to author 27.05.22; revised version received 21.07.22; accepted 25.11.22; published 27.01.23.

Please cite as:
Yount KM, Cheong YF, Bergenfeld I, Trang QT, Sales JM, Li Y, Minh TH
Impacts of GlobalConsent, a Web-Based Social Norms Edutainment Program, on Sexually Violent Behavior and Bystander Behavior Among University Men in Vietnam: Randomized Controlled Trial
JMIR Public Health Surveill 2023;9:e35116
URL: https://publichealth.jmir.org/2023/1/e35116
doi:10.2196/35116
PMID:36705965

Short Paper

Online Help-Seeking Among Youth Victims of Sexual Violence Before and During COVID-19 (2016-2021): Analysis of Hotline Use Trends

Kimberly L Goodman[1*], PhD; Kristyn Kamke[1*], PhD; Tara M Mullin[1], MA

Rape, Abuse, and Incest National Network, Washington, DC, United States

[*] these authors contributed equally

Corresponding Author:
Kimberly L Goodman, PhD
Rape, Abuse, and Incest National Network
1220 L St NW
Suite 505
Washington, DC, 20009
United States
Phone: 1 202 587 5347
Email: kimberlyg@rainn.org

Abstract

Background: Three years since the onset of COVID-19, pandemic-related trends in child sexual abuse (CSA) remain poorly understood. Common administrative surveillance metrics may have underestimated abuse during the pandemic, given youths' limited access to mandatory reporters. Research using anonymous service-use data showed increased violence-related online help-seeking but overlooked youth-specific help-seeking for CSA during COVID-19. Understanding pandemic-related trends in CSA can inform abuse detection practices and mental health service provision for youth victims.

Objective: The purpose of this study was to harness anonymous help-seeking data from the National Sexual Assault Online Hotline (NSAOH) to glean insights about CSA occurrence in the United States during the COVID-19 pandemic.

Methods: We used an archival sample of victims who contacted NSAOH from 2016 to 2021 (n=41,561). We examined differences in the proportion of youth and adult victims contacting NSAOH during the first COVID-19 year (March 2020 to February 2021) compared to the prior year (March 2019 to February 2020; n=11,719). Further, we compared key characteristics of hotline interactions among youth victims during the first COVID-19 year to the prior year (n=5913). Using joinpoint regression analysis, we examined linear trends in the number of monthly sampled youth and adult victims (excluding victims of unknown age) from 2016 to 2021 who discussed any victimization event (n=26,904) and who discussed recent events (ie, events occurring during the pandemic; n=9932).

Results: Most youth victims were abused by family members prior to (1013/1677, 60.4%) and after (2658/3661, 72.6%) the onset of COVID-19. The number of youth victims contacting NSAOH spiked in March 2020 and peaked in November 2020 for all youth (slope=28.2, 95% CI 18.7-37.7) and those discussing recent events (slope=17.4, 95% CI 11.1-23.6). We observed a decline in youth victims into spring 2021 for all youth (slope=–56.9, 95% CI –91.4 to –22.3) and those discussing recent events (slope=–33.7, 95% 47.3 to -20.0). The number of adult victims discussing any victimization event increased steadily from January 2018 through May 2021 (slope=3.6; 95% CI 2.9-4.2) and then declined (slope=–13.8, 95% CI –22.8 to –4.7). Trends were stable for adults discussing recent events.

Conclusions: This study extends the use of hotline data to understand the implications of the pandemic on CSA. We observed increased youth help-seeking through the NSAOH coinciding with the onset of COVID-19. Trends persisted when limiting analyses to recent victimization events, suggesting increased help-seeking reflected increased CSA during COVID-19. These findings underscore the utility of anonymous online services for youth currently experiencing abuse. Further, the findings support calls for increased youth mental health services and efforts to incorporate online chat into youth-targeted services.

(JMIR Public Health Surveill 2023;9:e44760) doi:10.2196/44760

KEYWORDS

child sexual abuse; hotlines; COVID-19; help-seeking; online services; child abuse; mental health well-being; child support; sexual abuse; mental health service; sexual violence

Introduction

As we approach a postpandemic world, understanding trends in child maltreatment during COVID-19 remains vital for informing practices to detect abuse and ameliorate the effects of trauma. Evidence of pandemic-related trends in child maltreatment is conflicting, resulting in part from limitations of traditional surveillance strategies. Apparent declines in child maltreatment reflected in emergency room visits [1] and official reports [2] may be attributed to avoidance of emergency rooms [3] and decreased access to mandatory reporters [4] during the pandemic. These shortcomings underscore the importance of examining child maltreatment during COVID-19 with consideration of youths' context and perspectives [5,6].

While youths' perspectives are often assessed through self-report survey methodologies, indicators of online help-seeking are potentially valuable. Prior to the pandemic, youth actively sought help for experiences of abuse via online and text-based hotlines [7,8], which they prefer over and perceive as more private than telephone-based hotlines [8,9]. Online services offer enhanced privacy, autonomy over disclosure, and easy and immediate support, which make them acceptable and accessible for youth experiencing crisis [10], especially during a pandemic.

Recent work has explored trends in violence-related online help-seeking during the pandemic. Google searches related to child maltreatment increased during the pandemic [11,12] but may not represent help-seeking by victims. Similarly, unsupervised learning of tweets revealed how family violence was discussed during the pandemic but could not delineate whether the source was victims, perpetrators, or organizations [13]. Online help-seeking via victim-centered resources, such as online hotlines (or helplines), offer an important complementary metric.

A study of multinational child helplines [14] lacked consensus as to the presence and direction of trends in violence-related contacts early in the pandemic but did not delineate phone-based from online- or text-based hotlines. Notably, outreach to ChildHelp increased for text-based but not phone-based contacts early in the pandemic [15]. While these studies demonstrated the promise of hotline data for understanding child maltreatment during the pandemic, their limited focus on overall hotline utilization conflated victim and nonvictim contacts. Research examining hotline use specific to youth victims during the pandemic is needed.

Amid these challenges, research on pandemic-related child maltreatment has largely overlooked child sexual abuse (CSA). However, concerns about victims' physical and social isolation and accessibility to perpetrators during stay-at-home orders [16] are particularly applicable to CSA [17]. To address this gap in understanding of pandemic-related trends in CSA, we examine archival data from the National Sexual Assault Online Hotline (NSAOH).

We advance knowledge by addressing limitations of previous work on 3 fronts. First, we examine online help-seeking for sexual violence among youth victims specifically. Second, we incorporate 5 years of observations (2016-2021), providing a broader picture of help-seeking patterns across the pre- and postpandemic onset time periods. Third, we delineate recent events to distinguish violence that occurred during the pandemic.

Methods

Data Collection

The NSAOH is a US-based online hotline providing 24/7 anonymous crisis intervention services via desktop and mobile internet chat. RAINN (Rape, Abuse & Incest National Network), an anti–sexual violence organization in the United States, created and has operated NSAOH since 2006 to serve victims of sexual violence.

Data collection via an in-depth online assessment (called the "session assessment") serves an integral role in hotline operations to identify areas for service improvement and staff training. Directly following the first chat session of their shift, staff record session information, such as event characteristics and topics discussed. Because information is not requested for the purpose of completing the assessment, data are often considered unknown (ie, missing). Staff receive instructions for assessment completion both within the assessment and through supplemental training.

Data in this study represent a sample of all hotline chats in which the visitor discussed an experience of sexual or other interpersonal violence (ie, victims) between January 2016 and December 2021 (N=41,561). Additional information about hotline procedures, the organizational context, and sample inclusion criteria are available in Multimedia Appendix 1.

Ethical Considerations

Data used in this study were originally collected as part of ongoing internal program evaluation; thus, informed consent was not required. Analysis of this archival data for the purposes of generating generalizable knowledge of online hotline users and sexual violence experiences was deemed exempt by the Advarra institutional review board under category 4 of the Revised Common Rule.

Statistical Analysis

We examined differences in the proportion of youth (ie, those younger than 18 years) and adult (aged 18 years and older) victims during the first COVID-19 year (March 2020 through February 2021) compared to the year prior (March 2019 through February 2020; n=11,719), with victims of unknown age excluded (5178/16,897, 31%). Among youth (n=5913), we compared perpetrator type (family vs nonfamily), perpetrator living status (currently living with vs not currently living with the victim), event timeframe (whether the event occurred within the last month, 1 month to 1 year ago, or over 1 year ago), and event frequency (repeated abuse vs single occurrence) during the first COVID-19 year compared to the year prior. These analyses used the chi-squared test with pairwise exclusion of missing data; analyses used SPSS (version 28.0; IBM Corp).

To contextualize trends in youth help-seeking, we used Joinpoint (version 4.9.1.0; National Cancer Institute) to examine linear trends in the number of monthly sampled victims by age group from 2016 to 2021 among all youth and adult victims (n=26,904; Multimedia Appendix 1 provides additional information about joinpoint analyses). Cases of unknown age were excluded (n=14,657; 22.7%-50.3% missingness per month). We repeated this analysis among youth and adult victims known to have discussed recent events that occurred within a month of contacting the hotline (n=9932). We also graphed the proportion of monthly sampled youth victims by year among all victims and those discussing recent events. Data were analyzed in spring 2022.

Results

Sample Characteristics

We observed a significant increase in the proportion of youth victims during the first COVID-19 year relative to the year prior (1871/4560, 41% vs 4042/7159, 56.5%). Among youth, we noted statistically significant but small shifts in the proportion of chats that involved family-perpetrated assaults (1013/1677, 60.4% vs 2658/3661, 72.6%), perpetrators currently living with the victim (1028/1733, 59.3% vs 2815/3841, 73.3%), recent assaults (ie, within the last month; 1017/1313, 77.5% vs 2381/2774, 85.8%), and repeated assaults (1397/1764, 79.2% vs 3250/3773, 86.1%). Table 1 shows full sample characteristics and missingness, as well as characteristics and comparisons using pairwise deletion.

Table 1. Comparison of victim and event characteristics in the pre-COVID year (March 1, 2019-February 29, 2020) and the COVID year (March 1, 2020-February 28, 2021). Information about response option coding for table categories is available in Multimedia Appendix 1.

	Full sample (n=16,897)		Pairwise deletion (n=11,719)			
	Pre-COVID year (n=7139), n (%)	COVID year (n=9758), n (%)	Pre-COVID year (n=4560), n (%)	COVID year (n=7159), n (%)	Effect size (Cramér's V)[a]	*P* value[a]
Victim age group					0.15	<.001
Youth	1871 (26.2)	4042 (41.4)	1871 (41)	4042 (56.5)		
Adult	2689 (37.7)	3117 (31.9)	2689 (59)	3117 (43.5)		
Unknown	2579 (36.1)	2599 (26.6)	N/A[b]	N/A		
Characteristics among youth (n=5913)						
Perpetrator					0.12	<.001
Family member	1013 (54.1)	2658 (65.8)	1013 (60.4)	2658 (72.6)		
Non–family member	664 (35.5)	1003 (24.8)	664 (39.6)	1003 (27.4)		
Unknown	194 (10.4)	381 (9.4)	N/A	N/A		
Victim living with perpetrator					0.14	<.001
Yes, currently[c]	1028 (54.9)	2815 (69.6)	1028 (59.3)	2815 (73.3)		
Not currently[c]	705 (37.7)	1026 (25.4)	705 (40.7)	1026 (26.7)		
Unknown	138 (7.4)	201 (5)	N/A	N/A		
Event timeframe					0.10	<.001
Within the last month	1017 (54.4)	2381 (58.9)	1017 (77.5)	2381 (85.8)		
1 month to 1 year ago	132 (7.1)	183 (4.5)	132 (10.1)	183 (6.6)		
Over a year ago	164 (8.8)	210 (5.2)	164 (12.5)	210 (7.6)		
Unknown	558 (29.8)	1268 (31.4)	N/A	N/A		
Event frequency					0.09	<.001
Repeated	1397 (74.7)	3250 (80.4)	1397 (79.2)	3250 (86.1)		
Single occurrence	367 (19.6)	523 (12.9)	367 (20.8)	523 (13.9)		
Unknown	107 (5.7)	269 (6.7)	N/A	N/A		

[a] Chi-square comparisons; unknown (“missing”) data were excluded in these comparisons.

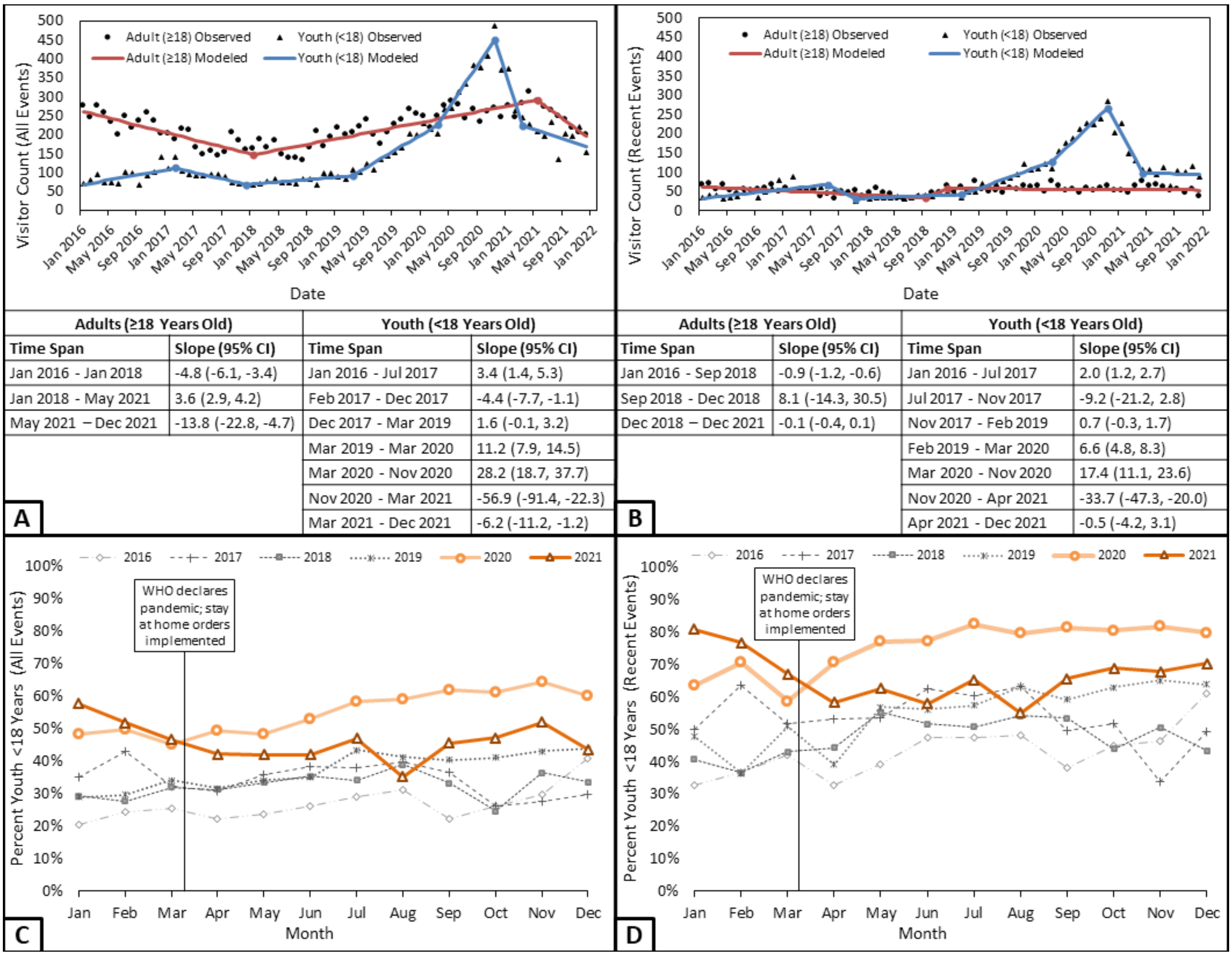
[b] N/A: not applicable.

[b] “Currently” means at the time of contacting the hotline.

Joinpoint Regression Analysis

Inflection points reveal variability in the strength or direction (or both) of trends in the number of victims by age in our sample (Figure 1). Prior to 2019, the number of youth victims fluctuated. In the year prior to COVID onset, the number of all youth victims and those discussing recent events increased by an average of 11.2 (95% CI 7.9-14.5) and 6.6 (95% CI 4.8, 8.3) per month, respectively. The most striking trend commenced in March 2020 at the onset of the pandemic: there were average monthly increases of 28.2 (95% CI 18.7-37.7) and 17.4 (95% CI 11.1-23.6) for all youth victims and youth discussing recent events, respectively. Both the number and proportion of youth victims peaked in November 2020. Subsequently, youth victims declined into early spring 2021, both overall (slope=–56.9; 95% CI –91.4 to –22.3) and for recent events (slope=–33.7; 95% CI –47.3 to –20.0). We observed an increase in all adult victims starting in January 2018 (slope=3.6; 95% CI 2.9-4.2), but this trend was consistent through May 2021, after which adult victims declined into December 2021 (slope=–13.8, 95% CI –22.8 to –4.7). The number of adult victims discussing recent events was largely consistent from 2016 to 2021.

Figure 1. Linear trends in National Sexual Assault Online Hotline use among youth. Analyses were limited to hotline visitors who identified as victims, with the sample reflecting the first chat of a staff member's shift. Panels A and C display all events, whereas panels B and D reflect chats discussing recent events (ie, occurring within the last month). Data are from January 2016 to December 2021 and were analyzed using Joinpoint (version 4.9.1.0; National Cancer Institute). WHO: World Health Organization.



A

Adults (≥18 Years Old)		Youth (<18 Years Old)	
Time Span	Slope (95% CI)	Time Span	Slope (95% CI)
Jan 2016 - Jan 2018	-4.8 (-6.1, -3.4)	Jan 2016 - Jul 2017	3.4 (1.4, 5.3)
Jan 2018 - May 2021	3.6 (2.9, 4.2)	Feb 2017 - Dec 2017	-4.4 (-7.7, -1.1)
May 2021 – Dec 2021	-13.8 (-22.8, -4.7)	Dec 2017 - Mar 2019	1.6 (-0.1, 3.2)
		Mar 2019 - Mar 2020	11.2 (7.9, 14.5)
		Mar 2020 - Nov 2020	28.2 (18.7, 37.7)
		Nov 2020 - Mar 2021	-56.9 (-91.4, -22.3)
		Mar 2021 - Dec 2021	-6.2 (-11.2, -1.2)

B

Adults (≥18 Years Old)		Youth (<18 Years Old)	
Time Span	Slope (95% CI)	Time Span	Slope (95% CI)
Jan 2016 - Sep 2018	-0.9 (-1.2, -0.6)	Jan 2016 - Jul 2017	2.0 (1.2, 2.7)
Sep 2018 - Dec 2018	8.1 (-14.3, 30.5)	Jul 2017 - Nov 2017	-9.2 (-21.2, 2.8)
Dec 2018 – Dec 2021	-0.1 (-0.4, 0.1)	Nov 2017 - Feb 2019	0.7 (-0.3, 1.7)
		Feb 2019 - Mar 2020	6.6 (4.8, 8.3)
		Mar 2020 - Nov 2020	17.4 (11.1, 23.6)
		Nov 2020 - Apr 2021	-33.7 (-47.3, -20.0)
		Apr 2021 - Dec 2021	-0.5 (-4.2, 3.1)

Discussion

Principal Findings

This study is the first, to our knowledge, to reveal increases in youth help-seeking for sexual violence coinciding with the onset of COVID-19. These increases peaked in November 2020 and were present for all youth victims and for youth victims discussing recent events. Although the number of adult victims increased, this trend was more gradual, preceded the onset of COVID-19, and was not present for those discussing recent events. Our finding that trends were sustained for recent events suggests an increase in youth affected by sexual violence during the pandemic rather than increased help-seeking among all victims for current and past events due to limited alternative support options.

While our findings seemingly conflict with declines in abuse-related hospital visits and official reports [1,2], evidence of pandemic-related trends in child maltreatment must be interpreted in light of the pandemic context [6]. Stay-at-home orders limited youths' access to formal services and mandated reporters, but online hotlines remained accessible to youth victims seeking help. Other studies used overall hotline utilization data to estimate pandemic-related child maltreatment [14,15] but conflated help-seeking among youth victims with that of adult victims and other hotline users (eg, supporters, general information seekers, or prank visitors) whose hotline use may have been differentially impacted by COVID-19. We also examined help-seeking for recent victimization, strengthening inferences that increased hotline contacts reflected help-seeking for violence experienced during the pandemic rather than past victimization. These procedures provide a model for how anonymous online hotline data can be used as a surveillance indicator for child maltreatment.

Most youth who contacted the NSAOH—both prior to and during the pandemic—discussed repeated experiences of violence perpetrated by family members or someone else living with them. During the pandemic, we observed small proportional increases in youths' chats that featured these characteristics, consistent with elevated violence driven by victims' increased accessibility to perpetrators in the home amid limited contact with trusted adults (eg, teachers). Future work should examine to what extent increased youth help-seeking reflected increased frequency and severity of ongoing violence predating the pandemic or cases of violence that began during the pandemic.

The number of chats from youth began to decrease after November 2020. This decline could relate to youth having a greater ability to avoid perpetrators or opportunities to disclose to alternative sources of support as schools reopened in the summer and fall of 2020. However, because reopening policies were variable across the United States, driven by state and local government, this trend should be interpreted with caution.

While reported trends were transitory, the effects of sexual abuse are not. Our findings reflect increased help-seeking by youth during the pandemic, an indication of rising demand for mental health services to treat trauma and its sequelae among youth. Given the national shortage of child mental health providers that predates COVID-19 [18], there is a pressing need to expand youth services to address the complex trauma resulting from CSA and other effects of COVID-19 [4].

Limitations

Joinpoint regression analysis allowed us to identify changes in the linear trend of victims' hotline use from 2016 to 2021. While we interpreted the reason for some of these changes, their causes cannot be definitively ascertained from our analysis. We could not differentiate repeat hotline users from new users, which would help delineate new cases of abuse. Our data collection protocol prioritized the delivery of victim-centered services. We did not solicit information from visitors for assessment purposes, resulting in some missing data and limited demographic information. Future work could incorporate anonymous collection of demographics prior to service provision to enhance the utility of hotline data as a surveillance indicator of child maltreatment. Understanding the interplay of hotline supply and demand will also be useful because changes in staff availability (ie, supply) could affect the accessibility and use of services.

Conclusions

Our study demonstrates the utility of anonymous online hotline data as a complementary surveillance indicator for child maltreatment that is transferrable to other public health priority areas (eg, mental health). Increased help-seeking amid decreased safety at home also speaks to the utility of online support options for youth experiencing abuse. Preparing for future public health emergencies necessitates consideration of additional online communication options that allow youth to privately disclose abuse to trusted adults. Future work should explore how to safely incorporate these communication options into existing services, such as online schooling or telemedicine platforms.

Acknowledgments

The authors' efforts in preparation of the manuscript were supported by RAINN (Rape, Abuse & Incest National Network). No external funding was received for this study.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the sensitivity of the data but aggregate data used in joinpoint analyses are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Expanded description of study methodology.
[DOCX File , 43 KB - publichealth_v9i1e44760_app1.docx]

References

1. Sege R, Stephens A. Child physical abuse did not increase during the pandemic. JAMA Pediatr 2022 Apr 01;176(4):338-340. [doi: 10.1001/jamapediatrics.2021.5476] [Medline: 34928320]
2. Marmor A, Cohen N, Katz C. Child maltreatment during COVID-19: Key conclusions and future directions based on a systematic literature review. Trauma Violence Abuse 2023 Apr;24(2):760-775 [FREE Full text] [doi: 10.1177/15248380211043818] [Medline: 34488521]

3. Jeffery MM, D'Onofrio G, Paek H, Platts-Mills TF, Soares WE, Hoppe JA, et al. Trends in emergency department visits and hospital admissions in health care systems in 5 states in the first months of the COVID-19 pandemic in the US. JAMA Intern Med 2020 Oct 01;180(10):1328-1333 [FREE Full text] [doi: 10.1001/jamainternmed.2020.3288] [Medline: 32744612]
4. Cénat JM, Dalexis RD. The complex trauma spectrum during the COVID-19 pandemic: A threat for children and adolescents' physical and mental health. Psychiatry Res 2020 Nov;293:113473 [FREE Full text] [doi: 10.1016/j.psychres.2020.113473] [Medline: 33198045]
5. Krause KH, Verlenden JV, Szucs LE, Swedo EA, Merlo CL, Niolon PH, et al. Disruptions to school and home life among high school students during the COVID-19 pandemic - adolescent behaviors and experiences survey, United States, January-June 2021. MMWR Suppl 2022 Apr 01;71(3):28-34 [FREE Full text] [doi: 10.15585/mmwr.su7103a5] [Medline: 35358164]
6. Korczak DJ, Madigan S, Vaillancourt T. Data divide-disentangling the role of the COVID-19 pandemic on child mental health and well-being. JAMA Pediatr 2022 Jul 01;176(7):635-636. [doi: 10.1001/jamapediatrics.2022.0791] [Medline: 35467721]
7. Schwab-Reese L, Kanuri N, Cash S. Child maltreatment disclosure to a text messaging-based crisis service: Content analysis. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e11306 [FREE Full text] [doi: 10.2196/11306] [Medline: 30907745]
8. Watling D, Batchelor S, Collyer B, Mathieu S, Ross V, Spence SH, et al. Help-seeking from a national youth helpline in Australia: An analysis of kids helpline contacts. Int J Environ Res Public Health 2021 Jun 03;18(11):6024 [FREE Full text] [doi: 10.3390/ijerph18116024] [Medline: 34205148]
9. Mathieu SL, Uddin R, Brady M, Batchelor S, Ross V, Spence SH, et al. Systematic review: The state of research into youth helplines. J Am Acad Child Adolesc Psychiatry 2021 Oct;60(10):1190-1233. [doi: 10.1016/j.jaac.2020.12.028] [Medline: 33383161]
10. Pretorius C, Chambers D, Coyle D. Young people's online help-seeking and mental health difficulties: Systematic narrative review. J Med Internet Res 2019 Nov 19;21(11):e13873 [FREE Full text] [doi: 10.2196/13873] [Medline: 31742562]
11. Riddell CA, Neumann K, Santaularia NJ, Farkas K, Ahern J, Mason SM. Excess Google searches for child abuse and intimate partner violence during the COVID-19 pandemic: Infoveillance approach. J Med Internet Res 2022 Jun 13;24(6):e36445 [FREE Full text] [doi: 10.2196/36445] [Medline: 35700024]
12. Riem MME, De Carli P, Guo J, Bakermans-Kranenburg MJ, van IJzendoorn MH, Lodder P. Internet searches for terms related to child maltreatment during COVID-19: Infodemiology approach. JMIR Pediatr Parent 2021 Jul 13;4(3):e27974 [FREE Full text] [doi: 10.2196/27974] [Medline: 34174779]
13. Xue J, Chen J, Chen C, Hu R, Zhu T. The hidden pandemic of family violence during COVID-19: Unsupervised learning of tweets. J Med Internet Res 2020 Nov 06;22(11):e24361 [FREE Full text] [doi: 10.2196/24361] [Medline: 33108315]
14. Petrowski N, Cappa C, Pereira A, Mason H, Daban RA. Violence against children during COVID-19: Assessing and understanding change in use of helplines. Child Abuse Negl 2021 Jun;116(Pt 2):104757 [FREE Full text] [doi: 10.1016/j.chiabu.2020.104757] [Medline: 33593632]
15. Ortiz R, Kishton R, Sinko L, Fingerman M, Moreland D, Wood J, et al. Assessing child abuse hotline inquiries in the wake of COVID-19: Answering the call. JAMA Pediatr 2021 Aug 01;175(8):859-861 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.0525] [Medline: 33938944]
16. Silliman Cohen RI, Bosk EA. Vulnerable youth and the COVID-19 pandemic. Pediatrics 2020 Jul;146(1):e20201306. [doi: 10.1542/peds.2020-1306] [Medline: 32345686]
17. Wortley R, Smallbone S. Applying situational principles to sexual offenses against children. In: Wortley R, Smallbone S. editors. Situational Prevention of Child Sexual Abuse. Monsey, NY: Criminal Justice Press; 2006:7-35.
18. Whitney DG, Peterson MD. US National and State-Level Prevalence of Mental Health Disorders and Disparities of Mental Health Care Use in Children. JAMA Pediatr 2019 Apr 01;173(4):389-391. [Medline: 30742204]

Abbreviations

CSA: child sexual abuse
NSAOH: National Sexual Assault Online Hotline
RAINN: Rape, Abuse & Incest National Network

Edited by A Mavragani, T Sanchez; submitted 02.12.22; peer-reviewed by G Lasiuk, YY Chen; comments to author 23.02.23; revised version received 22.03.23; accepted 16.06.23; published 11.08.23.

Please cite as:
Goodman KL, Kamke K, Mullin TM
Online Help-Seeking Among Youth Victims of Sexual Violence Before and During COVID-19 (2016-2021): Analysis of Hotline Use Trends
JMIR Public Health Surveill 2023;9:e44760
URL: https://publichealth.jmir.org/2023/1/e44760
doi:10.2196/44760
PMID:37566446

Original Paper

Firearm Possession Rates in Home Countries and Firearm Suicide Rates Among US- and Foreign-Born Suicide Decedents in the United States: Analysis of Combined Data from the National Violent Death Reporting System and the Small Arms Survey

In Han Song[1,2,3], PhD; Jin Hyuk Lee[4], PhD; Jee Soo Shin[5], MA

[1]ICONS Convergence Academy, Yonsei University, Seoul, Republic of Korea
[2]Graduate School of Social Welfare, Yonsei University, Seoul, Republic of Korea
[3]Faculty of Medicine, Vilnius University, Vilnius, Lithuania
[4]Department of Social Welfare, Ewha Womans University, Seoul, Republic of Korea
[5]Interdisciplinary Program of Social Welfare Policy, The Graduate School, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
In Han Song, PhD
ICONS Convergence Academy
Yonsei University
Appenzeller Hall #205
Seoul, 03722
Republic of Korea
Phone: 82 221236217
Email: isong@yonsei.ac.kr

Abstract

Background: Suicide by firearms is a serious public health issue in the United States. However, little research has been conducted on the relationship between cultural backgrounds and suicide by firearms, specifically in those born and raised in the United States compared to those who have immigrated to the United States.

Objective: To better understand the relationship between cultural backgrounds and suicide, this study aimed to examine firearm suicide rates among US- and foreign-born suicide decedents based on the firearm possession rate in the decedent’s home country.

Methods: Multivariate logistic regression was performed to analyze data of 28,895 suicide decedents from 37 states obtained from the 2017 National Violent Death Reporting System data set. The firearm possession rate in the home countries of foreign-born suicide decedents was obtained from the 2017 Small Arms Survey.

Results: The firearm suicide rate was about twice as high among US-born suicide decedents compared to their foreign-born counterparts. Meanwhile, suicide by hanging was about 75% higher among foreign-born compared to US-born suicide decedents. Those from countries with a low-to-medium firearm possession rate were significantly less likely to use firearms compared to US-born suicide decedents (adjusted odds ratio [AOR]=0.45, 95% CI 0.31-0.65, and AOR=0.46, 95% CI 0.39-0.53, respectively). Meanwhile, firearm suicide rates were not different between US- and foreign-born suicide decedents from countries with a similarly high firearm possession rate.

Conclusions: The results suggest that there is an association between using firearms as a means of suicide and the firearm possession rate in the decedent’s home country. Suicide by firearms in the United States needs to be understood in the sociocultural context related to firearm possession.

(JMIR Public Health Surveill 2023;9:e44211) doi:10.2196/44211

KEYWORDS

firearm suicide; US born; foreign born; means of suicide; firearm possession rate; suicide decedents

Introduction

Currently, suicide by firearms is a serious public health issue in the United States. According to the annual mortality statistics of the Centers for Disease Control and Prevention (CDC) [1], about 100 people die by suicide every day in the United States, with around 50% of them using firearms, 28.6% using poison, and 12.9% using hanging as a means of suicide. Firearm suicide rates in the United States have steadily increased over the past few years [2,3]. The firearm possession rate is defined as an estimate of civilian firearms per 100 individuals in a country [4]. States, regions, and countries with higher firearm possession rates are associated with higher firearm suicide rates [5]. According to Miller et al [6] and Stack [7], firearm possession rates are positively correlated with firearm suicide rates in the United States. In fact, the most common means of suicide in the United States is death by firearms. Other countries with high firearm possession rates, such as Uruguay and Montenegro, have also reported high of suicide by firearms [4,8]. Meanwhile, in countries that do not have high firearm possession rates, such as Asian or Nordic countries, poisoning by pesticides is the most common means of suicide. Similarly, suicide by hanging is the most common means of suicide in Eastern Europe [9]. One evidence-based suicide prevention strategy is to reduce the social acceptance and availability of lethal means for those at risk of suicide [10,11]. Investigating the context around suicide, such as accessibility to specific means, can provide important information when devising suicide prevention interventions [12]. This is because restricting the means of suicide is known to be effective in preventing suicides [13,14].

According to Wong et al [15], the social acceptance and availability of the means of suicide are associated with which means of suicide are more likely to be chosen. In this study, *social acceptance* refers to the cultural beliefs surrounding firearms, familiarity with firearms, the normality of firearm possession, and the role that firearms play in daily life. These factors may influence an individual’s own perceptions about firearms and increase or decrease the likelihood of their owning their own personal firearm [16]. Meanwhile, *availability* in this study is defined as the widespread availability of and accessibility to firearms. Using firearms as the most common means of suicide is associated with the high household firearm possession rate in the United States [17]. However, compared to the volume of research on the availability of firearms and firearm suicide rates, little research has been conducted to understand the relationship between the social acceptance of firearms and firearm suicide rates. The United States is a nation made up of immigrants from many diverse countries. Because of this, it is important to be cognizant of different demographics and their relationships and associations with firearms and, as a result, the likelihood to die by firearm suicide. According to Liu et al [18], US-born individuals have a significantly higher firearm suicide rate, while foreign-born individuals have a significantly higher rate of suicide by hanging, suffocation, jumping from heights, and sharp instruments. Moreover, it is suggested that immigrants are influenced by the culture surrounding firearms in their home countries [19]. Assuming that there is an association between the cultural familiarity with and social acceptance of firearms and the probability of an individual using a firearm as a means of suicide, this cultural distinction may be important in analyzing the difference between US- and foreign-born suicide decedents. Therefore, in this research, we compiled information about firearm possession rates in the home countries of immigrant populations in order to analyze the relationship between firearm familiarity and suicide behaviors of choosing a means of suicide. Despite the critical importance of identifying factors that may impact suicidal behaviors, to the best of our knowledge, little research has been conducted on the firearm possession rate by home country to understand the relationship between cultural/ethnic backgrounds and the means of suicide. Thus, our study aimed to better understand the cultural differences around firearms and how these may impact the likelihood of an individual dying by firearm suicide.

With the goal to determine the relationship between cultural familiarity and the means of suicide, we compared the means of suicide between US- and foreign-born suicide decedents. Focusing on suicides by firearms, in particular, we divided foreign-born individuals by the firearm possession rate in their home countries and compared them with US-born suicide decedents in order to understand the connection between the gun culture of the home country and the familiarity with firearms as a means of suicide. Our findings can serve as baseline data for devising suicide prevention strategies that focus on the restriction of access to various means of suicide, consider sociocultural backgrounds related to specific means, and customize the strategies to the sociocultural backgrounds of target groups.

Methods

Study Design

The National Violent Death Reporting System (NVDRS) is a state-based active surveillance system that provides a detailed account of violent deaths that occur in participating states. The NVDRS provides abundant data on the details of violent deaths in the United States, and it is the first national system to collect comprehensive information about suicides [20]. Moreover, the NVDRS is the nation’s only reporting system for complex events, such as homicide-suicide and mass killings, and offers far richer information about all other violent deaths compared to other data sources [21]. With support from the CDC, the system collects data from multiple sources, including death certificates, police reports, medical examiner/coroner reports, and, sometimes, crime lab information [22].

This study used restricted data of decedents provided by the 2017 NVDRS and examined only deaths resulting from suicide. In 2017, 37 states (Alaska, Arizona, California, Colorado, Connecticut, Delaware, District of Columbia, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, Utah, Vermont, Virginia, Washington, West Virginia, and Wisconsin) contributed data to the NVDRS. The 2017 NVDRS data include decedents who died from violent deaths in 2017. Specifically, the data include 33,257 individuals who died by suicide; data on the means of suicide and the home country are missing for 4362 (13.1%) decedents. Our final sample consisted of 28,895 decedents after removing individuals with missing data. So, this study used complete case analysis (CCA) to deal with missing values.

Measurements

The means of suicide were used as the outcome variable in this study. In the NVDRS data, the means of suicide include firearms, hanging or suffocation, poisoning, jumping from a height, sharp instruments, drowning, and other. However, this study used firearms, hanging or suffocation, and poisoning as the main outcome variables. Nativity, the independent variable, was divided into US- and foreign-born suicide decedents. People born in the United States, Puerto Rico, and a US Island Area (Guam, the Commonwealth of the Northern Mariana Islands, or the US Virgin Islands) were defined as US born, and those born outside the United States were defined as foreign born [23].

The firearm possession rate by home country was obtained from the estimate of civilian firearms per 100 individuals per country provided by the 2017 Small Arms Survey (SAS) [4]. The SAS is an independent research project located at the Graduate Institute of International and Development Studies in Geneva, Switzerland. The SAS uses data from various sources, such as official documents, studies, questionnaires, public opinion surveys, news reports, and expert correspondence, to estimate civilian firearm ownership. It systematically integrates these data to generate total estimates for each country, while excluding extreme high and low numbers as outliers, when feasible. This data set is used because it provides information about firearm possession rates for most countries. We matched the information about firearm possession rates to the home countries of the foreign-born individuals. A country with a firearm possession rate of <1 SD from the mean value was coded as "low," of >1 SD from the mean value was coded as "high," and of 1 SD from the mean value was coded as "medium." Sociodemographic variables included race, age, gender, region, education, and marital status.

Ethical Considerations

This study was exempt from human subject research ethics by the Institutional Review Board of Yonsei University (7001988-202010-HR-1013-01E) due to secondary analysis of research data. The data were derived from routine injury mortality surveillance. The NVDRS is incident based and links all victims and alleged perpetrators (suspects) associated with a given incident in 1 record. To fully characterize incidents, states collect information about deaths from numerous data sources. These sources include death certificates, coroner/medical examiner reports, and law enforcement reports [22].

Statistical Analysis

A descriptive analysis was conducted on the general characteristics of suicide decedents. Rates per 100,000 people (by nativity, race, sex, and age group) were calculated using the census data of the 2017 American Community Survey. Logistic regression was performed to analyze the association between nativity and the means of suicide. The independent variables of this study were nativity (US born/foreign born) and the firearm possession rate in the home country. The dependent variable was the choice of a means of suicide (firearm, hanging, and poisoning). Other means of suicide were excluded from the logistic analysis in order to explain the 3 major means of suicide.

First, logistic regression was performed to compare the correlation between sociodemographic variables and the means of suicide. Second, we analyzed the association between nativity and the means of suicide after adjusting for gender, age, place of residence, education, and marital status. Next, we analyzed the association between the firearm possession rate in the home country and the means of suicide. We used the US-born suicide decedents as a reference group for the 2 independent variables and examined how the choice of a means of suicide for the 2 independent variables is different compared to the US-born suicide decedents. The adjusted odds ratio (AOR) and 95% CI were presented, and statistical significance at the α level of .05 was indicated by 95% CIs that did not include 1. All analyses were performed using STATA, version 13.

Results

Demographic Characteristics of Suicide Decedents

Tables 1 and 2 show the demographic characteristics of suicide decedents in 37 states presented by the 2017 NVDRS. Of the total 28,895 deaths by suicide, the firearm suicide rate was 5.9 per 100,000 people, the hanging suicide rate was 3.7 per 100,000 people, the poisoning suicide rate was 1.6 per 100,000 people, and the rate of suicide by other means was 1.0 per 100,000 people. Firearm suicide rates were higher than hanging and poisoning suicide rates among US-born, White, Black, male, and 18-year-old or older decedents. The hanging suicide rates were higher compared to firearm and poisoning suicide rates among were foreign born; belonged to Hispanic, Asian/Pacific Islander, or other ethnic groups; and were less than 18 years old. The poisoning suicide rates were slightly higher than firearm and hanging suicide rates among women.

Table 1. Demographics of suicide decedents by means of suicide: 2017 NVDRS[a].

Characteristics	Decedents (N=28,895), n (%)	Firearms (n=14,041, 48.6%), n (%)	Hanging (n=8679, 30.0%), n (%)	Poisoning (n=3900, 13.5%), n (%)	Other[b] (n=2275, 7.9%), n (%)
Nativity					
US born	27,252 (94.3)	13,561 (96.6)	7922 (91.3)	3716 (95.3)	2053 (90.2)
Foreign born	1643 (5.7)	480 (3.4)	757 (8.7)	184 (4.7)	222 (9.8)
Race					
White	23,681 (82.0)	12,062 (85.9)	6534 (75.3)	3414 (87.5)	1671 (73.5)
Black	1812 (6.3)	874 (6.2)	547 (6.3)	183 (4.7)	208 (9.1)
Hispanic	2120 (7.3)	668 (4.8)	1024 (11.8)	186 (4.8)	242 (10.6)
Asian/Pacific Islander	566 (2.0)	148 (1.1)	265 (3.1)	52 (1.3)	101 (4.4)
Other	716 (2.5)	289 (2.1)	309 (3.6)	65 (1.7)	53 (2.3)
Sex					
Male	22,454 (77.7)	12,164 (86.6)	6723 (77.5)	1890 (48.5)	1677 (73.7)
Female	6441 (22.3)	1877 (13.4)	1956 (22.5)	2010 (51.5)	598 (26.3)
Age (years)					
<18	1092 (3.8)	415 (3.0)	551 (6.3)	72 (1.8)	54 (2.4)
18-24	3149 (10.9)	1498 (10.7)	1159 (13.4)	224 (5.7)	268 (11.8)
25-44	9443 (32.7)	3922 (27.9)	3668 (42.3)	1101 (28.2)	752 (33.1)
45-64	10,125 (35.0)	4750 (33.8)	2638 (30.4)	1848 (47.4)	889 (39.1)
≥65	5086 (17.6)	3456 (24.6)	663 (7.6)	655 (16.8)	312 (13.7)

[a] NVDRS: National Violent Death Reporting System.

[b] “Other” includes jumping from heights, sharp instruments, drowning, and other means.

Table 2. Rate of suicide per 100,000 people by means of suicide: 2017 NVDRS[a].

Characteristics	Decedents, rate/100,000	Firearms, rate/100,000	Hanging, rate/100,000	Poisoning, rate/100,000	Other[b], rate/100,000
Total	12.2	5.9	3.7	1.6	1.0
Nativity					
US born	13.3	6.6	3.9	1.8	1.0
Foreign born	5.0	1.5	2.3	0.6	0.7
Race					
White	15.8	8.1	4.4	2.3	1.1
Black	6.8	3.3	2.0	0.7	0.8
Hispanic	5.5	1.7	2.6	0.5	0.6
Asian/Pacific Islander	3.8	1.0	1.8	0.4	0.7
Other	9.0	3.6	3.9	0.8	0.7
Sex					
Male	19.2	10.4	5.7	1.6	1.4
Female	5.3	1.6	1.6	1.7	0.5
Age (years)					
<18	2.0	0.8	1.0	0.1	0.1
18-24	13.6	6.5	5.0	1.0	1.2
25-44	15.0	6.2	5.8	1.8	1.2
45-64	16.2	7.6	4.2	2.9	1.4
≥65	14.5	9.8	1.9	1.9	0.9

[a] NVDRS: National Violent Death Reporting System.

[b] "Other" includes jumping from heights, sharp instruments, drowning, and other means.

First, we examined the correlation between demographic variables and the choice of a means of suicide. The results are presented in Tables 3-5. Compared to Whites, Blacks were more likely to use firearms but less likely to use hanging or poisoning as a means of suicide. Compared to Whites, Hispanic, Asian/Pacific Islander, and other ethnic minorities were more likely to use hanging or poisoning and less likely to select firearms as a means of suicide. Older people were more likely to select firearms or poisoning and less likely to choose hanging as a means of suicide. Compared to men, women selected poisoning more than firearms. Urban dwellers were less likely to choose firearms and more likely to use hanging or poisoning. The higher the education level, the fewer the number of suicides by hanging and the more the number of suicides by poisoning. Those with spouses used firearms but not as much as poisoning as a means of suicide.

Next, we tested the association between nativity and the means of suicide after adjusting for demographic variables. The results are shown in Tables 3-5 and Figure 1. We compared the choice of a means of suicide between US- and foreign-born suicide decedents. Compared to US-born suicide decedents, their foreign-born counterparts were less likely to use firearms and more likely to select hanging as a means of suicide. Next, we compared the means of suicide between US- and foreign-born suicide decedents based on the firearm possession rate in the home countries of the latter. People who immigrated from countries with a low firearm possession rate selected firearms more than they did hanging as a means of suicide compared to US-born suicide decedents. People who were from countries with a medium firearm possession rate chose firearms less than they did hanging as a means of suicide compared to US-born suicide decedents. However, the difference between foreign-born suicide decedents from countries with a high firearm possession rate and their US-born counterparts was not statistically significant.

Table 3. Statistics of suicide by firearms among suicide decedents from the United States and other countries with different levels of the firearm possession rate: 2017 NVDRS[a].

Characteristics	Model 1		Model 2	
	AOR[b] (95% CI)	*P* value	AOR (95% CI)	*P* value
Foreign born (reference: US born)	0.52 (0.46-0.59)	<.001	N/A[c]	N/A
Firearm possession rate in the home country (reference: US born)				
Low	N/A	N/A	0.45 (0.31-0.65)	<.001
Medium	N/A	N/A	0.46 (0.39-0.53)	<.001
High	N/A	N/A	0.83 (0.65-1.08)	.162
Race (reference: White)				
Black	1.20 (1.09-1.33)	<.001	1.21 (1.1-1.35)	<.001
Hispanic	0.58 (0.52-0.64)	<.001	0.60 (0.54-0.66)	<.001
Asian/Pacific Islander	0.62 (0.51-0.77)	<.001	0.67 (0.54-0.83)	<.001
Other	0.77 (0.66-0.91)	.002[d]	0.78 (0.66-0.91)	.002[d]
Age	1.01 (1.01-1.02)	<.001	1.01 (1.01-1.02)	<.001
Female	0.35 (0.33-0.37)	<.001	0.35 (0.33-0.37)	<.001
Urban dweller	0.65 (0.61-0.69)	<.001	0.65 (0.61-0.69)	<.001
Education level	0.99 (0.97-1.01)	.160	0.99 (0.97-1.01)	.144
Married	1.40 (1.32-1.48)	<.001	1.40 (1.33-1.48)	<.001

[a] NVDRS: National Violent Death Reporting System.

[b] AOR: adjusted odds ratio.

[c] N/A: not applicable.

[d] *P*<.01.

Table 4. Statistics of suicide by hanging among suicide decedents from the United States and other countries with different levels of the firearm possession rate: 2017 NVDRS[a].

Characteristics	Model 1		Model 2	
	AOR[b] (95% CI)	*P* value	AOR (95% CI)	*P* value
Foreign born (reference: US born)	1.75 (1.56-1.97)	<.001	N/A[c]	N/A
Firearm possession rate in the home country (reference: US born)				
Low	N/A	N/A	2.90 (2.10-4.00)	<.001
Medium	N/A	N/A	1.83 (1.60-2.10)	<.001
High	N/A	N/A	1.12 (0.85-1.48)	.435
Race (reference: White)				
Black	0.80 (0.71-0.89)	<.001	0.79 (0.70-0.88)	<.001
Hispanic	1.62 (1.47-1.79)	<.001	1.60 (1.45-1.77)	<.001
Asian/Pacific Islander	1.36 (1.12-1.64)	.002[d]	1.21 (0.99-1.47)	.062
Other	1.45 (1.24-1.70)	<.001	1.44 (1.23-1.68)	<.001
Age	0.98 (0.97-0.98)	<.001	0.98 (0.97-0.98)	<.001
Female	1.01 (0.95-1.08)	.718	1.01 (0.95-1.08)	.108
Urban dweller	1.27 (1.18-1.36)	<.001	1.27 (1.18-1.36)	<.001
Education level	0.95 (0.93-0.97)	<.001	0.95 (0.93-0.97)	<.001
Married	0.95 (0.89-1.01)	.120	0.95 (0.89-1.01)	.108

[a] NVDRS: National Violent Death Reporting System.

[b] AOR: adjusted odds ratio.

[c] N/A: not applicable.

[d] *P*<.01.

Table 5. Statistics of suicide by poisoning among suicide decedents from the United States and other countries with different levels of the firearm possession rate: 2017 NVDRS[a].

Characteristics	Model 1		Model 2	
	AOR[b] (95% CI)	*P* value	AOR (95% CI)	*P* value
Foreign born (reference: US born)	0.85 (0.71-1.03)	.095	N/A[c]	N/A
Firearm possession rate in the home country (reference: US born)				
Low	N/A	N/A	0.64 (0.36-1.15)	.135
Medium	N/A	N/A	0.87 (0.70-1.09)	.223
High	N/A	N/A	0.87 (0.61-1.24)	.434
Race (reference: White)				
Black	0.69 (0.58-0.82)	<.001	0.69 (0.59-0.82)	<.001
Hispanic	0.69 (0.58-0.81)	<.001	0.68 (0.58-0.81)	<.001
Asian/Pacific Islander	0.55 (0.39-0.75)	<.001	0.57 (0.41-0.80)	.001[d]
Other	0.67 (0.51-0.88)	.004[d]	0.67 (0.51-0.88)	.004[d]
Age	1.01 (1.01-1.02)	<.001	1.01 (1.01-1.02)	<.001
Female	5.05 (4.69-5.43)	<.001	5.05 (4.69-5.44)	<.001
Urban dweller	1.32 (1.20-1.46)	<.001	1.32 (1.20-1.46)	<.001
Education level	1.05 (1.02-1.07)	<.001	1.05 (1.02-1.07)	<.001
Married	0.67 (0.62-0.73)	<.001	0.67 (0.62-0.73)	<.001

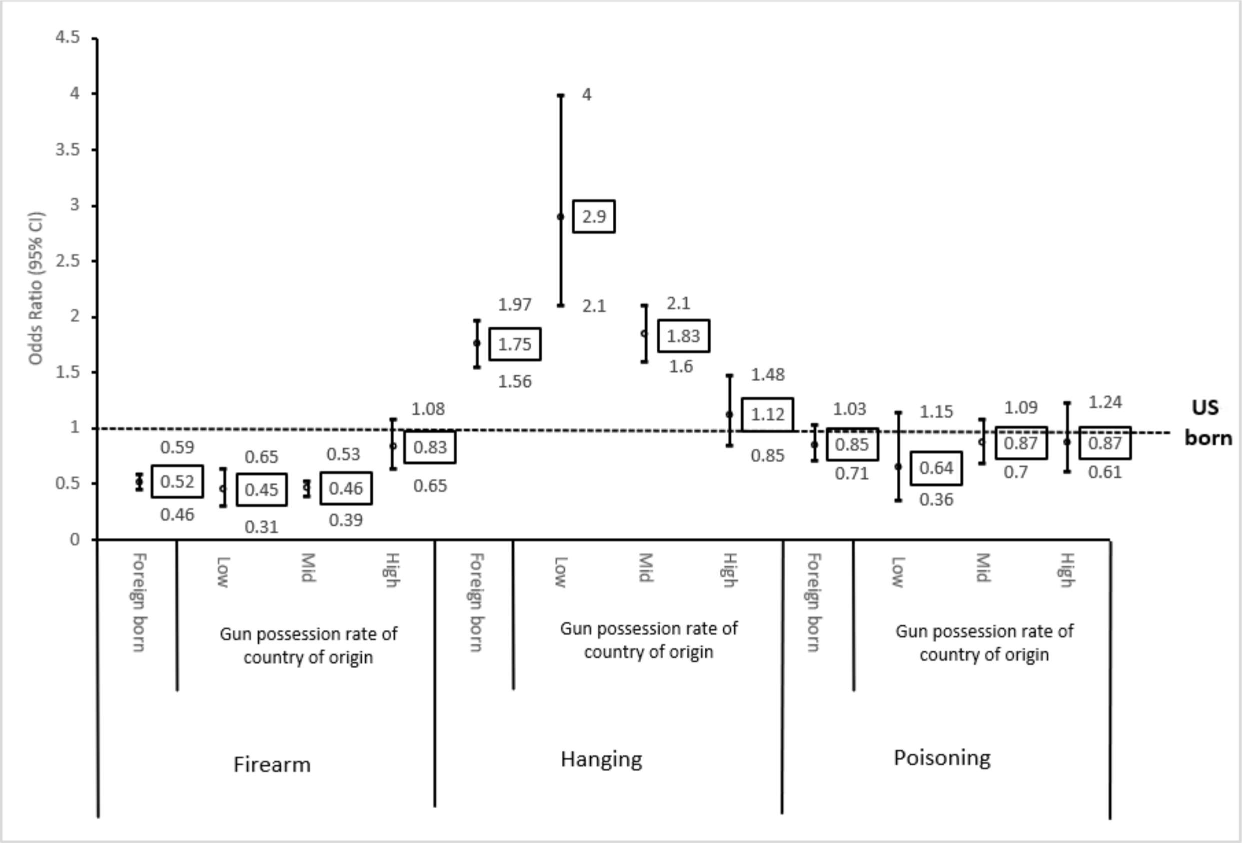
[a] NVDRS: National Violent Death Reporting System.

[b] AOR: adjusted odds ratio.

[c] N/A: not applicable.

[d] $P<.01$.

Figure 1. Means of suicide among suicide decedents from the United States and other countries with different levels of the firearm possession rate: 2017 National Violent Death Reporting System.



Discussion

Principal Findings

This research examined the data of 28,895 suicide decedents from 37 states included in the 2017 NVDRS with the goal of comparing the means of suicide between US- and foreign-born American suicide decedents and analyzing whether these differences could be explained by the culture surrounding firearms in their home countries, specifically firearm possession rates that illustrate a cultural familiarity with firearms. We found statistically significant differences in the means of suicide between US- and foreign-born individuals. The firearm suicide rate was about twice as high among US-born suicide decedents compared to their foreign-born counterparts in 2017, while the hanging suicide rate was about 75% higher among foreign-born suicide decedents compared to their US-born counterparts. We divided the home countries of the foreign-born suicide decedents into 3 groups by their firearm possession rate (low, medium, and high) and compared their means of suicide with those of the US-born suicide decedents. Those from countries with low and medium firearm possession rates were significantly less likely to use firearms compared to the US-born suicide decedents (AOR=0.45, 95% CI 0.31-0.65, and AOR=0.46, 95% CI 0.39-0.53, respectively). However, there was no statistical difference in the use of firearms between foreign-born suicide decedents from countries with a high firearm possession rate and their US-born counterparts. Foreign-born individuals from countries with a low or medium firearm possession rate were significantly more likely to use hanging compared to a firearm (AOR=2.90, 95% CI2.10-4.00, and AOR=1.83, 95% CI 1.60-2.10, respectively). In the case of suicide by poisoning, there was no statistical difference between US-born and all foreign-born suicide decedents. Compared to US-born individuals, immigrants from countries with a high firearm possession rate had no significant difference in firearm suicide rates and immigrants from countries with a low firearm possession rate had a significantly lower firearm suicide rate.

The findings of this research are in line with those of Liu et al [18], who found that US-born suicide decedents have a greater firearm suicide rate and a lower hanging suicide rate compared to their foreign-born counterparts and that there is no significant difference in the poisoning suicide rate between US- and foreign-born Americans. Furthermore, our results were confirmed by Wong et al [15], who found that US-born individuals have a significantly greater firearm suicide rate, a lower hanging suicide rate, and no significant difference in the poisoning suicide rate compared to foreign-born Asian/Pacific Islander Americans. In addition, this research compared cultural differences regarding firearms for the various home countries of immigrants residing in the United States who died by suicide. This confirms that there is an association between the firearm possession rate in the home country and the prevalence of various means of suicide. Furthermore, sociocultural familiarity with firearms is associated with a greater likelihood to die by firearm suicide [15]. Those born in a country with a high firearm possession rate are more likely to be exposed to and familiar with firearms, becoming more accepting of guns as a means of suicide, compared to those born in a country with a low firearm possession rate. This argument is supported by a number of studies: Conner et al [24] concluded that the means of suicide is influenced by the culture and gun policy of the home country, Kaplan and Geling [25] described the cultural origin of firearm suicide, Price et al [26] argued that firearm suicide is related to the gun laws of the country where an individual lives, Miller et al [27] supported that both firearm suicide and total suicide rates are associated with state and regional gun prevalence, and Saunders et al [19] suggested that immigrants are influenced by the firearm possession culture of their home country.

Considering that cultural perceptions of firearms can make certain means of suicide appear more manageable [28], it is understood that firearms are most commonly used for suicide in the United States because of their availability [29]. In contrast, hanging is the most common means of suicide in several Asian countries, including Japan, South Korea, and the Philippines [17]. In addition, this research found that the rate of suicide by hanging is high among immigrants from countries with a low firearm possession rate compared to US firearm possession rates.

It is necessary to devise suicide prevention strategies that focus on restricting access to dangerous items for at-risk individuals and that also consider the sociocultural implications of this topic. This is important as certain cultural backgrounds may be more sensitive to the exposure of the means of suicide. Suicide prevention efforts targeting war veterans include the establishment of suicide prevention campaigns and laws, such as the Joshua Omvig Veterans Suicide Prevention Act. In addition, the Veterans Agency and the Department of Defense provide free gun-safe storage to veterans at risk of dying by firearm suicide [30]. This measure is expected to be beneficial, given prior research that has proven the efficacy of proper gun safety and storage in reducing suicides among veterans [31]. Furthermore, it is imperative that suicide prevention campaigns be customized to best fit the needs of various sociocultural backgrounds. The World Health Organization’s National Strategy of Suicide Prevention notes the importance of restricting access to the means of suicide but fails to consider the way in which gun violence disproportionately affects certain sociocultural and ethnic backgrounds. There should be further research on this intersection between suicide rates and various sociocultural and ethnic groups in order to develop more effective suicide prevention campaigns. Wong et al [15] suggest that given the prevalence of suicide by firearms among the US-born population and suicide by hanging among immigrants, public health professionals should advise individuals with suicidal thoughts to remove potentially dangerous items from their environments. Based on our results, there seems to be an association between the culture surrounding firearms that an individual grew up in and the likelihood of them choosing suicide by firearms as opposed to other means. Therefore, public health professionals may need to follow culturally sensitive approaches to establish suicide prevention policies and strategies by reflecting on the cultural backgrounds of the population. In addition, when devising gun safety campaigns, states with high proportions of foreign-born individuals should consider their cultural backgrounds. Foreign-born individuals from countries with high firearm possession rates and US-born individuals are priority populations. In the case of both US- and foreign-born individuals from cultures highly familiar with firearms at risk of suicide, reducing access to lethal means is an evidence-based suicide prevention initiative [10,11]. Ideally, firearm accessibility should be limited in the United States to suppress the innate and widespread culture of firearm possession that has arisen in the country and has led to increased gun violence throughout the country. To alleviate this, federal gun policies must be enforced to reduce gun violence in the United States, including suicide by firearms. Such policies could include banning guns in public spaces, requiring a stricter process prior to buying a gun, banning high-capacity magazines and assault rifles, enforcing that all gun owners must store their guns in an approved gun storage unit, and requiring gun safety courses for everyone who wishes to own a gun. These measures will evidently reduce gun violence and particularly suicide by firearms [22]. The study is meaningful because it shows empirical results that the cultural background is related to suicide by firearms. Although the causal relationship is limited, the significant correlation means that cultural perspectives must be considered in planning firearm suicide prevention policies and interventions in a multicultural society, such as the United States.

Limitations

There are several limitations of this research. First, the NVDRS data used in this study had 13% missing values due to difficulties in obtaining some sensitive information about suicide decedents. To deal with the missing values, this study used CCA. This is because compared to multiple imputation, CCA has been reported to provide unbiased estimates in cases where the missing percentage is less than 20%, and this missing percentage refers to missing completed at random (MCAR) or missing completed random (MAR) [32]. Second, this research was premised around the cultural attitudes of firearms of the home countries of various immigrant suicide decedents. However, when analyzing these decedents, other crucial demographic factors were not considered regarding their culture. Because of this, other exogenous factors, such as the length of exposure to the original culture, the maintenance of the original culture after immigration to the United States, and the length of exposure to US culture, will need to be considered in future research. Third, we were unable to examine the physical availability of or accessibility to firearms at the regional level in the United States due to a lack of relevant data, such as the firearm possession rate at the state level. Fourth, the findings of this research cannot be generalized to the entire country, because the 2017 NVDRS, which contains the most recent scientific data that include detailed information about suicide, obtained data from only 37 states. Fifth, the foreign-born suicide decedents in this research include immigrants with diverse citizenship/residency statuses, including permanent residents, temporary workers, students, visitors, and illegal immigrants. This indicates that their use of a means of suicide may have been impacted by other extenuating factors, such as practical barriers to purchasing guns, rather than cultural familiarity. Sixth, this cross-national comparative study did not consider state-specific differences in the United States, because national-level data were analyzed. Further studies on suicide by firearms in the United States will benefit from considering state-specific gun policies or contexts. Lastly, there was a lack of data on statistics that internationally compare suicide trends by country with regard to US- and foreign-born residents of the country and their suicide rates. This study used the firearm possession rates in the home countries of suicide decedents to indirectly identify potential associations. If such statistics exist and are used in future studies, more direct and robust implications can be drawn.

Conclusion

Despite limitations, this research has contributed to expanding our understanding of how cultural influences affect higher firearm suicide rates in the United States. Our research shows that nativity is related to firearm suicide rates. Those from countries with high firearm possession rates show a similar propensity for firearm suicide as US-born Americans. The firearm suicide rate is significantly lower among immigrants from countries with a low firearm possession rate. Based on these findings, it is arguable that the high prevalence of firearm possession in the United States contributes to the country's high firearm suicide rate by promoting public familiarity toward firearms. Consequentially, US-born individuals are desensitized to gun violence, including suicide by firearms. This is based on the finding that foreign-born individuals from countries with low firearm possession rates have a lower tendency to use firearms to die by suicide compared to US-born individuals. However, immigrants from countries with high firearm possession rates are more likely to die by firearm suicide, which supports the association between cultural norms surrounding firearms and the likelihood of dying by firearm suicide [24]. This study has furthered our understanding of the cultural background that contributes to the process by which individuals choose their means of suicide. We hope that this research will contribute to an effective reduction in firearm suicide in the United States as well as improve suicide prevention efforts throughout the country.

Acknowledgments

The corresponding author thanks Professor David Hemenway and Harvard Injury Control Research Center for the opportunity to conduct research on National Violent Death Reporting System (NVDRS) while at Harvard University T.H. Chan School of Public Health, Boston, MA, USA.

Data Availability

The data used in this analysis are available upon request from the Centers for Disease Control and Prevention (CDC) through its restricted-access data process. Our use of these restricted-access National Violent Death Reporting System (NVDRS) data is governed by a data use agreement (DUA) with the CDC. This DUA legally prohibits us from sharing these data with outside investigators. Any investigator can gain access to these restricted-access NVDRS data by contacting the CDC and following the procedures outlined [33].

Conflicts of Interest

None declared.

References

1. Fatal injury reports: national, regional, and state (restricted), 1999–2018. Centers for Disease Control and Prevention. 2018. URL: https://webappa.cdc.gov/sasweb/ncipc/mortrate.html [accessed 2020-09-15]
2. Lam E, Moreno M, Bennett E, Rowhani-Rahbar A. Receptiveness and responsiveness toward using social media for safe firearm storage outreach: mixed methods study. J Med Internet Res 2021 Jun 18;23(6):e24458 [FREE Full text] [doi: 10.2196/24458] [Medline: 34142974]
3. Fowler KA, Dahlberg LL, Haileyesus T, Gutierrez C, Bacon S. Childhood firearm injuries in the United States. Pediatrics 2017 Jul;140(1):e20163486 [FREE Full text] [doi: 10.1542/peds.2016-3486] [Medline: 28630118]
4. Karp A. Estimating global civilian-held firearms numbers: briefing paper. Small Arms Survey. 2018 Jun. URL: http://www.smallarmssurvey.org/fileadmin/docs/Weapons_and_Markets/Tools/Firearms_holdings/SAS-BP-Civilian-held-firearms-annexe.pdf [accessed 2023-09-20]
5. National Research Council, Division of Behavioral and Social Sciences and Education, Committee on Law and Justice, Committee to Improve Research Information and Data on Firearms, Wellford CF, Pepper JV, et al. Firearms and violence: a critical review. National Academies. 2005. URL: https://nap.nationalacademies.org/catalog/10881/firearms-and-violence-a-critical-review [accessed 2023-09-20]
6. Miller M, Azrael D, Hemenway D. Firearm availability and suicide, homicide, and unintentional firearm deaths among women. J Urban Health 2002 Mar;79(1):26-38 [FREE Full text] [doi: 10.1093/jurban/79.1.26] [Medline: 11937613]
7. Stack S. Contributing factors to suicide: political, social, cultural and economic. Prev Med 2021 Nov;152(Pt 1):106498. [doi: 10.1016/j.ypmed.2021.106498] [Medline: 34538366]
8. Ilic I, Zivanovic Macuzic I, Kocic S, Ilic M. Worldwide suicide mortality trends by firearm (1990-2019): a joinpoint regression analysis. PLoS One 2022 May 25;17(5):e0267817 [FREE Full text] [doi: 10.1371/journal.pone.0267817] [Medline: 35613099]

9. Ajdacic-Gross V, Weiss MG, Ring M, Hepp U, Bopp M, Gutzwiller F, et al. Methods of suicide: international suicide patterns derived from the WHO mortality database. Bull World Health Organ 2008 Sep 01;86(9):726-732 [FREE Full text] [doi: 10.2471/blt.07.043489] [Medline: 18797649]
10. Betz ME, Knoepke CE, Simpson S, Siry BJ, Clement A, Saunders T, et al. An interactive web-based lethal means safety decision aid for suicidal adults (Lock to Live): pilot randomized controlled trial. J Med Internet Res 2020 Jan 29;22(1):e16253 [FREE Full text] [doi: 10.2196/16253] [Medline: 32012056]
11. Mann JJ, Michel CA. Prevention of firearm suicide in the United States: what works and what is possible. Am J Psychiatry 2016 Oct 01;173(10):969-979. [doi: 10.1176/appi.ajp.2016.16010069] [Medline: 27444796]
12. National suicide prevention strategies: progress, examples and indicators. World Health Organization. 2018 Nov 30. URL: https://www.who.int/publications/i/item/national-suicide-prevention-strategies-progress-examples-and-indicators [accessed 2023-09-20]
13. Yip PS, Caine E, Yousuf S, Chang S, Wu KC, Chen Y. Means restriction for suicide prevention. Lancet 2012 Jun;379(9834):2393-2399. [doi: 10.1016/s0140-6736(12)60521-2]
14. Clarke R, Lester D. Suicide: Closing the Exits. New York, NY: Springer-Verlag; 1989.
15. Wong YJ, Deng K, Lee CS, Grimes J, Li PFJ. Asian Pacific Islander Americans' and White Americans' suicide methods. Asian Am J Psychol 2018 Dec;9(4):318-326. [doi: 10.1037/aap0000127]
16. Houtsma C, Butterworth SE, Anestis MD. Firearm suicide: pathways to risk and methods of prevention. Curr Opin Psychol 2018 Aug;22:7-11. [doi: 10.1016/j.copsyc.2017.07.002] [Medline: 30122279]
17. Haser G, Yousuf S, Turnock B, Sheehan K. Promoting safe firearm storage in an urban neighborhood: the views of parents concerning the role of health care providers. J Community Health 2020 Apr 23;45(2):338-341. [doi: 10.1007/s10900-019-00748-0] [Medline: 31549353]
18. Liu RT, Kraines MA, Puzia ME, Massing-Schaffer M, Kleiman EM. Sociodemographic predictors of suicide means in a population-based surveillance system: findings from the National Violent Death Reporting system. J Affect Disord 2013 Nov;151(2):449-454. [doi: 10.1016/j.jad.2013.06.023] [Medline: 23845386]
19. Saunders NR, Lee H, Macpherson A, Guan J, Guttmann A. Risk of firearm injuries among children and youth of immigrant families. CMAJ 2017 Mar 27;189(12):E452-E458 [FREE Full text] [doi: 10.1503/cmaj.160850] [Medline: 28385861]
20. Hemenway D, Barber CW, Gallagher SS, Azrael DR. Creating a national violent death reporting system: a successful beginning. Am J Prev Med 2009 Jul;37(1):68-71. [doi: 10.1016/j.amepre.2009.03.005] [Medline: 19524145]
21. Barber C, Azrael D, Hemenway D. A truly national National Violent Death Reporting System. Inj Prev 2013 Aug 25;19(4):225-226. [doi: 10.1136/injuryprev-2013-040817] [Medline: 23710067]
22. Hemenway D. Firearms data, and an ode to data systems. CHANCE 2018 Feb 15;31(1):7-11. [doi: 10.1080/09332480.2018.1438703]
23. About the foreign-born population. United States Census Bureau. URL: https://www.census.gov/topics/population/foreign-born/about.html [accessed 2023-09-20]
24. Conner KR, Huguet N, Caetano R, Giesbrecht N, McFarland BH, Nolte KB, et al. Acute use of alcohol and methods of suicide in a US national sample. Am J Public Health 2014 Jan;104(1):171-178. [doi: 10.2105/ajph.2013.301352]
25. Kaplan MS, Geling O. Sociodemographic and geographic patterns of firearm suicide in the United States, 1989-1993. Health Place 1999 Jun;5(2):179-185. [doi: 10.1016/s1353-8292(99)00007-6] [Medline: 10670999]
26. Price JH, Thompson AJ, Dake JA. Factors associated with state variations in homicide, suicide, and unintentional firearm deaths. J Community Health 2004 Aug;29(4):271-283. [doi: 10.1023/b:johe.0000025326.89365.5c] [Medline: 15186014]
27. Miller M, Azrael D, Hemenway D. Household firearm ownership and suicide rates in the United States. Epidemiology 2002 Sep;13(5):517-524. [doi: 10.1097/00001648-200209000-00006] [Medline: 12192220]
28. Wu KC, Chen Y, Yip PSF. Suicide methods in Asia: implications in suicide prevention. Int J Environ Res Public Health 2012 Apr 28;9(4):1135-1158 [FREE Full text] [doi: 10.3390/ijerph9041135] [Medline: 22690187]
29. Miller M, Warren M, Hemenway D, Azrael D. Firearms and suicide in US cities. Inj Prev 2015 Apr 03;21(e1):e116-e119. [doi: 10.1136/injuryprev-2013-040969] [Medline: 24302479]
30. US Department of Veterans Affairs. National Strategy for Preventing Veteran Suicide: 2018–2028. Washington, DC: US Department of Veterans Affairs; 2018.
31. Bryan CJ, Bryan AO, Anestis MD, Khazem LR, Harris JA, May AM, et al. Firearm availability and storage practices among military personnel who have thought about suicide. JAMA Netw Open 2019 Aug 02;2(8):e199160 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.9160] [Medline: 31418802]
32. Lee JH, Huber JC. Evaluation of multiple imputation with large proportions of missing data: how much is too much? Iran J Public Health 2021 Jul 05;50(7):1372-1380 [FREE Full text] [doi: 10.18502/ijph.v50i7.6626] [Medline: 34568175]
33. NVDRS data access. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/violenceprevention/datasources/nvdrs/dataaccess.html [accessed 2023-09-20]

Abbreviations

AOR: adjusted odds ratio

CCA: complete case analysis
CDC: Centers for Disease Control and Prevention
NVDRS: National Violent Death Reporting System
SAS: Small Arms Survey

Edited by A Mavragani; submitted 10.11.22; peer-reviewed by A Le Jeannic, H Szlyk; comments to author 12.06.23; revised version received 02.07.23; accepted 05.09.23; published 29.09.23.

Please cite as:
Song IH, Lee JH, Shin JS
Firearm Possession Rates in Home Countries and Firearm Suicide Rates Among US- and Foreign-Born Suicide Decedents in the United States: Analysis of Combined Data from the National Violent Death Reporting System and the Small Arms Survey
JMIR Public Health Surveill 2023;9:e44211
URL: https://publichealth.jmir.org/2023/1/e44211
doi:10.2196/44211
PMID:37773604

Original Paper

Systematic Violence Monitoring to Reduce Underreporting and to Better Inform Workplace Violence Prevention Among Health Care Workers: Before-and-After Prospective Study

Giovanni Veronesi[1], PhD; Marco Mario Ferrario[1], MD; Emanuele Maria Giusti[1], PhD; Rossana Borchini[2], MD; Lisa Cimmino[1], MS; Monica Ghelli[3], MS; Alberto Banfi[4], MS; Alessandro Luoni[5], MD; Benedetta Persechino[3], PhD; Cristina Di Tecco[3], PhD; Matteo Ronchetti[3], MS; Francesco Gianfagna[1,6], MD; Sara De Matteis[7], MD; Gianluca Castelnuovo[8,9], PhD; Licia Iacoviello[1,10], MD

[1]Research Center in Epidemiology and Preventive Medicine, Department of Medicine and Surgery, University of Insubria, Varese, Italy
[2]Occupational and Preventive Medicine, Azienda Socio-Sanitaria Territoriale Lariana, Como, Italy
[3]Department of Occupational and Environmental Medicine, Epidemiology and Hygiene, Italian Workers Compensations Authority (INAIL), Roma, Italy
[4]Struttura Complessa Qualità, Risk Management e Accreditamento, Azienda Socio-Sanitaria Territoriale dei Sette Laghi, Varese, Italy
[5]School of Specialization in Occupational Medicine, University of Insubria, Varese, Italy
[6]Mediterranea Cardiocentro, Napoli, Italy
[7]Department of Health Sciences, University of Milan, Milan, Italy
[8]Psychology Research Laboratory, Istituto Auxologico Italiano, Istituto di Ricovero e Cura a Carattere Scientifico, Milano, Italy
[9]Department of Psychology, Catholic University of the Sacred Heart, Milano, Italy
[10]Department of Epidemiology and Prevention, Istituto di Ricovero e Cura a Carattere Scientifico Neuromed, Pozzilli, Italy

Corresponding Author:
Giovanni Veronesi, PhD
Research Center in Epidemiology and Preventive Medicine
Department of Medicine and Surgery
University of Insubria
Via Monte Generoso 71
Varese, 21100
Italy
Phone: 39 0332 217221
Email: giovanni.veronesi@uninsubria.it

Abstract

Background: Monitoring workplace violence (WPV) against health care workers (HCWs) through incident reporting is crucial to drive prevention, but the actual implementation is spotty and experiences underreporting.

Objective: This study aims to introduce a systematic WPV surveillance in 2 public referral hospitals in Italy and assess underreporting, WPV annual rates, and attributes “before” (2016-2020) and “after” its implementation (November 2021 to 2022).

Methods: During 2016-2020, incident reporting was based on procedures and data collection forms that were neither standardized between hospitals nor specific for aggressions. We planned and implemented a standardized WPV surveillance based on (1) an incident report form for immediate and systematic event notification, adopting international standards for violence definitions; (2) second-level root cause analysis with a dedicated psychologist, assessing violence determinants and impacts and offering psychological counseling; (3) a web-based platform for centralized data collection; and (4) periodic training for workforce coordinators and newly hired workers. We used data from incident reports to estimate underreporting, defined as an observed-to-expected (from literature and the “before” period) WPV ratio less than 1, and the 12-month WPV rates (per 100 HCWs) in the “before” and “after” periods. During the latter period, we separately estimated WPV rates for first and recurrent events.

Results: In the “before” period, the yearly observed-to-expected ratios were consistently below 1 and as low as 0.27, suggesting substantial violence underreporting of up to 73%. WPV annual rates declined in 1 hospital (from 1.92 in 2016 to 0.57 in 2020) and rose in the other (from 0.52 to 1.0), with the divergence being attributable to trends in underreporting. Available data were poorly informative to identify at-risk HCW subgroups. In the “after” period, the observed-to-expected ratio rose to 1.14 compared to literature and 1.91 compared to the “before” period, consistently in both hospitals. The 12-month WPV rate was 2.08 (95% CI 1.79-2.42; 1.52 and 2.35 in the 2 hospitals); one-fifth (0.41/2.08, 19.7%) was due to recurrences. Among HCWs, the youngest group (3.79; *P*<.001), nurses (3.19; *P*<.001), and male HCWs (2.62; *P*=.008) reported the highest rates. Emergency departments and psychiatric wards were the 2 areas at increased risk. Physical assaults were more likely in male than female HWCs (45/67, 67.2% vs 62/130, 47.7%; *P*=.01), but the latter experienced more mental health consequences (46/130, 35.4% vs 13/67, 19.4%; *P*=.02). Overall, 40.8% (53/130) of female HWCs recognized sociocultural (eg, linguistic or cultural) barriers as contributing factors for the aggression, and 30.8% (40/130) of WPV against female HCWs involved visitors as perpetrators.

Conclusions: A systematic WPV surveillance reduced underreporting. The identification of high-risk workers and characterization of violence patterns and attributes can better inform priorities and contents of preventive policies. Our evaluation provides useful information for the large-scale implementation of standardized WPV-monitoring programs.

(JMIR Public Health Surveill 2023;9:e47377) doi:10.2196/47377

KEYWORDS

workplace; work; workers; worker; occupational health; safety; report; reporting; incident; abuse; health care workers; HCW; violence; surveillance; underreporting; risk; guidelines; incident report; Italy; prevention; workplace violence; hospital setting; assault; physical assaults; mental health; risk management

Introduction

In 2019, the International Labour Organization (ILO) adopted the Violence and Harassment Convention to “promote and realize the right of everyone to a world of work free from violence and harassment” ([1], art 4). The European Foundation for the Improvement of Living and Working Conditions identifies the health and social work sector as the one with the highest prevalence of adverse social behaviors at work in Europe [2]. The official statistics from the Italian Workers Compensations Authority (INAIL) reported a mean of 2500 workplace violence (WPV) cases per year determining a work injury among health care workers (HCWs) over the 2016-2020 period; of these, 75% were directed toward women [3]. Due to the high prevalence and the consequences on the affected HCWs, WPV against HCW is a public health concern [4,5]. A meta-analysis of the studies published up to 2018—including mostly cross-sectional surveys—estimated for Europe a 12-month prevalence of exposure to nonphysical and physical WPV at 36.6% and 20.1%, respectively [4], with a high degree of heterogeneity across studies. These figures did not decline after the SARS-CoV-2 outbreak [5], exacerbating the impact of the pandemic on the mental health of HCWs [6]. In a more recent systematic review of Italian surveys, the 12-month prevalence of WPV ranged between 11.9% and 93.3% and between 27.5% and 50.3% for verbal and physical violence, respectively [7]. Such a large amount of heterogeneity reflects several methodological concerns, and it challenges the interpretation of results and their usability to guide preventive strategies [8]. First, most studies have been conducted on selected, high-risk wards and thus are poorly representative of the entire population. Second, the negative correlation between survey participation rate and reported violence prevalence [7] suggests the presence of self-selection bias. Finally, the retrospective evaluation of violence and the lack of a clear standardized definition of violence increase the risk of recall bias and measurement error [4,5,7,8].

Complementary to retrospective survey data, active monitoring and surveillance based upon systematic incident reporting is important not only to provide estimates of violence rates but also to investigate the circumstances in which WPV occurs, to identify violence determinants and attributes, and to quantify violence consequences on the assaulted workers [8-10]. Of note, these may also vary according to the assaulted HCW: for instance, evidence from Italy suggests the presence of sex-specific patterns in violence [11]. All this information is crucial to better inform prevention and mitigation strategies by health care organizations [8-11]. However, the implementation of such systems is still scanty and spotty [10]: to the best of our knowledge, standardized and regulatory-based violence monitoring programs implemented at a comprehensive regional- or state-level are present only in Australia [9] and in California [12,13]. The Lombardia Region—the most populated region in Italy with about 10 million inhabitants, located in the north of the country—published in 2019 new guidelines for WPV reporting, risk assessment, and management for regional public and private hospitals, to overcome the recognized WPV underreporting in ongoing registration systems [14]. Stemming from such document, our research group designed the study “Determinants of Violence against the Health care Workers” (*Determinanti Violenze Operatori Sanitari* [DeVOS]) to develop and implement a guideline-based incident report protocol for systematic WPV risk monitoring and management. With this paper, we aim to assess data completeness, WPV underreporting, and rates “before-and-after” the implementation of the new standard in 2 public hospitals (about 9000 employees) in the region. In addition, we report on violence attributes, contributing factors, and consequences and assess the presence of sex-specific patterns in these violence characteristics.

Methods

The DeVOS Study

The DeVOS study started in June 2020 in the only 2 publicly funded, referral hospitals (Azienda Socio-Sanitaria Territoriale [ASST]) serving the provinces of Varese (ASST Sette Laghi, hospital 1) and Como (ASST Lariana, hospital 2). Located in the Lombardia Region, the 2 provinces include about 1.5 million inhabitants, corresponding to 15% of the regional population.

Each ASST includes hub-and-spoke hospitals, hospitals dedicated to mothers and children and rehabilitations, and outpatient clinics. To address the feasibility of our protocol in structures at different underlying WPV risks, we included all the hospitals comprising the ASSTs. The study had 4 main objectives: (1) to quantify WPV before the implementation of the new protocol (2016-2020) and to document the extent by which the available information on WPV episodes was compliant with the guidelines [14] across different hospitals; (2) to design and implement a new protocol for WPV reporting and management (Figure S1 in Multimedia Appendix 1), comprising of a standardized and easy-to-access incident report procedure gathering first information on violence attributes, consequences and contributing factors; of a root cause analysis for more in-depth assessment of violence determinants and impacts; and of a web-based platform for comprehensive event management and data collection; (3) to estimate WPV prevalence in the entire HCW population and in different subpopulations and wards after the implementation of the new standard; and (4) to assess the role of work organizational factors by estimating the association between turn-over, downsizing, sickness leaves, night shift working (defined according to a published method [15]), and WPV occurrence. The specific aims of this paper are related to objectives 1 to 3, whereas the role of organizational factors, as well as the psychological impact of violence, will be addressed in dedicated works. In accordance with regional guidelines [14] and international standards [16], WPV was defined as any form of verbal abuse, threats, physical assaults (to persons or things) and sexual harassment occurring at the workplace and perpetrated by hospital patients or visitors or hospital employees.

WPV Reporting Before the Implementation of the Study Protocol (2016-2020)

During 2016-2020, incident reporting was based upon procedures and data collection forms that were neither standardized between hospitals nor specific for aggressions. The affected HCW notified the violence episode to the risk manager in 1 ASST and to the safety personnel in the other. The incident report data collection forms differed across hospitals and were also used for incidents other than WPV (eg, treatments adverse events, “near miss”), and several versions were adopted during the 5 years following changes in hospital’s organizations and managements. A root cause analysis was not implemented in either hospital.

Development and Implementation of the Project’s WPV Reporting Protocol

Figure S1 in Multimedia Appendix 1 depicts the flow chart for the WPV reporting and management protocol of the DeVOS project. Assaulted HCWs were required to notify the WPV to the risk management office (hospital 1) or to the safety personnel (hospital 2) within 72 hours of occurrence by using a standardized, WPV-specific incident report data collection form, available on the hospitals’ intranet. The form was developed taking into account previous experience by the risk managers of the participating hospitals, regional guidelines [14], and existing literature on the topic [9,10,17,18], and it collected information on violence attributes (form, hospital ward, involved HCW, perpetrator, violence date and time, and environmental factors), consequences (physical, psychological, reactions, feelings during the aggression, and work injury), and contributing factors (sociocultural, structural, organizational, relational, and clinical). More details are in Table S1 in Multimedia Appendix 1. Most fields were not mutually exclusive, and the HCW could fulfill more than 1 choice. In the case of WPV involving multiple HCWs, the risk management office was required to check that each assaulted HCW filled in 1 incident report. HCWs signing the specific consent were further contacted by the psychologist for the root cause analysis, comprising of a structured interview including the Modified Overt Aggression Scale [19] and the Broset Violence Checklist [20]. Then, the HCW received to an email address of his or her choice a 1-time access link to the project web platform for a safe and easy completion of the questionnaires assessing the impact of the WPV event (presence of cognitive, emotional, and somatic symptoms [21]; the Maslach Burnout Inventory [22]; and the General Health Questionnaire-12 [23]) and psychosocial work conditions (the Italian version of the Health and Safety Executive’s Management Standards Indicator Tool [24]). These questionnaires were validated for use with Italian HCWs [21,24,25]. Furthermore, the psychologist contacted the HCW supervisor for a guided interview using a checklist on work content and context factors. During periodic training meetings with workforce coordinators, the risk management officers explained the new protocol and provided information on violence notification, field definitions on the incident report form to standardize data collection (including contributing factors), and the rationale and motivation for participation in the root cause analysis. The coordinators were then asked to instruct the HCW of their ward or unit. Additionally, in hospital 1, all newly hired HCWs were targeted by the same training. In both ASSTs, the new protocol became effective on November 1, 2021. One ASST (hospital 2) gradually introduced the protocol to the outpatient departments during the year 2022. Therefore, for that ASST, we included only the hub and the main spoke hospitals in the current analyses, corresponding to 71% of the workforce.

Web-Based Platform for Data Collection and Risk Management

As part of the study, we developed a new web-based platform for centralized data collection, with customized access to data visualization or modification for risk management office personnel, the psychologist, and the data analyst. Due to ethical and confidentiality issues, we collected minimal sensitive information on the HCW and the aggressor. Furthermore, the easy visualization of questionnaire results by the psychologist allowed an immediate identification of high-risk workers for referral to supportive or mitigation resources.

Ethical Considerations

The study received approval by the Ethical Committee of Insubria (IDs 82/2021 and 90/2021). Participants signed an informed consent on data confidentiality and protection and a separate informed consent to participate in the interview with the psychologist. The web-based platform for centralized data collection complied with the European Union legislation on data protection. The platform applied a complete anonymization of personal data, allowing the tracing of repeated events on the same worker. HCWs received no compensation for their participation in the study.

Statistical Analysis

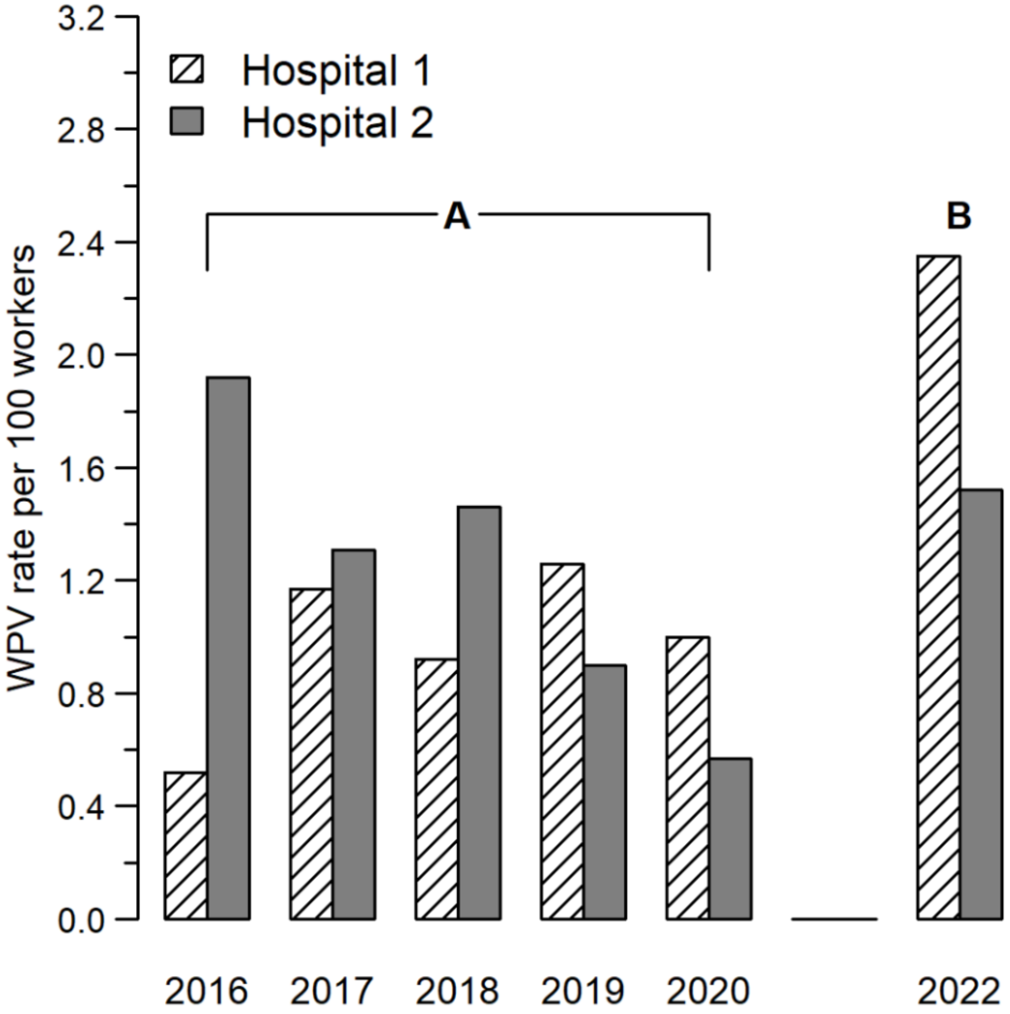
To be consistent with the literature mostly reporting WPV prevalence for a 12-month interval, for each year in the "before" study period (2016-2020) and by hospital, we estimated the 12-month WPV rate (per 100 workers) as the ratio between the number of reported WPV in the year and the number of HCWs in the payroll administrative records at the beginning of each year. The 95% CI for the yearly rate was estimated using the exact binomial distribution [26]. We estimated underreporting as the ratio between the observed and expected number of violent incidents. For the latter, we used the WPV rate estimated in 2015-2017 in a large public University hospital in Northern Italy by Viottini and colleagues [27] as 210 observed incidents each year over a population of 10,970 HCWs, along with the number of HCWs, again from payroll administrative records. The 95% CI for the ratio was estimated from the Poisson exact method [26]. To date, we completed the first 14 months since the implementation of the new reporting standard (November 1, 2021 to December 31, 2022). For the entire period, we reported on data completeness as missing data prevalence on the incident report data collection forms. To describe the violence attributes, we first assessed the co-occurrence of different types of violence in the same episode, by reporting percentages and mean number of violence types by episode. Then, we reported the distribution of major violence characteristics, consequences, and contributing factors, in the overall sample and by sex of the assaulted HCWs, to identify the presence of sex-related patterns. The association between violence characteristics and sex was formally tested through chi-square tests. The analyses of WPV underreporting and rate were restricted to a 12-month period (year 2022, n=166 WPV) adopting the same methods described above for the "before" period and using the number of HCWs in the payroll records at the beginning of the implementation of the new standard. The overall rate was then broken down as the sum of 2 mutually exclusive components: 1 being the rate of the first violent episode and the second being the rate of recurrent violent episodes on the same worker. Statistical analyses were performed using SAS software (version 9.4; SAS Institute Inc).

Results

Data Completeness and WPV Rate in the "Before" Period

Between 2016 and 2020, a total of 400 WPV reports were notified in the 2 study hospitals. Data completeness was scanty: in 160 (40%) of the reports, the affected HCW was not identified (eg, replaced by a "generic" description of job title or by expressions such as "everyone present"), and when identifiable, in 70.8% (170/240) of reports, there was at least 1 missing information on age, sex, or job title. This prevented us from providing specific prevalence estimates by subgroups. Similarly, environmental details, consequences, and contributing factors were not available in 59% (236/400), 18% (72/400), and 50% (200/400) of reports. Yearly WPV rates were diverging between the 2 study hospitals, increasing from 0.52 (95% CI 0.34-0.77) per 100 HCWs to 1.0 (95% CI 0.75-1.31) in 1 hospital and declining from 1.92 (95% CI 1.42-2.54) to 0.57 (95% CI 0.32-0.93) in the other (Figure 1A). The yearly observed to expected ratio was significantly below 1 for most years, suggesting substantial (up to 73%) underreporting. In particular, the ratio from 2016 to 2020 changed from 0.27 to 0.52 (eg, reduction in underreporting) in the hospital with an increasing WPV rate and from 1.0 to 0.30 (eg, increase in underreporting) in the other.

Figure 1. Twelve-month WPV rate observed in the study hospitals, before (period A) and after (period B) the implementation of the study protocol on incident report. The study protocol on incident report was effective since November 1, 2021. The WPV rate (per 100 health care workers) is the number of observed WPV incidents divided by the number of health care workers at the beginning of each year. WPV: workplace violence.



Data Completeness and WPV Rate in the "After" Period

During the first 14 months of implementation of the study standard, a total number of 205 WPV were reported, half (50%) of them on the same day of occurrence, and 75% within the first 72 hours. Data completeness was optimal: only in 2% (4/205) of reports the affected HCW was not identified, and when identifiable, only in another 2% (4/201) there was at least 1 missing information on demographics or job title information (Table S1 in Multimedia Appendix 1). The observed to expected ratio was 1.91 (95% CI 1.63-2.21) as compared to the "before" period in the same hospitals, which was consistent in the 2 hospitals (1.96 and 1.77). When using the reference rate from literature to estimate the expected number of WPVs, the observed to expected ratio was 1.14 (95% CI 0.98-1.32), being slightly lower in hospital 2 where the first notification was to safety personnel (0.83, 95% CI 0.60-1.13) rather than to the risk management office (1.28, 95% CI 1.07-1.52) as in hospital 1. The 12-month WPV rate was 2.08 (95% CI 1.79-2.42) per 100 HCWs (Table 1), being slightly lower in hospital 2 than in hospital 1 (1.52 vs 2.35; *P*=.02; Figure 1B). WPV rate (Table 1) declined with age, with the youngest group being at the highest risk (*P*<.001); it was higher in male than female HCWs (2.82 vs 1.83; *P*=.008); and it was more than doubled in nurses (3.49; 95% CI 2.93-4.15), as compared to nurse assistants (1.74) and physicians (1.08; *P*<.001). Finally, psychiatric wards (14.3) and emergency departments (13.3) were at higher WPV rates than the remaining wards; about one-fifth (0.41/2.08, 19.7%) of the rates are due to recurrences. Analyses by hospital (Table S2 in Multimedia Appendix 1) suggested higher rates in the youngest group in hospital 1 and in psychiatric wards in hospital 2, whereas associations by sex and job title were homogeneous across the hospitals.

Table 1. Twelve-month WPV[a] rate (with 95% CI) for the year 2022 by affected HCW[b] demographic and work characteristics[c].

Characteristics	HCWs (n=7982), n (%)	WPV (n=166), n (%)[d]	12-month WPV rate, (95% CI)			
			First WPV[e]	Recurrent WPV[f]	Overall WPV[g]	*P* value[h]
All WPV	7982 (100)	166 (100)	1.67 (1.41-1.97)	0.41 (0.29-0.58)	2.08 (1.79-2.42)	—[i]
Age group (years)						<.001
<30	829 (10.4)	33 (19.9)	2.90 (1.95-4.28)	1.09 (0.57-2.07)	3.98 (2.85-5.54)	
30-50	3652 (45.8)	88 (53)	1.86 (1.47-2.36)	0.55 (0.35-0.85)	2.41 (1.96-2.96)	
>50	3501 (43.9)	45 (27.1)	1.17 (0.86-1.59)	0.11 (0.04-0.30)	1.29 (0.96-1.72)	
Sex						.008
Female	5960 (74.7)	109 (65.7)	1.51 (1.23-1.85)	0.32 (0.20-0.50)	1.83 (1.52-2.20)	
Male	2022 (25.3)	57 (34.3)	2.13 (1.58-2.86)	0.69 (0.41-1.17)	2.82 (2.18-3.64)	
Job title						<.001
Physician	1296 (16.2)	14 (8.4)	0.93 (0.53-1.62)	0.15 (0.04-0.61)	1.08 (0.64-1.82)	
Nurse	3498 (43.8)	122 (73.5)	2.69 (2.20-3.28)	0.80 (0.55-1.12)	3.49 (2.93-4.15)	
Nurse assistant	1149 (14.4)	20 (12.0)	1.48 (0.92-2.37)	0.26 (0.08-0.81)	1.74 (1.13-2.68)	
Health care technician	833 (10.4)	2 (1.21)	0.24 (0.06-0.95)	—	0.24 (0.06-0.95)	
Administrative clerk	824 (10.3)	1 (0.60)	0.12 (0.002-0.86)	—	0.12 (0.002-0.86)	
Other	382 (4.8)	7 (4.2)	1.83 (0.88-3.79)	—	1.83 (0.88-3.79)	
Hospital ward						<.001
Psychiatry and mental health departments	322 (4)	46 (27.7)	10.6 (7.64-14.4)	3.73 (2.13-6.44)	14.3 (10.9-18.6)	
Emergency department	592 (7.4)	79 (47.6)	10.1 (7.95-12.8)	3.21 (2.06-4.98)	13.3 (10.8-16.3)	
Other wards	7068 (88.5)	41 (24.7)	0.55 (0.40-0.75)	0.03 (0.001-0.11)	0.58 (0.43-0.79)	

[a] WPV: workplace violence.

[b] HCW: health care worker.

[c] Twelve-month rate is the ratio between the number of WPV and the number of HCWs, per 100 HCWs.

[d] Excluding 8 WPVs without data on age, sex, and job title of the affected HCW.

[e] First WPV: number of first WPVs in the 12-month period.

[f] Recurrent WPV: number of WPVs from the second case onward, in the 12-month period.

[g] The overall violence rate is the sum of first and recurrent violence rates.

[h] Wald chi-square test on the overall violence reporting rate, with a number of degrees of freedom equal to the number of classes in the independent variable – 1.

[i] Not applicable (no groups to be compared).

Violence Attributes and Patterns in the “After” Period

Table 2 reports the co-occurrence of violence forms in the incident reports. Verbal abuse was the largest majority (186/205, 90.7%) Of these, verbal abuse was the only form of violence in 26 (14%) out of 186 incidents, whereas 134 (72%) also reported threats, 98 (52.7%) reported physical assaults, and 15 (8%) reported sexual harassment. The mean number of reported violence forms was 2.33. We report similar figures for the other violence forms. Table 3 reports the distribution of major violence characteristics and risk factors, by sex of the assaulted HCW. Male HCWs were more likely than female HCWs to be subject to physical assault (45/67, 67.2% vs 62/130, 47.7%; *P*=.01). Either alone or with a patient, visitors were more likely to assault female than male HCWs (40/130, 30.7% vs 8/67, 12%; *P*=.02). No sex-related differences were observed on violence time or location. Female HCWs were less likely than male HCWs to report a work injury to the insurance authority (4/130, 3.1% vs 15/67, 22.4%; *P*<.001) and were more likely than men to report psychological consequences (46/130, 35.4% vs 13/67, 19.4%; *P*=.02) and life-threatening feelings during the aggression (23/130, 17.7% vs 5/67, 7.5%; *P*=.05). Finally, 47.8% (32/67) of WPV in male HCWs was related to the clinical conditions of the perpetrator, which is greater than that in female HCWs (40/130, 30.8%; *P*=.02). In female HCWs, the single most frequently reported factor was the sociocultural one (eg, related to linguistic or cultural barriers or behaviors; 53/130, 40.8%). Of note, structural factors including safety (the lack of barriers, alarm systems, or escape or poor lightning) were among the least prevalent identified contributing factors, both in men and in women.

Table 2. Co-occurrence of different types of violence observed in the "after" period.

Type of violence (number of reports)	Co-occurrence, n (%)[a]				No co-occurrence, n (%)[b]	Number of violence types in incident reports, mean
	VA[c]	T[d]	PA[e]	SH[f]		
VA (n=186)	—[g]	134 (72)	98 (52.7)	15 (8)	26 (14)	2.33
T (n=139)	134 (96.4)	—	77 (55.4)	12 (8.6)	2 (1.4)	2.60
PA (n=115)	98 (85.3)	77(67)	—	10 (8.7)	14 (12.2)	2.61
SH (n=15)	15 (100)	12 (80)	10 (66.7)	—	0 (0)	3.46

[a] Percent of incident reports with the row type of violence that also reports the column type.

[b] Percent of incident reports with only the row type of violence.

[c] VA: verbal abuse.

[d] T: threat.

[e] PA: physical assault.

[f] SH sexual harassment.

[g] Not relevant.

Table 3. Distribution of major violence attributes, consequences, and contributing factors, by sex of the assaulted HCW[a], in the "after" period.

Attributes	All WPV[b] (n=197), n (%)	WPV against female HCWs (n=130), n (%)	WPV against male HCWs (n=67), n (%)	*P* values
Type of violence[c]				
Verbal abuse	179 (90.7)	120 (92.3)	59 (88.1)	.33
Threat	132 (67.0)	82 (63.1)	50 (74.6)	.10
Physical assault	107 (54.3)	62 (47.7)	45 (67.2)	.01
Sexual harassment	14 (7.1)	8 (6.2)	6 (9.0)	.47
Perpetrator				.02
Patient only	138 (70.1)	82 (63.1)	56 (83.6)	
Visitor only	38 (19.3)	31 (23.9)	7 (10.5)	
Patient and visitor	10 (5.1)	9 (6.8)	1 (1.5)	
Coworker	11 (5.6)	8 (6.9)	3 (4.5)	
Violence during a night shift	68 (34.7)	43 (33.3)	25 (37.3)	.58
When during the work shift				.73
Beginning	25 (12.8)	16 (12.4)	9 (13.4)	
During	162 (82.7)	106 (81.2)	56 (83.6)	
End	9 (4.6)	7 (5.4)	2 (3.0)	
Location[c]				
Patient's bedroom	45 (22.8)	29 (22.3)	16 (23.9)	.80
Waiting examination room	108 (54.8)	71 (54.6)	37 (55.2)	.94
Communal location[d]	54 (27.4)	35 (26.9)	19 (28.4)	.83
External areas	20 (10.2)	15 (11.5)	5 (7.5)	.37
Other[e]	6 (3.1)	5 (3.9)	1 (1.5)	.36
Consequences[c]				
Physical	39 (19.8)	22 (16.9)	17 (25.4)	.16
Psychological	59 (30.0)	46 (35.4)	13 (19.4)	.02
Life-threatening feeling	28 (14.2)	23 (17.7)	5 (7.5)	.05
Work injury report[f]	19 (9.6)	4 (3.1)	15 (22.4)	<.001
Contributing factors[c,g]				
Sociocultural	80 (40.6)	53 (40.8)	27 (40.3)	.95
Structural	30 (15.2)	18 (13.9)	12 (17.9)	.45
Organizational	60 (30.5)	41 (31.5)	19 (28.4)	.65
Relational	55 (27.9)	37 (28.5)	18 (26.9)	.81
Clinical	72 (36.6)	40 (30.8)	32 (47.8)	.02
Other	18 (9.1)	15 (11.5)	3 (4.5)	.10
Not identified	28 (14.2)	18 (13.9)	10 (14.9)	.84

[a] HCW: health care worker.

[b] WPV: workplace violence.

[c] More than 1 answer was possible, so the total does not sum up to 100%.

[d] For example, corridors, stairs, and elevators.

[e] Including web-based violence via email or telephone.

[f] Violence determining a work injury must be reported to the Italian Workers Compensations Authority.

[g] Sociocultural (eg, linguistic barriers, behavior conditioned by cultural elements such as education or country of origin, the inadequacy of social

behaviors); structural (eg, lack of barriers/alarm systems/escape, poor lightning); organizational (eg, related to work organization); relational (eg, related to HCW experience and communication abilities); and clinical (eg, related to the conditions of the perpetrator, including psychiatric disorders and substance abuse).

Discussion

The implementation of the ILO convention [1] and its subsequent ratification in state members, including Italy (Law n. 4 of January 15, 2021), calls for a systematic and standardized registration of WPV in HCWs implemented on a large scale to inform prevention, protection, and enforcement actions [7]. In Italy so far, national statistics cover only more severe cases, such as those exiting in a work injury [3], whereas academic data obtained from convenient samples and retrospective investigations exhibit significant heterogeneity, rendering them uninformative for guiding prevention efforts effectively [7]. This study documents underreporting and WPV prevalence estimates in 2 large public referral hospitals before and after the implementation of a standardized, guidelines-based program for systematic violence monitoring. Similar experiences [9,12,13] have not yet documented such temporal changes. WPV data in HCWs are affected by a substantial amount of underreporting [28,29], up to 80% to 90% according to 1 US study [29]. In line with this previous knowledge, in the 2016-2020 period, we estimated a yearly observed to expected ratio mostly below 1 in the 2 study hospitals and as low as 0.27, corresponding to 73% underreporting toward expected prevalence from the literature [27]. Such a large amount of underreporting hampers a meaningful interpretation and comparison of time trends in violence prevalence between the structures. After the implementation of the new standard, the 12-month observed to expected ratio increased to be larger than 1 as compared to the same benchmark and almost 2 as compared to the "before" period in the same hospitals. In addition, only 9.8% (20/205) of the reported WPV determined a work injury, suggesting that we were fairly able to detect less serious incidents in which "no one was hurt," which are generally overlooked [30]. Finally, although verbal abuse was present in 90.7% (186/205) of reports, the affected HCWs were able to describe and report the complex co-occurrence of violence forms in the same episode. However, in the study hospital in which the first notification was to safety personnel, rather than to the risk management office, we found a lower observed to expected ratio, a lower 12-month overall WPV rate but with a higher peak in psychiatric wards, where episodes are more likely to be related to the psychiatric conditions of the perpetrator. Based on these findings, notification to dedicated personnel is recommended in future applications to enhance violence reporting.

Several reasons have been advocated for WPV underreporting, including the feeling that violence is "part of the job" [31], the lack of supervisor or coworker support, the fear of blame, and the belief that reporting would not lead to positive changes [29]. We may speculate that our protocol could have contrasted some of these barriers, through the introduction of a new reporting standard specifically dedicated to WPV; the periodic training of the new procedure with the workforce and newly hired workers; the easiness of reporting by the HCW along with a close support contact with the risk management office during the first 72 hours since the episode; the availability of a dedicated psychologist for counseling, also outside the HCW working hours; and the periodic sharing of reporting with workers' safety representatives for health. These aspects, specifically related to our protocol, might have reinforced the positive impact on violence reporting related to increased awareness of the issue due to the perception that the institution was "giving attention to it." Taken together, these put into light the crucial role of a participative approach to the management of WPV that, starting from standardized and systematic data collection, requires the inclusion and active involvement of the organization, workers, and occupational health figures (risk management, occupational physicians, psychologist, and workers' safety representatives).

The 12-month violence rates we estimated should be compared with caution with those from survey studies, since ours are referring to the entire HCW population and are based on incident reporting rather than on recollection. We add important pieces of information that can help to identify high-risk workers. First, while confirming the highest violence rates for psychiatric wards, emergency departments, and nurses, we add the notion that about 25% of their rate is due to the reoccurrence of aggressions on the same workers, as previously observed in Italy [27,32], Norway [33], and Denmark [34]. As current knowledge on the factors related to multiple aggressions is so far limited [33,34], future analyses are required to elucidate the role of personal and work-related characteristics on multiple aggressions. Second, in identifying high-risk subpopulations, it is important to consider that age, sex, and job title might have an unbalanced distribution in the specific health care workforce. Although in our data the large majority (66%, 130/167) of incident reports involved female HCWs, confirming institutional data [3], these do not appear to be at higher WPV risk as compared to male HCWs once their disproportional number in the workforce has been considered, as in other prospective studies [33,35]. Finally, our data can be used to well-characterize patterns in WPV attributes and consequences related to HCW's characteristics, of which the sex-related one is an exemplification. To this extent, we confirmed the highest prevalence of physical violence in male HCWs recently observed in Italy [11], at the same time expanding knowledge by reporting sex-specific patterns in consequences, contributing factors, and conditions of the perpetrator. Taken together, this information can both guide priorities of interventions and a better design of the contents of primary, secondary, and tertiary prevention policies [36,37]. Systematic incident reporting is important to provide a "time zero" to evaluate the efficacy of future interventions.

Study limitations include the short time period of implementation of the new standard; since it is still in use by the 2 hospitals, we will be able to monitor in future reports its sustainability over time. Our study cannot document changes in WPV prevalence before and after the SARS-CoV-2 pandemic outbreak, given that the change in the reporting system became effective only at the end of 2021. Our underreporting metric is based upon a violence rate derived from literature, but assuming it remained constant during the study periods (2016-2020 and 2022). On the other hand, due to the substantial heterogeneity documented in the introduction, the validity of the metric depends upon the choice of a comparable benchmark for study design, period, and setting, since violence prevalence in Italy is also heterogeneous across public versus private hospitals [32] and country area [38]. The study we chose [27] satisfies most of these requirements. In addition, we used the annual count of HCWs to mitigate the impact on the expected number of incidents due to changes in the workforce. Then, our WPV rates can provide an updated benchmark for assessing underreporting in Italy, using a similar observed-to-expected metric. Our study protocol was implemented and tested in 2 large public general hospitals, each with a complex organizational structure including hub-and-spoke hospitals, as well as rehabilitation and outpatient clinics. We adopted recognized standards for violence definition, attributes, consequences, and determinants [10], to enhance the comparison of our findings with other studies.

In conclusion, a guidelines-based protocol can mitigate the underreporting of violence episodes against HCWs and provide accurate information to identify high-risk workers and describe violence attributes and patterns. The standardization across hospitals can better inform priorities and contents of preventive policies, at both a local and a large scale. To this extent, our evaluation can provide useful information for large-scale implementation of guidelines-based monitoring programs, as well as in other contexts.

Acknowledgments

The authors wish to thank all the health care workers who have participated in the study. The authors are grateful to Davide Trevisan, for the development of the web-based data platform for data collection; to Giacomo Radici and Patrick Tayoun, who helped with bibliographic research; to Eleonora Lusito, for her contribution to data handling and analysis; and to Simona Boneschi, who contributed to the selection of interview or questionnaires to be included in the root cause analysis. The authors also thank the management of the participating organizations, as well as people contributing with their work in the quality and safety and the occupational medicine units, namely Giovanna Martinelli, Roberto Guarnaschelli, Alessandra Tropia, Angela Trentin, Sara Colombo, Roberta Somaschini, Gloria Simonelli, Francesca Cereda, Francesca Ricci, Mariavalentina Giordano, Cristina Curioni, and Andrea Frontini. Preliminary results were presented at the EPICOH 2023 conference, and the abstract was published in a dedicated supplementary issue of Occupational and Environmental Medicine. This study was supported by a grant provided by the Italian Workers Compensations Authority (INAIL; Bando BRIC 2019 - ID29A).

Data Availability

The data sets generated and analyzed during this study are not publicly available due to restrictions imposed by the informed consent obtained for this study and by the legislation of the European Union, but they can be made available from the corresponding author on reasonable request. Approval by each management of the Azienda Socio-Sanitaria Territoriale that participated in the study will be required for the release of the data.

Authors' Contributions

GV and MMF contributed to the conceptualization and design of the study, drafted the original paper, and are the principal investigators of the *Determinanti Violenze Operatori Sanitari* (DeVOS) Project. GV is the study guarantor. GV and EMG were responsible for statistical analyses. RB, LC, AB, and AL contributed to data acquisition and paper editing. EMG, FG, and SDM critically contributed to the interpretation of results, and to review and editing of the paper. MG, BP, CDT, MR, GC, and LI critically reviewed the paper.

Conflicts of Interest

None declared

Multimedia Appendix 1

Study flow chart, data fields in the report form and prevalence of missing data, and workplace violence rates by study hospital.

[DOCX File , 152 KB - publichealth_v9i1e47377_app1.docx]

References

1. C190—Violence and harassment convention, 2019 (No. 190). International Labour Organization. 2019. URL: https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:C190 [accessed 2023-01-04]
2. Violence and harassment in European workplaces: causes, impacts and policies. Eurofound. Dublin; 2015. URL: https://zenodo.org/records/16387 [accessed 2023-01-04]
3. Rossi P, Mele A, Punziano A. Gli episodi di violenza nei confronti degli esercenti le professioni sanitarie Inail. Inail. 2022. URL: https://www.inail.it/cs/internet/docs/alg-factsheet-violenza-professioni-sanitarie-inail-2022_6443174670061.pdf [accessed 2023-01-04]

4. Liu J, Gan Y, Jiang H, Li L, Dwyer R, Lu K, et al. Prevalence of workplace violence against healthcare workers: a systematic review and meta-analysis. Occup Environ Med 2019;76(12):927-937. [doi: 10.1136/oemed-2019-105849] [Medline: 31611310]
5. Ramzi ZS, Fatah PW, Dalvandi A. Prevalence of workplace violence against healthcare workers during the COVID-19 pandemic: a systematic review and meta-analysis. Front Psychol 2022;13:896156 [FREE Full text] [doi: 10.3389/fpsyg.2022.896156] [Medline: 35712196]
6. Giusti EM, Veronesi G, Callegari C, Borchini R, Castelnuovo G, Gianfagna F, et al. Pre-pandemic burnout and its changes during the COVID-19 outbreak as predictors of mental health of healthcare workers: a lesson to be learned. Psychiatry Res 2023;326:115305 [FREE Full text] [doi: 10.1016/j.psychres.2023.115305] [Medline: 37331071]
7. Civilotti C, Berlanda S, Iozzino L. Hospital-based healthcare workers victims of workplace violence in Italy: a scoping review. Int J Environ Res Public Health 2021;18(11):5860 [FREE Full text] [doi: 10.3390/ijerph18115860] [Medline: 34072551]
8. Phillips JP. Workplace violence against health care workers in the United States. N Engl J Med 2016;375(7):e14 [FREE Full text] [doi: 10.1056/NEJMc1606816] [Medline: 27532859]
9. Morphet J, Griffiths D, Innes K. The trouble with reporting and utilization of workplace violence data in health care. J Nurs Manag 2019;27(3):592-598. [doi: 10.1111/jonm.12717] [Medline: 30223311]
10. Arnetz JE, Hamblin L, Ager J, Aranyos D, Essenmacher L, Upfal MJ, et al. Using database reports to reduce workplace violence: perceptions of hospital stakeholders. Work 2015;51(1):51-59. [doi: 10.3233/wor-141887]
11. Acquadro Maran D, Cortese CG, Pavanelli P, Fornero G, Gianino MM. Gender differences in reporting workplace violence: a qualitative analysis of administrative records of violent episodes experienced by healthcare workers in a large public Italian hospital. BMJ Open 2019;9(11):e031546 [FREE Full text] [doi: 10.1136/bmjopen-2019-031546] [Medline: 31712339]
12. Gooch PP. Hospital workplace violence prevention in California: new regulations. Workplace Health Saf 2018;66(3):115-119 [FREE Full text] [doi: 10.1177/2165079917731791] [Medline: 28980509]
13. Odes R, Chapman S, Ackerman S, Harrison R, Hong O. Differences in hospitals' workplace violence incident reporting practices: a mixed methods study. Policy Polit Nurs Pract 2022;23(2):98-108 [FREE Full text] [doi: 10.1177/15271544221088248] [Medline: 35317690]
14. Giunta Regione Lombardia. Allegato C. Deliberazione XI/1986 del 23/07/2019. URL: https://tinyurl.com/36e9h9jx [accessed 2023-02-06]
15. Ferrario MM, Bertù L, Cimmino L, Conti M, Borsani A, Veronesi G. A multi-dimensional evaluation of job stress integrating organizational and subjective risk indicators: one application to health care workers in Northern Italy. G Ital Med Lav Ergon 2011;33(3 Suppl):315-318. [Medline: 23393865]
16. National Institute of Occupational Safety and Health. Violence occupational hazards in hospitals. U.S. Department of Health and Human Services: 2002-2101. 2002. URL: https://www.cdc.gov/niosh/docs/2002-101/ [accessed 2023-10-27]
17. Pompeii L, Dement J, Schoenfisch A, Lavery A, Souder M, Smith C, et al. Perpetrator, worker and workplace characteristics associated with patient and visitor perpetrated violence (Type II) on hospital workers: a review of the literature and existing occupational injury data. J Safety Res 2013;44:57-64. [doi: 10.1016/j.jsr.2012.09.004] [Medline: 23398706]
18. Arnetz JE, Hamblin L, Essenmacher L, Upfal MJ, Ager J, Luborsky M. Understanding patient-to-worker violence in hospitals: a qualitative analysis of documented incident reports. J Adv Nurs 2015;71(2):338-348 [FREE Full text] [doi: 10.1111/jan.12494] [Medline: 25091833]
19. Margari F, Matarazzo R, Casacchia M, Roncone R, Dieci M, Safran S, et al. Italian validation of MOAS and NOSIE: a useful package for psychiatric assessment and monitoring of aggressive behaviours. Int J Methods Psychiatr Res 2005;14(2):109-118 [FREE Full text] [doi: 10.1002/mpr.22] [Medline: 16175880]
20. Woods P, Almvik R. The Brøset Violence Checklist (BVC). Acta Psychiatr Scand Suppl 2002(412):103-105. [doi: 10.1034/j.1600-0447.106.s412.22.x] [Medline: 12072138]
21. Sanza M, Ceccarelli P, Ballanti L, Bruno A, Iorio G, Loddo C. La valutazione dell’impatto delle aggressioni sul benessere psicologico del personale sanitario. uno studio osservazionale. Politiche Sanitarie 2018;19(1):43-50 [FREE Full text]
22. Maslach C, Jackson SE. The measurement of experienced burnout. J Organ Behav 1981;2(2):99-113 [FREE Full text] [doi: 10.1002/job.4030020205]
23. Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. Psychol Med 1979;9(1):139-145. [doi: 10.1017/s0033291700021644] [Medline: 424481]
24. Veronesi G, Giusti EM, D'Amato A, Gianfagna F, Borchini R, Castelnuovo G, et al. The North Italian Longitudinal study assessing the mental health effects of SARS-CoV-2 pandemic on health care workers-part I: study design and psychometric structural validity of the HSE indicator tool and work satisfaction scale. Int J Environ Res Public Health 2022;19(15):9514 [FREE Full text] [doi: 10.3390/ijerph19159514] [Medline: 35954871]
25. Giusti EM, Veronesi G, Callegari C, Castelnuovo G, Iacoviello L, Ferrario MM. The North Italian Longitudinal study assessing the mental health effects of SARS-CoV-2 pandemic health care workers-part II: structural validity of scales assessing mental health. Int J Environ Res Public Health 2022;19(15):9541 [FREE Full text] [doi: 10.3390/ijerph19159541] [Medline: 35954915]

26. Guidelines for using confidence intervals for public health assessment. Washington State Department of Health. 2012. URL: https://doh.wa.gov/sites/default/files/legacy/Documents/1500//ConfIntGuide.pdf [accessed 2023-01-04]
27. Viottini E, Politano G, Fornero G, Pavanelli PL, Borelli P, Bonaudo M, et al. Determinants of aggression against all health care workers in a large-sized university hospital. BMC Health Serv Res 2020;20(1):215 [FREE Full text] [doi: 10.1186/s12913-020-05084-x] [Medline: 32178674]
28. Campbell CL, Burg MA, Gammonley D. Measures for incident reporting of patient violence and aggression towards healthcare providers: a systematic review. Aggress Violent Behav 2015;25:314-322. [doi: 10.1016/j.avb.2015.09.014]
29. Arnetz JE, Hamblin L, Ager J, Luborsky M, Upfal MJ, Russell J, et al. Underreporting of workplace violence: comparison of self-report and actual documentation of hospital incidents. Workplace Health Saf 2015;63(5):200-210 [FREE Full text] [doi: 10.1177/2165079915574684] [Medline: 26002854]
30. Copeland D, Henry M. Workplace violence and perceptions of safety among emergency department staff members: experiences, expectations, tolerance, reporting, and recommendations. J Trauma Nurs 2017;24(2):65-77 [FREE Full text] [doi: 10.1097/JTN.0000000000000269] [Medline: 28272178]
31. Dubb SS. It doesn't "come with the job": violence against doctors at work must stop. BMJ 2015;350:h2780. [doi: 10.1136/bmj.h2780] [Medline: 26013379]
32. Firenze A, Santangelo OE, Gianfredi V, Alagna E, Cedrone F, Provenzano S, et al. Violence on doctors. an observational study in Northern Italy. Med Lav 2020;111(1):46-53 [FREE Full text] [doi: 10.23749/mdl.v111i1.8795] [Medline: 32096772]
33. Nøland ST, Taipale H, Mahmood JI, Tyssen R. Analysis of career stage, gender, and personality and workplace violence in a 20-year nationwide cohort of physicians in Norway. JAMA Netw Open 2021;4(6):e2114749 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.14749] [Medline: 34181010]
34. Hogh A, Sharipova M, Borg V. Incidence and recurrent work-related violence towards healthcare workers and subsequent health effects. a one-year follow-up study. Scand J Public Health 2008;36(7):706-712. [doi: 10.1177/1403494808096181] [Medline: 18775833]
35. Johansen IH, Baste V, Rosta J, Aasland OG, Morken T. Changes in prevalence of workplace violence against doctors in all medical specialties in Norway between 1993 and 2014: a repeated cross-sectional survey. BMJ Open 2017;7(8):e017757 [FREE Full text] [doi: 10.1136/bmjopen-2017-017757] [Medline: 28801441]
36. Kumari A, Sarkar S, Ranjan P, Chopra S, Kaur T, Baitha U, et al. Interventions for workplace violence against health-care professionals: a systematic review. Work 2022;73(2):415-427 [FREE Full text] [doi: 10.3233/WOR-210046] [Medline: 35431213]
37. Fricke J, Siddique SM, Douma C, Ladak A, Burchill CN, Greysen R, et al. Workplace violence in healthcare settings: a scoping review of guidelines and systematic reviews. Trauma Violence Abuse 2023;24(5):3363-3383. [doi: 10.1177/15248380221126476] [Medline: 36341578]
38. La Torre G, Firenze A, Di Gioia LP, Perri G, Soncin M, Cremonesi D, et al. Workplace violence among healthcare workers, a multicenter study in Italy. Public Health 2022;208:9-13. [doi: 10.1016/j.puhe.2022.04.008] [Medline: 35660281]

Abbreviations

ASST: Azienda Socio-Sanitaria Territoriale
DeVOS: Determinanti Violenze Operatori Sanitari (Determinants of Violence against the Health care Workers)
HCW: health care worker
ILO: International Labour Organization
INAIL: Italian Workers Compensations Authority
WPV: workplace violence

Edited by A Mavragani, T Sanchez; submitted 17.03.23; peer-reviewed by R Odes, N Berti, J Vilchis Gil, A Mannocci; comments to author 07.07.23; revised version received 14.07.23; accepted 26.09.23; published 13.11.23.

Please cite as:
Veronesi G, Ferrario MM, Giusti EM, Borchini R, Cimmino L, Ghelli M, Banfi A, Luoni A, Persechino B, Di Tecco C, Ronchetti M, Gianfagna F, De Matteis S, Castelnuovo G, Iacoviello L
Systematic Violence Monitoring to Reduce Underreporting and to Better Inform Workplace Violence Prevention Among Health Care Workers: Before-and-After Prospective Study
JMIR Public Health Surveill 2023;9:e47377
URL: https://publichealth.jmir.org/2023/1/e47377
doi:10.2196/47377
PMID:37955961

Original Paper

Hospital-Level Implementation Barriers, Facilitators, and Willingness to Use a New Regional Disaster Teleconsultation System: Cross-Sectional Survey Study

Tehnaz Boyle[1], MD, PhD; Krislyn Boggs[2], MPH; Jingya Gao[2], MS; Maureen McMahon[3], BSN, MS, RN; Rachel Bedenbaugh[4], MPH; Lauren Schmidt[4], MPH; Kori Sauser Zachrison[2], MSc, MD; Eric Goralnick[5], MD, MS; Paul Biddinger[4], MD; Carlos A Camargo Jr[2], MD, DrPH

[1]Department of Pediatrics, Boston Medical Center, Boston, MA, United States
[2]Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, United States
[3]Department of Emergency Management, Boston Medical Center, Boston, MA, United States
[4]Center for Disaster Medicine, Massachusetts General Hospital, Boston, MA, United States
[5]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
Tehnaz Boyle, MD, PhD
Department of Pediatrics
Boston Medical Center
801 Albany Street
Rm 4025
Boston, MA, 02119
United States
Phone: 1 617 414 3682
Email: tehnaz.boyle@bmc.org

Abstract

Background: The Region 1 Disaster Health Response System project is developing new telehealth capabilities to provide rapid, temporary access to clinical experts across US jurisdictions to support regional disaster health response.

Objective: To guide future implementation, we identified hospital-level barriers, facilitators, and willingness to use a novel regional peer-to-peer disaster teleconsultation system for disaster health response.

Methods: We used the National Emergency Department Inventory-USA database to identify all 189 hospital-based and freestanding emergency departments (EDs) in New England states. We digitally or telephonically surveyed emergency managers regarding notification systems used for large-scale no-notice emergency events, access to consultants in 6 disaster-relevant specialties, disaster credentialing requirements before system use, reliability and redundancy of internet or cellular service, and willingness to use a disaster teleconsultation system. We examined state-wise hospital and ED disaster response capability.

Results: Overall, 164 (87%) hospitals and EDs responded—126 (77%) completed telephone surveys. Most (n=148, 90%) receive emergency notifications from state-based systems. Forty (24%) hospitals and EDs lacked access to burn specialists; toxicologists, 30 (18%); radiation specialists, 25 (15%); and trauma specialists, 20 (12%). Among critical access hospitals (CAHs) or EDs with <10,000 annual visits (n=36), 92% received routine nondisaster telehealth services but lacked toxicologist (25%), burn (22%), and radiation (17%) specialist access. Most hospitals and EDs (n=115, 70%) require disaster credentialing of teleconsultants before system use. Among 113 hospitals and EDs with written disaster credentialing procedures, 28% expected completing disaster credentialing within 24 hours, and 55% within 25-72 hours, which varied by state. Most (n=154, 94%) reported adequate internet or cellular service for video-streaming; 81% maintained cellular service despite internet disruption. Fewer rural hospitals and EDs reported reliable internet or cellular service (19/22, 86% vs 135/142, 95%) and ability to maintain cellular service with internet disruption (11/19, 58% vs 113/135, 84%) than urban hospitals and EDs. Overall, 133 (81%) were somewhat or very likely to use a regional disaster teleconsultation system. Large-volume EDs (annual visits ≥40,000) were less likely to use the service than smaller ones; all CAHs and nearly all rural hospitals or freestanding EDs were likely to use disaster consultation services. Among hospitals and EDs somewhat or very unlikely to use the system (n=26), sufficient consultant access (69%) and reluctance to use new technology or systems (27%) were common barriers. Potential delays (19%), liability (19%), privacy (15%), and hospital information system security restrictions (15%) were infrequent concerns.

Conclusions: Most New England hospitals and EDs have access to state emergency notification systems, telecommunication infrastructure, and willingness to use a new regional disaster teleconsultation system. System developers should focus on ways to improve telecommunication redundancy in rural areas and use low-bandwidth technology to maintain service availability to CAHs and rural hospitals and EDs. Policies and procedures to accelerate and standardize disaster credentialing are needed for implementation across jurisdictions.

(JMIR Public Health Surveill 2023;9:e44164) doi:10.2196/44164

KEYWORDS

disaster medicine; disaster; telemedicine; telehealth; eHealth; teleconsultation; remote consultation; health care delivery; e-consult; notification; alert; emergency; health system; hospital management

Introduction

Background

Recent large-scale, no-notice disasters such as mass shootings and tornadoes, and public health emergencies like the COVID-19 pandemic, have exposed how even the most prepared health care systems can be challenged to function effectively in a crisis [1]. To address this problem, the Administration for Strategic Preparedness and Response in the US Department of Health and Human Services launched a new program to develop regional disaster health response systems (RDHRS) as part of the national strategic response plan to support catastrophic disaster medical care [2]. The goal of an RDHRS is to create partnerships among hospitals and health care facilities that support regional (ie, multistate) health care delivery when existing local and state capacity and capability are exceeded by catastrophic events. RDHRS efforts include novel adaptations of telehealth tools and systems to rapidly expand access to highly specialized, disaster-relevant clinical subspecialists on a regional or national level to support medical response in the immediate aftermath of unusual hazards or catastrophic events.

Importance

Regionalization of routine specialty care, like trauma and pediatric emergency care, has concentrated the available expert workforce in urban settings and academic medical centers [3], exacerbating access disparities in underserved areas [4-6]. Telehealth services have been used to bridge access to specialty care in underserved, often rural, communities, and in low-resource settings [7-10]. Despite technological advancements, legal, regulatory, and administrative barriers and variability in access to reliable high-speed telecommunications infrastructure have continued to limit the development and use of regional telehealth care delivery models in the United States [11-13]. Variability in medical licensing and credentialing procedures has also hindered the ability to leverage clinical expertise across state lines and jurisdictions [12]. Telehealth services have demonstrated use in military disaster settings [14]. However, civilian telehealth systems to support trauma, burn, pediatric, and other highly regionalized specialty care remain generally underdeveloped and inaccessible to many community hospitals that may be overwhelmed in disaster events [15,16].

Goals of This Investigation

The Region 1 RDHRS, which covers 6 New England states, is developing a new peer-to-peer disaster teleconsultation system to support disaster health response across state lines [17]. This system is designed to rapidly expand regional access to disaster-relevant medical experts (eg, burn surgeons), who may be in limited supply immediately following a large-scale, no-notice emergency event, such as a mass casualty incident. The goal is not to replace existing telehealth services but to supplement overwhelmed services or provide new ones where they are otherwise unavailable when needed in a crisis. The RDHRS disaster teleconsultation system aims to provide local bedside providers access to electronic consultation services with remote clinical specialists via telephone, video, or other platforms compliant with regulatory standards governing protected health information within hours of an event [18]. This would allow disaster-relevant clinical experts to participate early in the acute response phase before in-person disaster teams can be mobilized and deployed to the field. To guide development and future implementation, the objectives of this study were to identify barriers and facilitators to implementation and to determine willingness (primary outcome) to use a new regional disaster teleconsultation system at a hospital level in New England.

Methods

Study Design and Setting

We conducted a cross-sectional survey study of all hospital-based and freestanding emergency departments (EDs; hospitals and EDs) open in March 2021 in the 6 New England states (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont).

Ethical Considerations

For this study, we collected only facility-level information about hospitals and EDs. No compensation was provided for the completion of the study. All data are securely stored on a password-protected server accessible to the study staff. The Mass General Brigham Human Research Committee reviewed this project and classified it as exempt.

Survey Administration

Staff from the Emergency Medicine Network (Boston, Massachusetts) used the National Emergency Department

Inventory (NEDI)-USA database to identify all 189 New England hospitals and EDs open during the study period. Briefly, NEDI-USA contains information about all nonfederal, nonspecialty US EDs [19]. Between March 1 and June 30, 2021, we called the main telephone number for each hospital or ED and requested to speak to the emergency manager for survey completion by structured interview. We selected emergency managers as key respondents, as these individuals are responsible for and most knowledgeable about the disaster preparedness and response capabilities of their facilities.

Three staff members with formal training and prior experience with contacting hospitals and EDs for other surveys about facility characteristics conducted the surveys. All used a standardized protocol and interview guide with scripted explanations in case respondents requested clarification of terms and definitions. Telephone survey responses were directly entered into a secure database by the interviewer, using Research Electronic Data Capture (REDCap; Vanderbilt University) tools hosted at Massachusetts General Hospital [20]. REDCap is a secure, web-based software platform designed to support data capture for research studies. We also created a web-based version of the survey administered via REDCap as an alternate means of participation. The web-based survey link was active throughout the entire course of data collection.

Participants were given the option to either complete the interview during the first call, schedule a time for callback, or receive the survey by email. All respondents were able to ask clarifying questions by telephone or via email before completing the survey regardless of the mode of survey administration. Survey respondents also had the option to leave questions unanswered or consult other hospital staff (eg, compliance officers) to confirm answers if needed.

Measurements

The survey questions were developed by the multidisciplinary region 1 RDHRS Telehealth Working Group, which includes experts in adult and pediatric emergency medicine, disaster medicine, public health, hospital emergency system management, and telehealth implementation. Survey questions were reviewed by a hospital credentialing and compliance officer, 2 directors of emergency preparedness, and 2 emergency management specialists to ensure face validity.

The survey collected information about (1) how each hospital and ED receives notification of large-scale or state-wide no-notice emergency events, (2) the availability of 6 disaster-relevant clinical specialist types (toxicology, radiation or nuclear medicine, trauma, burn, high-consequence infectious disease, and critical care) for consultation in-person, via telephone, or via telehealth within 4 hours of an event, (3) disaster credentialing requirements of remote specialists before system use, (4) availability of internet or cellular service to support video streaming in clinical spaces, and (5) whether each hospital and ED has a contingency plan if internet or cellular service is disrupted. To assess hospital-level willingness (primary outcome), we asked respondents how likely it would be for their hospital and ED to use a new disaster teleconsultation system to access specialists (very likely, somewhat likely, somewhat unlikely, or very unlikely). Among those who were "somewhat unlikely" or "very unlikely" to use this system, we asked about potential barriers to use (Multimedia Appendix 1).

Additionally, using the NEDI-USA data set, we identified each ED's total annual visit volume (<10,000; 10,000-19,000; 20,000-39,000; ≥40,000), whether or not each ED was a designated critical access hospital (CAH) [21] or freestanding ED [22], whether each hospital and ED was in an urban or rural location based on location within or outside of a core based statistical area, and whether the ED received routine (nondisaster) telehealth services [23].

Primary Data Analysis

We used descriptive and bivariate statistics to describe and compare the disaster response characteristics of New England hospitals and EDs by state. Data are presented as frequencies with percentages. A chi-square test and the Fisher exact test were used to examine interstate variation in ED disaster response capability. Fisher exact test was used alternatively for low counts. P values of <.05 were considered statistically significant. We also created maps to demonstrate the geographic differences in the hospital or ED disaster response infrastructure. Data analysis and mapping were performed using SAS (version 9.4; SAS Institute).

Results

Of the 189 hospitals and EDs open in New England in 2021, a total of 164 (87%) responded to the survey. Among these, 126 (77%) completed the survey by phone interview, and 38 (23%) completed it digitally. The response rate in each of the 6 New England states was ≥83%. There were no differences in ED characteristics between responders and nonresponders.

Table 1 summarizes the emergency notification mechanisms, whereas Table 2 summarizes gaps in access to disaster-relevant specialists for consultation in New England hospitals and EDs. While the percentage of hospitals and EDs notified by state-level systems or emergency medical services (EMS) and public safety systems was similar across all states, notification by regional- or county-level systems or via media mechanisms was more varied (Table 1). Most New England hospitals and EDs (97%) could access at least 1 type of disaster-relevant specialist for consultation in-person, via telephone, or via telehealth within the first 4 hours of a large-scale, no-notice emergency event (Table 2). There was no significant difference in access to specialists by state.

ED characteristics are summarized in Table 3. Because small volume EDs (<10,000 annual visits), CAHs, and rural hospitals may lack access to specialists, we further examined this subset (Table S1 in Multimedia Appendix 1). Eighteen of the 22 EDs with <10,000 annual visits were also designated CAHs. Among CAHs or EDs with <10,000 annual visits (n=36), 92% received routine nondisaster telehealth services, and most reported access to specialists in trauma (92%), high-consequence infectious disease (92%), and critical care (97%). However, 25% lacked access to toxicologists, 22% to burn specialists, and 17% to radiation or nuclear medicine specialists.

Next, we asked if hospitals and EDs would require disaster credentialing of teleconsultants before permitting system use. Overall, 70% (115/164) of New England hospitals and EDs reported that they would require disaster credentialing and 113 of these had written procedures in their medical staff bylaws. Among those with written procedures (n=113), disaster credentialing (n=105, 93%) was used more frequently than credentialing-by-proxy (n=16, 14%) or other procedures (n=5, 4%). Over half (n=62, 55%) expected disaster credentialing to take 25 to 72 hours to complete and varied by state. Only 32 (28%) expected to complete these procedures within 24 hours and 17 (15%) within 4 hours. Figure 1 depicts the variability in expected time to complete disaster credentialing procedures in hospitals and EDs within and between states in New England. Notably, 43 (38%) had used disaster credentialing procedures during a major event in the past 20 years, of which 53% had successfully completed it within 4 hours. Approximately two-thirds of hospitals and EDs requiring disaster credentialing of teleconsultants were willing to use a third-party verification system to complete this procedure. More hospitals and EDs in Rhode Island (100%), Maine (94%), and New Hampshire (86%) were willing to use a third-party verification system than in Vermont (57%), Connecticut (52%), and Massachusetts (47%).

We then examined the reliability and redundancy of hospital telecommunication infrastructure as this could affect adoption and use of a disaster teleconsultation system. Overall, 154 (94%) hospitals and EDs reported reliable internet or cellular service connectivity for video streaming in clinical areas, and 142 (87%) had a contingency plan for loss of telecommunication services. Of the 154 hospitals and EDs with reliable connectivity, 124 (81%) could maintain cellular service if there was an internet disruption. Hospitals and EDs in Maine, New Hampshire, and Vermont had less reliable telecommunication and lower capability of maintaining services during disruption (Figure 2). Fewer rural hospitals and EDs reported reliable internet or cellular service (19/22, 86% vs 135/142, 95%) and ability to maintain cellular service with internet disruption (11/19, 58% vs 113/135, 84%) than urban hospitals and EDs.

Most (n=133, 81%) New England hospitals and EDs were very or somewhat likely to use RDHRS disaster telehealth services to access specialists if a large-scale no-notice event affected their facility. Hospitals and EDs in Vermont, New Hampshire, and Rhode Island reported higher willingness to use the system than those in Maine, Massachusetts, and Connecticut (Figure 3). Among the 26 (16%) hospitals and EDs that were somewhat or very unlikely to use the system, the leading barriers were sufficient access to specialists in-house or within the health system (n=18, 69%) and reluctance to use new technology or systems during a disaster (n=7, 27%). Potential time delays (n=5, 19%), liability (n=5, 19%), privacy (n=4, 15%), and hospital information system security restrictions (n=4, 15%) were other reported concerns.

Table 3 examines the willingness to use disaster teleconsultation services by ED characteristics. While overall willingness was high, large volume EDs (annual visits ≥40,000) were less likely to use the service than smaller ones. In contrast, all CAHs and nearly all rural hospitals or freestanding EDs were likely to use disaster consultation services. Eighty percent of hospitals and EDs that were likely to use the system also received nondisaster telehealth services.

Table 1. Overall and by-state comparison of mechanisms by which New England hospitals and emergency departments are notified of large-scale or state-wide no-notice emergency events.

Disaster response capability	Total (n=164), n (%)	Connecticut (n=33), n (%)	Massachusetts (n=57), n (%)	Maine (n=30), n (%)	New Hampshire (n=24), n (%)	Rhode Island (n=8), n (%)	Vermont (n=12), n (%)	*P* value
EMS[a] or public safety system notification	142 (87)	32 (97)	47 (82)	27 (90)	21 (88)	7 (88)	8 (67)	.13
Hospital network emergency notification system activated at state level	148 (90)	33 (100)	48 (84)	27 (90)	20 (83)	8 (100)	12 (100)	.10
Hospital network emergency notification system activated at regional or county level within a state	133 (81)	28 (85)	48 (84)	28 (93)	17 (71)	6 (75)	6 (50)	.02
Media notification	80 (49)	29 (88)	24 (42)	12 (40)	8 (33)	2 (25)	5 (42)	<.001
Other[b]	13 (8)	5 (15)	5 (9)	3 (10)	0 (0)	0 (0)	0 (0)	.36

[a] EMS: emergency medical service.

[b] "Other" includes receiving notifications from a dedicated third party, internal notification system, national alert network such as from Centers for Disease Control, and so forth.

Table 2. Frequency and percentage of New England hospitals and emergency departments reporting no access to specialists for consultation in-person, via telephone, or via telehealth within first 4 hours of large-scale, no-notice emergency event, overall and by state.

Disaster response capability	Total (n=164), n (%)	Connecticut (n=33), n (%)	Massachusetts (n=57), n (%)	Maine (n=30), n (%)	New Hampshire (n=24), n (%)	Rhode Island (n=8), n (%)	Vermont (n=12), n (%)	*P* value
Toxicology	30 (18)	4 (12)	11 (19)	3 (10)	7 (29)	3 (38)	2 (17)	.27
Radiation or nuclear medicine	25 (15)	3 (9)	8 (14)	7 (23)	4 (17)	2 (25)	1 (8)	.61
Trauma	20 (12)	2 (6)	11 (19)	3 (10)	1 (4)	2 (25)	1 (8)	.19
Burn	40 (24)	3 (9)	20 (35)	8 (27)	5 (21)	2 (25)	2 (17)	.08
Infectious disease	7 (4)	1 (3)	1 (2)	4 (13)	1 (4)	0 (0)	0 (0)	.26
Critical care	6 (4)	1 (3)	2 (4)	3 (10)	0 (0)	0 (0)	0 (0)	.59
No access to any specialist	2 (1)	1 (3)	0 (0)	1 (3)	0 (0)	0 (0)	0 (0)	.52

Table 3. Emergency department (ED) characteristics by willingness to use a regional disaster teleconsultation system.[a]

Characteristics	Total (N=164), n (%)	Very likely (N=47), n (%)	Somewhat likely (N=86), n (%)	Somewhat unlikely (N=23), n (%)	Very unlikely (N=3), n (%)
Annual ED visit volume					
<10,000	22 (13)	8 (17)	12 (14)	1 (4)	0 (0)
10,000-19,999	46 (28)	8 (17)	32 (37)	4 (17)	1 (33)
20,000-39,999	48 (29)	21 (45)	20 (23)	6 (26)	1 (33)
≥40,000	48 (29)	10 (21)	22 (26)	12 (52)	1 (33)
Critical access hospital					
Yes	32 (20)	12 (26)	19 (22)	0 (0)	0 (0)
No	132 (80)	35 (74)	67 (78)	23 (100)	3 (100)
Freestanding ED					
Yes	12 (7)	2 (4)	9 (10)	0 (0)	1 (33)
No	152 (93)	45 (96)	77 (90)	23 (100)	2 (67)
Urban status					
Urban	142 (87)	41 (87)	72 (84)	22 (96)	3 (100)
Rural	22 (13)	6 (13)	14 (16)	1 (4)	0 (0)
Access to routine (nondisaster) telehealth services					
Yes	132 (80)	41 (87)	65 (76)	18 (78)	3 (100)
No	17 (10)	2 (4)	11 (13)	4 (17)	0 (0)
Missing	15 (9)	4 (9)	10 (12)	1 (4)	0 (0)

[a] 5 participants skipped the willingness question.

Figure 1. By-state comparison of expected timeline to complete disaster/emergency credentialing and privileging procedures among New England hospitals and emergency departments that would require disaster credentialing and had written procedures in medical staff bylaws (n=113).

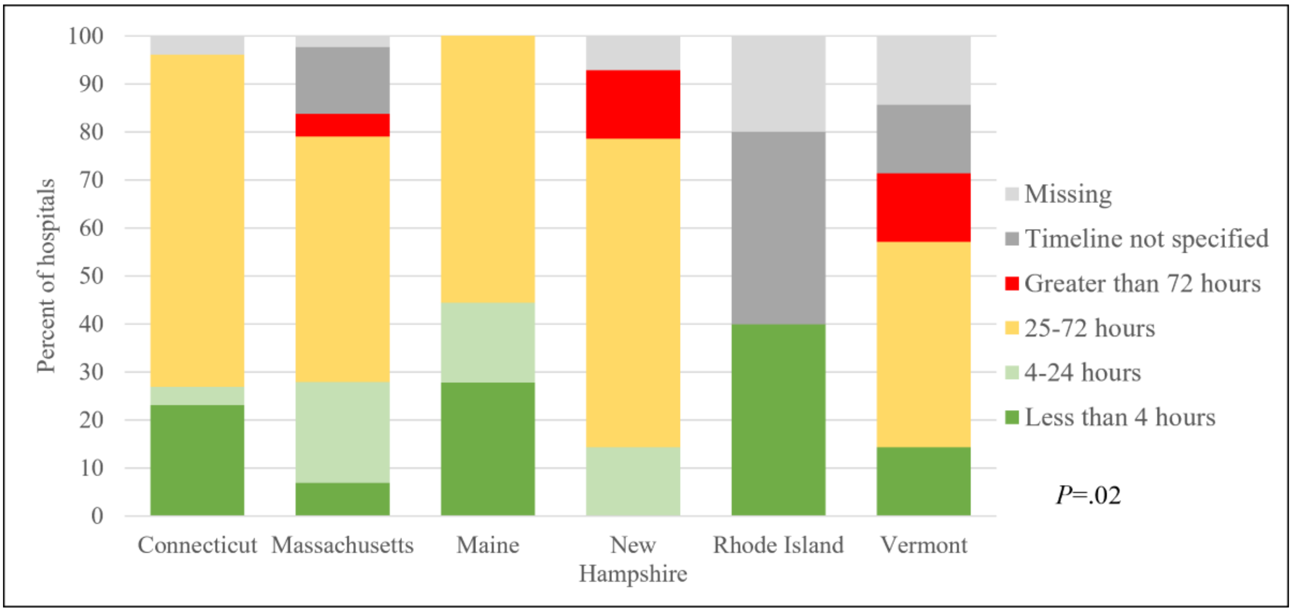


Figure 2. Reliability and redundancy of telecommunications infrastructure. (A) The percentage of emergency departments with adequate internet and cellular service to support video-streaming (reliability). (B) The percentage of emergency departments that maintain cellular service despite internet disruption (redundancy). ED: emergency department.

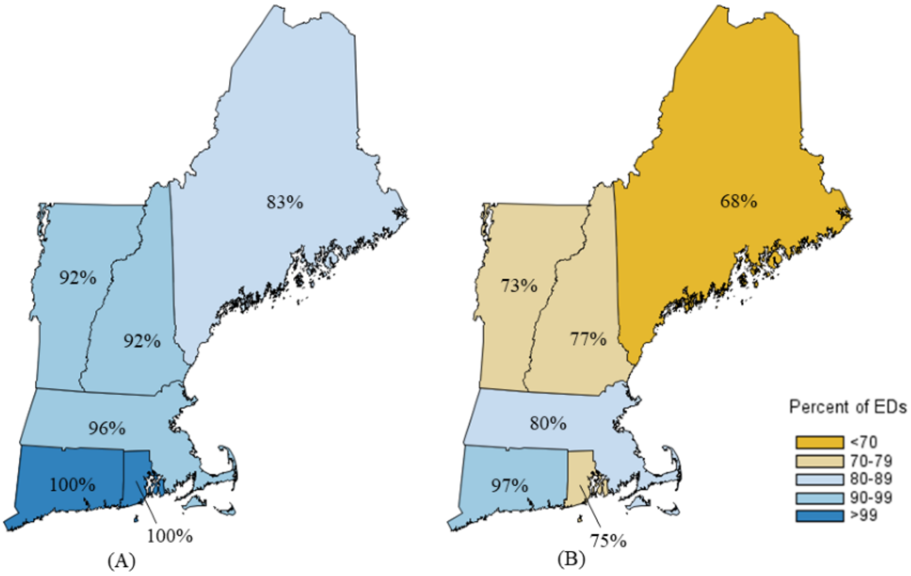
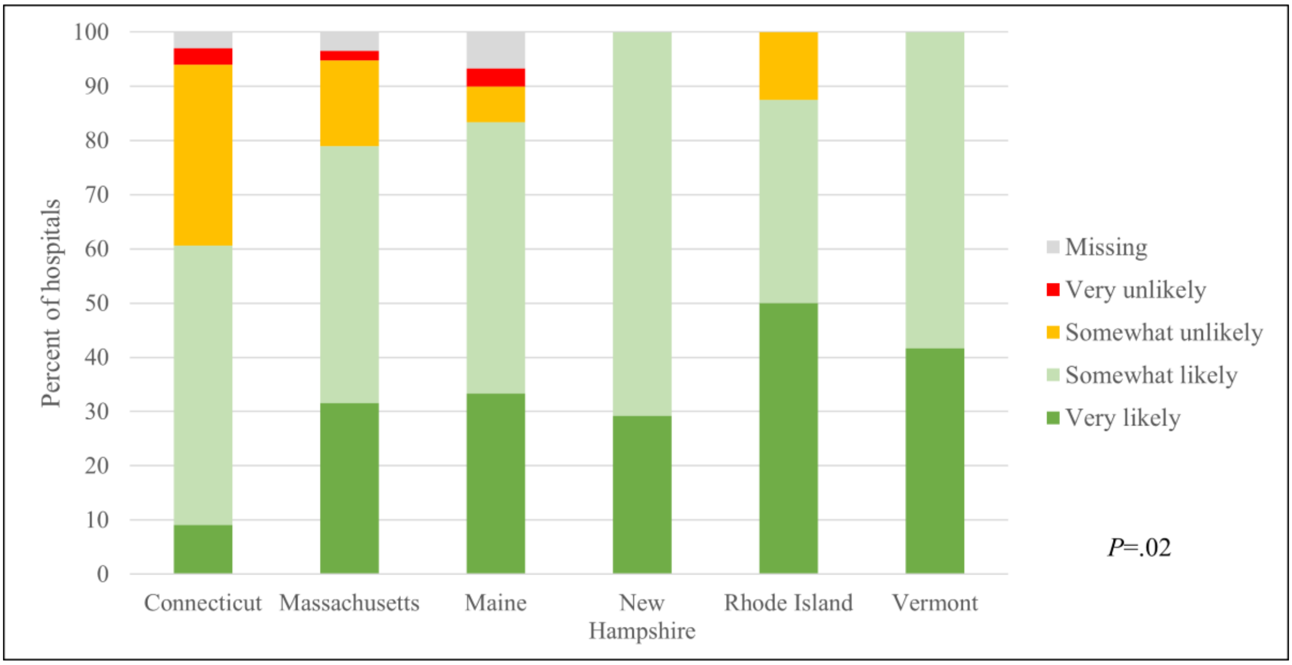


Figure 3. Willingness of New England hospitals and emergency departments to use a regional disaster teleconsultation system to access specialists by state (n=164).



Discussion

Principal Results

In this regional study, we found that most New England hospitals and EDs had the necessary emergency notification mechanisms, telecommunication infrastructure, and willingness to support the future regional implementation of a disaster teleconsultation system. Although many EDs had access to at least 1 disaster-relevant specialist, 1 in 4 New England hospitals and EDs lacked access to burn specialists, identifying an important gap to target in future service delivery. Rural hospitals, freestanding EDs, and CAHs demonstrated the highest willingness to use this type of service but may require improvements in telecommunication infrastructure to support use in conditions where connectivity is disrupted or infrastructure destroyed. Despite limiting disaster telehealth services to teleconsultation only, 7 in 10 hospitals and EDs would require disaster credentialing before using the service and less than a third of these sites could complete this procedure within 24 hours. This could significantly delay postdisaster access to specialists required for time-critical patient care.

Comparison With Prior Work

Identifying which emergency notification mechanisms are used most frequently by regional hospitals can help RDHRS programs prioritize the most effective communication channels. RDHRS programs are designed to build on local and state assets and emergency response capabilities and provide regional coordination of key health care resources, such as burn centers and pediatric hospitals. Our findings suggest that integrating RDHRS communication strategies with hospital emergency network notification systems activated by state, EMS, and public safety systems would be the most efficient mechanisms for rapid regional communication. Such channels could be used to notify hospitals of the availability of regional disaster services in the aftermath of an event. Formalizing RDHRS programs as regional response entities within the national strategic response plan would facilitate future integration with existing emergency notification systems at the local and state level.

Depending on regional need and availability, RDHRS programs may need to create new disaster telehealth service lines or facilitate mechanisms to expand access to existing ones. Understanding gaps in regional access to disaster-relevant specialty care is necessary to target which service lines need the most urgent development, especially as rural and urban practitioners may perceive service need and use differently [24]. Among New England hospitals and EDs, we found larger gaps in access to burn, toxicology, radiation, and trauma specialists than to infectious disease and critical care specialists. Burn and trauma care are notable examples where disparities in routine access may be worsened by a disaster. For example, after a mass casualty incident, interfacility patient transfers may be delayed due to limited availability of emergency medical services for transportation, and overwhelming volume may limit capacity at receiving specialty hospitals. Regional disparities in access to burn centers have been described [25,26]. Despite the ability to alter triage, care, and transfer decisions [27-29], many burn telehealth programs remain localized without regional or national coordination due to barriers in sustainable funding and license portability [30,31]. Regional variation in access to trauma centers impacts injury mortality and emergency resource usage [32,33]. While trauma telehealth programs support providers who are uncomfortable or unfamiliar with specialty trauma care and improve transfer efficiency for acutely injured patients [34-36], adoption has lagged.

Mass casualty incidents involving radiological and chemical agents also require immediate access to experts in radiological exposure, hazardous materials, and medical toxicology. Although federal guidance emphasizes early mass decontamination during the emergency response to such incidents [37], a qualitative study across 3 US regions found ED clinicians perceived a lack of hospital preparedness for radiological terrorism [38]. Teleconsultation with hazardous material or radiation specialists could be used to support field providers and receiving hospitals, although studies examining the use of telehealth for such incidents are sparse. Teleconsultation with patients for counseling and risk communication after exposure could add public health benefit, as described after the 2011 Fukushima nuclear disaster [39]. Regional poison centers can also be integrated into disaster response infrastructure to leverage available medical toxicology expertise [40]. Like other service lines with low-frequency use, the adoption of routine teleconsultation with medical toxicologists has been limited by reimbursement barriers [41]. Dedicated funding to ensure the correct people, equipment, and administrative supports are maintained during the preparedness portion of the disaster cycle will be essential to ensure effective national strategic response for future events.

Historically, concerns over the cost of technology and service use, and administrative barriers related to interstate license portability and hospital credentialing and privileging requirements have slowed the expansion of telehealth in the United States [12,42-45]. Licensure refers to the process of securing the authority to practice medicine within a state. Credentialing refers to the administrative process of validating the competency of a health care provider by verifying their license, education, certification, and other information to ensure they meet the practice standards required by a hospital. Disaster credentialing is an expedited procedure that can only be used when a hospital's emergency operations plan is activated to facilitate rapid access to outside assistance (eg, medical professionals) to sustain patient care. Our findings highlight how variability in regulatory and administrative practices at the state and hospital level could hobble efforts to integrate telehealth tools into regional disaster medical response. Despite limiting disaster services to peer-to-peer teleconsultation with certified expert providers, 70% of New England hospitals and EDs would require disaster credentialing of teleconsultants before permitting service use with many requiring up to 72 hours to complete this process. This administrative barrier could delay implementation by days to weeks and negate the potential benefits of accessing specialists within hours of a no-notice event, like a mass casualty incident, where interventions and outcomes are critically time sensitive.

Our findings are surprising as peer-to-peer teleconsultation for the purpose of providing expert advice is not subject to the same legal and regulatory requirements as telehealth services that deliver direct patient care [11]. Common licensure exceptions include physician-to-physician consultations, public health services, medical emergencies, and natural disasters [11]. For example, Regional Poison Control Centers represent a widely accepted model where episodic, emergency consultations are delivered via telephone by certified expert personnel across jurisdictions [46]. These consultations are freely and routinely accessible to providers without administrative requirements for hospitals using the service. Whether survey respondents perceived expert teleconsultations via a regional disaster health response system as distinct from telephone consultations via a regional poison control center is unclear. Finally, in the absence of clear guidance on acceptable standards, hospital administrators may vary in their approach to disaster credentialing to avoid potential regulatory penalties and minimize medicolegal risk [12]. We encourage future research to understand the reasons underlying the variability in hospital credentialing practices, as this could have important implications for policy makers.

Creation of a national licensure standard, incentivizing the adoption of existing mechanisms to streamline interstate licensure such as the Interstate Medical Licensure Compact [47], or permitting limited exceptions to telehealth licensure laws could all support the ability to maintain a rapidly deployable network of disaster-relevant specialists at the ready. The Centers for Medicare and Medicaid Services could also incentivize hospitals to adopt mechanisms to streamline disaster credentialing procedures by using rapid third-party verification mechanisms. For example, Provider Bridge, a web-based platform to speed verification of out-of-state health care volunteers, was developed by the Federation of State Medical Boards with temporary funding from the Health Resources and Services Administration and the Coronavirus Aid, Relief, and Economic Security Act [48]. However, cessation of grant funding has left the future of such programs uncertain. We strongly support continued funding and intentional integration of effective pilot programs to ensure investments made to support pandemic response are available for future disaster response.

Successful application of telehealth requires reliable access to high-speed broadband networks and functional telecommunications infrastructure. Broadband refers to various technologies (eg, cable, wireless, satellite, wireless, and mobile) that provide high-speed data transmission and connection to the internet. Prior studies have described disparities in access to broadband coverage in the United States, particularly in rural communities and tribal lands [49,50]. By the most recent Federal Communications Commission estimates, 77% of Americans in rural areas and 72% in tribal lands have broadband access via fixed terrestrial (eg, cable) and mobile (cellular) networks at speeds that can support 2-way video-conferencing as compared to 99% of Americans in urban areas [51]. In this study, most New England hospitals and EDs reported adequate broadband access via fixed and mobile networks to support video streaming in clinical spaces. However, 1 in 5 rural hospitals in New England may lose this capability if fixed networks are disrupted because they have inadequate access to mobile networks within facility walls. More research is required to understand regional capabilities to mobilize satellite or mobile towers when infrastructure is lost.

Telehealth platforms that function at low bandwidth speeds and operate on both fixed and mobile networks will be necessary for resiliency in disasters where telecommunications infrastructure may be damaged or destroyed, or networks overwhelmed by surging demand. Additionally, asynchronous communication mechanisms that remain operational in limited bandwidth settings could also be sufficient to serve some disasters and areas. Federal investment in closing the "digital divide" should include strategies to ensure reliability and redundancy of telecommunications infrastructure in vulnerable facilities in rural areas and tribal lands. A similar approach could support telehealth applications for humanitarian response in resource-limited settings and different cultural or geographic environments.

Widespread adoption and use of telehealth during the COVID-19 pandemic has resulted in broad technology acceptance within the global health care community. In the postpandemic environment, cost may become a lesser barrier as clinical needs push providers and organizations to adopt new models of remote care delivery [52]. We found high willingness to use regional disaster teleconsultation services to access key specialist groups among New England hospitals and EDs. This survey was conducted in 2021 after the first year of the pandemic, when many hospitals were still struggling with large patient surges and workforce shortages and using emergency operations [53,54]. While this may have biased respondents toward expressing greater willingness, their attitudes are arguably informed by real-world experience rather than conjecture. Whether frontline providers and clinicians would express similar levels of willingness to use disaster teleconsultation services is unclear, although simulation-based studies suggest this may hold true [17]. Future research examining attitudes and drivers of the adoption of disaster telehealth services on a provider level, mechanisms to disseminate awareness of service availability, and integration with existing emergency management plans will be needed to ensure service lines will actually be used. Finally, even the hospitals and EDs that are willing to use disaster telehealth services may experience barriers to implementation. Thus, system developers should not only indicate potential demotivating factors, barriers, or weaknesses to implementation but also identify critical acceptance criteria to implement disaster telehealth services on a hospital level.

The motivators and challenges to the implementation of regional disaster telehealth systems for cross-jurisdictional response in the United States parallel applications of telehealth for international humanitarian response [55]. The Sendai Framework for Disaster Risk Reduction 2015-2030 was developed to guide international efforts in disaster risk reduction through coordinated implementation of various measures to reduce vulnerabilities, increase preparedness for response and recovery, and strengthen resilience [56]. In 2022, rapidly deployable disaster telehealth services coordinated by an international, nongovernmental collaborative of digital health, telehealth, disaster response, and medical experts and staffed by clinical volunteers across the globe provided access to health care for Ukrainians displaced by the Russia-Ukraine conflict [57]. This was a powerful demonstration of how a disaster telehealth system could leverage a remote international workforce to support large-scale humanitarian response and supplement in-person field response. However, in the United States, many insurers and state licensing boards only permit telemedical practice with oversight due to the risk of potential abuse. Furthermore, lack of adoption of standards or agreements prevents efficient use of telehealth solutions to support timely "surge capacity" to deliver health services during emergencies because expert volunteers are mired in cumbersome and complicated processes of entering into agreements with their home jurisdictions before they can be digitally "deployed" across jurisdictions. We call for the uniform adoption of national (eg, Uniform Emergency Volunteer Health Practitioners Act) and international standards governing licensure and verification of medical volunteers practicing across jurisdictions to facilitate engagement and avoid implementation delays.

Limitations

This study has potential limitations. First, we defined disaster teleconsultation services as electronic consultations between providers as recognized by the Health Resources and Services Administration [18]. Here, teleconsultants provide peer-to-peer advice only and the bedside clinician makes final patient care decisions. However, some respondents may not have distinguished this from nonconsultative services or other types of routine, nondisaster telehealth services where a remote clinician directs patient care. One possibility is that emergency managers lack knowledge of the credentialing procedures that are the purview of compliance officers and provided erroneous answers. Another possibility is that despite our efforts to define the limited scope of disaster teleconsultation services, respondents were confused by or disagreed with our definitions. If respondents did not distinguish between teleconsultation for advice and telehealth to provide direct patient care, they could perceive credentialing procedures as necessary. Confusion about the service type offered could overestimate variability in disaster credentialing procedures. However, respondents were provided ample opportunity to ask clarifying questions regarding terms and definitions during survey administration, and we did not perceive any confusion to this end.

Second, the data were self-reported by 1 person per site, which may have introduced information bias. However, we mitigated this by contacting emergency managers as respondents, as they are responsible for and knowledgeable about disaster operations and plans for their facilities. Third, the attitudes of hospital emergency managers may not reflect those of frontline clinicians or other hospital administrators, so it is possible our estimates of willingness to use regional disaster telehealth services are inaccurate. Future research should examine the attitudes of additional groups involved in the adoption and implementation of telehealth services for disaster response on a hospital level. Last, this survey was administered in New England, which may not be representative of the other regions. However, diversity in geography and hospital and ED types within the region supports the generalizability of the findings.

Conclusions

In summary, we found that most hospitals and EDs in New England have the necessary access to statewide emergency notification mechanisms, telecommunications infrastructure, and willingness to use a new regional disaster teleconsultation system to access specialists if a large-scale event affected their facility. Even in a relatively resource-rich region, access to burn care was the most limited among numerous disaster-relevant specialty types and should be an early focus for service line development. We also identified significant variability and lack of standardization in hospital credentialing practices, which could delay timely access to teleconsultants and diminish potential benefits for no-notice disasters with time-sensitive outcomes. These findings can inform how health systems and disaster response organizations can coordinate strategic plans to implement and integrate new telehealth tools for regional disaster medical response in New England. We encourage future research on interregional differences in clinical needs, gaps in telecommunications infrastructure, and reasons for variability in disaster credentialing practices at a hospital level. This research is critical to build an effective telehealth capability that supports real-world needs for regional disaster health care response across the United States.

Acknowledgments

The authors would like to thank Lily Y Lu, Sabrina N Robichaud, and Cordelia Zhong for their assistance with data collection, and Gerard Dubois and the Northeast Telehealth Resource Center for expert guidance. This work is funded by Administration for Strategic Preparedness and Response (grant #HITEP180042). TB is funded by the National Institutes of Health (NIH; grant #1K23HL145126) and the Health Resources and Services Administration/Maternal Child and Health Bureau (grant #U1I45814). The contents herein are solely the authors’ responsibility and do not necessarily represent the official views of the Administration for Strategic Preparedness and Response or NIH.

Data Availability

Study data were collected with the promise that individual hospital or emergency department data would not be made publicly available. Thus, the data sets generated during and analyzed during this study have restricted access and are not publicly available. Limited data are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey and supplemental table.
[DOCX File , 34 KB - publichealth_v9i1e44164_app1.docx]

References

1. Balsari S, Kiang MV, Buckee CO. Data in crisis: rethinking disaster preparedness in the United States. N Engl J Med 2021;385(16):1526-1530 [FREE Full text] [doi: 10.1056/NEJMms2104654] [Medline: 34469643]
2. Regional disaster health response system: building regional solutions. Administration for Strategic Preparedness and Response. URL: https://www.phe.gov/Preparedness/planning/RDHRS/Pages/default.aspx [accessed 2023-05-26]
3. Bennett CL, Espinola JA, Sullivan AF, Boggs KM, Clay CE, Lee MO, et al. Evaluation of the 2020 pediatric emergency physician workforce in the US. JAMA Netw Open 2021;4(5):e2110084 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.10084] [Medline: 34003272]
4. Carr BG, Bowman AJ, Wolff CS, Mullen MT, Holena DN, Branas CC, et al. Disparities in access to trauma care in the United States: a population-based analysis. Injury 2017;48(2):332-338 [FREE Full text] [doi: 10.1016/j.injury.2017.01.008] [Medline: 28069138]
5. Wei R, Mann NC, Dai M, Hsia RY. Injury-based geographic access to trauma centers. Acad Emerg Med 2019;26(2):192-204 [FREE Full text] [doi: 10.1111/acem.13518] [Medline: 30019802]
6. Vali Y, Rashidian A, Jalili M, Omidvari AH, Jeddian A. Effectiveness of regionalization of trauma care services: a systematic review. Public Health 2017;146:92-107. [doi: 10.1016/j.puhe.2016.12.006] [Medline: 28404479]
7. Delaigue S, Bonnardot L, Steichen O, Garcia DM, Venugopal R, Saint-Sauveur JF, et al. Seven years of telemedicine in médecins sans frontières demonstrate that offering direct specialist expertise in the frontline brings clinical and educational value. J Glob Health 2018;8(2):020414 [FREE Full text] [doi: 10.7189/jogh.08.020414] [Medline: 30574293]
8. Uscher-Pines L, Fischer S, Tong I, Mehrotra A, Malsberger R, Ray K. Virtual first responders: the role of direct-to-consumer telemedicine in caring for people impacted by natural disasters. J Gen Intern Med 2018;33(8):1242-1244 [FREE Full text] [doi: 10.1007/s11606-018-4440-8] [Medline: 29691713]
9. Litvak M, Miller K, Boyle T, Bedenbaugh R, Smith C, Meguerdichian D, et al. Telemedicine use in disasters: a scoping review. Disaster Med Public Health Prep 2022;16(2):791-800 [FREE Full text] [doi: 10.1017/dmp.2020.473] [Medline: 33750505]
10. Kim TJ, Arrieta MI, Eastburn SL, Icenogle ML, Slagle M, Nuriddin AH, et al. Post-disaster Gulf Coast recovery using telehealth. Telemed J E Health 2013;19(3):200-210 [FREE Full text] [doi: 10.1089/tmj.2012.0100] [Medline: 23427981]
11. Becker CD, Dandy K, Gaujean M, Fusaro M, Scurlock C. Legal perspectives on telemedicine part 1: legal and regulatory issues. Perm J 2019;23:18-293 [FREE Full text] [doi: 10.7812/TPP/18-293] [Medline: 31314722]

12. Baker-Whitcomb A, Harvey J. Benefits and barriers to telehealth credentialing by proxy. Telemed J E Health 2018;24(11):922-926. [doi: 10.1089/tmj.2017.0129] [Medline: 29620971]
13. Tedeschi C. Ethical, legal, and social challenges in the development and implementation of disaster telemedicine. Disaster Med Public Health Prep 2021;15(5):649-656 [FREE Full text] [doi: 10.1017/dmp.2020.118] [Medline: 32364096]
14. Nettesheim N, Powell D, Vasios W, Mbuthia J, Davis K, Yourk D, et al. Telemedical support for military medicine. Mil Med 2018;183(11-12):e462-e470 [FREE Full text] [doi: 10.1093/milmed/usy127] [Medline: 30496581]
15. Lurie N, Carr BG. The role of telehealth in the medical response to disasters. JAMA Intern Med 2018;178(6):745-746. [doi: 10.1001/jamainternmed.2018.1314] [Medline: 29710200]
16. Madanian S, Norris T, Parry D. Disaster eHealth: scoping review. J Med Internet Res 2020;22(10):e18310 [FREE Full text] [doi: 10.2196/18310] [Medline: 33112244]
17. Boyle TP, Ludy S, Meguerdichian D, Dugas JN, Drainoni ML, Litvak M, et al. Feasibility and acceptability of a model disaster teleconsultation system for regional disaster health response. Telemed J E Health 2023;29(4):625-632. [doi: 10.1089/tmj.2022.0103] [Medline: 36036805]
18. Telehealth for emergency departments: E-consults. Health Resources & Services Administration. URL: https://telehealth.hhs.gov/providers/telehealth-for-emergency-departments/e-consults/ [accessed 2023-05-26]
19. Sullivan AF, Richman IB, Ahn CJ, Auerbach BS, Pallin DJ, Schafermeyer RW, et al. A profile of US emergency departments in 2001. Ann Emerg Med 2006;48(6):694-701. [doi: 10.1016/j.annemergmed.2006.08.020] [Medline: 17067721]
20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
21. Critical access hospital locations list. Flex Monitoring Team. URL: https://www.flexmonitoring.org/data/critical-access-hospital-locations/ [accessed 2023-05-26]
22. Herscovici DM, Boggs KM, Sullivan AF, Camargo CAJ. What is a freestanding emergency department? Definitions differ across major United States data sources. West J Emerg Med 2020;21(3):660-664 [FREE Full text] [doi: 10.5811/westjem.2020.3.46001] [Medline: 32421516]
23. Lu LY, Boggs KM, Espinola JA, Sullivan AF, Cash RE, Camargo CA. Development of a unified national database of burn centers with colocated emergency departments, 2020. J Burn Care Res 2022;43(5):1066-1073. [doi: 10.1093/jbcr/irab238] [Medline: 34893840]
24. Oest SER, Swanson MB, Ahmed A, Mohr NM. Perceptions and perceived utility of rural emergency department telemedicine services: a needs assessment. Telemed J E Health 2020;26(7):855-864 [FREE Full text] [doi: 10.1089/tmj.2019.0168] [Medline: 31580783]
25. Klein MB, Kramer CB, Nelson J, Rivara FP, Gibran NS, Concannon T. Geographic access to burn center hospitals. JAMA 2009;302(16):1774-1781 [FREE Full text] [doi: 10.1001/jama.2009.1548] [Medline: 19861669]
26. Carmichael H, Wiktor AJ, McIntyre RC, Wagner AL, Velopulos CG. Regional disparities in access to verified burn center care in the United States. J Trauma Acute Care Surg 2019;87(1):111-116. [doi: 10.1097/TA.0000000000002259] [Medline: 30865160]
27. den Hollander D, Mars M. Smart phones make smart referrals: the use of mobile phone technology in burn care: a retrospective case series. Burns 2017;43(1):190-194 [FREE Full text] [doi: 10.1016/j.burns.2016.07.015] [Medline: 27575675]
28. Martinez R, Rogers AD, Numanoglu A, Rode H. The value of WhatsApp communication in paediatric burn care. Burns 2018;44(4):947-955 [FREE Full text] [doi: 10.1016/j.burns.2017.11.005] [Medline: 29395403]
29. Garber RN, Garcia E, Goodwin CW, Deeter LA. Pictures do influence the decision to transfer: outcomes of a telemedicine program serving an eight-state rural population. J Burn Care Res 2020;41(3):690-694 [FREE Full text] [doi: 10.1093/jbcr/iraa017] [Medline: 32044972]
30. Holt B, Faraklas I, Theurer L, Cochran A, Saffle JR. Telemedicine use among burn centers in the United States: a survey. J Burn Care Res 2012;33(1):157-162. [doi: 10.1097/BCR.0b013e31823d0b68] [Medline: 22105096]
31. Russell KW, Saffle JR, Theurer L, Cochran AL. Transition from grant funding to a self-supporting burn telemedicine program in the western United States. Am J Surg 2015;210(6):1037-1044 discussion 1042. [doi: 10.1016/j.amjsurg.2015.08.003] [Medline: 26434619]
32. Hashmi ZG, Jarman MP, Uribe-Leitz T, Goralnick E, Newgard CD, Salim A, et al. Access delayed is access denied: relationship between access to trauma center care and pre-hospital death. J Am Coll Surg 2019;228(1):9-20. [doi: 10.1016/j.jamcollsurg.2018.09.015] [Medline: 30359830]
33. Newgard CD, Fu R, Bulger E, Hedges JR, Mann NC, Wright DA, et al. Evaluation of rural vs urban trauma patients served by 9-1-1 emergency medical services. JAMA Surg 2017;152(1):11-18 [FREE Full text] [doi: 10.1001/jamasurg.2016.3329] [Medline: 27732713]
34. Mohr NM, Vakkalanka JP, Harland KK, Bell A, Skow B, Shane DM, et al. Telemedicine use decreases rural emergency department length of stay for transferred North Dakota trauma patients. Telemed J E Health 2018;24(3):194-202. [doi: 10.1089/tmj.2017.0083] [Medline: 28731843]

35. Wood T, Freeman S, Banner D, Martin-Khan M, Hanlon N, Flood F. Factors associated with teletrauma utilization in rural areas: a review of the literature. Rural Remote Health 2021;21(1):6354 [FREE Full text] [doi: 10.22605/RRH6354] [Medline: 33721502]
36. Kim PT, Falcone RA. The use of telemedicine in the care of the pediatric trauma patient. Semin Pediatr Surg 2017;26(1):47-53 [FREE Full text] [doi: 10.1053/j.sempedsurg.2017.01.008] [Medline: 28302285]
37. Collins S, James T, Carter H, Symons C, Southworth F, Foxall K, et al. Mass casualty decontamination for chemical incidents: research outcomes and future priorities. Int J Environ Res Public Health 2021;18(6):3079 [FREE Full text] [doi: 10.3390/ijerph18063079] [Medline: 33802722]
38. Becker SM, Middleton SA. Improving hospital preparedness for radiological terrorism: perspectives from emergency department physicians and nurses. Disaster Med Public Health Prep 2008;2(3):174-184. [doi: 10.1097/DMP.0b013e31817dcd9a] [Medline: 18813129]
39. Yashima S, Chida K. Telephone consultations on exposure to nuclear disaster radiation. Nihon Hoshasen Gijutsu Gakkai Zasshi 2014;70(3):242-249 [FREE Full text] [doi: 10.6009/jjrt.2014_jsrt_70.3.242] [Medline: 24647062]
40. Jacobitz KL, Nailon RE. The role of a regional poison center in assessing and managing disaster risk. Health Secur 2020;18(3):195-204. [doi: 10.1089/hs.2019.0124] [Medline: 32559158]
41. Crane PW, Wiegand TJ, Kamali M, Reif M, Wratni R, Montante R, et al. Telemedicine delivery and successful reimbursement in toxicology. J Med Toxicol 2018;14(3):242-247 [FREE Full text] [doi: 10.1007/s13181-018-0665-5] [Medline: 29845507]
42. Zachrison KS, Boggs KM, Hayden EM, Espinola JA, Camargo CA. Understanding barriers to telemedicine implementation in rural emergency departments. Ann Emerg Med 2020;75(3):392-399. [doi: 10.1016/j.annemergmed.2019.06.026] [Medline: 31474481]
43. Shachar C, Engel J, Elwyn G. Implications for telehealth in a postpandemic future: regulatory and privacy issues. JAMA 2020;323(23):2375-2376 [FREE Full text] [doi: 10.1001/jama.2020.7943] [Medline: 32421170]
44. Rogove HJ, Amoateng B, Binner J, Demaerschalk BM, Sanders RB. A survey and review of telemedicine license portability. Telemed J E Health 2015;21(5):374-381. [doi: 10.1089/tmj.2014.0116] [Medline: 25671795]
45. Merchant KAS, Ward MM, Mueller KJ, Rural Policy Research Institute, RUPRI Center for Rural Health Policy Analysis, University of Iowa College of Public Health, Department of Health Management and Policy. Hospital views of factors affecting telemedicine use. Rural Policy Brief 2015(2015 5):1-4. [doi: 10.1007/978-1-4757-3310-5_1] [Medline: 26793811]
46. McIntire MS, Angle CR. Regional poison-control centers improve patient care. N Engl J Med 1983;308(4):219-221. [doi: 10.1056/NEJM198301273080412] [Medline: 6848924]
47. Interstate Medical Licensure Compact. URL: https://www.imlcc.org/ [accessed 2023-05-26]
48. Providerbridge. URL: https://www.providerbridge.org/ [accessed 2023-05-26]
49. Serrano KJ, Thai CL, Greenberg AJ, Blake KD, Moser RP, Hesse BW. Progress on broadband access to the internet and use of mobile devices in the United States. Public Health Rep 2017;132(1):27-31 [FREE Full text] [doi: 10.1177/0033354916679365] [Medline: 28005473]
50. Bauerly BC, McCord RF, Hulkower R, Pepin D. Broadband access as a public health issue: the role of law in expanding broadband access and connecting underserved communities for better health outcomes. J Law Med Ethics 2019;47(2_suppl):39-42 [FREE Full text] [doi: 10.1177/1073110519857314] [Medline: 31298126]
51. 2020 broadband deployment report. Federal Communications Commission. 2020. URL: https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2020-broadband-deployment-report [accessed 2023-05-26]
52. Kruse C, Heinemann K. Facilitators and barriers to the adoption of telemedicine during the first year of COVID-19: systematic review. J Med Internet Res 2022;24(1):e31752 [FREE Full text] [doi: 10.2196/31752] [Medline: 34854815]
53. Wahlster S, Sharma M, Lewis AK, Patel PV, Hartog CS, Jannotta G, et al. The coronavirus disease 2019 pandemic's effect on critical care resources and health-care providers: a global survey. Chest 2021;159(2):619-633 [FREE Full text] [doi: 10.1016/j.chest.2020.09.070] [Medline: 32926870]
54. Gray BM, Vandergrift JL, Barnhart BJ, Reddy SG, Chesluk BJ, Stevens JS, et al. Changes in stress and workplace shortages reported by U.S. critical care physicians treating coronavirus disease 2019 patients. Crit Care Med 2021;49(7):1068-1082 [FREE Full text] [doi: 10.1097/CCM.0000000000004974] [Medline: 33730741]
55. He S, Marzouk S, Balk A, Boyle T, Lee J. The telehealth advantage: supporting humanitarian disasters with remote solutions. Am J Disaster Med 2022;17(2):95-99. [doi: 10.5055/ajdm.2022.0423] [Medline: 36494880]
56. Aitsi-Selmi A, Murray V. The sendai framework: disaster risk reduction through a health lens. Bull World Health Organ 2015;93(6):362 [FREE Full text] [doi: 10.2471/BLT.15.157362] [Medline: 26240454]
57. Lee J, Petrea-Imenokhoeva M, Naim H, Balk A, He S. Rapid deployment of telehealth in a conflict zone: supporting the humanitarian needs in Ukraine. NEJM Catalyst 2023;4(3):CAT-22. [doi: 10.1056/cat.22.0413]

Abbreviations

CAH: critical access hospital
ED: emergency department
EMS: emergency medical services

NEDI: National Emergency Department Inventory
RDHRS: Regional Disaster Health Response System
REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 10.11.22; peer-reviewed by S Fischer, K Ray, R Halkes, C Calyx; comments to author 23.02.23; revised version received 20.04.23; accepted 03.05.23; published 27.06.23.

Please cite as:
Boyle T, Boggs K, Gao J, McMahon M, Bedenbaugh R, Schmidt L, Zachrison KS, Goralnick E, Biddinger P, Camargo Jr CA
Hospital-Level Implementation Barriers, Facilitators, and Willingness to Use a New Regional Disaster Teleconsultation System: Cross-Sectional Survey Study
JMIR Public Health Surveill 2023;9:e44164
URL: https://publichealth.jmir.org/2023/1/e44164
doi:10.2196/44164
PMID:37368481

Original Paper

COVID-19 Vaccination Preferences Among Non-Chinese Migrants in Hong Kong: Discrete Choice Experiment

Saba Asim[1*], MBBS, MPH; Kailu Wang[2*], PhD; Elena Nichini[1], PhD; Faustina Fu Yip[1], MA; Liling Zhu[1], MSW; Hin Chung Eddy Fung[3], BSc; Yan Zeng[1], MA; Zhilan Fang[1], MA; Annie Wai-Ling Cheung[2], MPhil; Eliza Lai-yi Wong[2], PhD; Dong Dong[2], PhD; Eng-Kiong Yeoh[2], MBBS

[1]The Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Shatin, Hong Kong

[2]Centre for Health Systems and Policy Research, The Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Shatin, Hong Kong

[3]Faculty of Law, The Chinese University of Hong Kong, Shatin, Hong Kong

[*]these authors contributed equally

Corresponding Author:
Dong Dong, PhD
Centre for Health Systems and Policy Research
The Jockey Club School of Public Health and Primary Care
The Chinese University of Hong Kong
4/F School of Public Health Building
Prince of Wales Hospital
Shatin
Hong Kong
Phone: 852 22528461
Email: dongdong@cuhk.edu.hk

Abstract

Background: Studies have shown increasing COVID-19 vaccination hesitancy among migrant populations in certain settings compared to the general population. Hong Kong has a growing migrant population with diverse ethnic backgrounds. Apart from individual-level factors, little is known about the migrants' preference related to COVID-19 vaccines.

Objective: This study aims to investigate which COVID-19 vaccine–related attributes combined with individual factors may lead to vaccine acceptance or refusal among the migrant population in Hong Kong.

Methods: An online discrete choice experiment (DCE) was conducted among adults, including Chinese people, non-Chinese Asian migrants (South, Southeast and Northeast Asians), and non-Asian migrants (Europeans, Americans, and Africans) in Hong Kong from February 26 to April 26, 2021. The participants were recruited using quota sampling and sent a link to a web survey. The vaccination attributes included in 8 choice sets in each of the 4 blocks were vaccine brand, safety and efficacy, vaccine uptake by people around, professionals' recommendation, vaccination venue, and quarantine exemption for vaccinated travelers. A nested logistic model (NLM) and a latent-class logit (LCL) model were used for statistical analysis.

Results: A total of 208 (response rate 62.1%) migrant participants were included. Among the migrants, those with longer local residential years (n=31, 27.7%, for ≥10 years, n=7, 20.6%, for 7-9 years, n=2, 6.7%, for 4-6 years, and n=3, 9.7%, for ≤3 years; *P*=.03), lower education level (n=28, 28.3%, vs n=15, 13.9%, *P*=.01), and lower income (n=33, 25.2%, vs n=10, 13.2%, *P*=.04) were more likely to refuse COVID-19 vaccination irrespective of vaccination attributes. The BioNTech vaccine compared with Sinovac (adjusted odds ratio [AOR]=1.75, 95% CI 1.14-2.68), vaccine with 90% (AOR=1.44, 95% CI 1.09-1.91) and 70% efficacy (AOR=1.21, 95% CI 1.03-1.44) compared with 50% efficacy, vaccine with fewer serious adverse events (1/100,000 compared with 1/10,000; AOR=1.12, 95% CI 1.00-1.24), and quarantine exemption for cross-border travelers (AOR=1.14, 95% CI 1.01-1.30) were the vaccine attributes that could increase the likelihood of vaccination among migrants. For individual-level factors, full-time homemakers (AOR=0.44, 95% CI 0.29-0.66), those with chronic conditions (AOR=0.61, 95% CI 0.41-0.91) and more children, and those who frequently received vaccine-related information from the workplace (AOR=0.42, 95% CI 0.31-0.57) were found to be reluctant to accept the vaccine. Those with a higher income (AOR=1.79, 95% CI 1.26-2.52), those knowing anyone infected with COVID-19 (AOR=1.73, 95% CI 1.25-2.38), those having greater perceived susceptibility of COVID-19 infection (AOR=3.42, 95% CI 2.52-4.64), those who received the influenza vaccine (AOR=2.15, 95% CI 1.45-3.19), and those who frequently received information from social media (AOR=1.52, 95% CI 1.12-2.05) were more likely to accept the vaccine.

Conclusions: This study implies that migrants have COVID-19 vaccination preference heterogeneity and that more targeted and tailored approaches are needed to promote vaccine acceptance for different subgroups of the migrant population in Hong Kong. Vaccination promotion strategies are needed for low-education and low-income migrant groups, migrants with chronic diseases, the working migrant population, homemakers, and parents.

(JMIR Public Health Surveill 2023;9:e40587) doi:10.2196/40587

KEYWORDS

COVID-19 vaccination; migrants; discrete choice experiment; ethnic minorities; vaccine attributes; Hong Kong; COVID-19

Introduction

COVID-19 vaccination has been implemented worldwide as a crucial public health tool to reduce morbidity and mortality and eventually put an end to social distancing measures and the pandemic as a whole [1]. As of December 29, 2021, about 8.68 billion vaccine doses had been administered worldwide [2], first targeting priority groups and gradually extending vaccination to the general population.

For this campaign to be successful, public acceptance and the willingness to undergo vaccination are needed across the whole population [1,3]; yet research indicates that migrant populations in certain settings are particularly hesitant toward COVID-19 vaccines [4,5]. However, little is known about the migrants' preference related to COVID-19 vaccines [6]; this undermines health equity, since as vulnerable populations, migrants are at a higher risk of COVID-19 infection, hospitalization, and mortality compared to the general population [7].

Previous studies have mainly focused on individual-level factors. They have reported that migrants' socioeconomic conditions and gender have an impact on COVID-19 vaccine acceptance [3,8,9]. Ethnicity-related factors can influence vaccine hesitancy as well [4,9-11]. Moreover, the literature on migrant health suggests that migration history can differently affect access to primary medicine and vaccination uptake [12].

Finally, low vaccine coverage among migrants may be due to limited knowledge of the overall health care system [4]; in addition, during the pandemic, barriers within health care systems usually identified as possible factors affecting migrants' lower access to preventive care and vaccinations have increased [5,9]. Trust in the health care system or in vaccines may affect vaccine uptake as well [4,9,11,13].

Hong Kong is an international city with a growing migrant and ethnic population, which reached 8% of the overall population in 2016, representing a sharp increase of about 70% compared to 2006 [14]. Despite its facade of an international and inclusive city, Hong Kong holds a hegemonic Chinese culture where exclusion is widespread [15]; this may affect migrant population access and attitudes toward vaccines as well.

Although research has focused on individual variables, little is known about how vaccine-related attributes may influence migrants' attitudes toward COVID-19 vaccination. Similar to previous research conducted on the general population [16,17], this study aims to investigate COVID-19 vaccine preferences among the migrant population by identifying which individual factors combined with vaccine-related attributes may lead to vaccine acceptance or refusal. Although most of the literature focuses on the general population, this analysis will provide a more nuanced understanding of migrants' overall attitudes toward vaccines and guide public health efforts to promote COVID-19 vaccination among more diverse groups.

Methods

Study Sample and Data Collection

An online discrete choice experiment (DCE) was conducted among all adults, including Chinese people and non-Chinese migrants in Hong Kong, China, from February 26 to April 26, 2021, the first 2 months after the commencement of the citywide COVID-19 vaccination program of 2 major vaccines. Chinese or non-Chinese residents of Hong Kong aged 18 years or above were eligible to participate in the survey, while those who had a history of being diagnosed with COVID-19 or had received any COVID-19 vaccine were excluded. The participants were recruited using quota sampling for people with different ethnicities according to Hong Kong census data [18], including Chinese, non-Chinese Asians (South Asians, Southeast Asians and Northeast Asians), and non-Asians (Europeans, Americans, and Africans). A web link to the survey was sent to potential participants through the network established in a previous survey, through local nongovernmental organizations that provide social services to socially disadvantaged individuals and migrants, and through persons familiar with the migrant communities. The distribution of ethnicities in the sample was monitored online during the survey to adjust the sampling strategies for the remaining participants, and ceiling limits (quotas) were set for the number of participants with different ethnicities based on the ethnical distribution in the population. The survey was originally designed in English and then forward- and back-translated to other languages and thus was also available in Bahasa Indonesia, Nepali, Urdu, Thai, and traditional Chinese.

Experimental Design

A DCE requires participants to make choices from a series of choice sets described by a number of attributes. The attributes describing the COVID-19 vaccination plans included vaccine brand, efficacy, safety, uptake of vaccine by people around, recommendations from professionals, venue for vaccination, and exemption of quarantine for vaccinated travelers (Table 1), which were generated from prior individual interviews with 45

Hong Kong residents with diverse demographical characteristics and health conditions [19]. In this qualitative study, the confidence in and concerns about the benefits and side effects of the vaccines, vaccine origins and brands, recommendations from health care professionals, social influences of family members and friends, expectations of ease of travel restrictions, and logistics arrangement of vaccination (eg, locations of receiving the vaccines) were frequently reported as factors influencing the participants' willingness to accept the vaccines [19]. The levels of the attributes were determined based on the existing studies on vaccine effectiveness and safety [20-23] and the vaccination practice in Hong Kong [24]. For vaccine brands, all 3 (Sinovac, BioNtech, and AstraZeneca) were planned to be used for vaccination in Hong Kong at the time of the study, although only the former 2 of them were actually launched for the public eventually.

The full factorial design of the 7 attributes involved 648 combinations and 209,628 pairwise choice sets, so it was impossible to adopt the full factorial design in the DCE. The choice sets were designed using the D-optimality algorithm, and a total of 32 choice sets were generated. To further reduce the cognitive burden on participants, these choice sets were divided into 4 blocks, with 8 choice sets in each. The participants were randomly assigned to 1 of the 4 blocks: participants only provided their response to the choice sets in the block they were randomized to. In each choice set, participants were asked to choose 1 of the 3 choices ("vaccination plans"), 2 of which were choices to get a COVID-19 vaccine with different attributes, while the third choice was an opt-out option for accepting neither vaccination plan ("no vaccination"). An example of a choice set is shown in Table 2. The questionnaire was validated in the Chinese language. In total, 12 eligible adults were invited to a pilot survey for refinement of the translation in other languages.

Table 1. Attributes and levels for the DCE[a].

Attribute	Levels
Brand	• Sinovac • BioNTech • AstraZeneca
Probability of COVID-19 infection (efficacy)	• Reduce 50% infections • Reduce 70% infections • Reduce 90% infections
Probability of serious adverse event (safety)	• 1/10,000 • 1/100,000
Vaccine uptake by people around	• Nobody • Friends/colleagues • Family members
Recommendations from professionals	• Recommended by general physicians • Recommended by the government expert panel
Venue for vaccination	• Community hall • Health care facilities • Housing estate/workplace
Quarantine arrangement for vaccinated travelers	• At least 14-day compulsory quarantine • Exempted from the 14-day quarantine

[a] DCE: discrete choice experiment.

Table 2. Example of choice sets.

Attribute	Vaccination plan 1	Vaccination plan 2	Do not receive any vaccination
Brand	Sinovac	BioNTech	None
Probability of COVID-19 infection	Reduce 50% infections	Reduce 90% infections	No reduction
Probability of serious adverse event	1/100,000	1/10,000	No serious adverse event
Vaccine uptake by people around	Family members received the vaccine	No known people received the vaccine	N/A[a]
Recommendations from professionals	Government expert advisory panel	General physicians	N/A
Venue for vaccination	Housing estate/workplace	Community hall	N/A
Quarantine arrangement for vaccinated traveler	At least 14-day compulsory quarantine	Exempted from the 14-day quarantine	At least 14-day compulsory quarantine
Which vaccination plan would you choose?	Plan 1: ☐	Plan 2: ☐	Neither plan: ☐

[a] N/A: not applicable.

Measurement

In addition to measuring the participants' choice of COVID-19 vaccine via the DCE, the questionnaire also collected data on (1) experience, knowledge, and behaviors during the pandemic, such as whether the participant knew anyone infected with COVID-19, and the perceived susceptibility and severity of COVID-19 infection; (2) experience and perceptions related to vaccination, including previous uptake of influenza vaccination and information sources on COVID-19 vaccination; and (3) demographic and socioeconomic status of the participants, including their ethnicity, years of residence in Hong Kong, income level, education level, and employment status.

Ethical Considerations

The study was approved by the Survey and Behavioral Research Ethics Committee of the anonymous university (ref. no. SBRE-20-540). At the start of the survey, an electronic informed consent form was provided to the participants with details of the study purpose, data anonymity, and confidentiality. Those who agreed to join signed it electronically before moving to the questionnaire. Later, a supermarket coupon of HK $100 (US $12.74) was sent to each participant through mail. All the responses were anonymized and contained no personal information about any participant. The data were stored online and secured with passwords.

Statistical Analysis

To find out the preference of all participants for vaccination and its heterogeneity across ethnic groups, a latent-class logit (LCL) model was applied to the entire sample. LCL allocated participants in the survey sample with a similar preference for the vaccination attributes into the same latent class and provided estimates of the preference for each of the classes. The number of latent classes was determined based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) of the LCL model with the different number of classes. The difference in the preference between the Chinese and non-Chinese participants was explored. Subsequently, the analysis focused on the preference pattern within the non-Chinese migrants due to a substantial difference found between Chinese people and non-Chinese migrants, where the evidence for the latter group was limited.

For migrants, COVID-19 "vaccine refusal" irrespective of vaccine attributes was considered when a participant consistently chose "no vaccination" throughout all 8 choice sets. The vaccine refusal rate across different socialdemographic characteristics and the perception and experience during the COVID-19 pandemic were summarized using cross-tabulation and the chi-square test. To find out the influence of vaccination attributes, a nested logistic model (NLM) was adopted to simulate the choices over the vaccination plans in the DCE survey, which involved 2 decisions, namely (1) whether to accept a COVID-19 vaccine and (2) which COVID-19 vaccine to accept. NLM allowed estimation of the influences of both vaccination attributes and individual-level factors on the vaccine acceptance. The dependent variable was the binary choice (0="not choose", 1="choose") made for each of the alternatives in the DCE choice sets, and the independent variables were vaccination attributes and individual-level factors. The individual vaccine acceptance probability was estimated based on NLM outcomes [25,26]. In addition, sensitivity analysis was performed using the LCL model among migrant participants to find out whether there was any preference heterogeneity.

Results

Sample Characteristics and COVID-19 Vaccine Refusal

In total, 2892 Chinese people and non-Chinese migrants were invited, of which 2392 (82.7%) were eligible. Among them, 434 (18.1%) refused to participate in the survey and 261 (10.9%) did not complete the questionnaire; hence, 2032 (84.9%) valid responses were received.

Figure 1 shows the participant flow of the migrant group. A total of 462 people received the invitation, and 335 (72.5%) were eligible to the survey. Among them, 109 (32.5%) did not agree to participate and 18 (5.4%) did not complete the questionnaire; therefore, 208 (response rate 62.1%) participants were recruited. The characteristics of the participants can be found in Table 3. Of the 208 participants, 143 (68.8%) were female, and 67 (32.2%) were aged 18-29 years, 103 (49.5%) were aged 30-44 years, and 38 (18.3%) were aged 45 years or more. Most of them (n=180, 86.5%) reported their ethnicity as Asian. In addition, 108 (51.9%) of them attained a bachelor’s degree or above. Almost half of them (n=100, 48.1%) had a full-time job, and 33 (15.9%) were full-time homemakers. Furthermore, 76 (36.5%) of them had more than HK$ 30,000 (US $3821.68) as their monthly household income, which was the approximate median monthly household income in Hong Kong in 2020. Only 19 (9.1%) had chronic conditions. For years of residence, 112 (53.8%) had lived in Hong Kong for 10 years or more, while 31 (14.9%) of them had lived here for 3 years or less.

Regarding COVID-19 vaccine refusal irrespective of vaccination attributes (Table 3), participants with longer local residential years, lower education level, and lower income were more likely to refuse COVID-19 vaccination. With regard to ethnicity, Asian participants were more likely to refuse vaccination than non-Asian participants, although this difference was marginal. Full-time homemakers/housewives, those who had chronic conditions, those who did not know anyone having COVID-19, and those who did not receive influenza vaccines were more likely to refuse vaccination irrespective of the attributes, although the differences were not statistically significant. On the contrary, those with greater perceived susceptibility of COVID-19 infection were less likely to refuse vaccination.

Figure 1. Participant selection flowchart.

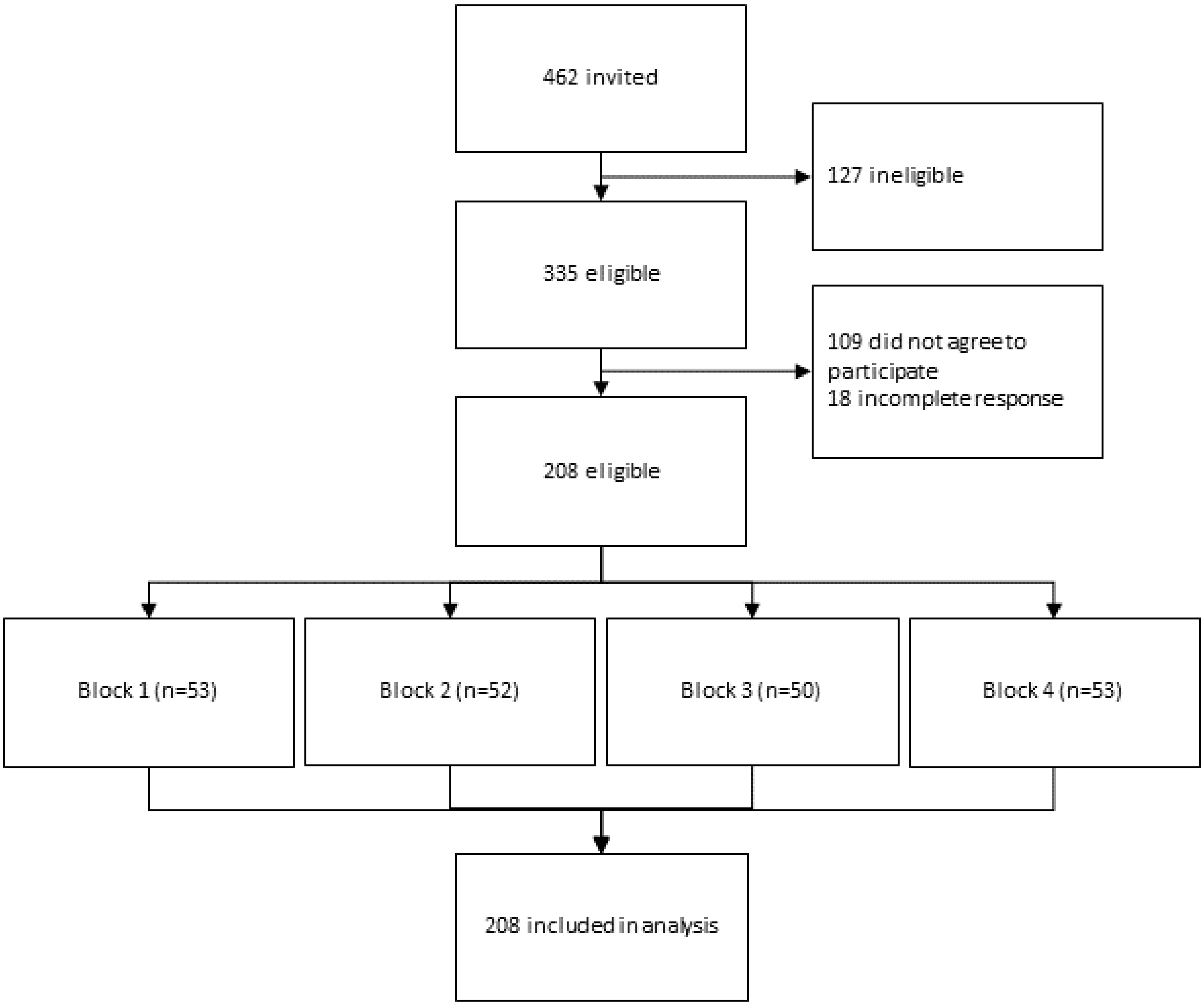


Table 3. Sample characteristics and COVID-19 vaccine refusal irrespective of vaccine attributes.

Characteristics	No refusal, n (%)	Refusal irrespective of attributes, n (%)	Migrants (N=208), n (%)	Total participant (N=2032), n (%)
Age (years), *P*=.26				
18-29	56 (83.6)	11 (16.4)	67 (32.2)	832 (40.9)
30-44	83 (80.6)	20 (19.4)	103 (49.5)	706 (34.7)
45-59	25 (69.4)	11 (30.6)	36 (17.3)	361 (17.8)
≥60	1 (50.0)	1 (50.0)	2 (1.0)	133 (6.6)
Sex, *P*=.84				
Male	51 (78.5)	14 (21.5)	65 (31.2)	769 (37.8)
Female	114 (79.7)	29 (20.3)	143 (68.8)	1263 (62.2)
Ethnicity[a], *P*=.06				
Asian	139 (77.2)	41 (22.8)	180 (86.5)	2004 (98.6)
Non-Asian	26 (92.9)	2 (7.1)	28 (13.5)	28 (1.4)
Years of residence in Hong Kong, *P*=.03				
≤3 years	28 (90.3)	3 (9.7)	31 (14.9)	N/A[b]
4-6 years	28 (93.3)	2 (6.7)	30 (14.4)	N/A
7-9 years	27 (79.4)	7 (20.6)	34 (16.3)	N/A
≥10 years	81 (72.3)	31 (27.7)	112 (53.8)	N/A
Missing	N/A	N/A	1 (0.4)	N/A
Education, *P*=.01				
Below bachelor's degree	71 (71.7)	28 (28.3)	99 (47.8)	923 (45.5)
Bachelor's degree or above	93 (86.1)	15 (13.9)	108 (52.2)	1108 (54.6)
Employment, *P*=.11				
Full-time	81 (81.0)	19 (19.0)	100 (48.1)	1353 (68.6)
Part time	28 (80.0)	7 (20.0)	35 (16.8)	
Unemployed	12 (80.0)	3 (20.0)	15 (7.2)	83 (4.2)
Students/interns	23 (92.0)	2 (8.0)	25 (12.0)	271 (13.7)
Full-time homemakers/housewives	21 (63.6)	12 (36.4)	33 (15.9)	175 (8.9)
Retired	N/A	N/A	N/A	91 (4.6)
Monthly household income (HK $/US $)[c], *P*=.04				
<30,000/<3821.68	98 (74.8)	33 (25.2)	131 (63.3)	953 (46.9)
≥30,000/≥3821.68	66 (86.8)	10 (13.2)	76 (36.7)	1078 (53.1)
Chronic condition, *P*=.07				
No	153 (81.0)	36 (19.1)	189 (90.9)	1802 (88.7)
Yes	12 (63.2)	7 (36.8)	19 (9.1)	230 (11.3)
Know anyone diagnosed with COVID-19, *P*=.07				
No	115 (76.2)	36 (23.8)	151 (72.6)	1676 (91.9)
Yes	50 (87.7)	7 (12.3)	57 (27.4)	356 (19.5)
Perceived "likely/very likely" to be infected, *P*<.001				
No	86 (69.9)	37 (30.1)	123 (59.1)	1163 (63.8)
Yes	79 (92.9)	6 (7.1)	85 (40.9)	869 (47.6)
Perceived "slightly severe/very severe" if get infected COVID-19, *P*=.82				
No	80 (80.0)	20 (20.0)	100 (48.1)	978 (53.6)

Characteristics	No refusal, n (%)	Refusal irrespective of attributes, n (%)	Migrants (N=208), n (%)	Total participant (N=2032), n (%)
Yes	85 (78.7)	23 (21.3)	108 (51.9)	1054 (57.8)
Previous influenza vaccination, *P*=.44				
No	138 (78.4)	38 (21.6)	176 (84.6)	1636 (89.7)
Yes	27 (84.4)	5 (15.6)	32 (15.4)	396 (21.7)
Often received vaccine information from social media, *P*=.60				
No	77 (77.8)	22 (22.2)	99 (47.6)	956 (52.4)
Yes	88 (80.7)	21 (19.3)	109 (52.4)	1076 (59.0)
Often received vaccine information from family/friends, *P*=.06				
No	114 (83.2)	23 (16.8)	137 (65.9)	1237 (67.8)
Yes	51 (71.8)	20 (28.2)	71 (34.1)	795 (43.6)
Often received vaccine information from the workplace, *P*=.33				
No	94 (77.1)	28 (23.0)	122 (58.7)	1339 (73.4)
Yes	71 (82.6)	15 (17.4)	86 (41.3)	693 (38.0)
Often received vaccine information from the government's official source, *P*=.45				
No	97 (77.6)	28 (22.4)	125 (60.1)	1292 (70.8)
Yes	68 (81.9)	15 (18.1)	83 (39.9)	740 (40.6)
Total	165 (79.3)	43 (20.7)	208 (100.0)	1824 (100.0)

[a] Ethnicity groups included Asians (South Asians, Southeast Asians, and Northeast Asians) and non-Asians (Europeans, Americans, and Africans).

[b] N/A: not applicable.

[c] HK $1=US $0.13.

Influence of COVID-19 Vaccination Attributes on Vaccine Acceptance Among the Entire Sample

A model with 6 latent classes was applied to test the preference heterogeneity between Chinese people and non-Chinese migrants (Table 4). Using class 6 as a reference, the migrants were much more likely to be assigned to class 1 based on their preference, followed by class 2 and class 4. In class 1, vaccine brand was the most important attribute that affected their choices, where BioNtech was the most preferable brand, followed by vaccine efficacy, venue of vaccination, vaccine uptake by others, and quarantine arrangement for vaccinated travelers. In class 2, vaccine brand, efficacy, safety, and quarantine arrangement for vaccinated travelers shared a similar level of perceived importance based on the respondents' choices, while vaccination at health care facilities and vaccination at the housing estate/workplace were preferable options as well. In class 4, higher vaccine efficacy was more important than the other attributes that lead to vaccine acceptance. The different preference pattern across classes 1, 2, and 4 implied that preference heterogeneity may be found within the migrant group, which is reported in the following sections. Compared to classes 3, 5, and 6, respondents in classes 1, 2, and 4 were more likely to be affected by the venue of vaccination. In class 3, the participants were efficacy oriented, preferred BioNtech and AstraZeneca, and were more likely to be influenced by general physicians with regard to vaccination decisions. In class 5, the participants preferred Sinovac, were more safety oriented, and presented a lower tendency to refuse vaccination than the other classes. Class 6 had the largest class share, where the participants shared similar preferences with participants in class 3 but were not influenced by recommendations from different health care professionals and vaccine uptake by their family members.

Table 4. Preference heterogeneity for COVID-19 vaccination plans in both Chinese people and non-Chinese migrants.

Attribute	Class 1, coefficient[a] (95% CI)	Class 2, coefficient[a] (95% CI)	Class 3, coefficient[a] (95% CI)	Class 4, coefficient[a] (95% CI)	Class 5, coefficient[a] (95% CI)	Class 6, coefficient[a] (95% CI)
Brand (reference=Sinovac)						
BioNtech	3.51[b] (2.97 to 4.06)	0.49[b] (0.33 to 0.65)	1.10[b] (0.87 to 1.32)	1.43[b] (1.05 to 1.81)	–3.59[b] (–4.12 to –3.07)	0.57[b] (0.01 to 1.12)
AstraZeneca	–0.29 (–0.73 to 0.16)	0.69[b] (0.52 to 0.85)	1.29[b] (1.06 to 1.53)	1.41[b] (0.94 to 1.89)	–4.04[b] (–4.58 to –3.50)	0.59[b] (0.03 to 1.15)
Efficacy (reference=50%)						
Reduce 70% infections	1.15[b] (0.83 to 1.48)	–0.03 (–0.16 to 0.10)	0.93[b] (0.69 to 1.16)	2.10[b] (1.73 to 2.48)	–0.01 (–0.44 to 0.41)	0.20 (–0.38 to 0.78)
Reduce 90% infections	0.56[b] (0.05 to 1.07)	0.74[b] (0.58 to 0.91)	2.32[b] (2.06 to 2.59)	3.43[b] (2.93 to 3.94)	0.30 (–0.14 to 0.75)	1.12[b] (0.58 to 1.65)
Serious adverse event (reference=1/10,000 people)						
1/100,000 people	–0.20 (–0.49 to 0.08)	0.62[b] (0.51 to 0.72)	0.64[b] (0.49 to 0.80)	–0.33[b] (–0.55 to –0.11)	0.33[b] (0.01 to 0.64)	0.72[b] (0.26 to 1.18)
Vaccine uptake by others (reference=no known people take the vaccine)						
Friends/colleagues received	0.58[b] (0.03 to 1.13)	–0.05 (–0.18 to 0.08)	0.33[b] (0.14 to 0.52)	1.17[b] (0.84 to 1.51)	0.64[b] (0.24 to 1.04)	0.73[b] (0.19 to 1.28)
Family members received	0.63[b] (0.17 to 1.09)	0.08 (–0.07 to 0.23)	0.42[b] (0.21 to 0.63)	0.97[b] (0.59 to 1.34)	0.48[b] (0.03 to 0.94)	0.41 (–0.21 to 1.03)
Recommendations from experts (reference=from general physicians)						
From expert advisory panel of the government	–0.13 (–0.58 to 0.31)	0.04 (–0.06 to 0.14)	–0.21[b] (–0.36 to –0.06)	0.82[b] (0.53 to 1.11)	0.10 (–0.21 to 0.41)	–0.13 (–0.58 to 0.33)
Venue for vaccination (reference=community center)						
Health care facilities	–1.00[b] (–1.51 to –0.50)	0.23[b] (0.09 to 0.36)	–0.09 (–0.28 to 0.09)	–0.63[b] (–0.96 to –0.29)	–0.08 (–0.45 to 0.28)	0.05 (–0.47 to 0.57)
Housing estate/workplace	0.08 (–0.33 to 0.50)	0.25[b] (0.12 to 0.39)	–0.04 (–0.22 to 0.15)	–0.40[b] (–0.69 to –0.10)	0.24 (–0.26 to 0.74)	–0.13 (–0.69 to 0.44)
Quarantine arrangement for vaccinated traveler (reference=compulsory quarantine required)						
14-day compulsory quarantine can be exempted	–0.54[b] (–0.98 to –0.10)	0.67[b] (0.55 to 0.79)	0.44[b] (0.28 to 0.61)	0.19 (–0.03 to 0.42)	–0.11 (–0.46 to 0.25)	0.80[b] (0.35 to 1.25)
Opt out (reference=not opt out)						
Opt out/no vaccination	5.07[b] (2.35 to 7.80)	1.70[b] (0.87 to 2.54)	10.07[b] (8.72 to 11.42)	9.95[b] (7.43 to 12.48)	–8.24[b] (–10.27 to –6.21)	10.67[b] (7.49 to 13.85)
Class size	10.7% (N/A[c])	12.2% (N/A)	16.4% (N/A)	21.5% (N/A)	5.8% (N/A)	33.4% (N/A)
Membership						
Age: 30-44 years	1.19[b] (0.74 to 1.64)	0.31 (0.00 to 0.63)	0.28 (–0.04 to 0.60)	0.58[b] (0.19 to 0.97)	2.10[b] (1.21 to 2.99)	Reference
Age: 45-60 years	1.57[b] (1.02 to 2.11)	1.07[b] (0.68 to 1.46)	–0.04 (–0.58 to 0.50)	1.39[b] (0.93 to 1.86)	3.24[b] (2.35 to 4.14)	Reference
Age: >60 years	1.80[b] (1.04 to 2.56)	1.05[b] (0.43 to 1.66)	–0.15 (–1.03 to 0.74)	1.31[b] (0.60 to 2.02)	3.78[b] (2.80 to 4.76)	Reference
Bachelor's degree or higher	0.01 (–0.39 to 0.41)	–0.08 (–0.38 to 0.22)	0.77[b] (0.43 to 1.11)	0.26 (–0.11 to 0.62)	–0.58[b] (–1.08 to –0.07)	Reference
Monthly income: >HK $30,000/US $3821.68[d]	0.02 (–0.38 to 0.42)	–0.02 (–0.32 to 0.28)	–0.34[b] (–0.66 to –0.03)	–0.03 (–0.39 to 0.33)	–0.30 (–0.76 to 0.16)	Reference

Attribute	Class 1, coefficient[a] (95% CI)	Class 2, coefficient[a] (95% CI)	Class 3, coefficient[a] (95% CI)	Class 4, coefficient[a] (95% CI)	Class 5, coefficient[a] (95% CI)	Class 6, coefficient[a] (95% CI)
Migrants	1.48[b] (0.99 to 1.97)	0.86[b] (0.41 to 1.32)	–0.65 (–1.37 to 0.07)	0.86[b] (0.34 to 1.38)	–1.55 (–3.32 to 0.23)	Reference

[a] Coefficient of attributes on utility estimated by the latent-class logit (LCL) model. The values of coefficients are weightings of the attribute levels, showing the magnitude of contributions from the attribute levels to the overall utility of the COVID-19 vaccine option. The higher the value is, the greater the contribution the corresponding attribute level makes to the overall utility, and hence the more important and preferable the attribute is.

[b] *P*<.05.

[c] N/A: not applicable.

[d] HK $1=US $0.13.

Influence of COVID-19 Vaccination Attributes on Vaccine Acceptance Among Migrants

Given that substantial preference heterogeneity was found between Chinese people and non-Chinese migrants, the preference of the migrants was further modeled separately. Among the COVID-19 vaccination attributes, brand, efficacy, safety (probability of serious adverse event), and the quarantine arrangement of vaccinated people could affect vaccine acceptance (Tables 5 and 6). Independent from efficacy and safety characteristics, participants were more likely to accept the BioNtech vaccine compared with Sinovac. The vaccine efficacy also made a substantial difference. A vaccine with 90% and 70% efficacy would increase the likelihood of vaccination by 44% and 21%, respectively. Fewer serious adverse events (1/100,000 vs 1/10,000) could also improve the likelihood of vaccination. Apart from brand, efficacy, and safety, if the quarantine for cross-border travelers can be exempted for those who received the vaccine, the likelihood of vaccine acceptance could increase by around 14%.

Table 5. Influence of COVID-19 vaccination attributes on vaccine acceptance in the NLM[a].

Attribute	Vaccine acceptance, AOR[b] (95% CI)
Brand (reference=Sinovac)	
BioNtech	1.75[c] (1.14-2.68)
AstraZeneca	1.11 (0.97-1.26)
Efficacy (reference=50%)	
70%	1.21[c] (1.03-1.44)
90%	1.44[c] (1.09-1.91)
Serious adverse event (reference=1/10,000)	
1/100,000	1.12[c] (1.00-1.24)
Vaccine uptake by others (reference=no known people take the vaccine)	
Friends/colleagues received	1.11 (0.98-1.25)
Family members received	1.12 (0.97-1.28)
Recommendations from professionals (reference=general physicians)	
Government expert advisory panel	1.08 (0.99-1.17)
Venue for vaccination (reference=community center)	
Health care facilities	1.00 (0.91-1.10)
Housing estate/workplace	0.98 (0.90-1.07)
Quarantine arrangement for vaccinated traveler (reference=14-day quarantine)	
14-day compulsory quarantine can be exempted	1.14[c] (1.01-1.30)

[a] NLM: nested logistic model.

[b] AOR: adjusted odds ratio.

[c] *P*<.05.

Table 6. Influence of individual-level factors on vaccine acceptance in the NLM[a].

Attribute	Vaccine acceptance, AOR[b] (95% CI)
Age (years; reference=18-29 years)	
30-44	2.42[c] (1.60-3.64)
≥45	2.20[d] (1.34-3.60)
Female (reference=male)	1.17 (0.87-1.57)
Ethnicity (reference=Asian)	
Non-Asian	2.28 (0.49-10.64)
Years of local residence (reference=less than 3 years)	
4-6	0.97 (0.51-1.83)
7-9	0.24[c] (0.13-0.43)
≥10	0.41[c] (0.25-0.67)
Education (reference=below bachelor's degree)	
Bachelor's degree or above	0.76 (0.53-1.09)
Employment (reference=full-time)	
Part time	1.12 (0.77-1.63)
Unemployed	1.08 (0.64-1.81)
Students/interns	1.41 (0.81-2.43)
Full-time homemakers	0.44[c] (0.29-0.66)
Number of children (reference=0)	
1	0.52[d] (0.33-0.82)
2	0.53[d] (0.33-0.85)
≥3	0.36[c] (0.22-0.61)
Monthly household income>HK $30,000 (US $3821.68)	1.79[d] (1.26-2.52)
With any chronic condition	0.61[d] (0.41-0.91)
Know anyone diagnosed with COVID-19	1.73[d] (1.25-2.38)
Perceived "likely/very likely" to be infected	3.42[c] (2.52-4.64)
Perceived "slightly severe/very severe" if get infected COVID-19	0.90 (0.69-1.18)
Previous influenza vaccination	2.15[c] (1.45-3.19)
Often received vaccine information from social media	1.52[d] (1.12-2.05)
Often received vaccine information from the workplace	0.42[c] (0.31-0.57)
Often received vaccine information from family/friends	1.08 (0.80-1.46)
Often received vaccine information from the government's official source	1.27 (0.95-1.70)

[a] NLM: nested logistic model.

[b] AOR: adjusted odds ratio.

[c] *P*<.001.

[d] *P*<.05.

For individual-level factors, participants who had been in Hong Kong for more than 7 years were less likely to accept the COVID-19 vaccine compared with those who had been in Hong Kong for less than 3 years. Full-time homemakers, those with chronic conditions and more children, and those who frequently received vaccine-related information from the workplace were found to be reluctant to accept the vaccine. On the contrary, those with a higher income, those knowing anyone infected with COVID-19, those having greater perceived susceptibility of COVID-19 infection, those who received the influenza vaccine, and those who frequently received information from social media were more likely to accept the vaccine.

Based on the model outcomes, the individual probability of the acceptance of vaccines with different characteristics can be estimated [25,26]. According to the estimation, an Asian migrant aged 18-29 years has a 73.6% chance to accept the BioNtech vaccine with 90% efficacy and a 1/100,000 serious adverse event probability when the vaccine is available in a community center, recommended by the government, and received by family members and when the quarantine exemption is in place. The acceptance probability was estimated to be 48.6% for the Sinovac vaccine with 50% efficacy. A non-Asian aged 18-29 years was estimated to have an 87.1% and a 69.8% chance to accept BioNtech and Sinovac vaccines, respectively. The acceptance probability changes with increasing age. For a migrant aged 30-44 years, the acceptance probabilities were estimated to be 88.9% (BioNtech) and 73.0% (Sinovac) among Asian migrants and 95.1% (BioNtech) and 86.9% (Sinovac) among non-Asian migrants. The estimated probabilities were similar for those aged above 45 years, which were 88.2% (BioNtech) and 71.8% (Sinovac) among Asian migrants and 94.8% (BioNtech) and 86.1% (Sinovac) among non-Asian migrants.

Sensitivity Analysis Account for Preference Heterogeneity

Substantial preference heterogeneity was found using a 4-class LCL model (Table 7). The participants had a 50.2% probability of belonging to class 1, where they attached greater importance to brand, efficacy, and exemption of quarantine for vaccinated travelers than the other attributes, as found in the NLM model for the entire study sample, and it was considered the reference group for class membership as it had the largest class size. Compared to class 1 (the reference group), participants in class 2 (16.1% probability) were more likely to live in Hong Kong for 7-9 years and they valued safety and the venue for vaccination, in addition to the 3 attributes mentioned in class 1. They were also likely to refuse vaccination (ie, choosing “no vaccination”). Class 2 and class 4 shared similar characteristics, they were likely to have longer local living experience compared to class 1, while class 4 was slightly more likely to comprise full-time homemakers. In class 4 (22.7% probability), the participants did not have clear preferences for these vaccination attributes.

Preferences in class 3 were quite different from those in classes 1 and 2. In class 3 (11.0% probability), participants preferred AstraZeneca over Sinovac and BioNtech and vaccine uptake by family and friends/colleagues as well as recommendations from experts from the government panel could also improve their COVID-19 vaccine acceptance, while exemption of quarantine did not have an impact on it. Participants in class 3 were more likely to be non-Asians and much less likely to be full-time homemakers.

Table 7. Sensitivity analysis for preference heterogeneity for COVID-19 vaccination plans.

Attribute	Class 1, coefficient[a] (95% CI)	Class 2, coefficient[a] (95% CI)	Class 3, coefficient[a] (95% CI)	Class 4, coefficient[a] (95% CI)
Preference for brand (reference=Sinovac)				
BioNtech	1.61[b] (1.35 to 1.88)	2.13[b] (1.57 to 2.69)	–0.87 (–1.81 to 0.06)	–0.52 (–3.28 to 2.24)
AstraZeneca	0.36[c] (0.08 to 0.65)	–0.04 (–0.67 to 0.59)	1.21[c] (0.08 to 2.33)	0.94 (–1.69 to 3.57)
Efficacy (reference=50%)				
Reduce 70% infections	0.85[b] (0.57 to 1.12)	0.53[c] (0.01 to 1.06)	–1.52[c] (–2.78 to –0.27)	0.12 (–3.75 to 4.00)
Reduce 90% infections	1.15[b] (0.86 to 1.44)	0.62[c] (0.07 to 1.18)	0.63 (–0.13 to 1.39)	2.67 (–0.26 to 5.60)
Serious adverse event (reference=1/10,000 people)				
1/100,000 people	0.06 (–0.14 to 0.25)	0.68[c] (0.26 to 1.09)	1.54[c] (0.48 to 2.61)	–2.53 (–5.74 to 0.68)
Vaccine uptake by others (reference=no known people take the vaccine)				
Friends/colleagues received	0.13 (–0.15 to 0.41)	0.45 (–0.07 to 0.98)	1.11[c] (0.20 to 2.02)	1.55 (–1.85 to 4.94)
Family members received	0.22 (–0.09 to 0.53)	0.08 (–0.48 to 0.63)	1.63[c] (0.33 to 2.92)	2.85 (–1.20 to 6.90)
Recommendations from experts (reference=general physicians)				
From expert advisory panel of the government	0.19 (–0.01 to 0.39)	–0.39 (–0.83 to 0.04)	1.13[c] (0.36 to 1.90)	–0.37 (–2.29 to 1.55)
Venue for vaccination (reference=community center)				
Health care facilities	–0.16 (–0.46 to 0.13)	0.64[c] (0.12 to 1.17)	0.63 (–0.14 to 1.40)	0.95 (–1.36 to 3.26)
Housing estate/workplace	–0.03 (–0.30 to 0.25)	0.09 (–0.45 to 0.62)	0.38 (–0.37 to 1.12)	–0.16 (–3.08 to 2.76)
Quarantine arrangement for vaccinated traveler (reference=compulsory quarantine required)				
14-day compulsory quarantine can be exempted	0.33[c] (0.09 to 0.58)	0.73[c] (0.26 to 1.19)	–0.27 (–1.07 to 0.54)	0.58 (–1.97 to 3.13)
Opt out (reference=not opt out)				
Opt out/no vaccination	–5.96 (–274.19 to 262.27)	7.00[b] (4.16 to 9.85)	5.93[c] (0.67 to 11.19)	13.43 (–4.11 to 30.97)
Class size	50.2% (N/A[d])	16.1% (N/A)	11.0% (N/A)	22.7% (N/A)
Membership				
Age: 30-44 years	Reference	–0.93 (–1.92 to 0.06)	–1.23 (–2.46 to 0.00)	–0.35 (–1.27 to 0.56)
Age: ≥45 years	Reference	–0.73 (–2.04 to 0.59)	–0.74 (–2.17 to 0.69)	0.15 (–0.97 to 1.26)
Non-Asian	Reference	–0.78 (–2.42 to 0.85)	1.20[c] (0.05 to 2.36)	–0.89 (–2.26 to 0.48)
Residential years: 4-6 years	Reference	0.55 (–1.26 to 2.36)	0.79 (–1.26 to 2.84)	0.36 (–1.44 to 2.15)
Residential years: 7-9 years	Reference	1.71[c] (0.08 to 3.35)	0.21 (–2.11 to 2.54)	2.03[c] (0.48 to 3.57)
Residential years: ≥10 years	Reference	1.35 (–0.07 to 2.76)	1.36 (–0.28 to 3.00)	1.69[c] (0.33 to 3.05)
Full-time homemaker	Reference	0.48 (–0.75 to 1.70)	–80.06[b] (–102.11 to –58.01)	0.96[c] (0.00 to 1.91)

[a] Coefficient of attributes on utility estimated by the latent-class logit (LCL) model. The values of coefficients are weightings of the attribute levels, showing the magnitude of contributions from the attribute levels to the overall utility of the COVID-19 vaccine option. The higher the value is, the greater the contribution the corresponding attribute level makes to the overall utility, and hence the more important and preferable the attribute is.

[b] *P*<.001.

[c] *P*<.05.

[d] N/A: not applicable.

Discussion

Principal Findings

This study investigated both vaccine-related attributes and individual-level factors to fully understand migrants' COVID-19 vaccine preferences in Hong Kong. It also provided an insight into the heterogeneity of the vaccine preference of the migrant population by dividing them into 4 classes.

Attributes

Our DCE study found that brand, efficacy, and quarantine arrangement for vaccinated people to be the most relevant vaccine-related attributes for vaccine acceptance among the migrant population in Hong Kong. The heterogeneity analysis complemented these findings, as half of the participants belonging to class 1 made their preferences based on these attributes. In addition, participants in class 3 preferred AstraZeneca over Sinovac and BioNtech. This indicated that migrants are inclined to use the brand as a proxy of unobserved characteristics in decision-making for vaccine acceptance. Previous research conducted on the US general population similarly showed brand and, specifically, the country where the vaccine was developed as predictors of COVID-19 vaccine acceptance [17,27]. With regard to efficacy, our study showed that the higher the efficacy, the higher the willingness to get vaccinated against COVID-19; this is in line with what emerged from studies in the general and ethnic populations in the United States [27] and the United Kingdom [17], where most of these studies have been conducted. Furthermore, specific migrant and ethnic groups may be persuaded if vaccines are proved to reduce the risk of being infected with COVID-19, as shown by research conducted in the United Kingdom [11]. The exemption of quarantine for vaccinated travelers was a predictor of COVID-19 vaccine acceptancy; this result is understandable within the Hong Kong context, where, upon returning, unvaccinated residents undergo self-paid 21-day compulsory quarantine in designated hotels compared to only 7 days for those vaccinated [28]. Other studies did not mention the quarantine arrangement specifically but more generally confirmed the vaccine would be accepted if it enabled social and family life to return to normal [11]. In addition to Sinovac and efficacy, our study found that a lower likelihood of serious adverse events is an attribute for vaccine acceptance, particularly for class 2 and class 3 participants. Similarly, the literature showed that concerns about vaccine safety and side effects are a predictor of vaccine hesitancy in both general and migrant populations in the United Kingdom [6,10,11].

The vaccine uptake by friends, colleagues, and family members; information from family and friends; and recommendations from the government and health care professionals were insignificant for the entire population sample but significantly affected only class 3 participants' vaccine acceptance. These participants were mostly non-Asian. The existing literature suggests that such recommendations may be welcomed only in the presence of trust toward the government and the health care system [5,6]. In addition, American Hispanics in the United States, for instance, rely on their community network to make decisions about COVID-19 vaccines [29]. Similarly, migrants in the United Kingdom tend to look for information from peers, especially when the government and health care system are not trusted [5]. To promote COVID-19 vaccinations and improve their health care access and quality, efforts should be made to establish more trustworthy relationships between the local government and the health care system with all the migrant populations in Hong Kong.

In this study, based on vaccine attributes, we did not find significant preference differences between the migrant population and the general population. Instead, there was a substantial heterogeneity representing diversity in the preferences of the non-Chinese migrant population in Hong Kong. This implied that while designing COVID-19 vaccination promotion programs for migrants, this diversity should be considered and more targeted strategies should be used for different migrant subgroups. The venue of vaccination was the only significant factor for vaccine acceptance among migrants when compared to the general population, which is consistent with the literature that suggests an inconvenient location would influence hesitancy [30] and minimal travel or comfortable places would reduce hesitancy among undocumented migrants, asylum seekers, and refugees [6]. However, in the sensitivity analysis for preference heterogeneity among the migrant population, only a small proportion of participants in class 2 of our study had an attached preference with the venue of vaccination. These participants also shared characteristics with class 4 participants who did not have a clear preference for vaccine attributes. This means the choices made by these people were quite random, so they were probably overwhelmed by the choice tasks.

Individual-Level Factors

Our results also showed that irrespective of the vaccination attributes, longer residential years, lower education and income, and less perceived COVID-19 infection susceptibility were associated with increased COVID-19 vaccine refusal. These individual-level factors were also related to decreased COVID-19 vaccine acceptance among those whose preferences were linked to vaccine attributes. Participants with more than 7 years of migration history in Hong Kong were least likely to get vaccinated compared to those with a more recent migration history. This aligns with much of the literature on migrant health, where a longer history of migration and host country language proficiency are associated with acculturation in the host society and health care use similar to nationals [12]; vaccine hesitancy among Hong Kong local residents is in fact similarly widespread [31]. Therefore, effective policies and strategies that can boost vaccination acceptance among the broader public will aid in achieving high vaccination uptake among migrants as well. Previous research has also identified the educational level as a predictor of vaccine acceptance [9,11]. In line with most research, a higher income was also associated with higher vaccine acceptance, while a lower income was associated with vaccine resistance [10,27]. Similar to what emerged from the previous literature, knowing someone diagnosed with COVID-19 was a factor for vaccine acceptance [17] and so was high risk perception; conversely, low risk perception was associated with higher hesitancy [30].

In addition, our study also identified other individual-level factors that predict COVID-19 vaccine acceptance among individuals who preferred different vaccine programs. A previous influenza vaccination experience was relevant for COVID-19 vaccine acceptance; this is in line with studies conducted in migrant and general populations in other settings [8,10,17].

Our study also indicated that those participants who often received information from social media were more likely to accept COVID-19 vaccination. Regarding the role of this source of information, the literature is mixed, with varying degrees of results; some studies suggest social media is used to solve confusion from contradicting information [5], while other research suggests it is a source of misinformation among migrants [4,6]. Our study also found that those who received information from the workplace were less likely to get vaccinated. This might be because, on the one hand, these people are employed and healthy and therefore have low risk perception, leading to low vaccine acceptance. On the other hand, working people are exposed to different information sources, including local ones; discussions with their colleagues may also increase their vaccination hesitancy, as there is a decreasing trend in the willingness toward COVID-19 vaccination among general working people in Hong Kong [32]. The promotion of vaccine-related information through this source should be encouraged.

Older age was identified as a factor of vaccine acceptance; specifically, participants aged 30-44 years would be the most likely to get vaccinated, while the youngest cohort (18-29 years old) would be the least likely to join the vaccination campaign. A study in the United Kingdom confirmed the young group (16-24 years old) to be the most hesitant [11]; this hesitancy may be motivated by low risk perception compared to other age groups [30]. COVID-19 vaccination promotion campaigns should target young age group specifically among the migrant population in Hong Kong.

In line with the international literature, the Hong Kong Centre for Health Protection encourages individuals with chronic conditions to get vaccinated against COVID-19, since they are at a higher risk of morbidity and mortality due to COVID-19 infection [33]; however, our study highlighted that people with chronic diseases are less likely to accept vaccines. This result suggests the need to plan COVID-19 vaccination promotion campaigns targeting this high-risk group specifically.

Being a homemaker and having children were indicators of COVID-19 vaccine hesitancy; in the United Kingdom, those living with children are similarly less likely to accept the vaccine [10]. This might be because migrants have a lack of family and social support and, being parents, they are main carers of their children. During COVID-19, the care burden of parents, particularly homemakers and mothers, has increased due to lockdown, school closure, and home schooling. They have busy routines and no time for making appointments and going for vaccination. Moreover, there might be concerns about who will take care of their children in case they have vaccine-related side effects. So, along with providing information about vaccine safety, an outreach approach can be considered for this particular migrant subgroup. Qualitative research is needed to further address the reasons behind this attitude.

In this study, gender was not a predictor of COVID-19 vaccine hesitancy; this is in contrast with previous research in other settings that identified women as being more hesitant compared to men both in the general and in the migrant population [8,10,11,17].

Limitations

Our study had some limitations that should be clarified when interpreting the results. First, the preference of respondents for COVID-19 vaccination is subject to change with time and could be different as the vaccination campaign progresses. It may also have been affected by reported adverse events following immunization for COVID-19, which were not captured in this survey. Second, although we incorporated “Asian”/“non-Asian” as a variable for ethnicity in the analysis (Tables 5 and 6), there were no significant difference in vaccine acceptance between these 2 groups, and due to the limited sample size of ethnicity subgroups and statistical power, we could not further subdivide the ethnicity in our analysis. Instead, the LCL model was used to identify the preference heterogeneity (Table 7), where non-Asian migrant respondents tended to attach more importance to the vaccination of family and friends and preferred AstraZeneca to the other 2 vaccine brands compared to Asian migrants. However, the ethnicity of the respondents could not be further divided in the analysis due to the limited subgroup sample size. Third, although the survey was piloted and refined prior to the formal investigation, there were around 20% respondents (class 4 in the LCL model) who may have been overwhelmed by the cognitive burden of the DCE: their COVID-19 vaccination choice appears to be random. By using the LCL model, we were able to separate them from other groups of respondents and avoid affecting the results.

Conclusion

In summary, this study provided insight into how different COVID-19 vaccine attributes, preferences, and individual-level characteristics of the migrant population could influence their COVID-19 vaccine acceptance. Although vaccine brand, efficacy, safety, and quarantine exemption were the most preferred attributes for the majority of the migrants, we reiterate that there is COVID-19 vaccination preference heterogeneity among migrants and that more targeted and tailored approaches are needed to promote vaccine acceptance for different subgroups of the migrant population in Hong Kong. The increased willingness toward COVID-19 vaccination in the general population will reduce vaccine hesitancy among migrants with long years of living experience in Hong Kong. Vaccination promotional programs for low-education and low-income migrant groups should be designed in different languages. The dissemination of COVID-19 vaccine information, particularly about efficacy and safety, through social media platforms can be continued, and information dissemination at the workplace for the working migrant population should also be considered. Detailed and tailored educational programs should be provided to people with chronic diseases to promote COVID-19 vaccination in this high-risk group. An outreach approach for people with children, particularly homemakers and mothers, can be used. Further follow-up studies, including qualitative interviews, can be conducted with these participants to explore in-depth the reasons behind their COVID-19 vaccination preferences and refusal.

Acknowledgments

SA, KW, EN, FFY, and DD managed conceptualization; SA, KW, EN, and DD, methodology; SA, EN, FFY, LZ, HCEF, YZ, and ZF, data collection; SA and KW, data analysis; DD and EL-YW, data curation; SA, KW, EN, and FFY, writing—original draft preparation; KW, visualization; DD, writing—review and editing; DD, EL-YW, and E-KY, supervision; and EN, LZ, and AW-LC, project administration. All authors have read and agreed to the published version of the manuscript.

Data Availability

All the data used in this study are available upon request from the corresponding author.

Conflicts of Interest

None declared.

References

1. Williams L, Gallant A, Rasmussen S, Brown Nicholls LA, Cogan N, Deakin K, et al. Towards intervention development to increase the uptake of COVID-19 vaccination among those at high risk: outlining evidence-based and theoretically informed future intervention content. Br J Health Psychol 2020 Nov;25(4):1039-1054 [FREE Full text] [doi: 10.1111/bjhp.12468] [Medline: 32889759]
2. WHO coronavirus (COVID-19) dashboard with vaccination data. World Health Organization. 2022. URL: https://covid19.who.int/ [accessed 2022-03-22]
3. Bell S, Clarke R, Mounier-Jack S, Walker J, Paterson P. Parents' and guardians' views on the acceptability of a future COVID-19 vaccine: a multi-methods study in England. Vaccine 2020 Nov;38(49):7789-7798 [FREE Full text] [doi: 10.1016/j.vaccine.2020.10.027]
4. Crawshaw AF, Deal A, Rustage K, Forster AS, Campos-Matos I, Vandrevala T, et al. What must be done to tackle vaccine hesitancy and barriers to COVID-19 vaccination in migrants? J Travel Med 2021 Jun 01;28(4):2021 [FREE Full text] [doi: 10.1093/jtm/taab048] [Medline: 33772312]
5. Knights F, Carter J, Deal A, Crawshaw AF, Hayward SE, Jones L, et al. Impact of COVID-19 on migrants' access to primary care and implications for vaccine roll-out: a national qualitative study. Br J Gen Pract 2021;71(709):e583-e595 [FREE Full text] [doi: 10.1101/2021.01.12.21249692]
6. Deal A, Hayward S, Huda M, Knights F, Crawshaw AF, Carter J, ESCMID Study Group for Infections in Travellers and Migrants (ESGITM). Strategies and action points to ensure equitable uptake of COVID-19 vaccinations: a national qualitative interview study to explore the views of undocumented migrants, asylum seekers, and refugees. J Migr Health 2021 May 27;4:100050 [FREE Full text] [doi: 10.1016/j.jmh.2021.100050] [Medline: 34075367]
7. COVID-19 immunization in refugees and migrants: principles and key considerations: interim guidance, 31 August 2021. World Health Organization. 2021 Aug 31. URL: https://www.who.int/publications/i/item/covid-19-immunization-in-refugees-and-migrants-principles-and-key-considerations-interim-guidance-31-august-2021 [accessed 2023-03-09]
8. Alabdulla M, Reagu SM, Al-Khal A, Elzain M, Jones R. COVID-19 vaccine hesitancy and attitudes in Qatar: a national cross-sectional survey of a migrant-majority population. Influenza Other Respir Viruses 2021 May;15(3):361-370 [FREE Full text] [doi: 10.1111/irv.12847] [Medline: 33605010]
9. Malik AA, McFadden SM, Elharake J, Omer SB. Determinants of COVID-19 vaccine acceptance in the US. EClinicalMedicine 2020 Sep;26:100495 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100495] [Medline: 32838242]
10. Paul E, Steptoe A, Fancourt D. Attitudes towards vaccines and intention to vaccinate against COVID-19: implications for public health communications. Lancet Reg Health Eur 2021 Feb;1:100012 [FREE Full text] [doi: 10.1016/j.lanepe.2020.100012] [Medline: 33954296]
11. Robertson E, Reeve KS, Niedzwiedz CL, Moore J, Blake M, Green M, et al. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. Brain Behav Immun 2021 May;94:41-50 [FREE Full text] [doi: 10.1016/j.bbi.2021.03.008] [Medline: 33713824]
12. Rosano A. Access to Primary Care and Preventative Health Services of Migrants. Cham: Springer; 2018.
13. Ferdinand KC, Nedunchezhian S, Reddy TK. The COVID-19 and influenza "Twindemic": barriers to influenza vaccination and potential acceptance of SARS-CoV2 vaccination in African Americans. J Natl Med Assoc 2020 Dec;112(6):681-687. [doi: 10.1016/j.jnma.2020.11.001] [Medline: 33276969]
14. Thematic report: ethnic minorities. Census and Statistics Department. 2017 Dec. URL: https://www.statistics.gov.hk/pub/B11201002016XXXXB0100.pdf [accessed 2023-03-09]

15. Crabtree S, Wong H. 'Ah Cha'! The racial discrimination of Pakistani minority communities in Hong Kong: an analysis of multiple, intersecting oppressions. Br J Soc Work 2012 Apr 02;43(5):945-963 [FREE Full text] [doi: 10.1093/bjsw/bcs026]
16. Dong D, Xu RH, Wong ELY, Hung MC, Feng D, Feng Z, et al. Public preference for COVID-19 vaccines in China: a discrete choice experiment. Health Expect 2020 Dec;23(6):1543-1578 [FREE Full text] [doi: 10.1111/hex.13140] [Medline: 33022806]
17. Kreps S, Prasad S, Brownstein JS, Hswen Y, Garibaldi BT, Zhang B, et al. Factors associated with US adults' likelihood of accepting COVID-19 vaccination. JAMA Netw Open 2020 Oct 01;3(10):e2025594 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.25594] [Medline: 33079199]
18. Thematic report: ethnic minorities. Census and Statistics Department. 2017 Dec. URL: https://www.statistics.gov.hk/pub/B11201002016XXXXB0100.pdf [accessed 2023-03-09]
19. Wong CH, Zhong CC, Chung VC, Nilsen P, Wong EL, Yeoh EK. Barriers and facilitators to receiving the COVID-19 vaccination and development of theoretically-informed implementation strategies for the public: qualitative study in Hong Kong. Vaccines (Basel) 2022 May 12;10(5):1 [FREE Full text] [doi: 10.3390/vaccines10050764] [Medline: 35632520]
20. 20 VM, Clemens S, Madhi S, Weckx L, Folegatti P, Aley P. Safetyefficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa,the UK. Lancet 2021;397:99-111 [FREE Full text] [doi: 10.3410/f.739172196.793581032]
21. Polack FP, Thomas S, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020;383:2603-2615 [FREE Full text] [doi: 10.1056/nejmoa2034577]
22. Zhang Y, Zeng G, Pan H, Li C, Hu Y, Chu K, et al. Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine in healthy adults aged 18–59 years: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial. Lancet Infect Dis 2021 Feb;21(2):181-192 [FREE Full text] [doi: 10.1016/s1473-3099(20)30843-4]
23. Wu Z, Hu Y, Xu M, Chen Z, Yang W, Jiang Z, et al. Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine (CoronaVac) in healthy adults aged 60 years and older: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial. Lancet Infect Dis 2021 Jun;21(6):803-812 [FREE Full text] [doi: 10.1016/s1473-3099(20)30987-7]
24. COVID-19 vaccination programme. Government of the Hong Kong Special Administrative Region. URL: https://www.covidvaccine.gov.hk/en/ [accessed 2023-02-13]
25. Craig BM. United States COVID-19 vaccination preferences (CVP): 2020 hindsight. Patient 2021 May;14(3):309-318 [FREE Full text] [doi: 10.1007/s40271-021-00508-0] [Medline: 33783724]
26. Wang K, Wong E, Cheung A, Chung V, Wong C, Dong D, et al. Impact of information framing and vaccination characteristics on parental COVID-19 vaccine acceptance for children: a discrete choice experiment. Eur J Pediatr 2022 Nov;181(11):3839-3849 [FREE Full text] [doi: 10.1007/s00431-022-04586-6] [Medline: 36056176]
27. Pogue K, Jensen JL, Stancil CK, Ferguson DG, Hughes SJ, Mello EJ, et al. Influences on attitudes regarding potential COVID-19 vaccination in the United States. Vaccines (Basel) 2020 Oct 03;8(4):582 [FREE Full text] [doi: 10.3390/vaccines8040582] [Medline: 33022917]
28. Inbound travel. Government of the Hong Kong Special Administrative Region. URL: https://www.coronavirus.gov.hk/eng/inbound-travel.html [accessed 2023-03-09]
29. Opel DJ, Lo B, Peek ME. Addressing mistrust about COVID-19 vaccines among patients of color. Ann Intern Med 2021 May;174(5):698-700 [FREE Full text] [doi: 10.7326/m21-0055]
30. Razai MS, Chaudhry U, Doerholt K, Bauld L, Majeed A. Covid-19 vaccination hesitancy. BMJ 2021 May 20;373:n1138 [FREE Full text] [doi: 10.1136/bmj.n1138] [Medline: 34016653]
31. Chan PKS, Wong M, Wong ELY. Vaccine hesitancy and COVID-19 vaccination in Hong Kong. Hong Kong Med J 2021 Apr;27(2):90-91 [FREE Full text] [doi: 10.12809/hkmj215115] [Medline: 33879624]
32. Wang K, Wong E, Ho K, Cheung A, Yau P, Dong D, et al. Change of willingness to accept covid-19 vaccine and reasons of vaccine hesitancy of working people at different waves of local epidemic in Hong Kong, China: repeated cross-sectional surveys. Vaccines (Basel) 2021 Jan 18;9(1):1 [FREE Full text] [doi: 10.3390/vaccines9010062] [Medline: 33477725]
33. Interim guidance notes on common medical diseases and COVID-19 vaccination in primary care settings. Center for Health Protection. 2022 Sep 2. URL: https://www.covidvaccine.gov.hk/pdf/Guidance_Notes.pdf [accessed 2023-03-09]

Abbreviations

AOR: adjusted odds ratio
DCE: discrete choice experiment
LCL: latent-class logit
NLM: nested logistic model

Edited by A Mavragani, T Sanchez; submitted 30.06.22; peer-reviewed by O Rahaman, DMA Awal, Z Han; comments to author 06.01.23; revised version received 27.01.23; accepted 23.02.23; published 27.03.23.

Please cite as:
Asim S, Wang K, Nichini E, Yip FF, Zhu L, Fung HCE, Zeng Y, Fang Z, Cheung AWL, Wong ELY, Dong D, Yeoh EK
COVID-19 Vaccination Preferences Among Non-Chinese Migrants in Hong Kong: Discrete Choice Experiment
JMIR Public Health Surveill 2023;9:e40587
URL: https://publichealth.jmir.org/2023/1/e40587
doi:10.2196/40587
PMID:36848242

Original Paper

Using Natural Language Processing to Predict Fatal Drug Overdose From Autopsy Narrative Text: Algorithm Development and Validation Study

Leigh Anne Tang[1,2], MS; Jessica Korona-Bailey[2], MPH; Dimitrios Zaras[2], PhD; Allison Roberts[2], PhD; Sutapa Mukhopadhyay[2], PhD; Stephen Espy[2], MS; Colin G Walsh[1], MA, MD

[1]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States

[2]Office of Informatics and Analytics, Tennessee Department of Health, Nashville, TN, United States

Corresponding Author:
Leigh Anne Tang, MS
Department of Biomedical Informatics
Vanderbilt University Medical Center
2525 West End Avenue
Suite 1475
Nashville, TN, 37203
United States
Phone: 1 510 366 1022
Email: leigh.a.tang@vanderbilt.edu

Abstract

Background: Fatal drug overdose surveillance informs prevention but is often delayed because of autopsy report processing and death certificate coding. Autopsy reports contain narrative text describing scene evidence and medical history (similar to preliminary death scene investigation reports) and may serve as early data sources for identifying fatal drug overdoses. To facilitate timely fatal overdose reporting, natural language processing was applied to narrative texts from autopsies.

Objective: This study aimed to develop a natural language processing–based model that predicts the likelihood that an autopsy report narrative describes an accidental or undetermined fatal drug overdose.

Methods: Autopsy reports of all manners of death (2019-2021) were obtained from the Tennessee Office of the State Chief Medical Examiner. The text was extracted from autopsy reports (PDFs) using optical character recognition. Three common narrative text sections were identified, concatenated, and preprocessed (bag-of-words) using term frequency–inverse document frequency scoring. Logistic regression, support vector machine (SVM), random forest, and gradient boosted tree classifiers were developed and validated. Models were trained and calibrated using autopsies from 2019 to 2020 and tested using those from 2021. Model discrimination was evaluated using the area under the receiver operating characteristic, precision, recall, F_1-score, and F_2-score (prioritizes recall over precision). Calibration was performed using logistic regression (Platt scaling) and evaluated using the Spiegelhalter z test. Shapley additive explanations values were generated for models compatible with this method. In a post hoc subgroup analysis of the random forest classifier, model discrimination was evaluated by forensic center, race, age, sex, and education level.

Results: A total of 17,342 autopsies (n=5934, 34.22% cases) were used for model development and validation. The training set included 10,215 autopsies (n=3342, 32.72% cases), the calibration set included 538 autopsies (n=183, 34.01% cases), and the test set included 6589 autopsies (n=2409, 36.56% cases). The vocabulary set contained 4002 terms. All models showed excellent performance (area under the receiver operating characteristic ≥0.95, precision ≥0.94, recall ≥0.92, F_1-score ≥0.94, and F_2-score ≥0.92). The SVM and random forest classifiers achieved the highest F_2-scores (0.948 and 0.947, respectively). The logistic regression and random forest were calibrated (P=.95 and P=.85, respectively), whereas the SVM and gradient boosted tree classifiers were miscalibrated (P=.03 and P<.001, respectively). “Fentanyl” and “accident” had the highest Shapley additive explanations values. Post hoc subgroup analyses revealed lower F_2-scores for autopsies from forensic centers D and E. Lower F_2-score were observed for the American Indian, Asian, ≤14 years, and ≥65 years subgroups, but larger sample sizes are needed to validate these findings.

Conclusions: The random forest classifier may be suitable for identifying potential accidental and undetermined fatal overdose autopsies. Further validation studies should be conducted to ensure early detection of accidental and undetermined fatal drug overdoses across all subgroups.

(JMIR Public Health Surveill 2023;9:e45246) doi:10.2196/45246

KEYWORDS

fatal drug overdose; natural language processing; surveillance; Tennessee; State Unintentional Drug Overdose Reporting System; SUDORS

Introduction

Background

In recent years, incidences of fatal drug overdose have surged in the United States. During the COVID-19 pandemic, the per capita monthly overdose death rate increased by 60% in May 2020 compared with May 2019 [1]. At the state level, Tennessee recorded an increase of 98% from May 2019 to May 2020 and was surpassed by only West Virginia and Kentucky. Since the pandemic, the rate of increase in fatal overdoses has slowed; however, fatal drug overdoses remain a public health concern with a substantial economic burden [2,3]. Provisional national data estimated that 107,622 drug overdose deaths occurred in 2021, an increase of 15% from 2020 [2]. Meanwhile, Tennessee has continued to surpass national trends, recording 3814 overdose deaths in 2021, a 26% increase from 2020 [4]. This stark increase in overdose deaths calls for improved surveillance, intervention, and prevention work to confront the overdose crisis at the state level.

Fatal drug overdose surveillance provides critical information for prevention and intervention. Specifically, surveillance helps identify where fatal overdoses occur, what substances were involved, and whether those trends have changed over time, all of which help make response efforts more targeted and relevant [5]. Data timeliness enables more effective and immediate intervention. The Tennessee Department of Health (TDH) currently conducts fatal drug overdose surveillance through the State Unintentional Drug Overdose Reporting System (SUDORS). SUDORS captures the details associated with accidental and undetermined fatal overdoses using death certificates, death scene investigations, autopsies, toxicology reports, and prescription drug monitoring program data. SUDORS is nested within the National Violent Death Reporting System and is funded by the Overdose Data to Action grant from the Centers for Disease Control and Prevention [6].

Although SUDORS provides a wealth of knowledge pertaining to fatal drug overdoses, a delay of 6 months occurs from the time of death to data collection. The delay is needed for death certificate coding to be completed and autopsy reports to become available for abstraction [7,8]. Notably, recent efforts to facilitate rapid fatal drug overdose surveillance have leveraged natural language processing (NLP). Several studies have explored the application of NLP to social media data for more timely fatal drug overdose surveillance [9-11]. Although potentially available closer to real time, social media data are subject to other challenges such as selection bias (social media users vs nonusers), user privacy settings limiting access to posts, and observer effects altering user behaviors [12]. Other studies have applied NLP to more traditional data sources for surveillance. Using death certificates, some surveillance teams have identified drug overdose deaths from cause of death fields available before the final cause of death coding [13-15]. Other teams have used text from verbal autopsies, which describe interviews with witnesses or relatives, as opposed to forensic autopsies, which detail findings from extensive physical examinations and toxicology tests [16,17]. Text from forensic autopsy reports has been used for the automatic classification of causes of death unrelated to drug overdoses [18]. However, to the best of our knowledge, no studies have leveraged free text in forensic autopsy reports to specifically predict fatal overdoses. Autopsy narrative text is similar to scene evidence descriptions in medicolegal documents, which are available sooner than autopsies.

Objectives

Faster identification of fatal overdoses can facilitate timely prevention and response efforts. In partnership with a research team in Biomedical Informatics at Vanderbilt University Medical Center (VUMC), the TDH sought to use NLP to identify fatal drug overdose deaths using narrative text from forensic autopsy reports.

Methods

Tennessee Fatal Drug Overdose Data Abstraction Process

TDH operates within a decentralized medicolegal death investigation system. Each of the 95 counties reports to 1 of the 5 forensic centers, which send reports to the Tennessee Office of the State Chief Medical Examiner (OSCME). The OSCME manages a repository of autopsy reports from all the forensic centers. Using the International Classification of Diseases, Tenth Revision (ICD-10) codes and keywords in death certificates, the SUDORS team identifies potential fatal overdoses. The SUDORS team obtains access to relevant autopsy reports and manually reviews each report to abstract variables into a REDCap (Research Electronic Data Capture) database (hereafter referred to as the SUDORS database) [19,20]. The SUDORS database contains information such as basic demographics, location of death, cause of death, and death scene information. As part of routine work at the TDH, at least 2 SUDORS abstractors review each case to ensure accurate coding.

Data Sources

NLP models were developed using text from forensic autopsy reports describing all manners of death from 2019 to 2021. Autopsies included all fatal overdose deaths identified by the

SUDORS team (cases) and the remaining other-cause deaths (controls) in the OSCME repository from 2019 to 2021, as of August 31, 2022. Autopsies were in semistructured forms containing common headings (detailed below) and computer-entered text. Autopsies were received as faxes or scans (PDFs) and processed using Adobe Acrobat Pro optical character recognition (OCR; Adobe). The outcome of interest was whether a given autopsy described an accidental or undetermined (ie, could not be definitively declared as accidental) fatal drug overdose, as indicated by the SUDORS team in the SUDORS database.

Narrative sections were defined as free-text portions of the autopsy that included the death scene information and medical history. These sections were commonly preceded by headings such as "summary of case," "summary and interpretation," "summary and opinion," and "narrative summary." The SUDORS team deemed these narrative sections to be the most informative parts of the autopsy report for determining the cause of death. Other data elements in the autopsy report varied substantially in quality (scanned toxicology results from different laboratories) or contained information that was interpreted and summarized in narrative sections (internal examination findings, external examination findings, and toxicology results). As such, this approach was limited to narrative sections to improve generalizability to other data sources in the future, such as medical examiner scene investigation notes, which are similar in content to the autopsy narrative text. This study was conducted before medical examiner notes were made available to the SUDORS team.

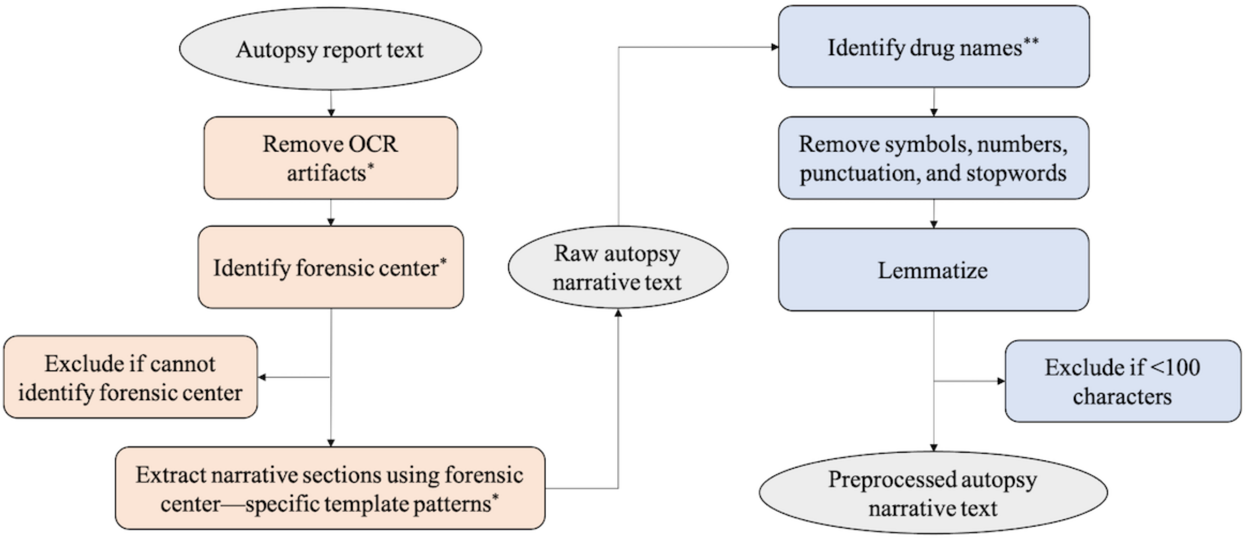
Preprocessing of Narrative Text

Autopsy PDFs were converted to a binary format and parsed page by page into text. Symbols that were likely artifacts of the OCR process (as determined by a manual review) were removed using regular expressions. The autopsy report templates were semistructured and unique to each forensic center. Template text patterns were identified by manual review and used to extract the forensic center issuing the autopsy, as well as relevant narrative sections. Autopsies for which a forensic center could not be identified using this templated language were removed from the data set.

To facilitate the comparison of information across report formats, narrative text was divided into three sections based on location within the autopsy report: (1) an initial narrative (beginning of report), (2) a case summary (middle of report and after internal and external examination findings), and (3) a summary of circumstances (end of report). Narrative sections within the same autopsy report were concatenated such that each report had only one combined narrative section.

The autopsy narrative text was extracted and preprocessed according to the steps outlined in Figure 1. A set of rules were developed using patterns found by manual review and enacted to remove additional artifacts resulting from the OCR process while ensuring the detection of drug names containing hyphens (eg, "4-ANPP" or 4-anilino-N-phenethylpiperidine) as one word rather than separate words ("4" and "ANPP"). Generalized regular expressions were used to identify potential references to drugs to account for novel drug combinations and varying OCR qualities. Words containing a number, hyphen, and letter were identified as potential drugs and combined ("4ANPP"). An exception to this rule was if the string contained "old" or "year," in which case the string likely referred to age rather than a drug ("35-year-old"). The numbers were removed if they were detected in words that were not identified as potential drug names. Punctuation and stop words (with the exception of "none" and "nowhere") were removed, and terms were lemmatized. Reports with fewer than 100 characters were likely not sufficiently processed via OCR and were thus removed from the data set.

Figure 1. Autopsy narrative text extraction (orange) and preprocessing (blue) steps. *Regular expression patterns determined by manual review with input from SUDORS team. **Words with a number, hyphen, and letter were combined (eg, "4-ANPP" or 4-anilino-N-phenethylpiperidine converted to "4ANPP"). An exception was if the string contained "old" or "year" (eg, "35-year-old"). OCR: optical character recognition; SUDORS: State Unintentional Drug Overdose Reporting System.



Modeling Plan

Autopsies from 2019 to 2020 were used for training (95%) and calibration (5%), and autopsies from 2021 were used for testing. This allowed model development to occur while the SUDORS team abstracted information from 2021 autopsies. Bag-of-words models were developed using term frequency–inverse document frequency scoring, which normalizes the word count in a given document based on the prevalence of the word among all documents, such that rarer words are given higher scores. Terms appearing in fewer than 20 documents in the training set or with fewer than 4 characters were excluded.

Logistic regression, support vector machine (SVM), random forest, and gradient boosted tree classifiers were generated. Hyperparameters for SVM, random forest, and gradient boosted trees were tuned with 5-fold cross-validation.

Model Metrics

The area under the receiver operating characteristic (AUROC), precision, recall, F_1-score (equal weighting of precision and recall), and F_2-score (recall weighted heavier than precision) were calculated for the corresponding hold-out data sets. Model error was evaluated using 100 bootstrapped samples, and 95% CIs were generated for each metric. Recall was identified as the most important metric that ensured that all potential cases of fatal overdoses would be reviewed by the surveillance team. Therefore, the F_2-score rather than F_1-score was primarily used to compare model performance.

The SVM, random forest, and gradient boosted tree classifiers achieving the highest F_2-scores were calibrated using Platt scaling, which applies logistic regression to normalize the predicted probabilities from the test set to the expected distribution of probabilities, as determined by the calibration set. The divergence between the predicted and expected probabilities of fatal overdoses was evaluated using the Spiegelhalter *z* test statistic. SHAP (Shapley additive explanations) values were generated for the random forest and gradient boosted tree classifiers [21].

Post Hoc Subgroup Analysis

The random forest classifier, which achieved the highest F_2-score and was calibrated, was used for post hoc subgroup analyses that assessed model performance within the test data set by forensic center, race, age at death, sex, and education level. Individuals in these subgroups were suspected to have varying levels of narrative text available for model prediction [22]. This analysis was also completed using the logistic regression model to ensure that the subgroup performances of the parametric (logistic regression) and nonparametric (random forest classifier) models were evaluated.

Subgroup information was extracted from the SUDORS database and death certificates (for non-SUDORS deaths or controls). In death certificates, race categories were mutually exclusive (ie, a decedent could only be coded as 1 race), whereas in the SUDORS database, more than 1 race could be indicated. To harmonize these data, race was coded as American Indian, Asian, Black, Other, Pacific Islander, or White. Individuals of multiple races were recoded as the most prevalent race within the population. For example, an individual coded as White (prevalence=0.73) and American Indian (prevalence=0.001) was recoded to White to reflect the more prevalent race in the data set. Age at death was recoded into ranges used by the National Center for Health Statistics: ≤14, 15-24, 25-34, 35-44, 45-54, 55-64, and ≥65 years [23]. The sex categories were female, male, and unknown. Education level was coded as 8th grade or less, 9th to 12th grade but no diploma, high school graduate or General Educational Development test, college but no degree, associate’s degree, bachelor’s degree, master’s degree, doctorate or professional degree, or unknown. Within each subgroup, AUROC, precision, recall, F_2-score, cases, controls, and median narrative text length were calculated.

Experimental Setup

This study was conducted in Python (version 3.9.7; Python Software Foundation) and R (version 4.2.0; R Foundation for Statistical Computing). Autopsy PDFs were converted to text using the *extract_text* function in the Python *pdfminer* package (version 20191125). Data preprocessing was performed using the *spaCy* package (version 3.2.4) and the *scispaCy* (version 0.5.0) *en_core_sci_sm* NLP pipeline, which was pretrained on biomedical data and had a vocabulary size of approximately 100K [24]. The models were constructed in Python using *scikit-learn* (version 1.1.1) and *xgboost* (version 1.6.1). The Spiegelhalter *z* test statistic was calculated using *rms* (version 6.3.0) in R, and the SHAP values were generated using SHAP (version 0.41.0) in Python.

Ethical Considerations

This study was a quality improvement project found exempt by the TDH institutional review board.

Results

Overview

An initial set of 17,521 autopsies was obtained for all manners of death from 2019 to 2021. A subset of autopsies (179/17,521, 1%) was removed because of unsuccessful identification of a specific forensic center or insufficient characters (<100 characters) after text preprocessing, including lemmatization. The final data set contained 17,342 autopsies (5912 cases and 11,430 controls). The most common manners of death were *accidental* among cases (5811/5912, 98%) and *natural* among controls (3994/11,430, 35%; Multimedia Appendix 1). Forensic centers were deidentified and denoted as A, B, C, D, and E. Approximately half (8743/17,342, 50.42%) of the total autopsies were obtained from the forensic center A (Table 1). The proportion of fatal overdoses was the highest among autopsy reports from the forensic center C (1229/3007, 40.87%).

Table 1. Cases and controls by forensic center (columns), entire data set.

	Forensic center, n (%)					Total (N=17,342), n (%)
	A (n=8743)	B (n=3717)	C (n=3007)	D (n=1452)	E (n=423)	
Cases	3095 (35.39)	1015 (27.3)	1229 (40.87)	417 (28.72)	156 (36.87)	5912 (34.09)
Controls	5648 (65.6)	2702 (72.69)	1778 (59.13)	1035 (71.28)	267 (63.12)	11,430 (65.91)

Narrative Text Sections

Across forensic centers, narrative text sections remained similar in content (each contained scene evidence and medical history) but had varying levels of missingness (Table 2). Notably, forensic centers A and B shared very similar formats, consisting of an initial narrative and a case summary. Autopsies from forensic centers C and D tended to have an initial narrative and summary of circumstances. Autopsies from forensic center E only presented a summary of circumstances.

Table 2. Sections of narrative text by forensic center in entire data set.

Forensic center	Initial narrative, n (%)	Case summary, n (%)	Summary of circumstances, n (%)	Total characters, median (IQR)
A (n=8743)	8675 (99.22)	8732 (99.87)	3 (0.03)	1287 (1053-1560)
B (n=3717)	3717 (100)	3687 (99.19)	0 (0)	1319 (1119-1566)
C (n=3007)	3007 (100)	0 (0)	2238 (74.43)	2852 (1855-4142)
D (n=1452)	1452 (100)	0 (0)	562 (38.71)	1074 (666-1511)
E (n=423)	0 (0)	0 (0)	423 (100)	462 (273-617)

Model Development

Of the 17,342 autopsies, a subset of 10,215 autopsies was used for training (3342/10,215, 32.72% cases), 538 for calibration (183/538, 34.01% cases), and 6589 for testing (2409/6589, 36.56% cases). A total of 4002 terms from the training data comprised the vocabulary set for the bag-of-words models.

Model Validation

All models achieved AUROC≥0.95, precision≥0.94, recall≥0.92, F_1-score ≥0.94, and F_2-score ≥0.92 (Table 3). The highest recall and F_2-scores were achieved using the SVM (recall=0.948, 95% CI 0.937-0.959; F_2-score =0.948, 95% CI 0.94-0.957) and random forest (recall=0.942, 95% CI 0.933-0.954; F_2-score =0.947, 95% CI 0.939-0.956). The corresponding 95% CIs for these metrics did not overlap with those for logistic regression, suggesting that SVM and random forest achieved significantly higher recall and F_2-scores than logistic regression ($P<.05$). The logistic regression and random forest were calibrated ($P=.95$ and $P=.85$, respectively). The SVM and gradient boosted tree classifiers were miscalibrated ($P=.03$ and $P<.001$, respectively).

The SHAP value analysis indicated that “fentanyl,” “accident,” “toxicity,” and “combined” were among the top features for the random forest (Figure 2) and gradient boosted trees classifiers (Multimedia Appendix 2). “Natural” was most strongly associated with controls.

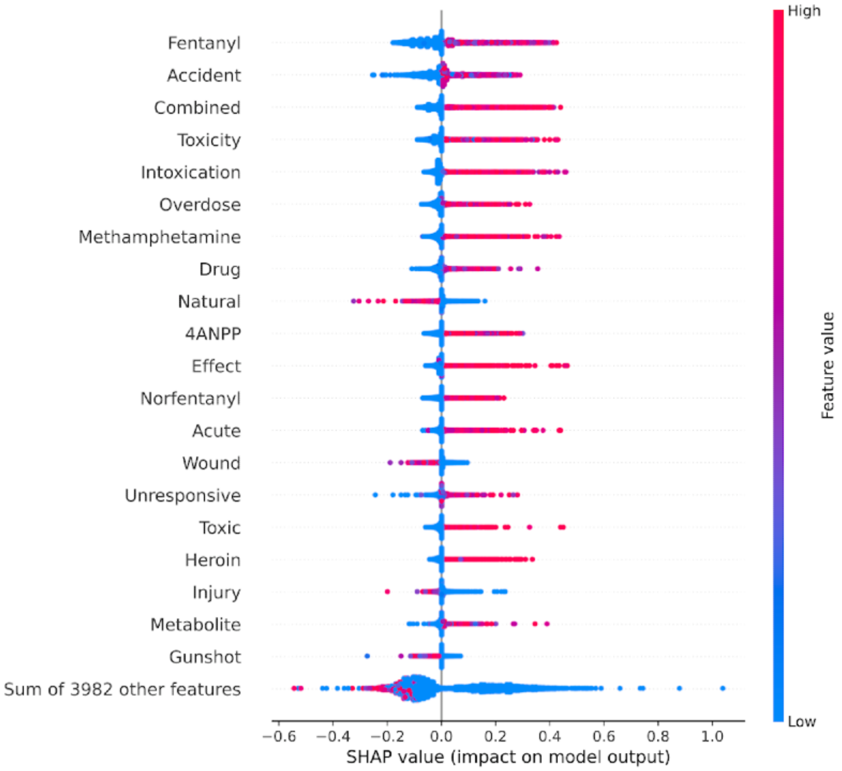
Table 3. Model discrimination and classification metrics across 100 bootstrapped samples.

	Area under the receiver operating characteristic, median (95% CI)	Precision, median (95% CI)	Recall, median (95% CI)	F_1-score, median (95% CI)	F_2-score[a], median (95% CI)	P value[b]
Logistic regression	0.95 (0.943-0.955)	0.963 (0.954-0.972)	0.92 (0.908-0.928)	0.941 (0.933-0.946)	0.928 (0.918-0.935)	−0.057 (.95)
Support vector machine	0.96 (0.954-0.966)	0.95 (0.936-0.962)	0.948 (0.937-0.959)	0.949 (0.942-0.956)	0.948 (0.94-0.957)	2.17 (.03)
Random forest	0.961 (0.955-0.967)	0.964 (0.954-0.973)	0.942 (0.933-0.954)	0.954 (0.945-0.96)	0.947 (0.939-0.956)	−0.189 (.85)
Gradient boosted trees	0.954 (0.946-0.962)	0.949 (0.935-0.964)	0.937 (0.92-0.953)	0.943 (0.935-0.953)	0.939 (0.926-0.951)	−6.79 (<.001)

[a] F_2-score prioritizes maximizing recall over precision.

[b] Spiegelhalter *z* test statistic.

Figure 2. Top 20 SHAP (Shapley additive explanations) values for the random forest classifier. 4ANPP: 4-anilino-N-phenethylpiperidine.



Subgroup Analysis

The subgroup analysis revealed similar results for the random forest classifier (Table 4) and logistic regression model (Multimedia Appendix 3), with a few exceptions. Across forensic centers for both models, forensic centers A, B, and C had the highest F_2-scores, whereas forensic centers D and E had the lowest F_2-scores. Across race categories for random forest, F_2-score was highest for Pacific Islander (F_2-score =1) and lowest for American Indian (F_2-score =0.833) and Asian (F_2-score =0.862). The logistic regression model achieved similar metrics but performed better for the American Indian race category (F_2-score =1). Across age groups for the random forest, F_2-score was highest for age groups 15-24 years (F_2-score =0.973) and lowest for ≤14 years (F_2-score =0.6) and ≥65 years (F_2-score =0.789). The same relative order of F_2-scores was preserved for logistic regression, but performance was considerably worse for age group ≤14 years (F_2-score =0.227). Models performed similarly across sex and education levels. Notably, several subgroups included fewer than 20 cases or controls.

Table 4. Random forest discrimination and classification metrics, cases, controls, and narrative section length for each subgroup in the test set (n=6589).

Subgroup	Area under the receiver operating characteristic	Precision	Recall	F_2-score	Cases, n (%)	Controls, n (%)	Length, median (IQR)
Forensic center							
A	0.978	0.971	0.973	0.973	1192 (18.09)	2038 (30.93)	1256 (1040-1497)
B	0.975	0.966	0.966	0.966	471 (7.15)	969 (14.71)	1344 (1137-1554)
C	0.956	0.946	0.953	0.951	510 (7.74)	699 (10.61)	3621 (2462-6025)
D	0.876	0.968	0.764	0.798	195 (2.96)	407 (6.18)	737 (512-1074)
E	0.846	0.967	0.707	0.747	41 (0.62)	67 (1.02)	686 (590-794)
Race							
American Indian	0.9	1	0.8	0.833	<20 (<0.30)	<20 (<0.30)	1224 (858-1593)
Asian[a]	0.917	1	0.833	0.862	<20 (<0.30)	30 (0.46)	1361 (1178-1705)
Black	0.973	0.965	0.961	0.962	492 (7.47)	1173 (17.80)	1299 (1092-1581)
Other[b]	0.989	1	0.979	0.983	47 (0.71)	32 (0.49)	1254 (1031-1478)
Pacific Islander[c]	1	1	1	1	<20 (<0.30)	<20 (<0.30)	837 (612-1110)
White	0.959	0.963	0.942	0.946	1858 (28.20)	2936 (44.56)	1339 (1027-1877)
Age range (years)							
≤14	0.796	0.6	0.6	0.6	<20 (<0.30)	246 (3.73)	1603 (1218-2279)
15-24	0.982	0.985	0.97	0.973	198 (3.01)	492 (7.47)	1183 (981-1523)
25-34	0.964	0.971	0.958	0.961	692 (10.50)	682 (10.35)	1279 (1039-1693)
35-44	0.954	0.957	0.957	0.957	805 (12.22)	717 (10.88)	1362 (1087-1823)
45-54	0.956	0.962	0.934	0.94	458 (6.95)	753 (11.43)	1344 (1064-1748)
55-64	0.951	0.98	0.907	0.921	216 (3.28)	679 (10.31)	1359 (1040-1761)
≥65	0.882	0.871	0.771	0.789	35 (0.53)	611 (9.27)	1355 (1020-1796)
Sex							
Female	0.959	0.961	0.943	0.947	736 (11.17)	1164 (17.67)	1381 (1081-1859)
Male	0.964	0.965	0.947	0.951	1673 (25.39)	3016 (45.77)	1307 (1039-1711)
Education							
8th grade or less	0.94	0.932	0.896	0.903	77 (1.17)	330 (5.01)	1485 (1153, 2156)
9th-12th grade but no diploma	0.968	0.959	0.963	0.962	460 (6.98)	691 (10.49)	1304 (1035-1709)
High school graduate or General Educational Development	0.961	0.967	0.944	0.948	1232 (18.70)	1808 (27.44)	1314 (1031-1726)
Some college but no degree	0.965	0.969	0.948	0.953	368 (5.59)	591 (8.97)	1297 (1052-1664)
Associate’s degree	0.946	0.935	0.926	0.928	108 (1.64)	209 (3.17)	1338 (1049-1745)
Bachelor’s degree	0.969	0.977	0.944	0.950	89 (1.35)	309 (4.69)	1340 (1097-1700)
Master’s degree	1	1	1	1	<20 (<0.30)	85 (1.29)	1324 (1119-1849)
Doctorate or professional degree	0.989	0.75	1	0.938	<20 (<0.30)	44 (0.67)	1369 (1037-2200)
Unknown	0.964	1	0.929	0.942	56 (0.85)	113 (1.71)	1507 (1207-2389)

[a] *Asian* includes Asian Indian, Chinese, Filipino, Korean, Vietnamese, and Other Asian.

[b] *Other* includes Other Race and Unknown.

[c] *Pacific Islander* includes Guamanian or Chamorro, Samoan, and Other Pacific Islander.

Discussion

Principal Findings

Narrative text was used to construct bag-of-words models (term frequency–inverse document frequency scoring) that predicted whether a given autopsy described a fatal overdose. The structure and details of the scene varied considerably across autopsies. Despite this heterogeneity in autopsy data, the SVM, random forest, and gradient boosted tree classifiers all showed excellent performance. Both SVM and random forest achieved the highest recall and F_2-score, but only the random forest was calibrated.

The terms most strongly associated with fatal overdoses in both the random forest and gradient boosted tree models were aligned with known state-level trends. From 2015 to 2019, the rate of fatal drug overdoses involving fentanyl increased from 2.7 to 16.8 per 100,000 residents in Tennessee [25]. More than half of fentanyl overdoses involved additional substances in 2019. Unsurprisingly, “fentanyl,” “norfentanyl,” “4ANPP” (ie, “4-ANPP” before text preprocessing; a precursor to illicit fentanyl), and “combined” were among the top terms most strongly associated with fatal overdose autopsies. Overdoses have also been increasingly linked to methamphetamine and heroin, both of which are highly predictive of fatal overdoses [25]. This rise in fentanyl- and stimulant-related overdose deaths has also been observed at the national level [23,26].

Forensic center subgroup analysis revealed lower model discrimination on data from forensic centers D and E. These 2 forensic centers contributed the fewest autopsies. It is possible that the models were not given ample opportunities to learn from autopsies contributed by these 2 centers. Model performance may improve with the oversampling of autopsies from these 2 forensic centers. The subgroup analyses by race and age indicated lower F_2-scores for the American Indian and Asian race groups as well as the ≤14 and ≥65 years age groups. These race and age subgroup results differed slightly in the logistic regression analysis, suggesting that the small sample sizes for these subgroups and others warrant additional study.

Subgroup differences in model performance and narrative length are worth additional study, particularly considering the findings of Mezuk et al [22]. In their study, Mezuk et al [22] identified shorter National Violent Death Reporting System narrative lengths for suicide and undetermined deaths among individuals who were male, were older, achieved lower education, and were part of certain racial and ethnic minority groups, compared with their counterparts. These findings suggest that models that use narrative text may generate biased predictions for individuals in certain subgroups. A similar detailed analysis is beyond the scope of this study. Our data did not suggest a noticeable difference in median narrative length across races. However, an analysis of narrative length using the Wilcoxon rank-sum test indicated a significant difference in narrative section length between the Black and White race subgroups ($P<.001$; Multimedia Appendix 4). These shorter sections did not appear to affect model performance, but further studies are needed.

Previous studies have used more complex machine learning methods, such as deep learning on autopsy text, to predict the cause of death ICD-10 codes [14,15,27]. Although such models may incorporate more information to perform generalized cause of death prediction, our models were specific to fatal overdose prediction. More specialized models may be better suited to identify fatal overdoses when the evidence is less clear-cut. For instance, drugs discovered at the death scene may not necessarily be linked to the primary cause of death. Alternatively, a fatal overdose may be the primary cause of death, but the death certificate may only reflect a contributing comorbidity. Our model performance metrics suggested that more advanced machine learning methods were not needed for this prediction task. As indicated by the relatively small size of the vocabulary set in this study (4002 terms across 10,215 autopsies) and subsequently verified by manual review, the language across autopsy reports was fairly standardized and uniform. Perhaps certain combinations of terms (eg, “fentanyl,” “accident,” and “toxicity”) were highly predictive of fatal overdose. Another study used NLP to predict fatal overdoses from free-text fields in Kentucky death certificates [13]. Similar to our study, bag-of-words–based machine learning models were used. However, whereas Kentucky relied on death certificate fields (as few as 2 words), our study used a different data source: autopsy narrative text. These sections ranged from hundreds to thousands of characters in length and provided additional contextual information in the form of scene evidence and medical history.

This work was completed in an academic-public partnership between the TDH and VUMC, where the SUDORS team at the TDH provided domain knowledge, and the Walsh Lab at VUMC provided modeling expertise. The goal was to develop a model that can be run and maintained by the SUDORS team. To this end, a simple modeling pipeline has several advantages. Adobe Acrobat Pro is easy to operate and does not require computer programming skills. Therefore, the OCR process can remain consistent, despite any changes in personnel that may occur. Sections of narrative text were identified using keywords that appeared as headings in autopsy reports. These rules can be modified to ensure the continued extraction of the appropriate autopsy sections, should autopsy templates change in the future. The bag-of-words models were agnostic to word order and less context dependent compared with more complex models such as word2vec. The simplicity of the bag-of-words approach translated into increased robustness to the variable quality of text extracted using OCR. Together, these aspects of the modeling pipeline facilitated the handoff of the code to the TDH SUDORS team and may inform modeling practices in other states.

Limitations

This study has several limitations. The OCR results depended on the quality of the scanned autopsy reports. The OCR was less accurate when the reports were scanned crooked (as reported in the results, no more than 1% of all autopsies were removed for this reason). Sometimes, spaces would be missed (eg, “decedentwas” instead of “decedent was”), misspellings would be introduced, or the OCR would add symbols (often <4 characters long). These tendencies informed the decision to exclude terms appearing in fewer than 20 documents or those having fewer than 4 characters. Although the Adobe Acrobat Pro OCR was sufficient for this study, careful manual reviews should be conducted when applying this tool to different data sources. In addition, the text was extracted using a simple Python function that did not always maintain proper word order when the text spanned multiple columns. As word order was occasionally altered during the OCR text extraction process, the tested models did not consider negation. Although word-order–agnostic models were sufficient for the prediction task presented in this study, more precise text extraction methods may yield at least a modest improvement in recall. Other studies may assess the importance of considering negation in autopsy reports.

Fatal overdoses were identified by the TDH surveillance team through an automated process that involved ICD-10 codes and keyword searches of death certificates, followed by a manual review. Although this process has been improved and tested over the years, it depends on death certificate codes and cause of death text, both of which may be delayed. The NLP-based approach in this study has the potential to overcome this limitation, as it relies on autopsy narrative text (a proxy for medicolegal scene investigation documentation), which is often available much sooner and can provide more contextual information than death certificates. Furthermore, the Tennessee medicolegal death investigation system is decentralized, thus allowing a wide range of variability in the level of detail included in the autopsy narratives. Despite this variation, the content of the information conveyed remained consistent. Finally, this study involved only data from one state health department. Although the studied classifiers performed well, data from other states may require additional preprocessing steps and potentially model retraining.

Future Work

The absence of a standardized autopsy template for all forensic centers added complexity to this prediction task and likely contributes to the additional time spent by the surveillance team in manually abstracting information. The recent adoption of the Medicolegal Death Investigation Log for some jurisdictions throughout the state is a step toward standardization, as this platform offers a common set of data fields for users to complete, in addition to a centralized location for documents to be uploaded. The utility of Medicolegal Death Investigation Log data in NLP models should be investigated in the future.

Efforts to standardize autopsy reports across forensic centers could facilitate the development of NLP tools specifically trained on text from autopsy reports. Such tools may identify scene evidence and naloxone administration more accurately than tools designed for electronic health record data, as described by Harris et al [28]. In the future, more specialized NLP tools could facilitate the automated extraction of public health–related variables for more complex NLP tasks.

Systematic characterization of the differential coverage of information related to scene evidence and medical history conveyed in autopsies from each forensic center may be worthwhile. This analysis may reveal factors contributing to poorer model performance and inform the standardization of statewide autopsy reports. As public health surveillance becomes more automated, it is increasingly important to consider variations conferred by nonstandardized documentation practices.

Given the findings of Mezuk et al [22] related to narrative lengths across different subgroups, future studies might seek to understand such patterns using data from all-cause deaths. Mezuk et al [22] also noted that missingness (ie, the presence or absence of a medical examiner narrative or law enforcement report) was associated with different social variables, such as homeless status, education, race, and ethnicity. These findings from Mezuk et al [22] are worth considering in future work because of the potential downstream effects of predictive models leveraging such biased data. In the case of public health surveillance, the result may be underreporting of deaths in certain groups, thus limiting intervention in communities that may be the most in need. This is particularly important, given the existing disparities exacerbated by the COVID-19 pandemic, which contributed to disproportionately high increases in overdose mortality rates among Black and American Indian or Alaska Native individuals compared with White individuals [29,30]. There is an urgent need to develop predictive models that ensure the appropriate allocation of resources during interventions.

Operationalizing the random forest classifier may facilitate rapid fatal overdose reporting in the future. OCR was performed using the batch processing option accessible through the Adobe Acrobat Pro user interface. The random forest classifier should be easy to scale to larger data sets and reduce overfitting by leveraging ensemble learning methods. Swift reporting can aid in the development of faster prevention and response activities at the community level, which can help curtail the growing number of fatal drug overdoses in Tennessee. Deploying this model could enhance the capacity of TDH to conduct fatal drug overdose surveillance.

Conclusions

Narrative text from Tennessee autopsy reports was used to develop NLP models that predicted the likelihood that a given autopsy described a fatal overdose. Simple bag-of-words–based models were sufficient for identifying potential fatal overdoses for public health surveillance. Additional studies are needed to ensure that the random forest classifier facilitates timely fatal overdose reporting for individuals across all subgroups.

Acknowledgments

The authors acknowledge Molly Golladay and the Office of the State Chief Medical Examiner for providing access to the requested autopsy reports. This work was fully supported by the American Public Health Association (APHA) and Centers for Disease Control and Prevention Data Science Demonstration Project Award to the Tennessee Department of Health. All findings and conclusions reported here are those of the authors alone and do not necessarily represent the views of the Centers for Disease Control and Prevention or APHA. In addition, CGW was supported by grants U54HG012510, Wellcome Leap, FDA WO2006, R01 MH121455, R01 MH120122, R01 MH116269, R01 MH118233, and 5RM1HG009034. LAT was supported by grant T15 LM007450.

Data Availability

The data sets analyzed in this study contain protected health information and are not publicly available; however, aggregate data may be available from the Office of Informatics and Analytics at the Tennessee Department of Health upon reasonable request [31].

Authors' Contributions

LAT, JKB, DZ, AR, SM, SE, and CGW contributed to study conception and design. LAT, JKB, and DZ collected the data. LAT, JKB, DZ, SM, and CGW contributed to analysis and interpretation of results. LAT, JKB, DZ, SM, and CGW contributed to draft manuscript preparation. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Manners of death among cases and controls, as indicated in death certificates.
[DOCX File , 13 KB - publichealth_v9i1e45246_app1.docx]

Multimedia Appendix 2
Top 20 SHAP (Shapley additive explanations) values for gradient boosted tree classifier.
[DOCX File , 272 KB - publichealth_v9i1e45246_app2.docx]

Multimedia Appendix 3
Logistic regression discrimination and classification metrics for each subgroup in the test set (n=6589).
[DOCX File , 20 KB - publichealth_v9i1e45246_app3.docx]

Multimedia Appendix 4
Comparison of narrative section lengths by race subgroup in the entire data set (N=17,342).
[DOCX File , 14 KB - publichealth_v9i1e45246_app4.docx]

References

1. Friedman J, Akre S. COVID-19 and the drug overdose crisis: uncovering the deadliest months in the United States, January July 2020. Am J Public Health 2021 Jul;111(7):1284-1291. [doi: 10.2105/ajph.2021.306256]
2. U.S. overdose deaths in 2021 increased half as much as in 2020 – but are still up 15%. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/202205.htm [accessed 2022-09-09]
3. Florence C, Luo F, Rice K. The economic burden of opioid use disorder and fatal opioid overdose in the United States, 2017. Drug Alcohol Depend 2021 Jan 01;218:108350 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108350] [Medline: 33121867]
4. 2021 Tennessee drug overdose deaths. Tennessee Department of Health Office of Informatics and Analytics. URL: https://www.tn.gov/content/dam/tn/health/documents/pdo/2021_Tennessee_Drug_Overdose_Deaths.pdf [accessed 2022-11-14]
5. Hoots BE. Opioid overdose surveillance : improving data to inform action. Public Health Rep 2021 Nov 02;136(1_suppl):5S-8S [FREE Full text] [doi: 10.1177/00333549211020275] [Medline: 34726970]
6. Korona-Bailey JA, Nechuta S, Golladay M, Moses J, Bastasch O, Krishnaswami S. Characteristics of fatal opioid overdoses with stimulant involvement in Tennessee: a descriptive study using 2018 State Unintentional Drug Overdose Reporting System Data. Ann Epidemiol 2021 Jun;58:149-155. [doi: 10.1016/j.annepidem.2021.03.004] [Medline: 33744415]
7. Instructions for classifying multiple causes of death, 2019 – section I. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nvss/manuals/2019/2b-sectioni-2019.htm [accessed 2022-09-09]
8. Ruiz L, Posey BM, Neuilly M, Stohr MK, Hemmens C. Certifying death in the United States. J Forensic Sci 2018 Jul 16;63(4):1138-1145. [doi: 10.1111/1556-4029.13689] [Medline: 29143322]
9. Simpson SS, Adams N, Brugman CM, Conners TJ. Detecting novel and emerging drug terms using natural language processing: a social media corpus study. JMIR Public Health Surveill 2018 Jan 08;4(1):e2 [FREE Full text] [doi: 10.2196/publichealth.7726] [Medline: 29311050]

10. Anwar M, Khoury D, Aldridge AP, Parker SJ, Conway KP. Using twitter to surveil the opioid epidemic in North Carolina: an exploratory study. JMIR Public Health Surveill 2020 Jun 24;6(2):e17574 [FREE Full text] [doi: 10.2196/17574] [Medline: 32469322]
11. Cuomo R, Purushothaman V, Calac AJ, McMann T, Li Z, Mackey T. Estimating county-level overdose rates using opioid-related twitter data: interdisciplinary infodemiology study. JMIR Form Res 2023 Jan 25;7:e42162 [FREE Full text] [doi: 10.2196/42162] [Medline: 36548118]
12. Fung IC, Tse ZT, Fu K. The use of social media in public health surveillance. Western Pac Surveill Response J 2015;6(2):3-6 [FREE Full text] [doi: 10.5365/WPSAR.2015.6.1.019] [Medline: 26306208]
13. Ward PJ, Rock PJ, Slavova S, Young AM, Bunn TL, Kavuluru R. Enhancing timeliness of drug overdose mortality surveillance: a machine learning approach. PLoS One 2019;14(10):e0223318 [FREE Full text] [doi: 10.1371/journal.pone.0223318] [Medline: 31618226]
14. Duarte F, Martins B, Pinto CS, Silva MJ. Deep neural models for ICD-10 coding of death certificates and autopsy reports in free-text. J Biomed Inform 2018 Apr;80:64-77 [FREE Full text] [doi: 10.1016/j.jbi.2018.02.011] [Medline: 29496630]
15. Falissard L, Morgand C, Roussel S, Imbaud C, Ghosn W, Bounebache K, et al. A deep artificial neural network-based model for prediction of underlying cause of death from death certificates: algorithm development and validation. JMIR Med Inform 2020 Apr 28;8(4):e17125 [FREE Full text] [doi: 10.2196/17125] [Medline: 32343252]
16. Jeblee S, Gomes M, Jha P, Rudzicz F, Hirst G. Automatically determining cause of death from verbal autopsy narratives. BMC Med Inform Decis Mak 2019 Jul 09;19(1):127 [FREE Full text] [doi: 10.1186/s12911-019-0841-9] [Medline: 31288814]
17. Bailo P, Gibelli F, Ricci G, Sirignano A. Verbal autopsy as a tool for defining causes of death in specific healthcare contexts: study of applicability through a traditional literature review. Int J Environ Res Public Health 2022 Sep 17;19(18):11749 [FREE Full text] [doi: 10.3390/ijerph191811749] [Medline: 36142022]
18. Mujtaba G, Shuib L, Raj RG, Rajandram R, Shaikh K, Al-Garadi MA. Automatic ICD-10 multi-class classification of cause of death from plaintext autopsy reports through expert-driven feature selection. PLoS One 2017;12(2):e0170242 [FREE Full text] [doi: 10.1371/journal.pone.0170242] [Medline: 28166263]
19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
20. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
21. Lundberg S, Lee S. A unified approach to interpreting model predictions. In: Proceedings of the 31st Conference on Neural Information Processing Systems. 2017 Presented at: 31st Conference on Neural Information Processing Systems; Dec 4 - 9, 2017; Long Beach, CA, USA URL: https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
22. Mezuk B, Kalesnikava VA, Kim J, Ko TM, Collins C. Not discussed: inequalities in narrative text data for suicide deaths in the national violent death reporting system. PLoS One 2021 Jul 16;16(7):e0254417 [FREE Full text] [doi: 10.1371/journal.pone.0254417] [Medline: 34270588]
23. Hedegaard H, Miniño A, Spencer M, Warner M. Drug Overdose Deaths in the United States, 1999–2020. Centers for Disease Control and Prevention. 2021 Dec. URL: https://www.cdc.gov/nchs/products/databriefs/db428.htm [accessed 2022-11-30]
24. Neumann M, King D, Beltagy I, Ammar W. ScispaCy: fast and robust models for biomedical natural language processing. In: Proceedings of the 18th BioNLP Workshop and Shared Task. 2019 Presented at: 18th BioNLP Workshop and Shared Task; Aug 1, 2019; Florence, Italy. [doi: 10.18653/v1/w19-5034]
25. Office of Informatics and Analytics. Tennessee's Annual Overdose Report 2021: Report on Epidemiologic Data and Projects to Address the Overdose Epidemic. Nashville, TN: Office of Informatics and Analytics. Tennessee Department of Health; Mar 2021.
26. Ciccarone D. The rise of illicit fentanyls, stimulants and the fourth wave of the opioid overdose crisis. Curr Opin Psychiatry 2021 Jul 01;34(4):344-350 [FREE Full text] [doi: 10.1097/YCO.0000000000000717] [Medline: 33965972]
27. Koopman B, Karimi S, Nguyen A, McGuire R, Muscatello D, Kemp M, et al. Automatic classification of diseases from free-text death certificates for real-time surveillance. BMC Med Inform Decis Mak 2015 Jul 15;15:53 [FREE Full text] [doi: 10.1186/s12911-015-0174-2] [Medline: 26174442]
28. Harris DR, Eisinger C, Wang Y, Delcher C. Challenges and barriers in applying natural language processing to medical examiner notes from fatal opioid poisoning cases. Proc IEEE Int Conf Big Data 2020 Dec;2020:3727-3736 [FREE Full text] [doi: 10.1109/bigdata50022.2020.9378443] [Medline: 35282306]
29. Friedman JR, Hansen H. Evaluation of increases in drug overdose mortality rates in the US by race and ethnicity before and during the COVID-19 pandemic. JAMA Psychiatry 2022 Apr 01;79(4):379-381 [FREE Full text] [doi: 10.1001/jamapsychiatry.2022.0004] [Medline: 35234815]

30. Han B, Einstein EB, Jones CM, Cotto J, Compton WM, Volkow ND. Racial and ethnic disparities in drug overdose deaths in the US during the COVID-19 pandemic. JAMA Netw Open 2022 Sep 01;5(9):e2232314 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.32314] [Medline: 36125815]
31. TDH Data Request Form. Office of Informatics & Analytics, Tennessee Department of Health. URL: https://www.surveygizmo.com/s3/5819792/TDH-Data-Request-Form [accessed 2023-05-01]

Abbreviations

AUROC: area under the receiver operating characteristic
ICD-10: International Classification of Diseases, Tenth Revision
NLP: natural language processing
OCR: optical character recognition
OSCME: Office of the State Chief Medical Examiner
REDCap: Research Electronic Data Capture
SHAP: Shapley additive explanations
SUDORS: State Unintentional Drug Overdose Reporting System
SVM: support vector machine
TDH: Tennessee Department of Health
VUMC: Vanderbilt University Medical Center

Edited by A Mavragani, T Sanchez; submitted 21.12.22; peer-reviewed by N Some, T Scherr; comments to author 30.01.23; revised version received 17.02.23; accepted 07.03.23; published 19.05.23.

Please cite as:
Tang LA, Korona-Bailey J, Zaras D, Roberts A, Mukhopadhyay S, Espy S, Walsh CG
Using Natural Language Processing to Predict Fatal Drug Overdose From Autopsy Narrative Text: Algorithm Development and Validation Study
JMIR Public Health Surveill 2023;9:e45246
URL: https://publichealth.jmir.org/2023/1/e45246
doi:10.2196/45246
PMID:37204824

Original Paper

Influenza-Associated Excess Mortality by Age, Sex, and Subtype/Lineage: Population-Based Time-Series Study With a Distributed-Lag Nonlinear Model

Li Li[1*], PhD; Ze-Lin Yan[1*], BSc; Lei Luo[2*], MD; Wenhui Liu[2], MM; Zhou Yang[1], BSc; Chen Shi[1], BSc; Bo-Wen Ming[1], BSc; Jun Yang[3], PhD; Peihua Cao[1,4], PhD; Chun-Quan Ou[1], PhD

[1]State Key Laboratory of Organ Failure Research, Department of Biostatistics, School of Public Health, Southern Medical University, Guangzhou, China

[2]Guangzhou Center for Disease Control and Prevention, Guangzhou, China

[3]School of Public Health, Guanghzou Medical University, Guangzhou, China

[4]Clinical Research Center, Zhujiang Hospital, Southern Medical University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Chun-Quan Ou, PhD
State Key Laboratory of Organ Failure Research, Department of Biostatistics
School of Public Health
Southern Medical University
No1023, South Shatai Road, Baiyun District
Guangzhou, 510515
China
Phone: 86 20 61649461
Email: ouchunquan@hotmail.com

Abstract

Background: Accurate estimation of the influenza death burden is of great significance for influenza prevention and control. However, few studies have considered the short-term harvesting effects of influenza on mortality when estimating influenza-associated excess deaths by cause of death, age, sex, and subtype/lineage.

Objective: This study aimed to estimate the cause-, age-, and sex-specific excess mortality associated with influenza and its subtypes and lineages in Guangzhou from 2015 to 2018.

Methods: Distributed-lag nonlinear models were fitted to estimate the excess mortality related to influenza subtypes or lineages for different causes of death, age groups, and sex based on daily time-series data for mortality, influenza, and meteorological factors.

Results: A total of 199,777 death certificates were included in the study. The average annual influenza-associated excess mortality rate (EMR) was 25.06 (95% empirical CI [eCI] 19.85-30.16) per 100,000 persons; 7142 of 8791 (81.2%) deaths were due to respiratory or cardiovascular mortality (EMR 20.36, 95% eCI 16.75-23.74). Excess respiratory and cardiovascular deaths in people aged 60 to 79 years and those aged ≥80 years accounted for 32.9% (2346/7142) and 63.7% (4549/7142) of deaths, respectively. The male to female ratio (MFR) of excess death from respiratory diseases was 1.34 (95% CI 1.17-1.54), while the MFR for excess death from cardiovascular disease was 0.72 (95% CI 0.63-0.82). The average annual excess respiratory and cardiovascular mortality rates attributed to influenza A (H3N2), B/Yamagata, B/Victoria, and A (H1N1) were 8.47 (95% eCI 6.60-10.30), 5.81 (95% eCI 3.35-8.25), 3.68 (95% eCI 0.81-6.49), and 2.83 (95% eCI –1.26 to 6.71), respectively. Among these influenza subtypes/lineages, A (H3N2) had the highest excess respiratory and cardiovascular mortality rates for people aged 60 to 79 years (20.22, 95% eCI 14.56-25.63) and ≥80 years (180.15, 95% eCI 130.75-227.38), while younger people were more affected by A (H1N1), with an EMR of 1.29 (95% eCI 0.07-2.32). The mortality displacement of influenza A (H1N1), A (H3N2), and B/Yamagata was 2 to 5 days, but 5 to 13 days for B/Victoria.

Conclusions: Influenza was associated with substantial mortality in Guangzhou, occurring predominantly in the elderly, even after considering mortality displacement. The mortality burden of influenza B, particularly B/Yamagata, cannot be ignored. Contrasting sex differences were found in influenza-associated excess mortality from respiratory diseases and from cardiovascular diseases; the underlying mechanisms need to be investigated in future studies. Our findings can help us better understand the magnitude and time-course of the effect of influenza on mortality and inform targeted interventions for mitigating the influenza mortality burden, such as immunizations with quadrivalent vaccines (especially for older people), behavioral campaigns, and treatment strategies.

(JMIR Public Health Surveill 2023;9:e42530) doi:10.2196/42530

KEYWORDS

influenza; disease burden; distributed-lag nonlinear model; excess mortality; harvesting effects

Introduction

Seasonal influenza has been associated with a large number of deaths, both in China and globally [1-5]. An accurate gauge of the mortality burden of the seasonal influenza epidemic is of great importance for understanding the impact of influenza, as well as for formulating and adjusting corresponding prevention and control measures [6]. This has been challenging, as laboratory diagnoses of influenza are not routinely made, and it is difficult to distinguish influenza virus infections from other respiratory pathogen infections based on nonspecific clinical symptoms [7]. Additionally, older adults are more likely to die from influenza-triggered complications, but influenza is seldom recorded as the cause of death in the mortality registration system [8]. Therefore, laboratory-confirmed deaths are an underestimate of the mortality burden of influenza [6].

Many studies have assessed the mortality burden of influenza with statistical models [3,6-19]. There have been substantial variations in estimates of influenza-associated excess mortality and in analytical strategies, including the definition of the lag period between influenza activity and mortality [20]. Previous studies have assumed that the lag was zero weeks, months, or years or one week; correspondingly, a simple influenza activity proxy was often included in the model, that is, a proxy lag of 1 week was used [20]. Lytras et al [21] examined the delayed effect of influenza activity on all-cause mortality using a distributed-lag nonlinear model (DLNM), allowing a flexible lag-response relationship between influenza and death. Interestingly, the study detected a short-term harvesting effect between influenza activity and all-cause mortality. Whether such findings can be generalized to other cause-of-death groupings and locations with different circulating influenza strains, health care systems, age structures, and contact patterns needs to be further explored. In addition, previous studies seldom reported potential disparities in influenza-associated excess mortality due to influenza B lineage and sex [19,22].

Guangzhou, a subtropical city in the western Pacific region of southern China (location N 23°8 , E 113°17), is an international transportation hub with a permanent population of approximately 15 million (in 2018) and an area of 7434.4 km^2. Free influenza vaccination has been available for older residents since 2021 in Guangzhou; however, the vaccination coverage rate has been unsatisfactory overall in China [23]. Assessment of the mortality burden of influenza would provide necessary data for evaluating this program. The current study aimed to estimate excess mortality associated with influenza virus subtypes and lineages for different cause-of-death groupings in Guangzhou from 2015 to 2018 across different age groups and sexes using DLNMs.

Methods

Ethical Considerations

This study was approved by the Research Ethics Committee of Southern Medical University (NFYKDX-ER2022012). The need for informed consent was waived because the data were deidentified and aggregated.

Data Sources

Individual data for deaths occurring between January 5, 2015, and December 30, 2018, in Guangzhou were obtained from the Guangzhou Center for Disease Control and Prevention (CDC). Individual death information included the underlying cause of death, age at death, and sex. The daily number of deaths was aggregated. Here, we considered 5 cause-of-death groupings associated with influenza virus infection: all causes (International Statistical Classification of Diseases and Related Health Problems, 10th Revision [ICD-10] codes A00-Z99), respiratory and cardiovascular diseases (ICD-10 codes I00-I99 and J00-J99), respiratory disease (ICD-10 codes J00-J99), pneumonia and influenza (ICD-10 codes J10-J18), and cardiovascular disease (ICD-10 codes I00-I99).

The Guangzhou CDC also provided influenza surveillance data, including the weekly proportions of specimens testing positive for influenza A (H1N1), A (H3N2), B/Yamagata, and B/Victoria, and the weekly proportion of consultations for influenza-like illness (ILI), that is, a body temperature ≥38 °C with cough or sore throat, among outpatient visits at sentinel hospitals in Guangzhou.

Data on annual population size were obtained from the Public Security Bureau of Guangzhou Municipality. We collected data on daily mean temperature and relative humidity from the website of the China Meteorological Data Service Center [24].

Statistical Analysis

We defined a weekly influenza virus activity proxy by multiplying the weekly proportion of consultations for ILI and the proportion of specimens testing positive for different subtypes and lineages, and then multiplying the resulting value by 1000 [25,26]. Then, we converted the weekly proxy to a daily proxy with cubic smoothing splines with 1 *df* per week, determined with generalized cross-validation [21]. The daily population size was estimated by linear interpolation. Daily mean temperature and relative humidity data were used to calculate daily absolute humidity [17], which has been suggested to be a better predictor of influenza virus transmission and survival than relative humidity [27].

A quasi-Poisson regression model was applied to estimate the excess mortality associated with influenza for different cause-of-death groupings, age groups (<60 years, 60-79 years, ≥80 years, and all ages), and sex. The associations of death counts with influenza and temperature were determined using a DLNM [28], as follows:

where Y_t and μ_t were the observed and expected number of deaths on day t, respectively. ϕ is the overdispersion parameter. The logarithm of population (Pop_t), with a fixed regression coefficient of 1, was used as an offset. The influenza activity proxy variables for A (H1N1), A (H3N2), B/Victoria, and B/Yamagata on day t were $Proxy_{H1,t}$, $Proxy_{H3,t}$, $Proxy_{BV,t}$, and $Proxy_{BY,t}$, respectively. Cross-basis $T_{t,l}(.)$ was constructed for the activity proxy variable of each influenza subtype/lineage, assuming a linear relationship between influenza virus activity and population mortality [3,6,13,14,21]. Cross-basis $T_{t,l}(.)$ was constructed for daily temperature ($Temp_t$) on day t with a natural cubic spline with 4 *df* [29]. It should be noted that the lag-response associations for the temperature and influenza virus activity proxies were set up with natural cubic splines with 3 knots placed at equally spaced log values of lags of 30 days to fully capture possible short-term mortality displacement effects (Multimedia Appendix 1) [21,30].

In addition, we controlled the effect of absolute humidity (Hum_t) using a natural cubic spline with 4 *df* [22]. Additionally, we included a categorical variable for the day of the week (Dow_t) and an indicator variable for holidays ($Holiday_t$) in the model. Additional adjustments in the model included a natural cubic spline with 5 *df* for the day of the year (Doy_t; ie, 1-366) and a categorical variable for calendar year ($Year_t$), to control for seasonality and the time trend in death counts, respectively [21].

We present the relative risk (RR) and cumulative RR of death associated with an increase of 10 units in the influenza virus activity proxy relative to an influenza activity proxy of 0 across a lag of 0 to 30 days. The daily number of excess deaths related to influenza was estimated as the difference between the estimated number of deaths given the observed influenza virus activity and the estimate given that the influenza virus did not circulate (ie, the influenza virus activity proxy was 0) [22]. The excess mortality rates (EMRs) associated with influenza for different age groups were estimated by dividing the influenza-related excess death numbers by the corresponding population sizes and then multiplying by 100,000. It should be noted that the EMRs associated with influenza in this study were in units per 100,000 persons per year. We estimated the EMRs that incorporated the lagged effects of influenza on mortality over 0 to 30 days. Meanwhile, we applied Monte Carlo simulations to estimate the 95% empirical CI (95% eCI) of the EMRs, which accounted for autocorrelation using the Newey-West method [31-33].

To detect whether there was a sex difference in influenza-related EMRs, we divided the male EMRs by female EMRs to obtain the male-to-female excess mortality ratios (MFRs) and derived the corresponding 95% CIs using the delta method [22].

To check the robustness of the results, we conducted sensitivity analyses by (1) changing the maximum number of lag days for lag-response relationships between the death and activity proxies of each influenza subtype and lineage, as well as for the relationship between death and temperature; (2) changing the way we controlled the temporal trend in the mortality rate; (3) ignoring the classifications of influenza virus types and influenza B lineages; (4) applying different cross-basis matrices to temperature; and (5) including the influenza virus activity proxy and temperature at a lag of 7 days in the model (Multimedia Appendix 1). All analyses were performed in R (version 4.1.1; R Foundation for Statistical Computing).

Results

There were 199,777 death certificates (86,440 for women and 113,337 for men; 34,229 for the 0-to-59-year age group, 73,552 for the 60-to-79-year age group, and 91,996 for the ≥80-year age group) included in this analysis, of which 105,998 were respiratory/cardiovascular disease deaths, 28,528 were respiratory disease deaths, and 12,255 were pneumonia/influenza deaths. In Guangzhou from 2015 to 2018, influenza viruses circulated every year. In 2015 and 2017, influenza activity peaked in the summer, while in 2016 and 2018, influenza activity was greater in the winter and early spring. A similar pattern was observed for the excess respiratory/cardiovascular disease mortality associated with influenza (Figure 1).

It was estimated that 2198 all-cause deaths, 1786 respiratory/cardiovascular disease deaths, 825 respiratory disease deaths, and 340 pneumonia/influenza deaths were attributable to influenza annually; correspondingly, the annual influenza-associated excess all-cause, respiratory/cardiovascular disease, respiratory disease, and pneumonia/influenza mortality rates were 25.06 (95% eCI 19.85-30.16), 20.36 (95% eCI 16.75-23.74), 9.41 (95% eCI 7.78-10.90) and 3.88 (95% eCI 2.80-4.84) per 100,000 persons, respectively (Table 1). Excess deaths associated with influenza accounted for 4.4% (8791/199,777), 6.7% (7142/105,998), 11.6% (3301/28,528), and 11.1% (1361/12,255) of all-cause, respiratory/cardiovascular disease, respiratory disease, and pneumonia/influenza deaths, respectively. Influenza-associated excess deaths due to respiratory/cardiovascular disease, respiratory disease, and pneumonia/influenza accounted for 81.2% (7142/8791), 37.6% (3301/8791), and 15.5% (1361/8791) of excess all-cause deaths, respectively.

The excess mortality due to various underlying causes associated with influenza varied across age groups. The EMR for each cause of death associated with influenza in people aged ≥80 years was higher than that in people aged 60 to 79 years. However, it should be noted that only influenza-related excess mortality due to respiratory/cardiovascular disease was statistically significant among deaths from all different causes in the 0-to-59-year age group. The EMRs for respiratory/cardiovascular disease for people aged 60 to 79 years and ≥80 years were 45.21 (95% eCI 33.31-56.74) and 452.63 (95% eCI 357.46-544.47) per 100,000 persons, respectively, which was higher than the EMR for people aged 0 to 59 years (91.50, 95% eCI 4.79-165.41). The burden in these two age groups represented 32.9% (2346/7142) and 63.7% (4549/7142) of the excess respiratory/cardiovascular disease deaths attributed to influenza, respectively (Table 1).

Figure 1. Influenza activity and weekly influenza-associated excess respiratory and cardiovascular deaths in Guangzhou, China, from 2015 to 2018. (A) Influenza virus activity proxy (this proxy does not have units). (B) Weekly influenza-associated excess respiratory/cardiovascular deaths. The bars in B represent the estimates of excess influenza-related deaths per week and the shaded areas are the corresponding 95% empirical CIs. The influenza virus activity proxy was calculated by multiplying the weekly proportion of consultations for influenza-like illness and the proportion of specimens testing positive for different subtypes and lineages, and then multiplying the resulting value by 1000.

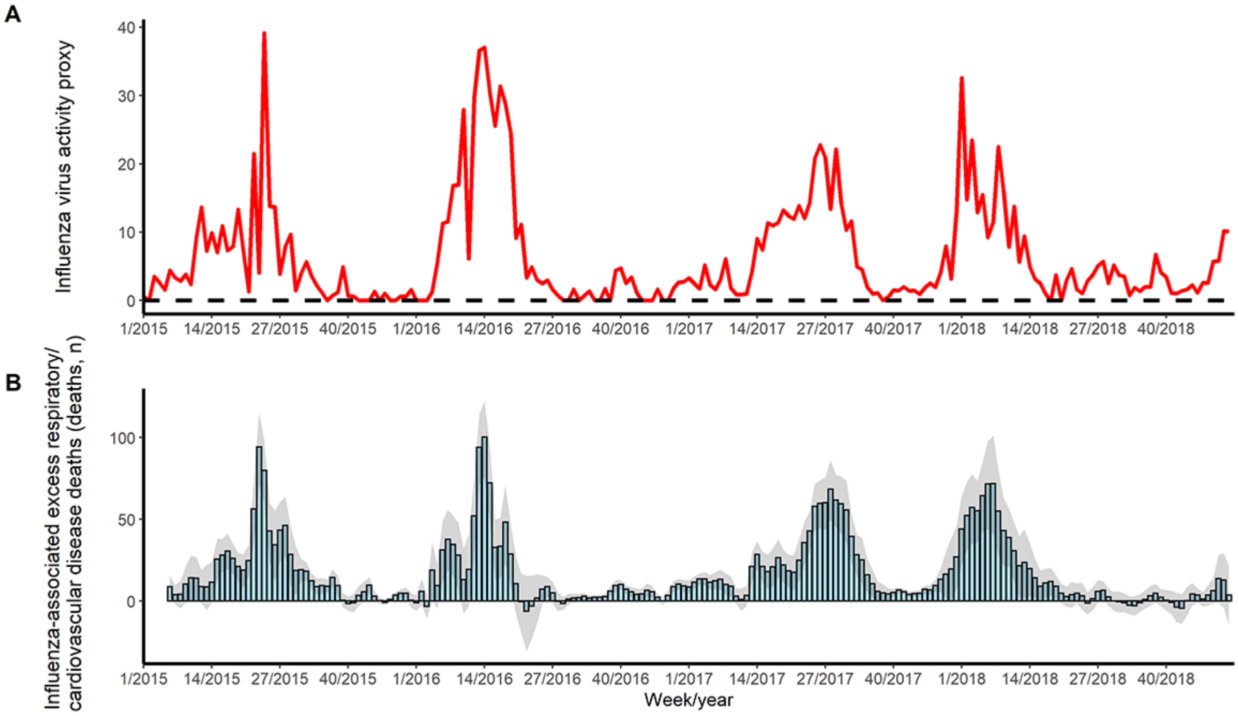


Table 1. Average annual excess mortality rates associated with influenza by cause-of-death grouping and age group in Guangzhou, China, from 2015 to 2018.

Age groups (years) and cause of death	Deaths, n (95% eCI[a])	Rate per 100,000 persons (95% eCI)
0-59		
All causes	64.05 (–105.30 to 226.41)	0.89 (–1.46 to 3.14)
Respiratory/cardiovascular disease	91.50 (4.79 to 165.41)	1.27 (0.07 to 2.29)
Respiratory disease	32.42 (–0.99 to 56.29)	0.45 (–0.01 to 0.78)
Pneumonia/influenza	16.80 (–13.76 to 35.73)	0.23 (–0.19 to 0.49)
Cardiovascular disease	55.53 (–23.09 to 122.85)	0.77 (–0.32 to 1.70)
60-79		
All causes	856.96 (621.47 to 1086.86)	66.06 (47.91 to 83.79)
Respiratory/cardiovascular disease	586.48 (432.06 to 735.96)	45.21 (33.31 to 56.74)
Respiratory disease	290.12 (223.20 to 344.15)	22.37 (17.21 to 26.53)
Pneumonia/influenza	103.51 (60.04 to 136.67)	7.98 (4.63 to 10.54)
Cardiovascular disease	283.58 (144.93 to 416.85)	21.86 (11.17 to 32.13)
≥80		
All causes	1305.53 (996.43 to 1596.4)	519.56 (396.54 to 635.31)
Respiratory/cardiovascular disease	1137.35 (898.22 to 1368.13)	452.63 (357.46 to 544.47)
Respiratory disease	509.81 (384.01 to 626)	202.89 (152.82 to 249.13)
Pneumonia/influenza	222.58 (140.17 to 293.36)	88.58 (55.78 to 116.75)
Cardiovascular disease	619.02 (433.25 to 796.09)	246.35 (172.42 to 316.82)
All		
All causes	2197.87 (1740.75 to 2644.78)	25.06 (19.85 to 30.16)
Respiratory/cardiovascular disease	1785.47 (1469.03 to 2082.24)	20.36 (16.75 to 23.74)
Respiratory disease	825.26 (682.51 to 955.75)	9.41 (7.78 to 10.90)
Pneumonia/influenza	340.34 (245.92 to 424.18)	3.88 (2.80 to 4.84)
Cardiovascular disease	940.06 (696.35 to 1171.28)	10.72 (7.94 to 13.36)

[a] eCI: empirical CI.

The difference in influenza-associated excess respiratory/cardiovascular disease mortality rate by sex was not statistically significant. However, the excess respiratory mortality was significantly higher in males than in females, with an MFR of 1.34 (95% CI 1.17-1.54). Conversely, the MFR of influenza-associated EMR for cardiovascular disease was 0.72 (95% CI 0.63-0.82; Table 2). Furthermore, the MFR varied by age group. For excess respiratory deaths related to influenza among those aged 60 to 79 and ≥80 years, the MFRs were 3.03 (95% CI 2.75-3.33) and 1.45 (95% CI 1.41-1.49), respectively, and the MFR was 0.45 (95% CI 0.23-0.89) for people aged 0 to 59 years (Multimedia Appendix 2). For people aged 0 to 59 and 60 to 79 years, the MFRs for excess cardiovascular deaths were 3.24 (95% CI 1.96-5.34) and 2.05 (95% CI 1.90-2.22), respectively. However, for influenza-associated excess cardiovascular deaths in people aged ≥80 years, the MFR was 0.65 (95% CI 0.64-0.67) (Multimedia Appendix 3).

Temporal trends of influenza virus activity varied by subtype and lineage (Figure 2). The EMR for respiratory/cardiovascular disease attributed to influenza A (H3N2) was 8.47 (95% eCI 6.60-10.30) per 100,000 persons, which was higher than the EMR due to influenza A (H1N1) (2.83, 95% eCI –1.26 to 6.71), B/Victoria (3.68, 95% eCI 0.81-6.49), and B/Yamagata (5.81, 95% eCI 3.35-8.25). For people aged 60 to 79 years and those aged ≥80 years, the excess respiratory/cardiovascular disease mortality associated with A (H3N2) was higher than that for other influenza subtypes or lineages. However, for people aged 0 to 59 years, excess influenza-related deaths were dominated by influenza A (H1N1), with excess mortality of 1.29 (95% eCI 0.07-2.32; Table 3).

Between 2015 and 2018, the predominant circulating influenza virus subtypes and lineages varied. Correspondingly, the EMRs per 100,000 persons for respiratory and cardiovascular disease associated with influenza A (H1N1), A (H3N2), B/Victoria, and B/Yamagata ranged between 0.21 (95% eCI –0.13 to 0.54) and 5.62 (95% eCI –2.74 to 13.19), between 0.15 (95% eCI 0.10-0.19) and 21.25 (95% eCI 16.61-25.59), between 0.32 (95% eCI 0.09-0.56) and 9.52 (95% eCI 2.58-16.29), and between 0.63 (95% eCI –0.62 to 1.77) and 10.84 (95% eCI 6.74-14.83), respectively (Figure 2).

The RR of respiratory/cardiovascular disease death associated with influenza virus subtypes/lineages changed with lag time. For influenza A (H1N1), the highest RR occurred on the current day, and there seemed to be a displacement effect during days 2 to 5. This effect lasted approximately 14 days. Similar patterns were observed in the RRs for influenza A (H3N2) and B/Yamagata. For B/Victoria, RR peaked at a lag of 3 days. The displacement effect was observed on days 5 to 13 (Figure 3).

Table 2. Average annual excess mortality associated with influenza by cause-of-death grouping and sex in Guangzhou, China, from 2015-2018.

Underlying cause	Male		Female		Male-to-female excess mortality ratio (95% CI)	*P* value
	Deaths, n (95% eCI[a])	Rate per 100,000 persons (95% eCI)	Deaths, n (95% eCI)	Rate per 100,000 persons (95% eCI)		
All causes	1060.79 (755.56-1358.70)	24.12 (17.18-30.89)	1132.64 (859.06-1385.33)	25.91 (19.65-31.69)	0.93 (0.86-1.01)	.09
Respiratory/ cardiovascular diseases	876.90 (674.64-1071.51)	19.94 (15.34-24.36)	906.37 (716.96-1086.57)	20.73 (16.40-24.85)	0.96 (0.88-1.06)	.41
Respiratory diseases	472.95 (362.98-571.77)	10.75 (8.25-13.00)	351.06 (263.27-428.41)	8.03 (6.02-9.80)	1.34 (1.17-1.54)	<.001
Pneumonia/ influenza	177.27 (109.82-232.72)	4.03 (2.50-5.29)	162.02 (100.26-216.15)	3.71 (2.29-4.94)	1.09 (0.88-1.34)	.46
Cardiovascular diseases	393.08 (227.15-553.88)	8.94 (5.16-12.59)	543.79 (375.64-705.39)	12.44 (8.59-16.14)	0.72 (0.63-0.82)	<.001

[a] eCI: empirical CI.

Figure 2. Annual influenza-associated excess respiratory/cardiovascular mortality rates and influenza activity in Guangzhou, China, from 2015 to 2018. (A) Annual influenza-associated excess respiratory/cardiovascular mortality rates per 100,000 persons by influenza subtype/lineage. (B) Influenza virus activity proxies by influenza subtype/lineage (this proxy does not have units). The dots in A indicate point estimates of excess respiratory/cardiovascular mortality, while vertical line segments indicate the corresponding 95% empirical CIs.

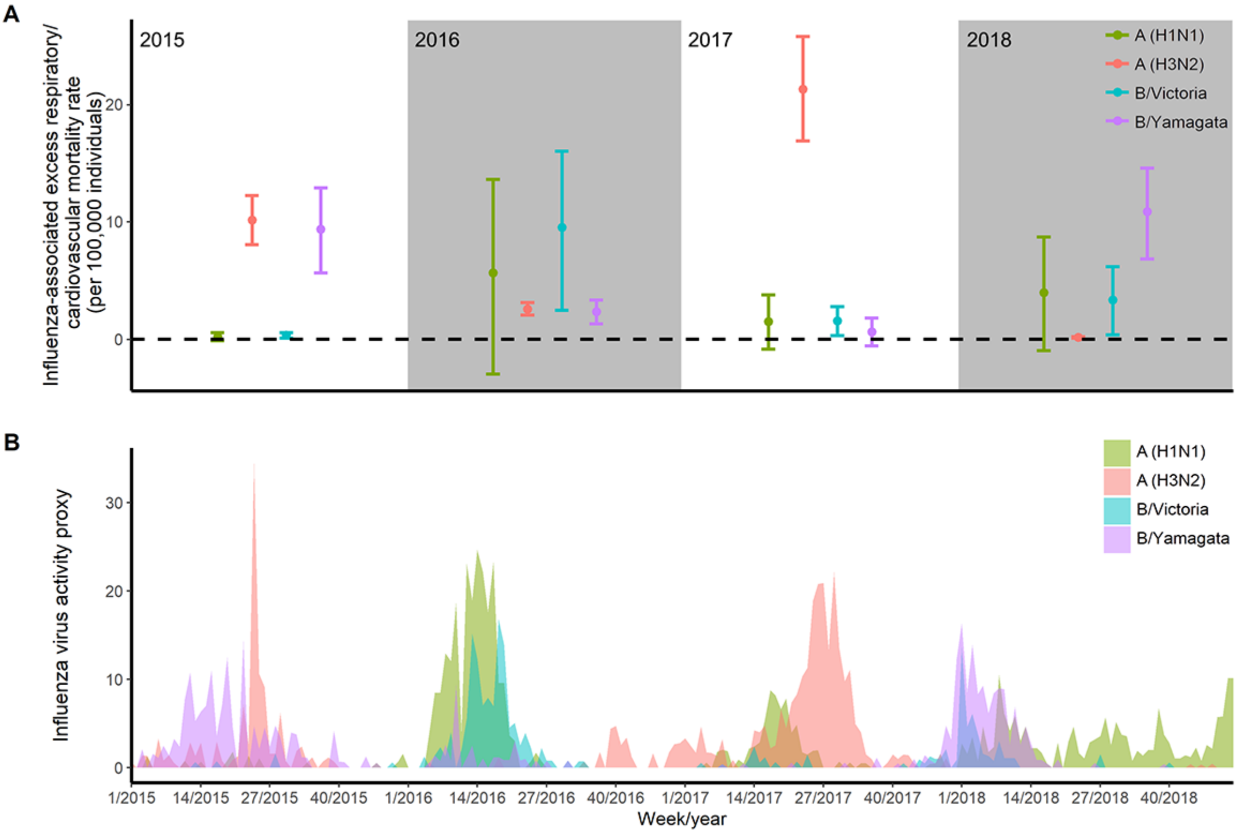
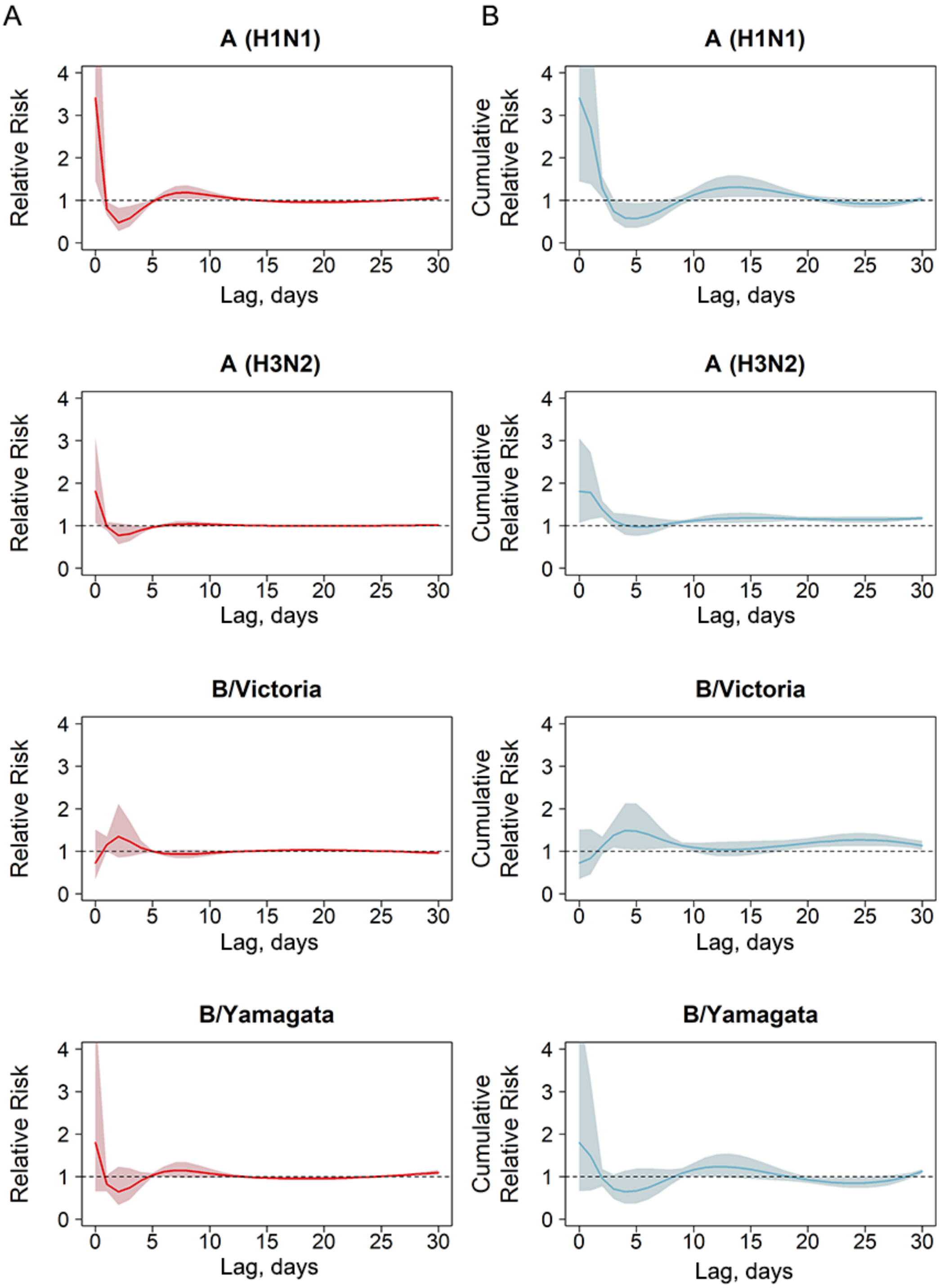


Table 3. Excess respiratory and cardiovascular mortality rates associated with different influenza subtypes and lineages.

Influenza type or lineage and age groups (years)	Rate per 100,000 persons (95% eCI[a])
A (H1N1)	
0-59	1.29 (0.07 to 2.32)
60-79	5.74 (–8.61 to 19.15)
≥80	55.75 (–60.87 to 166.10)
All ages	2.83 (–1.26 to 6.71)
A (H3N2)	
0-59	0.33 (–0.24 to 0.85)
60-79	20.22 (14.56 to 25.63)
≥80	180.15 (130.75 to 227.38)
All ages	8.47 (6.60 to 10.30)
B/Victoria	
0-59	–0.47 (–1.59 to 0.45)
60-79	7.36 (–2.87 to 16.65)
≥80	88.82 (6.18 to 168.44)
All ages	3.68 (0.81 to 6.49)
B/Yamagata	
0-59	0.02 (–0.78 to 0.74)
60-79	12.87 (4.60 to 20.74)
≥80	138.21 (70.53 to 204.40)
All ages	5.81 (3.35 to 8.25)
All influenza	
0-59	1.27 (0.07 to 2.29)
60-79	45.21 (33.31 to 56.74)
≥80	452.63 (357.46 to 544.47)
All ages	20.36 (16.75 to 23.74)

[a] eCI: empirical CI.

Figure 3. Relative risks and cumulative relative risks of death associated with influenza subtypes/lineages. (A) Relative risks of death on each single day and (B) cumulative relative risks of death associated with an increase of 10 units in influenza virus activity proxy for influenza A (H1N1), A (H3N2), B/Victoria, and B/Yamagata. The solid lines represent the estimates of relative risk over time, and the shaded areas are the corresponding 95% empirical CIs.



Estimates of influenza-associated excess respiratory and cardiovascular disease mortality rates did not change significantly after changing the method of controlling for the temporal trend in mortality rate, changing the maximum lag, or the function used for the lag-response dimension in the cross-basis matrices of influenza and temperature, ignoring the classification of influenza B lineages (Multimedia Appendix 4).

Discussion

Principal Results

In this study, we report excess mortality associated with influenza and its subtypes/lineages from 2015 to 2018 in Guangzhou, a subtropical city of China. We estimate that 1786 respiratory/cardiovascular disease deaths were attributable to influenza annually, accounting for 6.7% (7142/105,998) of all respiratory/cardiovascular disease deaths. We found that influenza-associated excess mortality was higher in the elderly, which is consistent with the conclusions of other studies [1,14,15,21,34].

Variations were observed in the estimates of influenza-associated excess mortality by subtype/lineage. The mortality burden of influenza B/Yamagata was higher than that of B/Victoria, which is consistent with previous studies [6,22,34]. The disparity between the lineages may be due to the elderly being more likely to be infected by influenza B/Yamagata than B/Victoria [35,36]. In 2018, influenza B/Yamagata was the predominant lineage in Guangzhou, and the vast majority of influenza-related excess deaths were attributed to B/Yamagata. Similarly, an outbreak dominated by influenza B/Yamagata was reported in Europe during the 2017-2018 influenza season, causing substantial excess mortality [34]. Together, the evidence implies that influenza B does not necessarily lead to mild illness in humans; the mortality burden of influenza B cannot be ignored. Furthermore, including influenza B lineages in active surveillance would help better understand the epidemiology of influenza B and reduce the excess mortality of influenza by enabling targeted interventions. Vaccination is the most effective way to prevent influenza; however, the vaccination coverage in China is low (9.4%) [23], partly because influenza vaccination is not government funded in many places [37]. During the 2018-2019 influenza season, the quadrivalent inactivated vaccine was approved for the first time in China [38]. In 2019, 30.78 million doses of influenza vaccines were distributed in China, but most of them were trivalent inactivated vaccines [37]. Compared with the trivalent influenza vaccine, the quadrivalent influenza vaccine has a higher antibody protection rate and antibody-positive seroconversion rate against the increased influenza B strain [39,40]. High-risk groups are recommended to receive quadrivalent influenza vaccines when vaccines are available, given the nonnegligible burden of influenza B.

We also noted that the mortality burden of influenza virus subtypes and lineages differed according to age group. For people aged ≥60 years, the fatality rate caused by influenza A (H3N2) was higher than that of A (H1N1), B/Victoria, and B/Yamagata. The risk of influenza virus infection and severe complications, including death, is increased by the stronger antigenic drift and virulence of influenza A (H3N2) [22,41]. However, for people aged 0 to 59 years, influenza A (H1N1) had a greater EMR than other influenza subtypes or lineages, which is in accord with previous studies [22]. The lower mortality burden of influenza A (H1N1) might be due to protection from early life exposure to this subtype [42].

Consistently with previous studies performed in Shanghai, we did not find a statistically significant difference in influenza-related excess respiratory/cardiovascular disease deaths between males and females [22]. However, males had a higher influenza-related excess respiratory mortality rate than females and a lower excess cardiovascular mortality rate. In addition, in people aged 60 to 79 years, it has been found that the influenza-related EMR for respiratory disease or cardiovascular disease in men is higher than in women, and the same situation was found in people aged 65 to 74 years in a US study [19]. However, for people aged ≥80 years, men had higher influenza-related excess respiratory mortality than women, but a lower excess cardiovascular mortality associated with influenza. In the US study, the results of a comparison of point estimates of EMR for the corresponding causes for males and females were consistent with our results, but the difference was not statistically significant [19]. Decreasing testosterone levels and higher respiratory mortality in older men [22,36], or the fact that women may have a different innate or acquired immune response than men due to discrepancies in sex steroid hormones [43], might lead to sex disparities in the severity of illness after influenza virus infection.

Our estimate of average annual influenza-associated excess all-cause mortality (25.06 per 100,000 individuals) was higher than that of previous studies in countries/cities and time periods including Greece from 2013 to 2017 (23.60) [21], Beijing from 2007 to 2013 (19.10) [14], Guangzhou from 2010 to 2012 (14.72) [15], and the United States from 1997 to 2007 (11.78) [19]. However, the estimate was relatively lower than the number of annual average influenza-related excess all-cause deaths in Europe for 2017-2018 (25.40) [34], Chongqing from 2012 to 2018 (33.50) [44], and Shanghai from 2010 to 2015 (27.66) [22]. Factors such as population structure, health status, health care capacity, circulating strains, host immunity, public health interventions (eg, influenza vaccination programs), and social distancing might contribute to these differences [20,22]. In addition to the abovementioned factors that may lead to heterogeneity in the results of different studies, some methodological differences need to be considered. The estimates of influenza-associated excess mortality varied with different analytical strategies, although the disparities were not significant in the sensitivity analysis. Influenza can be divided into different subtypes and lineages to explore the impact of influenza on the health of populations when relevant data are available, which also helps us understand the effects of different influenza subtypes and lineages on mortality.

Previous studies have estimated excess mortality based on the association between influenza activity proxies at a lag of 0, 1, or 2 weeks and mortality [20]. In this study, we used a DLNM to capture the association between influenza and mortality. A sensitivity analysis showed no significant differences between the influenza-associated excess respiratory/cardiovascular disease mortality rates estimated by a DLNM and a model that included linear terms of an influenza virus activity proxy and a natural cubic spline of temperature at a lag of 7 days. DLNMs allow us to consider a flexible lag-response relationship between influenza and death. The time course of influenza effects and potential mortality displacement can be examined with DLNMs [21,30]. We found that many influenza-related respiratory/cardiovascular disease deaths occurred during the initial exposure to influenza A (H1N1), A (H3N2), and B/Yamagata. Therefore, starting antiviral treatment as soon as possible, especially with neuraminidase inhibitors and other means of treatment for critically ill influenza patients, is likely to bring great survival benefits [35]. Taking into account the observed mortality displacement, we estimate that 743, 322, and 510 annual respiratory/cardiovascular disease deaths in Guangzhou were attributable to influenza A (H3N2), B/Victoria, and B/Yamagata, respectively; this indicates that the impact of influenza A (H3N2) and B on mortality is not limited to bringing forward deaths in the short term. Therefore, the mortality burden of influenza should not be overlooked. Efforts are needed to raise awareness of influenza as a severe disease, especially among high-risk groups.

Limitations

This study had some limitations. First, we did not use age-specific influenza virus activity proxies to estimate influenza-associated excess mortality, and a proxy for all-age activity was used instead. Second, the study period was relatively short (ie, 4 years), which may have influenced the comparison of the mortality burden of different influenza subtypes and lineages. This is mainly because the virology data included samples that tested positive for influenza B, but information on lineage was unavailable before 2015. Data from 2019 onwards were not collected. Moreover, we did not consider the influence of factors such as respiratory syncytial virus and vaccination coverage on the estimate of influenza-associated excess mortality, since such data were not available.

Conclusion

In conclusion, after considering the observed mortality displacement, influenza was associated with substantial mortality in Guangzhou, occurring predominantly in the elderly. The mortality burden of influenza B, particularly B/Yamagata, cannot be ignored. Contrasting sex differences were found in influenza-associated excess mortality from respiratory diseases and from cardiovascular diseases, and the underlying mechanisms need to be investigated in further studies. Our findings can help us better understand the magnitude and time-course of the effect of influenza on mortality and inform targeted interventions for mitigating the influenza mortality burden, such as immunizations with quadrivalent vaccines (especially for older individuals), behavioral campaigns, and treatment strategies.

Acknowledgments

We thank the staff members of the Guangzhou Center for Disease Control and Prevention for administrative work and data collection. This work was supported by the National Natural Science Foundation of China (grants 82003555 and 81903406).

Data Availability

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

CQO supervised the study. L Li, L Luo, and CQO conceived the study. ZLY, L Luo, WHL, ZY, CS, and BWM collected the data. L Li, ZLY, JY, and PHC analyzed the data. L Li and ZLY wrote the first draft of the manuscript. All authors contributed to the interpretation of the results and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional details of the methods.
[PDF File (Adobe PDF File), 427 KB - publichealth_v9i1e42530_app1.pdf]

Multimedia Appendix 2
Average annual excess respiratory mortality rates related to influenza by sex and age.
[PDF File (Adobe PDF File), 91 KB - publichealth_v9i1e42530_app2.pdf]

Multimedia Appendix 3
Average annual excess cardiovascular mortality rates related to influenza by sex and age.
[PDF File (Adobe PDF File), 92 KB - publichealth_v9i1e42530_app3.pdf]

Multimedia Appendix 4
Comparison of average annual excess respiratory and cardiovascular mortality rates attributable to all influenza estimated from different models in the sensitivity analysis.
[PDF File (Adobe PDF File), 100 KB - publichealth_v9i1e42530_app4.pdf]

References

1. Iuliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, Global Seasonal Influenza-associated Mortality Collaborator Network. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. Lancet 2018 Mar 31;391(10127):1285-1300 [FREE Full text] [doi: 10.1016/S0140-6736(17)33293-2] [Medline: 29248255]
2. Influenza (Seasonal). World Health Organization. URL: https://www.who.int/zh/news-room/fact-sheets/detail/influenza-(seasonal) [accessed 2022-05-01]
3. Li L, Liu Y, Wu P, Peng Z, Wang X, Chen T, et al. Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. Lancet Public Health 2019 Sep;4(9):e473-e481 [FREE Full text] [doi: 10.1016/S2468-2667(19)30163-X] [Medline: 31493844]
4. Grohskopf LA, Alyanak E, Ferdinands JM, Broder KR, Blanton LH, Talbot HK, et al. Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices, United States, 2021-22 influenza season. MMWR Recomm Rep 2021 Aug 27;70(5):1-28 [FREE Full text] [doi: 10.15585/mmwr.rr7005a1] [Medline: 34448800]
5. Coleman BL, Fadel SA, Fitzpatrick T, Thomas S. Risk factors for serious outcomes associated with influenza illness in high- versus low- and middle-income countries: Systematic literature review and meta-analysis. Influenza Other Respir Viruses 2018 Jan;12(1):22-29 [FREE Full text] [doi: 10.1111/irv.12504] [Medline: 29197154]
6. Yu X, Wang C, Chen T, Zhang W, Yu H, Shu Y, et al. Excess pneumonia and influenza mortality attributable to seasonal influenza in subtropical Shanghai, China. BMC Infect Dis 2017 Dec 07;17(1):756 [FREE Full text] [doi: 10.1186/s12879-017-2863-1] [Medline: 29212467]
7. Feng L, Shay DK, Jiang Y, Zhou H, Chen X, Zheng Y, et al. Influenza-associated mortality in temperate and subtropical Chinese cities, 2003-2008. Bull World Health Organ 2012 Apr 01;90(4):279-288B [FREE Full text] [doi: 10.2471/BLT.11.096958] [Medline: 22511824]
8. Liu X, Qin G, Li X, Zhang J, Zhao K, Hu M, et al. Excess mortality associated with influenza after the 2009 H1N1 pandemic in a subtropical city in China, 2010-2015. Int J Infect Dis 2017 Apr;57:54-60 [FREE Full text] [doi: 10.1016/j.ijid.2017.01.039] [Medline: 28167255]
9. Yu H, Feng L, Viboud CG, Shay DK, Jiang Y, Zhou H, et al. Regional variation in mortality impact of the 2009 A(H1N1) influenza pandemic in China. Influenza Other Respir Viruses 2013 Nov;7(6):1350-1360 [FREE Full text] [doi: 10.1111/irv.12121] [Medline: 23668477]
10. Hardelid P, Pebody R, Andrews N. Mortality caused by influenza and respiratory syncytial virus by age group in England and Wales 1999-2010. Influenza Other Respir Viruses 2013 Jan;7(1):35-45 [FREE Full text] [doi: 10.1111/j.1750-2659.2012.00345.x] [Medline: 22405488]
11. Green HK, Andrews N, Fleming D, Zambon M, Pebody R. Mortality attributable to influenza in England and Wales prior to, during and after the 2009 pandemic. PLoS One 2013;8(12):e79360 [FREE Full text] [doi: 10.1371/journal.pone.0079360] [Medline: 24348993]
12. van Asten L, van den Wijngaard C, van Pelt W, van de Kassteele J, Meijer A, van der Hoek W, et al. Mortality attributable to 9 common infections: significant effect of influenza A, respiratory syncytial virus, influenza B, norovirus, and parainfluenza in elderly persons. J Infect Dis 2012 Sep 01;206(5):628-639. [doi: 10.1093/infdis/jis415] [Medline: 22723641]
13. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA 2003 Jan 08;289(2):179-186. [doi: 10.1001/jama.289.2.179] [Medline: 12517228]
14. Wu S, Wei Z, Greene CM, Yang P, Su J, Song Y, et al. Mortality burden from seasonal influenza and 2009 H1N1 pandemic influenza in Beijing, China, 2007-2013. Influenza Other Respir Viruses 2018 Jan;12(1):88-97 [FREE Full text] [doi: 10.1111/irv.12515] [Medline: 29054110]
15. Wang H, Fu C, Li K, Lu J, Chen Y, Lu E, et al. Influenza associated mortality in Southern China, 2010-2012. Vaccine 2014 Feb 12;32(8):973-978. [doi: 10.1016/j.vaccine.2013.12.013] [Medline: 24370709]
16. Guo R, Zheng H, Ou C, Huang L, Zhou Y, Zhang X, et al. Impact of influenza on outpatient visits, hospitalizations, and deaths by using a time series Poisson generalized additive model. PLoS One 2016;11(2):e0149468 [FREE Full text] [doi: 10.1371/journal.pone.0149468] [Medline: 26894876]
17. Wu P, Goldstein E, Ho LM, Yang L, Nishiura H, Wu JT, et al. Excess mortality associated with influenza A and B virus in Hong Kong, 1998-2009. J Infect Dis 2012 Dec 15;206(12):1862-1871 [FREE Full text] [doi: 10.1093/infdis/jis628] [Medline: 23045622]
18. Thompson WW, Weintraub E, Dhankhar P, Cheng P, Brammer L, Meltzer MI, et al. Estimates of US influenza-associated deaths made using four different methods. Influenza Other Respir Viruses 2009 Jan;3(1):37-49 [FREE Full text] [doi: 10.1111/j.1750-2659.2009.00073.x] [Medline: 19453440]
19. Quandelacy TM, Viboud C, Charu V, Lipsitch M, Goldstein E. Age- and sex-related risk factors for influenza-associated mortality in the United States between 1997-2007. Am J Epidemiol 2014 Jan 15;179(2):156-167 [FREE Full text] [doi: 10.1093/aje/kwt235] [Medline: 24190951]
20. Li L, Wong JY, Wu P, Bond HS, Lau EHY, Sullivan SG, et al. Heterogeneity in estimates of the impact of influenza on population mortality: a systematic review. Am J Epidemiol 2018 Feb 01;187(2):378-388 [FREE Full text] [doi: 10.1093/aje/kwx270] [Medline: 28679157]

21. Lytras T, Pantavou K, Mouratidou E, Tsiodras S. Mortality attributable to seasonal influenza in Greece, 2013 to 2017: variation by type/subtype and age, and a possible harvesting effect. Euro Surveill 2019 Apr;24(14):11 [FREE Full text] [doi: 10.2807/1560-7917.ES.2019.24.14.1800118] [Medline: 30968823]
22. Jin S, Li J, Cai R, Wang X, Gu Z, Yu H, et al. Age- and sex-specific excess mortality associated with influenza in Shanghai, China, 2010-2015. Int J Infect Dis 2020 Sep;98:382-389 [FREE Full text] [doi: 10.1016/j.ijid.2020.07.012] [Medline: 32663600]
23. Wang Q, Yue N, Zheng M, Wang D, Duan C, Yu X, et al. Influenza vaccination coverage of population and the factors influencing influenza vaccination in mainland China: A meta-analysis. Vaccine 2018 Nov 19;36(48):7262-7269. [doi: 10.1016/j.vaccine.2018.10.045] [Medline: 30340886]
24. China Meteorological Data Service Centre. National Meteorological Information Centre. URL: http://data.cma.cn/en [accessed 2022-06-21]
25. Goldstein E, Cobey S, Takahashi S, Miller JC, Lipsitch M. Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: a statistical method. PLoS Med 2011 Jul;8(7):e1001051 [FREE Full text] [doi: 10.1371/journal.pmed.1001051] [Medline: 21750666]
26. Goldstein E, Viboud C, Charu V, Lipsitch M. Improving the estimation of influenza-related mortality over a seasonal baseline. Epidemiology 2012 Nov;23(6):829-838 [FREE Full text] [doi: 10.1097/EDE.0b013e31826c2dda] [Medline: 22992574]
27. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. Proc Natl Acad Sci U S A 2009 Mar 03;106(9):3243-3248 [FREE Full text] [doi: 10.1073/pnas.0806852106] [Medline: 19204283]
28. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. Stat Med 2010 Sep 20;29(21):2224-2234 [FREE Full text] [doi: 10.1002/sim.3940] [Medline: 20812303]
29. Guo Y, Gasparrini A, Armstrong BG, Tawatsupa B, Tobias A, Lavigne E, et al. Temperature variability and mortality: a multi-country study. Environ Health Perspect 2016 Oct;124(10):1554-1559 [FREE Full text] [doi: 10.1289/EHP149] [Medline: 27258598]
30. Zanobetti A, Wand MP, Schwartz J, Ryan LM. Generalized additive distributed lag models: quantifying mortality displacement. Biostatistics 2000 Sep;1(3):279-292. [doi: 10.1093/biostatistics/1.3.279] [Medline: 12933509]
31. Gasparrini A, Leone M. Attributable risk from distributed lag models. BMC Med Res Methodol 2014 Apr 23;14:55 [FREE Full text] [doi: 10.1186/1471-2288-14-55] [Medline: 24758509]
32. Bottomley C, Scott J, Isham V. Analysing interrupted time series with a control. Epidemiol Methods 2019;8(1):20180010 [FREE Full text] [doi: 10.1515/EM-2018-0010]
33. Newey WK, West KD. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 1987 May;55(3):703-708. [doi: 10.2307/1913610]
34. Nielsen J, Vestergaard LS, Richter L, Schmid D, Bustos N, Asikainen T, et al. European all-cause excess and influenza-attributable mortality in the 2017/18 season: should the burden of influenza B be reconsidered? Clin Microbiol Infect 2019 Oct;25(10):1266-1276 [FREE Full text] [doi: 10.1016/j.cmi.2019.02.011] [Medline: 30790685]
35. Caini S, Kusznierz G, Garate VV, Wangchuk S, Thapa B, de Paula Júnior FJ, Global Influenza B Study team. The epidemiological signature of influenza B virus and its B/Victoria and B/Yamagata lineages in the 21st century. PLoS One 2019;14(9):e0222381 [FREE Full text] [doi: 10.1371/journal.pone.0222381] [Medline: 31513690]
36. Skowronski DM, Chambers C, De Serres G, Sabaiduc S, Winter A, Dickinson JA, et al. Age-related differences in influenza b infection by lineage in a community-based sentinel system, 2010-2011 to 2015-2016, Canada. J Infect Dis 2017 Sep 15;216(6):697-702 [FREE Full text] [doi: 10.1093/infdis/jix393] [Medline: 28934439]
37. Tao Y, Li J, Hu Y, Hu Y, Zeng G, Zhu F. Quadrivalent influenza vaccine (Sinovac Biotech) for seasonal influenza prophylaxis. Expert Rev Vaccines 2021 Jan;20(1):1-11. [doi: 10.1080/14760584.2021.1875823] [Medline: 33434084]
38. Yang XK, Zhang YQ, Feng D, Xia Z, Fan SM, Zhao HT. Characteristics of influenza vaccines lot release in China, 2007 to 2020. Int J Virol 2021;28(05):364-368. [doi: 10.3760/cma.j.issn.1673-4092.2021.05.003]
39. Moa AM, Chughtai AA, Muscatello DJ, Turner RM, MacIntyre CR. Immunogenicity and safety of inactivated quadrivalent influenza vaccine in adults: A systematic review and meta-analysis of randomised controlled trials. Vaccine 2016 Jul 29;34(35):4092-4102 [FREE Full text] [doi: 10.1016/j.vaccine.2016.06.064] [Medline: 27381642]
40. Bekkat-Berkani R, Ray R, Jain VK, Chandrasekaran V, Innis BL. Evidence update: GlaxoSmithKline's inactivated quadrivalent influenza vaccines. Expert Rev Vaccines 2016;15(2):201-214. [doi: 10.1586/14760584.2016.1113878] [Medline: 26641539]
41. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. Nature 2008 May 29;453(7195):615-619 [FREE Full text] [doi: 10.1038/nature06945] [Medline: 18418375]
42. Gostic KM, Bridge R, Brady S, Viboud C, Worobey M, Lloyd-Smith JO. Childhood immune imprinting to influenza A shapes birth year-specific risk during seasonal H1N1 and H3N2 epidemics. PLoS Pathog 2019 Dec;15(12):e1008109 [FREE Full text] [doi: 10.1371/journal.ppat.1008109] [Medline: 31856206]

43. Furman D, Hejblum BP, Simon N, Jojic V, Dekker CL, Thiébaut R, et al. Systems analysis of sex differences reveals an immunosuppressive role for testosterone in the response to influenza vaccination. Proc Natl Acad Sci USA 2014 Jan 14;111(2):869-874 [FREE Full text] [doi: 10.1073/pnas.1321060111] [Medline: 24367114]
44. Qi L, Li Q, Ding X, Gao Y, Ling H, Liu T, et al. Mortality burden from seasonal influenza in Chongqing, China, 2012-2018. Hum Vaccin Immunother 2020 Jul 02;16(7):1668-1674 [FREE Full text] [doi: 10.1080/21645515.2019.1693721] [Medline: 32343618]

Abbreviations

CDC: Center for Disease Control and Prevention
DLNM: distributed-lag nonlinear model
EMR: excess mortality rate
ILI: influenza-like illness
MFR: male-to-female excess mortality ratio
RR: relative risk

Edited by A Mavragani; submitted 07.09.22; peer-reviewed by X Wang, H Jin; comments to author 26.10.22; revised version received 14.11.22; accepted 25.11.22; published 11.01.23.

Please cite as:
Li L, Yan ZL, Luo L, Liu W, Yang Z, Shi C, Ming BW, Yang J, Cao P, Ou CQ
Influenza-Associated Excess Mortality by Age, Sex, and Subtype/Lineage: Population-Based Time-Series Study With a Distributed-Lag Nonlinear Model
JMIR Public Health Surveill 2023;9:e42530
URL: https://publichealth.jmir.org/2023/1/e42530
doi:10.2196/42530
PMID:36630176

Original Paper

Seesaw Effect Between COVID-19 and Influenza From 2020 to 2023 in World Health Organization Regions: Correlation Analysis

Qing Wang[1], MPH; Mengmeng Jia[1], MD; Mingyue Jiang[1], MD; Wei Liu[2], BSc; Jin Yang[1], MM; Peixi Dai[3], BS; Yanxia Sun[1], MD; Jie Qian[1], MD; Weizhong Yang[1*], MD; Luzhao Feng[1*], MD, PhD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[2]Department of Statistics, Yunnan University, Kunming, China

[3]Division of Infectious Diseases, Chinese Center for Disease Control and Prevention, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Luzhao Feng, MD, PhD
School of Population Medicine and Public Health
Chinese Academy of Medical Sciences & Peking Union Medical College
No 9, Dongdan Santiao
Dongcheng District
Beijing, 100730
China
Phone: 86 10 65120716
Email: fengluzhao@cams.cn

Abstract

Background: Seasonal influenza activity showed a sharp decline in activity at the beginning of the emergence of COVID-19. Whether there is an epidemiological correlation between the dynamic of these 2 respiratory infectious diseases and their future trends needs to be explored.

Objective: We aimed to assess the correlation between COVID-19 and influenza activity and estimate later epidemiological trends.

Methods: We retrospectively described the dynamics of COVID-19 and influenza in 6 World Health Organization (WHO) regions from January 2020 to March 2023 and used the long short-term memory machine learning model to learn potential patterns in previously observed activity and predict trends for the following 16 weeks. Finally, we used Spearman correlation coefficients to assess the past and future epidemiological correlation between these 2 respiratory infectious diseases.

Results: With the emergence of the original strain of SARS-CoV-2 and other variants, influenza activity stayed below 10% for more than 1 year in the 6 WHO regions. Subsequently, it gradually rose as Delta activity dropped, but still peaked below Delta. During the Omicron pandemic and the following period, the activity of each disease increased as the other decreased, alternating in dominance more than once, with each alternation lasting for 3 to 4 months. Correlation analysis showed that COVID-19 and influenza activity presented a predominantly negative correlation, with coefficients above –0.3 in WHO regions, especially during the Omicron pandemic and the following estimated period. The diseases had a transient positive correlation in the European region of the WHO and the Western Pacific region of the WHO when multiple dominant strains created a mixed pandemic.

Conclusions: Influenza activity and past seasonal epidemiological patterns were shaken by the COVID-19 pandemic. The activity of these diseases was moderately or greater than moderately inversely correlated, and they suppressed and competed with each other, showing a seesaw effect. In the postpandemic era, this seesaw trend may be more prominent, suggesting the possibility of using one disease as an early warning signal for the other when making future estimates and conducting optimized annual vaccine campaigns.

(JMIR Public Health Surveill 2023;9:e44970) doi:10.2196/44970

KEYWORDS

COVID-19; influenza; negative correlation; seesaw effect; respiratory infectious disease; epidemiological trends

Introduction

Influenza was primarily characterized as a seasonal epidemic before the COVID-19 pandemic. Stringent public health and social interventions (PHSMs) were implemented in early 2020 to contain the spread of SARS-CoV-2, and influenza activity also declined [1-3]. However, during the latest winter and spring, influenza re-emerged and had a three-year-high level of activity in China [4] and other countries [5] at the same time that there was reduced virulence of Omicron and easing of PHSMs. Under the same natural conditions and PHSMs in a region, there seems to be a subtle relationship between these 2 respiratory infections. For example, they have been found to cocirculate and to coinfect people [6], but most study subjects have been COVID-19 patients [6,7]. General population-based epidemiological patterns and regularities have not yet been fully explored.

Member countries of the World Health Organization (WHO) regions update influenza activity data weekly in the FluNet, a web-based global influenza virology surveillance tool first introduced in 1997 [8]. The WHO divides member countries into 6 regions based on their location, including the African region (AFRO), Eastern Mediterranean region (EMRO), European region (EURO), Americas region (AMRO), Southeast Asian region (SEARO), and Western Pacific region (WPRO) [9]. Since January 2020, a subset of member countries in each region have also reported their national test positivity rates for SARS-CoV-2 on a weekly basis. Influenza sentinel surveillance systems were leveraged to integrate SARS-CoV-2 testing in specimens from influenza surveillance sources. The genomic sequencing of SARS-CoV-2 from representative and systematically sourced sentinel specimens has been expedited to monitor the trends and prevalence (ie, proportions) of existing and emerging circulating and cocirculating genetic variants (ie, clades) and to improve the geographic and demographic representativeness and timeliness of SARS-CoV-2 genetic-sequence data in publicly accessible databases to inform PHSMs [10].

Most studies have conducted epidemiological analyses and estimation of COVID-19 or influenza by leveraging infectious disease dynamics models [11-13] (eg, the susceptible-exposed-infected-removed model), but few have predicted trends based on longer time series and machine learning during the past 3 years. These models help tackle multimodal data and are increasingly being used in auxiliary diagnosis and other medical areas [14,15]. They are composed of different networks with various connections, referred to as neural networks or neural-like networks, to achieve automated learning and prediction by mimicking the neural network construction of the human brain. This method is highly adaptable and can adjust itself to diverse data sets, thus adapting itself to different application scenarios. Secondly, deep learning avoids human specification by automatically extracting data features and processing them. This is highly scalable and can improve simulation performance by increasing the network layers and nodes to cope with larger-scale data and more complex problems [16]. Therefore, more learning models could be introduced into infectious disease surveillance and studies.

In this study, we examined global changes in COVID-19 and influenza activity since 2020, aiming to assess the correlation between COVID-19 and influenza activity and estimate upcoming trends with the help of a deep learning model. The findings may provide a theoretical reference for the epidemiological patterns of these 2 respiratory infectious diseases and insights for future interventions.

Methods

Summary of the Study Design

In this study, we retrospectively described the dynamics of COVID-19 and influenza in the 6 WHO regions from January 2020 to March 2023 and used a neural network model to learn potential patterns of previously observed activity to predict trends for the next 16 weeks. Finally, based on an epidemiological perspective, Spearman correlation coefficients were used to assess the past and future correlations. The analysis was performed using Python (version 3.9.6; Python Software Foundation).

Data Source

The test positivity rate was used as an accurate indicator of disease activity; this method used the number of positive tests as the numerator and the specimens processed as the denominator. We extracted the weekly test positivity rates for influenza and SARS-CoV-2 from reports to FluNet between January 12, 2020, and March 26, 2023 [10], covering the AFRO, EMRO, EURO, AMRO, SEARO, and WPRO. The detailed raw data names and sources are presented in Multimedia Appendix 1, Table S1.

Model Construction

Based on the 3-year activity of COVID-19 and influenza, we used a long short-term memory (LSTM) machine learning model to estimate trends in the following 16 weeks. LSTM is widely applied to stock and sales forecasting, specializing in a unique memory cell structure that can process long time-series information more effectively than other net models. It can be trained directly on the series without smoothness requirements, reducing the need for artificial intervention and enhancing objectivity [16]. The LSTM transforms the traditional neurons of a neural network into store cells, which are capable of storing and transmitting the hidden information in a time series (Figure 1).

The LSTM converts conventional neurons into a storage unit containing a matrix of 3 gating devices: input gate, forgetting gate, and output gate. These combine to control the data information flow and store the hidden information in the time series [17], hence solving the gradient disappearance [18,19] (Figure 1).

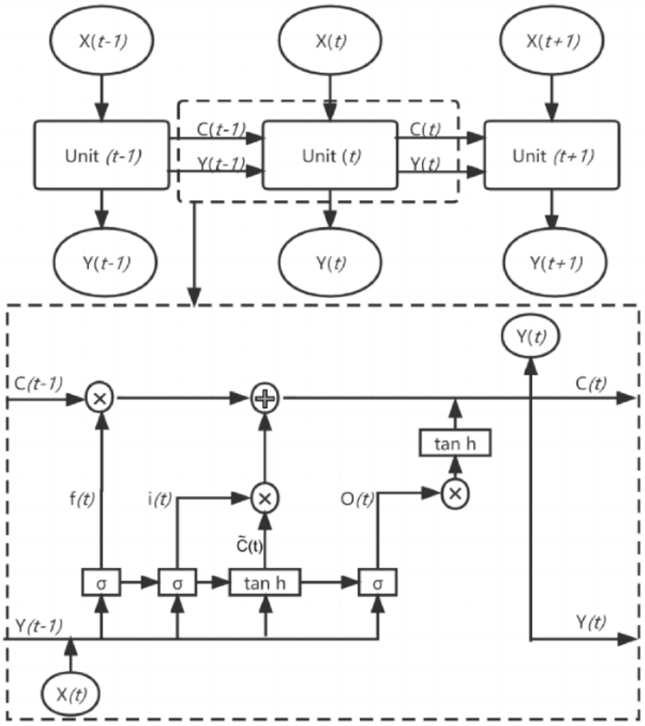
First, the forgetting gate filters the information in the unit C_{t-1} at the previous moment t–1, and how much of its C_{t-1} information is retained depends on the forget gate f_t that has the range [0, 1]; f_t is set as 1 or 0 and in between, with the information of C_{t-1} corresponding to all retained, all forgotten, and partially forgotten proportions, respectively. In the forgetting gate unit equation (equation 1), let W_f, X_t, y_{t-1}, and b_f denote the weight of the forgetting gate, the input at time t, the hidden state at time t–1, and the bias of the forget gate. The sigmoid activation function is σ (equation 2).

Second, the input gate represents the information acquired at the moment t. It determines how much information from the input X_t at the moment t can be saved in the unit C_t (equations 3 and 4). W_i and W_c are the weights of the input gates. The biases of the input gates are b_i and b_c, and tanh is the activation function (equation 5).

Third, the update of the unit status C_t (equation 6) depends on two parts: One is the product of the unit status C_t at the moment t–1 and the forget gate f_t, namely the retained information, and the other part is the product of the input unit status $\tilde{C}_t$ at the present moment t and the input gate i_t, namely the new information obtained. $\tilde{C}_t$ is the input node at time t.

Fourth, the output gate is used to output the hidden status y_t at the moment t (equations 7 and 8). W_o and b_o are the respective weights and biases of the output gates.

Figure 1. Schematic diagram of the long short-term memory neural network framework. The neural network consists of many units. The dashed box shows one of the unit structures, containing the forgetting gate, the input gate, and the output gate control system. The circles intersected by arrows in the dotted box denote multiplication or addition corresponding to the formula for the method.



Model Performance Evaluation and Selection of Optimal Parameters

To enhance the model prediction accuracy and scalability, we used the mean absolute percentage error and root mean squared error as evaluation indicators to determine the optimal proportion of training set and test set, as well as step size, layers, nodes, and algorithms [20,21], for different regions. The selection process is detailed in Multimedia Appendix 1, Tables S2-S6.

Correlation Analysis

We referred to the WHO list of the currently circulating variants of concern of SARS-CoV-2 [22] and the study objectives to uniformly divide all regions into 5 time periods to analyze the correlations: these periods covered the original strain (January 2020 to November 2020), cocirculation of a multivariant strain (referred to as “others”; December 2020 to April 2021), Delta (May 2021 to November 2021), Omicron (November 2021 to March 2023), and the period estimated by the models (March 2023 to July 2023). The starting and ending dates of variant pandemics in different regions differed slightly. In the 6 WHO regions, the Spearman nonparametric analysis of COVID-19 and influenza activity was conducted sequentially to identify whether the epidemiological patterns were correlated across the 5 time periods mentioned above. The correlation analysis was not intended to reveal causality but rather to identify and explain significant trends.

Ethical Statement

As only published data were used for this study, ethical approval was not necessary.

Results

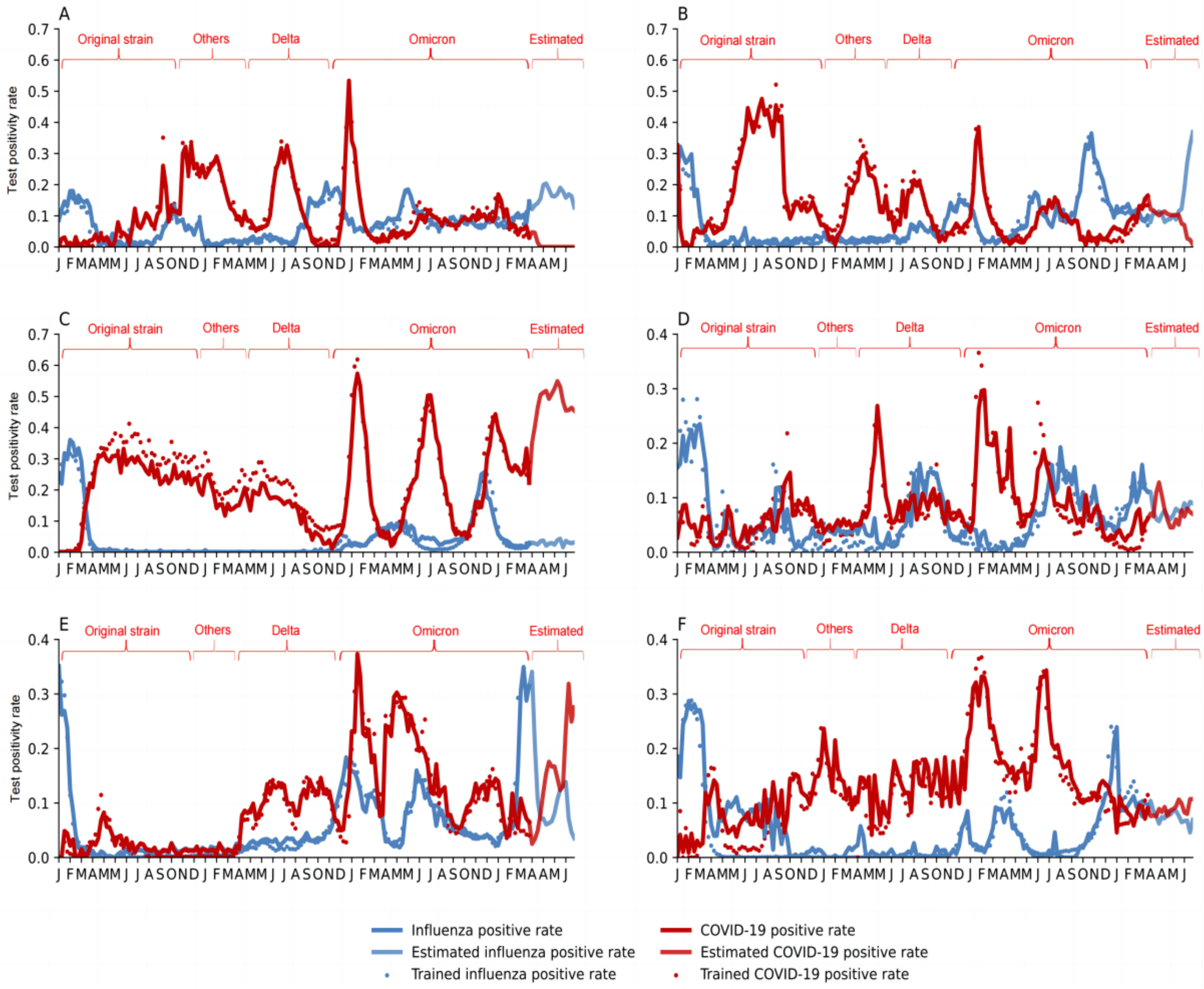
Description of Past Trends

Before the Delta variant emerged, the activity of the original strain and other variants of SARS-CoV-2 was higher than that of influenza in the 6 regions. Among the regions, SARS-CoV-2 activity was observed in AFRO, EMRO, and AMRO at a high level (peaking at over 30%), and influenza activity was at a low level (peaking at about less than 10%). Both fluctuated with low activity levels (less than 15% overall), but influenza activity was relatively lower in SEARO, WPRO, and EURO.

Following the emergence of Delta as the globally dominant strain, influenza activity began to increase in WHO regions, but in the AMRO and EURO, activity was less than 10% and remained below that of Delta. In the other 4 WHO regions, Delta activity reached a peak (over 30%) and then declined to below 10%, and the decline was accompanied by a step-up in influenza activity to 10% to 20% until the advent of Omicron.

During the first 3 months of the Omicron pandemic, influenza activity in the 6 regions first declined alongside soaring Omicron activity and then increased immediately when Omicron plummeted, whereas the influenza test positivity rate overall was consistently lower than that of Omicron. Thereafter, their trends were identical; each declined as the other rose. This occurred 1 to 2 more times, being more apparent in EMRO, AMRO, WPRO, and EURO; each alternation, when influenza activity was higher than Omicron or vice versa, lasted for about 3 to 4 months (Figure 2).

Figure 2. The observed and estimated activity of COVID-19 and influenza. Figures A-F correspond to the following regions of the World Health Organization: African, Eastern Mediterranean, Americas, Southeast Asia, Western Pacific, and European. Periods are set for the original strain (Jan 2020-Nov 2020), "others" (Dec 2020-Apr 2021), Delta (May 2021-Nov 2021), Omicron (Nov 2021-Mar 2023), and the model-estimated period (Mar 2023-July 2023). The horizontal coordinate intervals are divided by month and labeled with the first letter of the month.



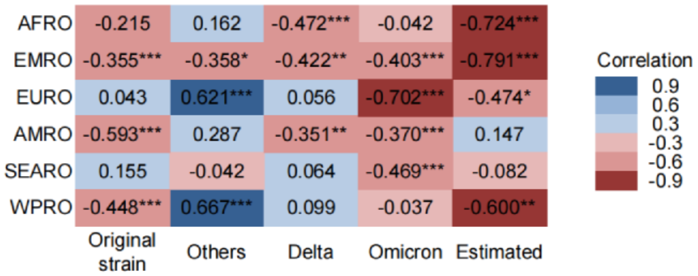
Estimated COVID-19 and Influenza Trends

We estimated the dynamics of COVID-19 and influenza activity for 16 weeks after the study observation end point (March 26, 2023). The estimation was that they would be alternatingly prevalent at less than 40% positivity rates, except for AMRO (over 40%). In SEARO and EURO, it was estimated that influenza and COVID-19 activity would be closely matched, with both hovering around 10% positivity rates; in AFRO and EMRO, influenza activity would rise from 10% to about 20% to 30%, while COVID-19 activity would decrease from similar levels to less than 5%, and COVID-19 activity would no longer be consistently higher than influenza. Conversely, influenza could decrease to below 10% in AMRO and WPRO as COVID-19 activity might go up to above 30% (Figure 2).

Correlation Between COVID-19 and Influenza Activity in Different Time Periods

The Spearman coefficients were calculated for different phases of COVID-19 and influenza activity. In our study, correlation coefficients less than 0.3, 0.3 to less than 0.6, and more than or equal to 0.6 were considered weak, moderate, and strong correlations, respectively. Statistically, COVID-19 and influenza were moderately and negatively correlated in EMRO, AMRO, and WPRO during the original-strain pandemic (–0.355, –0.593, –0.448) and also in EMRO (–0.358) during the multivariant mixed pandemic, while they were transiently, strongly, and positively correlated in EURO and WPRO (0.621 and 0.667). During the Delta pandemic, the 2 diseases were moderately and negatively correlated in AFRO, EMRO, and AMRO (–0.472, –0.422, –0.351). During the Omicron pandemic, they were moderately and negatively correlated in EMRO, AMRO, and SEARO (–0.403, –0.370, –0.469) and highly negatively correlated in EURO (–0.702). Similarly, in the estimated 16-week trend, they showed more significant and highly negative correlations in AFRO, EMRO, and WPRO (–0.724, –0.791, –0.600) and a moderate negative correlation in EURO (–0.474). Overall, COVID-19 and influenza activity were significantly negatively correlated, with coefficients greater than 0.3, especially during the Omicron pandemic and in the upcoming period. No significant correlation was found for other pandemic phases and WHO regions (Figure 3).

Figure 3. Spearman correlation coefficients for COVID-19 and influenza in different World Health Organization regions. Positive and negative correlations are shown in blue and red, respectively, with darker colors indicating a stronger correlation at the same significance level. AFRO: African region; AMRO: Americas region; EMRO: Eastern Mediterranean region; EURO: European region; SEARO: Southeast Asian region; WPRO: Western Pacific region. *$P<.01$, **$P<.05$, ***$P<.001$.



	Original strain	Others	Delta	Omicron	Estimated
AFRO	-0.215	0.162	-0.472***	-0.042	-0.724***
EMRO	-0.355***	-0.358*	-0.422**	-0.403***	-0.791***
EURO	0.043	0.621***	0.056	-0.702***	-0.474*
AMRO	-0.593***	0.287	-0.351**	-0.370***	0.147
SEARO	0.155	-0.042	0.064	-0.469***	-0.082
WPRO	-0.448***	0.667***	0.099	-0.037	-0.600**

Discussion

Principal Findings

This study analyzed the dynamics of COVID-19 and influenza in 6 WHO regions over the past 3 years and estimated them in the following 16 weeks. We found that early in the pandemic, when COVID-19 was emerging as an infectious disease, influenza activity stayed below 10% for more than 1 year in the 6 WHO regions. Subsequently, influenza activity gradually rose as Delta activity dropped, but still peaked at a level below that of Delta. Omicron alternated with influenza as the dominant disease. The trend of one disease declining as another rises has been named the seesaw effect [23] and was clearly apparent in the epidemiological patterns of the 2 infectious diseases in this study, characterized by increasing magnitude and frequency, with each alternation have a duration of about 3 to 4 months, suggesting a competitive relationship. The seesaw effect of alternating dominance is likely to become more conspicuous in the postpandemic era. Negative correlation coefficients for different WHO regions and time periods statistically complement the validation of this effect.

Our quantitative analysis based on epidemiological indicators (test positivity rates from the population) showed the dynamic patterns of the 2 respiratory diseases, which have similar transmission modes. The generation and variation of the seesaw effect could be related to numerous factors. On the one hand, from a biological perspective, respiratory viruses share the same host; therefore, viruses compete with each other for susceptible cells in the host. Cells invaded by one respiratory virus produce immune interference that drives uninfected neighboring cells to adopt a protectively antiviral status [24,25], making the host resistant or only partially susceptible to subsequent viruses. One example is the significant decrease in rhinovirus prevalence in patients during peak influenza activity [26]. On the other hand, this effect varies in magnitude and duration across regions, which could be driven by the different intensity of PHSMs adopted in different countries [27-30] and time periods [31], differences in vaccination and natural infection status, or variant strain properties. For example, the earliest adoption of strict nonpharmaceutical interventions in China not only controlled the spread and dissemination of the original strain [32], but also reduced influenza activity by 82%, and by 64% in the north and south [3]. Interestingly, after the PHSMs, influenza activity in China in the latest winter was at a low level because of the rapid rise in Omicron infections in the population at the same time, and after the Omicron test positivity rate dropped, influenza activity jumped sharply, to over 50%. This once again suggests that driving factors influence the magnitude of the seesaw effect and the timing of its onset. In addition, the higher virulence of the original strain attacked individuals more and longer [6], resulting in less opportunity for influenza virus infection; thus, absolute suppression occurred in the first year of the COVID-19 pandemic. The more moderate virulence and lower severity of Delta and Omicron allow the possibility of infecting the host with influenza, which may explain why influenza activity increased after the Delta pandemic and was able to dominate during the Omicron period.

During the multivariant pandemic period, the dynamics of COVID-19 and influenza had short-term positive correlations in EURO and WPRO, while negative correlations were not significant in other regions. Temporary phenomena mediated by complex factors, such as the sudden appearance of variants, require more evidence to unravel their underlying mechanisms. Similar phenomena have been observed in other studies: there are reports of coinfection with SARS-CoV-2 and influenza A in some countries [33-36], and meta-analyses based on the November 2019 to August 2021 period [37] all showed coinfection with both diseases. Nevertheless, most study subjects have been COVID-19 patients or laboratory animals rather than the general population [37,38], with a low proportion of observed coinfections (about 1%) or small sample sizes [39] before and during the Omicron pandemic [37,40]. Inevitably, the positive correlation for susceptibility to coinfection at the individual level or in some high-risk populations might lead to serious clinical outcomes [41]; therefore, much larger, sample-based meta-studies are warranted to explore the reasons for this phenomenon to prevent and control it or to meet the potential surge in hospital visits and demand. In addition, it is worth noting that other respiratory infectious diseases have been threatening humans. Further research on the interactions between respiratory pathogens other than COVID-19 and influenza remains to be done. Regardless of the fate of COVID-19, influenza, or other respiratory pathogens, necessary personal protection and vaccination should be implemented.

Our study provided qualitative and quantitative assessments of the interaction between COVID-19 and influenza. We identified a seesaw effect between COVID-19 and influenza activity based on a competitive effect from an epidemiological perspective, echoing viral antagonism from a historical pathogenic perspective. This finding can be used to guide disease surveillance, early warning, and intervention; for example, real-time surveillance of viral dynamics can be used to estimate trends in other diseases or stagger vaccinations in response to alternating epidemics or pandemics.

Limitations

There are some limitations to mention. First, the high uncertainty of viral variation, changes in influenza visit behavior triggered by early and rigorous interventions, the decline in SARS-CoV-2 tests in some countries, and other unforeseen factors all interfered with the observed positive test numbers for the 2 diseases studied. Hence, we used the test positivity rate rather than the number of positive tests as an indicator of activity and divided the time periods to minimize the negative impact. Second, the raw data may not fully represent the overall trend in each area because of the limited number of member countries that periodically report to the WHO. Nevertheless, there is no more authoritative data source for COVID-19 and influenza activity.

Conclusions

Influenza activity and former seasonal epidemiological patterns were shaken by the COVID-19 pandemic. Their activities were inversely correlated to a moderate or greater degree, and they suppressed and competed with each other, showing a seesaw effect. In the postpandemic era of COVID-19, seesaw trends will be more prominent, prompting the possibility of using one disease as an early warning signal for the other for future estimates; this could be used to conduct optimized annual vaccine campaigns.

Acknowledgments

We sincerely thank all the member of the Infectious Diseases Team at the School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College. This study was supported by the Non-profit Central Research Institute Fund of the Chinese Academy of Medical Sciences (2021-RC330-002) and the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2021-I2M-1-044, 2020-I2M-1-001). The funding sources had no role in data collection or interpretation or decisions relating to data publication.

Data Availability

The data are available in the FluNet repository [10]. The code generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

QW and LF conceptualized and designed the study. QW, JY, PD, YS, and JQ performed data collection. QW, M Jia, and WL performed data analysis and model construction. QW and M Jiang wrote the first draft. LF and WY supervised the project. All authors read and approved the final manuscript. Authors LF and WY (yangweizhong@cams.cn) are cocorresponding authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material and tables for the methods and results.
[DOCX File , 31 KB - publichealth_v9i1e44970_app1.docx]

References

1. Olsen SJ, Azziz-Baumgartner E, Budd AP, Brammer L, Sullivan S, Pineda RF, et al. Decreased influenza activity during the COVID-19 pandemic-United States, Australia, Chile, and South Africa, 2020. Am J Transplant 2020 Dec;20(12):3681-3685 [FREE Full text] [doi: 10.1111/ajt.16381] [Medline: 33264506]
2. Sakamoto H, Ishikane M, Ueda P. Seasonal influenza activity during the SARS-CoV-2 outbreak in Japan. JAMA 2020 May 19;323(19):1969-1971 [FREE Full text] [doi: 10.1001/jama.2020.6173] [Medline: 32275293]
3. Feng L, Zhang T, Wang Q, Xie Y, Peng Z, Zheng J, et al. Impact of COVID-19 outbreaks and interventions on influenza in China and the United States. Nat Commun 2021 May 31;12(1):3249 [FREE Full text] [doi: 10.1038/s41467-021-23440-1] [Medline: 34059675]
4. Influenza Weekly Report. Chinese National Influenza Center. URL: https://ivdc.chinacdc.cn/cnic/zyzx/lgzb/202304/t20230413_265257.htm [accessed 2023-04-13]
5. Global Influenza Programme. World Health Organization. URL: https://www.who.int/teams/global-influenza-programme/surveillance-and-monitoring/influenza-updates [accessed 2023-04-12]
6. Swets MC, Russell CD, Harrison EM, Docherty AB, Lone N, Girvan M, ISARIC4C Investigators, et al. SARS-CoV-2 co-infection with influenza viruses, respiratory syncytial virus, or adenoviruses. Lancet 2022 Apr 16;399(10334):1463-1464 [FREE Full text] [doi: 10.1016/S0140-6736(22)00383-X] [Medline: 35344735]
7. Kawai S, Fukushima K, Yomota M, Fukuda A, Fujiwara S, Tanaka M, et al. Number of patients with influenza and COVID-19 coinfection in a single Japanese hospital during the first wave. Jpn J Infect Dis 2021 Nov 22;74(6):570-572 [FREE Full text] [doi: 10.7883/yoken.JJID.2020.1009] [Medline: 33952766]
8. FluNet. World Health Organization. URL: https://www.who.int/tools/flunet [accessed 2023-03-10]
9. WHO regional offices. World Health Organization. URL: https://www.who.int/about/who-we-are/regional-offices [accessed 2023-03-10]
10. Influenza and SARS-CoV-2 surveillance data reported to FluNet. World Health Organization. URL: https://tinyurl.com/y6crsw4u [accessed 2023-04-10]
11. Franco N. COVID-19 Belgium: Extended SEIR-QD model with nursing homes and long-term scenarios-based forecasts. Epidemics 2021 Dec;37:100490 [FREE Full text] [doi: 10.1016/j.epidem.2021.100490] [Medline: 34482186]
12. Shin H. A multi-stage SEIR(D) model of the COVID-19 epidemic in Korea. Ann Med 2021 Dec;53(1):1159-1169 [FREE Full text] [doi: 10.1080/07853890.2021.1949490] [Medline: 34269629]
13. Wang Y, Wang P, Zhang S, Pan H. Uncertainty modeling of a modified SEIR epidemic model for COVID-19. Biology (Basel) 2022 Aug 02;11(8):1157 [FREE Full text] [doi: 10.3390/biology11081157] [Medline: 36009784]
14. Wu X, Wang H, Shi P, Sun R, Wang X, Luo Z, et al. Long short-term memory model - A deep learning approach for medical data with irregularity in cancer predication with tumor markers. Comput Biol Med 2022 May;144:105362 [FREE Full text] [doi: 10.1016/j.compbiomed.2022.105362] [Medline: 35299045]
15. Lu JY, Zhu J, Zhu J, Duong TQ. Long-short-term memory machine learning of longitudinal clinical data accurately predicts acute kidney injury onset in COVID-19: a two-center study. Int J Infect Dis 2022 Sep;122:802-810 [FREE Full text] [doi: 10.1016/j.ijid.2022.07.034] [Medline: 35872094]
16. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997 Nov 15;9(8):1735-1780. [doi: 10.1162/neco.1997.9.8.1735] [Medline: 9377276]
17. Shu X, Tang J, Qi G, Liu W, Yang J. Hierarchical long short-term concurrent memory for human interaction recognition. IEEE Trans Pattern Anal Mach Intell 2021 Mar;43(3):1110-1118. [doi: 10.1109/TPAMI.2019.2942030] [Medline: 31545711]
18. Jo Y, Lee L, Palaskar S. Combining LSTM and latent topic modeling for mortality prediction. ArXiv Preprint posted online Sep 8, 2017. [FREE Full text] [doi: 10.48550/arXiv.1709.02842]
19. DiPietro R, Hager G. Deep learning: RNNs and LSTM. In: Zhou SK, Rueckert D, Fichtinger G, editors. Handbook of Medical Image Computing and Computer Assisted Intervention. Cambridge, MA: Academic Press; 2020:503-519.
20. Zhang J, Nawata K. Multi-step prediction for influenza outbreak by an adjusted long short-term memory. Epidemiol Infect 2018 May;146(7):809-816 [FREE Full text] [doi: 10.1017/S0950268818000705] [Medline: 29606177]
21. Xu J, Wu Q. Prediction on influenza-like virus pathogen and its effects on prognosis of patients with community acquired pneumonia under long and short term memory neural network model. Results Phys 2021 May;24:104079 [FREE Full text] [doi: 10.1016/j.rinp.2021.104079]
22. Tracking SARS-CoV-2 variants. World Health Organization. URL: https://www.who.int/activities/tracking-SARS-CoV-2-variants [accessed 2023-04-10]
23. Ortwine JK, Werth BJ, Sakoulas G, Rybak MJ. Reduced glycopeptide and lipopeptide susceptibility in Staphylococcus aureus and the "seesaw effect": Taking advantage of the back door left open? Drug Resist Updat 2013;16(3-5):73-79. [doi: 10.1016/j.drup.2013.10.002] [Medline: 24268586]
24. Chan K, Carolan L, Korenkov D, Druce J, McCaw J, Reading P, et al. Investigating viral interference between influenza A virus and human respiratory syncytial virus in a ferret model of infection. J Infect Dis 2018 Jul 02;218(3):406-417 [FREE Full text] [doi: 10.1093/infdis/jiy184] [Medline: 29746640]
25. Gonzalez A, Ijezie E, Balemba O, Miura T. Attenuation of influenza A virus disease severity by viral coinfection in a mouse model. J Virol 2018 Dec 01;92(23):e00881-18 [FREE Full text] [doi: 10.1128/JVI.00881-18] [Medline: 30232180]

26. Nickbakhsh S, Mair C, Matthews L, Reeve R, Johnson PCD, Thorburn F, et al. Virus-virus interactions impact the population dynamics of influenza and the common cold. Proc Natl Acad Sci U S A 2019 Dec 26;116(52):27142-27150 [FREE Full text] [doi: 10.1073/pnas.1911083116] [Medline: 31843887]
27. Hong HC, Lee H, Lee SJ, Park C, Lee M. The determinants of adherence to public health and social measures against COVID-19 among the general population in South Korea: national survey study. JMIR Public Health Surveill 2023 Jan 17;9:e35784 [FREE Full text] [doi: 10.2196/35784] [Medline: 36446132]
28. Leung K, Wu JT, Leung GM. Effects of adjusting public health, travel, and social measures during the roll-out of COVID-19 vaccination: a modelling study. Lancet Public Health 2021 Sep;6(9):e674-e682 [FREE Full text] [doi: 10.1016/S2468-2667(21)00167-5] [Medline: 34388389]
29. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, Imperial College COVID-19 Response Team, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020 Aug;584(7820):257-261. [doi: 10.1038/s41586-020-2405-7] [Medline: 32512579]
30. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science 2021 Feb 19;371(6531):eabd9338 [FREE Full text] [doi: 10.1126/science.abd9338] [Medline: 33323424]
31. Carvalho T, Krammer F, Iwasaki A. The first 12 months of COVID-19: a timeline of immunological insights. Nat Rev Immunol 2021 Apr;21(4):245-256 [FREE Full text] [doi: 10.1038/s41577-021-00522-1] [Medline: 33723416]
32. Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature 2020 Sep;585(7825):410-413 [FREE Full text] [doi: 10.1038/s41586-020-2293-x] [Medline: 32365354]
33. Hashemi SA, Safamanesh S, Ghasemzadeh-Moghaddam H, Ghafouri M, Azimian A. High prevalence of SARS-CoV-2 and influenza A virus (H1N1) coinfection in dead patients in Northeastern Iran. J Med Virol 2021 Feb;93(2):1008-1012. [doi: 10.1002/jmv.26364] [Medline: 32720703]
34. Covin S, Rutherford G. Coinfection, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and influenza: an evolving puzzle. Clin Infect Dis 2021 Jun 15;72(12):e993-e994 [FREE Full text] [doi: 10.1093/cid/ciaa1810] [Medline: 33277660]
35. Cuadrado-Payán E, Montagud-Marrahi E, Torres-Elorza M, Bodro M, Blasco M, Poch E, et al. SARS-CoV-2 and influenza virus co-infection. Lancet 2020 May 16;395(10236):e84 [FREE Full text] [doi: 10.1016/S0140-6736(20)31052-7] [Medline: 32423586]
36. Wu X, Cai Y, Huang X, Yu X, Zhao L, Wang F, et al. Co-infection with SARS-CoV-2 and influenza a virus in patient with pneumonia, China. Emerg Infect Dis 2020 Jun;26(6):1324-1326 [FREE Full text] [doi: 10.3201/eid2606.200299] [Medline: 32160148]
37. Krumbein H, Kümmel LS, Fragkou PC, Thölken C, Hünerbein BL, Reiter R, et al. Respiratory viral co-infections in patients with COVID-19 and associated outcomes: A systematic review and meta-analysis. Rev Med Virol 2023 Jan;33(1):e2365 [FREE Full text] [doi: 10.1002/rmv.2365] [Medline: 35686619]
38. Bai L, Zhao Y, Dong J, Liang S, Guo M, Liu X, et al. Coinfection with influenza A virus enhances SARS-CoV-2 infectivity. Cell Res 2021 Apr;31(4):395-403 [FREE Full text] [doi: 10.1038/s41422-021-00473-1] [Medline: 33603116]
39. Adams K, Tastad KJ, Huang S, Ujamaa D, Kniss K, Cummings C, et al. Prevalence of SARS-CoV-2 and influenza coinfection and clinical characteristics among children and adolescents aged <18 years who were hospitalized or died with influenza - United States, 2021-22 influenza season. MMWR Morb Mortal Wkly Rep 2022 Dec 16;71(50):1589-1596 [FREE Full text] [doi: 10.15585/mmwr.mm7150a4] [Medline: 36520656]
40. Fujita DM, Dos Santos Soares G, Sartori GP, Henrique da Silva Nali L. COVID-19 and Influenza coinfection: The rise of Ômicron and H3N2 in Brazil - 2022. Travel Med Infect Dis 2022;46:102262 [FREE Full text] [doi: 10.1016/j.tmaid.2022.102262] [Medline: 35038569]
41. Kim E, Nguyen T, Casel MAB, Rollon R, Kim S, Kim Y, et al. Coinfection with SARS-CoV-2 and influenza A virus increases disease severity and impairs neutralizing antibody and CD4 T cell responses. J Virol 2022 Mar 23;96(6):e0187321 [FREE Full text] [doi: 10.1128/jvi.01873-21] [Medline: 35107382]

Abbreviations

AFRO: African region of WHO
AMRO: Americas region of the WHO
EMRO: Eastern Mediterranean region of the WHO
EURO: European region of the WHO
LSTM: long short-term memory
PHSM: public health and social measures
SEARO: Southeast Asia region of the WHO
WHO: World Health Organization
WPRO: Western Pacific region of the WHO

Edited by A Mavragani, T Sanchez; submitted 11.12.22; peer-reviewed by Y Wang, S Tang; comments to author 23.03.23; revised version received 15.04.23; accepted 16.05.23; published 12.06.23.

Please cite as:
Wang Q, Jia M, Jiang M, Liu W, Yang J, Dai P, Sun Y, Qian J, Yang W, Feng L
Seesaw Effect Between COVID-19 and Influenza From 2020 to 2023 in World Health Organization Regions: Correlation Analysis
JMIR Public Health Surveill 2023;9:e44970
URL: https://publichealth.jmir.org/2023/1/e44970
doi:10.2196/44970
PMID:37191650

Review

The Prevalence, Features, Influencing Factors, and Solutions for COVID-19 Vaccine Misinformation: Systematic Review

Sihong Zhao[1*], BSc; Simeng Hu[1*], BSc; Xiaoyu Zhou[1], BSc; Suhang Song[2], PhD; Qian Wang[1], MPH; Hongqiu Zheng[1], BSc; Ying Zhang[1], BSc; Zhiyuan Hou[1], PhD

[1]School of Public Health, Global Health Institute, National Health Commission Key Laboratory of Health Technology Assessment, Fudan University, Shanghai, China

[2]Department of Health Policy and Management, College of Public Health, University of Georgia, Athens, GA, United States

[*]these authors contributed equally

Corresponding Author:
Zhiyuan Hou, PhD
School of Public Health, Global Health Institute
National Health Commission Key Laboratory of Health Technology Assessment
Fudan University
220 Handan Road, Yangpu District
Shanghai, 200437
China
Phone: 86 21 6564 2222
Email: zyhou@fudan.edu.cn

Abstract

Background: During the COVID-19 pandemic, infodemic spread even more rapidly than the pandemic itself. The COVID-19 vaccine hesitancy has been prevalent worldwide and hindered pandemic exiting strategies. Misinformation around COVID-19 vaccines is a vital contributor to vaccine hesitancy. However, no evidence systematically summarized COVID-19 vaccine misinformation.

Objective: This review aims to synthesize the global evidence on misinformation related to COVID-19 vaccines, including its prevalence, features, influencing factors, impacts, and solutions for combating misinformation.

Methods: We performed a systematic review by searching 5 peer-reviewed databases (PubMed, Embase, Web of Science, Scopus, and EBSCO). We included original articles that investigated misinformation related to COVID-19 vaccines and were published in English from January 1, 2020, to August 18, 2022. We excluded publications that did not cover or focus on COVID-19 vaccine misinformation. The Appraisal tool for Cross-Sectional Studies, version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB 2), and Critical Appraisal Skills Programme Checklist were used to assess the study quality. The review was guided by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and registered with PROSPERO (CRD42021288929).

Results: Of the 8864 studies identified, 91 observational studies and 11 interventional studies met the inclusion criteria. Misinformation around COVID-19 vaccines covered conspiracy, concerns on vaccine safety and efficacy, no need for vaccines, morality, liberty, and humor. Conspiracy and safety concerns were the most prevalent misinformation. There was a great variation in misinformation prevalence, noted among 2.5%-55.4% in the general population and 6.0%-96.7% in the antivaccine/vaccine hesitant groups from survey-based studies, and in 0.1%-41.3% on general online data and 0.5%-56% on antivaccine/vaccine hesitant data from internet-based studies. Younger age, lower education and economic status, right-wing and conservative ideology, and having psychological problems enhanced beliefs in misinformation. The content, format, and source of misinformation influenced its spread. A 5-step framework was proposed to address vaccine-related misinformation, including identifying misinformation, regulating producers and distributors, cutting production and distribution, supporting target audiences, and disseminating trustworthy information. The debunking messages/videos were found to be effective in several experimental studies.

Conclusions: Our review provides comprehensive and up-to-date evidence on COVID-19 vaccine misinformation and helps responses to vaccine infodemic in future pandemics.

Trial Registration: PROSPERO CRD42021288929; https://tinyurl.com/2prejtfa

(JMIR Public Health Surveill 2023;9:e40201) doi:10.2196/40201

KEYWORDS

COVID-19; COVID-19 vaccine; misinformation; anti-vaccine; review; social media; survey

Introduction

Background

The COVID-19 pandemic has become the most threatening global health issue for almost 3 years [1]. As a cost-effective measure to protect people, governments have implemented various policies to promote the COVID-19 vaccination. However, as of early 2022, the global acceptance rate of the COVID-19 vaccination was only 67.8% [2]. Considering the powerful capability of the omicron variant to escape neutralizing antibodies elicited by current vaccines, the current vaccine acceptance rate is not enough to control the omicron variant [3,4]. It is thus necessary to investigate the negative factors that hinder the COVID-19 vaccination and take actions to increase vaccine coverage.

Vaccine acceptance is determined by contextual influences, individual/social group influences, and vaccine- and vaccination-specific issues [5]. Among all these factors, influences of infodemic and vaccine misinformation deserve more attention for the COVID-19 vaccination. An infodemic is an overabundance of information including misinformation in digital and physical environments, which makes it hard to find trustworthy sources and reliable guidance during a disease outbreak [6]. In the era of social media, the dissemination of information, especially misinformation, has been intensified [7]. During the COVID-19 pandemic, COVID-19 infodemic spread even more rapidly than the pandemic itself. The COVID-19 infodemic had jeopardized public trust in the pandemic response strategies such as vaccination and attracted attention from governments and health agencies across the world [8]. It is difficult but urgent to terminate and resolve the infodemic to promote the vaccination.

Misinformation is referred to false or inaccurate information deliberately intended to deceive [9]. It originates from rumors, websites and social media, works of fiction, governments, politicians, and vested interests [10,11]. Misinformation around COVID-19 vaccines is a noteworthy component of contextual influences on vaccine acceptance or hesitancy [12,13]. It can distort people's perception of COVID-19 vaccines [14], exaggerate the probability of adverse events following vaccination [15], and lead to extreme political sentiments [16]. COVID-19 vaccine hesitancy, partially driven by misinformation, heavily hindered the pandemic exiting strategies worldwide.

Some reviews have summarized COVID-19–related infodemic and misinformation during the pandemic, but there is a lack of systematic evidence focusing on COVID-19 vaccine misinformation. For instance, Ries [17] synthesized the mechanisms and impacts of COVID-19 infodemic. Gabarron and colleagues [18] summarized the types of COVID-19–related misinformation and its possible consequences. A few studies focused on misinformation about COVID-19 vaccines, but they mainly evaluated its influence on vaccine hesitancy [19,20]. However, no evidence systematically summarized the distribution of COVID-19 vaccine misinformation in the population, what features it has, and how to fight against it.

Objective

We aimed to synthesize global evidence on misinformation related to COVID-19 vaccines, including its prevalence, features, influencing factors, impacts, and solutions for combating misinformation around COVID-19 vaccines. Specifically, the following questions guided our inquiry: How prevalent was COVID-19 vaccine misinformation across regions and populations? What types and features did the misinformation have? Where did the misinformation come from? How was it distributed among the general population? What factors affected misinformation believing and spreading? How did the misinformation influence vaccine hesitancy and behaviors? How to fight against vaccine misinformation in a future pandemic? This systematic review would enrich the evidence regarding vaccine misinformation and inform response strategies when new vaccines are introduced in future pandemics.

Methods

Search Strategy and Selection Criteria

According to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, we conducted a systematic review of empirical articles on COVID-19 vaccine misinformation. The PRISMA checklist can be found in Multimedia Appendix 1. The review protocol was registered with PROSPERO (registration number CRD42021288929). This review was developed based on 5 peer-reviewed databases (PubMed, Embase, Web of Science, Scopus, and EBSCO). We used keywords related to COVID-19 vaccines and misinformation to identify empirical articles published in English from January 1, 2020, to August 18, 2022.

To identify the keywords of "misinformation," we referred to the codebook by Kata [10], a widely recognized study on the classification of vaccine misinformation. We also referred to the fact sheet of COVID-19 vaccine misinformation published by the US Centers for Disease Control and Prevention (CDC) and collected the keywords on their websites to match the emerging misinformation about COVID-19 vaccines [21]. Using the keywords above, we piloted literature search in PubMed and Web of Science, and further refined the keywords according to literature searching results. The search strategy (Multimedia Appendix 2) consisted of 2 major concepts: COVID-19 vaccine and misinformation, which contains the general descriptions of misinformation (such as "misinformation," "infodemic," "myth") and specified descriptions of some certain misinformation (such as "fertility," "toxic," "freedom").

Original observational or interventional articles that investigated misinformation related to COVID-19 vaccines were included.

We excluded studies that (1) investigated non-COVID-19 vaccines or the COVID-19 pandemic instead of COVID-19 vaccines, (2) did not investigate misinformation, and (3) did not focus on COVID-19 vaccine misinformation. We also excluded the following article types: conference abstract, editorial, letter, commentary, correspondence, study protocol, and review.

Data Screening and Extraction

We exported identified articles from databases, imported them into EndNote 20 (Clarivate Analytics), and removed the duplicates. Two reviewers (SH and SZ) first screened titles and abstracts independently to include articles meeting the inclusion criteria. The full texts of included studies after initial screening were scrutinized to assess the overall eligibility based on the inclusion and exclusion criteria by 2 independent reviewers (SH and SZ). When discrepancies in article inclusion emerged between the 2 reviewers, they engaged in discussion with a third researcher (XZ) to reach a consensus.

For eligible studies, data were independently extracted by 2 reviewers (SH and SZ), and inconsistencies or disagreements were reconciled in data extraction. Besides the study characteristics (region and period, study design, data sources, target population, sample size, and analysis methods), we extracted 4 outcomes of interest for each included study: (1) the types, sources, and prevalence of COVID-19 vaccine misinformation; (2) factors that affect the believing and spreading of COVID-19 vaccine misinformation; (3) the impact of misinformation on vaccine hesitancy and behaviors; and (4) proposed solutions for combating misinformation.

Classification of Study Types, Populations, and Phases

In this systematic review, we covered 4 types of study design: survey-based study, internet-based study, interview, and experiment. Survey-based study included cross-sectional or follow-up studies conducted among population using questionnaires; and internet-based study referred to studies that acquired publicly available information thought the internet. The prevalence and impact of COVID-19 vaccine misinformation might be different between the vaccine hesitant/refusal group and the general population. Thus, each type of studies was divided into 2 subgroups according to their study population's prior defined attitudes toward COVID-19 vaccines: the antivaccine/vaccine hesitant group and the general population group.

To investigate the potential change in vaccine misinformation in terms of its types, contents, prevalence, and impact, we used 2 key time points to define 3 phases in this review: prevaccination phase (phase 1), from the outbreak of COVID-19 to the first dose of COVID-19 vaccines being injected at December 8, 2020 [22]; postvaccination and pre-Omicron phase (phase 2), from the end of phase 1 to the date that the new variant Omicron was officially reported (November 26, 2021) [23]; post-Omicron phase (phase 3), after November 26, 2021. If the time frame in data collection covered the time point and the ending/beginning date was not close to the time point (over 1 month), the study phase was considered to have crossed 2 phases.

Framework of Classifying Misinformation Types and Contents

Like the identification process of misinformation-related keywords, we first adapted the codebook by Kata [10] to construct a framework to classify the types and contents of COVID-19 vaccine misinformation. To cover the emerging and evolving COVID-19 vaccine infodemic, we further referred to the CDC fact sheet to revise the framework [21]. Our primary framework consisted of main classifications and contents from Kata's work as well as contents from the CDC fact sheet. New types or contents of misinformation may have occurred when we extracted data from the included studies, and would be considered to finalize our framework. The original misinformation contents extracted from articles were rechecked using the final framework. Discrepancies in misinformation classification were discussed and reconciled.

Framework of Solutions for Combating Misinformation

For clarity and consistency in data extraction, we sorted solutions for combating misinformation into a 5-step framework, which was adapted from the disinfodemic policy brief by the United Nations Educational, Scientific and Cultural Organization (UNESCO) [24]. In this policy brief, solutions to debunk the misinformation consisted of 4-part responses: identifying misinformation, regulating producers and distributors, cutting production and distribution, and supporting the target audiences of misinformation. As the dissemination of trustworthy information deserved more attention, it was added as the fifth part of the framework.

Quality Assessment

As study designs varied across eligible studies, we used 3 kinds of quality assessment tools to assess their quality. The Appraisal tool for Cross-Sectional Studies (AXIS) was used to evaluate survey-based cross-sectional studies, and its modified version was applied to assess internet-based studies [25]. The version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB 2) tool and the Critical Appraisal Skills Programme Checklist were used to assess methodological quality in experimental studies and interviews, respectively [26,27]. Quality assessment was conducted by 2 reviewers (SH and SZ), and the risk of bias for each eligible study was classified as "low risk," "some concerns," or "high risk." The detailed scoring criteria of quality assessment are shown in Multimedia Appendix 3.

Patient and Public Involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Ethical Approval

Ethical approval was waived as this is a secondary analysis on the published articles.

Results

Basic Characteristics of Included Studies

In total, we identified 23,398 studies from 5 peer-reviewed databases (Figure 1). After removing duplicates, 8864 studies were initially screened based on their titles and abstracts, and after initial screening, 313 studies underwent full-text assessment for eligibility. A total of 102 studies met the inclusion criteria, including 91 observational studies and 11 interventional studies. According to the quality assessment criteria, the majority (90/102) of included studies were in the low risk of bias category. Eight studies were classified as “some concerns” and 4 were classified as “high risk.” The basic characteristics of included studies are presented in Table 1, and further details about each study are listed in Multimedia Appendix 4 (see also [28-71]).

Among the 91 observational studies, 38 conducted surveys, 42 analyzed data from the internet, 10 used interviews, and the remaining 1 performed mixed methods of cross-sectional survey and interview. In terms of study region, the American region was the most studied (n=25), followed by the European (n=19) and Eastern Mediterranean (n=11) regions. By contrast, the African (n=6), South-East Asian (n=5), and Western Pacific (n=4) regions were the less investigated.

According to the prior defined attitudes toward COVID-19 vaccines in study populations, survey- and internet-based studies can be divided into the following: survey-based study on general population (n=31), survey-based study on antivaccine/vaccine hesitant population (n=4), internet-based study on general online data (n=22), and internet-based study on antivaccine/vaccine hesitant data (n=19). A total of 3 survey-based studies and 1 internet-based study reported results from both general population/data and antivaccine/vaccine hesitant population/data.

Figure 1. Flow diagram.

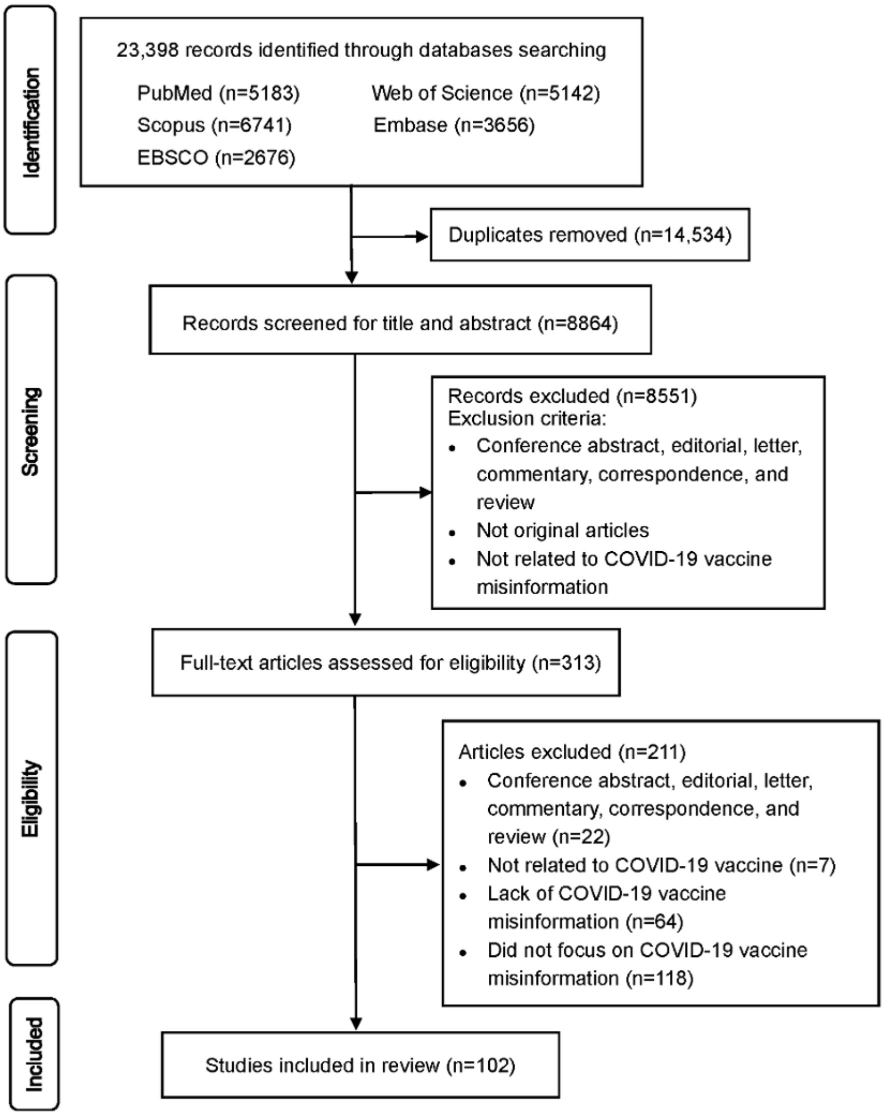


Table 1. Basic characteristics of included articles.

Study characteristics	Studies, n
Study type	
Observational study	
Survey-based study	38
Internet-based study	42
Interview	10
Mixed methods study	1
Experimental study	
Randomized trial	9
Quasi trial	2
Risk of bias	
Low risk	90
Some concern	8
High risk	4
Study phase[a]	
No report	8
Phase 1	28
Phase 2	49
Phase 3	1
Region	
Worldwide[b]	26
African	6
American	25
Eastern Mediterranean	11
European	19
South-East Asian	5
Western Pacific	4
More than 1 region[c]	6
Prior defined attitudes toward COVID-19 vaccines[d] **(n=80)**	
Survey-based study	
General population	31
Antivaccine/vaccine hesitant population	4
Both groups above	3
Internet-based study	
General online data	22
Antivaccine/vaccine hesitant data	19
Both groups above	1
Platform for internet-based studies (n=42)	
Social media	
Twitter	20
YouTube	5
Facebook	4

Study characteristics	Studies, n
TikTok	1
Parler	1
Multiple social media platforms[e]	3
Google Trends	1
Google Images	1
Internet news	1
Online article database	1
Multiple platforms[f]	1

[a] Study phase was defined as follows: phase 1, from November 2019 to December 8, 2020; phase 2, from December 8, 2020, to November 26, 2021; phase 3 from November 26, 2021, to the latest ending date of data collection in included studies.

[b] Studies covering worldwide, including internet-based study with no limit in data collection (n=24), and survey-based study in multiple countries (n=2).

[c] Studies covering more than 1 region: 3 studies on American, European, and Western Pacific regions; and another 3 on American and European regions.

[d] Study populations differ by their prior defined attitudes toward COVID-19 vaccines: studies reporting misinformation among the vaccine hesitancy or refusal group (“antivaccine/vaccine hesitant group” and “antivaccine/vaccine hesitant data” in the table), and studies which did not prior define the vaccine hesitancy or refusal group (“general population” and “general online data” in the table).

[e] Multiple social media platforms in 3 studies [28,72,73] were Facebook, Instagram, and Twitter; Instagram and Facebook; and YouTube, Twitter, Facebook, and Instagram.

[f] Multiple platforms included Google, Google Fact Check, Facebook, YouTube, Twitter, fact-checking agency websites, and websites of television and newspaper.

Most internet-based studies used social media platforms as data source, of which the most frequently discussed was Twitter (n=20), followed by YouTube (n=5), Facebook (n=4), TikTok (n=1), and Parler (n=1). Besides social media, 4 studies used general internet information, such as Google and online news database, as data source. A total of 4 studies used data from multiple platforms or multiple social media. Most internet-based studies did not limit the region of data source (n=24). Among those internet-based studies reporting study regions, the Americas (n=8) and Europe (n=6) were more commonly addressed, whereas the Western Pacific (n=2), Eastern Mediterranean (n=1), Africa (n=1) were less studied, leaving no internet-based study in South-East Asia. Most internet-based studies analyzed information written in English (n=33), whereas non-English information were less studied: 3 for Spanish, 2 for Arabic, 1 for both English and Spanish together, and 1 each for Chinese, Italian, and Turkish.

Types, Sources, and Prevalence of Misinformation

About 90% (91/102, 89.2%) of studies mentioned the types of COVID-19 vaccine–related misinformation. Using the coding framework mentioned in the “Methods” section, we divided vaccine-related misinformation into 7 types: conspiracy, concerns on vaccine safety and efficacy, no need for vaccines, morality, liberty, humor, and overstatement. We further divided these types into 54 different contents (Table 2). Conspiracy, being discussed in 77 studies, was the most commonly studied misinformation, and it could be further specified as vaccine existence conspiracy, political conspiracy, vaccine development and promotion conspiracy, and conspiracy related to a certain group. Concerns on vaccine safety and efficacy, emphasizing the unsubstantiated concerns on safety and efficacy, were the secondary most studied misinformation (n=63). Misinformation about the necessity of vaccines (no need for vaccines) was found in 23 studies, and its subtypes included simple claims such as vaccines are unnecessary, preference of natural immunity and protective behavior, claims about mild COVID-19, COVID-19 denial, underestimation of personal need, and overstatement of vaccine refusal. Misinformation on morality and liberty was found in 15 and 13 studies, respectively; 2 studies investigated humor and another 2 found the overstatement on the effect of COVID-19 vaccines.

Vaccine-related misinformation changed in its contents over time. We compared the misinformation among 3 phases, and found that some conspiracy theories and concerns on vaccine safety circulated in all 3 phases: depopulation, control people, microchip for monitor/control, and financial incentives behind vaccine development in terms of conspiracy; vaccines would alter DNA, and cause fertility, death, or other diseases; the worry of no one would be responsible for potential side effects; beliefs in alternatives such as natural immunity and protection behaviors; referring to vaccines as immoral human experiments and comprising fetal tissue; and individual freedom. Meanwhile, there were emerging and disappearing topics on misinformation. The claims that vaccines were a hoax/fraud only appeared in phase 1. With the progress of COVID-19 vaccination, especially after the massive immunization (phase 2), misinformation about registration (false claims that some vaccines were rejected), against government control (such as fake news that some government will make COVID-19 vaccination mandatory), and conspiracy about vaccine promotion (claims that physicians or people will get financial benefit through promoting vaccination, and that the news about celebrities getting vaccinated are fake: they did not really get vaccinated; rather, they were injected with saline) emerged.

Table 2. Types and contents of COVID-19 vaccine–related misinformation.

Type of misinformation and subtype	Explanation	Contents
Conspiracy		
Vaccine existence conspiracy	Conspiracy about the existence of COVID-19 vaccines	• No vaccine: COVID-19 vaccines are a conspiracy, hoax, or fraud.
Political conspiracy	Conspiracy about political purpose related to the vaccines	• Device for track/control: Vaccines contain microchip/nanochips and are used for tracking or control people. • Population control and new world order: The vaccines are used for the control of population or to create new world order. • Depopulation: Vaccines are used for reducing population or for genocide. • Bioweapon: Vaccines are a bioweapon. • Other conspiracy to a certain subject: Certain people or groups are behind the vaccines for their own good, such as government officials, pharmacy company, Bill Gates.
Vaccine development and promotion conspiracy	Conspiracy about COVID-19 in vaccine development and vaccination promotion	• COVID-19 for vaccines: COVID-19 vaccines existed before the virus, or COVID-19 was made to enforce vaccination. • Financial incentive in developing: Vaccines are created only for the profit of pharmaceutical companies and government. • Political incentive in developing: Vaccines are approved because of political pressure, or because pharmaceutical company has bought/coerced the government. • No enough evidence: Vaccines are untested or not tested enough. • Rush in development: Vaccines are rushed in development, thus cannot be trusted. • Fabricated vaccine efficacy: Data on vaccine efficacy/effectiveness are fabricated. • Cover-up side effect data: Data about the vaccine side effect or death are kept secret. • Rejection in registration: Vaccines (which are approved in fact) are rejected in registration for concern. • Paid for promotion: Physicians issue vaccines for financial profit. • Playacting in promotion: Public figures are vaccinated with inert substances (eg, saline).
Conspiracy related to a certain group	The conspiracy that vaccines will be harmful, used to control, or be tested in a specific group	• Conspiracy of a certain country/region: People in certain country or region are guinea pigs. Vaccines are less effective or is fake in these regions. (eg, African region, Muslim nations). • Conspiracy of ethnic minorities: Vaccines are harmful or for killing ethnic minorities; ethnic minorities are guinea pigs (eg, Asian, Black people). • Conspiracy of other certain groups: Vaccines are used to reduce these people (eg, elder people, less educated people, low-income group).
Concerns on vaccine safety and efficacy		
Effectiveness	Beliefs that vaccines are ineffective	• Vaccines would not work: Vaccines will not work (a simple claim without further explanation). • Ineffective in certain groups: Vaccines will not be effective on people with comorbidities such as diabetes, hypertension. • Ineffective claim from authority: Government officers or doctors admit vaccines would not work. • New strain for cover-up: The new variant of COVID-19 is a hoax and is a cover-up for ineffectiveness of the vaccines. • Fake vaccines: Vaccines are water. • Same as other vaccines: COVID-9 vaccines are no different from the flu vaccine. • Fail before vaccination: Vaccines cannot be preserved properly.

Type of misinformation and subtype	Explanation	Contents
Safety	Unsubstantiated safety concern	• Cause COVID-19: Vaccines will make people catch COVID-19. • Worse than COVID-19: Vaccines are more dangerous than the disease itself. • Alter DNA: Vaccines will change people’s DNA. • Infertility and offspring: Vaccines will make women infertile or affect their offspring. • Cause death: Vaccines kill people. • Cause other diseases: Vaccines make people get chronic disease, autism, autoimmune disease, paralysis, cancer, physical destruction, impotency, etc. • Change people: Vaccines make people magnetic, or turn people into robots, vampires, zombies, etc. • Poison: Vaccines contain poisonous materials, such as mercury, toxic ingredients, chemicals. • Live virus: COVID-19 vaccines contain a “live strain” of the virus. • Unspecified danger: Vaccines are dangerous (a simple claim without further explanation). • Side effect responsibility: No one is responsible for the potential side effects of the vaccines.
No need for vaccines		
Unnecessary	A simple claim that COVID-19 vaccines are unnecessary without explanation	• Unnecessary: COVID-19 vaccines are not needed.
Natural immunity and protective behavior	Beliefs that it is better to get natural immunity or get immunity through naturally protective ways, etc.	• Alternative: Prefer or believe in natural immunity, herd immunity, or protective behaviors.
Mild COVID-19	Believe COVID is a mild disease and no need for vaccines	• Mild COVID-19: COVID-19 is not dangerous; thus, vaccines are not needed.
COVID-19 denial	Denial of COVID-19 pandemic leading to COVID-19 vaccines	• COVID-19 denial: COVID-19 is a hoax, or a fraud, thus no vaccines are needed.
Personal need	Beliefs that it is unnecessary for the vaccines because of their health status	• Underlying disease: I get autoimmune disease, thus cannot take vaccines. • Immune from infection: I am already immune from a past COVID-19 infection. • Not a risk group: I am young/healthy/a certain blood type, thus have a low risk of getting COVID-19 or developing serious disease.
Overstatement of vaccine refusal	Exaggeration about vaccine refusal rate in other places	• Refusal exaggeration: 1 in 6 people will refuse COVID-19 vaccines.
Morality		
Religion	Vaccines go against religious belief, or are viewed as a sign of demon	• Religion: Vaccines may go against religious beliefs. Devil led people to receive vaccines.
Human experiment	Vaccine campaign is a human experiment	• Human experiment: It is a part of a secret research. Vaccinated ones are guinea pigs.
Fetal tissue	Fetal remains in vaccines	• Fetal tissue: COVID-19 vaccines are made from cells of aborted fetuses.
Liberty		
Against mandatory vaccines/control	False claims about mandatory vaccination, and refusal to obey “control”	• Mandatory vaccination: They are forcing us. • Against government control: Opposition to the government’s control.
Support for freedom	Claims that vaccine infringes individual freedom	• Freedom: Vaccines are an attempt to take away personal freedom.
Ignore consent	Concerns that children will be forced to get vaccinated without parents’ consent	• Ignore consent: Children getting vaccinated without parental consent.
Humor		

Type of misinformation and subtype	Explanation	Contents
N/A[a]	A humorous but exaggerated way to express unsubstantiated vaccine rumor	- Humor: Parody/meme of an adverse reaction.
Overstatement		
Overstatement of the protection/progress of vaccines	Exaggeration about the effect and the scientific progress in COVID-19 vaccines	- Overstatement: COVID-19 vaccines are ready; able to cure a patient within 3 hours. After getting the vaccines, we will not have the infection anymore. After getting the COVID-19 vaccines, one can stop wearing the mask and taking safety precaution. As vaccines for COVID-19 have been developed, we can make vaccines for the common cold, HIV, and others.

[a] N/A: not applicable.

Eight studies surveyed the sources of misinformation. Three studies reported social media as main source [74-76]. Two studies found that family and friends also played a role [74,77]. One study measured proportions of different sources: social media (57.6%), followed by family or friends (13.1%), and television (5.7%) [74]. On social media, misinformation was generated majorly by antivaccine groups or well-known antivaccine individuals [78-80]. Two studies found that online celebrity tended to contribute more misinformation [80,81]. One Twitter-based study showed that the highly polarized antivaccine information was mainly from political and nonmedical users, while health care workers were less engaged in COVID-19 vaccine conversation on online platforms [78].

A total of 57 studies reported the prevalence of misinformation (Table 3), including 27 survey-based studies and 30 internet-based studies; further details on the prevalence reported in each study are presented in Multimedia Appendix 5 (see also [29,30,34,36,37,41-43,46-50,52-57,60,62,67,68,70,71]). Among 27 surveys, some investigated both general population and antivaccine groups. The 24 surveys on general population reported a prevalence of general misinformation ranging from 2.5% to 55.4%; concerning the prevalence of specific types of misinformation, conspiracy beliefs varied from 2.5% to 48.4%, concerns about vaccine safety and efficacy from 2.78% to 55.4%, "no need for vaccine" from 3.8% to 28.1%, morality from 1.4% to 20.6%, and liberty from 6% to 36.3%. A total of 6 surveys [29-32,76,82] on antivaccine/vaccine hesitant groups reported a higher prevalence of misinformation: conspiracy, 6.0%-22%; concerns on vaccine safety and efficacy, 12.2%-96.7%; and no need for vaccines, 6.3%-70.4%. The 96.7% prevalence rate was simply driven by 1 study [30] that reported the prevalence on concerns and no need for vaccines; after excluding this outlier, the prevalence of misinformation varied between 6.0% and 20.1%.

Among 30 internet-based studies, 16 investigated general online data and reported the prevalence of general misinformation ranging from 0.1% to 41.3%; for the prevalence of specific types of misinformation, conspiracy ranged from 5.3% to 21.7%, concerns on vaccine safety and efficacy from 0.4% to 11.1%, humor at 26%, no need for vaccines at 10.1%, and vaccine morality from 3.9% to 20.6% (Table 3). Regarding social media platforms, YouTube showed the lowest prevalence of misinformation, ranging from 1.7% to 10.7%, and the prevalence on Twitter varied from 0.4% to 56%; the prevalence of conspiracy is 3% on TikTok and 12.44%-40% on Reddit. Humor accounted for up 26% of all COVID-19 vaccine–related videos in TikTok. In addition, antivaccine/hesitant online posts from the remaining 16 internet-based studies reported a 0.5%-56% prevalence of misinformation; specifically, conspiracy ranged from 3.9% to 55.4%, concerns on vaccine safety and efficacy from 1.3% to 44.8%, no need for vaccines from 0.5% to 3.7%, morality from 2% to 10.4%, and liberty from 5% to 46%.

Some studies reported that the prevalence of misinformation fluctuated over time, and both increasing and decreasing trends were noted [83,84]. One study showed that the change in misinformation was in close association with news or events related to vaccine developments [81].

Table 3. Prevalence of COVID-19 vaccine–related misinformation.

Study design and subtype of misinformation	Study, n	Prevalence, %
Surveys on general population	24	2.5-55.4
Conspiracy	22	2.5-48.4
Concerns on vaccine safety and efficacy	12	2.78-55.4
No need for vaccines	6	3.8-28.1
Morality	2	1.4-20.6
Liberty	2	6-36.3
Surveys on the antivaccine/vaccine hesitant group	6	6.0-96.7
Conspiracy	4	6.0-22
Concerns on vaccine safety and efficacy	4	12.2-96.7
No need for vaccines	5	6.3-70.4
Internet-based studies on general online data	16	0.1-41.3
Conspiracy	8	5.3-21.7
Concerns on vaccine safety and efficacy	3	0.4-11.1
No need for vaccines	1	10.1
Morality	2	3.9-20.6
Liberty	1	41.1
Humor	1	26
Internet-based studies on antivaccine/vaccine hesitant data	15	0.5-56
Conspiracy	10	3.9-55.4
Concerns on vaccine safety and efficacy	8	1.3-44.8
No need for vaccines	3	0.5-3.7
Morality	2	2-10.4
Liberty	5	5-46

Factors That Affect Misinformation Believing and Spreading

In total, 37 articles reported factors that affected the believing and spreading of misinformation. Textbox 1 summarizes individual characteristics and information-seeking behaviors that affect (enhance or reduce) beliefs in misinformation (n=25) as well as misinformation features that promote its spread (n=14).

In terms of geographical areas, 2 studies found a relatively higher prevalence of misinformation in the United States [85,86]. Another study found that Wyoming had the highest level of misinformation in the United States [87]. Living in a village instead of a city was also found to relate to misinformation [88]. Concerning demographic factors, a younger age was found to enhance beliefs in misinformation in 5 studies [89]. The role of sex remained controversial: 5 studies found that females were more likely to accept conspiracy theories than males, while 2 studies found that males were more fragile [90-92]. In the United States, 2 studies found ethnic minorities were related to beliefs in misinformation [93,94], whereas 1 study found that Whites were more susceptible to misinformation than racial/ethnic minorities [91]. Christians or those with a higher level of religiosity were more likely to be influenced by misinformation [91].

Social economic status and occupation also affected beliefs in misinformation. Most studies documented that lower education and economic status were linked to accepting misinformation [89,90,92-100], although 1 study in African and Middle East countries indicated that individuals with higher education levels believed rumors such as changes in human genome due to vaccines [101]. Medical workers were less susceptible to misinformation than the general population in the UK and Jordan, and among medical workers, juniors were more susceptible to misinformation [94,102]. In addition, the unemployed were less likely to trust misinformation than the employed in the UK [91].

For political orientation, right-wing and conservative ideology would increase the conspiracy belief [90,97], and in the United States, Republicans were more likely to accept vaccine conspiracy than Independents [82,103,104]. For disease and migration experience, having basic diseases, no experience of COVID-19 infection or vaccination, and migration could enhance beliefs in misinformation [97,101]. For psychological status and beliefs, depression, perceived ethnic discrimination, national narcissism, and general conspiracy-mindedness were more likely to accept misinformation [93,97,105,106].

Textbox 1. Factors that affect misinformation believing and spreading.

Individual characteristics that "enhance" beliefs in misinformation

Geographic areas:

- United States
- Living in a village instead of a city

Age:

- Younger adults

Sex:

- Mixed influence

Ethnicity:

- Mixed influence in the United States

Religion:

- Christians
- Higher level of religiosity

Socioeconomic status:

- Lower education level
- Lower income
- Lower social economic status

Occupation:

- Employed people

Political orientation:

- Republicans
- Conservatives
- Far-right
- Not being affected or vaccinated

Disease experience:

- Having basic diseases

Migration experience:

- Having migration experience

Psychology status and beliefs:

- Depression
- National narcissism
- Perceived ethnic discrimination
- General conspiracy-mindedness

Information-related behaviors that "enhance" beliefs in misinformation

Source of information:

- Social media
- Conservative media
- Family and friends

Awareness of information:

- Feeling less informed about science
- Perceive higher incidence of fake news

Online posting:

- Posting more online

Individual characteristics and information-related behaviors that “reduce” beliefs in misinformation

Occupation:

- Medical workers

Channels to accessing information:

- Taking lectures about COVID-19 vaccines
- Trust celebrities for information

Features of misinformation that “promote” its spread

Type of misinformation:

- Safety concern
- Conspiracy
- Efficacy concern

Content of misinformation:

- Positive valence (positive emotion)
- Concreteness

Format of misinformation:

- Number of hashtags
- Language or format that mimics news/scientific reports

Source of misinformation:

- Antivaccine group
- Social media influencer
- Health care workers
- Unregulated bot

Platform:

- Different platforms present different misinformation
- Newer social media

The following information-seeking behaviors also enhanced beliefs in misinformation: usage of social media and conservative media [1,92,107,108], posting more online [109], feeling less informed about science [97], and trusting friends or family for COVID-19 information [91]. People who take lectures about COVID-19 vaccines or trust celebrities for COVID-19 information tended to refuse misinformation [91].

In addition, some features of misinformation promoted its spread. Safety concerns, conspiracy, and efficacy concerns were reported as the most popular misinformation types [85,110]. Misinformation posts that had a higher level of positive valence and concreteness would spread [111]. The hashtags and the language or format mimicking news/scientific reports helped misinformation spread [72,110,111]. Misinformation from antivaccine groups, social media influencers, health care workers such as Sherri Tenpenny, and unregulated bots can speed its spread [79,81,112-114]. Each social media has its prevailing misinformation topics [115], and a newer social media platform was more likely to bolster vaccine conspiracy [110].

Impact of Misinformation on Vaccine Hesitancy and Behaviors

A total of 29 studies indicated that misinformation is related to vaccine hesitancy or negative vaccine perception, and 3 experimental studies also supported this finding [75,91,116].

Misinformation would ignite concerns and fears about the safety profile of vaccines and lead to vaccine hesitancy and refusal [94]. Four studies found a negative relationship between misinformation and vaccine uptake rate [84,87,89,99]. One experimental study in the UK and United States verified that the misinformation exposure significantly reduced both intentions of self-vaccination and vaccination to protect others by around 6% [91]. Skepticism also attenuated the effect of public service messages on promoting vaccination willingness [116]. Besides vaccination behaviors, misinformation reduced the uptake of self-protection behaviors such as mask wearing, distancing, and compliance with health guidance [1,90,107].

The impacts of misinformation may change across different study phases. Impacts of misinformation on protective behaviors were mainly studied in phase 1 (n=4). One study in phase 3 showed a negative relationship between some religious beliefs on COVID-19 vaccines and protective behaviors such as masks and distancing remained [99]. After mass vaccination (phase 2), the negative impact of misinformation on vaccine uptake rate was uncovered and lasted up to phase 3. The negative association of misinformation with attitudes and intention of taking vaccines was consistent among all 3 phases.

Interventions to Address Misinformation

Overview

A total of 65 studies proposed solutions to address COVID-19 vaccine misinformation, and 9 studies assessed the effects of various interventions to combat misinformation. Table 4 summarizes the proposed solutions according to our 5-step framework and specific solutions in each study are detailed in Multimedia Appendix 6 (see also [28, 31, 32, 37-41, 43-45, 48, 50, 52, 54, 56-58, 60, 63-66]).

Table 4. Proposed solutions to address COVID-19 vaccine-related misinformation.

Framework and meaning	Detailed solutions	No. of articles	Main actors	Targets
Identifying misinformation				
Identifying misinformation through diverse channels	• Routine monitoring and fact-checking • Investigation	5	Government and health officers	Misinformation
Regulating producers and distributors				
Regulating the source of misinformation	• Legislation and national campaign	3	Political power	Sources of misinformation
Cutting production and distribution				
Cutting the circulation of misinformation	• Technical response: more effective moderation policies	5	News media, social media	Misinformation
	• Economic response: The boycott of harmful content by advertisers	1	News media, social media	Misinformation
	• Curatorial response: dissemination of truth to debunk misinformation	4	News media, social media	The public
Supporting the target audiences of misinformation				
Improving target population's health literacy and helping the public identify misinformation	• Tailored intervention to improve health literacy by suggestion, advocacy, and health education	13	Multiple	Different misinformed groups
	• Empowerment by public awareness campaigns and credibility warning labels	3	Multiple	The public
Disseminating trustworthy information				
Implementing multitiered strategies to convey the message by multiple media	N/A[a]	23	Multiple: public health agencies, medical professionals, religious leaders, etc.	The public

[a] N/A: not applicable.

Identifying Misinformation

This step contained the routine fact-check and monitoring and investigation of misinformation. A total of 7 articles emphasized that government and health officers should develop a public health surveillance system to track the emergence of misinformation and the outlets of antivaccine groups through data mining applications [97].

Regulating Producers and Distributors

Producers and distributors of misinformation need to be regulated by political power. A total of 6 articles mentioned that policy and legal actions should be implemented by the government.

Cutting Production and Distribution

This step underlined the reaction to communication platforms, including news media and social media. A total of 11 articles in our review specified this part, which consisted of technical, economic, and curatorial responses. Technical response required social media companies to build and execute more effective moderation policies, such as checking information, altering keyword searches, redirecting individuals to correct sources, banning overt conspiracy groups, and flagging or rapidly removing misinformation [104]. The economic response could be either the boycott of harmful content by advertisers or the monetization limit of the channels producing misinformation [24,108]. The curatorial response emphasized that messages should be directly debunking misinformation. The “backfire effect” (ie, factual counterargument entrenches false beliefs) was found when using debunking messages to address misinformation [117,118]. Therefore, pre-debunking message (inoculation message) before public message communication should be piloted [73,119].

Supporting the Target Audiences of Misinformation

A total of 20 studies outlined the importance of supporting the target audiences. Tailored interventions targeting different misinformed groups were considered an important approach to counter the misinformation. Empowerment of the public was recommended to improve health literacy and awareness. Public awareness campaigns were also mentioned in 2 studies [74,102].

Disseminating Trustworthy Information

This step seems similar to the curatorial response, yet it focused not only on messages that directly debunk misinformation, but also aimed to pass correct and scientific information. Most studies (n=39) mentioned or exemplified how to disseminate evidence-based information, and they recommended to mobilize trusted medical professionals and scientists to engage in social media conversation. Health care workers and public health agencies should engage in social media and learn how to produce short scientific videos [78,102,120]. Further, collaboration with social media influencers allowed a wider reach to the public [121,122]. Both social media and mass media should be utilized to spread information [123], which can express vaccine support, emphasize scientific procedure, appeal to altruism, picture the meaning and importance of vaccination, create a sense of companionship in the battle of infodemic, or encourage participation through peer pressure. Evidence-based messages should be delivered in ways understandable to individuals from a variety of socioeconomic and educational background. Therefore, youth, religious leaders, community stakeholders, faith-based organizations, and schools could be engaged to co-design culturally compelling and context-appropriate messages [85].

Although many studies proposed solutions, only 9 assessed effects of some interventions to combat misinformation. One observational study detected a limited active impact of the policies developed by Twitter [113]. Among 8 experimental studies, 4 assessed the effect of debunking message/video, 3 assessed the effect of inoculation messages, and 1 assessed the effect of warning tag/cover. All 4 experiments found that debunking message/video could reduce the belief in misinformation [118,124-126]. Yet, 1 experiment found that this effect worked well for people without strong beliefs in misinformation; when it comes to people who strongly believe in misinformation, the result became counterproductive because it may evoke small backfiring effects of vaccination intention [118]. One experiment further showed that partisans exposed to ingroup media (media held the same political preference as participants) perceived debunking messages as more credible and held higher engagement [125]. For inoculation messages, 2 experiments found that simple inoculation message/video could protect people from misinformation [119,127], while the remaining 1 found it had no significant effect, but was useful when combined with viewing or writing comments on the inoculation message [128]. In terms of warning tag/cover, interstitial warnings or cover warnings, which require individuals to click through to continue, were found to be more effective to help participants identify misinformation while warning tags showed no effect [104].

Discussion

Principal Findings

This review revealed the features, influencing factors, impacts, and solutions for COVID-19 vaccine misinformation and provided evidence to combat vaccine-related misinformation. The included articles were predominantly from American and European regions, and there was less evidence from African, South-East Asian, and Western Pacific regions. Social media was considered as the main source of COVID-19 vaccine misinformation. Conspiracy, concerns on vaccine safety and efficacy, and no need for vaccines were the most prevalent types of misinformation.

Our review documented the high but wide prevalence of COVID-19 vaccine misinformation. The great variation may be due to the ambiguity in misinformation classification and the difference across social media platforms, regions, and study participants. On social media, antivaccine or misinformation tweets accounted for 0.1%-41.3% of all vaccine-related tweets. Its prevalence varied from 2.5% to 55.4% among the general population, which was consistent with a previous review on general COVID-19 misinformation [18]. Another systematic review also illustrated that the prevalence of misinformation on general vaccines ranged from 1% to 65% on social media between 2012 and 2018 [129]. The prevalence of misinformation posed challenges to the COVID-19 vaccination. As the most prevalent type of COVID-19 vaccine misinformation, conspiracies were usually presented as half-truths, which made it hard to recognize them as misinformation. Therefore, dealing with conspiracy or skepticism should be the priority to combat the prevalent vaccine misinformation.

Social media is considered a “double-edged sword” to inform the public [7]. The endorsement of COVID-19 misinformation was strongly associated with the information sources. People who rely on print media and mainstream print were less apt to endorse COVID-19 misinformation, while the use of social media was positively associated with misperceptions regarding COVID-19 facts. In our review, as the main source of COVID-19 vaccine misinformation, social media usage was related to misinformation believing. Because of its wide usage, social media should be used to debunk misinformation and disseminate trustworthy information [121]. Public health authorities and health professionals should change their low engagement status and be more actively engaged in COVID-19 vaccine conversation on online platforms [78,120]. Social media platforms should also make rules and policies to combat misinformation [113].

The global acceptance of COVID-19 vaccine remains at a low level [2]. Misinformation is negatively associated with vaccine acceptance and self-protection behaviors. In our review, geographic areas, demographic characteristics, education, occupation, political orientation, disease and migration experience, psychological status and beliefs, and information-seeking behaviors were found to influence the believing and spreading of vaccine misinformation. However, there is less clarity on the induced pathway between these factors and misinformation. For example, the higher prevalence of COVID-19 vaccine misinformation in African and West Mediterranean regions was associated with feeling less informed, and the less informed were further linked to poor socioeconomic status, low education level, and lack of information [130]. Future research should thus focus on how these factors influence the believing and spreading of misinformation, and verify targeted populations and intervention strategies to combat vaccine misinformation.

It is urgent to implement effective intervention strategies to combat COVID-19 vaccine misinformation. Through the solutions proposed by 65 included studies, we constructed a 5-step framework to address vaccine-related misinformation. The dissemination of trustworthy information was the most frequently mentioned, followed by supporting target audiences of misinformation and cutting its production and distribution, whereas identifying misinformation and regulating its producers and distributors were less mentioned. However, most proposed solutions were not verified regarding their effects on misinformation. With the limited experimental studies, the debunking message/video was considered as effective interventions against misinformation during the COVID-19 pandemic [118,124-126].

Combating misinformation is a persistent and complex work. More scientific evidence is needed to support misinformation surveillance, punishment of misinformation producers, and dissemination of trustworthy information [24]. As many countries reported the first case of mpox, another epidemic in the shadow of the COVID-19 pandemic is looming [131]. In this context, it is particularly important to synthesize potential intervention strategies to combat vaccine misinformation for future anti-infodemic campaign.

Limitations

Our review has several limitations. First, most included studies were observational, leading to low quality of evidence. More interventional studies are needed to evaluate the effect of misinformation interventions and identify effective interventions. Second, we only included publications in English, and social media platforms in non-English languages were not included in our review. The feature, prevalence, and impact of COVID-19 vaccine misinformation may differ by languages. Third, nearly one-half of the included articles were based on the internet; however, the internet penetration rates are low in low-income countries. Our review may thus not provide enough data and evidence of COVID-19 vaccine misinformation in low-income countries. Fourth, the included studies were published from January 2020 to August 2022, therefore this period may be insufficient to observe and explore the long-term impact of COVID-19 vaccine misinformation.

Conclusion

Our review provides comprehensive and up-to-date evidence on COVID-19 vaccine misinformation and helps responses to vaccine infodemic in future pandemics. Its prevalence was high but widely varied worldwide. The most frequent misinformation types were conspiracy and concerns about vaccine safety and efficacy. Information features, information-seeking behaviors, and demographic factors influenced the spreading and believing of misinformation. More evidence is needed to verify potential intervention strategies to combat vaccine misinformation.

Acknowledgments

This work was supported by the Soft Science Research Project of Shanghai (22692107600).

Authors' Contributions

ZH conceived the review. ZH, SZ, and XZ refined the search strategy. SZ and XZ searched for articles. SZ, SH, XZ, QW, HZ, YZ screened and extracted the articles. SZ and SH assessed the quality of the articles and analyzed data. SZ and SH wrote the first draft of the manuscript. SS revised the manuscript. ZH supervised the review process and prepared the final draft for submission. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[DOCX File , 25 KB - publichealth_v9i1e40201_app1.docx]

Multimedia Appendix 2

Search strategy for 5 peer-reviewed databases.

[DOCX File , 21 KB - publichealth_v9i1e40201_app2.docx]

Multimedia Appendix 3

Quality assessment criteria.

[DOCX File , 30 KB - publichealth_v9i1e40201_app3.docx]

Multimedia Appendix 4

Characteristics of included studies.

[DOCX File , 63 KB - publichealth_v9i1e40201_app4.docx]

Multimedia Appendix 5

Prevalence of COVID-19 vaccine–related misinformation by study design.

[DOCX File , 38 KB - publichealth_v9i1e40201_app5.docx]

Multimedia Appendix 6

Proposed solutions to address misinformation in each study.

[DOCX File , 36 KB - publichealth_v9i1e40201_app6.docx]

References

1. Romer D, Jamieson KH. Patterns of Media Use, Strength of Belief in COVID-19 Conspiracy Theories, and the Prevention of COVID-19 From March to July 2020 in the United States: Survey Study. J Med Internet Res 2021 Apr 27;23(4):e25215 [FREE Full text] [doi: 10.2196/25215] [Medline: 33857008]
2. Wang Q, Hu S, Du F, Zang S, Xing Y, Qu Z, et al. Mapping global acceptance and uptake of COVID-19 vaccination: A systematic review and meta-analysis. Commun Med (Lond) 2022 Sep 12;2(1):113 [FREE Full text] [doi: 10.1038/s43856-022-00177-6] [Medline: 36101704]
3. Lin C, Tu P, Beitsch LM. Confidence and Receptivity for COVID-19 Vaccines: A Rapid Systematic Review. Vaccines (Basel) 2020 Dec 30;9(1):16 [FREE Full text] [doi: 10.3390/vaccines9010016] [Medline: 33396832]
4. Lu L, Mok BWY, Chen LL, Chan JMC, Tsang OTY, Lam BHS, et al. Neutralization of Severe Acute Respiratory Syndrome Coronavirus 2 Omicron Variant by Sera From BNT162b2 or CoronaVac Vaccine Recipients. Clin Infect Dis 2022 Aug 24;75(1):e822-e826. [doi: 10.1093/cid/ciab1041] [Medline: 34915551]
5. Arede M, Bravo-Araya M, Bouchard É, Singh Gill G, Plajer V, Shehraj A, et al. Combating Vaccine Hesitancy: Teaching the Next Generation to Navigate Through the Post Truth Era. Front Public Health 2018;6:381 [FREE Full text] [doi: 10.3389/fpubh.2018.00381] [Medline: 30693276]
6. World Health Organization. Infodemic. 2020. URL: https://www.who.int/health-topics/infodemic [accessed 2022-11-11]
7. Cuan-Baltazar JY, Muñoz-Perez MJ, Robledo-Vega C, Pérez-Zepeda MF, Soto-Vega E. Misinformation of COVID-19 on the Internet: Infodemiology Study. JMIR Public Health Surveill 2020 Apr 09;6(2):e18444 [FREE Full text] [doi: 10.2196/18444] [Medline: 32250960]
8. Zhang S, Pian W, Ma F, Ni Z, Liu Y. Characterizing the COVID-19 Infodemic on Chinese Social Media: Exploratory Study. JMIR Public Health Surveill 2021 Feb 05;7(2):e26090 [FREE Full text] [doi: 10.2196/26090] [Medline: 33460391]
9. Pan American Health Organization. The potential of frequently used information technologies during the pandemic. 2020. URL: https://iris.paho.org/handle/10665.2/52021 [accessed 2022-11-11]
10. Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 2010 Feb 17;28(7):1709-1716. [doi: 10.1016/j.vaccine.2009.12.022] [Medline: 20045099]
11. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. Misinformation and Its Correction: Continued Influence and Successful Debiasing. Psychol Sci Public Interest 2012 Dec;13(3):106-131. [doi: 10.1177/1529100612451018] [Medline: 26173286]

12. Sabahelzain MM, Moukhyer M, Dubé E, Hardan A, van den Borne B, Bosma H. Towards a further understanding of measles vaccine hesitancy in Khartoum state, Sudan: A qualitative study. PLoS One 2019;14(6):e0213882 [FREE Full text] [doi: 10.1371/journal.pone.0213882] [Medline: 31220092]
13. Chou WS, Budenz A. Considering Emotion in COVID-19 Vaccine Communication: Addressing Vaccine Hesitancy and Fostering Vaccine Confidence. Health Commun 2020 Dec;35(14):1718-1722. [doi: 10.1080/10410236.2020.1838096] [Medline: 33124475]
14. Blane JT, Bellutta D, Carley KM. Social-Cyber Maneuvers During the COVID-19 Vaccine Initial Rollout: Content Analysis of Tweets. J Med Internet Res 2022 Mar 07;24(3):e34040 [FREE Full text] [doi: 10.2196/34040] [Medline: 35044302]
15. Herrera-Peco I, Jiménez-Gómez B, Romero Magdalena CS, Deudero JJ, García-Puente M, Benítez De Gracia E, et al. Antivaccine Movement and COVID-19 Negationism: A Content Analysis of Spanish-Written Messages on Twitter. Vaccines (Basel) 2021 Jun 15;9(6):656 [FREE Full text] [doi: 10.3390/vaccines9060656] [Medline: 34203946]
16. Timothy G, Axel B, Guangnan Z. Like a virus: the coordinated spread of coronavirus disinformation. 2020. URL: https://doi.org/APO-305864 [accessed 2022-11-11]
17. Gruber A, Ghiringhelli M, Edri O, Abboud Y, Shiti A, Shaheen N, et al. Literature Review and Knowledge Distribution During an Outbreak: A Methodology for Managing Infodemics. Acad Med 2021 Jul 01;96(7):1005-1009. [doi: 10.1097/ACM.0000000000004073] [Medline: 33788792]
18. Gabarron E, Oyeyemi SO, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bull World Health Organ 2021 Jun 01;99(6):455-463A [FREE Full text] [doi: 10.2471/BLT.20.276782] [Medline: 34108756]
19. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation About COVID-19 Vaccines on Social Media: Rapid Review. J Med Internet Res 2022 Aug 04;24(8):e37367 [FREE Full text] [doi: 10.2196/37367] [Medline: 35816685]
20. Lieneck C, Heinemann K, Patel J, Huynh H, Leafblad A, Moreno E, et al. Facilitators and Barriers of COVID-19 Vaccine Promotion on Social Media in the United States: A Systematic Review. Healthcare (Basel) 2022 Feb 08;10(2):321 [FREE Full text] [doi: 10.3390/healthcare10020321] [Medline: 35206935]
21. Centers FDC. Myths and Facts about COVID-19 Vaccines. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/facts.html [accessed 2022-11-11]
22. Taylor D. A Timeline of the Coronavirus Pandemic. 2021. URL: https://www.nytimes.com/article/coronavirus-timeline.html [accessed 2022-11-11]
23. World Health Organization. Classification of Omicron (B1.1.529): SARS-CoV-2 Variant of Concern. 2021. URL: https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern [accessed 2022-11-11]
24. Bontcheva K, Posetti J. Disinfodemic: dissecting responses to COVID-19 disinformation. URL: https://unesdoc.unesco.org/ark:/48223/pf0000374417 [accessed 2022-11-11]
25. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open 2016 Dec 08;6(12):e011458 [FREE Full text] [doi: 10.1136/bmjopen-2016-011458] [Medline: 27932337]
26. Critical Appraisal Skills Programme. CASP Qualitative Checklist. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2022-11-11]
27. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
28. Wong LP, Lin Y, Alias H, Bakar SA, Zhao Q, Hu Z. COVID-19 Anti-Vaccine Sentiments: Analyses of Comments from Social Media. Healthcare (Basel) 2021 Dec 09;9(11):1530 [FREE Full text] [doi: 10.3390/healthcare9111530] [Medline: 34828576]
29. Magadmi RM, Kamel FO. Beliefs and barriers associated with COVID-19 vaccination among the general population in Saudi Arabia. BMC Public Health 2021 Jul 21;21(1):1438. [doi: 10.1186/s12889-021-11501-5] [Medline: 34289817]
30. Yousuf H, van der Linden S, van Essen T, Gommers D, Scherder E, Narula J, et al. Dutch Perspectives Toward Governmental Trust, Vaccination, Myths, and Knowledge About Vaccines and COVID-19. JAMA Netw Open 2021 Dec 01;4(12):e2140529 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.40529] [Medline: 34967887]
31. Kricorian K, Civen R, Equils O. COVID-19 vaccine hesitancy: misinformation and perceptions of vaccine safety. Hum Vaccin Immunother 2022 Dec 31;18(1):1950504 [FREE Full text] [doi: 10.1080/21645515.2021.1950504] [Medline: 34325612]
32. Bíró-Nagy A, Szászi ÁJ. The roots of COVID-19 vaccine hesitancy: evidence from Hungary. J Behav Med 2022 May 14:1-16 [FREE Full text] [doi: 10.1007/s10865-022-00314-5] [Medline: 35567729]
33. Alliheibi FM, Omar A, Al-Horais N. Opinion Mining of Saudi Responses to COVID-19 Vaccines on Twitter: A Computational Linguistic Approach. International Journal of Advanced Computer Science and Applications 2021;12(6):72-78. [doi: 10.14569/IJACSA.2021.0120610]
34. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. Hum Vaccin Immunother 2021 Aug 03;17(8):2373-2377 [FREE Full text] [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]

35. Bishnoi HR, Avasthi RD, Sharma SK. A Qualitative Study To Myths And Reality About Covid-19 Vaccination Among People In Western Rajasthan. Journal of Cardiovascular Disease Research 2022 Apr;12(1):476-482 [FREE Full text]
36. Charquero-Ballester M, Walter J, Nissen I, Bechmann A. Different types of COVID-19 misinformation have different emotional valence on Twitter. Big Data & Society 2021 Sep 22;8(2):205395172110412. [doi: 10.1177/20539517211041279]
37. Dereje N, Tesfaye A, Tamene B, Alemeshet D, Abe H, Tesfa N, et al. COVID-19 vaccine hesitancy in Addis Ababa, Ethiopia: a mixed-method study. BMJ Open 2022 May 30;12(5):e052432 [FREE Full text] [doi: 10.1136/bmjopen-2021-052432] [Medline: 35636790]
38. Ebrahimi OV, Johnson MS, Ebling S, Amundsen OM, Halsøy Ø, Hoffart A, et al. Risk, Trust, and Flawed Assumptions: Vaccine Hesitancy During the COVID-19 Pandemic. Front Public Health 2021;9:700213 [FREE Full text] [doi: 10.3389/fpubh.2021.700213] [Medline: 34277557]
39. Garcia J, Vargas N, de la Torre C, Magana Alvarez M, Clark JL. Engaging Latino Families About COVID-19 Vaccines: A Qualitative Study Conducted in Oregon, USA. Health Educ Behav 2021 Dec;48(6):747-757 [FREE Full text] [doi: 10.1177/10901981211045937] [Medline: 34596462]
40. Gesualdo F, Parisi L, Croci I, Comunello F, Parente A, Russo L, et al. How the Italian Twitter Conversation on Vaccines Changed During the First Phase of the Pandemic: A Mixed-Method Analysis. Front Public Health 2022;10:824465 [FREE Full text] [doi: 10.3389/fpubh.2022.824465] [Medline: 35664110]
41. Griffith J, Marani H, Monkman H. COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework. J Med Internet Res 2021 Apr 13;23(4):e26874 [FREE Full text] [doi: 10.2196/26874] [Medline: 33769946]
42. Bou Hamdan M, Singh S, Polavarapu M, Jordan T, Melhem N. COVID-19 vaccine hesitancy among university students in Lebanon. Epidemiol. Infect 2021 Nov 02;149:e242. [doi: 10.1017/s0950268821002314]
43. Hernández-García I, Gascón-Giménez I, Gascón-Giménez A, Giménez-Júlvez T. Information in Spanish on YouTube about Covid-19 vaccines. Hum Vaccin Immunother 2021 Nov 02;17(11):3916-3921 [FREE Full text] [doi: 10.1080/21645515.2021.1957416] [Medline: 34375570]
44. Hossain MB, Alam MZ, Islam MS, Sultan S, Faysal MM, Rima S, et al. COVID-19 vaccine hesitancy among the adult population in Bangladesh: A nationwide cross-sectional survey. PLoS One 2021;16(12):e0260821. [doi: 10.1371/journal.pone.0260821] [Medline: 34882726]
45. Kant R, Varea RR, Titifanue J. COVID-19 vaccine online misinformation in Fiji: Preliminary findings. Pacific Journalism Review 2022;27(1-2):47-62 [FREE Full text] [doi: 10.24135/pjr.v27i1&2.1189]
46. Küçükali H, Ataç Ö, Palteki AS, Tokaç AZ, Hayran O. Vaccine Hesitancy and Anti-Vaccination Attitudes during the Start of COVID-19 Vaccination Program: A Content Analysis on Twitter Data. Vaccines (Basel) 2022 Jan 21;10(2):161 [FREE Full text] [doi: 10.3390/vaccines10020161] [Medline: 35214620]
47. Kulkarni M, Khurana K. Acceptance and myths regarding covid vaccination among general population. Medical Science 2022;26(124):ms255e2245 [FREE Full text] [doi: 10.54905/disssi/v26i124/ms255e2245]
48. Kumar N, Corpus I, Hans M, Harle N, Yang N, McDonald C, et al. COVID-19 vaccine perceptions in the initial phases of US vaccine roll-out: an observational study on reddit. BMC Public Health 2022 Mar 07;22(1):446 [FREE Full text] [doi: 10.1186/s12889-022-12824-7] [Medline: 35255881]
49. Laforet PE, Basch CH, Tang H. Understanding the content of COVID-19 vaccination and pregnancy videos on YouTube: An analysis of videos published at the start of the vaccine rollout. Hum Vaccin Immunother 2022 Nov 30;18(5):2066935 [FREE Full text] [doi: 10.1080/21645515.2022.2066935] [Medline: 35507867]
50. Lamptey E, Senkyire EK, Dorcas S, Benita DA, Boakye EO, Ikome T, et al. Exploring the myths surrounding the COVID-19 vaccines in Africa: the study to investigate their impacts on acceptance using online survey and social media. Clin Exp Vaccine Res 2022 May;11(2):193-208 [FREE Full text] [doi: 10.7774/cevr.2022.11.2.193] [Medline: 35799880]
51. Lee SK, Sun J, Jang S, Connelly S. Misinformation of COVID-19 vaccines and vaccine hesitancy. Sci Rep 2022 Aug 11;12(1):13681 [FREE Full text] [doi: 10.1038/s41598-022-17430-6] [Medline: 35953500]
52. Li HOY, Pastukhova E, Brandts-Longtin O, Tan MG, Kirchhof MG. YouTube as a source of misinformation on COVID-19 vaccination: a systematic analysis. BMJ Glob Health 2022 Mar;7(3):e008334 [FREE Full text] [doi: 10.1136/bmjgh-2021-008334] [Medline: 35264318]
53. Liao TF. Understanding Anti-COVID-19 Vaccination Protest Slogans in the US. Front. Commun 2022 Jun 30;7:941872 [FREE Full text] [doi: 10.3389/fcomm.2022.941872]
54. Lurie P, Adams J, Lynas M, Stockert K, Carlyle RC, Pisani A, et al. COVID-19 vaccine misinformation in English-language news media: retrospective cohort study. BMJ Open 2022 Jun 01;12(6):e058956 [FREE Full text] [doi: 10.1136/bmjopen-2021-058956] [Medline: 35649595]
55. Muric G, Wu Y, Ferrara E. COVID-19 Vaccine Hesitancy on Social Media: Building a Public Twitter Data Set of Antivaccine Content, Vaccine Misinformation, and Conspiracies. JMIR Public Health Surveill 2021 Nov 17;7(11):e30642 [FREE Full text] [doi: 10.2196/30642] [Medline: 34653016]
56. Neely SR, Eldredge C, Ersing R, Remington C. Vaccine Hesitancy and Exposure to Misinformation: a Survey Analysis. J Gen Intern Med 2022 Jan;37(1):179-187 [FREE Full text] [doi: 10.1007/s11606-021-07171-z] [Medline: 34671900]

57. Obreja DM. Narrative communication regarding the Covid-19 vaccine: a thematic analysis of comments on Romanian official Facebook page "RO Vaccinare". SN Soc Sci 2022;2(8):119 [FREE Full text] [doi: 10.1007/s43545-022-00427-3] [Medline: 35875608]
58. Okoro O, Kennedy J, Simmons G, Vosen EC, Allen K, Singer D, et al. Exploring the Scope and Dimensions of Vaccine Hesitancy and Resistance to Enhance COVID-19 Vaccination in Black Communities. J Racial Ethn Health Disparities 2022 Dec;9(6):2117-2130 [FREE Full text] [doi: 10.1007/s40615-021-01150-0] [Medline: 34553340]
59. Oleksy T, Wnuk A, Gambin M, Łyś A, Bargiel-Matusiewicz K, Pisula E. Barriers and facilitators of willingness to vaccinate against COVID-19: Role of prosociality, authoritarianism and conspiracy mentality. A four-wave longitudinal study. Pers Individ Dif 2022 May;190:111524 [FREE Full text] [doi: 10.1016/j.paid.2022.111524] [Medline: 35068638]
60. Sallam M, Dababseh D, Eid H, Hasan H, Taim D, Al-Mahzoum K, et al. Low COVID-19 Vaccine Acceptance Is Correlated with Conspiracy Beliefs among University Students in Jordan. Int J Environ Res Public Health 2021 Mar 01;18(5):2407 [FREE Full text] [doi: 10.3390/ijerph18052407] [Medline: 33804558]
61. Savolainen R. Assessing the credibility of COVID-19 vaccine mis/disinformation in online discussion. Journal of Information Science 2021 Aug 19:016555152110406. [doi: 10.1177/01655515211040653]
62. Thelwall M, Kousha K, Thelwall S. Covid-19 vaccine hesitancy on English-language Twitter. El Profesional de la Información 2021;30(2):1-13. [doi: 10.3145/epi.2021.mar.12]
63. Wang CW, de Jong EP, Faure JA, Ellington JL, Chen CHS, Chan CC. A matter of trust: a qualitative comparison of the determinants of COVID-19 vaccine hesitancy in Taiwan, the United States, the Netherlands, and Haiti. Hum Vaccin Immunother 2022 Nov 30;18(5):2050121 [FREE Full text] [doi: 10.1080/21645515.2022.2050121] [Medline: 35349382]
64. Watermeyer J, Scott M, Kapueja L, Ware LJ. To trust or not to trust: an exploratory qualitative study of personal and community perceptions of vaccines amongst a group of young community healthcare workers in Soweto, South Africa. Health Policy Plan 2022 Oct 12;37(9):1167-1176 [FREE Full text] [doi: 10.1093/heapol/czac060] [Medline: 35880606]
65. Wawrzuta D, Jaworski M, Gotlib J, Panczyk M. What Arguments against COVID-19 Vaccines Run on Facebook in Poland: Content Analysis of Comments. Vaccines (Basel) 2021 May 10;9(5):481 [FREE Full text] [doi: 10.3390/vaccines9050481] [Medline: 34068500]
66. Wonodi C, Obi-Jeff C, Adewumi F, Keluo-Udeke SC, Gur-Arie R, Krubiner C, et al. Conspiracy theories and misinformation about COVID-19 in Nigeria: Implications for vaccine demand generation communications. Vaccine 2022 Mar 18;40(13):2114-2121 [FREE Full text] [doi: 10.1016/j.vaccine.2022.02.005] [Medline: 35153088]
67. Wu W, Lyu H, Luo J. Characterizing Discourse about COVID-19 Vaccines: A Reddit Version of the Pandemic Story. Health Data Sci 2021;2021:9837856 [FREE Full text] [doi: 10.34133/2021/9837856] [Medline: 36405359]
68. Yang Z, Luo X, Jia H. Is It All a Conspiracy? Conspiracy Theories and People's Attitude to COVID-19 Vaccination. Vaccines (Basel) 2021 Sep 22;9(10):1051 [FREE Full text] [doi: 10.3390/vaccines9101051] [Medline: 34696159]
69. Yaseen MO, Saif A, Khan TM, Yaseen M, Saif A, Bukhsh A, et al. A qualitative insight into the perceptions and COVID-19 vaccine hesitancy among Pakistani pharmacists. Hum Vaccin Immunother 2022 Dec 31;18(1):2031455 [FREE Full text] [doi: 10.1080/21645515.2022.2031455] [Medline: 35192781]
70. Abbas SW, Zareen SF, Nisar S, Farooq A, Rasheed A, Saleem MU. COVID-19 Vaccines: Community Myths Vs Facts. Pakistan Armed Forces Medical Journal 2022;72(2):497-500 [FREE Full text] [doi: 10.51253/pafmj.v72i2.6970] [Medline: 157247889]
71. Al-Rawi A, Fakida A, Grounds K. Investigation of COVID-19 Misinformation in Arabic on Twitter: Content Analysis. JMIR Infodemiology 2022;2(2):e37007-e37500 [FREE Full text] [doi: 10.2196/37007]
72. Ngai CSB, Singh RG, Yao L. Impact of COVID-19 Vaccine Misinformation on Social Media Virality: Content Analysis of Message Themes and Writing Strategies. J Med Internet Res 2022 Jul 06;24(7):e37806. [doi: 10.2196/37806] [Medline: 35731969]
73. Hughes B, Miller-Idriss C, Piltch-Loeb R, Goldberg B, White K, Criezis M, et al. Development of a Codebook of Online Anti-Vaccination Rhetoric to Manage COVID-19 Vaccine Misinformation. Int J Environ Res Public Health 2021 Jul 15;18(14):7556 [FREE Full text] [doi: 10.3390/ijerph18147556] [Medline: 34300005]
74. Arshad MS, Hussain I, Mahmood T, Hayat K, Majeed A, Imran I, et al. A National Survey to Assess the COVID-19 Vaccine-Related Conspiracy Beliefs, Acceptability, Preference, and Willingness to Pay among the General Population of Pakistan. Vaccines (Basel) 2021 Jul 01;9(7):720 [FREE Full text] [doi: 10.3390/vaccines9070720] [Medline: 34358136]
75. Thaker J, Subramanian A. Exposure to COVID-19 Vaccine Hesitancy Is as Impactful as Vaccine Misinformation in Inducing a Decline in Vaccination Intentions in New Zealand: Results from Pre-Post Between-Groups Randomized Block Experiment. Front. Commun 2021 Aug 19;6:721982. [doi: 10.3389/fcomm.2021.721982]
76. Solís Arce JS, Warren SS, Meriggi NF, Scacco A, McMurry N, Voors M, et al. COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. Nat Med 2021 Aug;27(8):1385-1394 [FREE Full text] [doi: 10.1038/s41591-021-01454-y] [Medline: 34272499]
77. Magee L, Knights F, Mckechnie DGJ, Al-Bedaery R, Razai MS. Facilitators and barriers to COVID-19 vaccination uptake among ethnic minorities: A qualitative study in primary care. PLoS One 2022;17(7):e0270504 [FREE Full text] [doi: 10.1371/journal.pone.0270504] [Medline: 35802738]

78. Hernandez RG, Hagen L, Walker K, O'Leary H, Lengacher C. The COVID-19 vaccine social media : healthcare providers' missed dose in addressing misinformation and vaccine hesitancy. Hum Vaccin Immunother 2021 Sep 02;17(9):2962-2964 [FREE Full text] [doi: 10.1080/21645515.2021.1912551] [Medline: 33890838]
79. Jemielniak D, Krempovych Y. An analysis of AstraZeneca COVID-19 vaccine misinformation and fear mongering on Twitter. Public Health 2021 Nov;200:4-6 [FREE Full text] [doi: 10.1016/j.puhe.2021.08.019] [Medline: 34628307]
80. Calac AJ, Haupt MR, Li Z, Mackey T. Spread of COVID-19 Vaccine Misinformation in the Ninth Inning: Retrospective Observational Infodemic Study. JMIR Infodemiology 2022;2(1):e33587 [FREE Full text] [doi: 10.2196/33587] [Medline: 35320982]
81. Jiang LC, Chu TH, Sun M. Characterization of Vaccine Tweets During the Early Stage of the COVID-19 Outbreak in the United States: Topic Modeling Analysis. JMIR Infodemiology 2021;1(1):e25636 [FREE Full text] [doi: 10.2196/25636] [Medline: 34604707]
82. Ruiz JB, Bell RA. Predictors of intention to vaccinate against COVID-19: Results of a nationwide survey. Vaccine 2021 Feb 12;39(7):1080-1086 [FREE Full text] [doi: 10.1016/j.vaccine.2021.01.010] [Medline: 33461833]
83. An L, Russell DM, Mihalcea R, Bacon E, Huffman S, Resnicow K. Online Search Behavior Related to COVID-19 Vaccines: Infodemiology Study. JMIR Infodemiology 2021;1(1):e32127 [FREE Full text] [doi: 10.2196/32127] [Medline: 34841200]
84. Chen Y, Chen Y, Yang K, Lai F, Huang C, Chen Y, et al. The Prevalence and Impact of Fake News on COVID-19 Vaccination in Taiwan: Retrospective Study of Digital Media. J Med Internet Res 2022 Apr 26;24(4):e36830 [FREE Full text] [doi: 10.2196/36830] [Medline: 35380546]
85. Islam MS, Kamal AM, Kabir A, Southern DL, Khan SH, Hasan SMM, et al. COVID-19 vaccine rumors and conspiracy theories: The need for cognitive inoculation against misinformation to improve vaccine adherence. PLoS One 2021;16(5):e0251605 [FREE Full text] [doi: 10.1371/journal.pone.0251605] [Medline: 33979412]
86. Savoia E, Harriman NW, Piltch-Loeb R, Bonetti M, Toffolutti V, Testa MA. Exploring the Association between Misinformation Endorsement, Opinions on the Government Response, Risk Perception, and COVID-19 Vaccine Hesitancy in the US, Canada, and Italy. Vaccines (Basel) 2022 Apr 23;10(5):671 [FREE Full text] [doi: 10.3390/vaccines10050671] [Medline: 35632427]
87. Pierri F, Perry BL, DeVerna MR, Yang K, Flammini A, Menczer F, et al. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. Sci Rep 2022 Apr 26;12(1):5966 [FREE Full text] [doi: 10.1038/s41598-022-10070-w] [Medline: 35474313]
88. Hammad AM, Al-Qerem W, Abu Zaid A, Khdair SI, Hall FS. Misconceptions Related to COVID 19 Vaccines Among the Jordanian Population: Myth and Public Health. Disaster Med Public Health Prep 2022 Jun 08:1-8 [FREE Full text] [doi: 10.1017/dmp.2022.143] [Medline: 35673791]
89. Eshel Y, Kimhi S, Marciano H, Adini B. Conspiracy claims and secret intentions as predictors of psychological coping and vaccine uptake during the COVID-19 pandemic. J Psychiatr Res 2022 Jul;151:311-318 [FREE Full text] [doi: 10.1016/j.jpsychires.2022.04.042] [Medline: 35526447]
90. Hornsey MJ, Chapman CM, Alvarez B, Bentley S, Salvador Casara BG, Crimston CR, et al. To what extent are conspiracy theorists concerned for self versus others? A COVID-19 test case. Eur J Soc Psychol 2021 Mar;51(2):285-293 [FREE Full text] [doi: 10.1002/ejsp.2737] [Medline: 33821057]
91. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021 Mar;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
92. Sallam M, Dababseh D, Eid H, Al-Mahzoum K, Al-Haidar A, Taim D, et al. High Rates of COVID-19 Vaccine Hesitancy and Its Association with Conspiracy Beliefs: A Study in Jordan and Kuwait among Other Arab Countries. Vaccines (Basel) 2021 Jan 12;9(1):42 [FREE Full text] [doi: 10.3390/vaccines9010042] [Medline: 33445581]
93. Andrade G. Covid-19 vaccine hesitancy, conspiracist beliefs, paranoid ideation and perceived ethnic discrimination in a sample of University students in Venezuela. Vaccine 2021 Nov 16;39(47):6837-6842 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.037] [Medline: 34711439]
94. Manby L, Dowrick A, Karia A, Maio L, Buck C, Singleton G, et al. Healthcare workers’ perceptions and attitudes towards the UK’s COVID-19 vaccination programme: a rapid qualitative appraisal. BMJ Open 2022 Feb 15;12(2):e051775. [doi: 10.1136/bmjopen-2021-051775]
95. Mahmud MR, Bin Reza R, Ahmed SZ. The effects of misinformation on COVID-19 vaccine hesitancy in Bangladesh. GKMC 2021 Oct 24:ahead-of-print. [doi: 10.1108/gkmc-05-2021-0080]
96. Buturoiu R, Udrea G, Oprea D, Corbu N. Who Believes in Conspiracy Theories about the COVID-19 Pandemic in Romania? An Analysis of Conspiracy Theories Believers’ Profiles. Societies 2021 Nov 13;11(4):138. [doi: 10.3390/soc11040138]
97. Jensen EA, Pfleger A, Herbig L, Wagoner B, Lorenz L, Watzlawik M. What Drives Belief in Vaccination Conspiracy Theories in Germany? Front. Commun 2021 May 25;6:105. [doi: 10.3389/fcomm.2021.678335]
98. Caycho-Rodríguez T, Ventura-León J, Valencia PD, Vilca LW, Carbajal-León C, Reyes-Bossio M, et al. What Is the Support for Conspiracy Beliefs About COVID-19 Vaccines in Latin America? A Prospective Exploratory Study in 13 Countries. Front. Psychol 2022 May 6;13:105. [doi: 10.3389/fpsyg.2022.855713]

99. Exline JJ, Pait KC, Wilt JA, Schutt WA. Demonic and Divine Attributions around COVID-19 Vaccines: Links with Vaccine Attitudes and Behaviors, QAnon and Conspiracy Beliefs, Anger, Spiritual Struggles, Religious and Political Variables, and Supernatural and Apocalyptic Beliefs. Religions 2022 Jun 06;13(6):519. [doi: 10.3390/rel13060519]
100. Lin F, Chen X, Cheng EW. Contextualized impacts of an infodemic on vaccine hesitancy: The moderating role of socioeconomic and cultural factors. Information Processing & Management 2022 Sep;59(5):103013. [doi: 10.1016/j.ipm.2022.103013]
101. Asadi Faezi N, Gholizadeh P, Sanogo M, Oumarou A, Mohamed MN, Cissoko Y, et al. Peoples' attitude toward COVID-19 vaccine, acceptance, and social trust among African and Middle East countries. Health Promot Perspect 2021;11(2):171-178 [FREE Full text] [doi: 10.34172/hpp.2021.21] [Medline: 34195040]
102. Aloweidi A, Bsisu I, Suleiman A, Abu-Halaweh S, Almustafa M, Aqel M, et al. Hesitancy towards COVID-19 Vaccines: An Analytical Cross-Sectional Study. Int J Environ Res Public Health 2021 May 12;18(10):5111 [FREE Full text] [doi: 10.3390/ijerph18105111] [Medline: 34065888]
103. Pierri F, Perry B, DeVerna M, Yang K, Flammini A, Menczer F, et al. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. Sci Rep 2022 Apr 26;12(1):5966 [FREE Full text] [doi: 10.1038/s41598-022-10070-w] [Medline: 35474313]
104. Sharevski F, Alsaadi R, Jachim P, Pieroni E. Misinformation warnings: Twitter's soft moderation effects on COVID-19 vaccine belief echoes. Comput Secur 2022 Mar;114:102577 [FREE Full text] [doi: 10.1016/j.cose.2021.102577] [Medline: 34934255]
105. Marchlewska M, Hamer K, Baran M, Górska P, Kaniasty K. COVID-19: Why Do People Refuse Vaccination? The Role of Social Identities and Conspiracy Beliefs: Evidence from Nationwide Samples of Polish Adults. Vaccines (Basel) 2022 Feb 10;10(2):268 [FREE Full text] [doi: 10.3390/vaccines10020268] [Medline: 35214726]
106. Perlis RH, Ognyanova K, Santillana M, Lin J, Druckman J, Lazer D, et al. Association of Major Depressive Symptoms With Endorsement of COVID-19 Vaccine Misinformation Among US Adults. JAMA Netw Open 2022 Jan 04;5(1):e2145697 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.45697] [Medline: 35061036]
107. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. Soc Sci Med 2020 Oct;263:113356 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113356] [Medline: 32967786]
108. Jennings W, Stoker G, Bunting H, Valgarðsson VO, Gaskell J, Devine D, et al. Lack of Trust, Conspiracy Beliefs, and Social Media Use Predict COVID-19 Vaccine Hesitancy. Vaccines (Basel) 2021 Jun 03;9(6):593 [FREE Full text] [doi: 10.3390/vaccines9060593] [Medline: 34204971]
109. Kumar M, Madhumathi J, Gayathri K, A Rozario AG, Vijayaprabha R, Balusamy M, et al. Community voices around COVID-19 vaccine in Chennai, India: A qualitative exploration during early phase of vaccine rollout. Indian J Med Res 2022;155(5&6):451-460. [doi: 10.4103/ijmr.ijmr_668_22] [Medline: 35975352]
110. Baines A, Ittefaq M, Abwao M. #Scamdemic, #Plandemic, or #Scaredemic: What Parler Social Media Platform Tells Us about COVID-19 Vaccine. Vaccines (Basel) 2021 Apr 22;9(5):421 [FREE Full text] [doi: 10.3390/vaccines9050421] [Medline: 33922343]
111. Saini V, Liang L, Yang Y, Le HM, Wu C. The Association Between Dissemination and Characteristics of Pro-/Anti-COVID-19 Vaccine Messages on Twitter: Application of the Elaboration Likelihood Model. JMIR Infodemiology 2022;2(1):e37077 [FREE Full text] [doi: 10.2196/37077] [Medline: 35783451]
112. Lanyi K, Green R, Craig D, Marshall C. COVID-19 Vaccine Hesitancy: Analysing Twitter to Identify Barriers to Vaccination in a Low Uptake Region of the UK. Front Digit Health 2021;3:804855 [FREE Full text] [doi: 10.3389/fdgth.2021.804855] [Medline: 35141699]
113. Larrondo-Ureta A, Fernández S, Morales-i-Gras J. Desinformación, vacunas y Covid-19. Análisis de la infodemia y la conversación digital en Twitter. RLCS 2021 Jun 07(79):1-18. [doi: 10.4185/rlcs-2021-1504]
114. Kalichman S, Eaton L, Earnshaw V, Brousseau N. Faster than warp speed: early attention to COVD-19 by anti-vaccine groups on Facebook. J Public Health (Oxf) 2022 Mar 07;44(1):e96-e105 [FREE Full text] [doi: 10.1093/pubmed/fdab093] [Medline: 33837428]
115. Wawrzuta D, Klejdysz J, Jaworski M, Gotlib J, Panczyk M. Attitudes toward COVID-19 Vaccination on Social Media: A Cross-Platform Analysis. Vaccines (Basel) 2022 Jul 27;10(8):1190 [FREE Full text] [doi: 10.3390/vaccines10081190] [Medline: 35893839]
116. Jin Q, Raza SH, Yousaf M, Zaman U, Siang JMLD. Can Communication Strategies Combat COVID-19 Vaccine Hesitancy with Trade-Off between Public Service Messages and Public Skepticism? Experimental Evidence from Pakistan. Vaccines (Basel) 2021 Jul 07;9(7):757 [FREE Full text] [doi: 10.3390/vaccines9070757] [Medline: 34358173]
117. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020 Oct;7(10):201199 [FREE Full text] [doi: 10.1098/rsos.201199] [Medline: 33204475]
118. Helfers A, Ebersbach M. The differential effects of a governmental debunking campaign concerning COVID-19 vaccination misinformation. Journal of Communication in Healthcare 2022 Mar 17:1-9. [doi: 10.1080/17538068.2022.2047497]

119. Vivion M, Anassour Laouan Sidi E, Betsch C, Dionne M, Dubé E, Driedger SM, et al. Prebunking messaging to inoculate against COVID-19 vaccine misinformation: an effective strategy for public health. Journal of Communication in Healthcare 2022 Mar 04;15(3):232-242. [doi: 10.1080/17538068.2022.2044606]
120. Basch CH, Hillyer GC, Zagnit EA, Basch CE. YouTube coverage of COVID-19 vaccine development: implications for awareness and uptake. Hum Vaccin Immunother 2020 Nov 01;16(11):2582-2585 [FREE Full text] [doi: 10.1080/21645515.2020.1790280] [Medline: 32701403]
121. Chan C, Sounderajah V, Daniels E, Acharya A, Clarke J, Yalamanchili S, et al. The Reliability and Quality of YouTube Videos as a Source of Public Health Information Regarding COVID-19 Vaccination: Cross-sectional Study. JMIR Public Health Surveill 2021 Jul 08;7(7):e29942 [FREE Full text] [doi: 10.2196/29942] [Medline: 34081599]
122. Criss S, Nguyen TT, Norton S, Virani I, Titherington E, Tillmanns EL, et al. Advocacy, Hesitancy, and Equity: Exploring U.S. Race-Related Discussions of the COVID-19 Vaccine on Twitter. Int J Environ Res Public Health 2021 May 26;18(11):5693 [FREE Full text] [doi: 10.3390/ijerph18115693] [Medline: 34073291]
123. Ginossar T, Cruickshank IJ, Zheleva E, Sulskis J, Berger-Wolf T. Cross-platform spread: vaccine-related content, sources, and conspiracy theories in YouTube videos shared in early Twitter COVID-19 conversations. Hum Vaccin Immunother 2022 Dec 31;18(1):1-13 [FREE Full text] [doi: 10.1080/21645515.2021.2003647] [Medline: 35061560]
124. Yousuf H, van der Linden S, Bredius L, Ted van Essen GA, Sweep G, Preminger Z, et al. A media intervention applying debunking versus non-debunking content to combat vaccine misinformation in elderly in the Netherlands: A digital randomised trial. EClinicalMedicine 2021 May;35:100881 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100881] [Medline: 34124631]
125. Lu S, Zhong L. From Believing to Sharingxamining the Effects of Partisan Media's Correction of COVID-19 Vaccine Misinformation. INTERNATIONAL JOURNAL OF COMMUNICATION. 2022 2022 2022 Jun 15:16.
126. Talabi F, Ugbor I, Talabi M, Ugwuoke J, Oloyede D, Aiyesimoju A, et al. Effect of a social media-based counselling intervention in countering fake news on COVID-19 vaccine in Nigeria. Health Promot Int 2022 Apr 29;37(2):daab140. [doi: 10.1093/heapro/daab140] [Medline: 34510187]
127. Piltch-Loeb R, Su M, Hughes B, Testa M, Goldberg B, Braddock K, et al. Testing the Efficacy of Attitudinal Inoculation Videos to Enhance COVID-19 Vaccine Acceptance: Quasi-Experimental Intervention Trial. JMIR Public Health Surveill 2022 Jun 20;8(6):e34615 [FREE Full text] [doi: 10.2196/34615] [Medline: 35483050]
128. Dai Y, Jia W, Fu L, Sun M, Jiang LC. The effects of self-generated and other-generated eWOM in inoculating against misinformation. Telematics and Informatics 2022 Jul;71:101835. [doi: 10.1016/j.tele.2022.101835]
129. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of Health Misinformation on Social Media: Systematic Review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
130. Abdulai A, Tiffere A, Adam F, Kabanunye MM. COVID-19 information-related digital literacy among online health consumers in a low-income country. Int J Med Inform 2021 Jan;145:104322. [doi: 10.1016/j.ijmedinf.2020.104322] [Medline: 33157342]
131. Hasan S, Saeed S. Monkeypox Disease: An Emerging Public Health Concern in the Shadow of COVID-19 Pandemic: An Update. Trop Med Infect Dis 2022 Oct 03;7(10):283 [FREE Full text] [doi: 10.3390/tropicalmed7100283] [Medline: 36288024]

Abbreviations

AXIS: Appraisal tool for Cross-Sectional Studies
CDC: Centers for Disease Control and Prevention
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RoB 2: version 2 of the Cochrane risk-of-bias tool for randomized trials
UNESCO: United Nations Educational, Scientific and Cultural Organization

Edited by T Sanchez, A Mavragani; submitted 10.06.22; peer-reviewed by J Lyu, S Lin; comments to author 12.08.22; revised version received 10.11.22; accepted 29.11.22; published 11.01.23.

Please cite as:
Zhao S, Hu S, Zhou X, Song S, Wang Q, Zheng H, Zhang Y, Hou Z
The Prevalence, Features, Influencing Factors, and Solutions for COVID-19 Vaccine Misinformation: Systematic Review
JMIR Public Health Surveill 2023;9:e40201
URL: https://publichealth.jmir.org/2023/1/e40201
doi:10.2196/40201
PMID:36469911

Original Paper

The Impact of Mask Mandates on Face Mask Use During the COVID-19 Pandemic: Longitudinal Survey Study

Mawuena Binka[1,2], MPH, PhD; Prince Asumadu Adu[1,2], MA, MPH, PhD; Dahn Jeong[1,2], MSc; Nirma Khatri Vadlamudi[3], MPH, PhD; Héctor Alexander Velásquez García[1,2], MD, MS, MPH, PhD; Bushra Mahmood[1,4], MPH, PhD; Terri Buller-Taylor[1], PhD; Michael Otterstatter[1,2], MSc, PhD; Naveed Zafar Janjua[1,2,5], MBBS, MSc, DrPH

[1]British Columbia Centre for Disease Control, Vancouver, BC, Canada

[2]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

[3]Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

[4]Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

[5]Centre for Health Outcomes and Evaluation, St Paul's Hospital, Vancouver, BC, Canada

Corresponding Author:
Mawuena Binka, MPH, PhD
British Columbia Centre for Disease Control
655 W 12th Ave
Vancouver, BC, V5Z 4R4
Canada
Phone: 1 6047072400
Email: mawuena.binka@bccdc.ca

Abstract

Background: Face mask use has been associated with declines in COVID-19 incidence rates worldwide. A handful of studies have examined the factors associated with face mask use in North America during the COVID-19 pandemic; however, much less is known about the patterns of face mask use and the impact of mask mandates during this time. This information could have important policy implications, now and in the event of future pandemics.

Objective: To address existing knowledge gaps, we assessed face mask usage patterns among British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents and evaluated the impact of the provincial mask mandate on these usage patterns.

Methods: Between September 2020 and July 2022, adult British Columbia residents completed the web-based BC-Mix survey, answering questions on the circumstances surrounding face mask use or lack thereof, movement patterns, and COVID-19–related beliefs. Trends in face mask use over time were assessed, and associated factors were evaluated using multivariable logistic regression. A stratified analysis was done to examine effect modification by the provincial mask mandate.

Results: Of the 44,301 respondents, 81.9% reported wearing face masks during the 23-month period. In-store and public transit mask mandates supported monthly face mask usage rates of approximately 80%, which was further bolstered up to 92% with the introduction of the provincial mask mandate. Face mask users mostly visited retail locations (51.8%) and travelled alone by car (49.6%), whereas nonusers mostly traveled by car with others (35.2%) to their destinations—most commonly parks (45.7%). Nonusers of face masks were much more likely to be male than female, especially in retail locations and restaurants, bars, and cafés. In a multivariable logistic regression model adjusted for possible confounders, factors associated with face mask use included age, ethnicity, health region, mode of travel, destination, and time period. The odds of face mask use were 3.68 times greater when the provincial mask mandate was in effect than when it was not (adjusted odds ratio [aOR] 3.68, 95% CI 3.33-4.05). The impact of the mask mandate was greatest in restaurants, bars, or cafés (mandate: aOR 7.35, 95% CI 4.23-12.78 vs no mandate: aOR 2.81, 95% CI 1.50-5.26) and in retail locations (mandate: aOR 19.94, 95% CI 14.86-26.77 vs no mandate: aOR 7.71, 95% CI 5.68-10.46).

Conclusions: Study findings provide added insight into the dynamics of face mask use during the COVID-19 pandemic. Mask mandates supported increased and sustained high face mask usage rates during the first 2 years of the pandemic, having the greatest impact in indoor public locations with limited opportunity for physical distancing targeted by these mandates. These findings highlight the utility of mask mandates in supporting high face mask usage rates during the COVID-19 pandemic.

(JMIR Public Health Surveill 2023;9:e42616) doi:10.2196/42616

KEYWORDS

face mask; face covering; COVID-19; SARS-CoV-2; outbreak; public health; health policy; trend analysis; logistic regression

Introduction

The rapid spread of SARS-CoV-2 worldwide led to the declaration of a global pandemic by the World Health Organization in March 2020 [1,2]. SARS-CoV-2 infection causes COVID-19, which, in extreme cases, results in severe lung damage, multiorgan failure, and death. Person-to-person spread of SARS-CoV-2 is mediated through aerosolized droplets that are generated during activities such as talking, singing, coughing, or sneezing [1,3,4]. When worn appropriately, face masks and other face coverings limit the spread of aerosolized droplets by trapping them within their fibers [5,6]. The utility of face masks and other face coverings in reducing person-to-person transmission of SARS-CoV-2 has been demonstrated in epidemiological and laboratory-based studies, as well as in real-world settings [5,7-12]. This efficacy, alongside the widespread availability and ease of use of face masks has prompted public health officials worldwide to advocate for, or mandate, face mask use in indoor public spaces and in settings with limited opportunity for physical distancing, as part of efforts to control the spread of the virus [13,14].

Public health mandates provide a blanket order for the application of interventions to reduce disease transmission rather than providing a choice for the adoption of those interventions. Hence, these mandates may be perceived as infringing on freedom of choice in those settings. Consequently, mask mandates and recommendations were met with resistance from certain groups [15,16]. The lack of consensus among global political and public health leaders on the need for face masks during the early stages of the COVID-19 pandemic, as well as misinformation and disinformation regarding the utility of masks and potential adverse effects of face mask use, may also explain this resistance [14,16-19]. Specific reasons for the lack of face mask use in a survey conducted among participants from several Western countries included discomfort, difficulty breathing, and skepticism about the ability of face masks to prevent infection [20]. Nevertheless, mask mandates and recommendations have contributed to decreased incidence of COVID-19 cases and related deaths worldwide [21-23]. The advent of COVID-19 vaccines and increasing vaccination coverage has prompted the relaxation of mask mandates and recommendations in various countries worldwide [13,24]. However, recent resurgences in COVID-19 cases in regions where mask mandates were rescinded [25-27] underscore the continued need for the use of face masks in certain regions as global vaccination efforts progress and more is learned about the efficacy of current vaccines in reducing the transmission of new and highly contagious variants. Understanding factors associated with face mask use and quantifying the impact of mask mandates is, therefore, important for health communication and decision-making by public health leadership, especially in the context of repeated outbreaks. Recent studies investigating the factors associated with nonuse of face masks in Canada provided much needed information on the motivation and belief systems underlying face mask use in the country [28,29]. However, limited information is available on face mask usage patterns, with and without provincial mask mandates, during the COVID-19 pandemic in Canada. This information could have important policy implications, now and for future respiratory virus-driven pandemic(s). In this study, we bridge this knowledge gap by assessing face mask usage patterns in the presence and absence of the provincial mask mandate and the factors associated with mask use among respondents of a population-based survey in British Columbia (BC), Canada.

Methods

Context

Initial public health measures to control the spread of COVID-19 were introduced in BC, Canada, on March 18, 2020 [30]; however, the provincial mask mandate requiring face masks in all indoor public spaces did not come into effect until November 19, 2020 [31,32]. Nevertheless, major retail locations in the province mandated the use of face masks between July and August 2020 [33], prior to the provincial mask mandate, as did BC public transit on August 24, 2020 [34]. Due in part to increasing COVID-19 vaccination rates, the provincial mask mandate was lifted on July 1, 2021, although the mandatory requirement for face mask use remained in effect at major retail locations [35]. The provincial mask mandate was reinstated for select indoor public places on August 25, 2021, remaining in effect until March 11, 2022 [36-38]. By April 8, 2022, all other public health requirements, including proof of vaccination for admission to certain locations, were no longer mandated [38].

Study Population and Variable Definitions

The BC COVID-19 Population Mixing Patterns (BC-Mix) survey is an ongoing web-based survey developed to assess population mixing patterns during the COVID-19 pandemic among BC residents [39]. The survey, launched on September 4, 2020, is composed of 94 questions across six key domains: (1) demographic information; (2) COVID-19 testing and results, symptoms, and health behaviors; (3) activities and behavior in and outside of the home; (4) internet and social media use; (5) perceptions and attitudes around COVID-19; and (6) COVID-19 vaccine acceptance (added March 8, 2021). It is administered on the Qualtrics platform to English-speaking persons aged ≥18 years and residing in BC. Anonymous links to the survey were circulated via advertisements placed on Google and social media platforms, namely Instagram, Facebook, WhatsApp, YouTube, and Twitter. Detailed descriptions of survey design, domains, and recruitment methods have been published elsewhere [39]. Participants completed a baseline survey (for first-time respondents), and those who consented were invited to complete shorter follow-up surveys every 2 to 4 weeks.

This analysis was restricted to the baseline responses received between September 4, 2020, and July 31, 2022. Survey respondents who left home at least once the previous day (survey question: "How many times did you leave your home [or property, apartment] yesterday?") were asked whether or not they used a face mask ("Did you use a face mask yesterday?"). Survey respondents who provided valid answers to the face mask question ("Yes," "No," or "Prefer not to answer") were included in this analysis. As people who left their homes either did or did not wear a mask, those who answered "Prefer not to answer" either did not want to anonymously report not wearing masks or did not want to report wearing masks to researchers for some reason. Assuming the former formed the majority of this subgroup and wanting to capture as many types of nonusers of face masks as possible, responses to the face mask use question were recategorized as "Yes" and "No" ("No" + "Prefer not to answer") for the purpose of this study. Other questions addressed ethnicity, education, employment status, location of face mask use, duration of face mask use, number of trips outside the home, distance travelled, destination, and mode of travel (Table S1 in the Multimedia Appendix 1). Time period was grouped by calendar month; thus, the periods during which the provincial mask mandate were in effect were defined as from November 2020 to June 2021 and from September 2021 to February 2022.

Statistical Analyses

Descriptive analyses were done with and without sampling weights. Sampling weights were based on age, sex, geography (Health Authority region), and ethnicity; derived with a weighting adjustment technique [40] using available participant and BC 2016 Canadian Census data; and applied so that response frequencies were representative of the BC population. All comparisons between face mask users (face mask use="Yes") and nonusers of face masks (face mask use="No") were made with weighted data. Chi-square tests were used to ascertain between-group differences in variable distribution. Factors associated with face mask use were assessed with a multivariable logistic regression model, adjusting for time period, age group, sex, ethnicity, destination, number of trips taken, distance travelled, mode of travel, and Health Authority of residence—incorporating sampling weights. The association between the provincial mask mandate and face mask use was also assessed with multivariable logistic regression models, and a stratified analysis was done to examine effect modification by the mandate.

Data preparation, descriptive analyses, and data visualization were done with R statistical software (version 3.5.2; R Foundation for Statistical Computing) [41]. Weighted logistic regression modeling was done with SAS statistical software (version 9.4; SAS Institute) [42]. Statistical significance was assessed at the *P*<.05 level.

Ethical Approval

This study complied with the ethical standards of the Helsinki Declaration. Participation was voluntary and electronic informed consent was sought from all participants on the survey start page. Analytical data sets were deidentified and included no personally identifiable information. Ethical approval for this study was provided by the University of British Columbia Behavioral Research Ethics Board (H20-01785).

Results

Respondent Characteristics

A total of 44,301 respondents were eligible for inclusion in this analysis (see Table 1). Survey respondents who answered the face mask question were mostly male (52.4%), not part of a visible minority group (63.3%), aged 25-34 years (18.81%) and 45-54 years (18.6%), employed full time (33%), and residing in the Fraser Health region (26.3%; Table 1).

Table 1. Proportion of British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents by face mask use (yes, n=36,716; no, n=7585), stratified by demographic characteristics, from September 2020 to July 2022.

Characteristic	Unweighted values		Weighted values (distribution "across" groups)				*P* value[a]	Weighted values (distribution "within" each group)	
	Yes, n	No, n	Yes, n	No, n	Yes, %	No, %		Yes, %	No, %
Age group (years)									
18-24	1095	239	3244	822	8.4	9.6	.25	79.8	20.2
25-34	3459	696	7160	1673	18.6	19.6		81.1	18.9
35-44	5157	1066	6122	1385	15.9	16.2		81.6	18.4
45-54	6409	1159	7216	1553	18.7	18.2		82.3	17.7
55-64	9149	1969	6543	1370	17	16.1		82.7	17.3
65-74	8817	1886	6216	1324	16.1	15.5		82.4	17.6
≥75	2630	570	2008	411	5.2	4.8		83	17
Sex									
Female	29,833	5737	19,052	3362	49.5	39.4	<.001	85	15
Male	6883	1848	19,455	5176	50.5	60.6		79	21
Ethnicity									
Chinese	891	114	4298	541	11.2	6.3	<.001	88.8	11.2
South Asian	604	69	3186	395	8.3	4.6		89	11
Other visible minority	1267	155	2586	326	6.7	3.8		88.8	11.2
Not a visible minority	30,894	6439	23,922	5865	62.1	68.7		80.3	19.7
Other ethnicity	1916	430	3076	820	8	9.6		78.9	21.1
Prefer not to answer	1144	378	1441	591	3.7	6.9		70.9	29.1
Education									
Below high school	592	106	820	154	2.1	1.8	<.001	84.2	15.8
Below bachelor's degree	13,834	2946	13,096	2997	34	35.1		81.4	18.6
University degree	13,597	2676	14,558	2839	37.8	33.2		83.7	16.3
Prefer not to answer or missing	8693	1857	10,034	2549	26.1	29.9		79.7	20.3
Employment status									
Employed full time	10,402	1758	12,975	2544	33.7	29.8	<.001	83.6	16.4
Employed part time	2807	427	2767	434	7.2	5.1		86.4	13.6
Self-employed	2268	562	2522	650	6.5	7.6		79.5	20.5
Unemployed	992	231	1459	362	3.8	4.2		80.1	19.9
Full-time parent or homemaker	648	199	535	176	1.4	2.1		75.3	24.7
Retired	9893	2339	6684	1456	17.4	17.1		82.1	17.9
Student or pupil	475	94	1029	220	2.7	2.6		82.4	17.6
Long-term sickness or disabled	670	126	654	137	1.7	1.6		82.7	17.3
Prefer not to answer or missing	8561	1849	9882	2559	25.7	30		79.4	20.6
Occupation									
Essential workers	7990	1536	9303	2161	24.2	25.3	<.001	81.2	18.8
Nonessential workers	10,246	1898	10,518	1949	27.3	22.8		84.4	15.6
Others	2458	500	2756	577	7.2	6.8		82.7	17.3

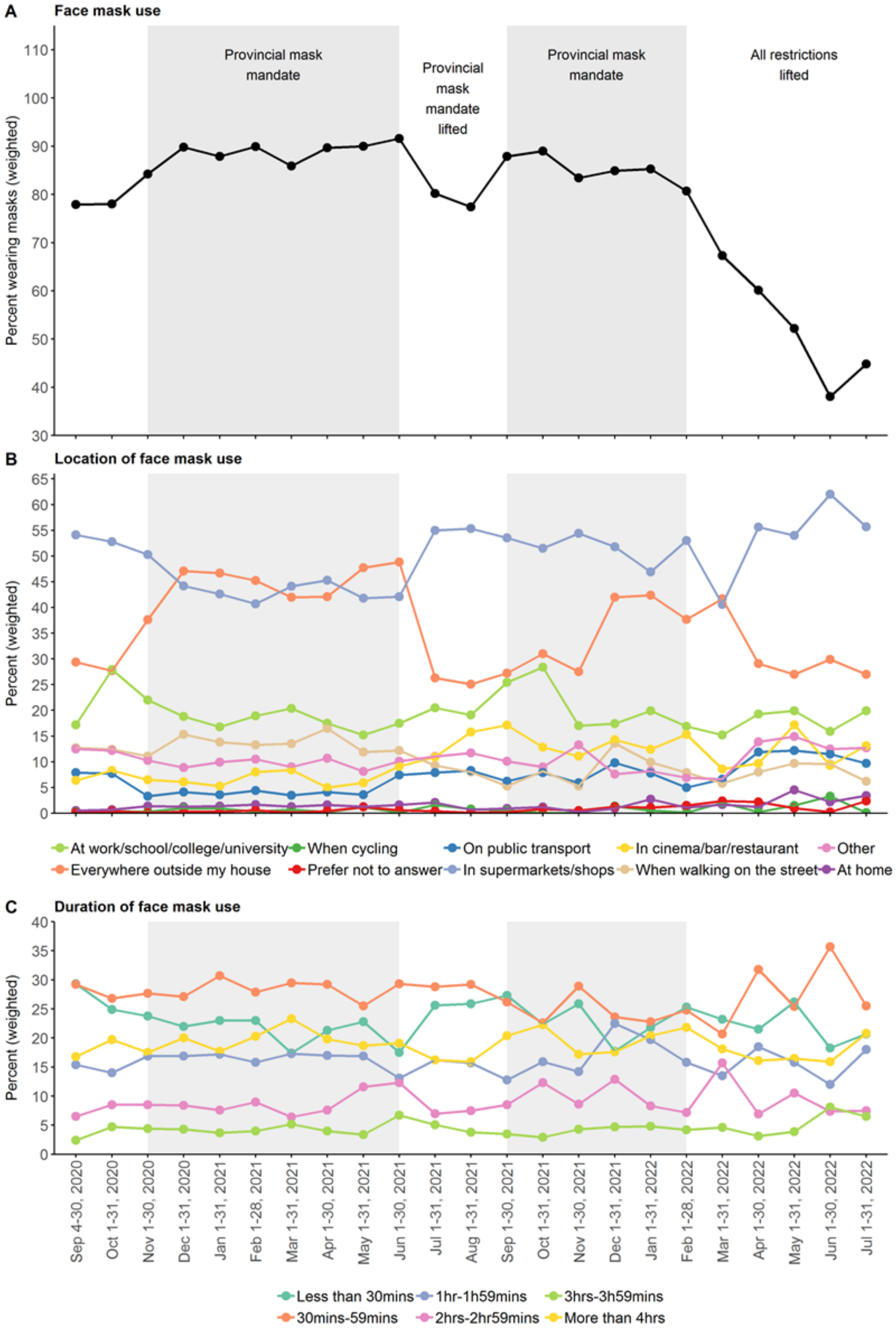
Characteristic	Unweighted values		Weighted values (distribution "across" groups)				P value[a]	Weighted values (distribution "within" each group)	
	Yes, n	No, n	Yes, n	No, n	Yes, %	No, %		Yes, %	No, %
Do not work	7282	1749	5751	1206	14.9	14.1		82.7	17.3
Prefer not to answer or missing	8740	1902	10,180	2645	26.4	31		79.4	20.6
Health Authority									
Fraser Health	7339	1333	10,335	2060	26.8	24.1	<.001	83.4	16.6
Interior Health	4896	1354	4252	1382	11	16.2		75.5	24.5
Northern Health	1378	356	1441	479	3.7	5.6		75.1	24.9
Vancouver Coastal Health	7079	1103	7835	1218	20.3	14.3		86.5	13.5
Vancouver Island Health	7274	1574	4851	1096	12.6	12.8		81.6	18.4
Missing	8750	1865	9794	2303	25.4	27		81	19

[a] Chi-square test.

Face Mask Usage Patterns and the Impact of the Provincial Mask Mandate

Between September 2020 and July 2022, 81.9% of survey respondents reported wearing a face mask outside their homes the day before completing the survey (face mask use: yes, n=36,716; no, n=7585 [*prefer not to answer*, n=155 + *no*, n=7430]). Face mask usage rates were approximately 78% between September and October 2020 when face masks and coverings were required in major retail locations in BC but not provincially mandated. Face mask usage rates increased following the introduction of the provincial mask mandate in November 2020 and remained at or above 84% each month thereafter until the mandate was first lifted in July 2021 (Figure 1A). After a 2-month decline in face mask usage rates to pre–provincial mandate levels, usage rates rebounded to 87.9% in September 2021 once the provincial mask mandate was reinstated at the end of August 2021. As before, face mask usage rates remained above 80% when the provincial mandate was in effect until March 2022, when the mandate was lifted a final time as an important step in the winding down of control measures across the province. Face mask usage rates fell rapidly soon afterward, reaching 38.1% in June 2022.

Figure 1. Face mask usage patterns among British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents by month, from September 2020 to July 2022. (A) Face mask use rates. (B) Location of face mask use: percentages calculated independently for each option provided. (C) Duration of face mask use. Shaded gray region: provincial mask mandate in effect.



Face mask usage patterns were generally consistent over the 23-month period (Figure 1A). Masks were primarily worn in supermarkets and shops (48.2% of face mask users) and everywhere outside the house (38.2% of face mask users; Figure 1B and Figure S1A in the Multimedia Appendix 1).The proportion of people who wore face masks everywhere outside their homes remained at or above 25%, with face mask users being more likely to have worn face masks everywhere outside their homes when provincial mask mandates were in effect. Regardless of time period, most face mask users reported wearing their mask for 59 minutes or less (Figure 1C and Figure S1B in the Multimedia Appendix 1).

Travel patterns were distinct between users and nonusers of face masks between September 2020 and July 2022. The majority of face mask users (53.2%) and nonusers (43.5%) left home only once the previous day, although nonusers of face masks were at least twice as likely to leave home 4 times or more (18.2% vs 7.3%; Figure S2A in the Multimedia Appendix 1 and Figure 2A). Retail locations including grocery stores, pharmacies, and liquor stores were the most frequented destinations for face mask users throughout the 23-month period (≥45%; Figure S2A in the Multimedia Appendix 1 and Figure 2A). Among nonusers of face masks, however, parks or public spaces were the most common destinations visited prior to June 2021, which was gradually surpassed by retail locations after the lifting of provincial mask mandates (Figure 2B). Face mask users mostly travelled alone in a car, although the mode of travel was more heterogeneous among nonusers of face masks during the 23-month period (Figure 2C and Figure S2C in the Multimedia Appendix 1). Statistically significant differences were observed in the distributions of face mask users and nonusers by sex, where larger proportions of females than males opted for wearing masks (85% vs 79%; *P*<.001; Table 1). This difference was more evident when survey respondents were grouped by destination (Table 2, Table 3). In terms of face mask group composition, males formed a large majority of people who opted out of wearing face masks when visiting retail locations (74.9%); restaurants, bars, or cafés (76.1%); or workplaces (78%).

Figure 2. Travel patterns of British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents by face mask use by month, from September 2020 to July 2022. (A) Number of trips taken outside the home. (B) Destination: percentages calculated independently for each option provided. (C) Mode of travel: percentages calculated independently for each option provided. Shaded gray region: provincial mask mandate in effect.

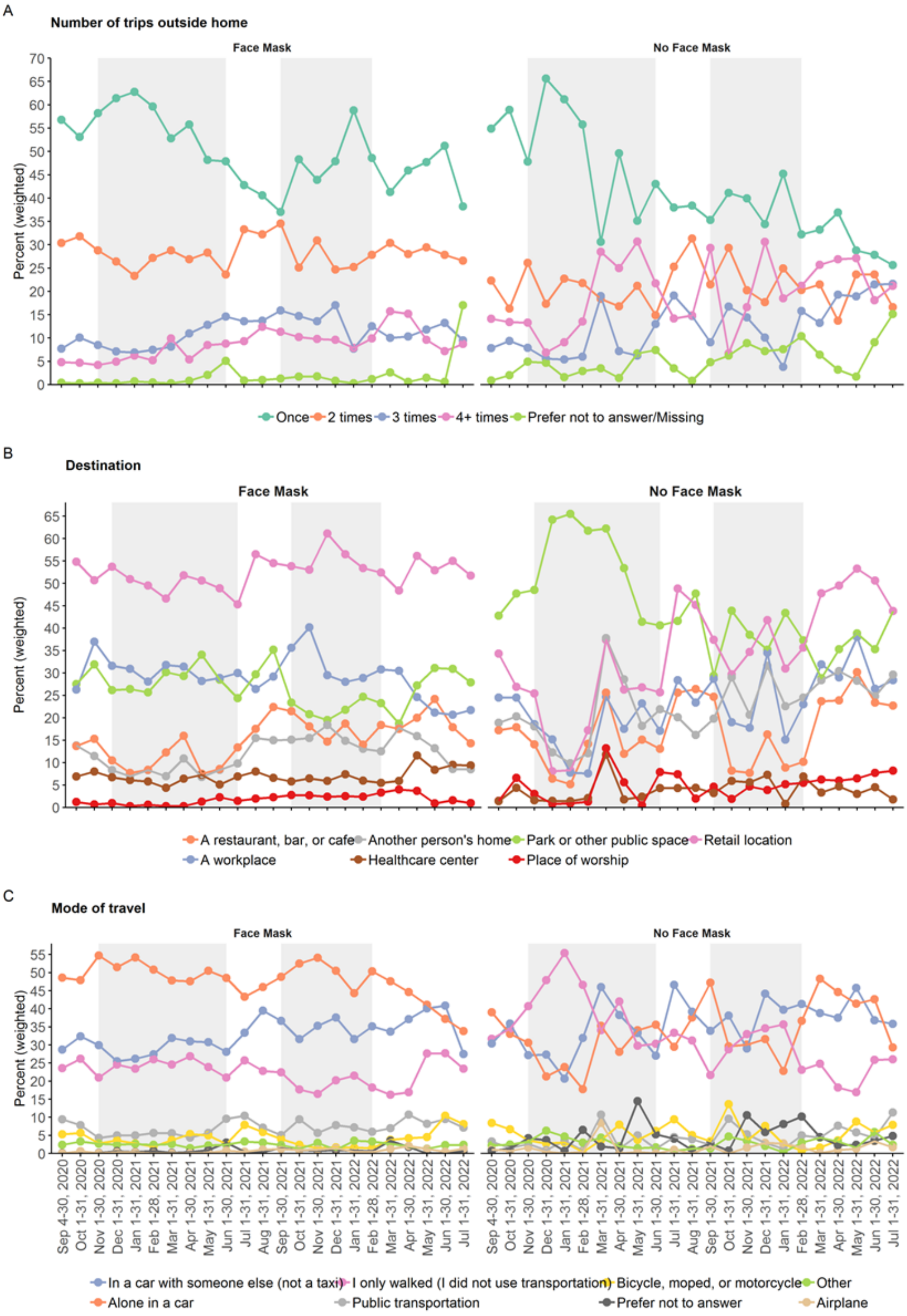


Table 2. Proportion of British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents by destination[a] and face mask use (yes or no), stratified “across” age groups and sex, from September 2020 to July 2022.

Characteristic	Retail location (n=21,732)		Restaurant, bar, or café (n=5183)		Workplace (n=10,226)		Park or other public space (n=14,577)	
	Yes (n=19,761), %[b]	No (n=1971), %[b]	Yes (n=4239), %[b]	No (n=944), %[b]	Yes (n=9120), %[b]	No (n=1106), %[b]	Yes (n=10,735), %[b]	No (n=3842), %[b]
Age group (years)								
18-24	6.4	9.5	10.4	13	11.8	14.6	7.4	7.7
25-34	16.8	20	21.7	20.1	24.8	22.1	21.8	21.3
35-44	14.6	18.2	15.8	17.2	19.2	18.1	18	15.3
45-54	19.3	19.4	18.1	20.4	22.9	24.9	16.9	18.2
55-64	17.5	15.2	14.4	12.7	15.9	14	15.1	15.9
65-74	19	13.8	14.6	12	4.7	5.1	16	16.8
≥75	6.3	4	4.9	4.4	0.6	1.1	4.8	4.8
Sex								
Female	48.1	25.1	42.3	23.9	46.7	22	52.7	43.9
Male	51.9	74.9	57.7	76.1	53.3	78	47.3	56.1

[a] Selected individually—percentages were calculated for each option provided.

[b] Weighted percentages.

Table 3. Proportion of British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents by destination[a], stratified as face mask users and nonusers “within” each age group and sex, from September 2020 to July 2022.

Characteristic	Retail location (n=21,732)		Restaurant, bar, or café (n=5183)		Workplace (n=10,226)		Park or other public space (n=14,577)	
	User (n=19,761), %[b]	Nonuser (n=1971), %[b]	User (n=4239), %[b]	Nonuser (n=944), %[b]	User (n=9120), %[b]	Nonuser (n=1106), %[b]	User (n=10,735), %[b]	Nonuser (n=3842), %[b]
Age group (years)								
18-24	82.3	17.7	72.8	27.2	82.7	17.3	72.5	27.5
25-34	85.3	14.7	78.3	21.7	86.9	13.1	73.5	26.5
35-44	84.6	15.4	75.4	24.6	86.3	13.7	76.2	23.8
45-54	87.2	12.8	74.9	25.1	84.5	15.5	71.6	28.4
55-64	88.8	11.2	79.1	20.9	87	13	72.1	27.9
65-74	90.5	9.5	80.2	19.8	84.5	15.5	72.1	27.9
≥75	91.5	8.5	78.5	21.5	76.6	23.4	73.2	26.8
Sex								
Female	92.9	7.1	85.5	14.5	92.6	7.4	76.5	23.5
Male	82.6	17.4	71.7	28.3	80.2	19.8	69.6	30.4

[a] Selected individually—percentages were calculated for each option provided.

[b] Weighted percentages.

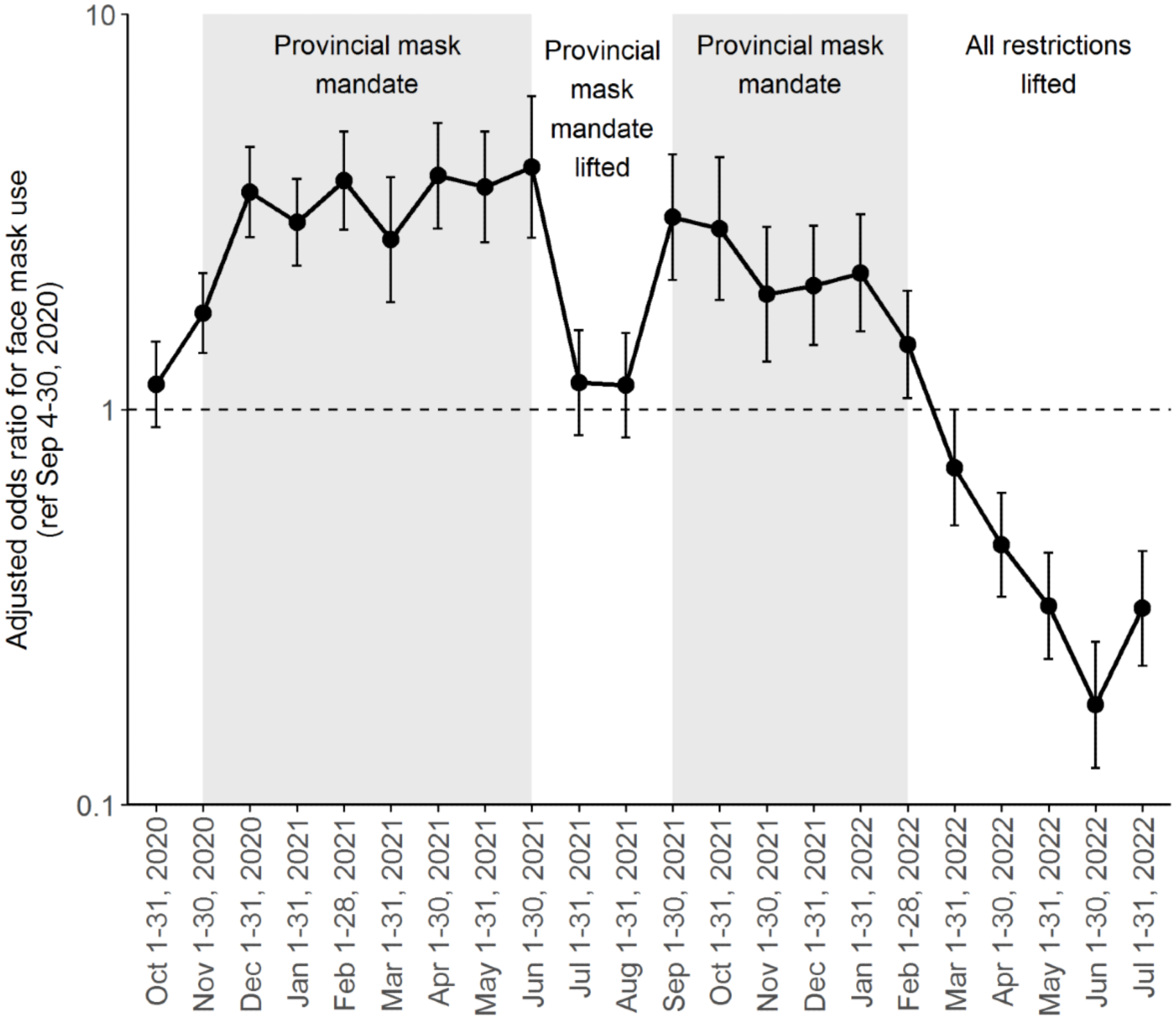
There were small differences in the demographic distributions of people who reported wearing or not wearing face masks in the presence and absence of the provincial mask mandate (Table S3 in the Multimedia Appendix 1). However, the shift toward decreased face mask use when the provincial mask mandate was not in effect was evident across people of all sexes, age groups, and ethnicities, regardless of their level of education, occupation, or employment status.

Factors Associated with Face Mask Use

In a multivariable logistic regression model, the odds of wearing a face mask were statistically significantly greater during the months when the mask mandate was in effect (all $P<.05$; Figure

3 and Table S2 in the Multimedia Appendix 1). The destination and mode of travel were associated with face mask use, where people going to retail locations, including grocery stores, pharmacies, and liquor stores, had greater odds of wearing face masks than those going to parks or other public spaces (adjusted odds ratio [aOR] 14.23, 95% CI 11.69-17.31), as did persons travelling alone in a car (aOR 2.15, 95% CI 1.86-2.50) or in a car with someone else (aOR 1.59, 95% CI 1.36-1.86) relative to those who only walked to their destinations (Table S2 in the Multimedia Appendix 1). Compared to people who were not part of a visible minority group, Chinese people (aOR 2.02, 95% CI 1.54-2.65), South Asian people (aOR 1.80, 95% CI 1.27-2.56), and others who were part of a visible minority group (aOR 1.89, 95% CI 1.49-2.40) had greater odds of wearing face masks. The odds of face mask use were also greater among females, people aged ≥65 years, and people living in the more populous health regions.

Figure 3. Odds ratios for face mask use among British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey respondents by time period, from September 2020 to July 2022.



The impact of the provincial mask mandates was even more evident when examined directly, with 3.68 times greater odds of face mask use when the provincial mask mandate was in effect (aOR 3.68, 95% CI 3.33-4.05; Table S4 in the Multimedia Appendix 1). The odds ratios for face mask use increased more than 2-fold among people whose destinations were indoor public spaces such as restaurants, bars, or cafés (aOR 7.35, 95% CI 4.23-12.78 vs aOR 2.81, 95% CI 1.50-5.26) or retail locations (aOR 19.94, 95% CI 14.86-26.77 vs aOR 7.71, 95% CI 5.68-10.46) with the mask mandate versus without. Slight shifts in odds ratios were also noted by mode of travel in the presence versus the absence of a provincial mask mandate.

Discussion

Principal Findings

Prior to the availability and high coverage of COVID-19 vaccines, masks and other nonpharmaceutical interventions were mainstays for preventing infection and reducing disease transmission, with the ultimate goal of reducing the impact of the COVID-19 pandemic on population health. Face mask use was mandated in certain settings across many countries to reduce SARS-CoV-2 transmission. Establishing the factors associated with and the patterns of face mask use, with or without mask mandates, is necessary to assess the impact of mask mandates and to inform health communication strategies and decision-making by public health leadership. In this study, based on survey responses from a voluntary sample of BC residents conducted between September 2020 and July 2022, 81.9% of respondents reported wearing a face mask during outings. Over the 23-month period, face masks were mostly worn for less than an hour, being primarily used in supermarkets and shops; at workplaces; and in schools, colleges, or universities. In a multivariable logistic regression model, factors associated with face mask use included age, sex, ethnicity, time period, destination, and mode of travel. Face mask usage rates were sustained by in-store mandates in the fall of 2020 and further boosted by the provincial mandates. The odds of face mask use increased 3-fold when the provincial mask mandate was in effect. These findings highlight the role of mask mandates in facilitating high levels of face mask use at the population level.

Studies based in the United States have shown an increased likelihood of face mask use in indoor public spaces, such as grocery stores, compared to outdoor public spaces, such as parks or beaches [43,44]. Similarly, retail locations and workplaces were among the major destinations associated with face mask use in BC, whereas parks were favored among people who opted against wearing masks. This finding was to be expected, as in-store and regional mask mandates were in effect in most of these locations for the majority of the study period. Differences in face mask use at retail locations have been reported in the United States, where females [45], individuals aged ≥65 years [45], non-Hispanic Black and Hispanic/Latino persons [46], and people shopping in urban or suburban locations [45] were among the most likely to have worn face masks prior to the introduction of mask mandates or recommendations. Similarly, females and people aged ≥55 years had greater odds of adopting face masks across Canada [28]. This was congruent with our findings in BC, where statistically significant differences in face mask use were noted by age and sex—males were more likely to be nonusers of face masks, especially in commonly frequented settings.

Our findings, and those of others, highlight the impact of mask mandates in promoting face mask use during the COVID-19 pandemic [45,47]. Province-wide in-store mask mandates sustained face mask usage rates in BC at approximately 80%, both before the introduction of the provincial mask mandate and during the 2-month period when the mandate was first lifted. Moreover, face mask usage rates were at or above 84% when provincial mask mandates were in effect, similar to findings in the United States and Australia [45,47]. The greatest impact of the mask mandate on the odds of face mask use was seen at key locations such as workplaces, restaurants, bars, cafes, grocery stores, liquor stores, and pharmacies. Once removed, alongside other control measures, face mask usage rates declined 2-fold to 38%, possibly reflecting baseline midpandemic mask usage rates in the absence of mask mandates.

Limitations

Study findings should be interpreted with the following limitations in mind. Data collection began after face masks were made mandatory on public transit and in many retail locations in BC; thus, we were unable to fully quantify premandate willingness to voluntarily wear face masks at these locations. Nevertheless, our data does contribute to the body of knowledge about (un)willingness to wear face masks in the face of regional or in-store mandates, as a sizeable proportion of respondents fell into this category. Our findings are also subject to selection bias, as survey respondents were recruited mainly on social media platforms (Instagram, Facebook, YouTube, and Twitter) and participated on a voluntary basis. Thus, persons who did not use these social media platforms would not have been able to participate without referral. Furthermore, we were not able to quantify nonparticipation as recruitment was done in a passive manner. In addition, we did not assess type of face masks used, which may have provided additional insights into the characteristics and behaviors of survey respondents. Nevertheless, our study provides valuable insight into the dynamics of face mask use during the COVID-19 pandemic.

Conclusions

Various studies have shown the association between face mask use and declines in SARS-CoV-2 transmission [21-23]. Thus, in the absence of vaccines for disease prevention and therapeutics for the treatment and prevention of severe disease, mask mandates were introduced during the COVID-19 pandemic to limit the spread of the disease and to reduce its impact on society. We found a pattern of high mask usage rates with retail location and public transit mask mandates in BC, which was further enhanced by the provincial mask mandate. These findings demonstrate the utility of mask mandates in sustaining high rates of face mask use during the COVID-19 pandemic and provide concrete evidence for their use in regions with low vaccination rates and recurrent surges in COVID-19 cases and in the event of future respiratory virus-driven pandemics or severe respiratory disease outbreaks. Lessons learned from the COVID-19 pandemic do suggest, however, that mask mandate imposition should require a sound ethical analysis beforehand to ensure that the benefits achieved with their use outweigh the harms related to infringement on individual choices.

Acknowledgments

We gratefully acknowledge the residents of British Columbia who participated in the British Columbia COVID-19 Population Mixing Patterns (BC-Mix) survey, for whom this work is intended to benefit.

This study was funded by the Michael Smith Foundation for Health Research COVID-19 Research Response Fund (COV-2020-1183) and the Canadian Institutes of Health Research (VR5-172683).

The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

Data are available upon reasonable request.

Authors' Contributions

NZJ, PAA , BM, and TBT designed the survey. NZJ and PAA contributed to survey dissemination and data collection. MB, PAA and NZJ conceptualized the study. MB carried out data analyses and drafted and revised the manuscript. All authors contributed to data interpretation and manuscript revision.

Conflicts of Interest

NZJ participated in advisory boards and has spoken for AbbVie and Gilead, not related to current work. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 637 KB - publichealth_v9i1e42616_app1.docx]

References

1. Güner R, Hasanoğlu I, Aktaş F. COVID-19: prevention and control measures in community. Turk J Med Sci 2020 Apr 21;50(SI-1):571-577 [FREE Full text] [doi: 10.3906/sag-2004-146] [Medline: 32293835]
2. Listings of WHO's response to COVID-19. World Health Organization. 2020 Jun 29. URL: https://www.who.int/news/item/29-06-2020-covidtimeline [accessed 2020-11-30]
3. Leung NHL, Chu DKW, Shiu EYC, Chan KH, McDevitt JJ, Hau BJP, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med 2020 May;26(5):676-680 [FREE Full text] [doi: 10.1038/s41591-020-0843-2] [Medline: 32371934]
4. Alsved M, Matamis A, Bohlin R, Richter M, Bengtsson P, Fraenkel C, et al. Exhaled respiratory particles during singing and talking. Aerosol Sci Technol 2020 Sep 17;54(11):1245-1248. [doi: 10.1080/02786826.2020.1812502]
5. Gandhi M, Marr LC. Uniting infectious disease and physical science principles on the importance of face masks for COVID-19. Med (N Y) 2021 Jan 15;2(1):29-32 [FREE Full text] [doi: 10.1016/j.medj.2020.12.008] [Medline: 33521753]
6. Gandhi M, Beyrer C, Goosby E. Masks do more than protect others during COVID-19: reducing the inoculum of SARS-CoV-2 to protect the wearer. J Gen Intern Med 2020 Oct;35(10):3063-3066 [FREE Full text] [doi: 10.1007/s11606-020-06067-8] [Medline: 32737790]
7. Science brief: community use of cloth masks to control the spread of SARS-CoV-2. Centers for Disease Control and Prevention. 2020 Nov 10. URL: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/masking-science-sars-cov2.html [accessed 2021-03-22]
8. Hendrix MJ, Walde C, Findley K, Trotman R. Absence of apparent transmission of SARS-CoV-2 from two stylists after exposure at a hair salon with a universal face covering policy - Springfield, Missouri, May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):930-932 [FREE Full text] [doi: 10.15585/mmwr.mm6928e2] [Medline: 32673300]
9. Lyu W, Wehby GL. Community use of face masks and COVID-19: evidence from a natural experiment of state mandates in the US. Health Aff (Millwood) 2020 Aug 01;39(8):1419-1425. [doi: 10.1377/hlthaff.2020.00818] [Medline: 32543923]
10. Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, van der Westhuizen H, et al. An evidence review of face masks against COVID-19. Proc Natl Acad Sci U S A 2021 Jan 26;118(4):e2014564118 [FREE Full text] [doi: 10.1073/pnas.2014564118] [Medline: 33431650]
11. Guy GP, Lee FC, Sunshine G, McCord R, Howard-Williams M, Kompaniyets L, CDC COVID-19 Response Team, Mitigation Policy Analysis Unit, CDC Public Health Law Program. Association of state-issued mask mandates and allowing on-premises restaurant dining with county-level COVID-19 case and death growth rates - United States, March 1-December 31, 2020. MMWR Morb Mortal Wkly Rep 2021 Mar 12;70(10):350-354 [FREE Full text] [doi: 10.15585/mmwr.mm7010e3] [Medline: 33705364]
12. Li Y, Liang M, Gao L, Ayaz Ahmed M, Uy JP, Cheng C, et al. Face masks to prevent transmission of COVID-19: a systematic review and meta-analysis. Am J Infect Control 2021 Jul;49(7):900-906 [FREE Full text] [doi: 10.1016/j.ajic.2020.12.007] [Medline: 33347937]
13. Mask mandates by nation: most still await a breath of fresh air. Bloomberg News. 2021 May 14. URL: https://www.bloomberg.com/news/articles/2021-05-14/mask-mandates-by-nation-most-still-await-a-breath-of-fresh-air [accessed 2021-07-21]
14. Feng S, Shen C, Xia N, Song W, Fan M, Cowling BJ. Rational use of face masks in the COVID-19 pandemic. Lancet Respir Med 2020 May;8(5):434-436 [FREE Full text] [doi: 10.1016/S2213-2600(20)30134-X] [Medline: 32203710]
15. Prasad R. Coronavirus: why is there a US backlash to masks? BBC News. 2020 May 05. URL: https://www.bbc.com/news/world-us-canada-52540015 [accessed 2021-07-22]

16. Jarry J. Why some people choose not to wear a mask. McGill Office for Science and Society. 2020 Sep 03. URL: https://www.mcgill.ca/oss/article/covid-19-health/why-some-people-choose-not-wear-mask [accessed 2021-07-23]
17. Mandating masks will remain the job of businesses and services, Dr. Bonnie Henry says. CBC News. 2020 Nov 16. URL: https://www.cbc.ca/news/canada/british-columbia/bc-mask-mandate-businesses-1.5804517 [accessed 2021-07-23]
18. Taylor S, Asmundson GJG. Negative attitudes about facemasks during the COVID-19 pandemic: the dual importance of perceived ineffectiveness and psychological reactance. PLoS One 2021 Feb 17;16(2):e0246317 [FREE Full text] [doi: 10.1371/journal.pone.0246317] [Medline: 33596207]
19. Haraf R, Faghy M, Carlin B, Josephson R. The physiological impact of masking is insignificant and should not preclude routine use during daily activities, exercise, and rehabilitation. J Cardiopulm Rehabil Prev 2021 Jan 01;41(1):1-5 [FREE Full text] [doi: 10.1097/HCR.0000000000000577] [Medline: 33351538]
20. Howard MC. Understanding face mask use to prevent coronavirus and other illnesses: development of a multidimensional face mask perceptions scale. Br J Health Psychol 2020 Nov;25(4):912-924 [FREE Full text] [doi: 10.1111/bjhp.12453] [Medline: 32588949]
21. Krishnamachari B, Morris A, Zastrow D, Dsida A, Harper B, Santella A. The role of mask mandates, stay at home orders and school closure in curbing the COVID-19 pandemic prior to vaccination. Am J Infect Control 2021 Aug;49(8):1036-1042 [FREE Full text] [doi: 10.1016/j.ajic.2021.02.002] [Medline: 33577824]
22. Van Dyke ME, Rogers TM, Pevzner E, Satterwhite CL, Shah HB, Beckman WJ, et al. Trends in county-level COVID-19 incidence in counties with and without a mask mandate - Kansas, June 1-August 23, 2020. MMWR Morb Mortal Wkly Rep 2020 Nov 27;69(47):1777-1781 [FREE Full text] [doi: 10.15585/mmwr.mm6947e2] [Medline: 33237889]
23. Mitze T, Kosfeld R, Rode J, Wälde K. Face masks considerably reduce COVID-19 cases in Germany. Proc Natl Acad Sci U S A 2020 Dec 22;117(51):32293-32301 [FREE Full text] [doi: 10.1073/pnas.2015954117] [Medline: 33273115]
24. Peeples L. What the science says about lifting mask mandates. Nature 2021 May 25;593(7860):495-498. [doi: 10.1038/d41586-021-01394-0] [Medline: 34035531]
25. Kresge N. Europe Can't Shake Off Covid as Variant Fuels Summer Spike. Bloomberg. 2022 Jul 16. URL: https://www.bloomberg.com/news/articles/2022-07-16/europe-can-t-shake-covid-as-ba-5-variant-fuels-summer-wave-of-cases [accessed 2022-08-29]
26. French authorities 'encourage' mask use as COVID cases surge. EuroNews. 2022 Jul 04. URL: https://www.euronews.com/2022/07/02/french-authorities-encourage-mask-use-as-covid-cases-surge [accessed 2022-08-29]
27. St. George D. School mask mandates return as latest coronavirus variants surge. The Washington Post. 2022 Jul 27. URL: https://www.washingtonpost.com/education/2022/07/27/school-mask-mandates-return/ [accessed 2022-08-29]
28. Lang R, Atabati O, Oxoby RJ, Mourali M, Shaffer B, Sheikh H, et al. Characterization of non-adopters of COVID-19 non-pharmaceutical interventions through a national cross-sectional survey to assess attitudes and behaviours. Sci Rep 2021 Nov 05;11(1):21751 [FREE Full text] [doi: 10.1038/s41598-021-01279-2] [Medline: 34741109]
29. Benham JL, Lang R, Kovacs Burns K, MacKean G, Léveillé T, McCormack B, et al. Attitudes, current behaviours and barriers to public health measures that reduce COVID-19 transmission: a qualitative study to inform public health messaging. PLoS One 2021 Feb 19;16(2):e0246941 [FREE Full text] [doi: 10.1371/journal.pone.0246941] [Medline: 33606782]
30. Government of British Columbia. COVID-19 Information and Resources. EmergencyInfoBC. 2021. URL: https://www.emergencyinfobc.gov.bc.ca/covid19 [accessed 2021-05-04]
31. Ministerial order no. M012. British Columbia Minister of Public Safety and Solicitor General. 2021 Jan 08. URL: https://www.bclaws.gov.bc.ca/civix/document/id/mo/mo/m0012_2021 [accessed 2021-07-23]
32. Weichel A. B.C. announces mask mandate, temporary social lockdown for entire province. CTV News. 2020 Nov 20. URL: https://bc.ctvnews.ca/b-c-announces-mask-mandate-temporary-social-lockdown-for-entire-province-1.5196004 [accessed 2021-07-23]
33. Where do you need to wear a mask in B.C.? Here are some places where they are mandatory. CBC News. 2020 Aug 31. URL: https://www.cbc.ca/news/canada/british-columbia/where-do-you-need-to-wear-a-mask-in-b-c-here-are-some-places-where-they-are-mandatory-1.5704700 [accessed 2021-07-23]
34. Ross A. Face masks will be mandatory on public transit across much of B.C. CBC News. 2020 Aug 06. URL: https://www.cbc.ca/news/canada/british-columbia/face-masks-mandatory-translink-1.5676780 [accessed 2021-07-23]
35. Zussman R. B.C. drops mandatory indoor mask mandate starting July 1. Global News. 2021 Jun 29. URL: https://globalnews.ca/news/7989652/bc-ending-mandatory-masks/ [accessed 2021-07-23]
36. Mask mandate to reduce transmission, protect people in public spaces. British Columbia Ministry of Health. 2021. URL: https://news.gov.bc.ca/releases/2021HLTH0053-001665 [accessed 2021-09-10]
37. Indoor mask mandate extended. British Columbia Ministry of Health. 2021. URL: https://news.gov.bc.ca/releases/2021HLTH0190-002077 [accessed 2021-11-09]
38. B.C. takes next step in balanced plan to lift COVID-19 restrictions. Government of British Columbia. 2022. URL: https://news.gov.bc.ca/releases/2022HLTH0081-000324 [accessed 2022-08-08]
39. Adu PA, Binka M, Mahmood B, Jeong D, Buller-Taylor T, Damascene MJ, et al. Cohort profile: the British Columbia COVID-19 Population Mixing Patterns Survey (BC-Mix). BMJ Open 2022 Aug 24;12(8):e056615 [FREE Full text] [doi: 10.1136/bmjopen-2021-056615] [Medline: 36002217]

40. Bethlehem J. Applied survey methods: A statistical perspective. New York, NY: John Wiley & Sons; 2009.
41. R Core Team. R: a language and environment for statistical computing. The R Foundation for Statistical Computing. URL: https://www.r-project.org [accessed 2021-04-06]
42. What is SAS? SAS Institute Inc. 2021 Nov 04. URL: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/lrcon/n0toa1y5c8wkpxn1duru2ltw8i52.htm [accessed 2021-04-06]
43. Knotek ESII, Schoenle RS, Dietrich AM, Müller GJ, Myrseth KOR, Weber M. Consumers and COVID-19: survey results on mask-wearing behaviors and beliefs. Federal Reserve Bank of Cleveland, Economic Commentary. 2020 Jul 16. URL: https://www.clevelandfed.org/publications/economic-commentary/ec-202020-survey-results-on-mask-wearing-behaviors-and-beliefs [accessed 2022-12-12]
44. Schoeni RF, Wiemers EE, Seltzer JA, Langa KM. Association between risk factors for complications from COVID-19, perceived chances of infection and complications, and protective behavior in the US. JAMA Netw Open 2021 Mar 01;4(3):e213984 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.3984] [Medline: 33787906]
45. Haischer MH, Beilfuss R, Hart MR, Opielinski L, Wrucke D, Zirgaitis G, et al. Who is wearing a mask? gender-, age-, and location-related differences during the COVID-19 pandemic. PLoS One 2020 Oct 15;15(10):e0240785 [FREE Full text] [doi: 10.1371/journal.pone.0240785] [Medline: 33057375]
46. Fisher KA, Barile JP, Guerin RJ, Vanden Esschert KL, Jeffers A, Tian LH, et al. Factors associated with cloth face covering use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):933-937 [FREE Full text] [doi: 10.15585/mmwr.mm6928e3] [Medline: 32673303]
47. Scott N, Saul A, Spelman T, Stoove M, Pedrana A, Saeri A, et al. The introduction of a mandatory mask policy was associated with significantly reduced COVID-19 cases in a major metropolitan city. PLoS One 2021 Jul 21;16(7):e0253510 [FREE Full text] [doi: 10.1371/journal.pone.0253510] [Medline: 34288910]

Abbreviations

aOR: adjusted odds ratio
BC: British Columbia
BC-Mix: British Columbia COVID-19 Population Mixing Patterns

Edited by A Mavragani, T Sanchez; submitted 12.09.22; peer-reviewed by C Sun, C Shen, M Mourali; comments to author 07.10.22; revised version received 04.11.22; accepted 29.11.22; published 11.01.23.

Please cite as:
Binka M, Adu PA, Jeong D, Vadlamudi NK, Velásquez García HA, Mahmood B, Buller-Taylor T, Otterstatter M, Janjua NZ
The Impact of Mask Mandates on Face Mask Use During the COVID-19 Pandemic: Longitudinal Survey Study
JMIR Public Health Surveill 2023;9:e42616
URL: https://publichealth.jmir.org/2023/1/e42616
doi:10.2196/42616
PMID:36446134

Original Paper

The Determinants of Adherence to Public Health and Social Measures Against COVID-19 Among the General Population in South Korea: National Survey Study

Hye Chong Hong[1], RN, PhD; Hyeonkyeong Lee[2], MSN, PhD; Suk Jeong Lee[1], MPH, PhD; Chang Park[3], PhD; Mikyung Lee[2], MPH, PhD

[1]Red Cross College of Nursing, Chung-Ang University, Seoul, Republic of Korea
[2]College of Nursing, Mo-Im Kim Nursing Research Institute, Yonsei University, Seoul, Republic of Korea
[3]Department of Population Health Nursing Science, University of Illinois Chicago, Chicago, IL, United States

Corresponding Author:
Hyeonkyeong Lee, MSN, PhD
College of Nursing
Mo-Im Kim Nursing Research Institute
Yonsei University
50-1 Yonsei-ro
Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 3373
Email: hlee39@yuhs.ac

Abstract

Background: The COVID-19 pandemic has created devastating health, social, economic, and political effects that will have long-lasting impacts. Public health efforts to reduce the spread of COVID-19 are the priority of national policies for responding to the pandemic globally. Public health and social measures (PHSMs) have been shown to be effective when used alone or in combination with other measures, reducing the risk of spreading COVID-19. However, there is insufficient evidence on the status of compliance with PHSMs in the general population for the prevention of COVID-19 in public areas, including Korea.

Objective: The aim of this study was to assess levels of compliance with the recommended PHSMs against SARS-CoV-2 infection and their predictors among the general population by using national data.

Methods: This study was a secondary data analysis of the National Survey of Infectious Disease Preventive Behaviors in Community, which was conducted by the Korea Centers for Disease Control and Prevention Agency (KDCA) between October 12 and October 30, 2020. The primary study was cross-sectional, using stratified sampling via an adjusted proportional allocation method to select representative samples and ensure the stability of samples. The data were collected through phone interviews conducted by trained enumerators using a structured questionnaire. PHSM adherence was measured using a 10-item comprehensive infectious disease prevention behavior (CIDPB) scale, and each sociocognitive factor, including perceived susceptibility to SARS-CoV-2 infection, perceived severity of SARS-CoV-2 infection, perceived confidence in performing preventive behaviors related to COVID-19, information comprehension ability, and trust in information from the KDCA, was measured. A total of 4003 participants were included in the final analysis. Tobit regression and a decision tree analysis were performed to identify the predictors of preventive measures and the target groups for intervention.

Results: We discovered that women scored 1.34 points higher on the CIDPB scale than men (*P*<.001). Compared to the group aged 19 to 29 years, those aged 50 to 59 years and those older than 60 years scored 1.89 and 2.48 points higher on the CIDPB scale (*P*<.001), respectively. The perceived severity of infection, confidence in preventive behaviors, information comprehension ability, and trust in information from the KDCA were significant positive determinants of CIDPBs (*P*<.001). The perceived susceptibility to infection showed a significant negative relationship with CIDPBs (*P*<.001).

Conclusions: Female sex, older age, lower income, and sociocognitive factors were found to be significant determinants of adhering to PHSMs. The findings suggest the need for tailored interventions for target groups; specifically, the age group that was the most active at work indicated the highest potential to spread infection. Adequate public health education and health communication for promoting adherence to PHSMs should be emphasized, and behavior change strategies for those with low perceived confidence in performing PHSMs should be prioritized.

(JMIR Public Health Surveill 2023;9:e35784) doi:10.2196/35784

KEYWORDS

COVID-19; preventive measures; health literacy; trust; national survey; Tobit regression

Introduction

The COVID-19 pandemic is a global health crisis. The pandemic has created devastating health, social, economic, and political effects that will have long-lasting impacts in many parts of the world [1,2]. SARS-CoV-2 is a novel virus causing an outbreak of mild to severe respiratory diseases. COVID-19 has affected millions of people around the world. As of September 2021, a total of 219 million cases and over 4.55 million deaths were reported. Korea is no exception to this devastating crisis. More than 276,000 total cases and 2367 deaths were reported as of September 2021 [3]. Although COVID-19 vaccines are now available in many countries, vaccination does not guarantee complete protection from contracting COVID-19, and it will take a while to achieve herd immunity, as a substantial proportion of the population needs to be vaccinated [4]. Research is still ongoing to find out how strong the vaccines are against COVID-19, how effectively they would reduce transmission, and how long the effects of vaccination last. Until herd immunity is achieved, preventive measures should still be important means for reducing the transmission of COVID-19 and protecting ourselves from contracting an evolving infectious disease in the future [4].

Public health efforts to reduce the spread of COVID-19 are still the priority of national policies for responding to the pandemic globally. Most countries have announced public health and social measures (PHSMs) against COVID-19. PHSMs include personal protective measures, such as handwashing and mask wearing in public areas; environmental measures, such as disinfection and ventilation; surveillance and response measures, such as contact tracing, isolation, and quarantine; physical distancing measures, such as limiting the size of gatherings and maintaining distance in public areas; and limitations on international travel [5]. These protective measures have been shown to be effective when used alone or in combination with several measures, reducing the risk of spreading COVID-19 [6-8]. The Korea Centers for Disease Control and Prevention Agency (KDCA) has also announced guidelines for PHSMs that are similar to WHO guidelines [9]. Nevertheless, there is insufficient evidence on the status of compliance with PHSMs in the general population for the prevention of COVID-19 in public areas, including Korea.

Researchers have reported the disparities in compliance with PHSMs by socioeconomic groups. In general, being male, having a lower education and a lower income, living in rural areas, and being younger are associated with performing fewer preventive behaviors. [10-13]. Mistrust toward the health care system and health literacy are other important factors that contribute to the disparities in PHSM compliance. People who do not trust the health care system are less likely to follow recommendations if the recommendations are from the system that they do not trust, especially when these recommendations are inconsistent and difficult to understand [14]. An effective and rapid response to COVID-19 requires the public to follow recommended behaviors, and public health information plays a crucial role [15]. However, people vary in terms of their abilities to obtain and understand health information and use services to make informed decisions (ie, health literacy) [16], and conflicting messages about PHSMs for COVID-19 have created confusion in the public, which has caused delays in following PHSMs [17]. In fact, McCaffery et al [15] reported disparities in COVID-19–related knowledge, attitudes, and behaviors by health literacy level. However, more studies are needed to confirm the relationships among mistrust in the health care system, health literacy, and preventive behaviors in the new era of the COVID-19 pandemic.

Health beliefs are closely related to health behaviors. According to the Rosenstock Health Belief Model [18], perceived susceptibility (beliefs about acquiring a disease), perceived severity (beliefs about the seriousness of risk), and self-efficacy (confidence in the ability to control a situation to achieve a goal) are important concepts that could influence people’s decisions to engage in health behaviors. Although sparse, recent research revealed that health beliefs are significant predictors of COVID-19 preventive health behaviors [13,19,20]. Public health efforts reduce the spread of COVID-19 both globally and within countries, including Korea. Examining the factors that influence preventive health behaviors would provide opportunities to target interventions toward those who are in need. To date, there is a lack of knowledge about factors related to PHSMs against COVID-19. Thus, the purpose of this study was to assess levels of compliance with the recommended preventive measures against SARS-CoV-2 infection and their predictors among the general population by using national data.

Methods

Data Source and Sample

This study was a secondary data analysis of the National Survey of Infectious Disease Preventive Behaviors in Community, which was conducted by the KDCA between October 12 and October 30, 2020. The survey was developed based on the following steps: (1) a literature review on previous studies related to infectious diseases and associated preventive behaviors; (2) in-depth interviews with experts in infectious diseases on preventive behaviors and their predictors; (3) the pooling of items based on the literature review and in-depth interviews with experts; (4) the development of a pilot questionnaire based on the pooled items; (4) content validity index calculation with 11 experts; and (5) a pilot test with 20 adults on the difficulty, relevance, and comprehensiveness of the questionnaire. The survey included data on adherence to PHSMs, socioeconomic characteristics, experiences of respiratory and intestinal infectious disease symptoms, experiences of isolation or hospitalization due to SARS-CoV-2 infection, perceived susceptibility to COVID-19, perceived severity of COVID-19, confidence in performing preventive behaviors, COVID-19–related information comprehension ability, and trust in information from the KDCA.

A representative sample of the population aged 19 to 69 was stratified based on sex (male and female) and age groups (19-29 years, 30-39 years, 40-49 years, 50-59 years, and 60-69 years) in 17 cities or provinces nationwide. Stratified sampling via an adjusted proportional allocation method was used to select representative samples and to ensure the stability of samples in 17 cities or provinces. First, 50 people from each city or province were allocated. Second, a proportional allocation method was used to reduce or increase the sample size based on the number of people residing in each city or province. Therefore, the number of people selected from each city or province ranged from 25 to 1050. To increase the representativeness of the sample, a quota ratio was used to allocate samples based on sex and age groups; thus, the characteristics of the samples selected were consistent with those of the population. Trained professional enumerators conducted computer-assisted telephone interviews with participants, using random digital dialing, between October 12 and October 30, 2020. The interviews were stopped when the target proportion was reached. A total of 421,428 phone calls were made, and 5025 individuals completed telephone interviews. We included 4003 participants in this study after carefully reviewing and excluding participants with missing data on the dependent variable.

Ethical Considerations

This study was a secondary analysis of a pre-existing data set, and ethical review and approval were exempted for this study by the institutional review board of the first author's university (Chung-Ang University). The primary study was conducted by the KDCA, and informed consent was obtained prior to the phone interviews. The primary data contained anonymous information.

Measures

Adherence to PHSMs

Adherence to PHSMs was measured by using a comprehensive infectious disease prevention behavior (CIDPB) scale. The preliminary items were developed based on the review of literature published within the last 5 years in PubMed, Embase, Cochrane Library, and the Research Information Sharing Service on prevention behaviors for infectious diseases; existing instruments for COVID-19 prevention behaviors [21]; and the COVID-19 prevention guidelines from the Centers for Disease Control in the United States and Korea. The 10 items, including handwashing, cough etiquette, and social distancing, were constructed after a review by experts and content validity testing. Examples of items include (1) washing hands thoroughly for 30 seconds with soap and running water, (2) covering the nose and mouth with sleeves when coughing or sneezing, and (3) maintaining a 2-m (at least 1 m) distance from others. Each item was answered on a 4-point Likert scale (1=always; 2=most of the time; 3=sometimes; 4=never) and was reverse coded. The participants who answered "not applicable" or "don't know" and those who refused to answer were excluded from the analysis. The total score ranged from 4 to 40, with higher scores indicating the practicing of more preventive behaviors. The Cronbach α was 0.76 (95% CI 0.747-0.770), supporting the instrument's internal consistency.

Sociocognitive Factors

The participants were asked to answer questions related to perceived susceptibility to SARS-CoV-2 infection, perceived severity of SARS-CoV-2 infection, and perceived confidence in performing preventive behaviors related to SARS-CoV-2 infection, using a single question for each factor. Each question was answered on a 7-point Likert scale (1=not at all; 7=very likely), with a higher score indicating higher perceived susceptibility to SARS-CoV-2 infection, perceived severity of SARS-CoV-2 infection, perceived confidence in performing preventive behaviors, and information comprehension ability with regard to COVID-19. Information comprehension ability was assessed by asking the following question: "How easy is it for you to understand the various information given during COVID-19?" Participants were asked to answer on a 7-point Likert scale (1=very hard; 7=very easy), with a higher score indicating the easier comprehension of COVID-19 information.

Trust in information from the KDCA was assessed by asking the following question: "How much trust in information provided by KDCA?" Participants were asked to respond on a 10-point Likert scale (1=never trust; 10=very trust), with a higher score indicating more trust in information from the KDCA.

Demographic Variables

The demographic variables included in this study were age, sex (ie, male or female), educational level (ie, middle school or lower, high school, or college or above), monthly household income (ie, less than US $5000 or greater than US $5000), and living areas (urban or rural areas). Further, 8 metropolitan cities and cities in 9 provinces were categorized as urban areas, and the villages in the provinces were categorized as rural areas. Age was categorized as 19 to 29 years, 30 to 39 years, 40 to 49 years, 50 to 59 years, or ≥60 years.

Statistical Analysis

We analyzed the data by using STATA 15.1 (StataCorp LLC), SPSS 25 (IBM Corp), and R (R Foundation for Statistical Computing). A weighted statistical analysis was performed to increase the representativeness of the data. Descriptive statistics, such as means, SDs, frequencies, and percentages, were used to describe the participants. To examine the relationships that demographics, sociocognitive factors, and trust in information from the KDCA had with CIDPBs, the Mann-Whitney *U* test and the Kruskal-Wallis H test were done, as the CIDPBs were not normally distributed. A post hoc analysis was done by using Bonferroni corrections [22]. We then performed Tobit regression to examine the effects of demographics, sociocognitive factors, and trust in information from the KDCA on CIDPBs. A Tobit regression model, which is also called a *censored regression model*, was performed to estimate linear relationships among variables when the dependent variable was censored to either the left or right [23]. The dependent variable in this study (the CIDPBs) was skewed to the right, and this analysis allows one to specify a threshold to censor the regression. Further, a decision tree analysis was conducted by using the classification and regression tree (CART) technique introduced by Breiman et al [24] via the SPSS statistical package. CART analysis is a nonparametric technique for identifying each predictor in order to identify the most important predictor at each step when a sample is divided into 2 mutually exclusive and homogeneous subpopulations. The starting group is referred to as the *root*, each split is referred to as a *branch*, and the data subset resulting from the split is called a *node*, while the terminal nodes are referred to as *leaves* [24]. A single configuration of hyperparameters was chosen to reduce the complexity of CARTs to ensure sensible bias-variance trade-off. The minimum number of cases in the parent node and child node was 100 and 50, respectively. The chosen parameters for growing trees were a maximum tree depth of 5 and a significance threshold of .05 for splitting. The “±1 SE rule” was applied to prune the tree.

Results

The characteristics of participants are summarized in Table 1. The average age of participants was 45.4 (SD 14.98) years. Of the 4003 participants, 1947 (48.6%) were male. Approximately 57.3% (2265/3953) of participants were educated at a college or higher level. About 58% (1941/3349) earned less than US $5000 per month, and 42% (1408/3349) earned more than US $5000 per month. Most of the participants (3741/4003, 93.5%) resided in the city. Bivariate analyses were performed to examine the association between general characteristics and the CIDPBs (Table 2).

The mean CIDPB scale score was 34.91 (SD 3.74) and ranged from 15 to 40. Sex, age, the level of education, monthly income, and living areas were significantly associated with the CIDPBs. The mean score of male participants was slightly lower than that of female participants (*P*<.001), and the mean score for participants residing in urban areas was lower than that for participants residing in rural areas. The Bonferroni post hoc test revealed that the older age group had more CIDPBs. Specifically, the participants aged 19 to 29 years had the lowest mean score compared to those of other age groups (*P*<.001). The participants aged 30 to 39 years had a lower mean score than those of the participants aged 40 to 49 years (*P*=.001) and the participants aged 50 years and older (*P*<.001). The participants aged 40 to 49 years had a lower mean score than those of the participants aged 50 to 59 years (*P*=.002) and the participants aged 60 years and older (*P*<.001). The participants aged 50 to 59 years had a lower mean score than that of the participants aged 60 years and older (*P*<.001).

Table 1. General characteristics (N=4003) of participants in the National Survey of Infectious Disease Preventive Behaviors in Community (conducted on October 2020 in South Korea).

Characteristics	Participants, n (%)
Sex	
Male	1947 (48.6)
Female	2056 (51.4)
Age (years)	
19-29	737 (18.4)
30-39	693 (17.3)
40-49	799 (20)
50-59	773 (19.3)
≥60	1001 (25)
Education	
Middle school or lower	350 (8.9)
High school	1338 (33.8)
College or higher	2265 (57.3)
Monthly household income (US $)	
<5000	1941 (58)
≥5000	1408 (42)
Location of residence	
Urban	3741 (93.5)
Rural	262 (6.5)

Table 2. Level of compliance by general characteristics (N=4003) of participants in the National Survey of Infectious Disease Preventive Behaviors in Community (conducted on October 2020 in South Korea).

Characteristics	Score, mean (SD)	*z* test	*P* value	Post hoc test[a]
Sex		−9.86	<.001	
Male	34.4 (3.97)			N/A[b]
Female	35.5 (3.40)			N/A
Age (years)		263.95	<.001	
19-29 (A)	33.4 (3.97)			A<B, C, D, E
30-39 (B)	34.2 (3.78)			B<C, D, E
40-49 (C)	34.8 (3.64)			C<D, E
50-59 (D)	35.5 (3.40)			D<E
≥60 (E)	36.1 (3.38)			
Education		67.56	<.001	
Middle school or lower (A)	36.3 (3.50)			
High school (B)	34.8 (3.87)			B<A
College or higher (C)	34.7 (3.66)			C<A
Monthly household income (US $)		−4.28	<.001	
<5000	35.2 (3.66)			N/A
≥5000	34.6 (3.69)			N/A
Location of residence		−3.09	.002	
Urban	34.9 (3.75)			N/A
Rural	35.6 (3.57)			N/A

[a] Bonferroni test order.

[b] N/A: not applicable.

A Tobit regression model analysis was performed to identify predictors of CIDPBs (Table 3). The results revealed that the CIDPB score was 1.34 (95% CI 1.09-1.59) points higher in female participants when compared to that of male participants. The participants who were aged between 50 and 59 years scored higher on the CIDPB scale (performed more CIDPBs) by 1.89 (95% CI 1.45-2.34) points, and the participants who were older than 60 years scored higher on the CIDPB scale (performed more CIDPBs) by 2.48 points (95% CI 2.02-2.93), when compared to those who were aged between 19 and 29 years. The perceived severity of infection, confidence in preventive behaviors, information comprehension ability, and trust in information from the KDCA were statistically significant and positive (*P*<.001). The perceived susceptibility to infection was statistically significant and negative (*P*=.02).

Figure 1 depicts the decision tree produced by the CART analysis. Perceived confidence in performing preventive behaviors (scores of >6.5) was the first classifying dimension, which was mostly associated with CIDPBs. The results suggested that the participants aged 40 to 69 years with trust in information from the KDCA (scores of >9.5) had higher scores on the CIDPB scale (range 34.91-37.21). Men aged 19 to 49 years with scores of <4.5 for perceived confidence in performing preventive behaviors were the least likely to practice CIDPBs, with CIDPB scale scores ranging from 34.91 to 31.10. Perceived confidence in performing preventive behaviors showed the highest normalized importance among variables for explaining CIDPBs, followed by the age group and trust in information from the KDCA variables (Figure 2).

Table 3. Factors associated with preventive behaviors based on a Tobit regression model[a] in a National Survey of Infectious Disease Preventive Behaviors in Community (conducted on October 2020 in South Korea).

Characteristics	Coefficient (95% CI)	*P* value
Sex		
Male	Reference	
Female	1.34 (1.09 to 1.59)	<.001
Age (years)		
19-29	Reference	
30-39	0.84 (0.37 to 1.30)	<.001
40-49	1.41 (0.96 to 1.85)	<.001
50-59	1.89 (1.45 to 2.34)	<.001
≥60	2.48 (2.02 to 2.93)	<.001
Education		
Middle school or lower	Reference	
High school	−0.35 (−0.90 to 0.21)	.22
College or higher	−0.30 (−0.87 to 0.26)	.29
Monthly household income (US $)		
<5000	Reference	
≥5000	−0.47 (−0.73 to−0.20)	<.001
Location of residence		
Urban	Reference	
Rural	0.38 (−0.07 to 0.83)	.10
Sociocognitive factors		
Perceived susceptibility to infection	−0.08 (−0.15 to −0.02)	.02
Perceived severity of infection	0.14 (0.06 to 0.22)	<.001
Perceived confidence in preventive behavior	0.84 (0.72 to 0.97)	<.001
Information comprehension ability	0.21 (0.09 to 0.33)	<.001
Trust in information from the KDCA[b]	0.22 (0.14 to 0.30)	<.001

[a] Log pseudo likelihood=−8416.22; $F_{14,3321}$=48.79; P<.001; pseudo R^2=0.04.

[b] KDCA: Korea Centers for Disease Control and Prevention Agency.

Figure 1. Key factors associated with the comprehensive infectious disease prevention behaviors in the National Survey of Infectious Disease Preventive Behaviors in Community (conducted in October 2020 in South Korea). The classification and regression tree analysis results show relationships with various independent variables. KDCA: Korea Centers for Disease Control and Prevention Agency.

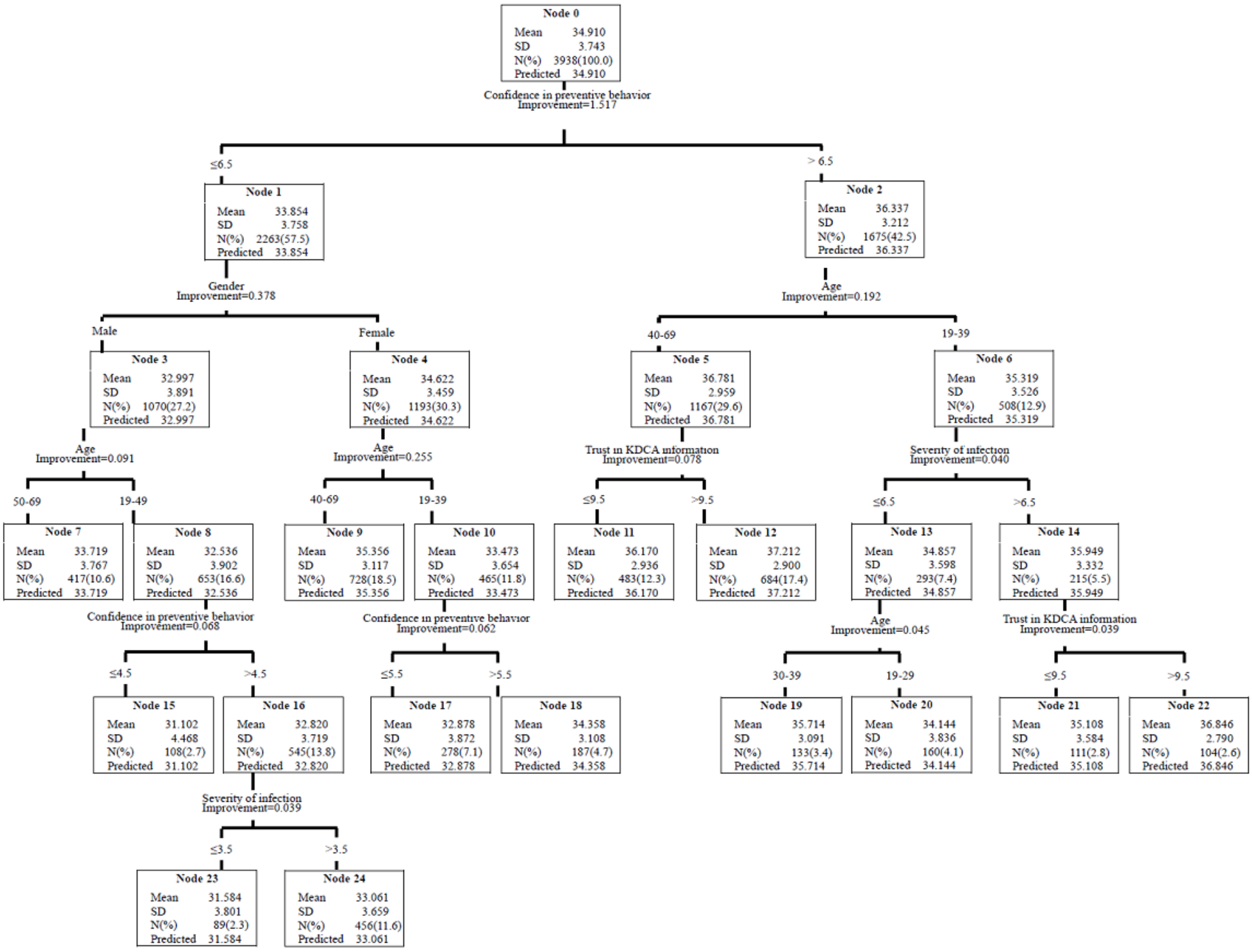
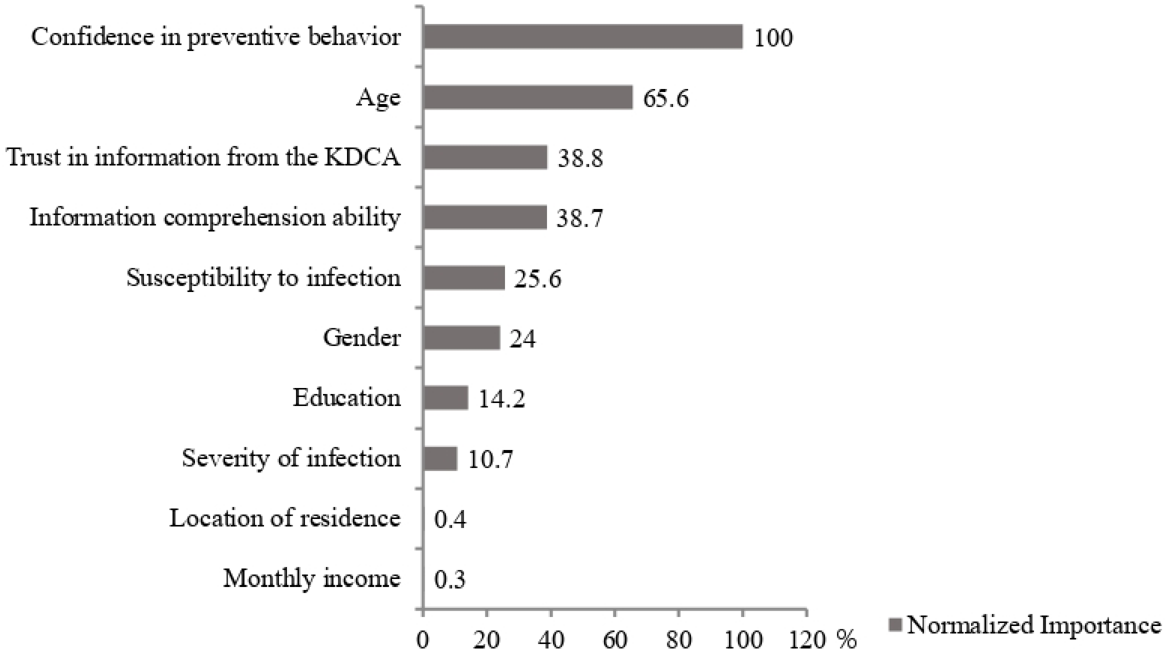


Figure 2. Normalized importance of independent variables based on the classification and regression tree results for the National Survey of Infectious Disease Preventive Behaviors in Community (conducted on October 2020 in South Korea). KDCA: Korea Centers for Disease Control and Prevention Agency.



Discussion

Principal Findings

The purpose of our study was to assess levels of compliance with the recommended PHSMs against SARS-CoV-2 infection and their predictors among the general population by using national data. There is a need for health care providers and public health officials to know which groups are in need of education on PHSMs. With regard to socioeconomic factors, sex (female), age, monthly income, and sociocognitive factors were significantly associated with more preventive behaviors. Consistent with previous research, female participants were more likely to perform preventive behaviors. During epidemics of infectious diseases, including the influenza, H1N1, and COVID-19 pandemics, it was shown that women are more sensitive to risk and are more motivated to perform health behaviors than men [25-27]. Interestingly, people who were older were more likely to perform preventive behaviors than those who were younger, which is consistent with a study by Raude et al [28]. The possible explanation for this finding is that people who are older may believe that they are more prone to infection and that they will be sicker and experience more complications once they are infected and thus engage in more preventive behaviors [29].

Surprisingly, those with lower incomes were more engaged in PHSMs. A previous study reported that people with higher incomes tend to stay at home more often when compared to those with lower incomes [29] and show higher compliance with preventive behaviors against COVID-19 when compared to those with a low economic status [11]. People with lower incomes may engage in jobs that are unstable and may work in environments that increase the likelihood of infection (eg, environments in which the proximity between people cannot be maintained or confined areas with limited ventilation), which may place them at risk for infection. However, the relationship between income and preventive behavior needs to be examined in future studies, as the exact pathway for why people with lower incomes are more engaged in preventive behaviors is not well established in the literature.

Consistent with previous findings [13,26,30,31], the influence of confidence in the ability to control a situation on health behavior change was found to be a strong determinant of adherence to PHSMs for preventing COVID-19 and the intention to follow preventive behaviors [32]. Moreover, in this study, we attempted to identify target groups that are in the most need of intervention to increase preventive behaviors, and we found that men aged between 19 and 49 years with lower scores for perceived confidence in performing preventive behaviors are in the most need of intervention. As this age group is the most active at work and in daily life, it should be given a higher priority in terms of their potential to spread infection to others. Given that the majority of the participants were employed (2,933/4003, 73%), policies and guidelines encouraging the practice of preventive behaviors and the provision of essential infection control materials should be established at work. In addition, men tended to dismiss preventive actions when compared to women in this study, which is consistent with the findings from a study of 8 countries [33]. Thus, gender perspectives are taken into account when designing education programs to promote preventive behaviors among a target population.

It is well believed that a person with high perceived susceptibility is more likely to take actions to reduce the risk of acquiring a disease. Thus, perceived susceptibility gained importance as a significant psychosocial factor related to COVID-19 preventive behaviors in earlier studies [32,34,35]. However, the temporal relationship between perceived susceptibility and preventive behaviors should be interpreted with caution. In contrast to these previous studies, we found a negative association between perceived susceptibility to infection and preventive behaviors. Since there was a negative association between perceived susceptibility and preventive behaviors, we may interpret the results as follows: those who do not perform preventive behaviors or perform fewer preventive behaviors may believe that they are at more risk of contracting diseases. In addition, as addressed in an earlier study on the use of mobile apps to trace COVID-19 cases [36], considering the relatively high adherence to PHSMs among the majority of participants in this study (34.91/40), another possible reason could be that the better the preventive action, the lower the perceived risk of contracting COVID-19. However, this study was cross-sectional; thus, we cannot assume this causal inference. Longitudinal studies are needed in the future to examine the causal relationship between sociocognitive factors and preventive behaviors.

The more a person believes that a given disease is serious, the more that person will try to reduce the chance of contracting that disease by performing preventive behaviors [37]. Health literacy is the knowledge and competence required to understand and process health information, and it influences individuals to make proper health decisions [16]. In fact, better health literacy, which we defined as information comprehension ability, was associated with more preventive behaviors in our study, which is consistent with previous studies [37,38]. The most important health information includes knowledge about COVID-19, skills for preventive behaviors, and government policies [38,39]. There is a lot of information about COVID-19 released by the media, which includes false information. Good health literacy enables people to differentiate between facts and myths regarding information about COVID-19, which influence people to perform appropriate preventive behaviors [17]. Health messages about COVID-19 are extremely confusing and change rapidly. People who do not trust information released by the health care system and the government are less likely to follow recommended behaviors [14]. In fact, our study results indicate that people with less trust in the information from the KDCA are less likely to perform preventive behaviors. Adequate government action and policy development are needed to ban the release of false information that may cause irrational panic and prevent people from practicing preventive behaviors. Moreover, demographic variables, such as low education and older age, are known to be associated with low health literacy levels [38,40]; thus, people with low educational levels and those who are older need to be targeted for intervention to increase health literacy.

Some limitations should be noted. First, our study was cross-sectional; thus, temporal inferences should be interpreted with caution. Second, the sample size may be small, although we included nationally representative samples by using stratified sampling and a proportional allocation method. Thus, longitudinal studies with larger sample sizes are needed to confirm the causality between preventive behaviors and predictors. Third, the primary study was conducted by using questionnaires; thus, recall bias may be present. Fourth, some of the variables, such as sociocognitive factors, were measured with a single item; therefore, there is a possibility that we did not capture the multidimensional concepts of each variable. In future studies, more reliable and valid instruments that capture all aspects of sociocognitive factors may need to be used. Fifth, we only included individual factors in the analysis. Thus, multilevel analyses may be necessary, as preventive behaviors could be affected by community and national factors. Sixth, this study was implemented before the outbreaks of COVID-19 variants, including the Omicron variants. Therefore, the study results need to be interpreted with caution in the context of the new COVID-19 variants. Lastly, this study did not include variables related to the vaccine awareness, as this study was conducted before the start of national vaccination in South Korea. Further studies are needed to assess the relationships between vaccination and preventive behaviors. Despite these limitations, our findings add new knowledge to previous literature, as we identified predictors of preventive behaviors by using nationally representative samples collected during the COVID-19 pandemic in Korea.

Conclusion

The purpose of this study was to assess levels of compliance with the recommended preventive measures against SARS-CoV-2 infection and their predictors among the general population by using national data. In addition to identifying predictors, we have provided insights into the specific groups that may need targeted interventions, such as men aged between 19 and 49 years with low perceived confidence in performing preventive behaviors. A theory-driven intervention may need to be carefully designed to meet the needs of the target population, and such an intervention should be delivered at appropriate places, such as workplaces, to reach a large number of people. Moreover, efforts to increase health literacy and trust in the health care system are needed at the community and national levels. However, the lack of a temporal relationship between preventive behaviors and predictors indicates the need for longitudinal studies in the future.

Conflicts of Interest

None declared.

References

1. Bonotti M, Zech ST. The human, economic, social, and political costs of COVID-19. In: Recovering Civility During COVID-19. Singapore: Palgrave Macmillan; Mar 03, 2021:1-36.
2. Osterrieder A, Cuman G, Pan-Ngum W, Cheah PK, Cheah PK, Peerawaranun P, et al. Economic and social impacts of COVID-19 and public health measures: results from an anonymous online survey in Thailand, Malaysia, the UK, Italy and Slovenia. BMJ Open 2021 Jul 20;11(7):e046863 [FREE Full text] [doi: 10.1136/bmjopen-2020-046863] [Medline: 34285007]
3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534 [FREE Full text] [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
4. Kim SY, Yeniova A. Global, regional, and national incidence and mortality of COVID-19 in 237 countries and territories, January 2022: a systematic analysis for World Health Organization COVID-19 Dashboard. Life Cycle 2022 May 12;2:e10 [FREE Full text] [doi: 10.54724/lc.2022.e10]
5. Maqbool A, Khan NZ. Analyzing barriers for implementation of public health and social measures to prevent the transmission of COVID-19 disease using DEMATEL method. Diabetes Metab Syndr 2020;14(5):887-892 [FREE Full text] [doi: 10.1016/j.dsx.2020.06.024] [Medline: 32563940]
6. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 27;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
7. Chen X, Ran L, Liu Q, Hu Q, Du X, Tan X. Hand hygiene, mask-wearing behaviors and its associated factors during the COVID-19 epidemic: A cross-sectional study among primary school students in Wuhan, China. Int J Environ Res Public Health 2020 Apr 22;17(8):2893 [FREE Full text] [doi: 10.3390/ijerph17082893] [Medline: 32331344]
8. Lahiri A, Jha SS, Bhattacharya S, Ray S, Chakraborty A. Effectiveness of preventive measures against COVID-19: A systematic review of in silico modeling studies in Indian context. Indian J Public Health 2020 Jun;64(Supplement):S156-S167 [FREE Full text] [doi: 10.4103/ijph.IJPH_464_20] [Medline: 32496248]
9. Coronavirus disease-19, Republic of Korea. Korea Centers for Disease Control and Prevention Agency (KDCA). URL: http://ncov.mohw.go.kr/en/ [accessed 2021-04-20]
10. Alsan M, Stantcheva S, Yang D, Cutler D. Disparities in coronavirus 2019 reported incidence, knowledge, and behavior among US adults. JAMA Netw Open 2020 Jun 01;3(6):e2012403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.12403] [Medline: 32556260]

11. Firouzbakht M, Omidvar S, Firouzbakht S, Asadi-Amoli A. COVID-19 preventive behaviors and influencing factors in the Iranian population; a web-based survey. BMC Public Health 2021 Jan 15;21(1):143 [FREE Full text] [doi: 10.1186/s12889-021-10201-4] [Medline: 33451303]
12. Papageorge NW, Zahn MV, Belot M, van den Broek-Altenburg E, Choi S, Jamison JC, et al. Socio-demographic factors associated with self-protecting behavior during the Covid-19 pandemic. J Popul Econ 2021;34(2):691-738 [FREE Full text] [doi: 10.1007/s00148-020-00818-x] [Medline: 33462529]
13. Shahnazi H, Ahmadi-Livani M, Pahlavanzadeh B, Rajabi A, Hamrah MS, Charkazi A. Assessing preventive health behaviors from COVID-19: a cross sectional study with health belief model in Golestan Province, Northern of Iran. Infect Dis Poverty 2020 Nov 17;9(1):157 [FREE Full text] [doi: 10.1186/s40249-020-00776-2] [Medline: 33203453]
14. Weiss BD, Paasche-Orlow MK. Disparities in adherence to COVID-19 public health recommendations. Health Lit Res Pract 2020 Aug 06;4(3):e171-e173 [FREE Full text] [doi: 10.3928/24748307-20200723-01] [Medline: 32926173]
15. McCaffery KJ, Dodd RH, Cvejic E, Ayrek J, Batcup C, Isautier JM, et al. Health literacy and disparities in COVID-19-related knowledge, attitudes, beliefs and behaviours in Australia. Public Health Res Pract 2020 Dec 09;30(4):30342012 [FREE Full text] [doi: 10.17061/phrp30342012] [Medline: 33294907]
16. Abdel-Latif MMM. The enigma of health literacy and COVID-19 pandemic. Public Health 2020 Aug;185:95-96 [FREE Full text] [doi: 10.1016/j.puhe.2020.06.030] [Medline: 32593054]
17. Spring H. Health literacy and COVID-19. Health Info Libr J 2020 Sep;37(3):171-172 [FREE Full text] [doi: 10.1111/hir.12322] [Medline: 32672399]
18. Rosenstock IM. Historical origins of the Health Belief Model. Health Educ Monogr 1974 Dec;2(4):328-335 [FREE Full text] [doi: 10.1177/109019817400200403]
19. Tong KK, Chen JH, Yu EWY, Wu AMS. Adherence to COVID-19 precautionary measures: Applying the Health Belief Model and generalised social beliefs to a probability community sample. Appl Psychol Health Well Being 2020 Dec;12(4):1205-1223 [FREE Full text] [doi: 10.1111/aphw.12230] [Medline: 33010119]
20. Yehualashet SS, Asefa KK, Mekonnen AG, Gemeda BN, Shiferaw WS, Aynalem YA, et al. Predictors of adherence to COVID-19 prevention measure among communities in North Shoa Zone, Ethiopia based on health belief model: A cross-sectional study. PLoS One 2021 Jan 22;16(1):e0246006 [FREE Full text] [doi: 10.1371/journal.pone.0246006] [Medline: 33481962]
21. Kim HS, Park JH. Predictors of MERS-related preventive behaviors performance among clinical practice students in a tertiary hospital. Journal of the Korea Academia-Industrial Cooperation Society 2018 Sep 30;19(9):174-185 [FREE Full text] [doi: 10.5762/KAIS.2018.19.9.174]
22. Lee SW. Methods for testing statistical differences between groups in medical research: statistical standard and guideline of Life Cycle Committee. Life Cycle 2022 Jan 24;2:e1 [FREE Full text] [doi: 10.54724/lc.2022.e1]
23. Tobin J. Estimation of relationships for limited dependent variables. Econometrica 1958 Jan;26(1):24-36. [doi: 10.2307/1907382]
24. Breiman L, Friedman JH, Olshen RA, Stone CJ. Regression trees. In: Classification and Regression Trees. New York: Routledge; 1984.
25. Cvetković VM, Nikolić N, Nenadić UR, Öcal A, Noji EK, Zečević M. Preparedness and preventive behaviors for a pandemic disaster caused by COVID-19 in Serbia. Int J Environ Res Public Health 2020 Jun 09;17(11):4124 [FREE Full text] [doi: 10.3390/ijerph17114124] [Medline: 32527056]
26. Kim S, Kim S. Analysis of the impact of health beliefs and resource factors on preventive behaviors against the COVID-19 pandemic. Int J Environ Res Public Health 2020 Nov 22;17(22):8666 [FREE Full text] [doi: 10.3390/ijerph17228666] [Medline: 33266386]
27. Prati G, Pietrantoni L, Zani B. A social-cognitive model of pandemic influenza H1N1 risk perception and recommended behaviors in Italy. Risk Anal 2011 Apr;31(4):645-656. [doi: 10.1111/j.1539-6924.2010.01529.x] [Medline: 21077927]
28. Raude J, Lecrique JM, Lasbeur L, Leon C, Guignard R, du Roscoät E, et al. Determinants of preventive behaviors in response to the COVID-19 pandemic in France: Comparing the sociocultural, psychosocial, and social cognitive explanations. Front Psychol 2020 Nov 30;11:584500 [FREE Full text] [doi: 10.3389/fpsyg.2020.584500] [Medline: 33329241]
29. Irigoyen-Camacho ME, Velazquez-Alva MC, Zepeda-Zepeda MA, Cabrer-Rosales MF, Lazarevich I, Castaño-Seiquer A. Effect of income level and perception of susceptibility and severity of COVID-19 on stay-at-home preventive behavior in a group of older adults in Mexico City. Int J Environ Res Public Health 2020 Oct 12;17(20):7418 [FREE Full text] [doi: 10.3390/ijerph17207418] [Medline: 33053788]
30. Duan T, Jiang H, Deng X, Zhang Q, Wang F. Government intervention, risk perception, and the adoption of protective action recommendations: Evidence from the COVID-19 prevention and control experience of China. Int J Environ Res Public Health 2020 May 13;17(10):3387 [FREE Full text] [doi: 10.3390/ijerph17103387] [Medline: 32414013]
31. Lau JTF, Griffiths S, Choi KC, Lin C. Prevalence of preventive behaviors and associated factors during early phase of the H1N1 influenza epidemic. Am J Infect Control 2010 Jun;38(5):374-380 [FREE Full text] [doi: 10.1016/j.ajic.2010.03.002] [Medline: 20569849]
32. Prasetyo YT, Castillo AM, Salonga LJ, Sia JA, Seneta JA. Factors affecting perceived effectiveness of COVID-19 prevention measures among Filipinos during enhanced community quarantine in Luzon, Philippines: Integrating protection motivation

theory and extended theory of planned behavior. Int J Infect Dis 2020 Oct;99:312-323 [FREE Full text] [doi: 10.1016/j.ijid.2020.07.074] [Medline: 32768695]
33. Galasso V, Pons V, Profeta P, Becher M, Brouard S, Foucault M. Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries. Proc Natl Acad Sci U S A 2020 Nov 03;117(44):27285-27291 [FREE Full text] [doi: 10.1073/pnas.2012520117] [Medline: 33060298]
34. Yıldırım M, Geçer E, Akgül Ö. The impacts of vulnerability, perceived risk, and fear on preventive behaviours against COVID-19. Psychol Health Med 2021 Jan;26(1):35-43. [doi: 10.1080/13548506.2020.1776891] [Medline: 32490689]
35. Lee M, You M. Psychological and behavioral responses in South Korea during the early stages of coronavirus disease 2019 (COVID-19). Int J Environ Res Public Health 2020 Apr 25;17(9):2977 [FREE Full text] [doi: 10.3390/ijerph17092977] [Medline: 32344809]
36. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: A Health Belief Model approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
37. Gautam V, Dileepan S, Rustagi N, Mittal A, Patel M, Shafi S, et al. Health literacy, preventive COVID 19 behaviour and adherence to chronic disease treatment during lockdown among patients registered at primary health facility in urban Jodhpur, Rajasthan. Diabetes Metab Syndr 2021;15(1):205-211 [FREE Full text] [doi: 10.1016/j.dsx.2020.12.023] [Medline: 33387854]
38. Wang H, Cheong PL, Wu J, Van IK. Health literacy regarding infectious disease predicts COVID-19 preventive behaviors: A pathway analysis. Asia Pac J Public Health 2021 Jul;33(5):523-529 [FREE Full text] [doi: 10.1177/10105395211013923] [Medline: 33938264]
39. Kamenidou IE, Stavrianea A, Liava C. Achieving a Covid-19 free country: Citizens preventive measures and communication pathways. Int J Environ Res Public Health 2020 Jun 27;17(13):4633 [FREE Full text] [doi: 10.3390/ijerph17134633] [Medline: 32605097]
40. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, HLS-EU Consortium. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058 [FREE Full text] [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]

Abbreviations

CART: classification and regression tree
CIDPB: comprehensive infectious disease prevention behavior
KDCA: Korea Centers for Disease Control and Prevention Agency
PHSM: public health and social measure

Edited by T Sanchez, A Mavragani; submitted 18.12.21; peer-reviewed by M Salimi, M Raimi, S Kalia, DK Yon; comments to author 06.10.22; revised version received 27.10.22; accepted 29.11.22; published 17.01.23.

Please cite as:
Hong HC, Lee H, Lee SJ, Park C, Lee M
The Determinants of Adherence to Public Health and Social Measures Against COVID-19 Among the General Population in South Korea: National Survey Study
JMIR Public Health Surveill 2023;9:e35784
URL: https://publichealth.jmir.org/2023/1/e35784
doi:10.2196/35784
PMID:36446132

Original Paper

Social Support and Technology Use and Their Association With Mental and Physical Health During the COVID-19 Pandemic Among Asian Americans: The COMPASS Cross-sectional Study

Linda G Park[1], MS, PhD; Oanh L Meyer[2], MAS, PhD; Marcelle M Dougan[3], MPH, MEng, SCD; Bethany Golden[4], MSN; Kevin Ta[2], BA; Bora Nam[1], MSN, PhD; Janice Y Tsoh[5,6,7], PhD; Marian Tzuang[1], MSW, PhD; Van M Ta Park[1,6,7], MPH, PhD

[1]Department of Community Health Systems, School of Nursing, University of California, San Francisco, San Francisco, CA, United States
[2]Department of Neurology, School of Medicine, University of California, Davis, Sacramento, CA, United States
[3]Department of Public Health and Recreation, San José State University, San Jose, CA, United States
[4]Department of Family Health Care Nursing, School of Nursing, University of California, San Francisco, San Francisco, CA, United States
[5]Department of Psychiatry and Behavioral Sciences, School of Medicine, University of California, San Francisco, San Francisco, CA, United States
[6]Asian American Research Center on Health, University of California, San Francisco, San Francisco, CA, United States
[7]Multiethnic Health Equity Research Center, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Linda G Park, MS, PhD
Department of Community Health Systems
School of Nursing
University of California, San Francisco
2 Koret Way
San Francisco, CA, 94143
United States
Phone: 1 415 502 6616
Email: Linda.Park@ucsf.edu

Abstract

Background: The global COVID-19 pandemic disproportionately affected Asian Americans and Pacific Islanders (AAPIs) and revealed significant health disparities with reports of increased discrimination and xenophobia. Among AAPIs, the pandemic exacerbated their social, linguistic, and geographic isolation. Social support may be especially important for AAPIs given the salience of collectivism as a cultural value. Another mechanism for support among AAPIs was technology use, as it is generally widespread among this population. However, older adults may not perceive the same benefits.

Objective: We examined social support and technology use and their relationships with mental and physical health outcomes through the COVID-19 pandemic among AAPIs.

Methods: Data were drawn from the COVID-19 Effects on the Mental and Physical Health of AAPI Survey Study (COMPASS) for the time period of October 2020 to February 2021. COMPASS was a cross-sectional, multilingual, national survey conducted online, by phone, and in person with AAPI adults who were ≥18 years of age, in collaboration with academic and community partners in the United States. Data were analyzed using multivariable linear regression using the outcome variables of mental and physical health with various predictors such as social support and technology use. We tested for interactions specific to age and ethnicity.

Results: Among 4631 AAPIs (mean age 45.9, SD 16.3 years; 2992/4631, 63.1% female), we found that (1) increased social support was associated with better physical health, (2) total social support was positively associated with better mental health, (3) higher technology use was associated with poorer mental health and inversely associated with poorer physical health, (4) the association of technology use with mental health was weaker among those with low social support (vs those with high social support), (5) adults younger than 60 years old (vs ≥60 years old) were more negatively affected with social support and mental health, and (6) Korean Americans appeared to be a high-risk group for poor physical health with increased technology use.

Conclusions: Our paper identified mental and physical health needs along with supportive therapies observed among AAPIs during the pandemic. Future research on how social support can be leveraged, especially among AAPIs younger than 60 years old, and how various types of technology are being utilized are important to guide the recovery efforts to address both mental and physical disparities across communities as a result of the COVID-19 pandemic.

(JMIR Public Health Surveill 2023;9:e35748) doi:10.2196/35748

KEYWORDS

health disparities; mental health; depression; anxiety; social support technology; COVID-19; pandemic; disparity; support; technology; physical health; race; survey; population; discrimination; outcome; AAPI

Introduction

The COVID-19 pandemic has had widespread health, social, and economic implications that the world has not experienced in modern history. Asian Americans and Pacific Islanders (AAPIs) are the most rapidly growing and heterogeneous racial group in the United States, with very limited research participation [1-3]. It has been brought to the forefront that significant health disparities, socioeconomic inequalities, and discrimination and xenophobia exist for AAPIs, both prior to and due to COVID-19 [4-6]. As a result of policies (eg, shelter-in-place [SIP], social distancing) that were implemented, persons and communities of color and those who are socially and technologically isolated were among the most vulnerable in terms of the adverse effects of COVID-19 [7,8]. Older AAPIs are especially vulnerable and more likely to be disproportionately affected by COVID-19 policies due to higher rates of poverty, lower educational attainment, limited English proficiency [9], and less access to digital technology [9,10] (eg, internet) or lower digital literacy [11], which has been an integral tool for social, educational, and health purposes during this pandemic [12].

The pandemic has exacerbated the social, linguistic, and geographic isolation of AAPIs [13-15]. Social isolation and the need for support are prevalent among older AAPIs whose English language proficiency is limited. Social support may be especially relevant for AAPIs given the salience of collectivism as a cultural value. Research has shown that having social connections and supportive resources can buffer the effects of social isolation and is a strong determinant of health and well-being [13,16]. Having emotional support lowers the adverse effects of stress, increases emotions that are positive, and decreases the duration and intensity of emotions that are negative [17]. Studies have shown the positive buffering effects of social support on health and well-being in AAPIs for culture and ethnicity-related stressors (eg, discrimination) [18,19].

The use of digital technology became a necessity to communicate and sustain daily activities during the COVID-19 pandemic [20,21]. Digital technology refers to a smartphone, computer, internet service, tablet, and television with cable. English-speaking Asian Americans had the highest technology adoption rates in the United States compared with other racial groups, with 62% of Asian Americans having access to an internet-connecting device such as a router (compared with 47% of Whites, 48% of Blacks, and 45% of Hispanics) and 97% owning a smartphone [22]. Although smartphone ownership and internet use have grown in popularity over the past decade, with the largest adoption rates among older adults, research is still lacking among older AAPIs, specifically those who have limited English or digital literacy and may have wide variations in income.

SIP and social distancing made the use of digital technology crucial for numerous reasons, ranging from health maintenance to social support to information gathering. Social media, for example, can provide individuals with opportunities for widening social circles, maintaining social ties, and increasing feelings of inclusion [23,24]. In examining digital technology, it is imperative to consider the ways in which technology is utilized, as different digital technology uses have different effects on mental and physical health. Nonpassive use of social media, such as interacting with others and posting online, is related with better well-being through social support [25,26]. On the other hand, passive use of digital technology, such as surfing the web, reading the news, and watching movies, is typically related to poorer well-being, as these activities require no interactions between people [27]. Additionally, consistent use of technology can create unhealthy habits (such as psychological dependence and addiction to technology use), which have been associated with higher perceived social isolation [28]. As a result of the pandemic, the use of digital technology became the only source of connection and social support available for many [29]. People have transitioned to Zoom and other video conferencing services for both work and maintaining personal relationships [30]. Patients have utilized telemedicine more than ever to receive care via phone or video [31,32].

The COVID-19 pandemic has greatly driven technology’s role as a source of social support [33-35]. For older adults, the use of digital technology has facilitated access to health resources, with the percentage of older adults who had ever participated in a telehealth visit rising sharply from 4% in 2019 to 30% in 2020 [36]. Similarly, the use of information and communication technology among older adults has facilitated different levels of social support such as making connections and emotional support [37]. Despite these findings, there is a lack of research on how AAPIs, as well as older AAPIs, utilize digital technology as a form of social support as a buffering effect alongside stress and isolation.

Given the importance of social support and technology use in the era of a global COVID-19 pandemic, we sought to collect data in a national survey (COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study [COMPASS]) that represent the experiences of AAPIs. The objective of this paper was to examine social support and use of digital technology and their associations with mental and physical health during the COVID-19 pandemic among AAPIs. This study sought to inform mental health providers, public health officials, and community partners about the role of social support in mental and physical health during the pandemic as well as the impact of the increased use of technology in AAPI communities.

Methods

Recruitment

There were 5418 participants who completed the survey from October 24, 2020, to April 10, 2021. The COMPASS survey was distributed through community organizations that serve AAPIs, personal and professional networks via email and listservs, social media, flyers, and directed ethnic media. The Collaborative Approach for Asian Americans and Pacific Islanders Research and Education (CARE) research recruitment registry [38,39] was also used to recruit via email. A US $10 gift card was offered upon survey completion. To encourage participation, the survey could be completed online on the COMPASS website [40], by phone, and limited in-person survey administration assistance by COMPASS' staff and community partners. Of the participants, 86.0% (4563/5304) completed the survey independently, and 14.0% (741/5304) had assistance from family, friends, or research staff. To maximize the use of available data, we excluded 787 participants who did not provide information on social support (n=729) or mental health (n=58) from the mental health analyses. Participants who did not provide information on technology use (n=106) or on their physical health (n=224) were further excluded from the analyses including these variables.

Ethical Approval

We obtained human subject research approval from the Institutional Review Board at University of California, San Francisco (Protocol #20-31925), and informed consent was obtained from the participants prior to completing the COMPASS cross-sectional online survey.

Study Eligibility and Procedures

Eligible participants were required to be 18 years of age or older; live within the United States; self-identify ethnically as an AAPI (full or combined with another race or ethnicity); and comprehend written English or Chinese (traditional or simplified Chinese), Korean, Samoan, or Vietnamese. The World Health Organization's process of adapting instruments [41], including backward and forward translation, was used to guide the multilanguage survey's scale and material development. Most participants completed the survey in English (3503/5418, 64.5%); however, there were 4.8% (261/5418) surveys completed in simplified Chinese, 4.7% (253/5418) completed in traditional Chinese, 12.7% (686/5418) completed in Korean, 3.5% (189/5418) completed in Samoan, and 9.7% (526/5418) completed in Vietnamese.

Measures

Mental health was assessed using the Patient Health Questionnaire-4 (PHQ-4) scale for depressive and anxiety symptoms [42]. The PHQ-4 consists of 2 items that measure depressive symptoms [43] and 2 items that measure anxiety symptoms [44]. Each item had response options ranging from 0 to 3, and a PHQ-4 score was created by summing the 4 items (range: 0-12), with higher scores indicating more severe symptoms. Mental health was measured by summing the responses across each of the 4 questions that generated a PHQ-4 score. PHQ-4 has demonstrated acceptable reliability (Cronbach alpha=.82) and construct validity [45] in the general population. In our population, the Cronbach alpha was .90, suggesting acceptable reliability. Scores of 0 to 2 indicated normal, 3 to 5 indicated mild, 6 to 8 indicated moderate, and 9 to 12 indicated severe depression or anxiety. In descriptive analyses by age, we also examined total PHQ-4 score as a categorical variable by comparing those older and younger than 60 years.

General perceived physical health was measured by asking participants to indicate their health "today" using the EuroQol EQ-5D [46,47]. The scale ranges from 0 (worst) to 100 (the best health you can imagine). EQ-5D has been translated and validated in many languages (including Chinese, Korean, and Vietnamese) [48].

Social support was measured using the 6-item Social Support Inventory from the National Latino Asian American Study. Total social support was examined as an average across 6 items: 3 related to family support and 3 related to friend support. The questions included: "How often do you talk on the phone or get together with relatives/friends?"; "How much can you rely on relatives/friends for help with a serious problem?"; and "How much can you open up to family/friends and talk about worries?" Corresponding points were added for each level of support (from 1 to 5 or 1 to 4, as appropriate), with the responses "Refused" and "Don't know" coded as missing. These were then reverse scored, such that higher values corresponded with more social support. Average scores were computed by summing all the individual items and dividing by 6. The Cronbach alpha across all 6 measures was .78.

The COMPASS team created a survey to assess patterns of technology use during the COVID-19 pandemic (Multimedia Appendix 1). The questions related to this study included "Did your use of technology increase during the COVID-19 crisis?" Response options were in the following 5 categories: "Did not increase," "Increased by 1-2 hours/day," "Increased by 3-4 hours/day," "Increased by 5-6 hours/day," and "Increased by 7+ hours/day." As we were interested primarily in the effects of increased technology use, we combined the 2 lowest categories of "Did not increase" and "Increased by 1-2 hours/day."

For sociodemographic and other COVID-19–related characteristics, participants were asked about their year of birth, sex, sexual orientation, race, ethnic or cultural group, marital status, country of birth, education, employment status, and annual household income in 2019. The question asking about the length of SIP orders was developed by the COMPASS team and asked: "How long was the shelter-in-place (or stay-at-home) order at where you live?" The response options included 0 (no order), 1 (<1 month), 2 (1-2 months), and 3 (≥2 months). The question about perceived severity of COVID-19 was also developed by the COMPASS team and asked the participants: "How would you rate the severity of COVID-19 outbreak where you live in comparison to other locations in the U.S.?" Response options were 1 (a lot less severe than most other places in the

U.S.), 2 (somewhat less severe), 3 (about the same), 4 (somewhat more severe), and 5 (a lot more severe). The effect of COVID-19 on family income or employment was measured by 1 of the items from the Coronavirus Impact Scale [49]. Participants were asked to rate how much the COVID-19 pandemic has changed their family income or employment. Response options included 0 (No change), 1 (Mild. Small change; able to meet all needs and pay bills), 2 (Moderate. Having to make cuts but able to meet basic needs and pay bills), and 3 (Severe. Unable to meet basic needs and/or pay bills).

Statistical Analysis

This study analyzed 2 outcome variables, specifically mental health and physical health. Mental health was measured using the PHQ-4 score (sum of 4 items) that was modeled as a continuous variable using multivariable linear regression analysis, with higher numbers indicating weaker mental health, which is consistent with how previous researchers have analyzed it [42,45]. A total of 3 or more for the first 2 PHQ-4 items is indicative of anxiety, and a total of 3 or more for the last 2 items is indicative of depression. Elevated depression or anxiety was indicated by a total of 3 or more across all 4 items. Physical health was also modeled as a continuous variable using multivariable linear regression analysis, with higher numbers indicating better physical health [46,47]. For the main independent variables of interest, social support was modeled as a continuous variable, while technology use was modeled as a categorical variable.

Descriptive statistics were used to report the sociodemographic characteristics of the sample. In the multivariable linear regression models, we adjusted for sociodemographic variables (age, cultural group, sex, country of birth, language proficiency, marital status, employment, education, income) since these variables are known to influence physical and mental health outcomes. We also adjusted for COVID-19–related variables (change in family income, length of SIP, perceived severity of COVID-19, and region) as they were contextual in the era of the pandemic and hypothesized to influence health outcomes. We examined multicollinearity using the variance inflation factor (VIF). The VIF for each of the variables in all physical and mental health models was <5, suggesting little or no multicollinearity between the variables.

To model interactions, we selected hypothesized modifiers of the relationship between our main independent variables and outcomes, specifically age and ethnicity. Age was categorized as a dichotomous variable representing age 60 years or older versus younger than 60 years old based on the age that is considered elderly by international standards [50]. For ethnicity, we selected the 3 largest groups to represent adequate sample sizes for interaction analyses: Ethnic Chinese (n=1560), Korean (n=1007), and Vietnamese (n=841). We examined interactions with age and ethnicity to detect age-related and ethnic or cultural differences, respectively, that may be related to all of the variables of interest (ie, social support, mental and physical health, technology use). We modeled interactions by multiplying the independent variable with the categorical variable (age and ethnicity) and used the Wald test for statistical significance. Tests for changes in technology use were conducted by modeling increase in technology use as a categorical variable with 5 categories ranging from decreased or stayed the same to ≥7–hour per-day increase, with decreased or staying the same as the reference group and using the Wald test for statistical significance. We also tested for a linear trend in technology use using an ordinal variable for technology use and assessing whether the *P* value was significant.

We sought to test the joint effect of social support and technology use on mental and physical health by modeling social support as both a continuous variable and a dichotomous variable (categorized below and above the median) and technology use as an ordinal variable. For the purposes of interpretability of the results, the interaction between social support and technology use was described using social support as a dichotomous variable (at or above the median vs below) and modeling technology use as an ordinal variable. We then used *P* values to assess whether the interaction was statistically significant for technology and high versus low social support.

All statistical tests were 2-sided with alphas for statistical significance set at ≤.05 for all tests. Statistical analyses were conducted using SAS software [51].

Results

Participant Characteristics

Table 1 describes the participant characteristics (n=4631). The major cultural groups were ethnic Chinese, Korean, and Vietnamese. Participants were mostly female (2992/4631, 63.1%). The mean age of participants was 45.9 (SD 16.3) years. Overall, the majority of participants were foreign-born (2976/4631, 64.3%); these foreign-born participants had lived in the United States an average of 24.7 (SD 15.1) years. The mean physical health score was 78.1 (SD 14.9) on a scale of 0 to 100.

Overall, based on the PHQ-4, 37.7% (1745/4631) of participants had elevated depression or anxiety. The prevalence of mild depression among those older than 60 years was lower than in those younger than 60 years (212/1044, 20.3% vs 1030/3587, 28.7%, respectively). Similarly, the prevalence of moderate (32/1044, 3.1% vs 320/3587, 8.9%) and severe depression (14/1044, 1.4% vs 137/3587, 3.8%) was lower among those aged 60 years and older compared with those younger than 60 years, respectively.

Table 1. Sociodemographic and background characteristics for all participants (n=4631).

Characteristics	Results
Cultural group, n (%)	
Asian-Indian	282 (6.1)
Ethnic Chinese[a]	1560 (33.7)
Filipino	165 (3.6)
Hmong	98 (2.1)
Japanese	202 (4.4)
Korean	1007 (21.7)
Native Hawaiian/Pacific Islander	246 (5.3)
Vietnamese	841 (18.2)
Other/mixed	232 (5.0)
Sex, n (%)	
Female	2992 (63.1)
Male	1669 (36.0)
Other/decline to answer	40 (0.9)
Sexual orientation, n (%)	
Heterosexual	4247 (91.7)
Not heterosexual	209 (4.5)
Decline to answer	175 (3.8)
Age (years), mean (SD)	45.9 (16.3)
Age (years), range	18-97
Age (years), n (%)	
18-30	967 (20.9)
30-39	797 (17.2)
40-49	825 (17.8)
50-59	998 (21.6)
≥60	1004 (22.5)
Country of birth, n (%)	
United States	1593 (34.4)
Foreign country	2976 (64.3)
Other	1 (0.02)
Don’t know	61 (1.3)
Limited English proficiency, n (%)	
Yes	1092 (23.6)
No	3539 (76.4)
Marital status, n (%)	
Single	1208 (26.1)
Married/living with partner	3062 (66.1)
Separated/divorced/widowed	326 (7.0)
Declined to answer	35 (0.8)
Employment status, n (%)	
Full time	2147 (46.4)
Part time	774 (16.7)

Characteristics	Results
Homemaker	372 (8.0)
Unemployed	507 (11.0)
Retired	541 (11.7)
Other/declined to answer	290 (6.3)
Education, n (%)	
High school or less	801 (17.5)
Some college or technical school	509 (11.1)
Bachelor’s degree	1648 (36.1)
Master’s degree or higher	1613 (35.3)
Household income (US $), n (%)	
≤25,000	869 (18.8)
>25,000-75,000	1230 (26.6)
>75,000-150,000	1135 (24.5)
>150,000	914 (19.7)
Declined to state	483 (10.4)
Length of SIP[b] order (n=5309, 17 missing responses), n (%)	
No order	438 (9.5)
<1 month	255 (5.5)
1 to <2 months	514 (11.1)
2 to <3 months	520 (11.3)
≥3 months	2545 (55.2)
Don't know	342 (7.4)
The severity of COVID-19 where you live (n=4613, 23 missing responses), n (%)	
A lot less	512 (11.1)
Somewhat less	794 (17.2)
About the same	984 (21.3)
Somewhat more	1337 (29.0)
A lot more	986 (21.4)
Technology use due to COVID-19 (n=4525, 106 missing responses)	
Decreased or stayed the same	913 (20.2)
Increased by 1-2 hours/day	1364 (30.1)
Increased by 3-4 hours/day	1416 (31.3)
Increased by 5-6 hours/day	547 (12.1)
Increased by ≥7 hours/day	285 (6.3)
Social support score, mean (SD)	3.01 (0.71)
Social support score, range	1-4.3
PHQ-4[c], range	0-12
PHQ-4, mean (SD)	
Entire sample	2.23 (2.62)
<60 years old	2.48 (2.71)
≥60 years old	1.36 (2.10)
Physical health score, range	0-100

Characteristics	Results
Physical health score, mean (SD)	
Entire sample	78.1 (14.9)
<60 years old	78.2 (15.0)
≥60 years old	77.6 (14.8)
Missing	236[d]

[a] Ethnic Chinese includes mainland Chinese, Hongkonger, Taiwanese, and Huaren.

[b] SIP: shelter in place.

[c] PHQ-4: Patient Health Questionnaire-4.

[d] Additional 12 missing technology use in addition to physical health; means and SD are similar for the available data.

Missing Data

There were up to 13% missing data on social support; however, only 5.4% of the physical health data were missing, and 1.6% of the mental health data were missing. Participants who were missing complete information on social support were similar to participants who were included (ie, age, sex, ethnic group) but more likely to have limited English proficiency (495/4072, 17.4% vs 234/1346, 12.2%; *P*<.001).

Social Support and Mental Health

Higher total social support was associated with better mental health. Total social support was inversely associated with PHQ-4 score in the crude and adjusted analyses. In the crude analysis, a 1-unit increase in average total support was associated with a lower total PHQ-4 score (=–0.51; 95% CI –0.61 to –0.40; *P*<.001). The association was attenuated in the adjusted analysis but remained significant: –0.35 (95% CI –0.45 to –0.25; *P*<.001; Table 2). We found evidence of a statistically significant interaction between age and social support in relation to mental health (*P*=.01), comparing those who were aged 60 years and older to those younger than 60 years old. In the stratified analyses by age, the association with total support was significantly stronger for those aged less than 60 years compared with those aged 60 years or older: =–0.563 (95% CI –0.69 to 0.44; *P*<.001) and =–0.17 (95% CI –0.34 to –0.01; *P*=.049), respectively ($P_{interaction}$=.008; Table 2).

Table 2. Social support in relation to mental health and physical health, overall and stratified by age, for all participants.

Social support	Crude		Adjusted[a]					
	Overall (n=4631), beta (95% CI)	*P* value	Overall (n=4631), beta (95% CI)	*P* value	18-60 years old (n=3587), beta (95% CI)	*P* value	≥60 years old (n=1044), beta (95% CI)	*P* value
Total support for mental health[b]	–0.51 (–0.61 to –0.40)[c]	<.001	–0.35 (–0.45 to –0.25)[d]	<.001	–0.56 (–0.69 to –0.44)[e]	<.001	–0.17 (–0.34 to –0.01)[f]	.049
Total support for physical support[g]	2.94 (2.32 to 3.56)[h]	<.001	2.81 (2.18 to 3.45)[i]	<.001	3.33 (2.60 to 4.07)[j]	<.001	0.95 (–0.36 to 2.26)[k]	.15

[a] Adjusted for age (18-30, 30-39, 40-49, 50-59, ≥60+ years), sex (male, female, other/don't know), marital status (single, married/living with partner, separate/divorced/widowed, declined to answer), employment status (full time, part-time home maker, unemployed, retired, other/declined to answer), cultural group (Asian Indian, ethnic Chinese, Filipino, Hmong, Japanese, Korean, Native Hawaiian and Pacific Islanders, Vietnamese, other/mixed), United States region (Midwest, Northeast, South, West), severity of COVID-19 in participant's location relative to the rest of the United States (a lot less, somewhat less, about the same, somewhat more, a lot more), annual household income (US $; ≤25,000, 25,001-75,000, 75,001-150,000, ≥150,001, declined to answer), education (high school or less, some college or technical school, bachelor's degree, master's degree or higher), length of shelter in place (no order, <1 month, 1 month to <2 months, 2-3 months, don't know), change in family income (no change, mild, moderate, severe), English proficiency (limited, not limited), country of birth (United States, outside the United States).

[b] $P_{interaction}$ for age x social support in relation to mental health=.008.

[c] Model R^2=0.02.

[d] Model R^2=0.21.

[e] Model R^2=0.14.

[f] Model R^2=0.12.

[g] $P_{interaction}$ for age x social support in relation to physical health=.01.

[h] Model R^2=0.02.

[i] Model R^2=0.04.

[j] Model R^2=0.05.

[k] Model R^2=0.06.

Social Support and Physical Health

Social support was positively associated with the physical health score. In the crude analysis, increased social support was associated with higher physical health scores. For every unit increase in mean total support score, the average physical health score was 2.94 (95% CI –2.32 to 3.56, $P<.001$). This association persisted in the adjusted analysis: A 1-unit increase in mean total support was associated with a 2.81 (95% CI 2.18 to 3.45; $P<.001$) increase in the physical health score (Table 2). This association appeared limited ($P_{interaction}$=.01) to those aged less than 60 years of age (=3.33, 95% CI 2.60 to 4.07; $P<.001$). The association between social support and physical health was not significant among those aged 60 years and older (Table 2).

Technology Use and Mental Health

Technology use at each level was associated with an increase in total PHQ-4 score in the crude models, an association that persisted in the adjusted models. Compared with those for whom technology use decreased or stayed the same during the pandemic, after adjusting for covariates, those whose technology increased by 7 hours or more per day had a 1.32-point (95% CI 1.00 to 1.64; *P* for the trend<.001) higher PHQ-4 score (Table 3). The differences by age group and ethnicity were not statistically significant (data not shown). Table S1 in Multimedia Appendix 2 presents the baseline technology use between ethnic groups.

Table 3. Mental health in relation to amount of technology use for all the participants.

Mental health	Crude[a], beta (95% CI)	*P* value	Adjusted[b,c], beta (95% CI)	*P* value
Decreased or stayed the same	Reference	N/A[d]	Reference	N/A
1-2 hour/day increase	0.33 (0.12 to 0.54)	.002	0.19 (–0.01 to 0.38)	.055
3-4 hour/day increase	1.11 (0.91 to 1.32)	<.001	0.63 (0.43 to 0.83)	<.001
5-6 hour/day increase	1.61 (1.34 to 1.87)	<.001	0.86 (0.61 to 1.12)	<.001
≥7 hour/day increase	2.15 (1.82 to 2.49)	<.001	1.32 (1.00 to 1.64)	<.001

[a] Model R^2=0.06; *P* for trend<.001.

[b] Adjusted for the same variables as in Table 2.

[c] Model R^2=0.22; *P* for trend<.001.

[d] N/A: not applicable.

Technology Use and Physical Health

Technology use was inversely associated with physical health. Compared with the reference group, those whose use increased by 7 hours or more a day reported a –3.48 (95% CI –5.50 to –1.46; *P* for the trend<.001) physical health score in the crude model. There was a statistically significant interaction by cultural group, with Koreans (*P* for the trend<.001) having significantly lower health scores for 7 or more hours of technology use daily compared with the same or decreased use; this was not true for ethnic Chinese (*P* for the trend=.48) or Vietnamese (*P* for the trend=.16; Table 4). The interaction by age was not significant (data not shown).

Table 4. Physical health in relation to technology use, overall and stratified by cultural group, for all participants.

Physical health	Crude[a]		Adjusted[b]							
	Overall, beta (95% CI)	*P* value	Overall[c], beta (95% CI)	*P* value	Ethnic Chinese[d,e], beta (95% CI)	*P* value	Korean[d,f], beta (95% CI)	*P* value	Vietnamese[d,g], beta (95% CI)	*P* value
Decreased or stayed the same	Reference	N/A[d]	Reference	N/A	Reference	N/A	Reference	N/A	Reference	N/A
1-2 hour/day increase	–0.80 (–2.05 to 0.44)	.21	–1.22 (–2.49 to 0.05)	.06	0.44 (–2.65 to 1.77)	.69	–1.49 (–4.94 to 1.96)	.40	0.91 (–1.54 to 3.36)	.56
3-4 hour/day increase	–2.18 (–3.41 to –0.94)	<.001	–2.55 (–3.85 to –1.25)	<.001	–0.83 (–3.03 to 1.37)	.46	–2.82 (–6.28 to 0.64)	.11	–1.68 (–4.37 to 1.00)	.22
5-6 hour/day increase	–2.87 (–4.46 to –1.28)	<.001	–3.43 (–5.10 to –1.76)	<.001	–1.58 (–4.24 to 1.09)	.25	–7.75 (–12.21 to –3.29)	.001	0.62 (–3.44 to 4.69)	.76
≥7 hour/day increase	–3.48 (–5.50 to –1.46)	<.001	–4.02 (–6.10 to –1.94)	<.001	–0.30 (–3.69 to 3.10)	.87	–9.39 (–15.02 to –3.76)	.001	–3.82 (–8.60 to 0.96)	.12

[a] Model R^2=0.02; *P* for trend<.001.

[b] Adjusted for the same variables as Table 2 except analyses stratified by cultural group

[c] Model R^2=0.04; *P* for trend<.001.

[d] $P_{interaction}$ for technology x cultural group in relation to physical health=.007

[e] Model R^2=0.03; *P* for trend=.48.

[f] Model R^2=0.04; *P* for trend<.001.

[g] Model R^2=0.08; *P* for trend=.16.

Social Support and Technology Use in Relation to Mental and Physical Health

In exploratory analyses, we examined the joint effects of social support and technology use in relation to mental and physical health. Social support and technology use were independently associated with each other in both models, and the interaction between the two was significant for both mental and physical health. We found that the effect of technology use on mental health was significantly weaker among those with low social support compared with among those with high social support (*P*=.01; Figure 1). Similarly, the effect of technology on physical health was significantly weaker among those with low social support (*P*=.02; Figure 2). The interaction finding with social support and technology use with both mental and physical health was tested for an overall effect but was not specific to certain length of time (hours) of technology use. Although there was a significant overall interaction for the measures, they were more pronounced at higher levels.

Figure 1. Linear regression analyses for joint effects of social support and technology use on mental health.

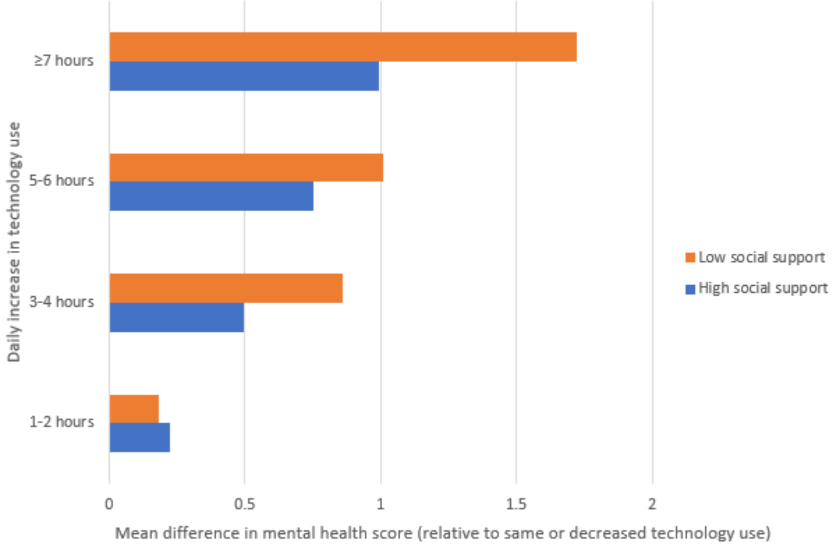
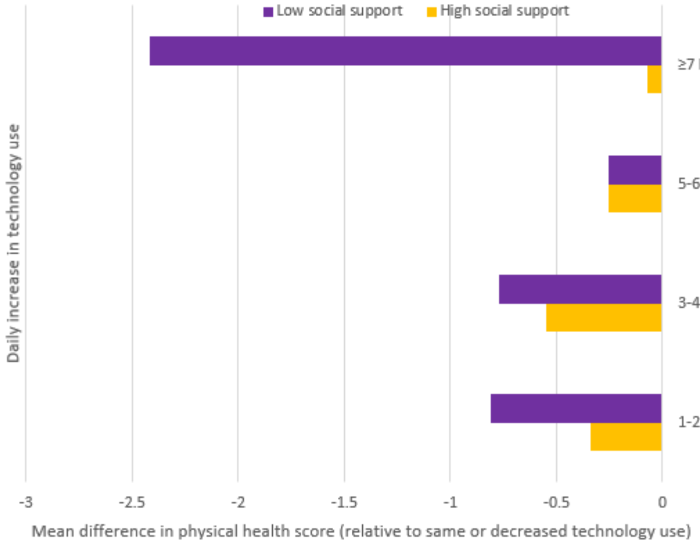


Figure 2. Linear regression analyses for joint effects of social support and technology use on physical health.



Discussion

Principal Findings

This paper presents the association of self-reported social support and technology use on the outcomes of mental and physical health during the COVID-19 pandemic as reported by over 4600 AAPI community members in a national survey. We found several significant results that represent 10 different AAPI ethnic groups; however, our paper presents stratified analyses in the 3 largest ethnic groups: Ethnic Chinese, Korean, and Vietnamese. Overall, 12.9% and 11.4% of our survey respondents reported depression and anxiety, respectively, which is markedly higher than prepandemic averages. According to the 2019 US Census Bureau assessments prior to the pandemic, only 2.9% and 3.2% of Asian Americans reported depressive and anxiety disorders, respectively (n=31,229) [52]. Among all US adults, there was a 4-fold increase in the likelihood of having depression or anxiety from 2019 to 2020, with the greatest increases among younger people, Asian Americans, and parents [52]. The average physical health score of 78.1 (out of 100) among AAPIs in our study was similar to what was reported in a prior study of over 11,500 diverse adults 18 years and older (eg, score of 79.2) [53].

In this cross-sectional study, we found that low social support was associated with more self-reported mental health symptoms (depression and anxiety) overall. Furthermore, although both age groups had these outcomes, when comparing the 2 age groups in the adjusted analyses, we found all types of social support were more significant for those younger than 60 years of age compared with those 60 years and older, suggesting greater impact of social support among younger adults with more mental health symptoms. A rise in mental health changes (ie, depression, fear, health-related anxiety) would be expected given the increased social isolation due to SIP orders and social distancing, as well as increased concern for self and loved ones (eg, health, financial stability) caused by the COVID-19 pandemic [54]. Reasons for the greater mental health impact in the younger group may be related to being accustomed to more socialization compared with those who are older who may be more resilient to social changes and less socialization [55]. No significant differences were found among AAPI ethnic groups. Potential reasons why depression and anxiety were lower in adults 60 years and older compared with younger adults include the public health message that isolation was protective against COVID-19; in addition to older adults being more used to not socializing as much as younger people, they were also likely more grateful to be alone, due to their awareness of being in an at-risk group. However, the literature on older adults, not specific to AAPI, shows that they were more likely at a high risk for mental illness from social isolation in general and social restrictions due to COVID-19 [56,57].

Beyond mental health outcomes, we found that high social support was associated with positive self-reported physical health scores. After adjusting for possible confounding variables, findings on social support and physical health scores were limited to those who were less than 60 years of age, unlike social support and mental health. Again, we did not detect significant differences among ethnic groups. By examining access and reliance on all forms of social support, our results confer our prepandemic understanding of coping in which forms of social support are used to improve health status and have a protective factor. Compared with this survey, a cross-sectional study of 3649 individuals aged 60 years and older concluded that the predictors of poor self-reported health included a lack of perceived social support in women and lack of social network involvement in men [58].

With regards to technology use during the pandemic, we found incremental increases in self-reported depressive and anxiety scores with increased use of technology; mental health symptoms (ie, higher PHQ-4 scores up to 2.15) were greater with technology use of 3 or more hours after adjusting for all possible confounders. These results are consistent with other research that showed that moderate or severe levels of depression were associated with more than 2-fold time spent watching TV and using a computer (>6 hours per day) [59]. Similarly, increased technology use was associated with poor physical health. Unlike social support, there was no significant interaction with the age groups. We did not detect significant differences among cultural groups.

When examining technology use and physical health, we detected significant changes in self-reported physical health scores with 3 or more hours per day of technology use after adjusting for all possible confounders. Consistent with our results, increased screen time has been linked to potentially harmful physical effects, as it negatively affects physical activity and sleep routine, thereby leading to headaches, neck pain, myopia, digital eye syndrome, and cardiovascular risk factors (obesity, high blood pressure, and insulin resistance) [60]. Additional analyses showed differences between cultural groups, with Korean American individuals reporting lower physical health scores than ethnic Chinese and Vietnamese individuals who had the same or lower amount of technology use. Our survey did not ask participants to specify the reason for using technology, so it is unclear whether this signals more technology use for work, leisure, or both. Based on increases in internet service use (from 40% to 100%) and video-conferencing services (10-fold increase) during the pandemic, we can assume a good portion was used for work purposes across all groups [61].

Comparisons With Prior Work

When examining the joint effects of both social support and technology use, we found significant independent associations with mental and physical health as well as a significant interaction between social support and technology use for both mental and physical health after accounting for all possible confounders. When considering technology, an important factor is social media use, with one study showing that social media exposure was associated with anxiety and combined depression and anxiety when compared with less social media exposure (1.72 and 1.91 greater odds, respectively) [12]. In terms of the interaction effect between social support and technology use on physical health, there are important potential benefits. In a study of 591 older adults, social connectedness via technology prior to the pandemic was associated with significant health benefits including better self-rated health, fewer chronic illnesses, higher subjective well-being, and fewer depressive symptoms [62].

Clinical Implications

There are important clinical implications related to our findings, particularly for mental health providers, public health officials, and community partners. Undoubtedly, the COVID-19 pandemic negatively impacted mental and physical health for the vast majority of adults and children; however, understanding the contributing factors and high-risk groups is critical. Adults younger than 60 years were more impacted by social support, as evidenced by stronger associations with lower mental health and higher physical health scores compared with those who were 60 years or older. In addition, Korean Americans appeared to be a high-risk group for poor physical health with increased technology use. Further work is needed to understand how social support is characterized and how to best operationalize it for each AAPI community to achieve the known mental and physical health benefits in times of emergency [63]. Although access to technology may have buffered loneliness and isolation [64], more research is needed on the potential for different uses of technology to improve self-care and well-being, along with the potential to bring harm to individuals.

Limitations

There are limitations that should be considered when interpreting these results, including a cross-sectional survey study design that did not allow us to assess dynamic changes over time during the pandemic. Another limitation was nonprobability sampling. We had significantly more female participants than male participants, although we adjusted for this in our analyses. Although this was a national survey, there were varying SIP and social distancing policies across regions and states, which may have affected the responses. Since the survey was not conducted in every AAPI language spoken, generalizability is limited and would be improved with more languages added to the survey. In addition, COMPASS is primarily an online survey, limiting its reach to those who lack access or are unable to utilize the internet. As mentioned, our inquiry on technology use did not distinguish between varying technologies nor whether the technologies were used for work or leisure.

Conclusions

The COVID-19 pandemic has taken its toll on both mental and physical health in many individuals and families in the United States and globally. In a large sample of AAPI recruited nationally, a significant portion of our study population reported probable anxiety and depression. The prevalence of moderate and severe depressive symptoms was lower among those aged 60 years and older compared with those younger than 60 years. Despite COVID-19–related deaths being highly concentrated in this older age group, this finding represents the resilience of older AAPI Americans during a pandemic, with less than 5% of older adults reporting moderate to severe depressive symptoms.

Importantly, this study offered new insights into contributing factors for mental and physical health status by examining perceived social support and changes in technology use among AAPI during the pandemic. The key findings suggested that greater social support during the pandemic was associated with lower levels of depression and anxiety and better physical health, whereas increased technology use during the pandemic was associated with poorer mental and physical health in AAPI. Such associations were particularly pronounced among adults younger than 60 years. Although we observed similar associations between social support and mental health across cultural groups, Korean Americans appeared to be more vulnerable to the effect of increased technology use on their physical health. These findings revealed new health disparities observed during the pandemic that require urgent attention on multiple fronts, including individual, social, and community levels. Future research on how social support can be leveraged, especially among those younger than 60 years, and how various types of technology are being utilized is important to guide the recovery efforts to address both mental and physical disparities across communities as a result of the negative impacts from the COVID-19 pandemic.

Acknowledgments

We sincerely appreciate the Asian Americans and Pacific Islander (AAPI) participants who completed the COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study (COMPASS) survey. We would like to thank the International Children Assistance Network (ICAN), National Asian Pacific Center on Aging (NAPCA), and other community partners who supported recruitment for this study.

This study was supported by the National Institutes of Health (NIH)/National Institute on Aging (NIA; 3R24AG063718-02S1, NIH/NIA R24AG063718, NIH/NIA R56AG069130). The funders had no role in study design, data collection, analysis, writing, and decision to publish. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

VMTP, LGP, OLM, MMD, and JYT designed the study and developed the methodology. LGP led the manuscript writing and drafted major sections of the paper. MMD and BG completed all statistical analyses and drafted the results, tables, and figures. VMTP, OLM, MMD, BG, KT, BN, JYT, and MT contributed to writing of the manuscript and reviewing tables and figures. VMTP acquired the research funding. All authors contributed to interpretation of results, reviewing drafts, and approval of the final manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Technology Use for Support During COVID-19 Survey.
[DOCX File , 20 KB - publichealth_v9i1e35748_app1.docx]

Multimedia Appendix 2
Baseline technology use between ethnic groups.
[DOCX File , 15 KB - publichealth_v9i1e35748_app2.docx]

References

1. Jones N, Marks R, Ramirez R, Rios-Vargas M. 2020 Census Illuminates Racial and Ethnic Composition of the Country. United States Census Bureau. 2021 Aug 12. URL: https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html [accessed 2022-12-12]
2. National Institutes of Health Revitalization Act of 1993: Clinical Research Equity Regarding Women and Minorities. National Institutes of Health. 1993. URL: https://orwh.od.nih.gov/resources/pdf/NIH-Revitalization-Act-1993.pdf [accessed 2022-12-12]
3. Chen MS, Lara PN, Dang JHT, Paterniti DA, Kelly K. Twenty years post-NIH Revitalization Act: enhancing minority participation in clinical trials (EMPaCT): laying the groundwork for improving minority clinical trial accrual: renewing the case for enhancing minority participation in cancer clinical trials. Cancer 2014 Apr 01;120 Suppl 7(0 7):1091-1096 [FREE Full text] [doi: 10.1002/cncr.28575] [Medline: 24643646]
4. Quach T, Đoàn LN, Liou J, Ponce NA. A rapid assessment of the impact of COVID-19 on Asian Americans: cross-sectional survey study. JMIR Public Health Surveill 2021 Jun 11;7(6):e23976 [FREE Full text] [doi: 10.2196/23976] [Medline: 34019478]
5. Park VT, Tsoh JY, Dougan M, Nam B, Tzuang M, Park LG, et al. Racial bias beliefs related to COVID-19 among Asian Americans, Native Hawaiians, and Pacific Islanders: findings from the COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study (COMPASS). J Med Internet Res 2022 Aug 09;24(8):e38443 [FREE Full text] [doi: 10.2196/38443] [Medline: 35658091]
6. Ta Park VM, Dougan MM, Meyer OL, Nam B, Tzuang M, Park LG, et al. Discrimination experiences during COVID-19 among a national, multi-lingual, community-based sample of Asian Americans and Pacific Islanders: COMPASS findings. Int J Environ Res Public Health 2022 Jan 14;19(2):924 [FREE Full text] [doi: 10.3390/ijerph19020924] [Medline: 35055744]
7. Ramsetty A, Adams C. Impact of the digital divide in the age of COVID-19. J Am Med Inform Assoc 2020 Jul 01;27(7):1147-1148 [FREE Full text] [doi: 10.1093/jamia/ocaa078] [Medline: 32343813]
8. Stokes EK, Zambrano LD, Anderson KN, Marder EP, Raz KM, El Burai Felix S, et al. Coronavirus Disease 2019 Case Surveillance - United States, January 22-May 30, 2020. MMWR Morb Mortal Wkly Rep 2020 Jun 19;69(24):759-765 [FREE Full text] [doi: 10.15585/mmwr.mm6924e2] [Medline: 32555134]

9. Ma KPK, Bacong AM, Kwon SC, Yi SS, Đoàn LN. The impact of structural inequities on older Asian Americans during COVID-19. Front Public Health 2021;9:690014 [FREE Full text] [doi: 10.3389/fpubh.2021.690014] [Medline: 34490181]
10. Rozenfeld Y, Beam J, Maier H, Haggerson W, Boudreau K, Carlson J, et al. A model of disparities: risk factors associated with COVID-19 infection. Int J Equity Health 2020 Jul 29;19(1):126 [FREE Full text] [doi: 10.1186/s12939-020-01242-z] [Medline: 32727486]
11. Shchafer D, Chun H, Lum W. The emerging needs of Asian American and Pacific Islander older adults. National Asian Pacific Center on Aging. 2017 Feb. URL: https://www.napca.org/wp-content/uploads/2017/10/NAPCA-The-Emerging-Needs-of-AAPI-Older-Adults_Final-Report_Feb2017.pdf [accessed 2022-12-12]
12. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
13. Muruthi JR, Lewis DC. Cambodian refugee families: impacts of immigration-related stressors on intergenerational relationships. Journal of Intergenerational Relationships 2017 May 16;15(2):125-142. [doi: 10.1080/15350770.2017.1294418]
14. Treas J, Mazumdar S. Older people in America's immigrant families. Journal of Aging Studies 2002 Aug;16(3):243-258. [doi: 10.1016/s0890-4065(02)00048-8]
15. The effects of COVID-19 related social isolation on the mental health of racialized communities. National Alliance on Mental Illness. URL: https://www.nami.org/Support-Education/NAMI-HelpLine/COVID-19-Information-and-Resources/COVID_Isolation_Guide.pdf [accessed 2022-12-12]
16. Webster NJ, Ajrouch KJ, Antonucci TC. Living healthier, living longer: The benefits of residing in community. Generations 2013;4:28-32 [FREE Full text]
17. Uchino BN, Cacioppo JT, Kiecolt-Glaser JK. The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. Psychol Bull 1996 May;119(3):488-531. [doi: 10.1037/0033-2909.119.3.488] [Medline: 8668748]
18. Lee S, Waters SF. Asians and Asian Americans' experiences of racial discrimination during the COVID-19 pandemic: Impacts on health outcomes and the buffering role of social support. Stigma and Health 2021 Feb;6(1):70-78. [doi: 10.1037/sah0000275]
19. Wei M, Yeh CJ, Chao RC, Carrera S, Su JC. Family support, self-esteem, and perceived racial discrimination among Asian American male college students. J Couns Psychol 2013 Jul;60(3):453-461. [doi: 10.1037/a0032344] [Medline: 23544839]
20. Salzano G, Passanisi S, Pira F, Sorrenti L, La Monica G, Pajno GB, et al. Quarantine due to the COVID-19 pandemic from the perspective of adolescents: the crucial role of technology. Ital J Pediatr 2021 Feb 22;47(1):40 [FREE Full text] [doi: 10.1186/s13052-021-00997-7] [Medline: 33618753]
21. Seifert A, Cotten SR, Xie B. A double burden of exclusion? digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):e99-e103 [FREE Full text] [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]
22. Marketers, It's Time to Engage Asian American Consumers. Nielsen. 2020 May. URL: https://www.nielsen.com/insights/2020/marketers-its-time-to-engage-asian-american-consumers/ [accessed 2022-12-12]
23. Ellison N, Lampe C, Steinfield C, Vitak J. With a little help from my friends: Social network sitessocial capital. In: Papacharissi Z, editor. A networked self: Identity, community and culture on social network sites. New York, NY: Routledge; 2010:124-145.
24. Seiter E. Practicing at Home: Computers, Pianos, and Cultural Capital. In: McPherson T, editor. Digital Youth, Innovation, and the Unexpected. Cambridge, MA: MIT Press; 2008:27-52.
25. Jung Y, Song H, Vorderer P. Why do people post and read personal messages in public? The motivation of using personal blogs and its effects on users' loneliness, belonging, and well-being. Computers in Human Behavior 2012 Sep;28(5):1626-1633. [doi: 10.1016/j.chb.2012.04.001]
26. Seo M, Kim J, Yang H. Frequent interaction and fast feedback predict perceived social support: using crawled and self-reported data of Facebook users. J Comput-Mediat Comm 2016 May 18;21(4):282-297. [doi: 10.1111/jcc4.12160]
27. Liu D, Baumeister RF, Yang CC, Hu B. Digital communication media use and psychological well-being: A meta-analysis. Journal of Computer-Mediated Communication 2019;27(1):zmab018. [doi: 10.1093/jcmc/zmz013]
28. Meshi D, Cotten S, Bender A. Problematic social media use and perceived social isolation in older adults: a cross-sectional study. Gerontology 2020;66(2):160-168 [FREE Full text] [doi: 10.1159/000502577] [Medline: 31522178]
29. Banerjee D, Rai M. Social isolation in Covid-19: The impact of loneliness. Int J Soc Psychiatry 2020 Sep;66(6):525-527 [FREE Full text] [doi: 10.1177/0020764020922269] [Medline: 32349580]
30. Branscombe M. The Network Impact of the Global COVID-19 Pandemic. The Newstack. 2020 Apr 14. URL: https://thenewstack.io/the-network-impact-of-the-global-covid-19-pandemic/ [accessed 2022-12-12]
31. Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
32. Pappot N, Taarnhøj GA, Pappot H. Telemedicine and e-Health solutions for COVID-19: patients' perspective. Telemed J E Health 2020 Jul;26(7):847-849. [doi: 10.1089/tmj.2020.0099] [Medline: 32329654]
33. Goldschmidt K. The COVID-19 pandemic: technology use to support the wellbeing of children. J Pediatr Nurs 2020 Jul;53:88-90 [FREE Full text] [doi: 10.1016/j.pedn.2020.04.013] [Medline: 32317129]

34. Drouin M, McDaniel BT, Pater J, Toscos T. How parents and their children used social media and technology at the beginning of the COVID-19 pandemic and associations with anxiety. Cyberpsychol Behav Soc Netw 2020 Nov;23(11):727-736. [doi: 10.1089/cyber.2020.0284] [Medline: 32726144]
35. Vargo D, Zhu L, Benwell B, Yan Z. Digital technology use during COVID‐19 pandemic: A rapid review. Human Behav and Emerg Tech 2020 Dec 28;3(1):13-24. [doi: 10.1002/hbe2.242]
36. Buis L, Singer D, Solway E, Kirch M, Kullgren J, Malani P. Telehealth Use Among Older Adults Before and During COVID-19. University of Michigan National Poll on Healthy Aging. 2020 Aug. URL: http://hdl.handle.net/2027.42/156253 [accessed 2022-12-12]
37. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018 May 08;58(3):467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]
38. Collaborative Approach for AAPI Research and Education (CARE). University of California San Francisco. URL: https://careregistry.ucsf.edu/ [accessed 2022-12-12]
39. Ta Park VM, Meyer OL, Tsoh JY, Kanaya AM, Tzuang M, Nam B, et al. The Collaborative Approach for Asian Americans and Pacific Islanders Research and Education (CARE): A recruitment registry for Alzheimer's disease and related dementias, aging, and caregiver-related research. Alzheimers Dement 2022 Apr 14:1. [doi: 10.1002/alz.12667] [Medline: 35420258]
40. COMPASS. URL: https://compass.ucsf.edu [accessed 2022-12-12]
41. Process of translation and adaptation of instruments. World Health Organization. 2018. URL: https://www.mhinnovation.net/sites/default/files/files/WHO%20Guidelines%20on%20Translation%20and%20Adaptation%20of%20Instruments.docx [accessed 2022-12-12]
42. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
43. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
44. Wild B, Eckl A, Herzog W, Niehoff D, Lechner S, Maatouk I, et al. Assessing generalized anxiety disorder in elderly people using the GAD-7 and GAD-2 scales: results of a validation study. Am J Geriatr Psychiatry 2014 Oct;22(10):1029-1038. [doi: 10.1016/j.jagp.2013.01.076] [Medline: 23768681]
45. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. J Affect Disord 2010 Apr;122(1-2):86-95. [doi: 10.1016/j.jad.2009.06.019] [Medline: 19616305]
46. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
47. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
48. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, present and future. Appl Health Econ Health Policy 2017 Apr;15(2):127-137 [FREE Full text] [doi: 10.1007/s40258-017-0310-5] [Medline: 28194657]
49. Stoddard J, Reynolds EK, Paris R, Haller S, Johnson S, Zik J, et al. The Coronavirus Impact Scale: construction, validation, and comparisons in diverse clinical samples. PsyArXiv: Preprint posted online June 10 2021 [FREE Full text] [doi: 10.31234/osf.io/kz4pg]
50. World Population Ageing 2019. United Nations Department of Economic and Social Affairs. 2019. URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf [accessed 2022-12-12]
51. SAS OnDemand for Academics. SAS Institute Inc. URL: https://support.sas.com/en/software/ondemand-for-academics-support.html [accessed 2022-12-12]
52. Twenge JM, McAllister C, Joiner TE. Anxiety and depressive symptoms in U.S. Census Bureau assessments of adults: Trends from 2019 to fall 2020 across demographic groups. J Anxiety Disord 2021 Oct;83:102455. [doi: 10.1016/j.janxdis.2021.102455] [Medline: 34332230]
53. Lubetkin EI, Jia H, Franks P, Gold MR. Relationship among sociodemographic factors, clinical conditions, and health-related quality of life: examining the EQ-5D in the U.S. general population. Qual Life Res 2005 Dec;14(10):2187-2196. [doi: 10.1007/s11136-005-8028-5] [Medline: 16328899]
54. Mertens G, Gerritsen L, Duijndam S, Salemink E, Engelhard IM. Fear of the coronavirus (COVID-19): Predictors in an online study conducted in March 2020. J Anxiety Disord 2020 Aug;74:102258 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102258] [Medline: 32569905]
55. Klaiber P, Wen JH, DeLongis A, Sin NL. The ups and downs of daily life during COVID-19: age differences in affect, stress, and positive events. J Gerontol B Psychol Sci Soc Sci 2021 Jan 18;76(2):e30-e37 [FREE Full text] [doi: 10.1093/geronb/gbaa096] [Medline: 32674138]
56. Gorenko JA, Moran C, Flynn M, Dobson K, Konnert C. Social isolation and psychological distress among older adults related to COVID-19: a narrative review of remotely-delivered interventions and recommendations. J Appl Gerontol 2021 Jan 11;40(1):3-13. [doi: 10.1177/0733464820958550] [Medline: 32914668]

57. Kotwal AA, Holt-Lunstad J, Newmark RL, Cenzer I, Smith AK, Covinsky KE, et al. Social isolation and loneliness among San Francisco Bay Area older adults during the COVID-19 shelter-in-place orders. J Am Geriatr Soc 2021 Jan 09;69(1):20-29 [FREE Full text] [doi: 10.1111/jgs.16865] [Medline: 32965024]
58. Caetano SC, Silva CM, Vettore MV. Gender differences in the association of perceived social support and social network with self-rated health status among older adults: a population-based study in Brazil. BMC Geriatr 2013 Nov 15;13(1):122 [FREE Full text] [doi: 10.1186/1471-2318-13-122] [Medline: 24229389]
59. Madhav K, Sherchand SP, Sherchan S. Association between screen time and depression among US adults. Prev Med Rep 2017 Dec;8:67-71 [FREE Full text] [doi: 10.1016/j.pmedr.2017.08.005] [Medline: 28879072]
60. Pandya A, Lodha P. Social connectedness, excessive screen time during COVID-19 and mental health: a review of current evidence. Front Hum Dyn 2021 Jul 22;3:A [FREE Full text] [doi: 10.3389/fhumd.2021.684137]
61. De' R, Pandey N, Pal A. Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. Int J Inf Manage 2020 Dec;55:102171 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102171] [Medline: 32836633]
62. Chopik W. The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol Behav Soc Netw 2016 Sep;19(9):551-556 [FREE Full text] [doi: 10.1089/cyber.2016.0151] [Medline: 27541746]
63. Uchino BN, Bowen K, de Grey RK, Mikel J, Fisher EB. Social Support and Physical Health: Models, Mechanisms, and Opportunities. In: Principles and Concepts of Behavioral Medicine. New York, NY: Springer; 2018:341-372.
64. Saltzman LY, Hansel TC, Bordnick PS. Loneliness, isolation, and social support factors in post-COVID-19 mental health. Psychol Trauma 2020 Aug;12(S1):S55-S57. [doi: 10.1037/tra0000703] [Medline: 32551762]

Abbreviations

AAPI: Asian Americans and Pacific Islander
CARE: The Collaborative Approach for Asian Americans and Pacific Islanders Research and Education
COMPASS: COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study
ICAN: International Children Assistance Network
NAPCA: National Asian Pacific Center on Aging
NIA: National Institute on Aging
NIH: National Institutes of Health
PHQ-4: Patient Health Questionnaire-4
SIP: shelter-in-place
VIF: variance inflation factor

Edited by T Sanchez, A Mavragani; submitted 16.12.21; peer-reviewed by N Charness, J Perales-Puchalt; comments to author 30.04.22; revised version received 19.08.22; accepted 28.10.22; published 23.01.23.

Please cite as:
Park LG, Meyer OL, Dougan MM, Golden B, Ta K, Nam B, Tsoh JY, Tzuang M, Park VMT
Social Support and Technology Use and Their Association With Mental and Physical Health During the COVID-19 Pandemic Among Asian Americans: The COMPASS Cross-sectional Study
JMIR Public Health Surveill 2023;9:e35748
URL: https://publichealth.jmir.org/2023/1/e35748
doi:10.2196/35748
PMID:36395324

Original Paper

The Association of Acute Signs and Symptoms of COVID-19 and Exacerbation of Depression and Anxiety in Patients With Clinically Mild COVID-19: Retrospective Observational Study

Sumi Sung[1,2], PhD; Su Hwan Kim[2], PhD; Changwoo Lee[1], BSc; Youlim Kim[1], BSc; Ye Seul Bae[1,3*], MD, MPH; Eui Kyu Chie[1,4,5*], MD, PhD

[1]Office of Hospital Information, Seoul National University Hospital, Seoul, Republic of Korea
[2]Biomedical Research Institute, Seoul National University Hospital, Seoul, Republic of Korea
[3]Department of Family Medicine, Seoul National University Hospital, Seoul, Republic of Korea
[4]Department of Radiation Oncology, College of Medicine, Seoul National University, Seoul, Republic of Korea
[5]Institute of Radiation Medicine, Medical Research Center, Seoul National University, Seoul, Republic of Korea
*these authors contributed equally

Corresponding Author:
Eui Kyu Chie, MD, PhD
Office of Hospital Information
Seoul National University Hospital
101 Daehak-ro
Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 2 2072 7600
Fax: 82 2 2072 0318
Email: ekchie93@snu.ac.kr

Abstract

Background: To date, the association between acute signs and symptoms of COVID-19 and the exacerbation of depression and anxiety in patients with clinically mild COVID-19 has not been evaluated.

Objective: This study was designed to assess the correlation between acute signs and symptoms of COVID-19 and the exacerbation of depression and anxiety in patients with clinically mild COVID-19 at a residential treatment center in South Korea.

Methods: This retrospective study assessed 2671 patients with COVID-19 admitted to 4 residential treatment centers operated by Seoul National University Hospital, South Korea, from March 2020 to April 2022. Depression and anxiety were assessed using the 2-item Patient Health Questionnaire (PHQ-2) and 2-item Generalized Anxiety Disorder (GAD-2) scale, respectively. The exacerbation of depression and anxiety symptoms was identified from the differences in PHQ-2 and GAD-2 scores between admission and discharge, respectively. The patients' clinical characteristics, including acute signs and symptoms of COVID-19, GAD-2 and PHQ-2 scores, were obtained from electronic health records. Demographic characteristics, a summary of vital signs, and COVID-19 symptoms were analyzed and compared between the patient groups with and those without exacerbated PHQ-2 and GAD-2 scores using the chi-square test. We applied logistic regression to identify the association between acute signs and symptoms of COVID-19 and the exacerbation of depression and anxiety.

Results: Sleep disorders were associated with exacerbated depression (odds ratio [OR] 1.09, 95% CI 1.05-1.13) and anxiety (OR 1.1, 95% CI 1.06-1.14), and the sore throat symptom was associated with exacerbated anxiety symptoms (OR 1.03, 95% CI 1.00-1.07). Patients with abnormal oxygen saturation during quarantine were more likely to have exacerbated depression (OR 1.27, 95% CI 1.00-1.62), and those with an abnormal body temperature during quarantine were more likely to experience anxiety (OR 1.08, 95% CI 1.01-1.16). As anticipated, patients who experienced psychological symptoms at admission were more likely to experience depression (OR 1.91, 95% CI 1.52-2.41) and anxiety (OR 1.98, 95% CI 1.54-2.53). Meanwhile, the PHQ-2 and GAD-2 scores measured at admission revealed that lower the score, higher the possibility of exacerbation of both depression (OR 0.15, 95% CI 0.11-0.22) and anxiety (OR 0.13, 95% CI 0.10-0.19).

Conclusions: Results from this study suggest the importance of further interventions for patients with abnormal oxygen saturation, abnormal body temperatures, sore throat, and sleep disorder symptoms or initial psychological symptoms to mitigate the exacerbation of depression and anxiety. In addition, this study highlights the usability of short and efficient scales such as the PHQ-2 and GAD-2 in the assessment of the mental health of patients with clinically mild COVID-19 symptoms who were quarantined at home during the pandemic era.

(JMIR Public Health Surveill 2023;9:e43003) doi:10.2196/43003

KEYWORDS

COVID-19; depression; anxiety; vital signs; symptoms; electronic health records

Introduction

The ongoing COVID-19 pandemic has not only affected physical health but also induced mental health issues such as depression, anxiety, stress, or insomnia [1-3]. A systematic review on the prevalence of psychological symptoms in the general population during the COVID-19 pandemic reported that the prevalence of depressive symptoms ranged from 14.6% to 48.3% and that of anxiety ranged from 6.33% to 50.9%, which is more than double the prevalence in the nonpandemic era [4]. South Korea is no exception because of its increased prevalence of psychological symptoms associated with depression and stress [5-7]. These remarkable increases in psychological symptoms are known to contribute to mental health problems such as anxiety, depression, insomnia, somatization, posttraumatic stress disorder, self-harm, and suicidal thoughts and behaviors. Thus, the importance of early detection and management of mental health disorders cannot be overemphasized [6,8].

Patients confirmed with a COVID-19 diagnosis most commonly experience physical symptoms such as fever, cough, and fatigue [9,10]. In particular, nonhospitalized patients with a clinically mild COVID-19 infection have fatigue, headache, and sneezing as the most common acute symptoms, and fatigue and reduced smell and taste as the most severe symptoms [11]. While transitioning to focusing on the post–COVID-19 condition, symptoms such as fatigue may persist as long as 12 weeks in the long-term phase, and these symptoms are known to affect patients' mental health and the quality of life [11,12].

Despite attempts to explore acute and persistent symptoms of COVID-19 and patients' mental health affected by persistent symptoms, little is known about the association between physical and psychological symptoms of acute COVID-19. Ismael et al [13] found that an increased number of COVID-19 symptoms may be associated with depression, anxiety, and posttraumatic symptoms. Perlis et al [14] reported that patients with COVID-19 who reported headaches during acute infection appeared to have an elevated risk of depression. During the transition to post–COVID-19 condition [15], it is necessary to examine the occurrence of acute signs and symptoms, changes in mental health, and their association while community-dwelling confirmed patients were quarantined.

The Korean government has operated a residential treatment center (RTC) in response to the COVID-19 pandemic [16]. In the early stages of the COVID-19 outbreak, patients with COVID-19 could not be admitted to the hospital because of a lack of negative-pressure isolation beds. Accordingly, the government admitted all confirmed patients with clinically mild conditions to RTCs and transferred them to the hospital if their medical condition deteriorated to require admission-based care. Since then, with the rapid increase in the number of patients with COVID-19, high-risk patient groups with mild conditions have been selectively isolated to RTCs because it was difficult to isolate all confirmed patients at RTCs.

As patients with confirmed COVID-19 without symptoms and those with mild symptoms were being monitored in quarantine at home, it was not easy to explore the status of their acute signs and symptoms and psychological diseases such as depression and anxiety. Therefore, it is important to ascertain the association between COVID-19–related acute signs and symptoms, and depression and anxiety, based on the accumulated retrospective data. With this background, we analyzed the association between acute signs and symptoms of COVID-19 and exacerbation of depression and anxiety in nonhospitalized patients with clinically mild COVID-19 symptoms who were admitted to RTCs.

Methods

Study Design

This retrospective study was performed at Seoul National University Hospital (SNUH).

Study Setting

The study setting was 4 RTCs operated by the SNUH in accordance with government guidelines [17]. The government RTC operating guidelines were periodically updated in consideration of the number of newly confirmed cases at the time the RTC operation and distribution of confirmed cases at the national level. The confirmed patients were assigned to each RTC by the government in accordance with this guideline. During the early phase of the pandemic, from December 2020 to November 2021, following the relevant Korean guidelines, all clinically mild cases among the confirmed cases in South Korea were quarantined at the RTC. With a rapid increase in the number of confirmed cases, all quarantine policies were modified to prioritize those who are more vulnerable to deterioration of their COVID-19 infection. Selection criteria for RTC admission include unvaccinated older individuals and patients in residential environments vulnerable to the spread of infection. All patients were quarantined independently in a separate room. All costs related to quarantine at an RTC were fully covered by the government. Quarantined patients requiring oxygen supply due to a persistent saturation of percutaneous oxygen (SpO_2) of <94%, or those with a clinical condition requiring medical intervention, were transferred to the hospital.

Table 1 shows the location, size in terms of the number of beds, operation period, and accumulated confirmed COVID-19 cases at the national level for operated RTCs in detail.

Table 1. Details of residential treatment centers operated by the Seoul National University Hospital.

Name	Location	Beds, n	Operation period (accumulated confirmed COVID-19 cases at the national level during the operational period, n) [17]
A	Near Daegu city in North Gyeongsang province	99	- March 2020 to April 2020 (7834)
B	Nowon-gu, Seoul	124	- August 2020 to October 2020 (12,206)
C	Seongnam city in Gyeonggi province	225	- August 2020 to September 2020 (9506)
D	Seongnam city in Gyeonggi province	334	- December 2020 to January 2021 (43,998) - July 2021 to April 2022 (17,079,907)

Study Participants

Adult patients with a confirmed COVID-19 diagnosis admitted to 4 RTCs operated by the SNUH from March 2020 to April 2022 were included in this study. Patients who tested positive for COVID-19 on the reverse transcriptase–polymerase chain reaction test but had clinically mild conditions were subjected to RTC admission. A total of 5199 patients were admitted during the study period. Patients with missing data were excluded from the analysis.

Data Sources

Using the SNUH information and communication technology–based remote patient monitoring (RPM) system [18], it was mandatory for every patient admitted to the RTC operated by the SNUH to use the mobile app and to self-report their past medical history, vital signs, subjective acute COVID-19 symptoms, and psychological symptoms. Reporting was done once on admission, twice a day during the quarantine period, and once on discharge. These data were automatically stored in the hospital information system at the SNUH. The records were monitored in real time by health care providers. For any missing data in the records, health care providers directly checked the patient and filled in the missing data. We finally extracted these electronic health records using the SNUH Patient Research Environment (also known as “SUPREME”)—a clinical data warehouse at the SNUH. The data sources in the final analysis included RTC admission notes, daily progress notes, and discharge notes in the electronic health record.

Data on patients’ depressive and anxiety symptoms were extracted from the RTC admission and discharge notes. Patients’ self-reported scores on the 2-item Patient Health Questionnaire (PHQ-2)—a subscale of the Patient Health Questionnaire 9 [19]—and 2-item Generalized Anxiety Disorder Scale (GAD-2)—a subscale of the Generalized Anxiety Disorder Scale (GAD-7) [20]—via a mobile app at admission and discharge. One of the PHQ-2 items was, “Over the last 2 weeks, how often have you been bothered by any of the following problems?” with response options of “Little interest or pleasure in doing things,” and “Feeling down, depressed, or hopeless” on a 4-point Likert scale. Scores ranging from 0 to 6, with a cutoff score of ≥3 indicating depressive symptoms. The PHQ-2 was reported to have a sensitivity of 0.79 and a specificity of 0.86 for any depressive disorders [21]. One of the GAD-2 items was, “Over the last 2 weeks, how often have you been bothered by any of the following problems?” with response options of “Feeling nervous, anxious, or on the edge” and “not being able to stop or control worrying” with a 4-point Likert scale. Scores range from 0 to 6, with a cutoff score of ≥3 indicating anxiety symptoms. The GAD-2 was reported to have a sensitivity of 0.65 and a specificity of 0.88 as a general screening tool for anxiety disorders [22].

Patient demographic characteristics (eg, age, sex, and admitted RTC site), medical history (diabetes mellitus, hypertension, cardiovascular disease, respiratory disease, sleep disorder, and psychological treatment), and initial psychological symptoms such as depression and panic were extracted from the RTCs’ admission notes. Patients’ smoking statuses were extracted from the RTCs’ discharge notes.

Patients’ self-measured vital signs and acute COVID-19 symptoms were extracted from the RTCs’ daily progress notes, which were collected via a mobile app at least twice a day. The vital signs included systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, body temperature (BT), and SpO_2. Acute COVID-19 symptoms consisted of the presence or absence of respiratory and nonrespiratory symptoms. Respiratory symptoms included cough, sputum production, fever, rhinorrhea, sore throat, dyspnea, and chest pain. Nonrespiratory symptoms include nausea, vomiting, abdominal discomfort, pain, constipation, diarrhea, abdominal pain, and sleep disorders.

The loss of taste and smell was extracted from the RTCs’ discharge notes. At discharge, participants reported the presence or loss of smell and taste.

Statistical Analysis

Data preparation was conducted before data analysis. The exacerbation of depressive symptoms’ and anxiety symptoms’ cases were identified from the differences between PHQ-2 scores and GAD-2 scores at admission and discharge, respectively. Criteria for abnormal vital signs, in consultation with an infection medicine specialist, are as follows: a systolic blood pressure above 140 mm Hg or below 60 mm Hg, a diastolic blood pressure above 90 mm Hg or below 30 mm Hg, a heart rate above 110 bpm or below 40 bpm, respiratory rate above 21/minute or below 8/minute, BT above 38.0 °C or below 35.0 °C, and an SpO_2 below 94%. The number of days with each abnormal vital sign and acute COVID-19 symptom was calculated.

Demographic characteristics, summary of vital signs, and COVID-19 symptoms were analyzed and compared between the patient groups with and those without exacerbated PHQ-2 and GAD-2 scores, using the chi-square test. We then applied logistic regression with exacerbation of depressive and anxiety symptoms as dependent variables, and sociodemographic features, past medical history, initial psychological symptoms, initial PHQ-2 or GAD-2 scores, average number of days with abnormal vital signs, and average number of days with acute COVID-19 symptoms as independent variables, using R statistical software (version 4.2; R Project for Statistical Computing). Statistical significance was tested at α=.01.

Ethical Considerations

This study was approved by the SNUH's institutional review board (H-2105-158-1221) and was conducted in accordance with the relevant guidelines and regulations. Informed consent was waived by the institutional review board considering the study design and adherence to the relevant guidelines. In particular, study data were deidentified to protect privacy and preserve the confidentiality of the study participants.

Results

Patients' Clinical Characteristics

In total, 2671 individuals reported COVID-19 infection without any missing values. The clinical characteristics of the patients are presented in Table 2. The mean age was 41.1 (SD 14.1) years, and the mean quarantine period was 7.07 days. Overall, 1242 (46.5%) participants were female, and 44 (1.6%) were admitted at RTC A, 324 (11.8%) at RTC B, 819 (30.7%) at RTC C, and 1494 (55.9%) at RTC D. A total of 746 (27.9%) patients were infected by the alpha variant of SARS-CoV-2, 1563 (58.5%) by the delta variant, and 362 (13.6%) by the omicron variant. Overall, 141 (5.3%) participants had a past medical history of diabetes mellitus, 361 (13.5%) had a medical history of hypertension, 168 (6.3%) had a medical history of sleep disorders, 158 (5.9%) had a history of psychological treatment. Upon admission, 1273 (47.7%) participants had initial psychological symptoms. Detailed clinical characteristics upon exacerbation of PHQ-2 and GAD-2 scores are presented in Multimedia Appendix 1.

Table 2. Patients' clinical characteristics (N=2671).

Clinical characteristics	Participants, n (%)
Demographic data	
Females	1242 (46.5)
Aged under 30 years	748 (28.0)
Aged 30 to 39 years	500 (18.7)
Ages 40 to 49 years	545 (20.4)
Aged 50 to 59 years	556 (20.8)
Aged 60 years or older	322 (12.1)
Admitted to residential treatment center A	44 (1.6)
Admitted to residential treatment center B	314 (11.8)
Admitted to residential treatment center C	819 (30.7)
Admitted to residential treatment center D	1494 (55.9)
Infected with the SARS-CoV-2 alpha variant	746 (27.9)
Infected with the SARS-CoV-2 delta variant	1563 (58.5)
Infected with the SARS-CoV-2 omicron variant	362 (13.6)
Medical history	
Diabetes mellitus	141 (5.3)
Hypertension	361 (13.5)
Cardiovascular disease	81 (3.0)
Respiratory disease	44 (1.6)
Sleep disorder	168 (6.3)
Psychological treatment	158 (5.9)
Initial psychological symptoms	1273 (47.7)
Initial score on the 2-item Patient Health Questionnaire	
0	1641 (61.4)
1	367 (13.7)
2 or above	663 (24.8)
Initial score on the 2-item Generalized Anxiety Disorder Scale	
0	1461 (54.7)
1	354 (13.3)
2 or above	856 (32.0)
Smoking status	
Never smoked	1600 (59.9)
Ex-smoker	565 (21.2)
Smoker	506 (18.9)
BMI category	
Underweight or normal	1182 (44.3)
Overweight	580 (21.7)
Obese	909 (34.0)
Abnormal vital signs	
Systolic blood pressure	564 (21.1)
Diastolic blood pressure	1073 (40.2)
Heart rate	86 (3.2)

Clinical characteristics	Participants, n (%)
Respiratory rate	730 (27.3)
Body temperature	375 (14.0)
Saturation of percutaneous oxygen	72 (2.7)
Acute respiratory COVID-19–related symptoms	
Cough	1759 (65.9)
Sputum	1543 (57.8)
Fever	544 (20.4)
Rhinorrhea	1276 (47.8)
Sore throat	1153 (43.2)
Dyspnea	250 (9.4)
Chest pain	421 (15.8)
Acute nonrespiratory COVID-19–related symptoms	
Nausea	207 (7.7)
Vomit	83 (3.1)
Abdominal discomfort	453 (17.0)
Pain	1047 (39.2)
Constipation	426 (15.9)
Diarrhea	490 (18.3)
Abdominal pain	186 (7.0)
Sleep disorder	624 (23.4)
Loss of smell	1264 (47.3)
Loss of taste	1243 (46.5)

PHQ-2 and GAD-2 Scores Measured at Admission and Discharge

The PHQ-2 and GAD-2 scores measured at admission and discharge are presented in Figure 1. At admission, the mean PHQ-2 score was 0.76 (SD 1.17). Of a total of 2671 patients, 663 (24.8%) had a PHQ-2 score of 2 or higher. The mean GAD-2 score was 0.93 (SD 1.27), and 856 (32.0%) had a score of 2 or higher. At discharge, the mean PHQ-2 score was 0.77 (SD 1.18), and 675 (25.3%) had a score of 2 or higher. The mean GAD-2 score was 0.84 (SD 1.21), and 744 (27.9%) had a score of 2 or higher.

A total of 535 patients showed exacerbation of depression. In fully adjusted models, an abnormal SpO_2 was associated with increased odds of exacerbation of depression (odds ratio [OR] 1.27, 95% CI 1.00-1.62). Among acute COVID-19 symptoms, presence of sleep disorder was associated with exacerbation of depression (OR 1.09, 95% CI 1.05-1.13). Women demonstrated higher odds of exacerbation of depression than men (OR 2.33, 95% CI 1.78-3.05) and higher odds of exacerbation of depression with increasing age (>60 years; OR 4.01, 95% CI 2.71-5.93). Patients admitted at RTC A demonstrated higher odds of exacerbation of depression than those who were admitted at RTCs B, C, and D (OR 2.93, 95% CI 1.24-6.93), and those diagnosed with COVID-19 due to the alpha variant of SARS-CoV-2 showed lower odds of symptom exacerbation than those with the delta variant (OR 0.57, 95% CI 0.36-0.90). Past medical history of sleep disorders demonstrated lower odds to associate with exacerbation (OR 0.60, 95% CI 0.36-0.98), whereas presence of initial psychological symptoms was associated with exacerbation (OR 1.91, 95% CI 1.52-2.41). The higher patients' PHQ-2 scores at admission demonstrated lower odds of exacerbation of depression (OR 0.15, 95% CI 0.11-0.22). Figure 2 illustrates the ORs from the regression models adjusted for the patients' clinical characteristics.

Figure 1. Scores on the 2-item Patient Health Questionnaire (PHQ-2) and the 2-item Generalized Anxiety Disorder Scale (GAD-2) measured at admission and discharge.

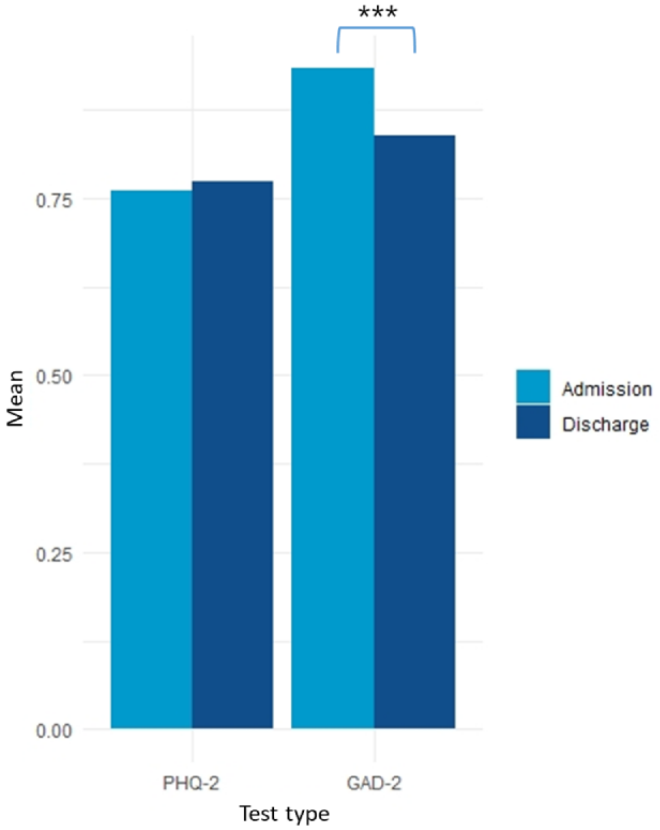
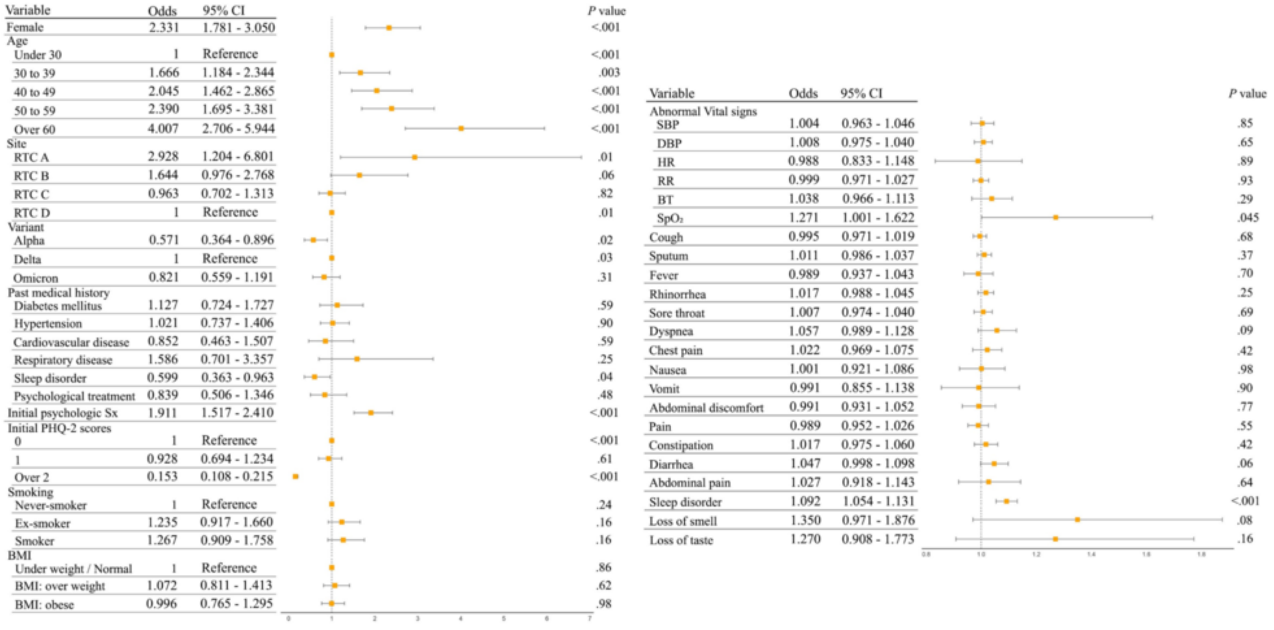


Figure 2. Forest plot showing the association of acute signs and symptoms of COVID-19 and exacerbation of depression. BT: body temperature; DBP: diastolic blood pressure; HR: heart rate; PHQ-2: 2-item Patient Health Questionnaire; RR: respiratory rate; RTC: residential treatment center; SBP: systolic blood pressure; SpO_2: saturation of percutaneous oxygen.



Variable	Odds	95% CI	P value
Female	2.331	1.781 - 3.050	<.001
Age			
Under 30	1	Reference	<.001
30 to 39	1.666	1.184 - 2.344	.003
40 to 49	2.045	1.462 - 2.865	<.001
50 to 59	2.390	1.695 - 3.381	<.001
Over 60	4.007	2.706 - 5.944	<.001
Site			
RTC A	2.928	1.204 - 6.801	.01
RTC B	1.644	0.976 - 2.768	.06
RTC C	0.963	0.702 - 1.313	.82
RTC D	1	Reference	.01
Variant			
Alpha	0.571	0.364 - 0.896	.02
Delta	1	Reference	.03
Omicron	0.821	0.559 - 1.191	.31
Past medical history			
Diabetes mellitus	1.127	0.724 - 1.727	.59
Hypertension	1.021	0.737 - 1.406	.90
Cardiovascular disease	0.852	0.463 - 1.507	.59
Respiratory disease	1.586	0.701 - 3.357	.25
Sleep disorder	0.599	0.363 - 0.963	.04
Psychological treatment	0.839	0.506 - 1.346	.48
Initial psychologic Sx	1.911	1.517 - 2.410	<.001
Initial PHQ-2 scores			
0	1	Reference	<.001
1	0.928	0.694 - 1.234	.61
Over 2	0.153	0.108 - 0.215	<.001
Smoking			
Never-smoker	1	Reference	.24
Ex-smoker	1.235	0.917 - 1.660	.16
Smoker	1.267	0.909 - 1.758	.16
BMI			
Under weight / Normal	1	Reference	.86
BMI: over weight	1.072	0.811 - 1.413	.62
BMI: obese	0.996	0.765 - 1.295	.98

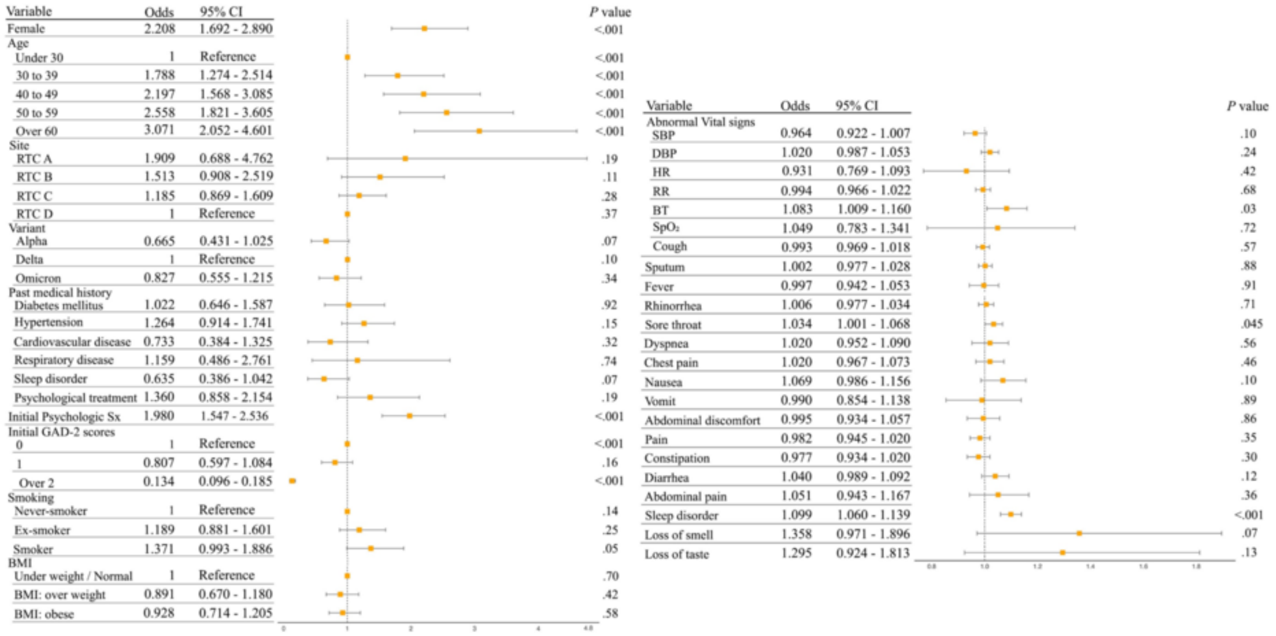
Variable	Odds	95% CI	P value
Abnormal Vital signs			
SBP	1.004	0.963 - 1.046	.85
DBP	1.008	0.975 - 1.040	.65
HR	0.988	0.833 - 1.148	.89
RR	0.999	0.971 - 1.027	.93
BT	1.038	0.966 - 1.113	.29
SpO_2	1.271	1.001 - 1.622	.045
Cough	0.995	0.971 - 1.019	.68
Sputum	1.011	0.986 - 1.037	.37
Fever	0.989	0.937 - 1.043	.70
Rhinorrhea	1.017	0.988 - 1.045	.25
Sore throat	1.007	0.974 - 1.040	.69
Dyspnea	1.057	0.989 - 1.128	.09
Chest pain	1.022	0.969 - 1.075	.42
Nausea	1.001	0.921 - 1.086	.98
Vomit	0.991	0.855 - 1.138	.90
Abdominal discomfort	0.991	0.931 - 1.052	.77
Pain	0.989	0.952 - 1.026	.55
Constipation	1.017	0.975 - 1.060	.42
Diarrhea	1.047	0.998 - 1.098	.06
Abdominal pain	1.027	0.918 - 1.143	.64
Sleep disorder	1.092	1.054 - 1.131	<.001
Loss of smell	1.350	0.971 - 1.876	.08
Loss of taste	1.270	0.908 - 1.773	.16

Association of Acute Signs and Symptoms of COVID-19 and Exacerbation of Anxiety

A total of 523 patients presented exacerbation of anxiety. Among abnormal vital signs, abnormal BT was associated with exacerbation of anxiety (OR 1.08, 95% CI 1.01-1.16). Among acute COVID-19 symptoms, the presence of sleep disorders and sore throat were associated with exacerbation of anxiety (sleep disorders: OR 1.1, 95% CI 1.06-1.14; sore throat: OR 1.03, 95% CI 1.00-1.07). Women showed higher odds of exacerbation of anxiety than men (OR 2.21, 95% CI 1.69-2.88) and higher odds of exacerbation of anxiety with increased age

(>60 years; OR 3.07, 95% CI 2.05-4.60). Presence of initial psychological symptoms was associated with the exacerbation of anxiety (OR 1.98, 95% CI 1.54-2.53). The higher patients' GAD-2 scores at admission indicated lower odds of exacerbation of anxiety (OR 0.13, 95% CI 0.10-0.19). Figure 3 illustrates the ORs from the regression models adjusted for patients' clinical characteristics.

Figure 3. Association of acute signs and symptoms of COVID-19 and exacerbation of anxiety. BT: body temperature; DBP: diastolic blood pressure; HR: heart rate; PHQ-2: 2-item Patient Health Questionnaire; RR: respiratory rate; RTC: residential treatment center; SBP: systolic blood pressure; SpO_2: saturation of percutaneous oxygen.



Discussion

Principal Findings

To the best of our knowledge, this study is the first to explore the association between acute signs and symptoms of COVID-19 and the exacerbation of depression and anxiety in patients with clinically mild COVID-19 [13-15]. Among sociodemographic features, we found that female sex and older age affected the likelihood of exacerbation of depression and anxiety in patients with clinically mild COVID-19, which is consistent with previous studies [23-26]. In addition, RTC A, located in Daegu and Gyeongbuk province, where patients in the early phase of COVID-19 with social stigma were admitted, was found to be highly associated with exacerbation of anxiety.

The most common acute COVID-19 symptoms are cough, sputum, rhinorrhea, sore throat, and loss of smell and taste. The proportion of these symptoms was ~40%-65%, which was relatively lower than that reported in previous studies on acute and persistent symptoms in nonhospitalized patients with COVID-19 [11]. This may reflect the results of triage, which incorporates multiple symptoms in accordance with the hospitalization policy of the Korea Disease Control and Prevention Agency. The types of acute COVID-19 symptoms most commonly complained of were consistent with those reported in previous studies in that fever, cough, myalgia, and fatigue were described as common symptoms of COVID-19 [9,10] and loss of smell and taste as prevalent and relatively discriminative symptoms of COVID-19 [27-29]. Sputum production has been regarded as a less common symptom of early COVID-19 [30,31] but was reported as one of the most common symptoms in this study. The reason for the high proportion of sputum production seems to be that study participants included patients infected with the omicron variant of SARS-CoV-2 whose complains of sputum symptoms were relatively high [32].

Among acute COVID-19 symptoms, patients who experienced sleep disorder symptoms more frequently during quarantine showed higher odds of exacerbation of depression (*P*<.001) and anxiety (*P*<.001) than those who did not. Sleep disorders are a core symptom of depression and anxiety, and this study shows a strong association between them, which is consistent with those reported in previous studies [33-35]. We also found that those who had experienced sore throat more frequently during quarantine had an elevated risk of anxiety symptoms (*P*<.05). The more the patient experiences sore throat, which is one of the representative upper respiratory symptoms of COVID-19 infection with the omicron variant, the higher the risk of exacerbation of anxiety, which means that early active intervention for mental health may be necessary. We expected that other similarly frequently reported acute symptoms such as cough, sputum production, rhinorrhea, and loss of smell and taste would also be associated with exacerbation of depression and anxiety, but this correlation was not found. Although it is difficult to generalize the results of this study, the association between physical symptoms and depression and anxiety is well known, and it is also well known that depression and anxiety accompany somatic symptoms. Thus, in the acute phase within 1 week of COVID-19 infection, health care providers should examine both physical and mental health during the treatment course of patients with acute COVID-19.

On the contrary, the vital signs of COVID-19 were stable, with a small proportion of patients presenting abnormal ranges during quarantine at the RTC. These results indicate that most patients admitted to the RTC had clinically mild conditions and did not require inpatient treatment. Patients who presented abnormal SpO_2 during quarantine demonstrated higher odds of exacerbation of depression ($P<.05$) than those who did not. In addition, patients who had an abnormal BT during quarantine demonstrated higher odds of exacerbation of anxiety ($P<.05$). It is important to note that patients who often had unstable vital signs during quarantine showed the possibility of exacerbation to depression and anxiety due to exposure.

As anticipated, patients who had experienced psychological symptoms at admission had higher odds of exacerbation of depression ($P<.001$) and anxiety ($P<.001$) than those who did not. Meanwhile, the PHQ-2 and GAD-2 scores measured at admission showed that lower the score, higher the association between exacerbation of both depression ($P<.001$) and anxiety ($P<.001$). In addition, patients who had received previous psychological treatment were less likely to experience exacerbation of depression ($P<.001$). The group that received psychological treatment might have had a higher PHQ-2 or GAD-2 score at admission. Thus, these results can be interpreted as follows: higher the PHQ-2 or GAD-2 score at admission, narrower the range of scores that can be exacerbated.

The novelty of this study is the use of both contactless RPM systems and screening scales. First, data were obtained through a contactless RPM system for patients with COVID-19. More than 80% of patients with COVID-19 had clinically mild conditions, and these patients were quarantined at home or RTCs, not medical institutions, and monitored with a contactless RPM system. The establishment of a contactless RPM system for patients with clinically mild conditions who do not require intensive care unit care has been gradually attempted in various studies [36-38], and the SNUH has also established a contactless RPM system [18]. The data analyzed in this study were also the results of establishing the system at 4 RTCs operated by the SNUH.

Second, the results of this study provide supporting evidence that depression and anxiety can be assessed by the PHQ-2 and GAD-2, which are useful and efficient scales for patients with COVID-19 quarantined at home. As the prevalence of mood disorders has increased after the onset of the COVID-19 pandemic [4-7], concerns about individuals' mental health are growing. Patients with COVID-19 who are self-isolated at home or at an RTC might be in a clinically mild condition and are not fully aware of the need for mental health assessment. Since they are not hospitalized and have limited access to medical personnel, it is necessary to assess the mental health of patients using a readily accessible simple convenient tool. In that sense, the PHQ-2 and GAD-2, which have already been validated as having high sensitivity and specificity as the PHQ-9 and GAD-7, may be regarded as useful screeners and efficient tools for monitoring treatment progress and outcomes in primary medical settings [21,39-43]. Furthermore, the results from this study suggest the importance of considering further interventions that might mitigate the exacerbation of depression and anxiety associated with acute signs and symptoms during the post–COVID-19 condition [15].

Limitations

This study is not without its limitations. First, this was a retrospective observational study; therefore, causal relationships cannot be inferred, and further analysis of additional data such as socioeconomic factors, not collected in the hospital information system at the SNUH, was not possible. Second, the data analyzed were retrieved from a single tertiary hospital operating 4 RTCs of different sizes in accordance with government regulations. The possibility of influence or correlation with various unmeasured variables such as, but not limited to, lack of social interactions, financial loss, and vaccination status may limit the generalizability of our findings. Third, sample bias was possible as patients with missing data were excluded. Fourth, there were no follow-up longitudinal data after discharge from the RTC. Thus, it was not possible to explore the associations between COVID-19–related acute signs and symptoms and the exacerbation of depression and anxiety, considering the acute and chronic periods in response to post–COVID-19 condition. Fifth, there may be a social desirability bias as patients self-reported their signs and symptoms and were not objectified by the caregivers. Lastly, while the mean quarantine period of participants was approximately 7 days, the PHQ-2 and GAD-2 measures depression and anxiety symptoms over the last 2 weeks. These timeline issues should be considered when interpreting the results. Further, other tools may be used in future studies.

Conclusions

This study explored the association between acute signs and symptoms of COVID-19 and the exacerbation of depression and anxiety in patients with clinically mild COVID-19 who were admitted at RTCs. Among the acute signs and symptoms of COVID-19, exacerbation of depressive symptoms was associated with an abnormal SpO_2 (OR 1.27, 95% CI 1.00-1.62) and sleep disorder (OR 0.60, 95% CI 0.36-0.98). Exacerbation of anxiety symptoms was associated with an abnormal BT (OR 1.08, 95% CI 1.01-1.16), sleep disorder (OR 1.1, 95% CI 1.06-1.14), and sore throat (OR 1.03, 95% CI 1.00-1.07). As anticipated, patients who had experienced initial psychological symptoms at admission were more likely to experience depression (OR 1.91, 95% CI 1.52-2.41) and anxiety (OR 1.08, 95% CI 1.01-1.16) than those who did not. Meanwhile, the PHQ-2 and GAD-2 scores measured at admission showed that lower the score, higher the possibility of exacerbation of both depression (OR 0.15, 95% CI 0.11-0.22) and anxiety (OR 0.13, 95% CI 0.10-0.19). Our results suggest the importance of considering further interventions that might mitigate the exacerbation of depression and anxiety in patients with an abnormal SpO_2, abnormal BT, sore throat, and sleep disorders or those with initial psychological symptoms. In addition, we propose the use of the PHQ-2 and GAD-2, which are short and efficient scales, in the assessment of mental health for patients with clinically mild COVID-19 who are quarantined at home.

Acknowledgments

This work was supported by an Institute of Information & Communications Technology Planning & Evaluation grant funded by the Korean government (2021-0-00312: Development of Non-Face-to-face Patient Infection Activity Prediction and Protection Management SW Technology at Home and Community Treatment Centers for Effective Response to Infectious Disease).

Data Availability

The data sets generated or analyzed during this study are not publicly available in accordance with the hospital's regulations, adhering to the National Privacy Act and relevant guidelines. However, they may be available from the corresponding author on reasonable request after review and approval of the institutional review board and the institutional data steering committee.

Authors' Contributions

SS, YSB, and EKC conceptualized the study and reviewed and edited the manuscript. SS, SHK, and YSB developed the study methodology and drafted the manuscript. SS, SHK, CL, YK, and YSB carried out the formal analysis. EKC supervised the study and acquired funding for the study. All authors reviewed the results and approved the final version of the manuscript. YSB and EKC contributed equally to this work as the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed clinical characteristics by exacerbation of PHQ-2 and GAD-2 scores.

[DOCX File , 19 KB - publichealth_v9i1e43003_app1.docx]

References

1. Brown E, Gray R, Lo Monaco S, O'Donoghue B, Nelson B, Thompson A, et al. The potential impact of COVID-19 on psychosis: a rapid review of contemporary epidemic and pandemic research. Schizophr Res 2020 Aug;222:79-87 [FREE Full text] [doi: 10.1016/j.schres.2020.05.005] [Medline: 32389615]
2. Huang Y, Zhao N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. Psychiatry Res 2020 Jun;288:112954 [FREE Full text] [doi: 10.1016/j.psychres.2020.112954] [Medline: 32325383]
3. Oh T, Park H, Song I. Risk of psychological sequelae among coronavirus disease ‐ 2019 survivors: a nationwide cohort study in South Korea. Depress Anxiety 2020 Dec 07;38(2):247-254. [doi: 10.1002/da.23124]
4. Xiong J, Lipsitz O, Nasri F, Lui LMW, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord 2020 Dec 01;277:55-64 [FREE Full text] [doi: 10.1016/j.jad.2020.08.001] [Medline: 32799105]
5. Min J, Kwon D, Lee W, Kang C, Park C, Heo S, et al. Individual and regional characteristics associated with changes in mental health before and during the COVID-19 pandemic in South Korea. Sci Rep 2022 Aug 19;12(1):14189 [FREE Full text] [doi: 10.1038/s41598-022-18493-1] [Medline: 35986083]
6. Jeong H, Park S, Kim J, Oh K, Yim H. Mental health of Korean adults before and during the COVID-19 pandemic: a special report of the 2020 Korea National Health and Nutrition Examination Survey. Epidemiol Health 2022;44:e2022042 [FREE Full text] [doi: 10.4178/epih.e2022042] [Medline: 35538698]
7. Kim S, Park I, Kim M, Park A, Jhon M, Kim J, et al. Risk and protective factors of depression in the general population during the COVID-19 epidemic in Korea. BMC Psychiatry 2021 Sep 08;21(1):445 [FREE Full text] [doi: 10.1186/s12888-021-03449-y] [Medline: 34496823]
8. Cénat JM, Felix N, Blais-Rochette C, Rousseau C, Bukaka J, Derivois D, et al. Prevalence of mental health problems in populations affected by the Ebola virus disease: a systematic review and meta-analysis. Psychiatry Res 2020 Jul;289:113033. [doi: 10.1016/j.psychres.2020.113033] [Medline: 32388176]
9. Grant M, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. PLoS One 2020;15(6):e0234765 [FREE Full text] [doi: 10.1371/journal.pone.0234765] [Medline: 32574165]
10. Mizrahi B, Shilo S, Rossman H, Kalkstein N, Marcus K, Barer Y, et al. Longitudinal symptom dynamics of COVID-19 infection. Nat Commun 2020 Dec 04;11(1):6208 [FREE Full text] [doi: 10.1038/s41467-020-20053-y] [Medline: 33277494]

11. Bliddal S, Banasik K, Pedersen OB, Nissen J, Cantwell L, Schwinn M, et al. Acute and persistent symptoms in non-hospitalized PCR-confirmed COVID-19 patients. Sci Rep 2021 Jun 23;11(1):13153 [FREE Full text] [doi: 10.1038/s41598-021-92045-x] [Medline: 34162913]
12. Jacobs LG, Gourna Paleoudis E, Lesky-Di Bari D, Nyirenda T, Friedman T, Gupta A, et al. Persistence of symptoms and quality of life at 35 days after hospitalization for COVID-19 infection. PLoS One 2020 Dec 11;15(12):e0243882 [FREE Full text] [doi: 10.1371/journal.pone.0243882] [Medline: 33306721]
13. Ismael F, Bizario J, Battagin T, Zaramella B, Leal FE, Torales J, et al. Post-infection depressive, anxiety and post-traumatic stress symptoms: a prospective cohort study in patients with mild COVID-19. Prog Neuropsychopharmacol Biol Psychiatry 2021 Dec 20;111:110341 [FREE Full text] [doi: 10.1016/j.pnpbp.2021.110341] [Medline: 33940097]
14. Perlis RH, Ognyanova K, Santillana M, Baum MA, Lazer D, Druckman J, et al. Association of acute symptoms of COVID-19 and symptoms of depression in adults. JAMA Netw Open 2021 Mar 01;4(3):e213223 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.3223] [Medline: 33710287]
15. The Lancet. Facing up to long COVID. Lancet 2020 Dec;396(10266):1861. [doi: 10.1016/s0140-6736(20)32662-3]
16. Kang E, Lee SY, Jung H, Kim MS, Cho B, Kim YS. Operating protocols of a community treatment center for isolation of patients with coronavirus disease, South Korea. Emerg Infect Dis 2020 Oct;26(10):2329-2337 [FREE Full text] [doi: 10.3201/eid2610.201460] [Medline: 32568665]
17. Coronavirus (COVID-19) Republic of Korea, 2022. Korea Disease Control and Prevention Agency. URL: https://ncov.kdca.go.kr/ [accessed 2022-12-22]
18. Bae Y, Kim K, Choi S, Ko T, Jeong CW, Cho B, et al. Information technology-based management of clinically healthy COVID-19 patients: lessons from a living and treatment support center operated by Seoul National University Hospital. J Med Internet Res 2020 Jun 12;22(6):e19938 [FREE Full text] [doi: 10.2196/19938] [Medline: 32490843]
19. Kroenke K, Spitzer R, Williams J. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
20. Spitzer R, Kroenke K, Williams J, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
21. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res 2005 Feb;58(2):163-171. [doi: 10.1016/j.jpsychores.2004.09.006] [Medline: 15820844]
22. Kroenke K, Spitzer R, Williams J, Monahan P, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
23. Mazza C, Ricci E, Biondi S, Colasanti M, Ferracuti S, Napoli C, et al. A nationwide survey of psychological distress among Italian people during the COVID-19 pandemic: immediate psychological responses and associated factors. Int J Environ Res Public Health 2020 May 02;17(9):3165 [FREE Full text] [doi: 10.3390/ijerph17093165] [Medline: 32370116]
24. Moccia L, Janiri D, Pepe M, Dattoli L, Molinaro M, De Martin V, et al. Affective temperament, attachment style, and the psychological impact of the COVID-19 outbreak: an early report on the Italian general population. Brain Behav Immun 2020 Jul;87:75-79 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.048] [Medline: 32325098]
25. Moghanibashi-Mansourieh A. Assessing the anxiety level of Iranian general population during COVID-19 outbreak. Asian J Psychiatr 2020 Jun;51:102076 [FREE Full text] [doi: 10.1016/j.ajp.2020.102076] [Medline: 32334409]
26. Nemani K, Li C, Olfson M, Blessing EM, Razavian N, Chen J, et al. Association of psychiatric disorders with mortality among patients with COVID-19. JAMA Psychiatry 2021 Apr 01;78(4):380-386 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4442] [Medline: 33502436]
27. Yan CH, Faraji F, Prajapati DP, Boone CE, DeConde AS. Association of chemosensory dysfunction and COVID-19 in patients presenting with influenza-like symptoms. Int Forum Allergy Rhinol 2020 Jul;10(7):806-813 [FREE Full text] [doi: 10.1002/alr.22579] [Medline: 32279441]
28. Peyrony O, Marbeuf-Gueye C, Truong V, Giroud M, Rivière C, Khenissi K, et al. Accuracy of emergency department clinical findings for diagnosis of coronavirus disease 2019. Ann Emerg Med 2020 Oct;76(4):405-412 [FREE Full text] [doi: 10.1016/j.annemergmed.2020.05.022] [Medline: 32563600]
29. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul 11;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
30. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020 Mar;395(10229):1054-1062. [doi: 10.1016/s0140-6736(20)30566-3]
31. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA 2020 Mar 17;323(11):1061-1069 [FREE Full text] [doi: 10.1001/jama.2020.1585] [Medline: 32031570]
32. Kim M, Lee B, Choi YY, Um J, Lee K, Sung HK, et al. Clinical characteristics of 40 patients infected with the SARS-CoV-2 omicron variant in Korea. J Korean Med Sci 2022 Jan 17;37(3):e31 [FREE Full text] [doi: 10.3346/jkms.2022.37.e31] [Medline: 35040299]

33. Nutt D, Wilson S, Paterson L. Sleep disorders as core symptoms of depression. Dialogues Clin. Neurosci 2022 Apr 01;10(3):329-336. [doi: 10.31887/dcns.2008.10.3/dnutt]
34. Krystal AD. Psychiatric disorders and sleep. Neurol Clin 2012 Nov;30(4):1389-1413 [FREE Full text] [doi: 10.1016/j.ncl.2012.08.018] [Medline: 23099143]
35. Freeman D, Sheaves B, Waite F, Harvey AG, Harrison PJ. Sleep disturbance and psychiatric disorders. Lancet Psychiat 2020 Jul;7(7):628-637. [doi: 10.1016/s2215-0366(20)30136-x]
36. Coffey JD, Christopherson LA, Glasgow AE, Pearson KK, Brown JK, Gathje SR, et al. Implementation of a multisite, interdisciplinary remote patient monitoring program for ambulatory management of patients with COVID-19. NPJ Digit Med 2021 Aug 13;4(1):123 [FREE Full text] [doi: 10.1038/s41746-021-00490-9] [Medline: 34389787]
37. Annis T, Pleasants S, Hultman G, Lindemann E, Thompson JA, Billecke S, et al. Rapid implementation of a COVID-19 remote patient monitoring program. J Am Med Inform Assoc 2020 Aug 01;27(8):1326-1330 [FREE Full text] [doi: 10.1093/jamia/ocaa097] [Medline: 32392280]
38. Gordon WJ, Henderson D, DeSharone A, Fisher HN, Judge J, Levine DM, et al. Remote patient monitoring program for hospital discharged COVID-19 patients. Appl Clin Inform 2020 Oct 25;11(5):792-801 [FREE Full text] [doi: 10.1055/s-0040-1721039] [Medline: 33241547]
39. Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Ann Fam Med 2010 Jul 19;8(4):348-353 [FREE Full text] [doi: 10.1370/afm.1139] [Medline: 20644190]
40. Manea L, Gilbody S, Hewitt C, North A, Plummer F, Richardson R, et al. Identifying depression with the PHQ-2: a diagnostic meta-analysis. J Affect Disord 2016 Oct;203:382-395. [doi: 10.1016/j.jad.2016.06.003] [Medline: 27371907]
41. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
42. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a primary care setting. Cureus 2020 May 21;12(5):e8224 [FREE Full text] [doi: 10.7759/cureus.8224] [Medline: 32582485]
43. Staples L, Dear B, Gandy M, Fogliati V, Fogliati R, Karin E, et al. Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: the PHQ-2, GAD-2, and K-6. Gen Hosp Psychiatry 2019;56:13-18 [FREE Full text] [doi: 10.1016/j.genhosppsych.2018.11.003] [Medline: 30508772]

Abbreviations

BT: body temperature
GAD-2: 2-item Generalized Anxiety Disorder
OR: odds ratio
PHQ-2: 2-item Patient Health Questionnaire
RPM: remote patient monitoring
RTC: residential treatment center
SNUH: Seoul National University Hospital
SpO_2: saturation of percutaneous oxygen

Edited by T Sanchez, A Mavragani; submitted 28.09.22; peer-reviewed by N Guo, CY Lin; comments to author 12.12.22; revised version received 30.12.22; accepted 12.01.23; published 30.01.23.

Please cite as:
Sung S, Kim SH, Lee C, Kim Y, Bae YS, Chie EK
The Association of Acute Signs and Symptoms of COVID-19 and Exacerbation of Depression and Anxiety in Patients With Clinically Mild COVID-19: Retrospective Observational Study
JMIR Public Health Surveill 2023;9:e43003
URL: https://publichealth.jmir.org/2023/1/e43003
doi:10.2196/43003
PMID:36645439

Original Paper

Using Location Intelligence to Evaluate the COVID-19 Vaccination Campaign in the United States: Spatiotemporal Big Data Analysis

Qingfeng Li[1], PhD; James Cheng Peng[2], MSE; Diwakar Mohan[1], MD, DrPH; Brennan Lake[3], MBA; Alex Ruiz Euler[3], PhD; Brian Weir[4], PhD; Lena Kan[1], MSPH; Cui Yang[4], PhD; Alain Labrique[1], PhD

[1]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[2]Department of Applied Mathematics and Statistics, Johns Hopkins Whiting School of Engineering, Baltimore, MD, United States

[3]Cuebiq – Intelligence in Action, New York, NY, United States

[4]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Qingfeng Li, PhD
Department of International Health
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St
Baltimore, MD, 21205
United States
Phone: 1 4106141419
Email: qli28@jhu.edu

Abstract

Background: Highly effective COVID-19 vaccines are available and free of charge in the United States. With adequate coverage, their use may help return life back to normal and reduce COVID-19–related hospitalization and death. Many barriers to widespread inoculation have prevented herd immunity, including vaccine hesitancy, lack of vaccine knowledge, and misinformation. The Ad Council and COVID Collaborative have been conducting one of the largest nationwide targeted campaigns ("It's Up to You") to communicate vaccine information and encourage timely vaccination across the United States. More than 300 major brands, digital and print media companies, and community-based organizations support the campaigns to reach distinct audiences.

Objective: The goal of this study was to use aggregated mobility data to assess the effectiveness of the campaign on COVID-19 vaccine uptake.

Methods: Campaign exposure data were collected from the Cuebiq advertising impact measurement platform consisting of about 17 million opted-in and deidentified mobile devices across the country. A Bayesian spatiotemporal hierarchical model was developed to assess campaign effectiveness through estimating the association between county-level campaign exposure and vaccination rates reported by the Centers for Disease Control and Prevention. To minimize potential bias in exposure to the campaign, the model included several control variables (eg, age, race or ethnicity, income, and political affiliation). We also incorporated conditional autoregressive residual models to account for apparent spatiotemporal autocorrelation.

Results: The data set covers a panel of 3104 counties from 48 states and the District of Columbia during a period of 22 weeks (March 29 to August 29, 2021). Officially launched in February 2021, the campaign reached about 3% of the anonymous devices on the Cuebiq platform by the end of March, which was the start of the study period. That exposure rate gradually declined to slightly above 1% in August 2021, effectively ending the study period. Results from the Bayesian hierarchical model indicate a statistically significant positive association between campaign exposure and vaccine uptake at the county level. A campaign that reaches everyone would boost the vaccination rate by 2.2% (95% uncertainty interval: 2.0%-2.4%) on a weekly basis, compared to the baseline case of no campaign.

Conclusions: The "It's Up to You" campaign is effective in promoting COVID-19 vaccine uptake, suggesting that a nationwide targeted mass media campaign with multisectoral collaborations could be an impactful health communication strategy to improve progress against this and future pandemics. Methodologically, the results also show that location intelligence and mobile phone–based monitoring platforms can be effective in measuring impact of large-scale digital campaigns in near real time.

(JMIR Public Health Surveill 2023;9:e39166) doi:10.2196/39166

KEYWORDS

COVID-19 vaccine; vaccine campaign; big data; vaccination; vaccine; COVID-19; uptake; effectiveness; barriers; hesitancy; health information

Introduction

The COVID-19 pandemic has impacted every country in the world, and the United States was hit particularly hard. As of February 2022, more than 900,000 people had died in the United States after contracting the virus [1]. Several highly effective vaccines have been developed and authorized for emergency use by Food and Drug Administration [2,3]. In sharp contrast to the early days of the pandemic in the United States, there are fewer structural barriers that could interfere with vaccine coverage. However, attitudinal barriers remain prevalent [4]. According to a national poll conducted by the COVID Collaborative, although nearly 90% of Americans recognize the effectiveness of vaccines, only about a third plan to get vaccinated themselves [5]; and other surveys revealed similar patterns [6,7]. This mistrust is more prevalent among racial and ethnic minority communities. A nationwide poll found that only 14% of Black Americans and 34% of Latinx Americans trust the safety of the vaccines [8]. The massive distrust among racial and ethnic minority communities is the result, according to various analyses, of generational trauma and continued mixed messages about the pandemic and vaccines [9]. Latinx and African American communities have been disproportionately affected by the COVID-19 pandemic [10-12]. This needs to be addressed immediately since enhancing access and uptake of the COVID-19 vaccine among those most at risk is critical to control the epidemic, reduce the emergence of new variants, and promote health equity.

Similar hesitancy has been documented with nearly all vaccines [9,10]. Vaccine hesitancy is a multifactorial phenomenon, and therefore, requires multifaceted strategies. An effective and comprehensive campaign should be part of a national vaccination plan. Many states began dealing with vaccine hesitancy early in vaccine rollout, particularly in selected areas and among certain racial and ethnic groups. Information campaigns were developed to provide science-based information and culturally appropriate resources, either through existing medical networks or communication firms [11]. Some states also built or used partnerships with faith-based leaders and trusted community organizations to reach critical populations and increase vaccine acceptance among minority communities [11]. There are also other forms of media campaigns, primarily focusing on certain groups or areas [12].

At the national level, the Ad Council and the COVID Collaborative launched a nationwide vaccine education campaign called “It’s Up to You” in February 2021 [13]. As one of the largest public health efforts in US history, the campaign aims to educate the American public and build confidence in vaccines by communicating vaccine messages informed by the best science. In coordination with over 300 major brands, digital and print media companies, community-based organizations, medical experts, the Ad Council has been rolling out public service announcements across airwaves, publications, and social media. The public service announcements provide facts and correct information, answer questions, and preempt false narratives through platforms that the target groups regularly visit. To bridge the racial and ethnic gap in vaccine knowledge, the campaign particularly targets racial and ethnic minority communities with contents customized for distinct cultural and historical background. Campaign materials are available in 7 languages (English, Spanish, Simplified Chinese, Korean, Russian, Haitian Creole, and Vietnamese) on the campaign website. Interested readers are referred to Ad Council and the campaign website for more information about the campaign design and delivery.

As of September 19, 2021, the campaign received US $168 million in media support and related publicity and achieved 80 million engagements with messages via social and search [13]. Historically, the use of media campaigns to communicate public health knowledge and promote positive behaviors has shown varying degrees of effectiveness [14-16]. Without a systematic assessment, implementers lose the opportunity to identify potential gaps in campaign design and make strategic changes to message content and dissemination strategy. Therefore, our study aims to evaluate the effectiveness of this national campaign in promoting vaccine uptake, which could be used to inform the design and evaluation of other large-scale behavior change interventions.

Methods

Exposure Data

Monitoring the exposure to mass media campaigns is challenging due to the dynamic nature of the campaign rollout and massive coverage. Traditional monitoring approaches include surveys, interviews, and cohort studies [17,18]. However, the relatively slow turnaround makes the approaches more suitable for a retrospective evaluation but less able to generate timely and actionable feedback.

Emerging geospatial technologies present an opportunity for frequent and rapid data collection. We collect aggregated exposure data through Cuebiq’s privacy-preserving geospatial data and analytics platform. Cuebiq partners with over 100 location-centric smartphone apps, providing a path for users to opt in and provide informed consent for their anonymized and aggregated data to be used for research purposes. In addition to its General Data Protection Regulation and California Consumer Privacy Act–compliant data collection practices, Cuebiq applies additional privacy protections beyond simple anonymization to prevent the reidentification of individual users.

To measure campaign exposure, a digital breadcrumb known as a pixel is attached to creative media assets served through multiple advertisement media, including web, mobile, and in-app browsing. Cuebiq then generates impressions data and matches those impressions with its own panel of users in a privacy-preserving manner. Measurement data are accessed by researchers via the Cuebiq Workbench platform, an auditable sandbox environment that allows access for the querying of data and generation of aggregate, privacy-preserving outputs. The sandbox enables the creation of aggregate data at county levels, without the ability or need to create individual-level outputs.

The Cuebiq platform collects exposure information in a nearly continuous manner for each individual device. The exposure is then summarized by a binary indicator for each device, denoting whether or not it is exposed to the campaign. Finally, for each county, we calculate the number of exposed devices, the number of unexposed devices, and the exposure rate (ie, the proportion of exposed devices among all monitored devices) on a weekly basis. As described above, the "It's Up to You" campaign uses both digital and print media to deliver the materials. Our exposure data only measure the exposure to digital contents.

The final analytical data set covers 3104 counties in 48 states and the District of Columbia. Hawaii and Alaska are excluded because the spatial model requires counties to have geographically adjacent neighbors. Ad Council campaigns and Cuebiq's measurement campaign were launched in late February 2021, but the campaign did not reach national coverage until about one month later. Campaign activities and related exposures wound down rapidly after August. To avoid potential selection bias in early campaign rollout, our analyses only used data between March 29 and August 29, 2021. The panel's width and length are adequate for us to include complex model components, such as the spatiotemporal correlation structure.

Outcome Data

The outcome measure, vaccination rate, comes from Centers for Disease Control and Prevention's COVID Data Tracker, which provides county-level vaccination rates on a daily basis. To remain consistent with the exposure measures described above, weekly averages are taken at the county level.

Other Control Variables

Several socioeconomic and demographic indicators have been shown to influence vaccine uptake. This study explored the following: age (percentage of people 85 years or older); race and ethnicity (percentage of non-Hispanic White population); political affiliation (percentage of people who voted for democrats in the 2020 presidential election); income (median household income). The inclusion of covariates is based on literature review, data availability, and model diagnoses.

Statistical Model

Bayesian hierarchical mixed-effects models are used to estimate the correlation between exposure and outcome measures at the county level. As an important feature of infectious diseases, county-level pandemic measures exhibit spatial autocorrelation [19]. Nearer counties are more similar than distant counties in terms of COVID-19 pandemic and vaccination. We will account for the spatial autocorrelation in the regression. However, the residual may still display spatial autocorrelation, which violates a key regression modeling assumption. Similarly, panel data typically display temporal autocorrelation. Therefore, our model includes a conditional autoregressive (CAR) component to account for the spatiotemporal autocorrection in the data. Essentially, the CAR component supplements the main effects model with a set of spatiotemporally autocorrelated random effects [20].

Our final model can be formally expressed as follows:

Here, where $c=1,2,\ldots,c$ indexes counties; Y_{ct} denotes the vaccination rate (weekly and not cumulative) in week t in county c; X_{ct} denotes regressors in week t in county c. Of note, some variables may be time invariant; β denotes a column vector of regression slopes.

A key model component is ψ_{ct}, which induces the spatiotemporal autocorrelation for county-level measures after accounting for the covariate effects. In particular, we use a first-order autoregressive process where correlation is ρ^w, $\rho\in(0,1)$ for Us that are w time units apart. Of particular interest is the spatial model that allows the U_{ct} to be correlated with correlation depending on a distance metric. As for the first-order autoregressive time-series model, spatial models induce correlation among the U_{ct} for a fixed t.

County coordinates are used to compute Euclidean distances. With relatively few nearby counties, a parameter model will estimate nearby correlations based on correlations between distant counties (an extrapolation). There are several model forms, including (powered) exponential, Gaussian, Spherical, Matern, and CAR models [21]. Due to the large number of contiguous counties in our data set, the CAR model worked well and was included in our analyses.

The model was fitted in R using the CARBayesST package [22]. The final results are extracted from 10,000 samples after discarding 10,000 burn-ins and thinning by 10 to reduce potential sample autocorrelation.

Ethical Considerations

Study data are anonymous and deidentified, and the study did not constitute human subjects. Therefore, ethical approval was not required for this study.

Results

The campaign was launched in February 2021, starting with a small number of counties in the first few weeks. It reached national coverage in week 13 (March 29 to April 4, 2021), when about 99% (3058/3104) of counties in the Cuebiq platform detected exposure to the campaign. Campaign activities started winding down quickly from week 34 (August 23-29, 2021). As a result, this study covered a period of 22 weeks. Cuebiq continued monitoring a panel of about 17 million mobile devices (Figure 1). The panel remained stable during this period, ensuring the temporal comparability of the observed metrics.

Since this is not a randomized controlled trial, it is important to assess the potential selection issues in the data. Selection could have occurred in two stages. First, certain areas and groups may be overrepresented in the Cuebiq platform. This may be caused by different smartphone penetration rates by state and county. Second, the dissemination of Ad Council campaigns may be disproportionally concentrated in certain areas and groups. To assess the severity of selection, we plotted Cuebiq coverage rate and campaign exposure rate against race, income, and political affiliation at the county level (Figure 2). The Cuebiq coverage is higher in counties with a larger percentage of non-Hispanic White population. Campaign exposure rate has a similar correlation with race. A similar correlation is observed for household income. Richer counties tend to be overrepresented in the Cuebiq platform and get more exposure to the campaigns than poorer counties. For political affiliation, more democratic counties appear to be less represented in the Cuebiq platform, and there is not a strong pattern between campaign exposure and political affiliation. In sum, there may be potential selection issues in both Cuebiq platform coverage and campaign exposure. We will minimize the bias through accounting for these factors in the regression model.

The overall proportion of Cuebiq-enrolled devices exposed to the Ad Council campaigns started at 3.17% in the first week of the study period and gradually declined to 1.28% by the end of August 2021. A large number of channels were used to disseminate the campaign messages. Top channels in terms of exposure include The Trade Desk, Q Digital, Buzzfeed, Pandora, and Red Ventures.

During the study period, the national vaccination rate increased from 16.4% to 47.2%, but the progress began slowing down (Figure 3) [23]. Overall, the weekly vaccination rate peaked in week 15 when 3.34% of Americans completed full vaccination that week; it dropped below 1% in later weeks. The box plots in Figure 3 illustrate the substantial heterogeneity in weekly vaccination rates across counties.

Temporal autocorrelation in weekly vaccination rate is evident. Figure 4 illustrates the strong geographic clustering or spatial correlation in weekly vaccination rates. That spatial correlation may not fully be explained by the 3 variables in the model.

All indicators included in the Bayesian hierarchical model vary greatly across the 3104 counties analyzed in the study. The average percentage of the population aged 85 years and older is 2.3% (SD 0.01%) with a minimum of 0.0% and a maximum of 8.0% (Table 1). The ranges of the percentage of non-Hispanic White population and the percentage who voted for democrats are much wider at 5.1%-99.8% and 3.1%-82.1%, respectively. The poorest county earns an average of 36.5 thousand US dollars per household, while the richest county earns 243.8 thousand.

Table 2 presents our main results from the Bayesian hierarchical model with a spatiotemporally correlated residual structure. The coefficient for the exposure rate is 2.2 (95% uncertainty interval: 2.0-2.4). The effectiveness is impressive; exposing everyone in the county to the campaign may boost the vaccination uptake by 2.2% on a weekly basis. This effectiveness size is substantial, given the slow weekly vaccination rate of less than 1% at the end of the study period. It is worth noting that the effectiveness was assessed based on observed data. The exact magnitude of the effectiveness will likely change when the overall vaccination rates grow substantially higher in the future.

All factors in the model also have statistically significant impacts on vaccine uptake. Proportions of older population (≥85 years), non-Hispanic White population, democratic affiliation, and household income are all positively associated with vaccine rollout. This is consistent with findings from previous studies.

Figure 1. Number of opted-in, deidentified devices (in million) monitored by the Cuebiq platform and the percentage exposed to the Ad Council Campaign.

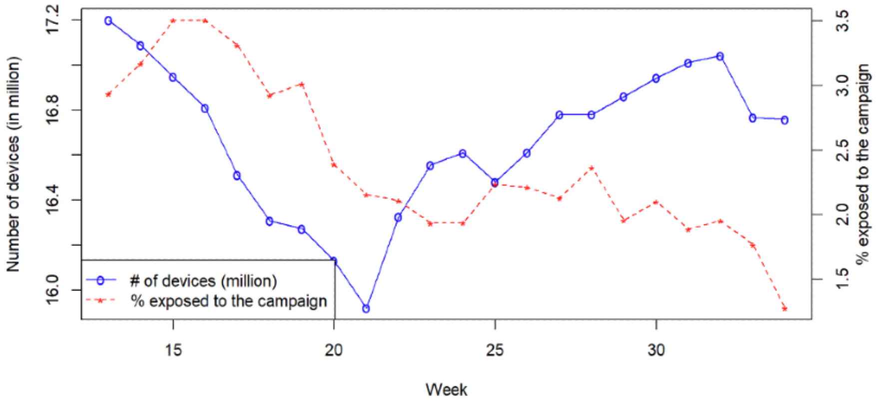


Figure 2. Cuebiq coverage and campaign exposure rate versus race, household income, and political affiliation during the first week of the study period. (A) Cuebiq coverage vs race; (B) Campaign exposure vs race; (C) Cuebiq coverage vs household income; (D) Campaign exposure vs household income; (E) Cuebiq coverage vs political affiliation; (F) Campaign exposure vs political affiliation.

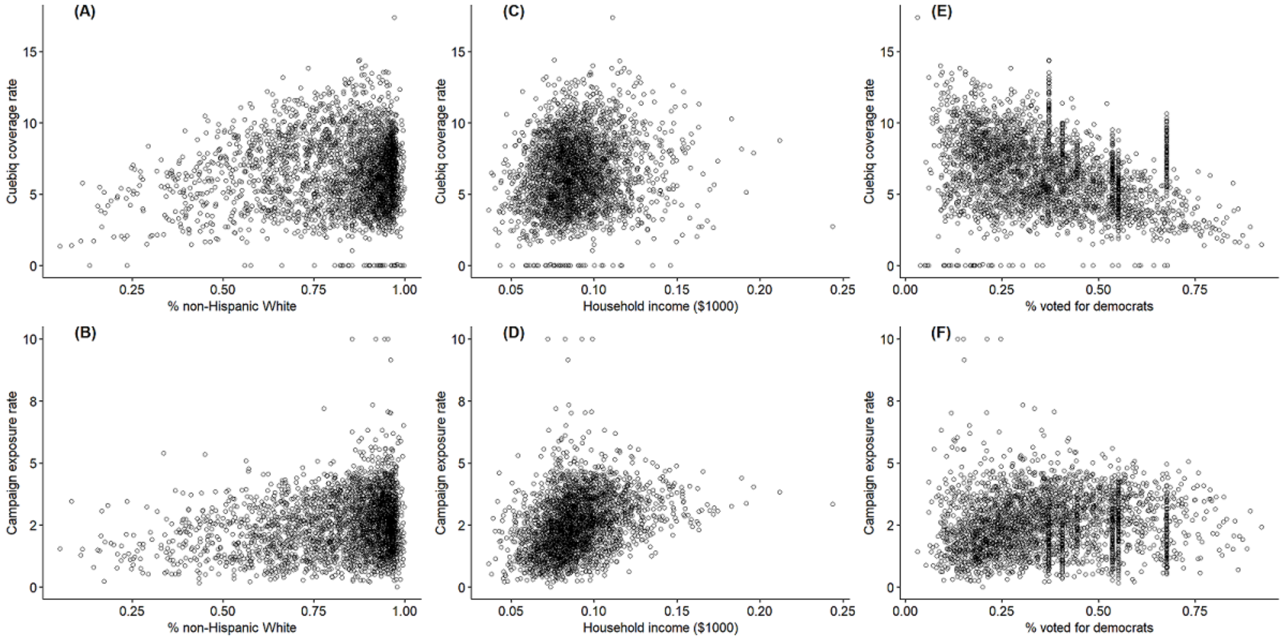


Figure 3. Distribution of county-level Centers for Disease Control and Prevention–reported weekly vaccination rate by week.

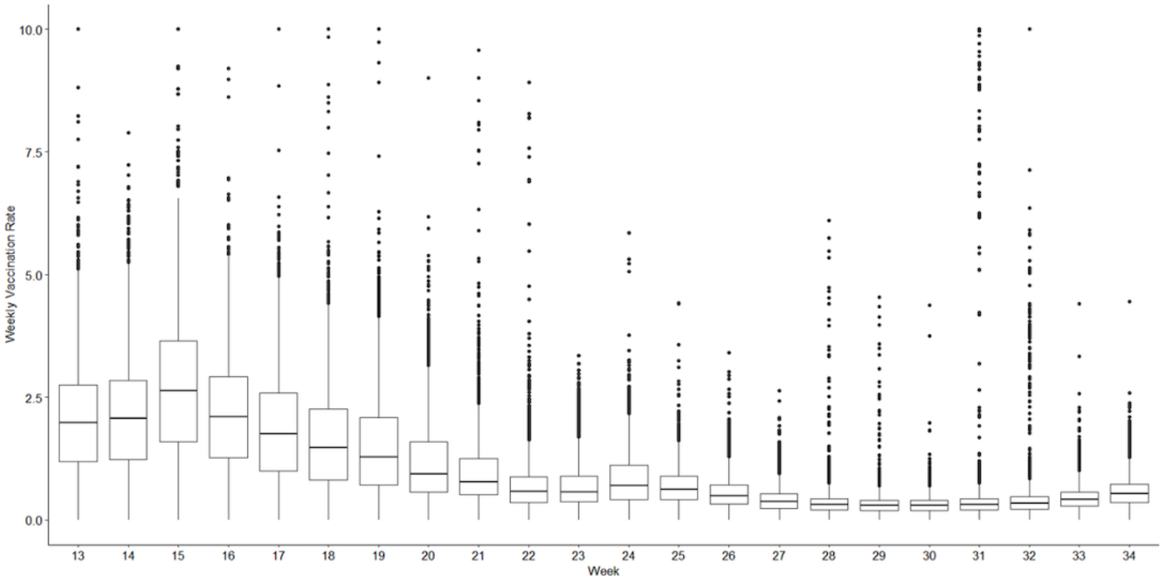


Figure 4. Exposure to the campaign and weekly vaccination rate by county during the first and last weeks of the study period. (A) Exposure rate (number of devices exposed to the campaign or total number of devices) to the campaign during the first week of the study period (week 13 or March 29 to April 4, 2021); (B) weekly vaccination rate during the first week of the study period; (C) exposure rate to the campaign during the last week of the study period (week 34 or August 12-29, 2021); (D) weekly vaccination rate during the last week of the study period.

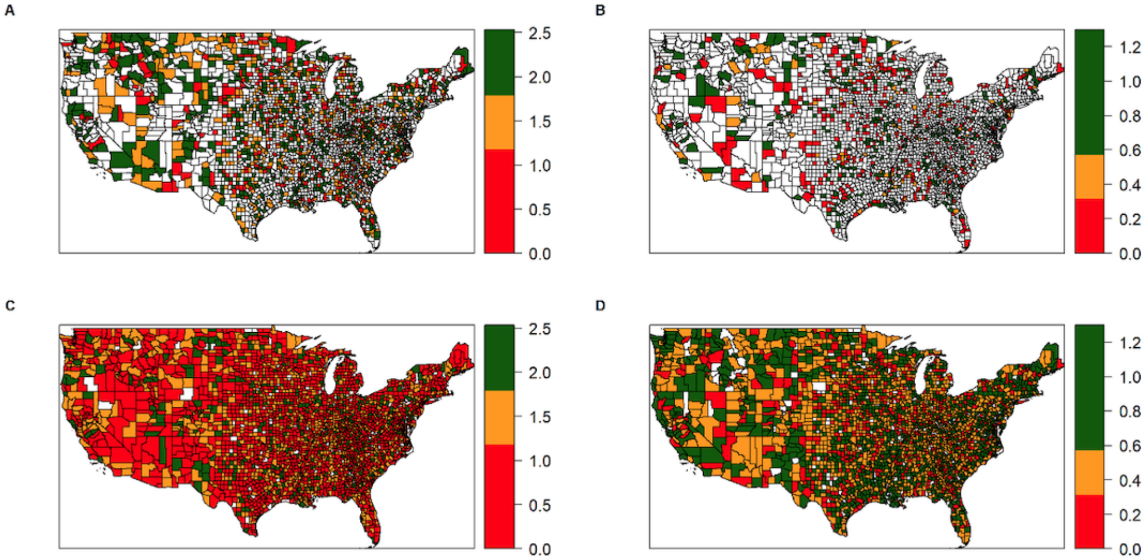


Table 1. Summary statistics of control variables for the 3104 counties in the study.

Variable	Mean (SD)	Min	Max
Population aged ≥85 years (%)	2.3 (0.01)	0.0	8.0
Non-Hispanic White (%)	83.3 (0.16)	5.1	99.8
Voted for democrats (%)	36.9 (0.17)	3.1	82.1
Household income (US $1000)	89.2 (20.0)	36.5	243.8

Table 2. Results from the Bayesian hierarchical model. Weekly data collected from 3104 counties during March 29 to August 29, 2021 (22 weeks); 95% uncertainty intervals were calculated from the posterior distributions.

Variable	Mean	95% uncertainty interval
Exposed to the campaign (%)	0.022	0.02 to 0.024
Population aged ≥85 years (%)	0.031	0.021 to 0.042
Non-Hispanic White (%)	0.004	0.003 to 0.004
Voted for democrats (%)	0.014	0.013 to 0.014
Median household income (US $1000)	0.058	0.054 to 0.062
Intercept	–0.005	–0.005 to –0.004

Discussion

The study is one of the first rigorous evaluations of the “It’s Up to You” campaign, the largest effort in promoting COVID-19 vaccine awareness and uptake in the United States. We used Cuebiq’s nationwide location intelligence platform to monitor campaign exposure and then estimated its association with vaccine uptake data tracked by US Centers for Disease Control and Prevention. Results from the Bayesian hierarchical model support the effectiveness of the campaign in promoting vaccine uptake.

Under Operation Warp Speed, the COVID-19 vaccine research and development achieved great success at unprecedented scale and speed, thanks to the coordinated effort from governments, pharmaceutical companies, and other stakeholders [24]. The US Food and Drug Administration granted Emergency Use Authorization to 2 highly effective COVID-19 vaccines in late 2020. Since then, barriers to achieving herd immunity have shifted from medical to social.

On May 25, 2021, the United States marked the milestone of inoculating half of the adults in the country. Vaccinating the remaining population will be harder. Even before the milestone, vaccinations had been plateauing, and the official data had been chronicling a decreasing daily vaccination rate. Research and expert consensus suggest that as high as 80%-90% of the total population needs to be vaccinated to achieve herd immunity after considering recent factors, such as more contagious variants [25]. Given the high prevalence of vaccine hesitancy and uneven distribution across regions and populations, convincing the hesitant groups to get inoculated is necessary.

Overall, the already vaccinated population is better informed, less vaccine hesitant and skeptical, and more eager to be vaccinated than the remaining unvaccinated group. As the country approaches the vaccination rate required for herd immunity, vaccine resistance is also becoming more prominent in the remaining unvaccinated population. Reaching herd immunity requires vaccinating a significant proportion of the resistant population in addition to vaccinating all nonresistant populations. Thus, effective messaging campaigns are urgently needed. The root causes of vaccine hesitancy, such as fear, mistrust, and misconceptions, need to be addressed through targeted communication strategies.

The study has several limitations. First, the campaign was not designed to be a randomized control trial. People (or devices) were not randomly chosen to be exposed to the campaign messages. However, it is unlikely that there were systematic self-selection biases, given the passive nature of campaign exposure. Second, vaccine uptake requires a complex decision-making process that involves a potentially large number of factors. Due to limited data availability, our model may have missed some important variables that affect vaccine uptake. Third, as a county-level ecological analysis, it is challenging to make a causal inference. Lastly, the Cuebiq platform only monitors exposure to digital campaigns. Exposure to traditional media (eg, print publication, billboards, and posters) is not captured in the study.

Despite the limitations, the study is among the first efforts to evaluate the nation’s largest public health campaign. We built a fully validated model and established a statistically significant association between campaign exposure and vaccine uptake at the county level. The sizable effectiveness suggests that Ad Council is a sound investment of public and private resources. Given the low exposure and slow vaccination rate as of the end of the study period, expanding the messaging campaign may help accelerate vaccination progress.

The United States is no outlier when it comes to vaccine hesitancy, which was reported for nearly every vaccine and named one of the top ten threats to global health by the World Health Organization [26]. Misinformation about vaccines and antivax messages are global concerns. For example, a factor for Israel’s successful vaccination rollout is its multiprolonged educational campaigns that provide information, allay fears, and overcome hesitancy [27]. The successful implementation of the “It is Up to You” campaign and the documented effects in promoting vaccine uptake may generate valuable experience for other countries facing similar challenges.

Data Availability

Cuebiq data are available upon request from the Cuebiq Data for Good Program. All other data used in the study are publicly available from respective sources.

Authors' Contributions

QL conceptualized the study, developed the models, visualized the results, and wrote the manuscript. JCP and ARE extracted the data from the Cuebiq platform. AL, CY, and BL secured funding. All authors reviewed and revised the manuscript.

Conflicts of Interest

None declared.

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534. [doi: 10.1016/S1473-3099(20)30120-1]
2. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416. [doi: 10.1056/nejmoa2035389]
3. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
4. Zhang Y, Fisk RJ. Barriers to vaccination for coronavirus disease 2019 (COVID-19) control: experience from the United States. Glob Health J 2021 Mar;5(1):51-55 [FREE Full text] [doi: 10.1016/j.glohj.2021.02.005] [Medline: 33585053]
5. Americans' views on a COVID-19 vaccine. Covid Collaborative. URL: https://www.covidcollaborative.us/resources/americans-views-on-a-covid-19-vaccine [accessed 2023-01-25]
6. Coustasse A, Kimble C, Maxik K. COVID-19 and vaccine hesitancy: a challenge the United States must overcome. J Ambul Care Manage 2021;44(1):71-75. [doi: 10.1097/JAC.0000000000000360] [Medline: 33165121]
7. Khubchandani J, Sharma S, Price JH, Wiblishauser MJ, Sharma M, Webb FJ. COVID-19 vaccination hesitancy in the United States: a rapid national assessment. J Community Health 2021 Apr 03;46(2):270-277 [FREE Full text] [doi: 10.1007/s10900-020-00958-x] [Medline: 33389421]

8. Coronavirus vaccine hesitancy in Black and Latinx communities. Covid Collaborative. URL: https://www.covidcollaborative.us/content/vaccine-treatments/coronavirus-vaccine-hesitancy-in-black-and-latinx-communities [accessed 2023-01-25]
9. Savoia E, Piltch-Loeb R, Goldberg B, Miller-Idriss C, Hughes B, Montrond A, et al. Predictors of COVID-19 vaccine hesitancy: socio-demographics, co-morbidity, and past experience of racial discrimination. Vaccines (Basel) 2021 Jul 09;9(7):767 [FREE Full text] [doi: 10.3390/vaccines9070767] [Medline: 34358184]
10. Koslap-Petraco M. Vaccine hesitancy: not a new phenomenon, but a new threat. J Am Assoc Nurse Pract 2019 Nov;31(11):624-626. [doi: 10.1097/jxx.0000000000000342]
11. Tewarson H, Greene K, Fraser MR. State strategies for addressing barriers during the early US COVID-19 vaccination campaign. Am J Public Health 2021 Jun;111(6):1073-1077. [doi: 10.2105/ajph.2021.306241]
12. Jain R, Kelly CA, Mehta S, Tolliver D, Stewart A, Perdomo J. A trainee-led social media advocacy campaign to address COVID-19 inequities. Pediatrics 2021 Mar;147(3):e2020028456. [doi: 10.1542/peds.2020-028456] [Medline: 33334921]
13. COVID-19 vaccine education initiative. Ad Council. 2021. URL: https://www.adcouncil.org/covid-vaccine [accessed 2023-01-25]
14. Hornik R. Public Health Communication. Evidence for Behavior Change. New York, NY: Routledge; 2002.
15. Bertrand J, Anhang R. The effectiveness of mass media in changing HIV/AIDS-related behaviour among young people in developing countries. World Health Organ Tech Rep Ser 2006;938:205-41; discussion 317. [Medline: 16921921]
16. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. The Lancet 2010 Oct;376(9748):1261-1271. [doi: 10.1016/s0140-6736(10)60809-4]
17. Bleakley A, Jordan A, Mallya G, Hennessy M, Piotrowski JT. Do you know what your kids are drinking? evaluation of a media campaign to reduce consumption of sugar-sweetened beverages. Am J Health Promot 2018 Jul 14;32(6):1409-1416. [doi: 10.1177/0890117117721320] [Medline: 28805073]
18. Farrelly MC, Duke JC, Nonnemaker J, MacMonegle AJ, Alexander TN, Zhao X, et al. Association between the real cost media campaign and smoking initiation among youths - United States, 2014-2016. MMWR Morb Mortal Wkly Rep 2017 Jan 20;66(2):47-50 [FREE Full text] [doi: 10.15585/mmwr.mm6602a2] [Medline: 28103214]
19. Kang D, Choi H, Kim J, Choi J. Spatial epidemic dynamics of the COVID-19 outbreak in China. Int J Infect Dis 2020 May;94:96-102 [FREE Full text] [doi: 10.1016/j.ijid.2020.03.076] [Medline: 32251789]
20. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math 1991 Mar;43(1):1-20. [doi: 10.1007/bf00116466]
21. Vallejos R, Osorio F. Effective sample size of spatial process models. Spat Stat 2014 Aug;9:66-92. [doi: 10.1016/j.spasta.2014.03.003]
22. Lee D, Rushworth A, Napier G. Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. J.Stat Soft 2018;84(9):1-39. [doi: 10.18637/jss.v084.i09]
23. Hughes MM, Wang A, Grossman MK, Pun E, Whiteman A, Deng L, et al. County-level COVID-19 vaccination coverage and social vulnerability - United States, December 14, 2020-March 1, 2021. MMWR Morb Mortal Wkly Rep 2021 Mar 26;70(12):431-436 [FREE Full text] [doi: 10.15585/mmwr.mm7012e1] [Medline: 33764963]
24. Slaoui M, Hepburn M. Developing safe and effective Covid vaccines — operation warp speed's strategy and approach. N Engl J Med 2020 Oct 29;383(18):1701-1703. [doi: 10.1056/nejmp2027405]
25. Barker P. Rethinking herd immunity: managing the COVID-19 pandemic in a dynamic biological and behavioral environment. NEJM Catalyst Innovations in Care Delivery. 2021. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.21.0288 [accessed 2023-01-25]
26. Solís Arce JS, Warren SS, Meriggi NF, Scacco A, McMurry N, Voors M, et al. COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. Nat Med 2021 Aug;27(8):1385-1394 [FREE Full text] [doi: 10.1038/s41591-021-01454-y] [Medline: 34272499]
27. Rosen B, Waitzberg R, Israeli A. Israel's rapid rollout of vaccinations for COVID-19. Isr J Health Policy Res 2021 Jan 26;10(1):6 [FREE Full text] [doi: 10.1186/s13584-021-00440-6] [Medline: 33499905]

Abbreviations

CAR: conditional autoregressive

Edited by T Sanchez, A Mavragani; submitted 01.05.22; peer-reviewed by J Lee, B Eshrati; comments to author 11.11.22; revised version received 04.12.22; accepted 22.12.22; published 16.02.23.

Please cite as:
Li Q, Peng JC, Mohan D, Lake B, Euler AR, Weir B, Kan L, Yang C, Labrique A
Using Location Intelligence to Evaluate the COVID-19 Vaccination Campaign in the United States: Spatiotemporal Big Data Analysis
JMIR Public Health Surveill 2023;9:e39166
URL: https://publichealth.jmir.org/2023/1/e39166
doi:10.2196/39166
PMID:36626835

Original Paper

Substance Use From Social Distancing and Isolation by US Nativity During the Time of COVID-19: Cross-sectional Study

Francisco Alejandro Montiel Ishino[1*], MPH, CPH, PhD; Kevin Villalobos[2*], BS; Faustine Williams[2*], MPH, MS, PhD

[1]Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Durham, NC, United States
[2]Division of Intramural Research, National Institute on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD, United States
[*]all authors contributed equally

Corresponding Author:
Francisco Alejandro Montiel Ishino, MPH, CPH, PhD
Division of Intramural Research
National Institute of Environmental Health Sciences
National Institutes of Health
111 TW Alexander Dr
Durham, NC, 27709
United States
Phone: 1 6143643314
Email: francisco.montielishino@nih.gov

Abstract

Background: The COVID-19 pandemic had many unprecedented secondary outcomes resulting in various mental health issues leading to substance use as a coping behavior. The extent of changes in substance use in a US sample by nativity has not been previously described.

Objective: This study aimed to design a web-based survey to assess the social distancing and isolation issues exacerbated by the COVID-19 pandemic to describe substance use as a coping behavior by comparing substance use changes before and during the pandemic.

Methods: A comprehensive 116-item survey was designed to understand the impact of COVID-19 and social distancing on physical and psychosocial mental health and chronic diseases. Approximately 10,000 web-based surveys were distributed by Qualtrics LLC between May 13, 2021, and January 09, 2022, across the United States (ie, continental United States, Hawaii, Alaska, and territories) to adults aged ≥18 years. We oversampled low-income and rural adults among non-Hispanic White, non-Hispanic Black, Hispanic or Latino, and foreign-born participants. Of the 5938 surveys returned, 5413 (91.16%) surveys were used after proprietary expert review fraud detection (Qualtrics) and detailed assessments of the completion rate and the timing to complete the survey. Participant demographics, substance use coping behaviors, and substance use before and during the pandemic are described by the overall US resident sample, followed by US-born and foreign-born self-reports. Substance use included the use of tobacco, e-cigarettes or nicotine vapes, alcohol, marijuana, and other illicit substances. Marginal homogeneity based on the Stuart-Maxwell test was used to assess changes in self-reported substance use before and during the pandemic.

Results: The sample mostly included White (2182/5413, 40.31%) and women participants (3369/5406, 62.32%) who identified as straight or heterosexual (4805/5406, 88.88%), reported making ≥US $75,000 (1405/5355, 26.23%), and had vocational or technical training (1746/5404, 32.31%). Similarities were observed between the US-born and the foreign-born participants on increased alcohol consumption: from no alcohol consumption before the pandemic to consuming alcohol once to several times a month and from once to several times per week to every day to several times per day. Although significant changes were observed from no prior alcohol use to some level of increased use, the opposite was also observed and was more pronounced among foreign-born participants. That is, there was a 5.1% overall change in some level of alcohol use before the pandemic to no alcohol use during the pandemic among foreign-born individuals, compared with a 4.3% change among US-born individuals.

Conclusions: To better prepare for the inadvertent effects of public health policies meant to protect individuals, we must understand the mental health burdens that can precipitate into substance use coping mechanisms that not only have a deleterious effect on physical and mental health but also exacerbate morbidity and mortality in a disease like COVID-19.

(JMIR Public Health Surveill 2023;9:e38163) doi:10.2196/38163

KEYWORDS

substance use; COVID-19; US nativity

Introduction

Background

The COVID-19 pandemic, as of March 2022, has >456 million recorded cases and 6 million reported deaths worldwide [1]. Although COVID-19 is caused by SARS-CoV-2, increased morbidity and mortality are associated with multiple direct and indirect social, physiological, and environmental factors [2,3]. There were, however, indirect health effects of the pandemic that affected mental health [4] and exacerbated maladaptive coping mechanisms such as substance use [5,6].

Coping strategies provide the ability to manage external and internal demands, given an individual's resources [7]. When demands are exceeded owing to high levels of physiological and psychological stress, as observed during the COVID-19 pandemic, it can lead to the adoption of maladaptive coping behaviors [8,9]. As such, mental health issues were exacerbated by multiple social and environmental stressors, social distancing, and isolation during the COVID-19 pandemic that in turn affected coping mechanisms leading to changes in substance use patterns [10,11].

Statistics for June 2020, when compared with 2019, as reported by the Centers for Disease Control and Prevention [5], show that 13% of US adults aged ≥18 years started or increased substance use to cope with stress from the COVID-19 pandemic. Reports of increased use were found for both licit and illicit substances [12,13] such as alcohol [14] and opioids [15], respectively. There were also mixed findings regarding the use of substances such as tobacco [16-18] and marijuana [19] as well as their respective methods of use through vaping [13,20,21]. Moreover, using licit and illicit substances can cause increased morbidity and mortality, especially when considering the physiological effects of COVID-19, in addition to worsening mental health [5].

The effects of social distancing and isolation early in the pandemic were compounded by multiple issues associated with access to health, mental health, and related telehealth services [22]. Multiple deficits and barriers to mental health services, especially among underserved and underrepresented communities, precipitated health care access and treatment disparities [6,10,23]. These deficits and barriers were then exacerbated by the COVID-19 pandemic and may have led to substance use as a prevalent coping mechanism during a period when substance treatment options were already limited [10,22,23]. Nonetheless, studies on substance use as a coping strategy during the COVID-19 pandemic are limited.

Objectives

Further research is needed to understand the multiple and varying sociodemographic and socioeconomic factors to fill this gap. In addition, there is a critical need to include both US-born and foreign-born individuals in substance use research during the COVID-19 pandemic as these studies are still limited. A limited number of studies have reported mixed findings, that is, increased and decreased substance use during the pandemic [24,25]. Nevertheless, the psychological effects of substance use may be detrimental and synergize disparities among racial and ethnic foreign-born minorities [24,26,27]. As such, our purpose was to describe substance use as a coping mechanism to COVID-19–induced social distancing and isolation by comparing substance use before and during the pandemic by US-born and foreign-born individuals.

Methods

Overview

Our study *Understanding the Impact of the Novel Coronavirus (COVID-19) and Social Distancing on Physical and Psychosocial (Mental) Health and Chronic Diseases* created a comprehensive 116-item web-based survey that was nationally distributed in the United States. The survey modules included (1) general health status; (2) COVID-19 symptoms, testing, and prevention; (3) chronic illness management; (4) social distancing; (5) mental health; (6) pandemic economic impact; (7) discrimination; and (8) sociodemographics.

The target population comprised adults aged ≥18 years residing in the United States. The US resident sample included both US-born and foreign-born participants. Qualtrics LLC was contracted to facilitate the recruitment and distribution of the web-based survey to both US-born and foreign-born racial and ethnic groups. The US-born racial and ethnic groups included Hispanic or Latino, White, Black, Asian, American Indian and Alaskan Native, and Native Hawaiian and Pacific Islander participants. Foreign-born racial and ethnic groups included African, Middle Eastern, Hispanic or Latino, and Asian participants. Qualtrics then used proprietary consumer panels to randomly sample White participants that matched demographic characteristics with other racial and ethnic groups. We oversampled adults with low income (<US $25,000 annual household income) who resided in rural areas (self-reported and cross-referenced with zip codes already collected by Qualtrics) among non-Hispanic White, non-Hispanic Black, Hispanic, and foreign-born participants. The survey was available only in English.

A total of 10,000 surveys were distributed between May 13, 2021, and January 09, 2022. The initial surveys received by Qualtrics were assessed via expert review fraud detection to prevent multiple submissions and detect "bots" to protect the integrity of the data. After the assessment, 5938 surveys were received by the research team from Qualtrics. Information Management Services, Inc, a research support firm that provides analytic services, was given the task to clean and manage the deidentified survey data.

To improve study integrity, initial data cleaning by the Information Management Services included flagging surveys based on the completion rate and the timing to complete the survey. Participants were flagged and removed from the analysis if they completed <80% of the survey based on 102 questions after accounting for skip pattern items or took <5 minutes to complete the survey. In total, 125 surveys were removed at this stage, giving us 5813. Our study ended with a total of 5413 surveys based on the completed responses in the social distancing module of the *Understanding the Impact of the Novel Coronavirus (COVID-19) and Social Distancing on Physical and Psychosocial (Mental) Health and Chronic Diseases* survey. The survey can be requested from the principal investigator, Faustine Williams, PhD, MPH, MS, from the National Institute on Minority Health and Health Disparities of the National Institutes of Health (NIH) in Bethesda, Maryland.

Ethical Considerations

Qualtrics recruited study participants, and web-based informed consent was provided before the survey. Participants were asked to participate in a voluntary research study titled *Understanding the Impact of the Novel Coronavirus (COVID-19) and Social Distancing on Physical and Psychosocial (Mental) Health and Chronic Diseases* conducted by Faustine Williams. If they chose to participate in the study, they were assured that their responses would be kept confidential. Participants were reminded that the study was voluntary; as such, they could change their minds at any time after starting the survey and opt out with no fear of repercussions. Participants could skip any questions that they did not want to answer. All answers were kept confidential, and Qualtrics assured the participants that no personal identifiers would be shared with the National Institute on Minority Health and Health Disparities research team. Participants were to receive an incentive of a US $10 gift card after the completion of their survey, which would take approximately 30 minutes of their time. At the end of the survey, participants were also asked if they would be interested in participating in another follow-up survey. Then, Qualtrics would follow-up with those who responded that they were interested. Participants were provided the contact information of the principal investigator (Faustine Williams) as well as the phone number of the NIH Institutional Review Board.

The research protocol for this study was reviewed by the NIH, Intramural Research Program Institutional Review Board, Human Research Protection Program, and Office of Human Subjects Research Protections and received an exemption on December 23, 2020 (IRB#000308). The NIH, Intramural Research Program Institutional Review Board, Human Research Protection Program, and Office of Human Subjects Research Protections determined that our protocol did not involve human participants and was excluded from the institutional review board review.

Descriptors

Sociodemographic Variables

All sociodemographic items were self-reported and allowed for either the selection of multiple categories or provided a free response if they selected a blank or other category. Nativity was categorized by country of birth as either US born or foreign born. US-born nativity was based on respondents' self-reported births in the 48 contiguous states, Washington, the District of Columbia, Alaska, Hawaii, and other US territories such as Puerto Rico. Foreign-born nativity was based on respondents' self-reported births occurring in another country outside the United States based on the US birth classification. Racial and ethnic categories included selecting ≥1 of the following options: White, Black or African American, Asian, American Indian or Alaskan Native, Hawaiian or Pacific Islander, African, Middle Eastern, and multiracial or multiethnic. If respondents selected ≥2 racial or ethnic groups, they were classified as multiracial and multiethnic. Gender categories included men, women, nonbinary, transgender people, and others. Sexual orientation included straight or heterosexual, lesbian, gay, bisexual, and other. The lesbian and gay categories were combined. Age was self-reported starting from 18 to ≥85 years. Age categories were then constructed as follows: 18 to 35 years, 36 to 55 years, and 56 to ≥85 years. Annual household income was reported as <US $25,000, US $25,000 to US $34,999, US $35,000 to US $49,999, US $50,000 to US $74,999, and US ≥$75,000. Educational attainment was categorized by self-reported highest schooling that included (1) less than high school or General Language Development (ie, did not attend school; elementary education, 6 years or less; more than elementary to junior high school; or some high school), (2) high school diploma or General Language Development, (3) some college or vocational or technical schooling, (4) bachelor's degree, and (5) master's degree or above (ie, master's degree or doctoral, professional, or postgraduate degree). Employment status was assessed using multiple survey items. Current employment (ie, employed, self-employed, unpaid or voluntary work, apprenticeship or vocational training, disabled, caretaker or looking after family or home, in school, retired, or unemployed) and if considered an essential worker (ie, no or yes). Unemployed or nontraditional work was categorized as being disabled, a caretaker or looking after family or home, in school, retired, or unemployed. Nonessential workers were categorized as not considered an essential worker and employed, self-employed, unpaid or voluntary work, or apprenticeship or vocational training. Essential workers were categorized as being considered an essential worker and employed, self-employed, unpaid or voluntary work, or apprenticeship or vocational training.

Substance Use

Coping behaviors for social distancing and isolation during the COVID-19 pandemic were assessed by asking questions regarding substance use in the social distancing module of the survey. Questions specific to exclusively using the following substances to cope were used: cigarettes or vaping, increased alcohol use, marijuana use, and illicit substance use. To assess cases of substance use during the pandemic, we asked "During the past month, how often did you" (1) smoke cigarettes or other tobacco products for tobacco use, (2) smoke e-cigarettes or other nicotine vaping products, (3) have a drink containing alcohol for alcohol use, (4) use marijuana, and (5) use illicit drugs. Illicit drugs were defined as other substances that were not previously listed (ie, tobacco, nicotine, alcohol, or marijuana) and could include but not be limited to opiates, hallucinogens, cocaine, or amphetamines. A follow-up control question was used for each aforementioned question category that asked, "Compared to before the pandemic, this is or was... .” Responses to each question had the following levels: not at all, once during the month, several times a month, once a week, several times a week, almost every day or every day, and several times a day. Responses were collapsed to (1) not at all, (2) once to several times per month, (3) once to several times per week, and (4) every day to several times per day.

Analytic Procedure

Descriptive statistics of survey sample sociodemographics and substance use behaviors during the COVID-19 pandemic were assessed by the nativity of respondents, that is, US born and foreign born. Descriptives for the overall survey sample before and during the pandemic were assessed, followed by a more detailed assessment by nativity. A Stuart-Maxwell test was used to examine whether substance use before the COVID-19 pandemic was equal to substance use during the pandemic among survey respondents. If significant differences were found in substance use, we tested for differences based on participant self-reported nativity. The Stuart-Maxwell test is an ideal nonparametric test to examine asymptotic symmetry and marginal homogeneity on matched-pair controls (ie, before the COVID-19 pandemic) and cases (ie, during the COVID-19 pandemic) with various discrete levels of substance use (ie, not at all, once several times per month, once to several times per week, and every day to several times per day). All analytical procedures were conducted using Stata/MP (version 16.1; StataCorp LLC). All analytical files are available upon reasonable request.

Results

Overview

The overall sample was primarily racially and ethnically White, Black or African American, and Hispanic or Latin. Most of the participants self-reported as women (3369/5406, 62.32%); were straight or heterosexual (4805/5406, 88.88%); had a household annual income of US ≥$75,000 (1405/5355, 26.23%); had some college, vocational, or technical training (1746/5404, 32.31%); and were unemployed or participated in nontraditional work (2401/5405, 44.42%). The age of the sample was between 18 and 35 years (1839/5119, 35.92%) and 36 and 55 years (1971/5119, 38.5%). The sample self-reported primarily using the following substances in the following order based on proportion: marijuana use (109/5404, 2.02%), increased alcohol use (87/5404, 1.61%), e-cigarette or nicotine vape (83/5404, 1.54%), and illicit substance use (24/5404, 0.44%). A more detailed breakdown of the sociodemographic profiles and substance use by nativity is shown in Table 1.

Table 2 describes substance use before and during the COVID-19 pandemic for the overall sample. Most of the participants who reported substance use before the COVID-19 pandemic used tobacco every day to several times per day (940/5130, 18.32%), e-cigarettes or nicotine vapes once to several times per month (366/5130, 7.13%), alcohol once to several times per month (1563/5130, 30.47%), marijuana every day to several times per day (471/5130, 9.18%), and other illicit substances once to several times per month (225/5130, 4.38%). During the COVID-19 pandemic, the same participants used tobacco every day to several times per day (664/3802, 17.46%), e-cigarettes or nicotine vapes once to several times per week (269/3802, 7.07%), alcohol once to several times per month (1051/3802, 27.64%), marijuana every day to several times per day (336/3802, 8.84%), and other illicit substances once to several times per week (148/3802, 3.89%).

In Table 3, we describe substance use before and during the COVID-19 pandemic by US-born and foreign-born individuals. If reporting substance use before the COVID-19 pandemic, the US-born sample primarily reported use of tobacco every day to several times per day (846/3903, 21.68%), e-cigarettes or nicotine vapes once to several times per week (298/3889, 7.66%), alcohol once to several times per month (1207/3903, 30.92%), marijuana every day to several times per day (420/3873, 10.84%), and other illicit substances once to several times per week (196/3887, 5.04%). The same US-born sample during the COVID-19 pandemic, if reporting substance use, used tobacco every day to several times per day (600/2912, 20.60%), e-cigarettes or nicotine vapes once to several times per week (231/2899, 7.97%), alcohol once to several times per month (812/2921, 27.80%), marijuana every day to several times per day (300/2901, 10.34%), and other illicit substances once to several times per week (131/2896, 4.52%).

The foreign-born sample’s substance use before the COVID-19 pandemic indicated tobacco use every day to several times per day (94/1179, 7.97%), e-cigarettes or nicotine vapes once to several times per month (75/1176, 6.38%), alcohol once to several times per month (356/1187, 29.99%), marijuana once to several times per month (62/1170, 5.3%), and other illicit substances once to several times per month (44/1180, 3.73%). The same foreign-born sample during the COVID-19 pandemic, if reporting substance use, used tobacco every day to several times per day (67/860, 7.79%), e-cigarettes or nicotine vapes once to several times per week (39/857, 4.55%), alcohol once to several times per month (242/860, 28.14%), marijuana once to several times per month (41/855, 4.79%), and other illicit substances once to several times per month (24/852, 2.82%).

Table 1. Sample descriptives.

Participant characteristics	US born (n=4166), n (%)	Foreign born (n=1247), n (%)	Total (n=5413), n (%)
Race and ethnicity (n=5413)			
White	2053 (49.28)	129 (10.34)	2182 (40.31)
Black or African American	1024 (24.58)	181 (14.51)	1205 (22.26)
Hispanic or Latino	551 (13.23)	435 (34.88)	986 (18.22)
Asian	217 (5.21)	338 (27.11)	555 (10.25)
American Indian or Alaska Native	137 (3.29)	8 (0.64)	145 (2.68)
Hawaiian or Pacific Islander	47 (1.13)	14 (1.12)	61 (1.13)
African	28 (0.67)	26 (2.09)	54 (1)
Middle Eastern	7 (0.17)	21 (1.68)	28 (0.52)
Multiracial or multiethnic	102 (2.45)	95 (7.62)	197 (3.64)
Gender (n=5406)			
Man	1493 (35.88)	412 (33.09)	1905 (35.24)
Woman	2589 (62.22)	780 (62.65)	3369 (62.32)
Nonbinary	17 (0.41)	24 (1.93)	41 (0.76)
Transgender people	10 (0.24)	8 (0.64)	18 (0.33)
Other	52 (1.25)	21 (1.69)	73 (1.35)
Sexual orientation (n=5382)			
Straight or heterosexual	3728 (89.96)	1077 (87)	4805 (89.28)
Lesbian or gay	148 (3.57)	49 (3.96)	197 (3.66)
Bisexual	226 (5.45)	83 (6.7)	309 (5.74)
Other	42 (1.01)	29 (2.34)	71 (1.32)
Age (years; n=5119)			
18 to 35	1417 (35.5)	476 (42.24)	1893 (36.98)
36 to 55	1579 (39.55)	392 (34.78)	1971 (38.5)
56 to ≥85	996 (24.95)	259 (22.98)	1255 (24.52)
Household income (n=5355)			
<US $25,000	1029 (24.96)	274 (22.22)	1303 (24.33)
US $25,000 to $34,999	645 (15.65)	174 (14.11)	819 (15.29)
US $35,000 to $49,999	638 (15.48)	195 (15.82)	833 (15.56)
US $50,000 to $74,999	755 (18.32)	240 (19.46)	995 (18.58)
US ≥$75,000	1055 (25.59)	350 (28.39)	1405 (26.24)
Educational attainment (n=5404)			
Less than high school	221 (5.31)	99 (7.97)	320 (5.92)
High school or general education diploma	1008 (24.22)	237 (19.08)	1245 (23.04)
Some college, vocational or technical	1451 (34.86)	295 (23.75)	1746 (32.31)
Bachelor’s degree	1031 (24.77)	375 (30.19)	1406 (26.02)
Master’s degree or above	451 (10.84)	236 (19)	687 (12.71)
Employment (n=5405)			
Unemployed or nontraditional work	1819 (43.7)	582 (46.82)	2401 (44.42)
Nonessential worker	1300 (31.23)	393 (31.62)	1693 (31.32)
Essential worker	1043 (25.06)	268 (21.56)	1311 (24.26)
Substance use coping behaviors (n=5404)			

Participant characteristics	US born (n=4166), n (%)	Foreign born (n=1247), n (%)	Total (n=5413), n (%)
Other coping behaviors	3895 (93.63)	1206 (96.94)	5101 (94.39)
E-cigarette or nicotine vape use	75 (1.8)	8 (0.64)	83 (1.54)
Increased alcohol use	68 (1.63)	19 (1.53)	87 (1.61)
Marijuana use	99 (2.38)	10 (0.8)	109 (2.02)
Illicit substance use	23 (0.55)	1 (0.08)	24 (0.44)

Table 2. Substance use before and during the COVID-19 pandemic.

	Before the COVID-19 pandemic (n=5130), n (%)	During the COVID-19 pandemic (n=3802), n (%)
Tobacco use		
Not at all	3418 (67.26)	2612 (69.32)
Once to several times per month	366 (7.2)	232 (6.16)
Once to several times per week	358 (7.04)	260 (6.9)
Every day to several times per day	940 (18.5)	664 (17.62)
E-cigarette or nicotine vape use		
Not at all	4017 (79.31)	3028 (80.7)
Once to several times per month	363 (7.17)	219 (5.84)
Once to several times per week	352 (6.95)	269 (7.17)
Every day to several times per day	333 (6.57)	236 (6.29)
Alcohol use		
Not at all	2140 (42.04)	1625 (43.02)
Once to several times per month	1563 (30.71)	1051 (27.83)
Once to several times per week	1083 (21.28)	814 (21.55)
Every day to several times per day	304 (5.97)	287 (7.6)
Marijuana use		
Not at all	3778 (74.92)	2857 (76.15)
Once to several times per month	460 (9.12)	313 (8.34)
Once to several times per week	334 (6.62)	246 (6.56)
Every day to several times per day	471 (9.34)	336 (8.96)
Illicit substance use		
Not at all	4470 (88.29)	3342 (89.26)
Once to several times per month	225 (4.44)	136 (3.63)
Once to several times per week	216 (4.27)	148 (3.95)
Every day to several times per day	152 (3)	118 (3.15)

Table 3. Substance use before and during the COVID-19 pandemic by US-born and foreign-born participants.

	US born (n=5130), n (%)		Foreign born (n=3802), n (%)	
	Before the COVID-19 pandemic	During the COVID-19 pandemic	Before the COVID-19 pandemic	During the COVID-19 pandemic
Tobacco use				
Not at all	2456 (62.93)	1903 (65.35)	962 (81.59)	710 (82.56)
Once to several times per month	299 (7.66)	191 (6.56)	67 (5.68)	41 (4.77)
Once to several times per week	302 (7.74)	218 (7.49)	56 (4.75)	42 (4.88)
Every day to several times per day	846 (21.68)	600 (20.6)	94 (7.97)	67 (7.79)
E-cigarette or nicotine vape use				
Not at all	3017 (77.58)	2277 (78.54)	1000 (85.03)	752 (87.75)
Once to several times per month	288 (7.41)	183 (6.31)	75 (6.38)	37 (4.32)
Once to several times per week	298 (7.66)	231 (7.97)	54 (4.59)	39 (4.55)
Every day to several times per day	286 (7.35)	208 (7.17)	47 (4)	29 (3.38)
Alcohol use				
Not at all	1554 (39.82)	1190 (40.74)	586 (49.37)	436 (50.7)
Once to several times per month	1207 (30.92)	812 (27.8)	356 (29.99)	242 (28.14)
Once to several times per week	894 (22.91)	680 (23.28)	189 (15.92)	134 (15.58)
Every day to several times per day	248 (6.35)	239 (8.18)	56 (4.72)	48 (5.58)
Marijuana use				
Not at all	2761 (71.29)	2107 (72.63)	1017 (86.92)	752 (87.95)
Once to several times per month	398 (10.28)	273 (9.41)	62 (5.30)	41 (4.8)
Once to several times per week	294 (7.59)	221 (7.62)	40 (3.42)	26 (3.04)
Every day to several times per day	420 (10.84)	300 (10.34)	51 (4.36)	36 (4.21)
Illicit substance use				
Not at all	3380 (86.96)	2551 (88.09)	1094 (92.71)	793 (93.07)
Once to several times per month	181 (4.66)	112 (3.87)	44 (3.72)	24 (2.82)
Once to several times per week	196 (5.04)	131 (4.52)	20 (1.69)	19 (2.23)
Every day to several times per day	130 (3.34)	102 (3.52)	22 (1.86)	16 (1.88)

Differences in Substance Use Before and During the COVID-19 Pandemic

Using the Stuart-Maxwell test of asymptotic symmetry and marginal homogeneity, we found significant differences in alcohol use in the overall sample (Table 4). In Table 5, we see that the largest contribution to χ^2 symmetry (χ^2_3=20.2) was owing to differences in no alcohol use before the COVID-19 pandemic to once to several times per month during the COVID-19 pandemic (ie, a 1.71% change) and once to several times per month before the COVID-19 pandemic to not at all during the COVID-19 pandemic (ie, a 3.37% change). The second largest contribution (χ^2_3=15.6) was observed with alcohol use once to several times per week before the COVID-19 pandemic to every day to several times per day during the COVID-19 pandemic (ie, a 2.25% change) and every day to several times per day before the COVID-19 pandemic to once to several times per week during to the COVID-19 pandemic (ie, a 1.07% change). All other substance use changes reported before and during the pandemic were not significant but can be found in Multimedia Appendix 1.

Then, we assessed differences in substance use before and during the pandemic by nativity and found them to be significantly different between US-born and foreign-born individuals (Table 5). Among the US-born participants, the largest contribution to the symmetry χ^2 (χ^2_3=13.3) was owing to differences in no alcohol use before the COVID-19 pandemic to once to several times per month during the COVID-19 pandemic (ie, a 1.77% change) and once to several times per month before the COVID-19 pandemic to not at all during the COVID-19 pandemic (ie, a 3.28% change). The second largest contribution among US-born participants (χ^2_3=13.2) was in alcohol use once to several times per week before the COVID-19 pandemic to every day to several times per day during the COVID-19 pandemic (ie, a 2.22% change) and every day to several times per day before the COVID-19 pandemic to once to several times per week during to the COVID-19 pandemic (ie, a 1.00% change). See Table 6 for more detail.

As seen in Table 7 among foreign-born participants, the largest contribution to the symmetry χ^2 (χ^2_3=7.4) was owing to differences in no prior alcohol use before the COVID-19 pandemic to once to several times per month during the COVID-19 pandemic (ie, a 1.5% change) and once to several times per month before the COVID-19 pandemic to not at all during the COVID-19 pandemic (ie, a 3.7% change). The second largest χ^2 contribution was alcohol use once to several times per month to every day to several times a day (χ^2_3=3.8). Of the 843 respondents, 23 (2.7%) shifted from alcohol use once to several times per month before the pandemic to every day to several times per day during the pandemic, compared with 18 (2.1%) who shifted from alcohol use every day to several times per day before the COVID-19 pandemic to once to several times per month during the COVID-19 pandemic. The third largest contribution (χ^2_3=2.6) was in alcohol use once to several times per week before the COVID-19 pandemic to every day to several times per day during the COVID-19 pandemic (ie, a 2.4% change) and every day to several times per day before the COVID-19 pandemic to once to several times per week during the COVID-19 pandemic (ie, a 1.3% change). Refer to Table 5 for all *the* χ^2 contributions to symmetry.

Table 4. Alcohol use before the COVID-19 pandemic compared with during the pandemic[a].

Before the COVID-19 pandemic	During the COVID-19 pandemic				
	Not at all, n (Δ%)	Once to several times per month, n (Δ%)	Once to several times per week, n (Δ%)	Every day to several times per day, n (Δ%)	Total
Not at all	1432 (N/A[b])	64 (1.71)	22 (0.59)	4 (0.11)	1522
Once to several times per month	126 (3.37)	826 (N/A)	136 (3.64)	15 (0.40)	1103
Once to several times per week	29 (0.77)	127 (3.40)	611 (N/A)	84 (2.25)	851
Every day to several times per day	12 (0.32)	21 (0.56)	40 (1.07)	182 (N/A)	255
Total	1599 (N/A)	1038 (N/A)	809 (N/A)	285 (N/A)	3731

[a] Symmetry (asymptotic) was based on χ^2_6=42.1 and $P\leq.001$; marginal homogeneity was based on the Stuart-Maxwell test (χ^2_3=30.0; $P\leq.001$).

[b] N/A: not applicable (as these are the references to compare and contrast contributions to symmetry).

Table 5. Contribution to symmetry χ^2 from alcohol use before and during the COVID-19 pandemic.

Change	Before or during COVID-19	Overall χ^2 (*df*)	US born χ^2 (*df*)	Foreign born χ^2 (*df*)
Not at all	Once to several times per month	20.2 (3)	13.3 (3)	7.4 (3)
Not at all	Once to several times per week	1.0 (3)	0.1 (3)	1.7 (3)
Not at all	Every day to several times per day	4.0 (3)	2.6 (3)	2.0 (3)
Once to several times per month	Once to several times per week	0.3 (3)	0.1 (3)	0.6 (3)
Once to several times per month	Every day to several times per day	1.0 (3)	0.0 (3)	3.8 (3)
Once to several times per week	Every day to several times per day	15.6 (3)	13.2 (3)	2.6 (3)

Table 6. Alcohol use before the COVID-19 pandemic compared with during the pandemic among US-born participants[a].

Before the COVID-19 pandemic	During the COVID-19 pandemic				
	Not at all, n (Δ%)	Once to several times per month, n (Δ%)	Once to several times per week, n (Δ%)	Every day to several times per day, n (Δ%)	Total
Not at all	1049 (N/A[b])	51 (1.77)	17 (0.59)	4 (0.14)	1121
Once to several times per month	95 (3.28)	630 (N/A)	113 (3.91)	12 (0.42)	850
Once to several times per week	19 (0.66)	109 (3.77)	518 (N/A)	64 (2.22)	710
Every day to several times per day	10 (0.34)	11 (0.38)	29 (1.00)	157 (N/A)	207
Total	1173 (N/A)	801 (N/A)	677 (N/A)	237 (N/A)	2888

[a] Marginal homogeneity based on Stuart-Maxwell χ^2_3 of 22.6 (P>.001).

[b] N/A: not applicable (as these are the references to compare and contrast contributions to symmetry).

Table 7. Alcohol use before the COVID-19 pandemic compared with during the pandemic among foreign-born participants[a].

Before the COVID-19 pandemic	During the COVID-19 pandemic				
	Not at all, n (Δ%)	Once to several times per month, n (Δ%)	Once to several times per week, n (Δ%)	Every day to several times per day, n (Δ%)	Total
Not at all	383 (N/A[b])	13 (1.5)	5 (0.6)	0 (0)	401
Once to several times per month	31 (3.7)	196 (N/A)	23 (2.7)	3 (0.4)	253
Once to several times per week	10 (1.2)	18 (2.1)	93 (N/A)	20 (2.4)	141
Every day to several times per day	2 (0.2)	10 (1.2)	11 (1.3)	25 (N/A)	48
Total	426 (N/A)	237 (N/A)	132 (N/A)	48 (N/A)	843

[a] Marginal homogeneity based on Stuart-Maxwell test χ^2_3 of 210.4 (P=.02).

[b] N/A: not applicable (as these are the references to compare and contrast contributions to symmetry).

Discussion

Principal Findings

We assessed changes in the use of combustible tobacco, e-cigarette and nicotine vape, alcohol, marijuana, and other illicit substances. Although use of alcohol was found to have significant changes before and during the pandemic in our overall sample, we did not observe significant changes in the use of tobacco, e-cigarette and nicotine vape, marijuana, or other illicit substances (Multimedia Appendix 1). Then, we examined alcohol use changes by comparing US-born and foreign-born participants and found significant changes in each group. The χ^2 contributions to symmetry indicated that the largest contributors to significant changes in substance use before and during the COVID-19 pandemic were similar for the overall sample in the United States and the US-born sample. Similarities were observed across US-born and foreign-born samples; however, changes in increased alcohol use among these groups were observed on (1) no alcohol use before the pandemic to using alcohol once to several times a month and (2) once to several times per week to every day to several times per day. Increases in alcohol use may indicate maladaptive coping with the effects of social distancing and isolation [8,9]. The increase from weekly alcohol use to daily use indicates increased physiological and psychological risk as well as a risk of developing possible alcohol use disorders [28].

Although we found that significant changes were observed from no prior alcohol use to some level of increased use, we also observed the opposite in both the US-born and foreign-born groups. The decrease in alcohol use was slightly more pronounced among foreign-born participants. That is, there was a 5.1% overall change in some level of alcohol use before the pandemic to no alcohol use during the pandemic among foreign-born participants, compared with a 4.3% change among US-born participants. In our findings, the largest shift was not associated with increased alcohol use but with decreased alcohol use. This decrease in substance use may also be indicative of isolation [26].

Comparison With Prior Work

The use of both licit and illicit substances can have deleterious effects not only on mental health but also on physiological health and physical functioning, as well as damage organ systems that can increase morbidity and mortality from COVID-19 [29]. First, given our findings, the synergistic effects of isolation and alcohol use must be considered. Although participants may use alcohol to cope with the deleterious effects of isolation and psychological distress caused by the COVID-19 pandemic, the use of substances in themselves has been reported to be maladaptive [9,26]. One could argue that the use of alcohol has social connotations, whereas others would argue that substance use creates isolation owing to its taboo in the social context [26]. In the context of our findings, the increase from weekly alcohol use to daily use may be interpreted as increased personal use, not during social gatherings, given the context of social distancing practices. Conversely, decreases in alcohol use may also be tracked to the social nature of general alcohol consumption that the pandemic disrupted [30,31].

When specifically seeking to understand the differences in alcohol use between US-born and foreign-born participants during the COVID-19 pandemic, the available studies were limited. Overall, when foreign-born immigrants were assessed for substance use during the pandemic, they reported less substance use when compared with their US-born counterparts [24,26,27]. Our findings parallel what the limited studies have found, yet we did find that changes in decreased alcohol use were higher in foreign-born individuals compared with US-born individuals, that is, from once to several times per month to not at all during the pandemic.

Nevertheless, decreased alcohol use may be observed in our sample owing to a wide range of socioeconomic status factors such as low income and unemployment and environmental factors such as access and scarcity [32]. This may also be an artifact of our oversampling of lower socioeconomic status groups. Regardless, more information is needed to assess the deleterious effects on physical and mental health based on the frequency and number of drinks consumed as well as the level of risk of developing possible alcohol use disorders.

Our findings at this stage may indicate small percentage changes but may reveal the ideal point of intervention to mitigate the effect of use disorders. Moreover, we must acknowledge the synergetic effects of mental health and physiological health, especially in the context of COVID-19. As such, the extremes of general alcohol consumption and use disorders must continue to be monitored not only for concomitant health effects, such as alcohol-associated liver disease and mortality [28,29], but also to better understand the social and environmental effects of the COVID-19 pandemic and the policies that affect changes in alcohol use [33]. Future research should focus on polysubstance, mental health, coping behaviors, and prior substance use to identify the groups most at risk and design the most appropriate intervention strategies.

Limitations

Our study had some limitations. First, although the survey was anonymous, the possibility of bias in recall and responses must be considered. There is a possible recall bias from self-reports, as we asked about behaviors before COVID-19 owing to the data collection starting in May 2021 and ending in January 2022. This may be magnified by asking sensitive questions concerning substance use behaviors before and during the pandemic, increasing the response bias. Second, the data were obtained from a cross-sectional survey that provided a descriptive analysis of substance use change in a large US sample. However, this sample is not representative of the United States. Moreover, because this was a cross-sectional sample, we could not discern causality or temporal directionality. The data and descriptive analysis provided a solid foundation for further examination and identification of substance use patterns across a diverse US sample.

Conclusions

Pandemics are predicted to increase in frequency in the near future. To better prepare for the indirect effects of public health policies meant to protect the health of individuals, we must also prepare for their indirect effects on mental health and related coping mechanisms. Substance use affects both physical and mental health and will therefore require a multimodal approach to efficiently and effectively address and intervene on the deleterious effects, especially for underserved and underrepresented communities.

Acknowledgments

The efforts of KV, FAMI, and FW were supported by the Division of Intramural Research, National Institute on Minority Health and Health Disparities, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily reflect the views of the National Institutes of Health.

Authors' Contributions

FAMI and FW were responsible for the concept and design. FAMI and KV were responsible for the acquisition, analysis, or interpretation of data and drafting of the manuscript. FAMI, KV, and FW were responsible for critical revision of the manuscript for important intellectual content. FAMI was responsible for statistical analysis. FW was responsible for providing administrative, technical, or material support and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Full study comparisons of substance use prior to and during the pandemic.
[DOCX File , 20 KB - publichealth_v9i1e38163_app1.docx]

References

1. COVID-19 Dashboard. Johns Hopkins University. URL: https://coronavirus.jhu.edu/map.html [accessed 2022-04-04]
2. Islam N, Lacey B, Shabnam S, Erzurumluoglu AM, Dambha-Miller H, Chowell G, et al. Social inequality and the syndemic of chronic disease and COVID-19: county-level analysis in the USA. J Epidemiol Community Health 2021 Jan 05;75(6):496-500. [doi: 10.1136/jech-2020-215626] [Medline: 33402397]
3. Rod MH, Hulvej Rod N. Towards a syndemic public health response to COVID-19. Scand J Public Health 2021 Feb 07;49(1):14-16. [doi: 10.1177/1403494820982862] [Medline: 33407002]
4. Xie Y, Xu E, Al-Aly Z. Risks of mental health outcomes in people with covid-19: cohort study. BMJ 2022 Feb 16;376:e068993 [FREE Full text] [doi: 10.1136/bmj-2021-068993] [Medline: 35172971]
5. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
6. Czeisler M, Board A, Thierry JM, Czeisler CA, Rajaratnam SM, Howard ME, et al. Mental health and substance use among adults with disabilities during the COVID-19 pandemic - United States, February-March 2021. MMWR Morb Mortal Wkly Rep 2021 Aug 27;70(34):1142-1149 [FREE Full text] [doi: 10.15585/mmwr.mm7034a3] [Medline: 34437518]
7. Willis T, Hirky A. Coping and substance abuse: a theoretical model and review of the evidence. In: Handbook of coping: Theory, Research, Applications. Oxford, England: John Wiley & Sons; 1996.
8. Romano I, Patte KA, de Groh M, Jiang Y, Wade TJ, Bélanger RE, et al. Substance-related coping behaviours among youth during the early months of the COVID-19 pandemic. Addict Behav Rep 2021 Dec;14:100392 [FREE Full text] [doi: 10.1016/j.abrep.2021.100392] [Medline: 34754898]
9. Greenglass E, Chiacchia D, Fiskenbaum L. Investigating COVID-19 stress and coping: substance use and behavioural disengagement. Int J Psychol 2022 Jun 18;57(3):325-335 [FREE Full text] [doi: 10.1002/ijop.12820] [Medline: 34792195]
10. Chacon NC, Walia N, Allen A, Sciancalepore A, Tiong J, Quick R, et al. Substance use during COVID-19 pandemic: impact on the underserved communities. Discoveries (Craiova) 2021 Dec 31;9(4):e141 [FREE Full text] [doi: 10.15190/d.2021.20] [Medline: 35261922]
11. Okafor CN, Bautista KJ, Asare M, Opara I. Coping in the time of COVID-19: buffering stressors with coping strategies. J Loss Trauma 2022 Apr 30;27(1):83-91 [FREE Full text] [doi: 10.1080/15325024.2021.1914987] [Medline: 35173557]
12. Wang QQ, Kaelber DC, Xu R, Volkow ND. COVID-19 risk and outcomes in patients with substance use disorders: analyses from electronic health records in the United States. Mol Psychiatry 2021 Jan 14;26(1):30-39 [FREE Full text] [doi: 10.1038/s41380-020-00880-7] [Medline: 32929211]
13. Patrick ME, Parks MJ, Fairlie AM, Kreski NT, Keyes KM, Miech R. Using substances to cope with the COVID-19 pandemic: U.S. National Data at Age 19 Years. J Adolesc Health 2022 Feb;70(2):340-344 [FREE Full text] [doi: 10.1016/j.jadohealth.2021.11.006] [Medline: 34916126]
14. Hanson JD, Noonan C, Harris A, Oziel K, Sarche M, MacLehose RF, et al. Alcohol consumption during COVID among women with an existing alcohol-use disorder. Int J Environ Res Public Health 2021 Sep 08;18(18):9460 [FREE Full text] [doi: 10.3390/ijerph18189460] [Medline: 34574387]
15. Bolinski RS, Walters S, Salisbury-Afshar E, Ouellet LJ, Jenkins WD, Almirol E, et al. The impact of the COVID-19 pandemic on drug use behaviors, fentanyl exposure, and harm reduction service support among people who use drugs in rural settings. Int J Environ Res Public Health 2022 Feb 16;19(4):2230 [FREE Full text] [doi: 10.3390/ijerph19042230] [Medline: 35206421]
16. Giovenco DP, Spillane TE, Maggi RM, Lee EY, Philbin MM. Multi-level drivers of tobacco use and purchasing behaviors during COVID-19 "lockdown": a qualitative study in the United States. Int J Drug Policy 2021 Aug;94:103175 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103175] [Medline: 33662811]
17. Gendall P, Hoek J, Stanley J, Jenkins M, Every-Palmer S. Changes in tobacco use during the 2020 COVID-19 lockdown in New Zealand. Nicotine Tob Res 2021 May 04;23(5):866-871 [FREE Full text] [doi: 10.1093/ntr/ntaa257] [Medline: 33515223]
18. Rigotti NA, Chang Y, Regan S, Lee S, Kelley JH, Davis E, et al. Cigarette smoking and risk perceptions during the COVID-19 pandemic reported by recently hospitalized participants in a smoking cessation trial. J Gen Intern Med 2021 Dec 07;36(12):3786-3793 [FREE Full text] [doi: 10.1007/s11606-021-06913-3] [Medline: 34100230]
19. Brenneke SG, Nordeck CD, Riehm KE, Schmid I, Tormohlen KN, Smail EJ, et al. Trends in cannabis use among U.S. adults amid the COVID-19 pandemic. Int J Drug Policy 2022 Feb;100:103517 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103517] [Medline: 34894469]
20. Nguyen N, Mathur Gaiha S, Halpern-Felsher B. Self-reported changes in cannabis vaping among US adolescents and young adults early in the COVID-19 pandemic. Prev Med Rep 2021 Dec;24:101654 [FREE Full text] [doi: 10.1016/j.pmedr.2021.101654] [Medline: 34976701]

21. Kalkhoran SM, Levy DE, Rigotti NA. Smoking and e-cigarette use among U.S. adults during the COVID-19 pandemic. Am J Prev Med 2022 Mar;62(3):341-349 [FREE Full text] [doi: 10.1016/j.amepre.2021.08.018] [Medline: 34756629]
22. Dunlop A, Lokuge B, Masters D, Sequeira M, Saul P, Dunlop G, et al. Challenges in maintaining treatment services for people who use drugs during the COVID-19 pandemic. Harm Reduct J 2020 May 06;17(1):26 [FREE Full text] [doi: 10.1186/s12954-020-00370-7] [Medline: 32375887]
23. Penner F, Rajesh A, Kinney KL, Mabus KL, Barajas KG, McKenna KR, et al. Racial and demographic disparities in emergency department utilization for mental health concerns before and during the COVID-19 pandemic. Psychiatry Res 2022 Apr;310:114442 [FREE Full text] [doi: 10.1016/j.psychres.2022.114442] [Medline: 35219262]
24. Lou NM, Noels KA, Zhang YS, Kurl S. Ethnic minority, immigrants, and Indigenous people's well-being disparities in Canada during the COVID-19 pandemic: the mediating role of threat perceptions. Int J Intercult Relat 2022 May;88:148-156 [FREE Full text] [doi: 10.1016/j.ijintrel.2022.04.006] [Medline: 35475126]
25. Barajas-Gonzalez RG, Brooks-Gunn J. Substance use differences among U.S.- versus foreign-born adolescents. Hispanic J Behav Sci 2014 Oct 09;36(4):506-521. [doi: 10.1177/0739986314551939]
26. Roe L, Proudfoot J, Tay Wee Teck J, Irvine RD, Frankland S, Baldacchino AM. Isolation, solitude and social distancing for people who use drugs: an ethnographic perspective. Front Psychiatry 2020 Jan 13;11:623032 [FREE Full text] [doi: 10.3389/fpsyt.2020.623032] [Medline: 33519561]
27. McKnight-Eily LR, Okoro CA, Strine TW, Verlenden J, Hollis ND, Njai R, et al. Racial and ethnic disparities in the prevalence of stress and worry, mental health conditions, and increased substance use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2021 Feb 05;70(5):162-166 [FREE Full text] [doi: 10.15585/mmwr.mm7005a3] [Medline: 33539336]
28. Bhalla S, Sharma B, Smith D, Boley R, McCluskey C, Ilyas Y, et al. Investigating unhealthy alcohol use as an independent risk factor for increased COVID-19 disease severity: observational cross-sectional study. JMIR Public Health Surveill 2021 Nov 05;7(11):e33022 [FREE Full text] [doi: 10.2196/33022] [Medline: 34665758]
29. Julien J, Ayer T, Tapper EB, Barbosa C, Dowd WN, Chhatwal J. Effect of increased alcohol consumption during COVID-19 pandemic on alcohol-associated liver disease: a modeling study. Hepatology 2022 Jun 24;75(6):1480-1490. [doi: 10.1002/hep.32272] [Medline: 34878683]
30. Lechner WV, Laurene KR, Patel S, Anderson M, Grega C, Kenne DR. Changes in alcohol use as a function of psychological distress and social support following COVID-19 related University closings. Addict Behav 2020 Nov;110:106527 [FREE Full text] [doi: 10.1016/j.addbeh.2020.106527] [Medline: 32679435]
31. Boschuetz N, Cheng S, Mei L, Loy VM. Changes in alcohol use patterns in the United States during COVID-19 pandemic. WMJ 2020 Sep;119(3):171-176 [FREE Full text] [Medline: 33091284]
32. Grossman ER, Benjamin-Neelon SE, Sonnenschein S. Alcohol consumption during the COVID-19 pandemic: a cross-sectional survey of US adults. Int J Environ Res Public Health 2020 Dec 09;17(24):9189 [FREE Full text] [doi: 10.3390/ijerph17249189] [Medline: 33316978]
33. Rehm J, Kilian C, Ferreira-Borges C, Jernigan D, Monteiro M, Parry CD, et al. Alcohol use in times of the COVID 19: implications for monitoring and policy. Drug Alcohol Rev 2020 May 02;39(4):301-304 [FREE Full text] [doi: 10.1111/dar.13074] [Medline: 32358884]

Abbreviations

NIH: National Institutes of Health

Edited by A Mavragani, G Eysenbach; submitted 21.03.22; peer-reviewed by D Smith, R De Boni; comments to author 25.07.22; revised version received 17.08.22; accepted 18.10.22; published 17.02.23.

Please cite as:
Montiel Ishino FA, Villalobos K, Williams F
Substance Use From Social Distancing and Isolation by US Nativity During the Time of COVID-19: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e38163
URL: https://publichealth.jmir.org/2023/1/e38163
doi:10.2196/38163
PMID:36265162

Original Paper

Changes in Mobile Health Apps Usage Before and After the COVID-19 Outbreak in China: Semilongitudinal Survey

Le Yang[1*], MD; Jiadong Wu[1*], BSc; Xiaoxiao Mo[1*], BSc; Yaqin Chen[2], BSc; Shanshan Huang[2], BSc; Linlin Zhou[1], BSc; Jiaqi Dai[1], BSc; Linna Xie[1], BSc; Siyu Chen[1], BSc; Hao Shang[1], BSc; Beibei Rao[1], BSc; Bingtao Weng[1], BSc; Ayiguli Abulimiti[1], BSc; Siying Wu[1], MD; Xiaoxu Xie[1], MD

[1]School of Public Health, Fujian Medical University, Fuzhou, China
[2]School of Nursing, Fujian Medical University, Fuzhou, China
[*]these authors contributed equally

Corresponding Author:
Xiaoxu Xie, MD
School of Public Health
Fujian Medical University
No. 1 Xuefu North Road
Fuzhou University New District
Fuzhou, 350122
China
Phone: 86 059122862023
Email: xiexiaoxu@aliyun.com

Abstract

Background: Mobile health (mHealth) apps are rapidly emerging technologies in China due to strictly controlled medical needs during the COVID-19 pandemic while continuing essential services for chronic diseases. However, there have been no large-scale, systematic efforts to evaluate relevant apps.

Objective: We aim to provide a landscape of mHealth apps in China by describing and comparing digital health concerns before and after the COVID-19 outbreak, including mHealth app data flow and user experience, and analyze the impact of COVID-19 on mHealth apps.

Methods: We conducted a semilongitudinal survey of 1593 mHealth apps to study the app data flow and clarify usage changes and influencing factors. We selected mHealth apps in app markets, web pages from the Baidu search engine, the 2018 top 100 hospitals with internet hospitals, and online shopping sites with apps that connect to smart devices. For user experience, we recruited residents from a community in southeastern China from October 2019 to November 2019 (before the outbreak) and from June 2020 to August 2020 (after the outbreak) comparing the attention of the population to apps. We also examined associations between app characteristics, functions, and outcomes at specific quantiles of distribution in download changes using quantile regression models.

Results: Rehabilitation medical support was the top-ranked functionality, with a median 1.44 million downloads per app prepandemic and a median 2.74 million downloads per app postpandemic. Among the top 10 functions postpandemic, 4 were related to maternal and child health: pregnancy preparation (ranked second; fold change 4.13), women's health (ranked fifth; fold change 5.16), pregnancy (ranked sixth; fold change 5.78), and parenting (ranked tenth; fold change 4.03). Quantile regression models showed that rehabilitation (P_{75}, P_{90}), pregnancy preparation (P_{90}), bodybuilding (P_{50}, P_{90}), and vaccination (P_{75}) were positively associated with an increase in downloads after the outbreak. In the user experience survey, the attention given to health information (prepandemic: 249/375, 66.4%; postpandemic: 146/178, 82.0%; P=.006) steadily increased after the outbreak.

Conclusions: mHealth apps are an effective health care approach gaining in popularity among the Chinese population following the COVID-19 outbreak. This research provides direction for subsequent mHealth app development and promotion in the postepidemic era, supporting medical model reformation in China as a reference, which may provide new avenues for designing and evaluating indirect public health interventions such as health education and health promotion.

(JMIR Public Health Surveill 2023;9:e40552) doi:10.2196/40552

KEYWORDS

application; China; COVID-19; mHealth; health management; mobile health; technology; app; survey; data; user; user experience; vaccination; download; healthcare; development

Introduction

Background

In the wake of the COVID-19 outbreak, digital health technologies to assist medical service systems and people [1], such as telemedicine, have moved from a convenience to a demand [2-5]. Mobile health (mHealth) apps are a novel platform that uses mobile devices to acquire data across wellness and disease diagnosis, prevention, and management [6,7].

Before the COVID-19 outbreak, several studies investigated the characteristics of apps in China; these studies focused on only specific health domains, such as disease management, women's health, and sports, instead of elaborating on the overall mHealth app situation [8-11]. However, in the context of COVID-19, the development of mHealth apps has become a hot topic [12]. Only a few studies have investigated the mHealth app market status, focusing on the assessment of functional distribution but without refined classifications of mHealth apps and lacking information integrity [13]. Therefore, the classifications for mHealth apps have not been elaborated, and detailed descriptive research on mHealth apps is lacking.

After the outbreak, apps directly related to COVID-19 that were used to track high-risk groups and assist in diagnosis were the most studied [14-16]. One study provided an overview and classifications of mHealth apps currently available on the market to combat COVID-19, based on differences in basic functions and purpose [17]. However, in the face of the significant challenges posed by the pandemic, app types other than those related to COVID-19 were also considered valuable tools, such as easing the burden on hospitals, providing access to reliable information, tracking individuals' symptoms and mental health, and discovering new predictors [18]. Previous studies have shown that mHealth apps can improve people's life needs (such as fertility), which promote their use [19]. However, the development trends for many types of mHealth apps have been complicated due to the pandemic, which requires attention.

At the same time, unlike other countries, China implemented normalized pandemic prevention and control and rarely used contact tracing apps in a personal form; thus, the apps developed directly as a result of the pandemic were not the focus of this research. There was a spike in the volume of phone calls asking medical questions after the outbreak with a great increase in online demand in China [17]. The unbalanced allocation of health resources between the east and west, due to the vast size of China, also led to the demand for telemedicine. During the COVID-19 pandemic, a multimodal telemedicine network in Sichuan Province in Western China was activated immediately, which was demonstrated to be feasible, acceptable, and effective [20]. Moreover, due to the unique national conditions of the 1-child policy and the aging population, changes in mHealth app fields before and after the outbreak are different from those in other countries [21,22]. Therefore, mHealth apps, as emerging tools, that focus on changes in various health functions before and after the outbreak are worth studying in China and could be important for prevention, diagnosis, treatment, and management decisions for future public health emergencies.

However, it is presently unclear what mHealth app characteristics, if any, have been influenced and appropriately deployed in the pre and postpandemic periods.

Objectives

Here, we conducted a nationwide study of mHealth apps in China. The aim of this study was (1) to describe and compare digital health concerns before and after the COVID-19 outbreak, including mHealth app data flow and user experience, and (2) to analyze the impact of COVID-19 on mHealth apps.

Methods

Data Acquisition

Before the COVID-19 outbreak, we conducted a comprehensive electronic search of 4 sources up to October 25, 2019: (1) apps on leaderboards of health-related categories in the 6 largest app markets in China, including the top 50 on the Huawei Android Market (Huawei Technologies Co Ltd; Shenzhen, China), top 50 on the OPPO Android Market (Oppo Electronics Co Ltd; Dongguan, Guangdong, China), top 100 on the Vivo Android Market (Vivo Co Ltd; Dongguan, Guangdong, China), top 50 on the Tencent Android Market (Tencent Holdings Limited; Shenzhen, China), top 100 on the 360 Android Market (Qihoo 360 Technology Co Ltd; Beijing, China), and top 100 on the Apple iTunes store for China (Apple; Cupertino, CA) [23]; (2) the mHealth apps on the first 20 web pages of the Baidu search engine, which is the largest search engine in China; (3) apps that can connect to smart devices from the 4 large online shopping sites (Tmall, JD, Pinduoduo, and Suning); and (4) apps affiliated with internet hospitals on the list of the top 100 Chinese hospitals in 2018. The exclusion criteria were (1) duplicated apps, (2) app descriptions irrelevant to health, (3) apps not available in the Chinese language, (4) apps that were not available for download through the official Android and Apple app stores or the Baidu search engine, and (5) apps that could not be opened or used due to technical problems. Apps with patient and clinician versions were evaluated as different items. Search terms, sample quantities, and data collection times for each source are provided in Table S1 in Multimedia Appendix 1.

We collected 4 types of data: (1) basic app characteristics from the description interface, including the size of apps, number of app downloads, and target users; (2) app developers' information from the largest commercial inquiry platform, the Tianyancha website [24], including transaction amount, registered capital, number of staff, operating status, establishment date, and geographic location; (3) app permission listing data from the permission interface; and (4) app functions from app trials, except for apps only open to internal users. As the iOS App Store does not display the number of downloads, we replaced total downloads in iOS with the number of reviews and used statements of "downloads" in the following paragraphs. If an app existed on multiple platforms, the total number of app downloads was calculated as the sum of app downloads from all platforms. The app trial included at least two investigators who downloaded selected applications and independently identified application functionality according to a clear functional definition by using iPhones and Android phones. See Table S2 in Multimedia Appendix 1 for details on function definitions.

Following the outbreak, the same mHealth apps were investigated a second time in April 2021 as semilongitudinal samples with download data, and we determined whether COVID-19 content had been added. A download change was defined as the difference between post and prepandemic downloads.

User Experience Survey

We recruited residents through a community health checkup program to explore the user experience with mHealth apps among a large community of more than 20,000 people with a balanced age distribution in southeastern China. We used an offline questionnaire to survey 400 participants from October 2019 to November 2019 before the COVID-19 outbreak and 200 participants from June 2020 to August 2020 after the outbreak. A total of 553 (553/600, 92.2%) participants completed the survey: 375 (375/400, 93.8%) before the outbreak and 178 (178/200, 89.0%) after the outbreak. A predesigned, structured questionnaire was provided to potential participants in the waiting areas of the medical examination center in this community. The questionnaire was designed to collect information on participants' attention to health information, various aspects of mHealth technology usage, willingness to consume mHealth technology, and health status and demographic characteristics. Trained research assistants who were fluent in Chinese administered the questionnaire and provided verbal instructions about how to complete them.

Ethical Considerations

Before taking the survey, informed consent was obtained from each participant. This study was approved by the Biomedical Research Ethics Committee of Fujian Medical University (2018 number 11). All procedures were performed using the relevant guidelines and regulations.

Data Analysis

A descriptive analysis was conducted for mHealth app characteristics, developers, permission, functions, and user experiences in China. Data are presented using frequencies and percentages, bar charts, statistical maps, Venn diagrams, and heat maps. Continuous variables are presented as the mean and SD or the median and IQR, while categorical variables are presented as the frequency and percentage. Mann-Whitney *U* tests or chi-square tests were used to assess differences among variables.

We also compared the post and prepandemic app downloads of each category using paired *t* tests. To account for multiple comparisons, we calculated a Bonferroni-corrected *P* value criterion of .05/28. Therefore, $P<1.79\times10^{-3}$ was considered statistically significant.

Quantile regression (QR) models were used to explore the relationship between modeling covariates and quartiles of the outcome variable of interest [25]. Because of the wide range and non-normal distribution of download changes, we used QR models to examine associations between app characteristics, functions, and outcomes at specific quantiles of distribution in download changes. This analysis does not make assumptions about the residual distribution and is more robust to outliers in the outcome [26]. Unlike an ordinary linear regression model, which models the mean of only 1 dependent variable, QR examines the effect of covariates at different points of the conditional distribution of the response variable and gives more comprehensive results. QR models provided a more detailed view of associations of app characteristics, functions, and outcomes with download changes. We obtained estimates and plotted them at the following quantiles: 10th, 25th, 50th (median), 75th, and 90th. In the QR analysis, significance was assessed at the 10% level.

All analyses were prespecified and performed using SPSS 25.0 (IBM Corp; Armonk, NY) and Stata version 13 (StataCorp; College Station, TX).

Results

App Characteristics

A total of 1593 mHealth apps were included in the analysis (see Figure 1). During the COVID-19 pandemic, various app downloads showed an overall upward trend (change in median: 61,561), and approximately 10% (182/1593, 11.4%) of the apps had functions or content added for COVID-19 (see Table 1). Figures 2 and 3 show the target users and geographic distribution of the apps. Target users included healthy people (921/1593, 57.8%), patients (513/1593, 32.2%), and health care professionals (393/1593, 24.7%). The geographical distribution of mHealth app developers included in the study was concentrated in megacities (ie, Beijing, Shanghai, Guangzhou) and southeast China (coastal areas with a developed economy; see Figure 3). Additional permission requests were found in app markets (see Table S3 in Multimedia Appendix 1).

Figure 1. Flow diagram of app selection.

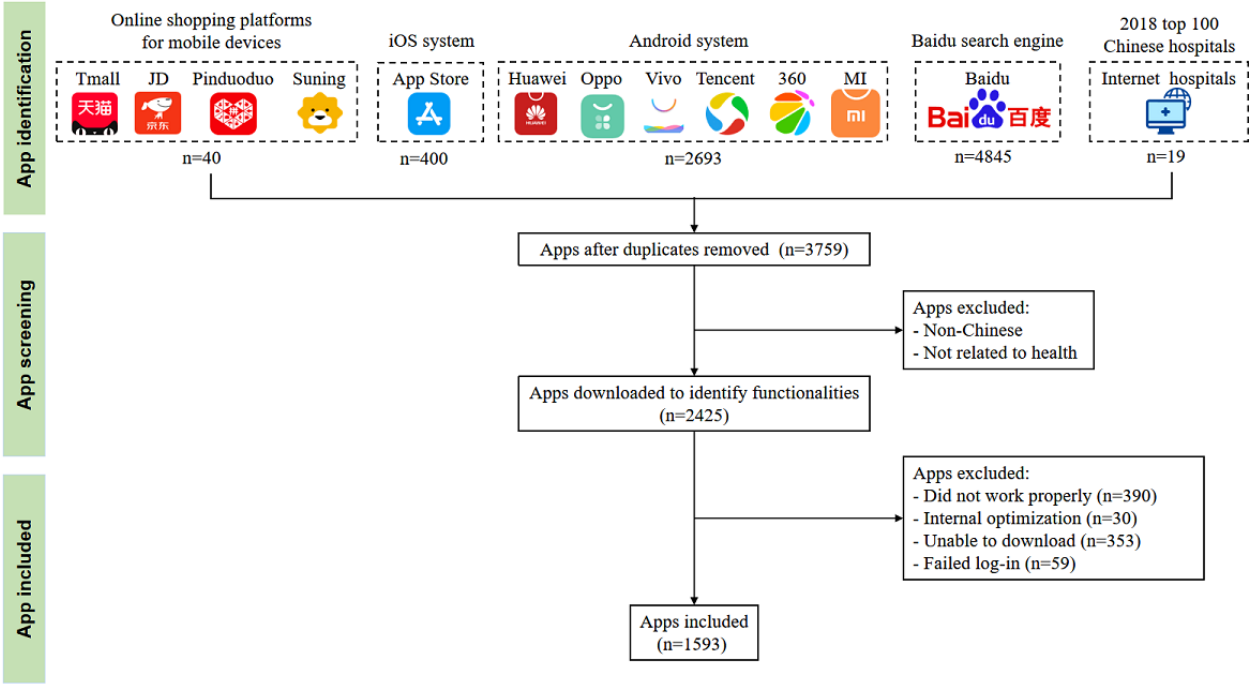


Table 1. Characteristics and developers of mobile health (mHealth) apps available on the Chinese market (1593 apps and 1196 developers).

Characteristics	Results
Basic information	
Size of app (MB), mean (SD)	37.26 (75.46)
Size of app (MB), median (IQR)	24.90 (13.50-39.88)
Download price, n (%)	
Free download	1530 (96.0)
Paid	63 (4.0)
If download is free, in-app purchase available, n (%)	1444 (90.6)
In-app advertisement, n (%)	175 (11.0)
Target users, n (%)	
Medical researchers	72 (4.5)
Medical personnel	100 (6.3)
Patients	513 (32.2)
Healthy people	921 (57.8)
Has a membership system, n (%)	101 (6.3)
Connects to smart devices, n (%)	292 (18.3)
User rating score, mean (SD)	6.99 (2.33)
User rating score, median (IQR)	7.00 (6.00-9.20)
Total downloads (x10^3)	
Prepandemic, mean (SD)	8967.33 (118,905.97)
Prepandemic, median (IQR)	44.09 (1.47-787.00)
Postpandemic, mean (SD)	17,342.13 (275,177.12)
Postpandemic, median (IQR)	223.12 (10.93-2252.00)
Number of functions, mean (SD)	2.72 (2.25)
Number of functions, median (IQR)	2 (1-4)
Added functions or content for the pandemic, n (%)	182 (11.4)
Type of function or content added for the pandemic, n (%)	
Pandemic consultation	105 (6.6)
Pandemic prevention knowledge	108 (6.8)
Pandemic dynamics	9 (0.6)
App developers	
Transaction amount (¥; x10^6)[a], mean (SD)	267.66 (812.75)
Transaction amount (¥; x10^6)[a], median (IQR)	50.00 (20.50-150.00)
Registered capital (¥; x10^6)[a], mean (SD)	188.10 (204.15)
Registered capital (¥; x10^6)[a], median (IQR)	937.50 (166.00-210.53)
Number of staff, mean (SD)	262 (3016)
Number of staff, median (IQR)	17 (2-62)
Operating status, n (%)	
In business	482 (30.3)
Remainder enterprise	697 (43.8)
Closed down	15 (1.0)

[a] A currency exchange rate of ¥1=US $0.15 is applicable.

Figure 2. Sankey diagram of flow direction in apps. The width of the colored boxes and their connecting gray bands are directly proportional to the frequency of apps from every data source (left side) and flow quantities of these apps to the attributable user communities (right side). Hcps: health care professionals.

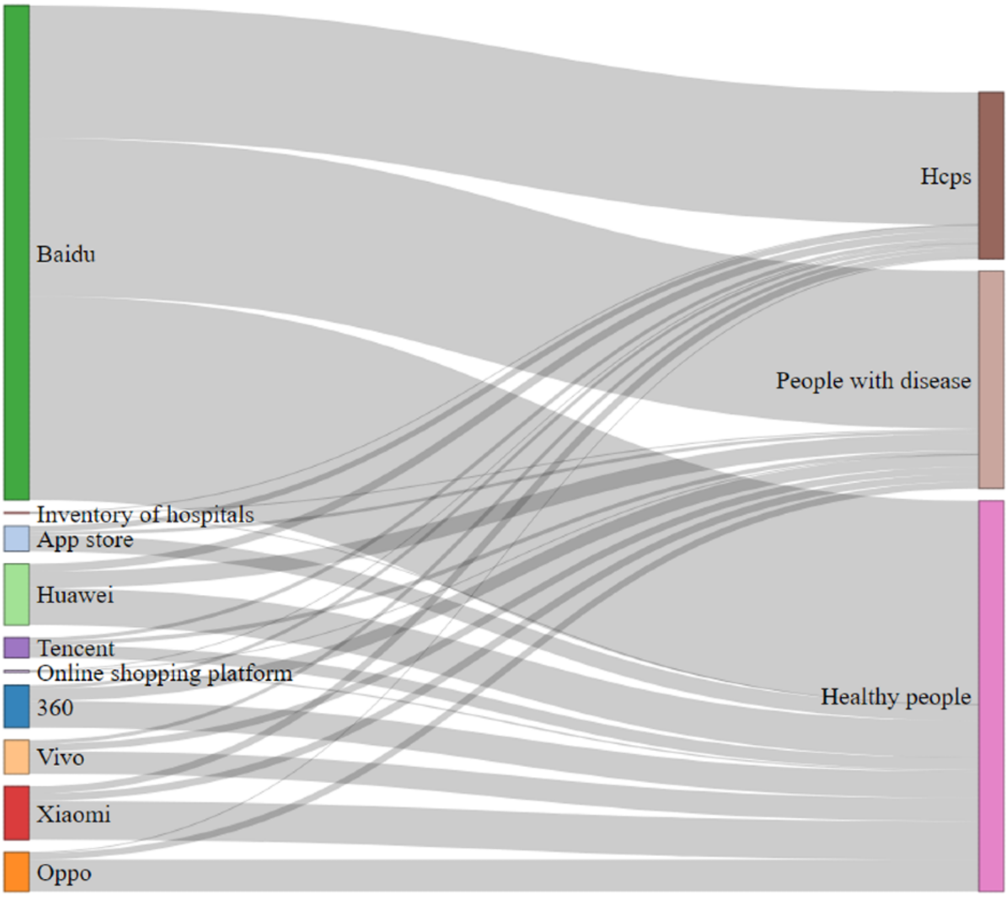
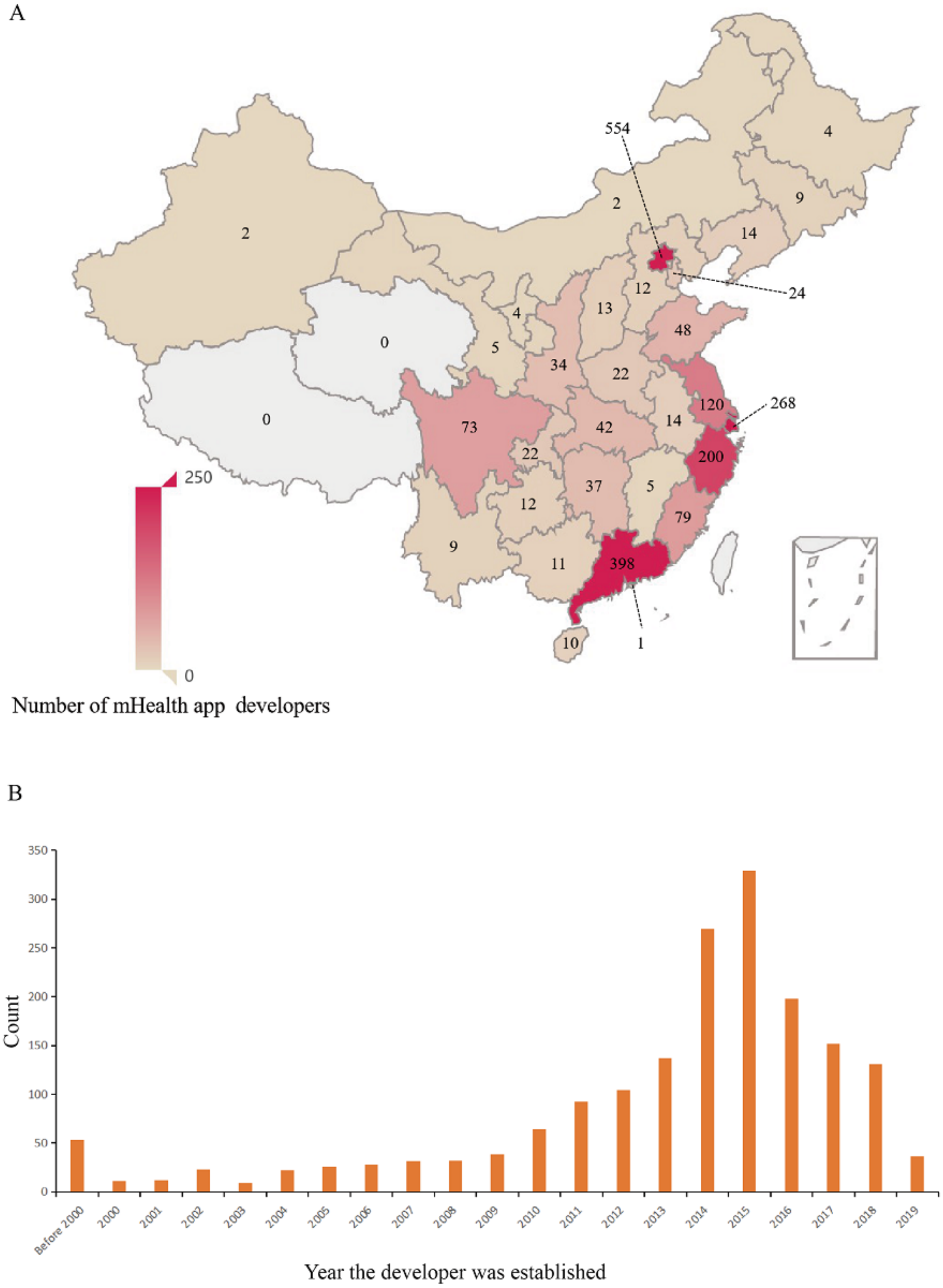
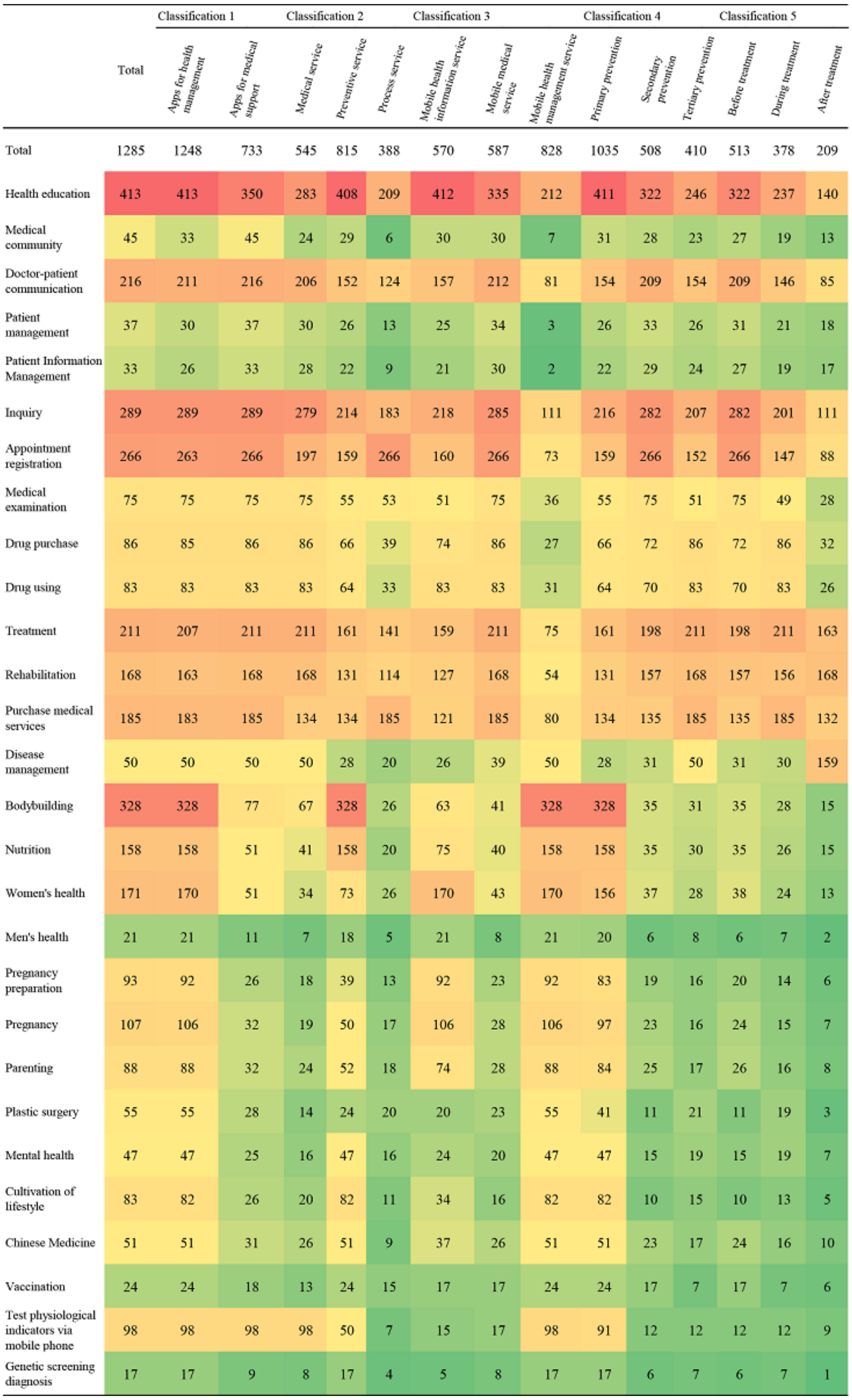


Figure 3. (A) Geographical distribution and (B) distribution of the time of establishment of mobile health (mHealth) app developers in China. There is another developer in Canada.



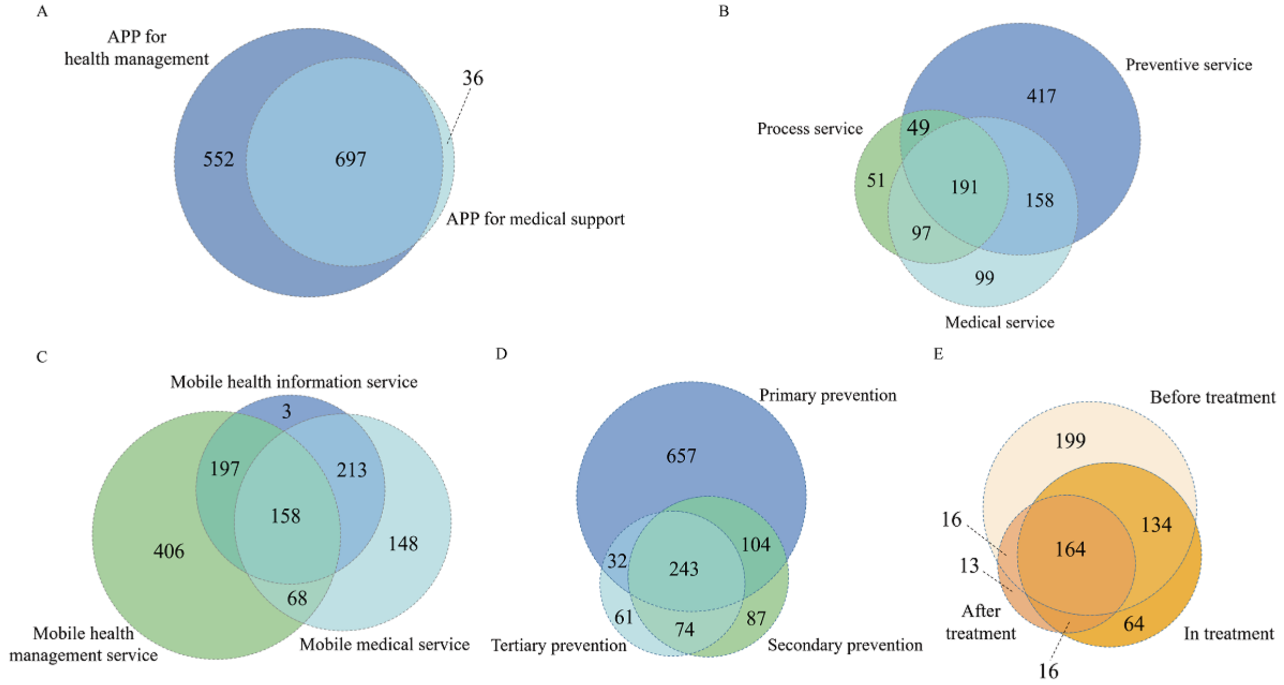
Of the 1593 mHealth apps, 1285 (80.7%) apps had full functionality available to conduct the app trial, including apps for health management (1248/1593, 78.3%) and apps for medical support (697/1593, 43.8%). The frequency of each function available in Chinese mHealth apps is shown in Figure 4. Figure 5 shows the associations between the 5 app classifications, with the circles scaled to the number of apps in each app classification. The rating concepts for these 5 different classifications in app trials are available in Table S4 in Multimedia Appendix 1 [27].

Figure 4. For each classification of mobile health (mHealth) apps in China: (A) frequency of app functions for 5 classifications; (B) ranking of the frequency of app functions is displayed on a color scale ranging from green (lowest charge rates) to orange, yellow, and red (highest charge rates).



		Classification 1		Classification 2			Classification 3			Classification 4			Classification 5		
	Total	Apps for health management	Apps for medical support	Medical service	Preventive service	Process service	Mobile health information service	Mobile medical service	Mobile health management service	Primary prevention	Secondary prevention	Tertiary prevention	Before treatment	During treatment	After treatment
Total	1285	1248	733	545	815	388	570	587	828	1035	508	410	513	378	209
Health education	413	413	350	283	408	209	412	335	212	411	322	246	322	237	140
Medical community	45	33	45	24	29	6	30	30	7	31	28	23	27	19	13
Doctor-patient communication	216	211	216	206	152	124	157	212	81	154	209	154	209	146	85
Patient management	37	30	37	30	26	13	25	34	3	26	33	26	31	21	18
Patient Information Management	33	26	33	28	22	9	21	30	2	22	29	24	27	19	17
Inquiry	289	289	289	279	214	183	218	285	111	216	282	207	282	201	111
Appointment registration	266	263	266	197	159	266	160	266	73	159	266	152	266	147	88
Medical examination	75	75	75	75	55	53	51	75	36	55	75	51	75	49	28
Drug purchase	86	85	86	86	66	39	74	86	27	66	72	86	72	86	32
Drug using	83	83	83	83	64	33	83	83	31	64	70	83	70	83	26
Treatment	211	207	211	211	161	141	159	211	75	161	198	211	198	211	163
Rehabilitation	168	163	168	168	131	114	127	168	54	131	157	168	157	156	168
Purchase medical services	185	183	185	134	134	185	121	185	80	134	135	185	135	185	132
Disease management	50	50	50	50	28	20	26	39	50	28	31	50	31	30	159
Bodybuilding	328	328	77	67	328	26	63	41	328	328	35	31	35	28	15
Nutrition	158	158	51	41	158	20	75	40	158	158	35	30	35	26	15
Women's health	171	170	51	34	73	26	170	43	170	156	37	28	38	24	13
Men's health	21	21	11	7	18	5	21	8	21	20	6	8	6	7	2
Pregnancy preparation	93	92	26	18	39	13	92	23	92	83	19	16	20	14	6
Pregnancy	107	106	32	19	50	17	106	28	106	97	23	16	24	15	7
Parenting	88	88	32	24	52	18	74	28	88	84	25	17	26	16	8
Plastic surgery	55	55	28	14	24	20	20	23	55	41	11	21	11	19	3
Mental health	47	47	25	16	47	16	24	20	47	47	15	19	15	19	7
Cultivation of lifestyle	83	82	26	20	82	11	34	16	82	82	10	15	10	13	5
Chinese Medicine	51	51	31	26	51	9	37	26	51	51	23	17	24	16	10
Vaccination	24	24	18	13	24	15	17	17	24	24	17	7	17	7	6
Test physiological indicators via mobile phone	98	98	98	98	50	7	15	17	98	91	12	12	12	12	9
Genetic screening diagnosis	17	17	9	8	17	4	5	8	17	17	6	7	6	7	1

Figure 5. Venn diagrams illustrating the associations between app classifications: (A) user communities, (B) mobile health service function, (C) content or services provided by apps, (D) tertiary prevention, and (E) service time provided by apps.



Comparison of Download Changes in Apps With Different Functions Between Pre and Postpandemic

Overall upward trends in app downloads during the pandemic were driven by some key app functions (see Figures 6 and 7). Rehabilitation (prepandemic median 1,437,500; postpandemic median 2,741,890; *P*<.001) and pregnancy preparation (prepandemic median 480,520; postpandemic median 1,982,490; *P*<.001) were the most often used functions, with statistically significant differences between pre and postpandemic. The 9 most important drivers for increasing downloads were divided into 2 categories. Three functions related to children and maternal health were observed: women's health (fold change 5.16), pregnancy (fold change 5.78), and parenting (fold change 4.03). Increased downloads of these apps (women’s health, *P*<.001; pregnancy, *P*<.001; parenting, *P*<.001) were also observed between pre and postpandemic. The other 6 most popular functions regarding treatment needs and maintenance were treatment (fold change 5.31), plastic surgery (fold change 8.77), drug use (fold change 5.92), patient information management (fold change 5.06), nutrition (fold change 8.79), and Chinese medicine (fold change 10.04). Vaccination, which was most relevant to COVID-19, saw a large spike in download rates, with an increase of 6.26 times (absolute change in mean: 7,233,800) postpandemic compared with prepandemic.

The 6 functions that decreased in ranking the most in the postpandemic period were health education, genetic screening, medical service purchases, drug purchases, inquiries, and physiological testing via mobile phone; 4 additional functions that declined after the COVID-19 outbreak included medical examination, medical community, patient management, and disease management. Bodybuilding, which was closely related to outdoor activities, also declined in ranking, with an increase of only 2.71 times prepandemic rates, but its absolute change in mean was 28,247,800. All these functions with declining rankings were growing, albeit at lower speeds relative to high-ranking functions. For example, the increase in drug purchases was 2.05 times (absolute change in mean: 7,388,400) the prepandemic rate, and patient management downloads were 4.98 times (absolute change in mean: 3,039,900) the prepandemic rate.

Figure 6. The number of occurrences of each function (ie, function frequency) in the 1593 mobile health (mHealth) apps with multiple functions in China.

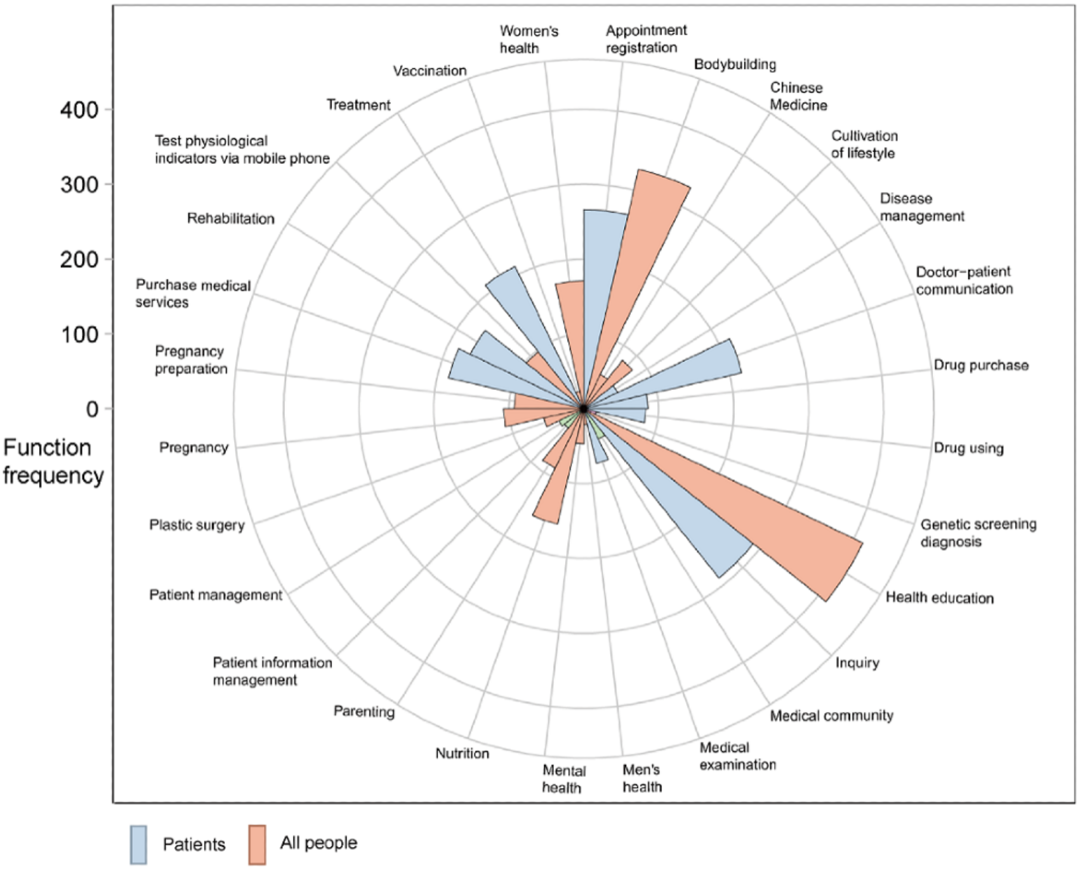
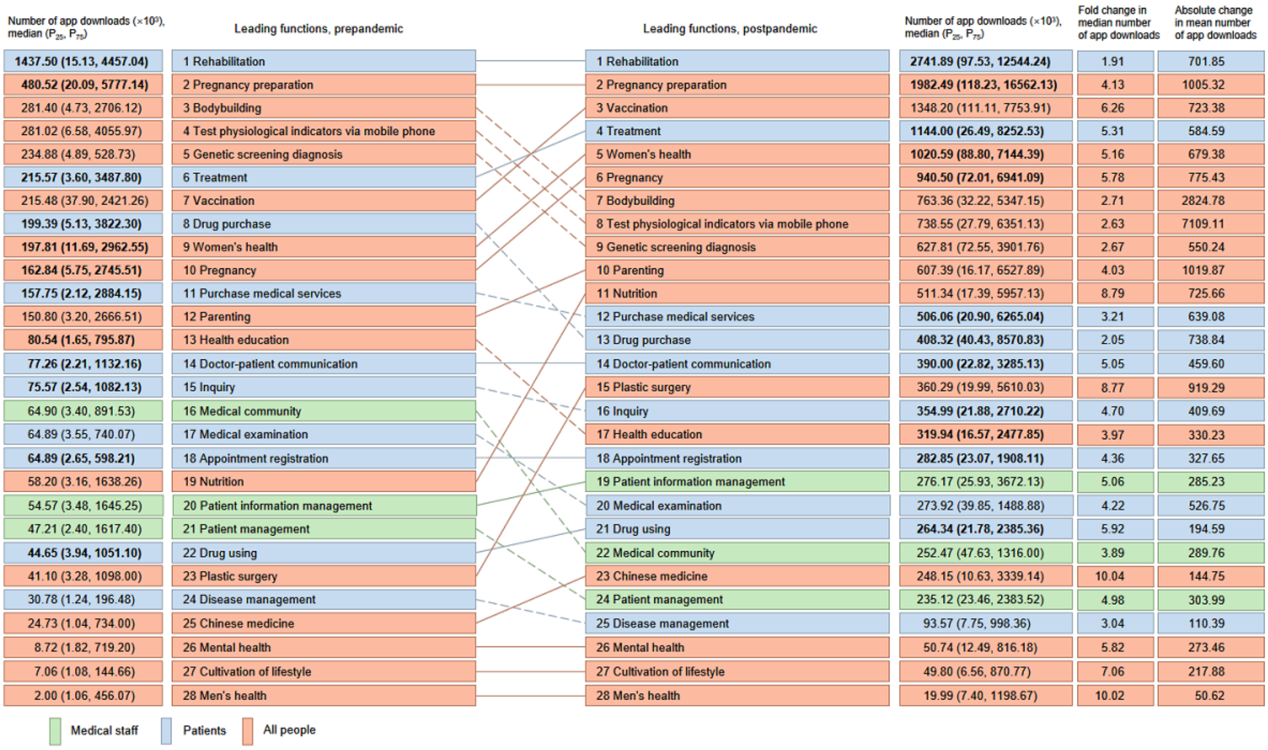


Figure 7. Leading functions of mobile health (mHealth) apps in the pre and postpandemic periods in China, connected by lines between the periods to show increased (solid line) or decreased (dashed line) ranking, while bold number indicate significant changes between the periods as determined using paired *t* tests with Bonferroni-corrected *P* values <1.79x10^{-3}. Fold change in median of the number of app downloads=(Medpost-Medpre)/Medpre. Absolute change in mean of the number of app downloads=∑(npost-npre)/n.



Number of app downloads (×10³), median (P_{25}, P_{75})	Leading functions, prepandemic	Leading functions, postpandemic	Number of app downloads (×10³), median (P_{25}, P_{75})	Fold change in median number of app downloads	Absolute change in mean number of app downloads
1437.50 (15.13, 4457.04)	1 Rehabilitation	1 Rehabilitation	**2741.89 (97.53, 12544.24)**	1.91	701.85
480.52 (20.09, 5777.14)	2 Pregnancy preparation	2 Pregnancy preparation	**1982.49 (118.23, 16562.13)**	4.13	1005.32
281.40 (4.73, 2706.12)	3 Bodybuilding	3 Vaccination	1348.20 (111.11, 7753.91)	6.26	723.38
281.02 (6.58, 4055.97)	4 Test physiological indicators via mobile phone	4 Treatment	**1144.00 (26.49, 8252.53)**	5.31	584.59
234.88 (4.89, 528.73)	5 Genetic screening diagnosis	5 Women's health	**1020.59 (88.80, 7144.39)**	5.16	679.38
215.57 (3.60, 3487.80)	6 Treatment	6 Pregnancy	**940.50 (72.01, 6941.09)**	5.78	775.43
215.48 (37.90, 2421.26)	7 Vaccination	7 Bodybuilding	763.36 (32.22, 5347.15)	2.71	2824.78
199.39 (5.13, 3822.30)	8 Drug purchase	8 Test physiological indicators via mobile phone	738.55 (27.79, 6351.13)	2.63	7109.11
197.81 (11.69, 2962.55)	9 Women's health	9 Genetic screening diagnosis	627.81 (72.55, 3901.76)	2.67	550.24
162.84 (5.75, 2745.51)	10 Pregnancy	10 Parenting	607.39 (16.17, 6527.89)	4.03	1019.87
157.75 (2.12, 2884.15)	11 Purchase medical services	11 Nutrition	511.34 (17.39, 5957.13)	8.79	725.66
150.80 (3.20, 2666.51)	12 Parenting	12 Purchase medical services	**506.06 (20.90, 6265.04)**	3.21	639.08
80.54 (1.65, 795.87)	13 Health education	13 Drug purchase	**408.32 (40.43, 8570.83)**	2.05	738.84
77.26 (2.21, 1132.16)	14 Doctor-patient communication	14 Doctor-patient communication	**390.00 (22.82, 3285.13)**	5.05	459.60
75.57 (2.54, 1082.13)	15 Inquiry	15 Plastic surgery	360.29 (19.99, 5610.03)	8.77	919.29
64.90 (3.40, 891.53)	16 Medical community	16 Inquiry	**354.99 (21.88, 2710.22)**	4.70	409.69
64.89 (3.55, 740.07)	17 Medical examination	17 Health education	**319.94 (16.57, 2477.85)**	3.97	330.23
64.89 (2.65, 598.21)	18 Appointment registration	18 Appointment registration	**282.85 (23.07, 1908.11)**	4.36	327.65
58.20 (3.16, 1638.26)	19 Nutrition	19 Patient information management	276.17 (25.93, 3672.13)	5.06	285.23
54.57 (3.48, 1645.25)	20 Patient information management	20 Medical examination	273.92 (39.85, 1488.88)	4.22	526.75
47.21 (2.40, 1617.40)	21 Patient management	21 Drug using	**264.34 (21.78, 2385.36)**	5.92	194.59
44.65 (3.94, 1051.10)	22 Drug using	22 Medical community	252.47 (47.63, 1316.00)	3.89	289.76
41.10 (3.28, 1098.00)	23 Plastic surgery	23 Chinese medicine	248.15 (10.63, 3339.14)	10.04	144.75
30.78 (1.24, 196.48)	24 Disease management	24 Patient management	235.12 (23.46, 2383.52)	4.98	303.99
24.73 (1.04, 734.00)	25 Chinese medicine	25 Disease management	93.57 (7.75, 998.36)	3.04	110.39
8.72 (1.82, 719.20)	26 Mental health	26 Mental health	50.74 (12.49, 816.18)	5.82	273.46
7.06 (1.08, 144.66)	27 Cultivation of lifestyle	27 Cultivation of lifestyle	49.80 (6.56, 870.77)	7.06	217.88
2.00 (1.06, 456.07)	28 Men's health	28 Men's health	19.99 (7.40, 1198.67)	10.02	50.62

Relationship Between App Characteristics and Download Changes

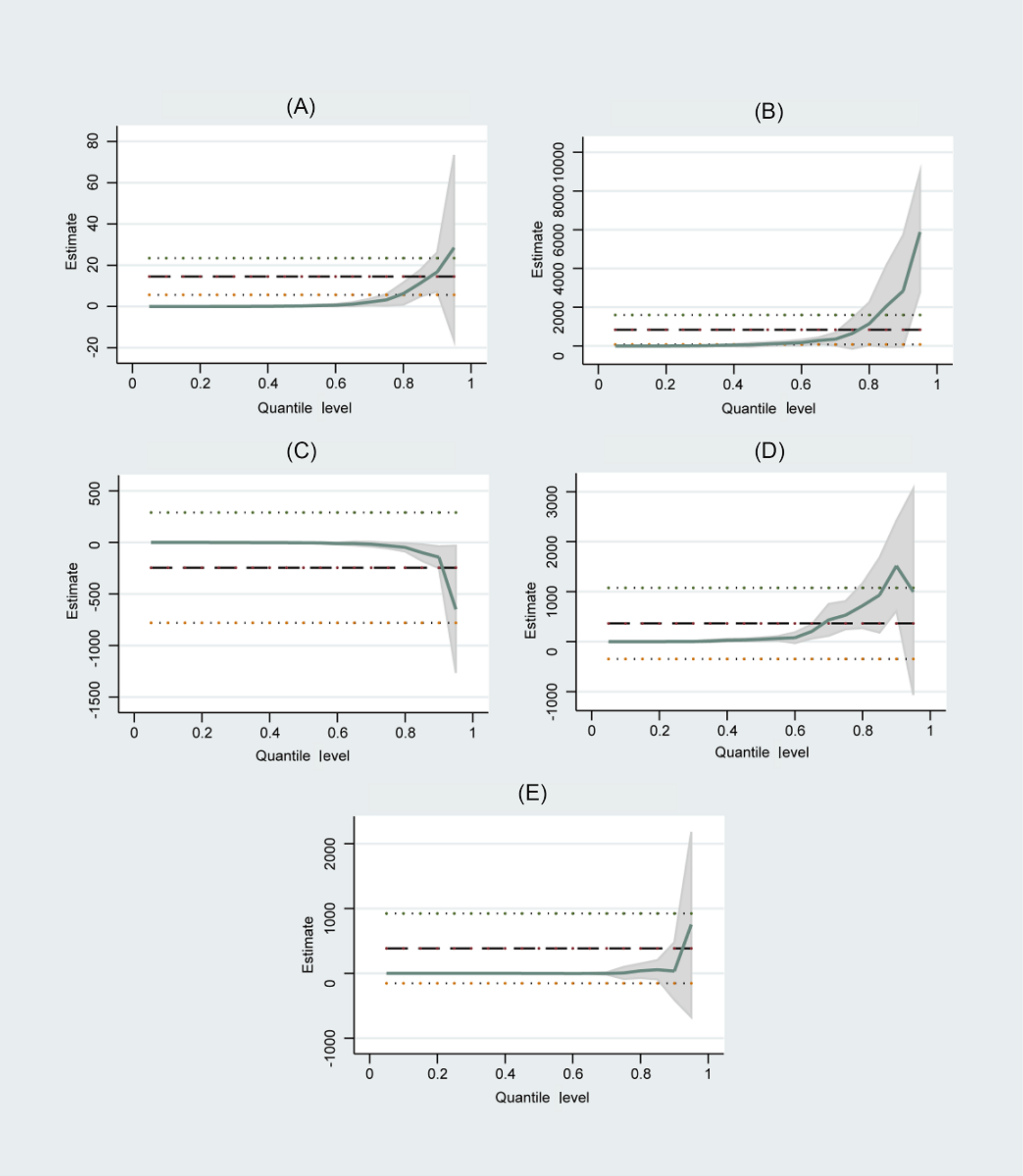
Based on the QR analysis, a positive effect for adding COVID-19 function and content on apps (P_{10}: *P*=.001; P_{50}: *P*=.01; P_{75}: *P*=.009; P_{90}: *P*=.01) was observed across download changes in most quantiles, with the largest effect at the 90th quantile (see Table 2, Figure 8, and Multimedia Appendix 2). Positive effects of the size of apps on download changes were also observed at high quantiles (P_{75}: *P*=.03; P_{90}: *P*=.002). Other characteristics of the QR are shown in Table 2.

Table 2. Quantile regression coefficients of app characteristics for changes in downloads.

App characteristics	Quantiles				
	0.10	0.25	0.50	0.75	0.90
Operating status					
Coefficient (x10^4; 95% CI)	0.000142 (–0.151 to 0.151)	–0.0428 (–0.500 to 0.414)	0.809 (–2.786 to 4.404)	–4.951 (–28.311 to 18.409)	–20.304 (–120.276 to 79.669)
P value[a]	>.99	.85	.66	.68	.69
Size of app					
Coefficient (x10^4; 95% CI)	0.0000367 (–0.005 to 0.005)	0.0102 (–0.018 to 0.038)	0.256 (–0.041 to 0.554)	3.173 (0.239 to 6.108)	16.728 (6.124 to 27.332)
P value[a]	.99	.45	.09	.03	.002
In-app advertisement					
Coefficient (x10^4; 95% CI)	–0.0857 (–0.321 to 0.150)	1.072 (–11.297 to 13.441)	101.809 (15.949 to 187.668)	648.489 (–267.967 to 1564.944)	2840.221 (–213.048 to 5898.490)
P value[a]	.48	.87	.02	.17	.07
Number of functions					
Coefficient (x10^4; 95% CI)	–0.0504 (–0.252 to 0.151)	–0.691 (–1.340 to –0.041)	–3.501 (–9.198 to 2.196)	–31.571 (–69.032 to 5.889)	–143.767 (–321.698 to 34.165)
P value[a]	.62	.04	.23	.098	.11
Prevention classification					
Coefficient (x10^4; 95% CI)	0.0258 (–0.019 to 0.071)	0.0593 (–0.147 to 0.265)	1.675 (–0.448 to 3.798)	7.928 (–10.998 to 26.855)	66.478 (–22.249 to 155.205)
P value[a]	.26	.57	.12	.41	.14
Connect to smart devices					
Coefficient (x10^4; 95% CI)	0.337 (0.008 to 0.665)	0.166 (–0.323 to 0.655)	–1.576 (–6.213 to 3.062)	5.069 (–51.167 to 61.305)	34.596 (–256.520 to 325.712)
P value[a]	.045	.51	.51	.86	.82
Added functions or contents for pandemic					
Coefficient (x10^4; 95% CI)	0.450 (0.190 to 0.710)	1.788 (–3.121 to 6.698)	46.627 (9.800 to 83.453)	529.365 (132.535 to 926.194)	1517.683 (352.991 to 2682.376)
P value[a]	.001	.48	.01	.009	.01

[a] Significance was assessed at the 10% level.

Figure 8. Statistically significant differential effects of app characteristics by quantile: (A) app size, (B) in-app advertisement, (C) number of functions, (D) functions or content added for the pandemic, (E) Internet of Things.



Relationship Between App Functions and Download Changes

We found that 4 of 28 functions, including rehabilitation (P_{75}: *P*=.003; P_{90}: *P*=.02), pregnancy preparation (P_{90}: *P*=.09), bodybuilding (P_{50}: *P*=.07; P_{90}: *P*=.08), and vaccination (P_{75}: *P*=.06), were positively associated with download changes, mainly in the higher quantiles (see Table 3, Figure 9, and Multimedia Appendix 3). Health education (P_{75}: *P*=.09; P_{90}: *P*=.09), drug use (P_{90}: *P*=.08), cultivation of lifestyle (P_{25}: *P*=.02; P_{50}: *P*=.06; P_{75}: *P*=.07), men's health (P_{90}: *P*=.07), and disease management (P_{75}: *P*=.03) had negative correlations with download changes. Other functions in the QR are shown in Table S5 in Multimedia Appendix 1.

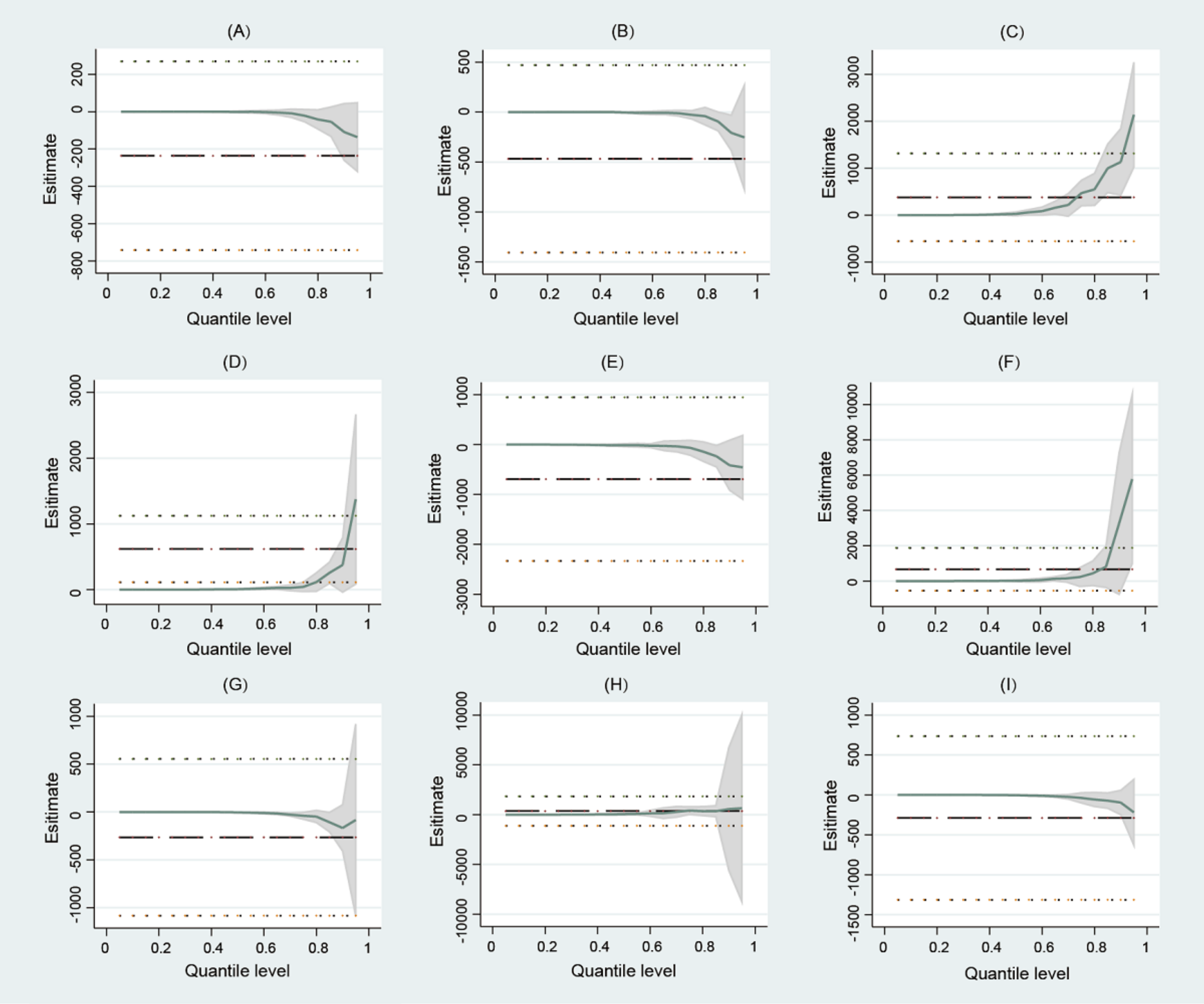
Table 3. Quantile regression coefficients of app functions for changes in downloads.

App functions	Quantiles				
	0.10	0.25	0.50	0.75	0.90
Health education					
Coefficient ($x10^4$; 95% CI)	0.0002 (–0.0687 to 0.0691)	0.0292 (–0.308 to 0.367)	–1.036 (–4.364 to 2.293)	–21.668 (–46.305 to 2.969)	–108.990 (–234.145 to 16.166)
P value[a]	>.99	.87	.54	.09	.09
Drug use					
Coefficient ($x10^4$; 95% CI)	0.0739 (–0.388 to 0.536)	0.133 (–0.866 to 1.132)	–3.450 (–11.716 to 4.816)	–26.942 (–103.928 to 50.045)	–207.444 (–436.652 to 21.764)
P value[a]	.75	.79	.41	.49	.08
Rehabilitation					
Coefficient ($x10^4$; 95% CI)	0.0992 (–0.218 to 0.417)	1.624 (–0.807 to 4.055)	26.345 (–10.549 to 63.239)	470.025 (160.786 to 779.265)	1132.3840 (172.086 to 2092.682)
P value[a]	.54	.19	.16	.003	.02
Bodybuilding					
Coefficient ($x10^4$; 95% CI)	0.0004 (–0.0324 to 0.0332)	0.0219 (–0.375 to 0.419)	7.194 (–0.472 to 14.861)	41.692 (–39.050 to 122.433)	376.090 (–49.408 to 801.587)
P value[a]	.98	.91	.07	.31	.08
Men's health					
Coefficient ($x10^4$; 95% CI)	–0.0047 (–0.0324 to 0.0332)	–1.497 (–5.075 to 2.081)	–11.928 (–45.420 to 21.565)	–65.017 (–242.440 to 112.407)	–415.092 (–869.710 to 39.527)
P value[a]	>.99	.41	.49	.47	.07
Pregnancy preparation					
Coefficient ($x10^4$; 95% CI)	0.1684 (–1.938 to 2.275)	3.037 (–1.692 to 7.765)	16.135 (–26.886 to 59.157)	226.703 (–248.826 to 702.233)	3257.672 (–446.489 to 6961.834)
P value[a]	.88	.21	.46	.35	.09
Cultivation of lifestyle					
Coefficient ($x10^4$; 95% CI)	8.35[b] (–0.0418 to 0.0418)	–0.445 (–0.803 to –0.0872)	–4.078 (–8.238 to 0.083)	–39.345 (–81.525 to 2.834)	–166.124 (–461.988 to 129.740)
P value[a]	>.99	.02	.06	.07	.27
Vaccination					
Coefficient ($x10^4$; 95% CI)	0.0046 (–5.995 to 6.004)	7.758 (–18.734 to 34.251)	44.707 (–139.634 to 229.048)	410.179 (–18.195 to 838.553)	566.790 (–5041.161 to 6174.740)
P value[a]	>.99	.57	.63	.06	.84
Disease management					
Coefficient ($x10^4$; 95% CI)	0.0418 (–0.377 to 0.460)	–0.102 (–0.856 to 0.652)	–4.011 (–9.643 to 1.622)	–37.508 (–71.617 to –3.398)	–95.521 (–237.203 to 46.161)
P value[a]	.85	.79	.16	.03	.19

[a] Significance was assessed at the 10% level.

[b] 10^{-17}.

Figure 9. Statistically significant differential effects of app functions by quantile: (A) health education, (B) drug use, (C) rehabilitation, (D) bodybuilding, (E) men’s health, (F) pregnancy preparation, (G) cultivation of lifestyle, (H) vaccination, (I) disease management.



User Experience

No significant difference was found in the sex (*P*=.41) or mean age (*P*=.52) of the participants. The attention given to health information (prepandemic: 249/375, 66.4%; postpandemic: 146/178, 82.0%; *P*=.006) and the percentage of people owning smartphones (prepandemic: 186/375, 49.7%; postpandemic: 108/178, 60.7%; *P*=.02) steadily increased after the outbreak (see Table 4). The vast majority of individuals (prepandemic: 129/141, 92.1%; postpandemic: 89/90, 98.9%) used social networks (eg, WeChat) to obtain health information online. Furthermore, the ratio of internet hospital use rose dramatically (prepandemic: 6/375, 1.6%; post-pandemic: 23/178, 12.9%; *P*<.001). Other characteristics of the user experience analysis are shown in Table 4.

Table 4. Population use of mobile health (mHealth) apps in the pre and postpandemic periods.

Characteristics	Before the outbreak (n=375)	After the outbreak (n=178)	*P* value
Age (years), median (IQR)	70 (66-74)	70 (66-75)	.52
Age (years), mean (SD)	70.46 (6.329)	70.85 (7.680)	—[a]
Sex, n (%)			.41
Male	166 (44.3)	85 (48.0)	
Female	209 (55.7)	92 (52.0)	
Attention to health information, n (%)			.006
Never	125 (33.3)	32 (18.0)	
Sometimes	131 (34.9)	75 (42.1)	
Often	73 (19.5)	57 (32.0)	
Always	45 (12.0)	14 (7.9)	
The ability to use electronic products (mobile phone, tablet, computer), n (%)	186 (49.7)	108 (60.7)	.02
Getting health information offline in the past 6 months, n (%)	108 (28.8)	104 (58.4)	<.001
Getting health information through the internet in the past 6 months, n (%)	141 (37.6)	90 (50.6)	.004
Ways to obtain health information, n (%)			
Social network (eg, WeChat, QQ)	129 (92.1)	89 (98.8)	.008
Portal web	36 (25.9)	19 (22.6)	.63
mHealth apps	6 (4.3)	2 (2.5)	.71
Search engine	36 (25.5)	22 (26.5)	.87
Internet hospitals	6 (1.6)	23 (12.9)	<.001
Consuming intention for internet health information per month (¥[b]), n (%)			.70
0	316 (84.3)	150 (84.3)	
1-10	10 (2.7)	5 (2.8)	
11-50	19 (5.1)	9 (5.1)	
51-100	3 (0.8)	0 (0)	
101-200	2 (0.5)	1 (0.6)	
201-500	1 (0.3)	0 (0)	
>500	0 (0)	0 (0)	
The actual cost for internet health information last month (¥[b]), n (%)			.45
0	348 (92.8)	164 (92.1)	
1-10	0 (0)	2 (1.1)	
11-50	0 (0)	2 (1.1)	
51-100	1 (0.3)	0 (0)	
101-200	1 (0.3)	0 (0)	
201-500	1 (0.3)	0 (0)	
>500	1 (0.3)	0 (0)	
Intention to avoid some unhealthy behaviors after obtaining health information from the internet, n (%)			.67
Never	15 (4.0)	0 (0)	
Sometimes	12 (3.2)	16 (9.0)	
Often	65 (17.3)	47 (26.4)	
Always	25 (6.7)	9 (5.0)	

[a] Age was not normally distributed, so differences in the median value were assessed using a nonparametric test.

[b] A currency exchange rate of ¥1=US $0.15 is applicable.

Discussion

Principal Findings

Our study demonstrates that the usage and population utilization of mHealth applications increased after the COVID-19 outbreak. As a powerful tool for providing health care services, functions closely related to the pandemic, including rehabilitation, treatment, drug use, and vaccination, were positively associated with changes in app downloads. The high growth of app use related to maternal and child health, including pregnancy preparation and women's health, shows the potentially increased desire for family among the Chinese population in the postpandemic era. Moreover, the user experience and high use of health management apps also reflect great attention to self-care. Overall, mHealth apps assist with health improvement against the background of normalized pandemic control and may improve fertility.

COVID-19–Related Apps

The usage of COVID-19 pandemic-related apps, such as vaccination, increased in rank. Furthermore, adding pandemic-related functions positively correlated with increased downloads. The likely reason behind this rise was that apps inherently related to the pandemic can easily capture the attention of the public as a means of obtaining information. Some apps with larger user groups may also add COVID-19 modules to respond to normalized pandemic prevention and control policies [14].

Medical Support Apps

Unprecedented large-scale quarantine measures and shortages of medical resources have made telemedicine care an important and real demand during the pandemic [28]. In our research, rehabilitation, including apps for long-term care or chronic disease management, was the most widely used function before and after the outbreak, overcoming the interruption of personal health care services caused by COVID-19. mHealth apps are used for many rehabilitation purposes [25]**.** One study showed that effective rehabilitation apps helped patients increase their health and happiness index during the pandemic period [29]. Furthermore, the use of mHealth apps can help improve adherence to treatment [30]. Use of apps with functions for rehabilitation, treatment, and drug use increased significantly after the outbreak, providing stable medical services to reduce the negative impact of home isolation. China's emerging internet model provides the basis for remote pharmacy services [31], assisting patients who cannot always go to the pharmacy. At the same time, personalized medical plans are also an important part of precision medicine [32]. The city of Taizhou, China, had a successful experience using telemedicine to prevent and treat COVID-19 [33]. Consequently, mHealth apps are an effective medical tool in the context of COVID-19.

Health Management Apps

Most apps were designed for health management by all people, mainly focusing on bodybuilding and nutrition. In our research, app flow and user experience surveys both showed that the rankings of most functions related to health management were rising, which reflects great attention to self-care postpandemic.

It is interesting that there was increased use of maternal and child health apps, including pregnancy preparation, women's health, pregnancy, and parenting, after the outbreak, showing potentially increased desire for fertility among the Chinese population.

One study concluded that the pandemic is affecting people's desire to become parents, which was consistent with our results [34]. Another previous study conducted a cross-sectional survey of 285 apps to analyze the current situation of maternity apps in Italy, finding that high-quality, targeted, and effective apps for pregnancy and postnatal health care had relevant implications in terms of maternal and newborn health prevention and promotion [35]. mHealth apps have the potential to be used extensively in improving maternal well-being [36]. The high growth of maternal and child health-related apps such as pregnancy preparation and women's health in our study coincides with this. Before the pandemic, the progression of population aging, decline of women in the childbearing period, and increase in high-risk pregnancies had led to the continuous decline in China's fertility rate [21]. Women have suffered significant reproductive health disruption since the beginning of the COVID-19 pandemic [37,38], which may have aggravated the decline in fertility [39]. However, with long-term home isolation and 2-child and 3-child policies proposed by the government, people are giving more attention to family matters, including pregnancy preparation and raising children, which may slow the decline in fertility to a certain extent. Therefore, maternal and child health apps can effectively assist health management and fill the vacancy of in-person perinatal health care services [40], despite interrupted pregnancy checkups [39]. It has been suggested that a model of health management combined with continuous care using the WeChat platform can significantly improve patients' postoperative medication compliance and quality of life by requiring them to complete their rehabilitation tasks in the WeChat group, which is worth applying and promoting [41]. In China, pregnant women are accustomed to using WeChat groups recommended by hospitals to discuss pregnancy health information and may promote apps in it, which may be a way to increase app usage.

For bodybuilding apps, this function had positive changes in downloads during the COVID-19 pandemic. One study found that the keyword "mHealth" was closely associated with "physical activity" and "ehealth" in the last 2 decades of research on digital health behavior change technologies [42]. Closed gyms and restricted outdoor activities during the pandemic reduced physical activity levels; however, the use of physical activity apps may counteract the decline in exercise [43].

Many kinds of apps attempted to provide health education, which was the most widely available function in our study. Although this function declined during the pandemic and had negative changes in downloads, the population's attention to health information increased. The probable cause behind this phenomenon is that new media platforms in China, such as WeChat and Weibo, have been vigorously promoted as important means for pandemic-related health information dissemination, which may decrease interest in acquiring apps when the information is readily available on these platforms. The high levels of knowledge of the Chinese public about COVID-19 prevention mainly comes from WeChat [44]. This is consistent with the results of our study that people use WeChat most often to obtain health information.

The COVID-19 pandemic has caused health anxiety at the population level. Digital intervention by mHealth apps is suitable for alleviating such sociopsychological consequences [45]. However, it may be because of the vigorous development of the psychological counseling hotline project in China during the pandemic that mental health apps were not of great concern. Furthermore, a study that conducted a systematic assessment of self-guided cognitive behavioral therapy (CBT)–based apps concluded that only a few self-guided CBT-based apps offer comprehensive CBT programs or suicide risk management resources, which may also be one of the reasons for not getting much attention [46].

Willingness to Use mHealth Apps

Consistent with the report of the rapid increase in older adult internet users during the COVID-19 pandemic in China, we found that people over 50 years old paid more attention to mHealth apps after the outbreak [47-50]. In conjunction with the increased interest of older adults in the mHealth space, the results of a previous study suggested that person-centered mHealth apps can be used to create mHealth solutions with positive outcomes for older adults [51]**.** Coupled with the higher probability of chronic diseases or other conditions, older individuals should be a primary and adaptable group for mHealth apps [52]. However, the digital divide makes optimizing mHealth apps among older individuals difficult [53]. The Chinese people strongly support public health measures proposed due to COVID-19, which makes it possible to develop the ability to use smart devices, reduce the digital divide, and provide QR codes for pass certificates [54]. It is worth noting that a larger software file size was positively associated with download changes, indicating that excessive software size has little effect on the willingness to use the software. Having the right function is what users are most concerned about, and the causal relationship needs to be further studied [55].

Limitations

Our study should be considered in the context of important limitations. First, this study excluded apps that were used internally by medical staff because we could not log in as an internal account holder; this dilutes the results of mHealth apps designed for health care professionals. However, our research focused on apps for patients and healthy people rather than internal apps. Future work will be conducted with apps used internally in medical care. Second, the absence of newly emerging apps made it impossible to provide an overview of the mHealth market after the COVID-19 outbreak. Therefore, we compared the changes in downloads during the pre and postpandemic periods to explore the relationship between various types of apps and the pandemic. Third, as a semilongitudinal survey, this study measured exposure and outcome, and it was difficult to derive causal relationships from the analysis; thus, we only made assumptions based on the status quo.

Strengths

The study has practical implications and applications. This study is the first to investigate the relationship between COVID-19 and population-level utilization of mHealth apps through a semilongitudinal study of app markets' data and a field questionnaire, combined with the results of both the web-based survey and the population user experience survey. As China is one of the few countries to adopt more active public health prevention and control measures, this study, which involves a multilevel and broad research scope, can provide strong data support for future comparative studies between different countries and regions. In the user experience survey, we explored the changing attitudes of the population toward digital health technology, suggesting that there is a good development environment for mHealth apps in the postpandemic era. This study, with consistent definitions of variables and processes, allowed the investigators to consistently classify mHealth apps and ensure data integrity, underpinning its strength. Our research clarified the relationship between various types of apps and usage changes by conducting investigations in the pre and postpandemic periods. We believe our results provide a good reference for the subsequent development of future mHealth apps. In addition to the increasing number of COVID-19–related apps prompted by pandemic policies, app developers should be aware that maternal, child, and self-care management are app functions about which the population is concerned.

Perspective

mHealth apps utilize information and telecommunications technology to transfer medical information for diagnosis, therapy, and education and played a significant role following the COVID-19 outbreak. The pandemic made people aware of the value of mHealth in promoting universal health coverage, which promotes stronger management of self-care. Against the backdrop of an increased desire to raise a family among the Chinese population in the postpandemic era, maternal and child health apps, as a health education tool, promote a healthy lifestyle for women's self-management in the antenatal and postpartum periods. Further research is needed to understand the users' requirements for these apps, which will influence their adoption. The explicit design of apps is another potential factor that can facilitate or hinder user engagement and requires further investigation.

Conclusion

mHealth apps are an effective approach to providing health care in the context of COVID-19. This study clarifies the increasing usage of different apps during the pre and postpandemic periods, showing greater attention to self-care and the Chinese population's increasing desire to raise a family. Moreover, our research provides direction for subsequent mHealth app development and promotion in the postepidemic era, supporting medical model reformation in China as a reference. This may provide new avenues for designing and evaluating indirect public health interventions such as health education and health promotion. Further research is needed to investigate the functions in each kind of app, which will contribute to the personalized development and specific improvement measures of mHealth apps as a health promotion strategy.

Acknowledgments

We thank the study participants. This research was supported by Fujian Medical University Talent Research Funding (XRCZX2019031), National Natural Science Foundation of China Youth Program (82203989), Natural Science Foundation of Fujian (2021J01729), Fujian Province Students’ Innovative Entrepreneurial Training Plan Program (S202010392017; S202010392019), Fujian Medical University College Students’ Innovative Entrepreneurial Training Plan Program (C21118, C22024), and Fujian Medical University Public Health School College Students’ Innovative Entrepreneurial Training Plan Program (xy202010015, xy202010018, xy202110005).

Data Availability

Scientists wishing to use mobile health app study data for noncommercial purposes can obtain the data set by contacting the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 34 KB - publichealth_v9i1e40552_app1.docx]

Multimedia Appendix 2
Differential effect of app functions by quantile (insignificant).
[PNG File , 202 KB - publichealth_v9i1e40552_app2.png]

Multimedia Appendix 3
Differential effect of app functions by quantile (insignificant).
[PNG File , 1172 KB - publichealth_v9i1e40552_app3.png]

References

1. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
2. O’Connell J, O’Keeffe DT. Contact tracing for Covid-19 — a digital inoculation against future pandemics. N Engl J Med 2021 Aug 05;385(6):484-487. [doi: 10.1056/nejmp2102256]
3. Cencetti G, Santin G, Longa A, Pigani E, Barrat A, Cattuto C, et al. Digital proximity tracing on empirical contact networks for pandemic control. Nat Commun 2021 Mar 12;12(1):1655 [FREE Full text] [doi: 10.1038/s41467-021-21809-w] [Medline: 33712583]
4. Dzau VJ, McClellan MB, McGinnis JM, Marx JC, Sullenger RD, ElLaissi W. Vital directions for health and health care: priorities for 2021. Health Aff (Millwood) 2021 Feb;40(2):197-203. [doi: 10.1377/hlthaff.2020.02204] [Medline: 33476192]
5. Murray CJL, Alamro NMS, Hwang H, Lee U. Digital public health and COVID-19. Lancet Public Health 2020 Sep;5(9):e469-e470 [FREE Full text] [doi: 10.1016/S2468-2667(20)30187-0] [Medline: 32791051]
6. Check Hayden E. Mobile-phone health apps deliver data bounty. Nature 2016 Mar 24;531(7595):422-423. [doi: 10.1038/531422a] [Medline: 27008946]
7. Shuren J, Patel B, Gottlieb S. FDA regulation of mobile medical apps. JAMA 2018 Jul 24;320(4):337-338. [doi: 10.1001/jama.2018.8832] [Medline: 29971339]
8. Wang Y, Wang Y, Greene B, Sun L. An analysis and evaluation of quality and behavioral change techniques among physical activity apps in China. Int J Med Inform 2020 Jan;133:104029. [doi: 10.1016/j.ijmedinf.2019.104029] [Medline: 31759245]
9. Xiao Q, Lu S, Wang Y, Sun L, Wu Y. Current status of cardiovascular disease-related smartphone apps downloadable in China. Telemed J E Health 2017 Mar;23(3):219-225. [doi: 10.1089/tmj.2016.0083] [Medline: 27356156]
10. Shang J, Wei S, Jin J, Zhang P. Mental health apps in China: analysis and quality assessment. JMIR Mhealth Uhealth 2019 Nov 07;7(11):e13236 [FREE Full text] [doi: 10.2196/13236] [Medline: 31697245]
11. Li Y, Ding J, Wang Y, Tang C, Zhang P. Nutrition-related mobile apps in the China App Store: assessment of functionality and quality. JMIR Mhealth Uhealth 2019 Jul 30;7(7):e13261 [FREE Full text] [doi: 10.2196/13261] [Medline: 31364606]

12. El-Sherif DM, Abouzid M. Analysis of mHealth research: mapping the relationship between mobile apps technology and healthcare during COVID-19 outbreak. Global Health 2022 Jun 28;18(1):67 [FREE Full text] [doi: 10.1186/s12992-022-00856-y] [Medline: 35765078]
13. Hsu J, Liu D, Yu YM, Zhao HT, Chen ZR, Li J, et al. The top Chinese mobile health apps: a systematic investigation. J Med Internet Res 2016 Aug 29;18(8):e222 [FREE Full text] [doi: 10.2196/jmir.5955] [Medline: 27573724]
14. Perlman A, Vodonos Zilberg A, Bak P, Dreyfuss M, Leventer-Roberts M, Vurembrand Y, et al. Characteristics and symptoms of app users seeking COVID-19-related digital health information and remote services: retrospective cohort study. J Med Internet Res 2020 Oct 20;22(10):e23197 [FREE Full text] [doi: 10.2196/23197] [Medline: 32961527]
15. Wang CJ. Contact-tracing app curbed the spread of COVID in England and Wales. Nature 2021 Jun 25;594(7863):336-337. [doi: 10.1038/d41586-021-01354-8] [Medline: 34035507]
16. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav 2021 Feb;5(2):247-255. [doi: 10.1038/s41562-020-01044-x] [Medline: 33479505]
17. Almalki M, Giannicchi A. Health apps for combating COVID-19: descriptive review and taxonomy. JMIR Mhealth Uhealth 2021 Mar 02;9(3):e24322 [FREE Full text] [doi: 10.2196/24322] [Medline: 33626017]
18. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. J Med Internet Res 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: 10.2196/23170] [Medline: 33197234]
19. Ahlers-Schmidt CR, Hervey AM, Neil T, Kuhlmann S, Kuhlmann Z. Concerns of women regarding pregnancy and childbirth during the COVID-19 pandemic. Patient Educ Couns 2020 Sep 24;103(12):2578-2582 [FREE Full text] [doi: 10.1016/j.pec.2020.09.031] [Medline: 33010997]
20. Hong Z, Li N, Li D, Li J, Li B, Xiong W, et al. Telemedicine during the COVID-19 pandemic: experiences from western China. J Med Internet Res 2020 May 08;22(5):e19577 [FREE Full text] [doi: 10.2196/19577] [Medline: 32349962]
21. Vollset SE, Goren E, Yuan C, Cao J, Smith AE, Hsiao T, et al. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. Lancet 2020 Oct 17;396(10258):1285-1306 [FREE Full text] [doi: 10.1016/S0140-6736(20)30677-2] [Medline: 32679112]
22. Peng PY. A complete collection of China family planning. Beijing, China: China Population Press; 1997.
23. 2018-2019 China Mobile App Store Market Monitoring Report. iiMedia Research. 2018 Dec 24. URL: https://www.iimedia.cn/c400/63240.html [accessed 2023-01-26]
24. Tianyancha. URL: https://www.tianyancha.com/ [accessed 2023-01-29]
25. Koenker R, Bassett G. Regression quantiles. Econometrica 1978 Jan;46(1):33. [doi: 10.2307/1913643]
26. Koenker R, Hallock KF. Quantile regression. Journal of Economic Perspectives 2001 Nov;15(4):143-156 [FREE Full text] [doi: 10.1257/jep.15.4.143]
27. Yin M. Research on the current situation and innovation development strategy of China Mobile Health Service (in Chinese). Journal of Medical Information 2019;40:2-6. [doi: 10.3969/j.issn.1673-6036.2019.02.001]
28. Dorsey ER. The new platforms of health care. NPJ Digit Med 2021 Jul 15;4(1):112 [FREE Full text] [doi: 10.1038/s41746-021-00478-5] [Medline: 34267312]
29. Hermans ANL, van der Velden RMJ, Gawalko M, Verhaert DVM, Desteghe L, Duncker D, TeleCheck-AF investigators. On-demand mobile health infrastructures to allow comprehensive remote atrial fibrillation and risk factor management through teleconsultation. Clin Cardiol 2020 Nov;43(11):1232-1239 [FREE Full text] [doi: 10.1002/clc.23469] [Medline: 33030259]
30. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile Apps for Increasing Treatment Adherence: Systematic Review. J Med Internet Res 2019 Jun 18;21(6):e12505 [FREE Full text] [doi: 10.2196/12505] [Medline: 31215517]
31. Chen Z, Liang W, Liu Q, He R, Chen Q, Li Q, et al. Use of a remote oncology pharmacy service platform for patients with cancer during the COVID-19 pandemic: implementation and user acceptance evaluation. J Med Internet Res 2021 Jan 21;23(1):e24619 [FREE Full text] [doi: 10.2196/24619] [Medline: 33395398]
32. Ang IYH, Tan KXQ, Tan C, Tan CH, Kwek JWM, Tay J, et al. A personalized mobile health program for type 2 diabetes during the COVID-19 pandemic: single-group pre-post study. JMIR Diabetes 2021 Jul 09;6(3):e25820 [FREE Full text] [doi: 10.2196/25820] [Medline: 34111018]
33. Shao C, Li S, Zhu F, Zhao D, Shao H, Chen H, et al. Taizhou's COVID-19 prevention and control experience with telemedicine features. Front Med 2020 Aug 29;14(4):506-510 [FREE Full text] [doi: 10.1007/s11684-020-0811-8] [Medline: 32862396]
34. Micelli E, Cito G, Cocci A, Polloni G, Russo GI, Minervini A, et al. Desire for parenthood at the time of COVID-19 pandemic: an insight into the Italian situation. J Psychosom Obstet Gynaecol 2020 Sep;41(3):183-190. [doi: 10.1080/0167482X.2020.1759545] [Medline: 32379999]
35. Brunelli L, De Vita C, Cenedese F, Cinello M, Paris M, Samogizio F, et al. Gaps and future challenges of Italian apps for pregnancy and postnatal care: systematic search on app stores. J Med Internet Res 2021 Aug 10;23(8):e29151 [FREE Full text] [doi: 10.2196/29151] [Medline: 34383668]
36. Chan KL, Chen M. Effects of social media and mobile health apps on pregnancy care: meta-analysis. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e11836 [FREE Full text] [doi: 10.2196/11836] [Medline: 30698533]

37. Phelan N, Behan LA, Owens L. The impact of the COVID-19 pandemic on women's reproductive health. Front Endocrinol (Lausanne) 2021 Mar 22;12:642755 [FREE Full text] [doi: 10.3389/fendo.2021.642755] [Medline: 33841334]
38. Fuentes-Afflick E, Perrin JM, Moley KH, Díaz Á, McCormick MC, Lu MC. Optimizing health and well-being for women and children. Health Aff (Millwood) 2021 Feb;40(2):212-218. [doi: 10.1377/hlthaff.2020.01504] [Medline: 33476200]
39. Watson C. Stillbirth rate rises dramatically during pandemic. Nature 2020 Sep 15;585(7826):490-491. [doi: 10.1038/d41586-020-02618-5] [Medline: 32934376]
40. Radin JM, Steinhubl SR, Su AI, Bhargava H, Greenberg B, Bot BM, et al. The Healthy Pregnancy Research Program: transforming pregnancy research through a ResearchKit app. NPJ Digit Med 2018;1:45 [FREE Full text] [doi: 10.1038/s41746-018-0052-2] [Medline: 31304325]
41. Xu M, Yang X, Liu L, Dai Y, Xu M, Lin S. Effect of the WeChat platform health management and refined continuous nursing model on life quality of patients with acute myocardial infarction after PCI. J Healthc Eng 2021;2021:5034269 [FREE Full text] [doi: 10.1155/2021/5034269] [Medline: 34880976]
42. Taj F, Klein MCA, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: 10.2196/13311] [Medline: 31833836]
43. Eapen ZJ, Peterson ED. Can Mobile Health Applications Facilitate Meaningful Behavior Change?: Time for Answers. JAMA 2015 Mar 21;314(12):1236-1237. [doi: 10.1001/jama.2015.11067] [Medline: 26393844]
44. Ma Z, Idris S, Pan Q, Baloch Z. COVID-19 knowledge, risk perception, and information sources among Chinese population. World J Psychiatry 2021 May 19;11(5):181-200 [FREE Full text] [doi: 10.5498/wjp.v11.i5.181] [Medline: 34046314]
45. Rauschenberg C, Schick A, Hirjak D, Seidler A, Paetzold I, Apfelbacher C, et al. Evidence Synthesis of Digital Interventions to Mitigate the Negative Impact of the COVID-19 Pandemic on Public Mental Health: Rapid Meta-review. J Med Internet Res 2021 Mar 10;23(3):e23365 [FREE Full text] [doi: 10.2196/23365] [Medline: 33606657]
46. Martinengo L, Stona A, Griva K, Dazzan P, Pariante CM, von Wangenheim F, et al. Self-guided cognitive behavioral therapy apps for depression: systematic assessment of features, functionality, and congruence with evidence. J Med Internet Res 2021 Jul 30;23(7):e27619 [FREE Full text] [doi: 10.2196/27619] [Medline: 34328431]
47. The 44th Statistical Report on Internet Development in China. China Internet Network Information Center (CNNIC). 2019 Aug. URL: https://www.cnnic.com.cn/IDR/ReportDownloads/201911/P020191112539794960687.pdf [accessed 2023-01-26]
48. The 45th Statistical Report on Internet Development in China. China Internet Network Information Center (CNNIC). 2020 Apr. URL: https://www.cnnic.com.cn/IDR/ReportDownloads/202008/P020200827549953874912.pdf [accessed 2023-01-26]
49. The 46th Statistical Report on Internet Development in China. China Internet Network Information Center (CNNIC). 2020 Sep. URL: https://www.cnnic.com.cn/IDR/ReportDownloads/202012/P020201201530023411644.pdf [accessed 2023-01-26]
50. The 47th Statistical Report on Internet Development in China. China Internet Network Information Center (CNNIC). 2021 Feb. URL: https://www.cnnic.com.cn/IDR/ReportDownloads/202104/P020210420557302172744.pdf [accessed 2023-01-26]
51. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e29512 [FREE Full text] [doi: 10.2196/29512] [Medline: 35029535]
52. Shengshou H, Runlin G, Lisheng L, Manlu Z, Wen W, Yongjun W, Zhaosu W, Huijun L, Dongfeng G, Yuejin Y, Zhe Z, Weiwei C. Summary of the 2018 Report on Cardiovascular Diseases in China. Chinese Circulation Journal 2019 Mar 24;34(03):209-220. [doi: 10.3969/j.issn.1000-3614.2019.03.001]
53. Zhou X, Chen L. Digital health care in China and access for older people. The Lancet Public Health 2021 Dec;6(12):e873-e874. [doi: 10.1016/s2468-2667(21)00051-7]
54. Lang R, Benham JL, Atabati O, Hollis A, Tombe T, Shaffer B, et al. Attitudes, behaviours and barriers to public health measures for COVID-19: a survey to inform public health messaging. BMC Public Health 2021 Apr 21;21(1):765 [FREE Full text] [doi: 10.1186/s12889-021-10790-0] [Medline: 33882896]
55. Lazard AJ, Babwah Brennen JS, Belina SP. App designs and interactive features to increase mHealth adoption: user expectation survey and experiment. JMIR Mhealth Uhealth 2021 Nov 04;9(11):e29815 [FREE Full text] [doi: 10.2196/29815] [Medline: 34734829]

Abbreviations

CBT: cognitive behavioral therapy
mHealth: mobile health
QR: quantile regression

Edited by T Sanchez, A Mavragani; submitted 27.06.22; peer-reviewed by S Badawy, MM Patwary, E Limijadi, F Petracca; comments to author 06.10.22; revised version received 26.10.22; accepted 12.01.23; published 22.02.23.

Please cite as:
Yang L, Wu J, Mo X, Chen Y, Huang S, Zhou L, Dai J, Xie L, Chen S, Shang H, Rao B, Weng B, Abulimiti A, Wu S, Xie X
Changes in Mobile Health Apps Usage Before and After the COVID-19 Outbreak in China: Semilongitudinal Survey
JMIR Public Health Surveill 2023;9:e40552
URL: https://publichealth.jmir.org/2023/1/e40552
doi:10.2196/40552
PMID:36634256

Original Paper

Dining-Out Behavior as a Proxy for the Superspreading Potential of SARS-CoV-2 Infections: Modeling Analysis

Ka Chun Chong[1], PhD; Kehang Li[1], MSc; Zihao Guo[1], MSc; Katherine Min Jia[2], MSc; Eman Yee Man Leung[1], PhD; Shi Zhao[1], PhD; Chi Tim Hung[1], MBBS; Carrie Ho Kwan Yam[1], PhD; Tsz Yu Chow[1], BSc; Dong Dong[1], PhD; Huwen Wang[1], MSc; Yuchen Wei[1], PhD; Eng Kiong Yeoh[1], MBBS

[1]Centre for Health Systems and Policy Research, Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, New Territories, Hong Kong

[2]Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, United States

Corresponding Author:
Eng Kiong Yeoh, MBBS
Centre for Health Systems and Policy Research, Jockey Club School of Public Health and Primary Care
The Chinese University of Hong Kong
Room 415, School of Public Health Building
Prince of Wales Hospital, Shatin
New Territories
Hong Kong
Phone: 852 22528702
Email: yeoh_ek@cuhk.edu.hk

Abstract

Background: While many studies evaluated the reliability of digital mobility metrics as a proxy of SARS-CoV-2 transmission potential, none examined the relationship between dining-out behavior and the superspreading potential of COVID-19.

Objective: We employed the mobility proxy of dining out in eateries to examine this association in Hong Kong with COVID-19 outbreaks highly characterized by superspreading events.

Methods: We retrieved the illness onset date and contact-tracing history of all laboratory-confirmed cases of COVID-19 from February 16, 2020, to April 30, 2021. We estimated the time-varying reproduction number (R_t) and dispersion parameter (k), a measure of superspreading potential, and related them to the mobility proxy of dining out in eateries. We compared the relative contribution to the superspreading potential with other common proxies derived by Google LLC and Apple Inc.

Results: A total of 6391 clusters involving 8375 cases were used in the estimation. A high correlation between dining-out mobility and superspreading potential was observed. Compared to other mobility proxies derived by Google and Apple, the mobility of dining-out behavior explained the highest variability of k (ΔR-sq=9.7%, 95% credible interval: 5.7% to 13.2%) and R_t (ΔR-sq=15.7%, 95% credible interval: 13.6% to 17.7%).

Conclusions: We demonstrated that there was a strong link between dining-out behaviors and the superspreading potential of COVID-19. The methodological innovation suggests a further development using digital mobility proxies of dining-out patterns to generate early warnings of superspreading events.

(JMIR Public Health Surveill 2023;9:e44251) doi:10.2196/44251

KEYWORDS

COVID-19; contact tracing; unlinked; superspreading; dispersion; public health; surveillance; digital health surveillance; digital surveillance; disease spread

Introduction

The COVID-19 pandemic has caused nearly 300 million confirmed cases and over 5 million attributable deaths worldwide during 2020-2021. As human mobility is a key driver of SARS-CoV-2 transmission [1], nonpharmaceutical interventions (NPI) aimed at reducing human movements and contacts, including travel restrictions, case detection, isolation, quarantine of close contacts, and social distancing, have been effective in flattening incidence curves and protecting the health care system from being overwhelmed [2]. Of the human activities, indoor dining at eateries exposes people to a high SARS-CoV-2 infection risk, especially since mask mandate does not apply in those occasions. Several investigations have shown that permitting on-premises dining was associated with increased risk of COVID-19 infections [3,4].

In the current pandemic, digital data encoding human mobility information have been highly leveraged for purposes of assessing risk of infectious disease transmission, quantifying the effectiveness and compliance of NPI, and enabling early detection early detection of cases to prompt timely interventions. To facilitate the control of the COVID-19 pandemic, several multinational companies with GPS-related services released their mobility data, such as the community mobility reports from Google LLC [5], mobility trends report from Apple Inc [6], and the COVID-19 mobility data network from Meta Data for Good program [7]. There are some more fine-grained digital proxies applied to construct the COVID-19 mobility networks, which entailed hourly movements to specific points of interest, including restaurants [8,9].

As a densely populated metropolitan, Hong Kong was vulnerable to considerable COVID-19 outbreak risk and has kept community transmissions at a containable level through case detection and isolation, intense contact tracing, and quarantine in the early phases of the pandemic [10,11]. However, the superspreading events (SSE), where a large number of secondary cases were generated by a few primary cases, presented challenges for the effectiveness of these measures. Previous studies indicated that approximately 20% of cases seeded 80% of all local transmissions in the first and second wave of COVID-19 epidemics [12,13]. Similar observations were also reported in other Asia-Pacific settings such as South Korea [14]. The government has rolled out different levels of measures to prevent transmissions by restricting capacity, limiting the number of diners per table, and shortening the dine-in hours. Eateries were found to be settings where major clusters with SSE occurred, even though corresponding restriction measures such as social distancing and limited dine-in hours were in place [15,16].

Many studies assessed the reliability of different digital mobility metrics as a proxy of the SARS-CoV-2 transmission [17-21]. For example, Kurita et al [19] demonstrated that Apple mobility data were useful for the short-term prediction of COVID-19 transmissibility. Nevertheless, to our knowledge, none have examined the relationship between the trend of dining-out behaviors and the superspreading potential of COVID-19. A good understanding of the relationship is essential for informing and evaluating the effectiveness of NPI. In this work, we aim to examine the association between the trend of dining-out behaviors and disease transmissibility in Hong Kong, a setting with COVID-19 outbreaks characterized by SSE.

Methods

Epidemiological Data

Surveillance data on local COVID-19 cases were provided by the Centre for Health Protection, Department of Health of the Government of the Hong Kong Special Administrative Region. All COVID-19 infections were confirmed by testing with polymerase chain reaction. We retrieved the illness onset date and contact-tracing history of the confirmed cases. Based on the contact history, we reconstructed the transmission clusters, defined as a number of cases with the same source of infection (ie, a primary case) or epidemiologically linked [22]. Cases without source cases that did not have epidemiological linkage with other confirmed cases were defined as sporadic cases (ie, cases with untraceable contacts). A sporadic case was considered as a cluster equal to one. The study period was from February 16, 2020, to April 30, 2021, when the third wave of transmissions was brought under control.

Digital Mobility Proxy of Dining-Out Patterns

A web crawler was developed to retrieve user comments from OpenRice, the most commonly used restaurant catalog for people to search and provide feedback in Hong Kong. The website covers restaurant information and user comments for more than 28,000 eateries and is open for public access without a log-in requirement. We obtained the daily total number of comments as a proxy of dining out in eateries from February 16, 2020, to April 30, 2021. The daily counts were normalized using the baseline means of total comments between November 1 and December 31, 2018.

Other Digital Mobility Proxies

We obtained anonymized and aggregated human mobility data released by Google and Apple during the COVID-19 pandemic by locations and transport modes [5,6]. The Google COVID-19 community mobility reports included the daily change of visitor numbers to 6 locations (ie, retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential) compared to the median value in baseline days (from January 3, 2020, to February 6, 2020). The commuting information was collected through GPS linked with the Google Maps app. Similarly, mobility data from Apple mobility trends were also obtained, which reports on the daily ratio of trip numbers to the baseline (January 13, 2020). The data sets were generated from users' devices connected to the Apple Maps service for directions and were categorized by driving and walking.

Estimation of the Time-Varying Reproduction Number

Following Cori et al [23], we estimated the time-varying reproduction number (R_t) over the study period. In this approach, the daily number of cases by illness onset date was modeled through a branching process, where the incidence at time t (ie, I_t) is Poisson distributed with the mean given by , where ω_s is the discretized probability distribution of the serial interval. Epidemiologically, a serial interval approximated the infectiousness of a COVID-19 case at time s after the symptom onset. We assumed the mean and SD of the serial interval during the study period to be 6.5 and 4.1 days, respectively, based on a previous study conducted in Hong Kong [24]. As the empirical serial interval estimates may be biased due to the sampling and selection bias within the contact-tracing data collected during an ongoing epidemic [25], we also tested another assumption of serial interval distribution (mean 4.6, SD 4.9 days) estimated from a modeling study that corrected for such bias [26] as a sensitivity analysis. We only considered cases that acquired the infection locally in the estimation. The R_t was estimated by fitting the time-series data to the Poisson distributions in a sliding window process, assuming the R_t remained constant within each window.

Estimation of Superspreading Potential

Given the stochastic effect of transmissions, the transmission dynamics can be described by an offspring distribution (ie, the distribution of the number of secondary cases generated by the primary case). Following Lloyd et al [27], we assumed a negative-binomial offspring distribution, which was parametrized by a reproduction number and a dispersion parameter (*k*). When *k* is sufficiently low (ie, less than 1), SSE are more likely to occur. Applying the branching process theory, *k* could be estimated by fitting the transmission cluster data to a cluster size distribution [28], which describes the probability of clusters with a size of *z* seeded by *u* primary cases. We only included transmission clusters that were initiated by local cases. For an ongoing epidemic, the clusters' sizes are likely to grow after the time of estimation. To this end, we also considered right censoring in the cluster size distributions [22,28,29]. Thus, the likelihood function is:

Here, Pr(.) is the probability mass function of the cluster size distribution assuming a negative binomial offspring distribution following previous studies [22]. The term *c* is an indicator of censoring, where *c*=0 if the cluster is censored and *c*=1 if the cluster is considered as self-limited. A cluster is censored if it has new confirmed cases within 11 days before the day of estimation [22]. We determined the time-varying *k* during the study period in a sliding window process.

The window length was fixed at 30 days, which could cover 4 transmission generations based on the upper bound of the estimated generation interval of COVID-19 [30]. Within each window period, we estimated *k* using the Markov chain Monte Carlo method. Gamma distribution and half-normal distribution were used as prior distributions for R_t and *k*, respectively. For each Markov chain Monte Carlo chain, we obtained 10,000 thinned posterior samples from 500,000 iterations.

Associations Between Mobility Proxies and Transmission Dynamics

Time series of the mobility proxies were smoothed by a 7-day rolling average. The Pearson correlation coefficients (*r*) between the mobility proxies and R_t, and between the mobility proxies and *k*, were computed, respectively. The R-square was determined to quantify the proportion of variance explained by a mobility proxy using linear regression. A full regression model was built including all the mobility proxies. R_t and *k* were log-transformed and negative log-transformed in the regression models, respectively. To quantify the relative contribution of a specific mobility proxy in predicting the R_t or *k*, we determined the difference in *R*-squares (ΔR-sq) between a full model and the model without the proxy [31]. We calculated the means and 95% credible intervals (CrIs) for the estimates by sampling 10,000 times from the posterior distributions of the estimates; 3 and 7 days of lags were tested to examine the lag effects of mobility on the outcomes of disease transmissibility ensuring the robustness of the study findings.

The analysis was conducted in R statistical software (version 4.0.3; The R Foundation). The programming code is available upon request.

Ethics Approval

Ethics approval was obtained from the Survey and Behavioral Research Ethics Committee, The Chinese University of Hong Kong (SBRE-20-581). Because this study was a modeling analysis using secondary data with no personal information, the requirement for obtaining informed consent was waived.

Results

The relationship between and among social distancing policies, disease transmission dynamics, and the mobility proxies from February 16, 2020, to April 30, 2021, is illustrated in Figure 1 and Figure 2. The estimated R_t decreased from 1.94 (95% CrI 1.63 to 2.29) to 0.17 (95% CrI 0.07 to 0.34), from 2.88 (95% CrI 2.59 to 3.20) to 0.59 (95% CrI 0.55 to 0.63), and from 1.63 (95% CrI 1.29 to 2.01) to 0.72 (95% CrI 0.66 to 0.78) at 3 intervention phases, respectively.

A total of 6391 clusters involving 8375 cases were reported during the study period. Among the identified clusters, there were 4527 clusters with a size equal to 1 (ie, sporadic cases). The distribution of the number of secondary cases generated by the primary cases for the whole study period is showed in Multimedia Appendix 1. Of the 7194 linked cases, 1096 (15.23%) cases generated at least 2 secondary cases, and 1 case even initiated a transmission cluster composing 394 cases. The cluster data were used to estimate *k*, which was negative log-transformed, so the larger the value, the higher the superspreading potential. At the 3 intervention phases, the negative-log *k* decreased from 3.49 (95% CrI 3.12 to 3.83) to 2.34 (95% CrI –0.70 to 3.30), from 3.02 (95% CrI 2.63 to 3.37) to 0.69 (95% CrI 0.28 to 1.06), and from 2.51 (95% CrI 1.61 to 3.17) to 0.68 (95% CrI 0.09 to 1.19), respectively (Figure 2).

Similar decreasing trends were observed for the mobility proxies (Figure 2 and Table 1). A high correlation between dining-out behavior and *k* was observed (*r*=0.46, 95% CrI 0.32 to 0.56). While several other mobility proxies were also highly correlated with *k*, including visiting retail and recreation (*r*=0.41, 95% CrI 0.24 to 0.54), parks (*r*=0.36, 95% CrI 0.17 to 0.51), and workplaces (*r*=0.37, 95% CrI 0.24 to 0.46), mobility proxies of transit stations, driving, and walking had a weaker correlation with *k* (*r*<0.15). Correlations between mobility proxies and R_t were generally weaker. Only the mobility of dining-out behavior in eateries (*r*=0.16, 95% CrI 0.14 to 0.18) and workplace (*r*=0.13, 95% CrI 0.10 to 0.16) maintained a larger correlation with R_t.

The mobility proxies could explain a higher percent of variability of *k* than that of R_t. In the full model, including all the mobility proxies, the *R*-squares were 54.3% (95% CrI 30.3% to 75.2%) and 23.7% (95% CrI 21.5% to 25.9%) for regressing *k* and R_t, respectively. We used ΔR-sq to quantify how much of the variabilities of the k and R_t were explained by each of the mobility proxies (Figure 3 and Table 1). Among the proxies, the mobility of dining-out behavior explained the most of variability of k (ΔR-sq=9.7%, 95% CrI 5.7% to 13.2%) and R_t (ΔR-sq=15.7%, 95% CrI 13.6% to 17.7%). Other mobility proxies result in a mean ΔR-sq of less than 5% when regressing k. For R_t, the proxies of driving and transit stations had a ΔR-sq of 7.8% (95% CrI 6.3% to 9.3%) and 3.3% (95% CrI 2.3% to 4.5%), respectively. Other mobility proxies had the mean ΔR-sq<3% when regressing R_t.

We examined the effects of 3-day and 7-day lags of mobility proxies on k and R_t (Multimedia Appendix 2). In general, dining-out behavior still explained the highest variance of the outcomes. The ΔR-sq of a 3-day lag of the metric was 7.8% (95% CrI 4.2% to 11.4%) and 13.1% (95% CrI 11.1% to 15.2%) for k and R_t, respectively, whereas that of a 7-day lag of the metric was 4.1% (95% CrI 2.1% to 6.8%) and 11.7% (95% CrI 9.9% to 13.7%) for k and R_t, respectively. The changes in ΔR-sq for other mobility proxies were also not apparent when their lag effects were considered. Apart from that, a change of the assumed serial interval (mean 4.6, SD 4.9 days) only resulted in a minor change in the ΔR-sq for R_t with the primary findings remaining unaffected (Multimedia Appendix 3).

Figure 1. Number of reported cases and estimated time-varying reproduction number (Rt) in Hong Kong from February 16, 2020, to April 30, 2021. The number of reported cases is indicated by the gray bars (left axis), and the posterior median estimate of Rt is indicated by the red line, with shading indicating the 95% credible interval (right axis). The purple shaded areas show the intervention phases at different periods. Phase 1: A table of restaurants was limited to 4 people, and 6 types of premises must close from 6 PM from March 28, 2020, and bars must close from 6 PM from April 3, 2020. The measures started to be relaxed from May 5, 2020 (left); Phase 2: Restaurant dine-in services at night were banned, and no more than 2 persons may be seated together at 1 table in restaurants since July 13, 2020. The measures started to be relaxed since August 28, 2020 (middle). Phase 3: Business hours of restaurants, bars, and clubs were shortened; the number of people allowed to be seated together at 1 table was reduced; and the number of people participating in any 1 banquet in catering premises would be limited to 40 since November 16, 2020. Dine-in services at restaurants from 6 PM to 5 AM of the next day were banned, and the number of people participating in a banquet was further limited to 20 since December 10, 2020. The measures started to be relaxed since February 18, 2021 (right).

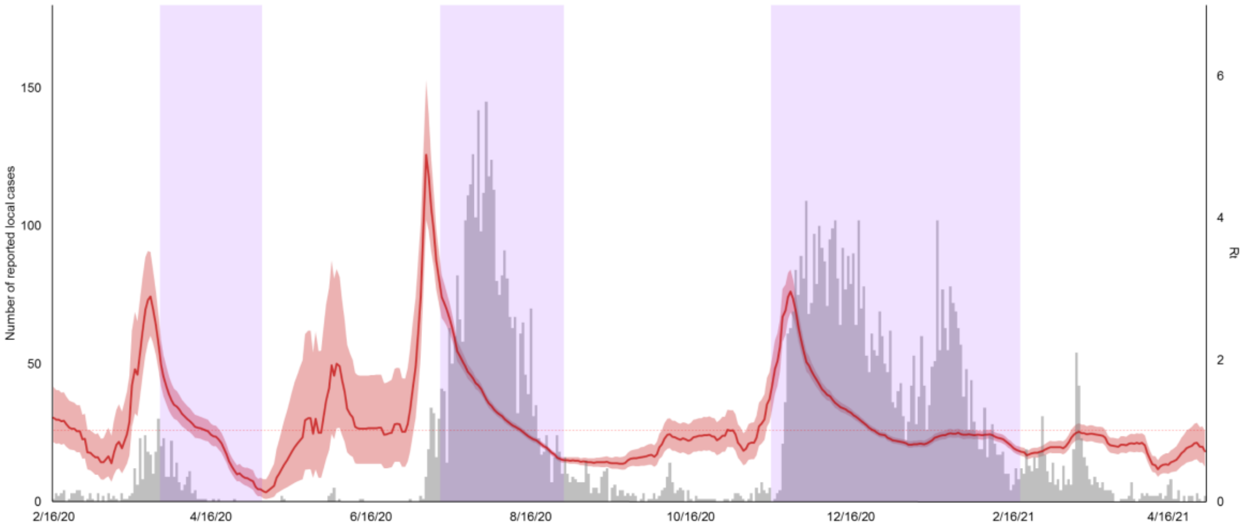


Figure 2. Negative logarithm of dispersion parameter (k) and mobility proxies. A. The negative logarithm of *k* is indicated by the blue line, with shading indicating the 95% credible interval (left axis), and the normalized counts of dining-out mobility is indicated by the orange line (right axis). B. The Google Mobility Index by types. C. The Apple Mobility Index by types.

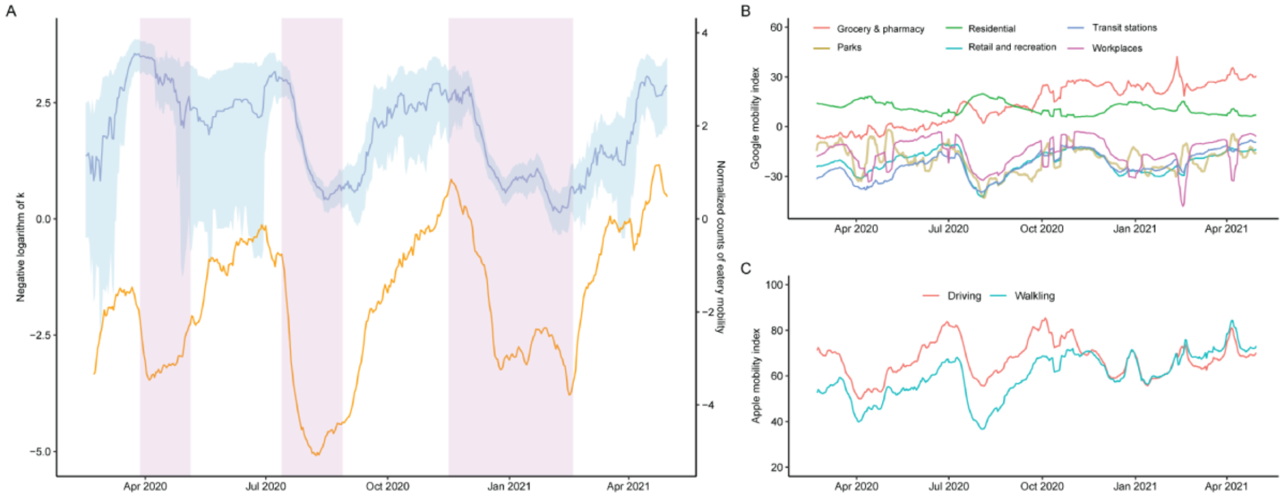


Table 1. Associations between mobility proxies and transmission dynamics[a].

Mobility proxies	R_t [b]		k [c]	
	r [d]	ΔR-sq[e] (%)	r	ΔR-sq (%)
Dining-out behavior	0.16 (0.14 to 0.18)	15.7 (13.6 to 17.7)	0.46 (0.32 to 0.56)	9.7 (5.7 to 13.2)
Retail and recreation	0.00 (–0.02 to 0.02)	0.8 (0.4 to 1.4)	0.41 (0.24 to 0.54)	3.7 (1.0 to 7.3)
Grocery and pharmacy	–0.01 (–0.04 to 0.03)	2.4 (1.5 to 3.5)	–0.23 (–0.36 to –0.04	1.3 (0.0 to 4.1)
Parks	–0.10 (–0.12 to –0.08)	0.8 (0.3 to 1.4)	0.36 (0.17 to 0.51)	0.4 (0.0 to 1.6)
Transit stations	0.02 (–0.01 to 0.04)	3.3 (2.3 to 4.5)	0.16 (0.06 to 0.24)	0.7 (0.0 to 2.3)
Workplaces	0.13 (0.10 to 0.16)	1.3 (0.8 to 1.9)	0.37 (0.24 to 0.46)	0.1 (0.0 to 0.5)
Residential	–0.03 (–0.05 to –0.01)	0.1 (0.0 to 0.2)	–0.21 (–0.29 to –0.11)	1.0 (0.2 to 2.0)
Driving	0.02 (0.00 to 0.05)	7.8 (6.3 to 9.3)	0.11 (–0.01 to 0.22)	0.4 (0.0 to 1.8)
Walking	–0.05 (–0.08 to –0.03)	0.0 (0.0 to 0.1)	0.01 (–0.08 to 0.08)	2.7 (1.0 to 4.7)

[a] The results were presented using means and 95% credible intervals of the estimates from the posterior distributions. Proxy of dining-out behavior was retrieved from OpenRice. Proxies of retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential were retrieved from Google COVID-19 community mobility reports. Proxies of driving and walking were retrieved from Apple mobility trends.

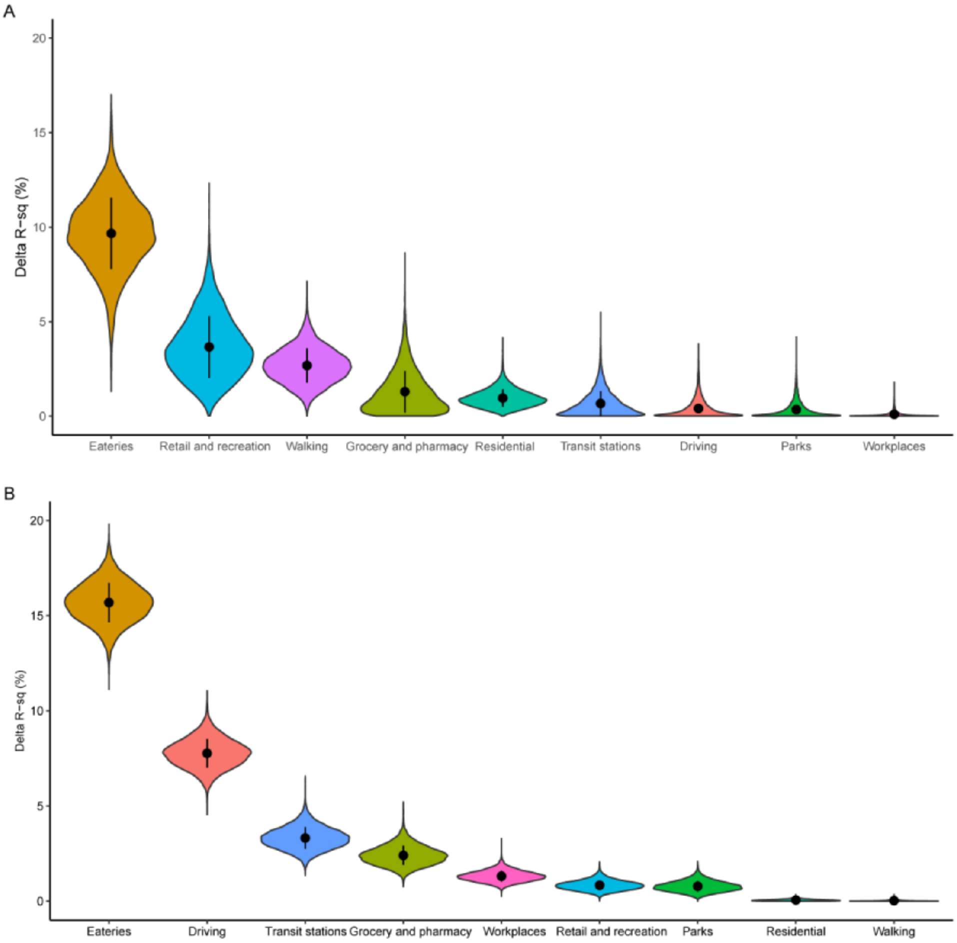
[b] R_t: time-varying reproduction number.

[c] k: dispersion parameter of superspreading potential.

[d] r: Pearson correlation coefficient.

[e] ΔR-sq: difference in *R*-squares between a full model and the model without the proxy.

Figure 3. Relative contributions of the mobility proxies in predicting dispersion parameter (k) and time-varying reproduction number (*R*t). A, B: The relative contributions of each of the mobility proxies, including eateries, retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, residential, driving, and walking, in predicting (A) *k* and (B) *R*t were quantified by the differences in *R*-squares (ΔR-sq). The black dot indicates the mean estimate of ΔR-sq interpolated with an SD.



Discussion

Principal Findings

Digital data on human mobility play an important role in tracking public compliance with NPI as well as monitoring infectious disease dynamics. While many studies evaluated the reliability of different digital mobility metrics as a proxy of the SARS-CoV-2 transmission [17-21], none have examined the relationship between the trend of dining-out behaviors and the superspreading potential of COVID-19. In this study, we evaluated the reliability of dining-out activities at eateries as a proxy for SARS-CoV-2 transmission risk in Hong Kong with outbreak highly characterized by SSE. According to our results, dining out activities were associated with the change in disease transmissibility as well as the superspreading potential of COVID-19. This finding is consistent with the fact that eatery venues were high-risk settings for frequent virus exposure and superspreading, as they tend to be indoors, populated, and without the mandate of mask wearing, having the highest proportion of COVID-19 infection counts than other places [32,33]. According to the COVID-19 spread pattern in Hong Kong, as characterized by previous research, eateries had a higher outbreak potential of unlinked cases than other settings and accounted for the second highest number of linked transmissions after the households [15,16], as dining-out activities gather individuals that are not socially connected in adjacent and common areas of the restaurant, widening the dispersion of infected persons and subsequently sustaining case clusters propagated across different settings. Similar to our results, the county-level COVID-19 growth in the United States increased from 1.1% to 1.2% when dining-out restrictions were removed, and a 55% decline in new COVID-19 cases was found after imposing bans on indoor on-premises dining [3,4].

During the COVID-19 epidemic, Hong Kong did not impose any stringent measures on eatery settings (eg, compulsory banning of indoor dining and closures of eatery venues), which has been adopted in a number of countries such as the United Kingdom and Singapore, before the availability of COVID-19 vaccines [34,35]. Nevertheless, as demonstrated in this study, less stringent measures such as shortening the business hours and restricting the capacity of eateries were still able to decrease the related mobility as well as the superspreading potential, particularly in the second and the third intervention phases. The finding supports the effectiveness of these measures, even though the related effect on transmission has been questioned [36]. However, we cannot factor in the effects of other social distancing measures implemented, although by including other mobility proxies in our analyses, we have adjusted for some secular effects from the other measures.

Of the mobility proxies, mobile device data including the community mobility reports from Google LLC and mobility trend reports from Apple Inc have been well studied as proxies for changes in human activities and SARS-CoV-2 transmissibility [17-19]. For instance, a Japanese investigation demonstrated that Apple mobility data were reliable for a short-term prediction of COVID-19 transmission [19], whereas an Italian investigation showed the changes in transmissibility of SARS-CoV-2 were associated with both social distancing measures and Google mobility [18]. While the studies assured the reliability of these metrics, we showed that dining-out behavior had a greater contribution to the COVID-19 transmissibility compared with the digital proxies generated by Google and Apple, likely because their data are not specific to eateries. Nevertheless, our analysis showed that the movement metrics including driving and transit stations still accounted for a certain proportion of variability of the community transmission. This observation echoes a previous study in Hong Kong using digital transactions in public transport and social mixing data to forecast the COVID-19 epidemics [20]. Similar movement metrics provided by Baidu Huiyan were also well demonstrated as useful data to improve the temporal and spatial resolution of COVID-19 transmissions [20,21].

In this study, estimating superspreading potential required comprehensive data of rapid contact tracing. In the Omicron epidemic, public health resources for conventional contact tracing were overwhelmed in Hong Kong due to a rapid and huge surge in infections. With insufficient data on case linkages, the superspreading potential cannot be determined during the Omicron period. It suggests a need for novel digital tools for contact tracing, for example, by app-based tools that can notify users instantaneously when their contacts are confirmed positive [37,38]. The deployment of app-based tools at a population level not only improves the timeliness of tracing but also avoids the recall bias in contact tracing. While digital contact tracing is common in many places [39], residents in Hong Kong were skeptical of this type of app due to ethical and privacy concerns [40]. As long as the digital data are available with a large population coverage and adequate compliance, they could be incorporated in our analytical framework that requires comprehensive contact tracing data.

With the comprehensive information of rapid contact tracing before an Omicron outbreak in Hong Kong, the major strength of this study is the capability of using the corresponding data for the estimation of time-varying superspreading potential, which were unlikely to be available in other settings without intense contact tracing during the COVID-19 pandemic. In addition, while a majority of studies related the mobility proxies to reproduction numbers [17,19], our study is the first investigation that linked the mobility pattern to disease superspreading, especially for the COVID-19 transmission, which was found to be highly heterogeneous [12-14,28,41].

Limitations

Nonetheless, this study had several limitations. First, individual-level data on demographics, duration of indoor dining, and the number of people dining together were not available, thus lowering the resolution of transmission risk imposed on the individual in an eatery. Still, the aggregated information provided a simple, inexpensive, and readily available data source for informing decisions for pandemic control and predicting the superspreading potential. Second, dining-out behavior may be different by geographic locations given the differential distributions of eateries. Due to a lack of the geographic information of the mobility proxies, the modeling analysis was unable to account the spatial variation, and the study findings may thus suffer from bias. Third, our study did not intend to develop a forecast model given limited mobility and social mixing data [20]. We expect that an increased coverage of app-based tools would offer high-resolution data for further development of a prediction model. Fourth, a single data source of proxy for the dining-out behavior was used for this study. While OpenRice was the most common web page for commenting and rating different eateries, we were unable to include diners who preferred using other web page or apps such as Facebook.

Conclusion

In conclusion, we demonstrated that there was a strong link between dining-out behaviors and the superspreading potential of COVID-19, and the metric was more reliable than other common mobility data derived by Google and Apple Inc. The findings suggest that the mobility proxy of dining-out behavior is able to help health officials for monitoring the public compliance of social distancing measures as well as the cluster outbreak potential. For example, officials could adjust the intensity of social distancing measures (eg, restriction of the maximum number of seats in a restaurant and mandatory closure of eateries) based on the disease transmissibility inferred by the mobility proxies. In addition, our methodological innovation recommends a further development of using digital mobility proxies of dining-out patterns to generate early warnings of SSE, thereby assisting in resource planning on corresponding measures.

Acknowledgments

We would like to thank Hospital Authority and Department of Health, Hong Kong Government, for providing the data for this study, and HyperLab for developing the web crawler. The Centre for Health Systems and Policy Research funded by the Tung Foundation is acknowledged for the support throughout the conduct of this study.

This research was supported by Health and Medical Research Fund (grant numbers COVID190105, COVID19F03, and INF-CUHK-1), Collaborative Research Fund of University Grants Committee (grant number C4139-20G), National Natural Science Foundation of China (grant number 71974165), and Group Research Scheme from The Chinese University of Hong

Kong. The funder of the study had no role in study design, data collection, data analysis, data interpretation, writing of the manuscript, or the decision to submit for publication. All authors had full access to all the data in the study and were responsible for the decision to submit the manuscript for publication.

Data Availability

The sharing of data is restricted by the Department of Health, Hong Kong.

Authors' Contributions

KCC, KL, ZG, and SZ were responsible for the study design and conceptualization. YW, CY, TYC, ZG, and EKY carried out data collection and preprocessing. KCC, KL, ZG, YW, and KJ carried out analysis and interpretation. KCC, KL, and ZG prepared the manuscript. KJ, EL, SZ, CTH, TYC, DD, HW, YW, and EKY were responsible for the critical revision of the study. All authors contributed to the revision and review of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The observed offspring distribution with fitted curve for the transmission chains, February 16, 2020, to April 30, 2021. Dotted dark blue curve is the estimated probability density function of the negative binomial distribution.

[PDF File (Adobe PDF File), 56 KB - publichealth_v9i1e44251_app1.pdf]

Multimedia Appendix 2

Relative contributions of (A, B) 3-day and (C, D) 7-day lag effects of mobility proxies in predicting dispersion parameter (k) and time-varying reproduction number (*R*t). The relative contributions of each of the mobility proxies, including eateries, retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, residential, driving, and walking, in predicting (A, C) *k* and (B, D) *R*t were quantified by the differences in *R*-squares (ΔR-sq). The black dot indicates the mean estimate of ΔR-sq interpolated with an SD.

[PDF File (Adobe PDF File), 159 KB - publichealth_v9i1e44251_app2.pdf]

Multimedia Appendix 3

Relative contributions of the mobility proxies in predicting time-varying reproduction number (Rt) when the mean and SD of the assumed serial interval distribution were changed to 4.6 and 4.9 days, respectively. The relative contributions of each of the mobility proxies, including eateries, retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, residential, driving, and walking, in predicting *R*t were quantified by the differences in *R*-squares (ΔR-sq). The black dot indicates the mean estimate of ΔR-sq interpolated with an SD.

[PDF File (Adobe PDF File), 46 KB - publichealth_v9i1e44251_app3.pdf]

References

1. Kraemer MUG, Yang C, Gutierrez B, Wu C, Klein B, Pigott DM, Open COVID-19 Data Working Group, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 2020 May 01;368(6490):493-497 [FREE Full text] [doi: 10.1126/science.abb4218] [Medline: 32213647]
2. Islam N, Sharp SJ, Chowell G, Shabnam S, Kawachi I, Lacey B, et al. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. BMJ 2020 Jul 15;370:m2743 [FREE Full text] [doi: 10.1136/bmj.m2743] [Medline: 32669358]
3. Schnake-Mahl A, O'Leary G, Mullachery P, Vaidya V, Connor G, Rollins H, et al. The Impact of Keeping Indoor Dining Closed on COVID-19 Rates Among Large US Cities: A Quasi-Experimental Design. Epidemiology 2022 Mar 01;33(2):200-208. [doi: 10.1097/EDE.0000000000001444] [Medline: 34799474]
4. Guy GP, Lee FC, Sunshine G, McCord R, Howard-Williams M, Kompaniyets L, CDC COVID-19 Response Team, Mitigation Policy Analysis Unit, CDC Public Health Law Program. Association of State-Issued Mask Mandates and Allowing On-Premises Restaurant Dining with County-Level COVID-19 Case and Death Growth Rates - United States, March 1-December 31, 2020. MMWR Morb Mortal Wkly Rep 2021 Mar 12;70(10):350-354 [FREE Full text] [doi: 10.15585/mmwr.mm7010e3] [Medline: 33705364]
5. COVID-19 community mobility reports. Google LLC. URL: https://www.google.com/covid19/mobility/ [accessed 2022-04-26]
6. Mobility trends reports. Apple Inc. URL: https://www.apple.com/covid19/mobility [accessed 2022-04-13]
7. Data for good. Meta. URL: https://dataforgood.facebook.com/dfg/covid-19 [accessed 2022-04-26]

8. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature 2021 Jan 10;589(7840):82-87. [doi: 10.1038/s41586-020-2923-3] [Medline: 33171481]
9. Holtz D, Zhao M, Benzell SG, Cao CY, Rahimian MA, Yang J, et al. Interdependence and the cost of uncoordinated responses to COVID-19. Proc Natl Acad Sci U S A 2020 Aug 18;117(33):19837-19843 [FREE Full text] [doi: 10.1073/pnas.2009522117] [Medline: 32732433]
10. Lai C, Ng R, Wong M, Chong K, Yeoh Y, Chen Z, et al. Epidemiological characteristics of the first 100 cases of coronavirus disease 2019 (COVID-19) in Hong Kong Special Administrative Region, China, a city with a stringent containment policy. Int J Epidemiol 2020 Aug 01;49(4):1096-1105 [FREE Full text] [doi: 10.1093/ije/dyaa106] [Medline: 32601677]
11. Wong MCS, Ng RWY, Chong KC, Lai CKC, Huang J, Chen Z, et al. Stringent containment measures without complete city lockdown to achieve low incidence and mortality across two waves of COVID-19 in Hong Kong. BMJ Glob Health 2020 Oct 07;5(10):e003573 [FREE Full text] [doi: 10.1136/bmjgh-2020-003573] [Medline: 33028700]
12. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat Med 2020 Nov 17;26(11):1714-1719. [doi: 10.1038/s41591-020-1092-0] [Medline: 32943787]
13. Guo Z, Zhao S, Lee S, Mok C, Wong N, Wang J, et al. Superspreading potential of COVID-19 outbreak seeded by Omicron variants of SARS-CoV-2 in Hong Kong. J Travel Med 2022 Sep 17;29(6):taac049 [FREE Full text] [doi: 10.1093/jtm/taac049] [Medline: 35435992]
14. Guo Z, Zhao S, Ryu S, Mok CKP, Hung CT, Chong KC, et al. Superspreading potential of infection seeded by the SARS-CoV-2 Omicron BA.1 variant in South Korea. J Infect 2022 Sep;85(3):e77-e79 [FREE Full text] [doi: 10.1016/j.jinf.2022.05.041] [Medline: 35659549]
15. Chong KC, Jia K, Lee SS, Hung CT, Wong NS, Lai FTT, et al. Characterization of Unlinked Cases of COVID-19 and Implications for Contact Tracing Measures: Retrospective Analysis of Surveillance Data. JMIR Public Health Surveill 2021 Nov 16;7(11):e30968 [FREE Full text] [doi: 10.2196/30968] [Medline: 34591778]
16. Wong NS, Lee SS, Kwan TH, Yeoh E. Settings of virus exposure and their implications in the propagation of transmission networks in a COVID-19 outbreak. Lancet Reg Health West Pac 2020 Nov;4:100052 [FREE Full text] [doi: 10.1016/j.lanwpc.2020.100052] [Medline: 34013218]
17. Nouvellet P, Bhatia S, Cori A, Ainslie KEC, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission. Nat Commun 2021 Feb 17;12(1):1090 [FREE Full text] [doi: 10.1038/s41467-021-21358-2] [Medline: 33597546]
18. Manica M, Guzzetta G, Riccardo F, Valenti A, Poletti P, Marziano V, et al. Impact of tiered restrictions on human activities and the epidemiology of the second wave of COVID-19 in Italy. Nat Commun 2021 Jul 27;12(1):4570 [FREE Full text] [doi: 10.1038/s41467-021-24832-z] [Medline: 34315899]
19. Kurita J, Sugishita Y, Sugawara T, Ohkusa Y. Evaluating Apple Inc Mobility Trend Data Related to the COVID-19 Outbreak in Japan: Statistical Analysis. JMIR Public Health Surveill 2021 Feb 15;7(2):e20335 [FREE Full text] [doi: 10.2196/20335] [Medline: 33481755]
20. Leung K, Wu JT, Leung GM. Real-time tracking and prediction of COVID-19 infection using digital proxies of population mobility and mixing. Nat Commun 2021 Mar 08;12(1):1501 [FREE Full text] [doi: 10.1038/s41467-021-21776-2] [Medline: 33686075]
21. Gibbs H, Liu Y, Pearson C, Jarvis C, Grundy C, Quilty B. Changing travel patterns in China during the early stages of the COVID-19 pandemic. Nature communications 2020;11(1):1-9. [doi: 10.1101/2020.05.14.20101824]
22. Guo Z, Zhao S, Lee SS, Hung CT, Wong NS, Chow TY, et al. A statistical framework for tracking the time-varying superspreading potential of COVID-19 epidemic. Epidemics 2023 Jan 24;42:100670 [FREE Full text] [doi: 10.1016/j.epidem.2023.100670] [Medline: 36709540]
23. Cori A, Ferguson N, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol 2013 Nov 01;178(9):1505-1512 [FREE Full text] [doi: 10.1093/aje/kwt133] [Medline: 24043437]
24. Chan YD, Flasche S, Lam TT, Leung MJ, Wong M, Lam H, et al. Transmission dynamics, serial interval and epidemiology of COVID-19 diseases in Hong Kong under different control measures. Wellcome Open Res 2020 Nov 9;5:91. [doi: 10.12688/wellcomeopenres.15896.2]
25. Park SW, Sun K, Champredon D, Li M, Bolker BM, Earn DJD, et al. Forward-looking serial intervals correctly link epidemic growth to reproduction numbers. Proc Natl Acad Sci U S A 2021 Jan 12;118(2):e2011548118 [FREE Full text] [doi: 10.1073/pnas.2011548118] [Medline: 33361331]
26. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis 2020 Apr;93:284-286 [FREE Full text] [doi: 10.1016/j.ijid.2020.02.060] [Medline: 32145466]
27. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. Nature 2005 Nov 17;438(7066):355-359 [FREE Full text] [doi: 10.1038/nature04153] [Medline: 16292310]
28. Endo A, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Res 2020 Jul 10;5:67 [FREE Full text] [doi: 10.12688/wellcomeopenres.15842.3] [Medline: 32685698]

29. Farrington C, Kanaan M, Gay N. Branching process models for surveillance of infectious diseases controlled by mass vaccination. Biostatistics 2003 Apr;4(2):279-295. [doi: 10.1093/biostatistics/4.2.279] [Medline: 12925522]
30. Zhao S, Tang B, Musa SS, Ma S, Zhang J, Zeng M, et al. Estimating the generation interval and inferring the latent period of COVID-19 from the contact tracing data. Epidemics 2021 Sep;36:100482 [FREE Full text] [doi: 10.1016/j.epidem.2021.100482] [Medline: 34175549]
31. Chong KC, Ran J, Lau SYF, Goggins WB, Zhao S, Wang P, et al. Limited role for meteorological factors on the variability in COVID-19 incidence: A retrospective study of 102 Chinese cities. PLoS Negl Trop Dis 2021 Feb 24;15(2):e0009056 [FREE Full text] [doi: 10.1371/journal.pntd.0009056] [Medline: 33626051]
32. Ning H, Li Z, Qiao S, Zeng C, Zhang J, Olatosi B, et al. Revealing geographic transmission pattern of COVID-19 using neighborhood-level simulation with human mobility data and SEIR model: A Case Study of South Carolina. medRxiv 2022 Aug 17:8809 [FREE Full text] [doi: 10.1101/2022.08.16.22278809] [Medline: 36032968]
33. Liu C, Huang J, Chen S, Wang D, Zhang L, Liu X, et al. The impact of crowd gatherings on the spread of COVID-19. Environ Res 2022 Oct;213:113604 [FREE Full text] [doi: 10.1016/j.envres.2022.113604] [Medline: 35691382]
34. New regulations created by Secretary of State for business closure (COVID-19). GOV.UK. URL: https://www.gov.uk/government/news/new-regulations-created-by-secretary-of-state-for-business-closure-covid-19 [accessed 2022-09-30]
35. Circuit breaker to minimise further spread of COVID-19. Ministry of Health, Singapore. URL: https://www.moh.gov.sg/news-highlights/details/circuit-breaker-to-minimise-further-spread-of-covid-19 [accessed 2022-09-30]
36. Ho F, Tsang TK, Gao H, Xiao J, Lau EH, Wong JY, et al. Restaurant-Based Measures to Control Community Transmission of COVID-19, Hong Kong. Emerg Infect Dis 2022 Mar;28(3):759-761 [FREE Full text] [doi: 10.3201/eid2803.211015] [Medline: 35202535]
37. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D, et al. The epidemiological impact of the NHS COVID-19 app. Nature 2021 Jun 12;594(7863):408-412. [doi: 10.1038/s41586-021-03606-z] [Medline: 33979832]
38. Salathé M, Althaus CL, Anderegg N, Antonioli D, Ballouz T, Bugnon E, et al. Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. Swiss Med Wkly 2020 Dec 14;150(5153):w20457 [FREE Full text] [doi: 10.4414/smw.2020.20457] [Medline: 33327003]
39. No authors listed. Show evidence that apps for COVID-19 contact-tracing are secure and effective. Nature 2020 Apr;580(7805):563. [doi: 10.1038/d41586-020-01264-1] [Medline: 32350479]
40. Public health concerns more than privacy. The Standard. 2020. URL: https://www.thestandard.com.hk/section-news/section/17/225095/Public-health-concerns-more-than-privacy [accessed 2021-05-26]
41. Yeoh EK, Chong KC, Chiew CJ, Lee VJ, Ng CW, Hashimoto H, et al. Assessing the impact of non-pharmaceutical interventions on the transmissibility and severity of COVID-19 during the first five months in the Western Pacific Region. One Health 2021 Jun;12:100213 [FREE Full text] [doi: 10.1016/j.onehlt.2021.100213] [Medline: 33506086]

Abbreviations

CrI: credible interval
NPI: nonpharmaceutical interventions
SSE: superspreading events

Edited by A Mavragani, T Sanchez; submitted 11.11.22; peer-reviewed by Z Luan, C Zeng; comments to author 20.01.23; revised version received 01.02.23; accepted 14.02.23; published 07.03.23.

Please cite as:
Chong KC, Li K, Guo Z, Jia KM, Leung EYM, Zhao S, Hung CT, Yam CHK, Chow TY, Dong D, Wang H, Wei Y, Yeoh EK
Dining-Out Behavior as a Proxy for the Superspreading Potential of SARS-CoV-2 Infections: Modeling Analysis
JMIR Public Health Surveill 2023;9:e44251
URL: https://publichealth.jmir.org/2023/1/e44251
doi:10.2196/44251
PMID:36811849

Original Paper

Enhancing Public Health Communication Regarding Vaccine Trials: Design and Development of the Pan-European VACCELERATE Toolkit

Christos D Argyropoulos[1], MSci, MSc, PhD; Janina Leckler[2,3], PhD; Jon Salmanton-García[2,3], MSc, PhD; Marinos Constantinou[1], BSc; Alexandra Alexandrou[1], BSc; Sophia Themistocleous[1], MSc; Evgenia Noula[1]; George Shiamakkides[1], MSc; Andria Nearchou[1]; Fiona A Stewart[2,3], MSc, PhD; Kerstin Albus[2,3], MSc, PhD; Markela Koniordou[4], MRes; Ioannis Kopsidas[4], MD, PhD; Orly Spivak[5], MSc; Margot Hellemans[6], MSc; Greet Hendrickx[6], MSc; Ruth Joanna Davis[7], MSc; Anna Maria Azzini[7], MD; Paula Valle Simon[8,9], MSc; Antonio Javier Carcas-Sansuan[8,9], MD, PhD; Helena Hervius Askling[10,11], MD, PhD; Sirkka Vene[10,11], MSc; Jana Baranda Prellezo[12], MSc; Elena Álvarez-Barco[13], PhD; Alan J Macken[13], MSc; Romina Di Marzo[14], BSc; Catarina Luís[14], MSc, PhD; Ole F Olesen[14], PhD; Jesus A Frias Iniesta[8,9], MSc, PhD; Imre Barta[15], PhD; Krisztina Tóth[15], MD; Murat Akova[16], MD; Marc M J Bonten[17,18], MD, PhD; Miriam Cohen-Kandli[5], PhD; Rebecca Jane Cox[19], PhD; Lenka Součková[20,21,22], PhD; Petr Husa[20,21,22], MUDr, PhD; Ligita Jancoriene[23,24], MD, PhD; Odile Launay[25,26], MD; Jens Lundgren[27], MD, DMSc; Patrick Mallon[13], PhD; Charis Armeftis[1], MD; Laura Marques[28], MD; Pontus Naucler[10,11], MD, PhD; Jordi Ochando[12], PhD; Evelina Tacconelli[7], MD, PhD; Pierre van Damme[6], MD, PhD; Theoklis Zaoutis[4], MD, PhD; Sanne Hofstraat[29], MSc; Patricia Bruijning-Verhagen[29], MD, PhD; Markus Zeitlinger[30], MD; Oliver A Cornely[2,3,31,32,33], MD, PhD; Zoi Dorothea Pana[1], MD, MSc, PhD

[1]School of Medicine, European University Cyprus, Nicosia, Cyprus

[2]Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[3]Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf and Excellence Center for Medical Mycology, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[4]Collaborative Center for Clinical Epidemiology and Outcomes Research, Athens, Greece

[5]Ministry of Health of Israel, Jerusalem, Israel

[6]VAXINFECTIO, Centre of Evaluation of Vaccination, Faculty of Medicine and Health Science, Universiteit Antwerpen, Antwerp, Belgium

[7]Infectious Diseases, Department of Diagnostic and Public Health, University of Verona, Verona, Italy

[8]Hospital La Paz Institute for Health Research, Madrid, Spain

[9]Servicio Madrileño de Salud, Madrid, Spain

[10]Department of Infectious Diseases, Karolinska University Hospital, Stockholm, Sweden

[11]Department of Medicine, Division of Infectious Diseases, Karolinska Institutet, Stockholm, Sweden

[12]Centro Nacional de Microbiología, Instituto de Salud Carlos III, Madrid, Spain

[13]Centre for Experimental Pathogen Host Research, School of Medicine, University College Dublin, Dublin, Ireland

[14]European Vaccine Initiative, Heidelberg, Germany

[15]National Koranyi Institute for Pulmonology, Budapest, Hungary

[16]Department of Infectious Diseases and Clinical Microbiology, Faculty of Medicine, Haceteppe University, Ankara, Turkey

[17]Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, Utrecht, Netherlands

[18]Department of Medical Microbiology, University Medical Centre Utrecht, Utrecht, Netherlands

[19]Influenza Centre, Department of Clinical Science, University of Bergen, Bergen, Norway

[20]Masaryk University, Brno, Czech Republic

[21]University Hospital Brno, Brno, Czech Republic

[22]Czech Clinical Research Infrastructure Network, Brno, Czech Republic

[23]Institute of Clinical Medicine, Medical Faculty, Vilnius University, Vilnius, Lithuania

[24]Vilnius University Hospital Santaros klinikos, Medical Faculty, Vilnius University, Vilnius, Lithuania

[25]Institut National de la Santé et de la Recherche Médicale - France Recherche Nord & Sud Sida- HIV Hépatites Maladies Infectieuses Émergentes, Paris, France

[26]Assistance Publique Hopitaux de Paris, Université Paris Cité, Paris, France

[27]Centre of Excellence for Health, Immunity and Infections, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
[28]Centro Hospitalar Universitário do Porto, Porto, Portugal
[29]University Medical Centre Utrecht, Utrecht University, Utrecht, Netherlands
[30]Medizinische Universität Wien, Vienna, Austria
[31]Clinical Trials Centre Cologne, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[32]Center for Molecular Medicine Cologne, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[33]German Centre for Infection Research, Partner Site Bonn-Cologne, Cologne, Germany

Corresponding Author:
Zoi Dorothea Pana, MD, MSc, PhD
School of Medicine
European University Cyprus
6 Diogenis Str
Egkomi
Nicosia, 2404
Cyprus
Phone: 357 94049474
Email: z.pana@euc.ac.cy

Abstract

Background: The pan-European VACCELERATE network aims to implement the first transnational harmonized and sustainable vaccine trial Volunteer Registry, being a single entry point for potential volunteers of large-scale vaccine trials across Europe. This work exhibits a set of harmonized vaccine trial–related educational and promotional tools for the general public, designed and disseminated by the pan-European VACCELERATE network.

Objective: This study primarily aimed to design and develop a standard toolkit to increase positive attitudes and access to trustworthy information for better access and increased recruitment to vaccine trials for the public. More specifically, the produced tools are focused on inclusiveness and equity, and are targeting different population groups, including underserved ones, as potential volunteers for the VACCELERATE Volunteer Registry (older individuals, migrants, children, and adolescents). The promotional and educational material is aligned with the main objectives of the Volunteer Registry to increase public literacy and awareness regarding vaccine-related clinical research or trials and trial participation, including informed consent and legal issues, side effects, and frequently asked questions regarding vaccine trial design.

Methods: Tools were developed per the aims and principles of the VACCELERATE project, focusing on trial inclusiveness and equity, and are adjusted to local country-wise requirements to improve public health communication. The produced tools are selected based on the cognitive theory, inclusiveness, and equity of differently aged and underrepresented groups, and standardized material from several official trustworthy sources (eg, COVID-19 Vaccines Global Access; the European Centre for Disease Prevention and Control; the European Patients' Academy on Therapeutic Innovation; Gavi, the Vaccine Alliance; and the World Health Organization). A team of multidisciplinary specialists (infectious diseases, vaccine research, medicine, and education) edited and reviewed the subtitles and scripts of the educational videos, extended brochures, interactive cards, and puzzles. Graphic designers selected the color palette, audio settings, and dubbing for the video story-tales and implemented QR codes.

Results: This study presents the first set of harmonized promotional and educational materials and tools (ie, educational cards, educational and promotional videos, extended brochures, flyers, posters, and puzzles) for vaccine clinical research (eg, COVID-19 vaccines). These tools inform the public about possible benefits and disadvantages of trial participation and build confidence among participants about the safety and efficacy of COVID-19 vaccines and the health care system. This material has been translated into several languages and is intended to be freely and easily accessible to facilitate dissemination among VACCELERATE network participant countries and the European and global scientific, industrial, and public community.

Conclusions: The produced material could help fill knowledge gaps of health care personnel, providing the appropriate future patient education for vaccine trials, and tackling vaccine hesitancy and parents' concerns for potential participation of children in vaccine trials.

(JMIR Public Health Surveill 2023;9:e44491) doi:10.2196/44491

KEYWORDS

vaccine trials; volunteer registry; educational material; promotional material; patient education; health communication; health promotion; public health; COVID-19; coronavirus; SARS-CoV-2; pandemic; vaccine; vaccination; hesitancy; campaign; misinformation

Introduction

VACCELERATE [1] is an independent, innovative, and transparent Pan-European academic network with the aim of harmonizing multinational vaccine trial initiatives and conducting capacity mapping of vaccine clinical trials sites and laboratories with standardized methods and protocols across Europe. The network identifies and provides access to state-of-the-art vaccine trial sites to accelerate the development of vaccines and recruit volunteers for vaccine trials through the VACCELERATE Volunteer Registry [2].

The main goal of the VACCELERATE Volunteer Registry [2] is to establish the first transnational harmonized trial participation platform serving as single entry point for the European region. The VACCELERATE Volunteer Registry [2], under the mandate of the European Union’s Health Emergency Preparedness and Response Authority (HERA) [3], provides a sustainable platform source for the recruitment of potential volunteers for clinical studies for the current COVID-19 pandemic and future epidemics.

According to the literature, one of the main challenges of a volunteer registry is the registration of a large number of potential trial participants with great diversity [4,5] and dedication in order to be able to recruit and match suitable candidates for the selected clinical trial. Therefore, a great amount of effort is spent looking for strategies to improve communication [6] and convince [7] potential volunteers to take part in vaccine trials.

The enrollment of volunteers in vaccine trials is also influenced by factors such as complacency, conspiracy theories [8] on social media regarding COVID-19, convenience and confidence (the “Three Cs” model of vaccine hesitancy) [7,9], disinformation [10], fake news [11], fear [12], and increasing level of health misinformation [13-15]. The latter leads to an infodemic [16] of misleading information with immediate effects on public health, refusal, and vaccine hesitancy [9] for participation in clinical trials [17,18]. It should be noted that health misinformation is responsible for increased fear and lack of trust, which are major factors for enrollment also in cancer clinical trials [19-21]. Recently, Luís et al [22] reported that there is a lack of public information about COVID-19 vaccine trials among several population groups (eg, minorities and underrepresented populations or individuals), while the available tools for trial participation were also sparse.

The main purpose of this study is the design and development of a standardized toolkit based on the Social Cognitive Theory [23,24] to increase positive attitudes, access to trustworthy information for better access and increased recruitment to vaccine trials. Our secondary objective is to foster engagement of the community through the VACCELERATE Volunteer Registry and smart technology for participating in forthcoming clinical trials [25-28]. The VACCELERATE toolkit has been translated into several languages and is freely and easily accessible to facilitate dissemination among the participating countries of the VACCELERATE Consortium [1].

Methods

The development of the VACCELERATE educational and promotional tools included the following steps: (1) conceptualization of the idea and context based on the selected educational aims and objectives; (2) selection of targeted population groups (eg, older individuals, children, adolescents, and minorities); (3) prioritization of trustworthy sources and websites for seeking information about COVID-19 vaccine trials for the developed materials; (4) performing graphical design (eg, visualization, audio, product, print, and animation design); and (5) acquisition of a digital object identifier (DOI) for each produced material from a DOI registration agency. More details on the developed tools are presented in the *Results* section.

The context of the educational and promotional tools was created and reviewed by a group of experts in the field of infectious diseases and vaccine and pedagogical research. Two graphic designers developed and selected the appropriate color palette and audio (music, speaker voice, and dubbing) for the video story-tales and they adapted the produced material into an efficient pedagogical, attractive, and eye-catching way. The developed tools were (1) icons related to the COVID-19 pandemic (eg, use of a face mask, syringe, microscope, vaccine, handwashing, virus, etc); (2) different versions of logotypes for the VACCELERATE project; (3) typefaces; (4) samples of fonts with different sizes; (5) color palettes (primary and secondary colors); (6) imagery templates and guidelines for social media (eg, Twitter and Instagram); (7) human characters (ie, medical staff, minorities, and older individuals) as prototypes for any future VACCELERATE promotional and or educational material. Special attention was also paid to foster inclusiveness among the characters (different age groups, different ethic and other minority groups, citizens with disabilities, migrants, etc). The provided tools also include QR codes with a URL. The selected URL can be a website (as shown on the VACCELERATE website [29]) or selected animation videos from trustworthy sources of information (eg, COVID-19 Vaccines Global Access [ie, COVAX]; the European Centre for Disease Prevention and Control; the European Patients’ Academy on Therapeutic Innovation; the European Vaccine Initiative; Gavi, the Vaccine Alliance; the United Nations International Children's Emergency Fund; and the World Health Organization).

In addition, the produced puzzles test memory in a team-based manner and help develop hand-eye coordination.

Finally, the developed tools have been designed in a such a way that can be adjusted in accordance with local country-wise requirements and requirements to build confidence among participants about the safety and efficacy of COVID-19 vaccines and the health care system, which is a key parameter to improve recruitment among minority communities.

Results

VACCELERATE Flyers

The flyers (see Multimedia Appendix 1) provide the appropriate amount of information to the public regarding the registration process and the VACCELERATE Volunteer Registry in general. The flyers are A3-sized (29.7×43.18 cm) and include a short description of the VACCELERATE Volunteer Registry objectives and volunteer selection process. In addition, they contain the appropriate contact details (email ID) and link (along with a QR code) for registration in the Volunteer Registry [30]. The registration link directs the interested volunteer to a brief survey (which takes 5-10 minutes to complete) that collects comprehensive information required for clinical trials. The flyer invites all citizens to register, including those with prior COVID-19 vaccination, with specific questions regarding COVID-19 infection, and the invitations to eligible volunteers for future vaccine trials will be sent out by the Coordination Office of VACCELERATE (University Hospital Cologne, Cologne, Germany). Finally, the flyers have been translated into 9 languages (Cypriot Greek, English, French, German, Greek, Hebrew, Italian, Portuguese, and Spanish).

VACCELERATE Posters

The developed posters (see Multimedia Appendix 1) are A0-sized (84.1×118.9 cm) and depict human characters (ie, medical staff, minorities, and older individuals, among others) and a QR code including a URL for the VACCELERATE Volunteer Registry website [30], as well as catchy slogans such as “Be the missing piece” and “Together we can tackle the COVID-19 pandemic.” Finally, the posters have been translated into 7 languages, namely English, French, German, Greek, Hebrew, Italian, and Spanish, in collaboration with individual partners and national coordinators of the VACCELERATE Consortium. The promotion of the VACCELERATE Volunteer Registry for participating in COVID-19 vaccine trials could be achieved by affixing these prototype posters to a wall at sites or in buildings with high visibility (eg, hospitals, universities, malls, and public buildings).

VACCELERATE Extended Brochure (Booklet)

The extended VACCELERATE brochure (see Multimedia Appendix 1) comprises valuable information regarding the VACCELERATE Volunteer Registry, such as “What is the VACCELERATE Network,” the mission and objectives, the current promotion activities, the importance of registration, and a description of the registration process. The brochure also contains key information for the volunteers, such as the possibility to withdraw their registration at any time. Finally, the extended brochure also includes a QR code that directs the interested reader to the VACCELERATE Volunteer Registry website [30] after scanning it with a smart device.

VACCELERATE Interactive Educational Cards

The educational cards enhance general knowledge and address the misinformation about COVID-19 vaccines, vaccine trials, and COVID-19 vaccination, as well as voluntary trial participation of children and adults. Each set of educational cards represents an educational aim and comprises a specific number of cards and target groups, as shown in Table 1.

The interactive educational cards (Multimedia Appendix 1) have been prepared with the help of 2 graphic designers and under the guidance of a pediatrician and several VACCELERATE participants. The educational cards include a QR code with a URL for promoting educational material on the one side and additional text script and characters on the other. The URL can be a website (eg, the one showed on the VACCELERATE website [29]) or selected videos from trustworthy sources (eg, COVID-19 Vaccines Global Access [ie, COVAX]; the European Centre for Disease Prevention and Control; the European Patients’ Academy on Therapeutic Innovation; the European Vaccine Initiative; Gavi, the Vaccine Alliance; the United Nations International Children's Emergency Fund; and the World Health Organization) and produced animation videos from VACCELERATE.

Table 1. Characteristics of interactive educational cards.

Set	Cards, n	Target group	Aim
1	25	Individuals aged >12 years	Increase literacy about vaccine trials, trial registries, participation in vaccine trials, and participant procedures in trials
2	10	Individuals aged ≥12 years	Increase literacy about COVID-19 vaccination and messenger RNA vaccines
3	7	Individuals aged ≥12 years	Increase literacy about vaccine manufacturing, the vaccine approval process, safety monitoring, and awareness against vaccine misinformation
4	4	All individuals irrespective of age	Improve awareness and information about the VACCELERATE Volunteer Registry
5	12	All individuals irrespective of age	Increase literacy about COVID-19 vaccination, expand awareness for COVID-19 vaccine equity, and utilitarianism

VACCELERATE Puzzles

The VACCELERATE puzzles (Multimedia Appendix 1) were developed using graphic techniques under the guidance of the VACCELERATE participants, including a pediatrician, with focus on younger children (aged <12 years) in order to raise their awareness of clinical trials, vaccines, and participation in trials and to promote inclusiveness. The link to the produced puzzles is provided in Multimedia Appendix 1.

One of the developed puzzles comprises 25 pieces and depicts a variety of human characters (ie, medical staff, minorities, older individuals, and children, among others) obtained from the VACCELERATE toolkit, including the appropriate link to the VACCELERATE Volunteer Registry [30]. The concept herein is to include as many human characters as possible to highlight the importance of inclusiveness and diversity in vaccine trials, especially for underserved population groups such as racial and ethnic minorities and older individuals.

The other 2 smaller puzzles (8 pieces each) were developed to familiarize children with symbols and icons related to the COVID-19 pandemic, as well as to understand and recognize the effectiveness of vaccines in minimizing the risk of infection and harmful effects of COVID-19.

VACCELERATE Animation Videos

Educational Videos (Adult and Pediatric)

The first produced educational video (see Multimedia Appendix 1) was mainly focused on adults and designed using advanced animation techniques, appropriate subtitles based on a storyboard and educational aims, and audio settings (eg, music and a speaker voice). In addition, the video included the following sections: (1) highlight and increase information regarding clinical trials and their necessity and usefulness for public health (eg, good clinical practices, vaccine trial phases and monitoring, and applied safety protocols); (2) the contribution of volunteers (ie, citizens, patient advocacy groups, and underserved populations) to this effort for capacity mapping and building of registries; and (3) volunteer safety, benefits, and any potential risks during a clinical trial.

The second educational video (see Multimedia Appendix 1) was mainly aimed at children and adolescents. Special attention was given to contain informative animations while delivering the learning aim and keeping each video segment short. The produced video included the following sections: (1) provide information about the significance of vaccine trials; (2) share knowledge and information about the high value of vaccination as the most powerful “weapon” to prevent morbidity and mortality associated with infectious diseases; (3) acknowledge the advancements in technology and safety procedures, which have contributed significantly to the development, evaluation, approval, and monitoring of safe and effective vaccines; and (4) highlight the contribution and importance of pediatric volunteers in vaccine clinical trials and ensure their safety and the benefits and risks of participating in clinical trials.

Promotional Video

The promotional video (see Multimedia Appendix 1) presents short, recorded videos by the VACCELERATE work package leaders and the National Coordinator’s pictures to promote, inform, and describe the VACCELERATE Volunteer Registry through specific text scripts. The aim of the video is the inclusiveness and representation of all participating European Union countries and European Union–associated countries of the VACCELERATE Consortium. Currently, we have active representation from 17 countries (Austria, Belgium, Cyprus, Czech Republic, Germany, Greece, Ireland, Israel, Italy, Hungary, Lithuania, the Netherlands, Norway, Portugal, Spain, Sweden, and Turkey) in the VACCELERATE Volunteer Registry.

Discussion

Principal Findings

This paper presents the first set of harmonized promotional and educational tools of the VACCELERATE Volunteer Registry [2]. The produced material, based on the selected population group, includes flyers, posters, interactive educational cards, puzzles, and animation videos. The tools can be downloaded from the official website of VACCELERATE [29] and appropriate references, based on their own DOIs. It should be noted that the provided educational material can be adopted to the user needs and may also be used as base-prototypes for further improvement and production of related tools.

To ensure high rates of vaccination in populations [31] and a large number of participants in vaccine trials, it is necessary to ensure effective, targeted, and trustworthy educational and promotional tools regarding vaccines and vaccine trials, guaranteeing high diversity and inclusiveness in vaccine trials without underrepresentation of minorities.

The developed standardized toolkit is one example of information from trustful sources that can be used against health misinformation and disinformation, lack of trust, fear, and fake news. Furthermore, it can be also used to increase literacy and understanding of the importance of vaccine clinical trials in the general population as part of an emergency preparedness plan.

In addition, the promotional and educational material will be able to bridge the gap in public information (eg, communication approach, target audiences, and type of media produced) regarding COVID-19 vaccine trials, which is covered in only 2.51% of the collected media and only in 4 languages (English, French, German, and Czech) [22]. Another challenge is using visual material and stories in order to inform children about the procedures, importance of, and volunteer participation in vaccine trials and about vaccination.

Limitations

The impact and assessment of the promotional and educational material are beyond the scope of this study. Nevertheless, we are currently in the process of evaluating them for vaccine trials during this year by another expert group (WP4) of the VACCELERATE Consortium.

Conclusions

Communicating with trustworthy information to the public, related to public health issues and vaccine research or trials, constitutes the most important preventive strategy against infectious diseases and threats of future emerging infection diseases.

Future initiatives could involve further improvement of the dissemination process for the developed tools in order to increase the number of volunteer registrations and in particular among population groups, which remain underrepresented in vaccine trials. Another initiative could be the design of an innovative video game [31,32] targeting training and educational material for vaccine clinical trials, addressing vaccine hesitancy and refusal.

Other prospective initiatives could be the conduct of studies based on citizen science methods (eg, community-based participatory action research [28]) for vaccine research [25-27,33] and future epidemics by engaging community members and service providers as partners in the research process and providing them the appropriate educational tools for tackling current and future public health issues. Such actions will eventually contribute to reduce the knowledge gaps and hesitancy regarding participation in vaccine trials in the general population.

Acknowledgments

This project has received funding from the European Union’s Horizon 2020 research and innovation program (grant agreement 101037867; VACCELERATE). The produced educational and promotional tools were developed as part of the VACCELERATE Consortium. For further information, please refer to the official website of the VACCELERATE network [29].

The research leading to these results was conducted as part of the VACCELERATE consortium. For further information please refer to [29].

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information file (Multimedia Appendix 1).

Conflicts of Interest

JS-G has received speaker honoraria from Gilead and Pfizer. The rest of the authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Multimedia Appendix 1
Information on the promotional and educational material including DOI number, format, title, and language.
[DOCX File , 28 KB - publichealth_v9i1e44491_app1.docx]

References

1. VACCELERATE - European Corona Vaccine Trial Accelerator Platform. European Commission. CORDIS EU Research Results. URL: https://cordis.europa.eu/project/id/101037867 [accessed 2023-03-24]
2. Salmanton-García J, Stewart FA, Heringer S, Koniordou M, Álvarez-Barco E, Argyropoulos CD, VACCELERATE consortium. VACCELERATE Volunteer Registry: a European study participant database to facilitate clinical trial enrolment. Vaccine 2022 Jul 29;40(31):4090-4097 [FREE Full text] [doi: 10.1016/j.vaccine.2022.05.022] [Medline: 35659449]
3. Health Emergency Preparedness and Response (HERA). European Commission. URL: https://health.ec.europa.eu/health-emergency-preparedness-and-response-hera_en [accessed 2023-03-24]
4. McCarthy CR. Historical background of clinical trials involving women and minorities. Acad Med 1994 Sep;69(9):695-698. [doi: 10.1097/00001888-199409000-00002] [Medline: 8074757]
5. Jaklevic MC. Researchers strive to recruit hard-hit minorities into COVID-19 vaccine trials. JAMA 2020 Sep 01;324(9):826-828. [doi: 10.1001/jama.2020.11244] [Medline: 32789501]
6. Langford AT. Health communication and decision making about vaccine clinical trials during a pandemic. J Health Commun 2020 Oct 02;25(10):780-789 [FREE Full text] [doi: 10.1080/10810730.2020.1864520] [Medline: 33719877]
7. Finset A, Bosworth H, Butow P, Gulbrandsen P, Hulsman RL, Pieterse AH, et al. Effective health communication - a key factor in fighting the COVID-19 pandemic. Patient Educ Couns 2020 May;103(5):873-876 [FREE Full text] [doi: 10.1016/j.pec.2020.03.027] [Medline: 32336348]
8. Earnshaw V, Eaton L, Kalichman S, Brousseau N, Hill E, Fox A. COVID-19 conspiracy beliefs, health behaviors, and policy support. Transl Behav Med 2020 Oct 08;10(4):850-856 [FREE Full text] [doi: 10.1093/tbm/ibaa090] [Medline: 32910819]
9. MacDonald N, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: Definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
10. Gottlieb M, Dyer S. Information and disinformation: social media in the COVID-19 crisis. Acad Emerg Med 2020 Jul 24;27(7):640-641 [FREE Full text] [doi: 10.1111/acem.14036] [Medline: 32474977]
11. Naeem SB, Bhatti R, Khan A. An exploration of how fake news is taking over social media and putting public health at risk. Health Info Libr J 2021 Jun 12;38(2):143-149 [FREE Full text] [doi: 10.1111/hir.12320] [Medline: 32657000]
12. Tannenbaum MB, Hepler J, Zimmerman RS, Saul L, Jacobs S, Wilson K, et al. Appealing to fear: A meta-analysis of fear appeal effectiveness and theories. Psychol Bull 2015 Nov;141(6):1178-1204 [FREE Full text] [doi: 10.1037/a0039729] [Medline: 26501228]

13. Gisondi MA, Barber R, Faust JS, Raja A, Strehlow MC, Westafer LM, et al. A deadly infodemic: social media and the power of COVID-19 misinformation. J Med Internet Res 2022 Feb 01;24(2):e35552 [FREE Full text] [doi: 10.2196/35552] [Medline: 35007204]
14. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
15. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
16. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic [accessed 2023-03-24]
17. Lee SK, Sun J, Jang S, Connelly S. Misinformation of COVID-19 vaccines and vaccine hesitancy. Sci Rep 2022 Aug 11;12(1):13681 [FREE Full text] [doi: 10.1038/s41598-022-17430-6] [Medline: 35953500]
18. Surajit B. Vaccine hesitancy and misinformation about clinical trials. Indian J Surg 2021 Jun 24;83(3):615-616 [FREE Full text] [doi: 10.1007/s12262-021-03011-w] [Medline: 34188367]
19. Ginossar T, Diaz Fuentes C, Oetzel J. Understanding willingness to participate in cancer clinical trials among patients and caregivers attending a minority-serving academic cancer center. J Cancer Educ 2022 Feb 14;37(1):179-187. [doi: 10.1007/s13187-020-01802-5] [Medline: 32666504]
20. Habr D, Ferdinand R. Addressing racial/ethnic disparities in cancer clinical trials: everyone has a role to play. Cancer 2021 Sep 15;127(18):3282-3289 [FREE Full text] [doi: 10.1002/cncr.33600] [Medline: 33904590]
21. Jennings W, Stoker G, Bunting H, Valgarðsson VO, Gaskell J, Devine D, et al. Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy. Vaccines (Basel) 2021 Jun 03;9(6):593 [FREE Full text] [doi: 10.3390/vaccines9060593] [Medline: 34204971]
22. Luís C, Di Marzo VR, Kaur M, Argyropoulos C, Devane D, Stewart FA, et al. Vaccination rates in Europe are not associated with online media intensity. JCOM 2022;21(05):A05 [FREE Full text] [doi: 10.22323/2.21050205]
23. DiClemente R, Crosby R, Kegler M. Understanding and applying theory in health promotion practice and research. In: Emerging theories in health promotion practice and research strategies for improving public health. San Francisco, CA: Jossey-Bass; 2002.
24. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
25. Beatty AL, Peyser ND, Butcher XE, Carton TW, Olgin JE, Pletcher MJ, et al. The COVID-19 Citizen Science Study: protocol for a longitudinal digital health cohort study. JMIR Res Protoc 2021 Aug 30;10(8):e28169 [FREE Full text] [doi: 10.2196/28169] [Medline: 34310336]
26. Marcus GM, Olgin JE, Peyser ND, Vittinghoff E, Yang V, Joyce S, et al. Predictors of incident viral symptoms ascertained in the era of COVID-19. PLoS One 2021 Jun 17;16(6):e0253120 [FREE Full text] [doi: 10.1371/journal.pone.0253120] [Medline: 34138915]
27. Paul K, Palfinger T. Walking the (argumentative) talk using citizen science: involving young people in a critical policy analysis of vaccination policy in Austria. Evid Policy 2020;16(2):229-247. [doi: 10.1332/174426419X15752578285791]
28. Hills M, Mullett J, Carroll S. Community-based participatory action research: transforming multidisciplinary practice in primary health care. Rev Panam Salud Publica 2007;21(2-3):125-135. [doi: 10.1590/s1020-49892007000200007] [Medline: 17565799]
29. VACCELERATE Official Website. URL: https://vaccelerate.eu/ [accessed 2023-03-24]
30. Volunteer-Registry. Vaccelerate. URL: https://vaccelerate.eu/volunteer-registry-2/ [accessed 2023-03-24]
31. Ohannessian R, Yaghobian S, Verger P, Vanhems P. A systematic review of serious video games used for vaccination. Vaccine 2016 Aug 31;34(38):4478-4483 [FREE Full text] [doi: 10.1016/j.vaccine.2016.07.048] [Medline: 27498211]
32. Ricciardi F, De Paolis LT. A comprehensive review of serious games in health professions. Int J Comput Games Technol 2014;2014:1-11. [doi: 10.1155/2014/787968]
33. Lin A, Vittinghoff E, Olgin J, Peyser N, Aung S, Joyce S, et al. Predictors of incident SARS-CoV-2 infections in an international prospective cohort study. BMJ Open 2021 Sep 21;11(9):e052025 [FREE Full text] [doi: 10.1136/bmjopen-2021-052025] [Medline: 34548363]

Abbreviations

DOI: Digital Object Identifier
ECDC: European Centre for Disease Prevention and Control
HERA: Health Emergency Preparedness and Response Authority

Edited by A Mavragani, T Sanchez; submitted 21.11.22; peer-reviewed by T Ginossar, B Eshrati; comments to author 20.01.23; revised version received 09.02.23; accepted 28.02.23; published 03.04.23.

Please cite as:
Argyropoulos CD, Leckler J, Salmanton-García J, Constantinou M, Alexandrou A, Themistocleous S, Noula E, Shiamakkides G, Nearchou A, Stewart FA, Albus K, Koniordou M, Kopsidas I, Spivak O, Hellemans M, Hendrickx G, Davis RJ, Azzini AM, Simon PV, Carcas-Sansuan AJ, Askling HH, Vene S, Prellezo JB, Álvarez-Barco E, Macken AJ, Di Marzo R, Luís C, Olesen OF, Frias Iniesta JA, Barta I, Tóth K, Akova M, Bonten MMJ, Cohen-Kandli M, Cox RJ, Součková L, Husa P, Jancoriene L, Launay O, Lundgren J, Mallon P, Armeftis C, Marques L, Naucler P, Ochando J, Tacconelli E, van Damme P, Zaoutis T, Hofstraat S, Bruijning-Verhagen P, Zeitlinger M, Cornely OA, Pana ZD
Enhancing Public Health Communication Regarding Vaccine Trials: Design and Development of the Pan-European VACCELERATE Toolkit
JMIR Public Health Surveill 2023;9:e44491
URL: https://publichealth.jmir.org/2023/1/e44491
doi:10.2196/44491
PMID:36878478

©Christos D Argyropoulos, Janina Leckler, Jon Salmanton-García, Marinos Constantinou, Alexandra Alexandrou, Sophia Themistocleous, Evgenia Noula, George Shiamakkides, Andria Nearchou, Fiona A Stewart, Kerstin Albus, Markela Koniordou, Ioannis Kopsidas, Orly Spivak, Margot Hellemans, Greet Hendrickx, Ruth Joanna Davis, Anna Maria Azzini, Paula Valle Simon, Antonio Javier Carcas-Sansuan, Helena Hervius Askling, Sirkka Vene, Jana Baranda Prellezo, Elena Álvarez-Barco, Alan J Macken, Romina Di Marzo, Catarina Luís, Ole F Olesen, Jesus A Frias Iniesta, Imre Barta, Krisztina Tóth, Murat Akova, Marc M J Bonten, Miriam Cohen-Kandli, Rebecca Jane Cox, Lenka Součková, Petr Husa, Ligita Jancoriene, Odile Launay, Jens Lundgren, Patrick Mallon, Charis Armeftis, Laura Marques, Pontus Naucler, Jordi Ochando, Evelina Tacconelli, Pierre van Damme, Theoklis Zaoutis, Sanne Hofstraat, Patricia Bruijning-Verhagen, Markus Zeitlinger, Oliver A Cornely, Zoi Dorothea Pana. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 03.04.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

COVID-19 Contact Tracing as an Indicator for Evaluating a Pandemic Situation: Simulation Study

Manuel Marques-Cruz[1,2,3,4], MSc, MD; Diogo Nogueira-Leite[1,2,3], MSc; João Miguel Alves[1,2,3], MSc, PharmD; Francisco Fernandes[5], MSc, MD; José Miguel Fernandes[6], MSc, MD; Miguel Ângelo Almeida[7], MD; Patrícia Cunha Correia[8], MSc, MD; Paula Perestrelo[3], MSc, MD; Ricardo Cruz-Correia[1,2], PhD; Pedro Pita Barros[3], PhD

[1]Department of Community Medicine, Information and Decision in Health, Faculty of Medicine, University of Porto, Porto, Portugal

[2]Center for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal

[3]NOVA Health Economics & Management Knowledge Center, NOVA School of Business and Economics, NOVA University of Lisbon, Lisboa, Portugal

[4]Public Health Unit Marão e Douro Norte, Agrupamento de Centros de Saúde Marão e Douro Norte, Administração Regional de Saúde do Norte, Vila Real, Portugal

[5]Public Health Unit Baixo Mondego, Agrupamento de Centros de Saúde Baixo Mondego, Administração Regional de Saúde do Centro, Figueira da Foz, Portugal

[6]Public Health Unit Alto Ave, Agrupamento de Centos de Saúde Alto Ave, Administração Regional de Saúde do Norte, Fafe, Portugal

[7]Public Health Unit Póvoa de Varzim, Agrupamento de Centros de Saúde Póvoa de Varzim - Vila do Conde, Administração Regional de Saúde do Norte, Vila do Conde, Portugal

[8]Public Health Unit of Local Health Unit of Guarda, Unidade Local de Saúde da Guarda, Guarda, Portugal

Corresponding Author:
Manuel Marques-Cruz, MSc, MD
Department of Community Medicine, Information and Decision in Health
Faculty of Medicine
University of Porto
Faculdade de Medicina da Universidade do Porto
Rua Dr. Plácido da Costa
Porto, 4200-450
Portugal
Phone: 351 225513622
Email: up201000048@up.pt

Abstract

Background: Contact tracing is a fundamental intervention in public health. When systematically applied, it enables the breaking of chains of transmission, which is important for controlling COVID-19 transmission. In theoretically perfect contact tracing, all new cases should occur among quarantined individuals, and an epidemic should vanish. However, the availability of resources influences the capacity to perform contact tracing. Therefore, it is necessary to estimate its effectiveness threshold. We propose that this effectiveness threshold may be indirectly estimated using the ratio of COVID-19 cases arising from quarantined high-risk contacts, where higher ratios indicate better control and, under a threshold, contact tracing may fail and other restrictions become necessary.

Objective: This study assessed the ratio of COVID-19 cases in high-risk contacts quarantined through contact tracing and its potential use as an ancillary pandemic control indicator.

Methods: We built a 6-compartment epidemiological model to emulate COVID-19 infection flow according to publicly available data from Portuguese authorities. Our model extended the usual susceptible-exposed-infected-recovered model by adding a compartment Q with individuals in mandated quarantine who could develop infection or return to the susceptible pool and a compartment P with individuals protected from infection because of vaccination. To model infection dynamics, data on SARS-CoV-2 infection risk (IR), time until infection, and vaccine efficacy were collected. Estimation was needed for vaccine data to reflect the timing of inoculation and booster efficacy. In total, 2 simulations were built: one adjusting for the presence and absence of variants or vaccination and another maximizing IR in quarantined individuals. Both simulations were based on a set of 100 unique parameterizations. The daily ratio of infected cases arising from high-risk contacts (q estimate) was calculated. A theoretical effectiveness threshold of contact tracing was defined for 14-day average q estimates based on the classification of

COVID-19 daily cases according to the pandemic phases and was compared with the timing of population lockdowns in Portugal. A sensitivity analysis was performed to understand the relationship between different parameter values and the threshold obtained.

Results: An inverse relationship was found between the *q* estimate and daily cases in both simulations (correlations >0.70). The theoretical effectiveness thresholds for both simulations attained an alert phase positive predictive value of >70% and could have anticipated the need for additional measures in at least 4 days for the second and fourth lockdowns. Sensitivity analysis showed that only the IR and booster dose efficacy at inoculation significantly affected the *q* estimates.

Conclusions: We demonstrated the impact of applying an effectiveness threshold for contact tracing on decision-making. Although only theoretical thresholds could be provided, their relationship with the number of confirmed cases and the prediction of pandemic phases shows the role as an indirect indicator of the efficacy of contact tracing.

(JMIR Public Health Surveill 2023;9:e43836) doi:10.2196/43836

KEYWORDS

COVID-19; public health; public health surveillance; quarantine; infection transmission; epidemiological models

Introduction

Background

Contact tracing is a fundamental activity in public health and is the process of identifying, triaging, and monitoring individuals exposed to a communicable disease to prevent secondary transmission [1]. When systematically applied, contact tracing leads to the breaking of transmission chains, and its valuable role in controlling COVID-19 transmission is widely recognized in the context of low transmission and active community transmission [2]. It is also acknowledged that the earlier the identification of infection cases and individuals to trace, the higher the likelihood of the pandemic situation being controlled even in instances where contact identification is incomplete [1,3,4].

In Portugal, contact tracing and the imposing of quarantine measures are tasks specific to public health units [5]. The capacity of local public health units to perform contact tracing is constrained by the availability of technological and workforce resources [3]. Several simulation studies have demonstrated that contact tracing could only be effective if a combination of high adherence to quarantine measures, minor delays from symptom onset to isolation of cases, and an increased number of contacts traced occurred [3,6-8].

Such conditions were more likely to be verified in moments with a lower number of confirmed cases (ie, corresponding to the troughs of the epidemic curve) [9]. Moreover, the initial Portuguese guidelines for tackling the pandemic predicted that contact tracing efforts would be abandoned when the pandemic entered a community transmission phase [10]. Therefore, it should be deemed necessary to estimate a threshold beyond which contact tracing and potential individual quarantine measures cease to be effective in breaking infection chains and, consequently, diminishing SARS-CoV-2 infection rates in the community [11]. This information may be used as a decision-making support tool for imposing generalized population containment measures and offers important lessons concerning population control of viral infections with characteristics similar to those of SARS-CoV-2 [12].

The effectiveness of contact tracing as a mechanism for breaking transmission chains occurs because of the quarantine of contacts of confirmed infected cases. Should these contacts develop the disease, they will not transmit it to other community members [1]. Thus, the hypothetical identification of all contacts of all SARS-CoV-2–infected individuals would necessarily lead to the breaking of all transmission chains. In other words, in a scenario where all the contacts of all infection cases were identified, new infection cases would only occur in those contacts. Hence, the effectiveness threshold of a contact tracing and quarantine strategy could be derived from the proportion of infection cases that arise in contacts of confirmed cases.

Hence, the effectiveness threshold of contact tracing is the point at which its utility as a health intervention that includes quarantine in controlling and breaking transmission chains is defined. The proportions of confirmed cases from quarantined individuals above this threshold could be indicative of effective contact tracing, and the pandemic combat strategy could rely mainly on this intervention. In contrast, proportions below the threshold could indicate that contact tracing is not effective and that further interventions may be necessary to stop transmission.

The proportions of confirmed cases from quarantined individuals lower than the effectiveness threshold may have occurred at different stages throughout the pandemic, mainly in periods of case surges, situations in which there was a need to implement more restrictive measures (namely, general or selective confinements). Imposing confinements has been demonstrated to affect pandemic control regarding case numbers, hospital admissions, and deaths because of COVID-19 when implemented at least 14 days before the peak of a case surge [13,14]. In Portugal, the decision to impose a population lockdown was primarily based on the 14-day incidence rate; the transmissibility rate (Rt); and, more recently, the critical care bed occupancy rate [15].

Despite the public availability of several global databases on COVID-19, especially concerning the number of new cases, deaths, tests performed, and vaccination data, there is no information on the number of infected cases coming from individuals identified as high-risk contacts through contact tracing. Furthermore, data on quarantined individuals are scarce. Portuguese data were until recently an exception in that the Directorate-General of Health (DGS) reported in a daily bulletin on COVID-19 the number of high-risk contacts identified, defined as individuals in quarantine by mandate from health authorities [16]. The Data Science for Social Good (DSSG) initiative developed a data repository [17] for COVID-19 in which it gathered, compiled, and curated the data made public by the DGS in these bulletins.

Infection dynamics and the effectiveness of contact tracing are not only influenced by the vaccinated population (for which the DSSG also kept curated data). They may also be affected by the prevalence of different SARS-CoV-2 variants at different points in time. The European Centre for Disease Prevention and Control keeps a data repository on the prevalence of different variants in Europe reported through The European Surveillance System in an open-access database [18].

Objectives

The aim of this study was to identify the ratio of COVID-19 cases that occurred in individuals in quarantine mandated by health authorities and the potential use of this ratio as a proof of concept of an indicator for assessing pandemic control in parallel with other established criteria such as the transmissibility index (Rt) and incidence.

Methods

Study Design

In this study, we collected data on COVID-19 for Portugal and built an expanded structure to a susceptible-exposed-infected -recovered (SEIR) compartmental model to emulate the pandemic. All compartment data came from collected data except for those of protected individuals. The purpose of the model was to estimate the daily number of quarantined individuals who became infected and the daily number of susceptible and vaccinated individuals who were quarantined as data regarding those values were lacking.

We input different values for each known strain of SARS-CoV-2 (infection risk [IR] and time until infection and until the end of quarantine) and immunity from vaccination and ran the model for 690 days. Consequently, we estimated the daily ratio of cases arising from daily quarantined individuals (*q* estimate) using the number of daily confirmed cases as the denominator. We measured the correlation between this *q* estimate and data from confirmed cases.

According to the epidemic case curve, we defined 3 pandemic phases in the Portuguese data: interpandemic phase, alert phase, and pandemic phase. As a proof of concept, we estimated the best hypothetical cutoff for our *q* estimates to distinguish the interpandemic and alert phases and compared that theoretical cutoff with the timing of population confinement measures. Finally, we performed a 2-part sensitivity analysis. Initially, we ran a multiple linear regression on the thresholds that each set of parameters conveyed to assess how each parameter would change the threshold value. In addition, we fixed the maximum and minimum values for each parameter and measured the correlation with the main simulation results and threshold values obtained by changing all other parameters.

Table 1 includes all model inputs, values, and sources of each input. We also describe the only 2 outputs of the model, namely, the number of daily quarantined individuals who develop infection and the theoretical *q* estimate.

The compartment transition dynamics were in accordance with the following equations (note that only equations 1 and 3 were estimated, corresponding to compartments S and E, for which real data were not available. All other compartment data were directly collected from official reports [17]):

$$S' = \kappa Q + \rho P - (\phi + \psi + \varepsilon)S \quad \textbf{(1)}$$

$$Q' = \phi S + \chi P - (\gamma + \kappa + \varsigma)Q \quad \textbf{(2)}$$

$$E' = I - \gamma Q \quad \textbf{(3)}$$

$$I' = \gamma Q + \iota E - \pi I \quad \textbf{(4)}$$

$$R = \pi I \quad \textbf{(5)}$$

Table 1. Data sources and model inputs and outputs.

Parameter	Value	Source
Susceptible (S)	Base case: 10 million	Estimated
Protected through vaccination (P)	P (Protection) × V	Estimated
Quarantined (Q)	Base case: 0	DSSG[a] [17]
Exposed not traced (E)	Base case: 0	Estimated
Infected (I)	Base case: 2	DSSG [17]
Recovered from infection (R)	Base case: 0	DSSG [17]
Vaccinated (V)	Base case: 0	DSSG [17]
Inoculation vaccine efficacy (Ve_0)	10%-60% (uniform distribution)	Hall et al [19] and Polack et al [20]
Maximum vaccine efficacy (Ve_{max})	Ve_0-95% (uniform distribution)	Hall et al [19] and Polack et al [20]
Waned vaccine efficacy (Ve_{waned})	$(Ve_0 + Ve_{max})/2$	Estimated
Inoculation booster dose efficacy (Be_0)	Ve_{waned}-80%	Estimated
Maximum booster dose efficacy (Be_{max})	Be_0-95%	Estimated
Waned booster dose efficacy (Be_{waned})	$(Be_0 + Be_{max})/2$	Estimated
Time until maximum efficacy (Δte_{max})	15 d	Hall et al [19] and Polack et al [20]
Time until waned efficacy (Δte_{waned})	180 d	Hall et al [19]
Variant prevalence	Base case: other=100%; alpha, beta, gamma, delta, and omicron=0%	ECDC[b] [18]
Maximum time until γ ($\Delta t\gamma_{max}$)	2-14 days (uniform distribution)	DGS[c] [5,21] and Wu et al [22]
Average time until γ ($\Delta t\gamma_{mean}$)	2-$\Delta t\gamma_{max}$ days (uniform distribution)	DGS [5,21] and Wu et al [22]
Maximum time until κ and ς ($\Delta t\kappa\varsigma_{max}$)	2-14 days (uniform distribution)	Estimated
Average time until κ and ς ($\Delta t\kappa\varsigma_{mean}$)	2-$\Delta t\kappa\varsigma_{max}$ days (uniform distribution)	Estimated
IR[d] simulation A	10%-50% (uniform distribution)	ECDC [1], Karumanagoundar et al [23], and Tang et al [24]
IR simulation B	0.1%-2.5%	Calibration
Quarantined infected	γQ	Estimated
Ratio of cases from quarantined (*q*)	$\gamma Q/(\gamma Q + \iota E)$	Estimated

[a] DSSG: Data Science for Social Good.

[b] ECDC: European Centre for Disease Prevention and Control.

[c] DGS: Directorate-General of Health.

[d] IR: infection risk.

Data Collection

The sources of data are described in full in Table 1. The primary data source to meet this study's aim was the DSSG COVID-19 data repository [17]. The data included the number of confirmed cases, daily new cases, people under surveillance, people fully vaccinated, and individuals with vaccine booster doses (using any of the vaccines available in Portugal) between the first confirmed case of COVID-19 in Portugal (March 2, 2020) and January 20, 2022. Open-access data from The European Surveillance System regarding the weekly prevalence of SARS-CoV-2 variants were also collected through the European Centre for Disease Prevention and Control open repository [18] for the same period. In the absence of specific data on the number of cases of COVID-19 among high-risk contacts, we built a simulation model to calculate the ratio of daily cases from those contacts using the available data.

Epidemiological Compartmental Model

Data were inserted into an epidemiological model based on compartmental models already applied to COVID-19 and other epidemiological contexts [25,26]. Several expanded models to an SEIR model have been attempted, including either one or more quarantine compartments and one or more protected individuals through vaccination [27]. In addition, the concept of high-risk exposure and exposed individuals (frequently included in compartment E) did not entirely comprise the DGS definition of high-risk contact in that, in all models, an exposed individual could not return to being susceptible [28]. Furthermore, in most models, a quarantined individual could have come from being susceptible and either return to being susceptible or progress to being infected [29]. Other quarantine definitions, different from the idea of prophylactic isolation, came from infected individuals [30].

Moreover, 2 transitions have not been reported in the reviewed literature on expanded SEIR models, namely, the direct transition between compartments S and I and the transition between vaccinated or protected compartments and quarantine or exposure compartments. As our model needed to comprise these transitions, several differences from already published models had to be introduced. To keep with the base SEIR structure, compartment E included only individuals who would progress to compartment I without previous contact tracing. The model was extended with compartments Q and P. Figure 1 illustrates the compartmental model used and the transitions between compartments and subcompartments.

Compartment S (susceptible) is the initial compartment of the model (ie, the starting point for all individuals). From compartment S, individuals can progress to compartments P, Q, and E.

Compartment P (protected) refers to the group of individuals who are immune to SARS-CoV-2 infection because of vaccination. Although the main effect of vaccines is protection against severe disease and not protection against infection [20], short-term efficacy in infection prevention has been proven [19]. This preventive effect is considerably reduced 6 months after the date of inoculation. When classified as high-risk contacts through contact tracing, individuals in this compartment could be quarantined regardless of their vaccination status in line with Portuguese norms that only lifted these compulsory measures for fully vaccinated individuals on January 10, 2022 [5].

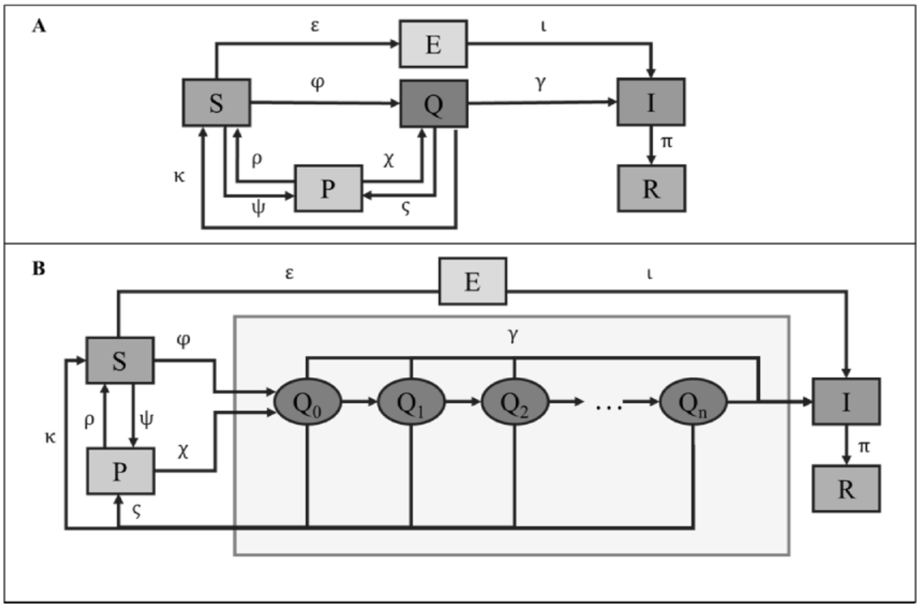
Compartment Q (quarantined or exposed with tracing) includes all high-risk contacts of SARS-CoV-2–infected individuals for whom quarantine was mandated by a public health authority according to norms and guidelines [5] and the definition by the Portuguese DGS as “individuals under surveillance” applicable at each moment. Individuals in this compartment could return to compartment S. This compartment is further divided into subcompartments, each representing a day in which an individual stayed in the exposed compartment. Thus, the transitions ϕ and κ, χ and ς, and γ are the pooled transitions from each subcompartment inside compartment Q and compartments S, P, and I, respectively.

Compartment E (exposed without tracing) includes exposed individuals who will develop COVID-19 and have not been traced by health authorities. Compartment I consists of all the confirmed cases of SARS-CoV-2 infection.

Compartment R (recovered) includes all individuals who recovered from infection. Compartment R is the terminal compartment as the model is not circular. This is because an individual previously infected with COVID-19 cannot be considered a high-risk contact before 180 days have passed from the date of infection or when not presenting symptoms suggesting SARS-CoV-2 infection [21]. In other words, until the natural immunity gained from infection fades, there can be no transition between the recovered (R) and exposed (E) or quarantined (Q) compartments. When such immunity wanes, the individual should return to compartment S. However, and assuming a large enough number of individuals in compartment S (equal to or greater than the total number of confirmed cases in the analysis period), there is no need to consider transitions between compartments R and S—even in case of reinfection, each new case of COVID-19 can be regarded as a case from a different individual.

The daily ratio of cases arising from individuals in quarantine mandated by health authorities was given by the quotient of the daily transitions between compartments Q and I (represented by γ) and the daily total confirmed cases.

Figure 1. Compartmental model without subcompartments (A) and with subcompartments (B). E: individuals exposed to SARS-CoV-2 who will develop COVID-19 and have not been traced by health authorities; I: infected individuals; P: subset of susceptible individuals protected from SARS-CoV-2 infection through vaccination; Q: quarantined individuals; Q0, Q1, Q2,..., Qn: subcompartments of the quarantine compartment (each number represents the number of days since exposure); R: recovered individuals; S: susceptible individuals.



Parameterization and Transitions

The compartmental model was run for 690 days (from March 2, 2020, to January 20, 2022). The data collected populated compartments Q, I, and R daily. In the absence of data for compartment S, an initial value of 10 million individuals was assumed, which is a frequently used approximation for the Portuguese population. This value ensured that the model was not circular and obviated the need to consider the possibility of reinfection. Data on compartments P and E were calculated, and expected values for other parameters were defined within each simulation run (Table 1).

Regarding daily transitions between compartments and subcompartments, the model was built in 3 steps: transitions from subcompartments inside Q (step 1), S-E-I transitions (step 2), and S-Q and P-Q transitions (step 3). Despite the existence of data for all compartments and some model parameters, owing to a lack of data in each of these steps regarding the actual number of individuals that transition in each iteration of the model, some parameters were input by defining expected values for the parameters that govern those transitions. Only those parameters were estimated (Table 1).

In step 1, Q-S, Q-P, and Q-I transitions are the sums of the number of individuals who each day (represented by each subcompartment) transition to compartments S, P, or I, respectively. Each day, transitions from each subcompartment are governed by the probabilities of staying in compartment Q (transition to the following subcompartment Q_{i+1}), making the transition γ to compartment I, or returning to compartment S or P (transitions κ and ς). The probabilities of transition of each subcompartment were defined according to the IR of a high-risk contact of a confirmed COVID-19 case and to the mean and maximum periods an individual remains exposed until the person either develops the infection or is considered susceptible again. Any of these 5 parameters can vary with the viral variant to which the individuals were exposed and with the quarantine period. Therefore, in the context of the described model, the probability of transitioning from any subcompartment to compartment S, P, or I is given by the sum of the product of the transition probability of each specific variant and the number of individuals exposed to that variant. Step 1 is concluded after calculating the ratio of new daily cases of infection that came from compartment Q (resulting from transition γ) to the total number of new daily cases (*q* estimate).

Steps 2 and 3 aim to keep the model closed (a necessary condition for using compartmental models), allowing the model to simulate the following day. The number of individuals that follow the ε and ι transitions in step 2 is given by the difference between the number of new daily infections and the number of individuals in transition γ. The ϕ and χ transitions in step 3 are calculated as the difference between the number of exposed individuals (high-risk contacts) on the following simulation day and the number of individuals inside compartment Q after steps 1 and 2. The value of this transition corresponds to the number of individuals inside subcompartment Q_0 at the beginning of step 1 on the following simulation day.

Finally, compartment P was defined according to 5 parameters, namely, vaccine efficacy after inoculation, maximum efficacy, and efficacy after a waning period as well as the elapsed time between vaccination and maximum efficacy and between maximum efficacy and waned efficacy. This compartment was calculated as the product of the probability of each vaccinated individual being protected from SARS-CoV-2 infection and the number of vaccinated individuals.

Assumptions of the Simulation

To operationalize the proposed model, several assumptions were made, namely regarding the transitions involving compartments Q and P.

A Markov chain Monte Carlo simulation defined the transitions between the Q subcompartments. Consequently, to compute the transitions between subcompartments, a series of steps were taken (Textbox 1).

Regarding compartment P, and particularly vaccine efficacy, a constant rate was assumed between two events: (1) efficacy after inoculation and maximum efficacy and (2) maximum efficacy and waned efficacy. Hence, vaccinated individuals were probabilistically placed in compartment P each day according to the time elapsed since inoculation. After defining new values for booster efficacy, the same procedure was replicated for individuals with booster doses. Individuals in compartment P could transition to compartment Q, but their IR was set at 0. Individuals in compartments S and P would transition to compartment Q according to the ratio between those compartments.

Owing to the inherent variability in the timing of contact tracing and imposing quarantine, it is impossible to accurately define certain parameters in the model for each variant of SARS-CoV-2 and each vaccine. Therefore, we decided to simulate a set of different parameterizations for these parameters. For each variant, the maximum time elapsed for the transition to compartments I and S was assumed to range from 2 to 14 days. The mean time for these transitions ranged from 2 days to the maximum time previously set for each variant. IR was set to range from 10% to 50% [1,23,24], and we considered 6 different variants, namely, “alpha,” “beta,” “gamma,” “delta,” “omicron,” and “others” (according to collected data). We assumed that the weekly prevalence of each variant [18] was equal to the prevalence on each day of that week, and individuals in the ϕ transition would be distributed to each variant according to that prevalence.

Regarding vaccination, we defined an efficacy between 10% and 60% after inoculation and a maximum efficacy of up to 95% [19,20]. We assumed a waned efficacy as the mean value between both efficacies. Booster vaccines had an assumed efficacy between the waned efficacy previously defined and 80% after inoculation and a maximum efficacy of up to 95%. A period of 15 days until maximum vaccine efficacy, and a period of 180 days until waned efficacy were assumed [19,20]. Individuals from compartment P entered the model as exposed to variant “others” with an IR set at 0%—in other words, they were set to follow the Poisson distribution estimated for returning to compartments S and P for the variant “others.”

Textbox 1. Steps taken to compute the transitions between subcompartments.

- To calculate the different transition probabilities, we assumed that everyone who transitions to compartment Q has a predetermined probability of becoming infected, called infection risk (IR). IR is specific to each variant.
- For each variant included, 2 different Poisson distributions were applied to the individuals who would become infected and those who would return to compartments S and P. Each distribution had a maximum and mean time until transition.
- From the combined transition probabilities of each distribution and according to the IR, a transition table was constructed, including the pooled transition probabilities from each subcompartment to the following subcompartment, to compartments S and P, and to compartment I.
- A Monte Carlo simulation was then applied to this transition table. Therefore, on each simulation day, an individual in a subcompartment of the Q compartment transitions probabilistically to the following subcompartment, to compartments S and P, or to compartment I.

Description of the Simulations

In total, 2 different simulations were performed, approaching the objective in 2 different ways. Table 2 summarizes the main characteristics and differences between the simulations. We first defined that each simulation should have a run time of ≤9 hours. An iteration of the model was tested on a computer with an Intel64 Family 6 Model 126 Stepping 5 GenuineIntel processor with a maximum clock frequency of 1498.0 MHz with 4 physical cores, 8 logical cores, and 15.60 GB of RAM using the operating system Windows 10.0.19041. With this hardware and software combination, the response time (time taken from start to end of the process) for this iteration was 10 seconds [31].

Simulation A aimed to be as approximate as possible to what we could expect in real life by accounting for the potential influence of variants and vaccination. However, it assumed that the risk of infection was fixed for all variants at the beginning of the simulation. For this simulation, 100 parameterizations were established, and these were quadrupled to reflect the presence and absence of the influence of variants and vaccination. Each parameterization was iterated 100 times, totaling 40,000 model iterations with an estimated run time of 8 hours and 20 minutes.

Simulation B ignored the presence of different variants (we assumed that all cases came from the variant “others”) and estimated the maximum value of IR for which the model kept the q estimate at <1. However, as in simulation A, the risk of infection in exposed individuals was fixed for each parameterization and did not fluctuate throughout each iteration. This simulation used the same set of 100 parameters used in simulation A, doubled to account for the presence or absence of vaccination. The IR started at 100% with each iteration interrupted, and this value was reset and diminished by 0.1 percentage points every time the daily q estimate was >1. We assumed that this approach might lead, in extremis, to circa 1000 iterations of each parameterization. Although mostly incomplete because of the described procedure to calibrate the IR, this number of iterations would lead model B to take 20 times longer to run than model A. Therefore, the research team opted to conduct only 10 iterations of each parameterization, which, jointly with half of the total parameterizations of simulation A, kept the estimated simulation running time at 8 hours and 20 minutes.

Table 2. Main characteristics of the implemented simulations.

	Simulation A	Simulation B
Used information on vaccines	Yes	Yes
Used information on variants	Yes	No
Number of different parameterizations, n	400 (100 × 4)	200 (100 × 2)
Ceiling value for the ratio of cases (*q* estimate)	No	Yes
Parameters included fixed infection risk	Yes	Yes
Number of iterations per parameterization, n	100	Approximately 1000
Total model iterations, n	40,000	200,000
Estimated run time	8 h and 20 min	8 h and 20 min

Data Analysis

The parameters' characteristics were aggregated in means and medians and described for each simulation. For each simulation, the distribution of the ratio of daily cases resulting from the γ transition (*q* estimate) for each parameterization was combined into a single distribution according to the methodology implemented by Hill [32]. We also combined the results of each simulation into a single distribution according to the presence or absence of both vaccination and different variants in the model when applicable. Although the *q* estimate intended to represent the proportion of cases arising from high-risk contacts, certain parameter combinations yielded values of the *q* estimate of >1 (Textbox 2).

Since other COVID-19 measures for lockdown were presented in a 14-day moving average, we applied a 14-day moving average to the daily *q* estimate for each simulation and to the gathered data on daily cases. The Spearman correlation was calculated between the 14-day *q* estimate and the 14-day average number of cases.

Local maxima for 14-day data on confirmed infected cases were computed, and local minima immediately before and after the computed local maxima were maintained. By using an adaptation of the method used by Vázquez-Seisdedos et al [33] and applied by Gianquintieri et al [34] to data on COVID-19 cases, we determined the inflection point of the case curve that occurs between the local minimum and the local maximum. This inflection point corresponds to the beginning of a pandemic phase. Each pandemic phase was defined, in the context of this study, as the interval between an inflection point and the following local minimum, which had a local maximum between them.

It was assumed that, by definition, a contact tracing effectiveness threshold would have to be surpassed during a pandemic phase. Specifically, the value of the effectiveness threshold would have to occur in an alert phase, defined as the critical period between the wave's inflection point and up until 14 days before the day when the peak number of cases was registered. In addition, it was assumed that contact tracing as a main strategy for pandemic combat was reset after the end of each pandemic phase (ie, local minimum after case peak) and the model entered an interpandemic phase. For this reason, an effectiveness threshold was calculated considering the interpandemic and alert phases. In other words, we included the entire period analyzed except for periods between the end of an alert phase and the end of each pandemic phase. The data between a local minimum and the next inflection point (interpandemic phase) were considered as controls for determining the threshold. As there was a more considerable period of control data than the period in the alert phase, we oversampled the test data. This classification into interpandemic, alert, and pandemic phases was loosely based on the World Health Organization pandemic phases for influenza [35].

This threshold estimation from the *q* estimates was a theoretical demonstration of the impact of applying the same procedure to the actual proportion of cases arising from high-risk contacts in anticipation of the need for populational lockdown measures that were ultimately imposed.

Effectiveness threshold values were estimated for each simulation using a receiver operating characteristic model applying the Youden method. The sensitivity, specificity, and positive predictive value (PPV) were calculated for the theoretical effectiveness threshold. Sensitivity represented the proportion of days in the alert phase in which the *q* estimate was below the effectiveness threshold. Specificity represented the proportion of days in the interpandemic phase in which the *q* estimate was above the effectiveness threshold. The PPV was the quotient of the number of alert phase days in which the *q* estimate was below the effectiveness threshold and the number of days in which the *q* estimate was below the effectiveness threshold.

Finally, we implemented a 2-part sensitivity analysis. First, we estimated a hypothetical threshold for each of the different parameterizations used (both in simulations A and B) and ran a multiple linear regression with the estimated thresholds as the dependent variable. We included 10% increases in IR of a high-risk contact of a confirmed COVID-19 case and mean and maximum periods during which an individual remains quarantined until they either develop the infection or are considered susceptible in days, for each strain, and 10% increases in maximum and waned efficacies for complete vaccination and booster doses as independent variables. The second part consisted of fixing the maximum and minimum values for parameters included in simulation B (all but those related to virus variants) and estimating both the thresholds obtained and the Spearman correlation coefficients with the results of simulation B. We presented the range of thresholds and correlations obtained.

The adapted compartmental model was constructed, the simulations were run, and the data were analyzed using the statistical computing and graphics software R (version 4.0.2; R Foundation for Statistical Computing) in the integrated development environment RStudio (version 2022.07.1+554; Posit). Packages *zoo* (version 1.8-10), *pROC* (version 1.18.0), and *unbalanced* (version 2.1) were used. The 95% CIs were calculated for all punctual estimates.

Textbox 2. Scenarios for different q estimates.

- The *q* estimate derived from the constructed models represents the ratio between the daily expected number of cases arising from quarantined high-risk contacts and the total number of daily cases. Several parameterizations might lead to different numbers of expected cases resulting from applying the modeled transitions to the real number of quarantined individuals.
- Simplifying the presented model to a single-day transition (all quarantined individuals must progress to being infected or return to being susceptible or protected), if on any given day 100 individuals were in compartment Q and the total number of new cases was 20 and…:
  - ...the infection risk (IR) was defined as 10%, the q estimate would be 0.5 (10% × 100/20). In other words, we would expect half the cases to come from quarantined individuals.
  - ...the IR was defined as 20%, the q estimate would be 1.0 (20% × 100/20). In other words, we would expect all cases to come from quarantined individuals.
  - ...the IR was defined as 30%, the q estimate would be 1.5 (30% × 100/20). In other words, we would expect more cases from quarantined individuals than the total number of new cases.
- Therefore, the *q* estimate determined by this model is not a proportion as *q* estimates of >1 are possible within the constructed model (see the Discussion section).

Ethical Considerations

Our study used publicly available aggregated secondary data with no characteristics that allowed for individual identification. Therefore, the research team considered that there were no relevant data protection and privacy issues to report.

Results

Descriptive Analysis of the Parameterizations

The characteristics of the parameterizations used in the model and for each simulation are described in Table 3. The median of the maximum transition periods ($\Delta t\gamma_{max}/\Delta t\kappa\varsigma_{max}$) was 8 days (IQR 6 for the γ transition and IQR 7 for the κ and ς transitions) and 4 days for the average transition periods ($\Delta t\gamma_{mean}/\Delta t\kappa\varsigma_{mean}$; IQR 4). The main difference between the simulations is in the IR, with an average of 28.9% (SD 13.6%) verified in simulation A, which is much higher than that of 0.8% (SD 0.5%) for simulation B.

Table 3. Descriptive analysis of simulation parameters.

	Simulation A	Simulation B
Variants		
$\Delta t\gamma_{max}$[a] (d), median (IQR)	8 (6)	8 (6)
$\Delta t\gamma_{mean}$[b] (d), median (IQR)	4 (4)	4 (4)
$\Delta t\kappa\varsigma_{max}$[c] (d), median (IQR)	8 (7)	9 (7)
$\Delta t\kappa\varsigma_{mean}$[d] (d), median (IQR)	4 (4)	4 (4)
IR[e] (%), mean (SD)	28.9 (13.6)	0.8 (0.5)
Complete vaccination (%), mean (SD)		
Ve_0[f]	33.7 (16.3)	33.7 (16.3)
Ve_{max}[g]	64.8 (21.8)	64.8 (21.8)
Ve_{waned}[h]	49.2 (14.9)	49.2 (14.9)
Booster doses (%), mean (SD)		
Be_0[i]	65.1 (14.0)	65.1 (14.0)
Be_{max}[j]	80.1 (12.3)	80.1 (12.3)
Be_{waned}[k]	49.2 (14.9)	49.2 (14.9)

[a] $\Delta t\gamma_{max}$: maximum time until γ.

[b] $\Delta t\gamma_{mean}$: average time until γ.

[c] $\Delta t\kappa\varsigma_{max}$: maximum time until κ and ς.

[d] $\Delta t\kappa\varsigma_{mean}$: average time until κ and ς.

[e] IR: infection risk.

[f] Ve_0: vaccine efficacy.

[g] Ve_{max}: maximum vaccine efficacy.

[h] Ve_{waned}: waned vaccine efficacy.

[i] Be_0: booster dose efficacy.

[j] Be_{max}: maximum booster dose efficacy.

[k] Be_{waned}: waned booster dose efficacy.

Descriptive Analysis of the Simulations

The daily value of the *q* estimate is plotted in Figures 2-5 for simulations A and B. This estimate exceeded the value of 1 at some moments in simulation A, indicating that, according to the model's parameters, it would be expected that more infection cases resulting from quarantined individuals had occurred than the total number of new cases of COVID-19 reported for those specific days (Textbox 2).

An inverse relationship between the values of the *q* estimate and the total number of new cases of COVID-19 is ascertainable through analysis of Figures 2 and 4. Indeed, simulation A presented a correlation of −0.71 (95% CI −0.74 to −0.67), and simulation B showed a correlation of −0.76 (95% CI −0.79 to −0.73).

Despite the different procedures underlying the 2 simulations (A and B), the maximum *q* estimate in both simulations was reached on day 161. After that, the same *q* estimate never reached >50% of that maximum value.

The analysis of all iteration results according to the presence or absence of different variants and vaccination's protective effect (Figures 3 and 5) shows the considerable overlap between the curves in either simulation A or B.

Figure 2. Results of simulation A in comparison with the daily total number of new COVID-19 cases at scale.

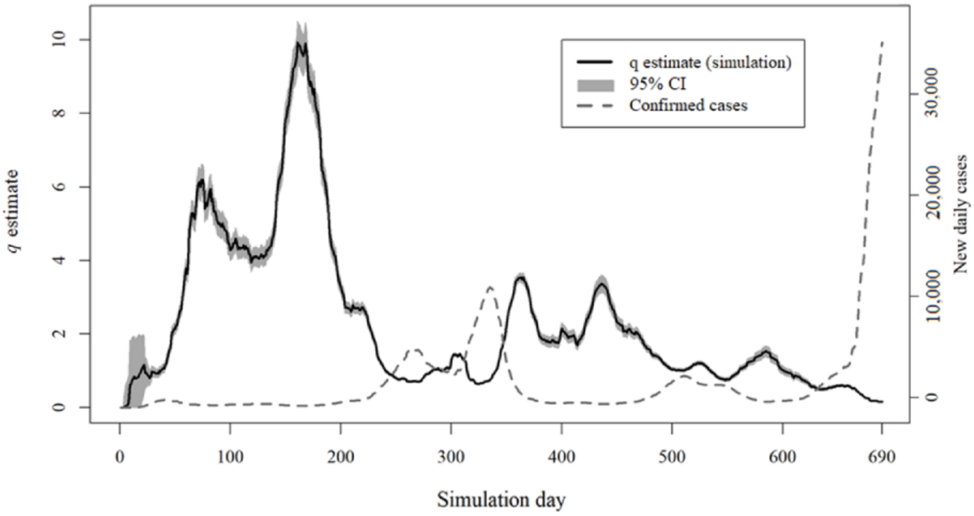


Figure 3. Results of simulation A separated according to inclusion or exclusion of vaccines or SARS-CoV-2 variants.

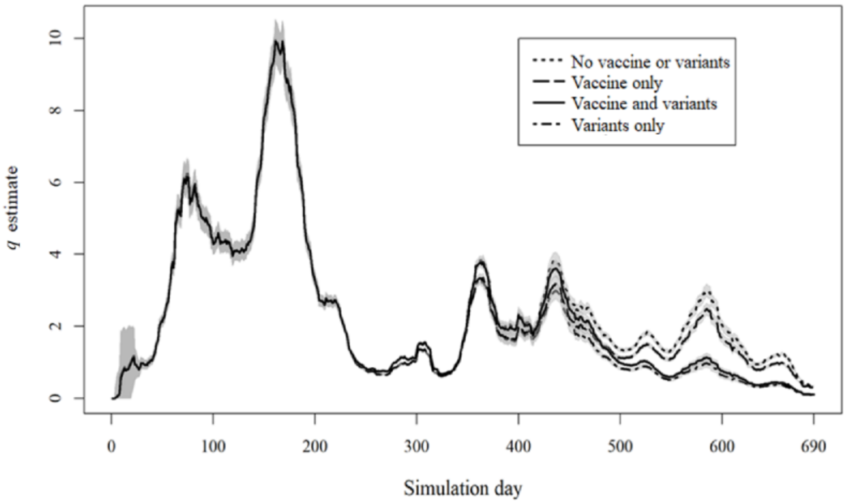


Figure 4. Results of simulation B in comparison with the daily total number of new COVID-19 cases at scale.

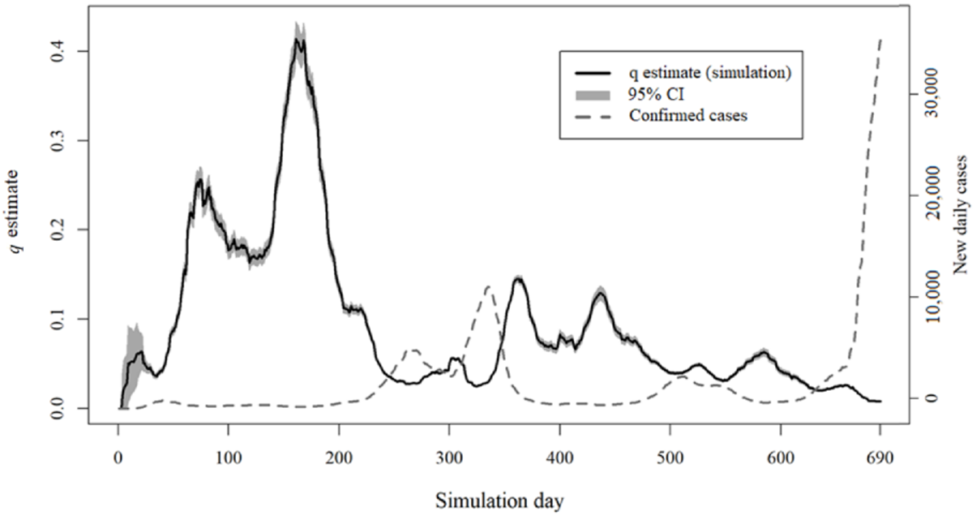
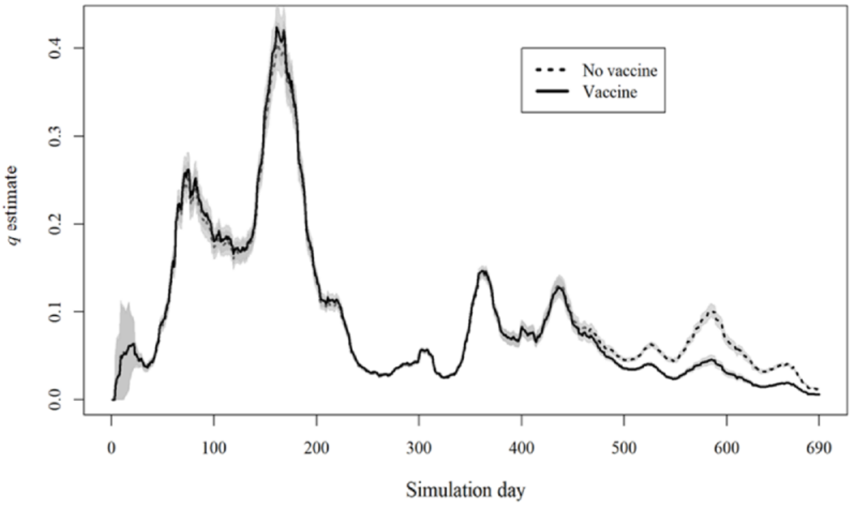


Figure 5. Results of simulation B separated according to the inclusion or exclusion of vaccines.



Defining the Effectiveness Threshold

Figure 6 shows the distribution of the number of daily cases throughout the 690 days included in the compartmental model, including the local maxima; local minima; inflection points; and pandemic, alert, and interpandemic phases. Case peaks occurred on days 40, 264, 335, and 511 of the study period. We also considered that, on day 690 (end of the study period), we were at a new peak because of the number of cases. The inflection points for each of the 5 peaks identified show that alert phases occurred between days 18 and 26 for the first pandemic phase, between days 220 and 250 for the second pandemic phase, between days 310 and 321 for the third pandemic phase, between days 470 and 497 for the fourth pandemic phase, and between days 660 and 676 for the fifth and final pandemic phase. The period included in calculating the effectiveness threshold encompassed 403 days (between days 1 and 26, days 82 and 250, days 302 and 321, days 400 and 497, and days 588 and 676 in accordance with each of the 5 pandemic phases). Table 4 shows the testing periods of the theoretical thresholds derived from both models and the classification of those periods in the alert (test-positive) and interpandemic (test-negative) phases.

For simulation A, we observed that the value of the *q* estimate that would most likely define the theoretical effectiveness threshold of contact tracing would be 1.93, with 85.7% (258/301; 95% CI 82.1%-89.7%) of alert phase days in which the *q* estimate was below the effectiveness threshold (ie, sensitivity), 66.1% (199/301; 95% CI 60.5%-71.1%) of interpandemic phase days in which the *q* estimate was above the effectiveness threshold (ie, specificity), and a proportion of alert days among days with the *q* estimate below the effectiveness threshold (ie, PPV) of 71.7% (258/360; 95% CI 68.4%-75.1%). For simulation B, the hypothetical effectiveness threshold was 0.07, with 87% (262/301; 95% CI 83.4%-90.7%) sensitivity, 66.8% (201/301; 95% CI 61.5%-72.1%) specificity, and a 72.5% (262/362; 95% CI 69.2%-75.9%) PPV.

In Portugal, population lockdowns were imposed during all pandemic phases except the fifth one [36]. In the model days, lockdowns occurred on days 18, 236, 313, and 488. Both simulations would, by default, start below the effectiveness threshold from day 1 in the model until the first pandemic phase. Unsurprisingly, by day 18, both would still stay below the effectiveness threshold. Neither of the simulations' *q* estimates went above threshold levels between the second and third pandemic phases, which may indicate that this hypothetical estimate could be used as an argument against lifting lockdown measures, thus avoiding an interpandemic phase of just 8 days. Both models kept *q* estimates below the effectiveness threshold after the fourth pandemic phase. For the 2 remaining pandemic phases (second and fourth), simulation A broke the threshold within the respective alert phases 7 and 17 days before the actual lockdowns (days 229 and 471), and simulation B broke the threshold 4 and 16 days before the actual lockdowns (days 232 and 472).

Figure 6. Daily COVID-19 case number evolution with identification of wave peaks and periods included in the effectiveness threshold for contact tracing. Vertical dotted lines represent the imposing of population lockdown measures by Portuguese authorities in each pandemic phase during the period of analysis.

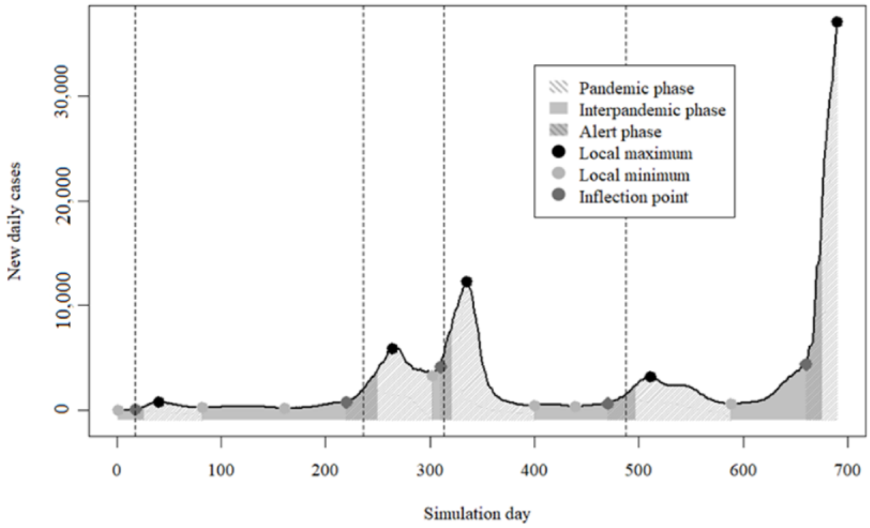


Table 4. Testing periods for theoretical effectiveness threshold.

Days	Phase	Classification
1-17	Interpandemic	Negative
18-26	Alert	Positive
82-219	Interpandemic	Negative
220-250	Alert	Positive
302-309	Interpandemic	Negative
310-321	Alert	Positive
400-469	Interpandemic	Negative
470-497	Alert	Positive
588-659	Interpandemic	Negative
660-676	Alert	Positive

Sensitivity Analysis

Tables 5 and 6 show the results of the sensitivity analysis. On the basis of the multiple linear regression results (Table 5), a 10% increase in IR for each variant was associated with an increase in the estimated threshold value of at least 0.01. The influence was maximal for the variant “other” (β=.52; *P*<.001). The influence on the result of the model for the remaining variant parameters was inconsistent, with some parameters leading to an increase in the hypothetical threshold for a certain variant and the same parameter leading to a decrease in the estimated threshold for other variants. Of these parameters, maximal and mean times until κ and ς for variant “other” were negatively related to the estimated threshold (β=−.10 and *P*<.001, and β=−.08 and *P*=.02, respectively), as were the mean times until κ and ς for the alpha and delta variants (β=−.08 and *P*=.04, and β=−.08 and *P*=.048, respectively). In contrast, an increase in the mean time until κ and ς for the omicron variant was related to an increased threshold (β=.08; *P*=.02). Regarding vaccination, only an increase in efficacy at inoculation for booster doses was negatively related to the estimated threshold (β=−.11; *P*=.02). All other vaccine efficacies did not present a significant relationship with the threshold attained.

Regarding the correlation and threshold analyses (Table 6), we obtained a minimum correlation coefficient of 0.87 and a maximum correlation coefficient of 0.99. Threshold values were highly variable except for those resulting from IR and when vaccine efficacy at inoculation (Ve_0) was set to 60%, where the ratio between the maximum and minimum thresholds was <20. This result corroborates the results obtained in the multiple linear regression regarding the impact of IR on the threshold estimates and, consequently, on *q* estimates.

Table 5. Multiple linear regression model parameter coefficients for threshold determination.

Variant and model parameter	Coefficient (95% CI)	*P* value
Other variant		
$\Delta t\gamma_{max}$[a]	0.02 (−0.04 to 0.08)	.51
$\Delta t\gamma_{mean}$[b]	−0.08 (−0.16 to 0.01)	.07
$\Delta t\kappa\varsigma_{max}$[c]	−0.10 (−0.16 to −0.05)	<.001
$\Delta t\kappa\varsigma_{mean}$[d]	−0.08 (−0.15 to −0.02)	.02
IR[e]	0.52 (0.41 to 0.64)	<.001
Alpha		
$\Delta t\gamma_{max}$	−0.03 (−0.09 to 0.03)	.35
$\Delta t\gamma_{mean}$	−0.04 (−0.12 to 0.03)	.25
$\Delta t\kappa\varsigma_{max}$	0.03 (−0.02 to 0.09)	.26
$\Delta t\kappa\varsigma_{mean}$	−0.08 (−0.15 to 0.00)	.04
IR	0.05 (−0.07 to 0.18)	.41
Beta		
$\Delta t\gamma_{max}$	0 (−0.07 to 0.06)	.99
$\Delta t\gamma_{mean}$	0.01 (−0.07 to 0.09)	.79
$\Delta t\kappa\varsigma_{max}$	0.03 (−0.04 to 0.10)	.41
$\Delta t\kappa\varsigma_{mean}$	0.01 (−0.07 to 0.09)	.78
IR	0.01 (−0.11 to 0.13)	.84
Gamma		
$\Delta t\gamma_{max}$	−0.04 (−0.11 to 0.04)	.31
$\Delta t\gamma_{mean}$	0.07 (−0.01 to 0.16)	.08
$\Delta t\kappa\varsigma_{max}$	0.03 (−0.02 to 0.09)	.23
$\Delta t\kappa\varsigma_{mean}$	0.01 (−0.06 to 0.09)	.75
IR	0.15 (0.03 to 0.28)	.02
Delta		
$\Delta t\gamma_{max}$	0.04 (−0.02 to 0.10)	.19
$\Delta t\gamma_{mean}$	−0.04 (−0.12 to 0.04)	.32
$\Delta t\kappa\varsigma_{max}$	0.02 (−0.04 to 0.09)	.47
$\Delta t\kappa\varsigma_{mean}$	−0.08 (−0.16 to 0.00)	.048
IR	0.11 (−0.04 to 0.26)	.16
Omicron		
$\Delta t\gamma_{max}$	0.02 (−0.05 to 0.08)	.59
$\Delta t\gamma_{mean}$	0.00 (−0.07 to 0.06)	.89
$\Delta t\kappa\varsigma_{max}$	−0.04 (−0.09 to 0.01)	.16
$\Delta t\kappa\varsigma_{mean}$	0.08 (0.01 to 0.15)	.02
IR	0.08 (−0.03 to 0.20)	.16
Ve_0[f]	0.02 (−0.04 to 0.09)	.48
Ve_{max}[g]	0.01 (−0.05 to 0.06)	.85

Variant and model parameter	Coefficient (95% CI)	*P* value
Ve_{waned}[h]	N/A[i]	N/A
Be_0[j]	−0.11 (−0.20 to −0.02)	.02
Be_{max}[k]	0.01 (−0.09 to 0.11)	.84
Be_{waned}[l]	N/A	N/A

[a] $\Delta t\gamma_{max}$: maximum time until γ.

[b] $\Delta t\gamma_{mean}$: average time until γ.

[c] $\Delta t\kappa\varsigma_{max}$: maximum time until κ and ς.

[d] $\Delta t\kappa\varsigma_{mean}$: average time until κ and ς.

[e] IR: infection risk.

[f] Ve_0: vaccine efficacy.

[g] Ve_{max}: maximum vaccine efficacy.

[h] Ve_{waned}: waned vaccine efficacy.

[i] N/A: not applicable; coefficients for waned efficacy were not calculated as this parameter was input in the model as the mean of the inoculation and maximum efficacies (Table 1).

[j] Be_0: booster dose efficacy.

[k] Be_{max}: maximum booster dose efficacy.

[l] Be_{waned}: waned booster dose efficacy.

Table 6. Correlation and threshold estimates for fixed parameters.

Variable and value	Correlation[a] (range)	Threshold (range)
$\Delta t\gamma_{max}$[b] (d)		
2	0.87-0.88	0.003-13.372
14	0.88-0.99	0.002-11.057
$\Delta t\gamma_{mean}$[c] (d)		
2	0.92-0.98	0.037-2.965
14	0.89-0.95	0.002-2.696
$\Delta t\kappa\varsigma_{max}$[d] (d)		
2	0.91-0.99	0.016-13.372
14	0.88-0.99	0.002-11.278
$\Delta t\kappa\varsigma_{mean}$[e] (d)		
2	0.95-0.99	0.030-7.385
14	0.88-0.99	0.002-3.137
IR[f]		
0.001	0.87-0.99	0.002-0.031
0.500	0.88-0.99	1.248-13.372
Ve_0[g]		
0.10	0.87-0.99	0.002-13.372
0.60	0.90-0.96	0.046-0.448
Ve_{max}[h]		
0.10	0.89-0.99	0.002-13.372
0.95	0.87-0.96	0.002-10.744
Be_0[i]		
0.10	0.87-0.99	0.002-13.372
0.80	0.89-0.99	0.002-13.144
Be_{max}[j]		
0.10	0.89-0.99	0.002-13.372
0.95	0.89-0.99	0.030-8.352

[a] Compared with simulation B results.

[b] $\Delta t\gamma_{max}$: maximum time until γ.

[c] $\Delta t\gamma_{mean}$: average time until γ.

[d] $\Delta t\kappa\varsigma_{max}$: maximum time until κ and ς.

[e] $\Delta t\kappa\varsigma_{mean}$: average time until κ and ς.

[f] IR: infection risk.

[g] Ve_0: vaccine efficacy.

[h] Ve_{max}: maximum vaccine efficacy.

[i] Be_0: booster dose efficacy.

[j] Be_{max}: maximum booster dose efficacy.

Discussion

Principal Findings

In this study, we developed a compartmental model to describe the relationship among the official data on COVID-19 for Portugal, namely regarding the DGS definition of quarantined individuals. The main aim of this study was to estimate the daily ratio of cases that occurred in individuals in imposed quarantine through contact tracing (*q* estimate). Despite being a compartmental model, only transitions to and from compartment Q (quarantined individuals) were modeled. Daily data regarding all the other transitions and daily data on all compartments except compartments S (susceptible) and E (exposed individuals not traced) were already available and, therefore, not processed.

To the best of our knowledge, this study is the first attempt to use COVID-19 contact tracing data to define the infection dynamics of SARS-CoV-2. The main reason for this might be that the DGS was, as far as the authors are aware, the only national health authority with public data regarding individuals under quarantine systematically imposed by health authorities.

We obtained *q* estimates for 2 different simulations, both grounded on published evidence and some degree of parameter estimation. Simulation A was intended to reflect the presence of multiple strains with different infectious characteristics, whereas simulation B was intended to calibrate the IR of SARS-CoV-2 in quarantined individuals. The *q* estimates obtained by the 2 simulations were highly negatively correlated with the number of confirmed daily cases of COVID-19, which shows that a diminishing ratio of cases from quarantined individuals indirectly indicates the waning effectiveness of contact tracing. Assuming that we have proven this concept, it may indicate epidemiological contexts in which contact tracing as an epidemic combat measure would be insufficient, thus prompting the need to consider implementing complementary pandemic combat measures that result in further social restrictions, namely, general lockdowns.

To demonstrate the impact of applying this theoretical *q* estimate, we tried to establish, for each simulation, a hypothetical effectiveness threshold for contact tracing and relate it to the key moment for implementing general lockdowns, either by defining the phases of the pandemic and estimating a PPV (ie, the quotient of the number of alert phase days in which the *q* estimate was below the effectiveness threshold and the number of days in which the *q* estimate was below the effectiveness threshold) of each simulation or by comparing with actual lockdown measures in Portugal. We found a PPV of >70% for both simulations, and an anticipation of at least 4 days in the second and fourth lockdowns was the result of the theoretical simulations used for decision-making.

In addition, a sensitivity analysis of the parameters used to model compartment Q transitions showed that only the IR for any strain and the vaccine efficacy at booster dose inoculation did significantly affect the *q* estimates. Despite having a residual influence, the parameters input do not allow for a direct interpretation of the value of the *q* estimate, especially in simulation A as this simulation at times obtained values of >1. The scenarios where this situation could occur in our model were explained in Textbox 2. The main reasons for these estimates are more likely related to an overestimated IR in compartment Q and may include misclassification of low-risk exposures as high-risk exposures (eg, in schools when the first guidelines initially considered all same-class students as high-risk contacts regardless of mask or social distancing measures implemented [37]), delayed diagnosis, or different timing for testing in quarantined individuals (such as changes in guidelines for testing in the beginning, middle, or end of imposed quarantine [5,21,38]) and different isolation timing (some individuals might only be traced near the end of their period of isolation, which in turn leads to a reduced time in compartment Q not controlled by our model design).

Therefore, our study should be regarded as a demonstration of the application of an effectiveness threshold for contact tracing as a measure of pandemic control. Consequently, the *q* estimates and effectiveness thresholds derived from both simulations are hypothetical values but likely to be highly correlated with the true proportion of cases from quarantined individuals, as shown by our sensitivity analysis.

We deem it necessary to draw attention to the importance of open data and data sharing as a way to catalyze research and accelerate innovation and development that shortens the time between the detection of potentially epidemic pathogens and the development of appropriate containment and mitigation measures [39,40]. The advantages of data sharing will only be fully attained if data are reported on an everyday basis of definitions and procedures. This factor was not observed during the COVID-19 pandemic. This research team expects that this lesson, too, can be derived from this pandemic and that, consequently, definitions and procedures will be adapted in conformity [41].

Limitations

The main limitation of our study arises from implementing a compartmental model and simulation methods to estimate transitions between compartments, potentially leading to results that only partially reflect reality as it occurred at each moment of the period analyzed. Furthermore, it is not possible to exclude the influence that the implementation of confinement measures or other measures implemented during each pandemic wave may have had on the proportion of new cases from exposed individuals. However, the impact of considering different parameterizations of variants or vaccination efficacy, demonstrated by the visual overlap between the results of the different simulations as well as the conducted sensitivity analysis, strengthens the potential robustness of the results and the model.

In addition, the parameters chosen include other limitations as they are based on data collected in studies that were run under controlled conditions [42]. Most of those studies also focused on the disease's transmission dynamics [22], analyzing mainly those persons who developed the disease and the time that elapsed until infection. Therefore, we lacked estimates of the duration of epidemiological surveillance for exposed individuals who did not develop the infection. To mitigate this limitation, we assumed that this time would be similar to the time required for individuals to develop the disease.

Few studies have compared the risk of infection per virus variant throughout the quarantine period. Therefore, although we acknowledge that some variants might have been more transmissible than others, it did not allow us to determine whether this variable translated only to a higher IR or a lower time until the transition between exposed and infected (or any other hypotheses, for that matter). To mitigate the impact of this limitation, different values for parameters (which assumed that all variants could have variable IR and time to infection) had the potential to depict variant differences concerning the degree of contagion and infection. The residual impact of different variants, except for their respective IRs, on the simulation results may attenuate this limitation.

It is not possible to exclude the hypothesis that the effect of implementation of confinement measures (or other measures implemented during each pandemic phase) on the ratio of new cases from quarantined individuals may have led the model to present *q* estimates that were different from those expected without such measures. Thus, the definition of a value for the ratio of cases of quarantined individuals that could work as a threshold for lifting those more restrictive measures or restarting contact tracing as the primary pandemic combat strategy is beyond the scope of this study. This “lift measures threshold” could define the moment where the influence of confinement measures stopped manifesting. In our model, we assume that the contact tracing strategy is maintained throughout the simulation period and only use data from the interpandemic and alert phases to define the hypothetical effectiveness threshold.

Finally, given the format in which the data were collected, the analysis could be conducted only at the national level. This lack of data granularity demanded an implicit assumption that any change in the number of cases or individuals under quarantine imposed by health authorities affects national, regional, and local levels simultaneously and proportionally. Greater data detail, both geographical and sectorial, would potentially allow for determining effectiveness thresholds—and the consequent implementation of pandemic containment measures—at those differentiated levels, with a more apparent effort-benefit ratio that could be better understood and better accepted by the population. This would most likely result in higher population adherence to containment measures and finer-tuned control of both the sanitary and socioeconomic impacts of the pandemic.

Future Work

As we have stated, this work is a starting point for using data on high-risk contacts of COVID-19 cases to define transmission dynamics of SARS-CoV-2, which could lead to further studies addressing the aforementioned limitations and discussions of alternative hypotheses and perspectives left unexplored in this study. An immediate suggestion could be the case of a scenario in which the variation in IR for compartment Q occurs within each iteration. Such dynamic IR should reflect the implementation of pandemic combat measures that have been proven to change the IR (namely, large-scale testing [11], scaling up of contact tracing [1], and population lockdowns [13,14]). Caution should be taken when applying this dynamic IR to these models to tackle *q* estimates of >1. Such an approach, without accounting for other measures, would lead to a maximization of IR in quarantined individuals, which would mean that, in pandemic phases in which cases are expected to occur mostly in non–contact-traced individuals, we would contrarily have a maximization of the *q* estimate.

Another scenario to tackle in future work—perhaps also resorting to simulation methods—would be to investigate what would happen should contact tracing come to a halt and what would be the impact of fully transferring resources allocated to this public health task to other activities regarding epidemiological dynamics and the pandemic’s impact on a country’s sanitary and economic dimensions.

As mentioned previously, the theoretical definition of contact tracing effectiveness thresholds that would allow for the relief of restrictive measures is beyond the scope of this work. Nonetheless, it might be equally valuable to define thresholds for restrictive measure relief as it could be helpful for communication by public officials during pandemic scenarios as well as to balance the scales between the implicit health-economy dichotomy that so often arose in narratives with global reach. There is also a relevant economic component to this work as it is intertwined with the efficiency of these pandemic control measures and their direct and indirect impacts on countries’ health and economy. This work might be very helpful in drawing lessons that better enable us to adapt our society to events such as pandemics.

Finally, and from a health economics standpoint, it would be interesting to understand how the population perceives contact tracing as an instrument used by health authorities to contain pandemics and what the perception would be, for example, in terms of adherence to pandemic control measures should contact tracing cease.

Conclusions

Our work provides important information for policy and decision makers, namely in terms of epidemiology and pandemic/crisis management, as well as for public health professionals. It also constitutes a relevant source of information and an objective acknowledgment of the importance of contact tracing, which was widely used worldwide in all stages of the COVID-19 pandemic.

Concretely, this study, although using secondary data, allowed us to (1) provide a proof concept for using estimates of an effectiveness threshold for contact tracing as a primary measure of pandemic containment and (2) consider the potential use of this effectiveness threshold as a decision variable for imposing more restrictive measures, namely, lockdowns, along with indicators that were used to assess the pandemic situation (such as the transmissibility indexes and case incidence). Our results are consistent with this last possibility and, notwithstanding the presented limitations, reveal a path that, we hope, deserves further exploration.

Acknowledgments

This paper was based on the scientific efforts of the research team developed upon the execution of the research protocol that won the first edition of the Scholarship Award Amélia Leitão, promoted by the Portuguese National Association of Public Health Doctors and Pfizer Laboratories. This work would have been impossible without the generous contribution of Data Science for Social Good in compiling and curating the data reported in PDF format by the Directorate-General of Health. This is our way to give back to the community as part of the benefits we received through them.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request. The code we developed and on which these simulations are based is available on a GitHub repository [43].

Conflicts of Interest

None declared.

References

1. Contact tracing for COVID-19: current evidence, options for scale-up and an assessment of resources needed. European Centre for Disease Prevention and Control. 2020 May 5. URL: https://www.ecdc.europa.eu/en/publications-data/contact-tracing-covid-19-evidence-scale-up-assessment-resources [accessed 2022-10-25]
2. Fetzer T, Graeber T. Measuring the scientific effectiveness of contact tracing: evidence from a natural experiment. Proc Natl Acad Sci U S A 2021 Aug 17;118(33):e2100814118 [FREE Full text] [doi: 10.1073/pnas.2100814118] [Medline: 34385318]
3. Thomas Craig KJ, Rizvi R, Willis VC, Kassler WJ, Jackson GP. Effectiveness of contact tracing for viral disease mitigation and suppression: evidence-based review. JMIR Public Health Surveill 2021 Oct 06;7(10):e32468 [FREE Full text] [doi: 10.2196/32468] [Medline: 34612841]
4. Zweig SA, Zapf AJ, Xu H, Li Q, Agarwal S, Labrique AB, et al. Impact of public health and social measures on the COVID-19 pandemic in the United States and other countries: descriptive analysis. JMIR Public Health Surveill 2021 Jun 02;7(6):e27917 [FREE Full text] [doi: 10.2196/27917] [Medline: 33975277]
5. Norma 015/2020 de 24/07/2020, COVID-19 - Rastreio de Contactos. Direção Geral da Saúde. 2020 Jul 24. URL: https://www.dgs.pt/normas-orientacoes-e-informacoes/normas-e-circulares-normativas/norma-n-0152020-de-24072020-pdf.aspx [accessed 2022-01-05]
6. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). J Epidemiol Community Health 2020 Oct;74(10):861-866 [FREE Full text] [doi: 10.1136/jech-2020-214051] [Medline: 32576605]
7. Kucharski AJ, Klepac P, Conlan AJ, Kissler SM, Tang ML, Fry H, CMMID COVID-19 working group. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. Lancet Infect Dis 2020 Oct;20(10):1151-1160 [FREE Full text] [doi: 10.1016/S1473-3099(20)30457-6] [Medline: 32559451]
8. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health 2020 Apr;8(4):e488-e496 [FREE Full text] [doi: 10.1016/S2214-109X(20)30074-7] [Medline: 32119825]
9. Clark E, Chiao EY, Amirian ES. Why contact tracing efforts have failed to curb coronavirus disease 2019 (COVID-19) transmission in much of the United States. Clin Infect Dis 2021 May 04;72(9):e415-e419 [FREE Full text] [doi: 10.1093/cid/ciaa1155] [Medline: 32761123]
10. Plano nacional de preparação e resposta à doença por novo coronavírus (COVID-19). Direção Geral da Saúde. 2020. URL: https://www.dgs.pt/documentos-e-publicacoes/plano-nacional-de-preparacao-e-resposta-para-a-doenca-por-novo-coronavirus-covid-19-pdf.aspx [accessed 2022-10-25]
11. Yuan HY, Blakemore C. The impact of contact tracing and testing on controlling COVID-19 outbreak without lockdown in Hong Kong: an observational study. Lancet Reg Health West Pac 2022 Mar;20:100374 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100374] [Medline: 35072128]
12. Morgan O. How decision makers can use quantitative approaches to guide outbreak responses. Philos Trans R Soc Lond B Biol Sci 2019 Jul 08;374(1776):20180365 [FREE Full text] [doi: 10.1098/rstb.2018.0365] [Medline: 31104605]
13. Vinceti M, Filippini T, Rothman KJ, Ferrari F, Goffi A, Maffeis G, et al. Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. EClinicalMedicine 2020 Aug;25:100457 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100457] [Medline: 32838234]
14. Oraby T, Tyshenko MG, Maldonado JC, Vatcheva K, Elsaadany S, Alali WQ, et al. Modeling the effect of lockdown timing as a COVID-19 control measure in countries with differing social contacts. Sci Rep 2021 Feb 08;11(1):3354 [FREE Full text] [doi: 10.1038/s41598-021-82873-2] [Medline: 33558571]

15. Peralta-Santos A, Nunes B, Mateiro Gomes B, Nunes C, Antunes C, Gouveia J, et al. Linhas Vermelhas: Epidemia de infeção por SARS-CoV-2/COVID-19. Direção Geral de Saúde, Instituto Nacional de Saúde Doutor Ricardo Jorge, Escola Nacional de Saúde Pública, Faculdade de Ciências da Universidade de Lisboa, Faculdade de Ciências da Universidade do Porto. 2021 Mar. URL: https://static.poder360.com.br/2021/03/Linhas-Vermelhas-pandemia-dgs-portugal.pdf [accessed 2022-10-20]
16. Relatório de Situação. Direção Geral da Saúde. 2022. URL: https://covid19.min-saude.pt/relatorio-de-situacao/ [accessed 2022-10-20]
17. Dados relativos à pandemia COVID-19 em Portugal. GitHub. Lisbon, Portugal: Associação Portuguesa de Ciência de Dados para o Bem Social (DSSG); 2023. URL: https://github.com/dssg-pt/covid19pt-data [accessed 2022-10-25]
18. Data on SARS-CoV-2 variants in the EU/EEA. European Centre for Disease Prevention and Control. 2023. URL: https://www.ecdc.europa.eu/en/publications-data/data-virus-variants-covid-19-eueea [accessed 2022-09-01]
19. Hall V, Foulkes S, Insalata F, Kirwan P, Saei A, Atti A, SIREN Study Group. Protection against SARS-CoV-2 after Covid-19 vaccination and previous infection. N Engl J Med 2022 Mar 31;386(13):1207-1220 [FREE Full text] [doi: 10.1056/NEJMoa2118691] [Medline: 35172051]
20. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
21. Norma 004/2020 de 23/03/2020 - Abordagem das Pessoas com Suspeita ou Confirmação de COVID-19. Direção Geral da Saúde. 2020 Mar 23. URL: https://www.dgs.pt/normas-orientacoes-e-informacoes/normas-e-circulares-normativas/norma-n-0042020-de-230320201.aspx [accessed 2022-01-05]
22. Wu Y, Kang L, Guo Z, Liu J, Liu M, Liang W. Incubation period of COVID-19 caused by unique SARS-CoV-2 strains: a systematic review and meta-analysis. JAMA Netw Open 2022 Aug 01;5(8):e2228008 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.28008] [Medline: 35994285]
23. Karumanagoundar K, Raju M, Ponnaiah M, Kaur P, Viswanathan V, Rubeshkumar P, et al. Secondary attack rate of COVID-19 among contacts and risk factors, Tamil Nadu, March-May 2020: a retrospective cohort study. BMJ Open 2021 Nov 05;11(11):e051491 [FREE Full text] [doi: 10.1136/bmjopen-2021-051491] [Medline: 34740930]
24. Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. J Clin Med 2020 Feb 07;9(2):462 [FREE Full text] [doi: 10.3390/jcm9020462] [Medline: 32046137]
25. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc Lond A Math Phys Sci 1927 Aug 01;115(772):700-721 [FREE Full text] [doi: 10.1098/rspa.1927.0118]
26. Leontitsis A, Senok A, Alsheikh-Ali A, Al Nasser Y, Loney T, Alshamsi A. SEAHIR: a specialized compartmental model for COVID-19. Int J Environ Res Public Health 2021 Mar 06;18(5):2667 [FREE Full text] [doi: 10.3390/ijerph18052667] [Medline: 33800896]
27. Kong L, Duan M, Shi J, Hong J, Chang Z, Zhang Z. Compartmental structures used in modeling COVID-19: a scoping review. Infect Dis Poverty 2022 Jun 21;11(1):72 [FREE Full text] [doi: 10.1186/s40249-022-01001-y] [Medline: 35729655]
28. Stanojevic S, Ponjavic M, Stanojevic S, Stevanovic A, Radojicic S. Simulation and prediction of spread of COVID-19 in The Republic of Serbia by SEAIHRDS model of disease transmission. Microb Risk Anal 2021 Aug;18:100161 [FREE Full text] [doi: 10.1016/j.mran.2021.100161] [Medline: 33723516]
29. Ding Y, Gao L. An evaluation of COVID-19 in Italy: a data-driven modeling analysis. Infect Dis Model 2020 Jul 09;5:495-501 [FREE Full text] [doi: 10.1016/j.idm.2020.06.007] [Medline: 32766461]
30. Ryu S, Ali ST, Lim JS, Chun BC. Estimation of the excess COVID-19 cases in Seoul, South Korea by the students arriving from China. Int J Environ Res Public Health 2020 Apr 29;17(9):3113 [FREE Full text] [doi: 10.3390/ijerph17093113] [Medline: 32365703]
31. Harris-Birtill R, Harris-Birtill D. Understanding computation time: a critical discussion of time as a computational performance metric. In: Misztal A, Harris PA, Parker JA, editors. Time in Variance. Leiden, The Netherlands: Brill; Sep 20, 2021:220-248.
32. Hill TP. Conflations of probability distributions. Trans Amer Math Soc 2011 Nov 06;363(6):3351-3372 [FREE Full text] [doi: 10.1090/s0002-9947-2011-05340-7]
33. Vázquez-Seisdedos CR, Neto JE, Marañón Reyes EJ, Klautau A, Limão de Oliveira RC. New approach for T-wave end detection on electrocardiogram: performance in noisy conditions. Biomed Eng Online 2011 Sep 09;10:77 [FREE Full text] [doi: 10.1186/1475-925X-10-77] [Medline: 21906317]
34. Gianquintieri L, Brovelli MA, Pagliosa A, Dassi G, Brambilla PM, Bonora R, et al. Mapping spatiotemporal diffusion of COVID-19 in Lombardy (Italy) on the base of emergency medical services activities. Int J Geo Inf 2020 Oct 27;9(11):639. [doi: 10.3390/ijgi9110639]
35. Pandemic influenza risk management: a WHO guide to inform and harmonize national and international pandemic preparedness and response. World Health Organization. 2017. URL: https://apps.who.int/iris/handle/10665/259893 [accessed 2022-10-25]

36. Mathieu E, Ritchie H, Rodés-Guirao L, Appel C, Giattino C, Hasell J, et al. Coronavirus Pandemic (COVID-19). Our World in Data. 2020. URL: https://ourworldindata.org/coronavirus [accessed 2022-10-25]
37. Referencial escolas: controlo da transmissão de COVID-19 em contexto escolar. Direção Geral de Saúde. 2020. URL: https://www.portugal.gov.pt/download-ficheiros/ficheiro.aspx?v=%3D%3DBQAAAB%2BLCAAAAAAABAAzNDAxNgQAzLafLAUAAAA%3D [accessed 2022-10-25]
38. Nogueira-Leite D, Alves JM, Marques-Cruz M, Cruz-Correia R. A cautionary tale on using COVID-19 data for machine learning. In: Proceedings of the 19th International Conference on Artificial Intelligence in Medicine. 2021 Jun Presented at: AIME '21; June 15-18, 2021; Virtual p. 265-275. [doi: 10.1007/978-3-030-77211-6_30]
39. Yozwiak NL, Schaffner SF, Sabeti PC. Data sharing: make outbreak research open access. Nature 2015 Feb 26;518(7540):477-479. [doi: 10.1038/518477a] [Medline: 25719649]
40. Harrison PW, Lopez R, Rahman N, Allen SG, Aslam R, Buso N, et al. The COVID-19 data portal: accelerating SARS-CoV-2 and COVID-19 research through rapid open access data sharing. Nucleic Acids Res 2021 Jul 02;49(W1):W619-W623 [FREE Full text] [doi: 10.1093/nar/gkab417] [Medline: 34048576]
41. Garcia M, Lipskiy N, Tyson J, Watkins R, Esser ES, Kinley T. Centers for Disease Control and Prevention 2019 novel coronavirus disease (COVID-19) information management: addressing national health-care and public health needs for standardized data definitions and codified vocabulary for data exchange. J Am Med Inform Assoc 2020 Jul 01;27(9):1476-1487 [FREE Full text] [doi: 10.1093/jamia/ocaa141] [Medline: 32940705]
42. Feikin DR, Higdon MM, Abu-Raddad LJ, Andrews N, Araos R, Goldberg Y, et al. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. Lancet 2022 Mar 05;399(10328):924-944 [FREE Full text] [doi: 10.1016/S0140-6736(22)00152-0] [Medline: 35202601]
43. Marques da Cruz M, Alves JM, Nogueira Leite D, Fernandes F, Almeida M. Simulation of COVID-19 cases arising from high-risk contacts. GitHub. 2022 Oct 26. URL: https://github.com/marquescruz/covid19-contact-tracing-cases-simulation [accessed 2022-10-26]

Abbreviations

DGS: Directorate-General of Health
DSSG: Data Science for Social Good
IR: infection risk
PPV: positive predictive value
SEIR: susceptible-exposed-infected-recovered

Edited by A Mavragani, T Sanchez; submitted 26.10.22; peer-reviewed by A Benis, Z Zrubka; comments to author 03.01.23; revised version received 24.01.23; accepted 01.03.23; published 06.04.23.

Please cite as:
Marques-Cruz M, Nogueira-Leite D, Alves JM, Fernandes F, Fernandes JM, Almeida MÂ, Cunha Correia P, Perestrelo P, Cruz-Correia R, Pita Barros P
COVID-19 Contact Tracing as an Indicator for Evaluating a Pandemic Situation: Simulation Study
JMIR Public Health Surveill 2023;9:e43836
URL: https://publichealth.jmir.org/2023/1/e43836
doi:10.2196/43836
PMID:36877958

Original Paper

The Role of Community Cohesion in Older Adults During the COVID-19 Epidemic: Cross-sectional Study

Ying Li[1,2], MD, PhD; XiWen Ding[1], MPH; Ayizuhere Aierken[1], MPH; YiYang Pan[1], MA; Yuan Chen[1,3], MA; DongBin Hu[4], MD

[1]Department of Social Medicine, School of Public Health, Zhejiang University, HangZhou, China
[2]School of Medicine, Zhejiang University, HangZhou, China
[3]Department of Quality Surveillance of Zhejiang Provincial People's Hospital, HangZhou, China
[4]The Second Affiliated Hospital, Harbin Medical University, Harbin, China

Corresponding Author:
Ying Li, MD, PhD
Department of Social Medicine, School of Public Health
Zhejiang University
866 Yu-hang-tang Road
HangZhou, 310058
China
Phone: 86 0571(8820)8590
Fax: 86 0571(8795)2233
Email: liying2012@zju.edu.cn

Abstract

Background: The community environment plays a vital role in the health of older adults. During the COVID-19 epidemic, older adults, who were considered the most impacted and most vulnerable social group, were confined to their homes during the implementation of management and control measures for the epidemic. In such situations, older adults may have to contend with a lack of resources and experience anxiety. Therefore, identifying the environmental factors that are beneficial for their physical and mental health is critical.

Objective: This study aimed to assess the association between community cohesion and the physical and mental health of older adults and to identify the related community services and environmental factors that may promote community cohesion.

Methods: This community-based cross-sectional study was designed during the COVID-19 epidemic. A multistage sampling method was applied to this study. A total of 2036 participants aged ≥60 years were sampled from 27 locations in China. Data were collected through face-to-face interviews. The neighborhood cohesion instrument consisting of scales on 3 dimensions was used to assess community cohesion. Self-efficacy and life satisfaction, cognitive function and depression, and community services and environmental factors were also measured using standard instruments. Statistical analyses were restricted to 99.07% (2017/2036) of the participants. Separate logistic regression analysis was conducted to assess the association among community cohesion and physical and mental health factors, related community services, and environmental factors among older adults.

Results: The results showed that high levels of community cohesion were associated with good self-perceived health status and life satisfaction (odds ratio [OR] 1.27, 95% CI 1.01-1.59 and OR 1.20, 95% CI 1.15-1.27, respectively) and high levels of self-efficacy and psychological resilience (OR 1.09, 95% CI 1.05-1.13 and OR 1.05, 95% CI 1.03-1.06, respectively). The length of stay in the community and the level of physical activity were positively associated with community cohesion scores, whereas the education level was negatively associated with community cohesion scores (*P*=.009). Community cohesion was also associated with low levels of depression and high levels of cognitive function. Community cohesion was significantly associated with community services and environmental factors on 4 dimensions. High levels of community cohesion were associated with transportation services and rehabilitation equipment rental services as well as high levels of satisfaction with community physicians' technical expertise and community waste disposal (OR 3.14, 95% CI 1.87-5.28; OR 3.62, 95% CI 2.38-5.52; OR 1.37, 95% CI 1.08-1.73; and OR 1.23, 95% CI 1.01-1.50, respectively).

Conclusions: Community cohesion was found to be associated with the physical and mental health of older adults. Our research suggests that enhancing community services and environmental resources may be an effective strategy to increase community cohesion during major infectious disease epidemics.

(JMIR Public Health Surveill 2023;9:e45110) doi:10.2196/45110

KEYWORDS

community cohesion; physical and mental health; community services; environmental resources; COVID-19 epidemic; older adults

Introduction

Background

Older adults become more vulnerable to mental diseases and environmental challenges as their age increases [1]. The community environment plays a vital role in their health, especially their mental health, as they age [2]. They usually wish to remain living in the community as long as possible, but their health status may deteriorate during the last stage of their lives to the point where they can no longer live alone. Therefore, the World Health Organization suggests the provision of a familiar environment and an emotional bond to support *aging in place* to enable older adults to live in their community independently and comfortably by experiencing a sense of attachment, connectedness, security, and identity [3-5].

During the COVID-19 pandemic, older adults were considered the most impacted and vulnerable social group because they had a high risk of being infected by SARS-CoV-2 and dying [6]. Several epidemic prevention and control measures had changed to a large extent the way of living they were accustomed to and seriously interfered with their access to public resources. During the initial stage of the pandemic, they were alone and unattended at home; thus, they could not purchase masks and obtain any information on disease prevalence; they also felt anxious, helpless, and scared [7-9]. Some studies showed that the deterioration in physical functioning and the reduction in social contact may have caused many older adults to experience poor mental health [10,11]. During COVID-19–related lockdowns imposed by the authorities, older adults were cut off from the outside world, which also hindered them from receiving social support [12]. In such circumstances, older adults can experience serious physical and mental health consequences, and, as a result, they become more dependent on community services and environment resources.

The Importance of Community Cohesion

Community cohesion, which is an important element of the community, includes 3 essential dimensions. It involves social relations, identification with the geographical unit, and orientation toward the common good [13]. Individual-level community cohesion is assessed in a geographically bounded neighborhood; this is known as the community or neighborhood cohesion assessment. In recent years, interest in the study of community cohesion has escalated, probably because of the sharp increase in older populations. A study suggested that community cohesion is an important factor that affects physical and psychological health outcomes among older adults [14]. Many studies have shown that community cohesion may positively influence both friendship and well-being, which act as buffers against serious psychological distress among older adults [15-17]. A lack of community cohesion is associated with depression, loneliness, suicidal ideation, and poor mental health [18-20]. Community cohesion is also an important protective factor for preventing posttraumatic stress disorder symptoms when older adults experience negative events and disasters [21]. Clinical and epidemiological studies have found a lack of community cohesion to be associated with disability, hypertension, myocardial infarction, stroke, and mortality [22-25]. Empirical studies have confirmed that low levels of community cohesion can accelerate epigenetic aging, and high levels of community cohesion can reduce the risks of diseases associated with genetic risk factors [26,27].

The COVID-19 pandemic has had an enormous and long-lasting negative impact on the physical and mental health of older adults worldwide. However, the unprecedented COVID-19 outbreak and special prevention and control measures have resulted in a limited number of studies being conducted on the association of community cohesion with the physical and mental health of older adults. We found several similar studies on the relationship between community cohesion and mental health, but these studies have a small sample size, resulting in low statistical power [28,29]. In addition, the difficulty in data collection during the COVID-19 pandemic resulted in only a single study (with a lack of theoretical models) being conducted on the association of community cohesion with the mental health of older adults. A multiscale evaluation study on the complex impact of the COVID-19 pandemic on the physical and mental health of older adults has not yet been conducted. To our knowledge, no research has assessed community cohesion factors, such as community services and community resources, that affected the physical and mental health of older adults and the possible community public strategies that could have been used to improve their physical and mental health during the COVID-19 pandemic while at the emergency prevention and control stage. Through this research, we expected to collect considerable amounts of community resources information as well as physical and mental health measurement data so that we could fully explore the effects of community cohesion on the physical and mental health of older adults at the individual, community, and social levels during the COVID-19 pandemic based on the social-ecological model and investigate the complex relationship among community cohesion, community services, environmental resources, and the physical and mental health of older adults and the potential interaction mechanism.

We hypothesized that a good community environment and satisfactory health service facilities can improve community cohesion and enhance community residents' perception of social support and self-efficacy in the effective use of community services and environmental resources. This study aimed to explore the effects of community cohesion on the physical and mental health of older adults living in a community during the COVID-19 epidemic, identify the association between community cohesion and perceived social support and self-efficacy, and clarify and define the community services and environmental factors that may promote community cohesion.

Methods

Study Design

A cross-sectional study was conducted among community-dwelling older adults to assess the related factors of community cohesion and the association between community cohesion and the physical and mental health of older adults during the COVID-19 pandemic. This study was supported by a big data–driven community mental health management model as well as an accessibility evaluation of health-related resource projects. Two community-based surveys that aimed to assess health status and accessibility of community health services among older adults living in the community were conducted to discover the main psychological and mental health challenges and explore ways to promote community health based on the theory of social ecology.

Participants

A total of 2036 community residents were selected through a multistage sampling design according to comprehensive geographic location attributes and fully considering the regional aging degree. Sampling was performed at 27 locations in 4 provinces in China from July 15, 2020, to August 31, 2022. All participants were aged ≥60 years. Participants with severe physical dysfunction and mental disorders who would not have been able to complete the questionnaire were not eligible to participate in this study.

Ethics Approval and Participation

The study was approved by the institutional review board at the School of Public Health, Zhejiang University (ZGL2020-010), and it was performed in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent; the secondary analysis was allowed without additional consent. All data were anonymous and kept confidential to protect the privacy of participants. This study was required to pay ≥CN¥40 (US $5.8) to each participant.

Data Collection

Well-trained investigators conducted face-to-face interviews with all participants in the presence of family members or in the community. Before the formal survey, community administrators made a home visit or telephone call to older adults living in the community. The community administrators briefly explained the basic situation and recruitment requirements for the project to the older adults and asked them whether they would participate in the survey. Participants who agreed were then scheduled to take the survey. On the day of the survey, the investigator explained the purpose, content, method, and other information regarding the survey to the eligible participants. Subsequently, the investigator received the participants' informed consent. Most of the participants (1835/2036, 90.13%) completed the interview in 50-60 minutes. The questionnaires consisted of 16 parts and included 578 items. The main contents included the following information: demographic characteristics; behavioral habits and general health status; community services and environmental resources; community cohesion; social support and psychological resilience; self-efficacy; neurocognition and social cognition; and assessment of depression, personality disorder, and activities of daily living.

The following general characteristics of the participants were collected: age, sex, education level, individual monthly income, marital status, physical activity level, and dietary habit. Community services and environmental resources were assessed in terms of 4 dimensions: community service facilities, medical resources, nursing resources, and welfare resources. Smoking status was examined using the question *Do you smoke?* If the response was yes, the question *What kind of cigarette do you smoke?* was asked, followed by the question *How many cigarettes do you smoke in a day?* Smoking history and age of starting smoking were also included in the questions. Physical activity status was examined using the question *Do you regularly participate in sports, such as hiking, jogging, playing ball, or swimming?* If the response was yes, details regarding time spent and frequency of participation in physical activity were obtained. Sleep was assessed using the following questions: *How many hours of actual sleep did you have per night during the past month?* and *When do you usually go to bed and get up?* Regarding daily living habits, participants were asked the following questions: *Do you eat at regular times every day? Do you have the habit of drinking tea?* and *Do you have a mobile phone?* Self-perceived health status was assessed using the question *Would you say your health is very good, good, fair, poor, or very poor?* Individuals with chronic diseases were identified using the question *Have you ever been diagnosed with chronic diseases by a doctor?* A total of 28 chronic diseases were included in the options.

Public service facilities in the community were assessed by using the questions *Does your community provide a meal service? Are activity centers for older adults available near your community?* and *Are you satisfied with the clearing and disposal of garbage?* The answer options were *Yes* and *No*. Community medical service resources were assessed using the questions *How far is your home from the nearest medical institution?* and *Does your community provide a family doctor contract service?* Nursing resources were assessed using the questions *Does your community provide day care services?* and *Does your community provide nursing care for patients with severe chronic disease?* The questions used to determine welfare resources included *Does your community organize regular outdoor activities for older adults?* and *Does your community provide transportation services?*

Measurements

Community Cohesion

The neighborhood cohesion instrument was applied to assess community cohesion using 18 items [30]. The neighborhood cohesion instrument consists of scales on 3 dimensions to measure the synthesis of community cohesion concepts. The 3 items—*Overall, I am very attracted to living in this neighborhood*; *Given the opportunity, I would like to move out of this neighborhood*; and *I rarely have neighbors come over to my house to visit*—measure attraction to the neighborhood.

Six items, including *If I needed advice about something I could go to someone in my neighborhood*, *I borrow things and exchange favors with my neighbors*, and *I rarely have neighbors come over to my house to visit*, measure the degree of neighboring. Nine items, including *I feel like I belong to this neighborhood*, *I think I agree with most people in my neighborhood about what is important in life*, and *I feel loyal to the people in my neighborhood*, measure the psychological sense of community. The response options for each item are (1) strongly agree, (2) agree, (3) neither agree nor disagree, (4) disagree, and (5) strongly disagree based on a 5-point Likert scale. The estimated Cronbach α coefficients range from .86 to .95.

Other Measurements

The general self-efficacy scale was used to measure the level of self-efficacy [31]. The general self-efficacy scale consists of 10 items with a total score ranging from 10 to 40 points; the higher the score, the higher the level of self-efficacy. The Chinese version of the Older American Resources and Services (OARS) scale was used to assess the level of social support [32]. The OARS scale consists of 3 dimensions: social interaction, family support, and interpersonal relations. High OARS scores indicate high levels of social support. The Chinese version of the 21-item Dementia Assessment Sheet for Community-Based Integrated Care System (DASC-21) was used to assess individual cognitive functions [33]. A score of ≥27 points suggests possible dementia, and a high score is recognized as indicative of a low level of cognitive function. The Eysenck Personality Questionnaire-Revised Short Scale was used to assess the personality characteristics of participants [34]. Community services and environmental resources included 44 items. However, the item *Help with using the toilet* was removed, given that participants with severe physical dysfunction were not eligible to participate in the study [35].

Statistical Analysis

The data obtained from participants with complete questionnaires (N=2017) were used for statistical analysis. Descriptive statistics were used to report the general characteristics of participants. Frequencies and percentages were computed for the variables.

The means and SDs of community cohesion scores were calculated using participant characteristics. A 2-tailed *t* test was used to compare the mean values between 2 groups, and 1-factor ANOVA was used to compare multiple groups. We conducted a normality test and a variance homogeneity test before performing the *t* test and 1-factor ANOVA.

Three separate logistic regression analyses were conducted to evaluate the association between community cohesion and physical and mental health factors by area (urban vs rural). Community cohesion scores were treated as a binary dependent variable and added to the logistic regression model. Self-perceived health and life satisfaction, cognitive function and depression, and self-efficacy and psychological resilience were added as independent variables to the 3 logistic regression models. Age, sex, marital status, education level, individual monthly income, physical activity level, and dietary habit were adjusted in all models, and the Eysenck Personality Questionnaire-Revised Short Scale scores for psychological factors were also controlled.

Binary logistic regression analysis was used to assess the association between community cohesion and self-perceived social support. The scores of the 3 dimensions for social support and the total scores were added to the logistic regression model. The association of community cohesion with community environmental factors was assessed on 3 dimensions—attraction to neighborhood, degree of neighboring, and psychological sense of community—using binary logistic regression analysis. The total community cohesion score, which was divided into 2 categories, was used as a dependent variable to establish the logistic regression model. For each logistic regression model, we used the stepwise regression method to select the variables. A variable has a power of 80% at a significance level of .05 for univariate analyses that can be added to the logistic regression model. The community cohesion scores among different levels of community services and environmental resources were calculated, and a radar map was drawn.

The association among community cohesion and self-efficacy, social support, cognitive function, community services and environmental resources, and physical and mental health was evaluated through mediating and moderating effect tests based on step-by-step general linear regression model. Statistical significance was set at $P<.05$ for the 2-tailed test. All data analyses were performed using SAS for Windows (version 9.4; SAS Institute Inc).

Results

Participant Characteristics

The general characteristics of participants by sex are shown in Table 1. Of the 2017 participants, 855 (42.39%) were aged ≥70 years, 771 (38.22%) were male, and 1246 (61.77%) were female. A total of 20.48% (413/2017) of the participants reported that they had completed ≥13 years of education. In terms of marital status, 9.5% (73/771) of the male participants and 23.43% (292/1246) of the female participants were unmarried. Overall, 72.93% (1471/2017) of the participants self-reported that they had ≥1 chronic diseases. More than half of the participants (1108/2017, 54.93%) had lived in their community for ≥30 years.

Table 1. Characteristics of study participants by sex (N=2017).

Variable categories	Male participants (n=771), n (%)	Female participants (n=1246), n (%)
Age (years)		
<70	399 (51.8)	763 (61.23)
≥70	372 (48.2)	483 (38.76)
Years of education		
0-6	72 (9.3)	231 (18.53)
7-9	314 (40.7)	464 (37.23)
10-12	216 (28.1)	307 (24.63)
≥13	169 (21.9)	244 (19.58)
Individual monthly income (CN¥ [US $])		
0-1999 (0-291)	256 (33.2)	362 (29.05)
2000-3999 (292-584)	281 (36.5)	621 (49.83)
≥4000 (585)	234 (30.3)	186 (14.92)
Marital status		
Married	698 (90.5)	954 (76.56)
Unmarried	73 (9.5)	292 (23.43)
Smoking status		
Smoker	364 (47.2)	19 (1.52)
Nonsmoker	407 (52.8)	1227 (98.48)
Alcohol use		
Yes	419 (54.4)	152 (12.19)
No	234 (45.6)	1094 (87.80)
Physical activity		
Yes	390 (50.6)	794 (63.72)
No	381 (49.4)	452 (36.28)
Chronic disease status		
Yes	579 (75.1)	892 (71.59)
No	192 (24.9)	354 (28.41)
Years lived in the community		
<30	299 (38.8)	610 (48.96)
≥30	472 (61.2)	636 (51.04)

Community Cohesion Scores Based on Participant Characteristics

Table 2 shows the means and SDs of community cohesion scores based on the characteristics of the participants. According to the logistic regression analysis, education level was negatively associated with the community cohesion score: the lower the education level, the higher the community cohesion score. In addition, we used a generalized linear regression model, and we observed a strong statistically significant linear trend (*P*<.001). Participants' self-reported levels of physical activity were positively associated with community cohesion scores (*P*=.003). The community cohesion scores were higher (*P*<.001) when individuals lived in the community longer.

Table 2. The means and SDs of community cohesion scores by participant characteristics.

Variable categories	Values, mean (SD)	*P* value
Age (years)		.39
<70	68.4 (9.5)	
≥70	69.1 (8.3)	
Sex		.01
Male	67.3 (8.2)	
Female	69.3 (9.2)	
Years of education		<.001
0-6	71.6 (9.3)	
7-9	70.5 (8.2)	
10-12	68.1 (9.6)	
≥13	65.6 (7.9)	
Individual monthly income (CN¥ [US $])		<.001
0-1999 (0-291)	65.3 (8.8)	
2000-3999 (292-584)	70.4 (8.7)	
≥4000 (585)	66.9 (8.8)	
Marital status		.43
Married	68.9 (9.0)	
Unmarried	68.1 (8.5)	
Smoking status		.31
Smoker	69.7 (8.6)	
Nonsmoker	68.6 (9.0)	
Alcohol use		.07
Yes	67.6 (8.7)	
No	69.2 (9.0)	
Physical activity		.003
Yes	69.3 (9.1)	
No	66.5 (8.1)	
Chronic disease status		.41
Yes	68.6 (8.8)	
No	69.3 (9.3)	
Years lived in the community		<.001
<30	67.2 (8.9)	
≥30	70.8 (8.6)	

Community Cohesion–Related Physical and Mental Health Status

The evaluation of community cohesion–related physical and mental health status by area (urban vs rural) is presented in Table 3. The logistic regression analysis showed that high levels of community cohesion were positively associated with life satisfaction, psychological resilience, levels of self-efficacy, and self-perceived health status (odds ratio [OR] 1.21, 95% CI 1.14-1.28; OR 1.05, 95% CI 1.03-1.07; OR 1.11, 95% CI 1.06-1.16; and OR 1.36, 95% CI 1.04-1.78, respectively, in urban areas; and OR 1.15, 95% CI 1.03-1.28; OR 1.07, 95% CI 1.02-1.11; OR 1.02, 95% CI 1.01-1.07; and OR 1.14, 95% CI 0.90-1.83, respectively, in rural areas). High levels of community cohesion were associated with good levels of cognitive function and low levels of depression, and community cohesion was negatively associated with cognitive function and depression scores (OR 0.92, 95% CI 0.87-0.98 and OR 0.80, 95% CI 0.71-0.90, respectively, in urban areas; and OR 0.88, 95% CI 0.79-0.91 and OR 0.92, 95% CI 0.90-0.99, respectively, in rural areas).

Table 3. Community cohesion–related physical and mental health factors by logistic regression analysis.

Variables	Multivariable adjusted odds ratio (95% CI)	*P* value
Urban areas		
Life satisfaction	1.21 (1.14-1.28)	<.001
Cognitive function	0.92 (0.87-0.98)	<.001
Depression	0.80 (0.71-0.90)	<.001
Self-efficacy	1.11 (1.06-1.16)	.002
Psychological resilience	1.05 (1.03-1.07)	<.001
Self-perceived health	1.36 (1.04-1.78)	.03
Rural areas		
Life satisfaction	1.15 (1.03-1.28)	<.001
Cognitive function	0.88 (0.79-0.91)	.02
Depression	0.92 (0.90-0.99)	.04
Self-efficacy	1.02 (1.01-1.07)	.05
Psychological resilience	1.07 (1.02-1.11)	.02
Self-perceived health	1.14 (0.90-1.83)	.06

Association Between Community Cohesion and Self-perceived Social Support

Table 4 shows a positive association between community cohesion and self-perceived social support after education level, dietary habit, age, sex, marital status, physical activity level, and individual monthly income were adjusted (OR 1.27, 95% CI 1.15-1.40; *P*<.001). Community cohesion was significantly associated with the 3 dimensions of social support: social interaction, family support, and interpersonal relations (OR 1.39, 95% CI 1.18-1.65; OR 1.34, 95% CI 1.05-1.72; and OR 1.31, 95% CI 1.13-1.52, respectively).

Table 4. The association between community cohesion and self-perceived social support.

Variables	Multivariable adjusted odds ratio (95% CI)	*P* value
Total scores of social support	1.27 (1.15-1.40)	<.001
Social interaction	1.39 (1.18-1.65)	<.001
Family support	1.34 (1.05-1.72)	.02
Interpersonal relations	1.31 (1.13-1.52)	<.001
Education level	0.72 (0.56-0.92)	.009
Dietary habit	0.85 (0.57-1.27)	.41
Sex	1.15 (0.70-1.89)	.59
Age	1.02 (0.98-1.06)	.32
Marital status	0.63 (0.33-1.19)	.16
Physical activity	1.06 (0.63-1.80)	.83
Individual monthly income	0.84 (0.63-1.11)	.21

Association Between Community Cohesion and Community Services and Environmental Factors

The community cohesion scores from different dimensions associated with community services and environmental factors by logistic regression analysis are shown in Table 5. In model 1, attraction to neighborhood was significantly associated with levels of satisfaction with community physicians' attitude, rehabilitation equipment rental services, caregiver guidance, and community waste disposal (OR 1.55, 95% CI 1.18-2.05; OR 3.91, 95% CI 2.39-6.37; OR 1.85, 95% CI 1.08-3.17; and OR 1.37, 95% CI 1.10-1.71, respectively). In model 2, the degree of neighboring was associated with re-employment assistance, rehabilitation equipment rental services, and levels of satisfaction with community physicians' technical expertise (OR 2.49, 95% CI 1.49-4.18; OR 1.63, 95% CI 1.10-2.42; and OR 1.24, 95% CI 1.01-1.53, respectively). In model 3, psychological sense of community was associated with transportation services, rehabilitation equipment rental services, and health assessment (OR 2.96, 95% CI 1.78-4.92; OR 3.55, 95% CI 2.34-5.39; and OR 2.27, 95% CI 1.28-4.04, respectively). In model 4, the overall levels of community cohesion were associated with transportation services and rehabilitation equipment rental services as well as levels of satisfaction with community physicians' technical expertise and community waste disposal (OR 3.14, 95% CI 1.87-5.28; OR 3.62, 95% CI 2.38-5.52; OR 1.37, 95% CI 1.08-1.73; and OR 1.23, 95% CI 1.01-1.50, respectively).

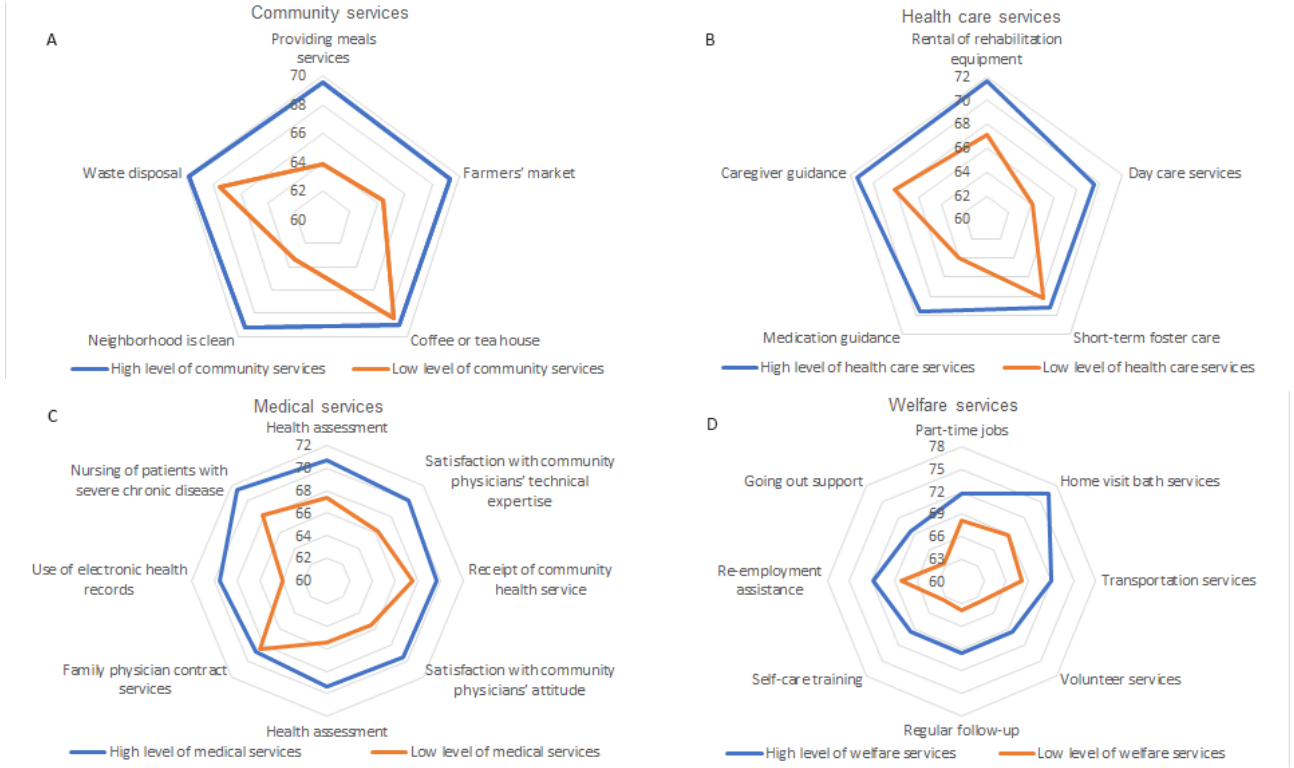
The mean scores of community cohesion from the 4 dimensions of community services and environmental resources are shown in Figure 1. In the community services environment, provision of meals, use of farmers' markets, and neighborhood showed high community cohesion scores (Figure 1A). In health care services, high community cohesion scores were observed in the following services: day care service, medication guidance, rehabilitation equipment rental services, and caregiver guidance (Figure 1B). In medical services, health assessment, satisfaction with community physicians' attitude, receipt of community health service, satisfaction with community physicians' technical expertise, nursing of patients with severe chronic disease, and use of electronic health records showed high community cohesion scores (Figure 1C). In welfare services, home visit bath service, transportation service, volunteer service, regular follow-ups, self-care training, re-employment assistance, going out support, and part-time jobs were associated with high community cohesion scores (Figure 1D).

Table 5. The odds ratios of community environment and other factors related to community cohesion by logistic regression model.

Variables	Multivariable adjusted odds ratio (95% CI)	*P* value
Model 1 (attraction to neighborhood)		
Levels of satisfaction with community physicians' attitude	1.55 (1.18-2.05)	.001
Levels of satisfaction with community waste disposal	1.37 (1.10-1.71)	.005
Rehabilitation equipment rental services	3.91 (2.39-6.37)	<.001
Caregiver guidance	1.85 (1.08-3.17)	.03
Depression scores	0.85 (0.76-0.94)	.001
EPQ-RSS[a] scores	1.09 (1.03-1.16)	.004
Model 2 (degree of neighboring)		
Levels of satisfaction with community physicians' technical expertise	1.24 (1.01-1.53)	.049
Rehabilitation equipment rental services	1.63 (1.10-2.42)	.02
Re-employment assistance	2.49 (1.49-4.18)	<.001
Depression scores	0.90 (0.82-0.99)	.04
Age	0.96 (0.94-0.98)	.006
Model 3 (psychological sense of community)		
Rehabilitation equipment rental services	3.55 (2.34-5.39)	.005
Transportation services	2.96 (1.78-4.92)	<.001
Health assessment	2.27 (1.28-4.04)	<.001
Depression scores	0.86 (0.77-0.96)	.005
EPQ-RSS scores	1.07 (1.01-1.13)	.03
Model 4 (total model)		
Levels of satisfaction with community physicians' technical expertise	1.37 (1.08-1.73)	.008
Levels of satisfaction with community waste disposal	1.23 (1.01-1.50)	.048
Rehabilitation equipment rental services	3.62 (2.38-5.52)	<.001
Transportation services	3.14 (1.87-5.28)	<.001
Depression scores	0.84 (0.75-0.93)	.001
EPQ-RSS scores	1.09 (1.04-1.16)	.001
Self-efficacy scores	1.05 (1.02-1.08)	.004

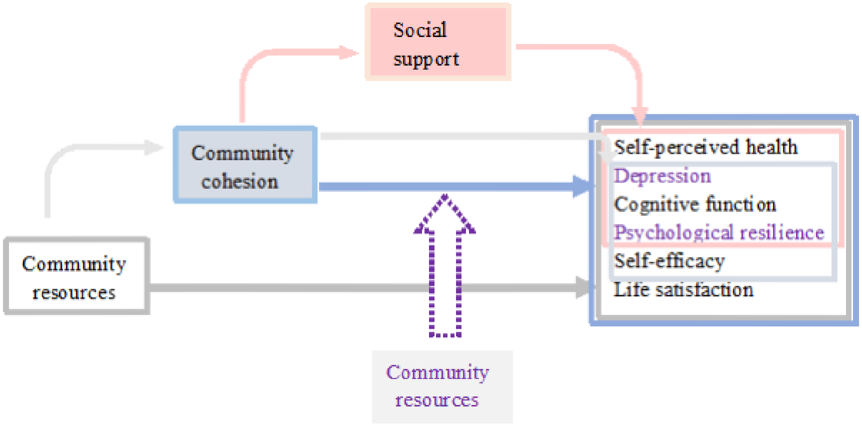
[a] EPQ-RSS: Eysenck Personality Questionnaire-Revised Short Scale.

Figure 1. Mean community cohesion scores. (A) Community cohesion scores for public services. (B) Community cohesion scores for health care services. (C) Community cohesion scores for medical services. (D) Community cohesion scores for welfare services.



The results of the association among community cohesion and self-efficacy, social support, cognitive function, community services and environmental resources, and physical and mental health are shown in Figure 2. Community cohesion, community resources, and social support were directly associated with physical and mental health. Community cohesion was associated with physical and mental health (self-perceived health, depression, cognitive function, and psychological resilience), which may be partially mediated by social support. Community resources were associated with physical and mental health (depression, cognitive function, psychological resilience, and self-efficacy), which may be partially mediated by community cohesion. In addition, the association between community cohesion and psychological resilience and depression may be moderated by community resources.

Figure 2. The association among community cohesion and social support, community resources, and physical and mental health of older adults.



Discussion

Principal Findings

In this study, we found that community cohesion had an important effect on the physical and mental health of older adults during the COVID-19 pandemic. Moreover, high levels of community cohesion were associated with good self-perceived health status and high life satisfaction. Participants who reported high levels of community cohesion stated that they had high levels of self-efficacy, psychological resilience, and self-perceived social support. Community cohesion was positively associated with transportation and rehabilitation equipment rental services as well as levels of satisfaction with community physicians' technical expertise and community waste disposal.

Some studies have reported the beneficial effects of community cohesion on the mental health of older adults; however, studies on the association between community cohesion and physical and mental health during the COVID-19 pandemic, especially studies on community services and environment factors that can improve community cohesion during major infectious disease epidemics, are limited [36,37]. The COVID-19 pandemic has brought about major changes in the daily lives and environment of older adults. Social support from children and relatives was unavailable because of COVID-19–related restrictions, regular visits to medical facilities were banned, and public transportation was inaccessible because they could not operate intelligent public transportation apps on their mobile phones. As they could not acquire and show health codes, as a result, they will be refused public transportation, which would lead them to feel more panicked and helpless. The restrictions on the use of public resources also caused them to become more dependent on the community. However, the implementation of COVID-19 prevention and control measures increased the interaction between the older adults and community staff, and this proved to be of considerable help to the older adults. Free testing to assess the prevalence of COVID-19 in the community made them feel safer. The community became a big presence in their lives. These pieces of evidence from the field supported our findings.

Studies have shown that rural areas have more neighborhood networks than urban areas [38], but community cohesion for health benefits between urban and rural areas demonstrates no difference among the general population. Our study showed that, during the COVID-19 pandemic, there were no differences between individuals with high levels of community cohesion association and good levels of physical and mental health in different areas (urban vs rural). However, the association between community cohesion and self-perceived health was inconsequential in rural areas. The prevalence of COVID-19 has resulted in the return of people to their community, which supports the observation on the effects of community cohesion on the physical and mental health of older adults during the COVID-19 pandemic. Our results showed that high levels of community cohesion were associated with high levels of self-perceived health status and life satisfaction. However, older adults were highly anxious about their health, and their lives were grossly circumscribed. During the pandemic, because family and social support for older adults was hindered, they became emotionally dependent on their community, and the rising levels of community cohesion enabled them to experience a high level of social support. Several studies have suggested that community cohesion can be viewed as a pattern of social support, which might affect individual mental health through enhancing mutual trust and emotional support as well as reducing stress levels [39].

Community cohesion can be linked to individual health benefits, but the underlying mechanism is still not fully understood. Studies have suggested that community cohesion enhances subjective well-being by promoting positive emotions and purpose in life, which acts as a barrier to psychopathology factors, resulting in better physical health and longevity [40-42]. Community cohesion can strengthen collective advocacy for resources, which promotes the dissemination of health-related information and increases awareness of chronic disease [43]. However, there is limited quantitative research on the association between community cohesion and physical and mental health among older adults during the COVID-19 pandemic. Our results showed that high levels of community cohesion were associated with low levels of depression and high levels of psychological resilience and cognitive function. This result can be easily explained by qualitative observation. During the pandemic, community prevention and control measures resulted in increased communication between individuals and community staff. There was also an increase in the exchange of information among neighbors on disease prevalence and prevention. In these circumstances, community cohesion and individual sense of security were enhanced, promoting the spread of health-related information and improving the recognition for diseases among older adults. A study has reported on the association between community cohesion and the use of preventive health services [44]. The study suggested that high levels of community cohesion could be positively associated with an increase in the use of influenza vaccine and cholesterol tests for community individuals through 4 hypothesized mechanisms: increasing the diffusion of information, providing emotional support, advocating for environmental resources, and maintaining healthy behaviors through informal social control.

There is growing interest in studying the effects of the environment on health and health-related outcomes, such as identifying the environmental resources and community services that may enhance community cohesion. In this study, we found that transportation and rehabilitation equipment rental services as well as levels of satisfaction with community physicians' technical expertise and community waste disposal were positively associated with community cohesion. These factors are extremely important for improving community cohesion. During the pandemic, operating intelligent public transportation apps on their mobile phones was a huge barrier for older adults when taking public transport. Being confined to the home for long periods and the restrictions on going out led to difficulties in physical functioning, which increased the use of, and demand for, transportation and rehabilitation equipment rental services among older adults. The COVID-19 outbreak made older adults more dependent on community physicians. A high level of technical expertise demonstrated by community physicians can make older adults living in the community feel safe and reduce their fear of disease. At the initial stage of the pandemic, normal garbage disposal was also disrupted. The normalization of community garbage disposal and the clean and tidy community environment once again made the community attractive to older adults. These factors directly or indirectly affected the physical and mental health of older adults during the COVID-19 pandemic.

The main strength of this study is that it was based on the use of big data platforms to study the mental health of older adults during the COVID-19 pandemic. We collected complete data, which enabled us to consider more indicators of mental health measurement and community cohesion–related factors for the analysis and control of many potential confounding factors. The large amount of data also ensured the reliability of the study results. This study included ≥10 mental health measurement instruments and identified the association between community cohesion and the physical and mental health of older adults at the individual, community, and social levels based on the social-ecological model. Using moderation and mediation analysis, this study further revealed community cohesion to be a potential interaction mechanism linking social support, self-efficacy, environmental resources, and the physical and mental health of older adults. This study also identified specific community services and environmental factors that can increase community cohesion and that can be considered when designing public health policies to deal with major infectious disease pandemics. Furthermore, when the COVID-19 pandemic initially broke out in China, the country's public health administration and disease control departments implemented a series of unprecedented prevention and control measures, including a *dynamic clearing* epidemic prevention policy and home isolation. This provided us with a rare opportunity to study the effects of community cohesion on the physical and mental health of older adults.

Study Limitations

Our study includes several limitations. First, the causality could not be determined, given that this study adopted a cross-sectional design. However, the associations between community cohesion and community services and environmental factors were observed through relatively large sample sizes, and the inference was based on some previous studies. Second, all participants in this study were older adults. During the first stage of the study, the face-to-face interviews took approximately 45 minutes to complete. Each participant was compensated with CN¥40 (US $5.8), based on local research payment standards and as required by the ethics committee. After the items to be investigated were added, the survey time was extended to approximately 60 minutes, and participants were offered higher remuneration accordingly. It is possible that the compensation offered could have led prospective participants to deceive the researchers regarding their eligibility, and the information they provided could have been biased [45]. Third and last, the study was conducted during the COVID-19 pandemic. We selected the participants according to the inclusion and exclusion criteria; however, we excluded some older adults with severe physical dysfunction and mental disorders, which could have resulted in potential exclusion bias.

Conclusions

We obtained important evidence on the effects of community services and environmental factors on community cohesion during the COVID-19 pandemic. We found that community cohesion is substantially positively associated with transportation and rehabilitation equipment rental services as well as high levels of satisfaction with community physicians' technical expertise and community waste disposal. Community cohesion is also directly or indirectly associated with the physical and mental health of older adults. Our research suggests that enhancing community services and environmental resources is an effective strategy to increase community cohesion during major infectious disease epidemics. Improving community cohesion will also help to promote the physical and mental health of older adults.

Acknowledgments

This work was supported by the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang province (2019R01007) and in part by the Key Laboratory of Intelligent Preventive Medicine of Zhejiang province (2020E10004).

Data Availability

The data sets generated and analyzed during this study are not publicly available to ensure the privacy of participants but are available from the corresponding author on reasonable request.

Authors' Contributions

XWD, AA, and YL were responsible for the study design. YYP, YC, and DBH contributed to the study selection and data extraction. XWD, AA, YYP, and YL contributed to writing the manuscript. All authors have reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Walsh JJ. Meeting compliance under Medicare and Medicaid emergency preparedness Final Rule requirements. J Bus Contin Emer Plan 2020 Jan 01;14(2):178-188. [Medline: 33239148]
2. Jopp DS, Boerner K, Ribeiro O, Rott C. Life at age 100: an international research agenda for centenarian studies. J Aging Soc Policy 2016 Jul;28(3):133-147. [doi: 10.1080/08959420.2016.1161693] [Medline: 26984376]
3. Jeste DV, Blazer 2nd DG, Buckwalter KC, Cassidy KL, Fishman L, Gwyther LP, et al. Age-friendly communities initiative: public health approach to promoting successful aging. Am J Geriatr Psychiatry 2016 Dec;24(12):1158-1170 [FREE Full text] [doi: 10.1016/j.jagp.2016.07.021] [Medline: 27742528]

4. Baron M, Fletcher C, Riva M. Aging, health and place from the perspective of elders in an Inuit community. J Cross Cult Gerontol 2020 Jun;35(2):133-153. [doi: 10.1007/s10823-020-09398-5] [Medline: 32409899]
5. Vanleerberghe P, De Witte N, Claes C, Schalock RL, Verté D. The quality of life of older people aging in place: a literature review. Qual Life Res 2017 Nov;26(11):2899-2907. [doi: 10.1007/s11136-017-1651-0] [Medline: 28707047]
6. Hippisley-Cox J, Coupland CA, Mehta N, Keogh RH, Diaz-Ordaz K, Khunti K, et al. Risk prediction of COVID-19 related death and hospital admission in adults after COVID-19 vaccination: national prospective cohort study. BMJ 2021 Sep 17;374:n2244 [FREE Full text] [doi: 10.1136/bmj.n2244] [Medline: 34535466]
7. Demetriou L, Chatzi V, Panaoura R. The impact of COVID-19 on anxiety and loneliness levels of elderly people and the role of counseling on their wellbeing. Psychol Res 2021 Sep 28;11(9):369-385 [FREE Full text] [doi: 10.17265/2159-5542/2021.09.001]
8. Dahlberg L. Loneliness during the COVID-19 pandemic. Aging Ment Health 2021 Jul;25(7):1161-1164. [doi: 10.1080/13607863.2021.1875195] [Medline: 33491474]
9. Wilson JM, Lee J, Shook NJ. COVID-19 worries and mental health: the moderating effect of age. Aging Ment Health 2021 Jul;25(7):1289-1296. [doi: 10.1080/13607863.2020.1856778] [Medline: 33291948]
10. Nakagawa T, Hülür G. Social integration and terminal decline in life satisfaction among older Japanese. J Gerontol B Psychol Sci Soc Sci 2020 Nov 13;75(10):2122-2131. [doi: 10.1093/geronb/gbz059] [Medline: 31074829]
11. Guida JL, Holt CL, Dallal CM, He X, Gold R, Liu H. Social relationships and functional impairment in aging cancer survivors: a longitudinal social network study. Gerontologist 2020 May 15;60(4):607-616 [FREE Full text] [doi: 10.1093/geront/gnz051] [Medline: 31050729]
12. Cugmas M, Ferligoj A, Kogovšek T, Batagelj Z. The social support networks of elderly people in Slovenia during the COVID-19 pandemic. PLoS One 2021 Mar 03;16(3):e0247993 [FREE Full text] [doi: 10.1371/journal.pone.0247993] [Medline: 33657172]
13. Schiefer D, van der Noll J. The essentials of social cohesion: a literature review. Soc Indic Res 2016 Apr 6;132(2):579-603 [FREE Full text] [doi: 10.1007/s11205-016-1314-5]
14. Gan DR, Fung JC, Cho IS. Neighborhood atmosphere modifies the eudaimonic impact of cohesion and friendship among older adults: a multilevel mixed-methods study. Soc Sci Med 2021 Feb;270:113682. [doi: 10.1016/j.socscimed.2021.113682] [Medline: 33461036]
15. Elliott J, Gale CR, Parsons S, Kuh D, HALCyon Study Team. Neighbourhood cohesion and mental wellbeing among older adults: a mixed methods approach. Soc Sci Med 2014 Apr;107:44-51 [FREE Full text] [doi: 10.1016/j.socscimed.2014.02.027] [Medline: 24602970]
16. Holmes LM, Marcelli EA. Neighborhood social cohesion and serious psychological distress among Brazilian immigrants in Boston. Community Ment Health J 2020 Jan;56(1):149-156 [FREE Full text] [doi: 10.1007/s10597-019-00468-8] [Medline: 31535346]
17. Miao J, Wu X, Sun X. Neighborhood, social cohesion, and the elderly's depression in Shanghai. Soc Sci Med 2019 May;229:134-143. [doi: 10.1016/j.socscimed.2018.08.022] [Medline: 30194018]
18. Yu R, Leung G, Chan J, Yip BH, Wong S, Kwok T, et al. Neighborhood social cohesion associates with loneliness differently among older people according to subjective social status. J Nutr Health Aging 2021;25(1):41-47. [doi: 10.1007/s12603-020-1496-z] [Medline: 33367461]
19. Kim J, Park MJ. Multilevel effect of neighborhood social cohesion and characteristics on suicidal ideation among Korean older adults. Community Ment Health J 2021 Apr;57(3):522-528. [doi: 10.1007/s10597-020-00678-5] [Medline: 32676878]
20. Gyasi RM, Yeboah AA, Mensah CM, Ouedraogo R, Addae EA. Neighborhood, social isolation and mental health outcome among older people in Ghana. J Affect Disord 2019 Dec 01;259:154-163. [doi: 10.1016/j.jad.2019.08.024] [Medline: 31445342]
21. Hikichi H, Aida J, Tsuboya T, Kondo K, Kawachi I. Can community social cohesion prevent posttraumatic stress disorder in the aftermath of a disaster? A natural experiment from the 2011 Tohoku earthquake and tsunami. Am J Epidemiol 2016 May 15;183(10):902-910 [FREE Full text] [doi: 10.1093/aje/kwv335] [Medline: 27026337]
22. Saju MD, Benny AM, Preet Allagh K, Joseph B, Amuthavalli Thiyagarajan J. Relationship between neighbourhood cohesion and disability: findings from SWADES population-based survey, Kerala, India. F1000Res 2020 Jul 13;9:700 [FREE Full text] [doi: 10.12688/f1000research.25073.1] [Medline: 32832072]
23. Mujahid MS, Diez Roux AV, Morenoff JD, Raghunathan TE, Cooper RS, Ni H, et al. Neighborhood characteristics and hypertension. Epidemiology 2008 Jul;19(4):590-598. [doi: 10.1097/EDE.0b013e3181772cb2] [Medline: 18480733]
24. Kim ES, Hawes AM, Smith J. Perceived neighbourhood social cohesion and myocardial infarction. J Epidemiol Community Health 2014 Nov;68(11):1020-1026 [FREE Full text] [doi: 10.1136/jech-2014-204009] [Medline: 25135074]
25. Clark CJ, Guo H, Lunos S, Aggarwal NT, Beck T, Evans DA, et al. Neighborhood cohesion is associated with reduced risk of stroke mortality. Stroke 2011 May;42(5):1212-1217 [FREE Full text] [doi: 10.1161/STROKEAHA.110.609164] [Medline: 21493914]
26. Martin CL, Ward-Caviness CK, Dhingra R, Zikry TM, Galea S, Wildman DE, et al. Neighborhood environment, social cohesion, and epigenetic aging. Aging (Albany NY) 2021 Mar 14;13(6):7883-7899 [FREE Full text] [doi: 10.18632/aging.202814] [Medline: 33714950]

27. Robinette JW, Boardman JD, Crimmins E. Perceived neighborhood social cohesion and cardiometabolic risk: a gene × environment study. Biodemography Soc Biol 2020 Jan;65(1):1-15. [doi: 10.1080/19485565.2019.1568672] [Medline: 32065540]
28. Zhang W, Liu S, Zhang K, Wu B. Neighborhood social cohesion, resilience, and psychological well-being among Chinese older adults in Hawai'i. Gerontologist 2020 Feb 24;60(2):229-238. [doi: 10.1093/geront/gnz104] [Medline: 31322675]
29. Carter ED, Cordero ML. Salir Adelante: social capital and resilience during the COVID-19 pandemic in Argentina. Health Place 2022 Sep;77:102870 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102870] [Medline: 35933852]
30. Buckner JC. The development of an instrument to measure neighborhood cohesion. Am J Community Psychol 1988 Dec;16:771-791 [FREE Full text] [doi: 10.1007/bf00930892]
31. Leung DY, Leung AY. Factor structure and gender invariance of the Chinese general self-efficacy scale among soon-to-be-aged adults. J Adv Nurs 2011 Jun;67(6):1383-1392. [doi: 10.1111/j.1365-2648.2010.05529.x] [Medline: 21129011]
32. Fu D, Shen Y, Xia Z. The reliability of Chinese version of the OARS. J Pre Med Inf 1997;13:193-197.
33. Chen Y, Cui P, Pan Y, Li Y. Use of the dementia assessment sheet for community-based integrated care system 21 items among Chinese community-dwelling older adults. Geriatr Gerontol Int 2021 Aug;21(8):705-711. [doi: 10.1111/ggi.14215] [Medline: 34164889]
34. Qian M, Wu G, Zhu R, Zhang S. Development of the revised eysenck personality questionnaire short scale for Chinese (EPQ-RSC). Psychol 2000;32(3):317-323 [FREE Full text]
35. Wang ZZ, Lan Y, Wang L. Research on the service accessibility of home care community: the study based on the survey data of Suzhou city. Popul Dev 2016;3:90-99.
36. Diener E, Chan MY. Happy people live longer: subjective well-being contributes to health and longevity. Appl Psychol 2010;3(1):1-43 [FREE Full text] [doi: 10.1037/e675972011-001]
37. Cramm JM, Nieboer AP. Social cohesion and belonging predict the well-being of community-dwelling older people. BMC Geriatr 2015 Mar 21;15:30 [FREE Full text] [doi: 10.1186/s12877-015-0027-y] [Medline: 25879773]
38. Han Y, Chung RY. The role of neighborhood social capital on health and health inequality in rural and urban China. Prev Med 2022 Mar;156:106989. [doi: 10.1016/j.ypmed.2022.106989] [Medline: 35150751]
39. Rosenblatt AM, Crews DC, Powe NR, Zonderman AB, Evans MK, Tuot DS. Association between neighborhood social cohesion, awareness of chronic diseases, and participation in healthy behaviors in a community cohort. BMC Public Health 2021 Sep 03;21(1):1611 [FREE Full text] [doi: 10.1186/s12889-021-11633-8] [Medline: 34479522]
40. Kim ES, Kawachi I. Perceived neighborhood social cohesion and preventive healthcare use. Am J Prev Med 2017 Aug;53(2):e35-e40 [FREE Full text] [doi: 10.1016/j.amepre.2017.01.007] [Medline: 28214249]
41. Bekalu MA, McCloud RF, Minsky S, Viswanath K. Association of social participation, perception of neighborhood social cohesion, and social media use with happiness: evidence of trade-off (JCOP-20-277). J Community Psychol 2021 Mar;49(2):432-446 [FREE Full text] [doi: 10.1002/jcop.22469] [Medline: 33169391]
42. Salvatore MA, Grundy E. Area deprivation, perceived neighbourhood cohesion and mental health at older ages: a cross lagged analysis of UK longitudinal data. Health Place 2021 Jan;67:102470 [FREE Full text] [doi: 10.1016/j.healthplace.2020.102470] [Medline: 33212393]
43. Hetherington E, McDonald S, Wu M, Tough S. Risk and protective factors for mental health and community cohesion after the 2013 Calgary flood. Disaster Med Public Health Prep 2018 Aug;12(4):470-477. [doi: 10.1017/dmp.2017.91] [Medline: 28770699]
44. Yu R, Cheung O, Leung J, Tong C, Lau K, Cheung J, et al. Is neighbourhood social cohesion associated with subjective well-being for older Chinese people? The neighbourhood social cohesion study. BMJ Open 2019 May 10;9(5):e023332 [FREE Full text] [doi: 10.1136/bmjopen-2018-023332] [Medline: 31079078]
45. Fernandez Lynch H, Joffe S, Thirumurthy H, Xie D, Largent EA. Association between financial incentives and participant deception about study eligibility. JAMA Netw Open 2019 Jan 04;2(1):e187355 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.7355] [Medline: 30681707]

Abbreviations

DASC-21: 21-item Dementia Assessment Sheet for Community-Based Integrated Care System
OARS: Older American Resources and Services
OR: odds ratio

Edited by A Mavragani, T Sanchez; submitted 16.12.22; peer-reviewed by O Navarro, J Hu; comments to author 19.01.23; revised version received 25.01.23; accepted 14.03.23; published 14.04.23.

Please cite as:
Li Y, Ding X, Aierken A, Pan Y, Chen Y, Hu D
The Role of Community Cohesion in Older Adults During the COVID-19 Epidemic: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e45110
URL: https://publichealth.jmir.org/2023/1/e45110
doi:10.2196/45110
PMID:36921236

Original Paper

Maintaining Adherence to COVID-19 Preventive Practices and Policies Pertaining to Masking and Distancing in the District of Columbia and Other US States: Systematic Observational Study

Monica S Ruiz[1*], MPH, PhD; Mercedes V McMahon[1*], MPH; Hannah Latif[1*], MPH; Amita Vyas[1*], PhD

Department of Prevention and Community Health, Milken Institute School of Public Health, George Washington University, Washington, DC, United States

[*] all authors contributed equally

Corresponding Author:
Monica S Ruiz, MPH, PhD
Department of Prevention and Community Health
Milken Institute School of Public Health
George Washington University
950 New Hampshire Ave NW
Suite 300
Washington, DC, 20052
United States
Phone: 1 202 994 3676
Email: msruiz@gwu.edu

Abstract

Background: Prior to the development of effective vaccines against SARS-CoV-2, masking and social distancing emerged as important strategies for infection control. Locations across the United States required or recommended face coverings where distancing was not possible, but it is unclear to what extent people complied with these policies.

Objective: This study provides descriptive information about adherence to public health policies pertaining to mask wearing and social distancing and examines differences in adherence to these policies among different population groups in the District of Columbia and 8 US states.

Methods: This study was part of a national systematic observational study using a validated research protocol for recording adherence to correct mask wearing and maintaining social distance (6 feet/1.83 meters) from other individuals. Data were collected from December 2020 to August 2021 by research team members who stationed themselves in outdoor areas with high pedestrian traffic, observed individuals crossing their paths, and collected data on whether individuals' masks were present (visible or not visible) or worn (correctly, incorrectly, not at all) and whether social distance was maintained if other individuals were present. Observational data were entered electronically into Google Forms and were exported in Excel format for analysis. All data analyses were conducted using SPSS. Information on local COVID-19 protection policies (eg, mask wearing requirements) was obtained by examining city and state health department websites for the locations where data were being collected.

Results: At the time these data were collected, most locations in our study required (5937/10,308, 57.6%) or recommended (4207/10,308, 40.8%) masking. Despite this, more than 30% of our sample were unmasked (2889/10136, 28.5%) or masked incorrectly (636/10136, 6.3%). Masking policy was significantly related to correct masking with locations that required or recommended masking (66% correct masking vs 28/164, 17.1% in locations that did not require masking, $P<.001$). Participants who maintained social distance from others were more likely to be correctly masked than those who were not ($P<.001$). Adherence to masking policy by location was significant ($P<.001$); however, this was driven by 100% compliance in Georgia, which did not require masks at any point during the data collection period. When the same analysis was conducted for compliance with mask requirements and recommendations, there was no significant difference by location. Overall adherence to masking policies was 66.9%

Conclusions: Despite a clear relationship between mask policies and masking behavior, one-third of our sample was nonadherent to those policies, and approximately 23% of our sample did not have any mask, either on or visible. This may speak to the confusion surrounding "risk" and protective behaviors, as well as pandemic fatigue. These results underscore the importance of clear public health communication, particularly given variations in public health policies across states and localities.

(JMIR Public Health Surveill 2023;9:e40138) doi:10.2196/40138

KEYWORDS

COVID-19; mask adherence; social distancing; public health; health policy; public health mandates

Introduction

Since the novel COVID-19 pandemic became known in January 2020, the United States has experienced over 103 million cases and over 1.1 million deaths (cumulative as of February 25, 2023) [1]. From the beginning of the pandemic, certain communities have been shown to be disproportionately affected by COVID-19; a recent systematic review of disparities in COVID-19 infections, hospitalizations, and deaths showed that African Americans/Blacks and Hispanic/Latinx populations had disproportionately higher rates of SARS-CoV-2 infection, hospitalization, and COVID-19–related mortality compared with non-Hispanic/Latinx White populations [2]. Other data show that older age and presence of health comorbidities—including presence of obesity and diabetes, cardiovascular disease, and chronic kidney disease—are associated with higher risk of COVID-19 morbidity and mortality [3].

Although the development and widespread dissemination of effective vaccines to prevent COVID-19 infection have helped to reduce infection rates, voluntary uptake of the vaccine has been slower than expected. National surveillance data show that, although 81.1% of the total US population—including 92% of adults over the age of 18 years—have had at least 1 dose of the SARS-CoV-2 vaccine, only 69.2% completed the primary series and only 16% have received an updated bivalent booster dose [4]. Nonetheless, vaccine uptake, combined with continued preventive behaviors (ie, mask wearing and social distancing), remain the most effective ways to minimize infection and transmission risk [5-9].

The importance of mask wearing continues to be an effective way to protect against infection against COVID-19 and other respiratory infections. There have been numerous studies of the effectiveness of mask wearing against SARS-CoV-2 infection. For example, in a cross-sectional, web-based study of US-based individuals, Rader et al [10] found that a reported 10% increase in mask wearing was associated with more than a 3-fold increase in infection control and that communities with high reported mask wearing and social distancing had the highest predicted probability of transmission control. In another study from Japan that surveyed close contacts of SARS-CoV-2–infected patients, Sugimura et al [11] found that the individuals who indicated that they did not wear face masks were infected at a higher rate than mask wearers (16.4% compared with 7.1%). A comprehensive narrative review published in 2021 by Howard et al [12] found support for the effectiveness of public mask wearing, stating that "...public mask wearing is most effective at reducing spread of the virus when compliance is high...in conjunction with existing hygiene, distancing, and contact tracing strategies." They further recommended that "public officials and governments strongly encourage the use of widespread face masks in public, including the use of appropriate regulation" [12]. The importance of continued masking is underscored by the presence of highly transmissible variants such as the Omicron variants, which, combined, accounted for 93.8% of infections in the United States (as of February 25, 2023) [13] and are contagious even among those who are fully vaccinated and boosted.

The issue of regulation—such as the enforcement of vaccine and mask wearing mandates—for the protection of public health has historically been and continues to be contentious [14]. In the way the COVID-19 pandemic unfolded in the United States, general vaccine hesitancy, politicization of governmental mandates, and the politicization of the pandemic itself served to hamper timely uptake and adherence to public health recommendations for mask wearing at a time when infections were rapidly spreading and vaccines had not yet become available. The politicization of the pandemic and mask wearing by conservative governmental leaders was particularly noteworthy; in the absence of a national mask wearing mandate, individual states implemented legislation requiring the wearing of face masks in public. In August 2020, while the epidemic was sweeping through the United States, 34 states and the District of Columbia had mask mandates, while 16 states—each of which had a Republican governor—did not [14]. A study that examined adherence to state-level masking policies and SARS-CoV-2 infection rates in those states at around this same time period in 2020 found that 93% of the states that had no mask wearing policies for the general public reported high COVID-19 rates, whereas none of the 8 states that had at least 75% adherence to masking policies reported high infection rates [15].

The effectiveness of a public health mandate to protect against disease spread may also be dependent on the level of mobilization of people in and and out of the community. For example, efforts to prevent SARS-CoV-2 infection and transmission in metropolitan regions may be especially tricky given the movement of people into and out of cities and their surrounding suburban areas, and Black and Indigenous People of Color (BIPOC) communities may have elevated levels of vulnerability to COVID-19 because of their participation in sectors of the workforce that are deemed essential. For example, the pandemic picture in the Washington, DC, region—which includes the District of Columbia and immediate surrounding jurisdictions in Maryland and Virginia—mirrors that of the rest of the nation in terms of its impact on disproportionately affected populations [16-18]. Because of its location between Maryland and Virginia, the District has very fluid boundaries and experiences a significant amount of population migration to and from these states as people enter and exit the District for employment, commerce, and tourism. As a result, the level of infection risk for SARS-CoV-2 in the District may be affected by factors in its neighboring jurisdictions, including the R0 (basic reproduction number), rate of vaccination uptake, and levels of social and community vulnerability that might increase SARS-CoV-2 infection risk. For these reasons, adherence to preventive behaviors in urban centers is of paramount importance.

Although there are data about the adherence to these preventive behaviors for other locations in the country, little is known about the levels of adherence to these behaviors in the metropolitan Washington, DC area and how this region compares with other metropolitan centers. The purpose of this paper was to provide descriptive information about adherence to mask wearing and social distancing in the District of Columbia and 8 other US states and to examine differences in adherence among different population groups within these locations.

Methods

This study was conducted by the George Washington University team that was part of the larger, national-level Systematic Observation of Mask Adherence and Distancing (SOMAD) study [19]. Although the methodological details pertaining to this study are provided elsewhere [20], we will also explain them in the following sections to provide clarity about the study procedures.

Student Researcher Training

The research team included 13 student researchers who were trained by the SOMAD principal investigator (via Zoom meeting) on the study protocol [21] and data collection procedures. The student researchers spent approximately 1 week practicing the data collection procedures to ensure comfort with and ability to adhere to the study protocol. Issues that arose during the "practice data collection" sessions were discussed with our university's faculty team leaders to ensure that study procedures were clear and that data collection practices were adherent with the master SOMAD protocol. Student researchers were advised that, although it may be difficult to ascertain exact ages, genders, and race/ethnicity of persons observed, they should use their best judgment based on the way the person is presenting.

Data Collection Methods

Our study was conducted from December 1, 2020, to August 1, 2021. During this time, the university was operating virtually, so student researchers were attending classes and research team meetings from their home locations. Student researchers chose places close to home as their observation sites. Although most data were collected in the metropolitan Washington, DC area (which includes parts of Maryland and Northern Virginia), we did have team members who were collecting data in other states. As such, in addition to data from the metropolitan Washington, DC area, data were collected from Georgia, Missouri, Arizona, California, Nevada, and Illinois.

As these data were collected prior to the wider availability of COVID-19 vaccines, all observation locations were outdoors and in high pedestrian traffic areas, such as parks, multi-use paths and trails (eg, walking, running, and biking trails), city intersections, and shopping areas that had outdoor pedestrian walkways and spaces. Student researchers wore masks and maintained social distance from others while collecting data to protect themselves from infection risk.

Per the SOMAD protocol [21], the student researcher (acting as the observer) selected a specific spot for data collection. As individuals passed by that spot, the student researcher recorded the characteristics of that person, entered those data into the online data collection form, and then continued to observe and enter data for the next person who passed by. If persons were walking in groups, student researchers would assess the general characteristics of 4 persons within each group (in order to retain as much detailed information as possible) and enter those data into the online form for each person.

Data were collected for a minimum of 2 hours per week, with 1 observation session occurring on a weekday and 1 session occurring on a weekend. Both observation sessions occurred at the same time of day to ensure comparability between weekdays and weekends. Data were collected using the national SOMAD team's online data collection form. Data were then cleaned by the national team's statistician and returned to our university team for analyses.

Measures

Data were collected pertaining to a variety of characteristics related to the outdoor space where data collection occurred and the persons being observed. These variables are described in the following sections.

Location and Setting Variables

At the start of each observation session, the student researcher recorded the date, time, city, address, and zip code in which the observation was occurring. Additionally, student researchers recorded the type of location where the observation was occurring (eg, commercial street, neighborhood park, trail) and, if applicable, the official name of the location (eg, the specific name of the park or trail). Student researchers also noted if the location had a masking policy and whether there was sufficient space in the location for social distancing, which is defined as being able to maintain 6 feet or 1.83 meters from another person.

Demographic Variables

With regard to the persons being observed, it was understood that it is often difficult to ascertain with certainty the demographic characteristics of individuals simply by looking at physical characteristics. Such judgments may be made even more difficult if individuals are wearing face coverings. Nonetheless, student researchers were asked to do the best they could to ascertain the approximate age, gender, and race/ethnicity of the persons being observed. Options for age group included infant or toddler (0-2 years of age), child (3-12 years of age), adolescent or teen (13-19 years of age), adult (20-59 years of age), and older adult (60 years of age and older). With regard to gender and race/ethnicity, student researchers were asked to note the gender (male, female, unable to judge) and apparent race/ethnicity (White, Black/African American, Asian, Latinx, unable to determine) of each observed individual.

Additionally, as these observations were occurring in outdoor spaces, student researchers were asked to record the physical activity levels of the individuals. Individuals could be categorized as being sedentary (eg, sitting, laying down, or being carried in the case of a child), moderately active (eg, walking or biking slowly), or vigorously active (eg, running, biking, climbing). Similarly, individuals using modes of transportation were recorded as being either "on wheels" (ie, being on a bicycle or skateboard or in a wheelchair or stroller) or "not on wheels" (ie, being on foot).

COVID-19 Protective Practices

Student researchers collected data on the adherence to COVID-19 protective practices of mask wearing and social distancing. For mask wearing, individuals were noted as having their masks on correctly (ie, covering the nose and mouth), wearing their masks incorrectly (eg, mask covering mouth but not nose or vice versa), having visible masks present but not on their face (eg, masks in hand, dangling from an ear, or being worn under the chin), or having no face covering at all. With regard to social distancing, student researchers recorded if observed individuals were traveling alone or in a group (defined as 2 or more people together) and whether individuals (including those traveling in groups) were maintaining social distance from others.

Policy Variables

In addition to the notation about masking policies of specific locations where observations occurred, we collected data pertaining to state and local masking and social distancing policies through examination of city and state health department websites for the locations where data were being collected.

Analyses

Observational data were entered electronically into Google Forms and were then exported in Excel format for analysis. Univariate and bivariate analyses were conducted to obtain descriptive information about the observational sample. Pearson's chi-square analyses were used to test for significant associations between study variables. All data analyses were conducted using SPSS version 24.

Ethical Considerations

Prior to engaging in data collection, the study protocol was submitted to the George Washington University's Institutional Review Board for review. A waiver of informed consent was requested due to (1) the nature of the study being purely observational and not requiring personal interaction with prospective participants (ie, those persons being observed) and (2) the fact that no protected personally identifying information or protected health data were collected from individuals who were observed as part of the study. Because no personally identifying information were collected, all data from persons observed were anonymous and analyzed in aggregate. Finally, as these data were collected without involving interpersonal interaction, individuals whose behaviors were observed as part of this study were not offered compensation. Based on these aspects of the study, the George Washington University Institutional Review Board determined that this study constituted minimal risk to research participants and therefore was exempt from review (IRB# NCR213396).

Results

Sample Demographics

The findings presented here represent data collected between November 29, 2020, and August 1, 2021, in multiple locations across 8 states and the District of Columbia. The geographic locations of our study sites in the United States are shown in Figure 1. In addition to geographic location, we have provided the daily average of new COVID-19 cases in each location (using data from the New York Times COVID-19 Tracker) [22] in order to demonstrate how the pandemic was affecting each state or, where more specific data were available, each city at the approximate date when data collection started.

Demographic data about the individuals observed are presented in Table 1. A total of 10,308 observations were collected. The sample was relatively evenly split between persons observed to be male and persons observed to be female (4539/10,081, 45% vs 5542/10,081, 55%), with most persons observed being adults between the ages of 20 years and 59 years (7490/10,118, 74%), of White or Black race (5271/10,308, 51.1% and 2623/10,308, 25.4% respectively), and of non-Latinx ethnicity (8928/9822, 90.9%). Of the observations, 47% (4836/10,291) were collected in the metropolitan Washington, DC area, and 38% (3942/10,291) came from 3 sites in the metropolitan Atlanta area in Georgia. Most data collection locations had policies requiring (5937/10,308, 57.6%) or recommending (4207/10,308, 40.8%) masking in outdoor locations.

Figure 1. Study locations and daily average new COVID-19 cases (as of December 2020).

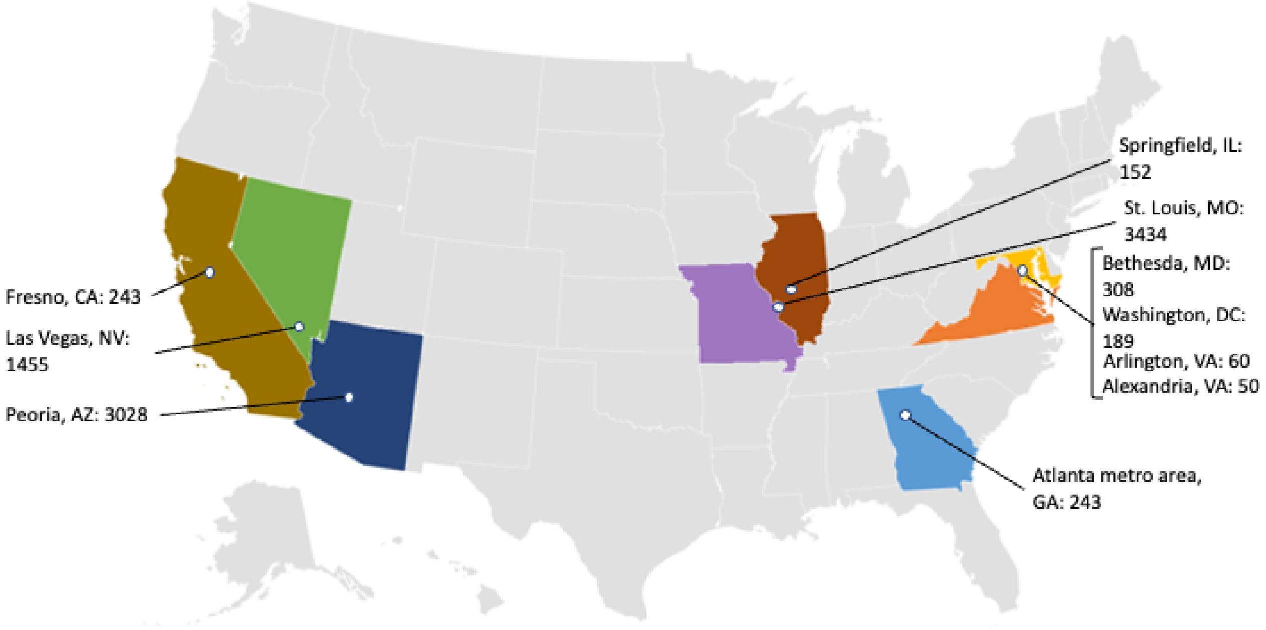


Table 1. Sample demographics (n=10,308).

Variable	Results[a], n (%)
Sex[b,c]	
Male	4539 (45)
Female	5542 (55)
Age[b,d] (years)	
0-2	110 (1.1)
3-12	542 (5.4)
13-19	802 (7.9)
20-59	7490 (74)
≥60	1174 (11.6)
Race[b]	
White	5271 (51.1)
Black	2623 (25.4)
Asian	1034 (10)
Other/unknown[e]	1380 (13.4)
Ethnicity[b,f]	
Latinx	894 (9.1)
Non-Latinx	8928 (90.9)
Data collection location[g]	
Washington, DC	858 (8.3)
Alexandria, VA	1578 (15.3)
Arlington, VA	660 (6.4)
Bethesda, MD	1740 (16.9)
Atlanta, GA	167 (1.6)
Lawrenceville, GA	3628 (35.3)
Dunwoody, GA	147 (1.4)
Las Vegas, NV	1201 (11.7)
Peoria, AZ	12 (0.1)
Springfield, IL	180 (1.7)
St Louis, MO	90 (0.9)
Fresno, CA	30 (0.3)
Outdoor masking policy	
Mask required	5937 (57.6)
Mask recommended	4207 (40.8)
Mask not required	164 (1.6)

[a] The total sample sizes for the variables may not match due to missing data.

[b] Indicated by observer report based on appearance.

[c] n=10,081.

[d] n=10,118.

[e] Observations marked as unknown have been combined with missing.

[f] n=9822.

[g] n=10,291.

Adherence to COVID-19 Preventive Behaviors

Regarding adherence to social distancing, the sample was almost evenly split between those correctly social distancing (maintaining a distance between persons of >6 feet/1.83 meters; 5185/10,113, 51.3%) and those incorrectly social distancing (standing <6 feet//1.83 meters from others; 4928/10,113, 48.7%). Approximately 65% (6611/10,136, 65.2%) of participants observed were wearing masks correctly (ie, fully covering the nose and mouth), while another 23.2% (2353/10,136) were not wearing masks at all. These data are shown in Table 2.

When examining correlates of correct masking behavior, several interesting findings emerged. These data are presented in Table 3. Persons who were correctly masked were more likely to be those observed as female (3828/5538, 69.1%; $\chi^2_1=74.8$, $P<.001$), of Asian race (793/1033, 76.8%; $\chi^2_2=129.6$, $P<.001$), and aged 60 years or older (790/1172, 67.4%; $\chi^2_4=206.7$, $P<.001$). Those correctly masked were also more likely to be adherent to social distancing guidelines (3556/5180, 68.6%; $\chi^2_1=55.7$, $P<.001$).

Correct masking behavior was more likely to be observed outside of the metropolitan Washington, DC area compared with within the metropolitan Washington, DC area (3585/5401, 66.4% vs 3018/4718, 64%; $\chi^2_1=6.4$, $P=.01$) and in locations with required or recommended outdoor masking policies (3796/5791, 65.5% and 2787/4181, 66.7%, respectively; $\chi^2_2=171.7$, $P<.001$). We further examined potential differences in adherence to masking policy by state or region, focusing on the 4 states and regions for which we had the most data: the metropolitan Washington, DC area; Georgia; Nevada; and Illinois. These data are presented in Table 4. In the time during which these data were collected, the metropolitan Washington, DC area; Illinois; and Nevada had policies that mandated masking in both indoor and outdoor locations, while Georgia did not have a mask mandate in place. However, despite the absence of a state-level mask mandate, the businesses in Georgia where those data were collected did require or recommend that customers entering those indoor spaces wear masks. When we examined masking behavior and adherence to state policies where masking was mandated, we found that, overall, there was significant adherence to masking policies ($\chi^2_3=1584.1$, $P<.001$), and the significance was largely driven by Georgia's 100% adherence to a state policy that did not require masking. We then examined masking behavior and adherence to policies where masking was either mandated (eg, state policy) or required or recommended (eg, business policies in GA). When combining these 2 conditions, we found that, although overall levels of masking adherence were close to 66%, the comparison was no longer statistically significant ($\chi^2_3=3.2$, $P=.37$).

Table 2. COVID-19 preventive behaviors among the study sample (n=10,308).

Variable	Results[a], n (%)
Social distancing behavior[b]	
Correctly social distancing	5185 (51.3)
Incorrectly social distancing	4928 (48.7)
Masking behavior[c,d]	
Mask on	6611 (65.2)
Mask partially on (eg, worn below nose)	636 (6.3)
Mask visible	536 (5.3)
No mask	2353 (23.2)

[a] The total sample sizes for the variables may not match due to missing data.

[b] n=10,113.

[c] Categories are mutually exclusive.

[d] n=10,136.

Table 3. Correlates of correct masking behavior (n=10,136).

Variable	Correct masking, n (%)	Incorrect or no masking, n (%)	χ^2 (*df*)	*P* value
Location[a]			6.4 (1)	.01
Metropolitan Washington, DC area	3018 (64)	1700 (36)		
Outside of Metropolitan Washington, DC area	3585 (66.4)	1816 (33.6)		
Sex[b,c]			74.8 (1)	<.001
Male	2762 (60.9)	1774 (39.1)		
Female	3828 (69.1)	1710 (30.9)		
Race[c,d]			129.6 (2)	<.001
White	3194 (60.6)	2074 (39.4)		
Black	1820 (69.4)	802 (30.6)		
Asian	793 (76.8)	240 (23.2)		
Ethnicity[c,e]			1.7 (1)	.20
Latinx	562 (62.9)	331 (37.1)		
Non-Latinx	5807 (65.1)	3116 (34.9)		
Age[f] **(years)**			206.7 (4)	<.001
0-2	9 (8.2)	101 (91.8)		
3-12	283 (52.8)	259 (47.8)		
13-19	528 (65.8)	274 (34.2)		
20-59	4983 (66.6)	2503 (33.4)		
≥60	790 (67.4)	382 (32.6)		
Distancing[g]			55.7 (1)	<.001
Correctly social distancing	3556 (68.6)	1624 (31.4)		
Incorrectly social distancing	3033 (61.6)	1893 (38.4)		
Mask policy			171.7 (2)	<.001
Required	3796 (65.5)	1995 (34.5)		
Recommended	2787 (66.7)	1394 (33.3)		
Not required	28 (17.1)	136 (82.9)		

[a] n=10,119.

[b] n=10,075.

[c] Unknown/other has been designated as missing.

[d] n=8923.

[e] n=9816.

[f] n=10,112.

[g] n=10,106.

Table 4. Adherence to masking policy by state or region.

State or region	Nonadherence with masking policy, n (%)	Adherence with masking policy, n (%)	χ^2 (*df*)	*P* value
Policy: masking mandated			1584.1 (3)	<.001
Metropolitan Washington, DC area	1540 (32.6)	3178 (67.4)		
Georgia	0 (0)	3942 (100)		
Nevada	377 (31.4)	823 (68.6)		
Illinois	34 (22.2)	119 (77.8)		
Total	1951 (19.5)	8062 (80.5)		
Policy: masking required or recommended			3.2 (3)	.37
Metropolitan Washington, DC area	1588 (33.7)	3130 (66.3)		
Georgia	1300 (33.2)	2619 (66.8)		
Nevada	377 (31.4)	823 (68.6)		
Illinois	44 (33.1)	106 (66.9)		
Total	3309 (33.1)	6678 (66.9)		

Discussion

Principal Findings

Our data show that, in a sample that included observations from 8 states and the District of Columbia, individuals who were observed to be female, older than 60 years, and of Asian race were more likely to be correctly adherent to public health guidelines in terms of the correct utilization of face masks when outdoors in public spaces. Moreover, we found that those who had engaged in correct masking behavior were also more likely to be adherent to social distancing guidelines. The findings from this study support the findings from the other SOMAD studies that also observed higher levels of adherence to correct mask usage among female individuals and persons of older age, as well as persons of Asian ancestry [20,23].

Comparisons With Prior Work

It is of interest to note that, while most of the data collection sites had policies in place that either required or recommended masking outdoors, we found that approximately 23% of the sample were not wearing masks at the time they were observed. This may be due to the perception—which is supported by epidemiological data—that SARS-CoV-2 transmission risk is significantly lower in outdoor spaces [24]. To that point, we did find a 66% adherence to state or local masking policies. This finding supports the results of another national-level systematic observation study that found that mask mandates result in a 3-fold increase in adherence to mask wearing [25]. Additionally, our finding showed that local policies—such as those implemented by private businesses in Georgia requiring or recommending mask wearing by their customers even though there was no mask mandate in place at the state level—may be influential in prompting preventive behaviors even in broader environmental and legislative contexts where such behaviors are not deemed important. Together, these findings support the idea that policies promoting or mandating public health protective behaviors can be useful in helping states and localities mitigate the impact of disease spread. It is likely that future observational studies of indoor mask wearing may show different levels of adherence to ongoing recommendations pertaining to social distancing and mask wearing, particularly in light of rising rates of vaccination among the general population, changing state and local policies pertaining to indoor mask usage, and waning rates of infection in the general population.

Limitations

There are several significant limitations of this research that should be noted. As mentioned previously, all data collected were observational. Because most participants were masked, it may have been more difficult than it would be under normal circumstances to ascertain persons' race, ethnicity, sex, and age. Similarly, because of the diversity of people observed and general difficulties in guessing people's ages, race/ethnicity, and gender, there is likely some amount of error in our data pertaining to our participants' demographic characteristics. This lack of precision is a limitation of the data collection procedures. Also, because data collectors were frequently observing individuals as they moved through spaces (eg, crossing a street or walking on a sidewalk), it may be possible that observers were unable to clearly determine individuals' demographic characteristics or mask usage. This may have contributed to misclassification of individuals into demographic and behavioral adherence categories.

Conclusions

Despite these limitations, we were able to collect a substantial amount of data from a variety of locations across the country, thereby providing a representative picture of COVID-19 preventive behaviors in a diverse population. These data provide some significant takeaways that can inform public health practice as the world moves into the third year of addressing the COVID-19 pandemic. The first of these points is that, despite significant gains in vaccination uptake, there is still a need for adherence to COVID-19 preventive behaviors, particularly in places (eg, rural communities) and among populations (eg, communities of color) where disparities in vaccine access and uptake exist and where the presence of significant health comorbidities increase risk of severe disease if persons do become infected. Similarly, with the continued emergence of new SARS-CoV-2 variants, maintaining high levels of adherence to protective behaviors will be critical to preventing further surges in the numbers of COVID-19 cases, including both new and breakthrough infections. As the pandemic becomes endemic and more normalized, maintaining adherence to preventive behaviors is challenging, as the presence of effective vaccines may cause the public to believe that the pandemic is over or that preventive behaviors are no longer relevant. Future research should focus on developing more innovative ways to promote continued adherence to masking and social distancing guidance where necessary and appropriate, with particular attention to how to mitigate "pandemic fatigue" [26].

Our findings also highlight the importance of clear and effective public health communication regarding the need for adherence to preventive behaviors other than or in addition to vaccination. This is particularly necessary considering variations in state and local adherence with federal guidance, as well as the abundance of misinformation about COVID-19 that continues to circulate in the news media and on social media. A recent Kaiser Family Foundation report found that, in a sample of 1519 adults 18 years of age and older sampled through random digit dialing, approximately 78% indicated that they believed or were unsure about at least one false statement about the COVID-19 pandemic or about vaccination that they had seen in the news media [27]. Future research should focus not only on improving the health messages that are disseminated to the public about COVID-19 and disease prevention but also on improving health literacy among the general population so that individuals are better able to determine the veracity of information they hear through the news media so that they can more easily spot "fake news" about the pandemic. Moreover, given the continued evolution of highly transmissible and contagious SARS-CoV-2 variants, it is critical that public health campaigns effectively communicate the continued benefit of protective behaviors such as vaccination, masking, and distancing as strategies to reduce the morbidity and mortality associated with the COVID-19 pandemic.

Acknowledgments

The study team would like to thank Dr. Deborah Cohen and the National Systematic Observation of Mask Adherence and Distancing (SOMAD) Research Team for giving us the opportunity to be a part of this study. We would also like to acknowledge the contributions to this research made by the SOMAD@GW team of "citizen scientist" data collectors: Rhiana Brennan, Jackie Carlton, Gabby Carmichael, Giulia Folli, Tsion Gashaw, Agnes Halm-Lutterodt, Misaal Irfan, Catherine Jones, Nicole Joseph, Hannah Latif, Elizabeth Mendoza-Tirado, Zena Osman, and Lizzie Zinn.

Data Availability

The data sets generated or analyzed during the current study are available from the corresponding author (MSR) on reasonable request. Data requests for the national SOMAD study should be made to Dr. Deborah A Cohen (Deborah.A.Cohen@kp.org).

Conflicts of Interest

None declared.

References

1. COVID Data Tracker. Centers for Disease Control and Prevention. 2022 Jun 02. URL: https://covid.cdc.gov/covid-data-tracker [accessed 2022-06-02]
2. Mackey K, Ayers CK, Kondo KK, Saha S, Advani SM, Young S, et al. Racial and ethnic disparities in COVID-19-related infections, hospitalizations, and deaths: a systematic review. Ann Intern Med 2021 Mar;174(3):362-373 [FREE Full text] [doi: 10.7326/M20-6306] [Medline: 33253040]
3. Carethers J. Insights into disparities observed with COVID-19. J Intern Med 2021 Apr 06;289(4):463-473 [FREE Full text] [doi: 10.1111/joim.13199] [Medline: 33164230]
4. COVID-19 Vaccinations in the United States. Centers for Disease Control and Prevention. 2021 Sep 15. URL: https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-rate-total [accessed 2021-09-15]
5. Lerner AM, Folkers GK, Fauci AS. Preventing the spread of SARS-CoV-2 with masks and other "low-tech" interventions. JAMA 2020 Nov 17;324(19):1935-1936. [doi: 10.1001/jama.2020.21946] [Medline: 33104157]
6. Leung CC, Cheng KK, Lam TH, Migliori GB. Mask wearing to complement social distancing and save lives during COVID-19. Int J Tuberc Lung Dis 2020 Jun 01;24(6):556-558. [doi: 10.5588/ijtld.20.0244] [Medline: 32553007]
7. Science Brief: Community Use of Masks to Control the Spread of SARS-CoV-2. Centers for Disease Control and Prevention. 2020 Oct. URL: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/masking-science-sars-cov2.html [accessed 2023-03-24]
8. Li Y, Liang M, Gao L, Ayaz Ahmed M, Uy JP, Cheng C, et al. Face masks to prevent transmission of COVID-19: A systematic review and meta-analysis. Am J Infect Control 2021 Jul;49(7):900-906 [FREE Full text] [doi: 10.1016/j.ajic.2020.12.007] [Medline: 33347937]

9. COVID-19: How to Protect Yourself and Others. Centers for Disease Control and Prevention. 2023 Jan 26. URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2023-03-24]
10. Rader B, White LF, Burns MR, Chen J, Brilliant J, Cohen J, et al. Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study. The Lancet Digital Health 2021 Mar;3(3):e148-e157. [doi: 10.1016/s2589-7500(20)30293-4]
11. Sugimura M, Chimed-Ochir O, Yumiya Y, Ohge H, Shime N, Sakaguchi T, et al. The association between wearing a mask and COVID-19. Int J Environ Res Public Health 2021 Aug 30;18(17):9131 [FREE Full text] [doi: 10.3390/ijerph18179131] [Medline: 34501719]
12. Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, van der Westhuizen H, et al. An evidence review of face masks against COVID-19. Proc Natl Acad Sci U S A 2021 Jan 26;118(4):e2014564118 [FREE Full text] [doi: 10.1073/pnas.2014564118] [Medline: 33431650]
13. Variant Proportions: Monitoring Variant Proportions. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker/#variant-proportions [accessed 2023-02-25]
14. Kahane LH. Politicizing the mask: political, economic and demographic factors affecting mask wearing behavior in the USA. East Econ J 2021 Jan 05;47(2):163-183 [FREE Full text] [doi: 10.1057/s41302-020-00186-0] [Medline: 33424048]
15. Fischer CB, Adrien N, Silguero JJ, Hopper JJ, Chowdhury AI, Werler MM. Mask adherence and rate of COVID-19 across the United States. PLoS One 2021 Apr 14;16(4):e0249891 [FREE Full text] [doi: 10.1371/journal.pone.0249891] [Medline: 33852626]
16. COVID-19 Surveillance. District of Columbia Department of Health. URL: https://coronavirus.dc.gov/data [accessed 2021-12-01]
17. Tracking COVID-19 Disparities. Virginia Department of Health. URL: https://www.vdh.virginia.gov/coronavirus/2021/11/15/tracking-covid-19-disparities/ [accessed 2021-12-01]
18. COVID-19 Statistics in Maryland. Maryland Department of Health. URL: https://coronavirus.maryland.gov [accessed 2021-12-01]
19. Cohen, DA. Systematic Observation of Mask Adherence and Distancing (SOMAD). Kaiser Permanente Southern California Department of Research & Evaluation. URL: https://www.kp-scalresearch.org/somad/ [accessed 2023-03-24]
20. Cohen DA, Talarowski M, Awomolo O, Han B, Williamson S, McKenzie TL. Systematic observation of mask adherence and distancing (SOMAD): Findings from Philadelphia. Prev Med Rep 2021 Sep;23:101449 [FREE Full text] [doi: 10.1016/j.pmedr.2021.101449] [Medline: 34155476]
21. Cohen D, Talarowski M, Awomolo O, McKenzie T, Han B, Williamson S, et al. National Study of Mask Adherence and Distancing: Manual of Procedures for Systematic Observation of Masking Adherence and Distancing (SOMAD). Kaiser Permanente Southern California Department of Research & Evaluation. 2020 Aug 27. URL: https://www.kp-scalresearch.org/documents/SOMADprotocol8-27-2020.pdf [accessed 2023-03-24]
22. Coronavirus in the U.S.: Latest Map and Case Count. The New York Times. URL: https://www.nytimes.com/interactive/2021/us/covid-cases.html [accessed 2023-02-23]
23. Pecoraro J, Bakour C, Oberne A, Mehmood A. Masking behaviors in the absence of local mandate-an observational study from Hillsborough County, Florida. Int J Environ Res Public Health 2022 Nov 30;19(23):15982 [FREE Full text] [doi: 10.3390/ijerph192315982] [Medline: 36498054]
24. Bulfone TC, Malekinejad M, Rutherford GW, Razani N. Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review. J Infect Dis 2021 Feb 24;223(4):550-561 [FREE Full text] [doi: 10.1093/infdis/jiaa742] [Medline: 33249484]
25. Puttock EJ, Marquez J, Young DR, Shirley AM, Han B, McKenzie TL, et al. Association of masking policies with mask adherence and distancing during the SARS-COV-2 pandemic. Am J Infect Control 2022 Sep;50(9):969-974 [FREE Full text] [doi: 10.1016/j.ajic.2022.04.010] [Medline: 35545151]
26. Parrish C. How to Deal with Coronavirus Burnout and Pandemic Fatigue. Johns Hopkins Medicine. 2020 Aug 11. URL: https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/how-to-deal-with-coronavirus-burnout-and-pandemic-fatigue [accessed 2021-12-01]
27. Hamel L, Lopes L, Kirzinger A, Sparks G, Stokes M, Brodie M. KFF COVID-19 Vaccine Monitor: Media and Misinformation. Kaiser Family Foundation. URL: https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-media-and-misinformation/ [accessed 2023-12-23]

Abbreviations

BIPOC: Black and Indigenous People of Color
R0: basic reproduction number
SOMAD: Systematic Observation of Mask Adherence and Distancing

Edited by A Mavragani, T Sanchez; submitted 07.06.22; peer-reviewed by J Wang, S Pesälä, D Hinthorn, CR Telles; comments to author 08.11.22; revised version received 28.02.23; accepted 07.03.23; published 25.04.23.

Please cite as:
Ruiz MS, McMahon MV, Latif H, Vyas A
Maintaining Adherence to COVID-19 Preventive Practices and Policies Pertaining to Masking and Distancing in the District of Columbia and Other US States: Systematic Observational Study
JMIR Public Health Surveill 2023;9:e40138
URL: https://publichealth.jmir.org/2023/1/e40138
doi:10.2196/40138
PMID:36888910

Original Paper

The Effects of COVID-19 Pandemic Policy on Social Needs Across the State of Kansas and Western Missouri: Paired Survey Response Testing

Dinesh Pal Mudaranthakam[1], MBA, PhD; Sam Pepper[1], MS; Tanner Fortney[2], MS; Alexander Alsup[1], MS; Jennifer Woodward[3], MD, MPH; Kevin Sykes[4], MPH, PhD; Elizabeth Calhoun[2], MEd, PhD

[1]Department of Biostatistics & Data Science, University of Kansas Medical Center, Kansas City, KS, United States
[2]Department of Population Health, University of Kansas Medical Center, Kansas City, KS, United States
[3]Department of Family Medicine and Community Health, University of Kansas Medical Center, Kansas City, KS, United States
[4]Department of Otolaryngology-Head and Neck Surgery, University of Kansas Medical Center, Kansas City, KS, United States

Corresponding Author:
Dinesh Pal Mudaranthakam, MBA, PhD
Department of Biostatistics & Data Science
University of Kansas Medical Center
3901 Rainbow Blvd
Kansas City, KS, 66160
United States
Phone: 1 9139456922
Email: dmudaranthakam@kumc.edu

Abstract

Background: Studying patients’ social needs is critical to the understanding of health conditions and disparities, and to inform strategies for improving health outcomes. Studies have shown that people of color, low-income families, and those with lower educational attainment experience greater hardships related to social needs. The COVID-19 pandemic represents an event that severely impacted people’s social needs. This pandemic was declared by the World Health Organization on March 11, 2020, and contributed to food and housing insecurity, while highlighting weaknesses in the health care system surrounding access to care. To combat these issues, legislators implemented unique policies and procedures to help alleviate worsening social needs throughout the pandemic, which had not previously been exerted to this degree. We believe that improvements related to COVID-19 legislature and policy have positively impacted people’s social needs in Kansas and Missouri, United States. In particular, Wyandotte County is of interest as it suffers greatly from issues related to social needs that many of these COVID-19–related policies aimed to improve.

Objective: The research objective of this study was to evaluate the change in social needs before and after the COVID-19 pandemic declaration based on responses to a survey from The University of Kansas Health System (TUKHS). We further aimed to compare the social needs of respondents from Wyandotte County from those of respondents in other counties in the Kansas City metropolitan area.

Methods: Social needs survey data from 2016 to 2022 were collected from a 12-question patient-administered survey distributed by TUKHS during a patient visit. This provided a longitudinal data set with 248,582 observations, which was narrowed down into a paired-response data set for 50,441 individuals who had provided at least one response before and after March 11, 2020. These data were then bucketed by county into Cass (Missouri), Clay (Missouri), Jackson (Missouri), Johnson (Kansas), Leavenworth (Kansas), Platte (Missouri), Wyandotte (Kansas), and Other counties, creating groupings with at least 1000 responses in each category. A pre-post composite score was calculated for each individual by adding their coded responses (yes=1, no=0) across the 12 questions. The Stuart-Maxwell marginal homogeneity test was used to compare the pre and post composite scores across all counties. Additionally, McNemar tests were performed to compare responses before and after March 11, 2020, for each of the 12 questions across all counties. Finally, McNemar tests were performed for questions 1, 7, 8, 9, and 10 for each of the bucketed counties. Significance was assessed at $P<.05$ for all tests.

Results: The Stuart-Maxwell test for marginal homogeneity was significant ($P<.001$), indicating that respondents were overall less likely to identify an unmet social need after the COVID-19 pandemic. McNemar tests for individual questions indicated that after the COVID-19 pandemic, respondents across all counties were less likely to identify unmet social needs related to food availability (odds ratio [OR]=0.4073, *P*<.001), home utilities (OR=0.4538, *P*<.001), housing (OR=0.7143, *P*<.001), safety among cohabitants (OR=0.6148, *P*<.001), safety in their residential location (OR=0.6172, *P*<.001), child care (OR=0.7410, *P*<0.01), health care access (OR=0.3895, *P*<.001), medication adherence (OR=0.5449, *P*<.001), health care adherence (OR=0.6378, *P*<.001), and health care literacy (0.8729, *P*=.02), and were also less likely to request help with these unmet needs (OR=0.7368, *P*<.001) compared with prepandemic responses. Responses from individual counties were consistent with the overall results for the most part. Notably, no individual county demonstrated a significant reduction in social needs relating to a lack of companionship.

Conclusions: Post-COVID-19 responses showed improvement across almost all social needs–related questions, indicating that the federal policy response possibly had a positive impact on social needs across the populations of Kansas and western Missouri. Some counties were impacted more than others and positive outcomes were not limited to urban counties. The availability of resources, safety net services, access to health care, and educational opportunities could play a role in this change. Future research should focus on improving survey response rates from rural counties to increase their sample size, and to evaluate other explanatory variables such as food pantry access, educational status, employment opportunities, and access to community resources. Government policies should be an area of focused research as they may affect the social needs and health of the individuals considered in this analysis.

(JMIR Public Health Surveill 2023;9:e41369) doi:10.2196/41369

KEYWORDS

social determinants of health; COVID-19; food assistance program; public health; quality of life; well-being; health disparity; health inequity; health policy; Kansas; social work; socioeconomic

Introduction

When people think of health care, facilities such as hospitals, primary care clinics, and other direct health care services typically come to mind. The direct services provided by health care professionals at these institutions are generally considered to constitute the health care system in its entirety. However, this is only one piece of the health care system. An individual's social needs outside of health care facilities also directly impact their health, and are important in understanding the causes of many health conditions and health disparities and in determining the actions society should take to improve overall health [1]. For example, obesity is influenced by the consumption of foods with poor nutritive value, a decrease in physical activity, and many other factors, which in turn lead to negative long-term health impacts [2,3].

One area of scientific research focuses on the neighborhood effect, which involves the interaction of social and environmental factors and how these interactions impact housing, crime, and violence [2]. The Robert Wood Johnson Foundation formed the Commission to Build a Healthier America, a nonpartisan group focused on improving public health outcomes for vulnerable populations, and released a report in 2014 comparing average life expectancy by county demonstrating wide variations in outcomes based on race/ethnicity, gender, income, and other factors [2,4]. A geospatial analysis showed that children with asthma residing in certain hotspot counties considered to be underserved have higher social vulnerabilities and are associated with longer hospital stay durations [5]. Another study found that social risk factors were associated with persistent functional disability caused by juvenile idiopathic arthritis [6].

Among the 105 counties in Kansas, United States, Wyandotte County is ranked 102nd according to county health rankings [7]. Wyandotte County residents are very diverse, with a distribution of 24.4% African American and 26.9% Latino [8]. Despite having one of the largest populations in the state (165,265 people), Wyandotte County is ranked among the lowest 25% of counties in Kansas regarding health outcomes and health factors [7]. Among adults, Wyandotte has a smoking rate of 20%, an obesity rate of 40%, a physical inactivity rate of 39%, and an uninsured rate of 17% compared to state averages of 17%, 36%, 26%, and 11%, respectively. Such vulnerable populations have been disproportionately impacted by the COVID-19 pandemic, creating the added burdens of stress and family challenges, both of which have impacted mental well-being [9].

The COVID-19 pandemic, declared by the World Health Organization on March 11, 2020, created challenges not only for the public health system but also for children and families suffering from housing and food insecurity [10]. COVID-19 had a different impact on each population group. For example, children experienced an increase in food insecurity due to the rising costs of food and reduced access to in-person school lunches [10]. In addition, people of color have experienced higher rates of COVID-19 infections, fatalities, and lower vaccination rates [11]. A lack of access to health care has been shown to cause food insecurity as health complications lead to or exacerbate the issue [12]. Lower-income households purchased shelf-stable processed foods and were more reliant on these types of foods than fresh food [13].

Income and education are other significant contributors to food insecurity, with lower-income families and those with lower educational attainment experiencing greater levels of insecurity [14]. Both income and education were also found to be major factors contributing to food security in Arkansas, which is one of the most food-insecure states in the United States [14]. Moreover, the COVID-19 pandemic exacerbated food insecurity for groups such as unemployed students [15,16].

During the first surge of the pandemic, an interview conducted with 20 nurses indicated that nearly every patient they were serving was a person of color, especially individuals of Hispanic ethnicity, who constituted a majority of the COVID-19 patients [17]. Many patients were suffering from untreated comorbidities, especially cases of diabetes and hypertension, that made patients more susceptible to the more serious outcomes associated with COVID-19 [17]. Some nurses also described cultural differences; for example, many Hispanic patients would not seek treatment when they began experiencing symptoms, and only sought medical attention when they became seriously ill at which time they were admitted directly to the intensive care unit (ICU) [17].

Differences in geographic location, income, education, and employment could explain a large amount of the disparity in vaccination coverage between Hispanic and white populations [18,19]. Models developed by Williams et al [20] indicated that health insurance in particular accounted for a 0.9 percentage-point disparity, and if the Hispanic population had a 13 percentage-point higher health insurance coverage to match that of white respondents, their predicted vaccination coverage would be an additional 1% higher. Their findings suggested that lack of insurance, unfamiliarity with the health care system, and not knowing that vaccines are provided free of charge may explain part of the vaccination coverage difference between these groups [20]. Compared with individuals identifying as white, minority groups such as Black, Asian, and Native American populations were more likely to have higher rates of ICU admissions, hospitalizations, and deaths when adjusted for age [21]. Research studies differ on the exact level of disparity; however, some minority groups experienced hospitalization rates that were nearly 9-times higher and infection rates that were 27-times higher than those of white individuals [21-23]. Based on data collected during the pandemic, life expectancy projections indicate that individuals of all races will experience a decreased life expectancy; however, Black and Latino individuals are expected to lose 2 and 3 years of life expectancy, respectively [24].

Social needs have a serious impact on a patient's access to care and vulnerability to the COVID-19 virus; this causes disparities in health outcomes between counties within the United States. Access to healthy food, health care services, and willingness to accept COVID-19 vaccines play a significant role in population health outcomes. To help minimize the impact of COVID-19, the United States implemented several policies. Moratoriums on evictions were implemented nationwide [25,26], meal plans for children and vulnerable populations were expanded [27], and stimulus funds were made available to nearly all citizens. These policies may have prevented many families from suffering dire consequences associated with business closures and COVID-19–related medical expenses. These programs aimed to help alleviate the elevated social needs issues across the nation.

Despite the vast amount of research aimed at examining the effects of COVID-19 on human health in the United States, very little research has been conducted to examine the outcomes associated with funds provided to households to combat food insecurity, the moratorium implemented to prevent evictions, and other federal policies on the population. This is an important gap in the literature that should be addressed as socials impacts could play an important role in directing federal policy during the next pandemic.

With this in mind, the research question we set out to address is whether there is a significant difference in social needs before and after the COVID-19 pandemic declaration. Additionally, we wanted to explore how the social needs of individuals compare across counties. A notable county of interest is Wyandotte County, Kansas. Compared to other counties in the Kansas City metropolitan area, Wyandotte has a high poverty rate of 16.9%, a large minority population, and a high crime rate [28,29]. Therefore, we expected Wyandotte County to be impacted the most by COVID-19 and the resulting polices. If we can better understand how COVID-19 policies and programs have impacted individuals' social needs, we will gain a better understanding of how to address social needs from both legislative and individual perspectives.

Our study thus aimed to illustrate the differences in social needs between populations living in the different counties throughout the state of Kansas and western Missouri. These counties represent a combination of rural and urban counties throughout the two states; however, our research focuses primarily on the regions of Johnson (Kansas), Jackson (Missouri), and Wyandotte (Kansas), since these jurisdictions are urban/suburban, are located near one another geographically, and are part of the same metropolitan area. The socioeconomic factors are also quite different among these jurisdictions, which may be a cause of disparities in patient responses and experiences.

Methods

Study Design

We used a survey developed at The University of Kansas Health System (TUKHS), which is a modified version of the validated social needs survey built from the Health Leads Toolkit [30]. The updates were specific to the separation of personal violence from community violence as those resources may differ [31]. The survey responses are used by the 12 social workers at TUKHS who aid patients who answer "yes" to any of the social needs questions. The data were retrieved from TUKHS informatics data warehouse HERON (aka. i2b2). The data set consists of patient demographic information along with the social needs responses. R studio with R version 4.1.2 was used to perform the descriptive and statistical analyses [32].

Ethical Considerations

The study protocol was approved by the Institutional Review Board at the University of Kansas Medical Center (STUDY00148041).

Data Sources

The primary data element utilized for this study was the patient-administered social needs survey, which is completed by patients during their primary care visit at TUKHS. The survey included 12 questions related to social needs, which are listed in Textbox 1.

Textbox 1. Items of the social needs survey.

Question 1: In the last 12 months, did you ever eat less than you should because there wasn't enough? (Food)

Question 2: In the last 12 months, has your utility company shut off your service for not paying your bill? (Utility)

Question 3: Are you worried that in the next 2 months, you may not have stable housing? (Housing)

Question 4: Are you afraid you might be hurt in your home by someone you know? (Safe home)

Question 5: Are you afraid you might be hurt in your apartment building or neighborhood? (Safe area)

Question 6: Do problems getting child care make it difficult for you to work or study? (Child care)

Question 7: In the last 12 months, have you needed to see a doctor, but could not because of cost? (Health care access)

Question 8: In the last 12 months, did you skip medications to save money? (Medication)

Question 9: In the last 12 months, have you ever had to go without health care? (Skip health care)

Question 10: Do you have problems understanding what is told to you about your medical conditions? (Health literacy)

Question 11: Do you often feel that you lack companionship? (Support)

Question 12: If you answered YES to any questions above, would you like to discuss help? (Need help in general)

Data (responses to the survey) were abstracted from the electronic medical record along with the zip code, city, and county of the patient on January 4, 2022. After deidentification of the patient information, data from patients not in Kansas or western Missouri were removed. This provided a longitudinal data set with 248,582 observations (820 from 2017, 47,877 from 2018, 58,376 from 2019, 74,618 from 2020, and 66,891 from 2021) representing the catchment area of TUKHS. In addition to the survey questions, 2020 county population estimates were obtained from the May 2021 Annual Resident Population Estimates of the US Census Bureau; population division area deprivation index by zip code was obtained from the 2019 release for both Kansas and Missouri, and 2013 Rural-Urban Continuum Code (RUCC) classifications were obtained from the update on December 10, 2020. RUCC codes 1, 2, and 3 were classified as urban, and codes 4, 5, 6, 7, 8, 9, and 10 were classified as rural [33,34].

Data Analysis

For the pre-post COVID-19 analysis, the longitudinal data set was subset down to a paired-response data set for 50,441 individuals who had at least one response before March 11, 2020, and at least one response after. If an individual had multiple responses before the pandemic declaration, the response closest to March 11, 2020, was kept, and if an individual had multiple responses after, the most recent response was kept (ie, that closest to January 4, 2022). A contingency table and bar chart were constructed to explore the data responses. An overall pre-post composite score was calculated by adding the responses coded as yes=1 and no=0 across the 12 questions. The 3M Social Determinants of Health (SDoH) score in combination with individual metrics has been used to model health care utilization within a population; SDoH scores were highly impactful in these models when compared to the other variables [35]. Recent studies have proposed a similar approach, referring to these composite scores as polysocial risk scores or health indices [36,37]. A comprehensive risk score helps with a holistic approach to conduct statistical evaluation, addressing outliers or patients with enormous social needs. For example, patients with multiple social conditions would have a higher composite score compared with that of patients under relatively minimal social conditions such as a lack of child support. The composite score allows researchers to evaluate the impact of risk scores on patients' health outcomes [37].

A Stuart-Maxwell marginal homogeneity test was run to compare the pre and post composite scores (relative to the pandemic announcement on March 11, 2020) across all counties. McNemar tests were then run to evaluate whether there is a statistically significant difference for each question when compared to paired pre-post responses across all counties. County-level McNemar tests were performed for questions 1, 7, 8, 9, and 10 for the county groupings of Cass (Missouri), Clay (Missouri), Jackson (Missouri), Johnson (Kansas), Leavenworth (Kansas), Platte (Missouri), Wyandotte (Kansas), and Other.

Results

The most recent responses were divided into a subset and the data set was visualized to determine the number of respondents per county. Figure 1 shows that the majority of patients in the data set reside in the immediate surrounding counties of TUKHS. The number of patients by county was then divided by their corresponding 2020 census county population estimate to determine the percentage of the county captured. Figure 2 shows that when adjusting for county populations, most of the respondents in the data set were still based in the immediate surrounding counties of TUKHS. Because of the low percentage of counties not immediately near TUKHS, we decided to classify the respondents into the counties of Cass, Clay, Jackson, Johnson, Leavenworth, Platte, Wyandotte, and Other, which created groupings with at least 1000 responses in each.

The yes/no counts by group for the most recent responses are provided in Table 1, showing that Wyandotte County consistently had a higher proportion of "yes" responses when compared to the other counties. Table 1 also shows that the proportion of "yes" response was below 6% for the most recent responses, demonstrating the rarity of this response. The county buckets of Cass, Clay, Jackson, Johnson, Leavenworth, and Platte all had a "yes" response rate below 4%.

The Stuart-Maxwell test showed a significant (*P*<.001) difference in the composite scores between the pre and post responses. Unadjusted *P* values of the McNemar test for each of the questions across all counties showed significance at an α level of .05 for all the questions except question 11 (Support). However, it should be noted that statistical significance does not necessarily mean that a meaningful difference exists given the rarity in the event of a respondent answering "yes." This scenario is particularly apparent for questions 2 (Utility), 3 (Housing), 4 (Safe home), 5 (Safe area), and 6 (Child care). Therefore, questions 1 (Food), 7 (Health care access), 8 (Medication), 9 (Skip health care), and 12 (Need help in general) are the notable questions to consider further (Table 2).

Notable results from the county-level McNemar tests were related to questions 1, 7, 8, and 9 from Wyandotte County, which had a relatively high percentage of respondents who answered "yes" in the pre-COVID-19 category and subsequently answered "no" in the post-COVID-19 category. The unadjusted *P* values from the McNemar test were significant for these questions. When looking at the McNemar test results by county, significant reductions in respondents answering "yes" for question 1 were found in Clay, Jackson, Johnson, Wyandotte, and Other counties. Significant reductions in respondents answering "yes" for question 7 were found in Cass, Clay, Jackson, Johnson, Platte, Wyandotte, and Other counties. Significant reductions in respondents answering "yes" for question 8 were found in Cass, Clay, Jackson, Johnson, Platte, Wyandotte, and Other counties. Significant reductions in respondents answering "yes" for question 9 were found in Jackson, Johnson, and Wyandotte. Finally, no county buckets demonstrated a significant change in the response to question 10 (Table 2).

Figure 1. Heat map of the raw response count by counties under The University of Kansas Cancer Center catchment area; yellow indicates lower counts and red indicates higher counts. ©Mapbox ©OpenStreetMap.

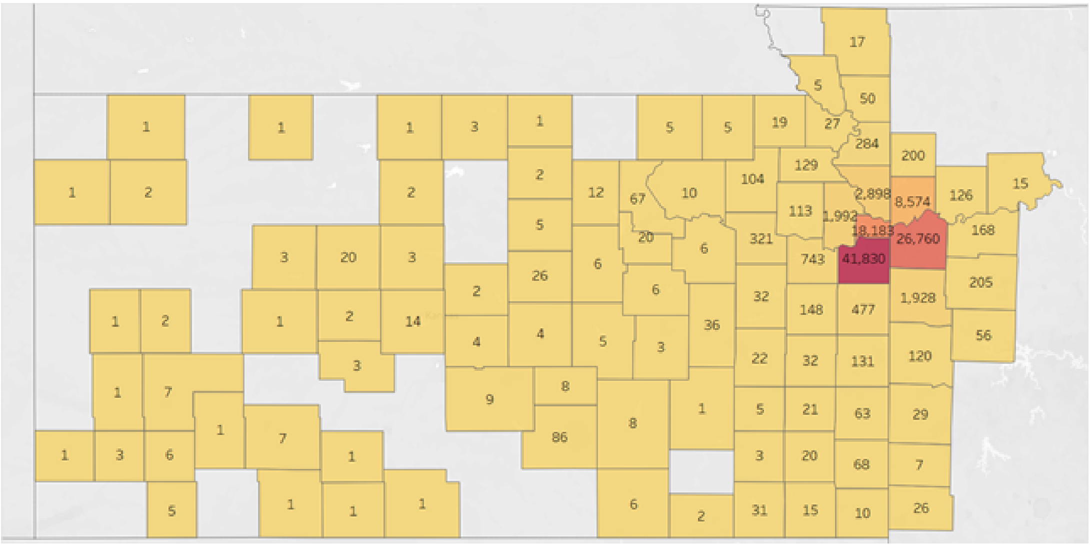


Figure 2. Heat map of the percentage of the total county population to responses by county in The University of Kansas Cancer Center catchment area; yellow indicates lower counts and red indicates higher counts. ©Mapbox ©OpenStreetMap.

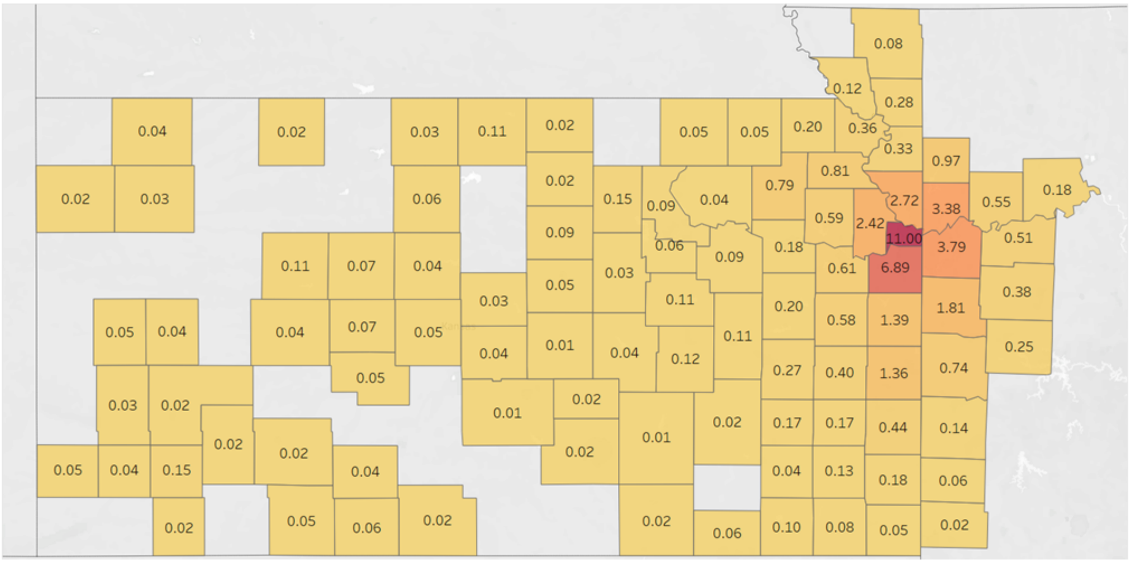


Table 1. Response distribution by counties with a minimum of 1000 patient responses.

Question[a]	Cass, MO[b]	Clay, MO	Jackson, MO	Johnson, KS[c]	Leavenworth, KS	Platte, MO	Wyandotte, KS	Other
1: Food, n (%)								
No	1837 (98.29)	8235 (98.32)	25,655 (98.01)	36,331 (98.74)	1717 (97.95)	27,535 (98.78)	15,088 (95.63)	3897 (96.82)
Yes	32 (1.71)	<25 (1.68)	520 (1.99)	463 (1.26)	36 (2.05)	34 (1.22)	689 (4.37)	128 (3.18)
2: Utility, n (%)								
No	1853 (99.09)	8322 (99.31)	25,905 (98.95)	36,621 (99.42)	1741 (99.20)	2773 (99.36)	15,324 (97.01)	3961 (98.48)
Yes	<25 (0.91)	58 (0.69)	275 (1.05)	215 (0.58)	<25 (0.80)	<25 (0.64)	473 (2.99)	61 (1.52)
3: Housing, n (%)								
No	1859 (99.41)	8270 (98.62)	25,820 (98.62)	36,523 (99.15)	1723 (98.23)	2763 (99.03)	15,363 (97.26)	3945 (98.09)
Yes	<25 (0.59)	110 (1.31)	360 (1.38)	313 (0.85)	31 (1.77)	27 (0.97)	432 (2.74)	77 (1.91)
4: Safe home, n (%)								
No	1863 (99.68)	8367 (99.86)	26,115 (99.76)	36,755 (99.81)	1751 (99.77)	2784 (99.78)	15,739 (99.65)	4003 (99.53)
Yes	<25 (0.32)	<25 (0.14)	62 (0.24)	70 (0.19)	<25 (0.23)	<25 (0.22)	55 (0.35)	<25 (0.47)
5: Safe area, n (%)								
No	1867 (99.84)	8347 (99.59)	25,963 (99.20)	36,720 (99.72)	1747 (99.54)	2780 (99.64)	15,593 (98.73)	3988 (99.15)
Yes	<25 (0.16)	34 (0.41)	210 (0.80)	102 (0.28)	<25 (0.46)	<25 (0.36)	201 (1.27)	34 (0.85)
6: Child care, n (%)								
No	1857 (99.46)	8331 (99.49)	26,044 (99.54)	36,649 (99.57)	1740 (99.20)	2772 (99.50)	15,603 (98.93)	3989 (99.35)
Yes	<25 (0.54)	43 (0.51)	120 (0.46)	160 (0.43)	<25 (0.80)	<25 (0.50)	168 (1.07)	26 (0.65)
7: Health care access, n (%)								
No	1819	8069	25,160	36,045	1703	2722	14,945	3865
Yes	49 (2.62)	306 (3.65)	1016 (3.88)	792 (2.15)	51 (2.91)	69 (2.47)	842 (5.33)	153 (3.81)
8: Medication, n (%)								
No	1830 (97.38)	8134 (97.08)	25,410 (97.09)	36,179 (98.22)	1707 (97.21)	2733 (97.92)	15,102 (95.64)	3868 (96.17)
Yes	40 (2.14)	245 (2.92)	761 (2.91)	655 (1.78)	49 (2.79)	58 (2.08)	689 (4.36)	154 (3.83)
9: Skip health care, n (%)								
No	1849 (98.93)	8275 (98.76)	25,750 (98.39)	36,466 (99.01)	1726 (98.35)	2765 (99.10)	15,173 (96.10)	3899 (97.01)
Yes	<25 (1.07)	104 (1.24)	421 (1.61)	364 (0.99)	29 (1.65)	25 (0.90)	616 (3.90)	120 (2.99)
10: Health literacy, n (%)								
No	1837 (98.24)	8227 (98.19)	25,692 (98.17)	36,198 (98.30)	1698 (96.75)	2752 (98.71)	15,063 (95.44)	3876 (96.39)
Yes	33 (1.76)	152 (1.81)	479 (1.83)	627 (1.70)	57 (3.25)	36 (1.29)	719 (4.56)	145 (3.61)
11: Support, n (%)								
No	1805 (98.31)	7992 (98.19)	24,685 (96.19)	35,376 (97.00)	1674 (96.60)	2676 (97.24)	14,631 (94.66)	3718 (94.44)
Yes	31 (1.69)	291 (3.51)	979 (3.81)	1094 (3.00)	59 (3.40)	76 (2.76)	826 (5.34)	219 (5.56)
12: Need help in general, n (%)								
No	1807 (98.74)	8155 (98.19)	25,068 (97.64)	35,538 (98.52)	1690 (97.52)	2714 (98.44)	14,791 (95.98)	3790 (96.91)
Yes	23 (1.26)	150 (1.81)	607 (2.36)	535 (1.48)	43 (2.48)	43 (1.56)	620 (4.02)	121 (3.09)

[a] For detailed questions, see Textbox 1.

[b] MO: Missouri.

[c] KS: Kansas.

Table 2. McNemar test results for responses before and after the COVID-19 pandemic declaration overall and by county.

Question[a]	OR[b] (95% CI)	*P* value
Q1: Food		
Overall	0.4073 (0.3577-0.4629)	<.001
Cass, MO[c]	0.3636 (0.0844-1.2272)	.12
Clay, MO	0.2381 (0.1259-0.4230)	<.001
Jackson, MO	0.4038 (0.3106-0.5210)	<.001
Johnson, KS[d]	0.4365 (0.3348-0.5652)	<.001
Leavenworth, KS	1.6000 (0.4615-6.2161)	.58
Other	0.4375 (0.2157-0.8435)	.02
Platte, MO	0.4500 (0.1804-1.0339)	.06
Wyandotte, KS	0.4014 (0.3201-0.5006)	<.001
Q2: Utility (Overall)	0.4538 (0.3853-0.5331)	<.001
Q3: Housing (Overall)	0.7143 (0.6183-0.8245)	<.001
Q4: Safe home (Overall)	0.6148 (0.4548-0.8263)	<.001
Q5: Safe area (Overall)	0.6172 (0.5029-0.7555)	<.001
Q6: Child care (Overall)	0.7410 (0.5820-0.9412)	.01
Q7: Health care access		
Overall	0.3895 (0.3498-0.4331)	<.001
Cass, MO	0.1667 (0.0420-0.4851)	<.001
Clay, MO	0.4359 (0.3074-0.6106)	<.001
Jackson, MO	0.4091 (0.3346-0.4980)	<.001
Johnson, KS	0.4441 (0.3632-0.5409)	<.001
Leavenworth, KS	0.5333 (0.1958-1.3400)	.21
Other	0.3148 (0.1711-0.5515)	<.001
Platte, MO	0.3684 (0.1844-0.6956)	.001
Wyandotte, KS	0.3158 (0.2484-0.3985)	<.001
Q8: Medication		
Overall	0.5449 (0.4928-0.6021)	<.001
Cass, MO	0.2857 (0.1125-0.6434)	.001
Clay, MO	0.6735 (0.4853-0.9295)	.02
Jackson, MO	0.5464 (0.4492-0.6627)	<.001
Johnson, KS	0.5076 (0.4186-0.6138)	<.001
Leavenworth, KS	1.3000 (0.5266-3.3119)	.68
Other	0.5556 (0.3265-0.9255)	.02
Platte, MO	0.4815 (0.2281-0.9658)	.04
Wyandotte, KS	0.5451 (0.4441-0.6671)	<.001
Q9: Skip health care		
Overall	0.6378 (0.5574-0.7291)	<.001
Cass, MO	0.2857 (0.0290-1.5006)	.18
Clay, MO	1.0769 (0.4929-1.9118)	.89
Jackson, MO	0.5436 (0.4095-0.7173)	<.001
Johnson, KS	0.6825 (0.5127-0.9050)	.01

Question[a]	OR[b] (95% CI)	*P* value
Leavenworth, KS	0.7778 (0.2462-2.3470)	.80
Other	0.7419 (0.4131-1.3144)	.34
Platte, MO	0.7500 (0.2791-1.9394)	.66
Wyandotte, KS	0.6068 (0.4819-0.7615)	<.001
Q10: Health literacy		
Overall	0.8729 (0.7823-0.9737)	.02
Cass, MO	0.6667 (0.2678-1.5863)	.42
Clay, MO	0.7458 (0.4929-1.1208)	.17
Jackson, MO	0.9125 (0.7241-1.1492)	.46
Johnson, KS	0.9510 (0.7772-1.1633)	.65
Leavenworth, KS	1.1000 (0.4241-2.8891)	>.99
Other	0.7879 (0.4525-1.3583)	.44
Platte, MO	0.7857 (0.3228-1.8630)	.69
Wyandotte, KS	0.8263 (0.6729-1.0137)	.07
Q11: Support (Overall)	0.9337 (0.8583-1.0157)	.11
Q12: Need help in general (Overall)	0.7368 (0.6619-0.8198)	<.001

[a] For detailed questions, see Textbox 1.

[b] OR: odds ratio.

[c] MO: Missouri.

[d] KS: Kansas.

Discussion

Principal Findings

An interesting takeaway from the pre-post COVID-19 comparison is that the post-COVID-19 responses across the questions showed an improvement in social needs. Less respondents answered “yes” to the social needs questions in the post-COVID-19 survey period. At this point, we can only speculate on the reasons for this improvement. One possibility is that this was due to the government policies issued during the pandemic [38]. Across society, changes were made regarding economic relief, food availability, and the ability to evict tenants; thus, it is possible that these changes are directly impacting individuals’ social needs [39,40]. This trend was maintained when breaking down the analysis to the county level, notably in Wyandotte County. Based on the results, it is possible that Wyandotte County was affected by COVID-19 differently than Johnson County or Jackson County. It certainly appears that some of the responses to the social needs survey improved in the post-COVID-19 period to the greatest extent in Wyandotte County. Further investigation should be performed with respect to Wyandotte County to try and explain this observation. When looking at the most recent responses, Wyandotte County ranked higher in the proportion of “yes” responses to all questions except questions 4 and 11, and differences in outcomes were also significant when comparing all counties in the analysis.

Among the counties examined in this analysis, Johnson County and Jackson County have the two largest populations; however, Johnson County had the lowest affirmative response rate for nearly every question, while Jackson County had an affirmative response rate second only to that of Wyandotte County. This suggests that a county’s population is not proportional to the resources available for its residents to weather a pandemic-related storm and that large communities may have worse social needs outcomes than smaller communities. Similarly, both Johnson County and Jackson County are urban communities and yet their outcomes are strikingly different, suggesting that the urban/suburban population density may not be a factor in the social needs for a given population. The primary difference may lie in the socioeconomic factors associated with the various counties. Johnson County is a more socially affluent community with respect to nearly all metrics, with low unemployment rates, a strong social safety net, and good access to health care and educational opportunities. These factors should be considered in future research.

As observed in this study, the usage and expansion of assistance programs may lead to a reduction in social needs issues. Across the state of Kansas and in western Missouri, we found improvements in TUKHS survey responses after COVID-19 policies had been implemented. When broken down by county, it became clear that these programs impacted each county differently. As a society, we must better utilize assistance programs and policies that are in place and understand the different and unique needs of our diverse populations. Health and social workers should be well versed and keep their patients informed on programs that they are eligible for. These programs are not useful without public knowledge of their existence.

Limitations

This study has several important limitations that should be considered when implementing policy recommendations and determining future research. First, the data are limited only to patients of TUKHS, located in the Kansas City Metropolitan Statistical Area. Over 98% of the unique responses were from individuals residing in counties that are part of the metropolitan area and this population may not be representative of other counties within the states of Kansas or Missouri. Due to socioeconomic conditions prevalent in rural areas, such as high rates of poverty, unemployment, and lack of social services, patients located in these counties may not see similar rates of improvements in social needs following the COVID-19 pandemic declaration.

Second, due to the large sample size and unadjusted *P* values, even small differences in outcomes will be statistically significant, but these results are not necessarily practically significant. A determination should be made as to what level of improvement is considered *practically* significant before policy recommendations can be made using this analysis. In addition, the sample size is much smaller for rural counties (<2% of responses), which could skew the results, and larger sampling would potentially correct for the rural-urban disparity.

Third, the data set does not include uninsured patients. Although we do not have data on household income for the individuals that responded to this survey, the presence of insurance could be used as a proxy variable, whereas those with insurance have higher incomes than those who do not. Because this study only considered those with insurance, we can presume that these individuals are less likely to answer "yes" to the questions presented in the survey, which could underrepresent the baseline scores as well as the level of improvement.

Moreover, the economic and societal hardships brought about by COVID-19 could have changed public perceptions of social needs. What was once considered an issue before the pandemic may not have been considered an issue after. Perhaps a societal shift in the perception or willingness to share social needs accounts for our findings. Patients may be less likely to disclose their social needs to health care professionals now that they have seen the government willingly address them. Perhaps the social needs survey was distributed with less importance during the pandemic as COVID-19 screening surveys took priority during patient visits. Survey responses are a subjective measurement that can be impacted by self-reported biases, as each respondent has different thoughts, behaviors, and feelings in the context of social needs. Some of their social needs could be relative to their perception based on past encounters.

Due to the limitations of the data and disparities in the patient population, the generalizability of this study may be limited to populations in urban and suburban areas. However, despite the small sample size, significant differences were found when aggregate patient responses from rural counties were compared to the urban counties, and previous research on rural poverty and lack of health care in these areas suggest a larger sample will not likely change the results [41]. By contrast, additional data from rural patients may increase the disparity in these populations and confirm the differences we identified.

Policy Implications and Future Research

Further research should focus on certain counties, particularly those in rural areas to increase the number of patients from these areas. Since education has been identified as an important indicator of poverty and income, exploring educational programs in poor-performing counties could improve outcomes. Social workers and navigators play an important role in helping individuals access food pantries, educational opportunities, employment, and other resources that influence these outcomes; thus, further study should be directed toward outcomes associated with interactions that reduce barriers to care, and unit-level acuity metrics should be evaluated and reviewed frequently to understand the correlation between health outcomes and patient success rates in accessing these resources [42].

Finally, unlike the responses to many viruses that have spread throughout the world in recent decades, the United States took unprecedented steps to not only curtail the spread of COVID-19 but also to minimize the economic impact on the country and ensure the financial well-being of its people. Although it is unclear whether the country would implement similar policies in the event of a future pandemic, the impact of these policies should be thoroughly analyzed in the context of social needs outcomes, as they could be a significant reason why many individuals are not experiencing severe negative impacts.

Acknowledgments

We would like to extend our gratitude to Matthew S Mayo and Byron Gajewski for their constructive feedback on the manuscript.

Data Availability

Detailed data cannot be shared publicly to protect the privacy of individual participants. Information to support the findings of these analyses can be made available by contacting the corresponding author, who, upon reasonable request and understanding of the intended use of the data, will provide the requested information in a manner that continues to protect individual patient information.

Conflicts of Interest

None declared.

References

1. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 2014;129 Suppl 2(Suppl 2):19-31 [FREE Full text] [doi: 10.1177/00333549141291S206] [Medline: 24385661]
2. Knickman J, Elbel B. Jonas and Kovner's Health care delivery in the United States, 12th edition. New York: Springer Publishing Company; 2019.
3. Maziak W, Ward KD, Stockton MB. Childhood obesity: are we missing the big picture? Obes Rev 2008 Jan;9(1):35-42. [doi: 10.1111/j.1467-789X.2007.00376.x] [Medline: 18154601]
4. Robert Wood Johnson Foundation Commission to Build a Healthier America. Time to act: Investing in the health of our children and communities. Robert Wood Johnson Foundation. 2014. URL: https://www.rwjf.org/en/insights/our-research/2014/01/recommendations-from-the-rwjf-commission-to-build-a-healthier-am.html [accessed 2023-04-10]
5. Grunwell JR, Opolka C, Mason C, Fitzpatrick AM. Geospatial analysis of social determinants of health identifies neighborhood hot spots associated with pediatric intensive care use for life-threatening asthma. J Allergy Clin Immunol Pract 2022 Apr;10(4):981-991 [FREE Full text] [doi: 10.1016/j.jaip.2021.10.065] [Medline: 34775118]
6. Soulsby WD, Balmuri N, Cooley V, Gerber LM, Lawson E, Goodman S, CARRA Registry Investigators. Social determinants of health influence disease activity and functional disability in polyarticular juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2022 Mar 07;20(1):18 [FREE Full text] [doi: 10.1186/s12969-022-00676-9] [Medline: 35255941]
7. County snapshot: Wyandotte, Kansas. University of Wisconsin Population Health Institute. 2022. URL: https://www.countyhealthrankings.org/explore-health-rankings/kansas/wyandotte?year=2022 [accessed 2022-12-09]
8. Fawcett SB, Sepers CE, Jones J, Jones L, McKain W. Participatory evaluation of a community mobilization effort to enroll Wyandotte County, Kansas, residents through the Affordable Care Act. Am J Public Health 2015 Jul;105(Suppl 3):S433-S437. [doi: 10.2105/AJPH.2014.302505] [Medline: 25905820]
9. Lee D, Paul C, Pilkington W, Mulrooney T, Diggs N, Kumar D. Examining the effects of social determinants of health on COVID-19 related stress, family's stress and discord, and personal diagnosis of COVID-19. J Affect Disord Rep 2021 Jul;5:100183 [FREE Full text] [doi: 10.1016/j.jadr.2021.100183] [Medline: 34151315]
10. Abrams EM, Greenhawt M, Shaker M, Pinto AD, Sinha I, Singer A. The COVID-19 pandemic: adverse effects on the social determinants of health in children and families. Ann Allergy Asthma Immunol 2022 Jan;128(1):19-25 [FREE Full text] [doi: 10.1016/j.anai.2021.10.022] [Medline: 34699969]
11. Cardona S, Felipe N, Fischer K, Sehgal NJ, Schwartz BE. Vaccination disparity: quantifying racial inequity in COVID-19 vaccine administration in Maryland. J Urban Health 2021 Aug 17;98(4):464-468 [FREE Full text] [doi: 10.1007/s11524-021-00551-0] [Medline: 34142349]
12. Fitzpatrick KM, Willis DE. Homeless and hungry: food insecurity in the land of plenty. Food Secur 2021 Nov 04;13(1):3-12 [FREE Full text] [doi: 10.1007/s12571-020-01115-x] [Medline: 33169087]
13. Leddy AM, Weiser SD, Palar K, Seligman H. A conceptual model for understanding the rapid COVID-19-related increase in food insecurity and its impact on health and healthcare. Am J Clin Nutr 2020 Nov 11;112(5):1162-1169 [FREE Full text] [doi: 10.1093/ajcn/nqaa226] [Medline: 32766740]
14. MacDonald A. Food insecurity and educational attainment at the University of Arkansas. Thesis. University of Arkansas. 2016. URL: https://scholarworks.uark.edu/scwkuht/3/ [accessed 2023-04-10]
15. Kendrick A, Fantasia HC, Morse B, Willis DE. Food insecurity in college students: a concept analysis. Nurs Forum 2022 Sep;57(5):898-904. [doi: 10.1111/nuf.12737] [Medline: 35616363]
16. Owens MR, Brito-Silva F, Kirkland T, Moore CE, Davis KE, Patterson MA, et al. Prevalence and social determinants of food insecurity among college students during the COVID-19 pandemic. Nutrients 2020 Aug 20;12(9):2515 [FREE Full text] [doi: 10.3390/nu12092515] [Medline: 32825251]
17. Shostak S, Bandini J, Cadge W, Donahue V, Lewis M, Grone K, et al. Encountering the social determinants of health on a COVID-19 ICU: frontline providers' perspectives on inequality in a time of pandemic. SSM Qual Res Health 2021 Dec;1:100001 [FREE Full text] [doi: 10.1016/j.ssmqr.2021.100001] [Medline: 34870264]
18. Barker L, Luman ET, McCauley MM, Chu SY. Assessing equivalence: an alternative to the use of difference tests for measuring disparities in vaccination coverage. Am J Epidemiol 2002 Dec 01;156(11):1056-1061. [doi: 10.1093/aje/kwf149] [Medline: 12446263]
19. Ndugga N, Pham O, Hill L, Artiga S, Alam R, Parker N. Latest data on COVID-19 vaccinations race/ethnicity. Kaiser Family Foundation. 2021. URL: https://www.kff.org/coronavirus-covid-19/issue-brief/latest-data-on-covid-19-vaccinations-by-race-ethnicity/ [accessed 2021-11-01]
20. Williams AM, Clayton HB, Singleton JA. Racial and ethnic disparities in COVID-19 vaccination coverage: the contribution of socioeconomic and demographic factors. Am J Prev Med 2022 Apr;62(4):473-482 [FREE Full text] [doi: 10.1016/j.amepre.2021.10.008] [Medline: 34872772]
21. Acosta AM, Garg S, Pham H, Whitaker M, Anglin O, O'Halloran A, et al. Racial and ethnic disparities in rates of COVID-19-associated hospitalization, intensive care unit admission, and in-hospital death in the United States from March 2020 to February 2021. JAMA Netw Open 2021 Oct 01;4(10):e2130479 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.30479] [Medline: 34673962]

22. Boserup B, McKenney M, Elkbuli A. Disproportionate impact of COVID-19 pandemic on racial and ethnic minorities. Am Surg 2020 Dec;86(12):1615-1622 [FREE Full text] [doi: 10.1177/0003134820973356] [Medline: 33231496]
23. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among Black patients and white Patients with Covid-19. N Engl J Med 2020 Jun 25;382(26):2534-2543 [FREE Full text] [doi: 10.1056/NEJMsa2011686] [Medline: 32459916]
24. Andrasfay T, Goldman N. Association of the COVID-19 pandemic with estimated life expectancy by race/ethnicity in the United States, 2020. JAMA Netw Open 2021 Jun 01;4(6):e2114520 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.14520] [Medline: 34165582]
25. Walensky R. Temporary halt in residential evictions in communities with substantial or high levels of community transmission to prevent the further spread of COVID-19. Centers for Disease Control and Prevention Department of Health and Human Services. 2021 Aug 03. URL: https://stacks.cdc.gov/view/cdc/108498 [accessed 2023-04-10]
26. Leifheit K, Linton SL, Raifman J, Schwartz GL, Benfer EA, Zimmerman FJ, et al. Expiring eviction moratoriums and COVID-19 incidence and mortality. Am J Epidemiol 2021 Dec 01;190(12):2503-2510 [FREE Full text] [doi: 10.1093/aje/kwab196] [Medline: 34309643]
27. Dunn CG, Kenney E, Fleischhacker SE, Bleich SN. Feeding low-income children during the Covid-19 pandemic. N Engl J Med 2020 Apr 30;382(18):e40. [doi: 10.1056/NEJMp2005638] [Medline: 32227759]
28. QuickFacts Wyandotte County, Kansas. United States Census Bureau. 2021. URL: https://www.census.gov/quickfacts/wyandottecountykansas [accessed 2022-12-09]
29. The safest and most dangerous places in Wyandotte County, KS: crime maps and statistics. Crime Grade. 2022. URL: https://crimegrade.org/safest-places-in-wyandotte-county-ks/ [accessed 2022-12-09]
30. The Health Leads Screening Toolkit. Health Leads. 2022. URL: https://healthleadsusa.org/resources/the-health-leads-screening-toolkit/ [accessed 2022-12-09]
31. Brady E, Bridges K, Murray M, Cheng H, Liu B, He J, et al. Relationship between a comprehensive social determinants of health screening and type 2 diabetes mellitus. Prev Med Rep 2021 Sep;23:101465 [FREE Full text] [doi: 10.1016/j.pmedr.2021.101465] [Medline: 34194965]
32. The R Project for Statistical Computing. URL: https://www.R-project.org/ [accessed 2021-11-01]
33. Population and housing unit estimates. United States Census Bureau. URL: https://www.census.gov/programs-surveys/popest.html [accessed 2021-11-01]
34. Kind A, Buckingham WR. Making neighborhood-disadvantage metrics accessible - The Neighborhood Atlas. N Engl J Med 2018 Jun 28;378(26):2456-2458 [FREE Full text] [doi: 10.1056/NEJMp1802313] [Medline: 29949490]
35. Butterfield R, Gottschalk M, Soloway L, LaBrec P. The relationship between social determinants of health and healthcare utilization. SAS Institute. 2019. URL: https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3563-2019.pdf [accessed 2022-12-09]
36. Friedman C. The Social Determinants of Health Index. Rehabil Psychol 2020 Feb;65(1):11-21. [doi: 10.1037/rep0000298] [Medline: 31724414]
37. Figueroa JF, Frakt AB, Jha AK. Addressing social determinants of health: time for a polysocial risk score. JAMA 2020 Apr 28;323(16):1553-1554. [doi: 10.1001/jama.2020.2436] [Medline: 32242887]
38. Béland D, Dinan S, Rocco P, Waddan A. Social policy responses to COVID-19 in Canada and the United States: explaining policy variations between two liberal welfare state regimes. Soc Policy Adm 2020 Nov;55(2):280-294. [doi: 10.1111/spol.12656]
39. Kim D. Financial hardship and social assistance as determinants of mental health and food and housing insecurity during the COVID-19 pandemic in the United States. SSM Popul Health 2021 Dec;16:100862 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100862] [Medline: 34692973]
40. Sandoval-Olascoaga S, Venkataramani AS, Arcaya MC. Eviction moratoria expiration and COVID-19 infection risk across strata of health and socioeconomic status in the United States. JAMA Netw Open 2021 Aug 02;4(8):e2129041 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.29041] [Medline: 34459904]
41. Douthit N, Kiv S, Dwolatzky T, Biswas S. Exposing some important barriers to health care access in the rural USA. Public Health 2015 Jun;129(6):611-620. [doi: 10.1016/j.puhe.2015.04.001] [Medline: 26025176]
42. Juvé-Udina ME, González-Samartino M, López-Jiménez MM, Planas-Canals M, Rodríguez-Fernández H, Batuecas Duelt IJ, et al. Acuity, nurse staffing and workforce, missed care and patient outcomes: a cluster-unit-level descriptive comparison. J Nurs Manag 2020 Nov 19;28(8):2216-2229 [FREE Full text] [doi: 10.1111/jonm.13040] [Medline: 32384199]

Abbreviations

ICU: intensive care unit
RUCC: rural-urban continuum code
SDoH: Social Determinants of Health
TUKHS: The University of Kansas Health System

Edited by A Mavragani, T Sanchez; submitted 23.07.22; peer-reviewed by B Padhi, Y Guo, N Mathews, A Steeves-Reece; comments to author 23.11.22; revised version received 21.12.22; accepted 23.03.23; published 25.04.23.

Please cite as:
Mudaranthakam DP, Pepper S, Fortney T, Alsup A, Woodward J, Sykes K, Calhoun E
The Effects of COVID-19 Pandemic Policy on Social Needs Across the State of Kansas and Western Missouri: Paired Survey Response Testing
JMIR Public Health Surveill 2023;9:e41369
URL: https://publichealth.jmir.org/2023/1/e41369
doi:10.2196/41369
PMID:36977199

Original Paper

Impact of Human Mobility on COVID-19 Transmission According to Mobility Distance, Location, and Demographic Factors in the Greater Bay Area of China: Population-Based Study

Jizhe Xia[1,2], PhD; Kun Yin[3], PhD; Yang Yue[1,2], PhD; Qingquan Li[1,2], PhD; Xiling Wang[4], PhD; Dongsheng Hu[5], PhD; Xiong Wang[6], PhD; Zhanwei Du[7,8], PhD; Ben J Cowling[7,8], PhD; Erzhen Chen[9], PhD; Ying Zhou[9], PhD

[1]Department of Urban Informatics, Guangdong Key Laboratory for Urban Informatics, Shenzhen University, Shenzhen, China

[2]Guangdong Laboratory of Artificial Intelligence and Digital Economy, Shenzhen, China

[3]School of Global Health, Chinese Center for Tropical Diseases Research, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]School of Public Health and Key Laboratory of Public Health Safety, Fudan University, Shanghai, China

[5]School of Public health, Shenzhen University, Shenzhen, China

[6]Institute for Advanced Study, Shenzhen University, Shenzhen, China

[7]World Health Organisation Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, Hong Kong, China (Hong Kong)

[8]Laboratory of Data Discovery for Health, Hong Kong Science and Technology Park, Hong Kong, China (Hong Kong)

[9]Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Corresponding Author:
Ying Zhou, PhD
Ruijin Hospital
Shanghai Jiao Tong University School of Medicine
Ruijin Second Road 197
HuangPu District
Shanghai, 200000
China
Phone: 86 64370045 ext 600603
Email: zy12941@rjh.com.cn

Abstract

Background: Mobility restriction was one of the primary measures used to restrain the spread of COVID-19 globally. Governments implemented and relaxed various mobility restriction measures in the absence of evidence for almost 3 years, which caused severe adverse outcomes in terms of health, society, and economy.

Objective: This study aimed to quantify the impact of mobility reduction on COVID-19 transmission according to mobility distance, location, and demographic factors in order to identify hotspots of transmission and guide public health policies.

Methods: Large volumes of anonymized aggregated mobile phone position data between January 1 and February 24, 2020, were collected for 9 megacities in the Greater Bay Area, China. A generalized linear model (GLM) was established to test the association between mobility volume (number of trips) and COVID-19 transmission. Subgroup analysis was also performed for sex, age, travel location, and travel distance. Statistical interaction terms were included in a variety of models that express different relations between involved variables.

Results: The GLM analysis demonstrated a significant association between the COVID-19 growth rate ratio (GR) and mobility volume. A stratification analysis revealed a higher effect of mobility volume on the COVID-19 GR among people aged 50-59 years (GR decrease of 13.17% per 10% reduction in mobility volume; *P*<.001) than among other age groups (GR decreases of 7.80%, 10.43%, 7.48%, 8.01%, and 10.43% for those aged ≤18, 19-29, 30-39, 40-49, and ≥60 years, respectively; *P*=.02 for the interaction). The impact of mobility reduction on COVID-19 transmission was higher for transit stations and shopping areas (instantaneous reproduction number [R_t] decreases of 0.67 and 0.53 per 10% reduction in mobility volume, respectively) than for workplaces, schools, recreation areas, and other locations (R_t decreases of 0.30, 0.37, 0.44, and 0.32, respectively; *P*=.02 for the interaction). The association between mobility volume reduction and COVID-19 transmission was lower with decreasing mobility distance as there was a significant interaction between mobility volume and mobility distance with regard to R_t (*P*<.001

for the interaction). Specifically, the percentage decreases in R_t per 10% reduction in mobility volume were 11.97% when mobility distance increased by 10% (Spring Festival), 6.74% when mobility distance remained unchanged, and 1.52% when mobility distance declined by 10%.

Conclusions: The association between mobility reduction and COVID-19 transmission significantly varied according to mobility distance, location, and age. The substantially higher impact of mobility volume on COVID-19 transmission for longer travel distance, certain age groups, and specific travel locations highlights the potential to optimize the effectiveness of mobility restriction strategies. The results from our study demonstrate the power of having a mobility network using mobile phone data for surveillance that can monitor movement at a detailed level to measure the potential impacts of future pandemics.

(JMIR Public Health Surveill 2023;9:e39588) doi:10.2196/39588

KEYWORDS

COVID-19; mobility restriction; mobility distance; demographic factors; locations

Introduction

Background

The COVID-19 pandemic had led to over 630 million infections and 6 million deaths worldwide by November 2022 [1]. Human mobility is critical to the transmission of airborne diseases. Reduction in human mobility was considered as one of the main tools to suppress and mitigate the pandemic, and has been applied widely and even repeatedly in multiple waves globally [2]. However, frequent mobility restrictions globally have caused unprecedented public health, social, and economic challenges. For example, mobility reduction was associated with indirect health outcomes, such as increased cerebrovascular excess deaths [3] and increased mental health issues [4]. Substantial economic and social burdens of lockdown measures were reflected in regions with different financial capacities [5]. The vaccine program may fail to help reach herd immunity, and easing mobility restrictions should be carefully considered against the risk of new outbreaks [6]. When variants with higher transmissibility circulate globally, mobility restriction could be one of the primary tools to control the pandemic. Therefore, it is important to determine how to optimize such mobility restriction measures to balance the indirect adverse outcomes in the society and economy, and the benefits of controlling transmission risk for COVID-19.

In almost 3 years of the COVID-19 pandemic, unprecedented efforts have been made to explore the association between human mobility and COVID-19 transmission [7]. Mobile-based mobility data can be used to assess mobility reduction and the effectiveness of social distancing for mitigating disease spread [8,9]. Most studies used mobility volume (ie, number of trips) for investigating the impact of mobility restriction policies on the transmission of COVID-19, but the correlation between reductions in mobility volume and COVID-19 transmission became weaker and even disappeared in the middle stage of the pandemic according to 2 studies in the United States [10,11]. Another study using the data of 52 countries indicated that the association between mobility volume and the reproduction number R changed over time [12]. After public health interventions were relaxed, the association between mobility and COVID-19 transmission decoupled in most countries. More detailed research is needed to better understand such a relation. Very few studies have explored the association between mobility distance and COVID-19 transmission [13], and it is unclear whether the weaker impact of mobility volume was influenced by a disproportional change in mobility distance. COVID-19 is known to exhibit age-related severity, and exposure rates are strongly dependent on age [14]. Similar mobility reductions among <18, 18-64, and ≥65 years age groups were reported in France [15], but it is unclear whether such reductions in various age groups have similar effects on COVID-19 transmission. Governments around the world have different policies for implementing and easing social distancing measures in various locations in the absence of evidence. These locations may have different transmission risks and social benefits [16]. It is critical to differentiate the impact of mobility reduction for various locations when making decisions to shut down specific locations. The impact of mobility reduction on COVID-19 transmission according to travel distance, location, and demographic factors is of great importance for understanding mobility restriction policies, but it has been poorly explored.

Objectives

In this study, we integrated anonymized geolocalized mobile phone data with census and demographic data in the Greater Bay Area of China. We aimed to analyze the impact of mobility reduction on COVID-19 transmission according to travel distance (long and short), location (workplaces, schools, recreation areas, shopping areas, transit stations, and other areas), age (≤18, 19-29, 30-39, 40-49, 50-59, and ≥60 years), and sex. The coverage rate of mobile phone use among the population aged 15-65 years was almost 100% [17]. This analysis could provide evidence to optimize mobility restriction policies for balancing the adverse outcomes of mobility reduction and the benefits of limiting community transmission of COVID-19.

Methods

Data Source

Mobility Data From the Greater Bay Area, China

Large volumes of anonymized aggregated mobile phone position data between January 1 and February 24, 2020, were collected for the 9 megacities of Guangzhou, Shenzhen, Foshan, Huizhou, Dongguan, Zhongshan, Zhaoqing, Zhuhai, and Jiangmen in Guangdong-Hong Kong-Macao Greater Bay Area, China. The Greater Bay Area is the most populated and largest urban area, and is 1 of the 4 largest bay areas in the world. Mobile phone data were provided by 1 of the 3 leading mobile phone service providers. Origin-destination matrices were constructed by computing the number of people that move between different locations on an hourly basis, as done previously [18]. The mobility volume was calculated as the number of trips between various locations. The mobility distance was determined by calculating the great circle distance between movement origins and destinations [19]. Aggregated sex- and age-specific daily mobility volumes were also obtained. Daily mobility volumes to various destinations (eg, workplaces, schools, recreation areas, shopping areas, transit stations, and other areas) were calculated by integrating the origin-destination matrices with the land use type of the trip destination (Table S1 in Multimedia Appendix 1). Official estimates of the total, sex-specific, and age-specific populations in the Greater Bay Area were retrieved from relevant government websites [20].

COVID-19 Transmission Data

Daily incidences of COVID-19 were obtained from official governmental reports in the Greater Bay Area, China [21]. Data on country-level estimates of the instantaneous reproduction number (R_t) were obtained from the EpiForecasts project by the London School of Hygiene & Tropical Medicine (London, UK) [22], which was calculated based on the daily number of COVID-19 infections by the European Centre for Disease Prevention and Control. The timings and details for public health interventions were taken as published on relevant government websites in the Greater Bay Area, China [23].

Ethics Approval

We used anonymized and aggregated mobile phone data at the population level without individual travel patterns for strict protection of personal privacy. All data were obtained in an anonymous format without personal identifying information. This study was approved by the Institutional Review Board of Shenzhen University, China (review number PN-202300030).

Statistical Analysis

The growth rate ratio (GR) of COVID-19 was computed as the average number of new cases per day over the previous 3 days to that over the previous 7 days. The static correlation and dynamic correlation between the GR and mobility volume were determined by Pearson correlation and rolling correlation. We tested the different day lags between the GR and mobility in the correlation analysis, as there may be a time lag between reported cases and true community infections. A generalized linear model (GLM) was established to test the association between mobility volume and COVID-19 transmission. Subgroup analysis was also performed for sex, age, travel location, and travel distance. The statistical interaction terms were included in a variety of models that express different relations between the involved variables. The R_t in Shenzhen by time for real transmission was estimated according to the likelihood-based estimation method [24].

An interaction analysis was performed by including the interaction terms of mobility volume ratio (VR) and distance ratio (DR) in the GLM analysis to determine whether the impact of mobility volume on COVID-19 transmission differs by mobility distance in the Greater Bay Area, China. The VR and DR were defined for each day (t), which quantified the change in mobility patterns and were similar to previous studies [25]. The baseline dates for the VR and DR were the normal days in 2020, which are here defined as the days before *Chunyun* (Spring Festival; January 1-10, 2020; Multimedia Appendix 1) for the Greater Bay Area, China. The VR is the sum of the total trips between various locations on a given day divided by the same measure on the baseline day, which reflects the change in the number of individual trips made to each area per day. VR_{it} was calculated as follows:

where V_{ij}^t represents the number of trips between areas i and j on day t, and t_0 represents the baseline measure. Using this function, VR values of 0, 0.5, and 1.0 indicate no trips, half the number of trips relative to baseline, and no change compared with baseline, respectively.

The DR represents the change in the distance of individual trips made to each area per day, relative to ordinary behavioral patterns (ie, before COVID-19), which was calculated as follows:

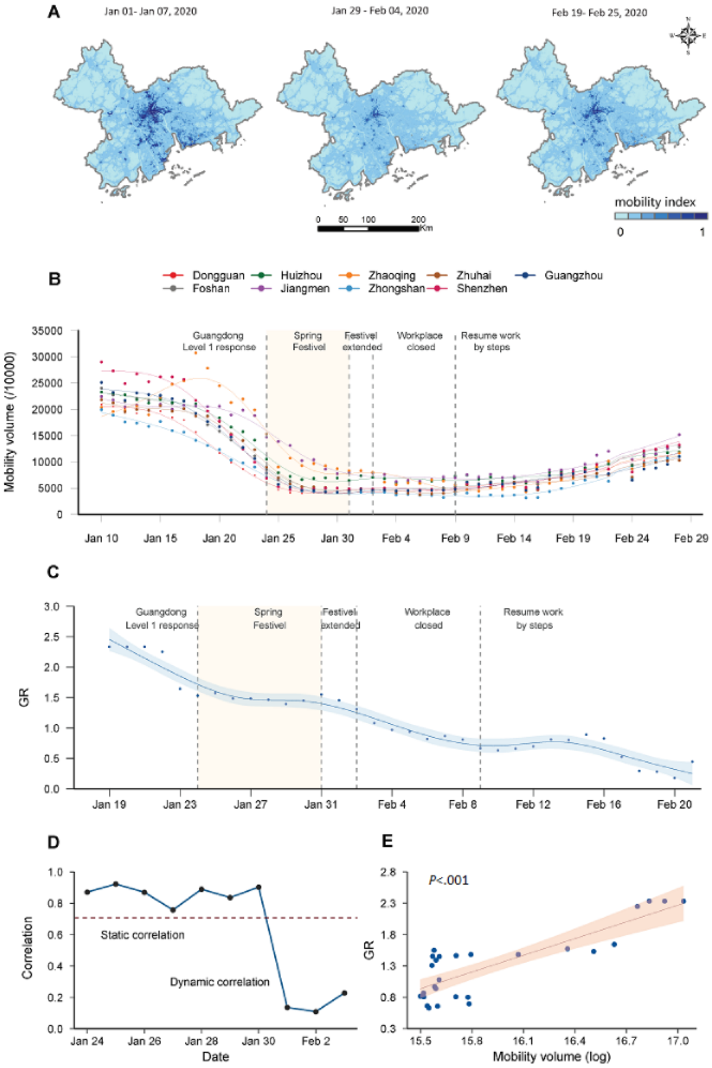
where D_{ij}^t represents the distance of trips between areas i and j on day t, and t_0 represents the baseline measure. DR values of 0, 0.5, and 1.0 indicate no trips, half of the mobility distance at baseline, and no change compared with baseline, respectively. Any value above 1 indicates that the mobility distance increased from baseline. The R_t ratio was defined as the R_t on a given day compared to the first R_t at the beginning of the study. ANOVA was used to evaluate the model with and without the interaction term. All statistical analyses were performed using R version 3.6.3 (R Foundation for Statistical Computing).

Results

The Effect of Mobility Volume on COVID-19 Transmission

The public health interventions in the Greater Bay Area, China during the study period are illustrated in Multimedia Appendix 2. The population mobility volume decreased by approximately 75%-85% and then was maintained at such a substantially low level for approximately 2 weeks in January 2020 in the Greater Bay Area, China (Figure 1A and 1B). The GR gradually decreased from 2.33 on January 19, 2020, to below 0.50 on February 20, 2020 (Figure 1C). The static correlation between the GR and mobility volume was 0.71 (*P*<.001; Figure 1D), considering the time lag of 2 days representing the highest correlation between the GR and mobility volume (Multimedia Appendix 3). The rolling correlation coefficients were >0.75 from January 24, 2020, to January 31, 2020 (all *P*<.01) but declined in February 2020. The GLM analysis demonstrated a significant association between the GR and mobility volume, with a 7.56% decrease in the GR per 10% reduction in mobility volume (95% CI 6.13%-9.00%; *P*<.001; Figure 1E).

Figure 1. Relationship between mobility volume and the GR in the Greater Bay Area, China. (A) Average mobility volume based on the mobile phone position data from January to February 2020 in the Greater Bay Area, China. (B) The change in mobility volume for 9 cities from January to February 2020 (the dots represent the observed data, and the plotted lines are smoothed by a generalized additive model). (C) The GR of COVID-19 from January to February 2020 (the line represents GLM fit to the data, and the shadow represents the 95% CI). (D) Correlation between mobility volume and the GR. (E) Association between mobility volume and the GR (the line represents GLM fit to the data, and the shadow represents the 95% CI). The dashed lines represent the main public health interventions. GLM: generalized linear model; GR: growth rate ratio.

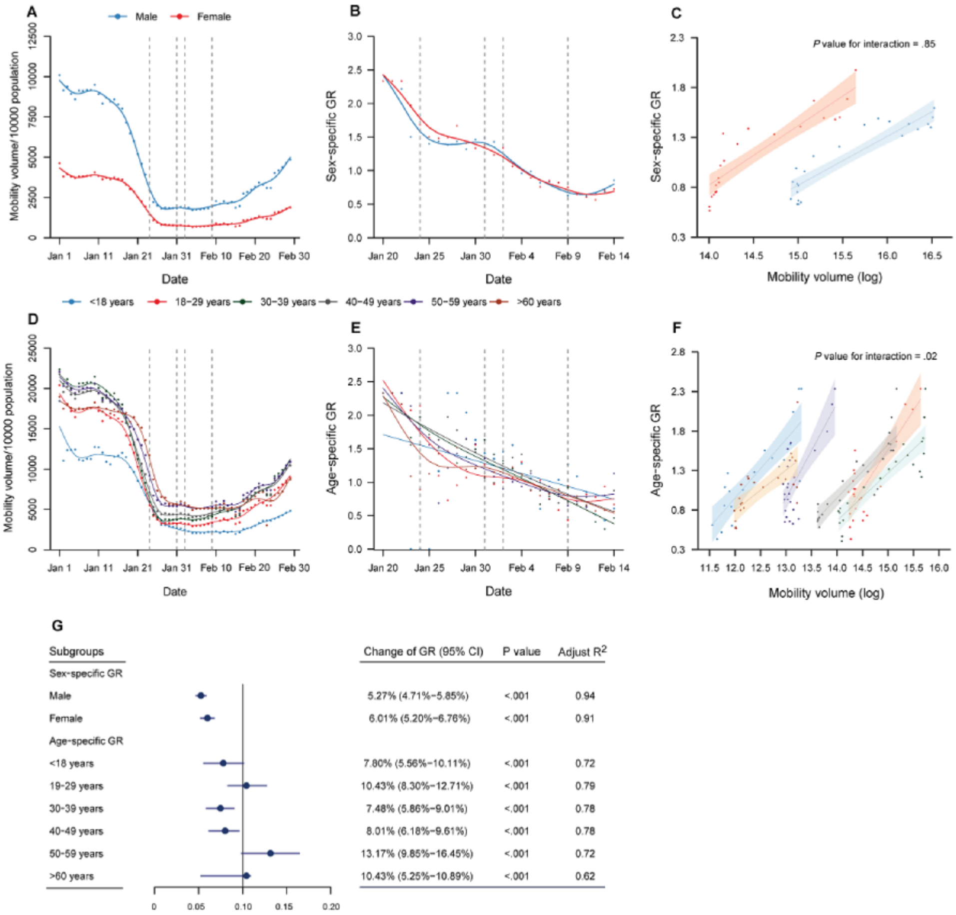


Demographic Disparities in the Impact of a Reduction in Mobility Volume on COVID-19 Transmission

We next investigated the impact of a reduction in mobility volume on COVID-19 transmission in various subgroups of sex and age. Although the mobility volume for females was less than that for males before, during, and after the COVID-19 pandemic (Figure 2A), a similar magnitude of decrease was observed (77.4% for females and 77.8% for males during the pandemic; Multimedia Appendix 4). The sex-specific GR declined with similar trends for both sexes (Figure 2B). The slopes of the association between the sex-specific mobility volume and sex-specific GR (Figure 2C and 2G) were similar in the GLM analysis for females and males.

Among different age groups, those aged 50-59 and ≥60 years exhibited the lowest levels of reduction in mobility volume of 72.9% and 67.0%, respectively, during the pandemic. The percentage reductions in mobility volume were 78.1%, 81.5%, 81.4%, and 77.7% for those aged ≤18, 19-29, 30-39, and 40-49 years, respectively (Figure 2D; Multimedia Appendix 4). The slopes of the association between the age-specific mobility volume and age-specific GR were significantly different among various age groups. The percentage decreases in the GR per 10% reduction in mobility volume were 7.80% for those aged ≤18 years (*P*<.001), 10.43% for those aged 19-29 years (*P*<.001), 7.48% for those aged 30-39 years (*P*<.001), 8.01% for those aged 40-49 years (*P*<.001), 13.17% for those aged 50-59 years (*P*<.001), and 10.43% for those aged ≥60 years (*P*<.001) (*P*=.02 for the interaction) (Figure 2F and 2G). The age group of 50-59 years had the highest slope in the GLM analysis for the age-specific GR and age-specific mobility volume.

Figure 2. Relationship between mobility volume and the GR by sex and age. (A) Time series of the daily average mobility volume and (B) the sex-specific GR for males and females. The dots represent the raw data, while the plotted lines are smoothed by a generalized additive model. (C) Relationship between the sex-specific GR and sex-specific mobility volume. The line represents GLM fit to the data, and the shadow represents the 95% CI. (D) Time series of the daily average mobility volume and (E) the age-specific GR for various age groups. The dots represent the raw data, while the plotted lines are smoothed by a generalized additive model. (F) Relationship between the age-specific GR and age-specific mobility volume. The line represents GLM fit to the data, and the shadow represents the 95% CI. The dashed lines represent the main public health interventions. (G) The change in the demographic-specific GR per 10% reduction in mobility volume. GLM: generalized linear model; GR: growth rate ratio.

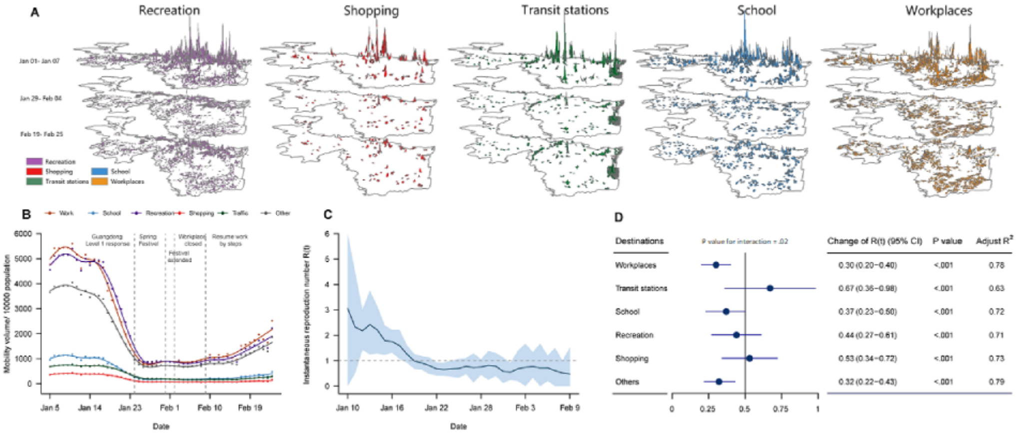


Subgroups	Change of GR (95% CI)	P value	Adjust R^2
Sex-specific GR			
Male	5.27% (4.71%–5.85%)	<.001	0.94
Female	6.01% (5.20%–6.76%)	<.001	0.91
Age-specific GR			
<18 years	7.80% (5.56%–10.11%)	<.001	0.72
19-29 years	10.43% (8.30%–12.71%)	<.001	0.79
30-39 years	7.48% (5.86%–9.01%)	<.001	0.78
40-49 years	8.01% (6.18%–9.61%)	<.001	0.78
50-59 years	13.17% (9.85%–16.45%)	<.001	0.72
>60 years	10.43% (5.25%–10.89%)	<.001	0.62

Location Disparities in the Impact of a Reduction in Mobility Volume on COVID-19 Transmission

The distribution of mobility volume at 6 types of locations at different time periods in the city of Shenzhen (one of the Greater Bay Area cities) is illustrated in Figure 3A. The mobility volume for workplaces was notably higher than other locations. The mobility volume at the 6 types of locations had declined during the pandemic, with average reductions of 83.8% for workplaces, 75.1% for transit stations, 83.5% for shopping areas, 82.9% for recreation areas, 83.7% for schools, and 81.8% for other locations (Figure 3B). The impact of mobility volume on the R_t was higher for the locations of transit stations and shopping areas, with the highest slope in the GLM analysis. The decreases in the R_t per 10% reduction in mobility volume were 0.67 for transit stations (*P*<.001), 0.53 for shopping areas (*P*<.001), 0.30 for workplaces (*P*<.001), 0.37 for schools (*P*<.001), 0.44 for recreation areas (*P*<.001), and 0.32 for other locations (*P*<.001) (*P*=.02 for the interaction; Figure 3D).

Figure 3. Relationship between mobility volume and COVID-19 transmission by various destinations. (A) The distribution of mobility volume for various destinations in the city of Shenzhen at different time periods in the Greater Bay Area, China (the peak of 3D bars represents the mobility volume in the given period), and (B) the time series of mobility volume at these destinations. (C) Time series of the instantaneous reproduction number (Rt) in Shenzhen. The shadow represents the 95% CI. (D) The change in the Rt per 10% reduction in mobility volume for a certain destination. Adjusted R^2 represents the goodness of fit for the models.

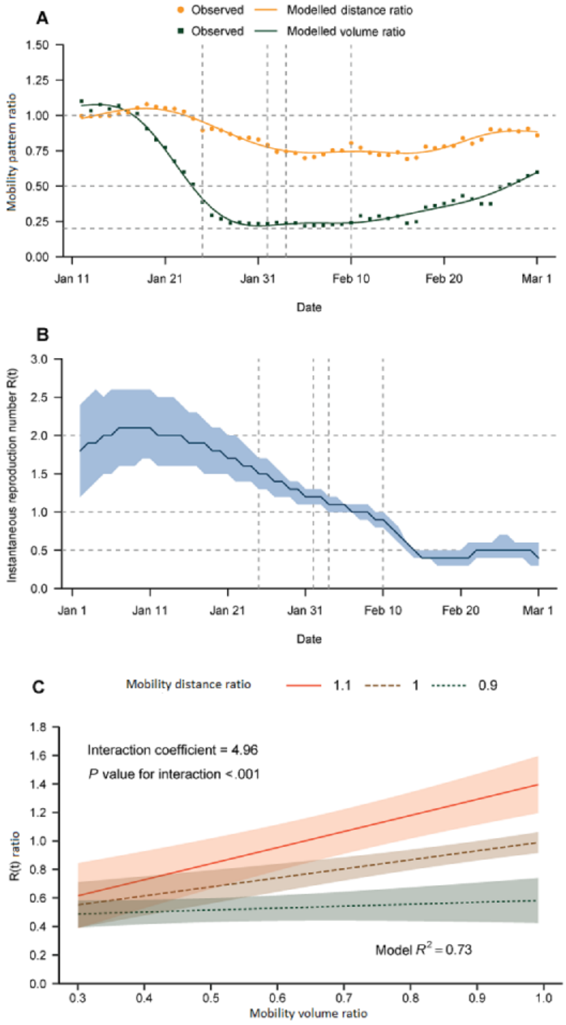


Mobility Distance Disparities in the Impact of a Reduction in Mobility Volume on COVID-19 Transmission

The change in mobility volume occurred earlier and was greater than the change in distance during the pandemic in the Greater Bay Area, China (Figure 4A). The population mobility volume had an average reduction of 76.1% in February compared with the baseline period of 2020 in the Greater Bay Area, China, and the mobility distance declined by an average of 18.6% (Figure 4A). The R_t had reduced to <1 by February 8, 2020, in China (Figure 4B). Exploratory analyses showed statistical evidence of an interaction between the VR and DR for the COVID-19 R_t ratio. In the analysis of Chinese cities, the impact of a reduction in mobility volume on the R_t was lower with a decreasing mobility distance. The percentage decreases in the R_t per 10% reduction in mobility volume were 11.97% when the average mobility distance increased by 10% compared with the baseline period, such as in the *Chunyun* period when people travelled longer distances in the Spring Festival (*P*<.001), 6.74% when the average mobility distance remained unchanged (*P*<.001), and 1.52% when the distance declined by 10% (*P*<.001) (*P*<.001 for the interaction; Figure 4C).

ANOVA showed that removing the interaction did significantly affect the fit of the model (*P*<.001). The R^2 for the model without an interaction was 0.65, which was lower than that for the model with an interaction (0.73 in the China model), increasing the variance of the dependent variable explained by the predictors. Finally, we validated the mobile phone mobility data for the Greater Bay Area, China by comparing it to Baidu mobility data for the whole country [26]. Here, the trend of mobility volume reduction was quite similar to the change in the mobility index from the Baidu Qianxi map (Multimedia Appendix 5).

Figure 4. Relationship between mobility volume and COVID-19 transmission by mobility distance. (A) Time series of the VR and DR in the Greater Bay Area, China. The dots represent the observed data, and the plotted lines are smoothed by a generalized additive model. (B) Time series of the instantaneous reproduction number (Rt) in the Greater Bay Area, China. The shadow represents the 95% CI. (C) Relationship between the VR and Rt ratio by different DRs in the Greater Bay Area, China. The line represents GLM fit to the data, and the shadow represents the 95% CI. DR: distance ratio; GLM: generalized linear model; VR: volume ratio.



Discussion

Principal Findings

In this study, we found that the impact of reductions in human mobility on COVID-19 transmission significantly varied by travel distance, location, and age. There was a significant positive interaction between mobility volume and mobility distance regarding COVID-19 transmission, with steeper slopes for the association (larger coefficient in the regression analysis) between mobility volume and COVID-19 transmission with increasing mobility distance. We found a significantly steeper slope for the association between the reduction in mobility volume and COVID-19 transmission among persons aged 50-59 years than among other age groups. Furthermore, the slope for the association was steeper for the locations of transit stations and shopping areas, compared with workplaces, schools, recreation areas, and other locations.

Our study indicated that the introduction of mobility restrictions in the Greater Bay Area, China led to a marked decrease in COVID-19 transmissibility. The time lag between mobility reduction and decline in the GR was estimated as 2 days at the very beginning of the pandemic in this area, which is shorter than the time of around 2 weeks in a similar study in the United States reported by Badr et al [11]. Our study also showed that the R_t in Shenzhen City, one of the megacities in this area, had declined from >3 to <1 within around 2 weeks, which means that the public health interventions in this area worked very fast. Therefore, the shorter time lag of 2 days between mobility reduction and change in the GR may be possible in the Greater Bay Area, China. The reduction in mobility volume with longer travel distances was associated with a greater reduction in COVID-19 transmission than shorter travel distances in our study. The decreasing correlations of mobility volume with COVID-19 transmission observed in our study were consistent with previous studies in the United States. Badr et al [11] found a strong correlation between mobility volume and COVID-19 case growth rates in the early stage of March to April 2020, but a strong linear association was absent after April 2020. Gatalo et al [10] also identified a strong correlation between March 27, 2020, and April 20, 2020, and only a weak correlation at later time periods (April 21, 2020, to May 24, 2020). The significant positive interaction between mobility volume and distance may provide a possible explanation for the time-driven relationship between mobility volume and COVID-19 transmission. The changes in mobility distance and volume are not necessarily synchronized. Our study indicated that there is an interaction, and changes in moving distance may affect the impact of mobility volume on COVID-19 transmission. This provides evidence for stricter restrictions in long-distance travel for better control of COVID-19 transmission. However, the different start times and levels of reduction in mobility volume and distance during the pandemic need to be further explored.

The mobility reduction was associated with a greater reduction in the GR for the age group of 50-59 years than the other age groups in our study. The slope of the association between mobility volume and the GR for those aged 50-59 years was steeper than that for the other age groups. Those aged 50-59 years showed lower mobility reduction than those aged ≤49 years, but were vulnerable to COVID-19 infection with a higher proportion of underlying chronic diseases compared with young people [27]. Those aged 50-59 years faced a higher risk of exposure to SARS-CoV-2 with more traveling for occupational and behavioral reasons compared with those aged ≥60 years [28]. These factors may provide possible explanations for the finding that most infections of COVID-19 in the initial wave in China were among those aged 50-59 years (22.4%) [29]. The highest proportion of COVID-19 cases was also among those aged 50-59 years during the initial wave from January to May 2020, but shifted toward younger people from June to August 2020 in the United States [30]. A similar age shift was observed in Europe, with the median age of patients with COVID-19 declining from 54 years at the beginning to 39 years in the later time periods [31]. Whether such an age shift is associated with a change in the mobility pattern for various age groups at the different stages of the pandemic needs further exploration. In addition, our data suggested that the declined magnitude of mobility volume and the effect of mobility volume on COVID-19 transmission were quite similar for both males and females, which is consistent with the conclusion of no sex variation for COVID-19 infections in China [32].

The mobility reduction for transit stations and shopping centers was associated with a greater reduction in COVID-19 transmission in the whole city compared with the findings for workplaces, schools, and recreation areas. The slopes for the association between mobility to transit stations and shopping centers and the R_t were steeper than the slopes for the association involving other locations. These high-contact environments are more crowded and therefore have a higher risk. However, the mobility reduction for transit stations was less than that for other locations. Many governments applied various policies for mobility restriction at specific locations, since there was not enough evidence available regarding which locations should be closed or which locations should remain open. Our study provides evidence that controlling mobility to a small number of locations could reduce transmission in the entire city. Therefore, location-specific mobility restrictions should be taken into consideration for precise interventions and reopening strategies with substantially lower economic costs [33].

Strengths and Limitations

Our results have public health and policy implications. First, we analyzed the relationship between mobility responses and COVID-19 transmission using mobility data of only a certain travel destination, travel distance, or demographic subgroup (age and sex groups) to gain more insightful knowledge. Our study provides evidence to identify hotspots of transmission and guide policy interventions for specific age groups or mobility patterns associated with higher risks of mobility-related COVID-19 transmission. Second, mobile phone data at fine spatial and temporal resolutions provide strong added value for explaining variations in COVID-19 transmission. The results from our study demonstrate the power of having mobility networks using mobile phone data that monitor movement at a detailed level across cities to measure the potential impacts of public health events. The network can be regularly updated and used to identify populations and travel characteristics at risk of adverse impacts during future pandemics or other crises. It is also suggested to set up a national mobility network that captures human mobility habits, which can form the basis for longitudinal studies.

It is important to note that our study has several limitations. First, we illustrated the detailed structural change in mobility patterns using mobile phone data from users in the Greater Bay Area, China, and these patterns may not be fully representative of other locations in China. However, the mobility change in the Greater Bay Area based on the mobile phone was quite similar to the change in the Baidu mobility index for the whole country, owing to nationally unified public health interventions. We believe that this analysis for the Greater Bay Area is an intuitive and representative estimate of the structural change in mobility patterns in China, but future extension of this analysis for the whole country should be further explored. Second, we focused on quantifying the relationship between mobility patterns and COVID-19 transmission; there is extensive evidence indicating that population-wide social distancing and other potential mitigating factors (eg, wearing face masks and washing hands) all contribute to achieving control of the COVID-19 pandemic [34,35]. Further studies are needed to evaluate the synergistic effect of other public health interventions along with mobility change in the environment confounded by climate, urban space, and other potential factors [36]. Third, our mobility estimates may also be biased toward the populations included in the mobile phone data, as the consumer location history feature is only available for smartphone users (young infants or very old people may be excluded from the data). However, despite these limitations, our fine-grained study facilitates a step toward using multiple data sets to capture population-level mobility patterns and provides important insights into the complex effects of mobility reduction for policymakers and future research.

Conclusions

The COVID-19 pandemic was the first time in human history that human mobility showed a large-scale decline after mobility restrictions to prevent and control the infectious disease. Our study demonstrated that the impact of reductions in human mobility on COVID-19 transmission was significantly modified by travel distance, travel location, and age. The higher impact of mobility reduction on COVID-19 transmission for longer distances, certain age groups, and specific locations highlights the potential to optimize mobility restriction policies to balance adverse health, society, and economic outcomes and the benefits of controlling the spread of COVID-19. It is of great significance to understand the impact of mobility reduction on the spread of infectious diseases in detail, and it provides evidence for the prevention and control of future pandemics.

Acknowledgments

We thank Mr Renzhe Xu for helpful suggestions, and Mr Juecheng Li and Ms Gongjian Wu for technical assistance. This work was supported by the National Key Research and Development (R&D) Program of China (grant number: 2018YFB2100704), Research on Prevention and Control of COVID-19 in the Guangdong Education Department (grant number: 2020KZDZX1171), Shenzhen Basic Research Fund (grant number: JCYJ20190808174209308), and National Natural Science Foundation of China (grant numbers: 42171400, 2021A1515011324, 82103945, and 71961137003). The funding body was not involved in the collection, analysis, or interpretation of the data, the writing of the article, or the decision to submit the manuscript for publication.

Data Availability

We purchased the mobile phone data (January to March 2020) from a service provider (China Unicom). Our data purchase agreement with China Unicom prohibits us from sharing the data with third parties, but interested parties can contact China Unicom to obtain the same data. We collated epidemiological data from publicly available data sources (government websites). All epidemiological information that was used has been documented in the article.

Authors' Contributions

YZ ran the study, analyzed and interpreted the data, wrote the paper, and obtained funding for the study. JX extracted the data, analyzed and interpreted the data, and obtained funding for the study. QL participated in discussions and obtained funding for the study. KY, Xiling W, Xiong W, DH, YY, BJC, EC, and ZD participated in discussions on the undertaking of the study. All authors reviewed the paper for content and approved the final report.

Conflicts of Interest

BJC has received honoraria from AstraZeneca, Fosun Pharma, GSK, Haleon, Moderna, Roche, and Sanofi Pasteur.

Multimedia Appendix 1
Supplementary materials and methods.
[DOC File , 468 KB - publichealth_v9i1e39588_app1.doc]

Multimedia Appendix 2
Main public health interventions in the Greater Bay Area, China during the study period.
[DOC File , 44 KB - publichealth_v9i1e39588_app2.doc]

Multimedia Appendix 3
Correlations between mobility volume and growth rate ratio at different time lags (in days).
[DOC File , 2335 KB - publichealth_v9i1e39588_app3.doc]

Multimedia Appendix 4
The magnitude of mobility volume change for various demographic groups.
[DOC File , 3146 KB - publichealth_v9i1e39588_app4.doc]

Multimedia Appendix 5
The time series of mobility volume from mobile phone data in the Greater Bay Area, China and the mobility index from Baidu for the whole country of China.
[DOC File , 781 KB - publichealth_v9i1e39588_app5.doc]

References

1. COVID-19 Data in Motion. Johns Hopkins Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/ [accessed 2022-11-03]

2. Flaxman S, Mishra S, Gandy A, Unwin H, Mellan T, Coupland H, Imperial College COVID-19 Response Team, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020 Jun 08;584(7820):257-261. [doi: 10.1038/s41586-020-2405-7] [Medline: 32512579]
3. Sharma R, Kuohn LR, Weinberger DM, Warren JL, Sansing LH, Jasne A, et al. Excess Cerebrovascular Mortality in the United States During the COVID-19 Pandemic. Stroke 2021 Jan;52(2):563-572 [FREE Full text] [doi: 10.1161/STROKEAHA.120.031975] [Medline: 33430638]
4. Devaraj S, Patel PC. Change in psychological distress in response to changes in reduced mobility during the early 2020 COVID-19 pandemic: Evidence of modest effects from the U.S. Soc Sci Med 2021 Mar;270:113615 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113615] [Medline: 33352476]
5. Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, et al. Economic and social consequences of human mobility restrictions under COVID-19. Proc Natl Acad Sci U S A 2020 Jul 07;117(27):15530-15535 [FREE Full text] [doi: 10.1073/pnas.2007658117] [Medline: 32554604]
6. Leung K, Wu JT, Leung GM. Effects of adjusting public health, travel, and social measures during the roll-out of COVID-19 vaccination: a modelling study. Lancet Public Health 2021 Sep;6(9):e674-e682 [FREE Full text] [doi: 10.1016/S2468-2667(21)00167-5] [Medline: 34388389]
7. Alessandretti L. What human mobility data tell us about COVID-19 spread. Nat Rev Phys 2022;4(1):12-13 [FREE Full text] [doi: 10.1038/s42254-021-00407-1] [Medline: 34877474]
8. Kishore K, Jaswal V, Verma M, Koushal V. Exploring the Utility of Google Mobility Data During the COVID-19 Pandemic in India: Digital Epidemiological Analysis. JMIR Public Health Surveill 2021 Aug 30;7(8):e29957 [FREE Full text] [doi: 10.2196/29957] [Medline: 34174780]
9. Kurita J, Sugishita Y, Sugawara T, Ohkusa Y. Evaluating Apple Inc Mobility Trend Data Related to the COVID-19 Outbreak in Japan: Statistical Analysis. JMIR Public Health Surveill 2021 Mar 15;7(2):e20335 [FREE Full text] [doi: 10.2196/20335] [Medline: 33481755]
10. Gatalo O, Tseng K, Hamilton A, Lin G, Klein E. Associations between phone mobility data and COVID-19 cases. The Lancet Infectious Diseases 2021 May;21(5):e111. [doi: 10.1016/s1473-3099(20)30725-8]
11. Badr HS, Gardner LM. Limitations of using mobile phone data to model COVID-19 transmission in the USA. The Lancet Infectious Diseases 2021 May;21(5):e113. [doi: 10.1016/s1473-3099(20)30861-6]
12. Nouvellet P, Bhatia S, Cori A, Ainslie KEC, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission. Nat Commun 2021 Mar 17;12(1):1090 [FREE Full text] [doi: 10.1038/s41467-021-21358-2] [Medline: 33597546]
13. Schlosser F, Maier BF, Jack O, Hinrichs D, Zachariae A, Brockmann D. COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. Proc Natl Acad Sci U S A 2020 Dec 29;117(52):32883-32890 [FREE Full text] [doi: 10.1073/pnas.2012326117] [Medline: 33273120]
14. Davies N, Klepac P, Liu Y, Prem K, Jit M, CMMID COVID-19 working group, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat Med 2020 Aug;26(8):1205-1211. [doi: 10.1038/s41591-020-0962-9] [Medline: 32546824]
15. Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. The Lancet Digital Health 2020 Dec;2(12):e638-e649. [doi: 10.1016/s2589-7500(20)30243-0]
16. Benzell S, Collis A, Nicolaides C. Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. Proc Natl Acad Sci U S A 2020 Jun 30;117(26):14642-14644 [FREE Full text] [doi: 10.1073/pnas.2008025117] [Medline: 32522870]
17. Chinese consumers at the forefront of digital technologies. Deloitte. URL: https://www2.deloitte.com/cn/en/pages/about-deloitte/articles/pr-deloitte-2018-mobile-consumer-survey.html [accessed 2021-07-26]
18. Zhou Y, Xu R, Hu D, Yue Y, Li Q, Xia J. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. The Lancet Digital Health 2020 Aug;2(8):e417-e424. [doi: 10.1016/s2589-7500(20)30165-5]
19. Porcu E, Bevilacqua M, Genton MG. Spatio-Temporal Covariance and Cross-Covariance Functions of the Great Circle Distance on a Sphere. Journal of the American Statistical Association 2016 Aug 18;111(514):888-898. [doi: 10.1080/01621459.2015.1072541]
20. Bulletin of Main Statistics from the National Census in Shenzhen. Shenzhen Statistics Department. URL: http://tjj.sz.gov.cn/ztzl/zt/sjfb/ [accessed 2020-03-03]
21. COVID-19 reports. Guangdong Health and Health Commission. URL: http://wsjkw.gd.gov.cn/xxgzbdfk/ [accessed 2020-08-12]
22. Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Res 2020 Dec 8;5:112. [doi: 10.12688/wellcomeopenres.16006.2]
23. Prevention and control reports. Guangdong Health and Health Commission. URL: http://wsjkw.gd.gov.cn/xxgzbdfk/fkdt/ [accessed 2020-08-12]

24. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol 2004 Sep 15;160(6):509-516 [FREE Full text] [doi: 10.1093/aje/kwh255] [Medline: 15353409]
25. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. The Lancet Infectious Diseases 2020 Nov;20(11):1247-1254. [doi: 10.1016/s1473-3099(20)30553-3]
26. Baidu qianxi map. Baidu Inc. URL: http://qianxi.baidu.com/ [accessed 2020-04-06]
27. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. J Infect 2020 Aug;81(2):e16-e25 [FREE Full text] [doi: 10.1016/j.jinf.2020.04.021] [Medline: 32335169]
28. Wynants L, Van Calster B, Collins G, Riley R, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020 Apr 07;369:m1328 [FREE Full text] [doi: 10.1136/bmj.m1328] [Medline: 32265220]
29. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) - China, 2020. China CDC Wkly 2020 Mar 21;2(8):113-122 [FREE Full text] [Medline: 34594836]
30. Boehmer TK, DeVies J, Caruso E, van Santen KL, Tang S, Black CL, et al. Changing Age Distribution of the COVID-19 Pandemic - United States, May-August 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 02;69(39):1404-1409 [FREE Full text] [doi: 10.15585/mmwr.mm6939e1] [Medline: 33001872]
31. Coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK – eleventh update: resurgence of cases. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-rapid-risk-assessment-20200810.pdf [accessed 2021-07-12]
32. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA 2020 Apr 07;323(13):1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]
33. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature 2021 Jan 10;589(7840):82-87. [doi: 10.1038/s41586-020-2923-3] [Medline: 33171481]
34. Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, van der Westhuizen H, et al. An evidence review of face masks against COVID-19. Proc Natl Acad Sci U S A 2021 Jan 26;118(4):e2014564118 [FREE Full text] [doi: 10.1073/pnas.2014564118] [Medline: 33431650]
35. Przekwas A, Chen Z. Washing hands and the face may reduce COVID-19 infection. Med Hypotheses 2020 Nov;144:110261 [FREE Full text] [doi: 10.1016/j.mehy.2020.110261] [Medline: 33254560]
36. Roberto Telles C, Lopes H, Franco D. SARS-COV-2: SIR Model Limitations and Predictive Constraints. Symmetry 2021 Apr 14;13(4):676. [doi: 10.3390/sym13040676]

Abbreviations

DR: distance ratio
GLM: generalized linear model
GR: growth rate ratio
Rt: instantaneous reproduction number
VR: volume ratio

Edited by T Sanchez, A Mavragani; submitted 20.05.22; peer-reviewed by B Rader, CR Telles, U Scholz, J Wherton; comments to author 05.10.22; revised version received 27.11.22; accepted 23.02.23; published 26.04.23.

Please cite as:
Xia J, Yin K, Yue Y, Li Q, Wang X, Hu D, Wang X, Du Z, Cowling BJ, Chen E, Zhou Y
Impact of Human Mobility on COVID-19 Transmission According to Mobility Distance, Location, and Demographic Factors in the Greater Bay Area of China: Population-Based Study
JMIR Public Health Surveill 2023;9:e39588
URL: https://publichealth.jmir.org/2023/1/e39588
doi:10.2196/39588
PMID:36848228

Original Paper

Participatory Surveillance for COVID-19 Trend Detection in Brazil: Cross-sectional Study

Salome Wittwer[1], MA; Daniela Paolotti[2], PhD; Guilherme Lichand[1], PhD; Onicio Leal Neto[3], PhD

[1]Department of Economics, University of Zurich, Zurich, Switzerland

[2]Data Science for Social Impact and Sustainability, ISI Foundation, Turin, Italy

[3]Institute for Information Security, Department of Computer Science, ETH Zürich, Zurich, Switzerland

Corresponding Author:
Onicio Leal Neto, PhD
Institute for Information Security
Department of Computer Science
ETH Zürich
Universitätstrasse 6
Zurich, 8092
Switzerland
Phone: 41 44 632 50 94
Email: onicio.batistalealneto@inf.ethz.ch

Abstract

Background: The ongoing COVID-19 pandemic has emphasized the necessity of a well-functioning surveillance system to detect and mitigate disease outbreaks. Traditional surveillance (TS) usually relies on health care providers and generally suffers from reporting lags that prevent immediate response plans. Participatory surveillance (PS), an innovative digital approach whereby individuals voluntarily monitor and report on their own health status via web-based surveys, has emerged in the past decade to complement traditional data collection approaches.

Objective: This study compared novel PS data on COVID-19 infection rates across 9 Brazilian cities with official TS data to examine the opportunities and challenges of using PS data, and the potential advantages of combining the 2 approaches.

Methods: The TS data for Brazil are publicly accessible on GitHub. The PS data were collected through the Brazil Sem Corona platform, a Colab platform. To gather information on an individual's health status, each participant was asked to fill out a daily questionnaire on symptoms and exposure in the Colab app.

Results: We found that high participation rates are key for PS data to adequately mirror TS infection rates. Where participation was high, we documented a significant trend correlation between lagged PS data and TS infection rates, suggesting that PS data could be used for early detection. In our data, forecasting models integrating both approaches increased accuracy up to 3% relative to a 14-day forecast model based exclusively on TS data. Furthermore, we showed that PS data captured a population that significantly differed from a traditional observation.

Conclusions: In the traditional system, the new recorded COVID-19 cases per day are aggregated based on positive laboratory-confirmed tests. In contrast, PS data show a significant share of reports categorized as potential COVID-19 cases that are not laboratory confirmed. Quantifying the economic value of PS system implementation remains difficult. However, scarce public funds and persisting constraints to the TS system provide motivation for a PS system, making it an important avenue for future research. The decision to set up a PS system requires careful evaluation of its expected benefits, relative to the costs of setting up platforms and incentivizing engagement to increase both coverage and consistent reporting over time. The ability to compute such economic tradeoffs might be key to have PS become a more integral part of policy toolkits moving forward. These results corroborate previous studies when it comes to the benefits of an integrated and comprehensive surveillance system, and shed light on its limitations and on the need for additional research to improve future implementations of PS platforms.

(JMIR Public Health Surveill 2023;9:e44517) doi:10.2196/44517

KEYWORDS

participatory surveillance; COVID-19; digital epidemiology; coronavirus; infectious disease; epidemic; pandemic; SARS-CoV-2; forecast; trend; reporting; self-report; surveillance

Introduction

The global COVID-19 pandemic in March 2020 has had unprecedented consequences around the world. It caused widespread illness and deaths, as well as worldwide economic, political, and social repercussions [1]. The occurrence of such an extraordinary event emphasizes the need for well-functioning disease surveillance systems to detect and monitor disease outbreaks and epidemics. Countries use disease monitoring systems to assess, predict, and mitigate infectious disease outbreaks [2,3]. Reliable and timely data are critical to protect populations and build the foundation for governments, policy makers, and officials to intervene and prevent widespread infections [4]. As such, developing and improving on existing surveillance methods remains a rapidly growing and emerging field [5].

Most current surveillance systems, listed here as traditional surveillance (TS) systems, rely largely on traditional health care institutions, such as clinics, hospitals, and laboratories, to systematically collect data from practitioners as a public health monitoring tool [2]. Health care providers send reports to public health officials with certain regional or national data aggregation. In a few cases (usually, only a small percentage), these reports are then confirmed by laboratory analysis. Reported cases are then accounted as official disease cases [6-8]. Since the data are sourced from different institutions, aggregation often involves several time lags throughout the chain of data collection, reducing the timeliness of measures and actions [4,9,10]. Moreover, the true burden of the disease is often underestimated. In industrialized countries, healthy adults with no previous conditions usually do not visit a doctor if their symptoms remain mild. In emerging countries, socioeconomically weak communities lack access or financial resources to seek medical aid and thus are overlooked by TS [11]. Over the last 2 decades, new digital disease surveillance approaches have emerged to supplement traditional data collection, such as participatory surveillance (PS). Participatory disease surveillance is understood as an approach that directly engages the public in providing health data. Individuals monitor and assess their own health status and are encouraged to submit self-reports through digital platforms using mobile apps, websites, and phone-based surveys via SMS text messages or automated calls (interactive voice response [IVR]) [12]. Any individual can register on the platform (if they live in a country where a PS platform is deployed) and participate on a voluntary basis. As such, digital crowdsourced data can be aggregated and analyzed in large numbers [13-15]. Users are asked to regularly complete a questionnaire aimed at gathering information about their current (lack of) symptoms, access to health care, risk exposure, and medications. In the case of COVID-19, the main difference compared to TS is the lack of laboratory-confirmed testing; positive cases are categorized as such only based on the reporting of certain symptoms, with a so-called syndromic surveillance approach. The added value of PS systems is that they can even reach individuals who do not engage with health care providers because of financial or cultural factors, because they live too remotely to have access to health care facilities, or because their symptoms are too mild to cause concern [12]. The rapid global increase in the use of mobile phones and wide access to the internet are largely responsible for the rise of such digital surveillance systems [3]. By integrating an additional subset of the population not covered by TS, the complementary data provide an additional layer of surveillance, potentially enabling more accurate detection of ongoing infections as well as anticipation of trend changes [13,14].

PS systems have so far proven to be accurate and reliable for influenza-like illness (ILI) surveillance [15,16]. Most recently, as testing capacities were exhausted across many countries in the context of the pandemic, PS systems were implemented to support traditional systems in monitoring and controlling COVID-19 infections [17,18]. So far, Brazil is the first and only Latin American country that has implemented a PS system on a large scale to carry out syndromic surveillance, specifically during the 2014 FIFA World Cup [19] and the 2016 Olympic Games [20]. Lately, a local Brazilian health authority has used a PS platform to complement the traditional system, with the goal of optimizing the targeting of test areas during the COVID-19 pandemic [17]. This PS system has been shown to be beneficial in identifying risk clusters for infections in this context; in particular, it was able to cover blind spots of the TS system, showcasing the potential to increase its sensitivity by complementing it with additional data, and to allocate scarce resources more efficiently by prioritizing certain areas for the distribution of test kits. Notably, in Europe and the United States, health agencies and governments are increasingly prone to using the innovative and digital PS approach as a complementary source of disease surveillance.

Low public funding and suboptimal resource allocation persist in the Brazilian health sector [21]. The limitations of the TS system stress the need to reduce the burden of diseases among vulnerable and socioeconomically weak communities. This motivates the study to examine the opportunities and challenges of a PS system in the context of COVID-19 case detection across 9 Brazilian cities. During 7 months of the global pandemic in 2020, several city-level governments implemented the *Brazil Sem Corona* PS platform to gather additional insights on the spread of the disease. *Brazil Sem Corona* was an initiative led by Colab in partnership with several local governments from Brazilian cities, with the purpose of leveraging PS data to complement TS systems in order to mitigate COVID-19 risk at the local level. For our analysis, we focused on 9 cities with the largest PS participation across Brazil. Our objective was to investigate the capability of PS data to approximately mirror traditional infection rates captured through TS, as well as the relevance of citizen participation for the identification of trends in COVID-19 cases. Furthermore, we investigated the potential benefits of combining the PS and TS systems for forecasting case trends. Among the benefits, we showed that the PS system captures a part of the population that is so far overseen by traditional sources.

Methods

Goal

In this work, the goal was to compare daily official COVID-19 infections at the municipality level with daily PS infection numbers.

TS Data

TS data for Brazil are publicly accessible on GitHub [22]. These data aggregate the official laboratory-confirmed daily new COVID-19 cases at the municipality level. Using the 2020 population size estimates for each city [23], we calculated the infection rate per 100,000 inhabitants as follows:

where *i* denotes the city and *t* denotes the day.

PS Data

PS data for this study were collected through the *Brazil Sem Corona* platform, a Colab participatory platform developed previously [24]. To gather information on an individual's health status, each participant was asked to fill out a daily questionnaire in the Colab app on symptoms as well as exposure. The app was available on the Apple Store as well as the Google Play Store. The list of symptoms was based on the COVID-19 case definition and contained the following: fever, cough, shortness of breath, runny nose, sore throat, headache, fatigue, nausea, rash, joint pain, chills, diarrhea, and loss of taste. Additionally, participants were asked to report about medication intake and whether they sought a health care facility for their symptoms. For this study, Colab subsequently provided access to the anonymized data set of *Brazil Sem Corona*.

Ethical Considerations

Before filling out the questionnaire, participants were asked to agree with an informed consent form in the registration phase. The form described the study and the purpose of the project, and provided information about how the data could be used by third parties for research analysis purposes. Access to the data and study was approved by the Colab Institutional Management Board. All methods were carried out in accordance with guidelines and regulations, including but not limited to the *Lei Geral da Proteção de Dados – LGPD*, the official regulation on data privacy and protection valid in Brazil. All collected study data were anonymous and deidentified. No compensation was provided to the participants. The data used in this study can be made available through a formal request to the Colab team.

Data Collection

The platform was set up online on March 20, 2020, right after the World Health Organization officially declared COVID-19 as a public health issue with pandemic implications on March 11, 2020 [25]. During the period following this official statement, Brazilian newspapers and magazines extensively advertised the *Brazil Sem Corona* platform, leading to an increase in self-report submissions [26-28].

The increase in the number of participants allowed the creation of a database of symptom-based reports that could be used for the analysis. Each report was categorized into one of the following categories: (1) no symptoms, (2) light symptoms, (3) suspected COVID-19 case, (4) severe suspected COVID-19 case, and (5) confirmed case. In order to be categorized as a suspected case, the user must report fever together with at least one other symptom. If along with these symptoms either medication intake is reported or a health care facility is sought, the report is labeled as severe suspected case [29]. Only those users reporting a positive laboratory COVID-19 test result were categorized as confirmed cases. All the suspected, severe suspected, and confirmed cases were then treated as *possible COVID-19 cases*, while light and no symptom cases were treated as *negative cases*. The inclusion of not only confirmed cases but also suspected cases is one of the key differences with respect to case counts performed in TS settings.

For each city in which data were collected, we aggregated the reports submitted for each day and calculated the daily infection rate:

where *i* denotes the city and *t* denotes the day.

Data sparsity and fluctuations were addressed by applying a simple but powerful tool called LOESS to the PS data. It fits smooth lines to empirical data using a nonparametric approach [30]. For all 9 cities, the smoothed PS and TS infection rates were then compared.

Pearson Correlation Calculation

To measure the statistical relationship between the 2 time series, we calculated the Pearson correlation (PC_t) coefficient for the original time series. Additionally, we used a 7-day and 14-day lagged PS series to calculate PC_{t-7} and PC_{t-14}. Finally, we determined the coefficients over a reduced 4-month observation period during which public engagement and participation rates were the largest.

Finally, we looked at the proportion among self-reported cases of confirmed cases and the percentage of confirmed cases seeking health care assistance.

Forecasting Models

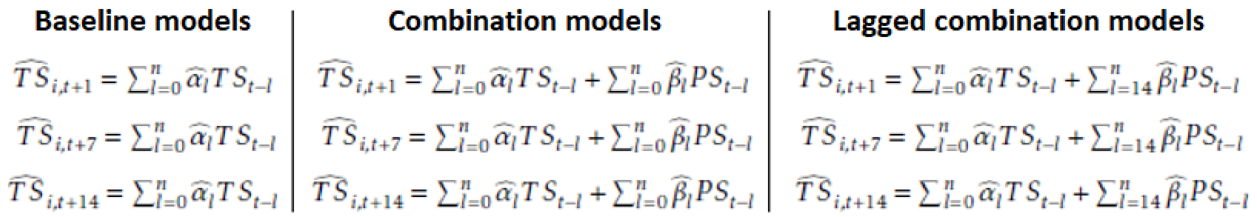
In order to further assess the added value of insights generated by PS systems, we used the PS data to inform 3 different forecasting models to predict incidence rates for COVID-19 with a time horizon varying between 1 and 3 weeks (Figure 1). TS data were used to generate a baseline model that represents the "ground truth" against which the forecasting models were compared.

Thus, the first model (further referred to as the *baseline model*) was a univariate model based on only TS data. The second and third models (referred to as the *combination model* and *lagged combination model*, respectively) were bivariate models integrating both TS and PS data. The third model also used a 14-day lag in the PS incidence rate series. For all 3 models, we used a linear autoregression function with *n* daily lagged components. For each city, the optimal number of independent variables was selected based on the Akaike Information Criterion (AIC) that estimates the prediction error and quality of statistical models. Thereby, the explained part of the variation was maximized while using only the lowest possible amount of time lag.

To evaluate the performance of the forecasting models, we calculated the root mean squared error (RMSE) and mean absolute error (MAE). The RMSE measures the difference between predicted and true values as follows:

The MAE measures the average of the absolute errors between predicted and true values as follows:

Figure 1. Model overview. stands for the estimation, whereas TS stands for the true value. By weighing n past components, 1-, 7-, and 14-day forecasts are estimated. The model parameters and are estimated based on a training subsample that contains the first 80% of the data. The out-of-sample forecasting accuracy is then calculated based on the remaining 20% of the data. Thus, the first 80% of the observation period is used to fit the models, while the remaining 20% is used to evaluate its performance. PS: participatory surveillance; TS: traditional surveillance.



Baseline models	Combination models	Lagged combination models
$\widehat{TS}_{i,t+1} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l}$	$\widehat{TS}_{i,t+1} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l} + \sum_{l=0}^{n} \widehat{\beta}_l PS_{t-l}$	$\widehat{TS}_{i,t+1} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l} + \sum_{l=14}^{n} \widehat{\beta}_l PS_{t-l}$
$\widehat{TS}_{i,t+7} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l}$	$\widehat{TS}_{i,t+7} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l} + \sum_{l=0}^{n} \widehat{\beta}_l PS_{t-l}$	$\widehat{TS}_{i,t+7} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l} + \sum_{l=14}^{n} \widehat{\beta}_l PS_{t-l}$
$\widehat{TS}_{i,t+14} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l}$	$\widehat{TS}_{i,t+14} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l} + \sum_{l=0}^{n} \widehat{\beta}_l PS_{t-l}$	$\widehat{TS}_{i,t+14} = \sum_{l=0}^{n} \widehat{\alpha}_l TS_{t-l} + \sum_{l=14}^{n} \widehat{\beta}_l PS_{t-l}$

Results

Reports were collected between March 20, 2020, and October 20, 2020. The analysis covered the whole 7-month observation period. Even though the platform was accessible all around Brazil, nearly 65% of the reports were submitted from 9 cities. Therefore, the study focused only on those cities with the largest number of submitted reports. People from Teresina, Caruaru, Santo Andre, Niteroi, Recife, Porto Alegre, Campinas, Sao Paulo, and Rio de Janeiro submitted an aggregated total of 83,005 reports (13,582 individuals).

To assess the importance of participation, we compared the capability of the PS data to mirror the infection rates reported by the TS systems implemented in the 9 cities. In Table 1, we present the 9 cities with the largest PS participation (the highest number of submitted reports across Brazil). As the population size varied between the cities, we additionally weighted the number of submitted reports by the population size displayed by the variable “number of reports by 100.” As participants could report several times throughout the observation period, there was a relatively large number of submitted reports compared with the number of participants. Even though São Paulo recorded a relatively large number of submitted reports, it was ranked only second to last due to its large population size. Teresina, Caruaru, and Santo André were ranked as the top 3 based on the variable “number of reports by 100.” The large participation in those 3 cities was not coincidental but rather driven by massive local media campaigns led by local officials [26-28]. Those social media campaigns seem to have impacted the participation behavior, underlining the government’s role. Additionally, we displayed the variable “share of zeros,” which captures the share of observation days on which no possible COVID-19 case was recorded on the platform within a city. This problem is often referred to as *zero inflation* (as the sample contains an excess of zeros) and is equivalent to censored data. Those numbers were negatively correlated with the level of engagement, as the cities at the bottom of the ranking displayed a striking share of zeros (as high as 96%).

As mentioned in the Methods section, every submitted report was categorized as either a negative COVID-19 case or suspected COVID-19 case based on the symptoms reported in the survey. In Figure 2, we show the TS and PS incidence rates over the entire 7-month observation period to compare differences in detected infection rates between the 2 surveillance methods. The 9 cities have been arranged according to the participation ranking (Table 1) from top left to bottom right. The top row displays the 3 cities with the largest community engagement (Teresina, Caruaru, and Santo André). Here, infections identified through PS and TS followed a relatively similar pattern, at least between March and August 2020. For most cities, the trend of detected cases diverged between the 2 methods toward the end of the observation period, which can be explained by the relatively low number of submitted reports over that interval. More details on the daily number of submitted reports can be found in Figure S1 in Multimedia Appendix 1. From the graphical analysis, it follows that the capability of the PS system to potentially point to the trend of infections seemed to decrease with low participation rates. This finding is supported by Table 2, which reports the Pearson correlation coefficients between the PS and TS time series within each city.

Table 2 confirms that the largest correlation between the 2 data sources was indeed noted for the cities with the largest participation rates (namely, Teresina, Caruaru, and Santo André). The Pearson correlation coefficients for those 3 cities were above 0.5, indicating a moderate to high positive correlation between the 2 data sources.

We next analyzed the participation behavior of individuals across all 9 cities. A significant share of 40%-50% of individuals participated only once throughout the observation period, while roughly 6%-13% were considered as frequent participants, with at least 15 submitted reports. Detailed numbers can be found in Table S1 in Multimedia Appendix 1.

In order to assess the effect of lag on the time series correlation, we also determined the Pearson correlation coefficients for the 3 key cities with the largest public engagement, using 7-day and 14-day lags in the PS data series. Table 3 shows the lagged Pearson correlations for the full observation period, while Table 4 presents the results for the reduced period from April 1 to July 31, 2020.

In Table 3, an increased correlation can only be found for Teresina, where the coefficient increased from 0.82 to 0.89 for the 14-day lag. For Caruaru and Santo André, the coefficients declined from 0.79 to 0.71 and from 0.56 to 0.45, respectively. However, the correlations for the 7-day lag remained at roughly the same level. While these results are ambiguous and rather weak, the picture appears clearer in Table 4.

In Table 4, the coefficients remain either constant or show an increase using 7-day and 14-day lags across all 3 cities. By removing the tails of the observation period, the focus was set on the months with the strongest engagement, as shown by the higher daily average of submitted reports in Table 4.

Table 5 presents the proportion of confirmed COVID-19 cases, proportion of cases for which health care assistance was sought, and proportion of cases involving medication intake among all submitted reports that were categorized as potential COVID-19 cases.

Table 1. Participation ranking for the 9 Brazilian cities under examination.

City	Number of reports	Number of reports per 100[a]	Number of participants	Share of zeros[b], %
Teresina	27,558	3.17	4449	7.0
Caruaru	10,279	2.81	1363	32.0
Santo Andre	14,207	1.97	2866	46.7
Niteroi	6757	1.31	897	62.9
Recife	3895	0.24	492	73.9
Porto Alegre	2779	0.19	336	96.1
Campinas	2006	0.17	307	92.9
Sao Paulo	12,453	0.10	2360	70.7
Rio de Janeiro	3071	0.05	734	74.6

[a] The ranking is based on the variable “number of reports per 100,” from the largest to the smallest. This variable shows the number of submitted reports per 100 inhabitants and is calculated as the number of submitted reports divided by population size times 100.

[b] The variable “share of zeros” shows the share of observation days on which no possible COVID-19 case was recorded within a city.

Figure 2. Daily PS and TS infection rates per city. The graphs are arranged according to the weighted participation from top left to bottom right. PS: participatory surveillance; TS: traditional surveillance.

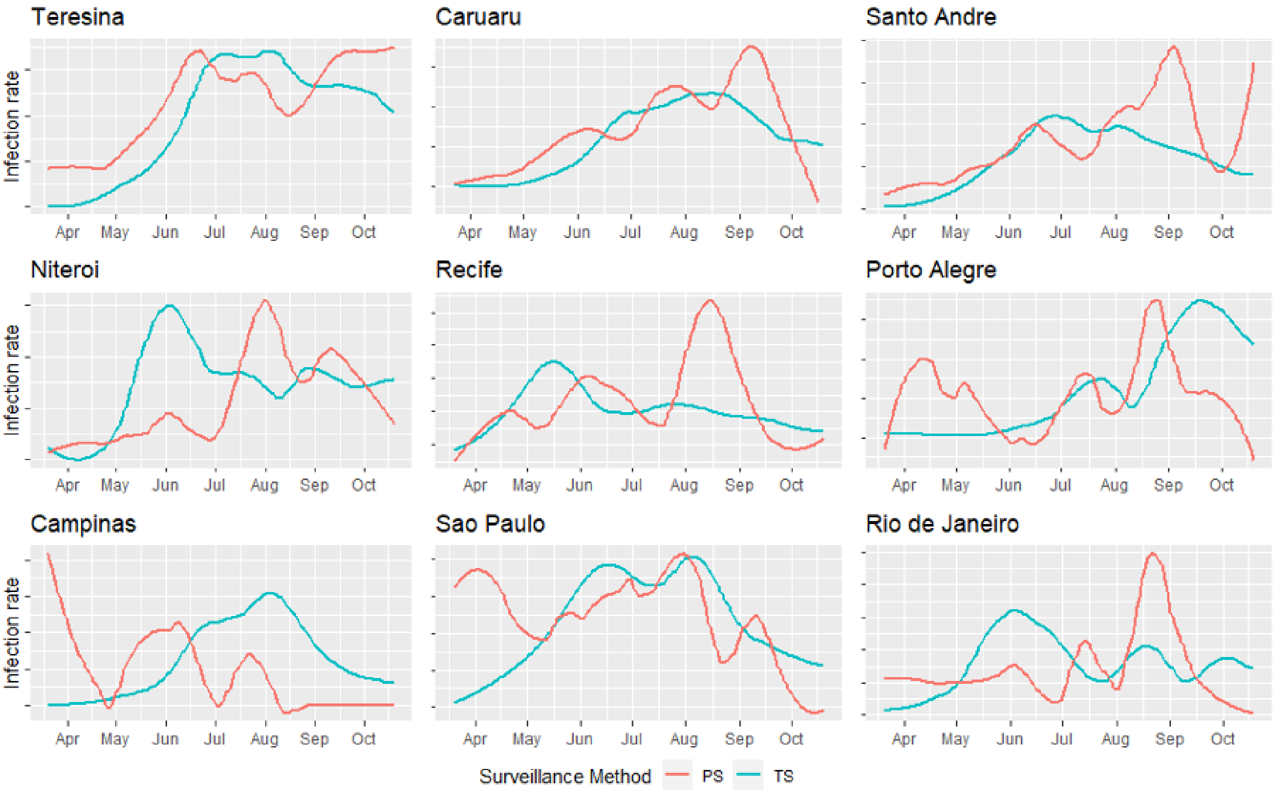


Table 2. Pearson correlations for each study city.

City[a]	Pearson correlation	*P* value
Teresina	0.82	<.001
Caruaru	0.79	<.001
Santo Andre	0.56	<.001
Niteroi	0.24	<.001
Recife	0.23	<.001
Porto Alegre	0.18	<.001
Campinas	−0.34	<.001
Sao Paulo	0.32	<.001
Rio de Janeiro	0.09	.19

[a] The cities are ranked according to population-weighted participation from the largest to the smallest.

Table 3. Lagged Pearson correlations for the full observation period (full 7-month period).

City	PC[a]	PC (7-day lag)	PC (14-day lag)	Average daily reports
Teresina	0.82	0.87	0.89	128
Caruaru	0.79	0.77	0.71	49
Santo Andre	0.56	0.56	0.45	66

[a] PC: Pearson correlation coefficient.

Table 4. Lagged Pearson correlations for the reduced observation period (April to July 2020).

City	PC[a]	PC (7-day lag)	PC (14-day lag)	Average daily reports
Teresina	0.86	0.91	0.93	193
Caruaru	0.85	0.85	0.88	76
Santo Andre	0.82	0.89	0.88	88

[a] PC: Pearson correlation coefficient.

Table 5. Characteristics of potential COVID-19 case reports.

Variable	Teresina (N=987), n (%)	Caruaru (N=356), n (%)	Santo Andre (N=189), n (%)
Confirmed cases	362 (36.7)	152 (42.7)	108 (57.1)
Cases seeking health care	270 (27.4)	112 (31.4)	65 (34.6)
Cases involving medication intake	495 (50.2)	185 (50.2)	110 (58.0)

Forecasting Models

For the 3 cities with an at least moderate (higher than 0.8) positive Pearson correlation coefficient, we compared the RMSEs and MAEs from the baseline model with those from the combination as well as lagged combination model (Tables 6-8). We show the results for 1-day, 7-day, and 14-day forecasts.

Even though improvements were only modest, there was a pattern that could be identified across all 3 cities. The combination model outperformed the baseline model for the 14-day forecast by up to 2.7%. The results for the 7-day forecast were ambiguous; while there were improvements of up to 4.1%, only 2 out of the 3 cities showed reduced RMSEs.

Table 6. Forecasting errors using different models for 1-day, 7-day, and 14-day forecasts in the city of Teresina.

Forecast period and model[a]	RMSE[b]	MAE[c]
1-day forecast		
Baseline model	0.0195	0.0133
Combination model	0.0198	0.0140
Lagged combination model	0.0193	0.0137
7-day forecast		
Baseline model	0.0410	0.0302
Combination model	0.0393	0.0281
Lagged combination model	0.0409	0.0310
14-day forecast		
Baseline model	0.0347	0.0245
Combination model	0.0338	0.0246
Lagged combination model	0.0340	0.0249

[a] The models used n=13 lagged components as independent variables.

[b] RMSE: root mean squared error.

[c] MAE: mean absolute error.

Table 7. Forecasting errors using different models for 1-day, 7-day, and 14-day forecasts in the city of Caruaru.

Forecast period and model[a]	RMSE[b]	MAE[c]
1-day forecast		
Baseline model	0.0214	0.0162
Combination model	0.0212	0.0160
Lagged combination model	0.0290	0.0220
7 -day forecast		
Baseline model	0.0316	0.0234
Combination model	0.0318	0.0237
Lagged combination model	0.0329	0.0244
14-day forecast		
Baseline model	0.0293	0.0190
Combination model	0.0288	0.0187
Lagged combination model	0.0292	0.0190

[a] The models used n=5 lagged components as independent variables.

[b] RMSE: root mean squared error.

[c] MAE: mean absolute error.

Table 8. Forecasting errors using different models for 1-day, 7-day, and 14-day forecasts in the city of Santo Andre.

Forecast period and model[a]	RMSE[b]	MAE[c]
1-day forecast		
Baseline model	0.0137	0.0094
Combination model	0.0236	0.0181
Lagged combination model	0.0203	0.0157
7 -day forecast		
Baseline model	0.0225	0.0169
Combination model	0.0219	0.0174
Lagged combination model	0.0321	0.0245
14-day forecast		
Baseline model	0.0292	0.0245
Combination model	0.0284	0.0236
Lagged combination model	0.0299	0.0251

[a] The models used n=14 lagged components as independent variables.

[b] RMSE: root mean squared error.

[c] MAE: mean absolute error.

Discussion

The goal of this study was to show how PS systems, already well established as a routine surveillance monitoring approach for ILI, can also complement and enhance TS systems for monitoring pandemic diseases such as COVID-19. The results showed that the validity of the approach strongly depends on the rate of participation among the general population.

We found that the PS infection rates from the 3 cities with the largest participation approximately mirrored the TS infection rates, even though the representativeness of the PS subpopulation was most likely biased in terms of sex and age. The 3 cities were able to engage a large group of individuals due to social media campaigns promoted by local officials and governments. The insights from the other cities were less reliable as the data were not able to represent the traditional infection rates due to low engagement and zero inflation. In Campinas, the city with the lowest total number of submitted reports, PS performed so poorly that its data displayed a *negative* correlation with TS data. With a total of only 2006 submitted reports over a period of 210 days (Table 1), a daily average of self-reports of just under 10 was observed, implying that the sample size was far too small. The same was true for the other cities at the low end of the ranking, with Pearson correlations below 0.25. The only exception was São Paulo, whose correlation was smaller than the correlations of the top 3 cities. Even though it was ranked as second to last regarding population-weighted participation, the total number of submitted reports for São Paulo was relatively large at 12,453 reports. This corresponded to an average of nearly 60 submissions per day, resulting in a Pearson correlation coefficient of 0.32. Overall, low participation is considered one of the main limitations for the capability of PS data to mirror traditional infection rates. Even though it is known for ILI that participation of 1%-2% is reasonable in epidemic periods, additional work is necessary to better estimate a reasonable cutoff for COVID-19.

A stronger correlation between the TS and lagged PS time series supports the theory of early trend detection, which is one goal pursued by PS. However, this was only conducted for 3 cities, which showed a strong positive PC_t above at least 0.5. It was found that toward the end of the observation period, only few people participated on the platform, such that around 80% of the reports were submitted within the period from April 1 to July 31, 2020.

The results from the forecasting analysis were also consistent with this conclusion. The errors for the 1-day ahead forecasts were generally smaller than the errors found in the 7-day and 14-day forecasts for all 3 cities. This is not surprising, as uncertainty rises with longer forecast horizons, and hence, the forecast accuracy of a model is lower. When looking at the 1-day forecasts, the baseline model seemed to outperform the combination model, at least for Teresina and Santo André. As mentioned before, uncertainty grows with longer horizons, which might explain the lack of value addition that comes from integrating PS data in a 1-day forecast. In Teresina, however, both the 7-day and 14-day forecasts from the combination and lagged combination models performed better than the baseline model, with slightly lower RMSEs. The combination model showed a 4.2% reduction in the RMSE for the 7-day horizon and a 2.6% reduction for the 14-day horizon, while the lagged combination model showed reductions of 0.2% and 2.1%, respectively. In Caruaru, the combination model performed slightly worse for the 7-day forecast relative to the baseline model, with a 0.5% increase in the RMSE, but an improvement was noted for the 14-day forecast, with a 1.7% reduction in the RMSE. Similar results were found for the lagged combination model, whereby the 7-day forecast showed a 4.1% increase in RMSE, while the 14-day forecast showed an improvement by 0.4%. The combination model applied on the data from Santo André again reduced the RMSEs for both the 7-day and 14-day forecasts by 2.8% and 2.7%, respectively. However, the lagged combination model was not able to improve the forecast for any of the applied horizons. This shows how complementing traditional disease surveillance systems may further increase the possibility for the early identification of outbreaks under the condition of sufficiently large participation. Slight improvements in the forecasting accuracy of up to 3% were identified for the models integrating data from both surveillance sources compared to the baseline model relying entirely on traditional data. Even if forecasting improvements are only weak, the detection of infections is improved as PS can cover an additional subset of the population that is overseen by the traditional system. The above findings contribute to a deeper understanding of the benefits of a complementary digital surveillance layer. They corroborate previous literature, emphasizing that the 2 approaches can be complements for timely health threat identification [3,14,16]. Even though some of the results indicate only small improvements in accuracy, the possibility of enhancing case detection through broader coverage cannot be neglected.

Some limitations related to the PS approach should be mentioned. Studies conducted in the United States and Western Europe suggest that the population subgroup captured by the *Brazil Sem Corona* PS system does not necessarily represent the general Brazilian population [31-33]. In previous PS data, female participants were significantly overrepresented compared to the general population. Besides that, age groups below 30 years and above 80 years were underrepresented. Moreover, the average participant most likely held a higher educational degree than the average population [34]. Furthermore, individuals living in bigger cities were more likely to participate, leading to clusters around more urban areas and information gaps in more remote regions. This is supported by the geocoordinates from the submitted reports on the *Brazil Sem Corona* platform, which indicate that a significant share of participants live in or around urban areas. Targeting the underrepresented population may prospectively improve the complementary benefit of the platform, particularly among those not seeking medical attention. Despite the biases found in the population covered by PS, studies highlight that PS systems that engage a sufficiently large group of participants can still adequately capture TS infection trends [33,35]. Ideally, PS tracks individuals throughout the season. Knowing that nearly half of the participants did not continue their engagement after the first report submission stresses the need for efforts to ensure more frequent participation in the future. Previous studies from the United States and Canada have found significant differences in participation across age groups, along with a 25% lower probability of frequent participation by women relative to men [33].

Another important aspect of complementing disease surveillance systems is the goal of early trend detection to identify disease outbreaks. Integrating multiple data sources not only aims at improving data insights and detecting more cases but also ideally leads to the detection of outbreaks at an earlier stage. The finding of greater coefficients in the lagged Pearson correlation indicates that the timeliness of the PS system helps in identifying slightly preceding trends. This supports the idea that a PS system, which engages a larger volunteer network, is more likely to depict infection trends, resulting in better data insights and, eventually, in earlier anticipation of trend changes. Recognizing outbreak patterns only slightly in advance might already have great benefits for health agencies when it comes to fighting pandemics. Improvements in 14-day forecasts allow health officials to respond more quickly and prioritize certain areas identified as more likely to suffer from rising infection numbers. This may be of particular importance for low-income countries or for regions that suffer from a considerable scarcity in health services. The PS system brings value by producing information that can be used to reduce uncertainty in allocation decisions. Furthermore, infection monitoring can be improved thanks to the geolocation information provided by PS data, which allows for a typically higher spatial resolution. Aside from preventing local transmissions, improvements in the surveillance system may result in externalities, such as lower infections in other regions [36].

We hypothesize that the severity of COVID-19 likely influenced individuals' motivation to engage in a voluntary surveillance system. Governments have adjusted their behavior during the ongoing pandemic and made significant efforts to improve the timeliness of traditional data reporting. Many countries, including Brazil, have overcome administrative burdens and were ultimately able to report daily new infections. As such, the timeliness that usually distinguishes a PS system from a TS system has most likely vanished. We see this as a reasonable explanation for why the magnitude of the results from the forecasting is weaker compared to the findings of previous studies conducted in the field of ILI [14,15]. Besides that, ILI tracking allows for data collection across several flu seasons, increasing the amount of available data and allowing for consistency checks across seasons, whereas the novelty of COVID-19 only allows for a single 7-month observation period in Brazil.

We emphasize the successful collaboration among governments, locals, and professionals when it comes to maximizing the use and gain of the *Brazil Sem Corona* platform. The 3 cities with extensive social media campaigns clearly demonstrated the capabilities of the PS system, under the condition of sufficient participation. Defining sufficient participation remains relevant for future research; so far, previous studies only highlighted the number of reports as a critical element and claimed the need to maintain sufficient coverage without setting a certain threshold [13,32,37]. Factors, such as population density, urbanization, and area size, most likely cause regional variations, making it significantly more difficult to determine a specific number that is generally valid.

The implementation of a proper surveillance system remains an important challenge in developing countries. While the PS system aims to address disparities in health outcomes, it currently reaches more urban regions and relatively more educated people. Therefore, voluntary crowdsourced information likely contains a population bias. Assessing the potential benefits for more rural communities remains open to future research in order to realize the system's full potential. Engaging and motivating a greater diversity of individuals remain key challenges that must be addressed in future PS platform implementations. Obstacles, such as lack of access to modern technologies, illiteracy, and simply lack of awareness of the benefits from participation, might hinder progress toward a more diverse reporting population.

Nearly half of the participants in Brazil submitted only a single report. This prevents monitoring the health status of individuals over a longer observation period, and as a result, reduces the insights that can be gained from PS systems. This stresses the need to evaluate incentives to induce more frequent participation. Additionally, in-depth research is needed to determine the reasons that lead to discontinuation of participation after the first submission. Addressing the above issues in future PS platform implementations is likely to lead to even greater benefits.

The unavailability of data on participants' sociodemographic characteristics in the *Brazil Sem Corona* platform prevents an important analysis of the representativeness of participants versus the general Brazilian population. Performing further tests and extending the method could confirm the validity of the results and deepen the insights. The forecasting models were kept rather simple on purpose, with only the TS and PS data as independent variables. Adding mobility data or contact tracing data, for instance, could improve accuracy, leading to a better understanding of transmission. There are further limitations to keep in mind, such as the impact of the criteria of categorization of a submitted report into either a *suspected COVID-19case* or a *negative case*, as this can influence the number of reports included in the case counts and the likelihood of false positives. Furthermore, the applied local regression ("LOESS") remains an approximation for smoothing the highly sparse data. Alternative smoothing techniques might produce different results, which could lead to divergent interpretations.

Public engagement remains a challenge. Future research needs to identify determinants of participation and proper incentives to induce larger coverage and higher diversity of participants. As the quality of data insights improves, PS benefits may further expand. Deeper insights allow for greater acceptance and credibility among governments and health authorities. Expanding collaboration among researchers, officials, and health authorities is needed to leverage data insights for timely response plans, which can ultimately lead to better health outcomes.

Acknowledgments

DP was supported by EU Horizon 2020 grants PANDEM-2 (883285) and EpiPose (101003688) and Horizon Europe VERDI (101045989) as well as grants from the Lagrange Project of the Institute for Scientific Interchange Foundation (ISI Foundation) funded by Fondazione Cassa di Risparmio di Torino (Fondazione CRT).

Data Availability

The data that support the findings of this study are available from Colab, but restrictions apply to the availability of these data. The data were used under license for this study, and thus, they are not publicly available. The data are however available from the authors upon reasonable request and with permission from Colab.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data description.
[PDF File (Adobe PDF File), 551 KB - publichealth_v9i1e44517_app1.pdf]

References

1. Ingram A. Outbreak, epidemic, pandemic: The politics of global health events. A review of Sara E. Davies' Containing Contagion. Singapore Journal of Tropical Geography 2021 May 17;42(2):350-352. [doi: 10.1111/sjtg.12367]
2. Mostashari F, Hartman J. Syndromic surveillance: a local perspective. J Urban Health 2003 Jun 01;80(2 Suppl 1):i1-i7 [FREE Full text] [doi: 10.1093/jurban/jtg042] [Medline: 12892064]
3. Salathé M. Digital Pharmacovigilance and Disease Surveillance: Combining Traditional and Big-Data Systems for Better Public Health. J Infect Dis 2016 Dec 01;214(suppl_4):S399-S403 [FREE Full text] [doi: 10.1093/infdis/jiw281] [Medline: 28830106]
4. Smith P, Hadler JL, Stanbury M, Rolfs RT, Hopkins RS, CSTE Surveillance Strategy Group. "Blueprint version 2.0": updating public health surveillance for the 21st century. J Public Health Manag Pract 2013;19(3):231-239. [doi: 10.1097/PHH.0b013e318262906e] [Medline: 22759985]
5. Abat C, Chaudet H, Rolain J, Colson P, Raoult D. Traditional and syndromic surveillance of infectious diseases and pathogens. Int J Infect Dis 2016 Jul;48:22-28 [FREE Full text] [doi: 10.1016/j.ijid.2016.04.021] [Medline: 27143522]
6. Event-based Surveillance. CDC. URL: https://www.cdc.gov/globalhealth/healthprotection/gddopscenter/how.html [accessed 2023-04-04]
7. Institute of Medicine. Managed Care Systems and Emerging Infections Challenges and Opportunities for Strengthening Surveillance, Research, and Prevention: Workshop Summary. Washington, DC: The National Academies Press; 2000.
8. Starnini M, Aleta A, Tizzoni M, Moreno Y. Impact of data accuracy on the evaluation of COVID-19 mitigation policies. Data & Policy 2021 Oct 28;3:e28. [doi: 10.1017/dap.2021.25]
9. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, et al. Chapter 53: Public Health Surveillance: A Tool for Targeting and Monitoring Interventions. In: Jamison DT, Breman JG, Measham AR, editors. Disease Control Priorities in Developing Countries. 2nd edition. Washington, DC: The International Bank for Reconstruction and Development/The World Bank; 2006.
10. Ramalingam B. Real-time Monitoring in Disease Outbreaks: Strengths, Weaknesses and Future Potential. Brighton, United Kingdom: Institute of Development Studies; 2016.
11. Peters D, Garg A, Bloom G, Walker DG, Brieger WR, Rahman MH. Poverty and access to health care in developing countries. Ann N Y Acad Sci 2008;1136:161-171. [doi: 10.1196/annals.1425.011] [Medline: 17954679]
12. Leal Neto O, Haenni S, Phuka J, Ozella L, Paolotti D, Cattuto C, et al. Combining Wearable Devices and Mobile Surveys to Study Child and Youth Development in Malawi: Implementation Study of a Multimodal Approach. JMIR Public Health Surveill 2021 Mar 05;7(3):e23154 [FREE Full text] [doi: 10.2196/23154] [Medline: 33536159]
13. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory Disease Surveillance: Engaging Communities Directly in Reporting, Monitoring, and Responding to Health Threats. JMIR Public Health Surveill 2017 Oct 11;3(4):e62 [FREE Full text] [doi: 10.2196/publichealth.7540] [Medline: 29021131]
14. Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, et al. Participatory Syndromic Surveillance of Influenza in Europe. J Infect Dis 2016 Dec 01;214(suppl_4):S386-S392. [doi: 10.1093/infdis/jiw280] [Medline: 28830105]
15. Perrotta D, Tizzoni M, Paolotti D. Using Participatory Web-based Surveillance Data to Improve Seasonal Influenza Forecasting in Italy. In: WWW '17: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: 26th International Conference on World Wide Web; April 3-7, 2017; Perth, Australia. [doi: 10.1145/3038912.3052670]
16. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. Clin Microbiol Infect 2014 Jan;20(1):17-21 [FREE Full text] [doi: 10.1111/1469-0691.12477] [Medline: 24350723]
17. Leal-Neto O, Santos F, Lee J, Albuquerque J, Souza W. Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning. Int J Med Inform 2020 Nov;143:104263 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104263] [Medline: 32877853]
18. Mahmud AS, Chowdhury S, Sojib KH, Chowdhury A, Quader MT, Paul S, et al. Participatory syndromic surveillance as a tool for tracking COVID-19 in Bangladesh. Epidemics 2021 Jun;35:100462 [FREE Full text] [doi: 10.1016/j.epidem.2021.100462] [Medline: 33887643]
19. Leal Neto O, Dimech GS, Libel M, de Souza WV, Cesse E, Smolinski M, et al. Saúde na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil. JMIR Public Health Surveill 2017 May 04;3(2):e26 [FREE Full text] [doi: 10.2196/publichealth.7313] [Medline: 28473308]

20. Leal Neto O, Cruz O, Albuquerque J, Nacarato de Sousa M, Smolinski M, Cesse P, et al. Participatory Surveillance Based on Crowdsourcing During the Rio 2016 Olympic Games Using the Guardians of Health Platform: Descriptive Study. JMIR Public Health Surveill 2020 Apr 07;6(2):e16119 [FREE Full text] [doi: 10.2196/16119] [Medline: 32254042]
21. Massuda A, Hone T, Leles FAG, de Castro MC, Atun R. The Brazilian health system at crossroads: progress, crisis and resilience. BMJ Glob Health 2018 Jul 03;3(4):e000829 [FREE Full text] [doi: 10.1136/bmjgh-2018-000829] [Medline: 29997906]
22. Cota W. Monitoring the number of COVID-19 cases and deaths in Brazil at municipal and federative units level. SciELO Preprint posted online May 7, 2020.
23. Cities and States. Instituto Brasileiro de Geografia e Estatística. URL: https://www.ibge.gov.br/en/cities-and-states.html?view=municipio [accessed 2023-04-05]
24. Brasil Sem Corona. URL: https://www.brasilsemcorona.com.br/ [accessed 2023-04-05]
25. Rolling updates on coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen [accessed 2023-04-05]
26. Apple e Google se unem para alertar sobre contágio de COVID-19 (Apple and Google team up to warn of COVID-19 contagion). CNN Brazil. URL: https://www.cnnbrasil.com.br/tecnologia/apple-e-google-se-unem-para-alertar-contagio-de-covid-19/ [accessed 2023-04-05]
27. Startups lançam aplicativo para monitorar riscos de infecção por coronavírus no (Startups launch application to monitor risks of coronavirus infection in Brazil). Época Negócios. URL: https://epocanegocios.globo.com/Empresa/noticia/2020/03/startups-lancam-aplicativo-para-monitorar-riscos-de-infeccao-por-coronavirus-no-brasil.html [accessed 2023-04-05]
28. Com colaboração de cidadãos, aplicativo Colab vai detectar risco de casos de coronavírus em todo o Brasil (With the collaboration of citizens, the Colab application will detect the risk of coronavirus cases throughout Brazil). Fórum de Saúde Digital. URL: https://forumsaudedigital.com.br/com-colaboracao-de-cidadaos-aplicativo-colab-vai-detectar-risco-de-casos-de-coronavirus-em-todo-o-brasil/ [accessed 2023-04-05]
29. Kennedy B, Fitipaldi H, Hammar U, Maziarz M, Tsereteli N, Oskolkov N, et al. App-based COVID-19 syndromic surveillance and prediction of hospital admissions in COVID Symptom Study Sweden. Nat Commun 2022 Apr 21;13(1):2110 [FREE Full text] [doi: 10.1038/s41467-022-29608-7] [Medline: 35449172]
30. Jacoby WG. Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral Studies 2000 Dec;19(4):577-613. [doi: 10.1016/s0261-3794(99)00028-1]
31. Cantarelli P, Debin M, Turbelin C, Poletto C, Blanchon T, Falchi A, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. BMC Public Health 2014 Sep 20;14(1):984 [FREE Full text] [doi: 10.1186/1471-2458-14-984] [Medline: 25240865]
32. Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T, et al. Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. PLoS One 2013 Sep 11;8(9):e73675 [FREE Full text] [doi: 10.1371/journal.pone.0073675] [Medline: 24040020]
33. Baltrusaitis K, Santillana M, Crawley AW, Chunara R, Smolinski M, Brownstein JS. Determinants of Participants' Follow-Up and Characterization of Representativeness in Flu Near You, A Participatory Disease Surveillance System. JMIR Public Health Surveill 2017 Apr 07;3(2):e18 [FREE Full text] [doi: 10.2196/publichealth.7304] [Medline: 28389417]
34. Bajardi P, Vespignani A, Funk S, Eames KT, Edmunds WJ, Turbelin C, et al. Determinants of follow-up participation in the Internet-based European influenza surveillance platform Influenzanet. J Med Internet Res 2014 Mar 10;16(3):e78 [FREE Full text] [doi: 10.2196/jmir.3010] [Medline: 24613818]
35. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. PLoS Comput Biol 2015 Oct 29;11(10):e1004513 [FREE Full text] [doi: 10.1371/journal.pcbi.1004513] [Medline: 26513245]
36. de Vries L, Koopmans M, Morton A, van Baal P. The economics of improving global infectious disease surveillance. BMJ Glob Health 2021 Sep 02;6(9):e006597 [FREE Full text] [doi: 10.1136/bmjgh-2021-006597] [Medline: 34475025]
37. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg Themes Epidemiol 2014 Jun 20;11(1):7 [FREE Full text] [doi: 10.1186/1742-7622-11-7] [Medline: 24991229]

Abbreviations

ILI: influenza-like illness
MAE: mean absolute error
PS: participatory surveillance
RMSE: root mean squared error
TS: traditional surveillance

Edited by A Mavragani, T Sanchez; submitted 22.11.22; peer-reviewed by A Husnayain, Y Wang, S Bezerra, P Susumpao; comments to author 09.01.23; revised version received 25.01.23; accepted 07.03.23; published 26.04.23.

Please cite as:
Wittwer S, Paolotti D, Lichand G, Leal Neto O
Participatory Surveillance for COVID-19 Trend Detection in Brazil: Cross-sectional Study
JMIR Public Health Surveill 2023;9:e44517
URL: https://publichealth.jmir.org/2023/1/e44517
doi:10.2196/44517
PMID:36888908

Original Paper

Hesitancy to Undergo SARS-CoV-2 Rapid Antigen Testing in China: Nationwide Cross-sectional Study

Zhen Lu[1], MSc; Leiwen Fu[1], MSc; Luoyao Yang[1], BD; Tian Tian[1], PhD; Yanxiao Gao[1], MSc; Xiaojun Meng[2], MSc; Huachun Zou[1], PhD

[1]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China
[2]Wuxi Municipal Center for Disease Control and Prevention, Wuxi, China

Corresponding Author:
Huachun Zou, PhD
School of Public Health (Shenzhen)
Sun Yat-sen University
No. 66, Gongchang Road, Guangming District
Shenzhen, 518107
China
Phone: 86 20 8733 5651
Email: zouhuachun@mail.sysu.edu.cn

Abstract

Background: SARS-CoV-2 rapid antigen testing (RAT) could be a useful supplementary test to diagnose larger numbers of acute asymptomatic infections and alleviate the limitations of polymerase chain reaction testing. However, hesitancy to undergo SARS-CoV-2 RAT may compromise its implementation.

Objective: We aimed to understand the prevalence and correlates of hesitancy to undergo RAT among adults not infected with SARS-CoV-2 in mainland China.

Methods: A nationwide cross-sectional survey on hesitancy to undergo SARS-CoV-2 RAT was conducted among adults not infected with SARS-CoV-2 in mainland China between April 29, 2022, and May 10, 2022. Participants completed an online questionnaire that covered the following COVID-19–related factors: sociodemographic characteristics, experiences of COVID-19 restrictions and knowledge of COVID-19, and attitude toward COVID-19 and its screening. This study was a secondary analysis of data from the survey. We compared the characteristics of participants by hesitancy to undergo SARS-CoV-2 RAT. Thereafter, logistic regression with a sparse group minimax concave penalty was used to identify correlates of hesitancy to undergo RAT.

Results: We recruited 8856 individuals with diverse demographic, socioeconomic, and geographic characteristics in China. Eventually, 5388 participants (valid response rate of 60.84%; 52.32% [2819/5388] women; median age 32 years) were included in the analysis. Among the 5388 participants, 687 (12.75%) expressed hesitancy to undergo RAT and 4701 (87.25%) were willing to undergo RAT. Notably, those who were from the central region (adjusted odds ratio [aOR] 1.815, 95% CI 1.441-2.278) and those who received COVID-19 information from traditional media (aOR 1.544, 95% CI 1.279-1.863) were significantly more likely to report hesitancy to undergo RAT (both $P<.001$). However, those who were women (aOR 0.720, 95% CI 0.599-0.864), were older (aOR 0.982, 95% CI 0.969-0.995), had postgraduate education (aOR 0.612, 95% CI 0.435-0.858), had children (<6 years old) and elders (>60 years old) in the family (aOR 0.685, 95% CI 0.510-0.911), had better knowledge about COVID-19 (aOR 0.942, 95% CI 0.916-0.970), and had mental health disorders (aOR 0.795, 95% CI 0.646-0.975) were less likely to report hesitancy to undergo RAT.

Conclusions: Hesitancy to undergo SARS-CoV-2 RAT was low among individuals who were not yet infected with SARS-CoV-2. Efforts should be made to improve the awareness and acceptance of RAT among men, younger adults, individuals with a lower education or salary, families without children and elders, and individuals who access COVID-19 information via traditional media. In a reopening world, our study could inform the development of contextualized mass screening strategies in general and the scale-up of RAT in particular, which remains an indispensable option in emergency preparedness.

(JMIR Public Health Surveill 2023;9:e43555) doi:10.2196/43555

KEYWORDS

COVID-19; SARS-CoV-2; vaccine; hesitancy; rapid antigen testing; China

Introduction

Multiple variants of SARS-CoV-2 have occurred [1], of which the Omicron variant has spread swiftly across the world and has become the progressive strain globally [2-5]. During the period of the COVID-19 pandemic, policy makers from various countries and regions have been focusing on efforts to mitigate the pandemic's devastating impacts on public health, economy, and social development. COVID-19 vaccination and nonpharmaceutical interventions (NPIs) are expected to contain the pandemic to a large extent. Across the world, as of August 11, 2022, there were 34 COVID-19 vaccines that had been authorized for use [6]. In addition, many countries implemented NPIs, including face mask wearing, hand washing, and physical distancing, in response to the ongoing COVID-19 pandemic [7].

When the genetic sequence of SARS-CoV-2 was published in January 2020 [8], 3 types of diagnostic tests for COVID-19 were implemented, including polymerase chain reaction (PCR) testing, rapid antigen testing (RAT), and serology testing. While PCR testing and RAT could be used to diagnose an acute infection involving COVID-19, serology testing provides indirect evidence of infection 1-2 weeks after the onset of symptoms [9] and thus was not applied to large-scale screening in the early phase. PCR testing is highly sensitive and specific, and requires the support of laboratory facilities and highly trained staff. Its results become available in the time range of less than 2 hours to up to 7 days [10]. However, a delay of several days in receiving testing results is unacceptable for mass screening, as it may lead to missed diagnosis of many cases, which would cause secondary transmission. RAT, on the other hand, is a useful supplementary testing approach that detects viral proteins (eg, spike and nucleocapsid proteins) to diagnose larger numbers of acute asymptomatic SARS-CoV-2 infections, and it can overcome the limitations of PCR testing in the early phase of infection in mass screening. Of note, RAT requires minimal effort of medical personnel training and is much faster (within 15 minutes) at providing results [11,12]. Hence, RAT could confirm or exclude infection with SARS-CoV-2 at the individual level for case management or self-isolation, and at the population level for large-scale screening and emergency responses.

To tackle sporadic COVID-19 outbreaks, testing for large-scale screening should be one of the backbones of the response [13]. China has adopted the general strategy of "guarding against imported cases and preventing a resurgence of the outbreak at home" and the policy of "dynamic zero-COVID," which include large-scale screening in a timely manner [14]. Furthermore, in March 2022, China implemented RAT to successfully identify infected individuals in the early phase in Shanghai. RAT has been used at universities and other congregate settings in the United States [12]. Germany, France, and the United Kingdom have already adopted RAT in their own strategies [15]. RAT is very useful in surveillance and emergency responses; however, the public may lack knowledge of RAT, that is, when and how to undergo RAT, or may be hesitant to undergo the procedure. In the case of outbreaks of COVID-19, especially those involving the Omicron variant, optimization of strategies of RAT for individuals in order to exert a dramatic effect on COVID-19 spread should be carefully considered. To address the transition from a pandemic to an endemic, evidence-based optimization of diagnostic strategies should be considered in the global agenda. Many studies have assessed reasons for COVID-19 vaccination hesitancy [16,17]; however, research characterizing hesitancy and the factors influencing hesitancy to undergo RAT for COVID-19 is scarce. Our study aimed to understand the prevalence and correlates of hesitancy to undergo RAT for COVID-19 among adults uninfected with SARS-CoV-2 in mainland China. For the COVID-19 pandemic and outbreaks of novel infectious diseases, the findings of this study could inform the future decision-making of the scale-up of RAT.

Methods

Study Participants

We conducted a nationwide online cross-sectional survey on hesitancy to undergo RAT among mainland Chinese adults uninfected with SARS-CoV-2 between April 29, 2022, and May 10, 2022. Participants were recruited through WeChat, which is a popular social media platform in China with over 1 billion monthly active users [18]. An online, self-administered, anonymous questionnaire was developed via Wenjuanxing, an online survey platform.

Eligibility criteria were as follows: (1) age of 18 years or older; (2) no previous infection with SARS-CoV-2; (3) residing in mainland China; (4) ability to complete the survey in Chinese; and (5) willingness to participate in the survey. To improve the quality of collected data, responses were deleted if (1) there were logically contradictory answers and (2) the time to complete the questionnaire was less than 2 minutes.

This survey based on a convenience sampling method [19,20] (assuming a 60% response rate) had access to respondents with diverse demographic, socioeconomic, and geographic characteristics in China. The questionnaire was comprised of closed-ended questions. Respondents were required to respond to all closed-ended questions so that there were no missing values for those questions. To avoid unmotivated, dishonest, inattentive, or duplicate responses, we tried to improve the quality of the questionnaire responses by deleting such responses. Moreover, in the process of data cleaning, we set rigorous standards for valid responses from participants.

Ethical Considerations

The nationwide cross-sectional survey was conducted with the approval of the Ethics Committee of the School of Public Health (Shenzhen), Sun Yat-sen University (approval number: SYSU-SPH2022020). Respondents were recruited without compensation. Respondents had to provide informed consent before proceeding to the questionnaire response page in the survey.

This study was a secondary analysis of data from the survey; therefore, this study received an exemption from the Ethics Committee of the School of Public Health (Shenzhen), Sun Yat-sen University. The study data were deidentified.

Measures

The questionnaire was adapted based on past literature on similar topics [16,21,22] and was revised by a panel of epidemiologists. A pilot study was conducted to help improve the questionnaire. All closed-ended questions had a single-answer or multiple-answer format, including “yes/no” scale, nominal and ordinal scale, and Likert scale questions. The questionnaire covered the following 3 parts: (1) sociodemographic characteristics; (2) experiences of COVID-19 restrictions and knowledge of COVID-19; and (3) attitude toward COVID-19 and its screening.

Sociodemographic Characteristics

The sociodemographic variables considered in this study were gender, age, ethnicity, socioeconomic status, residence, education level, marital status, age of other family members, occupation type, monthly salary during the COVID-19 pandemic, change in monthly salary during the COVID-19 pandemic, presence of chronic diseases (cardiovascular diseases, respiratory diseases, diabetes, cancers, liver diseases, and renal diseases, among others), and frailty status. Participants from 31 provinces, autonomous regions, and municipalities in mainland China were reclassified into the following 3 groups of socioeconomic status [23]: high (eastern region), medium (central region), and low (western region). The 5-item FRAIL scale (fatigue, resistance, ambulation, illness, and loss of weight) was used to identify frail persons at risk of developing disability, as well as decline in health functioning and mortality [24]. FRAIL scale scores range from 0 to 5 (0, best; 5, worst) and represent robust (0), prefrail (1-2), and frail (3-5) health statuses.

Experiences of COVID-19 Restrictions and Knowledge of COVID-19

Experiences of COVID-19 restrictions and knowledge of COVID-19 included experience of any NPIs, number of roommates, self-reported number of close contacts (physical contact distance of <1 meter) daily, COVID-19 vaccination status (COVID-19 vaccines include inactivated vaccines administered on a 2-dose schedule, recombinant subunit vaccines administered on a 3-dose schedule, or recombinant adenovirus type 5–vectored vaccines administered as a single dose), number of PCR tests for COVID-19 in the last month, frequency of attention to information about COVID-19, and information sources and health literacy about COVID-19. Health literacy about COVID-19 specifically refers to an individual’s ability to access, comprehend, and apply information related to COVID-19. Having adequate health literacy about COVID-19 is essential for individuals to take appropriate measures to prevent the spread of the virus, protect themselves and their loved ones, and make informed decisions about COVID-19–related health care options. Examples of health literacy about COVID-19 include understanding how the virus spreads, knowing the symptoms and when to seek medical attention, following public health guidelines, such as wearing masks and practicing physical distancing, and understanding the safety and efficacy of COVID-19 vaccines. In addition, we used a 5-point Likert scale (strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree) to assess items relating to health literacy about COVID-19.

Attitude Toward COVID-19 and Its Screening

Questions about attitude toward COVID-19 and its screening were asked, including perceived risk of infection of COVID-19, worries about COVID-19 infection, self-assessed mental health status, perceived burden or stress during the COVID-19 pandemic, and hesitancy about RAT. Mental health disorders related to COVID-19 refer to a range of psychological and emotional conditions that have arisen as a result of the COVID-19 pandemic. We asked respondents if they experienced mental health disorders related to the pandemic in the survey questionnaire. Respondents were asked whether they were unsure or unwilling to undergo SARS-CoV-2 RAT. The question we used in the questionnaire was as follows: “When SARS-CoV-2 RAT becomes available, will you take it?” The response options were “yes,” “not sure,” and “no.” Based on the answers, respondents were classified into the following 2 groups: willing and hesitant (including the answers of “not sure” and “no”).

Statistical Analysis

To tackle the question of the age of other family members with the multiple-answer format, the categorical variable was reclassified as follows: the age of other family members was reclassified by whether a participant had children (<6 years old) or elders (>60 years old). For other questions with the multiple-answer format, categorical variables were converted into binary dummy variables. Likert scale–type questions relating to health literacy about COVID-19 were treated as continuous variables. We calculated proportions or medians with interquartile ranges for all variables, stratified by hesitancy to undergo RAT. Chi-square tests (for proportions) and Wilcoxon rank sum tests (for medians) were used to assess statistically significant differences between the willing and hesitant groups.

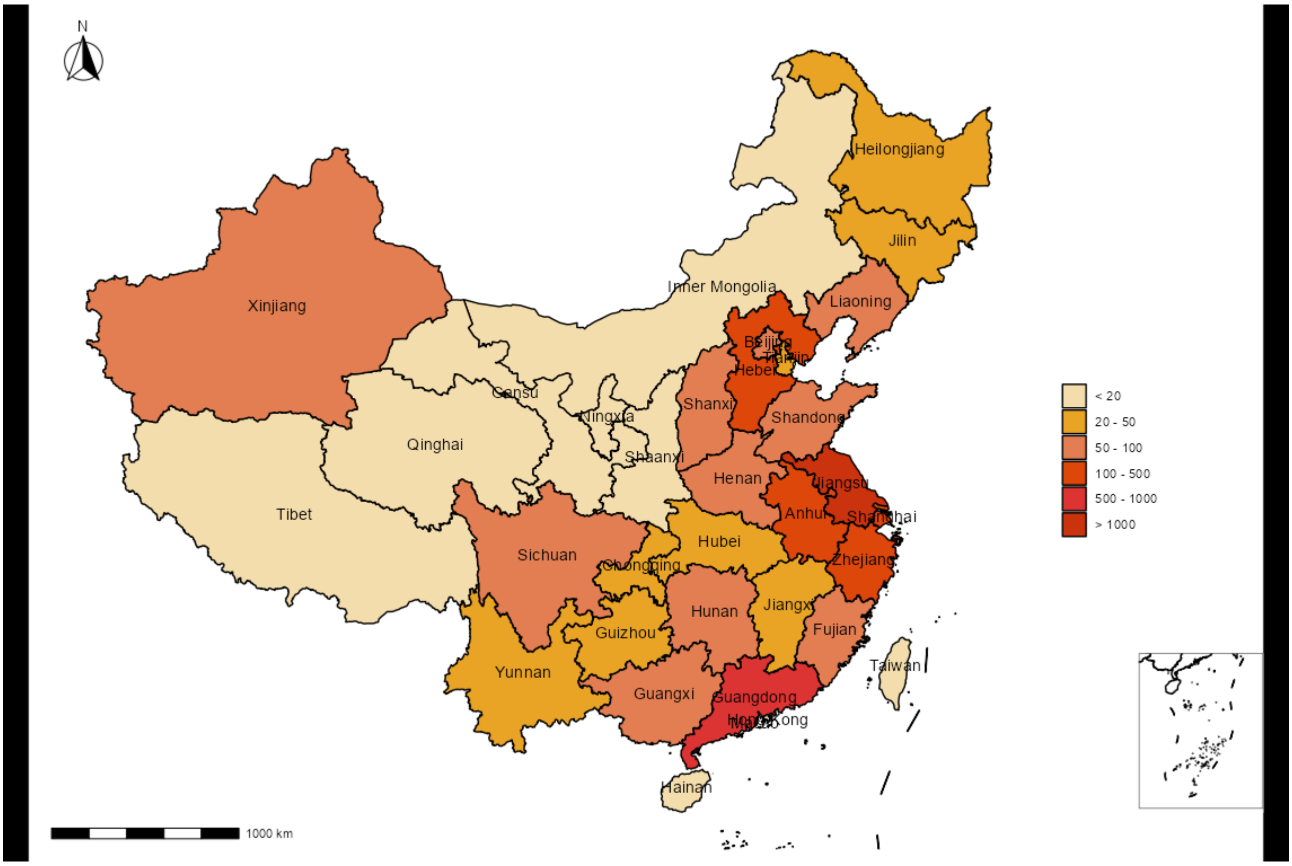
Logistic regression models were fitted to the data to assess correlates of hesitancy to undergo RAT. The outcome variable was whether a participant was willing to undergo RAT (coded as 0) or was hesitant (coded as 1). In addition, the questionnaire data had a group structure for all categorical variables, including 25 groups (ie, questions or covariates) that consisted of 44 variables (ie, the sum of [the number of answer choices for each binary dummy variable − 1], the number of other categorical variables, and the number of continuous variables). Due to the multicollinearity among covariates in regression analysis, we employed a variable selection step to determine the optimal covariate sets. In this study, to control multicollinearity effectively by assessing covariates with groups of variables as opposed to independent variables and to select the optimum group of variables by inducing sparsity in the number of covariates, we used the sparse group minimax concave penalty (MCP), which has been proven to outperform other penalties, such as the group LASSO (least absolute shrinkage and selection operator) [25]. Once the optimal penalty was estimated, the final logistic regression models were refitted without the penalty to debias the shrunken estimates in the coefficients due to the penalty [26]. All statistical analyses were performed using R software version 4.2.1 (R Core Team).

Results

Characteristics of the Study Participants

This survey had access to 8856 respondents with diverse demographic, socioeconomic, and geographic characteristics in China. In the process of data cleaning, we set rigorous standards for valid responses from participants. Ultimately, a total of 5388 participants (response rate of 60.84%; adults nationwide) was set as the fixed sample size. Figure 1 shows the geographical distribution (province level) of study participants in China. Among the participants, 2819 were women, and the proportions of women and men were 52.32% (2819/5388) and 47.68% (2569/5388), respectively. The median age of all participants was 32 years. In this survey, 88.33% (4759/5388) of participants were from urban areas, 80.31% (4327/5388) had an undergraduate education or above, 28.64% (1543/5388) were professional technicians, and 26.90% (1449/5388) had a monthly salary of at least 10,000 RMB (1 RMB=0.14 USD). With regard to health status, 86.06% (4637/5388) had absence of chronic diseases. Moreover 56.01% (3018/5388) had experienced any NPIs and 73.20% (3944/5388) were receiving or had completed the booster vaccination against COVID-19.

Figure 1. Geographical distribution of the study participants.



In this survey, among the 5388 participants, 687 (12.75%) expressed hesitancy to undergo RAT for COVID-19 and 4701 (87.25%) were willing to undergo RAT. The characteristics of study participants in the RAT willing and hesitant groups are shown in Tables 1-3. The 2 groups were roughly similar with respect to ethnicity, monthly salary, and COVID-19 vaccination status. Among 4701 participants in the willing group, 2529 (53.80%) were women, while among 687 participants in the hesitant group, 290 (42.21%) were women. The willing group was significantly older than the hesitant group (*P*<.001). Those in the willing group were more likely to be in the eastern region (3940/4701, 83.81% vs 479/687, 69.72%), be in urban areas (4186/4701, 89.04% vs 573/687, 83.41%), have a postgraduate education (901/4701, 19.17% vs 79/687, 11.50%), be married (2843/4701, 60.48% vs 387/687, 56.33%), and have elders in the family (2513/4701, 53.45% vs 248/687, 36.10%) than those in the hesitant group. In addition, the other variables (except ethnicity, monthly salary, and COVID-19 vaccination status) were correlated with willingness to undergo RAT.

Table 1. Sociodemographic characteristics of study participants by hesitancy to undergo rapid antigen testing.

Variable	Overall (N=5388)	Willing group (n=4701)	Hesitant group (n=687)	*P* value
Gender, n (%)				<.001
Men	2569 (47.68)	2172 (46.20)	397 (57.79)	
Women	2819 (52.32)	2529 (53.80)	290 (42.21)	
Age, median (IQR)	32 (26-40)	32 (26-40)	29 (23-36)	<.001
Ethnicity, n (%)				.64
Han	5163 (95.82)	4507 (95.87)	656 (95.49)	
Others	225 (4.18)	194 (4.13)	31 (4.51)	
Socioeconomic status[a], n (%)				<.001
High (eastern)	4419 (82.02)	3940 (83.81)	479 (69.72)	
Medium (central)	654 (12.14)	487 (10.36)	167 (24.31)	
Low (western)	315 (5.85)	274 (5.83)	41 (5.97)	
Residence, n (%)				<.001
Urban	4759 (88.33)	4186 (89.04)	573 (83.41)	
Rural	629 (11.67)	515 (10.96)	114 (16.59)	
Education, n (%)				<.001
High school or below	1061 (19.69)	891 (18.95)	170 (24.75)	
Bachelor’s degree	3347 (62.12)	2909 (61.88)	438 (63.76)	
Master’s degree or above	980 (18.19)	901 (19.17)	79 (11.50)	
Marital status, n (%)				.002
Single	1982 (36.79)	1694 (36.03)	288 (41.92)	
Married	3230 (59.95)	2843 (60.48)	387 (56.33)	
Divorced/widowed	176 (3.27)	164 (3.49)	12 (1.75)	
Age of other family members, n (%)				<.001
Not having children (<6 years old) or elders (>60 years old)	1952 (36.23)	1640 (34.89)	312 (45.41)	
Having children (<6 years old)	675 (12.53)	548 (11.66)	127 (18.49)	
Having elders (>60 years old)	1919 (35.62)	1748 (37.18)	171 (24.89)	
Having children (<6 years old) and elders (>60 years old)	842 (15.63)	765 (16.27)	77 (11.21)	
Occupation, n (%)				<.001
Company employee	789 (14.64)	652 (13.87)	137 (19.94)	
Health care professional	278 (5.16)	251 (5.34)	27 (3.93)	
Professional technician	1543 (28.64)	1350 (28.72)	193 (28.09)	
Student	588 (10.91)	508 (10.81)	80 (11.64)	
Public servant	781 (14.50)	713 (15.17)	68 (9.90)	
Self-employed	291 (5.40)	260 (5.53)	31 (4.51)	
Teacher	264 (4.90)	211 (4.49)	53 (7.71)	
Service worker	151 (2.80)	138 (2.94)	13 (1.89)	
Unemployed	194 (3.60)	161 (3.42)	33 (4.80)	
Retiree	202 (3.75)	184 (3.91)	18 (2.62)	
Others	307 (5.70)	273 (5.81)	34 (4.95)	
Monthly salary (RMB)[b], n (%)				.17
<5000	1174 (21.79)	1030 (21.91)	144 (20.96)	

Variable	Overall (N=5388)	Willing group (n=4701)	Hesitant group (n=687)	*P* value
5000-10,000	1849 (34.32)	1611 (34.27)	238 (34.64)	
10,001-15,000	739 (13.72)	647 (13.76)	92 (13.39)	
15,001-20,000	314 (5.83)	281 (5.98)	33 (4.80)	
>20,000	396 (7.35)	354 (7.53)	42 (6.11)	
No fixed salary	916 (17.00)	778 (16.55)	138 (20.09)	
Change in monthly salary, n (%)				<.001
No change	2602 (48.29)	2318 (49.31)	284 (41.34)	
Decrease by 10% or less	696 (12.92)	581 (12.36)	115 (16.74)	
Decrease by more than 10%	1976 (36.67)	1708 (36.33)	268 (39.01)	
Increase by 10% or less	56 (1.04)	44 (0.94)	12 (1.75)	
Increase by more than 10%	58 (1.08)	50 (1.06)	8 (1.16)	
Presence of chronic diseases (multiple answers), n (%)				
None	4637 (86.06)	4061 (86.39)	576 (83.84)	.07
Cardiovascular diseases	393 (7.29)	327 (6.96)	66 (9.61)	.01
Respiratory diseases	155 (2.88)	117 (2.49)	38 (5.53)	<.001
Diabetes	161 (2.99)	128 (2.72)	33 (4.80)	.003
Cancer	84 (1.56)	63 (1.34)	21 (3.06)	.001
Liver diseases	45 (0.84)	32 (0.68)	13 (1.89)	.001
Renal diseases	43 (0.80)	35 (0.74)	8 (1.16)	.25
Other diseases	133 (2.47)	126 (2.68)	7 (1.02)	.009
Frailty status, n (%)				<.001
Robust	2770 (51.41)	2458 (52.29)	312 (45.41)	
Prefrail	2462 (45.69)	2134 (45.39)	328 (47.74)	
Frail	156 (2.90)	109 (2.32)	47 (6.84)	

[a] The eastern region includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The central region includes Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, and Hunan. The western region includes Sichuan, Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang, Inner Mongolia, and Guangxi.

[b] A currency exchange rate of 1 RMB=0.14 USD is applicable.

Table 2. Experiences of COVID-19 restrictions and knowledge of COVID-19 by hesitancy to undergo rapid antigen testing.

Variable	Overall (N=5388)	Willing group (n=4701)	Hesitant group (n=687)	*P* value
Experience of any NPIs[a], n (%)				<.001
No	2370 (43.99)	2155 (45.84)	215 (31.30)	
Yes	3018 (56.01)	2546 (54.16)	472 (68.70)	
Number of roommates, n (%)				.02
0	664 (12.32)	573 (12.19)	91 (13.25)	
1	1218 (22.61)	1046 (22.25)	172 (25.04)	
2-5	3371 (62.56)	2973 (63.24)	398 (57.93)	
>5	135 (2.51)	109 (2.32)	26 (3.78)	
Self-reported number of close contacts daily[b], n (%)				<.001
0-5	2113 (39.22)	1813 (38.57)	300 (43.67)	
6-10	1358 (25.20)	1166 (24.80)	192 (27.95)	
11-20	782 (14.51)	684 (14.55)	98 (14.26)	
21-30	313 (5.81)	288 (6.13)	25 (3.64)	
>30	822 (15.26)	750 (15.95)	72 (10.48)	
COVID-19 vaccination status[c], n (%)				.10
Basic vaccination phase	1246 (23.13)	1066 (22.68)	180 (26.20)	
Booster vaccination phase	3944 (73.20)	3458 (73.56)	486 (70.74)	
No	198 (3.67)	177 (3.77)	21 (3.06)	
Number of PCR[d] tests for COVID-19 in the last month, n (%)				<.001
0-5	1155 (21.44)	855 (18.19)	300 (43.67)	
6-10	1691 (31.38)	1504 (31.99)	187 (27.22)	
11-20	1657 (30.75)	1528 (32.50)	129 (18.78)	
>20	885 (16.43)	814 (17.32)	71 (10.33)	
Frequency of attention to information about COVID-19, n (%)				<.001
Often	4641 (86.14)	4089 (86.98)	552 (80.35)	
Sometimes	622 (11.54)	508 (10.81)	114 (16.59)	
Rarely	108 (2.00)	92 (1.96)	16 (2.33)	
Never	17 (0.32)	12 (0.26)	5 (0.73)	
Information sources (multiple answers), n (%)				
Not interested in any information	17 (0.32)	12 (0.26)	5 (0.73)	.04
Internet media	4955 (91.96)	4359 (92.72)	596 (86.75)	<.001
Local authorities	3410 (63.29)	2985 (63.50)	425 (61.86)	.41
Traditional media	2409 (44.71)	2016 (42.88)	393 (57.21)	<.001
Friends or family members	2069 (38.40)	1804 (38.37)	265 (38.57)	.92
Others	439 (8.15)	397 (8.45)	42 (6.11)	.04
Health literacy about COVID-19, median (IQR)	11.0 (7.0-13.0)	12.0 (8.0-13.0)	10.0 (6.0-12.0)	<.001

[a] NPI: nonpharmaceutical intervention.

[b] Close contacts refer to contacts with a physical contact distance of <1 meter.

[c] COVID-19 vaccines include inactivated vaccines administered on a 2-dose schedule, recombinant subunit vaccines administered on a 3-dose schedule, or recombinant adenovirus type 5–vectored vaccines administered as a single dose.

[d] PCR: polymerase chain reaction.

Table 3. Attitude toward COVID-19 and its screening by hesitancy to undergo rapid antigen testing.

Variable	Overall (N=5388)	Willing group (n=4701)	Hesitant group (n=687)	*P* value
Perceived risk of infection of COVID-19, n (%)				.005
Low	1684 (31.25)	1440 (30.63)	244 (35.52)	
Medium	2700 (50.11)	2359 (50.18)	341 (49.64)	
High	1004 (18.63)	902 (19.19)	102 (14.85)	
Worry about COVID-19 if infected (multiple answers), n (%)				
Worry about health	4098 (76.06)	3583 (76.22)	515 (74.96)	.47
Worry about discrimination	1955 (36.28)	1696 (36.08)	259 (37.70)	.41
Worry about others	682 (12.66)	634 (13.49)	48 (6.99)	<.001
No worry	574 (10.65)	489 (10.40)	85 (12.37)	.12
Self-assessed mental health status, n (%)				<.001
General	1822 (33.82)	1623 (34.52)	199 (28.97)	
Good	334 (6.20)	276 (5.87)	58 (8.44)	
Better	271 (5.03)	210 (4.47)	61 (8.88)	
Relatively poor	2595 (48.16)	2291 (48.73)	304 (44.25)	
Poor	366 (6.79)	301 (6.40)	65 (9.46)	
Perceived burden and stress (multiple answers), n (%)				
None	242 (4.49)	217 (4.62)	25 (3.64)	.25
Postal and delivery services	2628 (48.78)	2295 (48.82)	333 (48.47)	.87
Nationwide travel restrictions	2643 (49.05)	2335 (49.67)	308 (44.83)	.02
Financial insecurity	2267 (42.07)	1971 (41.93)	296 (43.09)	.57
Medical services	1410 (26.17)	1209 (25.72)	201 (29.26)	.05
Mental health disorders	1999 (37.10)	1810 (38.50)	189 (27.51)	<.001
Burden of work	2267 (42.07)	1971 (41.93)	296 (43.09)	.57
Social isolation	1818 (33.74)	1599 (34.01)	219 (31.88)	.27
Others	164 (3.04)	156 (3.32)	8 (1.16)	.002

Correlates of Hesitancy to Undergo RAT

In total, 21 groups that consisted of 40 variables selected by the MCP-penalized group regression were included in the final model. Only 13 groups were significantly associated with developing hesitancy to undergo RAT for COVID-19 (Table 4). The forest plot given in Figure 2 reveals the difference in the adjusted odds ratios (aORs) of the risk factors of hesitancy to undergo SARS-CoV-2 RAT between the willing group and hesitant group. Actually, Figure 2 shows the forest plot summarizing only the significant variables that were associated with hesitancy to undergo RAT. The complete table containing all variables left in the final model is presented in Multimedia Appendix 1.

Six sociodemographic variables were selected to be included in the model. Those who were women (aOR 0.720, 95% CI 0.599-0.864), were older (aOR 0.982, 95% CI 0.969-0.995), had a postgraduate education (aOR 0.612, 95% CI 0.435-0.858), and had children and elders in the family (aOR 0.685, 95% CI 0.510-0.911) showed willingness to undergo RAT. However, those who were from the central region (aOR 1.815, 95% CI 1.441-2.278) and had a decrease in the monthly salary by more than 10% (aOR 1.449, 95% CI 1.123-1.875) were significantly more likely to report hesitancy to undergo RAT for COVID-19 (*P*<.001 and *P*=.005, respectively). Five variables concerning experiences of COVID-19 restrictions and knowledge of COVID-19 were also selected to be included in the model. Those who had 21-30 close contacts daily (aOR 0.621, 95% CI 0.386-0.960), had >20 PCR tests for COVID-19 (aOR 0.279, 95% CI 0.205-0.377), and had more knowledge about COVID-19 (aOR 0.942, 95% CI 0.916-0.970) showed greater ease in accepting RAT. In addition, those who experienced any NPIs (aOR 1.613, 95% CI 1.322-1.973) and received COVID-19 information from traditional media (aOR 1.544, 95% CI 1.279-1.863) were significantly more likely to report hesitancy to undergo RAT for COVID-19 (both *P*<.001). Two psychological and attitudinal variables were also found to be significantly associated with developing hesitancy to undergo RAT for COVID-19. Those who had a better self-assessed mental health status (aOR 1.943, 95% CI 1.352-2.765) were significantly more likely to report hesitancy to undergo RAT for COVID-19 (*P*<.001), while those who had mental health disorders (aOR 0.795, 95% CI 0.646-0.975) showed higher wiliness to undergo RAT.

Table 4. Correlates of hesitancy to undergo rapid antigen testing.

Variable	aOR[a] (95% CI)	*P* value
Sociodemographic characteristics		
Gender		
Men	Reference	N/A[b]
Women	0.720 (0.599-0.864)	<.001
Age	0.982 (0.969-0.995)	.006
Socioeconomic status[c]		
High (eastern)	Reference	N/A
Medium (central)	1.815 (1.441-2.278)	<.001
Low (western)	0.787 (0.531-1.140)	.22
Education		
High school or below	Reference	N/A
Bachelor’s degree	0.878 (0.699-1.107)	.27
Master’s degree or above	0.612 (0.435-0.858)	.005
Age of other family members		
Not having children (<6 years old) or elders (>60 years old)	Reference	N/A
Having children (<6 years old)	1.149 (0.879-1.497)	.31
Having elders (>60 years old)	0.659 (0.530-0.817)	<.001
Having children (<6 years old) and elders (>60 years old)	0.685 (0.510-0.911)	.01
Change in monthly salary		
No change	Reference	N/A
Decrease by 10% or less	1.577 (1.172-2.117)	.002
Decrease by more than 10%	1.449 (1.123-1.875)	.005
Increase by 10% or less	1.234 (0.554-2.575)	.59
Increase by more than 10%	1.546 (0.634-3.350)	.30
Experiences of COVID-19 restrictions and knowledge of COVID-19		
Experience of any NPIs[d]		
No	Reference	N/A
Yes	1.613 (1.322-1.973)	<.001
Self-reported number of close contacts daily[e]		
0-5	Reference	N/A
6-10	0.981 (0.790-1.217)	.86
11-20	0.989 (0.754-1.290)	.94
21-30	0.621 (0.386-0.960)	.04
>30	0.755 (0.558-1.010)	.06
Number of PCR[f] tests for COVID-19 in the last month		
0-5	Reference	N/A
6-10	0.370 (0.296-0.463)	<.001
11-20	0.262 (0.203-0.336)	<.001
>20	0.279 (0.205-0.377)	<.001
Information sources (multiple answers)		
Not interested in any information	Reference	N/A

Variable	aOR[a] (95% CI)	*P* value
Internet media	0.780 (0.577-1.064)	.11
Local authorities	0.852 (0.699-1.040)	.11
Traditional media	1.544 (1.279-1.863)	<.001
Friends or family members	1.073 (0.881-1.304)	.48
Others	0.863 (0.597-1.219)	.42
Health literacy about COVID-19	0.942 (0.916-0.970)	<.001
Attitude toward COVID-19 and its screening		
Self-assessed mental health status		
General	Reference	N/A
Good	1.583 (1.109-2.232)	.01
Better	1.943 (1.352-2.765)	<.001
Relatively poor	1.114 (0.897-1.387)	.33
Poor	1.337 (0.930-1.904)	.11
Perceived burden and stress (multiple answers)		
None	Reference	N/A
Postal and delivery services	1.058 (0.876-1.277)	.56
Nationwide travel restrictions	0.947 (0.779-1.150)	.58
Financial insecurity	0.875 (0.716-1.069)	.19
Medical services	1.104 (0.898-1.353)	.34
Mental health disorders	0.795 (0.646-0.975)	.03
Burden of work	1.142 (0.935-1.392)	.19
Social isolation	1.149 (0.934-1.413)	.19
Others	0.614 (0.268-1.221)	.20

[a] aOR: adjusted odds ratio.

[b] N/A: not applicable.

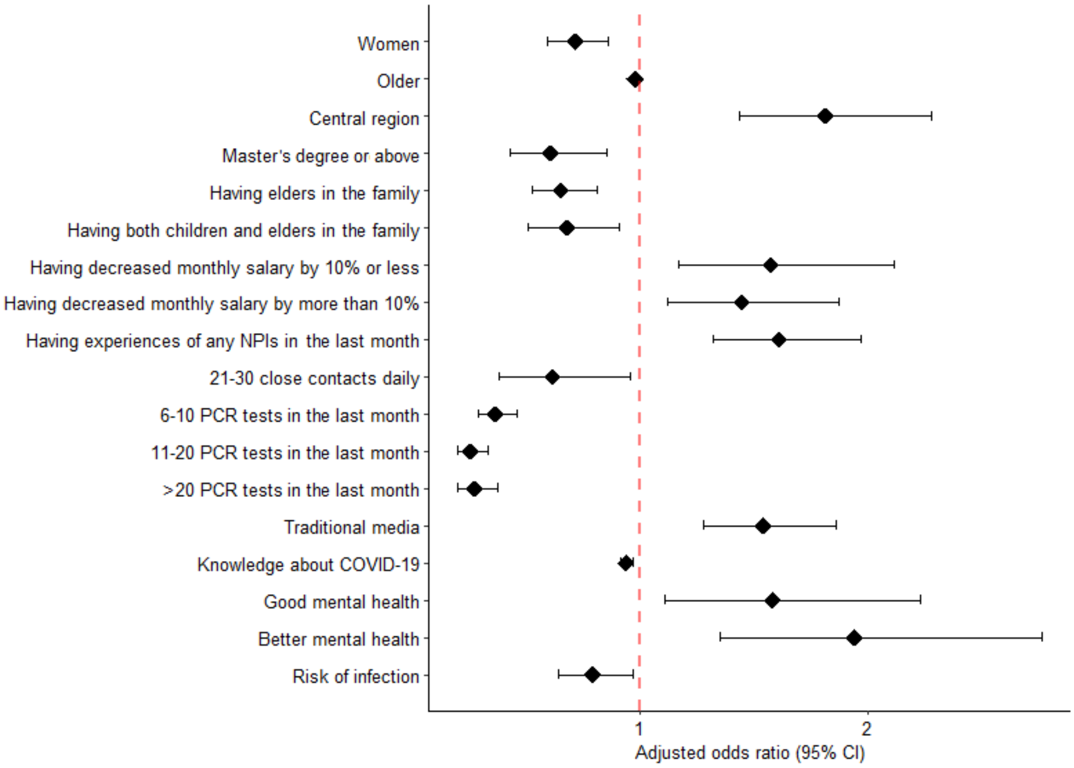
[c] The eastern region includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The central region includes Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, and Hunan. The western region includes Sichuan, Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang, Inner Mongolia, and Guangxi.

[d] NPI: nonpharmaceutical intervention.

[e] Close contacts refer to contacts with a physical contact distance of <1 meter.

[f] PCR: polymerase chain reaction.

Figure 2. Forest plot summarizing only the significant variables that were associated with hesitancy to undergo SARS-CoV-2 rapid antigen testing. The central region includes Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, and Hunan. Close contacts refer to contacts with a physical contact distance of <1 meter. NPI: nonpharmaceutical intervention; PCR: polymerase chain reaction.



Discussion

In this nationwide cross-sectional survey, we found that 12.75% (687/5388) of the participants were hesitant to undergo RAT for COVID-19, while 87.25% (4701/5388) were willing to undergo RAT. Several characteristics were found to be associated with hesitancy to undergo RAT, which could help identify subgroups to improve the awareness and acceptance of RAT.

Hesitancy about COVID-19 in China was low but may compromise the scale-up of RAT. A cross-sectional survey involving residents of Greece and Cyprus found that 79% (196/248) reported willingness to self-test in fighting the COVID-19 pandemic [27]. In another representative survey of the Greek population conducted in the first half of August 2021, two-thirds of responders characterized COVID-19 self-tests as unreliable and two-fifths believed they were dangerous [28]. An Indonesia study found that, if rapid COVID-19 antigen self-tests were available, 62.70% (395/630) would use them when necessary [29]. In addition to positive attitudes toward COVID-19 vaccination [30], this relatively high percentage of willingness to undergo RAT in China might reflect public trust in the government, and health authorities in China spared no effort to offer accessible health care services to increase safety and the perceptions of safety for the public. As the COVID-19 pandemic continues and in the postpandemic era, it would be important to monitor and promote public willingness toward RAT further in order to identify infected individuals in the early phase and keep COVID-19 incidence low, as well as improve the coverage of routine RAT for COVID-19.

Hesitancy to undergo RAT was found to be associated with being male, which is similar to the epidemiologic characteristic in the Greek population [28]. In addition, younger individuals reported low willingness to undergo RAT compared to their older counterparts. This association was also found in a representative survey of 17 countries on COVID-19 vaccine hesitancy [31]. There could be some possible explanations. Younger individuals may perceive themselves as being at lower risk of experiencing a severe illness associated with COVID-19, as the disease tends to affect older people more severely. This perception of lower risk could lead them to believe that RAT is unnecessary. In addition, younger individuals may have lower incomes or be more financially insecure, and the cost of RAT may be a barrier for them. They may also be more concerned about missing work or school if they test positive and need to isolate. Moreover, we found that socioeconomic status and education might play a role in determining hesitancy to undergo RAT. Areas further to the east in China have a better socioeconomic status and higher education level, and a better socioeconomic status and higher education level have been previously found to be associated with better health literacy and lower vaccine hesitancy [32,33]. This might have resulted in a situation where those who were from the central region and had a high school education or lower were significantly more likely to report hesitancy to undergo RAT, similar to vaccination for COVID-19. A change in the monthly salary during the COVID-19 lockdown was much more likely to be associated with hesitancy to undergo SARS-CoV-2 RAT compared with the current monthly salary. This could indicate that those who experience job loss, reduced income, and increased debt during the COVID-19 pandemic should receive much more attention to improve the awareness and acceptance of SARS-CoV-2 RAT. The pandemic has also highlighted existing social inequalities, with communities, who experienced heavier financial burden, being disproportionately affected by the virus. A study also found that individuals with lower salary were more likely to have higher similar mistrust about vaccines [33]. In addition, we found that not having children or elders in the family was correlated with hesitancy to undergo RAT. This may be due to the desire to protect children and elders in the family and prevent chains of intrahousehold transmission. Those having elders or having both children and elders in the family showed willingness to undergo RAT.

We identified 5 variables concerning experiences of COVID-19 restrictions and knowledge of COVID-19, which were significantly associated with hesitancy to undergo RAT. Those who experienced any NPIs and had 0-5 close contacts daily were more likely to report hesitancy. A prior study has found that more social contacts and large households increased the risk of SARS-CoV-2 infection [34,35]. The finding for the experience of any NPIs or the self-reported number of close contacts daily might be because the subgroup believed there was a low risk of SARS-CoV-2 infection and had stronger faith in the health condition, compared with those who did not experience NPIs for COVID-19 or had more close contacts in daily life. Moreover, those who had 0-5 PCR tests for COVID-19 in the last month were more likely to report hesitancy. PCR testing remains the gold standard in all countries, with the majority also employing it for diagnostic and surveillance purposes [36]. Fewer PCR tests might deepen hesitancy to undergo RAT. Furthermore, those who received COVID-19 information from traditional media were more likely to report hesitancy. Over the course of fighting the COVID-19 pandemic, health-related information and knowledge have played large roles in shaping vaccine hesitancy [37]. Increased likelihood of vaccine acceptance was seen among those who obtained COVID-19 information from digital media or nonhealth-related sources like influencers [38], and this likely was related to hesitancy to undergo RAT among participants in our survey. Traditional media like newspapers are not prone to misinformation, but do not have the same tremendous and popular health information available as other information sources like internet media. Moreover, inadequate knowledge has been found to be associated with vaccine hesitancy [39], which might be in line with our finding that those who had less knowledge about COVID-19 were significantly more likely to report hesitancy to undergo RAT.

We identified 2 psychological and attitudinal variables that were significant predictors of hesitancy to undergo RAT. Those who had good or better self-assessed mental health and did not perceive a high risk for SARS-CoV-2 infection were more likely to report hesitancy, which might be attributable to the belief that one ought to undergo RAT if they are at high risk for SARS-CoV-2 infection. This finding suggests that the government should offer appropriate recommendations on RAT, with the aim of identifying infected individuals in the early phase and keeping COVID-19 incidence low to tackle outbreaks in the future and especially to safeguard the shift from a pandemic phase to an endemic phase of the COVID-19 global response.

To reliably scale up RAT further for case management or self-isolation at the individual level and for large-scale screening and emergency responses at the population level, the government should offer more accurate recommendations on RAT to the target population. First, the public should be provided with the test as widely and freely as possible, which could be a strategy to improve the availability of RAT and the willingness to undergo RAT without additional financial cost for the public. Second, trust about RAT should be enhanced, including enough knowledge about COVID-19 and clear instructions about RAT via multiple media. Furthermore, the public should be provided with complete guidelines after undergoing RAT. For example, accessible PCR testing could be used to validate the results of RAT owing to the risk of false-positive results. If positive results are confirmed, guidelines about case management or self-isolation should be put in place.

SARS-CoV-2 has infected up to 900 million people as of January 11, 2023, in Mainland China after lifting of the COVID-19 policy [40]. The emergence of new variants (the Omicron sublineage XBB [41]) of the virus has also raised concerns about their potential to evade existing vaccines and treatments, highlighting the importance of ongoing surveillance and research. Given the reopening of society, vaccines; medication; and public health measures, such as masking, social distancing, and screening remain important tools in controlling the spread of the virus. Furthermore, by fully using the lessons learned from the mass screening (including timely scale-up of RAT) during the COVID-19 pandemic, this study could help to develop more resilient and prepared mass screening strategies, and could inform future responses to other novel infectious diseases.

In this study, we used group regression with an MCP penalty to discern variables that significantly shaped hesitancy to undergo RAT. It was found that several sociodemographic variables, variables concerning experiences of COVID-19 restrictions and knowledge of COVID-19, and psychological and attitudinal variables were associated with hesitancy to undergo RAT for COVID-19. Notably, variables, such as residence, marital status, and occupation type, were not selected for inclusion in the final model despite statistically significant differences between the study groups, which suggests that group regression with an MCP penalty was successful in not only selecting significant variables but also encouraging model sparsity. Some limitations of this study are worth acknowledging. In this survey, 88.33% of participants were from urban areas and 80.31% had an undergraduate education or above. Thus, it is very likely that individuals from rural areas and those with lower education are underrepresented. Individuals with SARS-CoV-2 infection were not included in the analysis. In addition, the sample size in this study was small compared to the large population size of China, which is another limitation of the study. Therefore, our findings should be interpreted and generalized with caution. Moreover, regarding hesitancy to undergo RAT, a single question was asked to assess hesitancy. Despite its simplicity and efficiency, more multidimensional items should be adopted to assess hesitancy in future studies.

In addition, the general bias for a cross-sectional design could not be dismissed, which limited our capacity to statistically discern causal relationships. Furthermore, given the reopening of society, the aim was no longer completely relevant, considering the lifting of global and Chinese preventive and control measures for COVID-19. However, the real-world evidence in this study could inform future responses to other novel infectious diseases.

In conclusion, our study found a low but potentially problematic level of hesitancy to undergo RAT. Although hesitancy to undergo SARS-CoV-2 RAT in China was low, it may influence the scale-up of RAT. Special efforts should be made to improve the awareness and acceptance of RAT among men, younger adults, individuals with a lower education or salary, families without children and elders, and individuals who access COVID-19 information via traditional media. Given the gradual reopening of society, this study could inform our responses to future novel infectious diseases and help to develop more resilient and prepared mass screening strategies (including timely scale-up of RAT).

Acknowledgments

This study was supported by the Special Support Plan for High-Level Talents of Guangdong Province (2019TQ05Y230) and the Programme of Science and Technology Development Plan of Wuxi Municipal Science and Technology Bureau (Y20222006). The funding parties did not have any role in the design of the study or in the explanation of the data.

Data Availability

The deidentified data set used in this study and the R codes are publicly available and can be accessed through the following GitHub repository: [42].

Authors' Contributions

HZ conceived the study and designed the protocol. ZL, LF, LY, TT, YG, and XM conducted the survey. ZL contributed to the statistical analyses and interpretations of the data. ZL and HZ drafted the manuscript, with all authors critically revising the manuscript. All authors have read and approved the final version of the manuscript. XM and HZ are co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
All variables left in the final model for correlates of hesitancy to undergo rapid antigen testing.
[DOCX File , 41 KB - publichealth_v9i1e43555_app1.docx]

References

1. Singh H, Dahiya N, Yadav M, Sehrawat N. Emergence of SARS-CoV-2 New Variants and Their Clinical Significance. Can J Infect Dis Med Microbiol 2022;2022:7336309 [FREE Full text] [doi: 10.1155/2022/7336309] [Medline: 35669528]
2. Tracking of VOC Omicron. Global Initiative of Sharing All Influenza Data (GISAID). URL: https://www.gisaid.org/hcov19-variants/ [accessed 2022-08-03]
3. Iuliano AD, Brunkard JM, Boehmer TK, Peterson E, Adjei S, Binder AM, et al. Trends in Disease Severity and Health Care Utilization During the Early Omicron Variant Period Compared with Previous SARS-CoV-2 High Transmission Periods - United States, December 2020-January 2022. MMWR Morb Mortal Wkly Rep 2022 Jan 28;71(4):146-152 [FREE Full text] [doi: 10.15585/mmwr.mm7104e4] [Medline: 35085225]
4. Coronavirus (COVID-19) latest insights. UK Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19/latestinsights [accessed 2022-08-01]
5. Press conference held on situation regarding strict prevention and control of COVID-19 epidemic. The State Council Information Office PRC. URL: http://www.gov.cn/xinwen/gwylflkjz193/index.htm [accessed 2022-08-19]
6. Coronavirus (COVID-19) vaccinations. Our World in Data. URL: https://ourworldindata.org/covid-vaccinations [accessed 2022-08-11]
7. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, et al. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. Lancet Child Adolesc Health 2020 May;4(5):397-404 [FREE Full text] [doi: 10.1016/S2352-4642(20)30095-X] [Medline: 32272089]
8. FIND evaluations of SARS-CoV-2 assays. FIND. URL: https://www.finddx.org/covid-19/find-evaluations-of-sars-cov-2-assays/ [accessed 2022-08-10]
9. Peeling RW, Heymann DL, Teo Y, Garcia PJ. Diagnostics for COVID-19: moving from pandemic response to control. Lancet 2022 Feb 19;399(10326):757-768 [FREE Full text] [doi: 10.1016/S0140-6736(21)02346-1] [Medline: 34942102]

10. McGarry BE, SteelFisher GK, Grabowski DC, Barnett ML. COVID-19 Test Result Turnaround Time for Residents and Staff in US Nursing Homes. JAMA Intern Med 2021 Apr 01;181(4):556-559 [FREE Full text] [doi: 10.1001/jamainternmed.2020.7330] [Medline: 33125044]
11. Hui DSC, Zumla A. Advances in the epidemiology, clinical features, diagnosis, clinical management and prevention of coronavirus disease 2019. Curr Opin Pulm Med 2022 May 01;28(3):166-173. [doi: 10.1097/MCP.0000000000000875] [Medline: 35266905]
12. Guidance for Antigen Testing for SARS-CoV-2 for Healthcare Providers Testing Individuals in the Community. CDC. URL: https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antigen-tests-guidelines.html [accessed 2022-08-10]
13. WHO Director-General's opening remarks at the media briefing on COVID-19 - 16 March 2020. WHO. URL: https://tinyurl.com/yffywfe3 [accessed 2023-08-10]
14. Liu J, Liu M, Liang W. The Dynamic COVID-Zero Strategy in China. China CDC Wkly 2022 Jan 28;4(4):74-75 [FREE Full text] [doi: 10.46234/ccdcw2022.015] [Medline: 35186372]
15. Options for the use of rapid antigen tests for COVID-19 in the EU/EEA - first update. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/publications-data/options-use-rapid-antigen-tests-covid-19-eueea-first-update [accessed 2022-08-01]
16. Chen M, Li Y, Chen J, Wen Z, Feng F, Zou H, et al. An online survey of the attitude and willingness of Chinese adults to receive COVID-19 vaccination. Hum Vaccin Immunother 2021 Jul 03;17(7):2279-2288 [FREE Full text] [doi: 10.1080/21645515.2020.1853449] [Medline: 33522405]
17. Ghaznavi C, Yoneoka D, Kawashima T, Eguchi A, Murakami M, Gilmour S, et al. Factors associated with reversals of COVID-19 vaccination willingness: Results from two longitudinal, national surveys in Japan 2021-2022. Lancet Reg Health West Pac 2022 Oct;27:100540 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100540] [Medline: 35892009]
18. Montag C, Becker B, Gan C. The Multipurpose Application WeChat: A Review on Recent Research. Front Psychol 2018;9:2247 [FREE Full text] [doi: 10.3389/fpsyg.2018.02247] [Medline: 30618894]
19. Hays RD, Liu H, Kapteyn A. Use of Internet panels to conduct surveys. Behav Res Methods 2015 Sep;47(3):685-690 [FREE Full text] [doi: 10.3758/s13428-015-0617-9] [Medline: 26170052]
20. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
21. Rey D, Fressard L, Cortaredona S, Bocquier A, Gautier A, Peretti-Watel P, Baromètre santé 2016 group. Vaccine hesitancy in the French population in 2016, and its association with vaccine uptake and perceived vaccine risk-benefit balance. Euro Surveill 2018 Apr;23(17):17-00816 [FREE Full text] [doi: 10.2807/1560-7917.ES.2018.23.17.17-00816] [Medline: 29717693]
22. Zeng Y, Yuan Z, Yin J, Han Y, Chu C, Fang Y. Factors affecting parental intention to vaccinate kindergarten children against influenza: A cross-sectional survey in China. Vaccine 2019 Mar 07;37(11):1449-1456. [doi: 10.1016/j.vaccine.2019.01.071] [Medline: 30765168]
23. Zhao Y, Di X, Li S, Zeng X, Wang X, Nan Y, et al. Prevalence, frequency, intensity, and location of cigarette use among adolescents in China from 2013-14 to 2019: Findings from two repeated cross-sectional studies. Lancet Reg Health West Pac 2022 Oct;27:100549 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100549] [Medline: 35923777]
24. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. J Nutr Health Aging 2012 Jul;16(7):601-608 [FREE Full text] [doi: 10.1007/s12603-012-0084-2] [Medline: 22836700]
25. Mazumder R, Friedman JH, Hastie T. : Coordinate Descent With Nonconvex Penalties. J Am Stat Assoc 2011;106(495):1125-1138 [FREE Full text] [doi: 10.1198/jasa.2011.tm09738] [Medline: 25580042]
26. Murphy KP. Machine Learning: A Probabilistic Perspective. Cambridge, MA: MIT Press; 2012.
27. Goggolidou P, Hodges-Mameletzis I, Purewal S, Karakoula A, Warr T. Self-Testing as an Invaluable Tool in Fighting the COVID-19 Pandemic. J Prim Care Community Health 2021;12:21501327211047782 [FREE Full text] [doi: 10.1177/21501327211047782] [Medline: 34583571]
28. Mouliou DS, Pantazopoulos I, Gourgoulianis KI. Societal Criticism towards COVID-19: Assessing the Theory of Self-Diagnosis Contrasted to Medical Diagnosis. Diagnostics (Basel) 2021 Sep 27;11(10):1777 [FREE Full text] [doi: 10.3390/diagnostics11101777] [Medline: 34679475]
29. Thomas C, Shilton S, Thomas C, Batheja D, Goel S, Mone Iye C, et al. Values and preferences of the general population in Indonesia in relation to rapid COVID-19 antigen self-tests: A cross-sectional survey. Trop Med Int Health 2022 May;27(5):522-536 [FREE Full text] [doi: 10.1111/tmi.13748] [Medline: 35332616]
30. Wang J, Jing R, Lai X, Zhang H, Lyu Y, Knoll MD, et al. Acceptance of COVID-19 Vaccination during the COVID-19 Pandemic in China. Vaccines (Basel) 2020 Aug 27;8(3):482 [FREE Full text] [doi: 10.3390/vaccines8030482] [Medline: 32867224]
31. Khan H, Dabla-Norris M, Lima F, Sollaci A. Who Doesn’t Want to be Vaccinated? Determinants of Vaccine Hesitancy During COVID-19. International Monetary Fund. URL: https://www.imf.org/en/Publications/WP/Issues/2021/05/06/Who-Doesnt-Want-to-be-Vaccinated-Determinants-of-Vaccine-Hesitancy-During-COVID-19-50244 [accessed 2022-08-10]

32. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 2021 Apr;6(4):e210-e221 [FREE Full text] [doi: 10.1016/S2468-2667(21)00012-8] [Medline: 33556325]
33. Paul E, Steptoe A, Fancourt D. Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications. Lancet Reg Health Eur 2021 Feb;1:100012 [FREE Full text] [doi: 10.1016/j.lanepe.2020.100012] [Medline: 33954296]
34. Damialis A, Gilles S, Sofiev M, Sofieva V, Kolek F, Bayr D, COVID-19/POLLEN study group, et al. Higher airborne pollen concentrations correlated with increased SARS-CoV-2 infection rates, as evidenced from 31 countries across the globe. Proc Natl Acad Sci U S A 2021 Mar 23;118(12):e2019034118 [FREE Full text] [doi: 10.1073/pnas.2019034118] [Medline: 33798095]
35. Gillies CL, Rowlands AV, Razieh C, Nafilyan V, Chudasama Y, Islam N, et al. Association between household size and COVID-19: A UK Biobank observational study. J R Soc Med 2022 Apr;115(4):138-144 [FREE Full text] [doi: 10.1177/01410768211073923] [Medline: 35118908]
36. Kevadiya BD, Machhi J, Herskovitz J, Oleynikov MD, Blomberg WR, Bajwa N, et al. Diagnostics for SARS-CoV-2 infections. Nat Mater 2021 May;20(5):593-605 [FREE Full text] [doi: 10.1038/s41563-020-00906-z] [Medline: 33589798]
37. Robinson E, Jones A, Lesser I, Daly M. International estimates of intended uptake and refusal of COVID-19 vaccines: A rapid systematic review and meta-analysis of large nationally representative samples. Vaccine 2021 Apr 08;39(15):2024-2034 [FREE Full text] [doi: 10.1016/j.vaccine.2021.02.005] [Medline: 33722411]
38. Shah S, Gui H, Chua PEY, Tan J, Suen LK, Chan SW, et al. Factors associated with COVID-19 vaccination intent in Singapore, Australia and Hong Kong. Vaccine 2022 May 09;40(21):2949-2959 [FREE Full text] [doi: 10.1016/j.vaccine.2022.03.062] [Medline: 35414417]
39. Luk TT, Zhao S, Wu Y, Wong JY, Wang MP, Lam TH. Prevalence and determinants of SARS-CoV-2 vaccine hesitancy in Hong Kong: A population-based survey. Vaccine 2021 Jun 16;39(27):3602-3607 [FREE Full text] [doi: 10.1016/j.vaccine.2021.05.036] [Medline: 34034950]
40. Zhong Y, Huang J, Zhang W, Li S, Gao Y. Addressing psychosomatic issues after lifting the COVID-19 policy in China: A wake-up call. Asian J Psychiatr 2023 Apr;82:103517 [FREE Full text] [doi: 10.1016/j.ajp.2023.103517] [Medline: 36791610]
41. Link-Gelles R, Ciesla AA, Roper LE, Scobie HM, Ali AR, Miller JD, et al. Early Estimates of Bivalent mRNA Booster Dose Vaccine Effectiveness in Preventing Symptomatic SARS-CoV-2 Infection Attributable to Omicron BA.5- and XBB/XBB.1.5-Related Sublineages Among Immunocompetent Adults - Increasing Community Access to Testing Program, United States, December 2022-January 2023. MMWR Morb Mortal Wkly Rep 2023 Feb 03;72(5):119-124 [FREE Full text] [doi: 10.15585/mmwr.mm7205e1] [Medline: 36730051]
42. Hesitancy-about-SARS-CoV-2-rapid-antigen-testing-in-China. URL: https://github.com/Leslie-Lu/Hesitancy-about-SARS-CoV-2-rapid-antigen-testing-in-China

Abbreviations

aOR: adjusted odds ratio
MCP: minimax concave penalty
NPI: nonpharmaceutical intervention
PCR: polymerase chain reaction
RAT: rapid antigen testing

Edited by A Mavragani, T Sanchez; submitted 15.10.22; peer-reviewed by B Liang, W Chen, H Jiang; comments to author 23.02.23; revised version received 27.02.23; accepted 07.03.23; published 01.05.23.

Please cite as:
Lu Z, Fu L, Yang L, Tian T, Gao Y, Meng X, Zou H
Hesitancy to Undergo SARS-CoV-2 Rapid Antigen Testing in China: Nationwide Cross-sectional Study
JMIR Public Health Surveill 2023;9:e43555
URL: https://publichealth.jmir.org/2023/1/e43555
doi:10.2196/43555
PMID:36888911

Review

Knowledge and Practice of Personal Protective Measures Against COVID-19 in Africa: Systematic Review

Joseph Kawuki[1], MPH; Paul Shing-fong Chan[1], MEd; Yuan Fang[2], PhD; Siyu Chen[1], MPH; Phoenix K H Mo[1], PhD; Zixin Wang[1], PhD

[1] Centre for Health Behaviours Research, Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Shatin, NT, China (Hong Kong)

[2] Department of Health and Physical Education, the Education University of Hong Kong, New Territories, China (Hong Kong)

Corresponding Author:
Zixin Wang, PhD
Centre for Health Behaviours Research
Jockey Club School of Public Health and Primary Care
The Chinese University of Hong Kong
Room 508, School of Public Health, Prince of Wales Hospital
Shatin, NT
China (Hong Kong)
Phone: 852 22528740
Email: wangzx@cuhk.edu.hk

Abstract

Background: With COVID-19 being a newly evolving disease, its response measures largely depend on the practice of and compliance with personal protective measures (PPMs).

Objective: This systematic review aimed to examine the knowledge and practice of COVID-19 PPMs in African countries as documented in the published literature.

Methods: A systematic search was conducted on the Scopus, PubMed, and Web of Science databases using appropriate keywords and predefined eligibility criteria for the selection of relevant studies. Only population-based original research studies (including qualitative, quantitative, and mixed methods studies) conducted in Africa and published in the English language were included. The screening process and data extraction were performed according to a preregistered protocol in PROSPERO (CRD42022355101) and followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The quality of the included studies was assessed using the Mixed Methods Appraisal Tool. Thematic analysis was used to systematically summarize the studies into 4 predefined domains: knowledge and perception of PPMs, mask use, social and physical distancing, and handwashing and hand hygiene, including their respective levels and associated factors.

Results: A total of 58 studies across 12 African countries were included, published between 2019 and 2022. African communities, including various population groups, had varying levels of knowledge and practice of COVID-19 PPMs, with the lack of personal protective equipment (mainly face masks) and side effects (among health care workers) being the major reasons for poor compliance. Lower rates of handwashing and hand hygiene were particularly noted in several African countries, especially among low-income urban and slum dwellers, with the main barrier being the lack of safe and clean water. Various cognitive (knowledge and perception), sociodemographic, and economic factors were associated with the practice of COVID-19 PPMs. Moreover, there were evident research inequalities at the regional level, with East Africa contributing 36% (21/58) of the studies, West Africa contributing 21% (12/58), North Africa contributing 17% (10/58), Southern Africa contributing 7% (4/58), and no single-country study from Central Africa. Nonetheless, the overall quality of the included studies was generally good as they satisfied most of the quality assessment criteria.

Conclusions: There is a need to enhance local capacity to produce and supply personal protective equipment. Consideration of various cognitive, demographic, and socioeconomic differences, with extra focus on the most vulnerable, is crucial for inclusive and more effective strategies against the pandemic. Moreover, more focus and involvement in community behavioral research are needed to fully understand and address the dynamics of the current pandemic in Africa.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42022355101; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022355101

(JMIR Public Health Surveill 2023;9:e44051) doi:10.2196/44051

KEYWORDS

personal protective measures; mask use; social distancing; hand hygiene; COVID-19; Africa; nonpharmaceutical interventions

Introduction

Background

After its emergence in December 2019, COVID-19 was declared a pandemic by the World Health Organization on March 11, 2020, and it has spread to almost all countries and regions, including Africa [1,2]. Spreading to the continent through travelers returning from hot spots in Asia, Europe, and the United States, COVID-19 was first recorded in Africa in Egypt on February 14, 2020, and within a few months, the virus had spread throughout the continent [1,3]. As of March 20, 2023, a total of 12,804,191 cumulative cases and 258,623 deaths have been reported in Africa compared with 682,546,389 cases and 6,819,835 deaths across the globe, showing a disproportionately low case fatality rate of COVID-19 in Africa [4].

As in the rest of the world, various response measures were implemented in different African countries to curb the spread of the virus, including statewide lockdowns, restrictions on movement, bans on social gatherings, and school closures [5,6]. Although the continent appears to have a lower absolute number of cases and deaths compared with other regions [7], which might also be related to the lower number of tests administered, the pandemic has had a deep impact on the socioeconomic systems of African countries [8,9]. The pandemic has also strained the weak and fragmented health systems, as shown by the lack of personal protective equipment (PPE), testing kits, and other treatment necessities, especially for patients with COVID-19 who are critically ill [6,8].

With COVID-19 being a newly evolving disease, its less-defined outcomes and unprecedented prevention, treatment, and control modes largely require indisputable collaboration among various stakeholders in the community [9]. Nonpharmaceutical interventions play an important role in the control and prevention of pandemics, including the COVID-19 pandemic, especially in its early phase and wave. Despite the availability of approved vaccines against COVID-19, response measures toward this pandemic still largely depend on the practice of and compliance with personal protective measures (PPMs), including face mask use, social and physical distancing, and hand hygiene [10]. Moreover, knowledge and perceptions of PPMs have been reported as among the key determinants of practice and compliance with PPMs against COVID-19 as they influence people’s behavior [10,11]. These were also considered in this study in the African context.

The pandemic has had a broad range of impacts and challenges across regions, and different communities have responded differently. However, given the diversity of social systems across regions and countries, preparedness and the search for a country- or region-specific practical solution to the pandemic require a better understanding of the challenges of practicing PPMs and hard-learned experiences through comprehensive research [9]. There has been a high research output documenting COVID-19 characteristics, clinical outcomes, response, and impact throughout the world but with much less research coming from African countries [12,13]. The unavailability of research information might be seen as a barrier to successful prevention and further as a sign of inequity between high- and low-income countries and regions [14]. This scant literature poses knowledge gaps on how African countries are responding to the pandemic in terms of PPMs. Nonetheless, a recent review by Nwagbara et al [15] reported that most communities in sub-Saharan Africa had a positive attitude toward and good practices regarding COVID-19. Notably, this review considered studies only from sub-Saharan Africa and those conducted in the first stages of the pandemic, so it lacked insights into the overall practice of PPMs in Africa.

Objectives

Thus, this systematic review aimed to examine COVID-19 PPM research from African countries as documented in the published literature. On the basis of specific keywords, the review looked at the levels and associated factors of (1) knowledge and perception of PPMs and (2) practice of COVID-19 PPMs in various populations (including face mask use, physical and social distancing, and handwashing and hand hygiene).

Methods

Study Design

This systematic review was conducted according to a preregistered protocol in PROSPERO (CRD42022355101) and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Multimedia Appendix 1) [16]. This systematic review considered literature concerning PPMs from African countries. Literature was sourced from the following databases: Scopus, PubMed, and Web of Science. These databases were considered as they sufficiently cover most of the key journals, including most African journals. In addition, 2 of these databases (Scopus and Web of Science) could refine the search based on countries and regions, unlike other databases, which enabled us to specifically assess publications from African countries only.

Search Strategy

We conducted a comprehensive search using a set of appropriate keywords and Medical Subject Heading terms to identify studies reporting on PPMs. For consistency and precision, similar keywords and Medical Subject Heading terms were used and searched for in the article titles across the databases. A comprehensive search of the published literature was performed in each of the 3 selected databases using combinations of key terms and Boolean operators (Textbox 1). These included “mask,” “nose covering,” “personal protective equipment,” “handwashing,” “hand washing,” “hand sanitizer,” “hand sanitiser,” “sanitation,” “hygiene,” “social distance,” “social distancing,” “physical distance,” “physical distancing,” “social acceptance," "COVID-19," "2019-nCoV," "coronavirus disease," "SARS-CoV-2," and "corona virus disease 2019."

Textbox 1. Key terms or Boolean operators used for the search.

- "Mask" OR "nose covering" OR "personal protective equipment" OR "handwashing" OR "hand washing" OR "hand sanitizer" OR "hand sanitiser" OR "sanitation" OR "hygiene" OR "social distance" OR "social distancing" OR "physical distance" OR "physical distancing" OR "social acceptance" AND "COVID-19" OR "2019-nCoV" OR "coronavirus disease" OR "SARS-CoV-2" OR "corona virus disease 2019"
- "Mask" OR "nose covering" OR "personal protective*" OR "hand wash*" OR "hand-wash*" OR "hand sanitize*" OR "hand sanitise*" OR "sanitation*" OR "hygiene*" OR "social distance*" OR "physical distance*" OR "social accept*" OR "social acceptance" AND "COVID-19" OR "COVID*"
- "Mask" OR "nose covering" OR "personal protective*" OR "hand wash*" OR "hand-wash*" OR "hand sanitize*" OR "hand sanitise*" OR "sanitation*" OR "hygiene*" OR "social distance*" OR "physical distance*" OR "social accept*" OR "social acceptance" AND "coronavirus*" OR "corona*"
- "Mask" OR "nose covering" OR "personal protective*" OR "hand wash*" OR "hand sanitize*" OR "hand sanitise*" OR "sanitation*" OR "hygiene*" OR "social distance*" OR "physical distance*" OR "social accept*" OR "social acceptance" AND "SARS-CoV-2*" OR "2019-nCoV*"

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria are listed in Table 1. Only population-based original research studies (including qualitative, quantitative, and mixed methods studies) conducted in Africa, published in English, and reporting on PPMs against COVID-19 were considered in the full review. In addition, multicountry studies were considered if they included an African country as part of their study population. Only English-language articles published between November 1, 2019, and March 4, 2022, were considered.

Table 1. Summary of the inclusion and exclusion criteria.

Parameter	Inclusion criteria	Exclusion criteria
Article or study type	• Population-based original research studies • Qualitative, quantitative, and mixed methods studies • Multicountry studies	• Reviews, commentaries, and editorials • Dissertations, government reports, newspaper articles, textbooks, book chapters, and protocols • Gray literature and preprints • Laboratory studies, model and framework studies, and validation studies
Language	• English language	• All other non-English languages
Publication period	• November 1, 2019, to March 4, 2022	• All periods outside November 2019 to March 2022
Study setting	• African countries	• All non-African country settings

Data Extraction

After screening, data from the relevant studies were independently extracted by 2 reviewers (JK and PSC) onto a structured data extraction template, and a consensus was reached through discussion in case of disagreements on the extracted data. The following variables were extracted: first author, year of publication, study location, study design, key measurements, study population, sample size. and main findings.

Quality Assessment and Analysis

We assessed the information from the included articles using the Mixed Methods Appraisal Tool (version 2018) with detailed descriptions of the rating [17]. In total, 2 reviewers also independently assessed the quality of the included studies, and in case of discrepancies, a consensus was reached through discussion.

This study used thematic analysis, and the literature in this review was used to understand the practice of PPMs against COVID-19 in African countries. The studies were classified according to four main themes: (1) knowledge and perception of PPMs, (2) mask use, (3) social and physical distancing, and (4) handwashing and hand hygiene, including their respective levels or prevalence and associated factors. In addition, various population groups and challenges faced in practicing COVID-19 PPMs were examined under each of the main themes based on the available information in the analyzed studies. The 4 themes were drafted by a panel of public health experts after a series of discussions to reach a consensus.

The analysis process involved a six-step data synthesis process: (1) in total, 2 reviewers (JK and PSC) extracted relevant information on knowledge and practice of PPMs from each article independently; (2) after extraction, they discussed to reach a consensus on the key information identified in each article; (3) the extracted information was coded under the 4 predefined domains by the 2 reviewers independently; (4) after completing the coding independently, they discussed the results, where any discrepancies were resolved through discussion; (5) the revised coding results were read and checked by the 2 reviewers independently to ensure that all the extracted information was mapped to the 4 domains correctly; and (6) all the information in the codebook was adapted into a tabular format.

Results

Selection of Studies Conducted in Africa

The number of studies identified, reviewed, and selected, including the reasons for exclusion, is summarized in Figure 1. A total of 58 studies were selected through this process and further analyzed [18-75]. The information and main findings extracted from all included studies is detailed in Table 2.

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the selected studies. WoS: Web of Science.

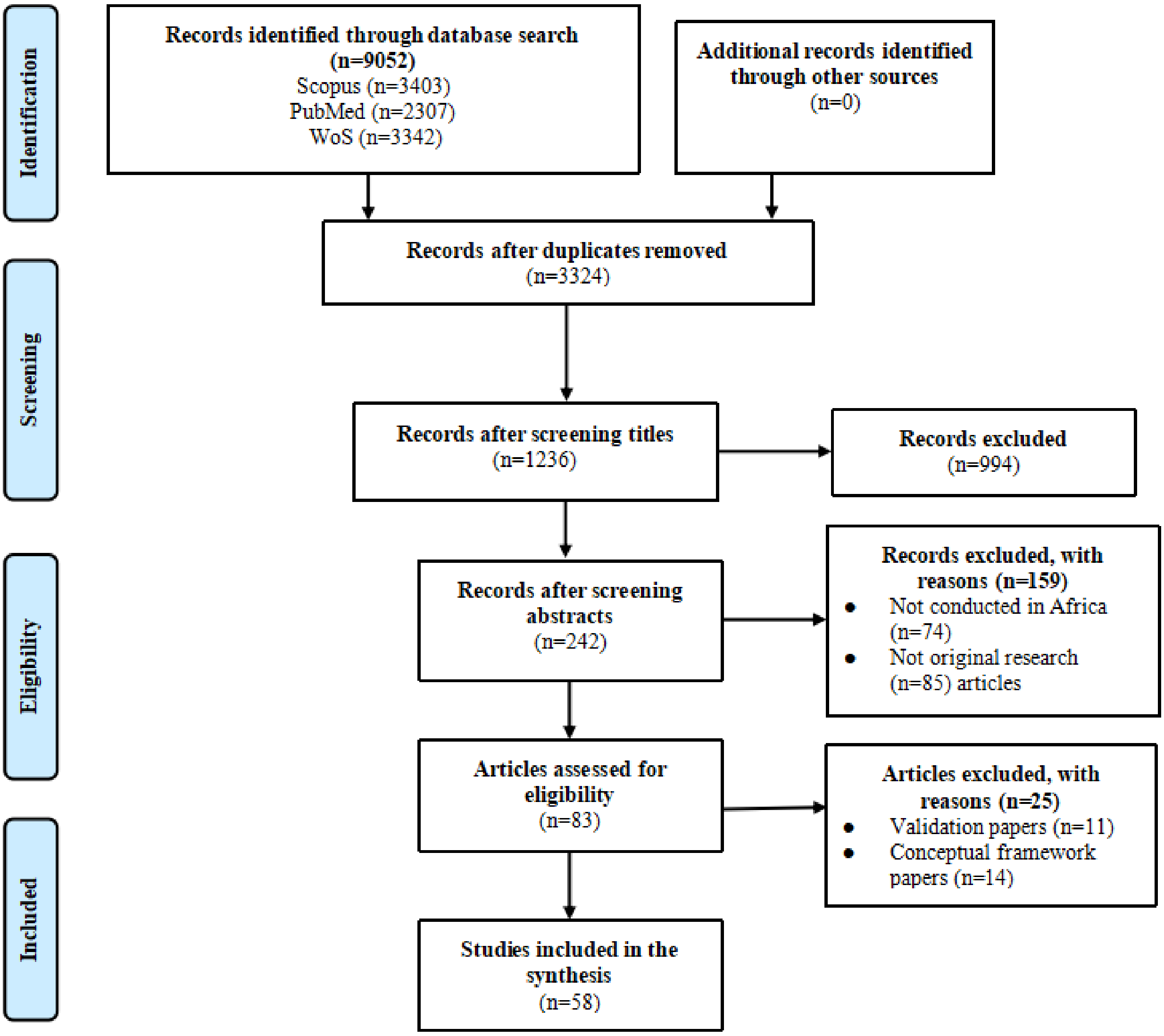


Table 2. Characteristics of the included studies.

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
Sikakulya et al [18], 2021	Uganda	Cross-sectional; knowledge, attitudes, and practices regarding proper use of face masks	Community	1114	51.5% had poor mask use	• Most participants (60.1%) had satisfactory knowledge of the use of face masks, and this was greater among participants with tertiary educational levels. • Regarding attitude, 69.4% were confident enough to correctly put on a face mask, 83.4% believed that a face mask can protect against COVID-19, and 75.9% had never shared their face mask. Most (95.2%) agreed that wearing face masks in public places was important to protect themselves against COVID-19, and 60.3% reported washing their hands before wearing and after removing the face mask.
Hailu et al [20], 2021	Ethiopia	Cross-sectional and mixed methods; compliance with social distancing	Community	401	Overall, 55.4% had poor compliance with social distancing measures	• Most (63.84%) reported that they went to crowded places without putting on a face mask, but 60.6% and 76.3% had good knowledge of COVID-19 transmission and prevention, respectively. • Only age was associated with social distancing measures, with older persons more likely than younger persons to comply with social distancing guidelines.
Bakry and Waly [21], 2020	Egypt	Cross-sectional; perception and practice of social distancing	Community	1036	82% were not strictly practicing social distancing	• Most (70%) perceived that social distancing measures reduced the transmission of COVID-19. • There was a significant association between the practice of social distancing and some sociodemographic factors such as sex, age, education, working status, place of residence, and community of residence.
Tadesse et al [22], 2020	Ethiopia	Cross-sectional; predictors of preventive practices	Community employees	628	68.8% had poor COVID-19 prevention practice	• Most (>50%) had high perceived susceptibility, severity, benefit, barriers, cues to action, and self-efficacy regarding COVID-19 prevention practice. • Employees with a low level of perceived barriers were less likely to have a poor practice of COVID-19 prevention compared with employees with a high level of perceived barriers. • Moreover, employees with low cues to action and employees with a low level of self-efficacy practiced COVID-19 prevention measures to a lesser extent compared with those with high cues to action and high levels of self-efficacy.
Iyamu et al [23], 2022	MCP[a]—6 countries: Botswana, Kenya, Malawi, Nigeria, Zambia, and Zimbabwe	Cross-sectional; face mask use perception and social media	Community	1988	—[b]	• A total of 58.8% used social media as their main source of information, whereas 85% agreed that face masks were effective against COVID-19. • Respondents who used social media were more likely to agree that face masks were effective compared with those who did not.
Bukuluki and Kisaakye [24], 2021	Uganda	Cross-sectional; face mask wearing in public places	Community	1054	52% and 78% wore face masks sometimes inside in public spaces and always outside in public spaces, respectively	• Approximately 90% of respondents agreed that wearing a mask inside or outside in public spaces can prevent COVID-19 infection. • Age and frequency of face mask wearing inside or outside in public spaces were significantly related to belief in face mask efficacy.

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
Nnama-Okechukwu et al [25], 2020	Nigeria	Qualitative study; knowledge of and compliance with preventive measures	Community	36	—	• Findings revealed that most of the respondents believed that the COVID-19 pandemic was more of a hoax than a reality. • Other findings showed that this poor knowledge negatively affected their compliance with preventive measures to curb the spread of coronavirus.
Kajiita and Kang'ethe [26], 2021	8 African countries	Cross-sectional and qualitative; social distancing perceptions	Community	20	—	• Results revealed varied conceptualizations and interpretations of the disease and social distancing. Notably, COVID-19 regulations such as social distancing and face mask wearing were perceived as an imported policy, a misconception responsible for nonadherence to COVID-19 protocols. • Furthermore, the study underscored that the disease and policies related to it disrupted ways of social life, infringed on people's social-cultural rights, and had adverse health consequences.
Fodjo et al [27], 2020	10 countries; DRC[c], Uganda, Mozambique, and Somalia	Multicountry web-based survey; compliance with mask use	General public	206,729	Face mask use—DRC: 43.2%; Uganda: 32.7%; Mozambique: 93.9%; Somalia: 51.2%	• Adherence rates were higher in countries where masking was mandatory or highly encouraged by the government during the early phases of the COVID-19 outbreak. • Reusable cloth masks (more cost-beneficial and environmentally friendly than surgical masks) were the most frequent, accounting for 51.1% of all mask types. • There were differential rates of mask uptake and use between sexes and age groups observed in different countries. Even in countries where no preexisting culture of mask use existed, high uptake of mass masking was feasible.
Sewpaul et al [28], 2021	South Africa	Cross-sectional; compliance with and determinants of social distancing	Community	17,563	20.3% reported having not left home	• A total of 50.6% were in close physical distance with 1-10 people, 21.1% were in close physical distance with 11-50 people, and 8% were in close physical distance with >50 people. • Larger household sizes and incorrect knowledge about the importance of social distancing were associated with being in contact with >50 people. Male sex, younger age, and being in the White and non-White population groups were significantly associated with being in contact with 1-10 people but not with larger numbers of people. • Employment, at least a secondary school education, the lack of self-efficacy in being able to protect oneself from infection, and moderate or high risk perception of becoming infected were all associated with increased odds of close contact with 1-10, 11-50, and >50 people relative to remaining at home.
Wondimu et al [29], 2020	Ethiopia	Cross-sectional; predictors of preventive practices	Community	803	Generally, 59.4% had good prevention practices for COVID-19	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						- Approximately 64.7% had a history of going to crowded places, whereas only 30.3% of the participants had a history of wearing a mask when leaving home. - A total of 64.4% had a history of maintaining their distance at 2 meters, and 64.8% washed their hands with soap and water or used alcohol-based hand sanitizers. - Urban residence, family size, good knowledge, positive attitude, intention to seek care, and perceived mortality were positively associated with good prevention practices.
Fikrie et al [30], 2021	Ethiopia	Cross-sectional; social distancing and associated factors	Community	410	38.3% had good social distancing practices	- Younger age (26-35 years) and being employed were positively associated with good social distancing practice, whereas poor knowledge, negative attitude, and low perceived susceptibility had a negative association.
Mejjad et al [31], 2021	Morocco	Cross-sectional; mask use and disposal behavior	Community	185	70% used face masks at least once a day	- A total of 70% of the respondents threw their discarded masks and gloves in the house trash or trash bins after their first use, whereas nearly 30% of respondents admitted that they did not wear masks as they did not leave their homes during the lockdown.
Burger et al [32], 2022	South Africa	Longitudinal survey; predictors of mask wearing	Community	7074	74% wore face masks when in public	- Self-efficacy, the prevalence of others' mask wearing in the same district, and affluence were positively associated with reported mask wearing. - Those who reported staying at home were significantly less likely to report wearing a mask. - Despite having a higher mortality risk, older adults had significantly lower odds of mask wearing. - The prevalence of mask wearing increased significantly from May 2020 to August 2020 (from 50% to 74%) as COVID-19 cases increased and lockdown restrictions were eased, but staying at home, physical distancing, and social distancing decreased.
Amuakwa-Mensah et al [33], 2021	MCP—12 sub-Saharan African countries	Cross-sectional; handwashing and COVID-19 concerns	Community	4788	54.6% washed their hands for 20 seconds >5 times a day, and 4.2% did not wash their hands at all	- The level of concern about the spread of the virus increased the likelihood of washing hands with soap under running water for a minimum of 20 seconds at least 5 times a day. - Heterogeneous effects across gender and age groups, locality, and various water sources were noted.
Szczuka et al [19], 2021	MCP—Gambia	Observational study; handwashing adherence	Community	6064	—	- Higher handwashing adherence was associated with more frequent exposure to handwashing guidelines, being a health care professional, being older, being female, and being married. - Stricter containment and health policies were associated with lower handwashing adherence.
Iwuoha and Aniche [34], 2020	Nigeria	Cross-sectional and qualitative; impact of physical distancing policies	Slum residents	49	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• The study demonstrated that the peculiar and adverse conditions of low-income urban dwellers were not considered in the formulation of the COVID-19 lockdown and physical distancing policies in Nigeria. Thus, such policies worsened the living conditions of extremely low-income urban or suburban slum dwellers in Nigeria. • There is a need to engender an indigenous (Afro-centered) approach to the containment of the pandemic.
McCreesh et al [35], 2021	South Africa	Longitudinal survey; impact of social distancing regulations	Community	1704	—	• Extrahousehold social contact fell substantially following the imposition of COVID-19 distancing regulations in that there were substantial declines in close physical and conversational contacts and also in beyond-household sharing of indoor space. • However, there was ongoing contact within intergenerational households, highlighting a potential limitation of social distancing measures in protecting older adults.
De Backer et al [36], 2020	MCP—38 countries; Uganda, South Africa, and Egypt	Cross-sectional; impact of social distancing on healthy meals	Community	37,207	—	• Increases in planning, selecting, and preparing healthy foods were found for women and men and were positively related to perceived time availability and stay-at-home policies. • Psychological distress was a barrier for women and an enabler for men, whereas financial stress was a barrier and enabler depending on various sociodemographic variables.
Kim et al [37], 2022	Kenya	Cross-sectional; WASH[d] accessibility	Slum dwellers	647	—	• A total of 77.4% and 60.6% of people living in Kibera and Mathare, respectively, had limited WASH facility accessibility or opportunity. • Overall accessibility and opportunity were better in Mathare than in Kibera.
Ag Ahmed et al [38], 2021	Mali	Qualitative study; adoption of physical distancing measures	Internally displaced people	68	—	• The main challenges concerning the implementation and adoption of physical distancing measures included the proximity in which internally displaced people live, their beliefs and values, the lack of toilets and safe water on sites, internally displaced people's habits and economic situation, humanitarian actors' lack of financial resources and authority, and social pressure from religious leaders. • Implemented mitigation measures included the building of new shelters or their compartmentalization, the creation of income-generating activities and food banks, psychosocial support, promotion of awareness of internally displaced people, and night police patrols and surveillance to discourage internally displaced people from going out.
Mhlanga-Gunda et al [39], 2022	Zimbabwe	Qualitative study; social distancing and prevention measures	Prisoners and staff	80	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• Outdated infrastructure, severe congestion, interrupted water supply, and inadequate hygiene and sanitation were conducive to ill health and the spread of disease. • Health professionals had been well trained regarding COVID-19 control measures, and COVID-19 awareness among prisoners was generally adequate. • There was no routine COVID-19 testing in place beyond thermal scanning. • Access to health care was good, but standards were hindered by inadequate medicine and PPE[e] supply, and isolation measures were compromised by accommodation capacity issues. • The flow of prison entries constituted a transmission risk, and social distancing was impossible during meals and at night.
Assefa et al [40], 2021	Ethiopia	Cross-sectional; knowledge, attitude, practice, and challenges regarding hand hygiene	HCWs[f]	96	76% had good hand hygiene practices with alcohol-based hand sanitizers	• All the HCWs practiced different COVID-19 prevention methods, and most were knowledgeable (93.8%) and had a favorable attitude (74%). • However, 84.5% of the respondents faced challenges during alcohol-based hand sanitizer use owing to it being unavailable (68.8%) or expensive (52.1%); forgetting (11.5%); and experiencing health-associated risks such as skin irritation (28.1%), skin dryness (62.5%), and ocular irritation (11.5%).
Seid Yimer and Gebrehana Belay [41], 2021	Ethiopia	Hospital-based cross-sectional study; knowledge and practice of proper face mask use	HCWs	422	59.5% practiced proper mask use	• The overall good knowledge score of health care providers regarding proper face mask use was 65.8%. • Of them, 67.3% knew that face masks were worn with the white side facing in, 62.6% knew that face masks had 3 layers, and 78.4% knew that surgical face masks were worn for up to 8 hours.
Ahmed Sayed et al [42], 2021	Egypt	Cross-sectional; preparedness and attitude toward PPE	HCWs	254	—	• Only 28.74% of the house officers had good preparedness, whereas 85.83% had a good PPE attitude. • Preparedness and willingness were significantly associated with the overall worry related to the pandemic (fear of contracting COVID-19 and having family members at risk of severe COVID-19). • The house officers with good preparedness and willingness to deal with COVID-19 seemed to have a good PPE attitude.
Alao et al [43], 2020	Nigeria	Cross-sectional; knowledge, attitudes, beliefs, and use of PPE	HCWs	272	—	• Only 25.7% had adequate knowledge about PPE. Of the respondents who presumed that they had adequate knowledge about donning and doffing PPE, 94 (56%) were incorrect. • The predictors of good knowledge were younger age (<45 years) and practice location.
Daghmouri et al [44], 2020	Tunisia	Cross-sectional and institution-based; PPE use	HCWs	723	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• There was a likely overuse of PPE in addition to a high rate of side effects caused by PPE. • A total of 57.3% of participants reported a lack of PPE, which could be extremely stressful and detrimental to them. • A total of 72.5% of respondents reused disposable PPE, especially facial protective shields and FFP2^g. • Only 37.8% of frontline HCWs had received official training on the correct use of PPE, especially on how to fit FFP2 masks (only 32.6%). • A total of 89.1% of participants believed that they needed additional training.
Kassie et al [45], 2020	Ethiopia	Cross-sectional; preventive practices	HCWs	630	38.7% (95% CI 34.8%-42.5%) good preventive practice against COVID-19	• Being a male health care provider, having work experience of 6-10 years, and having a poor attitude toward COVID-19 were found to be significantly associated with poor COVID-19 preventive practices among HCWs.
Keleb et al [46], 2021	Ethiopia	Cross-sectional; PPE use and hand hygiene and associated factors	HCWs	489	32% and 22.3% were compliant with PPE use and hand hygiene practice, respectively	• Feedback for safety, training on COVID-19 prevention, and perception of infection risk were significant factors of good compliance with PPE use.
Birhanu et al [47], 2021	Ethiopia	Cross-sectional; PPE use	HCWs	418	37.6% had good practice of PPE use	• Being male, being a nurse or midwifery professional, regularly sanitizing hands and medical equipment, having national COVID-19 management guidelines, taking COVID-19 training, and the feeling of eventually contracting COVID-19 at the workplace had a positive association with PPE use. • However, not feeling safe at work when using standard precautions was negatively associated with PPE use.
El-Sokkary et al [48], 2021	Egypt	Cross-sectional; mask use and compliance	HCWs	404	53.2% were noncompliant with mask use	• Most reported a shortage in N95 respirators (91.3%) and practiced extended PPE use (88.1%). Better compliance with proper PPE use was reported among female individuals, physicians, and medical specialties with <10 years of work experience and working >8 hours per day. • Significant predictors of compliance were receiving previous training on the proper use of PPE, exposure to patients with COVID-19, and performing procedures that posed a high risk of exposure to COVID-19 to HCWs.
Afemikhe et al [49], 2020	Nigeria	Cross-sectional; transmission-based precaution practices	Nurses	367	85.6% maintained a good level of preventive practices, and 89.1% performed hand hygiene	• Academic qualification was a significant predictor of good practice in favor of respondents with a degree in Nursing. • Challenges identified were the lack of financial motivation, fear of infecting family members, and fear of contracting the virus (93.5%).
Elhadi et al [50], 2021	Libya	Cross-sectional; mask wearing	General population and HCWs	15,087	68.1% had mask wearing adherence	—

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
Tabah et al [51], 2020	90 countries; Libya, Egypt, Morocco, and Tunisia	A cross-sectional, international survey; PPE use	HCWs	2711	For routine care, 58% used FFP2 or N95 masks, waterproof long-sleeved gowns (67%), and face shields or visors (62%)	• Powered air-purifying respirators were used routinely and for intubation only by 7% and 13% of respondents, respectively. • Surgical masks were used for routine care and for intubations by 15% and 2% of respondents, respectively. • At least one piece of standard PPE was unavailable for 1402 (52%) respondents, and 30% reported reusing single-use PPE. • PPE was worn for a median of 4 (IQR 2-5) hours. Adverse effects of PPE were associated with longer shift durations and included heat (51%), thirst (47%), pressure areas (44%), headaches (28%), inability to use the bathroom (27%), and extreme exhaustion (20%).
Mahmoud et al [52], 2021	Egypt and Saudi Arabia	Comparative and cross-sectional; effect of sanitizers and PPE use	HCWs	428	—	• The most affected areas from wearing PPE were the hands (49.8% and 54.5%), followed by the auricular area (44% and 40.9%), the nasal bridge (28.9% and 22.7%), the cheeks (16.9% and 13.6%), and the whole face (15.6% and 25%) among Saudi and Egyptian HCWs, respectively. • Approximately 70% of Egyptian HCW participants used 70% alcohol in the form of a gel as sanitizer, which was significantly higher than Saudi HCWs (59.1%). • The most adverse reaction to using sanitizers was skin dryness (55.1% and 63.6% among Saudi and Egyptian HCWs, respectively).
Shadi et al [53], 2022	MCP that included Egypt	Cross-sectional; PPE use and hand hygiene	HCWs	154	66.9% used N95, N98, or a surgical mask, and 86.4% had good hand hygiene	• A total of 77.3% had all the PPE and protective measures they needed. • A total of 66.2% had been recently educated on infection control. • A total of 20.8% always used a standard surgical mask and thought that it was sufficiently protective, 12.3% used either mask according to their availability, and none of the participants refrained from using masks.
Deressa et al [54], 2021	Ethiopia	Cross-sectional; availability and use of PPE and satisfaction with PPE	HCWs	1134	—	• Most (77%) of the HCWs reported that their hospital did not have adequate PPE. A critical shortage of N95 respirators was particularly reported; it only increased from 13% to 24% from before to during the COVID-19 pandemic. • The use of N95 increased from 9% to 21% from before to during the COVID-19 pandemic. Almost 72% of the respondents were dissatisfied with the availability and use of PPE in their hospital. • The independent predictors of the respondents' satisfaction levels with PPE were HCWs who reported that PPE was adequately available in the hospital and preparedness to provide care to COVID-19 cases.
Oladele et al [55], 2021	Nigeria	Cross-sectional and mixed methods; availability and use of PPE	HCWs	258	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• Only 22.1% of HCWs had regular access to PPE, and only 20.6% had access to N95 face masks compared with other PPEs. Male HCWs and those working at secondary or tertiary facilities had access to N95 face masks. • Facilitators of PPE use were the leadership quality of the hospital head and donation of PPE to the facilities, whereas barriers to PPE use included a limited supply of PPE as well as the facility's infrastructural and operational challenges.
Ashour et al [56], 2021	MCP—Egypt and Morocco	Cross-sectional; challenges and difficulties of using PPE	Ophthalmologists	172	—	• The analysis of the responses showed that most ophthalmologists used face masks without substantial problems during their examinations, whereas face shields followed by protective goggles were the most inconvenient PPE in the current ophthalmic practice. • Moreover, most (77.3%) noticed an increase in their examination time when using PPE. A considerable proportion (40.7%) stopped using one or more of the PPE because of inconvenience or discomfort.
Foula et al [57], 2021	Egypt	Cross-sectional; effect of wearing PPE on performance and decision-making	Physicians	272	—	• Results indicated that comfort, vision, and communication were significantly reduced because of PPE wearing in all physician groups (81.1%, 88.7%, and 75.5%, respectively). • In contrast, the handling of instruments was not significantly affected in the second group only. Moreover, decision-making and the rate of complications were not significantly affected.
Hajjij et al [58], 2020	Morocco	Cross-sectional; PPE and headaches	HCWs	155	—	• The overall prevalence of headaches related to PPE was 62%. It was experienced de novo by 32.9% of participants, whereas it was an aggravation of a preexisting headache in 29% of participants. • Working >8 hours per shift during the pandemic was correlated to de novo headache (*P*=.008). The profession of physician and working >12 hours per shift were correlated with aggravated headaches. HCWs experienced moderate discomfort, blurred vision, and reduced concentration. They judged their professional performance to be mildly reduced by the use of PPE.
Nwosu et al [59], 2021	Nigeria	Cross-sectional; impact of different face masks on comfort	HCWs	66	—	• HCWs wore masks for periods ranging from 68 to 480 minutes. The discomfort experienced with the use of the N95 mask was greater than with the surgical mask. • No significant change in arterial oxygen saturation was observed with the use of either mask type, and the tight strapping of the N95 mask was perceived as a contributor to the discomfort experienced with mask use.
Marraha et al [60], 2021	Morocco	Cross-sectional; skin reactions to PPE use	HCWs	273	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• A total of 80% of HCWs had adverse reactions, including skin problems, after wearing goggles (58%), wearing surgical masks and respirators (57%), handwashing and wearing gloves (45%), wearing a face shield (23%), and wearing protective clothing (11%). • Bleach immersion was highly significantly associated with hand reactions, whereas hand cream use more than twice daily was associated with fewer reactions. • The skin reactions were related to goggle use, wearing masks and N95 respirators was significantly associated with longer use duration, and adverse reactions to regular use of protective clothing were related to the frequency of its use per shift.
Jazieh et al [61], 2020	6 countries; Egypt, Algeria, and Morocco	Multicountry survey; behavioral response	Patients with cancer	1012	Adherence to handwashing (77%), keeping distance from others (67%), mask use (77%), and hand hygiene with hand sanitizer (69%) and soap (81%)	• Patients were worried about contracting the virus strongly (33%) or mildly (48%), and most (>80%) reported avoiding the following actions: hand shaking, hugging and kissing, social gatherings, meeting friends, and visiting markets. • Some reported adopting healthier diets (35%), using dietary supplements (18%), and reciting the Quran (61%) or supplications (75%). • Approximately 23% would choose not to show up for a scheduled medical appointment, and 43% had appointment cancellations at the request of the medical team (31%) or the patients themselves (12%). • Moreover, 84% preferred web-based medical appointments over regular visits.
Andarge et al [62], 2020	Ethiopia	Cross-sectional and facility-based; intention and practice of PPMs[h]	Adults with chronic conditions	806	52% and 76.3% intended to practice and had ever practiced PPMs	• Participants' subjective norms and perceived behavioral control were the factors associated with their intention. • Good knowledge and a positive attitude were found to be significant factors associated with the participants' actual practice of PPMs among other independent factors.
Mostafa and Hegazy [63], 2020	Egypt	Cross-sectional, observational study	Patients of dermatology	62	—	• There was an overall satisfaction and future use score of 91% among the interviewed patients who received teledermatology services; a usefulness score of 93.7%; interface and interaction quality scores of 85.9% and 87%, respectively; ease and use learnability score of 87.8%; and a reliability score of 86.7%.
Larebo and Abame [64], 2021	Ethiopia	Cross-sectional; face mask use and associated factors	University students	764	89.5% had good practice of face mask use	• Overall knowledge of the students was 29.2%, and their attitude was 88.1%. • Students from the College of Natural and Computational Sciences and students having good knowledge were found to be independently associated with face mask use.
Nalunkuma et al [65], 2022	Uganda	Cross-sectional; patterns of double mask use	Medical students	348	Only 20.5% reported double masking	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• A total of 68.7% believed that double masking was superior to single masking for COVID-19 prevention and control. • Those with a past COVID-19 positive test and those who believed that double masks had a superior protective advantage were more likely to double mask. • The lack of trust in the quality of masks (46.5%) was the most frequent motivation for double masking, whereas excessive sweating (68.4%), high cost of masks (66.4%), and difficulty in breathing (66.1%) were the major barriers.
Aronu et al [66], 2020	Nigeria	Cross-sectional; perception of masking in children	Mothers	387	—	• Only 44.7% of the mothers perceived masking in children as an appropriate measure for the prevention of COVID-19, and the frequent reasons for the inappropriateness of face masks in children given by most (55.3%) of the mothers included perceived difficulty in breathing (38.5%) and the child's readiness to take the masks off (29.3%). • A significantly higher proportion of children whose mothers were aged ≥35 years would wear face masks (64.2%) when compared with 31.7% of those whose mothers were aged <30 years. Similarly, 51% of the children who were aged >1 year would wear a face mask compared with 20.5% of those aged 8 days to 1 year. • The children whose mothers were aged <30 years were approximately 4 times less likely to wear a face mask when compared with those whose mothers were aged ≥35 years. The children whose fathers had attained tertiary education were approximately twice less likely to wear face masks when compared with those whose fathers had attained a secondary education or lower.
Haftom and Petrucka [67], 2021	Ethiopia	Cross-sectional; face mask use	Quarantined adults	331	46% did not wear a face mask when leaving home	• Face mask use was significantly associated with the knowledge score, employment status, gender, age, and educational status of the study participants.
Deressa et al [68], 2021	Ethiopia	Cross-sectional; social distancing and preventive measures	Government employees	1573	96% wore face masks, 94.5% practiced frequent handwashing, and 89.5% practiced physical distancing	• A total of 94.8% avoided close contact with people, including hand shaking; 95.6% consistently followed government recommendations; 88.1% avoided mass gatherings and crowded places; 71.8% restricted movement and traveling; and 35.6% stayed home. • A total of 80% perceived that consistently wearing a face mask was highly effective in preventing coronavirus infection, and the perception varied by region (Oromia residents being less likely to have good perceptions). • A total of 57% perceived that the policy measures in response to the pandemic were inadequate.
Dzisi and Dei [69], 2020	Ghana	Cross-sectional, roadside observer survey; adherence to social distancing and mask use	Commuters	850	98% of buses complied with the social distancing guidelines	• The policy on face masks was complied with only partially in most vehicles. • A total of 12.6% of the vehicles had <3 commuters without face masks, whereas 21.3% of buses had <3 people with face masks.
	Ghana			500	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
Agyemang et al [70], 2021		Cross-sectional; perception and mask use	Commercial drivers			- Most drivers had a high vulnerability perception to COVID-19. It further emerged that older drivers in particular consistently wore face masks and insisted on other persons in their commercial vehicles to follow suit. - Sociodemographic factors and the need to ensure one's safety and that of loved ones were critical determinants of face mask use among surveyed drivers.
Natnael et al [71], 2021	Ethiopia	Cross-sectional; knowledge, attitude, and frequent hand hygiene practices	Taxi drivers	417	66.4% had good frequent hand hygiene practices	- Good knowledge and positive attitude were reported in 69.8% and 67.6% of the drivers, respectively. - Educational level, place of residence, and attitude toward COVID-19 prevention were factors associated with good knowledge about COVID-19. Furthermore, age of >30 years, a secondary education or higher, income, and knowledge about COVID-19 in taxi drivers were factors associated with a positive attitude toward COVID-19 prevention. Moreover, attitude toward COVID-19 and educational level were the factors associated with good frequent hand hygiene practices.
Mboowa et al [72], 2021	Uganda	Cross-sectional; knowledge, attitudes, and practices regarding face mask use	High-risk groups	644	—	- Most had heard about COVID-19 (99.7%) and believed that face masks were protective against it (87.3%), whereas 67.9% reported having received information on face mask use. - Food market vendors and those with no formal education were 0.5 and 0.3 times less likely to have received information about face mask use than hospital workers and those who had completed secondary school, respectively. - Those who had received information on face mask use were 2.9 and 1.8 times more likely to own face masks and perceive them as protective, respectively. Food market vendors were 3.9 times more likely to reuse their face masks than hospital workers.
Fielmua et al [73], 2021	Ghana	Cross-sectional, observational study; hand hygiene and safety behaviors	Shoppers and shopkeepers	751	91.3% of the customers did not practice handwashing, and 84.2% did not wear face masks	- It was observed that adherence to COVID-19 safety protocols at shopping centers was very poor, and in 78% of the shops observed, no shop attendant wore a mask. - Despite the provision of handwashing facilities and widespread advocacy to minimize COVID-19 infections, the citizenry, especially the youth, demonstrated a poor attitude toward safety measures. Nonadherence to COVID-19 protocols was higher in shops where there was no pressure to conform to the protocols.
Ameme et al [74], 2021	Ghana	Observational study; hand hygiene and face mask wearing practices	Shop patrons	800	81.6% wore face masks, 12.3% performed hand hygiene, and 11.5% adhered to both measures	- A total of 72.3% of patrons wore face masks appropriately, whereas appropriate handwashing was recorded among only 10.1%. - Compared with inappropriate handwashing, appropriate handwashing was negatively associated with adherence to infection and control guidelines.
Yigzaw et al [75], 2021	Ethiopia	Observational cross-sectional study; handwashing practice	Bank visitors	415	—	

Study	Study setting or country	Study type and key measures	Study population	Sample size, N	Adherence or noncompliance rates	Relevant findings
						• Most (93.5%) heard and watched proper handwashing practice. The proportion of proper handwashing performance was 21.4% before the demonstration, but after the demonstration, it increased to 82.2%. • Older age, being married, and higher education were associated with proper handwashing practice. Overall, there was a significant change in handwashing practice after the demonstration.

[a] MCP: multicountry paper.

[b] Not available.

[c] DRC: Democratic Republic of the Congo.

[d] WASH: water, sanitation, and hygiene.

[e] PPE: personal protective equipment.

[f] HCW: health care worker.

[g] FFP2: filtering face piece 2.

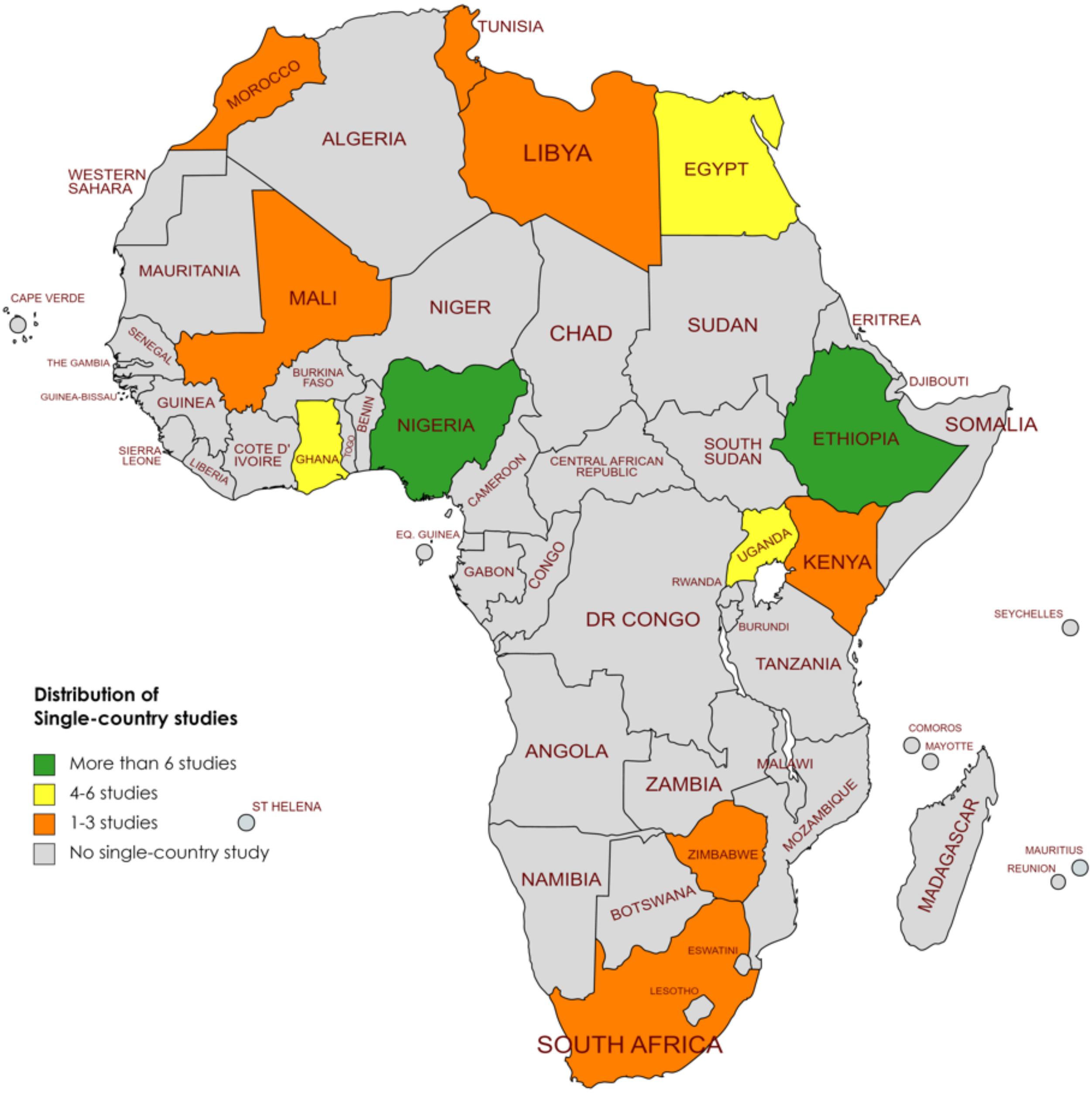
[h] PPM: personal protective measure.

Distribution and Characteristics of the Studies

Of the 58 analyzed studies that primarily reported on PPMs against COVID-19 in Africa, 47 (81%) were single-country studies and were conducted in only 12 of the 54 African countries, whereas the remaining 11 (19%) involved multiple countries (Figure 2). Of the single-country studies, 36% (21/58) were from East Africa, 21% (12/58) were from West Africa, 17% (10/58) were from North Africa, 7% (4/58) were from Southern Africa, and none were from Central Africa. Ethiopia (16/58, 28%), Nigeria (7/58, 12%), and Egypt (5/58, 9%) were the top contributors and altogether produced 48% (28/58) of the analyzed studies. Thus, no single-country studies regarding COVID-19 PPMs had been conducted in 42 African countries at the time of our literature search (Figure 2).

The 58 analyzed studies included 51 (88%) quantitative studies, 5 (9%) qualitative studies, and 2 (3%) mixed methods studies, and their sample sizes ranged from 20 to 206,729. The 4 themes were represented as follows: knowledge and perception of PPMs (21/58, 36%), mask use (37/58, 64%), physical and social distancing (17/58, 29%), and handwashing and hand hygiene (19/58, 33%), considering that most studies covered more than one theme. Moreover, 34% (20/58) of the analyzed studies were conducted among health care workers (HCWs), 33% (19/58) were conducted among the general public, 5% (3/58) were conducted among patients with comorbidities, 3% (2/58) were conducted among university students, and 22% (13/58) were conducted among other groups. The studies were published between 2019 and 2022, and their overall quality was generally good, meaning that the included studies satisfied most of the criteria. However, lower scores on item 4 (nonresponse bias) and item 5 (appropriateness of statistical methods used) were noted among several quantitative studies (9/51, 18% and 13/51, 26%, respectively), with a similar trend observed among qualitative studies, as detailed in Multimedia Appendix 2.

Figure 2. Country distribution of single-country studies. DR: Democratic Republic.



Knowledge and Perception of PPMs and Associated Factors

The results of the knowledge and perception of PPMs and other domains are summarized and presented in Table 3. Among the general public, higher rates of knowledge of COVID-19 preventive measures (>60%) were reported in Western Uganda [18] and Northwest Ethiopia [20]. Higher rates of good attitudes and perceptions were also reported in Western Uganda [18] and Egypt [21], but a lower rate of perceived benefits of preventive measures was reported in Addis Ababa, Ethiopia [22]. Similar findings of good perception were also reported in the Greater Kampala Metropolitan area of Uganda and in other 6 countries (Botswana, Kenya, Malawi, Nigeria, Zambia, and Zimbabwe), where a great majority of the residents believed that face masks were effective against COVID-19 spread and infection [23,24]. However, most individuals in Nigeria believed that COVID-19 was more of a hoax than a reality, and in several countries, preventive measures such as social distancing and face masking were perceived as imported policies that negatively affected their compliance with preventive measures to curb the spread of the disease [25,26].

Table 3. Levels and associated factors of personal protective measure (PPM) knowledge and perceptions and the practice of face mask use, social and physical distancing, and hand hygiene.

		Knowledge and perceptions of PPMs	Face mask use	Social and physical distancing	Handwashing and hand hygiene
General public					
	Levels	• Higher rates (>60%) of knowledge of PPMs [18,20] • Higher rates (>60%) of good attitude and perception [18,21,23,24] • Lower rates (<60%) of good attitude and perception [22,25,26]	• Higher rates (>60%) [24,27,31,32] • Lower rate (<60%) [18,20,22,27,29,39]	• Higher rate (>60%) information not available • Lower rates (<60%) [20-22,28-30,39]	• Higher rate (>60%) information not available • Lower rates (<60%) [33,37-39]
	Associated factors	• Educational level, age, frequency of PPE[a] practice, and social media exposure [18,23,24]	• Gender, age, educational level, marital status, working status, profession, place or community of residence, knowledge and attitude, strictness of containment and health policies, perceived risk and barriers, cues to action, and self-efficacy, among others [22,27,29,32]	• Gender, age, educational level, working status, place or community of residence, family size, knowledge and attitude, strictness of containment and health policies, perceived risk and barriers, cues to action, and self-efficacy, among others [20-22,28-30]	• Gender, age, educational level, marital status, profession, place or community of residence, knowledge and attitude, exposure to handwashing guidelines, strictness of containment and health policies, type and availability of water sources, and perceived risk and barriers, among others [19,22,29,33]
Health care workers					
	Levels	• Higher rates (>60%) of knowledge and attitude [40-42] • Lower knowledge rate (<50%) [43]	• Higher rates (>60%) [49-53] • Lower rates (<60%) [41,45-48]	—[b]	• Higher rates (>60%) [40,49,52,53] • Lower rates (<60%) [45-47]
	Associated factors	• Age, practice location, PPE attitude, and perceived threat [42,43]	• Gender, educational level, work experience, medical specialty (being a nurse or midwifery professional), hours of work, previous training on COVID-19 prevention and PPE use, perceived risk and attitude toward COVID-19, feedback on safety, having COVID-19 management guidelines, and ease and safety when using standard precautions, among others [44-49,55]	—	• Educational level, gender, work experience, medical specialty (being a nurse or midwifery professional), previous training on COVID-19 prevention and PPE use, perceived risk and attitude toward COVID-19, feedback on safety, having COVID-19 management guidelines, and ease and safety when using standard precautions, among others [45-47,49]
Other groups					
	Levels	• Higher rates of knowledge, good attitude, and perception among university students [64,65], patients with chronic diseases [61], government employees [68], taxi drivers [70,71], and market vendors [72] • Poor perception among mothers [66]	• Higher rates (>60%) among patients with chronic diseases [61,62], university students [64], government employees [68], patrons of convenience shops [74], and taxi drivers [70] • Lower rate (<60%) among medical students [65], children [66], quarantined individuals [67], commuters [69], and shopping centers [73]	• Higher rates (>60%) among patients with chronic diseases [61,62], government employees [68], and commuters [69] • Lower rate (<60%) information not available	• Higher rates (>60%) among patients with chronic diseases [61,62], government employees [68], taxi drivers [71], and bank visitors [75] • Lower rate (<60%) among shopping centers [73] and patrons of convenience shops [74]

	Knowledge and perceptions of PPMs	Face mask use	Social and physical distancing	Handwashing and hand hygiene
Associated factors	• Age, educational level, income, and residence (for taxi drivers and government employees) [68,70,71]	• Age, gender, knowledge, attitude, field of study (for university students), educational level, employment status, history of having COVID-19, and perceived benefit and susceptibility, among others [62,64-67]	• Knowledge and attitude [62]	• Age, marital status, knowledge, educational level, attitude, and adherence to infection and control guidelines, among others [62,71,74]

[a] PPE: personal protective equipment.

[b] Not available.

Among African HCWs, higher rates of knowledge and attitude regarding the use of PPE were reported in Ethiopia [40,41] and Egypt [42]. However, a lower knowledge rate of PPE use (<30%) was reported in Nigeria, and in the early stages of the pandemic, a large number of frontline HCWs in Tunisia had not received official training on the correct use of PPE as many believed that they needed additional training [44]. In Egypt, although a substantial proportion (>80%) of house officers (fresh medical graduates doing their 1-year training in different specialties) had good PPE attitudes, <30% had good preparedness and willingness to participate in COVID-19 management and care [42].

Among other groups, only 3% of patients with chronic conditions such as cancer in Egypt, Algeria, and Morocco knew someone who had a COVID-19 infection, but most were worried about contracting the virus [61]. In Ethiopia, although university students had higher rates (>85%) of good attitude toward face mask use, their overall knowledge about mask use was very low (<30%) [64], which contrasts with medical students in Uganda with a better knowledge rate and where close to 70% agreed on the superiority of double masking over single masking for COVID-19 prevention and control [65]. In Nigeria, >50% of mothers perceived masking in children as not an appropriate preventive measure against COVID-19 because of the perceived difficulty in breathing and discomfort among children [66]. Notably, government employees in Ethiopia were reported to have higher rates of good perception of COVID-19 PPMs, whereby approximately 80% perceived consistent mask wearing as highly effective against COVID-19 spread and infection [68]. High COVID-19 vulnerability perception was also reported among commercial drivers in Ghana, mostly among older drivers, who consistently wore face masks and insisted on other persons in their vehicles doing the same [70]. Similar findings were reported among Ethiopian drivers who had good knowledge and positive attitudes [71]. Furthermore, most of the high-risk individuals in Uganda, including market vendors, had received information on face mask use and believed that face masks were protective against COVID-19. Moreover, those who had received information on face mask use were more likely to own face masks and perceive them as protective despite market vendors being more likely to reuse face masks than hospital workers [72].

As possible predictors, educational level, age, income, residence, frequency of PPE practice, PPE attitude, social media exposure, and perceived threat were associated with knowledge and perception of PPMs among the general population, HCWs, and other groups (taxi drivers and government employees) [18,23,24,42,43,68,70,71]. Moreover, practice location was notably significantly associated among HCWs [43] (Table 3).

Mask Use and Associated Factors

Among the general public, largely lower adherence rates (range 20.3%-59.4%) of face mask use were reported in various African countries, including Uganda, the Democratic Republic of the Congo, Somalia [18,27], and Ethiopia [20,22,29]. Notably, adherence rates to mask use were reported to be higher (>60%) after the lifting of lockdown restrictions and in countries where mask use was mandatory, such as Mozambique [27], Uganda [24], Morocco [31], and South Africa [32]. In addition, reusable cloth masks, which are more cost-beneficial and environmentally friendly, were the most used mask type [27]. Interestingly, in South Africa, the prevalence of close others' mask wearing was reported to affect mask use, and older adults had poor mask use practices despite having a higher mortality risk [32]. In addition, the prevalence of mask wearing was noted to have increased substantially (50% to 74%) from May 2020 to August 2020 as COVID-19 cases increased and lockdown restrictions were eased, but staying at home, physical distancing, and social distancing decreased [32]. Regarding used mask disposal, poor disposal practices were reported in several African countries, including Morocco, where most threw their used masks and gloves in their house trash or trash bins, posing a transmission risk to sanitary workers or stray animals [31].

Among HCWs, varying rates of PPE use were reported in different regions and countries. Generally, lower rates (<60%) were reported in the Northwest [45], Northeast [46], Eastern [47], and Amhara regions [41] of Ethiopia, as well as in Egypt [48]. Nevertheless, higher rates (>60%) were reported in Nigeria [49], Libya [50], and 3 multicountry surveys [51-53]. The main challenges reported were inadequate PPE and side effects. The lack of PPE and, thus, the reuse of single-use PPE, especially facial protective shields and masks, were reported in several African countries, including Tunisia [44], Ethiopia [40,54], Egypt [48], and Nigeria [55], and in a multicountry survey including 4 North African countries [51].

Several side effects because of the use of PPE were also reported among African HCWs, including skin problems; heat; thirst; pressure areas; headaches; inability to use the bathroom; extreme exhaustion; discomfort; and reduced vision, concentration, and performance during or after wearing PPE [40,44,51,56-60].

Moreover, such side effects were associated with longer shift durations, the frequency of use, and medical specialty [51,58-60], and the most affected body areas from wearing PPE were the hands, the auricular area, the nasal bridge, the cheeks, and the whole face [52]. Notably, the most reported adverse reactions particularly because of using sanitizers were skin dryness, skin irritation, and ocular irritation [40,52]. Moreover, bleach immersion was reported to be highly associated with hand reactions, whereas hand cream use more than twice daily was associated with fewer reactions [60]. In contrast, a recent multicountry survey that included Egypt indicated that >70% of HCWs had all the PPE and protective measures they needed, >60% had been recently educated on COVID-19 infection control, and none of the interviewed HCWs refrained from using face masks [53].

Among other population groups, strict adherence to face mask use in public areas was reported among patients with cancer in a multicountry survey that included Egypt, Algeria, and Morocco [61]. In Ethiopia, >50% of adults with chronic conditions intended to practice and had ever practiced the recommended personal preventive measures against COVID-19 [62]. University students in Ethiopia were reported to have a higher adherence rate (>80%) of mask use [64]. Furthermore, approximately 20% of medical students in Uganda practiced double masking, where the lack of trust in the quality of masks was the most compelling factor for double masking [65]. Nonetheless, excessive sweating, the high cost of face masks, and difficulty in breathing were the major barriers to double masking among these medical students [65]. Moreover, poor adherence to mask use was highlighted among quarantined individuals in Ethiopia, where nearly half of them did not wear a face mask when leaving home [67]. However, high rates (>80%) of mask use were documented among government employees in Ethiopia [68]. Similar findings of high mask use (>70%) were also observed among taxi drivers [70] and patrons of convenience shops in Ghana [74] but with contrasting observations among commuters and in shopping centers, where less compliance with face mask use was reported [69,73].

Adherence to mask use was associated with gender, age, educational level, marital status, working status, profession, place or community of residence, knowledge and attitude, history of having COVID-19, perceived benefit, strictness of containment and health policies, subjective norms, perceived risk, barriers, cues to action, and self-efficacy among the general public, HCWs, and other groups (Table 3) [22,27,29,32,44-49,55,62,64,65,67,70]. Moreover, work experience, medical specialty (being a nurse or midwifery professional), hours of work, previous training on COVID-19 prevention and PPE use, having COVID-19 management guidelines, and ease and safety when using standard precautions were outstanding predictors among HCWs [44-49,55], and the field of study was a strong predictor of mask use among university students [64]. Surprisingly, mask use among children in Nigeria was highly dependent on the mother's opinions and characteristics, whereby it was associated with the mother's age, the age of the child, and the parental level of education [66].

Social and Physical Distancing and Associated Factors

Generally, lower adherence rates (range 18%-59%) of social distancing were reported in various African countries, including Egypt [21], South Africa [28], and Ethiopia [20,22,29,30]. Nonetheless, physical distancing policies disrupted social life and infringed on people's sociocultural rights, causing adverse socioeconomic and health consequences, especially for low-income urban or suburban slum dwellers [26,34]. Although the imposition of COVID-19 distancing regulations led to a substantial decrease in extrahousehold social contacts (close physical and conversational contacts) in several African countries, including South Africa, there was ongoing contact within intergenerational households, highlighting a potential limitation of social distancing measures in protecting older adults [35]. In contrast, such restrictive policies improved feeding habits through increased meal planning and selection and preparation of healthy foods among residents of various countries [36].

Regarding the implementation and adoption of physical distancing measures, despite the implementation of various mitigation measures, the internally displaced people in Mali still faced several challenges, including the proximity in which internally displaced people live, the lack of toilets and safe water, and the lack of financial resources [38]. Similar findings were reported among prisons in Zimbabwe, where there were several challenges in the adoption of COVID-19 PPMs, such as severe congestion, interrupted water supply, outdated infrastructure, and inadequate hygiene and sanitation [39]. Moreover, although prisoners had adequate COVID-19 awareness and prison health professionals received training on COVID-19 control measures, PPE supply was inadequate, with no routine COVID-19 testing in place beyond thermal scanning; isolation measures were compromised by accommodation capacity issues; and social distancing was impossible during meals and at night [39].

Among other population groups, strict adherence to social and physical distancing was documented among patients with cancer, and most preferred web-based medical appointments over regular visits. In addition, some adopted healthier diets, used dietary supplements, and recited the Quran or supplications [61]. Similar findings of good practice of social distancing were also reported among patients with chronic conditions in Ethiopia [62]. Moreover, in Egypt, patients preferred teledermatology services to the usual physical clinic visits as they perceived them as reliable and safe during the pandemic [63]. Government employees in Ethiopia also had higher rates (>80%) of good practice of physical distancing [68], and the same applied to commuters in Ghana, who had high compliance rates with social distancing guidelines [69].

Adherence to social and physical distancing was associated with gender, age, educational level, working status, place or community of residence, family size, knowledge and attitude, strictness of containment and health policies, perceived risk and barriers, cues to action, and self-efficacy among the general public and other groups (patients with chronic diseases; Table 3) [20-22,28-30,62].

Handwashing and Hand Hygiene and Associated Factors

Regarding community adherence to hand hygiene, lower rates (<60%) were reported in 12 sub-Saharan countries, where the likelihood of handwashing mainly varied with the level of concern about COVID-19 [33]. In resource-restricted settings, a recent study indicated that >60% of the slum dwellers in Nairobi, Kenya, had limited water, sanitation, and hygiene facility accessibility and opportunity, making adherence to COVID-19 PPMs impossible [37].

Concerning hand hygiene adherence among African HCWs, varying rates were reported in different regions and countries. Generally, lower rates (<60%) were reported in the Northwest [45], Northeast [46], and Eastern [47] regions of Ethiopia. Nevertheless, higher rates (>60%) were reported in Nigeria [49], Southwest Ethiopia [40], and 2 multicountry surveys [52,53].

Regarding other groups, strict adherence to proper hand hygiene was reported among patients with chronic conditions in various African countries, including Egypt, Algeria, Morocco [61], and Ethiopia [62]. A similar observation was made among government employees [68] and taxi drivers [71] in Ethiopia, both of whom had higher rates (>60%) of frequent handwashing and hand hygiene as a means of protection against COVID-19. However, poor adherence to COVID-19 safety protocols at shopping centers in Ghana was reported, whereby, although shops complied with providing handwashing facilities, most of the customers did not practice handwashing before entering the shops and did not wear face masks during shopping, and neither did the shop attendants [73]. Similarly, a very low rate (10%) of appropriate handwashing was reported among patrons of convenience shops in Accra, Ghana [74]. In contrast, an increase in proper handwashing performance was reported among bank visitors in Ethiopia after watching a handwashing demonstration [75].

Handwashing and hand hygiene during the COVID-19 pandemic was associated with gender, age, educational level, marital status, profession, place or community of residence, knowledge and attitude, exposure and adherence to handwashing guidelines, strictness of containment and health policies, type and availability of water sources, and perceived risk and barriers among the general public, HCWs, and other groups (Table 3) [19,22,29,33,45-47,49,62,71,74,75]. Moreover, work experience, medical specialty (being a nurse or midwifery professional), previous training on COVID-19 prevention and PPE use, feedback on safety, having COVID-19 management guidelines, and ease and safety when using standard precautions were notable predictors among HCWs [45-47,49].

Discussion

Principal Findings

To our knowledge, this is the first systematic review to evaluate PPMs against COVID-19 among various population groups in Africa. This systematic review has important implications as it reflects cognitive behavioral issues (knowledge and practice) regarding PPMs in some African countries during an infectious disease outbreak. Future outbreaks or waves of COVID-19 may force people to use PPMs again. The review used a multidimensional approach involving the systematic evaluation of evidence based on region, country, and population group. Moreover, comprehensive coverage of the literature was attained, and a reproducible search methodology was applied using a predefined framework, all of which are strengths of this review.

Among the general community, the review showed varying levels of knowledge, attitudes, and perceptions, which in turn influenced the practice levels of and compliance with COVID-19 PPMs, especially face mask use, hand hygiene, and physical and social distancing. This finding is in agreement with a previous study from sub-Saharan Africa [15], and similar findings have been reported in other regions where communities' cognition directly affected the practice and uptake of COVID-19 PPMs [76]. Nonetheless, the observed difference in the practice and adherence to PPMs across African countries may be due to the differences in COVID-19 control policies, income (gross domestic product), and the situation of the pandemic among the countries. Notably, the compliance rates of face mask use reported in most African communities were generally lower compared with those reported in studies from high-income countries [77,78]. This may partly be explained by the inability to afford to buy face masks and the differences in the strictness of such preventive measures [22,29,79]. Nevertheless, poor adherence to face mask use was also reported in some high-income countries such as Australia, Norway, and Sweden [80], the reasons for which may be other than just the inability to afford face masks. Moreover, lower rates of handwashing and hand hygiene were also reported in several African communities, especially among low-income urban and slum dwellers. This was partly due to a lack of safe and clean water in slum communities [37]. Moreover, buying soap or hand sanitizers was an additional financial constraint for low-income urban dwellers and, thus, may be seen as a luxury.

The results indicate a reduction in the rates of PPM practice (mainly mask use and social and physical distancing) noted in several African countries following the lifting of restrictive lockdown measures and the rollout of COVID-19 vaccination programs. This can be partly explained by pandemic fatigue as more people become demotivated and exhausted to follow the recommended infection prevention and control measures owing to the prolonged impact and existence of COVID-19 [81]. As COVID-19 PPMs complement the vaccination protective advantage, this implies a need for continued community sensitization and education programs to rectify the reluctance to practice PPMs amid the relaxation of preventive restrictions. Moreover, prompt management of infodemics in the current and future infectious outbreaks is needed to address the misinformation about PPMs [82].

Among African HCWs, generally good knowledge of PPE use was reported but with varying levels of practicing PPMs, and the low practice rates were attributed mainly to the lack of PPE and the side effects of prolonged PPE use. With HCWs being at the frontline of screening and managing suspects and patients with COVID-19, the lack of PPE increases the risk of infection when doing their work. Nonetheless, the lack of PPE has also been documented in other countries and regions outside Africa [83]. Furthermore, this review showed that most patients with comorbidities in Africa reported strict adherence to COVID-19 PPMs, which may be due to their perceived high vulnerability to COVID-19 infection and complications. Other studies outside Africa have reported similar findings among patients with comorbidities [84,85].

The study findings show that several cognitive (including knowledge, attitude, and perception), demographic, and socioeconomic factors were associated with the practice of and compliance with COVID-19 PPMs among African communities. COVID-19 being a newly evolving disease with varying cross-cutting impacts implies a need for consideration of such cognitive, demographic, and socioeconomic differences in the design of targeted response measures against the pandemic. Nonetheless, similar findings on the association of sociodemographics with the practice of COVID-19 PPMs have been reported in other regions outside Africa [80].

This review has some practical recommendations for improving COVID-19 control programs in Africa. Efforts are needed to improve the local capacity to produce and supply PPE, especially to HCWs, as the lack of PPE was the main barrier to PPE use. In the early phase of the pandemic, most countries were overwhelmed by the increased demand for PPE, which disrupted the global supply chain, and this had dire consequences for countries with inadequate local manufacturing and supply capacity [83]. In addition, providing free or subsidized face masks and soap, especially to low-income earners, would be a helpful strategy for improving PPM practice and adherence. Moreover, the consideration of vulnerable groups such as low-income urban dwellers and internally displaced people and targeted responses tailored to their socioeconomic dynamics are paramount for effective pandemic control programs. Knowledge and perception influenced the practice of PPMs, implying a need for continuous infodemic management, community education, and sensitization, and this should be tailored to address the existing misconceptions and barriers to PPM adherence. Notably, although several of the analyzed studies (11/58, 19%) evaluated the association between age and the practice of COVID-19 PPMs and showed varying rates and results, no single study focused on exploring COVID-19 PPMs among the older adult population of Africa. Given the known vulnerability of older people to severe COVID-19, efforts are needed to explore this special group to help fully understand their behavioral response to the pandemic, which is vital for guiding targeted responses.

The review also reveals substantial inequalities in terms of research output from different regions of Africa, with PPM studies mostly coming from East, West, and North Africa and only 3 countries (Ethiopia, Nigeria, and Egypt) producing >40% (28/58, 48%) of all the studies. This finding coincides with the study by Nwagbara et al [15], which showed the dominance of East and West Africa in COVID-19 research. The high PPM research output from North and West Africa could be because they were the first regions to record COVID-19 cases in the continent [86]. Although South Africa is known to lead African research with sound and more vibrant research institutions in the continent [14], its contribution to COVID-19 PPM research is far lower, as indicated by the study results. Regarding other African countries, the observed pattern may be explained by the differences in research capabilities and resources. Nonetheless, such research inequalities pose gaps in understanding how such countries and regions respond to the COVID-19 pandemic. This implies a need for more focus, funding, and involvement in behavioral health research, which is as important as clinical research and vital in guiding evidence-based and country-specific or tailored policies and responses in addressing the dynamics of the current COVID-19 pandemic.

Limitations

This systematic review has some limitations. Although we used a comprehensive keyword search strategy, some relevant studies might have been missed as only 3 databases and only English-language articles were considered. In addition, we did not consider gray literature and preprints in this review; thus, they should be considered in future or updated reviews on PPM practice for a more comprehensive search. Although a comprehensive search was performed, no relevant studies were found from 42 of the 54 African countries; thus, the findings might not provide a comprehensive picture of the knowledge and practice of PPMs across the entire continent. Moreover, the findings and conclusions of this review are based on studies that were mostly web-based surveys, which, although this was inevitable because of the restrictive preventive measures and lockdowns, are prone to selection bias based on internet accessibility. Owing to the self-report nature of these surveys, recall and social desirability bias cannot be overlooked. Moreover, assessments of statistical analyses of associations with the practice of COVID-19 preventive measures, as well as meta-analyses, were not performed as these were not the main focus of this review. Despite these limitations, this study provides valuable insights into the facilitators of and barriers to the practice of PPMs in Africa.

Conclusions

This review evaluated the knowledge and practice of COVID-19 PPMs in African countries. The findings, conclusions, and recommendations of this review specifically apply to the included countries and, thus, should be interpreted with caution. The results indicate that African communities, including various population groups, have varying levels of practice and compliance with COVID-19 PPMs, with the lack of PPE (mainly face masks) and side effects of PPE use being the major reasons for poor compliance, especially among HCWs. In addition, various cognitive, sociodemographic, and economic factors were associated with the practice of COVID-19 PPMs. Therefore, this review highlights a need for enhancing the local capacity to produce and supply PPE. The consideration of various cognitive, demographic, and socioeconomic differences, with extra focus on low-income urban dwellers and those who are less advantaged, is vital for inclusive and more effective strategies against the pandemic. Moreover, more focus, involvement, and funding of community behavioral (including protective measures) research is needed to fully understand and address the dynamics of the current pandemic in Africa.

Acknowledgments

This study was supported by internal funding from the Centre for Health Behaviours Research, Chinese University of Hong Kong.

Data Availability

All data generated or analyzed during this study are included in this published paper (and its supplementary information files).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Filled-in PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 27 KB - publichealth_v9i1e44051_app1.docx]

Multimedia Appendix 2
Quality assessment of the included studies using the Mixed Methods Appraisal Tool—version 2018.
[DOCX File , 36 KB - publichealth_v9i1e44051_app2.docx]

References

1. Mehtar S, Preiser W, Lakhe NA, Bousso A, TamFum JM, Kallay O, et al. Limiting the spread of COVID-19 in Africa: one size mitigation strategies do not fit all countries. Lancet Glob Health 2020 Jul;8(7):e881-e883 [FREE Full text] [doi: 10.1016/S2214-109X(20)30212-6] [Medline: 32530422]
2. Mbow M, Lell B, Jochems SP, Cisse B, Mboup S, Dewals BG, et al. COVID-19 in Africa: dampening the storm? Science 2020 Aug 07;369(6504):624-626. [doi: 10.1126/science.abd3902] [Medline: 32764055]
3. Nachega J, Seydi M, Zumla A. The late arrival of coronavirus disease 2019 (COVID-19) in Africa: mitigating pan-continental spread. Clin Infect Dis 2020 Jul 28;71(15):875-878 [FREE Full text] [doi: 10.1093/cid/ciaa353] [Medline: 32227121]
4. Coronavirus update (Live): reported cases and deaths by country, territory, or conveyance. Worldometers. URL: https://www.worldometers.info/coronavirus/ [accessed 2023-03-20]
5. Dyer O. COVID-19: Africa records over 10 000 cases as lockdowns take hold. BMJ 2020 Apr 08;369:m1439. [doi: 10.1136/bmj.m1439] [Medline: 32269023]
6. Wadvalla BA. How Africa has tackled COVID-19. BMJ 2020 Jul 16;370:m2830. [doi: 10.1136/bmj.m2830] [Medline: 32675053]
7. Maeda JM, Nkengasong JN. The puzzle of the COVID-19 pandemic in Africa. Science 2021 Jan 01;371(6524):27-28. [doi: 10.1126/science.abf8832] [Medline: 33384364]
8. El-Sadr WM, Justman J. Africa in the path of COVID-19. N Engl J Med 2020 Jul 16;383(3):e11. [doi: 10.1056/NEJMp2008193] [Medline: 32302075]
9. Ataguba JE. COVID-19 pandemic, a war to be won: understanding its economic implications for Africa. Appl Health Econ Health Policy 2020 Jun;18(3):325-328 [FREE Full text] [doi: 10.1007/s40258-020-00580-x] [Medline: 32249362]
10. Constant A, Conserve D, Gallopel-Morvan K, Raude J. Cognitive factors associated with public acceptance of COVID-19 nonpharmaceutical prevention measures: cross-sectional study. JMIRx Med 2022 May 13;3(2):e32859 [FREE Full text] [doi: 10.2196/32859] [Medline: 35648730]
11. Pan Y, Fang Y, Xin M, Dong W, Zhou L, Hou Q, et al. Self-reported compliance with personal preventive measures among Chinese factory workers at the beginning of work resumption following the COVID-19 outbreak: cross-sectional survey study. J Med Internet Res 2020 Sep 29;22(9):e22457 [FREE Full text] [doi: 10.2196/22457] [Medline: 32924947]
12. Mohadab ME, Bouikhalene B, Safi S. Bibliometric method for mapping the state of the art of scientific production in COVID-19. Chaos Solitons Fractals 2020 Oct;139:110052 [FREE Full text] [doi: 10.1016/j.chaos.2020.110052] [Medline: 32834606]
13. Lou J, Tian SJ, Niu SM, Kang XQ, Lian HX, Zhang LX, et al. Coronavirus disease 2019: a bibliometric analysis and review. Eur Rev Med Pharmacol Sci 2020 Mar;24(6):3411-3421 [FREE Full text] [doi: 10.26355/eurrev_202003_20712] [Medline: 32271460]
14. Volmink J, Dare L. Addressing inequalities in research capacity in Africa. BMJ 2005 Oct 01;331(7519):705-706 [FREE Full text] [doi: 10.1136/bmj.331.7519.705] [Medline: 16195259]
15. Nwagbara UI, Osual EC, Chireshe R, Bolarinwa OA, Saeed BQ, Khuzwayo N, et al. Knowledge, attitude, perception, and preventative practices towards COVID-19 in sub-Saharan Africa: a scoping review. PLoS One 2021 Apr 19;16(4):e0249853 [FREE Full text] [doi: 10.1371/journal.pone.0249853] [Medline: 33872330]

16. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1 [FREE Full text] [doi: 10.1186/2046-4053-4-1]
17. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. Educ Inf 2018 Dec 18;34(4):285-291 [FREE Full text] [doi: 10.3233/efi-180221]
18. Sikakulya FK, Ssebuufu R, Mambo SB, Pius T, Kabanyoro A, Kamahoro E, et al. Use of face masks to limit the spread of the COVID-19 among western Ugandans: knowledge, attitude and practices. PLoS One 2021 Mar 24;16(3):e0248706 [FREE Full text] [doi: 10.1371/journal.pone.0248706] [Medline: 33760882]
19. Szczuka Z, Abraham C, Baban A, Brooks S, Cipolletta S, Danso E, et al. The trajectory of COVID-19 pandemic and handwashing adherence: findings from 14 countries. BMC Public Health 2021 Oct 05;21(1):1791 [FREE Full text] [doi: 10.1186/s12889-021-11822-5] [Medline: 34610808]
20. Hailu W, Derseh L, Hunegnaw MT, Tesfaye T, Angaw DA. Compliance, barriers, and facilitators to social distancing measures for prevention of coronavirus disease 2019 in northwest Ethiopia, 2020. Curr Ther Res Clin Exp 2021;94:100632 [FREE Full text] [doi: 10.1016/j.curtheres.2021.100632] [Medline: 33967298]
21. Bakry HM, Waly EH. Perception and practice of social distancing among Egyptians in COVID-19 pandemic. J Infect Dev Ctries 2020 Aug 31;14(8):817-822 [FREE Full text] [doi: 10.3855/jidc.13160] [Medline: 32903223]
22. Tadesse T, Alemu T, Amogne G, Endazenaw G, Mamo E. Predictors of coronavirus disease 2019 (COVID-19) prevention practices using health belief model among employees in Addis Ababa, Ethiopia, 2020. Infect Drug Resist 2020 Oct 22;13:3751-3761 [FREE Full text] [doi: 10.2147/IDR.S275933] [Medline: 33122922]
23. Iyamu I, Apantaku G, Yesufu Z, Oladele EA, Eboreime E, Afirima B, et al. Is social media, as a main source of information on COVID-19, associated with perceived effectiveness of face mask use? Findings from six sub-Saharan African countries. Glob Health Promot 2022 Sep;29(3):86-96 [FREE Full text] [doi: 10.1177/17579759211065489] [Medline: 35081834]
24. Bukuluki PM, Kisaakye P. The decision to wear a face mask as a protective behavioral measure against COVID-19: survey results from greater Kampala metropolitan area, Uganda. Front Public Health 2021 Oct 21;9:675734 [FREE Full text] [doi: 10.3389/fpubh.2021.675734] [Medline: 34746072]
25. Nnama-Okechukwu CU, Chukwu NE, Nkechukwu CN. COVID-19 in Nigeria: knowledge and compliance with preventive measures. Soc Work Public Health 2020 Sep 01;35(7):590-602. [doi: 10.1080/19371918.2020.1806985] [Medline: 32970541]
26. Kajiita RM, Kang'ethe SM. Social distancing during the sars-Cov2 (COVID-19) pandemic: interpretations and implication in the African context. Perspect Global Dev Technol 2021 Aug 12;20(3):289-312 [FREE Full text] [doi: 10.1163/15691497-12341596]
27. Siewe Fodjo JN, Pengpid S, Villela EF, Van Thang V, Ahmed M, Ditekemena J, et al. Mass masking as a way to contain COVID-19 and exit lockdown in low- and middle-income countries. J Infect 2020 Sep;81(3):e1-e5 [FREE Full text] [doi: 10.1016/j.jinf.2020.07.015] [Medline: 32682733]
28. Sewpaul R, Mabaso M, Dukhi N, Naidoo I, Vondo N, Davids AS, et al. Determinants of social distancing among south Africans from 12 days into the COVID-19 lockdown: a cross sectional study. Front Public Health 2021 May 24;9:632619 [FREE Full text] [doi: 10.3389/fpubh.2021.632619] [Medline: 34109143]
29. Wondimu W, Ejigu A, Ayenew M, Kidnau AW, Niguse W, Geremaw M, et al. Factors associated with coronavirus disease 2019 prevention practices in three zones of southwest Ethiopia: community-based cross-sectional study. Int J Gen Med 2020 Dec 18;13:1603-1611 [FREE Full text] [doi: 10.2147/IJGM.S288067] [Medline: 33376385]
30. Fikrie A, Amaje E, Golicha W. Social distancing practice and associated factors in response to COVID-19 pandemic at west Guji zone, southern Ethiopia, 2021: a community based cross-sectional study. PLoS One 2021 Dec 20;16(12):e0261065 [FREE Full text] [doi: 10.1371/journal.pone.0261065] [Medline: 34928975]
31. Mejjad N, Cherif EK, Rodero A, Krawczyk DA, El Kharraz J, Moumen A, et al. Disposal behavior of used masks during the COVID-19 pandemic in the Moroccan community: potential environmental impact. Int J Environ Res Public Health 2021 Apr 20;18(8):4382 [FREE Full text] [doi: 10.3390/ijerph18084382] [Medline: 33924217]
32. Burger R, Christian C, English R, Maughan-Brown B, Rossouw L. Predictors of mask-wearing during the advent of the COVID-19 pandemic: evidence from South Africa. Transl Behav Med 2022 Jan 18;12(1):ibab132 [FREE Full text] [doi: 10.1093/tbm/ibab132] [Medline: 34865174]
33. Amuakwa-Mensah F, Klege RA, Adom PK, Köhlin G. COVID-19 and handwashing: implications for water use in Sub-Saharan Africa. Water Resour Econ 2021 Oct;36:100189 [FREE Full text] [doi: 10.1016/j.wre.2021.100189] [Medline: 34745865]
34. Iwuoha VC, Aniche ET. COVID-19 lockdown and physical distancing policies are elitist: towards an indigenous (Afro-centred) approach to containing the pandemic in sub-urban slums in Nigeria. Local Environ 2020;25(8):631-640 [FREE Full text] [doi: 10.1080/13549839.2020.1801618]
35. McCreesh N, Dlamini V, Edwards A, Olivier S, Dayi N, Dikgale K, et al. Impact of the COVID-19 epidemic and related social distancing regulations on social contact and SARS-CoV-2 transmission potential in rural south Africa: analysis of repeated cross-sectional surveys. BMC Infect Dis 2021 Sep 08;21(1):928 [FREE Full text] [doi: 10.1186/s12879-021-06604-8] [Medline: 34496771]

36. De Backer C, Teunissen L, Cuykx I, Decorte P, Pabian S, Gerritsen S, Corona Cooking Survey Study Group. An evaluation of the COVID-19 pandemic and perceived social distancing policies in relation to planning, selecting, and preparing healthy meals: an observational study in 38 countries worldwide. Front Nutr 2021 Feb 04;7:621726 [FREE Full text] [doi: 10.3389/fnut.2020.621726] [Medline: 33614693]
37. Kim J, Hagen E, Muindi Z, Mbonglou G, Laituri M. An examination of water, sanitation, and hygiene (WASH) accessibility and opportunity in urban informal settlements during the COVID-19 pandemic: evidence from Nairobi, Kenya. Sci Total Environ 2022 Jun 01;823:153398 [FREE Full text] [doi: 10.1016/j.scitotenv.2022.153398] [Medline: 35092785]
38. Ag Ahmed MA, Ly BA, Diarra NH, Traore FB, Diarra D, Kande IF, et al. Challenges to the implementation and adoption of physical distancing measures against COVID-19 by internally displaced people in Mali: a qualitative study. Confl Health 2021 Dec 04;15(1):88 [FREE Full text] [doi: 10.1186/s13031-021-00425-x] [Medline: 34863236]
39. Mhlanga-Gunda R, Rusakaniko S, Chinyuku AN, Pswarayi VF, Robinson CS, Kewley S, et al. "We sleep 10cm apart so there is no social distancing": COVID-19 preparedness in a Zimbabwean prison complex. Int J Prison Health (forthcoming) 2022 Feb 01. [doi: 10.1108/IJPH-10-2021-0101] [Medline: 35089667]
40. Assefa D, Melaku T, Bayisa B, Alemu S. Knowledge, attitude and self-reported performance and challenges of hand hygiene using alcohol-based hand sanitizers among healthcare workers during COVID-19 pandemic at a tertiary hospital: a cross-sectional study. Infect Drug Resist 2021 Jan 29;14:303-313 [FREE Full text] [doi: 10.2147/IDR.S291690] [Medline: 33542637]
41. Seid Yimer T, Gebrehana Belay H. Knowledge and practice of health care providers towards proper face mask utilization to minimize the extent of COVID-19 infection in Amhara region referral hospitals, Ethiopia. J Multidiscip Healthc 2021 Jul 28;14:1583-1591 [FREE Full text] [doi: 10.2147/JMDH.S306253] [Medline: 34234446]
42. Ahmed Sayed A, Mostafa Ahmed M, Talaat Elsayed I, Saeed S, Inas A, Nouran E, et al. Preparedness and attitude toward personal protective equipment among house officers during COVID-19 pandemic in Egypt. J Public Health (Oxf) 2021 Dec 10;43(4):754-762 [FREE Full text] [doi: 10.1093/pubmed/fdab078] [Medline: 33765145]
43. Alao MA, Durodola AO, Ibrahim OR, Asinobi OA. Assessment of health workers' knowledge, beliefs, attitudes, and use of personal protective equipment for prevention of COVID-19 infection in low-resource settings. Adv Public Health 2020 Aug 24;2020:1-10 [FREE Full text] [doi: 10.1155/2020/4619214]
44. Daghmouri MA, Akremi S, Amouri N, Ouanes S, Jaoua H, Ben Fadhel K. Personal protective equipment: a cross-sectional study in frontline healthcare workers during COVID-19 outbreak in Tunisia. Tunis Med 2020 Aug;98(8-9):633-638. [Medline: 33480018]
45. Kassie BA, Adane A, Abebe Kassahun E, Ayele AS, Kassahun Belew A. Poor COVID-19 preventive practice among healthcare workers in northwest Ethiopia, 2020. Adv Public Health 2020 Oct 16;2020:1-7 [FREE Full text] [doi: 10.1155/2020/7526037]
46. Keleb A, Ademas A, Lingerew M, Sisay T, Berihun G, Adane M. Prevention practice of COVID-19 using personal protective equipment and hand hygiene among healthcare workers in public hospitals of south wollo zone, Ethiopia. Front Public Health 2021 Dec 02;9:782705 [FREE Full text] [doi: 10.3389/fpubh.2021.782705] [Medline: 34926394]
47. Birhanu A, Balis B, Assebe Yadeta T, Bayu M. Personal protective equipment utilization practice and psychological preparedness of health care workers against COVID-19 pandemic in eastern Ethiopia. SAGE Open Med 2021 Oct 10;9:20503121211051925 [FREE Full text] [doi: 10.1177/20503121211051925] [Medline: 34659769]
48. El-Sokkary RH, Khater WS, El-Kholy A, Mohy Eldin S, Gad DM, Bahgat S, et al. Compliance of healthcare workers to the proper use of personal protective equipment during the first wave of COVID-19 pandemic. J Infect Public Health 2021 Oct;14(10):1404-1410 [FREE Full text] [doi: 10.1016/j.jiph.2021.07.017] [Medline: 34344624]
49. Afemikhe JA, Esewe RE, Enuku CA, Ehwarieme TA. Transmission based precaution practices among nurses in Edo state, Nigeria during COVID-19 pandemic. Afr J Reprod Health 2020 Jun;24(s1):98-107. [doi: 10.29063/ajrh2020/v24i2s.14] [Medline: 34077059]
50. Elhadi M, Alsoufi A, Alhadi A, Hmeida A, Alshareea E, Dokali M, et al. Knowledge, attitude, and acceptance of healthcare workers and the public regarding the COVID-19 vaccine: a cross-sectional study. BMC Public Health 2021 May 20;21(1):955 [FREE Full text] [doi: 10.1186/s12889-021-10987-3]
51. Tabah A, Ramanan M, Laupland KB, Buetti N, Cortegiani A, Mellinghoff J, PPE-SAFE contributors. Personal protective equipment and intensive care unit healthcare worker safety in the COVID-19 era (PPE-SAFE): an international survey. J Crit Care 2020 Oct;59:70-75 [FREE Full text] [doi: 10.1016/j.jcrc.2020.06.005] [Medline: 32570052]
52. Mahmoud MR, Almuaili JA, Shaheen SM, Alatifi MN, Abdelaziz OG, Abdelaziz SG. Comparative study upon using sanitizers and personal protective equipment during COVID-19 pandemic between Saudi and Egyptian health care providers. J Pharm Res Int 2021 Mar 19;33(13):61-73 [FREE Full text] [doi: 10.9734/jpri/2021/v33i1331267]
53. Shadi MS, Geneid A, Rubin JS, Ibrahim RA. Infection control, hand hygiene practice and PPE use among phoniatricians and ENT specialists during the COVID-19 pandemic, a UEP survey. Egypt J Otolaryngol 2022 Jan 24;38(1):14 [FREE Full text] [doi: 10.1186/s43163-022-00205-1]
54. Deressa W, Worku A, Abebe W, Gizaw M, Amogne W. Availability and use of personal protective equipment and satisfaction of healthcare professionals during COVID-19 pandemic in Addis Ababa, Ethiopia. Arch Public Health 2021 Aug 17;79(1):146 [FREE Full text] [doi: 10.1186/s13690-021-00668-3] [Medline: 34404464]

55. Oladele DA, Idigbe IE, Musa AZ, Gbaja-Biamila T, Bamidele T, Ohihoin AG, et al. Self-reported use of and access to personal protective equipment among healthcare workers during the COVID-19 outbreak in Nigeria. Heliyon 2021 May;7(5):e07100 [FREE Full text] [doi: 10.1016/j.heliyon.2021.e07100] [Medline: 34031646]
56. Ashour DM, Elkitkat RS, Gabr H, Yosef M, Singh Grewal D, Saleh MI. Challenges of personal protective equipment use among ophthalmologists during the COVID-19 pandemic: a multicenter international study. Eur J Ophthalmol 2022 May;32(3):1398-1405 [FREE Full text] [doi: 10.1177/11206721211028037] [Medline: 34180262]
57. Foula MS, Nwesar FA, Oraby EH, Foula A, Alarfaj MA, Foula HS, et al. Does wearing personal protective equipment affect the performance and decision of physicians? A cross-sectional study during the COVID-19 pandemic. Ann Med Surg (Lond) 2021 Jul;67:102488 [FREE Full text] [doi: 10.1016/j.amsu.2021.102488] [Medline: 34127939]
58. Hajjij A, Aasfara J, Khalis M, Ouhabi H, Benariba JF, El Kettani C. Personal protective equipment and headaches: cross-sectional study among Moroccan healthcare workers during COVID-19 pandemic. Cureus 2020 Dec 13;12(12):e12047 [FREE Full text] [doi: 10.7759/cureus.12047] [Medline: 33447477]
59. Nwosu AD, Ossai EN, Onwuasoigwe O, Ahaotu F. Oxygen saturation and perceived discomfort with face mask types, in the era of COVID-19: a hospital-based cross-sectional study. Pan Afr Med J 2021 Jul 16;39:203 [FREE Full text] [doi: 10.11604/pamj.2021.39.203.28266] [Medline: 34603584]
60. Marraha F, Al Faker I, Charif F, Chahoub H, Benyamna Y, Rahmani N, et al. Skin reactions to personal protective equipment among first-line COVID-19 healthcare workers: a survey in northern Morocco. Ann Work Expo Health 2021 Oct 09;65(8):998-1003 [FREE Full text] [doi: 10.1093/annweh/wxab018] [Medline: 33876214]
61. Jazieh AR, Benbrahim Z, Al-Sukhun S, Abusanad AM, Alorabi MO, El Kinge AR, et al. 1717P behavioral practices of cancer patients during COVID-19 pandemic: a middle east and north Africa study. Ann Oncol 2020 Sep;31:S1009 [FREE Full text] [doi: 10.1016/j.annonc.2020.08.1781]
62. Andarge E, Fikadu T, Temesgen R, Shegaze M, Feleke T, Haile F, et al. Intention and practice on personal preventive measures against the COVID-19 pandemic among adults with chronic conditions in southern Ethiopia: a survey using the theory of planned behavior. J Multidiscip Healthc 2020 Dec 03;13:1863-1877 [FREE Full text] [doi: 10.2147/JMDH.S284707] [Medline: 33299323]
63. Mostafa PI, Hegazy AA. Dermatological consultations in the COVID-19 era: is teledermatology the key to social distancing? An Egyptian experience. J Dermatolog Treat 2022 Mar;33(2):910-915. [doi: 10.1080/09546634.2020.1789046] [Medline: 32602763]
64. Larebo YM, Abame DE. Knowledge, attitudes, and practices of face mask utilization and associated factors in COVID-19 pandemic among wachemo university students, southern Ethiopia: a cross-sectional study. PLoS One 2021 Sep 20;16(9):e0257609 [FREE Full text] [doi: 10.1371/journal.pone.0257609] [Medline: 34543358]
65. Nalunkuma R, Abila DB, Ssewante N, Kiyimba B, Kigozi E, Kisuza RK, et al. Double face mask use for COVID-19 infection prevention and control among medical students at Makerere university: a cross-section survey. Risk Manag Healthc Policy 2022 Jan 20;15:111-120 [FREE Full text] [doi: 10.2147/RMHP.S347972] [Medline: 35087291]
66. Aronu AE, Chinawa JM, Nduagubam OC, Ossai EN, Chinawa AT, Igwe WC. Maternal perception of masking in children as a preventive strategy for COVID-19 in Nigeria: a multicentre study. PLoS One 2020 Nov 19;15(11):e0242650 [FREE Full text] [doi: 10.1371/journal.pone.0242650] [Medline: 33211761]
67. Haftom M, Petrucka PM. Determinants of face mask utilization to prevent COVID-19 pandemic among quarantined adults in Tigrai region, northern Ethiopia, 2020. Clin Nurs Res 2021 Sep;30(7):1107-1112 [FREE Full text] [doi: 10.1177/10547738211013219] [Medline: 33955280]
68. Deressa W, Worku A, Abebe W, Getachew S, Amogne W. Social distancing and preventive practices of government employees in response to COVID-19 in Ethiopia. PLoS One 2021 Sep 07;16(9):e0257112 [FREE Full text] [doi: 10.1371/journal.pone.0257112] [Medline: 34492089]
69. Dzisi EK, Dei OA. Adherence to social distancing and wearing of masks within public transportation during the COVID 19 pandemic. Transp Res Interdiscip Perspect 2020 Sep;7:100191 [FREE Full text] [doi: 10.1016/j.trip.2020.100191] [Medline: 34173465]
70. Agyemang E, Agyei-Mensah S, Kyere-Gyeabour E. Face mask use among commercial drivers during the COVID-19 pandemic in Accra, Ghana. J Community Health 2021 Dec;46(6):1226-1235 [FREE Full text] [doi: 10.1007/s10900-021-01004-0] [Medline: 34156593]
71. Natnael T, Adane M, Alemnew Y, Andualem A, Hailu F. COVID-19 knowledge, attitude and frequent hand hygiene practices among taxi drivers and associated factors in urban areas of Ethiopia. PLoS One 2021;16(8):e0253452 [FREE Full text] [doi: 10.1371/journal.pone.0253452] [Medline: 34359068]
72. Mboowa G, Musoke D, Bulafu D, Aruhomukama D. Face-masking, an acceptable protective measure against COVID-19 in Ugandan high-risk groups. Am J Trop Med Hyg 2020 Dec 14;104(2):502-513 [FREE Full text] [doi: 10.4269/ajtmh.20-1174] [Medline: 33319741]
73. Fielmua N, Guba BY, Mwingyine DT. Hand hygiene and safety behaviours at shopping centres in COVID-19: an observation in Wa township in Ghana. J Water Sanit Hyg Dev 2021;11(3):442-452 [FREE Full text] [doi: 10.2166/washdev.2021.240]

74. Ameme DK, Odikro MA, Baidoo A, Dsane-Aidoo P, Nuvey FS, Jackson DG, et al. Hand hygiene and face mask wearing practices for COVID-19 prevention: a non-intrusive observation of patrons of community convenience shops in Accra, Ghana. Pan Afr Med J 2021 Dec 01;40:195 [FREE Full text] [doi: 10.11604/pamj.2021.40.195.26195] [Medline: 35096222]
75. Yigzaw N, Ayalew G, Alemu Y, Tesfaye B, Demilew D. Observational study on hand washing practice during COVID-19 pandemic among bank visitors in Gondar Town, Northwest Ethiopia. J Hum Behav Soc Environ 2021 Jul 23;32(6):697-706 [FREE Full text] [doi: 10.1080/10911359.2021.1943592]
76. Puspitasari IM, Yusuf L, Sinuraya RK, Abdulah R, Koyama H. Knowledge, attitude, and practice during the COVID-19 pandemic: a review. J Multidiscip Healthc 2020 Jul 30;13:727-733 [FREE Full text] [doi: 10.2147/JMDH.S265527] [Medline: 32801735]
77. MacIntyre CR, Nguyen PY, Chughtai AA, Trent M, Gerber B, Steinhofel K, et al. Mask use, risk-mitigation behaviours and pandemic fatigue during the COVID-19 pandemic in five cities in Australia, the UK and USA: a cross-sectional survey. Int J Infect Dis 2021 May;106:199-207 [FREE Full text] [doi: 10.1016/j.ijid.2021.03.056] [Medline: 33771668]
78. Martín-Sánchez M, Lim WW, Yeung A, Adam DC, Ali ST, Lau EH, et al. COVID-19 transmission in Hong Kong despite universal masking. J Infect 2021 Jul;83(1):92-95 [FREE Full text] [doi: 10.1016/j.jinf.2021.04.019] [Medline: 33895227]
79. Aloui-Zarrouk Z, El Youssfi L, Badu K, Francis Fagbamigbe A, Matoke-Muhia D, Ngugi C, et al. The wearing of face masks in African countries under the COVID-19 crisis: luxury or necessity? AAS Open Res 2020 Aug 5;3:36 [FREE Full text] [doi: 10.12688/aasopenres.13079.1]
80. Badillo-Goicoechea E, Chang TH, Kim E, LaRocca S, Morris K, Deng X, et al. Global trends and predictors of face mask usage during the COVID-19 pandemic. BMC Public Health 2021 Nov 15;21(1):2099 [FREE Full text] [doi: 10.1186/s12889-021-12175-9] [Medline: 34781917]
81. Madziva R, Murewanhema G, Musuka G, Mapingure MP, Chingombe I, Herrera H, et al. Fighting COVID-19 pandemic fatigue and complacency in Zimbabwe. Public Health Pract (Oxf) 2022 Jun;3:100236 [FREE Full text] [doi: 10.1016/j.puhip.2022.100236] [Medline: 35169760]
82. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res 2020 Jun 29;22(6):e21820 [FREE Full text] [doi: 10.2196/21820] [Medline: 32589589]
83. Burki T. Global shortage of personal protective equipment. Lancet Infect Dis 2020 Jul;20(7):785-786 [FREE Full text] [doi: 10.1016/S1473-3099(20)30501-6] [Medline: 32592673]
84. Islam JY, Vidot DC, Camacho-Rivera M. Determinants of COVID-19 preventive behaviours among adults with chronic diseases in the USA: an analysis of the nationally representative COVID-19 impact survey. BMJ Open 2021 Feb 09;11(2):e044600 [FREE Full text] [doi: 10.1136/bmjopen-2020-044600] [Medline: 33563624]
85. Huynh G, Nguyen MQ, Tran TT, Nguyen VT, Nguyen TV, Do TH, et al. Knowledge, attitude, and practices regarding COVID-19 among chronic illness patients at outpatient departments in Ho Chi Minh city, Vietnam. Risk Manag Healthc Policy 2020 Sep 14;13:1571-1578 [FREE Full text] [doi: 10.2147/RMHP.S268876] [Medline: 32982515]
86. Anjorin AA. The coronavirus disease 2019 (COVID-19) pandemic: a review and an update on cases in Africa. Asian Pac J Trop Med 2019;13(5):199-203 [FREE Full text]

Abbreviations

HCW: health care worker
PPE: personal protective equipment
PPM: personal protective measure
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Mavragani, T Sanchez; submitted 04.11.22; peer-reviewed by JF Fuertes-Bucheli, CR Telles; comments to author 16.03.23; revised version received 29.03.23; accepted 10.04.23; published 16.05.23.

Please cite as:
Kawuki J, Chan PSF, Fang Y, Chen S, Mo PKH, Wang Z
Knowledge and Practice of Personal Protective Measures Against COVID-19 in Africa: Systematic Review
JMIR Public Health Surveill 2023;9:e44051
URL: https://publichealth.jmir.org/2023/1/e44051
doi:10.2196/44051
PMID:37058578

Original Paper

The COVID-19 Pandemic and Daily Steps in the General Population: Meta-analysis of Observational Studies

Ziying Wu[1*], MD; Yilun Wang[1*], MD, PhD; Yuqing Zhang[2,3], DSC; Kim L Bennell[4], PhD; Daniel K White[5], PT, SCD, MSc; Liusong Shen[1], BS; Wei Ren[1], BS; Jie Wei[6,7], PhD; Chao Zeng[1,7,8], MD, PhD; Guanghua Lei[1,7,8], MD, PhD

[1]Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, China

[2]Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[3]The Mongan Institute, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[4]Centre for Health, Exercise and Sports Medicine, The University of Melbourne, Melbourne, Australia

[5]Department of Physical Therapy, University of Delaware, Newark, DE, United States

[6]Health Management Center, Xiangya Hospital, Central South University, Changsha, China

[7]Hunan Key Laboratory of Joint Degeneration and Injury, Xiangya Hospital, Central South University, Changsha, China

[8]National Clinical Research Center of Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Guanghua Lei, MD, PhD
Department of Orthopaedics
Xiangya Hospital
Central South University
87 Xiangya Road
Changsha, 410008
China
Phone: 86 073184327326
Fax: 86 073184327326
Email: lei_guanghua@csu.edu.cn

Abstract

Background: The COVID-19 pandemic has the potential to accelerate another pandemic: physical inactivity. Daily steps, a proxy of physical activity, are closely related to health. Recent studies indicate that over 7000 steps per day is the critical physical activity standard for minimizing the risk of all-cause mortality. Moreover, the risk of cardiovascular events has been found to increase by 8% for every 2000 steps per day decrement.

Objective: To quantify the impact of the COVID-19 pandemic on daily steps in the general adult population.

Methods: This study follows the guidelines of the MOOSE (Meta-analysis Of Observational Studies in Epidemiology) checklist. PubMed, EMBASE, and Web of Science were searched from inception to February 11, 2023. Eligible studies were observational studies reporting monitor-assessed daily steps before and during the confinement period of the COVID-19 pandemic in the general adult population. Two reviewers performed study selection and data extraction independently. The modified Newcastle-Ottawa Scale was used to assess the study quality. A random effects meta-analysis was conducted. The primary outcome of interest was the number of daily steps before (ie, January 2019 to February 2020) and during (ie, after January 2020) the confinement period of COVID-19. Publication bias was assessed with a funnel plot and further evaluated with the Egger test. Sensitivity analyses were performed by excluding studies with low methodological quality or small sample sizes to test the robustness of the findings. Other outcomes included subgroup analyses by geographic location and gender.

Results: A total of 20 studies (19,253 participants) were included. The proportion of studies with subjects with optimal daily steps (ie, ≥7000 steps/day) declined from 70% before the pandemic to 25% during the confinement period. The change in daily steps between the 2 periods ranged from –5771 to –683 across studies, and the pooled mean difference was –2012 (95% CI –2805 to –1218). The asymmetry in the funnel plot and Egger test results did not indicate any significant publication bias. Results remained stable in sensitivity analyses, suggesting that the observed differences were robust. Subgroup analyses revealed that the decline in daily steps clearly varied by region worldwide but that there was no apparent difference between men and women.

Conclusions: Our findings indicate that daily steps declined substantially during the confinement period of the COVID-19 pandemic. The pandemic further exacerbated the ever-increasing prevalence of low levels of physical activity, emphasizing the necessity of adopting appropriate measures to reverse this trend. Further research is required to monitor the consequence of long-term physical inactivity.

Trial Registration: PROSPERO CRD42021291684; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=291684

(JMIR Public Health Surveill 2023;9:e40650) doi:10.2196/40650

KEYWORDS

COVID-19; daily steps; physical activity; meta-analysis

Introduction

Physical activity has been crucial for preventing and treating multiple chronic health conditions [1,2] and premature mortality [3,4], while the number of daily steps is a primary proxy measure for physical activity [5]. Daily step goals can be tailored to meet individual needs and when used with a monitor are effective in promoting physical activity [6]. Previous studies have found that higher daily steps are associated with multiple health benefits, such as lower risk of cardiovascular disease [7], diabetes [8,9], and all-cause mortality [10-13]. For example, the risk of all-cause mortality in adults shows a declining trend with an increase in daily steps and plateaus at approximately 7000 steps per day [12,13]. In contrast, a decrease in daily steps is responsible for a range of unhealthy outcomes, such as increased risk of cardiovascular disease [14], compromised muscle metabolism [15], and elevated systemic inflammation [15,16]. Specifically, the risk of cardiovascular events was reported to increase by 8% for every 2000 steps per day decrement [14]. A sharp reduction in daily steps might also reduce leg lean mass and induce impairments of myofibrillar protein synthesis in healthy elderly individuals, accompanied by increased circulating inflammatory markers [15,16].

Although global responses were far from homogeneous, most countries adopted restriction measures in varying forms to limit the transmission of COVID-19, such as physical distancing, working from home, and closing schools [17]. These measures have resulted in unprecedented changes in all aspects of daily life [18], and physical activity is certainly one of the aspects being most affected. Recently, several studies reported a significant decline in daily steps during the confinement period of the COVID-19 pandemic [19-23]. One study found that the daily steps of young adults in Singapore decreased by 42% on average during the lockdown period compared with the time before the pandemic [20]. Another study reported a sharp, 30% reduction in daily steps in Chinese citizens and a significant increase in the proportion of adults with frequent low daily steps (≤1500 steps/day for ≥14 days) during the confinement period [19]. By contrast, one study reported only a 5% reduction in daily steps 30 days after the start of lockdown in Australia [22]. These findings indicate the possibility of geographically specific variations in the relationship between COVID-19–related restrictions and daily steps, which limits the generalization of conclusions drawn from single country–based analyses. Therefore, a systematic review is needed to consolidate the current understanding of the impact of COVID-19 on daily steps in order to enlighten future interventions and research directions. In addition, a few studies have also evaluated whether the impact of COVID-19 on physical activity varied by gender. One study suggested that men had more reduction in time spent in moderate-vigorous physical activities and a larger increase in sedentary behaviors than women during the pandemic [24]. However, another study failed to confirm this finding [25].

To fill this knowledge gap, we conducted a meta-analysis of observational studies to quantify both global and regional impacts of confinement during the COVID-19 pandemic on daily steps in the general adult population, and we also examined whether the decline in daily steps differed between men and women.

Methods

Protocol and Registration

The MOOSE (Meta-analysis Of Observational Studies in Epidemiology) [26] and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 [27] guidelines were followed for reporting this study (Multimedia Appendix 1). The protocol for this study was prospectively registered in the PROSPERO database (CRD42021291684).

Search Strategy

A preliminary search was conducted on PubMed, EMBASE, and Web of Science to retrieve studies that reported daily step data both before and during the confinement period of the COVID-19 pandemic in the general adult population published before November 22, 2021. Further, an updated search was performed on February 11, 2023, to add recently published articles. The detailed search strategy is listed in Multimedia Appendix 2 [28-31]. The International Standard Randomized Controlled Trial Number Register (ISRCTN) and the World Health Organization COVID-19 database were also searched for potential gray literature. The reference lists and citing articles of relevant literature on this topic were also checked to find possible studies that could be included in this research. The language of publication was not restricted.

Eligibility Criteria and Study Selection

We included observational studies that reported daily steps before and during the confinement period of the COVID-19 pandemic in the general adult population. The reference period was restricted to ≤2 years before the time points when the government implemented the confinement measures. Further eligibility criteria are listed in Textbox 1.

Textbox 1. Eligibility criteria for study screening.

Inclusion criteria

Study type: Observational studies

Participants: General adult population (average age over 20 years)

Outcomes: Daily steps, measured using monitoring devices including wearable physical activity trackers (eg, pedometers, accelerometers, smartwatches, and smart bands) and smartphones mobile apps (eg, iPhone Health, Google Fit, WeChat, and Exercise Health)

Other criteria: Daily step data collected both before and during the confinement period of COVID-19. The reference period was restricted to ≤2 years before the time points when the government implemented the confinement measures

Exclusion criteria

Study type: Any non–observational study (eg, reviews, randomized controlled trials, and case reports), conference abstracts

Participants: Children, adolescents, athletes, specific patient groups, participants with particular conditions (eg, pregnant women or elderly individuals at high risk of falls)

Outcomes: Data other than daily steps (eg, physical activity measured using questionnaires alone or data measured by wearable devices that are classified as light, moderate, or vigorous physical activity)

Other criteria: Studies offering financial incentives that may promote physical activity (eg, rewards for individuals who achieve predetermined goals through monitors), studies lacking available data for meta-analysis

Two independent reviewers (LS and WR) screened all titles and abstracts. The full text of potentially relevant literature was further evaluated. A third reviewer (ZW) arbitrated the decision if a consensus was not reached.

Data Extraction

Two authors (LS and WR) extracted the following data independently: study design, country and continent (Africa, Asia, Europe, North America, South America, and Oceania) where the study was conducted, the sample size used for daily step analysis, participant characteristics (ie, mean age and female ratio), monitoring devices used to measure daily steps, the algorithm used to analyze daily steps, data collection duration, and COVID-19 restriction measures by the government at the time of data collection. The outcomes of interest were the total number of daily steps before and during the confinement period of COVID-19 and the percentage change between the 2 periods. If not reported, the mean (SD) of the outcome was estimated based on the methods provided by the Cochrane Handbook [32]. If data needed to be extracted from graphs, the GetData digitizer was used [33]. This tool has been validated and has good accuracy and precision for data extraction; it has been used in meta-analysis by previous studies [34]. Disagreements, if any, were resolved by consulting with a third reviewer (ZW). When any desired information related to daily steps was missing or unclear, the corresponding author was contacted via email to request or confirm the relevant data, as appropriate. If the data could not be obtained, the study was excluded from the meta-analysis.

Risk of Bias and Quality of Individual Studies

The modified Newcastle-Ottawa Scale (NOS) for applicable observational studies was used to assess the methodological quality and risk of bias in each study [28-30]. Specifically, the modified NOS consists of 3 domains: *selection* (maximum 3 stars), *comparability* (maximum 2 stars), and *outcome* (maximum 2 stars). Therefore, 1 study can be awarded a maximum of 7 stars. Studies with 6 to 7 stars were rated as high quality, those with 5 stars were rated as moderate quality, and those with 4 or fewer stars were rated as low quality [29]. Two reviewers (LS and ZW) rated the methodological quality of each study independently, and discrepancies, if any, were discussed for consensus or resolved by consulting a third reviewer (YW).

Data Synthesis and Analysis

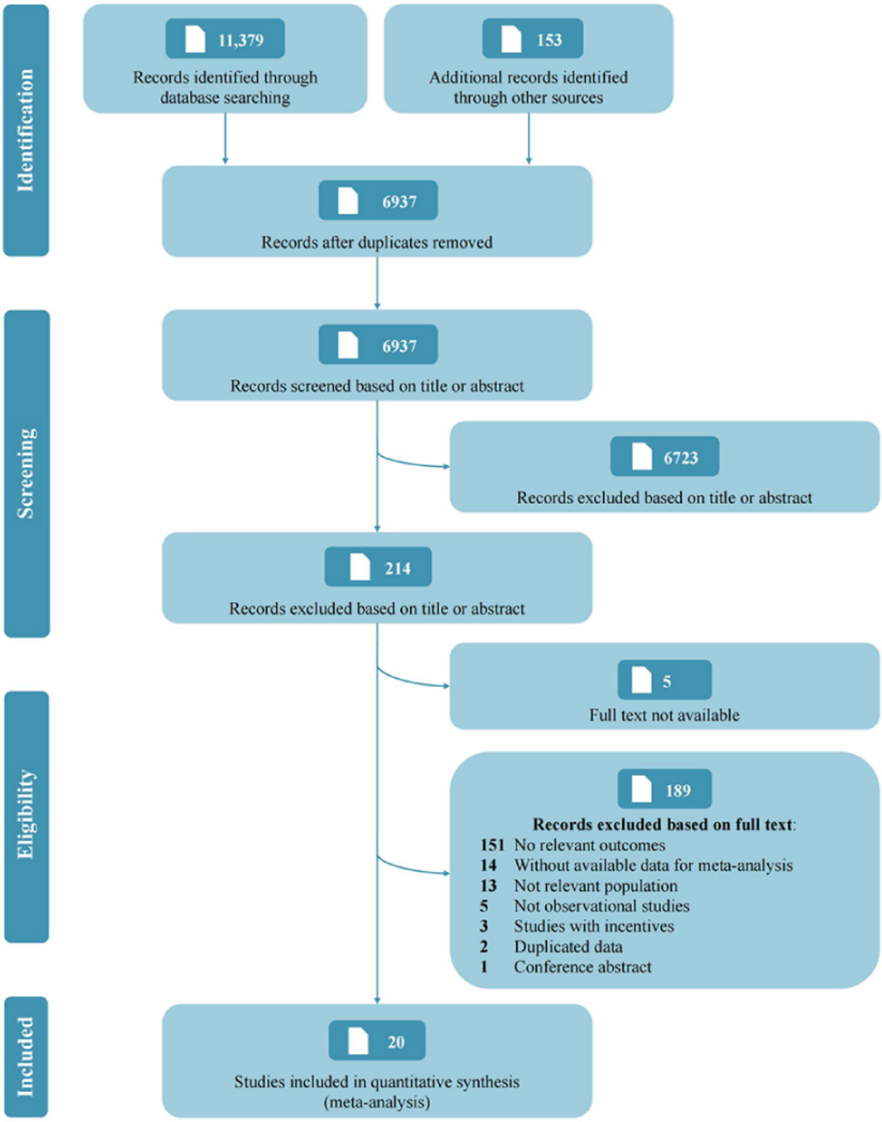
A random effects meta-analysis was conducted using Stata software (version 14.0; StataCorp). The change in daily steps was calculated using the mean difference (MD) with the 95% CI between the mean (SD) of daily steps before and during the confinement period of COVID-19. The Cochran Q statistic of the inconsistency index (I^2) was used to test statistical heterogeneity across studies [32,35]. An I^2 value of more than 75% indicated high heterogeneity [32]. Publication bias was assessed with a funnel plot and further evaluated with the Egger test [36]. To test the robustness of the findings on the primary outcome (ie, the number of daily steps), we performed 3 sensitivity analyses. First, we conducted an analysis that eliminated each study one at a time. Second, we performed an analysis that removed the studies rated as having low methodological quality. Third, we performed an analysis that excluded the studies with a small sample size (less than 140) based on the recommendations of Dunton and colleagues [37]. Finally, we also conducted subgroup analyses according to country, continent, and gender.

Results

Study Selection

A total of 6937 unduplicated records were preliminarily retrieved, and 214 potentially eligible studies were identified after the title and abstract screening. Then, 5 studies were excluded because the full text was unavailable, and full text assessments were made of the remaining 209 articles. Another 189 articles were eliminated. Of these, 151 studies did not report relevant outcomes (ie, daily steps), 14 did not provide necessary data for meta-analysis, 13 focused on irrelevant populations, 5 were not observational studies, 3 were conducted with financial incentives, 2 reported duplicated data, and 1 was a conference abstract. The remaining 20 studies were included in the meta-analysis [19,20,37-54]. Figure 1 shows the PRISMA flow diagram.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Study Characteristics

The included studies were conducted in Australia [38], Canada [39], China [19,40-43], France [38], Italy [44], Japan [45-48], Norway [49], Singapore [20], South Korea [50], Spain [51], the United Kingdom [38,52], and the United States [37,38,53,54], and all of them investigated the change in daily steps in the early stage of COVID-19 (from January to September 2020). The reference pre–COVID-19 period was from January 2019 to March 2020. The detailed duration of observation for each study is shown in Table 1. Nine studies measured daily steps with smartphones [19,37,40,41,43,45,46,50,54], 8 with wearable activity trackers [20,38,39,47,48,51-53], and 2 with both wearable activity trackers and smartphones [44,49]; 1 study did not specify the measurement device used in the survey [42]. One study provided 4 groups of data from 4 countries separately [38]. Therefore, the 20 studies included a total of 19,253 participants, and 23 sets of data were included in the meta-analysis. Additional study characteristics are summarized in Table 1.

Table 1. Characteristics of studies and participants included in the meta-analysis.

Study	Country/Continent	Participants, n	Age (years)	Female, %	Device	Algorithm	Definition of restriction measures	Observation period	Reference period
Azuma et al [45] 2021	Japan/Asia	530[a]	—[b]	75.1	Smartphone (iPhone Health)	Averaged data from 9 weekly averages	Declared emergency (closed schools and "stay at home" recommendation issued)	Feb 24 to Apr 26, 2020	Feb 24 to Apr 26, 2019
Bird et al[c] [52] 2021	UK/Europe	190	Range 18-85	77.9	Mobile device (eg, smartwatch)	—	Lockdown	March to May 2020	Before March 2020
Buoite Stella et al [44] 2021	Italy/Europe	197	—	—	Smart technology devices (eg, smartphone, smart band, and smartwatch)	Weekly average	Lockdown	Mar 23 to Mar 29, 2020	One week in Jan 2020
Ding et al[c] [40] 2021	China/Asia	301	—	76.4	Smartphone (WeChat)	Averaged data from 14-day data	Lockdown	Mar 11 to Mar 24, 2020	Dec 28, 2019, to Jan 10, 2020
Dunton et al [37] 2020	US/North America	149	Range 18-74	—	Smartphone (iPhone Health)	Monthly average	Declared "shelter in place" and "stay at home" orders	Apr 2020	Feb 2020
Gollwitzer et al [54] 2022	US/North America	220	Mean 35.2	41.4	Smartphone (iPhone Health)	Weekly average	Social distancing	Apr 3 to Apr 9, 2020	Apr 3 to Apr 9, 2019
He et al [41] 2020	China/Asia	339	Women: mean 37.6; men: mean 36.4	53.4	Smartphone (eg, Exercise Health, Keep)	Average data during the observation period	Semilockdown (assembly was suspended, and curfew and quarantine were implemented in some districts)	Jan 27 to Mar 1, 2020	Dec 23, 2019, to Jan 26, 2020
Henriksen et al[d] [49] 2021	Norway/Europe	113	Mean 40.6	56.2	Wearable activity trackers (Apple, Fitbit, Garmin, Oura, Polar, Samsung, and Withings), and smartphones (Google Fit and Apple Health)	Monthly average	Lockdown	Mar 13 to Mar 31, 2020	Mar 2019
Hudgins et al [53] 2021	US/North America	80	Mean 32.2	—	Wearable activity trackers	Monthly average	Academic break in 2020, during which the university transitioned to remote learning	30 days after spring break 2020	30 days before spring break 2020
Ji et al [42] 2022	China/Asia	781	—	—	—	—	—	—	—
Karageorghis et al [38] 2021	US/North America	1029	Mean 40.7	76.4	Electronic devices	Self-reported daily steps based on devices' records	Declared "stay at home" order	Mar 21 to Apr 7, 2020	Before Mar 21, 2020
Karageorghis et al [38] 2021	UK/Europe	392	Mean 51.2	80.1			National lockdown	After Mar 23, 2020	Before Mar 23, 2020
Karageorghis et al [38] 2021	France/Europe	734	Mean 37.7	76.0			National lockdown	After Mar 16 2020	Before Mar 16 2020
Karageorghis et al [38] 2021	Australia/Oceania	386	Mean 42.5	72.5			Interstate border closures[e]	After Mar 19, 2020	Before Mar 19, 2020

Study	Country/Continent	Participants, n	Age (years)	Female, %	Device	Algorithm	Definition of restriction measures	Observation period	Reference period
Obuchi et al [46] 2021	Japan/Asia	2587	—	76.7	Smartphone	Weekly average	Declared emergency (closed schools and issued “stay at home” recommendation)	Apr 21 to Apr 27, 2020	Apr 30 to May 6, 2019
Ong et al[c] [20] 2021	Singapore/Asia	1824	Mean 30.9	51.6	Wearable activity tracker (Fitbit)	Averaged data from 3-week data	Lockdown	Apr 7 to Apr 27, 2020	Jan 2 to Jan 22, 2020
Park et al [50] 2021	South Korea/Asia	1163[a]	Mean 23.7	45.6	Smartphone (iPhone Health)	Monthly average	Level 2.5 (gatherings of more than 50 people indoors and more than 100 people outdoors were prohibited)	Sep 2020	Jan to Dec 2019
Sañudo et al [51] 2020	Spain/Europe	20	Mean 22.6	47	Xiaomi Mi Band 2 wrist-worn accelerometer	Weekly average	Lockdown	One week, between Mar 24 and Apr 3, 2020	One week in Feb 2020
Sato et al [47] 2022	Japan/Asia	2846	Mean 45.9	59.6	Accelerometer (Panasonic)	Average data during the observation period	Declared emergency (closed schools and issued “stay at home” recommendation)	Apr 7 to May 13, 2020	Jan 1 to Feb 29, 2019
Wang et al[c] [19] 2020	China/Asia	3544	Mean 51.6	34.6	Smartphone (WeChat)	Monthly average	Declared emergency (physical distancing measures were issued)	Jan 22 to Feb 20, 2020	Feb 2 to Mar 3, 2019
Woodruff et al [39] 2021	Canada/North America	121	Mean 36.2	80	Wearable activity trackers (Apple, Fitbit, Samsung, Garmin)	—	Declared emergency (physical distancing measures were issued)	Based on the date each person began physically distancing/self-isolating	Based on the date each person began physically distancing/self-isolating
Yamada et al [48] 2023	Japan/Asia	678	—	—	Wearable activity tracker (EW-NK63)	Monthly average	Declared emergency (closed schools and issued “stay at home” recommendation)	April 2020	April 2019
Zhu et al [43] 2021	China/Asia	1029	Range 18-76	69.1	Smartphone (WeChat)	Self-reported daily steps based on smartphone’s records	Lockdown	Three months after entering isolated life	Jan 21 to Jan 25, 2020

[a] Observations are reported instead of participants for these studies, because of uncertainty about the number of participants who were double counted (before and after the confinement period).

[b] Not available.

[c] Data for daily steps were obtained from the authors for these studies.

[d] Only 106 of 113 participants provided characteristics in this study.

[e] “Interstate border closures” means that Australia closed its borders to all noncitizens and nonresidents. This started on March 20, 2020, with no exceptions for Australian citizens, permanent residents, or their immediate families. Furthermore, a lockdown policy started in Australia on March 23, 2020.

Quality Assessment

Thirteen studies (65%) were rated as having moderate to high methodological quality (Multimedia Appendix 2, Table S1). Considering the widespread weaknesses of small sample size, imbalanced gender ratio, limited age range of the study sample, and poor nationwide representativeness, only 2 studies were rated as having national representativeness of the exposed cohort (ie, samples during the confinement period of COVID-19). Thus, 10% of the included studies received 3 stars in the *selection* dimension, 30% received 2 stars in the *comparability* dimension, and 60% received 2 stars in the *outcome* dimension.

Quantitative Synthesis

As shown in Figure 2, the proportion of studies with subjects with optimal daily steps (ie, ≥7000 steps per day) declined from 70% before the pandemic to 25% during the confinement period. The change in daily steps between the 2 periods ranged from –5771 to –683. The pooled MD from the random effects pooled analysis was –2012 (95% CI –2805 to –1218; Figure 3), with significant heterogeneity among studies ($P<.001$; I^2=97.6%).

The funnel plot shows a symmetrical horizontal distribution for each study relative to the vertical line, suggesting the absence of reporting bias (Multimedia Appendix 2, Figure S1). The Egger publication bias indicated a lack of publication bias (Multimedia Appendix 2, Figure S2). Sensitivity analysis, performed by omitting each study one by one, revealed that the change in daily steps was not significantly altered by any individual study (Multimedia Appendix 2, Figure S3). Moreover, the results remained robust in sensitivity analyses after removing low-quality studies (MD=–1714, 95% CI –2809 to –618) and after removing small-sample studies (MD=–2024, 95% CI –2836 to –1212). The percentage change in daily steps varied from –69.6% to –10.6% across different studies (Figure 4).

Figure 2. Average daily steps around the period before and during the confinement period across studies.

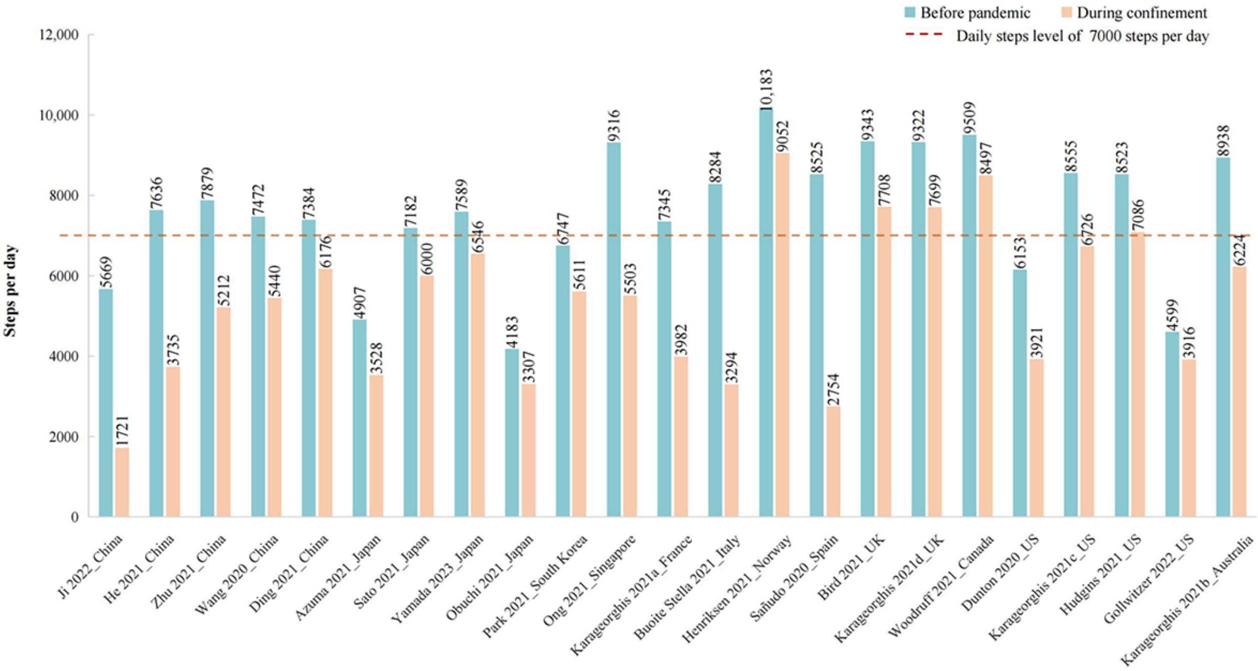


Figure 3. Forest plot of reduction in daily steps during the confinement period. DL: DerSimonian and Laird method.

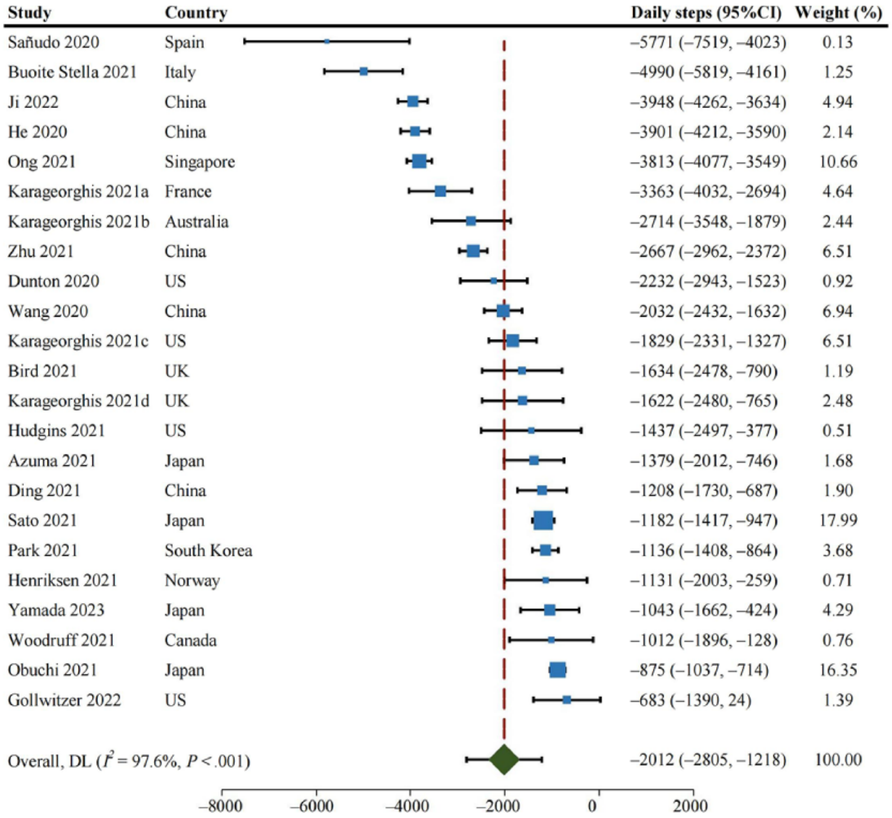
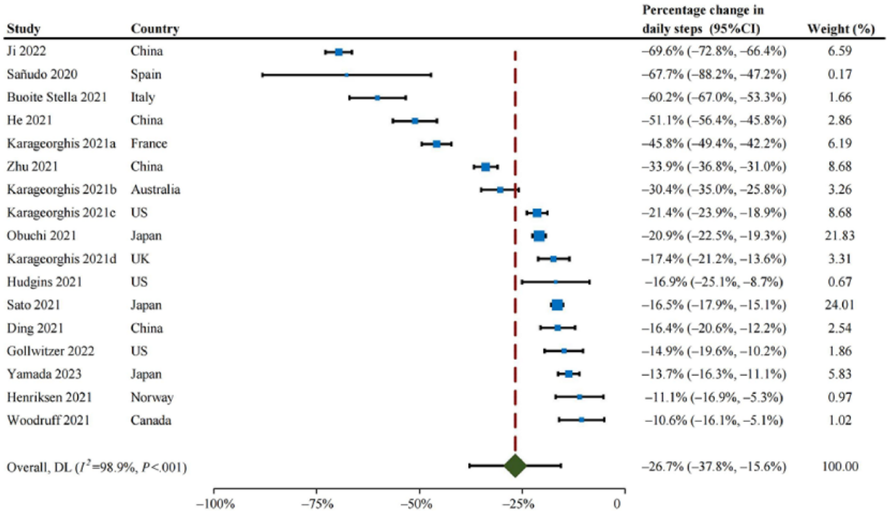
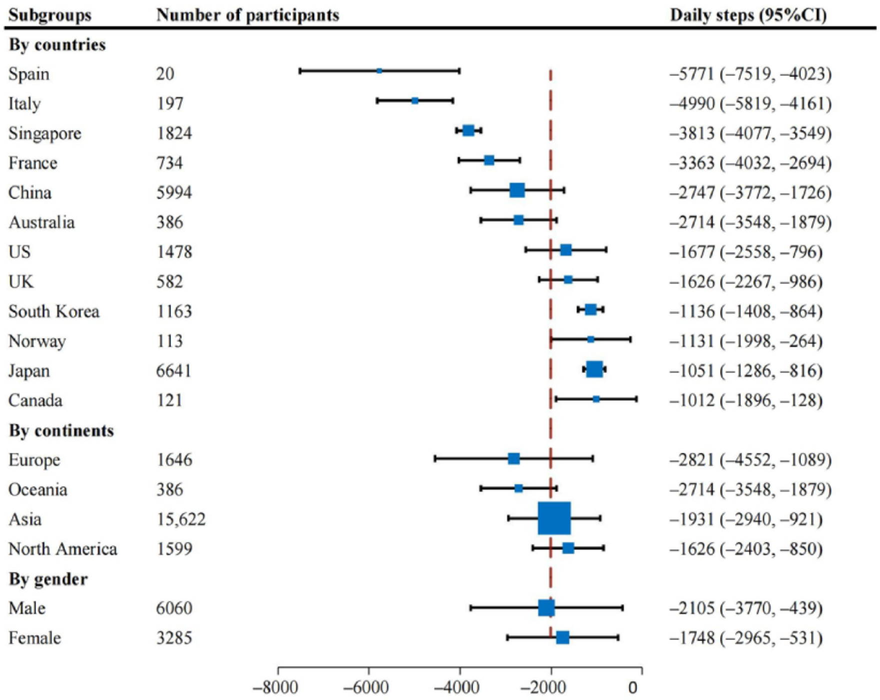


Figure 4. Percentage reduction in daily steps during the confinement period of the COVID-19 pandemic. DL: DerSimonian and Laird method.



Country-specific analyses (Figure 5) showed that the most significant decline in daily steps occurred in subjects from Spain (MD=–5771, 95% CI –7519 to –4023), followed by Italy (MD=–4990, 95% CI –5819 to –4161) and Singapore (MD=–3813, 95% CI –4077 to –3549). The decline in daily steps was less in subjects from Canada (MD=–1012, 95% CI –1896 to –128), Japan (MD=–1051, 95% CI –1286 to –816), and Norway (MD=–1131, 95% CI –1998 to –264). In terms of continent, European subjects presented the most significant reduction in daily steps but with a wide CI (MD=–2821, 95% CI –4552 to –1089), followed by subjects from Oceania (MD=–2714, 95% CI –3548 to –1879) and Asia (MD=–1931, 95% CI –2940 to –921), while subjects from North America presented the smallest decline in daily steps (MD=–1626, 95% CI –2403 to –850). Of the studies that provided gender-specific data [20,41,43,45-47,52], no apparent difference was observed in the change in daily steps between men (MD=–2105, 95% CI –3770 to –439) and women (MD=–1748, 95% CI –2965 to –531).

Figure 5. Pooled subgroup results of reduction in daily steps during the confinement period of the COVID-19 pandemic.



Discussion

Principal Results

Our study quantified the change in daily steps during the confinement period of COVID-19. The proportion of studies with subjects with optimal daily steps (ie, ≥7000 steps/day) declined from 70% before the pandemic to 25% during the confinement period, and the average number of daily steps was reduced by 2000. Our results also suggest a possible inequality in physical activity across countries, whether before or during the pandemic. However, given that most included studies lacked sufficient nationwide representativeness, this difference should be interpreted cautiously. Furthermore, there was no apparent gender difference in the decline in daily steps before and during the confinement period of COVID-19.

Possible Explanations and Clinical Implications

The population-level trends in daily steps during the confinement period of COVID-19 may reflect public panic about the virus and adherence to confinement measures (ie, physical distancing and sheltering in place) [55,56]. Our findings are not surprising given that COVID-19 restriction measures deprived the general population of many opportunities for physical activity, such as commuting, shopping, sports, and vacations or trips. In addition, the observed differences in the change in daily steps across studies may reflect regional variations in enforcement and behavioral changes [57,58]. Socioeconomic inequalities might also influence such variations among regions and disparities in the ability to engage in or access recreational physical activity within areas [59].

Daily steps are strongly associated with health outcomes [7-16]. Recent studies indicate that over 7000 steps per day is the critical physical activity standard for minimizing the risk of all-cause mortality [12,13]. Moreover, the risk of cardiovascular events increases by 8% for every 2000 steps per day decrement [14]. We found that daily steps declined by 2000, and half of the included studies reported that average daily steps declined to below 7000 during the confinement period of the COVID-19 pandemic. Such findings imply that the pandemic further exacerbated the ever-increasing prevalence of low-level physical activity [60,61], emphasizing the necessity of adopting appropriate measures to reverse this trend. Furthermore, this meta-analysis focused on observational studies based on monitoring devices. Since monitoring devices may encourage users to engage in physical activities [6], the number of daily steps during the confinement period might have been even lower in the population than was currently observed in the samples. In addition, the impact of COVID-19 on daily steps may persist in the long term. Although some researchers report that daily steps gradually recovered to the prepandemic level after the lifting of restriction measures [49,62], more studies argue that worldwide daily steps remain depressed compared with the prepandemic level (ie, the immediate 2 years before the outbreak of COVID-19) [59,63-66]. Thus, the consequences of long-term physical inactivity, which may be further exacerbated, should be addressed.

Comparison With Previous Work

Our finding of a reduction in daily steps during the confinement period of the pandemic is consistent with findings from previous studies using different physical activity metrics [67,68]. An earlier meta-analysis found that sedentary time significantly increased, by 126.9 minutes per day in adults and by 46.9 minutes per day in the elderly, during the confinement period of the pandemic [67]. Another systematic review reported that most of the included studies revealed a decrease in physical activity in most participants [68]. All studies showed an increased sedentary time in healthy adults and those with medical conditions during the lockdown period [68]. According to this evidence, public health messages about staying active during the pandemic appear to have had little effect on the population's engagement in physical activity [69]. In addition, 2 studies explored gender differences in the change in physical activity during the COVID-19 pandemic but had inconsistent conclusions [24,25]. One study found that men had more clearly decreased moderate-vigorous physical activity and increased sedentary behaviors than women during the pandemic [24]. However, another study failed to show significant differences between men and women in self-reported decline in physical activity [25]. Our meta-analysis adopted a primary indicator of physical activity, ie, daily steps, and found that daily steps declined substantially during the COVID-19 confinement period. Such a decline varied by region but not by gender. These results are consistent with previous studies [24,25,67,68].

Based on data collected from 187 countries and districts by a smartphone app, Tison and colleagues [59] found that mean daily steps declined by 5.5% (287 steps) within 10 days of the pandemic declaration by the World Health Organization and by 27.3% (1432 steps) within 30 days. The authors also reported noticeable differences across regions; for example, Italy, France, Brazil, and Iran exhibited a more than 40% maximum decrease, while Sweden exhibited a maximum reduction of daily steps of less than 10% [59]. Furthermore, regional differences were significant. For example, several studies showed that the decline in daily steps during confinement varied considerably among Chinese adults from different provinces [40,41,43]. One nationwide study in Japan found that daily steps declined significantly after the declaration of a state of emergency in urban areas but not rural areas [62].

Limitations

Several limitations should be acknowledged. First, daily step data are lacking from Africa and South America, limiting the generalizability of our findings. Second, when comparing regional differences across countries, we could not rule out potential sampling bias, because most of the included studies lacked sufficient nationwide representativeness. Therefore, future studies are needed to verify our findings. Third, since long-term data from the period that confinement was lifted are not suitable for inclusion in a meta-analysis, we could not directly assess the long-term impact of COVID-19 on physical activity levels using meta-analysis techniques. As more published data become available [59,63-66], this will be an important direction for future studies. Fourth, we did not examine the impact of reduced daily steps on health outcomes, owing to a lack of relevant data. Future studies are expected to address this crucial issue. Finally, there was considerable study heterogeneity. Potential explanations include discrepancies in confinement measures across countries or districts and in the adherence to policy of residents, as well as unclear or different time frames over which the average daily steps were calculated, the possibility that monitoring devices were changed in the different reporting periods, and different durations of observation.

Conclusion

In summary, we found that daily steps declined substantially during the COVID-19 confinement period. Our findings indicate that COVID-19 further exacerbated the ever-increasing prevalence of low levels of physical activity, emphasizing the necessity of adopting appropriate measures to reverse this trend. Further research is required to monitor the consequence of long-term physical inactivity.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (81930071, 81902265, and 82072502), the National Natural Science Foundation Regional Innovation and Development Joint Fund (U21A20352), the National Key Research and Development Program (2022YFC3601900 and 2022YFC3601901), the Natural Science Foundation of Hunan Province (2022JJ40821), the Project Program of the National Clinical Research Center for Geriatric Disorders (Xiangya Hospital; 2020LNJJ03 and 2021LNJJ06), and the Youth Science Foundation of Xiangya Hospital (2021Q14). KLB is supported by a National Health and Medical Research Council Investigator Fellowship (1058440). No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

ZW, YW, CZ, and GL were responsible for the conception and design of the study. ZW, YW, LS, and WR contributed to study retrieval. ZW and LS contributed to quality assessment. ZW, LS, and WR contributed to data collection. ZW, YW, and JW contributed to statistical analysis. All authors assisted with interpretation. ZW and YW drafted the manuscript. All authors contributed to the revision of the manuscript. All authors read and approved the final manuscript. CZ and GL supervised the study. Authors CZ (zengchao@csu.edu.cn) and GL (lei_guanghua@csu.edu.cn) are joint corresponding authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MOOSE checklist.
[DOCX File , 17 KB - publichealth_v9i1e40650_app1.docx]

Multimedia Appendix 2
Supplementary methods and results.
[DOCX File , 426 KB - publichealth_v9i1e40650_app2.docx]

References

1. Martinez-Gomez D, Esteban-Cornejo I, Lopez-Garcia E, García-Esquinas E, Sadarangani KP, Veiga OL, et al. Physical activity less than the recommended amount may prevent the onset of major biological risk factors for cardiovascular disease: a cohort study of 198 919 adults. Br J Sports Med 2020 Feb 15;54(4):238-244. [doi: 10.1136/bjsports-2018-099740] [Medline: 30554146]
2. Iso-Markku P, Kujala UM, Knittle K, Polet J, Vuoksimaa E, Waller K. Physical activity as a protective factor for dementia and Alzheimer's disease: systematic review, meta-analysis and quality assessment of cohort and case-control studies. Br J Sports Med 2022 Jun 17;56(12):701-709 [FREE Full text] [doi: 10.1136/bjsports-2021-104981] [Medline: 35301183]
3. Tarp J, Fagerland MW, Dalene KE, Johannessen JS, Hansen BH, Jefferis BJ, et al. Device-measured physical activity, adiposity and mortality: a harmonised meta-analysis of eight prospective cohort studies. Br J Sports Med 2022 Jul;56(13):725-732 [FREE Full text] [doi: 10.1136/bjsports-2021-104827] [Medline: 34876405]
4. Zhao M, Veeranki SP, Li S, Steffen LM, Xi B. Beneficial associations of low and large doses of leisure time physical activity with all-cause, cardiovascular disease and cancer mortality: a national cohort study of 88,140 US adults. Br J Sports Med 2019 Nov;53(22):1405-1411. [doi: 10.1136/bjsports-2018-099254] [Medline: 30890520]
5. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Department of Health and Human Services. URL: https://health.gov/sites/default/files/2019-09/PAG_Advisory_Committee_Report.pdf [accessed 2020-11-03]
6. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
7. Lin H, Sardana M, Zhang Y, Liu C, Trinquart L, Benjamin EJ, et al. Association of habitual physical activity with cardiovascular disease risk. Circ Res 2020 Oct 23;127(10):1253-1260 [FREE Full text] [doi: 10.1161/CIRCRESAHA.120.317578] [Medline: 32842915]
8. Fretts AM, Howard BV, McKnight B, Duncan GE, Beresford SAA, Calhoun D, et al. Modest levels of physical activity are associated with a lower incidence of diabetes in a population with a high rate of obesity: the strong heart family study. Diabetes Care 2012 Aug;35(8):1743-1745 [FREE Full text] [doi: 10.2337/dc11-2321] [Medline: 22723343]

9. Dwyer T, Ponsonby A, Ukoumunne OC, Pezic A, Venn A, Dunstan D, et al. Association of change in daily step count over five years with insulin sensitivity and adiposity: population based cohort study. BMJ 2011 Jan 13;342:c7249. [doi: 10.1136/bmj.c7249] [Medline: 21233153]
10. Lee I, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of step volume and intensity with all-cause mortality in older women. JAMA Intern Med 2019 Aug 01;179(8):1105-1112 [FREE Full text] [doi: 10.1001/jamainternmed.2019.0899] [Medline: 31141585]
11. Saint-Maurice PF, Troiano RP, Bassett DR, Graubard BI, Carlson SA, Shiroma EJ, et al. Association of daily step count and step intensity with mortality among US adults. JAMA 2020 Mar 24;323(12):1151-1160 [FREE Full text] [doi: 10.1001/jama.2020.1382] [Medline: 32207799]
12. Paluch AE, Bajpai S, Bassett DR, Carnethon MR, Ekelund U, Evenson KR, Steps for Health Collaborative. Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. Lancet Public Health 2022 Mar;7(3):e219-e228 [FREE Full text] [doi: 10.1016/S2468-2667(21)00302-9] [Medline: 35247352]
13. Paluch AE, Gabriel KP, Fulton JE, Lewis CE, Schreiner PJ, Sternfeld B, et al. Steps per day and all-cause mortality in middle-aged adults in the coronary artery risk development in young adults study. JAMA Netw Open 2021 Sep 01;4(9):e2124516 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.24516] [Medline: 34477847]
14. Yates T, Haffner SM, Schulte PJ, Thomas L, Huffman KM, Bales CW, et al. Association between change in daily ambulatory activity and cardiovascular events in people with impaired glucose tolerance (NAVIGATOR trial): a cohort analysis. Lancet 2014 Mar 22;383(9922):1059-1066. [doi: 10.1016/S0140-6736(13)62061-9] [Medline: 24361242]
15. Breen L, Stokes KA, Churchward-Venne TA, Moore DR, Baker SK, Smith K, et al. Two weeks of reduced activity decreases leg lean mass and induces "anabolic resistance" of myofibrillar protein synthesis in healthy elderly. J Clin Endocrinol Metab 2013 Jun;98(6):2604-2612. [doi: 10.1210/jc.2013-1502] [Medline: 23589526]
16. Bowden Davies KA, Norman JA, Thompson A, Mitchell KL, Harrold JA, Halford JCG, et al. Short-term physical inactivity induces endothelial dysfunction. Front Physiol 2021 Apr 9;12:659834 [FREE Full text] [doi: 10.3389/fphys.2021.659834] [Medline: 33897466]
17. Coronavirus: the world in lockdown in maps and charts. British Broadcasting Corporation. URL: https://www.bbc.com/news/world-52103747 [accessed 2023-04-13]
18. Coronavirus disease (COVID-19) pandemic. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed 2021-12-23]
19. Wang Y, Zhang Y, Bennell K, White DK, Wei J, Wu Z, et al. Physical distancing measures and walking activity in middle-aged and older residents in Changsha, China, during the COVID-19 epidemic period: longitudinal observational study. J Med Internet Res 2020 Oct 26;22(10):e21632 [FREE Full text] [doi: 10.2196/21632] [Medline: 33027035]
20. Ong JL, Lau T, Massar SAA, Chong ZT, Ng BKL, Koek D, et al. COVID-19-related mobility reduction: heterogenous effects on sleep and physical activity rhythms. Sleep 2021 Feb 12;44(2):a179 [FREE Full text] [doi: 10.1093/sleep/zsaa179] [Medline: 32918076]
21. Giuntella O, Hyde K, Saccardo S, Sadoff S. Lifestyle and mental health disruptions during COVID-19. Proc Natl Acad Sci U S A 2021 Mar 02;118(9):e2016632118 [FREE Full text] [doi: 10.1073/pnas.2016632118] [Medline: 33571107]
22. To QG, Duncan MJ, Van Itallie A, Vandelanotte C. Impact of COVID-19 on physical activity among 10,000 steps members and engagement with the program in Australia: prospective study. J Med Internet Res 2021 Jan 25;23(1):e23946 [FREE Full text] [doi: 10.2196/23946] [Medline: 33449907]
23. Pépin JL, Bruno RM, Yang R, Vercamer V, Jouhaud P, Escourrou P, et al. Wearable activity trackers for monitoring adherence to home confinement during the COVID-19 pandemic worldwide: data aggregation and analysis. J Med Internet Res 2020 Jun 19;22(6):e19787 [FREE Full text] [doi: 10.2196/19787] [Medline: 32501803]
24. Castañeda-Babarro A, Arbillaga-Etxarri A, Gutiérrez-Santamaría B, Coca A. Physical activity change during COVID-19 confinement. Int J Environ Res Public Health 2020 Sep 21;17(18):6878 [FREE Full text] [doi: 10.3390/ijerph17186878] [Medline: 32967091]
25. Wilke J, Mohr L, Tenforde AS, Edouard P, Fossati C, González-Gross M, et al. A Pandemic within the pandemic? physical activity levels substantially decreased in countries affected by COVID-19. Int J Environ Res Public Health 2021 Feb 24;18(5):2235 [FREE Full text] [doi: 10.3390/ijerph18052235] [Medline: 33668262]
26. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis Of Observational Studies in Epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000 Apr 19;283(15):2008-2012. [doi: 10.1001/jama.283.15.2008] [Medline: 10789670]
27. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160 [FREE Full text] [doi: 10.1136/bmj.n160] [Medline: 33781993]
28. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. 2009. URL: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2023-05-12]

29. Ceban F, Nogo D, Carvalho IP, Lee Y, Nasri F, Xiong J, et al. Association between mood disorders and risk of COVID-19 infection, hospitalization, and death: a systematic review and meta-analysis. JAMA Psychiatry 2021 Oct 01;78(10):1079-1091 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.1818] [Medline: 34319365]
30. Smith KN, Baynard T, Fischbach PS, Hankins JS, Hsu LL, Murphy PM, et al. Safety of maximal cardiopulmonary exercise testing in individuals with sickle cell disease: a systematic review. Br J Sports Med 2022 Jul;56(13):764-769 [FREE Full text] [doi: 10.1136/bjsports-2021-104450] [Medline: 34285054]
31. Li L, Shen J, Bala MM, Busse JW, Ebrahim S, Vandvik PO, et al. Incretin treatment and risk of pancreatitis in patients with type 2 diabetes mellitus: systematic review and meta-analysis of randomised and non-randomised studies. BMJ 2014 Apr 15;348:g2366 [FREE Full text] [doi: 10.1136/bmj.g2366] [Medline: 24736555]
32. Deeks J, Higgins J, Altman D. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5. London, UK: Cochrane Collaboration; 2011:1-44.
33. GetData Graph Digitizer. URL: https://www.getdata-graph-digitizer.com/ [accessed 2022-03-02]
34. Wojtyniak J, Britz H, Selzer D, Schwab M, Lehr T. Data digitizing: accurate and precise data extraction for quantitative systems pharmacology and physiologically-based pharmacokinetic modeling. CPT Pharmacometrics Syst Pharmacol 2020 Jun 16;9(6):322-331 [FREE Full text] [doi: 10.1002/psp4.12511] [Medline: 32543786]
35. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]
36. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
37. Dunton GF, Wang SD, Do B, Courtney J. Early effects of the COVID-19 pandemic on physical activity locations and behaviors in adults living in the United States. Prev Med Rep 2020 Dec;20:101241 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101241] [Medline: 33173751]
38. Karageorghis CI, Bird JM, Hutchinson JC, Hamer M, Delevoye-Turrell YN, Guérin SMR, et al. Physical activity and mental well-being under COVID-19 lockdown: a cross-sectional multination study. BMC Public Health 2021 May 27;21(1):988 [FREE Full text] [doi: 10.1186/s12889-021-10931-5] [Medline: 34039306]
39. Woodruff SJ, Coyne P, St-Pierre E. Stress, physical activity, and screen-related sedentary behaviour within the first month of the COVID-19 pandemic. Appl Psychol Health Well Being 2021 May;13(2):454-468 [FREE Full text] [doi: 10.1111/aphw.12261] [Medline: 33645893]
40. Ding D, Cheng M, Del Pozo Cruz B, Lin T, Sun S, Zhang L, et al. How COVID-19 lockdown and reopening affected daily steps: evidence based on 164,630 person-days of prospectively collected data from Shanghai, China. Int J Behav Nutr Phys Act 2021 Mar 17;18(1):40 [FREE Full text] [doi: 10.1186/s12966-021-01106-x] [Medline: 33731132]
41. He M, Xian Y, Lv X, He J, Ren Y. Changes in body weight, physical activity, and lifestyle during the semi-lockdown period after the outbreak of COVID-19 in China: an online survey. Disaster Med Public Health Prep 2021 Apr;15(2):e23-e28 [FREE Full text] [doi: 10.1017/dmp.2020.237] [Medline: 32660669]
42. Ji Y, Shi Y, Zhou J, Li X, Qin R, Zhu Q. Analysis on the change of college students' life pattern and its impact during the COVID-19 outbreak in China. Am J Health Behav 2022 Jun 23;46(3):218-230. [doi: 10.5993/AJHB.46.3.2] [Medline: 35794758]
43. Zhu Y, Wang Z, Maruyama H, Onoda K, Huang Q, Hu C, et al. Effect of the COVID-19 lockdown period on the physical condition, living habits, and physical activity of citizens in Beijing, China. J Phys Ther Sci 2021 Sep;33(9):632-636 [FREE Full text] [doi: 10.1589/jpts.33.632] [Medline: 34539065]
44. Buoite Stella A, AjČeviĆ M, Furlanis G, Cillotto T, Menichelli A, Accardo A, et al. Smart technology for physical activity and health assessment during COVID-19 lockdown. J Sports Med Phys Fitness 2021 Mar;61(3):452-460 [FREE Full text] [doi: 10.23736/S0022-4707.20.11373-2] [Medline: 33092330]
45. Azuma K, Nojiri T, Kawashima M, Hanai A, Ayaki M, Tsubota K, TRF-Japan Study Group. Possible favorable lifestyle changes owing to the coronavirus disease 2019 (COVID-19) pandemic among middle-aged Japanese women: An ancillary survey of the TRF-Japan study using the original "Taberhythm" smartphone app. PLoS One 2021;16(3):e0248935 [FREE Full text] [doi: 10.1371/journal.pone.0248935] [Medline: 33765024]
46. Obuchi SP, Kawai H, Ejiri M, Ito K, Murakawa K. Change in outdoor walking behavior during the coronavirus disease pandemic in Japan: A longitudinal study. Gait Posture 2021 Jul;88:42-46 [FREE Full text] [doi: 10.1016/j.gaitpost.2021.05.005] [Medline: 33990001]
47. Sato K, Sakata R, Murayama C, Yamaguchi M, Matsuoka Y, Kondo N. Changes in work and life patterns associated with depressive symptoms during the COVID-19 pandemic: an observational study of health app (CALO mama) users. Occup Environ Med 2021 Sep;78(9):632-637 [FREE Full text] [doi: 10.1136/oemed-2020-106945] [Medline: 33619124]
48. Yamada Y, Namba H, Date H, Kitayama S, Nakayama Y, Kimura M, et al. Regional difference in the impact of COVID-19 pandemic on domain-specific physical activity, sedentary behavior, sleeping time, and step count: web-based cross-sectional nationwide survey and accelerometer-based observational study. JMIR Public Health Surveill 2023 Feb 20;9:e39992 [FREE Full text] [doi: 10.2196/39992] [Medline: 36634262]

49. Henriksen A, Johannessen E, Hartvigsen G, Grimsgaard S, Hopstock LA. Consumer-based activity trackers as a tool for physical activity monitoring in epidemiological studies during the COVID-19 pandemic: development and usability study. JMIR Public Health Surveill 2021 Apr 23;7(4):e23806 [FREE Full text] [doi: 10.2196/23806] [Medline: 33843598]
50. Park J, Yoo E, Kim Y, Lee J. What happened pre- and during COVID-19 in South Korea? Comparing physical activity, sleep time, and body weight status. Int J Environ Res Public Health 2021 May 29;18(11):5863 [FREE Full text] [doi: 10.3390/ijerph18115863] [Medline: 34072563]
51. Sañudo B, Fennell C, Sánchez-Oliver AJ. Objectively-assessed physical activity, sedentary behavior, smartphone use, and sleep patterns pre- and during-COVID-19 quarantine in young adults from Spain. Sustainability 2020 Jul 22;12(15):5890 [FREE Full text] [doi: 10.3390/su12155890]
52. Bird JM, Karageorghis CI, Hamer M. Relationships among behavioural regulations, physical activity, and mental health pre- and during COVID-19 UK lockdown. Psychol Sport Exerc 2021 Jul;55:101945 [FREE Full text] [doi: 10.1016/j.psychsport.2021.101945] [Medline: 34518758]
53. Hudgins BL, Kurti SP, Edwards ES, Hargens TA. The impact of the COVID-19 pandemic on physical activity habits at a residential university. J Am Coll Health 2021 Dec 14:1-6. [doi: 10.1080/07448481.2021.2016772] [Medline: 34905716]
54. Gollwitzer A, McLoughlin K, Martel C, Marshall J, Höhs JM, Bargh JA. Linking self-reported social distancing to real-world behavior during the COVID-19 pandemic. Soc Psychol Personal Sci 2021 Jun 23;13(2):656-668 [FREE Full text] [doi: 10.1177/19485506211018132]
55. Luo F, Ghanei Gheshlagh R, Dalvand S, Saedmoucheshi S, Li Q. Systematic review and meta-analysis of fear of COVID-19. Front Psychol 2021;12:661078 [FREE Full text] [doi: 10.3389/fpsyg.2021.661078] [Medline: 34177712]
56. Bargain O, Aminjonov U. Trust and compliance to public health policies in times of COVID-19. J Public Econ 2020 Dec;192:104316 [FREE Full text] [doi: 10.1016/j.jpubeco.2020.104316] [Medline: 33162621]
57. Pak A, McBryde E, Adegboye OA. Does high public trust amplify compliance with stringent COVID-19 government health guidelines? A multi-country analysis using data from 102,627 individuals. Risk Manag Healthc Policy 2021;14:293-302 [FREE Full text] [doi: 10.2147/RMHP.S278774] [Medline: 33542664]
58. Fridman I, Lucas N, Henke D, Zigler CK. Association between public knowledge about COVID-19, trust in information sources, and adherence to social distancing: cross-sectional survey. JMIR Public Health Surveill 2020 Sep 15;6(3):e22060 [FREE Full text] [doi: 10.2196/22060] [Medline: 32930670]
59. Tison GH, Avram R, Kuhar P, Abreau S, Marcus GM, Pletcher MJ, et al. Worldwide effect of COVID-19 on physical activity: a descriptive study. Ann Intern Med 2020 Nov 03;173(9):767-770. [doi: 10.7326/m20-2665]
60. Physical activity. World Health Organization. URL: https://www.who.int/health-topics/physical-activity#tab=tab_1 [accessed 2021-11-02]
61. Hall G, Laddu DR, Phillips SA, Lavie CJ, Arena R. A tale of two pandemics: How will COVID-19 and global trends in physical inactivity and sedentary behavior affect one another? Prog Cardiovasc Dis 2021;64:108-110 [FREE Full text] [doi: 10.1016/j.pcad.2020.04.005] [Medline: 32277997]
62. Yamada Y, Yoshida T, Nakagata T, Nanri H, Miyachi M. Letter to the editor: Age, sex, and regional differences in the effect of COVID-19 pandemic on objective physical activity in Japan: A 2-year nationwide longitudinal study. J Nutr Health Aging 2021;25(8):1032-1033 [FREE Full text] [doi: 10.1007/s12603-021-1662-y] [Medline: 34545926]
63. Tison GH, Barrios J, Avram R, Kuhar P, Bostjancic B, Marcus GM, et al. Worldwide physical activity trends since COVID-19 onset. Lancet Glob Health 2022 Oct;10(10):e1381-e1382 [FREE Full text] [doi: 10.1016/S2214-109X(22)00361-8] [Medline: 36057269]
64. Lawhun Costello V, Chevance G, Wing D, Mansour-Assi SJ, Sharp S, Golaszewski NM, et al. Impact of the COVID-19 pandemic on objectively measured physical activity and sedentary behavior among overweight young adults: yearlong longitudinal analysis. JMIR Public Health Surveill 2021 Nov 24;7(11):e28317 [FREE Full text] [doi: 10.2196/28317] [Medline: 34665759]
65. Cornelius T, Denes A, Webber KT, Guest C, Goldsmith J, Schwartz JE, et al. Relationship quality and objectively measured physical activity before and after implementation of COVID-19 stay-home orders. J Health Psychol 2022 Sep;27(10):2390-2401. [doi: 10.1177/13591053211042075] [Medline: 34435516]
66. Barbieri PN, Giuntella O, Saccardo S, Sadoff S. Lifestyle and mental health 1 year into COVID-19. Sci Rep 2021 Dec 02;11(1):23349 [FREE Full text] [doi: 10.1038/s41598-021-02702-4] [Medline: 34857806]
67. Runacres A, Mackintosh KA, Knight RL, Sheeran L, Thatcher R, Shelley J, et al. Impact of the COVID-19 pandemic on sedentary time and behaviour in children and adults: a systematic review and meta-analysis. Int J Environ Res Public Health 2021 Oct 27;18(21):11286 [FREE Full text] [doi: 10.3390/ijerph182111286] [Medline: 34769800]
68. Stockwell S, Trott M, Tully M, Shin J, Barnett Y, Butler L, et al. Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: a systematic review. BMJ Open Sport Exerc Med 2021;7(1):e000960 [FREE Full text] [doi: 10.1136/bmjsem-2020-000960] [Medline: 34192010]
69. Staying physically active during the COVID-19 pandemic. American College of Sports Medicine. URL: https://www.acsm.org/read-research/newsroom/news-releases/news-detail/2020/03/16/staying-physically-active-during-covid-19-pandemic [accessed 2022-04-30]

Abbreviations

ISRCTN: International Standard Randomized Controlled Trial Number Register
MD: mean difference
MOOSE: Meta-analysis Of Observational Studies in Epidemiology
NOS: Newcastle-Ottawa Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis

---

Edited by T Sanchez, A Mavragani; submitted 30.06.22; peer-reviewed by Y Yamada, S Amagasa; comments to author 10.02.23; revised version received 23.03.23; accepted 27.04.23; published 30.05.23.

Please cite as:
Wu Z, Wang Y, Zhang Y, Bennell KL, White DK, Shen L, Ren W, Wei J, Zeng C, Lei G
The COVID-19 Pandemic and Daily Steps in the General Population: Meta-analysis of Observational Studies
JMIR Public Health Surveill 2023;9:e40650
URL: https://publichealth.jmir.org/2023/1/e40650
doi:10.2196/40650
PMID:37252779

Original Paper

Adoption of a National Prophylactic Anticoagulation Guideline for Hospitalized Pregnant Women With COVID-19: Retrospective Cohort Study

Yeon-Mi Hwang[1,2], MSc; Samantha Piekos[1], PhD; Tanya Sorensen[2,3], MD; Leroy Hood[1], MD, PhD; Jennifer Hadlock[1,2], MD

[1]Institute for Systems Biology, Seattle, WA, United States
[2]University of Washington, Seattle, WA, United States
[3]Swedish Medical Center, Providence Swedish, Seattle, WA, United States

Corresponding Author:
Jennifer Hadlock, MD
Institute for Systems Biology
401 Terry Ave N
Seattle, WA, 98109
United States
Phone: 1 206 732 1359
Email: jhadlock@systemsbiology.org

Abstract

Background: Both COVID-19 and pregnancy are associated with hypercoagulability. Due to the increased risk for thrombosis, the United States National Institute of Health’s recommendation for prophylactic anticoagulant use for pregnant patients has expanded from patients hospitalized for severe COVID-19 manifestation to all patients hospitalized for the manifestation of COVID-19 (no guideline: before December 26, 2020; first update: December 27, 2022; second update: February 24, 2022-present). However, no study has evaluated this recommendation.

Objective: The objective of this study was to characterize prophylactic anticoagulant use among hospitalized pregnant people with COVID-19 from March 20, 2020, to October 19, 2022.

Methods: This was a retrospective cohort study in large US health care systems across 7 states. The cohort of interest was pregnant patients who were hospitalized with COVID-19, without previous coagulopathy or contraindication to anticoagulants (n=2767). The treatment group consisted of patients prescribed prophylactic dose anticoagulation between 2 days before and 14 days after COVID-19 treatment onset (n=191). The control group was patients with no anticoagulant exposure between 14 days before and 60 days after COVID-19 treatment onset (n=2534). We ascertained the use of prophylactic anticoagulants with attention to the updates in guidelines and emerging SARS-CoV-2 variants. We propensity score matched the treatment and control group 1:1 on the most important features contributing to the prophylactic anticoagulant administration status classification. Outcome measures included coagulopathy, bleeding, COVID-19–related complications, and maternal-fetal health outcomes. Additionally, the inpatient anticoagulant administration rate was validated in a nationwide population from Truveta, a collective of 700 hospitals across the United States.

Results: The overall administration rate of prophylactic anticoagulants was 7% (191/2725). It was lowest after the second guideline update (no guideline: 27/262, 10%; first update: 145/1663, 8.72%; second update: 19/811, 2.3%; *P*<.001) and during the omicron-dominant period (Wild type: 45/549, 8.2%; Alpha: 18/129, 14%; Delta: 81/507, 16%; and Omicron: 47/1551, 3%; *P*<.001). Models developed on retrospective data showed that the variable most associated with the administration of inpatient prophylactic anticoagulant was comorbidities prior to SARS-CoV-2 infection. The patients who were administered prophylactic anticoagulant were also more likely to receive supplementary oxygen (57/191, 30% vs 9/188, 5%; *P*<.001). There was no statistical difference in a new diagnosis of coagulopathy, bleeding, or maternal-fetal health outcomes between those who received treatment and the matched control group.

Conclusions: Most hospitalized pregnant patients with COVID-19 did not receive prophylactic anticoagulants across health care systems as recommended by guidelines. Guideline-recommended treatment was administered more frequently to patients with greater COVID-19 illness severity. Given the low rate of administration and differences between treated and untreated cohorts, efficacy could not be assessed.

(JMIR Public Health Surveill 2023;9:e45586) doi:10.2196/45586

KEYWORDS

anticoagulants stillbirth; COVID-19; electronic health record; low birth weight; NIH antithrombotic therapy guideline; pregnancy; preterm birth; SARS-CoV-2; small for gestational age; thromboprophylaxis

Introduction

Both COVID-19 and pregnancy are associated with thrombosis [1-13]. Severe COVID-19 is accompanied by Virchow's triad of endothelial injury, stasis, and hypercoagulable state: categories of factors contributing to thrombosis [1-10]. The related virus, SARS-CoV-2, directly invades endothelial cells, leading to endothelial injury, exocytosis, and endotheliitis [1-4,8,14,15]. Hospitalization causes stasis of blood flow due to decreased mobility regardless of the presence of the SARS-CoV-2 virus [16]. COVID-19–induced changes in prothrombotic factors are linked to an increased risk of hypercoagulability [5-7,17-26]. Studies have reported elevated factor VIII, elevated fibrinogen, circulating prothrombotic microparticles, neutrophil extracellular traps, and hyperviscosity among severely ill patients with COVID-19 [5-7,17-20]. Multiple meta-analyses studies report high rates of coagulopathy and thrombosis across multiple tissues and organ sites among patients infected with COVID-19 since early in the pandemic [21-26].

Hypercoagulability is a physiologically adaptive mechanism that prevents bleeding from miscarriage, childbirth, and postpartum hemorrhage. Women in pregnancy or the postpartum period are at a 4- to 5-fold higher risk of thromboembolism than nonpregnant women [11-13]. Normal pregnancy is accompanied by increased concentration of factors VII, VIII, X, and von Willebrand factor and by pronounced increases in fibrinogen [12,27,28]. Due to the combination of increased risk of hypercoagulability, the National Institutes of Health (NIH) COVID-19 Treatment Guideline Panel recommends using a prophylactic dose of anticoagulation for pregnant patients who are hospitalized for COVID-19 unless a contraindication exists, since February 24, 2022 [29].

We hypothesized that we would observe a high prophylactic anticoagulant administration rate after the second update in NIH COVID-19 treatment guidelines. As prophylactic anticoagulant was recommended to all hospitalized pregnant patients with COVID-19, we expected a minimal difference in clinical condition between the treatment and control group, leading to minimized confounding by indication. This would allow us to properly evaluate the impact of prophylactic anticoagulants on inpatient pregnant patients with COVID-19. No study yet has evidence for or against the guideline on recommending prophylactic anticoagulants on inpatient pregnant patients with COVID-19 [29].

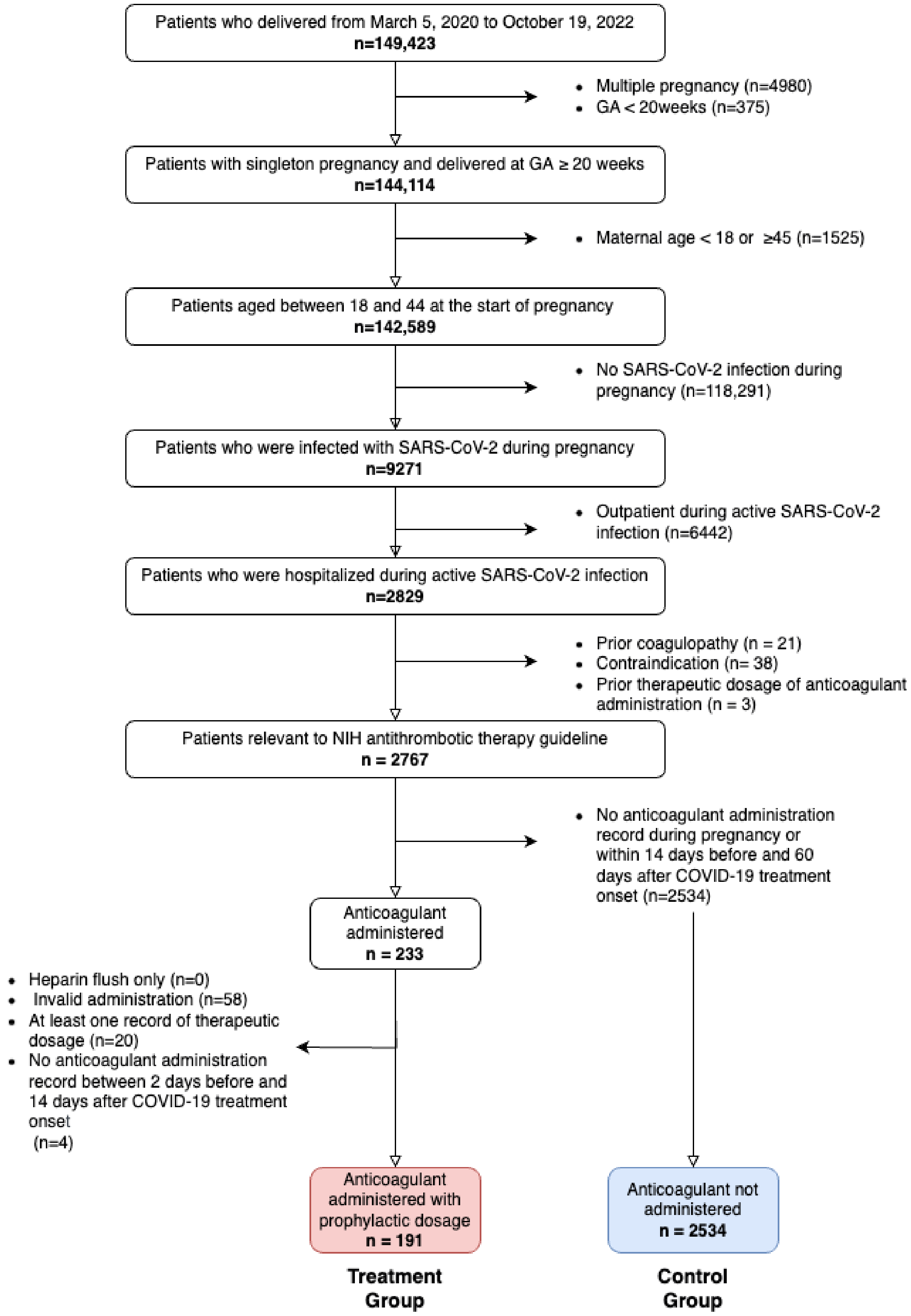
Here, we aim to characterize prophylactic anticoagulant use among hospitalized pregnant patients with COVID-19. We assessed the difference in the prophylactic anticoagulant use across the timeline of NIH guideline changes and SARS-CoV-2 variants. We evaluated the association between prophylactic anticoagulant use and risks of coagulopathy, COVID-19, and maternal-fetal health outcomes, after addressing the discrepancy in variables that could influence outcomes.

Methods

Study Setting and Population

Providence Health and Services and Affiliates (PHSA) is an integrated not-for-profit US community health care system that provides care in urban and rural settings across 7 states: Alaska, California, Montana, Oregon, New Mexico, Texas, and Washington. PHSA service includes 52 hospitals, 1085 clinics, and 120,000 caregivers. We used PHSA electronic health records of pregnant patients who delivered from January 26, 2020, through October 19, 2022 (n=149,423). Multimedia Appendix 1 ([30-36]) provides supplementary methods, tables, and figures. Figure S1 in Multimedia Appendix 1 and Figure 1 describe the cohort selection. Table S1 in Multimedia Appendix 1 defines variables and terminology. We excluded multiple pregnancies and deliveries with a gestational age of less than 20 weeks (n=126,261). We limited our analyses to pregnant patients aged between 18 and 45 years (n=144,114). Our inclusion criterion was COVID-19 diagnosis during the pregnancy period (n=9271). We excluded patients who were not hospitalized with COVID-19 (n=2829). The definition of this exclusion criterion was no overlap of hospitalization stay with an infection period between 7 days before and 14 days after the COVID-19 diagnosis date. This infection period was determined based on previous studies on incubation and symptomatic period [30]. We excluded patients with any coagulopathy event before COVID-19 diagnosis, contraindication to anticoagulant administration, or a record of therapeutic anticoagulant dosage [31] in the past 2 years (Tables S2 and S3 in Multimedia Appendix 1). We defined the remaining patients as our cohort of interest, relevant to the NIH antithrombotic therapy guideline (n=2767). This study was reported following STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [37].

Figure 1. Cohort selection flowchart. We excluded patients with any history of coagulopathy before the time of SARS-CoV-2 infection. We excluded patients with any contradiction to anticoagulant if the diagnosis was between 2 years before the pregnancy and the time of SARS-CoV-2 infection. Contraindications included major bleeding, peptic ulcer, stage 2 hypertension, esophageal varices, intracranial mass, end-stage liver diseases, aneurysm, proliferative retinopathy, and bleeding disorders. We considered administration to be valid unless the administration status was “Canceled Entry,” “Held,” “Missed,” “Automatically Held,” “Held by Provider,” “MAR Hold,” “Stopped,” or “Paused.” GA: gestational age; NIH: National Institutes of Health.



Variables

Exposures

Figure S1 in Multimedia Appendix 1 and Figure 1 describe definitions of exposure groups. From the cohort of interest (n=2767), patients with no anticoagulant administration record during pregnancy or within 14 days before and 60 days after the SARS-CoV-2 infection date comprised the control group (no anticoagulant administration group, n=2534). We defined the treatment group (prophylactic anticoagulant administration group, n=191) as patients who received prophylactic anticoagulation between 2 days before and 14 days after the potential COVID-19 treatment onset (Figure S1 and Tables S1, S3, and S4 in Multimedia Appendix 1). Our definition of prophylactic included prophylactic and intermediate dosages (Table S4 in Multimedia Appendix 1) [31].

Outcomes

Variables and diagnoses are defined in Tables S1 and S2 in Multimedia Appendix 1. Our primary outcome of interest was coagulopathy. We additionally observed diagnoses relevant to the consequences of COVID-19-associated coagulopathy [8,38,39]. These included thrombosis, pulmonary embolism, thromboembolism, myocardial infarction, skin necrosis or purpura, and stroke (Table S1 in Multimedia Appendix 1). Secondary outcomes were COVID-19 severity and maternal-fetal health outcomes. COVID-19 complications included the need for supplemental oxygen, vasopressor usage, mortality, length of hospital stay, unique diagnosis count, and unique medication count. Maternal-fetal outcomes were stillbirth, preterm birth (PTB), low birth weight (LBW), and small for gestational age. We also assessed the risk of bleeding during pregnancy and postpartum hemorrhage to evaluate the safety of anticoagulant use.

Covariates

Maternal characteristics were evaluated, including parity, gravidity, history of preterm delivery, age, race, ethnicity, insurance, pregravid BMI, smoking, illegal drug use, Centers for Disease Control and Prevention Social Vulnerability Index (CDC-SVI), and rural-urban classification (Table S1 in Multimedia Appendix 1). The count of unique diagnoses before COVID-19 was used to reflect the patient's previous comorbidity. Missing values were imputed using the cohort median. COVID-19–related features assessed were vaccination, previous SARS-CoV-2 infection, presumed SARS-CoV-2 variant, timing of SARS-CoV-2 infection, and stage of NIH COVID-19 antithrombotic therapy guidelines (Supplementary Methods and Table S5 in Multimedia Appendix 1). We defined the stages of guidelines based on major updates in NIH antithrombotic therapy guidelines on pregnant women [29]. Up until December 20, 2020, there was no specific guideline for pregnant women. On December 21, 2020, there was an update to recommend the prophylactic use of anticoagulants among pregnant patients with severe COVID-19. On February 24, 2022, a second update expanded the recommendation to any hospitalized patient with COVID-19.

Analyses

Descriptive Statistics

P value was calculated using the chi-square test and the 2-sample *t* test. Multiple testing error was corrected using Bonferroni correction. We performed pairwise Pearson correlation between variables using python library scipy (version 1.6.2). We ranked the count of diagnoses each exposure group received between COVID-19 diagnosis and delivery.

Classification Model, Feature Importance, and Propensity Score Matching

We conducted a classification model and feature importance analysis to investigate the differences between the treatment and control groups, and to select variables to be addressed in propensity score matching. Using 28 covariates (Table S1 in Multimedia Appendix 1), we built classification models, including logistic regression, random forest, and gradient boosting machine (GBM) models (Supplementary Methods in Multimedia Appendix 1). We applied random undersampling to address a class imbalance between the control and treatment groups. Due to the small sample size, we leveraged leave-one-out cross-validation. We then used Gini feature importance and SHAP (Shapley additive explanation) [40] to identify which covariates were most important for classifying the administration of prophylactic anticoagulant during a COVID-19–related hospitalization in the best-performing model (Supplementary Methods in Multimedia Appendix 1). We evaluated whether the model trained with these top features achieved comparable performance to the original model with 28 features. We then performed propensity score matching to control for these top features between the treatment and control groups. Compared to other propensity score methods and covariate adjustment methods, propensity score matching provided exceptional covariant balance across most circumstances [41]. Given the extreme weights between treatment and control, the inverse probability treatment weighting method (IPTW) was considered unsuitable for this analysis. K-nearest neighbors (k=1) were used to match with replacement across covariates using propensity logit and caliper of 0.2 using Python library PsmPy (version 0.2.8) to identify patients most similar on the top features and generate a matched control group [42]. The number of neighbors (k) and caliper threshold value were selected based on recommendations from previous studies [43,44]. The effect size of these matched features was evaluated using Cohen *d* score before and after matching with a score of <0.2 indicating a small effect size [45]. We evaluated outcome differences between the treatment and matched control groups using Fisher exact test for categorical variables and the Mann-Whitney *U* test for continuous variables.

Sensitivity Analysis

We assessed the influence of illness severity at the time of SARS-CoV-2 infection using the count of medication 3 days before and 3 days after the potential COVID-19 treatment onset. These counts of medications are measures of relative clinical severity in a generally healthy patient population when patients are unlikely to need mechanical ventilation. Although they are imperfect proxies, they have been shown to provide valuable insight, especially in COVID-19–related research. The date range for collecting the medication count was selected based on the distribution of the time gap between the anticoagulant administration date and the COVID-19 treatment onset date (Figure S3 in Multimedia Appendix 1). We conducted propensity score matching on this medication count at the time of SARS-CoV-2 infection, in addition to the best-performing features. We also addressed the influence of maternal age and pregravid BMI, as these variables are known risk factors for coagulopathy.

Investigation of Inpatient Anticoagulant Administration Rate Across Multiple Health Care Systems Using Truveta

We calculated the rate of the inpatient anticoagulant administration rate in the Truveta patient population (Figure S2 in Multimedia Appendix 1). Truveta is a consortium of 28 health care systems, including PHSA, providing patient care in over 20,000 clinics and 700 hospitals across 43 states [46]. Similar data fields across systems are mapped following the common schema referred to as the Truveta Data Model [47]. Among pregnant patients aged between 18 and 45 years and delivered singleton from January 26, 2020, through October 19, 2022, we identified patients with a COVID-19 diagnosis during pregnancy. We excluded patients who did not have any inpatient encounters during the active COVID-19 infection. We further excluded patients who had any anticoagulant use before pregnancy and defined the remaining patients as our Truveta cohort of interest, relevant to NIH antithrombotic therapy guidelines. From this analytic cohort, we calculated patients who were administered inpatient anticoagulants.

Ethical Considerations

All procedures were reviewed and approved by the institutional review board at the PHSA through expedited review on November 4, 2020 (study number STUDY2020000196). Consent was waived because the disclosure of protected health information for the study involved no more than minimal risk to the privacy of individuals. Patient data in the Truveta Studio were deidentified.

Results

Our cohort of interest was composed of 2767 patients (Figure S1 in Multimedia Appendix 1 and Figure 1). A total of 191 and 2534 patients were in the treatment and control groups, respectively. The treatment group was older (mean age 30.9 years; *P*=.02), less likely to be vaccinated (33/191, 17%; *P*<.001), less likely to be infected during the third trimester (139/191, 72.8%; *P*<.001), and more likely to be infected with Delta variant (81/191, 42%; *P*<.001). The treatment group was exposed to enoxaparin and heparin only (Table 1). The date difference between the anticoagulant administration date and potential COVID-19 treatment onset (anticoagulant administration date – COVID-19 treatment onset) ranged from –2 to 14 days. A total of 74.9% (143/191) of the treatment group received anticoagulant prescriptions during the first 3 days of COVID-19 treatment (Figure S3 in Multimedia Appendix 1).

Figure 2 displays the overall timeline of the prophylactic anticoagulant administration status from March 20, 2020, to October 19, 2022. The overall prevalence rate of prophylactic anticoagulant administration was 7% (191/2725). The administration rate was the lowest when the Omicron variant was dominant (47/1551, 3%) and after the second guideline update (19/811, 2.3%). It was the highest when the Delta variant was dominant (81/507, 16%) and when there was no guideline (27/262, 10%). The top 10 diagnoses of the control group were subcategories of COVID-19 and pregnancy (Figure S4 in Multimedia Appendix 1). The treatment group additionally included COVID-19 complications, such as pneumonia, acute respiratory failure with hypoxia, and hypokalemia (Figure S4 in Multimedia Appendix 1).

Figure S2 in Multimedia Appendix 1 displays the cohort selection procedure in the Truveta patient population. We identified 14,075 patients as our Truveta cohort of interest. Among these patients, 973 (6.9%) patients were administered inpatient anticoagulants.

The GBM model had the best performance with an area under the receiver operating characteristics curve (AUC-ROC) of 0.84 (95% CI 0.81-0.87; Figure S5 and Table S6 in Multimedia Appendix 1). Pre–COVID-19 diagnoses count, variant-omicron, socioeconomic status, third-trimester infection, housing type and transportation vulnerability, minority status and language vulnerability, and household composition and disability vulnerability were the most important 7 features (Figures S5 in Multimedia Appendix 1). The GBM model trained with these 7 features reached an AUC-ROC of 0.85 (95% CI 0.83-0.89; Figure S5 in Multimedia Appendix 1 and Figure 3). In the SHAPley figure, red and blue dots indicate high and low feature values. Red dots of pre–COVID-19 diagnoses count were clustered in the positive end of the SHAP axis. Red dots of third-trimester infection and variant Omicron were clustered in the negative SHAP axis. The matched control group (n=188) was generated by propensity score matching on the 7 most important features. These features had small effect sizes after matching (Cohen *d* values <0.2 [45] Table S7 in Multimedia Appendix 1).

The GBM model trained with the 7 top features and initial medication count reached an AUC-ROC of 0.93 (95% 0.91-0.95; Figure S6 in Multimedia Appendix 1). The feature importance showed that initial medication count was the most important feature classifying prophylactic anticoagulant administration status (Figure S6 in Multimedia Appendix 1). The red and blue dots were spread out across extreme negative SHAP values and up to 0.2 SHAP value, but the end of the positive SHAP axis (SHAP value >0.2) was clustered with red dots (Figure S6 in Multimedia Appendix 1).

Compared with the control group, the treatment group had an increased risk of maternal death (relative risk [RR] not calculable due to the small number of cases; 3/191, 1.6% vs 0/2545, 0%; *P*<.001), O_2 assistance (RR 9.3, 95% CI 6.9-12.3; 57/191, 33% vs 91/2545, 3.6%; *P*<.001), LBW (RR 2.1, 95% CI 1.5-2.8; 36/190, 19% vs 233/2539, 9.2%; *P*<.001), PTB (RR 2.1, 95% CI 1.5-2.8; 47/191, 25% vs 295/2545, 12%; *P*<.001), and bleeding (RR 2.5, 95% CI 1-6.3; 5/191, 2.6% vs 27/2545, 1.1%; *P*=.08). After matching, risk of LBW (RR 1.2, 95% CI 0.8-1.8; *P*=.50), PTB (RR 1.2, 95% CI 0.8-1.8; *P*=.33), and bleeding (RR 1.0, 95% CI 0.3-3.3; *P*>.99) were reduced and no longer statistically significant. We also did not observe an increase or decrease in risk between the treatment coagulopathy and other maternal-fetal health outcomes (Tables S8 and S9 in Multimedia Appendix 1 and Figure 4). However, the treatment group had significantly worse outcomes regarding COVID-19 illness severity even after the matching. The treatment group had a higher likelihood of supplemental oxygen (RR 6.9, 95% CI 3.5-13.4; 57/191, 30% vs 9/188, 5%; *P*<.001), longer inpatient stay (median 6 vs 3; *P*<.001), higher medication (median 51 vs 39; *P*<.001), and diagnoses count (median 7 vs 0; *P*<.001). The statistical significance remained even in our sensitivity analysis (Tables S8 and S9 in Multimedia Appendix 1) where we additionally addressed the initial illness severity. Maternal age and pregravid BMI did not influence the association between the treatment and outcomes of interest.

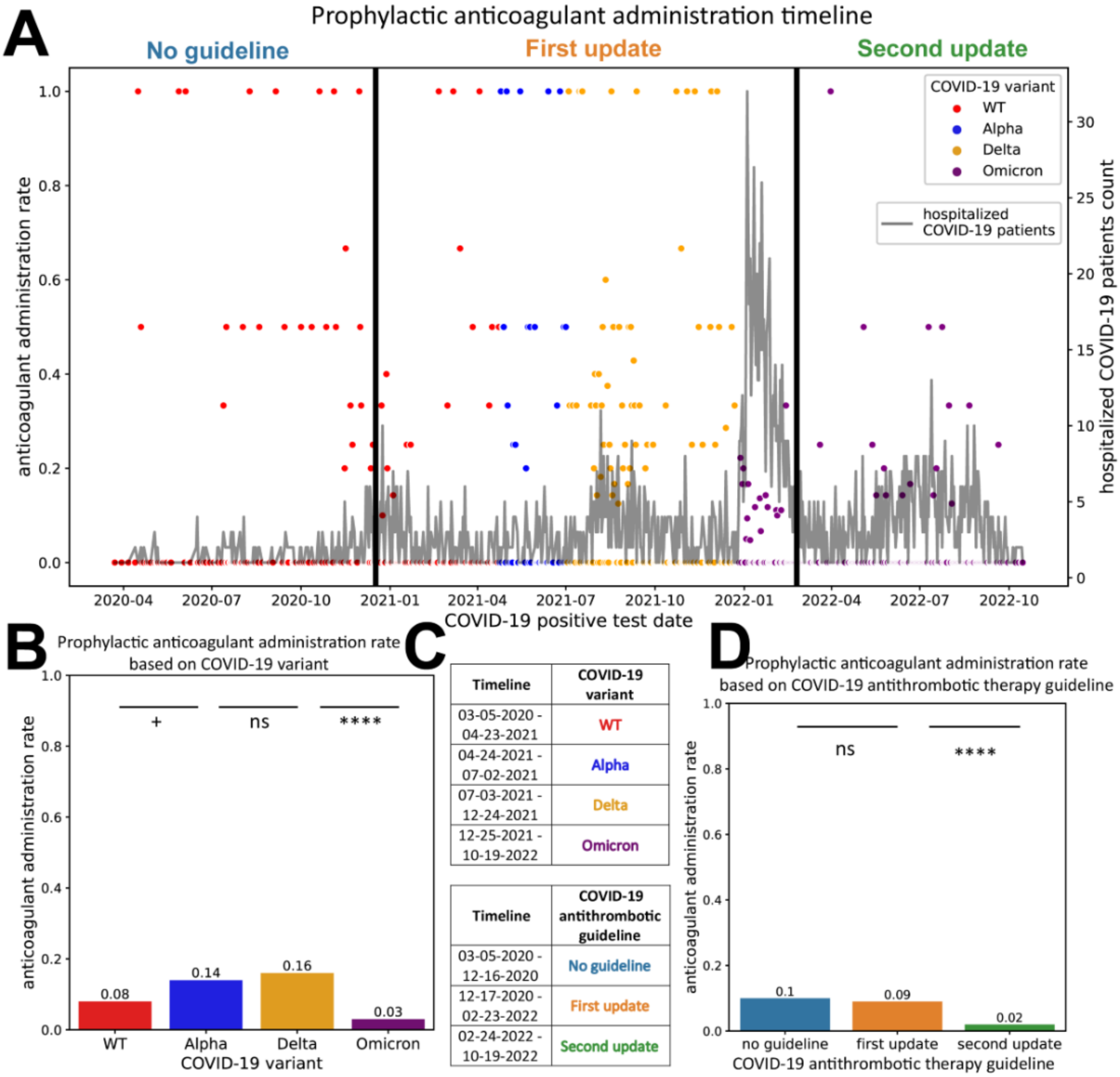
Table 1. Descriptive statistics of treatment and control group. Variables are defined in Table S1 in Multimedia Appendix 1. This table was generated using tableone PyPI package. *P* value was calculated using the chi-square test and the 2-sample t test. Multiple testing error was corrected using Bonferroni correction.

	Missing, n	Overall (N=2736)	Control group (n=2545)	Treatment group (n=191)	*P* value (adjusted)
Maternal age (years), mean (SD)	0	29.7 (5.7)	29.6 (5.7)	30.9 (5.3)	.02
Maternal age group (years), n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	.01
18-24		647 (23.6)	619 (24.3)	28 (14.7)	
25-29		758 (27.7)	713 (28.0)	45 (23.6)	
30-34		797 (29.1)	718 (28.2)	79 (41.4)	
35-40		435 (15.9)	402 (15.8)	33 (17.3)	
41-44		99 (3.6)	93 (3.7)	6 (3.1)	
Unknown		0 (0.0)	0 (0.0)	0 (0.0)	
Race group, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	.02
American Indian or Alaska Native		45 (1.6)	42 (1.7)	3 (1.6)	
Asian		166 (6.1)	146 (5.7)	20 (10.5)	
Black or African American		126 (4.6)	111 (4.4)	15 (7.9)	
Native Hawaiian or other Pacific Islander		46 (1.7)	38 (1.5)	8 (4.2)	
White or Caucasian		1466 (53.6)	1381 (54.3)	85 (44.5)	
Multiracial		150 (5.5)	138 (5.4)	12 (6.3)	
Other		620 (22.7)	580 (22.8)	40 (20.9)	
Unknown		117 (4.3)	109 (4.3)	8 (4.2)	
Ethnic group, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	.74
Hispanic or Latino		1077 (39.4)	1013 (39.8)	64 (33.5)	
Not Hispanic or Latino		1567 (57.3)	1444 (56.7)	123 (64.4)	
Unknown		92 (3.4)	88 (3.5)	4 (2.1)	
Pregravid BMI, n (%)	1888	848 (100.0)	792 (100.0)	56 (100.0)	.15
Underweight (below 18.5)		11 (1.3)	11 (1.4)	0 (0.0)	
Healthy weight (18.5-24.9)		326 (38.4)	314 (39.6)	12 (21.4)	
Overweight (25-29.9)		241 (28.4)	224 (28.3)	17 (30.4)	
Obesity (>30)		270 (31.8)	243 (30.7)	27 (48.2)	
Gravidity, n (%)	24	2712 (100.0)	2521 (100.0)	191 (100.0)	>.99
1-5		2529 (93.3)	2354 (93.4)	175 (91.6)	
>5		13 (6.7)	167 (6.6)	16 (8.4)	
Parity, n (%)	24	2712 (100.0)	2521 (100.0)	191 (100.0)	>.99
0		381 (14.0)	355 (14.1)	26 (13.6)	
1-5		2298 (84.7)	2136 (84.7)	162 (84.8)	
>5		33 (1.2)	30 (1.2)	3 (1.6)	
Preterm history, n (%)	433	2303 (100.0)	2145 (100.0)	158 (100.0)	>.99
0		2012 (87.4)	1879 (87.6)	133 (84.2)	
1		291 (12.6)	266 (12.4)	25 (15.8)	
Insurance status, n (%)	1	2736 (100.0)	2545 (100.0)	191 (100.0)	>.99
Commercial		959 (35.1)	890 (35.0)	69 (36.1)	
Medicaid		1771 (64.8)	1649 (64.8)	122 (63.9)	
Medicare		3 (0.1)	3 (0.1)	0 (0.0)	

	Missing, n	Overall (N=2736)	Control group (n=2545)	Treatment group (n=191)	*P* value (adjusted)
Uninsured self-pay		2 (0.1)	2 (0.1)	0 (0.0)	
Smoking status, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	0.49
No		2488 (90.9)	2322 (91.2)	166 (86.9)	
Yes		248 (9.1)	223 (8.8)	25 (13.1)	
Illegal drug use status, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	>.99
No		2399 (87.7)	2226 (87.5)	173 (90.6)	
Yes		337 (12.3)	319 (12.5)	18 (9.4)	
Rural-urban classification, n (%)	538	2198 (100.0)	2029 (100.0)	166 (100.0)	>.99
Metropolitan		1997 (90.9)	1846 (91.0)	151 (89.3)	
Micropolitan		129 (5.9)	114 (5.6)	15 (8.9)	
Rural		31 (1.4)	30 (1.5)	1 (0.6)	
Small town		41 (1.9)	39 (1.9)	2 (1.2)	
Socioeconomic status vulnerability, mean (SD)	455	0.5 (0.3)	0.5 (0.3)	0.5 (0.3)	>.99
Household composition and disability vulnerability, mean (SD)	452	0.4 (0.3)	0.4 (0.3)	0.5 (0.3)	>.99
Minority status and language vulnerability, mean (SD)	452	0.7 (0.3)	0.7 (0.3)	0.6 (0.2)	.15
Housing type and transportation vulnerability, mean (SD)	455	0.6 (0.3)	0.6 (0.3)	0.6 (0.3)	>.99
Vaccination status, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	<.001
No		1803 (65.9)	1645 (64.6)	158 (82.7)	
Yes		933 (34.1)	900 (35.4)	33 (17.3)	
Previous SARS-CoV-2 infection, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	>.99
No		2690 (98.3)	2500 (98.2)	190 (99.5)	
Yes		46 (1.7)	45 (1.8)	1 (0.5)	
Trimester of SARS-CoV-2 infection, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	<.001
First trimester		25 (0.9)	18 (0.7)	7 (3.7)	
Second trimester		120 (4.4)	75 (2.9)	45 (23.6)	
Third trimester		2591 (94.7)	2452 (96.3)	139 (72.8)	
SARS-CoV-2 variant, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	<.001
Wild type		549 (20.1)	504 (19.8)	45 (23.6)	
Alpha		129 (4.7)	111 (4.4)	18 (9.4)	
Delta		507 (18.5)	426 (16.7)	81 (42.4)	
Omicron		1551 (56.7)	1504 (59.1)	47 (24.6)	
NIH[a] antithrombotic therapy guideline, n (%)	0	2736 (100.0)	2545 (100.0)	191 (100.0)	<.001
No guideline		262 (9.6)	235 (9.2)	27 (14.1)	
First update		1663 (60.8)	1518 (59.6)	145 (75.9)	
Second update		811 (29.6)	792 (31.1)	19 (9.9)	
Diagnosis count before SARS-CoV-2 infection, mean (SD)	0	1.9 (3.6)	1.9 (3.7)	2 (3.3)	>.99

[a] NIH: National Institutes of Health.

Figure 2. Timeline of prophylactic anticoagulant administration among hospitalized COVID-19 pregnant patients from March 20, 2020, to October 19, 2022. The anticoagulant administration rate is defined as the count of the patients who belong to the treatment group divided by the sum of patients who belong to the treatment group and control group. (A) Timeline of prophylactic anticoagulant administration rate among hospitalized COVID-19 pregnant patients from March 20, 2020, to October 19, 2022. Marked time points are described in part C of the figure. Timeline definitions: red, blue, yellow, and purple dots indicate the prophylactic anticoagulant administration rate among COVID-19 wild type, Alpha, Delta, and Omicron variants, respectively. Grayline chart indicates the count of hospitalized pregnant patients with COVID-19. The first (left) vertical line is on December 17, 2020, the first update date in COVID-19 antithrombotic therapy guidelines to recommend the administration of prophylactic anticoagulants on pregnant patients hospitalized with severe COVID-19 manifestation. The second (right) vertical line is on February 24, 2022, the second update date to expand the recommendation to all pregnant patients hospitalized with COVID-19 manifestation. The overall administration rate of prophylactic anticoagulants was 7% (191/2725) (B) Prophylactic anticoagulant administration rate based on SARS-CoV-2 variant. The SARS-CoV-2 variant was determined based on the period during which each variant was the dominant variant accounting for >50% of cases as part of the Centers for Disease Control and Prevention (CDC) genomic surveillance for SARS-CoV-2 in Region 10 (Alaska, Idaho, Oregon, and Washington; CDC 2022).(C) Prophylactic anticoagulant administration rate was lowest during the omicron-dominant period (Wild type: 45/549, 8.2%; Alpha: 18/129, 14%; Delta: 81/507, 16%; and Omicron: 47/1551, 3%) (C) Timeline definition. Text colors indicate the SARS-CoV-2 variant and COVID-19 antithrombotic therapy guideline time period in parts A, B, and D of the figure. (D) Prophylactic anticoagulant administration rate based on COVID-19 antithrombotic therapy guideline. The first update was to recommend the administration of prophylactic anticoagulants on pregnant patients hospitalized with severe COVID-19 manifestations. The second update expanded the recommendation to all pregnant patients hospitalized with COVID-19 manifestation. Prophylactic anticoagulant administration rate was lowest after the second guideline update (no guideline: 27/262, 10%; first update: 145/1663, 8.7%; second update: 19/811, 2.3%). WT: wild type. ^{ns}P>.05, *P≤.05, **P≤.01, ***P≤.001, ****P≤.0001.



Timeline	COVID-19 variant
03-05-2020 - 04-23-2021	WT
04-24-2021 - 07-02-2021	Alpha
07-03-2021 - 12-24-2021	Delta
12-25-2021 - 10-19-2022	Omicron

Timeline	COVID-19 antithrombotic guideline
03-05-2020 - 12-16-2020	No guideline
12-17-2020 - 02-23-2022	First update
02-24-2022 - 10-19-2022	Second update

Figure 3. Performance of best-performing classification model and feature importance. (A) Performance of original and limited gradient boosted machine learning model classifying prophylactic anticoagulant administration status among hospitalized pregnant patients with COVID-19. The limited model was trained with the 7 most important features from the original model. The feature importance result of the original model is in Figure S5 in Multimedia Appendix 1. The 95% CI of the original and limited models were (0.81-0.87) and (0.83-0.89), respectively. (B) Feature importance ranking of the limited model. Variables are defined in Table S1 in Multimedia Appendix 1. (C) Shapley permutation explainer of the feature contribution. SHAP value reflects the contribution of the 7 most important features from the gradient boosting models toward classifying prophylactic anticoagulant administration status. SHAP value is the average marginal contribution of a feature value across all permutations of features. Each row represents an individual feature, and the dot represents a sample. The dot color reflects the value of the feature of the sample relative to all samples. The evaluation was done on the sample set composed of a treatment group and a 1:1 randomly undersampled matched control group (n=382). AUC: area under the curve; GBM: gradient boosting machine.

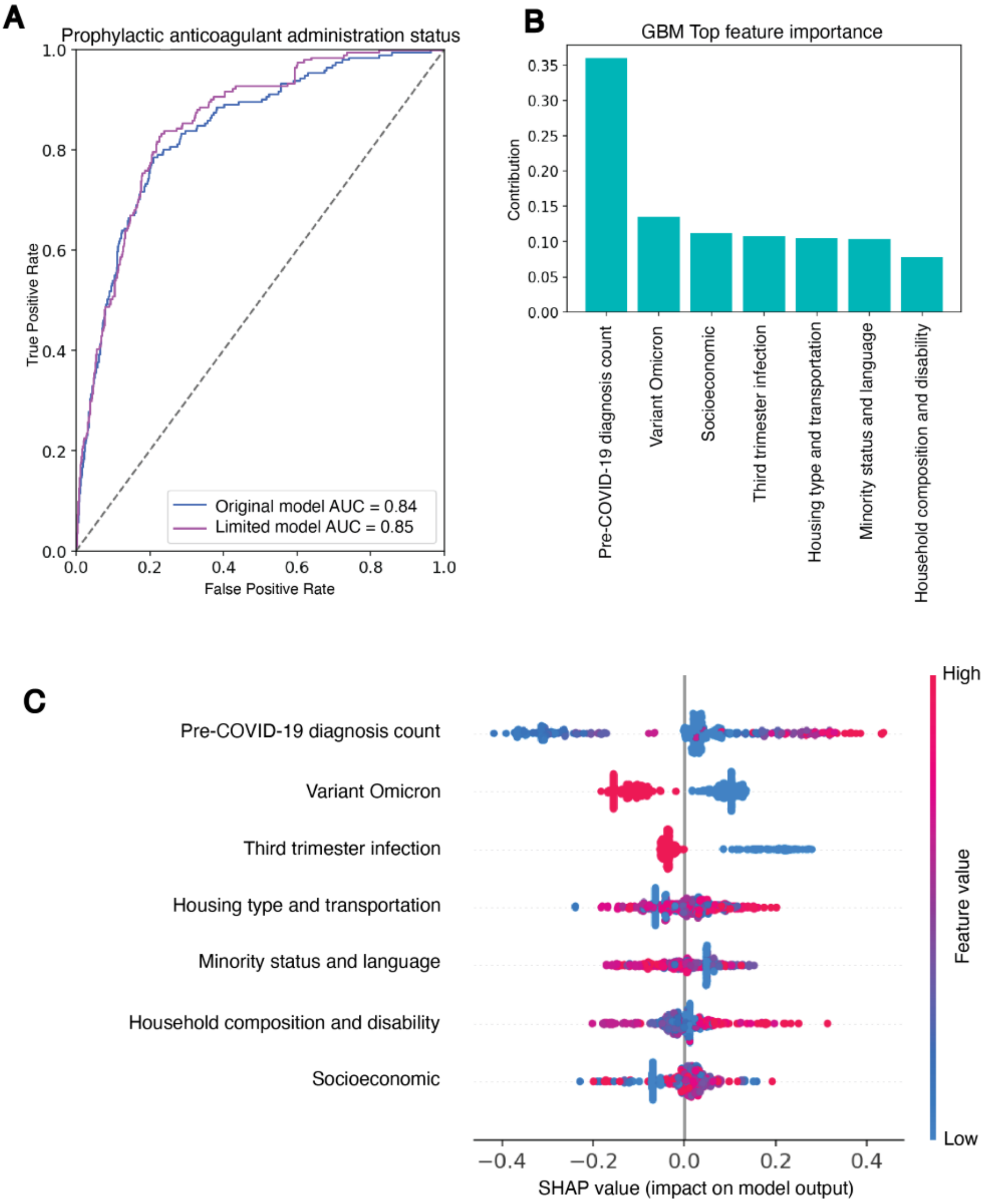
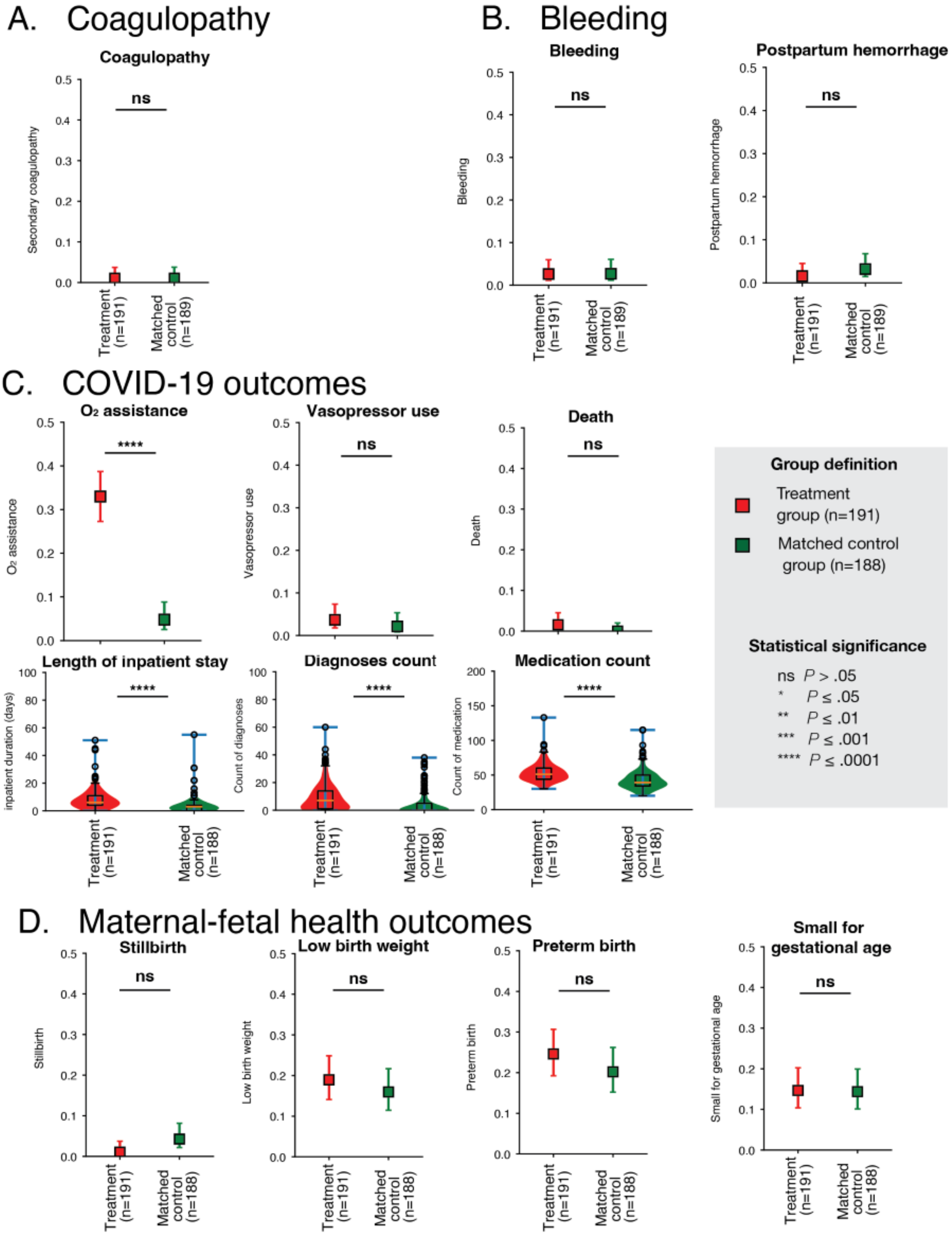


Figure 4. Clinical outcomes of the treatment and matched control. Results are presented in Tables S7 and S8 in Multimedia Appendix 1. Unless specified, the observation cutoff was the delivery date. The observation start point was the anticoagulant exposure date and potential COVID-19 treatment onset date, respectively. (A) Coagulopathy prevalence rate of the treatment and matched control. SNOMED codes we used for coagulopathy diagnosis are listed in Table S1 in Multimedia Appendix 1. (B) Bleeding and postpartum hemorrhage prevalence rate of the treatment and matched control. SNOMED codes we used for bleeding and postpartum hemorrhage diagnoses are listed in Table S1 in Multimedia Appendix 1. Postpartum hemorrhage was collected from the delivery date to +12 weeks postpartum. (C) O_2 assistance use, vasopressor use, maternal death prevalence rates for the treatment, and matched control groups. Violin plots of the length of inpatient stay (days), medication counts 48 hours after the potential COVID-19 treatment onset, and diagnoses count after the potential COVID-19 treatment onset. The cutoff point was the delivery date. (D) Stillbirth, low birth weight, preterm birth, small for gestational age prevalence rate of the treatment, matched control, and sensitivity analysis matched control group.



Discussion

Principal Results

To our knowledge, this was the first study evaluating the NIH antithrombotic therapy guideline on pregnant women. We initially hypothesized to observe a high prophylactic anticoagulant administration rate among hospitalized COVID-19 pregnant patients and minimal difference in patient characteristics between the treatment and control groups. However, we identified most patients relevant to NIH antithrombotic therapy guidelines, pregnant women hospitalized with COVID-19, did not receive prophylactic anticoagulant dosage. Instead, patients with higher comorbidity levels, infected in the first or second trimester, and infected with a non-Omicron variant were more likely to receive prophylactic anticoagulant. Despite having more severe COVID-19, the treatment group did not have an elevated risk of coagulopathy, bleeding, and maternal-fetal health outcomes.

Comparisons With Previous Works, Interpretation, and Implication

The most important and interesting finding was an unexpectedly low prophylactic anticoagulant administration rate among hospitalized COVID-19 pregnant patients across health care systems despite the guideline. A possible explanation for this observation is that clinicians administered prophylactic anticoagulants based on the patient's COVID-19 illness severity. The initial medication count, a proxy variable for COVID-19 illness severity at COVID-19 treatment onset, was the most important feature in the sensitivity analysis model classifying prophylactic anticoagulant administration status. It also enhanced the area under the curve score by 0.09 reaching 0.94. Besides, we noticed the treatment group was enriched with diagnoses indicative of severe COVID-19 symptoms, such as pneumonia, acute respiratory failure with hypoxia, and hypokalemia. In contrast, the control group was not. We presume clinicians did not follow the guideline because it was not strongly recommended and lacked studies to support it. The rating scheme of the NIH guideline is BIII. B and III, respectively, indicate a moderate recommendation and expert opinion evidence. This means no study yet has evidence for or against this recommendation [42]. Few studies describing COVID-19-positive pregnant patients have reported prophylactic use of anticoagulants during hospitalization. In these studies, the sample size was small, ranging from 9 to 20, and complications following the treatment were not evaluated [48-50].

We noted an elevated risk of respiratory assistance among the treatment group. This was inconsistent with previous studies. Although there was no study on pregnant patients, there have been multiple observational studies supporting prophylactic anticoagulant use on nonpregnant hospitalized patients with COVID-19. Studies showed that prophylactic anticoagulants improved freedom from intubation and lowered mortality, indicating better COVID-19 outcomes [51-53]. In these studies with a nonpregnant population, including men and an older population, the rate of prophylactic anticoagulant administration ranged from 45% (1959/4389) to 84% (3627/4297), whereas that of our study was 7% (191/2736). In addition, in our study, prophylactic anticoagulant administration was highly correlated with the comorbidity level before the SARS-CoV-2 infection and the initial illness severity. We do not think the increased risk in respiratory assistance in our study was influenced by the exposure to anticoagulants, but rather a byproduct of failure to overcome the confounding by indication.

Although we could only partially address the confounding by indication, we did not observe elevated risks of coagulopathy, bleeding, and maternal-fetal health outcomes among the treatment group. Given that severe COVID-19 is associated with coagulopathy and adverse pregnancy outcomes [1-10,54,55], this finding was somewhat promising as we expected worse outcomes in coagulopathy and maternal-fetal health outcomes. Based on our findings, we cannot support the benefit of the guideline, but can, at least, support its potential safety regarding maternal-fetal health outcomes. The treatment group had a similar likelihood to deliver low birthweight, preterm, and small for gestational age babies. This observation was less limited with a small sample size, compared to rare outcomes of interest, due to the high incidence rate of cases across groups. Considering this is the first study to evaluate the guideline, our study can contribute to our knowledge in treating hospitalized pregnant patients with COVID-19. Future research should be conducted in larger and various study settings and minimize the confounding by indication to understand better the risk, safety, and benefit of the treatment.

Strengths and Limitations

The small percentage of patients receiving guideline-recommended treatment was both an interesting observation and a major limitation. We expected to observe a high prophylactic anticoagulant administration rate after the second guideline update as all hospitalized COVID-19 pregnant patients were relevant, but only 7% (191/2736) were administered prophylactic anticoagulants. This led to confounding by indication. The treatment group had higher pre-COVID-19 comorbidity levels and included more critically ill patients. Although we attempted to minimize the bias using propensity score matching, we could not overcome it due to the unexpectedly small sample size and lack of appropriate variables reflecting COVID-19 illness severity at the time of infection. This study was conducted on retrospective structured data, which is insufficient to determine medical reasoning. In addition, individuals had a wide range of anticoagulant administration start times before and after the potential COVID-19 treatment onset. As covariates that occurred after the exposure are inappropriate for the propensity score method, we excluded them from the main matching model. Nevertheless, we performed sensitivity analysis on the impact of the initial medication count to partially address the COVID-19 illness severity at the time of treatment onset. Last, we did not verify whether the heightened risk of O_2 assistance among the treatment group resulted from confounding by indication or not. We assumed it was due to confounding by indication based on the previous studies [51-53] on hospitalized patients with COVID-19. More generally, electronic health record data are known to contain errors; inpatient medication administration records are carefully tracked and unlikely to be omitted, but it is not uncommon for comorbidities to be undiagnosed, unreported, or listed in more than one way. To address all of these limitations, a similar study should be conducted in a larger prospective cohort, with appropriate variables to assess COVID-19 illness severity at the time of admission.

This was the first study assessing the prophylactic use of anticoagulants on hospitalized patients after the guideline update on February 24, 2022. Although the guideline recommends on all hospitalized patients regardless of COVID-19 severity, no study to date has assessed the exposure of anticoagulants among COVID-19 hospitalized patients. Another strength of our study includes adjustments for social and economic risk factors. We adjusted for differences between CDC-SVI scores and rural and urban classifications in the treatment and control groups based on the census tract level. Social vulnerability and rural health disparities are significant risk factors for adverse pregnancy outcomes and severe COVID-19 symptoms [56-63]. As the census tract is more granular than the county level, it captures the environmental factors to which patients have been exposed more accurately. Indeed, all 4 CDC-SVI indexes were important predictors of anticoagulant administration status. This study was conducted on a COVID-19 maternity population that has been investigated by other researchers [64,65], allowing deeper insight into the study setting and population.

We validated the number of inpatient anticoagulant administration rates in Truveta's patient population, comprised of various health care systems. Overall inpatient anticoagulant administration rates in Truveta were similar to those observed in PHSA. This indicates that the unexpectedly low adaptation rate of NIH recommendations was not limited to the PHSA health care system. Here, we could not completely replicate our analysis as Truveta and Providence data models were different, and the leading researcher of this study had limited access to Truveta Studio. We did not exclude patients who were contraindicated to anticoagulants and did not differentiate between prophylactic and therapeutic dosage. Nevertheless, we expect our final estimation of the inpatient anticoagulant administration rate on the Truveta patient population would not drastically change based on the number we observed from the cohort selection procedure in the PHSA population.

Conclusions

We found that in a population of patients from multiple health care systems nationwide, only a small percentage of hospitalized pregnant patients received the prophylactic anticoagulation recommended by the NIH antithrombotic therapy guideline. Guideline recommended prophylactic anticoagulants were administered more frequently to patients with greater COVID-19 illness severity. We were not able to show the efficacy of prophylactic anticoagulation. However, anticoagulation was not associated with elevated risks of coagulopathy, bleeding, and maternal-fetal health outcomes despite the worse health conditions of the treatment group. This study cannot make a recommendation to administer prophylactic anticoagulant to hospitalized COVID-19 pregnant patients. However, as this was the first study characterizing prophylactic anticoagulant use after the NIH guideline update on February 24, 2022, these results can help inform patient-specific clinical decisions. An observational study with a larger sample size should be considered to further evaluate outcomes with prophylactic anticoagulation.

Acknowledgments

This work was funded by a grant from the National Institutes of Health (HD091527). We thank Providence Health and Services and Affiliates for sharing their data engineering expertise and computational resources. We would also like to acknowledge SNOMED International for developing and maintaining SNOMED-CT.

Data Availability

Results have been aggregated and reported within this paper to the extent possible while maintaining privacy from personal health information as required by law. Providence Health and Services and Affiliates data are archived within Providence Health and Services and Affiliates in a Health Insurance Portability and Accountability Act (HIPAA)–secure audited computer environment. Truveta data are available to Truveta subscribers. Requests to access these data sets can be directed to Truveta Inc., info@truveta.com. All clinical logic for data extraction has been shared within the paper and Multimedia Appendix 1. The code is publicly available on GitHub [66].

Authors' Contributions

YMH contributed to conceptualization, methodology, software, formal analysis, investigation, writing original draft preparation, and visualization. SP was involved in conceptualization, methodology, software, investigation, writing, reviewing, and editing. TS was responsible for supervision, writing, reviewing, and editing. LH contributed to conceptualization and supervision. JH was involved in conceptualization, supervision, writing, reviewing, and editing.

Conflicts of Interest

LH is a scientific advisor for Sera Prognostics, a company that develops diagnostic tests for pregnancy complications. The company is not associated with this study or any of the findings. JH reports grants from Pfizer and Novartis and research contracts from Janssen, Gilead Scientific, and Bristol Myers Squibb for work unrelated to this study. The remaining authors report no conflict of interest.

Multimedia Appendix 1
Supplementary methods, tables, and figures.
[DOCX File , 2659 KB - publichealth_v9i1e45586_app1.docx]

References

1. Otifi HM, Adiga BK. Endothelial dysfunction in Covid-19 infection. Am J Med Sci 2022;363(4):281-287 [FREE Full text] [doi: 10.1016/j.amjms.2021.12.010] [Medline: 35093394]
2. Teuwen LA, Geldhof V, Pasut A, Carmeliet P. COVID-19: the vasculature unleashed. Nat Rev Immunol 2020;20(7):389-391 [FREE Full text] [doi: 10.1038/s41577-020-0343-0] [Medline: 32439870]
3. Lowenstein CJ, Solomon SD. Severe COVID-19 is a microvascular disease. Circulation 2020;142(17):1609-1611 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.120.050354] [Medline: 32877231]
4. Libby P, Lüscher T. COVID-19 is, in the end, an endothelial disease. Eur Heart J 2020;41(32):3038-3044 [FREE Full text] [doi: 10.1093/eurheartj/ehaa623] [Medline: 32882706]
5. Panigada M, Bottino N, Tagliabue P, Grasselli G, Novembrino C, Chantarangkul V, et al. Hypercoagulability of COVID-19 patients in intensive care unit: a report of thromboelastography findings and other parameters of hemostasis. J Thromb Haemost 2020;18(7):1738-1742 [FREE Full text] [doi: 10.1111/jth.14850] [Medline: 32302438]
6. Ranucci M, Ballotta A, Di Dedda U, Baryshnikova E, Dei Poli M, Resta M, et al. The procoagulant pattern of patients with COVID-19 acute respiratory distress syndrome. J Thromb Haemost 2020;18(7):1747-1751 [FREE Full text] [doi: 10.1111/jth.14854] [Medline: 32302448]
7. Maier CL, Truong AD, Auld SC, Polly DM, Tanksley C, Duncan A. COVID-19-associated hyperviscosity: a link between inflammation and thrombophilia? Lancet 2020;395(10239):1758-1759 [FREE Full text] [doi: 10.1016/S0140-6736(20)31209-5] [Medline: 32464112]
8. Conway EM, Mackman N, Warren RQ, Wolberg AS, Mosnier LO, Campbell RA, et al. Understanding COVID-19-associated coagulopathy. Nat Rev Immunol 2022;22(10):639-649 [FREE Full text] [doi: 10.1038/s41577-022-00762-9] [Medline: 35931818]
9. Levy JH, Iba T, Gardiner EE. Endothelial injury in COVID-19 and acute infections: putting the pieces of the puzzle together. Arterioscler Thromb Vasc Biol 2021;41(5):1774-1776 [FREE Full text] [doi: 10.1161/ATVBAHA.121.316101] [Medline: 33792345]
10. Huertas A, Montani D, Savale L, Pichon J, Tu L, Parent F, et al. Endothelial cell dysfunction: a major player in SARS-CoV-2 infection (COVID-19)? Eur Respir J 2020 Jul;56(1):2001634 [FREE Full text] [doi: 10.1183/13993003.01634-2020] [Medline: 32554538]
11. Heit JA, Kobbervig CE, James AH, Petterson TM, Bailey KR, Melton LJ. Trends in the incidence of venous thromboembolism during pregnancy or postpartum: a 30-year population-based study. Ann Intern Med 2005 Nov 15;143(10):697-706. [doi: 10.7326/0003-4819-143-10-200511150-00006] [Medline: 16287790]
12. James AH. Pregnancy-associated thrombosis. Hematology Am Soc Hematol Educ Program 2009:277-285. [doi: 10.1182/asheducation-2009.1.277] [Medline: 20008211]
13. Pomp ER, Lenselink AM, Rosendaal FR, Doggen CJM. Pregnancy, the postpartum period and prothrombotic defects: risk of venous thrombosis in the MEGA study. J Thromb Haemost 2008;6(4):632-637 [FREE Full text] [doi: 10.1111/j.1538-7836.2008.02921.x] [Medline: 18248600]
14. Montiel V, Lobysheva I, Gérard L, Vermeersch M, Perez-Morga D, Castelein T, et al. Oxidative stress-induced endothelial dysfunction and decreased vascular nitric oxide in COVID-19 patients. EBioMedicine 2022;77:103893 [FREE Full text] [doi: 10.1016/j.ebiom.2022.103893] [Medline: 35219085]
15. Xu SW, Ilyas I, Weng JP. Endothelial dysfunction in COVID-19: an overview of evidence, biomarkers, mechanisms and potential therapies. Acta Pharmacol Sin 2023;44(4):695-709 [FREE Full text] [doi: 10.1038/s41401-022-00998-0] [Medline: 36253560]
16. Hospitalization and Blood Clots. Centers for Disease Control and Prevention Internet. 2022. URL: https://www.cdc.gov/ncbddd/dvt/hospitalization-infographic.html [accessed 2022-12-03]
17. Gorog DA, Storey RF, Gurbel PA, Tantry US, Berger JS, Chan MY, et al. Current and novel biomarkers of thrombotic risk in COVID-19: a consensus statement from the international COVID-19 thrombosis biomarkers colloquium. Nat Rev Cardiol 2022;19(7):475-495 [FREE Full text] [doi: 10.1038/s41569-021-00665-7] [Medline: 35027697]
18. Semeraro N, Colucci M. The prothrombotic state associated with SARS-CoV-2 infection: pathophysiological aspects. Mediterr J Hematol Infect Dis 2021;13(1):e2021045 [FREE Full text] [doi: 10.4084/MJHID.2021.045] [Medline: 34276914]
19. Acherjee T, Behara A, Saad M, Vittorio TJ. Mechanisms and management of prothrombotic state in COVID-19 disease. Ther Adv Cardiovasc Dis 2021;15:17539447211053470 [FREE Full text] [doi: 10.1177/17539447211053470] [Medline: 34693818]
20. von Meijenfeldt FA, Havervall S, Adelmeijer J, Lundström A, Rudberg AS, Magnusson M, et al. Prothrombotic changes in patients with COVID-19 are associated with disease severity and mortality. Res Pract Thromb Haemost 2021;5(1):132-141 [FREE Full text] [doi: 10.1002/rth2.12462] [Medline: 33537537]

21. Malas MB, Naazie IN, Elsayed N, Mathlouthi A, Marmor R, Clary B. Thromboembolism risk of COVID-19 is high and associated with a higher risk of mortality: a systematic review and meta-analysis. EClinicalMedicine 2020;29:100639 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100639] [Medline: 33251499]
22. Tan BK, Mainbourg S, Friggeri A, Bertoletti L, Douplat M, Dargaud Y, et al. Arterial and venous thromboembolism in COVID-19: a study-level meta-analysis. Thorax 2021;76(10):970-979 [FREE Full text] [doi: 10.1136/thoraxjnl-2020-215383] [Medline: 33622981]
23. Suh YJ, Hong H, Ohana M, Bompard F, Revel MP, Valle C, et al. Pulmonary embolism and deep vein thrombosis in COVID-19: a systematic review and meta-analysis. Radiology 2021;298(2):E70-E80 [FREE Full text] [doi: 10.1148/radiol.2020203557] [Medline: 33320063]
24. Katsoularis I, Fonseca-Rodríguez O, Farrington P, Jerndal H, Lundevaller EH, Sund M, et al. Risks of deep vein thrombosis, pulmonary embolism, and bleeding after covid-19: nationwide self-controlled cases series and matched cohort study. BMJ 2022;377:e069590 [FREE Full text] [doi: 10.1136/bmj-2021-069590] [Medline: 35387772]
25. Xiao D, Tang F, Chen L, Gao H, Li X. Cumulative evidence for the association of thrombosis and the prognosis of COVID-19: systematic review and meta-analysis. Front Cardiovasc Med 2021;8:819318 [FREE Full text] [doi: 10.3389/fcvm.2021.819318] [Medline: 35146009]
26. Kollias A, Kyriakoulis KG, Lagou S, Kontopantelis E, Stergiou GS, Syrigos K. Venous thromboembolism in COVID-19: a systematic review and meta-analysis. Vasc Med 2021;26(4):415-425 [FREE Full text] [doi: 10.1177/1358863X21995566] [Medline: 33818197]
27. Brenner B. Haemostatic changes in pregnancy. Thromb Res 2004;114(5-6):409-414. [doi: 10.1016/j.thromres.2004.08.004] [Medline: 15507271]
28. Battinelli EM, Marshall A, Connors JM. The role of thrombophilia in pregnancy. Thrombosis 2013;2013:516420 [FREE Full text] [doi: 10.1155/2013/516420] [Medline: 24455235]
29. NIH COVID-19 Treatment Guidelines. National Institute of Health. URL: https://www.covid19treatmentguidelines.nih.gov/therapies/antithrombotic-therapy/ [accessed 2023-06-20]
30. Wu Y, Kang L, Guo Z, Liu J, Liu M, Liang W. Incubation period of COVID-19 caused by unique SARS-CoV-2 strains: a systematic review and meta-analysis. JAMA Netw Open 2022;5(8):e2228008 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.28008] [Medline: 35994285]
31. Bauer KA. Use of anticoagulants during pregnancy and postpartum. UpToDate. Waltham, MA: UpToDate; 2022. URL: https://www.uptodate.com/contents/use-of-anticoagulants-during-pregnancy-and-postpartum [accessed 2023-06-20]
32. Geospatial Research, Analysis, and Services Program. CDC/ATSDR Social Vulnerability Index. Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry. 2020. URL: https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html [accessed 2022-12-07]
33. U.S. Department of Agriculture’s Economic Research Service (ERS). Rural-Urban Commuting Area Codes. 2010. URL: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/ [accessed 2022-12-07]
34. COVID Data Tracker. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker/#variant-proportions [accessed 2022-11-15]
35. Symptoms of COVID-19. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html [accessed 2023-04-25]
36. Sun X, Xu W. Fast Implementation of DeLong’s Algorithm for Comparing the Areas Under Correlated Receiver Operating Characteristic Curves. IEEE Signal Process Lett 2014;21(11):1389-1393 [FREE Full text] [doi: 10.1109/LSP.2014.2337313]
37. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596):1453-1457 [FREE Full text] [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
38. Leentjens J, van Haaps TF, Wessels PF, Schutgens REG, Middeldorp S. COVID-19-associated coagulopathy and antithrombotic agents-lessons after 1 year. Lancet Haematol 2021;8(7):e524-e533 [FREE Full text] [doi: 10.1016/S2352-3026(21)00105-8] [Medline: 33930350]
39. Görlinger K, Dirkmann D, Gandhi A, Simioni P. COVID-19-associated coagulopathy and inflammatory response: what do we know already and what are the knowledge gaps? Anesth Analg 2020;131(5):1324-1333 [FREE Full text] [doi: 10.1213/ANE.0000000000005147] [Medline: 33079850]
40. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Red Hook, NY, United States: Curran Associates Inc; 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4 - 9, 2017; Long Beach California USA URL: https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
41. Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, et al. Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies. J Am Coll Cardiol 2017;69(3):345-357 [FREE Full text] [doi: 10.1016/j.jacc.2016.10.060] [Medline: 28104076]
42. Kline A, Luo Y. PsmPy: a package for retrospective cohort matching in python. 2022 Presented at: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 11-15 July 2022; Glasgow, Scotland, United Kingdom p. 1354-1357. [doi: 10.1109/EMBC48229.2022.9871333]

43. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011;46(3):399-424 [FREE Full text] [doi: 10.1080/00273171.2011.568786] [Medline: 21818162]
44. Austin PC. Statistical criteria for selecting the optimal number of untreated subjects matched to each treated subject when using many-to-one matching on the propensity score. Am J Epidemiol 2010;172(9):1092-1097 [FREE Full text] [doi: 10.1093/aje/kwq224] [Medline: 20802241]
45. Sullivan GM, Feinn R. Using effect size-or why the P value is not enough. J Grad Med Educ 2012;4(3):279-282 [FREE Full text] [doi: 10.4300/JGME-D-12-00156.1] [Medline: 23997866]
46. Truveta's health system members. Truveta. 2022. URL: https://www.truveta.com/members/ [accessed 2023-03-30]
47. Althouse BM, Baker C, Smits PD, Gratzl S, Lee RH, Cartwright BMG, et al. Racial inequality in COVID-treatment and in-hospital length of stay in the US over time. Front Public Health 2022;10:1074775 [FREE Full text] [doi: 10.3389/fpubh.2022.1074775] [Medline: 36711416]
48. Koyuncu K, Sakin Ö, Aktas HA, Sahin K, Aygün T, Angin AD, et al. Thromboprophylaxis in Covid-19 positive pregnant women. South Clin Ist Euras 2020;31(3):281-286 [FREE Full text] [doi: 10.14744/scie.2020.86548]
49. Pierce-Williams RAM, Burd J, Felder L, Khoury R, Bernstein PS, Avila K, et al. Clinical course of severe and critical coronavirus disease 2019 in hospitalized pregnancies: a United States cohort study. Am J Obstet Gynecol MFM 2020;2(3):100134 [FREE Full text] [doi: 10.1016/j.ajogmf.2020.100134] [Medline: 32391519]
50. Jevtic SD, Malinowski AK, Othman M, Kadir RAA. Physician experiences in management of COVID-19-associated coagulopathy in pregnancy: communication from the ISTH SSC subcommittee on women's health issues in thrombosis and haemostasis. J Thromb Haemost 2021;19(10):2539-2545 [FREE Full text] [doi: 10.1111/jth.15462] [Medline: 34260818]
51. Nadkarni GN, Lala A, Bagiella E, Chang HL, Moreno PR, Pujadas E, et al. Anticoagulation, bleeding, mortality, and pathology in hospitalized patients with COVID-19. J Am Coll Cardiol 2020;76(16):1815-1826 [FREE Full text] [doi: 10.1016/j.jacc.2020.08.041] [Medline: 32860872]
52. Rentsch CT, Beckman JA, Tomlinson L, Gellad WF, Alcorn C, Kidwai-Khan F, et al. Early initiation of prophylactic anticoagulation for prevention of coronavirus disease 2019 mortality in patients admitted to hospital in the United States: cohort study. BMJ 2021;372:n311 [FREE Full text] [doi: 10.1136/bmj.n311] [Medline: 33574135]
53. Vaughn VM, Yost M, Abshire C, Flanders SA, Paje D, Grant P, et al. Trends in venous thromboembolism anticoagulation in patients hospitalized with COVID-19. JAMA Netw Open 2021;4(6):e2111788 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.11788] [Medline: 34115129]
54. Smith ER, Oakley E, Grandner GW, Ferguson K, Farooq F, Afshar Y, Perinatal COVID PMA Study Collaborators. Adverse maternal, fetal, and newborn outcomes among pregnant women with SARS-CoV-2 infection: an individual participant data meta-analysis. BMJ Glob Health 2023;8(1):e009495 [FREE Full text] [doi: 10.1136/bmjgh-2022-009495] [Medline: 36646475]
55. DeSisto CL, Wallace B, Simeone RM, Polen K, Ko JY, Meaney-Delman D, et al. Risk for stillbirth among women with and without COVID-19 at delivery hospitalization - United States, march 2020-september 2021. MMWR Morb Mortal Wkly Rep 2021;70(47):1640-1645 [FREE Full text] [doi: 10.15585/mmwr.mm7047e1] [Medline: 34818318]
56. Shah LM, Varma B, Nasir K, Walsh MN, Blumenthal RS, Mehta LS, et al. Reducing disparities in adverse pregnancy outcomes in the United States. Am Heart J 2021;242:92-102. [doi: 10.1016/j.ahj.2021.08.019] [Medline: 34481757]
57. Kent ST, McClure LA, Zaitchik BF, Gohlke JM. Area-level risk factors for adverse birth outcomes: trends in urban and rural settings. BMC Pregnancy Childbirth 2013;13:129 [FREE Full text] [doi: 10.1186/1471-2393-13-129] [Medline: 23759062]
58. Sharma G, Ying W, Vaught AJ. Understanding the rural and racial disparities in pre-pregnancy hypertension: important considerations in maternal health equity. J Am Coll Cardiol 2020;76(22):2620-2622 [FREE Full text] [doi: 10.1016/j.jacc.2020.09.602] [Medline: 33183895]
59. Oates GR, Juarez LD, Horswell R, Chu S, Miele L, Fouad MN, et al. The association between neighborhood social vulnerability and COVID-19 testing, positivity, and incidence in Alabama and Louisiana. J Community Health 2021;46(6):1115-1123 [FREE Full text] [doi: 10.1007/s10900-021-00998-x] [Medline: 33966116]
60. Raja R, Mukherjee I, Sarkar BK. A machine learning-based prediction model for preterm birth in rural India. J Healthc Eng 2021;2021:6665573 [FREE Full text] [doi: 10.1155/2021/6665573] [Medline: 34234931]
61. Mah JC, Andrew MK. Social vulnerability indices: a pragmatic tool for COVID-19 policy and beyond. Lancet Reg Health Eur 2022;14:100333 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100333] [Medline: 35194590]
62. Biggs EN, Maloney PM, Rung AL, Peters ES, Robinson WT. The relationship between social vulnerability and COVID-19 incidence among Louisiana census tracts. Front Public Health 2020;8:617976 [FREE Full text] [doi: 10.3389/fpubh.2020.617976] [Medline: 33553098]
63. Salmanian B, McKinney JR, Moufarrij S, Grace R, Sangi-Haghpeykar H, Eppes C, et al. 1011 Pregnancy affected by COVID-19, patient characteristics and outcomes based on social vulnerability index (SVI). Am J Obstet Gynecol 2021;224(2):S626-S627 [FREE Full text] [doi: 10.1016/j.ajog.2020.12.1036]
64. Piekos SN, Roper RT, Hwang YM, Sorensen T, Price ND, Hood L, et al. The effect of maternal SARS-CoV-2 infection timing on birth outcomes: a retrospective multicentre cohort study. Lancet Digit Health 2022;4(2):e95-e104 [FREE Full text] [doi: 10.1016/S2589-7500(21)00250-8] [Medline: 35034863]

65. Piekos SN, Hwang YM, Roper RT, Sorensen T, Price ND, Hood L, et al. The effect of COVID-19 vaccination and booster on maternal-fetal outcomes: a retrospective multicenter cohort study. medRxiv Preprint posted online on August 18, 2022. [FREE Full text] [doi: 10.1101/2022.08.12.22278727] [Medline: 36032974]
66. COVID19 Maternity Inpatient Anticoagulant Public. GitHub. URL: https://github.com/Hadlock-Lab/COVID19_Maternity_Inpatient_Anticoagulant [accessed 2023-07-03]

Abbreviations

AUC-ROC: area under the receiver operating characteristics curve
CDC-SVI: Centers for Disease Control and Prevention Social Vulnerability Index
GBM: gradient boosting machine
LBW: low birth weight
NIH: National Institutes of Health
PHSA: Providence Health and Services and affiliates
PTB: preterm birth
RR: relative risk
SHAP: Shapley additive explanation
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 09.01.23; peer-reviewed by V Chauhan, S Middeldorp; comments to author 13.04.23; revised version received 05.05.23; accepted 13.06.23; published 28.07.23.

Please cite as:
Hwang YM, Piekos S, Sorensen T, Hood L, Hadlock J
Adoption of a National Prophylactic Anticoagulation Guideline for Hospitalized Pregnant Women With COVID-19: Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e45586
URL: https://publichealth.jmir.org/2023/1/e45586
doi:10.2196/45586
PMID:37311123

Original Paper

Using Bandit Algorithms to Maximize SARS-CoV-2 Case-Finding: Evaluation and Feasibility Study

Michael F Rayo[1], PhD; Daria Faulkner[2], MPH; David Kline[3], PhD; Thomas Thornhill IV[4], MPH; Samuel Malloy[5], MA; Dante Della Vella[1], MSc; Dane A Morey[1], PhD; Net Zhang[5], BSc; Gregg Gonsalves[4], MPhil, PhD

[1]Department of Integrated Systems Engineering, College of Engineering, The Ohio State University, Columbus, OH, United States

[2]College of Public Health, The Ohio State University, Columbus, OH, United States

[3]Department of Biostatistics and Data Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[4]Public Health Modeling Unit, Department of the Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, United States

[5]Battelle Center for Science, Engineering, and Public Policy, John Glenn College of Public Affairs, The Ohio State University, Columbus, OH, United States

Corresponding Author:
Gregg Gonsalves, MPhil, PhD
Public Health Modeling Unit
Department of the Epidemiology of Microbial Diseases
Yale School of Public Health
350 George Street, 3rd Floor
New Haven, CT, 06511
United States
Phone: 1 2036069149
Email: gregg.gonsalves@yale.edu

Abstract

Background: The Flexible Adaptive Algorithmic Surveillance Testing (FAAST) program represents an innovative approach for improving the detection of new cases of infectious disease; it is deployed here to screen and diagnose SARS-CoV-2. With the advent of treatment for COVID-19, finding individuals infected with SARS-CoV-2 is an urgent clinical and public health priority. While these kinds of Bayesian search algorithms are used widely in other settings (eg, to find downed aircraft, in submarine recovery, and to aid in oil exploration), this is the first time that Bayesian adaptive approaches have been used for active disease surveillance in the field.

Objective: This study's objective was to evaluate a Bayesian search algorithm to target hotspots of SARS-CoV-2 transmission in the community with the goal of detecting the most cases over time across multiple locations in Columbus, Ohio, from August to October 2021.

Methods: The algorithm used to direct pop-up SARS-CoV-2 testing for this project is based on Thompson sampling, in which the aim is to maximize the average number of new cases of SARS-CoV-2 diagnosed among a set of testing locations based on sampling from prior probability distributions for each testing site. An academic-governmental partnership between Yale University, The Ohio State University, Wake Forest University, the Ohio Department of Health, the Ohio National Guard, and the Columbus Metropolitan Libraries conducted a study of bandit algorithms to maximize the detection of new cases of SARS-CoV-2 in this Ohio city in 2021. The initiative established pop-up COVID-19 testing sites at 13 Columbus locations, including library branches, recreational and community centers, movie theaters, homeless shelters, family services centers, and community event sites. Our team conducted between 0 and 56 tests at the 16 testing events, with an overall average of 25.3 tests conducted per event and a moving average that increased over time. Small incentives—including gift cards and take-home rapid antigen tests—were offered to those who approached the pop-up sites to encourage their participation.

Results: Over time, as expected, the Bayesian search algorithm directed testing efforts to locations with higher yields of new diagnoses. Surprisingly, the use of the algorithm also maximized the identification of cases among minority residents of underserved communities, particularly African Americans, with the pool of participants overrepresenting these people relative to the demographic profile of the local zip code in which testing sites were located.

Conclusions: This study demonstrated that a pop-up testing strategy using a bandit algorithm can be feasibly deployed in an urban setting during a pandemic. It is the first real-world use of these kinds of algorithms for disease surveillance and represents a key step in evaluating the effectiveness of their use in maximizing the detection of undiagnosed cases of SARS-CoV-2 and other infections, such as HIV.

(JMIR Public Health Surveill 2023;9:e39754) doi:10.2196/39754

KEYWORDS

active surveillance; bandit algorithms; infectious disease; community health; reinforcement learning; COVID-19; SARS-CoV-2

Introduction

Background

Columbus, Ohio—like much of the world during the first 3 years of the pandemic—experienced multiple waves of infections in which the number of SARS-CoV-2 infections overwhelmed and outstripped the capacity of efforts to detect new cases [1,2]. In fact, estimates suggest that even at the height of the Omicron surge, a significant number of infections went undetected by case-based surveillance systems and the rise of at-home testing [3]. As the epidemic continues, particularly as public testing sites are scaled back and closed, the ability to quickly identify individuals with SARS-CoV-2 and link them to treatment has become difficult, risking patient and population health [4,5]. Given the integral role of active surveillance efforts in identifying cases, linking patients to care, and managing COVID-19, local public health departments require ways to maximize the resources they have at their disposal to continue to find new cases of disease in their jurisdictions even as resources for these efforts continue to shrink [4,6,7].

How to optimize resource allocation over time is a well-studied problem in sequential decision-making and reinforcement learning. The introduction of a spatial component to these kinds of dilemmas has been applied in a variety of settings, from military search and rescue to oil exploration [8]. We have previously described the use of one set of tools, bandit algorithms, to address these kinds of problems for detection of HIV and SARS-CoV-2 in the community [8-10]. Up until now, these methods have only been evaluated in computer simulations. The study described here represents the first-ever real-world implementation that investigates these tools for active infectious disease surveillance.

An academic-governmental partnership—the Flexible Adaptive Algorithmic Surveillance Testing (FAAST) program—between Yale University, Ohio State University (OSU), Wake Forest University (WFU), the Ohio Department of Health (ODH), the Ohio National Guard (ONG), and the Columbus Metropolitan Libraries (CML) conducted a study of bandit algorithms to maximize the detection of new cases in SARS-CoV-2 in Columbus in 2021.

This initiative began as the Delta variant of SARS-CoV-2 established itself in Ohio in the summer of 2021, ending a period of decreasing case numbers in the state. From mid-August to mid-September, the daily case count rose from over 2000 to over 9000 [11]. Around the same time, the ODH started distributing home rapid antigen tests (RATs) for COVID-19 to allow people to monitor their own health status. These tests were provided to community-based organizations in Ohio, including over 246 library locations, which received more than 53,000 rapid tests in late summer 2021 [12]. While the high demand meant that these libraries quickly ran out of tests to distribute, the community's relationship and trust in libraries as a testing resource for COVID-19 enabled us to partner with the CML for this study [13]. The strong uptake of RATs is consistent with evidence that voluntary use of these tests is high in settings where they have been widely distributed [14], but it is important to note that decisions to opt in to population-level voluntary testing programs appear to be influenced by factors such as socioeconomic status rather than infection risk alone [15]. Thus, active surveillance has an important role in both prioritizing limited testing resources and in reaching individuals not seeking voluntary testing.

Working with our partners, we were able to offer an option to maximize the effectiveness of the scarce testing resources in Columbus by feeding our daily testing information into the bandit algorithm to target the next day's testing efforts. Our "learn by doing" method adaptively targeted locations in the CML system, with tests supplied by ODH and performed by ONG with student and faculty support from OSU, WFU, and Yale.

As the COVID-19 pandemic continues, we will further evaluate the tools in Columbus and the state of Ohio and elsewhere as opportunities arise. Given the flexibility of these methods, we can tailor them to identify new cases more efficiently in underserved communities, which may be at higher risk of transmission and serious clinical disease either due to undervaccination and the prevalence of underlying conditions or due to specific kinds of workplaces and high-risk settings (eg, skilled nursing facilities). Importantly, the potential utility of these tools is not limited to SARS-CoV-2. Evaluation of these tools for HIV, hepatitis C virus, and other sexually transmitted diseases is also being considered.

Goals of the Study

The purpose of this study was to evaluate the use of bandit algorithms to maximize the yield of testing for SARS-CoV-2 across multiple community sites in Columbus, Ohio, over a 3-month period from August to October 2021.

We sought to understand (1) if an algorithm-guided site selection would be operationally feasible for the branches of the CML system and ONG, since locations could shift from day to day, requiring staff to be ready and prepared for quick deployment and setup of testing efforts; (2) if shifting locales would present difficulties for uptake of testing services by the community, given the fact that the announcement of testing locations would only occur a few days in advance; and (3) how algorithm-guided site selection impacted the recruitment of minority residents of underserved communities in terms of number of tests conducted at each event among key demographic groups, particularly African-Americans, in Columbus.

In general, we were interested in a better understanding of the operational performance of the algorithm—as a formal analysis of the effectiveness of this algorithm-guided site selection approach to active disease surveillance would require a large cluster-randomized trial. Specifically, we chose 2 a priori indicators that the algorithm was homing in on hotspots for detecting new cases. Our goal was to sustain a testing positivity rate that was greater than (1) the estimated prevalence of the surrounding community and (2) the positivity rates of other types of testing events in the same zip codes. It is important to clarify that we were not attempting to estimate local prevalence of disease or positivity rates in this study nor making inferences about the effectiveness of the algorithm in practice but using these 2 metrics as a guide as to its basic performance as we assessed its initial feasibility in this research. The algorithm is a sequential decision-making tool; in this case, it was designed to guide resource allocation decisions for SARS-CoV-2 testing efforts.

Methods

Statistical Approach

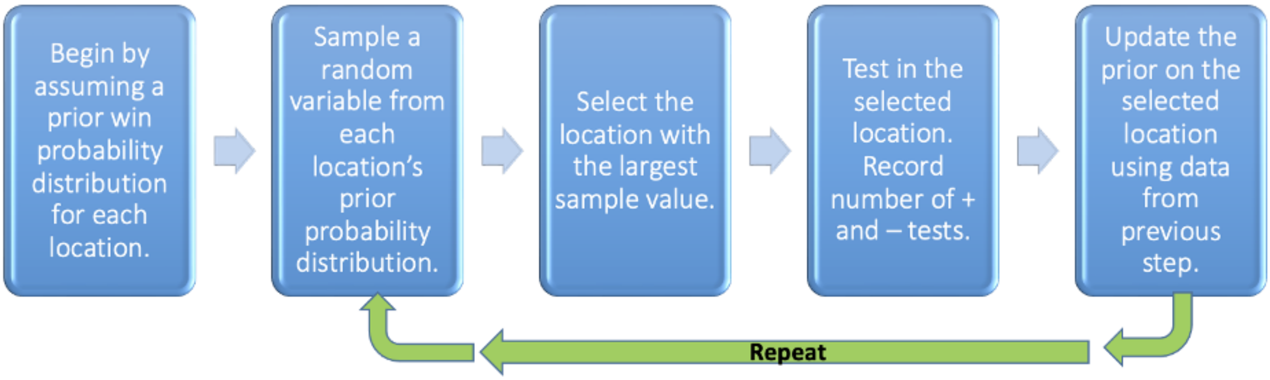
The algorithm used to direct pop-up SARS-CoV-2 testing for this project has been described in detail elsewhere [8-10]. The algorithm is based on Thompson sampling, which uses a Bayesian updating process involving iteratively sampling from prior probability distributions of all potential testing sites—the set of all locations at which testing is being considered—to home in on those with the highest probability over the long run in finding new cases of SARS-CoV-2 [16,17]. Multimedia Appendix 1 describes Thompson samplng in more detail. This approach is not being used to estimate the local prevalence or test positivity at each site in the community but to maximize success in finding new cases over time. This algorithm is not a sampling strategy in which we seek to learn about an underlying population parameter (eg, prevalence) but simply a process by which we can direct testing efforts most efficiently. In fact, the reason this algorithm is not designed to estimate local prevalence is because this metric, while related to testing yield, may be confounded by other factors at work in a given location. That is, while testing yield is indeed a function of prevalence in part, prevalence does not provide a one-to-one proxy for the number of tests obtainable at a given location. For instance, a given site may have a lower prevalence than another but be a place in which more new diagnoses are garnered for other reasons (eg, higher health-seeking behavior, fewer options for testing in other settings); thus, this algorithm is designed to maximize testing yields alone.

Testing sites are determined a priori and serve as a fixed list of potential locations for deploying testing. Initial prior probability distributions for each site can be assigned to be noninformative (ie, uniform), reflecting a lack of information about where yield might be highest, or informative based on local knowledge of the pandemic, previous testing efforts, or social and economic characteristics of a given neighborhood. While informative priors can aid convergence on hotspots if the information is accurate, they can delay convergence if the information is incorrect, so one should use caution when specifying the initial prior distributions.

For this study, we chose to specify informative prior distributions for each site based primarily on results from previous feasibility testing events that we conducted at each site prior to use of the algorithm, with slight modifications to account for the current pandemic conditions. To estimate the baseline values for the parameters for our prior distributions for each potential testing site, we consulted with experts on the local epidemiological conditions and considered the prior testing event results, the reported rate of SARS-CoV-2 infections, and the vaccination rate in the neighborhoods surrounding each site. Taken together, this process yielded assumed β distributions for the positivity rate at each site where the α parameter represented the assumed number of positive tests and the β parameter the assumed number of negative tests. The assumed positivity was $\alpha / \alpha + \beta$ and the total number of tests (ie, the sum of the α and β parameters). Details of the derived values for each site are shown in the Results section. While the true values of these parameters are unknown, we decided that given the short duration of our study period and the quality of the empirical and expert knowledge on these testing sites, estimating the values for these initial α and β parameters for each location was a reasonable choice.

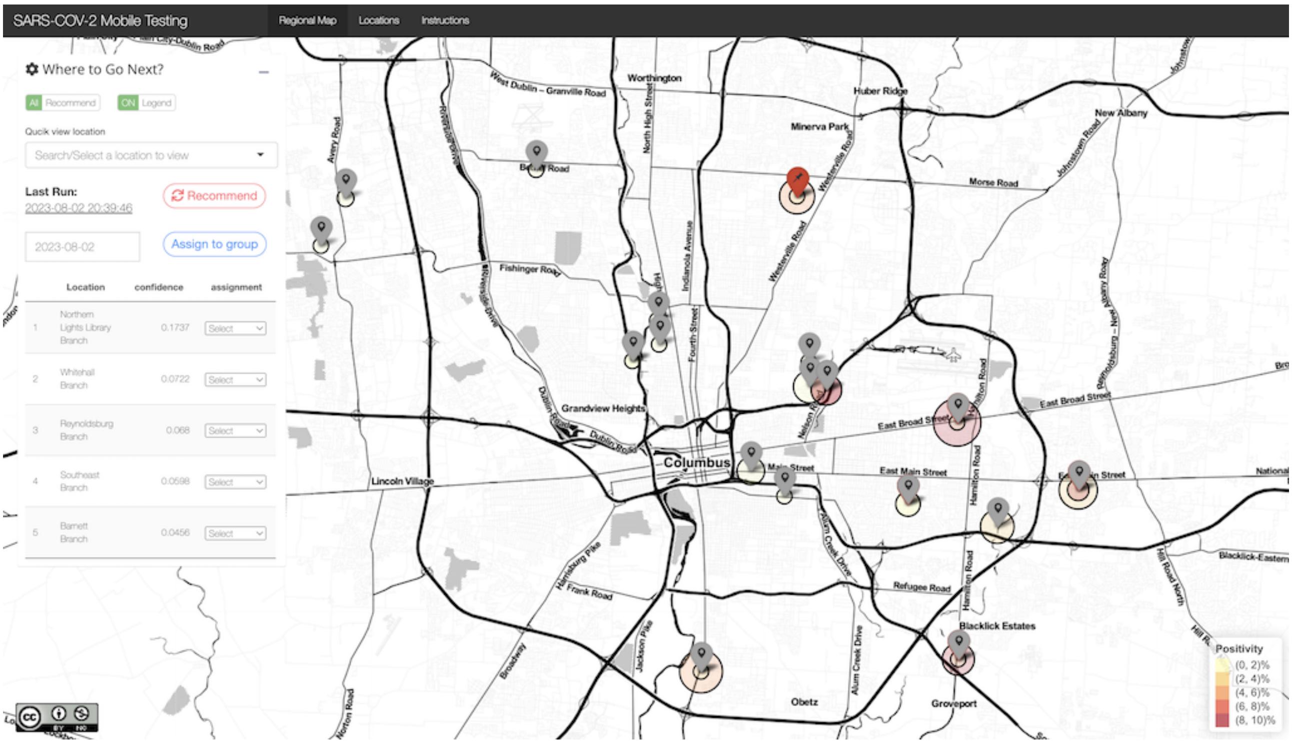
Before assigning a location for a given day's testing efforts, the algorithm randomly samples from the probability distributions of all the potential testing sites. Then, the site which has the largest realized value from these random draws is selected for the day's testing deployment [18]. The data on the number of tests at that day's site and the number of new diagnoses among them is then used to update the prior probability distribution for the site. The resulting posterior distribution after that day's testing then becomes the updated prior distribution for the next day's testing deployment. Before the next outing, a random draw from all the sites' updated probability distributions happens again. This process is then repeated so the algorithm updates with each subsequent testing deployment. The entire procedure is described in Figure 1. The intuition here is that over time, using the knowledge accrued by each outing's successes or failures (positive and negative tests), the algorithm learns the expected yields at each site and refines the probability distributions used for the random draws. Though all sites will have a nonzero chance of being chosen, the Bayesian updating processes here will be drawn toward sites where successes are occurring in previous rounds of testing. What is important is that other sites can still be chosen by chance and offer the opportunity to keep gathering information on the entire landscape of testing sites, but through the algorithm, we will prioritize the ones with the current greatest yields. If a low-priority site gets picked by random draw and turns out to have a high yield of positive tests, that site's probability distribution now will shift its likelihood of being chosen in subsequent rounds.

Figure 1. Steps in the implementation of the Thompson sampling algorithm.



Based on this algorithm, a web app was created for this project [19]. The underlying code is also available [20]. The app requires data input from each location where a testing event took place. Based on the information received, the tool then suggests testing locations likely to yield the most undetected positive cases per test for the next event. Views of the app's regional and location pages are presented in Figures 2 and 3, respectively. Figure 2 shows a map of Columbus with pinned testing locations. After each testing event, data from the outing were entered into the algorithm in Figure 3 (eg, date and time, the total number of tests conducted, and number of positives). The app then integrates this data into the bandit algorithm, which delivers a new testing location for the next outing. Given conjugacy of the underlying Bayesian model, the computational cost of this process is minimal, as it only requires the ability to generate random draws from a beta distribution.

Figure 2. Regional map page of the web app for targeting SARS-CoV-2 testing with mobile units.



Where to Go Next?

Qucik view location

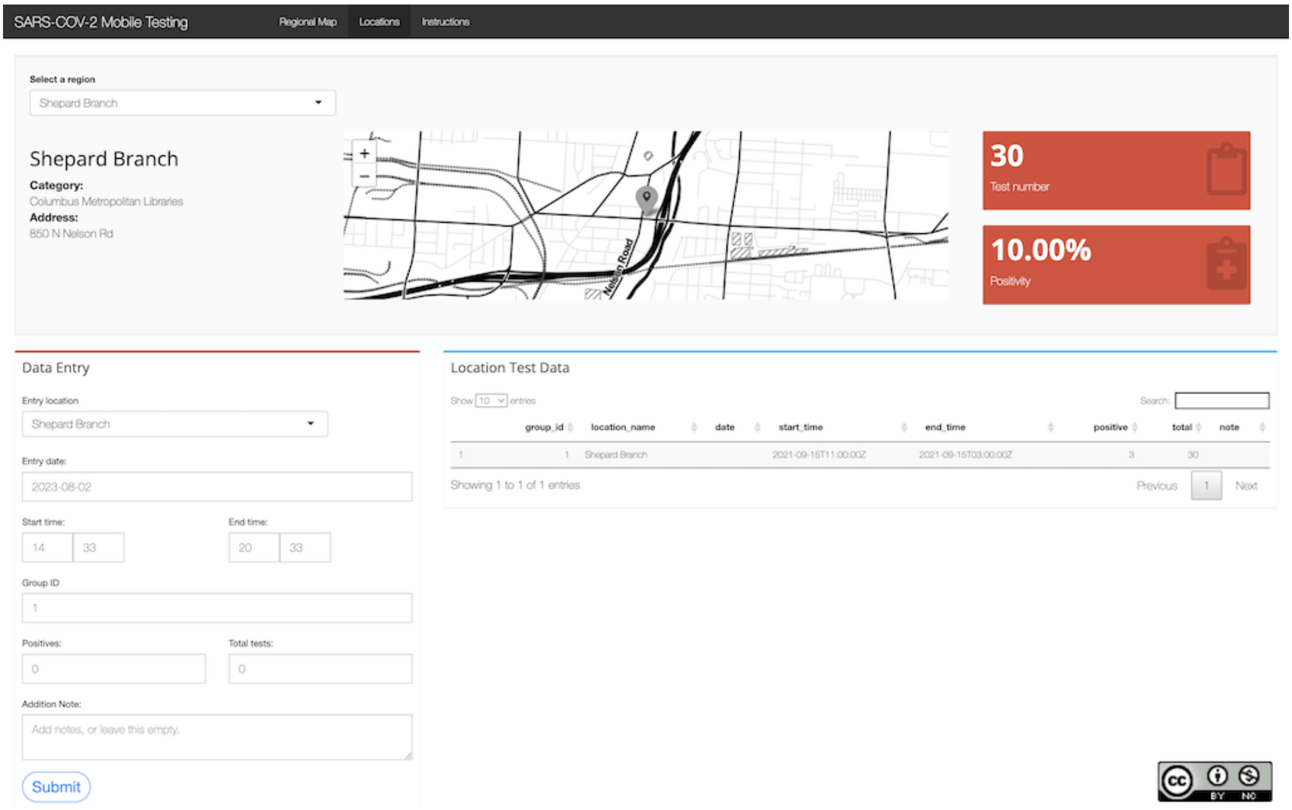
Search/Select a location to view

Last Run:
2023-08-02 20:39:46

2023-08-02

	Location	confidence	assignment
1	Northern Lights Library Branch	0.1737	Select
2	Whitehall Branch	0.0722	Select
3	Reynoldsburg Branch	0.068	Select
4	Southeast Branch	0.0598	Select
5	Barnett Branch	0.0456	Select

Figure 3. Locations page of the web app for targeting SARS-CoV-2 testing with mobile units.



Select a region: Shepard Branch

Shepard Branch

Category: Columbus Metropolitan Libraries

Address: 850 N Nelson Rd

30 Test number

10.00% Positivity

Data Entry

Entry location: Shepard Branch

Entry date: 2023-08-02

Start time: 14 33

End time: 20 33

Group ID: 1

Positives: 0

Total tests: 0

Addition Note: Add notes, or leave this empty.

Location Test Data

Show 10 entries

Search:

	group_id	location_name	date	start_time	end_time	positive	total	note
1	1	Shepard Branch		2021-09-15T11:00:00Z	2021-09-15T03:00:00Z	3	30	

Showing 1 to 1 of 1 entries

Previous 1 Next

Program Design and Implementation Overview

The logic model for program design and implementation is shown in Textbox 1. The logic model depicts the program events and expected outcomes through the integration of resources, the execution of activities, and the participation of communities and community organizations, resulting in measurable near-term and long-term outcomes. The inputs include program partners and the resources they provided to the project. The activities represent the main programmatic tasks accomplished with the participation of those tested in underserved communities in Columbus and with the assistance of local community-based organizations. The achievements of the program are denoted in short-term, measurable outcomes and benefits to participants and the larger community, as well as long-term, large-scale outcomes and benefits for the city and beyond.

The project team established partnerships with ODH, ONG, the CML system, and the Columbus Department of Neighborhoods (CDN). ODH supported the project by providing BinaxNOW RATs (Abbott), while ONG became a clinical partner who deployed their clinical staff to the sites to conduct testing. The neighborhood liaisons from CDN consulted the team regarding the initial location selection and promoted testing events in their designated areas of oversight. Finally, CML provided 7 of their branch locations as sites for pop-up testing and promoted the testing events through the CML network.

The study lasted 3 months, from August to October 2021. The duration of each testing event was about 4 hours each day. We conducted a total of 16 testing events at 13 zip codes. Figure 4 shows dates of testing events, tests conducted, number of positive SARS-CoV-2 diagnoses, and the site positivity rate. Multimedia Appendix 2, Table S1 contains information about the sites and the zip codes where they were located.

Textbox 1. Logic model for the program and its implementation.

Inputs

- Resources
  - Research and statistical design
  - Funding
  - Staffing
  - Mobile van
  - Logistics
  - Testing sites
  - Rapid tests
  - Personal protective equipment
  - Clinical support
- Partners
  - Ohio State University
  - Yale University
  - Wake Forest University
  - Columbus Metropolitan Libraries
  - Ohio Department of Health
  - Ohio National Guard

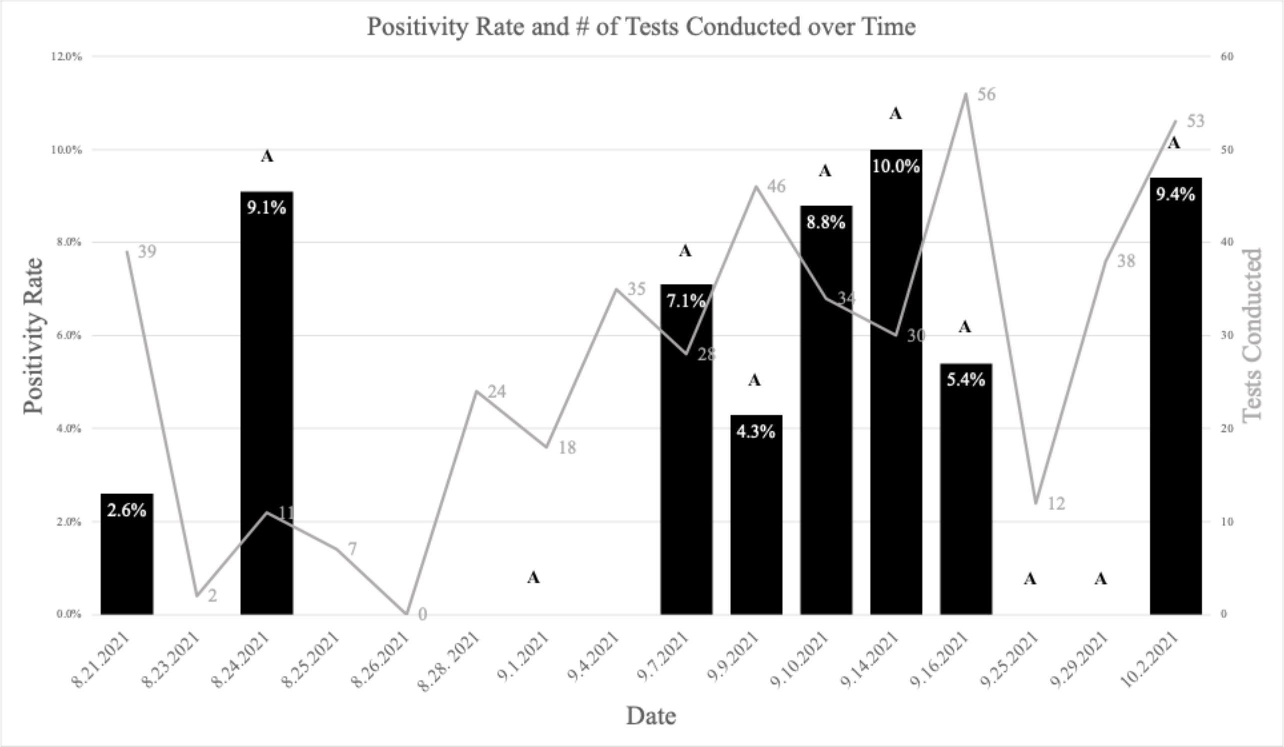
Processes

- Activities
  - Build the application
  - Conduct pop-up COVID-19 testing events
  - Distribute COVID-19 home test kits
  - Disseminate information about COVID-19
- Participants
- Target population: underserved communities in Columbus
- Community organizations (eg libraries, parks, and recreation centers)

Outcomes

- Near-term
  - Consistent detection of undiagnosed cases with the positivity rate being close to the state’s rate
  - Increased community access to COVID-19 rapid home tests
  - Increased community awareness about COVID-19 testing opportunities
- Long-term
  - Convenient and effective testing in low-opportunity areas in Columbus, Ohio
  - Prevention of transmission of COVID-19 in Columbus
  - Use of bandit approach in other infectious disease surveillance efforts

Figure 4. Positivity rate (black bars) and tests conducted (grey lines) for each testing session. The letter A is used to indicate testing at a library location. Starting on September 14, 2021, all testing locations were chosen by the algorithm.



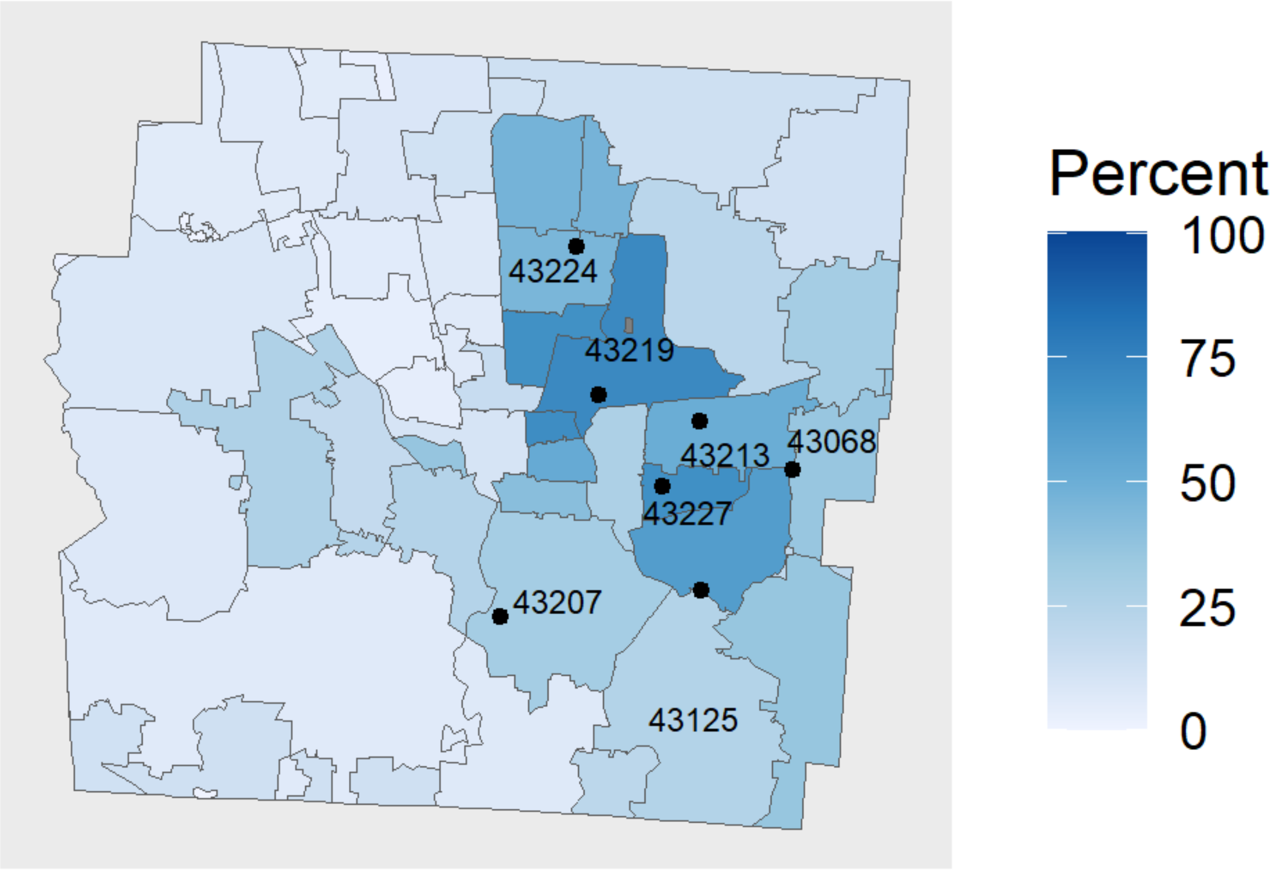
The initial choice of testing sites was derived from a set of highly trafficked candidate locations based on raw cellphone data from the UberMedia COVID-19 recovery data set spatially joined with Loveland Landgrid parcel data to produce indices of contacts and unique contacts per parcel ranked by volume [21,22]. From this list, 30 of the most highly trafficked locations from the data set were selected for further analysis. A team of volunteers from OSU then visited each location to assess the viability of each site in terms of the availability of the venue (whether it was private or public, its accessibility, its hours of operation, and the availability of onsite parking). The final step of the selection process was a discussion of each site with the neighborhood liaisons from the CDN, who shared their insights about the proposed locations and the potential for attracting participants to come forward for testing at each site.

Each testing event required the deployment of 2 volunteers from OSU and 2 ONG staff members. At the start of the study, events took place 2 to 3 times a week at community centers, movie theaters, parks, community events, shelters for the homeless, and libraries. These early weeks were dedicated to exploring the viability of the sites, constructing prior probability distributions for the algorithm for each site, and establishing an efficient operational workflow. In the last 2 weeks of the study, we deployed the bandit algorithm to choose sites for testing, restricting our sites to the branches of the CML, which turned out to be the most suitable sites for SARS-CoV-2 testing in the previous weeks. In these last weeks, the algorithm chose to send the team to sites 2, 4, 5, and 7, in that order, from September 16 to October 2, 2021 (site attributes can be found in Table 1). CML became an active partner in these efforts, promoting the testing events throughout their system with posters announcing the upcoming testing events distributed among the libraries in English, French, and Spanish. Figure 5 shows the location of the CML branches and underlying demographic information about these zip codes.

Table 1. Parameters for the assumed prior probability distributions for each testing site.

Site	Positivity	Tests, n	α	β
1	0.100	60	6	54
2	0.055	55	3.025	51.975
3	0.080	20	1.6	18.4
4	0.095	40	3.8	36.2
5	0.100	60	6	54
6	0.090	40	3.6	36.4
7	0.090	30	2.7	27.3

Figure 5. Location and zip codes of final testing sites with percentage of Black or African American residents.



Testing efforts were focused on reaching underserved communities through strategic selection of potential sites; however, testing was open to everyone. To invite undocumented immigrants to be tested as well, participants were not required to present identification documents. Results of the tests were entered in SimpleReport—a free tool built by the US Centers for Disease Control and Prevention that makes it easy for COVID-19 testing sites to record results for rapid point-of-care tests and quickly report required data to public health departments [23].

Ethical Considerations

The Office of Responsible Research Practices and the institutional review board at the Ohio State University deemed this project (2021E0496) exempt and not subject to regulations requiring institutional review board review and approval.

Results

We set a baseline goal of testing approximately 20 people at each event through consultation with other institutions that have deployed mobile pop-up testing sites (eg, New York City Health and Hospitals Corporation, University of Cincinnati Medical Center). Of 16 testing events conducted, 10 testing events met the baseline goal of testing over 20 participants. The total number of people tested over the course of the study was 405 (the mean number of tests conducted at each event was 25.3), producing 20 positive tests. This gave us an average positivity rate of 4.9% across all sites, with the highest being 9.5%. The positivity rates observed at our testing sites were comparable to or greater than the officially documented positivity rates in those same areas in central Ohio on the days of our testing events [24-26]. In particular, the proportion of symptomatic positive tests in our study exceeded those obtained from other providers in the state, which prioritized symptomatic individuals for diagnosis [27]. The rolling 2-week average positivity rate for all tests conducted in our study ranged from 2.9% (3/105) to 6.5% (9/138); for asymptomatic individuals, positivity ranged from 0% to 5% (3/57); and for symptomatic individuals, the range was from 0% to 21.4% (3/14). In each week of testing, and in each rolling 2-week period, the positivity of symptomatic testing was greater than the daily rate range of the surrounding county. In the final 2 weeks, when the bandit algorithm was used to select testing event sites, the average positivity rate for symptomatic testing was 17.4 (4/23)%, which nearly doubled the highest recorded daily positivity rate in Franklin county at that time (Table 1) [24-26]. Details are shown in Table 2.

Underrepresented communities were reached at testing events at rates higher than or near their proportion of the population in the zip codes we tested in. The percentage of Black or African American people tested was larger than the corresponding percentage of African Americans in 3 of 5 zip codes in which our testing sites were located [28]. Details are shown in Table 3.

Table 2. Study positivity rates compared to tests conducted by other providers in Franklin County, Ohio.

	Positivity for rolling 2-week periods in 2021							
	Aug 12-25	Aug 19-Sept 1	Aug 26-Sept 8	Sept 2-15	Sept 9-22	Sept 16-29	Sept 23-Oct[a]	Total
Flexible Adaptive Algorithmic Surveillance Testing clinic , n/N (%)								
All tests	3/65 (5)	5/96 (5.2)	4/99 (4)	8/146 (5.5)	9/138 (6.5)	3/105 (2.9)	6/106 (5.7)	22/408 (5.4)
Asymptomatic individuals	3/57 (5)	4/82 (4.9)	1/85 (1.2)	3/116 (2.6)	3/105 (2.7)	0/78 (0)	2/83 (2.4)	9/330 (2.7)
Symptomatic individuals	0/8 (0)	1/14 (7.1)	3/14 (21.4)	5/30 (16.7)	6/33 (18.2)	3/27 (11.1)	4/23 (17.4)	13/78 (16.7)
Franklin County daily rate[b] (%), range	5.5-7.2	5.5-9	7-9.9	9.5-10.2	9.5-10.4	9.3-10.4	8.9-10	5.5-10.4

[a] Time period when bandit algorithm was used to select testing event sites.

[b] Absolute numbers were not available for county daily rate.

Table 3. Study representation of Black or African American people by zip code.

	Zip code, n/N (%)				
	43227	43125	43207	43068	43213
Black or African American participants in Flexible Adaptive Algorithmic Surveillance Testing clinics	18/24 (75)	17/31 (55)	8/31 (26)	13/29 (45)	13/28 (46)
Black or African American population of Franklin county (US Census; 2021 American Community Survey 5-year data)	16,740/24,779 (67.6)	3429/13,756 (24.9)	13,936/46,308 (30.3)	20,658/58,426 (35.4)	18,205/35,954 (50.6)

Throughout the study, 55% (156/284) of the population tested identified themselves as African American, and 36% (103/284) self-identified as White. Almost half of the population tested (105/227, 46%) stated their level of education was high school or below, and 47% (106/225) stated they worked full-time. Data about vaccination status against COVID-19 was collected as well: 56% (126/224) were vaccinated, compared to the vaccination rate in Franklin County of 54% [29].

Our survey results indicate that 27% (41/153) of those who attended our testing events found out about the testing event from others, 22% (34/153) by communicating with the library, and 29% (45/153) by walking or driving by.

Discussion

Principal Findings

This study demonstrated that a pop-up testing strategy using a bandit algorithm can be feasibly deployed in an urban setting during a pandemic. Although this was a limited roll-out of the strategy, positivity rates in this feasibility study were comparable or greater to those obtained by other testing initiatives primarily targeted at symptomatic individuals at the same time in these same areas of central Ohio on the days of our testing events, even though the bulk of those tested at our sites were asymptomatic, providing encouraging, though only preliminary, evidence that this bandit algorithm may be useful in improving case detection efforts [24-26]. It is the first real-world use of these kinds of algorithms for disease surveillance and represents the first step in evaluating the effectiveness of their use in maximizing the detection of undiagnosed cases of SARS-CoV-2 and other infections, such as HIV. In addition, the study showed that a simple, scaled-down testing approach with a total budget of under US $10,000 could carry out a program that could test up to 60 people in a single 4-hour event during its 3-month run. This shows that a flexible, adaptive approach to SARS-CoV-2 surveillance can be run efficiently with support for basic supplies (eg, tents, chairs, tables), and partnerships with organizations who could supply the tests (ODH) and carry them out (ONG) [30].

The program's ability to reach underserved communities is especially notable for several reasons. The quick-pivot nature of the program required that we give communities very little notice of upcoming events, potentially taxing their ability to hear about and be able to travel to locations to participate. Accessibility and transportation are barriers to health services in some of these communities [31,32]. In addition, there have been nationwide and state-wide trends of COVID-19 resources being allocated inequitably toward wealthier, predominantly White communities [33]. This is particularly troubling as both general determinants of health as well as the specific impact of COVID-19 make communities of color and other underserved communities more vulnerable to COVID-19 infection and the morbidity and mortality associated with the disease [34,35].

Libraries are important resources for communities in the context of population health [36,37]. A novel discovery during the study was the importance of the CML system in our active disease surveillance efforts. We learned that the CML is known and trusted by the communities of the city as a source of information and a site for social services. Subsequent studies in Columbus will now be done in partnership with the CML, as it excels at community outreach, has a network of facilities strategically placed in targeted neighborhoods, communicates well within its network, and is eager to work with public health partners. Our testing partner, ONG, was suitable for this study, but because of other commitments, we could not deploy the algorithm on less than 3 days' notice with them, and we are seeking out partners (a local federally-qualified health center) for the next phase of this work that have greater flexibility and can be available on short notice. While a major strength of our approach is its ability to deploy testing teams to locations based on emerging data in real time, it requires an operational nimbleness from partners, which may not be possible for some.

This study uses a novel approach to implementation science in that it addresses feasibility and acceptability of a new strategy for active surveillance of infectious diseases as a precursor to additional evaluation of the effectiveness of these algorithms in the field. Randomized controlled trials (RCTs) are expensive to mount and time consuming for both researchers and participants—this kind of preliminary research on implementation challenges is critical to understand the contextual factors that might present difficulties for the conduct of RCTs or real-world use of the intervention under study [38].

Limitations

This study has a number of limitations. First, it was conducted over a comparatively short period of time at a specific point of increasing concern in the pandemic. As such, the engagement of the community and the effectiveness of the algorithm must be set within that context. In addition, this study was only conducted in one midsized Midwestern city and may not be generalizable to other settings.

Conclusions

Through the implementation of a bandit algorithm in this study, we demonstrated the feasibility of such approaches to guide community testing for SARS-CoV-2. In addition, we established an efficient workflow and operational plan that can be extended to other organizations conducting mobile testing. In particular, we were pleased that the program appears to be well-suited to reach historically underserved communities in Columbus. As the COVID-19 pandemic persists in the United States, using our approach not only for active surveillance but as a core component of "test-to-treat" or "test-to-vaccinate" efforts targeted at the communities most at risk of the disease should be considered as well [39,40]. Finally, this algorithmic approach is pathogen-agnostic—it can be used for other infectious disease efforts. It was initially developed for maximizing the yield of undiagnosed HIV infections but could be used similarly for other infections to maximize case detection in the community (eg, for sexually transmitted diseases) or as an adjunct to other efforts, beginning with case identification (eg, test-to-treat for SARS-COV-2, HIV, hepatitis C virus, and sexually transmitted infections). The important next step is to evaluate the effectiveness of this algorithm in its ability to maximize the number of cases of disease detected in comparison to standard operating procedures for such deployment by public health practitioners. As an interim step, we have established a new partnership with a public health department in one of the largest US cities and will be evaluating the performance of Thompson sampling using historical data provided by it to compare how yields of testing using this algorithm would have differed from the actual number of positive tests collected for SARS-CoV-2 over the past 3 years with their fleet of mobile health vans. Depending on the outcome of this analysis, we will pursue funding for a prospective study of our approach in the context of a cluster randomized trial.

Acknowledgments

We thank Mark Hamlin and Philip Ennen (Ohio Department of Health) for providing rapid COVID-19 tests to support our operations; Captain Rodshana Ulery (Ohio National Guard) for providing clinical staff to administer COVID-19 testing; and Shaunessy Everett (Columbus Metropolitan Library) for support in organization of testing events at the library locations. We thank the College of Public Health at Ohio State University and the Yale School of Public Health for funding the pilot and Ohio State University students—volunteers Amanda Bleichrodt, Sophia Padgett, Aaron Cochran, John Hansen, Mitali Dalwalla, Lauren Putz, Manar Taleb, Maria Krantz, Ryan Borne, Lily Liu, Samantha Malone, and Jack Long—for their help with executing testing events.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Thompson sampling algorithm.
[DOCX File , 13 KB - publichealth_v9i1e39754_app1.docx]

Multimedia Appendix 2

Table S1. Location data from the conducted testing events.
[DOCX File , 17 KB - publichealth_v9i1e39754_app2.docx]

References

1. The COVID-19 testing debacle. Nat Biotechnol 2020 Jun;38(6):653. [doi: 10.1038/s41587-020-0575-3] [Medline: 32493981]
2. Lim D, Ollstein A. Next testing debacle: The fall virus surge. Politico. 2020 Jun 17. URL: https://www.politico.com/news/2020/06/17/coronavirus-testing-fall-surge-327017 [accessed 2022-06-20]
3. Qasmieh SA, Robertson MM, Rane MS, Shen Y, Zimba R, Picchio CA, et al. The importance of incorporating at-home testing into SARS-CoV-2 point prevalence estimates: findings from aa US national cohort, February 2022. JMIR Public Health Surveill 2022 Dec 27;8(12):e38196 [FREE Full text] [doi: 10.2196/38196] [Medline: 36240020]
4. Bosman J, Kasakove S. States close mass test and vaccine sites, but virus may swell anew. New York Times. 2022 Mar 30. URL: https://www.nytimes.com/2022/03/30/us/covid-vaccine-testing-states.html [accessed 2023-07-05]
5. Gandhi RT, Malani PN, Del Rio C. COVID-19 therapeutics for nonhospitalized patients. JAMA 2022 Feb 15;327(7):617-618. [doi: 10.1001/jama.2022.0335] [Medline: 35029659]
6. Tolbert J, Artiga S. Implications of the lapse in federal Covid-19 funding on access to Covid-19 testing, treatment, and vaccines. KFF. 2022 Mar 28. URL: https://www.kff.org/coronavirus-covid-19/issue-brief/implications-of-the-lapse-in-federal-covid-19-funding-on-access-to-covid-19-testing-treatment-and-vaccines/ [accessed 2023-07-05]
7. Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J Infect Public Health 2020 Nov;13(11):1630-1638 [FREE Full text] [doi: 10.1016/j.jiph.2020.07.019] [Medline: 32855090]
8. Gonsalves GS, Copple JT, Paltiel AD, Fenichel EP, Bayham J, Abraham M, et al. Maximizing the efficiency of active case finding for SARS-CoV-2 using bandit algorithms. Med Decis Making 2021 Nov;41(8):970-977 [FREE Full text] [doi: 10.1177/0272989X211021603] [Medline: 34120510]
9. Gonsalves G, Copple J, Johnson T, Paltiel A, Warren J. Bayesian adaptive algorithms for locating HIV mobile testing services. BMC Med 2018 Sep 03;16(1):155 [FREE Full text] [doi: 10.1186/s12916-018-1129-0] [Medline: 30173667]
10. Gonsalves GS, Crawford FW, Cleary PD, Kaplan EH, Paltiel AD. An adaptive approach to locating mobile HIV testing services. Med Decis Making 2018 Feb;38(2):262-272 [FREE Full text] [doi: 10.1177/0272989X17716431] [Medline: 28699382]
11. Washington J, Drown H. COVID-19 summer surge may have hit highest point, but will numbers continue to drop in the fall? Advance Local. 2021 Jun 21. URL: https://www.cleveland.com/coronavirus/2021/09/covid-19-summer-surge-may-have-hit-highest-point-but-will-numbers-continue-to-drop-in-the-fall.html [accessed 2022-06-20]
12. Free rapid at-home COVID-19 tests available for Ohioans. Ohio Department of Health. 2021 Sep 7. URL: https://odh.ohio.gov/media-center/odh-news-releases/odh-news-release-09-07-21 [accessed 2022-06-20]
13. Walker M, Gordon K, Henry M. Greater Columbus libraries run out of COVID-19 test kits as demand increases. Columbus Dispatch. 2021 Sep 11. URL: https://www.dispatch.com/story/news/2021/09/11/covid-19-test-kits-columbus-libraries-demand-increases/8273918002/ [accessed 2022-06-20]
14. Kwan TH, Wong NS, Chan CP, Yeoh EK, Wong SY, Lee SS. Mass screening of SARS-CoV-2 with rapid antigen tests in a receding Omicron wave: population-based survey for epidemiologic evaluation. JMIR Public Health Surveill 2022 Nov 09;8(11):e40175 [FREE Full text] [doi: 10.2196/40175] [Medline: 36240027]
15. Stillman S, Tonin M. Communities and testing for COVID-19. Eur J Health Econ 2022 Jun;23(4):617-625 [FREE Full text] [doi: 10.1007/s10198-021-01385-y] [Medline: 35169950]
16. Scott SL. Multi-armed bandit experiments in the online service economy. Appl Stochastic Models Bus Ind 2015 Jan 16;31(1):37-45. [doi: 10.1002/asmb.2104]
17. Thompson WR. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika 1933 Dec 01;25(3-4):285-294. [doi: 10.1093/biomet/25.3-4.285]
18. Gelman A. Prior distribution. In: El-Shaarawi AH, Piegorsch WW, editors. Encyclopedia of Environmetrics, Vol 3. Chichester, UK: John Wiley & Sons; 2002:1634-1637.
19. Bandit Demo. ShinyApps. URL: https://netzissou.shinyapps.io/BanditDemo/ [accessed 2023-07-05]
20. NetZissou/Bandit. GitHub. URL: https://github.com/NetZissou/Bandit [accessed 2023-07-11]
21. COVID-19 recovery insights. UberMedia. URL: https://COVID19.ubermedia.com/COVID-19-recovery-insights/ [accessed 2021-10-02]
22. Landgrid. Loveland Technologies. URL: https://landgrid.com/ [accessed 2021-10-02]
23. SimpleReport: A better way to report COVID-19 tests. US Centers for Disease Control and Prevention. URL: https://www.simplereport.gov/ [accessed 2021-10-02]
24. Franklin County, OH. Act Now Coalition. URL: https://covidactnow.org/us/ohio-oh/county/franklin_county/?s=45423293 [accessed 2023-07-05]
25. Tracking coronavirus in Franklin County, Ohio: latest map and count. New York Times. URL: https://www.nytimes.com/interactive/2021/us/franklin-ohio-covid-cases.html [accessed 2023-07-05]

26. Archived Statewide and County Case Rates per 100,000 Residents | COVID-19. Ohio Department of Health. URL: https://coronavirus.ohio.gov/dashboards/overview/statewide-and-county-case-rates [accessed 2021-06-20]
27. Turner AN, Kline D, Norris A, Phillips WG, Root E, Wakefield J, et al. Prevalence of current and past COVID-19 in Ohio adults. Ann Epidemiol 2022 Mar;67:50-60 [FREE Full text] [doi: 10.1016/j.annepidem.2021.11.009] [Medline: 34921991]
28. American Community Survey 5-Year Data (2017-2021). United States Census Bureau. URL: https://www.census.gov/data/developers/data-sets/acs-5year.html [accessed 2023-07-05]
29. COVID-19 vaccine dashboard. Ohio Department of Health. URL: https://coronavirus.ohio.gov/dashboards/covid-19-vaccine/covid-19-vaccination-dashboard [accessed 2023-07-05]
30. Della Vella D, Rayo MF. The final inch: What pop-up COVID testing tells us about community engagement. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care 2022 Oct 06;11(1):76-81 [FREE Full text] [doi: 10.1177/2327857922111015]
31. Brakefield WS, Ammar N, Olusanya OA, Shaban-Nejad A. An urban population health observatory system to support COVID-19 pandemic preparedness, response, and management: Design and development study. JMIR Public Health Surveill 2021 Jun 16;7(6):e28269 [FREE Full text] [doi: 10.2196/28269] [Medline: 34081605]
32. Chen KL, Brozen M, Rollman JE, Ward T, Norris KC, Gregory KD, et al. How is the COVID-19 pandemic shaping transportation access to health care? Transp Res Interdiscip Perspect 2021 Jun;10:100338 [FREE Full text] [doi: 10.1016/j.trip.2021.100338] [Medline: 34514368]
33. Williams DR, Cooper LA. COVID-19 and health equity-a new kind of "herd immunity". JAMA 2020 Jun 23;323(24):2478-2480. [doi: 10.1001/jama.2020.8051] [Medline: 32391852]
34. Magesh S, John D, Li WT, Li Y, Mattingly-App A, Jain S, et al. Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: A systematic-review and meta-analysis. JAMA Netw Open 2021 Nov 01;4(11):e2134147 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.34147] [Medline: 34762110]
35. Bollyky T, Castro E, Aravkin A, Bhangdia K, Dalos J, Hulland EN, et al. Assessing COVID-19 pandemic policies and behaviours and their economic and educational trade-offs across US states from Jan 1, 2020, to July 31, 2022: an observational analysis. Lancet 2023 Apr 22;401(10385):1341-1360 [FREE Full text] [doi: 10.1016/S0140-6736(23)00461-0] [Medline: 36966780]
36. Philbin MM, Parker CM, Flaherty MG, Hirsch JS. Public libraries: a community-level resource to advance population health. J Community Health 2019 Feb;44(1):192-199 [FREE Full text] [doi: 10.1007/s10900-018-0547-4] [Medline: 29995303]
37. Sollenberger JF, Holloway RG. The evolving role and value of libraries and librarians in health care. JAMA 2013 Sep 25;310(12):1231-1232. [doi: 10.1001/jama.2013.277050] [Medline: 24065006]
38. Taylor SP, Kowalkowski MA. Using implementation science-guided pilot studies to assess and improve the informativeness of clinical trials. J Gen Intern Med 2021 Feb;36(2):533-536 [FREE Full text] [doi: 10.1007/s11606-020-06220-3] [Medline: 32918201]
39. Abdul-Mutakabbir JC, Tillman F, Marcelin JR, Saunders IM, Arya V. Slowed progression: the utility of Test to Treat initiatives in improving the neglected inequities of COVID-19 among racially/ethnically minoritized groups. J Am Pharm Assoc (2003) 2023;63(1):424-429 [FREE Full text] [doi: 10.1016/j.japh.2022.10.025] [Medline: 36470731]
40. Wroe EB, Seung KJ, Baker BK, Farmer PE. Test and treat: a missing link in the global fight against COVID-19. Lancet Glob Health 2022 Feb;10(2):e181-e182 [FREE Full text] [doi: 10.1016/S2214-109X(21)00568-4] [Medline: 34953515]

Abbreviations

CDN: Columbus Department of Neighborhoods
CML: Columbus Metropolitan Libraries
FAAST: Flexible Adaptive Algorithmic Surveillance Testing
ODH: Ohio Department of Health
ONG: Ohio National Guard
OSU: The Ohio State University
RAT: rapid antigen test
RCT: randomized controlled trial
WFU: Wake Forest University

Edited by T Sanchez, A Mavragani; submitted 27.06.22; peer-reviewed by J Wang, Y Xie; comments to author 09.02.23; revised version received 13.04.23; accepted 15.06.23; published 15.08.23.

Please cite as:
Rayo MF, Faulkner D, Kline D, Thornhill IV T, Malloy S, Della Vella D, Morey DA, Zhang N, Gonsalves G
Using Bandit Algorithms to Maximize SARS-CoV-2 Case-Finding: Evaluation and Feasibility Study
JMIR Public Health Surveill 2023;9:e39754
URL: https://publichealth.jmir.org/2023/1/e39754
doi:10.2196/39754
PMID:37581924

Original Paper

Comparison of Four Active SARS-CoV-2 Surveillance Strategies in Representative Population Sample Points: Two-Factor Factorial Randomized Controlled Trial

Andreas Deckert[1], PhD; Simon Anders[2], Prof Dr; Ivonne Morales[3], PhD; Manuela De Allegri[1], Prof Dr; Hoa Thi Nguyen[1], PhD; Aurélia Souares[1], PhD; Shannon McMahon[1], PhD; Matthias Meurer[2], Dipl Ing; Robin Burk[2], PhD; Dan Lou[2], PhD; Lucia Brugnara[4], MSc; Matthias Sand[5], PhD; Lisa Koeppel[3], PhD; Lena Maier-Hein[6], Prof Dr; Tobias Ross[6]; Tim J Adler[6], PhD; Stephan Brenner[1], PhD; Christopher Dyer[7], PhD; Konrad Herbst[2], PhD; Svetlana Ovchinnikova[2], PhD; Michael Marx[4], Prof Dr; Paul Schnitzler[8], Prof Dr; Michael Knop[2], Prof Dr; Till Bärnighausen[1], Prof Dr; Claudia M Denkinger[3], PD Dr

[1]Heidelberg Institute of Global Health, Heidelberg, Germany
[2]Center for Molecular Biology Heidelberg, Heidelberg, Germany
[3]Division of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany
[4]evaplan GmbH at the University Hospital, Heidelberg, Germany
[5]GESIS Leibniz-Institute for the Social Sciences, Mannheim, Germany
[6]Division of Computer Assisted Medical Interventions, German Cancer Research Centre, Heidelberg, Germany
[7]Gesundheitsamt des Rhein-Neckar-Kreises, Heidelberg, Germany
[8]Center of Infectious Diseases, Virology, Heidelberg University Hospital, Heidelberg, Germany

Corresponding Author:
Andreas Deckert, PhD
Heidelberg Institute of Global Health
INF 324
Heidelberg, 69120
Germany
Phone: 49 6221 4537260
Email: a.deckert@uni-heidelberg.de

Related Article:

This is a corrected version. See correction statement: https://publichealth.jmir.org/2024/1/e57203

Abstract

Background: The COVID-19 pandemic is characterized by rapid increases in infection burden owing to the emergence of new variants with higher transmissibility and immune escape. To date, monitoring the COVID-19 pandemic has mainly relied on passive surveillance, yielding biased epidemiological measures owing to the disproportionate number of undetected asymptomatic cases. Active surveillance could provide accurate estimates of the true prevalence to forecast the evolution of the pandemic, enabling evidence-based decision-making.

Objective: This study compared 4 different approaches of active SARS-CoV-2 surveillance focusing on feasibility and epidemiological outcomes.

Methods: A 2-factor factorial randomized controlled trial was conducted in 2020 in a German district with 700,000 inhabitants. The epidemiological outcome comprised SARS-CoV-2 prevalence and its precision. The 4 study arms combined 2 factors: individuals versus households and direct testing versus testing conditioned on symptom prescreening. Individuals aged ≥7 years were eligible. Altogether, 27,908 addresses from 51 municipalities were randomly allocated to the arms and 15 consecutive recruitment weekdays. Data collection and logistics were highly digitized, and a website in 5 languages enabled low-barrier registration and tracking of results. Gargle sample collection kits were sent by post. Participants collected a gargle sample at home and mailed it to the laboratory. Samples were analyzed with reverse transcription loop-mediated isothermal amplification (RT-LAMP); positive and weak results were confirmed with real-time reverse transcription–polymerase chain reaction (RT-PCR).

Results: Recruitment was conducted between November 18 and December 11, 2020. The response rates in the 4 arms varied between 34.31% (2340/6821) and 41.17% (2043/4962). The prescreening classified 16.61% (1207/7266) of the patients as COVID-19 symptomatic. Altogether, 4232 persons without prescreening and 7623 participating in the prescreening provided 5351 gargle samples, of which 5319 (99.4%) could be analyzed. This yielded 17 confirmed SARS-CoV-2 infections and a combined prevalence of 0.36% (95% CI 0.14%-0.59%) in the arms without prescreening and 0.05% (95% CI 0.00%-0.108%) in the arms with prescreening (initial contacts only). Specifically, we found a prevalence of 0.31% (95% CI 0.06%-0.58%) for individuals and 0.35% (95% CI 0.09%-0.61%) for households, and lower estimates with prescreening (0.07%, 95% CI 0.0%-0.15% for individuals and 0.02%, 95% CI 0.0%-0.06% for households). Asymptomatic infections occurred in 27% (3/11) of the positive cases with symptom data. The 2 arms without prescreening performed the best regarding effectiveness and accuracy.

Conclusions: This study showed that postal mailing of gargle sample kits and returning home-based self-collected liquid gargle samples followed by high-sensitivity RT-LAMP analysis is a feasible way to conduct active SARS-CoV-2 population surveillance without burdening routine diagnostic testing. Efforts to improve participation rates and integration into the public health system may increase the potential to monitor the course of the pandemic.

Trial Registration: Deutsches Register Klinischer Studien (DRKS) DRKS00023271; https://tinyurl.com/3xenz68a

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-021-05619-5

(JMIR Public Health Surveill 2023;9:e44204) doi:10.2196/44204

KEYWORDS

COVID-19; SARS-CoV-2; randomized controlled trial; multiarm; cluster sampling; surveillance; effectiveness; pandemic

Introduction

Background

Epidemiological surveillance of populations using serological or nucleic acid–based diagnostics is a well-known public health preparedness tool and is, for example, implemented in the global influenza surveillance network. Here, a system of World Health Organization collaborating centers acts mainly reactively to seasonally adjust influenza vaccines depending on emerging variants in representative samples [1,2]. In contrast, the COVID-19 pandemic has been characterized by less seasonality and a comparatively faster periodicity of episodes in which the burden of infection in the population rapidly increased because of the emergence of new variants with higher transmissibility and immune escape. The ever-present threat of overburdening the health care systems resulted in the implementation of containment measures based on limited forecasting tools [3].

Both presymptomatic and asymptomatic SARS-CoV-2 carriers are able to transmit; thus, the proportion of asymptomatic cases plays a key role in the spread of the virus, challenging the containment of this highly contagious disease [4-7]. With respect to the SARS-CoV-2 alpha variant (B.1.1.7, prevalent at the time of the study), asymptomatic or minimally symptomatic SARS-CoV-2 carriers accounted for approximately 40% of all unvaccinated infected individuals [8]. A surveillance system that randomly samples a representative population to test for active infection is therefore needed for pandemic preparedness. If combined with longitudinal serological studies to monitor historical exposure and thus the extent of immunity, the prevalence estimates could be incorporated into a prognostic model to effectively monitor the pandemic and predict the further course of the outbreak. This would enable evidence-based decision-making, especially as strict lockdowns and curfews can only serve as temporary measures [9]. Furthermore, active disease surveillance could assess symptomatology in combination with downstream genome sequencing [10].

Most countries mainly used results from different testing purposes in the so-called passive surveillance as a surrogate for active systems to restrictedly monitor the course of the pandemic [11,12]. This involves recording persons presenting to the health system with COVID-19–like symptoms to initiate their treatment, as well as symptomatically and asymptomatically infected persons identified through targeted screening (eg, workplaces and retirement homes), supplemented by contacts of detected cases who tested positive with the aim of isolating these individuals and containing the spread of the infection [13]. Others extended existing primary care influenza surveillance systems to virological and serological SARS-CoV-2 surveillance [14]. However, these largely symptom-driven surveillance systems do not provide representative prevalence estimates and are therefore of limited use in terms of informing decision makers [11]. Similarly, surveillance studies that rely on syndromic surveillance rather than direct testing focus on symptomatic cases only and are further limited as they do not differentiate between respiratory pathogens [15,16]. In contrast, active surveillance would test a representative proportion of the population without necessarily assuming the presence of COVID-19 symptoms [17].

However, testing for SARS-CoV-2 for surveillance in communities requires extensive, complex testing capacities and different strategies compared with testing for individual health purposes and screening with respect to sampling, turnaround time, and reporting [18,19]. Real-time reverse transcription–polymerase chain reaction (RT-PCR), a laboratory technology used to detect viral RNA, is considered the gold standard for SARS-CoV-2 detection during acute infection and up to 2 to 3 weeks thereafter [20]. However, RT-PCR is costly and only applicable in population screening if pooling techniques are applied [21,22]. Alternatively, reverse transcription loop-mediated isothermal amplification (RT-LAMP)–based method allows for the detection of viral RNA with sensitivity similar to that of RT-PCR but with reduced cost and straightforward application, not requiring a thermal cycler, hence enabling scale-up testing [23-26].

In addition, there is sufficient evidence that saliva, gargle solution, and expectorated mucus are suitable sources of biological material for the sensitive detection of SARS-CoV-2 [27-32]. Saliva samples are very stable as their viral load hardly changes even over several days at room temperature (20-25 °C), which makes it easily transportable even by post [33]. Therefore, the combination of self-collected saliva samples shipped by post and analyzed by RT-LAMP appears to be a promising approach for active population surveillance.

However, there is limited evidence regarding the use of comprehensive and effective active surveillance strategies. Most of the active surveillance studies conducted in the general population have exhibited a repeated cross-sectional design. To date, the best-described active surveillance systems of similar type have been implemented in the United Kingdom using RT-PCR assays and self-administered swabs [19,34]. These 2 attempts have demonstrated the importance and viability of implementing active surveillance for SARS-CoV-2 in affluent areas by providing significant SARS-COV-2 prevalence estimates in the community during various time points. Some studies have observed fixed cohorts over time [35,36]. Although these studies have shown the feasibility of active surveillance, no study to date has compared different surveillance approaches to assess their comparative feasibility, effectiveness, and cost. Nonetheless, the combination of these factors is relevant to public health decision makers facing resource constraints, with testing capacity depending mainly on external factors (eg, production of reagents and test materials), independent of the testing strategy.

The question for decision makers in health policy is therefore which active surveillance strategy is most effective in monitoring the actual progress of an epidemic or pandemic with a limited number of available tests and without limiting the capacity of passive surveillance.

Objectives

The objective of this study was to compare 4 different approaches to population-level surveillance of active SARS-CoV-2 infection and to report the epidemiological outcomes and feasibility. The cost-effectiveness was reported separately [37].

Methods

This study followed the principles of CONSORT (Consolidated Standards of Reporting Trials; extension for multiarm parallel-group randomized trials) statement (Multimedia Appendix 1) [38]. Details regarding the methods and study design are provided in the study protocol by Deckert et al [39].

Ethics Approval

The trial was approved by the ethics committee at the University of Heidelberg on November 2, 2020 (amendment November 9, 2020; file number S790/2020).

Participation and Informed Consent

The study applied the Declaration of Helsinki ethical standards and adhered to the legal requirements for research on humans in Germany as stated in the guidelines for good clinical practice and the Medical Devices Act (Medizinproduktegesetz) issued by the Ministry of Health of Germany and implemented by the Federal Institute for Drugs and Medical Devices.

With the invitation letter, participants were informed about the study purpose and procedures in an age group–appropriate manner by means of a detailed brochure; all web-based information was also available in 5 languages. Children and adolescents were asked to assent their participation with a mandatorily complemented consent of their legal representatives. Consent or assent to participate could be given on the web or on paper, whereas additional consent or assent to get the gargle sample analyzed had to be provided on paper (Multimedia Appendix 2 [20-26,28-33,39-48]). Participants' addresses were provided by the registers of residents based on a German law (Bundesmeldegesetz, § 46 Gruppenauskunft) that allows research institutions to draw collective address data for research purposes; consent was sought from the population register offices.

Data protection measures were implemented according to the Basic Data Protection Regulation (article 6, paragraph 1, letter a Datenschutz-Grundverordnung) and the Federal Data Protection Act. The manner of data collection and scope and content of data collected were authorized by the data protection officer of Heidelberg University. The consent forms provided information about the data protection measures and pseudonymization of data to identify individuals who tested positive; data analyses were conducted with anonymized data only. Withdrawal from study participation and deletion of data were possible at any time during the study. Participants were informed that the data could be used beyond the study for other research purposes related to SARS-CoV-2 or other respiratory viruses; however, the gargle samples were destroyed after study completion.

Compensation for study participation was not provided.

Trial Registration

The trial was registered (November 30, 2020) on the German Clinical Trials Register (registration number DRKS00023271), and the study protocol has been published accordingly [49].

Study Design

The study was designed as a 2-factor factorial, multiarm parallel randomized controlled trial. The 4 study arms represented all combinations of two factors: i) testing unconditional (A) versus testing under the condition of upstream COVID-19 symptom prescreening (B), and ii) testing individuals (1) versus households (2). The designation of the 4 study arms A1, A2, B1, and B2 depict these combinations.

Individuals aged ≥7 years who consented (by legal guardian in the case of minors) were eligible and provided a self-collected gargle sample (all 4 study arms) after completion of a prescreening questionnaire (B1 and B2).

Study Setting

The trial was conducted during the second SARS-CoV-2 wave in fall 2020 in Germany in Heidelberg and the Rhine-Neckar district, which is the location of Heidelberg University and the catchment area of the closely cooperating district health authority. The public health authority of the Rhine-Neckar district, also responsible for Heidelberg city, is responsible for 56 municipalities with approximately 700,000 inhabitants. Containment measures were tightened by mid-December 2020, and vaccination was not yet available at that time.

Outcomes

We considered effectiveness, feasibility, and costs to comprehensively inform policy makers as trade-offs are to be expected when the evidence is reviewed by an evidence-to-decision framework, for example, Grading of Recommendations, Assessment, Development, and Evaluations [49]. Here, we report the feasibility and epidemiological outcomes and review the cost-effectiveness separately. The epidemiological outcomes included the following: (1) cumulative SARS-CoV-2 prevalence and (2) prevalence precision.

The secondary outcomes were the participation rate (number of gargle samples; A1 and A2), the prescreening results or gargle samples (B1 and B2) divided by the number of contacts, the population prevalence estimation, and the symptom frequencies.

Sample Size

As high-quality test capacity is a limiting resource in active surveillance, we designed the study so that each arm would end up with the same number of laboratory tests, competing for feasibility, effectiveness, and cost-effectiveness. However, a composite end point was not possible, and a priori assumptions regarding cost-effectiveness could only be made arbitrarily. Therefore, we powered the effectiveness part to finally receive the same number of gargle samples in each arm to estimate a specific prevalence with a certain precision, based on the estimator theoretically closest to the true prevalence in study arm A1 (randomly selected individuals). The costs required to achieve this equal number of gargle samples might differ by design in the study arms, as may the number of positive cases among the gargle samples. Strategy A2 was designed to account for the likelihood of higher infection rates within households of infected individuals. In contrast, the detection rates of strategies B1 and B2 were expected to differ because of the different effects of the upstream prescreening mechanisms.

At the time of planning the study, the 4-week cumulative SARS-CoV-2 prevalence was estimated to be 0.25% in the region, based on data from the first surge [50], mainly sourced from passive surveillance (ie, symptomatic persons). Assuming a similarly high proportion of asymptomatic SARS-CoV-2 carriers, the overall prevalence was estimated to be approximately 0.5%. Thus, in the most representative study arm A1, which used a purely random sample to directly reflect the pandemic events, a total of 2500 gargle samples could estimate such a prevalence with a relative precision of −0.35% to +0.35% (95% CI) and a power of 95% [51]. Hence, assuming a 50% response rate after one-time prompting, at least 5000 addresses were required for strategy A1. For strategy A2, assuming an average household size of 2 and a response rate of 50% after prompting, a gross sample of at least 2500 addresses was necessary [52]. For strategy B1, we assumed an 80% response to the sampling request after positive prescreening; hence, at least 3125 sample kits had to be mailed out to yield 2500 tested samples. On the basis of this and assuming a SARS-CoV-2 prevalence of 0.5% and a prescreening tool sensitivity of 90% and specificity of 70%, at least 10,313 participants had to complete the initial prescreening. Furthermore, expecting a 50% response to the prescreening request and 50% response to prompting required an initial gross sample of at least 13,750 addresses. Similarly, for strategy B2, a gross sample of at least 6875 addresses was required (average household size of 2). Consequently, a total gross sample of 28,125 addresses was required, and the participants were allocated in a ratio of 10:5:28:14 to the study arms (Figure S1 in Multimedia Appendix 2) [39].

Randomization

A stratified general population representative sample with 2 strata (Rhine-Neckar district plus Heidelberg city) and 2 different sampling approaches (Heidelberg—simple random sample and Rhine-Neckar—2 stage cluster sampling) were used to draw 3 weekly batches to avoid negative impacts on the response rate owing to uncertain external circumstances (such as holidays or a higher rate of infections). Because of the considerably larger population of Heidelberg city, the stratification was used to avert a decrease in statistical power for estimation in rural areas. The batches were then divided among 15 consecutive recruitment weekdays and randomly assigned to the 4 study arms. The gross sample included 21,739 individuals in the Rhine-Neckar district and 6386 individuals in the Heidelberg stratum. Potential participants were randomly drawn from municipalities’ (51/53, 96%) population registers as samples proportional to the population size of the municipalities (Multimedia Appendix 2) [39].

The 2-stage sampling was designed by the Leibnitz Institute for Social Sciences in Mannheim, Germany. The Heidelberg Institute of Global Health requested the addresses from the municipalities.

Recruitment and Study Materials

Study recruitment started on November 18, 2020, and ended on December 11, 2020, and one-time reminders were sent until December 16, 2020. The general implementation (including testing activities and hotlines) lasted until December 23, 2020.

The study website displayed information in 5 languages (German, Turkish, English, Russian, and Italian) and contained a gargling demonstration video. The study information (brochure and website) and the consent forms were separately presented in adequate language for children (aged 7-11 years), adolescents (aged 12-17 years), parents, and adults. In addition, a hotline was implemented and made accessible during the study period (weekdays 7-11 AM and 2-6 PM).

A prescreening questionnaire was provided to all initially contacted individuals. Its completion was voluntary in A1 and A2 and mandatory in B1 and B2 study arms.

Logistics

The study arms A1 and A2 immediately received gargle sample collection kits, whereas the study arms B1 and B2 had to complete the prescreening questionnaire first. The differences in logistics in are discussed in detail as follows:

1. Study arm A1: Randomly selected individuals (hereinafter referred to as initially contacted individuals) received an invitation letter by post containing the study information, a sample collection kit, and a stamped biohazard return envelope (Multimedia Appendix 3).
2. Study arm A2: The procedures were the same as in A1 with the difference that the initially contacted individuals received 4 kits to sample household members. Additional kits could be ordered by email or via the hotline if needed.
3. Study arm B1: Initially contacted individuals first received an invitation letter with the study information, a stamped return envelope, and the prescreening questionnaire, asking about 16 typical COVID-19 symptoms to be completed web-based or in print. Next, the questionnaire was evaluated using a random forest algorithm that had been trained on data sets of patients with COVID-19 to classify participants into COVID-19–free and potentially sick individuals. If the questionnaire was flagged by the algorithm, the individual received a second envelope containing the sample collection kit and a stamped biohazard return envelope.
4. Study arm B2: The procedures were the same as in B1 with the difference that the initially contacted individuals who were classified as having COVID-19 by the prescreening algorithm subsequently received 4 sample collection kits for household members.

In each study arm, participants who received the sample collection kits were asked to collect a gargle sample themselves after gargling 5 mL of saline at home and return it to the laboratory (Multimedia Appendices 2 and 3). The participants were expected to take the sample latest 1 day after they received the sample collection kit.

Laboratory and Blinding

To enable scale-up of testing while maintaining sensitivity, an RT-LAMP was used for analyzing the liquid gargle samples. Weakly (<3 replicates positive) and clearly positive gargle samples (all replicates positive) were subsequently analyzed using RT-PCR as a confirmation test [24-26].

Laboratory staff who conducted the RT-LAMP and the RT-PCR confirmation tests were blinded with regard to the study arms. The vials carried a unique barcode only, and the laboratory data were kept on a separate server; linkage of laboratory data with participants’ data was performed after the laboratory analysis was completed.

Data Processing

Data cleaning was performed using SAS software (version 9.4 TS1M4; SAS Institute Inc). Statistical analyses were performed using R (version 4.1.1; R Foundation for Statistical Computing) and Stata (version 15.1; StataCorp).

Statistical Methods

Prescreening Questionnaire

We developed a symptom screening algorithm using machine learning. The underlying data sets were from various settings with and without SARS-CoV-2 infected patients, including samples from a general population screen and persons tested for SARS-CoV-2 infection because of symptoms or high-risk exposure (Multimedia Appendix 2). The threshold of the algorithm was optimized toward sensitivity (80% sensitivity and 75% specificity).

Descriptive Statistics

Differences in demographics, frequency of COVID-19 symptoms, and epidemiological variables between the trial arms were assessed using the chi-square test and ANOVA (no adjustment for multiple testing). The significance level α was set at 5%.

Prevalence Estimation

The total number of SARS-CoV-2 cases was calculated using the Horvitz-Thomson estimator to account for unequal selection probabilities in the 2-stage sampling [40]. The individual probability to get selected for participation was based on the number of primary sampling units per municipality, the population in this municipality obtained from the Leibniz-Institute for the Social Sciences, and was normalized for each study arm separately to match the number of performed tests. The Brewer approximation was used to calculate the estimate’s variance [41]. We refrained from calculating cluster-adjusted prevalence in A2 and B2 as the number of positive cases was too small to take the within household transmission rates into account. To compare the results of our study with the reported number of cases, we estimated the reported prevalence based on official data from the Robert Koch Institute (RKI; Multimedia Appendix 2).

Sensitivity Analysis

The trial was implemented in parallel to the existing passive surveillance system; hence, persons could have been captured by both systems. Some hotline callers refused to participate because they had already tested positive in the passive surveillance. Therefore, we conducted a sensitivity analysis and recalculated the prevalence estimates to account for missed positive cases owing to overlap with passive surveillance (Multimedia Appendix 2).

Results

Recruitment

Altogether, 30,629 addresses were provided by 51 municipalities’ registration offices. After excluding duplicates, 27,908 (99.23%) addresses out of the planned sample size of 28,125 were randomly allocated to the arms (Figure 1; Table

S1 in Multimedia Appendix 2). Considering the longer individual study course in B1 and B2 because of the prescreening questionnaire, shipment of study materials and recruitment of participants started on November 18, 2020, for B1 and B2 and on November 23, 2020, for A1 and A2 and stopped on December 8, 2020, for B1 and B2 and on December 11, 2020, for A1 and A2 (Figure 2). Similar gargle sampling days should be achieved for all study arms. In total, 11,855 participants were recruited.

Figure 1. Study flowchart. A1: individuals without mandatory prescreening; A2: households without mandatory prescreening; B1: individuals with mandatory prescreening; B2: households with mandatory prescreening; LAMP: loop-mediated isothermal amplification.

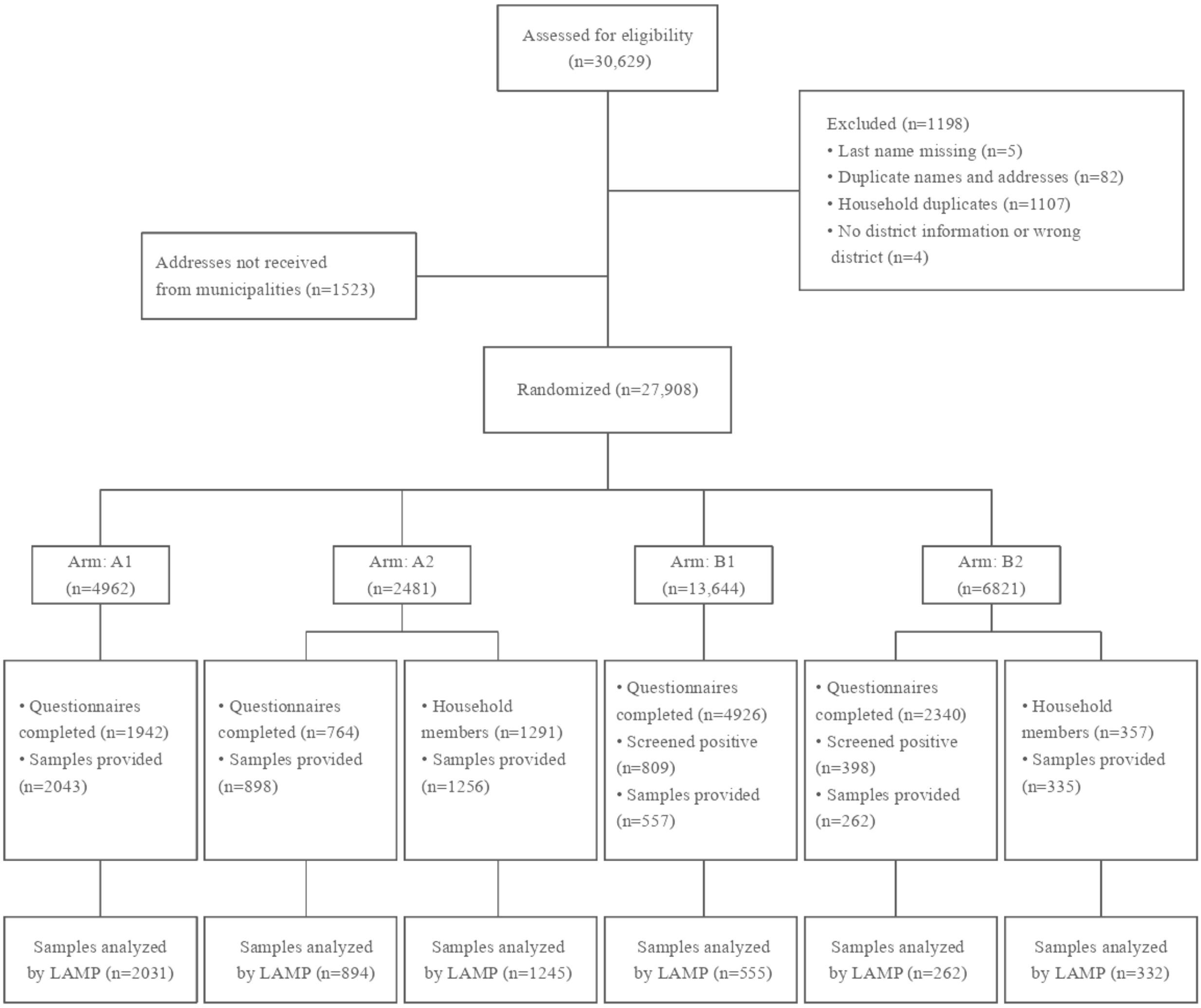
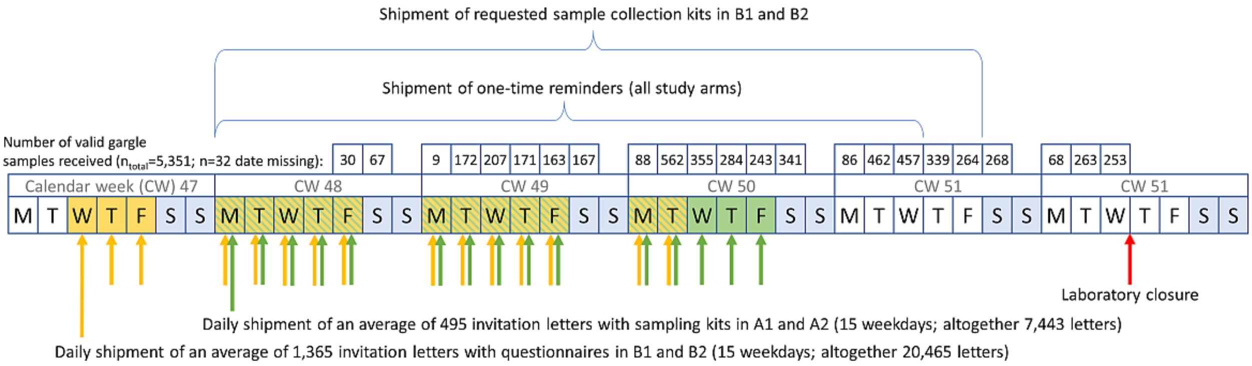


Figure 2. Study timeline. A1: individuals without mandatory prescreening; A2: households without mandatory prescreening; B1: individuals with mandatory prescreening; B2: households with mandatory prescreening.



Demographics

The participants were slightly older than the general population (mean age 44.5 years in 2019; [53]), mainly because of the exclusion of children aged <7 years (Table S2 in Multimedia Appendix 2; Table 1). The study arms differed with regard to the education level of participants who provided a valid questionnaire (Table 2).

Table 1. Demographics of initially contacted individuals who decided to participate (initially contacted individuals who either provided a sample [A1 and A2] or a questionnaire [B1 and B2]).

Demographics	A1[a] (n=2043)	A2[b] (n=898)	B1[c] (n=4926)	B2[d] (n=2340)	Total (n=10,207)	*P* value
Sex[e], n (%)						.64[f]
Female	1066 (52.18)	471 (52.45)	2615 (53.09)	1265 (54.06)	5417 (53.07)	
Male	977 (47.82)	426 (47.44)	2311 (46.91)	1075 (45.94)	4789 (46.92)	
Missing[g]	0 (0)	1 (0.11)	0 (0)	0 (0)	1 (0.01)	
Age (years)[e]						.52[h]
Value, mean (SD; range)	46.8 (20.4; 7-99)	47.5 (20.6; 7-93)	47.2 (20.7; 7-98)	47.7 (20.6; 7-98)	47.3 (20.6; 7-99)	
Missing, n (%)[g]	0 (0)	1 (0.11)	0 (0)	27 (1.15)	28 (0.27)	

[a] A1: individual participants without prescreening.

[b] A2: initially contacted individuals in households who decided to participate without prescreening.

[c] B1: individual participants with prescreening.

[d] B2: initially contacted individuals in households with prescreening.

[e] Age and sex were provided by the registration offices.

[f] Pearson chi-square test.

[g] Number of participants with missing information.

[h] Linear model ANOVA.

Table 2. Demographics of participants who provided a valid questionnaire.

Demographics	A1[a] (n=1942)	A2[b,c] (n=824)	B1[d] (n=4926)	B2[e,f] (n=2346)	Total (n=10,038)	*P* value
Sex[g], n (%)						.68[h]
Female	1019 (52.47)	427 (52.07)	2615 (53.09)	1267 (54.05)	5328 (53.11)	
Male	923 (47.53)	393 (47.93)	2311 (46.91)	1077 (45.95)	4704 (46.89)	
Missing[i]	0 (0)	4 (0.49)	0 (0)	2 (0.08)	6 (0.06)	
Age (years)[g]						.33[j]
Values, mean (SD; range)	46.7 (20.5; 7-99)	46.6 (20.7; 7-93)	47.2 (20.7; 7-98)	47.7 (20.5; 7-98)	47.2 (20.6; 7-99)	
Missing, n (%)[i]	0 (0)	6 (0.73)	0 (0)	2 (0.09)	8 (0.08)	
Education, n (%)						<.001[h]
Doctor of Philosophy	121 (6.97)	49 (6.55)	286 (6.29)	150 (6.94)	606 (6.59)	
Master studies or diploma studies	416 (23.95)	182 (24.33)	1041 (22.9)	538 (24.9)	2177 (23.68)	
Bachelor studies	171 (9.84)	60 (8.02)	393 (8.64)	177 (8.19)	801 (8.71)	
Master craftsman training	99 (5.7)	39 (5.21)	236 (5.19)	99 (4.58)	473 (4.15)	
High school or technical or economic high school	267 (15.37)	144 (19.25)	775 (17.05)	327 (15.13)	1513 (16.46)	
Vocational school	236 (13.59)	97 (12.97)	541 (11.9)	244 (11.29)	1118 (12.16)	
Secondary school	330 (19.0)	126 (16.84)	1072 (23.58)	532 (24.62)	2060 (22.41)	
No degree or still in school education	97 (5.58)	51 (6.82)	202 (4.44)	94 (4.35)	444 (4.83)	
Missing[i]	205 (10.56)	76 (9.22)	380 (7.71)	185 (7.89)	846 (8.43)	
Job description, n (%)						.04[h]
Managers	78 (4.72)	26 (3.61)	215 (5.14)	108 (5.38)	427 (4.99)	
Professionals	428 (25.92)	174 (24.13)	907 (21.68)	478 (23.83)	1987 (23.21)	
Technicians and associate professionals	169 (10.24)	84 (11.65)	467 (11.16)	232 (11.57)	952 (11.12)	
Clerical support workers	90 (5.45)	36 (4.99)	230 (5.5)	89 (4.44)	445 (5.2)	
Services and sales workers	135 (8.18)	48 (6.66)	289 (6.91)	130 (6.48)	602 (7.03)	
Skilled agricultural, forestry, and fishery workers	3 (0.18)	6 (0.83)	15 (0.36)	7 (0.35)	31 (0.36)	
Craft and related trades workers	65 (3.94)	31 (4.3)	156 (3.73)	71 (3.54)	323 (3.77)	
Plant and machine operators and assemblers	15 (0.91)	2 (0.28)	40 (0.96)	31 (1.55)	88 (1.03)	
Elementary occupations	29 (1.76)	15 (2.08)	63 (1.51)	37 (1.84)	144 (1.68)	
Not part of the labor force	577 (34.95)	267 (37.03)	1614 (38.58)	746 (37.19)	3204 (37.42)	
Other	9 (0.55)	5 (0.69)	27 (0.65)	10 (0.5)	51 (0.6)	
Ambiguous or incomplete responses	53 (3.21)	27 (3.74)	161 (3.85)	67 (3.34)	308 (3.6)	
Missing[i]	291 (14.98)	103 (12.5)	742 (15.06)	340 (14.49)	1476 (14.70)	
Employment status, n (%)						.26[h]
Employed but not specified	849 (51.52)	356 (49.44)	1973 (47.26)	987 (49.28)	4165 (48.74)	
Self-employed paid work	23 (1.4)	7 (0.97)	76 (1.82)	35 (1.75)	141 (1.65)	

Demographics	A1[a] (n=1942)	A2[b,c] (n=824)	B1[d] (n=4926)	B2[e,f] (n=2346)	Total (n=10,038)	*P* value
Nonself-employed paid work	122 (7.4)	46 (6.39)	302 (7.23)	156 (7.79)	626 (7.33)	
Service contract	1 (0.06)	0 (0)	2 (0.05)	0 (0)	3 (0.04)	
Incapacitated for work	2 (0.12)	0 (0)	6 (0.14)	3 (0.15)	11 (0.13)	
Working from home	5 (0.3)	1 (0.14)	16 (0.38)	4 (0.2)	26 (0.3)	
Parental leave	13 (0.79)	14 (1.94)	39 (0.93)	22 (1.1)	88 (1.03)	
Unpaid family workers	10 (0.61)	4 (0.56)	19 (0.46)	9 (0.45)	42 (0.49)	
Pupil	47 (2.85)	26 (3.61)	155 (3.71)	59 (2.95)	287 (3.36)	
Student	108 (6.55)	45 (6.25)	289 (6.92)	121 (6.04)	563 (6.59)	
Internship	3 (0.18)	0 (0)	5 (0.12)	3 (0.15)	11 (0.13)	
Further education (full time, long term)	1 (0.06)	2 (0.28)	3 (0.07)	1 (0.05)	7 (0.08)	
Doctor of Philosophy or doctorate	9 (0.55)	3 (0.42)	23 (0.55)	8 (0.4)	43 (0.5)	
Traineeship in a school or law firm	1 (0.06)	2 (0.28)	6 (0.14)	4 (0.2)	13 (0.15)	
Vocational training	16 (0.97)	12 (1.67)	76 (1.82)	26 (1.3)	130 (1.52)	
Pensioners	38 (2.31)	19 (2.64)	77 (1.84)	34 (1.7)	168 (1.97)	
Retirement	328 (19.9)	158 (21.94)	881 (21.1)	412 (20.57)	1779 (20.82)	
Partial retirement	12 (0.73)	2 (0.28)	20 (0.48)	13 (0.65)	47 (0.55)	
Unemployed	49 (2.97)	16 (2.22)	179 (4.29)	93 (4.64)	337 (3.94)	
Other (eg, military, civilian service, holiday, or illness)	4 (0.24)	3 (0.42)	10 (0.24)	2 (0.1)	19 (0.22)	
Undetermined	7 (0.42)	4 (0.56)	18 (0.43)	11 (0.55)	40 (0.47)	
Missing[i]	294 (15.14)	104 (12.62)	751 (15.24)	343 (14.62)	1492 (14.86)	
Household size, mean (SD)	2.64 (1.29)	2.73 (1.24)	2.66 (1.38)	2.65 (1.43)	2.66 (1.39)	.04[k]

[a] A1: individual participants without mandatory prescreening who voluntarily completed the prescreening questionnaire.

[b] A2: household members without mandatory prescreening who voluntarily completed the prescreening questionnaire.

[c] n=764 initially contacted individuals (n=730+34 initially contacted individuals who did not provide a valid sample or consent but answered or consented the questionnaire) plus 60 household members who additionally filled out a questionnaire.

[d] B1: individual participants with mandatory prescreening.

[e] B2: initially contacted individuals in households with mandatory prescreening.

[f] n=2340 initially contacted individuals plus 6 household members who additionally filled out a questionnaire.

[g] Age and sex were available for initially contacted individuals from the registration offices and for household members from the consent form.

[h] Pearson chi-square test.

[i] Number of participants with missing information.

[j] Linear model ANOVA.

[k] Welch 2-sample *t* test (2-tailed) between study arms A2 and B2.

Response Rates

Overall, 36.57% (10,207/27,908) of the individuals who were contacted responded by either directly providing a gargle sample (A1 and A2) or by completing the prescreening questionnaire (B1 and B2; Table 3). The response rates (A1: 41.2%, A2: 36.2%, B1: 36.1%, and B2: 34.3%) differed significantly among the study arms (*P*<.001; chi-square test), with the study arms involving household members responding the least. The lowest variation in response rates across the study region could be observed in A1 and B1 (Figure 3). Nonresponders were younger (44.2 years vs 46.6 years; *P*<.001) and comprised a larger proportion of males (50.5% vs 47.3%; *P*<.001). However, there were no additional study arm–related selection effects (responders' age: ANOVA *P*=.52; sex: chi-square *P*=.64; Table 1).

Table 3. Response rates and test results by study arm.

	No prescreening questionnaire		Prescreening questionnaire		Total
	Arm A1[a]	Arm A2[b]	Arm B1[c]	Arm B2[d]	
People contacted/group, n	4962	2481	13,644	6821	27,908
Participation rate					
Number of responders/total number of people contacted, n/N (%)[e]	2043/4962 (41.17)	898/2481 (36.19)	4926/13,644 (36.10)	2340/6821 (34.31)	10,207/27,908 (36.57)
Household members, n	N/A[f]	1291	N/A	357	1648
Questionnaires completed by initially contacted individuals					
Number completed/total number of responders, n/N (%)	1894/2043 (92.71)[g]	730/898 (81.29)[g]	4926/4926 (100)	2340/2340 (100)	9890/10,207 (96.89)
Initially contacted individuals classified as having COVID-19 in the prescreening					
Number classified as having COVID-19/total number of responders, n/N (%)	N/A	N/A	809/4926 (16.42)	398/2340 (17.01)	1207/7266 (16.61)
Samples provided by initially contacted individuals					
Number of samples provided/number of responders or classified as having COVID-19 in the prescreening, n/N (%)	2043/2043 (100)	898/898 (100)	557/809 (68.85)	262/398 (65.83)	3760/4148 (90.65)
Samples provided by household members					
Number of samples/total number of participating household members, n/N (%)	N/A	1256/1291 (97.29)	N/A	335/357 (93.84)	1591/1648 (96.54)
Samples tested for SARS-CoV-2 by LAMP[h] assay					
Number of samples tested/total number of samples received, n/N (%)	2031/2043 (99.41)	2139/2154 (99.30)[i]	555/557 (99.64)	594/597[j] (99.5)	5319/5351 (99.40)
Samples with (weakly) positive result by the LAMP assay[k]					
Number of samples with (weakly) positive LAMP results/number of samples tested, n/N (%)	26/2031 (1.28)	30/2139 (1.40)	9/555 (1.62)	10/594 (1.68)	75/5319 (1.41)
Samples with a positive result by the LAMP assay[l]					
Number of samples with positive LAMP result/number of samples with (weakly) positive LAMP results, n/N (%)	5/26 (19.23)	5/30 (16.67)	3/9 (33.33)	0/10 (0)	13/75 (17.33)
Total positive cases confirmed by PCR[m], n	6	7	3	1	17
Thereof weakly positive LAMP results confirmed positive by PCR[n]					
Number of positive PCR results/number of samples with (weakly) positive LAMP results, n/N (%)	1/26 (3.85)	2/30 (6.67)	0/9 (0)	1/10 (10)[o]	4/75 (5.33)
Thereof LAMP positive samples confirmed positive by PCR[p]					
Number of positive PCR results/number of samples with positive LAMP result, n/N (%)	5/5 (100)	5/5 (100)[q]	3/3 (100)	0/0 (0)	13/13 (100)
Asymptomatic cases[r]					

	No prescreening questionnaire		Prescreening questionnaire		Total
	Arm A1[a]	Arm A2[b]	Arm B1[c]	Arm B2[d]	
Number of asymptomatic cases/number of cases with prescreening data, n/N (%)	2/6 (33.33)	1/2 (50)	0/3 (0)	0/0 (0)	3/17 (17.65)

[a] A1: individuals without mandatory prescreening.

[b] A2: households without mandatory prescreening.

[c] B1: individuals with mandatory prescreening.

[d] B2: households with mandatory prescreening.

[e] Defined as the number of people initially contacted who provided a gargle liquid sample (arms A1 and A2) or who answered the prescreening questionnaire (arms B1 and B2), referred to as "responders," out of the total number of people contacted in each arm.

[f] N/A: not applicable.

[g] 7 questionnaires in A1 and A2 could not be considered in the random forest algorithm because of missing values in the paper version.

[h] LAMP: loop-mediated isothermal amplification.

[i] The denominator consists of the number of samples provided by initially contacted individuals in study arm A2 and their household members. The numerator consists of the samples tested by LAMP for initially contacted individuals (n=898) plus the samples tested by LAMP for household members (n=1241).

[j] The denominator consists of the number of samples provided by initially contacted individuals in study arm B2 and their household members. The numerator consists of samples tested by LAMP for initially contacted individuals (n=262) plus the samples tested by LAMP for household members (n=332).

[k] Samples with an initial weakly positive SARS-CoV-2 result by the LAMP assay.

[l] Samples that initially had a (weakly) positive SARS-CoV-2 result by the LAMP assay and that tested LAMP positive for SARS-CoV-2 upon retesting.

[m] PCR: polymerase chain reaction.

[n] Samples that initially had a weakly positive SARS-CoV-2 result by the LAMP assay, followed by a PCR positive test without an intermediate LAMP positive test.

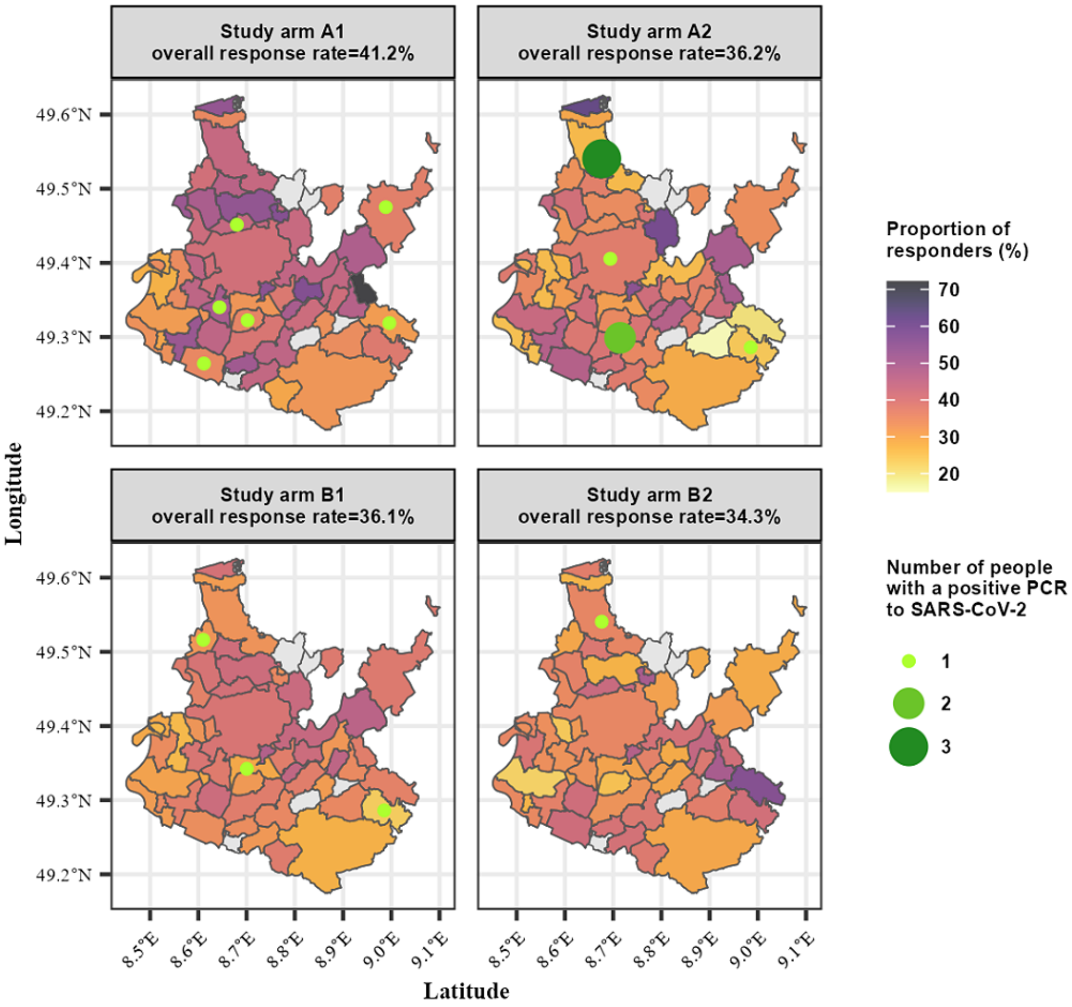
[o] Out of 1 PCR positive sample, 1 is from a household member.

[p] Samples that tested LAMP positive for SARS-CoV-2 and were then confirmed positive by PCR.

[q] Of the 5 PCR positives for SARS-CoV-2, a total of 3 were from household members.

[r] Cases without any of the following symptoms among all positive cases: fever, cough, cough with sputum, sore throat, difficulty breathing, exhaustion, headache, runny nose, muscle ache, chest pain, diarrhea, nausea, no taste or smell, chills, short breath, and confused; denominator: total number of PCR samples with questionnaire data available.

Figure 3. Response rates in the 51 municipalities; newly detected SARS-CoV-2 infections by study arm (the gray areas indicate municipalities randomly excluded in the cluster sampling approach). A1: individuals without mandatory prescreening; A2: households without mandatory prescreening; B1: individuals with mandatory prescreening; B2: households with mandatory prescreening; PCR: polymerase chain reaction.



Outcomes

Although the prescreening was not compulsory in A1 and A2, approximately 92.71% (1894/2043) of the participants in A1 and 81.3% (730/898) of the initially contacted individuals in A2 completed the optional questionnaire. The random forest algorithm classified approximately 16.61% (1207/7266) of the initially contacted individuals in B1 and B2 to have suspicious COVID-19 symptoms (Table 3). Of those who were classified as having COVID-19 in the prescreening, only 68.9% (557/809) provided a valid gargle sample upon subsequent request in B1 and 65.8% (262/398) provided a valid gargle sample upon subsequent request in B2. However, of the household members, 97.29% (1256/1291) provided valid gargle samples in A2 and 93.8% (335/357) provided valid gargle samples in B2.

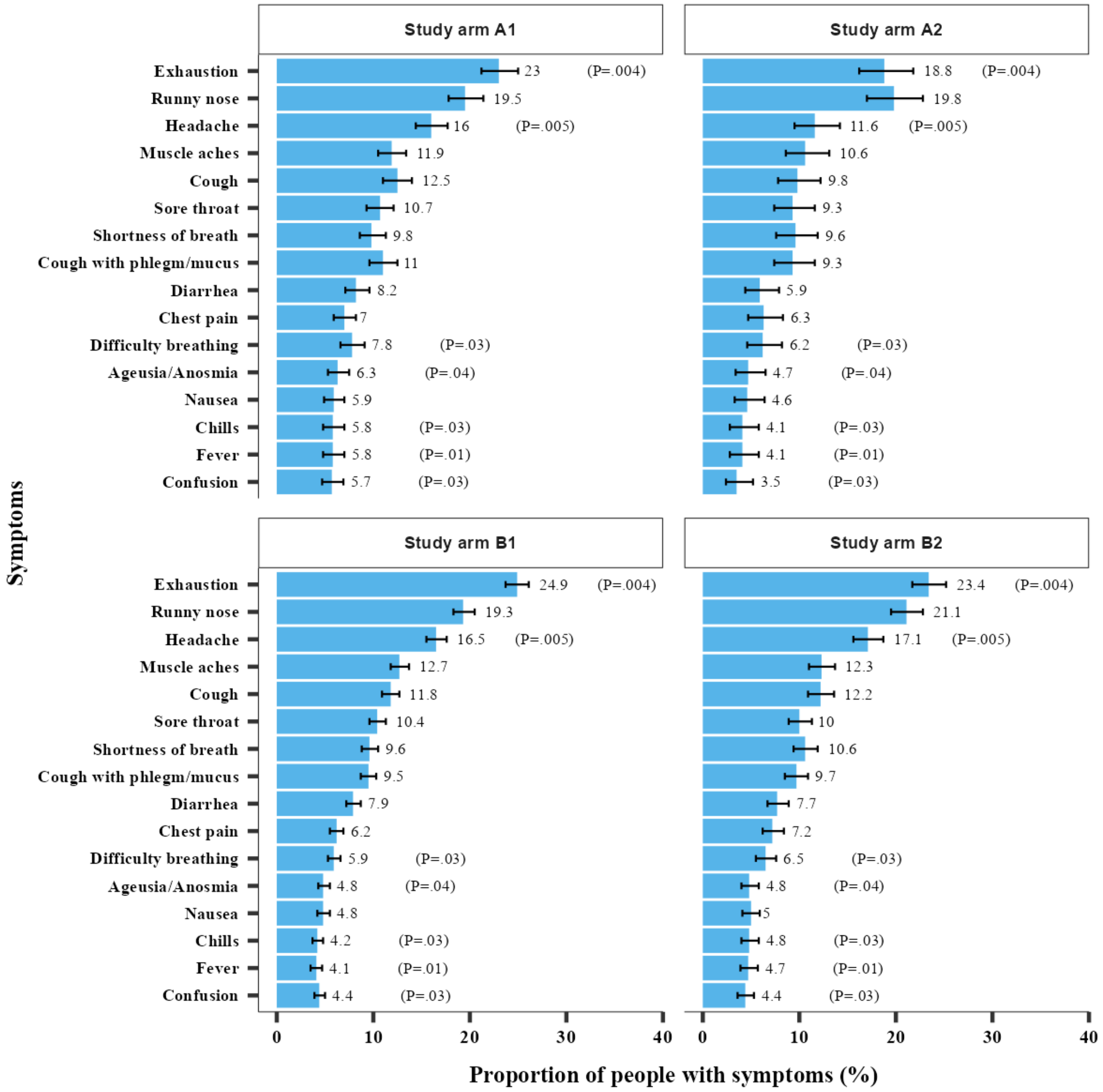
In all study arms, 99.40% (5319/5351) of the gargle samples were analyzed; only 0.60% (32/5351) of the samples could not be processed because of the viscosity being too high. All positive LAMP results were confirmed by RT-PCR. Of the weakly positive LAMP results, approximately 5% yielded a positive result in the PCR analysis. Overall, 17 SARS-CoV-2 infections were detected in this study (Figure 3; Table 3).

Of the 17 positive cases, 8 (47%) were symptomatic and 3 (18%; n=2, 67% adults in A1 and n=1, 33% child in A2) were completely asymptomatic SARS-CoV-2 carriers. Of the 8 symptomatic cases, 6 (75%) had <3 acute COVID-19 symptoms (n=4, 67% A1 and A2 and n=2, 33% B2), 1 (13%) had 8 symptoms (A1), and 1 (13%) had all 16 symptoms (B2). COVID-19 symptom data were missing for 5 participants in A2 (voluntary questionnaire) and 1 participant in B2 (household member); hence, some asymptomatic cases may not have been counted.

Overall, 8 hotline callers reported having recently tested positive and declined to provide a gargle sample. Of the 8 hotline callers, 5 (63%) received the sample collection kit or called the hotline within 14 days of testing positive elsewhere. These cases might have been detected in our study as well. In addition, as recorded at the local health authority, 51 of the nonresponders had been tested positive elsewhere within 14 days before our invitation (not considered in the sensitivity analysis).

The arms differed substantially regarding some COVID-19 symptoms (Figure 4; Table S3 and Figures S2 and S3 in Multimedia Appendix 2). Approximately 20.60% (2054/9972) either worked with children or in the medical field (Table S4 in Multimedia Appendix 2).

Figure 4. Frequency of COVID-19–related symptoms among those who filled out the prescreening questionnaire (initially contacted individuals only; *P* values based on Pearson chi-square test). A1: individuals without mandatory prescreening; A2: households without mandatory prescreening; B1: individuals with mandatory prescreening; B2: households with mandatory prescreening.



Prevalence Estimates

The weighted prevalence estimates (initial cases only) differed considerably between combined arms A with 0.36% (95% CI 0.14%-0.59%) and B with 0.05% (95% CI 0.00%-0.10%). However, we did not detect a significant difference between the single arms (household members in A2 and B2 included). In A1, the prevalence was 0.32% (95% CI 0.06%-0.58%), and in A2, the prevalence was 0.35% (95% CI 0.09%-0.61%), with lower estimates in B1 and B2 (0.07%, 95% CI 0.00%-0.15% and 0.02%, 95% CI 0.00%-0.06%; Figure 5; Table S5 in Multimedia Appendix 2). We also compared our daily case estimates with the RKI data (Figure S4 in Multimedia Appendix 2). Assuming that the participants remained positive on average 14 days, the estimated prevalence range based on RKI data (0.23%-0.48%) stayed within the 95% CI 0.14-0.59 of our estimate for the combined arms A. Assuming an average duration of 10 days, the RKI-based prevalence range (0.16%-0.36%) was completely within the lower half of our 95% CI (Figure 6). Figure S5 in Multimedia Appendix 2 depicts the minimum number of positive cases that should have been detected in the trial, with the lower bound determined by the prevalence estimate based on RKI data (Figure 6). If we additionally considered the hotline callers that tested positive elsewhere, the prevalence increased further to 0.37% (95% CI 0.10%-0.65%) in A1, 0.35% (95% CI 0.09%-0.61%) in A2, 0.14% (95% CI 0.03%-0.26%) in B1, and 0.07% (95% CI 0.00%-0.15%) in B2 (Figure S6 and Table S5 in Multimedia Appendix 2).

Figure 5. Weighted prevalence estimates in study arms A1 (individuals without mandatory prescreening) and A2 (households without mandatory prescreening) combined, B1 (individuals with mandatory prescreening) and B2 (households with mandatory prescreening) combined, and each arm separately. Light blue indicates the fraction of positive cases among all the performed tests. For combined arms, only the initially contacted participants are included to avoid household bias in arms A2 and B2.

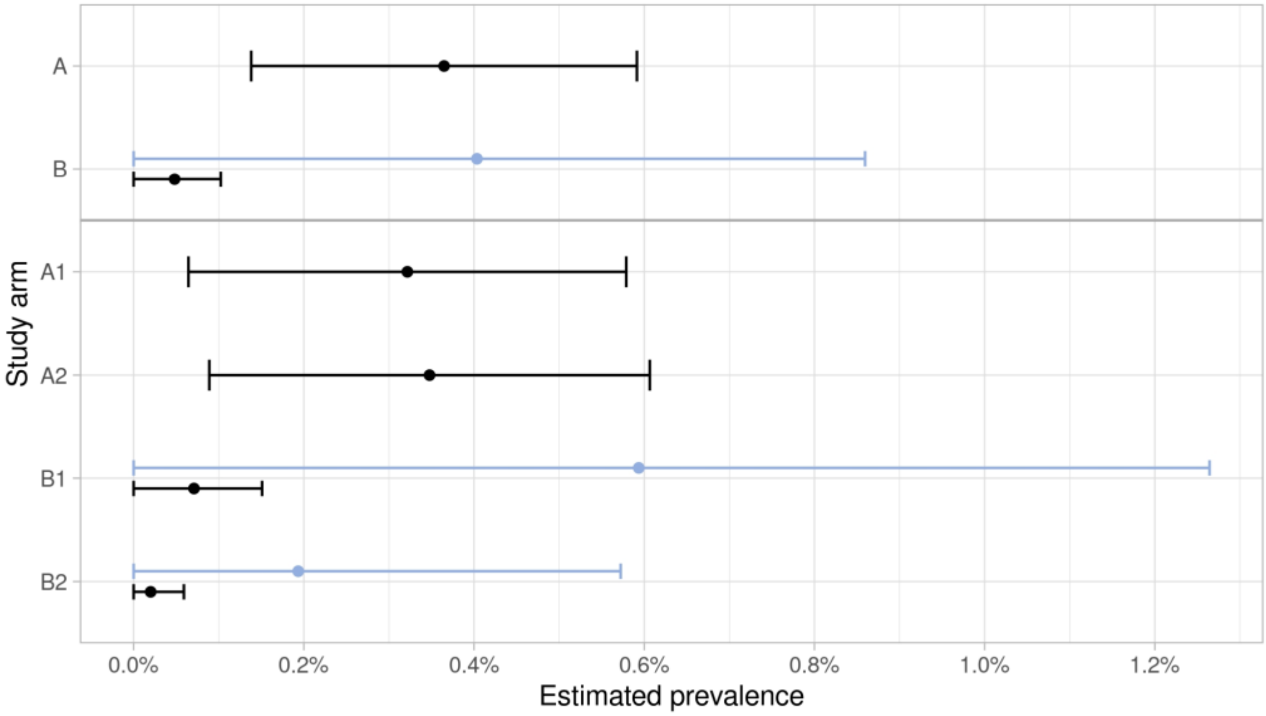
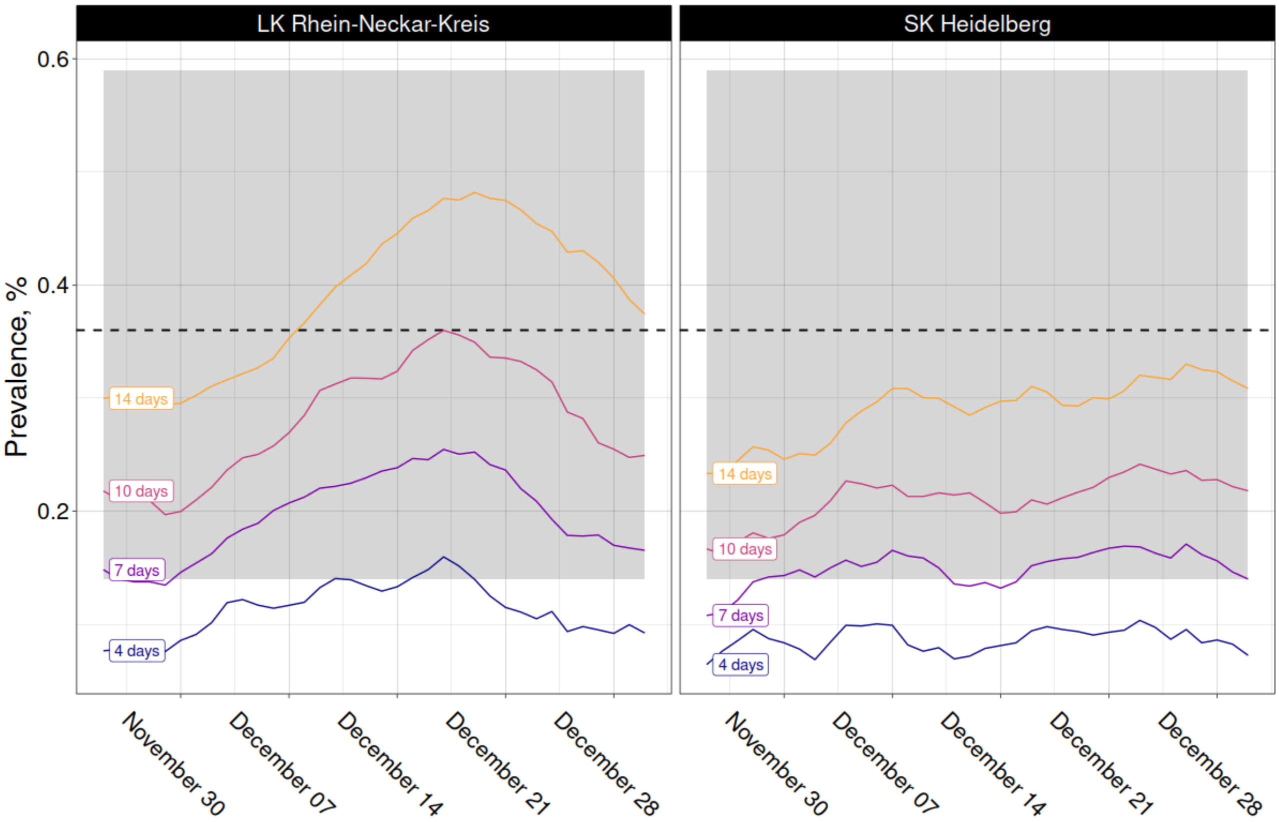


Figure 6. Variation in the prevalence estimate derived from Robert Koch Institute incidence data, depending on the assumed duration of polymerase chain reaction positivity (each prevalence curve is labeled by the corresponding duration in days; horizontal dashed line is the estimated prevalence based on the combined results from arms A1 and A2, with 95% CI indicated by the shaded area).



Discussion

Principal Findings

To the best of our knowledge, this is the only study to date that has simultaneously tested different approaches to active SARS-CoV-2 surveillance for the general public based on daily renewed random and population representative samples in a seamless chronology. In this paper, we focused on the feasibility and epidemiological outcomes of the trial.

Our study demonstrated that in general, the combination of postal mailing of sample collection kits and return of gargle liquid samples followed by high-sensitivity RT-LAMP analysis is a feasible way to conduct active SARS-CoV-2 surveillance in the general population without placing an additional burden on the capacity of PCR analysis and routine diagnostic testing [26]. Furthermore, our study suggests that the primarily symptom-based prescreening did not add benefit to case finding but rather complicated processes. Even if optimized, such a tool is unlikely to capture presymptomatic or asymptomatic cases and might have to be adapted with each new SARS-CoV-2 variant. The surveillance strategies could detect additional cases, some of them being asymptomatic, which were not captured by passive surveillance.

We were able to design and implement this complex trial in just 8 weeks by using innovative digital tools, agile development, and flat communication structures, which resulted in highly digitized processes and ensured that the study logistics were highly functional from the program outset. We are, therefore, convinced that a similar active pandemic surveillance system deployed outside a scientific study could be set up even faster, provided appropriate laboratory infrastructure and open-source software are held at the ready and standards are introduced [54].

Active SARS-CoV-2 surveillance should capture a sufficient proportion of symptomatic and asymptomatic cases to allow continuous estimation of the true prevalence trend with sufficient precision. However, despite the random drawing of potential participants, the nonresponders in our study were not missing at random, as selection owing to voluntary self-enrollment occurred, which was reflected in the younger and more male responders on average. Taking this into account, study arm A1 (random selection of individuals) was most likely to yield the best estimate of the true prevalence, although it was slightly biased because of voluntary self-enrollment. Investing in measures (eg, incentives and analytical corrections) to minimize the proportion of nonresponders may further reduce bias and increase accuracy.

The prevalence estimates derived from the RKI data based on passive surveillance constituted the lower bound of the prevalence estimates. As expected, the prevalence detected in our study tended to be higher, and the estimates of A1 and A2 were the most accurate. However, we cannot make a definitive statement regarding the most effective strategy because the number of positive cases was too low because of the actual sample size and the stage of the pandemic. It is important to note that this study was a research project not embedded in the public health system with data fusion and reconciliation. The inclusion of theoretically detectable cases who refused to participate but reported a recent infection to the study hotline upwardly adjusted the prevalence estimates. If other nonresponders who tested positive shortly before the invitation had been considered, the prevalence estimate would have been well above the official figures.

Study Participation

The overall participation rate of 36.57% (10,207/27,908) was lower than the envisaged 50%. First, the complexity of the information material necessary for a 4-arm trial might have been difficult to understand, thereby hindering a timely and appropriate participation [55]. Second, the different demands on the participants in a difficult overall situation (ie, lockdown) and an increasing pandemic fatigue before Christmas 2020 might have influenced the response [55]. One complicating aspect, the requested sampling before breakfast, seems to be less relevant based on new data (Multimedia Appendix 2).

Complementary qualitative data from approximately 80 responders and nonresponders shed light on these barriers. For example, some of the nonresponders thought they were selected because they participated in other studies or because acquaintances had to quarantine themselves. The study invitation also triggered fear of being infected in a few selected persons. The amount of information and perceived test complexity overwhelmed many [55].

Our study participation rates were similar to or exceeded those of other studies. In the United Kingdom, repeated community-based RT-PCR testing based on self-administered throat and nasal swabs estimated prevalence on a monthly basis. The number of invitation letters sent between May 2020 and December 2021 increased from 395,000 to 804,000, but the response rate (number of tests/letters) decreased from 30.5% to 12.1% and the PCR test return rate (number of tests performed/kits requested) varied between 67.2% and 78.9% [56]. During the same period as our study, 28.4% of the invitees requested sample collection kits, and the actual response rate was 22.4% [57]. We provided sample collection kits immediately along with the invitation letter in A1 and A2. In addition, we requested a gargle sample instead of a swab sample, which resulted in response rates of 41.2% and 36.2%, respectively. This may indicate that a low-barrier approach and an alternative to throat and nasal swabs could unlock further potential for participation.

Some methodological changes could have increased participation. Government support, media campaigns, and adding incentives could foster participation [58,59]. This might be particularly important if rapid antigen tests are widely available; therefore, adequate but not overly complicated information should educate participants about test quality and the purpose of active surveillance. Furthermore, the methodology was tailored to the ethical requirements for scientific trials, the mere extent of information materials might have deterred, and the label of a scientific study might have unsettled potential participants [55]. In contrast, routine implementation accompanied by a legal framework would only have to follow the general data protection laws, simplifying paperwork, and could be integrated into pandemic containment strategies, eliminating the chance for double sampling by linking data with passive surveillance systems.

Acceptability of Self-Sampling

In our study, mail-in gargle self-sampling proved to be feasible. The method was perceived as pleasant, allowed for high test accuracy in asymptomatic patients, and reduced the risk of spreading infections. Liquid gargle sampling was valued as more convenient than nasopharyngeal swabs, and although self-sampling was cumbersome for some older participants, many respondents found accurate self-sampling with gargling easier. The acceptability was mainly driven by communication (content of information, letters, support system, and rapidity of test results) and trust in the sampling or laboratory-based testing method [55]. Gargling a saline solution is a low-barrier sampling method applicable in almost all age groups [60] and suitable not only for LAMP but also for PCR analysis. Independently, the RT-LAMP analysis allows high-throughput testing with low costs and can be combined with nasopharyngeal, nasal, or throat swab samples without compromising sensitivity. Sample pooling can further multiply the testing capacity [21,22].

Cost-Effectiveness

From an epidemiological perspective, we identified A2 as the most cost-effective strategy, closely followed by A1, based on evidence from a parallel economic evaluation [37]. We postulate that surveillance strategies similar to A1 and A2 likely yield the best estimates and are most efficient. They are also logistically simpler than the strategies with a prescreening tool.

Limitations

This study had several limitations. First, the actual execution of the trial was limited to 3 weeks; time constraints limited adequate testing of the prescreening tool and impeded pilot studies. Second, random samples from most municipalities yielded a representative pool of potential participants; however, Heidelberg and its surrounding area have a higher socioeconomic level than an average German municipality. Voluntary participation likely correlates with gradients in wealth; thus, the participation rate could have been different elsewhere in Germany. As high socioeconomic strata are simultaneously less susceptible to SARS-CoV-2 infection and more likely to participate in studies, this possibly led to slightly underestimated prevalence estimates [61]. Third, participants in each study arm were expected to return samples no later than 1 day after receiving the sample collection kits. However, for some participants, the instructions in the invitation letter were not clear enough as this could not be tested in a pilot study owing to time constraints. Therefore, some participants collected samples on a day that was more convenient for them [55]. However, these untimely samples were most likely random, which may explain the lower sample return at the beginning and some late samples. Clear communication could yield timely sample return and thus more stable return numbers throughout the surveillance period. Finally, the calculation of the sample size and thus the accuracy of the prevalence were limited by the available financial resources and various uncertainties in the assumptions.

We expected to receive a similar number of gargle samples in each arm to enable systematic bias estimation. Because B1 uses a symptom-based prescreening that cannot capture asymptomatic cases, the prevalence should be systematically underestimated. However, B2 could also capture asymptomatic household members, and depending on the intrahousehold transmission rate, B2 could systematically overestimate the prevalence in certain scenarios. However, the average time between positive prescreening and LAMP result in B1 and B2 was 8.4 days; many of these participants had to be reminded to send the sample back. Hence, although an average duration of PCR positivity for 10 to 14 days seems reasonable, some people may have been positive when first receiving the study materials but cleared the infection before sampling [62]. In addition, the number of samples in B1 and B2 was considerably lower than that in A1 and A2 (1152 vs 4179 samples), which resulted in inaccurate prevalence estimates. Furthermore, the lower number of positive cases in B1 and B2 may have been because of the poor performance of the prescreening tool. The data sets used to train the random forest algorithm may not have been optimal to develop a sensitive discriminator. Including all household members in the prescreening could increase the chance of detecting symptomatic persons and might increase acceptance in parallel. However, this was omitted because of increased complexity and time constraints.

Active Versus Passive Surveillance

In Germany, the RKI's COVID-19 figures stem from passive surveillance, aiming to capture all positive test results (tested symptomatic persons and positively tested contact persons) to monitor the pandemic's progress. However, data are transmitted through multiple tiers from physicians and laboratories to the RKI via local and state health departments, causing time delays that affect the official reporting of case numbers. The RKI therefore applies the nowcasting method for estimation by means of assumptions and imputation of missing values [3]. Asymptomatic cases are also underrepresented, making it challenging to assess the situation accurately. In contrast, an active surveillance system would not claim completeness of cases but test population representative random samples that represent the actual occurrence of infections. Furthermore, it would not rely on bottom-up reporting but sample infections directly with less delays and yielding more accurate prevalence estimates. Decentralized active surveillance systems could also detect regional differences timely. Furthermore, a Bayesian sliding-window estimator could enable the prediction of the number of hospitalizations above a certain prevalence threshold.

Conclusions

The pandemic developed in a highly dynamic manner in most countries, with scarcely comparable trends. The succession of gradually increasing bouts of infection has been interrupted by periods of low infection activity. There is no reason to believe that future pandemics with similar hazard potential will evolve in a less chaotic manner; therefore, SARS-CoV-2 can be seen as a blueprint for future outbreaks. Therefore, rapidly deployable outbreak surveillance systems should be developed and put on standby.

Continuous active SARS-CoV-2 surveillance based on general population representative random samples of individuals or households and postal mailing of gargle sample collection kits for low-barrier home-based self-sampling, combined with a rapid, highly sensitive, and specific nucleic acid amplification test, is a feasible way to detect presymptomatic and asymptomatic cases without overburdening PCR capacities. However, additional measures are required to improve participation. Information material and consent processes should be tailored to real-life routine rather than scientific trials. In addition, the information material should be more concise and tested in pilot studies.

Acknowledgments

The authors would like to thank the local health authority Rhine-Neckar for their valuable support and the successful collaboration, with special thanks to Dr Britta Knorr, Dr Rainer Schwertz, Dr Andreas Welker, and Doreen Kuss. The authors would also like to thank Daniel Ürge and Victoria Witte for their contribution to the reverse transcription loop-mediated isothermal amplification analysis. Claudiu Grozea from the European Molecular Biology Laboratory produced the gargling and sampling demonstration video for the study website. In addition, the authors would like to thank the cohort of students who supported this study with the assembly and shipment of sample collection kits as well as sample processing, their service in the hotline, and support in data processing and analyses. Nicholas Kyei and Miriam Compton kindly proofread the manuscript. Finally, the authors would like to thank all the individuals and households of the Rhine-Neckar region in the state Baden-Württemberg in Germany who responded to their request and provided the survey data and gargle samples. The Cov Surv study was embedded in the nationwide research network "Applied Surveillance and Testing" (Bundesweites Forschungsnetz "Angewandte Surveillance und Testung," B-FAST) and the University Medicine's Network (Netzwerk Universitätsmedizin, NUM) for COVID-19 research in Germany. The trial was financed by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF). The laboratory robot was funded within the framework of the Excellence Strategy of the Federal and State Governments of Germany and by the Ministry of Science, Research and the Arts Baden-Württemberg. For the publication fee we acknowledge financial support by Deutsche Forschungsgemeinschaft (DFG) within the funding program "Open Access Publikationskosten" as well as by Heidelberg University. The funders had no role in the study design, data collection, data analysis, interpretation of data, or writing of the manuscript.

Data Availability

Anonymized data and a data dictionary will be made available in a public repository. Consent forms, invitation letters, and information material for the study are made available upon official request to the Network University Medicine for COVID-19 research in Germany. The study protocol has been published.

Authors' Contributions

CD, AD, and TB envisioned the study, developed the concept, and applied for the grant. SA and MK developed the sampling procedures and established the laboratory processes (collection and analyses of the samples) and the data processing. MDA and HTN were responsible for the cost-effectiveness measurement, with support from SB. AS and SM contributed the qualitative study component. MS accounted for the calculations with regard to the random sampling of addresses from the registration offices. LK together with LMH, CD, TR, and TJA designed, developed, and validated the prescreening tool. LB, RB, and M Marx set up the sample shipment logistics. M Meurer, KH, RB, DL, SA, and SO developed data processing pipelines and set up the server. KH developed the questionnaire scanner software. M Meurer, DL, and RB validated the sampling strategy and the sample analysis procedures. PS was responsible for confirmation of positive samples by real-time reverse transcription–polymerase chain reaction. SB registered the trial and was responsible for publishing the study protocol. AD, IM, and SO accessed and verified the data presented in this manuscript and performed the epidemiological analyses. CD performed data linkage with the health authority and supported the sensitivity analyses. AD, MDA, MK, CD, HTN, and SO drafted the manuscript. All authors contributed to the interpretation of the results and manuscript writing and read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT (Consolidated Standards of Reporting Trials) 2010 checklist.
[PDF File (Adobe PDF File), 106 KB - publichealth_v9i1e44204_app1.pdf]

Multimedia Appendix 2
Supplementary materials, containing details on trial operationalization and certain calculations as well as supplementary tables and figures.

[PDF File (Adobe PDF File), 776 KB - publichealth_v9i1e44204_app2.pdf]

Multimedia Appendix 3
Study presentation, web-based conference on February 8, 2021.
[PDF File (Adobe PDF File), 2324 KB - publichealth_v9i1e44204_app3.pdf]

References

1. Ampofo WK, Azziz-Baumgartner E, Bashir U, Cox NJ, Fasce R, Giovanni M, WHO Writing Group. Strengthening the influenza vaccine virus selection and development process: report of the 3rd WHO informal consultation for improving influenza vaccine virus selection held at WHO headquarters, Geneva, Switzerland, 1-3 April 2014. Vaccine 2015 Aug 26;33(36):4368-4382 [FREE Full text] [doi: 10.1016/j.vaccine.2015.06.090] [Medline: 26148877]
2. Petrova VN, Russell CA. The evolution of seasonal influenza viruses. Nat Rev Microbiol 2018 Jan;16(1):47-60. [doi: 10.1038/nrmicro.2017.118] [Medline: 29081496]
3. Hamouda O. Nowcasting und R-Schätzung: Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland. Robert Koch-Institut. 2020. URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/Nowcasting.html [accessed 2023-03-30]
4. Jones TC, Biele G, Mühlemann B, Veith T, Schneider J, Beheim-Schwarzbach J, et al. Estimating infectiousness throughout SARS-CoV-2 infection course. Science 2021 Jul 09;373(6551):eabi5273 [FREE Full text] [doi: 10.1126/science.abi5273] [Medline: 34035154]
5. Johansson MA, Quandelacy TM, Kada S, Prasad PV, Steele M, Brooks JT, et al. SARS-CoV-2 transmission from people without COVID-19 symptoms. JAMA Netw Open 2021 Jan 04;4(1):e2035057 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.35057] [Medline: 33410879]
6. Gao M, Yang L, Chen X, Deng Y, Yang S, Xu H, et al. A study on infectivity of asymptomatic SARS-CoV-2 carriers. Respir Med 2020 Aug;169:106026 [FREE Full text] [doi: 10.1016/j.rmed.2020.106026] [Medline: 32513410]
7. Sah P, Fitzpatrick MC, Zimmer CF, Abdollahi E, Juden-Kelly L, Moghadas SM, et al. Asymptomatic SARS-CoV-2 infection: a systematic review and meta-analysis. Proc Natl Acad Sci U S A 2021 Aug 24;118(34):e2109229118 [FREE Full text] [doi: 10.1073/pnas.2109229118] [Medline: 34376550]
8. Tabata S, Imai K, Kawano S, Ikeda M, Kodama T, Miyoshi K, et al. Clinical characteristics of COVID-19 in 104 people with SARS-CoV-2 infection on the diamond princess cruise ship: a retrospective analysis. Lancet Infect Dis 2020 Sep;20(9):1043-1050 [FREE Full text] [doi: 10.1016/S1473-3099(20)30482-5] [Medline: 32539988]
9. Nicholson G, Lehmann B, Padellini T, Pouwels KB, Jersakova R, Lomax J, et al. Improving local prevalence estimates of SARS-CoV-2 infections using a causal debiasing framework. Nat Microbiol 2022 Jan;7(1):97-107 [FREE Full text] [doi: 10.1038/s41564-021-01029-0] [Medline: 34972825]
10. Oude Munnink BB, Worp N, Nieuwenhuijse DF, Sikkema RS, Haagmans B, Fouchier RA, et al. The next phase of SARS-CoV-2 surveillance: real-time molecular epidemiology. Nat Med 2021 Sep;27(9):1518-1524. [doi: 10.1038/s41591-021-01472-w] [Medline: 34504335]
11. Subramanian SV, James KS. Use of the demographic and health survey framework as a population surveillance strategy for COVID-19. Lancet Glob Health 2020 Jul;8(7):e895 [FREE Full text] [doi: 10.1016/S2214-109X(20)30213-8] [Medline: 32530423]
12. Mina MJ, Andersen KG. COVID-19 testing: one size does not fit all. Science 2021 Jan 08;371(6525):126-127. [doi: 10.1126/science.abe9187] [Medline: 33414210]
13. Mercer T, Salit M. Testing at scale during the COVID-19 pandemic. Nat Rev Genet 2021 Jul;22(7):415-426 [FREE Full text] [doi: 10.1038/s41576-021-00360-w] [Medline: 33948037]
14. de Lusignan S, Lopez Bernal J, Zambon M, Akinyemi O, Amirthalingam G, Andrews N, et al. Emergence of a novel coronavirus (COVID-19): protocol for extending surveillance used by the royal college of general practitioners research and surveillance centre and public health England. JMIR Public Health Surveill 2020 Apr 02;6(2):e18606 [FREE Full text] [doi: 10.2196/18606] [Medline: 32240095]
15. Güemes A, Ray S, Aboumerhi K, Desjardins MR, Kvit A, Corrigan AE, et al. A syndromic surveillance tool to detect anomalous clusters of COVID-19 symptoms in the United States. Sci Rep 2021 Feb 25;11(1):4660 [FREE Full text] [doi: 10.1038/s41598-021-84145-5] [Medline: 33633250]
16. Mehl A, Bergey F, Cawley C, Gilsdorf A. Syndromic surveillance insights from a symptom assessment app before and during COVID-19 measures in Germany and the United Kingdom: results from repeated cross-sectional analyses. JMIR Mhealth Uhealth 2020 Oct 09;8(10):e21364 [FREE Full text] [doi: 10.2196/21364] [Medline: 32997640]
17. Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J Infect Public Health 2020 Nov;13(11):1630-1638 [FREE Full text] [doi: 10.1016/j.jiph.2020.07.019] [Medline: 32855090]
18. Hilborne LH, Wagner Z, Cabreros I, Brook RH. Linking statistics with testing policy to manage COVID-19 in the community. Am J Clin Pathol 2020 Jul 07;154(2):142-148 [FREE Full text] [doi: 10.1093/ajcp/aqaa099] [Medline: 32520340]

19. Riley S, Wang H, Eales O, Haw D, Walters CE, Ainslie KE, The COVID-19 Genomics UK (COG-UK) Consortium, et al. REACT-1 round 12 report: resurgence of SARS-CoV-2 infections in England associated with increased frequency of the Delta variant. medRxiv Preprint posted online on June 21, 2021 [FREE Full text] [doi: 10.1101/2021.06.17.21259103]
20. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 2020 Jan;25(3):2000045 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.3.2000045] [Medline: 31992387]
21. Deckert A, Bärnighausen T, Kyei NN. Simulation of pooled-sample analysis strategies for COVID-19 mass testing. Bull World Health Organ 2020 Sep 01;98(9):590-598 [FREE Full text] [doi: 10.2471/BLT.20.257188] [Medline: 33012859]
22. Mutesa L, Ndishimye P, Butera Y, Souopgui J, Uwineza A, Rutayisire R, et al. A pooled testing strategy for identifying SARS-CoV-2 at low prevalence. Nature 2021 Jan;589(7841):276-280 [FREE Full text] [doi: 10.1038/s41586-020-2885-5] [Medline: 33086375]
23. Kellner MJ, Ross JJ, Schnabl J, Dekens MP, Matl M, Heinen R, et al. A rapid, highly sensitive and open-access SARS-CoV-2 detection assay for laboratory and home testing. Front Mol Biosci 2022 Apr 01;9:801309 [FREE Full text] [doi: 10.3389/fmolb.2022.801309] [Medline: 35433827]
24. Zhang Y, Ren G, Buss J, Barry AJ, Patton GC, Tanner NA. Enhancing colorimetric loop-mediated isothermal amplification speed and sensitivity with guanidine chloride. Biotechniques 2020 Sep;69(3):178-185 [FREE Full text] [doi: 10.2144/btn-2020-0078] [Medline: 32635743]
25. Dao Thi VL, Herbst K, Boerner K, Meurer M, Kremer LP, Kirrmaier D, et al. A colorimetric RT-LAMP assay and LAMP-sequencing for detecting SARS-CoV-2 RNA in clinical samples. Sci Transl Med 2020 Aug 12;12(556):eabc7075 [FREE Full text] [doi: 10.1126/scitranslmed.abc7075] [Medline: 32719001]
26. Lou D, Meurer M, Ovchinnikova S, Burk R, Denzler A, Herbst K, et al. Scalable RT-LAMP-based SARS-CoV-2 testing for infection surveillance with applications in pandemic preparedness. EMBO Rep 2023 May 04;24(5):e57162 [FREE Full text] [doi: 10.15252/embr.202357162] [Medline: 36951170]
27. Beyene GT, Alemu F, Kebede ES, Alemayehu DH, Seyoum T, Tefera DA, et al. Saliva is superior over nasopharyngeal swab for detecting SARS-CoV2 in COVID-19 patients. Sci Rep 2021 Nov 22;11(1):22640 [FREE Full text] [doi: 10.1038/s41598-021-02097-2] [Medline: 34811429]
28. Goldfarb DM, Tilley P, Al-Rawahi GN, Srigley JA, Ford G, Pedersen H, et al. Self-collected saline gargle samples as an alternative to health care worker-collected nasopharyngeal swabs for COVID-19 diagnosis in outpatients. J Clin Microbiol 2021 Mar 19;59(4):e02427-20 [FREE Full text] [doi: 10.1128/JCM.02427-20] [Medline: 33514627]
29. Malecki M, Lüsebrink J, Teves S, Wendel AF. Pharynx gargle samples are suitable for SARS-CoV-2 diagnostic use and save personal protective equipment and swabs. Infect Control Hosp Epidemiol 2021 Feb;42(2):248-249 [FREE Full text] [doi: 10.1017/ice.2020.229] [Medline: 32389132]
30. Lee RA, Herigon JC, Benedetti A, Pollock NR, Denkinger CM. Performance of saliva, oropharyngeal swabs, and nasal swabs for SARS-CoV-2 molecular detection: a systematic review and meta-analysis. J Clin Microbiol 2021 Apr 20;59(5):e02881-20 [FREE Full text] [doi: 10.1128/JCM.02881-20] [Medline: 33504593]
31. Nasiri MJ, Haddadi S, Tahvildari A, Farsi Y, Arbabi M, Hasanzadeh S, et al. COVID-19 clinical characteristics, and sex-specific risk of mortality: systematic review and meta-analysis. Front Med (Lausanne) 2020 Jul 21;7:459 [FREE Full text] [doi: 10.3389/fmed.2020.00459] [Medline: 32793620]
32. Ibrahimi N, Delaunay-Moisan A, Hill C, Le Teuff G, Rupprecht J, Thuret J, et al. Screening for SARS-CoV-2 by RT-PCR: saliva or nasopharyngeal swab? Rapid review and meta-analysis. PLoS One 2021 Jun 10;16(6):e0253007 [FREE Full text] [doi: 10.1371/journal.pone.0253007] [Medline: 34111196]
33. Ott IM, Strine MS, Watkins AE, Boot M, Kalinich CC, Harden CA, Yale IMPACT Research team3. Stability of SARS-CoV-2 RNA in nonsupplemented saliva. Emerg Infect Dis 2021 Apr;27(4):1146-1150 [FREE Full text] [doi: 10.3201/eid2704.204199] [Medline: 33754989]
34. Pouwels KB, House T, Pritchard E, Robotham JV, Birrell PJ, Gelman A, COVID-19 Infection Survey Team. Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS coronavirus infection survey. Lancet Public Health 2021 Jan;6(1):e30-e38 [FREE Full text] [doi: 10.1016/S2468-2667(20)30282-6] [Medline: 33308423]
35. Aziz NA, Corman VM, Echterhoff AK, Müller MA, Richter A, Schmandke A, et al. Seroprevalence and correlates of SARS-CoV-2 neutralizing antibodies from a population-based study in Bonn, Germany. Nat Commun 2021 Apr 09;12(1):2117 [FREE Full text] [doi: 10.1038/s41467-021-22351-5] [Medline: 33837204]
36. Klein C, Borsche M, Balck A, Föh B, Rahmöller J, Peters E, et al. One-year surveillance of SARS-CoV-2 transmission of the ELISA cohort: a model for population-based monitoring of infection risk. Sci Adv 2022 Apr 15;8(15):eabm5016 [FREE Full text] [doi: 10.1126/sciadv.abm5016] [Medline: 35427158]
37. Nguyen HT, Denkinger CM, Brenner S, Koeppel L, Brugnara L, Burk R, et al. Cost and cost-effectiveness of four different SARS-CoV-2 active surveillance strategies: evidence from a randomised control trial in Germany. Eur J Health Econ 2023 Jan 19:1-15 [FREE Full text] [doi: 10.1007/s10198-022-01561-8] [Medline: 36656403]
38. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. J Pharmacol Pharmacother 2010 Jul;1(2):100-107 [FREE Full text] [doi: 10.4103/0976-500X.72352] [Medline: 21350618]

39. Deckert A, Anders S, De Allegri M, Nguyen HT, Souares A, McMahon S, et al. Effectiveness and cost-effectiveness of four different strategies for SARS-CoV-2 surveillance in the general population (CoV-Surv Study): study protocol for a two-factorial randomized controlled multi-arm trial with cluster sampling. Trials 2021 Sep 26;22(1):656 [FREE Full text] [doi: 10.1186/s13063-021-05619-5] [Medline: 34565421]
40. Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. J Am Stat Assoc 1952 Dec;47(260):663-685. [doi: 10.1080/01621459.1952.10483446]
41. Brewer K. Combined Survey Sampling Inference: Weighing Basu's Elephants. Oxford, UK: Oxford University Press; 2002.
42. Häder S. Stichproben in der Praxis (GESIS survey guidelines). Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften. 2015. URL: https://www.gesis.org/gesis-survey-guidelines/statistics/stichproben/stichproben-in-der-praxis [accessed 2023-07-17]
43. Brynildsrud O. COVID-19 prevalence estimation by random sampling in population - optimal sample pooling under varying assumptions about true prevalence. BMC Med Res Methodol 2020 Jul 23;20(1):196 [FREE Full text] [doi: 10.1186/s12874-020-01081-0] [Medline: 32703158]
44. Wyllie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, et al. Saliva or nasopharyngeal swab specimens for detection of SARS-CoV-2. N Engl J Med 2020 Sep 24;383(13):1283-1286 [FREE Full text] [doi: 10.1056/NEJMc2016359] [Medline: 32857487]
45. Nascimento JACJ, Santos AM, Oliveira AMS, Guimarães AG, Quintans-Júnior LJ, Coutinho HDM, et al. Trends in MERS-CoV, SARS-CoV, and SARS-CoV-2 (COVID-19) Diagnosis Strategies: A Patent Review. Front Public Health 2020 Oct 27;8:563095 [FREE Full text] [doi: 10.3389/fpubh.2020.563095] [Medline: 33194964]
46. CSV mit den aktuellen Covid-19 Infektionen pro Tag (Zeitreihe). Robert Koch-Institut. 2021. URL: https://www.arcgis.com/home/item.html?id=f10774f1c63e40168479a1feb6c7ca74 [accessed 2023-07-17]
47. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. Lancet Microbe 2021 Jan;2(1):e13-e22 [FREE Full text] [doi: 10.1016/S2666-5247(20)30172-5] [Medline: 33521734]
48. Levenshtein VI. Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady 1966:707-710 [FREE Full text]
49. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 2016 Jun 28;353:i2016. [doi: 10.1136/bmj.i2016] [Medline: 27353417]
50. Herrmann Y, Starck T, Brindl N, Kitchen PJ, Rädeker L, Sebastian J, et al. Description and analysis of representative COVID-19 cases-a retrospective cohort study. PLoS One 2021 Jul 30;16(7):e0255513 [FREE Full text] [doi: 10.1371/journal.pone.0255513] [Medline: 34329364]
51. DasGupta A, Cai TT, Brown LD. Interval Estimation for a Binomial Proportion. Statist Sci 2001 May;16(2):101-133 [FREE Full text] [doi: 10.1214/ss/1009213286]
52. Zahl der Privathaushalte und durchschnittliche Haushaltsgröße in Deutschland (1991-2035). Bundesinstitut für Bevölkerungsforschung. 2017. URL: https://www.bib.bund.de/DE/Fakten/Fakt/L50-Privathaushalte-Haushaltsgroesse-1991-Vorausberechnung.html [accessed 2023-03-30]
53. Durchschnittsalter der Bevölkerung in Deutschland nach Staatsangehörigkeit am 31. Statista. 2019. URL: https://de.statista.com/statistik/daten/studie/723069/umfrage/durchschnittsalter-der-bevoelkerung-in-deutschland-nach-staatsangehoerigkeit/ [accessed 2023-03-30]
54. Mercer T, Almond N, Crone MA, Chain PS, Deshpande A, Eveleigh D, et al. The coronavirus standards working group's roadmap for improved population testing. Nat Biotechnol 2022 Nov;40(11):1563-1568 [FREE Full text] [doi: 10.1038/s41587-022-01538-1] [Medline: 36323792]
55. Röhr F, Uellner F, Deckert A, Anders S, Burk R, Knop M, et al. From disgusting and complicated to simple and brilliant: implementation perspectives and lessons learned from users and rejectors of mail-in SARS-CoV-2 gargle tests. Front Public Health 2022 Jan 05;10:1024525 [FREE Full text] [doi: 10.3389/fpubh.2022.1024525] [Medline: 36684995]
56. REal-time Assessment of Community Transmission (REACT) - study summary of sample extraction and fieldwork dates, and response rates, by round. Imperial College London. 2020. URL: https://www.imperial.ac.uk/media/imperial-college/institute-of-global-health-innovation/REACT1-R19-REACT2-R6---Fieldwork-info-table.pdf [accessed 2023-03-30]
57. Riley S, Walters CE, Wang H, Eales O, Ainslie KC, Atchinson C, et al. REACT-1 round 7 updated report: regional heterogeneity in changes in prevalence of SARS-CoV-2 infection during the second national COVID-19 lockdown in England. medRxiv Preprint posted online on December 16, 2020 [FREE Full text] [doi: 10.1101/2020.12.15.20248244]
58. Stead M, Angus K, Langley T, Katikireddi SV, Hinds K, Hilton S, et al. Mass media to communicate public health messages in six health topic areas: a systematic review and other reviews of the evidence. National Health Service. 2019 Apr. URL: https://pubmed.ncbi.nlm.nih.gov/31046212/ [accessed 2023-07-04]
59. Hofmann B, Stanak M. Nudging in screening: literature review and ethical guidance. Patient Educ Couns 2018 Sep;101(9):1561-1569. [doi: 10.1016/j.pec.2018.03.021] [Medline: 29657111]

60. Gertler M, Krause E, van Loon W, Krug N, Kausch F, Rohardt C, et al. Self-collected oral, nasal and saliva samples yield sensitivity comparable to professionally collected oro-nasopharyngeal swabs in SARS-CoV-2 diagnosis among symptomatic outpatients. Int J Infect Dis 2021 Sep;110:261-266 [FREE Full text] [doi: 10.1016/j.ijid.2021.07.047] [Medline: 34302962]
61. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
62. Wang Y, Zhang L, Sang L, Ye F, Ruan S, Zhong B, et al. Kinetics of viral load and antibody response in relation to COVID-19 severity. J Clin Invest 2020 Oct 01;130(10):5235-5244 [FREE Full text] [doi: 10.1172/JCI138759] [Medline: 32634129]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
RKI: Robert Koch Institute
RT-LAMP: reverse transcription loop-mediated isothermal amplification
RT-PCR: real-time reverse transcription–polymerase chain reaction

Edited by A Mavragani, T Sanchez; submitted 14.11.22; peer-reviewed by S Kuhlmann Berenzon, G Dalla Costa; comments to author 15.02.23; revised version received 30.03.23; accepted 24.05.23; published 17.08.23.

Please cite as:
Deckert A, Anders S, Morales I, De Allegri M, Nguyen HT, Souares A, McMahon S, Meurer M, Burk R, Lou D, Brugnara L, Sand M, Koeppel L, Maier-Hein L, Ross T, Adler TJ, Brenner S, Dyer C, Herbst K, Ovchinnikova S, Marx M, Schnitzler P, Knop M, Bärnighausen T, Denkinger CM
Comparison of Four Active SARS-CoV-2 Surveillance Strategies in Representative Population Sample Points: Two-Factor Factorial Randomized Controlled Trial
JMIR Public Health Surveill 2023;9:e44204
URL: https://publichealth.jmir.org/2023/1/e44204
doi:10.2196/44204
PMID:37235704

Original Paper

Wastewater Surveillance for SARS-CoV-2 at Long-Term Care Facilities: Mixed Methods Evaluation

James W Keck[1], MPH, MD; Jess Lindner[2], BA; Matthew Liversedge[3], BA; Blazan Mijatovic[3], MS; Cullen Olsson[3], MS; William Strike[4], BS; Anni Noble[5], BS; Reuben Adatorwovor[6], PhD; Parker Lacy[7], BS; Ted Smith[8], PhD; Scott M Berry[4,5], MBA, PhD

[1]Department of Family & Community Medicine, University of Kentucky, Lexington, KY, United States
[2]College of Medicine - Northern Kentucky Campus, University of Kentucky, Highland Heights, KY, United States
[3]Department of Family and Community Medicine, University of Kentucky, Lexington, KY, United States
[4]Department of Biomedical Engineering, University of Kentucky, Lexington, KY, United States
[5]Department of Mechanical Engineering, University of Kentucky, Lexington, KY, United States
[6]Department of Biostatistics, University of Kentucky, Lexington, KY, United States
[7]Trilogy Health Services, Louisville, KY, United States
[8]Christina Lee Brown Envirome Institute, University of Louisville, Louisville, KY, United States

Corresponding Author:
James W Keck, MPH, MD
Department of Family & Community Medicine
University of Kentucky
2195 Harrodsburg Rd
Suite 125
Lexington, KY, 40504
United States
Phone: 1 907 729 2924
Email: jameswkeck@gmail.com

Abstract

Background: Wastewater surveillance provided early indication of COVID-19 in US municipalities. Residents of long-term care facilities (LTCFs) experienced disproportionate morbidity and mortality early in the COVID-19 pandemic. We implemented LTCF building-level wastewater surveillance for SARS-CoV-2 at 6 facilities in Kentucky to provide early warning of SARS-CoV-2 in populations considered vulnerable.

Objective: This study aims to evaluate the performance of wastewater surveillance for SARS-CoV-2 at LTCFs in Kentucky.

Methods: We conducted a mixed methods evaluation of wastewater surveillance following Centers for Disease Control and Prevention (CDC) guidelines for evaluating public health surveillance systems. Evaluation steps in the CDC guidelines were engaging stakeholders, describing the surveillance system, focusing the evaluation design, gathering credible evidence, and generating conclusions and recommendations. We purposively recruited stakeholders for semistructured interviews and undertook thematic content analysis of interview data. We integrated wastewater, clinical testing, and process data to characterize or calculate 7 surveillance system performance attributes (simplicity, flexibility, data quality, sensitivity and positive predictive value [PPV], timeliness, representativeness, and stability).

Results: We conducted 8 stakeholder interviews. The surveillance system collected wastewater samples (N=811) 2 to 4 times weekly at 6 LTCFs in Kentucky from March 2021 to February 2022. Synthesis of credible evidence indicated variable surveillance performance. Regarding simplicity, surveillance implementation required moderate human resource and technical capacity. Regarding flexibility, the system efficiently adjusted surveillance frequency and demonstrated the ability to detect additional pathogens of interest. Regarding data quality, software identified errors in wastewater sample metadata entry (110/3120, 3.53% of fields), technicians identified polymerase chain reaction data issues (140/7734, 1.81% of reactions), and staff entered all data corrections into a log. Regarding sensitivity and PPV, using routine LTCF SARS-CoV-2 clinical testing results as the *gold standard*, a wastewater SARS-CoV-2 signal of >0 RNA copies/mL was 30.6% (95% CI 24.4%-36.8%) sensitive and 79.7% (95% CI 76.4%-82.9%) specific for a positive clinical test at the LTCF. The PPV of the wastewater signal was 34.8% (95% CI 27.9%-41.7%) at >0 RNA copies/mL and increased to 75% (95% CI 60%-90%) at >250 copies/mL. Regarding timeliness, stakeholders received surveillance data 24 to 72 hours after sample collection, with delayed reporting because of the lack of weekend laboratory staff. Regarding representativeness, stakeholders identified challenges delineating the population contributing to LTCF wastewater because of visitors, unknown staff toileting habits, and the use of adult briefs by some residents preventing their waste from entering the sewer system. Regarding stability, the reoccurring cost to conduct 1 day of wastewater surveillance at 1 facility was approximately US $144.50, which included transportation, labor, and materials expenses.

Conclusions: The LTCF wastewater surveillance system demonstrated mixed performance per CDC criteria. Stakeholders found surveillance feasible and expressed optimism regarding its potential while also recognizing challenges in interpreting and acting on surveillance data.

(JMIR Public Health Surveill 2023;9:e44657) doi:10.2196/44657

KEYWORDS

wastewater surveillance; wastewater-based epidemiology; evaluation; long-term care facility; COVID-19; SARS-CoV-2

Introduction

Background

Two years into the COVID-19 pandemic, myriad public health challenges remain. Preventing disease in populations considered vulnerable is among the most persistent and pressing issues. Older adults, particularly those living in long-term care facilities (LTCFs), experienced disproportionate morbidity and mortality during the COVID-19 pandemic [1]. An estimated 23% of the COVID-19 deaths in the United States occurred among LTCF residents or staff as of early 2022 [2]. A combination of age-related relative immunosuppression, high prevalence of comorbid conditions, and increased exposure in congregate living settings contributes to a greater burden of infection and mortality among LTCF residents [1].

Early efforts in the pandemic to prevent and mitigate COVID-19 at LTCFs focused on limiting exposure (physical distancing and personal protective equipment) and symptom screening to identify potential cases [3]. There was limited availability of clinical testing; for example, in June 2020, all LTCF staff and residents in Kentucky underwent a single comprehensive round of SARS-CoV-2 testing, and, at the time, it was unclear whether more frequent clinical testing was feasible. We identified wastewater surveillance for SARS-CoV-2 as a promising approach to direct the limited clinical testing resources and potentially mitigate the spread of infection at LTCFs.

Wastewater surveillance analyzes samples of wastewater for the presence of disease markers, such as SARS-CoV-2 RNA [4]. Many individuals who have been infected with SARS-CoV-2 shed the virus in their stool [5], and there is evidence that wastewater viral trends precede clinical case detection by several days [6]. Monitoring wastewater for SARS-CoV-2 RNA can identify previously unknown cases in a population [7] and guide clinical testing and infection prevention activities to mitigate new infections [7,8].

Wastewater surveillance has potential advantages over clinical testing. Foremost, it measures disease presence in a population via a single sample, making it more efficient than individual surveillance [9]. It also has the potential to detect SARS-CoV-2 from asymptomatic individuals who might not otherwise seek clinical testing [6]. It is noninvasive, which means that it may be more acceptable to individuals than repeated clinical testing [10]. Wastewater surveillance may provide more timely information on disease presence and trends than clinical testing owing to the earlier detection of SARS-CoV-2 shed by asymptomatic or presymptomatic individuals in the population [6]. However, there is little evidence to guide the implementation of wastewater surveillance at LTCFs.

Objectives

We share our experience of implementing wastewater surveillance for SARS-CoV-2 at 6 LTCFs in Kentucky during the second year of the COVID-19 pandemic. We evaluated the LTCF wastewater surveillance according to the framework put forth by the Centers for Disease Control and Prevention (CDC) Guidelines Working Group [11] for evaluating public health surveillance systems.

Methods

Evaluation Approach and Manuscript Organization

The surveillance evaluation followed CDC guidelines for evaluating public health surveillance systems [11]. We followed the guidelines' evaluation steps, including (1) engaging stakeholders, (2) describing the surveillance system, (3) focusing the evaluation design, (4) gathering credible evidence, and (5) generating conclusions and recommendations. We describe step 1 (engaging stakeholders) and how we gathered and analyzed credible evidence (steps 4 and 5) in this *Methods* section. In the *Results* section, we describe the surveillance system (step 2) and share credible evidence (step 4). In accordance with CDC surveillance evaluation guidelines, we gathered evidence on 7 general performance attributes of surveillance systems (simplicity, flexibility, data quality, sensitivity and positive predictive value [PPV], timeliness, representativeness, and stability). We incorporated additional concerns that arose during stakeholder interviews into the evaluation. We provide conclusions and recommendations (step 5) in the *Discussion* section.

Stakeholder Interviews and Qualitative Analysis

The CDC guidelines identified stakeholder groups potentially relevant to surveillance system evaluation to include public health practitioners, health care providers, data providers and users, representatives of affected communities, governments, and professional and private organizations. The study team identified the following stakeholder groups to focus the evaluation design (step 3) and provide credible evidence (step

4) relevant for the evaluation: LTCF management (encompassing the categories of data provider and user, health care provider, representative of affected community, and private organization), public health practitioners, and wastewater testing field and laboratory staff (another category of data provider). We purposively recruited individuals from these groups based on their involvement with LTCF wastewater surveillance and jurisdictional relationship (eg, local public health department) for semistructured interviews tailored to stakeholder roles. Interview questions asked about LTCF wastewater surveillance usefulness, resources, challenges, communication, concerns, and actions. The semistructured interview guides are presented in Multimedia Appendix 1. Author JL, a medical student not involved in implementing the wastewater surveillance system, conducted the interviews remotely using videoconferencing. Stakeholder interviewees provided written informed consent before the interview; they did not receive compensation for their participation. Each stakeholder was interviewed once, and the interviews lasted 18 to 35 minutes. We recorded and transcribed the interviews for qualitative analysis. We analyzed interview content using the rigorous and accelerated data reduction approach for qualitative analysis [12]. The rigorous and accelerated data reduction method involved identifying key points and moving pertinent quotes or thoughts into a central spreadsheet. After each interview transcript underwent this process, we identified common themes and moved these data into a secondary spreadsheet. The data went through several rounds of this reduction process to isolate the most pertinent ideas and generate representative quotations. We used information from the stakeholder interviews to focus the evaluation design (step 3).

Other Credible Data Sources

Credible evidence (step 4 of the evaluation process) came from several sources. The project's LTCF partner shared deidentified staff and resident clinical testing results from the 6 facilities during the wastewater surveillance period. LTCF clinical testing occurred per state, CDC, and Centers for Medicare and Medicaid Services guidance and varied from twice weekly to monthly screening of unvaccinated staff per local disease transmission levels. Wastewater testing data came from the study team. The stakeholder interviews provided qualitative data. Additional data came from the project team's administrative records.

Wastewater Surveillance Performance

We calculated the sensitivity and specificity of wastewater surveillance by comparing wastewater positivity with clinical test positivity for SARS-CoV-2 at wastewater thresholds ranging from 0 to 250 copies/mL of SARS-CoV-2 RNA. The study team obtained wastewater measurements using a method developed by our team and described in the *Surveillance System Description* subsection of the *Results* section. We compared the wastewater data with clinical test results during the week that followed a wastewater measurement in a 2×2 contingency table (true positives, true negatives, false positives, and false negatives) at multiple wastewater SARS-CoV-2 threshold values. We then estimated the sensitivity and specificity of wastewater surveillance at each threshold and plotted a receiver operating characteristic curve. To evaluate whether wastewater testing identified SARS-CoV-2 at LTCFs earlier than routine clinical screening, we conducted a lead-lag time correlational analysis. We estimated the correlation between the average wastewater RNA concentration and the number of known SARS-CoV-2 infections at a specific LTCF by shifting the wastewater result 1 to 7 days before and after the clinical test collection date. We calculated the Kendall rank correlation coefficient because this estimator is conservative compared with the Pearson and Spearman correlation coefficients. We used Microsoft Excel for descriptive analyses of administrative data and SAS 9.4 (SAS Institute Inc) to estimate the wastewater surveillance performance metrics of sensitivity and PPV, and the correlation of the wastewater RNA signal with clinical test data.

Ethics Approval

The University of Kentucky Institutional Review Board reviewed and approved the study protocol (62384).

Results

Surveillance System Description

Overview

The study team collaborated with Trilogy Health Services, an LTCF organization that manages >100 facilities across the Midwest in the United States. The study team identified 6 LTCFs in Lexington and Louisville, Kentucky, for wastewater surveillance. Proximity to the University of Kentucky campus and the ability to access facility-specific wastewater effluent guided facility site selection. The 6 participating LTCFs each had 67 to 160 residents and 76 to 117 staff. Wastewater sampling began on March 19, 2021, and ended in Louisville on December 16, 2021, and in Lexington on February 23, 2022.

Wastewater Collection

The study team visited each LTCF to identify wastewater effluent access points with facility management. For each facility, there was manhole access to a sewer pipe that contained the entirety of the facility's wastewater effluent and did not contain effluent from neighboring buildings. Field technicians obtained LTCF wastewater effluent samples using autosamplers suspended under the manhole covers that collected 100 mL of wastewater every 20 minutes during a 24-hour period (ie, a composite sample). Rechargeable batteries provided power for the autosamplers. Ice packed around the autosampler jug refrigerated the composite sample to minimize RNA degradation. The technicians collected three to four 24-hour composite samples each week from the Lexington LTCFs and 2 to 3 samples each week from the Louisville LTCFs. Field technicians recorded wastewater composite sample collection date, location, volume, and temperature using tablet computers and a custom REDCap (Research Electronic Data Capture; Vanderbilt University) database. The field technicians transported 250 mL from each 24-hour composite sample to the University of Kentucky in Lexington for analysis. The team collected 811 composite wastewater samples across the 6 facilities during the study period.

Laboratory Analysis

Laboratory analysis of the wastewater included (1) heat-mediated viral lysis, (2) nucleic acid extraction using paramagnetic particles with exclusion-based sample extraction [13], and (3) SARS-CoV-2 RNA quantification using CDC-recommended N1 primers and real-time qualitative polymerase chain reaction (RT-qPCR) analysis. Laboratory staff processed 8 aliquots from each sample and reported the average SARS-CoV-2 RNA concentration across the 8 aliquots. Strike et al [14] provide a detailed description of the laboratory method. Laboratory quality control measures included negative RT-qPCR controls, positive RT-qPCR controls, spiked SARS-CoV-2 samples to assess RNA extraction efficiency, and a visual inspection of RT-qPCR readouts for concurrence with automated cycle threshold reads. In addition, the laboratory measured crAssphage DNA concentrations in 2 aliquots from each wastewater composite sample as an indicator of the sample's fecal load using the same general laboratory method as for SARS-CoV-2 quantification. crAssphage is a human gut bacteriophage that is ubiquitous in human stool at high concentrations.

Data Architecture and Communication

The study team developed a custom computer program to join the wastewater sample data stored in the REDCap database with the RT-qPCR wastewater RNA data, which resulted in a text file. The program incorporated multiple data validity checks and generated an error log that triggered study team members to investigate and rectify data quality issues. The team created a web-based password-protected data visualization dashboard that provided up-to-date results of wastewater surveillance to Trilogy management. The study team also shared wastewater surveillance results with Trilogy facility leadership via email and telephone. Trilogy decided how to respond to the wastewater surveillance data, including any enhanced clinical testing of staff or residents.

Wastewater Surveillance Performance

In accordance with CDC surveillance evaluation guidelines, we present credible evidence from stakeholder interviews (N=8), administrative records, wastewater analysis, and clinical testing results to address 7 general performance attributes of public health surveillance systems.

Simplicity: How Cumbersome Was the Process of Collecting Samples, Processing Them, and Sharing Surveillance Data With Stakeholders?

Testing wastewater samples 3 to 4 times per week at 6 facilities across 2 municipalities required 1 full-time laboratory technician, 1 full-time field technician, and 1 part-time field technician. Several part-time student laboratory assistants worked ≤10 hours a week to conduct wastewater laboratory analyses under the supervision of the full-time laboratory technician. The field technicians had no relevant experience at surveillance onset and acquired wastewater collection skills through apprenticeship during several field site visits.

The method of communicating wastewater results evolved over the study. Data sharing transitioned from emails and telephone calls to a web-based data dashboard, complemented by emails and telephone calls for results with perceived urgency. The dashboard displayed levels of SARS-CoV-2 RNA in wastewater by LTCF location and included clinical testing data from Trilogy's public-facing COVID-19 dashboard.

Flexibility: Did Wastewater Surveillance Adapt to the Changing Realities of the Pandemic, and If So, How Much Effort Was Needed, and How Successful Was the Transition?

Wastewater surveillance demonstrated flexibility in sampling frequency and duration, population under surveillance, and pathogens targeted. Initially, the field team collected wastewater samples 3 times weekly (Lexington) and twice weekly (Louisville). The team increased the sampling frequency to 4 times weekly (Lexington) and 3 times weekly (Louisville) to optimize the timely identification of new infections in the LTCF population. In addition, the study team reported that they efficiently resumed wastewater surveillance in January 2022 (after planned completion in December 2021) in Lexington because of concern for infections related to the SARS-CoV-2 Omicron variant in the LTCF population.

Although the focus of the wastewater surveillance system was SARS-CoV-2, the study team piloted testing LTCF wastewater for *Clostridioides difficile*, a pernicious colonic bacterium that can cause severe gastrointestinal illness. The testing of LTCF wastewater identified pathogenic (toxin producing) *C difficile* DNA, suggesting the flexibility of LTCF wastewater surveillance for pathogens beyond SARS-CoV-2.

Data Quality: How Complete Were the Wastewater Surveillance Data? How Was Data Quality Assured?

The wastewater surveillance system obtained 24-hour composite wastewater samples from the LTCFs 2 to 4 days per week. In Lexington, wastewater effluent was tested 47.3% (160/338) of the days in the study period. In Louisville facilities, wastewater testing covered 35.8% (98/274) to 37.2% (102/274) of the days in the study period. The subsection describing system stability provides additional information regarding wastewater sample collection issues that may have affected data quality. Of the 811 wastewater samples collected, 31 (3.8%) were not processed owing to reagent shortages (n=21, 68%), processing delays after winter storms (n=9, 29%), and contamination during laboratory extraction (n=1, 3%). Of the 780 samples analyzed for SARS-CoV-2, 40 (5.1%) had no detectable concentration of crAssphage or were not analyzed for crAssphage. The absence of detectable crAssphage in a wastewater sample suggests low or negligible presence of fecal material, which makes the detection of SARS-CoV-2 RNA less likely.

Custom software identified data inconsistencies in the REDCap database and the polymerase chain reaction (PCR) output files. In addition, laboratory personnel visually reviewed PCR curves and data generated by the PCR instrument to identify erroneous machine-estimated results. To maintain data integrity, technicians entered corrections into an Excel workbook. The software applied these corrections to the REDCap data before joining the data with the PCR results. The frequencies of identified data issues are presented in Table 1.

Table 1. Frequency of identified quality issues in long-term care facility wastewater surveillance data.

Data source and data quality issue	Values, n (%)
REDCap[a] (n=780)	
Inaccurate sample ID	89 (11.4)
Inaccurate location ID	16 (2.1)
Inaccurate collection method	0 (0)
Inaccurate sample date	5 (0.6)
RT-qPCR[b] (n=7734)	
False-positive PCR[c] value	91 (1.2)
False-negative PCR value	2 (0)
Inaccurate PCR value	19 (0.2)

[a] REDCap: Research Electronic Data Capture.

[b] RT-qPCR: real-time qualitative polymerase chain reaction.

[c] PCR: polymerase chain reaction.

Sensitivity and PPV: What Proportion of the Time When There Was a Known SARS-CoV-2 Case Was There a Positive Wastewater SARS-CoV-2 Signal? What Was the Probability That a Positive Wastewater Signal Indicated an Active SARS-CoV-2 Infection?

We used Trilogy's clinical testing data to estimate the performance of wastewater surveillance for detecting SARS-CoV-2 infections. Trilogy primarily used antigen-based SARS-CoV-2 tests. Trilogy tested LTCF staff per CDC and state guidance, which recommended PCR testing of staff who were symptomatic and every-other-week to twice-a-week testing of unvaccinated staff based on COVID-19 activity in the facility's county. Residents received SARS-CoV-2 testing when symptomatic and twice weekly after the identification of a case in the facility until no new cases were identified for 2 weeks. Trilogy did not require visitors to test.

Wastewater surveillance sensitivity for detecting the presence of identified clinical cases was 30.6% (95% CI 24.4%-36.8%) using a signal threshold of >0 RNA copies/mL and decreased to 11.5% (95% CI 7.2%-15.8%) with a signal threshold of >250 RNA copies/mL (Figure 1). Wastewater surveillance specificity ranged from 79.7% (95% CI 76.4%-82.9%; >0 RNA copies/mL) to 98% (95% CI 97.7%-99.6%; >250 RNA copies/mL). When limiting clinical test data to residents, wastewater surveillance sensitivity at a signal threshold of >0 RNA copies/mL improved to 48% (95% CI 36.5%-59.4%), with a specificity of 80% (95% CI 77%-82.9%; Figure 1). The PPV of wastewater surveillance ranged from 34.8% (95% CI 27.9%-41.7%) to 75% (95% CI 60%-90%) for wastewater signal thresholds of >0 copies/mL and >250 copies/mL, respectively, when including all clinical data and ranged from 19.7% (95% CI 13.8%-25.5%) to 39.3% (95% CI 21.2%-57.4%) when considering only clinical tests from residents (Figure 2).

Figure 1. Receiver operating characteristic curves depicting the sensitivity and specificity of wastewater surveillance at varying SARS-CoV-2 wastewater signal thresholds to discriminate the presence of staff and residents with a positive SARS-CoV-2 clinical test at 6 long-term care facilities. Note: panel A includes resident and staff clinical test results; panel B includes only resident clinical test results. cp: copies.

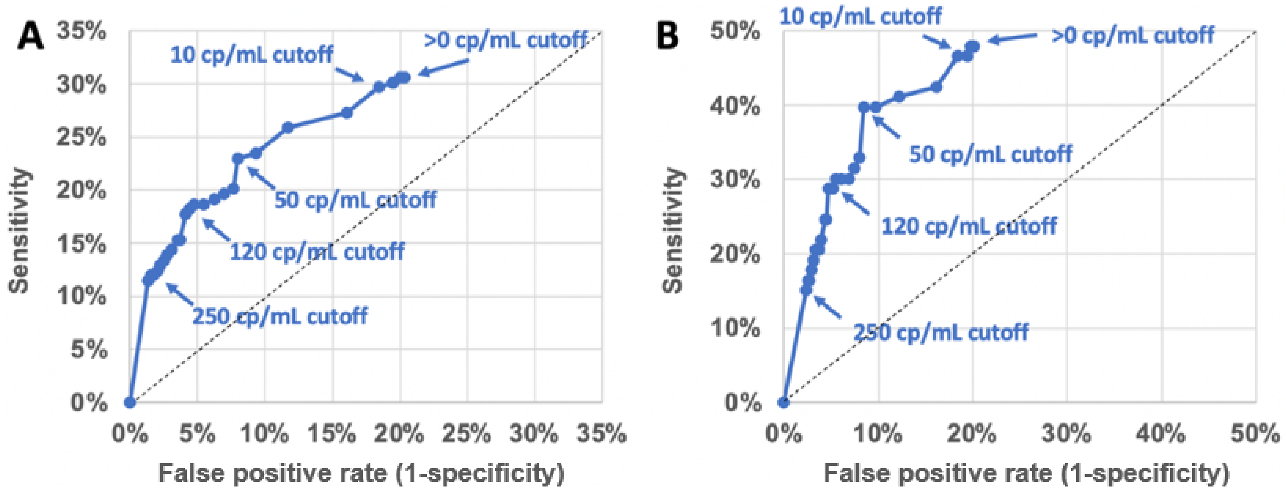
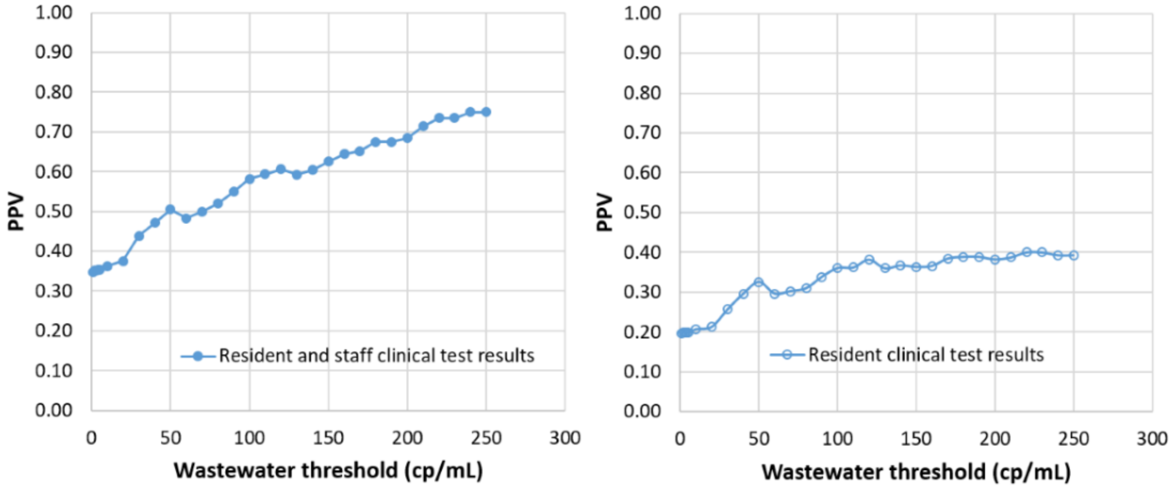


Figure 2. Positive predictive value (PPV) of a SARS-CoV-2 wastewater signal at various threshold values for the presence of a long-term care facility staff member or resident with a positive SARS-CoV-2 clinical test. cp: copies.



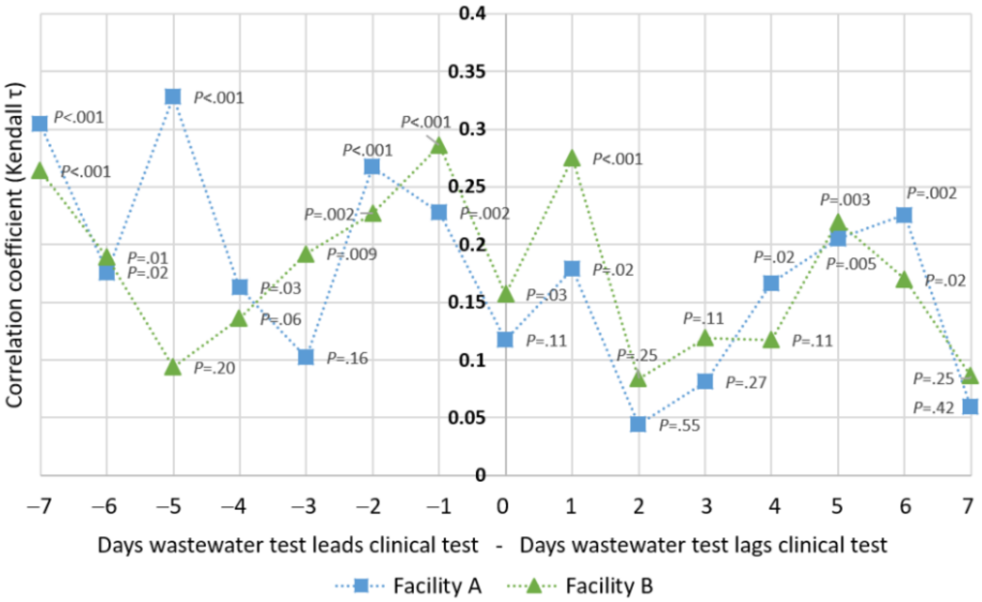
Limitations of the clinical testing data likely affected the estimates of wastewater surveillance performance. Rapid antigen test sensitivity is poor in asymptomatic individuals [15], and there was incomplete and variable testing of residents and staff. There was likely underascertainment of SARS-CoV-2 infections in the LTCF population, which would falsely deflate the estimates of wastewater surveillance specificity and PPV and inflate the estimates of sensitivity.

Timeliness: How Much Time Elapsed Between Collecting a Wastewater Sample, Analytic Results, and Infection Prevention Action by the Facility? Did Facility-Level Wastewater Surveillance Detect the Presence of SARS-CoV-2 Sooner Than Routine Clinical Testing or Testing Triggered by Symptom Screening?

Wastewater surveillance data were typically available for stakeholders 24 to 72 hours after sample collection. Variation in data timeliness was due primarily to personnel availability for sample processing and data entry. In general, turnaround time was faster for samples collected on Tuesdays, Wednesdays, and Thursdays than for samples collected on Fridays because laboratory staff did not work on weekends. Median turnaround times for results for samples collected on Tuesdays, Wednesdays, and Thursdays were 26, 37, and 28 hours, respectively. For results for samples collected on Fridays, the median turnaround time was 74 hours.

A lagged correlation analysis between the wastewater signal and clinical case detection showed variable correlation across the 6 facilities. At the 2 Lexington facilities with the greatest number of clinical cases, the wastewater signal correlated temporally with clinical testing results (Figure 3). Although wastewater positivity generally led clinical positivity (rather than lagged), there was not a specific lead time that outperformed all others because significant correlations were seen for lead times ranging from 1 (P<.001) to 7 (P<.001) days.

Figure 3. Lagged correlational analysis of wastewater SARS-CoV-2 positivity with SARS-CoV-2 clinical test positivity at 2 long-term care facilities. The x-axis indicates the number of days the wastewater data are offset relative to the clinical data, with negative values indicating the number of days that wastewater data lead, and positive values indicating the number of days that wastewater data lag, the clinical test data.



Representativeness: How Thoroughly Did Wastewater Testing Monitor the Population in the LTCFs?

Wastewater surveillance was representative of the population contributing to the wastewater at LTCFs because it examined building-wide sewer effluent over 24 hours. However, there were important caveats. The LTCF population that potentially contributed to wastewater at the facility was dynamic: there were resident admissions and discharges, staff worked at ≥1 locations, and the LTCFs reinstated resident visitation during the surveillance period. In addition, members of the LTCF population may not have contributed to the sewer effluent because they used toilets outside of the facility (staff) or because of incontinence and the use of adult briefs (residents). As previously mentioned, the inclusion of staff clinical results alongside resident results substantially affected the sensitivity (lower) and PPV (higher) of wastewater surveillance.

During a cluster of 10 identified resident cases at 1 LTCF, the surveillance team noted an intermittently positive wastewater signal. An investigation of this unexpected variability revealed that 6 (60%) of the 10 residents were fully or partially incontinent and using adult briefs. Their feces did not enter the wastewater stream and therefore did not contribute to a measurable SARS-CoV-2 wastewater signal.

Stability: What Issues Arose With Surveillance Equipment, Processes, Data Collection, or Reporting, and How Did This Affect Surveillance? How Much Did the System Cost to Operate?

Global supply chain issues owing to the COVID-19 pandemic required flexibility. At the study onset, autosamplers were unavailable, which delayed surveillance implementation by a month. Intermittent vendor shortages of laboratory supplies forced reuse protocols when possible and resulted in a reduction of replicates analyzed in 9.5% (74/780) of the samples. Several environmental and infrastructure factors challenged wastewater collection. Autosampler logs identified periods of low flow in the wastewater effluent stream, which resulted in smaller sample volumes and less representative composite samples at 1 facility in particular. We did not observe clogging of the strainers at the end of the autosampler tubing; however, intermittent clogging may have occurred and contributed to the low flow measured by the autosamplers. The results from the days and locations with smaller composite volumes may not represent conditions from the 24-hour collection period. In addition, sewer architecture resulted in a likely false-positive wastewater signal at 1 LTCF (Textbox 1).

Wastewater composite samples exceeded the goal temperature threshold of 4 °C (for optimal RNA stability) during summer months. Decreasing the composite sample volume (to 6000 mL) and increasing the volume of ice used in the autosamplers increased the frequency with which the composite wastewater samples stayed below the target temperature.

The reoccurring cost to conduct 1 day of wastewater surveillance at an LTCF was approximately US $144.50, or approximately US $18.06 per replicate. This estimate included transportation, labor, and materials based on wastewater testing at 6 sites (Table 2). Additional 1-time equipment expenses, such as laboratory equipment (notably a PCR machine) and autosamplers, contribute to start-up costs for wastewater surveillance. Amortized over the lifetime of the equipment, these expenses will increase per-sample costs by approximately US $5 to US $10. There are potential economies of scale related to labor because there are time efficiencies when obtaining and processing samples from multiple facilities.

Textbox 1. Adapting wastewater sampling to mitigate sewer architecture challenges.

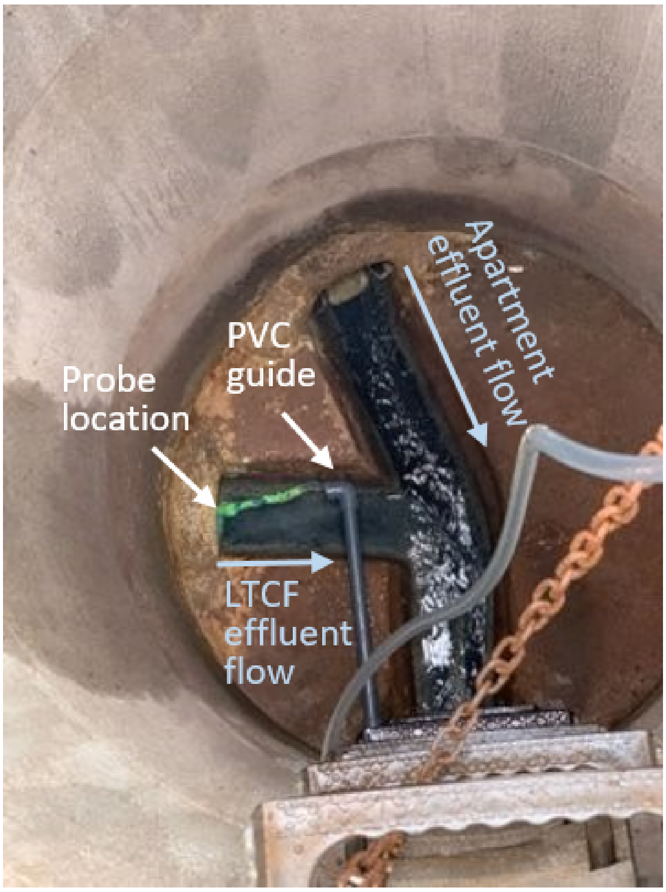
During routine wastewater surveillance, the team detected a strongly positive wastewater signal at a facility with no known COVID-19 infections. This happened on a Friday, and the surveillance team notified long-term care facility (LTCF) leadership shortly before the close of business. Over the subsequent 2 days, the LTCF leadership tested all residents and staff at the facility but did not identify an individual with SARS-CoV-2 infection. LTCF site management shared that there were COVID-19 cases at a neighboring apartment building. An examination of sewer architecture revealed that wastewater effluent from the LTCF and the apartment building converged at our sampling point (Figure 4). Although the field technician placed the autosampler probe into the LTCF wastewater effluent channel, the technician noted intermittent probe migration into the convergent channel. We presume that the strongly positive wastewater signal came from the apartment building. We addressed this sampling issue by designing an autosampler probe guide made using a polyvinyl chloride pipe and rigid wire. This custom device ensured that the probe remained in the channel containing only LTCF wastewater.

Table 2. Cost estimates for wastewater surveillance at long-term care facilities by expense category.

	Cost/replicate (US $)	Cost/sample (US $)	Cost per capita[a]/sample (US $)
Reagents	6.47	51.78	0.28
Labor	7.30	58.38	0.31
Transportation	4.29	34.33	0.18
Total	18.06	144.50	0.77

[a] On the basis of the total population (N=1124) of residents and staff at the 6 long-term care facilities at the conclusion of wastewater surveillance.

Figure 4. Autosampler probe guide designed to address the sampling issue at a long-term care facility (LTCF). PVC: polyvinyl chloride.



Additional Wastewater Surveillance Considerations

Stakeholders mentioned that the relative newness of wastewater surveillance added complexity to surveillance implementation. The collaboration between university researchers and the LTCF organization raised data sharing and privacy issues related to resident and staff COVID-19 testing results. A formal data use agreement clarified expectations.

All interviewed stakeholders raised the issue of how best to interpret, and respond to, the wastewater data. LTCF and local health department leadership expressed a desire for a clear course of action. They asked questions about the strength of a wastewater signal that should trigger clinical testing and about how wastewater RNA concentrations correlated to the numbers of infections at the facility. Study scientists stressed that, at best, wastewater data indicated trends and not absolute numbers of infections. Local health department officials could not foresee incorporating wastewater surveillance into daily practice without a high level of evidence and guidance supporting its adoption.

Several stakeholders mentioned the importance of goal setting between LTCF leadership and the study team. Scientists expressed a desire to define ideal parameters such as target sensitivity and PPV, turnaround time, and methods of communication. LTCF leadership prioritized a streamlined approach to wastewater surveillance, such as an *industry build out* with established resources to easily set up and maintain systems *off the shelf*, such as commercially available fire protection systems.

Discussion

Principal Findings

Our evaluation of wastewater surveillance for SARS-CoV-2 at LTCFs found the approach feasible and timely, with results available to stakeholders within approximately 24 to 72 hours. Wastewater surveillance demonstrated decent specificity but poor sensitivity for identifying known SARS-CoV-2 clinical cases, although sensitivity improved when using only clinical test results from LTCF residents (excluding staff). LTCF wastewater surveillance was stable; solutions to environmental challenges were manageable, and there were minimal disruptions in sampling.

Stakeholders consistently raised questions about wastewater surveillance performance and results interpretation. The limitations of clinical testing data, such as less frequent testing of residents and the use of rapid antigen tests, affected the team’s ability to accurately estimate the performance of wastewater surveillance (ie, there was not a gold standard on which to base sensitivity and specificity calculations). Uncertainties also existed regarding the population under wastewater surveillance. Family visits resumed during the surveillance period, staff bowel habits were unknown, and residents with fecal incontinence used adult briefs, meaning that their feces did not enter the wastewater stream. Staff with COVID-19 isolated at home for a variable period before returning to work and potentially contributing to the wastewater stream. These limitations eroded confidence in wastewater surveillance performance.

Stakeholders wanted to quantify the number of COVID-19 cases represented by a specific wastewater signal intensity. The relationship between wastewater signal intensity and number of cases was not clear, owing in part to uncertainties regarding the population under surveillance described in the preceding paragraph. In addition, viral shedding of SARS-CoV-2 is variable in frequency, intensity, and duration [16]. These parameters likely vary among SARS-CoV-2 variants and across host characteristics (age, disease intensity, and comorbid conditions) [17]. Additional facility and environmental considerations may have affected measured wastewater SARS-CoV-2 concentrations, such as variations in wastewater flow, presence of inhibitor substances (cleaning agents and disinfectants), ambient temperature, and the heterogeneous nature of wastewater [18,19]. However, water quality differences among facilities may not correlate with changes in SARS-CoV-2 concentrations [20]. Because of these factors, precisely delimiting the relationship between wastewater SARS-CoV-2 concentration and COVID-19 cases was difficult and may not be realistic at the facility level. Furthermore, once there is an established case in a building, identifying new infections using wastewater measurements remains a hurdle because convalescing individuals continue to shed virus [21]. SARS-CoV-2 RNA present in wastewater from convalescing individuals may decrease the estimated specificity and PPV of the wastewater signal for detecting new SARS-CoV-2 infections.

Variability in turnaround time from sample collection to result dissemination was primarily because of staffing and the day of sample collection. Samples collected on Tuesday through Thursday underwent analysis the same day, with results typically reported the next morning, whereas samples collected on Friday had delayed reporting of results owing to limited weekend staffing. Turnaround times observed in this evaluation were comparable with those reported by a wastewater laboratory in Berkeley, California [22]. Wastewater testing results turnaround time, regardless of the day of the week, was quicker than typical turnaround times for PCR clinical testing (48-72 h) but slower than rapid antigen tests (15 min to 1 h) [7]. Adding staff in the evening or during the weekend could mitigate the observed variability in results turnaround times and support wastewater surveillance as an effective early warning system.

Wastewater surveillance at LTCFs may provide an early warning of infection at a facility before detection by routine screening practices as suggested by the time-shifted correlation analysis. A positive wastewater signal could trigger enhanced clinical testing and infection prevention actions. To maximize the potential lead time provided by the wastewater signal, an LTCF would need clear operating procedures in place. These include capacity to test, perform contact tracing, and isolate individuals who have been potentially infected, as demonstrated on university campuses that used wastewater surveillance to trigger disease mitigation responses [23-25]. Uncertainties in how to interpret the wastewater data and how best to communicate this information within our academic-private partnership limited actions taken in response to wastewater data and require additional investigation in how best to use wastewater surveillance as an infection-detection modality.

Facility wastewater surveillance was flexible, as demonstrated by the rapid resumption of surveillance at the onset of the SARS-CoV-2 Omicron variant wave in January 2022; however, there are structural limitations to this surveillance. Sewer architecture is an important consideration when evaluating facility suitability for wastewater surveillance [26]. Situations where facility-level wastewater surveillance is less suitable include facilities where sewer access is limited to manholes in busy streets, there is low effluent flow, and convergent wastewater streams prohibit the selective sampling of wastewater from the facility of interest. Labor availability or lack of access to a laboratory skilled in RT-qPCR might prolong results turnaround time to a point where wastewater surveillance may not provide an early warning of disease [25].

Most stakeholders expressed enthusiasm for wastewater surveillance, recognizing its cost-effective noninvasive nature and potential for monitoring SARS-CoV-2 variants of concern and other pathogens, including methicillin-resistant *Staphylococcus aureus* and *C difficile*, as well as viruses that cause respiratory infections such as influenza. The specific use cases for wastewater surveillance will vary based on stakeholder priorities, institutional wastewater infrastructure, and available resources. Wastewater surveillance may also serve purposes beyond disease surveillance, such as maintaining relationships with external stakeholders [27]. In our case, the leadership at the LTCF organization were interested in wastewater surveillance as an emerging technology that could protect the well-being of their residents and demonstrate their commitment as an industry leader in safety.

Stakeholders in the LTCF industry cited the lack of easily deployable wastewater surveillance systems as a major hurdle to wider adoption. They described a desire for *off-the-shelf, ready-to-go* wastewater surveillance systems, akin to fire protection systems. The COVID-19 pandemic is driving rapid innovation of wastewater surveillance technology, and as the industry develops and technology improves, LTCF managers may be open to broader adoption of this tool. Alternatively, LTCFs could partner with local utilities or health departments to build sustainable wastewater surveillance systems.

Although reports describing the use of wastewater surveillance at LTCFs are sparse, there are other facility-based settings where wastewater surveillance flourished during the COVID-19 pandemic. Universities, in particular, were early adopters and implemented building-level wastewater testing, while developing wastewater sampling strategies, laboratory assays, and public health response measures [25]. Schools were another congregate setting where wastewater surveillance demonstrated utility in identifying and potentially mitigating SARS-CoV-2 infection [28,29]. Correctional settings have also used and evaluated wastewater surveillance [30], and former inmates have voiced a preference for wastewater surveillance over individual testing for SARS-CoV-2 [31]. Many of the lessons learned from this evaluation of wastewater surveillance at LTCFs may apply to wastewater surveillance in other facility-level settings.

This surveillance evaluation had limitations in addition to the challenges we have already described. Sparse data from relatively few positive clinical tests during the surveillance period reduced the power of the statistical analyses and may have affected surveillance performance estimates. The stakeholders interviewed for the evaluation shared perspectives that may not represent the opinions and experiences of individuals involved in wastewater surveillance in other settings, meaning that the findings from this evaluation may not generalize to other LTCF-based wastewater surveillance settings. However, by adhering to the CDC guidelines for evaluating public health surveillance systems, we used reproducible methods and provide credible data and insight into the performance of wastewater surveillance for SARS-CoV-2 at 6 LTCFs in Kentucky.

Conclusions and Recommendations

Facility-level wastewater surveillance can monitor populations considered vulnerable for the presence of infectious diseases. Stakeholders found the surveillance feasible and expressed optimism regarding its potential while also recognizing challenges in interpreting, and acting on, the data. Further studies of the performance of facility-level wastewater surveillance will improve the interpretation of wastewater data and increase the utility of this emerging surveillance modality.

Specific recommendations based on this evaluation are to (1) investigate the relationship between wastewater SARS-CoV-2 RNA levels and the number of individuals infected and convalescing at a facility (fecal shedding studies may help elucidate this relationship); (2) synthesize facility-based wastewater surveillance data across projects to develop and test guidance on data interpretation; and (3) foster relationships among academic partners, LTCF organizations, and public health officials to clarify and strengthen communication practices to understand the priorities and limitations of wastewater surveillance.

Acknowledgments

The authors would like to thank Steven Roggenkamp for creating the data architecture for this study. The authors acknowledge and thank the facilities staff at the Trilogy Health Services facilities where the wastewater surveillance occurred. This work was funded by the National Institutes of Health (P30 ES026529) and the Centers for Disease Control and Prevention (BAA 75D301-20-R-68024). The contents are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health or Centers for Disease Control and Prevention.

Data Availability

The wastewater and interview data evaluated for this study are available from the corresponding author on reasonable request. Clinical test data are available on the Trilogy COVID-19 dashboard [32].

Authors' Contributions

JWK and SMB conceived of, and secured funding for, this study. JL, ML, BM, CO, WS, and AN generated data for this study. PL provided access to data. JL, RA, JWK, SMB, and ML analyzed the data. JL and JWK drafted the manuscript. PL, SMB, TS, and ML reviewed and edited the manuscript. JWK is the guarantor of this manuscript.

Conflicts of Interest

SMB reports ownership interest in Salus Discovery, LLC, which licenses the technology used for the wastewater analysis described in the text. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Key informant semistructured interview guides.
[DOCX File , 14 KB - publichealth_v9i1e44657_app1.docx]

References

1. Lau-Ng R, Caruso LB, Perls TT. COVID-19 deaths in long-term care facilities: a critical piece of the pandemic puzzle. J Am Geriatr Soc 2020 Sep;68(9):1895-1898 [FREE Full text] [doi: 10.1111/jgs.16669] [Medline: 32501537]
2. Chidambaram P. Over 200,000 residents and staff in long-term care facilities have died from COVID-19. Kaiser Family Foundation. 2022 Feb 03. URL: https://www.kff.org/policy-watch/over-200000-residents-and-staff-in-long-term-care-facilities-have-died-from-covid-19/ [accessed 2022-05-03]
3. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, Public Health–Seattle and King County and CDC COVID-19 Investigation Team. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. N Engl J Med 2020 May 28;382(22):2081-2090 [FREE Full text] [doi: 10.1056/NEJMoa2008457] [Medline: 32329971]
4. Orive G, Lertxundi U, Barcelo D. Early SARS-CoV-2 outbreak detection by sewage-based epidemiology. Sci Total Environ 2020 Aug 25;732:139298 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.139298] [Medline: 32417555]
5. Chen Y, Chen L, Deng Q, Zhang G, Wu K, Ni L, et al. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. J Med Virol 2020 Jul;92(7):833-840. [doi: 10.1002/jmv.25825] [Medline: 32243607]

6. Olesen SW, Imakaev M, Duvallet C. Making waves: defining the lead time of wastewater-based epidemiology for COVID-19. Water Res 2021 Sep 01;202:117433 [FREE Full text] [doi: 10.1016/j.watres.2021.117433] [Medline: 34304074]
7. Betancourt WQ, Schmitz BW, Innes GK, Prasek SM, Pogreba Brown KM, Stark ER, et al. COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. Sci Total Environ 2021 Jul 20;779:146408 [FREE Full text] [doi: 10.1016/j.scitotenv.2021.146408] [Medline: 33743467]
8. Kirby AE, Walters MS, Jennings WC, Fugitt R, LaCross N, Mattioli M, et al. Using wastewater surveillance data to support the COVID-19 response - United States, 2020-2021. MMWR Morb Mortal Wkly Rep 2021 Sep 10;70(36):1242-1244 [FREE Full text] [doi: 10.15585/mmwr.mm7036a2] [Medline: 34499630]
9. Karthikeyan S, Nguyen A, McDonald D, Zong Y, Ronquillo N, Ren J, et al. Rapid, large-scale wastewater surveillance and automated reporting system enable early detection of nearly 85% of COVID-19 cases on a university campus. mSystems 2021 Aug 31;6(4):e0079321. [doi: 10.1128/msystems.00793-21]
10. LaJoie AS, Holm RH, Anderson LB, Ness HD, Smith T. Nationwide public perceptions regarding the acceptance of using wastewater for community health monitoring in the United States. PLoS One 2022 Oct 11;17(10):e0275075 [FREE Full text] [doi: 10.1371/journal.pone.0275075] [Medline: 36219594]
11. German R, Lee L, Horan J, Milstein RL, Pertowski CA, Waller MN, Guidelines Working Group Centers for Disease Control and Prevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-CE7. [Medline: 18634202]
12. Watkins DC. Rapid and rigorous qualitative data analysis: the “RADaR” technique for applied research. Int J Qual Methods 2017 Jan;16(1):160940691771213 [FREE Full text] [doi: 10.1177/1609406917712131]
13. Berry SM, Pezzi HM, Williams ED, Loeb JM, Guckenberger DJ, Lavanway AJ, et al. Using exclusion-based sample preparation (ESP) to reduce viral load assay cost. PLoS One 2015 Dec 2;10(12):e0143631 [FREE Full text] [doi: 10.1371/journal.pone.0143631] [Medline: 26630135]
14. Strike W, Amirsoleimani A, Olaleye A, Noble A, Lewis K, Faulkner L, et al. Development and validation of a simplified method for analysis of SARS-CoV-2 RNA in university dormitories. ACS ES T Water 2022 May 09;2(11):1984-1991. [doi: 10.1021/acsestwater.2c00044] [Medline: 37552725]
15. Oran DP, Topol EJ. The proportion of SARS-CoV-2 infections that are asymptomatic: a systematic review. Ann Intern Med 2021 May;174(5):655-662 [FREE Full text] [doi: 10.7326/M20-6976] [Medline: 33481642]
16. Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L, Byrne P, et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. J Infect 2020 Sep;81(3):357-371 [FREE Full text] [doi: 10.1016/j.jinf.2020.06.067] [Medline: 32615199]
17. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. Lancet Microbe 2021 Jan;2(1):e13-e22 [FREE Full text] [doi: 10.1016/S2666-5247(20)30172-5] [Medline: 33521734]
18. Buonerba A, Corpuz MV, Ballesteros F, Choo KH, Hasan SW, Korshin GV, et al. Coronavirus in water media: analysis, fate, disinfection and epidemiological applications. J Hazard Mater 2021 Aug 05;415:125580 [FREE Full text] [doi: 10.1016/j.jhazmat.2021.125580] [Medline: 33735767]
19. Kitajima M, Ahmed W, Bibby K, Carducci A, Gerba CP, Hamilton KA, et al. SARS-CoV-2 in wastewater: state of the knowledge and research needs. Sci Total Environ 2020 Oct 15;739:139076 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.139076] [Medline: 32758929]
20. Sharkey ME, Kumar N, Mantero AM, Babler KM, Boone MM, Cardentey Y, et al. Lessons learned from SARS-CoV-2 measurements in wastewater. Sci Total Environ 2021 Dec 01;798:149177 [FREE Full text] [doi: 10.1016/j.scitotenv.2021.149177] [Medline: 34375259]
21. Colosi LM, Barry KE, Kotay SM, Porter MD, Poulter MD, Ratliff C, et al. Development of wastewater pooled surveillance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from congregate living settings. Appl Environ Microbiol 2021 Jun 11;87(13):e0043321. [doi: 10.1128/aem.00433-21]
22. Kantor RS, Greenwald HD, Kennedy LC, Hinkle A, Harris-Lovett S, Metzger M, et al. Operationalizing a routine wastewater monitoring laboratory for SARS-CoV-2. PLoS Water 2022 Feb 15;1(2):e0000007 [FREE Full text] [doi: 10.1371/journal.pwat.0000007]
23. Scott LC, Aubee A, Babahaji L, Vigil K, Tims S, Aw TG. Targeted wastewater surveillance of SARS-CoV-2 on a university campus for COVID-19 outbreak detection and mitigation. Environ Res 2021 Sep;200:111374 [FREE Full text] [doi: 10.1016/j.envres.2021.111374] [Medline: 34058182]
24. Gibas C, Lambirth K, Mittal N, Juel MA, Barua VB, Roppolo Brazell L, et al. Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. Sci Total Environ 2021 Aug 15;782:146749 [FREE Full text] [doi: 10.1016/j.scitotenv.2021.146749] [Medline: 33838367]
25. Harris-Lovett S, Nelson KL, Beamer P, Bischel HN, Bivins A, Bruder A, et al. Wastewater surveillance for SARS-CoV-2 on college campuses: initial efforts, lessons learned, and research needs. Int J Environ Res Public Health 2021 Apr 22;18(9):4455 [FREE Full text] [doi: 10.3390/ijerph18094455] [Medline: 33922263]

26. Reeves K, Liebig J, Feula A, Saldi T, Lasda E, Johnson W, et al. High-resolution within-sewer SARS-CoV-2 surveillance facilitates informed intervention. Water Res 2021 Oct 01;204:117613 [FREE Full text] [doi: 10.1016/j.watres.2021.117613] [Medline: 34500183]
27. Harris-Lovett S, Nelson KL, Kantor R, Korfmacher KS. Wastewater surveillance to inform public health decision making in residential institutions. J Public Health Manag Pract 2023 May;29(3):317-321 [FREE Full text] [doi: 10.1097/PHH.0000000000001636] [Medline: 36214654]
28. Hassard F, Vu M, Rahimzadeh S, Castro-Gutierrez V, Stanton I, Burczynska B, et al. Wastewater monitoring for detection of public health markers during the COVID-19 pandemic: near-source monitoring of schools in England over an academic year. PLoS One 2023 May 30;18(5):e0286259 [FREE Full text] [doi: 10.1371/journal.pone.0286259] [Medline: 37252922]
29. Wolken M, Sun T, McCall C, Schneider R, Caton K, Hundley C, et al. Wastewater surveillance of SARS-CoV-2 and influenza in preK-12 schools shows school, community, and citywide infections. Water Res 2023 Mar 01;231:119648 [FREE Full text] [doi: 10.1016/j.watres.2023.119648] [Medline: 36702023]
30. Klevens RM, Young CC, Olesen SW, Osinski A, Church D, Muten J, et al. Evaluation of wastewater surveillance for SARS-CoV-2 in Massachusetts correctional facilities, 2020–2022. Front Water 2023 Jan 26;5 [FREE Full text] [doi: 10.3389/frwa.2023.1083316]
31. Riback LR, Dickson P, Ralph K, Saber LB, Devine R, Pett LA, et al. Coping with COVID in corrections: a qualitative study among the recently incarcerated on infection control and the acceptability of wastewater-based surveillance. Health Justice 2023 Feb 07;11(1):5 [FREE Full text] [doi: 10.1186/s40352-023-00205-0] [Medline: 36749465]
32. Trilogy Health Services. URL: https://trilogyhs.com/covid-19/ [accessed 2023-08-19]

Abbreviations

CDC: Centers for Disease Control and Prevention
LTCF: long-term care facility
PCR: polymerase chain reaction
PPV: positive predictive value
REDCap: Research Electronic Data Capture
RT-qPCR: real-time qualitative polymerase chain reaction

Edited by A Mavragani, T Sanchez; submitted 29.11.22; peer-reviewed by S Harris-Lovett, K Mendez; comments to author 06.04.23; revised version received 26.04.23; accepted 18.07.23; published 29.08.23.

Please cite as:
Keck JW, Lindner J, Liversedge M, Mijatovic B, Olsson C, Strike W, Noble A, Adatorwovor R, Lacy P, Smith T, Berry SM
Wastewater Surveillance for SARS-CoV-2 at Long-Term Care Facilities: Mixed Methods Evaluation
JMIR Public Health Surveill 2023;9:e44657
URL: https://publichealth.jmir.org/2023/1/e44657
doi:10.2196/44657
PMID:

Original Paper

Checkpoint Travel Numbers as a Proxy Variable in Population-Based Studies During the COVID-19 Pandemic: Validation Study

Jennifer M Kreslake[1], MPH, PhD; Kathleen Aarvig[1], MPH; Hope Muller-Tabanera[1], MS; Donna M Vallone[1,2,3], MPH, PhD; Elizabeth C Hair[1,2], PhD

[1]Schroeder Institute, Truth Initiative, Washington, DC, United States

[2]Department of Health, Behavior and Society, Johnks Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]College of Global Public Health, New York University, New York, NY, United States

Corresponding Author:
Jennifer M Kreslake, MPH, PhD
Schroeder Institute
Truth Initiative
900 G Street Northwest
Floor 4
Washington, DC, 20001
United States
Phone: 1 2024545782
Email: jkreslake@truthinitiative.org

Abstract

Background: The COVID-19 pandemic had wide-ranging systemic impacts, with implications for social and behavioral factors in human health. The pandemic may introduce history bias in population-level research studies of other health topics during the COVID-19 period.

Objective: We sought to identify and validate an accessible, flexible measure to serve as a covariate in research spanning the COVID-19 pandemic period.

Methods: Transportation Security Administration checkpoint travel numbers were used to calculate a weekly sum of daily passengers and validated against two measures with strong face validity: (1) a self-reported item on social distancing practices drawn from a continuous tracking survey among a national sample of youths and young adults (15-24 years) in the United States (N=45,080, approximately 280 unique respondents each week); and (2) Google's Community Mobility Reports, which calculate daily values at the national level to represent rates of change in visits and length of stays to public spaces. For the self-reported survey data, an aggregated week-level variable was calculated as the proportion of respondents who did not practice social distancing that week (January 1, 2019, to May 31, 2022). For the community mobility data, a weekly estimate of change was calculated using daily values compared to a 5-week prepandemic baseline period (January 3, 2020, to February 6, 2020). Spearman rank correlation coefficients were calculated for each comparison.

Results: Checkpoint travel data ranged from 668,719 travelers in the week of April 8, 2020, to nearly 15.5 million travelers in the week of May 18, 2022. The weekly proportion of survey respondents who did not practice social distancing ranged from 18.1% (n=42; week of April 15, 2020) to 70.9% (n=213; week of May 25, 2022). The measures were strongly correlated from January 2019 to May 2022 (ρ=0.90, *P*<.001) and March 2020 to May 2022 (ρ=0.87, *P*<.001). Strong correlations were observed when analyses were restricted to age groups (15-17 years: ρ=0.90; *P*<.001; 18-20 years: ρ=0.87; *P*<.001; 21-24 years: ρ=0.88; *P*<.001), racial or ethnic minorities (ρ=0.86, *P*<.001), and respondents with lower socioeconomic status (ρ=0.88, *P*<.001). There were also strong correlations between the weekly change from the baseline period for checkpoint travel data and community mobility data for transit stations (ρ=0.92, *P*<.001) and retail and recreation (ρ=0.89, *P*<.001), and moderate significant correlations for grocery and pharmacy (ρ=0.68, *P*<.001) and parks (ρ=0.62, *P*<.001). A strong negative correlation was observed for places of residence (ρ=−0.78, *P*<.001), and a weak but significant positive correlation was found for workplaces (ρ=0.24, *P*<.001).

Conclusions: The Transportation Security Administration's travel checkpoint data provide a publicly available flexible time-varying metric to control for history bias introduced by the pandemic in research studies spanning the COVID-19 period in the United States.

(JMIR Public Health Surveill 2023;9:e44950) doi:10.2196/44950

KEYWORDS

research methods; public health; data quality; psychosocial factors; history; COVID-19; social; behavioral; validation; social distancing; tracking survey; survey; pandemic

Introduction

On March 11, 2020, the outbreak of the disease resulting from SARS-CoV-2 (COVID-19) was declared a pandemic by the World Health Organization [1]. As cases surged in the United States, public health and government officials engaged in unprecedented efforts to limit the spread of the virus. Without effective pharmacological interventions and vaccines, social distancing emerged as the primary mitigation strategy [2,3]. Furthermore, many governments enacted stay-at-home mandates and the closing of nonessential businesses, with the reasoning that with fewer people traveling outside their homes, there would be less close physical contact between infected and uninfected individuals [4]. According to a growing body of research evaluating the effectiveness of these interventions, these actions were instrumental in reducing infection rates [5].

The COVID-19 pandemic, and subsequent control measures, had wide-ranging systemic impacts with implications for social and behavioral factors in human health. The time-dependent nature of the pandemic introduces history bias into longitudinal and time series analyses, with the pandemic's many impacts acting as confounders in population-level research studies. A commentary by Tuttle et al [6] noted that potential bias introduced by the COVID-19 pandemic has important considerations for analysis and interpretation in clinical research, and suggested that potential confounding may be addressed by examining pre-and postpandemic data with a control for COVID-19 behaviors. Such a control variable should have validity within a variety of target populations addressed by public health (eg, by age, socioeconomic status, race, or ethnicity).

Mobility data have the ability to provide nearly real-time information about changes in patterns of human movement as measured by activity on mobile phones, GPS tracking, and social media platforms. For the past several years, researchers have used mobility data to inform epidemiological modeling, situational awareness, and resource allocation during public health crises [7-14]. Thus, most of the available literature leverages mobility data to track the movements of individuals during the COVID-19 pandemic, to identify areas of high transmission, or evaluate the efficacy of interventions, such as social distancing [7-12]. In contrast, few studies have applied these data to other areas of public health research concurrent with the COVID-19 pandemic. It was 1 study from the United Kingdom that used mobility data to assess the impact of COVID-19 pandemic restrictions on physical activity levels [14]. Another in the United States established an association between county-level social distancing patterns during the COVID-19 pandemic and other county-level indicators of health behavior (eg, smoking, obesity, flu vaccination, and mammography screening) using data from GPS-enabled mobile devices [13].

The purpose of this study is to identify a suitable control variable for longitudinal or time series social and behavioral studies that span the COVID-19 pandemic period. We sought to determine whether air travel metrics can be used as a time-varying endogenous indicator of COVID-19 impacts on societal activities. The airline industry was detrimentally impacted by the COVID-19 pandemic. Passenger air travel was the largest contributor to industry losses during the pandemic, when the number of air travel passengers dropped by approximately 60% in 2020 [15,16]. Thus, we believe this may be a suitable control variable to integrate retrospectively into data sets that span the COVID-19 period. We will validate this measure against 2 time-varying population-level measures with strong face validity: the first is calculated by aggregating individual-level self-reported social distancing practices at a weekly level, drawn from a continuous tracking survey of youths and young adults in the United States; and the second is from Google's COVID-19 Community Mobility Reports, which calculate daily values at the national level to represent rates of change in visits and length of stays to public spaces [9].

Methods

Data and Measures

Checkpoint Travel Data

Publicly available data from the Transportation Security Administration (TSA) was used to construct an exogenous indicator of COVID-19's population-level impacts. TSA maintains a daily count of travelers screened at TSA airport checkpoints. As of this writing, data have been made available by TSA on the internet from 2019 to the present [17]. A weekly checkpoint travel variable was constructed by summing the TSA's daily checkpoint travel numbers within each week (Wednesday-Tuesday), which was used for the analysis correlating this measure to self-reported social distancing (Multimedia Appendix 1). For analyses comparing checkpoint travel data to community mobility data, the weekly checkpoint travel data were calculated as a weekly change score compared to the median checkpoint travel value for the corresponding day of the week, during the 5-week prepandemic period spanning January 3, 2020, to February 6, 2020.

Self-Reported Social Distancing Data

The self-reported measure of social distancing used data from a cross-sectional, continuous tracking survey of youths and young adults (aged 15-24 years) drawn from the Dynata national web-based panel. The survey is administered continuously (ie, sampling occurs every day) and receives responses from approximately 280 unique participants per week. Participants who completed a survey between January 1, 2019, to May 31, 2022 (N=45,080) were included in the analysis. Sampling quotas were used and survey weights were applied using age, race or ethnicity, and sex based on demographic benchmarks from the US Census to approximate a national sample.

A self-reported, single-response item on the frequency of social distancing during the past week was added to the survey instrument on March 25, 2020, and remained throughout the study period. Respondents were presented with the following question: “Social distancing means people stay home as much as possible to prevent catching and spreading the virus, and if they go outside, they stay at least 6 feet away at all times from any person who they don’t live with. Other terms for social distancing that some people use are self-quarantine and self-isolation. In the past week, have you done social distancing…?” Responses were coded as 0=at all times, 1=usually, 2=sometimes, 3=very little, and 4=not at all. Responses to the self-reported measure on social distancing were dichotomized (0=at all times or usually, 1=sometimes or very little or not at all). An aggregated week-level variable was calculated as the proportion of respondents that week who reported social distancing sometimes or very little or not at all (ie, coded as “1” for the dichotomized variable; Multimedia Appendix 2). In weeks prior to March 9, 2020, respondents were coded as “1” to reflect a lack of social distancing during the pre–COVID-19 period. The survey did not ask about social distancing during the first 2 weeks of the pandemic (weeks of March 9 and March 16), so the Next Observation Carried Backward method was used to impute values for the weeks of March 9 and March 16 using the proportions from the week of March 23. The aggregated values in weeks prior to March 9 were 100%, representing no social distancing across all respondents (n=280) within each of those weeks.

We constructed 2 additional aggregate measures of self-reported social distancing using only data from population subgroups (ie, racial or ethnic minorities and lower socioeconomic status) to determine if the correlation to the endogenous measure is similar for target populations commonly found in public health research. For racial or ethnic minorities, we limited the aggregation of the self-reported measure to those from respondents who self-identified as being non-White (ie, Black or African American, Asian, Native Hawaiian or other Pacific Islander, American Indian or Alaska Native, or multiracial) or Hispanic (n=14,657). For socioeconomic status, we used a subjective measure of household income that has been validated among young adults. Young adult respondents are asked: “Considering your own income and the income from any people who help you, how would you describe your own overall financial situation? Would you say you… (live comfortably; meet needs with a little left over; just meet basic expenses with nothing left over; don’t meet basic expenses).” This measure was adapted to respondents aged 15 to 17 years, who are asked: “Considering all income earners in your family, how would you describe your family’s overall financial situation, would you say you…” with the same response options as the original measure. The aggregated measure for low socioeconomic status was limited to respondents who reported having their basic needs met with nothing left over or not having their basic needs met (n=30,054).

COVID-19 Community Mobility Data

Google’s COVID-19 Community Mobility Data compare how visits and length of stay in various sites within communities changed compared to a 5-week prepandemic baseline period (January 3, 2020, to February 6, 2020) [18]. Data are drawn from mobile devices where users are signed into their Google accounts. Data collection is limited to Google users who have elected to activate their Location History; the default setting is off. To be comparable to the TSA’s checkpoint travel data, the analysis for this paper used country-level data for the United States.

A daily change score is calculated using values from each day compared to the median value for the corresponding day of the week during the 5-week baseline period. For the purposes of this study, we calculated weekly change scores for the TSA data from the same baseline period.

Google classifies community locales using the following categories: retail and recreation (eg, restaurants and cafes, shopping centers, theme parks, museums, libraries, and movie theaters), grocery and pharmacies (eg, grocery markets, food warehouses, farmer’s markets, specialty food shops, drug stores, and pharmacies), parks (eg, local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens), transit stations (eg, subway, bus, and train stations), workplaces, and places of residence.

Ethics Approval

The study was conducted in accordance with the Declaration of Helsinki. Checkpoint travel data and community mobility report data were obtained from anonymous, aggregated publicly available sources (TSA and Google, respectively). The protocol for the survey providing self-reported social distancing data was approved by the Advarra Institutional Review Board (Protocol ID: Pro0010120). Informed consent was obtained from all survey respondents aged 18 years or older, or from a parent or legal guardian for all respondents under 18 years of age. Respondent incentives were in accordance with their membership in a web-based panel maintained by Dynata.

Statistical Analysis

The Shapiro-Wilk test for normality determined that measures were not normally distributed for weekly air travel (W=0.919, $P<.001$), aggregated weekly social distancing (W=0.942, $P<.001$), and community mobility data (grocery: W=0.892, $P<.001$; recreation: W=0.877, $P<.001$; transit: W=0.986, $P<.001$) during the pandemic period, likely due to the high level of social distancing and reduction in travel observed during a short duration at the beginning of the pandemic period.

Spearman rank correlation coefficient was calculated for the weekly sum of travel checkpoint data, as were (1) the weekly proportion of survey respondents who did not practice social distancing and (2) the weekly percent change in visits to public spaces compared to the prepandemic baseline.

A secondary analysis was conducted using first-order differencing to remove the effects of the overall trend spanning the study period and determine the checkpoint travel data’s usage in controlling for weekly change.

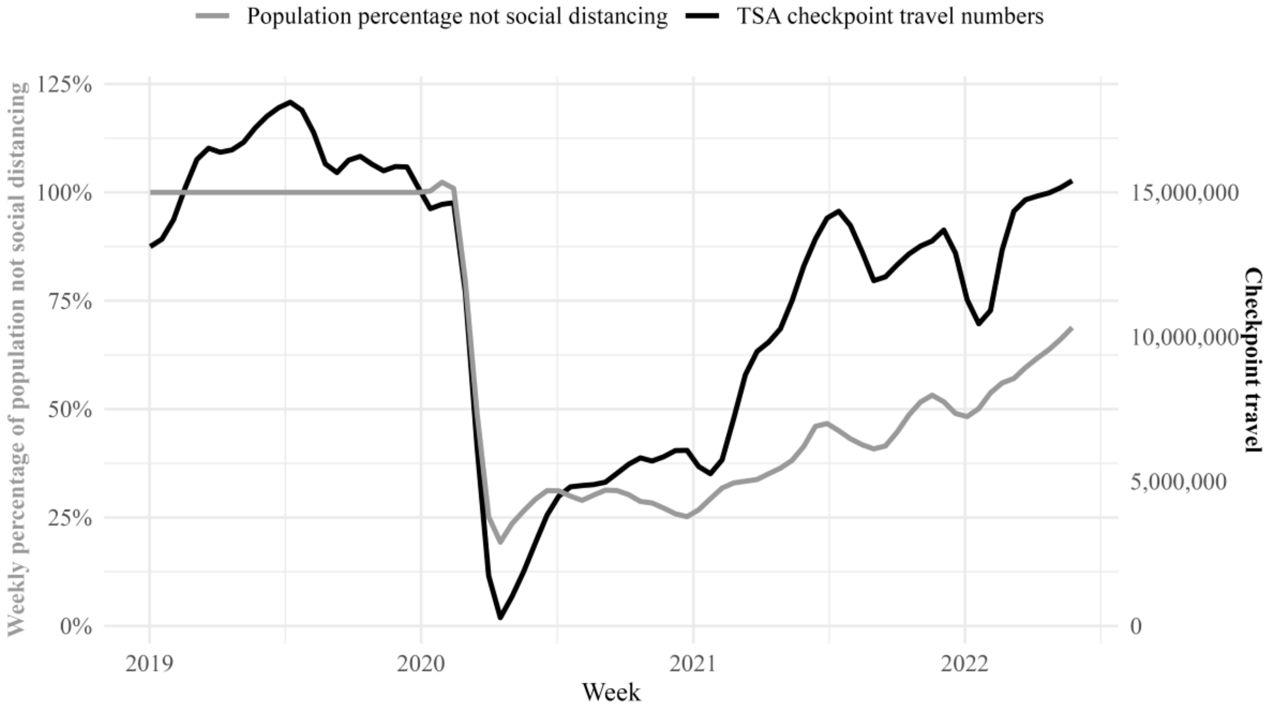
All analyses were conducted using RStudio (version 2021.09.0), developed by Posit, formerly known as RStudio Inc.

Results

Correlations Between Checkpoint Travel Data and Self-Reported Social Distancing

The relationship between the checkpoint travel data and self-reported social distancing measure (Figure 1) was visually inspected. The plot suggested a positive linear relationship between checkpoint travel numbers with no significant outliers when examining the prepandemic period and pandemic periods separately. Additional graphs were created to inspect whether a positive relationship could also be observed between checkpoint travel data and population subgroups (Multimedia Appendix 3) and whether checkpoint travel varied in relation to surges in COVID-19 variants (Multimedia Appendix 4).

Figure 1. Weekly checkpoint travel volume and proportion of youths and young adults reporting not social distancing before and during the COVID-19 pandemic. TSA: Transportation Security Administration.



Checkpoint travel data showed the minimum weekly sum of travelers (n=668,719) occurring the week of April 8, 2020, peaking at nearly 15.5 million the week of May 18, 2022 (Multimedia Appendix 1). The weekly proportion of survey respondents who did not practice social distancing followed a similar trend, ranging from 18.1% (n=42; week of April 15, 2020) to 70.9% (n=213; week of May 25, 2022; Multimedia Appendix 2). The measures were strongly and significantly correlated from January 2019 to May 2022 and March 2020 to May 2022 (Table 1). Strong correlations were observed when analyses were restricted to age groups (ie, 15-17, 18-20, and 21-24 years), racial or ethnic minorities, and respondents with lower socioeconomic status, all of which were significant (Table 1). First-order differencing correlations between weekly checkpoint travel data and self-reported social distancing were not significant for either the overall sample, neither for racial or ethnic minorities nor for respondents with lower socioeconomic status. A weak positive and significant correlation was observed for first-order differences of weekly checkpoint travel data and self-reported social distancing among youths aged 15-17 years (Table 1), but no other correlations using first-order differencing were significant for other age groups.

Table 1. Correlation of Transportation Security Administration's weekly checkpoint travel volume with other weekly indicators of behavioral impacts of the COVID-19 pandemic, January 2019 to May 2022.

	ρ	*P* value
Self-reported social distancing[a]		
Total population (aged 15-24 years)	0.90	<.001
Racial or ethnic minorities	0.86	<.001
Low SES[b]	0.88	<.001
Age 15-17 years	0.90	<.001
Age 18-20 years	0.87	<.001
Age 21-24 years	0.88	<.001
Change in community mobility[c]		
Transit stations	0.92	<.001
Retail and recreation	0.89	<.001
Grocery and pharmacy	0.68	<.001
Parks	0.62	<.001
Workplaces	0.24	<.001
Places of residence	−0.78	<.001

[a] Aggregated weekly percent reporting socially distancing sometimes or very little or not at all in the past week.

[b] SES: socioeconomic status.

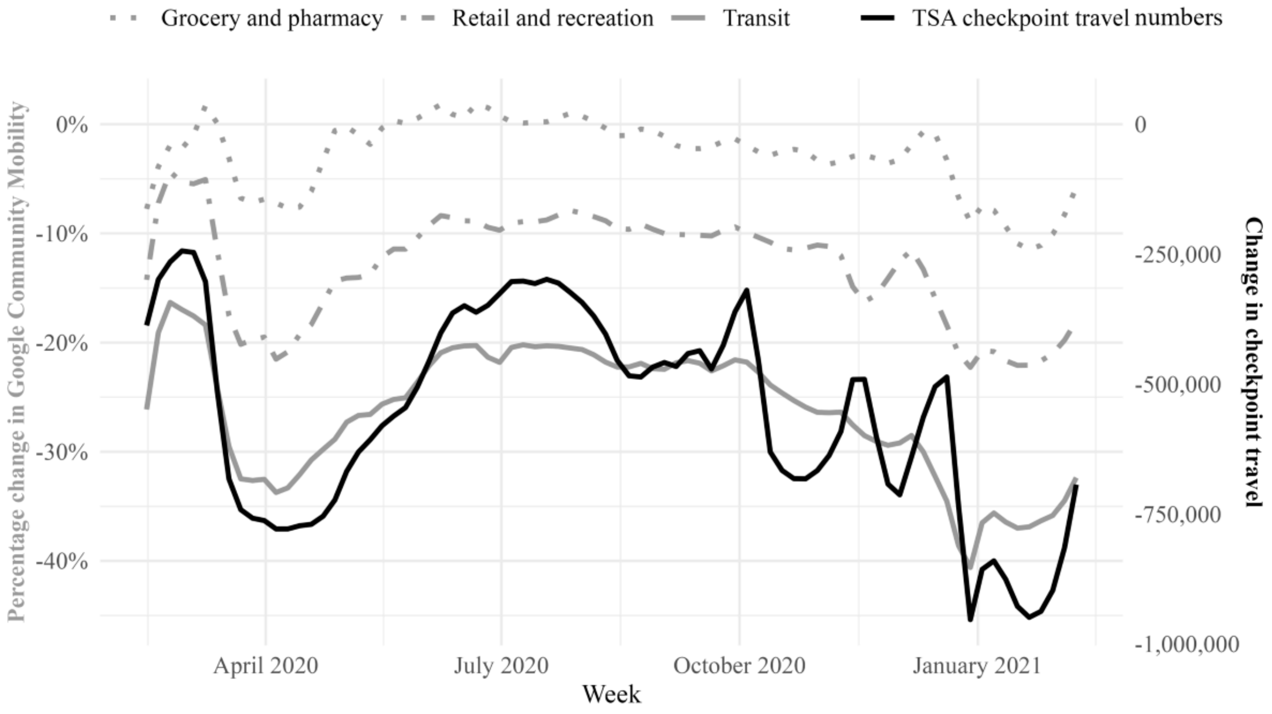
[c] Aggregated weekly change score compared to baseline weeks (January 3, 2020, to February 6, 2020).

A consistent visual pattern was observed when we inspected checkpoint travel data during the 10-week period before and after the emergence of new variants or surge in cases over the course of the study period (Multimedia Appendix 3). New variants included Alpha B.1.1.7 (emerged December 29, 2020), Delta B.1.617.2 (emerged June 1, 2021, and surged July 27, 2021), and Omicron B.1.1.529 (emerged December 1, 2021), with surges of both Omicron and Delta variants identified on January 1, 2022 [19]. Increased checkpoint travel numbers reflected a rise in air travel during the winter holiday and summer months. A rise in checkpoint travel numbers directly preceded surges in new variants of COVID-19: Alpha B.1.1.7 (December 2020), Delta B.1.617.2 (July 2021), and Omicron B.1.1.529 and Delta B.1.617.2 (December 2021). A drop in checkpoint travel numbers immediately followed each surge before rising again in subsequent months.

Correlations Between Checkpoint Travel Data and Community Mobility Data

The relationship between the checkpoint travel measure and community mobility measure was visually inspected (Figure 2).

Figure 2. Change in weekly checkpoint travel volume and community mobility estimates compared to baseline weeks preceding the COVID-19 pandemic (January 3, 2020, to February 6, 2020). TSA: Transportation Security Administration.



There was a strong positive, significant correlation between the weekly change from the baseline period for checkpoint travel data and community mobility data for transit stations and retail and recreation, and moderate significant correlations for grocery and pharmacy as well as parks (Table 1). A strong negative correlation was observed for places of residence, and a weak but significant positive correlation was found for workplaces (Table 1).

First-order differencing of changes in weekly checkpoint travel and community mobility from baseline had moderate positive and significant correlations for transit stations, retail and recreation, grocery and pharmacy, and parks. A weak but significant negative correlation was observed for places of residence and workplaces (Table 1).

Discussion

Principal Findings

The TSA's daily travel checkpoint data provide a publicly available, flexible metric that may be easily incorporated into national data sets to serve as a proxy for the intensity of COVID-19 activity in the population at a given time interval. As of this writing, these data are available on the internet and span back to 2019. It is unknown whether these data will continue to be tracked and published on the internet indefinitely.

Future data collection efforts may consider including a self-reported measure on social distancing in population-based surveys for as long as new strains of COVID-19 are expected to impact social dynamics to a varying extent over time. Public health officials have noted that society continues to be vulnerable to other emergent and resurgent diseases with pandemic potential [20]. This measure may be applicable to other pandemic diseases requiring self-isolation and quarantine at a broad scale.

Changes in checkpoint travel data used in our study correlated in the expected direction with changes in community mobility, with strong correlations observed for transit stations, and retail and recreation sites. A moderate correlation was observed with places to access essential goods and services (ie, grocery stores and pharmacies) as well as parks, which were popular outdoor recreational options during the pandemic. A weak correlation was observed for workplaces, likely reflecting the adoption of remote work policies by many employers. As expected, a strong negative correlation was observed between checkpoint travel data and being at home. Google ceased publication of its COVID-19 Community Mobility Reports (used in this study) as of October 15, 2022, yet daily checkpoint travel data remain available. Mobility data are subject to proprietary data processing methods and policies, including changes in parameters used by Google in the change score calculation over time (for example, to account for shifts in regional population size or retailer density). For this reason, Google cautions against comparing time periods that span longer than 6 months. By contrast, checkpoint travel data do not rely on third parties to generate estimates and can be applied to data sets spanning multiple years. Checkpoint travel data are available in the years preceding the COVID-19 pandemic, whereas the mobility data used in our study could only be viewed as a percent change from the prepandemic baseline determined by the company.

First-order differencing analysis at a weekly level found few significant correlations between checkpoint travel data and self-reported social distancing, and weak to moderate correlations with community mobility data. These results suggest that the measure is less useful as a control for time intervals with high granularity, such as time series studies examining week-to-week changes within the COVID-19 period. Similarly, Google cautions against using its Community Mobility Reports to compare day-to-day activity and calculates its change score from a median value from a pre-COVID baseline period spanning several weeks. The use of checkpoint travel data as a proxy for COVID-19 activity is, therefore, best applied to analyses that need to control for time-dependent trends attributable to the COVID-19 pandemic, such as analyses that encompass the prepandemic and pandemic periods, or the first several months of the pandemic when lockdowns occurred compared to periods later in the pandemic after societal activities began to resume.

Limitations

This study was subject to limitations. First, the daily air travel metrics of interest to this study were validated against a national sample of youths and young adults, rather than a nationally representative data set of the general population. While we are encouraged that the strength of the correlation was similar within defined subgroups in our sample, these subgroups remain restricted to people younger than 25 years. It is unclear how similar the social distancing practices of young people in our sample would be to those of the general adult population.

Second, the metric against which checkpoint travel data were validated (Google Community Mobility Data) is only available for Google account users with Location History settings activated and thus is also subject to an unknown degree and direction of bias.

Finally, while initially checkpoint travel numbers declined sharply following the emergence of COVID-19 and society's attempts to limit its transmission, later in the pandemic a predictive pattern was evident between a rise in checkpoint travel numbers and subsequent COVID-19 variant surges. A rise in seasonal travel (ie, holidays and summer vacations) appeared to prompt surges in the incidence of emerging COVID-19 variants. The steep drop in checkpoint travel numbers immediately after each COVID-19 surge may have been driven in part by individuals' desire to avoid further transmission, but it is also likely that these are normal seasonal fluctuations after a popular travel period. These data illustrate the extent to which air travel not only reflects but almost certainly contributes to COVID-19 transmission. Therefore, although they may serve as an appropriate proxy for COVID-19 dynamics in the population, checkpoint travel numbers are not a completely independent indicator of this phenomenon.

Conclusions

We recommend using the TSA's checkpoint travel data in longitudinal or time series studies that span the COVID-19 pandemic period. This variable can control for the potential confounding effect of COVID-19 activity in population-based studies of a range of public health topics and within demographic subgroups.

Acknowledgments

JMK contributed to this paper's conceptualization, analytic plan, interpretation of the data analysis, and writing and editing of this paper. KA contributed to the literature review and writing and editing of this paper. HMT contributed to the development of the analytic plan, data analysis, interpretation of data analysis, and writing of this paper. ECH provided critical reviews of the results. ECH and DMV supervised the project. The authors received no financial support for the research, authorship, or publication of this paper.

Data Availability

The Transportation Security Administration's checkpoint travel data and Google's COVID-19 Community Mobility Reports are publicly available [1,2]. Investigators seeking access to the self-reported social distancing survey data used in this study should make a written request to Truth Initiative authors and submit a detailed research plan including the purpose of the proposed research, required variables, duration of the analysis phase, institutional review board approval with Federalwide Assurance information, and investigator training in human subjects. A data sharing agreement is required for use of all data. Truth Initiative does not share data with tobacco industry representatives or affiliated researchers. Approved investigators must access data sets via an analytic Portal owned and administered by Truth Initiative.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Weekly sum of travelers screened at Transportation Security Administration (TSA) checkpoints, January 2019-May 2022.
[DOCX File , 20 KB - publichealth_v9i1e44950_app1.docx]

Multimedia Appendix 2
Proportion of youth and young adults reporting social distancing during the past week (approximately 280 unique respondents per week, N=45,080), March 2020-May 2022.
[DOCX File , 27 KB - publichealth_v9i1e44950_app2.docx]

Multimedia Appendix 3

Weekly percent of subgroups of respondents not social distancing: (A) age (15-17 years, 18-20 years, 21-24 years), (B) lower socioeconomic status, and (C) racial/ethnic minorities.

[PNG File , 129 KB - publichealth_v9i1e44950_app3.png]

Multimedia Appendix 4

Weekly checkpoint travel volume and the emergence of new COVID-19 variants, 10 weeks before and after surges of (A) Alpha B.1.1.7, (B) Delta B.1.617.2, and (C) Omicron B.1.1.529 and Delta B.1.617.2.

[PNG File , 150 KB - publichealth_v9i1e44950_app4.png]

References

1. WHO director-general's opening remarks at the media briefing on COVID-19-11 March 2020. World Health Organization. 2020. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2022-08-24]
2. Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. Lancet Infect Dis 2020;20(6):678-688 [FREE Full text] [doi: 10.1016/S1473-3099(20)30162-6] [Medline: 32213332]
3. West R, Michie S, Rubin GJ, Amlôt R. Applying principles of behaviour change to reduce SARS-CoV-2 transmission. Nat Hum Behav 2020;4(5):451-459 [FREE Full text] [doi: 10.1038/s41562-020-0887-9] [Medline: 32377018]
4. States that issued lockdown and stay-at-home orders in response to the coronavirus (COVID-19) pandemic, 2020. Ballotpedia. 2020. URL: https://ballotpedia.org/States_that_issued_lockdown_and_stay-at-home_orders_in_response_to_the_coronavirus_(COVID-19)_pandemic,_2020 [accessed 2022-07-19]
5. Auger KA, Shah SS, Richardson T, Hartley D, Hall M, Warniment A, et al. Association Between Statewide School Closure and COVID-19 Incidence and Mortality in the US. JAMA 2020;324(9):859-870 [FREE Full text] [doi: 10.1001/jama.2020.14348] [Medline: 32745200]
6. Tuttle KR. Impact of the COVID-19 pandemic on clinical research. Nat Rev Nephrol 2020;16(10):562-564 [FREE Full text] [doi: 10.1038/s41581-020-00336-9] [Medline: 32760016]
7. Buckee CO, Balsari S, Chan J, Crosas M, Dominici F, Gasser U, et al. Aggregated mobility data could help fight COVID-19. Science 2020;368(6487):145-146. [doi: 10.1126/science.abb8021] [Medline: 32205458]
8. Halford E, Dixon A, Farrell G, Malleson N, Tilley N. Crime and coronavirus: social distancing, lockdown, and the mobility elasticity of crime. Crime Sci 2020;9(1):11 [FREE Full text] [doi: 10.1186/s40163-020-00121-w] [Medline: 32834925]
9. Sulyok M, Walker M. Community movement and COVID-19: a global study using Google's Community Mobility Reports. Epidemiol Infect 2020;148:e284 [FREE Full text] [doi: 10.1017/S0950268820002757] [Medline: 33183366]
10. Kishore N, Kiang MV, Engø-Monsen K, Vembar N, Schroeder A, Balsari S, et al. Measuring mobility to monitor travel and physical distancing interventions: a common framework for mobile phone data analysis. Lancet Digit Health 2020;2(11):e622-e628 [FREE Full text] [doi: 10.1016/S2589-7500(20)30193-X] [Medline: 32905027]
11. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020;20(11):1247-1254 [FREE Full text] [doi: 10.1016/S1473-3099(20)30553-3] [Medline: 32621869]
12. McKee KL, Crandell IC, Hanlon AL. County-Level Social Distancing and Policy Impact in the United States: A Dynamical Systems Model. JMIR Public Health Surveill 2020;6(4):e23902 [FREE Full text] [doi: 10.2196/23902] [Medline: 33296866]
13. Bourassa KJ, Sbarra DA, Caspi A, Moffitt TE. Social Distancing as a Health Behavior: County-Level Movement in the United States During the COVID-19 Pandemic Is Associated with Conventional Health Behaviors. Ann Behav Med 2020;54(8):548-556 [FREE Full text] [doi: 10.1093/abm/kaaa049] [Medline: 32608474]
14. Dobbie LJ, Hydes TJ, Alam U, Tahrani A, Cuthbertson DJ. The Impact of the COVID-19 Pandemic on Mobility Trends and the Associated Rise in Population-Level Physical Inactivity: Insights From International Mobile Phone and National Survey Data. Front Sports Act Living 2022;4:773742 [FREE Full text] [doi: 10.3389/fspor.2022.773742] [Medline: 35368411]
15. Taking stock of the pandemic’s impact on global aviation. McKinsey & Company. URL: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/taking-stock-of-the-pandemics-impact-on-global-aviation [accessed 2022-09-29]
16. COVID-19 takes a bite out of 2020 passenger transportation. Bureau of Transportation Statistics. URL: https://www.bts.gov/data-spotlight/covid-19-takes-bite-out-2020-passenger-transportation [accessed 2022-09-29]
17. TSA checkpoint travel numbers (current year versus prior year(s)/same weekday). Transportation Security Administration. URL: https://www.tsa.gov/coronavirus/passenger-throughput [accessed 2022-09-29]
18. COVID-19 community mobility reports. Google LLC. URL: https://www.google.com/covid19/mobility?hl=en [accessed 2023-03-17]

19. CDC museum COVID-19 timeline. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/museum/timeline/covid19.html [accessed 2022-09-02]
20. Morens DM, Fauci AS. Emerging Pandemic Diseases: How We Got to COVID-19. Cell 2020;182(5):1077-1092 [FREE Full text] [doi: 10.1016/j.cell.2020.08.021] [Medline: 32846157]

Abbreviations

TSA: Transportation Security Administration

Edited by A Mavragani, T Sanchez; submitted 09.12.22; peer-reviewed by C Kakalou, H Du; comments to author 25.01.23; revised version received 30.03.23; accepted 16.05.23; published 29.08.23.

Please cite as:
Kreslake JM, Aarvig K, Muller-Tabanera H, Vallone DM, Hair EC
Checkpoint Travel Numbers as a Proxy Variable in Population-Based Studies During the COVID-19 Pandemic: Validation Study
JMIR Public Health Surveill 2023;9:e44950
URL: https://publichealth.jmir.org/2023/1/e44950
doi:10.2196/44950
PMID:37191643

Original Paper

Unmet Psychosocial Needs of Health Care Professionals in Europe During the COVID-19 Pandemic: Mixed Methods Approach

Svenja Hummel[1]; Ina Michelsen[1], BA; Ali Zafar[1], MA; Steffen Moritz[2], Prof Dr; Charles Benoy[3], PhD; Cédric Lemogne[4], MD, PhD; Rosa Almeida[5], MSc; Raquel Losada[5], MSc; Oscar Ribeiro[6], PhD; Vincenza Frisardi[7], MD, MBA; Ilaria Tarricone[8], MD, PhD; Silvia Ferrari[9], MD, PhD; Frieder Dechent[10], MD; Christian G Huber[10], Prof Dr; Steffi Weidt[11], MD, PD; Gwendolyn Mayer[1], Dr sc hum; Jobst-Hendrik Schultz[1], MD, PD

[1]Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg, Germany

[2]Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[3]Centre Hospitalier Neuro-Psychiatrique Luxembourg, Ettelbrück, Luxembourg

[4]Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), Service de Psychiatrie de l'adulte, AP-HP, Hôpital Hôtel-Dieu, Paris, France

[5]INTRAS Research, Innovation and Knowledge Unit, Intras Foundation, Valladolid, Spain

[6]Center for Health Technology and Services Research (CINTESIS) at the Associate Laboratory RISE - Health Research Network, Department of Education and Psychology, University of Aveiro, Aveiro, Portugal

[7]Geriatric Unit, IRCCS-AOU BO, Policlinico San't Orsola, Bologna, Italy

[8]Department of Medical and Surgical Sciences, Bologna University, Bologna, Italy

[9]Department of Biomedical, Metabolic and Neural Sciences, University of Modena & Reggio Emilia, Modena, Italy

[10]University Psychiatric Clinics Basel, University of Basel, Basel, Switzerland

[11]Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric University Hospital Zurich and University of Zurich, Zurich, Switzerland

Corresponding Author:
Svenja Hummel
Department of General Internal Medicine and Psychosomatics
Heidelberg University Hospital
Im Neuenheimer Feld 410
Heidelberg
Germany
Phone: 49 62215634170
Email: svenja.hummel@med.uni-heidelberg.de

Abstract

Background: The COVID-19 pandemic severely affected everyday life and working conditions for most Europeans, particularly health care professionals (HCPs). Over the past 3 years, various policies have been implemented in various European countries. Studies have reported on the worsening of mental health, work-related stress, and helpful coping strategies. However, having a closer look is still necessary to gain more information on the psychosocial stressors and unmet needs of HCPs as well as nonmedical staff.

Objective: This study aimed to obtain quantitative information on job-related stressors of physicians and nurses and the coping strategies of HCPs and nonmedical staff at 2 periods of the COVID-19 pandemic. By further analyzing qualitative comments, we wanted to gain more information on the psychosocial stressors and unmet needs of HCPs as well as nonmedical staff on different levels of experience.

Methods: A cross-sectional survey was conducted at 2 time points during the COVID-19 pandemic in several European countries. The first study period (T1) lasted between April 1 and June 20, 2020, and the second study period (T2) lasted between November 25, 2021, and February 28, 2022. On a quantitative level, we used a questionnaire on stressors for physicians and nurses and a questionnaire on coping strategies for HCPs and nonmedical staff. Quantitative data were descriptively analyzed for mean values and differences in stressors and coping strategies. Qualitative data of free-text boxes of HCPs and nonmedical staff were analyzed via thematic analysis to explore the experiences of the individuals.

Results: T1 comprised 609 participants, and T2 comprised 1398 participants. Overall, 296 participants made 438 qualitative comments. The uncertainty about when the pandemic would be controlled (T1: mean 2.28, SD 0.85; T2: mean 2.08, SD 0.90) and the fear of infecting the family (T1: mean 2.26, SD 0.98; T2: mean 2.02, SD 1.02) were the most severe stressors identified by physicians and nurses in both periods. Overall, the use of protective measures (T1: mean 2.66, SD 0.60; T2: mean 2.66, SD 0.60) and acquiring information about COVID-19 (T1: mean 2.29, SD 0.82; T2: mean 1.99, SD 0.89) were identified as the most common coping strategies for the entire study population. Using thematic analysis, we identified 8 themes of personal experiences on the micro, meso, and macro levels. Measures, working conditions, feelings and emotions, and social climate were frequently mentioned topics of the participants. In T1, feelings of isolation and uncertainty were prominent. In T2, feelings of exhaustion were expressed and vaccination was frequently discussed. Moreover, unmet psychosocial needs were identified.

Conclusions: There is a need for improvement in pandemic preparedness. Targeted vocational education measures and setting up of web-based mental health support could be useful to bridge gaps in psychosocial support needs in future crises.

(JMIR Public Health Surveill 2023;9:e45664) doi:10.2196/45664

KEYWORDS

COVID-19; mental health; health care professionals; health care workers; pandemic preparedness; mixed methods; coping; stressors; psychosocial

Introduction

Background

The COVID-19 pandemic had a drastic impact on daily life and working conditions in Europe. By mid-March 2020, preventive measures such as physical distancing were implemented across European countries such as France, Italy, Spain, the United Kingdom, and Belgium. In many countries, a shortage of health care resources required reasonable distribution and sustainable use of personal protective equipment (PPE), human resources, and intensive care unit (ICU) beds. Triaging patients with COVID-19 became necessary. The collapse of the health care system proved serious in some countries such as Italy. Spain was more affected than neighboring countries, such as Portugal, which had fewer reported cases in March 2020. Before COVID-19, there were 9.7 ICU beds per 100,000 people in Spain versus 33.9 ICU beds per 100,000 people in Germany [1]. At the beginning of the pandemic, Austria, Germany, and Switzerland were successful in managing the crisis and reported lower mortality rates than Belgium, the United Kingdom, France, Spain, and Italy. Early efforts were made to contain the number of cases, such as the implementation of information websites. In Germany, for example, health care system preparedness was improved by rapidly creating new ICU resources rather than repurposing them, as was the case in Switzerland and Austria. Furthermore, there was a large increase in video consultations in terms of telehealth. However, in some cases, policies in different federal states within the country were not consistent [2].

As virus transmission subsided, eventually, these policies were eased [3]. From June to September 2020, borders were reopened [4] and detailed testing protocols were established for health care professionals (HCPs) [5]. Italy was one of the first countries to mandate vaccinations for HCPs in May 2021. By September 2021, mandatory vaccination protocols for HCPs were adopted by Greece and France [6], followed by Germany [7]. Austria became the first country to implement mandatory vaccinations for all people. Italy and Greece followed the mandatory vaccination for people above a certain age [8]. In some European countries such as Germany, France, and Italy, only completely vaccinated individuals or those who had already been infected with COVID-19 had access to particular public spaces [7]. Emergency measures such as curfews, school closures, and a shift to working from home or job losses are a few examples of the challenges that people had to face. With the introduction of vaccines, the topic of vaccine hesitancy became a focus of attention [9]. As many new policies were introduced during the COVID-19 pandemic, it was not surprising that the general population in Europe was distressed and, in many cases, had mental health problems [10]. On the one hand, a COVID-19 infection directly impacted mental health [11], whereas on the other hand, there was the deterioration of mental health among the general population [12] as an indirect effect of the pandemic, for example, in terms of psychosocial problems. A psychosocial problem can originate in a variety of domains (eg, environmental problems, occupational problems, inadequate social support, educational problems, inadequate access to health and other services, and interpersonal losses), which can have a negative influence on the individual and might cause mental illnesses [13].

The COVID-19 pandemic posed major challenges for health care systems and HCPs worldwide.

For European HCPs, there was an increase in workload [14]. In the nursing profession, staff shortages and increased workloads were already commonplace before the pandemic [15] and, during COVID-19, the burnout of nurses was identified [16]. The negative impact of mental health on HCPs due to COVID-19 has already been proven [17,18], whereby nurses and frontline HCPs were particularly strained psychologically. Nevertheless, in a study by Hummel et al [19], we found significantly lower mean scores for depression and anxiety among HCPs than among nonmedical staff.

As nurses and frontline HCPs (eg, physicians) seem to be the most affected groups during the pandemic, it is worth focusing on these professions regarding work-related stressful events (hereafter referred to as stressors and psychosocial burden). Stress occurs when the demands overwhelm the person [20]. Several stressors of HCPs were identified during the severe acute respiratory syndrome (SARS) outbreak in 2003 [21] and during the COVID-19 pandemic [22]. In 1 study, nurses rated the fear of infection as the most important stressor, followed by death [23]. Coping can be understood as an adaptation that helps the individual deal with challenges that exceed their capacities [20]. Coping strategies are, in our context, to be understood as profession-independent ways of dealing with the strain of the COVID-19 pandemic in general. Taking personal protective measures was a frequently used coping strategy by nurses and physicians along with talking to family and friends or doing relaxation activities in a study on the COVID-19 pandemic by Rose et al [22].

By adding qualitative data, one might gain more insights into the participants' way of thinking than by using only the results of quantitative data. Responses to such questions can differ in length and detail, and the participants have the opportunity to explain themselves in detail. Open-ended survey questions help to supplement the quantitative results by providing additional information [24], which can be helpful in exploring psychosocial factors and unmet needs. According to social work research, individuals can be studied at a macro, meso, and micro level. The micro level describes the level of the individual, the meso level describes the interaction between groups, and the macro level addresses social structures and institutions [25]. Considering the deterioration of HCPs and the nonmedical staff's mental health during the COVID-19 pandemic, assessments are needed with different levels of individual experiences (macro, meso, and micro) to identify existing gaps in psychosocial care.

Objectives

This study aimed to gain quantitative information on work-related stressors of physicians and nurses and the coping strategies of HCPs and nonmedical staff during 2 periods of the COVID-19 pandemic. By further analyzing the qualitative comments, we obtained more information on the psychosocial stressors and unmet needs of HCPs as well as nonmedical staff at different levels of experience.

Methods

Study Design

A cross-sectional web-based survey was conducted at 2 different stages during the COVID-19 pandemic. The first period lasted from April 1 to June 20, 2020, during the first lockdown in Europe [26]. The second phase lasted between November 25, 2021, and February 28, 2022, when the second and third lockdowns ended in Europe [27,28].

Ethics Approval

The study was approved by the ethics committee of the Heidelberg University Medical Faculty (S-361/2020).

Data Collection, Informed Consent, and Participation

Data collection was conducted in compliance with the European General Data Protection Regulation. The survey questionnaire was distributed in 10 European countries, and all health care workers and associated staff at hospitals as well as nonmedical staff were eligible to participate. Consent to participate was included in the web-based questionnaire and obtained without a signature. No allowance was provided for participation in the survey. All the questionnaires were completed anonymously. Data security was granted using the Secure Sockets Layer–encrypted platform, SoSci Survey [29].

Questionnaire

A questionnaire comprising 3 sections on demographics, stress factors, and coping strategies was administered. As no COVID-19–specific questionnaires were available at the beginning of the pandemic, we derived our instrument from a study by Lee et al [21] on SARS. The original 23-item questionnaire focused on work-related stressors for nurses during SARS. Therefore, we modified the questionnaire to include physicians and nurses working in hospitals [19]. For the stress factors section, participants indicated on a Likert scale from 0 (not at all) to 4 (very much) how often they thought about or were concerned about 23 specific stressors in their everyday lives or at clinical work. Only physicians and nurses were eligible to complete this section.

The section on coping strategies was also derived from the study by Lee et al [21]. Participants responded to 12 items on a scale from 0 (almost never) to 3 (almost always) to assess how frequently these 12 coping strategies were applied in their everyday lives. This section was available to all participants, as it did not focus on medical issues. Considering the utility of gaining qualitative data, at the end of the questionnaire, we included an open-ended question with free-text boxes for the participants to leave one or more comments at the end of the survey. To ensure neutral responses within the context of the study objectives, the question was framed as, "If there is anything else you would like to tell us, you can do so here."

Recruitment

The questionnaire was translated by native speakers or professional translators into several European languages, including English, Italian, German, French, Spanish, and Portuguese, to obtain a larger sample population. Accordingly, the questionnaire was also distributed in the respective countries. It was made available on the web via the platform SoSci Survey [29]. The link to the survey was distributed via email to our personal and professional networks using the snowball sampling method. Invitation emails were sent to colleagues at the Heidelberg University Hospital and further distributed to related institutions and to European contacts with their partner organizations, hospitals, and professional associations. Participants were also recruited via personal networks or public social networking groups, such as Twitter, LinkedIn, and Facebook.

Participants

The HCPs included physicians and nurses, as well as a group of *other job in health care* comprising health care psychologists, physiotherapists, other nursing professions, laboratorians, technicians, and occupational therapists. The nonmedical staff was a heterogeneous group, which, for example, consisted of administrators, managers, engineers, teachers, retired persons, or secretaries. However, not all nonmedical participants revealed their actual profession.

Data Analysis

Overview

Quantitative and qualitative data were analyzed independently. In the first study period, the sample comprised 609 participants, of whom 78 submitted 111 qualitative comments. In the second phase, 1398 participants completed the survey, with 218 participants submitting 327 qualitative comments. Overall, 438 qualitative comments were made by 296 participants, of whom 31.4% (n=93) were nurses, 17.2% (n=51) were physicians, and 19.9% (n=59) were nonmedical staff. As the stressors' questionnaire was specific for people working at hospitals, we only analyzed the answers from physicians and nurses for this questionnaire. The coping questionnaire was analyzed for the entire study population. For not fulfilling the inclusion criteria, 69 cases were excluded from the data analysis of the first phase, whereas 41 cases were excluded from the second period. The population demographics and quantitative results for the first study period were published earlier [19].

Quantitative Data

Missing data occurred only for the stressors section during the first study period. Questionnaires with >10% missing data per participant were excluded using a pragmatic and rigor approach following previous recommendations [30]. In total, we analyzed 346 questionnaires for stressors in the first period and 696 in the second period. For coping strategies, 609 questionnaires were analyzed in the first period and 1398 in the second period.

Means and SDs were calculated for all items on stressors for physicians and nurses and coping strategies for all participants. We used Cronbach α to determine the internal consistency of the stressors and coping scales for both study periods. Cronbach α >.80 was considered a threshold for acceptance [31,32]. For the stressors questionnaire, the internal consistency of the first study period was assessed from a Cronbach α of .92 and that of the second study period was assessed from a Cronbach α of .93. For the coping questionnaire, the internal consistency of the first study period was assessed from a Cronbach α of .74 and of the second study period was assessed from a Cronbach α of .72. *t* tests (2-tailed) were conducted for independent groups to compare the mean values of stressors and coping items of the 2 study periods. In all analyses, *P* values <.05 were considered statistically significant. Quantitative data were analyzed using SPSS (version 26; IBM Corp) [33].

Qualitative Data

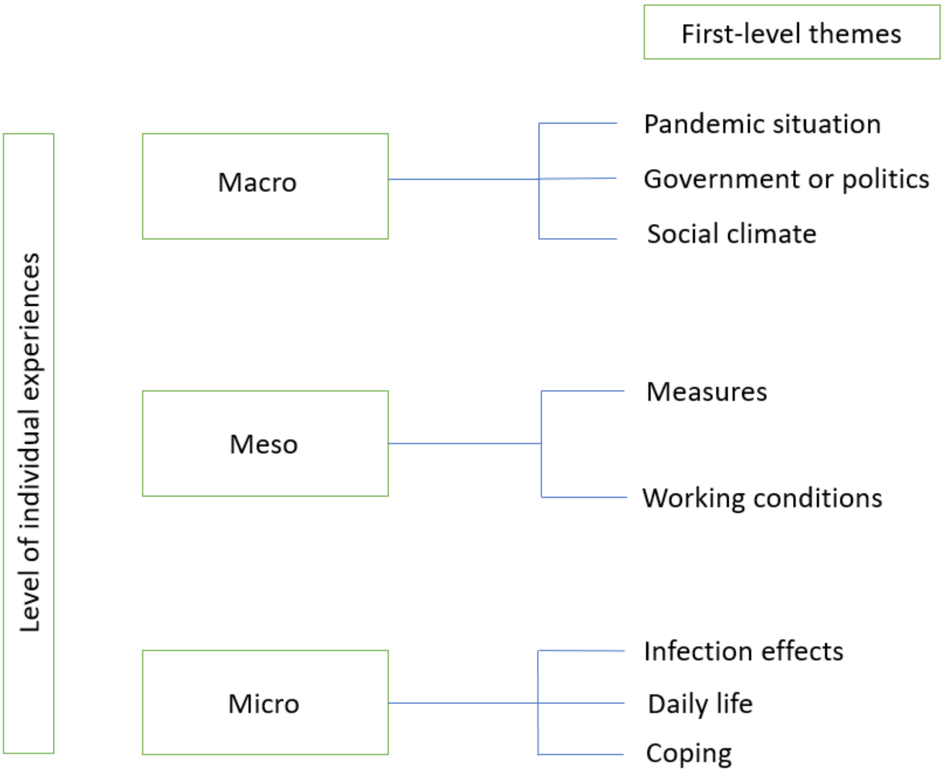
All comments in the free-text comment boxes by HCPs and nonmedical staff were considered for analysis. We provided the demographics of our qualitative data with means and SDs for age, frequencies, and percentages. The length of the comments made by the participants ranged from 3 words (eg, "not enough staff") to >100 words. The comments were translated and analyzed for each study period separately using inductive thematic analysis [34] to identify themes that tell a story about what concerns people beyond quantitative data. We used the 6-step approach developed by Braun and Clarke [35].

Codes and resulting themes were elaborated by 2 authors (SH and IM). After each step, the authors discussed the results to create a common theme map. Second-level and third-level themes were developed when the first-level theme revealed multiple aspects. The key themes identified were located at different levels in relation to the individual based on the levels of experience. Therefore, the themes were divided into *macro*, *meso*, and *micro* levels based on recommendations for social work research [25].

As the *pandemic situation*, *government or politics*, and *social climate* themes described the overarching sociopolitical impacts of COVID-19, these themes were categorized to the macro level. The topics *measures* and *working conditions* could be located at a more individual impact level and were assigned the meso level. The micro level included topics dealing with the individual and included the themes *daily life*, *coping*, and *infection effects* (Figure 1).

The 8 key themes and the most important second- and third-level themes are presented in this paper, with relevant quotes to illustrate the themes. The complete list of themes is available in Multimedia Appendix 1 [36].

Figure 1. First-level themes at their assigned levels of individual experience.



Results

Quantitative Data

Stressors for Physicians and Nurses

For both study periods, *Uncertainty about when the epidemic will be under control* was rated as the most stressful factor, followed by *Worry about inflicting COVID-19 on family*. In the first period, *Worry about nosocomial spread* and *Frequent modification of infection control procedures* were the third and fourth biggest stressors, respectively. In the second period, *Worry about lack of manpower* was the third and *Frequent modification of infection control procedures* was the fourth biggest stressor. The highest significant mean difference ($t_{752.58}$=10.09, *P*<.001) between first and second study period was found for *Worry about nosocomial spread* (first study period: mean 2.04, SD 0.91; second study period: mean 1.41, SD 1.01; mean difference 0.63, 95% CI 0.50-0.75).

Another item with the most significant mean difference ($t_{656.97}$=7.39, *P*<.001) was *Worry about lack of proper knowledge and equipment* (first study period: mean 1.66, SD 1.04; second study period: mean 1.17, SD 1.17; mean difference 0.50, 95% CI 0.36-0.63). The mean values and 2-tailed *t* test results for all the items are presented in Table 1.

Table 1. Stressors for nurses and physicians during COVID-19 for both study periods (first n=346 and second n=696).

Items[a]	First study period		Second study period		*t* test (*df*)	*P* value
	Values, n (%)	Values, mean (SD)	Values, n (%)	Values, mean (SD)		
Uncertainty about when the epidemic will be under control	345 (99.7)[b]	2.28 (0.85)[b]	696 (100)[b]	2.08 (0.90)[b]	3.48 (1039)	.001[c]
Worry about inflicting COVID-19 on family	346 (100)[b]	2.26 (0.98)[b]	696 (100)[b]	2.02 (1.02)[b]	3.60 (1040)	<.001[c]
Worry about nosocomial spread	345 (99.7)[b]	2.04 (0.91)[b]	696 (100)	1.41 (1.01)	10.09 (752.58)	<.001[c]
Frequent modification of infection control procedures	345 (99.7)[b]	2.03 (0.88)[b]	696 (100)[b]	1.73 (0.96)[b]	5.03 (741.61)	<.001[c]
Protective gears cause physical discomfort	345 (99.7)	1.75 (1.02)	696 (100)	1.50 (1.06)	3.57 (1039)	<.001[c]
Deterioration of patients' condition	345 (99.7)	1.70 (1.00)	696 (100)	1.52 (1.04)	2.70 (1039)	.01[c]
Worry about lack of proper knowledge and equipment	346 (100)	1.66 (1.04)	696 (100)	1.17 (0.99)	7.39 (656.97)	<.001[c]
Worry about being negligent and endangering patients	346 (100)	1.66 (1.07)	696 (100)	1.25 (1.04)	6.03 (1040)	<.001[c]
Worry about getting infected	346 (100)	1.62 (1.03)	696 (100)	1.41 (1.00)	3.10 (1040)	.002[c]
Patients' emotional reaction	346 (100)	1.57 (0.96)	696 (100)	1.53 (0.99)	0.61 (1040)	.54
Worry about lack of manpower	346 (100)	1.57 (1.05)	696 (100)[b]	1.97 (1.04)[b]	−5.86 (1040)	<.001[c]
Documentation and reporting procedures unclear	345 (99.7)	1.54 (1.01)	696 (100)	1.26 (1.00)	4.17 (1039)	<.001[c]
Patient families' emotional reaction	345 (99.7)	1.52 (1.01)	696 (100)	1.44 (1.01)	1.26 (1039)	.21
Coworkers being emotionally unstable	346 (100)	1.53 (0.97)	696 (100)	1.59 (1.00)	−0.92 (1040)	.36
Being without properly fitted environment	346 (100)	1.51 (1.08)	696 (100)	1.27 (1.06)	3.47 (1040)	.001[c]
Conflict between duty and safety	346 (100)	1.48 (1.07)	696 (100)	1.30 (1.00)	2.66 (649.35)	.01[c]
Worry about being negligent and endangering coworkers	346 (100)	1.48 (1.04)	696 (100)	0.98 (0.97)	7.44 (652.43)	<.001[c]
Be infected by the colleagues	346 (100)	1.30 (1.01)	696 (100)	1.10 (0.97)	3.08 (662.88)	.002[c]
Protective gears being a drag in providing quality care	346 (100)	1.28 (1.05)	696 (100)	1.15 (1.04)	1.93 (1040)	.05
Coworkers displaying COVID-19–like symptoms	346 (100)	1.25 (0.97)	696 (100)	1.09 (0.96)	2.54 (1040)	.01[c]
Equivocal definition of the responsibility between physicians and nurses	346 (100)	1.19 (1.04)	696 (100)	1.08 (1.02)	1.57 (1040)	.12
Yourself displaying COVID-19-like symptoms	346 (100)	1.12 (1.04)	696 (100)	1.05 (0.98)	1.00 (1040)	.32
Blaming from commanding officers	345 (99.7)	0.70 (0.95)	696 (100)	0.85 (1.04)	−2.38 (1039)	.02[c]

[a] Mean and SD for the question: "When you think about COVID-19 in your life and work, how often did you think or worry about the following things?" (0=not at all and 3=very much).

[b] Frequently used stressor.

[c] *P* value <.05 was considered statistically significant.

Coping Strategies of the Whole Population

Taking protective measures (washing hands, wearing mask, taking own temperature, etc) was the most often used coping strategy, followed by *Actively acquiring more knowledge about COVID-19 (symptoms, transmission pathway, etc)* in both study periods. In the first period, the third most reported coping strategy was *Video-chatting with family and friends by phone to share concerns and support*, whereas in the second period, it was *Engaging in recreational activities (web-based shopping, social media, internet surfing...)*.

The highest significant mean difference ($t_{1218.57}$=11.14, *P*<.001) between the first and second periods was found for *Video-chatting with family and friends by phone to share concerns and support* (first study period: mean 1.85, SD 0.88; second study period: mean 1.36, SD 0.93; mean difference 0.49, 95% CI 0.40-0.57). Another item with a significant mean difference (t_{2005}=7.09, *P*<.001) was *Actively acquiring more knowledge about COVID-19 (symptoms, transmission pathway, etc*; first study period: mean 2.29, SD 0.82; second study period: mean 1.99, SD 0.89; mean difference 0.30, 95% CI 0.22-0.38). The mean values and 2-tailed *t* test results for all the items are presented in Table 2.

Table 2. Coping strategies during COVID-19 for the whole study population of both study periods (first n=609 and second n=1398).

Items[a]	First study period		Second study period		*t* test (*df*)	*P* value
	Values, n (%)	Values, mean (SD)	Values, n (%)	Values, mean (SD)		
Taking protective measures (washing hands, wearing mask, taking own temperature, etc)	609 (100)[b]	2.66 (0.60)[b]	1398 (100)[b]	2.66 (0.60)[b]	0.06 (2005)	.95
Actively acquiring more knowledge about COVID-19 (symptoms, transmission pathway, etc)	609 (100)[b]	2.29 (0.82)[b]	1398 (100)[b]	1.99 (0.89)[b]	7.09 (2005)	<.001[c]
Video-chatting with family and friends by phone to share concerns and support	609 (100)[b]	1.85 (0.88)[b]	1398 (100)	1.36 (0.93)	11.14 (1218.57)	<.001[c]
Engaging in recreational activities (web-based shopping, social media, internet surfing...)	609 (100)	1.67 (0.94)	1398 (100)[b]	1.57 (0.90)[b]	2.18 (2005)	.03[c]
Engaging in health-promoting behaviors (more rest, exercise, balanced diet, etc)	609 (100)	1.63 (0.97)	1398 (100)	1.41 (0.99)	4.57 (2005)	<.001[c]
Switching thoughts and facing the situations with positive attitude	609 (100)	1.58 (0.88)	1398 (100)	1.45 (0.92)	2.93 (2005)	.003[c]
Limiting oneself from watching too much news about COVID-19	609 (100)	1.42 (0.98)	1398 (100)	1.55 (0.99)	−2.71 (2005)	.01[c]
Distracting oneself from thinking about COVID-19 issues by suppression or keeping busy	609 (100)	1.35 (0.94)	1398 (100)	1.28 (0.94)	1.62 (2005)	.11
Acquiring mental health knowledge and information	609 (100)	1.10 (0.97)	1398 (100)	1.13 (0.97)	−0.63 (2005)	.53
Practicing relaxation methods (meditation, yoga, Tai Chi, etc)	609 (100)	0.57 (0.91)	1398 (100)	0.76 (0.97)	−4.04 (1226.25)	<.001[c]
Venting emotions by crying, screaming, smashing things, and so on	609 (100)	0.54 (0.85)	1398 (100)	0.44 (0.76)	2.48 (1065.30)	.01[c]
Using alcohol or drugs	609 (100)	0.35 (0.65)	1398 (100)	0.39 (0.66)	−1.45 (1203.59)	.15

[a] Mean and SD for the question “When you think about COVID-19 in your life and work. How often did you use or try to use the following methods to handle the situation?” (from 1=almost never to 4=almost ever) for all participants in the order of their frequency of use.

[b] Frequently used coping strategy.

[c] *P* value <.05 was considered statistically significant.

Qualitative Results

Study Participants

Table 3 provides the sociodemographic characteristics of the study population for the qualitative analysis of each study period. For the overall population, the age range was 22 to 76 years.

Table 3. Demographic characteristics of the study participants in the first period, second period, and total.

Characteristics	First study period (n=78)	Second study period (n=218)	Total (n=296)
Age (years), mean (SD)	46.12 (12.74)	44.53 (11.09)	45.33 (11.92)
Gender, n (%)			
Man	26 (33.3)	48 (22)	74 (25)
Woman	52 (66.7)	169 (77.5)	221 (74.7)
Nonbinary	0 (0)	1 (0.5)	1 (0.3)
Country, n (%)			
United Kingdom	11 (14.1)	0 (0)	11 (3.7)
Germany	18 (23.1)	35 (16.1)	53 (17.9)
Austria	5 (6.4)	0 (0)	5 (1.7)
Switzerland	5 (6.4)	25 (11.5)	30 (10.1)
France	10 (12.8)	76 (34.9)	86 (29.1)
Italy	12 (15.4)	7 (3.2)	19 (6.4)
Spain	13 (16.7)	4 (1.8)	17 (5.7)
Portugal	4 (5.1)	2 (0.9)	6 (2)
Belgium	0 (0)	16 (7.3)	16 (5.4)
Luxemburg	0 (0)	53 (24.3)	53 (17.9)
Language, n (%)			
English	12 (15.4)	1 (0.5)	13 (4.4)
German	27 (34.6)	84 (38.5)	111 (37.5)
Italian	12 (15.4)	7 (3.2)	19 (6.4)
Spanish	13 (16.7)	4 (1.8)	17 (5.7)
Portuguese	4 (5.1)	3 (1.4)	7 (2.4)
French	10 (12.8)	119 (54.6)	129 (43.6)
Profession, n (%)			
Health care professional			
Physicians	20 (25.6)	31 (14.2)	51 (17.2)
Volunteers	1 (1.3)	0 (0)	1 (0.3)
Dentist	1 (1.3)	0 (0)	1 (0.3)
Nurse	17 (21.8)	76 (34.9)	93 (31.4)
Other job in health care	13 (16.7)	78 (35.8)	91 (30.7)
Nonmedical	26 (33.3)	33 (15.1)	59 (19.9)

Themes

From the thematic analysis, 8 key themes emerged, namely *pandemic situation*, *government or politics*, *social climate*, *measures*, *working conditions*, *infection effects*, *daily life*, and *coping*. Among these key themes, 27 second-level themes were identified. For some second-level themes, third-level themes were also identified. A total of 3 themes were assigned to the macro level of experience, 2 referred to the meso level, and 3 referred to the micro level of experience. Most topics were related to the societal impact of the COVID-19 pandemic; one topic addressed the impact of the COVID-19 infection itself on the individual.

Macro Level

Pandemic Situation

Some participants talked about aspects that came along with the pandemic situation in general. One aspect was uncertainty in terms of not knowing how the pandemic would develop. Another point is the feeling of isolation. Comments referring to isolation revolved mostly around social isolation due to social distancing. Many participants voiced criticism of the vaccine itself, for example, on its effectiveness. Other criticisms targeted mandatory vaccinations.

> *I have a feeling of emptiness, seeing my life just floating by and a lot of uncertainty in the future...a*

> *lot of loneliness and a strong feeling of not being able to move on because there is no security in the future.* [Female, aged 57 years, nonmedical staff; clerk, first study period]

Government or Politics

Many participants criticized the management of the pandemic by their governments or politicians. Frequently mentioned issues were the lack of trust in politicians and dissatisfaction with political decisions. On the other hand, there were participants who were satisfied with the governmental crisis management:

> *...Anger at executives for failing to act proactively, particularly over mask distribution and lack of testing. Anger at leaders for conflicting announcements, late, vague pronouncements or actions aimed at getting us back to work at full steam to get the economy back on its feet before the pandemic is quite over.* [Female, aged 69 years, nonmedical staff; legal secretary, first study period]

Social Climate

The participants frequently discussed the social climate. The social climate was indirectly reflected in the comments of many participants, who either directly described it or expressed its effects through comments such as fear, anger, frustration, or annoyance. Many participants talked about a deterioration in emotional well-being unrelated to their working conditions. The unmet needs were mostly discussed during the second study period. Societal split as a social response was directly mentioned by some participants, primarily in relation to the decision to vaccinate. The topic of radicalization was also related to attitudes toward vaccination. Many participants described the phenomenon of questionable information—we call this an "infodemic," according to the World Health Organization (WHO) [36]—within society, for example, through the media.

> *People make me sick!* [Female, aged 33 years, nonmedical staff; administrator, first study period]
>
> *I'm so annoyed and can't hear the word corona anymore.* [Female, aged 57 years, nurse, second study period]

Meso Level

Measures

Measures in general other than vaccination were frequently discussed by the participants. Topics included the implementation of and reaction to the measures. Participants voiced their complaints about side effects, for example, of wearing masks. There were also comments on the impact of the measures on daily life or work regarding their implementation and inconsistencies.

> *Wearing the mask for more than 8 hours a day doesn't exactly help, but we have to live with that for now.* [Female, aged 47 years, other job in health care; clinical pharmaceutical laboratory analyst, second study period]

Working Conditions

Working conditions were a frequently mentioned topic. There were comments on feelings or emotions such as the feeling of pressure, for example, due to high workload and lack of appreciation. Structural changes at work due to the pandemic were mentioned by all profession groups. The HCPs reported on how patient care had changed. Some reported thoughts of quitting their jobs or having already done so. The need for protection and supply was mentioned, for example, in terms of psychological support.

As the *other job in health care* group was very heterogeneous, there were also comments on nursing experiences given by this group. Participants from all professions reported how well or poorly they felt they had been cared for by their employers.

> *The pressure on the nursing staff has increased greatly as a result of the pandemic and there is little support from politicians and employers.* [Male, aged 40 years, nurse, second study period]
>
> *I miss citizen support and acceptance. It cannot be that nursing staff is denied access to grocery stores.* [Female, aged 56 years, nurse, first study period]
>
> *Recognizing of caregivers only at the peak of the pandemic, and then when the pandemic ends, it is over.* [Female, aged 52 years, nurse, second study period]

Micro Level

Infection Effects

The experiences of participants who had a COVID-19 infection or experiences about the infection of acquaintances or relatives also emerged, including the long-term effects of a past infection.

> *To date I have been diagnosed with long COVID. It's been a year now since I was infected. I still have after-effects [loss of taste and smell, shortness of breath, cough, fatigue, joint pain].* [Female, aged 46 years, nurse, second study period]

Daily Life

Many participants wrote about the upcoming challenges in their daily lives due to the COVID-19 pandemic, mentioning caring for children at home, dealing with vaccine opponents in everyday life and reduced social life in general. However, positive aspects, such as an increase in family activities, were also reported.

> *I'm in a privileged situation, have a good family life, a big house with lots of outdoor space...* [Female, aged 52 years, dentist, first study period]

Coping

Some participants reported on how they tried to deal with the pandemic. Specifically, exercising and positive thinking seemed to have helped the participants.

> *In terms of positive feelings, I have been able to adapt by maintaining a healthy lifestyle and exercising at home, taking time to read and after a few weeks appreciating a slower pace of life [before I was very*

> *frantic and out of the house all day], I have not felt any fear of contracting the disease, nor have I become obsessed with the subject [I have followed recommended prevention guidelines].* [Female, aged 34 years, nonmedical staff; occupational therapist, first study period]

First Versus Second Period

Supporting quotes for the 2 periods are available in Multimedia Appendix 2. For the first study period, participants talked about isolation due to social distancing and voiced a lack of support and protection, for example, regarding PPE. The sense of uncertainty and insecurity was evident in many comments, but there were also comments on the advantages of curfews. During the second study period, vaccination and its effect on society were frequently mentioned topics. The implementation and impact of the COVID-19 pandemic measures were also discussed. Working conditions, measures, social climate, and feelings and emotions were mentioned during both the first and second study periods. All second-level and third-level themes that emerged in the data of the first study period could also be found in the second study period, but some of the second-level themes came up only in the second phase, for example, vaccination and splitting and radicalization of the society. In addition, unmet psychological needs mostly emerged during the second study period.

> *I don’t feel adequately protected by the protective clothing that we currently have and I am afraid of the day when we will no longer have any protective clothing.* [Female, nurse, aged 27 years, first study period]
>
> *OP's started again...letting everything go back to normal during this time means more work for the same thing.* [Female, aged 30 years, nurse, second period]

Discussion

Principal Findings

In this study, we investigated work-related stressors for physicians and nurses and the coping strategies of both HCPs and nonmedical staff at 2 different periods during the COVID-19 pandemic. In addition, qualitative data were analyzed to gain more information on psychosocial stress factors and unmet needs at different levels of experience of the population during the COVID-19 pandemic crisis.

On the quantitative scale, the uncertainty about when the pandemic will end and the fear of infecting the family were the worst stressors for physicians and nurses at both periods. The most frequent coping strategies for the entire study population were the use of protective measures and active acquisition of knowledge about COVID-19. Significant mean differences were found for stressors such as *Worry about nosocomial spread* ($t_{752.58}$=10.09, P<.001) and coping strategies such as *Video-chatting with family and friends by phone to share concerns and support* ($t_{1218.57}$=11.14, P<.001).

On the qualitative level, at the beginning of the pandemic, the feeling of insecurity and uncertainty was present, and participants talked about protective equipment and how to handle the social distancing situation. During the second study period, unmet psychosocial needs, exhaustion, and vaccination were frequently mentioned. In the second phase, the topics were more widely spread, as many participants used the free-text fields to “process” their experiences in the previous years.

Overall, these results provide a deeper insight into the perceived problems, burdens, and challenges that our participants faced during the pandemic. By sorting the themes to different levels of experience, the identification of targets for possible necessary psychosocial support was easier. It turned out that the need for psychosocial interventions existed and was even possible to implement at the macro level (eg, more governmental support), at the meso level (eg, working conditions and support of employers), and at the micro level (eg, psychological support).

Changing of Themes Over the Study Periods

Analyzing the themes over 2 study periods, a shift in topics within the population was observed. Social climate, measures, emotions, and working conditions were frequently discussed during both study periods, but there were different focuses in some places. When the pandemic situation was new to the people and protective gear was in short supply [14], the themes of social distancing reflected feelings of isolation and a general feeling of insecurity and uncertainty. In the second stage, the sentiments of exhaustion and annoyance were more prominent. Vaccination, in general, was a frequently discussed topic in the second period, which in some cases can be connected to comments on the increasing splitting and radicalization of society. An earlier longitudinal study on the challenges of the pandemic found a kind of indignation among their participants in the later stages of the study, whereas at the beginning, participants tended to express a sense of tiredness and monotony [37].

Comments from the second study period mentioned the consequences of the pandemic, such as side effects of the measures or their implementation. Apart from vaccinations, unmet needs, such as additional psychosocial support, mostly emerged in the second study period. A possible explanation might be that in the first study period, the whole situation was new and unknown, so people were busier trying to handle the situation rather than thinking about what could be helpful to them.

This shift in topics illustrates the usefulness of studies at different periods when attempting to understand the psychosocial situation of society during the course of a pandemic. People were overwhelmed by the situation and were more concerned with basic care issues at the beginning, whereas psychosocial burdens and needs seemed to emerge later in time.

Work Stressors of HCPs and Working Conditions of All Participants

Uncertainty about control over the pandemic and possible infection of family members was the most relevant stressor on a quantitative level for physicians and nurses during both study periods. For the first period, the third and fourth most relevant stressors were concerns about nosocomial spread and constant change in infection control procedures. The measures and their implementation were also discussed in the qualitative comments, especially in the second study period. Considering how often measures, guidelines, and policies have changed [38-40] during the COVID-19 pandemic following the rising or falling infection rates, this concern is easy to understand. Several federal states within a country with differing policies have made this even more complicated [2]. The stressors *Worry about lack of proper knowledge and equipment* and *Worry about nosocomial spread* seemed to be less important for the participants in the second study period, as in the first study period, PPE was in short supply [14].

Working conditions were frequently mentioned in the comments, emphasizing the importance of this theme to participants. As the *other job in health care* group was very heterogeneous, this group also commented on their nursing experiences. One theme was the lack of appreciation, which was frequently mentioned by nursing people. Appreciation at work was proven to be especially important when there were interruptions at work [41], which might be the case in terms of the COVID-19 pandemic. An increased workload, as well as a lack of support from employers, was reported. Feelings of exhaustion, pressure, and thoughts about quitting the job were mostly reported by HCPs. An increase in workload due to the COVID-19 pandemic for HCPs has already been discussed [14], as well as the decline in mental health [17]. Some participants reported how caring for patients has changed or how work processes, in general, have changed due to the pandemic. In addition, telehealth and telework were discussed as working from home increased, especially during the first period of the pandemic [9]. Working from home can indicate a decrease in mental well-being status [42], and it also seemed to be a challenge for some participants in our study. Another European study reported that there were more people who wanted to work from home from the end of 2020 to the beginning of 2021 [9]. However, at this point, the opinions of our participants differed, as many of them also expressed criticism of telemedicine or teleworking regarding the irreplaceability of personal contacts. Staff shortages and high workloads have been a longstanding problem in nursing [15], and there was even more workload during the pandemic [14], as it was also reflected in our study. The fear of staff shortages was also represented on a quantitative level as the third most frequent stressor in the second study period. This feeling might be aggravated by the high COVID-19 infection rates of HCPs in general [3], which worsened again during the omicron wave at the beginning of 2022 [43]. As nurses have been identified as a vulnerable group regarding mental health during the pandemic [17], our results indicated that this professional group needs to be heard for its unmet psychological support.

Vaccination in general and mandatory vaccinations for HCPs was a huge topic in the second study period. This was associated with feelings of being forced and having limited freedom. Regarding working conditions, the emerging themes in our study represent the areas where better preparedness is necessary, for example, proper support by employers, reliable information, employee protection, for example, in the sense of PPE but also regarding workload and work processes, appreciation, more options to combine family and career, and proper psychological support. A European study on working conditions reported a decrease in mental well-being, a decrease in work-life balance for young parents, an increase in people in a financially fragile situation, tiredness with regard to home schooling, and job losses due to the pandemic [9].

Coping Strategies

When considering the results of the questionnaire on coping strategies, low Cronbach α, which is below the threshold of acceptance, must be considered. The most frequently rated stressors were reflected in the most frequently rated coping strategies. The best coping strategies for the entire study population were to protect themselves and gather more information about the virus. By comparing quantitative data with qualitative comments on coping, the importance of digital offers became clear. The third most often used coping strategy in the first study period was *Video-chatting with family and friends by phone to share concerns and support*, and in the second study period, it was *Engaging in recreational activities (web-based shopping, social media, internet surfing...)*. The use of video-chatting may have decreased during the pandemic, as we found a significant decrease in the mean values of the first and second study periods. One reason for this could be the possibility of returning to personal contact during the pandemic. However, it should be noted that few participants also emphasized the non-substitutability of personal contacts by digital contact. On a qualitative level, some participants shared how they dealt with the pandemic situation, for example, by going out, trying to relax, doing hobbies on the web, doing sports, positive thinking, or being with the family. Gaining information was one of the most often used coping strategy on a quantitative level and not being sufficiently informed was also often voiced by participants on a qualitative level, which might be another important approach. Clear information is important for mental health, especially for HCPs [44]. In this context, vocational and educational training for pandemic preparedness of HCPs might be a helpful approach that has already been initiated by various organizations such as the WHO [45].

Psychosocial Approaches for Future Pandemic Preparedness

Our results emphasize the necessity of an improvement regarding *pandemic preparedness*, which is defined by the WHO as having plans and resources in place to actively respond to a pandemic. It includes, for example, prevention, detection, and containment measures but also plans to respond to any shortages that may arise. Preparing for possible future crises beforehand using the lessons learned from the current pandemic is necessary to face future challenges [46]. In this study, some participants wished for more psychological support. One participant reported that no mental or relaxing therapies were offered at work, whereas another voiced that they were not supported on a psychological level. The stated needs are in line with findings from another study on the psychosocial impact of the COVID-19 pandemic, where the authors argued that support services for future pandemics should be implemented [47].

Previous studies have reported similar needs for HCPs, calling for mental health support services for HCPs in particular [17]. A participant in our study mentioned video- or web-based interventions to help HCPs dealing with stress. The emerging use of e-mental health might be one way to address these psychosocial needs in future crises, as the effectiveness of e-mental health interventions has already been proven [48], and the pandemic has accelerated further advancements [49]. However, it should be noted that few participants also emphasized the non-substitutability of personal contacts by digital contact. Regarding the need for sufficient information, vocational and educational training may be an important approach for future pandemic preparedness. *Infodemic* can have adverse side effects [36] and clear and rapid information for the general population [46] and for HCPs [44], in particular, is important. For future psychometric approaches, attention to the different macro, meso, and micro levels is relevant when collecting psychosocial quantitative data, which might be realized by the combination of qualitative and quantitative approaches.

Limitations

Several limitations of this study should be considered when interpreting these findings. First, as the questions in our study questionnaire focused on the negative aspects or problems of the pandemic and the open-ended question was very unspecific, the open comments at the end of the study may also be biased toward negative statements or inappropriate statements. Second, the link to the web-based survey was distributed via social media and via the personal and professional networks of the authors. As the contact networks in the individual European countries were not equally strong and web-based distribution was difficult to control, the number of participants for each country was different. Despite the large sample size, these results cannot be generalized for Europe as our participants came from only 10 countries out of the 27 European Union countries. Furthermore, only a small subset of the entire population provided qualitative feedback. Third, the different groups within the HCP group of nurses, physicians, and other professionals in health care organizations were somewhat heterogeneous, as well as the group of nonmedical staff, of whom not all revealed their actual profession. We have attempted to explain these groups in as much detail as possible.

Moreover, the study phases did not always occur at the exact time of the lockdowns or peaks in infection rates in individual European countries, which may have influenced the results. At best, there was a partial overlap between respondents in the first and second study periods. The validation of questionnaires on stressors and coping should be part of future studies.

Conclusions

This mixed methods study provides valuable insights into the individual and psychosocial problems faced by European HCPs as well as nonmedical staff over the course of 3 years of the COVID-19 pandemic. Open-ended questions supplemented our quantitative surveys to identify the core problems on different levels of experience of the individuals. Using thematic analysis of qualitative data, we identified 8 major themes of experiences in the pandemic situation: *government or politics*, *social climate*, *measures*, *working conditions*, *infection effects*, *daily life*, and *coping mechanisms*. These different levels of experiences could be assigned to the macro, meso, and micro levels of social structures.

Uncertainty about the end of the pandemic and the fear of infecting family members were identified as the most important stress factors for physicians and nurses. The most commonly used coping strategies of the entire study population were the use of protective measures and acquiring up-to-date information on the COVID-19 pandemic. In the first study period, social distancing, uncertainty, and the need for protection were the topics mentioned. During the second study period, topics such as vaccination, exhaustion, and unmet psychosocial needs emerged. Working conditions were frequently discussed by all participants during both periods. An improvement in pandemic preparedness, with emphasis on vulnerable groups such as HCP in general and nurses in particular, is needed. Several psychosocial approaches should be considered for future research, for example, the development of easily accessible digital psychosocial services and educational and vocational training.

Acknowledgments

The authors thank all the participants of the 2 study periods for their willingness to share their personal views and experiences with them. For the publication fee, the authors acknowledge financial support from Deutsche Forschungsgemeinschaft within the funding program “Open Access Publikationskosten” as well as by Heidelberg University. In addition, the authors are thankful to Mrs Eva Quintero for her support in translation and data organization and Mr Maximilian Selic for his support in data analysis.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of the themes with supporting quotes.
[DOCX File , 24 KB - publichealth_v9i1e45664_app1.docx]

Multimedia Appendix 2
Supporting quotes for the first versus second study period.
[DOCX File , 22 KB - publichealth_v9i1e45664_app2.docx]

References

1. Miralles O, Sanchez-Rodriguez D, Marco E, Annweiler C, Baztan A, Betancor É, et al. Unmet needs, health policies, and actions during the COVID-19 pandemic: a report from six European countries. Eur Geriatr Med 2021 Feb;12(1):193-204 [FREE Full text] [doi: 10.1007/s41999-020-00415-x] [Medline: 33057981]
2. Desson Z, Lambertz L, Peters JW, Falkenbach M, Kauer L. Europe's Covid-19 outliers: German, Austrian and Swiss policy responses during the early stages of the 2020 pandemic. Health Policy Technol 2020 Dec;9(4):405-418 [FREE Full text] [doi: 10.1016/j.hlpt.2020.09.003] [Medline: 33520639]
3. Rapid risk assessment: Coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK– ninth update. European Centre for Disease Prevention and Control. 2020. URL: https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-coronavirus-disease-2019-covid-19-pandemic-ninth-update [accessed 2023-03-29]
4. Report offers first comprehensive analysis of the travel restrictions and border closures taken globally to battle COVID-19, looks ahead as mobility resumes. Migration Policy Institute. 2021 Apr 08. URL: https://www.migrationpolicy.org/news/comprehensive-analysis-covid19-state-of-global-mobility [accessed 2023-03-29]
5. COVID-19 testing strategies and objectives. European Centre for Disease Prevention and Control. 2020 Sep 18. URL: https://www.ecdc.europa.eu/en/publications-data/covid-19-testing-strategies-and-objectives [accessed 2023-03-29]
6. Walter JD. Mandatory COVID vaccines stoke controversy across Europe. Deutsche Welle. 2021 Jun 11. URL: https://www.dw.com/en/mandatory-covid-vaccines-a-controversy-across-europe/a-59742720 [accessed 2023-03-29]
7. Legal issues surrounding compulsory COVID-19 vaccination. Think Tank European Parliament. 2022 Mar 14. URL: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)729309 [accessed 2023-03-29]
8. Buchholz K. The countries where COVID-19 vaccination is mandatory. Statista. 2022 Feb 08. URL: https://www.statista.com/chart/25326/obligatory-vaccination-against-covid-19/ [accessed 2023-03-29]
9. Ahrendt D, Mascherini M, Nivakoski S, Sándor E. Living, working and COVID-19 (update April 2021): mental health and trust decline across EU as pandemic enters another year. European Foundation for the Improvement of Living and Working Conditions. 2021. URL: https://www.eurofound.europa.eu/publications/report/2021/living-working-and-covid-19-update-april-2021-mental-health-and-trust-decline-across-eu-as-pandemic [accessed 2023-03-29]
10. Santini ZI, Koyanagi A. Loneliness and its association with depressed mood, anxiety symptoms, and sleep problems in Europe during the COVID-19 pandemic. Acta Neuropsychiatr 2021 Jun;33(3):160-163. [doi: 10.1017/neu.2020.48] [Medline: 33427135]
11. Vindegaard N, Benros ME. COVID-19 pandemic and mental health consequences: systematic review of the current evidence. Brain Behav Immun 2020 Oct;89:531-542 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.048] [Medline: 32485289]
12. Hossain MM, Tasnim S, Sultana A, Faizah F, Mazumder H, Zou L, et al. Epidemiology of mental health problems in COVID-19: a review. F1000Res 2020 Jun 23;9:636 [FREE Full text] [doi: 10.12688/f1000research.24457.1] [Medline: 33093946]
13. Skodol AE. Axis IV: a reliable and valid measure of psychosocial stressors? Compr Psychiatry 1991 Nov;32(6):503-515 [FREE Full text] [doi: 10.1016/0010-440x(91)90029-c] [Medline: 1778077]
14. Winkelmann J, Webb E, Williams GA, Hernández-Quevedo C, Maier CB, Panteli D. European countries' responses in ensuring sufficient physical infrastructure and workforce capacity during the first COVID-19 wave. Health Policy 2022 May;126(5):362-372 [FREE Full text] [doi: 10.1016/j.healthpol.2021.06.015] [Medline: 34311982]
15. Griffiths P, Dall'Ora C, Simon M, Ball J, Lindqvist R, Rafferty AM, RN4CAST Consortium. Nurses' shift length and overtime working in 12 European countries: the association with perceived quality of care and patient safety. Med Care 2014 Nov;52(11):975-981 [FREE Full text] [doi: 10.1097/MLR.0000000000000233] [Medline: 25226543]
16. Galanis P, Vraka I, Fragkou D, Bilali A, Kaitelidou D. Nurses' burnout and associated risk factors during the COVID-19 pandemic: a systematic review and meta-analysis. J Adv Nurs 2021 Aug;77(8):3286-3302 [FREE Full text] [doi: 10.1111/jan.14839] [Medline: 33764561]
17. Shaukat N, Ali DM, Razzak J. Physical and mental health impacts of COVID-19 on healthcare workers: a scoping review. Int J Emerg Med 2020 Jul 20;13(1):40 [FREE Full text] [doi: 10.1186/s12245-020-00299-5] [Medline: 32689925]
18. Vizheh M, Qorbani M, Arzaghi SM, Muhidin S, Javanmard Z, Esmaeili M. The mental health of healthcare workers in the COVID-19 pandemic: a systematic review. J Diabetes Metab Disord 2020 Oct 26;19(2):1967-1978 [FREE Full text] [doi: 10.1007/s40200-020-00643-9] [Medline: 33134211]

19. Hummel S, Oetjen N, Du J, Posenato E, Resende de Almeida RM, Losada R, et al. Mental health among medical professionals during the COVID-19 pandemic in eight European countries: cross-sectional survey study. J Med Internet Res 2021 Jan 18;23(1):e24983 [FREE Full text] [doi: 10.2196/24983] [Medline: 33411670]
20. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. New York, NY: Springer; 1984.
21. Lee SH, Juang YY, Su YJ, Lee HL, Lin YH, Chao CC. Facing SARS: psychological impacts on SARS team nurses and psychiatric services in a Taiwan general hospital. Gen Hosp Psychiatry 2005 Sep;27(5):352-358 [FREE Full text] [doi: 10.1016/j.genhosppsych.2005.04.007] [Medline: 16168796]
22. Rose S, Hartnett J, Pillai S. Healthcare worker's emotions, perceived stressors and coping mechanisms during the COVID-19 pandemic. PLoS One 2021 Jul 09;16(7):e0254252 [FREE Full text] [doi: 10.1371/journal.pone.0254252] [Medline: 34242361]
23. Lorente L, Vera M, Peiró T. Nurses´ stressors and psychological distress during the COVID-19 pandemic: the mediating role of coping and resilience. J Adv Nurs 2021 Mar;77(3):1335-1344 [FREE Full text] [doi: 10.1111/jan.14695] [Medline: 33210768]
24. Jackson KM, Trochim WM. Concept mapping as an alternative approach for the analysis of open-ended survey responses. Organ Res Methods 2002 Oct;5(4):307-336 [FREE Full text] [doi: 10.1177/109442802237114]
25. Mauldin RL. Foundations of Social Work Research. Arlington, TX: Mavs Open Press; 2020.
26. Dambeck H. So unterschiedlich meisterte Europa die erste Welle. Spiegel. 2020 Aug 15. URL: https://www.spiegel.de/wissenschaft/corona-pandemie-so-unterschiedlich-meisterte-europa-die-erste-welle-a-dca7cabf-8a3b-4bbc-a776-50384285969a [accessed 2023-05-11]
27. Bletzinger M. Coronavirus: Dritte Welle! Hier schlägt sie bereits um sich - Neuer Anstieg jetzt wohl noch viel gefährlicher. Merkur.de. 2023 Jul 24. URL: https://www.merkur.de/welt/coronavirus-dritte-welle-deutschland-europa-infektion-zahlen-mutation-tschechien-oesterreich-ungarn-zr-90239302.html [accessed 2023-05-11]
28. Merlot J, Stotz P. Die zweite Welle rollt. Spiegel. 2020. URL: https://www.spiegel.de/wissenschaft/medizin/coronavirus-in-europa-die-zweite-welle-rollt-a-1d5b12a1-162d-48a3-8e1e-40235c996080 [accessed 2023-05-11]
29. SoSci survey – the solution for professional online questionnaires. SoSci survey. 2023. URL: https://www.soscisurvey.de [accessed 2022-03-01]
30. Bannon WJ. Missing data within a quantitative research study: how to assess it, treat it, and why you should care. J Am Assoc Nurse Pract 2015 Apr;27(4):230-232. [doi: 10.1002/2327-6924.12208] [Medline: 25676704]
31. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572 [FREE Full text] [doi: 10.1136/bmj.314.7080.572] [Medline: 9055718]
32. Schmitt N. Uses and abuses of coefficient alpha. Psychol Assess 1996 Dec;8(4):350-353 [FREE Full text] [doi: 10.1037/1040-3590.8.4.350]
33. IBM: IBM SPSS statistics for Windows, version 26. IBM Corp. 2019.
34. Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. Nurse Educ Today 2017 Sep;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
36. Managing the COVID-19 infodemic: promoting healthy behaviours and mitigating the harm from misinformation and disinformation. World Health Organization. 2020. URL: https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation [accessed 2023-04-04]
37. Maison D, Jaworska D, Adamczyk D, Affeltowicz D. The challenges arising from the COVID-19 pandemic and the way people deal with them. A qualitative longitudinal study. PLoS One 2021 Oct 11;16(10):e0258133 [FREE Full text] [doi: 10.1371/journal.pone.0258133] [Medline: 34634054]
38. Corona-Maßnahmen. Neue Lockerungen treten in Kraft. Presse- und Informationsdienst der Bundesregierung. 2022. URL: https://www.bundesregierung.de/breg-de/themen/coronavirus/corona-lockerungen-2009902 [accessed 2023-05-01]
39. COVID in Europe: Netherlands to reopen shops, gyms, and hairdressers. Euronews. 2022. URL: https://www.euronews.com/2022/01/13/covid-19-spike-felt-across-europe-as-vaccination-remains-stagnant [accessed 2023-05-01]
40. Chadwick L. How COVID-19 upended life in Europe throughout 2020. Euronews. 2020 Dec 31. URL: https://www.euronews.com/2020/12/31/how-covid-19-upended-life-in-europe-throughout-2020 [accessed 2023-05-01]
41. Stocker D, Keller AC, Meier LL, Elfering A, Pfister IB, Jacobshagen N, et al. Appreciation by supervisors buffers the impact of work interruptions on well-being longitudinally. Int J Stress Manag 2019 Nov;26(4):331-343 [FREE Full text] [doi: 10.1037/str0000111]
42. Xiao Y, Becerik-Gerber B, Lucas G, Roll SC. Impacts of working from home during COVID-19 pandemic on physical and mental well-being of office workstation users. J Occup Environ Med 2021 Mar 01;63(3):181-190 [FREE Full text] [doi: 10.1097/JOM.0000000000002097] [Medline: 33234875]

43. Assessment of the further spread and potential impact of the SARS-CoV-2 Omicron variant of concern in the EU/EEA, 19th update. European Centre for Disease Prevention and Control. 2022. URL: https://www.ecdc.europa.eu/en/publications-data/covid-19-omicron-risk-assessment-further-emergence-and-potential-impact [accessed 2023-03-29]
44. Giorgi G, Lecca LI, Alessio F, Finstad GL, Bondanini G, Lulli LG, et al. COVID-19-related mental health effects in the workplace: a narrative review. Int J Environ Res Public Health 2020 Oct 27;17(21):7857 [FREE Full text] [doi: 10.3390/ijerph17217857] [Medline: 33120930]
45. Coronavirus disease (COVID-19) training: online training. World Health Organization. 2023. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/training/online-training [accessed 2023-04-07]
46. Burwell SM, Townsend FF, Bollyky TJ, Patrick SM. Independent task force report no. 78 - improving pandemic preparedness, lessons from COVID-19. Council on Foreign Relations. URL: https://www.cfr.org/report/pandemic-preparedness-lessons-COVID-19/pdf/TFR_Pandemic_Preparedness.pdf [accessed 2023-04-07]
47. Dubey S, Biswas P, Ghosh R, Chatterjee S, Dubey MJ, Chatterjee S, et al. Psychosocial impact of COVID-19. Diabetes Metab Syndr 2020 Sep;14(5):779-788 [FREE Full text] [doi: 10.1016/j.dsx.2020.05.035] [Medline: 32526627]
48. Mayer G, Schultz JH. Depression per App behandeln? Möglichkeiten und Grenzen digitaler Gesundheitsanwendungen (DiGAs). Psychotherapie im Dialog 2021 Aug 27;22(03):54-58 [FREE Full text] [doi: 10.1055/a-1215-1563]
49. Ellis LA, Meulenbroeks I, Churruca K, Pomare C, Hatem S, Harrison R, et al. The application of e-mental health in response to COVID-19: scoping review and bibliometric analysis. JMIR Ment Health 2021 Dec 06;8(12):e32948 [FREE Full text] [doi: 10.2196/32948] [Medline: 34666306]

Abbreviations

HCP: health care professional
ICU: intensive care unit
PPE: personal protective equipment
SARS: severe acute respiratory syndrome
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 11.01.23; peer-reviewed by C Wu, A AL-Asadi; comments to author 20.03.23; revised version received 18.05.23; accepted 26.06.23; published 06.09.23.

Please cite as:
Hummel S, Michelsen I, Zafar A, Moritz S, Benoy C, Lemogne C, Almeida R, Losada R, Ribeiro O, Frisardi V, Tarricone I, Ferrari S, Dechent F, Huber CG, Weidt S, Mayer G, Schultz JH
Unmet Psychosocial Needs of Health Care Professionals in Europe During the COVID-19 Pandemic: Mixed Methods Approach
JMIR Public Health Surveill 2023;9:e45664
URL: https://publichealth.jmir.org/2023/1/e45664
doi:10.2196/45664
PMID:37672320

Original Paper

Risk Factors for Not Completing a 2-Dose Primary Series of Messenger RNA COVID-19 Vaccination in a Large Health Care System in Southern California: Retrospective Cohort Study

Stanley Xu[1,2], PhD; Vennis Hong[1], MPH; Lina S Sy[1], MPH; Katia J Bruxvoort[1,3], PhD; Bruno Lewin[1], MD; Bing Han[1], PhD; Kimberly J Holmquist[1], MPH; Lei Qian[1], PhD

[1]Department of Research & Evaluation, Southern California Permanente Medical Group, Pasadena, CA, United States
[2]Department of Health Systems Science, Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, CA, United States
[3]School of Public Health, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:
Stanley Xu, PhD
Department of Research & Evaluation
Southern California Permanente Medical Group
100 S Los Robles
Pasadena, CA, 91101
United States
Phone: 1 626 564 3958
Fax: 1 626 564 3409
Email: stan.xu@kp.org

Abstract

Background: COVID-19 vaccination is crucial in combating the COVID-19 pandemic. Messenger RNA COVID-19 vaccines were initially authorized as a 2-dose primary series and have been widely used in the United States; completing the 2-dose primary series offers protection against infection, severe illness, and death. Understanding the risk factors for not completing the 2-dose primary series is critical to evaluate COVID-19 vaccination programs and promote completion of the 2-dose primary series.

Objective: This study examined potential risk factors for not completing a 2-dose primary series of mRNA COVID-19 vaccination.

Methods: We conducted a retrospective cohort study among members aged ≥18 years from a large integrated health care system, Kaiser Permanente Southern California, from December 14, 2020, to June 30, 2022. Noncompletion of the 2-dose primary series was defined as not completing the second dose within 6 months after receipt of the first dose. Crude noncompletion rates were estimated overall and by demographic characteristics, health care use patterns, comorbidity, and community-level socioeconomic factors. A Poisson regression model was fit to examine associations of individual-level and community-level risk factors with noncompletion of the 2-dose primary series.

Results: Among 2.5 million recipients of ≥1 dose of mRNA COVID-19 vaccines, 3.3% (n=81,202) did not complete the second dose within 6 months. Members aged 25-44 years, 65-74 years, and ≥75 years were less likely to not complete the 2-dose primary series than those aged 18-24 years, while members aged 45-64 years were more likely to not complete the 2-dose primary series (adjusted risk ratio [aRR] 1.13, 95% CI 1.10-1.15). Male sex was associated with a higher risk of noncompletion (aRR 1.17, 95% CI 1.15-1.19). Hispanic and non-Hispanic Black race/ethnicity were associated with a lower risk of noncompletion (range aRR 0.78-0.91). Having Medicaid and prior influenza vaccination were associated with a higher risk of noncompletion. Having SARS-CoV-2 infection, experiencing an adverse event, or having an inpatient and emergency department visit during the minimum recommended dose intervals were associated with a higher risk of not completing the 2-dose primary series (aRR 1.98, 95% CI 1.85-2.12; 1.99, 95% CI 1.43-2.76; and 1.85, 95% CI 1.77-1.93, respectively). Those who received the first dose after June 30, 2021, were more likely to not complete the 2-dose primary series within 6 months of receipt of the first dose.

Conclusions: Despite limitations such as being a single-site study and the inability to consider social factors such as employment and vaccine attitudes, our study identified several risk factors for not completing a 2-dose primary series of mRNA vaccination, including being male; having Medicaid coverage; and experiencing SARS-CoV-2 infection, adverse events, or inpatient and emergency department visits during the minimum recommended dose intervals. These findings can inform future efforts in developing effective strategies to enhance vaccination coverage and improve the completion rate of necessary doses.

(JMIR Public Health Surveill 2023;9:e46318) doi:10.2196/46318

KEYWORDS

mRNA COVID-19 vaccines; 2-dose primary series; vaccines; SARS-CoV-2; coronavirus; respiratory; infectious; communicable; immunization; immunize; noncompletion; risk factors; multiple Poisson regression model; COVID-19; vaccination; vaccine; dose; dosing; regression; risk; risks; health outcome; health outcomes; retrospective; cohort; dosage; United States; community; inoculation

Introduction

The COVID-19 pandemic started in December 2019 and has tragically resulted in millions of deaths worldwide. In combating the pandemic, various mitigation measures have been implemented, including social distancing, wearing masks, promoting hygiene practices, testing, and using therapeutic interventions. Among these measures, vaccination has played a crucial role in our collective efforts to combat and control the impact of the COVID-19 pandemic. Among 4 COVID-19 vaccines (BNT162b2, mRNA-1273, Ad26.COV2.S, and NVX-CoV2373) authorized in the United States since December 2020, the 2 (messenger RNA) mRNA COVID-19 vaccines (BNT162b2, manufactured by Pfizer-BioNTech, and mRNA-1273, manufactured by Moderna) have been the most widely used [1]. These COVID-19 vaccines are effective at preventing infection, hospitalization, and death due to SARS-CoV-2 infection [2-5]. Despite these benefits, a major obstacle to COVID-19 vaccination has been vaccine hesitancy. This hesitancy has been linked to concerns regarding vaccine safety, a lack of knowledge about novel COVID-19 vaccines, demographic factors, limited health care access, attitudes toward COVID-19 vaccines, and use and trust in social media [6-11].

Two mRNA vaccines, BNT162b2 and mRNA-1273, were initially authorized as a 2-dose primary series. Optimal protection depends on completing the initial 2-dose primary series of mRNA vaccines according to recommended dose intervals, as well as receiving booster shots to counteract waning immunity and new variants [5,12-19]. As of December 8, 2022, 89.5% (n=231,012,435) of the US population aged 18 years and older received at least 1 COVID-19 vaccine dose, and 76.8% (n=198,386,322) had completed the primary series [1]. Although the vast majority of those who received an mRNA vaccine dose completed the 2-dose primary series, some did not complete the primary series and thus remained underprotected against severe COVID-19 illness and death.

Understanding the risk factors for not completing a 2-dose primary series is critical to evaluate COVID-19 vaccination programs and to promote completion of the 2-dose primary series. Although some studies examined the completion of multidose series of hepatitis A and B vaccines [20,21] and recombinant herpes zoster vaccine [22,23], studies of risk factors associated with not completing a 2-dose primary series of mRNA COVID-19 vaccines are limited. Nguyen et al [24] found associations of sociodemographic characteristics with each of 3 COVID-19 vaccination categories: receipt of at least 1 dose, receipt of the full primary series, and receipt of a booster dose after the primary series. Although sampling methods and data weighting were designed to produce nationally representative results, the sample size was moderate with 74,995 responses from the Household Pulse Survey from December 29, 2021, to January 10, 2022 [25]. Vaccination status for respondents was self-reported and was subject to social desirability bias.

The goal of this study was to examine potential risk factors for not completing a 2-dose primary series of mRNA COVID-19 vaccination in a large health care system in the United States. We considered a range of risk factors including individual-level demographics, clinical comorbidities, prior health care usage, and community-level socioeconomic characteristics, as well as health care usage, adverse events, and SARS-CoV-2 infection during the minimum recommended interval between doses.

Methods

Sample and Data

We conducted a retrospective cohort study using electronic health records of members aged ≥18 years from a large integrated health care system, Kaiser Permanente Southern California (KPSC). KPSC serves 4.7 million members of diverse sociodemographic, racial, and ethnic backgrounds at 15 medical centers [26]. To evaluate potential risk factors for not completing a 2-dose primary series of mRNA COVID-19 vaccination, we required members to receive their first dose of either mRNA COVID-19 vaccine from December 14, 2020, to December 31, 2021. We also required members to have continuous membership (allowing up to a 31-day gap) from at least 1 year before to 6 months after the first dose. The study outcome was assessed by June 30, 2022.

Ethics Approval

Ethics approval for this study was obtained from the KPSC Institutional Review Board On July 5, 2022 (approval 13270). In accordance with 45CFR 46.116, the need for informed consent was waived by the institutional review board because the research activities (secondary analyses of electronic health records data) presented no more than minimal risk to subjects. To protect the privacy and confidentiality of human subjects, all staff working on the research study were trained in procedures to protect the privacy of medical record information. All research data are stored behind a firewall in a password-protected network within the Department of Research & Evaluation at KPSC. Study participants were not compensated given the observational nature of the study.

Measures of Variables

We assessed the completion rates of 2 mRNA vaccines, BNT162b2 and mRNA-1273. The outcome, noncompletion of the 2-dose primary series, was defined as not completing 2 doses of mRNA vaccination (2 doses of BNT162b2, 2 doses of mRNA-1273 or 1 dose of each) within 6 months after receipt of the first dose. The minimum recommended intervals between dose 1 and dose 2 are 17 days and 24 days (allowing for a 4-day grace period) for BNT162b2 and mRNA-1273, respectively [27,28].

We considered individual-level risk factors including age at receipt of the first dose, sex, race and ethnicity, Medicaid status, health care usage (number of outpatient visits, virtual visits, inpatient visits, and emergency department [ED] visits) within 1 year prior to the first dose date, inpatient or ED visit within 7 days prior to the first dose date (yes or no), Charlson Comorbidity Index within 1 year prior to the first dose date, and receipt of influenza vaccine in the 2 years prior to the first dose date. We also considered community-level risk factors such as neighborhood median household income and neighborhood education level defined as 50% or more than or less than 50% of the neighborhood attaining more than high school education.

Additional individual-level risk factors that were measured during the minimum recommended interval between doses included SARS-CoV-2 infection, encounters with prespecified serious adverse events (AEs), any inpatient or ED visit, and any outpatient visit. SARS-CoV-2 infection was determined by a positive laboratory test for SARS-CoV-2 or a COVID-19 diagnosis. In identifying encounters with serious AEs, 21 types of prespecified AEs were evaluated based on their inclusion as outcomes of interest in COVID-19 vaccine safety studies [29].

In April 2021, COVID-19 vaccinations were expanded from a phased allocation to all eligible adults. To consider the impact of this expansion, we created an indicator for the time of receipt of the first dose of mRNA vaccine before and after June 30, 2021.

Model and Data Analysis Procedure

Characteristics of completers and noncompleters of the 2-dose primary series were described and compared with chi-square test and 2-tailed *t* test for categorical and continuous risk factors, respectively. To examine the association between potential risk factors and not completing the 2-dose primary series, we fit a multiple Poisson regression model with noncompletion of the 2-dose primary series within 6 months after the first dose as the dependent variable and potential risk factors as the independent variables. The specification of the Poisson regression model is as follows:

where μ_i is the expected value of y_i given a set of risk factors; y_i is the outcome for the ith subject, $y_i=1$ for not completing and $y_i=0$ for completing the 2-dose primary series; α is the intercept; X_i is a row vector of individual- and community-level risk factors, and β is a column vector of corresponding coefficients to be estimated. Because we required members to have continuous membership for at least 6 months after the first dose and the outcome was assessed within 6 months after the first dose, the follow-up period was identical across all individuals; therefore, there was no need to adjust for person time as an offset in the Poisson model. Exponentials of coefficients from the Poisson model can be interpreted as risk ratios.

We reported unadjusted risk ratios (RRs), adjusted risk ratios (aRRs), and 95% CIs. All analyses were conducted using SAS Enterprise Guide (version 8.2; SAS Institute).

Results

Among more than 4 million members of KPSC who received their first dose of mRNA COVID-19 vaccine during December 14, 2020, to December 31, 2021, 3.3 million (83.4%) were aged ≥18 years (Figure 1). Among members aged ≥18 years who received ≥1 dose of mRNA COVID-19 vaccine, 2.5 million met the membership requirement, and 3.3% (n=81,202) of them did not complete the 2-dose primary series within 6 months after receipt of the first dose (Table 1). Members who were <65 years old, male, or Hispanic were more likely not to complete the primary series. Noncompletion rates were slightly higher among members in communities with a median household income <US $60,000 and lower education level than those in communities with a higher median household income and higher education level. Noncompleters of a 2-dose primary series of mRNA COVID-19 vaccination were less likely to have had influenza vaccine in the 2 years prior to receipt of dose 1 than completers (n=48,886, 60.2% vs n=1,721,833, 72.5%). Noncompleters were more likely to receive the first dose after June 30, 2021, than completers (n=45,216, 55.7% vs n=255,354, 10.8%).

Figure 1. Study population flowchart. mRNA: messenger RNA.

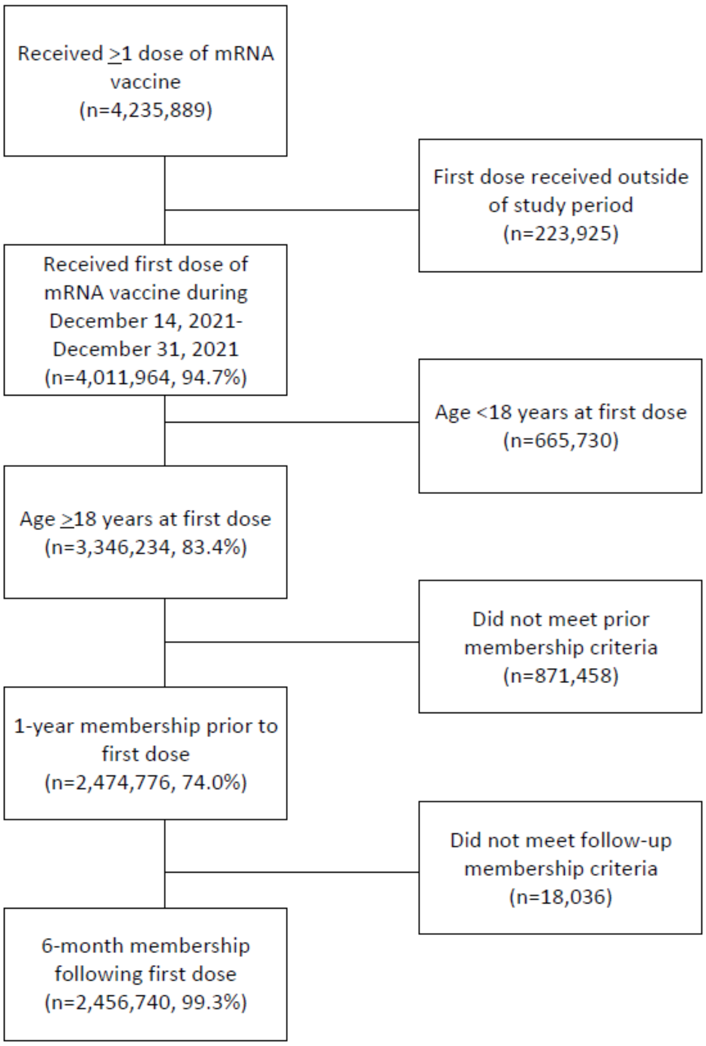


Table 1. Characteristics of completers and noncompleters of a 2-dose primary series of messenger RNA vaccination among members of Kaiser Permanente Southern California from December 14, 2020, to December 31, 2021.

Characteristic	Completion of 2-dose primary series within 6 months of first dose		Total (n=2,456,740)	*P* value
	Yes (n=2,375,538)	No (n=81,202)		
Age (years), n (%)				<.001[a]
18-24	231,728 (9.8)	11,401 (14)	243,129 (10)	
25-44	748,748 (31.5)	29,126 (35.9)	777,874 (31.7)	
45-64	818,200 (34.4)	29,658 (36.5)	847,858 (34.5)	
65-74	350,870 (14.8)	7066 (8.7)	357,936 (14.6)	
≥75	225,992 (9.5)	3951 (4.9)	229,943 (9.4)	
Age, mean (SD)	49.8 (18.3)	45.4 (17.1)	49.7 (18.2)	<.001[b]
Sex, n (%)				<.001[a]
Female	1,290,402 (54.3)	39,584 (48.8)	1,329,986 (54.1)	
Male	1,085,136 (45.7)	41,618 (51.2)	1,126,754 (45.9)	
Race and ethnicity, n (%)				<.001[a]
White	760,981 (32)	24,284 (29.9)	785,265 (32)	
Hispanic	945,650 (39.8)	33,611 (41.4)	979,261 (40)	
Asian or Pacific Islander	332,258 (14)	9981 (12.3)	342,239 (13.9)	
Black	180,116 (7.6)	6321 (7.8)	186,437 (7.6)	
Other/multiple/unknown	156,533 (6.6)	7005 (8.6)	163,538 (6.7)	
Number of outpatient visits in 1 year prior to receipt of dose 1, n (%)				<.001[a]
0	501,929 (21.1)	20,122 (24.8)	522,051 (21.2)	
1-4	1044,438 (44)	36,833 (45.4)	1,081,271 (44)	
5-9	489,910 (20.6)	15,225 (18.8)	505,135 (20.6)	
≥10	339,261 (14.3)	9022 (11.1)	348,283 (14.2)	
Number of virtual encounters in 1 year prior to receipt of dose 1, n (%)				<.001[a]
0	740,458 (31.2)	32,370 (39.9)	772,828 (31.5)	
1-4	1,061,283 (44.7)	33,976 (41.8)	1,095,259 (44.6)	
5-9	350,673 (14.8)	8977 (11.1)	359,650 (14.6)	
≥10	223,124 (9.4)	5879 (7.2)	229,003 (9.3)	
Number of inpatient visits in 1 year prior to receipt of dose 1, n (%)				<.001[a]
0	2,266,685 (95.4)	77,599 (95.6)	2,344,284 (95.4)	
1	92,717 (3.9)	2917 (3.6)	95,634 (3.9)	
≥2	16,136 (0.7)	686 (0.8)	16,822 (0.7)	
Number of ED[c] visits in 1 year prior to receipt of dose 1, n (%)				<.001[a]
0	2,059,842 (86.7)	69,647 (85.8)	2,129,489 (86.7)	
1-2	284,371 (12)	10,163 (12.5)	294,534 (12)	
≥3	31,325 (1.3)	1392 (1.7)	32,717 (1.3)	
Charlson Comorbidity Index in 1 year prior to receipt of dose 1, n (%)				<.001[a]
0	1,669,133 (70.3)	63,373 (78)	1,732,506 (70.5)	
1	342,218 (14.4)	9757 (12)	351,975 (14.3)	
2	155,730 (6.6)	3632 (4.5)	159,362 (6.5)	

Characteristic	Completion of 2-dose primary series within 6 months of first dose		Total (n=2,456,740)	*P* value
	Yes (n=2,375,538)	No (n=81,202)		
≥3	208,457 (8.8)	4440 (5.5)	212,897 (8.7)	
Had inpatient or ED visit in the 7 days prior to dose 1, n (%)				<.001[a]
Yes	9073 (0.4)	487 (0.6)	9560 (0.4)	
No	2,366,465 (99.6)	80,715 (99.4)	2,447,180 (99.6)	
Had Medicaid in 1 year prior to receipt of dose 1, n (%)				<.001[a]
Yes	165,639 (7)	7453 (9.2)	173,092 (7)	
No	2,209,899 (93)	73,749 (90.8)	2,283,648 (93)	
Had influenza vaccine in the 2 years prior to dose 1, n (%)				<.001[a]
Yes	1,721,833 (72.5)	48,886 (60.2)	1,770,719 (72.1)	
No	653,705 (27.5)	32,316 (39.8)	686,021 (27.9)	
Neighborhood median household income, n (%), US ($)				<.001[a]
<$40,000	84,852 (3.6)	3442 (4.2)	88,294 (3.6)	
$40,000-$59,999	422,975 (17.8)	15,956 (19.7)	438,931 (17.9)	
$60,000-$79,999	558,709 (23.5)	19,533 (24.1)	578,242 (23.5)	
$80,000-$99,999	533,012 (22.4)	17,347 (21.4)	550,359 (22.4)	
≥$100,000	772,181 (32.5)	24,678 (30.4)	796,859 (32.4)	
Missing	3809 (0.2)	246 (0.3)	4055 (0.2)	
Neighborhood-level education, n (%)				<.001[a]
Less than 50% of the neighborhood attaining greater than a high school education	717,881 (30.2)	26,299 (32.4)	744,180 (30.3)	
50% or more of the neighborhood attaining greater than a high school education	1,653,747 (69.6)	54,655 (67.3)	1,708,402 (69.5)	
Missing	3910 (0.2)	248 (0.3)	4158 (0.2)	
Had SARS-CoV-2 infection during minimum recommended dose interval, n (%)				<.001[a]
Yes	5783 (0.2)	858 (1.1)	6641 (0.3)	
No	2,369,755 (99.8)	80,344 (98.9)	2,450,099 (99.7)	
Had adverse event during minimum recommended dose interval, n (%)				<.001[a]
Yes	410 (0.02)	36 (0.04)	446 (0.02)	
No	2,375,128 (99.98)	81,166 (99.96)	2,456,294 (99.98)	
Had inpatient or ED visit during minimum recommended dose interval, n (%)				<.001[a]
Yes	35,052 (1.5)	2363 (2.9)	37,415 (1.5)	
No	2,340,486 (98.5)	78,839 (97.1)	2,419,325 (98.5)	
Had outpatient visit during minimum recommended dose interval, n (%)				<.001[a]
Yes	1,339,444 (56.4)	34,644 (42.7)	1,374,088 (55.9)	
No	1,036,094 (43.6)	46,558 (57.3)	1,082,652 (44.1)	
Receipt of first dose after June 30, 2021, n (%)				<.001[a]
Yes	255,354 (10.8)	45,216 (55.7)	300,570 (12.2)	

Characteristic	Completion of 2-dose primary series within 6 months of first dose		Total (n=2,456,740)	*P* value
	Yes (n=2,375,538)	No (n=81,202)		
No	2,120,184 (89.2)	35,986 (44.3)	2,156,170 (87.8)	

[a] Chi-square *P* value.

[b] *t* test *P* value.

[c] ED: emergency department.

Among the 2.5 million members in the study cohort, 6641 (0.3%) members had SARS-CoV-2 infection during the minimum recommended dose interval (Table 1), of whom 858 (12.9%) did not complete the primary series in the 6 months after receipt of the first dose, higher than the overall noncompletion rate of 3.3% (n=81,202). Among 446 (0.02%) members who experienced an AE during the minimum recommended dose interval, 36 (8.1%) did not complete the primary series in the 6 months after receipt of the first dose. Among 37,415 (1.5%) members who had an ED or inpatient visit during the minimum recommended dose interval, 2363 (6.3%) did not complete the primary series in the 6 months after receipt of the first dose.

Unadjusted RRs for noncompletion of the 2-dose primary series are displayed in Figure S1 in Multimedia Appendix 1. The multiple Poisson regression showed that after adjusting for other potential risk factors, members aged 25-44 years, 65-74 years, and ≥75 years were more likely to complete the 2-dose primary series than those aged 18-24 years (aRR for noncompletion 0.94, 95% CI 0.92-0.96; aRR 0.90, 95% CI 0.87-0.93; and 0.92, 95% CI 0.88-0.96, respectively), while members aged 45-64 years were more likely to not complete the 2-dose primary series (aRR 1.13, 95% CI 1.10-1.15) (Figure S2 in Multimedia Appendix 1). Male sex was associated with a higher risk of noncompletion (aRR 1.17, 95% CI 1.15-1.19).

Hispanic and non-Hispanic Black race and ethnicities were associated with lower risk of noncompletion (aRR 0.91, 95% CI 0.90-0.93; aRR 0.78, 95% CI 0.76-0.80, respectively), while Asian and Pacific Islander race was associated with a slightly higher risk of noncompletion (aRR 1.07, 95% CI 1.04-1.09). Having Medicaid in 1 year prior to receipt of dose 1, having influenza vaccine in the 2 years prior to receipt of dose 1, and certain health care use patterns (1-4 and 5-9 outpatient visits, ≥2 inpatient visits, and ≥1 ED visits in the year prior to receipt of dose 1) were associated with a higher risk of noncompletion.

Members with SARS-CoV-2 infection during the minimum recommended dose interval were more likely to be noncompleters (aRR 1.98, 95% CI 1.85-2.12) (Figure S2 in Multimedia Appendix 1). Experiencing an AE during the minimum recommended dose interval was also associated with not completing the primary series (aRR 1.99, 95% CI 1.43-2.76). In addition, members with an inpatient or ED visit during the minimum recommended dose interval were more likely to be noncompleters (aRR 1.85, 95% CI 1.77-1.93), but having an outpatient visit during the minimum recommended dose interval significantly decreased the risk of not completing the primary series (aRR 0.53, 95% CI 0.52-0.53). Receipt of the first dose after June 30, 2021, was associated with a higher risk of not completing the primary series (aRR 9.78, 95% CI 9.63-9.93).

Discussion

Principal Findings

Acting in synergy with community pharmacy services and local government mass vaccination efforts, the KPSC health care system was very successful in promoting COVID-19 vaccination and completion of the 2-dose primary series. During the pandemic, KPSC implemented a community-oriented and geographically targeted vaccine strategy aimed at identifying specific zip codes that required additional resources. This proactive approach proved helpful in enhancing COVID-19 vaccination coverage and effectively addressing the disparities in vaccine uptake within underserved communities [30]. Several other aspects should also be considered in designing effective public policy to address future pandemic crises such as Peltzman effects in vaccinations, the emergence of new variants of concern, investment in health care, and the impact of environmental pollution and climate factors [31].

This study showed some impact of sociodemographic characteristics on the completion of a 2-dose primary series of mRNA COVID-19 vaccination in KPSC. For example, after adjusting for other risk factors, the characteristics of being aged ≥65 years, female, and Black or Hispanic, as well as living in a neighborhood with a lower education level, were associated with a lower risk of not completing the 2-dose primary series within 6 months of receipt of the first dose. Members who had ≥2 ED visits in the year prior to receipt of dose 1 were less likely to complete the primary series (aRR for noncompletion 1.23, 95% CI 1.16-1.30). Having SARS-CoV-2 infection, experiencing an AE, or having an inpatient or ED visit during the minimum recommended dose intervals were associated with a higher risk of not completing the 2-dose primary series (aRR 1.98, 95% CI 1.85-2.12; 1.99, 95% CI 1.43-2.76; and 1.85, 95% CI 1.77-1.93, respectively). Those who received the first dose before June 30, 2021, were more likely to complete the 2-dose primary series within 6 months of receipt of the first dose.

In this study, male sex was associated with a higher risk of not completing the 2-dose primary series, while a study by Nguyen et al [24] did not find an association between sex and noncompletion. Although the association of age with noncompletion of the 2-dose primary series in unadjusted analyses (Figure S1 in Multimedia Appendix 1) was largely consistent with those of Nguyen et al [24], the strength of association attenuated after adjusting for other risk factors, from unadjusted RRs of 0.37-0.80 to aRRs of 0.90-1.13. Our finding of lower risk of not completing the 2-dose primary series among

Black and Hispanic race and ethnicity was consistent with the study by Nguyen et al [24]; for Asian race and ethnicity, our study found a weak association with not completing the 2-dose primary series (aRR 1.07, 95% CI 1.04-1.09), while Nguyen et al [24] found that Asian race/ethnicity was strongly associated with completion of the 2-dose primary series. Those with lower household income had unadjusted RR ≥1 of not completing the 2-dose primary series (Figure S1 in Multimedia Appendix 1), which is similar to Nguyen et al [24]; however, after adjusting for other risk factors, the association became null. The 2 studies differed in several aspects including the study population (insured vs a general population), study design (retrospective cohort study using electronic health records vs survey study), and risk factors considered. These differences may have contributed to the different findings. In comparison to mRNA COVID-19 vaccination, completion rates for hepatitis A and B vaccinations were observed to be lower, ranging from 40% to 65% across a broad age range [20]. Notably, individuals with Medicaid coverage were found to be associated with a lower completion rate for both the 2-dose primary series of mRNA COVID-19 vaccination in our study and the multidose vaccinations of hepatitis A (2 doses) and B (3 doses) in the study by Nelson et al [20].

With more than 12 months of first-dose data from a large health care organization, this study examined a broad range of potential risk factors including individual- and community-level sociodemographic characteristics, Medicaid status, comorbidities, health care usage, and medical encounters during minimum recommended dose intervals. In a study examining potential factors associated with completing 2 doses of recombinant zoster vaccine among individuals aged ≥50 years, local and systemic reactions after receipt of the first dose were examined in addition to sociodemographic characteristics and health care usage; however, medical encounters with serious AEs were not examined [23]. A unique strength of our study is that we examined the association between medical encounters with serious AEs during minimum recommended dose intervals and noncompletion of the mRNA COVID-19 vaccine 2-dose primary series. Only 87.1% (n=5783) of those who had SARS-CoV-2 infection during the minimum recommended dose intervals completed the primary series, which is significantly lower than the overall completion rate of 96.7% (n=2,375,538). The noncompleters might have felt that it was not necessary to get the second dose because of natural immunity gained from infection. Only 8.1% (n=36) of those who experienced an AE during the minimum recommended dose intervals did not complete the primary series.

There are some limitations in this study. First, the study population is an insured population from a large health care system. The completion rate of the 2-dose primary series among KPSC members aged 18 years and older is 96.7% (n=2,375,538). In contrast, the national completion rate for the primary series, which includes individuals from various health care providers and uninsured individuals, is notably lower at 76.8% (n=198,386,322). The findings in this study may not be generalizable to other health care systems and uninsured populations. Second, data on other risk factors such as members' employment were not available. Mandatory completion of a 2-dose primary series for employment would decrease the likelihood of noncompletion. Third, we did not consider the requirement of 3 doses for the primary series among immunocompromised individuals [27]. Finally, while we included SARS-CoV-2 infection during the minimum recommended dose intervals as a risk factor for completing the second dose, the impact of SARS-CoV-2 infection prior to receipt of the first dose was not considered.

Conclusions

To adequately prepare for the pandemic, it is crucial to attain a sufficient level of vaccination coverage, contributing to herd immunity. This will effectively halt the spread of potential virus, thereby protecting people against severe illness and death [32]. Despite limitations such as a study of a single site of an insured population and inability to consider social factors such as employment and attitude to vaccination, our study identified several risk factors for not completing a 2-dose primary series of mRNA vaccination. These factors include being male; having Medicaid coverage; and experiencing SARS-CoV-2 infection, AEs, or inpatient and ED visits during the minimum recommended dose intervals. These findings can inform future efforts in developing effective strategies to enhance vaccination coverage and improve the completion rate of necessary doses.

Acknowledgments

Funding for this study was provided by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01 AI168209. The findings and conclusions in this report are those of the authors and do not necessarily represent the official views of the National Institutes of Health.

Data Availability

Individual-level data reported in this study involving human research participants are not publicly shared due to potentially identifying or sensitive patient information. Upon request to the corresponding author, and subject to review and approval of an analysis proposal, KPSC may provide the deidentified aggregate-level data that support the findings of this study within 6 months. Anonymized data (deidentified data including participant data as applicable) that support the findings of this study may be made available from the investigative team in the following conditions: (1) agreement to collaborate with the study team on all publications, (2) provision of external funding for administrative and investigator time necessary for this collaboration, (3) demonstration that the external investigative team is qualified and has documented evidence of training for human subjects protections, and (4) agreement to abide by the terms outlined in data use agreements between institutions.

Conflicts of Interest

LSS reports research support from Moderna, GlaxoSmithKline, and Dynavax for unrelated studies. LQ received research support from Moderna for a COVID-19 vaccine effectiveness study and GlaxoSmithKline and Dynavax for unrelated studies. VH reports research funding from Pfizer sent directly to their institution and unrelated to this study. KJB reports research funding from Dynavax, GlaxoSmithKline, Moderna, and Pfizer unrelated to this study. All other authors declare no conflicts.

Multimedia Appendix 1
Risk ratios (95% CI) of not completing the 2-dose primary series of mRNA COVID-19 vaccination within 6 months of the first dose given among members of Kaiser Permanente Southern California during December 14, 2020 to December 31, 2021.
[PDF File (Adobe PDF File), 738 KB - publichealth_v9i1e46318_app1.pdf]

References

1. COVID-19 vaccinations in the United States. Centers for Disease Control and Prevention. 2022. URL: https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-rate-total [accessed 2022-12-11]
2. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
3. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021;384(5):403-416 [FREE Full text] [doi: 10.1056/NEJMoa2035389] [Medline: 33378609]
4. Magazzino C, Mele M, Coccia M. A machine learning algorithm to analyse the effects of vaccination on COVID-19 mortality. Epidemiol Infect 2022;150:e168 [FREE Full text] [doi: 10.1017/S0950268822001418] [Medline: 36093862]
5. Rosero-Bixby L. The effectiveness of pfizer-biontech and Oxford-astrazeneca vaccines to prevent severe COVID-19 in Costa Rica: nationwide, ecological study of hospitalization prevalence. JMIR Public Health Surveill 2022;8(5):e35054 [FREE Full text] [doi: 10.2196/35054] [Medline: 35483079]
6. Griffith J, Marani H, Monkman H. COVID-19 vaccine hesitancy in Canada: content analysis of Tweets using the theoretical domains framework. J Med Internet Res 2021;23(4):e26874 [FREE Full text] [doi: 10.2196/26874] [Medline: 33769946]
7. Yuan Y, Melde C, Zhang N, Pagidipati P. Race, ethnicity, psychological factors, and COVID-19 vaccine hesitancy during the COVID-19 pandemic. Psychol Health Med 2023;28(2):427-438. [doi: 10.1080/13548506.2022.2084123] [Medline: 35638115]
8. Matas JL, Landry LG, Lee L, Hansel S, Coudray MS, Mata-McMurry LV, et al. Demographic determinants and geographical variability of COVID-19 vaccine hesitancy in underserved communities: cross-sectional study. JMIR Public Health Surveill 2023;9:e34163 [FREE Full text] [doi: 10.2196/34163] [Medline: 36811869]
9. Moon I, Han J, Kim K. Determinants of COVID-19 vaccine hesitancy: 2020 California health interview survey. Prev Med Rep 2023;33:102200 [FREE Full text] [doi: 10.1016/j.pmedr.2023.102200] [Medline: 37034299]
10. Jones LF, Bonfield S, Farrell J, Weston D. Understanding the public's attitudes toward COVID-19 vaccines in Nottinghamshire, United Kingdom: qualitative social media analysis. J Med Internet Res 2023;25:e38404 [FREE Full text] [doi: 10.2196/38404] [Medline: 36812390]
11. Carrieri V, Guthmuller S, Wübker A. Trust and COVID-19 vaccine hesitancy. Sci Rep 2023;13(1):9245 [FREE Full text] [doi: 10.1038/s41598-023-35974-z] [Medline: 37286569]
12. Livingston EH. Necessity of 2 doses of the pfizer and moderna COVID-19 vaccines. JAMA 2021;325(9):898 [FREE Full text] [doi: 10.1001/jama.2021.1375] [Medline: 33533893]
13. Martínez-Baz I, Trobajo-Sanmartín C, Miqueleiz A, Guevara M, Fernández-Huerta M, Burgui C, Working Group for the Study of COVID-19 in Navarre, Investigators, other members of the Working Group for the Study of COVID-19 in Navarre. Product-specific COVID-19 vaccine effectiveness against secondary infection in close contacts, Navarre, Spain, April to August 2021. Euro Surveill 2021;26(39):2100894 [FREE Full text] [doi: 10.2807/1560-7917.ES.2021.26.39.2100894] [Medline: 34596016]
14. Goldberg Y, Mandel M, Bar-On YM, Bodenheimer O, Freedman L, Haas EJ, et al. Waning immunity after the BNT162b2 vaccine in Israel. N Engl J Med 2021;385(24):e85 [FREE Full text] [doi: 10.1056/NEJMoa2114228] [Medline: 34706170]
15. Tartof SY, Slezak JM, Fischer H, Hong V, Ackerson BK, Ranasinghe ON, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet 2021;398(10309):1407-1416 [FREE Full text] [doi: 10.1016/S0140-6736(21)02183-8] [Medline: 34619098]
16. Bruxvoort KJ, Sy LS, Qian L, Ackerson BK, Luo Y, Lee GS, et al. Real-world effectiveness of the mRNA-1273 vaccine against COVID-19: interim results from a prospective observational cohort study. Lancet Reg Health Am 2022;6:100134 [FREE Full text] [doi: 10.1016/j.lana.2021.100134] [Medline: 34849505]
17. Tseng HF, Ackerson BK, Luo Y, Sy LS, Talarico CA, Tian Y, et al. Effectiveness of mRNA-1273 against SARS-CoV-2 Omicron and Delta variants. Nat Med 2022;28(5):1063-1071 [FREE Full text] [doi: 10.1038/s41591-022-01753-y] [Medline: 35189624]

18. Korves C, Izurieta HS, Smith J, Zwain GM, Powell EI, Balajee A, et al. Relative effectiveness of booster vs. 2-dose mRNA Covid-19 vaccination in the Veterans Health Administration: self-controlled risk interval analysis. Vaccine 2022;40(33):4742-4747 [FREE Full text] [doi: 10.1016/j.vaccine.2022.06.047] [Medline: 35773122]
19. Rosenblum HG, Wallace M, Godfrey M, Roper LE, Hall E, Fleming-Dutra KE, et al. Interim recommendations from the advisory committee on immunization practices for the use of bivalent booster doses of COVID-19 vaccines - United States, october 2022. MMWR Morb Mortal Wkly Rep 2022;71(45):1436-1441 [FREE Full text] [doi: 10.15585/mmwr.mm7145a2] [Medline: 36355612]
20. Nelson JC, Bittner RCL, Bounds L, Zhao S, Baggs J, Donahue JG, et al. Compliance with multiple-dose vaccine schedules among older children, adolescents, and adults: results from a vaccine safety datalink study. Am J Public Health 2009;99(Suppl 2):S389-S397 [FREE Full text] [doi: 10.2105/AJPH.2008.151332] [Medline: 19797753]
21. Bruxvoort K, Slezak J, Huang R, Ackerson B, Sy LS, Qian L, et al. Association of number of doses with hepatitis B vaccine series completion in US adults. JAMA Netw Open 2020;3(11):e2027577 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.27577] [Medline: 33252692]
22. Patterson BJ, Chen CC, McGuiness CB, Glasser LI, Sun K, Buck PO. Early examination of real-world uptake and second-dose completion of recombinant zoster vaccine in the United States from October 2017 to September 2019. Hum Vaccin Immunother 2021;17(8):2482-2487 [FREE Full text] [doi: 10.1080/21645515.2021.1879579] [Medline: 33849373]
23. Ackerson B, Qian L, Sy LS, Bruxvoort K, Wu J, Luo Y, et al. Completion of the two-dose recombinant zoster vaccine series in adults 50 years and older. Vaccine 2021;39(6):926-932 [FREE Full text] [doi: 10.1016/j.vaccine.2020.12.076] [Medline: 33441234]
24. Nguyen KH, Chen Y, Huang J, Allen JD, Beninger P, Corlin L. Who has not been vaccinated, fully vaccinated, or boosted for COVID-19? Am J Infect Control 2022;50(10):1185-1189 [FREE Full text] [doi: 10.1016/j.ajic.2022.05.024] [Medline: 35688229]
25. Source of the Data and Accuracy of the Estimates for the Household Pulse Survey – Phase 3.3. US Census Bureau. 2022. URL: https://www2.census.gov/programs-surveys/demo/technical-documentation/hhp/Phase3-3_Source_and_Accuracy_Week41.pdf [accessed 2022-12-20]
26. Koebnick C, Langer-Gould AM, Gould MK, Chao CR, Iyer RL, Smith N, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. Perm J 2012;16(3):37-41 [FREE Full text] [doi: 10.7812/TPP/12-031] [Medline: 23012597]
27. Use of COVID-19 vaccines in the United States. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/vaccines/covid-19/clinical-considerations/covid-19-vaccines-us.html [accessed 2022-12-29]
28. Interim clinical considerations for use of COVID-19 vaccines in the United States. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/vaccines/covid-19/clinical-considerations/interim-considerations-us.html#timing-spacing-interchangeability [accessed 2023-01-12]
29. Xu S, Hong V, Sy LS, Glenn SC, Ryan DS, Morrissette KL, et al. Changes in incidence rates of outcomes of interest in vaccine safety studies during the COVID-19 pandemic. Vaccine 2022;40(23):3150-3158 [FREE Full text] [doi: 10.1016/j.vaccine.2022.04.037] [Medline: 35465977]
30. Swope M, Alem AC, Russo SC, Gin NE, Chevez SG, Haque R. Developing a community-oriented and place-based strategy to improve COVID-19 vaccine accessibility. Perm J 2023;27(1):103-112 [FREE Full text] [doi: 10.7812/TPP/22.097] [Medline: 36911890]
31. Coccia M. COVID-19 vaccination is not a sufficient public policy to face crisis management of next pandemic threats. Public Organization Review. 2022. URL: https://link.springer.com/article/10.1007/s11115-022-00661-6 [accessed 2023-09-12]
32. Coccia M. Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: a global analysis. Environ Res 2022;204(Pt C):112314 [FREE Full text] [doi: 10.1016/j.envres.2021.112314] [Medline: 34736923]

Abbreviations

AE: adverse event
aRR: adjusted risk ratio
ED: emergency department
KPSC: Kaiser Permanente Southern California
mRNA: messenger RNA
RR: risk ratio

Edited by A Mavragani, T Sanchez; submitted 06.02.23; peer-reviewed by M Coccia, J Hollister; comments to author 04.07.23; revised version received 22.07.23; accepted 29.08.23; published 04.10.23.

Please cite as:
Xu S, Hong V, Sy LS, Bruxvoort KJ, Lewin B, Han B, Holmquist KJ, Qian L
Risk Factors for Not Completing a 2-Dose Primary Series of Messenger RNA COVID-19 Vaccination in a Large Health Care System in Southern California: Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e46318
URL: https://publichealth.jmir.org/2023/1/e46318
doi:10.2196/46318
PMID:37792452

Original Paper

Prevalence of COVID-19 Mitigation Behaviors in US Adults (August-December 2020): Nationwide Household Probability Survey

Travis Sanchez[1]; Eric Hall[1], PhD; Aaron J Siegler[1], PhD; Radhika Prakash-Asrani[1], MPH; Heather Bradley[2], PhD; Mansour Fahimi[3], PhD; Benjamin Lopman[1]; Nicole Luisi[1]; Kristin N Nelson[1]; Charles Sailey[4]; Kayoko Shioda[1]; Mariah Valentine-Graves[1]; Patrick S Sullivan[1], DVM, PhD

[1]Rollins School of Public Health, Emory University, Atlanta, GA, United States
[2]School of Public Health, Georgia State University, Atlanta, GA, United States
[3]Marketing Systems Group, Horsham, PA, United States
[4]Molecular Testing Labs, Vancouver, WA, United States

Corresponding Author:
Travis Sanchez
Rollins School of Public Health
Emory University
1518 Clifton Road NE
Atlanta, GA, 30322
United States
Phone: 1 404 727 8403
Email: Travis.Sanchez@emory.edu

Abstract

Background: COVID-19 mitigation behaviors, such as wearing masks, maintaining social distancing, and practicing hand hygiene, have been and will remain vital to slowing the pandemic.

Objective: This study aims to describe the period prevalence of consistent mask-wearing, social distancing, and hand hygiene practices during the peak of COVID-19 incidence (August-December 2020) and just before COVID-19 vaccine availability, overall and in demographic subgroups.

Methods: We used baseline survey data from a nationwide household probability sample to generate weighted estimates of mitigation behaviors: wearing masks, maintaining social distancing, and practicing hand hygiene. Weighted logistic regression explored differences in mitigation behaviors by demographics. Latent class analysis (LCA) identified patterns in mitigation behaviors.

Results: Among 4654 participants, most (n=2727, 58.6%) were female, were non-Hispanic White (n=3063, 65.8%), were aged 55 years or older (n=2099, 45.1%), lived in the South (n=2275, 48.9%), lived in metropolitan areas (n=4186, 89.9%), had at least a bachelor's degree (n=2547, 54.7%), had an income of US $50,000-$99,000 (n=1445, 31%), and were privately insured (n=2734, 58.7%). The period prevalence of consistent mask wearing was 71.1% (sample-weighted 95% CI 68.8-73.3); consistent social distancing, 42.9% (95% CI 40.5-45.3); frequent handwashing, 55.0% (95% CI 52.3-57.7); and frequent hand sanitizing, 21.5% (95% CI 19.4-23.8). Mitigation behaviors were more prevalent among women, older persons, Black or Hispanic persons, those who were not college graduates, and service-oriented workers. LCA identified an optimal-mitigation class that consistently practiced all behaviors (n=2656, 67% of US adults), a low-mitigation class that inconsistently practiced all behaviors (n=771, 20.6%), and a class that had optimal masking and social distancing but a high frequency of hand hygiene (n=463, 12.4%).

Conclusions: Despite a high prevalence of COVID-19 mitigation behaviors, there were likely millions who did not consistently practice these behaviors during the time of the highest COVID-19 incidence. In future infectious disease outbreak responses, public health authorities should also consider addressing disparities in mitigation practices through more targeted prevention messaging.

(JMIR Public Health Surveill 2023;9:e37102) doi:10.2196/37102

KEYWORDS

COVID-19; mask; social distancing; handwashing; hand sanitizer; public health; pandemic; mitigation behavior; risk factor; disease prevention; health policy; latent class analysis; hygiene

Introduction

The COVID-19 pandemic was first documented in China in December 2019, with cases identified in the United States shortly afterward [1,2]. As of December 2020, there were more than 20 million COVID-19 cases and more than 360,000 COVID-19–related deaths in the United States [3]. Because the vaccine did not become available to select groups until late December 2020, the primary means of preventing COVID-19 infection during the first year of the pandemic included the consistent use of mitigation behaviors of mask wearing, social distancing, and hand hygiene [4].

Multiple published studies have explored the prevalence of mitigation behaviors among US adults during the first year of the pandemic, but many of these are convenience samples or from limited geographic areas. Those that explored mask usage in public settings have reported a highly variable prevalence (40%-90%) of usage by the type of public setting for mask wearing and the timing of local mask mandates [5-10]. Social distancing, which can encompass limiting exposure to persons outside of one’s household and keeping at least 6 feet apart from others when outside the home, has been reported by a few previous studies, and estimates have also been prone to variation (70%-87%) due to rapidly changing local mandates for limiting travel and closure of public venues [11,12]. Recommended hand hygiene practices, such as frequent handwashing or the use of alcohol-based hand sanitizers, have been infrequently reported by studies of the general population but have more consistently shown high variation (74%-93%) [13-15].

The COVID Impact Survey is one of the most comprehensive and well-published studies of mitigation behaviors using cross-sectional samples and emailed surveys at multiple time points [12]. The study found that mask usage grew significantly from 78% in April 2020 to 89% in June 2020. The authors also found reductions in attending public venues, and >80% of participants reported keeping at least 6 feet apart from others, although there was a decreasing trend in these behaviors between April and July 2020, likely associated with the easing of local travel restrictions and business closures. Finally, the authors found that the proportion of US adults who frequently washed or sanitized hands was the highest (90%) in April 2020 but declined through July 2020 [12].

Just as there has been substantial heterogeneity in COVID-19 cases, morbidity, and mortality among US adults from various demographic backgrounds [16-19], there have also been some reported differences in the uptake of mitigation practices during the COVID-19 pandemic. Mask wearing may be more prevalent among women, non-White persons, and those who do not live in rural areas [5,9]. Handwashing may also be more prevalent among women, older adults, and those who identify as Black or Hispanic/Latinx [15]. Evidence from ecological analyses of mobility data found that social distancing was higher in counties with lower levels of poverty, a larger proportion of Black residents, and a higher population density [20].

Despite these high-quality studies of COVID-19 mitigation practices, there remain substantial gaps in our knowledge. Currently published estimates of the prevalence of mitigation practices were from before the peak of COVID-19 incidence, and few were from probability-based samples of the general US population. Having minimally biased information at multiple time points during the pandemic is critical to understanding the ongoing needs for public health communications regarding these mitigation practices. Having additional reliable estimates from later time points during the pandemic are also useful parameters for COVID-19 modeling activities.

We collected and analyzed baseline assessment data of a prospective cohort of a representative household-based sample of US adults. The main objective of this study was to describe the period prevalence of consistent mask-wearing, social distancing, and hand hygiene practices during the peak of COVID-19 incidence (August-December 2020) and just before COVID-19 vaccine availability, overall and in demographic subgroups. We also explored whether people engage in mitigation practices as a set of activities or as individual unconnected behaviors.

Methods

Participants and Procedures

COVIDVu is a prospective observational cohort study with a nationwide household probability sample of US adults using sampling methods that have been previously described [21]. A total of 39,500 US households were sampled using addresses derived from the US Postal Service Computerized Delivery Sequence File, including oversampling of households with census tracts comprising >50% Black residents and households with surnames likely to represent Hispanic ethnicity. The sampled households also include oversamples in California (16.5%) and Georgia (30.4%) to allow state-level estimation. All sampled households were shipped a study kit each. One adult resident enumerated the number of household members and each person’s age. One enumerated household member aged ≥18 years was randomly selected and offered participation in the study. Consenting participants were asked to complete an online survey, self-collect an anterior nares swab and a dried blood spot card, and return the specimens to a central laboratory via a prepaid mailer. Participants who returned the specimens were compensated at least US $40. Using procedures previously described for this study, sample and design weights were applied to estimate unbiased measures for noninstitutionalized, housed US adults in 2020 [21].

Ethical Considerations

Informed consent was obtained from each participant in the study. The study was conducted in compliance with federal regulations governing protection of human subjects and was reviewed and approved by Emory University's Institutional Review Board (protocol 00000695).

Measures

The following primary dependent measures were used in this study: mask wearing, social distancing, handwashing, and hand sanitizer use. Mask wearing was defined by the question "When you go out, do you wear a face mask?" with response options of always, often, sometimes, rarely, and never. Social distancing was defined by the question "How often are you trying to keep at least 6 feet between you and other people you don't live with to avoid spreading illness?" with response options of always, often, sometimes, rarely, and never. Those who reported always wearing masks or maintaining social distancing were defined as consistently practicing these behaviors. Handwashing was defined by the question "In the past 24 hours, about how many times did you wash your hands with soap and water?" Hand sanitizer use was defined by the question "In the past 24 hours, about how many times did you use an alcohol-based hand sanitizer spray, gel, or wipes?" Based on a prior study examining handwashing effectiveness at preventing seasonal coronavirus [22], we categorized handwashing and hand sanitizer use as 0-5 times per day (the prior study referent group), 6-10 times per day (the prior study effective intervention group), and ≥11 times per day (the prior study intervention group that was not effective). For dichotomous analyses, those in the categories of 6-10 and ≥11 times per day were defined as frequently practicing hand hygiene.

Independent analysis measures included standard demographic characteristics of gender, race/ethnicity, age group, US Census region, ZIP code–based urbanicity, highest education, annual household income, and current health insurance. For those working, the job type was collected using 2018 US Bureau of Labor Statistics major job categories [23]. Some types of jobs, such as food service, education, health care, retail, and transportation services, may have had additional recommendations about mask wearing and hand hygiene [24]. These job types were differentiated from others in the analyses. For those who leave home for work, we assessed whether their jobs were completely indoors or outdoors/mix/other.

Analyses

We developed sample weights to represent noninstitutionalized, housed adults (aged ≥18 years, US population). In brief, hierarchical hot deck imputation was performed to ensure no participants were missing data for key variables needed for weighting, such as gender, education, race, ethnicity, and marital status; each had <3% missingness [25]. Design weights, adjusted with classification and regression tree (CART) analysis for a differential nonresponse, were developed to facilitate population inference. A raking procedure aligned weighted distributions to the observed distributions from the Census along the lines including age, race-ethnicity, education, and income [26]. To address outlier weights, those at the 99th percentile of each distribution side were trimmed.

Using the sampling weights, we estimated the weighted prevalence and 95% modified Wilson score confidence limits for (1) consistent mask wearing, (2) social distancing, (3) handwashing, and (4) hand sanitizer use. Prevalence estimates were descriptively summarized by sociodemographic factors (race, sex, age, region, urbanicity, highest level of education, annual income, health insurance, job type), personal behaviors (leaving home for work), knowledge of mitigation behaviors, and month of sampling. To identify significant differences in the prevalence of mitigation behaviors by sociodemographic factors, prevalence ratios (PRs) and corresponding 95% CIs were estimated using weighted logistic regression procedures. All prevalence analyses were performed using SAS v9.4 (SAS Institute) and SUDAAN (RTI International).

People may follow all public health recommendations for mitigation practices similarly, or there may be individual variations in mitigation practices. This information may be useful for understanding how different groups respond to multicomponent prevention messages. We therefore conducted latent class analysis (LCA) with polytomous outcomes variables to classify participants based on their responses to the 4 primary dependent measures (mask wearing, social distancing, handwashing, and hand sanitizer use). Each measure was included as a single item. Considering we did not know the number of classes represented by these data, we fit several models with a different number (1-6) of classes. All models were fit using the *polCA* package in R v4.1.0 (R Foundation for Statistical Computing), which uses maximum likelihood parameter estimation with robust SEs. Each model was estimated with 30 different sets of starting values and allowed a maximum of 3000 iterations for convergence. To select the final model, we compared fit statistics (ie, Bayesian information criteria [BIC], Akaike information criteria [AIC], adjusted BIC) and accuracy statistics (eg, entropy) [27]. Each participant was classified into a latent class (ie, mitigation) group by the largest posterior probability for belonging to each class indicated by the final model. We estimated the weighted prevalence of each latent class group in the entire sample and by sociodemographic characteristics. Finally, as a minimal internal validity check, we examined whether class membership was associated with answering the following statements with "true": "Consistently wearing a face mask will provide me with 95% or better protection from getting infected with the new coronavirus" and "It is not necessary for children and young adults to take measures to prevent infection by the COVID-19 virus." PRs and corresponding 95% CIs were estimated using weighted logistic regression procedures in SUDAAN to identify any differences by sociodemographic characteristics.

Results

Participants

A total of 4654 participants completed baseline enrollment procedures and were included in this study (Table 1). These participants represented 242,875,582 US adults in 2020. Most were female, were non-Hispanic White, aged 55 years or older, lived in the South, lived in metropolitan areas, had at least a bachelor's degree, had an income of US $50,000-$99,000, and were privately insured. The highest monthly enrollment occurred in November 2020. Among the available response options in this study for job types, most respondents worked in health care and social services; however, most had other job types not available for selection in the study. Most of those whose jobs required leaving home worked completely indoors.

Table 1. Prevalence of consistent mask wearing among a household probability sample of 4654 US adults (August-December 2020).

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
Overall	3351/4654 (72.0)	172,749,029/242,875,582 (71.1); 68.8-73.3	N/A[b]
Sex			
Male	1315/1927 (68.2)	76,820,491/115,613,214 (66.4); 62.7-70.0	Reference (N/A)
Female	2036/2727 (74.7)	95,928,538/127,262,368 (75.4); 72.6-77.9	1.15 (1.08-1.22)
Race/ethnicity			
Hispanic	475/607 (78.3)	31,465,187/40,277,007 (78.1); 72.1-83.1	1.15 (1.07-1.25)
Non-Hispanic Black	540/683 (79.1)	21,964,396/27,643,982 (79.5); 71.2-85.8	1.16 (1.05-1.28)
Non-Hispanic White	2113/3063 (69.0)	104,453,228/153,881,404 (67.9); 65.1-70.5	Reference (N/A)
Other	223/301 (74.1)	14,866,219/21,073,189 (70.5); 61.9-77.9	1.06 (0.94-1.18)
Age (years)			
18-34	699/1013 (69.0)	45,256,658/67,946,989 (66.6); 61.5-71.3	0.83 (0.77-0.91)
35-44	528/777 (68.0)	26,925,908/40,347,844 (66.7); 60.9-72.1	0.85 (0.77-0.93)
45-54	528/765 (69.0)	27,180,059/39,524,761 (68.8); 63.0-74.0	0.87 (0.80-0.95)
55-64	685/926 (74.0)	30,521,571/41,638,646 (73.3); 68.2-77.9	0.93 (0.87-1.01)
≥65	911/1173 (77.7)	42,864,834/53,417,341 (80.2); 76.4-83.6	Reference (N/A)
US Census region			
Northeast	359/476 (75.4)	32,878,801/42,937,799 (76.6); 71.2-81.2	1.05 (0.96-1.14)
Midwest	381/591 (64.5)	31,994,649/51,141,237 (62.6); 57.1-67.8	0.86 (0.78-0.96)
South	1607/2275 (70.6)	65,323,037/90,171,242 (72.4); 68.7-75.9	0.99 (0.92-1.07)
West	1004/1312 (76.5)	42,552,543/58,625,304 (72.6); 68.1-76.6	Reference (N/A)
Urbanicity			
Micropolitan/small town/rural	282/468 (60.3)	17,718,426/32,292,975 (54.9); 47.8-61.8	Reference (N/A)
Metropolitan	3069/4186 (73.3)	155,030,603/210,582,607 (73.6); 71.2-75.9	1.30 (1.14-1.48)
Education			
High school/General Educational Development (GED) or less	482/698 (69.1)	60,604,634/85,965,483 (70.5); 65.4-75.1	1.01 (0.93-1.10)
Some college/associate's degree	992/1409 (70.4)	48,068,329/69,226,861 (69.4); 65.6-73.0	0.98 (0.91-1.06)
Bachelor's degree	1036/1430 (72.4)	39,476,277/55,756,279 (70.8); 67.0-74.3	Reference (N/A)
Graduate degree	841/1117 (75.3)	24,599,790/31,926,958 (77.1); 73.3-80.4	1.07 (1.00-1.15)
Annual income (US $)			
0-24,999	512/721 (71.0)	21,167,549/29,566,723 (71.6); 65.2-77.2	1.04 (0.94-1.14)
25,000-49,999	659/916 (71.9)	28,730,742/41,443,877 (69.3); 63.3-74.7	0.99 (0.89-1.09)
50,000-99,999	1054/1445 (72.9)	51,366,352/73,211,031 (70.2); 65.8-74.2	Reference (N/A)
100,000-199,999	817/1125 (72.6)	48,935,593/67,795,060 (72.2); 67.8-76.2	1.02 (0.94-1.10)
≥200,000	309/447 (69.1)	22,548,792/30,858,891 (73.1); 66.8-78.5	1.03 (0.93-1.13)
Health insurance			
No health insurance	173/263 (65.8)	8,652,878/13,358,208 (64.8); 53.5-74.7	Reference (N/A)
Medicare/Medicaid/other	992/1352 (73.4)	50,432,470/66,230,875 (76.1); 72.2-79.7	1.20 (1.01-1.42

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
Private insurance/parent’s plan	1958/2734 (71.6)	101,517,515/147,299,448 (68.9); 65.9-71.8	1.08 (0.91-1.28)
Do not know	228/305 (74.8)	12,146,167/15,987,051 (76.0); 67.7-82.7	1.18 (0.98-1.44)
Month of sample collection			
August	830/1195 (69.5)	68,382,130/98,937,128 (69.1); 65.3-72.6	Reference (N/A)
September	301/406 (74.1)	23,861,095/33,460,432 (71.3); 64.3-77.4	1.03 (0.93-1.15)
October	596/812 (73.4)	40,495,885/55,101,083 (73.5); 68.7-77.8	1.06 (0.98-1.14)
November	1569/2165 (72.5)	39,080,513/53,835,755 (72.6); 68.1-76.6	1.06 (0.98-1.14)
December	55/76 (72.4)	929,406/1,541,184 (60.3); 35.2-81.0	0.99 (0.68-1.44)
Job type[c]			
Accommodation and food services	61/86 (70.9)	3,947,552/6,572,047 (60.1); 43.8-74.4	0.87 (0.66-1.14)
Educational services	261/334 (78.1)	8,107,735/10,277,744 (78.9); 71.9-84.5	1.12 (1.02-1.23)
Health care and social assistance	308/433 (71.1)	15,317,980/21,216,123 (72.2); 65.8-77.8	1.03 (0.93-1.14)
Retail trade	91/131 (69.5)	6,651,265/11,147,959 (59.7); 45.8-72.2	0.89 (0.71-1.12)
Transportation and warehousing	70/104 (67.3)	2,959,476/6,340,795 (46.7); 31.6-62.4	0.67 (0.47-0.96)
Other	1118/1606 (69.6)	60,001,477/86,625,204 (69.3); 65.3-72.9	Reference (N/A)
Work location[d]			
Completely indoors	735/1032 (71.2)	38,924,489/57,197,182 (68.1); 63.2-72.6	1.18 (1.03-1.35)
Completely outdoor/mixture/other	362/571 (63.4)	18,622,285/32,017,239 (58.2); 51.1-64.9	Reference (N/A)

[a] PR: prevalence ratio.

[b] N/A: not applicable.

[c] Among those who were employed.

[d] Among those who were employed and left home for work.

Mask Wearing

The estimated national period prevalence of consistently wearing a mask from August through December 2020 was 71.1% (95% CI 68.8-73.3; Table 1). Consistent mask wearing was significantly more prevalent among those who were female; were Hispanic or non-Hispanic Black; lived in a metropolitan area; had a graduate degree, were insured through Medicare, Medicaid, or other public health insurance; worked in educational services; or worked completely indoors (among those who were working from somewhere other than at home). Consistent mask wearing was significantly less prevalent among those who were less than 65 years old, lived in the Midwest (compared to the West), and worked in transportation or warehouse services.

Social Distancing

The estimated national prevalence of consistently practicing social distancing was 42.9% (95% CI 40.5-45.3; Table 2). Consistent social distancing was significantly more prevalent among those who were female, were Hispanic or non-Hispanic Black, lived in the South (compared to the West), had a graduate degree or some college education, or had an annual income of less than US $25,000. Consistent social distancing was significantly less prevalent among those aged 18-34 years (compared to ≥65 years) or lived in the Midwest (compared to the West).

Table 2. Prevalence of consistent social distancing among a household probability sample of 4654 US adults (August-December 2020).

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
Overall	2138/4654 (45.9)	104,253,682/242,875,582 (42.9); 40.5-45.3	N/A[b]
Sex			
Male	833/1927 (43.2)	46,015,009/115,613,214 (39.8); 36.2-43.5	Reference (N/A)
Female	1305/2727 (47.9)	58,238,673/127,262,368 (45.8); 42.6-48.9	1.15 (1.03-1.29)
Race/ethnicity			
Hispanic	274/607 (45.1)	18,911,041/40,277,007 (47.0); 40.5-53.5	1.19 (1.02-1.39)
Non-Hispanic Black	441/683 (64.6)	16,316,928/27,643,982 (59.0); 50.3-67.2	1.49 (1.27-1.76)
Non-Hispanic White	1298/3063 (42.4)	60,824,958/153,881,404 (39.5); 36.8-42.3	Reference (N/A)
Other	125/301 (41.5)	8,200,755/21,073,189 (38.9); 31.0-47.4	0.98 (0.78-1.23)
Age (years)			
18-34	351/1013 (34.6)	21,618,778/67,946,989 (31.8); 27.2-36.8	0.67 (0.56-0.81)
35-44	314/777 (40.4)	17,333,281/40,347,844 (43.0); 37.3-48.8	0.91 (0.77-1.08)
45-54	346/765 (45.2)	17,811,654/39,524,761 (45.1); 39.3-51.0	0.95 (0.81-1.13)
55-64	499/926 (53.9)	22,270,261/41,638,646 (53.5); 48.0-58.9	1.13 (0.98-1.31)
≥65	628/1173 (53.5)	25,219,708/53,417,341 (47.2); 42.4-52.1	Reference (N/A)
US Census region			
Northeast	201/476 (42.2)	18,423,748/42,937,799 (42.9); 37.1-48.9	1.03 (0.86-1.22)
Midwest	233/591 (39.4)	17,432,852/51,141,237 (34.1); 29.4-39.2	0.81 (0.68-0.97)
South	1107/2275 (48.7)	43,864,240/90,171,242 (48.6); 44.5-52.8	1.16 (1.01-1.33)
West	597/1312 (45.5)	24,532,842/58,625,304 (41.8); 37.5-46.3	Reference (N/A)
Urbanicity			
Micropolitan/small town/rural	219/468 (46.8)	13,688,195/32,292,975 (42.4); 35.7-49.4	Reference (N/A)
Metropolitan	1919/4186 (45.8)	90,565,487/210,582,607 (43.0); 40.5-45.6	1.01 (0.85-1.20)
Education			
High school/General Educational Development (GED) or less	351/698 (50.3)	37,216,103/85,965,483 (43.3); 38.2-48.5	1.13 (0.97-1.32)
Some college/associate's degree	641/1409 (45.5)	30,576,917/69,226,861 (44.2); 40.1-48.3	1.16 (1.01-1.32)
Bachelor's degree	605/1430 (42.3)	21,335,935/55,756,279 (38.3); 34.6-42.1	Reference (N/A)
Graduate degree	541/1117 (48.4)	15,124,727/31,926,958 (47.4); 43.0-51.8	1.23 (1.08-1.41)
Annual income (US $)			
0-24,999	380/721 (52.7)	15,386,978/29,566,723 (52.0); 45.4-58.6	1.22 (1.03-1.43)
25,000-49,999	418/916 (45.6)	17,864,754/41,443,877 (43.1); 37.3-49.1	1.01 (0.85-1.20)
50,000-99,999	671/1445 (46.4)	31,156,215/73,211,031 (42.6); 38.3-47.0	Reference (N/A)
100,000-199,999	484/1125 (43.0)	26,369,580/67,795,060 (38.9); 34.5-43.5	0.91 (0.78-1.06)
≥200,000	185/447 (41.4)	13,476,156/30,858,891 (43.7); 37.1-50.5	1.02 (0.85-1.23)
Health insurance			
No health insurance	118/263 (44.9)	6,131,940/13,358,208 (45.9); 34.9-57.3	Reference (N/A)
Medicare/Medicaid/other	727/1352 (53.8)	33,026,689/66,230,875 (49.9); 45.2-54.5	1.08 (0.83-1.42)
Private insurance/parent's plan	1155/2734 (42.2)	58,149,331/147,299,448 (39.5); 36.5-42.5	0.86 (0.66-1.11)
Do not know	138/305 (45.2)	6,945,721/15,987,051 (43.4); 34.3-53.0	0.95 (0.68-1.33)
Month of sample collection			

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
August	527/1195 (44.1)	43,165,763/98,937,128 (43.6); 39.8-47.6	Reference (N/A)
September	196/406 (48.3)	15,505,163/33,460,432 (46.3); 39.5-53.3	1.06 (0.89-1.26)
October	348/812 (42.9)	21,917,341/55,101,083 (39.8); 34.9-44.9	0.92 (0.78-1.07)
November	1028/2165 (47.5)	23,238,793/53,835,755 (43.2); 38.7-47.7	0.99 (0.86-1.14)
December	39/76 (51.3)	426,622/1,541,184 (27.7); 14.1-47.2	0.64 (0.33-1.23)
Job type[c]			
Accommodation and food services	31/86 (36.0)	1,889,223/6,572,047.14 (28.7); 17.2-43.9	0.72 (0.43-1.18)
Educational services	148/334 (44.3)	4,010,043/10,277,744.49 (39.0); 31.1-47.6	0.97 (0.77-1.24)
Health care and social assistance	171/433 (39.5)	7,690,760/21,216,123.32 (36.2); 30.0-43.0	0.91 (0.74-1.12)
Retail trade	57/131 (43.5)	5,088,591/11,147,959.41 (45.6); 33.1-58.8	1.14 (0.83-1.56)
Transportation and warehousing	35/104 (33.7)	1,687,619/6,340,794.69 (26.6); 16.1-40.7	0.66 (0.41-1.09)
Other	663/1606 (41.3)	34,585,460/86,625,204.41 (39.9); 36.1-43.9	Reference (N/A)
Work location[d]			
Completely indoors	384/1032 (37.2)	19,740,169/57,197,182 (34.5); 30.2-39.1	0.96 (0.77-1.21)
Completely outdoor/mixture/other	210/571 (36.8)	11,466,330/32,017,239 (35.8); 29.4-42.7	Reference (N/A)

[a] PR: prevalence ratio.

[b] N/A: not applicable.

[c] Among those who were employed.

[d] Among those who were employed and left home for work.

Handwashing

Among the 4090 participants who were administered the hand hygiene questions, the average number of times the participants washed hands in the past 24 hours was 8.8 (SE 0.3). The estimated national prevalence of individuals frequently washing hands was 55.0% (95% CI 52.3-57.7; Table 3). Frequent handwashing was significantly more prevalent among those who were female, were Hispanic or non-Hispanic Black, were aged 35-54 years (compared to ≥65 years), were enrolled in November 2020 (compared to August 2020), or worked in health care and social assistance services.

Table 3. Prevalence of frequent handwashing among a household probability sample of 4090 US adults (August-December 2020).

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
Overall	2226/4090 (54.4)	107,258,747/195,041,917 (55.0); 52.3-57.7	N/A[b]
Sex			
Male	746/1682 (44.4)	40,507,668/92,525,526 (43.8); 39.7-47.9	Reference (N/A)
Female	1480/2408 (61.5)	66,751,079/102,516,391 (65.1); 61.8-68.3	1.48 (1.34-1.64)
Race/ethnicity			
Hispanic	341/584 (58.4)	21,843,430/36,294,202 (60.2); 53.5-66.5	1.17 (1.04-1.32)
Non-Hispanic Black	376/661 (56.9)	14,430,982/23,137,584 (62.4); 53.6-70.4	1.17 (1.01-1.36)
Non-Hispanic White	1367/2578 (53.0)	62,577,192/119,543,448 (52.3); 49.1-55.6	Reference (N/A)
Other	142/267 (53.2)	8,407,143/16,066,683 (52.3); 43.1-61.4	1.02 (0.85-1.22)
Age (years)			
18-34	479/900 (53.2)	29,369,641/54,991,956 (53.4); 47.7-59.0	1.05 (0.90-1.21)
35-44	405/685 (59.1)	19,647,980/32,573,737 (60.3); 54.0-66.3	1.23 (1.07-1.41)
45-54	379/656 (57.8)	17,923,525/30,308,444 (59.1); 52.4-65.5	1.18 (1.01-1.37)
55-64	442/825 (53.6)	18,533,469/34,713,815 (53.4); 47.4-59.3	1.07 (0.92-1.24)
≥65	521/1024 (50.9)	21,784,132/42,453,965 (51.3); 46.0-56.6	Reference (N/A)
US Census region			
Northeast	159/301 (52.8)	13,690,592/26,689,317 (51.3); 43.8-58.8	0.98 (0.83-1.15)
Midwest	272/462 (58.9)	23,078,467/40,022,544 (57.7); 51.6-63.5	1.08 (0.95-1.22)
South	1081/2018 (53.6)	38,443,937/69,829,637 (55.1); 50.4-59.7	1.02 (0.91-1.14)
West	714/1309 (54.5)	32,045,751/58,500,420 (54.8); 50.3-59.2	Reference (N/A)
Urbanicity			
Micropolitan/small town/rural	213/403 (52.9)	13,506,913/26,235,378 (51.5); 43.7-59.2	Reference (N/A)
Metropolitan	2013/3687 (54.6)	93,751,834/168,806,540 (55.5); 52.7-58.4	1.06 (0.91-1.24)
Education			
High school/General Educational Development (GED) or less	309/609 (50.7)	36,508,095/69,755,426 (52.3); 46.5-58.1	1.00 (0.88-1.13)
Some college/associate's degree	729/1268 (57.5)	33,497,775/56,483,039 (59.3); 54.9-63.6	1.08 (0.97-1.20)
Bachelor's degree	660/1256 (52.5)	24,259,383/43,783,023 (55.4); 51.2-59.5	Reference (N/A)
Graduate degree	528/957 (55.2)	12,993,495/25,020,429 (51.9); 47.0-56.9	0.93 (0.82-1.04)
Annual income (US $)			
0-24,999	353/640 (55.2)	12,398,320/22,961,654 (54.0); 46.8-61.1	1.03 (0.88-1.20)
25,000-49,999	428/811 (52.8)	18,172,778/33,333,054 (54.5); 47.9-60.9	1.05 (0.91-1.21)
50,000-99,999	720/1286 (56.0)	32,722,037/61,651,585 (53.1); 48.2-57.9	Reference (N/A)
100,000-199,999	519/966 (53.7)	29,925,499/52,932,785 (56.5); 51.3-61.6	1.04 (0.92-1.18)
≥200,000	206/387 (53.2)	14,040,112/24,162,839 (58.1); 50.8-65.0	1.07 (0.92-1.24)
Health insurance			
No health insurance	125/246 (50.8)	5,896,607/11,801,478 (50.0); 38.0-61.9	Reference (N/A)
Medicare/Medicaid/other	616/1191 (51.7)	26,735,424/52,283,277 (51.1); 46.0-56.2	1.04 (0.81-1.35)
Private insurance/parent's plan	1348/2393 (56.3)	68,204,026/118,431,180 (57.6); 54.2-60.9	1.14 (0.89-1.46)
Do not know	137/260 (52.7)	6,422,690/12,525,982 (51.3); 40.6-61.8	1.05 (0.76-1.45)

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
Month of sample collection			
August	352/655 (53.7)	26,438,562/52,551,995 (50.3); 45.0-55.6	Reference (N/A)
September	224/392 (57.1)	17,794,291/32,683,426 (54.4); 47.3-61.4	1.06 (0.91-1.25)
October	429/806 (53.2)	30,795,988/54,773,380 (56.2); 51.0-61.3	1.07 (0.93-1.22)
November	1176/2161 (54.4)	31,727,178/53,491,933 (59.3); 54.8-63.7	1.15 (1.02-1.31)
December	45/76 (59.2)	502,728/1,541,184 (32.6); 16.4-54.4	0.71 (0.38-1.31)
Job type[c]			
Accommodation and food services	53/79 (67.1)	3,700,939/5,866,155 (63.1); 45.6-77.7	1.19 (0.90-1.57)
Educational services	171/291 (58.8)	4,875,979/7,963,246 (61.2); 51.8-69.9	1.14 (0.96-1.35)
Health care and social assistance	251/384 (65.4)	12,160,332/17,500,465 (69.5); 62.5-75.7	1.30 (1.14-1.47)
Retail trade	67/106 (63.2)	4,897,239/8,228,603 (59.5); 44.0-73.3	1.23 (0.97-1.55)
Transportation and warehousing	58/100 (58.0)	3,279,732/5,854,707 (56.0); 39.4-71.4	1.05 (0.77-1.44)
Other	726/1412 (51.4)	37,061,432/70,086,074 (52.9); 48.4-57.3	Reference (N/A)
Work location[d]			
Completely indoors	568/907 (62.6)	29,785,987/47,435,209 (62.8); 57.4-67.8	1.11 (0.96-1.29)
Completely outdoor/mixture/other	277/507 (54.6)	14,909,587/26,354,333 (56.6); 49.0-63.9	Reference (N/A)

[a] PR: prevalence ratio.

[b] N/A: not applicable.

[c] Among those who were employed.

[d] Among those who were employed and left home for work.

Hand Sanitizer Use

Among the 4090 participants who were administered the hand hygiene questions, the average number of times the participants used a hand sanitizer in the past 24 hours was 4.99 (SE 0.2). The estimated national prevalence of frequently using a hand sanitizer was 21.5% (95% CI 19.4-23.8; Table 4). Frequent use of a hand sanitizer was significantly more prevalent among those who were female, Hispanic or non-Hispanic Black, less than 65 years of age, lived in the South (compared to the West), had an annual income of less than US $25,000, were enrolled in September or November 2020 (compared to August 2020), or worked in accommodation, food services, health care, social assistance, retail trade, or transportation/warehouse services. Frequent use of a hand sanitizer was significantly less prevalent among those who had an annual income of US $100,000-$199,000 or were insured through Medicare, Medicaid, or other public health insurance (compared to those uninsured).

Table 4. Prevalence of frequent hand sanitizer use among a household probability sample of 4090 US adults (August-December 2020).

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
Overall	846/4090 (20.7)	41,964,720/195,041,917 (21.5); 19.4-23.8	N/A[b]
Sex			
Male	293/1682 (17.4)	17,369,555/92,525,526 (18.8); 15.7-22.3	Reference (N/A)
Female	553/2408 (23.0)	24,595,165/102,516,391 (24.0); 21.1-27.1	1.26 (1.02-1.56)
Race/ethnicity			
Hispanic	176/584 (30.1)	11,525,753/36,294,202 (31.8); 25.8-38.4	1.92 (1.51-2.44)
Non-Hispanic Black	197/661 (29.8)	7,287,691/23,137,584 (31.5); 23.8-40.3	1.84 (1.37-2.48)
Non-Hispanic White	422/2578 (16.4)	20,054,567/119,543,448 (16.8); 14.6-19.2	Reference (N/A)
Other	51/267 (19.1)	3,096,709/16,066,683 (19.3); 13.5-26.8	1.16 (0.80-1.68)
Age (years)			
18-34	233/900 (25.9)	14,220,072/54,991,956 (25.9); 21.2-31.1	2.14 (1.48-3.11)
35-44	177/685 (25.8)	9,327,426/32,573,737 (28.6); 23.4-34.5	2.48 (1.71-3.60)
45-54	150/656 (22.9)	6,848,252/30,308,444 (22.6); 17.5-28.6	1.92 (1.28-2.88)
55-64	163/825 (19.8)	6,460,726/34,713,815 (18.6); 14.6-23.4	1.62 (1.09-2.41)
≥65	123/1024 (12.0)	5,108,245/42,453,965 (12.0); 8.7-16.4	Reference (N/A)
US Census region			
Northeast	59/301 (19.6)	5,660,463/26,689,317 (21.2); 15.8-27.9	1.12 (0.80-1.57)
Midwest	73/462 (15.8)	6,519,188/40,022,544 (16.3); 12.5-20.9	0.83 (0.60-1.13)
South	478/2018 (23.7)	18,042,211/69,829,637 (25.8); 21.9-30.2	1.31 (1.03-1.68)
West	236/1309 (18.0)	11,742,858/58,500,420 (20.1); 16.6-24.0	Reference (N/A)
Urbanicity			
Micropolitan/small town/rural	81/403 (20.1)	4,709,547/26,235,378 (18.0); 13.0-24.3	Reference (N/A)
Metropolitan	765/3687 (20.7)	37,255,173/168,806,540 (22.1); 19.7-24.6	1.21 (0.87-1.69)
Education			
High school/General Educational Development (GED) or less	141/609 (23.2)	16,518,098/69,755,426 (23.7); 19.2-28.9	1.69 (1.27-2.25)
Some college/associate's degree	309/1268 (24.4)	13,288,148/56,483,039 (23.5); 20.1-27.4	1.60 (1.24-2.06)
Bachelor's degree	210/1256 (16.7)	6,499,265/43,783,023 (14.8); 12.1-18.0	Reference (N/A)
Graduate degree	186/957 (19.4)	5,659,209/25,020,429 (22.6); 18.4-27.5	1.50 (1.13-1.99)
Annual income (US $)			
0-24,999	156/640 (24.4)	6,956,140/22,961,654 (30.3); 23.8-37.6	1.36 (1.01-1.81)
25,000-49,999	188/811 (23.2)	7,970,726/33,333,054 (23.9); 18.8-29.9	1.07 (0.79-1.43)
50,000-99,999	267/1286 (20.8)	13,913,036/61,651,585 (22.6); 18.6-27.1	Reference (N/A)
100,000-199,999	179/966 (18.5)	8,941,815/52,932,785 (16.9); 13.6-20.8	0.73 (0.55-0.96)
≥200,000	56/387 (14.5)	4,183,004/24,162,839 (17.3); 12.3-23.8	0.75 (0.51-1.09)
Health insurance			
No health insurance	69/246 (28.0)	3,625,449/11,801,478 (30.7); 20.6-43.1	Reference (N/A)
Medicare/Medicaid/other	196/1191 (16.5)	9,517,303/52,283,277 (18.2); 14.4-22.8	0.62 (0.40-0.97)
Private insurance/parent's plan	510/2393 (21.3)	25,573,342/118,431,180 (21.6); 18.9-24.5	0.71 (0.48-1.05)
Do not know	71/260 (27.3)	3,248,626/12,525,982 (25.9); 18.2-35.5	0.88 (0.53-1.46)
Month of sample collection			

Characteristics	Unweighted	Weighted	
	Prevalence, n/N (%)	Prevalence, n/N (%); 95% CI	PR[a] (95% CI)
August	101/655 (15.4)	9,075,148/52,551,995 (17.3); 13.6-21.7	Reference (N/A)
September	106/392 (27.0)	8,760,306/32,683,426 (26.8); 21.0-33.5	1.51 (1.09-2.10)
October	144/806 (17.9)	11,226,992/54,773,380 (20.5); 16.6-25.0	1.12 (0.82-1.53)
November	479/2161 (22.2)	12,787,219/53,491,933 (23.9); 20.1-28.2	1.34 (1.01-1.79)
December	16/76 (21.1)	115,054/1,541,184 (7.5); 2.6-19.6	0.46 (0.15-1.42)
Job type[c]			
Accommodation and food services	26/79 (32.9)	2,105,92/5,866,155 (35.9); 21.9-52.8	1.86 (1.14-3.06)
Educational services	72/291 (24.7)	1,926,548/7,963,246 (24.2); 17.5-32.4	1.24 (0.86-1.78)
Health care and social assistance	164/384 (42.7)	7,894,448/17,500,465 (45.1); 37.9-52.6	2.35 (1.84-3.00)
Retail trade	37/106 (34.9)	3,108,150/8,228,603 (37.8); 24.4-53.4	2.15 (1.40-3.29)
Transportation and warehousing	32/100 (32.0)	2,187,560/5,854,707 (37.4); 22.7-54.8	1.94 (1.18-3.18)
Other	262/1412 (18.6)	13,321,475/70,086,074 (19.0); 15.7-22.8	Reference (N/A)
Work location[d]			
Completely indoors	297/907 (32.7)	16,448,119/47,435,209 (34.7); 29.8-39.9	1.14 (0.87-1.49)
Completely outdoor/mixture/other	143/507 (28.2)	8,086,420/26,354,333 (30.7); 24.2-38.1	Reference (N/A)

[a] PR: prevalence ratio.

[b] N/A: not applicable.

[c] Among those who were employed.

[d] Among those who were employed and left home for work.

Mitigation Classifications

The final classification model identified 3 latent classes: (1) optimal mitigation, (2) optimal mitigation with additional hand hygiene, and (3) lowest mitigation. Optimal mitigation was consistent mask wearing, consistent social distancing, and handwashing or hand sanitizer use 6-10 times per day. Optimal mitigation with additional hand hygiene was consistent mask wearing, consistent social distancing, and handwashing or hand sanitizer use >11 times per day. The lowest mitigation was inconsistent mask wearing, inconsistent social distancing, and handwashing or hand sanitizer use 0-5 times per day. There were no classes that had suboptimal use of only some mitigation strategies but optimal use of others. All participants were categorized into these classes. Two-thirds (n=2656, 67%) practiced optimal mitigation by consistently wearing a mask, consistently following social distancing, and frequently washing their hands or using a hand sanitizer (Tables 5-7). Furthermore, 1 in 5 (n=771, 20.6%) practiced the lowest mitigation by inconsistently or infrequently engaging in all mitigation practices. The final class made up the remainder (n=463, 12.4%) who consistently wore masks and maintained social distance but had the highest frequency of handwashing or sanitizer use (>11 times per day). Compared to optimal mitigation practices, the likelihood of being in the lowest-mitigation class was significantly greater among those who were male, less than 65 years of age, lived in the Midwest (compared to the West), lived outside a metropolitan area, had no health insurance (compared to Medicare, Medicaid, or other public insurance), worked in transportation or warehouse services, or worked somewhere other than completely indoors (Table 8). Compared to just the optimal-mitigation class, the likelihood of being in the class with optimal mask wearing and social distancing but with additional handwashing or sanitizer use was significantly greater among those who were Hispanic or non-Hispanic Black, were less than 65 years of age, had less than a bachelor’s degree, or worked in any of the selected job types (compared to other jobs).

Compared to the optimal-mitigation class, those in the lowest-mitigation class were 20% less likely (PR 0.80, CI 0.79-0.90) to agree that masks provide 95% or better protection against COVID-19 and were twice as likely (PR 2.00, CI 1.27-3.15) to state that it was not necessary for youth to take measures to prevent COVID-19 infection (Table 8). There were no significant differences between the optimal-mitigation class and the class with additional hand hygiene for both the mask-wearing (PR 0.93, CI 0.82-1.06) and youth prevention (PR 1.70, CI 0.92-3.16) questions.

Table 5. “Optimal mitigation” latent class of combined strategies to prevent COVID-19 among a household probability sample of 4090 US adults (August-December 2020).

Characteristics	Total sample		Optimal mitigation (consistent masking and social distancing, hand hygiene 6-10 times/day)	
	Unweighted, N	Weighted, N	Unweighted prevalence, n (%)	Weighted prevalence, n (%)
Overall	3863	183,171,244	2656 (68.8)	122,800,910 (67.0)
Sex				
Male	1603	86,348,193	1102 (68.7)	56,182,229 (65.1)
Female	2260	96,823,051	1554 (68.8)	66,618,681 (68.8)
Race/ethnicity				
Hispanic	551	33,539,313	364 (66.1)	21,961,242 (65.5)
Non-Hispanic Black	607	22,219,194	423 (69.7)	15,006,373 (67.5)
Non-Hispanic White	2454	112,449,529	1678 (68.4)	74,776,252 (66.5)
Other	251	14,963,209	191 (76.1)	11,057,042 (73.9)
Age (years)				
18-34	869	52,755,525	538 (61.9)	32,268,979 (61.2)
35-44	643	29,840,688	405 (63.0)	17,937,534 (60.1)
45-54	608	28,183,232	400 (65.8)	17,959,527 (63.7)
55-64	774	31,700,429	542 (70.0)	21,137,563 (66.7)
≥65	969	40,691,370	771 (79.6)	33,497,307 (82.3)
US Census region				
Northeast	279	24,075,803	203 (72.8)	17,676,409 (73.4)
Midwest	438	37,524,032	285 (65.1)	23,288,505 (62.1)
South	1896	65,519,953	1249 (65.9)	43,782,348 (66.8)
West	1250	56,051,456	919 (73.5)	38,053,648 (67.9)
Urbanicity				
Micropolitan/small town/rural	374	24,397,071	231 (61.8)	14,006,081 (57.4)
Metropolitan	3489	158,774,173	2425 (69.5)	108,794,829 (68.5)
Education				
High school/General Educational Development (GED) or less	543	63,033,498	350 (64.5)	39,876,092 (63.3)
Some college/associate’s degree	1189	53,702,228	766 (64.4)	34,863,092 (64.9)
Bachelor’s degree	1203	42,013,058	862 (71.7)	29,934,515 (71.3)
Graduate degree	928	24,422,460	678 (73.1)	18,127,211 (74.2)
Annual income (US $)				
0-24,999	586	21,039,489	396 (67.6)	13,895,746 (66.0)
25,000-49,999	756	30,682,885	504 (66.7)	18,888,597 (61.6)
50,000-99,999	1220	57,414,158	841 (68.9)	37,446,727 (65.2)
100,000-199,999	934	50,966,442	663 (71.0)	35,542,740 (69.7)
≥200,000	367	23,068,271	252 (68.7)	17,027,100 (73.8)
Health insurance				
No health insurance	230	11,173,450	138 (60.0)	6,510,927 (58.3)
Medicare/Medicaid/other	1101	47,572,527	824 (74.8)	35,945,737 (75.6)
Private insurance/parent’s plan	2294	112,869,879	1535 (66.9)	72,421,450 (64.2)
Do not know	238	11,555,388	159 (66.8)	7,922,797 (68.6)

Characteristics	Total sample		Optimal mitigation (consistent masking and social distancing, hand hygiene 6-10 times/day)	
	Unweighted, N	Weighted, N	Unweighted prevalence, n (%)	Weighted prevalence, n (%)
Month of sample collection				
August	619	48,004,288	430 (69.5)	32,876,252 (68.5)
September	372	30,679,522	242 (65.1)	18,957,076 (61.8)
October	775	52,803,927	557 (71.9)	36,692,789 (69.5)
November	2026	50,370,899	1376 (67.9)	33,317,935 (66.1)
December	71	1,312,608	51 (71.8)	956,860 (72.9)
Job type[a]				
Accommodation and food services	77	5,669,023	44 (57.1)	3,208,728 (56.6)
Educational services	283	7,801,317	195 (68.9)	5,538,379 (71.0)
Health care and social assistance	363	16,769,318	193 (53.2)	8,084,708 (48.2)
Retail trade	98	7,256,895	57 (58.2)	3,524,372 (48.6)
Transportation and warehousing	93	5,661,875	56 (60.2)	2,458,090 (43.4)
Other	1354	66,741,897	925 (68.3)	44,833,174 (67.2)
Work location[b]				
Completely indoors	868	44,926,729	519 (59.8)	25,143,068 (56.0)
Completely outdoor/mixture/other	485	24,905,513	276 (56.9)	11,660,499 (46.8)

[a] Among those who were employed.

[b] Among those who were employed and left home for work.

Table 6. "Optimal mitigation plus additional hand hygiene" latent class of combined strategies to prevent COVID-19 among a household probability sample of 4090 US adults (August-December 2020).

Characteristics	Total sample		Optimal mitigation plus additional hand hygiene (consistent masking and social distancing, hand hygiene ≥11 times/day)	
	Unweighted, N	Weighted, N	Unweighted prevalence, n (%)	Weighted prevalence, n (%)
Overall	3863	183,171,244	436 (11.3)	22,548,164 (12.3)
Sex				
Male	1603	86,348,193	118 (7.4)	8,424,160 (9.8)
Female	2260	96,823,051	318 (14.1)	14,124,004 (14.6)
Race/ethnicity				
Hispanic	551	33,539,313	103 (18.7)	6,328,829 (18.9)
Non-Hispanic Black	607	22,219,194	105 (17.3)	4,443,359 (20.0)
Non-Hispanic White	2454	112,449,529	205 (8.4)	10,067,066 (9.0)
Other	251	14,963,209	23 (9.2)	1,708,910 (11.4)
Age (years)				
18-34	869	52,755,525	121 (13.9)	7,686,991 (14.6)
35-44	643	29,840,688	89 (13.8)	4,992,781 (16.7)
45-54	608	28,183,232	86 (14.1)	4,417,403 (15.7)
55-64	774	31,700,429	90 (11.6)	3,907,870 (12.3)
≥65	969	40,691,370	50 (5.2)	1,543,119 (3.8)
US Census region				
Northeast	279	24,075,803	28 (10.0)	2,367,057 (9.8)
Midwest	438	37,524,032	40 (9.1)	3,709,602 (9.9)
South	1896	65,519,953	240 (12.7)	9,277,429 (14.2)
West	1250	56,051,456	128 (10.2)	7,194,076 (12.8)
Urbanicity				
Micropolitan/small town/rural	374	24,397,071	44 (11.8)	3,079,673 (12.6)
Metropolitan	3489	158,774,173	392 (11.2)	19,468,491 (12.3)
Education				
High school/General Educational Development (GED) or less	543	63,033,498	75 (13.8)	9,177,714 (14.6)
Some college/associate's degree	1189	53,702,228	178 (15.0)	7,573,166 (14.1)
Bachelor's degree	1203	42,013,058	94 (7.8)	3,420,744 (8.1)
Graduate degree	928	24,422,460	89 (9.6)	2,376,540 (9.7)
Annual income (US $)				
0-24,999	586	21,039,489	86 (14.7)	4,030,880 (19.2)
25,000-49,999	756	30,682,885	99 (13.1)	4,365,638 (14.2)
50,000-99,999	1220	57,414,158	141 (11.6)	7,137,152 (12.4)
100,000-199,999	934	50,966,442	85 (9.1)	5,067,508 (9.9)
≥200,000	367	23,068,271	25 (6.8)	1,946,986 (8.4)
Health insurance				
No health insurance	230	11,173,450	32 (13.9)	1,547,869 (13.9)
Medicare/Medicaid/other	1101	47,572,527	96 (8.7)	4,518,681 (9.5)
Private insurance/parent's plan	2294	112,869,879	272 (11.9)	14,983,739 (13.3)
Do not know	238	11,555,388	36 (15.1)	1,497,874 (13.0)

Characteristics	Total sample		Optimal mitigation plus additional hand hygiene (consistent masking and social distancing, hand hygiene ≥11 times/day)	
	Unweighted, N	Weighted, N	Unweighted prevalence, n (%)	Weighted prevalence, n (%)
Month of sample collection				
August	619	48,004,288	51 (8.2)	4,344,792 (9.1)
September	372	30,679,522	59 (15.9)	5,007,293 (16.3)
October	775	52,803,927	70 (9.0)	5,658,534 (10.7)
November	2026	50,370,899	246 (12.1)	7,502,823 (14.9)
December	71	1,312,608	10 (14.1)	34,721 (2.6)
Job type[a]				
Accommodation and food services	77	5,669,023	18 (23.4)	920,635 (16.2)
Educational services	283	7,801,317	40 (14.1)	952,513 (12.2)
Health care and social assistance	363	16,769,318	101 (27.8)	5,864,653 (35.0)
Retail trade	98	7,256,895	27 (27.6)	2,517,580 (34.7)
Transportation and warehousing	93	5,661,875	15 (16.1)	1,174,633 (20.7)
Other	1354	66,741,897	104 (7.7)	5,333,381 (8.0)
Work location[b]				
Completely indoors	868	44,926,729	157 (18.1)	8,661,383 (19.3)
Completely outdoor/mixture/other	485	24,905,513	72 (14.8)	4,798,204 (19.3)

[a] Among those who were employed.

[b] Among those who were employed and left home for work.

Table 7. “Lowest mitigation” latent class of combined strategies to prevent COVID-19 among a household probability sample of 4090 US adults (August-December 2020).

Characteristics	Total sample		Lowest mitigation (inconsistent masking and social distancing, hand hygiene 0-5 times/day)	
	Unweighted, N	Weighted, N	Unweighted prevalence, n (%)	Weighted prevalence, n (%)
Overall	3863	183,171,244	771 (20.0)	37,822,170 (20.6)
Sex				
Male	1603	86,348,193	383 (23.9)	21,741,804 (25.2)
Female	2260	96,823,051	388 (17.2)	16,080,366 (16.6)
Race/ethnicity				
Hispanic	551	33,539,313	84 (15.2)	5,249,241 (15.7)
Non-Hispanic Black	607	22,219,194	79 (13.0)	2,769,462 (12.5)
Non-Hispanic White	2454	112,449,529	571 (23.3)	27,606,210 (24.5)
Other	251	14,963,209	37 (14.7)	2,197,256 (14.7)
Age (years)				
18-34	869	52,755,525	210 (24.2)	12,799,555 (24.3)
35-44	643	29,840,688	149 (23.2)	6,910,373 (23.2)
45-54	608	28,183,232	122 (20.1)	5,806,302 (20.6)
55-64	774	31,700,429	142 (18.3)	6,654,996 (21.0)
≥65	969	40,691,370	148 (15.3)	5,650,944 (13.9)
US Census region				
Northeast	279	24,075,803	48 (17.2)	4,032,338 (16.7)
Midwest	438	37,524,032	113 (25.8)	10,525,924 (28.1)
South	1896	65,519,953	407 (21.5)	12,460,175 (19.0)
West	1250	56,051,456	203 (16.2)	10,803,732 (19.3)
Urbanicity				
Micropolitan/small town/rural	374	24,397,071	99 (26.5)	7,311,317 (30.0)
Metropolitan	3489	158,774,173	672 (19.3)	30,510,853 (19.2)
Education				
High school/General Educational Development (GED) or less	543	63,033,498	118 (21.7)	13,979,692 (22.2)
Some college/associate’s degree	1189	53,702,228	245 (20.6)	11,265,970 (21.0)
Bachelor’s degree	1203	42,013,058	247 (20.5)	8,657,800 (20.6)
Graduate degree	928	24,422,460	161 (17.3)	3,918,708 (16.0)
Annual income (US $)				
0-24,999	586	21,039,489	104 (17.7)	3,112,863 (14.8)
25,000-49,999	756	30,682,885	153 (20.2)	7,428,650 (24.2)
50,000-99,999	1220	57,414,158	238 (19.5)	12,830,279 (22.3)
100,000-199,999	934	50,966,442	186 (19.9)	10,356,193 (20.3)
≥200,000	367	23,068,271	90 (24.5)	4,094,185 (17.7)
Health insurance				
No health insurance	230	11,173,450	60 (26.1)	3,114,654 (27.9)
Medicare/Medicaid/other	1101	47,572,527	181 (16.4)	7,108,109 (14.9)
Private insurance/parent’s plan	2294	112,869,879	487 (21.2)	25,464,690 (22.6)
Do not know	238	11,555,388	43 (18.1)	2,134,717 (18.5)

Characteristics	Total sample		Lowest mitigation (inconsistent masking and social distancing, hand hygiene 0-5 times/day)	
	Unweighted, N	Weighted, N	Unweighted prevalence, n (%)	Weighted prevalence, n (%)
Month of sample collection				
August	619	48,004,288	138 (22.3)	10,783,244 (22.5)
September	372	30,679,522	71 (19.1)	6,715,153 (21.9)
October	775	52,803,927	148 (19.1)	10,452,603 (19.8)
November	2026	50,370,899	404 (19.9)	9,550,142 (19.0)
December	71	1,312,608	10 (14.1)	321,027 (24.5)
Job type[a]				
Accommodation and food services	77	5,669,023	15 (19.5)	1,539,660 (27.2)
Educational services	283	7,801,317	48 (17.0)	1,310,424 (16.8)
Health care and social assistance	363	16,769,318	69 (19.0)	2,819,957 (16.8)
Retail trade	98	7,256,895	14 (14.3)	1,214,943 (16.7)
Transportation and warehousing	93	5,661,875	22 (23.7)	2,029,152 (35.8)
Other	1354	66,741,897	325 (24.0)	16,575,342 (24.8)
Work location[b]				
Completely indoors	868	44,926,729	192 (22.1)	11,122,278 (24.8)
Completely outdoor/mixture/other	485	24,905,513	137 (28.2)	8,446,810 (33.9)

[a] Among those who were employed.

[b] Among those who were employed and left home for work.

Table 8. Comparison of participant characteristics by latent classes of combined strategies to prevent COVID-19 among a household probability sample of 4090 US adults (August-December 2020).

Characteristics	Additional hand hygiene[a] vs optimal mitigation[b], PR[c] (95% CI)	Lowest mitigation[d] vs optimal mitigation, PR (95% CI)
Sex		
Male	Reference (N/A[e])	Reference (N/A)
Female	1.34 (0.97-1.86)	0.70 (0.56-0.87)
Race/ethnicity		
Hispanic	1.89 (1.33-2.68)	0.72 (0.51-1.01)
Non-Hispanic Black	1.93 (1.27-2.92)	0.58 (0.33-1.00)
Non-Hispanic White	Reference (N/A)	Reference (N/A)
Other	1.13 (0.66-1.93)	0.61 (0.37-1.01)
Age (years)		
18-34	4.37 (2.45-7.80)	1.97 (1.42-2.72)
35-44	4.94 (2.76-8.84)	1.93 (1.36-2.72)
45-54	4.48 (2.47-8.14)	1.69 (1.16-2.47)
55-64	3.54 (1.92-6.53)	1.66 (1.16-2.36)
≥65	Reference (N/A)	Reference (N/A)
US Census region		
Northeast	0.74 (0.44-1.25)	0.84 (0.57-1.25)
Midwest	0.86 (0.55-1.35)	1.41 (1.06-1.87)
South	1.10 (0.78-1.56)	1.00 (0.76-1.32)
West	Reference (N/A)	Reference (N/A)
Urbanicity		
Micropolitan/small town/rural	Reference (N/A)	Reference (N/A)
Metropolitan	0.84 (0.54-1.32)	0.64 (0.49-0.84)
Education		
High school/General Educational Development (GED) or less	1.82 (1.21-2.75)	1.16 (0.87-1.54)
Some college/associate’s degree	1.74 (1.21-2.51)	1.09 (0.86-1.38)
Bachelor’s degree	Reference (N/A)	Reference (N/A)
Graduate degree	1.13 (0.74-1.72)	0.79 (0.60-1.05)
Annual income (US $)		
0-24,999	1.40 (0.92-2.14)	0.72 (0.50-1.02)
25,000-49,999	1.17 (0.77-1.78)	1.11 (0.80-1.53)
50,000-99,999	Reference (N/A)	Reference (N/A)
100,000-199,999	0.78 (0.51-1.18)	0.88 (0.66-1.18)
≥200,000	0.64 (0.36-1.13)	0.76 (0.54-1.06)
Health insurance		
No health insurance	Reference (N/A)	Reference (N/A)
Medicare/Medicaid/other	0.58 (0.29-1.16)	0.51 (0.32-0.81)
Private insurance/parent’s plan	0.89 (0.48-1.67)	0.80 (0.53-1.22)
Do not know	0.83 (0.37-1.84)	0.66 (0.37-1.18)
Month of sample collection		

Characteristics	Additional hand hygiene[a] vs optimal mitigation[b], PR[c] (95% CI)	Lowest mitigation[d] vs optimal mitigation, PR (95% CI)
August	Reference (N/A)	Reference (N/A)
September	1.79 (1.12-2.85)	1.06 (0.75-1.50)
October	1.14 (0.73-1.80)	0.90 (0.67-1.21)
November	1.57 (1.04-2.38)	0.90 (0.68-1.19)
December	0.30 (0.11-0.78)	1.02 (0.33-3.11)
Job type[f]		
Accommodation and food services	2.10 (1.05-4.20)	1.20 (0.65-2.22)
Educational services	1.38 (1.05-4.20)	0.71 (0.47-1.08)
Health care and social assistance	3.95 (2.70-5.79)	0.96 (0.68-1.36)
Retail trade	3.92 (2.28-6.74)	0.95 (0.45-1.99)
Transportation and warehousing	3.04 (1.46-6.35)	1.68 (1.04-2.70)
Other	Reference (N/A)	Reference (N/A)
Work location[g]		
Completely indoors	0.88 (0.60-1.28)	0.73 (0.55-0.97)
Completely outdoor/mixture/other	Reference (N/A)	Reference (N/A)

[a] Optimal mitigation plus additional hand hygiene is consistent masking and social distancing, as well as hand hygiene ≥11 times/day.

[b] Optimal mitigation is consistent masking and social distancing, as well as hand hygiene 6-10 times/day.

[c] PR: prevalence ratio.

[d] The lowest mitigation is inconsistent masking and social distancing, as well as hand hygiene 0-5 times/day.

[e] N/A: not applicable.

[f] Among those who were employed.

[g] Among those who were employed and left home for work.

Discussion

Principal Findings

We report the first national probability survey estimates of the prevalence of COVID-19 mitigation strategies among US adults. During the 2020 peak of COVID-19 incidence, nearly three-quarters of adults consistently wore a mask when going out, about half consistently practiced social distancing or frequently washed their hands, and about a quarter frequently used a hand sanitizer. There were 3 distinct patterns of use of these mitigation practices. Two-thirds practiced optimal mitigation, with consistent and frequent use of all mitigation strategies; about 1 in 5 practiced the poorest mitigation practices, with inconsistent or infrequent use of all mitigation strategies; and about 1 in 9 consistently wore a mask and practiced social distancing and may have followed excessive hand hygiene practices. Finally, all mitigation practices and grouping of practices varied substantially among people with different demographic characteristics.

The prevalence of consistently wearing a mask in our population-based study was similar to earlier estimates from polls and convenience samples [5,9,10], but the estimate from our population-based sample was substantially lower than the 89% reported by the online convenience sampling–based COVID Impact Survey in June 2020 [12]. The difference in prevalence could be due to selection bias in the convenience sampling–based COVID Impact Survey if those who were more likely to wear masks were also more likely to respond to the survey. Our survey items were also slightly different, with our survey stipulating mask wearing when going out and asking about the frequency of mask usage, whereas the COVID Impact Survey asked a general question about mask wearing, without regard to context or frequency. It is also possible that mask wearing decreased between June and August 2020 (the beginning of our study), as public facilities reopened and mask requirements in each jurisdiction became more complex and possibly confusing.

There was a similar discrepancy between our findings and those of the COVID Impact Survey for practicing social distancing and washing hands, but for these practices, our estimates were even lower than those of the COVID Impact Survey, which reported >80% prevalence for both [12]. Although the COVID Impact Survey reported a slightly decreased prevalence of social distancing and hand hygiene in June compared to April 2020, a simple extrapolation of that decreasing trend would not explain the difference we found later in 2020. Similar selection biases and differences in survey items for these practices could explain part, but not all, of the difference between the findings of our study and the COVID Impact Survey. ConsumerStyles panel surveys that examined handwashing practices in October 2019 and June 2020 in specific contexts (eg, before eating, after sneezing, or after coughing) also found a substantially higher prevalence of handwashing than we did, providing further evidence that the survey question type (eg, making the questions conditional on situations in which handwashing is recommended even outside of COVID-19 times) can substantially affect prevalence estimates. We structured our questions based on the only published study on the effectiveness of hand hygiene for preventing seasonal coronavirus infection [22]. As we would expect during the COVID-19 pandemic with frequent communications about the importance of hand hygiene, our prevalence estimate of 57.7% who washed their hands 6-10 times in the past 24 hours was substantially higher than the 39.5% reported in the UK study conducted between 2006 and 2009. The prevalence of use of a hand sanitizer in our study was also substantially lower (21.5% vs 70.7%) compared to only 1 other previous paper, by Czeisler et al [14], that reported this as a separate behavior from handwashing. This difference in prevalence was likely due to context-specific differences in behaviors, where Czeisler et al [14] assessed hand sanitizer usage only after contact with high-touch public surfaces.

The distinct sets of mitigation practices (optimal mitigation, lowest mitigation, and optimal mitigation with additional hand hygiene) were also novel findings of our study. Those in the optimal-mitigation and optimal-mitigation-with-additional-hand-hygiene groups frequently wear masks and practice social distancing when they go out in public. Although there was a clear distinction between these groups based on the frequency of handwashing predetermined based on Beale et al’s [22] effectiveness study, there were no differences between these groups in their agreement with the mask-wearing and youth prevention questions. Our findings did indicate that the optimal-mitigation and optimal-mitigation-with-additional-hand-hygiene groups differed on multiple demographic characteristics, which supports the idea that these groups may have fundamental differences in their approaches toward COVID-19 prevention. Further study on the context of hand hygiene practices may clarify some of these issues, and we are now implementing a context-specific set of mitigation practice questions in our 3- and 6-month follow-up surveys with this cohort. The lowest-mitigation group, which was inconsistent in all mitigation practices, comprised an unfortunately large proportion of 1 in 5 US adults. The demographic differences between the optimal- and lowest-mitigation groups were even more pronounced, emphasizing the demographic disparities in COVID-19 mitigation practices.

This heterogeneity in COVID-19 mitigation practices among demographic groups in our study has also been partly reported in other published papers for individual practices [5,9,12,15,20], and those prior published findings are reasonably consistent with the demographic heterogeneity we found. Our study goes a step further to illustrate how persons from various backgrounds combine the individual mitigation recommendations in practice. The demographic heterogeneity in these empirically determined grouping of mitigation practices is even more evident than in individual practices. Compared to men and younger adults, women and older adults are much more likely to optimally use all mitigation practices. These differences in patterns of use may reflect greater risk perception, more exposure to COVID-19 prevention messages, or other contextual factors, such as leaving the home or living in group settings. US adults who were Black or Hispanic (compared to White, non-Hispanic), had no college degree, or worked in service-oriented jobs were more likely to report excessive hand hygiene, while also consistently wearing a mask and maintaining social distancing. These differences might also reflect greater risk perception and prevention message exposure but are more likely due to other contextual factors, such as hand hygiene requirements of their jobs. Finally, US adults who live outside metropolitan areas were likely to engage in all mitigation practices inconsistently. This might be due to differing risk perceptions or exposure to prevention messages in less densely populated areas [28].

Limitations

This study has several limitations. First, there was a lack of contextual information for some mitigation practices that could better clarify whether people are engaging in practices in only some settings or situations but not in others. These situational assessments have been added to our follow-up surveys, which were completed by mid-2021, and will be included in subsequent analyses. Second, some demographic heterogeneity could be explained by confounding, which could be elucidated with additional modeling. Multivariable modeling is planned for follow-up survey analyses. Third, we did not assess the quality of the mitigation behaviors, such as correctly wearing masks, or the effectiveness of those behaviors on preventing COVID-19 infection. The prospective component of the study will directly examine these associations. Fourth, the enrollment and baseline surveys occurred during a 5-month period of substantial changes in the COVID-19 pandemic and response. There may be time frame heterogeneity in the mitigation behaviors during these changes, but we were unable to analyze these baseline data as cross-sectional time series due to sampling method changes and prioritization of the entire survey sample weighting for national estimates [21]. Finally, although household probability sampling methods and weighting allowed for national estimation of these essential mitigation practices, there is likely still selection bias due to nonresponse.

Conclusion

Although the prevalence of consistently wearing a mask was relatively high among US adults, there were still millions who were not doing so during the time of the highest COVID-19 incidence to date in the pandemic. Even greater numbers of US adults did not consistently practice social distancing outside their homes and did not frequently practice hand hygiene. These practices remained crucial to blunting the surge of COVID-19 infections, especially since we had not yet achieved sufficient vaccine coverage to stop the pandemic. Despite clear public health evidence of their importance, the implementation of these practices was further undermined by a confusing array of local jurisdiction messages about mask requirements and restrictions on public gatherings. In future infectious disease outbreak responses, monitoring mitigation practices in a context of changing mandates and messages will help us refine communication strategies to increase the adoption and persistence of effective mitigation behaviors. This monitoring will also help ensure that disparities in mitigation practices do not widen further, leading to even greater disparities in infectious disease incidence and continuation of high-level community transmission.

Acknowledgments

The study was funded by a grant from the National Institutes of Health (3R01AI143875-02S1).

Conflicts of Interest

TS, PSS, and HB are members of the Editorial Board of JMIR Public Health and Surveillance. However, they had no involvement in the editorial decisions for this manuscript. The manuscript was reviewed and handled by an independent editor.

References

1. Coronavirus disease 2019 (COVID-19): situation report, 51. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/331475 [accessed 2023-07-06]
2. COVIDView: a weekly surveillance summary of U.S. COVID-19 activity. Centers for Disease Control and Prevention. 2020 Apr 6. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview.html [accessed 2023-07-06]
3. Trends in number of COVID-19 cases and deaths in the US reported to CDC. Centers for Disease Control and Prevention. 2020. URL: https://covid.cdc.gov/covid-data-tracker/#trends_dailytrendscases [accessed 2021-06-22]
4. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2021-06-22]
5. Haischer MH, Beilfuss R, Hart MR, Opielinski L, Wrucke D, Zirgaitis G, et al. Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic. PLoS One 2020 Oct 15;15(10):e0240785 [FREE Full text] [doi: 10.1371/journal.pone.0240785] [Medline: 33057375]
6. Rader B, White LF, Burns MR, Chen J, Brilliant J, Cohen J, et al. Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study. Lancet Digit Health 2021 Mar;3(3):e148-e157 [FREE Full text] [doi: 10.1016/S2589-7500(20)30293-4] [Medline: 33483277]
7. Katz J, Sanger-Katz M, Quealy K. A detailed map of who is wearing masks in the US. New York Times. 2020 Jul 17. URL: https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html [accessed 2023-07-06]
8. Brennan M. Americans' face mask usage varies greatly by demographics. Gallup. 2020 Jul 13. URL: https://news.gallup.com/poll/315590/americans-face-mask-usage-varies-greatly-demographics.aspx [accessed 2023-06-07]
9. Fisher KA, Barile JP, Guerin RJ, Vanden Esschert KL, Jeffers A, Tian LH, et al. Factors associated with cloth face covering use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):933-937 [FREE Full text] [doi: 10.15585/mmwr.mm6928e3] [Medline: 32673303]
10. Kramer S. More Americans say they are regularly wearing masks in stores and other businesses. Pew Research Center. 2020. URL: https://www.pewresearch.org/fact-tank/2020/08/27/more-americans-say-they-are-regularly-wearing-masks-in-stores-and-other-businesses/ [accessed 2021-06-22]
11. Ciancio A, Kämpfen F, Kohler IV, Bennett D, Bruine de Bruin W, Darling J, et al. Know your epidemic, know your response: early perceptions of COVID-19 and self-reported social distancing in the United States. PLoS One 2020 Sep 4;15(9):e0238341 [FREE Full text] [doi: 10.1371/journal.pone.0238341] [Medline: 32886671]
12. Hutchins HJ, Wolff B, Leeb R, Ko JY, Odom E, Willey J, et al. COVID-19 mitigation behaviors by age group - United States, April-June 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1584-1590 [FREE Full text] [doi: 10.15585/mmwr.mm6943e4] [Medline: 33119562]
13. Brown LG, Hoover ER, Barrett CE, Vanden Esschert KL, Collier SA, Garcia-Williams AG. Handwashing and disinfection precautions taken by U.S. adults to prevent coronavirus disease 2019, spring 2020. BMC Res Notes 2020 Dec 04;13(1):550 [FREE Full text] [doi: 10.1186/s13104-020-05398-3] [Medline: 33276806]
14. Czeisler M, Garcia-Williams AG, Molinari N, Gharpure R, Li Y, Barrett CE, et al. Demographic characteristics, experiences, and beliefs associated with hand hygiene among adults during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 16;69(41):1485-1491 [FREE Full text] [doi: 10.15585/mmwr.mm6941a3] [Medline: 33056951]
15. Haston JC, Miller GF, Berendes D, Andújar A, Marshall B, Cope J, et al. Characteristics associated with adults remembering to wash hands in multiple situations before and during the COVID-19 pandemic - United States, October 2019 and June 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 09;69(40):1443-1449 [FREE Full text] [doi: 10.15585/mmwr.mm6940a2] [Medline: 33031363]
16. COVID-19 racial and ethnic health disparities. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/racial-ethnic-disparities/index.html [accessed 2023-07-06]
17. Van Dyke ME, Mendoza MCB, Li W, Parker EM, Belay B, Davis EM, et al. MMWR Morb Mortal Wkly Rep 2021 Mar 19;70(11):382-388 [FREE Full text] [doi: 10.15585/mmwr.mm7011e1] [Medline: 33735165]

18. Romano SD, Blackstock AJ, Taylor EV, El Burai Felix S, Adjei S, Singleton C, et al. Trends in racial and ethnic disparities in COVID-19 hospitalizations, by region - United States, March-December 2020. MMWR Morb Mortal Wkly Rep 2021 Apr 16;70(15):560-565 [FREE Full text] [doi: 10.15585/mmwr.mm7015e2] [Medline: 33857068]
19. Pennington A, Kompaniyets L, Summers A, Danielson M, Goodman A, Chevinsky J, et al. Risk of clinical severity by age and race/ethnicity among adults hospitalized for COVID-19-United States, March-September 2020. Open Forum Infect Dis 2021 Feb;8(2):ofaa638 [FREE Full text] [doi: 10.1093/ofid/ofaa638] [Medline: 33553477]
20. Garnier R, Benetka JR, Kraemer J, Bansal S. Socioeconomic disparities in social distancing during the COVID-19 pandemic in the United States: observational study. J Med Internet Res 2021 Jan 22;23(1):e24591 [FREE Full text] [doi: 10.2196/24591] [Medline: 33351774]
21. Siegler AJ, Sullivan PS, Sanchez T, Lopman B, Fahimi M, Sailey C, et al. Protocol for a national probability survey using home specimen collection methods to assess prevalence and incidence of SARS-CoV-2 infection and antibody response. Ann Epidemiol 2020 Sep 10;49:50-60 [FREE Full text] [doi: 10.1016/j.annepidem.2020.07.015] [Medline: 32791199]
22. Beale S, Johnson AM, Zambon M, Watch Group F, Hayward AC, Fragaszy EB. Hand hygiene practices and the risk of human coronavirus infections in a UK community cohort. Wellcome Open Res 2020;5:98 [FREE Full text] [doi: 10.12688/wellcomeopenres.15796.2] [Medline: 34250260]
23. 2018 Standard occupational classification system. Bureau of Labor Statistics. 2018. URL: https://www.bls.gov/soc/2018/major_groups.htm [accessed 2021-08-18]
24. COVID-19: health and safety steps for specific occupations. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/coronavirus/2019-ncov/community/workplaces-businesses/specific-industries.html#food-services [accessed 2021-08-18]
25. Andridge RR, Little RJA. A review of hot deck imputation for survey non-response. Int Stat Rev 2010 Apr;78(1):40-64 [FREE Full text] [doi: 10.1111/j.1751-5823.2010.00103.x] [Medline: 21743766]
26. Lavrakas P. Raking. In: Encyclopedia of Survey Research Methods. Thousand Oaks, CA: SAGE Publications; 2008:672.
27. Berlin KS, Williams NA, Parra GR. An introduction to latent variable mixture modeling (part 1): overview and cross-sectional latent class and latent profile analyses. J Pediatr Psychol 2014 Mar;39(2):174-187. [doi: 10.1093/jpepsy/jst084] [Medline: 24277769]
28. Boyle J, Brassell T, Dayton J. Rural Americans feel less threatened by COVID-19 than urban and suburban Americans do, but still view mitigation as important. ICF International. 2020 Jul 22. URL: https://www.icf.com/insights/health/covid-19-survey-rural-vs-urban-threat [accessed 2021-10-11]

Abbreviations

LCA: latent class analysis
PR: prevalence ratio

Edited by A Mavragani; submitted 07.02.22; peer-reviewed by C Sun, M Lee, S Badawy; comments to author 25.03.22; revised version received 09.05.23; accepted 12.05.23; published 06.12.23.

Please cite as:
Sanchez T, Hall E, Siegler AJ, Prakash-Asrani R, Bradley H, Fahimi M, Lopman B, Luisi N, Nelson KN, Sailey C, Shioda K, Valentine-Graves M, Sullivan PS
Prevalence of COVID-19 Mitigation Behaviors in US Adults (August-December 2020): Nationwide Household Probability Survey
JMIR Public Health Surveill 2023;9:e37102
URL: https://publichealth.jmir.org/2023/1/e37102
doi:10.2196/37102
PMID:38055314

Review

Predisposing Factors of Nosocomial Infections in Hospitalized Patients in the United Kingdom: Systematic Review

Sharon Shivuli Isigi[1], MSc; Ali Davod Parsa[1], PhD; Ibrahim Alasqah[2,3], PhD; Ilias Mahmud[3,4], PhD; Russell Kabir[1], PhD

[1]School of Allied Health, Anglia Ruskin University, Essex, United Kingdom
[2]Department of Public Health, College of Public Health and Health Informatics, Qassim University, Al Bukairiyah, Saudi Arabia
[3]School of Health, University of New England, Armidale, Australia
[4]BRAC James P Grant School of Public Health, BRAC University, Dhaka, Bangladesh

Corresponding Author:
Ibrahim Alasqah, PhD
Department of Public Health
College of Public Health and Health Informatics
Qassim University
8848 King Abdulaziz Road
Al Bukairiyah, 52741
Saudi Arabia
Phone: 966 163020518
Email: i.alasqah@qu.edu.sa

Abstract

Background: Nosocomial infections are infections incubating or not present at the time of admission to a hospital and manifest 48 hours after hospital admission. The specific factors contributing to the risk of infection during hospitalization remain unclear, particularly for the hospitalized population of the United Kingdom.

Objective: The aim of this systematic literature review was to explore the risk factors of nosocomial infections in hospitalized adult patients in the United Kingdom.

Methods: A comprehensive keyword search was conducted through the PubMed, Medline, and EBSCO CINAHL Plus databases. The keywords included "risk factors" or "contributing factors" or "predisposing factors" or "cause" or "vulnerability factors" and "nosocomial infections" or "hospital-acquired infections" and "hospitalized patients" or "inpatients" or "patients" or "hospitalized." Additional articles were obtained through reference harvesting of selected articles. The search was limited to the United Kingdom with papers written in English, without limiting for age and gender to minimize bias. The above process retrieved 377 articles, which were further screened using inclusion and exclusion criteria following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The retained 9 studies were subjected to critical appraisal using the Critical Appraisal Skills Programme (cohort and case-control studies) and Appraisal Tool for Cross-Sectional Studies (cross-sectional studies) checklists. Finally, 6 eligible publications were identified and used to collect the study findings. A thematic analysis technique was used to analyze data extracted on risk factors of nosocomial infections in hospitalized patients in the United Kingdom.

Results: The risk factors for nosocomial infections that emerged from the reviewed studies included older age, intrahospital transfers, cross-infection, longer hospital stay, readmissions, prior colonization with opportunistic organisms, comorbidities, and prior intake of antibiotics and urinary catheters. Nosocomial infections were associated with more extended hospital stays, presenting with increased morbidity and mortality. Measures for controlling nosocomial infections included the use of single-patient rooms, well-equipped wards, prior screening of staff and patients, adequate sick leave for staff, improved swallowing techniques and nutritional intake for patients, improved oral hygiene, avoiding unnecessary indwelling plastics, use of suprapubic catheters, aseptic techniques during patient care, and prophylactic use.

Conclusions: There is a need for further studies to aid in implementing nosocomial infection prevention and control.

(JMIR Public Health Surveill 2023;9:e43743) doi:10.2196/43743

KEYWORDS

hospital-acquired infections; nosocomial infections; infection risk; systematic review; hospitalized patients; public health

Introduction

Nosocomial infections are recognized public health issues worldwide with a prevalence of 3.0%-20.7% and an incidence rate of 5%-10% [1]. These infections are not present or incubating at the time of admission to a hospital but are acquired after hospitalization and manifest 48 hours after admission; 30% of nosocomial infections are estimated to be preventable [2]. Nosocomial infections are a potential risk to the patients, staff, and community, with many of the microorganisms isolated known to be drug-resistant, resulting in increased admissions associated with a longer hospital stay, mortality, and increased hospital expenses [3]. A recent study [4] found that nosocomial infections account for 80% of all hospital infections, which include catheter-associated urinary tract infections (UTIs), surgical site infections (SSIs), ventilator-associated pneumonia, hospital-acquired pneumonia, and *Clostridium difficile* infections. Major symptoms of nosocomial infections are productive cough, abdominal pain, shortness of breath, altered mental status, rebound tenderness, palpitations, suprapubic pain, dysuria, polyuria, and costovertebral angle tenderness [2].

The risk factors associated with nosocomial infections are dependent on infection control procedures at the facility, the prevalence of pathogens in the community, and patients' immune status [1]. Research-based known risk factors include immunosuppression, older age, period of hospitalization, underlying comorbidities, frequency of hospital visits, mechanical ventilator support, invasive procedures, indwelling devices, longer stay in the intensive care unit (ICU), lack of infection control measures, and environmental hygiene [5]. The increased risk of nosocomial infections in older patients is due to frailty, suppressed immunity, and preexisting conditions [6].

Nosocomial outbreaks can be a cause of hospital closure, increased morbidity, and mortality. Therefore, to prevent and minimize nosocomial infections, several hospitals have implemented infection control procedures such as disinfection, steaming, proper waste management, isolation of affected staff until 48 hours after recovery, and aseptic hand washing before and after patient care, alongside prior identification of patients at risk of nosocomial infections and multidrug resistance [5].

Freeman and McGowan [6] found that risk factor identification allows for evaluating per-patient component risk, epidemiologic comparisons between hospital populations, and cost-effectiveness of infection prevention and control interventions. Calculation of risk per admission justifies findings on length of hospital stay as an increased risk for nosocomial infections [7]. This calculated per-day risk will account for the daily overall risk of infection per patient [6].

This systematic review was conducted to fill the gap in the absence of reviews on risk factors for nosocomial infections in hospitalized patients in the United Kingdom without limiting age and gender. The unique health care system in the United Kingdom funded by the National Health Service (NHS) is always at the frontline of adopting innovations and technologies in health care, which provides a guide to the international community on emerging innovations around infection control and prevention. Therefore, understanding predisposing factors of nosocomial infections in the United Kingdom can help inform local policies and practices, serving as a reference point for other countries considering revising their health care guidelines and policies.

The aims and objectives of this systematic review were focused on exploring the risk factors of nosocomial infections in hospitalized patients in the United Kingdom, the population at risk of nosocomial infections, impacts associated with nosocomial infections, and organisms associated with the spread of nosocomial infections.

Methods

Search Strategy

The search strategy was formulated based on the Population, Exposure, and Outcome framework, which is suitable for investigating the likelihood of developing a condition in the presence of a factor and providing practical strategies for controlling the outcome [8].

The PubMed and Google Scholar databases were searched for any existing or ongoing systematic reviews published between 2017 and 2021 regarding risk factors of nosocomial infections in hospitalized patients. Although systematic reviews were identified, none of these focused on the risk factors of nosocomial infections in hospitalized patients.

The review was focused on a comprehensive search using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for systematic reviews of 2020 [9]. Headings and keywords were used through a database search, including the PubMed, EBSCO CINAHL Plus, and EBSCO Medline databases. Although searching multiple databases identified duplicate studies, this approach reduced the chances of missing out on important publications.

Boolean operators "OR" and "AND" were utilized to translate the research question into research strings to retrieve relevant and focused results. Medical Subject Headings in databases and keywords from other research papers were used to identify synonyms of the keywords for obtaining more focused results. The keywords included "risk factors" OR "contributing factors" OR "predisposing factors" OR "cause" OR "vulnerability factors" AND "nosocomial infections" OR "hospital-acquired infections" AND "hospitalized patients" OR "inpatients" OR "patients" OR "hospitalized." Additional articles were obtained through reference harvesting of selected articles. The search was limited to the United Kingdom with papers in English, without limiting for age and gender to minimize bias.

Study Selection

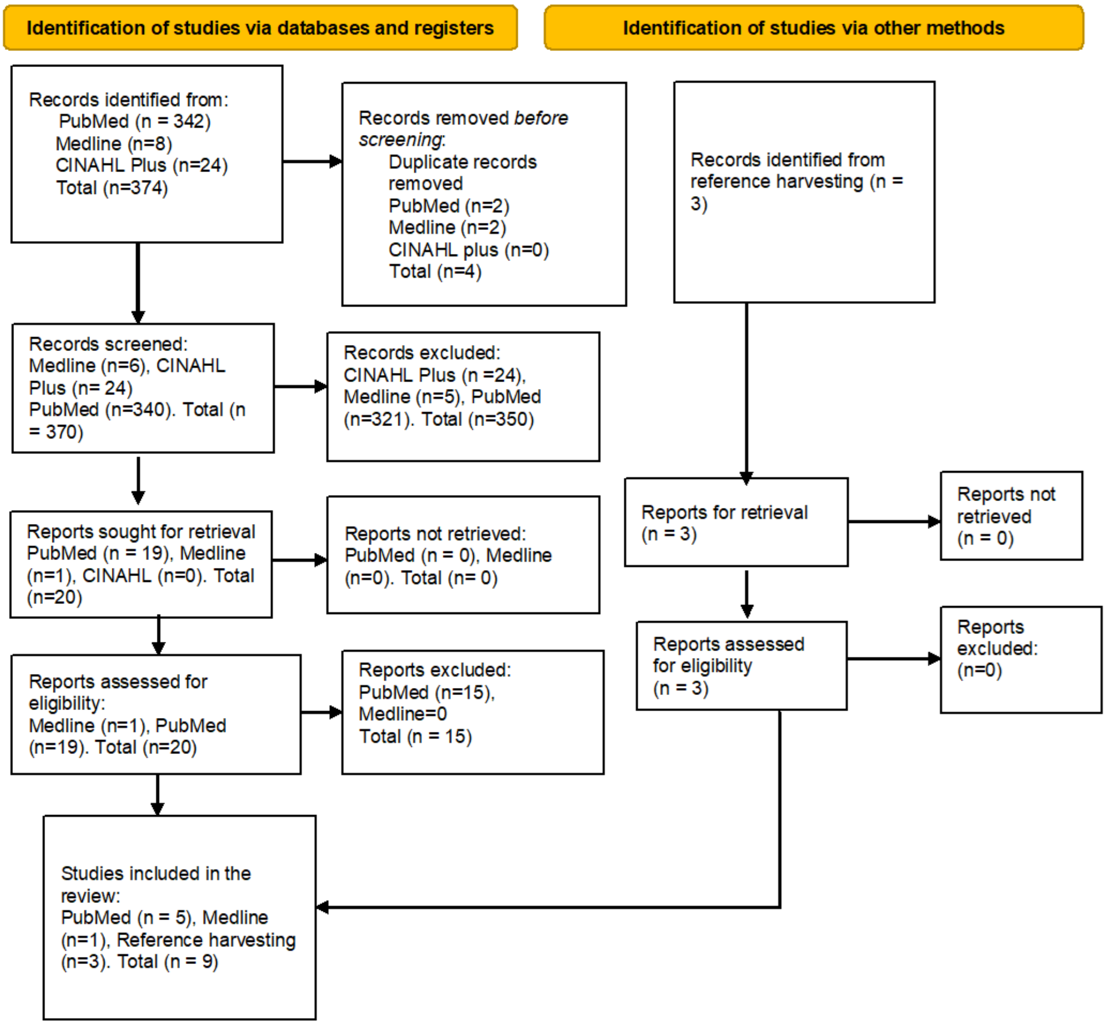
The inclusion and exclusion criteria for the systematic review are shown in Table 1.

The database search resulted in 374 articles plus 3 articles retrieved from reference harvesting (Figure 1).

Table 1. Inclusion and exclusion criteria.

Category	Inclusion criteria	Exclusion criteria
Study type	Primary research studies, 2011-2021 time frame, peer-reviewed articles, English language, full-text articles, qualitative and quantitative studies	Systematic reviews, non-English language, articles published before 2011, not peer-reviewed articles, abstract only
Population	Hospitalized patients/patients/inpatients in the United Kingdom	Hospitalized patients outside the United Kingdom
Exposure	All hospitals in the United Kingdom	Hospitals outside the United Kingdom
Outcome	Nosocomial infections in the United Kingdom	Any other articles that did not collect information on nosocomial infections in the United Kingdom

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for study selection.



RefWorks software was used to remove duplicate articles and titles were read manually to avoid duplication bias, resulting in 370 articles from all databases.

Title and abstract screening based on the inclusion and exclusion criteria resulted in 20 included articles and 350 excluded articles. The reasons for exclusion were that titles and abstracts were not giving relevant information on risk factors of nosocomial infections, guidelines, providing information on community-acquired infections in non-UK countries, or systematic reviews. Full-text article screening resulted in zero articles being excluded. In addition, 15 studies were excluded because of providing inadequate information on risk factors of nosocomial infections. After applying these criteria, nine articles, including the three obtained from reference harvesting, were subject to critical appraisal. A scoping review was carried out and no relevant papers were identified before 2011.

Data Extraction

The data extracted were tabulated using Microsoft Word. The data drawn out included authors' details, study design, the context of the study, sample size, aim of the study related to the risk of nosocomial infections, key findings, limitations, and conclusions of the identified studies.

Data Analysis

Thematic analysis was conducted on the included studies following the Braun and Clarke [10] guidelines, who describe a thematic analysis as a useful and flexible method for qualitative research. Thematic analysis classifies themes and patterns that connect to data, enabling researchers to identify and detect variables influencing any issues in study participants [10]. The themes were identified by one author and then agreed upon via discussion with all authors.

Critical Appraisal

Critical appraisal was applied to the selected 9 articles to check for the validity of the research if the results provided answers to the research question and the relevance of the information while minimizing bias, with the answers being either yes (+), no (–), or can't tell (+/–). Included case-control (Table 2) and cohort (Table 3) studies were appraised using the Critical Appraisal Skills Programme (CASP) tool and cross-sectional studies (Table 4) were appraised using the Appraisal Tool for Cross-Sectional Studies (AXIS) checklist. The CASP tool is a checklist that uses focused questions to help researchers evaluate the validity and reliability of identified research articles. This provides appraisal skills and tools that facilitate checking research resources for trustworthiness, relevance, and results [11]. The AXIS tool was used to evaluate the reporting quality, study design, and risk of bias in cross-sectional studies [12]. This tool focuses on presented methods and results of studies, which helps researchers assess whether the findings are reliable and credible and whether they relate to the aims, methods, and analysis of what is reported in identified studies. Following critical appraisal, three articles were excluded due to lack of ethical approval, low internal validity, or reliability of results, resulting in a total of six included articles.

Table 2. Critical appraisal of the selected case-control studies using the Critical Appraisal Skills Programme tool [11].

Critical appraisal questions	Boncea et al [13]	Alrawi et al [14]	Ousey et al [15]
Did the study address a clearly focused issue?	+[a]	+	+
Did the authors use an appropriate method to answer their question?	+/–[b]	+	+/–
Were the cases recruited acceptably?	+	+	+/–
Were the controls selected acceptably?	+	+	+/–
Was the exposure accurately measured to minimize bias?	+/–	+	+
Aside from the experimental intervention were the groups treated equally?	+	+/–	+/–
Have the authors taken into account the potential confounding factors in the design and/or in their analysis?	+/–	–[c]	+/–
How large was the treatment effect?	+	–	+/–
How precise was the estimate of the treatment effect?	+	–	+
Do you believe the results?	+/–	–	+/–
Can the results be applied to the local population?	–	+/–	+/–
Do the results of this study fit with other available evidence?	+	+	+/–

[a] +:Yes.

[b] +/–:Unclear.

[c] –:No.

Table 3. Critical appraisal of the selected cohort studies using the Critical Appraisal Skills Programme tool [11].

Critical appraisal questions	Mo et al [16]	Behar et al [17]	Ewan et al [18]	Melzer and Welch [19]	Wloch et al [20]
Did the study address a clearly focused issue?	+[a]	+	+	+	+
Was the cohort recruited acceptably?	+	+/–[b]	+/–	+	+
Was the exposure accurately measured to minimize bias?	+/–	+	+	+	+
Was the outcome accurately measured to minimize bias?	+/–	+	+/–	+	+/–
Have the authors identified all important confounding factors?	+	+	+	+	+
Have they taken into account the confounding factors in the design and/or analysis?	+	+	+	+	+/–
Was the follow-up of subjects complete enough?	+	+	+	+	+
Was the follow-up of subjects long enough?	+	+	+	+	+/–
What are the results of the study?	+	+	+	+	+
How precise are the results?	+	+	+	+	+
Did you believe the results?	+	+	+	+	+
Can the results be applied to the local population?	+	+	+/–	–[c]	+
Do the results of this study fit with other available evidence?	+	+	+/–	+	+
What are the implications of the study for practice?	+	+	+	+	+

[a] +:Yes.

[b] +/–:Unclear.

[c] –:No.

Table 4. Critical appraisal of the selected cross-sectional study using the Appraisal Tool for Cross-Sectional Studies [12].

Critical appraisal questions	Ledwoch et al [21]
Was the aim/objective of the study clear?	+[a]
Was the study design appropriate for the stated aim?	+
Was the sample size justified?	+
Was the reference population clearly defined? (Is it clear who the research was about?)	+
Was the sample frame taken from an appropriate population to closely represent the reference population under investigation?	+
Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	+
Were measures undertaken to address and categorize nonresponders?	–[b]
Were the risk factor and outcome variables measured appropriate to the aim of the study?	+
Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialed, piloted, or published previously?	+
Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	–
Were the basic data adequately described?	+
Does the response rate raise concerns about nonresponse bias?	+
If appropriate, was information about nonresponders described?	–
Were the results internally consistent?	+
Were the results presented for all the analyses described in the methods?	+
Were the authors' discussion and conclusions justified by the results?	+
Were the limitations of the study discussed?	+
Was there funding or conflicts of interest that may affect the authors' interpretation of results?	+
Was ethical approval or consent of participants attained?	–

[a] +:Yes.

[b] –:No.

Results

Characteristics of Included Studies

This systematic review identified 5 cohort studies [16-20], 3 case-control studies [13-15], and 1 cross-sectional study [21]. After critical appraisal, 5 cohort studies [16-20] and 1 case-control study [13] were included for analysis. The included studies covered countries within the United Kingdom, with the majority covering England. Most of the studies identified older adults and patients with preexisting conditions at increased risk of nosocomial infections. The organisms associated with increased nosocomial infections according to the majority of the included studies were *C. difficile*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Enterococcus* spp, *Klebsiella pneumoniae*, *Staphylococcus aureus*, coliform species, anaerobic cocci, methicillin-resistant *S. aureus* (MRSA), coagulase-negative *Staphylococcus*, Enterobacteriaceae, and streptococci (Table 5).

Table 5. Data extracted from the included studies.

Reference	Study design	Context	Sample size	Study aim	Key findings	Limitations	Conclusion
Alrawi et al [14]	Case-control study	United Kingdom	Adult patients admitted to a burn center over 12 months	Determine the presence of different microbes in a UK burn center, and examine the relationships between bacterial colonization, burn size, length of hospital stay, and delayed referral	*Staphylococcus aureus* was common with 79% positive patients, and a direct link was found between an increased incidence of bacteria colonization, delay in referral >24 hours, length of hospital stay, and large burn size	None	A burn wound creates a good medium for bacteria colonization and proliferation; in understanding the sources of bacteria and patients' susceptibility, wound care clinicians can form better management and treatment to reduce mortality and morbidity from burn wound sepsis
Behar et al [17]	Cohort study	England	727 patients admitted to an elderly medicine ward	Establish the risk factors for *Clostridium difficile* in hospitalized patients	9.8% of patients carried toxigenic *C. difficile* and ribotype 027 was not identified, while ribotype 106 was identified 3 times and in 7 others twice. Independent factors of colonization included previous *C. difficile* infection (OR[a] 4.53, 95% CI 15.48) and malnutrition (OR 3.29, 95% CI 1.47-7.35)	Overestimation of carriers because of a higher rate in sampled patients than in unsampled patients. No certainty whether the infection was community-acquired or hospital-acquired. The study may be prone to type 2 errors. Failure to calculate sample size due to lack of information on the frequency of colonization before the start of the study.	The study identifies the common carriage of *C. difficile* in older patients. Patients with previous *C.difficile* infections are to be evaluated as a cost-effective intervention to reduce symptomatic *C. difficile* infection in hospitalized patients.
Boncea et al [13]	Case-control study	United Kingdom	3 hospitals with 24,240 older patients (>65 years)	Explore how intrahospital transfer influences the odds of developing an HAI[b] in an urban hospital network	72.2% of the hospital transfer cases had at least one intrahospital transfer and each additional transfer increased the odds of acquiring a nosocomial infection by 9% in older patients (OR 1.09, 95% CI 1.05-1.13)	Lack of physiological data. Use of OPCS-4[c], which lacks hierarchy in the invasiveness of procedures and failure to record minor medical devices. Unavailable information on prescribed antibiotics and proton pump inhibitors may result in confounding effects. Lack of staffing level; unclear if infection spread is due to staff or patient casual movement. Missing symptom information. Findings not generalizable to the younger population. Lack of time stamp inaccuracies and diagnostic coding errors.	Intrahospital transfers increased the odds of developing nosocomial infections. Strategies for minimizing transfer should be considered. Further studies needed to identify unnecessary transfers.

Reference	Study design	Context	Sample size	Study aim	Key findings	Limitations	Conclusion
Ewan et al [18]	Cohort study	Northeast England	90 patients with lower limb fractures aged 65-101 years in a general hospital	Investigate the association between hospital-acquired pneumonia preceding heavy dental plaque and oral carriage of respiratory pathogens in older patients with limb fractures to determine the target for intervention studies	Hospital-acquired pneumonia was not associated with dentate, tooth number, or heavy dental plaques, but was associated with increased hospital stay and prior oral carriage of *Escherichia coli/S. aureus/Pseudomonas aeruginosa*/MRSA[d] (OR 9.48, 95% CI 2.28-38.78; *P*=.002)	Results are not generalizable to medical patients. Biased sample toward "well patients," leading to underestimation of exposure and outcome variables. Overestimated hospital-acquired pneumonia incidence. Overprediction of significance when the colonization index was larger than zero.	Lower limb fracture patients colonized with *E. coli*, *S. aureus*, MRSA, or *P. aeruginosa* after 5 days in hospital were at greater risk of hospital-acquired pneumonia (*P*=.002). Methods to implement and deliver good oral hygiene at the ward level in the health system need to be investigated to minimize the use of antibiotics and length of hospital stay.
Ledwoch et al [21]	Cross-sectional study	Wales, Scotland, England	3 hospitals and 1 dental practice	Investigate the presence of DSBs[e] on 52 routinely cleaned keyboards from 4 hospitals across the United Kingdom	Hospital keyboards outside the patient area harbor DSBs, representing potential reservoirs for transferrable pathogens with 31% vancomycin-resistant *Enterococcus*, 17% MDR[f]*Acinetobacter* spp, and 72% MRSA recovered from almost half of the 45 samples	Selective plates and quality control were not selective to a single species from the DSB	Hospital keyboards are a source of infection for HAIs with a need for further studies on DSBs to find products that can control DSBs while effectively preventing bacteria transfer
Melzer and Welch [19]	Cohort study	London	500 patients in Royal London Hospital	To determine host factors that can predict severe sepsis in a bacteremic cohort	Bacteremia largely occurred in patients aged over 50 years (64.3%) and in men (58.2%); common isolates were *E. coli* (34.8%) and MRSA (9.6%). In the multivariable logistic regression, the site of infection was associated with severe sepsis and catheter-associated UTIs[g], significant after adjustment of age, sex, Charlson comorbidities index, and where the infection was acquired	Findings are not generalizable to other severe sepsis cohorts with bacteremia	Urinary catheter increases the risk of severe sepsis; if clinically indicated, should be used and removal dates indicated unless long-term use is required

Reference	Study design	Context	Sample size	Study aim	Key findings	Limitations	Conclusion
Mo et al [16]	Cohort study	Oxfordshire, United Kingdom	4 teaching hospitals	Quantify the importance of different transmission pathways of SARS-CoV-2 in a hospital setting	For susceptible patients, 1 day in the same ward with another patient with nosocomial SARS-CoV-2 infection was associated with 7.5 infections per 1000 susceptible patients per day, while exposure to a patient with community-acquired SARS-CoV-2 infection or to an infectious health care worker had a lower risk of 2 per 1000 susceptible patients	No clarity during phase 3 onward when weekly screening was implemented. Recall bias from staff self-reporting. Staff were assumed to be absent after the first positive PCR[h] result. Genomic sequencing of SARS-CoV-2 is not considered to confirm the transmission pathway.	Exposure to patients with nosocomial SARS-CoV-2 infection poses a risk to health care workers and hospitalized patients. Further investigation to enhance infection control and prevention around these patients is required.
Ousey et al [15]	Case- control study	England	206 patients with emergency cesarean section in an acute district general hospital	Assess the incidence risk and factors contributing to an acquired SSI[i] after an emergency cesarean section	BMI was associated with SSI (OR 1.17, 95% CI 1.11-1.24; *P*<.001). Nonsignificant links were identified between SSIs and age and vaginal swab status.	Stigmatization of pregnant women. Use of wound swabs leading to false-positive or false-negative results. Missing or inaccurate recordings. Small sample size. Use of BMI to calculate body composition when its accuracy has been questioned.	BMI was the main risk factor for SSI because of excessive adipose tissue. Diabetes status, age, and preoperative vaginal swab status were not associated with SSI.

[a] OR: odds ratio.

[b] HAI: hospital-acquired infection.

[c] OPCS-4: Office of Population Census and Surveys.

[d] MRSA: methicillin-resistant *Staphylococcus aureus*.

[e] DSB: dry surface biofilm.

[f] MDR: multidrug resistant.

[g] UTI: urinary tract infection.

[h] PCR: polymerase chain reaction.

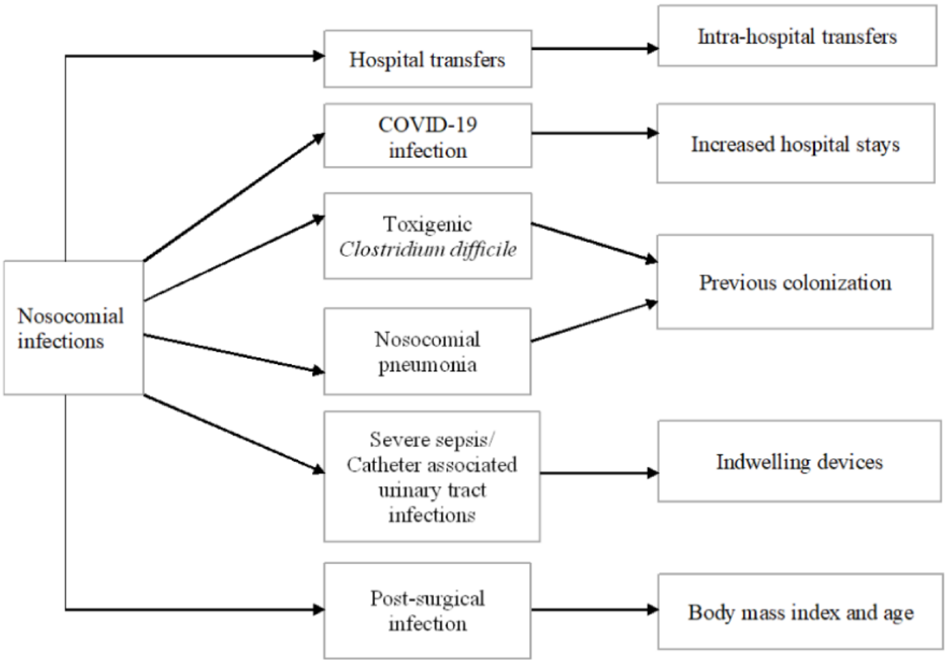
[i] SSI: surgical site infection.

Themes

Overview

The themes identified included intrahospital transfers, increased hospital stay, previous colonization, indwelling devices, and BMI and age, as shown in Figure 2.

Figure 2. Identified themes.



Theme 1: Intrahospital Transfers

According to Boncea et al [13], older patients acquired nosocomial infections associated with ICU admissions and in-hospital deaths (*P*<.001) more frequently than younger patients 48 hours after admission (median 79, IQR 73-86 vs median 79, IQR 72-85; *P*=.004) and also had higher comorbidities (mean 4.0, SD 2.0 vs mean 3.5, SD 1.9; difference=0.5; *P*<.001). The ICU admissions cases (6.7%) and in-hospital deaths (13.3%) were significantly higher (*P*<.001) among patients 48 hours after admission. The findings demonstrated that an additional intrahospital transfer was associated with a 9% increase in the chances of developing a nosocomial infection (odds ratio [OR] 1.09, 95% CI 1.05-1.13).

Overall, 11.9% of the patients developed a nosocomial infection (49.6% men and 50.4% women) with a median age of 79 (IQR 72-86) years and a mean of 3.5 (SD 1.9) Elixhauser comorbidities, an index used to predict hospital length of stay, hospital charges, and in-hospital mortality.

In addition, 76.0% of patients with a positive lab culture collected at least 48 hours after hospitalization experienced at least one transfer compared to 71.7% of controls (patients who remained infection-free throughout the admission), with cardiology patients experiencing more movement (median 2, IQR 1-2). Logistic regression analysis was used to rule out the influence of ethnicity, gender, weekend admissions, and readmissions within 30 days. The most commonly isolated organisms from the patient group were *C. difficile*, *E. coli*, *P. aeruginosa*, *Enterococcus* species, *K. pneumoniae*, *S. aureus*, coliform species, MRSA, and coagulase-negative *Staphylococcus*.

Theme 2: Increased Hospital Stay

Mo et al [16] observed that an increased hospital stay increases the risk of nosocomial infections among health care workers and patients. The evidence indicated that newly admitted COVID-19 patients were associated with a risk of onward transmission to other patients in a hospital setting.

Approximately 62.1% of patients had positive SARS-CoV-2 results after admission and most of them were older. These older patients had more extended hospital stays and more readmissions. Susceptible patients were at a higher risk of acquiring COVID-19 from a patient with nosocomial COVID-19 (adjusted OR 1.76, 95% CI 1.51-2.04) than from a patient with community-acquired COVID-19 (adjusted OR 1.12, 95% CI 0.96- 1.26) or from a health care worker with COVID-19 (OR 1.45, 95% CI 1.22-1.71). These results indicated that the risk of patients transmitting COVID-19 to other patients within the same facility was higher as a result of increased hospital stay. Other infections can also be transmitted because of interacting with the same health care workers and patients for prolonged periods. Mo et al [16] observed that exposure to nosocomial infections was linked with a substantial infection risk to other hospitalized patients.

These findings are consistent with those of Ewan et al [18], who found that increased hospital stay was significantly associated with nosocomial infections (Fisher exact test *P*=.001). Similarly, Boncea et al [13] found that increased length of hospital stay was associated with nosocomial infections even when the researchers controlled for severity of illness.

Theme 3: Previous Colonization

According to Behar et al [17], risk factors for toxigenic *C. difficile* colonization among patients were a previous diagnosis of symptomatic *C. difficile* infection (OR 8.48, 95% CI 2.77-25.96), hospital stay of 3 months, and frailty as assessed by the Malnutrition Universal Screening Tool (MUST) and Barthel score. Multivariable analysis proved colonization to be associated with previous *C. difficile* infection (OR 4.53, 95% CI 1.33-15.48) and a MUST score indicating a higher level of frailty (≥2) (OR 3.22, 95% CI 1.47-7.06); the association with a recent hospital stay was borderline significant (OR 2.18, 95% CI 0.99-4.78).

Overall, 9.8% of the patients were carriers of *C. difficile* after their first stool test presenting with diarrhea. The most isolated ribotypes were 126 (×3), (009, 018, 020, 023, 028, 038, and 039) ×2, and ribotype 027 strains were undetected.

Ewan et al [18] found that prior carriage with opportunistic *S. aureus*, *E. coli*, *P. aeruginosa*, and MRSA (OR 9.41, 95% CI 2.28-38.78; *P*=.002) was significantly associated with nosocomial pneumonia with relative risk of 6.44 (95% CI 2.04-20.34; *P*=.002). A combination of pathogens was identified, and the highest colonization rate was identified with *Streptococcus pneumoniae* (n=27 patients) due to persistent carriage, followed by *Hemophilus influenzae* (n=8), *S. aureus* (n=6), *P. aeruginosa* (n=4), MRSA (n=3), and *E. coli* and *Acinetobacter* spp (n=2). Ewan et al [18] also identified an association between nosocomial pneumonia with admission from the hospital, active cancer, witnessed aspiration episode, increased frailty, antibiotic intake before admission, and reduced mobility.

Theme 4: Indwelling Devices

Melzer and Welch [19] sought to establish whether the presence of a urinary catheter predicts severe sepsis as a nosocomial infection within a bacteremic cohort. The researchers identified 64.3% bacteremic episodes in patients >50 years old, 58.2% of whom were men. The episodes were community-acquired, health care–associated, and hospital-acquired infections, but community-acquired episodes were not associated with severe sepsis (Pitt score≥2). Site of infection, individual patient, critical care admission, and catheter-associated UTI were significantly associated with severe sepsis. After multivariate analysis, the site of infection was significantly associated with severe sepsis, with a stronger association for catheter-associated UTIs (OR 3.87, 95% CI 1.82-8.22), which remained significant after adjustment for age and where the infection was acquired (OR 3.94, 95% CI 1.70-9.11). These results indicate that indwelling devices such as a urinary catheter are a risk factor that can predict nosocomial infections such as severe sepsis.

Theme 5: BMI and Age

Wloch et al [20] found that 10 days following a hospital admission was the median time for all SSIs and 8 days of admission was the median time for deep and organ infections alone. The mentioned risk factors for these nosocomial infections included BMI, which was strongly associated with superficial infections (*P*<.001), and deep organ space (*P*<.003). After multivariate analysis, women with a BMI of 25-30 (overweight) had an estimated OR of 1.6 (95% CI 1.2-2.2) compared to that of women within a normal BMI range. The results also indicated that women with a BMI >30 (obese) had a 2.4-times greater chance of nosocomial infection (95% CI 1.7-3.4).

After multivariate analysis, age had a stronger association for SSIs and women <20 years of age (*P*=.04) had a weaker association for SSIs. *S. aureus* was the commonly associated organism at 40.4%; 17.7% were methicillin-resistant organisms and the rest included anaerobic cocci at 23.2%, Enterobacteriaceae at 13.3%, and streptococci at 7.4%.

Boncea et al [13] observed that increased age was a risk factor for nosocomial infections in a study focused on hospital-acquired infections in an urban UK hospital network. Patients with hospital-acquired infections were older than those who did not acquire the infections (median 79, IQR 73-86 vs median 79, IQR 72-85; *P*<.001). Boncea et al [13] also found that older patients had a higher number of comorbidities than younger patients (mean 4.0, SD 2.0 vs mean 3.5, SD 1.9; difference=0.5; *P*<.001).

Discussion

Principal Findings

The findings of this systematic review provided answers to the study objectives. The main risk factors of nosocomial infections identified in this study include intrahospital transfers, increased hospital stay, previous colonization, indwelling devices, and BMI and age.

Blay et al [22] demonstrated that intrahospital transfers increased the risk of developing nosocomial infections, with the majority of patients in the study undergoing 2.4 transfers per year. This finding has been supported by Boncea et al [13], who reported that an additional intrahospital transfer was associated with a 9% higher risk of acquiring infections in a hospital. Boncea et al [13] also found that older patients were more likely to develop nosocomial infections. The prevalence of nosocomial infections was high in the older population due to frailty and immunosuppression. Intrahospital transfers occurred as a result of lesser equipped wards along with procedure rooms and bed shortages that expose susceptible patients to infectious hospital surfaces, staff, and other patients, thus increasing their risk of nosocomial infection. The findings suggested that minimizing infection risk to be a priority when transferring patients. Kulshrestha and Singh [23] mentioned that patient intratransfers should aim at maintaining patient care to minimize morbidity and mortality. Single-patient rooms can be used to lower pathogen transmission and the number of intrahospital transfers, thereby reducing interactions between staff, hospital surfaces, and patients to eventually reduce the incidence of nosocomial infections [13].

Mo et al [16] observed that COVID-19 transmission between patients and staff had a stronger association for nosocomial COVID-19 than community-acquired COVID-19 [16]. The daily risk of COVID-19 was associated with increased exposure to newly infected patients and staff. Medical professionals can address this challenge by providing adequate sick pay leave for symptomatic staff for better complete recovery and regular screening for early detection to reduce cross-infection and control for the incidence of nosocomial COVID-19. Furthermore, early symptom presentation and regular screening of staff and patients is a better strategy for infection prevention and control, where infected patients and staff should be isolated immediately.

According to Behar et al [17], older patients were largely colonized by *C. difficile* presenting with diarrhea and contributing to transmission to other patients. The risk of colonization with asymptomatic *C. difficile* was due to frailty associated with suppressed immunity, which resulted in increased mortality. The risk of colonization was also associated with previous diagnoses and longer hospital stays. Early detection and isolation upon admission can reduce the further spread of nosocomial *C. difficile*. In a study by Dubberke et al [24], patients showed acquisition of *C. difficile* while hospitalized. The initial intake of antibiotics was associated with disease, whereas susceptibility to colonization with *C. difficile* was associated with suppressed immunity and an altered intestinal flora composition, gender, preexisting comorbidities, and frailty-determined mortality rate [17]. Risks such as previous hospital admission, antibiotic therapy, and immune suppression were well documented and explained in a similar study on *C. difficile* infection in hospitalized children in the United States [25].

Nosocomial pneumonia occurred after admission, with most samples positive for colonization with *S. aureus*, MRSA, *P. aeruginosa*, and *E. coli*, which resulted in extended hospital stays of 30 days, similar to patients with ventilator-associated pneumonia and aspiration pneumonia in medical patients [18]. Colonization was identified as the key risk factor for nosocomial pneumonia.

The results of this systematic review indicate that detection using molecular techniques resulted in a high incidence rate of nosocomial pneumonia in the exposed patient group. Ewan et al [13] failed to identify *K. pneumoniae* and Enterobacteriaceae as prevalent organisms and a risk of nosocomial pneumonia in frailer and unwell patients [26]. Oral colonization began within 72 hours of admission in a majority of the cases, with the possible explanation that some have been acquired from the community or upon admission; therefore, interventions should start within this timeframe. Delivery of better oral hygiene at the ward level in a resource-scarce health system can control for antibiotic use and length of hospital stay associated with nosocomial pneumonia.

Melzer and Welch [19] identified the presence of urinary catheters to be strongly associated with severe sepsis, with a stronger association for gram-negative multidrug-resistant *E. coli*. Furthermore, underlying diseases and obstruction history are also identifiable risks for severe sepsis. *E. coli*, a gram-negative bacterium gastrointestinal colonizer, is known to commonly cause UTIs and septicemia [26], which was also associated with delayed administration of antibiotics. To account for this, patients exhibiting any signs of sepsis should be given immediate admission or management in critical care areas, and implementation of a critical care bundle approach would result in reduced admissions due to severe sepsis. A critical care bundle would aid in the documentation of reasons for catheter insertion, removal dates, and aseptic techniques during and after insertion to control for infection. Medical professionals can increase patient satisfaction and lower infection risks by issuing catheter passports to long-term users, guiding community nurses on the use of prophylactic antibiotics when changing catheters, and using suprapubic catheters.

SSIs were strongly associated with an inpatient stay, readmissions, diabetes, surgeons' grade, and women <20 years and >45 years of age. The most commonly isolated organisms from the skin or genital flora were *S. aureus* (40.4%, 17.1% of which were MRSA), anaerobic cocci (23.2%), Enterobacteriaceae (13.3%), and streptococci (7.4%). BMI was strongly associated with SSIs due to poor prescription techniques that exclude patients' BMI, longer incision time, and slow transit of antibiotics due to thick adipose tissue that result in impaired wound healing and decreased immunity. This view on BMI has been supported by Ousley et al [15]. According to Chu et al [27], younger age in women is associated with an increased risk for SSIs due to underdeveloped immunity to fight infections. Older age is associated with frailty and a suppressed immune system, which has been supported by Boncea et al [13], who also mentioned readmission and increased hospital stay resulting in exposure of patients to hospital surfaces with pathogenic organisms, thereby increasing the risk for nosocomial infections. Another study confirmed diabetes to be a risk of postsurgical infections, which is a marker for other conditions such as vascular changes and white blood cell malfunction, resulting in suppressed immunity that exposes a patient to multiple infections during and after surgery [28]. In general, the analyzed studies identified impacts of nosocomial infections to be associated with increased admissions, increased hospital stay, and costs on health care services. Other studies have reached similar findings and identified that older age, readmissions, and longer hospital stay are associated with increased infection risk due to exposure to infectious pathogens, frailty, and suppressed immunity. Özdemir and Dizbay [29] found that older patients have a higher risk of nosocomial infections resulting from impaired immune defense, long-term hospital stay, immunosuppressive treatments, and underlying chronic illnesses.

Study Strengths and Weaknesses

There were no systematic reviews on risk factors of nosocomial infections in hospitalized patients, and the study identified both qualitative and quantitative data relevant to the research topic. As another strength, this review pointed out key risk factors associated with nosocomial infections and the patient populations at risk, additionally suggesting measures for infection control and prevention, providing an advantage to the NHS whose priority is controlling health care–associated infections [30]. In this review, non-English articles were excluded, which may have limited capturing information in other languages.

Conclusions

Patients with previous colonization and underlying conditions are at increased risk of nosocomial infections due to frailty, extended hospital stay, and suppressed immunity. The use of indwelling devices, increased age, high BMI, and intrahospital transfers increase the susceptibility of patients to nosocomial infections.

Intrahospital transfers expose patients to other infectious patients, staff, and hospital environments, thereby increasing their risk for infection. In addition, the number of staff per patient has been positively associated with an increased risk of infection. Readmissions, underlying conditions and/or active cancer, prior antibiotic intake associated with drug-resistant microorganisms, and colonization with opportunistic organisms contribute to nosocomial infections. Indwelling devices such as urinary catheters increase the risk of sepsis, while obesity and age are strongly associated with SSIs. The most common organisms associated with nosocomial infections included *C. difficile*, *E. coli*, *P. aeruginosa*, *Enterococcus* spp, *K. pneumoniae*, *S. aureus*, coliform species, anaerobic cocci, MRSA, coagulase-negative *Staphylococcus*, Enterobacteriaceae, and streptococci.

Implemented interventions should consider the time frame before and after admission to control for organisms acquired during both periods. Further studies are needed to identify the risk factors of nosocomial infections and drug-resistant pathogens targeting a larger patient group for better infection prevention and control.

Acknowledgments

The researchers would like to thank the Deanship of Scientific Research, Qassim University, for funding the publication of this project.

Authors' Contributions

Conceptualization: SSI, ADP, IM, IA, and RK; methodology: SSI, ADP, IA, IM, and RK; software: SSI; validation: IA, ADP, and RK; formal analysis: SSI, RK, ADP, IA, and IM; investigation: SSI, ADP, and RK; data curation: SSI; writing—original draft preparation: SSI; writing—review and editing: ADP, IM, IA, and RK; visualization: SSI; supervision: ADP, IM, and RK; project administration: SSI, ADP, and RK; funding acquisition: IA. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None.

Multimedia Appendix 1
PRISMA 2020 Checklist.
[DOCX File , 69 KB - publichealth_v9i1e43743_app1.docx]

References

1. Samuel S, Kayode O, Musa O, Nwigwe G, Aboderin A, Salami T, et al. Nosocomial infections and the challenges of control in developing countries. Af J Clin Exp Micro 2010 Apr 20;11(2):102-110. [doi: 10.4314/ajcem.v11i2.53916]
2. Lobdell KW, Stamou S, Sanchez JA. Hospital-acquired infections. Surg Clin North Am 2012 Feb;92(1):65-77. [doi: 10.1016/j.suc.2011.11.003] [Medline: 22269261]
3. Kaye KS, Marchaim D, Chen T, Baures T, Anderson DJ, Choi Y, et al. Effect of nosocomial bloodstream infections on mortality, length of stay, and hospital costs in older adults. J Am Geriatr Soc 2014 Feb;62(2):306-311 [FREE Full text] [doi: 10.1111/jgs.12634] [Medline: 24438554]
4. Llewellyn C, Ayers S, McManus C, Newman S, Petrie K, Revenson T. Cambridge Handbook of Psychology, Health and Medicine, 3rd edition. Cambridge, UK: Cambridge University Press; 2019.
5. Sadique Z, Lopman B, Cooper BS, Edmunds WJ. Cost-effectiveness of ward closure to control outbreaks of norovirus infection in United Kingdom National Health Service hospitals. J Infect Dis 2016 Feb 01;213(Suppl 1):S19-S26. [doi: 10.1093/infdis/jiv410] [Medline: 26744428]
6. Freeman J, McGowan JE. Risk factors for nosocomial infection. J Infect Dis 1978 Dec;138(6):811-819. [doi: 10.1093/infdis/138.6.811] [Medline: 105056]
7. Forster AJ, Taljaard M, Oake N, Wilson K, Roth V, van Walraven C. The effect of hospital-acquired infection with Clostridium difficile on length of stay in hospital. CMAJ 2012 Jan 10;184(1):37-42 [FREE Full text] [doi: 10.1503/cmaj.110543] [Medline: 22143235]
8. Sheldon L. Chapter 42: Exposure framework. In: Krieger R, editor. Hayes' handbook of pesticide toxicology, 3rd edition. New York: Academic Press; 2010:971-976.
9. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
10. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
11. Long HA, French DP, Brooks JM. Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. Res Meth Med Health Sci 2020 Aug 06;1(1):31-42. [doi: 10.1177/2632084320947559]

12. Ma L, Wang Y, Yang Z, Huang D, Weng H, Zeng X. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? Mil Med Res 2020 Feb 29;7(1):7 [FREE Full text] [doi: 10.1186/s40779-020-00238-8] [Medline: 32111253]
13. Boncea EE, Expert P, Honeyford K, Kinderlerer A, Mitchell C, Cooke GS, et al. Association between intrahospital transfer and hospital-acquired infection in the elderly: a retrospective case-control study in a UK hospital network. BMJ Qual Saf 2021 Jun;30(6):457-466 [FREE Full text] [doi: 10.1136/bmjqs-2020-012124] [Medline: 33495288]
14. Alrawi M, Crowley TP, Pape SA. Bacterial colonisation of the burn wound: a UK experience. J Wound Care 2014 May;23(5):274-277. [doi: 10.12968/jowc.2014.23.5.274] [Medline: 24810312]
15. Ousey K, Blackburn J, Stephenson J, Southern T. Incidence and risk factors for surgical site infection following emergency cesarean section: a retrospective case-control study. Adv Skin Wound Care 2021 Sep 01;34(9):482-487. [doi: 10.1097/01.ASW.0000767368.20398.14] [Medline: 34415252]
16. Mo Y, Eyre DW, Lumley SF, Walker TM, Shaw RH, O'Donnell D, Oxford COVID Infection Review Team, et al. Transmission of community- and hospital-acquired SARS-CoV-2 in hospital settings in the UK: a cohort study. PLoS Med 2021 Oct;18(10):e1003816 [FREE Full text] [doi: 10.1371/journal.pmed.1003816] [Medline: 34637439]
17. Behar L, Chadwick D, Dunne A, Jones CI, Proctor C, Rajkumar C, et al. Toxigenic Clostridium difficile colonization among hospitalised adults; risk factors and impact on survival. J Infect 2017 Jul;75(1):20-25 [FREE Full text] [doi: 10.1016/j.jinf.2017.04.006] [Medline: 28435086]
18. Ewan VC, Sails AD, Walls AWG, Rushton S, Newton JL. Dental and microbiological risk factors for hospital-acquired pneumonia in non-ventilated older patients. PLoS One 2015 Apr 29;10(4):e0123622 [FREE Full text] [doi: 10.1371/journal.pone.0123622] [Medline: 25923662]
19. Melzer M, Welch C. Does the presence of a urinary catheter predict severe sepsis in a bacteraemic cohort? J Hosp Infect 2017 Apr;95(4):376-382 [FREE Full text] [doi: 10.1016/j.jhin.2017.01.003] [Medline: 28202189]
20. Wloch C, Wilson J, Lamagni T, Harrington P, Charlett A, Sheridan E. Risk factors for surgical site infection following caesarean section in England: results from a multicentre cohort study. BJOG 2012 Oct;119(11):1324-1333. [doi: 10.1111/j.1471-0528.2012.03452.x] [Medline: 22857605]
21. Ledwoch K, Dancer SJ, Otter JA, Kerr K, Roposte D, Maillard J. How dirty is your QWERTY? The risk of healthcare pathogen transmission from computer keyboards. J Hosp Infect 2021 Jun;112:31-36. [doi: 10.1016/j.jhin.2021.02.021] [Medline: 33647376]
22. Blay N, Roche MA, Duffield C, Gallagher R. Intrahospital transfers and the impact on nursing workload. J Clin Nurs 2017 Dec;26(23-24):4822-4829. [doi: 10.1111/jocn.13838] [Medline: 28382638]
23. Kulshrestha A, Singh J. Inter-hospital and intra-hospital patient transfer: recent concepts. Indian J Anaesth 2016 Jul;60(7):451-457 [FREE Full text] [doi: 10.4103/0019-5049.186012] [Medline: 27512159]
24. Dubberke ER, Reske KA, Seiler S, Hink T, Kwon JH, Burnham CD. Risk factors for acquisition and loss of Clostridium difficile colonization in hospitalized patients. Antimicrob Agents Chemother 2015 Aug;59(8):4533-4543 [FREE Full text] [doi: 10.1128/AAC.00642-15] [Medline: 25987626]
25. Nylund CM, Goudie A, Garza JM, Fairbrother G, Cohen MB. Clostridium difficile infection in hospitalized children in the United States. Arch Pediatr Adolesc Med 2011 May;165(5):451-457 [FREE Full text] [doi: 10.1001/archpediatrics.2010.282] [Medline: 21199971]
26. Khan HA, Ahmad A, Mehboob R. Nosocomial infections and their control strategies. Asian Pacific J Trop Biomed 2015 Jul;5(7):509-514. [doi: 10.1016/j.apjtb.2015.05.001] [Medline: 38058587]
27. Chu K, Maine R, Trelles M. Cesarean section surgical site infections in sub-Saharan Africa: a multi-country study from Medecins Sans Frontieres. World J Surg 2015 Feb;39(2):350-355 [FREE Full text] [doi: 10.1007/s00268-014-2840-4] [Medline: 25358418]
28. Martin ET, Kaye KS, Knott C, Nguyen H, Santarossa M, Evans R, et al. Diabetes and risk of surgical site infection: a systematic review and meta-analysis. Infect Control Hosp Epidemiol 2016 Jan;37(1):88-99 [FREE Full text] [doi: 10.1017/ice.2015.249] [Medline: 26503187]
29. Özdemir K, Dizbay M. Nosocomial infection and risk factors in elderly patients in intensive care units. J Microbiol Infect Dis 2015;5(1):38-43. [doi: 10.5799/ahinjs.02.2015.01.0174]
30. Duerden B. Controlling healthcare-associated infections in the NHS. Clin Med 2008 Apr;8(2):140-143 [FREE Full text] [doi: 10.7861/clinmedicine.8-2-140] [Medline: 18478854]

Abbreviations

AXIS: Appraisal Tool for Cross-Sectional Studies
CASP: Critical Appraisal Skills Programme
ICU: intensive care unit
MRSA: methicillin-resistant Staphylococcus aureus
MUST: Malnutrition Universal Screening Tool
NHS: National Health Service

OR: odds ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SSI: surgical site infection
UTI: urinary tract infection

Edited by A Mavragani, T Sanchez; submitted 23.10.22; peer-reviewed by R Mirzaei, Z Xu; comments to author 12.06.23; revised version received 04.09.23; accepted 28.11.23; published 19.12.23.

Please cite as:
Isigi SS, Parsa AD, Alasqah I, Mahmud I, Kabir R
Predisposing Factors of Nosocomial Infections in Hospitalized Patients in the United Kingdom: Systematic Review
JMIR Public Health Surveill 2023;9:e43743
URL: https://publichealth.jmir.org/2023/1/e43743
doi:10.2196/43743
PMID:38113098

Original Paper

Syndromic Surveillance Using Structured Telehealth Data: Case Study of the First Wave of COVID-19 in Brazil

Viviane S Boaventura[1,2], MD, PhD; Malú Grave[3,4,5], DSc; Thiago Cerqueira-Silva[2,6], MD, PhD; Roberto Carreiro[7], MSc; Adélia Pinheiro[8], MD, PhD; Alvaro Coutinho[3], DSc; Manoel Barral Netto[2,6], MD, PhD

[1]Laboratório de Doenças Infecciosas Transmitidas por Vetores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[2]Faculdade de Medicina, Federal University of Bahia, Salvador, Brazil
[3]Department of Civil Engineering, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[4]Fundação Oswaldo Cruz, Salvador, Brazil
[5]Department of Civil Engineering, Fluminense Federal University, Niterói, Brazil
[6]Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[7]Centre for Data and Knowledge Integration for Health, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[8]Departamento de Ciências da Saúde, Universidade Estadual de Santa Cruz, Salvador, Brazil

Corresponding Author:
Viviane S Boaventura, MD, PhD
Laboratório de Doenças Infecciosas Transmitidas por Vetores
Instituto Gonçalo Moniz, Fundação Oswaldo Cruz
Rua Waldemar Falcao, 121
Salvador, 40296-710
Brazil
Phone: 55 71 9619 4245
Email: viviane.boaventura@fiocruz.br

Abstract

Background: Telehealth has been widely used for new case detection and telemonitoring during the COVID-19 pandemic. It safely provides access to health care services and expands assistance to remote, rural areas and underserved communities in situations of shortage of specialized health professionals. Qualified data are systematically collected by health care workers containing information on suspected cases and can be used as a proxy of disease spread for surveillance purposes. However, the use of this approach for syndromic surveillance has yet to be explored. Besides, the mathematical modeling of epidemics is a well-established field that has been successfully used for tracking the spread of SARS-CoV-2 infection, supporting the decision-making process on diverse aspects of public health response to the COVID-19 pandemic. The response of the current models depends on the quality of input data, particularly the transmission rate, initial conditions, and other parameters present in compartmental models. Telehealth systems may feed numerical models developed to model virus spread in a specific region.

Objective: Herein, we evaluated whether a high-quality data set obtained from a state-based telehealth service could be used to forecast the geographical spread of new cases of COVID-19 and to feed computational models of disease spread.

Methods: We analyzed structured data obtained from a statewide toll-free telehealth service during 4 months following the first notification of COVID-19 in the Bahia state, Brazil. Structured data were collected during teletriage by a health team of medical students supervised by physicians. Data were registered in a responsive web application for planning and surveillance purposes. The data set was designed to quickly identify users, city, residence neighborhood, date, sex, age, and COVID-19–like symptoms. We performed a temporal-spatial comparison of calls reporting COVID-19–like symptoms and notification of COVID-19 cases. The number of calls was used as a proxy of exposed individuals to feed a mathematical model called "susceptible, exposed, infected, recovered, deceased."

Results: For 181 (43%) out of 417 municipalities of Bahia, the first call to the telehealth service reporting COVID-19–like symptoms preceded the first notification of the disease. The calls preceded, on average, 30 days of the notification of COVID-19 in the municipalities of the state of Bahia, Brazil. Additionally, data obtained by the telehealth service were used to effectively reproduce the spread of COVID-19 in Salvador, the capital of the state, using the "susceptible, exposed, infected, recovered, deceased" model to simulate the spatiotemporal spread of the disease.

Conclusions: Data from telehealth services confer high effectiveness in anticipating new waves of COVID-19 and may help understand the epidemic dynamics.

(JMIR Public Health Surveill 2023;9:e40036) doi:10.2196/40036

KEYWORDS

telehealth; telemedicine; disease surveillance; mathematical model; COVID-19; prediction; cases; detection; monitoring; surveillance; computational modeling; spread; transmission; disease; infectious diseases; syndromic

Introduction

Telehealth encompasses the distinct ways of interaction between patients and their health care providers. The growing popularity of virtual visits stems from the possibility of faster contact, elimination of transportation time [1], and by providing initial care in remote areas without adequate health care [2], making health care more efficient. Telehealth services can be used for diagnosis, treatment, follow-up, and screening purposes [3]. During the COVID-19 pandemic, telehealth services have been widely used for screening suspected cases [4,5], as limited physical contact reduces everyone's exposure to COVID-19. This successful strategy safely provides access to health care services and expands assistance to remote rural areas and underserved communities in situations of shortage of specialized health professionals [3]. Using health information platforms, health care workers can systematically collect qualified data on suspected cases. Such data can be used as a proxy for the spread of infectious diseases for health surveillance [6-11].

The mathematical modeling of epidemics is a well-established field that has been successfully used for tracking the spread of SARS-CoV-2 infection, supporting the decision-making process on diverse aspects of public health response to the COVID-19 pandemic. Such models are usually defined as compartmental models. The population under study is divided into compartments based on qualitative characteristics, with different assumptions about the nature and rate of transfer across compartments. The urgency of the COVID-19 pandemic has motivated the need for more research in this area, with several models for this pandemic outbreak being presented in the last few years [12-18]. The response of the current models depends on the quality of input data, particularly the transmission rate, initial conditions, and other parameters present in compartmental models. Telehealth systems may feed numerical models developed to model virus spread in a specific region.

Here, we report using a high-quality data set obtained from a state-based telehealth service for disease surveillance both for forecasting the geographical spread of new cases of COVID-19 and feeding an advanced computational model that tries to reproduce the virus spread dynamics. We used data from a toll-free telehealth service of the 14.8 million population from the state of Bahia, Brazil [19], to simulate the spatiotemporal spread of COVID-19 using a compartmental model with diffusion in Salvador, the state capital.

Methods

Telehealth

Service and Data Collection

Telecoronavirus was a toll-free phone-teletriage service offered to the population of the state of Bahia, Brazil (population of 14.8 million, 417 municipalities, located in the northeast of the country; Multimedia Appendix 1) [19]. Risk screening was provided for cases suspected of COVID-19 from March 24, 2020, to July 31, 2020, during the first epidemic wave, starting 18 days after the first confirmed case of COVID-19 in the state. This toll-free service operated 7 days a week and was publicized statewide by the state government through television, the internet, social media, billboards, radio, and newspapers. In total, 77% (320/417) of all cities in Bahia and all sanitary districts of the state capital, Salvador, accessed the telehealth service.

Structured data were collected during teletriage by a health team of medical students supervised by physicians. Data were registered in a responsive web application for planning and surveillance purposes. The data set was designed to quickly identify users, city, residence neighborhood, date, sex, age, and COVID-19–like symptoms (fever, cough, breathlessness, rhinorrhea, and gustatory or olfactory disorder). For this study, the call reporting at least one of those symptoms was considered a suspected case of COVID-19.

Strategies to improve the service quality were applied, including educational and technical support for the health care team, updating the online application based on clinical protocols, and monitoring the quality of information registered by the health care team, as previously described [19]. Data of telehealth service were categorized according to the phone call date, the type of COVID-19–like symptom reported by the user, and the city of residence. For Salvador, the state capital, calls were categorized according to the COVID-19–like symptoms and the user's health district. The COVID-19 cases' notification date and their related geolocation in the state of Bahia were obtained through the Brazilian Ministry of Health. For Salvador, notification data were obtained through the Salvador Municipal Health Department. We further quantified the cities of the state of Bahia and the districts of Salvador, where phone calls reporting symptoms of COVID-19 preceded the first notification of the disease. The lag in days between the first call and the first notification dates was also annotated.

Ethical Considerations

Anonymized consolidated data were provided by the Bahia State Health Secretary for research purposes. The ethics committee of the School of Medicine of Bahia approved the project (approval number 4459774, on 12/13/2020). The requirement for obtaining informed consent was waived due to the characteristics of the research.

Data Source and Preparation for Study

We extracted Telecoronavirus data for Salvador inhabitants (population: 2.88 million; surface area: 692.818 km²) for use in the prediction model. The call records were grouped in epidemiological weeks and residential areas informed by Telecoronavirus users. Additionally, they were grouped in sanitary districts using the list of neighborhoods comprising each of the 13 sanitary districts of the city of Salvador. Official numbers of COVID-19 confirmed cases were also obtained by sanitary districts.

Use of Telehealth Data for Feeding Epidemiological Modeling

We further tested the hypothesis that telehealth systems can help feed numerical models developed to predict a virus spread in a specific region using a spatiotemporal model presented in [12,20-22] and explained as follows.

The Susceptible, Exposed, Infected, Recovered, Deceased Model for COVID-19

The COVID-19 dynamics may be modeled as compartmental models, in which the population under study is divided into compartments and has assumptions about the nature and time rate of transfer from one compartment to another [23]. These models have been used extensively in biological, ecological, and chemical applications [24-26]. They allow for an understanding of the processes at work and predict the dynamics of the epidemic. One of the simplest compartmental models is the "susceptible, infected, removed" model proposed in 1927 by Kermack and McKendrick [27], in which the population is divided into susceptible, infected, and recovered compartments. This basic "susceptible, infected, removed" model can be extended in several ways by enriching the number of compartments, as the "susceptible, exposed, infectious, recovered, deceased" model. Here, we work with a spatiotemporal "susceptible, exposed, infectious, recovered, deceased" model, presented in [20-22], given by,

Where $s(x, t)$, $e(x, t)$, $i(x, t)$, $r(x, t)$, and $d(x, t)$ denote the densities of the *susceptible*, *exposed*, *infected*, *recovered*, and *deceased* populations, respectively. The sum of all the compartments, except for $d(x, t)$, is represented by n_{pop}, which is the total living population. β_i and β_e denote the transmission rates between symptomatic and susceptible individuals and asymptomatic and susceptible individuals, respectively (units 1/days), a denotes the incubation period (units 1/days), γ_e corresponds to the asymptomatic recovery rate (units 1/days), γ_i the symptomatic recovery rate (units 1/days), δ represents the mortality rate (units 1/days), and ν_s, ν_e, ν_i, ν_r are the diffusion parameters of the different population groups as denoted by the subscripted letters: units km²/(persons.days).

To use this model, we need to define all parameters that govern the system of equations and the initial population of each compartment. Note that all these parameters can be considered time- and space-dependent. We need to infer several hypotheses about the initial conditions, especially those related to the exposed compartment, consisting of asymptomatic cases, which are more challenging to estimate.

Model Construction for Salvador, Bahia

We define the beginning of the simulation as April 1, 2020, and simulate 180 days. The initial infected population is set according to the 7 days moving average data provided by the Brazilian Ministry of Health, following the procedure in [12] and [28]. The susceptible population is based on the estimation of the population of each Sanitary district, given by the Brazilian Institute for Geography and Statistics. Recovered and deceased populations start with zero assigned cases (ie, we consider that nobody died or recovered from COVID-19 at the beginning of the simulation).

The exposed compartment is the most nontrivial compartment to assign an initial condition. In previous works, the estimation was based on the amount of the infected population. It was considered that the exposed population could be about 10 times the number of the infected [29]. Here, we consider the 7-day moving average data from telehealth on April 1, 2020, as the initial exposed population. The telehealth data indicated rises in cases approximately 2 weeks before the notification system. Therefore, the calls may indicate a better estimation than the relation with the infected, especially at the epidemic's beginning.

Each compartment's population is divided by the area of each sanitary district and distributed in the 12 areas as people/km². Table 1 shows the differences between how the initial exposed population is considered in this work (related to the telehealth system) and how it would be represented when using the old approach (by multiplying the number of infected individuals by 10). Without the need for any calculation, it is possible to see that there is no correlation between the numbers of each approach.

The biological parameters of the simulation are defined based on the literature, as $\alpha = 1/7\ day^{-1}$, $\gamma_i = 1/24\ day^{-1}$, $\gamma_e = 1/6\ day^{-1}$, $\delta = 1/160\ day^{-1}$ [21]. On the other hand, the contact rate and the diffusion coefficient have to be estimated. They are based on the social distancing estimation, representing the homestay rate for Bahia.

Table 1. Initial exposed population (people/km2) based on 10 times the number of infected and the telehealth calls.

Sanitary district	Ten times the number of infected	Telehealth
Barra Rio Vermelho	27.14	53.00
Brotas	8.57	34.57
Centro Histórico	2.86	13.42
Liberdade	4.29	17.70
Boca do Rio	5.71	23.42
Cabula/Beiru	8.57	43.57
Itapuã	10.00	36.14
Pau da Lima	4.29	25.85
Subúrbio Ferroviário	7.14	30.28
Itapagipe	2.86	18.42
São Caetano or Valéria	7.14	26.57
Cajazeiras	2.86	10.28

Results

The telehealth service received a total of 111,795 calls, 83,175 (74%) of which reporting at least one COVID-19 symptom (fever, cough, breathlessness, rhinorrhea, and gustatory or olfactory disorder) during the first 4 months of the COVID-19 epidemic in the state of Bahia (Figure 1). Olfactory or gustatory dysfunction (smell and taste change), considered highly specific for COVID-19 during the first wave of the pandemic [30-32], was reported in 47% of the calls, suggesting a high frequency of COVID-19 cases among users of the telehealth service.

Calls were registered by 320 (77%) out of 417 cities of Bahia State. The majority of these users of the telehealth service were female (48,873, 60%), with a median age of 38 (IQR 28-49) years. The demand for the service progressively increased between the first and ninth week, with a peak in the 23rd week.

In the 4th month of operation, we observed a reduction in the number of daily calls to values close to those at the beginning of the service (Table 2).

During the first month of operation (April 2020), calls reporting COVID-19–like symptoms were registered in 205 cities, including areas from the north and east of the state. By this time, COVID-19 cases were notified in 136 cities (Figure 2). The first call to the telehealth service preceded the first COVID-19 notification in 181 (43%) out of 417 municipalities of Bahia. In these municipalities, the call occurred on average 30 (IQR 11-42) days before the notification. Additionally, for 68 (38%) of these 181 municipalities, the symptoms registered in the telehealth service included olfactory and gustatory disorders. These symptoms specific for COVID-19 were reported on average 14 (IQR 6-31) days before the first notified COVID-19 case (Multimedia Appendix 2).

Figure 1. Flowchart of calls for Telecoronavirus service according to local and presence of COVID-19–like symptoms.

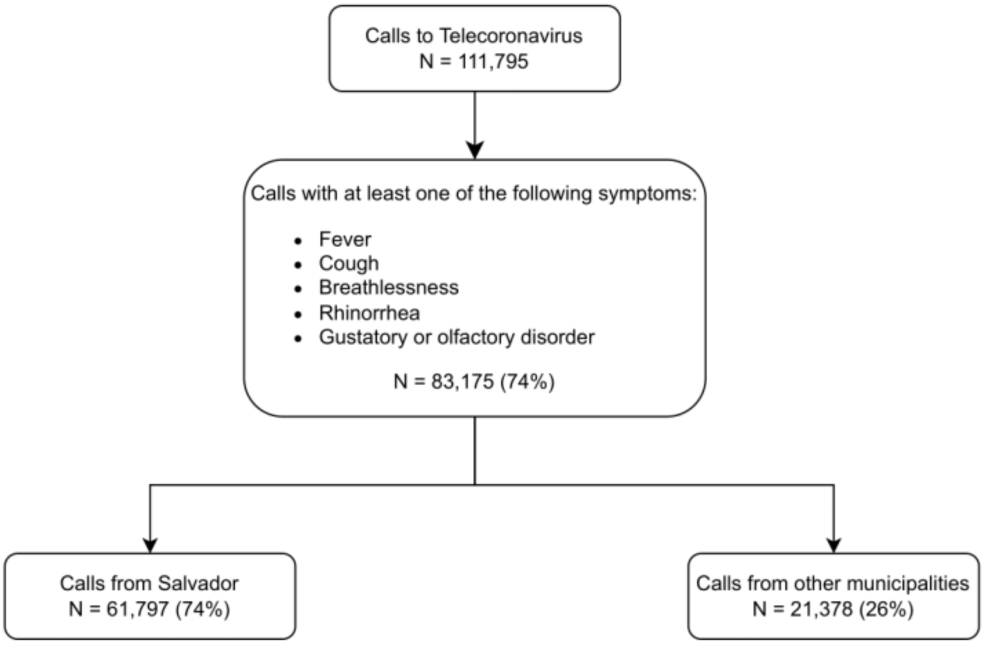


Table 2. Symptoms reported by users and the number of monthly calls.

Characteristic	Salvador (n=61,797)	Other cities (n=21,378)	State (n=83,175)
Age (years), median (IQR)	39 (29, 50)	37 (27, 48)	38 (28, 49)
Month, n (%)			
March	178 (0.3)	49 (0.2)	227 (0.3)
April	10,009 (16)	3464 (16)	13,473 (16)
May	23,227 (38)	4822 (23)	28,049 (34)
June	20,743 (34)	8634 (40)	29,377 (35)
July	7640 (12)	4409 (21)	12,049 (14)
Sex: female, n (%)	36,315 (60)	12,558 (60)	48,873 (60)
Fever, n (%)	30,501 (49)	10,468 (49)	40,969 (49)
Cough, n (%)	30,307 (49)	11,287 (53)	41,594 (50)
Shortness of breath, n (%)	15,518 (25)	6054 (28)	21,572 (26)
Smell or taste change, n (%)	28,741 (47)	9248 (44)	37,989 (47)

We also analyzed the spatial-temporal distribution of calls and notified cases in the districts of Salvador, the state capital. In all districts, an increase in the number of calls reporting COVID-19–like symptoms preceded an elevation in the number of notifications of COVID-19 (Figure 3).

Next, we evaluated if the data obtained with the telehealth service could be helpful in feeding a mathematical model to predict disease spread in Salvador. We simulated 180 days of the epidemic. To validate the results of our model with the available data, we compared the number of notified COVID-19 cases. First, we showed the values obtained for the whole city (Figure 4). Then, we integrated the values of the sanitary districts and plotted them in time. The simulation provides curves of accumulated infection similar to the data of notified cases.

Figure 2. Spatial distribution of calls related to COVID-19–like symptoms (orange) and notifications (green) by months of operation of the Telecoronavirus in Bahia state.

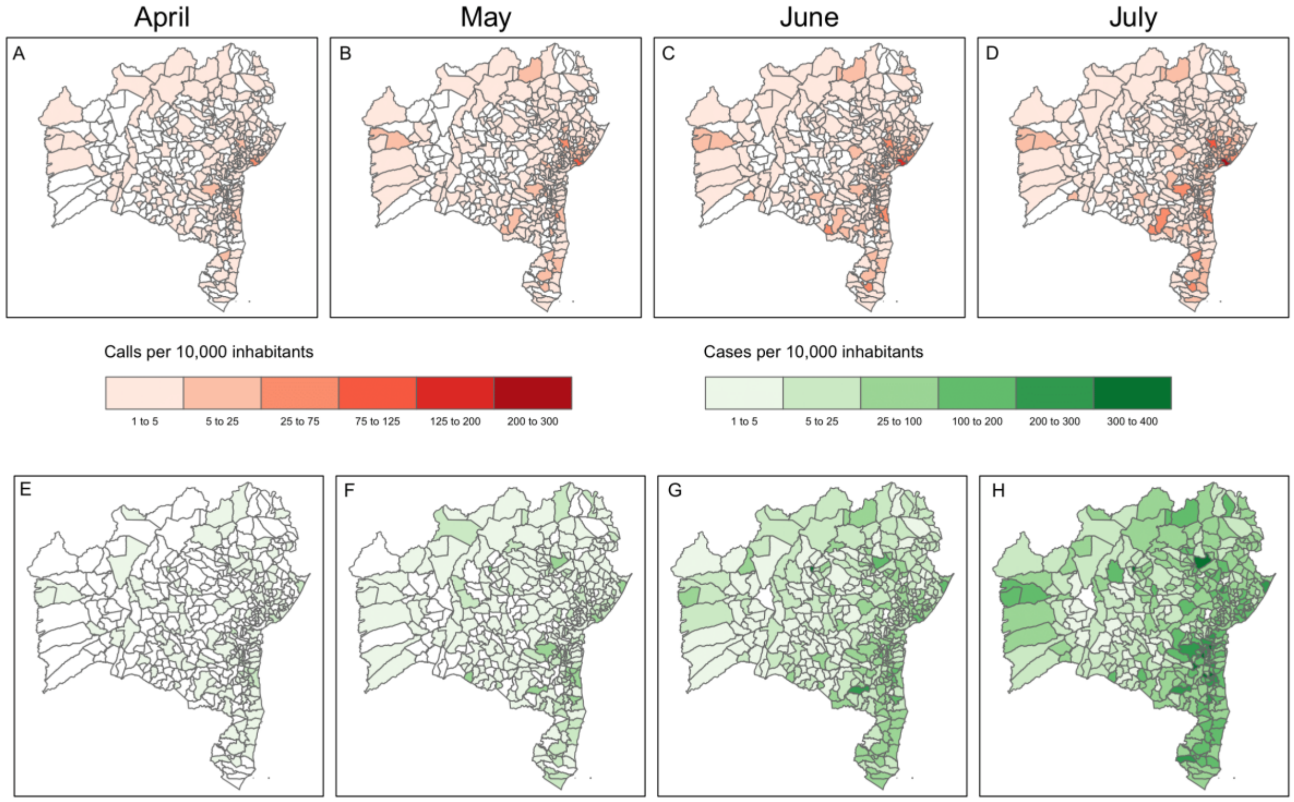


Figure 3. Seven-day rolling average of number of calls reporting COVID-19–like symptoms (blue line) and number of confirmed cases of COVID-19 (red line) in the period of operation of the telehealth service in Salvador, Bahia state, Brazil. (A) All cities. (B) Stratified by sanitary district.

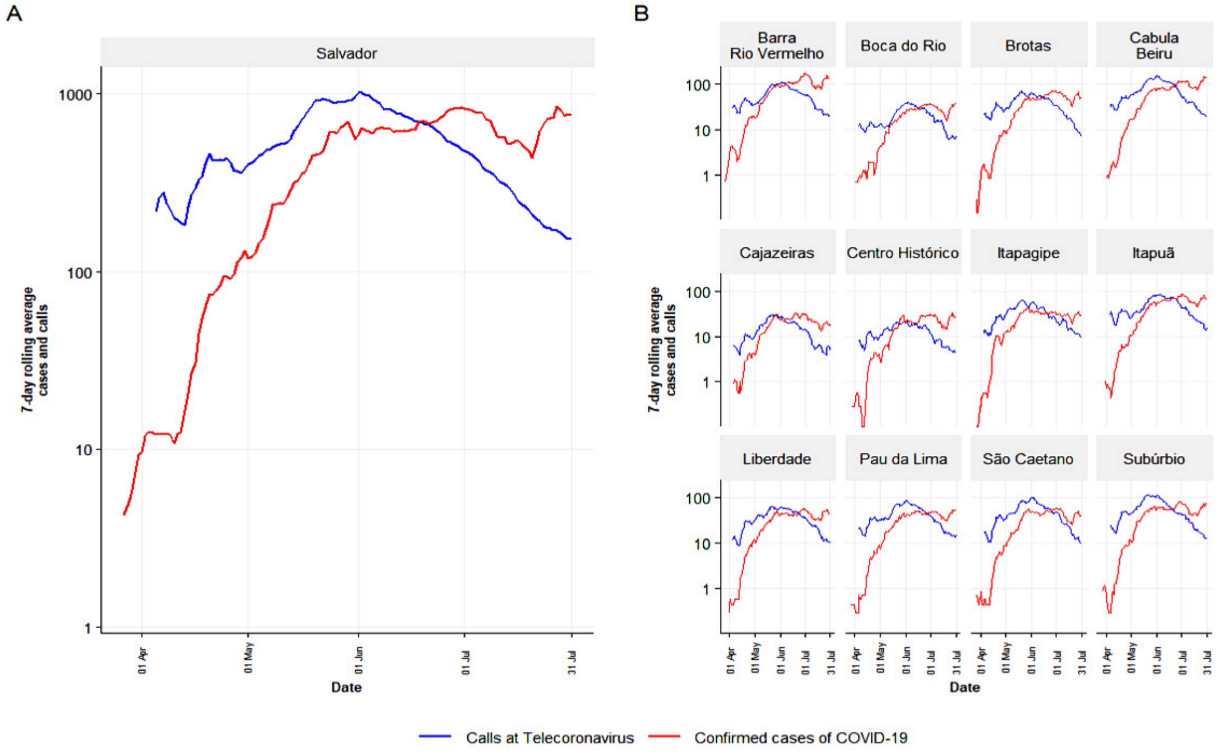
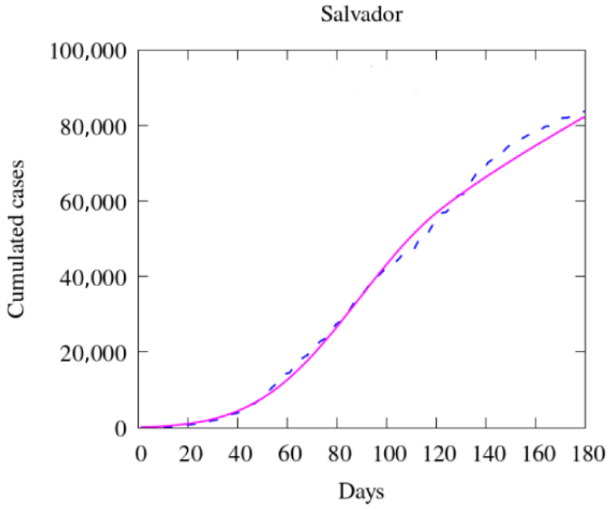


Figure 4. Comparison between simulation and real data of cases at Salvador (total). Dashed blue lines: real data. Pink line: simulation with exposed population based on telehealth calls.



Discussion

Principal Findings

A critical challenge in predicting the spread of cases, especially in locals with a low test rate, is to timely estimate the number of suspected cases. Here, we demonstrated that curated data obtained from a telehealth service could be helpful in anticipating areas with new cases of COVID-19. A rise in calls referring to COVID-19–like illness preceded, in about two weeks, the increase in the notification of COVID-19 in the majority of Bahia state cities demonstrating the potential utility of this instrument for syndromic surveillance in the early phase of the epidemic curve. Additionally, it contains relevant information about the disease under investigation, such as the profile of the affected population, type, and frequency of symptoms, which can help characterize susceptible populations, which is important in the context of epidemics caused by new infectious agents. Calls captured in the telehealth service were also valuable in feeding a numerical disease spread model. Data obtained by a toll-free telehealth service can achieve the goals required for surveillance purposes, such as the following: (1) to be collected early after disease onset, (2) to represent the majority of the population in the covered area, and (3) to be rapidly accessed by the decision-making officials [5,33].

First, in an outbreak caused by a pathogen of airborne transmission, such as COVID-19, the fear of contamination reduces seeking face-to-face medical assistance. The majority of individuals with mild symptoms avoid being exposed in clinical care facilities, reducing the number of individuals who get tested, and therefore confirmed infections may be underestimated. From the patient perspective, telehealth assistance becomes more attractive. Accordingly, a substantial increase in telemedicine service was detected in the first weeks of the pandemic [5,34]. Consequently, data from telehealth services become more suitable for surveillance purposes than laboratory-confirmed infection or electronic medical records obtained from health care units.

Second, telehealth services can cover a population from a wide geographic area. The service offered via telephone dismisses reading or writing skills, digital literacy, or internet access and can be used by all age groups and social strata. If the service is toll-free, most of the population may be covered, which is especially important for low- to middle-income countries, as observed in this case.

Third, data sets containing time-sensitive information collected from telehealth services can be accessed simultaneously by public health authorities for syndromic surveillance. Automating data extraction and real-time interpretation of information in the context of relevant public health emergencies is still challenging. Structured surveys designed for telescreening assistance, as described here, can serve this purpose and offer daily reports for the stakeholders. Although several aspects, such as technical or operational, ethical, and legal points, should be addressed to use these data for disease surveillance, analysis of real-time data obtained from telehealth services is feasible. It has already been performed for contact tracing, screening, and monitoring clinical conditions in emergency response to epidemics [35] and can support decision-making on public health policy.

Our result suggests that the data set from a telehealth service may also be helpful as input to a mathematical model to predict COVID-19 spread. The initial conditions of the exposed and infected populations are essential to short-term predictions. The estimate of exposure may be challenging since this population has no symptoms. Therefore, the hypothesis of assuming the number of exposed as the number of calls might be better than an estimation based on the number of infected, since the telehealth data indicated rises in cases approximately 2 weeks before the notification system.

The simulation for a city of about 2.9 million inhabitants showed that most sanitary districts had a good agreement between the actual data and the simulation. The simulation overestimated the cases in 1 district area and underestimated them in 2 areas. Other districts had an excellent agreement until day 100 and lost accuracy afterward. A loss of accuracy in a fast-changing infectious disease is expected since it is challenging to reproduce long-term predictions. The simulation may still be improved due to several aspects needing to be considered. For example, we used estimates for parameters such as transmission rate and diffusion, and it would be possible to use a machine learning mechanism to find the best ones for this case. Moreover, we define that all sanitary districts would receive the same input parameters, and we know that different regions might have performed differently regarding the population’s behavior and restriction policies. Besides, we used only spatial spread mechanisms due to diffusion and further use of different approaches, as convection and source terms could have offered better estimates. However, the central idea of this report is to show that telehealth data impact a robust algorithm that worked well to simulate the COVID-19 behavior of different regions (Italy, Brazil, and the United States [12]).

Collecting real-time data may help predictors forecast new surges and prepare the population, authorities, and health systems. Moreover, we may infer that our hypothesis works since our simulation with the telehealth data have provided promising results.

Limitations

Our study has some limitations. We used COVID-19–like symptoms as a proxy for confirmed infection. Although this is not accurate, the number of phone calls and the trend in service assessment may suggest the direction and amount of disease spread early on. It shall also be stressed that confirmatory tests may not be available in the early stages of a new epidemic infection. In such cases, the use of the syndromic approach, as employed here, is a critical element for predicting infection spreading. Another significant limitation is that the telehealth service was not equally publicized in all cities, which may have implications for interpreting surveillance data in some areas.

Conclusion

In conclusion, data from telehealth services help model COVID-19 spread and may be helpful in other health situations. Telehealth data and digital health technologies for monitoring disease spread may be especially useful considering the resurgence of new SARS-CoV-2 variants [36]. Data collected from primary health care systems can be used for monitoring the dynamics of COVID-19 cases and the geographic localization of cases. Considering the continuous expansion of telehealth and telemedicine tools in the health care system, the availability of such data may prove a critical tool for modern epidemiology.

Acknowledgments

The authors thank all TeleCoronaVirus volunteers, the Government of the state of Bahia, and all participating institutions.

This study is part of the Alert-Early System of Outbreaks with Pandemic Potential (AESOP) program funded by Fundação Oswaldo Cruz and the Rockefeller Foundation. During the period of analyses, TC-S was a PhD student at the Post-Graduation Program in Health Sciences, Universidade Federal da Bahia, which is supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil, finance code 001. MBN, VSB, and AC are research fellows from the National Council for Scientific and Technological Development (CNPq).

Data Availability

Our agreement with the Bahia State Health Secretary for accessing the database denies authorization of access to a third party. Request for accessing the database should be addressed to the state government, Secretaria de Saúde do Estado da Bahia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Brazil’s map showing the state of Bahia and the capital Salvador.

[PNG File , 62 KB - publichealth_v9i1e40036_app1.png]

Multimedia Appendix 2

Date of first confirmed case, first call, and first call reporting smell or taste disorder for each municipality of Bahia state, Brazil.

[DOCX File , 49 KB - publichealth_v9i1e40036_app2.docx]

References

1. Rogers DG, Santamaria K, Seng EK, Grinberg AS. Behavioral Health, Telemedicine, and Opportunities for Improving Access. Curr Pain Headache Rep 2022 Dec;26(12):919-926 [FREE Full text] [doi: 10.1007/s11916-022-01096-w] [Medline: 36418847]
2. Karamagi HC, Muneene D, Droti B, Jepchumba V, Okeibunor JC, Nabyonga J, et al. eHealth or e-Chaos: The use of Digital Health Interventions for Health Systems Strengthening in sub-Saharan Africa over the last 10 years: A scoping review. J Glob Health 2022 Dec 03;12:04090 [FREE Full text] [doi: 10.7189/jogh.12.04090] [Medline: 36462201]
3. Mbunge E, Batani J, Gaobotse G, Muchemwa B. Virtual healthcare services and digital health technologies deployed during coronavirus disease 2019 (COVID-19) pandemic in South Africa: a systematic review. Glob Health J 2022 Jun;6(2):102-113 [FREE Full text] [doi: 10.1016/j.glohj.2022.03.001] [Medline: 35282399]
4. Khoshrounejad F, Hamednia M, Mehrjerd A, Pichaghsaz S, Jamalirad H, Sargolzaei M, et al. Telehealth-Based Services During the COVID-19 Pandemic: A Systematic Review of Features and Challenges. Front Public Health 2021;9:711762 [FREE Full text] [doi: 10.3389/fpubh.2021.711762] [Medline: 34350154]
5. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the Use of Telehealth During the Emergence of the COVID-19 Pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
6. Katayama Y, Kiyohara K, Komukai S, Kitamura T, Ishida K, Hirose T, et al. The relationship between seasonal influenza and telephone triage for fever: A population-based study in Osaka, Japan. PLoS One 2020;15(8):e0236560 [FREE Full text] [doi: 10.1371/journal.pone.0236560] [Medline: 32760164]
7. Hughes SL, Morbey RA, Elliot AJ, McEwen SA, Greer AL, Young I, et al. Monitoring telehealth vomiting calls as a potential public health early warning system for seasonal norovirus activity in Ontario, Canada. Epidemiol Infect 2019 Jan;147:e112 [FREE Full text] [doi: 10.1017/S0950268818003357] [Medline: 30869009]
8. Pivette M, Mueller JE, Crépey P, Bar-Hen A. Drug sales data analysis for outbreak detection of infectious diseases: a systematic literature review. BMC Infect Dis 2014 Nov 18;14:604 [FREE Full text] [doi: 10.1186/s12879-014-0604-2] [Medline: 25403237]
9. Rodrigues DLG, Belber GS, Padilha FVDQ, Spinel LF, Moreira FR, Maeyama MA, et al. Impact of Teleconsultation on Patients With Type 2 Diabetes in the Brazilian Public Health System: Protocol for a Randomized Controlled Trial (TELEconsulta Diabetes Trial). JMIR Res Protoc 2021 Jan 21;10(1):e23679 [FREE Full text] [doi: 10.2196/23679] [Medline: 33475516]
10. Der-Martirosian C, Wyte-Lake T, Balut M, Chu K, Heyworth L, Leung L, et al. Implementation of Telehealth Services at the US Department of Veterans Affairs During the COVID-19 Pandemic: Mixed Methods Study. JMIR Form Res 2021 Sep 23;5(9):e29429 [FREE Full text] [doi: 10.2196/29429] [Medline: 34477554]
11. King K, Ford D, Haschker M, Harvey J, Kruis R, McElligott J. Clinical and Technical Considerations of an Open Access Telehealth Network in South Carolina: Definition and Deployment. J Med Internet Res 2020 May 22;22(5):e17348 [FREE Full text] [doi: 10.2196/17348] [Medline: 32441657]
12. Grave M, Viguerie A, Barros GF, Reali A, Coutinho ALGA. Assessing the Spatio-temporal Spread of COVID-19 via Compartmental Models with Diffusion in Italy, USA, and Brazil. Arch Comput Methods Eng 2021;28(6):4205-4223 [FREE Full text] [doi: 10.1007/s11831-021-09627-1] [Medline: 34335018]
13. Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, CMMID COVID-19 working group. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. Lancet Infect Dis 2020 Oct;20(10):1151-1160 [FREE Full text] [doi: 10.1016/S1473-3099(20)30457-6] [Medline: 32559451]

14. Ilyin SO. A Recursive Model of the Spread of COVID-19: Modelling Study. JMIR Public Health Surveill 2021 Apr 19;7(4):e21468 [FREE Full text] [doi: 10.2196/21468] [Medline: 33871381]
15. McMahon A, Robb NC. Reinfection with SARS-CoV-2: Discrete SIR (Susceptible, Infected, Recovered) Modeling Using Empirical Infection Data. JMIR Public Health Surveill 2020 Nov 16;6(4):e21168 [FREE Full text] [doi: 10.2196/21168] [Medline: 33052872]
16. Deng Q. Dynamics and Development of the COVID-19 Epidemic in the United States: A Compartmental Model Enhanced With Deep Learning Techniques. J Med Internet Res 2020 Aug 21;22(8):e21173 [FREE Full text] [doi: 10.2196/21173] [Medline: 32763892]
17. Utamura M, Koizumi M, Kirikami S. An Epidemiological Model Considering Isolation to Predict COVID-19 Trends in Tokyo, Japan: Numerical Analysis. JMIR Public Health Surveill 2020 Dec 16;6(4):e23624 [FREE Full text] [doi: 10.2196/23624] [Medline: 33259325]
18. Shapiro MB, Karim F, Muscioni G, Augustine AS. Adaptive Susceptible-Infectious-Removed Model for Continuous Estimation of the COVID-19 Infection Rate and Reproduction Number in the United States: Modeling Study. J Med Internet Res 2021 Apr 07;23(4):e24389 [FREE Full text] [doi: 10.2196/24389] [Medline: 33755577]
19. Cerqueira-Silva T, Carreiro R, Nunes V, Passos L, Canedo BF, Andrade S, Cruz, et al. Bridging Learning in Medicine and Citizenship During the COVID-19 Pandemic: A Telehealth-Based Case Study. JMIR Public Health Surveill 2021 Mar 04;7(3):e24795 [FREE Full text] [doi: 10.2196/24795] [Medline: 33630746]
20. Viguerie A, Veneziani A, Lorenzo G, Baroli D, Aretz-Nellesen N, Patton A, et al. Diffusion-reaction compartmental models formulated in a continuum mechanics framework: application to COVID-19, mathematical analysis, and numerical study. Comput Mech 2020;66(5):1131-1152 [FREE Full text] [doi: 10.1007/s00466-020-01888-0] [Medline: 32836602]
21. Viguerie A, Lorenzo G, Auricchio F, Baroli D, Hughes TJR, Patton A, et al. Simulating the spread of COVID-19 a spatially-resolved susceptible-exposed-infected-recovered-deceased (SEIRD) model with heterogeneous diffusion. Appl Math Lett 2021 Jan;111:106617 [FREE Full text] [doi: 10.1016/j.aml.2020.106617] [Medline: 32834475]
22. Grave M, Coutinho ALGA. Adaptive mesh refinement and coarsening for diffusion-reaction epidemiological models. Comput Mech 2021;67(4):1177-1199 [FREE Full text] [doi: 10.1007/s00466-021-01986-7] [Medline: 33649692]
23. Brauer F, Castillo-Chavez C, Feng Z. Mathematical models in epidemiology. New York, US: Springer; 2019.
24. Brauer F, Castillo-Chavez C, Castillo-Chavez C. Mathematical models in population biology and epidemiology. New York, US: Springer; 2012.
25. Keller JP, Gerardo-Giorda L, Veneziani A. Numerical simulation of a susceptible-exposed-infectious space-continuous model for the spread of rabies in raccoons across a realistic landscape. J Biol Dyn 2013;7 Suppl 1(Suppl 1):31-46 [FREE Full text] [doi: 10.1080/17513758.2012.742578] [Medline: 23157180]
26. Jourdan N, Neveux T, Potier O, Kanniche M, Wicks J, Nopens I, et al. Compartmental Modelling in chemical engineering: A critical review. Chemical Engineering Science 2019 Dec;210:115196. [doi: 10.1016/j.ces.2019.115196]
27. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. A 1997 Jan;115(772):700-721. [doi: 10.1098/rspa.1927.0118]
28. Grave M, Viguerie A, Barros GF, Reali A, Andrade RFS, Coutinho ALGA. Modeling nonlocal behavior in epidemics via a reaction-diffusion system incorporating population movement along a network. Comput Methods Appl Mech Eng 2022 Nov 01;401:115541 [FREE Full text] [doi: 10.1016/j.cma.2022.115541] [Medline: 36124053]
29. Reis RF, de Melo Quintela B, de Oliveira Campos J, Gomes JM, Rocha BM, Lobosco M, et al. Characterization of the COVID-19 pandemic and the impact of uncertainties, mitigation strategies, and underreporting of cases in South Korea, Italy, and Brazil. Chaos Solitons Fractals 2020 Jul;136:109888 [FREE Full text] [doi: 10.1016/j.chaos.2020.109888] [Medline: 32412556]
30. Xydakis MS, Dehgani-Mobaraki P, Holbrook EH, Geisthoff UW, Bauer C, Hautefort C, et al. Smell and taste dysfunction in patients with COVID-19. Lancet Infect Dis 2020 Sep;20(9):1015-1016 [FREE Full text] [doi: 10.1016/S1473-3099(20)30293-0] [Medline: 32304629]
31. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
32. Altwairqi RG, Aljuaid SM, Alqahtani AS. Effect of tonsillectomy on humeral and cellular immunity: a systematic review of published studies from 2009 to 2019. Eur Arch Otorhinolaryngol 2020 Jan;277(1):1-7. [doi: 10.1007/s00405-019-05672-6] [Medline: 31664514]
33. Gunasekeran DV, Tseng RMWW, Tham Y, Wong TY. Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. NPJ Digit Med 2021 Feb 26;4(1):40 [FREE Full text] [doi: 10.1038/s41746-021-00412-9] [Medline: 33637833]
34. Roy B, Nowak RJ, Roda R, Khokhar B, Patwa HS, Lloyd T, et al. Teleneurology during the COVID-19 pandemic: A step forward in modernizing medical care. J Neurol Sci 2020 Jul 15;414:116930 [FREE Full text] [doi: 10.1016/j.jns.2020.116930] [Medline: 32460041]

35. Gao Y, Liu R, Zhou Q, Wang X, Huang L, Shi Q, COVID-19 EvidenceRecommendations Working Group. Application of telemedicine during the coronavirus disease epidemics: a rapid review and meta-analysis. Ann Transl Med 2020 May;8(10):626 [FREE Full text] [doi: 10.21037/atm-20-3315] [Medline: 32566563]
36. Newcomb K, Bilal S, Michael E. Combining predictive models with future change scenarios can produce credible forecasts of COVID-19 futures. PLoS One 2022;17(11):e0277521 [FREE Full text] [doi: 10.1371/journal.pone.0277521] [Medline: 36378674]

Edited by T Sanchez, A Mavragani; submitted 02.06.22; peer-reviewed by S Yan, E Afari-kumah, H Tran Ngoc; comments to author 14.11.22; revised version received 24.11.22; accepted 27.12.22; published 24.01.23.

Please cite as:
Boaventura VS, Grave M, Cerqueira-Silva T, Carreiro R, Pinheiro A, Coutinho A, Barral Netto M
Syndromic Surveillance Using Structured Telehealth Data: Case Study of the First Wave of COVID-19 in Brazil
JMIR Public Health Surveill 2023;9:e40036
URL: https://publichealth.jmir.org/2023/1/e40036
doi:10.2196/40036
PMID:36692925

Original Paper

Regional Difference in the Impact of COVID-19 Pandemic on Domain-Specific Physical Activity, Sedentary Behavior, Sleeping Time, and Step Count: Web-Based Cross-sectional Nationwide Survey and Accelerometer-Based Observational Study

Yosuke Yamada[1], PhD; Hideyuki Namba[2], PhD; Heiwa Date[3], PhD; Shinobu Kitayama[4], PhD; Yui Nakayama[1], MD; Misaka Kimura[5], PhD; Hiroyuki Fujita[5], PhD; Motohiko Miyachi[6], PhD

[1]Department of Physical Activity Research, National Institute of Biomedical Innovation, Health and Nutrition, Osaka, Japan
[2]Department of Liberal Arts and Science, College of Science and Technology, Nihon University, Tokyo, Japan
[3]Faculty of Data Science, Shiga University, Shiga, Japan
[4]Department of Psychology, University of Michigan, Ann Arbor, MI, United States
[5]Institute for Active Health, Kyoto University of Advanced Science, Kyoto, Japan
[6]Faculty of Sport Sciences, Waseda University, Saitama, Japan

Corresponding Author:
Yosuke Yamada, PhD
Department of Physical Activity Research
National Institute of Biomedical Innovation, Health and Nutrition
3-17, Senrioka-shinmachi
Settsu-shi
Osaka, 5660002
Japan
Phone: 81 9078318291
Email: yamaday@nibiohn.go.jp

Abstract

Background: Physical activity (PA) and sedentary behavior (SB) have been affected by the COVID-19 pandemic and its restrictive environments, such as social distancing and lockdown measures. However, regional differences in the changes in domain-specific PA and SB in response to the COVID-19 pandemic are not clearly understood.

Objective: This study aimed to examine regional differences in domain-specific PA and SB, as well as sleeping time in response to the COVID-19 pandemic in Japan.

Methods: A web-based cross-sectional nationwide survey and an accelerometer-based longitudinal observation were conducted. In the web-based survey, we recruited 150 Japanese men and 150 Japanese women for each of the following age groups: 20s, 30s, 40s, 50s, 60s, and 70s (n=1800). A total of 1627 adults provided valid responses to web-based surveillance from June to July 2020. Participants were recruited from urban (Greater Tokyo Area, n=1028), urban-rural (regional core cities, n=459), or rural (regional small and medium cities, n=140) areas. They answered sociodemographic and health-related questions and retrospectively registered the PA data of their average day before and during the COVID-19 pandemic in a web-based PA record system. In the accelerometer-based observation, PA and step count data were obtained using a triaxial accelerometer on people living in urban (n=370) and rural (n=308) areas.

Results: Before the COVID-19 pandemic, there were no significant differences between these 3 regions in the time spent sleeping, staying at home, working or studying, and exercising (*P*>.05). By contrast, people living in urban areas had a longer duration of SB and transportation and a shorter duration of moderate-to-vigorous PA and lying or napping time compared with people living in rural areas (*P*>.05). During the COVID-19 pandemic, a significant decrease was observed in transportation time in urban (–7.2 min/day, *P*<.001) and urban-rural (–2.0 min/day, *P*=.009) areas but not in rural (–0.4 min/day, *P*=.52) areas. The moderate-to-vigorous PA was decreased in urban (–31.3 min/day, *P*<.001) and urban-rural (–30.0 min/day, *P*<.001) areas but not in rural areas (–17.3 min/day, *P*=.08). A significant increase was observed in time spent sleeping in urban (+22.4 min/day, *P*<.001) and urban-rural (+24.2 min/day, *P*<.001) but not in rural areas (+3.9 min/day, *P*=.74). Lying or napping was increased in urban (+14.9 min/day, *P*<.001) but not in rural areas (−6.9 min/day, *P*=.68). PA and step count obtained using an accelerometer significantly decreased in urban (*P*<.05) but not in rural areas (*P*>.05).

Conclusions: The effect of the COVID-19 pandemic on PA and SB was significantly dependent on living area, even in a single country. The effects of PA and SB were greater in the Greater Tokyo Area and regional core cities but were not observed in regional small and medium cities in Japan.

(JMIR Public Health Surveill 2023;9:e39992) doi:10.2196/39992

KEYWORDS

web-based survey; social distancing measure; transportation; physical activity record system; physical activity; sedentary; sleep; sleeping time; COVID-19; impact; pandemic; sleeping pattern; surveillance; demographic; regional; differences

Introduction

Human lifestyles have changed in recent years, with an increasing number of people living a sedentary lifestyle, particularly in the countries with higher human development index [1]. Adequate amounts of physical activity (PA), limited sedentary behavior (SB), and proper sleeping time (ST) are essential for maintaining health [2-4]. Physical inactivity increases the risk of many adverse health conditions, such as coronary heart disease, type 2 diabetes, colon and breast cancers, mobility disabilities, and mortality [5].

The COVID-19 pandemic has changed the world and strongly affected the health of the people, with more than 640 million people with COVID-19, causing more than 6.6 million deaths worldwide [6]. Studies have indicated that the COVID-19 pandemic and its restrictive environment induced by the government's statements about social distancing and lockdown measures affect PA, SB, and ST in children, adolescents, and young and older adults worldwide [7-15]. Stockwell et al [16] conducted a systematic review of the changes in PA and SB from before to during the COVID-19 pandemic restrictive environment. Most studies showed a decreased PA and an increased SB in the 66 included studies. They noted that different degrees of *lockdown* in different countries, even regions within a country, make it difficult to objectively assess how different degrees of restrictive environments impact behaviors. They also noted that most studies report PA without investigating in detail the types and durations of PA engaged before and during the lockdown. Thus, it would be beneficial to investigate these because the magnitude of changes will impact the effects on health.

For the Japanese population, a web-based survey using the short version of the International Physical Activity Questionnaire (IPAQ) found that PA decreased in April and recovered in June 2020 in older Japanese adults [7,8]. Another web-based survey using the Global Physical Activity Questionnaire (GPAQ) of Japanese workers found a significant increase in ST and a significant decrease in leisure-related moderate-to-vigorous physical activity (MVPA) during the COVID-19 outbreak (July 2020) compared to before the outbreak (February 2019) [13]. The IPAQ and GPAQ are considered world-standard questionnaires, with the advantage of comparing the results with worldwide data [17]. However, their external validity against the doubly labeled water (DLW) method is poor [18] compared with objective accelerometers [19,20]. The most recent meta-analysis also indicated that the criterion validity of IPAQ and GPAQ against objective smart trackers was low (weighted mean of r=0.23) [17]. Thus, more reliable systems to assess PA and SB in daily life are required for a stronger conclusion. In addition, it is essential to examine regional differences in PA and SB in response to the COVID-19 pandemic.

This study assessed a 24-hour domain-specific PA, SB, and ST in Japanese adults nationwide using a previously validated web-based simplified physical activity record (sPAR) system. In addition, we obtained longitudinal PA and step count data by a triaxial accelerometer on people living in urban and rural areas. We hypothesized that the impact of the COVID-19 pandemic and its restrictions (including government social distancing measures) on domain-specific PA, SB, ST, or step count vary by region of residence, even within a single country.

Methods

Data Source

The web-based surveillance was conducted through an internet research company (Cross Marketing Inc). Cross Marketing Inc and its partner companies have active panels with over 5 million people who registered their sociodemographic information in the company's database and responded to at least one survey within the last year [21]. The company recruited 1800 adults from their active panels from June to July 2020; in total, 150 men and 150 women were recruited for each of the following age groups: 20s, 30s, 40s, 50s, 60s, and 70s. Individuals completed the web-based sPAR system, which was included before and during the COVID-19 pandemic. The company provided the URL via email and requested each participant to complete a survey. All the participants signed a web-based informed consent form. The research company removed participants' names, addresses, contact information, and any other details that might be used to identify individuals prior to transferring the data to researchers from the data. A total of 1627 adults provided valid responses for this web-based survey. Self-reported age, height, and weight data were obtained. The participants were categorized by gender and residential area, which were divided into urban (Greater Tokyo Area), urban-rural (regional core cities), and rural (regional small and medium cities), based on the Fourth National Comprehensive Development Plan, Japanese National Land Agency [22].

Ethical Considerations

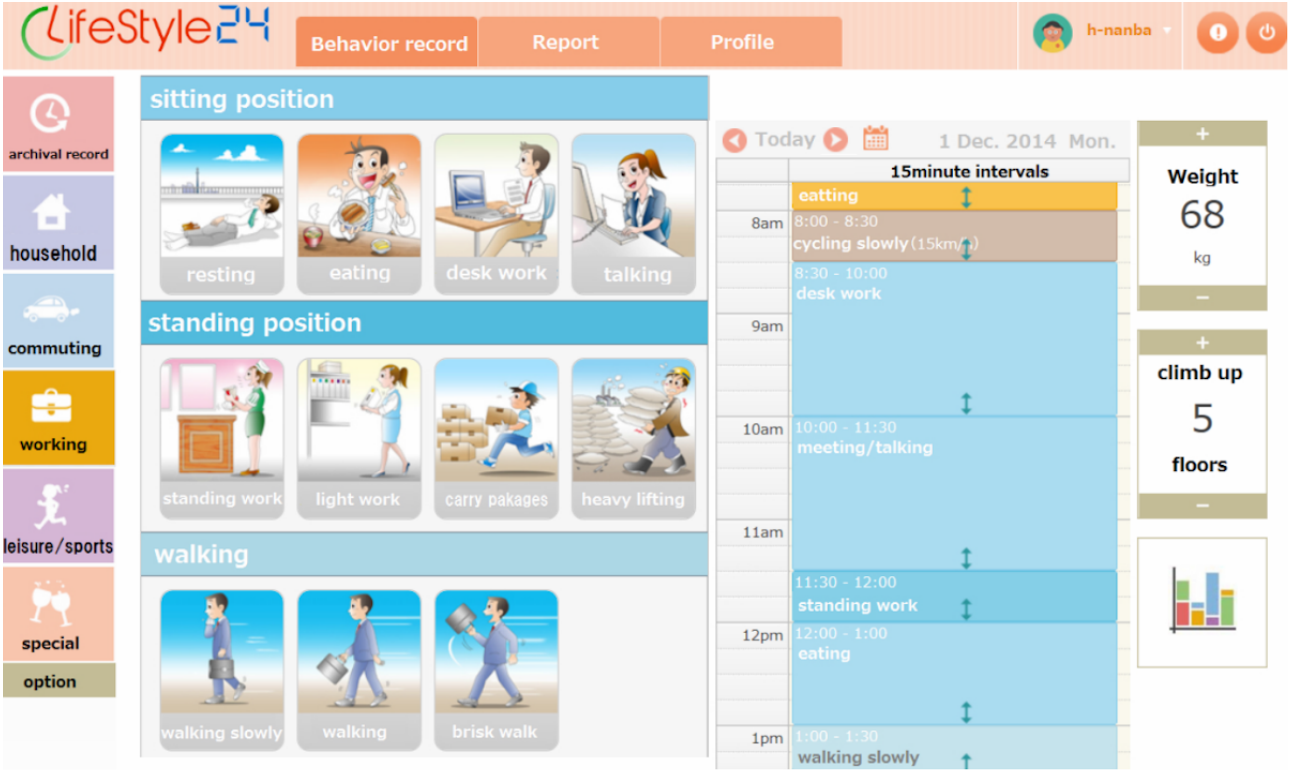
This study was conducted with approval from the ethics committees of Kyoto University of Advanced Science (KUAS no.20-3) and the National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN no. 202). Participants were provided with an information sheet on the landing page of the web-based survey. Participants were only allowed to continue participating if they acknowledged that they had read the information sheet and agreed to give informed consent.

Web-Based PA Record

The sPAR is a unique system for assessing 24-hour domain-specific PA, SB, and ST [23]. The external validity is moderate to high (*r*=0.8 for total energy expenditure and *r*=0.63 for PA level) [23-25]. One study developed a web-based sPAR system and found that external and criterion validity were also moderate to high (*r*=0.68 for activity energy expenditure [AEE] measured by DLW and *r*=0.87 for total energy expenditure; *r*=0.77 for AEE; *r*=0.71 for PA level assessed by a validated triaxial accelerometer) [26,27]. The external validity from another study of a triaxial accelerometer and sPAR against the DLW method was similar and high [24].

A representative screenshot of the web-based sPAR system is shown in Figure 1. The system collected recalled activities over 24 hours. The participants were instructed to respond to questions about a typical day that is not a weekend or holiday. The participants selected one of the preset activities performed in each of the four categories (household, transportation, work-related, and leisure or sports) for 15-minute intervals. In this system, 91 activities among the four categories were preset. The details of the web-based sPAR system have been previously described [26,28]. The intensity of each activity was entered into the system as metabolic equivalent of tasks (METs) based on the 2011 Compendium of Physical Activities [29]. MVPA (≥3.0 METs) and SB (1.0-1.5 METs) were calculated based on the MET value of each activity.

Figure 1. Screenshot of the web-based simplified physical activity record (sPAR) system (Namba et al [28]).



eatting

8:00 - 8:30 cycling slowly (15km/h)

8:30 - 10:00 desk work

10:00 - 11:30 meeting/talking

11:30 - 12:00 standing work

12:00 - 1:00 eating

1:00 - 1:30 walking slowly

Weight 68 kg

climb up 5 floors

Accelerometer

In addition to this web-based survey, we also used the data set of a triaxial accelerometer that had been previously published as a short research paper [11]. The original data contained 628,111 observation days from January 1, 2019, to January 1, 2021, from 1167 unique users of the accelerometer nationwide in Japan. From the 1167 unique users' data, people who lived in the Greater Tokyo Area (n=370) and those who lived in rural areas (n=308) were selected to compare urban and rural differences in this study based on the information about the prefecture in which they lived [11]. The average step counts and AEEs for January 2019, April 2019, January 2020, and April 2020 were calculated. Step counts and AEE were monitored using a triaxial accelerometer (EW-NK63; Panasonic). This accelerometer was manufactured based on the Actimarker (EW4800, Panasonic), an accelerometer for research use, as a low-cost version of the one for the general public [30]. Daily step counts and AEE were stored on a server via participants' smartphones. All participants read the explanation of the study and indicated their understanding before providing informed consent. The details of the METs calculations have been previously described [31,32]. The output of the accelerometer was highly correlated (R^2=0.86) with the METs while walking or running at 7 speeds ranging from 40 to 160 m/min^{-1} and during the following 7 daily activities: performing self-care while standing, changing clothes, cooking, simulating eating supper, washing dishes, doing laundry, and using a vacuum cleaner [31,32]. AEE was calculated from 24-hour average METs, and this accelerometer highly correlates with the DLW method [20,24].

Statistical Analyses

Data are shown as means and standard deviations in the tables and as means and standard errors in the figures. One-way ANOVA was performed to compare age, and a chi-square test was conducted to compare the percentage of either gender between residential areas. Socioeconomic status, smoking, alcohol consumption, and health-related information, including medical history, were extracted from the web-based survey. Psychological distress was assessed by the Kessler Psychological Distress Scale (K6) screening scales [33-35]. One-way analysis of covariance was performed to compare other values between residential areas with the variables in Tables 1 and 2. A 2-tailed paired *t* test was conducted to compare values before and during the COVID-19 pandemic. The results were considered significant at a *P*<.05. SPSS version 22 (IBM Corp) was used for all statistical analyses.

Table 1. Physical characteristics of the participants (n=1627).

Characteristics	Urban (n=1028)	Urban-rural (n=459)	Rural (n=140)	*P* value
Men, n (%)	542 (52.7)	267 (58.2)	79 (56.4)	.21
Age (years), mean (SD)	49 (17)	48 (16)	46 (15)	.20
Height (cm), mean (SD)	164 (8)	165 (11)	164 (9)	.77
Weight (kg), mean (SD)	61 (20)	61 (13)	62 (14)	.68

Table 2. Social and health status of the participants (N=1627).

Social and health status	Urban (n=1028), n (%)	Urban-rural (n=459), n (%)	Rural (n=140), n (%)	*P* value
Current worker or student	701 (68.2)	315 (68.6)	108 (77.1)	.096
Alcohol drinker	609 (59.2)	250 (54.5)	53 (37.9)	*<.001*
Current smoker	210 (20.4)	92 (20.0)	35 (25.0)	.42
Living alone	191 (18.6)	73 (15.9)	20 (14.3)	.27
Living with pets	220 (21.4)	119 (25.9)	39 (27.9)	.07
Household income ≥10 million yen (≥US $77,000)	117 (11.4)	28 (6.1)	10 (7.1)	*.004*
Good subjective economic status	541 (52.6)	183 (39.9)	62 (44.3)	*<.001*
Education ≥13 years	821 (79.9)	296 (64.5)	83 (59.3)	*<.001*
Good subjective health	836 (81.3)	350 (76.3)	107 (76.4)	.05
Psychological Distress (K6[a] score≥5)	433 (42.1)	221 (48.1)	62 (44.3)	.096
Having purpose in life	725 (70.5)	310 (67.5)	96 (68.6)	.496
No medication	694 (67.5)	293 (63.8)	93 (66.4)	.38
Hypertension	158 (15.4)	82 (17.9)	14 (10.0)	.08
Hyperlipidemia	74 (7.2)	27 (5.9)	5 (3.6)	.21
Diabetes	52 (5.1)	24 (5.2)	5 (3.6)	.72
Heart disease	17 (1.7)	10 (2.2)	5 (3.6)	.29
Cancer	16 (1.6)	9 (2.0)	1 (0.7)	.58
Depression	39 (3.8)	26 (5.7)	9 (6.4)	.15

[a] K6: Kessler Psychological Distress Scale.

Results

The physical characteristics of the participants are shown in Table 1. No significant differences were observed in age, gender, height, or weight among the 3 residential area categories (*P*>.05). Table 2 shows the social and health status of the participants. The urban area had a higher percentage of alcohol drinkers, people with higher household income and subjective economic status, and those with higher education than other areas (*P*<.05). However, other variables were not significantly different between areas (*P*>.05).

In Table 3, Before the COVID-19 pandemic, there were no significant differences in sleeping time, time spent at home, working or studying time, and exercise time (*P*>.05) in the 3 regions. By contrast, people living in urban areas had a longer duration of SB and transportation and a shorter duration of MVPA and lying or napping time compared with people living in rural areas (*P*<.05). Significant and negative correlation was observed between work duration and ST (*r*= 0.373, *P*<.001), and between transportation time and ST (*r*= 0.289, *P*<.001).

Table 4 and Figure 2 show changes in ST, SB, transportation, and MVPA before and during the COVID-19 pandemic. A significant decrease was observed in transportation time (−7.2 min/day, *P*<.001) and urban-rural (−2.0 min/day, *P*=.009) areas but not in rural (−0.4 min/day, *P*=.52) areas during the pandemic. The MVPA was decreased in urban (−31.3 min/day, *P*<.001) and urban-rural (−30.0 min/day, *P*<.001) areas but not in rural areas (−17.3 min/day, *P*=.08). A significant increase was observed in time spent sleeping in urban (+22.4 min/day, *P*<.001) and urban-rural (+24.2 min/day, *P*<.001) but not in rural areas (+3.9 min/day, *P*=.74). Lying or napping was increased in urban (+14.9 min/day, *P*<.001) but not in rural areas (−6.9 min/day, *P*=.68).

Figure 3 shows the relationship between the changes in PA and SB during the COVID-19 pandemic. The change in the duration of transportation was significantly and negatively correlated with the change in the duration of staying at home (*P*<.05). The change in MVPA duration was significantly and negatively correlated with the change in the duration of lying or napping and the change of the duration of SB (*P*<.05).

Figure 4 shows the results of the accelerometer monitoring. The step counts and AEE obtained by an accelerometer significantly decreased during the COVID-19 pandemic in urban areas (*P*<.05) but not in rural areas (*P*>.05).

Table 3. Physical activities and sedentary behavior of the participants (n=1627).

Variables	Urban (n=1028), mean (SD)	Urban-rural, (n=459), mean (SD)	Rural, (n=140), mean (SD)	*P* value
Before the COVID-19 pandemic				
Sleeping time (min/day)	453 (139)	462 (145)	460 (137)	.51
Lying or napping time (min/day)	99 (153)	111 (151)	132 (194)	*.045*
SB[a] (min/day)	610 (248)	592 (242)	528 (252)	*.001*
MVPA[b] (min/day)	168 (181)	169 (185)	214 (212)	*.02*
Staying time at home (min/day)	639 (301)	650 (284)	613 (275)	.41
Transportation (min/day)	65 (85)	48 (72)	41 (49)	*<.001*
Working or studying time (min/day)	265 (262)	266 (260)	309 (258)	.17
Exercise time (min/day)	10 (30)	6 (21)	10 (62)	.07
During the COVID-19 pandemic				
Sleeping time (min/day)	476 (165)[c]	487 (181)[c]	464 (141)	.32
Lying or napping time (min/day)	114 (179)[c]	121 (175)	124 (208)	.68
SB (min/day)	625 (259)[d]	595 (251)	551 (273)	*.002*
MVPA (min/day)	137 (161)[c]	139 (168)[c]	197 (205)	*<.001*
Staying time at home (min/day)	658 (299)[c]	654 (286)	630 (279)	.57
Transportation (min/day)	57 (83)[c]	46 (72)[d]	41 (49)	*.007*
Working or studying time (min/day)	249 (261)[c]	260 (258)[e]	296 (261)	.12
Exercise time (min/day)	10 (29)	5 (19)	10 (62)	*.03*

[a] SB: sedentary behavior.

[b] MVPA: moderate-to-vigorous physical activity.

[c] Significant difference versus before the COVID-19 pandemic: *P*<.001.

[d] Significant difference versus before the COVID-19 pandemic: *P*<.01.

[e] Significant difference versus before the COVID-19 pandemic: *P*<.05.

Table 4. Change in score of physical activity and sedentary behavior in the participants (N=1928).

Variable	Urban (n=1028), mean (IQR)	*P* value	Urban-rural (n=459), mean (IQR)	*P* value	Rural (n=140), mean (IQR)	*P* value
Sleeping time (min/day)	22.4 (14.8, 30.0)	*<.001*	24.2 (11, 37.4)	*<.001*	3.9 (–19, 26.7)	.74
Lying or napping time (min/day)	14.9 (6.2, 23.7)	*.001*	10.0 (–2.7, 22.7)	.12	–6.9 (–40.1, 26.4)	.68
SB[a] (min/day)	14.9 (4.1, 25.8)	*.007*	2.9 (–14.6, 20.3)	.75	22.5 (–7.5, 52.5)	.14
MVPA[b] (min/day)	–31.3 (–38.8, –23.7)	*<.001*	–30.0 (–40.8, –19.1)	*<.001*	–17.3 (–36.3, 1.8)	.08
Staying time at home (min/day)	18.8 (12.1, 25.4)	*<.001*	3.7 (–3.6, 11.0)	.33	17.3 (–4.0, 38.5)	.11
Transportation (min/day)	–7.2 (-9.4, -5.0)	*<.001*	–2.0 (–3.5, –0.5)	*.009*	–0.4 (–1.7, 0.9)	.52
Working or studying time (min/day)	–16.3 (–21.7, –11)	*<.001*	–5.6 (–10.1, –1.2)	*.01*	–13.4 (–28.7, 1.9)	.09
Exercise time (min/day)	0.0 (–0.9, 0.9)	.98	–0.6 (–1.3, 0.1)	.09	–0.2 (–0.7, 0.3)	.42

[a] SB: sedentary behavior.

[b] MVPA: moderate-to-vigorous physical activity.

Figure 2. The changes in sleeping time, sedentary behavior, and physical activity from before to during the COVID-19 pandemic were assessed by a web-based physical activity record system (mean and SE). Significant differences were observed between the urban (Greater Tokyo Area) or urban-rural areas and the rural area in changes in the duration of sleeping, lying or napping, transportation, and moderate-to-vigorous physical activity (MVPA). ***Signficantly changed from before the COVID-19 pandemic (*P*<.001).

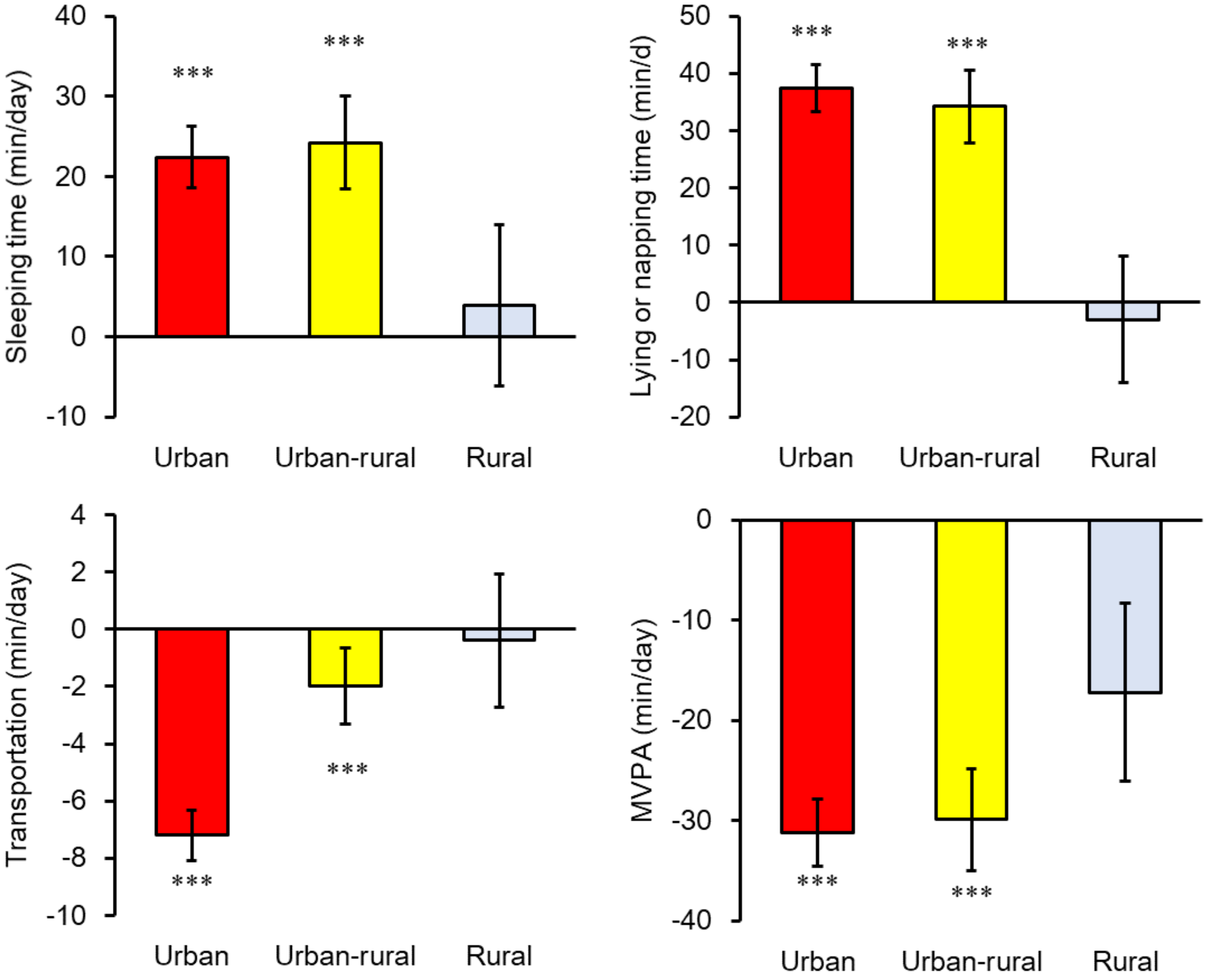


Figure 3. The relationship between changes in physical activity and sedentary behavior (SB) during the COVID-19 pandemic. The change in moderate-to-vigorous physical activity (MVPA) duration was significantly and negatively correlated with the change in the duration of lying or napping and SB. The change in transport duration was significantly and negatively correlated with the change in the duration of staying home.

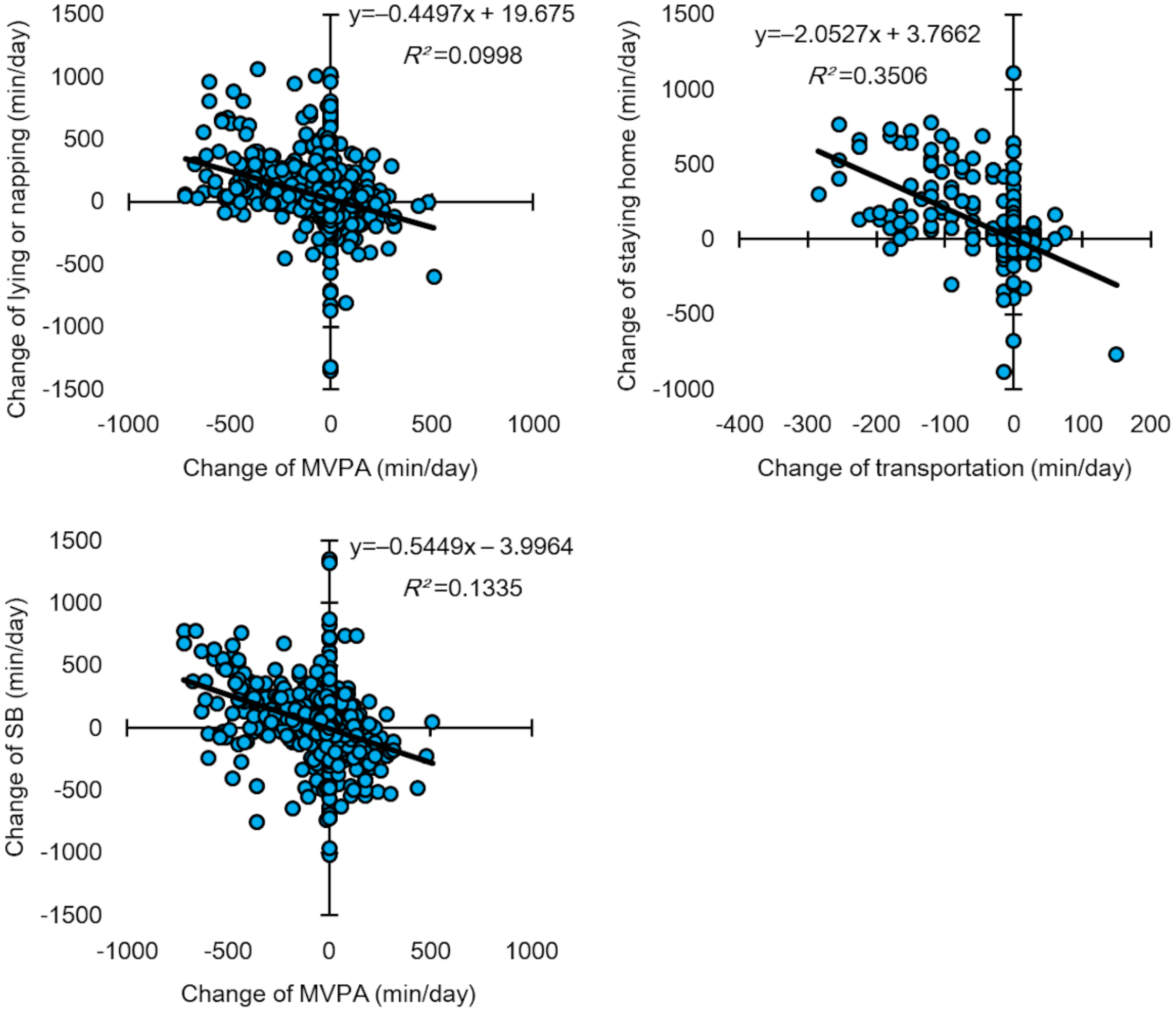
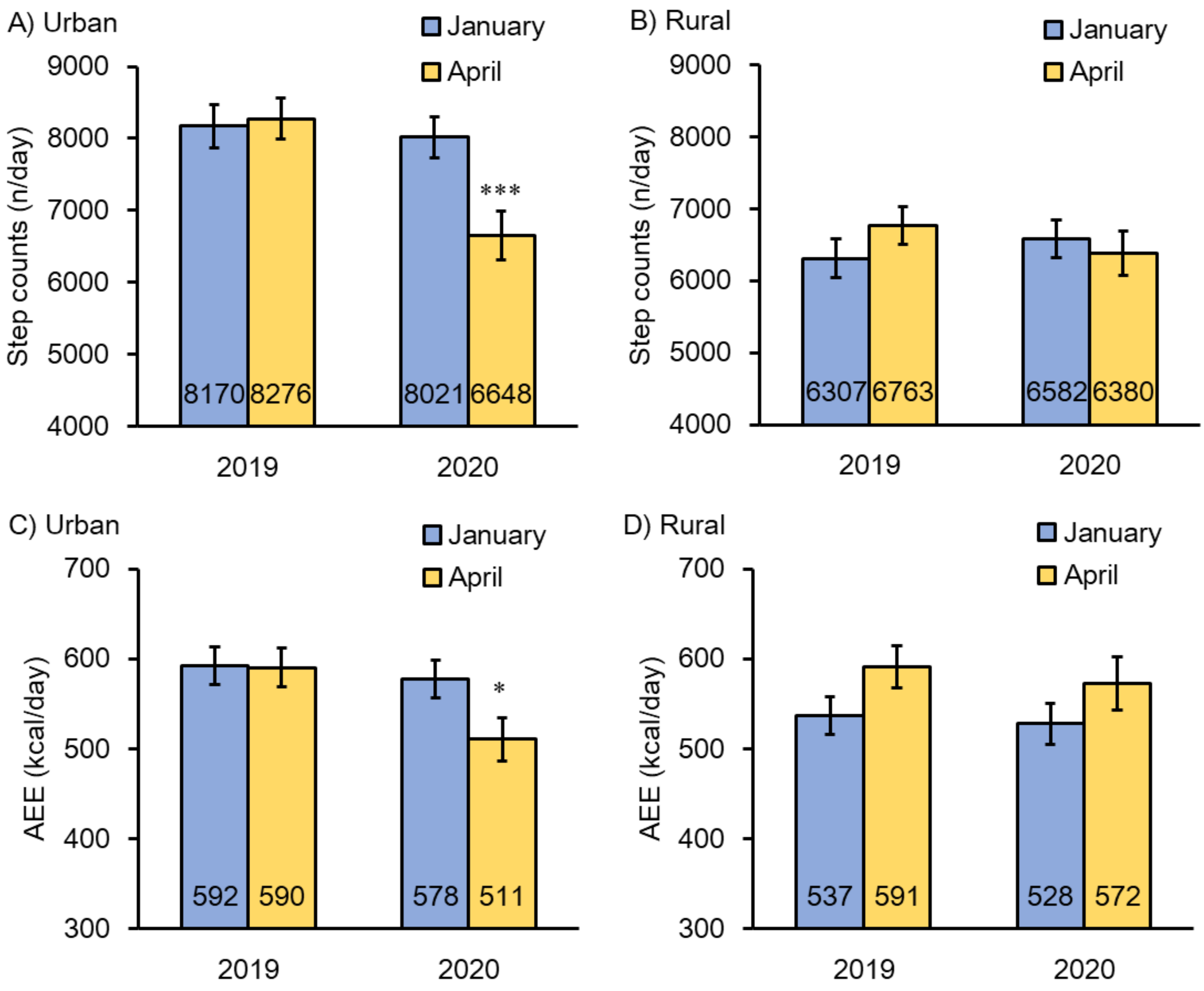


Figure 4. Objectively measured step counts and activity energy expenditure (AEE) by triaxial accelerometer (mean and SE). There is no significant difference between January and April 2019 in step counts and AEE in urban and rural prefectures. The step counts (*P*<.001) and AEE (*P*<.05) in April 2020 were significantly decreased in urban prefectures when January and April 2019 are compared to January 2020.



Discussion

Principal Findings

This study clearly shows regional differences in domain-specific PA, SB, and ST changes in response to the COVID-19 pandemic. The Japanese government first declared a state of emergency in response to the COVID-19 outbreak on April 7, 2020 [36,37]. The effects of the COVID-19 outbreak and the government-declared state of emergency were significantly larger in the urban (Greater Tokyo Area) and urban-rural (regional core cities) areas than in rural areas (regional small and medium cities). Particularly, MVPA decreased by 31 min/day (–18.4%), from 168 min/day before to 137 min/day during the COVID-19 pandemic. Sleeping time (+26 min/day on average) and lying or napping time (+36 min/day on average) increased significantly during the COVID-19 pandemic in the urban and urban-rural areas. In addition, the duration of transportation was significantly decreased by 7 min/day, from 65 min/day before to 57 min/day during the COVID-19 pandemic in the urban area. In addition, we confirmed these regional differences using the triaxial accelerometer as an objective method.

Several previous studies have observed similar patterns in other population groups [7,13,38,39]. However, many surveys that addressed domain-specific PA used PA questionnaires, such as the IPAQ and GPAQ. The external and criterion validity of the PA questionnaires was relatively low when compared with a gold standard method, such as the DLW method, or with objective accelerometers [17-20]. Thus, more reliable systems to assess PA and SB in daily life are required for a stronger conclusion. The web-based sPAR system is a unique method for obtaining PA and SB data with moderate or high external validity [23,26,27].

In this study, Japanese living in urban or urban-rural areas were found to have spent more time lying and sleeping during the COVID-19 pandemic than before the pandemic. The average ST of the participants was 455 min/day before the COVID-19 pandemic. This is consistent with the Organization for Economic Co-operation and Development (OECD) statistics in 2019, which shows 442 min/day as the average sleep duration for individuals in Japan [40]. Japan has the shortest ST compared to other countries [40]. OECD Better Life Index stated that “the percentage of employees working very long hours is higher in Japan than the OECD average of 11%, and full-time workers devote less of their day to personal care and leisure than the

OECD average" [41]. This survey found a significant and negative correlation between working duration and ST (P<.05).

In Japan, several companies have introduced and supported some level of teleworking to avoid denseness in response to government statements [18]. Our study found a significant and negative correlation between transportation time and ST. The average transportation time in the Greater Tokyo Area was 65 min/day, which was significantly longer than that in the rural areas. Teleworking leads to less time traveling for work, and people might sleep more than they did before the COVID-19 pandemic. This might be a favorable effect of teleworking. However, people have spent more time in a lying position during the COVID-19 pandemic. A recent study found that an increase in SB was significantly associated with an increase in the motivation and PA aspects of fatigue [18]. Therefore, people who live in urban or urban-rural areas and spend more SB time might experience increased fatigue. However, further research is required [42].

MVPA is defined as any activity ≥3 METs, including sports, exercise, and nonexercise PA, such as walking or bicycling to supermarkets or active periods of childcare. In this study, significantly decreased exercise time was not observed in all areas (P>.05); however, MVPA duration significantly and largely decreased during the COVID-19 pandemic (P<.001). Our study indicated that nonexercise MVPA decreased during the COVID-19 pandemic by about 30 min/day in the urban and urban-rural areas but not in rural areas in Japan.

One of the factors contributing to the regional differences is the difference in population density. The Tokyo metropolitan area has a very high population density. Thus, requests for self-restraint from the government were more severe than in smaller regional cities. It is also likely that a stronger response was seen in terms of people's behavior to avoid crowds. Another factor that may have contributed to the regional differences is the occupational characteristics of the residents. In rural small and medium-sized cities, a smaller percentage of residents are office workers, university faculty, and people from other professional occupations, while a higher percentage comprise blue-collar workers and agricultural and workers in the fishing industry. We believe that differences in population density in the places where they work or live and whether they can be replaced by teleworking also had an impact on the change of PA and SB.

Limitations

Our study has some methodological limitations. First, our web-based survey and accelerometer-based study used convenient samples without random sampling; thus, we did not calculate sample sizes and effect sizes in this study. Those not using the internet or accelerometers, especially older people, could not be included in this study. Therefore, this study's results do not reflect the Japanese population characteristics; participants could be more health conscious than the general population, and our study potentially includes selection bias. Second, possible inconsistency of PA levels on workdays and weekends may exist. That is, the 24-hour sPAR data from a single acquisition may not represent a typical week for the PA level of the study participants. Particularly, the current system cannot examine weekly or monthly exercise habits. In addition, while we obtained the physical and mental health status in the web-based survey, we could not obtain the physical and mental health status in the accelerometer study, which may affect the results. Finally, as the main limitation of this study is based on the study design, it cannot be argued that such differences are based on the COVID-19 pandemic or the aging effect.

Conclusion

This research, using a web-based sPAR system, found several lifestyle changes, indicating that people became less active and spent a longer amount of time in the lying position in urban and urban-rural areas during the COVID-19 pandemic. In addition, we found that Japanese living in urban and urban-rural areas had increased sleeping time during the COVID-19 pandemic. Longer sleep duration during the COVID-19 pandemic may favor the population's health, as Japan has the shortest sleeping time worldwide. However, the less time spent on MVPA and the longer time spent on SB, particularly in the lying position, would have an unfavorable effect on health. The sociopsychological aspects of the effect of such lifestyle change should be examined in future studies.

Acknowledgments

This research is supported by Grants of Japan Society for the Promotion of Science (JSPS) KAKENHI for YY (18H03164).

Conflicts of Interest

None declared.

References

1. Dumith SC, Hallal PC, Reis RS, Kohl HW. Worldwide prevalence of physical inactivity and its association with human development index in 76 countries. Prev Med 2011;53(1-2):24-28. [doi: 10.1016/j.ypmed.2011.02.017] [Medline: 21371494]
2. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020. URL: https://apps.who.int/iris/rest/bitstreams/1315866/retrieve [accessed 2023-01-31]
3. Watanabe D, Yamada Y, Yoshida T, Watanabe Y, Hatamoto Y, Fujita H, et al. Association of the interaction between physical activity and sitting time with mortality in older Japanese adults. Scand J Med Sci Sports 2022 Dec;32(12):1757-1767 [FREE Full text] [doi: 10.1111/sms.14234] [Medline: 36112073]

4. Watanabe D, Yoshida T, Watanabe Y, Yamada Y, Miyachi M, Kimura M, et al. Combined use of sleep quality and duration is more closely associated with mortality risk among older adults: a population-based Kyoto-Kameoka prospective cohort study. J Epidemiol 2022 Sep 24:1-27 [FREE Full text] [doi: 10.2188/jea.JE20220215] [Medline: 36155361]
5. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
6. COVID Coronavirus Pandemic. Worldometer. 2022. URL: https://www.worldometers.info/coronavirus/ [accessed 2022-11-19]
7. Yamada M, Kimura Y, Ishiyama D, Otobe Y, Suzuki M, Koyama S, et al. Effect of the COVID-19 Epidemic on Physical Activity in Community-Dwelling Older Adults in Japan: A Cross-Sectional Online Survey. J Nutr Health Aging 2020 Jun 23;24(9):948-950. [doi: 10.1007/s12603-020-1501-6]
8. Yamada M, Kimura Y, Ishiyama D, Otobe Y, Suzuki M, Koyama S, et al. Recovery of Physical Activity among Older Japanese Adults Since the First Wave of the COVID-19 Pandemic. J Nutr Health Aging 2020 Sep 08;24(9):1036-1037. [doi: 10.1007/s12603-020-1516-z]
9. Bates LC, Zieff G, Stanford K, Moore JB, Kerr ZY, Hanson ED, et al. COVID-19 Impact on Behaviors across the 24-Hour Day in Children and Adolescents: Physical Activity, Sedentary Behavior, and Sleep. Children (Basel) 2020 Sep 16;7(9):138 [FREE Full text] [doi: 10.3390/children7090138] [Medline: 32947805]
10. Zenic N, Taiar R, Gilic B, Blazevic M, Maric D, Pojskic H, et al. Levels and Changes of Physical Activity in Adolescents during the COVID-19 Pandemic: Contextualizing Urban vs. Rural Living Environment. Applied Sciences 2020 Jun 09;10(11):3997. [doi: 10.3390/app10113997]
11. Yamada Y, Yoshida T, Nakagata T, Nanri H, Miyachi M. Letter to the Editor: Age, Sex, and Regional Differences in the Effect of COVID-19 Pandemic on Objective Physical Activity in Japan: A 2-Year Nationwide Longitudinal Study. J Nutr Health Aging 2021 Jul 30;25(8):1032-1033 [FREE Full text] [doi: 10.1007/s12603-021-1662-y] [Medline: 34545926]
12. Tison GH, Avram R, Kuhar P, Abreau S, Marcus GM, Pletcher MJ, et al. Worldwide Effect of COVID-19 on Physical Activity: A Descriptive Study. Ann Intern Med 2020 Dec 03;173(9):767-770 [FREE Full text] [doi: 10.7326/M20-2665] [Medline: 32598162]
13. Koohsari MJ, Nakaya T, McCormack GR, Shibata A, Ishii K, Oka K. Changes in Workers' Sedentary and Physical Activity Behaviors in Response to the COVID-19 Pandemic and Their Relationships With Fatigue: Longitudinal Online Study. JMIR Public Health Surveill 2021 Mar 26;7(3):e26293 [FREE Full text] [doi: 10.2196/26293] [Medline: 33727211]
14. Wang Y, Zhang Y, Bennell K, White DK, Wei J, Wu Z, et al. Physical Distancing Measures and Walking Activity in Middle-aged and Older Residents in Changsha, China, During the COVID-19 Epidemic Period: Longitudinal Observational Study. J Med Internet Res 2020 Oct 26;22(10):e21632 [FREE Full text] [doi: 10.2196/21632] [Medline: 33027035]
15. He M, Xian Y, Lv X, He J, Ren Y. Changes in Body Weight, Physical Activity, and Lifestyle During the Semi-lockdown Period After the Outbreak of COVID-19 in China: An Online Survey. Disaster Med Public Health Prep 2021 May;15(2):e23-e28 [FREE Full text] [doi: 10.1017/dmp.2020.237] [Medline: 32660669]
16. Stockwell S, Trott M, Tully M, Shin J, Barnett Y, Butler L, et al. Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: a systematic review. BMJ Open Sport Exerc Med 2021;7(1):e000960 [FREE Full text] [doi: 10.1136/bmjsem-2020-000960] [Medline: 34192010]
17. Meh K, Jurak G, Sorić M, Rocha P, Sember V. Validity and Reliability of IPAQ-SF and GPAQ for Assessing Sedentary Behaviour in Adults in the European Union: A Systematic Review and Meta-Analysis. Int J Environ Res Public Health 2021 May 26;18(9):4602 [FREE Full text] [doi: 10.3390/ijerph18094602] [Medline: 33926123]
18. Sasai H, Nakata Y, Murakami H, Kawakami R, Nakae S, Tanaka S, et al. Simultaneous Validation of Seven Physical Activity Questionnaires Used in Japanese Cohorts for Estimating Energy Expenditure: A Doubly Labeled Water Study. J Epidemiol 2018 Oct 05;28(10):437-442 [FREE Full text] [doi: 10.2188/jea.JE20170129] [Medline: 29709888]
19. Murakami H, Kawakami R, Nakae S, Nakata Y, Ishikawa-Takata K, Tanaka S, et al. Accuracy of Wearable Devices for Estimating Total Energy Expenditure: Comparison With Metabolic Chamber and Doubly Labeled Water Method. JAMA Intern Med 2016 May 01;176(5):702-703. [doi: 10.1001/jamainternmed.2016.0152] [Medline: 26999758]
20. Murakami H, Kawakami R, Nakae S, Yamada Y, Nakata Y, Ohkawara K, et al. Accuracy of 12 Wearable Devices for Estimating Physical Activity Energy Expenditure Using a Metabolic Chamber and the Doubly Labeled Water Method: Validation Study. JMIR Mhealth Uhealth 2019 Aug 02;7(8):e13938 [FREE Full text] [doi: 10.2196/13938] [Medline: 31376273]
21. Internet research (Internet-based research, web surveys). Cross Marketing. 2022. URL: https://www.cross-m.co.jp/en/service/netresearch/ [accessed 2022-12-05]
22. Fourth National Comprehensive Development Plan. Japanese National Land Agency. 1989. URL: https://www.mlit.go.jp/common/001135927.pdf [accessed 2022-12-05]
23. Koebnick C, Wagner K, Thielecke F, Moeseneder J, Hoehne A, Franke A, et al. Validation of a simplified physical activity record by doubly labeled water technique. Int J Obes (Lond) 2005 Mar 11;29(3):302-309. [doi: 10.1038/sj.ijo.0802882] [Medline: 15672111]

24. Yamada Y, Hashii-Arishima Y, Yokoyama K, Itoi A, Adachi T, Kimura M. Validity of a triaxial accelerometer and simplified physical activity record in older adults aged 64-96 years: a doubly labeled water study. Eur J Appl Physiol 2018 Oct;118(10):2133-2146. [doi: 10.1007/s00421-018-3944-6] [Medline: 30019086]
25. Yamada Y, Noriyasu R, Yokoyama K, Osaki T, Adachi T, Itoi A, et al. Association between lifestyle and physical activity level in the elderly: a study using doubly labeled water and simplified physical activity record. Eur J Appl Physiol 2013 Oct;113(10):2461-2471. [doi: 10.1007/s00421-013-2682-z] [Medline: 23801238]
26. Namba H, Yamaguchi Y, Yamada Y, Tokushima S, Hatamoto Y, Sagayama H, et al. Validation of Web-based physical activity measurement systems using doubly labeled water. J Med Internet Res 2012 Oct 25;14(5):e123 [FREE Full text] [doi: 10.2196/jmir.2253] [Medline: 23010345]
27. Namba H, Kurosaka Y, Minato K, Yamada Y, Kimura M. Validation of web-based physical activity measurement system using a tri-axial accelerometer (Japanese, with English abstract). Research in Exercise Epidemiology 2015;17(1):19-28.
28. Namba H, Yamada Y, Ishida M, Takase H, Kimura M. Use of a Web-based physical activity record system to analyze behavior in a large population: cross-sectional study. J Med Internet Res 2015 Mar 19;17(3):e74 [FREE Full text] [doi: 10.2196/jmir.3923] [Medline: 25794109]
29. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011 Aug;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
30. Watanabe D, Yoshida T, Watanabe Y, Yamada Y, Kimura M, Group KS. Objectively Measured Daily Step Counts and Prevalence of Frailty in 3,616 Older Adults. J Am Geriatr Soc 2020 Oct;68(10):2310-2318 [FREE Full text] [doi: 10.1111/jgs.16655] [Medline: 33269469]
31. Matsumura Y, Yamamoto M, Kitado T, Nakamura H, Kidera K, Fujimoto S. High-accuracy physical activity monitor utilizing three-axis accelerometer. National Technical Report 2008;56(2):60-66.
32. Yamada Y, Yokoyama K, Noriyasu R, Osaki T, Adachi T, Itoi A, et al. Light-intensity activities are important for estimating physical activity energy expenditure using uniaxial and triaxial accelerometers. Eur J Appl Physiol 2009 Jan 21;105(1):141-152. [doi: 10.1007/s00421-008-0883-7] [Medline: 18853176]
33. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
34. Furukawa TA, Kessler RC, Slade T, Andrews G. The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. Psychol Med 2003 Mar;33(2):357-362. [doi: 10.1017/s0033291702006700] [Medline: 12622315]
35. Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, et al. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res 2008;17(3):152-158 [FREE Full text] [doi: 10.1002/mpr.257] [Medline: 18763695]
36. Yoneoka D, Tanoue Y, Kawashima T, Nomura S, Shi S, Eguchi A, et al. Large-scale epidemiological monitoring of the COVID-19 epidemic in Tokyo. Lancet Reg Health West Pac 2020 Oct;3:100016 [FREE Full text] [doi: 10.1016/j.lanwpc.2020.100016] [Medline: 34173599]
37. Nagata S, Nakaya T, Adachi Y, Inamori T, Nakamura K, Arima D, et al. Mobility Change and COVID-19 in Japan: Mobile Data Analysis of Locations of Infection. J Epidemiol 2021 Jul 05;31(6):387-391 [FREE Full text] [doi: 10.2188/jea.JE20200625] [Medline: 33814508]
38. Moore SA, Faulkner G, Rhodes RE, Brussoni M, Chulak-Bozzer T, Ferguson LJ, et al. Impact of the COVID-19 virus outbreak on movement and play behaviours of Canadian children and youth: a national survey. Int J Behav Nutr Phys Act 2020 Jul 06;17(1):85 [FREE Full text] [doi: 10.1186/s12966-020-00987-8] [Medline: 32631350]
39. López-Bueno R, Calatayud J, Andersen LL, Balsalobre-Fernández C, Casaña J, Casajús JA, et al. Immediate Impact of the COVID-19 Confinement on Physical Activity Levels in Spanish Adults. Sustainability 2020 Jul 16;12(14):5708. [doi: 10.3390/su12145708]
40. Shut-Eye Deficit: OECD Survey Reveals Japan Most Lacking in Sleep. Nippon. 2019. URL: https://www.nippon.com/en/japan-data/h00424/shut-eye-deficit-oecd-survey-reveals-japan-most-lacking-in-sleep.html [accessed 2022-01-03]
41. OECD Better Life Index. URL: https://www.oecdbetterlifeindex.org/countries/japan/ [accessed 2022-03-01]
42. Hu Z, Lin X, Chiwanda Kaminga A, Xu H. Impact of the COVID-19 Epidemic on Lifestyle Behaviors and Their Association With Subjective Well-Being Among the General Population in Mainland China: Cross-Sectional Study. J Med Internet Res 2020 Aug 25;22(8):e21176 [FREE Full text] [doi: 10.2196/21176] [Medline: 32759103]

Abbreviations

AEE: activity energy expenditure
DLW: doubly labeled water
GPAQ: Global Physical Activity Questionnaire
IPAQ: International Physical Activity Questionnaire

MET: metabolic equivalent of task
MVPA: moderate-to-vigorous physical activity
OECD: Organization for Economic Co-operation and Development
PA: physical activity
SB: sedentary behavior
sPAR: simplified physical activity record
ST: sleeping time

Edited by T Sanchez, A Mavragani; submitted 31.05.22; peer-reviewed by C Zeng, J Moore, YL Chen, W Liang; comments to author 25.07.22; revised version received 06.12.22; accepted 12.01.23; published 20.02.23.

Please cite as:
Yamada Y, Namba H, Date H, Kitayama S, Nakayama Y, Kimura M, Fujita H, Miyachi M
Regional Difference in the Impact of COVID-19 Pandemic on Domain-Specific Physical Activity, Sedentary Behavior, Sleeping Time, and Step Count: Web-Based Cross-sectional Nationwide Survey and Accelerometer-Based Observational Study
JMIR Public Health Surveill 2023;9:e39992
URL: https://publichealth.jmir.org/2023/1/e39992
doi:10.2196/39992
PMID:36634262

Original Paper

Decisions and Decisional Needs of Canadians From all Provinces and Territories During the COVID-19 Pandemic: Population-Based Cross-sectional Surveys

Dawn Stacey[1,2], RN, PhD; Claire Ludwig[3], PhD; Patrick Archambault[4,5,6], MD, MSc; Maureen Smith[3,7], MEd; Monica Taljaard[2,8], PhD; Meg Carley[2], BSc; Karine Plourde[4], PhD; Laura Boland[2], PhD; Amédé Gogovor[6], DVM, MSc, PhD; Ian Graham[2,8], PhD; Daniel Kobewka[2,9,10], MSc, MD; Robert K D McLean[11], MSc, PhD; Michelle L A Nelson[12,13], PhD; Brandi Vanderspank-Wright[1,2], PhD, RN; France Légaré[4,6], MD, MSc, PhD, CCFP

[1]School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada
[2]Centre for Implementation Research, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[3]Patient Partner, Ottawa, ON, Canada
[4]Department of Family Medicine and Emergency Medicine, Université Laval, Quebec City, QC, Canada
[5]Centre de recherche intégrée pour un système apprenant en santé et services sociaux, Université Laval, Lévis, QC, Canada
[6]VITAM - Centre de recherche en santé durable, Université Laval, Quebec City, QC, Canada
[7]Cochrane Consumer Network Executive, Ottawa, ON, Canada
[8]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
[9]Department of Medicine, University of Ottawa, Ottawa, ON, Canada
[10]General Internal Medicine, The Ottawa Hospital, Ottawa, ON, Canada
[11]International Development Research Centre, Ottawa, ON, Canada
[12]Sinai Health System, University of Toronto, Toronto, ON, Canada
[13]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Dawn Stacey, RN, PhD
School of Nursing
Faculty of Health Sciences
University of Ottawa
451 Smyth Road
Ottawa, ON, K1H 8M5
Canada
Phone: 1 613 562 5800 ext 8413
Email: dstacey@uottawa.ca

Abstract

Background: Never before COVID-19 had Canadians faced making health-related decisions in a context of significant uncertainty. However, little is known about which type of decisions and the types of difficulties that they faced.

Objective: We sought to identify the health-related decisions and decisional needs of Canadians.

Methods: Our study was codesigned by researchers and knowledge users (eg, patients, clinicians). Informed by the CHERRIES (the Checklist for Reporting Results of Internet E-Surveys) reporting guideline, we conducted 2 online surveys of random samples drawn from the Leger consumer panel of 400,000 Canadians. Eligible participants were adults (≥18 years) who received or were receiving any health services in the past 12 months for themselves (adults) or for their child (parent) or senior with cognitive impairment (caregiver). We assessed decisions and decisional needs using questions informed by the Ottawa Decision Support Framework, including decisional conflict and decision regret using the Decision Conflict Scale (DCS) and the Decision Regret Scale (DRS), respectively. Descriptive statistics were conducted for adults who had decided for themselves or on behalf of someone else. Significant decisional conflict (SDC) was defined as a total DCS score of >37.5 out of 100, and significant decision regret was defined as a total DRS score of >25 out of 100.

Results: From May 18 to June 4, 2021, 14,459 adults and 6542 parents/caregivers were invited to participate. The invitation view rate was 15.5% (2236/14,459) and 28.3% (1850/6542); participation rate, 69.3% (1549/2236) and 28.7% (531/1850); and completion rate, 97.3% (1507/1549) and 95.1% (505/531), respectively. The survey was completed by 1454 (97.3%) adults and 438 (95.1%) parents/caregivers in English (1598/1892, 84.5%) or French (294/1892, 15.5%). Respondents from all 10 Canadian provinces and the northern territories represented a range of ages, education levels, civil statuses, ethnicities, and annual household income. Of 1892 respondents, 541 (28.6%) self-identified as members of marginalized groups. The most frequent decisions were (adults vs parents/caregivers) as follows: COVID-19 vaccination (490/1454, 33.7%, vs 87/438, 19.9%), managing a health condition (253/1454, 17.4%, vs 47/438, 10.7%), other COVID-19 decisions (158/1454, 10.9%, vs 85/438, 19.4%), mental health care (128/1454, 8.8%, vs 27/438, 6.2%), and medication treatments (115/1454, 7.9%, vs 23/438, 5.3%). Caregivers also reported decisions about moving family members to/from nursing or retirement homes (48/438, 11.0%). Adults (323/1454, 22.2%) and parents/caregivers (95/438, 21.7%) had SDC. Factors making decisions difficult were worrying about choosing the wrong option (557/1454, 38.3%, vs 184/438, 42.0%), worrying about getting COVID-19 (506/1454, 34.8%, vs 173/438, 39.5%), public health restrictions (427/1454, 29.4%, vs 158/438, 36.1%), information overload (300/1454, 20.6%, vs 77/438, 17.6%), difficulty separating misinformation from scientific evidence (297/1454, 20.4%, vs 77/438, 17.6%), and difficulty discussing decisions with clinicians (224/1454, 15.4%, vs 51/438, 11.6%). For 1318 (90.6%) adults and 366 (83.6%) parents/caregivers who had decided, 353 (26.8%) and 125 (34.2%) had significant decision regret, respectively. In addition, 1028 (50%) respondents made their decision alone without considering the opinions of clinicians.

Conclusions: During COVID-19, Canadians who responded to the survey faced several new health-related decisions. Many reported unmet decision-making needs, resulting in SDC and decision regret. Interventions can be designed to address their decisional needs and support patients facing new health-related decisions.

(JMIR Public Health Surveill 2023;9:e43652) doi:10.2196/43652

KEYWORDS

health care decisions; decisional conflict; decision regret; shared decision-making; COVID-19; older adults; caregivers; parents; public health decision; health care; health outcome; pandemic preparedness; public health policy

Introduction

During the COVID-19 pandemic, never before had Canadians made health-related decisions in the context of so much uncertainty. Health care decisions were complex, with limited, rapidly changing evidence and evolving public health directives [1,2]. Concurrently, health services shifted from in-person to virtual delivery and emergency department visits increased after the first wave [3-5]. These circumstances are perfectly fit for shared decision-making, a process between patients and clinicians that relies on the best evidence available and what matters most to patients. However, due to unique challenges with communicating risk, lack of nonverbal communication, and less meaningful patient involvement experienced during the pandemic, it may have been possible that Canadians were inadequately involved in health-related decisions [2,6,7]. When patients are inadequately involved in health decisions, there are more harms from choosing ineffective options, health care system waste, poor patient experiences, more litigation, and higher inequities [8,9]. This is why identifying Canadians decision-making needs was of uttermost importance in order to be able to provide person-centered care, achieve improved health outcomes, and inform future pandemic preparedness.

Decisional needs are deficits that can adversely affect the quality of decisions [10]. A quality decision is informed with the best-available evidence and based on patients' values for features and outcomes of options. A previous systematic review of 45 decisional needs assessment studies included 2 population surveys that identified 75 decisions (including 16 social decisions) and 43 clinical studies that focused on 29 specific health decisions [10]. Common decisional needs included decisional conflict, inadequate knowledge of the options (including benefits, harms), unclear values, and limited support and resources [10,11]. Decisional conflict refers to "uncertainty about a course of action when choice among competing options involves risk, regret, or challenge to personal life values" [12]. Those who experience more difficultly with decision-making are living with a serious or chronic illness, are immunocompromised, belong to linguistic minorities, have lower education, are passive in decision-making, or make a decision on behalf of someone else (parents and caregivers) [13-19]. None of the 45 studies from 7 countries were focused on decision-making needs during pandemics [10]. A recent cross-sectional study of 4905 Canadians aged 18-40 years reported that factors associated with vaccine hesitancy are negative attitudes toward vaccines in general; COVID-19 conspiracy theory beliefs; distrust of the government; and a low income, low education, or unemployment [20]. Canada was no exception, as little was known about its health-related decision-making experience during the COVID-19 pandemic. Therefore, we sought to determine the type of decisions and decisional needs of Canadians during the first year of the COVID-19 pandemic.

Methods

Study Design

Our team of researchers and knowledge users (patients, clinicians) conducted 2 population-based cross-sectional surveys using Leger's consumer panel. Having patients and clinicians as equal partners on the team was done to ensure the study yielded relevant findings [21]. Patients were on the executive research team coleading the study (author CL who is immunocompromised), and author MS stimulated the research topic. Both were engaged in all study aspects. Together, we identified the research objectives, outcomes, procedures, and deliverables for the funded proposal. Next, they were involved in guiding the study. We used the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guideline [22].

Setting

The survey was conducted in Canada during the third wave of COVID-19 [23]. During the first wave of COVID-19 in January-June 2020, 80% of the COVID-19–related deaths occurred in long-term-care homes, and governments issued stay-at-home orders and travel restrictions [24]. Masks became mandated for use in indoor spaces in July 2020. The second wave of COVID-19 started in November 2020 and the third wave in March 2021. When the hospitals became overwhelmed during the third wave, with a high impact on intensive care units, governments reissued a stay-at-home order [23]. COVID-19 vaccines were initially approved by Health Canada in December 2020, with priority for health care workers, Indigenous peoples, the elderly living in group settings, and others at higher risk (eg, people with cancer, with organ transplants, or undergoing dialysis). Vaccination gradually opened to adults in spring 2021 and to children (ages 5-12 years) in fall 2021 [25]. Although the AstraZeneca vaccine was approved for use on February 26, 2021, the rate of vaccine-induced immune thrombotic thrombocytopenia was of concern (1 in 60,000; April 2021). Governments continued to recommend the AstraZeneca vaccine until June 2021 because Canada was in the third wave of infections and there was insufficient supply of messenger RNA (mRNA) vaccines [23].

Recruitment

We recruited 2 groups of participants: (1) adults aged 18 years or older who received or were receiving any health services in the past 12 months for themselves (labeled "adults") and (2) adults aged 18 years or older who were responsible for children or seniors aged >65 years with cognitive impairment who received or were receiving any health services in the past 12 months (labeled "parents/caregivers"). Leger recruited participants through its consumer panel titled "Leger Opinion (LEO) Panel," which has about 400,000 Canadians across 10 provinces and 3 territories. The panel includes data on age, gender, and region that can be used for sampling and quota management. At the onset, specific quotas by age, gender, and region were set based on the Canadian population data published by Statistics Canada [26]. Leger monitored throughout the recruitment phase, and sampling was adjusted to ensure the data collected were representative. For example, if a certain cohort was underrepresented, sampling was adjusted to recruit more respondents falling within that cohort. There was no weighting applied to the data after being collected. Respondents received a personalized email containing a unique URL link to the survey. The email invite said, "LEO wants to hear from you!" and it did not state the exact survey topic.

Survey Instrument

Adapted from previous surveys in Canada [13,14,27], questions were based on the Ottawa Decision Support Framework [11]. Questions from instruments with good reliability and validity included the Decisional Conflict Scale (DCS) [12], the Decision Regret Scale (DRS) [28], and Strull's roles in decision-making [13]. The survey enquired about health-related decisions adults and parents/caregivers faced during the past 12 months, factors influencing decision-making, trusted information sources, and sociodemographics according to the PROGRESS (place of residence, race/ethnicity/culture/language, occupation, gender/sex, religion, education, socioeconomic status, and social capita) framework [29,30]. To collect data on a broader range of participants' characteristics that may stratify health opportunities and outcomes due to discrimination [29,30], we also asked respondents to self-identify if they had lived experience as a member of a marginalized group defined as disabled or caregiver of a person with a disability, gender diverse (eg, agender, nonbinary, transgender, cisgender), intersex, LGBTQ+ (lesbian, gay, bisexual, pansexual, transgender, queer, two-spirited, questioning), Indigenous, racialized (eg, person of color), neurodivergent (eg, attention deficit hyperactivity disorder [ADHD], autism, dyslexia), or a marginalized group not listed. Respondents were initially asked to list all the decisions made in the past year and then pick 1 difficult decision for the remaining questions. A difficult decision was defined as having more than 1 option and no clear best option. The online survey, available in English or French, screened eligibility and included 32 questions (Multimedia Appendix 1). Two random test questions were used to ensure respondents were not trying to speed through the survey, and the survey stopped for those who clicked a wrong response on both.

Respondents received 1 question per screen and pressed "continue" to proceed. They could not return to previous screens to change responses. To avoid missing responses, no questions could be skipped. Up to 3 reminders were sent to those who started but did not complete the survey. To avoid multiple entries, respondents were assigned a unique identifier linked to their LEO account; if they tried to complete the survey again, they received an error message.

The English and French versions were pretested by members of our team, including patient partners. On May 18, 2021, we field-tested the surveys with 139 adults and 26 parents/caregivers to monitor completion time (adults: ~10 minutes; parents/caregivers: ~9 minutes) and potential problems with surveys, including predefined skip logic. Given no changes were made to the surveys based on field testing, the collected data were retained.

Sample Size

Based on our previous research [13,31,32], sample sizes of 1500 adults and 500 parents/caregivers were adequate to estimate proportions with significant decisional conflict (SDC) associated with decision delay. SDC is defined as a total score of >37.5 out of 100 on the DCS based on 253 studies [33] using a 2-sided 95% CI with a margin of error of +1.80% or +3.85% when the estimated proportions are 0.15 or 0.26, respectively [34,35]. This is the most severe cutoff; sensitivity analyses were performed with a more relaxed cutoff of 25, also sometimes seen in the literature (data not shown), but we privileged a more severe cutoff. This means it could have been showing more people with SDC. To account for an anticipated 10% response rate, target sample sizes of adults and parents/caregivers were 15,000 and 5000, respectively.

Statistical Analysis

We classified respondents into adults who made a health-related decision for themselves (adults) or made a decision on behalf of a child or a senior with cognitive impairment (parents/caregivers). Next, we used descriptive statistics to identify decisions and described decisional needs using the Ottawa Decision Support Framework for each group (adults and parents/caregivers) [10]. We classified respondents as having SDC [36], and for those who had made the decision, we classified respondents as having significant decision regret based on a cutoff of >25 out of 100 on the DRS [37].

Ethical Considerations

The University of Ottawa research ethics board approved our study (H-03-21-6752). Invitees provided study consent at the start of the online survey and only initiated the survey questions after clicking the link consenting to participate. Respondents voluntarily answered questions and were guaranteed confidentiality. Leger offered an incentive of 2000 points (equivalent to CA $1.60, or US $1.18) to complete the survey.

Results

Respondent Details

From May 18 to June 4, 2021, 14,459 adults and 6542 parents/caregivers were invited to participate (Figure 1). The view rate (unique survey visitors/unique survey invitees) was 15.5% (2236/14,459) and 28.3% (1850/6542), respectively. The participation rate (unique visitors who consented to participate/unique survey visitors) was 69.3% (1549/2236) and 28.7% (531/1850). The completion rate (users who completed the survey/users who consented to participate) was 97.3% (1507/1549) and 95.1% (505/531), respectively. Among those who completed the surveys, we removed respondents if they did not identify a difficult health decision. After reading open text describing options, 36 were moved to the adult database and 16 were moved to the parent/caregiver database.

There were 1454 adults and 438 parents/caregivers included in the analysis. The survey was completed in English (1598/1892, 84.5%) or French (294/1892, 15.5%). Respondents represented a range of ages, education levels, civil statuses, ethnicities, and annual household income (Table 1). Of 1892 respondents, 541 (28.6%) self-identified as members of marginalized groups.

Figure 1. Cross-sectional survey recruitment of Canadians during the COVID-19 pandemic.

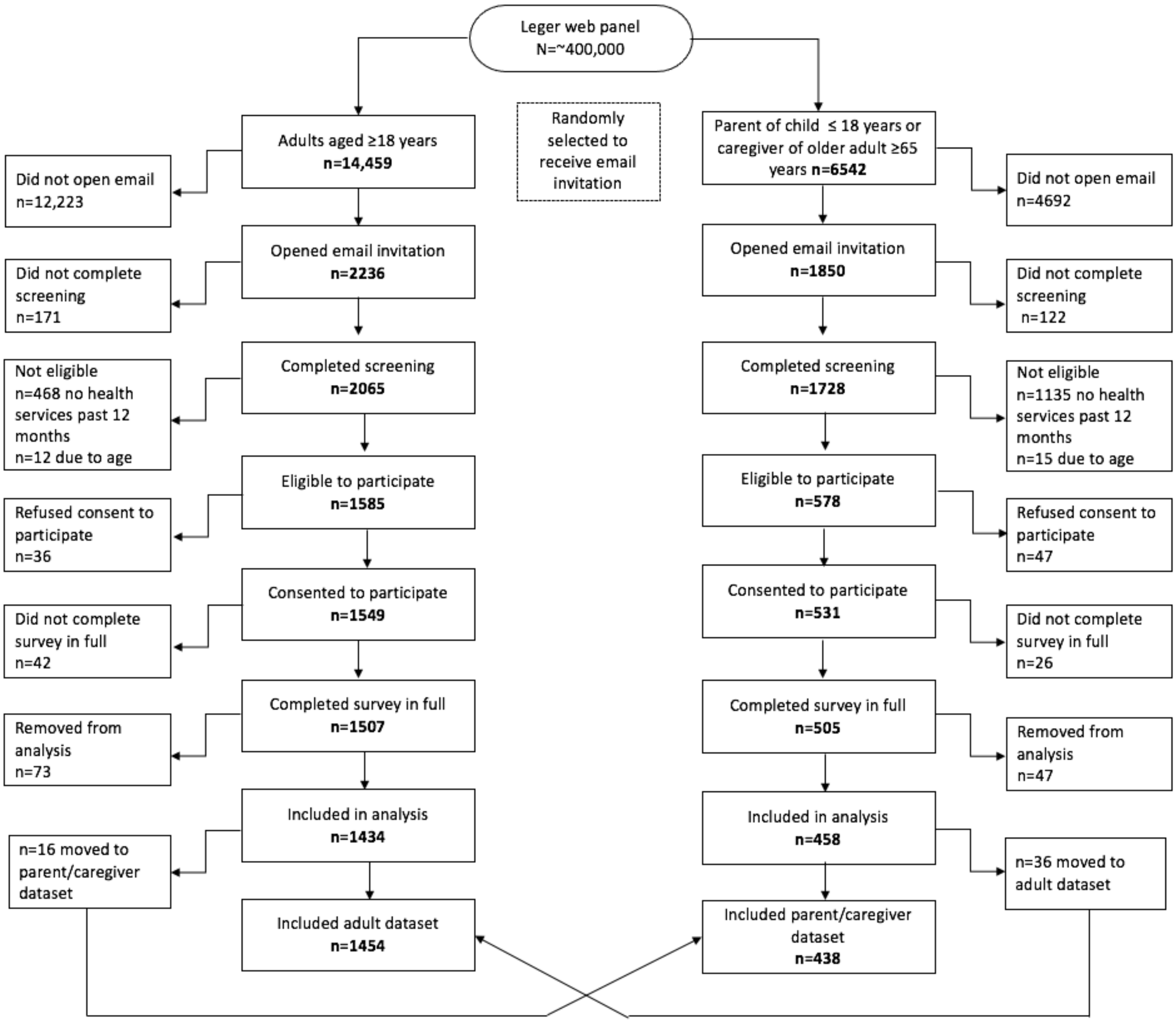


Table 1. Demographic characteristics of Canadian cross-sectional survey respondents during the COVID-19 pandemic.

Variable	Adults (N=1454)	Parents/caregivers (N=438)
Age (years), n (%)		
18-29	264 (18.2)	46 (10.5)
30-39	238 (16.4)	117 (26.7)
40-49	259 (17.8)	126 (28.8)
50-59	248 (17.1)	92 (21.0)
60-69	250 (17.2)	41 (9.4)
≥70	195 (13.4)	16 (3.7)
Provinces/territories, n (%)		
Ontario	558 (38.4)	165 (37.7)
Quebec	264 (18.2)	84 (19.2)
Prairie Provinces	269 (18.5)	87 (19.9)
British Columbia	241 (16.6)	69 (15.8)
Atlantic Provinces	115 (7.9)	30 (6.8)
Northern Territories	7 (0.5)	3 (0.7)
Geographical area, n (%)		
Urban	1269 (87.3)	393 (89.7)
Rural	185 (12.7)	45 (10.3)
Sex, n (%)		
Female	789 (54.3)	219 (50.0)
Male	664 (45.7)	218 (49.8)
Prefer not to say	1 (0.1)	1 (0.2)
Gender, n (%)		
Woman	781 (53.7)	219 (50.0)
Man	662 (45.5)	218 (49.8)
Other[a] or prefer not to say	11 (0.8)	1 (0.2)
Language first learned and still understood[b], n (%)		
English	1023 (70.4)	290 (66.2)
French	270 (18.6)	92 (21.0)
Mandarin or Cantonese	46 (3.2)	13 (3.0)
Other[c]	124 (8.5)	43 (9.8)
Highest level of education, n (%)		
High school or less	303 (20.9)	67 (15.3)
Certificate/diploma below bachelor’s level	528 (36.3)	157 (35.8)
Bachelor’s degree	372 (25.6)	136 (31.1)
University degree above bachelor’s level	249 (17.1)	75 (17.1)
Prefer not to say	2 (0.1)	3 (0.7)
Cultural/ethnic background[b,d], n (%)		
White	1149 (79.0)	314 (71.7)
Asian	206 (14.2)	84 (19.2)
North American Indigenous	53 (3.6)	17 (3.9)
Black	22 (1.5)	17 (3.9)

Variable	Adults (N=1454)	Parents/caregivers (N=438)
Latin American	20 (1.4)	10 (2.3)
European	25 (1.7)	4 (0.9)
Other	26 (1.8)	7 (1.6)
Prefer not to say	17 (1.2)	7 (1.6)
Lived experience as member of a marginalized group[b,e], n (%)		
Disabled or caregiver of a person with a disability	142 (9.8)	68 (15.5)
LGBTQ+[f]	133 (9.1)	32 (7.3)
Gender diverse: eg, agender, nonbinary, trans-/cis-gender	26 (1.8)	8 (1.8)
Racialized as person of color	109 (7.5)	44 (10.0)
Neurodivergent: eg, ADHD[g], autistic, dyslexic	60 (4.1)	13 (3.0)
Indigenous	29 (2.0)	13 (3.0)
Other	26 (1.8)	5 (1.1)
None of the above	1038 (71.4)	283 (64.6)
Prefer not to say	25 (1.7)	5 (1.1)
Civil status, n (%)		
Married or common law	838 (57.6)	321 (73.3)
Single, divorced, or separated	570 (39.2)	108 (24.7)
Widowed	38 (2.6)	7 (1.6)
Prefer not to say	8 (0.6)	2 (0.5)
Number in household, mean (SD; range)	2.6 (1.3; 1-11)	3.3 (1.3; 1-8)
Annual household income (CA $/US $)[h], n (%)		
<50,000/<36761.75	325 (22.4)	56 (12.8)
50,000-99,999/36761.75-73522.76	554 (38.1)	214 (48.9)
≥100,000/≥73523.50	450 (30.9)	146 (33.3)
Prefer not to say	125 (8.6)	22 (5.0)
Quality of life[i], mean (SD)	4.8 out of 7 (1.4)	4.7 out of 7 (1.4)

[a] Examples of other genders specified included bisexual, demiboy, gender fluid, transgender, and 2-spirited.

[b] Respondents sometimes specified more than 1 response option.

[c] Other languages included Arabic, Armenian, Bengali, Croatian, Czech, Danish, Dutch, Filipino, Finnish, Fukien, Friulan, German, Greek, Haitian Creole, Hakka, Hindi, Irish, Italian, Japanese, Kannada, Konkani, Korean, Laotian, Latvian, Macedonian, Malayalam, Mongolian, Norwegian, Pashto, Polish, Portuguese, Punjabi, Romanian, Russian, Serbian, Slovak, Spanish, Tagalog, Taiwanese, Tamil, Teochew, Ukrainian, Urdu, Vietnamese, Yoruba, and Zulu.

[d] Other cultural/ethnic backgrounds included Acadian, African, Canadian, Caribbean, Fijian, Guyanese, Jewish, mixed, and West Indian.

[e] Other reasons respondents indicated having lived experience as a member of a marginalized group included intersex, their gender (eg, "a woman"), occupation, social capital, socioeconomic status, religion, philosophical or political beliefs, physical or mental health condition, and not specified.

[f] LGBTQ+: lesbian, gay, bisexual, pansexual, transgender, queer, two-spirited, questioning.

[g] ADHD: attention deficit hyperactivity disorder.

[h] CA $1=US $0.74.

[i] Overall quality of life was measured on a 7-point scale: from 1 (life is very distressing, and it is difficult to imagine how it could get much worse) to 7 (life is great, and it is difficult to imagine how it could get much better); 4, life is so-so, neither good nor bad.

Decisions

Health-related decisions, in order of frequency, were about COVID-19 vaccination, managing a health condition, social COVID-19 decisions (masking, limiting contacts), mental health care or addiction treatment, medication, surgery, pain management, health care for COVID-19 (testing, seeking care for symptoms), stay at home or move to an assisted care facility (eg, nursing home), pregnancy, birth control, staying safe at home (eg, adapt or retrofit home), and end-of-life care (Table 2). Decisions were similar when respondents were asked to pick a single decision described as more difficult, with the exception of health care for COVID-19, which a few rated as difficult. Other single decisions for caregivers were about moving a family member to/from a retirement or nursing home and stopping them from driving a car. Decisions (adults vs parents/caregivers) were within the past month (318/1318, 24.1%, vs 70/366, 19.1%), 1-6 months (524/1318, 39.8%, vs 134/366, 36.6%), 6-12 months (307/1318, 23.3%, vs 97/366, 26.5%), or no response (169/1318, 12.8%, vs 65/366, 17.8%).

Table 2. Health-related decisions of Canadian cross-sectional survey respondents during the COVID-19 pandemic.

Decisions	Adults, n/N (%)		Parents/caregivers, n/N (%)	
	All decisions	One difficult decision[a]	All decisions	One difficult decision[a]
COVID-19 vaccination	1180/1454 (81.2)	490/1454 (33.7)	235/438 (53.7)	87/438 (19.9)
COVID-19 social decisions: eg, masking, limiting contacts	975/1454 (67.1)	158/1454 (10.9)	243/438 (55.5)	85/438 (19.4)
Health care for COVID-19	350/1454 (24.1)	39/1454 (2.7)	114/438 (26.0)	29/438 (6.6)
Testing	N/A[b]	27/39 (69.2)	N/A	18/29 (62.1)
Seeking care of COVID-19 symptoms	N/A	9/39 (23.1)	N/A	2/29 (6.9)
Participating in clinical trials	N/A	1/39 (2.6)	N/A	0
Not specified	N/A	2/39 (5.1)	N/A	9/29 (31.0)
COVID-19: delaying medical treatment/visits	N/A	120/1454 (8.3)	N/A	29 (6.6)
Health condition	N/A	87/120 (72.5)	N/A	15/29 (51.7)
Surgery	N/A	20/120 (16.7)	N/A	9/29 (31.0)
Mental health care	N/A	13/120 (10.8)	N/A	3/29 (10.3)
Chemotherapy	N/A	0	N/A	2/29 (6.9)
Pregnancy and childbirth	N/A	23/1454 (1.6)	N/A	N/A
Move temporarily from nursing or retirement home	13/1454 (0.9)	1/1454 (0.1)	30/438 (6.8)	6/438 (1.4)
Options to stay safe at home or move to have proper support and assistance	N/A	N/A	N/A	8/438 (1.8)
Managing a health condition	641/1454 (44.1)	166/1454 (11.4)	159/438 (36.3)	30/438 (6.8)
Delaying medical treatment/visits	N/A	92/166 (55.4)	N/A	7/30 (23.3)
Having tests	N/A	27/166 (16.3)	N/A	7/30 (23.3)
New treatments	N/A	14/166 (8.4)	N/A	4/30 (13.3)
Dental visit	N/A	26/166 (15.7)	N/A	1/30 (3.3)
Admission to hospital	N/A	1/166 (0.6)	N/A	1/30 (3.3)
Other	N/A	2/166 (1.2)	N/A	2/30 (6.7)
Not specified	N/A	4/166 (2.4)	N/A	8/30 (26.7)
Mental health care decisions				
Treatment	326/1454 (22.4)	115/1454 (7.9)	78/438 (17.8)	24/438 (5.5)
Addiction or overdose	20/1454 (1.4)	2/1454 (0.1)	26/438 (5.9)	2/438 (0.5)
Medication decisions	477/1454 (32.8)	115/1454 (7.9)	138/438 (31.5)	23/438 (5.3)
Antibiotics	N/A	1/115 (0.9)	N/A	1/23 (4.3)
Lower cholesterol	N/A	6/115 (5.2)	N/A	0
Control blood sugar	N/A	3/115 (2.6)	N/A	1/23 (4.3)
Prevent heart burn	N/A	4/115 (3.5)	N/A	0
Sleeping pills	N/A	10/115 (8.7)	N/A	1/23 (4.3)
Other	N/A	10/115 (8.7)	N/A	3/23 (13.0)
Not specified	N/A	81/115 (70.4)	N/A	17/23 (73.9)
Pain management	327/1454 (22.5)	88/1454 (6.1)	83/438 (18.9)	15/438 (3.4)
Surgery decisions	153/1454 (10.5)	69/1454 (4.7)	54/438 (12.3)	12/438 (2.7%)
Delay surgery	N/A	56/69 (81.2)	N/A	10/12 (83.3)
Joint replacement	N/A	7/69 (10.1)	N/A	0
Prostate cancer	N/A	2/69 (2.9)	N/A	0
Back surgery	N/A	4/69 (5.8)	N/A	0

Decisions	Adults, n/N (%)		Parents/caregivers, n/N (%)	
	All decisions	One difficult decision[a]	All decisions	One difficult decision[a]
Breast cancer	N/A	0	N/A	1/12 (8.3)
Not specified	N/A	0	N/A	1/12 (8.3)
Pregnancy or childbirth	101/1454 (6.9)	24/1454 (1.7)	7/438 (1.6)	N/A
Planning a pregnancy	N/A	17/24 (70.8)	N/A	N/A
Prenatal testing	N/A	1/24 (4.2)	N/A	N/A
Childbirth type or setting	N/A	3/24 (12.5)	N/A	N/A
Methods of feeding	N/A	2/24 (8.3)	N/A	N/A
Unplanned pregnancy	N/A	1/24 (4.2)	N/A	N/A
Birth control	162/1454 (11.1)	34/1454 (2.3)	4/438 (0.9)	N/A
End of life	14/1454 (1.0)	4/1454 (0.3)	28/438 (6.4)	13/438 (3.0)
Mechanical ventilator	N/A	1/4 (25.0)	N/A	3/13 (23.1)
Palliative care	N/A	0	N/A	2/13 (15.4)
Medical Assistance in Dying (MAiD)	N/A	2/4 (50.0)	N/A	1/13 (7.7)
Advanced care planning	N/A	0	N/A	1/13 (7.7)
Location of care	N/A	1/4 (25.0)	N/A	6/13 (46.2)
Other				
Stay home or move (eg, nursing home)	6/1454 (0.4)	N/A	92/438 (21.0)	42/438 (9.6)
Best option to stay safe at home	6/1454 (0.4)	N/A	103/438 (23.5)	20/438 (4.6)
Stop driving car	5/1454 (0.3)	2/1454 (0.1)	52/438 (11.9)	13/438 (3.0)
Participating in clinical trials	42/1454 (2.9)	1/1454 (0.1)	1/438 (0.2)	N/A
Smoking cessation	N/A	1/1454 (0.1)	N/A	N/A
Being more active/eating healthy	N/A	1/1454 (0.1)	N/A	N/A
Moving to another location	N/A	1/1454 (0.1)	N/A	N/A

[a] Identify 1 specific difficult health care decision that you faced or are facing.

[b] N/A: not applicable.

Decisional Needs

Of 1454 adults and 438 parents/caregivers, 323 (22.2%) and 95 (21.7%), respectively, had SDC (Tables 3 and 4). Decisions, in order of frequency, for adults with SDC were mental health care, managing a health condition, taking medications, pain management, and COVID-19 vaccination (Figure 2). Decisions for parents/caregivers with SDC were COVID-19 vaccination, managing a health condition, health care for COVID-19, and mental health care.

Table 3. Decisional needs and factors influencing decision-making of Canadian cross-sectional survey respondents during the COVID-19 pandemic (DCS[a]).

Variables	Adults (N=1454)	Parents/caregivers (N=438)
DCS score>37.5 out of 100	323 (22.2)	95 (21.7)
Total decisional conflict (DCS)[b]		
Mean (SD)	25.5 (17.3)	26.2 (16.9)
Median (Q1, Q3)[c]	25.0 (12.5, 36.3)	25.0 (14.1, 35.9)
DCS uncertain subscale		
Mean (SD)	31.8 (23.6)	32.2 (21.9)
Median (Q1, Q3)	25.0 (16.7, 50.0)	29.2 (16.7, 50.0)
DCS uninformed subscale		
Mean (SD)	23.5 (19.2)	23.0 (17.4)
Median (Q1, Q3)	25.0 (8.3, 33.3)	25.0 (8.3, 33.3)
DCS unclear values subscale		
Mean (SD)	23.4 (19.6)	24.1 (19.8)
Median (Q1, Q3)	25.0 (0.0, 33.3)	25.0 (8.3, 33.3)
DCS unsupported subscale		
Mean (SD)	26.9 (20.8)	28.8 (21.4)
Median (Q1, Q3)	25.0 (8.3, 41.7)	25.0 (16.7, 41.7)
DCS ineffective decision subscale		
Mean (SD)	22.9 (19.5)	23.7 (18.7)
Median (Q1, Q3)	25.0 (6.3, 31.3)	25.0 (6.3, 31.3)
Worried about choosing the wrong option, n (%)	557 (38.3)	184 (42.0)
Worried about getting COVID-19, n (%)	506 (34.8)	173 (39.5)
Public health restrictions due to COVID-19, n (%)	427 (29.4)	158 (36.1)
Overloaded with information, n (%)	300 (20.6)	77 (17.6)
Difficulty separating misinformation from scientific evidence, n (%)	297 (20.4)	77 (17.6)
Difficulty discussing the decision with important others (eg, spouse, family, friends), n (%)	192 (13.2)	89 (20.3)
Difficulty discussing the decision with clinicians, n (%)	224 (15.4)	51 (11.6)
No or limited access to information on the decision or options, n (%)	173 (11.9)	62 (14.2)
Difficulty in believing scientific evidence, n (%)	158 (10.9)	44 (10.0)
No skills or ability for making this type of decision, n (%)	91 (6.3)	40 (9.1)
Other (eg, unable to see the doctor in person to manage the health condition, side effects of the COVID-19 vaccine), n (%)	195 (13.4)	32 (7.3)
Considered the costs related to the options, n (%)	356 (24.5)	134 (30.6)

[a] DCS: Decisional Conflict Scale.

[b] Respondents sometimes specified more than 1 response option.

[c] Q1: quartile 1; Q3: quartile 3.

Table 4. Decisional needs and factors influencing decision-making of Canadian cross-sectional survey respondents during the COVID-19 pandemic (DRS[a]).

Variables	Adults who made a decision (N=1318)	Parents/caregivers who made a decision (N=366)
Decisional regret[b]		
Mean (SD)	18.8 (18.2)	21.3 (18.4)
Median (Q1, Q3)[c]	20.0 (0.0, 35.0)	25.0 (5.0, 40.0)
No decisional regret=0, n (%)	367 (27.8)	89 (24.3)
Low decisional regret=1 to ≤25, n (%)	598 (45.4)	152 (41.5)
Decisional regret>25, n (%)	353 (26.8)	125 (34.2)
Decision made alone, n (%)	843 (58.0)	185 (50.5)
Preferred option chosen, n (%)	996 (75.6)	289 (79.0)

[a] DRS: Decision Regret Scale.

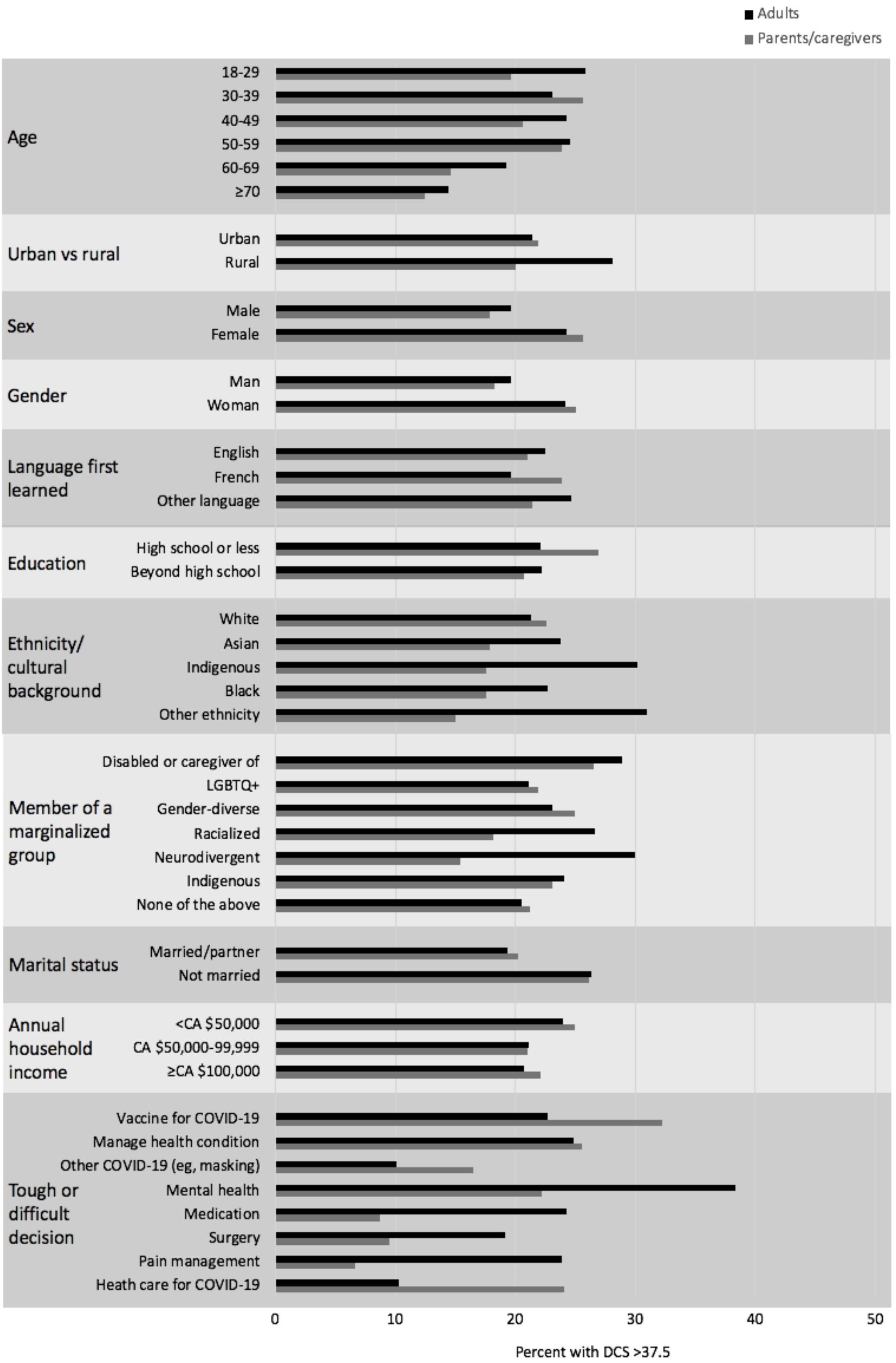
[b] Values were standardized out of 100.

[c] Q1: quartile 1; Q3: quartile 3.

Respondents experiencing SDC were more likely to feel uninformed (mean 44, SD 18.6, vs mean 18, SD 14.3, out of 100), have unclear values (mean 45, SD 20.3, vs mean 18, SD 14.6), feel they made an ineffective decision (mean 48, SD 16.1, vs mean 16, SD 13.3), and have decision regret (mean 41, SD 17.5, vs mean 16, SD 15.7) compared to those not experiencing SDC. In addition to feeling worried about choosing the wrong option (193/418, 46.2%), respondents experiencing SDC identified the following factors as making the decisions more difficult: public health restrictions due to COVID-19 (133/418, 31.8%), difficulty separating misinformation from scientific evidence (113/418, 27.0%), information overload (107/418, 25.6%), and no or limited access to information on the decision (97/418, 23.2%); see Table 3 for all respondents. Adults identified difficulty discussing the decision with clinicians (85/323, 26.3%). Parents/caregivers identified difficulty discussing the decision with significant others (eg, family, friends; 22/95, 23.2%).

Of 1318 (90.6%) adults and 366 (83.6%) parents/caregivers who made a decision, 353 (26.8%) and 125 (34.2%), respectively, had decision regret (Table 4). Those facing mental health care decisions had higher decision regret (mean 32.0%, SD 46.8%). Adults (322/1318, 24.4%) and parents/caregivers (77/366, 21.0%) reported not getting their preferred option.

Figure 2. Proportion of Canadian cross-sectional survey respondents with clinical SDC>37.5 out of 100 during the COVID-19 pandemic. CA $1=US $0.74. DCS: Decisional Conflict Scale; LGBTQ+: lesbian, gay, bisexual, pansexual, transgender, queer, two-spirited, questioning; SDC: significant decisional conflict.



Role in Decision-Making

Most respondents indicated an active role in making the decision either alone (768/1318, 58.3%, adults vs 181/366, 49.5%, parents/caregivers) or after considering the opinion of their clinician (348/1318, 26.4%, adults vs 105/366, 28.7%, parents/caregivers). Few decided together with their clinician (148/1318, 11.2%, adults vs 63/366, 17.2%, parents/caregivers) or deferred to their clinician (54/1318, 4.1%, adults vs 17/366, 4.6%, parents/caregivers). If asked to make the decision again, most preferred an active role (1067/1318, 81.0%, adults vs 264/366, 72.1%, parents/caregivers) or together with their clinician (218/1318, 16.5%, adults vs 83/366, 22.7%, parents/caregivers).

Trusted Information Sources

Respondents indicated that trustworthy information sources were health professionals (1190/1454, 81.8%, adults vs 323/438, 73.7%, parents/caregivers), Health Canada (931/1454, 64.0%, adults vs 273/438, 62.3%, parents/caregivers), and provincial health departments (755/1454, 51.9%, adults vs 211/438, 48.2%, parents/caregivers). Fewer respondents trusted information from specific health organizations (583/1454, 40.1%, adults vs 167/438, 38.1%, parents/caregivers), consumer/patient organizations (265/1454, 18.2%, adults vs 79/438, 18.0%, parents/caregivers), companies that produce health information (199/1454, 13.7%, adults vs 49/438, 11.2%, parents/caregivers), or health insurance companies (99/1454, 6.8%, adults vs 43/438, 9.8%, parents/caregivers). Those experiencing SDC and decision regret were less trusting of these sources (Table 5).

Table 5. Trusted information sources for Canadian cross-sectional survey respondents during the COVID-19 pandemic.

Variable	Adults				Parents/caregivers			
	DCS[a]≤37.5 (n=1131)	DCS>37.5 (n=323)	DRS[b]≤25 (n=965)	DRS>25 (n=353)	DCS≤37.5 (n=343)	DCS>37.5 (n=95)	DRS≤25 (n=241)	DRS>25 (n=125)
Health professional	943 (83.4)	247 (76.5)	828 (85.8)	264 (74.8)	261 (76.1)	62 (65.3)	211 (87.6)	76 (60.8)
Health Canada	763 (67.5)	168 (52.0)	676 (70.1)	186 (52.7)	226 (65.9)	47 (49.5)	166 (68.9)	73 (58.4)
Provincial health departments	622 (55.0)	133 (41.2)	551 (57.1)	152 (43.1)	177 (51.6)	34 (35.8)	141 (58.5)	51 (40.8)
Specific health organizations	463 (40.9)	120 (37.2)	400 (41.5)	130 (36.8)	135 (39.4)	32 (33.7)	104 (43.2)	43 (34.4)
Consumer/patient associations	204 (18.0)	61 (18.9)	170 (17.6)	66 (18.7)	64 (18.7)	15 (15.8)	41 (17.0)	21 (16.8)
Companies that produce health information	150 (13.3)	49 (15.2)	123 (12.7)	57 (16.1)	39 (11.4)	10 (10.5)	27 (11.2)	16 (12.8)
Health insurance companies	81 (7.2)	18 (5.6)	66 (6.8)	23 (6.5)	36 (10.5)	7 (7.4)	20 (8.3)	13 (10.4)

[a] DCS: Decisional Conflict Scale.

[b] DRS: Decision Regret Scale.

Discussion

Principal Findings

In the first year of the pandemic, the most frequent decisions identified by Canadians were about COVID-19 vaccination, managing a health condition, social COVID-19 decisions, mental health care, medication treatments, and caregiver decisions about moving family members to or from residential facilities. One in five respondents had SDC, and a third reported decision regret. Factors making decisions more difficult were public health restrictions due to COVID-19, information overload, difficulty separating misinformation from scientific evidence, and difficulty discussing decisions with clinicians. Demographics of and the types of decisions made by respondents indicating iniquities above a level of 30% differences were ethnicity and mental health, respectively. The most trusted information sources were health care professionals and governmental sources. Our results led to the following observations:

- Respondents described 2 broad types of decisions, COVID-19–specific decisions and "routine" health-related decisions influenced by the pandemic or pandemic-related changes to health care services (eg, virtual care). When respondents were asked to focus on 1 difficult decision, some COVID-19–related decisions were selected less often and these decisions were more likely to have been influenced by mandatory public health regulations (eg, mask wearing in public indoor spaces, COVID-19 testing before an exposed child goes to school) [1,6]. However, the common decision about COVID-19 vaccination, selected by those experiencing SDC, would have been influenced by emerging scientific evidence and changing recommendations from Canadian public health officials [38]. COVID-19 vaccination became available to Canadian adults 5 months prior to the survey, with priority for frontline health care workers, the elderly in residential care, and Indigenous peoples [39]. In addition, 2 months prior to the survey, 4 vaccines were approved in Canada and public health officials recommended vaccination for adults and encouraged pregnant or breastfeeding persons to make shared decisions with their clinicians [38]. Decisions were being made when COVID-19 misinformation was spreading rapidly on social media, health care services were mostly virtual, and risk communication was also challenged with reports that the AstraZeneca vaccine caused rare blood clots [2-4,40-43]. In fact, misinformation also influenced uptake of the influenza vaccine in the United Kingdom [43].
- Many respondents reported decisional conflict and were worried about choosing the wrong option, a known manifestation of decisional conflict [33,44]. The 22.2% who reported SDC in our survey using the 16-item measure of decisional conflict were less than the 59% who reported being unsure about what to choose (1-item measure of decisional conflict) in the 1999 population-based study of 635 Canadians [17]. Another study conducted in March 2020 prior to the COVID-19 pandemic reported that 14.6% of 460 Canadian adults aged 65 years and older receiving home care services had SDC and common difficult decisions were about housing [14]. The highest proportion of respondents experiencing SDC in our study were making mental health care decisions. This was not surprising, given mental health has been impacted the most during COVID-19 [45,46]. Respondents reporting SDC highlighted unmet

decisional needs requiring targeted support to address underlying modifiable factors [11]. Effective interventions for addressing decisional conflict are patient decision aids [11,47]. During our survey, 3 publicly available decision aids specific to COVID-19 were available (eg, moving a loved one from a retirement or nursing home, vaccination for persons who were pregnant or breastfeeding) [48-50]. However, not all information is accessible. Racialized and Indigenous Canadians described that the barriers to understanding COVID-19 public health information were the use of unfamiliar medical terminology, limited to English or French, and requiring technology to access the information (eg, the internet, television) [51]. Hence, our findings indicate the importance of monitoring and better supporting those experiencing SDC with plain-language interventions designed to ensure accessibility for all Canadians.

- For respondents who made a decision, a third reported decision regret. Decision regret was previously reported in 54% of 932 Canadian caregivers of seniors receiving health care services in their home in February 2020 [27]. Those with decision regret in both studies had SDC as well as a mismatch between their preferred option and the decision made. Decision regret is associated with lower satisfaction, lower quality of life, and decisional conflict [28]. Although regret can be managed by using interventions to ensure a realistic understanding of options and expectations of outcomes [11,52], our findings showed that respondents who were experiencing decision regret had lower levels of trust in the information provided by national, provincial, or local organizations. In addition, their decisions were influenced by information overload, difficulty separating misinformation from scientific evidence, and difficulty believing scientific evidence. Another survey reported that vaccine-hesitant Canadians aged 18-40 years were influenced by conspiracy theories and a distrust of governments [20]. Furthermore, protests against public health regulations indicate high levels of distrust of COVID-19 information [53]. This highlights the need to find ways to correct misinformation and help people make informed decisions [42,43]. More importantly, as we know that decision regret may lead to more litigation/complaints, we cannot completely exclude that this may be a mechanism by which there is an increasingly fractured society, with increasingly more Canadians feeling deceived by public officials, including public health officials.
- Our surveys found more patient-controlled decision-making compared to previous surveys [6,17]. We are unsure whether these findings are related specifically to the context of the COVID-19 pandemic or a change in Canadians' roles. Half of the respondents made their decisions alone, but some preferred a more collaborative role with their clinicians. Our findings are higher than the 29% who decided alone in the 1999 Canadian survey [17] and different from a survey of 1061 Germans who preferred clinician-led decisions for hypothetical COVID-19 situations [6]. A few respondents (<6%) in our study had their clinicians make the decision for them or their clinicians make the decision after considering the respondents' opinion. Another consideration is that most health care services were virtual during the pandemic [3,7] and likely impacted roles in decision-making.
- Lastly, we found demographics and types of health-related decisions showing iniquities among Canadians, mostly ethnicity and mental health. As our health care system is espousing the quintuple objective (better patient outcomes, better patient experiences, better efficiency, better health team well-being, and more equity [54,55]), this should be resonating in public health policies in terms of future interventions targeting the most vulnerable people in our society.

Strengths and Limitations

The involvement of knowledge users on the study team was beneficial, and they guided all aspects of this study, including survey design, interpretation of findings, and drafting of results. No negative effects of involving knowledge users were encountered. Our view rate of 15.5% for adults and 28.3% for parents/caregivers was comparable to the typical rate for online surveys of 10%-20% [27,56-58]. In addition, the geographic distribution across provinces and territories was consistent with Canadian distribution [59].

With regard to limitations, our data may be subject to recall bias, given respondents were asked to identify a difficult decision made within the previous year. Second, it would have been easier to compare findings with the 1999 survey had we used "facing a complex health decision" as an eligibility criterion [17]. However, we were interested in all decisions (including those with any level of SDC or decision regret) for those who had interactions within the health care system. Compared to Canadian census data, fewer respondents had high school education or less (33.7% invited, 20.5% participated, 44.8% census) and were from the Canadian territories (0.2%-0.4% invited, 0.5% from Yukon participated, 0.1% census); see Multimedia Appendix 2. Finally, the question about costs influencing the decision preceded the question asking to list options, and some only listed options related to costs.

Conclusion

Our survey of Canadians identified COVID-19–related decisions that emerged during the first year of the pandemic and how the pandemic influenced other health-related decisions. Many adults, parents, and caregivers had unmet decision-making needs, resulting in SDC and decision regret. Factors making decisions more difficult were public health restrictions due to COVID-19, information overload, difficulty separating misinformation from scientific evidence, and difficulty discussing decisions with clinicians. Most Canadians made the decisions on their own, with few sharing the decisions with others. Canadians need their decision-making needs recognized for providing person-centered care to achieve improved health outcomes and to inform future pandemic preparedness.

Acknowledgments

All authors contributed substantially to conception and design, acquisition of data, analysis and interpretation of data; drafted the paper or revised it critically for important intellectual content; gave final approval of the version to be published; and agreed to act as guarantors of the work (ensuring that questions related to any part of the work are appropriately investigated and resolved). The study was funded by the Canadian Institutes of Health Research Priority Announcement: Patient-Oriented Research (#PJK-175386).

Data Availability

The data sets used and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supporting Canadians making health decisions: a decisional needs assessment during the COVID-19 pandemic – questionnaire.

[DOCX File , 64 KB - publichealth_v9i1e43652_app1.docx]

Multimedia Appendix 2

Comparison of respondents to those invited and the Canadian census.

[DOCX File , 16 KB - publichealth_v9i1e43652_app2.docx]

References

1. Carley S, Horner D, Body R, Mackway-Jones K. Evidence-based medicine and COVID-19: what to believe and when to change. Emerg Med J 2020 Sep 10;37(9):572-575. [doi: 10.1136/emermed-2020-210098] [Medline: 32651176]
2. Abrams EM, Shaker M, Oppenheimer J, Davis RS, Bukstein DA, Greenhawt M. The challenges and opportunities for shared decision making highlighted by COVID-19. J Allergy Clin Immunol Pract 2020 Sep;8(8):2474-2480.e1 [FREE Full text] [doi: 10.1016/j.jaip.2020.07.003] [Medline: 32679348]
3. Wong A, Bhyat R, Srivastava S, Boissé Lomax L, Appireddy R. Patient care during the COVID-19 pandemic: use of virtual care. J Med Internet Res 2021 Jan 21;23(1):e20621 [FREE Full text] [doi: 10.2196/20621] [Medline: 33326410]
4. Bhatia RS, Chu C, Pang A, Tadrous M, Stamenova V, Cram P. Virtual care use before and during the COVID-19 pandemic: a repeated cross-sectional study. CMAJ Open 2021 Feb 17;9(1):E107-E114 [FREE Full text] [doi: 10.9778/cmajo.20200311] [Medline: 33597307]
5. Hohl CM, Rosychuk RJ, Hau JP, Hayward J, Landes M, Yan JW, Canadian COVID-19 Rapid Response Network (CCEDRRN) investigators for the Network of Canadian Emergency Researchers, for the Canadian Critical Care Trials Group. Treatments, resource utilization, and outcomes of COVID-19 patients presenting to emergency departments across pandemic waves: an observational study by the Canadian COVID-19 Emergency Department Rapid Response Network (CCEDRRN). CJEM 2022 Jun 01;24(4):397-407 [FREE Full text] [doi: 10.1007/s43678-022-00275-3] [Medline: 35362857]
6. Köther AK, Siebenhaar KU, Alpers GW. Shared decision making during the COVID-19 pandemic. Med Decis Making 2021 Mar 30;41(4):430-438. [doi: 10.1177/0272989x211004147] [Medline: 33783266]
7. Verma P, Kerrison R. Patients' and physicians' experiences with remote consultations in primary care during the COVID-19 pandemic: a multi-method rapid review of the literature. BJGP Open 2022 Jan 14;6(2):BJGPO.2021.0192. [doi: 10.3399/bjgpo.2021.0192] [Medline: 35031558]
8. A coordinated global research roadmap: 2019 novel coronavirus. World Health Organization. 2020 Mar 12. URL: https://www.who.int/publications/m/item/a-coordinated-global-research-roadmap [accessed 2023-03-02]
9. Legare F, Harter M, Stiggelbout A, Thomson R, Stacey D. Choosing treatments and the role of shared decision-making. In: Nolte E, Merkur S, Anell A, editors. Achieving Person-Centred Health Systems. Cambridge, UK: Cambridge University Press; 2020.
10. Hoefel L, O'Connor AM, Lewis KB, Boland L, Sikora L, Hu J, et al. 20th Anniversary Update of the Ottawa Decision Support Framework Part 1: A Systematic Review of the Decisional Needs of People Making Health or Social Decisions. Med Decis Making 2020 Jul;40(5):555-581. [doi: 10.1177/0272989X20936209] [Medline: 32659154]
11. Stacey D, Légaré F, Boland L, Lewis KB, Loiselle M, Hoefel L, et al. 20th Anniversary Ottawa Decision Support Framework: Part 3 Overview of Systematic Reviews and Updated Framework. Med Decis Making 2020 Apr;40(3):379-398. [doi: 10.1177/0272989X20911870] [Medline: 32428429]
12. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995;15(1):25-30. [doi: 10.1177/0272989X9501500105] [Medline: 7898294]

13. Haesebaert J, Adekpedjou R, Croteau J, Robitaille H, Légaré F. Shared decision-making experienced by Canadians facing health care decisions: a Web-based survey. CMAJ Open 2019;7(2):E210-E216 [FREE Full text] [doi: 10.9778/cmajo.20180202] [Medline: 30948650]
14. Toi KA. Évaluation des besoins décisionnels en matière de santé et de services sociaux des aînés de 65 ans et plus recevant des soins à domicile au Canada Université Laval; 2021. Université Laval. 2021. URL: http://hdl.handle.net/20.500.11794/67976 [accessed 2023-03-02]
15. COVID-19 rapid guideline: managing symptoms (including at the end of life) in the community. National Institute for Health and Care Excellence (NICE). 2020 Oct 13. URL: https://www.nice.org.uk/guidance/ng163 [accessed 2023-03-02]
16. COVID-19: find information about COVID-19, how to protect yourself, your family and your commnity and what to do if you suspect you have the virus. British Columbia Centre for Disease Control. URL: http://www.bccdc.ca/health-info/diseases-conditions/covid-19 [accessed 2022-10-19]
17. O'Connor AM, Drake E, Wells G, Tugwell P, Laupacis A, Elmslie T. A survey of the decision-making needs of Canadians faced with complex health decisions. Health Expect 2003 Jun;6(2):97-109 [FREE Full text] [doi: 10.1046/j.1369-6513.2003.00215.x] [Medline: 12752738]
18. Adisso EL, Taljaard M, Stacey D, Brière N, Zomahoun HTV, Durand PJ, et al. Shared decision-making training for home care teams to engage frail older adults and caregivers in housing decisions: stepped-wedge cluster randomized trial. JMIR Aging 2022 Sep 20;5(3):e39386 [FREE Full text] [doi: 10.2196/39386] [Medline: 35759791]
19. Hoffman A, Crocker L, Mathur A, Holman D, Weston J, Campbell S, et al. Patients' and providers' needs and preferences when considering fertility preservation before cancer treatment: decision-making needs assessment. JMIR Form Res 2021 Jun 07;5(6):e25083 [FREE Full text] [doi: 10.2196/25083] [Medline: 34096871]
20. Santavicca T, Ngov C, Frounfelker R, Miconi D, Levinsson A, Rousseau C. COVID-19 vaccine hesitancy among young adults in Canada. Can J Public Health 2023 Feb 07;114(1):10-21 [FREE Full text] [doi: 10.17269/s41997-022-00693-x] [Medline: 36207641]
21. Bowen S, Graham I. Integrated knowledge translation. In: Straus S, Tetroe J, Graham ID, editors. Knowledge Translation in Health Care: Moving from Evidence to Practice. New York, NY: Wiley Blackwell; 2013:14-23.
22. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
23. Detsky AS, Bogoch II. COVID-19 in Canada: experience and response to waves 2 and 3. JAMA 2021 Sep 28;326(12):1145-1146. [doi: 10.1001/jama.2021.14797] [Medline: 34424275]
24. Detsky AS, Bogoch II. COVID-19 in Canada: experience and response. JAMA 2020 Aug 25;324(8):743-744. [doi: 10.1001/jama.2020.14033] [Medline: 32790824]
25. Detsky AS, Bogoch II. COVID-19 in Canada-the fourth through seventh waves. JAMA Health Forum 2022 Nov 04;3(11):e224160 [FREE Full text] [doi: 10.1001/jamahealthforum.2022.4160] [Medline: 36399350]
26. Annual demographic estimates: Canada, provinces and territories, 2020. Statistics Canada. 2020. URL: https://www150.statcan.gc.ca/n1/pub/91-215-x/91-215-x2020001-eng.htm [accessed 2022-10-19]
27. Lognon T, Gogovor A, Plourde KV, Holyoke P, Lai C, Aubin E, et al. Predictors of decision regret among caregivers of older Canadians receiving home care: a cross-sectional online survey. MDM Policy Pract 2022 Aug 11;7(2):23814683221116304 [FREE Full text] [doi: 10.1177/23814683221116304] [Medline: 35983319]
28. Brehaut JC, O'Connor AM, Wood TJ, Hack TF, Siminoff L, Gordon E, et al. Validation of a decision regret scale. Med Decis Making 2003;23(4):281-292. [doi: 10.1177/0272989X03256005] [Medline: 12926578]
29. Bauer GR, Scheim AI. Methods for analytic intercategorical intersectionality in quantitative research: discrimination as a mediator of health inequalities. Soc Sci Med 2019 Apr;226:236-245 [FREE Full text] [doi: 10.1016/j.socscimed.2018.12.015] [Medline: 30674435]
30. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol 2014 Jan;67(1):56-64. [doi: 10.1016/j.jclinepi.2013.08.005] [Medline: 24189091]
31. Abbasgholizadeh Rahimi S, Lépine J, Croteau J, Robitaille H, Giguere AM, Wilson BJ, et al. Psychosocial factors of health professionals' intention to use a decision aid for Down syndrome screening: cross-sectional quantitative study. J Med Internet Res 2018 Apr 25;20(4):e114 [FREE Full text] [doi: 10.2196/jmir.9036] [Medline: 29695369]
32. Delanoë A, Lépine J, Turcotte S, Leiva Portocarrero ME, Robitaille H, Giguère AM, et al. Role of psychosocial factors and health literacy in pregnant women's intention to use a decision aid for Down syndrome screening: a theory-based web survey. J Med Internet Res 2016 Oct 28;18(10):e283 [FREE Full text] [doi: 10.2196/jmir.6362] [Medline: 27793792]
33. Garvelink MM, Boland L, Klein K, Nguyen DV, Menear M, Bekker HL, et al. Decisional Conflict Scale Findings among Patients and Surrogates Making Health Decisions: Part II of an Anniversary Review. Med Decis Making 2019 May;39(4):315-326. [doi: 10.1177/0272989X19851346] [Medline: 31142205]
34. Adekpedjou R, Stacey D, Brière N, Freitas A, Garvelink MM, Turcotte S, et al. "Please listen to me": a cross-sectional study of experiences of seniors and their caregivers making housing decisions. PLoS One 2018 Aug 30;13(8):e0202975 [FREE Full text] [doi: 10.1371/journal.pone.0202975] [Medline: 30161238]

35. Adekpedjou R, Haesebaert J, Stacey D, Brière N, Freitas A, Rivest L, et al. Variations in factors associated with healthcare providers' intention to engage in interprofessional shared decision making in home care: results of two cross-sectional surveys. BMC Health Serv Res 2020 Mar 12;20(1):203 [FREE Full text] [doi: 10.1186/s12913-020-5064-3] [Medline: 32164669]
36. O'Connor AM, Tugwell P, Wells GA, Elmslie T, Jolly E, Hollingworth G, et al. Randomized trial of a portable, self-administered decision aid for postmenopausal women considering long-term preventive hormone therapy. Med Decis Making 1998;18(3):295-303. [doi: 10.1177/0272989X9801800307] [Medline: 9679994]
37. Becerra Pérez MM, Menear M, Brehaut JC, Légaré F. Extent and Predictors of Decision Regret about Health Care Decisions: A Systematic Review. Med Decis Making 2016 Aug;36(6):777-790. [doi: 10.1177/0272989X16636113] [Medline: 26975351]
38. Zipursky JS, Greenberg RA, Maxwell C, Bogler T. Pregnancy, breastfeeding and the SARS-CoV-2 vaccine: an ethics-based framework for shared decision-making. CMAJ 2021 Mar 01;193(9):E312-E314 [FREE Full text] [doi: 10.1503/cmaj.202833] [Medline: 33504561]
39. McAlister FA, Bushnik T, Leung AA, Saxinger L. Informing COVID-19 vaccination priorities based on the prevalence of risk factors among adults in Canada. CMAJ 2021 Apr 26;193(17):E617-E621 [FREE Full text] [doi: 10.1503/cmaj.210529] [Medline: 33837040]
40. Hunter PR. Thrombosis after COVID-19 vaccination. BMJ 2021 Apr 14;373:n958. [doi: 10.1136/bmj.n958] [Medline: 33853865]
41. Pai M, Chan B, Stall NM, Grill A, Ivers N, Maltsev A, Ontario COVID-19 Science Advisory Table. Vaccine-induced immune thrombotic thrombocytopenia (VITT) following adenovirus vector COVID-19 vaccination. Ontario COVID-19 Science Advisory Table. 2021 May 7. URL: https://covid19-sciencetable.ca/sciencebrief/vaccine-induced-immune-thrombotic-thrombocytopenia-vitt-following-adenovirus-vector-covid-19-vaccination/ [accessed 2023-03-02]
42. Duong D. How can clinicians counter viral misinformation? CMAJ 2021 May 03;193(18):E651-E652 [FREE Full text] [doi: 10.1503/cmaj.1095937] [Medline: 33941526]
43. Bachtiger P, Adamson A, Chow JJ, Sisodia R, Quint JK, Peters NS. The impact of the COVID-19 pandemic on the uptake of influenza vaccine: UK-wide observational study. JMIR Public Health Surveill 2021 Apr 14;7(4):e26734 [FREE Full text] [doi: 10.2196/26734] [Medline: 33651708]
44. Garvelink MM, Boland L, Klein K, Nguyen DV, Menear M, Bekker HL, et al. Decisional Conflict Scale Use over 20 Years: The Anniversary Review. Med Decis Making 2019 May;39(4):301-314. [doi: 10.1177/0272989X19851345] [Medline: 31142194]
45. Khan KS, Mamun MA, Griffiths MD, Ullah I. The mental health impact of the COVID-19 pandemic across different cohorts. Int J Ment Health Addict 2022 Jul 09;20(1):380-386 [FREE Full text] [doi: 10.1007/s11469-020-00367-0] [Medline: 32837440]
46. Gadermann AC, Thomson KC, Richardson CG, Gagné M, McAuliffe C, Hirani S, et al. Examining the impacts of the COVID-19 pandemic on family mental health in Canada: findings from a national cross-sectional study. BMJ Open 2021 Jan 12;11(1):e042871 [FREE Full text] [doi: 10.1136/bmjopen-2020-042871] [Medline: 33436472]
47. Stacey D, Légaré F, Lewis KB. Patient decision aids to engage adults in treatment or screening decisions. JAMA 2017 Aug 15;318(7):657-658. [doi: 10.1001/jama.2017.10289] [Medline: 28810006]
48. A to Z inventory of decision aids. Patient Decision Aids Research Group, Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/AZinvent.php [accessed 2022-10-19]
49. I am pregnant or breastfeeding. Should I get the COVID-19 vaccine? Version 5. Provincial Council for Maternal and Child Health. 2022. URL: https://www.pcmch.on.ca/wp-content/uploads/2022/02/PCMCH-COVID-19-Vaccine-Patient-Information-Sheet-2022_01_25_ENGLISH.pdf [accessed 2022-10-19]
50. Stacey D, Ludwig C, Lavoie J, Sinha S. During the COVID-19 pandemic, should I go to live elsewhere or stay in my retirement/assisted living home? Ottawa Hospital Research Institute, Patient Decision Aids Research Group; 2020. Ottawa Hospital Research Institute, National Institute on Ageing. 2020 Apr 14. URL: https://decisionaid.ohri.ca/docs/das/COVID-MoveFromRetirementHome.pdf [accessed 2023-03-02]
51. Aylsworth L, Manca T, Dubé È, Labbé F, Driedger SM, Benzies K, Canadian Immunization Research Network (CIRN) investigators. A qualitative investigation of facilitators and barriers to accessing COVID-19 vaccines among racialized and indigenous peoples in Canada. Hum Vaccin Immunother 2022 Nov 30;18(6):2129827 [FREE Full text] [doi: 10.1080/21645515.2022.2129827] [Medline: 36218335]
52. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4(4):CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
53. Dyer O. Covid-19: Ottawa declares emergency as truckers' protest continues. BMJ 2022 Feb 09;376:o352. [doi: 10.1136/bmj.o352] [Medline: 35140070]
54. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014 Nov;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]

55. Sikka R, Morath JM, Leape L. The Quadruple Aim: care, health, cost and meaning in work. BMJ Qual Saf 2015 Oct 02;24(10):608-610. [doi: 10.1136/bmjqs-2015-004160] [Medline: 26038586]
56. Aitken C, Power R, Dwyer R. A very low response rate in an on-line survey of medical practitioners. Aust N Z J Public Health 2008 Jun;32(3):288-289 [FREE Full text] [doi: 10.1111/j.1753-6405.2008.00232.x] [Medline: 18578832]
57. Martinez RN, Gordon EJ, Tully S, Silva A, Tarlov E, French DD, et al. A mixed-methods study of Veterans Affairs providers' experiences communicating with patients about the Affordable Care Act. Mil Med 2017 May;182(5):e1715-e1723. [doi: 10.7205/milmed-d-16-00354]
58. Surdam J, Daly B, Fulton S, Margevicius S, Schluchter M, Flocke S, et al. Recruitment strategies for nurse enrollment in an online study. Nurs Res 2020;69(1):69-73 [FREE Full text] [doi: 10.1097/NNR.0000000000000393] [Medline: 31804433]
59. Table 17-10-0009-01 Population estimates, quarterly. Statistics Canada. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901 [accessed 2022-10-19]

Abbreviations

ADHD: attention deficit hyperactivity disorder
DCS: Decisional Conflict Scale
DRS: Decision Regret Scale
LEO: Leger Opinion
LGBTQ+: lesbian, gay, bisexual, pansexual, transgender, queer, two-spirited, questioning
SDC: significant decisional conflict

Edited by A Mavragani, T Sanchez; submitted 20.10.22; peer-reviewed by P van der Nat, CY Lin; comments to author 05.12.22; revised version received 21.12.22; accepted 20.01.23; published 21.03.23.

Please cite as:
Stacey D, Ludwig C, Archambault P, Smith M, Taljaard M, Carley M, Plourde K, Boland L, Gogovor A, Graham I, Kobewka D, McLean RKD, Nelson MLA, Vanderspank-Wright B, Légaré F
Decisions and Decisional Needs of Canadians From all Provinces and Territories During the COVID-19 Pandemic: Population-Based Cross-sectional Surveys
JMIR Public Health Surveill 2023;9:e43652
URL: https://publichealth.jmir.org/2023/1/e43652
doi:10.2196/43652
PMID:36688986

Original Paper

Impact of the COVID-19 Pandemic on the Public Perceptions of the Roles and Functions of Community Pharmacies in South Korea: Updated Cross-Sectional Self-Reported Web-Based Survey

Dong-Wook Yang[1], PhD; Kyung-Bok Son[2], PhD

[1]School of Pharmacy, Sungkyunkwan University, Suwon, Gyeonggi-do, Republic of Korea

[2]College of Pharmacy, Hanyang University, Ansan, Gyeonggi-do, Republic of Korea

Corresponding Author:
Kyung-Bok Son, PhD
College of Pharmacy
Hanyang University
55 Hanyangdeahak-ro
Sangnok-gu
Ansan, Gyeonggi-do, 15588
Republic of Korea
Phone: 82 31 400 5812
Fax: 82 31 400 5958
Email: sonkyungbok@gmail.com

Abstract

Background: Community pharmacists confronted dual burdens in response to the COVID-19 pandemic by expanding the scope of pharmaceutical practices.

Objective: This study aimed to assess the perceived roles and functions of community pharmacies during the pandemic and to explore their updated roles after the pandemic began.

Methods: We conducted a self-reported web-based survey in October 2022. Based on Korean census data, we recruited the study participants (n=1000) through quota sampling stratified by age, sex, and region, yielding a 7.45% (1000/13,423) response rate. The questionnaires were composed of 3 sections: demographics, the roles and functions of community pharmacies during the pandemic, and the updated roles of community pharmacies during disasters. Each question in the second and third sections was rated on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree), and each item's mean scores and SDs were reported. The study participants were categorized into 2 groups: individuals who had a family pharmacy and those who did not. A chi-square test and ordered logistic regression analyses were conducted.

Results: Out of 1000 respondents, 418 (41.8%) had a history of COVID-19, and 639 (63.9%) had a family pharmacy. Assigning specific roles and functions to community pharmacies during the pandemic contributed to positive assessments. Respondents gave higher scores to community pharmacies that had responded appropriately (a mean Likert score of 3.66, SD .077 out of 5) and provided continuous pharmaceutical services (mean 3.67, SD 0.87) during the pandemic. The pandemic served as an opportunity to positively recognize the role of community pharmacies (mean 3.59, SD 0.83). In the ordered logistic model, having a family pharmacy was consistently associated with positive perceptions. Respondents perceived that community pharmacies collaborated with general practitioners and health authorities. However, community pharmacies need to function appropriately in terms of knowledge. The mean score of the 4 domains of community pharmacy functions was the highest for collaboration (mean 3.66, SD 0.83), followed by communication (mean 3.57, SD 0.87), responsiveness (mean 3.54, SD 0.87), and knowledge (mean 3.41, SD 0.91).

Conclusions: The pandemic resulted in interprofessional collaboration between community pharmacists and general practitioners. Family pharmacies could be a valuable asset to the comprehensive case management of patients. However, community pharmacists should have the expertise to build solid interprofessional collaborations and fulfill their expanded and updated roles.

(JMIR Public Health Surveill 2023;9:e46723) doi:10.2196/46723

KEYWORDS

community pharmacy; pharmaceutical practice; COVID-19; South Korea; pandemic; survey; pharmacists; pharmacy; primary care; medicine-centered services

Introduction

Background

Pharmacists are the third-largest group of health care professionals after nurses and physicians [1,2]. The most common type of pharmacy is a community pharmacy [3]. Community pharmacies include retail pharmacies that provide over-the-counter drugs and other health-related products to a specific community [4] and outpatient pharmacies that are considered an essential part of the primary care system [5]. Community pharmacists are the most accessible and visited health care professionals, and community pharmacies are the public's first access to their medication. Given their roles and functions, community pharmacists are integral health care professionals. However, community pharmacies are not integrated into the current primary care system [6,7]. Community pharmacists operate in a retail environment, selling or dispensing medications independently [8], and remain underestimated in their roles and functions as an integral part of the health care system [9].

In an aging population, community pharmacies are increasingly required to expand the scope of traditional pharmaceutical practices [10-12]. The health care demands of patients exceed physicians' capacity to address them [13]. Pharmaceutical practices can be categorized into medicine-centered services, patient-centered care, and public health services [14-16]. Medicine-centered services indicate traditional pharmaceutical care, such as promoting the safe use of medicines, providing information on medicines, and dispensing prescriptions. Patient-centered care emerged as an essential element of high-quality health care [17], and patient-centered pharmaceutical care includes counseling services and managing specific disease states [18-20]. Public health services may include pharmacovigilance, monitoring infectious diseases, and education in disease prevention [15,21]. Several policies have been implemented to expand the scope of pharmaceutical practice care from medicine-centered services to patient-centered care and public health services [12-15].

Community pharmacies are currently compelled to be on the frontline against COVID-19 [22,23], and the expanded role of community pharmacies has gained momentum within a short time frame [24]. As the spread of the COVID-19 pandemic continues, unprecedented measures, such as social distancing restrictions, mandated stay-at-home rules, and business closures, have been implemented to contain the transmission of the virus, control people who have been infected, and prevent overwhelming the health system [25-28]. The increased number of people with COVID-19 has caused a surge in the demand for health care services [29-31]. Further, the need for pharmaceutical services during the pandemic is beyond the scope of pharmaceutical services in the nonpandemic situation. According to the 4 disaster phases, the updated need for pharmaceutical services could be divided into prevention, preparedness, response, and recovery [32,33]. The public needs to be informed about adequate personal protection against the infection and community transmission of COVID-19. Patients must be treated for the disease and other related symptoms at home with rational medication use.

Pharmaceutical Practice in South Korea

Pharmacy practice in South Korea needs to catch up to practices in other countries. The concept of pharmaceutical care was first introduced in 1994 [34]. However, pharmacy practices are still limited to medicine-centered services. Pharmacists in South Korea do not have the authority to direct refill reviews and approvals and administer vaccinations [11]. South Korea introduced a law separating prescribing and dispensing medicine in 2000 [35]. Under the law, only physicians could prescribe medicines, and pharmacists could dispense the medicine by prescription. The law aimed to enhance the rational use of medicine and manage pharmaceutical expenditures [36]. However, the separate systems relegated community pharmacies to delivering medicine-centered services, only dispensing medicines rather than performing other pharmacy practices [11]. Thus, community pharmacies are located near clinics or hospitals where they receive and dispense prescriptions [37]. Patients visit community pharmacies near clinics or hospitals to fill prescriptions; they usually receive prescriptions when they visit clinics or hospitals and then fill those prescriptions at different community pharmacies. Thus, providing centralized and consistent pharmaceutical practices is difficult in South Korea. The Korean Pharmaceutical Association has suggested a family pharmacy system to the government to provide effective and comprehensive pharmaceutical services [37]. However, a family pharmacy is still not a legally established concept in the Korean health care system.

In South Korea, the COVID-19 pandemic has changed the roles and functions of community pharmacies in 2 ways. The expanded role of traditional community pharmacies requested by the community has been accelerating quickly. Furthermore, community pharmacies must develop new functions to address the pandemic, particularly in their preparedness and responses to COVID-19. We considered the interaction between the 2 burdens—increased and updated demand for pharmaceutical practice—placed on the current community pharmacies. The primary objective of this study was to assess the perceived roles and functions of community pharmacies during the pandemic and explore their updated roles, not to test a hypothesis but to measure the public perceptions toward community pharmacies. The secondary objective was to investigate the role of family pharmacies in the public perceptions of community pharmacies.

Methods

Overview

We conducted this study as part of a more extensive study on the public perceptions of the roles and functions of community pharmacies in South Korea. Public perceptions before the COVID-19 outbreak have been published elsewhere [11]. This study investigated the changes in the public perceptions of community pharmacies during the pandemic. To this end, we developed a self-reported web-based survey. The survey was conducted from October 2 to October 4, 2022, with technical support from Realmeter, a survey agency. The characteristics of the previous and current studies are provided in Multimedia Appendix 1.

Study Design and Participant Recruitment

Assuming a 95% CI, 3.1% margin of error, and 0.5% SD, a representative sample size (n=1000) for the entire population was obtained. A quota sampling method was implemented to recruit study participants. The participants were registered in the survey agency's panel and agreed to respond to the survey in advance. Quotas for the sample were stratified by sex, age, and region based on Korean census data. The study population was Korean individuals aged 19 years or older. For the quota sampling process, regions were categorized into *Si* and *Do*. *Si* indicates an urban area in South Korea and includes 8 regions: Seoul, Busan, Daegu, Incheon, Kwangju, Daejeon, Ulsan, and Sejong. Do indicates a rural area in South Korea and includes 9 regions: Gyeonggi, Gangwon, Chungbuk, Chungnam, Jeonbuk, Jeonnam, Gyeongbuk, Gyeongnam, and Jeju. *Do* indicates a rural area in Korean and includes 9 regions: Gyeonggi, Gangwon, Chungbuk, Chungnam, Jeonbuk, Jeonnam, Gyeongbuk, Gyeongnam, and Jeju. Invitations were sent randomly for the stratified quotas until the number of completed surveys reached the predetermined percentage.

The survey was conducted as follows. We explained the purpose of the study. Study participants who consented to participate in this study were directed to the encrypted website to complete the survey. Those who completed the survey received a voucher worth US $4. Out of 13,423 invitations, 1000 completed the survey, yielding a 7.45% response rate.

Ethics Approval

The study protocol was reviewed and approved by the Hanyang University Institutional Review Board (HYIRB-202209-006).

Questionnaires

The survey comprised 3 sections: demographics, the roles and functions of the community pharmacy during the pandemic, and the updated roles of the community pharmacy during disasters. The questions in the second and third sections were derived from the literature on pharmaceutical care, pharmacy practice, and the role of pharmacies during disasters. They were also modified to reflect the Korean context. The content of the developed questionnaire was pretested with members of the public, registered pharmacists, and researchers.

The demographic variables were sex, age, education, occupation, and region. For the statistical analyses, regions (8 *Si*s and 9 *Do*s) were recategorized into metropolitan, urban, and rural areas. Metropolitan areas included South Korea's capital city, Seoul, and its surrounding areas, Incheon and Gyeonggi. Urban areas included the 6 *Si*s excluding Seoul and Incheon. Rural areas included the 8 *Do*s excluding Gyeonggi. Two additional items were added to the demographics: experience with COVID-19 and having a family pharmacy. In this study, *a family pharmacy* was defined as a pharmacy where a respondent would usually fill a prescription or ask for advice about his or her health problem [38-40]. The second section was composed of 28 questions regarding the roles and functions of community pharmacies during the pandemic. We presented 12 questions to assess the roles of community pharmacies. For the functions of community pharmacies, we categorized 16 questions into 4 domains: communication, knowledge, responsiveness, and collaboration. The third section includes 19 questions regarding the roles of community pharmacies during disasters. To clarify the updated roles of the community pharmacy during disasters, 4 disaster phases—prevention, preparedness, responses, and recovery—were used. Questions in the second and third sections were rated on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Detailed survey questions are provided in Multimedia Appendix 2.

Statistical Analyses

Two types of analysis were applied: descriptive analysis and ordered logistic regression. Descriptive analysis was used to present the distribution of rated scores, mean scores, and SDs. A chi-square test compared categorical variables, including having a family pharmacy and experiencing COVID-19 [41]. Cronbach α was calculated to measure the internal consistency of survey items. An ordered logistic regression was conducted to elucidate the association between having a family pharmacy and positive perceptions toward community pharmacies. In the ordered logistic model, the dependent variable was ordinal scale data, and the independent variables were having a family pharmacy, experiencing COVID-19, sex, age, region, and education. We tested the proportional odds assumption using the Brant test, and if it was not satisfied, we conducted a generalized ordered logistic regression [42]. The results of the ordered logistic regression presented adjusted odds ratios (aORs) with 95% CIs. Statistical significance was considered at a 2-sided *P* value <.05. All statistical analyses were performed using Stata software (version 16; StataCorp).

Results

Survey Respondents

Table 1 presents the basic characteristics of the 1000 respondents to this study. Of the respondents, 418 (41.8%) had a history of COVID-19, and 639 (63.9%) had a family pharmacy. We separated the respondents into 2 groups: those with or without a family pharmacy. The distribution of respondents in the 2 groups was not significantly different in terms of sex, age, region, education, occupation, or COVID-19 infection (all *P*>.05).

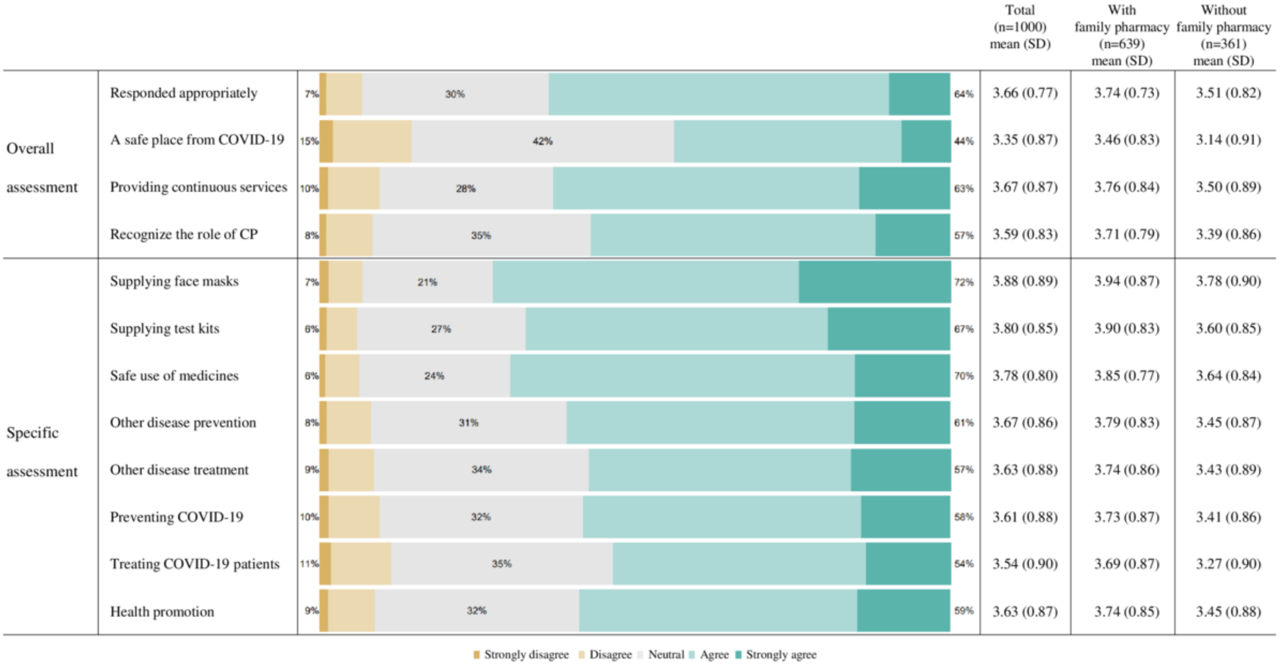
Table 1. Characteristics of the survey respondents.

Characteristics	Total respondents (n=1000), n (%)	Respondents with a family pharmacy (n=639), n (%)	Respondents without a family pharmacy (n=361), n (%)	*P* value
Sex				.39
Male	496 (49.6)	324 (50.7)	172 (47.6)	
Female	504 (50.4)	315 (49.3)	189 (52.4)	
Age range (years)				.07
19-29	162 (16.2)	89 (13.9)	73 (20.2)	
30-39	151 (15.1)	94 (14.7)	57 (15.8)	
40-49	188 (18.8)	120 (18.8)	68 (18.8)	
50-59	198 (19.8)	131 (20.5)	67 (18.6)	
60 and older	301 (30.1)	205 (32.1)	96 (26.6)	
Region				.14
Metropolitan	504 (50.4)	317 (49.6)	187 (51.8)	
Urban	190 (19)	133 (20.8)	57 (15.8)	
Rural	306 (30.6)	189 (29.6)	117 (32.4)	
Education				.36
Below 12 years	221 (22.1)	135 (21.1)	86 (23.8)	
13 years and above	779 (77.9)	504 (78.9)	275 (76.2)	
Occupation				.14
White-collar worker	450 (45)	288 (45.1)	162 (44.9)	
Blue-collar worker	141 (14.1)	91 (14.2)	50 (13.9)	
Self-employed	90 (9)	62 (9.7)	28 (7.8)	
Housewife	159 (15.9)	106 (16.6)	53 (14.7)	
Student	25 (2.5)	10 (1.6)	15 (4.2)	
Others	135 (13.5)	82 (12.8)	53 (14.7)	
COVID-19 infection				.65
Yes	418 (41.8)	271 (42.4)	147 (40.7)	
No	582 (58.2)	368 (57.6)	214 (59.3)	

Assessment of the Roles of Community Pharmacies During the Pandemic

Figure 1 describes the overall and specific assessments of the roles of community pharmacies during the pandemic. We reported the distribution of rated scores, mean scores, and SDs. A higher score indicates that the roles of community pharmacies are perceived as positive by the respondents. Regarding the overall assessment, respondents gave higher scores for community pharmacies that had responded appropriately (a mean Likert score of 3.66, SD 0.77 out of 5) and provided continuous pharmaceutical services (mean 3.67, SD 0.87). The pandemic served as an opportunity to positively recognize the roles of community pharmacies (mean 3.59, SD 0.83). However, respondents gave lower scores for community pharmacies as a safe place from COVID-19 (mean 3.35, SD 0.87). Respondents gave higher scores for community pharmacies playing an essential role in the supply of publicly distributed face masks (mean 3.88, SD 0.89) and COVID-19 rapid antigen test kits (mean 3.80, SD 0.85) and in the safe use of medicines (mean 3.78, SD 0.80). However, respondents gave lower scores for community pharmacies playing an essential role in treating patients with COVID-19 (mean 3.54, SD 0.90) and preventing COVID-19 (mean 3.61, SD 0.88). When we separated the respondents into 2 groups, higher scores were observed consistently in the subgroup of those with a family pharmacy than those without one.

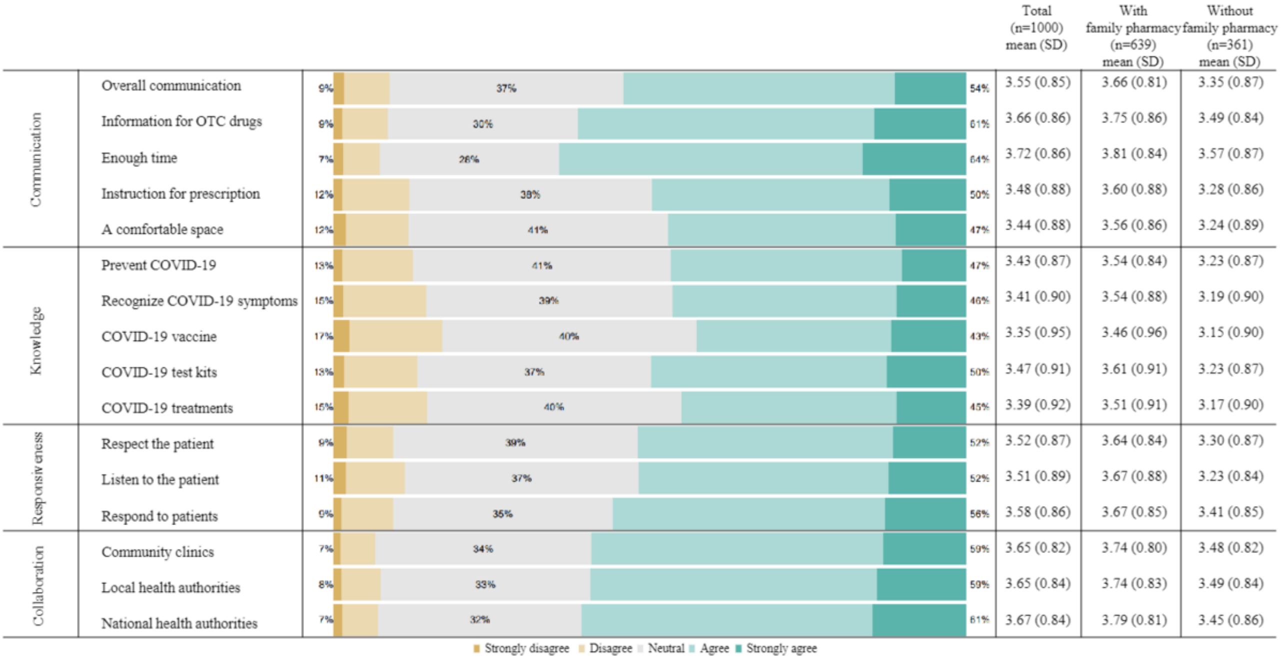
Figure 1. The assessment of the roles of community pharmacies (CP) during the pandemic.



Perceived Functions of Community Pharmacies During the Pandemic

Figure 2 presents the perceived functions of community pharmacies during the pandemic. The mean score of the 4 domains was the greatest for collaboration (mean 3.66, SD 0.83), followed by communication (mean 3.57, SD 0.87), responsiveness (mean 3.54, SD 0.87), and knowledge (mean 3.41, SD 0.91). Respondents gave higher scores for community pharmacies collaborating with community clinics and hospitals (mean 3.65, SD 0.82), local health authorities (mean 3.65, SD 0.84), and national health authorities (mean 3.67, SD 0.84). However, respondents gave lower scores for the expertise of community pharmacies regarding the COVID-19 vaccine (mean 3.35, 0.95), COVID-19 treatment (mean 3.39, SD 0.92), and recognizing COVID-19 symptoms (mean 3.41, SD 0.90). Cronbach α values for the communication, knowledge, responsiveness, and collaboration domains were .88, .92, .83, and .85, respectively. When we separated the respondents into 2 groups, higher scores were observed consistently in the subgroup of those with a family pharmacy than those without one.

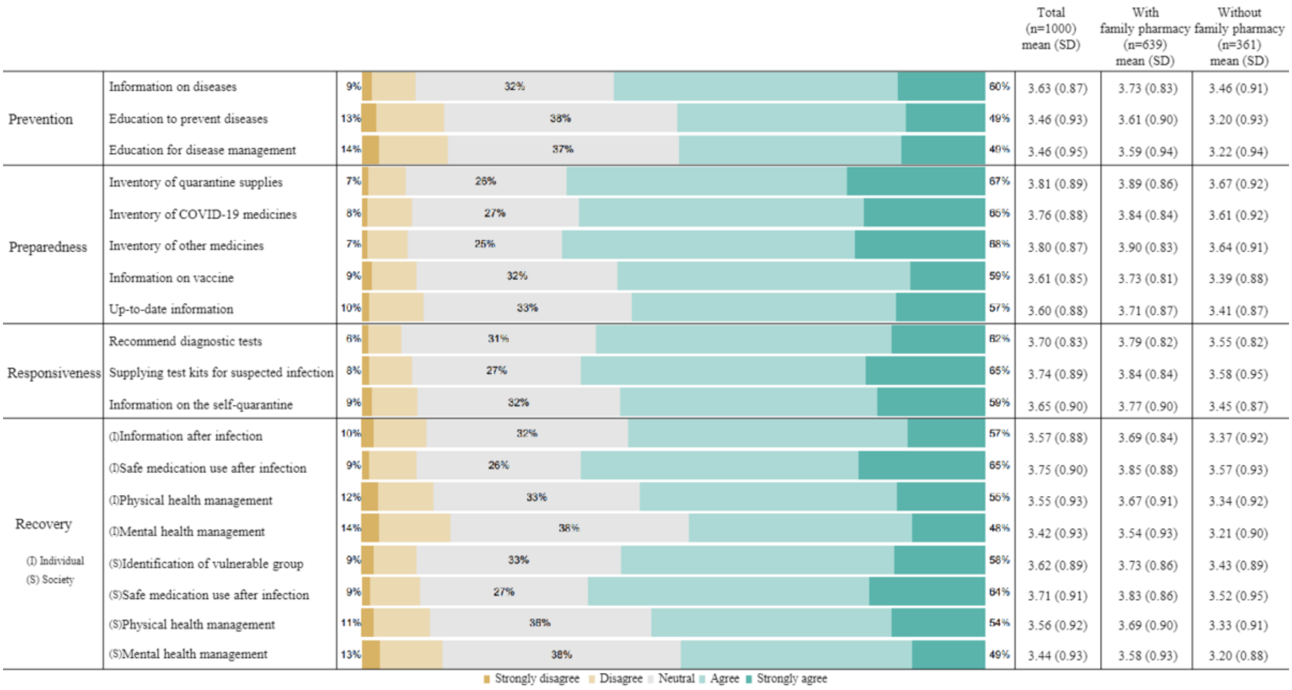
Figure 2. The perceived functions of community pharmacies during the pandemic. OTC: over-the-counter.



Updated Roles of the Community Pharmacy During Disasters

Figure 3 describes the updated roles of community pharmacies during disasters. The 4 phases in descending order according to mean scores were preparedness (mean 3.72, SD 0.88), responsiveness (mean 3.70, SD 0.87), recovery (mean 3.58, SD 0.92), and prevention (mean 3.52, SD 0.92). Respondents gave higher scores for the items belonging to the preparedness phase: an inventory of quarantine supplies (mean 3.81, SD 0.89), other medicines, including over-the-counter drugs (mean 3.80, SD 0.87), and medicines used for treating COVID-19 (mean 3.76, 0.88). Respondents gave higher scores for pharmacies supplying COVID-19 rapid antigen test kits (mean 3.74, SD 0.89) under the responsiveness phase and managing safe medication use after infection (mean 3.75, SD 0.90) under the recovery phase. However, respondents gave lower scores to pharmacies for providing education to prevent infectious disease (mean 3.46, SD 0.93) and education for infectious disease management (mean 3.46, 0.95) under the prevention phase. Cronbach α values for the prevention, preparedness, responsiveness, individual recovery, and social recovery domains were .86, .88, .81, .87, and .88, respectively. Higher scores were consistently observed in the subgroup of those with a family pharmacy than those without one.

Figure 3. The updated roles of community pharmacies during disasters.



					Total (n=1000) mean (SD)	With family pharmacy (n=639) mean (SD)	Without family pharmacy (n=361) mean (SD)
Prevention	Information on diseases	9%	32%	60%	3.63 (0.87)	3.73 (0.83)	3.46 (0.91)
	Education to prevent diseases	13%	38%	49%	3.46 (0.93)	3.61 (0.90)	3.20 (0.93)
	Education for disease management	14%	37%	49%	3.46 (0.95)	3.59 (0.94)	3.22 (0.94)
Preparedness	Inventory of quarantine supplies	7%	26%	67%	3.81 (0.89)	3.89 (0.86)	3.67 (0.92)
	Inventory of COVID-19 medicines	8%	27%	65%	3.76 (0.88)	3.84 (0.84)	3.61 (0.92)
	Inventory of other medicines	7%	25%	68%	3.80 (0.87)	3.90 (0.83)	3.64 (0.91)
	Information on vaccine	9%	32%	59%	3.61 (0.85)	3.73 (0.81)	3.39 (0.88)
	Up-to-date information	10%	33%	57%	3.60 (0.88)	3.71 (0.87)	3.41 (0.87)
Responsiveness	Recommend diagnostic tests	6%	31%	62%	3.70 (0.83)	3.79 (0.82)	3.55 (0.82)
	Supplying test kits for suspected infection	8%	27%	65%	3.74 (0.89)	3.84 (0.84)	3.58 (0.95)
	Information on the self-quarantine	9%	32%	59%	3.65 (0.90)	3.77 (0.90)	3.45 (0.87)
Recovery Ⓘ Individual Ⓢ Society	ⒾInformation after infection	10%	32%	57%	3.57 (0.88)	3.69 (0.84)	3.37 (0.92)
	ⒾSafe medication use after infection	9%	26%	65%	3.75 (0.90)	3.85 (0.88)	3.57 (0.93)
	ⒾPhysical health management	12%	33%	56%	3.55 (0.93)	3.67 (0.91)	3.34 (0.92)
	ⒾMental health management	14%	38%	48%	3.42 (0.93)	3.54 (0.93)	3.21 (0.90)
	ⓈIdentification of vulnerable group	9%	33%	58%	3.62 (0.89)	3.73 (0.86)	3.43 (0.89)
	ⓈSafe medication use after infection	9%	27%	64%	3.71 (0.91)	3.83 (0.86)	3.52 (0.95)
	ⓈPhysical health management	11%	36%	54%	3.56 (0.92)	3.69 (0.90)	3.33 (0.91)
	ⓈMental health management	13%	38%	49%	3.44 (0.93)	3.58 (0.93)	3.20 (0.88)

Having a Family Pharmacy and Positive Perceptions

Table 2 presents factors associated with a positive assessment of the roles of community pharmacies during the pandemic. Having a family pharmacy was consistently associated with a positive assessment of all items about the roles of community pharmacies during the pandemic. In particular, having a family pharmacy (aOR 1.67, 95% CI 1.23-2.15) and higher educational attainment (aOR 1.40, 95% CI 1.02-1.93) were associated with a positive assessment that community pharmacies responded appropriately during the pandemic.

Table 2. Ordered logistic regression analysis for each item in the assessment of the roles of community pharmacies (CP) during the pandemic.

Dependent variable[a]	Independent variable (reference), odds ratio (95% CI)						
	Family pharmacy: yes (no)	COVID-19 infection: yes (no)	Sex: female (male)	Age (years)	Education: ≥13 years (<13 years)	Region: metropolitan (rural)	Region: urban (rural)
Overall assessment							
Responded appropriately[b]	1.669[c] (1.298-2.147)	1.060 (0.831-1.352)	1.058 (0.830-1.348)	1.000 (0.991-1.010)	1.401[d] (1.015-1.933)	1.277 (0.952-1.714)	0.766 (0.400-1.464)
A safe place from COVID-19	1.970[c] (1.531-2.536)	0.846 (0.667-1.071)	0.867 (0.684-1.100)	0.987[e] (0.979-0.996)	1.366[d] (1.010-1.847)	1.250 (0.894-1.746)	0.914 (0.701-1.191)
Providing continuous services	1.790[c] (1.396-2.295)	1.006 (0.796-1.270)	1.021 (0.805-1.295)	1.000 (0.991-1.008)	1.610[e] (1.179-2.199)	1.105 (0.799-1.528)	1.099 (0.835-1.447)
Recognize the role of CP	2.041[c] (1.589-2.623)	1.087 (0.858-1.377)	0.824 (0.649-1.046)	1.001 (0.992-1.009)	1.323 (0.989-1.769)	0.932 (0.677-1.281)	0.889 (0.676-1.169)
Specific assessment							
Supplying face masks	1.425[e] (1.114-1.823)	0.825 (0.650-1.046)	1.136 (0.899-1.436)	1.001 (0.993-1.010)	1.272 (0.920-1.758)	0.928 (0.664-1.296)	0.846 (0.637-1.123)
Supplying test kits	1.968[c] (1.543-2.509)	0.935 (0.739-1.183)	0.940 (0.744-1.189)	0.995 (0.956-1.003)	1.141 (0.849-1.534)	1.198 (0.855-1.678)	0.886 (0.674-1.164)
Safe use of medicines	1.676[c] (1.282-2.114)	0.947 (0.743-1.207)	1.005 (0.788-1.281)	1.000 (0.991-1.009)	1.281 (0.934-1.758)	1.016 (0.724-1.424)	0.930 (0.702-1.233)
Other disease prevention[b]	2.516[c] (1.924-3.290)	0.942 (0.743-1.195)	0.978 (0.772-1.238)	1.006 (0.997-1.016)	1.195 (0.886-1.612)	1.302 (0.926-1.829)	1.138 (0.848-1.527)
Other disease treatment	1.945[c] (1.524-2.482)	0.989 (0.785-1.246)	0.997 (0.789-1.259)	0.997 (0.989-1.006)	1.345 (0.993-1.820)	0.971 (0.704-1.338)	1.027 (0.783-1.347)
Preventing COVID-19	1.981[c] (1.560-2.514)	0.886 (0.701-1.119)	0.961 (0.756-1.220)	1.000 (0.992-1.008)	1.362 (1.008-1.842)	1.097 (0.797-1.510)	0.759 (0.577-0.997)
Treating COVID-19 patients	2.460[c] (1.926-3.143)	1.049 (0.832-1.322)	0.987 (0.780-1.249)	0.991[d] (0.983-1.000)	1.131 (0.831-1.540)	1.059 (0.761-1.474)	0.981 (0.743-1.294)
Health promotion[b]	1.795[c] (1.407-2.290)	1.057 (0.836-1.337)	0.859 (0.679-1.086)	1.008 (0.999-1.018)	1.401[d] (1.046-1.876)	1.420 (0.977-2.063)	1.271 (0.947-1.705)

[a] Value of dependent variable: 1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, and 5=Strongly agree.

[b] Generalized ordered logistic regression (1, 2, and 3 vs 4 and 5).

[c] *P*<.001.

[d] *P*<.05.

[e] *P*<.01.

Discussion

Study Overview

This study assessed the perceived roles and functions of community pharmacies during the pandemic and explored their updated roles after the pandemic. The expanded roles of community pharmacies requested by the community accelerated quickly during the pandemic. Furthermore, community pharmacies began to develop new functions to address the pandemic. Findings from this study highlight how to expand and update the roles and functions of community pharmacies in countries where pharmacies are still limited to medicine-centered services.

Principal Findings in Pharmacy Practice

Three findings are noteworthy. First, in the overall assessment, more than 50% of participants agreed that community pharmacies responded appropriately, and the pandemic served as an opportunity to recognize the role of community pharmacies. When we asked for specific assessments, the proportion of participants who agreed was the highest for the question that pharmacies that had played an essential role in supplying face masks, followed by COVID-19 test kits and advice about the safe use of medicines. Assigning specific roles and functions to community pharmacies would contribute to the public's positive assessment of community pharmacies. Similarly, community pharmacies in other countries have shown effort and courage in providing face masks and COVID-19 test kits and administering vaccinations [6,43,44]. The public assessed effort and courage positively.

Second, we conducted this study as part of a more extensive study on the public perceptions of community pharmacies in South Korea. By comparing the previous and current studies, we could evaluate the changed roles and functions of community pharmacies before and during the pandemic. As explained earlier, these 2 studies implemented a quota sampling method stratified by age, sex, and region. Both studies conducted a self-reported web-based survey for 1000 adults aged older than 19 years. Furthermore, the questionnaires in the 2 studies were closely related. We provided 4 domains to understand the functions of community pharmacies. Changes in the public perceptions toward the functions of community pharmacies were noted. In this study, the mean Likert scores of the 4 domains in descending order were collaboration (3.66), communication (3.57), responsiveness (3.54), and knowledge (3.41). In contrast, the mean Likert scores of the 4 domains in descending order were knowledge (3.73), responsiveness (3.47), communication (3.45), and collaboration (3.44) in the previous study [11]. The mean Likert score for collaboration, communication, and responsiveness increased by 0.22, 0.12, and 0.07 points, respectively, whereas the score for knowledge decreased by 0.32 points. Knowledge has been essential to interprofessional collaboration between community pharmacists and general practitioners [45,46]. However, community pharmacists still need to develop their knowledge, particularly about COVID-19, during the pandemic. In contrast, the public responded that community pharmacies collaborated with community clinics and health authorities during the pandemic. The expertise of community pharmacists is required to build solid interprofessional collaboration between community pharmacists and general practitioners after the pandemic.

Third, we asked about the updated roles of community pharmacies in 4 disaster phases—prevention, preparedness, responsiveness, and recovery. Respondents gave higher scores for the items in preparedness and responsiveness, whereas they gave lower scores for the items in prevention and recovery. The preparedness and responsiveness phases were key areas during the pandemic [47]. However, community pharmacies' role in prevention and recovery should not be underestimated [48,49]. The role of community pharmacies in the infection prevention phase is to take measures to reduce the health risk of disasters, including providing information and education about infections and ensuring that patients are aware of the increased risk of adverse health outcomes [24,32]. The role of community pharmacies in the recovery phase involved checking on the local community's health needs, identifying and prioritizing susceptible patients, and participating in postdisaster reports [32,47]. In European countries, a comprehensive range of pharmacy practices, from prevention to recovery, has been implemented to address the demands during a pandemic [24]. By doing so, community pharmacies in other countries could be a bridge between medical care and broader community services. Furthermore, community pharmacies could be embedded in the comprehensive case management of patients with COVID-19 and other susceptible groups [50,51]. In this study, having a family pharmacy was consistently associated with positive assessments by the public. The consistent effect implies that family pharmacies could be a platform for embedding community pharmacies into the comprehensive case management of patients.

Policy Implications: Barriers and Facilitators in Expanding the Roles of Community Pharmacies

It is necessary to review the mean score of each survey item. The scores ranged from 3.35 to 3.88 points, implying that they were less than "agree" or "positive." The government and researchers in South Korea praised the roles of community pharmacies in distributing face masks in the early phase of the pandemic [52-54]. However, in the survey, respondents gave ratings of less than "agree (4.00 pts)" when asked if community pharmacies had played an essential role in the supply of publicly distributed face masks. Which factors caused the difference in assessment between the public and the government and researchers? In addition, what caused the public to perceive the roles of community pharmacies during the pandemic as less than positive?

First, we should note the timing factor. We conducted this survey in October 2022, and community pharmacies supplied face masks from March 2020 to July 2020. A more than 2-year gap between the 2 events could have caused the public's neutral stance. The lack of expertise is another factor. Similar to retailers, community pharmacies supplied face masks to the public. Pharmacies, post offices, and Nonghyup Hanaro Mart—a grocery store—supplied face masks during the initial stage of the pandemic. However, the public did not expect the distribution of face masks to be a substantial role of community pharmacies. The public also responded that community pharmacies needed to gain knowledge. In the questionnaire, the public gave the lowest score to community pharmacists who provided expertise on the COVID-19 vaccine. Community pharmacists in Egypt expressed an unmet need for pandemic-related knowledge as a barrier to the expanded roles of community pharmacies [44]. The lack of expertise requires measures such as providing continuous guidance on the pandemic and educational materials to community pharmacists. Space is the final factor. Our previous research raised time and space issues in the communication between pharmacists and patients [11]. Before the pandemic, we noted the availability of private and comfortable spaces within community pharmacies as barriers to expanding the roles of community pharmacies. During the pandemic, the public raised space issues regarding coronavirus transmission within the pharmacy and responded that community pharmacies were unsafe from COVID-19. Improper infection control measures in community pharmacies, reported in Australia, provided clues to interpret this finding [55]. In South Korea, approximately 60% of community pharmacies had a total space of less than 66 m^2, implying that a separate space for infection control was impossible in most cases [56].

This study also highlights the association between community pharmacies' roles and functions and pharmacists' expertise. Pharmacists in South Korea do not have the authority to direct refill reviews and approvals and administer vaccinations [11]. During the pandemic, the roles of community pharmacists were limited to supplying face masks and COVID-19 test kits. Their roles did not expand to more active and professional ones, such as vaccinating the public and finding, preventing, and treating COVID-19 cases. In this context, the public might not expect the expertise of community pharmacists in areas they have yet to experience. Similarly, community pharmacists have not been incentivized to develop their expertise in those areas [57]. Assigning new roles to community pharmacies would contribute to developing the expertise of community pharmacists. Similarly, developing the expertise of community pharmacists would successfully link to expanding the roles and functions of community pharmacies.

Strengths and Limitations

This study has several strengths. First, we analyzed public perceptions toward the roles and functions of community pharmacies. Many studies have captured the roles and functions of community pharmacies from the perspective of health care providers. However, public perceptions are essential in designing patient-centered community pharmacies. Second, we conducted a large-scale (n=1000) web-based questionnaire that represented the entire population in South Korea in terms of sex, age, and region. Third, we conducted this study as part of a more extensive study on the public perceptions of the roles and functions of community pharmacies. Before the COVID-19 pandemic, we studied the public perceptions of the roles and functions of community pharmacies in South Korea. This study analyzed the public perceptions of the roles and functions of community pharmacies during the COVID-19 pandemic. By comparing the different study results between the prepandemic and pandemic periods, we could investigate the effect of the pandemic on the public perceptions of the roles and functions of community pharmacies.

This study has several limitations. First and foremost, this study used a web-based survey and applied a stratified sampling method to obtain samples that represented the entire population in South Korea. The response rate was low for some quotas, and the total response rate was 7.45%. A low response rate implies that the study participants are a self-selected minority, and their participation might be linked to other personal characteristics. There would be selection bias in the sampling process, and the conclusion would be biased. Second, we investigated public perceptions to assess the roles and functions of community pharmacies during the pandemic and to explore the updated roles of community pharmacies after the pandemic began. However, measuring perceptions is problematic when allowing respondents to rate different items on the questionnaire. Third, this study was a cross-sectional survey, which may hinder the interpretation of the findings as a causal relationship. Longitudinal design and qualitative studies are needed to understand the contexts and mechanisms of public perceptions.

Conclusions

The COVID-19 pandemic has changed the roles and functions of community pharmacies. Assigning specific roles and functions to community pharmacies contributed to the public’s positive assessment. Furthermore, having a family pharmacy was consistently associated with the public’s positive assessments. The pandemic resulted in interprofessional collaboration between community pharmacists and general practitioners. In particular, family pharmacies could be a valuable platform in which to embed community pharmacies in the comprehensive case management of patients. However, community pharmacists should have the expertise to build solid interprofessional collaborations and fulfill their expanded and updated roles.

Acknowledgments

This work was supported by the National Research Foundation of Korea (grant NRF-2022R1F1A1071338). However, the funder was not involved in the study design; the collection, analysis, and interpretation of data; the writing of this paper; or the decision to submit it for publication.

Authors' Contributions

KBS and DWY conceived and designed the analysis, performed the analysis, and wrote the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The characteristics of the previous and current studies.
[PPTX File , 41 KB - publichealth_v9i1e46723_app1.pptx]

Multimedia Appendix 2
Detailed survey questions.
[DOCX File , 18 KB - publichealth_v9i1e46723_app2.docx]

References

1. Smith M, Giuliano MR, Starkowski MP. In Connecticut: improving patient medication management in primary care. Health Aff (Millwood) 2011 Apr;30(4):646-654. [doi: 10.1377/hlthaff.2011.0002] [Medline: 21471485]

2. Mossialos E, Courtin E, Naci H, Benrimoj S, Bouvy M, Farris K, et al. From "retailers" to health care providers: transforming the role of community pharmacists in chronic disease management. Health Policy 2015 May;119(5):628-639. [doi: 10.1016/j.healthpol.2015.02.007] [Medline: 25747809]
3. OECD. Health at a Glance 2021: OECD Indicators. Paris, France: OECD Publishing; 2021.
4. Ibrahim MIM. Chapter 18 - assessment of medication dispensing and extended community pharmacy services. In: Social and Administrative Aspects of Pharmacy in Low- and Middle-Income Countries: Present Challenges and Future Solutions. Cambridge, Massachusetts: Academic Press, Elsevier Inc; 2018:295-309.
5. Zheng S, Yang L, Zhou P, Li H, Liu F, Zhao R. Recommendations and guidance for providing pharmaceutical care services during COVID-19 pandemic: a China perspective. Res Social Adm Pharm 2021 Jan;17(1):1819-1824 [FREE Full text] [doi: 10.1016/j.sapharm.2020.03.012] [Medline: 32249102]
6. Piquer-Martinez C, Urionagüena A, Benrimoj SI, Calvo B, Martinez-Martinez F, Fernandez-Llimos F, et al. Integration of community pharmacy in primary health care: the challenge. Res Social Adm Pharm 2022 Aug;18(8):3444-3447. [doi: 10.1016/j.sapharm.2021.12.005] [Medline: 35016847]
7. Dawoud D, Chen AM, Rossing CV, Garcia-Cardenas V, Law AV, Aslani P, et al. Pharmacy practice research priorities during the COVID-19 pandemic: recommendations of a panel of experts convened by FIP Pharmacy Practice Research Special Interest Group. Res Social Adm Pharm 2021 Jan;17(1):1903-1907 [FREE Full text] [doi: 10.1016/j.sapharm.2020.08.020] [Medline: 32912829]
8. Hindi AMK, Schafheutle EI, Jacobs S. Community pharmacy integration within the primary care pathway for people with long-term conditions: a focus group study of patients', pharmacists' and GPs' experiences and expectations. BMC Fam Pract 2019 Feb 08;20(1):26 [FREE Full text] [doi: 10.1186/s12875-019-0912-0] [Medline: 30736732]
9. Anderson C, Sharma R. Primary health care policy and vision for community pharmacy and pharmacists in England. Pharm Pract (Granada) 2020 Jan;18(1):1870 [FREE Full text] [doi: 10.18549/PharmPract.2020.1.1870] [Medline: 32256901]
10. Tannenbaum C, Tsuyuki RT. The expanding scope of pharmacists' practice: implications for physicians. CMAJ 2013 Oct 01;185(14):1228-1232 [FREE Full text] [doi: 10.1503/cmaj.121990] [Medline: 23959280]
11. Son K, Choi S, Kim D. Public perceptions of the roles and functions of community pharmacies in the era of expanding scopes of pharmaceutical practice: a questionnaire survey in South Korea. Health Soc Care Community 2019 Jul 06;27(4):1095-1101. [doi: 10.1111/hsc.12727] [Medline: 30729603]
12. Hindi AMK, Jacobs S, Schafheutle EI. Solidarity or dissonance? a systematic review of pharmacist and GP views on community pharmacy services in the UK. Health Soc Care Community 2019 May 26;27(3):565-598. [doi: 10.1111/hsc.12618] [Medline: 30047617]
13. Bollen A, Harrison R, Aslani P, van Haastregt JC. Factors influencing interprofessional collaboration between community pharmacists and general practitioners-a systematic review. Health Soc Care Community 2019 Jul 20;27(4):e189-e212. [doi: 10.1111/hsc.12705] [Medline: 30569475]
14. Agomo CO, Ogunleye J, Portlock J. A survey to identify barriers in the public health role of community pharmacists. J Pharm Health Serv Res 2016 Jul 25;7(4):253-261. [doi: 10.1111/jphs.12153]
15. Eades CE, Ferguson JS, O'Carroll RE. Public health in community pharmacy: a systematic review of pharmacist and consumer views. BMC Public Health 2011 Jul 21;11(1):582 [FREE Full text] [doi: 10.1186/1471-2458-11-582] [Medline: 21777456]
16. Wolters M, van Hulten R, Blom L, Bouvy ML. Exploring the concept of patient centred communication for the pharmacy practice. Int J Clin Pharm 2017 Dec 9;39(6):1145-1156 [FREE Full text] [doi: 10.1007/s11096-017-0508-5] [Medline: 28887610]
17. Woods TM, Lucas AJ, Robke JT. Making a case for a patient-centered integrated pharmacy practice model. Am J Health Syst Pharm 2011 Feb 01;68(3):259-263. [doi: 10.2146/ajhp100013] [Medline: 21258031]
18. Kibicho J, Owczarzak J. A patient-centered pharmacy services model of HIV patient care in community pharmacy settings: a theoretical and empirical framework. AIDS Patient Care STDS 2012 Jan;26(1):20-28. [doi: 10.1089/apc.2011.0212] [Medline: 22149903]
19. Raiche T, Pammett R, Dattani S, Dolovich L, Hamilton K, Kennie-Kaulbach N, et al. Community pharmacists' evolving role in Canadian primary health care: a vision of harmonization in a patchwork system. Pharm Pract (Granada) 2020 Oct;18(4):2171 [FREE Full text] [doi: 10.18549/PharmPract.2020.4.2171] [Medline: 33149795]
20. Hepler CD, Strand LM. Opportunities and responsibilities in pharmaceutical care. Am J Hosp Pharm 1990 Mar;47(3):533-543. [doi: 10.1093/ajhp/47.3.533] [Medline: 2316538]
21. Kiles TM, Hagemann T, Felts B, Crill C. Development of a remote public health advanced pharmacy practice experience in response to COVID-19. J Pharm Pract 2022 Oct 01;35(5):783-790. [doi: 10.1177/08971900211004830] [Medline: 33789494]
22. Cadogan CA, Hughes CM. On the frontline against COVID-19: community pharmacists' contribution during a public health crisis. Res Social Adm Pharm 2021 Jan;17(1):2032-2035 [FREE Full text] [doi: 10.1016/j.sapharm.2020.03.015] [Medline: 32245691]

23. Carpenter DM, Hastings T, Westrick S, Rosenthal M, Mashburn P, Kiser S, et al. Rural community pharmacies' preparedness for and responses to COVID-19. Res Social Adm Pharm 2021 Jul;17(7):1327-1331 [FREE Full text] [doi: 10.1016/j.sapharm.2020.10.008] [Medline: 34155979]
24. Costa S, Romão M, Mendes M, Horta MR, Rodrigues AT, Carneiro AV, et al. Pharmacy interventions on COVID-19 in Europe: mapping current practices and a scoping review. Res Social Adm Pharm 2022 Aug;18(8):3338-3349 [FREE Full text] [doi: 10.1016/j.sapharm.2021.12.003] [Medline: 34924315]
25. Venkatesh A, Edirappuli S. Social distancing in COVID-19: what are the mental health implications? BMJ 2020 Apr 06;369:m1379. [doi: 10.1136/bmj.m1379] [Medline: 32253182]
26. Tull MT, Edmonds KA, Scamaldo KM, Richmond JR, Rose JP, Gratz KL. Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life. Psychiatry Res 2020 Jul;289:113098 [FREE Full text] [doi: 10.1016/j.psychres.2020.113098] [Medline: 32434092]
27. Sameer A, Khan M, Nissar S, Banday M. Assessment of mental health and various coping strategies among general population living under imposed COVID-lockdown across world: a cross-sectional study. Ethics Med Public Health 2020 Oct;15:100571 [FREE Full text] [doi: 10.1016/j.jemep.2020.100571] [Medline: 32838000]
28. Son K, Lee T, Hwang S. Disease severity classification and COVID-19 outcomes, Republic of Korea. Bull World Health Organ 2021 Jan 01;99(1):62-66 [FREE Full text] [doi: 10.2471/BLT.20.257758] [Medline: 33658735]
29. Miller IF, Becker AD, Grenfell BT, Metcalf CJE. Disease and healthcare burden of COVID-19 in the United States. Nat Med 2020 Aug 16;26(8):1212-1217. [doi: 10.1038/s41591-020-0952-y] [Medline: 32546823]
30. Wang X, Pasco RF, Du Z, Petty M, Fox SJ, Galvani AP, et al. Impact of social distancing measures on coronavirus disease healthcare demand, Central Texas, USA. Emerg Infect Dis 2020 Oct;26(10):2361-2369 [FREE Full text] [doi: 10.3201/eid2610.201702] [Medline: 32692648]
31. Mian M, Sreedharan S, Giles S. Increased dispensing of prescription medications in Australia early in the COVID-19 pandemic. Med J Aust 2021 May 21;214(9):428-429 [FREE Full text] [doi: 10.5694/mja2.51029] [Medline: 33884634]
32. Watson KE, Singleton JA, Tippett V, Nissen LM. Defining pharmacists' roles in disasters: a Delphi study. PLoS One 2019 Dec 26;14(12):e0227132 [FREE Full text] [doi: 10.1371/journal.pone.0227132] [Medline: 31877194]
33. Watson KE, Tippett V, Singleton JA, Nissen LM. Disaster health management: do pharmacists fit in the team? Prehosp Disaster Med 2019 Feb;34(1):30-37. [doi: 10.1017/S1049023X18001152] [Medline: 30604658]
34. Hong K. Pharmaceutical care in Korea. Article in Korean. J Kor Soc Health-Syst Pharm 1995 Sep 30;12(3):263-269 [FREE Full text]
35. Jones RS. Health - care reform in Korea. In: OECD Economics Department Working Papers, No. 797. Paris, France: OECD Publishing; 2010.
36. Kwon S. Pharmaceutical reform and physician strikes in Korea: separation of drug prescribing and dispensing. Soc Sci Med 2003 Aug;57(3):529-538. [doi: 10.1016/s0277-9536(02)00378-7] [Medline: 12791494]
37. Lee JH, Kim H, Han HS, Park I, Sohn HS. Needs for a family pharmacy and expected family pharmacy services. Yakhak Hoeji 2023 Feb 28;67(1):58-65. [doi: 10.17480/psk.2023.67.1.58]
38. Shearer SW, Sarlin MD, Burns WR. Family pharmacists. JAMA 1982 Nov 19;248(19):2450-2451. [doi: 10.1001/jama.1982.03330190020016]
39. Helling DK. Family practice pharmacy service: part I. Drug Intell Clin Pharm 1981 Dec 30;15(12):971-977. [doi: 10.1177/106002808101501210] [Medline: 6121696]
40. Helling DK. Family practice pharmacy service: part II. Drug Intell Clin Pharm 1982 Jan 30;16(1):35-48. [doi: 10.1177/106002808201600106] [Medline: 7053952]
41. Lee SW. Methods for testing statistical differences between groups in medical research: statistical standard and guideline of Life Cycle Committee. Life Cycle 2022 Jan 24;2:e1. [doi: 10.54724/lc.2022.e1]
42. Williams R. Understanding and interpreting generalized ordered logit models. J Math Sociol 2016 Jan 29;40(1):7-20. [doi: 10.1080/0022250x.2015.1112384]
43. Hoti K, Jakupi A, Hetemi D, Raka D, Hughes J, Desselle S. Provision of community pharmacy services during COVID-19 pandemic: a cross sectional study of community pharmacists' experiences with preventative measures and sources of information. Int J Clin Pharm 2020 Aug 11;42(4):1197-1206 [FREE Full text] [doi: 10.1007/s11096-020-01078-1] [Medline: 32524513]
44. Bahlol M, Dewey RS. Pandemic preparedness of community pharmacies for COVID-19. Res Social Adm Pharm 2021 Jan;17(1):1888-1896 [FREE Full text] [doi: 10.1016/j.sapharm.2020.05.009] [Medline: 32417070]
45. Hansen RN, Nørgaard LS, Hedegaard U, Søndergaard L, Servilieri K, Bendixen S, et al. Integration of and visions for community pharmacy in primary health care in Denmark. Pharm Pract (Granada) 2021 Jan 22;19(1):2212 [FREE Full text] [doi: 10.18549/PharmPract.2021.1.2212] [Medline: 33520039]
46. Williams S, Brad L, Jones D, Hayes J. Pharmacy professionals are part of the solution to reduce GPs' workload. BMJ 2021 May 14;373:n1176. [doi: 10.1136/bmj.n1176] [Medline: 33990342]
47. Aruru M, Truong H, Clark S. Pharmacy Emergency Preparedness and Response (PEPR): a proposed framework for expanding pharmacy professionals' roles and contributions to emergency preparedness and response during the COVID-19

pandemic and beyond. Res Social Adm Pharm 2021 Jan;17(1):1967-1977 [FREE Full text] [doi: 10.1016/j.sapharm.2020.04.002] [Medline: 32389631]

48. Hess K, Bach A, Won K, Seed SM. Community pharmacists roles during the COVID-19 pandemic. J Pharm Pract 2022 Jun 15;35(3):469-476. [doi: 10.1177/0897190020980626] [Medline: 33317371]
49. Liu S, Luo P, Tang M, Hu Q, Polidoro JP, Sun S, et al. Providing pharmacy services during the coronavirus pandemic. Int J Clin Pharm 2020 Apr 28;42(2):299-304 [FREE Full text] [doi: 10.1007/s11096-020-01017-0] [Medline: 32222911]
50. Hayden JC, Parkin R. The challenges of COVID-19 for community pharmacists and opportunities for the future. Ir J Psychol Med 2020 Sep 21;37(3):198-203 [FREE Full text] [doi: 10.1017/ipm.2020.52] [Medline: 32434603]
51. Alexander GC, Stoller KB, Haffajee RL, Saloner B. An epidemic in the midst of a pandemic: opioid use disorder and COVID-19. Ann Intern Med 2020 Jul 07;173(1):57-58 [FREE Full text] [doi: 10.7326/M20-1141] [Medline: 32240283]
52. Kim J. President Moon, “ Remeber the hard work of pharmacists.” Article in Korean. Yonhap News. 2020 Jul 14. URL: https://www.yna.co.kr/view/AKR20200713158100017 [accessed 2023-06-20]
53. Moon JI. We are deeply appreciate for the cooperation of frontline pharmacists. Article in Korean. Office of the President, Republic of Korea. 2020 Mar 6. URL: https://tinyurl.com/y7eyvbx2 [accessed 2023-06-20]
54. Choi H, Kim M, Ryu H, Mo Y. COVID-19 mask policy, role of pharmacists and pharmacies. Article in Korean. HSS21 2021 Jun 30;12(3):1829-1842. [doi: 10.22143/hss21.12.3.127]
55. Sum ZZ, Ow CJ. Community pharmacy response to infection control during COVID-19. a cross-sectional survey. Res Social Adm Pharm 2021 Jan;17(1):1845-1852 [FREE Full text] [doi: 10.1016/j.sapharm.2020.06.014] [Medline: 33317763]
56. Choi J. What is the pharmacist's preferred pharmacy area? Article in Korean. K-Pharm News. 2021 Feb 23. URL: https://www.kpanews.co.kr/article/show.asp?table=article&idx=220252&category= [accessed 2023-06-20]
57. Svensberg K, Kälvemark Sporrong S, Håkonsen H, Toverud EL. 'Because of the circumstances, we cannot develop our role': Norwegian community pharmacists' perceived responsibility in role development. Int J Pharm Pract 2015 Aug;23(4):256-265. [doi: 10.1111/ijpp.12154] [Medline: 25359665]

Abbreviations

aOR: adjusted odds ratio

Edited by A Mavragani, T Sanchez; submitted 26.02.23; peer-reviewed by IH Lee, DK Yon; comments to author 12.04.23; revised version received 29.05.23; accepted 29.06.23; published 13.07.23.

Please cite as:
Yang DW, Son KB
Impact of the COVID-19 Pandemic on the Public Perceptions of the Roles and Functions of Community Pharmacies in South Korea: Updated Cross-Sectional Self-Reported Web-Based Survey
JMIR Public Health Surveill 2023;9:e46723
URL: https://publichealth.jmir.org/2023/1/e46723
doi:10.2196/46723
PMID:37390391

Original Paper

A Spatiotemporal Solution to Control COVID-19 Transmission at the Community Scale for Returning to Normalcy: COVID-19 Symptom Onset Risk Spatiotemporal Analysis

Chengzhuo Tong[1], PhD; Wenzhong Shi[1], PhD; Anshu Zhang[1], PhD; Zhicheng Shi[2], PhD

[1]Otto Poon Charitable Foundation Smart Cities Research Institute and Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China

Corresponding Author:
Zhicheng Shi, PhD
Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University
3688 Nanhai Avenue, Nanshan District
Shenzhen, 518060
China
Phone: 86 0755 26531244
Email: shizhic@hotmail.com

Abstract

Background: Following the recent COVID-19 pandemic, returning to normalcy has become the primary goal of global cities. The key for returning to normalcy is to avoid affecting social and economic activities while supporting precise epidemic control. Estimation models for the spatiotemporal spread of the epidemic at the refined scale of cities that support precise epidemic control are limited. For most of 2021, Hong Kong has remained at the top of the "global normalcy index" because of its effective responses. The urban-community-scale spatiotemporal onset risk prediction model of COVID-19 symptom has been used to assist in the precise epidemic control of Hong Kong.

Objective: Based on the spatiotemporal prediction models of COVID-19 symptom onset risk, the aim of this study was to develop a spatiotemporal solution to assist in precise prevention and control for returning to normalcy.

Methods: Over the years 2020 and 2021, a spatiotemporal solution was proposed and applied to support the epidemic control in Hong Kong. An enhanced urban-community-scale geographic model was proposed to predict the risk of COVID-19 symptom onset by quantifying the impact of the transmission of SARS-CoV-2 variants, vaccination, and the imported case risk. The generated prediction results could be then applied to establish the onset risk predictions over the following days, the identification of high–onset-risk communities, the effectiveness analysis of response measures implemented, and the effectiveness simulation of upcoming response measures. The applications could be integrated into a web-based platform to assist the antiepidemic work.

Results: Daily predicted onset risk in 291 tertiary planning units (TPUs) of Hong Kong from January 18, 2020, to April 22, 2021, was obtained from the enhanced prediction model. The prediction accuracy in the following 7 days was over 80%. The prediction results were used to effectively assist the epidemic control of Hong Kong in the following application examples: identified communities within high–onset-risk always only accounted for 2%-25% in multiple epidemiological scenarios; effective COVID-19 response measures, such as prohibiting public gatherings of more than 4 people were found to reduce the onset risk by 16%-46%; through the effect simulation of the new compulsory testing measure, the onset risk was found to be reduced by more than 80% in 42 (14.43%) TPUs and by more than 60% in 96 (32.99%) TPUs.

Conclusions: In summary, this solution can support sustainable and targeted pandemic responses for returning to normalcy. Faced with the situation that may coexist with SARS-CoV-2, this study can not only assist global cities in responding to the future epidemics effectively but also help to restore social and economic activities and people's normal lives.

(JMIR Public Health Surveill 2023;9:e36538) doi:10.2196/36538

KEYWORDS

return to normalcy; precise prevention and control; risk prediction; COVID-19 symptom onset; symptom; COVID-19

Introduction

After more than a year of the COVID-19 pandemic, people are beginning to look forward to the return of normalcy in the forms of no masks, no isolation, and no social distancing [1]. A global normalcy index has been compiled by the Economist to measure the current daily activities relative to the prepandemic level in the 50 largest economies in the world [2]. The index comprises 8 indicators, split into 3 domains, as follows: transport and travel, recreation and entertainment, and retailing and work [2]. According to the data as of December 31, 2021, Hong Kong is at the top of the normalcy index [2]. Hong Kong is one of the most densely populated cities and ranks one of the top 5 territories for population density in the world. The high population density together with the highly developed international transportation networks make it vulnerable to importation as well as the local spread of SARS-CoV-2 variants [3-5]. Despite the severe challenges [6-8], Hong Kong's figures on confirmed and fatal cases are among the lowest in the world [9]. Even in the face of the rapidly spreading variant of Omicron, the local epidemic situation in Hong Kong has remained stable, and no local cases have been recorded for 80 consecutive days [9,10]. Hong Kong's economy is exhibiting major signs of recovery with people's lives subsequently being back to normal [11].

It is of note that from the outbreak of the COVID-19 epidemic on January 18, 2020, only strict social distancing and crowd control measures were adopted in Hong Kong, as opposed to extreme measures, such as citywide lockdowns or curfews [12]. The overall aim of this new direction was to formulate sustainable and targeted pandemic responses and hence more effectively enable better control of the epidemic situations [13]. A spatiotemporal solution has been developed to enable the prediction of COVID-19 symptom onset risk and provide support for Hong Kong's new direction of pandemic prevention.

The improved intercity-scale [14,15] and urban-community–scale [16] weighted kernel density estimation (WKDE) models have been applied to predict spatiotemporal COVID-19 symptom onset risk [17,18]. The above extended WKDE models provide competitive advantages based on high viral load around the date of symptom onset [19,20] as well as the delay between the onset dates and subsequent confirmation report dates [21-23]. Thus, the tracking of the COVID-19 symptom onset risk better reflects the COVID-19 transmission at the urban-community scale. However, the urban-community–scale WKDE model needs further improvement regarding the new normal of long-term coexistence with SARS-CoV-2 [24]. Consequently, the urban-community–scale WKDE model has been further enhanced by introducing the reproduction number for local cases, the number of passenger arrivals at all ports of entry, time-varying vaccination rate, and vaccination efficiency.

The enhanced urban-community–scale WKDE model has been used to predict the risk of COVID-19 symptom onset in each of the 291 tertiary planning units (TPUs) in Hong Kong. It generates the following urban-community–scale prediction results: the daily onset risk prediction results during the following few days and the daily onset risk prediction enhanced by the population distributions within 291 TPUs during the following few days. Hence, these applications are able to be implemented to support precise prevention and control as well as the recovery of socioeconomic activities in Hong Kong. A web-based visualization platform has been further developed to present the above prediction and application results. The study was conducted based on a spatiotemporal data set of 7184 local onset cases [25] with community-level locations in Hong Kong from January 18, 2020, to April 22, 2021, including, so far, all four epidemic waves of COVID-19 epidemics in Hong Kong [26].

Methods

Data Sources

In addition to asymptomatic cases, imported cases under mandatory quarantine, and cases with unknown location information, a total of 7184 local onset cases from January 18, 2020, to April 22, 2021, have been used in this study. These local onset cases include information regarding the date of the symptom onset and consequent report as well as the community-level location of these onset cases prior to diagnosis. Currently, the main COVID-19 vaccines used in Hong Kong are mRNA (BNT162b2) vaccines [27] and inactivated (CoronaVac) SARS-CoV-2 vaccines [28]. Thus, to measure the impact of vaccination on the COVID-19 epidemic, daily data [29] on people with full vaccination were selected, including the daily number of people receiving the second dose of the BNT162b2 vaccine [29] and the daily number of people receiving the second dose of the CoronaVac vaccine [29]. The daily traffic flow data covering all Hong Kong strategic routes from January 18, 2020, to April 22, 2021, were used in this study [30]. To enable the measurement of the risk of the imported cases, the daily number of passenger arrivals at all 16 ports of entry during the same period was also obtained [31]. Moreover, the reproductive number for local cases during the same period was generated by an enhanced Susceptible-Infectious-Removed model [22].

An Enhanced Urban-Community–Scale WKDE Model for Predicting the Onset Risk of COVID-19 Symptoms

As a further development of the original urban-community–scale WKDE model, the enhanced urban-community–scale WKDE model was used in this study [14-16]. The details of this model are presented in Multimedia Appendix 1. Finally, the original onset risk prediction in each location was divided by the city's maximum predicted risk, on a specific date, and thereby, standardized to a value between 0 and 1 [14-16]. Following the idea of a hit rate, the accuracy of the enhanced urban-community–scale WKDE model is set as the percentage of onset cases, on the date of prediction, that occur in the areas with predicted onset risk higher than 0.8 (known as hotspots) [14-16]. The population distributions in different communities were further used to enhance the original onset risk (details are presented in Multimedia Appendix 1).

Ethical Considerations

For this study, no ethics approval was required because the granularity of the data was at the city level.

Results

Daily Spatiotemporal Risk Prediction of COVID-19 Symptom Onset by the Enhanced Urban-Community–Scale WKDE Model

Based on the onset cases data in 291 TPUs of Hong Kong from January 18, 2020, to April 24, 2021, the daily spatiotemporal onset risk predictions of COVID-19 symptoms were obtained by the enhanced urban-community–scale WKDE model developed in this study. During the following first week after predicting the onset risk, the median prediction accuracy of the urban-community–scale WKDE model was over 80% (Figure 1). The prediction accuracy during the following second week, as possibly could be expected, was lower due to the accumulation of prediction errors over time.

In all four epidemic waves caused by different SARS-CoV-2 variants, the daily spatiotemporal distributions of urban-community–scale COVID-19 symptom onset risk from January 18, 2020, to April 24, 2021, were used to measure urban-community–scale heterogeneity (Figure 2). Communities with the onset risk higher than 0.8 were always concentrated in Central and Western, Wan Chai, Kowloon City, and Yau Tsim Mong, all of which are central shopping and commercial areas with a very high traffic flow. Conversely, the onset risk was relatively low in suburban communities, far away from the main transportation network. The Q statistic results conducted by Geodetector showed that the variations in mobility had a very strong effect on variations in onset risk in 291 TPUs, with a Q value of 0.98 (*P*=.005). Technological details of Geodetector is presented in Multimedia Appendix 1. Moreover, the trend on the daily regional shift regarding different onset risk levels could be visually displayed.

Figure 1. Accuracy of the predicted risk of COVID-19 symptom onset by the enhanced urban-community–scale weighted kernel density estimation model.

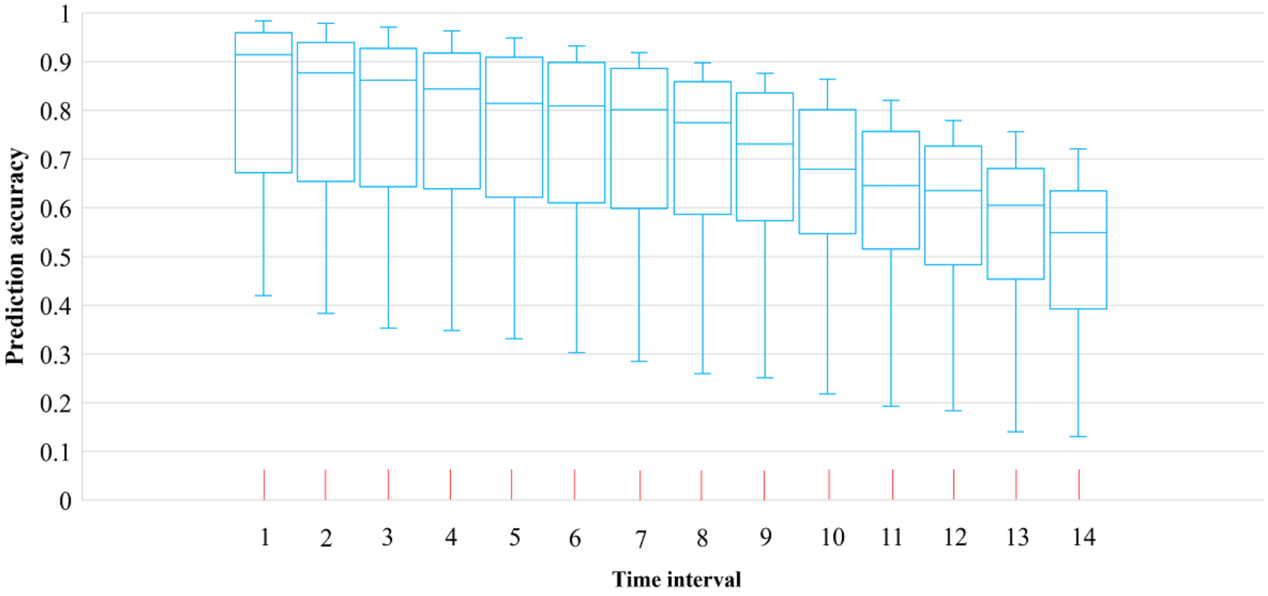
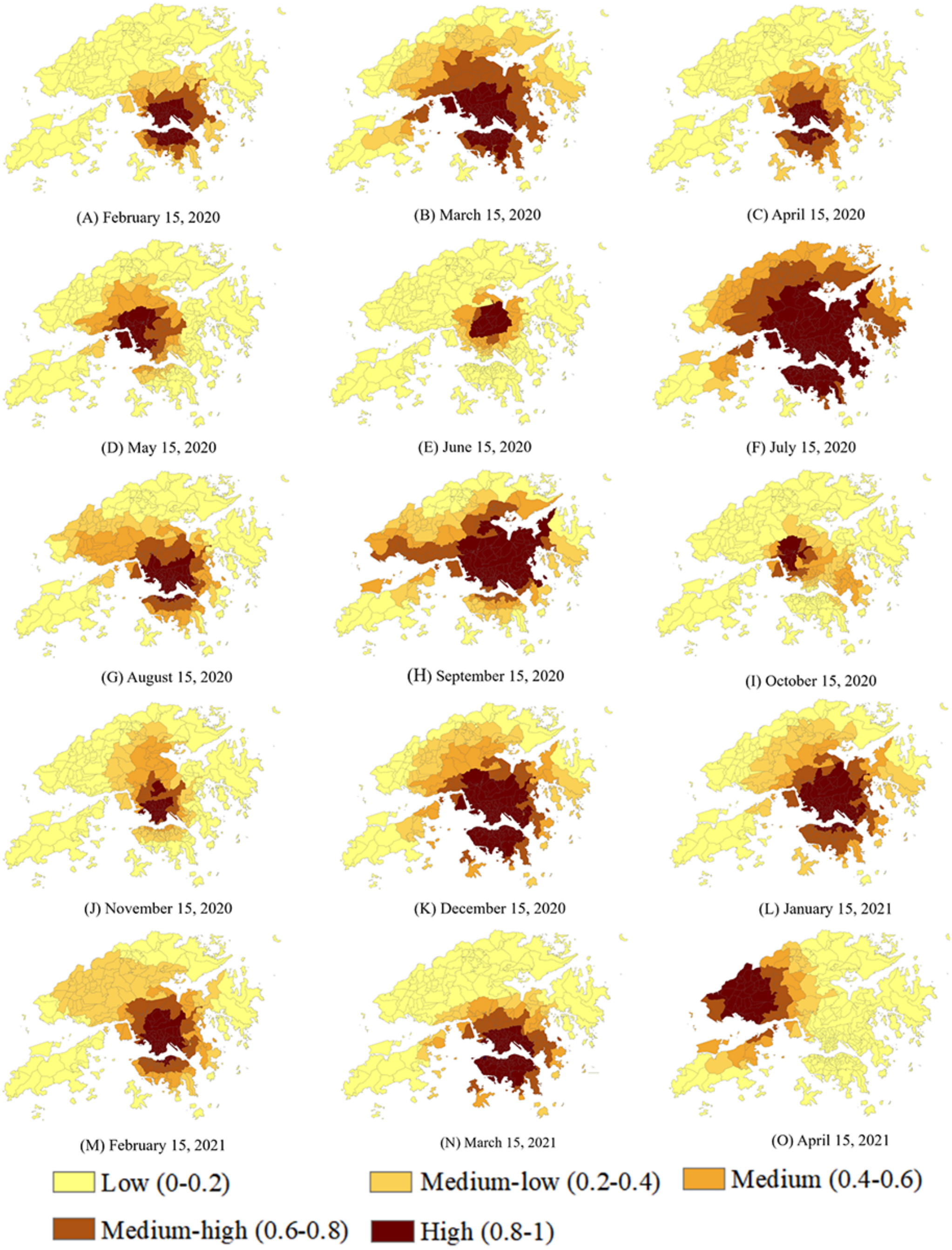


Figure 2. Predicted risk of original COVID-19 symptom onset risk across Hong Kong (A-O) in all four epidemic waves from January 18, 2020, to April 24, 2021.



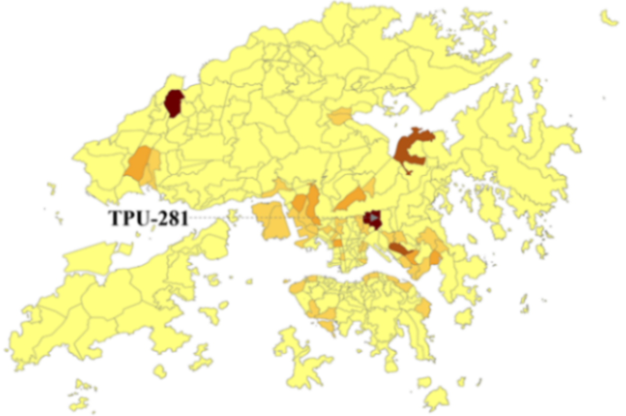
(A) February 15, 2020
(B) March 15, 2020
(C) April 15, 2020
(D) May 15, 2020
(E) June 15, 2020
(F) July 15, 2020
(G) August 15, 2020
(H) September 15, 2020
(I) October 15, 2020
(J) November 15, 2020
(K) December 15, 2020
(L) January 15, 2021
(M) February 15, 2021
(N) March 15, 2021
(O) April 15, 2021

Identification of High–Onset-Risk Communities in Three Epidemiologic Setting Scenarios

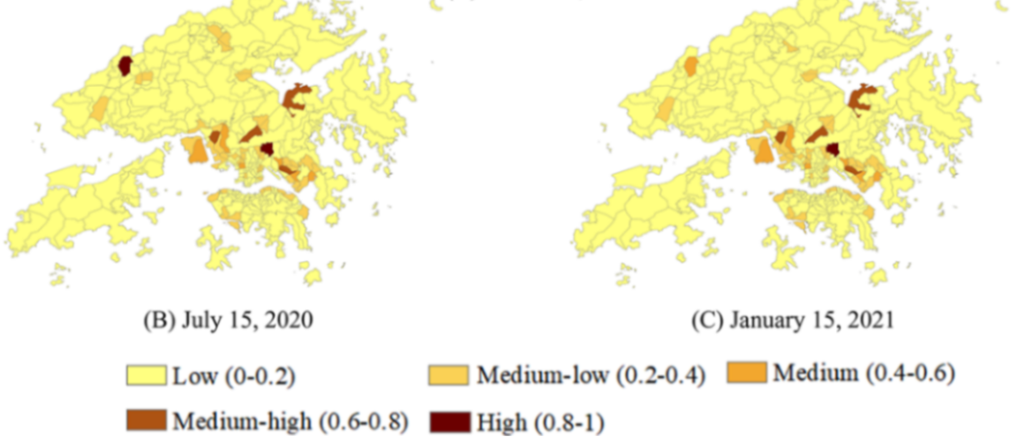
From the identified results of high–onset-risk communities based on the daily prediction of the original onset risk, the number of TPUs within high–onset-risk areas always only accounted for 2%-25% under the two epidemiologic setting scenarios of no cases and sporadic and clusters of cases. This enabled only the most stringent measures needing to be implemented in a very limited area within the high onset risk to reduce the impact on normal activities in other areas. In addition, such an identification of high–onset-risk communities also supports the allocation of limited COVID-19 medical supplies in Hong Kong.

However, in the community transmission scenario, the high–onset-risk communities identified directly from the original symptom onset risk prediction did not seem to be controlled in a small area. For example, on July 15, 2020—the peak of the third wave of the epidemic in Hong Kong—the high–onset-risk communities identified accounted for 155 (53.26%) of all 291 TPUs (Figure 2F). Limited COVID-19 medical supplies and human resources for health could not be allocated to the communities that needed them most. Based on the enhanced identification of high–onset-risk communities by population distributions, even during the peak period of the four epidemic waves, communities within the high onset risk could be reduced to a limited area (Figure 3A-C).

Figure 3. Predicted COVID-19 symptom onset risk enhanced by the population distributions within 291 tertiary planning units (TPUs) of Hong Kong (A-C).



(A) March 15, 2020


(B) July 15, 2020

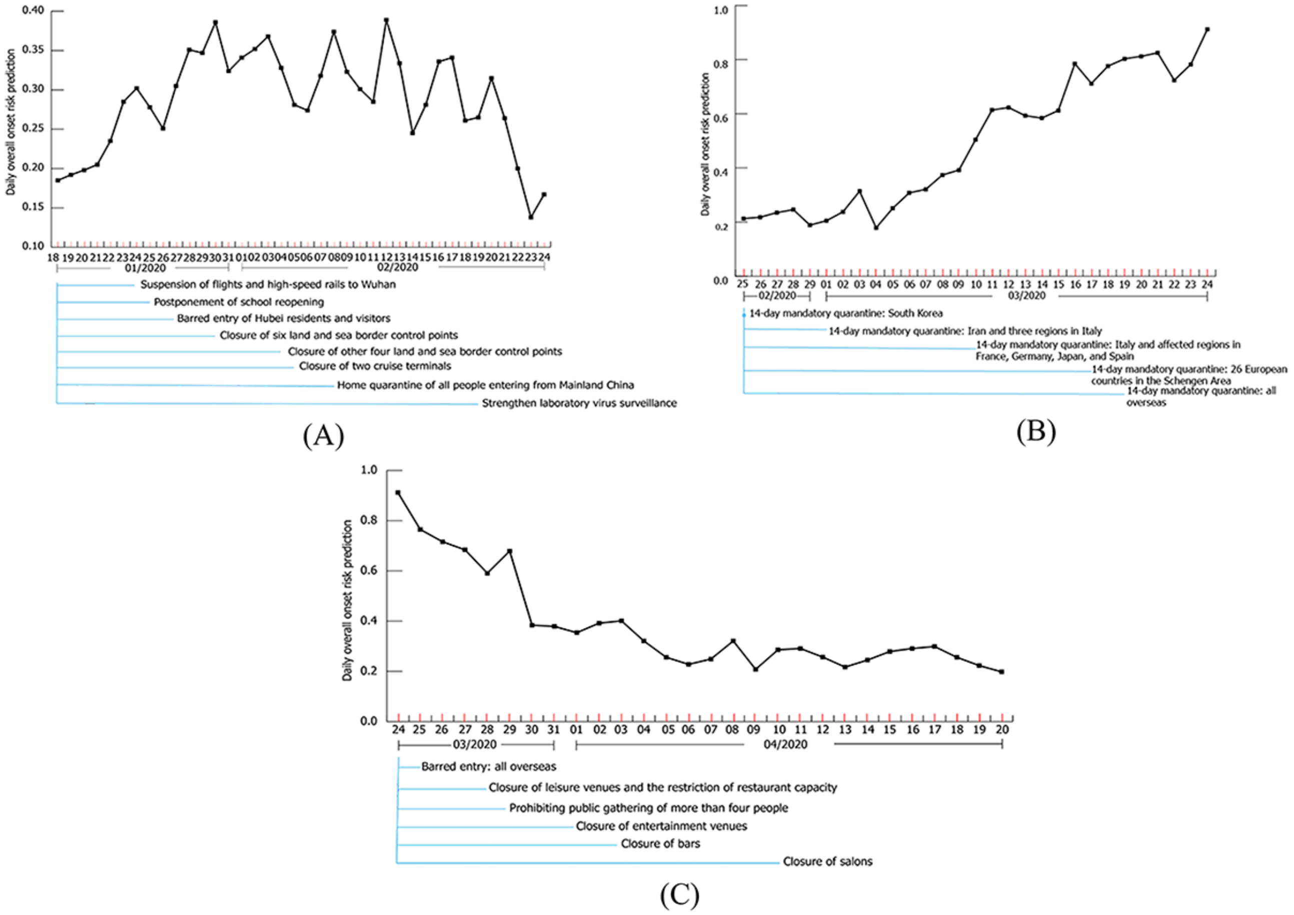
(C) January 15, 2021

Assistance in the Effect Analysis and Implementation of COVID-19 Response Measures

During the early stages of the epidemic, a series of previously untried COVID-19 response measures were implemented to suppress the spatiotemporal spread of SARS-CoV-2, including stringent border control measures to guard against imported cases and social distancing measures to prevent the local spread of the virus. Thus, it was necessary to evaluate the effect of these COVID-19 response measures on the spatiotemporal onset risk so as to adjust the COVID-19 response measures during subsequent stages. Thus, the effects of border control measures and social distancing measures already implemented in the first and second waves during the early stage of the epidemic were evaluated by the overall onset risk variations in 291 TPUs of Hong Kong within one week of the implementation (Figure 4). For border control measures, barred entry of all non-Hong Kong residents seemed to be the most effective during the peak period of the epidemic caused by imported cases. After its implementation, the overall onset risk was reduced by 28.82% within one week (Figure 4C). For local social distancing measures, prohibiting public gatherings of more than 4 people achieved the most significant reduction effect, by reducing the overall onset risk by 45.91% (Figure 4C). Other effective measures have been the closure of bars, the closure of leisure venues, the restriction of restaurant capacity, and the strengthening of laboratory virus surveillance, enabling the overall onset risk to be reduced by 33.33%, 30.17%, and 16.26%, respectively (Figure 4).

As the epidemic enters the normalization stage of prevention and control, COVID-19 response measures need to be more targeted than during the early stage of the pandemic. For the new response measures to be implemented, it is necessary to analyze the effects of new measures and assist in a more targeted implementation. For example, during the peak period of the recent fourth epidemic wave, the new measure—compulsory testing—was planned to be adopted from January 23, 2021. Before the implementation of the compulsory testing measures, based on the results of symptom onset risk prediction before January 23, 2021, “specified areas” with medium-high risk and “restricted areas” with high risk were quickly delineated by identifying the presence of TPUs within areas with medium-high onset risk and high onset risk (Figure 5). Moreover, in response to the public’s possible doubts about the controlling effects of the new measures, the simulation of the spatiotemporal effects of the new measures could also be provided, such as the simulation of the effect of the compulsory testing measure. The comparison of simulation results (Figure 6) showed that the compulsory testing not only suppressed the spread of SARS-CoV-2 in 291 TPUs but also reduced the onset risk related to each TPU. The onset risk was reduced by more than 80% in 42 (14.43%) TPUs and by more than 60% in 96 (32.99%) TPUs (Figure 6E; Table 1). Moreover, the compulsory testing was the most significant regarding the risk of those TPUs that had been originally at low or medium risk.

Figure 4. The average daily variation of the risk of COVID-19 symptom onset from January 18 to April 20, 2020; (A) average daily onset risk in all Hong Kong tertiary planning units (TPUs) from January 18 to February 24, 2020 (during this period, all COVID-19 response measures implemented were marked); (B) average daily onset risk in all Hong Kong TPUs from February 25 to March 24, 2020 (during this period, all COVID-19 response measures implemented were marked); (C) average daily onset risk in all Hong Kong TPUs from March 24 to April 20, 2020 (during this period, all COVID-19 response measures implemented were marked).



(A)

(B)

(C)

Figure 5. Predicted risk of COVID-19 symptom onset on January 23, 2021, for assisting the compulsory testing in Hong Kong.

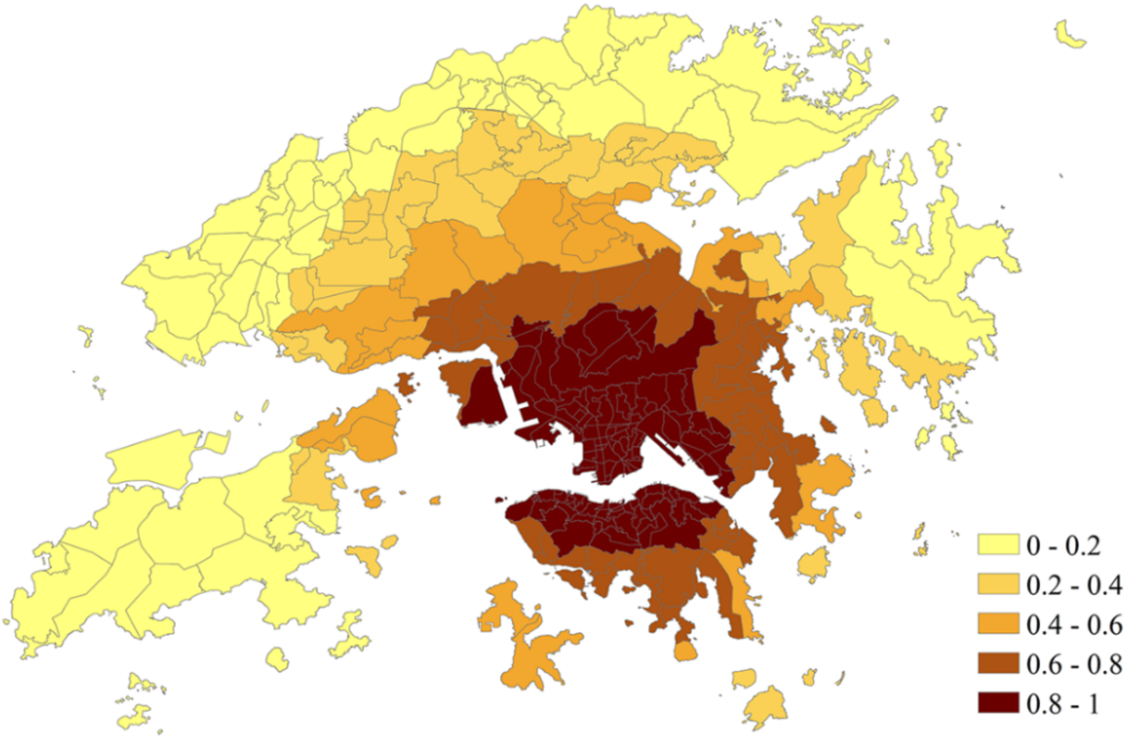
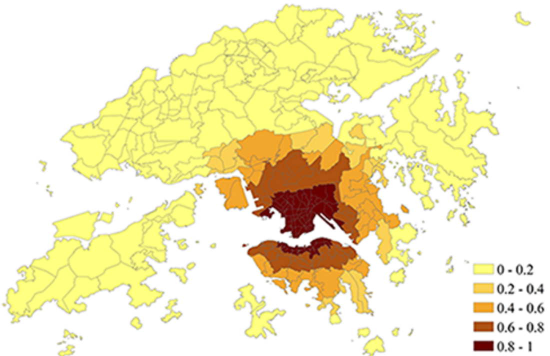
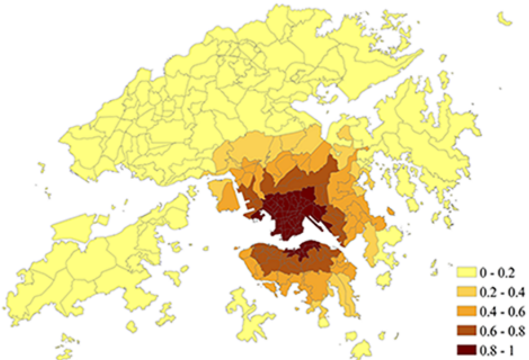


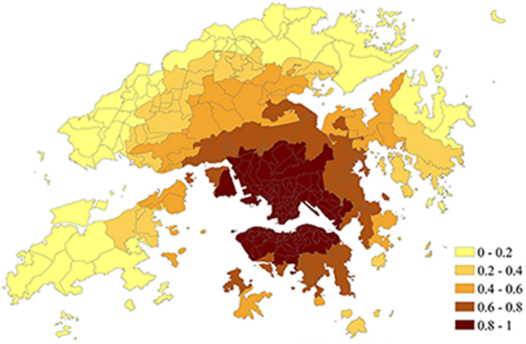
Figure 6. The risk of COVID-19 symptom onset under two scenarios (ie, with and without the compulsory testing measure) from January 24, 2021, to January 25, 2021. (A-D) The predicted symptom onset risk with (A-B) and without (C-D) the compulsory testing; (E) average daily percentage reduction in the onset risk in 291 tertiary planning units (TPUs) of Hong Kong in the compulsory testing scenario, compared with the noncompulsory testing scenario.



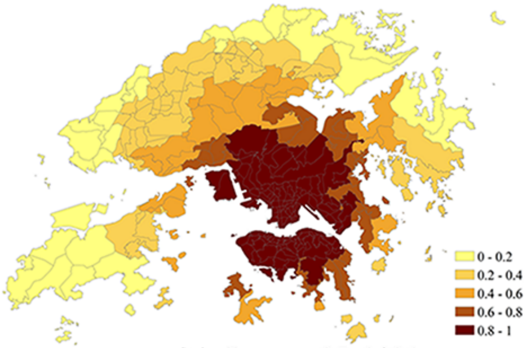
(A) January 24, 2021 with compulsory testing



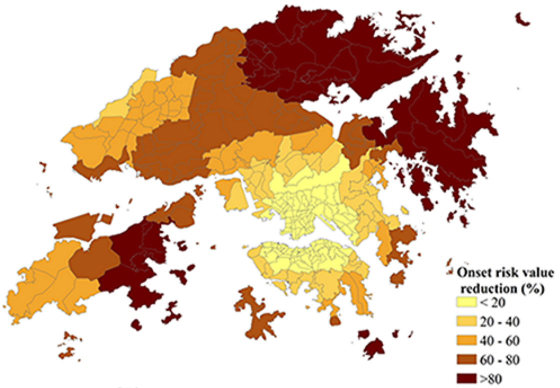
(B) January 25, 2021 with compulsory testing



(C) January 24, 2021 without compulsory testing



(D) January 25, 2021 without compulsory testing



(E) Average daily percentage reduction in the onset risk in 291 TPUs of Hong Kong

Table 1. The number and percentage of tertiary planning units (TPUs) at different reduction levels in the onset risk in the compulsory testing scenario, compared with the noncompulsory testing scenario.

Onset risk reduction (%)	Number of TPUs, n (%)
0-20	108 (37.11)
20-40	40 (13.75)
40 60	47 (16.15)
60-80	54 (18.56)
>80	42 (14.43)

The Developed Web-Based Platform to Inform Governmental Agencies and Populations Based on the Symptom Onset Risk Predictions

The precise prevention and control in Hong Kong rely on effective government response and public cooperation. Based on the above onset risk prediction and analysis results, a web-based platform has been developed in this study. Through the services of this platform, the government and the public can obtain the prediction and analysis results of the risk of COVID-19 symptom onset within the next three days. Advanced deployment of epidemic prevention measures can be achieved especially when large-scale events are held. In addition, for the public cooperation, an intuitive understanding of the onset risk predictions in the surrounding communities and potential travel communities helped in the adjustment of their daily activity plans in advance.

Discussion

Principal Findings

Global cities are taking swift actions in regard to building a sustainable postpandemic recovery [32]. COVID-19 is likely to stay with us as an endemic disease and cannot be eradicated globally or even eliminated regionally [33]. Thus, it is crucial for the recovery to identify resource mobilization needs, strengthening targeted measures and public cooperation in the new normalcy of long-term coexistence with SARS-CoV-2. Preemptively estimating the virus spread trend should become part of future epidemic prevention. Hence, in line with this aim, a spatiotemporal solution that supports more sustainable pandemic responses for returning to normalcy in the context of living with COVID-19 has been developed and is presented in this study. Since the COVID-19 epidemic outbreak, the proposed spatiotemporal solution has been supporting precise prevention and control in Hong Kong.

Firstly, the daily onset risk prediction results enables the spread of SARS-CoV-2 in communities to be timely assessed. It can therefore support the implementation of targeted and differentiated measures, especially at large events such as university entrance exams.

Secondly, under three main epidemiologic setting scenarios, high–onset-risk communities could be identified from the original daily onset risk prediction result and daily onset risk prediction enhanced by population distributions. It could limit the high–onset-risk communities in a small area in all three epidemiologic setting scenarios. The strictest COVID-19 response measures were only implemented in limited high–onset-risk communities to reduce the impact on normal social and economic activities.

Third, the effect of COVID-19 response measures already implemented by the government could be analyzed. These proven effective measures in this study have been playing a key role in Hong Kong's epidemic prevention. This was important in the early stage when the characteristics of the spread of SARS-CoV-2 variants were not clear. This could help the government adjust previous measures in subsequent stages.

Fourth, the implementation of the new COVID-19 response measures could be assisted, such as the new compulsory testing measure in the fourth epidemic wave. Before this measure was implemented first, the high–onset-risk restriction area could be determined through onset risk prediction results.

Fifth, in response to the public's possible doubts about the effects of the new COVID-19 response measures, the effects of the new COVID-19 response measures could be simulated. This is important with respect to the enhancement of policy transparency and also in building public trust in epidemic prevention.

Sixth, the prediction and set of analysis results related to onset risk have been integrated into a web-based visualized platform. Visualized information of spatiotemporal prediction and analysis enables the government and people to more easily and even more intuitively understand the nature of the potential spread of onset risk, leading to further understanding of symptom onset risk information related to each community. Logically, such understanding can better strengthen the appropriate response.

Comparison With Prior Work

To the best of our knowledge, this is the first attempt to use spatiotemporal onset cases data to explore the spatiotemporal variations of onset risk at the urban-community scale in the context of the emergence of different variants since the epidemic outbreak. It can further support effective epidemic control at the urban-community level for returning to normal. The Space-Time Scan Statistical Analysis Method [34], the improved Susceptible Exposure Infection and Recovery model [35-37], the spatiotemporal Bayesian inference model [38], and the hierarchical clustering analysis [39] were used in the previous studies to investigate the impact of the spatiotemporal spread of COVID-19. Consistent with the conclusions of these studies, results of this study also show that precise epidemic prevention and control measures can help effectively control the epidemic. Similarly, when epidemic prevention and control measures are relaxed, high-risk areas will also expand significantly, especially in the early stages of the epidemic. However, the vast majority of previous modeling studies have focused on the national, state, and regional levels [40]. The strength of this study is that building an onset risk prediction model at the urban-community level can facilitate accurate and differentiated public health responses at the finer spatial scale [40]. This allows the vast majority of areas in the city to maintain normal social and economic activities while maintaining strict prevention and control measures in a very small number of areas. Meanwhile, it can also facilitate the more effective allocation of limited medical resources to the limited areas with high onset risk that need it. In addition, this study simulates and predicts the spatiotemporal spread of COVID-19 under different measures or population mobility based on the symptom onset risk model. Hence, epidemic prevention measures and people's daily travel can be adjusted in advance in a timely manner. The proposed solution in this study can be used to support the enhancement need for community-based interventions and services in a multinational Delphi consensus to end the COVID-19 public health threat [41].

Limitations

We acknowledge the following potential limitations: First, limited by the information of the current official data, the location information of the cases used in this study only includes community-level locations of residence. If more information about the work locations and visited locations are combined, urban-community–level onset risk associated with their spatiotemporal activities can be estimated more comprehensively. Second, the applied data of human mobility mainly include the traffic flow data on the main roads, which fails to more comprehensively reflect the population flow under multiple modes of transportation. In addition, various data that can reflect viral load of SARS-CoV-2 at different locations, such as sewage monitoring, can be incorporated into our proposed solution in the future.

Conclusions

In the new normalcy of long-term coexistence with SARS-CoV-2, it is crucial to explore how to take the necessary response actions to reduce the impacts of the future epidemics and increase urban pandemic resilience for global cities [32]. Therefore, based on the enhanced prediction of COVID-19 symptom onset risk, a spatiotemporal solution has been proposed in this study. Such a solution has been used to support healthier and more sustainable pandemic responses in 291 communities of Hong Kong. The related web-based platform has served both the government agencies and Hong Kong residents. This solution not only assists global cities in responding to the future epidemic effectively but also helps to restore the social and economic activities and people’s normal lives. Based on this solution, we hope to make global cities more sustainable and hence promote the United Nations Sustainable Development Goals [42].

Acknowledgments

This study was supported by Hong Kong Research Grants Council (C5079-21G), the Innovation and Technology Commission, Hong Kong Special Administrative Regions (HKSAR) Government (ITP/041/21LP), Otto Poon Charitable Foundation Smart Cities Research Institute, and The Hong Kong Polytechnic University (work program CD03).

The authors would like to thank Katelyn Ma from York University for her advice during the preparation of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The technical details about the enhanced weighted kernel density estimation (WKDE) model and the Geodetector.
[DOCX File , 95 KB - publichealth_v9i1e36538_app1.docx]

References

1. Acuto M, Larcom S, Keil R, Ghojeh M, Lindsay T, Camponeschi C, et al. Seeing COVID-19 through an urban lens. Nat Sustain 2020 Sep 16;3(12):977-978. [doi: 10.1038/s41893-020-00620-3]
2. The global normalcy index. The Economist. URL: https://www.economist.com/graphic-detail/tracking-the-return-to-normalcy-after-covid-19 [accessed 2021-10-01]
3. Callaway E. The mutation that helps Delta spread like wildfire. Nature 2021 Aug;596(7873):472-473. [doi: 10.1038/d41586-021-02275-2] [Medline: 34417582]
4. Reardon S. How the Delta variant achieves its ultrafast spread. Nature 2021 Jul 21. [doi: 10.1038/d41586-021-01986-w] [Medline: 34290423]
5. Scudellari M. How the coronavirus infects cells - and why Delta is so dangerous. Nature 2021 Jul;595(7869):640-644. [doi: 10.1038/d41586-021-02039-y] [Medline: 34321669]
6. O'Dowd A. Covid-19: Cases of delta variant rise by 79%, but rate of growth slows. BMJ 2021 Jun 21;373:n1596. [doi: 10.1136/bmj.n1596] [Medline: 34154997]
7. Delta variant, a warning the COVID-19 virus is getting 'fitter and faster'. United Nations. URL: https://news.un.org/en/story/2021/07/1096792 [accessed 2021-10-01]
8. Mallapaty S. Delta threatens rural regions that dodged earlier COVID waves. Nature 2021 Aug;596(7872):325-326. [doi: 10.1038/d41586-021-02146-w] [Medline: 34354278]
9. Building double barriers against COVID-19. The Government of the Hong Kong Special Administrative Region. URL: https://www.ceo.gov.hk/eng/pdf/article20210827.pdf [accessed 2021-10-01]
10. Tong C, Shi W, Zhang A, Shi Z. Tracking and controlling the spatiotemporal spread of SARS-CoV-2 Omicron variant in South Africa. Travel Med Infect Dis 2022;46:102252 [FREE Full text] [doi: 10.1016/j.tmaid.2021.102252] [Medline: 34973454]
11. Countries / areas with reported cases of Coronavirus Disease-2019. The Government of the Hong Kong Special Administrative Region. URL: https://www.chp.gov.hk/files/pdf/statistics_of_the_cases_novel_coronavirus_infection_en.pdf [accessed 2021-10-01]

12. Fighting the epidemic for one year consolidating our experience. The Government of the Hong Kong Special Administrative Region. URL: https://www.ceo.gov.hk/eng/pdf/article20210127.pdf [accessed 2021-10-01]
13. A new direction in fighting the pandemic. The Government of the Hong Kong Special Administrative Region. URL: https://www.ceo.gov.hk/eng/pdf/article20210426.pdf [accessed 2021-10-01]
14. Tong C, Shi W, Zhang A, Shi Z. Tracking and controlling the spatiotemporal spread of SARS-CoV-2 lineage B.1.1.7 in COVID-19 reopenings. Geohealth 2021 Dec;5(12):e2021GH000517 [FREE Full text] [doi: 10.1029/2021GH000517] [Medline: 34938933]
15. Shi W, Tong C, Zhang A, Wang B, Shi Z, Yao Y, et al. An extended Weight Kernel Density Estimation model forecasts COVID-19 onset risk and identifies spatiotemporal variations of lockdown effects in China. Commun Biol 2021 Jan 25;4(1):126 [FREE Full text] [doi: 10.1038/s42003-021-01677-2] [Medline: 33495509]
16. Shi W, Tong C, Zhang A, Shi Z. A spatial and dynamic solution for allocation of COVID-19 vaccines when supply is limited. Commun Med (Lond) 2021;1:23 [FREE Full text] [doi: 10.1038/s43856-021-00023-1] [Medline: 35602195]
17. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. Proc Natl Acad Sci U S A 2020 May 12;117(19):10484-10491. [doi: 10.1073/pnas.2004978117] [Medline: 32327608]
18. Rader B, Scarpino SV, Nande A, Hill AL, Adlam B, Reiner RC, et al. Crowding and the shape of COVID-19 epidemics. Nat Med 2020 Dec;26(12):1829-1834. [doi: 10.1038/s41591-020-1104-0] [Medline: 33020651]
19. Kawasuji H, Takegoshi Y, Kaneda M, Ueno A, Miyajima Y, Kawago K, et al. Transmissibility of COVID-19 depends on the viral load around onset in adult and symptomatic patients. PLoS One 2020;15(12):e0243597 [FREE Full text] [doi: 10.1371/journal.pone.0243597] [Medline: 33296437]
20. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020 May;26(5):672-675. [doi: 10.1038/s41591-020-0869-5] [Medline: 32296168]
21. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat Med 2020 Nov;26(11):1714-1719. [doi: 10.1038/s41591-020-1092-0] [Medline: 32943787]
22. Leung K, Wu JT, Leung GM. Real-time tracking and prediction of COVID-19 infection using digital proxies of population mobility and mixing. Nat Commun 2021 Mar 08;12(1):1501 [FREE Full text] [doi: 10.1038/s41467-021-21776-2] [Medline: 33686075]
23. Riccardo F, Ajelli M, Andrianou XD, Bella A, Del Manso M, Fabiani M, COVID-19 working group. Epidemiological characteristics of COVID-19 cases and estimates of the reproductive numbers 1 month into the epidemic, Italy, 28 January to 31 March 2020. Euro Surveill 2020 Dec;25(49) [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.49.2000790] [Medline: 33303064]
24. Hu B, Guo H, Zhou P, Shi Z. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 2021 Mar;19(3):141-154 [FREE Full text] [doi: 10.1038/s41579-020-00459-7] [Medline: 33024307]
25. Onset cases in Hong Kong. Centre for Health Protection. URL: https://www.chp.gov.hk/en/media/116/index.html# [accessed 2021-10-01]
26. Chan W, Ip JD, Chu AW, Tse H, Tam AR, Li X, et al. Phylogenomic analysis of COVID-19 summer and winter outbreaks in Hong Kong: An observational study. Lancet Reg Health West Pac 2021 May;10:100130 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100130] [Medline: 33778795]
27. Report on evaluation of safety , efficacy and quality of comirnaty COVID-19 mRNA vaccine (BNT162b2) concentrate for dispersion for injection. Food and Health Bureau of Hong Kong. URL: https://www.fhb.gov.hk/download/our_work/health/201200/e_evaluation_report.pdf [accessed 2021-10-01]
28. Report on evaluation of safety, efficacy and quality of CoronaVac COVID-19 vaccine (Vero Cell) inactivated. Food and Health Bureau of Hong Kong. URL: https://www.fhb.gov.hk/download/our_work/health/201200/e_evaluation_report_CoronaVac.pdf [accessed 2021-10-01]
29. Hong Kong vaccination dashboard. The Government of the Hong Kong Special Administrative Region. URL: https://www.covidvaccine.gov.hk/en/dashboard [accessed 2021-10-01]
30. 737 closed circuit televisions (CCTV) and traffic detectors in Hong Kong. Transport Department of Hong Kong. URL: https://data.gov.hk/en-data/dataset/hk-td-tis_2-traffic-snapshot-images [accessed 2021-10-01]
31. Statistics on passenger traffic. Immigration Department. URL: https://www.immd.gov.hk/eng/message_from_us/stat_menu.html [accessed 2021-10-01]
32. Contestabile M. A sustainable post-COVID future. Nat Sustain 2021 Mar 15;4(6):464-465. [doi: 10.1038/s41893-021-00703-9]
33. Telenti A, Arvin A, Corey L, Corti D, Diamond MS, García-Sastre A, et al. After the pandemic: perspectives on the future trajectory of COVID-19. Nature 2021 Aug;596(7873):495-504. [doi: 10.1038/s41586-021-03792-w] [Medline: 34237771]
34. Luo W, Liu Z, Zhou Y, Zhao Y, Li YE, Masrur A, et al. Investigating linkages between spatiotemporal patterns of the COVID-19 Delta variant and public health interventions in Southeast Asia: prospective space-time scan statistical analysis method. JMIR Public Health Surveill 2022 Aug 09;8(8):e35840 [FREE Full text] [doi: 10.2196/35840] [Medline: 35861674]

35. Shapiro MB, Karim F, Muscioni G, Augustine AS. Adaptive Susceptible-Infectious-Removed Model for continuous estimation of the COVID-19 infection rate and reproduction number in the United States: modeling study. J Med Internet Res 2021 Apr 07;23(4):e24389 [FREE Full text] [doi: 10.2196/24389] [Medline: 33755577]
36. Lin R, Lin S, Yan N, Huang J. Do prevention and control measures work? Evidence from the outbreak of COVID-19 in China. Cities 2021 Nov;118:103347 [FREE Full text] [doi: 10.1016/j.cities.2021.103347] [Medline: 34312572]
37. Khadadah F, Al-Shammari AA, Alhashemi A, Alhuwail D, Al-Saif B, Alzaid SN, et al. The effects of non-pharmaceutical interventions on SARS-CoV-2 transmission in different socioeconomic populations in Kuwait: a modeling study. BMC Public Health 2021 May 26;21(1):990 [FREE Full text] [doi: 10.1186/s12889-021-10984-6] [Medline: 34039289]
38. Ge Y, Zhang W, Liu H, Ruktanonchai CW, Hu M, Wu X, et al. Impacts of worldwide individual non-pharmaceutical interventions on COVID-19 transmission across waves and space. Int J Appl Earth Obs Geoinf 2022 Feb;106:102649 [FREE Full text] [doi: 10.1016/j.jag.2021.102649] [Medline: 35110979]
39. Castro MC, Kim S, Barberia L, Ribeiro AF, Gurzenda S, Ribeiro KB, et al. Spatiotemporal pattern of COVID-19 spread in Brazil. Science 2021 May 21;372(6544):821-826 [FREE Full text] [doi: 10.1126/science.abh1558] [Medline: 33853971]
40. Zhou S, Zhou S, Zheng Z, Lu J, Song T. Risk assessment for precise intervention of COVID-19 epidemic based on available big data and spatio-temporal simulation method: Empirical evidence from different public places in Guangzhou, China. Appl Geogr 2022 Jun;143:102702 [FREE Full text] [doi: 10.1016/j.apgeog.2022.102702] [Medline: 35469327]
41. Lazarus JV, Romero D, Kopka CJ, Karim SA, Abu-Raddad LJ, Almeida G, COVID-19 Consensus Statement Panel. A multinational Delphi consensus to end the COVID-19 public health threat. Nature 2022 Nov;611(7935):332-345 [FREE Full text] [doi: 10.1038/s41586-022-05398-2] [Medline: 36329272]
42. The sustainable development goals report 2021. United Nations. URL: https://unstats.un.org/sdgs/report/2021/ [accessed 2021-10-01]

Abbreviations

TPU: tertiary planning unit
WKDE: weighted kernel density estimation

Edited by T Sanchez, A Mavragani; submitted 17.01.22; peer-reviewed by J Wang, S Pesälä; comments to author 10.05.22; revised version received 27.05.22; accepted 09.12.22; published 06.01.23.

Please cite as:
Tong C, Shi W, Zhang A, Shi Z
A Spatiotemporal Solution to Control COVID-19 Transmission at the Community Scale for Returning to Normalcy: COVID-19 Symptom Onset Risk Spatiotemporal Analysis
JMIR Public Health Surveill 2023;9:e36538
URL: https://publichealth.jmir.org/2023/1/e36538
doi:10.2196/36538
PMID:36508488

Original Paper

Resident Willingness to Participate in Digital Contact Tracing in a COVID-19 Hotspot: Findings From a Detroit Panel Study

Lydia Wileden[1,2], PhD; Denise Anthony[3,4,5], PhD; Celeste Campos-Castillo[6], PhD; Jeffrey Morenoff[5,7,8], PhD

[1]Mansueto Institute for Urban Innovation, University of Chicago, Chicago, IL, United States
[2]Division of the Social Sciences, University of Chicago, Chicago, IL, United States
[3]Department of Health Management & Policy, School of Public Health, University of Michigan, Ann Arbor, MI, United States
[4]School of Information, University of Michigan, Ann Arbor, MI, United States
[5]Department of Sociology, University of Michigan, Ann Arbor, MI, United States
[6]Department of Media & Information, Michigan State University, East Lansing, MI, United States
[7]Gerald R. Ford School of Public Policy, University of Michigan, Ann Arbor, MI, United States
[8]Population Studies Center, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Lydia Wileden, PhD
Mansueto Institute for Urban Innovation
University of Chicago
1155 E 60th St
Chicago, IL
United States
Phone: 1 773 702 7894
Email: lwileden@uchicago.edu

Abstract

Background: Digital surveillance tools and health informatics show promise in counteracting diseases but have limited uptake. A notable illustration of the limits of such tools is the general failure of digital contact tracing in the United States in response to COVID-19.

Objective: We investigated the associations between individual characteristics and the willingness to use app-based contact tracing in Detroit, a majority-minority city that experienced multiple waves of COVID-19 outbreaks and deaths since the start of the pandemic. The aim of this study was to examine variations among residents in the willingness to download a contact tracing app on their phones to provide public health officials with information about close COVID-19 contact during summer 2020.

Methods: To examine residents' willingness to participate in digital contact tracing, we analyzed data from 2 waves of the Detroit Metro Area Communities Study, a population-based survey of Detroit, Michigan residents. The data captured 1873 responses from 991 Detroit residents collected in June and July 2020. We estimated a series of multilevel logit models to gain insights into differences in the willingness to participate in digital contact tracing across a variety of individual attributes, including race/ethnicity, degree of trust in the government, and level of education, as well as interactions among these variables.

Results: Our results reflected widespread reluctance to participate in digital contact tracing in response to COVID-19, as less than half (826/1873, 44.1%) of the respondents said they would be willing to participate in app-based contact tracing. Compared to White respondents, Black (odds ratio [OR] 0.45, 95% CI 0.23-0.86) and Latino (OR 0.32, 95% CI 0.11-0.99) respondents were significantly less willing to participate in digital contact tracing. Trust in the government was positively associated with the willingness to participate in digital contact tracing (OR 1.17, 95% CI 1.07-1.27), but this effect was the strongest for White residents (OR 2.14, 95% CI 1.55-2.93). We found similarly divergent patterns of the effects of education by race. While there were no significant differences among noncollege-educated residents, White college-educated residents showed greater willingness to use app-based contact tracing (OR 6.12, 95% CI 1.86-20.15) and Black college-educated residents showed less willingness (OR 0.46, 95% CI 0.26-0.81).

Conclusions: Trust in the government and education contribute to Detroit residents' wariness of digital contact tracing, reflecting concerns about surveillance that cut across race but likely arise from different sources. These findings point to the importance of a culturally informed understanding of health hesitancy for future efforts hoping to leverage digital contact tracing. Though contact tracing technologies have the potential to advance public health, unequal uptake may exacerbate disparate impacts of health crises.

(JMIR Public Health Surveill 2023;9:e39002) doi:10.2196/39002

KEYWORDS

COVID-19; contact tracing; surveillance; informatics; trust; racial disparities

Introduction

Throughout the COVID-19 pandemic, public health authorities have engaged in contact tracing to inform those exposed to COVID-19, monitor for signs and symptoms of the virus, aid in access to testing, and promote self-quarantine as necessary [1]. While these efforts generally involve traditional direct outreach by public health workers, technologists and public health advocates have highlighted the potential of smartphone-based digital technologies to assist in contact tracing [2,3]. A May 2020 guidance from the Centers for Disease Control and Prevention (CDC) acknowledged that digital contact tracing tools for COVID-19 can function as a “supplement to traditional public health contact tracing with the use of voluntary opt-in proximity or exposure notification tools” [4].

The proliferation of digital contact tracing apps in response to COVID-19 builds on advancements in health informatics and the rise of health apps and eHealth over the past decade [5,6]. This expansion of digital health surveillance tools has a variety of potential benefits, including aiding in managing health conditions [7-9], modifying health behaviors [5,10], and helping facilitate racial equity in health information and health care [11,12]. Additionally, recent studies have found that digital health tools with exposure notifications are effective in decreasing infections from the coronavirus [13,14].

Despite the promise of digital surveillance tools to mitigate the spread of COVID-19, app-based contact tracing has had limited success in the United States [15,16]. As with all contact tracing approaches, the success of app-based contact tracing depends on individuals’ willingness to participate. Such willingness appears to be low [17,18]. In late 2020, Pew Research Center found that about 40% of Americans said they were unlikely to talk to contact tracers, and 50% said they were uncomfortable sharing location data from their phones [19].

Studies on public engagement with digital contact tracing around the world have proliferated since the start of the COVID-19 pandemic [13,20-24]. However, fewer studies have focused on the United States and drawn on representative survey samples. A notable exception is the work of Camacho-Rivera et al [18], who conducted a representative survey of US households between April and June 2020 to examine attitudes toward mobile health tools for COVID, including location-based exposure tracking. They found that although more than half of the respondents reported they were not likely to use mobile health tools for COVID, some sociodemographic and health-related factors were associated with more willingness to engage with digital contact tracing. For example, non-White respondents were more likely than their counterparts to say they were willing to use app-based location exposure tracking for COVID [18]. Additionally, systematic reviews of engagement with contact tracing during COVID (and some other contagious health outbreaks) across various countries have observed a number of barriers to uptake, including concerns about privacy and surveillance, mistrust of the government, and mistrust of technology [20,21]. Thus, while past research has pointed to several factors that influence contact tracing, limited research has considered how these factors interact in meaningful ways.

Gaining insights into differences regarding who is willing to participate in contact tracing may be especially consequential given the disproportionate impact of COVID-19 on racial/ethnic minority populations in the United States [25]. Greater willingness or hesitancy to participate in contact tracing among minoritized groups could affect the ability to stop the virus’ localized spread and could heighten the risks of morbidity and mortality in communities of color. Related research on the limited uptake of digital health technologies suggests that differences in uptake, especially along racial/ethnic and socioeconomic lines, can exacerbate health disparities [26,27]. For example, some studies on the adoption of telehealth technologies indicate that Black and Latino patients are less likely to use telehealth than White patients, limiting their access to care [28,29]. However, other studies have reported opposite findings, with Black and Latino patients being more likely to use telehealth than White patients [30,31]. Additionally, there is evidence that individuals with lower levels of education are less likely to use digital health platforms and services compared to those with more than a high school education [18,26,32].

As noted in other studies, willingness to use digital health technologies in general and app-based contact tracing in particular is likely undermined by issues of trust. Studies have demonstrated that individuals’ trust in the medical system influences their use of health services [33,34]. Lack of trust among racial/ethnic minorities, including lower trust toward medical providers, research, and treatments, is commonly cited as a key mechanism underlying disparities in health care [24,35-37]. Such differences in trust in the health care system may be further exacerbated by concerns around privacy and security, which can adversely affect patients’ willingness to use eHealth tools and provide accurate health information to providers [38,39]. Prior research shows that there are socioeconomic differences in trust, as adults with a college degree are more likely than those with lower educational attainment to report avoiding products or services because of privacy concerns [40]. Further, widespread uptake of contact tracing may be hampered by concerns regarding data privacy and trust in the government [20-22,41]. Recent evidence suggests that contact tracing efforts face greater resistance in communities where people have low levels of trust in public officials or are worried about state surveillance [42,43].

Given previous findings regarding the role of trust and mixed findings on the differential uptake of health technologies, including COVID-19 apps, we set out to determine the individual factors associated with the willingness to participate in digital contact tracing in response to COVID-19 in the US city of Detroit, Michigan. By focusing on Detroit, we highlight the attitudes of residents in a majority-minority city that experienced multiple waves of COVID-19 outbreaks and deaths since the start of the pandemic. In Detroit, Black residents have been disproportionately impacted by COVID-19, representing over 75% of known diagnoses and 90% of deaths, and they are 10 times more likely than White residents to have friends or family who have died from COVID-19 [44]. Using data from a population-based survey of Detroit, Michigan residents, we examined not only how the factors of race/ethnicity, education, and trust are associated with individuals' willingness to download a contact tracing app during summer 2020, but also how those factors are intertwined in important ways. The findings of this study can help shed light on what factors need to be considered to encourage use of digital public health tools in the future, including how various subpopulations may respond to messages about or sources of those tools.

Methods

Data Source

To examine individual willingness to participate in app-based contact tracing, we analyzed data from 2 surveys conducted by the Detroit Metro Area Communities Study (DMACS) in 2020 (one in June and the other in July). DMACS is a panel study designed to regularly capture the perspectives and behaviors of a representative sample of Detroit adults [45,46]. DMACS launched in 2016 and recruited respondents from an address-based probability sample of Detroit households. It has refreshed its sample approximately once a year through additional address-based probability sampling. In response to the March 2020 declaration of COVID-19 as a national emergency and the emergence of Detroit as a "hotspot" of coronavirus cases, DMACS initiated a series of rapid response surveys on residents' experiences with COVID-19 [44,47,48]. Due to restrictions on human subject interactions during the pandemic, participation in these surveys was limited to survey panelists who had provided email addresses or phone numbers prior to the COVID-19 pandemic. In the June survey (fielded between May 28 and June 11, 2020), 1802 panelists were invited and 1173 completed surveys (66.1% response rate). In the July survey (fielded between July 15 and July 29, 2020), 1772 panelists were invited and 1137 completed surveys (64.2% response rate). In both waves, surveys were either self-administered online by respondents or interview administered by telephone. We limited the current analysis to panelists who provided complete responses to the items measuring key independent and dependent variables in the June and July surveys. In total, our results reflected 1873 responses from 991 Detroit residents.

Ethics Approval

Human subjects research was approved by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (HUM00187155).

Measures

Our dependent variable captured respondent willingness to participate in app-based contact tracing with the item, "Would you be willing to download an app on your cell phone that would provide information to public health officials about close contact you have with other people?" Respondents could answer yes, no, don't know, or "I do not own a cell phone." Because our primary interest was to examine what distinguishes those who are willing to participate in app-based contact tracing, we dichotomized responses to compare those who said yes to those who said no or don't know. We dropped the small number (n=18) of respondents who reported not owning a cell phone.

We examined the willingness to engage in app-based contact tracing in relation to self-reported race/ethnicity, gender, age, education, and income. We used multiple imputation with chained equations (MICE, implemented with Stata 17 software, StataCorp) and Rubin's combination rules to fill in missing values of these and other independent variables to avoid dropping respondents from the analyses [49,50]. To ensure sufficient data coverage, we restricted our analyses to Detroit's 3 largest racial/ethnic groups: non-Latino Black (hereafter Black), non-Latino White (hereafter White), and Latino. We defined "Latino" as any respondent who self-identified as being of "Hispanic, Latino, or Spanish origin," including those who selected another racial identity and Latino. "White" and "Black" refer to respondents who selected only those respective categories (and no other ethnoracial categories) and who did not identify as Latino. Respondents who identified on the survey using other ethnoracial categories alone or in combination (Asian, Native American, Hawaiian or Pacific Islander, multi-race, or other) were omitted from this analysis owing to their small sample sizes and to avoid combining segments of unlike populations. Age was captured as a continuous variable. Categories were created for annual household income (<US $10,000, US $10,000-29,999, US $30,000-49,999, US $50,000-100,000, and >US $100,000). Because income was commonly missing and unwillingness to provide income information may be associated with data privacy concerns and unwillingness to provide other personal information relevant to contact tracing, we included a dummy variable in our analysis that captured whether the respondents withheld their income data. Education level (less than college degree, or college degree or higher) and gender (male or female) were captured as binary variables.

Finally, we included in our analyses several measures of respondents' beliefs about and experiences with COVID-19. Specifically, we included a measure of respondent trust in the government (adapted from the ICF COVID Monitor Survey of US Adults [51]). The question asked was as follows: "How much do you trust the Michigan state government to deal with the COVID-19 pandemic, on a scale of 1 to 10, where 1 is not at all and 10 is entirely?" Responses were captured on a 10-point Likert scale and were mean-centered for interpretability. We focused on this measure of trust in the state government because our outcome variable made reference to sharing information with "public health officials," and most local COVID health efforts and digital contact tracing tools were coordinated at the state level at this time. We also included, as dichotomous variables, measures of whether the respondent had a close friend or family member who died from the coronavirus; had a close friend or family member who got ill from the coronavirus; had been diagnosed with COVID-19; and would say the COVID-19 pandemic is very serious for them personally. Because our outcome variable was focused on the use of a contact tracing app, we included a binary variable for the availability of reliable internet access as a proxy for technological use and access. We also controlled for the effect of time by including a binary variable for survey wave. We did not use sampling weights because they are incompatible with the multiple imputation package of Stata 17 when estimating multi-level models, but we controlled for survey sampling strata and individual characteristics used in the creation of sample weights.

Statistical Analyses

Because we were analyzing panel data with multiple observations per person, we used random effects logit models, which introduce a random intercept to help correct for person-level unobserved heterogeneity. We estimated a series of models to gain insights into respondents' willingness to engage in app-based contact tracing. Our first model examined differences in the willingness to contact trace across a variety of individual attributes, including race/ethnicity, degree of trust in the government, and education. Our second and third models added interactions among these key variables to examine if the effects of trust and level of education vary by race/ethnicity of the respondent. We tested a 3-way interaction between race/ethnicity, education, and trust in the government but found that the results were substantively similar to the results for the interaction between education and race/ethnicity, and thus, we omitted these results for parsimony.

Results

Descriptive Statistics

Summary statistics for respondents are presented in Table 1. Less than half (826/1873, 44.1%) of the respondents reported they would be willing to participate in app-based contact tracing. Black residents comprised the majority of our sample (1486/1873, 79.3%), while 14.2% (265/1873) of our sample was White and 6.5% (122/1873) was Latino, similar to the proportions in the city of Detroit. On average, DMACS respondents reported a relatively high degree of trust in the state government to deal with COVID-19 (7.25 out of 10 points). The majority of respondents did not attend college (1301/1873, 69.5%) and had an annual household income below US $50,000 (1219/1772, 68.8%). The average respondent was 49 years old. Like many surveys, our sample had a higher proportion of female respondents (1372/1873, 73.3%). Our respondents reflected the severe impact of the COVID-19 pandemic on those living in Detroit, with 71.4% (1337/1873) saying that the pandemic was very serious for them personally and many saying that they had a close friend or family member who got ill (1088/1870, 58.2%) or died (745/1871, 39.8%) from COVID-19. However, only 5.5% (102/1868) of respondents said that they had been diagnosed with COVID-19 as of July 2020. In general, respondents were technologically connected, with 86.6% (1622/1873) reporting that they had reliable access to the internet at home.

Table 1. Descriptive statistics of the analytical sample.

Variable	Value (N=1873)
Willing to participate in contact tracing, n (%)	826 (44.1)
Race/ethnicity, n (%)	
White	265 (14.2)
Black	1486 (79.3)
Latino	122 (6.5)
Trust in the state government (score range 0-10), mean (SD)	7.25 (0.25)
Education, n (%)	
Less than college	1301 (69.5)
College	572 (30.5)
Gender, n (%)	
Male	501 (26.8)
Female	1372 (73.3)
Age (years)[a]	
Mean (SD)	48.96 (15.44)
Range	18.16-93.43
Income (US$)[a], n (%)	
<10,000	380 (21.4)
10,000-29,999	441 (24.9)
30,000-49,999	398 (22.5)
50,000-100,000	405 (22.9)
>100,000	148 (8.4)
Income missing	101 (5.4)
Family member died from COVID[a], n (%)	745 (39.8)
Family member got ill from COVID[a], n (%)	1088 (58.2)
Pandemic was very serious personally, n (%)	1337 (71.4)
Diagnosed with COVID[a], n (%)	102 (5.5)
Internet use, n (%)	1622 (86.6)
Wave, n (%)	
June 2020 (wave 9)	987 (52.7)
July 2020 (wave 10)	886 (47.3)

[a] Variable imputed with multiple imputations. Within the data set of 1873 observations, income was missing in 101 cases, age was missing in 18 cases, information on whether a family member died from COVID was missing in 2 cases, information on whether a family member got ill from COVID was missing in 3 cases, and information on diagnosis of COVID was missing in 5 cases.

Analytic Results

Results from our models are presented in Table 2. Model 1 examined the relationships of individuals' demographics, trust in the government, and COVID-19 experiences with their willingness to download a contact tracing app. The results showed that individual characteristics were strong predictors of the willingness to participate in app-based contact tracing. Compared with White respondents, Black (odds ratio [OR] 0.45, 95% CI 0.23-0.86) and Latino (OR 0.32, 95% CI 0.11-0.99) respondents had significantly lower odds of being willing to participate in contact tracing. Similarly, the odds of women saying they would participate in app-based contact tracing (OR 0.44, 95% CI 0.27-0.70) were about half of the odds of men. The willingness to participate in contact tracing was associated with trust, whereby higher levels of trust in the state government's management of COVID-19 were associated with greater odds of downloading a contact tracing app (OR 1.17, 95% CI 1.07-1.27). Furthermore, respondents who reported their household income had roughly 4 times the odds of being willing to participate in contact tracing as those who elected to not report their income (OR 0.24, 95% CI 0.09-0.63), which is perhaps a measure of the wariness about sharing one's personal information. While it might be expected that individual experience with COVID-19 (being sick, having friends or family who have gotten sick or died from the virus, or feeling the pandemic is personally very serious) would increase the willingness to participate in preventative measures, including contact tracing, we found that only those who lost a close friend or family member to COVID-19 showed significantly greater odds of participating in app-based contract tracing (OR 1.72, 95% CI 1.12-2.62). Age, level of income, access to the internet, and time were not significant predictors of an individual's willingness to participate in contact tracing via an app.

Table 2. Predictors of the willingness to participate in app-based contact tracing among the respondents.

Variable	Model 1[a], OR[b] (95% CI)	Model 2[a], OR (95% CI)	Model 3[a], OR (95% CI)
Race/ethnicity			
White	Reference	Reference	Reference
Black	0.45 (0.23-0.86)	0.47 (0.24-0.92)	1.70 (0.67-4.30)
Latino	0.32 (0.11-0.99)	0.36 (0.12-1.10)	0.73 (0.21-2.62)
Trust in the state government (centered)	1.17 (1.07-1.27)	2.14 (1.55-2.93)	1.16 (1.07-1.26)
Race × trust			
White × trust	N/A[c]	Reference	N/A
Black × trust	N/A	0.51 (0.37-0.71)	N/A
Latino × trust	N/A	0.56 (0.36-0.87)	N/A
Education			
Less than college	Reference	Reference	Reference
College	0.73 (0.43-1.23)	0.73 (0.43-1.22)	6.12 (1.86-20.15)
Race × education			
White × college	N/A	N/A	Reference
Black × college	N/A	N/A	0.08 (0.02-0.27)
Latino × college	N/A	N/A	0.92 (0.98-1.01)
Gender			
Male	Reference	Reference	Reference
Female	0.44 (0.27-0.70)	0.43 (0.27-0.69)	0.44 (0.27-0.70)
Age[d]	0.99 (0.97-1.00)	0.99 (0.98-1.00)	0.99 (0.98-1.00)
Income (US$)[d]			
<10,000	Reference	Reference	Reference
10,000-29,999	1.47 (0.80-2.71)	1.47 (0.80-2.68)	1.54 (0.84-2.80)
30,000-49,999	0.79 (0.41-1.50)	0.79 (0.42-1.50)	0.79 (0.42-1.50)
50,000-100,000	0.65 (0.32-1.31)	0.61 (0.31-1.22)	0.69 (0.35-1.38)
>100,000	1.00 (0.37-2.72)	0.99 (0.37-2.67)	0.80 (0.30-2.18)
Income missing	0.24 (0.09-0.63)	0.23 (0.09-0.59)	0.28 (0.11-0.72)
Family/friend died from COVID[d]	1.72 (1.12-2.62)	1.73 (1.14-2.64)	1.79 (1.17-2.73)
Family/friend got ill from COVID[d]	0.79 (0.53-1.18)	0.79 (0.53-1.17)	0.76 (0.51-1.13)
Pandemic was very serious personally	1.30 (0.88-1.91)	1.33 (0.90-1.96)	1.36 (0.92-2.01)
Diagnosed with COVID[d]	1.04 (0.46-2.32)	1.07 (0.48-2.38)	1.08 (0.49-2.40)
Internet use	1.33 (0.73-2.44)	1.27 (0.70-2.32)	1.38 (0.76-2.50)
Wave			
June 2020 (wave 9)	Reference	Reference	Reference
July 2020 (wave 10)	0.89 (0.68-1.14)	0.88 (0.68-1.14)	0.88 (0.68-1.14)
Constant	2.93 (0.44-19.45)	2.95 (0.45-19.17)	0.71 (0.10-5.18)

[a] Results from multilevel logistic regression with random effects. We controlled for sampling strata in our models, but we have not presented the nonsignificant effect of these strata to keep the table short. Data reflect 1873 responses nested within 991 respondents across 2 survey waves.

[b] OR: odds ratio.

[c] N/A: not applicable.

[d] Variable imputed with multiple imputation. Sensitivity analyses without imputation reach substantively similar conclusions as those of the imputed models.

To understand what may be driving racial/ethnic differences in the willingness to participate in contact tracing, subsequent models estimated interactions between race/ethnicity and trust (Model 2), and race/ethnicity and education (Model 3). By allowing the effects of race to differ by the degree of trust and level of education, we can better understand whether resistance to contact tracing among racial/ethnic groups is uniform or varies within subgroups.

Examining the interaction between race/ethnicity and trust (Model 2), we found that for White respondents, higher levels of trust in the state government's management of COVID-19 were associated with greater likelihood of being willing to participate in contact tracing via an app (OR 2.14, 95% CI 1.55-2.93). This relationship was significantly weaker for Black and Latino respondents (Black: OR 0.51, 95% CI 0.37-0.71; Latino: OR 0.56, 95% CI 0.36-0.87). Figure 1 illustrates the relationship between race/ethnicity, trust, and contact tracing, showing the strong positive relationship between trust and the willingness to participate in contact tracing for White respondents and the negligible effect of trust on the willingness to participate in contact tracing for other respondents. Model 2 results also show that the effect of trust does not fully explain differences in the willingness to participate in contact tracing by race/ethnicity. Black respondents with an average degree of trust in the government remained significantly less likely than White respondents to be willing to engage in app-based contact tracing (OR 0.47, 95% CI 0.24-0.92). We also continued to find that women and those who did not report income were less likely to participate in contact tracing, while those who had a close acquaintance die of COVID remained more likely to participate.

Turning to the interaction between race/ethnicity and education (Model 3), the main effects showed no significant difference in the willingness to participate in contact tracing among noncollege-educated respondents by race. However, the interaction showed that the effect of college education differs significantly depending on whether the respondent is Black or White. The diverging effect of education on the likelihood to engage in app-based contact tracing among White and Black respondents is illustrated in Figure 2. Among White respondents, the probability of participating in app-based contract tracing was significantly higher among college-educated respondents (predicted probability 0.63, 95% CI 0.53-0.73) compared to noncollege-educated respondents (predicted probability 0.39, 95% CI 0.28-0.50). However, among Black respondents, college-educated respondents were less likely to participate in app-based contact tracing (predicted probability 0.36, 95% CI 0.30-0.41) compared to noncollege-educated respondents (predicted probability 0.46, 95% CI 0.42-0.50). There was no significant difference between college- and noncollege-educated Latino respondents in the probability of participating in app-based contact tracing.

Figure 1. Predicted probability of the willingness to participate in app-based contact tracing by race and trust in the government.

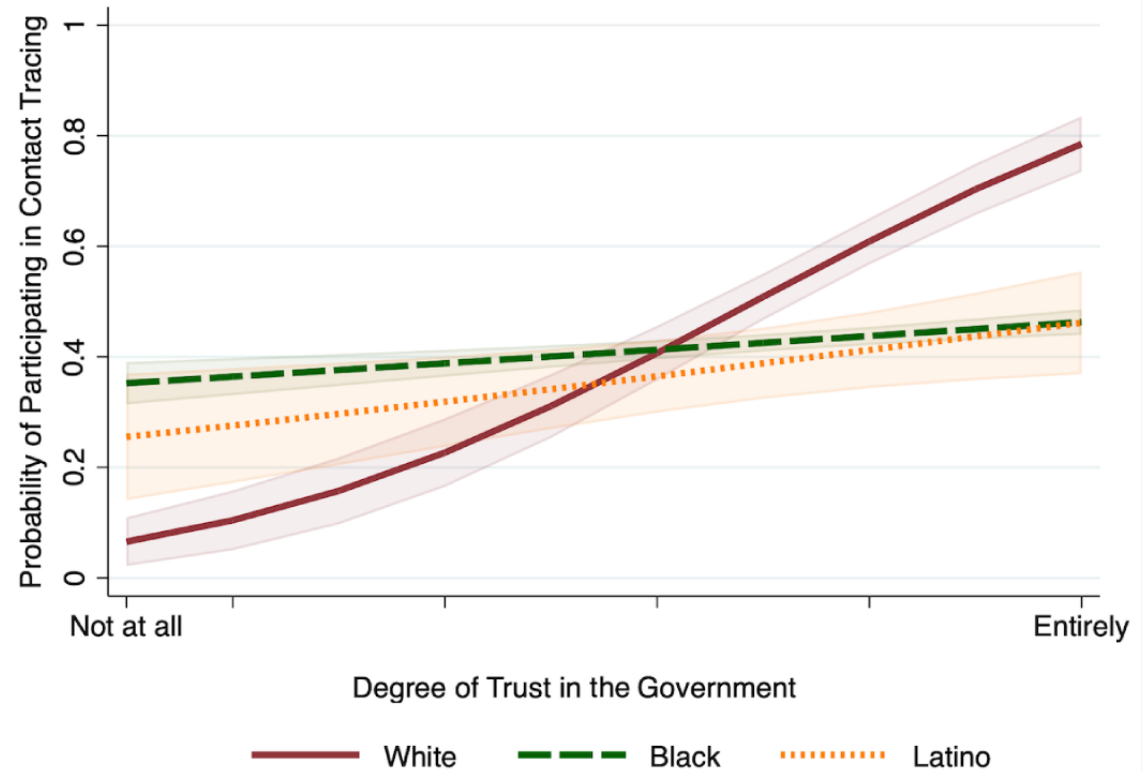
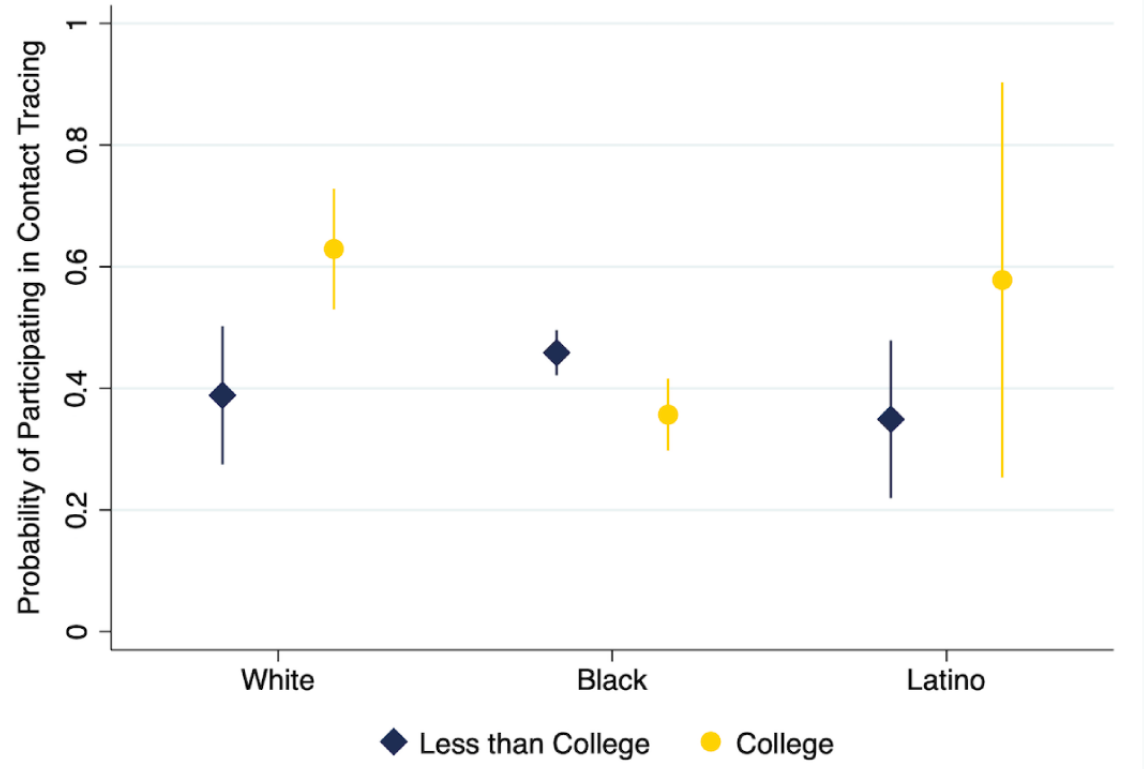


Figure 2. Predicted probability of the willingness to participate in app-based contact tracing by race and education level.



Discussion

Principal Findings

In this study, we found several factors that were associated with respondents' willingness to participate in contact tracing using a digital app. Black and Latino respondents were significantly less likely than White respondents to say they are willing to use an app for contact tracing, while those with higher trust in the government as well as those who had a close family member or friend die of COVID-19 were significantly more willing to use a contact tracing app than those with low trust or no close experiences. We also found that the relationship between trust in the government and willingness to participate in contact tracing was much stronger for White respondents than for Black and Latino respondents among whom trust in the government does not matter as much for predicting the willingness to participate in contact tracing. Given the widespread evidence of mistrust in the medical system among racial/ethnic minorities [35-37], this lack of association between trust in the government and willingness to adopt app-based contact tracing among Black and Latino respondents may reflect a general apprehension around digital health care that cannot be moderated by the level of trust. At the same time, the effect of trust in the state government on White residents' willingness or unwillingness to participate in contact tracing may reflect the partisan divide apparent nationally in response to COVID-19 [19].

In addition to our findings on trust, we also found that willingness to engage in app-based contact tracing varies substantially among racial/ethnic subpopulations depending on the level of education. While the likelihood of contact tracing among noncollege-educated respondents was relatively uniform across racial/ethnic groups, college-educated White respondents were substantially more likely to participate in contact tracing than noncollege-educated White respondents, and college- and noncollege-educated Black respondents. Moreover, for Black respondents, a college education significantly decreased the willingness to participate in contact tracing. In light of existing research indicating that education may influence one's use of digital health platforms [26,32] and online products or services because of privacy concerns [40], the interaction between race and education regarding the willingness to participate in app-based contact tracing suggests an understudied phenomenon whereby education may amplify the fears of minority populations when it comes to public health and may mitigate the fears of White populations.

Comparison to Prior Work

Our findings are consistent with prior research on public health and contact tracing apps, which has shown that factors, including personal experience with COVID-19, trust in the government, education level, and race and ethnicity, affect the willingness to use digital health tools [18,20,21]. Unlike previous work, however, we showed that some of these factors intersect in important ways. For example, while greater trust in the government was associated with increased willingness to use these digital tools among White respondents in our study, it had no significant effect among Black and Latino respondents.

Lower trust among racial and ethnic minorities is often linked to experiences of discrimination that reveal power imbalances and systemic inequalities in the practices of social institutions [52,53]. For example, systemic racism in health care is associated with disparities in access, quality of care, and health outcomes, as well as mistrust [54-59]. It is often racial minority individuals with higher levels of socioeconomic status who report more instances of discrimination in health care and in other institutions [52,55,60]. Upwardly mobile young adult Black and Latino individuals, for example, not only experience higher levels of discrimination than White individuals, but also are significantly more likely to experience acute and chronic discrimination compared to their socioeconomically stable counterparts [61]. Such research is consistent with our findings of an interaction between race/ethnicity and education, with college-educated Black respondents being significantly less willing to use a contact tracing app than either similarly educated White respondents or Black respondents with less than a college education. Awareness of and concerns about the extent of racial bias in digital technologies generally [62-66] may further contribute to reluctance to use these particular technologies among highly educated minoritized individuals.

Our findings are important because they suggest that factors that matter for some individuals and groups do not operate in the same way for others. Although trust in the government matters for the use of public health technologies for some individuals, it is likely to be interconnected with other social factors and conditions that must be considered. For most White individuals, trust in the government matters greatly, so understanding what influences that trust may be important for influencing the use of public health tools. For minoritized individuals, unequal experiences across social institutions may influence not only levels of trust but also more general opinions about and use of digital technologies. Future work must be attentive to not only how opinions, behaviors, and outcomes differ across minoritized groups, but also how actions and outcomes are associated with the lived experiences of individuals, stemming from institutionalized racism and other systems of inequality.

Limitations

A number of limitations should be considered when interpreting our findings. First, because the data were drawn from a single city, the findings may not be generalizable to digital contact tracing efforts in other communities or in the United States more broadly. Additional research is needed to understand the extent to which our findings on the willingness to participate in contact tracing extend to other communities with different demographic and socioeconomic compositions, including larger populations of other racial and gender minorities. As Detroit is a majority Black city in a majority White state, the effect of trust in the state government on individuals' decisions to participate in contact tracing may be more pronounced here than in other communities. At the same time, Detroit's experience, particularly early in the pandemic, as a COVID-19 hotspot may mean that residents are more willing to engage in contact tracing than other populations. Future research should investigate whether the patterns captured in our analysis hold in other communities and in the national population. Second, because our measure of contact tracing focused specifically on the willingness to download an app, it is possible that respondents' hesitancy is shaped more by discomfort with or privacy concerns regarding phone apps or access to smartphones and not contact tracing specifically. Though we controlled for internet access as a proxy for technology access and use, and only examined responses among cell phone owners, future research would benefit from examining the willingness to participate in digital contact tracing across a variety of modalities. Finally, our data were collected within the first 4 months of the coronavirus pandemic being declared a national emergency in the United States. In the ensuing months and years, COVID-19 has continued to spread and people's personal experiences of being infected with or knowing others infected with the coronavirus have changed. While digital contact tracing efforts in the United States have generally waned over time, it is possible that perspectives on contact tracing have shifted since the data were collected in response to the evolving nature of the pandemic and perspectives on public health generally.

Conclusions

The impact of COVID-19 has been disproportionately experienced by people of color [25]. In Detroit, Black residents experienced extremely high levels of disease, death, and loss of loved ones as a result of the pandemic [44]. Contact tracing is an important public health tool to limit the spread of infectious diseases like COVID-19, and digital contact tracing apps can positively contribute to such efforts [3]. Yet, throughout the COVID-19 pandemic, people's willingness to engage with digital contact tracing has varied [19].

Contact tracing technologies have the potential to advance public health efforts during the COVID-19 pandemic and beyond, but unequal uptake of these platforms across racial/ethnic groups and socioeconomic status may exacerbate the disparate impacts of health crises. Our findings show that residents' wariness of participating in contact tracing is associated with trust in the government and education levels, suggesting that distrust and concerns about systemic inequality and surveillance may drive reluctance to use such tools, especially among people of color [24]. These results point to the importance of being attuned to the ways intersectional identities influence public health behaviors and the need to develop culturally informed advocacy when promoting these platforms in the future [23,67].

Acknowledgments

We would like to thank the Detroit residents who took part in the survey. We would also like to acknowledge the staff of the Detroit Metro Area Communities Study and the funding for data collection from The Knight Foundation, The Ballmer Group, and Poverty Solutions at the University of Michigan. Additional support was provided though a Eunice Kennedy Shriver National Institute of Child Health and Human Development training grant to the Population Studies Center at the University of Michigan (T32 HD007339).

Conflicts of Interest

None declared.

References

1. Kalyanaraman N, Fraser MR. Containing COVID-19 Through Contact Tracing : A Local Health Agency Approach. Public Health Rep 2021 Nov 10;136(1):32-38 [FREE Full text] [doi: 10.1177/0033354920967910] [Medline: 33170094]
2. Almagor J, Picascia S. Exploring the effectiveness of a COVID-19 contact tracing app using an agent-based model. Sci Rep 2020 Dec 17;10(1):22235 [FREE Full text] [doi: 10.1038/s41598-020-79000-y] [Medline: 33335125]
3. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 Mobile Apps: A Systematic Review of the Literature. J Med Internet Res 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: 10.2196/23170] [Medline: 33197234]
4. Digital Contact Tracing Tools. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/digital-contact-tracing-tools.html [accessed 2023-01-06]
5. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of Mobile Apps to Promote Health and Manage Disease: Systematic Review and Meta-analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth 2021 Jan 11;9(1):e21563 [FREE Full text] [doi: 10.2196/21563] [Medline: 33427672]
6. O'Neill PH, Ryan-Mosley T, Johnson B. A flood of coronavirus apps are tracking us. Now it's time to keep track of them. MIT Technology Review. 2020. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/ [accessed 2022-12-28]
7. Hui CY, Walton R, McKinstry B, Jackson T, Parker R, Pinnock H. The use of mobile applications to support self-management for people with asthma: a systematic review of controlled studies to identify features associated with clinical effectiveness and adherence. J Am Med Inform Assoc 2017 May 01;24(3):619-632 [FREE Full text] [doi: 10.1093/jamia/ocw143] [Medline: 27694279]
8. Gandapur Y, Kianoush S, Kelli HM, Misra S, Urrea B, Blaha MJ, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes 2016 Oct 01;2(4):237-244 [FREE Full text] [doi: 10.1093/ehjqcco/qcw018] [Medline: 29474713]
9. Heisey-Grove DM, Church DR, Haney GA, Demaria A. Enhancing surveillance for hepatitis C through public health informatics. Public Health Rep 2011 Jan 01;126(1):13-18 [FREE Full text] [doi: 10.1177/003335491112600105] [Medline: 21337927]
10. Dounavi K, Tsoumani O. Mobile Health Applications in Weight Management: A Systematic Literature Review. Am J Prev Med 2019 Jun;56(6):894-903 [FREE Full text] [doi: 10.1016/j.amepre.2018.12.005] [Medline: 31003801]
11. Brewer LC, Fortuna KL, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the Future: Achieving Health Equity Through Health Informatics and Digital Health. JMIR Mhealth Uhealth 2020 Jan 14;8(1):e14512 [FREE Full text] [doi: 10.2196/14512] [Medline: 31934874]
12. Ray R, Sewell AA, Gilbert KL, Roberts JD. Missed Opportunity? Leveraging Mobile Technology to Reduce Racial Health Disparities. J Health Polit Policy Law 2017 Oct;42(5):901-924. [doi: 10.1215/03616878-3940477] [Medline: 28663182]
13. Huang Y, Sun M, Sui Y. How Digital Contact Tracing Slowed Covid-19 in East Asia. Harvard Business Review. 2020. URL: https://hbr.org/2020/04/how-digital-contact-tracing-slowed-covid-19-in-east-asia [accessed 2022-12-28]
14. Zeng K, Bernardo SN, Havins WE. The Use of Digital Tools to Mitigate the COVID-19 Pandemic: Comparative Retrospective Study of Six Countries. JMIR Public Health Surveill 2020 Dec 23;6(4):e24598 [FREE Full text] [doi: 10.2196/24598] [Medline: 33302255]
15. De La Garza A. Contact Tracing Apps Were Big Tech's Best Idea for Fighting COVID-19. Why Haven't They Helped? Time. URL: https://time.com/5905772/covid-19-contact-tracing-apps/ [accessed 2022-12-28]
16. Rich J. How our outdated privacy laws doomed contact-tracing apps. Brookings. URL: https://www.brookings.edu/blog/techtank/2021/01/28/how-our-outdated-privacy-laws-doomed-contact-tracing-apps/ [accessed 2022-12-28]
17. Maytin L, Maytin J, Agarwal P, Krenitsky A, Krenitsky J, Epstein RS. Attitudes and Perceptions Toward COVID-19 Digital Surveillance: Survey of Young Adults in the United States. JMIR Form Res 2021 Jan 08;5(1):e23000 [FREE Full text] [doi: 10.2196/23000] [Medline: 33347420]
18. Camacho-Rivera M, Islam JY, Rivera A, Vidot DC. Attitudes Toward Using COVID-19 mHealth Tools Among Adults With Chronic Health Conditions: Secondary Data Analysis of the COVID-19 Impact Survey. JMIR Mhealth Uhealth 2020 Dec 17;8(12):e24693 [FREE Full text] [doi: 10.2196/24693] [Medline: 33301415]
19. Mcclain C, Rainie L. The Challenges of Contact Tracing as U.S. Battles COVID-19. Pew Research Center. URL: https://www.pewresearch.org/internet/2020/10/30/the-challenges-of-contact-tracing-as-u-s-battles-covid-19/ [accessed 2022-12-28]
20. Pegollo L, Maggioni E, Gaeta M, Odone A. Characteristics and determinants of population acceptance of COVID-19 digital contact tracing: a systematic review. Acta Biomed 2021 Dec 10;92(S6):e2021444 [FREE Full text] [doi: 10.23750/abm.v92iS6.12234] [Medline: 34889313]

21. Megnin-Viggars O, Carter P, Melendez-Torres GJ, Weston D, Rubin GJ. Facilitators and barriers to engagement with contact tracing during infectious disease outbreaks: A rapid review of the evidence. PLoS One 2020;15(10):e0241473 [FREE Full text] [doi: 10.1371/journal.pone.0241473] [Medline: 33120402]
22. Touzani R, Schultz E, Holmes SM, Vandentorren S, Arwidson P, Guillemin F, et al. Early Acceptability of a Mobile App for Contact Tracing During the COVID-19 Pandemic in France: National Web-Based Survey. JMIR Mhealth Uhealth 2021 Jul 19;9(7):e27768 [FREE Full text] [doi: 10.2196/27768] [Medline: 34086589]
23. Shoji M, Ito A, Cato S, Iida T, Ishida K, Katsumata H, et al. Prosociality and the Uptake of COVID-19 Contact Tracing Apps: Survey Analysis of Intergenerational Differences in Japan. JMIR Mhealth Uhealth 2021 Aug 19;9(8):e29923 [FREE Full text] [doi: 10.2196/29923] [Medline: 34313601]
24. Dowthwaite L, Fischer J, Perez Vallejos E, Portillo V, Nichele E, Goulden M, et al. Public Adoption of and Trust in the NHS COVID-19 Contact Tracing App in the United Kingdom: Quantitative Online Survey Study. J Med Internet Res 2021 Sep 17;23(9):e29085 [FREE Full text] [doi: 10.2196/29085] [Medline: 34406960]
25. Risk for COVID-19 infection, hospitalization, and death by race/ethnicity. Centers for Disease Control and Prevention. URL: https://stacks.cdc.gov/view/cdc/107296 [accessed 2022-12-28]
26. Anthony DL, Campos-Castillo C, Lim PS. Who Isn't Using Patient Portals And Why? Evidence And Implications From A National Sample Of US Adults. Health Aff (Millwood) 2018 Dec;37(12):1948-1954. [doi: 10.1377/hlthaff.2018.05117] [Medline: 30633673]
27. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018 Aug 01;25(8):1080-1088 [FREE Full text] [doi: 10.1093/jamia/ocy052] [Medline: 29788380]
28. Pierce RP, Stevermer JJ. Disparities in the use of telehealth at the onset of the COVID-19 public health emergency. J Telemed Telecare 2020 Oct 21;29(1):3-9. [doi: 10.1177/1357633x20963893]
29. Nouri S, Khoong EC, Lyles CR, Karliner L. Addressing Equity in Telemedicine for Chronic Disease Management During the Covid-19 Pandemic. NEJM Catalyst Innovations in Care Delivery 2020:1-13 [FREE Full text] [doi: 10.1056/CAT.20.0123]
30. Campos-Castillo C, Anthony D. Racial and ethnic differences in self-reported telehealth use during the COVID-19 pandemic: a secondary analysis of a US survey of internet users from late March. J Am Med Inform Assoc 2021 Jan 15;28(1):119-125 [FREE Full text] [doi: 10.1093/jamia/ocaa221] [Medline: 32894772]
31. Reed ME, Huang J, Graetz I, Lee C, Muelly E, Kennedy C, et al. Patient Characteristics Associated With Choosing a Telemedicine Visit vs Office Visit With the Same Primary Care Clinicians. JAMA Netw Open 2020 Jun 01;3(6):e205873 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.5873] [Medline: 32585018]
32. Liaw WR, Jetty A, Coffman M, Petterson S, Moore MA, Sridhar G, et al. Disconnected: a survey of users and nonusers of telehealth and their use of primary care. J Am Med Inform Assoc 2019 May 01;26(5):420-428 [FREE Full text] [doi: 10.1093/jamia/ocy182] [Medline: 30865777]
33. Whetten K, Leserman J, Whetten R, Ostermann J, Thielman N, Swartz M, et al. Exploring Lack of Trust in Care Providers and the Government as a Barrier to Health Service Use. Am J Public Health 2006 Apr;96(4):716-721. [doi: 10.2105/ajph.2005.063255]
34. Mechanic D. The functions and limitations of trust in the provision of medical care. J Health Polit Policy Law 1998 Aug 01;23(4):661-686. [doi: 10.1215/03616878-23-4-661] [Medline: 9718518]
35. LaVeist TA, Nickerson KJ, Bowie JV. Attitudes about racism, medical mistrust, and satisfaction with care among African American and white cardiac patients. Med Care Res Rev 2000;57 Suppl 1:146-161. [doi: 10.1177/1077558700057001S07] [Medline: 11092161]
36. Institute of Medicine (US) Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. In: Smedley BD, Stith AY, Nelson AR, editors. Unequal Treatment Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: National Academies Press (US); 2003.
37. Sewell AA. Disaggregating ethnoracial disparities in physician trust. Soc Sci Res 2015 Nov;54:1-20. [doi: 10.1016/j.ssresearch.2015.06.020] [Medline: 26463531]
38. Hall JL, McGraw D. For telehealth to succeed, privacy and security risks must be identified and addressed. Health Aff (Millwood) 2014 Feb;33(2):216-221. [doi: 10.1377/hlthaff.2013.0997] [Medline: 24493763]
39. Campos-Castillo C, Anthony DL. The double-edged sword of electronic health records: implications for patient disclosure. J Am Med Inform Assoc 2015 Apr;22(e1):e130-e140. [doi: 10.1136/amiajnl-2014-002804] [Medline: 25059953]
40. Perrin A. Half of Americans have decided not to use a product or service because of privacy concerns. Pew Research Center. URL: https://www.pewresearch.org/fact-tank/2020/04/14/half-of-americans-have-decided-not-to-use-a-product-or-service-because-of-privacy-concerns/ [accessed 2022-12-28]
41. Jones K, Thompson R. To Use or Not to Use a COVID-19 Contact Tracing App: Mixed Methods Survey in Wales. JMIR Mhealth Uhealth 2021 Nov 22;9(11):e29181 [FREE Full text] [doi: 10.2196/29181] [Medline: 34698645]
42. Perry BL. Contact Tracing Could Exacerbate COVID-19 Health Disparities: The Role of Economic Precarity and Stigma. Am J Public Health 2021 May;111(5):778-781. [doi: 10.2105/ajph.2021.306244]
43. Schneider J, Love W, Rusie L, Flores A, Tadesse B, Hazra A, et al. COVID-19 Contact Tracing Conundrums: Insights From the Front Lines. Am J Public Health 2021 May;111(5):917-922. [doi: 10.2105/ajph.2021.306200]

44. Ray R, Morgan JF, Wileden L, Elizondo S, Wiley-Yancy D. Examining and addressing COVID-19 racial disparities in Detroit. Brookings Institution. 2021. URL: https://www.brookings.edu/research/examining-and-addressing-covid-19-racial-disparities-in-detroit/ [accessed 2022-12-28]
45. Marans RW, Gerber ER, Morenoff JD. Detroit Area Studies (DAS). In: Maggino F, editor. Encyclopedia of Quality of Life and Well-Being Research. Cham: Springer; 2021:1-6.
46. Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys. American Association for Public Opinion Research. 2016. URL: https://www-archive.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf [accessed 2022-12-28]
47. Wagner A, Wileden L, Shanks T, Goold S, Morenoff J, Sheinfeld Gorin SN. Mediators of Racial Differences in COVID-19 Vaccine Acceptance and Uptake: A Cohort Study in Detroit, MI. Vaccines (Basel) 2021 Dec 28;10(1):36 [FREE Full text] [doi: 10.3390/vaccines10010036] [Medline: 35062697]
48. Wagner AL, Sheinfeld Gorin S, Boulton ML, Glover BA, Morenoff JD. Effect of vaccine effectiveness and safety on COVID-19 vaccine acceptance in Detroit, Michigan, July 2020. Hum Vaccin Immunother 2021 Sep 02;17(9):2940-2945 [FREE Full text] [doi: 10.1080/21645515.2021.1917233] [Medline: 33998949]
49. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Hoboken, NJ: Wiley & Sons, Inc; 1987.
50. Stata Multiple Imputation Reference Manual. StataCorp. URL: https://www.stata.com/manuals/mi.pdf [accessed 2022-12-28]
51. Boyle J, Brassell T, Dayton J. As cases increase, American trust in COVID-19 information from federal, state, and local governments continues to decline. ICF International Inc. URL: https://www.icf.com/insights/health/covid-19-survey-american-trust-government-june [accessed 2022-12-28]
52. Nunnally SC. Trust in Black America: Race, Discrimination, and Politics. New York, NY: NYU Press; 2012.
53. Smith SS. Race and Trust. Annu. Rev. Sociol 2010 Jun 01;36(1):453-475. [doi: 10.1146/annurev.soc.012809.102526]
54. Armstrong K, Ravenell KL, McMurphy S, Putt M. Racial/ethnic differences in physician distrust in the United States. Am J Public Health 2007 Jul;97(7):1283-1289. [doi: 10.2105/AJPH.2005.080762] [Medline: 17538069]
55. Bleich SN, Findling MG, Casey LS, Blendon RJ, Benson JM, SteelFisher GK, et al. Discrimination in the United States: Experiences of black Americans. Health Serv Res 2019 Dec 29;54 Suppl 2(Suppl 2):1399-1408 [FREE Full text] [doi: 10.1111/1475-6773.13220] [Medline: 31663124]
56. Williams DR, Lawrence JA, Davis BA. Racism and Health: Evidence and Needed Research. Annu Rev Public Health 2019 Apr 01;40:105-125 [FREE Full text] [doi: 10.1146/annurev-publhealth-040218-043750] [Medline: 30601726]
57. 2019 National Healthcare Quality and Disparities Report. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/research/findings/nhqrdr/nhqdr19/index.html [accessed 2022-12-28]
58. Feagin J, Bennefield Z. Systemic racism and U.S. health care. Soc Sci Med 2014 Feb;103:7-14. [doi: 10.1016/j.socscimed.2013.09.006] [Medline: 24507906]
59. Nong P, Williamson A, Anthony D, Platt J, Kardia S. Discrimination, trust, and withholding information from providers: Implications for missing data and inequity. SSM Popul Health 2022 Jun;18:101092 [FREE Full text] [doi: 10.1016/j.ssmph.2022.101092] [Medline: 35479582]
60. Abramson CM, Hashemi M, Sánchez-Jankowski M. Perceived discrimination in U.S. healthcare: Charting the effects of key social characteristics within and across racial groups. Prev Med Rep 2015;2:615-621 [FREE Full text] [doi: 10.1016/j.pmedr.2015.07.006] [Medline: 26844127]
61. Colen CG, Ramey DM, Cooksey EC, Williams DR. Racial disparities in health among nonpoor African Americans and Hispanics: The role of acute and chronic discrimination. Soc Sci Med 2018 Feb;199:167-180 [FREE Full text] [doi: 10.1016/j.socscimed.2017.04.051] [Medline: 28571900]
62. Benjamin R. Race After Technology: Abolitionist Tools for the New Jim Code. Hoboken, NJ: Wiley & Sons, Inc; 2019.
63. Tynes BM, Giang MT, Williams DR, Thompson GN. Online racial discrimination and psychological adjustment among adolescents. J Adolesc Health 2008 Dec;43(6):565-569. [doi: 10.1016/j.jadohealth.2008.08.021] [Medline: 19027644]
64. Buolamwini J, Gebru T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: Proceedings of the 1st Conference on Fairness, Accountability and Transparency. 2018 Presented at: 1st Conference on Fairness, Accountability and Transparency; February 23-24, 2018; New York, NY p. 77-91.
65. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453 [FREE Full text] [doi: 10.1126/science.aax2342] [Medline: 31649194]
66. Laestadius LI, Craig KA, Campos-Castillo C. Perceptions of Alerts Issued by Social Media Platforms in Response to Self-injury Posts Among Latinx Adolescents: Qualitative Analysis. J Med Internet Res 2021 Aug 10;23(8):e28931 [FREE Full text] [doi: 10.2196/28931] [Medline: 34383683]
67. Witteveen D, de Pedraza P. The Roles of General Health and COVID-19 Proximity in Contact Tracing App Usage: Cross-sectional Survey Study. JMIR Public Health Surveill 2021 Aug 18;7(8):e27892 [FREE Full text] [doi: 10.2196/27892] [Medline: 34081602]

Abbreviations

DMACS: Detroit Metro Area Communities Study

OR: odds ratio

Edited by A Mavragani; submitted 25.04.22; peer-reviewed by A Triantafyllidis, C Biernesser; comments to author 17.05.22; revised version received 06.06.22; accepted 13.10.22; published 19.01.23.

Please cite as:
Wileden L, Anthony D, Campos-Castillo C, Morenoff J
Resident Willingness to Participate in Digital Contact Tracing in a COVID-19 Hotspot: Findings From a Detroit Panel Study
JMIR Public Health Surveill 2023;9:e39002
URL: https://publichealth.jmir.org/2023/1/e39002
doi:10.2196/39002
PMID:36240029

Original Paper

Adaptation and Utilization of a Postmarket Evaluation Model for Digital Contact Tracing Mobile Health Tools in the United States: Observational Cross-sectional Study

Kevin E Cevasco[1], MPH, MBA; Amira A Roess[1], MPH, PhD

[1] College of Health and Human Services, George Mason University, Fairfax, VA, United States

Corresponding Author:
Amira A Roess, MPH, PhD
College of Health and Human Services
George Mason University
4400 University Dr
Fairfax, VA, 22030
United States
Phone: 1 703 993 1923
Email: aroess@gmu.edu

Abstract

Background: Case investigation and contact tracing are core public health activities used to interrupt disease transmission. These activities are traditionally conducted manually. During periods of high COVID-19 incidence, US health departments were unable to scale up case management staff to deliver effective and timely contact-tracing services. In response, digital contact tracing (DCT) apps for mobile phones were introduced to automate these activities. DCT apps detect when other DCT users are close enough to transmit COVID-19 and enable alerts to notify users of potential disease exposure. These apps were deployed quickly during the pandemic without an opportunity to conduct experiments to determine effectiveness. However, it is unclear whether these apps can effectively supplement understaffed manual contact tracers.

Objective: The aims of this study were to (1) evaluate the effectiveness of COVID-19 DCT apps deployed in the United States during the COVID-19 pandemic and (2) determine if there is sufficient DCT adoption and interest in adoption to meet a minimum population use rate to be effective (56%). To assess uptake, interest and safe use covariates were derived from evaluating DCTs using the American Psychological Association App Evaluation Model (AEM) framework.

Methods: We analyzed data from a nationally representative survey of US adults about their COVID-19–related behaviors and experiences. Survey respondents were divided into three segments: those who adopted a DCT app, those who are interested but did not adopt, and those not interested. Descriptive statistics were used to characterize factors of the three groups. Multivariable logistic regression models were used to analyze the characteristics of segments adopting and interested in DCT apps against AEM framework covariates.

Results: An insufficient percentage of the population adopted or was interested in DCTs to achieve our minimum national target effectiveness rate (56%). A total of 17.4% (n=490) of the study population reported adopting a DCT app, 24.7% (n=697) reported interest, and 58.0% (n=1637) were not interested. Younger, high-income, and uninsured individuals were more likely to adopt a DCT app. In contrast, people in fair to poor health were interested in DCT apps but did not adopt them. App adoption was positively associated with visiting friends and family outside the home (odds ratio [OR] 1.63, 95% CI 1.28-2.09), not wearing masks (OR 0.52, 95% CI 0.38-0.71), and adopters thinking they have or had COVID-19 (OR 1.60, 95% CI 1.21-2.12).

Conclusions: Overall, a small percentage of the population adopted DCT apps. These apps may not be effective in protecting adopters' friends and family from their maskless contacts outside the home given low adoption rates. The public health community should account for safe use behavioral factors in future public health contact-tracing app design. The AEM framework was useful in developing a study design to evaluate DCT effectiveness and safety.

(JMIR Public Health Surveill 2023;9:e38633) doi:10.2196/38633

KEYWORDS

COVID-19; contact tracing; postmarketing; mobile apps; public health; digital; interventions; tool; adoption; effectiveness; prevention; application; transmission

Introduction

Background

During periods of high COVID-19 incidence, US health departments were unable to scale up case management staff to deliver effective and timely contact-tracing services [1]. This is expected considering that prior to the COVID-19 outbreak, US health departments were grossly understaffed and needed a minimum of 80,000 more full-time equivalent positions to provide adequate infrastructure and a minimum package of public health services [2].

Digital contact tracing (DCT) apps are intended to supplement health department case managers by automatically and more efficiently estimating the proximity and duration of an individual's exposure to patients diagnosed with COVID-19. A model developed to predict the impact of conventional and mobile DCT apps on the SARS-CoV-2 pandemic estimated that the disease transmission rate could be reduced by adding either contact tracers or effective widely used DCT apps (eg, Bluetooth, GPS) [3]. DCTs reduce disease spread by alerting the user if they have been near an infected person who is also using DCT, followed by user isolation and testing for infection. However, a key DCT effectiveness assumption is that there are sufficient people adopting the app to signal disease exposure. Ongoing monitoring of population uptake is necessary to address concerns that users are relying on DCT apps that have not met minimum population-level app adoption rates [4,5].

Health Technology Assessments

Health technology assessments (HTAs) of public health applications are uncommon and challenged by lack of data, conflicting stakeholder priorities, and methodological issues [6]. The surge in COVID-19 cases in the United States in the summer of 2021 (ie, the "third wave") driven by the SARS-CoV-2 Delta variant marks an urgent time to perform an assessment of DCT apps to inform future public health contact-tracing programs and app design [7]. By 2020, an estimated 2.6 billion people globally were expected to use mobile health (mHealth) apps, generating an estimated US $31 billion in revenue [8].

Despite the large market of mHealth apps, there is no US regulatory authority in place to evaluate if DCT apps are safe and beneficial [9]. Section 522 of the Federal Food, Drug, and Cosmetic Act provides the Food and Drug Administration (FDA) with the authority to require manufacturers to conduct postmarket surveillance to address important public health questions on the effectiveness and safety of a health device [10]. However, the FDA generally does not regulate products that are intended solely to track locations or contacts associated with public health surveillance [11]. Public health DCT developers could voluntarily implement HTAs using the FDA Software Precertification Program; however, the FDA guidance does not provide a framework to conduct the postmarket evaluations [12].

There are a wide variety of frameworks available for evaluating health applications, but they may not be suitable for HTAs [13,14]. The clinical applications marketplace conducts HTAs, but they have not been frequently performed in the public health market [6]. This study adapted the American Psychological Association (APA) App Evaluation Model (AEM) hierarchical rating system because it was developed by harmonizing application evaluation questions from 45 frameworks [15]. The AEM provides a comprehensive evaluation framework for clinicians and patients to identify high-quality mHealth apps. The framework may work well at a population level with hierarchical layers to explain adoption decisions [16,17].

DCT only functions when all people near each other are using the app. Individuals install DCT apps on their own mobile devices, and these apps must be adopted by a sufficient percentage of a population to be effective [16,17]. In this study, we set the adoption effectiveness rate at 56% of the national population [18]. We define our DCT adoption as the percentage of the population using a DCT plus those that are interested in using a DCT but have not yet adopted it. Our assumption is that public health campaigns could convert many of the interested users into adopters through public health communication campaigns.

Following public health safety guidance while using DCT is important, and the AEM framework provides health and technology evaluation elements to evaluate functional use. For example, DCT requires users to voluntarily report their positive COVID-19 test results to public health authorities; therefore, user behavior is part of DCT functionality. After DCT adoption, users should not cease their social distancing behaviors in a way that increases disease exposures.

DCT represents a new technology with a short research history, and the AEM provides a structure for measure selection [19]. The AEM has five categories for evaluation: Access and Background, Privacy and Safety, Clinical Foundation, Usability, and Data Integration Toward a Therapeutic Goal. The AEM provides a way to critically assess mHealth and does not include a minimum or maximum score to be considered "good" or "useful" [15,20].

Objective

The main objective of this study was to evaluate COVID-19 DCT apps deployed in the United States during the COVID-19 pandemic under the hypothesis that DCT requires a minimal level of uptake by a population to be effective and DCT users should follow social distancing policies. Accordingly, the specific study aims were as follows.

Aim 1: Determine if there is a sufficient percentage of people using or interested in DCT to reach the minimum critical mass (56%) [18].

Aim 2: Assess the effectiveness and safety of DCT apps using the APA AEM hierarchical rating system [16].

The AEM is a conceptual framework for developing key study measures and was used to select variables. The measures derived from the framework assess adoption rate characteristics and behaviors necessary for the safe use of DCT apps.

Methods

Data Source

We approximated a nationally representative sample of US adults (aged 18+ years), with a stratified nonprobability sample to survey 3853 online panel participants between December 22, 2020, and January 2, 2021. Respondents were recruited by Climate Nexus Polling, using several market research panels as described elsewhere [21]. Participants were recruited using stratified sampling methods [21]. Compensation for participants depended on the specific market research panel and respondents' preferences (eg, cash, gift cards, reward points) and was valued below US $4.00. Quotas were set to match the US Census Bureau's Voting and Registration Supplement to the Current Population Survey parameters for age, gender, race, educational attainment, census region, and Hispanic ethnicity. Sampling weights were used to account for small deviations from the preselected census parameters. The participation rate of the survey was 68.5% (qualified respondents that completed the survey). The 95% credibility interval for this survey was ±1.7%. The survey data are provided in Multimedia Appendix 1.

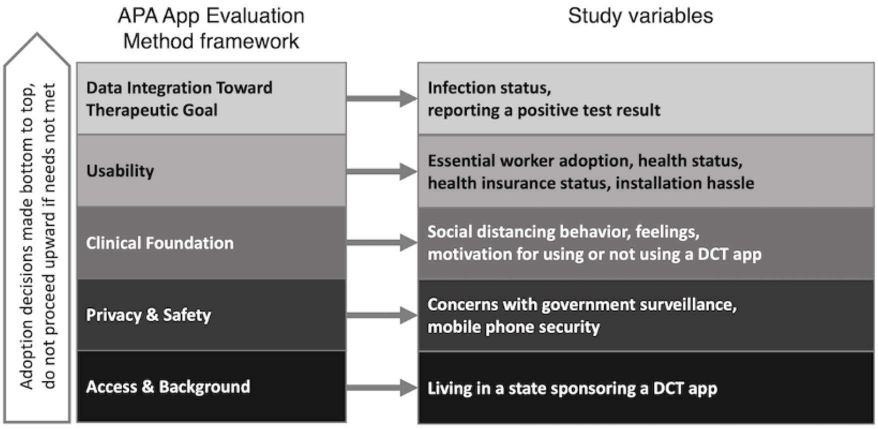
DCT apps are new technology and there were no established and validated survey instruments to measure their use; therefore, Climate Nexus Polling developed DCT app survey questions based on a previous Pew Research survey on digital privacy [22].

Key Study Measures Aligned to the APA AEM

Model Variable Selection

Demographic measures were based on a literature review to assess if app adoption applied to a diverse range of people [23,24]. The following background for each category links the AEM to the study's DCT postmarket evaluation. The AEM framework was used to select variables a priori from data available in the survey (Figure 1). The detailed process is described in Multimedia Appendix 2.

Figure 1. App Evaluation Model (AEM) with study variables. APA: American Psychological Association; DCT: digital contact tracing.



Access and Background

This category is concerned with the transparency of app governance and ensuring that the benefits of mHealth apps are available to a diverse range of people, regardless of background. The US DCT apps are based on mobile phone Bluetooth proximity tracking using the Google/Apple Exposure Notification (GAEN) platform [25], and are cosponsored by US state government public health organizations [25]. Adoption of DCT apps in the United States is voluntary, and varies by demographics and behavior [23]. Furthermore, DCT app adoption may be subject to health equity issues since mHealth adoption is generally associated with younger, educated populations [23,26].

The "state of residence" measure was used to demonstrate the role of a person's state government in driving local adoption [25,27,28].

Privacy and Safety

There is an expectation of confidentiality by DCT users, and privacy policies should ensure that COVID-19 exposure proximity measurement is accurate. Concern over privacy and data breaches is a genuine concern. Experts collaborated to develop secure decentralized contract-tracing platform architectures [29]; however, some GAEN implementations may have exposed user data. A class action lawsuit was filed in April of 2021 against Google, claiming that their COVID-19 contact-tracing platform exposed unwitting Android adopters' sensitive personal information [30].

Concerns over government surveillance and mobile device hacking measures were selected from the survey [23,28,30].

Clinical Foundation

There is no agreed-upon definition of effectiveness for measuring the clinical foundation of DCT. In January 2021, the National Institute of Standards and Technology held a workshop to discuss successes and challenges associated with implementing proximity detection technologies [31]. For public health, effectiveness is defined as reducing transmission and quicker notification of exposure. From an adopter perspective, effectiveness lies in addressing privacy concerns and finding a value proposition. From a technology perspective, effectiveness is how well the DCT apps measure proximity and if they can function at current adoption rates. A systematic review of DCT also found this lack of definition for effectiveness [18].

Also of potential concern is the "risk compensation" theory, which suggests that individuals will increase risky social distancing behavior after a COVID-19 public health intervention is implemented [32,33]. Researchers are concerned that DCT app adoption could create a false sense of security and safety, leading to behavior that increases infections [34]. Additionally, the personality of the adopter may be an issue because the extroversion personality type is associated with lower compliance with social distancing [35].

Clinical foundation measures were selected to inform future work on the definition of DCT public health effectiveness. Measures were selected to assess social distancing behavior (mask, visiting friends/family, attending religious services) [23,35]. Psychological distress that could impact adoption and measures were included to assess a user's worry about COVID-19 [36]. Measures on motivations for adoption and interest were selected from the survey [23,24].

Usability

Grounded in the principle of autonomy, an mHealth app should enable the user to take an educated and active role in their health decisions. For example, DCT apps could help essential workers take an active role in managing their exposures. However, a German study showed that there was not greater uptake of DCT apps by people working during the early pandemic [23]. In addition, the readability of DCT apps is higher than US reading levels and potentially less accessible for the general population [27]. There is also little known about DCT use by people in good versus poor health. Studies show that older people with chronic diseases adopt self-care and vitals mHealth apps [26]; however, their chronic disease management is a very different purpose than tracking infectious disease exposure.

Usability is especially important for populations with high exposure risk or at high risk of poor outcomes. Measures were selected for essential worker, health, and health insurance status [23,24].

Data Integration Toward a Therapeutic Goal

This category is related to the principle of shared decision-making through appropriate information sharing with public health departments and health care systems. Some DCT apps do offer adopters links to testing resources, local health department resources, and quarantine social support [27]. In the US GAEN DCT systems, adopters are solely in control of reporting their positive test results through the app; therefore, designers have limited ability to automate the contact-tracing reporting data path. Unfortunately, a Morbidity and Mortality Weekly Report found that 35% to 48% of individuals with COVID-19 did not report any of their contacts to public health department contact tracers [37]. DCT apps provide anonymous COVID-19–positive lab test reporting capability, but there is no evidence showing that DCT adoption improves the likelihood of reporting.

User infection status and sharing a positive test result measures assess if apps ultimately provide exposure data to public health authorities [24,37].

Statistical Analysis

Market segments of the survey population were divided into three categories: (1) the "adopted" category includes the population self-declaring use of a DCT app; (2) the "interested-not-adopted" category includes the population self-declaring interest in, but not using a DCT app; and (3) the "not interested" category describes the remaining population. All three categories were included in the descriptive statistics table using frequencies and survey weighted percentages. Only the adopted and interested-not-adopted ("interested") categories were included in the statistical analysis because these populations would be targets for uptake communications campaigns. Logistic regression models were applied to the adopted and interested-not-adopted segments to compare against the key measures. Regression results for the two models are presented with unweighted percentages and adjusted odds ratios (ORs) with 95% CIs. Both model results and key measures are presented in tables with variables organized by AEM categories. Health care workers (n=1029) were excluded from the analysis because health systems had contact-tracing protocols in place before the COVID-19 pandemic to meet health care–acquired infections regulations. Health care workers may answer survey questions inconsistently because they follow workplace infection contact-tracing protocols that are not applicable to the general public. Multicollinearity of adopted and interested-not-adopted models were checked through correlation analysis. Descriptive and statistical analyses were conducted using STATA version 16.1 (StataCorp, College Station, TX).

To mitigate the risk of model overfitting given the large number of measures, variables were selected a priori through a literature review and allocated to the APA framework as described in the Key Measures section above and Multimedia Appendix 2. By using measures with statistical significance from other data sets, we reduced the likelihood that our results were based on overfitting in our survey sample. Results were also analyzed for overfit risk by setting a minimum measure subsample frequency threshold (n>15) [38].

Ethical Considerations

This project was considered exempt for review by the George Mason University Institutional Review Board (1684418-3).

Results

Descriptive Analysis

Out of 3853 respondents, 2824 were included in the analysis, with 1029 excluded because they self-identified as health care workers. Among the total sample of 2824 people, 490 (17.4%) reported using a DCT app, 697 (24.7%) reported interest, and 1637 (58.0%) were not interested (Table 1). All measures exceed the minimum frequency threshold for assessing the risk of overfitting, with most of the measures having over 100 survey respondents.

Older age groups adopted the apps less frequently, where the ≥65-year-old group are at higher risk of poor COVID-19 outcomes but represent a lower rate of adoption (10.2%). Those not interested in adopting the app were disproportionately concerned with government surveillance postepidemic (43.2%).

Table 1. Frequencies of survey population and survey segments.

Survey segment	Study population (N=2824), n (wt%[a])	Adopted (n=490), n (%)	Interested-not-adopted (n=697), n (%)	Not interested (n=1637), n (%)
Access and background: State of residence offers a DCT[b] app (yes)	1534 (55.3)	310 (63.3)	376 (54.0)	848 (51.8)
Security and privacy				
Do not use because of concerns with government surveillance postepidemic	N/A[c]	N/A	236 (33.9)	707 (43.2)
Do not use because my phone might get hacked	N/A	N/A	155 (22.2)	405 (24.7)
Clinical foundation				
Social distancing				
Visit friends and family weekly (yes)	1321 (47.6)	302 (61.6)	286 (43.0)	733 (45.1)
Worn a mask in public to help protect yourself or others from getting sick? (yes)	2409 (84.5)	367 (74.9)	611 (87.7)	1431 (87.4)
Attends religious services (yes)	1309 (46.5)	285 (58.2)	333 (47.8)	691 (42.2)
Feelings				
Worried about losing your home (yes)	784 (28.5)	212 (43.3)	215 (30.9)	357 (21.8)
Feel afraid (yes)	893 (31.9)	204 (41.6)	271 (38.9)	418 (25.5)
Feel lonely (yes)	812 (28.6)	202 (41.2)	230 (33.0)	380 (23.2)
Motivation for use				
Use to protect family and friends	N/A	279 (56.9)	430 (61.7)	N/A
Use to help stop the pandemic/reduce deaths in older American adults	N/A	229 (46.7)	326 (46.8)	N/A
Use as a responsibility to my community	N/A	219 (44.7)	320 (45.9)	N/A
Use to let me know my risk of infection/help peace of mind/help me stay healthy	N/A	268 (54.7)	432 (62.0)	N/A
Not use because it would make me feel more anxious	N/A	N/A	227 (32.6)	552 (33.7)
Not use because I would not benefit/I won't be infected	N/A	N/A	122 (17.5)	446 (27.2)
Usability				
Essential worker	664 (24.5)	157 (32.0)	173 (24.8)	334 (20.4)
Health status				
Very good or excellent	1288 (45.7)	266 (54.3)	317 (45.5)	705 (43.1)
Good	967 (34.2)	151 (30.8)	226 (32.4)	590 (36.0)
Fair or poor	569 (20.2)	73 (14.9)	154 (22.1)	342 (20.9)
Has health insurance (yes)	2421 (85.3)	389 (79.4)	616 (88.4)	1416 (86.5)
No use because too much hassle to install	N/A	N/A	187 (26.8)	389 (23.8)
Therapeutic goal				
Would you tell your neighbors or friends if you become COVID-19–positive? (yes)	1949 (69.1)	330 (67.4)	502 (72.0)	1117 (68.2)
Do you think you have already had the coronavirus or currently have it? (yes)	490 (18.2)	140 (28.6)	115 (16.5)	235 (14.4)
Demographics				
Age (Gallup divisions, years)				
18-29	510 (19.0)	155 (31.6)	144 (20.7)	211 (12.9)
30-49	948 (33.2)	214 (43.7)	283 (40.6)	451 (27.6)

Survey segment	Study population (N=2824), n (wt%[a])	Adopted (n=490), n (%)	Interested-not-adopted (n=697), n (%)	Not interested (n=1637), n (%)
50-64	686 (24.2)	71 (14.5)	154 (22.1)	461 (28.2)
65+	680 (23.6)	50 (10.2)	116 (16.6)	514 (31.4)
Gender				
Female	1491 (51.7)	245 (50.0)	393 (56.4)	853 (52.1)
Male	1333 (48.3)	245 (50.0)	304 (43.6)	784 (47.9)
Race				
White	2126 (72.1)	319 (65.1)	506 (72.6)	1301 (79.5)
Black	322 (10.2)	69 (14.1)	84 (12.1)	169 (10.3)
Hispanic	232 (14.1)	68 (13.9)	68 (9.8)	96 (5.9)
Other	144 (3.7)	34 (6.9)	39 (5.6)	71 (4.3)
Political party affiliation				
Republican	1066 (38.1)	154 (31.4)	205 (29.4)	707 (43.2)
Democratic	1273 (44.5)	269 (54.9)	383 (55.0)	621 (37.9)
Independent	485 (17.4)	67 (13.7)	109 (15.6)	309 (18.9)
Annual income level (US $)				
<50,000	1606 (57.4)	244 (49.8)	366 (52.5)	996 (60.8)
50,000-99,000	778 (27.4)	137 (28.0)	201 (28.8)	440 (26.9)
≥100,000	440 (15.2)	109 (22.2)	130 (18.7)	201 (12.3)
Population density				
Urban	814 (29.5)	185 (37.8)	236 (33.9)	393 (24.0)
Semiurban	1326 (46.8)	218 (44.5)	322 (46.2)	786 (48.0)
Rural	684 (23.8)	87 (17.8)	139 (19.9)	458 (28.0)
Education				
No high school	176 (6.7)	38 (7.8)	44 (6.3)	94 (5.7)
At least some high school	718 (26.4)	107 (21.8)	147 (21.1)	464 (28.3)
Some college	1025 (37.7)	158 (32.2)	246 (35.3)	621 (37.9)
Bachelor’s degree or above	905 (29.3)	187 (38.2)	260 (37.3)	458 (28.0)

[a] wt%: weighted percentage.

[b] DCT: digital contact tracing.

[c] N/A: not applicable.

DCT app adopters more frequently visited friends and family outside the home each week (61.6%) compared to the overall survey rate (47.6%). Those interested in the app visited friends and family outside the home less frequently (43.0%). Adopters also had high rates of concern about losing their home (43.3%), being lonely (41.2%), and feeling afraid (41.6%). Despite having a higher rate of thinking they have or had the coronavirus infection (28.6%), their rate of telling neighbors and friends was similar to that of the base population (67.4%).

There was also a difference in the rate of wearing masks in the overall study population (84.5%) versus a lower rate in the adoption segment (74.9%). Adopters showed a higher rate of very good/excellent health (54.3%) compared to the overall population and other categories.

Logistic Regression Analysis

There were differences in the frequency and association between the adopted and interested segment populations across all five AEM categories (Table 2). Covariate-adjusted ORs show how the adoption and interest vary by study subject characteristics. The OR is interpreted as the odds of a covariate being the same (OR=1.0) for adopted versus not adopted groups. For example, people who live in a state that offers a DCT adopt the apps 1.47 times more often (OR>1.0) than people in states not offering DCTs. People who wore a mask in public to help protect themselves or others from getting sick adopt DCT apps 0.52 times less often (OR<1.0) than people that do not wear masks.

Table 2. Logistic regression results for adopted and interested-not-adopted segments.

Survey segments	Adopted		Interested-not-adopted	
	Respondents (n=2824), %[a]	Adjusted OR[b] (95% CI) (n=490)	Respondents (n=2334), %[a]	Adjusted OR (95% CI) (n=697)
Access and background: state of residence offers a DCT[c] app (yes)	*63.3*[d]	*1.47 (1.16-1.86)*	53.9	0.98 (0.81-1.19)
Security and privacy				
Do not use because of concerns with government surveillance postepidemic	N/A[e]	N/A	*25.0*	*0.67 (0.54-0.83)*
Do not use because my phone might get hacked	N/A	N/A	*27.7*	*0.78 (0.62-0.99)*
Clinical foundation				
Social distancing				
Visit friends and family weekly (yes)	*22.9*	*1.63 (1.28-2.09)*	*28.1*	*0.81 (0.66-0.99)*
Worn a mask in public to help protect yourself or others from getting sick? (yes)	*15.2*	*0.52 (0.38-0.71)*	29.9	1.05 (0.77-1.44)
Attends religious services (Yes)	21.8	1.26 (0.99-1.61)	*32.5*	*1.29 (1.06-1.58)*
Feelings				
Worried about losing your home (yes)	*27.0*	*1.52 (1.18-1.95)*	37.6	1.19 (0.94-1.49)
Feel afraid (yes)	22.8	1.04 (0.80-1.35)	*39.3*	*1.38 (1.10-1.73)*
Feel lonely (yes)	*24.9*	*1.31 (1.00-1.70)*	37.7	1.14 (0.90-1.44)
Motivation for use				
Use to protect family and friends	*39.4*	*2.65 (2.00-3.51)*	61.7	—[f]
Use to help stop the pandemic/reduce deaths in older American adults	*41.3*	*2.04 (1.52-2.73)*	46.8	—
Use as a responsibility to my community	*40.6*	*1.49 (1.10-2.03)*	45.9	—
Use to let me know my risk of infection/help peace of mind/help me stay healthy	*38.3*	*2.09 (1.57-2.78)*	62.0	—
Do not use because it would make me feel more anxious	N/A	N/A	*29.1*	*0.59 (0.47-0.73)*
Do not use because I would not benefit/I won’t be infected	N/A	N/A	21.5	1.14 (0.90-1.44)
Usability				
Essential worker (yes)	23.6	1.12 (0.87-1.46)	34.1	1.06 (0.84-1.34)
Health status				
Very good or excellent	20.7	1	31.0	1
Good	15.6	0.93 (0.72-1.22)	27.7	1.10 (0.88-1.38)
Fair or poor	12.8	0.73 (0.52-1.03)	*31.0*	*1.42 (1.08-1.85)*
Has health insurance (yes)	*16.1*	*0.65 (0.48-0.89)*	30.3	1.24 (0.92-1.68)
Do not use because too much hassle to install	N/A	N/A	32.5	0.93 (0.73-1.17)
Therapeutic goal				
Would you tell your neighbors or friends if you become COVID-19–positive? (yes)	67.4	0.84 (0.65-1.08)	72.0	1.20 (0.97-1.50)
Do you think you have already had the coronavirus or currently have it? (yes)	*28.6*	*1.60 (1.21-2.12)*	32.9	0.96 (0.74-1.26)
Demographics				
Age (years)				
18-29	30.4	1	40.6	1
30-49	22.6	0.93 (0.69-1.25)	38.6	0.86 (0.65-1.15)
50-64	*10.3*	*0.69 (0.47-1.00)*	*25.0*	*0.47 (0.34-0.64)*

Survey segments	Adopted		Interested-not-adopted	
	Respondents (n=2824), %[a]	Adjusted OR[b] (95% CI) (n=490)	Respondents (n=2334), %[a]	Adjusted OR (95% CI) (n=697)
65+	*7.4*	*0.63 (0.41-0.98)*	*18.4*	*0.30 (0.21-0.43)*
Gender: male	18.4	1.14 (0.90-1.44)	27.9	0.89 (0.73-1.08)
Race				
White	15.0	1	28.0	1
Black	21.4	1.06 (0.74-1.53)	33.2	0.87 (0.63-1.20)
Hispanic	29.3	1.25 (0.86-1.83)	41.5	1.30 (0.90-1.87)
Other	23.6	1.45 (0.90-2.34)	35.5	1.01 (0.65-1.57)
Political party affiliation				
Republican	14.4	1	22.5	1
Democratic	21.1	1.05 (0.80-1.39)	*38.1*	*1.62 (1.29-2.03)*
Independent	13.8	0.85 (0.59-1.22)	26.1	0.97 (0.73-1.30)
Annual income level (US $)				
<50,000	15.2	1	26.9	1
50,000-99,000	17.6	1.10 (0.82-1.47)	31.4	1.13 (0.89-1.43)
≥100,000	*24.8*	*1.45 (1.02-2.06)*	*39.3*	*1.42 (1.05-1.91)*
Population density				
Urban	22.7	1	37.5	1
Semiurban	16.4	0.98 (0.75-1.28)	29.1	0.86 (0.68-1.08)
Rural	12.7	0.93 (0.66-1.31)	*23.3*	*0.73 (0.55-0.96)*
Education				
No high school	21.6	1	31.9	1
Some high school	14.9	1.02 (0.62-1.67)	24.1	0.76 (0.49-1.16)
Some college	15.4	0.86 (0.53-1.40)	28.4	0.96 (0.63-1.46)
Bachelor's degree and above	20.7	0.92 (0.55-1.53)	36.2	1.24 (0.80-1.93)

[a] Frequency is presented as a percentage of adopted or interested-not-adopted against the total population of each variable.

[b] OR: odds ratio.

[c] DCT: digital contact tracing.

[d] Statistically significant associations are in italics.

[e] N/A: not applicable.

[f] Removed from model due to collinearity with being in the interested-not-adopted segment.

Essential worker status was not associated with being in the adoption or interest segments, and education was not a significant factor. Being an adopter was negatively associated with having health insurance (16.1%; OR 0.65, 95% CI 0.48-0.89) and being interested was associated with being in fair to poor health (31.0%; OR 1.42, 95% CI 1.08-1.85). Older age was negatively associated with both adoption and interest. Having an annual income over US $100,000 was associated with both adoption and interest. Political party affiliation was not a statistically significant factor in adoption, although those self-declared as Democratic party members had the highest adoption frequency.

Not adopting a DCT app due to concerns about government surveillance (25.0%; OR 0.65, 95% CI 0.52-0.81) and that their phone will be hacked (27.7%; OR 0.78, 95% CI 0.62-0.99) was negatively associated with interest. Adoption was positively associated with visiting friends and family (22.9%; OR 1.63, 95% CI 1.28-2.09), whereas interest was negatively associated with visits (28.1%; OR 0.81, 95% CI 0.66-0.99). Religious services attendance was associated with interest (32.5%; OR 1.29, 95% CI 1.06-1.58). Adopters tended to express feeling lonely (24.9%; OR 1.31, 95% CI 1.00-1.70), while being interested was positively associated with feeling afraid (39.3%; OR 1.38, 95% CI 1.10-1.73). Motivation measures showed the highest frequencies in the adoption (38.3%-41.3%) and interest segments.

Adopters were more likely to think they have or had the coronavirus (28.6%; OR 1.60, 95% CI 1.21-2.12). Adoption was not positively associated with telling friends and neighbors about a positive COVID-19 lab result.

Discussion

Principal Findings

Overall, each layer of the hierarchical AEM framework shows barriers to the ability of DCT to effectively and safely supplement understaffed manual contact tracers and protect the public. With respect to aim 1, we found that a small percentage of the eligible population reported adoption of the app (17.4%), which is lower than estimated adoption rates in the United Kingdom (28%) [39] and New Zealand (31%) [40]. Programs to encourage adoption among those interested in a DCT app (24.7%) could increase adoption, but still fall well below the minimum target threshold of adoption. The not-interested segment is large enough (58.0%) to prevent DCT apps from reaching the population adoption minimum. Low adoption and interest rates are problematic because in order for DCT apps to be effective, high population uptake is needed alongside other control measures [18]. With respect to aim 2, inconsistent associations across prior DCT studies may be due to variable selection differences, which shows the benefit of using the AEM framework to develop study measures a priori [19]. Our results show that DCT apps may not be clinically effective in protecting adopters' friends and family from their frequent maskless contact outside the home given low adoption rates and DCT technology false-negative detection issues. Adopter and interested segments were positively associated with motivations to protect friends and family, community, and the older population; however, this study found several areas of concern regarding beneficence.

Essential workers are at greater risk of exposure as part of their employment, but they adopted DCT apps at a low rate (23.6%). Essential workers were not associated with being part of either the adoption or interest segment, and other studies also found that they had lower intent to adopt and report positive COVID-19 lab tests [24]. Future DCT app design should consider that infection risk varies across subpopulations, and DCT solution design should focus on adoption with higher-risk subpopulations. A study on DCT app user sentiment suggested that DCT apps are largely passive in nature, and could be improved by adding features such as proactive reminder notifications and links to detailed disease-spread information [30].

The US public health department sponsorship of DCT apps was considered a positive attribute in the AEM Access and Background category; however, in the not-interested segment, this was a negative attribute due to the association with distrust of government surveillance. Other studies indicated concerns about privacy that varied by political affiliation [28], whereas concerns about postpandemic surveillance and hacking were not associated with adoption hesitancy in the interested segment. Resolving this conundrum requires further study of the not-interested segment.

DCT apps are available for free on Google Play and Apple Store, and the adoption rate is higher in the 26 US states sponsoring apps [25]. Early studies before DCTs were deployed could only survey people about their intent to adopt [24], and our findings show differences in adoption versus interest subpopulations across the AEM categories. The influence of age, gender, race, and income was found to be inconsistent across studies [24] and variation in culture is suggested to play a role [23]. Our results support previous findings that DCT adoption is associated with a younger population, but did not support the association with well-educated populations [23,26].

Safe use of DCT apps after adoption is a concern. The high volume of social visits warrants further investigation into the influence of risk compensation behavior and an extraversion personality trait. The COVID-19 pandemic has resulted in ongoing related psychological stress accompanying DCT exposure monitoring, reporting, quarantine, and isolation [41]. Adoption and interest are associated with mental health stressors, with worries of losing a home and feelings of loneliness. More research is needed to determine if there are causal links with loneliness, extroversion personality type, DCT adoption, and poor social distancing compliance [35]. Our study suggests that DCT apps should not be considered as passive public health infection reporting tools because of the psychological and behavioral associations [36]. Our findings support a prior recommendation to investigate DCT adoption psychological impacts [36].

Neither adoption nor interest was associated with telling neighbors or friends about testing positive for COVID-19. This is concerning considering that adoption was associated with adopters thinking they had been infected with the coronavirus. DCTs may not be an effective addition to manual contact tracing given that they are not associated with improving upon the rate of infected people sharing contacts with public health contact tracers. Lessons from other stigmatized diseases such as HIV may offer DCT designers insight to encourage the sharing of exposure information to slow transmission [42]. Findings from an HIV mHealth prevention and testing study recommended using evidence-based methods for app development, including links to geospatial prevention information and the ability to access self-testing kits [43]. There is a clinical trial underway for an HIV mHealth app that may produce useful information to DCT developers [44].

Limitations

Adoption and interest were measured by self-described use at the time of the survey and did not include questions about duration of use. The survey was national and did not ask the respondents which of the DCT apps they considered when answering survey questions, although most states use the same GAEN proximity detection platform. The study results occasionally varied from those reported by international DCT deployments, which suggests that local culture can play a role in adoption. The data are representative of the US adult population but not of all app adopters, and we were able to identify important differences between adopters and nonadopters. Importantly, this study is the first to create a postmarket assessment approach for DCTs, and points to the need to prioritize development of multidisciplinary teams at each step of the development, implementation, and evaluation continuum.

This study did not assess the political party control of DCT sponsoring states, which controls if a DCT is offered and how energetically it is marketed. An individual's political party affiliation was not associated with adoption, but political party in control of a regional government is likely a different factor than individual user political affiliation.

Conclusions

With low adoption rates nationally, the DCT apps may only be useful as part of pandemic management programs in smaller, more targeted populations such as company employees, university employees and students, or health system workers and clients.

The AEM highlighted key clinical foundation issues that should be addressed in exposure contact-tracing system design. For example, DCT app adoption is associated with frequent visits with friends and family members outside the home, not wearing masks, not having health insurance, feeling lonely, and adopters thinking they have/had the coronavirus. Designing for behavioral risks requires public health and technology organizations to codevelop multidisciplinary definitions of DCT effectiveness along with measurement methods.

The APA AEM framework combined with observational data analysis is suited to assess end-user mobile app experience and can help with future study measure selection. This postmarket beneficence and safety evaluation approach using the AEM and observational data could be applied to other COVID-19 pandemic public health applications that manage disease exposure, promote personal health tracking, monitor health, and raise awareness [45].

Acknowledgments

The authors wish to thank our colleagues at Climate Nexus for their support, including James Wyatt, Molly Fisch-Friedman, and Reece Rushing, and Dr. Gary Kreps for his insights on health communication. The authors would like to thank the Institute for Biohealth Innovation at George Mason University and the TD Charitable Foundation for financial support.

Data Availability

A subset of survey data that were used in this study and the data dictionary are available in Multimedia Appendix 1.

Authors' Contributions

KEC and AAR had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: KEC. Acquisition, analysis, and interpretation of data: KEC, AAR. Drafting of the manuscript: KEC. Critical revision of the manuscript for important intellectual content: AAR. Statistical analysis: KEC. Obtained funding: AAR. Study supervision: AAR.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey data used in a postmarket evaluation model for digital contract tracing (January 2021).
[XLSX File (Microsoft Excel File), 496 KB - publichealth_v9i1e38633_app1.xlsx]

Multimedia Appendix 2
App evaluation model variable selection process.
[DOCX File , 24 KB - publichealth_v9i1e38633_app2.docx]

References

1. Spencer KD, Chung CL, Stargel A, Shultz A, Thorpe PG, Carter MW, et al. COVID-19 case investigation and contact tracing efforts from health departments - United States, June 25-July 24, 2020. MMWR Morb Mortal Wkly Rep 2021 Jan 22;70(3):83-87. [doi: 10.15585/mmwr.mm7003a3] [Medline: 33476317]
2. Staffing Up: workforce levels needed to provide basic public health services for all Americans. de Beaumont Foundation. 2021 Oct. URL: https://debeaumont.org/wp-content/uploads/2021/10/Staffing-Up-FINAL.pdf [accessed 2022-12-18]
3. Kretzschmar ME, Rozhnova G, Bootsma MCJ, van Boven M, van de Wijgert JHHM, Bonten MJM. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. Lancet Public Health 2020 Aug;5(8):e452-e459 [FREE Full text] [doi: 10.1016/S2468-2667(20)30157-2] [Medline: 32682487]
4. Morley J, Cowls J, Taddeo M, Floridi L. Ethical guidelines for COVID-19 tracing apps. Nature 2020 Jun 28;582(7810):29-31. [doi: 10.1038/d41586-020-01578-0] [Medline: 32467596]
5. Trang S, Trenz M, Weiger WH, Tarafdar M, Cheung CM. One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. Eur J Inf Syst 2020 Jul 27;29(4):415-428. [doi: 10.1080/0960085x.2020.1784046]

6. Stojanovic J, Wübbeler M, Geis S, Reviriego E, Gutiérrez-Ibarluzea I, Lenoir-Wijnkoop I. Evaluating public health interventions: a neglected area in health technology assessment. Front Public Health 2020 Apr 22;8:106 [FREE Full text] [doi: 10.3389/fpubh.2020.00106] [Medline: 32391300]
7. Tarricone R, Petracca F, Ciani O, Cucciniello M. Distinguishing features in the assessment of mHealth apps. Expert Rev Pharmacoecon Outcomes Res 2021 Aug 04;21(4):521-526. [doi: 10.1080/14737167.2021.1891883] [Medline: 33663324]
8. mHealth app developer economics 2016: the current status and trends of the mHealth app market. Research2Guidance. 2016. URL: https://research2guidance.com/product/mhealth-app-developer-economics-2016/ [accessed 2022-12-18]
9. Shuren J, Patel B, Gottlieb S. FDA regulation of mobile medical apps. JAMA 2018 Jul 24;320(4):337-338. [doi: 10.1001/jama.2018.8832] [Medline: 29971339]
10. Postmarket surveillance under Section 522 of the Federal Food, Drug, and Cosmetic Act. US Food and Drug Administration. 2022. URL: https://www.fda.gov/media/81015/download [accessed 2022-12-18]
11. Digital health policies and public health solutions for COVID-19. US Food and Drug Administration. 2022 Apr 28. URL: https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/digital-health-policies-and-public-health-solutions-covid-19 [accessed 2022-12-18]
12. Developing a software precertification program: a working model v1. US Food and Drug Administration. 2019 Jan. URL: https://www.fda.gov/media/119722/download [accessed 2022-12-18]
13. Hensher M, Cooper P, Dona S, Angeles M, Nguyen D, Heynsbergh N, et al. Scoping review: development and assessment of evaluation frameworks of mobile health apps for recommendations to consumers. J Am Med Inform Assoc 2021 Jun 12;28(6):1318-1329 [FREE Full text] [doi: 10.1093/jamia/ocab041] [Medline: 33787894]
14. Moshi MR, Tooher R, Merlin T. Suitability of current evaluation frameworks for use in the health technology assessment of mobile medical applications: a systematic review. Int J Technol Assess Health Care 2018 Jan 11;34(5):464-475. [doi: 10.1017/S026646231800051X] [Medline: 30201060]
15. App Advisor: An American Psychiatric Association Initiative. The App Evaluation Model. American Psychiatric Association. URL: https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/the-app-evaluation-model [accessed 2022-12-18]
16. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. Lancet Digit Health 2019 Jun;1(2):e52-e54 [FREE Full text] [doi: 10.1016/S2589-7500(19)30013-5] [Medline: 33323229]
17. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98 [FREE Full text] [doi: 10.1002/wps.20592] [Medline: 30600619]
18. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. Lancet Digit Health 2020 Nov;2(11):e607-e621 [FREE Full text] [doi: 10.1016/S2589-7500(20)30184-9] [Medline: 32839755]
19. Hafermann L, Becher H, Herrmann C, Klein N, Heinze G, Rauch G. Statistical model building: background "knowledge" based on inappropriate preselection causes misspecification. BMC Med Res Methodol 2021 Sep 29;21(1):196 [FREE Full text] [doi: 10.1186/s12874-021-01373-z] [Medline: 34587892]
20. Lagan S, Aquino P, Emerson MR, Fortuna K, Walker R, Torous J. Actionable health app evaluation: translating expert frameworks into objective metrics. NPJ Digit Med 2020;3:100 [FREE Full text] [doi: 10.1038/s41746-020-00312-4] [Medline: 32821855]
21. Goldberg M, Gustafson A, Maibach E, Ballew M, Bergquist P, Kotcher J, et al. Mask-wearing increased after a government recommendation: a natural experiment in the US during the COVID-19 pandemic. Front Commun 2020 Jun 17;5:5. [doi: 10.3389/fcomm.2020.00044]
22. Auxier B. How Americans see digital privacy issues amid the COVID-19 outbreak. Pew Research Center. 2020 May 04. URL: https://pewrsr.ch/2SCKgoW [accessed 2020-03-01]
23. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav 2021 Feb;5(2):247-255. [doi: 10.1038/s41562-020-01044-x] [Medline: 33479505]
24. Li T, Cobb C, Yang J, Baviskar S, Agarwal Y, Li B, et al. What makes people install a COVID-19 contact-tracing app? Understanding the influence of app design and individual difference on contact-tracing app adoption intention. Pervasive Mob Comput 2021 Aug;75:101439 [FREE Full text] [doi: 10.1016/j.pmcj.2021.101439] [Medline: 36569467]
25. Johnson B. The Covid Tracing Tracker: What's happening in coronavirus apps around the world. MIT Technology Review. 2020 Dec 16. URL: https://www.technologyreview.com/2020/12/16/1014878/covid-tracing-tracker/ [accessed 2022-12-18]
26. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: a source of new digital inequalities? Inf Soc 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550] [Medline: 26281194]
27. Blacklow S, Lisker S, Ng M, Sarkar U, Lyles C. Usability, inclusivity, and content evaluation of COVID-19 contact tracing apps in the United States. J Am Med Inform Assoc 2021 Aug 13;28(9):1982-1989 [FREE Full text] [doi: 10.1093/jamia/ocab093] [Medline: 34022053]
28. Zhang B, Kreps S, McMurry N, McCain RM. Americans' perceptions of privacy and surveillance in the COVID-19 pandemic. PLoS One 2020 Dec 23;15(12):e0242652 [FREE Full text] [doi: 10.1371/journal.pone.0242652] [Medline: 33362218]
29. Rivest R, Canetti R. The PACT protocol specifications. 2020. URL: https://pact.mit.edu/wp-content/uploads/2020/04/The-PACT-protocol-specification-ver-0.1.pdf [accessed 2021-09-19]

30. Mitchell H. Lawsuit alleges Google left millions of users' contact tracing data public. Beckers Hospital Review. 2020 Apr. URL: https://www.beckershospitalreview.com/cybersecurity/lawsuit-alleges-google-left-millions-of-users-contact-tracing-data-public.html [accessed 2022-12-18]
31. Stephens M, Cala G, Greene K, Keenan K, Robinson A, Shi L, et al. Challenges for digital proximity detection in pandemics: Privacy, accuracy, and impact. 2001 Presented at: National Institute of Standards and Technology Workshop; January 26-28, 2001; online. [doi: 10.6028/nist.sp.1268]
32. Hedlund J. Risky business: safety regulations, risks compensation, and individual behavior. Inj Prev 2000 Jun;6(2):82-90 [FREE Full text] [doi: 10.1136/ip.6.2.82] [Medline: 10875661]
33. Yan Y, Bayham J, Richter A, Fenichel EP. Risk compensation and face mask mandates during the COVID-19 pandemic. Sci Rep 2021 Feb 04;11(1):3174. [doi: 10.1038/s41598-021-82574-w] [Medline: 33542386]
34. Vandamme AM, Nguyen T, Denis M, Depypere M, Kiekens A, Naesens L, on behalf of the Coronavirus Pandemic Preparedness team. Factors that should shape the decision-making to deploy contact tracing apps for pandemic containment measures. KU Leuven Institute for the Future. 2020. URL: https://rega.kuleuven.be/if/tracing-tools-for-pandemics [accessed 2022-12-18]
35. Carvalho LDF, Pianowski G, Gonçalves AP. Personality differences and COVID-19: are extroversion and conscientiousness personality traits associated with engagement with containment measures? Trends Psychiatry Psychother 2020 Jun;42(2):179-184. [doi: 10.1590/2237-6089-2020-0029] [Medline: 32294713]
36. Kawakami N, Sasaki N, Kuroda R, Tsuno K, Imamura K. The effects of downloading a government-issued COVID-19 Contact tracing app on psychological distress during the pandemic among employed adults: prospective study. JMIR Ment Health 2021 Jan 12;8(1):e23699 [FREE Full text] [doi: 10.2196/23699] [Medline: 33347424]
37. Lash RR, Donovan CV, Fleischauer AT, Moore ZS, Harris G, Hayes S, Contact Tracing Assessment Team, et al. COVID-19 contact tracing in two counties - North Carolina, June-July 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 25;69(38):1360-1363. [doi: 10.15585/mmwr.mm6938e3] [Medline: 32970654]
38. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 2004 May 01;66(3):411-421. [doi: 10.1097/01.psy.0000127692.23278.a9] [Medline: 15184705]
39. Lewis D. Contact-tracing apps help reduce COVID infections, data suggest. Nature 2021 Mar;591(7848):18-19. [doi: 10.1038/d41586-021-00451-y] [Medline: 33623147]
40. Gasteiger N, Gasteiger C, Vedhara K, Broadbent E. The more the merrier! Barriers and facilitators to the general public's use of a COVID-19 contact tracing app in New Zealand. Inform Health Soc Care 2022 Apr 03;47(2):132-143. [doi: 10.1080/17538157.2021.1951274] [Medline: 34260338]
41. Lohiniva A, Dub T, Hagberg L, Nohynek H. Learning about COVID-19-related stigma, quarantine and isolation experiences in Finland. PLoS One 2021 Apr 14;16(4):e0247962 [FREE Full text] [doi: 10.1371/journal.pone.0247962] [Medline: 33852581]
42. Eaton LA, Kalichman SC. Social and behavioral health responses to COVID-19: lessons learned from four decades of an HIV pandemic. J Behav Med 2020 Jun 25;43(3):341-345 [FREE Full text] [doi: 10.1007/s10865-020-00157-y] [Medline: 32333185]
43. Biello KB, Hill-Rorie J, Valente PK, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Development and evaluation of a mobile app designed to increase HIV testing and pre-exposure prophylaxis use among young men who have sex with men in the United States: open pilot trial. J Med Internet Res 2021 Mar 24;23(3):e25107 [FREE Full text] [doi: 10.2196/25107] [Medline: 33759792]
44. Liu A, Coleman K, Bojan K, Serrano PA, Oyedele T, Garcia A, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Jan 25;8(1):e10659 [FREE Full text] [doi: 10.2196/10659] [Medline: 30681964]
45. Almalki M, Giannicchi A. Health apps for combating COVID-19: descriptive review and taxonomy. JMIR Mhealth Uhealth 2021 Mar 02;9(3):e24322 [FREE Full text] [doi: 10.2196/24322] [Medline: 33626017]

Abbreviations

AEM: App Evaluation Model
APA: American Psychological Association
DCT: digital contact tracing
FDA: Food and Drug Administration
GAEN: Google/Apple exposure notification
HTA: health technology assessment
mHealth: mobile health
OR: odds ratio

Edited by T Sanchez, A Mavragani; submitted 10.04.22; peer-reviewed by A Teles, M Cozad; comments to author 14.12.22; revised version received 13.01.23; accepted 19.01.23; published 22.03.23.

Please cite as:
Cevasco KE, Roess AA
Adaptation and Utilization of a Postmarket Evaluation Model for Digital Contact Tracing Mobile Health Tools in the United States: Observational Cross-sectional Study
JMIR Public Health Surveill 2023;9:e38633
URL: https://publichealth.jmir.org/2023/1/e38633
doi:10.2196/38633
PMID:36947135

Original Paper

Regional Differences in Medical Costs of Chronic Kidney Disease in the South Korean Population: Marginalized Two-Part Model

Minah Park[1*], MPH; Choa Yun[2*], MS; Jae Hong Joo[3], PhD; Soo Hyun Kang[4], PhD; Sung Hoon Jeong[1,5], MPH, PhD; Chung-Mo Nam[6], PhD; Eun-Cheol Park[1,6], MD, PhD; Yoondae Han[7], MD; Sung In Jang[1,6], MD, PhD

[1]Institute of Health Services Research, Yonsei University, Seoul, Republic of Korea
[2]Department of Biostatistics and Computing, College of Medicine, Yonsei University, Seoul, Republic of Korea
[3]Department of Health Informatics and Biostatistics, Graduate School of Public Health, Yonsei University, Seoul, Republic of Korea
[4]Department of Preventive Medicine, College of Medicine, Gachon University, Incheon, Republic of Korea
[5]Department of Rehabilitation Medicine, Seoul National University Hospital, Seoul, Republic of Korea
[6]Department of Preventive Medicine, College of Medicine, Yonsei University, Seoul, Republic of Korea
[7]Division of Colorectal Surgery, Department of Surgery, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Sung In Jang, MD, PhD
Department of Preventive Medicine
College of Medicine
Yonsei University
50 Yonsei-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 1863
Fax: 82 2 392 8133
Email: jangsi@yuhs.ac

Abstract

Background: There are regional gaps in the access to medical services for patients with chronic kidney disease (CKD), and it is necessary to reduce those gaps, including the gaps involving medical costs.

Objective: This study aimed to analyze regional differences in the medical costs associated with CKD in the South Korean population.

Methods: This longitudinal cohort study included participants randomly sampled from the National Health Insurance Service-National Sample Cohort of South Korea. To select those who were newly diagnosed with CKD, we excluded those who were diagnosed in 2002-2003 and 2018-2019. A total of 5903 patients with CKD were finally included. We used a marginalized two-part longitudinal model to assess total medical costs.

Results: Our cohort included 4775 (59.9%) men and 3191 (40.1%) women. Of these, 971 (12.2%) and 6995 (87.8%) lived in medically vulnerable and nonvulnerable regions, respectively. The postdiagnosis costs showed a significant difference between the regions (estimate: –0.0152, 95% confidence limit: –0.0171 to –0.0133). The difference in medical expenses between the vulnerable and nonvulnerable regions showed an increase each year after the diagnosis.

Conclusions: Patients with CKD living in medically vulnerable regions are likely to have higher postdiagnostic medical expenses compared to those living in regions that are not medically vulnerable. Efforts to improve early diagnosis of CKD are needed. Relevant policies should be drafted to decrease the medical costs of patients with CKD disease living in medically deprived areas.

(JMIR Public Health Surveill 2023;9:e39904) doi:10.2196/39904

KEYWORDS

chronic kidney disease; cost analysis; medical expenses; medically vulnerable regions; kidney; public health; cost; economic; chronic disease; insurance; regional; longitudinal model

Introduction

Chronic kidney disease (CKD) has been known as a causative factor for early death and is being recognized as a global health problem [1]. In the United States, 1 in 7 adults are known to have CKD, and recent data show that the number of deaths due to CKD in the past 2 decades has increased. With the increased incidence of CKD, the cost of health care has increased substantially [2]. The most severe phase of CKD, which is end-stage renal disease (ESRD), has cost Medicare US $32.9 billion [2], with earlier stages costing Medicare approximately US $48 billion in 2010 [3]. This phenomenon is also evident in South Korea. In South Korea, the number of patients with CKD has increased by 8.7% annually over the last 5 years (from 2013 to 2017), and CKD was estimated to affect approximately 4.6 million patients as of 2017. The annual medical fee per patient with CKD is 8,361,000 won (US $6367.04), and the annual total medical cost of these patients exceeds 1.7 trillion won (1.2 billion dollars), which makes the per-patient treatment cost for CKD the highest compared with that for other diseases, such as dementia or cancer [4].

In the case of chronic diseases, the effective implementation of medical care, management, and prevention of these diseases are important for social integration [5]. However, in South Korea, despite the need for more medical help in provincial areas, resources and medical expenditure are concentrated in big cities. As of 2009, Seoul was home to 27.6% of the country's medical specialists and 52.4% of its physicians and dentists. In addition, medical expenditure is concentrated in the capital region. As of 2008, Seoul accounted for 26.9% of all insured medical bills, while regions outside Seoul and the capital region accounted for 36.2% and 14.5% of insured medical bills, respectively [6]. Patients with CKD living in rural areas have a higher risk of morbidity, hospitalization, and mortality compared to those living in urban areas [7]. In addition, rural residents are likely to have low income levels and need to travel farther to seek medical care, which can result in further disparities in CKD care [8]. As South Korea becomes an aging society, rural areas are observed to be aging more rapidly than urban areas [9]. Considering that age is an important factor for CKD [10], its treatment is even more challenging in rural areas with high aging rates. Since there are clear economic and regional gaps in access to services in the health system, it is necessary to make efforts to improve access to essential medical services for CKD.

Kim et al [11] studied the cost-utility data among patients with ESRD, although their analysis was limited due to the data being collected from a specific hospital. Another study [12] on the economic burden of CKD in South Korea used a national sample cohort; however, they found no regional differences. Therefore, our study aimed to estimate the regional differences using the position value for relative composite index of the medical costs of CKD in the South Korean population. We used the data available from the Korean Health Insurance Service.

Methods

Data and Study Population

The data for this study were obtained from the National Health Insurance Service- National Sample Cohort (NHIS-NSC). The NHIS-NSC data were collected by random sampling of medical claims, covering 2% of the South Korean population. The period during 2002 to 2003 was designated as a washout period to account for the effects of other existing diseases that might influence the results associated with the hypothesized relationship. To select those who were newly diagnosed with CKD, we excluded those who were diagnosed with CKD in 2002-2003 and 2018-2019. From these data, we extracted 10,019 cases diagnosed with CKD (International Classification of Diseases, 10th revision; code: N18). After excluding patients who were diagnosed with CKD during the excluded period, a total of 7966 individuals were included in the final study (vulnerable regions: n=971, 12.2%; nonvulnerable regions: n=6995, 87.8%).

Ethical Consideration

This study was reviewed and approved by the International Review Board of Yonsei University's Health System (Y-2020-0031) and adheres to the tenets of the Declaration of Helsinki. The NHIS-NSC data do not contain any identifying information; thus additional approval was not required.

Variables

The variable of interest in this study was the region. Position value for relative composite (PARC) indicators were used to categorize the health care level by region in South Korea; the analysis methods have been explained in detail in previous studies [13-18]. PARC is an objective indicator that identifies the relative health care level of a location compared with other regions. The PARC value ranges from −1 to 1, with a value of 1 as the best, 0 as the average, and −1 as the worst value, compared with the mean value of the entire region; the closer the value is to −1, the lower the health care level in the region compared with the average; on the other hand, the closer the value is to 1, the higher the level is compared with the average. In this study, when the PARC value was less than −0.33, the region was classified as a medically vulnerable region.

The dependent variable of this study was the total medical cost. Based on the time of diagnosis, the medical expenses were calculated monthly for 24 months before diagnosis and 24 months after diagnosis. Medical expenses before death were also included.

In addition, analyses that included sex, age, income, and social security type as independent variables were performed. Social security was categorized into health insurance (eg, corporate or regional) and medical aid. One of the social security systems in South Korea, National Health Insurance Service (NHIS), covers the entire population except for medical aid beneficiaries. Since the South Korean government provides insurance for the poor, it offers a medical aid program to people unable to pay for their own health care coverage. The medical aid program beneficiaries are financed by both the local and central government and are a part of the South Korean public assistance system [19].

Statistical Analyses

Marginalized Two-Part Longitudinal Model

To overcome the limitations posed by a conventional two-part model, we used a marginalized two-part (MTP) longitudinal model that directly parameterizes the marginal mean of Y_{ij} with covariates [20]. The MTP model has the same two-part structure as the conventional model, but instead of parametrizing the model with a conditionally positive-valued log-scale location parameter μ_{ij}, it models the marginal mean ν_{ij} that is conditional expectation value of Y_{ij} on b_i for medical users and nonusers combined. Thus, we take the following equation:

We redefined b_i to now have b'_i vector that includes a'_i and d'_i, the random effects are assumed to jointly follow a multivariate normal distribution; this allows cross-part correlation among random effects. Under the MTP parameterization, β_k is the incremental effect in the log of the overall mean, $E(Y_{ij} |b_i)$, that comprises the entire population, including both medical users and nonusers, $E(Y_{ij} |b_i)$, and corresponds to a unit increase in the *k*th covariate, x_{2kij}.

Model Details

The percentage of the study population with 0 expenditure over the study period was 29%, creating the need for a two-part model. To assess the effect of the regional difference on medical costs, we fit the MTP model with explanatory variables in the binary and overall mean components. In addition, we used segmented regression analysis for powerful estimation of intervention effects in an interrupted time series. The full model was constructed as follows:

In the above, $0 \leq time_{ij} \leq 48$ and $0 \leq After\ time_{ij} \leq 24$.

We fit the MTP model using the %MTPmle SAS macro program that correctly estimates model parameters through the likelihood function and can be maximized using SAS PROC NLMIXED. In addition, we chose a model type that used a continuous distribution in the second part of the MTP model: the gamma distribution.

Results

We analyzed the descriptive statistics of the medical costs of CKD by region (Table 1). The people living in the vulnerable region comprised 12.2% (971/7966) of participants, and those in the nonvulnerable region comprised 87.8% (6995/7966).

We analyzed the summary statistics for the semicontinuous outcome at each time point (Table 2). After the diagnosis of CKD, significant differences were found in region, sex, income, and Charlson comorbidity index between the groups.

We calculated the overall means and 95% confidence limits of the MTP model parameters (Table 3). There was a significant difference in difference in pre- and postdiagnosis medical costs between the regions (estimate: –0.0152, 95% CI 0.0171 to –0.0133).

We mathematically expressed the MTP model–estimated slope effects of differences in diagnosis and region (Table 4).

We used the MTP model to estimate the effect of region on medical expenses over time (Table 5). The first 2 years are the estimates, and the following years are predictions. The difference in medical expenses between vulnerable and nonvulnerable regions increased every year.

The overall mean of the values with log scale in Table 2 was calculated (Figure 1). Before the CKD diagnoses, there was a significant difference in the slope or intercept. There was also a clear difference in both the slope and intercept after the diagnosis.

Table 1. Descriptive statistics of the medical costs of chronic kidney disease analyzed by region.

Characteristics	Region			*P* value
	Total (N=7966, 100), n (%)	Vulnerable regions (n=971, 12.2), n (%)	Nonvulnerable regions (n=6995, 87.8), n (%)	
Sex				.40
Male	4775 (59.9)	570 (58.7)	4250 (60.1)	
Female	2790 (40.1)	401 (41.3)	2790 (39.9)	
Age (years)				<.001
<30	559 (7.0)	56 (5.8)	503 (7.2)	
30-39	742 (9.3)	79 (7.5)	669 (9.3)	
40-49	1303 (16.4)	124 (9.5)	1179 (16.9)	
50-59	1769 (22.2)	215 (22.1)	1554 (22.2)	
60-69	2451 (30.8)	343 (35.3)	2108 (30.1)	
70-79	1106 (13.9)	154 (15.9)	952 (13.6)	
>80	36 (0.5)	6 (0.6)	30 (0.4)	
Coverage type				.15
NHI[a], employed, or self-employed	3799 (47.7)	442 (45.5)	3357 (48)	
Medical aid	4167 (52.3)	529 (54.5)	3638 (52)	
Income				<.001
Low	1094 (13.7)	132 (13.6)	962 (13.8)	
Middle	3534 (44.4)	488 (50.3)	3046 (43.6)	
High	3338 (41.9)	351 (36.2)	2987 (42.7)	
CCI[b]				<.001
0	1645 (20.7)	163 (16.8)	1482 (21.2)	
1	2300 (28.9)	266 (27.4)	2034 (29.1)	
>2	4021 (50.5)	542 (55.8)	3479 (50)	

[a] NHI: National Health Insurance.

[b] CCI: Charlson comorbidity index.

Table 2. Summary statistics for the semicontinuous outcome (expenditures) at each time point. Monthly expenditures were calculated by converting South Korean won to US dollars.

Characteristics	Overall expenditures		Expenditures before diagnosis		Expenditures after diagnosis	
	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value
Region		.12		.67		.02
Vulnerable region	313 (411)		209 (303)		416 (682)	
Nonvulnerable region	291 (390)		214 (379)		368 (586)	
Sex		.10		.76		.05
Male	288 (382)		212 (377)		363 (570)	
Female	303 (408)		215 (361)		390 (639)	
Age (years)		.01		<.001		.13
<30	266 (414)		170 (426)		362 (614)	
30-39	266 (420)		162 (452)		370 (637)	
40-49	294 (441)		195 (427)		393 (666)	
50-59	311 (411)		230 (353)		392 (617)	
60-69	306 (369)		236 (339)		375 (573)	
70-79	274 (320)		216 (294)		332 (502)	
>80	250 (254)		237 (296)		262 (361)	
Coverage type		.11		.35		.13
NHI[a], employed, or self-employed	286 (392)		209 (365)		363 (597)	
Medical aid	300 (394)		217 (376)		384 (599)	
Income		.02		.52		.01
Low	318 (384)		225 (372)		411 (598)	
Middle	298 (396)		213 (356)		382 (609)	
High	282 (392)		210 (385)		353 (586)	
CCI[b]		<.001		<.001		<.001
0	173 (311)		108 (278)		237 (472)	
1	265 (385)		177 (316)		354 (622)	
>2	359 (413)		278 (418)		441 (621)	

[a] NHI: National Health Insurance.

[b] CCI: Charlson comorbidity index.

Table 3. Overall means and 95% CIs of the marginalized two-part model parameters.

Characteristics	Parameter	Estimate	95% confidence limit
Binary component			
Intercept	α_1	–0.4169	(–0.4501, –0.3835)
Diagnosis	α_2	–0.03515	(–0.0668, –0.0035)
Region	α_3	0.1875	(0.1533, 0.2218)
Time	α_4	0.0260	(0.0244, 0.0275)
Diagnosis×region	α_5	–0.1124	(–0.1839, –0.0408)
Diagnosis×after time	α_6	–0.0888	(–0.0910, –0.0867)
Region×after time	α_7	–0.0151	(–0.0172, –0.0130)
Diagnosis×region×after time	α_8	0.0112	(0.0091, 0.0133)
Age	γ_1	0.0116	(0.0161, 0.0171)
Sex	γ_2	–0.0816	(–0.0962, –0.0671)
CCI[a]	γ_3	0.3348	(0.3257, 0.3440)
Overall mean component			
Intercept	β_1	11.1894	(11.1561, 11.2226)
Diagnosis	β_2	0.0123	(–0.0119, 0.0366)
Region	β_3	–0.0356	(–0.0611, –0.0101)
Time	β_4	0.0652	(0.0641, 0.0664)
Diagnosis×region	β_5	0.0884	(0.0332, 0.1435)
Diagnosis×after time	β_6	–0.0825	(–0.08430, –0.0807)
Region×after time	β_7	0.0178	(0.0159, 0.0197)
Diagnosis×region×after time	β_8	–0.0152	(–0.0171, –0.0133)
Age	δ_1	–0.0002	(–0.0006, 0.0003)
Sex	δ_2	–0.0635	(–0.0762, –0.0507)
CCI	δ_3	0.3734	(0.3650, 0.3818)

[a] CCI: Charlson comorbidity index.

Table 4. Marginalized two-part model–estimated slope effects of differences in diagnosis and region (mathematical expression).

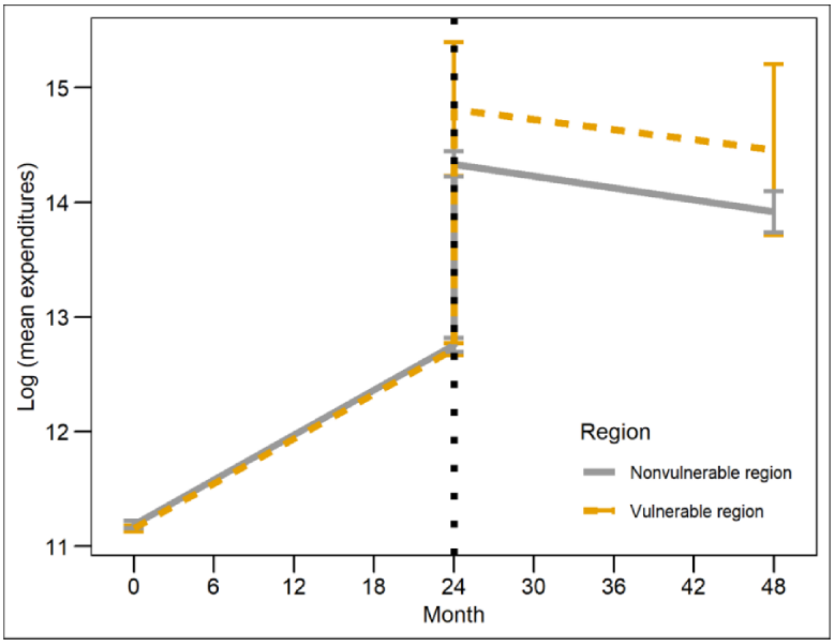
Characteristics	Before	After	After-before
Vulnerable region			
Nonvulnerable region			
Difference of regions			

Table 5. Marginalized two-part model–estimated effects of region over time.

	1 year	2 years	3 years[a]	4 years[a]	5 years[a]
Difference (%)	0.26	0.52	0.78	1.05	1.31

[a] The data used in the marginalized two-part model analysis are for 2 years before and 2 years after diagnosis. The estimated effects at 3, 4, and 5 years are predicted values.

Figure 1. Model of the estimated log (mean expenditures) and 95% CIs for the chronic kidney disease costs.



Discussion

In this study, we examined the differences in medical costs of CKD among the regions of South Korea. The study results demonstrated that there were differences in postdiagnosis medical costs for patients with CKD depending on the region. Patients living in medically vulnerable areas had higher medical costs than patients living in nonvulnerable areas.

These findings are consistent with those of previous studies on chronic obstructive pulmonary disease, as our analysis on this disease (Multimedia Appendix 1), which found that deprivation of health care resulted in an increase in medical costs, emergency health care use, and mortality [21]. As shown in our results, people in medically vulnerable areas were likely to spend more in terms of medical costs. This could be because people living in vulnerable areas discover their symptoms only when conditions have become severe, which results in higher medical costs. Since social deprivation has been found to affect the development of a kidney condition, it is likely that the stage of CKD may worsen [22]. Diagnosis of CKD tends to happen at later stages, resulting in treatment delay and poorer outcomes [23]. Many studies, including those conducted in Canada [24], have found that rural residents with CKD are likely to have less access to specialty care and are provided with low-quality care, which results in a higher prevalence of severe stages of CKD in these areas. As previous studies have shown [3,25], more severe symptoms cause a rapid increase in medical expenses.

Geographical location plays an important role in the treatment of CKD. Some treatments, such as hospital-based hemodialysis, may not be feasible in rural areas. For patients who are dependent on ambulance service, transportation also becomes an important issue [26]. There are many indicators, such as Index Multiple Deprivation, that are used in the United Kingdom to evaluate deprivation within a geographic area [27]. The PARC indicators measure many aspects of the health care level by region in South Korea, rather than simply classifying regions as rural or urban; therefore, they might provide more accurate indicators for evaluation of the actual deprivation of health care regarding the medical cost of CKD.

Since there is an increase in postdiagnosis medical expenses for CKD in vulnerable areas, necessary policies should be implemented to lower the burden of this condition. In addition, efforts to increase early diagnosis of CKD are needed. Currently, most cases of CKD are detected during the course of treating other health problems rather than because of any CKD symptoms. Often, the early stages of CKD show no symptoms, and discovery is made only when conditions become severe. As shown previously [3], early detection of CKD is needed to lower the medical expenses of the patient.

Our study has some important limitations. First, we were not able to perform a random slope analysis in the study because SAS software does not currently provide the necessary program. In future studies, using another statistical program would avoid this issue. Second, since the data set is a collection of medical-claim bills, it is highly likely that the actual number of patients with CKD and their actual burdens are higher than the reported numbers. In general, the number of patients with CKD can be said to represent a pyramid with ESRD at its peak, although the number of patients with CKD based on their treatment performance with health insurance and medical benefits shows the opposite picture [28]. In addition, data on socioeconomic costs that could affect medical expenses, such as transportation, privately hired caregivers, private health insurance, and health supplements, were not available for this study.

However, there are some major strengths to our study. To the best of our knowledge, this study is the first to investigate the regional differences in the medical costs of CKD in the South Korean population using a TPM model. In addition, since all South Korean citizens are obligated to enroll in the NHIS, the NHIS data sets provide nationally representative data.

Our findings suggest that patients with CKD living in medically vulnerable regions are more likely to have increased postdiagnosis medical expenses compared with those living in medically nonvulnerable regions. Over time, the differences in medical expenses are likely to increase substantially. Policies are needed to decrease the medical bills of patients with CKD living in medically deprived areas.

Acknowledgments

This research was supported by a grant (HI20C1130) of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, South Korea. This work was also supported by the National Research Foundation of Korea (NRF; grant 2022R1F1A1062794) funded by the Korea government (Ministry of Science and ICT [information and communication technology]).

The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

All data are available in the database of the Korean National Health Insurance Sharing Service and can be accessed upon reasonable request.

Authors' Contributions

MP and CY conceptualized and designed the study; CY contributed to the acquisition, analysis, and interpretation of data as well as statistical analysis; MP drafted the manuscript; SIJ, ECP, and CMN were in charge of the administrative, technical, or material support of the study; YH and SIJ supervised the study; all authors contributed to the critical revision of the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Overall means and 95% CIs of the marginalized two-part model parameters for chronic obstructive pulmonary disease.

[DOCX File , 15 KB - publichealth_v9i1e39904_app1.docx]

References

1. Park JI, Baek H, Jung HH. Prevalence of chronic kidney disease in Korea: the Korean national health and nutritional examination survey 2011-2013. J Korean Med Sci 2016 Jun;31(6):915-923 [FREE Full text] [doi: 10.3346/jkms.2016.31.6.915] [Medline: 27247501]
2. Hoerger TJ, Simpson SA, Yarnoff BO, Pavkov ME, Ríos Burrows N, Saydah SH, et al. The future burden of CKD in the United States: a simulation model for the CDC CKD initiative. Am J Kidney Dis 2015 Mar;65(3):403-411. [doi: 10.1053/j.ajkd.2014.09.023] [Medline: 25468386]
3. Honeycutt AA, Segel JE, Zhuo X, Hoerger TJ, Imai K, Williams D. Medical costs of CKD in the Medicare population. J Am Soc Nephrol 2013 Sep;24(9):1478-1483 [FREE Full text] [doi: 10.1681/ASN.2012040392] [Medline: 23907508]
4. Yoo C, Kim Y, Kim H. Impact of chronic kidney disease on depression, perceived health, and health-related quality of life: results from the Korea national health and nutrition examination survey in 2014 and 2016. Korean J Adult Nurs 2020;32(4):374-384 [FREE Full text] [doi: 10.7475/kjan.2020.32.4.374]
5. Yun H. Regional difference in outpatient service utilization for chronic diseases among the elderly. Health Policy Manag 2014 Jun 30;24(2):128-135 [FREE Full text] [doi: 10.4332/kjhpa.2014.24.2.128]
6. Hong E, Ahn BC. Income-related health inequalities across regions in Korea. Int J Equity Health 2011 Oct 03;10:41 [FREE Full text] [doi: 10.1186/1475-9276-10-41] [Medline: 21967804]
7. Scholes-Robertson NJ, Howell M, Gutman T, Baumgart A, SInka V, Tunnicliffe DJ, et al. Patients' and caregivers' perspectives on access to kidney replacement therapy in rural communities: systematic review of qualitative studies. BMJ Open 2020 Sep 23;10(9):e037529 [FREE Full text] [doi: 10.1136/bmjopen-2020-037529] [Medline: 32967878]
8. Rao MK, Morris CD, O'Malley JP, Davis MM, Mori M, Anderson S. Documentation and management of CKD in rural primary care. Clin J Am Soc Nephrol 2013 May;8(5):739-748 [FREE Full text] [doi: 10.2215/CJN.02410312] [Medline: 23371962]
9. Min Y, Lee M. Regional differences in population aging and policy responses of local governments. GRI Review 2013;15:113-140 [FREE Full text]

10. Prakash S, O'Hare AM. Interaction of aging and chronic kidney disease. Semin Nephrol 2009 Sep;29(5):497-503 [FREE Full text] [doi: 10.1016/j.semnephrol.2009.06.006] [Medline: 19751895]
11. Kim S, Kim Y, Kim M. A renal transplantation and hemodialysis cost-utility analysis in patients with end-stage renal disease. J Korean Soc Transplant 2010;24(3):173-181 [FREE Full text] [doi: 10.4285/jkstn.2010.24.3.173]
12. Kim S, Jo M, Go D, Ryu D, Park J. Economic burden of chronic kidney disease in Korea using national sample cohort. J Nephrol 2017 Dec;30(6):787-793. [doi: 10.1007/s40620-017-0380-3] [Medline: 28303461]
13. Chae W, Lee S, Park E. Position value for relative comparison of healthcare status of Korea among Organization for Economic Cooperation and Development countries, 2015. Health Policy Manag 2018;28(1):98-103 [FREE Full text]
14. Jang S, Nam J, Choi J, Park E. Disease management index of potential years of life lost as a tool for setting priorities in national disease control using OECD health data. Health Policy 2014 Mar;115(1):92-99. [doi: 10.1016/j.healthpol.2013.11.007] [Medline: 24365032]
15. Youn H, Lee H, Park E. Position value for relative comparison of healthcare status of Korea in 2018. Health Policy Manag 2021;31(2):217-224 [FREE Full text]
16. Lee H, Oh S, Park E. Position value for relative comparison of healthcare status of Korea in 2017. Health Policy Manag 2020;30(1):131-138 [FREE Full text]
17. Oh S, Park E. Position value for relative comparison of healthcare status of Korea in 2016. Health Policy Manag 2019;29(1):90-97 [FREE Full text]
18. Park M, Youn H, Park E. Position value for relative comparison of healthcare status of korea in 2019: comparison with countries of the organization for economic cooperation and development. Health Policy Manag 2021;32(1):113-121 [FREE Full text]
19. Jeong W, Kim H, Joo JH, Jang S, Park E. Association between depression and risk of Parkinson's disease in South Korean adults. J Affect Disord 2021 Sep 01;292:75-80 [FREE Full text] [doi: 10.1016/j.jad.2021.05.038] [Medline: 34102551]
20. Smith VA, Neelon B, Preisser JS, Maciejewski ML. A marginalized two-part model for longitudinal semicontinuous data. Stat Methods Med Res 2017 Aug;26(4):1949-1968. [doi: 10.1177/0962280215592908] [Medline: 26156962]
21. Collins PF, Stratton RJ, Kurukulaaratchy RJ, Elia M. Influence of deprivation on health care use, health care costs, and mortality in COPD. Int J Chron Obstruct Pulmon Dis 2018;13:1289-1296 [FREE Full text] [doi: 10.2147/COPD.S157594] [Medline: 29719384]
22. Merkin SS, Coresh J, Diez Roux AV, Taylor HA, Powe NR. Area socioeconomic status and progressive CKD: the Atherosclerosis Risk in Communities (ARIC) Study. Am J Kidney Dis 2005 Aug;46(2):203-213. [doi: 10.1053/j.ajkd.2005.04.033] [Medline: 16112038]
23. Bello AK, Peters J, Rigby J, Rahman AA, El Nahas M. Socioeconomic status and chronic kidney disease at presentation to a renal service in the United Kingdom. Clin J Am Soc Nephrol 2008 Sep;3(5):1316-1323 [FREE Full text] [doi: 10.2215/CJN.00680208] [Medline: 18579673]
24. Bello AK, Ronksley PE, Tangri N, Kurzawa J, Osman MA, Singer A, et al. Prevalence and demographics of CKD in canadian primary care practices: a cross-sectional study. Kidney Int Rep 2019 Apr;4(4):561-570 [FREE Full text] [doi: 10.1016/j.ekir.2019.01.005] [Medline: 30993231]
25. Wyld MLR, Lee CMY, Zhuo X, White S, Shaw JE, Morton RL, et al. Cost to government and society of chronic kidney disease stage 1-5: a national cohort study. Intern Med J 2015 Jul;45(7):741-747. [doi: 10.1111/imj.12797] [Medline: 25944415]
26. Loft R. Report reveals disparities in kidney health across the UK. Kidney Research UK. 2018 Mar 13. URL: https://www.kidneyresearchuk.org/2019/03/13/report-reveals-disparities-in-kidney-health-across-the-uk/ [accessed 2021-09-27]
27. Jordan H, Roderick P, Martin D. The Index of Multiple Deprivation 2000 and accessibility effects on health. J Epidemiol Community Health 2004 Mar;58(3):250-257 [FREE Full text] [doi: 10.1136/jech.2003.013011] [Medline: 14966241]
28. Yoo K, Choi J, Kim B, Kim T. Socioeconomic burden of chronic kidney disease in South Korea. Korean Public Health Res 2014;40:13-23. [doi: 10.22900/kphr.2014.40.3.002]

Abbreviations

CKD: chronic kidney disease
ESRD: end-stage renal disease
MTP: marginalized two-part
NHIS-NSC: National Health Insurance Service-National Sample Cohort
NHIS: National Health Insurance Service
PARC: position value for relative composite

Edited by G Eysenbach; submitted 27.05.22; peer-reviewed by M Singh, M Burnier; comments to author 11.11.22; revised version received 03.01.23; accepted 19.01.23; published 30.03.23.

Please cite as:
Park M, Yun C, Joo JH, Kang SH, Jeong SH, Nam CM, Park EC, Han Y, Jang SI
Regional Differences in Medical Costs of Chronic Kidney Disease in the South Korean Population: Marginalized Two-Part Model
JMIR Public Health Surveill 2023;9:e39904
URL: https://publichealth.jmir.org/2023/1/e39904
doi:10.2196/39904
PMID:36995749

Original Paper

Racial/Ethnic Disparity in Association Between Fetal Alcohol Syndrome and Alcohol Intake During Pregnancy: Multisite Retrospective Cohort Study

Sarah Soyeon Oh[1,2], PhD; Bada Kang[3], PhD; Jewel Park[4], MD, MPH; SangMin Kim[5], MD; Eun-Cheol Park[2,6], MD, PhD; Seung Hee Lee[7], MSN; Ichiro Kawachi[1], MBChB, PhD

[1]Department of Social and Behavioral Sciences, Harvard TH Chan School of Public Health, Boston, MA, United States
[2]Institute of Health Services Research, Yonsei University College of Medicine, Seoul, Republic of Korea
[3]Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, Seoul, Republic of Korea
[4]Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[5]Harvard Medical School, Boston, MA, United States
[6]Department of Preventive Medicine and Public Health, Yonsei University College of Medicine, Seoul, Republic of Korea
[7]College of Nursing and Brain Korea 21 Four Project, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Bada Kang, PhD
Mo-Im Kim Nursing Research Institute
Yonsei University College of Nursing
50-1 Yonsei-Ro
Seodaemun-Gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 3283
Email: bdkang@yuhs.ac

Related Article:

This is a corrected version. See correction statement: https://publichealth.jmir.org/2023/1/e48460

Abstract

Background: Alcohol consumption during pregnancy is associated with a range of adverse birth-related outcomes, including stillbirth, low birth weight, preterm birth, and fetal alcohol syndrome (FAS). With more than 10% of women consuming alcohol during pregnancy worldwide, it is increasingly important to understand how racial/ethnic variations affect FAS onset. However, whether race and ethnicity inform FAS risk assessment when daily ethanol intake is controlled for remains unknown.

Objective: This study aimed to assess racial/ethnic disparities in FAS risk associated with alcohol consumption during pregnancy.

Methods: We used data from a longitudinal cohort study (the Collaborative Initiative on Fetal Alcohol Spectrum Disorders) at 5 hospital sites around the United States of 595 women who consumed alcohol during pregnancy from 2007 to 2017. Questionnaires, in-person interviews, and reviews of medical, legal, and social service records were used to gather data on average alcoholic content (AAC) during pregnancy. Self-reports of maternal race (American Indian/Alaska Native [AI/AN], Asian, Native Hawaiian or other Pacific Islander, Black or African American, White, more than one race, and other) and ethnicity (Hispanic/Latino or not Hispanic/Latino), as well as FAS diagnoses based on standardized dysmorphological criteria, were used for analysis. Log-binomial regression was used to examine the risk of FAS associated with each 1-gram increase in ethanol consumption during pregnancy, stratified by race/ethnicity.

Results: A total of 3.4% (20/595) of women who reported consuming alcohol during pregnancy gave birth to a baby with FAS. Women who gave birth to a baby with FAS had a mean AAC of 32.06 (SD 9.09) grams, which was higher than that of women who did not give birth to a baby with FAS (mean 12.07, SD 15.87 grams). AI/AN mothers with FAS babies had the highest AAC (mean 42.62, SD 8.35 grams), followed by White (mean 30.13, SD 4.88 grams) and Black mothers (mean 27.05, SD 12.78 grams). White (prevalence ratio [PR] 1.10, 95% CI 1.03-1.19), Black (PR 1.13, 95% CI 1.04-1.23), and AI/AN (PR 1.10, 95% CI 1.00-1.21) mothers had 10% to 13% increased odds of giving birth to a baby with FAS given the same exposure to alcohol during pregnancy. Regardless of race, a 1-gram increase in AAC resulted in a 4% increase (PR 1.04, 95% CI 1.02-1.07) in the chance of giving birth to a baby with ≥2 facial anomalies (ie, short palpebral fissures, thin vermilion border of the upper lip, and smooth philtrum) and a 4% increase (PR 1.04, 95% CI 1.01-1.07) in the chance of deficient brain growth.

Conclusions: The risk of delivering a baby with FAS was comparable among White, Black, and AI/AN women at similar levels of drinking during pregnancy. Regardless of race, a 1-gram increase in AAC resulted in increased odds of giving birth to a baby with facial anomalies or deficient brain growth.

(JMIR Public Health Surveill 2023;9:e45358) doi:10.2196/45358

KEYWORDS

fetal alcohol syndrome; ethnic disparity; alcohol intake; pregnancy; health disparity; public health; minority population; vulnerable population; women's health; pediatrics; fetal health; pediatrics

Introduction

Alcohol consumption during pregnancy is associated with a range of adverse birth-related outcomes, including stillbirth, low birth weight, preterm birth, and fetal alcohol spectrum disorders (FASDs) [1]. FASDs range from physical defects to cognitive, social, behavioral, and emotional impairments [2]. Fetal alcohol syndrome (FAS) is the most severe form of FASDs, which are defined as a range of lifelong congenital anomalies, pre- or postnatal growth restrictions, and dysmorphic facial features originating from prenatal alcohol exposure [3].

Globally, around 10% of women consume alcohol during pregnancy, of whom 1 in 13 will deliver a child with a FASD, and 1 in 67, a child with FAS [4]. Previous studies have reported racial/ethnic variations in alcohol consumption during pregnancy [4]. Black, Asian/Pacific Islander, and Hispanic women drink less than White women during pregnancy [5], while non-Hispanic Black women are up to 48% and 62% less likely to drink alcohol during pregnancy and binge drink, respectively, compared to non-Hispanic White women [6]. Controlling for socioeconomic background, social support, paternal health behaviors, and maternal medical history, Black and Hispanic mothers are 41% and 58% less likely to drink during pregnancy, respectively, compared to White mothers [7].

However, although White women may drink more alcohol than Black/African American or Hispanic women before and during pregnancy [8], they are less likely than their counterparts to be at risk of adverse birth-related outcomes [9]. The fetal mortality rate among Black women (11.2%) is nearly twice that of White (5.9%) and Hispanic (5.1%) women in the United States [9]. Black women are also known to be at greater risk of both unintended pregnancies and poor pregnancy outcomes, such as preterm birth or low birth weight [10].

Concerning FAS and FASDs, Black infants have been reported to have a 7-fold higher risk of FAS compared to White infants, while infants born in some Southwestern Native American communities have prevalence rates of up to 17.9 per 1000 individuals [11] (US average: 2-7 per 1000 infants) [12]. To date, there is uncertainty regarding the racial disparities in FASDs, as there is a possibility of misestimation of prevalence. Black infants and children are more likely to be misdiagnosed or receive a later diagnosis due to limited access to care, stigmas, or lower socioeconomic status [13]. Contrastingly, concerns have been raised that clinician racial bias results in higher reports of prenatal substance exposure among Black children and lower reports among Hispanic children than White children [13]. Racial and ethnic differences in facial morphology will also result in an overestimation of FASDs among certain African American populations, especially in the absence of a race-specific lip/philtrum guide and 3D evaluations of philtrum height [14].

Advances in technology such as 3D ultrasound allow accurate assessment of dysmorphic facial features among fetuses with FASD [15]. Nevertheless, researchers are questioning whether the current reference standards for facial measurements in assessing FASDs account for such biases and racial/ethnic variations [16]. Current Institute of Medicine (IOM) diagnostic guidelines recommend that ethnic phenotypes should be considered, as what may seem dysmorphic in one racial/ethnic context (ie, lip/philtrum abnormalities or growth retardation, that is, height or weight ≤10th percentile) may be normal in another [17]. Hence, we sought to examine racial/ethnic variations in the prevalence of FASDs using high-quality diagnostic data in which each instance of a FASD was assessed using uniform criteria rather than an administrative assessment (eg, discharge codes on electronic health records).

Methods

Study Protocol

As part of a multisite research study of patients with FASD and mothers worldwide conducted since 2003, we analyzed data collected by the Collaborative Initiative on FASDs [18]. We restricted our analyses to sites within the United States, including (1) the Center for Behavioral Teratology at San Diego State University (San Diego, CA); (2) Emory University (Atlanta, GA); (3) 7 Northern Plains communities, including 6 Indian reservations; (4) the University of California, Los Angeles (UCLA; Los Angeles, CA); and (5) the University of Minnesota (Minneapolis, MN) [18]. Methods for obtaining information about prenatal alcohol exposure varied by site and ranged from questionnaires and in-person interviews to reviews of medical, legal, or social service records [18]. For all locations, histories of prenatal alcohol exposure were verified retrospectively [19]. At all sites, the participants were examined for FAS using standardized methodology, as reported by Mattson et al [18,20].

Although the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD) also included some women with no alcohol consumption during pregnancy or with unverified prenatal alcohol exposure histories, these samples were not representative of the source populations. Therefore, reliable estimates for average alcoholic content (AAC) could not be calculated. Hence, our analyses were restricted to subjects with at least one incident of verified alcohol consumption during pregnancy. Our objective was to estimate the risk of FASD per gram of pure ethanol intake among women who reported drinking during pregnancy.

Study Sites

Center for Behavioral Teratology Study Site

At this site, children suspected of alcohol exposure were referred to the principal investigator and local professionals for participation in this project [18]. Many patients were already attending this center before the initiation of the CIFASD project, including patients referred to the investigative team for meeting the traditional diagnostic criteria for FAS (eg, facial anomalies, growth retardation, evidence of central nervous system dysfunction (eg, microcephaly, mental retardation, or attentional deficits) [20]. Alcohol exposure histories were obtained via self-reports or professional reviews of medical, legal, or social service records of the biological mother. Parents or primary caregivers completed questionnaires regarding the children’s behavior, while the children were examined for facial features of FAS (ie, 2 of the 3 key facial features: short palpebral fissures [≤10th percentile], a thin vermilion border on the upper lip [rank 4 or 5 on a racially normed lip/philtrum guide], a smooth philtrum [rank 4 or 5 on a racially normed lip/philtrum guide], and signs of pre- or postnatal growth deficiency [head circumference, height, or weight ≤10th percentile]) [19].

Emory University Study Site

At Emory University, the Fetal Alcohol and Drug Exposure Clinic gathers data on a large sample of patients with FASDs while providing clinical services and facial evaluations at the Emory University Marcus Institute [21]. In the absence of direct reports, documentation of alcohol abuse or dependence by the biological mother in the form of medical, social services, or court records was reviewed [18]. Recruitment took place via clinical and community referrals. Parents or primary caregivers completed questionnaires and interviews, while patients with FASD were administered various neuropsychological tests over a 3-hour session [22].

Northern Plains Study Site

Seven communities, including 1 urban and 6 reservation sites throughout North Dakota, South Dakota, and Montana, participated in this study [21]. Children with FASDs were recruited via active case ascertainment methods and advertisements in tribal and community health centers [19]. Data on prenatal alcohol exposure were obtained from in-person interviews with the parent or primary caregiver to retrospectively obtain exact exposure histories and were also confirmed via reviews of medical records when available [19].

University of Minnesota Study Site

The Department of Psychiatry at the University of Minnesota collected data on prenatal alcohol exposure histories through several modalities, including medical reports, birth records, social service records, and, when available, maternal self-reports [21].

UCLA Study Site

Data were collected from children attending the Fetal Alcohol and Related Disorders Clinic at UCLA [23]. Participant recruitment was through local FASD clinic referrals, online advertisements, and word of mouth in caregiving communities [23]. All alcohol exposure histories were confirmed via in-person interviews and maternal reports of prenatal substance exposure or reviews of maternal medical records by a licensed medical doctor [23].

Study Participants

Among children with confirmed prenatal alcohol exposure, a diagnosis of FAS was made if 2 of the 3 key facial features of FAS (ie, short palpebral fissure, smooth philtrum, and thin vermillion border) were accompanied by either microcephaly or growth retardation [18]. Children were excluded when there were reports of known causes of mental deficiency, such as congenital hypothyroidism, neurofibromatosis, or chromosomal abnormalities [18].

Ethical Considerations

This study was approved by the institutional review board of the Harvard TH Chan School of Public Health (IRB21-1261). Primary data collection at each clinical location of the CIFASD received IRB approval, and informed consent was obtained from all adult participants or their legal guardians [24]. For secondary analysis of the data, our research team was provided with deidentified and anonymized data upon request and approval from CIFASD’s data committee. As a multisite study across several locations and time periods, compensation amount and type varied by collection site; for example, at Emory University, participants were given US $50 as monetary compensation [24,25], while at the University of Southern California and University of Minnesota, participants were compensated for their time with gift cards [26].

Measures

Alcohol Consumption

To calculate the grams of ethanol consumed per day during pregnancy for each pregnant woman, maternal reports of alcohol consumption frequency, quantity, and preferred alcoholic beverage were used. All women were asked to report the frequency (1-2 times during pregnancy, 3-5 times, once/month, 2-3 times/week, almost daily, or daily) and quantity (eg, 2) of alcohol exposure during pregnancy, as well as their preferred alcoholic beverage (eg, beer). Only cases of confirmed exposure were included. Each beverage was converted to grams of ethanol based on average alcoholic content (AAC), that is, 13.6 grams of pure alcohol or one standard drink: 340 mL of regular beer (approximately 5% alcohol), 150 mL of wine (approximately 12% alcohol), and 45 mL of distilled spirits (approximately 40% alcohol) [18]. Upon converting all frequency measures to a daily dose (eg, 1.5/280 days throughout the gestational period, with 1 to 2 incidents of consumption during their entire pregnancy), quantity and AAC were multiplied to determine the grams of pure ethanol consumed by each woman daily.

Race/Ethnicity

All participants were asked to self-report their ethnicity as Hispanic/Latino or not Hispanic/Latino. Then, they were asked to self-report their race as American Indian/Alaska Native (AI/AN), Asian, Native Hawaiian or other Pacific Islander, Black/African American, White, more than one race, or other. Owing to the limited number of participants in this investigation for whom there were available data on AAC, they were divided into White, Black, AI/AN, or other.

Outcome Assessment

At 4 of the sites (excluding the Northern plains), a dysmorphologist was trained to accurately diagnose FAS based on physical features as defined by the CIFASD dysmorphology core [27]. The CIFASD dysmorphology core has been tested in previous studies and has good interrater agreement for height and head circumference, moderate-to-fair agreement for facial anomalies, and poor agreement for the thin vermilion border on the upper lip [28]. Site differences for demographic variables including race, ethnicity, and age group were assessed via a 4-way interaction multivariate analysis of variance (MANOVA) but are not reported, as the effects of demographic variables varied by site [27].

Contrastingly, for the Northern Plains site, a team of physicians, teachers, and other representatives was trained to identify children with certain morphological characteristics of FASDs and other birth defects, as well as IQ and neuropsychologic traits; however, a pediatric dysmorphologist could not be consulted for verification. As “a single generalized set of nonspecific normative physical measures” was often insensitive to local ethnic variations in FASD morphology, syndromic features of FASDs were sometimes compared to normal controls within the same population in terms of weight, head circumference, fissure length, and other facial characteristics (eg, ptosis, and intercanthal distance) [29].

For all CIFASD sites, physical growth and dysmorphology features were recorded on a standardized and weighted checklist, and the IOM criteria for FASDs were used for FAS diagnosis: (1) evidence of a characteristic pattern of minor facial anomalies including at least ≥2 key facial features of FAS (palpebral fissures ≤10th percentile, thin vermilion border, or smooth philtrum), (2) evidence of pre- or postnatal growth retardation (height or weight ≤10th percentile), (3) evidence of deficient brain growth (structural brain anomalies or occipitofrontal circumference ≤10th percentile), and, if possible, (4) confirmation of maternal alcohol consumption directly from the mother or a knowledgeable collateral source [30].

For FAS severity, the IOM-recommended Diagnostic Criteria for FAS and Alcohol-Related Effects were used to determine severity of symptoms in terms of the number of FASD features and type of features (ie, facial anomalies, pre- or postnatal growth deficiency, deficient brain growth, and neurobehavioral impairment) for each patient [30]. FAS was considered the most severe form of FASD, with characteristic facial patterns, growth retardation, and neurodevelopmental abnormalities of the central nervous system (CNS). This was followed by partial FAS, defined as the presence of some components of FAS facial patterns and any of the aforementioned characteristics of FAS [30]; alcohol-related birth defects, which result from congenital defects and malformations or dysplasias of the heart, bone, kidney, vision, or hearing systems; and alcohol-related neurodevelopmental disorders, which include CNS neurodevelopmental abnormalities and complex behavioral and cognitive deficits [30].

Likewise, across all sites, the Child Behavior Checklist (CBCL) was completed by caregivers to collect information regarding child psychopathology and behavioral or emotional functioning [31]. All scores were compiled as a parent-completed survey assessment of their child to determine whether prenatal alcohol exposure affected the child’s ability to internalize or externalize difficult situations; adhere to rules and guidelines; and avoid social, thought, and attention problems [32].

Statistical Analysis

Descriptive statistics of the overall sample for mean AAC were calculated for participants according to race/ethnicity, maternal age at childbirth, tobacco and substance use (ie, marijuana, amphetamine, cocaine, hallucinogens, heroin, and methadone) during pregnancy, child sex, CBCL score, and clinical location of investigation. To examine the association between race/ethnicity and FAS, as well as FASDs, such as facial anomalies, growth retardation, and deficient brain growth, a log-binomial regression analysis was used while adjusting for child sex, child age, maternal age, and clinic location (as a dummy variable). Analyses were further stratified by race/ethnicity while accounting for a 1-gram unit increase in AAC. All analyses, including the GPLOT SAS procedure for displaying graphs, were performed using SAS software (version 9.4; SAS Institute).

Results

Table 1 shows the general characteristics of our sample by FAS diagnosis. Overall, 3.4% (20/595) of women who reported consuming alcohol during pregnancy gave birth to a baby with FAS. Women who gave birth to a baby with FAS had a mean AAC of 32.06 (SD 9.09) grams, which was higher than women who did not (mean 12.07, SD 15.87 grams). AI/AN mothers with FAS babies had the highest AAC (mean 42.62, SD 8.35 grams), followed by White (mean 30.13, SD 4.88 grams) and Black mothers (mean 27.05, SD 12.78 grams).

As seen in Figure 1, there was a dose-response relationship between AAC and FAS for exposure during different pregnancy trimesters. For women who only consumed alcohol in the first trimester, there was a 25% probability of giving birth to a baby with FAS when they consumed more than 40 grams of pure ethanol (ie, approximately 2.9 standard drinks). Contrastingly, for women who kept drinking alcohol throughout all 3 trimesters of their pregnancy, there was a 25% probability of giving birth to a baby with FAS when they consumed around 30 grams of pure ethanol (ie, approximately 2.1 standard drinks).

Table 1. General characteristics of sample by fetal alcohol syndrome diagnosis.

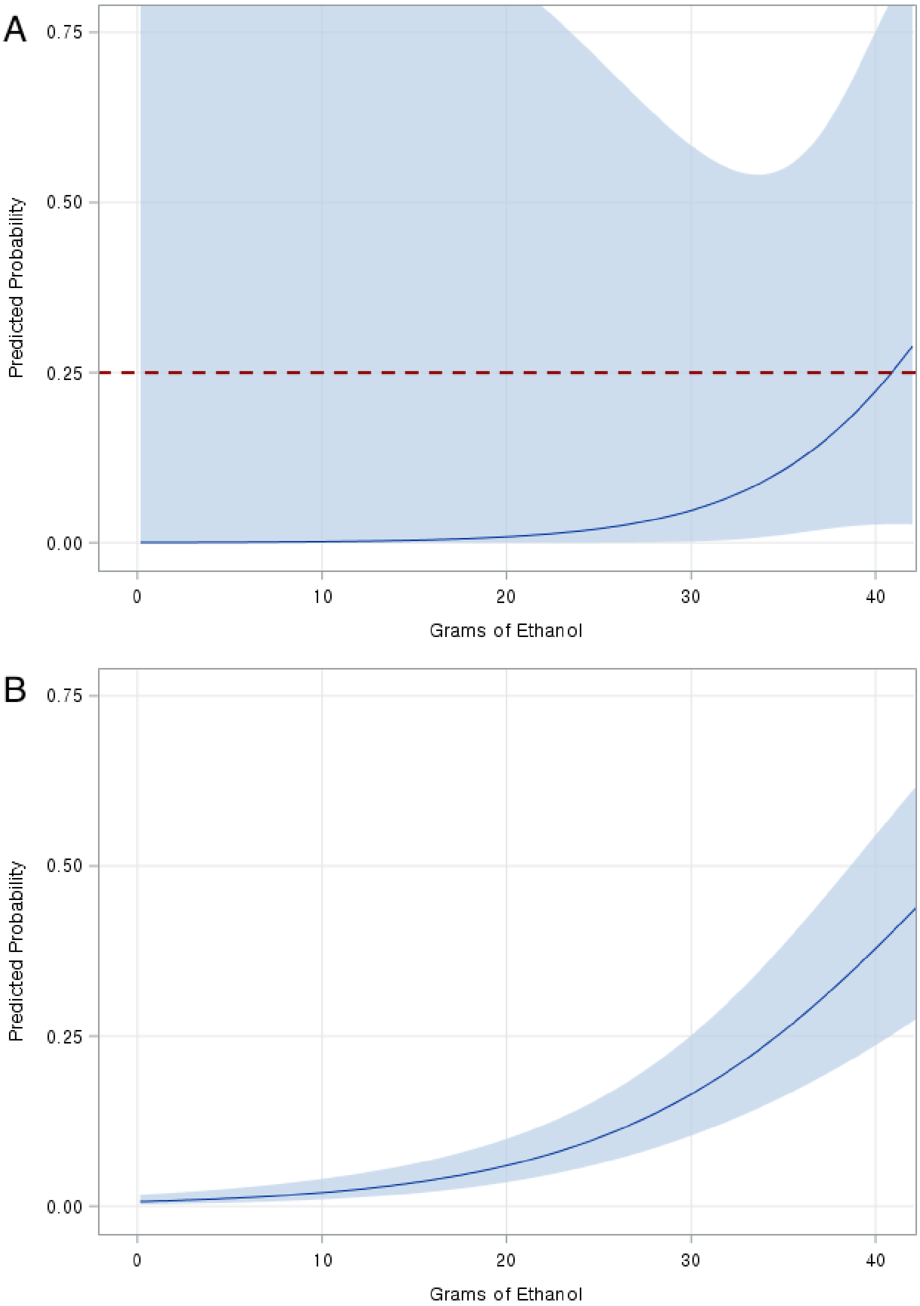
Characteristics	Gave birth to a baby with fetal alcohol syndrome		Did not give birth to a baby with fetal alcohol syndrome	
	Participants, n (%)	AAC[a] (grams/day), mean (SD)	Participants, n (%)	AAC (grams/day), mean (SD)
Race				
White (n=292)	11 (3.8)	30.13 (4.88)	281 (96.2)	12.29 (15.66)
Black (n=254)	4 (1.6)	27.05 (12.78)	250 (98.4)	11.81 (11.55)
American Indian/Alaska Native (n=23)	3 (13)	42.62 (8.35)	20 (87)	19.32 (26.13)
Other (n=26)	2 (7.7)	37.80[b]	24 (92.3)	10.63 (17.48)
Ethnicity				
Hispanic or Latino (n=517)	19 (3.7)	32.49 (9.21)	498 (96.3)	12.18 (13.71)
Not Hispanic or Latino (n=78)	1 (1.3)	25.20[b]	77 (98.7)	11.13 (15.43)
Maternal age at childbirth (years)				
<30 (n=297)	9 (3)	31.68 (12.43)	288 (97)	10.87 (13.88)
30-35 (n=95)	8 (8.4)	30.86 (5.21)	87 (91.6)	12.21 (15.05)
>35 (n=203)	3 (1.5)	37.80[b]	200 (98.5)	13.82 (13.36)
Tobacco used during pregnancy				
No (n=494)	4 (0.8)	35.20 (4.50)	490 (99.2)	8.60 (11.22)
Yes (n=101)	16 (15.8)	29.78 (8.04)	85 (84.2)	18.76 (11.56)
Substances used during pregnancy				
None (n=514)	9 (1.8)	28.03 (10.01)	505 (98.2)	9.97 (11.47)
Marijuana only (n=63)	7 (11.1)	33.60 (6.51)	56 (88.9)	16.96 (12.81)
Other (amphetamine, cocaine, hallucinogens, heroin, methadone) (n=18)	4 (22.2)	30.75 (5.22)	14 (77.8)	22.64 (9.86)
Child sex				
Male (n=331)	8 (2.4)	32.92 (9.51)	323 (97.6)	12.80 (9.71)
Female (n=264)	12 (4.6)	31.46 (9.25)	252 (95.5)	10.96 (14.21)
CBCL score (points)				
Normal (<65) (n=447)	7 (1.6)	36.31 (8.70)	440 (98.4)	9.40 (13.30)
Borderline (65-69) (n=65)	6 (9.2)	26.43 (9.94)	59 (90.8)	14.57 (12.52)
Clinical (>69) (n=83)	7 (8.4)	33.72 (5.84)	76 (91.6)	18.08 (14.71)
Clinical location				
San Diego University (n=101)	2 (2)	37.80[b]	99 (98)	17.88 (15.10)
Emory University (n=242)	10 (4.1)	31.82 (6.49)	232 (95.9)	10.73 (17.20)
Northern Plains (n=51)	5 (9.8)	39.47 (9.29)	46 (90.2)	24.22 (11.26)
Minnesota (n=165)	0 (0)	N/A[c]	165 (100)	9.91 (17.05)
University of California, Los Angeles (n=36)	3 (8.3)	21.00 (7.27)	33 (91.7)	11.19 (15.06)
Total (n=595)	20 (3.4)	32.06 (9.09)	575 (96.6)	12.07 (15.87)

[a] AAC: average alcoholic content.

[b] SD values missing due to the number of participants being too low.

[c] N/A: not applicable.

Figure 1. Dose-response effect of average alcoholic content from 0 grams to 45 grams per day on predicted probability of fetal alcohol syndrome. (A) First-trimester drinking only; (B) drinking throughout all trimesters.



As seen in Table 2, AI/AN (PR 9.44, 95% CI 1.78-50.09) mothers and mothers who used tobacco during pregnancy (PR 13.78, 95% CI 4.18-45.35) had significantly increased odds of giving birth to a baby with FAS relative to their counterparts.

As seen in Table 3, when accounting for a 1-gram increase in alcohol consumption during pregnancy, White (PR 1.10, 95% CI 1.03-1.19), Black (PR 1.13, 95% 1.04-1.23), and AI/AN (PR 1.10, 95% CI 1.00-1.21) mothers had 10% to 13% increased odds of giving birth to a baby with FAS given the same exposure to ethanol during pregnancy.

Regardless of race/ethnicity, as seen in Table 4, every 1-gram increase in alcohol consumption during pregnancy was associated with an 11% increased risk of FAS (PR 1.11, 95% CI 1.07-1.15) and 7% increased risk of alcohol-related brain damage (PR 1.07, 95% CI 1.02-1.13). Each 1-gram increase in AAC was also associated with a 4% increase in the chance of giving birth to a baby with ≥2 facial anomalies (ie, short palpebral fissures, thin vermilion border of the upper lip, or a smooth philtrum; PR 1.04, 95% CI 1.02-1.07) or deficient brain growth (PR 1.04, 95% CI 1.01-1.07).

Table 2. Log-binomial regression analysis of association between sociodemographic variables and fetal alcohol syndrome, adjusted for child sex, child age, maternal age, and clinical location.

Sociodemographic variables	Prevalence ratio (95% CI)	*P* value
Race		
White (reference)	1.00	
Black	0.44 (0.14-1.39)	.16
American Indian/Alaska Native	9.44 (1.78-50.09)	.008
Other, including Asian, Native Hawaiian or other Pacific Islander, and more than one race	0.40 (0.03-5.52)	.50
Ethnicity		
Hispanic or Latino (reference)	1.00	
Not Hispanic or Latino	1.94 (0.43-8.76)	.39
Maternal age at childbirth (years)		
<30 (reference)	1.00	
30-35	2.04 (0.69-6.08)	.20
>35	0.81 (0.28-2.36)	.69
Tobacco used during pregnancy		
No (reference)	1.00	
Yes	13.78 (4.18-45.35)	<.001
Substances used during pregnancy		
None (reference)	1.00	
Marijuana only	1.21 (0.37-3.93)	.75
Other (amphetamine, cocaine, hallucinogens, heroin, methadone)	3.49 (0.82-14.88)	.09
Child sex		
Male (reference)	1.00	
Female	3.64 (1.45-9.12)	.006
Child Behavior Checklist score (points)		
Normal (<65) (reference)	1.00	
Borderline (65-69)	3.32 (1.08-10.19)	.04
Clinical (>69)	1.89 (0.63-5.69)	.26
Clinical location		
San Diego University (reference)	1.00	
Emory University	3.19 (0.62-16.44)	.17
Northern Plains	0.97 (0.17-5.62)	.98
Minnesota	N/A[a]	N/A
University of California, Los Angeles	2.26 (0.38-13.58)	.37

[a] N/A: Not applicable (a low number of participants made stratification impossible).

Table 3. Log-binomial regression analysis of risk of fetal alcohol syndrome associated with a 1-gram increase in average alcoholic content.

Sociodemographic variables	Prevalence ratio (95% CI)	*P* value
Race		
White	1.10 (1.03-1.19)	.009
Black	1.13 (1.04-1.23)	.004
American Indian/Alaska Native	1.10 (1.00-1.21)	.05
Other, including Asian, Native Hawaiian or other Pacific Islander, and more than one race	N/A[a]	N/A
Ethnicity		
Hispanic/Latino	1.11 (1.07-1.16)	<.001
Not Hispanic/Latino	1.24 (0.02-78.08)	.92
Maternal age at childbirth (years)		
<30	1.08 (1.02-1.14)	.01
30-35	1.04 (0.93-1.17)	.49
>35	1.22 (1.07-1.40)	.004
Tobacco used during pregnancy		
No	1.14 (1.04-1.26)	.006
Yes	1.08 (1.04-1.12)	<.001
Substances used during pregnancy		
None	1.12 (1.01-1.18)	<.001
Marijuana only	1.10 (1.02-1.19)	.02
Other (amphetamine, cocaine, hallucinogens, heroin, methadone)	1.40 (0.25-7.89)	.70
Child sex		
Male	1.14 (1.06-1.22)	<.001
Female	1.12 (1.05-1.19)	<.001
Child Behavior Checklist score (points)		
Normal (<65)	1.14 (1.07-1.21)	<.001
Borderline (65-69)	2.01 (1.01-4.22)	.05
Clinical (>69)	1.10 (1.01-1.20)	.04
Clinical location		
San Diego University	2.62 (0.39-17.68)	.32
Emory University	1.17 (1.09-1.26)	<.001
Northern Plains	1.09 (1.01-1.17)	.03
Minnesota	N/A[a]	N/A
University of California, Los Angeles	0.99 (0.33-2.99)	.99

[a] N/A: not applicable (a low number of participants made stratification impossible).

Table 4. Log-binomial regression analysis of fetal alcohol spectrum disorder severity associated with a 1-gram increase in average alcoholic content.

FASD[a] characteristics	Participants, n (%)	PR (95% CI)	*P* value
FASD severity (n=43)			
FAS[b] (most severe)	20 (46.5)	1.11 (1.07-1.15)	<.001
Partial FAS	1 (2.3)	1.78 (0.59-5.43)	.31
Alcohol-related neurodevelopmental disorder	9 (20.9)	1.08 (0.96-1.22)	.21
Alcohol-related birth defects (least severe)	13 (30.2)	1.07 (1.02-1.13)	.008
Number of FASD features (n=291)			
1	209 (71.8)	1.13 (1.04-1.23)	.002
2	56 (19.2)	1.17 (1.06-1.30)	.002
3	21 (7.2)	1.38 (0.59-5.43)	.98
4	5 (1.7)	—[c]	—
Type of FASD features (n=404)			
≥2 Facial anomalies[c]	113 (28)	1.04 (1.02-1.07)	<.001
Pre- or postnatal growth deficiency	92 (22.8)	1.01 (0.98-1.04)	.52
Deficient brain growth	51 (12.6)	1.04 (1.01-1.07)	.01
Neurobehavioral impairment	148 (36.6)	1.00 (0.28-3.53)	.83

[a] FASD: fetal alcohol spectrum disorder.

[b] FAS: fetal alcohol syndrome.

[c] Not available.

[c] Facial anomalies included short palpebral fissures (≤10th percentile), a thin vermilion border on the upper lip (rank 4 or 5 on a racially normed lip/philtrum guide) and a smooth philtrum (rank 4 or 5 on a racially normed lip/philtrum guide).

Discussion

Principal Results

Our results suggest that although the average quantity of alcohol consumed during pregnancy varied by race/ethnicity, the risk of delivering a baby with FAS or FASDs did not vary by more than 3% among White, Black, and AI/AN women when equal doses of ethanol were consumed. There is a consensus that "for unclear reasons, there is an increased risk of FAS in those who belong to African American or Native American ethnicities" [33]. Some researchers have attributed this to genetic polymorphisms that result in an increased risk of FAS among certain races [34,35]. African American race/ethnicity is believed to increase FAS risk due to the ADH1B*3 functional polymorphism—found in approximately 33% of those with African ancestry and almost exclusively in this group, which has a 70-to-80-time higher conversion rate of ethanol to acetaldehyde [36]. However, a previous study attributed to the maternal CIF*3 allele found among African Americans a protective effect against adverse prenatal effects of alcohol [37]. The high prevalence of FAS in certain communities (eg, the Indigenous) is likely due to more intensive screening or other social and environmental risk factors (eg, lead exposure or poor diet) than innate racial characteristics [38].

Our findings also suggest that innate racial characteristics are likely unassociated with the risk of FASDs for any race, especially when AAC is controlled for. Regarding Hispanic/Latino women, our study also found no statistically significant association between ethnicity and risk of FAS when AAC was accounted for, and this was in alignment with a previous study reporting that the pattern of early-pregnancy alcohol consumption in Latinas of childbearing age residing in the United States was similar to that of other racial/ethnic groups [39].

As data on patients with FAS or FASDs and their exposure histories have remained sparse [40], it was difficult to adjust for AAC accurately before the public release of the CIFASD data. While existing nationally representative data sets, such as the National Birth Defects Prevention Study (NBDPS) [41] and Pregnancy Risk Assessment Monitoring System (PRAMS) [42] have some information on women's health-related behaviors during pregnancy, they do not have a detailed assessment of drinking behaviors, such as the exact type of alcoholic beverage consumed (eg, beer, wine, or distilled spirits) [43], which makes it impossible to calculate exact AACs.

Limitations

This study has limitations. Firstly, all data were self-reported; many studies have commented on the underreporting of alcohol consumption during pregnancy by mothers [44,45]. Our calculations of AAC need to be interpreted with caution, as the retrospective manner in which study participants were interviewed made it difficult to collect exact, day-to-day measurements of the quantity, frequency, and type of alcoholic beverage consumed.

As mentioned above, our sample size was also limited and not nationally representative of the United States, which may have resulted in sampling bias owing to the clinical locations where patients were recruited for the CIFASD project. This is likely the reason why the prevalence of pregnant drinkers among certain races and ethnicities in our sample differed from that of previous studies that used nationally representative data sets, such as the NBDPS or PRAMS [42]. Due to the manner in which the CIFASD collected data at certain sites (eg, word of mouth, advertisements in tribal and community health centers, and existing patients at referral clinics), the data on the prevalence of FAS and pregnant drinkers and nondrinkers by race or ethnicity presented in Table 1 should not be interpreted as representative of the general United States or AI/AN population. Among AI/AN people, FASDs are widely recognized as a “major public health priority” [46], since higher patterns of alcohol consumption result in a higher prevalence of FASDs compared to the general population [47].

The small number of FAS-diagnosed individuals (n=2) in the AI/AN sample (n=23) is reflected by the highly imprecise PR estimate for AI/AN individuals (PR 9.44, 95% CI 1.78-50.09). In the United States, AI/AN individuals are only around 2% of the population (for a total of approximately 5.2 million people) [48], and only 14 states (including New Mexico, where the CIFASD collected data for this study) have more than 100,000 AI/AN residents [49]. FAS remains a major concern for the AI/AN community, and previous studies have noted that AI individuals have 3 times the rate of fetal alcohol births [50], possibly because of polymorphisms related to aldehyde dehydrogenase and acetaldehyde dehydrogenase genes found in AI populations [51].

However, until more data are collected for this community, we believe that caution is warranted in ascribing racial disparities in FAS diagnosis to genetic or biological susceptibility [51]. For example, the high prevalence of alcohol use disorders in AI/AN groups has been shown to be driven by primarily nonbiological factors [52].

Furthermore, previous studies have highlighted the importance of considering immigration history and its interaction with race and ethnicity [53]. Trends in child health in the United States, especially with regard to Hispanic- or Asian-origin populations, may be influenced by immigration status and extent of acculturation [53]; for example, relative to United States–born women, it is believed that non–United States–born women are less likely to initiate early prenatal care (PNC) or receive adequate PNC due to lower rates of insurance coverage and patient-provider communication gaps [54]. The CIFASD data set does not include information on maternal immigration or health insurance status, and we encourage future studies to explore the interaction between such factors and FAS onset, as mothers of children who have FAS are more likely to be without PNC, be on Medicaid at childbirth, and potentially have different health-related behaviors, depending on naturalization status [55].

Because of the small sample size of individuals with FAS at different clinical locations, as seen in certain parts of our analyses (ie, the log-binomial analysis of FAS individuals from Minnesota University [n=0] and the subgroup analysis of other races [n=2] for FAS onset associated with a 1-gram increase in AAC), some PRs could not be calculated. In future studies with larger sample sizes, we hope that there will be more information on FAS incidence, so that quantitative analyses with various stratifications of variables are possible.

Conclusions

The current IOM guidelines for FASD diagnosis recommend that racial/ethnic variations in the expression of key features of FASDs be carefully examined, if possible, for the detection of minor physical variations [56]. However, more research that takes into consideration drinking behaviors and cultural drinking norms among each racial/ethnic group in the United States is warranted to determine whether the onset of FAS and related disabilities is biological or behavioral.

Acknowledgments

We thank the National Institute on Alcohol Abuse and Alcoholism Collaborative Initiative on Fetal Alcohol Spectrum Disorder for providing us with the dysmorphology (U24AA014815), neurobehavior (U01AA014834), and management (U01AA014809) data sets used in this project. We thank Leah Wetherill for providing us with her insight and guidance. SHL received a scholarship from Brain Korea 21 FOUR Project funded by National Research Foundation (NRF) of Korea, Yonsei University College of Nursing.

Data Availability

Our agreement with the Collaborative Initiative on Fetal Alcohol Spectrum Disorder for accessing the CIFASD database does not include authorization of access to a third party. Request for database access should be addressed to CIFASD’s data committee [57].

Conflicts of Interest

None declared.

References

1. Popova S, Lange S, Probst C, Shield K, Kraicer-Melamed H, Ferreira-Borges C, et al. Actual and predicted prevalence of alcohol consumption during pregnancy in the WHO African Region. Trop Med Int Health 2016 Oct;21(10):1209-1239 [FREE Full text] [doi: 10.1111/tmi.12755] [Medline: 27429168]
2. Paley B, O'Connor MJ. Intervention for individuals with fetal alcohol spectrum disorders: treatment approaches and case management. Dev Disabil Res Rev 2009;15(3):258-267. [doi: 10.1002/ddrr.67] [Medline: 19731383]
3. Denny L, Coles S, Blitz R. Fetal alcohol syndrome and fetal alcohol spectrum disorders. Am Fam Physician 2017 Oct 15;96(8):515-522 [FREE Full text] [Medline: 29094891]
4. Popova S, Lange S, Probst C, Gmel G, Rehm J. Estimation of national, regional, and global prevalence of alcohol use during pregnancy and fetal alcohol syndrome: a systematic review and meta-analysis. Lancet Glob Health 2017 Mar;5(3):e290-e299 [FREE Full text] [doi: 10.1016/S2214-109X(17)30021-9] [Medline: 28089487]
5. Tenkku LE, Morris DS, Salas J, Xaverius PK. Racial disparities in pregnancy-related drinking reduction. Matern Child Health J 2009 Sep;13(5):604-613. [doi: 10.1007/s10995-008-0409-2] [Medline: 18780169]
6. Ethen MK, Ramadhani TA, Scheuerle AE, Canfield MA, Wyszynski DF, Druschel CM, National Birth Defects Prevention Study. Alcohol consumption by women before and during pregnancy. Matern Child Health J 2009 Mar;13(2):274-285 [FREE Full text] [doi: 10.1007/s10995-008-0328-2] [Medline: 18317893]
7. Perreira KM, Cortes KE. Race/ethnicity and nativity differences in alcohol and tobacco use during pregnancy. Am J Public Health 2006 Sep;96(9):1629-1636. [doi: 10.2105/AJPH.2004.056598] [Medline: 16873756]
8. Elek E, Harris SL, Squire CM, Margolis M, Weber MK, Dang EP, et al. Women's knowledge, views, and experiences regarding alcohol use and pregnancy: Opportunities to improve health messages. Am J Health Educ 2013;44(4):177-190 [FREE Full text] [doi: 10.1080/19325037.2013.768906] [Medline: 28261370]
9. Pruitt SM, Hoyert DL, Anderson KN, Martin J, Waddell L, Duke C, et al. Racial and ethnic disparities in fetal deaths - United States, 2015-2017. MMWR Morb Mortal Wkly Rep 2020 Sep 18;69(37):1277-1282 [FREE Full text] [doi: 10.15585/mmwr.mm6937a1] [Medline: 32941410]
10. Orr ST, James SA, Reiter JP. Unintended pregnancy and prenatal behaviors among urban, black women in Baltimore, Maryland: the Baltimore preterm birth study. Ann Epidemiol 2008 Jul;18(7):545-551. [doi: 10.1016/j.annepidem.2008.03.005] [Medline: 18504137]
11. Russo D, Purohit V, Foudin L, Salin M. Workshop on Alcohol Use and Health Disparities 2002: a call to arms. Alcohol 2004 Jan;32(1):37-43. [doi: 10.1016/j.alcohol.2004.01.003] [Medline: 15066702]
12. May PA, Gossage JP, Kalberg WO, Robinson LK, Buckley D, Manning M, et al. Prevalence and epidemiologic characteristics of FASD from various research methods with an emphasis on recent in-school studies. Dev Disabil Res Rev 2009;15(3):176-192. [doi: 10.1002/ddrr.68] [Medline: 19731384]
13. Hairston DR, Gibbs TA, Wong SS, Jordan A. Clinician bias in diagnosis and treatment. In: Medlock M, Shtasel D, Trinh NH, Williams D, editors. Racism and Psychiatry. Cham, Switzerland: Humana Press; 2019:105-137.
14. Hoyme HE, Hoyme DB, Elliott AJ, Blankenship J, Kalberg WO, Buckley D, et al. A South African mixed race lip/philtrum guide for diagnosis of fetal alcohol spectrum disorders. Am J Med Genet A 2015 Apr;167A(4):752-755 [FREE Full text] [doi: 10.1002/ajmg.a.37023] [Medline: 25711340]
15. Lindinger NM, Jacobson JL, Warton CMR, Malcolm-Smith S, Molteno CD, Dodge NC, et al. Fetal alcohol exposure alters BOLD activation patterns in brain regions mediating the interpretation of facial affect. Alcohol Clin Exp Res 2021 Jan;45(1):140-152 [FREE Full text] [doi: 10.1111/acer.14519] [Medline: 33220071]
16. Geerts L, Meyer R, Nolan H, Nel L, Nel DG, Brink L, et al. Reference standards for facial measurements in early third trimester South African fetuses, and the effect of maternal and fetal characteristics. J Matern Fetal Neonatal Med 2022 Dec;35(25):8434-8442. [doi: 10.1080/14767058.2021.1977794] [Medline: 35042446]
17. Viljoen D, Louw JG, Lombard C, Olivier L. Comparing diagnostic outcomes of children with fetal alcohol syndrome in South Africa with diagnostic outcomes when using the updated Institute of Medicine diagnostic guidelines. Birth Defects Res 2018 Oct 16;110(17):1335-1342. [doi: 10.1002/bdr2.1399] [Medline: 30347134]
18. Mattson SN, Foroud T, Sowell ER, Jones KL, Coles CD, Fagerlund A, CIFASD. Collaborative initiative on fetal alcohol spectrum disorders: methodology of clinical projects. Alcohol 2010;44(7-8):635-641 [FREE Full text] [doi: 10.1016/j.alcohol.2009.08.005] [Medline: 20036488]
19. Bernes GA, Courchesne-Krak NS, Hyland MT, Villodas MT, Coles CD, Kable JA, CIFASD. Development and validation of a postnatal risk score that identifies children with prenatal alcohol exposure. Alcohol Clin Exp Res 2022 Jan;46(1):52-65 [FREE Full text] [doi: 10.1111/acer.14749] [Medline: 34806190]
20. Doyle LR, Coles CD, Kable JA, May PA, Sowell ER, Jones KL, CIFASD. Relation between adaptive function and IQ among youth with histories of heavy prenatal alcohol exposure. Birth Defects Res 2019 Jul 15;111(12):812-821 [FREE Full text] [doi: 10.1002/bdr2.1463] [Medline: 30719847]
21. Suttie M, Wetherill L, Jacobson SW, Jacobson JL, Hoyme HE, Sowell ER, CIFASD. Facial curvature detects and explicates ethnic differences in effects of prenatal alcohol exposure. Alcohol Clin Exp Res 2017 Aug;41(8):1471-1483 [FREE Full text] [doi: 10.1111/acer.13429] [Medline: 28608920]

22. Doyle LR, Glass L, Wozniak JR, Kable JA, Riley EP, Coles CD, CIFASD. Relation between oppositional/conduct behaviors and executive function among youth with histories of heavy prenatal alcohol exposure. Alcohol Clin Exp Res 2019 Jun;43(6):1135-1144 [FREE Full text] [doi: 10.1111/acer.14036] [Medline: 30908651]
23. Mattson SN, Roesch SC, Fagerlund A, Autti-Rämö I, Jones KL, May PA, Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD). Toward a neurobehavioral profile of fetal alcohol spectrum disorders. Alcohol Clin Exp Res 2010 Sep 01;34(9):1640-1650 [FREE Full text] [doi: 10.1111/j.1530-0277.2010.01250.x] [Medline: 20569243]
24. Kable JA, Mehta PK, Rashid F, Coles CD. Path analysis of the impact of prenatal alcohol on adult vascular function. Alcohol Clin Exp Res 2023 Jan;47(1):116-126. [doi: 10.1111/acer.14970] [Medline: 36330744]
25. Aragón AS, Kalberg WO, Buckley D, Barela-Scott LM, Tabachnick BG, May PA. Neuropsychological study of FASD in a sample of American Indian children: processing simple versus complex information. Alcohol Clin Exp Res 2008 Dec;32(12):2136-2148 [FREE Full text] [doi: 10.1111/j.1530-0277.2008.00802.x] [Medline: 18828799]
26. Uban KA, Herting MM, Wozniak JR, Sowell ER, CIFASD. Sex differences in associations between white matter microstructure and gonadal hormones in children and adolescents with prenatal alcohol exposure. Psychoneuroendocrinology 2017 Sep;83:111-121 [FREE Full text] [doi: 10.1016/j.psyneuen.2017.05.019] [Medline: 28609669]
27. Panczakiewicz AL, Glass L, Coles CD, Kable JA, Sowell ER, Wozniak JR, CIFASD. Neurobehavioral deficits consistent across age and sex in youth with prenatal alcohol exposure. Alcohol Clin Exp Res 2016 Sep;40(9):1971-1981 [FREE Full text] [doi: 10.1111/acer.13153] [Medline: 27430360]
28. Mattson SN, Roesch SC, Glass L, Deweese BN, Coles CD, Kable JA, CIFASD. Further development of a neurobehavioral profile of fetal alcohol spectrum disorders. Alcohol Clin Exp Res 2013 Mar;37(3):517-528 [FREE Full text] [doi: 10.1111/j.1530-0277.2012.01952.x] [Medline: 22974253]
29. May PA, Gossage JP, Smith M, Tabachnick BG, Robinson LK, Manning M, et al. Population differences in dysmorphic features among children with fetal alcohol spectrum disorders. J Dev Behav Pediatr 2010 May;31(4):304-316 [FREE Full text] [doi: 10.1097/DBP.0b013e3181dae243] [Medline: 20431397]
30. Hoyme HE, May PA, Kalberg WO, Kodituwakku P, Gossage JP, Trujillo PM, et al. A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 institute of medicine criteria. Pediatrics 2005 Jan;115(1):39-47 [FREE Full text] [doi: 10.1542/peds.2004-0259] [Medline: 15629980]
31. Bordin IA, Rocha MM, Paula CS, Teixeira MCTV, Achenbach TM, Rescorla LA, et al. Child Behavior Checklist (CBCL),Youth Self-Report (YSR) and Teacher's Report Form(TRF): an overview of the development of the original and Brazilian versions. Cad Saude Publica 2013 Jan;29(1):13-28 [FREE Full text] [doi: 10.1590/s0102-311x2013000100004] [Medline: 23370021]
32. Ware AL, O'Brien JW, Crocker N, Deweese BN, Roesch SC, Coles CD, CIFASD. The effects of prenatal alcohol exposure and attention-deficit/hyperactivity disorder on psychopathology and behavior. Alcohol Clin Exp Res 2013 Mar;37(3):507-516 [FREE Full text] [doi: 10.1111/j.1530-0277.2012.01953.x] [Medline: 22974279]
33. Dejong K, Olyaei A, Lo JO. Alcohol use in pregnancy. Clin Obstet Gynecol 2019 Mar;62(1):142-155 [FREE Full text] [doi: 10.1097/GRF.0000000000000414] [Medline: 30575614]
34. Green RF, Stoler JM. Alcohol dehydrogenase 1B genotype and fetal alcohol syndrome: a HuGE minireview. Am J Obstet Gynecol 2007 Jul;197(1):12-25. [doi: 10.1016/j.ajog.2007.02.028] [Medline: 17618743]
35. Warren KR, Li T. Genetic polymorphisms: impact on the risk of fetal alcohol spectrum disorders. Birth Defects Res A Clin Mol Teratol 2005 Apr;73(4):195-203. [doi: 10.1002/bdra.20125] [Medline: 15786496]
36. Edenberg HJ. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. Alcohol Res Health 2007;30(1):5-13 [FREE Full text] [Medline: 17718394]
37. Dodge NC, Jacobson JL, Jacobson SW. Protective effects of the alcohol dehydrogenase-ADH1B*3 allele on attention and behavior problems in adolescents exposed to alcohol during pregnancy. Neurotoxicol Teratol 2014;41:43-50 [FREE Full text] [doi: 10.1016/j.ntt.2013.11.003] [Medline: 24263126]
38. O'Leary CM. Fetal alcohol syndrome: diagnosis, epidemiology, and developmental outcomes. J Paediatr Child Health 2004;40(1-2):2-7. [doi: 10.1111/j.1440-1754.2004.00280.x] [Medline: 14717994]
39. Chambers CD, Hughes S, Meltzer SB, Wahlgren D, Kassem N, Larson S, et al. Alcohol consumption among low-income pregnant Latinas. Alcohol Clin Exp Res 2005 Nov;29(11):2022-2028. [doi: 10.1097/01.alc.0000187160.18672.f9] [Medline: 16340460]
40. O'Leary CM, Nassar N, Kurinczuk JJ, de Klerk N, Geelhoed E, Elliott EJ, et al. Prenatal alcohol exposure and risk of birth defects. Pediatrics 2010 Oct;126(4):e843-e850. [doi: 10.1542/peds.2010-0256] [Medline: 20876169]
41. Waller DK, Shaw GM, Rasmussen SA, Hobbs CA, Canfield MA, Siega-Riz A, National Birth Defects Prevention Study. Prepregnancy obesity as a risk factor for structural birth defects. Arch Pediatr Adolesc Med 2007 Aug;161(8):745-750. [doi: 10.1001/archpedi.161.8.745] [Medline: 17679655]
42. Shulman HB, D'Angelo DV, Harrison L, Smith RA, Warner L. The Pregnancy Risk Assessment Monitoring System (PRAMS): overview of design and methodology. Am J Public Health 2018 Oct;108(10):1305-1313. [doi: 10.2105/AJPH.2018.304563] [Medline: 30138070]

43. Mateja WA, Nelson DB, Kroelinger CD, Ruzek S, Segal J. The association between maternal alcohol use and smoking in early pregnancy and congenital cardiac defects. J Womens Health (Larchmt) 2012 Jan;21(1):26-34. [doi: 10.1089/jwh.2010.2582] [Medline: 21895513]
44. Ernhart CB, Morrow-Tlucak M, Sokol RJ, Martier S. Underreporting of alcohol use in pregnancy. Alcohol Clin Exp Res 1988 Aug;12(4):506-511. [doi: 10.1111/j.1530-0277.1988.tb00233.x] [Medline: 3056071]
45. Morrow-Tlucak M, Ernhart CB, Sokol RJ, Martier S, Ager J. Underreporting of alcohol use in pregnancy: relationship to alcohol problem history. Alcohol Clin Exp Res 1989 Jun;13(3):399-401. [doi: 10.1111/j.1530-0277.1989.tb00343.x] [Medline: 2665555]
46. Gonzales KL, Jacob MM, Mercier A, Heater H, Nall Goes Behind L, Joseph J, et al. An indigenous framework of the cycle of fetal alcohol spectrum disorder risk and prevention across the generations: historical trauma, harm and healing. Ethn Health 2021 Feb;26(2):280-298. [doi: 10.1080/13557858.2018.1495320] [Medline: 29999420]
47. Montag AC, Romero R, Jensen T, Goodblanket A, Admire A, Whitten C, et al. The prevalence of fetal alcohol spectrum disorders in an American Indian community. Int J Environ Res Public Health 2019 Jun 20;16(12):2179 [FREE Full text] [doi: 10.3390/ijerph16122179] [Medline: 31226736]
48. Roubidoux MA, Kaur JS, Rhoades DA. Health disparities in cancer among American Indians and Alaska Natives. Acad Radiol 2022 Jul;29(7):1013-1021. [doi: 10.1016/j.acra.2021.10.011] [Medline: 34802904]
49. Hicks JT, Burnett E, Matanock A, Khalil G, English K, Doman B, et al. Hospitalizations for COVID-19 among American Indian and Alaska Native Adults (≥ 18 years old) - New Mexico, March-September 2020. J Racial Ethn Health Disparities 2023 Feb;10(1):56-63 [FREE Full text] [doi: 10.1007/s40615-021-01196-0] [Medline: 35060084]
50. Spillane NS, Smith GT. A theory of reservation-dwelling American Indian alcohol use risk. Psychol Bull 2007 May;133(3):395-418. [doi: 10.1037/0033-2909.133.3.395] [Medline: 17469984]
51. Cordova F, Trujillo M, Walker R. American Indians/Alaskan Natives and alcohol: biology, nutrition, and positive programs. In: Watson R, Preedy V, Zibadi S, editors. Alcohol, Nutrition, and Health Consequences. Totowa, NJ: Humana Press; 2013:135-142.
52. Gampa V, Bernard K, Oldani MJ. Racialization as a barrier to achieving health equity for Native Americans. AMA J Ethics 2020 Oct 01;22(10):E874-E881 [FREE Full text] [doi: 10.1001/amajethics.2020.874] [Medline: 33103650]
53. Slopen N, Shonkoff JP, Albert MA, Yoshikawa H, Jacobs A, Stoltz R, et al. Racial disparities in child adversity in the U.S.: Interactions with family immigration history and income. Am J Prev Med 2016 Jan;50(1):47-56. [doi: 10.1016/j.amepre.2015.06.013] [Medline: 26342634]
54. Green TL, Bodas MV, Jones HA, Masho SW, Hagiwara N. Disparities in self-reported prenatal counseling: Does immigrant status matter? J Community Health 2018 Oct;43(5):864-873. [doi: 10.1007/s10900-018-0495-z] [Medline: 29516385]
55. Cannon MJ, Dominique Y, O'Leary LA, Sniezek JE, Floyd RL, FASSNet Team. Characteristics and behaviors of mothers who have a child with fetal alcohol syndrome. Neurotoxicol Teratol 2012;34(1):90-95. [doi: 10.1016/j.ntt.2011.09.010] [Medline: 22001355]
56. Calhoun F, Warren K. Fetal alcohol syndrome: historical perspectives. Neurosci Biobehav Rev 2007;31(2):168-171. [doi: 10.1016/j.neubiorev.2006.06.023] [Medline: 17224346]
57. Accessing CIFASD Research Data. Collaborative Initiative on Fetal Alcohol Spectrum Disorders. URL: https://cifasd.org/data-sharing [accessed 2023-04-11]

Abbreviations

AAC: average alcoholic content
AI/AN: American Indian/Alaska Native
CBCL: Child Behavior Checklist
CIFASD: Collaborative Initiative on Fetal Alcohol Spectrum Disorders
FAS: fetal alcohol syndrome
FASD: fetal alcohol spectrum disorder
IOM: Institute of Medicine
NBDPS: National Birth Defects Prevention Study
PNC: prenatal care
PR: prevalence ratio
PRAMS: Pregnancy Risk Assessment Monitoring System
UCLA: University of California, Los Angeles

Edited by A Mavragani, T Sanchez; submitted 05.01.23; peer-reviewed by JY Moon, SH Kim; comments to author 19.01.23; revised version received 15.03.23; accepted 23.03.23; published 21.04.23.

Please cite as:
Oh SS, Kang B, Park J, Kim S, Park EC, Lee SH, Kawachi I
Racial/Ethnic Disparity in Association Between Fetal Alcohol Syndrome and Alcohol Intake During Pregnancy: Multisite Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e45358
URL: https://publichealth.jmir.org/2023/1/e45358
doi:10.2196/45358
PMID:37083819

Original Paper

Associations Between Frailty and the Increased Risk of Adverse Outcomes Among 38,950 UK Biobank Participants With Prediabetes: Prospective Cohort Study

Xingqi Cao[1,2,3], MS; Xueqin Li[1,2,3], BS; Jingyun Zhang[1,2,3], PhD; Xiaoyi Sun[1,2,3], BS; Gan Yang[1,2,3], BS; Yining Zhao[1,2,3], BS; Shujuan Li[4], MD; Emiel O Hoogendijk[5], PhD; Xiaofeng Wang[6,7], MD, PhD; Yimin Zhu[8*], PhD; Heather Allore[9], PhD; Thomas M Gill[9], MD; Zuyun Liu[1,2,3*], PhD

[1] Center for Clinical Big Data and Analytics of the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

[2] Department of Big Data in Health Science School of Public Health, Zhejiang University School of Medicine, Hangzhou, China

[3] The Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province, Zhejiang University School of Medicine, Hangzhou, China

[4] Department of Neurology, Fuwai Hospital, National Center for Cardiovascular Diseases, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

[5] Department of Epidemiology & Data Science, Amsterdam Public Health Research Institute, Amsterdam University Medical Center, Amsterdam, Netherlands

[6] National Clinical Research Center for Aging and Medicine, Huashan Hospital, Shanghai, China

[7] Human Phenome Institute, Fudan University, Shanghai, China

[8] Department of Epidemiology & Biostatistics, School of Public Health, Zhejiang University School of Medicine, Hangzhou, China

[9] Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States

[*] these authors contributed equally

Corresponding Author:
Zuyun Liu, PhD
Center for Clinical Big Data and Analytics of the Second Affiliated Hospital
Zhejiang University School of Medicine
866 Yuhangtang Rd
Hangzhou, 310058
China
Phone: 86 057187077127
Email: zuyunliu@zju.edu.cn

Abstract

Background: Compared with adults with normal glucose metabolism, those with prediabetes tend to be frail. However, it remains poorly understood whether frailty could identify adults who are most at risk of adverse outcomes related to prediabetes.

Objective: We aimed to systematically evaluate the associations between frailty, a simple health indicator, and risks of multiple adverse outcomes including incident type 2 diabetes mellitus (T2DM), diabetes-related microvascular disease, cardiovascular disease (CVD), chronic kidney disease (CKD), eye disease, dementia, depression, and all-cause mortality in late life among middle-aged adults with prediabetes.

Methods: We evaluated 38,950 adults aged 40 years to 64 years with prediabetes using the baseline survey from the UK Biobank. Frailty was assessed using the frailty phenotype (FP; range 0-5), and participants were grouped into nonfrail (FP=0), prefrail ($1 \leq FP \leq 2$), and frail ($FP \geq 3$). Multiple adverse outcomes (ie, T2DM, diabetes-related microvascular disease, CVD, CKD, eye disease, dementia, depression, and all-cause mortality) were ascertained during a median follow-up of 12 years. Cox proportional hazards regression models were used to estimate the associations. Several sensitivity analyses were performed to test the robustness of the results.

Results: At baseline, 49.1% (19,122/38,950) and 5.9% (2289/38,950) of adults with prediabetes were identified as prefrail and frail, respectively. Both prefrailty and frailty were associated with higher risks of multiple adverse outcomes in adults with prediabetes (*P* for trend <.001). For instance, compared with their nonfrail counterparts, frail participants with prediabetes had a significantly higher risk (*P*<.001) of T2DM (hazard ratio [HR]=1.73, 95% CI 1.55-1.92), diabetes-related microvascular disease (HR=1.89, 95% CI 1.64-2.18), CVD (HR=1.66, 95% CI 1.44-1.91), CKD (HR=1.76, 95% CI 1.45-2.13), eye disease (HR=1.31, 95% CI 1.14-1.51), dementia (HR=2.03, 95% CI 1.33-3.09), depression (HR=3.01, 95% CI 2.47-3.67), and all-cause mortality (HR=1.81, 95% CI 1.51-2.16) in the multivariable-adjusted models. Furthermore, with each 1-point increase in FP score, the risk of these adverse outcomes increased by 10% to 42%. Robust results were generally observed in sensitivity analyses.

Conclusions: In UK Biobank participants with prediabetes, both prefrailty and frailty are significantly associated with higher risks of multiple adverse outcomes, including T2DM, diabetes-related diseases, and all-cause mortality. Our findings suggest that frailty assessment should be incorporated into routine care for middle-aged adults with prediabetes, to improve the allocation of health care resources and reduce diabetes-related burden.

(JMIR Public Health Surveill 2023;9:e45502) doi:10.2196/45502

KEYWORDS

frailty; adverse outcomes; diabetes; prediabetes; prospective study

Introduction

In 2021, the International Diabetes Federation estimated that there were more than 500 million adults with prediabetes among those aged 20 years to 79 years worldwide [1]. As an intermediate hyperglycemia state, prediabetes increases the risk of diabetes [2] and diabetes-related complications (eg, cardiovascular disease [CVD], chronic kidney disease [CKD], and dementia) [3]; the latter contributes to a large proportion of diabetes-related burden [4,5]. The latest guidelines from the American Diabetes Association (ADA) recommend annual diabetes screening for adults with prediabetes [6]. However, this is challenged by emerging evidence showing the very low rates of diabetes progression among older adults with prediabetes [7]. Conversely, middle-aged adults (ie, <65 years) with prediabetes should be monitored for adverse outcomes, which is of high value and appropriate [8].

Prediabetes is highly heterogeneous, impeding the application of a one-size-fits-all health management strategy. Recently, a simple health aging indicator—frailty—has been demonstrated to be able to predict the risk of adverse outcomes (eg, CVD and mortality) [9-12] even in the younger population [13]. Frailty is defined as a state of decreased reserve and resistance to stressors, characterized by functional decline in multiple systems [9]. Frailty and disorders of glucose metabolism share common physiological mechanisms, such as insulin resistance [14,15] and chronic inflammation [15,16]. Frailty has been found to be an important risk factor for disability [17], fracture [18], CVD [19,20], hospitalization [20], intensive care unit admission [20], and mortality [20,21] among adults with diabetes. A few studies have shown that frailty incidence is slightly higher in older adults with prediabetes compared with those with normal glucose metabolism [22]. Only 1 prospective study recently reported that frailty was positively associated with the progression of prediabetes to type 2 diabetes mellitus (T2DM), as well as higher risks of CVD and all-cause mortality, in middle-aged and older adults with prediabetes [23]. However, whether these positive associations remain in those aged less than 65 years is not yet clear. In addition, impaired glucose metabolism is also associated with higher risks of CKD [3], eye disease (eg, cataract) [24], dementia [3], and depression [25]. However, relatively little is known about whether frailty could identify middle-aged adults with prediabetes who are most at risk of these adverse outcomes.

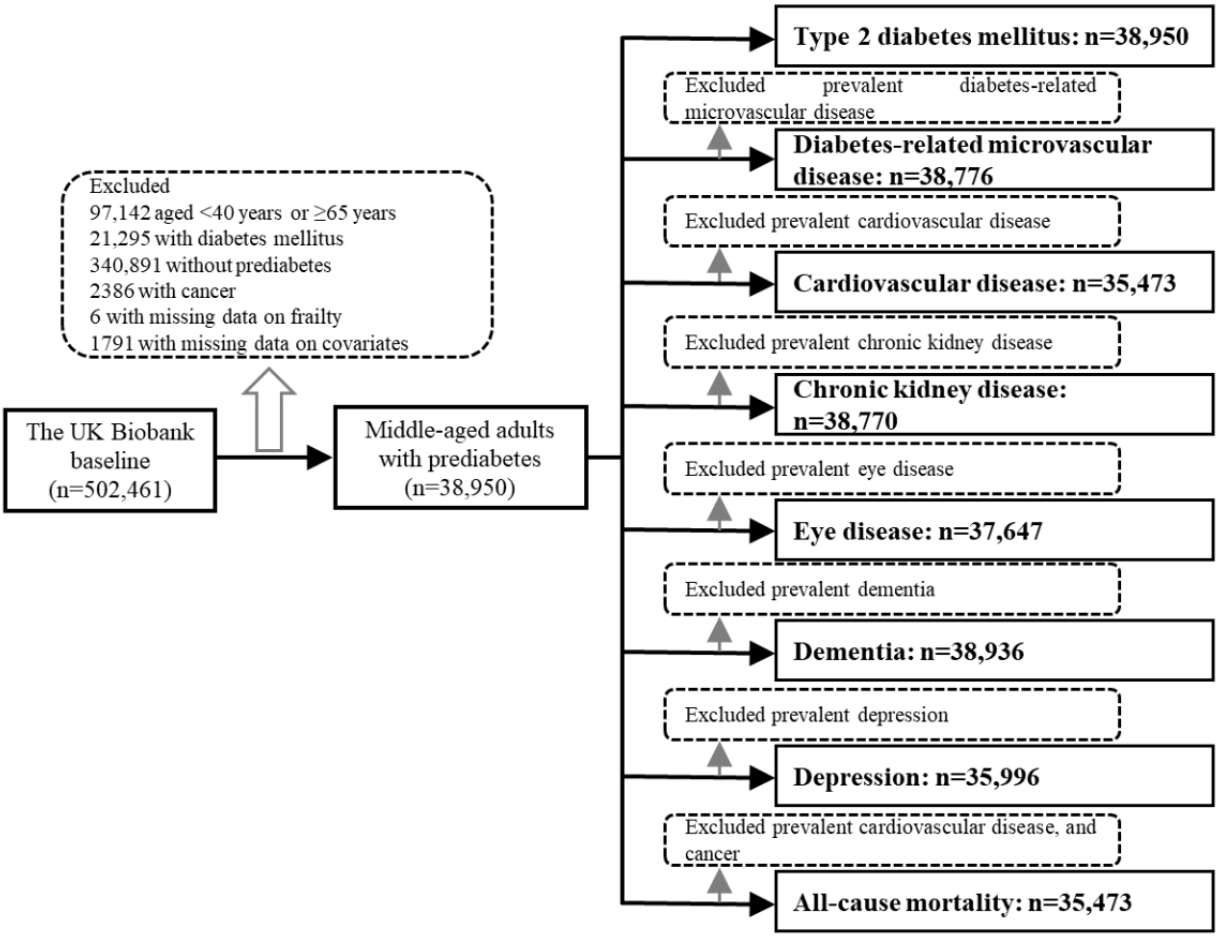
Therefore, we performed a prospective cohort study among 38,950 middle-aged adults with prediabetes from the UK Biobank (UKB). Using a widely validated frailty measurement—frailty phenotype (FP) [9]—the objective of this study was to systematically evaluate the associations of frailty with the risk of multiple adverse outcomes, including incident T2DM, diabetes-related microvascular disease, CVD, CKD, eye disease, dementia, depression, and all-cause mortality.

Methods

Study Participants

The UKB is a large-scale health research study with a long-term follow-up that began in 2006 to 2010 with the recruitment of approximately 500,000 adults in the United Kingdom [26]. Adults in the UKB were recruited through 22 assessment centers across England, Scotland, and Wales. Data were collected through a touch screen questionnaire and verbal interviews (eg, demographic, health, lifestyle variables), physical measures (eg, handgrip strength), and biological sample collection (eg, blood). Since recruitment, all adults have given consent for the UKB to follow up to determine the incidence of health outcomes through links to health-related records (eg, hospital inpatient episodes and death registrations), and only about 0.3% of the adults have been lost to follow-up because they left the United Kingdom or withdrew consent for future linkage. The protocol of the UKB is available online [27]. At baseline, there were 405,319 middle-aged adults (age: 40-64 years), of whom 43,133 had prediabetes. Prediabetes was defined by a nonfasting glycated hemoglobin (HbA_{1c}) level of 5.7% to 6.4% (39-47 mmol/mol) following the ADA criteria [6]. After the exclusion of adults with prevalent cancer (n=2386) or with missing data on frailty (n=6) and covariates (eg, ethnicity, educational level; n=1791), 38,950 middle-aged adults with prediabetes were included in the final analytic samples. Additionally, because the number of prevalent cases for each outcome varied, we assembled different analytic samples for each outcome (see details in Figure 1).

Figure 1. Flow chart of the sample for analyses.



Ethical Considerations

The UKB was approved by the North West Multi-Centre Research Ethics Committee (11/NW/0382). Written informed consent from all participants was obtained. The data used in this study were anonymized and de-identified for privacy and confidentiality protection.

Outcomes

In this study, the outcomes included T2DM, diabetes-related microvascular disease (including retinopathy, neuropathy, and nephropathy), diabetes-related macrovascular disease (ie, CVD including ischemic heart disease and stroke), CKD, eye disease (including cataract and glaucoma), dementia, depression, and all-cause mortality.

We defined prevalent and incident T2DM using a UKB algorithm that combined self-reported medical history and medication information (for the ascertainment of prevalent cases only), as well as linked hospital admissions records (Table S1 in Multimedia Appendix 1). In addition, according to the ADA criteria [6], undiagnosed prevalent T2DM cases were identified using random glucose (≥11.1 mmol/L) or HbA$_{1c}$ (≥6.5% [48 mmol/mol]) levels. We ascertained prevalent and incident cases of diabetes-related microvascular disease through linked hospital admissions records using the International Statistical Classification of Diseases and Related Health Problems, 9th version (ICD-9) and 10th version (ICD-10; Table S1 in Multimedia Appendix 1). We ascertained prevalent and incident cases of CVD, CKD, eye disease, dementia, and depression using self-reported medical history (for the ascertainment of prevalent cases only) and linked hospital admissions records using ICD-9 and ICD-10 (Table S1 in Multimedia Appendix 1). We ascertained death through linkage to national death registries. For analyses of each outcome, the time to event was calculated from the baseline (ie, the years 2006-2010) to the occurrence of the specific disease outcome, death, loss to follow-up, or end of follow-up (the year 2021), whichever came first. For instance, for analysis of incident T2DM, time to event was calculated from the baseline to the occurrence of T2DM, death, loss to follow-up, or end of follow-up, whichever came first.

Frailty Measurement

We used FP, a widely used physical frailty measurement proposed by Fried et al [9]. FP was evaluated using 5 criteria (unintentional weight loss, exhaustion, weakness, slow gait speed, and low physical activity) and was used previously in the UKB [28]. Of the 5 criteria, weakness was assessed using objectively measured handgrip strength; the other 4 criteria were assessed using a self-reported questionnaire (see details in Table 1). The FP score ranged from 0 to 5, with a higher score indicating greater frailty. Participants were categorized into nonfrail (FP score=0), prefrail (FP score≥1 and ≤2), and frail (FP score≥3), as done in previous studies [9,28].

Table 1. The 5 criteria for the frailty phenotype in the UK Biobank.

Number	Criteria description	Categories
1	Unintentional weight loss: Participants were asked “Compared with one year ago, has your weight changed?”	1: “Yes, loss weight”; 0: Others
2	Exhaustion: Participants were asked “Over the past 2 weeks, how often have you felt tired or had little energy?”	1: “More than half the days or nearly every day”; 0: Others
3	Weakness: Weakness was measured using grip strength with a Jamar J00105 hydraulic hand dynamometer (Lafayette Instrument). Participants were asked to complete a grip assessment for both hands once. The maximal value of the right and left hands was used.	1: (1) Men: ≤29 kg for BMI ≤24 kg/m^2; ≤30 kg for BMI 24.1-26 kg/m^2; ≤30 kg for BMI 26.1-28 kg/m^2; or ≤32 kg for BMI >28 kg/m^2; (2) Women: ≤17 kg for BMI ≤23 kg/m^2; ≤17.3 kg for BMI 23.1-26 kg/m^2; ≤18 kg for BMI 26.1-29 kg/m^2; or ≤21 kg for BMI >29 kg/m^2; 0: Others
4	Slow gait speed: Participants were asked “How would you describe your usual walking pace?”	1: “Slow pace”; 0: Others
5	Low physical activity: Participants were asked “In the last 4 weeks, did you spend any time doing light DIY[a] activity, heavy DIY activity, or strenuous sports?”	1: “None or light activity with a frequency of once per week or less”; 0: Others

[a] DIY: do it yourself.

Covariates

Baseline data on age, sex (female or male), ethnicity (White, mixed race, South Asian, Black, Chinese, or other background), educational level (high, intermediate, or low), occupational status (working, retired, or other), alcohol consumption (never or special occasions only, 1 to 3 times per month, 1 to 4 times per week, or daily or almost daily), smoking status (never, previous smoker, or current smoker), healthy diet (yes or no), and family history of disease (including diabetes, CVD, dementia, and depression) were collected through a questionnaire interview. The Townsend deprivation index (TDI) was calculated based on areas before participants were recruited in the UKB. BMI was calculated as measured weight/height^2 (kg/m^2).

Statistical Analyses

Baseline characteristics of the complete analyzed sample and by frailty status are presented as median (IQRs) and number (percentage) for continuous variables and categorical variables, respectively. Kruskal-Wallis tests and chi-square tests were used to compare the differences in characteristics by frailty status.

To evaluate the associations between frailty status (nonfrail, prefrail, and frail) and adverse outcomes, Cox proportional hazards regression models were performed. The Schoenfeld residuals test was used to verify the proportional hazard assumption, and no significant violation was found. We calculated hazard ratios (HRs) and corresponding 95% CIs using 2 models. Model 1 was adjusted for age and sex. Model 2 was further adjusted for ethnicity, educational level, occupational status, TDI, alcohol consumption, smoking status, healthy diet, BMI, and family history of disease based on Model 1. Additionally, we calculated HRs (95% CIs) for adverse outcomes per 1-point increase in FP score.

Several sensitivity analyses were conducted to confirm the robustness of the results. First, we compared the characteristics of included and excluded study participants. Second, to minimize the influence of reverse causality, we repeated the main analyses after excluding those without 2 years of follow-up. Third, to reduce the influence of poor health on frailty status, we repeated the main analyses after excluding participants with poor self-rated health status at baseline. Fourth, to account for the influence of missing data on results, we performed multiple imputations by chained equations [29] for missing values and repeated the primary analyses. Finally, we validated the associations between frailty and adverse outcomes among adults with T2DM. For adults with T2DM, HbA$_{1c}$ level (≥7.0% [≥53 mmol/mol] or <7.0% [<53 mmol/mol]), diabetes medication use (oral antidiabetes drug only, insulin, or neither), diabetes duration (in years), and prevalent diabetes-related microvascular disease (except for incident diabetes-related microvascular disease) were also included in Model 2.

We used SAS version 9.4 (SAS Institute) and R version 3.6.3 (2020-02-29) to conduct all statistical analyses. To account for multiple testing, we used Bonferroni correction in all analyses (P<.006).

Results

Baseline Characteristics

Among the 38,950 participants with prediabetes, the median age was 58.6 (IQR 53.1-62.0) years, and the majority were women (21,155/38,950, 54.3%) and White (34,705/38,950, 89.1%; Table 2). The prevalences of prefrailty and frailty were 49.1% (19,122/38,950) and 5.9% (2289/38,950), respectively. Prefrail and frail adults were more likely to be women, have a lower educational level, and have a higher level of TDI and BMI, compared with the nonfrail adults. Table 2 shows the detailed baseline characteristics by frailty status.

Table 2. Baseline characteristics of study participants with prediabetes by frailty status.

Variables	Total (n=38,950)	Nonfrail (n=17,539)	Prefrail (n=19,122)	Frail (n=2289)	*P* value[a]
Age (years), median (IQR)	58.6 (53.1 to 62.0)	59.0 (53.7 to 62.1)	58.3 (52.7 to 61.8)	58.3 (52.9 to 61.7)	<.001
Gender, n (%)					<.001
Female	21,155 (54.3)	8928 (50.9)	10,771 (56.3)	1456 (63.6)	
Male	17,795 (45.7)	8611 (49.1)	8351 (43.7)	833 (36.4)	
Ethnicity, n (%)					<.001
White	34,705 (89.1)	16,075 (91.7)	16,719 (87.4)	1911 (83.5)	
Mixed	339 (0.9)	137 (0.8)	180 (0.9)	22 (1.0)	
South Asian	1558 (4.0)	449 (2.6)	943 (4.9)	166 (7.3)	
Black	1474 (3.8)	563 (3.2)	796 (4.2)	115 (5.0)	
Chinese	261 (0.7)	104 (0.6)	139 (0.7)	18 (0.8)	
Other background	613 (1.6)	211 (1.2)	345 (1.8)	57 (2.5)	
Educational level[b], n (%)					<.001
High	11,198 (28.7)	5647 (32.2)	5156 (27.0)	395 (17.3)	
Intermediate	12,464 (32.0)	5728 (32.7)	6165 (32.2)	571 (24.9)	
Low	15,288 (39.3)	6164 (35.1)	7801 (40.8)	1323 (57.8)	
Occupational status, n (%)					<.001
Working	23,793 (61.1)	11,059 (63.1)	11,892 (62.2)	842 (36.8)	
Retired	10,407 (26.7)	5095 (29.0)	4710 (24.6)	602 (26.3)	
Other	4750 (12.2)	1385 (7.9)	2520 (13.2)	845 (36.9)	
Townsend deprivation index, median (IQR)	–1.7 (–3.5 to 1.2)	–2.2 (–3.7 to 0.3)	–1.4 (–3.2 to 1.6)	0.5 (–2.3 to 3.6)	<.001
BMI (kg/m^2), median (IQR)	28.5 (25.4 to 32.1)	27.5 (24.8 to 30.8)	29.2 (25.9 to 32.9)	31.6 (27.8 to 36.4)	<.001
Smoking status, n (%)					<.001
Never	19,301 (49.6)	8963 (51.1)	9366 (49.0)	972 (42.5)	
Previous	12,788 (32.8)	5929 (33.8)	6137 (32.1)	722 (31.5)	
Current	6861 (17.6)	2647 (15.1)	3619 (18.9)	595 (26.0)	
Alcohol consumption, n (%)					<.001
Never or special occasions only	9939 (25.5)	3308 (18.9)	5551 (29.0)	1080 (47.2)	
1 to 3 times per month	4919 (12.6)	2045 (11.7)	2587 (13.5)	287 (12.5)	
1 to 4 times per week	17,545 (45.0)	8674 (49.5)	8176 (42.8)	695 (30.4)	
Daily or almost daily	6547 (16.8)	3512 (20.0)	2808 (14.7)	227 (9.9)	
Healthy diet, n (%)					<.001
No	9146 (23.5)	3444 (19.6)	4930 (25.8)	772 (33.7)	
Yes	29,804 (76.5)	14,095 (80.4)	14,192 (74.2)	1517 (66.3)	
Glycated hemoglobin (mmol/mol), median (IQR)	40.4 (39.6 to 42.0)	40.3 (39.5 to 41.6)	40.5 (39.6 to 42.1)	40.9 (39.8 to 42.6)	<.001
Prevalent diseases, n (%)					
Cardiovascular disease	3477 (8.9)	1157 (6.6)	1835 (9.6)	485 (21.2)	<.001
Chronic kidney disease	180 (0.5)	55 (0.3)	88 (0.5)	37 (1.6)	<.001
Eye disease	1303 (3.3)	520 (3.0)	646 (3.4)	137 (6.0)	<.001
Dementia	14 (0.0)	4 (0.0)	7 (0.0)	3 (0.1)	.37

Variables	Total (n=38,950)	Nonfrail (n=17,539)	Prefrail (n=19,122)	Frail (n=2289)	*P* value[a]
Depression	2954 (7.6)	812 (4.6)	1636 (8.6)	506 (22.1)	<.001
Family history, n (%)					
Diabetes mellitus	11,197 (28.7)	4716 (26.9)	5720 (29.9)	761 (33.2)	<.001
Cardiovascular disease	23,633 (60.7)	10,448 (59.6)	11,711 (61.2)	1474 (64.4)	<.001
Dementia	4733 (12.2)	2164 (12.3)	2283 (11.9)	286 (12.5)	.44
Depression	5146 (13.2)	2087 (11.9)	2637 (13.8)	422 (18.4)	<.001

[a] Generated using chi-square and Kruskal-Wallis tests for categorical and continuous variables, respectively.

[b] Educational level was classified as high (college or university degree), intermediate (A/AS levels or equivalent, O levels/General Certificate of Secondary Education levels or equivalent), and low (none of the above).

Frailty and Risks of Adverse Outcomes in Middle-aged Adults With Prediabetes

During a median follow-up of 12 years, there were 5289 incident T2DM cases, 2657 incident diabetes-related microvascular disease cases, 3234 incident CVD cases, 1439 incident CKD cases, 3525 incident eye disease cases, 325 incident dementia cases, 1265 incident depression cases, and 2016 deaths. We found that frail participants developed more adverse outcomes than did their prefrail and nonfrail counterparts over the 12-year follow-up (Figure 2).

Table 3 shows the associations between frailty and the risks of multiple adverse outcomes in middle-aged adults with prediabetes. In the age- and sex-adjusted model, both prefrailty and frailty were associated with higher risks of all adverse outcomes (all *P* for trend <.001). After further adjusting for additional covariates, these associations remained statistically significant. When comparing prefrail participants with their nonfrail counterparts, the multivariable-adjusted HRs were 1.35 (95% CI 1.27-1.43) for T2DM, 1.29 (95% CI 1.18-1.40) for diabetes-related microvascular disease, 1.17 (95% CI 1.08-1.26) for CVD, 1.22 (95% CI 1.09-1.37) for CKD, 1.12 (95% CI 1.04-1.20) for eye disease, 1.57 (95% CI 1.23-2.01) for dementia, 1.48 (95% CI 1.30-1.68) for depression, and 1.25 (95% CI 1.14-1.38) for all-cause mortality. For frail participants, the multivariable-adjusted HRs were 1.73 (95% CI 1.55-1.92) for T2DM, 1.89 (95% CI 1.64-2.18) for diabetes-related microvascular disease, 1.66 (95% CI 1.44-1.91) for CVD, 1.76 (95% CI 1.45-2.13) for CKD, 1.31 (95% CI 1.14-1.51) for eye disease, 2.03 (95% CI 1.33-3.09) for dementia, 3.01 (95% CI 2.47-3.67) for depression, and 1.81 (95% CI 1.51-2.16) for all-cause mortality, compared with their nonfrail counterparts. Additionally, with each 1-point increase in FP score, the incidence risks of these adverse outcomes significantly increased by 10% to 42% (Model 2).

Figure 2. Age-adjusted incidence of adverse outcomes among UKB participants with prediabetes during 12 years of follow-up. UKB: UK Biobank.

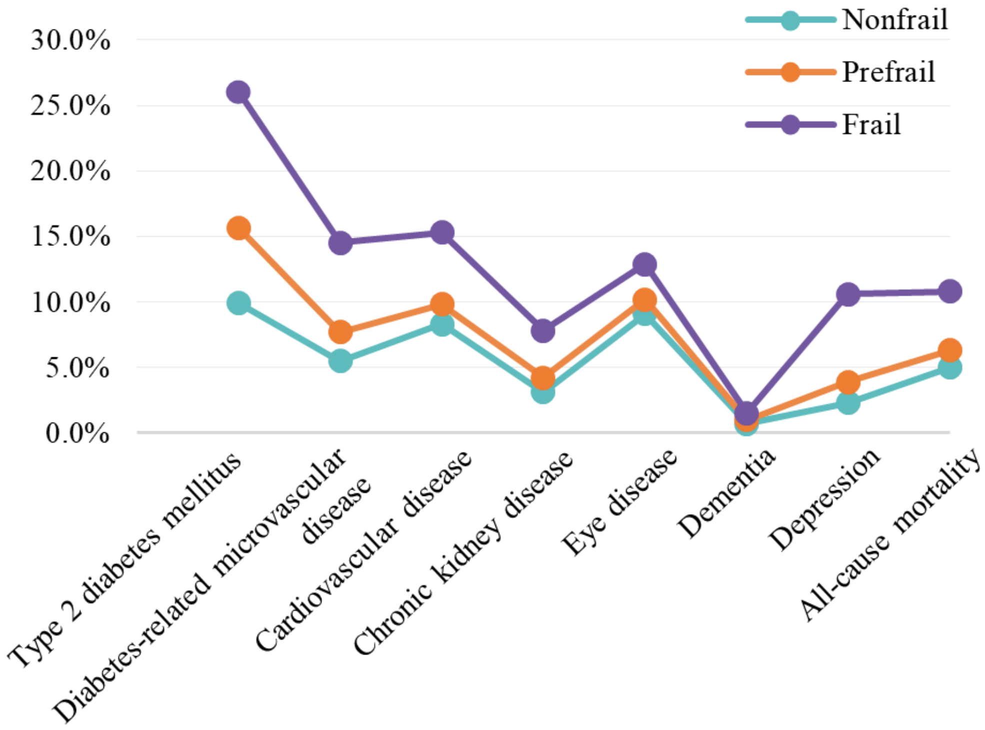


Table 3. Associations between frailty and adverse health outcomes among middle-aged adults with prediabetes.

Outcomes	Frailty status			*P* value for trend[a]	Hazard ratio (HR) per 1-point increase
	Nonfrail	Prefrail	Frail		
Type 2 diabetes mellitus (n=38,950)					
Number of events/person-years	1724/207,929	2965/218,522	600/23,890	__[b]	—
Model 1[c], HR (95% CI)	Reference	1.70 (1.61-1.81)	3.37 (3.07-3.71)	<.001	1.45 (1.42-1.49)
Model 2[d], HR (95% CI)	Reference	1.35 (1.27-1.43)	1.73 (1.55-1.92)	<.001	1.19 (1.16-1.23)
Diabetes-related microvascular disease (n=38,776)					
Number of events/person-years	926/212,240	1417/226,956	314/25,445	—	—
Model 1[c], HR (95% CI)	Reference	1.54 (1.42-1.67)	3.23 (2.84-3.68)	<.001	1.45 (1.40-1.50)
Model 2[d], HR (95% CI)	Reference	1.29 (1.18-1.40)	1.89 (1.64-2.18)	<.001	1.24 (1.19-1.29)
Cardiovascular disease (n=35,473)					
Number of events/person-years	1314/195,009	1651/201,855	269/20,099	—	—
Model 1[c], HR (95% CI)	Reference	1.31 (1.22-1.41)	2.39 (2.09-2.72)	<.001	1.29 (1.25-1.34)
Model 2[d], HR (95% CI)	Reference	1.17 (1.08-1.26)	1.66 (1.44-1.91)	<.001	1.16 (1.12-1.21)
Chronic kidney disease (n=38,770)					
Number of events/person-years	513/213,304	758/228,894	168/25,929	—	—
Model 1[c], HR (95% CI)	Reference	1.47 (1.31-1.64)	3.01 (2.53-3.58)	<.001	1.43 (1.36-1.50)
Model 2[d], HR (95% CI)	Reference	1.22 (1.09-1.37)	1.76 (1.45-2.13)	<.001	1.23 (1.16-1.30)
Eye disease (n=37,647)					
Number of events/person-years	1470/202,556	1792/216,687	263/24,132	—	—
Model 1[c], HR (95% CI)	Reference	1.20 (1.12-1.29)	1.62 (1.42-1.85)	<.001	1.17 (1.13-1.21)
Model 2[d], HR (95% CI)	Reference	1.12 (1.04-1.20)	1.31 (1.14-1.51)	<.001	1.10 (1.06-1.14)
Dementia (n=38,936)					
Number of events/person-years	111/215,549	181/232,270	33/26,881	—	—
Model 1[c], HR (95% CI)	Reference	1.69 (1.34-2.15)	2.87 (1.94-4.23)	<.001	1.41 (1.28-1.56)
Model 2[d], HR (95% CI)	Reference	1.57 (1.23-2.01)	2.03 (1.33-3.09)	<.001	1.29 (1.16-1.44)
Depression (n=35,996)					
Number of events/person-years	387/204,125	687/209,605	191/19,994	—	—
Model 1[c], HR (95% CI)	Reference	1.71 (1.51-1.94)	4.97 (4.18-5.92)	<.001	1.63 (1.55-1.71)
Model 2[d], HR (95% CI)	Reference	1.48 (1.30-1.68)	3.01 (2.47-3.67)	<.001	1.42 (1.34-1.50)
All-cause mortality (n= 35,473)					
Number of events/person-years	783/195,777	1047/204,710	186/20,940	—	—
Model 1[c], HR (95% CI)	Reference	1.39 (1.27-1.53)	2.65 (2.26-3.12)	<.001	1.35 (1.29-1.41)
Model 2[d], HR (95% CI)	Reference	1.25 (1.14-1.38)	1.81 (1.51-2.16)	<.001	1.21 (1.16-1.27)

[a] Calculated to test linear trend using frailty status (3 categories) as a continuous variable.

[b] Not applicable.

[c] Model 1 was adjusted for age and sex.

[d] Model 2 was further adjusted for ethnicity, educational level, occupational status, Townsend deprivation index, alcohol consumption, smoking status, healthy diet, BMI, and family history of disease based on Model 1.

Sensitivity Analyses

The differences in characteristics between included and excluded participants were observed. Those who were excluded were more likely to be older, women, non-White, and frail (Table S2 in Multimedia Appendix 1). Robust results were generally observed when excluding the participants with less than 2 years of follow-up (Table S3 in Multimedia Appendix 1), excluding the participants with poor self-rated health status at baseline (Table S4 in Multimedia Appendix 1), or imputing missing data on frailty and covariates (Table S5 in Multimedia Appendix 1). In addition, we confirmed that frailty was positively associated with the risks of diabetes-related microvascular disease, CVD, CKD, eye diseases, dementia, depression, and all-cause mortality in middle-aged adults with T2DM, and these associations were independent of factors related to diabetes severity at baseline (Table S6 in Multimedia Appendix 1).

Discussion

Principal Findings

In a large sample of UKB participants with prediabetes, we, for the first time, demonstrated that both prefrailty and frailty were associated with higher risks of multiple adverse outcomes, including T2DM, diabetes-related microvascular disease, CVD, CKD, eye disease, dementia, depression, and all-cause mortality. Our findings support the heterogeneity of prediabetes in middle-aged adulthood and suggest that assessing frailty status among middle-aged adults with prediabetes may help to identify those who were most at risk of subsequent adverse outcomes.

We observed a nearly twice higher prevalence of frailty among middle-aged adults with prediabetes (ie, 5.9%) in this study than that in general adults (ie, 3.3%) from the UKB as well [28]. Similarly, the prevalence of frailty among older adults with diabetes [30] is almost twice as high as that in those without diabetes (20.1% vs 12%) [31]. It seems that adults with glucose metabolism disorders are experiencing an accelerated aging process [32]. Multiple age-related metabolic disturbances are present in adults with prediabetes, including chronic inflammation, hyperglycemia, insulin resistance, and β-cell dysfunction [2,16], creating a pathophysiological environment that contributes to frailty. Given the sharp increase in frailty after the age of 65 years [33], our findings suggest that there is a need for early identification of frailty, an aging indicator, in this middle-aged population with prediabetes.

To the best of our knowledge, this study provided new evidence on the associations between frailty and higher risks of a series of adverse outcomes in middle-aged adults with prediabetes. A few studies on the relationship between frailty and adverse outcomes included middle-aged adults with diabetes as part of the study sample [19,20,34]. One prospective study of 998 African Americans aged 49 years to 65 years has shown that frail adults with diabetes had an increased risk of mortality [21]. Except for this study, only 1 study conducted in middle-aged and older adults with prediabetes found that frailty was associated with the progression of prediabetes to diabetes, as well as higher risks of CVD and all-cause mortality [23]. This large prospective study (n=38,950) showed that frailty was positively associated with higher risks of more outcomes including CKD, eye disease, and dementia in middle-aged adults with prediabetes.

This study draws attention to the accelerated aging process in adults with prediabetes, which may lead to rapid diabetes progression and contribute to the development of diabetes-related complications [32]. Nutritional and pharmacological anti-aging interventions have been revealed to help mitigate or reverse the accelerated aging process [35]. A recent review suggested that the most effective and easiest intervention strategy targeting frailty is to combine strength exercise and protein supplements in primary care [36]. Thus, our findings implicate that frailty assessment might help primary care providers identify the subpopulation at higher risk of adverse outcomes even in middle-aged adults with prediabetes in communities. It is worth noting that the application of technological solutions in assessing frailty is constantly expanding [37,38]. The major types of technologies include information and telecommunications technology–based platforms, smartphones, remote monitoring, and wearable sensors and devices [39]. For example, a frailty prediction model based on a points system and integrated into a mobile app for Android phones has been developed in the clinical setting, enabling professionals to identify frailty using clinical information and further improve decision-making [40]. With the aid of these technological tools, frailty screening becomes more convenient and flexible. Next, early preventive and interventive programs targeting frailty in adults with prediabetes are urgently needed. On the one hand, they may directly help reduce the occurrence of T2DM; on the other hand, they may indirectly help reduce diabetes-related burden. Meanwhile, pharmacologic interventions or other aggressive approaches to diabetes prevention are also encouraged [41,42]. Before formal implementation, considerably more research on the effectiveness and cost-effectiveness of interventional programs in this population is required.

Strengths and Limitations

The major strengths of this study were the large sample of middle-aged adults with prediabetes, the long follow-up time, rich phenotype data, and linked hospital admissions records, enabling us to systematically evaluate the prospective associations between frailty and multiple adverse outcomes. There were several potential limitations. First, the UKB was not representative of the sampling population, and the majority of included adults were White. Also, there were differences in baseline characteristics between included and excluded participants. Thus, selection bias existed in this study, and the results may not be generalizable to populations from other countries. Second, transitions in frailty status may occur over time [43], and evidence has suggested that transitions in frailty status were associated with adverse outcomes [44]. However, repeated measurements of frailty were lacking; thus, we were unable to estimate the influence of frailty transitions on the subsequent adverse outcomes in this study. Future longitudinal studies incorporating data on frailty transition are needed. Third, multiple outcomes were considered in this study, and thus, type Ⅰ errors inevitably increased. To reduce the possibility of chance findings, we used Bonferroni correction. Finally, because of the observational study design, we could not draw a causal inference.

Conclusion

In this prospective cohort study of middle-aged UKB participants with prediabetes, both prefrailty and frailty were significantly associated with increased risks of multiple adverse outcomes, including T2DM, diabetes-related microvascular disease, CVD, CKD, eye disease, dementia, depression, and all-cause mortality. The findings underscore the importance of frailty assessment in routine care for middle-aged adults with prediabetes. Detecting frailty at an early stage (ie, accelerated aging) and implementing timely targeted interventions may help to improve the allocation of health care resources and to reduce diabetes-related burden.

Acknowledgments

This research was conducted using the UK Biobank resource under application number 61856. We wish to acknowledge the UK Biobank participants who formed the sample that made the data available. This study was supported by a grant from the National Natural Science Foundation of China (82171584); the Fundamental Research Funds for the Central Universities, Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province (2020E10004); and Zhejiang University Global Partnership Fund (188170-11103). TMG is supported by the Claude D. Pepper Older Americans Independence Center at Yale School of Medicine from the National Institute on Aging (P30AG021342) and the National Center for Advancing Translational Sciences (UL1TR001863). The funders had no role in the study design; data collection, analysis, or interpretation; in the writing of the report; or in the decision to submit the article for publication.

Data Availability

The data sets analyzed during this study are available at [45].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables.

[DOCX File , 62 KB - publichealth_v9i1e45502_app1.docx]

References

1. IDF Diabetes Atlas. International Diabetes Federation. URL: https://diabetesatlas.org/ [accessed 2023-05-04]
2. Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: a high-risk state for diabetes development. Lancet 2012 Jun 16;379(9833):2279-2290 [FREE Full text] [doi: 10.1016/S0140-6736(12)60283-9] [Medline: 22683128]
3. Schlesinger S, Neuenschwander M, Barbaresko J, Lang A, Maalmi H, Rathmann W, et al. Prediabetes and risk of mortality, diabetes-related complications and comorbidities: umbrella review of meta-analyses of prospective studies. Diabetologia 2022 Feb;65(2):275-285 [FREE Full text] [doi: 10.1007/s00125-021-05592-3] [Medline: 34718834]
4. Harding JL, Pavkov ME, Magliano DJ, Shaw JE, Gregg EW. Global trends in diabetes complications: a review of current evidence. Diabetologia 2019 Jan;62(1):3-16 [FREE Full text] [doi: 10.1007/s00125-018-4711-2] [Medline: 30171279]
5. Whicher CA, O'Neill S, Holt RIG. Diabetes in the UK: 2019. Diabet Med 2020 Feb;37(2):242-247. [doi: 10.1111/dme.14225] [Medline: 31901175]
6. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, on behalf of the American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of care in diabetes-2023. Diabetes Care 2023 Jan 01;46(Suppl 1):S19-S40 [FREE Full text] [doi: 10.2337/dc23-S002] [Medline: 36507649]
7. Rooney MR, Rawlings AM, Pankow JS, Echouffo Tcheugui JB, Coresh J, Sharrett AR, et al. Risk of progression to diabetes among older adults with prediabetes. JAMA Intern Med 2021 Apr 01;181(4):511-519 [FREE Full text] [doi: 10.1001/jamainternmed.2020.8774] [Medline: 33555311]
8. Lam K, Lee SJ. Prediabetes-a risk factor twice removed. JAMA Intern Med 2021 Apr 01;181(4):520-521 [FREE Full text] [doi: 10.1001/jamainternmed.2020.8773] [Medline: 33555307]
9. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
10. Sergi G, Veronese N, Fontana L, De Rui M, Bolzetta F, Zambon S, et al. Pre-frailty and risk of cardiovascular disease in elderly men and women: the Pro.V.A. study. J Am Coll Cardiol 2015 Mar 17;65(10):976-983 [FREE Full text] [doi: 10.1016/j.jacc.2014.12.040] [Medline: 25766943]

11. Soysal P, Veronese N, Thompson T, Kahl KG, Fernandes BS, Prina AM, et al. Relationship between depression and frailty in older adults: A systematic review and meta-analysis. Ageing Res Rev 2017 Jul;36:78-87. [doi: 10.1016/j.arr.2017.03.005] [Medline: 28366616]
12. Aguayo GA, Vaillant MT, Donneau A, Schritz A, Stranges S, Malisoux L, et al. Comparative analysis of the association between 35 frailty scores and cardiovascular events, cancer, and total mortality in an elderly general population in England: An observational study. PLoS Med 2018 Mar;15(3):e1002543 [FREE Full text] [doi: 10.1371/journal.pmed.1002543] [Medline: 29584726]
13. Fan J, Yu C, Guo Y, Bian Z, Sun Z, Yang L, China Kadoorie Biobank Collaborative Group. Frailty index and all-cause and cause-specific mortality in Chinese adults: a prospective cohort study. Lancet Public Health 2020 Dec;5(12):e650-e660 [FREE Full text] [doi: 10.1016/S2468-2667(20)30113-4] [Medline: 33271078]
14. Clegg A, Hassan-Smith Z. Frailty and the endocrine system. Lancet Diabetes Endocrinol 2018 Sep;6(9):743-752 [FREE Full text] [doi: 10.1016/S2213-8587(18)30110-4] [Medline: 30017798]
15. Abdul-Ghani MA, Tripathy D, DeFronzo RA. Contributions of beta-cell dysfunction and insulin resistance to the pathogenesis of impaired glucose tolerance and impaired fasting glucose. Diabetes Care 2006 May;29(5):1130-1139. [doi: 10.2337/diacare.2951130] [Medline: 16644654]
16. Lü Q, Tong N, Liu Y, Li N, Tang X, Zhao J, et al. Community-based population data indicates the significant alterations of insulin resistance, chronic inflammation and urine ACR in IFG combined IGT group among prediabetic population. Diabetes Res Clin Pract 2009 Jun;84(3):319-324. [doi: 10.1016/j.diabres.2009.03.002] [Medline: 19442860]
17. Castro-Rodríguez M, Carnicero JA, Garcia-Garcia FJ, Walter S, Morley JE, Rodríguez-Artalejo F, et al. Frailty as a major factor in the increased risk of death and disability in older people with diabetes. J Am Med Dir Assoc 2016 Oct 01;17(10):949-955. [doi: 10.1016/j.jamda.2016.07.013] [Medline: 27600194]
18. Li G, Prior JC, Leslie WD, Thabane L, Papaioannou A, Josse RG, CaMos Research Group. Frailty and risk of fractures in patients with type 2 diabetes. Diabetes Care 2019 Apr;42(4):507-513. [doi: 10.2337/dc18-1965] [Medline: 30692240]
19. Pandey A, Khan MS, Garcia K, Simpson F, Bahnson J, Patel KV, et al. Association of baseline and longitudinal changes in frailty burden and risk of heart failure in type 2 diabetes-findings from the Look AHEAD Trial. J Gerontol A Biol Sci Med Sci 2022 Dec 29;77(12):2489-2497 [FREE Full text] [doi: 10.1093/gerona/glac094] [Medline: 35453142]
20. Chao C, Wang J, Chien K, COhort of GEriatric Nephrology in NTUH (COGENT) study group. Both pre-frailty and frailty increase healthcare utilization and adverse health outcomes in patients with type 2 diabetes mellitus. Cardiovasc Diabetol 2018 Sep 27;17(1):130 [FREE Full text] [doi: 10.1186/s12933-018-0772-2] [Medline: 30261879]
21. Chode S, Malmstrom TK, Miller DK, Morley JE. Frailty, diabetes, and mortality in middle-aged African Americans. J Nutr Health Aging 2016;20(8):854-859. [doi: 10.1007/s12603-016-0801-3] [Medline: 27709235]
22. Chhetri JK, Zheng Z, Xu X, Ma C, Chan P. The prevalence and incidence of frailty in pre-diabetic and diabetic community-dwelling older population: results from Beijing longitudinal study of aging II (BLSA-II). BMC Geriatr 2017 Feb 08;17(1):47 [FREE Full text] [doi: 10.1186/s12877-017-0439-y] [Medline: 28178934]
23. He D, Li J, Li Y, Zhu J, Zhou T, Xu Y, et al. Frailty is associated with the progression of prediabetes to diabetes and elevated risks of cardiovascular disease and all-cause mortality in individuals with prediabetes and diabetes: Evidence from two prospective cohorts. Diabetes Res Clin Pract 2022 Dec;194:110145. [doi: 10.1016/j.diabres.2022.110145] [Medline: 36356844]
24. Feldman-Billard S, Dupas B. Eye disorders other than diabetic retinopathy in patients with diabetes. Diabetes Metab 2021 Nov;47(6):101279. [doi: 10.1016/j.diabet.2021.101279] [Medline: 34534696]
25. Wang F, Wang S, Zong Q, Zhang Q, Ng CH, Ungvari GS, et al. Prevalence of comorbid major depressive disorder in type 2 diabetes: a meta-analysis of comparative and epidemiological studies. Diabet Med 2019 Aug;36(8):961-969. [doi: 10.1111/dme.14042] [Medline: 31127631]
26. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015 Mar;12(3):e1001779 [FREE Full text] [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
27. UK Biobank: Protocol for a large-scale prospective epidemiological resource. UK Biobank. 2007 Mar 21. URL: https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf [accessed 2023-05-04]
28. Hanlon P, Nicholl BI, Jani BD, Lee D, McQueenie R, Mair FS. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. The Lancet Public Health 2018 Jul;3(7):e323-e332. [doi: 10.1016/s2468-2667(18)30091-4]
29. van Buuren SV, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J. Stat. Soft 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
30. Kong L, Lyu Q, Yao H, Yang L, Chen S. The prevalence of frailty among community-dwelling older adults with diabetes: A meta-analysis. Int J Nurs Stud 2021 Jul;119:103952. [doi: 10.1016/j.ijnurstu.2021.103952] [Medline: 34022743]
31. O'Caoimh R, Sezgin D, O'Donovan MR, Molloy DW, Clegg A, Rockwood K, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. Age Ageing 2021 Jan 08;50(1):96-104. [doi: 10.1093/ageing/afaa219] [Medline: 33068107]

32. Palmer AK, Gustafson B, Kirkland JL, Smith U. Cellular senescence: at the nexus between ageing and diabetes. Diabetologia 2019 Oct;62(10):1835-1841 [FREE Full text] [doi: 10.1007/s00125-019-4934-x] [Medline: 31451866]
33. Raymond E, Reynolds CA, Dahl Aslan AK, Finkel D, Ericsson M, Hägg S, et al. Drivers of Frailty from Adulthood into Old Age: Results from a 27-Year Longitudinal Population-Based Study in Sweden. J Gerontol A Biol Sci Med Sci 2020 Sep 25;75(10):1943-1950 [FREE Full text] [doi: 10.1093/gerona/glaa106] [Medline: 32348465]
34. Liccini A, Malmstrom TK. Frailty and sarcopenia as predictors of adverse health outcomes in persons with diabetes mellitus. J Am Med Dir Assoc 2016 Sep 01;17(9):846-851. [doi: 10.1016/j.jamda.2016.07.007] [Medline: 27569712]
35. Ros M, Carrascosa JM. Current nutritional and pharmacological anti-aging interventions. Biochim Biophys Acta Mol Basis Dis 2020 Mar 01;1866(3):165612 [FREE Full text] [doi: 10.1016/j.bbadis.2019.165612] [Medline: 31816437]
36. Travers J, Romero-Ortuno R, Bailey J, Cooney M. Delaying and reversing frailty: a systematic review of primary care interventions. Br J Gen Pract 2019 Jan;69(678):e61-e69 [FREE Full text] [doi: 10.3399/bjgp18X700241] [Medline: 30510094]
37. Mugueta-Aguinaga I, Garcia-Zapirain B. Is technology present in frailty? Technology a back-up tool for dealing with frailty in the elderly: a systematic review. Aging Dis 2017 Apr;8(2):176-195 [FREE Full text] [doi: 10.14336/AD.2016.0901] [Medline: 28400984]
38. Gallucci A, Trimarchi PD, Abbate C, Tuena C, Pedroli E, Lattanzio F, et al. ICT technologies as new promising tools for the managing of frailty: a systematic review. Aging Clin Exp Res 2021 Jun;33(6):1453-1464. [doi: 10.1007/s40520-020-01626-9] [Medline: 32705589]
39. Cruz AM, Monsalve L, Ladurner A, Jaime LF, Wang D, Quiroga DA. Information and communication technologies for managing frailty: a systematic literature review. Aging Dis 2021 Jun;12(3):914-933 [FREE Full text] [doi: 10.14336/AD.2020.1114] [Medline: 34094651]
40. Aznar-Tortonda V, Palazón-Bru A, la Rosa DMF, Espínola-Morel V, Pérez-Pérez BF, León-Ruiz AB, et al. Detection of frailty in older patients using a mobile app: cross-sectional observational study in primary care. Br J Gen Pract 2020 Jan;70(690):e29-e35 [FREE Full text] [doi: 10.3399/bjgp19X706577] [Medline: 31685541]
41. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
42. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001 May 03;344(18):1343-1350. [doi: 10.1056/NEJM200105033441801] [Medline: 11333990]
43. Anker D, Carmeli C, Zwahlen M, Rodondi N, Santschi V, Henchoz Y, et al. How blood pressure predicts frailty transitions in older adults in a population-based cohort study: a multi-state transition model. Int J Epidemiol 2022 Aug 10;51(4):1167-1177 [FREE Full text] [doi: 10.1093/ije/dyab210] [Medline: 34652417]
44. Wang M, Li T, Li C, Liu C, Lin W, Lin C, et al. Frailty, transition in frailty status and all-cause mortality in older adults of a Taichung community-based population. BMC Geriatr 2019 Jan 28;19(1):26 [FREE Full text] [doi: 10.1186/s12877-019-1039-9] [Medline: 30691410]
45. UK Biobank. URL: https://www.ukbiobank.ac.uk/enable-your-research/register [accessed 2023-05-04]

Abbreviations

ADA: American Diabetes Association
CKD: chronic kidney disease
CVD: cardiovascular disease
FP: frailty phenotype
HbA_{1c}: glycated hemoglobin
HR: hazard ratio
ICD-9: International Statistical Classification of Diseases and Related Health Problems, 9th version
ICD-10: the International Statistical Classification of Diseases and Related Health Problems, 10th version
T2DM: type 2 diabetes mellitus
TDI: Townsend deprivation index
UKB: UK Biobank

Edited by A Mavragani, T Sanchez; submitted 05.01.23; peer-reviewed by S Thirunavukkarasu, V Buss, Y Jeem; comments to author 06.02.23; revised version received 17.03.23; accepted 23.03.23; published 18.05.23.

Please cite as:
Cao X, Li X, Zhang J, Sun X, Yang G, Zhao Y, Li S, Hoogendijk EO, Wang X, Zhu Y, Allore H, Gill TM, Liu Z
Associations Between Frailty and the Increased Risk of Adverse Outcomes Among 38,950 UK Biobank Participants With Prediabetes: Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e45502
URL: https://publichealth.jmir.org/2023/1/e45502
doi:10.2196/45502
PMID:37200070

Review

Willingness to Vaccinate Against Herpes Zoster and Its Associated Factors Across WHO Regions: Global Systematic Review and Meta-Analysis

Qiang Wang[1,2], MM; Liuqing Yang[1], MPH; Lan Li[3], MPH; Chang Liu[1], MM; Hui Jin[1,4], PhD; Leesa Lin[2,5,6], PhD

[1]Department of Epidemiology and Health Statistics, School of Public Health, Southeast University, Nanjing, China

[2]Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom

[3]Centre for Digital Public Health in Emergencies, Institute for Risk and Disaster Reduction, University College London, London, United Kingdom

[4]Key Laboratory of Environmental Medicine Engineering, Ministry of Education, School of Public Health, Southeast University, China

[5]Laboratory of Data Discovery for Health, Hong Kong Science Park, Hong Kong, China

[6]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China

Corresponding Author:
Leesa Lin, PhD
Department of Infectious Disease Epidemiology
London School of Hygiene and Tropical Medicine
Keppel Street
London, WC1E 7TH
United Kingdom
Phone: 44 617 632 6142
Email: Leesa.Lin@lshtm.ac.uk

Abstract

Background: A life-course immunization approach would enhance the quality of life across all age groups and improve societal well-being. The herpes zoster (HZ) vaccine is highly recommended for older adults to prevent HZ infection and related complications. The proportions of willingness to receive the HZ vaccine varies across countries, and various kinds of factors, including sociodemographics and individual perceptions, influence the willingness to vaccinate.

Objective: We aim to estimate the HZ vaccination willingness rate and identify factors associated with vaccine uptake willingness across all World Health Organization (WHO) regions.

Methods: A global systematic search was performed on PubMed, Web of Science, and the Cochrane Library for all papers related to the HZ vaccine published until June 20, 2022. Study characteristics were extracted for each included study. Using double arcsine transformation, vaccination willingness rates with 95% CIs were pooled and reported. The willingness rate and associated factors were analyzed by geographical context. Associated factors were also summarized based on Health Belief Model (HBM) constructs.

Results: Of the 26,942 identified records, 13 (0.05%) papers were included, covering 14,066 individuals from 8 countries in 4 WHO regions (Eastern Mediterranean Region, European Region, Region of the Americas, and Western Pacific Region). The pooled vaccination willingness rate was 55.74% (95% CI 40.85%-70.13%). Of adults aged ≥50 years, 56.06% were willing to receive the HZ vaccine. After receiving health care workers' (HCWs) recommendations, 75.19% of individuals were willing to get the HZ vaccine; without HCWs' recommendations, the willingness rate was only 49.39%. The willingness rate was more than 70% in the Eastern Mediterranean Region and approximately 55% in the Western Pacific Region. The willingness rate was the highest in the United Arab Emirates and the lowest in China and the United Kingdom. The perception of HZ severity and susceptibility was positively associated with vaccination willingness. The perceived barriers to vaccination willingness (main reasons for unwillingness) included low trust in the effectiveness of the HZ vaccine, concerns about safety, financial concerns, and being unaware of the HZ vaccine's availability. Older individuals, those having lower education, or those having lower income levels were less likely to willing to be vaccinated.

Conclusions: Only 1 in 2 individuals showed a willingness to be vaccinated against HZ. The willingness rate was the highest in the Eastern Mediterranean Region. Our findings show the critical role HCWs play in promoting HZ vaccination. Monitoring

HZ vaccination willingness is necessary to inform public health decision-making. These findings provide critical insights for designing future life-course immunization programs.

(JMIR Public Health Surveill 2023;9:e43893) doi:10.2196/43893

KEYWORDS

herpes zoster vaccine; willingness; associated factors; systematic review

Introduction

A life-course immunization approach would enhance the quality of life across all age groups and improve societal well-being through a healthier population [1,2]. Older adults are at high risk of being ill, being hospitalized, and dying from many vaccine-preventable diseases [3]. The World Health Organization (WHO) and health authorities in various countries have recommended several vaccines for older adults, including the seasonal influenza vaccine; pneumococcal vaccine; hepatitis B vaccine; the tetanus, diphtheria, and pertussis (Tdap) vaccine booster; and, most recently, COVID-19 vaccines [2]. The herpes zoster (HZ) vaccine is highly recommended for older adults for reducing the incidence of HZ [4].

Caused by varicella zoster virus (VZV), HZ is a distinctive syndrome occurring when immunity to VZV declines due to age or immunosuppression [5]. It presents a disproportionate risk to adults aged 50 years and over, with estimated incidence rates reaching 5.2-10.9 cases per 1000 person-years [6]. Due to the potential for serious complications, postherpetic neuralgia (PHN), neurological sequelae, HZ ophthalmicus with eye involvement, and disseminated disease, this disease has a significant impact on the quality of life [6,7]. It was reported that 5%-30% of patients with HZ are likely to develop PHN with persistent pain for more than 1 month [7].

Existing evidence supports the substantial efficacy of the HZ vaccine; a meta-analysis showed that the efficacy of the adjuvant recombinant subunit HZ vaccine is 94% when compared with the placebo [8]. Additionally, HZ vaccination was found to be cost-effective versus no vaccination for older adults in high-income countries based on health economics evaluations [9,10].

Currently, 2 types of HZ vaccines (Zostavax and Shingrix) are commonly used in over 60 countries, being recommended to adults aged ≥50 years or adults aged ≥18 years with an immunocompromised state [11]. Noticeably, HZ vaccination uptake is relatively low worldwide. In the United States, 24.1% of adults aged ≥50 years received an HZ vaccine in 2018 [12]. The coverage rates among older adults in Australia and Canada were reported to be lower than 10% [13,14]; only 1.8% of adults over 65 years old received the HZ vaccine in Turkey [15].

Vaccine uptake is positively associated with the willingness to vaccinate. The willingness to receive the HZ vaccine has been surveyed in previous studies [16-18]. However, the proportions of the willingness to receive the HZ vaccine varies across countries, ranging from 17% (China) to 90% (Australia) [16,17]. Importantly, vaccination willingness is influenced by multiple factors, which include sociodemographic, cognitive, psychologic, and politics-and-culture contexts [19]. According to recent evidence, various kinds of factors, including sociodemographics (age and income level) and individual perceptions (perception of disease susceptibility), influence HZ vaccination willingness [16-18]. Additionally, these associated factors may differ across different regions.

Summarizing and understanding the willingness to receive the HZ vaccine and its associated factors is thus essential for obtaining insights into intervention strategies to improve its uptake among older adults. It is also necessary to explore methods mitigating vaccine hesitancy among older adults, which is also an indispensable part of life-course immunization practices. However, there has only been scant attention paid to HZ vaccination willingness in the research literature. To fill this gap, this systematic review seeks to summarize the up-to-date evidence on the HZ vaccination willingness rate and identify its associated factors.

Methods

Study Design

The Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) statement was used for this systematic review [20].

Search Strategy and Selection Criteria

A systematic search was conducted in 3 electronic databases (PubMed, Web of Science, and the Cochrane Library) using the following search terms: ["Herpes Zoster" OR "Herpesvirus 3" OR shingles OR zoster OR "varicellovir*" OR hhv3 OR hhv-3 OR varicella-zoster OR postherpetic] AND ["vaccin*" OR "immuniz*" OR "inocul*"]. All studies published from the inception of the databases until June 20, 2022, were included. After combining all search records, duplicate records were identified and removed. Titles and abstracts were then screened for relevance, and full texts of the included records were retrieved and reviewed.

The screening process was independently performed by 2 reviewers (authors QW and LQY) with the following inclusion criteria: (1) cross-sectional surveys or cohort studies and (2) relevance to HZ vaccine attitudes, intentions, willingness, and acceptance. There were no restrictions on language or paper type. Papers were excluded if they met any of the following criteria: (1) willingness data were not reported, (2) insufficient data were provided for pooling (at least 2 of the 3 following topics were absent: total number of surveyed individuals, number of individuals willing to receive the HZ vaccine, and willingness rate), (3) they were duplicate data or surveys, or (4) they were experimental and observational studies that reported the willingness rate after interventions. Any disagreement was resolved by discussions with a third reviewer (author LL). The review protocol is available in the International Prospective Register of Systematic Reviews (PROSPERO id: CRD42022348426).

Data Abstraction and Quality Assessment

For each included study, we extracted the following data: title, first author, journal name, published date, paper type, sampling method, study setting, study period, study population, study location, and inquiry questions (were there any assumptions, for example, free vaccination was provided). The number of surveyed individuals, the number of individuals who accepted HZ vaccination, and the willingness rate were also extracted for pooling. Additionally, influencing factors examined in each study were abstracted.

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement was used to assess the quality of the included papers [21]. Data abstraction and quality assessment were conducted by QW and LQY independently. A third researcher, LL, resolved disputes when the data extracted or quality assessment scores for papers were not consistent between the 2 initial reviewers.

Statistical Analysis

Using double arcsine transformation on 2 variables, number of surveyed individuals and number of individuals accepting HZ vaccination, we calculated and reported the pooled willingness rate and the 95% CI [22,23]. If these data were not included in the study, the number of individuals accepting HZ vaccination could be calculated by multiplying the total number by the willingness rate. Due to high heterogeneity, DerSimonian-Laird random effects were used in the calculation.

Stratified subgroup analysis was subsequently performed to explore the causes of heterogeneity according to the study characteristics (eg, sampling method, study setting, study period, and study population). Studies that shared the same characteristics were grouped together (eg, surveyed using the random sampling method or the convenience sampling method); when the characteristic was unclear, the study was labeled as “not mentioned” in the subgroup analysis. For each subgroup, the willingness rate was pooled. To examine the potential impact of the COVID-19 pandemic on the willingness to receive the HZ vaccine, we divided the study period into 2 subperiods, before the pandemic and during the pandemic. The study locations were categorized by their WHO regions and economic levels [24,25]. Due to the limited availability of data, it was difficult to further estimate the willingness rate by age group. We only reported the willingness rate among individuals aged ≥50 years in further analysis of age.

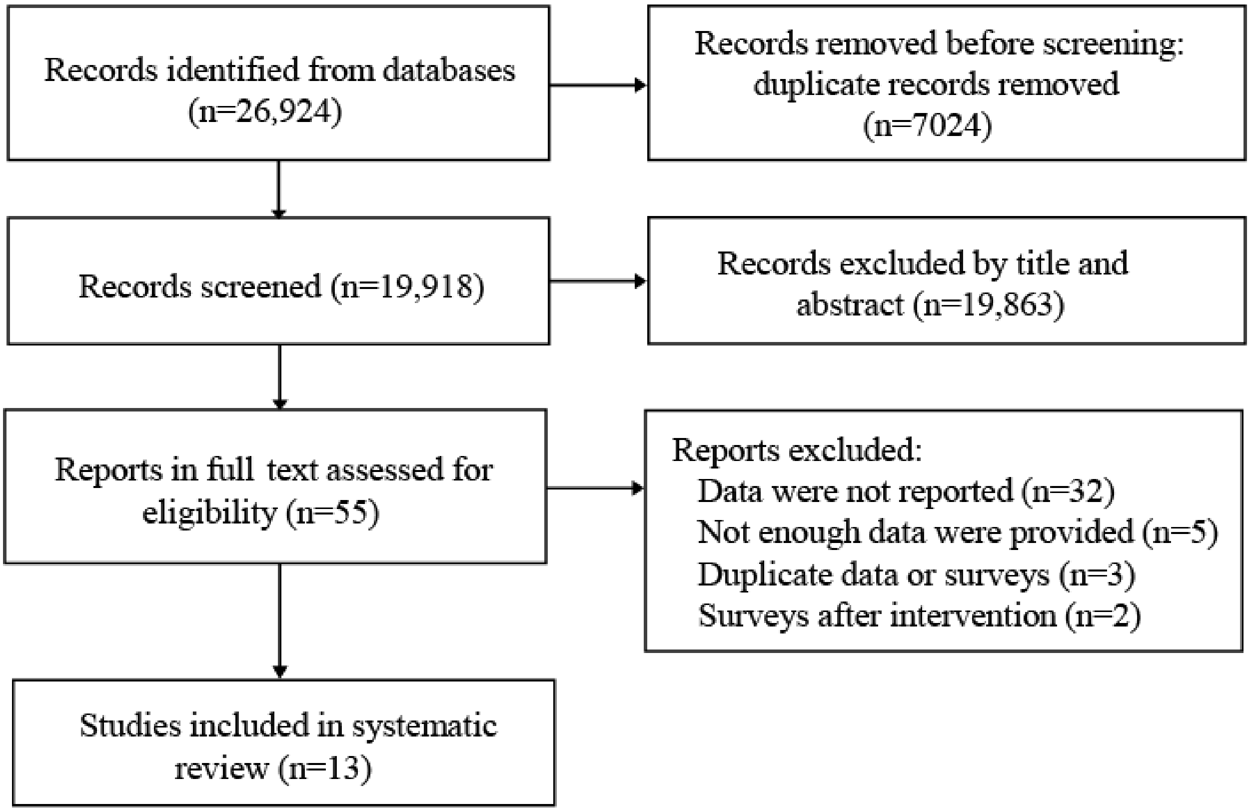
The willingness rate and its associated factors were examined by geographical context in order to explore potential differences. We also classified factors based on Health Belief Model (HBM) constructs. This method was previously used to identify factors associated with seasonal influenza vaccination, human papillomavirus vaccination, and COVID-19 vaccination [26-28]. Following the HBM construct, the influencing factors were summarized and classified according to the perceived susceptibility and severity of HZ, the perceived benefits of and barriers to vaccination willingness (reasons for vaccination willingness or unwillingness), modified factors, and cues to action. We evaluated the publication bias using the funnel plot [29]. The publication bias was not adjusted because of the presence of substantial heterogeneity [30]. All analyses were conducted in STATA version 14.0 (StataCorp) and Microsoft Excel 2016.

Results

Search Results and Quality Assessment

The initial search identified 26,942 papers. After deduplication and screening by title and abstract, 55 (0.20%) papers were screened by full text. A total of 13 (23.6%) papers were included in this review. The detailed selection process is illustrated in Figure 1. Of the 13 studies, 10 (76.9%) were empirical studies [16,18,31-38], 2 (15.4%) were meeting abstracts [17,39], and 1 (7.7%) was a letter [40]. All 13 studies belong to cross-section surveys. The final analysis included 14,066 individuals from 8 countries (Australia, Canada, China, France, South Korea, the United Arab Emirates, the United Kingdom, and the United States). The 13 studies were conducted in 4 WHO regions: the Eastern Mediterranean Region, the European Region, the Region of the Americas, and the Western Pacific Region. Study populations were adults aged ≥50 years in 10 (76.9%) studies, adults aged ≥18 years in 2 (15.4%) studies, and not specified in the remaining 1 (7.7%) study. The quality assessment scores of the included studies ranged from 10 to 20, with an average of 16.30 (SD 3.86). Detailed study characteristics and quality assessment results of the included studies are provided in Multimedia Appendix 1, Tables S1 and S2.

Figure 1. Flowchart of paper selection.



Vaccination Willingness Rate

The pooled HZ vaccination willingness rate was estimated to be 55.74% (95% CI 40.85%-70.13%). The asymmetry plots in the funnel plot indicated that there might be publication bias (Multimedia Appendix 1, Figure S1). It was estimated that 56.06% (95% CI 37.26%-74.02%) of adults aged ≥50 years were willing to receive an HZ vaccine. There were no significant differences between pooled willingness rates from surveys in the hospital context (58.29%, 95% CI 34.35%-80.33%) and surveys in the community context (53.45%, 95% CI 33.07%-73.26%), as seen in Multimedia Appendix 1, Figure S2. The willingness rate was 71.90% (95% CI 67.42%-75.99%) in the Eastern Mediterranean Region and 55.68% (95% CI 22.03%-86.67%) in the Western Pacific Region. The willingness rate was estimated to be 60.18% (95% CI 47.53%-72.18%) in high-income economies and 23.97% (95% CI 22.22%-25.77%) in upper-middle-income economies. Under health care workers' (HCWs) recommendations, 75.19% (95% CI 54.61%-91.07%) of individuals were willing to receive the HZ vaccine, which was higher than the willingness rate among individuals (49.39%, 95% CI 30.42%-68.45%) without HCWs' recommendations.

Rate and Associated Factors by Geographical Context

Eastern Mediterranean Region

Only 1 (7.7%) study was conducted in the United Arab Emirates [35], which reported a willingness rate of 71.9%, as shown in Multimedia Appendix 1, Figure S3. This study found that persons who were unsure about their chickenpox infection history were less likely to receive the HZ vaccine. In this study, the top 3 reasons respondents provided explaining their unwillingness were "prefers to take medication when sick," "side effects of vaccine," and "not at risk since I am healthy."

European Region

Of the 13 studies, 2 (15.4%) reported findings from surveys conducted in France (68.91%) [31] and the United Kingdom (34.62%) [36]. In France, women and persons in a relationship/married were more likely to vaccinate themselves against HZ. This study also identified that persons who agreed with the statements "HZ causes pain," "HZ is always severe," or "vaccination is a good prevention tool against HZ" preferred to receive the HZ vaccine. In the United Kingdom, Nicholls et al [36] reported the association between vaccine hesitancy using the Vaccination Attitudes Examination (VAX) scale and the 5C scale. Lower scores on "collective responsibility" on the 5C scale and higher scores on "concerns" on the VAX scale were associated with a lower likelihood of receiving the HZ vaccine.

Region of the Americas

A total of 4 (30.8%) studies were conducted in the Region of the Americas, including the United States (n=3, 75%) [33,34,40] and Canada (n=1, 25%) [32]. The willingness rates in the United States and Canada were 53.28% and 55.01%, respectively. People in the United States with a higher income were more likely to receive an HZ vaccine [34]. Two surveys in the United States reported reasons for the refusal of the HZ vaccine; the top 3 reasons in Lu et al's [34] study were "the vaccination was not needed," "not a risk, healthy, or high immune," and "not trusting doctors or medicine"; according to Funovits et al [40], the top 3 reasons were "believes vaccine is not important," "not covered by insurance," and "would like to obtain primary care physician recommendations."

Western Pacific Region

There were 6 (34.6%) surveys carried out in Australia (n=1, 16.7%), South Korea (n=1, 16.7%), and China (n=4, 66.6%). More than 85% of individuals were willing to accept the HZ vaccine in Australia and South Korea, considerably higher than that observed in China (35.13%). As shown in surveys in South Korea [37], younger people were more likely to accept the HZ vaccine. Individuals with a college education were more likely to accept the HZ vaccine than those with incomplete high school education. In this study, the top 3 reasons for unwillingness were "low perceived risk of developing HZ," "concerns about the adverse effects following immunization," and "concerns about the vaccination cost" [37].

Of the 4 studies in China, 2 (50%) were conducted in Hong Kong and 2 (50%) in Shanghai City. In the studies in Hong Kong, the most frequent reason listed for refusal was "they were unaware of its availability" [18]. According to Lu et al [16], in Shanghai City, younger people, people with higher monthly incomes, people who know the elderly are susceptible to HZ, and people with lower vaccine hesitancy levels were more willing to receive the vaccine. Qiu et al [38] reported that people who know about the HZ vaccine, are concerned with getting an HZ infection themselves, and who believe the vaccine can prevent HZ have a greater likelihood of receiving the HZ vaccine [38].

Associated Factors Using HBM Constructs

Information about how factors associated with vaccination willingness relate to HBM constructs is provided in Table 1 and Multimedia Appendix 1, Table S3.

Table 1. Description of influencing factors using HBM[a] constructs.

HBM framework	Factors	References
Individual perceptions		
Perceived susceptibility and severity of HZ[b]	• Perception of HZ infection susceptibility and severity	• [16,31,38]
Perceived barriers to vaccination willingness	• Low trust to the effectiveness of HZ vaccines, unawareness of HZ vaccine availability, concerns about side effects and safety, and cost	• [18,34,35,37,39]
Modifying factors		
Sociodemographics	• Age	• [16,31,33,34,37,38,40]
	• Education level	• [16,31,34,37,38,40]
	• Gender	• [31,34,37,38,40]
	• Income level	• [16,34,37,38]
	• Marital status	• [31,34,36]
	• Insurance	• [16,34,35]
	• Race	• [34]
	• Nationality	• [35]
	• Occupation	• [35]
Knowledge, attitude, beliefs, and prior experience	• Confidence in the effectiveness of the HZ vaccine	• [31,38]
	• Knowing the HZ vaccine	• [31,38]
	• Receiving influenza vaccination	• [34,37]
	• HZ infection history	• [31,35,37,38]
	• With chronic diseases	• [35,38]
Cues to action		
Interpersonal relationships	• Knowing someone with a history of HZ	• [31,37,38]
Community	• Doctors' recommendations	• [17]

[a] HBM: Health Belief Model.

[b] HZ: herpes zoster.

Perceived Susceptibility and Severity of Herpes Zoster

Of the 13 studies, 2 (15.4%) reported that the perception of disease severity is positively associated with vaccination willingness [31,38], while 1 (7.7%) study found a positive relationship between perceived infection susceptibility and vaccination willingness [16].

Perceived Benefits of and Barriers to Vaccination Willingness

Of the 13 studies, 5 (38.5%) explored reasons for vaccination unwillingness (perceived barriers to willingness). Low trust in the effectiveness of the HZ vaccine, concerns about safety, low perception of disease risk, financial concerns, and unawareness of the HZ vaccine's availability contributed to the unwillingness to receive the HZ vaccine [18,34,35,37,40].

Modified Factors

Modified factors included (1) sociodemographics and (2) knowledge, attitude, beliefs, and prior experience in the analysis.

Sociodemographics

Age (7/13, 53.8%), education level (6/13, 46.2%), gender (5/13, 38.5%), and income level (4/13, 30.8%) were the most frequent factors reported in the included studies. Of the 13 studies, 3 (23.1%) found a nonsignificant association between age and vaccination willingness [31,34,40], whereas 4 (30.8%) reported a significant relationship [16,33,37,38]. Nonsignificant relationships between education level and willingness were reported in 3 (23.1%) studies [31,34,40], while the other 3 (23.1%) studies found that a higher education level is associated with higher willingness rates [16,37,38]. In addition, 4 (30.8%) studies showed a nonsignificant association between gender and vaccination willingness [34,37,38,40]; however, 1 (7.7%) study reported that males are more likely to receive the HZ vaccine [31]. It was reported that individuals with higher income levels are more likely to receive the HZ vaccine [16,34], but 2 (15.4%) studies found no significant association between income level and willingness rate [37,38]. Additionally, 1 (7.7%) study reported a significant association between marital status and vaccination willingness [31], whereas 2 (15.4%) studies found no significant association between these factors [34,36].

Knowledge, Attitude, Beliefs, and Prior Experience

We found that confidence in the effectiveness of the HZ vaccine is positively associated with the willingness to receive the HZ vaccine [31,38]. Individuals who knew about the HZ vaccine were more likely to receive it [38], a finding that was not observed in another study [31]. Having a history of HZ infection was positively correlated with vaccination willingness [38]; however, 2 (15.4%) studies reported only nonsignificant associations between infection history and vaccination willingness [31,37]. Influenza vaccination history was not found to be an influencing factor for HZ vaccination acceptance [34,37]. In addition, 1 (7.7%) study reported that individuals without chronic diseases are more likely to receive the HZ vaccine [38], whereas another study did not find a significant association between these factors [35].

Cues to Action

Doctors' recommendations to receive the HZ vaccine were positively associated with vaccination willingness [17]. The experience of knowing someone with a history of HZ was positively associated with vaccination willingness [38], albeit this relationship was found to be insignificant in 2 (15.4%) studies [31,37].

Discussion

Principal Findings

This systematic review demonstrated that the pooled HZ vaccination willingness rate is 55.74% worldwide. The main reasons for the unwillingness to receive the HZ vaccine include low trust in the effectiveness of the HZ vaccine, concerns about safety, low perceptions of disease risk, financial concerns, and unawareness of the availability of the HZ vaccine. Our findings showed that HCWs' recommendations are correlated with a greater likelihood of receiving an HZ vaccine.

A large variability in HZ vaccination rates was found; however, certain patterns could still be deduced. Higher vaccination acceptance rates were observed in countries with an earlier license date of the HZ vaccine. From 2006 onward, countries with developed economies gradually licensed the HZ vaccine, such as the United States, France, Australia, South Korea, and the United Kingdom [11]. People living in these countries may have a greater awareness of HZ vaccine availability and may have received more frequent recommendations from HCWs to receive the vaccine. The willingness rate in China was considerably lower than that in these countries. HZ vaccines were first licensed in China in 2019 [11]. Less than one-third of the participants knew about the availability of the HZ vaccine in Shanghai, China [38]. The lack of awareness of HZ vaccine availability among the public may contribute to the low vaccination willingness rate in China. However, although the United Kingdom approved the HZ vaccine in 2006, the observed vaccination willingness rate remains low. The willingness rate might also be associated with the incidence of HZ across the countries. According to van Oorschot et al [6], the cumulative incidence of HZ in South Korea and Australia (more than 10 cases per 1000 population) is considerably higher than that in Canada, China, and France (between 2.9 and 8.67 cases per 1000 population) [6]. Thus, the higher incidence rate might result in individuals having an increased perception of susceptibility to HZ infection.

Recommendations from HCWs were associated with a higher HZ vaccination willingness rate; the impact of HCWs' recommendations on vaccine uptake has been observed in previous studies [41,42]. In turn, this channel can be an effective communication strategy to tackle vaccine hesitancy [43]. Glenton et al [44] summarized HCWs' views and experiences of communicating about vaccination with individuals over 50 years old and found that misinformation, fears, and concerns about vaccines are commonly voiced among older adults during their communications with HCWs. Hence, HCWs' attitudes and practices regarding the HZ vaccine are crucial. Reassuringly, HZ vaccination was reportedly recommended by 65.6% of HCWs in Italy [45]. A study in the United States reported that more than 90% of HCWs reported were willing to recommended HZ vaccination to their patients [46]. However, there remains a gap in relevant data in countries with upper-middle-income economies; therefore, more studies in upper-middle-income economies should be carried out to determine provider willingness to recommend an HZ vaccine.

Sociodemographic factors, while inconclusive, are essential to understanding and predicting HZ vaccination willingness rates. Individuals with lower education and income levels might be less likely to receive the HZ vaccine. Similar results have been observed in previous reviews of factors associated with influenza vaccination uptake and COVID-19 vaccination willingness among older adults [47,48]. Low socioeconomic status can be an important barrier to vaccination for older adults. Worryingly, the price of the HZ vaccine is considerably higher (about 10 times) than that of other vaccines, such as the influenza vaccine and the Tdap vaccine [49,50]. Along with a low perception of disease risk, low trust in the HZ vaccine, and awareness of HZ vaccine availability, the price of the vaccine might also represent a significant barrier to vaccination willingness.

Implications for Policy and Practice

Our findings have important implications for developing intervention strategies for increasing HZ vaccine acceptance across multiple levels. Countries with earlier licensing dates should pay attention to methods to improve individuals' confidence in the HZ vaccine and decrease their complacency. It is essential for countries with later licensing dates, such as China, to increase the awareness of the availability of the HZ vaccine first. Embracing the idea of communication strategies could be a potentially effective way to improve the acceptance of the HZ vaccine. One such strategy is engaging HCWs, as HCWs play a crucial role in older peoples' vaccination decisions [44]. Active recommendations from HCWs could be an effective way to increase vaccine awareness and willingness, as well as vaccine confidence. Using social media, such as posting content on Facebook or YouTube, is another feasible way to educate people about HZ vaccination [51]. At the individual level, our findings could facilitate greater awareness of people with identified factors for unwillingness (eg, age and income) and may help create tailored interventions for these populations using behavioral change strategies and techniques [52]. Our findings also indicate that some challenges exist toward the practice of life-course immunization, such as individuals being might face some challenges. Due to a lack of information, individuals might need to be aware of age-specific immunizations, which further emphasizes that the role of HCWs in promoting life-course immunization is vital and critical.

Limitations

Several limitations of this review were identified. First, the representativeness of the samples in the included studies was unclear. Only 3 of the 13 studies used a random sampling method. However, stratified subgroup analysis was adopted to mitigate this potential bias. Although no significant differences were observed among sampling methods, the results need to be interpreted with caution. Second, a lack of studies in this area limited the diversity and representativeness of our review findings. Since most studies were performed in countries with high-income economies, more studies need to be carried out in less economically developed countries. Additionally, some associated factors might not be reported, because the included studies did not explore them. For example, vaccination policy and smoking habits were determined to be associated with influenza vaccination willingness [53,54]; trust in the government and information exposure were associated with the COVID-19 vaccine [28,55]. These could be explored in future research.

Conclusion

Only 1 in 2 individuals showed a willingness to be vaccinated against HZ. The willingness rate was the highest in the Eastern Mediterranean Region. A higher acceptance rate could be generally observed in countries licensing HZ vaccines earlier or having higher HZ incidence, such as Australia and South Korea. The reasons for unwillingness included low trust in the effectiveness of the HZ vaccine, concerns about safety, low perceptions of disease risk, financial concerns, and unawareness of the availability of the HZ vaccine. Our findings showed the critical role HCWs play in promoting HZ vaccination. Countries with earlier vaccine licensing dates and high HZ incidence should improve individuals' confidence in the vaccine and decrease their complacency. It is necessary for countries with later licensing dates to prioritize the improvement of the awareness of HZ vaccine availability. Monitoring HZ vaccination willingness is necessary to inform public health decision-making. These findings will also be of interest to the practice of life-course immunization.

Acknowledgments

This work was supported by AIR@InnoHK administered by the Innovation and Technology Commission. Further, QW was also supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX20_0153), the Fundamental Research Funds for the Central Universities (3225002102D), and the China Scholarship Council (202106090165). LL (Leesa Lin) was supported by AIR@InnoHK administered by the Innovation and Technology Commission. LL (Lan Li) was partially supported by the China Scholarship Council (202008060009). The funder/sponsor did not participate in the work.

Data Availability

The data underlying this paper are available within the paper and in Multimedia Appendix 1.

Authors' Contributions

QW and LL (Leesa Lin) proposed and initiated this study. QW, LQY, CL, and LL (Lan Li) designed the data collection instruments and collected data. QW carried out the initial analyses. QW, LL (Lan Li), HJ, and LL (Leesa Lin) drafted the initial manuscript and reviewed and revised the manuscript. All authors commented on and revised drafts of the manuscript. All authors have read and approved the final report.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material.
[DOCX File , 383 KB - publichealth_v9i1e43893_app1.docx]

References

1. Tate J, Aguado T, Belie JD, Holt D, Karafillakis E, Larson HJ, et al. The life-course approach to vaccination: harnessing the benefits of vaccination throughout life. Vaccine 2019 Oct 16;37(44):6581-6583. [doi: 10.1016/j.vaccine.2019.09.016] [Medline: 31558327]
2. Philip RK, Attwell K, Breuer T, Di Pasquale A, Lopalco PL. Life-course immunization as a gateway to health. Expert Rev Vaccines 2018 Oct;17(10):851-864. [doi: 10.1080/14760584.2018.1527690] [Medline: 30350731]
3. Cunningham AL, McIntyre P, Subbarao K, Booy R, Levin MJ. Vaccines for older adults. BMJ 2021 Feb 22;372:n188. [doi: 10.1136/bmj.n188] [Medline: 33619170]
4. Berlinberg EJ, Deiner MS, Porco TC, Acharya NR. Monitoring interest in herpes zoster vaccination: analysis of Google Search data. JMIR Public Health Surveill 2018 May 02;4(2):e10180 [FREE Full text] [doi: 10.2196/10180] [Medline: 29720364]
5. Dworkin RH, Johnson RW, Breuer J, Gnann JW, Levin MJ, Backonja M, et al. Recommendations for the management of herpes zoster. Clin Infect Dis 2007 Jan 01;44 Suppl 1:S1-S26. [doi: 10.1086/510206] [Medline: 17143845]
6. van Oorschot D, Vroling H, Bunge E, Diaz-Decaro J, Curran D, Yawn B. A systematic literature review of herpes zoster incidence worldwide. Hum Vaccin Immunother 2021 Jun 03;17(6):1714-1732 [FREE Full text] [doi: 10.1080/21645515.2020.1847582] [Medline: 33651654]
7. Kawai K, Gebremeskel BG, Acosta CJ. Systematic review of incidence and complications of herpes zoster: towards a global perspective. BMJ Open 2014 Jun 10;4(6):e004833 [FREE Full text] [doi: 10.1136/bmjopen-2014-004833] [Medline: 24916088]
8. Tricco AC, Zarin W, Cardoso R, Veroniki AA, Khan PA, Nincic V, et al. Efficacy, effectiveness, and safety of herpes zoster vaccines in adults aged 50 and older: systematic review and network meta-analysis. BMJ 2018 Oct 25;363:k4029 [FREE Full text] [doi: 10.1136/bmj.k4029] [Medline: 30361202]
9. Chiyaka ET, Nghiem VT, Zhang L, Deshpande A, Mullen PD, Le P. Cost-effectiveness of herpes zoster vaccination: a systematic review. Pharmacoeconomics 2019 Feb;37(2):169-200. [doi: 10.1007/s40273-018-0735-1] [Medline: 30367401]
10. Szucs TD, Pfeil AM. A systematic review of the cost effectiveness of herpes zoster vaccination. Pharmacoeconomics 2013 Feb;31(2):125-136. [doi: 10.1007/s40273-012-0020-7] [Medline: 23335045]
11. Pan CX, Lee MS, Nambudiri VE. Global herpes zoster incidence, burden of disease, and vaccine availability: a narrative review. Ther Adv Vaccines Immunother 2022;10:25151355221084535 [FREE Full text] [doi: 10.1177/25151355221084535] [Medline: 35340552]
12. Lu PJ, Hung MC, Srivastav A, Grohskopf LA, Kobayashi M, Harris AM, et al. Surveillance of vaccination coverage among adult populations -United States, 2018. MMWR Surveill Summ 2021 May 14;70(3):1-26 [FREE Full text] [doi: 10.15585/mmwr.ss7003a1] [Medline: 33983910]
13. Lin J, Wood JG, Bernardo C, Stocks NP, Liu B. Herpes zoster vaccine coverage in Australia before and after introduction of a national vaccination program. Vaccine 2020 Apr 29;38(20):3646-3652. [doi: 10.1016/j.vaccine.2020.03.036] [Medline: 32245643]
14. Liu XC, Simmonds KA, Russell ML, Svenson LW. Herpes zoster vaccine (HZV): utilization and coverage 2009 - 2013, Alberta, Canada. BMC Public Health 2014 Oct 23;14:1098 [FREE Full text] [doi: 10.1186/1471-2458-14-1098] [Medline: 25342438]
15. Kizmaz M, Kumtepe Kurt B, Çetin Kargin N, Döner E. Influenza, pneumococcal and herpes zoster vaccination rates among patients over 65 years of age, related factors, and their knowledge and attitudes. Aging Clin Exp Res 2020 Nov;32(11):2383-2391. [doi: 10.1007/s40520-019-01423-z] [Medline: 31776859]
16. Lu X, Lu J, Zhang F, Wagner AL, Zhang L, Mei K, et al. Low willingness to vaccinate against herpes zoster in a Chinese metropolis. Hum Vaccin Immunother 2021 Nov 02;17(11):4163-4170 [FREE Full text] [doi: 10.1080/21645515.2021.1960137] [Medline: 34370590]
17. Litt J, Kim S, Woodman R, MacIntyre R, Cunningham T. Australian zoster study: GP and patient views about herpes zoster (shingles), its complications, and the likely acceptance of a zoster vaccine (Zostavax). Int J Infect Dis 2014 Apr;21:436-437 [FREE Full text] [doi: 10.1016/j.ijid.2014.03.1320]
18. Lam AC, Chan M, Chou H, Ho S, Li H, Lo C, et al. A cross-sectional study of the knowledge, attitude, and practice of patients aged 50 years or above towards herpes zoster in an out-patient setting. Hong Kong Med J 2017 Aug 7;23(4):365-373 [FREE Full text] [doi: 10.12809/hkmj165043] [Medline: 28684648]
19. Cascini F, Pantovic A, Al-Ajlouni Y, Failla G, Ricciardi W. Attitudes, acceptance and hesitancy among the general population worldwide to receive the COVID-19 vaccines and their contributing factors: A systematic review. EClinicalMedicine 2021 Oct;40:101113 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101113] [Medline: 34490416]

20. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
21. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Epidemiology 2007 Nov;18(6):800-804. [doi: 10.1097/EDE.0b013e3181577654] [Medline: 18049194]
22. Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. J Epidemiol Community Health 2013 Nov 01;67(11):974-978. [doi: 10.1136/jech-2013-203104] [Medline: 23963506]
23. Freeman MF, Tukey JW. Transformations related to the angular and the square root. Ann Math Stat 1950;21:607-611 [FREE Full text] [doi: 10.1214/aoms/1177729756]
24. World health statistics 2011. World Health Organization. 2011. URL: https://www.who.int/docs/default-source/gho-documents/world-health-statistic-reports/en-whs2011-full.pdf [accessed 2023-02-21]
25. Data: World Bank country and lending groups. World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2023-02-21]
26. Prematunge C, Corace K, McCarthy A, Nair RC, Pugsley R, Garber G. Factors influencing pandemic influenza vaccination of healthcare workers--a systematic review. Vaccine 2012 Jul 06;30(32):4733-4743. [doi: 10.1016/j.vaccine.2012.05.018] [Medline: 22643216]
27. Batista Ferrer H, Audrey S, Trotter C, Hickman M. An appraisal of theoretical approaches to examining behaviours in relation to human papillomavirus (HPV) vaccination of young women. Prev Med 2015 Dec;81:122-131 [FREE Full text] [doi: 10.1016/j.ypmed.2015.08.004] [Medline: 26314783]
28. Wang Q, Yang L, Jin H, Lin L. Vaccination against COVID-19: a systematic review and meta-analysis of acceptability and its predictors. Prev Med 2021 Sep;150:106694 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106694] [Medline: 34171345]
29. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
30. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. Stat Med 2003 Jul 15;22(13):2113-2126. [doi: 10.1002/sim.1461] [Medline: 12820277]
31. Del Signore C, Hemmendinger A, Khanafer N, Thierry J, Trépo E, Martin Gaujard G, et al. Acceptability and perception of the herpes zoster vaccine in the 65 and over population: a French observational study. Vaccine 2020 Aug 18;38(37):5891-5895. [doi: 10.1016/j.vaccine.2020.07.004] [Medline: 32703748]
32. MacDougall DM, Halperin BA, MacKinnon-Cameron D, Li L, McNeil SA, Langley JM, et al. The challenge of vaccinating adults: attitudes and beliefs of the Canadian public and healthcare providers. BMJ Open 2015 Sep 29;5(9):e009062 [FREE Full text] [doi: 10.1136/bmjopen-2015-009062] [Medline: 26419683]
33. Baalbaki NA, Fava JP, Ng M, Okorafor E, Nawaz A, Chiu W, et al. A community-based survey to assess knowledge, attitudes, beliefs and practices regarding herpes zoster in an urban setting. Infect Dis Ther 2019 Dec;8(4):687-694 [FREE Full text] [doi: 10.1007/s40121-019-00269-2] [Medline: 31583578]
34. Lu P, Euler GL, Jumaan AO, Harpaz R. Herpes zoster vaccination among adults aged 60 years or older in the United States, 2007: uptake of the first new vaccine to target seniors. Vaccine 2009 Feb 05;27(6):882-887. [doi: 10.1016/j.vaccine.2008.11.077] [Medline: 19071175]
35. Al-Khalidi T, Genidy R, Almutawa M, Mustafa M, Adra S, Kanawati NE, et al. Knowledge, attitudes, and practices of the United Arab Emirates population towards herpes zoster vaccination: A cross-sectional study. Hum Vaccin Immunother 2022 Nov 30;18(5):2073752 [FREE Full text] [doi: 10.1080/21645515.2022.2073752] [Medline: 35622980]
36. Nicholls LAB, Gallant AJ, Cogan N, Rasmussen S, Young D, Williams L. Older adults' vaccine hesitancy: psychosocial factors associated with influenza, pneumococcal, and shingles vaccine uptake. Vaccine 2021 Jun 11;39(26):3520-3527 [FREE Full text] [doi: 10.1016/j.vaccine.2021.04.062] [Medline: 34023136]
37. Yang TU, Cheong HJ, Song JY, Noh JY, Kim WJ. Survey on public awareness, attitudes, and barriers for herpes zoster vaccination in South Korea. Hum Vaccin Immunother 2015;11(3):719-726 [FREE Full text] [doi: 10.1080/21645515.2015.1008885] [Medline: 25785481]
38. Qiu J, Sun XD, Hu JY, Huang ZY, Guo X, Liang XF. Willingness to receive herpes zoster vaccine and factors influencing willingness among 50-yearold adults of Shanghai in May-June 2020. Chin J Vaccines Immun 2021;27(3):307-310.
39. Lau J, Lee S. Impedance of vaccination against influenza, pneumococcus and zoster among older adults in Hong Kong: a qualitative study. Int J Infect Dis 2020 Dec;101:480 [FREE Full text] [doi: 10.1016/j.ijid.2020.09.1256]
40. Funovits AL, Wagamon KL, Mostow EN, Brodell RT. Refusal of shingles vaccine: implications for public health. J Am Acad Dermatol 2012 Jun;66(6):1011-1012. [doi: 10.1016/j.jaad.2011.11.931] [Medline: 22583716]
41. Wang Q, Yue N, Zheng M, Wang D, Duan C, Yu X, et al. Influenza vaccination coverage of population and the factors influencing influenza vaccination in mainland China: a meta-analysis. Vaccine 2018 Nov 19;36(48):7262-7269. [doi: 10.1016/j.vaccine.2018.10.045] [Medline: 30340886]

42. Isenor JE, Edwards NT, Alia TA, Slayter KL, MacDougall DM, McNeil SA, et al. Impact of pharmacists as immunizers on vaccination rates: a systematic review and meta-analysis. Vaccine 2016 Nov 11;34(47):5708-5723. [doi: 10.1016/j.vaccine.2016.08.085] [Medline: 27765379]
43. Possenti V, Luzi AM, Colucci A, De Mei B. Communication and basic health counselling skills to tackle vaccine hesitancy. Ann Ist Super Sanita 2019;55(2):195-199 [FREE Full text] [doi: 10.4415/ANN_19_02_12] [Medline: 31264643]
44. Glenton C, Carlsen B, Lewin S, Wennekes MD, Winje BA, Eilers R, VITAL Consortium. Healthcare workers' perceptions and experiences of communicating with people over 50 years of age about vaccination: a qualitative evidence synthesis. Cochrane Database Syst Rev 2021 Jul 20;7(7):CD013706 [FREE Full text] [doi: 10.1002/14651858.CD013706.pub2] [Medline: 34282603]
45. Vezzosi L, Riccò M, Agozzino E, Odone A, Signorelli C. Knowledge, attitudes, and practices of general practitioners from the Province of Parma (Northern Italy) towards vaccinations in adults ≥ 65 year-old. Acta Biomed 2019 Sep 13;90(9-S):71-75 [FREE Full text] [doi: 10.23750/abm.v90i9-S.8704] [Medline: 31517892]
46. Hurley LP, Allison MA, Dooling KL, O'Leary ST, Crane LA, Brtnikova M, et al. Primary care physicians' experience with zoster vaccine live (ZVL) and awareness and attitudes regarding the new recombinant zoster vaccine (RZV). Vaccine 2018 Nov 19;36(48):7408-7414 [FREE Full text] [doi: 10.1016/j.vaccine.2018.09.018] [Medline: 30420121]
47. Veronese N, Saccaro C, Demurtas J, Smith L, Dominguez LJ, Maggi S, et al. Prevalence of unwillingness and uncertainty to vaccinate against COVID-19 in older people: a systematic review and meta-analysis. Ageing Res Rev 2021 Dec;72:101489 [FREE Full text] [doi: 10.1016/j.arr.2021.101489] [Medline: 34662744]
48. Roller-Wirnsberger R, Lindner S, Kolosovski L, Platzer E, Dovjak P, Flick H, et al. The role of health determinants in the influenza vaccination uptake among older adults (65+): a scope review. Aging Clin Exp Res 2021 Aug 15;33(8):2123-2132 [FREE Full text] [doi: 10.1007/s40520-021-01793-3] [Medline: 33587270]
49. Carrico J, Talbird SE, La EM, Poston S, Poirrier JE, DeMartino JK, et al. Cost-benefit analysis of vaccination against four preventable diseases in older adults: impact of an aging population. Vaccine 2021 Aug 23;39(36):5187-5197 [FREE Full text] [doi: 10.1016/j.vaccine.2021.07.029] [Medline: 34334253]
50. Le P, Rothberg MB. Cost-effectiveness of the recommendations of the advisory committee on immunization practices for the recombinant adjuvanted zoster subunit vaccine. JAMA Intern Med 2018 Sep 01;178(9):1277-1278 [FREE Full text] [doi: 10.1001/jamainternmed.2018.3200] [Medline: 29987323]
51. Dunnsiri T, Kawashita T, Lee S, Monga A, Woo B. Assessing YouTube as an educational tool for shingles: cross-sectional study. JMIR Dermatol 2020;3(1):e20338 [FREE Full text] [doi: 10.2196/20338]
52. Li L, Wood CE, Kostkova P. Vaccine hesitancy and behavior change theory-based social media interventions: a systematic review. Transl Behav Med 2022 Feb 16;12(2):243-272 [FREE Full text] [doi: 10.1093/tbm/ibab148] [Medline: 34850217]
53. Chen C, Liu X, Yan D, Zhou Y, Ding C, Chen L, et al. Global influenza vaccination rates and factors associated with influenza vaccination. Int J Infect Dis 2022 Dec;125:153-163 [FREE Full text] [doi: 10.1016/j.ijid.2022.10.038] [Medline: 36328290]
54. Kan T, Zhang J. Factors influencing seasonal influenza vaccination behaviour among elderly people: a systematic review. Public Health 2018 Mar;156:67-78 [FREE Full text] [doi: 10.1016/j.puhe.2017.12.007] [Medline: 29408191]
55. Zhao S, Hu S, Zhou X, Song S, Wang Q, Zheng H, et al. The prevalence, features, influencing factors, and solutions of COVID-19 vaccine misinformation: a systematic review. JMIR Public Health Surveill 2023 Jan 11;9:e40201 [FREE Full text] [doi: 10.2196/40201] [Medline: 36469911]

Abbreviations

HBM: Health Belief Model
HCW: health care worker
HZ: herpes zoster
PHN: postherpetic neuralgia
Tdap: tetanus, diphtheria, and pertussis
VAX: Vaccination attitudes examination
VZV: varicella zoster virus
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 28.10.22; peer-reviewed by L Li, W Trzebinski; comments to author 07.12.22; revised version received 25.12.22; accepted 19.01.23; published 09.03.23.

Please cite as:
Wang Q, Yang L, Li L, Liu C, Jin H, Lin L
Willingness to Vaccinate Against Herpes Zoster and Its Associated Factors Across WHO Regions: Global Systematic Review and Meta-Analysis
JMIR Public Health Surveill 2023;9:e43893
URL: https://publichealth.jmir.org/2023/1/e43893
doi:10.2196/43893
PMID:36892937

Review

National Initiatives on Salt Substitutes: Scoping Review

Bingxuan Kong[1,2], PhD; Shanyue Yang[2], MPH; Jiewei Long[2], MPH; Yuhan Tang[2], PhD; Yang Liu[1], MPH; Zeng Ge[3], PhD; Shuang Rong[4], PhD; Yangfeng Wu[5], PhD; Gangqiang Ding[1,6], PhD; Yuexin Yang[1,7], MPH; Ping Yao[2,8*], PhD; Chao Gao[1,6*], PhD

[1]National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention, Beijing, China

[2]Department of Nutrition and Food Hygiene, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[3]Cardiovascular Health Program, Vital Strategies (USA) Jinan Representative Office, Jinan, China

[4]Department of Nutrition, School of Public Health, Wuhan University, Wuhan, China

[5]Peking University Clinical Research Institute, Peking University, Beijing, China

[6]Key Laboratory of Trace Element Nutrition of National Health Commission, Beijing, China

[7]Chinese Nutrition Society, Beijing, China

[8]Ministry of Education Key Laboratory of Environment, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Chao Gao, PhD
National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention
29 Nanwei Road
Beijing, 100050
China
Phone: 86 010 66237206
Email: gaochao20090901@163.com

Abstract

Background: Salt substitutes are edible salts that contain less sodium chloride owing to its partial substitution with other minerals, which serve as an important and effective intervention and public health strategy targeting hypertension and its secondary diseases, despite a small degree of controversy.

Objective: This study aimed to identify the current salt substitute initiatives in various nations and intergovernmental organizations (IGOs) worldwide and summarize their types and characteristics.

Methods: A scoping review was performed based on Arksey and O'Malley's framework and the latest guidance from the Joanna Briggs Institute. Searches were conducted on Google; government websites on food, health, and other related topics; PubMed; Web of Science; and Google Scholar from January 2022 to May 2022. Initiatives related to salt substitutes that were included in the study focused on the involvement of governments or IGOs through the publication of standards, actions, collaborations, funding, and so on. Data were extracted into Microsoft Excel (version 2019; Microsoft Corp) based on predefined items and analyzed using narrative synthesis and frequency count methods.

Results: A total of 35 initiatives from 11 countries (of which 9 are high-income countries) and 3 IGOs were identified. We classified all salt substitute initiatives into 5 types, namely benefit-risk assessments and cautions; plans and actions; regulations and standards; labels; and food reformulation, cooperation with the food industry, and media. More than half (18/35, 51%) of the salt substitute initiatives were launched within the past 5 years. Except for regulations and standards, salt substitute initiatives are, in general, part of the salt reduction framework. No nation or IGO has yet reported on the monitoring and implications of the use of salt substitutes.

Conclusions: Despite the limited number of salt substitute initiatives worldwide at present, a review on the different types and characteristics of such initiatives could be helpful in providing a reference for policy makers and stakeholders. Given the great potential of salt substitutes in improving hypertension and stroke, we call on more nations to pay attention to these substitutes and propose salt substitute initiatives in line with their national conditions.

(JMIR Public Health Surveill 2023;9:e45266) doi:10.2196/45266

KEYWORDS

salt substitute; nutrition policy; initiatives; strategies

Introduction

Background

Hypertension is a common chronic disease, accounting for approximately 11 million deaths, or approximately one-fifth of all deaths, in 2019 [1]. Hypertension and cardiovascular disease continue to be the leading cause of disability and death among patients with COVID-19 during the pandemic, who are affected by the lack of health care resources caused by the spread of the pandemic [2,3].

Salt reduction is one of the most important public health strategies for the prevention and control of hypertension. In China, a high-sodium diet was identified as the third most significant risk factor for both the number of deaths and the percentage of disability-adjusted life-years in 2017, accounting for over 1.5 million fatalities [4]. The World Health Organization (WHO) advocated for a 30% decrease in population salt intake as one of the primary interventions to reduce premature death from noncommunicable diseases [5]. Although the WHO recommends that adults consume <5 g of salt or 2000 mg of sodium per day, most people consume up to twice the recommended amount: an average of 9 to 12 g of salt per day [6]. A systematic review of global salt reduction initiatives shows that as of 2019, a total of 96 countries or regions have adopted different structural or regulatory approaches to salt reduction [7]. At the same time, the WHO also recommended that adults consume at least 3510 mg (90 mmol) of potassium per day [8]. Potassium has a blood pressure control effect [9], and potassium intake has been shown in many randomized controlled trials and modeling studies to be a good initiative for blood pressure and cardiovascular disease control [10,11]. In addition, sodium-potassium interaction is an important mechanism of hypertension [12]. However, the amount of potassium intake is generally not enough, such as the 1997 mg/day reported by the United States [13]. It was also reported that potassium intake and 24-hour urinary potassium excretion were less than half of the current recommendations in China [14,15].

As an emerging approach to salt reduction, salt substitutes are gaining attention in more and more countries. Salt substitutes or low-sodium salts usually refer to table or cooking salts that either do not contain sodium chloride or contain lower levels of sodium chloride owing to its partial replacement with potassium chloride, magnesium sulfate, or other minerals [16]. Salt substitutes have a similar taste to that of regular table salt. With an extra amount of potassium added, it serves as an ideal solution to reduce sodium and increase potassium and other mineral intake at the same time. Studies have shown the benefits of salt substitutes in controlling blood pressure and reducing the risk of cardiovascular disease and mortality [17]. A global environmental scan of salt substitution products showed that as of September 2020, a total of 87 salt substitutes were available in 47 countries worldwide [16]. However, many people still have concerns about the risk of adverse reactions associated with the intake of salt substitutes, such as hyperkalemia [18].

This Review

Salt reduction is one of the most important methods of preventing and controlling noncommunicable diseases in most countries [7]. Among others, salt substitute, as a new means of salt reduction, is gradually gaining attention and being popularized. However, there is no research summarizing and analyzing national salt substitute initiatives. In this context, the objectives of this scoping review were to (1) search for and document existing national salt substitute initiatives (before May 2022), (2) analyze the types and characteristics of national salt substitute initiatives, and (3) provide recommendations on the promotion and regulation of salt substitute applications in different regions and countries.

Methods

Approach

The scoping review was developed based on Arksey and O'Malley's [19] framework and the latest guidance from the Joanna Briggs Institute [20]. We reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) recommendations [21]. The PRISMA-ScR checklist is presented in Multimedia Appendix 1 [21].

Definition of Salt Substitutes and National Initiatives on Salt Substitutes

The expression of salt substitutes and the content of sodium varied across countries and regions for different products. Salt substitutes and low-sodium salts were considered synonymous in this study and defined as table salts or cooking salts whose sodium chloride was replaced with other minerals such as potassium chloride or magnesium sulfate, including sodium-reduced and sodium-free salts.

National initiatives on salt substitutes referred to initiatives related to salt substitutes that national governments and subsidiary organizations were engaged in through cooperation, development, publication, and implementation, as described in the *Eligibility Criteria* section. Nations in this study refer to the 193 United Nations (UN) member states [22]. Notably, intergovernmental organizations (IGOs) were organizations with international legal personalities established through treaty or other agreement between member states [23,24]. The main IGOs searched included UN agencies and regional organizations related to health, food, and agriculture [25] (all nations and IGOs searched in this study are listed in Multimedia Appendix 2).

Eligibility Criteria

The inclusion criteria were as follows: (1) participation of the national government (through funding, commissioning, the publication of standards, etc); (2) a national document or statement stressing on salt reduction and including the use of salt substitutes as a strategy; (3) a salt substitute program developed by the government in partnership with the food or salt industries; (4) a scheme about the labels, slogans, ingredient lists, etc printed on the package of salt substitute or the dietetic product with salt substitutes; and (5) mainstream media initiatives, consumer education, and other means used to increase knowledge on, improve attitudes toward, and behaviors related to salt substitutes.

The exclusion criteria were as follows: (1) salt substitute initiatives without the participation of nations or IGOs, for instance, initiatives led by nongovernmental organizations or local governments, and (2) the definition of salt substitutes differs from that used in this study, for example, flavor enhancers such as pepper and lemon.

Search Strategy

Three methods were used to conduct the research: government website search, Google search, and literature search. The search concepts included salt substitutes, countries, IGOs, portal function, government portals, states, and initiatives. Multimedia Appendix 2 lists the specific search process, all the keywords used to address the different aspects of the search concepts, and all the countries and IGOs involved. The government websites were obtained by searching the name of the 193 UN member states and the keywords related to government portals on Google Chrome with Boolean operators. First, keywords related to concepts such as government portals and portal function and country names were used to sort out government portal websites on health, food, medicine, and other related topics on Google Chrome, followed by a more in-depth exploration of keywords related to salt substitutes on these websites. With regard to IGOs, the official websites of a total of 66 IGOs related to global or regional political, economic, health, and agricultural cooperation were searched on Google Chrome. In addition, Google Chrome was directly used for keywords related to the concepts of salt substitutes, initiatives, nations, and IGOs with suitable Boolean operators. The first 5 pages of each Google search result were evaluated for eligibility, which usually contain almost all the information needed. All UN member countries were searched, with no language limitations set. If a country's official language was not English, the built-in translation service in Google Chrome was used to translate this language into English [26]. Finally, literature searches on PubMed, Web of Science, and Google Scholar were carried out using the corresponding search keywords and search formulas shown in Multimedia Appendix 2. The language of the literature search was restricted to English. All information was collected between January 2022 and May 2022.

Data Selection and Extraction

After receiving training, BK, SY, and JL individually conducted data identification and retrieval based on the established search strategy and eligibility criteria under the supervision of CG. The collected data was subsequently summarized. The checks during the summarization phase primarily focused on identifying and removing duplicates. Discrepancies were reviewed independently by CG, YT, YL, and PY and resolved through discussion. Data accuracy was independently checked by CG, PY, ZG, SR, YW, GD, and YY. The data eventually included in this study were extracted by BK and YL into a standardized table in a Microsoft Excel (version 2019; Microsoft Corp) sheet, which was independently reviewed by CG and PY and discussed in case of discrepancies. For government documents, the key information extracted included the (1) name of the document, (2) initiative enactment time, (3) the nation which published the document, (4) ministry that mandated the document, and (5) key content. From academic literature, we collected the following information: (1) the type of literature (article, review, letter, etc), (2) the year of publication, (3) the methods used, (4) the study population, and (5) relevant results.

Data Analysis

Data analysis was mainly performed through narrative synthesis and frequency count [27]. The research team collated and categorized the information obtained, summarizing and analyzing them through words and text. Data were also categorized based on the date of the publication of the initiatives, the WHO regional categorization, and the countries' income levels in the 2022 World Bank classification. Software used included Microsoft Excel 2019, ArcMap (version 10.8; Environment Systems Research Institute), and GraphPad Prism (version 8.0.1; Graph Pad Software, Inc).

Ethical Considerations

This study did not involve human participants, human participant research ethics, or related secondary analyses; therefore, ethics approval and consent to participate were not required.

Results

Search Results

A total of 16,889 records were identified after removing duplicates. On the basis of titles and abstracts, 16,260 (92.28%) records were excluded. After full-text reading, of the remaining 629 records, 594 (94.4%) were excluded for the reasons listed in Figure 1, the PRISMA-ScR flow diagram. As a result, 35 initiatives from 11 countries and 3 IGOs were eligible. Multimedia Appendix 3 [28-63] contains all the national and IGO initiatives for salt substitutes that were included.

The 11 countries were the United Kingdom, Germany, Norway, Ireland, and Finland in Europe; the United States and Canada in North America; Singapore, China, and Australia in East Asia and the Pacific; and India in South Asia (Figure 2). Except for China, which is an upper-middle-income country, and India, which is a lower-middle-income country according to the recently updated (2022) World Bank income ranking, all countries are high-income countries [64]. The 3 IGOs are the European Union (EU), the Eurasian Economic Union (EAEU), and The Codex Alimentarius Commission (CAC). Salt substitute initiatives were divided into five types, namely (1) benefit-risk assessments and cautions; (2) plans and actions; (3) regulations and standards; (4) labels; and (5) food reformulation, cooperation with the food industry, and media, all of which are detailed in Multimedia Appendix 4.

The total number of countries classified as belonging to the given regions according to the WHO regional categorization is shown within parentheses. Some initiatives fall under several categories.

Figure 1. Flow diagram of salt substitute initiatives identification. IGO: intergovernmental organization.

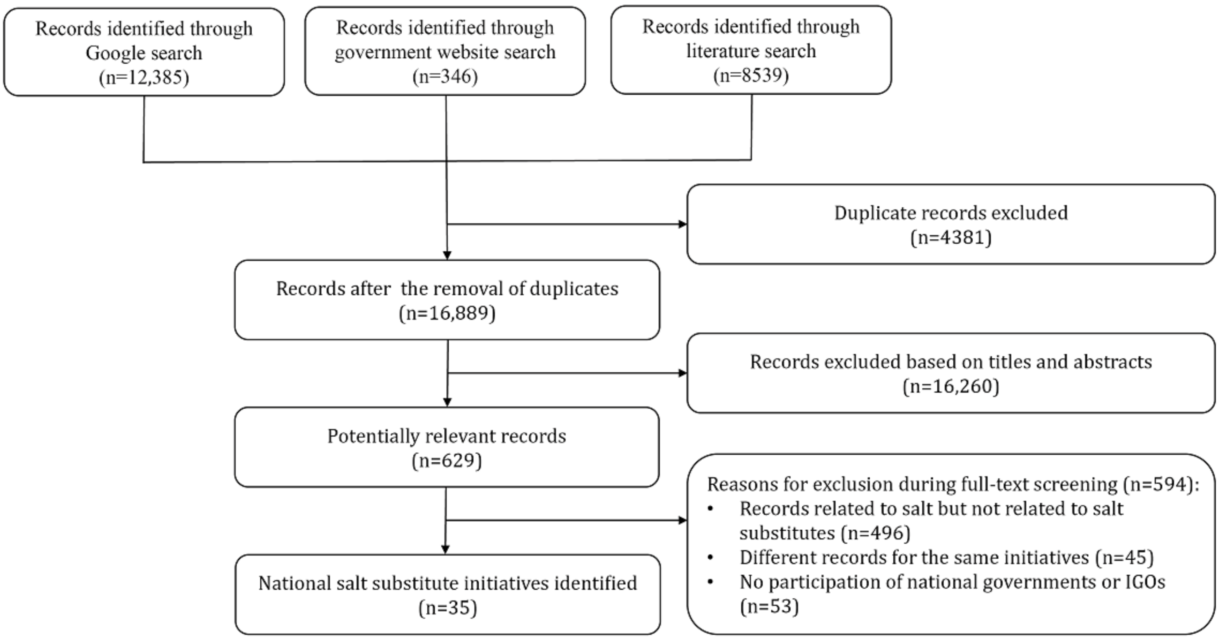
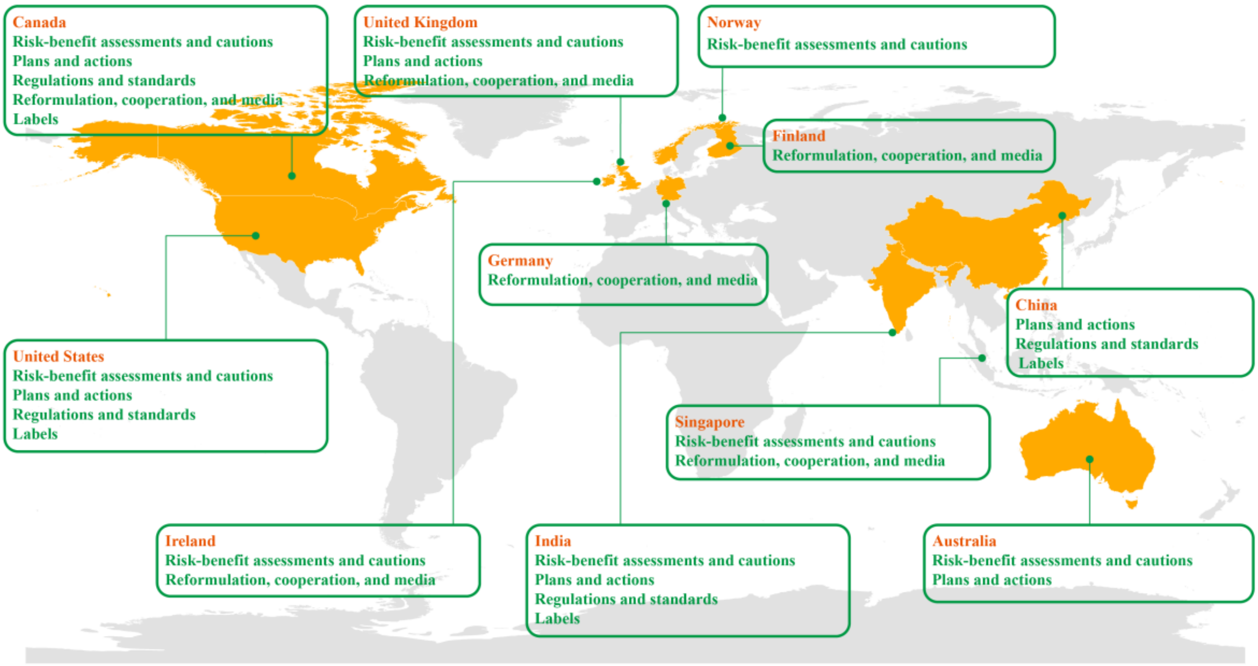


Figure 2. National salt substitute initiatives.



Types and Characteristics of National Salt Substitute Initiatives

Benefit-Risk Assessments and Cautions

The benefit-risk assessments of salt substitutes are usually carried out by a nutrition-related scientific committee appointed by the relevant government food safety department, providing recommendations on the use of salt substitutes. The benefit-risk assessments are an important part of determining national attitudes toward salt substitutes, which are a prerequisite for some countries to develop salt reduction–related initiatives. Assessments have been conducted in the United Kingdom [28], Ireland [29,30], Norway [31,32], and Australia [33].

In the United Kingdom, the Scientific Advisory Committee on Nutrition and the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment published *Potassium-Based Sodium Replacers: Assessment of the Health*

Benefits and Risks of Using Potassium-Based Sodium Replacers in Foods in the UK, which was an important evaluation of evidence for the application and promotion of salt substitutes [28]. The Scientific Advisory Committee on Nutrition and Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment applied the benefit-risk analysis for foods methodology to hypothesize the benefits and risks of reduced sodium intake and increased potassium intake in the general UK population by replacing 15% to 25% of sodium with potassium (potassium chloride, potassium carbonate, or potassium bicarbonate). This report recognized that the potential benefits of salt substitutes outweighed the potential risks and that although the impact at the individual level would be small, salt substitutes would have a beneficial effect on a large proportion of the UK population. The report recommended that the government encourage food businesses to use salt substitutes, but the level of potassium in foods and the types of foods in which salt substitutes are used needed to be monitored.

In addition to the United Kingdom, Norway conducted 2 assessments, one in 2014 and another in 2021. In 2014, the Norwegian Scientific Committee for Food and Environment, commissioned by the Norwegian Food Safety Authority and The Norwegian Directorate of Health, conducted a benefit-risk assessment by assuming the benefits and risks of 3 scenarios in which 30%, 50%, or 70% of sodium chloride was replaced by potassium, concluding that it was reasonable to anticipate that the percentage of persons likely to face an increased risk was far greater than the percentage of persons likely to benefit [31]. In 2021, the scientific committee was commissioned again to hypothesize the potential health effects of replacing 0% to 30% of sodium chloride with potassium chloride, and the report has not yet been published [32].

Ireland conducted *Salt and Health: Review of the Scientific Evidence and Recommendations for Public Policy in Ireland* in 2005 and 2016 (revision 1). In 2005, the scientific committee indicated that salt substitutes may increase susceptibility of the health of some populations (including those with type 1 diabetes, chronic renal insufficiency, end-stage renal disease, severe heart failure, or adrenal insufficiency) and did not contribute to lowering the population salt taste thresholds [29]. However, in 2016, the scientific committee concluded that the potassium intake at that time in the Irish diet was not enough and that the use of potassium-based salt replacement ingredients by the food industry could help supplement potassium intake among Irish people. However, the reduction of salt taste thresholds in the population was a major requirement for the food industry. Because the use of salt substitutes was not the preferred way to reduce the thresholds of salt taste, it should only be considered if reducing the sodium content independently could be detrimental to food safety or the physical or organoleptic properties of the food. Meanwhile, the scientific committee recommended that the Irish Food Safety Authority develop guidelines for the food industry regarding the use of potassium and other mineral salt replacements and make these guidelines available to all susceptible groups, that the sodium and potassium content of foods continue to be monitored, and that sodium and potassium intake be reassessed regularly [30].

The Australian Government Department of Health published *Food Reformulation: Risk Assessment* in 2020, which stated that salt substitutes did not necessarily bring about a change in population consumption patterns. In addition, potassium-based replacements could affect the health of susceptible people ("with major impairment of renal function because of chronic kidney disease or other morbidity, and those taking medications such as angiotensin-converting enzyme inhibitors and potassium-sparing diuretics that reduce renal excretion of potassium"). The report recommended that product monitoring and the monitoring of temporal trends in the incidence of hyperkalemia be performed when using salt substitutes and that clear labeling of relevant foods and sound physician recommendations be made for patients at risk [33].

To reduce the potential harm caused by salt substitutes for the higher-risk individuals mentioned earlier, some countries have proposed cautions for salt substitutes, including the recommendation that susceptible groups should use salt substitutes with caution or under professional supervision. For example, in Singapore, salt substitutes are included in the precautions for the use of potassium chloride [34] and the management of chronic heart failure [35]. In China, a warning label (shown in the *Labels* section) is required on the package of salt substitutes [36].

Despite the risk of hyperkalemia that potassium chloride may pose, since 1983, the US Code of Federal Regulations has announced that potassium chloride is generally recognized as safe as a direct human food ingredient when used in accordance with the current good manufacturing practice conditions of use (as a flavor enhancer, flavoring agent, nutrient supplement, pH control agent and stabilizer, or thickener) [37].

Plans and Actions

Plans and actions mainly refer to a series of guidelines and solution plans formulated and released by the government for specific problems, such as salt reduction, nutrition, and pollution. The use of salt substitutes in salt reduction plans or actions has been noted in 5 nations (the United Kingdom, Canada, Australia, China, and India).

Guidelines usually address 2 main aspects of salt substitute use. First, the processed food industry should consider reducing the salt content in their products to produce low-salt foods and then consider the use of salt substitutes when necessary (as proposed by the United Kingdom [38] and Canada [39]). Second, when using salt substitutes, attention should be paid to their safety and efficacy. Besides benefit-risk assessments, the disclosure of potassium content on nutrition labels, monitoring of potassium intake, and technological innovations are all possible approaches (as proposed by Canada [39] and Australia [40]). The focus of the salt reduction initiatives in China is different from that of the salt reduction initiatives in other countries, as the Chinese diet is predominantly homemade, especially in rural areas. Because of the high sodium intake, very low potassium intake, and high cardiovascular disease burden in these areas, the initiatives recommended adjusting the salt supply in the whole population while promoting and popularizing low-sodium and potassium-enriched salt. Baseline data collection and follow-up studies were conducted in a nationally representative population using dietary surveys and 24-hour urinary potassium excretion test [41]. In its book *Do You Eat Right*, the Food Safety and Standards Authority of India acknowledged the potential of salt substitutes for salt reduction but also noted that salt substitute brands were currently limited on the market and that salt substitutes should be used only under medical supervision [42].

In addition, although the US *Dietary Guidelines for Americans (2020-2025)* did not address potassium-rich salt substitutes, they stated that Americans' potassium intake was low and sodium intake should be limited and that the requirements for the nutrition fact label mandated the inclusion of the amount of potassium in packaged foods in the nutrition label [43].

Regulations and Standards

Regulations and standards provide reference for the production, packaging, trading, and use of salt substitutes and are issued by government departments to ensure safety and effectiveness. Regulations are mandatory and have legally binding effects. A total of 4 countries (the United States, Canada, China, and India; Table 1) have established regulations or standards for salt substitutes or have addressed salt substitutes in other regulations or standards.

First, the United States [44], India [45], and China [36] had clear but varied regulations on the composition of salt substitutes. For example, the commercial item description of salt substitutes developed by the US Department of Agriculture in 1997 divided salt substitutes into 2 categories, potassium chloride (type I) and potassium chloride with L-lysine (type II) [44]. China has established an industry standard (QB/*t* 2019-2020) on *low-sodium salt*, which stated that sodium chloride and potassium chloride should reach 65.0-80.0 g/100 g and 20.0-35.0 g/100 g on a dry basis, respectively [36]. In addition to active ingredients, these countries also imposed requirements on other ingredients, such as ingredients used for food grade bulking, anticaking agents, flavor enhancers, nutrition enhancers, and pollutants. For example, India not only requires salt substitutes to contain no more than 120 mg/100 g of sodium but also restricts the component of cations in some acidity regulators [45]. Canada has not developed standards for salt substitutes, but the *Labelling Requirements for Salt* issued by the Canadian Food Inspection Agency in 2019 stated that salt substitutes were generally sodium-reduced or sodium-free alternatives that usually contain potassium chloride. When a salt substitute was used as an ingredient in another food, it was an additive [46]. In addition, Goods and Services Identification Number of Salt Substitute released by the Canadian Food Inspection Agency classified salt substitutes as special dietary foods and food specialty preparations [47]. The regulations on salt substitute composition could be clearly seen in Table 1, whereas the regulations and standards on labeling are introduced in the next section.

Table 1. Regulations and standards for salt substitutes in countries.

Category and country	Title	Year	Department	Regulation and standard for salt substitutes
Regulations and standards of formula				
United States	Commercial Item Description—Salt Substitutes [44]	1997	The US Department of Agriculture	• Classification: the salt substitutes shall conform to the following list which shall be specified in the solicitation, contract, or purchase order: • Type I: potassium chloride • Type II: potassium chloride with L-lysine • All ingredients, including food grade bulking and anticaking agents, used in the preparation of the salt substitutes shall be of Food Chemicals Codex purity or US Pharmacopeia-National Formulary quality. The active ingredient in the salt substitutes shall be potassium chloride. Type II salt substitutes shall include L-lysine. The salt substitutes may contain flavor enhancers and anticaking agents such as, but not limited to monopotassium glutamate, glutamic acid hydrochloride, tricalcium phosphate, and calcium stearate.
China	Light Industry Standard of the People's Republic of China QB/*t* 2019-2020 Low Sodium Salt [36]	2020	Ministry of Industry and Information Technology of the People's Republic of China	• This standard specifies the sensory, physical, and chemical properties of; food additives used in; nutritional fortification of; and contaminants of salt substitutes. The dry basis of sodium chloride and potassium chloride should reach 65.0-80.0 g/100 g and 20.0-35.0 g/100 g, respectively.
India	Ministry of Health and Family Welfare (Food Safety and Standards Authority of India) Notification F. No. Stds/03/Notification (LS)/ FSSAI-2017 [45]	2017	Ministry of Health and Family Welfare (Food Safety and Standards Authority of India)	• Salt substitutes may contain:(a) Colloidal silica or calcium silicate: not more than 1% m/m of the salt substitute mixture, individually or in combination.(b) Diluents: safe and suitable nutritive foods as normally consumed namely, sugars, cereal flour. • The addition of iodine-containing compounds to salt substitutes shall be as per the Foods Safety and Standards Regulations, 2011. • The sodium content of salt substitutes shall be not more than 120 mg/100 g of the salt substitute mixture. • The composition of salt substitutes shall be as follows: • Potassium sulphate, potassium, calcium or ammonium salts of adipic; glutamic; carbonic; succinic; lactic; tartaric; citric; acetic; or hydro-, chloric, or orthophosphoric acids, and/or good manufacturing practice, except that phosphorus not to exceed 4% m/m and NH4+ 3% m/m of the salt substitute mixture • Magnesium salts of adipic; glutamic; carbonic; citric; succinic; acetic; tartaric; lactic; or hydro-, chloric, or orthophosphoric acids, mixed with other Mg-free salt substitutes as listed in 6.(1) (a), 6.(1) (c) and 6.(1) (d), and/or Mg++ to be not more than 20% m/m of the total of the cations K+, Ca++and NH4+ present in the salt substitute mixture and phosphorus not to exceed 4% m/m of the salt substitute mixture • Choline salts of acetic, carbonic, lactic, tartaric, citric, or hydrochloric acids, mixed with other choline-free salt substitutes as listed in 6.(1) (a), 6.(1) (b) and 6.(1) (d), and/or the choline content not to exceed 3% m/m of the salt substitute mixture • Free adipic, glutamic, citric, lactic, or malic acids—good manufacturing practice
Regulations and standards of labeling				
Canada	Labelling Requirements for Salt [46]	2019	Canadian Food Inspection Agency	• Salt substitutes do not have a prescribed standard and are generally a sodium reduced or sodium free alternative. Salt substitutes usually contain potassium chloride. Table salt substitutes, while not required to, are permitted to contain added iodine.
Regulations and standards of product classification				
Canada	Goods and Services Identification Number (GSIN)—Salt Substitute [47]	2018	Public Services and Procurement Canada	• GSIN[a] Category: Goods • GSIN Group Description: Subsistence • GSIN Class Description: Special Dietary Foods and Food Specialty Preparations

[a] GSIN: Goods and Services Identification Number.

Labels

Labels refer to the list of nutritional composition and health claims printed on packages, front-of-package labels, etc. Regulations on the labels of salt substitute products and the labels of packaged foods with added salt substitutes are part of the national strategy for salt substitutes to guide consumers in choosing or not choosing relevant products and are one of the important policy tools. Canada [46], the United States [48], Singapore [65], China [49], and India [45] have all regulated salt substitute labels (Table 2).

For salt substitutes, Canada required all the ingredients in the salt substitute to be labeled [46]. In China, salt substitute product labels must mention the potassium content and clearly indicate that salt substitutes should be used with caution by people for whom high potassium intake is not suitable, such as those who work at high temperatures, those who engage in high-intensity physical labor, those with kidney dysfunction, and those with hypertension who are taking antihypertensive medication [36,49].

For packaged foods containing salt substitutes, India required the labels to carry information on the presence of salt substitutes and the amount of potassium in them [45]. Because potassium chloride was not well known to consumers, research suggested that consumers may erroneously associate it with other chemicals and avoid consumption [66]. To change the negative public perception of potassium chloride and improve the health of the US population, NuTek Food Science, supported by food companies as well as public health organizations such as the Center for Science in the Public Interest and World Action on Salt & Health, submitted a citizen's petition proposing that the US Food and Drug Administration rename "potassium chloride" as "potassium salt" in food labels in 2016, and the proposal was passed after a public consultation in 2020 [48,67]. Thereafter, potassium chloride in packaged foods could be renamed to "potassium salt," but this is only a recommendation, not a mandatory responsibility, and is at the discretion of the food company. Canadian health promotion organizations such as Heart and Stroke and Hypertension Canada followed the lead of the United States and proposed to Health Canada and the Food Inspection Agency in 2019 to allow the use of the alternate name of "potassium chloride" [68]. The current Canadian law necessitates that the ingredient be identified by its common name along with "salt substitute" in parentheses. An example of this would be "potassium chloride (salt substitute)" [46]. Meanwhile, Canada has developed health claims for sodium and potassium, which can be made when processed foods meet the Food and Drugs Regulations' requirements, such as "a healthy diet containing foods high in potassium and low in sodium may reduce the risk of hypertension. [Name of the food] is a good source of potassium and is low in sodium." [50]. Health Canada's reforms to nutrition labeling has been considering the inclusion of potassium in the nutrition fact label [39].

In addition, Singapore's Healthier Choice Symbol (HCS) on packaged foods indicated to consumers which foods are healthier and more suitable choices for them [51]. The HCS label was found on approximately 4000 different foods and included 3 types of salt substitutes: Pagoda Less Sodium Mineral Salt, PanSalt, and GoodSalt [65].

Table 2. The regulations for the labels of salt substitute products and packaged foods with added salt substitutes.

Country	The regulation for labels
Salt substitute products	
Canada [46]	• All ingredients and their components must appear in the list of ingredients of table salt substitutes. This includes the declaration of iodide if present.
India [45]	• A declaration on the label as "low sodium salt substitute" or "low sodium dietetic salt" • A declaration on the label regarding the amount of cations, that is, sodium, potassium, calcium, magnesium, ammonium, and choline/100 g (m/m), in the salt substitute mixture
Packaged foods added salt substitutes	
United States [48]	• Exercise enforcement discretion for the declaration of "potassium salt" in the place of "potassium chloride" in the ingredient statement of food labels when potassium chloride is used as an ingredient in the food. • Potassium and sodium are listed on the nutrition fact label on packaged foods and beverages.
Canada [46]	• Salt substitute must be declared by its common name in the list of ingredients. The term "salt substitute" on its own would not be acceptable. However, it would be acceptable to declare the additive's function in brackets after the common name, for example, "potassium chloride (salt substitute)." • A salt substitute that meets the compositional and labeling criteria for a free of sodium or salt claim or for a low in sodium or salt claim may be represented as a food for special dietary use, such as "for salt free diets" or "for salt reduced diets."
China [49]	• "低钠盐的产品标签中应标示钾的含量，并应清晰标示：'高温作业者、重体力劳动强度者、肾功能障碍者及服用降压药物的高血压患者等不适宜高钾摄入的人群应慎用。'" • Translation: the product label of low-sodium salt should indicate the potassium content and should clearly indicate that "it should be used with caution by people who are not suitable for high potassium intake, such as those who work in high temperature, those who work with heavy physical strength, those who have kidney dysfunction and those who take antihypertensive drugs for hypertension."

Food Reformulation, Cooperation With the Food Industry, and Media

Many companies in the food industry are trying to reduce the amount of sodium in processed foods through food reformulation, that is, by replacing some or the entire amount of salt with salt substitutes and finding formulations that maintain good organoleptic, microbiological, physical, and chemical properties and the safety and quality of food. Some reformulations were carried out by professional technicians or companies commissioned or supported by the government. The German Federal Ministry of Food and Agriculture produced healthy and safe sodium-reduced fish [52,53], semihard cheese [54,55], raw and salted meat products [56], etc.

In other countries such as Canada and Singapore, the government encourages food companies to use salt substitutes as a means of reducing salt in processed foods. In Singapore, in addition to the HCS labeling, the government is also running the Healthier Ingredient Promotion Scheme based on the HCS, which supports a variety of food companies, including salt producers, to innovate and market healthier and more locally flavored products with government funding that covers (1) marketing and publicity, (2) trade promotion, and (3) ingredient thematic promotion [57]. The salt selected for the program was IMI Lifestyle Products' "GoodSalt." The sodium content in GoodSalt was replaced by essential minerals such as potassium, magnesium, lysine, and iodine [69]. Finland is an example of countries that successfully promoted salt substitutes through the media, with a salt substitute called "Pansalt" being reported and promoted by mainstream media such as Helsingin Sanomat. With salt substitutes becoming well known in Finland, the media has significantly contributed to the realization of salt reduction in Finland [58].

Salt Substitute Initiatives From 3 IGOs

Although IGOs are not countries, the initiatives proposed by them are often an important basis for regional or global food trade to ensure safety and health and are consciously abided by member states. These initiatives can be used as a basis for national legislation [24,70]. Therefore, this study also collected these initiatives on salt substitutes as a supplement to national initiatives (Multimedia Appendix 5 [59-62]).

Regarding the classification of salt substitutes, unlike Canada, the EU's food additive standard stated that when table salt substitutes and minerals were used for flavor, mouthfeel, or nutritional purposes, they were considered substances and not food additives [59]. In addition, in terms of food reformation, the *Institut National de la Recherche Agronomique* commissioned by the EU produced sausages, in which 20% to 30% of sodium was replaced with potassium [60].

The CAC and the EAEU regulated the composition and labeling of salt substitutes and have similar requirements [61,62]. In terms of ingredients, they limited the maximum quantity of different substitutable ingredients; for example, they required "the choline content not to exceed 3% m/m of the salt substitute mixture." For the labels of salt substitutes, the CAC required the mention of the name "low sodium salt substitute" or "low sodium dietetic salt," whereas the EAEU required the mention of "the substitute of salt with low content of sodium" or "dietetic salt with low content of sodium." Both regulations required companies to declare on the product label the full list of ingredients and the amount of cations in a 100 g mixture. For dietetic products containing salt substitutes, the presence of salt substitutes should be declared. The total potassium content and cation content in 100 g of the product should also be stated.

Discussion

Principal Findings

In this review, a total of 35 initiatives on salt substitutes from 11 countries and 3 IGOs were assessed in detail. Several major types of salt substitute initiatives, namely benefit-risk assessments and cautions; plans and actions; regulations and standards; labels; and food reformulation, cooperation with the food industry, and media, were summarized.

These initiatives had some significant features. First, the focus on salt substitutes reflects an interest in salt reduction. For most of the national initiatives that proposed salt substitutes as a means of salt reduction in the food industry, reducing the salt content of food has been regarded as a prerequisite. It has been shown that human preferences for salt taste could be shaped, with taste buds gradually adapting to saltier or less salty taste as dietary salt intake increases or decreases, respectively [71]. By contrast, replacing common salt with salt substitutes does not achieve the effects of lowering the taste threshold and changing the consumption pattern of the population. Salt substitutes as a means of salt reduction would only be considered in food categories if salt reduction is difficult because of microbial food safety, functional, and taste reasons.

Second, the mention of "potassium salt" or "potassium chloride (salt substitute)" as an alternative to "potassium chloride" on packaged food labels may be a new possible strategy. The International Food Information Council Foundation conducted a web-based survey of 1000 consumers and found that compared with potassium salt, consumers had a more negative association with potassium chloride, and their familiarity with "potassium" was very low [66]. However, there are still some issues to consider before implementation. For example, whether a name change would lead to a massive and rapid introduction of potassium salt products by the food industry, resulting in the lack of consideration of food safety and increased risk to susceptible groups, and the need to ensure appropriate mandatory nutrition labels, warning labels, and front-of-pack labels for potassium should be considered. Therefore, the implementation may need to be coupled with consumer education and monitoring.

Third, high-income countries are more concerned about the application and promotion of salt substitutes in the food industry, which is the industry where salt is mainly used. Several countries have funded research on innovative foods prepared with salt substitutes and cooperated with enterprises to encourage them to produce and sell packaged foods prepared with reduced salt and salt substitutes. However, some food companies have expressed the presence of barriers. For example, the effects of salt substitutes on product quality and taste have led to the frequent failure of companies' research on reformulation with salt substitutes as well as increased costs. Companies have also expressed concerns about the consumer acceptability of salt substitutes [63]. It is important to consider the complex industry chain that includes salt substitute manufacturers, processed food manufacturers, retailers, consumers, media, and many others [72]. All stakeholders should work together to implement locally adapted initiatives to promote salt substitutes, such as establishing public-private partnerships and promoting salt substitutes through subsidies, tax incentives, awareness campaigns, promotions, and consumer education.

In middle-income countries, such as China, the source of sodium chloride intake is more likely to be household salt, especially in rural areas. A large randomized controlled trial in a Chinese population of >20,000 older people from 600 villages with a history of stroke or hypertension demonstrated that salt substitutes significantly reduced the risk of stroke, major cardiovascular events, and death (rate ratios=0.86, 0.87, and 0.88, respectively) [73]. Moreover, if all Chinese households were to switch to salt substitutes, 450,000 lives could be saved each year [10]. However, there are still some issues with the use of salt substitutes in households. For example, the possibility of adding more salt substitutes for taste demand; the bitter taste of potassium chloride, which may discourage some people with sensitive tastes from choosing it; the need to consider the balance between salt substitute intake and iodine intake in areas where iodine intake is inadequate; and a lack of recommendations or standards for the use of salt substitutes in the household preparation of some foods, such as cured meats, salted fish, and spicy cabbages, which may pose microbiological risks.

Fourth, there is a controversy among the benefit-risk assessments carried out in different countries with regard to the percentage of sodium expected to be replaced by potassium. In the United Kingdom, the assumption of replacing 15% to 25% of sodium with potassium gave an overall positive conclusion [28]. However, in Norway, assumptions of 30%, 50%, and 70% substitutions translated into potentially higher-risk conclusions in 2014 [31]. A new assessment in Norway in 2021 changed the assumption to 0% to 30% substitution of sodium with potassium [32]. Considering the different demographic profiles, sodium and potassium intakes of the population, average blood pressure, and epidemiology of cardiovascular disease in various countries, this study recommends the establishment of appropriate assumptions and benefit-risk assessments based on national realities.

In fact, benefit-risk assessments have always been focused on the possible harm of the use of potassium in salt substitutes. People with limitation of renal potassium excretion, such as patients with renal failure; those with diabetes mellitus; and those using potassium-sparing diuretics, angiotensin-converting enzyme inhibitors, or angiotensin receptor blockers, may experience cardiac arrhythmia and have an increased level of death risk after consuming excessive potassium [74]. Several cases of hyperkalemia caused by the use of salt substitutes have been reported in high-risk participants [75,76]. Nevertheless, 2 clinical trials from China conducted salt substitute interventions among older people in 600 villages and 48 residential older adult care facilities. No significant increase in adverse clinical outcomes was found in both studies, although the second study found that the use of salt substitutes may result in more frequent hyperkalemia biochemically [73,77]. These studies provided high level of evidence on the safety of salt substitute when it comes to promotion. Furthermore, monitoring of the use of salt substitutes after national or regional promotion campaigns is very important, namely population consumption of salt substitutes, levels of sodium and potassium consumed and excreted, changes in blood pressure, cardiovascular disease and mortality, and cases of adverse effects.

Fifth, media campaigns are an important part of consumer education for salt reduction efforts. For example, in Finland, where salt reduction efforts began early and achieved good results, there was a long-standing media campaign on salt reduction with a strong promotion of salt substitute brands [58,78]. By contrast, a company in India with a wide range of salt products, including salt substitutes, became the subject of fake news stating that their products contained toxic ingredients. Widely circulated on social media, the fake news caused panic among consumers, only to be clarified with a statement from the government [79].

Sixth, there are currently no proposed measures on the price control and taxation of salt substitutes in various countries. As of September 2020, the price of salt substitutes ranged from US $0.46 to US $87.00 per kg worldwide, which is higher than the price of regular salt [16]. In non–high-income countries, lowering the price of salt substitutes is undoubtedly an important factor in promoting salt substitutes. In high-income countries, the cost of input is an equally important consideration in whether companies choose to reformulate high-sodium foods with salt substitutes. If countries wish to actively promote the use of salt substitutes, they may consider controlling the price by reducing the tax. As of 2019, a total of 5 countries around the world tax high-sodium foods [7], yet there are currently no policies to increase or reduce taxes on salt substitutes in any individual country. At the same time, governments could invest resources to support technology development aimed at reducing the cost of producing salt substitutes.

So far, there have been few reports issued by governments on the effects of the implementation of salt substitute policies, such as changes in sodium and potassium intakes, cardiovascular health, and potential risks. Most of the salt substitute initiatives were proposed in recent years. It is possible that such assessments of effects may have not been published or remain internal government documents. By contrast, governments usually reported the overall effect of salt reduction, including changes in salt intake and the salt content of foods [7]. Nevertheless, the process and results of policy development in one country have implications for other countries, thanks to the extensive communication among stakeholders. According to press reports, after food and health organizations in the United States submitted a petition to the Food and Drug Administration to use an alternate name of potassium chloride, stakeholders in Canada also took action. Eventually, they both contributed to the renaming of potassium chloride [68]. The benefit-risk assessment published in the United Kingdom also had a wide impact, being mentioned in published documents in both Norway [32] and Ireland [30].

On the basis of the classification of country regions by the World Bank and WHO, salt substitutes are of interest in high-income countries and European countries or organizations [64,80] (Multimedia Appendix 4). Despite the neutral, positive, or conservative attitudes toward salt substitutes shown in the published initiatives in different countries, the publication itself represents a concern for salt substitutes. With higher importance attached to salt reduction and health, salt substitutes will be circulated and popularized around the world. Therefore, we call on more countries to implement salt substitute initiatives. It is necessary for all countries to conduct benefit-risk assessments and develop regulations or standards for salt substitutes and their labeling according to the circumstances. In addition to this, depending on the source of dietary salt (homemade food or packaged foods), the inclusion of salt substitutes in national salt reduction initiatives, appropriate interventions, and consumer education are also important approaches.

This review has some strengths and limitations. We did not coordinate with country program leaders, global experts in salt substitutes, or regional WHO representatives. However, a literature search that covered the peer-reviewed and gray literature, Google search, and government website search were used to obtain as much information as possible from the web. Owing to the limitations and time lags in the information published on the web, it is possible that not all country-level salt substitute initiatives are properly included. In addition, instead of multilingual translators, Google Translate was used to identify and filter non-English information, which might have resulted in the omission and misinterpretation of relevant initiatives in non–English-speaking countries. Furthermore, this review focused on official national initiatives on salt substitutes and did not include stakeholders, academic institutions, or subnational recommendations and initiatives, such as the salt substitute promotion initiative of the China National Salt Group in Beijing [81]. These subnational initiatives also have high relevance to the application and promotion of salt substitutes in the world and in the corresponding region and are references for many salt substitute initiatives at the national level.

Conclusions

This scoping review has shown that although there are only a limited number of national initiatives focusing on salt substitutes at present, relevant initiatives are increasing year by year and involve multiple aspects, including regulations, guidelines, and cooperation with the industry. Given the great potential of salt substitutes in improving hypertension and stroke, we call on more countries to pay attention to and propose salt substitute initiatives. This requires the government to conduct a comprehensive evaluation and extensively cooperate with stakeholders, such as companies, media, and consumers.

Acknowledgments

This research was funded by Vital Strategies (United States) Jinan Representative Office; national financial project: Assessment and Application of Nutrients Requirement and Food Environment for Chinese Residents (131031107000210003); Chinese Nutrition Society Nutrition Science Foundation (CNS-YUM2019A10); and Danone Institute China Diet Nutrition Research and Communication Grant (DIC2020-06).

Data Availability

All data generated or analyzed during this study are included in this published paper (and in Multimedia Appendices 1 to 5).

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[DOCX File , 84 KB - publichealth_v9i1e45266_app1.docx]

Multimedia Appendix 2
Search strategies.
[DOCX File , 42 KB - publichealth_v9i1e45266_app2.docx]

Multimedia Appendix 3
The complete salt substitute initiatives from nations and intergovernmental organizations.
[DOCX File , 66 KB - publichealth_v9i1e45266_app3.docx]

Multimedia Appendix 4
Types of salt substitute initiatives.
[PNG File , 172 KB - publichealth_v9i1e45266_app4.png]

Multimedia Appendix 5
The salt substitute initiatives by intergovernmental organizations.
[DOCX File , 23 KB - publichealth_v9i1e45266_app5.docx]

References

1. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
2. Du Y, Zhou N, Zha W, Lv Y. Hypertension is a clinically important risk factor for critical illness and mortality in COVID-19: a meta-analysis. Nutr Metab Cardiovasc Dis 2021 Mar 10;31(3):745-755 [FREE Full text] [doi: 10.1016/j.numecd.2020.12.009] [Medline: 33549450]
3. Omboni S, McManus RJ, Bosworth HB, Chappell LC, Green BB, Kario K, et al. Evidence and recommendations on the use of telemedicine for the management of arterial hypertension: an international expert position paper. Hypertension 2020 Nov;76(5):1368-1383. [doi: 10.1161/HYPERTENSIONAHA.120.15873] [Medline: 32921195]
4. Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 Sep 28;394(10204):1145-1158 [FREE Full text] [doi: 10.1016/S0140-6736(19)30427-1] [Medline: 31248666]
5. Global action plan for the prevention and control of NCDs 2013–2020. World Health Organization. 2013 Nov 14. URL: https://www.who.int/publications/i/item/9789241506236 [accessed 2022-05-05]
6. Salt reduction. World Health Organization. 2020 Apr 29. URL: https://www.who.int/news-room/fact-sheets/detail/salt-reduction [accessed 2022-05-05]
7. Santos JA, Tekle D, Rosewarne E, Flexner N, Cobb L, Al-Jawaldeh A, et al. A systematic review of salt reduction initiatives around the world: a midterm evaluation of progress towards the 2025 global non-communicable diseases salt reduction target. Adv Nutr 2021 Oct 01;12(5):1768-1780 [FREE Full text] [doi: 10.1093/advances/nmab008] [Medline: 33693460]
8. Guideline: potassium intake for adults and children. World Health Organization. 2012 Dec 25. URL: https://www.who.int/publications/i/item/9789241504829 [accessed 2022-02-10]
9. Savica V, Bellinghieri G, Kopple JD. The effect of nutrition on blood pressure. Annu Rev Nutr 2010 Aug 21;30:365-401. [doi: 10.1146/annurev-nutr-010510-103954] [Medline: 20645853]
10. Marklund M, Singh G, Greer R, Cudhea F, Matsushita K, Micha R, et al. Estimated population wide benefits and risks in China of lowering sodium through potassium enriched salt substitution: modelling study. BMJ 2020 Apr 22;369:m824 [FREE Full text] [doi: 10.1136/bmj.m824] [Medline: 32321724]
11. Bernabe-Ortiz A, Sal Y Rosas VG, Ponce-Lucero V, Cárdenas MK, Carrillo-Larco RM, Diez-Canseco F, et al. Effect of salt substitution on community-wide blood pressure and hypertension incidence. Nat Med 2020 Mar;26(3):374-378 [FREE Full text] [doi: 10.1038/s41591-020-0754-2] [Medline: 32066973]
12. Adrogué HJ, Madias NE. Sodium and potassium in the pathogenesis of hypertension. N Engl J Med 2007 May 10;356(19):1966-1978. [doi: 10.1056/NEJMra064486] [Medline: 17494929]
13. Whelton PK. Sodium and potassium intake in US adults. Circulation 2018 Jan 16;137(3):247-249 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.117.031371] [Medline: 29335285]
14. Li Y, Zhang P, Wu J, Ma J, Xu J, Zhang X, et al. Twenty-four-hour urinary sodium and potassium excretion and their associations with blood pressure among adults in China: baseline survey of action on salt China. Hypertension 2020 Nov;76(5):1580-1588. [doi: 10.1161/HYPERTENSIONAHA.120.15238] [Medline: 32981359]
15. Du S, Wang H, Zhang B, Popkin BM. Dietary potassium intake remains low and sodium intake remains high, and most sodium is derived from home food preparation for Chinese adults, 1991-2015 trends. J Nutr 2020 May 01;150(5):1230-1239 [FREE Full text] [doi: 10.1093/jn/nxz332] [Medline: 31909790]
16. Yin X, Liu H, Webster J, Trieu K, Huffman MD, Miranda JJ, et al. Availability, formulation, labeling, and price of low-sodium salt worldwide: environmental scan. JMIR Public Health Surveill 2021 Jul 14;7(7):e27423 [FREE Full text] [doi: 10.2196/27423] [Medline: 33985938]
17. Greer RC, Marklund M, Anderson CA, Cobb LK, Dalcin AT, Henry M, et al. Potassium-enriched salt substitutes as a means to lower blood pressure: benefits and risks. Hypertension 2020 Feb;75(2):266-274. [doi: 10.1161/HYPERTENSIONAHA.119.13241] [Medline: 31838902]
18. Singal S, Parulkar A, Parulkar SS, Shin V. Watch what you eat: salt substitute causing life-threatening arrhythmia. J Am Coll Cardiol 2021 May;77(18):2042. [doi: 10.1016/s0735-1097(21)03398-2]
19. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
20. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Implement 2021 Mar;19(1):3-10. [doi: 10.1097/XEB.0000000000000277] [Medline: 33570328]
21. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]

22. Member states. United Nations. URL: https://www.un.org/en/about-us/member-states [accessed 2023-03-15]
23. Intergovernmental organizations (IGOs). Harvard Law School. URL: https://tinyurl.com/4p99da64/ [accessed 2023-03-13]
24. International organization. Wikipedia. URL: https://en.wikipedia.org/wiki/International_organization [accessed 2023-03-13]
25. List of intergovernmental organizations. Wikipedia. URL: https://en.wikipedia.org/wiki/List_of_intergovernmental_organizations [accessed 2023-03-13]
26. Change Chrome languages and translate webpages. Google Chrome Help. URL: https://support.google.com/chrome/answer/173424?hl=en&co=GENIE.Platform%3DDesktop&oco=0 [accessed 2022-01-02]
27. Soilemezi D, Linceviciute S. Synthesizing qualitative research: reflections and lessons learnt by two new reviewers. Int J Qual Methods 2018 Apr 11;17:1-14. [doi: 10.1177/1609406918768014]
28. Potassium-based sodium replacers: assessment of the health benefits and risks of using potassium based sodium replacers in foods in the UK. The Scientific Advisory Committee on Nutrition and the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/660526/SACN_COT_-_Potassium-based_sodium_replacers.pdf [accessed 2022-03-22]
29. Salt and health: review of the scientific evidence and recommendations for public policy in Ireland. The Scientific Committee of the Food Safety Authority of Ireland. 2005. URL: https://www.fsai.ie/uploadedFiles/Science_and_Health/salt_report-1.pdf [accessed 2022-03-22]
30. Salt and health: review of the scientific evidence and recommendations for public policy in Ireland (revision 1). The Scientific Committee of the Food Safety Authority of Ireland. 2016. URL: https://www.fsai.ie/science_and_health/salt_and_health/the_science_of_salt_and_health.html [accessed 2022-03-22]
31. Risk- benefit assessment of potassium chloride as replacement for sodium chloride. Scientific Steering Committee. 2014 Apr 24. URL: https://vkm.no/english/riskassessments/allpublications/riskbenefitassessmentofpotassiumchlorideasreplacementforsodiumchloride.4.27ef9ca915e07938c3b29bc0.html [accessed 2022-04-30]
32. Potassium chloride as substitute for salt - a benefit-risk assessment. The Norwegian Scientific Committee for Food and Environment. URL: https://vkm.no/english/riskassessments/allpublications/potassiumchlorideassubstituteforsaltabenefitriskassessment.4.20bfba7e17a31ff03a53dd7.html [accessed 2022-04-30]
33. Baines J. Food reformulation: risk assessment. Australian Government Department of Health and Aged Care. 2020 Jan. URL: https://www.health.gov.au/sites/default/files/documents/2021/03/partnership-reformulation-program-risk-assessment-report-food-reformulation-program-risk-assessment-report.pdf [accessed 2022-04-30]
34. Medication information leaflet-potassium chloride. Pharmaceutical Society of Singapore. URL: https://www.healthhub.sg/a-z/medications/368/Potassium-Chloride [accessed 2022-04-30]
35. Management of chronic heart failure. Ministry of Health Singapore. 2007. URL: https://www.moh.gov.sg/docs/librariesprovider 4/guidelines/management-of-chronic-heart-failure---booklet.pdf [accessed 2022-08-05]
36. Light Industry Standard of the People's Republic of China QB/T2019-2020 low sodium salt. Ministry of Industry and Information Technology of the People's Republic of China. 2020. URL: https://www.cssn.net.cn/cssn/productDetail/30482f65c42915ff132748edc92a8009 [accessed 2022-05-10]
37. Part 184 - direct food substances affirmed as generally recognized as safe. Code of Federal Regulations. 1977 Mar 15. URL: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-B/part-184 [accessed 2022-04-18]
38. Salt reduction: targets for 2024. Public Health England United Kingdom Government. 2020 Sep 7. URL: https://www.gov.uk/government/publications/salt-reduction-targets-for-2024 [accessed 2022-03-20]
39. Sodium reduction strategy for Canada - Recommendations of the Sodium Working Group. Health Canada. 2010 Jul. URL: https://publications.gc.ca/collections/collection_2010/sc-hc/H164-121-2010-eng.pdf [accessed 2022-04-25]
40. Partnership reformulation program – Implementation plan. Australian Government Department of Health and Aged Care. 2021 Apr 7. URL: https://www.health.gov.au/resources/publications/partnership-reformulation-program-implementation-plan [accessed 2022-04-30]
41. China's salt reduction action "2010-2020 action plan". Chinese Center of Disease Control. 2010 Jul 7. URL: https://www.docin.com/p-977861860.html [accessed 2022-05-25]
42. Do you eat right. The Food Safety and Standards Authority of India. 2020. URL: https://fssai.gov.in/upload/knowledge_hub/852185f89a7fc009c5Book_Do_You_Eat_Right_16_10_2020.pdf [accessed 2022-05-25]
43. Dietary guidelines for Americans 2020-2025. U.S. Departments of Agriculture and Health and Human Services. URL: https://www.dietaryguidelines.gov/ [accessed 2022-03-22]
44. Commercial item description: salt substitute. The U.S. Department of Agriculture (USDA). URL: https://www.ams.usda.gov/sites/default/files/media/CID%20Salt%20Substitute.pdf [accessed 2022-05-25]
45. F. No. Stds/03/Notification (LS)/FSSAI-2017. Food Safety and Standards Authority of India. 2017 Jun 19. URL: https://fssai.gov.in/upload/uploadfiles/files/Gazette_Notification_Special_Dietary_Food_22_06_2017.pdf [accessed 2023-08-17]
46. Labelling requirements for salt. Government of Canada. URL: https://inspection.canada.ca/food-labels/labelling/industry/salt/eng/1391790253201/1391795959629?chap=0 [accessed 2022-04-30]
47. Salt Substitute (replaced by gsin N8920ZC). Public Services and Procurement Canada. URL: https://buyandsell.gc.ca/procurement-data/goods-and-services-identification-number/gsin/N8940LA [accessed 2022-04-23]

48. Guidance for industry: the use of an alternate name for potassium chloride in food labeling. U.S. Food & Drug Administration. 2020 Dec. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-use-alternate-name-potassium-chloride-food-labeling [accessed 2022-03-22]
49. National standard of the People's Republic of China GB 2721-2015 national food safety standard edible salt. The Standardization Administration of the People's Republic of China. 2015. URL: https://www.cssn.net.cn/cssn/productDetail/009c1bf59b4bbf0dc137ccab0fed67bc [accessed 2022-05-25]
50. Guidance for the food industry on reducing sodium in processed foods. Government of Canada. 2012 Jun. URL: https://www.canada.ca/en/health-canada/services/food-nutrition/legislation-guidelines/guidance-documents/guidance-food-industry-reducing-sodium-processed-foods-2012.html#a33 [accessed 2022-05-08]
51. Healthier choice symbol. Health Promotion Board. URL: https://www.hpb.gov.sg/food-beverage/healthier-choice-symbol [accessed 2022-04-13]
52. Reduction of the salt content in selected fish products by the use of salt substitutes. FISA. URL: https://tinyurl.com/52pmdyth [accessed 2022-04-21]
53. Giese E, Meyer C, Ostermeyer U, Lehmann I, Fritsche J. Sodium reduction in selected fish products by means of salt substitutes. Eur Food Res Technol 2019 Apr 5;245(8):1651-1664. [doi: 10.1007/s00217-019-03277-1]
54. Strategies for salt reduction in semi-hard cheese. FISA. URL: https://tinyurl.com/3pupzxcz [accessed 2022-04-25]
55. Hoffmann W, Luzzi G, Steffens M, Clawin‐Rädecker I, Franz CM, Fritsche J. Salt reduction in film-ripened, semihard Edam cheese. Int J Dairy Technol 2019 Dec 08;73(1):270-282. [doi: 10.1111/1471-0307.12675]
56. Strategies for salt reduction in meat products (LOSS). FISA. URL: https://tinyurl.com/4j92c4zj [accessed 2022-04-20]
57. Healthier ingredient development scheme. Health Promotion Board. URL: https://www.hpb.gov.sg/healthy-living/food-beverage/healthier-ingredient-schemes/about-the-healthier-ingredient-promotion-scheme [accessed 2022-07-05]
58. Karppanen H, Mervaala E. Sodium intake and hypertension. Prog Cardiovasc Dis 2006;49(2):59-75. [doi: 10.1016/j.pcad.2006.07.001] [Medline: 17046432]
59. Regulation (EC) No 1333/2008 of the European Parliament and of the Council of 16 December 2008 on food additives (Text with EEA relevance). The European Union. 2008 Dec 31. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32008R1333 [accessed 2022-04-30]
60. TeRiFiQ project final report. CORDIS European Commission. URL: https://cordis.europa.eu/docs/results/289/289397/final1-final_report_terifiq_v1-0.pdf [accessed 2022-04-29]
61. Standard for special dietary foods with low-sodium content (including salt substitutes). Codex Alimentarius International Food Standards. URL: https://tinyurl.com/2fmk3ntx [accessed 2022-05-20]
62. Technical regulation of the Customs Union: on safety of specific types of specialized food products including the therapeutic and preventive dietary food. The Council of the Eurasian Economic Commission. 2012 Jun 15. URL: https://food.ec.europa.eu/system/files/2016-10/ia_eu-ru_sps-req_decision-34_annex_en.pdf [accessed 2022-04-28]
63. Appendix 1: summary of development of 2024 salt reduction targets. Public Health England. 2020 Sep. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/915402/2024_salt_reduction_targets_Appendix-1_070920-FINAL.pdf [accessed 2022-03-20]
64. World bank country and lending groups. The World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519 [accessed 2022-07-07]
65. HCS product list - by brand name (as of 31 March 2022). Health Promotion Board. URL: https://hpb.gov.sg/docs/default-source/default-document-library/hcs-website-list_as-of-31-march-2022.pdf?sfvrsn=1a353cbd_2 [accessed 2022-04-30]
66. Comment from International Food Information Council Foundation. Regulations.gov. URL: https://www.regulations.gov/comment/FDA-2019-D-0892-0056 [accessed 2022-05-16]
67. Citizen petition from NuTek food science LLC. Regulations.gov. 2016 Jun 28. URL: https://www.regulations.gov/document/FDA-2016-P-1826-0001 [accessed 2022-05-16]
68. Watson E. Health advocacy groups, top food brands,urge Health Canada to permit term 'potassium salt' on food labels. Food Navigator USA. 2019 May 6. URL: https://www.foodnavigator-usa.com/Article/2019/05/07/Health-advocacy-groups-top-food-brands-urge-Health-Canada-to-permit-term-potassium-salt-on-food-labels [accessed 2022-04-20]
69. Good salt. IMI Lifestyle. URL: https://www.imi.com.sg/products/goodsalt-low-sodium-cooking-salt [accessed 2022-04-27]
70. About codex alimentarius. Food and Agriculture Organization of the United Nations and World Health Organization. URL: https://www.fao.org/fao-who-codexalimentarius/about-codex/en/#c453333 [accessed 2022-04-30]
71. Micarelli A, Malacrida S, Strapazzon G, Mrakic-Sposta S, Micarelli B, Alessandrini N, et al. Impact of nutritional intervention on taste perception-a scoping review. Foods 2021 Nov 09;10(11):2747 [FREE Full text] [doi: 10.3390/foods10112747] [Medline: 34829027]
72. Hillman AJ, Hitt MA. Corporate political strategy formulation: a model of approach, participation, and strategy decisions. Acad Manage Rev 1999 Oct;24(4):825-842. [doi: 10.5465/amr.1999.2553256]
73. Neal B, Wu Y, Feng X, Zhang R, Zhang Y, Shi J, et al. Effect of salt substitution on cardiovascular events and death. N Engl J Med 2021 Sep 16;385(12):1067-1077. [doi: 10.1056/NEJMoa2105675] [Medline: 34459569]
74. Hunter RW, Bailey MA. Hyperkalemia: pathophysiology, risk factors and consequences. Nephrol Dial Transplant 2019 Dec 01;34(Suppl 3):iii2-iii11 [FREE Full text] [doi: 10.1093/ndt/gfz206] [Medline: 31800080]

75. Yip T, Wan W, Hui PC, Lui SL, Lo WK. Severe hyperkalemia in a peritoneal dialysis patient after consumption of salt substitute. Perit Dial Int 2012;32(2):206-208 [FREE Full text] [doi: 10.3747/pdi.2011.00198] [Medline: 22383720]
76. Ayach T, Nappo RW, Paugh-Miller JL, Ross EA. Life-threatening hyperkalemia in a patient with normal renal function. Clin Kidney J 2014 Feb;7(1):49-52 [FREE Full text] [doi: 10.1093/ckj/sft151] [Medline: 25859350]
77. Yuan Y, Jin A, Neal B, Feng X, Qiao Q, Wang H, et al. Salt substitution and salt-supply restriction for lowering blood pressure in elderly care facilities: a cluster-randomized trial. Nat Med 2023 Apr;29(4):973-981. [doi: 10.1038/s41591-023-02286-8] [Medline: 37055566]
78. Pietinen P, Männistö S, Valsta LM, Sarlio-Lähteenkorva S. Nutrition policy in Finland. Public Health Nutr 2010 Jun;13(6A):901-906 [FREE Full text] [doi: 10.1017/S1368980010001072] [Medline: 20513258]
79. Allegations against Tata Salt debunked by FSSAI and IMA. Food Safety and Standards Authority of India. 2019 Sep 5. URL: https://www.fssai.gov.in/upload/media/FSSAI_News_Allegation_Express_09_09_2019.pdf [accessed 2022-05-15]
80. World Health Organization homepage. World Health Organization. URL: https://www.who.int/ [accessed 2022-05-20]
81. "Low-sodium salt promotion action" will cover more than 200 communities in Beijing. Banyuetan. 2022 Mar 15. URL: http://www.banyuetan.org/qyzx/detail/20220315/1000200033137541647309880152591089_1.html [accessed 2022-05-20]

Abbreviations

CAC: Codex Alimentarius Commission
EAEU: Eurasian Economic Union
EU: European Union
HCS: Healthier Choice Symbol
IGO: intergovernmental organization
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
UN: United Nations
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 23.12.22; peer-reviewed by X Yin, K Kissock; comments to author 24.02.23; revised version received 29.04.23; accepted 15.06.23; published 17.11.23.

Please cite as:
Kong B, Yang S, Long J, Tang Y, Liu Y, Ge Z, Rong S, Wu Y, Ding G, Yang Y, Yao P, Gao C
National Initiatives on Salt Substitutes: Scoping Review
JMIR Public Health Surveill 2023;9:e45266
URL: https://publichealth.jmir.org/2023/1/e45266
doi:10.2196/45266
PMID:37435719

Original Paper

Preference and Willingness to Pay for the Regular COVID-19 Booster Shot in the Vietnamese Population: Theory-Driven Discrete Choice Experiment

Bach Xuan Tran[1,2*], PhD; Anh Linh Do[3*], MSc; Laurent Boyer[2], MD, PhD; Pascal Auquier[2], MD, PhD; Huong Thi Le[1], MD, PhD; Minh Ngoc Le Vu[4], BSc; Trang Huyen Thi Dang[1], BSc; Khuy Minh Cao[1], BSc; Linh Dieu Thi Le[1], BSc; Lam Tung Ngoc Cu[1], BSc; Bang Viet Ly[1], BSc; Duong Anh Thi Nguyen[1], BSc; Manh Duc Nguyen[1], BSc; Carl A Latkin[5], PhD; Roger C M Ho[6,7], PhD; Cyrus S H Ho[7], PhD; Melvyn W B Zhang[8], PhD

[1]Institute for Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam
[2]Research Centre on Health Services and Quality of Life, Aix Marseille University, Marseille, France
[3]SC Johnson College of Business, Cornell University, Ithaca, NY, United States
[4]Institute of Health Economics and Technology, Hanoi, Vietnam
[5]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
[6]Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
[7]Institute for Health Innovation and Technology (iHealthtech), National University of Singapore, Singapore, Singapore
[8]Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, Singapore
*these authors contributed equally

Corresponding Author:
Bach Xuan Tran, PhD
Institute for Preventive Medicine and Public Health
Hanoi Medical University
1 Ton That Tung Street, Kim Lien Ward, Dong Da District
Hanoi, 100000
Vietnam
Phone: 84 888288399
Email: bach.ipmph@gmail.com

Abstract

Background: The COVID-19 booster vaccination rate has declined despite the wide availability of vaccines. As COVID-19 is becoming endemic and charges for regular booster vaccination are being introduced, measuring public acceptance and the willingness to pay for regular COVID-19 boosters is ever more crucial.

Objective: This study aims to (1) investigate public acceptance for regular COVID-19 boosters, (2) assess the willingness to pay for a COVID-19 booster shot, and (3) identify factors associated with vaccine hesitancy. Our results will provide crucial insights into and implications for policy response as well as the development of a feasible and effective vaccination campaign during Vietnam's waning vaccine immunity period.

Methods: A cross-sectional study was conducted among 871 Vietnamese online participants from April to August 2022. An online questionnaire based on the discrete choice experiment (DCE) design was developed, distributed using the snowball sampling method, and subsequently conjointly analyzed on the Qualtrics platform. A history of COVID-19 infection and vaccination, health status, willingness to vaccinate, willingness to pay, and other factors were examined.

Results: Among the participants, 761 (87.4%) had received or were waiting for a COVID-19 booster shot. However, the willingness to pay was low at US $8.02, and most participants indicated an unwillingness to pay (n=225, 25.8%) or a willingness to pay for only half of the vaccine costs (n=222, 25.4%). Although information insufficiency and a wariness toward vaccines were factors most associated with the unwillingness to pay, long-term side effects, immunity duration, and mortality rate were the attributes the participants were most concerned with during the vaccine decision-making period. Participants who had children less than 18 years old in their homes infected with COVID-19 had a lower willingness to pay (odds ratio [OR] 0.54, 95% CI 0.39-0.74). Respondents who had children under 12 years old in their family who received at least 1 vaccine dose had a higher willingness to pay (OR 2.03, 95% CI 1.12-3.66). The burden of medical expenses (OR 0.33, 95% CI 0.25-0.45) and fear of the vaccine (OR 0.93, 95% CI 0.86-1.00) were negative factors associated with the level of willingness to pay.

Conclusions: A significant inconsistency between high acceptance and a low willingness to pay underscores the role of vaccine information and public trust. In addition to raising awareness about the most concerning characteristics of the COVID-19 booster, social media and social listening should be used in collaboration with health professionals to establish a 2-way information exchange. Work incentives and suitable mandates should continue to encourage workforce participation. Most importantly, all interventions should be conducted with informational transparency to strengthen trust between the public and authorities.

(JMIR Public Health Surveill 2023;9:e43055) doi:10.2196/43055

KEYWORDS

COVID-19; epidemic; vaccine; booster; willingness to take; willingness to pay; Vietnam; policy; feasibility; acceptability; infection; vaccination; social media; intervention

Introduction

The COVID-19 pandemic has profoundly affected numerous facets of life worldwide in the past 3 years, ranging from health, well-being, and socioeconomic status to the behavioral pattern and structural characteristics of societies [1-3]. Given its effectiveness in reducing virus severity and transmission, the COVID-19 vaccine has been used as the key pandemic control measure [4,5]. According to the World Health Organization (WHO), the COVID-19 vaccination decreased the number of deaths and severe disease from COVID-19, as well as decreasing COVID-19 transmission [6]. The Centers for Disease Control and Prevention (CDC) revealed that the protection by COVID-19 vaccination is up to 90% [7]. In Vietnam, the COVID-19 vaccine was highly accepted by the general population, including high-risk groups, such as pregnant women and the elderly [8,9]. As of July 2022, 82.2% of the Vietnamese population had received at least 2 doses of the COVID-19 vaccine [10]. However, vaccine effectiveness is dependent on the individual acceptance of the vaccine as well as the type of vaccine received, lasting on average up to a year with infection-induced antibodies and up to 6 months with vaccine-induced antibodies [11-14]. The resurgence of COVID-19 can be caused by several factors, including the decline in immunity to the COVID-19 vaccine, which emphasizes the importance of booster shots. Devenport et al [15] revealed that every 108 days or so, persons inoculated against COVID-19 would lose almost half of their protective antibodies [15]. Hence, immunizations that provided 90% protection against mild episodes of illness may only be 70% effective after 6 or 7 months [15]. Immunological studies have also shown a steady decline in antibody levels in people who have been vaccinated against COVID-19, and suggest an increased risk of an infection outbreak [16,17].

Given the effectiveness of COVID-19 booster shots, the current health focus of COVID-19 worldwide has shifted from disease control to booster regularity and long-term maintenance [18-21]. As COVID-19 progresses to its endemic phase, a pattern observed in previous major pandemics, the disease will likely not be eradicated but instead will be controlled by regular vaccinations [22,23]. Therefore, the promotion of booster vaccination is a key intervention in the new response scheme. However, unlike mandatory doses, the current COVID-19 booster shots record relatively low acceptance, even in the biggest COVID-19 vaccine–supplying countries, such as India (55.9%) [24] or the United States (61.8%) [23]. China's contribution to the COVID-19 vaccine cannot be ignored. Although China is among the top countries with the highest acceptance of the COVID-19 vaccine in the world [25], the acceptance to take COVID-19 booster shots rate has dropped to only 76.8% [26]. In Vietnam, this rate is below 60% compared to nearly 95% coverage for mandatory shots [10]. Many countries have also experienced a recent halt in booster administration rates due to concerns about their long-term effects and importance. In the coming endemic phase, the current vaccine policy in Vietnam will undergo major revisions, the most concerning of which is the transition from free-of-charge vaccines to with-charge and regular vaccination. Therefore, assessing the willingness of the population to pay for COVID-19 boosters is crucial to inform policy makers and public health practitioners on how to proceed with and promote vaccination costs.

Several frameworks have strived to describe the correlations between different factors for the acceptance of vaccines, most commonly the integrated framework of the Health Belief Model (HBM) and the theory of reasoned action (TRA) [27,28]. Indicators in this model have been consulted, such as the perception of severity and benefits, as well as the limitations of vaccination, attitude, and subjective norms, during previous pandemics [29]. During the COVID-19 outbreak in Malaysia, Ng et al [30] added contextual factors, such as susceptibility, trust, and vaccine preference, to improve the applicability of this framework to COVID-19 vaccine acceptance [30]. Integrated models were perfected through pandemics and effectively used to inform COVID-19 vaccine campaigns. However, the acceptance of initial shots and booster shots differs immensely in that multiple doses of COVID-19 boosters are required over time and depend on individual progress instead of single or 2-dose mass public vaccinations. Thus, conventional vaccine acceptance models cannot be easily applied to understand booster shot hesitancy [31].

Moreover, the effectiveness and side effects of booster shots have been well documented. Therefore, one's decision to receive a booster is influenced by a personal preference for known vaccine characteristics rather than a hypothetical evaluation of the vaccine's severity, benefits, or public attitude, as suggested by previous frameworks. This change in the nature of predictors calls for a new approach to understanding vaccine acceptance:

to investigate personal priorities for vaccine characteristics and assess the extent to which one is willing to take or pay in acknowledgment of risks.

Although the initial acceptance of the COVID-19 vaccine has been well studied for 3 years, little is known about the public's willingness to pay due to the free-of-charge policy for the COVID-19 vaccine in most countries worldwide and in Vietnam [26,32,33]. However, given current economic damages, it is unlikely that COVID-19 boosters can be provided under the same policy. Thus, the measurement of the willingness to pay for a COVID-19 booster has become a priority to inform an important transition of the vaccine implementation approach. We conducted this study during Vietnam's waning vaccine immunity period to (1) investigate public acceptance for regular COVID-19 boosters, (2) assess the willingness to pay for a COVID-19 booster shot, and (3) identify factors associated with vaccine hesitancy. Our results will provide crucial insights into and implications for policy response as well as the development of a feasible, effective, and sustainable vaccination campaign.

Methods

Study Design

The PREVENT (Preference For Vaccine Evaluation & Trail) study was Vietnam's nationwide assessment using a patient-centered design to inform health technology development and acceleration. Designed using the Qualtrics system, the PREVENT study included an online interactive questionnaire to target subjects across different regions of Vietnam from April to August 2022. Respondents were Vietnamese who live in Vietnam, who were aged 16 years and above (in Vietnam, people from the age of 16 years must take responsibility for their actions; moreover, they must bear penal liability like adults), and who were being referred to and agreed to complete the survey. The PREVENT study included 2 groups: (1) the general population and (2) health professionals and medical students. Snowball sampling was used for disseminating the survey, involving 20 "seeders" in all 3 regions: northern, middle, and southern. Participants took 30 minutes to complete the questionnaire and were encouraged to introduce more acquaintances and colleagues to the survey. Respondents were able to review and change their answers through a Back button in the Qualtrics system.

The PREVENT study included 5 topics of interest: (1) COVID-19 vaccine booster for adults, (2) monkeypox, (3) COVID-19 vaccine for children, (4) HIV vaccine, and (5) a hypothetical pandemic in the future. After answering general social demographic questions, each respondent was randomly assigned to 1 of these 5 topics, generating 5 separate data sets. In total, 5700 respondents were included in the PREVENT study. The substudy of the COVID-19 vaccine booster for adults included 871 complete records.

Ethical Considerations

Participants were informed of the benefits and risks of participating in this study, and they provided informed consent. Records were monitored and tracked using IP addresses using the Qualtrics system to ensure the validity of the data set; they were then extracted, analyzed, and stored safely and confidently and used merely for research purposes. The protocol was approved by ethical review committees designated by the Vietnam Ministry of Health (decision no. 164/GCN-HDDDNCYSH-DHYHN and 13/HDDD-DHDT).

Measurement and Instrument

We applied a standard procedure for generating the research instrument. Initially, a systematic review was conducted to identify important facets that emerged from previous studies. Next, we constructed the questionnaire, covering the breadth of measurements of interest. A group of experts in public health, infectious diseases, health services, econometrics, linguistics, representatives of target groups, and research assistants joined in the deliberation of translating, rephrasing, piloting, and shortening the questionnaire. Finally, the tool included 5 major sessions: (1) sociodemographics, (2) a history of COVID-19 infection and vaccination, (3) the willingness to take a COVID-19 vaccine booster for adults, and (4) the willingness to pay for it.

Outcome Variables

Willingness to Vaccinate

A short questionnaire was designed to determine the participants' willingness to take the COVID-19 vaccine booster for adults using the following 3 items:

- Will not vaccinate
- Intend to inject COVID-19 booster shots
- Have completed the COVID-19 booster shots

Willingness to Pay for the Vaccine

A questionnaire was designed to determine the participants' willingness to pay for vaccines using the following 5 options:

To what extent are you willing to pay for the vaccine?

- Unwilling to pay
- 20% of the cost
- 50% of the cost
- 80% of the cost
- Full cost

The actual cost of the vaccines was not known to participants. This question was used to examine the participants' willingness to pay when financial support/a discount was provided.

Socioeconomic Status

Participants responded to questions about their sociodemographics, including age, gender (male/female), marital status (single, others), job, monthly household income per capita, and area.

History of COVID-19 Infection and Vaccination

Participants responded to questions about their personal and family history of COVID-19, time since the last COVID-19 infection, health status, and personal and family history of COVID-19 vaccination.

Factors Affecting Vaccination to Prevent Disease

We used 10 items corresponding to 2 factors to measure the factors affecting vaccination to prevent disease. A 3-point Likert scale was used to evaluate the answers.

Factor 1: concerns about the vaccine and responsibility to the community

- Item 1. Concerned that the vaccine is newly developed.
- Item 2. Concerned about immediate side effects of the vaccine.
- Item 3. Concerned about long-term side effects of the vaccine.
- Item 4. Concerned about new components of the vaccine.
- Item 5. Concerned about the immunity duration of the vaccine.
- Item 9. Vaccinate so the community can maintain normal living and working conditions.
- Item 10. Vaccinate to fulfill personal responsibility of disease prevention.

Factor 1 included 7 questions. The total score was calculated by adding the scores of the 7 items and then converted to a 10-point scale.

Factor 2: fear of the vaccine

- Item 6. Fear of vaccines and injections in general.
- Item 7. Fear of insufficient information to make decisions.
- Item 8. Wait for others to vaccinate first.

Factor 2 included 3 questions. The total score was calculated by adding the scores of the 3 items and then converted to a 10-point scale.

Hence, the score of factors 1 and 2 ranged from 1 to 10. Participants with a higher score indicated a heightened concern about the vaccine and responsibility to the community and a heightened fear of the vaccine. The Cronbach α of factors 1 and 2 was good at .89 and .84, respectively.

Interpersonal Factors

We included 8 items to assess the interpersonal factors related to vaccination. A 3-point Likert scale was used to evaluate the answers. There were 2 domains included:

Factor 1: risks of infection and fear of the impact of the disease on health and economy

- Item 1. How is the risk of reinfection for you and your family members?
- Item 2. How is the risk of reinfection for children in your family?
- Item 3. How afraid are you of the spread of this disease?
- Item 4. Ho afraid are you of the impact on the health of this disease?
- Item 5. How afraid are you of the impact on the economy of this disease?

Factor 2: service satisfaction

- Item 6. How satisfied are you with vaccination services?
- Item 7. How satisfied are you with COVID-19 consultation services?
- Item 8. How satisfied are you with health care services for pandemics?

After summing the total score of each item, the total score of factors 1 and 2 was converted to a 10-point scale. Respondents with a higher score indicated a higher risk of infection and fear of the impact of the disease on health and economy or service satisfaction. The Cronbach α of factors 1 and 2 was good at .87 and .91, respectively.

Discrete Choice Experiment

We conducted a literature review of factors affecting the willingness to take and to pay for vaccination services and the designs of previous discrete choice experiments (DCEs) on adult vaccination. The results informed our selection of 6 major attributes of such services that influence an individual's preference, including effectiveness (<60%, 60%-90%, and >90%), immunity duration (3-6 months to a lifetime), side effects (minor [can maintain daily functioning] to major [severe fatigue and immobility]), mortality rate (0.001%-0.01%), limitations if not vaccinated (yes or no), cost of the vaccine (VND 100,000 [US $4.27] to VND 2,000,000 [US $85.33]) [34-39]. These attributes were then assigned 2-5 levels for choosing, contributing to a total of 31 × 41 × 23 × 51 = 480 possible alternatives (Table 1). Each participant was asked to respond to 7 different scenarios based on generated combinations by selecting which scenarios they preferred.

The sample size for the DCE was determined using Sawtooth software to determine the number of responses: sample size = (multiplier*c)/(t*a) = 357, where multiplier=1000, “c” is the largest number of levels across all features (n=5), “t” is the number of tasks or questions (n=7), and “a” is the number of alternatives or choices per question (n=2). This sample size was reached for both target groups of the survey, with 461 (52.9%) in the general population and 410 (47.1%) in health professionals and medical students.

Table 1. Vaccine attributes in the DCE[a].

Attributes	Response options
Effectiveness (%)	<60 60-90 >90
Immunity duration	3-6 months 6-12 months 1-3 years Lifetime
Side effects	Minor: can maintain daily functioning Major: severe fatigue and immobility
Mortality rate	0.001% 0.01%
Limitations if not vaccinated	No Yes: traveling banned and social gatherings restricted
Cost (VND/US $)[b]	100,000/4.27 200,000/8.53 500,000/21.33 1,000,000/42.67 2,000,000/85.33

[a] DCE: discrete choice experiment.

[b] VND 23,437.98=US $1.00.

Statistical Analysis

Statistical analysis was performed using Qualtrics and STATA software (version 15). Descriptive data were generated for all variables. With missing data, we used the listwise deletion method to clean data before analyzing them; 1071 records were collected, of which 871 (81.3%) records were complete, so 200 (18.7%) records were excluded from the analysis. Continuous variables were presented as the mean (SD), while categorical variables were presented as frequencies with percentages.

Factorial Structure

Exploratory factor analysis (EFA) using principal component analysis (PCA) was performed to evaluate the optimal structural model of the instrument according to the observed data. The number of factors was determined based on the Scree plot, and parallel analysis, along with eigenvalues and the proportion of variance explained (Multimedia Appendices 1 and 2). Items with a loading value of ≥0.5 were included in the relevant component.

Potential covariates for full models of the decision to take and the willingness to pay for COVID-19 boosters for adults included socioeconomic data, COVID-19 characteristics, related information regarding the COVID-19 vaccine, factors affecting the intention to vaccinate for disease prevention, and interpersonal factors (Multimedia Appendices 1 and 2). We used multivariate ordered logistic regression to identify factors related to the willingness to pay for COVID-19 booster and multinomial logistic regression to identify factors associated with the willingness to take a COVID-19 booster for adults. *P*<.05 was considered statistically significant.

In DCE data analysis, individual-based utility models were yielded using hierarchical Bayes estimation that uses Bayesian methods to probabilistically derive the relative value of each tested variable. The models estimated the optimal package of the preferred attributes of COVID-19 vaccination and the contribution of each attribute.

Results

Participant Characteristics

Table 2 demonstrates the demographic characteristics of the respondents. More than two-thirds of the participants (n=607, 69.7%) were females, and 661 (75.9%) were aged from 16 to 24 years. Most respondents finished college, university, postgraduate education (n=702, 80.6%) and were not married (n=707, 81.2%). Health care professionals and medical students (n=410, 47.1%) constituted the largest occupational group. The mean average monthly income per household was VND 3-10 million (US $128.00-$426.66; n=337, 38.7%).

Table 2. Demographic characteristics of respondents (N=871).

Characteristics	Participants, n (%)
Gender	
Male	264 (30.3)
Female	607 (69.7)
Age group (years)	
16-19	305 (35.0)
20-24	356 (40.9)
>25	210 (24.1)
Education level	
Not graduated from high school	95 (10.9)
Graduated from high school	74 (8.5)
College/university/postgraduate	702 (80.6)
Marital status	
Single/divorced/widowed	707 (81.2)
Married	164 (18.8)
Job	
Health care worker/medical student	410 (47.1)
Other students	190 (21.8)
Stable jobs	64 (7.3)
Other jobs	207 (23.8)
Monthly household income per capita (VND/US $)	
<1 million/<42.67	249 (28.6)
1-2.9 million/42.67-123.73	175 (20.1)
3-10 million/128.00-426.66	337 (38.7)
>10 million/>426.66	110 (12.6)
Children	
No children	710 (81.5)
Pregnant/have children	161 (18.5)
Area	
Hanoi	391 (45.0)
Other northern provinces/cities	102 (11.7)
Southern provinces/cities	101 (11.6)
Central provinces and Central Highlands	192 (22.1)
Other provinces	85 (9.6)

Table 3 presents the history of previous infections, vaccinations, and health characteristics of the respondents. Of 871 participants, 492 (56.3%) were infected with COVID-19 and 249 (28.5%) had children in the family infected with COVID-19. The most prevalent infection period was the recent 3-6 months. The largest proportion of participants (n=400, 46.1%) reported a moderate health status, followed by 359 (41.4%) who self-reported that they were completely healthy. Regarding the COVID-19 vaccination for self and family, two-thirds of the participants had taken the third, or booster, dose (n=582, 66.7%), while only 46 (5.3%) participants had children below 12 years old in their family who had been vaccinated. The most popular information sources were health care officials (n=592, 67.8%) and the media (n=457, 62.7%).

Table 3. Personal and family history of COVID-19 infection and vaccination of study subjects (N=871).

Characteristics	Participants, n (%)
Personal and family history of COVID-19	
Nobody in my family has been infected COVID-19.	128 (14.7)
I was infected with COVID-19.	492 (56.3)
Adults in my family were infected with COVID-19.	481 (55.1)
Children <18 years old in my family were infected with COVID-19.	249 (28.5)
Time since the last COVID-19 infection	
Not yet infected	381 (43.7)
1-3 months	154 (17.7)
3-6 months	283 (32.4)
>6 months	54 (6.2)
Health status	
Completely healthy (100%)	359 (41.4)
Relatively healthy (80%-100%)	400 (46.1)
Slightly compromised health (<80%)	109 (12.5)
History of COVID-19 vaccination of self and family	
I have received 2 doses.	213 (24.4)
I have received 3 doses or a booster or both.	582 (66.7)
Children under 12 years old in my family have received at least 1 vaccine dose.	46 (5.3)
Children 12-17 years old in my family have received at least 1 vaccine dose.	105 (12.0)
All adults in my family have received at least 2 doses.	436 (49.9)
Vaccine information sources	
Health care officials	592 (67.8)
Relatives, friends, neighbors	371 (42.5)
Media (health consultation switchboard, radio, newspaper, television)	547 (62.7)
Others	77 (8.8)

Willingness to Take and Willingness to Pay for a COVID-19 Booster

Table 4 and Multimedia Appendix 3 demonstrates the respondents' willingness to receive and their reasons for vaccine refusal. Nearly two-thirds of the participants received a booster shot (n=524, 60%), and only 61 (7%) participants did not want a booster vaccine. The most common reasons for vaccine hesitancy included insufficient time since the last shot (n=38, 34.6%) and the wait-and-see approach (n=31, 28.2%). In addition, 225 (25.8%) respondents were unwilling to pay for a COVID-19 booster, and 222 (25.4%) were willing to pay half of the vaccine cost (ie, the other half of the cost had to be covered by someone else). A considerable proportion (n=571, 65.6%) viewed the booster shot as an economic burden.

Table 4. Willingness to take and willingness to pay for a COVID-19 booster.

Characteristics	Participants, n (%)
Willingness to take a COVID-19 booster	
Will not vaccinate	61/871 (7.0)
Intend to inject COVID-19 booster shots	288/871 (33.0)
Have completed COVID-19 booster	524/871 (60.0)
Reasons for refusal to vaccinate	
Not enough time between COVID-19 shoots	38/109 (34.6)
Feeling unwell	9/109 (8.2)
Have underlying medical problems or other medical treatment	6/109 (5.5)
Allergic history	3/109 (3.6)
Am waiting for my turn to be vaccinated	14/109 (12.7)
Wait and see	31/109 (28.2)
No means of transport and money to travel to the place of injection	0
Suffered from severe side effects of the previous injection	17/109 (15.5)
Feel that booster is unnecessary as COVID-19 has become irrelevant	29/109 (26.4)
Have just been infected	28/109 (25.5)
Willingness to pay for a COVID-19 booster	
Unwilling to pay	225/871 (25.8)
20% of the cost	133/871 (15.2)
50% of the cost	222/871 (25.4)
80% of the cost	93/871 (10.7)
Full cost	200/871 (22.9)
Financial burden of COVID-19 booster	
No	300/871 (34.4)
Yes	571/871 (65.6)

Multivariable Analyses to Identify Factors Associated With the Decision to Take and the Willingness to Pay for a COVID-19 Vaccine

Table 5 reveals the factors associated with the decision to take the COVID-19 vaccine. People who had higher education (odds ratio [OR] 17.22, 95% CI 2.86-103.63), who had graduated from high school (OR 12.35, 95% CI 3.46-44.00), who had received 3 doses or a booster or both (OR 13.31, 95% CI 4.31-41.09), and in whose family all adults had received at least 2 doses (OR 2.88, 95% CI 1.31-6.31) were likely to have completed the COVID-19 booster. Relatively healthy participants were likely to intend to get the COVID-19 vaccine (OR 2.98, 95% CI 1.37-6.47).

Multivariate ordered logistic regression results to identify factors associated with the willingness to pay for the COVID-19 vaccine are presented in Table 6. Participants aged 20-24 years, who had a monthly household income per capita of >VND 10 million, in whose family adults were infected with COVID-19, in whose family children under 12 years old had received at least 1 vaccine dose, who used media to access information about the vaccine, and who had higher concerns about vaccine characteristics and responsibility to the community were likely to have a higher level of willingness to pay for the COVID-19 vaccine. Children <18 years old in the family infected with COVID-19 and the burden of medical expenses were harmful factors that reduced the level of willingness to pay for the COVID-19 vaccine.

Figure 1 presents the feature importance of each vaccine attribute generated by the conjoint analysis. The mortality rate had the most influence on the decision-making process, at 27 points, followed by immunity at 22.8 points and vaccine effectiveness at 22.4 points. The cost was the least weighted attribute, at 7.6 points.

Table 7 shows the willingness to pay based on current COVID-19 vaccine attributes: more than 90% effectiveness, 6-12 months’ immunity, minor side effects, insignificant mortality rate, and major limitations if not vaccinated. For this package, the generated willingness to pay was VND 188,000 (US $ 8.02).

Table 5. Factors associated with the decision to take the COVID-19 vaccine.

Factors	Willingness to get the vaccine	
	“Will not vaccinate” vs “have completed the COVID-19 booster,” OR (95% CI)	“Intend to inject” vs “have completed the COVID-19 booster,” OR (95% CI)
Education level (reference: not graduated from high school)		
Graduated from high school	0.06[a] (0.01-0.35)	0.20[b] (0.06-0.68)
College/university/postgraduate	0.08[a] (0.02-0.29)	0.21[a] (0.07-0.60)
City/province (vs Hanoi: reference)		
Northern provinces	0.69 (0.20-2.41)	1.04 (0.57-1.91)
Southern provinces	1.65 (0.60-4.55)	0.33[a] (0.16-0.68)
Central provinces and Central Highlands	0.72 (0.27-1.94)	0.70 (0.41-1.21)
Other provinces	0.19[c] (0.03-1.06)	0.59 (0.29-1.23)
Personal and family history of COVID-19		
Children <18 years old in my family were infected with COVID-19.	1.71 (0.76-3.81)	1.77 [b] (1.12-2.79)
Health status (vs completely healthy 100%: reference)		
Relatively healthy (80%-<100%)	0.50[c] (0.22-1.12)	1.48[c] (0.97-2.27)
Slightly compromised health	2.44[c] (0.92-6.45)	2.18[b] (1.17-4.05)
History of COVID-19 vaccination for yourself and your family (yes vs no: reference)		
I have received 2 doses.	7.03[a] (2.46-20.06)	5.76[a] (3.00-11.04)
I have received 3 doses or a booster or both.	0.08[a] (0.02-0.23)	0.14[a] (0.08-0.24)
All adults in my family have received at least 2 doses.	0.35[a] (0.16-0.76)	0.38[a] (0.24-0.60)

[a] *P*<.10.

[b] *P*<.05.

[c] *P*<.01.

Table 6. Factors associated with the willingness to pay for the COVID-19 vaccine.

Factors	Level of willingness to pay (1=unwilling to pay to 5=willing to pay the full cost), OR (95% CI)
Age (years; reference: 16-19 years)	
20-24	1.40[a] (1.03-1.91)
>25	1.20 (0.77-1.87)
Job (reference: health care worker/medical student)	
Other students	1.07 (0.77-1.50)
Stable jobs	1.41 (0.77-2.58)
Other jobs	1.70[b] (1.14-2.53)
Monthly household income per capita (VND/US $; reference: <VND 1 million/<US $42.67)	
1-2.9 million/42.67-123.73	0.88 (0.61-1.29)
3-10 million/128.00-426.66	1.22 (0.87-1.71)
>10 million/>426.66	2.36[b] (1.47-3.79)
Personal and family history of COVID-19 (vs nobody has been infected with COVID-19: reference)	
Adults in my family were infected with COVID-19.	1.33[a] (1.01-1.77)
Children <18 years old at my family were infected with COVID-19.	0.54[b] (0.40-0.74)
History of COVID-19 vaccination for yourself and your family (yes vs no: reference)	
Children under 12 years old in my family have received at least 1 vaccine dose.	1.94[a] (1.09-3.48)
Medical expense burden (yes vs no: reference)	
The burden of medical expenses	0.33[b] (0.25-0.45)
Resources to access information on vaccine (yes vs no: reference)	
Media (health consultation switchboard, radio, newspaper, television)	1.36[a] (1.03-1.80)
Factors affecting vaccination to prevent disease	
Concerns about vaccine characteristics and responsibility to the community (unit: score)	1.10[a] (1.01-1.20)
Fear of the vaccine (unit: score)	0.93[c] (0.86-1.00)
Interpersonal factors	
Service satisfaction (unit: score)	1.05 (0.98-1.13)

[a] *P*<.05.

[b] *P*<.10

[c] *P*<.01.

Figure 1. Feature importance of vaccine attributes during the decision-making process.

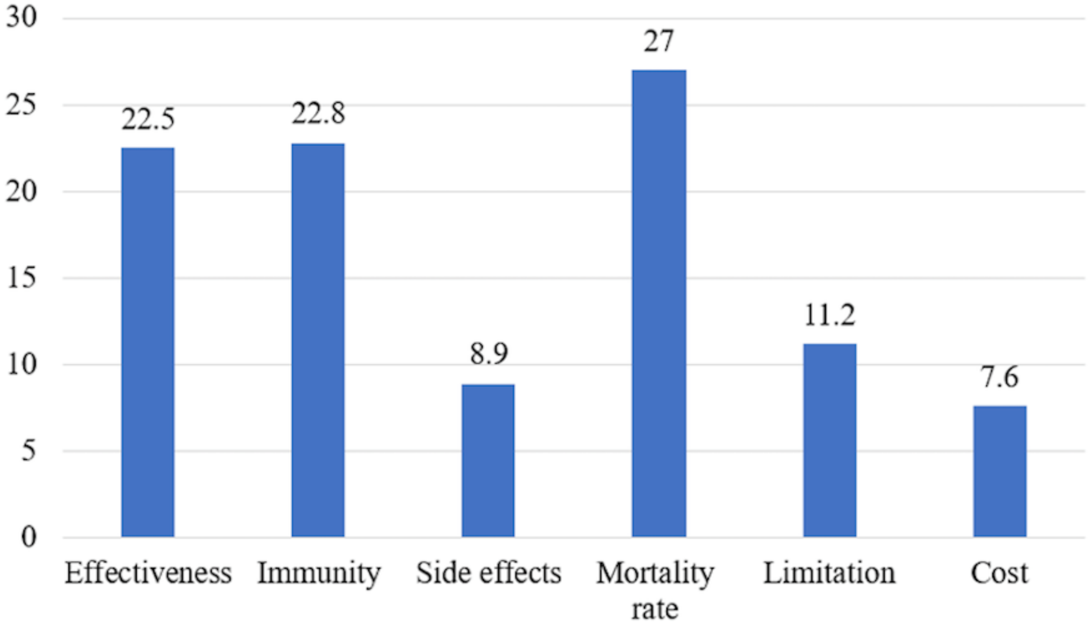


Table 7. Willingness to pay for a hypothetical COVID-19 vaccine booster for adults generated by the DCE[a].

Vaccine attribute	Value
Vaccine effectiveness	More than 90%
Immunity	6-12 months
Side effects	Minor, does not affect daily functioning
Mortality rate (/100,000 people)	1 death
Limitation if not vaccinated	Yes: traveling banned and social activity restricted
Cost (VND/US $)	188,000/8.02

[a] DCE: discrete choice experiment.

Discussion

Principal Findings

This study results indicated the participants' high willingness to take COVID-19 vaccine booster shots. Their willingness to pay, however, was scattered between price levels, the most common of which were willingness to pay the full price, willingness to pay half the price, and total unwillingness to pay at all. The reasons for an unwillingness to pay and booster hesitancy were concerns of long-term effects, immunity duration, information insufficiency, lack of confidence in the vaccine, unawareness of the risk of reinfection, and dissatisfaction with COVID-19 care services. Based on the identified issues of lack of information and weakening public trust in the vaccine, we proposed personal, interpersonal, and managerial solutions to ensure affordable vaccinations and consistent uptake. The burden of medical expenses and the fear of vaccines were negative factors associated with the willingness to pay; in contrast, participants who had children under 12 years old in their family who had received at least 1 vaccine dose used the media to access information about vaccines, while concerns about vaccine characteristics and responsibility to the community had the opposite impact.

In our study, 87% of the participants had taken a booster dose or were waiting to get vaccinated, which is generally higher than that in low-, middle-, and high-income countries (73.4% in low-income countries, 67.9% in middle-income countries, and 83.0% in high-income countries) [40].

When examined at the national level, the willingness to pay generated in our study accounted for a small fraction of the per capita monthly income of our sample (VND 3,000,000-10,000,000, or US $ 128.00-426.66). Since the effects of the COVID-19 vaccine typically start to diminish after 6 months, the actual monthly cost of a booster is even lower when spread out over months, making it a low financial burden. Indeed, a critically low willingness to pay despite a high acceptance rate might be a proxy for people’s underestimation of vaccine importance, suggesting gaps in communication between the general public and authorities. This is consistent with the fact that the most common reason for vaccine hesitancy was a lack of knowledge about the vaccine.

Public information about COVID-19 boosters does not cover the main factors highlighted by this study that influence the vaccine decision and willingness to pay, such as potential side effects, mortality rate, induced immunity, and effectiveness. Furthermore, as health care professionals and the media were deemed to provide the most trusted information, public education approaches should be more comprehensive and interactive, allowing for televised question-and-answer (Q&A) sessions with experts, hotline services staffed by trained practitioners, and open informational portals made available on social media. Social listening is another recommended approach to be used by governments and health organizations to understand people's perceptions and concerns, to refute false information and handle controversies, and to identify other social leaders or influencers who may have a positive impact on the general public.

Our findings emphasize the importance of building public trust when disseminating information about vaccinations. Contrary to traditional vaccines with a long history of development and trials, the COVID-19 vaccine was made available on an emergency warranty after preclinical studies and only 12-18 months of multiphase trials [41]. Despite ongoing efforts to map the long-term effects of the COVID-19 vaccine, the lack of longitudinal evidence leaves them unconfirmed to both clinicians and the general public. Although rarely, the COVID -19 vaccine can also have adverse side effects, such as myocarditis, Guillain-Barre syndrome, or thrombotic syndrome [42]. This information, if not communicated clearly and transparently, may cause public distrust, which in turn undermines the national sustainable vaccine uptake [43]. Making COVID-19 vaccinations a regular/seasonal requirement may also prove challenging in the absence of public transparency. During the COVID-19 outbreak, it was evident that public trust differed between the types of response, specifically between countries whose governments showcased true surveillance results and those whose governments intentionally provided incomplete information in an attempt to comfort the public [44,45]. The widespread availability and ease of use of social media platforms hastened the spread of false information online. This issue is further exacerbated if there exists low vaccine literacy in the country [46]. Public distrust, if it already exists, can be tackled through social listening, educating, and engaging with advocates, as in the case of the influenza outbreak [47].

Other significant factors promoting vaccine acceptance were fear of disrupting normal working and living conditions, reinfection, and re-emergence of the outbreak. In the work sphere, various interventions can be implemented to improve vaccine uptake, including paid time off for vaccinations or recovery time. A study among US workers observed that such incentives could improve vaccination rates, and they have been frequently implemented since the beginning of vaccination programs [48]. Financial plans that allow the cost of the COVID-19 vaccine to be deducted directly from one's salary and the payments to be spread over months should be encouraged to relieve the financial burden. The inclusion of the booster vaccine in the standard workplace health requirements or routine health check-ups must be endorsed by relevant institutions. Since COVID-19 has been identified as a massive public burden affecting the entire economy, the costs associated with its prevention could be shared by both public and private sectors.

Interventions drawing on the fear of a COVID-19 pandemic re-emergence are numerous. Nudge interventions, such as mobile phone personal reminders or fixated appointments, have been found to be effective in increasing vaccine uptake in Denmark [49]. In Vietnam, a new "Vaccine Reminder" extension can be added to the national COVID-19 portal, which was made available and effectively used for case surveillance during the pandemic peak. Restrictions regarding social gatherings and traveling should continue to be imposed to make boosters informally mandatory [50].

Strengths and Limitations

The strengths of our study included a large sample size and diverse demographics, which enabled high generalizability. Our DCE design and conjoint analysis allowed us to imitate real-life conditions of the willingness to pay and generate a more reliable result.

Numerous limitations also exist, including the inability to infer causal relationships due to a cross-sectional design, the influence of the current vaccine policy on respondents' perception, and personal bias due to self-reporting. Nonetheless, this study is 1 of the first and most significant in our country to provide early insights into the public acceptance and willingness to pay for COVID-19 boosters, thus making substantial contributions to Vietnam's efforts to integrate COVID-19 boosters into its regular health care requirements.

Conclusion

Our study showed a high willingness to take yet a low willingness to pay for a COVID-19 vaccine booster among Vietnamese participants. More importantly, the low willingness to pay was traced to insufficient information and underestimating vaccine importance. To resolve the lack of knowledge, we proposed a vaccine promotion approach with emphasis on side effects, mortality rate, immunity duration, and effectiveness, as well as social listening to counter misinformation. As the media and health experts are the most trusted information providers, their role in monitoring, directing, and resolving issues of public perception can be optimized through interactive interventions, such as Q&A sessions between influential health figures on television or an online platform, a COVID-19 vaccine–specialized hotline by health officials, and use of social listening insights. Work incentives, such as assistance for COVID-19 vaccine sick leave and technical convenience for vaccine reminders, can also be provided to encourage vaccine uptake. Above all, all vaccine promotion campaigns should be based on an open display of benefits as well as unwanted limitations of vaccines to achieve a nationally accepted regular COVID-19 vaccination in the long term.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Exploratory factor analysis about factors affecting vaccination to prevent disease.
[DOCX File , 14 KB - publichealth_v9i1e43055_app1.docx]

Multimedia Appendix 2
Exploratory factor analysis of interpersonal factors.
[DOCX File , 14 KB - publichealth_v9i1e43055_app2.docx]

Multimedia Appendix 3
Willingness to take and willingness to pay for a COVID-19 booster regarding 2 groups.
[DOCX File , 16 KB - publichealth_v9i1e43055_app3.docx]

References

1. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. Int J Soc Psychiatry 2020 Jun 31;66(4):317-320. [doi: 10.1177/0020764020915212] [Medline: 32233719]
2. Koon OE. The impact of sociocultural influences on the COVID-19 measures-reflections from Singapore. J Pain Symptom Manage 2020 Aug;60(2):e90-e92 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.04.022] [Medline: 32360238]
3. Arora T, Grey I. Health behaviour changes during COVID-19 and the potential consequences: a mini-review. J Health Psychol 2020 Aug 18;25(9):1155-1163. [doi: 10.1177/1359105320937053] [Medline: 32551944]
4. Andrews N, Stowe J, Kirsebom F, Toffa S, Rickeard T, Gallagher E. Covid-19 vaccine effectiveness against the Omicron (B1.1. 529) variant. New Engl J Med 2022;386(16):1532-1546. [doi: 10.1101/2022.03.22.22272691]
5. Chu D, Vu Ngoc S, Vu Thi H, Nguyen Thi Y, Ho T, Hoang V, et al. COVID-19 in Southeast Asia: current status and perspectives. Bioengineered 2022 Feb 26;13(2):3797-3809 [FREE Full text] [doi: 10.1080/21655979.2022.2031417] [Medline: 35081861]
6. World Health Organization. Statement for Healthcare Professionals: How COVID-19 Vaccines Are Regulated for Safety and Effectiveness (Revised March 2022). URL: https://www.who.int/news/item/17-05-2022-statement-for-healthcare-professionals-how-covid-19-vaccines-are-regulated-for-safety-and-effectiveness [accessed 2022-03-01]
7. Centers for Disease Control and Prevention. Benefits of Getting A COVID-19 Vaccine. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/vaccine-benefits.html [accessed 2022-08-17]
8. Huynh G, Nguyen TV, Nguyen DD, Lam QM, Pham TN, Nguyen HTN. Knowledge about COVID-19, beliefs and vaccination acceptance against COVID-19 among high-risk people in Ho Chi Minh City, Vietnam. IDR 2021 May;14:1773-1780. [doi: 10.2147/idr.s308446]
9. Nguyen LH, Hoang MT, Nguyen LD, Ninh LT, Nguyen HTT, Nguyen AD, et al. Acceptance and willingness to pay for COVID-19 vaccines among pregnant women in Vietnam. Trop Med Int Health 2021 Oct 23;26(10):1303-1313 [FREE Full text] [doi: 10.1111/tmi.13666] [Medline: 34370375]
10. Mathieu E, Ritchie H, Ortiz-Ospina E, Roser M, Hasell J, Appel C, et al. A global database of COVID-19 vaccinations. Nat Hum Behav 2021 Jul 10;5(7):947-953. [doi: 10.1038/s41562-021-01122-8] [Medline: 33972767]
11. Pilishvili T, Gierke R, Fleming-Dutra KE, Farrar JL, Mohr NM, Talan DA, et al. Effectiveness of mRNA Covid-19 vaccine among U.S. health care personnel. N Engl J Med 2021 Dec 16;385(25):e90. [doi: 10.1056/nejmoa2106599]
12. Katella K. How Long Will Your Coronavirus Vaccination Last?. URL: https://www.yalemedicine.org/news/how-long-will-coronavirus-vaccine-last [accessed 2023-01-16]
13. Jamshidi E, Asgary A, Shafiekhani P, Khajeamiri Y, Mohamed K, Esmaily H, et al. Longevity of immunity following COVID-19 vaccination: a comprehensive review of the currently approved vaccines. Hum Vaccin Immunother 2022 Nov 30;18(5):2037384 [FREE Full text] [doi: 10.1080/21645515.2022.2037384] [Medline: 35417285]
14. Centers for Disease Control and Prevention. Science Brief: SARS-CoV-2 Infection-Induced and Vaccine-Induced Immunity. URL: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html [accessed 2023-01-16]
15. Khoury DS, Cromer D, Reynaldi A, Schlub TE, Wheatley AK, Juno JA, et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. Nat Med 2021 Jul;27(7):1205-1211. [doi: 10.1038/s41591-021-01377-8] [Medline: 34002089]
16. Naaber P, Tserel L, Kangro K, Sepp E, Jürjenson V, Adamson A, et al. Dynamics of antibody response to BNT162b2 vaccine after six months: a longitudinal prospective study. Lancet Reg Health Eur 2021 Nov;10:100208 [FREE Full text] [doi: 10.1016/j.lanepe.2021.100208] [Medline: 34514454]
17. Thomas SJ, Moreira ED, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine through 6 months. N Engl J Med 2021 Nov 04;385(19):1761-1773 [FREE Full text] [doi: 10.1056/NEJMoa2110345] [Medline: 34525277]

18. Huebner ES, Suldo SM, Smith LC, McKnight CG. Life satisfaction in children and youth: empirical foundations and implications for school psychologists. Psychol Sch 2003 Dec 09;41(1):81-93. [doi: 10.1002/pits.10140]
19. Milas G, Martinović Klarić I, Malnar A, Saftić V, Šupe-Domić D, Slavich GM. The impact of stress and coping strategies on life satisfaction in a national sample of adolescents: a structural equation modelling approach. Stress Health 2021 Dec;37(5):1026-1034 [FREE Full text] [doi: 10.1002/smi.3050] [Medline: 33876570]
20. Moksnes UK, Løhre A, Lillefjell M, Byrne DG, Haugan G. The association between school stress, life satisfaction and depressive symptoms in adolescents: life satisfaction as a potential mediator. Soc Indic Res 2014 Dec 6;125(1):339-357. [doi: 10.1007/s11205-014-0842-0]
21. Suldo SM, Huebner ES. Does life satisfaction moderate the effects of stressful life events on psychopathological behavior during adolescence? Sch Psychol Q 2004;19(2):93-105. [doi: 10.1521/scpq.19.2.93.33313]
22. Our World in Data. COVID-19 Vaccine Boosters Administered per 100 People. URL: https://ourworldindata.org/grapher/covid-vaccine-booster-doses-per-capita?country=BGD~BRA~CHL~IND~ITA~PAK~RUS~SGP~USA~OWID_WRL [accessed 2023-01-16]
23. Yadete T, Batra K, Netski DM, Antonio S, Patros MJ, Bester JC. Assessing acceptability of COVID-19 vaccine booster dose among adult Americans: a cross-sectional study. Vaccines (Basel) 2021 Dec 02;9(12):1424 [FREE Full text] [doi: 10.3390/vaccines9121424] [Medline: 34960170]
24. Achrekar GC, Batra K, Urankar Y, Batra R, Iqbal N, Choudhury SA, et al. Assessing COVID-19 booster hesitancy and its correlates: an early evidence from India. Vaccines (Basel) 2022 Jun 30;10(7):1048 [FREE Full text] [doi: 10.3390/vaccines10071048] [Medline: 35891212]
25. Wang Q, Hu S, Du F, Zang S, Xing Y, Qu Z, et al. Mapping global acceptance and uptake of COVID-19 vaccination: a systematic review and meta-analysis. Commun Med (Lond) 2022 Sep 12;2(1):113 [FREE Full text] [doi: 10.1038/s43856-022-00177-6] [Medline: 36101704]
26. Wu F, Yuan Y, Deng Z, Yin D, Shen Q, Zeng J, et al. Acceptance of COVID-19 booster vaccination based on the protection motivation theory: a cross-sectional study in China. J Med Virol 2022 Sep 09;94(9):4115-4124 [FREE Full text] [doi: 10.1002/jmv.27825] [Medline: 35506329]
27. Fall E, Izaute M, Chakroun-Baggioni N. How can the health belief model and self-determination theory predict both influenza vaccination and vaccination intention? A longitudinal study among university students. Psychol Health 2018 Jun 05;33(6):746-764. [doi: 10.1080/08870446.2017.1401623] [Medline: 29132225]
28. Kan T, Zhang J. Factors influencing seasonal influenza vaccination behaviour among elderly people: a systematic review. Public Health 2018 Mar;156:67-78 [FREE Full text] [doi: 10.1016/j.puhe.2017.12.007] [Medline: 29408191]
29. Ajzen I, Fishbein M. Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs, NJ: Prentice-Hall; 1980.
30. Fan G, Song H, Yip S, Zhang T, He D. Impact of low vaccine coverage on the resurgence of COVID-19 in Central and Eastern Europe. One Health 2022 Jun;14:100402 [FREE Full text] [doi: 10.1016/j.onehlt.2022.100402] [Medline: 35611185]
31. Iguacel I, Álvarez-Najar JP, Vásquez PDC, Alarcón J, Orte M, Samatán E, et al. Citizen stance towards mandatory COVID-19 vaccination and vaccine booster doses: a study in Colombia, El Salvador and Spain. Vaccines (Basel) 2022 May 15;10(5):781 [FREE Full text] [doi: 10.3390/vaccines10050781] [Medline: 35632537]
32. Lai X, Zhu H, Wang J, Huang Y, Jing R, Lyu Y, et al. Public perceptions and acceptance of COVID-19 booster vaccination in China: a cross-sectional study. Vaccines (Basel) 2021 Dec 10;9(12):1461 [FREE Full text] [doi: 10.3390/vaccines9121461] [Medline: 34960208]
33. Tien NNA, Dung ND. Protecting right to health in developing countries during the pandemic of COVID-19. Int J Health Sci 2022 Apr 06:2428-2440. [doi: 10.53730/ijhs.v6nS2.5553]
34. Bonner KE, Ssekyanzi H, Sicsic J, Mueller JE, Toomey T, Ulrich AK, et al. What drives willingness to receive a new vaccine that prevents an emerging infectious disease? A discrete choice experiment among university students in Uganda. PLoS One 2022 May 19;17(5):e0268063 [FREE Full text] [doi: 10.1371/journal.pone.0268063] [Medline: 35587501]
35. de Bekker-Grob EW, Veldwijk J, Jonker M, Donkers B, Huisman J, Buis S, et al. The impact of vaccination and patient characteristics on influenza vaccination uptake of elderly people: a discrete choice experiment. Vaccine 2018 Mar 07;36(11):1467-1476 [FREE Full text] [doi: 10.1016/j.vaccine.2018.01.054] [Medline: 29426662]
36. Determann D, Korfage IJ, Lambooij MS, Bliemer M, Richardus JH, Steyerberg EW, et al. Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment. PLoS One 2014 Jul 24;9(7):e102505 [FREE Full text] [doi: 10.1371/journal.pone.0102505] [Medline: 25057914]
37. Luyten J, Kessels R, Atkins KE, Jit M, van Hoek AJ. Quantifying the public's view on social value judgments in vaccine decision-making: a discrete choice experiment. Soc Sci Med 2019 May;228:181-193 [FREE Full text] [doi: 10.1016/j.socscimed.2019.03.025] [Medline: 30925392]
38. Seanehia J, Treibich C, Holmberg C, Müller-Nordhorn J, Casin V, Raude J, et al. Quantifying population preferences around vaccination against severe but rare diseases: a conjoint analysis among French university students, 2016. Vaccine 2017 May 09;35(20):2676-2684. [doi: 10.1016/j.vaccine.2017.03.086] [Medline: 28408120]
39. Verelst F, Kessels R, Delva W, Beutels P, Willem L. Drivers of vaccine decision-making in South Africa: a discrete choice experiment. Vaccine 2019 Apr 03;37(15):2079-2089. [doi: 10.1016/j.vaccine.2019.02.056] [Medline: 30857931]

40. Ghazy RM, Abdou MS, Awaidy S, Sallam M, Elbarazi I, Youssef N, et al. Acceptance of COVID-19 vaccine booster doses using the Health Belief Model: a cross-sectional study in low-middle- and high-income countries of the East Mediterranean Region. Int J Environ Res Public Health 2022 Sep 25;19(19):12136 [FREE Full text] [doi: 10.3390/ijerph191912136] [Medline: 36231447]
41. Klein NP, Lewis N, Goddard K, Fireman B, Zerbo O, Hanson KE, et al. Surveillance for adverse events after COVID-19 mRNA vaccination. JAMA 2021 Oct 12;326(14):1390-1399 [FREE Full text] [doi: 10.1001/jama.2021.15072] [Medline: 34477808]
42. Centers for Disease Control and Prevention. Selected Adverse Events Reported after COVID-19 Vaccination. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/adverse-events.html [accessed 2023-01-16]
43. World Health Organization. Ethical Considerations in Developing a Public Health Response to Pandemic Influenza. URL: https://apps.who.int/iris/handle/10665/70006 [accessed 2023-01-16]
44. Lee Y, Li JQ. The role of communication transparency and organizational trust in publics' perceptions, attitudes and social distancing behaviour: a case study of the COVID‐19 outbreak. J Contingencies Crisis Manag 2021 Feb 07;29(4):368-384. [doi: 10.1111/1468-5973.12354]
45. Enria L, Waterlow N, Rogers NT, Brindle H, Lal S, Eggo RM, et al. Trust and transparency in times of crisis: results from an online survey during the first wave (April 2020) of the COVID-19 epidemic in the UK. PLoS One 2021 Feb 16;16(2):e0239247 [FREE Full text] [doi: 10.1371/journal.pone.0239247] [Medline: 33591985]
46. Engelbrecht MC, Kigozi NG, Heunis JC. Factors associated with limited vaccine literacy: lessons learnt from COVID-19. Vaccines (Basel) 2022 May 28;10(6):865 [FREE Full text] [doi: 10.3390/vaccines10060865] [Medline: 35746473]
47. Larson HJ, Heymann DL. Public health response to influenza A(H1N1) as an opportunity to build public trust. JAMA 2010 Jan 20;303(3):271-272. [doi: 10.1001/jama.2009.2023] [Medline: 20085957]
48. Hamel L, Lopes L, Kearney A, Sparks G, Stokes M, Brodie M. KFF COVID-19 Vaccine Monitor: June 2021. URL: https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-june-2021/ [accessed 2023-01-16]
49. Jensen UT, Ayers S, Koskan AM. Video-based messages to reduce COVID-19 vaccine hesitancy and nudge vaccination intentions. PLoS One 2022 Apr 6;17(4):e0265736 [FREE Full text] [doi: 10.1371/journal.pone.0265736] [Medline: 35385505]
50. Antonini M, Eid MA, Falkenbach M, Rosenbluth ST, Prieto PA, Brammli-Greenberg S, et al. An analysis of the COVID-19 vaccination campaigns in France, Israel, Italy and Spain and their impact on health and economic outcomes. Health Policy Technol 2022 Jun;11(2):100594 [FREE Full text] [doi: 10.1016/j.hlpt.2021.100594] [Medline: 34976711]

Abbreviations

DCE: discrete choice experiment
OR: odds ratio

Edited by Y Khader; submitted 28.09.22; peer-reviewed by YJ Bao, ECY Su; comments to author 28.11.22; revised version received 20.12.22; accepted 31.12.22; published 31.01.23.

Please cite as:
Tran BX, Do AL, Boyer L, Auquier P, Le HT, Le Vu MN, Dang THT, Cao KM, Le LDT, Cu LTN, Ly BV, Nguyen DAT, Nguyen MD, Latkin CA, Ho RCM, Ho CSH, Zhang MWB
Preference and Willingness to Pay for the Regular COVID-19 Booster Shot in the Vietnamese Population: Theory-Driven Discrete Choice Experiment
JMIR Public Health Surveill 2023;9:e43055
URL: https://publichealth.jmir.org/2023/1/e43055
doi:10.2196/43055
PMID:36599156

Original Paper

The Association Between Pediatric COVID-19 Vaccination and Socioeconomic Position: Nested Case-Control Study From the Pedianet Veneto Cohort

Erich Batzella[1], MSc; Anna Cantarutti[2,3], PhD; Nicola Caranci[4], MSc; Carlo Giaquinto[5], PhD; Claudio Barbiellini Amidei[1], MSc; Cristina Canova[1], PhD

[1]Unit of Biostatistics, Epidemiology and Public Health, Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padova, Padua, Italy

[2]Division of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

[3]National Centre for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

[4]Regional Health and Social Care Agency, Emilia-Romagna Region, Bologna, Italy

[5]Department for Women's and Children's Health, University of Padua, Padua, Italy

Corresponding Author:
Cristina Canova, PhD
Unit of Biostatistics, Epidemiology and Public Health
Department of Cardio-Thoraco-Vascular Sciences and Public Health
University of Padova
Via Loredan 18
Padua, 35131
Italy
Phone: 39 0498275391
Email: cristina.canova@unipd.it

Abstract

Background: The success of pediatric COVID-19 vaccination strongly depends on parents' willingness to vaccinate their children. To date, the role of socioeconomic position (SEP) in pediatric COVID-19 vaccination has not been thoroughly examined.

Objective: We evaluated the association between COVID-19 vaccination and SEP in a large pediatric cohort.

Methods: A case-control study design nested into a pediatric cohort of children born between 2007 and 2017, living in the Veneto Region and followed up to at least January 1, 2022, was adopted. Data on children were collected from the Pedianet database and linked with the regional COVID-19 registry. Each child vaccinated with at least one dose of any COVID-19 vaccine between July 1, 2021, and March 31, 2022, was matched by sex, year of birth, and family pediatrician to up to 5 unvaccinated children. Unvaccinated children with a positive outcome on the swab test within 180 days before the index date were excluded from the analyses. Children were geo-referenced to determine their area deprivation index (ADI)—a social and material deprivation measure calculated at the census block level and consisting of 5 socioeconomic items. The index was then categorized in quintiles based on the regional ADI level. The association between ADI quintiles and vaccination status was measured using conditioned logistic regression models to estimate odds ratios and the corresponding 95% CIs. Quantile-g-computation regression models were applied to develop a weighted combination of the individual items to estimate how much each component influenced the likelihood of vaccination. All analyses were stratified by age at vaccination (5-11 and 12-14 years).

Results: The study population consisted of 6475 vaccinated children, who were matched with 32,124 unvaccinated children. Increasing area deprivation was associated with a lower probability of being vaccinated, with approximately a linear dose-response relationship. Children in the highest deprivation quintile were 36% less likely to receive a COVID-19 vaccine than those with the lowest area deprivation (95% CI 0.59-0.70). The results were similar in the 2 age groups, with a slightly stronger association in 5-11–year-old children. When assessing the effects of the weighted combination of the individual items, a quintile increase was associated with a 17% decrease in the probability of being vaccinated (95% CI 0.80-0.86). The conditions that influenced the probability of vaccination the most were living on rent, being unemployed, and being born in single-parent families.

Conclusions: This study has shown a significant reduction in the likelihood of receiving a COVID-19 vaccine among children living in areas characterized by a lower SEP. Findings were robust among multiple analyses and definitions of the deprivation index. These findings suggest that SEP plays an important role in vaccination coverage, emphasizing the need to promote targeted public health efforts to ensure global vaccine equity.

(JMIR Public Health Surveill 2023;9:e44234) doi:10.2196/44234

KEYWORDS

SEP; socioeconomic position; quantile-g-computation; nested case-control study; COVID-19 vaccine; children; area deprivation index

Introduction

Given the high contagiousness of SARS-CoV-2 and the constant development and selection of new variants with increased pandemic potential, the success of population-based vaccination campaigns strongly depends on vaccination coverage [1]. It is largely accepted that COVID-19 vaccines are highly effective in reducing severe disease outcomes in adults, and numerous studies suggest they are also safe and effective in children [2-4]. In Italy, the first COVID-19 vaccines were approved for individuals aged between 12 and 15 years in May 2021 [5]. Since December 2021, the national drug regulation agency (Agenzia italiana del farmaco) has authorized the administration of a reduced dose of the vaccine also for children aged 5 to 11 years [5]. In December 2022, first dose coverage was 28.9% among children aged 5-11 years and 78.8% in those aged 12-19 years (information on children aged 12-14 years was not available) [6]. Before the end of the study period (March 2022), most of those first doses had been administered (with more than 95% coverage in both age groups) [6].

COVID-19 strongly impacts children from numerous perspectives, including interference with social interactions and mental health [7]. Children with COVID-19 are also at increased risk of developing multisystem inflammatory syndrome, and COVID-19 vaccines seem to reduce this risk [8,9]. Therefore, the importance of vaccination among children should not be ignored, in addition to the role played by the pediatric population in the COVID-19 transmission chain.

Depending on the characteristics of new SARS-CoV-2 variants and the mutating epidemiological scenario, future studies on vaccine effectiveness and safety in children will determine if vaccinating the pediatric population will remain a priority [1]. In 2019, the World Health Organization identified vaccine hesitancy as one of the 10 greatest threats to global health [10]. Some studies have suggested that a lower socioeconomic position (SEP) may be associated with a reduced chance of receiving a COVID-19 vaccine in adults [11-13]. However, evidence in children is scarce [14-22] also due to the shorter time since the approval of most COVID-19 vaccines. Furthermore, there are limitations to the generalizability of these works related to low numbers in the study populations and analyses exclusively based on referred data (ie, parental propensity to vaccinate their children or reported vaccine status), as opposed to actual recorded vaccination.

The surge of new variants could severely endanger public health by effectively eluding the immune system of previously immunized people, including children, increasing the population at risk of being infected and facing the consequences of COVID-19. If this were to occur, a prompt public health response would be required, and a vaccination campaign with booster doses or updated vaccines against new variants might have to rapidly reach very high portions of the population and potentially also the pediatric population. Therefore, identifying the socioeconomic determinants that most heavily affect the chances of children being vaccinated is crucial. Reducing these determinants' effects by increasing and diversifying vaccination campaign approaches and facilitating access to vaccination centers in specific at-risk groups should become a public health priority [7,23-26].

The area deprivation index (ADI) is a composite index that has been extensively used to examine socioeconomic status [27,28]. Having low education, being unemployed, living on rent, living in crowded households, and being born in single-parent families in the area are the items that compose the index and are among the most relevant socioeconomic determinants.

The aim of this study was to examine the association between ADI and its specific components with COVID-19 vaccine status in a large cohort of children aged 5-14 years in northern Italy. Given that socioeconomic variables are strongly correlated with one another, a weighted combination of the indicators (weighted deprivation index [WDI]) was also generated to assess which components most influence the likelihood of vaccination.

Methods

Setting and Study Design

A case-control study design nested into a large pediatric cohort of children living in the Veneto region, Italy, was adopted. Children were defined as being vaccinated if they had received at least one vaccine dose by the end of follow-up. Vaccination status in children aged 5-14 years was examined between July 1, 2021, and March 31, 2022. Each vaccinated child was matched by sex, year of birth (from 2007 to 2017), and family pediatrician (FP; among the 33 doctors included in the analyses) to up to 5 unvaccinated children. Based on the residential status of each child, an ADI was computed at the census block level.

Data on all children included in the study derive from the Veneto Pedianet database [29]. Pedianet is an Italian network of more than 400 FPs who use an established pediatric primary care database based on the Junior Bit software in their clinical practice. Data generated by Pedianet FPs are anonymized, in compliance with Italian regulations, stored in a protected "cloud" under a unique numerical identifier, and regularly checked for validation and quality control. The database includes patient demographic and clinical characteristics, containing diagnoses (free text or coded diagnoses using the International Classification of Diseases, Ninth Revision, Clinical Modification system), drug prescriptions (recorded in accordance with the Anatomical Therapeutic Chemical codes), health care copayment exemptions, specialist visits, diagnostic procedures, hospital admissions, growth parameters (repeated measures of height and weight in accordance with national indications), and free text to report symptoms or other medical observations related to the visit. Moreover, through the electronic health record of each child included in the cohort, we were able to link the Veneto Pedianet database with the regional COVID-19 registry—where results from COVID-19 swab tests are recorded—and the regional immunization database, for vaccination history.

Inclusion in the Pedianet database is voluntary; parents or legal guardians provided consent for their children's anonymized data to be used for research purposes in accordance with national and international regulations.

Pedianet is a large database rich in numerous clinical variables, but information on parental SEP is often missing or incomplete. SEP was first measured by means of an ADI, computed at the census block level retrieved from the 2011 Italian Census [27] (with a median number of residents of 66, IQR 28-181, per census block).

The ADI is based on 5 items that recurrently describe social and material deprivation: (1) low education, (2) unemployment, (3) living on rent, (4) crowded households, and (5) single-parent families [27,28]. The index is calculated as the sum of standardized indicators. The index was then categorized in quintiles based on the regional ADI level to ensure within-region appropriately represented categories.

The addresses of the children included in the study population were geo-referenced and linked to the census block of each Italian municipality. Through record linkage with the census block number, the ADI and all the variables defined at the census block level were retrieved for each address.

Ethical Considerations

This is an observational, retrospective, noninterventional study. According to a bylaw on the classification and implementation of observational drug-related research, as issued by the Italian National Drug Agency (an entity belonging to the Italian Ministry of Health), this study does not require approval by an ethics committee in Italy (Italian Drug Agency note on August 3, 2007). This study was conducted in accordance with the tenets of the Declaration of Helsinki and was compliant with the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance's *Guide on Methodological Standards in Pharmacoepidemiology*.

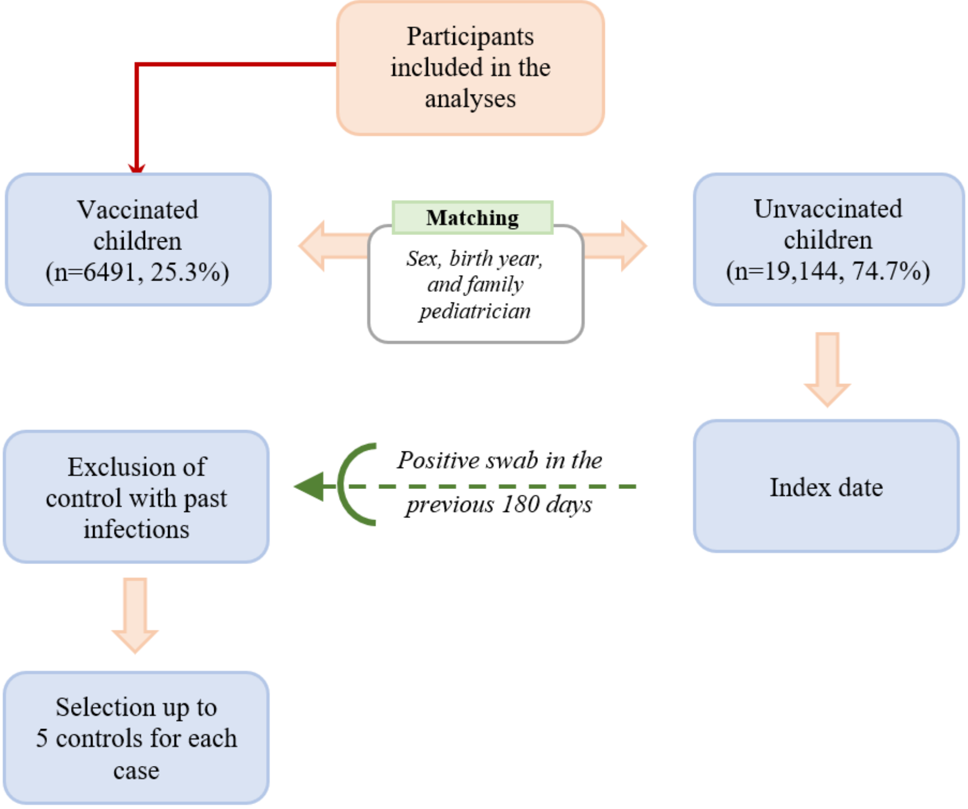
Ethical approval of the study and access to the database was approved by the Internal Scientific Committee of So.Se.Te. Srl, the legal owner of Pedianet.

Study Population

We identified 160,720 children from the Veneto Pedianet database. Exclusion criteria were (1) children born before 2007 or after 2017, (2) those who had the end of follow-up or the last contact with the FP before January 1, 2022, (3) those who were not properly geo-referenced or for whom it was not possible to determine an ADI, and (4) those for whom individual linkage with the COVID-19 registry was not possible (Figure S1 in Multimedia Appendix 1, details provided in Multimedia Appendix 2). Of the 25,635 children aged 5 to 14 years included in the study, 6491 (25.3%) received a COVID-19 vaccine between July 2021 and March 2022, and 19,144 (74.7%) did not. For the purpose of our analyses, we defined children as vaccinated if they received at least one dose of any COVID-19 vaccine.

Each vaccinated child was first matched by sex, year of birth, and FP to all possible available unvaccinated children. The same control was eligible for more than one case. Each control was then assigned an index date that corresponded to the date of COVID-19 vaccination of the matched case. Since matching by year of birth has led some controls to have a different age (±1 year) compared to cases, only controls who belonged to the same age group as cases at the index date (5-11 years and 12-14 years) were included. Given the indication of government agencies to vaccinate children with a previous SARS-CoV-2 infection only after 6 months, unvaccinated children with a positive swab within 180 days before the index date were excluded from the analyses. Finally, for each vaccinated child (case), up to 5 controls were randomly selected from the matched unvaccinated children (Figure 1). Matching was implemented to control for confounding and improve statistical efficiency.

Figure 1. Flowchart of the study design.



Statistical Analysis

Spearman rank correlation (ρ) was used to describe pair-wise relations between the ADI items.

Conditional logistic regression models were fitted to estimate odds ratios (ORs) and the corresponding 95% CIs for receiving a COVID-19 vaccine associated with ADI and the single items included in the index. All variables were treated as categorical with the first quintile (least deprived) as a reference. Age, sex, and FP confounding were controlled for by the study design.

To account for the strong correlation between the socioeconomic determinants that composed the ADI, we then applied a quantile-g-computation regression model to develop a weighted combination of socioeconomic variables (weighted deprivation index [WDI]) to examine its association with the outcome. Quantile-g-computation is a parametric, generalized linear model–based implementation of g-computation, which quantifies the expected change in the outcome, given a single quantile change in all exposures (the items included in the ADI) simultaneously while controlling for covariates. This method first transforms all exposures into quintiles and then fits a linear model with the dependent variables. Finally, weights are defined for each exposure, corresponding to the partial effect attributed to a specific exposure. Following these steps, the ADI items were combined to calculate a weighted index that was then categorized in quintiles. Both positive and negative relations between each indicator of the index and the outcome were allowed without any directional homogeneity assumption [30]. We defined as an item of concern any socioeconomic determinant contributing to the overall effect with a negative weight that exceeded the threshold of *1/p* (where *p* varied depending on the number of items with estimated negative weight in each stratum).

As a sensitivity analysis, the prevalence of foreigners in the area was also included among the area-based socioeconomic variables under study. All analyses were stratified by age at the index date (5-11 years and 12-14 years). Analyses were performed using SAS [31] and R [32] statistical software. We used *survival* [33] and *qgcomp* [30] packages to conduct conditional logistic regression and quantile-g-computation with R, respectively. Results with estimated *P* values of <.05 were considered significant.

Results

General Characteristics

The general characteristics of the study population are reported in Table S1 in Multimedia Appendix 1. Among the 25,635 children included in the study, the proportion of those who were vaccinated against COVID-19 was markedly higher for older children. The prevalence of vaccinated children varied largely based on the pediatrician and birth year (Table S1 in Multimedia Appendix 1). The distribution of the ADI quintiles of the children included in the study did not differ from the regional ADI distribution (Table S2 in Multimedia Appendix 1).

Among the specific items that composed the deprivation index, we observed strong correlations, especially between living on rent and being unemployed (ρ=0.31), followed by living on rent with crowded households (ρ=0.26). When including the presence of non-Italian ethnicity in the area, in the correlation matrix, an even stronger correlation was observed with living on rent (ρ=0.68), being unemployed (ρ=0.36), and living in crowded households (ρ=0.26).

The analyzed population consisted of 6475 vaccinated children, 3694 aged 5-11 years and 2781 aged 12-14 years, matched with 18,429 and 13,695 unvaccinated children, respectively.

Association With Vaccination Status

Results from the conditioned logistic regression models are shown in Table 1.

This analysis revealed negative associations between the probability of being vaccinated and increasing ADI quintiles, with approximately a linear dose-response relationship. Children in the highest deprivation quintile were 36% less likely to be vaccinated against COVID-19 than those living in the least deprived areas (OR 0.64, 95% CI 0.59-0.70). The results were similar in the 2 age groups, with more pronounced differences among 5-11–year-old children. For each quintile increase of ADI, the chances of being vaccinated against COVID-19 decreased by 10% (OR 0.90, 95% CI 0.88-0.91; Table 2).

Table 1. Association between the quintiles of area-deprivation index (ADI) and COVID-19 vaccination status, stratified by age groups (N=38,599).

Quintile of ADI	Vaccinated (n=6475), n (%)	Unvaccinated (n=32,124), n (%)	Odds ratio (95% CI)[a]
Total			
ADI 1	1501 (23.2)	6160 (19.2)	1 (reference)
ADI 2	1332 (20.6)	5909 (18.4)	0.93 (0.86-1.02)
ADI 3	1234 (19.1)	6291 (19.6)	0.81 (0.74-0.88)[b]
ADI 4	1281 (19.8)	6717 (20.9)	0.77 (0.71-0.84)[b]
ADI 5	1127 (17.4)	7047 (21.9)	0.64 (0.59-0.70)[b]
Age group of 5-11 years (n=22,123)			
ADI 1	866 (23.4)	3507 (19.0)	1 (reference)
ADI 2	764 (20.7)	3297 (17.9)	0.94 (0.85-1.05)
ADI 3	683 (18.5)	3383 (18.4)	0.82 (0.73-0.91)[b]
ADI 4	715 (19.4)	3971 (21.5)	0.72 (0.64-0.80)[b]
ADI 5	666 (18.0)	4271 (23.2)	0.62 (0.55-0.69)[b]
Age group of 12-14 years (n=16,476)			
ADI 1	635 (22.8)	2653 (19.4)	1 (reference)
ADI 2	568 (20.4)	2612 (19.1)	0.92 (0.81-1.05)
ADI 3	551 (19.8)	2908 (21.2)	0.80 (0.70-0.91)[b]
ADI 4	566 (20.4)	2746 (20.1)	0.86 (0.76-0.98)[b]
ADI 5	461 (16.6)	2776 (20.3)	0.67 (0.59-0.77)[b]

[a] Odds ratios and 95% CIs, estimated by means of a conditional logistic regression model.

[b] *P*<.001.

Table 2. Association between a quintile increase of deprivation index and COVID-19 vaccination status, stratified by age-group.

Interquintile increase in the deprivation index	Total (n=38,599), odds ratio (95% CI)	Age group of 5-11 years (n=22,123), odds ratio (95% CI)	Age group of 12-14 years (n=16,476), odds ratio (95% CI)
IQ ADI[a]	0.90 (0.88-0.92)[b]	0.88 (0.86-0.91)[b]	0.92 (0.89-0.95)[b]
IQ WDI[c]	0.83 (0.80-0.86)[b]	0.81 (0.77-0.85)[b]	0.84 (0.80-0.89)[b]
IQ WDI[d]	0.82 (0.79-0.85)[b]	0.81 (0.77-0.85)[b]	0.83 (0.79-0.88)[b]

[a] From the conditional logistic regression model; ADI: area deprivation index.

[b] *P*<.001.

[c] From the quantile g-computation model; WDI: weighted deprivation index.

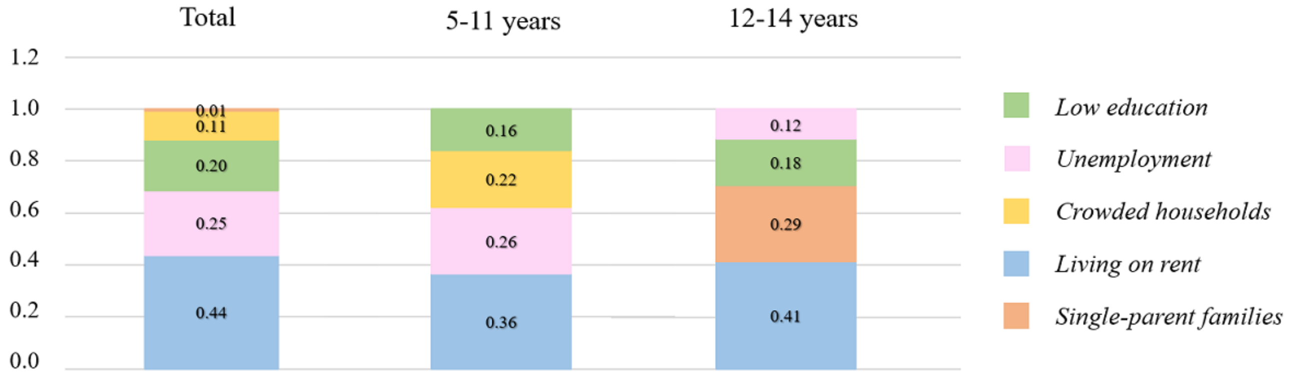
[d] From the quantile g-computation model including the presence of foreigners in the area.

This effect seems to be slightly stronger among children aged 5-11 years, where the chances of being vaccinated decreased by 12% (OR 0.88, 95% CI 0.86-0.91) for each quintile increase of the ADI, while the decrease corresponded to 8% among children aged 12-14 years (OR 0.92, 95% CI 0.89-0.95; Table 2).

When assessing the effects of the WDI, a quintile increase was associated with a decrease in the probability of being vaccinated of 17% (overall OR 0.83, 95% CI 0.80-0.86; Table 2). The most influential condition on the probability of being vaccinated was living on rent (weight –0.44), which showed the highest weight in both 5-11–year-old children (weight –0.36) and 12-14–year-old children (weight –0.41; Figure 2). Being unemployed was also influential both overall and among the 5-11–year-old children (weights –0.25 and –0.26, respectively), while among 12-14–year-old children, single-parent families weighed the most (weight –0.29; Figure 2).

When including the prevalence of foreigners in the area among the area-based socioeconomic variables, no significant differences were observed in the WDI estimates (overall OR 0.82, 95% CI 0.79-0.85; Table 2). Nevertheless, on observing the weights of single indicators, a stronger association was found among younger children; the multiracial area was always identified as an influential condition, with the greatest weight, compared to the other variables both in the overall population and among children aged 12-14 years (weights –0.45 and –0.64, respectively; Table S3 in Multimedia Appendix 1). WDI analysis confirmed that living on rent remained the greatest determinant in the association of socioeconomic deprivation with COVID-19 vaccination in 5-11–year-old children (weight –0.24).

Figure 2. Quantile-g-computation model regression index negative weights for the probability of being vaccinated.



Discussion

Principal Findings and Comparison With Prior Work

The findings of this study have shown a significant reduction in the likelihood of receiving a COVID-19 vaccine among children living in areas characterized by greater socioeconomic deprivation. The dose-response gradient was strongest among children aged 5-11 years, for whom social and material deprivation reduced the chances of being vaccinated the most. The robustness of these findings is confirmed by results that were consistent across multiple analyses and definitions of the deprivation index.

Previous studies, mainly conducted in the United States, have shown how social inequalities affect COVID-19 vaccination rates in children [14-21]. To the best of our knowledge, this is one of the first studies to examine the association between a robust and largely used ADI and the registered vaccination status against COVID-19. Most other studies examining this association were surveys and mainly focused on parental willingness to vaccinate their children [14-16,18,20,21]. Only a few studies examined the COVID-19 vaccination status of children [17-19,22], but in most cases, data were reported by the parents [17-19].

When examining single indicators of lower socioeconomic status, items that were most negatively associated with the chances that a child received a COVID-19 vaccine were lower education [16,18-20], low income levels [14,18,20], high unemployment rates [16], living on rent [15], as well as living in a multiethnic area [14,18].

Comparable results to ours were reported by a study conducted in the United Kingdom based on over 27,000 students aged 9-18 years, which showed that the hesitancy toward vaccinating their children among both undecided parents (n=1688) and those against vaccination (n=4874) was higher among children in the most deprived quintile [15]. Regarding studies on reported vaccination status, in a large US survey of 59,424 children aged 12-17 years, Nguyen et al [19] found greater vaccination coverage among children whose parents had more than college education (adjusted prevalence ratio 1.15, 95% CI 1.05-1.25) [19]. Furthermore, a recent study elucidated variables associated with COVID-19 vaccination among Israeli adolescents (n=43,919) using data from an electronic database, showing that adolescents of low SEP had a very low vaccination rate (42.3%) compared to those having a high SEP (80.4%), which was highly significant upon univariate analysis but did not reach statistical significance using a multivariate approach [22].

Vaccination of children and adolescents largely depends on the vaccine hesitancy or antivaccination positions of the parents. Lower parental SEP has been associated with a lower parental propensity to vaccinate their children. In Europe, over the past decades, the link between SEP and vaccine coverage has been weakening, and a much more relevant role seems to be played by the lack of vaccine confidence or alternative health beliefs [34]. In the present mutating context, the pandemic may have determined hardly predictable effects. The increase in the vaccine coverage rate may have been primarily driven by the fear of infection and long-term consequences of COVID-19, while vaccine hesitancy may have increased as a result of the fear of adverse effects for newly registered vaccines, which may have been greater due to the novelty of mRNA vaccines [35-37].

Public perception of the pandemic and of COVID-19 vaccines was extremely heterogeneous. Mass vaccination campaigns and government incentives such as vaccine passports, which had not been implemented for decades in high-income countries, may have also affected vaccine hesitancy in different ways [38,39]. Nonetheless, it remains unclear how these opposing tendencies may have affected the association between SEP and COVID-19 vaccination.

When assessing the association between interquintile changes in the deprivation index and the likelihood of receiving a COVID-19 vaccine, the effects estimated from the WDI were always stronger. The ADI treats each dimension as if they were of equal importance by summing up their relative contribution with equal weight [27], rather than weighing indicators in accordance with their association with the outcome. This approach also allowed us to assess the role of the ethnic composition in each area, which was not included in the ADI. This variable was strongly associated with the likelihood of being vaccinated when considered individually as a socioeconomic determinant, but its inclusion in the WDI did not change the magnitude of the association of each quintile with the chances of being vaccinated, although this component weighted more than the others. This helps confirming the robustness of the national deprivation index (ie, ADI) to the inclusion of such a component.

We observed large differences in vaccination coverage by age, with much higher percentages among older children. A reasonable explanation could be found in the Italian COVID-19 regulations, issued differently in accordance with age: all adolescents aged ≥12 years had to exhibit a vaccine passport (a certificate provided to all people immunized against COVID-19) in order to participate, to most public, leisure, and sport activities, while there were no restrictions for younger children. This may also have contributed to a reduction in the role of SEP on the likelihood of vaccination specifically in adolescents, flattening the differences in this group. We can expect that this effect could be more evident in older adolescents (aged 15-19 years), who exceeded the maximum age required to be followed up by a FP. As age increases, the parental willingness to vaccinate their children may play a weaker role, as well as the SEP itself, given a greater autonomy of the adolescent.

Our study also showed how FPs were among the most relevant elements in determining the chances of receiving a COVID-19 vaccine. This suggests that promoting vaccination campaigns and raising awareness of the importance of pediatric vaccinations may significantly affect vaccination coverage among the children they assist.

Strengths and Limitations

Our study has several strengths. First, the large cohort study was based on more than 25,000 children with real data on the vaccination status, regardless of the willingness or hesitancy to get a COVID-19 vaccine. Vaccinations were entirely managed by the National Health Service, and all vaccinations were promptly recorded in a centralized register, so data on vaccination coverage were complete and constantly updated. Data on previous SARS-CoV-2 infections that determined changes in the vaccination schedule based on government indications allowed us to perform more accurate analyses, thereby including only children who did not receive a COVID-19 vaccine for reasons likely not related to vaccine indications. Furthermore, we controlled by design the interpediatrician variability, which emerged as one of the main determinants of receiving a COVID-19 vaccination in our study population.

Another strength of this study was addressing the presence of intercorrelation between social determinants by using an ADI, both nonweighted and weighted, through a data-driven method for modeling the exposure to multiple socioeconomic determinants. This method provided a measure of the simultaneous effect of the included socioeconomic indicators after a quantile transformation, which makes the exposure more robust to extreme values and more comparable to other settings, as opposed to using value-based cutoffs [40]. This supervised approach empirically examines single contributions of each indicator to the association, allowing us to understand which condition most influences the likelihood of receiving a vaccine.

The main limitation of this work is the absence of comprehensive individual information on socioeconomic status, which is why a census block–level deprivation index was used. Still, we know that ADI represents a simplification that may not exactly correspond to the reality, but it acts as a good proxy of individual socioeconomic status or as a comparative measure of socioeconomic background [27].

A possible bias could derive from the exclusion of some FPs that did not adhere to the COVID-19 data registry linkage or that were missing up-to-date information about the children they assisted. However, we believe that this bias should be negligible, since the distribution of the ADI quintiles among the children included in the study did not differ from the regional ADI quintile distribution, suggesting that the study population is likely to be representative of the regional population.

Moreover, some residual confounders related to specific indications to not vaccinate children (ie, due to specific health conditions) cannot be entirely ruled out. However, the magnitude of the association between the likelihood of being vaccinated and the deprivation indexes we used makes an overturning of the results unlikely.

The relatively short study period also is unlikely to have affected our results, given the low vaccination rates in children from April 2022 onward.

Conclusions

In a rapidly evolving pandemic scenario, achieving good vaccination coverage in adults is crucial, and high vaccination coverage could become a priority among children as well. It is of utmost importance for public health experts to identify population strata at risk of not achieving adequate vaccine coverage in order to remove barriers to accessing vaccines and to promoting targeted campaigns for these groups at increased risk. The relevance of vaccine coverage applies to COVID-19 vaccines, but also to other vaccines such as those against seasonal influenza.

Acknowledgments

The authors gratefully acknowledge the contributions of Giorgia Stoppa (University of Padova), Francesco Pirotti (University of Padova), Luigi Cantarutti (Pedianet), and all Pedianet family pediatricians (Pedianet): Alongi Angelo, Andreola Barbara, Angelini Roberta, Apostolo Maria Grazia, Arcangeli Bruno, Argo Giovanna, Avarello Giovanni, Azzoni Lucia Barbazza, Maria Carolina, Barbieri Patrizia, Belluzzi Gabriele, Benetti Eleonora, Biasci Paolo, Boe Franca, Bollettini Stefano, Bonaiuto Francesco, Bontempelli Anna Maria, Brusaterra Ivana, Budassi Roberto, Caccini Massimo, Calì Laura, Cantalupi Laura, Cardarelli Chiara Carli Giovanna, Carnazza Sylvia, Castaldo Massimo, Castelli Stefano, Cavedagni Monica, Cera Giuseppe Egidio, Chillemi Chiara, Cichello Francesca, Cicione Giuseppe, Ciscato Carla, Clerici Schoeller Mariangela, Collacciani Giuseppe, Comaita Fabrizio, Conte Valeria, Conte Ugo Alfredo, Corro' Roberta, Costagliola Rosaria, Costanzo Nicola, Cozzani Sandra, Cuboni Giancarlo, Curia Giorgia, D'alia Caterina, D'Amanti Vito Francesco, D'Avino Antonio, De Clara Roberto, De Giovanni Lorenzo, De Marchi Annamaria, Del Ponte Gigliola, Di Giampietro Tiziana, Di Mauro Giuseppe, Di Santo Giuseppe, Di Saverio Piero, Dieli Mattea, Dolci Marco, Doria Mattia, El Mazloum Dania, Fadda Maria Carmen, Falco Pietro Fama Mario, Faraci Marco, Farina Maria Immacolata, Favilli Tania, Federico Mariagrazia, Felice Michele, Ferraiuolo Maurizio, Ferrara Enrico Ferretti Michele, Forcina Paolo, Foti Patrizia, Frison Ezio, Fusco Fabrizio, Gallo Giovanni, Galvagno Andrea, Genova Grazia, Gentilucci Pierfrancesco, Giampaolo Giuliana, Gianfredi Francesco, Girotto Silvia, Grelloni Mauro, Grugnetti Mirco, Isca Antonina, Lagrasta Urania, lisabetta Landi Massimo, Letta Maria Rosaria, Lietti Giuseppe, Lista Cinzia, Lucantonio Ricciardo, Luise Francesco, Macropodio Nadia, Marine Francesca Mariniello Lorenzo, Marostica Gabriella, Masotti Sergio, Meneghetti Stefano, Milani Massimo, Milone Stella Vittoria, Moggia Donatella, Monteleone Angela Maria, Mussinu Pierangela, Naccari Anna, Nicoloso Flavia, Olimpi Laura Maria, Ongaro Riccardo, Palma Maria Maddalena Pasinato Angela, Passarella Andrea, Pazzola Pasquale, Perin Monica, Perri Danilo, Pescosolido Silvana Rosa, Petrazzuoli Giovanni, Petrotto Giuseppe, Picco Patrizia, Pirola Ambrogina, Pisanello Lorena, Pittarello Daniele, Porro Elena, Porto Adolfo Francesco, Puma Antonino, Puocci Maria Paola, Righetti Andrea, Rizzari Rosaria, Rosas Paolo, Ruffato Bruno, Ruggieri Lucia, Russo Annarita, Salamone Pietro, Sambugaro Daniela, Saretta Luigi, Sarno Vittoria, Savio Valentina, Sciolla Nico Maria, Semenzato Rossella, Silvan Carla, Soldà Giorgia, Spanevello Valter Spedale Sabrina, Speranza Francesco, Stefani Sara, Storelli Francesco, Tamassia Gianni, Tambaro Paolo, Toffol Giacomo, Tonelli Gabriele Tulone Silvia, Tummarello Angelo Giuseppe, Venditti Sergio, Vitale Maria Grazia, Volpe Francescopaolo, Volpe Concetta, Vozzi Aldo Zanon Giulia, Zuccolo Maria Luisa.

Data Availability

The data that support the findings of this study are available on request from the corresponding author (CC). The data are not publicly available due to restrictions (containing information that could compromise the privacy of research participants).

Authors' Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by EB and CC. The first draft of the manuscript was written by EB, AC, CBA, and CC and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Figures and Tables.
[PDF File (Adobe PDF File), 268 KB - publichealth_v9i1e44234_app1.pdf]

Multimedia Appendix 2
Details on linkage with healthcare registries of the Veneto Region.
[DOCX File , 12 KB - publichealth_v9i1e44234_app2.docx]

References

1. World Health Organization. Strategy to achieve global Covid-19 vaccination by mid-2022. World Health Organization. URL: https://cdn.who.int/media/docs/default-source/immunization/covid-19/strategy-to-achieve-global-covid-19-vaccination-by-mid-2022.pdf [accessed 2023-01-25]
2. Li G, Cappuccini F, Marchevsky NG, Aley PK, Aley R, Anslow R, Bristol Clinical Research Nurse Team. Safety and immunogenicity of the ChAdOx1 nCoV-19 (AZD1222) vaccine in children aged 6–17 years: a preliminary report of COV006, a phase 2 single-blind, randomised, controlled trial. Lancet 2022 Jun;399(10342):2212-2225. [doi: 10.1016/s0140-6736(22)00770-x]

3. Tan SH, Cook AR, Heng D, Ong B, Lye DC, Tan KB. Effectiveness of BNT162b2 vaccine against omicron in children 5 to 11 years of age. N Engl J Med 2022 Aug 11;387(6):525-532. [doi: 10.1056/nejmoa2203209]
4. Walter EB, Talaat KR, Sabharwal C, Gurtman A, Lockhart S, Paulsen GC, C4591007 Clinical Trial Group. Evaluation of the BNT162b2 Covid-19 vaccine in children 5 to 11 years of age. N Engl J Med 2022 Jan 06;386(1):35-46 [FREE Full text] [doi: 10.1056/NEJMoa2116298] [Medline: 34752019]
5. Piano Vaccinale del Commissario straordinario. Governo Italiano Presidenza del Consiglio dei Ministri. URL: https://www.governo.it/it/dipartimenti/cscovid19-pianovaccini/16417 [accessed 2022-01-02]
6. Vaccinazioni Regione del Veneto Vaccinazioni Regione del Veneto. Regione del Veneto. URL: https://www.regione.veneto.it/dati-vaccinazioni/ [accessed 2022-12-01]
7. Xiao Y, Yip PS, Pathak J, Mann JJ. Association of social determinants of health and vaccinations with child mental health during the COVID-19 pandemic in the US. JAMA Psychiatry 2022 Jun 01;79(6):610-621 [FREE Full text] [doi: 10.1001/jamapsychiatry.2022.0818] [Medline: 35475851]
8. Levy M, Recher M, Hubert H, Javouhey E, Fléchelles O, Leteurtre S, et al. Multisystem inflammatory syndrome in children by COVID-19 vaccination status of adolescents in France. JAMA 2022 Jan 18;327(3):281-283 [FREE Full text] [doi: 10.1001/jama.2021.23262] [Medline: 34928295]
9. Shein SL, Carroll CL, Remy KE, Rogerson CM, McCluskey CK, Lin A, et al. Epidemiology and outcomes of SARS-CoV-2 infection or multisystem inflammatory syndrome in children vs influenza among critically ill children. JAMA Netw Open 2022 Jun 01;5(6):e2217217 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.17217] [Medline: 35704321]
10. Ten threats to global health in 2019. World Health Organization. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2023-01-25]
11. Caspi G, Dayan A, Eshal Y, Liverant-Taub S, Twig G, Shalit U, et al. Socioeconomic disparities and COVID-19 vaccination acceptance: a nationwide ecologic study. Clin Microbiol Infect 2021 Oct;27(10):1502-1506 [FREE Full text] [doi: 10.1016/j.cmi.2021.05.030] [Medline: 34111591]
12. Cesaroni G, Calandrini E, Balducci M, Cappai G, Di Martino M, Sorge C, et al. Educational inequalities in COVID-19 vaccination: a cross-sectional study of the adult population in the Lazio region, Italy. Vaccines (Basel) 2022 Feb 25;10(3):364 [FREE Full text] [doi: 10.3390/vaccines10030364] [Medline: 35334995]
13. Lee J, Huang Y. COVID-19 vaccine hesitancy: the role of socioeconomic factors and spatial effects. Vaccines (Basel) 2022 Feb 24;10(3):352 [FREE Full text] [doi: 10.3390/vaccines10030352] [Medline: 35334984]
14. Alfieri NL, Kusma JD, Heard-Garris N, Davis MM, Golbeck E, Barrera L, et al. Parental COVID-19 vaccine hesitancy for children: vulnerability in an urban hotspot. BMC Public Health 2021 Sep 13;21(1):1662 [FREE Full text] [doi: 10.1186/s12889-021-11725-5] [Medline: 34517848]
15. Fazel M, Puntis S, White SR, Townsend A, Mansfield KL, Viner R, et al. Willingness of children and adolescents to have a COVID-19 vaccination: results of a large whole schools survey in England. EClinicalMedicine 2021 Oct;40:101144 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101144] [Medline: 34608453]
16. Fedele F, Aria M, Esposito V, Micillo M, Cecere G, Spano M, et al. COVID-19 vaccine hesitancy: a survey in a population highly compliant to common vaccinations. Hum Vaccin Immunother 2021 Oct 03;17(10):3348-3354 [FREE Full text] [doi: 10.1080/21645515.2021.1928460] [Medline: 34096836]
17. Gray A, Fisher CB. Determinants of COVID-19 vaccine uptake in adolescents 12-17 years old: examining pediatric vaccine hesitancy among racially diverse parents in the United States. Front Public Health 2022 Mar 22;10:844310 [FREE Full text] [doi: 10.3389/fpubh.2022.844310] [Medline: 35392471]
18. McKinnon B, Quach C, Dubé È, Tuong Nguyen C, Zinszer K. Social inequalities in COVID-19 vaccine acceptance and uptake for children and adolescents in Montreal, Canada. Vaccine 2021 Dec 03;39(49):7140-7145 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.077] [Medline: 34763947]
19. Nguyen KH, Nguyen K, Geddes M, Allen JD, Corlin L. Trends in adolescent COVID-19 vaccination receipt and parental intent to vaccinate their adolescent children, United States, July to October, 2021. Ann Med 2022 Dec 03;54(1):733-742 [FREE Full text] [doi: 10.1080/07853890.2022.2045034] [Medline: 35238263]
20. Skeens M, Guttoo P, Stanek JR, Taylor K, Stratz E, Ardura MI, et al. An exploration of COVID-19 impact and vaccine hesitancy in parents of pediatric hematopoietic stem cell transplant (HCT) recipients. Bone Marrow Transplant 2022 Apr 24;57(4):547-553 [FREE Full text] [doi: 10.1038/s41409-022-01587-9] [Medline: 35075246]
21. Zhang KC, Fang Y, Cao H, Chen H, Hu T, Chen YQ, et al. Parental acceptability of COVID-19 vaccination for children under the age of 18 years: cross-sectional online survey. JMIR Pediatr Parent 2020 Dec 30;3(2):e24827 [FREE Full text] [doi: 10.2196/24827] [Medline: 33326406]
22. Shkalim Zemer V, Grossman Z, Cohen H, Hoshen M, Gerstein M, Richenberg Y, et al. Variables associated with COVID-19 vaccination among Israeli adolescents and the need for targeted interventions. Pediatr Infect Dis J 2022 Nov 01;41(11):927-932 [FREE Full text] [doi: 10.1097/INF.0000000000003664] [Medline: 35980828]
23. Abrams EM, Greenhawt M, Shaker M, Pinto AD, Sinha I, Singer A. The COVID-19 pandemic: adverse effects on the social determinants of health in children and families. Ann Allergy Asthma Immunol 2022 Jan;128(1):19-25 [FREE Full text] [doi: 10.1016/j.anai.2021.10.022] [Medline: 34699969]

24. Chhibber A, Kharat A, Duong K, Nelson RE, Samore MH, Wilson FA, et al. Strategies to minimize inequity in COVID-19 vaccine access in the US: implications for future vaccine rollouts. Lancet Reg Health Am 2022 Mar;7:100138 [FREE Full text] [doi: 10.1016/j.lana.2021.100138] [Medline: 34901919]
25. Liao TF. Social and economic inequality in coronavirus disease 2019 vaccination coverage across Illinois counties. Sci Rep 2021 Sep 16;11(1):18443 [FREE Full text] [doi: 10.1038/s41598-021-97705-6] [Medline: 34531435]
26. Liao TF, De Maio F. Association of social and economic inequality with coronavirus disease 2019 incidence and mortality across US counties. JAMA Netw Open 2021 Jan 04;4(1):e2034578 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.34578] [Medline: 33471120]
27. Rosano A, Pacelli B, Zengarini N, Costa G, Cislaghi C, Caranci N. [Update and review of the 2011 Italian deprivation index calculated at the census section level]. Epidemiol Prev 2020;44(2-3):162-170. [doi: 10.19191/EP20.2-3.P162.039] [Medline: 32631016]
28. Caranci N, Biggeri A, Grisotto L, Pacelli B, Spadea T, Costa G. [The Italian deprivation index at census block level: definition, description and association with general mortality]. Epidemiol Prev 2010;34(4):167-176. [Medline: 21224518]
29. The Pedianet Project. Pedianet. URL: http://www.pedianet.it/en [accessed 2023-01-25]
30. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 2020 Apr;128(4):047004. [doi: 10.1289/ehp5838]
31. SAS/ACCESS® 9.4 Interface to ADABAS: Reference. SAS Institute Inc. URL: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/acadbas/titlepage.htm [accessed 2023-01-25]
32. R: A Language and Environment for Statistical Computing. R Core Team. 2016. URL: https://www.R-project.org/ [accessed 2023-01-25]
33. Therneau T, Lumley T, Atkinson E, Crowson C. survival: Survival Analysis. R package version 3. 2023. URL: https://CRAN.R-project.org/package=survival [accessed 2023-01-25]
34. de Figueiredo A, Johnston I, Smith DMD, Larson HJ, Jones N. Changing socioeconomic determinants of childhood vaccines: a global analysis over three decades. Lancet Glob Health 2015 Mar;3:S20. [doi: 10.1016/s2214-109x(15)70139-7]
35. Cascini F, Pantovic A, Al-Ajlouni YA, Failla G, Puleo V, Melnyk A, et al. Social media and attitudes towards a COVID-19 vaccination: a systematic review of the literature. EClinicalMedicine 2022 Jun;48:101454 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101454] [Medline: 35611343]
36. Mertens G, Lodder P, Smeets T, Duijndam S. Fear of COVID-19 predicts vaccination willingness 14 months later. J Anxiety Disord 2022 May;88:102574 [FREE Full text] [doi: 10.1016/j.janxdis.2022.102574] [Medline: 35512598]
37. Prickett KC, Habibi H, Carr PA. COVID-19 vaccine hesitancy and acceptance in a cohort of diverse New Zealanders. Lancet Reg Health West Pac 2021 Sep;14:100241 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100241] [Medline: 34377999]
38. Hogan CM, Waselewski ME, Szachta P, Wolff C, Amaro X, Chang T. Perceptions of COVID-19 vaccine incentives among adolescents and young adults. JAMA Netw Open 2022 Jun 01;5(6):e2216628 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.16628] [Medline: 35675072]
39. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 2021 Apr;6(4):e210-e221. [doi: 10.1016/s2468-2667(21)00012-8]
40. Carrión D, Colicino E, Pedretti NF, Arfer KB, Rush J, DeFelice N, et al. Neighborhood-level disparities and subway utilization during the COVID-19 pandemic in New York City. Nat Commun 2021 Jun 17;12(1):3692 [FREE Full text] [doi: 10.1038/s41467-021-24088-7] [Medline: 34140520]

Abbreviations

ADI: area deprivation index
FP: family pediatricians
OR: odds ratio
SEP: socioeconomic position
WDI: weighted deprivation index

Edited by A Mavragani, T Sanchez; submitted 11.11.22; peer-reviewed by G Cesaroni, A Catalano; comments to author 25.11.22; revised version received 14.12.22; accepted 06.01.23; published 01.02.23.

Please cite as:
Batzella E, Cantarutti A, Caranci N, Giaquinto C, Barbiellini Amidei C, Canova C
The Association Between Pediatric COVID-19 Vaccination and Socioeconomic Position: Nested Case-Control Study From the Pedianet Veneto Cohort
JMIR Public Health Surveill 2023;9:e44234
URL: https://publichealth.jmir.org/2023/1/e44234
doi:10.2196/44234
PMID:36645419

Original Paper

Tracking Changes in Mobility Before and After the First SARS-CoV-2 Vaccination Using Global Positioning System Data in England and Wales (Virus Watch): Prospective Observational Community Cohort Study

Vincent Nguyen[1,2], MPhil; Yunzhe Liu[3], PhD; Richard Mumford[4], BSc; Benjamin Flanagan[4], BSc; Parth Patel[1], MBBS; Isobel Braithwaite[1], MBBS; Madhumita Shrotri[1], MBBS; Thomas Byrne[1], MPhil, MPH; Sarah Beale[1,2], MPhil, MSc; Anna Aryee[1], MBBS; Wing Lam Erica Fong[1], MSc; Ellen Fragaszy[1,5], PhD; Cyril Geismar[1,2], MSc; Annalan M D Navaratnam[1,2], MBBS; Pia Hardelid[6], PhD; Jana Kovar[2], PhD; Addy Pope[4], PhD; Tao Cheng[3], PhD; Andrew Hayward[2], MD, MBBS; Robert Aldridge[1], PhD, MBBS, MEng; Virus Watch Collaborative[7]

[1]Centre for Public Health Data Science, Institute of Health Informatics, University College London, London, United Kingdom
[2]Institute of Epidemiology and Health Care, University College London, London, United Kingdom
[3]SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering, University College London, London, United Kingdom
[4]Technical Research Department, Esri, Edinburgh, United Kingdom
[5]Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom
[6]Department of Population, Policy and Practice, University College London Great Ormond Street Institute of Child Health, London, United Kingdom
[7]See Acknowledgements

Corresponding Author:
Robert Aldridge, PhD, MBBS, MEng
Centre for Public Health Data Science
Institute of Health Informatics
University College London
222 Euston Road
London, NW1 2DA
United Kingdom
Phone: 44 20 3549 5541
Email: r.aldridge@ucl.ac.uk

Abstract

Background: Evidence suggests that individuals may change adherence to public health policies aimed at reducing the contact, transmission, and spread of the SARS-CoV-2 virus after they receive their first SARS-CoV-2 vaccination when they are not fully vaccinated.

Objective: We aimed to estimate changes in median daily travel distance of our cohort from their registered addresses before and after receiving a SARS-CoV-2 vaccine.

Methods: Participants were recruited into Virus Watch starting in June 2020. Weekly surveys were sent out to participants, and vaccination status was collected from January 2021 onward. Between September 2020 and February 2021, we invited 13,120 adult Virus Watch participants to contribute toward our tracker subcohort, which uses the GPS via a smartphone app to collect data on movement. We used segmented linear regression to estimate the median daily travel distance before and after the first self-reported SARS-CoV-2 vaccine dose.

Results: We analyzed the daily travel distance of 249 vaccinated adults. From 157 days prior to vaccination until the day before vaccination, the median daily travel distance was 9.05 (IQR 8.06-10.09) km. From the day of vaccination to 105 days after vaccination, the median daily travel distance was 10.08 (IQR 8.60-12.42) km. From 157 days prior to vaccination until the vaccination date, there was a daily median decrease in mobility of 40.09 m (95% CI –50.08 to –31.10; *P*<.001). After vaccination, there was a median daily increase in movement of 60.60 m (95% CI 20.90-100; *P*<.001). Restricting the analysis to the third national lockdown (January 4, 2021, to April 5, 2021), we found a median daily movement increase of 18.30 m (95% CI –19.20

to 55.80; *P*=.57) in the 30 days prior to vaccination and a median daily movement increase of 9.36 m (95% CI 38.6-149.00; *P*=.69) in the 30 days after vaccination.

Conclusions: Our study demonstrates the feasibility of collecting high-volume geolocation data as part of research projects and the utility of these data for understanding public health issues. Our various analyses produced results that ranged from no change in movement after vaccination (during the third national lock down) to an increase in movement after vaccination (considering all periods, up to 105 days after vaccination), suggesting that, among Virus Watch participants, any changes in movement distances after vaccination are small. Our findings may be attributable to public health measures in place at the time such as movement restrictions and home working that applied to the Virus Watch cohort participants during the study period.

(JMIR Public Health Surveill 2023;9:e38072) doi:10.2196/38072

KEYWORDS

COVID-19; SARS-CoV-2; vaccination; global positioning system; GPS; movement tracking; geographical tracking; mobile app; health application; surveillance; public health; mHealth; mobile surveillance; tracking device; geolocation

Introduction

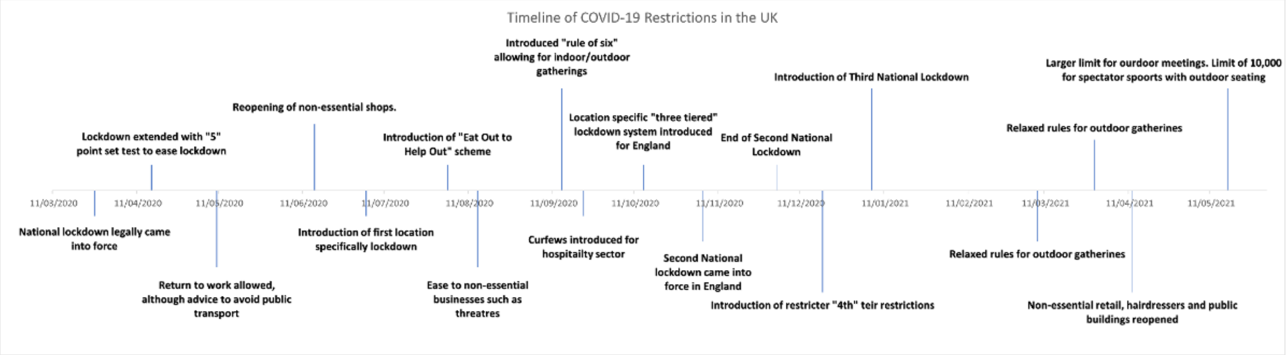
The UK response to the COVID-19 pandemic has included multiple rounds of restrictions on nonessential movement to reduce contacts and control transmission [1]. Examples of permissible travel included necessary shopping, exercise, medical need, or travel to and from essential work [2]. However, the restriction of movement can have a detrimental impact on a wide variety of outcomes such as reduced physical activity [3], mental health, domestic accidents, the economy, and education [4]. Adherence to travel-based public health interventions and those especially aimed at limiting nonhousehold contact has varied through time [5]. A brief timeline of the United Kingdom’s approach to restrictions can be found in Figure 1 [6].

The introduction of vaccinations against COVID-19 reduces SARS-CoV-2 transmission and disease [7] and, as a result, is a critical part of the strategies to allow more normal societal mixing due to increased immunization. However, in the UK context, there are concerns that misunderstandings about the effectiveness of the COVID-19 vaccine after the first dose, which may be leading to a reduction in adherence to other public health policies and increased exposure of partially protected individuals [8]. Preliminary research on vaccination in February 2021 found that 41% of individuals age >80 years who had their first dose of the vaccine had met at least one other person outside their household within 3 weeks. These contacts did not include care workers or members of their support bubble that were permitted by the restrictions in place at the time [8]. This finding is concerning as antibody levels would not have risen in the 1-2 weeks following the first dose of vaccine [9,10]. Further evidence also suggests that those aged >80 years are more likely to have a positive polymerase chain reaction test in the first 9 days after vaccination compared to a control group, which might be explained by increased mobility and contacts between people in the period following vaccination [11]. During the emergence of the Delta (B.1.617.2) SARS-CoV-2 variant, the effectiveness of both the Oxford-AstraZeneca and Pfizer-BioNTech vaccines were estimated to be 33% against symptomatic disease after a first dose [12], although protection against hospitalization appears to be much higher [13]. Therefore, if those who are not fully vaccinated increase their level of social contact and mobility after vaccination, their risk of becoming infected and infecting others may also be increased.

Understanding movement after the first vaccination is important, as it could help policy makers understand how perceived protection from the vaccination program may negatively offset the effectiveness of other policies designed to reduce transmission. Such arguments can also be extended beyond the first vaccination with the introduction of booster shots and variant-specific vaccines in relation to future SARS-CoV-2 variants. Although previous studies have attempted to investigate travel distances after vaccination [14], these studies were conducted using mobile call data based on cellular tower location, which is considered less accurate compared to GPS location. Alternative methods for movement tracking exist, such as tracking debit/credit card usage [15] and QR code check-in for venues [16]. Although such technology can be used to track specific activities, it can be limited as they require active interaction from users, and assessing travel distance can be limited if travel occurs beyond the use of such technology. GPS technology, on the other hand, uses satellites to allow users to passively submit longitude and latitude data globally, both indoors and outdoors [17].

In this analysis, we aimed to quantify the effect that the first SARS-CoV-2 vaccination had on travelling behavior, using mobile phone GPS data collected from study participants who consented and voluntarily downloaded the ArcGIS Tracker app onto their mobile phones.

Figure 1. Timeline of restriction events in the United Kingdom during the COVID-19 pandemic from March 2020 to May 2021 (adapted from the Institute for Government [6], which is published under Creative Commons Attribution 4.0 International License [CC BY-NC]).



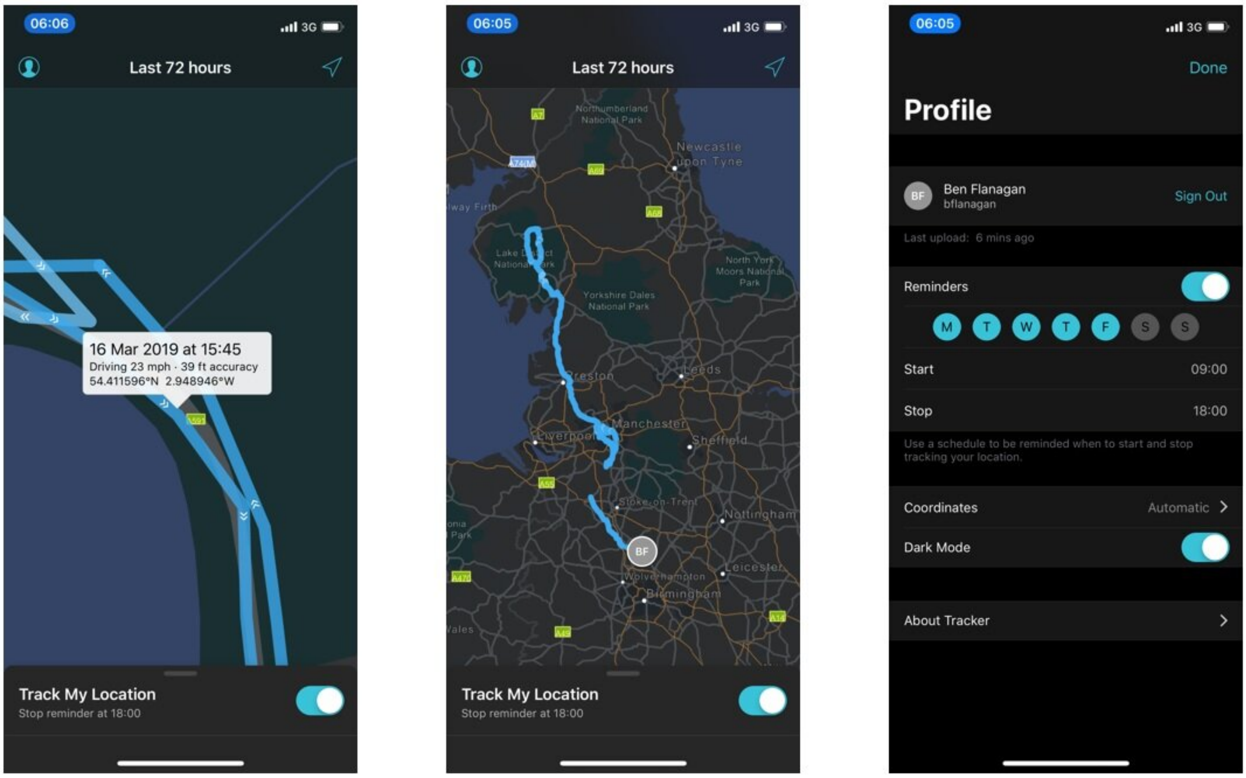
Methods

Study Design and Setting

The study design used prospective observational data from the Virus Watch cohort; a full description of the Virus Watch study has been published previously [18]. Briefly, households were recruited starting in mid-June 2020, which was aimed at creating a representative cohort of England and Wales. To rapidly recruit participants at the start of the pandemic, we used a range of methods aimed at creating a representative cohort of England and Wales. We used the Royal Mail Post Office Address File to generate a random list of residential addresses that were sent recruitment postcards (n=3914), placed social media advertisements on Facebook and Twitter (n=18,594), and sent SMS text messages (n=11,151) and letters to participants and households from their general practitioners (n=3803). We invited a random sample of eligible participants to our tracker subcohort who were adults (aged ≥18 years) on entry that agreed to participate in installing the ArcGIS Tracker on their compatible smartphone and who provided full details on gender, ethnicity, and a registered address.

The ArcGIS Tracker app is available for the Apple iOS (iOS 12 or later) and Google Android (version 5.0 or later) platforms and can be downloaded from their respective app stores [19]. The ArcGIS Tracker app requires users to log in (using provided credentials) and share their location using an on-off toggle button that allows users to control when and where they would like to share their location. The location app is designed to run in the background to collect data; however, on certain models of phones, this function had to be enabled in the smartphones' settings menu. Using the ArcGIS Tracker app, the subcohort provided the following data: date; time; longitude; latitude; travel mode; and GPS accuracy, which was defined by the phone manufacturer's GPS algorithm. See Figure 2 for screenshots of the ArcGIS Tracker app.

Figure 2. Screenshots of the ArcGIS Tracker app. Esri UK [20], Esri [21], HERE [22], Garmin [23], FAO [24], NOAA [25], USGS [26] are data providers for Esri basemaps.



16 Mar 2019 at 15:45
Driving 23 mph · 39 ft accuracy
54.411596°N 2.948946°W

Ben Flanagan
bflanagan

Last upload: 6 mins ago

Reminders
Start 09:00
Stop 18:00
Use a schedule to be reminded when to start and stop tracking your location.
Coordinates Automatic

Intervention

The date of receiving the first dose of vaccine was self-reported through the weekly Virus Watch questionnaire. We began collecting weekly vaccination status on January 11, 2021, and asked about any prior vaccination during the first 2 weekly surveys. Subsequently, participants were asked to provide a weekly update only. The options available were "Pfizer/BioNTech," "Oxford/AstraZeneca," "Moderna," and "Other/Can't remember."

Study Population

The study population included adults (aged ≥18 years on entry) in the subcohort of the Virus Watch study who were vaccinated and submitted at least 10 days of readings. Participants had to submit at least 5 days of readings before and after their self-reported vaccination date. We excluded readings that were outside of England and excluded analysis from days where there were fewer than 5 contributors. We only used location readings with an accuracy rating of less than 30 m.

Study Period

We started sending out invitations for the Tracker cohort between September 2020 and February 2021, with the data extraction for this analysis being undertaken in May 2021.

Outcomes

The unit of the analysis was the aggregated median daily travel distance, with the outcome being the change in median daily travel distance. To calculate the group's median daily travel distance, we aggregated the daily movement from each participant's registered address for each day. The median daily distance was chosen to account for the distribution of the cohort's daily travel patterns.

Analysis

For each individual and each day, we calculated the cumulative outdoor travel distance recorded using the ArcGIS Tracker app from their registered household address. This was calculated by summing up the distance (*d*) computed by the Euclidean distance method (equation 1) between the 2 sequential outdoor GPS records. Considering the accuracy of the GPS records, we set up a 25 m radius buffer zone (the average horizontal accuracy is 25 m) around a participant's home location. Points that fall within the buffer are considered as at-home travel activities and therefore considered as zero distance in analyses.

The equation to calculate the distance travelled by each participant is as follows:

where $d(p_i, p_j)$ is the Euclidean distance between 2 sequential GPS points (ie, p_i and p_j); the Cartesian coordinates are (p_{ix}, p_{iy}) for p_i and (p_{jx}, p_{jy}) for p_j. We used the British National Grid as the reference system.

Our statistical analysis was conducted using an interrupted time series, where we used segmented linear regression to estimate the trends in travel patterns, with the first segment estimating the median travel distance for the cohort before vaccination and the second segment estimating the median travel distance for the cohort after vaccination. Therefore, we defined the interruption time point in our analysis as the date of the first vaccination for each individual, with negative days denoting days prior to vaccination and positive days denoting days after vaccination; for each day, we then calculated the median travel distance.

To calculate the travel trajectory before vaccination, we conducted linear regression analysis using data before vaccination to estimate the sample's median daily travel distance from their home with time (days before vaccination) as the explanatory variable. To calculate the travel trajectory after vaccination, we conducted linear regression analysis using data after vaccination to estimate the sample's median daily travel distance from their home with time (days after vaccination) as the explanatory variable. For both models, each day represented 1 data point, with the points for each day being the median travel distance of those who submitted readings on that day. The segmented regression equations can be found in equation 2; linear regression was chosen a priori as we expected the limitations on movement to create a stable pattern in movement. Our alternative hypothesis was that after vaccination, we would see an increase in movement that would be expressed if $a_2 > a_1$ (a slope change) or $b_2 > b_1$ when $a_2 \geq a_1$ (a level change) [27] in equation 2.

Equation 2 uses segmented linear regression models with model (and subscript) 1 representing the trends before vaccination and model (and subscript) 2 representing the trends after vaccination; y_n represents the estimated median daily travel distance with coefficient a_n, x represents the days since vaccination (negative for model 1 and positive for model 2), and b_n is a constant:

$$y_1 = a_1x + b_1 \text{ for } x < 0 \quad \textbf{(2)}$$

$$y_2 = a_2x + b_2 \text{ for } x > 0$$

The UK vaccination program prioritized people by (older) age and clinical risk groups, which, in addition to differences in the socioeconomic backgrounds between those invited and accepting a vaccination, meant that selecting an appropriate control group for this analysis was not feasible.

Covariates

Due to the study design, which compared the same individuals' movement before and after vaccination, we did not use covariates for regression adjustment. For each eligible individual, we used the following data: days since vaccination and the total travel distance for the corresponding day.

Sensitivity Analyses

We performed various sensitivity analyses. First, after reviewing the data, we repeated the analyses with outliers removed. Three cutoff points were used: days with a median travel distance lower than 50 km, days with a median travel distance lower than 25 km, and days with a median travel distance lower than the median travel distance on the day of vaccination. This third cutoff point was used as the day of vaccination is the 1 day in which we were sure people had travelled (to be vaccinated).

Our second sensitivity analysis accounts for the effect of the removal of national restrictions on movement as alternative explanations for differences in movement after vaccination. We conducted a sensitivity analysis that limited travel and vaccination events to the third national lockdown. This period was from January 4, 2021, to April 5, 2021, and represents a time period when restrictions did not change in relation to rules regarding travel and social distancing.

Ethics Approval

The Virus Watch study was approved by the Hampstead National Health Service Health Research Authority Ethics Committee (20/HRA/2320). All members of participating households provided informed consent for themselves and, where relevant, for children that they were responsible for. To contribute to the tracker subcohort of Virus Watch, adults had to provide explicit consent during our registration process.

Information Governance

This research was registered with the University College London (UCL) data protection office and reviewed by the UCL information security and governance teams. The Virus Watch Data Privacy Impact Assessment can be found on the web [28]. During the consent and registration process, adult participants were invited to contribute geolocation data using the ArcGIS mobile phone tracker app. For those who chose to participate, we sent personal identifiable data from the UCL data safe haven to a secure memory stick on a UCL computer, from which we transferred the data via HTTPS into the ArcGIS Online (Esri UK) subscription. The purpose of this data transfer was to set up participants' tracker app accounts. The transferred data, along with the account passwords, were stored in North America. The UCL Virus Watch study team undertook the data transfer process and had access to the Participant Profile within the ArcGIS Online subscription. Only a small number of named ArcGIS employees had access to the participant profile area and only for the purposes of assisting the UCL study team when necessary. Once the tracker app accounts were created, the UCL study team emailed tracker app participants instructions on how to download the app and sign into the tracker app.

The geolocation data collected by the app were stored securely on a section of the ArcGIS Online subscription hosted in Europe, which is securely cleared every 30 days. Participants' geolocation data were transferred on a regular basis via HTTPS to a secure memory stick on a UCL machine and were then imported via a secure gateway technology onto the UCL secure memory stick and into the UCL data safe haven.

Geolocation data were linked with other participant study data in the data safe haven. Once analyzed, aggregate data (generated from geolocation and other study data) were exported from the data safe haven and published on the public study website and in research publications.

The aggregated data set used in this analysis will be securely destroyed after 20 years, in line with UCL's record retention policy. In line with policies developed for electronic health care research, we did not report any data with a cell containing <5 events, and where necessary, we protected these counts with secondary suppression.

Results

Cohort Demographics

By February 2021, Virus Watch recruited a total of 45,963 individuals, of whom 39,558 were at least 18 years old when registering. Of these 39,558 individuals, 79% (n=31,317) provided consent to install the ArcGIS Tracker app. Of the 31,317 participants, a sample of 13,120 adults (aged ≥18 years on entry) were chosen to be randomly invited based on having complete information on gender, ethnicity, and address details. Of these participants, 2193 contributed at least one GPS reading. After removing invalid data points, including those outside of England and those that did not submit accurate readings (eg, points that exhibit extremely high horizontal and vertical accuracy), 1376 participants were included. Of the 1376 individuals, 1244 individuals were vaccinated by May 2021. After removing individuals with fewer than 5 data points on either side of their vaccination date, 249 individuals were included in our final analysis. See Figure 3 for the Consolidated Standards of Reporting (CONSORT) diagram of how the cohort was analyzed.

Of the 249 participants, there were more women (n=141, 56.6%) than men, with a median age of 62 (IQR 55-67) years, which is older than the whole Virus Watch cohort. Individual residents in local super output areas in the 3 least deprived quintiles represented 79.1% (n=197) of the population. In all, 88.8% (n=221) of our cohort self-identified as “White—English/Welsh/Scottish/Northern Irish/British,” which is more than the whole Virus Watch cohort. See Table 1 for a sociodemographic breakdown of the cohort.

From the 249 participants, there were 157 days of eligible readings prior to the first vaccination, that is, 157 days before vaccination was the earliest day where there were at least 5 people who submitted readings; there were 105 days of readings under the same eligibility criteria after vaccination. The median number of people who contributed per day was 89 (IQR 34.75-135.50) people, with a median of 91 (IQR 35.00-132.00) people contributing before vaccination and a median of 88 (IQR 30.50-149.25) people contributing after vaccination. The median number of days contributed by the 249 participants was 87 (IQR 58.00-128.00) days, with a median contribution of 51 (IQR 26-77) days before vaccination and a median contribution of 36 (IQR 19.00-56.00) days after vaccination. Zero participants produced data for every day and 5 participants produced an equal amount of data before and after vaccination. Before vaccination (–157 days to –1 day), the median daily travel distance was 9.05 (IQR 8.06-10.09) km, and after vaccination (+1 day to 105 days), the median daily travel distance was 10.08 (IQR 8.60-12.42) km. The median travel distance on the day of vaccination was 19.1 (IQR 8.75-37.90) km.

During the first segment of the linear regression model, from 157 days before the first vaccination to the vaccination date, there was a median daily decline of 40.09 m (95% CI –50.08 to –31.10; *P*<.001) of movement with a constant of 6.90 m (95% CI 6.02-7.79; Figure 4A). During the second segment of the linear regression model, from the first vaccination date to 105 days after vaccination, there was a median daily increase of 60.6 m (95% CI 20.9-100; *P*<.001) of movement with a constant of 8.00 m (95% CI 5.59-10.40).

Figure 3. Consolidated Standards of Reporting Trials (CONSORT) diagram for how the cohort was derived.

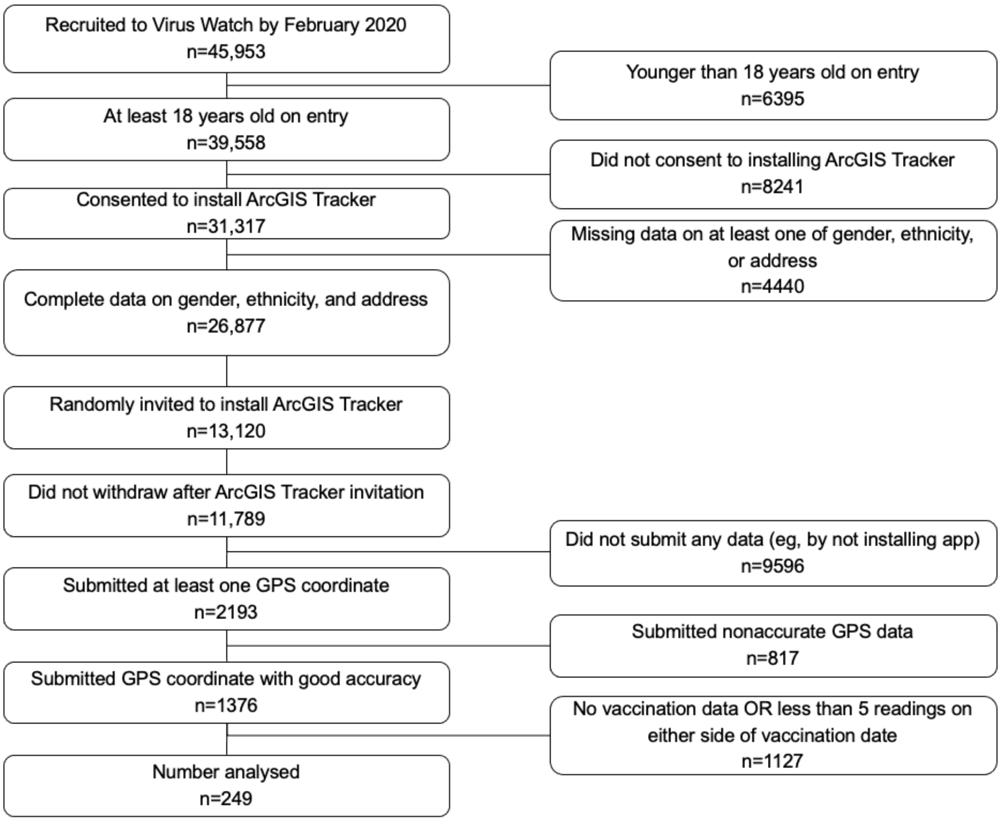
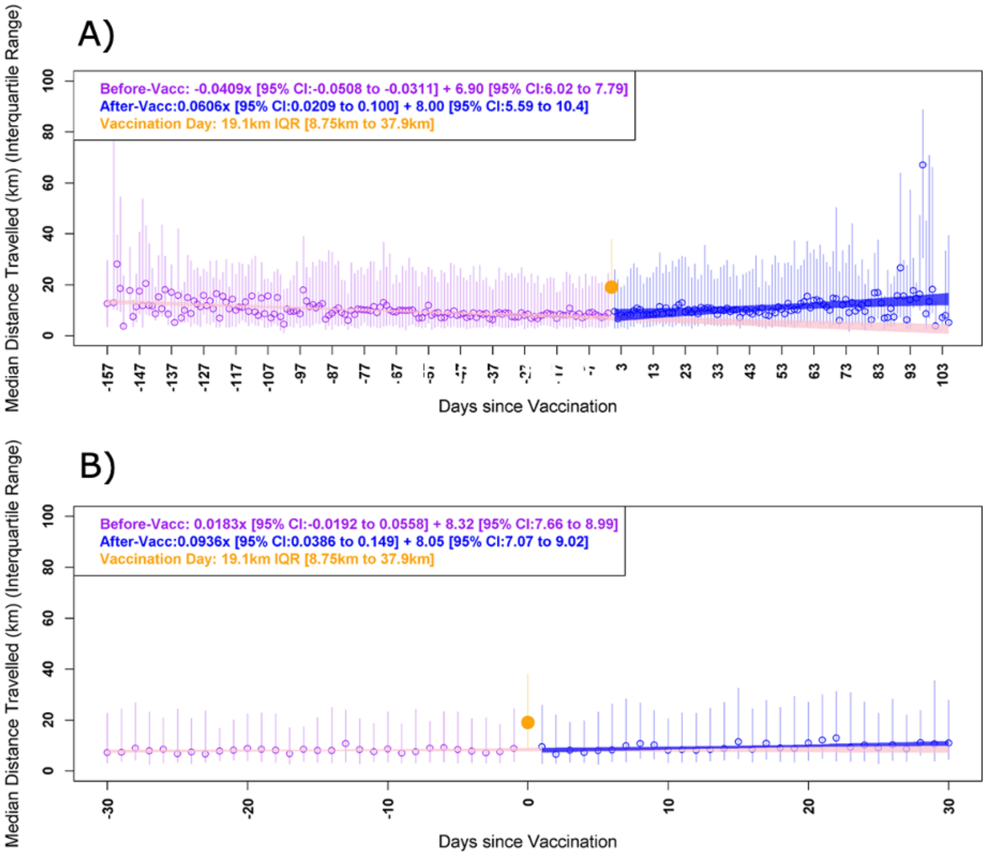


Table 1. Sociodemographic breakdown of the included cohort.

Characteristic	Virus Watch cohort by February 2021 (n=45,963)	Invited to tracker cohort and did not withdraw (n=11,789)	Analyzed (n=249)
Age on entry (years), median (IQR)	54 (34-66)	62 (52-69)	62 (55-67)
Sex, n (%)			
Female	21,625 (47)	6589 (55.9)	141 (56.6)
Intersex	56 (0.1)	21 (0.2)	0 (0)
Male	17,338 (37.7)	5,108 (43.3)	108 (43.4)
Missing	6906 (15)	0 (0)	0 (0)
Prefer not to say	38 (<0.1)	71 (0.6)	0 (0)
Region name, n (%)			
East Midlands	3678 (8)	1012 (8.6)	36 (14.5)
East of England	9052 (19.7)	2321 (19.7)	50 (20.1)
London	6299 (13.7)	1686 (14.3)	40 (16.1)
North East	2118 (4.6)	574 (4.9)	5 (2)
North West	4598 (10)	1320 (11.2)	23 (9.2)
South East	8058 (17.5)	2289 (19.4)	46 (18.5)
South West	2992 (6.5)	871 (7.4)	19 (7.6)
Wales	1038 (2.3)	306 (2.6)	0 (0)
West Midlands	2310 (5)	560 (4.8)	18 (7.2)
Yorkshire and the Humber	2049 (4.5)	634 (5.4)	12 (4.8)
Unknown	3771 (8.2)	216 (1.8)	0 (0)
Index of multiple deprivation quintile, n (%)			
1 (most deprived)	4145 (9)	1048 (8.9)	21 (8.4)
2	6625 (14.4)	1872 (15.9)	31 (12.4)
3	8585 (18.7)	2359 (20)	43 (17.3)
4	10,695 (23.3)	2884 (24.5)	74 (29.7)
5 (least deprived)	12,142 (26.4)	3410 (28.9)	80 (32.1)
Unknown	3771 (8.2)	216 (1.8)	0 (0)
Minority ethnicity status, n (%)			
Minority ethnicity	5381 (11.7)	1295 (11)	28 (11.2)
Missing	7101 (15.4)	0 (0)	0 (0)
Prefer not to say	109 (0.2)	84 (0.7)	0 (0)
White British	33,372 (72.6)	10,410 (88.3)	221 (88.8)

Figure 4. Graphical representation of the interrupted time series. Median daily travel distance from home in reference to vaccination date: (A) unrestricted (primary analysis); (B) only accounting for movement and vaccinations that occurred during the third national lockdown (January 4, 2021 to April 5, 2021) with a 30-day period either side of the vaccination date.



Sensitivity Analysis 1

After reviewing the data, we believe that certain median travel distances were outliers; therefore, we conducted sensitivity analyses to remove days that had readings higher than 50 km, 25 km, and the median distance travelled on the day of vaccination (19.1 km). Under these analyses, all median travel distances before and after vaccination, as well the coefficients and constants, stayed similar (ie, they did not cross the 95% CIs) when compared to the primary analysis. See Table 2 for the coefficients and median travel distances before and after vaccination. See Multimedia Appendices 1-3 for graphical representations of the interrupted time series for these sensitivity analyses.

Table 2. Results of various sensitivity analysis involving the removal of outlier points.

Maximum median distance	Median travel distance before vaccination (km; IQR)	Coefficient before vaccination (95% CI)	Constant before vaccination (95% CI)	Median travel distance after vaccination (km; IQR)	Coefficient after vaccination (95% CI)	Constant after vaccination (95% CI)
Unrestricted (primary analysis)	9.05 (8.06 to 11.34)	–0.041 (–0.051 to –0.031)	6.90 (6.02 to 7.79)	10.08 (8.60 to 12.42)	0.061 (0.021 to 0.100)	8.00 (5.59 to 10.40)
50 km	9.07 (8.06 to 11.34)	–0.041 (–0.051 to –0.031)	6.90 (6.02 to 7.79)	10.07 (8.59 to 12.21)	0.034 (0.014 to 0.055)	8.83 (7.63 to 10.00)
25 km	9.05 (8.05 to 11.32)	–0.037 (–0.046 to –0.028)	7.10 (6.29 to 7.90)	10.07 (8.58 to 12.07)	0.028 (0.011 to 0.046)	9.01 (7.95 to 10.10)
19.1 km (vaccination day)	9.05 (8.05 to 11.29)	–0.035 (–0.044 to –0.026)	7.18 (6.40 to 7.97)	10.07 (8.58 to 12.06)	0.029 (0.011 to 0.046)	9.01 (7.95 to 10.10)

Sensitivity Analysis 2

To account for the effect of national restrictions on movement, we conducted a sensitivity analysis that limited travel and vaccination events to the third national lockdown from January 4, 2021, to April 5, 2021. Due to the asymmetry of the number of data points before and after vaccination in this analysis, we analyzed movement 30 days before vaccination and 30 days after vaccination.

From 30 days prior to vaccination to the vaccination date, there was a median daily movement of 8.06 (IQR 7.49-8.51) km with a median daily movement increase of 18 m (95% CI –19 to 56; *P*=.57). From the vaccination date to the following 30 days (during the third national lockdown), there was a median daily movement of 9.16 (IQR 8.35-10.71) km with a median daily movement increase of 9.35 m (95% CI 39-149; *P*=.69). See Figure 4B for a graphical representation of the interrupted time series for this sensitivity analysis.

Discussion

Principal Findings

Our study demonstrates the feasibility of collecting high-volume geolocation data as part of research projects and the utility of these data for understanding public health issues. Our results require cautious interpretation. Our initial analysis found evidence of a modest increase in the rate of change in median daily distance travelled after participants received their first dose of SARS-CoV-2 vaccine, but when restricting our analysis to a period of lockdown, we did not find evidence of a difference in mobility following 1 vaccination dose. On balance, our results do not provide evidence that people increase the rate of their movements following the first vaccination, as the results are consistent with both an increase and no movement after vaccination and suggest that any change in mobility after vaccination is likely to be modest.

We used GPS data to measure the travel distance of vaccinated individuals. Not only does this improve accuracy over other methods of distance estimation by using the GPS system (as compared to cellular location), but it also reduces recall bias when compared with using self-reported data. Our interrupted time series study design aids in reducing the impact of non–time-varying confounders as the same individual's data are considered before and after vaccination.

Comparisons With Prior Work

Compared to prior work [14], our methodology builds on that work by using techniques that produce more accurate measures of distance by using the GPS system as compared to cell phone towers. That study took a different-in-difference approach with 1 week of data on either side of vaccination and found that there was an increase in movement after vaccination by 8.6% 1 week after vaccination when compared to the week prior to vaccination. Although our sensitivity analysis, where we restricted to 30 days before and after vaccination, found a 13.6% increase in median movement after vaccination, the CIs of our regression models did not support an overall change in movement and are reflective of the overlapping IQR ranges before (median 8.06, IQR 7.49-8.51 km) and after (median 9.16, IQR 8.35-10.71 km) vaccination.

Given that previous studies have suggested people increase their nonhousehold contacts after their first vaccination, further research on behavior change following vaccination is warranted. In the meantime, it is important that public health communications are clear about the differential protection against SARS-CoV-2 infection offered by the first and second doses of the vaccine, such that people can exercise sound personal judgement on how they alter their behavior following vaccination.

Limitations

Our studies' sample size means that we may be underpowered to detect small changes in mobility, particularly when restricting to a period of national lockdown. Our GPS collection was automated but could be switched on and off by participants and is more likely to have been switched off on days when participants stayed at home. The use of the app in this way would result in our analysis overestimating the median distance travelled per day, through the nonreporting of GPS data (eg, switching the app off) on days when participants stayed at home. Due to technological requirements of tracking apps, the results were skewed toward those who had access to a smartphone and were able to contribute their data plan toward research activities, leading to a low initial uptake rate. Furthermore, for those who did provide data, there was an inconsistent provision of data over time, which could be caused by participants switching off the app. With a draining effect on battery life from the GPS app used, the dropout rate from the tracker cohort of Virus Watch was relatively high. People taking part in Virus Watch are self-selecting and motivated to contribute to COVID-19 research, and therefore, their movement patterns may not be generalizable to all vaccinated groups.

An important limitation in our primary analysis is that it does not control for time-varying confounders such as changes in national physical distance rules that were likely to have led to an increase in the rate of change in median daily travel distance. Other time-varying SARS-CoV-2–related events that we could not control for included changes in infection rates, which may have influenced people's decision to travel. As we centered our analysis around the vaccination date, our analysis did not account for different time-varying public health interventions on different vaccination dates. It is also possible that other non–SARS-CoV-2 events such as weather changes may have affected the findings, with participants increasing mobility during this same time period after vaccination as a result of improvements in weather conditions.

Conclusions

Although previous research found that nearly half of those who were vaccinated (after 1 dose) met with others outside their households or support bubbles [8], our findings provide a mixed picture about movement after the first dose of vaccination, and we found no evidence of an increase in movement when we conducted our analyses during a period of national movement restrictions.

Acknowledgments

The research costs for the Virus Watch study have been supported by the Medical Research Council (MRC) grant (MC_PC 19070) awarded to University College London (UCL) on March 30, 2020, and MRC grant (MR/V028375/1) awarded on August 17, 2020. The study also received US $15,000 of Facebook advertising credit to support a pilot social media recruitment campaign on August 18, 2020. Access to the ArcGIS Tracker and supporting software was provided by Esri UK free of charge. The contributors associated with Virus Watch Collaborative are as follows: Alison Rodger, Institute for Global Health, University College London, London, UK; Susan Michie, Centre for Behaviour Change, University College London, London, UK; Linda Wijlaars, Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK; Eleni Nastouli, Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK and Francis Crick Institute, London, UK; Moira Spyer, Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK and Francis Crick Institute, London, UK; Ben Killingley, Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK, Health Protection and Influenza Research Group, Division of Epidemiology and Public Health, University of Nottingham School of Medicine, Nottingham, UK, and University College London Hospital, London, UK; Ingemar Cox, Department of Computer Science, University College London, London, UK; Vasileios Lampos, Department of Computer Science, University College London, London, UK; Rachel A McKendry, London Centre for Nanotechnology and Division of Medicine, London, University College London; Anne M Johnson, Institute for Global Health, University College London, London, UK; Jo Gibbs, Institute for Global Health, University College London, London, UK; and Richard Gilson, Institute for Global Health, University College London, London, UK.

Conflicts of Interest

AH serves on the UK New and Emerging Respiratory Virus Threats Advisory Group. RM, BF, and AP were employed by Esri UK at the time of the study, who provided the ArcGIS Tracker and supporting software free of charge. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Sensitivity Analysis 1: graphical representation of the interrupted time series once we removed days that had readings higher than 50 km.
[PNG File , 73 KB - publichealth_v9i1e38072_app1.png]

Multimedia Appendix 2
Sensitivity Analysis 1: graphical representation of the interrupted time series once we removed days that had readings higher than 25 km.
[PNG File , 73 KB - publichealth_v9i1e38072_app2.png]

Multimedia Appendix 3
Sensitivity Analysis 1: graphical representation of the interrupted time series once we removed days that had readings higher than the median distance travelled on the day of vaccination (19.1 km).
[PNG File , 73 KB - publichealth_v9i1e38072_app3.png]

References

1. Davies NG, Barnard RC, Jarvis CI, Russell TW, Semple MG, Jit M, Centre for Mathematical Modelling of Infectious Diseases COVID-19 Working Group, ISARIC4C investigators. Association of tiered restrictions and a second lockdown with COVID-19 deaths and hospital admissions in England: a modelling study. Lancet Infect Dis 2021 Apr;21(4):482-492 [FREE Full text] [doi: 10.1016/S1473-3099(20)30984-1] [Medline: 33357518]
2. Prime Minister's statement on coronavirus (COVID-19): 23 March 2020. GOV.UK. 2020 Mar 23. URL: https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020 [accessed 2022-08-04]
3. Paterson DC, Ramage K, Moore SA, Riazi N, Tremblay MS, Faulkner G. Exploring the impact of COVID-19 on the movement behaviors of children and youth: a scoping review of evidence after the first year. J Sport Health Sci 2021 Dec;10(6):675-689 [FREE Full text] [doi: 10.1016/j.jshs.2021.07.001] [Medline: 34237456]
4. Pieh C, Budimir S, Delgadillo J, Barkham M, Fontaine JRJ, Probst T. Mental health during COVID-19 lockdown in the United Kingdom. Psychosom Med 2021 May 01;83(4):328-337 [FREE Full text] [doi: 10.1097/PSY.0000000000000871] [Medline: 33009276]
5. Jarvis CI, Gimma A, van Zandvoort K, Wong KLM, CMMID COVID-19 working group, Edmunds WJ. The impact of local and national restrictions in response to COVID-19 on social contacts in England: a longitudinal natural experiment. BMC Med 2021 Feb 19;19(1):52 [FREE Full text] [doi: 10.1186/s12916-021-01924-7] [Medline: 33602244]
6. Timeline of UK government coronavirus lockdowns and restrictions, March 2020 to December 2021. The Institute for Government. 2021. URL: https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns [accessed 2022-08-04]
7. Levine-Tiefenbrun M, Yelin I, Katz R, Herzel E, Golan Z, Schreiber L, et al. Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine. Nat Med 2021 May 29;27(5):790-792. [doi: 10.1038/s41591-021-01316-7] [Medline: 33782619]

8. Coronavirus and vaccine attitudes and behaviours in England: over 80s population, 15 February to 20 February 2021. Office of National Statistics. 2021 Mar 04. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronavirusandvaccineattitudesandbehavioursinengland/over80spopulation15februaryto20february2021 [accessed 2021-06-11]
9. Eyre DW, Lumley SF, Wei J, Cox S, James T, Justice A, et al. Quantitative SARS-CoV-2 anti-spike responses to Pfizer-BioNTech and Oxford-AstraZeneca vaccines by previous infection status. Clin Microbiol Infect 2021 Oct;27(10):1516.e7-1516.e14 [FREE Full text] [doi: 10.1016/j.cmi.2021.05.041] [Medline: 34111577]
10. Shrotri M, Fragaszy E, Geismar C, Nguyen V, Beale S, Braithwaite I, et al. Spike-antibody responses to ChAdOx1 and BNT162b2 vaccines by demographic and clinical factors (Virus Watch study) Internet. medRxiv Preprint posted online on May 20, 2021. [doi: 10.1101/2021.05.12.21257102]
11. Bernal JL, Andrews N, Gower C, Stowe J, Robertson C, Tessier E. Early effectiveness of COVID-19 vaccination with BNT162b2 mRNA vaccine and ChAdOx1 adenovirus vector vaccine on symptomatic disease, hospitalisations and mortality in older adults in England Internet. medRxiv Preprint posted online on March 2, 2022. [doi: 10.1101/2021.03.01.21252652]
12. Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of COVID-19 Vaccines against the B.1.617.2 (Delta) variant. N Engl J Med 2021 Aug 12;385(7):585-594 [FREE Full text] [doi: 10.1056/NEJMoa2108891] [Medline: 34289274]
13. SARS-CoV-2 variants of concern and variants under investigation in England: technical briefing 15. Public Health England. 2021 Jun 11. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993198/Variants_of_Concern_VOC_Technical_Briefing.pdf [accessed 2022-11-24]
14. SPI-B: behavioural considerations for vaccine uptake in Phase 2 and beyond. Scientific Pandemic Insights Group on Behaviours (SPI-B). 2021 Mar 09. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/984351/S1162_SPI-B_-_Behavioural_Considerations_for_Vaccine_Uptake_in_Phase_2.pdf [accessed 2022-11-24]
15. Lauer J. Plastic surveillance: payment cards and the history of transactional data, 1888 to present. Big Data Soc 2020 Apr 03;7(1):205395172090763. [doi: 10.1177/2053951720907632]
16. Pandl KD, Thiebes S, Schmidt-Kraepelin M, Sunyaev A. How detection ranges and usage stops impact digital contact tracing effectiveness for COVID-19. Sci Rep 2021 May 03;11(1):9414 [FREE Full text] [doi: 10.1038/s41598-021-88768-6] [Medline: 33941793]
17. Shahroz M, Ahmad F, Younis MS, Ahmad N, Kamel Boulos MN, Vinuesa R, et al. COVID-19 digital contact tracing applications and techniques: a review post initial deployments. Transportation Engineering 2021 Sep;5:100072. [doi: 10.1016/j.treng.2021.100072]
18. Hayward A, Fragaszy E, Kovar J, Nguyen V, Beale S, Byrne T, et al. Risk factors, symptom reporting, healthcare-seeking behaviour and adherence to public health guidance: protocol for Virus Watch, a prospective community cohort study. BMJ Open 2021 Jun 23;11(6):e048042 [FREE Full text] [doi: 10.1136/bmjopen-2020-048042] [Medline: 34162651]
19. Requirements—ArcGIS Tracker | Documentation. Esri. URL: https://doc.arcgis.com/en/tracker/faq/requirements.htm [accessed 2022-11-24]
20. Esri UK and Ireland. URL: https://www.esriuk.com/en-gb/home [accessed 2023-02-15]
21. Esri . URL: https://www.esri.com/en-us/home [accessed 2023-02-15]
22. HERE. URL: https://www.here.com/platform/map-data [accessed 2023-02-15]
23. Garmin. URL: https://www.garmin.com/en-US/aboutgps/ [accessed 2023-02-15]
24. Food and Agriculture Organization of the United Nations (FAO). URL: https://www.fao.org/home/en [accessed 2023-02-15]
25. National Oceanic and Atmospheric Administration (NOAA). URL: https://www.noaa.gov/ [accessed 2023-02-15]
26. United States Geological Survey (USGS). URL: https://www.usgs.gov/ [accessed 2023-02-15]
27. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017 Feb 01;46(1):348-355 [FREE Full text] [doi: 10.1093/ije/dyw098] [Medline: 27283160]
28. Hayward A, Aldridge R. Virus Watch Study Data Privacy Impact Assessment. Virus Watch. 2021 Feb 19. URL: https://ucl-virus-watch.net/wp-content/uploads/2021/06/Virus-Watch-Study-Data-Privacy-Impact-Assessment-21-02-19.pdf [accessed 2022-07-19]

Abbreviations

CONSORT: Consolidated Standards of Reporting
UCL: University College London

Edited by T Sanchez, A Mavragani; submitted 17.03.22; peer-reviewed by L Kennedy-Shaffer, C Baxter; comments to author 02.05.22; revised version received 05.08.22; accepted 29.09.22; published 08.03.23.

Please cite as:
Nguyen V, Liu Y, Mumford R, Flanagan B, Patel P, Braithwaite I, Shrotri M, Byrne T, Beale S, Aryee A, Fong WLE, Fragaszy E, Geismar C, Navaratnam AMD, Hardelid P, Kovar J, Pope A, Cheng T, Hayward A, Aldridge R, Virus Watch Collaborative
Tracking Changes in Mobility Before and After the First SARS-CoV-2 Vaccination Using Global Positioning System Data in England and Wales (Virus Watch): Prospective Observational Community Cohort Study
JMIR Public Health Surveill 2023;9:e38072
URL: https://publichealth.jmir.org/2023/1/e38072
doi:10.2196/38072
PMID:36884272

Original Paper

COVID-19 Vaccine Acceptance and Uptake in Bangkok, Thailand: Cross-sectional Online Survey

Christopher Remmel[1*], MM; Gaurav Tuli[1*], MS, PhD; Tanner J Varrelman[1], PhD; Aimee R Han[1], MS; Pakkanan Angkab[2], MA; Hathairat Kosiyaporn[2], MSc, MD; Chanikarn Netrpukdee[2], BA; Supatnuj Sorndamrih[3], MM; Thaksaphon Thamarangsi[2], MD, PhD; John S Brownstein[1,4], PhD; Christina M Astley[1,4,5,6], MD, ScD

[1]Computational Epidemiology Lab, Boston Children's Hospital, Boston, MA, United States
[2]International Health Policy Program, Ministry of Public Health, Nonthaburi, Thailand
[3]ThaiHealth Promotion Foundation, Bangkok, Thailand
[4]Faculty of Medicine, Harvard Medical School, Cambridge, MA, United States
[5]Division of Endocrinology, Boston Children's Hospital, Boston, MA, United States
[6]Broad Institute of Harvard and MIT, Cambridge, MA, United States
[*]these authors contributed equally

Corresponding Author:
Christopher Remmel, MM
Computational Epidemiology Lab
Boston Children's Hospital
300 Longwood Ave
Boston, MA, 02115
United States
Phone: 1 617 355 6000
Email: Christopher.Remmel@childrens.harvard.edu

Abstract

Background: The third most severe COVID-19 wave in the middle of 2021 coincided with the dual challenges of limited vaccine supply and lagging acceptance in Bangkok, Thailand. Understanding of persistent vaccine hesitancy during the "608" campaign to vaccinate those aged over 60 years and 8 medical risk groups was needed. On-the-ground surveys place further demands on resources and are scale limited. We leveraged the University of Maryland COVID-19 Trends and Impact Survey (UMD-CTIS), a digital health survey conducted among daily Facebook user samples, to fill this need and inform regional vaccine rollout policy.

Objective: The aims of this study were to characterize COVID-19 vaccine hesitancy, frequent reasons for hesitancy, mitigating risk behaviors, and the most trusted sources of COVID-19 information through which to combat vaccine hesitancy in Bangkok, Thailand during the 608 vaccine campaign.

Methods: We analyzed 34,423 Bangkok UMD-CTIS responses between June and October 2021, coinciding with the third COVID-19 wave. Sampling consistency and representativeness of the UMD-CTIS respondents were evaluated by comparing distributions of demographics, 608 priority groups, and vaccine uptake over time with source population data. Estimates of vaccine hesitancy in Bangkok and 608 priority groups were tracked over time. Frequently cited hesitancy reasons and trusted information sources were identified according to the 608 group and degree of hesitancy. Kendall tau was used to test statistical associations between vaccine acceptance and vaccine hesitancy.

Results: The Bangkok UMD-CTIS respondents had similar demographics over weekly samples and compared to the Bangkok source population. Respondents self-reported fewer pre-existing health conditions compared to census data overall but had a similar prevalence of the important COVID-19 risk factor diabetes. UMD-CTIS vaccine uptake rose in parallel with national vaccination statistics, while vaccine hesitancy and degree of hesitancy declined (−7% hesitant per week). Concerns about vaccination side effects (2334/3883, 60.1%) and wanting to wait and see (2410/3883, 62.1%) were selected most frequently, while "not liking vaccines" (281/3883, 7.2%) and "religious objections" (52/3883, 1.3%) were selected least frequently. Greater vaccine acceptance was associated positively with wanting to "wait and see" and negatively with "don't believe I need (the vaccine)" (Kendall tau 0.21 and −0.22, respectively; adjusted $P<.001$). Scientists and health experts were most frequently cited as trusted COVID-19 information sources (13,600/14,033, 96.9%), even among vaccine hesitant respondents.

Conclusions: Our findings provide policy and health experts with evidence that vaccine hesitancy was declining over the study timeframe. Hesitancy and trust analyses among the unvaccinated support Bangkok policy measures to address vaccine safety and efficacy concerns through health experts rather than government or religious officials. Large-scale surveys enabled by existing widespread digital networks offer an insightful minimal-infrastructure resource for informing region-specific health policy needs.

(JMIR Public Health Surveill 2023;9:e40186) doi:10.2196/40186

KEYWORDS

COVID-19 vaccines; Thailand; survey; vaccines; COVID-19; pandemic; public health; health policy; epidemiology; social media; vaccine hesitancy

Introduction

The third COVID-19 wave in Thailand, spanning June to October 2021, presented a serious challenge to health infrastructure [1-3]. A rapid increase in COVID-19 severity and spread coincided with slow vaccine rollout to nonhealth care workers. Vaccine distribution was further impeded by limited supply and lagging demand. Thailand initially targeted vaccination of the “608” group that included high-risk people defined as those aged 60 years or older and those with 1 of 8 pre-existing health conditions (PHCs) such as diabetes, respiratory disease, or pregnancy [4,5]. The government offered expanded vaccine supply and options [6,7], and focused institutional support in conjunction with incentives and, in some instances, punitive measures [8-10]. Nevertheless, vaccine hesitancy remained a persistent public health problem in Thailand [11], and yet, studies of vaccine hesitancy trends in the region were sparse. There was an urgent need to understand these trends, understand the reasons for vaccine hesitancy, and understand how best to persuade the unvaccinated in Thailand.

Vaccine hesitancy literature on the psychological and sociological drivers of this public health issue identified common correlates such as lack of social pressure, complacency toward the disease, lack of trust in medicine, false beliefs about vaccination, lack of trust in the government, and conspiratorial thinking [12]. Specific reasons for vaccine hesitancy in Thailand were often connected to fears of side effects and of lack of benefits, most notably regarding the Sinovac vaccine offered by the government early in vaccine rollout [13,14]. However, these surveys were either narrowly focused on particular subgroups, such as health care workers or seniors, or were otherwise offered for short windows of time [13,15-20]. Policy and public health officials working to change sentiment and behaviors in the region might use data from more detailed longitudinal studies inclusive of broader population demographics.

The University of Maryland Global COVID Trends and Impact Survey (UMD-CTIS) is an innovative data stream covering the health attitudes and practices of a large population sample (the global Facebook active user base [FAUB]) from early in the COVID-19 pandemic (April 2020). A large proportion of the population of Thailand, especially Bangkok, is represented in the FAUB. As such, the UMD-CTIS was well suited to address the limitations of prior surveys and fill the need for context-specific information to combat vaccine hesitancy in the region. Leveraging UMD-CTIS responses from the residents of Bangkok, the most populous and most densely populated urban area with the highest COVID-19 burden [21], we sought to understand vaccine uptake, attitudes, and opportunities to better address vaccine messaging in the 608 vaccine rollout era.

Methods

Study Design

UMD-CTIS

This research is based on survey responses from the UMD-CTIS, which has been approved by the University of Maryland (UMD) Institutional Review Board (1587016-10) and has been described previously [22]. Briefly, the UMD-CTIS, in partnership with Facebook, is a cross-sectional survey of daily samples of the FAUB population. Sampled FAUB users were invited to participate in the UMD-CTIS through a special banner. Respondents aged ≥18 years who consented to participate in the UMD-CTIS study completed an online Qualtrics survey administered by the UMD [23]. Though statistical resampling of users from the FAUB over time is possible, these resampled FAUB users cannot be identified, and their UMD-CTIS responses cannot be linked longitudinally by design.

Survey Instrument

The UMD-CTIS instrument was designed to evaluate a range of public health topics through a single cross-sectional survey requiring limited time burden. Survey questions were updated by the UMD periodically in response to evolving epidemiologic needs. A consistent survey instrument (version 11) was used during the study period described below. This instrument covered vaccination status and vaccine hesitancy, in addition to questions on demographics, knowledge, attitudes, health practices, and health status. One of two additional survey modules was offered randomly to each respondent, either module A (media trust) or module B (PHCs). Survey logic and language (English and Thai) for version 11 are summarized in Multimedia Appendix 1, and are available online [24].

Because survey questions administered to each respondent target different subject matter by design and responses from the resampled FAUB cannot be linked, possible dependence of observations and measures of internal consistency are not available. Assuming that the large FAUB population, statistical sampling scheme, and response/participation rates are stable relative to the study timescale, respondents are a consistent sample of the FAUB population. Respondent characteristics (eg, age and gender) may be used as a measure of sampling response consistency for those characteristics. Trends in survey-estimated metrics (eg, vaccination proportion among

UMD-CTIS respondents) are assumed to be estimates of trends in the source population for the FAUB population. The UMD-CTIS compared to external benchmark data may be used as a measure of the representativeness of FAUB respondents relative to the source population of interest for those survey questions.

UMD-CTIS Study Population in Bangkok, Thailand

This study analyzed UMD-CTIS responses from self-identified residents of Bangkok, Thailand from June 14, 2021, to October 4, 2021. This period coincided with the third COVID-19 wave and the 608 campaign rollout. To evaluate for possible differences among those targeted in the 608 campaign, we separately analyzed 2 subgroups of survey responses representative of those groups, that is, respondents self-reporting older age and self-reporting at least one of the targeted PHCs.

The UMD-CTIS in conjunction with Facebook provided survey weights, which combine design weights (for disproportional population stratification sampling), nonresponse weights (inverse of response propensity), and poststratification weights (standardizing to regional age-gender distributions), to simultaneously adjust for regional demographics, nonresponse bias, and sampling bias, as described previously in greater detail [23]. The individual component weights are not provided by design. Thus, when evaluating subregional data (eg, PHC subgroups), it is not possible to apply only the design and nonresponse weights, without also applying poststratification weights for the region. Estimates with and without survey weights were nevertheless similar both regionally and in subgroups, and are presented for comparison in Multimedia Appendices 2-4.

Covariates and Outcome Measures

Demographics and 608 Campaign Groups

Respondents self-reported demographic characteristics, including 4 gender and 7 age categories. UMD-CTIS age categories were in 10-year brackets, which did not align with the 60-year-old threshold in the 608 campaign. To avoid including subjects aged 55 to 59 years, the elderly risk group was identified as those aged ≥65 years (ie, selected 65-74 or ≥75 years UMD-CTIS age categories). When evaluating age-gender distributions, gender was limited to binary responses ("male" and "female"). Those with self-reported targeted PHCs in the 608 campaign were identified from the subpopulation of respondents randomly offered survey module B. Those who received module B did not receive module A. We included surveys from those who received module B and reported at least one of the targeted PHCs on module B. Pregnancy was only queried among those not reporting male gender. The 608 campaign PHCs were mapped to reasonable UMD-CTIS proxies of those PHCs. These included pregnancy, diabetes, obesity, cancer, chronic kidney disease, chronic respiratory diseases (asthma and chronic obstructive pulmonary disease combined), and cardiovascular diseases (heart attack, heart disease, or other heart conditions). The 608 campaign PHCs of neurovascular diseases did not have reasonable proxies in the UMD-CTIS and were not evaluated.

Vaccine Uptake

We evaluated vaccine uptake in the UMD-CTIS in Bangkok and the 608 subpopulations relative to the publicly reported vaccine uptake measures. Self-reported vaccination status was ascertained using the questions "Have you had a COVID-19 vaccination?" and "How many COVID-19 vaccinations have you received?" Complete vaccination was assessed with the response "Vaccinated, two doses," which is consistent with the primary formulations available widely during the study period. Additionally, we evaluated uptake trends by assessing the responses "Vaccinated, one dose" and "Scheduled" (for vaccination) to the question "Do you have an appointment to receive a COVID-19 vaccine?" Vaccine uptake in the region was trended over the course of the study period.

Vaccine Acceptance and Hesitancy

The degree of vaccine acceptance was characterized as a nominal scale with decreasing willingness to be vaccinated. The 3 groups with the highest vaccine acceptance were those fully vaccinated ("Vaccinated, two doses"), partially vaccinated ("Vaccinated, one dose"), and about to be vaccinated ("Scheduled"). The remaining unvaccinated and unscheduled respondents were categorized into 4 groups of increasing hesitancy to become vaccinated according to their responses to the question "If a vaccine to prevent COVID-19 was offered to you today, would you choose to get vaccinated?" The 4 groups were as follows: "Definitely," "Probably," "Probably not," and "Definitely not." Multimedia Appendix 1 details the survey questions, response options, and survey logic for the following questions: "Have you had a COVID-19 vaccination?" "How many COVID-19 vaccinations have you received?" "Do you have an appointment to receive a COVID-19 vaccine?" and "If a vaccine to prevent COVID-19 was offered to you today, would you choose to get vaccinated?" Surveys with missing responses to enable categorization into these 7 groups were excluded.

Reasons for Vaccine Hesitancy

Reasons for vaccine hesitancy were examined among the 4 most vaccine hesitant subgroups ("Definitely," "Probably," "Probably not," and "Definitely not"). The question stem text varied slightly to align with the respondents' self-identified degree of willingness to be vaccinated (ie, "Which of the following, if any, are reasons that you definitely wouldn't/probably wouldn't/only probably would choose to get a COVID-19 vaccine?"). Hesitancy reasons were presented as multichoice responses and included a range of options from concerns about vaccine side effects to beliefs about vaccine efficacy (Multimedia Appendix 1).

Reasons for Not Believing Vaccination is Necessary

Hesitant respondents who endorsed "I don't believe I need a vaccine" were further questioned for their reasons for holding that belief. Proportions were calculated for each individual reason and for the incidence of co-selected pairs in this multichoice response.

Trusted Media for COVID-19 Information

Potential avenues for communicating with vaccine hesitant individuals in the region were evaluated among the 4 most vaccine hesitant subgroups and those who accepted vaccination in some form (scheduled or received any number of doses). Survey module A surveyed preferred sources of information relating to COVID-19. This module was presented to a random subset of respondents; those who received module A did not receive module B. Subjects were asked "How much do you trust the following sources to provide accurate news and information about COVID-19?" The 3 response options "Trust," "Somewhat trust," and "Do not trust" were dichotomized by combining the first two into a single "Trust" option. For each respondent, "trusted sources" are therefore sources of COVID-19 information that respondents indicated they "Somewhat trust" or "Trust."

External Data and Benchmarking

To compare the Bangkok UMD-CTIS estimates and trends to external benchmark measures of population-level statistics, we compiled population demographics, PHCs, and vaccine data from publicly available sources. Further methodology, data sources, and results are provided in Multimedia Appendix 5. To compare the Bangkok UMD-CTIS vaccine uptake relative to government-reported trends over the 608 campaign study period, we estimated the proportion of fully vaccinated (2-dose) respondents in the UMD-CTIS over time. The daily sample size of the Bangkok UMD-CTIS is small in comparison to the rate of change in the source population vaccinated proportion. Thus, to show temporal trends in vaccine uptake in the Bangkok UMD-CTIS study over time, 28-day rolling averages of the proportion vaccinated are shown in relation to the national daily reports. Bangkok vaccination data over time were obtained from the Thailand Department of Disease Control daily reports [25]. Complete vaccination per population was calculated using data on 2 doses administered as described above, consistent with the daily vaccination reports.

Statistical Analyses

Estimates and CIs for proportions were calculated with the Wilson Score method, using the count of affirmative responses per total survey in the subgroup of interest, unless noted otherwise. For multichoice questions (eg, PHCs and hesitancy reasons), we used the count for that choice item. Survey-weighted and raw proportions are shown in Multimedia Appendix 5, though subgroup demographics differed from Bangkok, Thailand demographics from which UMD-CTIS weights were, in part, derived. A test of the trend for demographic differences between the UMD-CTIS and census over time was conducted using linear regression and applying Bonferroni correction for multiple comparisons to generate an adjusted *P* value. Among hesitant respondents, we evaluated associations between the degree of vaccine acceptance and reasons for hesitancy by using Kendall tau and applying Bonferroni correction for multiple comparisons. We repeated this analysis for 608 campaign group subsets.

Ethics Approval

The Institutional Review Board of Boston Children's Hospital (P00023700) approved this study that used the UMD-CTIS.

Results

Bangkok UMD-CTIS

Demographics of the Bangkok UMD-CTIS

Table 1 shows that the Bangkok UMD-CTIS respondents (N=36,334) were more frequently male and under 65 years of age. There were 1265 (3.5%) respondents in the ≥65 years elderly risk group targeted by the 608 campaign. Of the 19,734 respondents who received the PHC survey module, 3407 (17.3%) reported having at least one of the 7 PHCs covered by the survey. Diabetes (1090/19,734, 5.5%) and obesity (1555/19,734, 7.9%) were the most frequently reported PHCs targeted in the 608 campaign. Comparisons of the Bangkok UMD-CTIS and 608 subgroups to benchmark census data, along with survey weight-adjusted estimates, are presented in Multimedia Appendix 5. Multimedia Appendix 6 shows that demographics were consistent over time, while in Multimedia Appendix 7, linear regression results for demographic differences between the UMD-CTIS and census showed that adjusted *P* values never met the threshold for significance and coefficients were always close to zero. While Bangkok UMD-CTIS respondents were similar to the Bangkok general population in terms of age, gender, and diabetes status, Bangkok UMD-CTIS respondents, who were selected from the Bangkok FAUB population, indicated that they were healthier than the general population with respect to other COVID-19 risk factors.

Table 1. Demographic and pre-existing health condition characteristics of survey respondents in the cross-sectional survey (June 14, 2021, to October 4, 2021).

Characteristic	Value, n (%)
Age group (years) (N=36,334)	
18-24	2759 (7.6)
25-34	7641 (21.0)
35-44	8907 (24.5)
45-54	6439 (17.7)
55-64	3569 (9.8)
65-74	1102 (3.0)
≥75	163 (0.4)
Not reported	5754 (15.8)
Sex (N=36,334)	
Male	16,308 (44.9)
Female	13,468 (37.1)
Other	331 (0.9)
Prefer not to answer	477 (1.3)
Not reported	5750 (15.8)
PHC[a,b] (N=19,734)	
At least one condition	3407 (17.3)
Obesity	1555 (7.9)
Diabetes	1090 (5.5)
Chronic respiratory diseases[c]	641 (3.2)
Asthma	575 (2.9)
Cardiovascular diseases	488 (2.5)
Chronic kidney disease	207 (1.0)
Cancer	171 (0.9)
Chronic lung diseases	155 (0.8)
Pregnancy[d]	92 (0.3)
Vaccine uptake and hesitancy (N=36,334)	
Vaccinated, 2 doses	8596 (23.7)
Vaccinated, 1 dose; unspecified dose	13,712 (37.7)
Scheduled	5505 (15.2)
Definitely willing	2909 (8.0)
Probably willing	2619 (7.2)
Probably not willing	652 (1.8)
Definitely not willing	522 (1.4)

[a] PHC: pre-existing health condition.

[b] Of the 36,334 respondents, 19,734 (54.3%) received the module about PHCs and 16,600 (45.7%) did not receive the module.

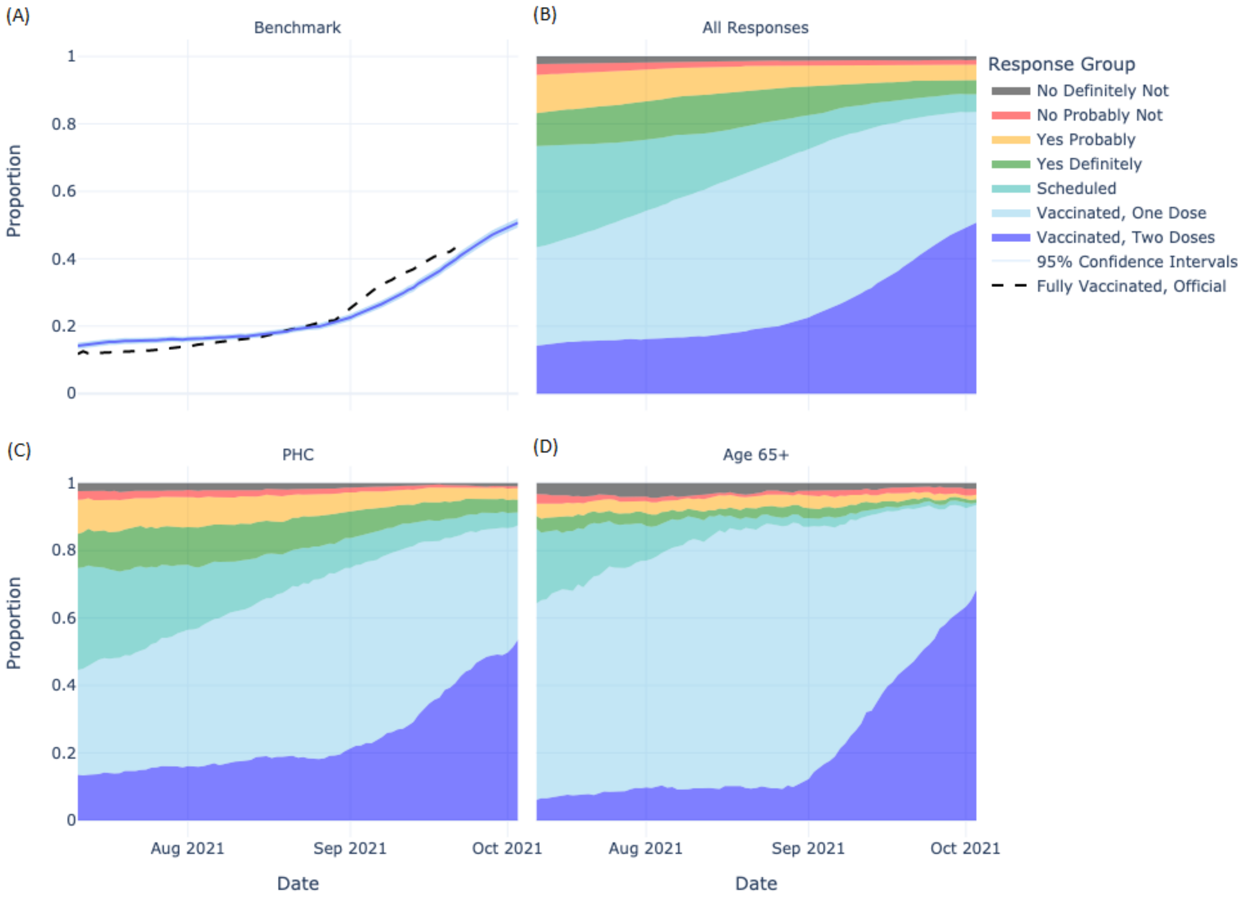
[c] Includes respondents who self-identified being diagnosed with asthma or chronic lung diseases such as chronic obstructive pulmonary disease, chronic bronchitis, or emphysema.

[d] A total of 7405 (25.0%) responded “No.”

Vaccine Uptake

In Figure 1, we show vaccine uptake tracked over the 608 campaign study period using the proportion of Bangkok UMD-CTIS respondents who indicated they were scheduled to or had received one or two doses of the COVID-19 vaccine. Vaccination trends in 2-dose vaccinated respondents closely mirrored Bangkok government statistics (Figure 1A). The proportion scheduled to be vaccinated declined (−14.0%/week) as the proportion vaccinated with at least one dose rose. While elderly people received their first dose more quickly than others (158/246, 64.2% vs 3305/7617, 43.4% with at least one dose at the start of the study period), uptake of the second dose increased more gradually during the summer. Uptake patterns among those with PHCs were not noticeably different from responses overall. The 95% CIs were very small owing to the large sample size (Multimedia Appendix 8). The UMD-CTIS–weighted trends were similar.

Figure 1. Vaccine uptake and acceptance in Bangkok over time and within 608 vaccination campaign priority groups from June 14, 2021, to October 4, 2021. (A) The 4-week moving average government vaccination uptake trend (dashed line) compared against the fully vaccinated trend in the COVID-19 Trends and Impact Survey overall (dark blue line) with 95% CIs (light blue line). (B-D) Stacked proportion of respondents indicating vaccine acceptance (2-dose vaccinated, dark blue; 1-dose vaccinated, blue; scheduled, turquoise; definitely, green; probably, yellow; probably not, red; definitely not, black) across all responses (B), and further faceted across pre-existing health condition (PHC) (C) and elderly (D) subgroups.



Vaccine Acceptance and Hesitancy

To evaluate vaccine hesitancy over time, we classified unvaccinated and unscheduled respondents according to their degree of vaccine acceptance. The proportions of those who said that they would probably, probably not, and definitely not get vaccinated steadily declined, though more gradually (−7.3%/week) compared to the rise in vaccine uptake (Figure 1). Vaccination hesitancy did not increase over the study period overall or within the 608 targeted groups (Figure 1). However, the rise in vaccine uptake appeared to mirror changes in those who were scheduled to receive vaccination and accepting of vaccination, rather than a decline in those hesitant to do so.

Reasons for Vaccine Hesitancy

We evaluated the reasons cited by the 3 groups of hesitant respondents who indicated they would probably, probably not, or definitely not get vaccinated (N=3883). Figure 2 shows that reasons relating to the risk-benefit of vaccination were more frequently selected among hesitant respondents. Respondents reported concerns over side effects (2434/3883, 62.7%), wanting to wait and see for longer to determine whether the vaccines are safe (2410/3883, 62.1%), and not feeling sure that the vaccines will be effective in protecting against COVID-19 (1407/3883, 36.2%). Concerns about side effects and lack of vaccine benefit were the most frequently selected together with wanting to wait and see (45.8% and 27.9%, respectively).

Figure 3 shows that when stratifying according to respondents' vaccine acceptance, risk-benefit concerns were similar in frequency. In Table 2, we see that the strength of the association of endorsing a specific concern with acceptance of vaccination was small, with Kendall tau mostly below ±0.2, even though adjusted *P* values were mostly <.001 from the large survey size. The most hesitant were more averse to vaccination ("Don't believe I need," tau=−0.22; "Don't like vaccines," tau=−0.13) and were less likely to consider vaccination later ("Wait and see," tau=+0.2). The most hesitant were nevertheless unsure of the benefits of vaccination, similar to the less hesitant groups ("Don't know if it will work," tau=−0.01). Those remaining unvaccinated and hesitant at the end of the study period (August 27 to October 3, 2021) selected "Don't believe I need" more frequently than at the start of the study period, while "Don't know if it will work" did not shift over time.

In Multimedia Appendix 9, we see that among the 608 campaign groups, the reasons for hesitancy patterns were broadly similar to the patterns for all respondents. The unvaccinated and hesitant respondents aged ≥65 years selected "Don't believe I need" more frequently but were similarly most often concerned about "side effects."

Figure 2. Frequency matrix of hesitancy reasons in the cross-sectional survey (June 14, 2021, to October 4, 2021). Intersection of 2 reasons is the frequency that both reasons were chosen by a respondent. Intersection of a reason with itself is the frequency for that single reason across all respondents.

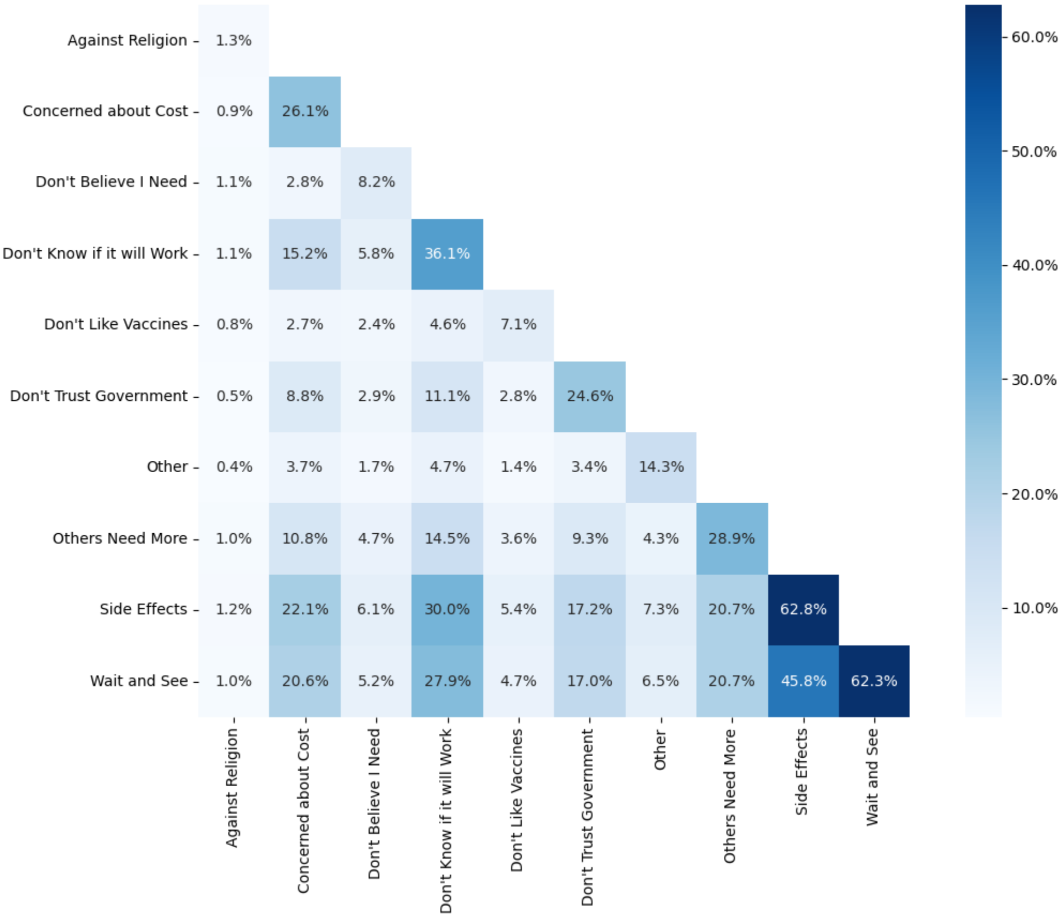


Figure 3. Reasons for hesitancy by degree of vaccine acceptance in the cross-sectional survey (June 14, 2021, to October 4, 2021).

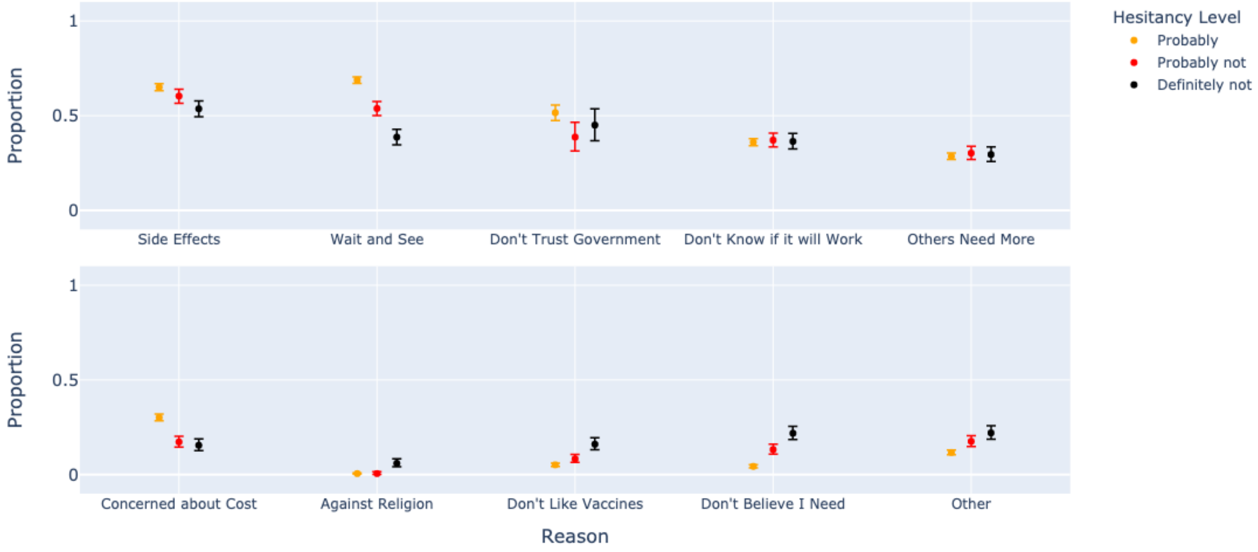


Table 2. Association between vaccine acceptance and individual reasons for hesitancy measured by Kendall tau in the cross-sectional survey (June 14, 2021, to October 4, 2021).

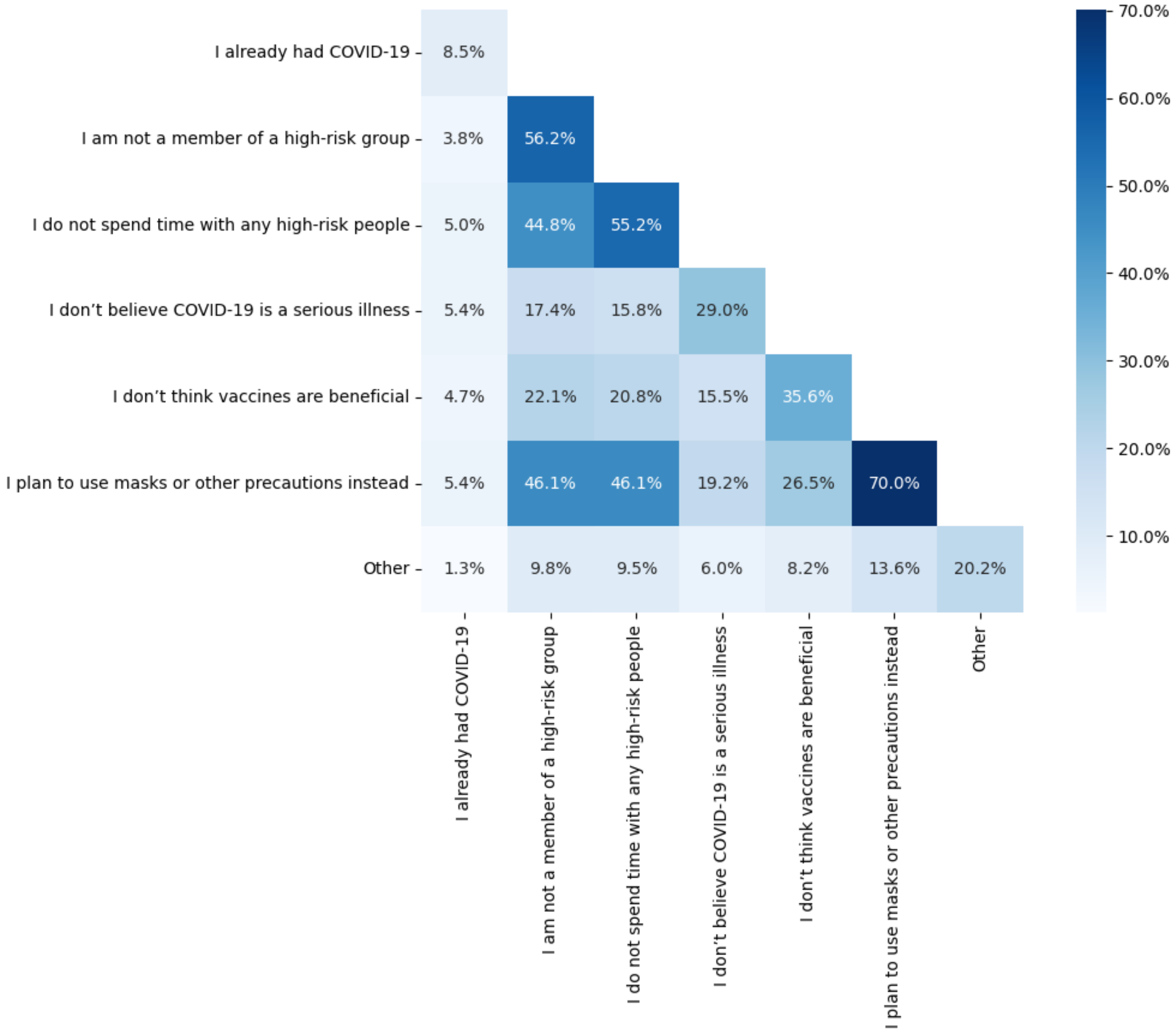
Hesitancy reason	Kendall tau	Adjusted *P* value
Wait and see	0.21	<.001
Concerned about cost	0.14	<.001
Side effects	0.08	<.001
Don’t know if it will work	−0.01	>.99
Others need it more	−0.01	>.99
Other	−0.1	<.001
Against religion	−0.11	<.001
Don’t like vaccines	−0.13	<.001
Don’t believe I need	−0.22	<.001

Reasons for Not Believing Vaccination is Necessary

Given the consistent endorsement of “I don’t believe I need a COVID-19 vaccine” among the unvaccinated and most hesitant respondents across risk groups and over time, we further evaluated the behaviors and attitudes of respondents who selected this reason for hesitancy. In Figure 4, we see that those who did not believe they needed a COVID-19 vaccine (N=317) generally endorsed other protective behaviors and beliefs overall and within the 608 campaign groups. They most frequently reported their intention to mask or pursue other mitigation strategies instead of vaccination (222/317, 70.0%), and least frequently indicated low confidence in vaccines (113/317, 35.6%) or in COVID-19 as a serious illness (92/317, 29.0%).

Over half (35/60, 58%) of unvaccinated hesitant respondents who did not believe they needed a COVID-19 vaccine did not endorse being a member of a high-risk group, even though they self-reported older age or a PHC in the 608 campaign. Nevertheless, Multimedia Appendix 10 shows that most respondents who indicated they were not in a risk group also planned to use protective measures other than vaccines, and infrequently indicated they did not think COVID-19 was a serious illness. Thus, while unvaccinated hesitant people in the 608 campaign groups may not be aware of their underlying COVID-19 risk, such as older age or diabetes status, they do appear to have an appreciation of the risk and an understanding of important risk mitigation measures.

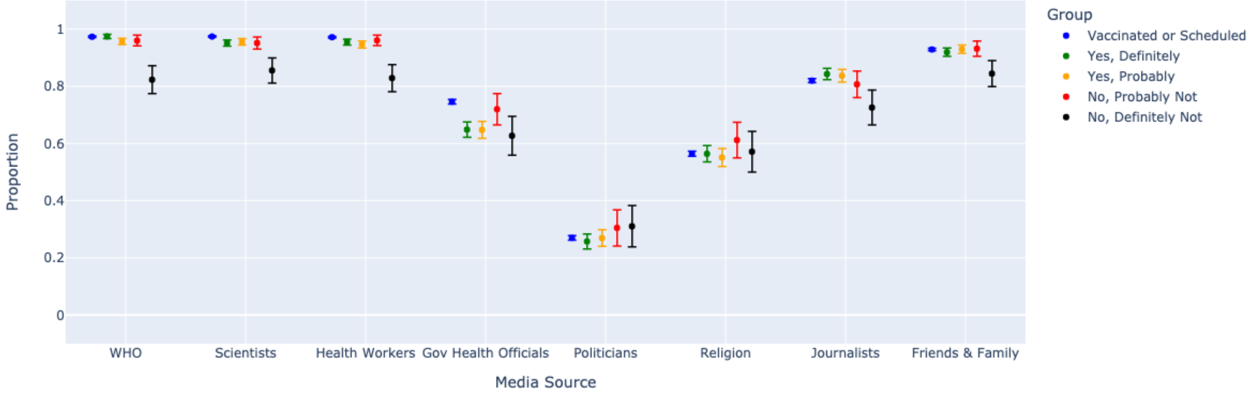
Figure 4. Frequency of selected multichoice responses for why individuals chose "I don't believe I need a COVID-19 vaccine" in the cross-sectional survey (June 14, 2021, to October 4, 2021).



Trusted Media for COVID-19 Information

Given vaccine hesitancy concerns and respondents indicating they might vaccinate later if there is more information, we sought to investigate what information sources would best reach this group. Medical and scientific representatives were the most trusted sources of information on COVID-19 among all respondents, regardless of the degree of hesitancy. Specifically, in Figure 5, we see that respondents who completed each of these questions most frequently indicated that they trust the World Health Organization (13,564/13,980, 97.0%), scientists/health experts (13,600/14,033, 96.9%), and local health workers (13,827/14,304, 96.7%) for news relating to the pandemic. Politicians (3723/13,794, 27.0%) and religious leaders (7668/13,794, 55.6%) were least frequently selected as trusted. The most hesitant respondents were less trusting overall but still trusted scientists and other health experts more than other sources.

Figure 5. Trusted sources of information on COVID-19 by grade of vaccine hesitancy. Frequency of selected multichoice options for trusted media sources, combining responses for most trusted and somewhat trusted.



Discussion

Principal Findings

Through leveraging daily online health surveys conducted among the residents of Bangkok, Thailand, we found that concerns about the risks and benefits of COVID-19 vaccines were paramount during the third wave, but trust in scientific sources remained intact. Globally, the pandemic and response polarized many populations. In many regions, this led to distrust of scientific evidence and interventions, and eroded the relationship among people, politicians, and public health officials [26]. We showed that, in contrast to the antiscience and antivaccine sentiment identified elsewhere [27], most unvaccinated hesitant respondents in Bangkok still believed in the seriousness of COVID-19, in the benefits of mitigation measures, and in the information relayed by health officials. The high professional trust in health experts and scientists in Thailand offered a valuable support mechanism for the government to effectively manage response efforts [28,29].

In a systematic review of the factors associated with vaccine hesitancy [26], the most important predictors included low perceived risk of infection or sequelae, low trust in institutions, low rates of influenza vaccination, and concerns regarding vaccine side effects. Prior surveys in Thailand identified unique concerns about the efficacy of vaccines from specific manufacturers, especially Sinovac [14,19]. Our findings of risk-benefit concerns among the unvaccinated echo these themes. Additionally, we identified that many hesitant respondents were planning to wait for additional information on vaccine safety and that the specific distrust of all vaccines is low (10% vs <20% in the United States [30]). Thus, messaging from health experts emphasizing vaccine safety and efficacy, and updates as more evidence on Thai-specific vaccine options becomes available, would be most effective in reaching the most hesitant unvaccinated respondents in our survey.

Among respondents within the ≥65 years and PHC risk groups, vaccine hesitancy and reasons for hesitancy were broadly similar to the general Bangkok population. Over half (35/60, 58%) of the unvaccinated respondents who provided reasons for not believing that they needed to be vaccinated and who self-reported older age and PHCs, however, did not self-identify as being at risk for COVID-19. While older age, pregnancy, and diabetes may be well-understood by a patient, other diagnoses may be less so. The burden of disease estimates for conditions, such as renal disease [31], may depend on access to testing (eg, biopsy) or may only be understood by the patient as a medical condition or risk factor when especially severe (eg, requiring medication or dialysis). This highlights the importance of 608 campaign messaging to educate citizens about the top priority vaccination groups, with the support of community health workers.

Being able to understand the unique needs of Bangkok in the context of the 608 campaign and third COVID-19 wave underscores the value of using insights from digital health surveys, such as the UMD-CTIS, to understand interventions, inform policy, and lessen the pandemic impact. Responses to health surveys may not be generalizable due to selection and response bias, and this may be compounded by sampling from the FAUB of social media users [32]. However, most Thai people have Facebook social media accounts [33], and we showed that the demographics of UMD-CTIS respondents aligned with the demographics of the Bangkok population. Thus, while there may be some bias toward respondents with more health access, awareness, or interest, the efficacy of daily sampling of the FAUB and the consistent survey instrument combined with the minimal need for on-the-ground infrastructure make this data stream a powerful resource. Additionally, observed slower uptake during the summer lines up with the time period when a longer gap between doses was recommended [34], which demonstrates the survey's capability to identify and measure the effects of health policy changes. This analysis provides region-specific supporting evidence of resilience and adaptability in the face of a rapidly evolving public health threat and serves as a helpful validation of government efforts with support from public health experts.

Limitations

The analyses presented here have several potential limitations. The global UMD-CTIS instrument was developed during the pandemic to measure the most critical epidemiologic measures, knowledge, attitudes, and practices in near real time. The instrument is limited by a lack of internal and external validation metrics and is thus susceptible to measurement bias [35]. Redundant questions were not included to conduct within-survey internal consistency measures. Repeat respondents could not be paired by design, and as such, the independence of observations cannot be guaranteed, which could make estimates of standard error overconfident. We found that the UMD-CTIS was consistent with 3 external benchmarks, and prior work validated other UMD-CTIS measures [36,37]. As discussed above, PHCs other than diabetes were likely undermeasured by the instrument or underrepresented in the survey sample. However, the notable variation in differences from population estimates across the PHCs suggests that other factors may also be contributing to underestimation, for example, patient understanding of the nuances of chronic kidney disease stages relative to that of the diagnosis of diabetes, which might also affect a subject's vaccine acceptance. While digital health surveys sampled from a social media platform user base may not be generalizable to all Bangkok residents, this limitation also applies to other previously used survey methods conducted in the region to understand vaccine hesitancy, such as telephone and door-to-door surveys [17]. There is possible selection bias as subgroups that are less comfortable with social media and consumer technology use, such as very elderly people, may be underrepresented, which would mean less coverage of especially vulnerable groups. In contrast, marginalized populations, such as migrant workers affected by the case surges in the third wave [38], may be fearful of in-person interviews, but may be more comfortable responding to an anonymous survey through social media. A cross-sectional survey, such as the UMD-CTIS, may be subject to ecological bias, especially if the source population changes over time or there is confounding by time of the response pairs. Because participation in public health surveys is voluntary, there may be response bias if respondents have more of an interest in public health than the population at large, and as such, attitudes may be shifted away from hesitancy or lack of interest in mitigation. We observed relatively stable trends in respondent demographics and hesitancy attitudes over time, even with continued uptake of vaccination. The sampling scheme of the UMD-CTIS, being nested within an existing social media user base, may be less susceptible to this if the membership in the FAUB and response rate are stable over the time period under investigation.

Conclusions

We showed that vaccine hesitancy in Bangkok during the third wave was connected to an interest in more health expert information about vaccine efficacy and safety. The UMD-CTIS findings were concordant with smaller studies highlighting skepticism about specific vaccine options available during rollout. Importantly, hesitancy among unvaccinated respondents in the UMD-CTIS was not associated with a broader distrust in science or the public health system. This may be a valuable resource for future studies to understand the regional landscapes of vaccine hesitancy in the context of specific policy interventions such as the 608 campaign in Thailand.

Acknowledgments

The survey data used were collected by the University of Maryland in partnership with Facebook and were accessed by Boston Children's Hospital investigators via the University of Maryland microdata portal. The study sponsors had no role in the analysis and interpretation of the data, in the writing of the report, and in the decision to submit the paper for publication. This work was supported by the Facebook Sponsored Research Agreement INB1116217 (CR, GT, TJV, ARH, JSB, and CMA).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of survey questions, answers, and logic for variables used in this study.

[DOCX File , 30 KB - publichealth_v9i1e40186_app1.docx]

Multimedia Appendix 2

Unweighted vs weighted vaccine uptake and acceptance in Bangkok over time. Four-week moving average for the stacked proportion of respondents indicating grade of vaccine uptake and acceptance from fully vaccinated (dark blue) to definitely would not vaccinate (black). Government vaccination uptake trend (dashed line) is also visible in the fully vaccinated trend overall and further faceted across unweighted and weighted data.

[PNG File , 56 KB - publichealth_v9i1e40186_app2.png]

Multimedia Appendix 3

Unweighted vs weighted hesitancy reasons among all respondents. Unweighted (circle) and weighted (open circle) frequencies of selected responses for vaccine hesitancy of multichoice items overall.

[PNG File , 46 KB - publichealth_v9i1e40186_app3.png]

Multimedia Appendix 4

Unweighted vs weighted trusted sources of information on COVID-19 by grade of vaccine hesitancy. Unweighted (circle) vs weighted (line) frequencies of selected multichoice options for trusted media sources, combining responses for most trusted and somewhat trusted.

[PNG File , 48 KB - publichealth_v9i1e40186_app4.png]

Multimedia Appendix 5

Demographic and pre-existing health condition characteristics of survey respondents and that of the general population in Bangkok.

[DOCX File , 21 KB - publichealth_v9i1e40186_app5.docx]

Multimedia Appendix 6

Demographic differences by week between census and University of Maryland COVID-19 Trends and Impact Survey data.

[PNG File , 86 KB - publichealth_v9i1e40186_app6.png]

Multimedia Appendix 7

Results for simple linear regression for week versus difference in the proportion of each of the 14 age-gender groups in University of Maryland COVID-19 Trends and Impact Survey versus Bangkok census demographics.

[DOCX File , 19 KB - publichealth_v9i1e40186_app7.docx]

Multimedia Appendix 8

Vaccine uptake by group with CIs. Four-week moving average for the proportion of respondents faceted by the grade of vaccine uptake and acceptance from fully vaccinated to definitely would not vaccinate. Grey lines denote the 95% CI.

[PNG File , 100 KB - publichealth_v9i1e40186_app8.png]

Multimedia Appendix 9

Reasons for hesitancy by grade of vaccine uptake and acceptance, and by the 608 vaccination campaign priority groups. Frequency of selected responses for vaccine hesitancy of multichoice items overall (circle) and within pre-existing health conditions (open diamond) or elderly people (open square), as well as within subgroups defined by degree of hesitancy (probably, orange; probably not, red; definitely not, black).

[PNG File , 80 KB - publichealth_v9i1e40186_app9.png]

Multimedia Appendix 10

Reasons for not needing vaccination in the 608 vaccination campaign priority groups. Frequency of selected multichoice responses for why individuals chose “I don't believe I need a COVID-19 vaccine” as a hesitancy reason. Nonexclusively grouped by all responses (circle), pre-existing health condition (open diamond), and age 65+ years (open square).

[PNG File , 64 KB - publichealth_v9i1e40186_app10.png]

References

1. Kunno J, Supawattanabodee B, Sumanasrethakul C, Wiriyasivaj B, Kuratong S, Kaewchandee C. Comparison of different waves during the COVID-19 pandemic: Retrospective descriptive study in Thailand. Adv Prev Med 2021;2021:5807056 [FREE Full text] [doi: 10.1155/2021/5807056] [Medline: 34659835]
2. How Thailand rode out COVID waves only to be hit by ‘virus tsunami’. Thai PBS World. 2021. URL: https://www.thaipbsworld.com/how-thailand-rode-out-covid-waves-only-to-be-hit-by-virus-tsunami/ [accessed 2022-02-18]
3. Thailand struggles to contain a new coronavirus wave. France 24. 2021. URL: https://www.france24.com/en/asia-pacific/20210613-thailand-struggles-to-contain-a-new-coronavirus-wave [accessed 2022-02-18]
4. Strategy: Managing the new wave of the Covid-19 Epidemic. Ministry of Public Health. 2021. URL: https://ddc.moph.go.th/viralpneumonia/eng/file/main/en_Thailand%20Covid-19%20plan_MOPH_2021.pdf [accessed 2021-12-07]
5. The COVID-19 vaccination guideline 2021 in Thailand (second revision). Ministry of Public Health. 2021. URL: https://ddc.moph.go.th/vaccine-covid19/getFiles/11/1628849610213.pdf [accessed 2021-12-07]
6. Thawan T. Thailand to import 5 million Moderna vaccine doses for private hospitals. Thaiger. 2021. URL: https://thethaiger.com/news/national/thailand-to-import-5-million-moderna-vaccine-doses-for-private-hospitals [accessed 2022-03-09]
7. Pfizer and Moderna set to be part of Thailand’s vaccine mix. Royal Thai Embassy. 2021. URL: https://thaiembdc.org/2021/06/23/pfizer-and-moderna-set-to-be-part-of-thailands-vaccine-mix/ [accessed 2022-03-09]
8. COVID-19 Vaccination Compulsory in Buriram Province. Thai News. 2021. URL: https://thainews.prd.go.th/en/news/detail/TCATG210516122632355 [accessed 2022-02-18]
9. Phuket Sandbox Programme. Royal Thai Embassy in the Hague, Netherlands. 2021. URL: https://web.archive.org/web/20211214143514/https://hague.thaiembassy.org/th/content/phuket-sandbox [accessed 2021-12-07]
10. Three schemes set to welcome visitors. Bangkok Post. URL: https://www.bangkokpost.com/business/2202063/three-schemes-set-to-welcome-visitors [accessed 2021-12-17]
11. Thai PM promises more vaccine amid criticism of slow rollout. Reuters. 2021. URL: https://www.reuters.com/world/asia-pacific/thai-pm-promises-more-vaccine-amid-criticism-slow-rollout-2021-04-21/ [accessed 2022-02-18]

12. Romate J, Rajkumar E, Gopi A, Abraham J, Rages J, Lakshmi R, et al. What contributes to COVID-19 vaccine hesitancy? A systematic review of the psychological factors associated with COVID-19 vaccine hesitancy. Vaccines (Basel) 2022 Oct 22;10(11):1777 [FREE Full text] [doi: 10.3390/vaccines10111777] [Medline: 36366286]
13. Thanapluetiwong S, Chansirikarnjana S, Sriwannopas O, Assavapokee T, Ittasakul P. Factors associated with COVID-19 Vaccine Hesitancy in Thai Seniors. Patient Prefer Adherence 2021;15:2389-2403 [FREE Full text] [doi: 10.2147/PPA.S334757] [Medline: 34754180]
14. Boon-Itt S, Rompho N, Jiarnkamolchurn S, Skunkan Y. Interaction between age and health conditions in the intention to be vaccinated against COVID-19 in Thailand. Hum Vaccin Immunother 2021 Dec 02;17(12):4816-4822 [FREE Full text] [doi: 10.1080/21645515.2021.1979378] [Medline: 34613887]
15. Sathitakorn O, Jantarathaneewat K, Weber DJ, Warren DK, Apisarnthanarak A. Factors associated with intensified infection prevention and vaccination practice among Thai health care personnel: A multicenter survey during COVID-19 pandemic. Am J Infect Control 2022 Jun;50(6):704-706 [FREE Full text] [doi: 10.1016/j.ajic.2021.12.011] [Medline: 34971716]
16. Osterrieder A, Cuman G, Pan-Ngum W, Cheah PK, Cheah P, Peerawaranun P, et al. Economic and social impacts of COVID-19 and public health measures: results from an anonymous online survey in Thailand, Malaysia, the UK, Italy and Slovenia. BMJ Open 2021 Jul 20;11(7):e046863 [FREE Full text] [doi: 10.1136/bmjopen-2020-046863] [Medline: 34285007]
17. Impact of COVID-19 on Thailand’s Households : Insights from a Rapid Phone Survey (English). World Bank Group. 2021. URL: http://documents.worldbank.org/curated/en/390821633313679563/Impact-of-COVID-19-on-Thailand-s-Households-Insights-from-a-Rapid-Phone-Survey [accessed 2022-02-28]
18. Sirikalyanpaiboon M, Ousirimaneechai K, Phannajit J, Pitisuttithum P, Jantarabenjakul W, Chaiteerakij R, et al. COVID-19 vaccine acceptance, hesitancy, and determinants among physicians in a university-based teaching hospital in Thailand. BMC Infect Dis 2021 Nov 22;21(1):1174 [FREE Full text] [doi: 10.1186/s12879-021-06863-5] [Medline: 34809607]
19. Kitro A, Sirikul W, Piankusol C, Rirermsoonthorn P, Seesen M, Wangsan K, et al. Acceptance, attitude, and factors affecting the intention to accept COVID-19 vaccine among Thai people and expatriates living in Thailand. Vaccine 2021 Dec 20;39(52):7554-7561 [FREE Full text] [doi: 10.1016/j.vaccine.2021.11.031] [Medline: 34819224]
20. Yoda T, Suksatit B, Tokuda M, Katsuyama H. Analysis of people's attitude toward COVID-19 vaccine and its information sources in Thailand. Cureus 2022 Feb;14(2):e22215 [FREE Full text] [doi: 10.7759/cureus.22215] [Medline: 35308704]
21. Coronavirus disease 2019 (COVID-19) WHO Thailand Situation Report 196 - 12 August 2021 [EN/TH]. ReliefWeb. 2021. URL: https://reliefweb.int/report/thailand/coronavirus-disease-2019-covid-19-who-thailand-situation-report-196-12-august-2021 [accessed 2022-03-08]
22. Kreuter F, Barkay N, Bilinski A, Bradford A, Chiu S, Eliat R, et al. Partnering with a global platform to inform research and public policy making: What needs to be in place to make a global COVID-19 survey work? Survey Research Methods 2020;14(2):159-163. [doi: 10.18148/srm/2020.v14i2.7761]
23. Barkay N, Cobb C, Eilat R, Galili T, Haimovich D, LaRocca S, et al. Weights and Methodology Brief for the COVID-19 Symptom Survey by University of Maryland and Carnegie Mellon University, in Partnership with Facebook. arXiv. 2020. URL: http://arxiv.org/abs/2009.14675 [accessed 2022-02-18]
24. COVID Map. URL: https://covidmap.umd.edu/fbsurvey/ [accessed 2022-02-18]
25. Department of Disease Control, Ministry of Public Health. URL: https://ddc.moph.go.th/dcd/ [accessed 2022-02-18]
26. Pires C. Global predictors of COVID-19 vaccine hesitancy: A systematic review. Vaccines (Basel) 2022 Aug 18;10(8):1349 [FREE Full text] [doi: 10.3390/vaccines10081349] [Medline: 36016237]
27. Lee KW, Gew LT, Siau CS, Peh SC, Chia YC, Yacob S, et al. COVID-19 vaccine hesitancy and its associated factors in Malaysia. PLoS One 2022 Sep 1;17(9):e0266925 [FREE Full text] [doi: 10.1371/journal.pone.0266925] [Medline: 36048822]
28. Saechang O, Yu J, Li Y. Public trust and policy compliance during the COVID-19 pandemic: The role of professional trust. Healthcare (Basel) 2021 Feb 02;9(2):151 [FREE Full text] [doi: 10.3390/healthcare9020151] [Medline: 33540762]
29. A breakdown in trust is eroding Thai social cohesion. Nikkei Asia. URL: https://asia.nikkei.com/Opinion/A-breakdown-in-trust-is-eroding-Thai-social-cohesion [accessed 2022-05-04]
30. COVID Data Tracker. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker [accessed 2022-11-04]
31. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020 Feb 29;395(10225):709-733 [FREE Full text] [doi: 10.1016/S0140-6736(20)30045-3] [Medline: 32061315]
32. Safdar N, Abbo LM, Knobloch MJ, Seo SK. Research methods in healthcare epidemiology: Survey and qualitative research. Infect Control Hosp Epidemiol 2016 Nov;37(11):1272-1277 [FREE Full text] [doi: 10.1017/ice.2016.171] [Medline: 27514583]
33. Leesa-Nguansuk S. Thailand makes top 10 in social media use. Bangkok Post. 2018. URL: https://www.bangkokpost.com/business/1420086/thailand-makes-top-10-in-social-media-use [accessed 2022-03-09]
34. Thailand to extend dose gap for AstraZeneca vaccine to 16 weeks. Reuters. 2021. URL: https://www.reuters.com/business/healthcare-pharmaceuticals/thailand-extend-dose-gap-astrazeneca-vaccine-16-weeks-2021-05-24/ [accessed 2022-02-28]
35. Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases? Nephrology (Carlton) 2020 Jun 27;25(6):435-441 [FREE Full text] [doi: 10.1111/nep.13706] [Medline: 32133725]

36. Astley CM, Tuli G, Mc Cord KA, Cohn EL, Rader B, Varrelman TJ, et al. Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. Proc Natl Acad Sci U S A 2021 Dec 21;118(51):e2111455118 [FREE Full text] [doi: 10.1073/pnas.2111455118] [Medline: 34903657]
37. Baltrusaitis K, Santillana M, Crawley AW, Chunara R, Smolinski M, Brownstein JS. Determinants of participants' follow-up and characterization of representativeness in flu near you, A participatory disease surveillance system. JMIR Public Health Surveill 2017 Apr 07;3(2):e18 [FREE Full text] [doi: 10.2196/publichealth.7304] [Medline: 28389417]
38. Quinley C. Thailand: Migrants plead for vaccines as COVID takes lives, jobs. Al Jazeera. 2021. URL: https://www.aljazeera.com/news/2021/9/14/thailand-migrants-plead-for-vaccines-as-covid-takes-lives-jobs [accessed 2023-03-17]

Abbreviations

FAUB: Facebook active user base
PHC: pre-existing health condition
UMD: University of Maryland
UMD-CTIS: University of Maryland COVID-19 Trends and Impact Survey

Edited by A Mavragani, T Sanchez; submitted 09.06.22; peer-reviewed by A Rovetta, T Cruvinel; comments to author 14.07.22; revised version received 04.11.22; accepted 22.02.23; published 13.04.23.

Please cite as:
Remmel C, Tuli G, Varrelman TJ, Han AR, Angkab P, Kosiyaporn H, Netrpukdee C, Sorndamrih S, Thamarangsi T, Brownstein JS, Astley CM
COVID-19 Vaccine Acceptance and Uptake in Bangkok, Thailand: Cross-sectional Online Survey
JMIR Public Health Surveill 2023;9:e40186
URL: https://publichealth.jmir.org/2023/1/e40186
doi:10.2196/40186
PMID:36811852

Original Paper

Accuracy of Self-Reported COVID-19 Vaccination Status Compared With a Public Health Vaccination Registry in Québec: Observational Diagnostic Study

Patrick M Archambault[1,2,3,4], MSc, MD; Rhonda J Rosychuk[5], PhD; Martyne Audet[2], MSc; Rajan Bola[6], MSc; Shabnam Vatanpour[7], PhD; Steven C Brooks[8], MHSc, MD; Raoul Daoust[9], MSc, MD; Gregory Clark[10], MD; Lars Grant[10,11], MD, PhD; Samuel Vaillancourt[12,13], MPH, MD; Michelle Welsford[14], MD; Laurie J Morrison[15,16], MSc, MD; Corinne M Hohl[6], MHSc, MD; Canadian COVID-19 Emergency Department Rapid Response Network (CCEDRRN) investigators[17]; Network of Canadian Emergency Researchers[18]; Canadian Critical Care Trials Group[19]

[1]Department of Family Medicine and Emergency Medicine, Université Laval, Québec, QC, Canada

[2]Centre de recherche intégrée pour un système de santé apprenant en santé et services sociaux, Centre intégré de santé et services sociaux de Chaudière-Appalaches, Lévis, QC, Canada

[3]VITAM - Centre de recherche en santé durable, Québec, QC, Canada

[4]Department of Anesthesiology and Intensive Care, Université Laval, Québec, QC, Canada

[5]Department of Pediatrics, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada

[6]Department of Emergency Medicine, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

[7]Center for Health Informatics, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[8]Department of Emergency Medicine, Queen's University, Kingston, ON, Canada

[9]Department of Family Medicine and Emergency Medicine, Université de Montréal, Montréal, QC, Canada

[10]Department of Emergency Medicine, Faculty of Medicine, McGill University, Montréal, QC, Canada

[11]Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, QC, Canada

[12]Department of Emergency Medicine, Li Ka Shing Knowledge Institute, St Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

[13]Department of Medicine, University of Toronto, Toronto, ON, Canada

[14]Division of Emergency Medicine, Department of Medicine, McMaster University, Hamilton, ON, Canada

[15]Division of Emergency Medicine, Department of Medicine, University of Toronto, Toronto, ON, Canada

[16]Department of Emergency Services, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

[17]University of British Columbia, Vancouver, BC, Canada

[18]Canadian Association of Emergency Physicians, Ottawa, ON, Canada

[19]Canadian Critical Care Trials Group, Toronto, ON, Canada

Corresponding Author:
Patrick M Archambault, MSc, MD
Department of Family Medicine and Emergency Medicine
Université Laval
Pavillon Vandry, Avenue de la Médecine
Québec, QC, G1V 0A6
Canada
Phone: 1 4189552552
Email: patrick.archambault@fmed.ulaval.ca

Abstract

Background: The accuracy of self-reported vaccination status is important to guide real-world vaccine effectiveness studies and policy making in jurisdictions where access to electronic vaccine registries is restricted.

Objective: This study aimed to determine the accuracy of self-reported vaccination status and reliability of the self-reported number of doses, brand, and time of vaccine administration.

Methods: This diagnostic accuracy study was completed by the Canadian COVID-19 Emergency Department Rapid Response Network. We enrolled consecutive patients presenting to 4 emergency departments (EDs) in Québec between March 24, 2020, and December 25, 2021. We included adult patients who were able to consent, could speak English or French, and had a proven COVID-19 infection. We compared the self-reported vaccination status of the patients with their vaccination status in the electronic Québec Vaccination Registry. Our primary outcome was the accuracy of the self-reported vaccination status (index test) ascertained during telephone follow-up compared with the Québec Vaccination Registry (reference standard). The accuracy was calculated by dividing all correctly self-reported vaccinated and unvaccinated participants by the sum of all correctly and incorrectly self-reported vaccinated and unvaccinated participants. We also reported interrater agreement with the reference standard as measured by unweighted Cohen κ for self-reported vaccination status at telephone follow-up and at the time of their index ED visit, number of vaccine doses, and brand.

Results: During the study period, we included 1361 participants. At the time of the follow-up interview, 932 participants reported at least 1 dose of a COVID-19 vaccine. The accuracy of the self-reported vaccination status was 96% (95% CI 95%-97%). Cohen κ for self-reported vaccination status at phone follow-up was 0.91 (95% CI 0.89-0.93) and 0.85 (95% CI 0.77-0.92) at the time of their index ED visit. Cohen κ was 0.89 (95% CI 0.87-0.91) for the number of doses, 0.80 (95% CI 0.75-0.84) for the brand of the first dose, 0.76 (95% CI 0.70-0.83) for the brand of the second dose, and 0.59 (95% CI 0.34-0.83) for the brand of the third dose.

Conclusions: We reported a high accuracy of self-reported vaccination status for adult patients without cognitive disorders who can express themselves in English or French. Researchers can use self-reported COVID-19 vaccination data on the number of doses received, vaccine brand name, and timing of vaccination to guide future research with patients who are capable of self-reporting their vaccination data. However, access to official electronic vaccine registries is still needed to determine the vaccination status in certain susceptible populations where self-reported vaccination data remain missing or impossible to obtain.

Trial Registration: Clinicaltrials.gov NCT04702945; https://clinicaltrials.gov/ct2/show/NCT04702945

(JMIR Public Health Surveill 2023;9:e44465) doi:10.2196/44465

KEYWORDS

electronic vaccination registry; self-reported vaccination status; COVID-19; accuracy; diagnostic study; interrater agreement

Introduction

Background

Since the beginning of the COVID-19 pandemic, vaccination campaigns have been instrumental in preventing mortality and morbidity related to COVID-19 worldwide [1]. With over 13 billion doses of vaccines administered worldwide [2], public health policy makers around the world have been relying on evidence from clinical trials to guide decision-making and develop vaccination strategies to prevent the spread and consequences of this pandemic [3]. In addition to evidence from controlled clinical trials, real-world evidence regarding vaccine effectiveness is important to guide policy making [4,5] and patient decision-making [6]. In particular, real-world evidence can help understand vaccine effectiveness against different disease severities in different susceptible populations often excluded from clinical trials [7], inform policy makers about the length of vaccine protection over time [8], and document rarer side effects [9]. Real-world data can also help understand the evolving prevalence and underlying causes of vaccine hesitancy over time [10] and shape future vaccination campaigns [11,12]. Understanding the role of herd immunity and how to promote vaccination are important questions that need reliable vaccination data [13-16]. Studying COVID-19 occupational hazards for professionals working with susceptible populations also requires timely and reliable access to vaccination data [17,18]. As new variants emerge, real-world evidence informs current and future strategies to protect against COVID-19 infection [19]. To measure real-world vaccine effectiveness and inform evidence-based health promotion strategies, researchers and policy makers alike need access to reliable vaccination data. Unfortunately, not all researchers and policy makers have access to official electronic public health vaccine registries [20-23], while others have substantial challenges in securing access to individual-level vaccination data even when they exist [24]. Many researchers and policy makers must then rely on patients' self-reported status to measure real-world vaccine effectiveness, and if this is inaccurate, it may bias vaccine effectiveness estimates in test-negative designs [25]. For these studies, not only are accurate data about vaccination status key, but, most importantly, reliable data about when vaccine doses were administered are essential to determine the effectiveness in preventing infections and health service use such as emergency department (ED) visits [26].

Objective

To inform the conduct of real-world vaccine effectiveness studies using patient self-reported vaccine information and to support the development of evidence-based public health promotion policies, this study aimed to compare the accuracy of self-reported vaccination status for COVID-19 with data from an official electronic vaccination registry in the province of Québec, Canada [27].

Methods

Study Design and Setting

The Canadian COVID-19 Emergency Department Rapid Response Network (CCEDRRN; pronounced "SED-rin") is a national collaboration created to harmonize data collection related to COVID-19 in over 50 EDs across 8 provinces in Canada (British Columbia, Alberta, Manitoba, Saskatchewan, Ontario, Québec, New-Brunswick, and Nova Scotia) [28]. The

CCEDRRN has created a national observational database of suspected and confirmed patients with COVID-19 who presented to acute care EDs. Data from this network are used in national studies to inform decision-making, emerging treatment guidelines, resource allocation, and planning in response to the pandemic. As of April 2023, CCEDRRN had captured data on over 208,000 patients across Canada [29]. In December 2020, CCEDRRN initiated data collection on vaccination status that supports this study.

This CCEDRRN study is a nested diagnostic accuracy study within the larger CCEDRRN cohort that was reported using the 2015 Standards for Reporting of Diagnostic Accuracy studies [30]. This study only used CCEDRRN data collected from 4 sites in the province of Québec because access to government electronic vaccine registries in the other provinces of Canada was unavailable. These 4 Québec CCEDRRN sites only collected data on patients with a positive COVID-19 test. The study enrolled consecutive consenting patients presenting to these 4 participating EDs between March 24, 2020, and December 25, 2021.

Vaccination in Québec started on December 14, 2020, with long-term care residents and health care workers [31]. On March 1, 2021, vaccination in the general population started in sequence based on their age group. On April 14, 2021, more than 2 million Québecers had received at least 1 dose, representing a quarter of the population. On May 18, 2021, more than half of the population had received at least 1 dose, and on June 6, more than 75% of the population aged >12 years had received at least 1 dose. On July 5, 2021, Québec reduced the dosing interval to 4 weeks to maximize 2-dose coverage, such that by September 30, 2021, a total of 75% of adults were considered fully vaccinated. On September 28, 2021, a third booster dose was recommended for residents in long-term care and assisted living. Between November 16, 2021, and December 29, 2021, the booster dose was progressively authorized for specific high-risk subgroups and then in sequence for the general population based on age groups. After each dose was administered in Québec, vaccinators were required to document the vaccine brand and the time of administration directly within the provincial electronic registry [32]. The identities of the vaccinated patients were validated using at least 2 personal identification questions. Another key public health measure to consider for this study was the mandate on September 1, 2021, to implement a mandatory vaccination passport that contained information about the number of doses, vaccine brand names, and vaccination dates to access certain public services in Québec [31].

Ethics Approval and Informed Consent

The ethics committee at the *Centre intégré de santé et de services sociaux de Chaudière-Appalaches* approved this project in February 2021 (MP-23-2021-766) with a waiver of informed consent for retrospective data collected; however, informed consent was obtained during telephone follow-up for all participants. The participants did not receive any financial compensation.

Participants

Patients were included if they had a confirmed COVID-19 infection (≥1 nucleic acid amplification tests positive for SARS-CoV-2 from specimens collected during the index ED visit, 14 days before the ED visit in the community, or within 14 days after admission) [28]. Included patients were aged ≥18 years, were able to communicate in French or English, and consented independently to telephone follow-up and the use of their health data. All included patients sought care at any of the 4 participating Québec sites (Hôtel-Dieu de Lévis, Hôpital Sacré-Coeur de Montréal, Royal Victoria Hospital, and the Montreal General Hospital). Patients were excluded if they were unable to provide consent (eg, had died since the ED encounter), were unable to communicate in French or English, could not be reached after 5 attempts, or did not have a record in the Québec Vaccination Registry (eg, out-of-province patients).

Data Collection and Outcome Measures

Trained research assistants collected data for our study from different sources and at different time periods. First, starting on July 9, 2020, we used a retrospective chart review to extract demographics, the time and date of ED visits, chief complaint, laboratory tests, treatments, admission status, and comorbidities. We retrospectively collected data for the index visits of all patients and any subsequent ED visits at least 30 days after the index visit. Second, starting on March 24, 2021, we initiated data collection from the Québec Vaccination Registry and phone follow-ups. A single research assistant at each site was designated to access the Québec Vaccination Registry to determine patients' vaccination status, vaccine brands received, the total number of doses, and dates of vaccination. Separate research assistants at each site were trained to only conduct telephone follow-ups. They were blinded to the reason for collecting self-reported vaccination data and did not access the information collected from the Québec Vaccination Registry. These research assistants contacted patients by telephone to obtain consent to use their data; to obtain additional sociodemographic data not available in their medical records (eg, income, self-identified race, and education level); and to collect their self-reported vaccination data (vaccination status, vaccine brands received, the total number of doses, and dates of vaccination). The median (IQR) time to contact the patients was 202 (108-352) days after the ED index visit. As patients were called at most 30 days after their data were consulted from the Québec Vaccination Registry, we only used self-reported vaccination data from before the date we consulted the Québec Vaccination Registry.

In accordance with CCEDRRN's national data management plan, after assigning a unique study identifier for each participant, deidentified data were collected and stored in REDCap (Research Electronic Data Capture; Vanderbilt University) at the University of British Columbia. After data verification, registry data were uploaded into CaraSpace, a secure private cloud for the storage and analysis of privacy-sensitive data [28]. Analysts accessed the cloud space via an encrypted virtual private network through a firewall and 2-factor authentication.

The primary outcome of this study was the accuracy of self-reported COVID-19 vaccination status at the time of the telephone follow-up compared with their vaccination status in the Québec Vaccination Registry. We considered the self-reporting of vaccination status via telephone as the index test and the Québec Vaccination Registry as the reference standard. The index test had 2 possible results: "no vaccine dose" (self-reported as not vaccinated) or "at least one vaccine dose" (self-reported as vaccinated). We calculated the accuracy of the self-reported vaccination status as follows:

> Accuracy = (correctly self-reported as vaccinated + correctly self-reported as not vaccinated) / (correctly self-reported as vaccinated + correctly self-reported as not vaccinated + incorrectly self-reported as vaccinated + incorrectly self-reported as not vaccinated) **(1)**

We also compared the clinical and demographic characteristics of patients with concordant and discordant self-reported vaccination status compared with the Québec Vaccination Registry.

Secondary outcomes were the sensitivity, calculated as

> (correctly self-reported as vaccinated / [correctly self-reported as vaccinated + incorrectly self-reported as not vaccinated]) **(2)**

and specificity, calculated as

> (correctly self-reported as not vaccinated / [correctly self-reported as not vaccinated + incorrectly self-reported as vaccinated]) **(3)**

of the self-reported vaccination status and vaccination status in the Québec Vaccination Registry. Relaxing the reference standard assumption about the Québec Vaccination Registry, we also measured the interrater agreement between self-reported vaccination status and vaccination status in the Québec Vaccination Registry, interrater agreement between self-reported COVID-19 vaccine brand received and the brand recorded in the Québec Vaccination Registry, and interrater agreement between self-reported number of COVID-19 vaccine doses received compared with the number of doses recorded in the vaccine registry.

An additional secondary outcome was the interrater agreement between vaccination status at the time of the index ED visit determined retrospectively by using the self-reported vaccination date compared with the vaccination status at the time of the index ED visit determined by the vaccination date in the Québec Vaccination Registry (Multimedia Appendix 1). Vaccination status was defined as fully, partially, and not vaccinated at the time of the index ED visit based on the number of doses and time since the last dose received to account for likely waning protection [33,34] (Multimedia Appendix 1). We defined fully vaccinated patients at the time of the index ED visit as patients who received ≥2 doses ≥14 days before the index ED visit but less than 6 months since the last dose. We defined partially vaccinated patients as having received only 1 dose before the index ED visit or having received their second or third dose <14 days before the index ED visit or having received their second or third dose >6 months before the index ED visit. Patients were considered not vaccinated if no doses had been received before their index ED visit. Patients who could not recall the date of their last vaccination were excluded from this analysis.

Analysis

The vaccination data in the Québec Vaccination Registry were considered the reference standard. As the interest was to compare the same information collected by telephone as in the Québec Vaccination Registry, the sample size calculations were based on a 1-sided binomial test for sensitivity [35,36]. A sample size of 613 patients was needed to achieve a 95% CI for sensitivity of half width 0.02 assuming a sensitivity of 0.98 and prevalence of at least 0.40.

Data were summarized by using descriptive statistics (eg, mean [SD], counts, and percentage). Accuracy, sensitivity, and specificity for COVID-19 vaccination status reported by patients compared with the reference standard (ie, the Québec Vaccination Registry) were calculated with associated 95% CIs. Secondary analyses relaxed the reference standard assumption and provided interrater agreement using the unweighted Cohen κ coefficient [37]. We chose to use the unweighted Cohen κ because we had 2 raters (ie, self-report by participant and report from the Québec Vaccination Registry), and we wanted disagreements to carry the same weight. All analyses were conducted in R (R Foundation for Statistical Computing) using the *rms* and *psych* packages [38-40]. To ensure patient privacy, a cell-size restriction policy prohibited us from reporting counts of <5.

Results

Participant Characteristics

During the study period, there were 5912 adult patients who presented to the Québec study sites, and 1482 met the inclusion criteria; of them, 1361 (23.02%) patients had the required data from follow-up and available in the Québec Vaccination Registry (Figure 1). Among the 4430 excluded patients, 1596 (36.03%) did not consent to participate or were not able to consent. There were also 62 (1.05%) patients who did not have any vaccination record in the Québec Vaccination Registry because they were patients from outside the province or did not have a provincial health insurance number. The baseline participant characteristics are shown in Table 1. Patients who did not or could not consent were slightly older than patients who participated (age: mean 58.6, SD 21.3 years vs mean 55.1, SD 17.4 years, respectively) and were more likely to be male individuals (nonconsenting male individuals: 871/1596, 54.57%, vs consenting male individuals: 686/1361, 50.4%). Three-quarters (1010/1361, 74.21%) of the patients presented to the ED during the second and third waves of the pandemic (August 23, 2020, to July 17, 2021).

Figure 1. Flow diagram of patients in the study.

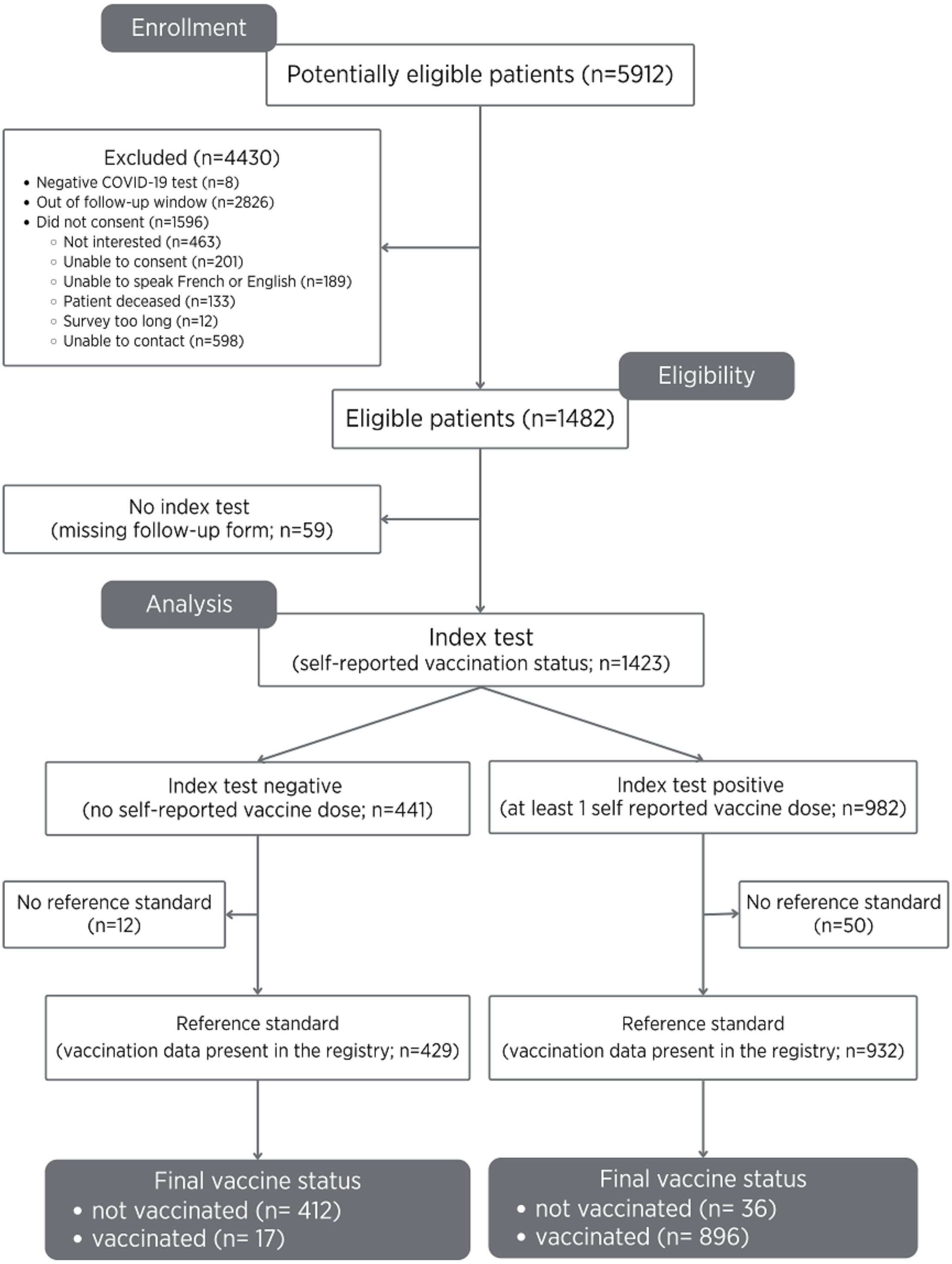


Table 1. Sociodemographic and clinical characteristics of study participants whose index visit occurred in a Québec emergency department between March 24, 2020, and December 25, 2021 (n=1361).

	Study participants
Age (years), mean (SD)	55.1 (17.4)
Age group (years), n (%)	
<24	45 (3.3)
25-39	230 (16.9)
40-64	685 (50.3)
65-79	272 (20)
≥80	129 (9.5)
Sex, n (%)	
Male	686 (50.4)
Female	675 (49.6)
Intersex	0 (0)
Site, n (%)	
Hôtel-Dieu de Lévis	293 (21.5)
Royal Victoria	501 (36.8)
Montreal General	132 (9.7)
Sacré-Coeur de Montréal	435 (32)
Index visit occurred during, n (%)	
Wave 1 (March 1, 2020, to August 22, 2020)	207 (15.2)
Wave 2 (August 23, 2020, to March 20, 2021)	725 (53.3)
Wave 3 (March 21, 2021, to July 17, 2021)	285 (20.9)
Wave 4 to early 5 (July 18, 2021, to December 31, 2021)	144 (10.6)
Index visit occurred before the start of vaccination campaign in Québec (December 14, 2020), n (%)	
Prevaccination campaign	490 (36)
Postvaccination campaign	871 (64)
Time elapsed from index visit to telephone follow-up (days), median (IQR)	202 (108-352)
Top 10 comorbid conditions, n (%)	
Hypertension	440 (32.3)
Dyslipidemia	314 (23.1)
Diabetes	241 (17.7)
Hypothyroidism	160 (11.8)
Asthma	144 (10.6)
Psychiatric condition or mental health diagnosis	113 (8.3)
Coronary artery disease	99 (7.3)
Past malignancy	73 (5.4)
Chronic neurological conditions	66 (4.8)
Rheumatologic disorder	66 (4.8)
ED[a] disposition, n (%)	
Discharged home	740 (54.4)
Admitted	576 (42.3)
Transferred to other hospital	38 (2.8)
Transfer to long-term care or rehab	<5[b] (<0.4)

	Study participants
Other	6 (0.4)

[a] ED: emergency department.

[b] Cells with fewer than 5 individuals could not be presented as requested by the ethics committee.

Vaccination Status

Over two-thirds (932/1361, 68.55%) of the participants reported at least 1 dose of a COVID-19 vaccine at the time of their follow-up interview (Table 2). The most administered vaccine type was Pfizer-BioNTech (689/1361, 50.6%) followed by Moderna (158/1361, 11.6%).

Table 2. Québec Vaccination Registry information and self-reported vaccination status during the initial telephone follow-up of participants whose emergency department index visit occurred between March 24, 2020, and December 25, 2021 (n=1361).

	Québec Vaccination Registry, n (%)	Self-reported vaccination status, n (%)
Vaccination status		
Yes (at least 1 dose)	913 (67.1)	932 (68.5)
No (no dose given)	448 (32.9)	429 (31.5)
Unknown or not assessed	0 (0)	0 (0)
First vaccine brand		
Pfizer-BioNTech	707 (51.9)	689 (50.6)
Moderna	162 (11.9)	158 (11.6)
AstraZeneca or Covishield	42 (3.1)	42 (3.1)
Johnson & Johnson	<5 (<0.4)	<5[a] (<0.4)
Unknown, but vaccinated	0 (0)	37 (2.7)
Unknown or not vaccinated	448 (32.9)	433 (31.8)
Number of vaccines doses		
0	448 (32.9)	429 (31.5)
1	366 (26.9)	368 (27)
2	501 (36.8)	511 (37.5)
3	46 (3.4)	53 (3.9)

[a] Cells with fewer than 5 individuals could not be presented as requested by the ethics committee.

Self-Reported Vaccination Status Accuracy and Interrater Agreement Between Self-Reported Vaccination Data and Québec Vaccination Registry Data

We reported high measures of accuracy (96%), sensitivity (98%), and specificity (92%) for self-reported vaccination status compared with the Québec Vaccination Registry as the reference standard (Tables 3 and 4). Moreover, Cohen κ was 0.91 (95% CI 0.89-0.93), indicating excellent agreement between patient self-report and the Québec Vaccination Registry.

With respect to vaccine brand type, Cohen κ for dose 1 brand was 0.80 (95% CI 0.75-0.84) indicating strong agreement (Table 5). However, this estimate decreased with subsequent doses: Cohen κ was 0.76 (95% CI 0.70-0.83) and 0.59 (95% CI 0.34-0.83) for dose 2 and dose 3 brands, respectively.

Most patients reported receiving 1 dose (368/1361, 27.04%) or 2 doses (511/1361, 37.56%) of COVID-19 vaccines at the time of the follow-up interview (Table 6). Only 3.89% (53/1361) of patients reported receiving 3 doses. One-third (429/1361, 31.52%) of the patients reported not having received the COVID-19 vaccine at the time of follow-up. Cohen κ for the number of self-reported doses compared with the Québec Vaccination Registry was very strong, at 0.89 (95% CI 0.87-0.91).

Cohen κ for self-reported vaccination status at the time of ED index visit was 0.85 (95% CI 0.77-0.92). However, many patients (774/1361, 56.87%) could not recall the exact date of their latest vaccine dose and had to be excluded from this analysis. Of those who did recall their vaccination dates, most (538/587, 91.7%) reported not being fully vaccinated at the time of the ED visit, with only 16 (2.7%) patients being fully vaccinated (Table 7).

The demographic, clinical, and sociocultural characteristics of patients with discordant (n=53) and concordant (n=1308) vaccination statuses are shown in Multimedia Appendix 2. Participants with discordant self-reported vaccination status compared to those with concordant self-reported vaccination status were younger (age: mean 51.1, SD 20.5 years, vs mean 55.3, SD 17.3 years, respectively), had lower self-reported income (5/53, 9% reported current income >CAD $73,701 (US

$56,693) per year vs 294/1308, 22.48%), had lower education level (5/53, 9% with a university degree vs 444/1308, 33.94%), and had a lower vaccination rate according to the Québec Vaccination Registry (17/53, 32% vs 896/1308, 68.5%). The median time elapsed between the ED index visit and the telephone follow-up to determine the patient's self-reported vaccination status was similar for discordant (median 94, IQR 53-191 days) and concordant cases (median 126, 62-196 days). Immigration status, race, and ethnicity did not influence the accuracy of our results.

Table 3. Confusion matrix comparing self-reported vaccination status with the Québec Vaccination Registry data for all participants whose emergency department index visit occurred between March 24, 2020, and December 25, 2021 (n=1361).

Self-reported vaccination status (index test)	Vaccination status in the Québec Vaccination Registry (reference standard), n	
	Yes	No
Yes	896	36
No	17	412

Table 4. Sensitivity, specificity and accuracy of self-reported vaccination status and interrater agreement of self-reported vaccination status compared with the Québec Vaccination Registry for all participants whose emergency department index visit occurred between March 24, 2020, and December 25, 2021 (n=1361).

	Measure (95% CI)
Sensitivity	0.98 (0.97-0.99)
Specificity	0.92 (0.89-0.94)
Accuracy	0.96 (0.95-0.97)
Cohen κ	0.91 (0.89-0.93)

Table 5. Interrater agreement for self-reported brand by dose compared with the Québec Vaccination Registry for all participants whose emergency department index visit occurred between March 24, 2020, and December 25, 2021.

Dose and self-reported vaccine brand	Québec Vaccination Registry vaccine brand, n				Cohen κ[a] (95% CI)
	Pfizer	Moderna	Astra Zeneca or Covishield	Johnson & Johnson	
1					0.80 (0.75-0.84)
Pfizer	645	19	<5	0	
Moderna	11	137	0	0	
AstraZeneca or Covishield	<5	0	38	0	
Johnson & Johnson	0	0	0	<5	
Unknown, but vaccinated	29	6	<5	0	
2					0.76 (0.70-0.83)
Pfizer	386	19	<5	0	
Moderna	5	84	0	0	
AstraZeneca or Covishield	0	0	9	0	
Unknown, but vaccinated	16	5	0	0	
3					0.59 (0.34-0.83)
Pfizer	24	<5	0	0	
Moderna	<5	7	0	0	
Unknown, but vaccinated	<5	<5	0	0	

[a] Cohen κ calculated without participants who answered "Unknown, but vaccinated" were 0.89 (95% CI 0.86-0.93) for first dose, 0.85 (95% CI 0.80-0.91) for second dose, and 0.70 (95% CI 0.43-0.97) for third dose.

Table 6. Interrater agreement for self-reported number of doses compared with the Québec Vaccination Registry for participants whose emergency department index visit occurred between March 24, 2020, and December 25, 2021 (n=1361).[a]

Self-reported number of doses	Number of doses in the Québec Vaccination Registry, n			
	0	1	2	3
0	412	7	8	<5
1	20	336	12	0
2	13	23	469	6
3	<5	0	12	38

[a] Cohen κ: 0.89 (95% CI 0.87-0.91).

Table 7. Interrater agreement for self-reported vaccination status at emergency department index visit date versus the Québec Vaccination Registry for participants whose emergency department index visit occurred between March 24, 2020, and December 25, 2021 (n=587).[a]

Self-reported vaccination status determined using the date of latest vaccination dose in relation to the ED[b] visit date[c]	Québec Vaccination Registry vaccination status determined in relation to the ED index visit date[d], n		
	Not vaccinated	Partially vaccinated	Fully vaccinated
Not vaccinated	527	5	6
Partially vaccinated	<5	29	0
Fully vaccinated	0	0	16

[a] Cohen κ: 0.85 (95% CI 0.77-0.92).

[b] ED: emergency department.

[c] Categories of vaccination status at the time of ED index visit (fully, partially, and not vaccinated) for this analysis are defined in Multimedia Appendix 1.

[d] A total of 774 (56.9%) participants were excluded from this analysis because they could not recall the date of their last dose.

Discussion

Principal Findings

We found that the participants' self-reported vaccination status was highly accurate at the time of the initial follow-up telephone call. The interrater agreement between the self-reported vaccine number of doses and the number of doses recorded in the vaccine registry was also strong. The agreement between self-reported vaccination status at the index visit and the patient's vaccination status as determined by the registry was strong when patients could provide vaccination dates. However, many patients could not recall their vaccination dates. Although there was a strong and moderate agreement for the self-reported vaccine brand for the first and second doses, respectively, the agreement dropped to weak for the vaccine brand of the third dose. A few discordant cases concerning the accuracy of patients' self-report of receiving at least 1 dose were found. The discordant cases involved patients of younger age groups, lower self-reported pre– and post–COVID-19 income, lower education levels, and lower vaccination rates in the Québec Vaccination Registry.

Our results are similar to those of a study conducted in the United States by Stephenson et al [41] that showed high levels of agreement between self-reported vaccination status and official computerized vaccination registry data (unweighted κ=0.9127, 95% CI 0.9109-0.9145). Similar to our study, they also found that discordant self-reported vaccination data were associated with lower education levels and unemployment. It could be that both studies found that lower education is associated with lower accuracy of self-reported vaccination status because information provided about vaccination status tends to be complex [42]. Comprehending and remembering this information requires certain health literacy skills [43]. Rolnick et al [44] also found that education and income level could influence the accuracy of self-reported vaccination status for 8 different vaccines (pneumococcal, influenza, tetanus diphtheria, tetanus diphtheria pertussis, human papillomavirus, hepatitis A, hepatitis B, and herpes zoster).

Contrary to our study, the study by Stephenson et al [41] found that the female sex increased the risk of discordant self-reported vaccination status. Sex and gender are important factors in understanding the real-world effectiveness of vaccines and postmarketing safety monitoring [45,46]. Sex and gender differences in vaccine hesitancy and immunization outcomes have been observed for COVID-19 and other vaccine-preventable diseases, with women typically reporting more vaccine hesitancy and higher adverse reactions than men [47-49]. Our results support the accuracy of self-reported vaccination data from both men and women.

Ulrich et al [50] studied the accuracy of the self-reported vaccination status by university students in a 2-dose meningococcal vaccine campaign. This study showed good self-reported vaccination status accuracy at 2 months (86%) and poor self-reported vaccination status at 20 months (69%). King et al [51] also reported a decrease in the agreement of patient self-reported vaccination status compared with a registry for prior season influenza vaccination (κ=0.86) versus the current season (κ=0.95). Although our study did not observe a time-dependent increase in discordant self-reported vaccination status, we did find that patients reporting a higher number of doses were more likely to forget the vaccine brand received and did not remember the dates of their latest vaccination dose. In patients who did remember the vaccine brand for each dose, the agreement with the registry was high. This limits the use of self-reported vaccination brand names for studies comparing the effectiveness of multiple doses of different types of vaccines [52].

The reliability of data on the timing of vaccination before being tested for a viral infection such as SARS-CoV-2 is essential for test-negative design vaccine effectiveness studies [53]. Our study was interested in determining the reliability of the self-reported vaccination data (ie, the date of their last vaccination and the total number of doses) used to retrospectively determine their vaccination status (not vaccinated, partially vaccinated, or fully vaccinated) at the time of their index ED visit. For patients who could remember their last vaccination date, our study found high levels of agreement between the self-reported vaccination status determined at the ED index visit date and the Québec Vaccination Registry. Unfortunately, more than half of our participants could not recall the date of their last dose, making this measure subject to an important selection bias when used in test-negative designs. This is an argument to support wider implementation and easy access to rigorous vaccination registries in all jurisdictions [21]. This is especially important to improve the care, research, and support of vulnerable populations presenting to the ED. In Canada, access to these registries is currently restricted by data custodians and administrative barriers that hinder timely knowledge creation and decision-making for a rapid-learning health system [54,55]. Although a study on the accuracy of self-reported influenza and pneumococcal vaccination status in a high-risk inflammatory bowel disease population showed high accuracy for self-reported influenza (97%) and pneumococcal (84%) vaccination status [56], our study did not find any difference between high-risk groups with different comorbidities. In concordance with Laurence et al [57], who found that older patients accurately self-reported their influenza (96%) and pneumococcal (87%) vaccine status, we also found that older adults adequately self-reported their COVID-19 vaccine status. Surprisingly, we found a higher proportion of discordant self-reported COVID-19 vaccination statuses among patients aged ≤34 years. This could be associated with the fact that COVID-19 presents a milder course in younger patients [57,58] and increased vaccination hesitancy in this age group [59-61].

There are several potential explanations for the high accuracy of self-reported vaccination status at telephone follow-up and moderate to strong agreement between the self-reported number of doses and the self-reported vaccination status determined at the time of the ED index visit. First, this high accuracy and strong agreement could have been caused by the implementation of Québec's mandatory vaccine passport, which registered dates, brands, and doses as a requirement to access the workplace and most commercial services. Although we did not ask participants to consult their passports before answering our phone questionnaire, our study personnel did not prevent participants from consulting their passports during our telephone follow-ups. Participants' recall of their own vaccination data could also have been improved by having to consult their passport frequently, as both federal and provincial governments in Canada granted substantial benefits to fully vaccinated citizens, such as lifting travel restrictions and access to event venues, fitness centers, and restaurants via the use of a vaccination passport [62]. Second, the impact of COVID-19 on the Canadian population has been substantial, and it may have improved the recall and the accuracy of the self-reported vaccination status. Third, the COVID-19 vaccination campaign; the omnipresence of the COVID-19 pandemic in the lay press; the many frequently reported statistics portraying the evolution of this pandemic (cases, hospitalizations, deaths, and population vaccination coverage); and the vigorous public debate about vaccination public health mandates may have also raised citizens' awareness of their own vaccination status. Fourth, this study only included patients with confirmed COVID-19 infections. This could have improved the recall of participants compared with patients who were not infected with COVID-19.

Our results indicate that when access to official electronic vaccination registries is not available, the reliance on patients' self-reported COVID-19 vaccination status, vaccine brand received, number of doses, and vaccination dates (when provided) can be used for research purposes. This is a substantial finding for researchers in jurisdictions that do not have access to electronic vaccine registries [20,63] and for others that have substantial challenges in securing timely access to individual-level vaccination data even when they exist [24]. Vaccine registries can also suffer from missing data for patients who have received their vaccines in different jurisdictions (as shown in our results), misclassification bias because of human error, and the timeliness of data entry into electronic vaccination registries [64]. Relying on timely and highly accurate self-reported vaccination status reduces the potential underestimation of real-world test-negative vaccine effectiveness studies introduced by vaccination-reporting bias [25]. Knowing the sensitivity and specificity of the self-reported vaccination status is also important to correct the estimation of vaccine effectiveness based on the observed risks for vaccination status misclassification [65].

This study has limitations. First, although we assumed that the Québec Vaccination Registry was the reference standard, we cannot rule out the possibility that some vaccination data were entered incorrectly. Second, our results do not apply to patients with cognitive disorders who cannot provide consent. This represents a selection bias that we could not avoid because of ethical constraints. Patients who consented to participate in our study were also more likely to be younger and female than those who did not consent or who could not consent. Understanding the impact of the pandemic on patients who cannot consent to provide their vaccination status because of cognitive issues will still necessitate access to linked vaccination registry data. Access to a vaccination registry also avoids systematic differences in the characteristics of consenting patients versus nonconsenting patients in studies that rely on self-reporting. Third, although we had many landed immigrants constituting an ethnically diverse sample, our results are potentially not applicable to patients who do not speak English or French. Other studies have shown that the self-reported vaccination status is inaccurate in certain susceptible and hard-to-follow-up racial and ethnic groups [44,50]. Although we had to exclude patients who could not speak English or French, immigration and race and ethnicity did not influence the accuracy of our results. Fourth, our study did not verify the accuracy of vaccination status provided at the time of ED care. Future prospective studies at the point of care would be needed to validate the accuracy of self-reported data for use in guiding clinical decision-making.

Conclusions

Our study found that relying on self-reported COVID-19 vaccination status is accurate for adult patients without cognitive disorders who can express themselves in English or French. Researchers can use self-reported COVID-19 vaccination data on the number of doses received, vaccine brand name, and timing of vaccination to guide future research with certain groups of patients who are capable of self-reporting their vaccination data. However, access to official electronic vaccine registries is still needed to determine the vaccination status in certain susceptible populations where self-reported vaccination data remain missing or impossible to obtain.

Acknowledgments

The authors would like to acknowledge the assistance of Ms Amber Cragg and Ms Josie Kanu in the preparation of this manuscript. The authors would like to thank the University of British Columbia clinical coordinating center staff; legal, ethics, privacy, and contract staff; and the research staff at each of the participating institutions in the network outlined in the attached supplement (Multimedia Appendix 3). In particular, the authors would like to thank Chantal Lanthier and David Iannuzzi for their support in collecting the data for this study. The network would not exist today without the dedication of these professionals. The authors would also like to thank all the patient partners who shared their lived experiences and perspectives to ensure that the knowledge we cocreate addresses the concerns of patients and the public. Creating the largest network of collaboration across Canadian emergency departments would not have been possible without the tireless efforts of emergency department chiefs, research coordinators, and research assistants at participating sites. Finally, the authors would like to express their most humble and sincere gratitude to all their colleagues in medicine, nursing, and the allied health professions who have been on the frontlines of this pandemic from day 1, staffing ambulances, emergency departments, intensive care units, and hospitals bravely facing the risks of COVID-19 to look after fellow citizens and after one another. The authors would like to dedicate this network to them. This network was funded by the Canadian Institutes of Health Research (447679, 464947, and 466880); Ontario Ministry of Colleges and Universities (C-655-2129); Saskatchewan Health Research Foundation (5357); Genome BC (COV024 and VAC007); Fondation du CHU de Québec (Octroi No. 4007); and Sero-Surveillance and Research (COVID-19 Immunity Task Force Initiative), who provided peer-reviewed funding. The BC Academic Health Science Network and BioTalent Canada provided non–peer-reviewed funding. These organizations are not for profit and had no role in conducting the study, analysis, or manuscript preparation.

The Canadian COVID-19 Emergency Department Rapid Response Network (CCEDRRN) consisted of the following members: Patrick Fok, Hana Wiemer, Samuel Campbell, Kory Arsenault, Tara Dahn, Kavish Chandra, Patrick Archambault, Joel Turner, Lars Grant, Greg Clark, Raoul Daoust, Éric Mercier, Sébastien Robert, Laurie Morrison, Steven Brooks, Ivy Cheng, Jeffrey Perry, Rob Ohle, Justin Yan, Rohit Mohindra, Megan Landes, Michelle Welsford, Tomislav Jelic, Phil Davis, Andrew McRae, Brian Rowe, Katie Lin, Stephanie VandenBerg, Jake Hayward, Jaspreet Khangura, Corinne Hohl, Daniel Ting, Maja Stachura, Frank Scheuermeyer, Balijeet Braar, John Taylor, Ian Martin, Sean Wormsbecker and Lee Graham.

Data Availability

Data are available on reasonable request. For investigators who wish to access Canadian COVID-19 Emergency Rapid Response Network data, proposals may be submitted to the network for review and approval by the network's peer review publication committee, the data access and management committee, and the executive committee as per the network's governance. Information regarding submitting proposals and accessing data may be found on the Canadian COVID-19 Emergency Department Rapid Response Network website [9].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Definitions used to determine the interrater agreement between the self-reported vaccination status and the vaccination status determined by the Québec Vaccination Registry at the time of emergency department index visit.

[PNG File , 288 KB - publichealth_v9i1e44465_app1.png]

Multimedia Appendix 2

Characteristics of all participants by concordant or discordant vaccination status.
[DOCX File , 32 KB - publichealth_v9i1e44465_app2.docx]

Multimedia Appendix 3
Contributors to the Canadian COVID-19 Emergency Department Rapid Response Network.
[DOCX File , 50 KB - publichealth_v9i1e44465_app3.docx]

References

1. Coccia M. Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: a global analysis. Environ Res 2022 Mar;204(Pt C):112314 [FREE Full text] [doi: 10.1016/j.envres.2021.112314] [Medline: 34736923]
2. COVID-19 map. Johns Hopkins Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2023-04-06]
3. Coccia M. COVID-19 vaccination is not a sufficient public policy to face crisis management of next pandemic threats. Public Organ Rev (forthcoming) 2022 Oct 17 [FREE Full text] [doi: 10.1007/s11115-022-00661-6]
4. Patel MK, Bergeri I, Bresee JS, Cowling BJ, Crowcroft NS, Fahmy K, et al. Evaluation of post-introduction COVID-19 vaccine effectiveness: summary of interim guidance of the World Health Organization. Vaccine 2021 Jul 05;39(30):4013-4024 [FREE Full text] [doi: 10.1016/j.vaccine.2021.05.099] [Medline: 34119350]
5. Patel MM, Jackson ML, Ferdinands J. Postlicensure evaluation of COVID-19 vaccines. JAMA 2020 Nov 17;324(19):1939-1940. [doi: 10.1001/jama.2020.19328] [Medline: 33064144]
6. Stacey D, Ludwig C, Archambault P, Smith M, Taljaard M, Carley M, et al. Decisions and decisional needs of Canadians from all provinces and territories during the COVID-19 pandemic: population-based cross-sectional surveys. JMIR Public Health Surveill 2023 Mar 21;9:e43652 [FREE Full text] [doi: 10.2196/43652] [Medline: 36688986]
7. Shea LL, Becker A, Lee BK, Miller KK, Cooper D, Anderson K, et al. Self-reported COVID-19 vaccination acceptance and hesitancy among autistic adults. Vaccine 2022 May 26;40(24):3288-3293 [FREE Full text] [doi: 10.1016/j.vaccine.2022.04.060] [Medline: 35484038]
8. Sasanami M, Fujimoto M, Kayano T, Hayashi K, Nishiura H. Projecting the COVID-19 immune landscape in Japan in the presence of waning immunity and booster vaccination. J Theor Biol 2023 Feb 21;559:111384 [FREE Full text] [doi: 10.1016/j.jtbi.2022.111384] [Medline: 36528092]
9. Wesselink AK, Hatch EE, Rothman KJ, Wang TR, Willis MD, Yland J, et al. A prospective cohort study of COVID-19 vaccination, SARS-CoV-2 infection, and fertility. Am J Epidemiol 2022 Jul 23;191(8):1383-1395 [FREE Full text] [doi: 10.1093/aje/kwac011] [Medline: 35051292]
10. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger J. Vaccine hesitancy: an overview. Hum Vaccin Immunother 2013 Aug;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
11. Yoon H, Choi BY, Seong WJ, Cho GJ, Na S, Jung YM, et al. COVID-19 vaccine acceptance during pregnancy and influencing factors in South Korea. J Clin Med 2022 Sep 28;11(19):5733 [FREE Full text] [doi: 10.3390/jcm11195733] [Medline: 36233601]
12. Coccia M. Improving preparedness for next pandemics: max level of COVID-19 vaccinations without social impositions to design effective health policy and avoid flawed democracies. Environ Res 2022 Oct;213:113566 [FREE Full text] [doi: 10.1016/j.envres.2022.113566] [Medline: 35660409]
13. Etienne D, Archambault P, Aziaka D, Chipenda-Dansokho S, Dubé E, Fallon CS, et al. A personalized avatar-based web application to help people understand how social distancing can reduce the spread of COVID-19: cross-sectional, observational, pre-post study. JMIR Form Res 2023 Apr 25;7:e38430 [FREE Full text] [doi: 10.2196/38430] [Medline: 36961787]
14. Hakim H, Bettinger JA, Chambers CT, Driedger SM, Dubé E, Gavaruzzi T, et al. A web application about herd immunity using personalized avatars: development study. J Med Internet Res 2020 Oct 30;22(10):e20113 [FREE Full text] [doi: 10.2196/20113] [Medline: 33124994]
15. Akinbami LJ, Kruszon-Moran D, Wang CY, Storandt RJ, Clark J, Riddles MK, et al. SARS-CoV-2 serology and self-reported infection among adults - national health and nutrition examination survey, United States, August 2021-May 2022. MMWR Morb Mortal Wkly Rep 2022 Dec 02;71(48):1522-1525 [FREE Full text] [doi: 10.15585/mmwr.mm7148a4] [Medline: 36454698]
16. Miao Y, Zhang W, Li Y, Wu J, Xu D, Gu J, et al. Association between lifestyle and COVID-19 vaccination: a national cross-sectional study. Front Public Health 2022 Oct 11;10:918743 [FREE Full text] [doi: 10.3389/fpubh.2022.918743] [Medline: 36304239]
17. Kronfli N, Dussault C, Maheu-Giroux M, Halavrezos A, Chalifoux S, Park H, et al. Importance of occupation for SARS-CoV-2 seroprevalence and COVID-19 vaccination among correctional workers in Quebec, Canada: a cross-sectional study. Front Public Health 2022 Nov 09;10:1021871 [FREE Full text] [doi: 10.3389/fpubh.2022.1021871] [Medline: 36438247]

18. Gormley MA, Nolan MS, Heo M, Litwin AH, Alier A, Daguise V. Differences in motivating factors for SARS-CoV-2 vaccination and perceptions of infection risk among healthcare and EMS personnel in South Carolina. South Med J 2022 Jun;115(6):381-387 [FREE Full text] [doi: 10.14423/SMJ.0000000000001406] [Medline: 35649524]
19. Coccia M. Sources, diffusion and prediction in COVID-19 pandemic: lessons learned to face next health emergency. AIMS Public Health 2023 Mar 02;10(1):145-168 [FREE Full text] [doi: 10.3934/publichealth.2023012] [Medline: 37063362]
20. Namageyo-Funa A, Samuel A, Bloland P, Macneil A. Considerations for the development and implementation of electronic immunization registries in Africa. Pan Afr Med J 2018 May 29;30:81 [FREE Full text] [doi: 10.11604/pamj.2018.30.81.11951] [Medline: 30344865]
21. Benati I, Coccia M. Global analysis of timely COVID-19 vaccinations: improving governance to reinforce response policies for pandemic crises. Int J Health Gov 2022 May 31;27(3):240-253 [FREE Full text] [doi: 10.1108/ijhg-07-2021-0072]
22. Coccia M. Pandemic prevention: lessons from COVID-19. Encycl 2021 May 31;1(2):433-444 [FREE Full text] [doi: 10.3390/encyclopedia1020036]
23. González MJ, Rodríguez FS. Self-reported COVID-19 vaccination coverage in Venezuela: results of national sample surveys (June-December 2021). Gac Med Caracas 2022 Jul 15;130(2S):S459-S469 [FREE Full text] [doi: 10.47307/gmc.2022.130.s2.17]
24. McRae AD, Archambault P, Fok P, Wiemer H, Morrison LJ, Herder M, Canadian COVID-19 Emergency Department Rapid Response Network Investigators. Reducing barriers to accessing administrative data on SARS-CoV-2 vaccination for research. CMAJ 2022 Jul 18;194(27):E943-E947 [FREE Full text] [doi: 10.1503/cmaj.211712] [Medline: 35851526]
25. Jackson ML. Use of self-reported vaccination status can bias vaccine effectiveness estimates from test-negative studies. Vaccine X 2018 Dec 29;1:100003 [FREE Full text] [doi: 10.1016/j.jvacx.2018.100003] [Medline: 33826685]
26. Feikin DR, Higdon MM, Abu-Raddad LJ, Andrews N, Araos R, Goldberg Y, et al. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. Lancet 2022 Mar 05;399(10328):924-944 [FREE Full text] [doi: 10.1016/S0140-6736(22)00152-0] [Medline: 35202601]
27. Québec vaccination registry. Gouvernement du Québec. URL: https://www.quebec.ca/en/health/your-health-information/quebec-vaccination-registry [accessed 2023-04-06]
28. Hohl CM, Rosychuk RJ, McRae AD, Brooks SC, Archambault P, Fok PT, Canadian COVID-19 Emergency Department Rapid Response Network investigators and for the Network of Canadian Emergency Researchers and the Canadian Critical Care Trials Group. Development of the Canadian COVID-19 emergency department rapid response network population-based registry: a methodology study. CMAJ Open 2021 Mar 17;9(1):E261-E270 [FREE Full text] [doi: 10.9778/cmajo.20200290] [Medline: 33731427]
29. Canadian COVID-19 Emergency Department Rapid Response Network. URL: https://www.ccedrrn.com/ [accessed 2023-05-28]
30. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015 Oct 28;351:h5527 [FREE Full text] [doi: 10.1136/bmj.h5527] [Medline: 26511519]
31. Ligne du temps COVID-19 au Québec. Institut national de santé publique du Québec. URL: https://www.inspq.qc.ca/covid-19/donnees/ligne-du-temps [accessed 2023-04-06]
32. Inscription obligatoire en tant que personne vaccinée. Gouvernement du Québec. 2018 Apr 20. URL: https://www.quebec.ca/sante/vos-informations-de-sante/registre-de-vaccination-du-quebec/inscription-obligatoire-en-tant-que-personne-vaccinee [accessed 2023-04-06]
33. Britton A, Fleming-Dutra KE, Shang N, Smith ZR, Dorji T, Derado G, et al. Association of COVID-19 vaccination with symptomatic SARS-CoV-2 infection by time since vaccination and delta variant predominance. JAMA 2022 Mar 15;327(11):1032-1041 [FREE Full text] [doi: 10.1001/jama.2022.2068] [Medline: 35157002]
34. Tartof SY, Slezak JM, Fischer H, Hong V, Ackerson BK, Ranasinghe ON, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet 2021 Oct 16;398(10309):1407-1416 [FREE Full text] [doi: 10.1016/S0140-6736(21)02183-8] [Medline: 34619098]
35. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform 2014 Apr;48:193-204 [FREE Full text] [doi: 10.1016/j.jbi.2014.02.013] [Medline: 24582925]
36. Sample size software. NCSS: Statistical Software. 2012. URL: https://www.ncss.com/software/pass/ [accessed 2023-05-28]
37. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
38. R Core Team. R: a language and environment for statistical computing. R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2023-05-28]
39. Harrell Jr FE. Regression modeling strategies. Comprehensive R Archive Network (CRAN). 2023 May 08. URL: https://CRAN.R-project.org/package=rms [accessed 2023-05-28]
40. Revelle W. Procedures for psychological, psychometric, and personality research. Comprehensive R Archive Network (CRAN). 2023 Mar 18. URL: https://CRAN.R-project.org/package=psych [accessed 2023-05-28]

41. Stephenson M, Olson SM, Self WH, Ginde AA, Mohr NM, Gaglani M, et al. Ascertainment of vaccination status by self-report versus source documentation: impact on measuring COVID-19 vaccine effectiveness. Influenza Other Respir Viruses 2022 Nov;16(6):1101-1111 [FREE Full text] [doi: 10.1111/irv.13023] [Medline: 35818721]
42. Lorini C, Santomauro F, Donzellini M, Capecchi L, Bechini A, Boccalini S, et al. Health literacy and vaccination: a systematic review. Hum Vaccin Immunother 2018 Feb 01;14(2):478-488 [FREE Full text] [doi: 10.1080/21645515.2017.1392423] [Medline: 29048987]
43. Rowlands G. Health literacy. Hum Vaccin Immunother 2014;10(7):2130-2135 [FREE Full text] [doi: 10.4161/hv.29603] [Medline: 25424830]
44. Rolnick SJ, Parker ED, Nordin JD, Hedblom BD, Wei F, Kerby T, et al. Self-report compared to electronic medical record across eight adult vaccines: do results vary by demographic factors? Vaccine 2013 Aug 20;31(37):3928-3935 [FREE Full text] [doi: 10.1016/j.vaccine.2013.06.041] [Medline: 23806243]
45. Vassallo A, Shajahan S, Harris K, Hallam L, Hockham C, Womersley K, et al. Sex and gender in COVID-19 vaccine research: substantial evidence gaps remain. Front Glob Womens Health 2021 Nov 01;2:761511 [FREE Full text] [doi: 10.3389/fgwh.2021.761511] [Medline: 34816252]
46. Heidari S, Goodman T. Critical sex and gender considerations for equitable research, development and delivery of COVID-19 vaccines. World Health Organization. 2021 Apr 18. URL: https://cdn.who.int/media/docs/default-source/immunization/sage/covid/gender-covid-19-vaccines-sage-background-paper.pdf?sfvrsn=899e8fca_15&download=true [accessed 2023-04-06]
47. Vijayasingham L, Heidari S, Munro J, Omer S, MacDonald N. Resolving sex and gender bias in COVID-19 vaccines R&D and beyond. Hum Vaccin Immunother 2022 Dec 31;18(1):2035142 [FREE Full text] [doi: 10.1080/21645515.2022.2035142] [Medline: 35143380]
48. Fischinger S, Boudreau CM, Butler AL, Streeck H, Alter G. Sex differences in vaccine-induced humoral immunity. Semin Immunopathol 2019 Mar;41(2):239-249 [FREE Full text] [doi: 10.1007/s00281-018-0726-5] [Medline: 30547182]
49. Zintel S, Flock C, Arbogast AL, Forster A, von Wagner C, Sieverding M. Gender differences in the intention to get vaccinated against COVID-19: a systematic review and meta-analysis. Z Gesundh Wiss 2022 Jan 07:1-25 [FREE Full text] [doi: 10.1007/s10389-021-01677-w] [Medline: 35018277]
50. Ulrich AK, McKearnan SB, Lammert S, Wolfson J, Pletcher J, Halloran ME, et al. Validity of university students' self-reported vaccination status after a meningococcal B outbreak. J Am Coll Health 2022 Apr;70(3):824-829 [FREE Full text] [doi: 10.1080/07448481.2020.1772270] [Medline: 32672510]
51. King JP, McLean HQ, Belongia EA. Validation of self-reported influenza vaccination in the current and prior season. Influenza Other Respir Viruses 2018 Nov;12(6):808-813 [FREE Full text] [doi: 10.1111/irv.12593] [Medline: 30028081]
52. Zheng C, Shao W, Chen X, Zhang B, Wang G, Zhang W. Real-world effectiveness of COVID-19 vaccines: a literature review and meta-analysis. Int J Infect Dis 2022 Jan;114:252-260 [FREE Full text] [doi: 10.1016/j.ijid.2021.11.009] [Medline: 34800687]
53. Skowronski DM, Febriani Y, Ouakki M, Setayeshgar S, El Adam S, Zou M, et al. Two-dose severe acute respiratory syndrome coronavirus 2 vaccine effectiveness with mixed schedules and extended dosing intervals: test-negative design studies from British Columbia and Quebec, Canada. Clin Infect Dis 2022 Nov 30;75(11):1980-1992 [FREE Full text] [doi: 10.1093/cid/ciac290] [Medline: 35438175]
54. Lamontagne F, Rowan KM, Guyatt G. Integrating research into clinical practice: challenges and solutions for Canada. CMAJ 2021 Jan 25;193(4):E127-E131 [FREE Full text] [doi: 10.1503/cmaj.202397] [Medline: 33298404]
55. Etheredge LM. A rapid-learning health system. Health Aff (Millwood) 2007 Mar;26(2):w107-w118. [doi: 10.1377/hlthaff.26.2.w107] [Medline: 17259191]
56. Smith R, Hubers J, Farraye FA, Sampene E, Hayney MS, Caldera F. Accuracy of self-reported vaccination status in a cohort of patients with inflammatory bowel disease. Dig Dis Sci 2021 Sep;66(9):2935-2941 [FREE Full text] [doi: 10.1007/s10620-020-06631-6] [Medline: 32995995]
57. Laurence A, Lewis P, Gately C, Dixon A. Influenza and pneumococcal vaccination: do older people know if they have been vaccinated? Aust N Z J Public Health 2016 Jun;40(3):279-280 [FREE Full text] [doi: 10.1111/1753-6405.12423] [Medline: 26261068]
58. Lu X, Zhang L, Du H, Zhang J, Li YY, Qu J, Chinese Pediatric Novel Coronavirus Study Team. SARS-CoV-2 Infection in Children. N Engl J Med 2020 Apr 23;382(17):1663-1665 [FREE Full text] [doi: 10.1056/NEJMc2005073] [Medline: 32187458]
59. Khan MS, Nguyen TX, Lal S, Watanapongvanich S, Kadoya Y. Hesitancy towards the third dose of COVID-19 vaccine among the younger generation in Japan. Int J Environ Res Public Health 2022 Jun 08;19(12):7041 [FREE Full text] [doi: 10.3390/ijerph19127041] [Medline: 35742292]
60. Diesel J, Sterrett N, Dasgupta S, Kriss JL, Barry V, Vanden Esschert K, et al. COVID-19 vaccination coverage among adults - United States, December 14, 2020-May 22, 2021. MMWR Morb Mortal Wkly Rep 2021 Jun 25;70(25):922-927 [FREE Full text] [doi: 10.15585/mmwr.mm7025e1] [Medline: 34166331]
61. Baack BN, Abad N, Yankey D, Kahn KE, Razzaghi H, Brookmeyer K, et al. COVID-19 vaccination coverage and intent among adults aged 18-39 years - United States, March-May 2021. MMWR Morb Mortal Wkly Rep 2021 Jun 25;70(25):928-933 [FREE Full text] [doi: 10.15585/mmwr.mm7025e2] [Medline: 34166337]

62. COVID-19 vaccine uptake and intent: Canadian community health survey (CCHS) insight. Public Health Agency of Canada. 2021. URL: https://www.canada.ca/en/public-health/services/publications/vaccines-immunization/covid-19-vaccine-uptake-intent-canadian-community-health-survey.html [accessed 2023-04-06]
63. Sharif A, Botlero R, Hoque N, Alif SM, Nazmul Karim M, Islam SM. A pragmatic approach to COVID-19 vaccine passport. BMJ Glob Health 2021 Oct;6(10):e006956 [FREE Full text] [doi: 10.1136/bmjgh-2021-006956] [Medline: 34610907]
64. Bosch-Capblanch X, Ronveaux O, Doyle V, Remedios V, Bchir A. Accuracy and quality of immunization information systems in forty-one low income countries. Trop Med Int Health 2009 Jan;14(1):2-10 [FREE Full text] [doi: 10.1111/j.1365-3156.2008.02181.x] [Medline: 19152556]
65. Baum U, Kulathinal S, Auranen K. Exposure misclassification bias in the estimation of vaccine effectiveness. PLoS One 2021 May 13;16(5):e0251622 [FREE Full text] [doi: 10.1371/journal.pone.0251622] [Medline: 33984065]

Abbreviations

CCEDRRN: Canadian COVID-19 Emergency Department Rapid Response Network
ED: emergency department
REDCap: Research Electronic Data Capture

Edited by A Mavragani, T Sanchez; submitted 24.11.22; peer-reviewed by S Olson, M Coccia; comments to author 15.03.23; revised version received 11.04.23; accepted 03.05.23; published 16.06.23.

Please cite as:
Archambault PM, Rosychuk RJ, Audet M, Bola R, Vatanpour S, Brooks SC, Daoust R, Clark G, Grant L, Vaillancourt S, Welsford M, Morrison LJ, Hohl CM, Canadian COVID-19 Emergency Department Rapid Response Network (CCEDRRN) investigators, Network of Canadian Emergency Researchers, Canadian Critical Care Trials Group
Accuracy of Self-Reported COVID-19 Vaccination Status Compared With a Public Health Vaccination Registry in Québec: Observational Diagnostic Study
JMIR Public Health Surveill 2023;9:e44465
URL: https://publichealth.jmir.org/2023/1/e44465
doi:10.2196/44465
PMID:37327046

Original Paper

COVID-19 Vaccine Hesitancy Among Patients Recovered From COVID-19 Infection in Wuhan, China: Cross-Sectional Questionnaire Study

Yiman Huang[1*], MPH; Ling Zhang[1*], MPH; Jiaqi Fu[1], MPhil; Yijin Wu[1], MPhil; Hao Wang[1], MPhil; Weijun Xiao[1], MPhil; You Xin[1], BA; Zhenwei Dai[1], MPH; Mingyu Si[1], PhD; Xu Chen[1], MPH; Mengmeng Jia[1], PhD; Zhiwei Leng[1], PhD; Dan Cui[2], PhD; Xiaoyou Su[1], PhD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
[2]National Clinical Research Center for Respiratory Diseases, China-Japan Friendship Hospital, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Xiaoyou Su, PhD
School of Population Medicine and Public Health
Chinese Academy of Medical Sciences & Peking Union Medical College
31 BeiJiGe San Tiao, Dongcheng District
Beijing, 100730
China
Phone: 86 65105830
Email: suxiaoyou@hotmail.com

Abstract

Background: Although patients recovered from COVID-19 already have immunity gained from natural infection, they are still at risk of reinfection due to the emergence of new variants of COVID-19 and the diminishing of naturally acquired immunity over time. Vaccination is associated with efficacious protection against COVID-19 infection and could boost infection-acquired immunity; however, various COVID-19 survivors have not been vaccinated due to vaccine hesitancy.

Objective: The aim of this study was to investigate COVID-19 vaccine hesitancy and related factors among COVID-19 survivors.

Methods: A cross-sectional questionnaire survey was conducted among patients who recovered from COVID-19 infection in Wuhan, China, between June 10 and July 25, 2021. The questionnaire included sociodemographic information, items on COVID-19 infection, the COVID-19 vaccine hesitancy scale based on the 3Cs (complacency, convenience, and confidence) model, trust in vaccine manufacturers and health facilities, and reasons for the decision to accept COVID-19 vaccination. Multivariate logistic regression analysis was used to assess the factors influencing COVID-19 vaccine hesitancy.

Results: Among the 1422 participants, 538 (37.8%) were not vaccinated against COVID-19. The COVID-19–recovered patients who self-reported having a current unhealthy status expressed more hesitancy about the COVID-19 vaccine than those who perceived themselves to be healthy (odds ratio [OR] 0.45, 95% CI 0.28-0.71). Compared to the asymptomatic patients, patients with mild symptoms were more likely to receive a COVID-19 vaccine (OR 1.67, 95% CI 1.02-2.82). Regarding the 3Cs model, high complacency (P=.005) and low convenience (P=.004) were significant negative factors for COVID-19 vaccination. Trust in vaccine manufacturers and health facilities was a significant positive factor for COVID-19 vaccination (OR 1.14, 95% CI 1.09-1.19). “Self-needs” was the main reason for patients to receive the COVID-19 vaccine, whereas “already have antibodies and do not need vaccination” was the main reason for patients to not receive the COVID-19 vaccine.

Conclusions: Among the three major factors of vaccine hesitancy, complacency proved to be the most notable among COVID-19–recovered patients. Therefore, educational campaigns can focus on raising the awareness of risk of infection and the benefits of vaccination to reduce complacency toward vaccination among this population. In particular, for individuals who have recovered from COVID-19, improving factors related to convenience such as transportation, the environment of vaccination, and providing door-to-door service was also deemed necessary to facilitate their vaccination. In addition, addressing the concerns about vaccination of COVID-19–recovered patients could foster trust and promote their uptake of vaccination.

(JMIR Public Health Surveill 2023;9:e42958) doi:10.2196/42958

KEYWORDS

COVID-19; COVID-19 survivors; vaccine hesitancy; complacency; confidence; convenience; cross-sectional questionnaire; health education; health promotion; public health

Introduction

The COVID-19 pandemic has caused a serious disease burden, as well as financial, psychological, and life hardship for people over the past 3 years. Currently, the high transmissibility, high pathogenicity, and high immune-evasion ability of the emerging new variants pose new challenges and uncertainties in preventing COVID-19 infection, as well as posing a serious threat to public health [1-3]. Previous studies have demonstrated that individuals naturally infected with COVID-19 might be at decreased future risk of COVID-19 infection due to the immunity induced by contracting COVID-19 to protect them from reinfection [4]. However, according to the World Health Organization (WHO), the presence of antibodies in recovered patients does not guarantee protection against reinfection, and the probability of protection against reinfection at 6 months was estimated to be 50.0% in people aged over 65 years [5-7]. Reinfection of COVID-19 could be as severe or even more severe than the first infection, including among patients with antibodies [6,8]. Therefore, since previous infection does not necessarily protect an individual from reinfection, it is necessary for patients who recovered from COVID-19 infection to protect against reinfection through COVID-19 vaccination [9].

Vaccination is considered to be one of the greatest achievements of public health. The ongoing COVID-19 pandemic can be mitigated by an efficacious vaccine, which reduces disease incidence, prevalence, new hospitalizations, and intensive care demand [10,11]. The COVID-19 vaccine has been shown to provide significant protection against COVID-19 infection and to significantly reduce the risk of symptomatic COVID-19 infection. Moreover, COVID-19 vaccination could boost infection-acquired immunity, and this increased immunity would remain high for more than 1 year after infection [12,13]. In 2021, five COVID-19 vaccines were approved for conditional marketing or emergency use in China, including inactivated vaccines, adenovirus vector vaccines, and recombinant protein vaccines, and the Chinese government provided nationwide COVID-19 vaccination for all populations free of charge. As of July 25, 2021, approximately 155 million cumulative doses of the COVID-19 vaccine had been administered in mainland China [14]. With the steady increase in COVID-19 vaccine supplies, vaccine hesitancy is becoming a barrier to high vaccine coverage [15]. There are still many people who delay or reject being vaccinated due to vaccine hesitancy even though they are recommended to accept the COVID-19 vaccination, including COVID-19 survivors. In the United States, vaccination coverage (>1 dose and full vaccination) was lower among those who ever had COVID-19 than among those who had no history of COVID-19 infection from July to August 2021 [16]. In Italy, 34.2% and 24.9% of COVID-19 patients were undecided or reluctant, respectively, to receive the COVID-19 vaccine from September to November 2020 [17].

The SAGE Working Group on Vaccine Hesitancy defines vaccine hesitancy as a delay in acceptance or refusal of vaccination despite the availability of vaccination services. Vaccine hesitancy is complex and context-specific, varying across time, place, and vaccines [18]. Factors such as the evolving epidemiological context and multiple waves of infection, trust in the health care system, attitudes toward vaccines, self-efficacy, and the presence of chronic disease have all been indicated to be associated with hesitancy to receive the COVID-19 vaccine [19-23]. In particular, COVID-19–recovered patients might experience not only physical disease symptoms but also psychological distress [24,25]. Some of these individuals might be hesitant to get vaccinated, fearing that they are not physically able to tolerate the vaccine. In addition, since other infectious diseases such as measles confer immunity for a longer duration or for life, some patients might consider that infection-acquired immunity with COVID-19 is the best type of immunity. Moreover, the concerns about the safety of vaccines developed in a short time might also be a factor contributing to COVID-19 vaccine hesitancy [26].

Vaccine hesitancy as a complex decision-making process has been interpreted based on various conceptual models such as the “3Cs” model, the health belief model, and the theory of planned behavior [27,28]. The “3Cs” model is one of the most widely known models, highlighting three dimensions: complacency, referring to the belief that perceived risks of vaccine-preventable diseases are low and that vaccination is not a necessary preventive action; convenience, referring to vaccine availability and accessibility; and confidence, referring to the trust in the effectiveness and safety of vaccines, the delivery system, and the motivations of vaccination policy makers [29]. Previous studies demonstrated that complacency, convenience, and confidence were equally significant factors influencing COVID-19 vaccine hesitancy [30-32].

Given the lack of data on COVID-19 vaccination status among COVID-19–recovered patients infected with the original SARS-CoV-2 strain in China, the aim of this study was to investigate COVID-19 vaccine hesitancy status and relevant factors among COVID-19–recovered patients. In particular, we focused on the role of the factors considered in the 3Cs model to identify the specific constructs that can influence COVID-19 vaccine hesitancy. In addition, the vaccination status and possible determinants of vaccine hesitancy in terms of sociodemographic information; clinical classification; history of infection in family members, relatives, and friends; and other vaccination history were investigated among COVID-19–recovered patients. This study will provide a basis for developing vaccination promotion activities and strategies for COVID-19–recovered patients to reduce their vaccine hesitancy and promote them to receive the COVID-19 vaccine.

Methods

Study Design and Sampling

This cross-sectional study was carried out among former COVID-19 patients in Jianghan District of Wuhan, China, from

June 10 to July 25, 2021. According to the electronic medical records of the Health Bureau of Jianghan District and the inclusion criteria, a total of 3059 recovered COVID-19 patients were eligible for the study who had all been infected with the original SARS-CoV-2 strain and were diagnosed between December 10, 2019, and April 20, 2020. Among them, 1601 COVID-19 survivors were invited to complete a questionnaire survey on their vaccine hesitancy status when they were receiving clinical reexamination. If they agreed to participate in this survey, they were invited to the "Survey Room" to complete the questionnaire. Self-administered electronic questionnaires were generated on Research Electronic Data Capture (REDCap), an online survey platform, from which patients could complete by themselves. To ensure the quality of the survey, our trained investigators stayed at the "Survey Room" to promptly answer the participants' questions.

The following inclusion requirements had to be met by participants: participants must be at least 18 years old, have a COVID-19 infection history, be able to work independently with the researcher to complete several scale assessments, and have access to a mobile communication device such as a cell phone with a WeChat account. People who fit one or more of the following criteria were excluded: (1) having significant cognitive impairment; (2) having life-threatening medical conditions such as heart, lung, kidney, liver diseases or cancers; and (3) finding it challenging to cooperate with the questionnaire study. In total, 1422 of the 1601 invited participants were chosen as the sample for this study based on the aforementioned standards and after discarding any incomplete questionnaires.

Ethical Considerations

Ethics approval for the study was obtained from the Ethics Review Committee of the Institute of Pathogen Biology, Chinese Academy of Medical Sciences, Beijing, China (IPB-2020-22). All individuals provided digital informed consent to ensure their voluntary participation, which also included consent that the study data could be analyzed when used anonymously. All data were deidentified. All data are stored in an account with a password and cannot be used without consent. All study participants were compensated for transportation to the questionnaire site.

Measurements

Sociodemographic Characteristics

Demographic characteristics included gender, age, place of residence, education level, marital status, smoking (whether or not they are habitual smokers), drinking (whether or not they habitually drink alcohol), and perceived current health status. We also investigated COVID-19 clinical classification by asking participants which type of diagnosis they had after their first admission (ie, asymptomatic, mild, moderate, clinically severe) and their acute-phase symptoms (ie, fever, respiratory symptoms, cardiovascular symptoms, gastrointestinal symptoms, other). Questions on COVID-19 infection and vaccination were also surveyed, including history of infection in family members, history of infection in relatives and friends, and other vaccination history (within the last 5 years).

COVID-19 Vaccine Hesitancy Scale Based on the 3Cs Model

On the basis of the flu vaccine hesitancy scale, we replaced "flu vaccine" in the items with "COVID-19 vaccine" to form the revised COVID-19 vaccine hesitancy scale. The flu vaccine hesitancy scale consists of 6 items and the 3 dimensions of the 3Cs model (complacency, confidence, and convenience) [33]. Complacency was measured by perceived necessity and importance of the vaccine, confidence was measured by perceived vaccine safety and effectiveness, and convenience was measured by perceived convenience and affordability of the vaccine. Participants rated each item on a 5-point Likert scale (1=strongly disagree; 5=strongly agree). The items were as follows: (1) necessity ("Thinking specifically about the COVID-19 vaccine, do you think the COVID-19 vaccine is necessary?"), (2) importance ("Thinking specifically about the COVID-19 vaccine, do you think the COVID-19 vaccine is important?"), (3) safety ("Thinking specifically about the COVID-19 vaccine, do you think the COVID-19 vaccine is safe?"), (4) effectiveness ("Thinking specifically about the COVID-19 vaccine, do you think the COVID-19 vaccine is effective?"), (5) convenience ("Thinking specifically about the COVID-19 vaccine, do you think the COVID-19 vaccine is convenient?"), and (6) affordability ("Thinking specifically about the COVID-19 vaccine, do you think the COVID-19 vaccine is affordable?"). The Cronbach α value for the whole scale was .92, and the Cronbach α values for the dimensions of complacency, convenience, and confidence were .92, .94, and .73, respectively. [34].

Trust

Two items were used to measure participants' trust in vaccine manufacturers and health facilities, respectively, which were each scored on a 10-point Likert scale (1=strongly disagree; 10=strongly agree). The Cronbach α value for the two items was .92.

Reasons for Accepting and Not Accepting the COVID-19 Vaccine

For participants who had received the COVID-19 vaccine, multiple-choice questions were designed to investigate the corresponding reasons, with answers including: (1) self-needs, (2) recommended by health agencies, (3) recommended by others (eg, relatives, friends, neighbors), (4) recommended by those who have been vaccinated, (5) internet information, (6) free vaccination, and (7) others.

For participants who had not received the COVID-19 vaccine, multiple-choice questions were set up to investigate the reasons, with answers including: (1) don't know how to get reliable information about COVID-19 vaccine, (2) have received negative media reports about COVID-19 vaccines, (3) have had a bad experience with the health clinic or provider, (4) have had a bad experience or adverse reactions from previous vaccinations, (5) have been told that others have had adverse reactions from some vaccines, (6) fear of injections, (7) already have antibodies and do not require vaccination, (8) long waiting time for vaccination, and (9) others.

Statistical Analysis

Confirmatory Factor Analysis of the COVID-19 Vaccine Hesitancy Scale

To validate the COVID-19 vaccine hesitancy scale, we performed a confirmatory factor analysis. Model fit was evaluated by various goodness-of-fit indices, including the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker Lewis index (TLI) [35-37]. RMSEA values close to 0.06 or below were regarded as a good fit, 0.07 to 0.08 as moderate fit, 0.08 to 0.10 as marginal fit, and >0.10 as poor fit [38]. For the CFI and TLI, values close to 0.95 or above were regarded as good fit, values close to 0.90 and 0.95 as acceptable fit, and values approaching 0 as poor fit [39,40]. Convergent validity was assessed by average variance extracted (AVE) and composite reliability (CR), and convergent validity was considered high if the AVE was greater than 0.50, CR was greater than 0.70, and CR was greater than AVE. Discriminant validity was considered satisfactory if the correlation between the factor scores was significant and the correlation coefficient was less than the square root of the corresponding AVE [41,42].

Univariate and Multivariate Analysis

Categorical variables are summarized as frequencies and proportions, whereas continuous variables are summarized as mean (SD). Respondents who were vaccinated against COVID-19 indicated that they were not hesitant about the COVID-19 vaccine, whereas those who were not vaccinated against COVID-19 indicated that they had COVID-19 vaccine hesitancy. To identify any differences in the distributions of variables between the vaccinated and unvaccinated groups, the χ^2 test and *t* test were used for categorical and continuous variables, respectively. Using the unvaccinated population as a reference, multiple logistic regression analysis was used to assess the association between the investigated factors and COVID-19 vaccine hesitancy. The associations between dependent and independent variables were determined using the odds ratio (OR) with 95% CI and a *P* value <.05 was deemed to indicate statistical significance of the independent variables. Data were analyzed using Stata version 16.0.

Results

Confirmatory Factor Analysis of the COVID-19 Vaccine Hesitancy Scale

Figure 1 depicts the results of the confirmatory factor analysis for the model built in the sample, including the three dimensions of the 3Cs model. Separate confirmatory factor analysis based on the original structure of the "complacency" dimension (2 items), "confidence" dimension (2 items), and "convenience" dimension (2 items) were performed. The RMSEA value (0.076) indicated a moderate fit, the CFI value (0.994) indicated a good fit, and the TLI value (0.986) indicated a good fit.

Table 1 shows the results of convergent validity and discriminant validity of the model. For convergent validity, the AVE values for all dimensions of the model were greater than 0.50 and the CR values for all three dimensions of the model were greater than 0.70. In addition, the CR values for all factors in the model were greater than the AVE. For discriminant validity, the correlation coefficients for any two dimensions were less than the square root of the corresponding AVE, except for the correlation coefficients for confidence and convenience, which were higher than the square root of the corresponding AVE.

Figure 1. Results of the confirmatory factor analysis model. Standardized covariances between dimensions are shown in bold, standardized path coefficients are italicized, and standardized variances are also shown.

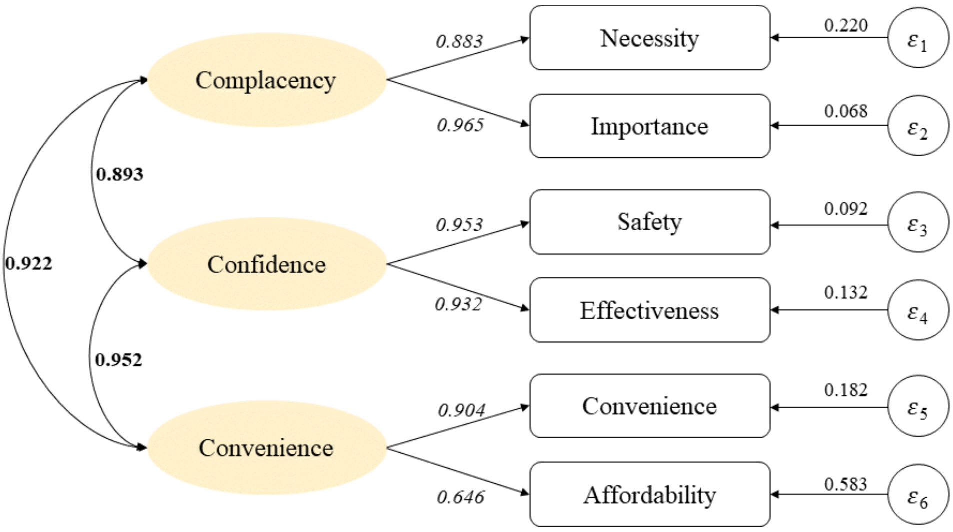


Table 1. Convergent and discriminant validity of the COVID-19 vaccine hesitancy model.

Dimensions	Pearson correlation coefficient			AVE[a]	CR[b]
	Complacency	Confidence	Convenience		
Complacency	*0.925*[c]	N/A[d]	N/A	0.855	0.922
Confidence	–0.827*	*0.942*	N/A	0.888	0.941
Convenience	–0.759*	0.796*	0.785	0.617	0.758

[a] AVE: average variance extracted.

[b] CR: composite reliability.

[c] Values on the diagonal are the square roots of each AVE value for comparison with other correlation coefficients; values in italics indicate a higher square root of the AVE value in each subscale than the correlation coefficients with other subscales.

[d] N/A: not applicable.

*P<.001.

Sociodemographic Characteristics of Participants

Among the 1422 participants, 538 (37.8%) were not vaccinated against COVID-19. There was a higher proportion of participants aged 41 to 60 years (634/1422, 45.2%), followed by the 61-80 years age group (611/1422, 43%). The great majority of the participants lived in urban areas (1245/1422, 87.6%), and 1197 (84.2%) participants were married. The majority (968/1422, 68.1%) had an education level of senior high school or below. Nearly half of the participants had a household income for 2020 lower than 60,000 Yuan (approximately US $8607) per year. Most participants self-reported that they were not current smokers (1233/1422, 86.7%) or drinkers (1038/1422, 73.0%), and 6.1% (87/1422) of the patients were asymptomatic, 70.8% (1007/1422) had mild symptoms of COVID-19, 10.3% (146/1422) had moderate symptoms, and 12.8% (182/1422) had critically severe symptoms. In the acute phase, 62.7% (891/1422) of patients had fever, 49.9% (709/1422) had respiratory symptoms, 18.2% (259/1422) had gastrointestinal symptoms, 6.4% (91/1422) had cardiovascular symptoms, and 15.5% (221/1422) had other symptoms. Overall, 24.1% (342/1422) were reinfected with COVID-19 after the first discharge, nearly half of the participants (654/1422, 46%) had family members who had been infected with COVID-19, 30.2% (429/1422) had relatives or friends who had been infected with COVID-19, and 17.2% (245/1422) of the participants had received other vaccines within the last 5 years (Table 2).

Table 2 also shows that the age gap between vaccinated and unvaccinated participants was statistically significant; participants aged ≤40 years had a higher hesitancy rate compared to the other three age groups (*P*<.001). The hesitancy rate of those living in urban areas (485/1245, 39.0%) was significantly higher than that of participants living in a village (53/177, 29.9%). Significantly, participants who were married had a lower hesitancy rate (428/1197, 35.8%) compared to that of participants who were single (23/54, 42.6%) and those with another marital status (87/171, 50.9%). Participants who perceived having an unhealthy status had a higher hesitancy rate than that of participants who self-reported being healthy (*P*<.001). Patients who had fever in the acute phase had a higher COVID-19 vaccine hesitancy rate (359/891, 40.3%) than those who did not have fever (179/531, 33.7%).

Table 2. Sociodemographic characteristics of participants.

Variables	Unvaccinated (n=538)	Vaccinated (n=884)	Total (N=1422)	*P* value
Gender, n (%)				.45
Male	221 (36.7)	381 (63.3)	602 (42.3)	
Female	317 (38.7)	503 (61.3)	820 (57.7)	
Age (years), n (%)				<.001
≤40	70 (48.3)	75 (51.7)	145 (10.2)	
41-60	240 (37.3)	403 (62.7)	643 (45.2)	
61-80	207 (33.9)	404 (66.1)	611 (43.0)	
≥81	21 (87.5)	2 (12.5)	23 (1.6)	
Residence place, n (%)				.02
Urban	485 (39.0)	760 (61.0)	1245 (87.6)	
Village	53 (29.9)	124 (70.1)	177 (12.4)	
Marital status, n (%)				.001
Single	23 (42.6)	31 (57.4)	54 (3.8)	
Married	428 (35.8)	769 (64.2)	1197 (84.2)	
Others	87 (50.9)	84 (49.1)	171 (12.0)	
Education level, n (%)				.02
Senior high school or below	347 (35.9)	621 (64.1)	968 (68.1)	
College and above	191 (42.1)	263 (57.9)	454 (31.9)	
Income for 2020 (CNY[a]/year), n (%)				.20
<60,000 Yuan	310 (36.9)	531 (63.1)	841 (59.1)	
60,000-120,000 Yuan	135 (36.5)	235 (63.5)	370 (26.0)	
130,000-300,000 Yuan	84 (44.9)	103 (55.1)	187 (13.2)	
>300,000 Yuan	9 (37.5)	15 (62.5)	24 (1.7)	
Have underlying diseases, n (%)				.06
Yes	315 (40.0)	473 (60.0)	788 (55.4)	
No	223 (35.2)	411 (64.8)	634 (44.6)	
Perceived current health status, n (%)				<.001
Healthy	482 (36.5)	840 (63.5)	1322 (93.0)	
Unhealthy	56 (56.0)	44 (44.0)	100 (7.0)	
Current smoker, n (%)				.38
Yes	77 (40.7)	112 (59.3)	189 (13.3)	
No	461 (37.4)	772 (62.6)	1233 (86.7)	
Alcohol use, n (%)				.31
Yes	137 (35.7)	247 (64.3)	384 (27.0)	
No	401 (38.6)	634 (61.4)	1038 (73.0)	
Clinical classification of COVID-19 patients, n (%)				.17
Asymptomatic	39 (44.8)	48 (55.2)	87 (6.1)	
Mild	363 (36.1)	644 (63.9)	1007 (70.8)	
Moderate	60 (41.1)	86 (58.9)	146 (10.3)	
Critically severe	76 (41.8)	106 (58.2)	182 (12.8)	
Acute-phase symptoms, n (%)				
Fever				.01

Variables	Unvaccinated (n=538)	Vaccinated (n=884)	Total (N=1422)	*P* value
Yes	359 (40.3)	532 (59.7)	891 (62.7)	
No	179 (33.7)	352 (66.3)	531 (37.3)	
Respiratory symptoms				.85
Yes	270 (38.1)	439 (61.9)	709 (49.9)	
No	268 (37.6)	445 (62.4)	713 (50.1)	
Gastrointestinal symptoms				.48
Yes	103 (39.8)	156 (60.2)	259 (18.2)	
No	435 (37.4)	728 (62.6)	1163 (81.8)	
Cardiovascular symptoms				.59
Yes	32 (35.2)	59 (64.8)	91 (6.4)	
No	506 (38.0)	825 (62.0)	1331 (93.6)	
Other symptoms				.59
Yes	80 (36.2)	141 (63.8)	221 (15.5)	
No	458 (38.1)	743 (61.9)	1201 (84.5)	
Reinfection of COVID-19 after the first discharge, n (%)				.23
Yes	139 (40.6)	203 (59.4)	342 (24.0)	
No	323 (37.9)	530 (62.1)	853 (60.0)	
Not hospitalized (not sure)	76 (33.5)	151 (66.5)	227 (16.0)	
Have family members infected with COVID-19, n (%)				.48
Yes	241 (36.9)	413 (63.1)	654 (46.0)	
No	297 (38.7)	471 (61.3)	768 (54.0)	
Have relatives or friends infected with COVID-19, n (%)				.22
Yes	152 (35.4)	277 (64.6)	429 (30.2)	
No	386 (38.9)	607 (61.1)	993 (69.8)	
Have received other vaccinations within the last 5 years, n (%)				.07
Yes	80 (32.7)	165 (67.3)	245 (17.2)	
No	458 (38.9)	719 (61.1)	1177 (82.8)	

[a] CNY: Chinese Yuan renminbi (1 CNY was equivalent to approximately US $6.98 in 2020).

Contributions of 3Cs Model Constructs and Trust to Vaccine Hesitancy

Figure 2 shows the distribution of answers on the COVID-19 vaccine hesitancy scale based on the 3Cs model among all participants. Table 3 shows the univariate analysis results of 3Cs model components and trust in vaccinated and unvaccinated groups. The unvaccinated group scored significantly higher on “complacency” related to COVID-19 vaccination than the vaccinated group. In contrast, participants who had received the COVID-19 vaccine scored higher on “confidence” and “convenience” than those who were not vaccinated. Moreover, the vaccinated group scored significantly higher than the unvaccinated group on trust (all *P*<.001).

Figure 2. Distribution of answers on the COVID-19 vaccine hesitancy scale according to the 3Cs model (complacency, convenience, confidence).

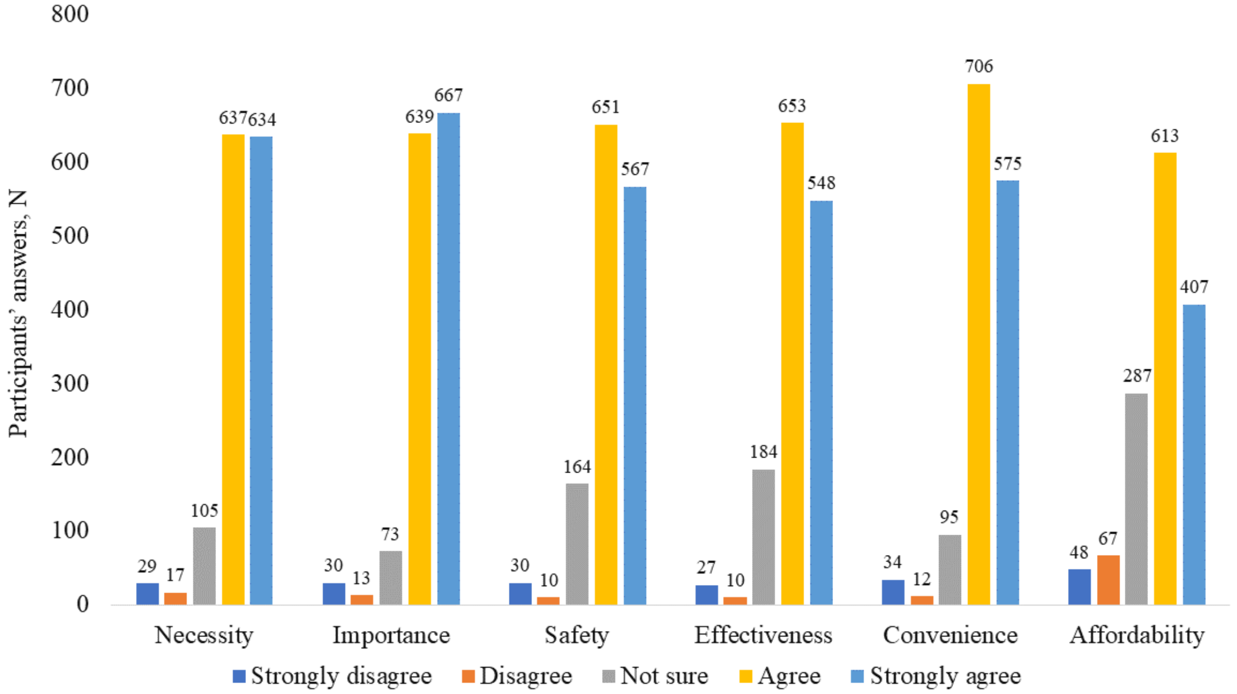


Table 3. Univariate analysis of the 3Cs (complacency, confidence, convenience) model constructs and trust in vaccinated and unvaccinated groups.

Variables	Scores, mean (SD)			*P* value
	Unvaccinated (n=538)	Vaccinated (n=884)	Total (N=1422)	
3Cs Complacency	3.82 (1.69)	3.11 (1.40)	3.38 (1.55)	<.001
3Cs Confidence	7.96 (1.65)	8.66 (1.52)	8.39 (1.61)	<.001
3Cs Convenience	7.66 (1.62)	8.43 (1.52)	8.14 (1.60)	<.001
Trust	17.02 (3.39)	18.42 (2.44)	17.89 (2.91)	<.001

Multivariate Predictors of COVID-19 Vaccine Hesitancy

Table 4 shows that participants aged 41-60 years (OR 2.12, 95% CI 1.34-3.36) and 61-80 years (OR 2.59, 95 %CI 1.56-4.30) were more likely to receive a COVID-19 vaccine than participants aged ≤40 years; however, participants aged >80 years (OR 0.14, 95% CI 0.03-0.66) were less likely to receive a COVID-19 vaccine than those aged <40 years. Participants who reported "other" marital status (ie, divorced, widowed) were more likely to be hesitant about receiving the COVID-19 vaccine than those who were single (OR 0.42, 95% CI 0.19-0.91). In addition, patients who perceived having an unhealthy status had more hesitancy about the COVID-19 vaccine than those who were healthy (OR 0.45, 95% CI 0.28-0.71). Compared to the asymptomatic patients, patients with mild symptoms were more likely to receive the COVID-19 vaccine (OR 1.67, 95% CI 1.02-2.82). Regarding the acute-phase symptoms, patients who did not have fever were less likely to receive the COVID-19 vaccine than those who had fever (OR 1.48, 95% CI 1.13-1.94). Not receiving any other vaccination within the last 5 years was negatively associated with receiving the COVID-19 vaccine (OR 0.69, 95% CI 0.50-0.95). Regarding the 3Cs model constructs, complacency was a significant negative factor for getting vaccinated for COVID-19 (OR 0.81, 95% CI 0.71-0.94), whereas convenience was a significant positive factor for COVID-19 vaccination (OR 1.21, 95% CI 1.06-1.37). Participants who had more trust were more likely to receive the COVID-19 vaccine (OR 1.14, 95% CI 1.09-1.19).

Table 4. Multivariate logistic regression analysis on factors contributing to COVID-19 vaccine hesitancy.[a]

Variables	OR[b] (SE)	95% CI	*P* value
Gender			
Male (reference)	N/A[c]	N/A	N/A
Female	0.85 (0.12)	0.63-1.13	.26
Age (years)			
≤40 (reference)	N/A	N/A	N/A
41-60	2.12 (0.50)	1.34-3.36	.001
61-80	2.59 (0.67)	1.56-4.30	<.001
≥81	0.14 (0.11)	0.03-0.66	.01
Place of residence			
Urban (reference)	N/A	N/A	N/A
Village	1.45 (0.29)	0.98-2.16	.07
Marital status			
Single (reference)	N/A	N/A	N/A
Married	0.73 (0.26)	0.37-1.45	.37
Others	0.42 (0.17)	0.19-0.91	.03
Education level			
Senior high school or below (reference)	N/A	N/A	N/A
College and above	0.85 (0.14)	0.63-1.18	.36
Income for 2020 (CNY[d]/year)			
<60,000 (reference)	N/A	N/A	N/A
60,000-120,000	1.00 (0.15)	0.75-1.34	>.99
130,000-300,000	0.77 (0.16)	0.52-1.15	.20
>300,000	0.88 (0.42)	0.35-2.22	.79
No underlying disease	1.32 (0.17)	1.02-1.70	.03
Perceived current health status			
Healthy (reference)	N/A	N/A	N/A
Unhealthy	0.45 (0.11)	0.28-0.71	.001
Nonsmoker	1.23 (0.24)	0.84-1.80	.28
No alcohol use	0.82 (0.13)	0.61-1.11	.20
Clinical classification of COVID-19			
Asymptomatic	N/A	N/A	N/A
Mild	1.70 (0.44)	1.02-2.821	.04
Moderate	1.52 (0.49)	0.81-2.85	.19
Critically severe	1.42 (0.44)	0.77-2.62	.27
Acute-phase symptoms (yes=reference)			
No fever	1.48 (0.21)	1.13-1.94	.01
No respiratory symptoms	0.97 (0.12)	0.75-1.24	.79
No gastrointestinal symptoms	1.07 (0.17)	0.78-1.47	.67
No cardiovascular symptoms	0.69 (0.18)	0.41-1.15	.15
No other symptoms	0.89 (0.15)	0.41-1.15	.15
No reinfection of COVID-19 after first discharge	1.07 (0.16)	0.81-1.43	.64

Variables	OR[b] (SE)	95% CI	*P* value
Not hospitalized (not sure)	1.50 (0.30)	1.01-2.22	.045
No family members infected with COVID-19	0.92 (0.12)	0.72-1.18	.53
No relatives or friends infected with COVID-19	0.83 (0.12)	0.63-1.10	.20
No other vaccinations within 5 years	0.69 (0.11)	0.50-0.95	.02
3Cs Complacency	0.81 (0.06)	0.71-0.94	.005
3Cs Confidence	0.92 (0.07)	0.79-1.06	.26
3Cs Convenience	1.21 (0.08)	1.06-1.37	.004
Trust	1.14 (0.03)	1.09-1.19	<.001

[a] Model fit indices: Pearson χ^2=1417.44, *P*=.25; Hosmer-Lemeshow χ^2=14.39, *P*=.07.

[b] OR: odds ratio.

[c] N/A: not applicable.

[d] CNY: Chinese Yuan renminbi (1 CNY was equivalent to approximately US $6.98 in 2020).

Reasons for Accepting and Not Accepting COVID-19 Vaccination

Among the reasons for getting vaccinated against COVID-19, "self-needs" was chosen most frequently (n=625), followed by "recommended by health agencies" (n=191) and "free vaccination" (n=149). The reason for "recommended by those who have been vaccinated" was selected 37 times, which was only higher than the "others" response option (n=11). Among the reasons for not getting vaccinated against COVID-19, "antibodies are already in existence and do not require vaccination" was chosen most frequently (n=88), followed by "don't know how to get reliable information about the COVID-19 vaccine" (n=53) and "others" (n=30). The reason "have had a bad experience with the health clinic or provider" was selected 4 times, which was the least frequently selected option (Table 5).

Table 5. Reasons for accepting and not accepting COVID-19 vaccination.

Reasons	Participants, n
Accepting COVID-19 vaccination	
Self-needs	625
Recommended by health agencies	191
Recommended by others (eg, relatives, friends, neighbors)	73
Recommended by those who have been vaccinated	37
Internet information	40
Free vaccination	149
Others	11
Not accepting COVID-19 vaccination	
Don't know how to get reliable information about COVID-19 vaccine	53
Have received negative media reports about COVID-19 vaccines	26
Have a bad experience with the health clinic or provider	4
Have a bad experience or adverse reactions from previous vaccinations	18
Have been told that others have had adverse reactions from vaccines	29
Fear of injections	25
Antibodies are already in existence and do not require vaccination	88
Long waiting time for vaccination	19
Others	30

Discussion

Principal Results

The WHO indicated that antibodies in COVID-19–recovered patients do not guarantee protection against reinfection, making it necessary to receive the COVID-19 vaccine after natural infection [7]. However, diverse and effective preventive measures such as wearing a mask provided an alternative to prevent COVID-19 and might undermine the perceived need for vaccination to the point of developing an attitude of COVID-19 vaccine hesitancy [43,44]. This study revealed that the rate of COVID-19 vaccine hesitancy among patients who recovered from COVID-19 was 37.8%. According to previous research, this vaccine hesitancy rate is higher than that determined in a similar time period among Chinese patients with chronic diseases, including HIV infection (27.5%), cancer (24.1%), and inflammatory bowel disease (27.4%) [45-47]. Even though COVID-19–recovered patients and patients with chronic diseases are both vulnerable to COVID-19 reinfection based on their poor health conditions, COVID-19–recovered patients have a lower perceived risk of COVID-19 reinfection due to their belief that they already have antibodies gained from previous infection compared to patients with chronic diseases. Several studies also suggested lower vaccination intention among COVID-19–recovered individuals compared to uninfected individuals [48,49]. Research has shown that acquired immunity from COVID-19 infection would diminish over 6-8 months and protection against the new variant (Omicron variant) might be inadequate [50]; thus, COVID-19–recovered patients are still at risk of reinfection. The high perceived risk of contracting diseases is an essential determinant in overcoming vaccine hesitancy; that is, individuals who perceived a high risk of COVID-19 infection were assumed to adopt more preventive health behaviors to avoid or minimize health risks [51]. As the threat of reinfection increases due to continuously emerging new mutant strains, raising COVID-19–recovered patients' awareness of the reinfection risk is essential to reduce their vaccine hesitancy.

In this study, COVID-19 patients with mild symptoms were more likely to receive the COVID-19 vaccine than asymptomatic patients. Similarly, previous studies have shown that people who have experienced severe COVID-19 disease have a lower rate of hesitancy compared to those with less severe disease [17]. Patients who have experienced the negative health effects of COVID-19 infection were willing to be vaccinated because they did not want to experience these symptoms again and would have a greater fear of reinfection. Previous studies also suggested that the severity perception of COVID-19 infection would directly affect the intention to vaccinate against COVID-19; patients who perceived COVID-19 infection to be severe were the most likely to be vaccinated [52,53]. In addition, patients with symptoms may develop sequelae symptoms and thus believe that they are vulnerable and that their perceived risk of COVID-19 infection will be higher, leading to the belief that they need to be vaccinated. Results of a cohort study showed that 61.4% of patients infected with the original SARS-CoV-2 strain had at least one sequelae symptom [54]. The significance of vaccination was not limited to preventing COVID-19 infection, as vaccination in COVID-19–recovered patients was found to also be effective at preventing sequelae symptoms [55]. Several studies have shown that the vaccinated group had a lower risk of developing sequelae symptoms compared to the unvaccinated group [56-58]. Emphasizing the benefits of vaccination among hesitant patients, such as prevention of reinfection and prevention of sequelae, will raise their awareness of the need for vaccination and thus motivate unvaccinated COVID-19–recovered patients to get vaccinated.

Complacency occurs when individuals have a lower perception of the need for a vaccination or a perceived low risk from diseases, which is influenced by general health beliefs [59,60]. Similar to other studies, complacency had a significant effect on increasing COVID-19 vaccine hesitancy in this study [61,62]. The immunity from previous infections might be one of the reasons for the complacency among the patients who recovered from COVID-19 infection. Meanwhile, misconceptions about the efficacy or safety of vaccination, misinformation in popular social media, and the perception that vaccines may not offer better protection than previous infections also could lead to complacency and result in underacceptance of the COVID-19 vaccine [63-65]. Additionally, the belief that vaccination of those around them is sufficient to prevent transmission and protect themselves from COVID-19 reinfection could result in a lower perception of COVID-19 reinfection risk among COVID-19–recovered patients. According to our study, among the reasons why COVID-19 survivors were willing to be vaccinated, most participants chose "self-needs," which could eventually lead to high vaccine uptake [66]. By reducing their beliefs about complacency, this group of COVID-19 survivors might realize that vaccination is necessary for them. Specifically, health systems and relevant authorities should provide valuable information highlighting the evidence that immunity gained from natural infection would diminish over time and the risks of failure to vaccinate to enhance their "self-needs." This would increase vaccination rates among COVID-19 survivors and ensure that they have sufficient protection against reinfection.

Convenience was a significant factor affecting vaccine hesitancy. Previous vaccination programs have shown that obstacles to vaccination include limited access to information, difficulties obtaining vaccines, unaffordable vaccine prices, and the long time (or distance) required to receive a vaccine [67-69]. Our results showed that "long waiting time for vaccination" and "have had a bad experience with the health clinic or provider" were among the convenience-related reasons participants stated for not receiving the COVID-19 vaccine. "Free vaccination" was among the reasons why participants received a COVID-19 vaccine, which is related to high convenience. Hence, to enhance COVID-19–recovered patients' convenience in getting vaccinated for COVID-19, relevant health facilities should optimize the vaccination process to shorten waiting or queueing times, further improve the vaccination environment, and train vaccination-related workers to make them knowledgeable about the vaccine and to be more patient when answering questions [65]. In addition, targeting COVID-19–recovered patients with mobility issues and offering door-to-door vaccinations to address accessibility barriers would be very effective measures to increase vaccination rates.

Trust is a strong driver in reducing vaccine hesitancy. People with a higher level of trust in health authorities would have a more positive perception of the COVID-19 vaccine [70,71]. This study utilized two questions to evaluate participants' trust, and the results also showed that trust was a motivating factor for patients who recovered from COVID-19 to get immunized. However, a crisis of trust in the COVID-19 vaccine, vaccine manufacturers, and health facilities unavoidably arose as a result of vaccine safety–related events and the dissemination of false information since the COVID-19 outbreak. Trust would build when people feel that health authorities possess knowledge and expertise; that the authorities take into account all relevant opinions; and that the authorities are transparent, honest, and open [72]. To assist in fostering trust and boosting confidence in the COVID-19 vaccine, identifying the concerns of patients who recovered from COVID-19, providing accurate information, and establishing communication channels with the health authorities are necessary. Ensuring appropriate, effective, and more specific education targeted at patients who recovered from COVID-19, while expressing the overall benefits and risks of the vaccine would also be beneficial to build trust to promote vaccination.

Since the availability of COVID-19 vaccines, we have dependable, high-quality evidence attesting to their safety, effectiveness, and value of protecting people from COVID-19 infection. However, our research has shown that patients are likely to be hesitant to receive vaccines because of complacency, convenience, trust, and other issues. In December 2022, a large-scale outbreak of COVID-19 infection occurred in China that lasted for 2-3 months, resulting in a high proportion of the population being in the category of COVID-19–recovered patients. These individuals will inevitably have to decide whether to get another dose (booster shot) if the COVID-19 epidemic persists and new variants continue to emerge. The study team conducted an online survey between January 5, 2023, and February 9, 2023, among people in seven geographic subdivisions of China regarding their willingness to receive a booster shot. A total of 7070 valid questionnaires were collected among these COVID-19–recovered patients and the COVID-19 vaccine hesitancy rate was 43.54% (data not shown). This high rate indicates that vaccine hesitancy still exists among COVID-19–recovered patients in China since the latest surge of the COVID-19 pandemic. The findings of this study may contribute to providing a research base and support for interventions to reduce future hesitancy related to COVID-19 booster shots among COVID-19–recovered patients, since future booster shots and the COVID-19 vaccination services and procedures that are currently in progress may converge.

Limitations

This study has several limitations. First, the study was conducted in Wuhan, China, and a convenience sampling approach was employed for the survey, which may result in potentially biased estimates and selection bias. Therefore, the generalizability of our results will still be limited in certain aspects. Second, this was a cross-sectional study using a self-report questionnaire, leading to the presence of information bias. Third, since these data were collected between June and July 2021, the investigation was carried out in a very dynamic and ever-changing context, and it is possible that current perceptions of vaccines have changed both in terms of the perceived disease threat to study participants and the development of the COVID-19 vaccine itself.

Conclusions

COVID-19–recovered patients may have a low perception of the risk of reinfection because they already have immunity acquired through natural infection, and this may lead to complacency, ultimately leading to COVID-19 vaccine hesitancy. Advocacy strategies based on scientific evidence to raise public awareness of the risk of reinfection and the superiority of immunity gained from vaccination over that gained from infection can be effective in reducing complacency and thus overcoming vaccine hesitancy. Improving convenience-related factors such as time, transportation, and environment to get vaccinated is also necessary to address accessibility barriers and facilitate vaccination uptake. Furthermore, education efforts targeted at individuals who recovered from COVID-19 based on solid and adequate knowledge are needed to address their concerns about vaccination, which could foster greater trust and promote their vaccination acceptance.

Acknowledgments

This research was funded by the Asian Regional Special Cooperation Fund of the National Health Commission of the People's Republic of China (BLXM01) and the Innovative Engineering Program sponsored by the Chinese Academy of Medical Sciences (2020-I2M-2-015).

Data Availability

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

YH, LZ, YW, and JF prepared the revised manuscript. YH and HW prepared the first draft. XS provided overall guidance and managed the overall project. WX, HW, YX, MS, JF, XC, MJ, ZL, and DC were responsible for the questionnaire survey and data management. YH analyzed the data. YH, LZ, and XS finalized the manuscript on the basis of comments from other authors. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Challen R, Brooks-Pollock E, Read JM, Dyson L, Tsaneva-Atanasova K, Danon L. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. BMJ 2021 Mar 09;372:n579 [FREE Full text] [doi: 10.1136/bmj.n579] [Medline: 33687922]
2. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science 2021 May 21;372(6544):815-821 [FREE Full text] [doi: 10.1126/science.abh2644] [Medline: 33853970]
3. Fan Y, Li X, Zhang L, Wan S, Zhang L, Zhou F. SARS-CoV-2 Omicron variant: recent progress and future perspectives. Signal Transduct Target Ther 2022 Apr 28;7(1):141. [doi: 10.1038/s41392-022-00997-x] [Medline: 35484110]
4. Harvey RA, Rassen JA, Kabelac CA, Turenne W, Leonard S, Klesh R, et al. Association of SARS-CoV-2 seropositive antibody test with risk of future infection. JAMA Intern Med 2021 May 01;181(5):672-679 [FREE Full text] [doi: 10.1001/jamainternmed.2021.0366] [Medline: 33625463]
5. Fiolet T, Kherabi Y, MacDonald C, Ghosn J, Peiffer-Smadja N. Comparing COVID-19 vaccines for their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: a narrative review. Clin Microbiol Infect 2022 Feb;28(2):202-221 [FREE Full text] [doi: 10.1016/j.cmi.2021.10.005] [Medline: 34715347]
6. Nguyen NN, Houhamdi L, Hoang VT, Delerce J, Delorme L, Colson P, et al. SARS-CoV-2 reinfection and COVID-19 severity. Emerg Microbes Infect 2022 Dec;11(1):894-901 [FREE Full text] [doi: 10.1080/22221751.2022.2052358] [Medline: 35264078]
7. COVID-19 natural immunity. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci_Brief-Natural_immunity-2021.1 [accessed 2023-05-13]
8. Wang J, Kaperak C, Sato T, Sakuraba A. COVID-19 reinfection: a rapid systematic review of case reports and case series. J Investig Med 2021 Aug;69(6):1253-1255. [doi: 10.1136/jim-2021-001853] [Medline: 34006572]
9. Letizia AG, Ge Y, Vangeti S, Goforth C, Weir DL, Kuzmina NA, et al. SARS-CoV-2 seropositivity and subsequent infection risk in healthy young adults: a prospective cohort study. Lancet Respir Med 2021 Jul;9(7):712-720 [FREE Full text] [doi: 10.1016/S2213-2600(21)00158-2] [Medline: 33865504]
10. Haynes BF, Corey L, Fernandes P, Gilbert PB, Hotez PJ, Rao S, et al. Prospects for a safe COVID-19 vaccine. Sci Transl Med 2020 Nov 04;12(568):eabe0948. [doi: 10.1126/scitranslmed.abe0948] [Medline: 33077678]
11. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger J. Vaccine hesitancy: an overview. Hum Vaccin Immunother 2013 Aug;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
12. Hall V, Foulkes S, Insalata F, Kirwan P, Saei A, Atti A, SIREN Study Group. Protection against SARS-CoV-2 after Covid-19 vaccination and previous infection. N Engl J Med 2022 Mar 31;386(13):1207-1220 [FREE Full text] [doi: 10.1056/NEJMoa2118691] [Medline: 35172051]
13. Shrestha NK, Burke PC, Nowacki AS, Terpeluk P, Gordon SM. Necessity of coronavirus disease 2019 (COVID-19) vaccination in persons who have already had COVID-19. Clin Infect Dis 2022 Aug 24;75(1):e662-e671 [FREE Full text] [doi: 10.1093/cid/ciac022] [Medline: 35028662]
14. COVID-19 vaccination status. National Health Commission of the People's Republic of China. 2021. URL: http://www.nhc.gov.cn/xcs/yqjzqk/202107/8348635aeddd4ac58dcef6276fdbcda8.shtml [accessed 2021-03-17]
15. Adhikari B, Cheah PY. Vaccine hesitancy in the COVID-19 era. Lancet Infect Dis 2021 Aug;21(8):1086 [FREE Full text] [doi: 10.1016/S1473-3099(21)00390-X] [Medline: 34217431]
16. Nguyen KH, Huang J, Mansfield K, Corlin L, Allen JD. COVID-19 vaccination coverage, behaviors, and intentions among adults with previous diagnosis, United States. Emerg Infect Dis 2022 Mar;28(3):631-638. [doi: 10.3201/eid2803.211561] [Medline: 35202522]
17. Gerussi V, Peghin M, Palese A, Bressan V, Visintini E, Bontempo G, et al. Vaccine hesitancy among Italian patients recovered from COVID-19 infection towards influenza and Sars-Cov-2 vaccination. Vaccines 2021 Feb 18;9(2):172 [FREE Full text] [doi: 10.3390/vaccines9020172] [Medline: 33670661]
18. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
19. Wake AD. The willingness to receive COVID-19 vaccine and its associated factors: "Vaccination refusal could prolong the war of this pandemic" - a systematic review. Risk Manag Healthc Policy 2021;14:2609-2623 [FREE Full text] [doi: 10.2147/RMHP.S311074] [Medline: 34188572]
20. Karafillakis E, Van Damme P, Hendrickx G, Larson HJ. COVID-19 in Europe: new challenges for addressing vaccine hesitancy. Lancet 2022 Feb 19;399(10326):699-701 [FREE Full text] [doi: 10.1016/S0140-6736(22)00150-7] [Medline: 35123665]

21. Utami A, Margawati A, Pramono D, Nugraheni A, Pramudo SG. Determinant factors of COVID-19 vaccine hesitancy among adult and elderly population in Central Java, Indonesia. Patient Prefer Adherence 2022;16:1559-1570 [FREE Full text] [doi: 10.2147/PPA.S365663] [Medline: 35789823]
22. Chen X, Lee W, Lin F. Infodemic, institutional trust, and COVID-19 vaccine hesitancy: a cross-national survey. Int J Environ Res Public Health 2022 Jun 30;19(13):8033 [FREE Full text] [doi: 10.3390/ijerph19138033] [Medline: 35805691]
23. Lazarus JV, Wyka K, White TM, Picchio CA, Rabin K, Ratzan SC, et al. Revisiting COVID-19 vaccine hesitancy around the world using data from 23 countries in 2021. Nat Commun 2022 Jul 01;13(1):3801. [doi: 10.1038/s41467-022-31441-x] [Medline: 35778396]
24. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet 2021 Jan 16;397(10270):220-232 [FREE Full text] [doi: 10.1016/S0140-6736(20)32656-8] [Medline: 33428867]
25. Sheehy LM. Considerations for postacute rehabilitation for survivors of COVID-19. JMIR Public Health Surveill 2020 May 08;6(2):e19462 [FREE Full text] [doi: 10.2196/19462] [Medline: 32369030]
26. Dubé E, MacDonald NE. COVID-19 vaccine hesitancy. Nat Rev Nephrol 2022 Jul;18(7):409-410 [FREE Full text] [doi: 10.1038/s41581-022-00571-2] [Medline: 35414006]
27. Shmueli L. Predicting intention to receive COVID-19 vaccine among the general population using the health belief model and the theory of planned behavior model. BMC Public Health 2021 Apr 26;21(1):804 [FREE Full text] [doi: 10.1186/s12889-021-10816-7] [Medline: 33902501]
28. Zampetakis LA, Melas C. The health belief model predicts vaccination intentions against COVID-19: A survey experiment approach. Appl Psychol Health Well Being 2021 May 26;13(2):469-484 [FREE Full text] [doi: 10.1111/aphw.12262] [Medline: 33634930]
29. González-Block M, Gutiérrez-Calderón E, Pelcastre-Villafuerte BE, Arroyo-Laguna J, Comes Y, Crocco P, et al. Influenza vaccination hesitancy in five countries of South America. Confidence, complacency and convenience as determinants of immunization rates. PLoS One 2020;15(12):e0243833 [FREE Full text] [doi: 10.1371/journal.pone.0243833] [Medline: 33306744]
30. Liu X, Mesch GS. The adoption of preventive behaviors during the COVID-19 pandemic in China and Israel. Int J Environ Res Public Health 2020 Sep 30;17(19):7170 [FREE Full text] [doi: 10.3390/ijerph17197170] [Medline: 33007985]
31. Gerretsen P, Kim J, Caravaggio F, Quilty L, Sanches M, Wells S, et al. Individual determinants of COVID-19 vaccine hesitancy. PLoS One 2021;16(11):e0258462 [FREE Full text] [doi: 10.1371/journal.pone.0258462] [Medline: 34788308]
32. Kwok KO, Li K, Wei WI, Tang A, Wong SYS, Lee SS. Editor's Choice: Influenza vaccine uptake, COVID-19 vaccination intention and vaccine hesitancy among nurses: A survey. Int J Nurs Stud 2021 Feb;114:103854 [FREE Full text] [doi: 10.1016/j.ijnurstu.2020.103854] [Medline: 33326864]
33. Quinn SC, Jamison AM, An J, Hancock GR, Freimuth VS. Measuring vaccine hesitancy, confidence, trust and flu vaccine uptake: Results of a national survey of White and African American adults. Vaccine 2019 Feb 21;37(9):1168-1173. [doi: 10.1016/j.vaccine.2019.01.033] [Medline: 30709722]
34. DeVellis R, Thorpe C. Scale development: Theory and applications. Thousand Oaks, CA: Sage Publications; 2021.
35. Steiger JH. Structural model evaluation and modification: an interval estimation approach. Multivariate Behav Res 1990 Apr 01;25(2):173-180. [doi: 10.1207/s15327906mbr2502_4] [Medline: 26794479]
36. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990 Mar;107(2):238-246. [doi: 10.1037/0033-2909.107.2.238] [Medline: 2320703]
37. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika 1973 Mar;38(1):1-10. [doi: 10.1007/BF02291170]
38. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equat Model 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
39. Gomez R, Stavropoulos V. Confirmatory factor analysis and exploratory structural equation modeling of the structure of attention-deficit/hyperactivity disorder symptoms in adults. Assessment 2021 Sep;28(6):1570-1582. [doi: 10.1177/1073191120905892] [Medline: 32062977]
40. Browne M, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park, CA: Sage Publications; 1993:136-162.
41. Bagozzi RP, Yi Y. Specification, evaluation, and interpretation of structural equation models. J Acad Market Sci 2011 Aug 21;40(1):8-34. [doi: 10.1007/s11747-011-0278-x]
42. Fornell C, Larcker D. Structural equation models with unobservable variables and measurement error: algebra and statistics. J Market Res 2018 Nov 28;18(3):382-388. [doi: 10.1177/002224378101800313]
43. Chan NN, Ong KW, Siau CS, Lee KW, Peh SC, Yacob S, et al. The lived experiences of a COVID-19 immunization programme: vaccine hesitancy and vaccine refusal. BMC Public Health 2022 Feb 14;22(1):296 [FREE Full text] [doi: 10.1186/s12889-022-12632-z] [Medline: 35164734]
44. Kumar D, Mathur M, Kumar N, Rana RK, Tiwary RC, Raghav PR, et al. Understanding the phases of vaccine hesitancy during the COVID-19 pandemic. Isr J Health Policy Res 2022 Mar 22;11(1):16 [FREE Full text] [doi: 10.1186/s13584-022-00527-8] [Medline: 35317859]

45. Liu Y, Han J, Li X, Chen D, Zhao X, Qiu Y, et al. COVID-19 vaccination in people living with HIV (PLWH) in China: a cross sectional study of vaccine hesitancy, safety, and immunogenicity. Vaccines 2021 Dec 09;9(12):1458 [FREE Full text] [doi: 10.3390/vaccines9121458] [Medline: 34960204]
46. Hu S, Liu J, Li S, Wu Q, Wang X, Xu D, et al. Patients with IBD have a more cautious attitude towards COVID-19 vaccination. Front Immunol 2022;13:1077308 [FREE Full text] [doi: 10.3389/fimmu.2022.1077308] [Medline: 36741375]
47. Hong J, Xu X, Yang J, Zheng J, Dai S, Zhou J, et al. Knowledge about, attitude and acceptance towards, and predictors of intention to receive the COVID-19 vaccine among cancer patients in Eastern China: A cross-sectional survey. J Integr Med 2022 Jan;20(1):34-44 [FREE Full text] [doi: 10.1016/j.joim.2021.10.004] [Medline: 34774463]
48. Noushad M, Nassani MZ, Koppolu P, Alsalhani AB, Samran A, Alqerban A, et al. Predictors of COVID-19 vaccine intention among the Saudi Arabian population: a cross-sectional survey. Vaccines 2021 Aug 12;9(8):892 [FREE Full text] [doi: 10.3390/vaccines9080892] [Medline: 34452017]
49. Kanyangarara M, McAbee L, Daguise VG, Nolan MS. Factors associated with COVID-19 vaccine intentions among South Carolina residents. Vaccines 2022 Jun 14;10(6):942 [FREE Full text] [doi: 10.3390/vaccines10060942] [Medline: 35746550]
50. Altarawneh HN, Chemaitelly H, Hasan MR, Ayoub HH, Qassim S, AlMukdad S, et al. Protection against the Omicron variant from previous SARS-CoV-2 infection. N Engl J Med 2022 Mar 31;386(13):1288-1290. [doi: 10.1056/nejmc2200133]
51. Du M, Tao L, Liu J. The association between risk perception and COVID-19 vaccine hesitancy for children among reproductive women in China: an online survey. Front Med 2021 Sep 8;8:741298 [FREE Full text] [doi: 10.3389/fmed.2021.741298] [Medline: 34568394]
52. Petravić L, Arh R, Gabrovec T, Jazbec L, Rupčić N, Starešinič N, et al. Factors affecting attitudes towards COVID-19 vaccination: an online survey in Slovenia. Vaccines 2021 Mar 12;9(3):247 [FREE Full text] [doi: 10.3390/vaccines9030247] [Medline: 33808958]
53. AlAwadhi E, Zein D, Mallallah F, Bin Haider N, Hossain A. Monitoring COVID-19 vaccine acceptance in Kuwait during the pandemic: results from a national serial study. Risk Manag Healthc Policy 2021;14:1413-1429 [FREE Full text] [doi: 10.2147/RMHP.S300602] [Medline: 33854390]
54. Cui D, Chen S, Feng L, Jia M, Wang Y, Xiao W, et al. Long-term outcomes in COVID-19 patients who recovered from the first wave of the pandemic. Natl Sci Rev 2022 Nov;9(11):nwac192 [FREE Full text] [doi: 10.1093/nsr/nwac192] [Medline: 36478734]
55. Gao P, Liu J, Liu M. Effect of COVID-19 vaccines on reducing the risk of long COVID in the real world: a systematic review and meta-analysis. Int J Environ Res Public Health 2022 Sep 29;19(19):12422 [FREE Full text] [doi: 10.3390/ijerph191912422] [Medline: 36231717]
56. Mohr NM, Plumb ID, Harland KK, Pilishvili T, Fleming-Dutra KE, Krishnadasan A, Project PREVENT Network. Presence of symptoms 6 weeks after COVID-19 among vaccinated and unvaccinated US healthcare personnel: a prospective cohort study. BMJ Open 2023 Feb 02;13(2):e063141 [FREE Full text] [doi: 10.1136/bmjopen-2022-063141] [Medline: 36731936]
57. Azzolini E, Levi R, Sarti R, Pozzi C, Mollura M, Mantovani A, et al. Association between BNT162b2 vaccination and long COVID after infections not requiring hospitalization in health care workers. JAMA 2022 Aug 16;328(7):676-678 [FREE Full text] [doi: 10.1001/jama.2022.11691] [Medline: 35796131]
58. Ayoubkhani D, Bermingham C, Pouwels KB, Glickman M, Nafilyan V, Zaccardi F, et al. Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study. BMJ 2022 May 18;377:e069676 [FREE Full text] [doi: 10.1136/bmj-2021-069676] [Medline: 35584816]
59. Zhang X, Guo Y, Zhou Q, Tan Z, Cao J. The mediating roles of medical mistrust, knowledge, confidence and complacency of vaccines in the pathways from conspiracy beliefs to vaccine hesitancy. Vaccines 2021 Nov 17;9(11):1342 [FREE Full text] [doi: 10.3390/vaccines9111342] [Medline: 34835273]
60. Larson HJ, Cooper LZ, Eskola J, Katz SL, Ratzan S. Addressing the vaccine confidence gap. Lancet 2011 Aug 06;378(9790):526-535. [doi: 10.1016/S0140-6736(11)60678-8] [Medline: 21664679]
61. Sheikh N, Touseef M, Sultan R, Cheema K, Cheema S, Sarwar A. Understanding COVID-19 vaccine hesitancy in Pakistan: the paradigm of confidence, convenience and complacency; a cross-sectional study. medRxiv. 2021. URL: https://www.medrxiv.org/content/10.1101/2021.07.12.21260409v1 [accessed 2023-04-18]
62. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, et al. Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. J Med Internet Res 2021 Jun 11;23(6):e27632 [FREE Full text] [doi: 10.2196/27632] [Medline: 34061757]
63. Al-Marshoudi S, Al-Balushi H, Al-Wahaibi A, Al-Khalili S, Al-Maani A, Al-Farsi N, et al. Knowledge, attitudes, and practices (KAP) toward the COVID-19 vaccine in Oman: a pre-campaign cross-sectional study. Vaccines 2021 Jun 04;9(6):602 [FREE Full text] [doi: 10.3390/vaccines9060602] [Medline: 34200001]
64. Troiano G, Nardi A. Vaccine hesitancy in the era of COVID-19. Public Health 2021 May;194:245-251 [FREE Full text] [doi: 10.1016/j.puhe.2021.02.025] [Medline: 33965796]
65. Finney Rutten LJ, Zhu X, Leppin AL, Ridgeway JL, Swift MD, Griffin JM, et al. Evidence-based strategies for clinical organizations to address COVID-19 vaccine hesitancy. Mayo Clin Proc 2021 Mar;96(3):699-707 [FREE Full text] [doi: 10.1016/j.mayocp.2020.12.024] [Medline: 33673921]

66. Ashwell D, Cullinane J, Croucher SM. COVID-19 vaccine hesitancy and patient self-advocacy: a statistical analysis of those who can and can't get vaccinated. BMC Public Health 2022 Jul 05;22(1):1296 [FREE Full text] [doi: 10.1186/s12889-022-13661-4] [Medline: 35790942]
67. Tostrud L, Thelen J, Palatnik A. Models of determinants of COVID-19 vaccine hesitancy in non-pregnant and pregnant population: review of current literature". Hum Vaccin Immunother 2022 Nov 30;18(6):2138047 [FREE Full text] [doi: 10.1080/21645515.2022.2138047] [Medline: 36345571]
68. González-Block M, Pelcastre-Villafuerte BE, Riva Knauth D, Fachel-Leal A, Comes Y, Crocco P, et al. Influenza vaccination hesitancy in large urban centers in South America. Qualitative analysis of confidence, complacency and convenience across risk groups. PLoS One 2021;16(8):e0256040 [FREE Full text] [doi: 10.1371/journal.pone.0256040] [Medline: 34383834]
69. Huang Y, Chen C, Wang L, Wu H, Chen T, Zhang L. HPV vaccine hesitancy and influencing factors among university students in China: a cross-sectional survey based on the 3Cs model. Int J Environ Res Public Health 2022 Oct 28;19(21):14025 [FREE Full text] [doi: 10.3390/ijerph192114025] [Medline: 36360905]
70. Lazarus JV, Ratzan SC, Palayew A, Gostin LO, Larson HJ, Rabin K, et al. A global survey of potential acceptance of a COVID-19 vaccine. Nat Med 2021 Feb;27(2):225-228 [FREE Full text] [doi: 10.1038/s41591-020-1124-9] [Medline: 33082575]
71. Servidio R, Malvaso A, Vizza D, Valente M, Campagna MR, Iacono ML, et al. The intention to get COVID-19 vaccine and vaccine uptake among cancer patients: an extension of the theory of planned behaviour (TPB). Support Care Cancer 2022 Oct;30(10):7973-7982 [FREE Full text] [doi: 10.1007/s00520-022-07238-5] [Medline: 35752690]
72. Vaccination and trust: how concerns arise and the role of communication in mitigating crisis. World Health Organization. URL: https://www.who.int/europe/publications/i/item/WHO-EURO-2017-2908-42666-59448 [accessed 2023-04-18]

Abbreviations

AVE: average variance extracted
CFI: comparative fit index
CR: composite reliability
OR: odds ratio
REDCap: Research Electronic Data Capture
RMSEA: root mean square error of approximation
TLI: Tucker Lewis index
WHO: World Health Organization
3Cs: complacency, convenience, confidence

Edited by T Sanchez, A Mavragani; submitted 26.09.22; peer-reviewed by A Ramaprasad, O Ham; comments to author 15.03.23; revised version received 25.04.23; accepted 26.05.23; published 03.07.23.

Please cite as:
Huang Y, Zhang L, Fu J, Wu Y, Wang H, Xiao W, Xin Y, Dai Z, Si M, Chen X, Jia M, Leng Z, Cui D, Su X
COVID-19 Vaccine Hesitancy Among Patients Recovered From COVID-19 Infection in Wuhan, China: Cross-Sectional Questionnaire Study
JMIR Public Health Surveill 2023;9:e42958
URL: https://publichealth.jmir.org/2023/1/e42958
doi:10.2196/42958
PMID:37247615

Original Paper

The Impact of Priority Settings at the Start of COVID-19 Mass Vaccination on Subsequent Vaccine Uptake in Japan: One-Year Prospective Cohort Study

Daisuke Hori[1], MD, PhD; Tsukasa Takahashi[1], MD, PhD; Akihiko Ozaki[2], MD, DrPH; Takahiro Tabuchi[3], MD, PhD

[1]Institute of Medicine, University of Tsukuba, Tsukuba, Japan

[2]Department of Breast and Thyroid Surgery, Jyoban Hospital of Tokiwa Foundation, Iwaki, Japan

[3]Cancer Control Center, Osaka International Cancer Institute, Osaka, Japan

Corresponding Author:
Daisuke Hori, MD, PhD
Institute of Medicine
University of Tsukuba
1-1-1 Tennodai
Tsukuba, 305-8575
Japan
Phone: 81 29 853 6025
Email: daisuke_hori@md.tsukuba.ac.jp

Abstract

Background: Distributing COVID-19 vaccines to the public was an important task for the governments of each country. Because of various limitations, priority settings for vaccination were determined at the time of mass vaccination. However, trends between vaccine intention and uptake, as well as reasons for getting vaccinated or not getting vaccinated, among these groups were understudied, undermining verification of the legitimacy of priority selection.

Objective: This study aims to illustrate a trend from prior COVID-19 vaccine intention, when the vaccine was not available, to the actual uptake within 1 year when all residents had access to the vaccine, to illustrate a change of reason for getting vaccinated or not getting vaccinated and to examine whether priority settings predicted subsequent vaccination uptake.

Methods: Prospective cohort, web-based, self-administered surveys were conducted in Japan at 3 time points: February 2021, September to October 2021, and February 2022. In total, 13,555 participants (age: mean 53.1, SD 15.9 years) provided valid responses, with a 52.1% follow-up rate. On the basis of the information obtained in February 2021, we identified 3 types of priority groups: health care workers (n=831), people aged ≥65 years (n=4048), and those aged 18 to 64 years with underlying medical conditions (n=1659). The remaining patients were treated as nonpriority (n=7017). Modified Poisson regression analysis with a robust error estimated the risk ratio for COVID-19 vaccine uptake after adjusting for socioeconomic background, health-seeking behavior, attitude toward vaccines, and COVID-19 infection history.

Results: In February 2021, a total of 5182 out of 13,555 (38.23%) respondents expressed their intention to get vaccinated. In February 2022, a total of 1570 out of 13,555 (11.6%) respondents completed the third dose and 10,589 (78.1%) respondents completed the second dose. Prior vaccine intention and subsequent vaccine coverage rates were higher in the priority groups. Protection of themselves and their families from potential infection was the most frequent reason for getting vaccinated, whereas concern about side effects was the most frequent reason for hesitation across the groups. Risk ratios for received, reserved, or intended for vaccination in February 2022 were 1.05 (95% CI 1.03-1.07) for the health care worker group, 1.02 (95% CI 1.005-1.03) for the older adult group, and 1.01 (95% CI 0.999-1.03) for the preexisting conditions group compared with the nonpriority group. Prior vaccine intention and confidence in vaccines were strong predictors of vaccine uptake.

Conclusions: The priority settings at the start of the COVID-19 vaccination program had a significant impact on vaccine coverage after 1 year. The priority group for vaccination achieved higher vaccination coverage in February 2022. There was room for improvement among the nonpriority group. The findings of this study are essential for policy makers in Japan and other countries to develop effective vaccination strategies for future pandemics.

(JMIR Public Health Surveill 2023;9:e42143) doi:10.2196/42143

KEYWORDS

cohort studies; SARS-CoV-2; COVID-19; Japan; vaccination; vaccination hesitancy; vaccines

Introduction

Background

Since 2020, the wave of COVID-19 infection caused by SARS-COV-2 has affected populations intermittently around the world. To address this global challenge, COVID-19 vaccines were quickly developed, tested, licensed, and delivered worldwide in a relatively short time frame [1]. Vaccines are the most effective way to reduce the risk of COVID-19 infection, to bring herd immunity, and eventually to end the pandemic [2-4]. Being vaccinated contributes not only to preventing infection but also to reducing the severity, hospitalization, and mortality rates of COVID-19 [5-8]. In addition to the direct benefits to vaccine recipients, there are indirect benefits to those who are protected from the spread of COVID-19 infection. At the initial stage of vaccination, due to limitations in supply, logistics, and workforce, it was impossible to provide an equal opportunity for vaccination to all citizens at the same time. Prioritizing health care workers (HCWs) and people with medical vulnerabilities associated with poorer COVID-19 outcomes, such as older adults or those with chronic diseases, was common practice across countries [9-13]. Certainly, it was necessary to prioritize these populations when new vaccines were introduced to maintain the health care system and social function. However, at the same time, this practice might create artificial health disparities between priority and nonpriority groups. Attention should be paid to the manner in which the setting of priorities has affected the subsequent vaccine uptake situation.

The following paragraph describes an overview of the COVID-19 vaccine situation in Japan. Regarding countermeasures against COVID-19, there were no strict restrictions that involved penalties, such as those imposed during the lockdown in the United Kingdom. Instead, the government declared a state of emergency and called for citizens to refrain from actions that could spread the infection. It should also be noted that Japan has a universal health care system. COVID-19 vaccination was free of charge and was not mandatory. On February 14, the Pfizer vaccine was approved in Japan. The first dose of vaccination began on February 17, 2021, targeting HCWs in the National Hospital Organization, particularly those on the frontlines of COVID-19 patient care. Nationwide vaccination for HCWs began in March 2021. Mass vaccination for the following groups started in April: those aged ≥65 years at the end of March, those aged 18 to 64 years with preexisting conditions, and those who were working in nursing homes [14]. The government approved the Takeda or Moderna vaccine in May and the AstraZeneca vaccine in August. Although Japan started mass vaccination several months later than Western countries, vaccination coverage grew steadily once it began [15,16]. The third dose of vaccination for HCWs started in December 2021. As of April 2022, Japan has depended on imported vaccines supplied by Pfizer, Takeda or Moderna, and AstraZeneca that require 2-dose shots. By April 2022, more than half of the total population had received a booster dose, although vaccination rates among those in their 20s and 30s were reported to be low [17]. Those aged ≥60 years and those with underlying conditions were given the opportunity for a fourth vaccination dose, starting in May 2022, before HCWs.

Researchers have investigated the sociodemographic, psychological, behavioral, and social factors associated with vaccine intention or uptake status. In Japan, the following factors were related to vaccine intention or uptake: age [18-25], sex [18,20-25], employment status [18,22], marital status [18,19], income [18-20,22,23], educational attainment [18,19,22], underlying medical conditions [18-21,25], influenza vaccine history [19,24-26], medical checkup history [26], and COVID-19–related fear or anxiety [18-20,25,26]. Regarding attitudes toward the COVID-19 vaccine, concerns about the side effects of vaccines [25,26], confidence in vaccine effectiveness and public authorities [18,20,25], and pseudoscience [26] were linked to vaccine intention. Concern about adverse reactions is one of the most common reasons for vaccine hesitancy [21,22,25]. Although most of the existing literature relies on cross-sectional data, longitudinal studies on vaccine intention and subsequent uptake status are emerging worldwide [27-32]. These longitudinal studies demonstrate that the shift from initial intention to COVID-19 vaccination to actual uptake was influenced by the information about vaccine safety and efficacy that became available to the public. However, to date, no longitudinal studies have simultaneously observed the shift between those who are eligible for priority vaccination and those who are not. This is mainly because of the difficulty in recruiting a sufficient number of respondents in cohort studies.

Objectives

Two series of web-based, prospective cohort studies are ongoing: the Japan Society and New Tobacco Internet Survey (JASTIS) and the Japan COVID-19 and Society Internet Survey (JACSIS). We derived 1-year longitudinal data from 3 time points of these surveys to address the following objectives: (1) to illustrate a trend from prior COVID-19 vaccine intention, when the vaccine was not available, to the actual uptake within 1 year when all residents had access to the vaccine; (2) to illustrate a change of reason for being vaccinated or not being vaccinated; and (3) to examine whether priority settings predicted vaccination uptake. The results are presented focusing on 3 types of vaccination priority and nonpriority groups at the rollout of mass vaccination. This approach allowed us to identify important trends over time and to examine whether the setting of priority groups was reasonable. Our study will also provide profound information for governments, health care policy makers, and health care providers to determine the target population for public health messages intended to promote vaccine uptake and how to tailor these messages for their specific audience.

Methods

Study Setting and Design

The data were derived from JASTIS from 2015 onward and JACSIS from 2020 onward. The study profiles have been described elsewhere [18,33]. Briefly, they comprised a series of web-based, self-administered cohort surveys that shared the same survey panel. The survey panel was drawn from approximately 2.2 million panelists at a Japanese internet research company (Rakuten Insight, Inc), which comprises individuals from diverse socioeconomic backgrounds (eg, marital status, educational level, and household income) to be a nationally representative sample of Japan. In JASTIS and JACSIS, sampling was stratified by sex, age, and prefecture of residence to be as nationally representative as possible. We used 3 time points to analyze the longitudinal data from the surveys conducted from February 8 to 25, 2021 (T1); September 27 to October 29, 2021 (T2); and February 1 to 28, 2022 (T3). All the respondents received a nominal incentive for survey completion at each time point.

Study Participants

At T1, an email containing a link to the survey website was distributed to 32,827 candidates who participated in the same survey series conducted between 2015 and 2020. In total, 73.29% (24,059/32,827) of the participants answered the questionnaire. The survey continued to gather new panelists from the same internet research company until the number of respondents reached the targeted sample size (n=26,000). The respondents who answered that they were HCWs were classified as HCW group. The remaining respondents were classified into 3 groups: those who were aged ≥65 years, those who were aged 18 to 64 years with at least one preexisting condition, and those who were aged 18 to 64 years with no preexisting conditions. Preexisting conditions were defined as the following 14 conditions: hospital attendance for hypertension, diabetes mellitus, asthma, pneumonia or bronchitis, angina pectoris, myocardial infarction, stroke (including cerebral infarction and cerebral hemorrhage), chronic obstructive pulmonary disease, chronic kidney disease, chronic liver disease (excluding fatty liver and hepatitis), immune disorders and other diseases that cause immune deficiency (including steroid use), or cancer (including malignant tumor); sleep apnea identified by a medical checkup or physician; and a BMI ≥30 kg/m^2. The exclusion criteria are summarized in Multimedia Appendix 1.

Ethics Approval, Informed Consent, and Participation

Ethics approval for this study was obtained from the ethics committee of the Osaka International Cancer Institute (#20084-6) and the ethics committee of the University of Tsukuba, Institute of Medicine (#1737). All procedures were performed in accordance with the ethical standards of the institutional or national research committees and the 1964 Declaration of Helsinki and its later amendments, or comparable ethical standards. All participants provided web-based informed consent at registration. The study data were deidentified. For the research incentives, the participants earned points that could be used for shopping but the amount was not disclosed.

Measures

Assessment of COVID-19 Vaccine Intention, Uptake, and Reasons for a Decision

In each survey, COVID-19 vaccine intention and uptake were assessed using a single-item question. The details are described in Multimedia Appendix 2. According to the response at T3, the participants were dichotomized into either a hesitant group or a received, reserved, or intended group. The hesitant group included participants who expressed that they did not want to get vaccinated or they would like to get vaccinated after waiting to see how it goes. The received, reserved, or intended group included participants who answered that they wanted to get vaccinated; had reserved a vaccination; or had received the first, second, and third (booster) dose of vaccine.

Each survey asked the received, reserved, or intended group for their reasons for getting vaccinated against COVID-19. The following sentence was only asked at T1: “It was recommended by SNS or the media.” The following sentences were only asked at T2 and T3: “It was recommended by the media,” “It was recommended by SNS,” and “It was recommended by a health care worker.” Each survey also asked the hesitant group about reasons for not getting vaccinated. The following sentences were only asked at T2 and T3: “Long-term side effects are unknown,” “I’m worried about short term side effects,” “I’m worried about death by the vaccines,” and “I don’t trust the vaccines’ approval process.” Multiple answers were allowed; therefore, the sum of the proportions for each reason did not necessarily add up to 100%.

Sociodemographic and Behavioral Factors and Vaccine-Related Scales

The following variables were derived from the T1 survey: age group, preexisting conditions, employment status, sex, marital status, educational background, occupation, annual household income, and influenza vaccination in the 2019/2020 season. The following variables were derived from the T2 survey: regular medical checkups before the COVID-19 pandemic, the vaccine hesitancy scale, the vaccine conspiracy beliefs scale, and the fear of COVID-19 scale. COVID-19 infection history and vaccine supplier details were derived from the T3 survey.

The vaccine hesitancy scale modified version [34,35] was used, which comprises 9 items, such as “Vaccines are important for my health,” scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). One item, “I do not need vaccines for diseases that are not that common anymore,” was omitted because COVID-19 was not convergent. The average scores for 2 subscales were calculated: 2 items for aversion to the risk of vaccination and 7 items for lack of confidence in the vaccine [36]. A higher average score, each ranging from 1 to 5, indicates a higher level of lack of confidence or risks. The Cronbach α was .62 for aversion to risk and .93 for lack of confidence.

The vaccine conspiracy beliefs scale was used, which comprises 7 items, such as “Vaccine safety data is often fabricated,” scored on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree) [35]. A higher average score, ranging from 1 to 7, indicates a higher level of conspiracy beliefs against the vaccine. The Cronbach α was .95.

The fear of COVID-19 scale was used, which comprises 7 items, such as “I am very afraid of coronavirus-19,” scored on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree) [37,38]. A higher average score, ranging from 1 to 5, indicates a higher level of fear of COVID-19. The Cronbach α was .84.

The score for the aversion to the risk of vaccination was classified into 2 groups by the median: “low” defined as a score lower than the median and “high” defined as a score equal to or higher than the median. Scores for the lack of confidence in the vaccine, conspiracy beliefs about the vaccine, and fear of COVID-19 were categorized into 3 groups by the tertile scores: “low” defined as a score below the lower tertile, “middle” defined as a score between the lower and higher tertiles, and “high” defined as a score higher than the higher tertile.

The participants were asked if they had been diagnosed with COVID-19 infection. Those who answered “Yes, within the last year” or “Yes, over a year ago” were defined as having a COVID-19 infection history. Participants were also asked if they had been hospitalized due to COVID-19.

Those who had received the vaccination at least once were asked which company supplied the vaccines they received. The possible answers were Pfizer, Takeda or Moderna, AstraZeneca, Johnson & Johnson, others, and “I don’t know.”

Statistical Analysis

To identify factors associated with COVID-19 vaccine uptake at T3, we calculated the risk ratios (RRs) and 95% CIs using a modified Poisson regression analysis with a robust error variance [39]. The objective variable was vaccine uptake; those who had received vaccination, reserved vaccination, or intended to receive vaccination were coded as 1, and the hesitant individuals were coded as 0. In model 1, unadjusted RRs were calculated for the priority setting. In model 2, the following variables assessed at T1 were used for adjustment: prior COVID-19 vaccination intention, age group, employment status, number of preexisting conditions, sex, marital status, educational background, annual household income, and influenza vaccination in the 2019/2020 season. Model 3 was additionally adjusted for the following variables assessed at T2: regular medical checkups before the pandemic, aversion to the risk of vaccination, lack of confidence in the vaccine, conspiracy beliefs about the vaccine, and fear of COVID-19. Model 4 was additionally adjusted for COVID-19 infection history at T3. A sensitivity analysis was conducted, including age group, employment status, and number of preexisting conditions instead of vaccination priority settings.

Before conducting the modified Poisson regression analysis, the multiple regression analysis confirmed that there was no multicollinearity among the explanatory variables. The 2-sided significance level was set at 5%, and SPSS Statistics (version 28; IBM Corp) was used for all the analyses.

Results

Eligible Participants for Analysis

The flowchart of selecting eligible participants for analysis is illustrated in Multimedia Appendix 3. After excluding participants who matched the exclusion criteria or were lost to follow-up at each time point, a total of 13,555 individuals (age: mean 53.1, SD 15.9 years; 6519/13,555, 48.09% female) remained for analysis. They were older and had more male individuals than the excluded 12,445 individuals (age: mean 46, SD 16.9 years; 6231/12,445, 50.1% female). The overall valid follow-up rate was 52.1%. A total of 6538 individuals were eligible for each of the 3 types of vaccination priority: 1659 (25.37%) aged <65 years with underlying medical conditions (age: mean 51.5, SD 10.2 years; 540/1659, 32.55% female), 4048 (61.91%) aged ≥65 years (age: mean 71.3, SD 4.1 years; 1965/4048, 48.54% female), and 831 (12.71%) HCWs (age: mean 44.7, SD 12.6 years; 379/831, 45.6% female). A total of 7017 individuals did not meet any of the requirements for vaccination priority (age: mean 44, SD 12.4 years; 3635/7017, 51.80% female).

Table 1 shows the characteristics of the participants according to the priority settings at T1. At T3, the proportion of the hesitant group was highest in the nonpriority group (13.3%) and lowest in the priority group (4.5%) of older adults. Table S1 in Multimedia Appendix 4 summarizes the breakdown of underlying medical conditions among the priority group of those with preexisting conditions. Hypertension (n=779) and sleep apnea (n=409) were the most common preexisting conditions. The proportion of the received, reserved, or intended group ranged from 86.5% among patients with angina pectoris to 98.7% among patients with immune deficiency. Table S2 in Multimedia Appendix 4 shows the breakdown of jobs among the priority group of HCWs. There were 110 nurses, 63 pharmacists, and 55 physicians. Midwives (n=7) exhibited relatively lower vaccine coverage of 71.4%. Tables S1 to S5 in Multimedia Appendix 5 summarize the COVID-19 vaccine coverage rate stratified by COVID-19 infection history or hospitalization over 1 year. Higher RRs indicated that those who had been infected with COVID-19 in the past were more likely to receive vaccines than those who had not been infected. Across all 4 groups, the RRs ranged from 0.90 to 1.10.

Table 2 demonstrates the kind of COVID-19 vaccines that were inoculated to the respondents by T3. Most people had been vaccinated with Pfizer vaccines.

Table 1. Characteristics of the respondents stratified by the COVID-19 vaccination priority settings at T1 (February 2021; the percentages in parentheses in each category vertically add up to 100).

Characteristics	Total (N=13,555), n (%)	Nonpriority (n=7017), n (%)	Aged 18-64 years with preexisting conditions, non-HCW[a] (n=1659), n (%)	Aged ≥65 years, non-HCW (n=4048), n (%)	HCW (n=831), n (%)
Prior COVID-19 vaccination intention[b]					
Intended	5182 (38.23)	1966 (28.02)	683 (41.17)	2130 (52.62)	403 (48.5)
Wait and see	6936 (51.17)	4133 (58.9)	829 (49.97)	1644 (40.61)	330 (39.71)
Refused	1437 (10.6)	918 (13.08)	147 (8.86)	274 (6.77)	98 (11.79)
Age group (years)[b]					
18-34	2035 (15.01)	1713 (24.41)	119 (7.17)	0 (0)	203 (24.42)
35-44	2052 (15.14)	1601 (22.82)	234 (14.1)	0 (0)	217 (26.11)
45-54	2774 (20.24)	2029 (28.91)	529 (31.89)	0 (0)	216 (25.99)
55-64	2600 (19.18)	1674 (23.86)	777 (46.83)	0 (0)	149 (17.93)
65-80	4094 (30.2)	0 (0)	0 (0)	4048 (100)	46 (5.53)
Employment status[b]					
Employed, non-HCW	7209 (53.18)	4997 (71.21)	1221 (73.6)	991 (24.48)	0 (0)
Employed, HCW	831 (6.13)	0 (0)	0 (0)	0 (0)	831 (100)
Unemployed	2034 (15)	488 (6.95)	187 (11.27)	1359 (33.57)	0 (0)
Not working (student, homemaker, or retiree)	3481 (25.68)	1532 (21.83)	251 (15.13)	1698 (41.95)	0 (0)
Number of preexisting conditions[b]					
0	9860 (72.74)	7017 (100)	0 (0)	2177 (53.78)	666 (80.14)
>1	3695 (27.26)	0 (0)	1659 (100)	1871 (46.22)	165 (19.86)
Sex[b]					
Male	7036 (51.91)	3382 (48.2)	1119 (67.45)	2083 (51.46)	452 (54.39)
Female	6519 (48.09)	3635 (51.8)	540 (32.55)	1965 (48.54)	379 (45.61)
Marital status[b]					
Single, divorced, or widowed	4945 (36.48)	3133 (44.65)	606 (36.53)	886 (21.89)	320 (38.51)
Married	8610 (63.52)	3884 (55.35)	1053 (63.47)	3162 (78.11)	511 (61.49)
Educational background[b]					
Others	6811 (50.25)	3324 (47.37)	829 (49.97)	2290 (56.57)	368 (44.28)
4-year college, university, or graduate	6744 (49.75)	3693 (52.63)	830 (50.03)	1758 (43.43)	463 (55.72)
Household income (million; yen)[b]					
<5	5530 (40.8)	2377 (33.87)	576 (34.72)	2347 (57.98)	230 (27.68)
5-10	4048 (29.86)	2369 (33.76)	571 (34.42)	779 (19.24)	329 (39.59)
≥10	1383 (10.2)	832 (11.86)	226 (13.62)	186 (4.59)	139 (16.73)
I do not know or prefer not to answer.	2594 (19.14)	1439 (20.51)	286 (17.24)	736 (18.18)	133 (16)
Influenza vaccination in the 2019/2020 season[b]					
No	8246 (60.83)	4762 (67.86)	1018 (61.36)	2130 (52.62)	336 (40.43)
Yes	5309 (39.17)	2255 (32.14)	641 (38.64)	1918 (47.38)	495 (59.57)
Regular medical checkups before the pandemic[c]					
No	4612 (34.02)	2711 (38.63)	487 (29.36)	1157 (28.58)	257 (30.93)

Characteristics	Total (N=13,555), n (%)	Nonpriority (n=7017), n (%)	Aged 18-64 years with preexisting conditions, non-HCW[a] (n=1659), n (%)	Aged ≥65 years, non-HCW (n=4048), n (%)	HCW (n=831), n (%)
Yes	8943 (65.97)	4306 (61.36)	1172 (70.64)	2891 (71.42)	574 (69.07)
Aversion to the risk of vaccination[c]					
Low, 1.00-3.00	6587 (48.59)	2984 (42.52)	801 (48.28)	2370 (58.55)	432 (51.99)
High, 3.01-5.00	6968 (51.4)	4033 (57.47)	858 (51.72)	1678 (41.45)	399 (48.01)
Lack of confidence in the vaccine[c]					
Low, 1.00-1.99	4080 (30.1)	1660 (23.66)	507 (30.56)	1666 (41.16)	247 (29.72)
Middle, 2.00-2.43	5458 (40.26)	2791 (39.77)	644 (38.82)	1735 (42.86)	288 (34.66)
High, 2.44-5.00	4017 (29.63)	2566 (36.57)	508 (30.62)	647 (15.98)	296 (35.62)
Conspiracy beliefs about the vaccine[c]					
Low, 1.00-3.27	4253 (31.37)	1810 (25.79)	504 (30.38)	1673 (41.33)	266 (32.01)
Middle, 3.28-4.00	6399 (47.21)	3566 (50.82)	784 (47.26)	1667 (41.18)	382 (45.97)
High, 4.01-5.00	2903 (21.42)	1641 (23.39)	371 (22.36)	708 (17.49)	183 (22.02)
Fear of COVID-19[c]					
Low, 1.00-2.27	3756 (27.71)	2135 (30.43)	398 (23.99)	973 (24.04)	250 (30.08)
Middle, 2.28-3.00	6638 (48.97)	3350 (47.74)	824 (49.67)	2082 (51.43)	382 (45.97)
High, 3.01-7.00	3161 (23.32)	1532 (21.83)	437 (26.34)	993 (24.53)	199 (23.95)
COVID-19 infection history[d]					
None	13,264 (97.85)	6848 (97.59)	1607 (96.87)	4018 (99.26)	791 (95.19)
Yes	291 (2.15)	169 (2.41)	52 (3.13)	30 (0.74)	40 (4.81)
Vaccine uptake status[d]					
Third dose vaccinated	1570 (11.58)	225 (3.21)	71 (4.28)	954 (23.57)	320 (38.51)
Second dose vaccinated	10,589 (78.12)	5804 (82.71)	1430 (86.2)	2906 (71.79)	449 (54.03)
First dose vaccinated	57 (0.42)	41 (0.58)	7 (0.42)	5 (0.12)	4 (0.48)
Reserved or intended	21 (0.15)	15 (0.21)	4 (0.24)	0 (0)	2 (0.24)
Wait and saw	443 (3.27)	334 (4.76)	42 (2.53)	50 (1.23)	17 (2.05)
Refused	875 (6.45)	598 (8.52)	105 (6.33)	133 (3.28)	39 (4.69)
Dichotomized vaccine uptake status[d]					
Received, reserved, or intended	12,237 (90.28)	6085 (86.72)	1512 (91.14)	3865 (95.48)	775 (93.26)
Hesitant	1318 (9.72)	932 (13.28)	147 (8.86)	183 (4.52)	56 (6.74)

[a] HCW: health care worker.

[b] The variables were derived from T1 (February 2021).

[c] The variables were derived from T2 (September to October 2021).

[d] The variables were derived from T3 (February 2022).

Table 2. Summary of the COVID-19 vaccine suppliers that the respondents received by T3 (February 2022)[a].

COVID-19 vaccine supplier	Total (n=12,216), n (%)	Nonpriority (n=6070), n (%)	18-64 years with preexisting conditions, non-HCW[b] (n=1508), n (%)	≥65 years, non-HCW (n=3865), n (%)	HCW (n=773), n (%)
Pfizer	9421 (77.12)	4146 (68.3)	1144 (75.86)	3482 (90.09)	649 (83.96)
Takeda or Moderna	2860 (23.41)	1834 (30.21)	359 (23.81)	545 (14.1)	122 (15.78)
AstraZeneca	47 (0.38)	24 (0.39)	7 (0.46)	8 (0.21)	8 (1.03)
Others	28 (0.23)	16 (0.26)	4 (0.26)	5 (0.13)	3 (0.38)
I don’t know	218 (1.78)	126 (2.07)	25 (1.66)	53 (1.37)	14 (1.81)

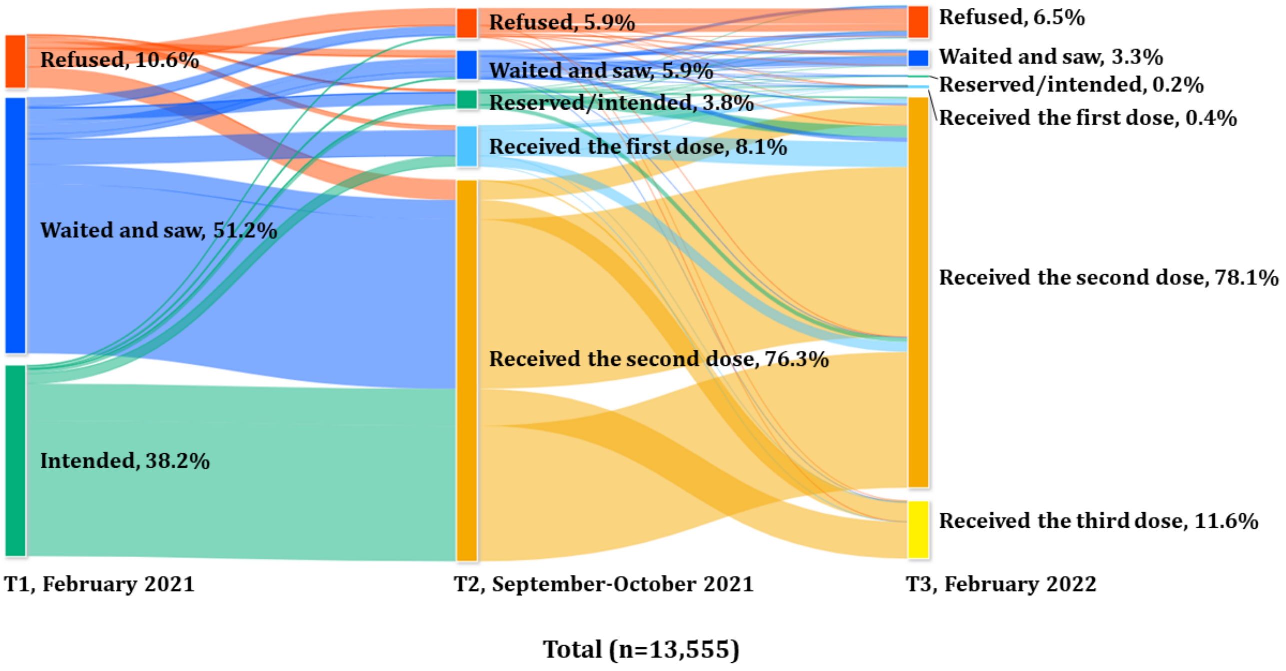
[a] Only those who had received at least 1 dose of the COVID-19 vaccine were asked to respond. Multiple answers were allowed.

[b] HCW: health care worker.

Shift From Prior COVID-19 Vaccine Intention to Actual Vaccine Uptake 1 Year Later

As illustrated in Figure 1, at T1, a total of 6936 out of 13,555 (51.17%) respondents answered, “I would like to get vaccinated after waiting to see how it goes.” At T2, a total of 10,339 out of 13,555 (76.27%) respondents had received 2 vaccination shots. At T3, a total of 10,589 out of 13,555 (78.12%) respondents completed 2 shots and 1570 (11.58%) respondents completed a third (booster) shot of vaccination. Figures S1 to S4 in Multimedia Appendix 6 showed the results stratified by the nonpriority group and the 3 priority groups. Prior intention to get vaccinated was lowest among the nonpriority group and highest in the older adult group. At T3, the percentage of those who had received the second dose/third dose of vaccine was 82.7%/3.2% in the nonpriority group, 86.2%/4.3% in the group with preexisting conditions, 71.8%/23.6% in the group aged ≥65 years, and 54%/38.5% in the HCW group.

Figure 1. Shift from prior COVID-19 vaccine intention in February 2021 to actual vaccine uptake in February 2022, including the nonpriority group and the 3 priority groups. ChartExpo™ was used to illustrate the Sunkey diagram.



Shift in the Reasons for Receiving or Not Receiving COVID-19 Vaccination

Figure 2 summarizes the shift in the reasons for getting vaccinated among those who received vaccination, who reserved vaccination, or who were intended for vaccination. The top reasons for getting vaccinated were worry about infecting themselves, their family, or others around them, followed by social norms. The percentage of those who expressed that they were worried about contracting COVID-19 gradually declined (63.6%, 51.4%, and 46.6%, respectively). Figures S1 to S4 in Multimedia Appendix 7 show the results for each priority group and the nonpriority group. There were intergroup characteristics in the reasons for getting vaccinated: the preexisting conditions and older adult groups were more likely to get vaccinated because they thought they had a high risk of contracting severe COVID-19 infection. The HCW group had the lowest number of reasons for getting vaccinated.

The reasons for not getting vaccinated among the hesitant group are shown in Figure 3. The most common reason for not getting vaccinated among the unvaccinated people was concerns about side effects at all 3 time points. The percentages were highest when the vaccination program had not yet begun at T1, declined at T2, and remained at the same level at T3 (81.7%, 51.9%, and 53.7%, respectively). Concerns about long-term side effects were raised more frequently than those about short-term side effects. The percentage who argued that the COVID-19 vaccine was ineffective gradually increased over the course of the year (9.2%, 21%, and 29.9%, respectively). Figures S1 to S4 in Multimedia Appendix 8 show the results for the nonpriority group and each priority group. At T3, the percentage of respondents who answered "I'm worried about side effects of the vaccines" was highest in the priority group with underlying medical conditions. The proportion of those who said "I don't trust the vaccines' approval process" was highest in the priority group of adults aged ≥65 years.

Figure 2. Shift in reason for getting vaccinated against COVID-19 from February 2021 to February 2022, including the nonpriority group and the 3 priority groups. Respondents were those who had intended or reserved to receive the vaccines or those who had received the vaccines at least once at each time point. The following sentences was asked only at T1: "It was recommended by SNS or the media." The following sentences were only asked at T2 and T3: "It was recommended by the media," "It was recommended by SNS," and "It was recommended by a health care worker.".

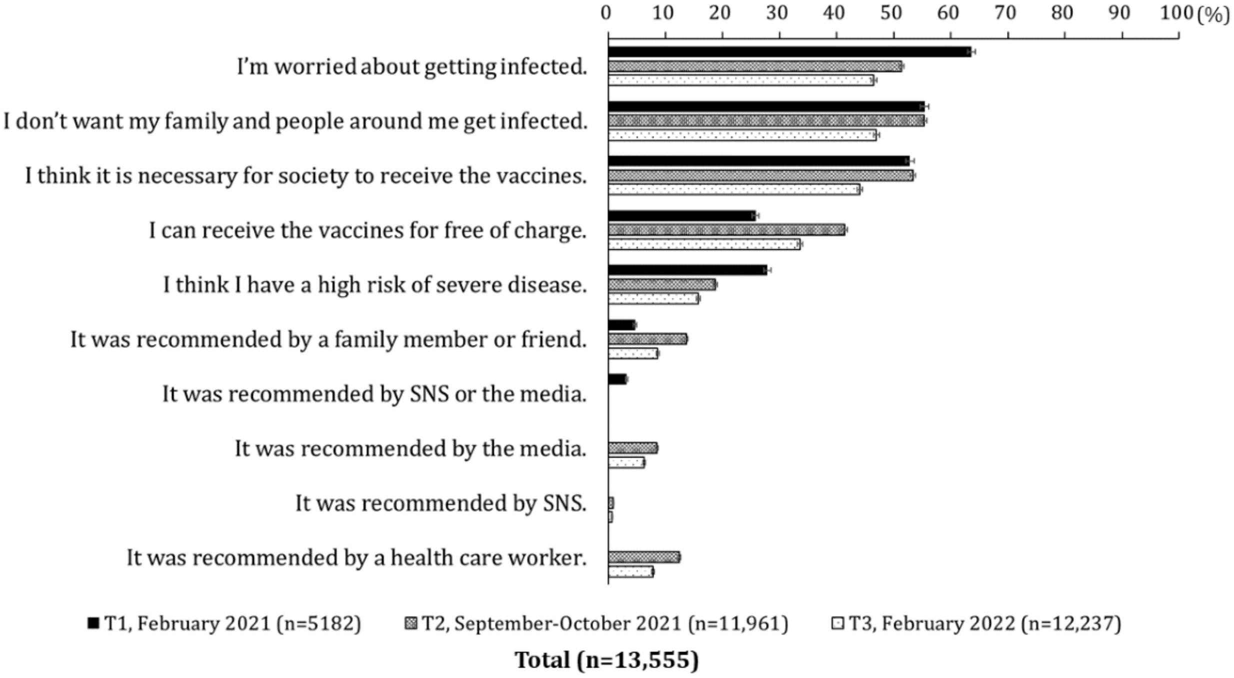
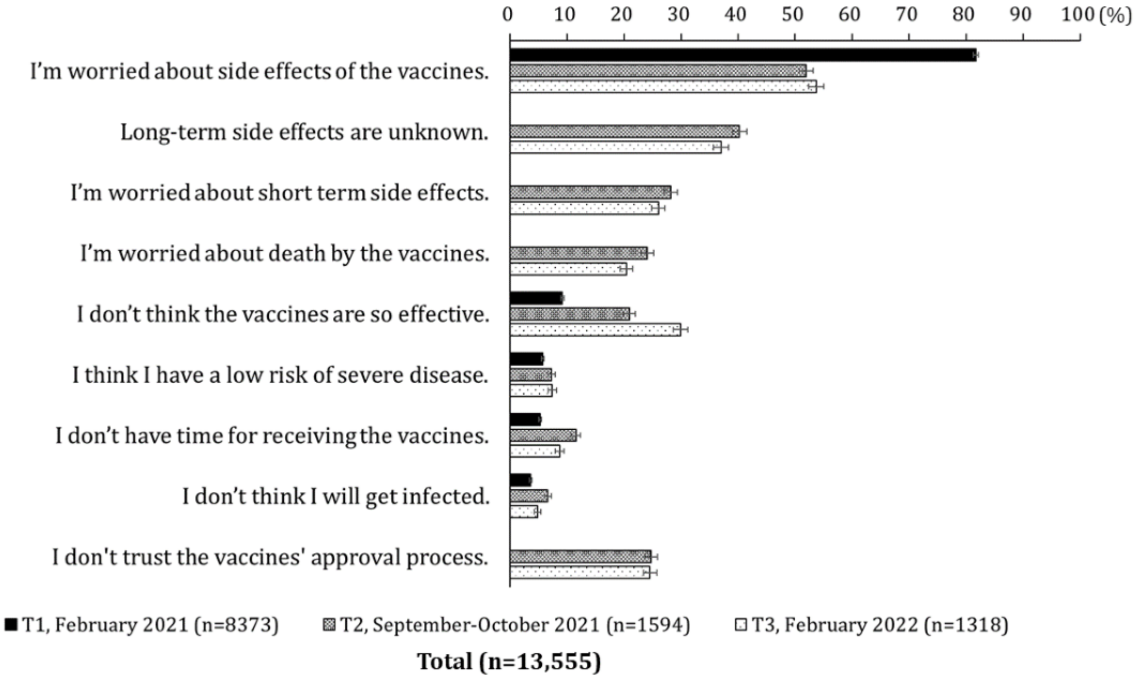


Figure 3. Shift in reason for not getting vaccinated against COVID-19 from February 2021 to February 2022, including the nonpriority group and the 3 priority groups. Respondents were those who chose to wait and see or refused to receive the vaccines at each time point.The following sentences was asked only at T2 and T3: “Long-term side effects are unknown,” “I’m worry about short term side effects,” “I’m worried about death by the vaccines,” and “I don’t trust the vaccines’ approval process.”.



Priority Settings at the Start of Mass Vaccination and Subsequent COVID-19 Vaccine Uptake 1 Year Later

Table 3 summarizes the results of the Poisson regression analyses. In model 4, significantly higher RRs for getting vaccinated were observed in the priority group of adults aged ≥65 years (RR 1.02, 95% CI 1.005-1.03) and HCWs (RR 1.05, 95% CI 1.03-1.07).

A sensitivity analysis was performed by replacing the priority group with age group, employment status, and number of preexisting conditions. The results are shown in Multimedia Appendix 9. In model 4, prior willingness for COVID-19 vaccination, a lower level of aversion to the risk of vaccination, a higher level of confidence in the vaccine, and a lower level of conspiracy beliefs about the vaccine exhibited higher RRs, which indicated a higher likelihood of getting immunization. Those who had the following characteristics were more likely to receive COVID-19 vaccination: being older, female, higher educational level, higher household income, having received influenza vaccination in the 2019/2020 season, and regular medical checkups before the pandemic. In contrast, statistically significant lower RRs were observed in those who were unemployed or not working and those who had a lower level of fear of COVID-19.

Table 3. Risk ratios (RRs) for COVID-19 vaccination uptake status at T3 (February 2022) by priority settings at T1 (February 2021).[a]

	Nonpriority, RR (95% CI)	18-64 years with preexisting conditions, non-HCW[b], RR (95% CI)	≥65 years, non-HCW, RR (95% CI)	HCW, RR (95% CI)
Model 1	Reference	1.05 (1.03-1.07)	1.10 (1.09-1.11)	1.08 (1.05-1.10)
Model 2	Reference	1.03 (1.01-1.04)	1.05 (1.04-1.07)	1.04 (1.02-1.06)
Model 3	Reference	1.02 (0.999-1.03)	1.02 (1.004-1.03)	1.05 (1.03-1.07)
Model 4	Reference	1.01 (0.999-1.03)	1.02 (1.005-1.03)	1.05 (1.03-1.07)

[a] Modified Poisson regression analysis with a robust error variance was performed. A higher RR indicated a higher likelihood of receiving, reserving, or intending to vaccinate at T3. Model 1 was unadjusted. Model 2 was adjusted for prior vaccination intention, sex, marital status, educational background, annual household income, and influenza vaccination during the 2019/2020 season. In addition to model 2, model 3 was adjusted for regular medical checkups before the pandemic, aversion to the risk of vaccination, lack of confidence in the vaccine, conspiracy beliefs about the vaccine, and fear of COVID-19. In addition to model 3, model 4 was adjusted for COVID-19 infection history.

[b] HCW: health care worker.

Discussion

Principal Findings

Prompt distribution of COVID-19 vaccines to the public was one of the global goals to be accomplished in 2021. In this study conducted in Japan, we reported on the shift from COVID-19 vaccine intention to actual uptake and the reasons for getting or refusing vaccination in a 1-year period. In February 2021, a total of 5182 out of 13,555 (38.23%) participants expressed that they intended to take the vaccine, 6936 (51.17%) chose to wait and see, and 1437 (10.6%) refused to get vaccinated. One year later, >90% (12,216/13,555) of the participants had received at least 1 dose of the vaccine. There were considerable differences in uptake across the nonpriority group and the 3 typical vaccine priority groups. We confirmed that COVID-19 vaccination uptake rates were higher among priority groups than the nonpriority group in line with recent reports [17]. This trend continued for 1 year after the start of mass vaccination. Overall, protection against infection for themselves, their families, and the people around them was cited as the main reason for seeking vaccination. In contrast, concern about side effects was cited as the main reason for avoiding vaccination. We also elucidated the factors that predicted, or were associated with, vaccine uptake. As hypothesized, prioritized groups at the initial stage of the COVID-19 mass vaccination program exhibited a higher vaccine uptake rate after adjusting for sociodemographic background, health-seeking behavior, and attitude toward the vaccine. To our knowledge, this is the first longitudinal study to illustrate these results simultaneously by priority groups and nonpriority groups with the advantage of a large cohort sample.

The vaccination coverage rate in our sample was higher than that reported in the official Japanese vaccination statistics, which stated that approximately 75% of the total population had received at least 1 dose by the end of February 2022 [40]. Therefore, caution should be exercised when interpreting and generalizing these results. Although stratified sampling was used to get as close to a nationally representative population as possible, the participants were registered monitors of an internet research company, and sampling bias was inevitable. For example, it is difficult to reach those who have no access to the internet, such as the institutionalized, poor, or homeless. In addition, the follow-up rate was slightly higher than 50%. It is also possible that the participants at the baseline died of COVID-19 and dropped out. Our current sample may comprise more older and health-oriented people than the actual population. These factors could lead to the overestimation of vaccination coverage. Nevertheless, a web-based survey targeting registered monitors was a feasible and frequently used research method during the COVID-19 pandemic. It should also be noted that this study focused on priority settings at the start of vaccination in the spring of 2021. Governments around the world made changes to their priority conditions as the vaccines became more widely distributed; for instance, lowering the age threshold for the older adult group [11]. The Japanese government did not set a priority for the third (booster) vaccination, which started in December 2021.

Overall, almost all of those who intended or chose to wait and see to receive the vaccine before the vaccine rollout actually received the vaccine after it became available. A large intention-to-action gap was observed in the past in relation to vaccination against H1N1 influenza [41,42]; however, the gap appears to be much smaller for COVID-19 vaccines [27,29,31]. Approximately half of those who had initially refused the vaccine continued to refuse, whereas the other half changed their intentions and accepted vaccination. The persistent refusers across the 3 time points were found more in the nonpriority group and the priority group with preexisting conditions. It should be noted that, among the hesitant individuals in these 2 groups, the percentage of those who refused vaccination increased between T2 and T3. Whether they remain unvaccinated or change their mind afterward remains to be observed. The highest vaccination coverage rates were found among the priority groups of older adults and HCWs. More than 90% of participants in these groups completed 2-dose vaccination within half a year. This tendency was also confirmed by the multivariable analysis. The older adult group was considered one of the most susceptible groups to severe COVID-19 infection, and participants in this group were the most eager to get vaccinated, even before the vaccination program began. This finding is in line with previous studies that have reported that older age groups were more willing to get vaccinated [18-26].

Concerns about side effects in general remained the most common reason for vaccine hesitancy or resistance within the year. This result is in line with previous studies in Japan [21,22,25] and in other countries [43-46]. We demonstrated that this trend continued from before the start of mass vaccination to 1 year later. We further illustrated that the number of participants who thought the COVID-19 vaccine was ineffective increased gradually in the vaccine hesitant group. This may be because the COVID-19 pandemic has not yet settled despite increased vaccination rates. This may also be due to outbreaks of the omicron variants, which were reported to be highly infectious and resistant to currently available vaccines [47,48]. There were some intergroup differences among the groups. In the older adults and preexisting conditions groups, more respondents cited fear of severe COVID-19 and awareness of the need for vaccination as their reason for getting vaccinated. Despite having the second highest coverage rate, the HCW group cited the fewest reasons for vaccination. This group might have benefited from the availability of vaccines at their workplace, although the current survey did not provide information on where they were vaccinated. Future studies should consider the convenience of time and location for vaccination, which can affect vaccine hesitancy [49,50].

There were slight differences in subsequent vaccine uptake between those who had a history of COVID-19 infection and those who had no history at T1. The same was true for a history of COVID-19 infection at T3. It should be noted that during the surveyed period, from February 2021 to February 2022, it is unclear from the current data which came first, the infection or vaccination. There would have been both prevaccination and breakthrough infections during the surveyed period. The current results do not support the existing literature in other countries

[31,51], which suggested that those who had previously been infected were less anxious about reinfection and had less interest in vaccination. However, vaccination after recovery from COVID-19 is reported to still be effective in preventing reinfection [52]. This inconsistency could be due to the small number of infected individuals in the current data set. Indeed, despite being a superaging society, Japan's COVID-19 mortality rate is among the lowest in the world [53,54]. In our sample, the highest percentage of COVID-19 infection history was found in the HCW group, who were considered to have the most frequent contact with patients with COVID-19 and also the most frequent infection testing. This finding underlines the need to prioritize vaccination for HCWs [55]. In contrast, the older adults group had the lowest percentage of COVID-19 infection history. Together with high vaccination coverage, it is fair to say that the vaccination campaign targeting the older adult population was successful. The nonpriority group had the largest number of infected cases, although the percentage was smaller than that of the HCW and preexisting conditions groups.

Those who were younger, were unemployed, were unmarried, had lower educational attainment, had lower income, had a presence of underlying medical conditions, had not had an influenza vaccination, and had not attended regular health checkups before the start of mass vaccination exhibited a lower vaccine uptake rate 1 year later. The current results corroborate a number of cross-sectional studies in Japan that have reported an association between these factors with vaccine hesitancy [18-26]. Our findings were consistent with a number of studies in other countries that have elucidated factors related to vaccine intention and uptake [56-58]. We have demonstrated the importance of conducting vaccine promotion campaigns for these populations. Prior favorable vaccine intention and more positive attitudes about the safety and effectiveness of the vaccine significantly improved the likelihood of COVID-19 vaccine uptake. Those who believed that there was a conspiracy behind vaccines had a lower incidence of immunization. The low level of fear of COVID-19 was also associated with being vaccine hesitant. These results were congruent with previous studies [27,29,31,56,57,59]. In particular, prior intention for vaccination and confidence in vaccines were the strongest predictors of subsequent vaccine uptake. These findings support the idea that bridging the vaccination intention-to-action gap could be accomplished by overcoming concerns regarding vaccine safety [1,60].

Clinical Implications

This study has important clinical implications. The public authorities need to start by acknowledging the disparities in vaccine uptake in the nonpriority group, including younger healthy adults. In Japan, the largest percentage of confirmed cases of COVID-19 was among individuals aged under 30 years [61], highlighting their significance as key players in COVID-19 transmission. Matrajt et al [62] claimed that switching the vaccine allocation target to the high-transmission groups from the high-risk groups would achieve high vaccine effectiveness, as is the case with influenza vaccines [63,64]. In other words, it is possible to reduce the number of patients with COVID-19 in the older adults group by controlling the number of patients in the young adults group. Therefore, it is essential to keep examining whether the vaccine priority settings were appropriate for maximizing vaccine effectiveness. In promoting vaccine uptake, the content of the message should be tailored to each population group, taking into account their attitudes toward vaccines. When targeting young and healthy groups, those responsible for delivering the message should recognize the concerns of these groups about vaccine safety and effectiveness. Changing the perception of vaccine risk would be a key strategy in promoting vaccination. Needless to say, the development of new vaccines that have fewer side effects and are more effective against new types of mutation is expected. Our results also suggest that it is important to begin the promotion of vaccine acceptance before the introduction of new vaccines. Even after COVID-19 has converged, other emerging infectious diseases could lead to pandemics in the future. One possible way to reduce vaccine hesitancy in introducing new vaccines is to increase the number of people who receive influenza vaccination or undergo medical checkups regularly. These behaviors would reduce the barriers to vaccination if a new virus emerged for which vaccines are required [65].

Limitations

Despite the abovementioned strengths of this study, there are some limitations that should be noted. First, vaccine-related scales such as the vaccine hesitancy scale were only assessed at T2. At that time, some respondents had already received vaccination, while others had not. Therefore, we were unable to draw conclusions regarding the direction of causality between these variables. Second, due to the nature of self-administered surveys, recall bias and reporting bias may have occurred. Third, the classification of the priority groups was not completely accurate. For example, although the severity of each disease was set as a condition for priority, information on this was not obtained. As a proxy, a hospital visit for the disease was used to define the group. In addition, individuals who reached the age of 65 years in March 2021 were misclassified into the nonpriority group; however, the number of relevant participants was estimated to be small enough to affect the result. Participants working in a nursing home, one of the vaccination priority groups, could not be identified. Finally, we could not distinguish within-group differences in vaccine priority, such as frontline workers or office workers in a hospital.

Conclusions

Our findings demonstrate that the priority settings at the start of the COVID-19 vaccination program had a significant impact on vaccine coverage 1 year later. Japan achieved a high rate of vaccine uptake, resulting in a low COVID-19 mortality rate at the start of 2022. However, things took a drastic turn in 2022 with the outbreak of the omicron variant, which was a complete change from the year 2021. At its peak in August 2022, there were >250,000 cases and >300 deaths per day. Japan experienced another surge in infections in January 2023. After its peak had passed, the government downgraded COVID-19 to category 5 of infectious disease, the same rank as seasonal influenza, in May 2023. Although the threat of COVID-19 appears to have subsided for the moment, there remains the possibility that new emerging infectious diseases will arise and cause unprecedented situations. Further research is necessary to determine the effectiveness of vaccine distribution procedures in preventing COVID-19 from 2022 to 2023. Such efforts are essential for reducing the impact of the next pandemic on individuals and society.

Acknowledgments

The authors appreciate the valuable feedback from the Japan COVID-19 and Society Internet Survey study group researchers. The authors thank Dr Julia Mortimer for her English language editing. The authors express their gratitude to Tomohiro Shinozaki for his advice on the methodology of analysis.

This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI grants (grant 17H03589, 19K10671, 19K10446, 18H03107, 18H03062, 19H03860, 21H04856, and 21H03203), the JSPS Grant-in-Aid for Young Scientists (grant 19K19439), the Research Support Program to Apply the Wisdom of the University to Tackle COVID-19 Related Emergency Problems, University of Tsukuba, Health Labour Sciences Research Grant (grant 19FA1005, 19FG2001, 19FA1012, and 21HA2016), the Japan Agency for Medical Research and Development (AMED; grant 2033648), and the Individual Research Allowance at Kanagawa University of Human Services. The findings and conclusions of this paper are the sole responsibility of the authors and do not represent the official views of the research funders.

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

AO received personal fees from MNES Inc and Kyowa Kirin Inc outside the submitted work.

Multimedia Appendix 1
Exclusion criteria for analysis.
[DOCX File , 20 KB - publichealth_v9i1e42143_app1.docx]

Multimedia Appendix 2
Summary of response options for COVID-19 vaccine intention or uptake for each time point, as well as classification of uptake status at T3 for the following modified Poisson regression analysis.
[DOCX File , 22 KB - publichealth_v9i1e42143_app2.docx]

Multimedia Appendix 3
Flowchart of eligible participants for analysis.
[XLSX File (Microsoft Excel File), 103 KB - publichealth_v9i1e42143_app3.xlsx]

Multimedia Appendix 4
Summary of preexisting conditions among the non–health care worker priority group aged 18-64 years with preexisting conditions, as well as the profession of those in the health care worker priority group.
[DOCX File , 26 KB - publichealth_v9i1e42143_app4.docx]

Multimedia Appendix 5
COVID-19 infection over 1 year and risk ratios for COVID-19 vaccine uptake.
[DOCX File , 32 KB - publichealth_v9i1e42143_app5.docx]

Multimedia Appendix 6
Shift from prior COVID-19 vaccine intention in February 2021 to actual vaccine uptake in February 2022, stratified by the nonpriority group and the 3 priority groups.
[PPTX File , 967 KB - publichealth_v9i1e42143_app6.pptx]

Multimedia Appendix 7
Shift in reason for getting vaccinated against COVID-19 from February 2021 to February 2022, stratified by the nonpriority group and the 3 priority groups.
[PPTX File , 267 KB - publichealth_v9i1e42143_app7.pptx]

Multimedia Appendix 8

Shift in reason for not getting vaccinated against COVID-19 from February 2021 to February 2022, stratified by the nonpriority group and the 3 priority groups.

[PPTX File , 250 KB - publichealth_v9i1e42143_app8.pptx]

Multimedia Appendix 9

Risk ratios for COVID-19 vaccination uptake status at T3 (February 2022) using sociodemographic factors, behavioral factors, and vaccine hesitancy–related scales.

[DOCX File , 35 KB - publichealth_v9i1e42143_app9.docx]

References

1. Wouters OJ, Shadlen KC, Salcher-Konrad M, Pollard AJ, Larson HJ, Teerawattananon Y, et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. Lancet 2021 Mar 13;397(10278):1023-1034 [FREE Full text] [doi: 10.1016/S0140-6736(21)00306-8] [Medline: 33587887]
2. Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. BNT162b2 mRNA COVID-19 vaccine in a nationwide mass vaccination setting. N Engl J Med 2021 Apr 15;384(15):1412-1423 [FREE Full text] [doi: 10.1056/NEJMoa2101765] [Medline: 33626250]
3. Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of COVID-19 vaccines against the B.1.617.2 (Delta) variant. N Engl J Med 2021 Aug 12;385(7):585-594 [FREE Full text] [doi: 10.1056/NEJMoa2108891] [Medline: 34289274]
4. Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. Lancet 2021 May 15;397(10287):1819-1829 [FREE Full text] [doi: 10.1016/S0140-6736(21)00947-8] [Medline: 33964222]
5. Arbel R, Sergienko R, Friger M, Peretz A, Beckenstein T, Yaron S, et al. Effectiveness of a second BNT162b2 booster vaccine against hospitalization and death from COVID-19 in adults aged over 60 years. Nat Med 2022 Jul;28(7):1486-1490. [doi: 10.1038/s41591-022-01832-0] [Medline: 35468276]
6. Patel MD, Rosenstrom E, Ivy JS, Mayorga ME, Keskinocak P, Boyce RM, et al. Association of simulated COVID-19 vaccination and nonpharmaceutical interventions with infections, hospitalizations, and mortality. JAMA Netw Open 2021 Jun 01;4(6):e2110782 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.10782] [Medline: 34061203]
7. Tenforde MW, Patel MM, Ginde AA, Douin DJ, Talbot HK, Casey JD, InfluenzaOther Viruses in the Acutely Ill (IVY) Network. Effectiveness of severe acute respiratory syndrome coronavirus 2 messenger RNA vaccines for preventing coronavirus disease 2019 hospitalizations in the United States. Clin Infect Dis 2022 May 03;74(9):1515-1524 [FREE Full text] [doi: 10.1093/cid/ciab687] [Medline: 34358310]
8. Tenforde MW, Self WH, Adams K, Gaglani M, Ginde AA, McNeal T, Influenza and Other Viruses in the Acutely Ill (IVY) Network. Association between mRNA vaccination and COVID-19 hospitalization and disease severity. JAMA 2021 Nov 23;326(20):2043-2054 [FREE Full text] [doi: 10.1001/jama.2021.19499] [Medline: 34734975]
9. Persad G, Peek ME, Emanuel EJ. Fairly prioritizing groups for access to COVID-19 vaccines. JAMA 2020 Oct 27;324(16):1601-1602. [doi: 10.1001/jama.2020.18513] [Medline: 32910182]
10. Epstein S, Ayers K, Swenor BK. COVID-19 vaccine prioritisation for people with disabilities. Lancet Public Health 2021 Jun;6(6):e361 [FREE Full text] [doi: 10.1016/S2468-2667(21)00093-1] [Medline: 33894137]
11. Cylus J, Panteli D, van Ginneken E. Who should be vaccinated first? Comparing vaccine prioritization strategies in Israel and European countries using the COVID-19 health system response monitor. Isr J Health Policy Res 2021 Feb 19;10(1):16 [FREE Full text] [doi: 10.1186/s13584-021-00453-1] [Medline: 33608023]
12. A global framework to ensure equitable and fair allocation of COVID-19 products and potential implications for COVID-19 vaccines. World Health Organization. 2020. URL: https://bit.ly/32rhHPb [accessed 2022-07-08]
13. Buckner JH, Chowell G, Springborn MR. Dynamic prioritization of COVID-19 vaccines when social distancing is limited for essential workers. Proc Natl Acad Sci U S A 2021 Apr 20;118(16):e2025786118 [FREE Full text] [doi: 10.1073/pnas.2025786118] [Medline: 33811185]
14. Shin-gata corona vaccine Q&A. Ministry of Health, Labour and Welfare. 2022. URL: https://www.cov19-vaccine.mhlw.go.jp/qa/0033.html [accessed 2023-03-04]
15. Mori H, Naito T. A rapid increase in the COVID-19 vaccination rate during the Olympic and Paralympic games 2021 in Japan. Hum Vaccin Immunother 2022 Dec 31;18(1):2010440 [FREE Full text] [doi: 10.1080/21645515.2021.2010440] [Medline: 34893009]
16. Kosaka M, Hashimoto T, Ozaki A, Tanimoto T, Kami M. Delayed COVID-19 vaccine roll-out in Japan. Lancet 2021 Jun 19;397(10292):2334-2335 [FREE Full text] [doi: 10.1016/S0140-6736(21)01220-4] [Medline: 34089658]
17. Shin-gata corona 3 kaime no vaccine sessyu zen-jinkou no hansu koeru. Nippon Hōsō Kyōkai. 2022. URL: https://www3.nhk.or.jp/news/html/20220425/k10013597731000.html [accessed 2022-07-04]

18. Okubo R, Yoshioka T, Ohfuji S, Matsuo T, Tabuchi T. COVID-19 vaccine hesitancy and its associated factors in Japan. Vaccines (Basel) 2021 Jun 17;9(6):662 [FREE Full text] [doi: 10.3390/vaccines9060662] [Medline: 34204465]
19. Ishimaru T, Okawara M, Ando H, Hino A, Nagata T, Tateishi S, CORoNaWork Project. Gender differences in the determinants of willingness to get the COVID-19 vaccine among the working-age population in Japan. Hum Vaccin Immunother 2021 Nov 02;17(11):3975-3981 [FREE Full text] [doi: 10.1080/21645515.2021.1947098] [Medline: 34213406]
20. Machida M, Nakamura I, Kojima T, Saito R, Nakaya T, Hanibuchi T, et al. Acceptance of a COVID-19 vaccine in Japan during the COVID-19 pandemic. Vaccines (Basel) 2021 Mar 03;9(3):210 [FREE Full text] [doi: 10.3390/vaccines9030210] [Medline: 33802285]
21. Yoda T, Katsuyama H. Willingness to receive COVID-19 vaccination in Japan. Vaccines (Basel) 2021 Jan 14;9(1):48 [FREE Full text] [doi: 10.3390/vaccines9010048] [Medline: 33466675]
22. Okamoto S, Kamimura K, Komamura K. COVID-19 vaccine hesitancy and vaccine passports: a cross-sectional conjoint experiment in Japan. BMJ Open 2022 Jun 16;12(6):e060829 [FREE Full text] [doi: 10.1136/bmjopen-2022-060829] [Medline: 35710243]
23. Kadoya Y, Watanapongvanich S, Yuktadatta P, Putthinun P, Lartey ST, Khan MS. Willing or hesitant? A socioeconomic study on the potential acceptance of COVID-19 vaccine in Japan. Int J Environ Res Public Health 2021 May 02;18(9):4864 [FREE Full text] [doi: 10.3390/ijerph18094864] [Medline: 34063327]
24. Hara M, Ishibashi M, Nakane A, Nakano T, Hirota Y. Differences in COVID-19 vaccine acceptance, hesitancy, and confidence between healthcare workers and the general population in Japan. Vaccines (Basel) 2021 Nov 24;9(12):1389 [FREE Full text] [doi: 10.3390/vaccines9121389] [Medline: 34960135]
25. Nomura S, Eguchi A, Yoneoka D, Kawashima T, Tanoue Y, Murakami M, et al. Reasons for being unsure or unwilling regarding intention to take COVID-19 vaccine among Japanese people: a large cross-sectional national survey. Lancet Reg Health West Pac 2021 Sep;14:100223 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100223] [Medline: 34368797]
26. Harada T, Watanabe T. Changes in vaccine hesitancy in Japan across five months during the COVID-19 pandemic and its related factors. Vaccines (Basel) 2021 Dec 26;10(1):25 [FREE Full text] [doi: 10.3390/vaccines10010025] [Medline: 35062686]
27. Wang J, Zhu H, Lai X, Zhang H, Huang Y, Feng H, et al. From COVID-19 vaccination intention to actual vaccine uptake: a longitudinal study among Chinese adults after six months of a national vaccination campaign. Expert Rev Vaccines 2022 Mar;21(3):385-395. [doi: 10.1080/14760584.2022.2021076] [Medline: 34974779]
28. Shiloh S, Peleg S, Nudelman G. Vaccination against COVID-19: a longitudinal trans-theoretical study to determine factors that predict intentions and behavior. Ann Behav Med 2022 Apr 02;56(4):357-367. [doi: 10.1093/abm/kaab101] [Medline: 34864833]
29. Smith LE, Sim J, Cutts M, Dasch H, Amlôt R, Sevdalis N, et al. Psychosocial factors affecting COVID-19 vaccine uptake in the UK: a prospective cohort study (CoVAccS - Wave 3). Vaccine X 2023 Apr;13:100276 [FREE Full text] [doi: 10.1016/j.jvacx.2023.100276] [Medline: 36819214]
30. Hilverda F, Vollmann M. The role of risk perception in students' COVID-19 vaccine uptake: a longitudinal study. Vaccines (Basel) 2021 Dec 24;10(1):22 [FREE Full text] [doi: 10.3390/vaccines10010022] [Medline: 35062683]
31. Lutrick K, Groom H, Fowlkes AL, Groover KD, Gaglani M, Rivers P, et al. COVID-19 vaccine perceptions and uptake in a national prospective cohort of essential workers. Vaccine 2022 Jan 24;40(3):494-502 [FREE Full text] [doi: 10.1016/j.vaccine.2021.11.094] [Medline: 34906392]
32. Griffin B, Conner M, Norman P. Applying an extended protection motivation theory to predict COVID-19 vaccination intentions and uptake in 50-64 year olds in the UK. Soc Sci Med 2022 Apr;298:114819 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114819] [Medline: 35245755]
33. Tabuchi T, Shinozaki T, Kunugita N, Nakamura M, Tsuji I. Study profile: the Japan "society and new tobacco" internet survey (JASTIS): a longitudinal internet cohort study of heat-not-burn tobacco products, electronic cigarettes, and conventional tobacco products in Japan. J Epidemiol 2019 Nov 05;29(11):444-450 [FREE Full text] [doi: 10.2188/jea.JE20180116] [Medline: 30318495]
34. Luyten J, Bruyneel L, van Hoek AJ. Assessing vaccine hesitancy in the UK population using a generalized vaccine hesitancy survey instrument. Vaccine 2019 Apr 24;37(18):2494-2501. [doi: 10.1016/j.vaccine.2019.03.041] [Medline: 30940484]
35. Shapiro GK, Holding A, Perez S, Amsel R, Rosberger Z. Validation of the vaccine conspiracy beliefs scale. Papillomavirus Res 2016 Dec;2:167-172 [FREE Full text] [doi: 10.1016/j.pvr.2016.09.001] [Medline: 29074176]
36. Thaker J. The persistence of vaccine hesitancy: COVID-19 vaccination intention in New Zealand. J Health Commun 2021 Feb 01;26(2):104-111. [doi: 10.1080/10810730.2021.1899346] [Medline: 33719898]
37. Ahorsu DK, Lin CY, Imani V, Saffari M, Griffiths MD, Pakpour AH. The fear of COVID-19 scale: development and initial validation. Int J Ment Health Addict 2022;20(3):1537-1545 [FREE Full text] [doi: 10.1007/s11469-020-00270-8] [Medline: 32226353]
38. Midorikawa H, Aiba M, Lebowitz A, Taguchi T, Shiratori Y, Ogawa T, et al. Confirming validity of the fear of COVID-19 scale in Japanese with a nationwide large-scale sample. PLoS One 2021 Feb 10;16(2):e0246840 [FREE Full text] [doi: 10.1371/journal.pone.0246840] [Medline: 33566868]

39. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004 Apr 01;159(7):702-706. [doi: 10.1093/aje/kwh090] [Medline: 15033648]
40. COVID-19 vaccines. Prime Minister of Japan and His Cabinet. 2022. URL: https://japan.kantei.go.jp/ongoingtopics/vaccine.html [accessed 2022-07-04]
41. Kwon Y, Cho HY, Lee YK, Bae GR, Lee SG. Relationship between intention of novel influenza a (H1N1) vaccination and vaccination coverage rate. Vaccine 2010 Dec 16;29(2):161-165. [doi: 10.1016/j.vaccine.2010.10.063] [Medline: 21055495]
42. Liao Q, Cowling BJ, Lam WW, Fielding R. Factors affecting intention to receive and self-reported receipt of 2009 pandemic (H1N1) vaccine in Hong Kong: a longitudinal study. PLoS One 2011 Mar 11;6(3):e17713 [FREE Full text] [doi: 10.1371/journal.pone.0017713] [Medline: 21412418]
43. Paul E, Steptoe A, Fancourt D. Attitudes towards vaccines and intention to vaccinate against COVID-19: implications for public health communications. Lancet Reg Health Eur 2021 Feb;1:100012 [FREE Full text] [doi: 10.1016/j.lanepe.2020.100012] [Medline: 33954296]
44. Robertson E, Reeve KS, Niedzwiedz CL, Moore J, Blake M, Green M, et al. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. Brain Behav Immun 2021 May;94:41-50 [FREE Full text] [doi: 10.1016/j.bbi.2021.03.008] [Medline: 33713824]
45. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrachi M, Zigron A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol 2020 Aug;35(8):775-779. [doi: 10.1007/s10654-020-00671-y] [Medline: 32785815]
46. Lin C, Tu P, Beitsch LM. Confidence and receptivity for COVID-19 vaccines: a rapid systematic review. Vaccines (Basel) 2020 Dec 30;9(1):16 [FREE Full text] [doi: 10.3390/vaccines9010016] [Medline: 33396832]
47. Araf Y, Akter F, Tang YD, Fatemi R, Parvez MS, Zheng C, et al. Omicron variant of SARS-CoV-2: genomics, transmissibility, and responses to current COVID-19 vaccines. J Med Virol 2022 May;94(5):1825-1832 [FREE Full text] [doi: 10.1002/jmv.27588] [Medline: 35023191]
48. Hachmann NP, Miller J, Collier AY, Ventura JD, Yu J, Rowe M, et al. Neutralization escape by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4, and BA.5. N Engl J Med 2022 Jul 07;387(1):86-88. [doi: 10.1056/nejmc2206576]
49. Betsch C, Schmid P, Heinemeier D, Korn L, Holtmann C, Böhm R. Beyond confidence: development of a measure assessing the 5C psychological antecedents of vaccination. PLoS One 2018 Dec 7;13(12):e0208601 [FREE Full text] [doi: 10.1371/journal.pone.0208601] [Medline: 30532274]
50. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
51. Hall VJ, Foulkes S, Saei A, Andrews N, Oguti B, Charlett A, et al. COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): a prospective, multicentre, cohort study. Lancet 2021;397(10286):1725-1735 [FREE Full text] [doi: 10.1016/S0140-6736(21)00790-X]
52. Hammerman A, Sergienko R, Friger M, Beckenstein T, Peretz A, Netzer D, et al. Effectiveness of the BNT162b2 vaccine after recovery from COVID-19. N Engl J Med 2022 Mar 31;386(13):1221-1229 [FREE Full text] [doi: 10.1056/NEJMoa2119497] [Medline: 35172072]
53. Iwasaki A, Grubaugh ND. Why does Japan have so few cases of COVID-19? EMBO Mol Med 2020 May 08;12(5):e12481 [FREE Full text] [doi: 10.15252/emmm.202012481] [Medline: 32275804]
54. Guilmoto CZ. COVID-19 death rates by age and sex and the resulting mortality vulnerability of countries and regions in the world. medRxiv. Preprint posted online May 20, 2020 2023 [FREE Full text] [doi: 10.1101/2020.05.17.20097410]
55. Naito T, Yan Y, Tabe Y, Seyama K, Deshpande GA. Real-world evidence for the effectiveness and breakthrough of BNT162b2 mRNA COVID-19 vaccine at a medical center in Japan. Hum Vaccin Immunother 2022 Dec 31;18(1):1-2 [FREE Full text] [doi: 10.1080/21645515.2021.1984124] [Medline: 34614387]
56. Ruiz JB, Bell RA. Predictors of intention to vaccinate against COVID-19: results of a nationwide survey. Vaccine 2021 Feb 12;39(7):1080-1086 [FREE Full text] [doi: 10.1016/j.vaccine.2021.01.010] [Medline: 33461833]
57. Freeman D, Loe BS, Chadwick A, Vaccari C, Waite F, Rosebrock L, et al. COVID-19 vaccine hesitancy in the UK: the Oxford coronavirus explanations, attitudes, and narratives survey (Oceans) II. Psychol Med 2022 Oct;52(14):3127-3141 [FREE Full text] [doi: 10.1017/S0033291720005188] [Medline: 33305716]
58. Murphy J, Vallières F, Bentall RP, Shevlin M, McBride O, Hartman TK, et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. Nat Commun 2021 Jan 04;12(1):29 [FREE Full text] [doi: 10.1038/s41467-020-20226-9] [Medline: 33397962]
59. Taylor S, Landry CA, Paluszek MM, Groenewoud R, Rachor GS, Asmundson GJ. A proactive approach for managing COVID-19: the importance of understanding the motivational roots of vaccination hesitancy for SARS-CoV2. Front Psychol 2020 Oct 19;11:575950 [FREE Full text] [doi: 10.3389/fpsyg.2020.575950] [Medline: 33192883]
60. Brewer NT, Chapman GB, Rothman AJ, Leask J, Kempe A. Increasing vaccination: putting psychological science into action. Psychol Sci Public Interest 2017 Dec;18(3):149-207. [doi: 10.1177/1529100618760521] [Medline: 29611455]
61. Visualizing the data: information on COVID-19 infections. Ministry of Health, Labour and Welfare. URL: https://covid19.mhlw.go.jp/en/ [accessed 2022-07-15]
62. Matrajt L, Eaton J, Leung T, Brown ER. Vaccine optimization for COVID-19: who to vaccinate first? Sci Adv 2020 Feb 03;7(6):eabf1374 [FREE Full text] [doi: 10.1126/sciadv.abf1374] [Medline: 33536223]

63. Medlock J, Galvani AP. Optimizing influenza vaccine distribution. Science 2009 Sep 25;325(5948):1705-1708. [doi: 10.1126/science.1175570] [Medline: 19696313]
64. Longini Jr IM, Ackerman E, Elveback LR. An optimization model for influenza a epidemics. Math Biosci 1978 Jan;38(1-2):141-157 [FREE Full text] [doi: 10.1016/0025-5564(78)90023-8]
65. Caserotti M, Girardi P, Rubaltelli E, Tasso A, Lotto L, Gavaruzzi T. Associations of COVID-19 risk perception with vaccine hesitancy over time for Italian residents. Soc Sci Med 2021 Mar;272:113688 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113688] [Medline: 33485215]

Abbreviations

HCW: health care worker
JACSIS: Japan COVID-19 and Society Internet Survey
JASTIS: Japan Society and New Tobacco Internet Survey
RR: risk ratio

Edited by T Sanchez, A Mavragani; submitted 26.08.22; peer-reviewed by O Rahaman, A Miyawaki, R Hirota; comments to author 03.03.23; revised version received 18.03.23; accepted 24.05.23; published 10.07.23.

Please cite as:
Hori D, Takahashi T, Ozaki A, Tabuchi T
The Impact of Priority Settings at the Start of COVID-19 Mass Vaccination on Subsequent Vaccine Uptake in Japan: One-Year Prospective Cohort Study
JMIR Public Health Surveill 2023;9:e42143
URL: https://publichealth.jmir.org/2023/1/e42143
doi:10.2196/42143
PMID:37235691

Original Paper

Dynamic Trends and Underlying Factors of COVID-19 Vaccine Booster Hesitancy in Adults: Cross-Sectional Observational Study

Jian Wu[1,2], Prof Dr, PhD; Mingze Ma[1,2], PhD; Quanman Li[1,2], PhD; Xinghong Guo[1,2], PhD; Clifford Silver Tarimo[1,3], PhD; Shiyu Jia[1,2], MPH, MD; Xue Zhou[4], Prof Dr, PhD; Meiyun Wang[5], Prof Dr, PhD; Jianqin Gu[6], Prof Dr, PhD; Yudong Miao[1,2], Prof Dr, PhD; Beizhu Ye[1,2], Prof Dr, PhD

[1]Department of Health Management, College of Public Health, Zhengzhou University, Zhengzhou, Henan, China
[2]Henan Province Engineering Research Center of Health Economy & Health Technology Assessment, Zhengzhou, China
[3]Department of Science and Laboratory Technology, Dar es Salaam Institute of Technology, Dar es Salaam, United Republic of Tanzania
[4]Department of Public Utilities Management, College of Health Management, Mudanjiang Medical University, Hei Longjiang, China
[5]Henan Provincial People's Hospital, Zhengzhou University, Zhengzhou, China
[6]School of Medicine, Southern University of Science and Technology, Shenzhen, China

Corresponding Author:
Beizhu Ye, Prof Dr, PhD
Department of Health Management
College of Public Health
Zhengzhou University
No.100, Science Avenue, Gaoxin District
Zhengzhou, Henan, 450001
China
Phone: 86 67739815
Fax: 86 67739815
Email: yebeizhu@zzu.edu.cn

Abstract

Background: COVID-19 vaccine hesitancy reduces vaccination rates, which is detrimental to building herd immunity and halting the spread of COVID-19 and its variations. Most researches have simply identified the reasons affecting COVID-19 vaccination reluctance without delving into its dynamics, which makes forecasting future trends difficult.

Objective: This study aimed to examine the current COVID-19 vaccine booster hesitancy rate in Chinese adults as well as the dynamics of vaccine hesitancy and its influencing factors. The results of this study will have practical implications for policy responses in mainland China, and effective COVID-19 booster vaccination in specific populations.

Methods: The web-based survey was completed by creating questionnaires and using a stratified random sampling method to collect information from adults (≥18 years old) among 2556 households in 4 geographical regions of China. We collected sociodemographic information, health status, awareness of COVID-19 and its vaccine, self-perceptions, trust in medical staff and vaccine developers, and so on. The odds ratios and 95% CI for the statistical associations were estimated using logistic regression models.

Results: Overall, 6659 participants (females: n=3540, 53.2%; males: n=3119, 46.8%) responded. In total, 533 (8%; 95% CI 7.4%-8.7%) participants presented a clear hesitancy in receiving the COVID-19 booster vaccination, while 736 (11.1%; 95% CI 10.3%-11.8%) expressed hesitancy in regular booster vaccination. A higher prevalence of vaccine hesitancy in both booster vaccination and regular booster vaccination was observed among participants with a history of allergies, experiencing chronic disease, lower levels of public health prevention measures or susceptibility or benefits or self-efficiency, higher levels of severity or barriers, and lower trust in both medical staff and vaccine developers ($P<.05$). The females and participants with higher education levels, higher levels of barriers, lower levels of susceptibility, and lower trust in vaccine developers preferred to have attitudinal changes from acceptance to hesitancy, while people with higher education levels, lower self-report health conditions, experiencing chronic disease, history of allergies, and lower trust in medical staff and developers were all positively associated with constant COVID-19 booster hesitancy.

Conclusions: The prevalence of COVID-19 vaccine booster hesitancy is not high in mainland China. However, there is a slight increment in hesitancy on regular booster vaccination. Conducting targeted information guidance for people with higher education levels and chronic diseases, as well as improving accessibility to booster vaccination and increasing trust in medical staff and vaccine producers may be highly effective in reducing vaccine hesitancy.

(JMIR Public Health Surveill 2023;9:e44822) doi:10.2196/44822

KEYWORDS

COVID-19 vaccine; vaccine hesitancy; COVID-19 booster vaccination; influencing factors; China

Introduction

Vaccinating a sufficient proportion of the population against COVID-19 is one of the most effective methods for achieving "herd immunity," lowering morbidity, and increasing population survival. It is also the most cost-effective and straightforward intervention to prevent the COVID-19 pandemic [1,2]. Prevention of the COVID-19 virus and its variants through COVID-19 vaccination can confer benefits related to morbidity and mortality [3], mainly including the protection of immunocompromised populations such as infants, young children, and the elderly [4,5]. According to the latest reports from the World Health Organization (WHO), two-thirds of the world's population is vaccinated, including 75% of health workers and older people. However, there are still wide disparities in vaccination rates [6]. The COVID-19 booster vaccination is an important tool for consolidating herd immunity and responding to the COVID-19 pandemic and deaths [7], and receiving the COVID-19 booster vaccine is essential to achieve adequate immunization coverage to control the global pandemic [8,9]. Understanding people's attitudes to COVID-19 booster vaccination aids in developing COVID-19 booster vaccination plans, which ensure the consolidation of herd immunity and minimize health inequalities.

Since the global outbreak of the COVID-19 epidemic, several vaccines have been developed, such as inactivated vaccines, live-attenuated vaccines adenovirus vector vaccines, and recombinant protein vaccines [10,11]. Although these multiple vaccines are considered safe and effective [3,4], and countries around the world are increasing the supply of vaccines, waiving vaccination fees [12], and actively promoting the COVID-19 booster vaccination, there are still concerns about vaccine hesitancy, which affects the vaccination rates of COVID-19 vaccines as well as booster vaccination. This situation may not be helpful to bolster herd immunity in order to prevent the further spread of COVID-19 and its variations. According to the WHO, "Vaccine hesitancy" has been considered as one of the top 10 public health problems across the world [13-15]. It refers to the delay or refusal to receive safe vaccination services, which can accompany the use of vaccines [3,15-17]. The danger of vaccine hesitancy is that it can lead to immunization of populations below the herd immunity threshold and ultimately fail to protect the populations. This is particularly true for people whose immunity is very weak to be vaccinated, such as infants, the elderly, or patients with serious illnesses, who are the first to be exposed to infectious diseases if the herd immunity barrier is breached [4,18]. The COVID-19 vaccine booster hesitancy is a type of vaccine hesitancy and is a continuation of people's hesitancy about the COVID-19 vaccine. Therefore, what factors influence people's hesitancy to receive the COVID-19 booster vaccine? An in-depth study of this issue would be of great relevance in guiding the next phase of the COVID-19 booster vaccination program.

The reasons for vaccine hesitancy are thought to be complex and case specific. The existing literature cites concern about COVID-19 vaccine safety and efficacy, vaccination accessibility, the severity of the regional epidemic, experience with routine vaccination, and guardians' willingness to vaccinate their children as the primary reasons for hesitancy. For example, in a study of Canadians, individuals with vaccine hesitancy had a higher prevalence of being concerned about vaccine side effects [19]. The antecedents of vaccine hesitancy range from a lack of knowledge and awareness to culturally rooted reservations [20]. The most commonly reported reasons parents chose not to vaccinate their child against COVID-19 were concerns about long-term adverse side effects and a negative reaction [21]. At the same time, concern about side effects is also the most common reason for COVID-19 vaccine hesitancy in low- and middle-income countries [3]. In addition, misinformation and disinformation can substantially result in vaccine hesitancy [22] and might influence the vaccine advice of health care providers [23]. These factors will probably influence people's hesitancy about the COVID-19 booster vaccination.

Understanding local barriers to vaccination and concerns about vaccination is critical to developing tailored interventions [24]. The current basic vaccination rate of COVID-19 vaccine in China exceeds 89% [12], which is probably one of the leading countries in the world not only because of its large population but also its extensive exchanges and closer cooperation with other regions [25,26]. However, the coverage of booster vaccination in China is not widespread, which is still a significant gap compared with the basic vaccination rate. Furthermore, compared to other countries around the world, the epidemic prevention and control policies in mainland China are adequate, and vaccines are supplied for free and in sufficient quantities [27], but the problem of vaccine hesitancy remains unresolved. Focusing on the problem of vaccine hesitancy in COVID-19 booster vaccination in mainland China will be highly representative. The existing research on booster vaccination hesitancy is limited, and hence, we designed a multicenter study focusing on COVID-19 vaccine booster vaccination and regular booster vaccination in mainland China. This study aimed to understand the current COVID-19 vaccine hesitancy in booster vaccination of the Chinese population and analyze the dynamic trend of vaccine hesitancy as well as its influencing factors.

Methods

Procedures and Participants

From June 29, 2022, to July 2, 2022, we used a self-designed questionnaire to perform a preliminary survey by recruiting web-based volunteers. We conducted The Dynamic Evolution of COVID-19 Vaccination Study, a national, multicenter, observational household tracking survey from China using a stratified random sampling method. Eastern, Central, Western, and Northeast geographical regions within mainland China were selected to form the sample for this study. Therefore, 4 cities were selected from the Eastern (Changzhou and Jiangsu), Central (Zhengzhou, Henan), Western (Xi’ning and Qinghai), and Northeast (Mudanjiang and Heilongjiang) regions.

We calculated the minimum sample size of each region according to the population proportion of China’s Seventh National Population Census. In each province, the sample size of urban and rural areas was determined according to the proportion of population in the Seventh National Population Census, more than 2 cities and 2 rural areas were randomly selected for sampling. For all of the cities and families in the sample, coding was carried out followed by random sampling. Finally, all members of the selected family have been involved in this survey (age≥18 years) and completed the web-based or offline questionnaire with the assistance of the investigators. A total of 2556 households from 4 geographic regions of China were enrolled in the survey.

The study was designed as a 4-stage survey. For the first stage, we conducted the survey from August 3, 2022, to August 14, 2022. Data collection is still ongoing for the second, third, and fourth stages. Here, we describe the results from the first stage of the study only.

Ethics Approval

Participants were informed of the benefits and risks of participating in this study, and they provided informed consent. All data were used merely for research purposes. This study was approved by the Life Science Ethics Review Committee of Zhengzhou University (2021-01-12-05).

Questionnaire Design and Data Collection

Web-based questionnaires were designed to collect data through a survey conducted through the Wenjuanxing platform. We also conducted a face-to-face interview with participants who cannot use a smartphone. The questionnaire covered five thematic areas: (1) sociodemographic characteristics: age, gender, nation, religion, marital status, educational status, subjective social status in China or community, smoking status, drinking status, and physical activity; (2) health conditions: chronic disease, the history of allergies, self-report health condition (measured by EuroQol 5 Dimensions Questionnaire [28-30]); (3) perception of COVID-19 and its vaccine: public health prevention measures, awareness of COVID-19 vaccines, the channel of accessing information, the way of accessing information, the risk of COVID-19 infection, the history of COVID-19 infection, the convenience of vaccination, and the uptake of COVID-19 vaccine; (4) self-perception (Kaiser-Meyer-Olkin=0.796): perceived severity (Cronbach α=.879), perceived susceptibility (Cronbach α=.865), perceived benefits (Cronbach α=.907), perceived barriers (Cronbach α=.927), and self-efficiency (Cronbach α=.913; Kaiser-Meyer-Olkin=0.886); and (5) trust in medical staff and developers. All questionnaires are shown in Multimedia Appendix 1.

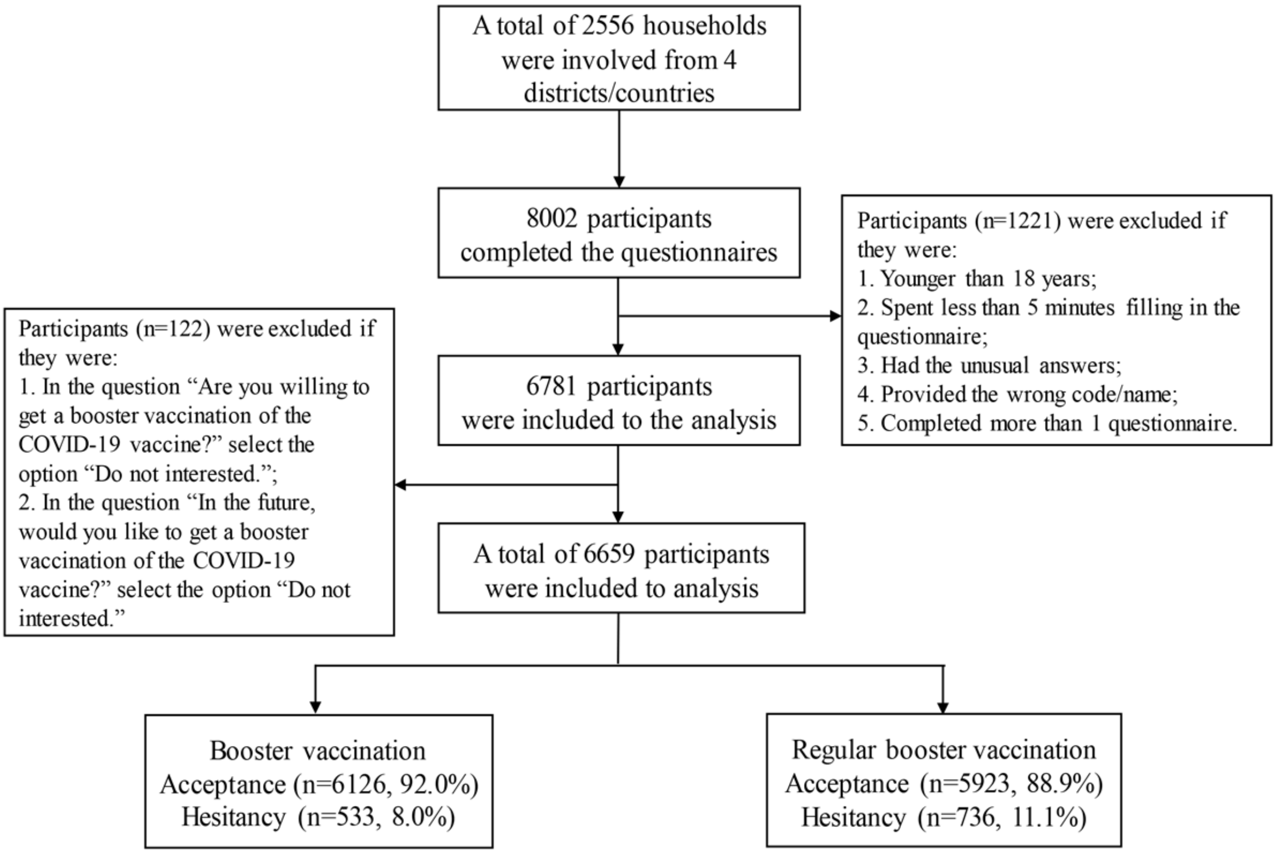
Assessments

Questionnaires with abnormal answers (contradictions and inconsistencies, eg, inconsistency or contradictions of sex) were excluded. In addition, we also excluded the participants who spent less than 5 minutes completing this questionnaire. Moreover, the remaining questionnaires were carefully reviewed by trained staff, and finally, 6659 questionnaires were deemed eligible for inclusion. The flowchart of participants is shown in Figure 1.

The primary outcome was COVID-19 vaccine hesitancy in both booster and regular booster vaccination, assessed by asking participants “Are you willing to get a booster vaccination of the COVID-19 vaccine?” and “In the future, would you like to get a booster vaccination of the COVID-19 vaccine?” Participants could choose 1 response from the options “(1) willing,” “(2) hesitant or delayed,” “(3) refused,” and “(4) do not interested” from the 2 questions. According to the definition of vaccine hesitancy, option 1 was regarded as “acceptance,” and options 2 and 3 were merged into “hesitancy.” In addition to this, participants were excluded if they selected the option “do not interested” (n=122).

A booster vaccination is a shot that is given after a period of time (which may be 6 months or more) after a population has completed a standard immunization program, and due to the gradual decay of antibodies in the body, the specific immunity in the body needs to be boosted with another booster vaccination. Regular booster vaccinations are administered after a period of booster vaccination to strengthen the immune system.

Figure 1. Flowchart of data processing and analysis.



Statistical Analysis

Chi-square tests were used to test differences in vaccine hesitancy between groups. Binary logistic regression analyses were carried out to examine factors associated with COVID-19 vaccine hesitancy in both booster and regular booster vaccination. A sensitivity analysis was performed by excluding the participants aged ≥80 years to test the robustness of results and assess the source of uncertainty. In stratified multistage sampling, post hoc stratification was used to weight the sample and the corresponding variables so that the sample's estimate of the total was unbiased. Odds ratios, 95% CI, and *P* values were calculated for each independent variable. All statistical analyses were carried out using SPSS (version 21.0; IBM Corp) and STATA (version 16.0; StataCorp). Differences were regarded as statistically significant if *P* values were less than .05.

Results

Characteristics of COVID-19 Vaccine Hesitancy in Booster Vaccination

Among the 8002 participants who completed the questionnaire, a total of 1343 participants who did not meet the inclusion criteria were subsequently excluded. Importantly, no statistically significant differences were found between the final 6659 participants included in the analysis and the overall sample of 8002 (*P*>.05) (Multimedia Appendix 2). The sociodemographic characteristics, awareness of COVID-19 vaccine, trust in the health care system as well as COVID-19 vaccine hesitancy in booster vaccination and regular booster vaccination of all study participants are displayed in Table 1. Overall, 533 adults (8%; 95% CI 7.4%-8.7%) showed a clear COVID-19 vaccine hesitancy in booster vaccination, while 6126 adults (92%; 95% CI 90.3%-94.2%) indicated that they were willing to vaccinate. Meanwhile, 736 participants (11.1%; 95% CI 10.3%-11.8%) expressed their hesitancy about regular booster vaccination of COVID-19 vaccine. Furthermore, we can see clearly that the participants of all ages are evenly distributed, the proportion of female participants (53.2%) is more than male (46.8%), and most of the participants (87.5%) were married.

A higher prevalence of vaccine hesitancy in both booster vaccination and regular booster vaccination was observed among participants with a history of allergies, experiencing chronic disease, lower levels of public health prevention measures or susceptibility or benefits or self-efficiency, higher levels of severity or barriers, and lower trust in both medical staff and vaccine developers (all *P*<.05; Figure 2).

Table 1. The characteristics and COVID-19 vaccine booster hesitancy of all study participants in China (n=6659).

Covariates	Total, n (%)	*P* value[a]	Vaccine hesitancy in booster vaccination (95% CI)[b]	*P* value	Vaccine hesitancy in regular booster vaccination (95% CI)	*P* value
Total participants	6659 (100)		8.0 (7.4-8.7)		11.1 (10.3-11.8)	
Age (years)		<.001		<.001		<.001
18-29	872 (13.1)		7.8 (6.0-9.6)		11.0 (8.9-13.1)	
30-39	1801 (27)		9.6 (8.2-10.9)		12.8 (11.3-14.5)	
40-49	1277 (19.2)		6.2 (4.9-7.7)		9.9 (8.2-11.5)	
50-59	1408 (21.1)		6.0 (4.7-7.2)		8.2 (6.8-9.7)	
≥60	1301 (19.5)		10.0 (8.4-11.6)		12.9 (11.1-14.7)	
Gender		<.001		.16		.02
Male	3119 (46.8)		7.5 (6.6-8.5)		10.1 (9.0-11.2)	
Female	3540 (53.2)		8.5 (7.5-9.4)		11.9 (10.9-13.0)	
Ethnic groups		<.001		.42		.27
Han	6472 (97.2)		8.1 (7.4-8.7)		11.1 (10.4-11.9)	
Minority	187 (2.8)		6.4 (2.9-9.9)		8.6 (4.6-12.6)	
Religion		<.001		.54		.88
Atheist	6362 (95.5)		8.1 (7.4-8.7)		11.1 (10.3-11.8)	
Others	297 (4.5)		7.1 (4.2-10.0)		10.8 (7.3-14.3)	
Marital status		<.001		.12		.43
Married	5829 (87.5)		8.2 (7.5-8.9)		11.2 (10.4-12.0)	
Others	830 (12.5)		6.6 (4.9-8.3)		10.2 (8.2-12.3)	
Educational status		<.001		<.001		<.001
Below high school	2882 (43.3)		6.6 (5.7-7.5)		8.4 (7.4-9.4)	
High school graduate	1661 (24.9)		8.1 (6.8-9.4)		11.9 (10.4-13.5)	
University graduate	2116 (31.8)		9.8 (8.6-11.1)		14.0 (12.5-15.5)	
Subjective social status in China		<.001		.01		.003
Level 1[c]	1942 (29.2)		7.8 (6.6-9.0)		9.9 (8.6-11.4)	
Level 2[c]	2825 (42.4)		9.0 (7.9-10.0)		12.4 (11.2-13.6)	
Level 3[c]	796 (12.0)		7.9 (6.0-9.8)		11.9 (9.7-14.2)	
Level 4[c]	1096 (16.5)		5.8 (4.5-7.2)		8.9 (7.2-10.5)	
Subjective social status in community		<.001		.01		<.001
Level 1[c]	1734 (26.0)		7.5 (6.3-8.8)		9.3 (7.9-10.7)	
Level 2[c]	2896 (43.5)		9.2 (8.1-10.2)		12.7 (11.5-13.9)	
Level 3[c]	858 (12.9)		7.6 (5.8-9.4)		11.9 (9.7-14.2)	
Level 4[c]	1171 (17.6)		6.2 (4.9-7.6)		9.0 (7.3-10.6)	
Self-report health condition (EQ-5D)[d]		0.01		<.001		<.001
Level 1[c]	1709 (25.7)		13.0 (11.4-14.7)		16.3 (14.5-18.1)	
Level 2[c]	1627 (24.4)		7.9 (6.6-9.3)		11.0 (9.5-12.6)	
Level 3[c]	1746 (26.2)		5.7 (4.6-6.9)		9.5 (8.1-10.9)	
Level 4[c]	1577 (23.7)		5.3 (4.2-6.4)		7.2 (6.0-8.5)	
Chronic disease		<.001		<.001		<.001

Covariates	Total, n (%)	*P* value[a]	Vaccine hesitancy in booster vaccination (95% CI)[b]	*P* value	Vaccine hesitancy in regular booster vaccination (95% CI)	*P* value
Yes	1030 (15.5)		6.9 (6.3-7.6)		17.1 (14.8-19.4)	
No	5269 (84.5)		13.9 (11.8-16.0)		10.0 (9.2-10.7)	
History of allergies		<.001		<.001		<.001
Yes	468 (7.0)		17.1 (13.7-20.5)		21.8 (18.1-25.5)	
No	5466 (82.1)		6.6 (5.9-7.3)		9.4 (8.6-10.1)	
Unclear	725 (10.9)		12.8 (10.4-15.3)		17.0 (14.2-19.7)	
Smoking status		<.001		.33		.05
Current smoker	1432 (21.5)		8.0 (6.6-9.5)		9.8 (8.2-11.3)	
Former smoker	402 (6.0)		10.0 (7.0-13.3)		13.9 (10.6-17.7)	
Never smoker	4825 (72.5)		7.9 (7.1-8.6)		11.2 (10.3-12.1)	
Drinking status		<.001		.04		.03
Current drinker	1706 (25.6)		7.8 (6.5-9.1)		10.7 (9.3-12.2)	
Former drinker	387 (5.8)		11.4 (8.2-14.5)		15.3 (11.7-18.8)	
Never drinker	4566 (68.6)		7.8 (7.0-8.6)		10.8 (9.9-11.7)	
Physical activity		<.001		<.001		<.001
High level	3399 (51.0)		6.4 (5.6-7.2)		9.0 (8.0-9.9)	
Middle level	2111 (31.7)		8.9 (7.7-10.1)		12.4 (11.0-13.8)	
Low level	1149 (17.3)		11.1 (9.3-13.1)		14.8 (12.7-16.9)	
Public health prevention measures		<.001		<.001		<.001
Low level	522 (7.8)		16.9 (13.7-20.1)		20.3 (16.9-23.8)	
Middle level	527 (7.9)		14.2 (11.3-17.2)		20.9 (17.4-24.3)	
High level	5610 (84.2)		6.6 (6.0-7.2)		9.3 (8.5-10.0)	
Awareness of COVID-19 vaccines		<.001		.03		.35
Level 1[c]	2533 (38.3)		9.2 (8.1-10.4)		11.5 (10.2-12.7)	
Level 2[c]	900 (13.5)		6.8 (5.1-8.4)		9.7 (7.7-11.6)	
Level 3[c]	1593 (23.9)		8.0 (6.6-9.3)		11.7 (10.1-13.3)	
Level 4[c]	1613 (24.2)		6.9 (5.7-8.1)		10.5 (9.0-12.0)	
Channel of vaccine information		<.001		.25		.15
We Media	1818 (27.3)		8.9 (7.6-10.2)		12.3 (10.8-13.9)	
Official media	704 (10.6)		7.7 (5.7-9.6)		10.2 (8.0-12.5)	
Others	4137 (62.1)		7.7 (6.9-8.5)		10.7 (9.7-11.6)	
Severity		<.001		<.001		<.001
Level 1[c]	1818 (27.3)		6.1 (5.0-7.2)		8.6 (7.3-9.9)	
Level 2[c]	1556 (23.4)		10.3 (8.8-11.8)		14.3 (12.6-16.2)	
Level 3[c]	2108 (31.7)		9.7 (8.5-11.1)		13.1 (11.7-14.6)	
Level 4[c]	1177 (17.7)		4.8 (3.6-6.1)		6.8 (5.4-8.2)	
Susceptibility		<.001		.006		<.001
Level 1[c]	1794 (26.9)		6.6 (5.5-7.8)		8.7 (7.4-10.0)	
Level 2[c]	1889 (28.4)		8.1 (6.9-9.3)		11.3 (9.9-12.7)	
Level 3[c]	2028 (30.5)		9.6 (8.3-10.9)		13.6 (12.1-15.1)	

Covariates	Total, n (%)	*P* value[a]	Vaccine hesitancy in booster vaccination (95% CI)[b]	*P* value	Vaccine hesitancy in regular booster vaccination (95% CI)	*P* value
Level 4[c]	948 (14.2)		7.1 (5.4-8.7)		9.6 (7.7-11.5)	
Benefits		<.001		<.001		<.001
Level 1[c]	1784 (26.8)		15.9 (14.2-17.6)		21.2 (19.3-23.1)	
Level 2[c]	2765 (41.5)		6.6 (5.6-7.5)		9.6 (8.5-10.7)	
Level 3[c]	2110 (31.7)		3.2 (2.5-4.0)		4.4 (3.5-5.3)	
Barriers		<.001		<.001		<.001
Level 1[c]	2026 (30.4)		2.6 (1.9-3.3)		3.0 (2.3-3.8)	
Level 2[c]	2453 (36.8)		4.4 (3.6-5.3)		6.8 (5.8-7.8)	
Level 3[c]	740 (11.1)		10.8 (8.6-13.1)		16.6 (13.9-19.3)	
Level 4[c]	1440 (21.6)		25.5 (23.3-27.6)		26.8 (24.5-29.1)	
Self-efficiency		<.001		<.001		<.001
Level 1[c]	4328 (65.0)		11.4 (10.4-12.4)		15.7 (14.6-16.8)	
Level 2[c]	191 (2.9)		6.8 (3.2-10.4)		10.5 (6.1-14.8)	
Level 3[c]	2140 (32.1)		1.3 (0.8-1.7)		1.7 (1.2-2.3)	
Trust in medical staff		<.001		<.001		<.001
Level 1[c]	2022 (30.4)		15.6 (14.0-17.2)		21.0 (19.2-22.7)	
Level 2[c]	1314 (19.7)		7.8 (6.4-9.3)		10.9 (9.2-12.8)	
Level 3[c]	1743 (26.2)		4.7 (3.7-5.7)		6.9 (5.7-8.1)	
Level 4[c]	1580 (23.7)		2.1 (1.9-2.8)		3.1 (2.3-4.0)	
Trust in developers		<.001		<.001		<.001
Level 1[c]	1783 (26.8)		17.2 (15.5-19.1)		23.6 (21.6-25.5)	
Level 2[c]	2152 (32.3)		6.2 (5.2-7.3)		9.3 (8.1-10.5)	
Level 3[c]	1194 (17.9)		4.4 (3.2-5.5)		6.0 (4.7-7.4)	
Level 4[c]	1530 (23.0)		2.6 (1.8-3.4)		2.9 (2.0-3.7)	

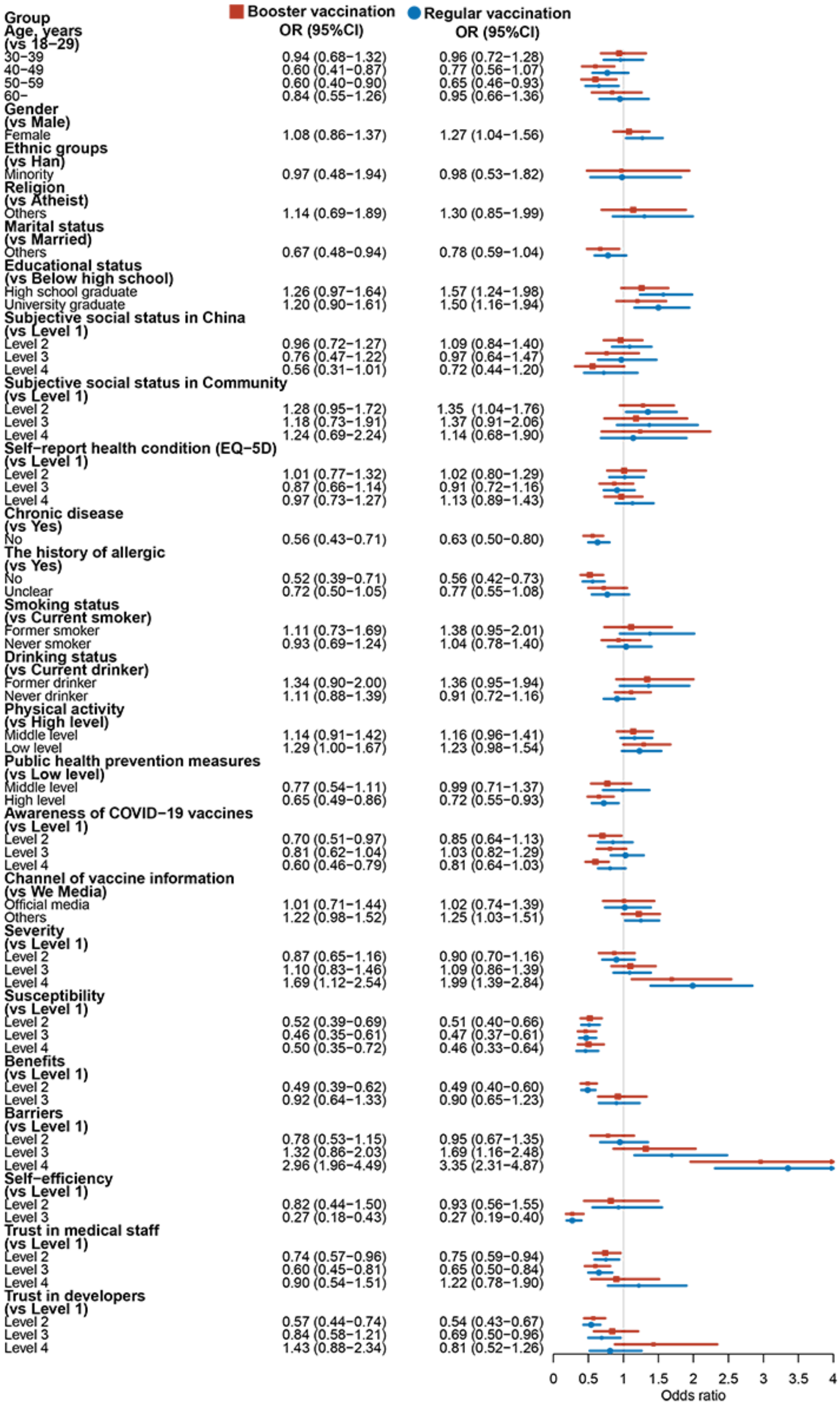
[a] Differences between categories within each variable.

[b] Row percentages derived from the total number in the corresponding row.

[c] Levels 1-4 indicate progressively higher degrees. The higher the degree, the higher level of social status in China or community, the better the self-assessment of health status, the more awareness of COVID-19 vaccine, the more severe or more barriers, the greater the susceptibility or benefits, the higher the self-efficacy and the more trust in medical staff and developers.

[d] EQ-5D: EuroQol 5 dimensions.

Figure 2. Influencing factors of vaccine hesitancy in both booster and regular booster vaccinations.



Group	🟥 Booster vaccination OR (95%CI)	● Regular vaccination OR (95%CI)
Age, years (vs 18−29)		
30−39	0.94 (0.68−1.32)	0.96 (0.72−1.28)
40−49	0.60 (0.41−0.87)	0.77 (0.56−1.07)
50−59	0.60 (0.40−0.90)	0.65 (0.46−0.93)
60−	0.84 (0.55−1.26)	0.95 (0.66−1.36)
Gender (vs Male)		
Female	1.08 (0.86−1.37)	1.27 (1.04−1.56)
Ethnic groups (vs Han)		
Minority	0.97 (0.48−1.94)	0.98 (0.53−1.82)
Religion (vs Atheist)		
Others	1.14 (0.69−1.89)	1.30 (0.85−1.99)
Marital status (vs Married)		
Others	0.67 (0.48−0.94)	0.78 (0.59−1.04)
Educational status (vs Below high school)		
High school graduate	1.26 (0.97−1.64)	1.57 (1.24−1.98)
University graduate	1.20 (0.90−1.61)	1.50 (1.16−1.94)
Subjective social status in China (vs Level 1)		
Level 2	0.96 (0.72−1.27)	1.09 (0.84−1.40)
Level 3	0.76 (0.47−1.22)	0.97 (0.64−1.47)
Level 4	0.56 (0.31−1.01)	0.72 (0.44−1.20)
Subjective social status in Community (vs Level 1)		
Level 2	1.28 (0.95−1.72)	1.35 (1.04−1.76)
Level 3	1.18 (0.73−1.91)	1.37 (0.91−2.06)
Level 4	1.24 (0.69−2.24)	1.14 (0.68−1.90)
Self−report health condition (EQ−5D) (vs Level 1)		
Level 2	1.01 (0.77−1.32)	1.02 (0.80−1.29)
Level 3	0.87 (0.66−1.14)	0.91 (0.72−1.16)
Level 4	0.97 (0.73−1.27)	1.13 (0.89−1.43)
Chronic disease (vs Yes)		
No	0.56 (0.43−0.71)	0.63 (0.50−0.80)
The history of allergic (vs Yes)		
No	0.52 (0.39−0.71)	0.56 (0.42−0.73)
Unclear	0.72 (0.50−1.05)	0.77 (0.55−1.08)
Smoking status (vs Current smoker)		
Former smoker	1.11 (0.73−1.69)	1.38 (0.95−2.01)
Never smoker	0.93 (0.69−1.24)	1.04 (0.78−1.40)
Drinking status (vs Current drinker)		
Former drinker	1.34 (0.90−2.00)	1.36 (0.95−1.94)
Never drinker	1.11 (0.88−1.39)	0.91 (0.72−1.16)
Physical activity (vs High level)		
Middle level	1.14 (0.91−1.42)	1.16 (0.96−1.41)
Low level	1.29 (1.00−1.67)	1.23 (0.98−1.54)
Public health prevention measures (vs Low level)		
Middle level	0.77 (0.54−1.11)	0.99 (0.71−1.37)
High level	0.65 (0.49−0.86)	0.72 (0.55−0.93)
Awareness of COVID−19 vaccines (vs Level 1)		
Level 2	0.70 (0.51−0.97)	0.85 (0.64−1.13)
Level 3	0.81 (0.62−1.04)	1.03 (0.82−1.29)
Level 4	0.60 (0.46−0.79)	0.81 (0.64−1.03)
Channel of vaccine information (vs We Media)		
Official media	1.01 (0.71−1.44)	1.02 (0.74−1.39)
Others	1.22 (0.98−1.52)	1.25 (1.03−1.51)
Severity (vs Level 1)		
Level 2	0.87 (0.65−1.16)	0.90 (0.70−1.16)
Level 3	1.10 (0.83−1.46)	1.09 (0.86−1.39)
Level 4	1.69 (1.12−2.54)	1.99 (1.39−2.84)
Susceptibility (vs Level 1)		
Level 2	0.52 (0.39−0.69)	0.51 (0.40−0.66)
Level 3	0.46 (0.35−0.61)	0.47 (0.37−0.61)
Level 4	0.50 (0.35−0.72)	0.46 (0.33−0.64)
Benefits (vs Level 1)		
Level 2	0.49 (0.39−0.62)	0.49 (0.40−0.60)
Level 3	0.92 (0.64−1.33)	0.90 (0.65−1.23)
Barriers (vs Level 1)		
Level 2	0.78 (0.53−1.15)	0.95 (0.67−1.35)
Level 3	1.32 (0.86−2.03)	1.69 (1.16−2.48)
Level 4	2.96 (1.96−4.49)	3.35 (2.31−4.87)
Self−efficiency (vs Level 1)		
Level 2	0.82 (0.44−1.50)	0.93 (0.56−1.55)
Level 3	0.27 (0.18−0.43)	0.27 (0.19−0.40)
Trust in medical staff (vs Level 1)		
Level 2	0.74 (0.57−0.96)	0.75 (0.59−0.94)
Level 3	0.60 (0.45−0.81)	0.65 (0.50−0.84)
Level 4	0.90 (0.54−1.51)	1.22 (0.78−1.90)
Trust in developers (vs Level 1)		
Level 2	0.57 (0.44−0.74)	0.54 (0.43−0.67)
Level 3	0.84 (0.58−1.21)	0.69 (0.50−0.96)
Level 4	1.43 (0.88−2.34)	0.81 (0.52−1.26)

Factors Associated With COVID-19 Vaccine Hesitancy in Booster Vaccination

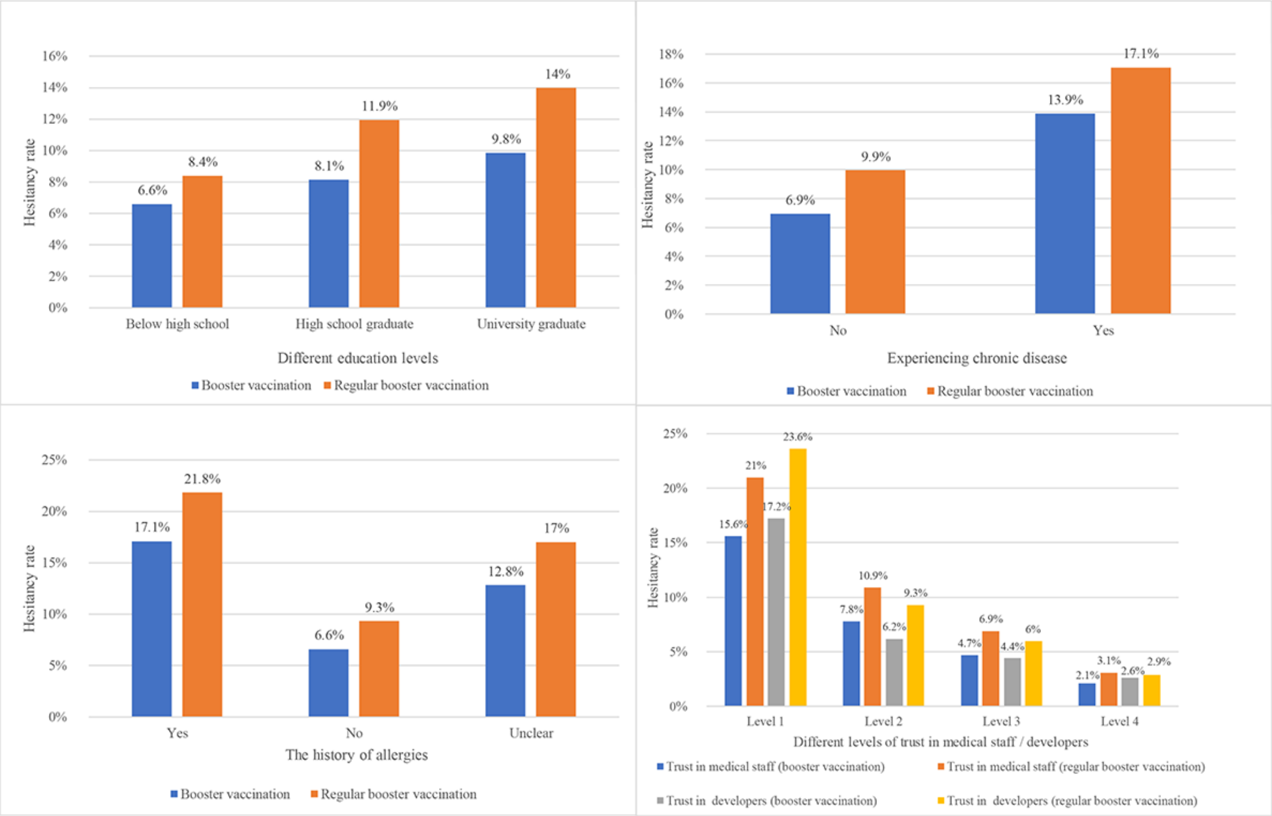
In the binary logistic regression model, subjective social status in China or community, chronic disease, history of allergies, drinking status, physical activity, public health prevention measures, awareness of COVID-19 vaccines, severity, susceptibility, benefits, barriers, self-efficiency, and trust in medical staff or developers were found to be all independently associated with COVID-19 vaccine hesitancy in both booster vaccination and regular booster vaccination (all *P*<.05).

After adjusting for potential confounders, we found that unmarried, experiencing chronic disease, history of allergies, low level of physical activity, lower level of public health prevention measures or susceptibility or benefits or self-efficiency, higher level of severity or barriers, and lower trust in medical staff or developers were all positively associated with COVID-19 vaccine hesitancy in booster vaccination (all

P<.05). Meanwhile, we observed that participants who were female, with higher educational level, higher level of subjective social status in community, experiencing chronic disease, history of allergies, lower level of public health prevention measures or susceptibility or benefits or self-efficiency, higher level of severity or barriers, lower trust in both medical staff and developers were all positively associated with COVID-19 vaccine hesitancy in regular booster vaccination after adjusting for potential confounders (all *P*<.05). Detailed results are all shown in Figure 2 and Multimedia Appendix 3.

In particular, participants with higher levels of education and lower levels of trust in medical staff or developers showed stronger COVID-19 vaccine hesitancy and higher rates of hesitancy at the following regular booster vaccination. The same trend could be observed in participants with a history of vaccination allergy and those with chronic disease (Figure 3).

Figure 3. COVID-19 vaccine hesitancy rates for different influencing factors in both booster and regular booster vaccination.



Dynamic Fluctuations of COVID-19 Vaccine Hesitancy in Booster Vaccination

The COVID-19 vaccine hesitancy rate among Chinese residents shows a trend of raising, which was from 8% (533/6659) to 11.1% (736/6659) between the booster vaccination and regular booster vaccination. In addition, some participants' attitudes toward COVID-19 vaccine booster vaccination were shown to change. There were 5870 and 480 participants who remained vaccinated and hesitant about routine booster vaccination respectively, while there were still 256 participants who would switch from acceptance to hesitancy.

Multimedia Appendix 4 lists 4 attitudinal changes toward COVID-19 booster vaccination, such as acceptance to acceptance, acceptance to hesitancy (ATH), hesitancy to acceptance, and hesitancy to hesitancy (HTH). The results demonstrated that participants who were female, with higher educational level, higher level of barriers, lower level of susceptibility or self-efficiency, and lower trust in developers were positively associated with ATH (all *P*<.05). Participants with higher levels of physical activity and self-efficiency as well as not experiencing chronic disease expressed their willingness from hesitancy to acceptance (all *P*<.001). Furthermore, participants who had higher educational level, lower self-report health condition, having from chronic disease, history of allergies, lower level of public health prevention measures or susceptibility or benefits or self-efficiency, higher level of barriers, lower trust in both medical staff and developers were all positively associated with HTH after adjusting for potential confounders (all *P*<.05). Detailed results are all presented in Figures 4 and 5.

Figure 4. Relevant influences on acceptance to acceptance and acceptance to hesitancy for all participants.

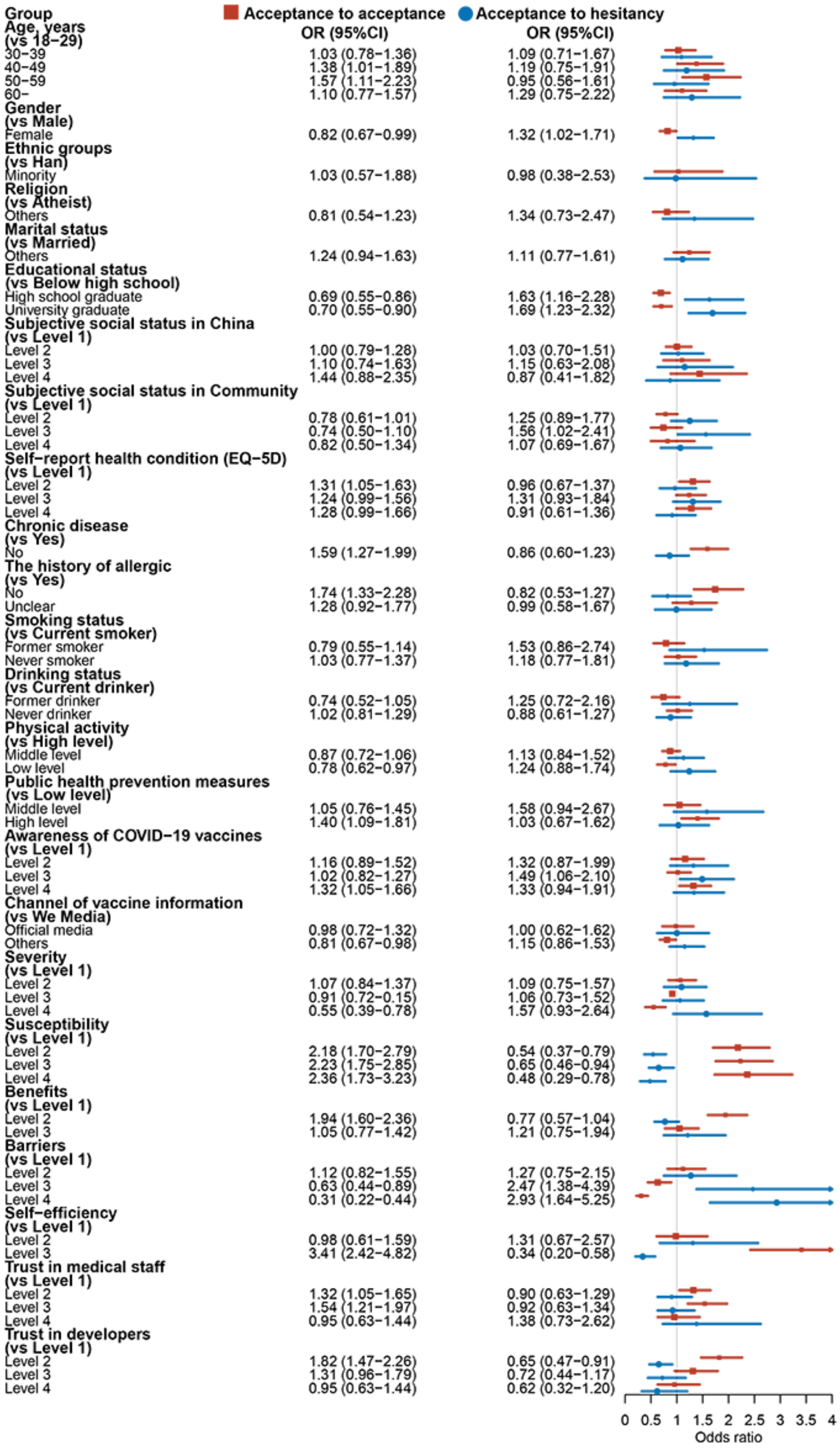
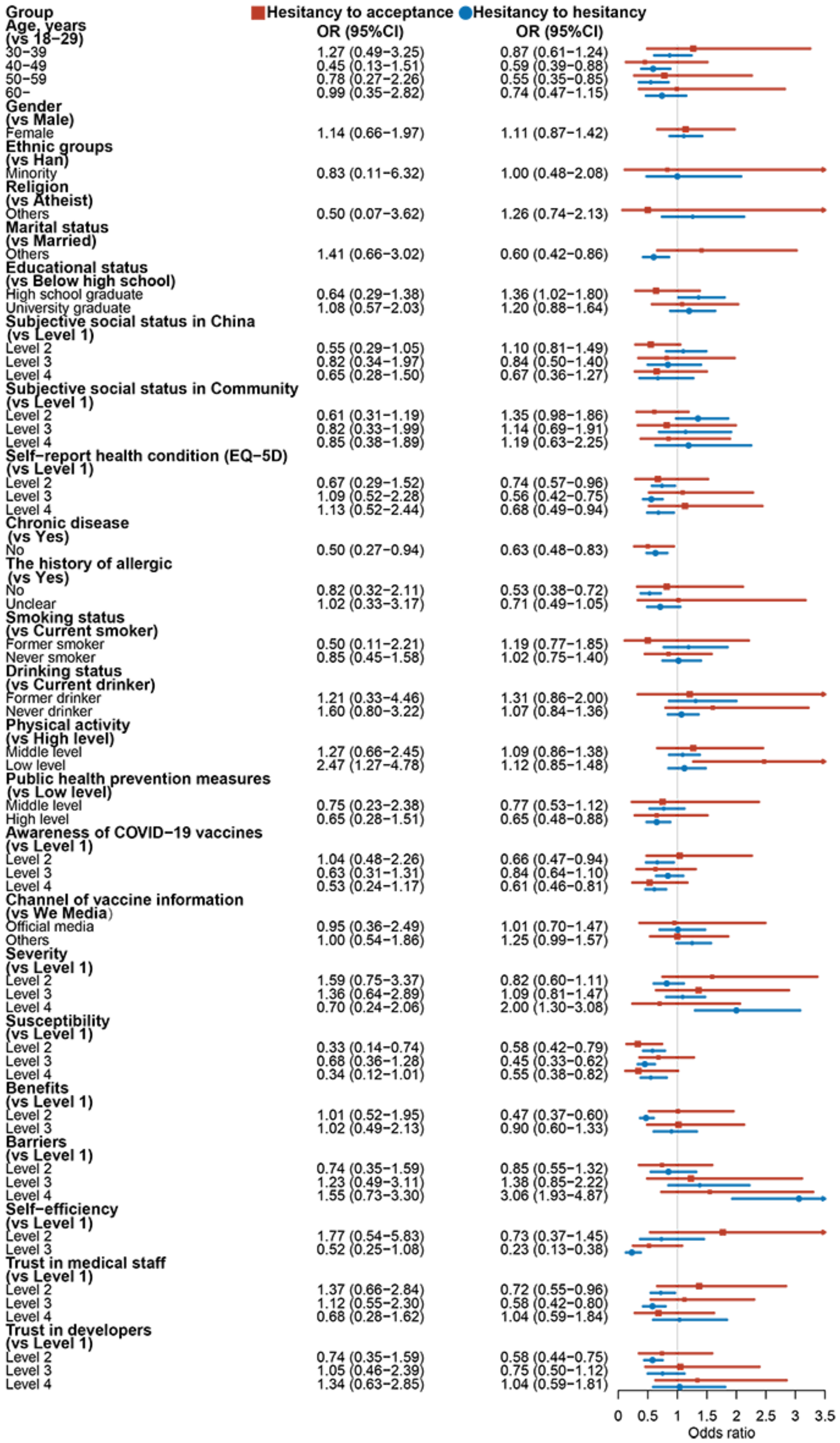


Figure 5. Relevant influences on hesitancy to acceptance and hesitancy to hesitancy for all participants.



Sensitivity Analysis

In sensitivity analyses, the exclusion of participants aged >80 years did not significantly alter the outcome of COVID-19 vaccine hesitancy. Effect estimates for the primary outcome remained similar (Multimedia Appendix 5).

Discussion

Principal Findings

This study focused on the hesitancy rate of COVID-19 vaccine booster in the sample from 4 provinces in mainland China, namely Heilongjiang, Jiangsu, Henan, and Qinghai. The vast majority of participants (6126/6659, 92%) indicated a willingness to receive the COVID-19 booster vaccine, with an additional 8% (533/6659) expressing vaccine hesitancy. The higher vaccination rate and lower vaccine hesitancy rate may be attributed to the following reasons: first, China has continued to strengthen postmarketing surveillance of vaccines, emphasizing their safety and efficacy while continuously tracking the incidence of vaccine-preventable diseases and public acceptance of vaccines [26]. Second, the Chinese government has increased the supply of COVID-19 vaccines and set up reasonably designated medical institutions to carry out booster vaccinations, which has improved the accessibility of vaccination services [31]. Finally, the government and communities have used instant communication media such as television, news, and radio to promote the idea that the overall benefits of COVID-19 vaccination outweigh the risks, and disseminated scientific knowledge about vaccination to reduce people's doubts. At the same time, community staff actively carried out door-to-door campaigns to mobilize residents actively participate in the booster vaccination.

Nevertheless, the hesitancy rate altered dynamically, increasing from 8% (533/6659) to 11% (736/6659) at the next regular COVID-19 booster vaccination. This may be related to the rapid mutation of the COVID-19 virus [10,32] and spot outbreaks in some cities such as Shanghai, Hainan, Zhengzhou, and so forth in mainland China [33]. People are beginning to wonder about the necessity of the COVID-19 booster vaccination and whether it is effective and safe. This result suggests that repeated outbreaks or spot outbreaks are destroying confidence in the COVID-19 vaccine, which will not be conducive to the control of the epidemic. Therefore, it would be relevant to explore the main influences on vaccine hesitancy and its dynamic fluctuations.

Previous studies have indicated that participants with a lower education level expressed a higher level of COVID-19 vaccine hesitancy in a global survey of 17 countries [34], adults in Ontario with less than a bachelor degree were more likely to report unwillingness to COVID-19 vaccination [35], as well as Chinese adults with lower education level endorsed clear vaccine hesitancy response in primary COVID-19 vaccination [26]. However, our findings suggest the opposite view. Those participants with higher levels of education were more reluctant to participate in the regular COVID-19 booster vaccination, showing a clear vaccine hesitancy. Previous studies have demonstrated similar results [36,37]. Likewise, a similar trend was observed in the regular and the booster vaccination attitude conversion (ATH and HTH). The reason for the contrary findings could be attributed to the fact that individuals with a higher level of education may have greater access to information regarding the COVID-19 outbreak and vaccine. This may lead them to believe that the epidemic will come to an end soon and that they will no longer require a booster shot for COVID-19 [5,20]. Alternatively, the highly educated population may have more resources to combat the risk of the COVID-19 epidemic [38,39].

It is worth noting that there was a significant association between experiencing chronic disease, history of previous vaccination allergies, and participants' hesitancy to receive booster vaccinations. This hesitancy was also found in the conversion of attitudes from booster vaccination to regular booster vaccination. Participants with chronic diseases (hypertension, diabetes, etc) expressed clear hesitancy to receive booster vaccinations and were consistently hesitant to receive regular booster vaccinations. The reason for this consequence may be that people with chronic disease were concerned about whether people with immunodeficiency and immunosuppression should be vaccinated [40,41]. They are confused as to whether having a chronic disease can exacerbate adverse reactions to COVID-19 booster vaccination [42,43]. The same trend can be observed in participants with a history of vaccination allergy. Previous experiences with vaccination allergies have made people wary of the COVID-19 vaccine [44]. They are unsure if they will develop allergic injuries after the COVID-19 booster vaccination [45], so hesitancy about the booster is inevitable. Consequently, widespread education of people would be essential to reduce vaccine hesitancy.

In this study, we found that participants with higher levels of epidemic severity, barriers, and lower levels of susceptibility, benefits, and self-efficacy were more hesitant to receive a booster vaccination. The severity of the epidemic influenced people's choice, with higher levels of severity instead promoting vaccine hesitancy, which is inconsistent with previous findings [12,46]. This may be due to people being more aware of self-protection, taking more adequate self-protection measures, going out less, or participating in timely nucleic acid testing when the epidemic is severe [47,48]. Although most people have confidence in the vaccine, this attitude does not automatically translate into vaccine use, as there are a variety of other barriers to vaccination [49,50]. While barriers to vaccination (eg, inconvenience and long queues) affect booster vaccination rates, a reasonable number of booster vaccination sites can facilitate vaccination [38], and the current high rate of COVID-19 vaccination in China can be attributed to a large number of vaccination sites. Therefore, in order to increase the booster vaccination rate, it is important to improve the accessibility of vaccination services. Besides, effective measures can be taken simultaneously, including ensuring the supply of vaccines, increasing the number of vaccination staff, improving the vaccination experience, and providing appointment services to reduce the waiting time for vaccination [26,50,51]. Consistent with the results of Włodarczyk and Ziętalewicz [52], higher self-efficacy and the benefits of COVID-19 vaccination influenced people to take part in vaccination, suggesting that we need to expand the perception that the benefits of vaccination outweigh the disadvantages and guide people positively to take part in vaccination with full self-efficacy.

Mistrust of medical staff and vaccine developers is an important factor influencing hesitancy in receiving COVID-19 booster vaccination. This has been demonstrated in previous studies of vaccine hesitancy [53-55]. Low levels of trust in medical staff and developers can lead to hesitancy in receiving COVID-19 booster vaccinations, even regular booster vaccinations. Both experiences of adverse reactions to vaccinations and even the influence of counterfeit vaccines in society have led to suspicion of medical staff and developers [56,57], which directly influences the COVID-19 booster vaccination. Therefore, it is imperative to strengthen the regulation of vaccine development, production, transportation, and vaccination. Strict adherence to industry guidelines and ethical standards will strongly ensure the standardization, safety, and efficacy of vaccines [16,58]. At the same time, it is necessary to strengthen the training of medical staff to enhance the communication between medical staff and residents, which aims to improve the quality of vaccination services. In addition, enhanced screening of patients with contraindications to COVID-19 vaccination is necessary [24,59], and avoiding or delaying vaccination of contraindicated patients will effectively reduce the occurrence and adverse effects of allergic events.

Strengths and Limitations

This is the first multicenter nationwide household-based survey to assess the hesitancy of COVID-19 booster vaccination, influencing factors, and its dynamic fluctuation trend with the help of a saturation sample, which is highly innovative and forward-looking. Throughout the survey process, dedicated staff were assigned to follow-up the survey sites in each of the 4 centers, and strict quality control was carried out to ensure that the survey data were complete and authentic. However, there are still several limitations of this study. First, this study is a cross-sectional observational study, the causality association of vaccine hesitancy cannot be established, and it needs to be supported by data from the follow-up surveys. Second, it may introduce reporting bias in the collecting data on COVID-19 booster vaccination. Finally, this survey was conducted on a household-based sample with a population of permanent residents, and the findings may not be applicable to the floating population.

Conclusions

The prevalence of COVID-19 vaccine booster hesitancy is not high in mainland China. However, there is a slight increment in hesitancy on regular booster vaccination. To reduce vaccine hesitancy, targeted information guidance for people with higher education levels and chronic diseases would be fruitful. Furthermore, improving the accessibility of booster vaccination and enhancing people’s trust in medical staff and vaccine developers would be highly effective in reducing vaccine hesitancy.

Acknowledgments

This work was funded by the National Social Science Fund of China (21BGL222); the Henan Provincial University Philosophy and Social Science Innovation Talent Support Plan (2023-CXRC-06); and the 2023 Graduate Independent Innovation Project of Zhengzhou University (20230440). The authors would like to thank all participants for their involvement.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

BY, YM, and JW contributed to the conceptualization. XZ, BY, QL, and MM contributed to the methodology. BY and JW contributed to the formal analysis. JW, YM, XZ, QL, and MM contributed to the investigation. YM, MM, XG, and SJ contributed to the data curation. MM and XG contributed to the software. BY, MW, and JG contributed to the funding acquisition. JW and YM contributed to the project administration. MM contributed to writing—original draft. JW, CST, and BY contributed to writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original questionnaire: data collection and explanatory variables.
[DOCX File , 64 KB - publichealth_v9i1e44822_app1.docx]

Multimedia Appendix 2
Analysis of variance between the participants (n=6659) included in the analysis and the total sample (N=8002).
[DOCX File , 17 KB - publichealth_v9i1e44822_app2.docx]

Multimedia Appendix 3
The effect factors of vaccine hesitancy in both the booster and regular booster vaccination (n=6659).
[DOCX File , 46 KB - publichealth_v9i1e44822_app3.docx]

Multimedia Appendix 4
Associations between COVID-19 booster vaccination acceptance, hesitancy transitions, and characteristics of all participants (n=6659).
[DOCX File , 47 KB - publichealth_v9i1e44822_app4.docx]

Multimedia Appendix 5

The sensitivity analyses of factors influence booster and regular booster vaccination (n=6239).

[DOCX File , 57 KB - publichealth_v9i1e44822_app5.docx]

References

1. Wang Z, Muecksch F, Schaefer-Babajew D, Finkin S, Viant C, Gaebler C, et al. Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. Nature 2021;595(7867):426-431 [FREE Full text] [doi: 10.1038/s41586-021-03696-9] [Medline: 34126625]
2. Burki TK. Herd immunity for COVID-19. Lancet Respir Med 2021;9(2):135-136 [FREE Full text] [doi: 10.1016/S2213-2600(20)30555-5] [Medline: 33245861]
3. Arce JSS, Warren SS, Meriggi NF, Scacco A, McMurry N, Voors M, et al. COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. Nat Med 2021;27(8):1385-1394 [FREE Full text] [doi: 10.1038/s41591-021-01454-y] [Medline: 34272499]
4. Adams K, Rhoads JP, Surie D, Gaglani M, Ginde AA, McNeal T, Influenza and other Viruses in the AcutelY ill (IVY) Network. Vaccine effectiveness of primary series and booster doses against covid-19 associated hospital admissions in the United States: living test negative design study. BMJ 2022;379:e072065 [FREE Full text] [doi: 10.1136/bmj-2022-072065] [Medline: 36220174]
5. Gupta PS, Mohareb AM, Valdes C, Price C, Jollife M, Regis C, et al. Mobile health services for COVID-19: counseling, testing, and vaccination for medically underserved populations. Am J Public Health 2022;112(11):1556-1559 [FREE Full text] [doi: 10.2105/AJPH.2022.307021] [Medline: 36223583]
6. WHO Director-General's opening remarks at the 13th meeting of the IHR Emergency Committee on COVID-19 pandemic. WHO. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-13th-meeting-of-the-ihr-emergency-committee-on-covid-19-pandemic---13-october-2022 [accessed 2023-06-22]
7. Kelly JD, Leonard S, Hoggatt KJ, Boscardin WJ, Lum EN, Moss-Vazquez TA, et al. Incidence of severe COVID-19 illness following vaccination and booster with BNT162b2, mRNA-1273, and Ad26.COV2.S vaccines. JAMA 2022;328(14):1427-1437 [FREE Full text] [doi: 10.1001/jama.2022.17985] [Medline: 36156706]
8. Mrak D, Sieghart D, Simader E, Tobudic S, Radner H, Mandl P, et al. Heterologous vector versus homologous mRNA COVID-19 booster vaccination in non-seroconverted immunosuppressed patients: a randomized controlled trial. Nat Commun 2022;13(1):5362 [FREE Full text] [doi: 10.1038/s41467-022-33036-y] [Medline: 36097029]
9. Gilboa M, Regev-Yochay G, Mandelboim M, Indenbaum V, Asraf K, Fluss R, et al. Durability of immune response after COVID-19 booster vaccination and association with COVID-19 omicron infection. JAMA Netw Open 2022;5(9):e2231778 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.31778] [Medline: 36107426]
10. Ioannou GN, Bohnert ASB, O'Hare AM, Boyko EJ, Maciejewski ML, Smith VA, COVID-19 Observational Research Collaboratory (CORC). Effectiveness of mRNA COVID-19 vaccine boosters against infection, hospitalization, and death: a target trial emulation in the omicron (B.1.1.529) variant era. Ann Intern Med 2022;175(12):1693-1706 [FREE Full text] [doi: 10.7326/M22-1856] [Medline: 36215715]
11. Abu-Raddad LJ, Chemaitelly H. The elusive goal of COVID-19 vaccine immunity. Lancet Respir Med 2023;11(2):115-117 [FREE Full text] [doi: 10.1016/S2213-2600(22)00394-0] [Medline: 36216010]
12. Wu J, Ma M, Miao Y, Ye B, Li Q, Tarimo CS, et al. COVID-19 vaccination acceptance among Chinese population and its implications for the pandemic: a national cross-sectional study. Front Public Health 2022;10:796467 [FREE Full text] [doi: 10.3389/fpubh.2022.796467] [Medline: 35211440]
13. Vaccine efficacy, effectiveness and protection. WHO. 2021. URL: https://www.who.int/zh/news-room/feature-stories/detail/vaccine-efficacy-effectiveness-and-protection [accessed 2021-07-14]
14. Berdzuli N, Datta SS. How to tackle inequitable access, vaccine hesitancy, and other barriers to achieve high vaccine uptake. BMJ 2022;377:o1094 [FREE Full text] [doi: 10.1136/bmj.o1094] [Medline: 35487528]
15. Cullen JP, Russ S, Russell HA. Theater for vaccine hesitancy-setting the stage for difficult conversations. JAMA 2022;328(11):1018-1019 [FREE Full text] [doi: 10.1001/jama.2022.14864] [Medline: 36053542]
16. Bussink-Voorend D, Hautvast JLA, Vandeberg L, Visser O, Hulscher MEJL. A systematic literature review to clarify the concept of vaccine hesitancy. Nat Hum Behav 2022;6(12):1634-1648 [FREE Full text] [doi: 10.1038/s41562-022-01431-6] [Medline: 35995837]
17. Draeger E, Bedford HE, Elliman DAC. Should measles vaccination be compulsory? BMJ 2019;365:l2359. [doi: 10.1136/bmj.l2359] [Medline: 31167804]
18. Trakarnvanich T, Ngamvichchukorn T, Phumisantiphong U, Pholtawornkulchai K, Phochanasomboon K, Manomaipiboon A. Immune response after COVID-19 vaccination among patients with chronic kidney disease and kidney transplant. Vaccine 2022;40(45):6499-6511 [FREE Full text] [doi: 10.1016/j.vaccine.2022.09.067] [Medline: 36202639]
19. Benham JL, Atabati O, Oxoby RJ, Mourali M, Shaffer B, Sheikh H, et al. COVID-19 vaccine-related attitudes and beliefs in Canada: national cross-sectional survey and cluster analysis. JMIR Public Health Surveill 2021;7(12):e30424 [FREE Full text] [doi: 10.2196/30424] [Medline: 34779784]

20. Reñosa MDC, Landicho J, Wachinger J, Dalglish SL, Bärnighausen K, Bärnighausen T, et al. Nudging toward vaccination: a systematic review. BMJ Glob Health 2021;6(9):e006237 [FREE Full text] [doi: 10.1136/bmjgh-2021-006237] [Medline: 34593513]
21. Byrne A, Thompson LA, Filipp SL, Ryan K. COVID-19 vaccine perceptions and hesitancy amongst parents of school-aged children during the pediatric vaccine rollout. Vaccine 2022;40(46):6680-6687 [FREE Full text] [doi: 10.1016/j.vaccine.2022.09.090] [Medline: 36220714]
22. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021;5(3):337-348 [FREE Full text] [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
23. Dubé E, MacDonald NE. COVID-19 vaccine hesitancy. Nat Rev Nephrol 2022;18(7):409-410 [FREE Full text] [doi: 10.1038/s41581-022-00571-2] [Medline: 35414006]
24. Gibson C, Schumann C, Neuschel K, McBride JA. COVID-19 vaccination coverage among people experiencing homelessness in a highly vaccinated Midwest County-Dane County, Wisconsin, 2021. J Infect Dis 2022;226(Suppl 3):S335-S339 [FREE Full text] [doi: 10.1093/infdis/jiac303] [Medline: 36208167]
25. Wu Q, Dudley MZ, Chen X, Bai X, Dong K, Zhuang T, et al. Evaluation of the safety profile of COVID-19 vaccines: a rapid review. BMC Med 2021;19(1):173 [FREE Full text] [doi: 10.1186/s12916-021-02059-5] [Medline: 34315454]
26. Wu J, Li Q, Tarimo CS, Wang M, Gu J, Wei W, et al. COVID-19 vaccine hesitancy among Chinese population: a large-scale national study. Front Immunol 2021;12:781161 [FREE Full text] [doi: 10.3389/fimmu.2021.781161] [Medline: 34912346]
27. Yin F, Wu Z, Xia X, Ji M, Wang Y, Hu Z. Unfolding the determinants of COVID-19 vaccine acceptance in China. J Med Internet Res 2021;23(1):e26089 [FREE Full text] [doi: 10.2196/26089] [Medline: 33400682]
28. Dennis A, Cuthbertson DJ, Wootton D, Crooks M, Gabbay M, Eichert N, et al. Multi-organ impairment and long COVID: a 1-year prospective, longitudinal cohort study. J R Soc Med 2023;116(3):97-112 [FREE Full text] [doi: 10.1177/01410768231154703] [Medline: 36787802]
29. Hümmert MW, Bütow F, Tkachenko D, Ayzenberg I, Pakeerathan T, Hellwig K, Neuromyelitis Optica Study Group (NEMOS). Effects of the COVID-19 pandemic on patients with NMO spectrum disorders and MOG-antibody-associated diseases: COPANMO(G)-study. Neurol Neuroimmunol Neuroinflamm 2023;10(2):e200082 [FREE Full text] [doi: 10.1212/NXI.0000000000200082] [Medline: 36693760]
30. Bahmer T, Borzikowsky C, Lieb W, Horn A, Krist L, Fricke J, NAPKON study group. Severity, predictors and clinical correlates of Post-COVID syndrome (PCS) in Germany: a prospective, multi-centre, population-based cohort study. EClinicalMedicine 2022;51:101549 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101549] [Medline: 35875815]
31. Ma L, Han X, Ma Y, Yang Y, Xu Y, Liu D, et al. Decreased influenza vaccination coverage among Chinese healthcare workers during the COVID-19 pandemic. Infect Dis Poverty 2022;11(1):105 [FREE Full text] [doi: 10.1186/s40249-022-01029-0] [Medline: 36209258]
32. McMenamin ME, Nealon J, Lin Y, Wong JY, Cheung JK, Lau EHY, et al. Vaccine effectiveness of one, two, and three doses of BNT162b2 and CoronaVac against COVID-19 in Hong Kong: a population-based observational study. Lancet Infect Dis 2022;22(10):1435-1443 [FREE Full text] [doi: 10.1016/S1473-3099(22)00345-0] [Medline: 35850128]
33. Update on the outbreak of novel coronavirus pneumonia as of 24 hours on 18 October. NHC. URL: http://www.nhc.gov.cn/xcs/yqtb/202210/24faf969a6db47ee88573499c8bf5eb3.shtml [accessed 2023-06-22]
34. Wong LP, Alias H, Danaee M, Ahmed J, Lachyan A, Cai CZ, et al. COVID-19 vaccination intention and vaccine characteristics influencing vaccination acceptance: a global survey of 17 countries. Infect Dis Poverty 2021;10(1):122 [FREE Full text] [doi: 10.1186/s40249-021-00900-w] [Medline: 34620243]
35. Syan SK, Gohari MR, Levitt EE, Belisario K, Gillard J, DeJesus J, et al. COVID-19 vaccine perceptions and differences by sex, age, and education in 1,367 community adults in Ontario. Front Public Health 2021;9:719665 [FREE Full text] [doi: 10.3389/fpubh.2021.719665] [Medline: 34631647]
36. Lazarus JV, Wyka K, White TM, Picchio CA, Rabin K, Ratzan SC, et al. Revisiting COVID-19 vaccine hesitancy around the world using data from 23 countries in 2021. Nat Commun 2022;13(1):3801 [FREE Full text] [doi: 10.1038/s41467-022-31441-x] [Medline: 35778396]
37. Huang X, Yu M, Fu G, Lan G, Li L, Yang J, et al. Willingness to receive COVID-19 vaccination among people living with HIV and AIDS in China: nationwide cross-sectional online survey. JMIR Public Health Surveill 2021;7(10):e31125 [FREE Full text] [doi: 10.2196/31125] [Medline: 34543223]
38. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrachi M, Zigron A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol 2020;35(8):775-779 [FREE Full text] [doi: 10.1007/s10654-020-00671-y] [Medline: 32785815]
39. Tsai R, Hervey J, Hoffman K, Wood J, Johnson J, Deighton D, et al. COVID-19 vaccine hesitancy and acceptance among individuals with cancer, autoimmune diseases, or other serious comorbid conditions: cross-sectional, internet-based survey. JMIR Public Health Surveill 2022;8(1):e29872 [FREE Full text] [doi: 10.2196/29872] [Medline: 34709184]
40. Pranata R, Henrina J, Raffaello WM, Lawrensia S, Huang I. Diabetes and COVID-19: the past, the present, and the future. Metabolism 2021;121:154814 [FREE Full text] [doi: 10.1016/j.metabol.2021.154814] [Medline: 34119537]

41. Tang LV, Hu Y. Hemophagocytic lymphohistiocytosis after COVID-19 vaccination. J Hematol Oncol 2021;14(1):87 [FREE Full text] [doi: 10.1186/s13045-021-01100-7] [Medline: 34088334]
42. Watanabe M, Balena A, Tuccinardi D, Tozzi R, Risi R, Masi D, et al. Central obesity, smoking habit, and hypertension are associated with lower antibody titres in response to COVID-19 mRNA vaccine. Diabetes Metab Res Rev 2022;38(1):e3465 [FREE Full text] [doi: 10.1002/dmrr.3465] [Medline: 33955644]
43. Schrezenmeier E, Bergfeld L, Hillus D, Lippert JD, Weber U, Tober-Lau P, et al. Immunogenicity of COVID-19 tozinameran vaccination in patients on chronic dialysis. Front Immunol 2021;12:690698 [FREE Full text] [doi: 10.3389/fimmu.2021.690698] [Medline: 34276681]
44. Cabanillas B, Novak N. Allergy to COVID-19 vaccines: a current update. Allergol Int 2021;70(3):313-318 [FREE Full text] [doi: 10.1016/j.alit.2021.04.003] [Medline: 33962863]
45. Laake I, Skodvin SN, Blix K, Caspersen IH, Gjessing HK, Juvet LK, et al. Effectiveness of mRNA booster vaccination against mild, moderate, and severe COVID-19 caused by the omicron variant in a large, population-based, Norwegian cohort. J Infect Dis 2022;226(11):1924-1933 [FREE Full text] [doi: 10.1093/infdis/jiac419] [Medline: 36259543]
46. Wester CT, Scheel-Hincke LL, Bovil T, Andersen-Ranberg K, Ahrenfeldt LJ, Hvidt NC. Prayer frequency and COVID-19 vaccine hesitancy among older adults in Europe. Vaccine 2022;40(44):6383-6390 [FREE Full text] [doi: 10.1016/j.vaccine.2022.09.044] [Medline: 36182618]
47. Bucyibaruta G, Blangiardo M, Konstantinoudis G. Community-level characteristics of COVID-19 vaccine hesitancy in England: a nationwide cross-sectional study. Eur J Epidemiol 2022;37(10):1071-1081 [FREE Full text] [doi: 10.1007/s10654-022-00905-1] [Medline: 36121531]
48. Kobayashi Y, Howell C, Heinrich T, Motta M. Investigating how historical legacies of militarized violence can motivate COVID-19 vaccine hesitancy: evidence from global dyadic survey. Soc Sci Med 2022;311:115346 [FREE Full text] [doi: 10.1016/j.socscimed.2022.115346] [Medline: 36108562]
49. Rich-Edwards JW, Rocheleau CM, Ding M, Hankins JA, Katuska LM, Kumph X, et al. COVID-19 vaccine uptake and factors affecting hesitancy among US nurses, March-June 2021. Am J Public Health 2022;112(11):1620-1629. [doi: 10.2105/AJPH.2022.307050] [Medline: 36223573]
50. Gbenonsi GY, Labat A, Oleffe A, Jidovsteff B, Servais O, Vermeulen N, et al. Factors associated with COVID-19 vaccination among individuals with vaccine hesitancy in French-speaking Belgium. JAMA Netw Open 2022;5(9):e2234433 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.34433] [Medline: 36112379]
51. Caserotti M, Girardi P, Rubaltelli E, Tasso A, Lotto L, Gavaruzzi T. Associations of COVID-19 risk perception with vaccine hesitancy over time for Italian residents. Soc Sci Med 2021;272:113688 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113688] [Medline: 33485215]
52. Włodarczyk D, Ziętalewicz U. Medics as a positive deviant in influenza vaccination: the role of vaccine beliefs, self-efficacy and contextual variables. Vaccines (Basel) 2022;10(5):723 [FREE Full text] [doi: 10.3390/vaccines10050723] [Medline: 35632479]
53. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 2021;6(4):e210-e221 [FREE Full text] [doi: 10.1016/S2468-2667(21)00012-8] [Medline: 33556325]
54. Padamsee TJ, Bond RM, Dixon GN, Hovick SR, Na K, Nisbet EC, et al. Changes in COVID-19 vaccine hesitancy among black and white individuals in the US. JAMA Netw Open 2022;5(1):e2144470 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.44470] [Medline: 35061038]
55. Salali GD, Uysal MS. COVID-19 vaccine hesitancy is associated with beliefs on the origin of the novel coronavirus in the UK and Turkey. Psychol Med 2020:1-3 [FREE Full text] [doi: 10.1017/S0033291720004067] [Medline: 33070804]
56. Allington D, McAndrew S, Moxham-Hall V, Duffy B. Coronavirus conspiracy suspicions, general vaccine attitudes, trust and coronavirus information source as predictors of vaccine hesitancy among UK residents during the COVID-19 pandemic. Psychol Med 2023;53(1):236-247 [FREE Full text] [doi: 10.1017/S0033291721001434] [Medline: 33843509]
57. Barello S, Nania T, Dellafiore F, Graffigna G, Caruso R. 'Vaccine hesitancy' among university students in Italy during the COVID-19 pandemic. Eur J Epidemiol 2020;35(8):781-783 [FREE Full text] [doi: 10.1007/s10654-020-00670-z] [Medline: 32761440]
58. Garcia-Grossman IR, Gransee L, Williams B. Strategies for addressing vaccine hesitancy within California state prisons in 2021 and beyond. Am J Public Health 2022;112(11):1543-1545. [doi: 10.2105/AJPH.2022.307042] [Medline: 36075008]
59. Agrawal U, Bedston S, McCowan C, Oke J, Patterson L, Robertson C, et al. Severe COVID-19 outcomes after full vaccination of primary schedule and initial boosters: pooled analysis of national prospective cohort studies of 30 million individuals in England, Northern Ireland, Scotland, and Wales. Lancet 2022;400(10360):1305-1320 [FREE Full text] [doi: 10.1016/S0140-6736(22)01656-7] [Medline: 36244382]

Abbreviations

ATH: acceptance to hesitancy
HTH: hesitancy to hesitancy

WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 05.12.22; peer-reviewed by Y Lai, M Fahimi; comments to author 17.01.23; revised version received 18.02.23; accepted 16.06.23; published 01.08.23.

Please cite as:
Wu J, Ma M, Li Q, Guo X, Tarimo CS, Jia S, Zhou X, Wang M, Gu J, Miao Y, Ye B
Dynamic Trends and Underlying Factors of COVID-19 Vaccine Booster Hesitancy in Adults: Cross-Sectional Observational Study
JMIR Public Health Surveill 2023;9:e44822
URL: https://publichealth.jmir.org/2023/1/e44822
doi:10.2196/44822
PMID:37526963

Original Paper

Immune-Boosting Effect of the COVID-19 Vaccine: Real-World Bidirectional Cohort Study

Ming Liu[1*], MM; Tianshuo Zhao[2,3,4,5*], MD; Qiuyue Mu[1], MPH; Ruizhi Zhang[1], MPH; Chunting Liu[1], MM; Fei Xu[1], MPH; Luxiang Liang[1], AD; Linglu Zhao[1], MM; Suye Zhao[1], BMed; Xianming Cai[2,3,4,5], MM; Mingting Wang[2,3,4,5], MPH; Ninghua Huang[2,3,4,5], MD; Tian Feng[1], MPH; Shiguang Lei[1], BMed; Guanghong Yang[1], MD; Fuqiang Cui[2,3,4,5], MD, MPH, MPM, PhD

[1]Guizhou Center for Disease Control and Prevention, Guiyang, China
[2]Department of Laboratorial Science and Technology, School of Public Health, Peking University, Beijing, China
[3]Vaccine Research Center, School of Public Health, Peking University, Beijing, China
[4]Center for Infectious Diseases and Policy Research & Global Health and Infectious Diseases Group, Peking University, Beijing, China
[5]Key Laboratory of Epidemiology of Major Diseases, Peking University, Ministry of Education, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Guanghong Yang, MD
Guizhou Center for Disease Control and Prevention
73 Bageyan Road
Yunyan District, Guizhou
Guiyang, 550004
China
Phone: 86 0851 86828805
Email: ghyang_gzmu@outlook.com

Abstract

Background: As the SARS-CoV-2 attenuates and antibodies from the COVID-19 vaccine decline, long-term attention should be paid to the durability of primary booster administration and the preventive effect of the second or multiple booster doses of the COVID-19 vaccine.

Objective: This study aimed to explore the durability of primary booster administration and the preventive effect of second or multiple booster doses of the COVID-19 vaccine.

Methods: We established a bidirectional cohort in Guizhou Province, China. Eligible participants who had received the primary booster dose were enrolled for blood sample collection and administration of the second booster dose. A retrospective cohort for the time of administration was constructed to evaluate antibody attenuation 6-12 months after the primary booster dose, while a prospective cohort on the vaccine effect of the second booster dose was constructed for 4 months after the second administration.

Results: Between September 21, 2022, and January 30, 2023, a total of 327 participants were included in the final statistical analysis plan. The retrospective cohort revealed that approximately 6-12 months after receiving the primary booster, immunoglobulin G (IgG) slowly declined with time, while immunoglobulin A (IgA) remained almost constant. The prospective cohort showed that 28 days after receiving the second booster, the antibody levels were significantly improved. Higher levels of IgG and IgA were associated with better protection against COVID-19 infection for vaccine recipients. Regarding the protection of antibody levels against post–COVID-19 symptoms, the increase of the IgG had a protective effect on brain fog and sleep quality, while IgA had a protective effect on shortness of breath, brain fog, impaired coordination, and physical pain.

Conclusions: The IgG and IgA produced by the second booster dose of COVID-19 vaccines can protect against SARS-CoV-2 infection and may alleviate some post–COVID-19 symptoms. Further data and studies on secondary booster administration are required to confirm these conclusions.

(JMIR Public Health Surveill 2023;9:e47272) doi:10.2196/47272

KEYWORDS

bidirectional cohort study; booster administration; COVID-19 vaccine; real-world study; SARS-CoV-2; vaccine efficacy; COVID-19

Introduction

Since the end of 2019, COVID-19 has been the cause of a global pandemic, placing a heavy burden on the global public health system [1]. With the widespread and continuous evolution of SARS-CoV-2, many variants of concern (VOCs), such as Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529), have emerged globally and have led to several infection waves [2-5]. At present, the Omicron variant, which has a higher transmissibility and immune escape ability, is the dominant variant in the world [6]. Previous studies have shown that the Omicron variant not only has resistance to serum antibodies of convalescent patients but also has certain resistance to the serum of individuals who have been fully vaccinated against COVID-19 [7-12]. Therefore, Omicron poses a serious threat to the control of the COVID-19 pandemic and disease treatment.

China has administered 3.491 billion doses of the COVID-19 vaccine. The coverage rate of the first dose and second dose for the entire population reached 92.9% and 90.6%, respectively [13], and more than 771 million booster injections have been administered [14]. In November 2022, China adjusted and optimized the prevention and control measures for COVID-19. The local epidemic quickly climaxed, which led to a depletion of medical resources. Febrifuge, antitussive, and other COVID-19–related drugs could not meet the exponential increase in the number of patients in the short term, and the more serious concern was overwhelming the availability of hospital beds [15,16]. Therefore, long-term attention should be paid to the preventive effect and clinical value of the second or multiple booster doses of the COVID-19 vaccine.

In China, the most commonly used vaccines for both primary and booster immunization against COVID-19 are inactivated vaccines produced by the China National Biotec Group and Sinovac Biotech. Inactivated vaccines are prepared by cultivating SARS-CoV-2 in vitro to render the virus noninfectious while preserving its antigenicity. Although homologous boosting is generally considered a standard practice, heterologous regimens have been proposed as a COVID-19 vaccine strategy to elicit stronger and broader, or longer-lasting, immunity [17,18]. A recombinant COVID-19 vaccine using adenovirus type 5 as a vector for inhalation was developed by CanSino Biologics (inhalant Ad5-nCoV). Ad5-nCoV inhalation involves the recombination of the spike glycoprotein gene of SARS-CoV-2 into the replication-deficient human type 5 adenovirus gene, which induces an immune response in the body. This inhalant is easy to administer and can stimulate mucosal immunity. In a previous phase 1 trial, Ad5-nCoV inhalation was found to be well tolerated. Further, compared with intramuscular vaccination, aerosol vaccination could trigger a higher ratio of neutralizing antibodies to total antibodies [19].

Few real-world studies have demonstrated the effect of a fourth dose of heterologous booster on Omicron, especially using inhalation vaccines. Here, we aimed to reveal the immunogenicity and persistence of the primary booster dose and the real-world immune effect of the secondary inhalation booster to assess immunogenicity and persistence and prevent sequelae of booster administration in the real world.

Methods

Study Design

Overview

A bidirectional cohort was established to investigate the efficacy of booster administration between September 21, 2022, and January 30, 2023, in Guiyang City, Guizhou Province, China. The cohort was retrospectively tracked to determine the effect and durability of primary booster administration and prospectively followed up to assess the immunogenicity and real-world protective effect of secondary booster heterologous immunization, with the time of receiving the second booster as the node.

Retrospective Study

Blood samples were collected from individuals for antibody testing at the time of enrollment. A questionnaire covering basic characteristics and immunization programs was required to evaluate the durability and effectiveness of the third booster dose.

Secondary Booster Administration

All enrolled individuals received the inhalant, Ad5-nCoV, as part of a secondary booster immunization program (the fourth dose).

Prospective Study

About 4 weeks (21-35 days) after receiving the second booster dose, blood samples were collected for antibody testing. After about 16 weeks (84-140 days), information on the infection and sequelae of COVID-19 was collected from participants through follow-up phone calls.

Participants

Participants were recruited by the Guizhou Center for Disease Control and Prevention. These individuals were aged 18 years or older and had received 3 doses of the COVID-19 vaccine before 6 months or above. The main exclusion criteria were individuals with a history of clinically or laboratory-confirmed COVID-19 or SARS-CoV-2 infection within the first 6 months of enrollment, a history of vaccination (any administration, including COVID-19 baseline or booster dose) within the first 6 months of enrollment, or an allergy to any component of the vaccine.

Procedures

Individuals from Guanshanhu District, Qingzhen City, and Baiyun District of Guiyang City were recruited for this study. All individuals completed the basic and primary booster immunization procedures with the inactivated vaccine from the China National Biotec Group or Sinovac Biotech.

Eligible participants received 1 dose of inhalant Ad5-nCoV (0.1 mL per dose) through a specific atomization device. Venous blood samples (5 mL) were collected before inhalation and 28 days after inhalation to detect immunoglobulin G (IgG) and immunoglobulin A (IgA) antibodies against SARS-CoV-2 in serum. Antibody detection was performed by the receptor-binding domain antibody test kit produced by Vazyme Biotech Co Ltd. The kit detects receptor-binding domains IgA and IgG antibodies against SARS-CoV-2 that are produced during incubation through an indirect enzyme-linked immunosorbent assay (ELISA). After processing and color development, the absorbance of the sample was measured at a wavelength of 450 nm. The absorbance of the sample was positively correlated with the antibody titers.

Survey Tool

Telephone follow-up was conducted with the participants to assess their status and the timing of contracting SARS-CoV-2 after inhalation and evaluate the persistent symptoms of post–COVID-19 using a scale. The scale comprised 49 items and was used to assess the severity of the post–COVID-19 impact using 8 indicators: fatigue, shortness of breath, brain fog, impaired coordination, physical pain, impaired sleep quality, depression, and impaired quality of life. These indicators were selected based on the common symptoms of post–COVID-19 condition (PCC), that is, a set of signs and symptoms that emerge during or after an infection consistent with COVID-19 and are not explained by an alternative diagnosis [20].

Each item was rated as “never occurred,” “slightly affected,” “moderately affected,” and “severely affected,” with scores of 0, 1, 2, and 3, respectively. The Cronbach α values of the 8 indicators ranged from .79 to .94, indicating acceptable reliability [21]. The detailed questionnaire and Cronbach α values are provided in Multimedia Appendix 1.

Statistical Analyses

The baseline characteristics are presented as means (SDs) for continuous variables and percentages for categorical variables. Missing values were treated and reported in all analyses.

The attenuation curves of the antibody and time were fitted through locally weighted scatterplot smoothing (LOWESS), a nonparametric method used in the analysis of local regression. The sample was divided into short intervals, and weighted polynomial fitting to the sample in each interval was conducted. Linear regressions were constructed to fit the curve stratified according to the preceding immune program.

Comparisons of geometric mean titers (GMTs) and geometric mean increases (GMIs) between the groups were performed using logarithmic conversion values. The differences in antibodies before and after administration were compared using a paired Student 2-tailed *t* test, and linear regression was used to compare GMIs among different groups. For positive seroconversion, the antibodies after administration should increase by 4-fold or more, according to the literature [19,22]. A logistic regression was used to compare the seroconversion rate. On the basis of the marginal forecast rates of each category estimated by the regression, we used the weighted average of the standard population to calculate the direct standardized seroconversion rate and GMI. All test criteria to confirm the hypothesis were bilateral, with a significance level of .05. The adjusted α was reduced to .017 when pairwise comparisons were made between the 3 groups.

Kaplan-Meier analysis was used to plot the uninfected curves and cumulative hazard curves of different antibody levels, and the log-rank test was used to compare the difference in infection time among individuals with different levels. A multivariate Cox proportional hazard regression model was used to adjust for the effects of confounding factors on the results. Finally, a linear regression model was constructed to analyze the correlation between different sequelae scores and antibodies.

All statistical analyses were performed using R (version 4.2.0; R Development Core Team) and Stata (version 17.0; Stata Corporation).

Ethics Approval

The protocol was approved by the institutional review board of the Guizhou Center for Disease Control and Prevention (approval number Q2023-03) and was performed in accordance with the Declaration of Helsinki and the Good Clinical Practice guidelines. All participants provided written informed consent before enrollment.

Results

Basic Characteristics of the Participants

A total of 327 participants who completed the vaccination and blood sampling were enrolled in the final statistical analysis. The specific participant entry and exit processes are outlined in Multimedia Appendix 1. A total of 234 female and 93 male participants were enrolled, with a mean age of 39.4 (SD 9.5) years. All participants received 3 doses of the COVID-19 vaccine before the survey, including 2 doses for basic immunization and 1 dose for booster immunization. Of the 327 participants, 166 received the BBIBP-CorV inactivated vaccine (SinoBio Pharmaceutical Ltd) and 161 received the CoronaVac inactivated vaccine produced by Sinovac Ltd. Among them, 233 participants worked in hospitals, 84 worked in nonhospital institutions, and the remaining 10 were reluctant to report their occupations and workplaces. Detailed information is provided in Table 1.

Table 1. The characteristics of participants (N=327).

Characteristics	Participants
Age (years), mean (SD)	39.4 (9.5)
Age (years), n (%)	
<30	63 (19.3)
30-40	94 (28.7)
40-50	118 (36.1)
≥50	52 (15.9)
Sex, n (%)	
Male	93 (28.4)
Female	234 (71.6)
Ethnicity, n (%)	
Han	279 (85.3)
Minority	48 (14.5)
Workplace[a], n (%)	
Hospital	233 (73.5)
Nonhospital	84 (26.5)
Primary booster vaccine	
BBIBP-CorV	166 (50.8)
CoronaVac	161 (49.2)

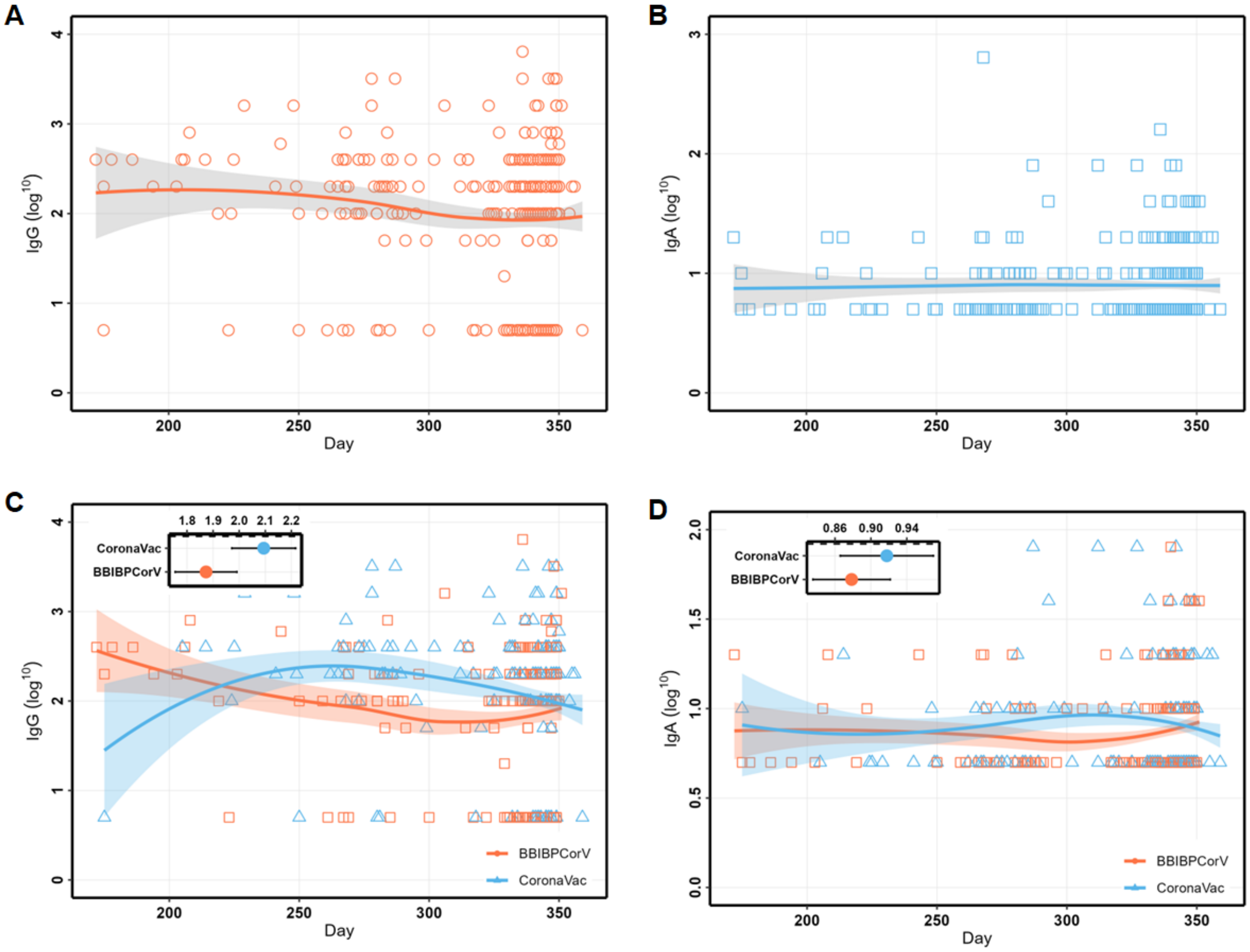
[a] n=317; 10 subjects were not willing to report their occupation or workplace.

Long-Term Durability of the Antibodies After Primary-Booster Administration (Third Dose)

Before receiving the second booster (fourth dose), we measured the antibody levels of the participants. Approximately 6-12 months after receiving the booster, the GMT of the IgG antibody was 95.9 (95% CI 78.8-116.7), while that of the IgA antibody was 7.9 (95% CI 7.3-8.5). The time interval between the last booster immunization and antibody detection was used as the independent variable, while the antibody level was used as the dependent variable for LOWESS segmented curve fitting. The fitting results showed that the IgG antibody level slowly declined with time after the primary booster, while the IgA antibody level remained almost constant (GMT 7.9; Figure 1A and B).

We proceeded to stratify the results based on the different immunization programs and perform linear fitting. Figure 1C shows that the declining trend for the IgG antibody level of participants administered the CoronaVac vaccine and those administered the BBIBP-CorV vaccine was consistent after 6 months of immunization, with the IgG antibody titer of CoronaVac (GMT 124.2; 95% CI 93.8-164.3) vaccine recipients being slightly higher than that of the BBIBP-CorV vaccine recipients (GMT 74.7; 95% CI 56.9-97.9). Figure 1D shows that the levels and trends of the 2 are almost identical.

Figure 1. The locally weighted scatterplot smoothing (LOWESS) curves of IgG and IgA over time during 6-12 months after primary booster administration. (A) The curves of IgG over time. (B) The curves of IgA over time. (C) The changing curves of IgG stratified by a booster vaccine (third dose). (D) The changing curves of IgA stratified by a booster vaccine (third dose). IgA: immunoglobulin A; IgG: immunoglobulin G.



Immunogenicity of the Second Booster (Fourth Dose)

At 28 days after receiving the inhalant vaccine as the second booster, the GMT of the IgG antibody was 5066.5 (95% CI 4418.1-5810.1), while that of the IgA antibody was 108.6 (95% CI 95.4-123.5). The GMI of IgG was 52.8 (95% CI 42.6-65.6), and the seroconversion rate reached 94.5% (95% CI 92-97). The GMI of IgA was 13.7 (95% CI 12-15.7), and the antibody seroconversion rate was 89.3% (95% CI 85.9-92.7).

The antibody levels of individuals with different sociodemographic characteristics and prevaccination programs showed significant improvement after vaccination. The line graphs of the pre and postvaccination GMTs for the IgG and IgA antibodies of different groups are shown in Figure 2.

We explored the crude and adjusted changes in antibody (GMI and seroconversion rate) among different groups based on age, sex, ethnicity, workplace, and type of primary booster vaccine, as detailed in Table 2. There was no difference in postvaccination GMI or seroconversion of IgG antibodies among the different demographic groups. However, for participants who received BBIBP-CorV (adjusted rate 97.4%; 95% CI 94.9-99.9) as their initial booster, the seroconversion rate of the IgG antibody was higher than that of those who received CoronaVac (adjusted rate 91.4%; 95% CI 87.2-95.7) after the secondary booster immunization.

Figure 2. The differences of the pre- and postvaccination GMTs for the IgG and IgA antibodies of different groups. (A-E) GMTs for the IgG. (A) Group by age. (B) Group by sex. (C) Group by ethnicity. (D) Group by type of primary booster vaccine. (E) Group by workplace. (F-J) GMTs for the IgA. (F) Group by age. (G) Group by sex. (H) Group by ethnicity. (I) Group by type of primary booster vaccine. (J) Group by workplace. GMT: geometric mean titer; IgA: immunoglobulin A; IgG: immunoglobulin G.

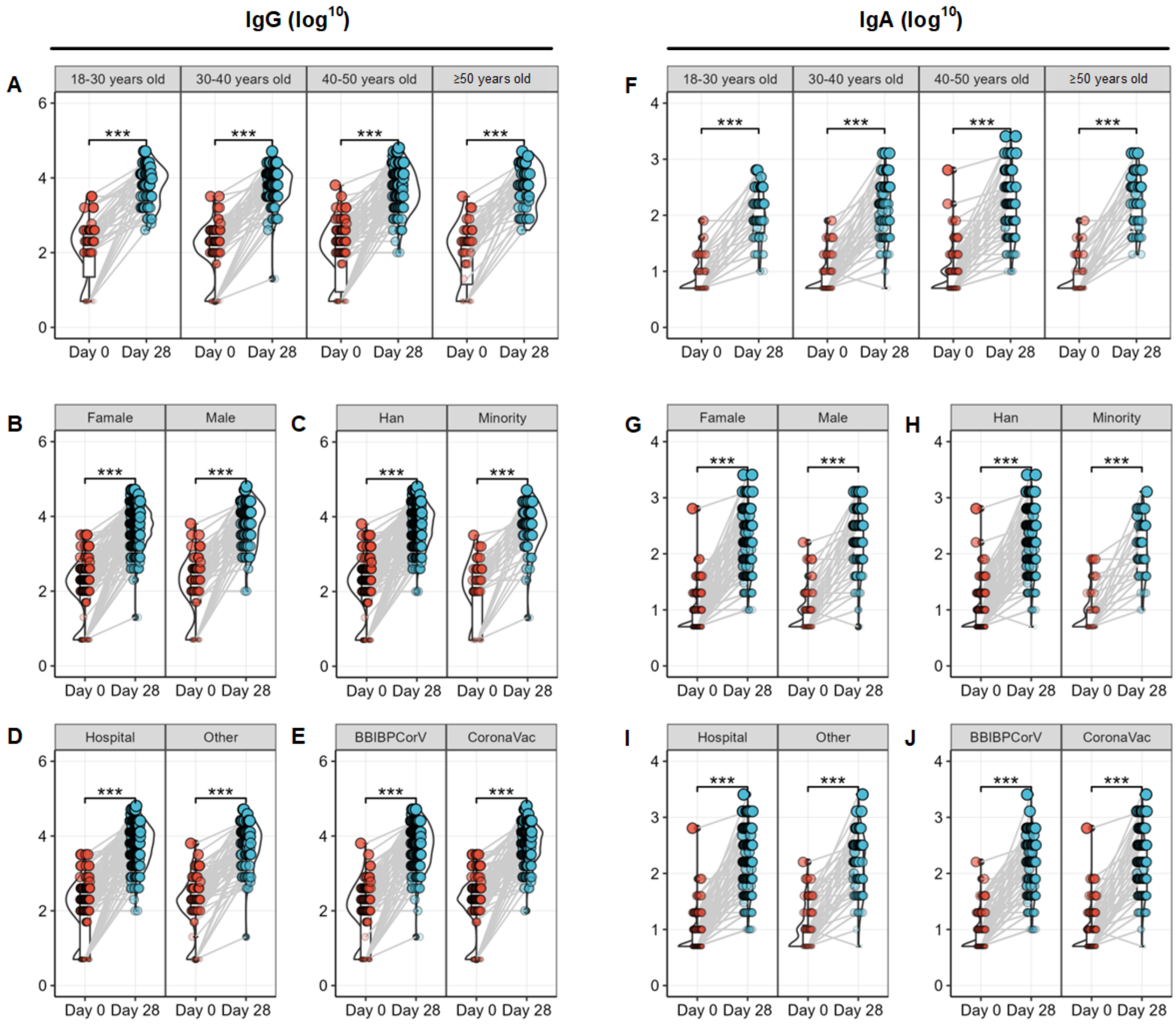


Table 2. The distribution and difference of geometric mean increase and positive seroconversion rate of immunoglobulin G and immunoglobulin A antibodies among groups.

Variable	Geometric mean increase, mean (95% CI)			Seroconversion rate, proportion (95% CI)		
	Crude	Adjusted	*P* value	Crude	Adjusted	*P* value
Immunoglobulin G						
Overall	52.8 (42.6-65.6)	N/A[a]	N/A	94.5 (92-97)	N/A	N/A
Age (years)						
<30	56.6 (33.9-94.7)	53.6 (32.2-89.1)	—[b]	93.7 (87.6-99.7)	92.3 (85-99.7)	—
30-40	49.1 (35.1-68.6)	52.2 (34.4-79.3)	.94	97.9 (94.9-100.8)	97.7 (94.5-100.9)	.16
40-50	48.4 (32.2-72.6)	49.9 (34.3-72.5)	.83	91.5 (86.5-96.6)	92.3 (87.6-96.9)	.99
≥50	67.7 (41-111.9)	69.8 (39.8-122.7)	.49	96.2 (90.9-101.5)	95.9 (90.5-101.4)	.45
Sex						
Male	52 (34.6-78.1)	57.6 (37.4-88.8)	—	92.5 (87.1-97.9)	93.3 (88.3-98.2)	—
Female	53.2 (41.2-68.6)	52.6 (40.3-68.6)	.73	95.3 (92.6-98)	94.8 (91.8-97.8)	.59
Ethnicity						
Han	52.8 (41.9-66.5)	53.5 (42.1-68)	—	94.3 (91.5-97)	94.2 (91.5-96.9)	—
Minority	52.9 (28.6-97.9)	56.8 (31.6-102.1)	.85	95.8 (90.1-101.6)	95 (88.3-101.7)	.84
Workplace						
Hospital	61.3 (47-80.1)	61.5 (47.4-79.8)	—	94.4 (91.5-97.4)	94.4 (91.5-97.3)	—
Nonhospital	37.9 (26-55.1)	37.6 (24.2-58.4)	.06	94 (88.9-99.2)	94.1 (89-99.2)	.92
Booster vaccine						
BBIBP-CorV	64.1 (48-85.6)	65.7 (48-89.9)	—	97.6 (95.2-99.9)	97.4 (94.9-99.9)	—
CoronaVac	43.3 (31.4-59.6)	44.3 (32.4-60.7)	.09	91.3 (86.9-95.7)	91.4 (87.2-95.7)	.03
Immunoglobulin A						
Overall	13.7 (12-15.7)	N/A	N/A	89.3 (85.9-92.7)	N/A	N/A
Age (years)						
<30	12.9 (9.8-16.8)	13.2 (9.7-18)	—	88.9 (81-96.7)	89.2 (81.5-96.9)	—
30-40	14.1 (11.2-17.7)	14.8 (11.5-19.1)	.58	92.6 (87.2-97.9)	93.7 (88.8-98.6)	.31
40-50	11.8 (9.3-14.9)	11.1 (8.9-14)	.38	83.9 (77.2-90.6)	82.5 (75.3-89.8)	.25
≥50	20.3 (14.6-28.2)	20.4 (14.5-28.7)	.07	96.2 (90.9-101.5)	95.9 (90.3-101.5)	.21
Sex						
Male	16.5 (12.6-21.5)	16.9 (13-21.9)	—	89.2 (82.9-95.6)	89.7 (83.4-95.9)	—
Female	12.8 (11-14.8)	12.6 (10.8-14.9)	.07	89.3 (85.3-93.3)	89.1 (85-93.2)	.88
Ethnicity						
Han	13.9 (12.1-16)	13.8 (11.9-15.9)	—	90.3 (86.8-93.8)	90.1 (86.6-93.6)	—
Minority	12.8 (8.6-19.2)	13.6 (9.5-19.3)	.94	83.3 (72.6-94)	83.9 (73.2-94.7)	.22
Workplace						
Hospital	14 (12-16.4)	14.3 (12.2-16.7)	—	89.7 (85.8-93.6)	89.8 (86-93.6)	—
Nonhospital	13 (9.8-17.1)	12.3 (9.4-16.1)	.36	88.1 (81.1-95.1)	87.8 (80.6-94.9)	.62
Booster vaccine						
BBIBP-CorV	13.3 (11-15.9)	13.1 (10.8-15.8)	—	88.6 (83.7-93.4)	87.6 (82.5-92.8)	—
CoronaVac	14.2 (11.8-17.2)	14.4 (11.9-17.5)	.48	90.1 (85.4-94.7)	90.8 (86.4-95.2)	.36

[a] N/A: not applicable.

[b] —: not available.

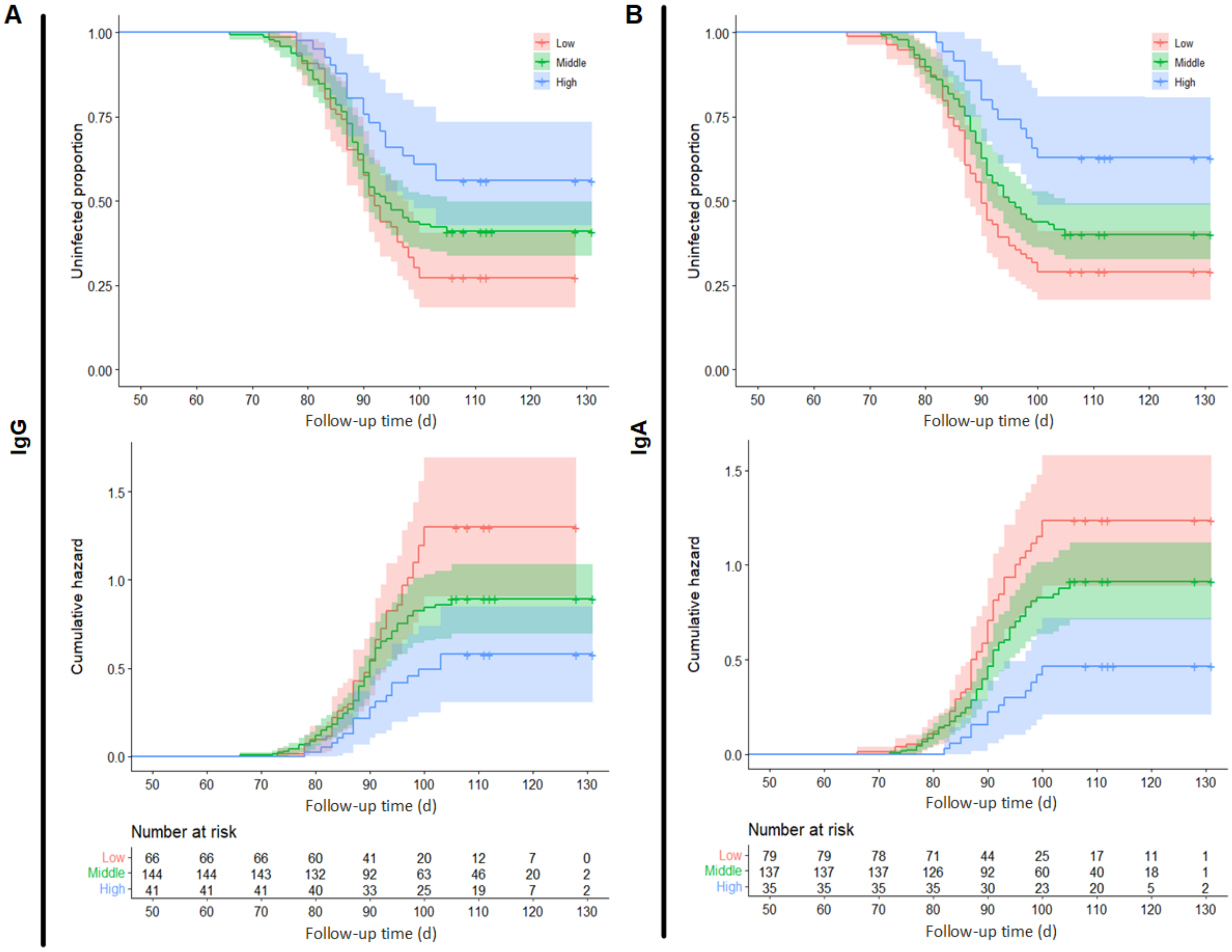
Real-World Protective Effect of Secondary Booster Administration Against Infection

After the second booster, we continued to track and follow up with the participants to derive the breakthrough infection rates. As of January 30, 2022, we followed up with 251 participants; the breakthrough infection rate was 60.2% (151/251).

After adjusting for confounding factors, Cox regression analysis revealed a significant correlation between postimmunization IgG antibody levels and breakthrough COVID-19 infection (hazard ratio 0.60; 95% CI 0.45-0.79); the same result was obtained for IgA antibody levels (hazard ratio 0.55; 95% CI 0.39-0.78). This result suggests that both IgG and IgA antibodies after the second booster can provide a certain degree of protection against COVID-19 and prevent infection. We categorized postimmunization IgG and IgA antibodies into 3 levels, high, medium, and low, based on 25th percentile and 75th percentile. Survival and cumulative risk curves were plotted for each level, as shown in Figure 3. Higher levels of postimmunization IgG antibodies were associated with better protection against COVID-19 infection and a lower cumulative risk for vaccine recipients (Figure 3A). Similarly, higher levels of postimmunization IgA antibodies were associated with better protection against COVID-19 infection and a lower cumulative risk for vaccine recipients (Figure 3B).

Kaplan-Meier curves for the risk of COVID-19 infection for age and sex subgroups were plotted, as shown in Figure 1. Differences between age groups were found: participants between 30 and 40 years of age had a higher hazard risk than other age groups. The differences were insignificant after adjusting for other sociodemographic factors. Women appear to have a higher risk of infection than men, but this did not reach statistical significance (*P*=.051).

Figure 3. The Kaplan-Meier and cumulative hazard curve of COVID-19 infection grouped by antibody levels. (A) Curves grouped by IgG levels (low, middle, and high). (B) Curves grouped by IgA levels (low, middle, and high). IgA: immunoglobulin A; IgG: immunoglobulin G.



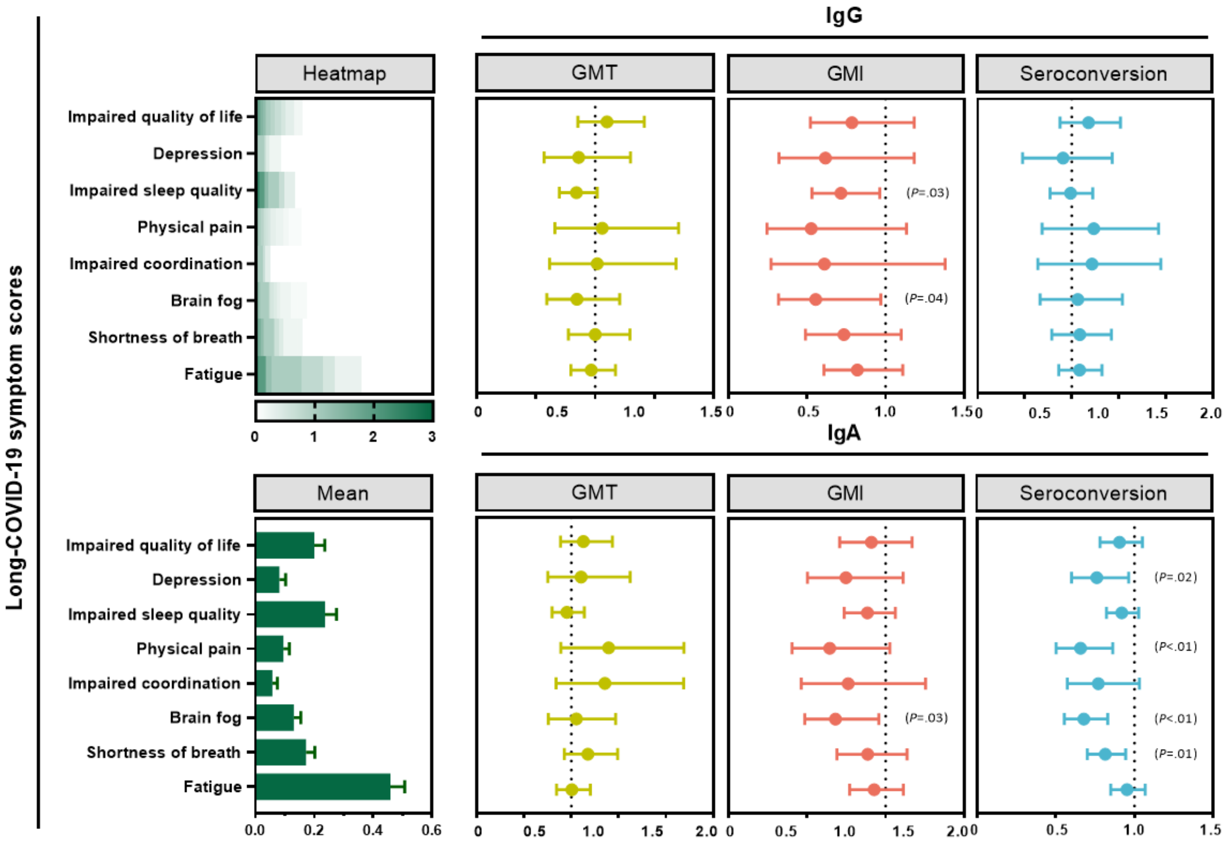
Real-World Protective Effect of Secondary Booster Administration Against Post–COVID-19 Symptoms

Finally, a questionnaire was used to evaluate the post–COVID-19 symptoms of 151 participants who experienced breakthrough infections, including fatigue, shortness of breath, brain fog, impaired coordination, physical pain, impaired sleep quality, depression, and impaired quality of life. The most common symptoms were fatigue, impaired sleep quality, and impaired quality of life, with mean scores of 0.46, 0.24, and 0.20, respectively.

We constructed a regression model with postimmunization antibody levels as the independent variable and questionnaire scores as the dependent variable. Based on the results, the GMI levels of postimmunization IgG antibodies had a protective effect on brain fog (odds ratio [OR] 0.56; 95% CI 0.32-0.97) and sleep quality (OR 0.71; 95% CI 0.53-0.96). The GMI levels of postimmunization IgA antibodies had a protective effect on brain fog (OR 0.68; 95% CI 0.49-0.96). Additionally, participants who were IgA antibody seropositive after infection had milder symptoms of shortness of breath (OR 0.81; 95% CI 0.7-0.95), brain fog (OR 0.68; 95% CI 0.55-0.83), impaired coordination (OR 0.66; 95% CI 0.50-0.86), and physical pain (OR 0.76; 95% CI 0.60-0.96; Figure 4).

The age-stratified protective effect of IgG and IgA after secondary enhancement on post–COVID-19 symptoms was reported. Unfortunately, no differences between age groups were found, and detailed results can be seen in Figure 2.

Figure 4. The adjusted regression models of antibodies on post–COVID-19 scores. The heatmap and bar chart on the left represent the discrete and central trends of the post–COVID-19 scores, and the scores of each subitem entered the model as a dependent variable. The 3 graphs on the top right are the results of entering different metrics of IgG (ie, GMT, GMI, and seroconversion rate) into the regression model as independent variables. The 3 graphs on the bottom right are the results of entering different metrics of IgA (ie, GMT, GMI, and seroconversion rate) into the regression model as independent variables. Each regression adjusted for basic sociodemographic characteristics. GMI: geometric mean increase; GMT: geometric mean titer; IgA: immunoglobulin A; IgG: immunoglobulin G.



Discussion

Overview

Based on existing evidence, there is a significant decline in neutralizing antibody titers 4-5 months after completion of the routine vaccination program [23,24]. The immune protection induced by the vaccine declines continuously with time, highlighting the urgent need for booster vaccination to enhance protection. According to statistics from the National Health Commission, approximately 850 million people in China have received their booster vaccine as of February 2023 [16]. The administration of a third dose of the same vaccine has been demonstrated to significantly increase neutralizing antibody levels and effectively reduce the symptomatic infection rate of the SARS-CoV-2 variant [25,26].

The literature suggests a decline or even disappearance of antibody levels within a short period of 3-6 months [27-30]. In this study, we tracked the antibody level owing to the first booster for 6-12 months without the interference of natural infection to evaluate long-term immunogenicity. We found that the decline in antibody level was slow after 6 months and maintained at a relatively low level, with a GMT of approximately 96 for IgG and approximately 8 for IgA. In confirmatory research, vaccine effectiveness was estimated to decline from approximately 70% one week after the booster dose to approximately 40% at 15 weeks or more [31].

The titers of postadministration antibodies vary according to the vaccine type. In China, inactivated vaccines are the most commonly used vaccines for basic and booster immunizations owing to their safety and stability. Based on evidence from the Chinese Center for Disease Control and Prevention, the GMTs owing to BBIBP-CorV were 25 at 1 month and 4 at 12 months, while those owing to CoronaVac were 20.2 and 4.1, respectively [32]. In this study, ELISA revealed that the IgG antibody titers from CoronaVac were slightly higher than those from BBIBP-CorV at 6-12 months after administration but converged at 12 months.

Breakthrough infections have become more common with the decline in antibody levels and the development of new VOCs with strong immunologic escape, despite the remarkable effect of primary boosting [33]. A fourth dose of the COVID-19 vaccine can boost cellular and humoral immunity, and the peak responses were found to be similar to the peak responses after the third dose [34]. According to some clinical trials, the adenovirus vector booster dose based on an inactivated vaccine could lead to higher neutralization antibodies than homologous boostering [35,36]. In China, nearly 47 million residents have now completed the sequential booster immunization since the start of its dissemination in November 2022 [14].

Based on the available evidence, a prospective study was conducted to evaluate heterologous secondary booster administration. To our knowledge, this study is the first real-world evaluation of the effectiveness of the second booster dose in China. Herein, the inhalant Ad5-nCoV vaccine was administered as a second booster dose. Inhalant Ad5-nCoV is homologous to injectable Ad5-nCoV but achieves protection through mucosal immunization through inhalation. Mucosal immunity is a critical component of the human immune system, with more than 90% of infections occurring in the mucosa, which comprises numerous dendritic cells with strong T-cell activation capacity that can induce an immune response. ELISA to detect the serum antibodies revealed that the GMT of the IgG antibody was 4978.2 and that of the IgA antibody was 107.8 at 28 days after inhalation. The seroconversion rates for IgG and IgA were 93.8% and 86.9%, respectively. The antibody titers of the primary booster against the Omicron variant were attenuated relative to those of other virus strains, such as the wild type and other VOCs. Therefore, the use of heterogeneous vaccines for the second-booster procedures is a concern for the prevention of the Omicron variant [37].

China suffered a COVID-19 epidemic between December 2022 and January 2023 owing to changes in health policies and the impact of the Omicron variant, with a peak in cases on December 22, 2022 [13]. In this study, we followed up with participants for 4 months after receiving their fourth vaccine dose to assess their real-world COVID-19 infection status. As of January 30, 2023, 60.2% (151/251) of participants self-reported that they had been infected with COVID-19. Notably, IgG and IgA provided strong protection against COVID-19 infection, as demonstrated by the high antibody titers after the fourth dose.

In addition to infection prevention, the long-term effects of COVID-19 are also concerning. Studies from high-income countries suggest that vaccination may alleviate post–COVID-19 or PCC [38]. However, based on other evidence, COVID-19 vaccination is not associated with improvement in PCC [39,40]. We sought to assess some nonspecific post–COVID-19 symptoms. Based on our results, the GMI and seroconversion rates of IgG and IgA may alleviate some of the symptoms after secondary booster administration, including sleep quality, shortness of breath, brain fog, impaired coordination, and physical pain. The GMTs had no statistical relationship, which may be explained by the insufficient statistical efficacy induced by the small sample size. However, its clinical value is still worth exploring. The varying levels of IgG and IgA suggest the mechanisms of immune protection after infection, suggesting that further tracking and research are warranted.

Our study had some limitations. First, the SARS-CoV-2 infection of participants was self-reported, which may have led to recall and reporting biases. Further, asymptomatic patients may not have been identified. Second, the antibody detection method used was an ELISA quantitative assay rather than the neutralization test. However, the results of both tests are highly correlated according to the literature. ELISA can be used as a substitute for the gold standard to assess immunogenicity [22]. Third, there was no control group setting without a second booster dose for estimating the vaccine effect of the second booster dose. Their use of antibody levels was reasonable in accordance with previous trials, however [41]. Finally, only follow-up data collected within approximately 1 month after infection were reported in this study. These findings may offer suggestions for future populations with post–COVID-19. Further studies are needed, including appropriate control measures for unvaccinated individuals, to confirm the trajectory of persistent symptoms after COVID-19 vaccination.

Conclusions

The IgG and IgA antibodies did not decrease significantly but remained at a relatively low level after administration of the second booster dose. The antibodies generated significant immunogenic protection against breakthrough infections and might partially alleviate post–COVID-19 symptoms. Further studies on secondary booster administration are needed to validate these correlations.

Acknowledgments

The authors would like to thank professors Fengcai Zhu and Aihua Zhang (Guizhou Medical University) for their great support and expertise in this study. We would also like to thank the following organizations for their full support in the implementation of the project: Guiyang Bureau of Health, Jinhua Community Healthcare Center of Guanshanhu District, Century-City Third Community Healthcare Center of Guanshanhu District, Qingzhen Center for Disease Control and Prevention, Baihua Community Healthcare Center of Qingzhen, Hongta Community Healthcare Center of Qingzhen, Baiyun Branch of Guizhou Provincial People's Hospital, the Affiliated Hospital of Guizhou Medical University (Baiyun Branch), Kaiyang Center for Disease Control and Prevention, Kaiyang COVID-19 Vaccination Site, Xifeng Bureau of Health, Xifeng Maternity and Child Healthcare Hospital, Xiuwen Center for Disease Control and Prevention, Baixin Hospital of Xiuwen, Center for Disease Control and Prevention of Yunyan District, and Qianling Hospital of Yunyan District. This study was supported by the National Key Research and

Development Program of China (2021YFC2301604) and Joint Research Fund for Beijing Natural Science Foundation and Haidian Original Innovation (L222029). The funders had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.

Data Availability

All data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

GY, FC, and SL designed the study. ML designed the survey instrument. QM, RZ, CL, FX, LL, LZ, SZ, and NH all contributed to the conceptual design of the research and to the data collection. XC and TZ conducted the statistical analysis. ML and TZ wrote the manuscript. GY (ghyang_gzmu@outlook.com), FC (cuifuq@bjmu.edu.cn), and SL (ShiguangLei193@foxmail.com) are co-corresponding authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Post-COVID symptom scale, Kaplan-Meier curves of COVID-19 infection, and age-stratified correlation between post-administrated antibody and post-COVID symptoms.
[DOCX File , 7592 KB - publichealth_v9i1e47272_app1.docx]

References

1. Haldane V, De Foo C, Abdalla SM, Jung A, Tan M, Wu S, et al. Health systems resilience in managing the COVID-19 pandemic: lessons from 28 countries. Nat Med 2021;27(6):964-980 [FREE Full text] [doi: 10.1038/s41591-021-01381-y] [Medline: 34002090]
2. Tracking SARS-CoV-2 variants. World Health Organization. 2022. URL: https://www.who.int/activities/tracking-SARS-CoV-2-variants [accessed 2022-02-24]
3. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature 2021;592(7854):438-443 [FREE Full text] [doi: 10.1038/s41586-021-03402-9] [Medline: 33690265]
4. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science 2021;372(6544):815-821 [FREE Full text] [doi: 10.1126/science.abh2644] [Medline: 33853970]
5. Dhar MS, Marwal R, Vs R, Ponnusamy K, Jolly B, Bhoyar RC, Indian SARS-CoV-2 Genomics Consortium (INSACOG), et al. Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. Science 2021;374(6570):995-999 [FREE Full text] [doi: 10.1126/science.abj9932] [Medline: 34648303]
6. Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 omicron variant in Southern Africa. Nature 2022;603(7902):679-686 [FREE Full text] [doi: 10.1038/s41586-022-04411-y] [Medline: 35042229]
7. Cele S, Jackson L, Khoury DS, Khan K, Moyo-Gwete T, Tegally H, NGS-SA, COMMIT-KZN Team, et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. Nature 2022;602(7898):654-656 [FREE Full text] [doi: 10.1038/s41586-021-04387-1] [Medline: 35016196]
8. Zhang L, Li Q, Liang Z, Li T, Liu S, Cui Q, et al. The significant immune escape of pseudotyped SARS-CoV-2 variant Omicron. Emerg Microbes Infect 2022;11(1):1-5 [FREE Full text] [doi: 10.1080/22221751.2021.2017757] [Medline: 34890524]
9. Lu L, Mok BWY, Chen LL, Chan JMC, Tsang OTY, Lam BHS, et al. Neutralization of severe acute respiratory syndrome coronavirus 2 omicron variant by sera from BNT162b2 or CoronaVac vaccine recipients. Clin Infect Dis 2022;75(1):e822-e826 [FREE Full text] [doi: 10.1093/cid/ciab1041] [Medline: 34915551]
10. Dejnirattisai W, Shaw RH, Supasa P, Liu C, Stuart AS, Pollard AJ, Com-COV2 study group. Reduced neutralisation of SARS-CoV-2 omicron B.1.1.529 variant by post-immunisation serum. Lancet 2022;399(10321):234-236 [FREE Full text] [doi: 10.1016/S0140-6736(21)02844-0] [Medline: 34942101]
11. Edara VV, Manning KE, Ellis M, Lai L, Moore KM, Foster SL, et al. mRNA-1273 and BNT162b2 mRNA vaccines have reduced neutralizing activity against the SARS-CoV-2 omicron variant. Cell Rep Med 2022;3(2):100529 [FREE Full text] [doi: 10.1016/j.xcrm.2022.100529] [Medline: 35233550]
12. Ai J, Zhang H, Zhang Y, Lin K, Zhang Y, Wu J, et al. Omicron variant showed lower neutralizing sensitivity than other SARS-CoV-2 variants to immune sera elicited by vaccines after boost. Emerg Microbes Infect 2022;11(1):337-343 [FREE Full text] [doi: 10.1080/22221751.2021.2022440] [Medline: 34935594]

13. The situation of the SARS-CoV-2 infection in China. Chinese Center for Disease Control and Prevention. 2023. URL: https://www.chinacdc.cn/jkzt/crb/zl/szkb_11803/jszl_13141/202302/t20230218_263807.html [accessed 2023-02-24]
14. Dedicated COVID-19 vaccination dashboard. World Health Organization. 2023. URL: https://covid19.who.int [accessed 2023-02-24]
15. Bai Y, Peng Z, Wei F, Jin Z, Wang J, Xu X, et al. Study on the COVID-19 epidemic in mainland China between November 2022 and January 2023, with prediction of its tendency. J Biosaf Biosecur 2023;5(1):39-44 [FREE Full text] [doi: 10.1016/j.jobb.2023.03.001] [Medline: 36992708]
16. Transcript of press conference under the Joint Prevention and Control Mechanism of The State Council on 23 Feb 2023. National Health Commission of China. URL: http://www.nhc.gov.cn/xcs/yqfkdt/202302/172708cde8fb4e40976e693443bbb596.shtml [accessed 2023-02-24]
17. WHO interim recommendations for heterologous COVID-19 vaccine schedules. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-vaccines-SAGE-recommendation-heterologous-schedules [accessed 2023-08-26]
18. Jin P, Guo X, Chen W, Ma S, Pan H, Dai L, et al. Safety and immunogenicity of heterologous boost immunization with an adenovirus type-5-vectored and protein-subunit-based COVID-19 vaccine (Convidecia/ZF2001): a randomized, observer-blinded, placebo-controlled trial. PLoS Med 2022;19(5):e1003953 [FREE Full text] [doi: 10.1371/journal.pmed.1003953] [Medline: 35617368]
19. Wu S, Huang J, Zhang Z, Wu J, Zhang J, Hu H, et al. Safety, tolerability, and immunogenicity of an aerosolised adenovirus type-5 vector-based COVID-19 vaccine (Ad5-nCoV) in adults: preliminary report of an open-label and randomised phase 1 clinical trial. Lancet Infect Dis 2021;21(12):1654-1664 [FREE Full text] [doi: 10.1016/S1473-3099(21)00396-0] [Medline: 34324836]
20. Fawzy NA, Shaar BA, Taha RM, Arabi TZ, Sabbah BN, Alkodaymi MS, et al. A systematic review of trials currently investigating therapeutic modalities for post-acute COVID-19 syndrome and registered on WHO International Clinical Trials Platform. Clin Microbiol Infect 2023;29(5):570-577 [FREE Full text] [doi: 10.1016/j.cmi.2023.01.007] [Medline: 36642173]
21. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: A Practical Guide. Cambridge, UK: Cambridge University Press; 2011.
22. Dolscheid-Pommerich R, Bartok E, Renn M, Kümmerer BM, Schulte B, Schmithausen RM, et al. Correlation between a quantitative anti-SARS-CoV-2 IgG ELISA and neutralization activity. J Med Virol 2022;94(1):388-392 [FREE Full text] [doi: 10.1002/jmv.27287] [Medline: 34415572]
23. Collier ARY, Yu J, McMahan K, Liu J, Chandrashekar A, Maron JS, et al. Differential kinetics of immune responses elicited by Covid-19 vaccines. N Engl J Med 2021;385(21):2010-2012 [FREE Full text] [doi: 10.1056/NEJMc2115596] [Medline: 34648703]
24. Suo S, Hongyang Y, Qian L, Yang Y, Jiayin S, Jianqing X, et al. Expert recommendations on booster immunization strategies of 2019-nCoV vaccines. Chin J Clin Indect Dis 2022;15(3):176-184 [FREE Full text] [doi: 10.3760/cma.j.issn.1674-2397.2022.03.003]
25. Cao Y, Wang X, Li S, Dong Y, Liu Y, Li J, et al. A third high dose of inactivated COVID-19 vaccine induces higher neutralizing antibodies in humans against the Delta and Omicron variants: a randomized, double-blinded clinical trial. Sci China Life Sci 2022;65(8):1677-1679 [FREE Full text] [doi: 10.1007/s11427-022-2110-1] [Medline: 35441932]
26. Zeng G, Wu Q, Pan H, Li M, Yang J, Wang L, et al. Immunogenicity and safety of a third dose of CoronaVac, and immune persistence of a two-dose schedule, in healthy adults: interim results from two single-centre, double-blind, randomised, placebo-controlled phase 2 clinical trials. Lancet Infect Dis 2022;22(4):483-495 [FREE Full text] [doi: 10.1016/S1473-3099(21)00681-2] [Medline: 34890537]
27. Cheng ZJ, Huang H, Zheng P, Xue M, Ma J, Zhan Z, et al. Humoral immune response of BBIBP COVID-19 vaccination before and after the booster immunization. Allergy 2022;77(8):2404-2414 [FREE Full text] [doi: 10.1111/all.15271] [Medline: 35255171]
28. Wisnewski AV, Luna JC, Redlich CA. Human IgG and IgA responses to COVID-19 mRNA vaccines. PLoS One 2021;16(6):e0249499 [FREE Full text] [doi: 10.1371/journal.pone.0249499] [Medline: 34133415]
29. Lyke KE, Atmar RL, Islas CD, Posavad CM, Szydlo D, Chourdhury RP, DMID 21-0012 Study Group. Rapid decline in vaccine-boosted neutralizing antibodies against SARS-CoV-2 omicron variant. Cell Rep Med 2022;3(7):100679 [FREE Full text] [doi: 10.1016/j.xcrm.2022.100679] [Medline: 35798000]
30. Thomas SJ, Moreira ED, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine through 6 months. N Engl J Med 2021;385(19):1761-1773 [FREE Full text] [doi: 10.1056/NEJMoa2110345] [Medline: 34525277]
31. Kirsebom FCM, Andrews N, Stowe J, Toffa S, Sachdeva R, Gallagher E, et al. COVID-19 vaccine effectiveness against the omicron (BA.2) variant in England. Lancet Infect Dis 2022;22(7):931-933 [FREE Full text] [doi: 10.1016/S1473-3099(22)00309-7] [Medline: 35623379]

32. Wang F, Huang B, Lv H, Feng L, Ren W, Wang X, et al. Factors associated with neutralizing antibody levels induced by two inactivated COVID-19 vaccines for 12 months after primary series vaccination. Front Immunol 2022;13:967051 [FREE Full text] [doi: 10.3389/fimmu.2022.967051] [Medline: 36159863]
33. Levin EG, Lustig Y, Cohen C, Fluss R, Indenbaum V, Amit S, et al. Waning immune humoral response to BNT162b2 Covid-19 vaccine over 6 months. N Engl J Med 2021;385(24):e84 [FREE Full text] [doi: 10.1056/NEJMoa2114583] [Medline: 34614326]
34. Regev-Yochay G, Gonen T, Gilboa M, Mandelboim M, Indenbaum V, Amit S, et al. Efficacy of a fourth dose of Covid-19 mRNA vaccine against Omicron. N Engl J Med 2022;386(14):1377-1380 [FREE Full text] [doi: 10.1056/NEJMc2202542] [Medline: 35297591]
35. Atmar RL, Lyke KE, Deming ME, Jackson LA, Branche AR, El Sahly HM, DMID 21-0012 Study Group. Homologous and heterologous Covid-19 booster vaccinations. N Engl J Med 2022;386(11):1046-1057 [FREE Full text] [doi: 10.1056/NEJMoa2116414] [Medline: 35081293]
36. Clemens SAC, Weckx L, Clemens R, Mendes AVA, Souza AR, Silveira MBV, RHH-001 study team. Heterologous versus homologous COVID-19 booster vaccination in previous recipients of two doses of CoronaVac COVID-19 vaccine in Brazil (RHH-001): a phase 4, non-inferiority, single blind, randomised study. Lancet 2022;399(10324):521-529 [FREE Full text] [doi: 10.1016/S0140-6736(22)00094-0] [Medline: 35074136]
37. Khong KW, Zhang R, Hung IFN. The four Ws of the fourth dose COVID-19 vaccines: why, who, when and what. Vaccines (Basel) 2022;10(11):1924 [FREE Full text] [doi: 10.3390/vaccines10111924] [Medline: 36423020]
38. Arnold DT, Milne A, Samms E, Stadon L, Maskell NA, Hamilton FW. Symptoms after COVID-19 vaccination in patients with persistent symptoms after acute infection: a case series. Ann Intern Med 2021;174(9):1334-1336 [FREE Full text] [doi: 10.7326/M21-1976] [Medline: 34029484]
39. Wisnivesky JP, Govindarajulu U, Bagiella E, Goswami R, Kale M, Campbell KN, et al. Association of vaccination with the persistence of Post-COVID symptoms. J Gen Intern Med 2022;37(7):1748-1753 [FREE Full text] [doi: 10.1007/s11606-022-07465-w] [Medline: 35266128]
40. Wynberg E, Han AX, Boyd A, van Willigen HDG, Verveen A, Lebbink R, RECoVERED Study Group. The effect of SARS-CoV-2 vaccination on post-acute sequelae of COVID-19 (PASC): a prospective cohort study. Vaccine 2022;40(32):4424-4431 [FREE Full text] [doi: 10.1016/j.vaccine.2022.05.090] [Medline: 35725782]
41. Tang R, Zheng H, Wang BS, Gou JB, Guo XL, Chen XQ, CanSino COVID-19 Study Group. Safety and immunogenicity of aerosolised Ad5-nCoV, intramuscular Ad5-nCoV, or inactivated COVID-19 vaccine CoronaVac given as the second booster following three doses of CoronaVac: a multicentre, open-label, phase 4, randomised trial. Lancet Respir Med 2023;11(7):613-623 [FREE Full text] [doi: 10.1016/S2213-2600(23)00049-8] [Medline: 36898400]

Abbreviations

ELISA: enzyme-linked immunosorbent assay
GMI: geometric mean increase
GMT: geometric mean titer
IgA: immunoglobulin A
IgG: immunoglobulin G
LOWESS: locally weighted scatterplot smoothing
OR: odds ratio
PCC: post–COVID-19 condition
VOC: variant of concern

Edited by A Mavragani, T Sanchez; submitted 14.03.23; peer-reviewed by L Rodewald, H Tian; comments to author 08.06.23; revised version received 25.06.23; accepted 08.08.23; published 11.10.23.

Please cite as:
Liu M, Zhao T, Mu Q, Zhang R, Liu C, Xu F, Liang L, Zhao L, Zhao S, Cai X, Wang M, Huang N, Feng T, Lei S, Yang G, Cui F
Immune-Boosting Effect of the COVID-19 Vaccine: Real-World Bidirectional Cohort Study
JMIR Public Health Surveill 2023;9:e47272
URL: https://publichealth.jmir.org/2023/1/e47272
doi:10.2196/47272
PMID:37819703

Review

Effects of SARS-CoV-2 Vaccines on Sperm Quality: Systematic Review

Guanjian Li[1,2*], MD, PhD; Rongqiu Zhang[3,4*], MD; Bing Song[1,2], MD, PhD; Chao Wang[1,2], MD; Qunshan Shen[5,6], MD, PhD; Xiaojin He[7], MD, PhD; Yunxia Cao[1,2], MD, PhD

[1]Reproductive Medicine Center, Department of Obstetrics and Gynecology, the First Affiliated Hospital of Anhui Medical University, Hefei, China

[2]National Health Commission Key Laboratory of Study on Abnormal Gametes and Reproductive Tract, Hefei, China

[3]Reproductive Medicine Center, the Affiliated Jinyang Hospital of Guizhou Medical University, Guiyang, China

[4]The Second People's Hospital of Guiyang, Guiyang, China

[5]Reproductive Medicine Center, Human Sperm Bank, Department of Obstetrics and Gynecology, the First Affiliated Hospital of Anhui Medical University, Hefei, China

[6]Key Laboratory of Population Health Across Life Cycle, Ministry of Education of the People’s Republic of China, Hefei, China

[7]Reproductive Medicine Center, Department of Obstetrics and Gynecology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Yunxia Cao, MD, PhD
Reproductive Medicine Center, Department of Obstetrics and Gynecology, the First Affiliated Hospital of Anhui Medical University
No 81 Meishan Road
Hefei, 230022
China
Phone: 86 15395104659
Email: caoyunxia6@126.com

Abstract

Background: The COVID-19 pandemic, caused by SARS-CoV-2, has triggered a global public health crisis of unprecedented proportions. SARS-CoV-2 vaccination is a highly effective strategy for preventing infections and severe COVID-19 outcomes. Although several studies have concluded that COVID-19 vaccines are unlikely to affect fertility, concerns have arisen regarding adverse events, including the potential impact on fertility; these concerns are plagued by limited and inconsistent evidence.

Objective: This review aims to provide a recent assessment of the literature on the impact of COVID-19 vaccines on male sperm quality. The possible impact of COVID-19 vaccines on fertility potential was also examined to draw a clearer picture and to evaluate the effects of COVID-19 on male reproductive health.

Methods: PubMed, Scopus, Web of Science, Embase, and Cochrane databases were searched from their inception to October 2023. Eligible studies included articles reporting SARS-CoV-2 vaccination and human semen quality and fertility, as well as the impact of vaccination on assisted reproductive technology treatment outcomes. The quality of cohort studies was assessed using the Newcastle-Ottawa Scale, and the quality of cross-sectional studies was assessed using the quality evaluation criteria recommended by the Agency for Healthcare Research and Quality. The systematic review followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Results: The initial literature search yielded 4691 records by searching 5 peer-reviewed databases (PubMed, Scopus, Web of Science, Embase, and Cochrane). Finally, 24 relevant studies were selected for our study. There were evident research inequalities at the regional level, with the United States and Western European countries contributing 38% (9/24) of the studies, Middle Eastern countries contributing 38% (9/24), China accounting for 21% (5/24), and Africa and South America accounting for none. Nonetheless, the overall quality of the included studies was generally good. Our results demonstrated that serious side effects of the COVID-19 vaccine are extremely rare, and men experience few problems with sperm parameters or reproductive potential after vaccination.

Conclusions: On the basis of the studies published so far, the COVID-19 vaccine is safe for male reproductive health. Obviously, vaccination is a wise option rather than experience serious adverse symptoms of viral infections. These instances of evidence may help reduce vaccine hesitancy and increase vaccination coverage, particularly among reproductive-age couples. As new controlled trials and prospective cohort studies with larger sample sizes emerge, the possibility of a negative effect of the COVID-19 vaccine on sperm quality must be further clarified.

(JMIR Public Health Surveill 2023;9:e48511) doi:10.2196/48511

KEYWORDS

COVID-19 vaccine; SARS-CoV-2; reproductive system; fertility; sperm quality

Introduction

Background

The COVID-19 pandemic, caused by SARS-CoV-2, has severely affected global public health, economics, and social life since 2019 [1-3]. To establish herd immunity and effectively reduce the risk of COVID-19 infection, the medical and research community conducted many research studies within a short period to develop the SARS-CoV-2 vaccines. Impressively, these efforts have ultimately led to the development of multiple efficacious vaccines in a short period, which is an unprecedented and important accomplishment [4]. Vaccination is proven to be an effective tool for reducing the risk of infection and severe COVID-19 outcomes and is now a global priority [5,6].

Earlier reports have confirmed that the testis is one of the tissues with the highest expression of angiotensin-converting enzyme 2 messenger RNA (mRNA) and protein, both of which are key targets for SARS-CoV-2 to enter host cells [7]. The inflammatory state caused by COVID-19 may undermine the integrity of the blood-testis barrier and further promote the infiltration of the virus. Therefore, in the early stages of the COVID-19 epidemic, some researchers argued that the testes might be susceptible to SARS-CoV-2. In fact, a series of studies have confirmed a short-term reduction in testosterone production and semen quality in patients with COVID-19. Semen parameters, including semen volume, sperm concentration, total sperm count, total sperm motility, and progressive sperm motility, were negatively affected by SARS-CoV-2 infection [8,9]. On the basis of the aforementioned potential damage to the reproductive system following SARS-CoV-2 infection, some researchers suspected that vaccines that mimic the virus could also affect fertility via a similar mechanism. Numerous types of disinformation and misinformation have emerged on social media, including questioning the relationship between COVID-19 vaccination and male fertility [10-12]. This public fear is partly owing to a lack of understanding of the newly developed COVID-19 vaccine [13,14].

It is also important to note that the current authoritative evidence on the impact of COVID-19 vaccines on male fertility is still limited, and the results are controversial. The lack of evidence for a definitive conclusion could discourage the public, fuel concerns about the potential impact of the vaccine on male fertility, and lead to vaccine hesitancy [15-17].

Objectives

As the pandemic continues to evolve, the number of newly approved vaccines and scientific studies will continue to increase [18]. In this study, we systematically reviewed the latest data and theoretical considerations on the impact of COVID-19 vaccines on semen quality in men to provide more conclusive information for policy makers, health care providers, media, and the public.

Methods

Search Strategy and Selection Criteria

According to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, we conducted a systematic review of empirical articles on SARS-CoV-2 vaccination, human semen quality and fertility, and the impact of vaccination on assisted reproductive technology (ART) treatment outcomes. The PRISMA checklist is provided in Multimedia Appendix 1. The PubMed, Scopus, Web of Science, Embase, and Cochrane databases were used to identify all relevant literature published up to October 2023. A keyword search strategy was developed with the following terms: “COVID-19,” “COVID19,” “covid19,” “covid-19,” “coronavirus,” “novel coronavirus,” “new coronavirus,” “SARS Coronavirus 2 Infection,” “2019 Novel Coronavirus Disease,” “2019 Novel Coronavirus Infection,” “SARS-CoV-2,” “severe acute respiratory syndrome coronavirus 2,” “COVID-19 Virus Disease,” “COVID-19 Virus Infection,” “Coronavirus Disease-19,” “Vaccines,” “vaccine,” “vaccination,” “COVID-19 Vaccines,” “COVID-19 Vaccines/adverse effects,” “SARS-CoV-2 Vaccine,” “Coronavirus Disease 2019 Vaccine,” “2019-nCoV Vaccine,” “SARS Coronavirus 2 Vaccine,” “BNT162b2 mRNA vaccine,” “COVID-19 mRNA Vaccine,” “Fertility,” “Fecundability,” “Fecundity,” “Infertility, Male,” “Infertility, Female,” “sterility,” “IVF,” “Fertilization in Vitro,” “reproductive function,” “ART,” “sperm,” “spermatozoa,” “semen analysis,” “spermatogenesis,” and “semen parameter.” The search strategies are provided in detail in Multimedia Appendix 2.

The search was restricted to articles on humans, published in any language, without additional restrictions. With regard to publication type, original studies were selected. Abstracts, comments, reviews (narrative or systematic), modeling only, case reports, case series (reporting data for <10 patients), and editorials were excluded. We evaluated the relevance and quality of the text. Two independent reviewers (GL and BS) screened all the studies by title, abstract, and keywords to identify potentially relevant articles. All articles obtained through the search strategy were imported into Rayyan (Qatar Computing Research Institute), and duplicates were removed. Papers that did not meet the inclusion criteria for the full-text evaluation were excluded. Subsequently, the full texts were screened to collect the data. Furthermore, the reference lists of the included studies were manually searched to identify additional eligible studies and to ensure that relevant studies were not missed.

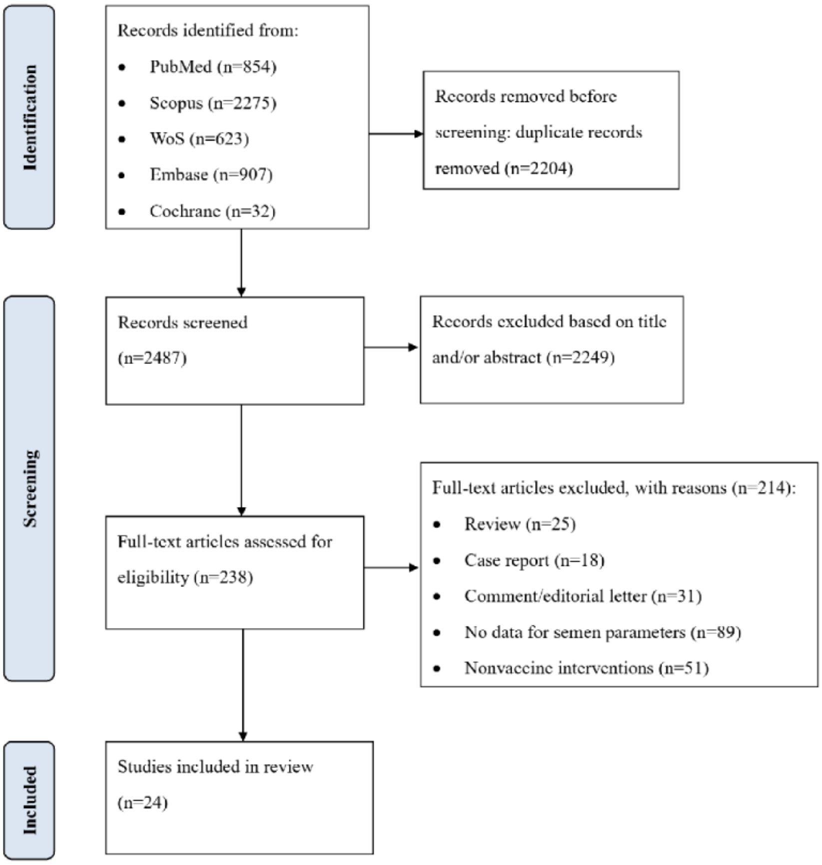
Eligible studies that were chosen to be included in this study were those evaluating the impacts of the COVID-19 vaccine on semen quality and male reproduction. Articles related to female vaccination and those without information on male vaccination were excluded. Data on the effect of the SARS-CoV-2 on male fertility were also excluded. The inclusion and exclusion criteria are presented in Table 1. Any disagreements were resolved through discussion among the study team members. The flow and results of the literature search are summarized in a PRISMA flowchart (Figure 1). Ethics approval was not required for this study because the data were derived from previously published studies.

The initial literature search yielded 4691 records. Of these, 2204 duplicate articles were removed, 2249 records were further excluded based on title or abstract screening, and 214 were excluded based on full-text assessment. Finally, 24 studies were selected for evaluation. We extracted the data using a pretested data extraction form created in Microsoft Excel 2019 (Microsoft Corporation). Article information (eg, PubMed Identifier, first author, year of publication, and country); study design; vaccine type; side effects; and changes in semen parameters were extracted from the articles.

Table 1. Summary of the inclusion and exclusion criteria.

Parameter	Inclusion criteria	Exclusion criteria
Article or study type	• Population-based original research studies	• Abstracts, comments, and editorials • Reviews and book chapters • Protocols and modeling only • Case reports and case series • Animal experiment
Language	• Any language	• Without additional restrictions
Publication period	• November 1, 2019, to October 1, 2023	• All dates outside the span from November 2019 to October 2023
Study content	• Studies assessing the impact of COVID-19 vaccines on semen quality and male reproduction	• Studies related to female vaccination without information on male vaccination • Data on the effects of COVID-19 on male fertility

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the screening and selection process. WoS: Web of Science.



Characteristics of the Included Literature

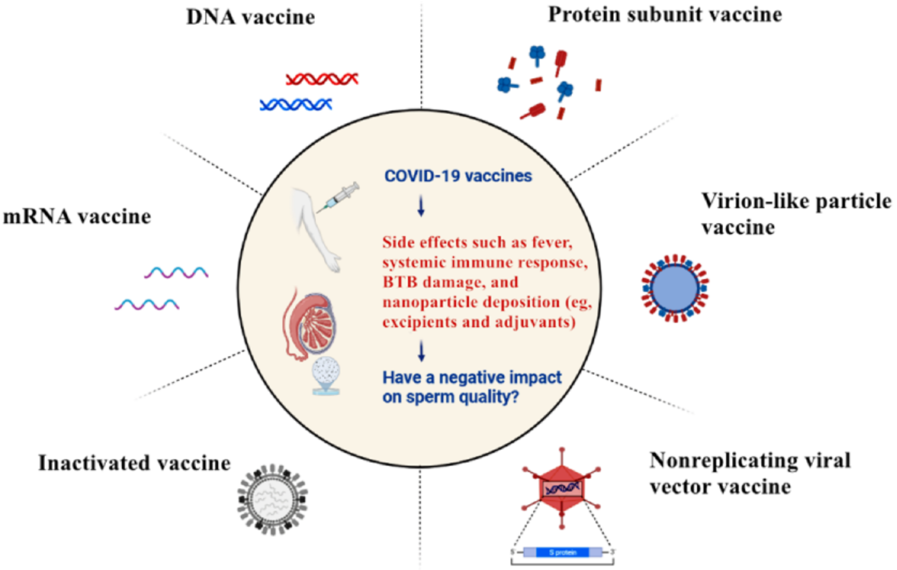
The methodological quality of the included studies was evaluated according to study design. Cohort studies were assessed using the Newcastle-Ottawa Scale, which provides a score between 0 and 9 and assesses the following quality parameters: comparability of study groups, selection of study groups, and ascertainment of outcomes [19]. Cross-sectional studies were evaluated using 11 entries from the quality evaluation criteria recommended by the Agency for Healthcare Research and Quality [20]. Each entry contained options "Yes," "No," and "Unclear." "Yes" was rated as 1 point, while "No" and "Unclear" were counted as 0. Scores of 0 to 3, 4 to 7, and 8 to 11 were considered as low-, medium- and high-quality studies, respectively.

Results

COVID-19 Vaccine Platform

Vaccines are critical and cost-effective tools to control the COVID-19 pandemic. Since the onset of the COVID-19 pandemic, research groups from various countries have made significant progress in developing vaccines in a short period using a variety of platform technologies, with >100 vaccine candidates currently undergoing preclinical development worldwide [21]. These vaccine platforms range from protein-based (eg, subunits and virus-like particles) and virus-based (eg, attenuated live and inactivated vaccines) strategies to gene-based strategies [22,23]. The different types of vaccine platforms are shown in Figure 2. In this study, we focused on several widely used vaccines that have been reported to affect semen quality.

Figure 2. Different platform types of COVID-19 vaccines and the possible mechanisms affecting semen quality. BTB: blood-testis barrier; mRNA: messenger RNA.



mRNA Vaccines

Several large biotechnology companies such as Pfizer, BioNTech, and Moderna have developed COVID-19 vaccines using advanced mRNA vaccine platforms. In the United States, the Food and Drug Administration approved emergency authorizations for the BNT162b2 (Pfizer-BioNTech) and mRNA - 1273 (Moderna) vaccines in December 2020 [24-26]. mRNA vaccines involve an innovative approach, borrowing the cell's transcriptional machinery to produce the desired set of particles. Vaccine mRNA can enter cells and can be converted into SARS-CoV-2 glycoproteins (spike proteins). The spike proteins elicit an immune response by activating helper T and B cells, eventually establishing immunity against SARS-CoV-2 [27]. These mRNA-based vaccines are noninfectious. They do not contain any live viral particles and are synthesized by in vitro transcription in the absence of microbial molecules.

Recombinant Viral-Vectored Vaccines

Recombinant viral-vectored vaccines use a bioengineered viral vector that can express and clone antigens from the target pathogen [28]. Lentiviruses, retroviruses, adenoviruses, and adeno-associated viruses are the common vectors used in this platform. Nonreplicating viral vaccines have been extensively studied and have a well-established track record for preventing infectious diseases because of their genetic malleability, safety, and ability to stimulate strong cellular immune responses without the use of adjuvants. Furthermore, a single dose of viral vector–based vaccines may provide adequate protection. Several COVID-19 vaccines based on recombinant viral vector technology including Sputnik V (Gamaleya Institute), ChAdOx1 (Oxford Astra Zeneca), and Ad26. COV2. S (Johnson & Johnson Janssen) have been approved for emergency use worldwide [29].

Inactivated Vaccines

Chemically or physically inactivated viral vaccines are among the oldest vaccine design approaches and have been used successfully against a variety of human viral infections such as hepatitis A, polio, and influenza [30]. This method involves injecting an inactivated virus into the host, inducing an immune response and promoting strong immunity against the virus. In contrast to their live attenuated counterparts, inactivated viral vaccines pose few safety concerns and deliver a wide range of native viral antigens, including surface antigens with conserved epitope conformations, which can stimulate conformation-dependent antibody responses. As completely inactive viruses do not replicate, repeated administration and adjuvants are required to stimulate the immune system and ensure that these vaccines work properly [31]. Currently, several vaccines, including CoronaVac (Sinovac) and BBIBP-CorV (Sinopharm), are being widely used in several countries, including China.

Effects of SARS-CoV-2 Vaccination on Semen Quality: Available Evidence

All available COVID-19 vaccines had acceptable efficacy and safety profiles in phase 3 clinical trials [32]. As of January 2023, >13 billion doses of the vaccine had been administered worldwide, with an excellent safety and efficacy profile. However, as reproductive toxicity has not been evaluated in clinical trials, one reason for vaccine hesitancy is its potential negative effects on fertility [33].

As shown in Table 2, we identified 24 studies that evaluated the effects of different COVID-19 vaccines on semen quality (some studies simultaneously evaluated >1 vaccine). All the studies were published between March 2020 and October 2023. The overall quality of both the cross-sectional and cohort studies is shown in Table 2.

All the studies were of reasonable quality. Of 24 studies, 9 (38%) studies were conducted in the United States and Western European countries [34-42], 5 (21%) in China [43-47], 1 (4%) in Indonesia [48], and 9 (38%) in the Middle Eastern countries [49-57]. In total, 3 (12%) studies used inactivated COVID-19 vaccines, 4 (17%) studies used viral-vectored vaccines, 8 (33%) studies used mRNA vaccines, and 3 (12%) studies included a combined analysis of multiple types of vaccines. A total of 11 (46%) studies involved patients with infertility undergoing ART, 11 (46%) involved healthy volunteers or sperm donors, and 2 (8%) involved other male populations with type 2 diabetes and nonserious chronic conditions. All studies described standard methods for sample collection and processing and followed the fifth edition of the World Health Organization laboratory manual. ’The study design, study population, vaccine type, result analysis, and adverse reaction reports of all 24 studies are listed in Table 3. These studies reported a variety of semen parameters: 9 (38%) reported total sperm count, 23 (96%) reported sperm concentration, 7 (29%) reported sperm morphology, 19 (79%) reported total sperm motility, 17 (71%) reported progressive sperm motility, and 11 (46%) reported total motile sperm count (TMSC).

Table 2. Summary of the inclusion and exclusion criteria.

Source	Country	Study design	AHRQ[a] or NOS[b] score	Interpretation
Diaz et al [34], 2021-2022	United States	Prospective cross-sectional	6	Medium quality
Gonzalez et al [35], 2020-2021	United States	Prospective cross-sectional	5	Medium quality
Wang et al [43], 2021	China	Prospective cross-sectional	6	Medium quality
Huang et al [44], 2021-2022	China	Retrospective cross-sectional	8	High quality
Zhu et al [45], 2020-2021	China	Retrospective cross-sectional	8	High quality
Xia et al [46], 2021	China	Retrospective cohort study	7	High quality
Dong et al [47], 2023	China	Retrospective cohort study	8	High quality
Abd et al [51], 2022	Iraq	Prospective cross-sectional	4	Medium quality
Al-Alami et al [50], 2021	Jordan	Retrospective cohort study	5	Medium quality
Alenzi et al [49], 2021	Saudi Arabia	Prospective cohort study	5	Medium quality
Lestari et al [48], 2020-2022	Indonesia	Prospective cohort study	8	Medium quality
Karavani et al [52], 2021-2022	Israel	Retrospective cross-sectional	7	Medium quality
Gat et al [53], 2021-2022	Israel	Retrospective cross-sectional	7	Medium quality
Lifshitz et al [54], 2021	Israel	Prospective cross-sectional	5	Medium quality
Safrai et al [55], 2021	Israel	Retrospective cross-sectional	9	High quality
Barda et al [56], 2020-2021	Israel	Prospective cross-sectional	4	Medium quality
Orvieto et al [57], 2021	Israel	Retrospective cohort study	6	Medium quality
Olana et al [36], 2021	Italy	Prospective cross-sectional	8	High quality
Reschini et al [37], 2022	Italy	Retrospective cross-sectional	5	Medium quality
Massarotti et al [38], 2021	Italy	Prospective cross-sectional	6	Medium quality
Chillon et al [42], 2021	Germany	Retrospective cross-sectional	7	Medium quality
Drapkina et al [39], 2021	Russia	Prospective cross-sectional	7	Medium quality
Esaulenko et al [40], 2021	Russia	Prospective cross-sectional	7	Medium quality
Elagin et al [41], 2021-2022	Russia	Prospective cross-sectional	5	Medium quality

[a] AHRQ: Agency for Healthcare Research and Quality.

[b] NOS: Newcastle-Ottawa Scale.

Table 3. The effect of COVID-19 vaccines on semen parameters.

Source	Country	Vaccine type	Sample size; population	Age (y)	Semen test before vaccination	Semen test after vaccination, days	Outcomes	Side effects
Diaz et al [34], 2021-2022	United States	58% Moderna vaccines and 42% Pfizer vaccines	12; normal volunteers	Mean 26; range 25-30	Before the first dose of vaccine	T1[a]: 3 months after the first dose and T2[b]: at least 9 months after the second dose	No significant differences	Not reported
Gonzalez et al [35], 2020-2021	United States	21 received BNT162b2 and 24 received mRNA-1273	45; normal volunteers	Mean 28; range 25-31	Before the first dose of vaccine	75 (IQR 70-86) days after the second dose	After vaccination, the median sperm concentration, median TMSC[c], semen volume, and sperm motility were significantly increased	Not reported
Wang et al [43], 2021	China	Inactivated vaccine: CoronaVac (Sinovac)	153; patients with infertility undergoing IVF[d] treatment	Mean 33; range 29-37	Before the first dose of vaccine	Median 71 (IQR 38-102) days after the second dose	After vaccination, the amount of semen increased, the vitality decreased, and the normal morphology of sperm decreased	Not reported
Huang et al [44], 2021-2022	China	Inactivated vaccine: BBIBP-CorV (Sinopharm) or CoronaVac (Sinovac)	128; men with previous semen examination in the medical center	Mean 31; range 29-35	Before the first dose of vaccine	Median 87.5 (IQR 52.0-137.5) days after the second dose	No significant differences	Not reported
Zhu et al [45] (2020-2021)	China	Inactivated vaccine	43; semen donors from sperm bank	Mean 28.6 (SD 5.9)	Before the first dose of vaccine	T1: mean 9.1 (SD 8.1) days after the first dose and T2: mean 30.1 (SD 23.3) days after the second dose	No significant differences	A sore arm for both the first and second injections (nearly 13.9% and 4.7%, respectively). One donor reported feeling tired
Xia et al [46], 2021	China	Inactivated vaccine: BBIBP-CorV (Sinopharm) or CoronaVac (Sinovac)	Vaccine group (n=105) and control group (n=155); patients undergoing IVF treatment	Vaccine group: mean 33.9 (SD 4.7) and control group: mean 33.3 (SD 4.4)	Not tested	Unvaccinated group and vaccinated group: average 80.6 days after the first dose	No significant differences	Pain at the injection site (11.43%), followed by fatigue (6.67%), headache (1.90%), nausea (0.95%), and low-grade fever (0.95%)
Dong et al [47], 2023	China	Inactivated vaccine	519; fertile men	Mean 35.3 (SD 4.3)	Not tested	Unvaccinated group and vaccinated group: ≤90 days or >90 days after the first dose	No significant differences	The main adverse effect was injection site pain or redness; percentage of participants with fever (first dose: 0.3%, second dose: 0.6%, and booster dose: 1.9%)
Abd et al [51], 2022	Iraq	Pfizer-BioNTech (BNT162b2)	60; men undergoing IVF treatment (infertility due to clear female factors)	Mean 37.2 (SD 5.3)	Before the first dose of vaccine	Mean 101 (SD 37) days after the first dose	Both total motility and progressive motility were slightly reduced after vaccination	Not reported

Source	Country	Vaccine type	Sample size; population	Age (y)	Semen test before vaccination	Semen test after vaccination, days	Outcomes	Side effects
Al-Alami et al [50], 2021	Jordan	Pfizer (44.1%), Sinopharm (31.4%), Astra Zeneca (7.9%), and Moderna (0.3%)	354; participants who visited 1 infertility unit	Mean 35.6 (SD 7.9)	Not tested	Unvaccinated group and vaccinated group: clear data were not available	The sperm concentration was higher among vaccinated participants than unvaccinated participants	Not reported
Alenzi et al [49], 2021	Saudi Arabia	Pfizer-BioNTech (44%), Oxford and Astra Zeneca (32%), and mixed (24%)	100; normal volunteers	Mean 37.89 (SD 10.3)	Before the first dose of vaccine	1 month after the second dose	There was a significant increase in progressive sperm motility: prevaccination motility: median 55.03 (IQR 42-61.75) vs postvaccination motility: median 56.5 (IQR 42.25-63)	Not reported
Lestari et al [48], 2020-2022	Indonesia	CoronaVac (Sinovac), Astra Zeneca, and Moderna: proportion is not clear	70; patients with infertility	Vaccine group: mean 29.8 (SD 3.5) and control group: mean 29.3 (SD 4.5)	Before the first dose of vaccine	2 to 4 weeks after the second dose	The viral-vector vaccine caused a decrease in morphology as well as an increase in DFI[e]	Not reported
Karavani et al [52], 2021-2022	Israel	Pfizer-BioNTech (BNT162b2)	58; men undergoing IVF treatment	Mean 38.0 (SD 6.1)	Before the first dose of vaccine	6 to 14 months after vaccination	No significant differences	Not reported
Gat et al [53], 2021-2022	Israel	Pfizer-BioNTech (BNT162b2)	37; semen donors from sperm banks	Mean 26.1 (SD 4.2)	Before the first dose of vaccine	T1, T2, and T3[f]: included 1 to 3 semen samples per donor who provided the samples 15 to 45 days, 75 to 125 days, and >145 days after vaccination	Compared with sperm concentration at T0[g], the concentration at T2 decreased by 15%, and total motor count decreased by 22%. Subsequently, T3 evaluation showed that the damage had recovered	Not reported
Lifshitz et al [54], 2021	Israel	Pfizer-BioNTech (BNT162b2)	75; normal volunteers, 40% of whom were medical staff	Mean 38.6 (SD 4.3)	Not tested	On average, 37 days after the second dose	Compared with the WHO[h] reference range, there are only 2 borderline sperm specimens	Fatigue (34.7%), pain at the injection site (13.3%), fever (9.3%), and chills (8%)
Safrai et al [55], 2021	Israel	Pfizer-BioNTech (BNT162b2)	72; patients undergoing IVF treatment	Mean 35.7; range 33-43	Before the first dose of vaccine	Median 71.0 (IQR 40.5-104.8) days after the first dose	No significant differences	Not reported
Barda et al [56], 2020-2021	Israel	Pfizer-BioNTech (BNT162b2)	33; semen donors from sperm banks	27	Before the first dose of vaccine	The average sperm parameters of multiple samples were used for comparison. For each donor, at least 1 sample was received ≥72 days after the second dose	Total sperm count and TMSC increased after the second dose of vaccine	Pain at injection site: 79% (T1) and 88% (T2) of participants, lethargy: 9% (T1) and 48% (T2) of participants, and fever: 1 case (T1)

Source	Country	Vaccine type	Sample size; population	Age (y)	Semen test before vaccination	Semen test after vaccination, days	Outcomes	Side effects
Orvieto et al [57], 2021	Israel	Pfizer-BioNTech (BNT162b2)	36; patients with infertility undergoing IVF treatment	Mean 40.1 (SD 4.8)	Before the first dose of vaccine	Mean 33.3 (SD 14.9) days after the second dose	No significant differences	Not reported
Olana et al [36], 2021	Italy	Pfizer-BioNTech (BNT162b2)	47; normal volunteers	Mean 29.3 (SD 6.0)	Before the first dose of vaccine	70 days from the second dose	No significant differences	Not reported
Reschini et al [37], 2022	Italy	Pfizer-BioNTech (69%), Moderna (19%), Oxford and Astra Zeneca (9%), Johnson & Johnson's Janssen (1%), and mixed (2%)	106; fertile men undergoing ART[i] programs	Mean 39; range 36-42	Before the first dose of vaccine	Median 75 (IQR 38-112) days after the first dose	No significant differences	45% of the patients reported mild, self-resolving adverse events after the vaccine, including pain at injection site, fever, fatigue, nausea, muscle pain, and diarrhea
Massarotti et al [38], 2021	Italy	76% received mRNA vaccines, 20% viral-vector vaccines, 2% a mixed formulation, and 2% were not clear about the type	101; men undergoing fertility treatments	Mean 37.5 (SD 5.5)	Before the first dose of vaccine	Mean 2.3 (SD 1.5) months after the second dose	After vaccination, the median volume of the sample decreased from 3.0 to 2.6 ml. Sperm concentration, progressive motility, and TMSC increased	Not reported
Chillon et al [42], 2021	Germany	Pfizer-BioNTech (not reported)	86; normal volunteers	Vaccine group: mean 38 (SD 5.5) and control group: mean 36 (SD 6.5)	Not tested	Unvaccinated group and vaccinated group: clear data were not available	SARS-CoV-2 vaccination parameters and vaccine-induced antibodies were not associated with sperm parameters	Not reported
Drapkina et al [39], 2021	Russia	Viral-vector vaccine: Sputnik V	45; normal volunteers	Mean 36; range 30-44	Before the first dose of vaccine	90 days after the first dose	No significant differences	Not reported
Esaulenko et al [40], 2021	Russia	Viral-vector vaccine: Sputnik V	30; patients with type 2 diabetes	Mean 46; range 42-48	Before the first dose of vaccine	90 days after the first dose	No significant differences	19 (63%) patients demonstrated a temperature rise and 26 (87%) patients reported tenderness at injections site

Source	Country	Vaccine type	Sample size; population	Age (y)	Semen test before vaccination	Semen test after vaccination, days	Outcomes	Side effects
Elagin et al [41], 2021-2022	Russia	Viral-vector vaccine: (Sputnik V)	44; normal volunteers	Mean 22.4 (SD 4.7)	Before the first dose of vaccine	10-12 days after the first dose and 32-36 days after the second dose	No significant differences	A total of 5 and 7 participants exhibited fever with headache after the first and second vaccine doses, respectively

[a] T1: time point 1.

[b] T2: time point 2.

[c] TMSC: total motile sperm count.

[d] IVF: in vitro fertilization.

[e] DFI: DNA fragmentation index.

[f] T3: time point 3.

[g] T0: time point 0.

[h] WHO: World Health Organization.

[i] ART: assisted reproductive technology.

Discussion

Principal Findings

Safrai et al [55] published the first study on the impact of the COVID-19 vaccine on sperm parameters, demonstrating that the BNT162b2 mRNA vaccine (Pfizer or BioNTech) was not associated with a decrease in sperm quality. The study enrolled 72 individuals (57 of whom had normal sperm parameters). Sperm parameters in men with normal and abnormal sperm analysis did not change significantly after vaccination. The authors noted that after COVID-19 vaccination, none of the sperm parameters changed significantly [55]. In February and March 2021, Lifshitz et al [54] evaluated the effects of the Pfizer COVID-19 vaccine in 75 men with proven fertility [54]. The primary outcome was the percentage of abnormal sperm parameters in men 1 to 2 months after the second injection of the COVID-19 vaccine, including abnormal sperm morphology, reduced percentage of motile spermatozoa, and oligozoospermia rates. The results showed that the semen parameters after COVID-19 vaccination were mostly within the normal reference ranges established by the World Health Organization and did not indicate any causative negative effects of the COVID-19 vaccination. However, the participants in this study were a relatively homogeneous group, consisting of fertile male individuals from high socioeconomic groups. Moreover, participants were only followed up for 1 to 2 months after receiving the second dose of vaccine; therefore, long-term results have not yet been reported. In contrast, in a retrospective, longitudinal, multicenter study recently published in *Andrology*, the effects of the COVID-19 BNT162b2 vaccine on semen parameters were evaluated in 37 semen donors at different time points, before and after immunization [53]. Compared with the prevaccination levels, there was a selective decrease in sperm concentration and TMSC 75 to 125 days after vaccination (P=.01 and P=.007, respectively). Normal levels of these parameters were restored 145 days after vaccination, and no changes in sperm volume or motility were observed. Similarly, Abd et al [51] observed that the total motility and progressive sperm activity decreased significantly after BNT162b2 injection. Notably, considering that all semen parameters were still within the normal range, the investigator considered that the vaccine had no deleterious effects on semen parameters [51]. An additional prospective cohort study evaluated the effects of an inactivated vaccine (CoronaVac) on semen quality and in vitro fertilization (IVF) outcomes. A total of 542 patients undergoing IVF were divided into the unexposed group (nonvaccination) and the exposed group (vaccination), and it was found that the semen parameters seemed to fluctuate in the vaccinated men, including increased semen volume, decreased normal forms of sperm, and lower motility, whereas the motile sperm counts were similar. In addition, all semen parameters were above the lower reference limit. Therefore, researchers jointly believed that inactivated vaccines may not negatively affect sperm indicators or embryo development potential in men [43]. However, although this was a cohort study with a relatively large sample size, the male participants were all infertile, limiting the generalizability of the findings.

Surprisingly, significant improvements in some sperm parameters were noted after the COVID-19 vaccination. Massarotti et al [38] conducted a prospective study in which 101 vaccinated men (who were undergoing fertility treatments) had semen testing before vaccination and 2.3 (SD 1.5) months after the second dose of vaccine (76% of the participants were vaccinated with mRNA vaccine and 20% with viral-vectored vaccine). The investigators found a significant reduction in the median volume of the sample (from 3.0 to 2.6 mL), whereas sperm concentration, progressive motility, and TMSC increased (from 25.0 to 43.0 million/mL, from 50% to 56%, and from 34.8 to 54.6 million, respectively) [38]. In another prospective cohort study published in JAMA, Gonzalez et al [35] evaluated the effect of mRNA vaccines in 45 healthy volunteers (21 Pfizer and 24 Moderna) and found no significant decrease in sperm parameters. Semen samples were collected before the first dose and a median of 75 days after the second dose [35]. The baseline median sperm concentration was 26 million/mL, and the TMSC was 36 million. After the second vaccine dose, the median sperm concentration increased to 30 million/mL, and the median TMSC concentration increased to 44 million. Sperm motility and semen volume also increased significantly. The authors suggested that the increase in sperm parameters might be explained by known individual variations in sperm and increased abstinence time before postvaccine sample collection. Similarly, Barda et al [56] evaluated the effect of 2 doses of the vaccine (BNT162b2) on the sperm parameters of 33 sperm donors who donated sperm on multiple occasions. Compared with the percentage of motile sperm in the sample before vaccination, the percentage after the second dose of the vaccine did not change. However, the total sperm and motile counts increased unexpectedly.[56]. One possible explanation is that some participants changed their lifestyle habits or were more health conscious than the retrospectively enrolled infertile population, which acted as a confounding factor and positively affected sperm quality [58].

Two previous meta-analyses examined the effect of COVID-19 vaccination on sperm parameters based on published data [59,60]. Both studies suggested that the vaccination had no discernible negative impact on sperm quality. However, the 2 meta-analyses had nonnegligible limitations: the included studies used different types of vaccines, administered different doses of vaccines, had mostly unknown time intervals from vaccination to sperm analysis, had large differences in the participating populations, and had a small sample size for each group after the subgroup analysis. In other words, the methodological quality of meta-analyses assessing the impact of COVID-19 vaccination on semen parameters is currently unsatisfactory. As new controlled trials and prospective cohort studies with larger sample sizes emerge, the possibility of a negative effect of the COVID-19 vaccine on sperm quality must be further clarified.

Commonly reported side effects of the COVID-19 vaccine in premarketing clinical trials and postmarketing surveillance programs include systemic reactions (eg, fatigue, headache, and muscle pain) as well as injection site reactions (eg, pain, redness, and swelling), with rare serious adverse events [61,62]. Furthermore, fever is considered one of the most common side effects of vaccination [63]. According to the literature, fever can temporarily impair sperm parameters, causing a decline in sperm count, quality, and DNA integrity [64]. It should be noted that a range of studies have demonstrated fever to be a common adverse effect following multiple vaccine injections including mRNA vaccines (eg, BNT162b2), inactivated vaccines (eg, CoronaVac), and viral-vector vaccines (eg, Ad26. Sputnik V). Therefore, male patients who experience fever after receiving the COVID-19 vaccine may experience fluctuations in their sperm parameters. However, this effect is similar to that caused by fever due to other registered vaccines [41,47,54].

According to a study by Gat et al [53], the systemic immune response following BNT162b2 immunization may contribute to the deterioration of selective temporary sperm concentration and TMSC; however, the long-term prognosis remains favorable [53]. It is important to note that similar studies did not have data on the detailed health and immune status of the participants before vaccination. Therefore, their conclusions should be interpreted with caution. Carto et al [65] examined data from a large US electronic health record database on the risk of orchitis, epididymitis, or both in men vaccinated against COVID-19 [65]. They discovered that the serum levels of antisperm antibodies and antiphospholipid antibodies against cardiolipin, phosphatidylserine, annexin V, and 2-glycoprotein-1 did not differ significantly before and after vaccination. Furthermore, the levels of these autoimmune antibodies did not correlate with the sperm parameters. Their findings even suggested that COVID-19 vaccination was linked to a lower risk of orchitis, epididymitis, or both (odds ratio 0.568, 95% CI 0.497-0.649). More directly, Chillon et al [42] examined antibody concentrations in the seminal plasma of 86 men vaccinated with the new Crown pneumonia vaccine and showed that the antibody levels in seminal plasma after COVID-19 vaccination were correlated with serum antibody titers, but not with sperm parameters.

In addition, BNT162b2 and mRNA-1273 are lipid nanoparticle-formulated vaccines against SARS-CoV-2. These excipients, adjuvants, or both may act as vehicles for therapeutic content delivery while also increasing the intensity of the immune response [22]. Several studies have shown that these nanoparticles can cross biological barriers and can be deposited in reproductive organs, such as the testes, impairing sperm quality by increasing inflammation, damaging DNA structure, decreasing mitochondrial function, and inducing apoptosis [66]. Therefore, some researchers are concerned that vaccine excipients, adjuvants, or both may induce testicular damage. However, Olana et al [36] found no significant differences in the levels of reactive oxygen metabolites, electrolytes, or interleukin-6 in the seminal plasma of participants before and after the BNT162b2 vaccination [36]. Indeed, the current basic research evidence and clinical data supporting the idea that a systemic immune response after COVID-19 vaccination can lead to impaired sperm quality are insufficient.

Although semen analysis serves as the foundation for evaluating male fertility, it should be noted that it is an imperfect predictor of fertility potential. Further persuasive evidence comes from ART and artificial insemination by partner outcomes in couples after receiving the COVID-19 vaccine [57,67]. In a prospective cohort study involving 735 infertile couples, Dong et al [47] demonstrated that the quality of the embryos and the pregnancy rate in IVF treatment were unaffected by the couples' vaccination status or vaccine type. Similarly, no differences in fertility or pregnancy outcomes were observed in the study by Orvieto et al [57]. In another multicenter prospective study of 4185 couples, Wang et al [68] found no association between COVID-19 vaccination status (eg, inactivated, adenoviral, and recombinant vaccines) and pregnancy rates in artificial insemination by partner treatment [68].

More directly, using data from an internet-based preconception cohort study, Wesselink et al [69] investigated the relationship between COVID-19 vaccination and SARS-CoV-2 infection with fertility in couples attempting spontaneous pregnancy [69]. Their results suggested that male SARS-CoV-2 infection may be associated with a short-term decline in fertility, whereas COVID-19 vaccination had no effect on fertility in either partner.

Limitations

It is important to acknowledge that this review has a few limitations. First, data from most relevant studies were observational and lacked complete follow-up and analysis of missing data. All the included studies assessed vaccination based on self-reports or with no descriptions. Although previous influenza vaccination validation studies found 97% agreement between self-report and medical records, the conclusion may not be directly generalizable to COVID-19 vaccination, and misclassification risk may still exist [70]. Second, there is a need for more robust studies, with more precise eligibility criteria, appropriate sample sizes, and a more representative population, not focusing only on specific groups, such as those undergoing IVF. Third, despite the transient alterations in sperm parameters found in some studies, these findings were often not well supported by histological or pathophysiological data. Furthermore, the potential implications for sperm biology, including mitochondrial function, and the possible effects on the regulatory systems of epigenetic changes have received little attention.

However, given the widespread and ongoing spread of COVID-19, the available data on decreased semen quality in survivors of COVID-19 outweigh concerns about the potential negative effects of the COVID-19 vaccine on sperm parameters [71,72]. From the perspective of reproduction, vaccination against COVID-19 is preferable for severe adverse symptoms of the SARS-CoV-2 infection. Large-scale prospective cohort studies and randomized controlled trials of existing COVID-19 vaccines and newly developed vaccines are needed to further confirm this review's conclusion.

Conclusions

In summary, the data in this review show that the COVID-19 vaccine is safe for male reproductive health. Serious side effects of the COVID-19 vaccine are extremely rare, and men experience few problems with sperm parameters or reproductive potential after vaccination. Considering that SARS-CoV-2 infection itself may be associated with impaired fertility, vaccination could serve as a potential tool to preserve male fertility by preventing COVID-19. Therefore, vaccination should be clearly recommended for all men wishing to have children unless there are additional contraindications.

Acknowledgments

The authors would like to thank the researchers from Anhui Medical University. This work was supported by The National Natural Science Foundation of China (number 82201777), the Key Project of Natural Science Research in the Universities of Anhui Province (number KJ2021A0296), and Basic and Clinical Collaborative Research Program of Anhui Medical University (number 2022xkjT020). YX C, XJ H, QS S, CW, and BS should be considered equal contributors.

Data Availability

All data generated or analyzed during this study are included in this published paper (and its supplementary information files).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Filled-in PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[DOCX File , 35 KB - publichealth_v9i1e48511_app1.docx]

Multimedia Appendix 2

Search strategy in electronic databases.

[DOCX File , 19 KB - publichealth_v9i1e48511_app2.docx]

References

1. Sachs JD, Karim SS, Aknin L, Allen J, Brosbøl K, Colombo F, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. Lancet 2022 Oct 08;400(10359):1224-1280 [FREE Full text] [doi: 10.1016/S0140-6736(22)01585-9] [Medline: 36115368]
2. Li G, Tang D, Song B, Wang C, Qunshan S, Xu C, et al. Impact of the COVID-19 pandemic on partner relationships and sexual and reproductive health: cross-sectional, online survey study. J Med Internet Res 2020 Aug 06;22(8):e20961 [FREE Full text] [doi: 10.2196/20961] [Medline: 32716895]
3. Han L, Zhan Y, Li W, Xu Y, Xu Y, Zhao J. Associations between the perceived severity of the COVID-19 pandemic, cyberchondria, depression, anxiety, stress, and lockdown experience: cross-sectional survey study. JMIR Public Health Surveill 2021 Sep 16;7(9):e31052 [FREE Full text] [doi: 10.2196/31052] [Medline: 34478402]
4. Dong Y, Dai T, Wang B, Zhang L, Zeng L, Huang J, et al. The way of SARS-CoV-2 vaccine development: success and challenges. Signal Transduct Target Ther 2021 Nov 09;6(1):387 [FREE Full text] [doi: 10.1038/s41392-021-00796-w] [Medline: 34753918]

5. Graña C, Ghosn L, Evrenoglou T, Jarde A, Minozzi S, Bergman H, et al. Efficacy and safety of COVID-19 vaccines. Cochrane Database Syst Rev 2022 Dec 07;12(12):CD015477 [FREE Full text] [doi: 10.1002/14651858.CD015477] [Medline: 36473651]
6. Roghani A. The influence of COVID-19 vaccination on daily cases, hospitalization, and death rate in Tennessee, United States: case study. JMIRx Med 2021;2(3):e29324 [FREE Full text] [doi: 10.2196/29324] [Medline: 34424255]
7. Kerslake R, Hall M, Randeva HS, Spandidos DA, Chatha K, Kyrou I, et al. Co-expression of peripheral olfactory receptors with SARS-CoV-2 infection mediators: potential implications beyond loss of smell as a COVID-19 symptom. Int J Mol Med 2020 Sep;46(3):949-956 [FREE Full text] [doi: 10.3892/ijmm.2020.4646] [Medline: 32705281]
8. Xie Y, Mirzaei M, Kahrizi MS, Shabestari AM, Riahi SM, Farsimadan M, et al. SARS-CoV-2 effects on sperm parameters: a meta-analysis study. J Assist Reprod Genet 2022 Jul;39(7):1555-1563 [FREE Full text] [doi: 10.1007/s10815-022-02540-x] [Medline: 35716338]
9. Corona G, Vena W, Pizzocaro A, Pallotti F, Paoli D, Rastrelli G, et al. Andrological effects of SARS-Cov-2 infection: a systematic review and meta-analysis. J Endocrinol Invest 2022 Dec;45(12):2207-2219 [FREE Full text] [doi: 10.1007/s40618-022-01801-x] [Medline: 35527294]
10. Yiannakoulias N, Darlington JC, Slavik CE, Benjamin G. Negative COVID-19 vaccine information on Twitter: content analysis. JMIR Infodemiology 2022 Aug 29;2(2):e38485 [FREE Full text] [doi: 10.2196/38485] [Medline: 36348980]
11. Himelein-Wachowiak M, Giorgi S, Devoto A, Rahman M, Ungar L, Schwartz HA, et al. Bots and misinformation spread on social media: implications for COVID-19. J Med Internet Res 2021 May 20;23(5):e26933 [FREE Full text] [doi: 10.2196/26933] [Medline: 33882014]
12. Zhao S, Hu S, Zhou X, Song S, Wang Q, Zheng H, et al. The prevalence, features, influencing factors, and solutions for COVID-19 vaccine misinformation: systematic review. JMIR Public Health Surveill 2023 Jan 11;9:e40201 [FREE Full text] [doi: 10.2196/40201] [Medline: 36469911]
13. -. COVID vaccine confidence requires radical transparency. Nature 2020 Oct;586(7827):8. [doi: 10.1038/d41586-020-02738-y] [Medline: 32994572]
14. Carpiano RM, Callaghan T, DiResta R, Brewer NT, Clinton C, Galvani AP, et al. Confronting the evolution and expansion of anti-vaccine activism in the USA in the COVID-19 era. Lancet 2023 Mar 18;401(10380):967-970 [FREE Full text] [doi: 10.1016/S0140-6736(23)00136-8] [Medline: 36871571]
15. Karafillakis E, Van Damme P, Hendrickx G, Larson HJ. COVID-19 in Europe: new challenges for addressing vaccine hesitancy. Lancet 2022 Feb 19;399(10326):699-701 [FREE Full text] [doi: 10.1016/S0140-6736(22)00150-7] [Medline: 35123665]
16. Lee M, You M. Direct and indirect associations of media use with COVID-19 vaccine hesitancy in South Korea: cross-sectional web-based survey. J Med Internet Res 2022 Jan 06;24(1):e32329 [FREE Full text] [doi: 10.2196/32329] [Medline: 34870605]
17. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, et al. Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. J Med Internet Res 2021 Jun 11;23(6):e27632 [FREE Full text] [doi: 10.2196/27632] [Medline: 34061757]
18. GeurtsvanKessel CH, de Vries RD. Evaluating novel COVID-19 vaccines in the current chapter of the pandemic. Lancet Infect Dis 2022 Dec;22(12):1652-1654 [FREE Full text] [doi: 10.1016/S1473-3099(22)00517-5] [Medline: 36075232]
19. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in metaanalyses. The Ottawa Hospital Research Institut. URL: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2023-11-01]
20. Zeng X, Zhang Y, Kwong JS, Zhang C, Li S, Sun F, et al. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. J Evid Based Med 2015 Feb 25;8(1):2-10. [doi: 10.1111/jebm.12141] [Medline: 25594108]
21. Luo WR, Wu XM, Wang W, Yu JL, Chen QQ, Zhou X, et al. Novel coronavirus mutations: vaccine development and challenges. Microb Pathog 2022 Dec;173(Pt A):105828 [FREE Full text] [doi: 10.1016/j.micpath.2022.105828] [Medline: 36243381]
22. Jafari A, Danesh Pouya F, Niknam Z, Abdollahpour-Alitappeh M, Rezaei-Tavirani M, Rasmi Y. Current advances and challenges in COVID-19 vaccine development: from conventional vaccines to next-generation vaccine platforms. Mol Biol Rep 2022 Jun;49(6):4943-4957 [FREE Full text] [doi: 10.1007/s11033-022-07132-7] [Medline: 35235159]
23. Rando HM, Lordan R, Lee AJ, Naik A, Wellhausen N, Sell E, et al. Application of traditional vaccine development strategies to SARS-CoV-2. mSystems 2023 Apr 27;8(2):e0092722 [FREE Full text] [doi: 10.1128/msystems.00927-22] [Medline: 36861991]
24. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
25. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416 [FREE Full text] [doi: 10.1056/NEJMoa2035389] [Medline: 33378609]

26. Lamb YN. BNT162b2 mRNA COVID-19 vaccine: first approval. Drugs 2021 Mar;81(4):495-501 [FREE Full text] [doi: 10.1007/s40265-021-01480-7] [Medline: 33683637]
27. Sahin U, Muik A, Derhovanessian E, Vogler I, Kranz LM, Vormehr M, et al. COVID-19 vaccine BNT162b1 elicits human antibody and T1 T cell responses. Nature 2020 Oct;586(7830):594-599. [doi: 10.1038/s41586-020-2814-7] [Medline: 32998157]
28. Baldo A, Leunda A, Willemarck N, Pauwels K. Environmental risk assessment of recombinant viral vector vaccines against SARS-Cov-2. Vaccines (Basel) 2021 May 03;9(5):453 [FREE Full text] [doi: 10.3390/vaccines9050453] [Medline: 34063733]
29. Dal-Ré R, Banzi R, Becker SL, Launay O, Pavli A. High heterogeneity on the accepted vaccines for COVID-19 certificates in European countries. Travel Med Infect Dis 2022 Jul;48:102321 [FREE Full text] [doi: 10.1016/j.tmaid.2022.102321] [Medline: 35390530]
30. Vellozzi C, Burwen DR, Dobardzic A, Ball R, Walton K, Haber P. Safety of trivalent inactivated influenza vaccines in adults: background for pandemic influenza vaccine safety monitoring. Vaccine 2009 Mar 26;27(15):2114-2120. [doi: 10.1016/j.vaccine.2009.01.125] [Medline: 19356614]
31. Khoshnood S, Arshadi M, Akrami S, Koupaei M, Ghahramanpour H, Shariati A, et al. An overview on inactivated and live-attenuated SARS-CoV-2 vaccines. J Clin Lab Anal 2022 May;36(5):e24418 [FREE Full text] [doi: 10.1002/jcla.24418] [Medline: 35421266]
32. Cheng H, Peng Z, Luo W, Si S, Mo M, Zhou H, et al. Efficacy and safety of COVID-19 vaccines in phase III trials: a meta-analysis. Vaccines (Basel) 2021 Jun 01;9(6):582 [FREE Full text] [doi: 10.3390/vaccines9060582] [Medline: 34206032]
33. Wang X, Wang H, Du A, Wang J, Shi J, Zhang Y, et al. COVID-19 vaccine hesitancy and associated factors among infertile couples undergoing assisted reproductive treatment. Front Immunol 2022 Sep 27;13:973600 [FREE Full text] [doi: 10.3389/fimmu.2022.973600] [Medline: 36238271]
34. Diaz P, Dullea A, Patel M, Blachman-Braun R, Reddy R, Khodamoradi K, et al. Long-term evaluation of sperm parameters after coronavirus disease 2019 messenger ribonucleic acid vaccination. F S Rep 2022 Sep;3(3):211-213 [FREE Full text] [doi: 10.1016/j.xfre.2022.07.007] [Medline: 35936108]
35. Gonzalez DC, Nassau DE, Khodamoradi K, Ibrahim E, Blachman-Braun R, Ory J, et al. Sperm parameters before and after COVID-19 mRNA vaccination. JAMA 2021 Jul 20;326(3):273-274 [FREE Full text] [doi: 10.1001/jama.2021.9976] [Medline: 34137808]
36. Olana S, Mazzilli R, Salerno G, Zamponi V, Tarsitano MG, Simmaco M, et al. 4BNT162b2 mRNA COVID-19 vaccine and semen: what do we know? Andrology 2022 Sep;10(6):1023-1029 [FREE Full text] [doi: 10.1111/andr.13199] [Medline: 35647664]
37. Reschini M, Pagliardini L, Boeri L, Piazzini F, Bandini V, Fornelli G, et al. COVID-19 vaccination does not affect reproductive health parameters in men. Front Public Health 2022 Feb 02;10:839967 [FREE Full text] [doi: 10.3389/fpubh.2022.839967] [Medline: 35186854]
38. Massarotti C, Stigliani S, Maccarini E, Bovis F, Ferraro MF, Gazzo I, et al. mRNA and viral vector COVID-19 vaccines do not affect male fertility: a prospective study. World J Mens Health 2022 Oct;40(4):561-569 [FREE Full text] [doi: 10.5534/wjmh.220055] [Medline: 36047075]
39. Drapkina YS, Dolgushina NV, Shatylko TV, Nikolaeva MA, Menzhinskaya IV, Ivanets TY, et al. Gam-COVID-Vac (Sputnik V) vaccine has no adverse effect on spermatogenesis in men. Akush Ginekol 2021 Jul 23;7_2021:88-94 [FREE Full text] [doi: 10.18565/aig.2021.7.88-94]
40. Esaulenko DI, Rozhivanov RR, Shishkina VV. Quality of ejaculate of male patients with type 2 diabetes mellitus (DM), vaccinated with GamCovidVac (Sputnik V). Diabetes mellitus 2022 Jan 04;24(5):422-426 [FREE Full text] [doi: 10.14341/dm12775]
41. Elagin V, Adamyan L, Vechorko V, Doroshenko D, Dashko A, Filippov O, et al. COVID-19 vaccine and male reproductive health (preliminary data). Probl Reprod 2021;27(4):17. [doi: 10.17116/repro20212704117]
42. Chillon TS, Demircan K, Weiss G, Minich WB, Schenk M, Schomburg L. Detection of antibodies to SARS-CoV-2 after vaccination in seminal plasma and their association to sperm parameters. Int J Infect Dis 2023 May;130:161-165 [FREE Full text] [doi: 10.1016/j.ijid.2023.03.018] [Medline: 36933610]
43. Wang M, Yang Q, Zhu L, Jin L. Investigating impacts of CoronaVac vaccination in males on in vitro fertilization: a propensity score matched cohort study. World J Mens Health 2022 Oct;40(4):570-579 [FREE Full text] [doi: 10.5534/wjmh.220017] [Medline: 36047069]
44. Huang J, Xia L, Tian L, Xu D, Fang Z, Lin J, et al. Comparison of semen quality before and after inactivated SARS-CoV-2 vaccination among men in China. JAMA Netw Open 2022 Sep 01;5(9):e2230631 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.30631] [Medline: 36074470]
45. Zhu H, Wang X, Zhang F, Zhu Y, Du M, Tao ZW, et al. Evaluation of inactivated COVID-19 vaccine on semen parameters in reproductive-age males: a retrospective cohort study. Asian J Androl 2022 Sep;24(5):441-444 [FREE Full text] [doi: 10.4103/aja202225] [Medline: 35532560]

46. Xia W, Zhao J, Hu Y, Fang L, Wu S. Investigate the effect of COVID-19 inactivated vaccine on sperm parameters and embryo quality in in vitro fertilization. Andrologia 2022 Jul;54(6):e14483 [FREE Full text] [doi: 10.1111/and.14483] [Medline: 35610731]
47. Dong Y, Li X, Li Z, Zhu Y, Wei Z, He J, et al. Effects of inactivated SARS-CoV-2 vaccination on male fertility: a retrospective cohort study. J Med Virol 2023 Jan;95(1):e28329. [doi: 10.1002/jmv.28329] [Medline: 36415120]
48. Lestari SW, Restiansyah G, Yunihastuti E, Pratama G. Comparison of sperm parameters and DNA fragmentation index between infertile men with infection and vaccines of COVID-19. Asian J Androl 2023 Apr 21;25(5):578-582 [FREE Full text] [doi: 10.4103/aja202310] [Medline: 37102509]
49. Alenzi MJ, Ikpi EE, Alshlash AS, Alkhaldi SM, Aljazi ZN, Alrawili M, et al. Do the COVID-19 vaccines affect the reproductive parameters of men?-a Northern Saudi experience. Niger J Clin Pract 2022 Nov;25(11):1812-1815 [FREE Full text] [doi: 10.4103/njcp.njcp_188_22] [Medline: 36412287]
50. Al-Alami ZM, Albeitawi S, ALNatsheh MS, Albakri K, Qublan H, Muhaidat N, et al. COVID-19 and semen fluid parameters, a retrospective study from infertility clinics. Life (Basel) 2022 Dec 10;12(12):2076 [FREE Full text] [doi: 10.3390/life12122076] [Medline: 36556441]
51. Abd ZH, Muter SA, Saeed RA, Ammar O. Effects of COVID-19 vaccination on different semen parameters. Basic Clin Androl 2022 Aug 02;32(1):13 [FREE Full text] [doi: 10.1186/s12610-022-00163-x] [Medline: 35915409]
52. Karavani G, Chill HH, Meirman C, Gutman-Ido E, Herzberg S, Tzipora T, et al. Sperm quality is not affected by the BNT162b2 mRNA SARS-CoV-2 vaccine: results of a 6-14 months follow-up. J Assist Reprod Genet 2022 Oct;39(10):2249-2254 [FREE Full text] [doi: 10.1007/s10815-022-02621-x] [Medline: 36114906]
53. Gat I, Kedem A, Dviri M, Umanski A, Levi M, Hourvitz A, et al. Covid-19 vaccination BNT162b2 temporarily impairs semen concentration and total motile count among semen donors. Andrology 2022 Sep;10(6):1016-1022 [FREE Full text] [doi: 10.1111/andr.13209] [Medline: 35713410]
54. Lifshitz D, Haas J, Lebovitz O, Raviv G, Orvieto R, Aizer A. Does mRNA SARS-CoV-2 vaccine detrimentally affect male fertility, as reflected by semen analysis? Reprod Biomed Online 2022 Jan;44(1):145-149 [FREE Full text] [doi: 10.1016/j.rbmo.2021.09.021] [Medline: 34815157]
55. Safrai M, Herzberg S, Imbar T, Reubinoff B, Dior U, Ben-Meir A. The BNT162b2 mRNA COVID-19 vaccine does not impair sperm parameters. Reprod Biomed Online 2022 Apr;44(4):685-688 [FREE Full text] [doi: 10.1016/j.rbmo.2022.01.008] [Medline: 35279377]
56. Barda S, Laskov I, Grisaru D, Lehavi O, Kleiman S, Wenkert A, et al. The impact of COVID-19 vaccine on sperm quality. Int J Gynaecol Obstet 2022 Jul;158(1):116-120 [FREE Full text] [doi: 10.1002/ijgo.14135] [Medline: 35128663]
57. Orvieto R, Noach-Hirsh M, Segev-Zahav A, Haas J, Nahum R, Aizer A. Does mRNA SARS-CoV-2 vaccine influence patients' performance during IVF-ET cycle? Reprod Biol Endocrinol 2021 May 13;19(1):69 [FREE Full text] [doi: 10.1186/s12958-021-00757-6] [Medline: 33985514]
58. Montano L, Ceretti E, Donato F, Bergamo P, Zani C, Viola GC, FASt study group. Effects of a lifestyle change intervention on semen quality in healthy young men living in highly polluted areas in Italy: the FASt randomized controlled trial. Eur Urol Focus 2022 Jan;8(1):351-359 [FREE Full text] [doi: 10.1016/j.euf.2021.01.017] [Medline: 33579652]
59. Omolaoye TS, Velilla PA, Moncada López JS, du Plessis SS, Cardona Maya WD. SARS-CoV-2 vaccine effects on semen parameters: a systematic review and meta-analysis. Arch Esp Urol 2022 Dec;75(10):831-843 [FREE Full text] [doi: 10.56434/j.arch.esp.urol.20227510.121] [Medline: 36651093]
60. Huang J, Fang Z, Huang L, Fan L, Liu Y, Xia L, et al. Effect of COVID-19 vaccination on semen parameters: a systematic review and meta-analysis. J Med Virol 2023 Jan;95(1):e28263 [FREE Full text] [doi: 10.1002/jmv.28263] [Medline: 36310390]
61. Al Khames Aga QA, Alkhaffaf WH, Hatem TH, Nassir KF, Batineh Y, Dahham AT, et al. Safety of COVID-19 vaccines. J Med Virol 2021 Dec 28;93(12):6588-6594 [FREE Full text] [doi: 10.1002/jmv.27214] [Medline: 34270094]
62. Rosenblum HG, Gee J, Liu R, Marquez PL, Zhang B, Strid P, et al. Safety of mRNA vaccines administered during the initial 6 months of the US COVID-19 vaccination programme: an observational study of reports to the Vaccine Adverse Event Reporting System and v-safe. Lancet Infect Dis 2022 Jun;22(6):802-812 [FREE Full text] [doi: 10.1016/S1473-3099(22)00054-8] [Medline: 35271805]
63. Kaplan RM, Milstein A. Influence of a COVID-19 vaccine's effectiveness and safety profile on vaccination acceptance. Proc Natl Acad Sci U S A 2021 Mar 09;118(10):e2021726118 [FREE Full text] [doi: 10.1073/pnas.2021726118] [Medline: 33619178]
64. Sergerie M, Mieusset R, Croute F, Daudin M, Bujan L. High risk of temporary alteration of semen parameters after recent acute febrile illness. Fertil Steril 2007 Oct;88(4):970.e1-970.e7 [FREE Full text] [doi: 10.1016/j.fertnstert.2006.12.045] [Medline: 17434502]
65. Carto C, Nackeeran S, Ramasamy R. COVID-19 vaccination is associated with a decreased risk of orchitis and/or epididymitis in men. Andrologia 2022 Mar;54(2):e14281 [FREE Full text] [doi: 10.1111/and.14281] [Medline: 34672002]
66. Alabi OA, Silva AH, Purnhagen LR, Souza GR, de Mello Júnior LJ, Filippin-Monteiro FB, et al. Genetic, reproductive and oxidative damage in mice triggered by co-exposure of nanoparticles: from a hypothetical scenario to a real concern. Sci Total Environ 2019 Apr 10;660:1264-1273. [doi: 10.1016/j.scitotenv.2019.01.036] [Medline: 30743921]

67. Dong M, Wu S, Zhang X, Zhao N, Qi J, Zhao D, et al. Effects of COVID-19 vaccination status, vaccine type, and vaccination interval on IVF pregnancy outcomes in infertile couples. J Assist Reprod Genet 2022 Aug;39(8):1849-1859 [FREE Full text] [doi: 10.1007/s10815-022-02543-8] [Medline: 35759062]
68. Wang C, Tang D, Liu J, Zhang S, Xu Y, Qiao J, et al. Association between COVID-19 vaccination and artificial insemination outcomes for couples experiencing infertility. JAMA Netw Open 2022 Dec 01;5(12):e2247216 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.47216] [Medline: 36525277]
69. Wesselink AK, Hatch EE, Rothman KJ, Wang TR, Willis MD, Yland J, et al. A prospective cohort study of COVID-19 vaccination, SARS-CoV-2 infection, and fertility. Am J Epidemiol 2022 Jul 23;191(8):1383-1395 [FREE Full text] [doi: 10.1093/aje/kwac011] [Medline: 35051292]
70. King JP, McLean HQ, Belongia EA. Validation of self-reported influenza vaccination in the current and prior season. Influenza Other Respir Viruses 2018 Nov;12(6):808-813 [FREE Full text] [doi: 10.1111/irv.12593] [Medline: 30028081]
71. Klepinowski T, Klepinowska M, Sagan L, Syrenicz A. Does SARS-CoV-2 affect human semen? A systematic review and meta-analysis. Arch Sex Behav 2023 Feb;52(2):669-677 [FREE Full text] [doi: 10.1007/s10508-022-02520-3] [Medline: 36602657]
72. Chen F, Zhu S, Dai Z, Hao L, Luan C, Guo Q, et al. Effects of COVID-19 and mRNA vaccines on human fertility. Hum Reprod 2021 Dec 27;37(1):5-13 [FREE Full text] [doi: 10.1093/humrep/deab238] [Medline: 34734259]

Abbreviations

ART: assisted reproductive technology
IVF: in vitro fertilization
mRNA: messenger RNA
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TMSC: total motile sperm count

Edited by A Mavragani, T Sanchez; submitted 26.04.23; peer-reviewed by H Zhu, B Ma; comments to author 02.06.23; revised version received 06.06.23; accepted 31.10.23; published 06.12.23.

Please cite as:
Li G, Zhang R, Song B, Wang C, Shen Q, He X, Cao Y
Effects of SARS-CoV-2 Vaccines on Sperm Quality: Systematic Review
JMIR Public Health Surveill 2023;9:e48511
URL: https://publichealth.jmir.org/2023/1/e48511
doi:10.2196/48511
PMID:37976132

Original Paper

Evaluation of Mpox Knowledge, Stigma, and Willingness to Vaccinate for Mpox: Cross-Sectional Web-Based Survey Among Sexual and Gender Minorities

Thiago Silva Torres[1*], PhD; Mayara Secco Torres Silva[1*], MD; Carolina Coutinho[1], PhD; Brenda Hoagland[1], MD, PhD; Emilia Moreira Jalil[1], MD, PhD; Sandra Wagner Cardoso[1], MD, PhD; Julio Moreira[2], BA; Monica Avelar Magalhaes[3], PhD; Paula Mendes Luz[1], MD, PhD; Valdilea G Veloso[1], MD, PhD; Beatriz Grinsztejn[1], MD, PhD

[1]Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

[2]Grupo Arco-Íris, Rio de Janeiro, Brazil

[3]Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz (ICICT-Fiocruz), Rio de Janeiro, Brazil

[*]these authors contributed equally

Corresponding Author:
Thiago Silva Torres, PhD
Instituto Nacional de Infectologia Evandro Chagas
Fundação Oswaldo Cruz
Av Brasil 4365 Manguinhos
Rio de Janeiro, 21040360
Brazil
Phone: 55 21996022998
Email: thiago.torres@ini.fiocruz.br

Abstract

Background: The 2022 multicountry mpox outbreak positioned the condition as a public health emergency of international concern. By May 2023, Brazil ranked second globally in the cumulative number of mpox cases and deaths. The higher incidence of mpox among gay and other men who have sex with men in the current mpox outbreak deepens the stigma and discrimination against sexual and gender minorities (SGM). This might worsen the structural barriers impacting access to health services, which ultimately leads to undertesting and underreporting of cases. There are no data available on mpox knowledge and stigma in Latin America.

Objective: We aimed to evaluate mpox knowledge, stigma, and willingness to vaccinate for mpox among SGM, and to describe sociodemographic and behavioral characteristics according to self-reported mpox diagnosis.

Methods: A cross-sectional, internet-based survey was conducted in a convenience sample of adults (aged >18 years) living in Brazil recruited through advertisements on dating apps, social media, referral institutions for infectious diseases websites, and mass media (October-November 2022). We compared participants' characteristics according to self-reported mpox diagnosis using chi-square test or Fisher exact test for qualitative variables and Kruskal-Wallis test for quantitative variables.

Results: We enrolled 6236 participants: 5685 (91.2%) were cisgender men; 6032 (96.7%) were gay, bisexual, or pansexual; 3877 (62.2%) were White; 4902 (78.7%) had tertiary education; and 4070 (65.2%) reported low or middle income. Most participants (n=5258, 84.4%) agreed or strongly agreed that "LGBTQIA+ individuals are being discriminated and stigmatized due to mpox." Mpox awareness was 96.9% (n=6044), and 5008 (95.1%) were willing to get vaccinated for mpox. Overall, 324 (5.2%) reported an mpox diagnosis. Among these, 318 (98.1%) reported lesions, 178 (56%) local pain, and 316 (99.4%) sought health care. Among participants not reporting a diagnosis, 288 (4.9%) had a suspicious lesion, but only 158 (54.9%) of these had sought health care. Compared to participants with no diagnosis, those reporting an mpox diagnosis were younger ($P<.001$), reported more sex partners ($P<.001$), and changes in sexual behavior after mpox onset ($P=.002$). Moreover, participants diagnosed with mpox reported more frequently being tested for HIV in the prior 3 months ($P<.001$), living with HIV ($P<.001$), currently using HIV pre-exposure prophylaxis ($P<.001$), and previous sexually transmitted infection diagnosis ($P<.001$).

Conclusions: Our results point to high mpox knowledge and willingness to vaccinate among SGM in Brazil. Participants self-reporting mpox diagnosis more frequently reported to be living with HIV, STI diagnosis, and current pre-exposure prophylaxis use, highlighting the importance of an mpox assessment that includes comprehensive sexual health screenings. Efforts to decrease stigma related to mpox among SGM are necessary to avoid mpox underdiagnosis.

(JMIR Public Health Surveill 2023;9:e46489) doi:10.2196/46489

KEYWORDS

sexual and gender minorities; mpox; Brazil; Latin America and Caribbean; community; LGBT; gay; lesbian; minority; minorities; survey; emergency; stigma; discrimination; Latin America; sociodemographic; behavioral; diagnosis; application; social media; infection; gender; MSM; monkeypox; infectious disease; awareness; patient education; health knowledge; patient knowledge; knowledge translation; access to care; viral

Introduction

The 2022 multicountry mpox outbreak positioned the condition as a public health emergency of international concern, as declared by the World Health Organization on July 23, 2022 [1]. Previously, mpox had been barely identified outside specific African countries with a known endemic epidemiology [2]. The first cases in nonendemic countries were announced in the United Kingdom, in May 2022, with over 87,000 confirmed mpox cases worldwide by May 23, 2023 [3,4]. The outbreak has disproportionally burdened vulnerable groups, especially gay, bisexual, and other cisgender men who have sex with men (MSM) [2,5]. Brazil ranked second globally in the cumulative number of mpox cases and deaths, with 10,941 confirmed diagnoses and 16 deaths as of May 23, 2023 [4]. Confirmed mpox cases in Brazil mostly occurred among cisgender MSM, with a potential role of sexual contact in the transmission dynamics [6-8].

Unfortunately, the higher incidence of mpox among vulnerable groups in the current mpox outbreak deepens the stigma and discrimination against MSM and other sexual and gender minorities (SGM) and might worsen the structural barriers impacting access to health services, which ultimately leads to undertesting and underreporting of cases [9-12]. Furthermore, the shame and fear related to societal judgment about individuals’ sexuality combined with the need for prolonged self-isolation have a significant impact on a person’s professional and daily life, thus undermining the mental health of the most vulnerable [13]. Moreover, this phenomenon might not be restricted to the current outbreak. Mpox-related stigma against SGM could be underestimated in countries with endemic mpox, as many of these still criminalize same-sex relationships, and data on cases among MSM might be inaccurate [14].

Efforts are in course to mitigate the impact of an mpox outbreak on the most vulnerable groups, and a stigma-free message that embraces sexual and gender-diverse people is key to establishing strong links connecting health services to the community. For this purpose, it is essential to analyze the knowledge among vulnerable populations about mpox so that communication and health actions can be tailored accordingly. A cross-sectional survey among 1932 respondents in the United Kingdom showed that, despite a vaccine acceptability rate as high as 96%, only 29% would seek a health service in case of mpox symptoms [15]. In the United States, 50% of MSM who responded to a web-based survey reported a 50% reduction in the number of sex partners, 1-time sexual encounters, group sex, and sex with a partner they met on a dating app or at a sex venue after the onset of the current mpox outbreak, although there was no change in condom use [16].

There are no data available on mpox knowledge in Latin America. This is the first survey in Brazil aiming to evaluate mpox knowledge, stigma, and willingness to vaccinate for mpox among SGM, and to describe sociodemographic and behavioral characteristics of study participants according to self-reported mpox diagnosis.

Methods

Study Design

This was a cross-sectional internet-based survey conducted in a convenience sample of adults (aged ≥18 years) living in Brazil, who were recruited through advertisements on dating apps (Grindr and Scruff), social media (Facebook and Instagram), referral institutions for infectious diseases websites or social media (Instituto Nacional de Infectologia Evandro Chagas and Fundação Oswaldo Cruz), radio, and mass media from October to November 2022. Banners in Scruff and Grindr and boosted posts on Facebook and Instagram, as in prior work [17,18], were used to recruit individuals to complete a voluntary open survey in Portuguese, with all items including a nonresponse option. The survey was programmed on Alchemer. Respondents were able to change or review answers and did not receive any compensation. Usability and technical functionality in different personal computers and mobile phone operating systems were confirmed before survey administration. We excluded respondents who did not reach the end of the survey and those who incorrectly answered any of the 5 attention questions [19]. Only 1 response per IP address was allowed.

Variables

Sociodemographic

Recruitment was categorized in Grindr, Scruff, Instagram, Facebook, and others (eg, WhatsApp and other websites). Age at the time of the survey was categorized in 4 intervals: 18-24; 25-30, 31-40, and >40 years. Gender categories were cisgender man, transgender man, cisgender woman, transgender woman, nonbinary or gender fluid, queer, and not declared. Sexual orientation categories were lesbian, gay, bisexual, pansexual, heterosexual, or other (eg, asexual and demisexual). Race was categorized as Asian (Japanese, Chinese, Korean, and among others), Black, Indigenous, Pardo (mixed race), White, or not declared. For schooling, the highest degree attained is considered (primary, secondary, or tertiary). Income reporting was based on the monthly income of the household in Brazilian minimum wages (MWs): low (up to 2 MW), middle (>2-6 MW), and high (>6 MW), in accordance with previous studies [18,20]. In 2022, MW per month was BRL 1212 (US $230). Participants were also asked whether they received any social support (eg, conditional cash transfers such as Auxílio Brasil) and the number of persons living in the same house (median, IQR). Brazilian regions were defined according to the Brazilian administrative division: North (7 states), Northeast (9 states), Central-west (3 states and Federal District), South (3 states), and Southeast (4 states). The state region was stratified into capital, metropolitan area, and other cities.

Mpox Awareness, Knowledge, Symptoms, and Contact

We defined mpox awareness as a positive answer to the question: “Have you ever heard about mpox?” Knowledge of mpox lesions was assessed with the question: “Do you know, or have you seen pictures of suspected mpox wounds/lesions?” Participants answered where they obtained information about mpox (eg, television, radio, Grindr, Hornet, Scruff, Google, other) and if they would follow social distancing if diagnosed with mpox (yes or no).

The survey also included questions about the participant’s health as it relates to any suspicious lesions related to mpox since the onset of mpox outbreak in Brazil (June 2022) (yes or no), location of the potential lesions, number of legions per location, if attended health facility to investigate lesions (yes or no), if lesions were painful, and graduation of pain from 0 (no pain) to 100 (worst pain). Participants were also asked if they had symptoms possibly related to mpox (predetermined list) since June 2022, and in case of any symptoms, if they attended a health facility to investigate them (yes or no).

Participants were asked about contact (sexual or nonsexual) with persons with suspected or confirmed mpox diagnosis, if traveled within Brazil or abroad since June 2022. Finally, for those self-reporting an mpox diagnosis, we inquired if they adhered to social distancing after receiving the diagnosis.

Sexual Behavior

Participants answered questions about their sexual behavior since June 2022, including gender of sex partners, number of sex partners (stratified in 0, 1, 2-5, 6-10, and >10), steady partners (yes or no and if only 1 steady partner), sex practices (insertive or receptive vaginal sex, anal sex, and oral sex), condomless sex for these sex practices, transactional sex, and frequency in sex venues (yes or no). Participants were also questioned about the changes in sexual behavior due to the mpox outbreak (“Have you changed your sexual behavior since mpox cases started in Brazil [June 2022]?”). If yes, they were asked about risk reduction measures adopted from a predetermined list, including the reduced number of sex partners and avoided sex parties.

Substance Use and Chemsex

Participants answered questions on substance use since June 2022. Binge drinking was evaluated with the question “Since June 2022, did you drink 5 or more drinks in a couple of hours?” (yes or no) [21]. Participants answered questions about illicit drug use from a predetermined list and an open field option for other substances, later dichotomized into yes or no. Individuals who reported binge drinking or any illicit drug use were inquired about use before or during sex (“binge drinking before or during sex” and “chemsex”) [21,22].

HIV Testing, Treatment, and Prevention, and Other Sexually Transmitted Infections

HIV self-reported status was derived from the question “Have you ever had an HIV test?” with potential responses positive, negative, or never (“unknown”). Participants were asked when the last HIV test was (≤3, >3-6, and >6 months). People living with HIV answered questions about the time since the diagnosis (≤6, >6-12, and >12 months), about antiretroviral therapy (ART) use, ART adherence using a validated scale [23], and if they currently had an undetectable HIV viral load (yes or no). Participants self-reporting negative or unknown status were asked about HIV pre-exposure prophylaxis (PrEP) use (never, current, or past) and adherence (rated from 0 [missed all doses in past 30 days] to 100 [no missing dose]). We dichotomized complete PrEP adherence (rating=100) into yes or no.

Participants were asked about syphilis, gonorrhea, and chlamydia diagnoses since June 2022 (yes or no), hepatitis B vaccination (no, 1-2 doses, 3 doses, and did not remember), and ever tested positive for hepatitis C (yes or no).

Adherence to Vaccination Campaigns, and Willingness to Use Vaccines and Treatment for Mpox

Participants responded to questions regarding participation in vaccination campaigns (always, sometimes, or never) and whether they had been vaccinated for COVID-19 and mpox. We used a 4-point Likert scale to access the willingness to use an mpox preventive vaccine, which was defined as responding highly likely to the question: “Would you vaccinate against mpox?” following other studies evaluating the willingness to use biomedical technologies [24,25]. We also accessed willingness to participate in research studies to evaluate: (1) mpox preventive vaccine; (2) mpox preventive drug; and (3) mpox therapeutic drug.

Internalized Lesbian, Gay, Bisexual, Transgender, Queer, Intersex, Asexual, Plus Phobia, Stigma, and Discrimination

The Reactions to Homosexuality Scale (RHS), originally developed to measure internalized homonegativity among MSM, includes 7 items measured on a 7-point Likert scale from strongly disagree to strongly agree, with total scores ranging from 0 to 42 [18,26]. For this analysis, we adapted the scale to allow completion by all SGM, as described in Table S1 in Multimedia Appendix 1. Higher RHS scores indicate higher internalized lesbian, gay, bisexual, transgender, queer, intersex, asexual, plus (LGBTQIA+) phobia. We also asked participants to rate mpox stigma toward LGBTQIA+ people on a 7-point Likert scale from strongly disagree to strongly agree: “LGBTQIA+ individuals are being discriminated and stigmatized due to mpox.”

Main Outcome

Our main outcome was self-reported mpox diagnosis, defined as a positive answer to the question: "Have you been diagnosed with mpox by a health professional?"

Spatial Analysis

We performed an exploratory analysis based on the kernel estimator, with adaptive radius and Gaussian function, to obtain an overview of the spatial distribution of the points corresponding to individuals with confirmed and nonconfirmed mpox diagnosis and identifying potential outbreaks of occurrence. The zip codes were considered for georeferencing.

Statistical Analysis

Continuous variables were described by their median and IQR and mean and SD. We compared sociodemographic, behavior, and clinical characteristics according to self-reported mpox diagnosis (yes vs no) using a chi-square test or Fisher exact test for qualitative variables and a Kruskal-Wallis test for quantitative variables. All analyses were performed using R software (version 4.2.2; R Foundation for Statistical Computing).

Ethics Approval

This study received approval from the institutional review board at the Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz (#CAAE 61290422.0.0000.5262). All study participants provided electronic informed consent before survey initiation.

Results

A total of 9798 individuals accessed the survey and 1219 (12.4%) did not meet the inclusion criteria (Figure 1). Of the 8579 individuals who initiated the survey, 6236 (72.7%) reached the final survey page and were included in this study. Recruitment mostly occurred on Instagram (n=2331, 37.4%), followed by Scruff (n=1561, 25.0%), Grindr (n=1246, 20.0%), and Facebook (n=692, 11.1%) (Table 1). Median age was 36 (IQR 31-44) years; 91.2% (n=5685) were cisgender men, 83.3% (n=5197) gay, 62.2% (n=3877) White, 78.7% (n=4902) had completed tertiary education, 41.2% (n=2571) reported middle income, 68.3% (n=4259) were from the Southeast region of Brazil, and 67.6% (n=4217) were living in a state capital city.

Figure 1. Study flowchart.

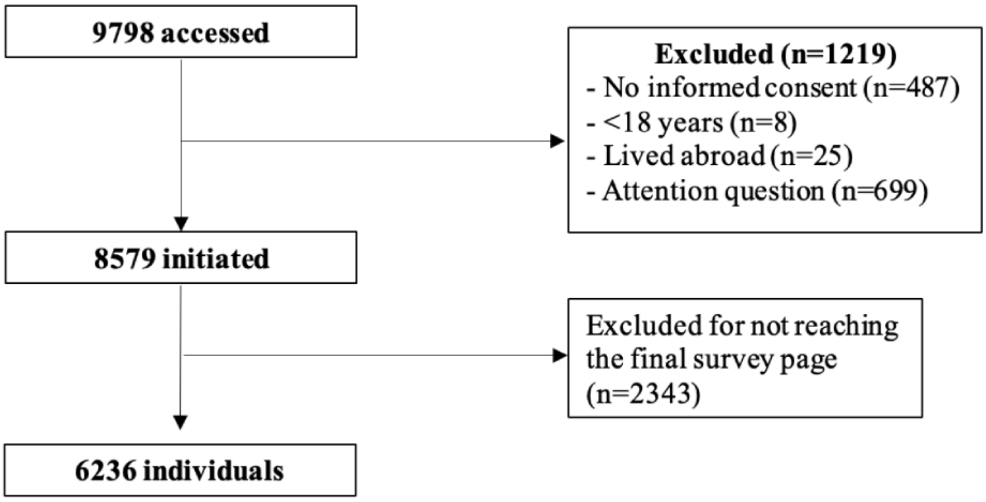


Table 1. Sociodemographic characteristics of the study population according to self-reported mpox diagnosis.

Characteristics	No (n=5912)	Yes (n=324)	Total (N=6236)	*P* value
Recruitment, n (%)				.009
Instagram	2180 (36.9)	151 (46.6)	2331 (37.4)	
Scruff	1499 (25.4)	62 (19.1)	1561 (25.0)	
Grindr	1187 (20.1)	59 (18.2)	1246 (20.0)	
Facebook	658 (11.1)	34 (10.5)	692 (11.1)	
Other	387 (6.5)	18 (5.6)	405 (6.5)	
Age (years)				
Median (IQR)	37 (31-44)	34 (30-39.2)	36 (31-44)	<.001
18-24, n (%)	363 (6.1)	10 (3.1)	373 (6)	<.001
25-30, n (%)	1071 (18.1)	77 (23.8)	1148 (18.4)	
31-40, n (%)	2369 (40.1)	171 (52.8)	2540 (40.7)	
>40, n (%)	2109 (35.7)	66 (20.4)	2175 (34.9)	
Gender, n (%)				.02
Cisgender man	5371 (90.8)	314 (96.9)	5685 (91.2)	
Cisgender woman	245 (4.1)	2 (0.6)	247 (4.0)	
Queer	156 (2.6)	6 (1.9)	162 (2.6)	
Nonbinary or gender diverse	89 (1.5)	1 (0.3)	90 (1.4)	
Transgender man	15 (0.3)	0 (0)	15 (0.2)	
Transgender woman	13 (0.2)	0 (0)	13 (0.2)	
Travesti[a]	2 (0)	0 (0)	2 (0)	
Not declared	21 (0.4)	1 (0.3)	22 (0.4)	
Sexual orientation, n (%)				<.001
Gay	4905 (83.0)	292 (90.1)	5197 (83.3)	
Bisexual	509 (8.6)	19 (5.9)	528 (8.5)	
Pansexual	305 (5.2)	2 (0.6)	307 (4.9)	
Heterosexual	139 (2.4)	11 (3.4)	150 (2.4)	
Lesbian	19 (0.3)	0 (0)	19 (0.3)	
Other	35 (0.6)	0 (0)	35 (0.6)	
Race, n (%)				.008
Asian	87 (1.5)	1 (0.3)	88 (1.4)	
Black	611 (10.3)	52 (16.0)	663 (10.6)	
Indigenous	30 (0.5)	3 (0.9)	33 (0.5)	
Pardo	1443 (24.4)	79 (24.4)	1522 (24.4)	
White	3689 (62.4)	188 (58.0)	3877 (62.2)	
Not declared	52 (0.9)	1 (0.3)	53 (0.8)	
Education, n (%)				.76
Primary	108 (1.8)	5 (1.5)	113 (1.8)	
Secondary	1157 (19.6)	59 (18.2)	1216 (19.5)	
Tertiary	4642 (78.6)	260 (80.2)	4902 (78.7)	
Income, n (%)				.08
Low	1425 (24.1)	74 (22.8)	1499 (24.0)	
Middle	2420 (40.9)	151 (46.6)	2571 (41.2)	

Characteristics	No (n=5912)	Yes (n=324)	Total (N=6236)	*P* value
High	1886 (31.9)	95 (29.3)	1981 (31.8)	
Did not want to answer	181 (3.1)	4 (1.2)	185 (3.0)	
Social support, n (%)				.47
No	5510 (93.2)	302 (93.2)	5812 (93.2)	
Yes	348 (5.9)	17 (5.2)	365 (5.9)	
Did not want to answer	54 (0.9)	5 (1.5)	59 (0.9)	
Persons living in the same house (n), median (IQR)	2 (1-3)	2 (1-3)	2 (1-3)	.13
Brazilian region, n (%)				.02
North	134 (2.3)	3 (0.9)	137 (2.2)	
Northeast	709 (12.0)	29 (9.0)	738 (11.8)	
Central-west	411 (7.0)	36 (11.1)	447 (7.2)	
Southeast	4038 (68.3)	221 (68.2)	4259 (68.3)	
South	620 (10.5)	35 (10.8)	655 (10.5)	
State region, n (%)				.003
Capital	3978 (67.3)	239 (73.8)	4217 (67.6)	
Metropolitan area	929 (15.7)	53 (16.4)	982 (15.7)	
Other cities	1005 (17.0)	32 (9.9)	1037 (16.6)	

[a] Female gender construction, which is identified in social, family, cultural, and interpersonal life through this identity.

Most participants had heard of mpox (n=6044, 96.9%), reported knowing how mpox lesions appear (n=5298, 85.0%) (Table 2), and that they would follow social distancing if diagnosed with mpox (n=6051, 97.0%). Information about mpox was mostly obtained on the internet (n=4793, 76.9%) and television (n=4519, 72.5%) (Table S2 in Multimedia Appendix 1).

Overall, 324 (5.6%) individuals reported an mpox diagnosis; 44.4% (n=144) of mpox cases were from São Paulo state, followed by 15.7% (n=51) in Rio de Janeiro (Table S3 in Multimedia Appendix 1). Hotspots of participants reporting an mpox diagnosis were concentrated in the metropolitan areas of São Paulo and Rio de Janeiro (Figure 2). A total of 304 (93.8%) respondents reported social distancing after mpox diagnosis. Compared to those with no diagnosis, participants reporting an mpox diagnosis were younger (age median 34, IQR 30-39.2 vs median 37, IQR 31-44 years; *P*<.001), more frequently cisgender men (314/324, 96.2% vs 5371/5912, 90.8%; *P*=.02), gay (292/324, 90.1% vs 4905/5912, 83%; *P*<.001), Black (52/324, 16% vs 611/5912, 10.3%; *P*=.008), and living in a state capital city (239/324, 73.8% vs 3978/5912, 67.3%; *P*=.003) (Table 1).

Among participants reporting an mpox diagnosis (n=324), 318 (98.1%) reported lesions, 178 (56%) had local pain, and 316 (99.4%) sought health facilities for diagnosis. The most common lesion sites were genital region (n=155, 48.7%), face (n=128, 40.3%), and anal region (n=118, 37.1%) (Table S4 in Multimedia Appendix 1). Sites with the highest mean numbers of lesions were anal region (4.4, SD 4.3), back (3.7, SD 4.4), genital region (3.2, SD 4.3), and arms (3.2, SD 3.1) (Table S5 in Multimedia Appendix 1). Mpox symptoms were reported by 95.4% (n=309/324) of the participants who reported mpox diagnosis. Among them, 96.8% (n=299/309) attended medical facilities to investigate these symptoms (Table 2), of which the most frequent were headache (n=211, 65.1%) and asthenia (n=205, 63.3%) (Table S6 in Multimedia Appendix 1).

Among participants with no mpox diagnosis, 4974 (84.1%) were aware of mpox lesions, 288 (4.9%) had a suspicious mpox lesion, and 4009 (67.8%) reported potential mpox symptoms. Among individuals with no diagnosis, only 54.9% (n=158/288) and 34.9% (n=1399/4009) attended a health facility to investigate a lesion or any symptoms, respectively (Table 2).

Participants with an mpox diagnosis compared to those with no diagnosis more frequently reported sex partners with suspicious or confirmed mpox (112/324, 34.6% vs 206/5912, 3.5%; *P*<.001), sex with a cisgender man (320/324, 98.8% vs 5132/5912, 86.8%; *P*<.001), more sexual partners (median 10, IQR 5-20 vs median 4, IQR 1-10; *P*<.001), and insertive or receptive anal or oral sex, including condomless sex (*P*<.001 for all) (Table 3). Binge drinking before or during sex (115/324, 35.5% vs 1565/5912, 26.5%; *P*<.001), chemsex (118/324, 36.4% vs 1429/5912, 24.2%; *P*<.001), and any illicit drug use (160/324, 49.4% vs 2354/5912, 39.8%; *P*<.001) were more frequent among those self-reporting an mpox diagnosis. Changes in sexual behavior after the outbreak onset were reported frequently regardless of mpox diagnosis but were higher among participants with a diagnosis (186/324, 57.4% vs 2878/5912, 48.7%; *P*=.002) (Table 4).

Table 2. Mpox awareness, knowledge, symptoms, and possible contact according to self-reported mpox diagnosis.

Characteristics	No (n=5912)	Yes (n=324)	Total (N=6236)	*P* value
Mpox awareness, n (%)	5724 (96.8)	320 (98.8)	6044 (96.9)	.048
Knowledge of mpox lesions, n (%)	4974 (84.1)	324 (100)	5298 (85.0)	<.001
Mpox lesions				
Mpox suspicious lesion, n (%)	288 (4.9)	318 (98.1)	606 (9.7)	<.001
Local pain (n=606), n (%)	78 (27.1)	178 (56.0)	256 (42.2)	<.001
Pain level (n=256), median (IQR)	40 (19-70.8)	75 (49-89.8)	67 (30-84.2)	<.001
Attended health facility to investigate lesions (n=606), n (%)	158 (54.9)	316 (99.4)	474 (78.2)	<.001
Health facility (n=473), n (%)				.14
Public health care	85 (53.8)	192 (61.0)	280 (58.6)	
Private health services	73 (46.2)	123 (39.0)	196 (41.4)	
Any mpox symptoms, n (%)	4009 (67.8)	309 (95.4)	4318 (69.2)	<.001
Attended medical service to investigate symptoms (n=4318), n (%)	1399 (34.9)	299 (96.8)	1698 (39.3)	<.001
Contact with person with suspicious or confirmed mpox, n (%)				
Any	421 (7.1)	112 (34.6)	533 (8.5)	<.001
Sex partner	206 (3.5)	112 (34.6)	318 (5.1)	<.001
Traveling history since June 2022 , n (%)				
Within Brazil	2459 (41.6)	150 (46.3)	2609 (41.8)	.09
Abroad	702 (11.9)	37 (11.4)	739 (11.9)	.80
Mpox vaccination, n (%)				.18
No	5866 (99.2)	323 (99.7)	6189 (99.2)	
1 dose	40 (0.7)	0 (0)	40 (0.6)	
2 doses	6 (0.1)	1 (0.3)	7 (0.1)	
Country of mpox vaccination, n (%)				.23
Canada	11 (24.4)	1 (100)	12 (26.1)	
United States	25 (55.6)	0 (0)	25 (54.3)	
France, Italy, or Spain	9 (20.0)	0 (0)	9 (19.6)	

Figure 2. Spatial distribution of individuals with negative and positive self-reported mpox diagnosis.

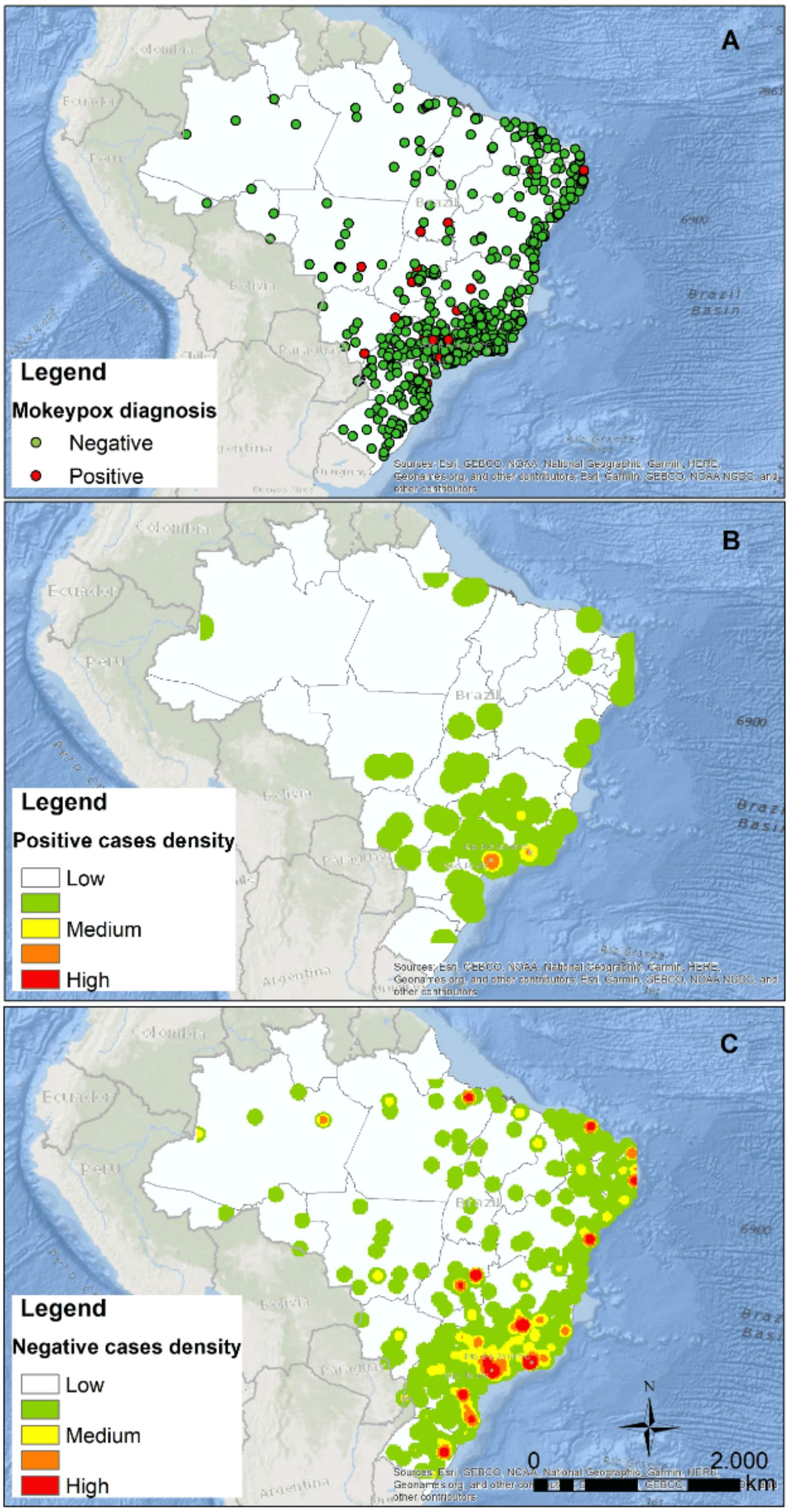


Table 3. Sexual behavior and substance use according to self-reported mpox diagnosis.

Characteristics	No (n=5912)	Yes (n=324)	Total (N=6236)	*P* value
Sex partners, n (%)				
Any gender	5330 (90.2)	323 (99.7)	5653 (90.7)	<.001
Cisgender man	5132 (86.8)	320 (98.8)	5452 (87.4)	<.001
Transgender man	108 (1.8)	5 (1.5)	113 (1.8)	.71
Cisgender woman	271 (4.6)	11 (3.4)	282 (4.5)	.32
Transgender man	43 (0.7)	2 (0.6)	45 (0.7)	.82
Travesti[a]	41 (0.7)	3 (0.9)	44 (0.7)	.63
Nonbinary	200 (3.4)	11 (3.4)	211 (3.4)	.99
Other gender	51 (0.9)	1 (0.3)	52 (0.8)	.29
Number of sex partners (n=5121)				
Median (IQR)	4 (1-10)	10 (5-20)	4 (1-10)	<.001
0, n (%)	475 (9.7)	1 (0.4)	476 (9.3)	<.001
1, n (%)	976 (20)	14 (5.6)	990 (19.3)	
2-5, n (%)	1716 (35.2)	54 (21.7)	1770 (34.6)	
6-10, n (%)	915 (18.8)	69 (27.7)	984 (19.2)	
>10, n (%)	790 (16.2)	111 (44.6)	901 (17.6)	
Steady sex partners, n (%)	2452 (50.3)	121 (48.6)	2573 (50.2)	.59
Only 1 steady sex partner, n (%)	1716 (70.4)	60 (51.3)	1776 (69.5)	<.001
Sex practices (n=5193), n (%)				
Insertive vaginal sex	278 (5.6)	10 (4.0)	288 (5.5)	.28
Receptive vaginal sex	157 (3.2)	2 (0.8)	159 (3.1)	.03
Insertive anal sex	2996 (60.6)	214 (85.9)	3210 (61.8)	<.001
Receptive anal sex	2874 (58.1)	187 (75.1)	3061 (58.9)	<.001
Insertive oral sex	3760 (76.1)	218 (87.6)	3978 (76.6)	<.001
Receptive oral sex	3662 (74.1)	220 (88.4)	3882 (74.8)	<.001
Condomless sex (n=5193), n (%)				
Any	3535 (72.9)	215 (87.4)	3750 (73.6)	<.001
Condomless insertive vaginal sex	172 (3.5)	3 (1.2)	175 (3.4)	.05
Condomless receptive vaginal sex	112 (2.3)	2 (0.8)	114 (2.2)	.12
Condomless insertive anal sex	2212 (44.7)	180 (72.3)	2392 (46.1)	<.001
Condomless receptive anal sex	2126 (43.0)	162 (65.1)	2288 (44.1)	<.001
Condomless insertive oral sex	1957 (39.6)	56 (22.5)	2013 (38.8)	<.001
Condomless receptive oral sex	2884 (58.3)	195 (78.3)	3079 (59.3)	<.001
Transactional sex (n=5064), n (%)	175 (3.6)	19 (7.9)	194 (3.8)	<.001
Frequented sex venues, n (%)	2173 (36.8)	206 (63.6)	2379 (38.1)	<.001
Substance use, n (%)				
Binge drinking	3465 (58.6)	200 (61.7)	3665 (58.8)	.27
Binge drinking before or during sex	1565 (26.5)	115 (35.5)	1680 (26.9)	<.001
Any illicit substance use	2354 (39.8)	160 (49.4)	2514 (40.3)	<.001
Cocaine	640 (10.8)	76 (23.5)	716 (11.5)	<.001
Marijuana	1770 (29.9)	116 (35.8)	1886 (30.2)	.02
Ecstasy	631 (10.7)	51 (15.7)	682 (10.9)	.004

Characteristics	No (n=5912)	Yes (n=324)	Total (N=6236)	*P* value
Ketamine	324 (5.5)	44 (13.6)	368 (5.9)	<.001
GHB[b]	233 (3.9)	31 (9.6)	264 (4.2)	<.001
Poppers	566 (9.6)	58 (17.9)	624 (10.0)	<.001
Chemsex	1429 (24.2)	118 (36.4)	1547 (24.8)	<.001

[a] Female gender construction, which is identified in social, family, cultural, and interpersonal life through this identity.

[b] GHB: gamma-hydroxybutyric acid.

Table 4. Changes in sexual behavior according to self-reported mpox diagnosis.

Characteristics	No (n=5912), n (%)	Yes (n=324), n (%)	Total (N=6236), n (%)	*P* value
Any change	2878 (48.7)	186 (57.4)	3064 (49.1)	.002
Reduced number of sex partners	2226 (37.7)	149 (46.0)	2375 (38.1)	.003
Avoided sex parties	1596 (27.0)	122 (37.7)	1718 (27.5)	<.001
Avoided bars and clubs	1050 (17.8)	70 (21.6)	1120 (18.0)	.08
Inspected partners for mpox lesions	860 (14.5)	66 (20.4)	926 (14.8)	.004
Asked partners about mpox symptoms	749 (12.7)	50 (15.4)	799 (12.8)	.15
Anal sex only with condoms	742 (12.6)	50 (15.4)	792 (12.7)	.13
Sex abstinence or avoided any type of sex	644 (10.9)	36 (11.1)	680 (10.9)	.90
Avoided kissing	409 (6.9)	26 (8.0)	435 (7.0)	.45
Web-based sex	382 (6.5)	12 (3.7)	394 (6.3)	.047
Washed bed linen and other (eg, towels) after sex	284 (4.8)	26 (8.0)	310 (5.0)	.009
Self-masturbation with partner in the space without touching him or her or they	90 (1.5)	2 (0.6)	92 (1.5)	.19
Wear face masks during sex	79 (1.3)	5 (1.5)	84 (1.3)	.75
Dated people only using clothing	80 (1.4)	2 (0.6)	82 (1.3)	.26
Oral sex with condoms only	63 (1.1)	2 (0.6)	65 (1.0)	.44
Used gloves for anal fisting or fingering	20 (0.3)	1 (0.3)	21 (0.3)	.93

HIV prevalence (122/324, 37.7% vs 1353/5912, 22.9%; *P*<.001), hepatitis C (ever tested positive; 23/324, 7.1% vs 238/5912, 4%; *P*=.007), and any diagnoses of sexually transmitted infections (STIs) (81/324, 25% vs 628/5912, 10.6%; *P*<.001) differed by mpox diagnosis report (Table 5). Among HIV-negative or unknown individuals, HIV testing in the prior 3 months (234/316, 74.1% vs 2335/5545, 42.1%; *P*<.001) and current PrEP use (96/201, 47.8% vs 1106/4537, 24.4%; *P*<.001) were higher among those reporting an mpox diagnosis (Table 5).

Most participants were willing to receive an mpox vaccine (n=5008, 95.1%). When asked about mpox research studies, 74.1% (n=4621) responded that they would participate in prevention vaccine trials, 69.2% (n=4288) in prevention drug trials, and 68.7% (n=4263) in therapeutic trials. In addition, 92% (n=5735) reported frequent participation in vaccination campaigns, with 95.8% (n=4734/4946) coverage of at least 3 doses of the COVID-19 vaccine. Nevertheless, only 47 (0.7%) participants received an mpox vaccine, which is not available in Brazil (Table 2).

Overall, 5258 (84.4%) participants agreed or strongly agreed with the statement: “LGBTQIA+ individuals are being discriminated and stigmatized due to mpox.” Median score of internalized LGBTQIA+ phobia measured by the adapted RHS scale was lower among participants reporting mpox compared to those with no diagnosis (6, IQR 2-11 vs 7, IQR 2-12; *P*<.001).

Table 5. HIV testing, treatment, and prevention, and other sexually transmitted infections according to self-reported mpox diagnosis.

Characteristics	No (n=5912), n (%)	Yes (n=324), n (%)	Total (N=6236), n (%)	*P* value
HIV status				<.001
Negative	4191 (70.9)	194 (59.9)	4385 (70.3)	
Positive	1353 (22.9)	122 (37.7)	1475 (23.7)	
Unknown	367 (6.2)	8 (2.5)	375 (6.0)	
Last HIV test (n=5861)				<.001
≤3 months	2335 (42.1)	234 (74.1)	2569 (43.8)	
>3-6 months	1039 (18.7)	36 (11.4)	1075 (18.3)	
>6 months	2171 (39.2)	46 (14.6)	2217 (37.8)	
Time living with HIV (n=1475)				.006
≤6 months	14 (1.0)	5 (4.1)	19 (1.3)	
>6-12 months	42 (3.1)	1 (0.8)	43 (2.9)	
>12 months	1297 (95.9)	116 (95.1)	1413 (95.8)	
HIV currently undetectable (n=1475)	1250 (92.4)	110 (90.2)	1360 (92.2)	.38
Adherence to ART[a] (n=1473)	821 (60.7)	74 (61.2)	895 (60.8)	.93
PrEP[b] use (n=4738)				<.001
Never	3132 (69.0)	85 (42.3)	3217 (67.9)	
Current	1106 (24.4)	96 (47.8)	1202 (25.4)	
Past	299 (6.6)	20 (10.0)	319 (6.7)	
Complete PrEP adherence (n=1202)	787 (70.6)	73 (75.3)	860 (71.0)	.33
STI[c] diagnoses				
Any	628 (10.6)	81 (25.0)	709 (11.4)	<.001
Syphilis	442 (7.5)	61 (18.8)	503 (8.1)	<.001
Gonorrhea	166 (2.8)	22 (6.8)	188 (3.0)	<.001
Chlamydia	94 (1.6)	16 (4.9)	110 (1.8)	<.001
Hepatitis B vaccination				.32
No vaccination	469 (7.9)	33 (10.2)	502 (8.1)	
1-2 doses	1395 (23.6)	82 (25.3)	1477 (23.7)	
3 doses	1739 (29.4)	95 (29.3)	1834 (29.4)	
Did not remember	2309 (39.1)	114 (35.2)	2423 (38.9)	
Hepatitis C (ever tested positive)	238 (4.0)	23 (7.1)	261 (4.2)	.007

[a] ART: antiretroviral therapy.

[b] PrEP: pre-exposure prophylaxis.

[c] STI: sexually transmitted infection.

Discussion

Principal Findings

This analysis describes the results of a large web-based survey in Brazil conducted during the mpox surge in 2022. Respondents were mostly MSM from metropolitan regions, with a 5.6% prevalence of self-reported mpox diagnosis, which is slightly higher than Brazilian surveillance numbers. This indicates that our survey may have reached populations most affected by mpox by the current outbreak [2,5,6]. There were no reported cases among transgender women, which highlights the low mpox burden in this group, in agreement with Brazilian official data available, disaggregated by gender [8,27]. Our findings point to a high awareness of mpox among SGM, with the internet and television as the most frequent information sources. This underscores the importance of adequate science communication through traditional and innovative means of communication and the need to stop fake news that contributes to misleading information as well as reinforces stigma [28]. Most participants agreed or strongly agreed that LGBTQIA+ individuals are being discriminated against and stigmatized due to mpox. To our knowledge, this is the first study to describe mpox awareness, willingness to vaccinate for mpox, and self-reported prevalence of mpox in Latin America.

Willingness to Use Mpox Vaccine

Willingness to use mpox vaccine and to follow social isolation if required were high, suggesting an adequate understanding of the importance of public health measures to successfully mitigate the impact of the mpox outbreak among SGM. While countries in North America and Western Europe initiated mpox vaccination of the most affected population (MSM) by mid-2022, mpox vaccination in Brazil only started in March 2023. Due to the low number of vaccine units available, vaccination in the country is restricted to very few groups such as people living with HIV, health professionals working with *Orthopoxvirus* (pre-exposure vaccination), and those who had contact with fluids and secretions of persons with suspected mpox (postexposure vaccination) [29].

Participants' Characteristics According to Self-Reported Mpox Diagnosis

In our study, individuals with self-reported mpox diagnosis more frequently referred to mucocutaneous, genital or anal lesions, coinfection with hepatitis C, and local pain. These clinical characteristics are in agreement with global and national data [2,6,30]. Local pain was the most frequent reason for hospitalization in a cohort in Rio de Janeiro State, Brazil [6]. This is critical in the context of poor outcomes related to pain control according to gender, sexual orientation, and race, which might deepen mpox-related stigma and increase the gap between the most vulnerable populations and health services [31,32].

Among respondents with no mpox diagnosis, a substantial proportion did not seek health care assistance despite reporting mpox suspicious lesions or symptoms. Avoidance of health services is common among SGM and directly linked to stigma, discrimination, and structural barriers faced by these groups [33,34]. This might lead to the underreporting of mpox cases in highly discriminatory settings, such as Brazil. In addition, participants with a self-reported mpox diagnosis also reported following a higher frequency of health measures (eg, HIV care and testing, STI diagnosis, and PrEP use) and lower rates of internalized LGBTQIA+ phobia. This might indicate that individuals who sought an mpox diagnosis most commonly have access to health services in general. Furthermore, these participants might have previous links with gender-competent health care services and more access to health information, all potentially related to a lower perceived gender and sexual orientation–based discrimination.

Although participants with an mpox diagnosis reported a higher number of sexual partners, binge drinking before sex, chemsex, substance use, and higher frequency to sex venues, they also adopted more changes in sexual behavior after the mpox outbreak, possibly due to higher awareness and feelings of fear and hopelessness related to mpox. This finding is in agreement with previous surveys conducted in the United States among the most vulnerable populations [16,35]. Moreover, most individuals with an mpox diagnosis in our survey also reported sex with other cisgender men, reinforcing the role of highly interconnected and dense sexual networks in the mpox transmission dynamics [36]. Even among participants who self-reported an mpox diagnosis, only a small proportion reported a partner with suspected or confirmed mpox. Although this observation might be related to underdiagnosed cases, it raises concerns about the potential role of subclinical infections in mpox transmission dynamics, while it also reflects the well-known challenges in STI network mapping [37-39].

People living with HIV enrolled in the current analysis reported low ART adherence (895/1473, 60.8%). Regardless of mpox diagnosis, this reinforces the urgent need to better monitor the HIV care continuum. In the context of mpox, advanced immunosuppression and poor ART adherence have been associated with worse mpox clinical outcomes, including mpox-related hospitalization and death [40]. The high rates of other STIs occurring concomitantly with mpox highlight the importance of a comprehensive screening at mpox assessment, in accordance with the well-established rationale that 1 STI diagnosis implies a higher likelihood of coinfection with an additional one [41].

Limitations

Our study has some limitations. First, the cross-sectional study design hinders identifying causal associations. In addition, all responses were self-reported, thus introducing the possibility of recall, response, or social desirability bias. Moreover, our sample constitutes a highly educated, middle or high socioeconomic status subset of Brazilian persons with access to a device compatible with GSN apps, although cellphones and internet connection have been shown to be widely available in all socioeconomic strata in Brazil [42]. Transgender women are often a highly vulnerable group in Brazil with lower access to the internet, which may have impacted their recruitment using web-based strategies as previously observed [17].

Conclusions

Our results point to high mpox knowledge and willingness to vaccinate for mpox among SGM in Brazil. Participants self-reporting mpox diagnosis more frequently reported to be living with HIV, STI diagnosis, and current PrEP use. Our findings highlight the importance of an mpox assessment that includes comprehensive sexual health screenings. Efforts to decrease stigma related to mpox among SGM is necessary to avoid mpox underdiagnosis.

Acknowledgments

Parts of these results were presented at CROI Conference in Seattle in February 2023. We would like to thank study participants and Leandra Sobral, Cláudio Nascimento, Gerson Santos, Priscila Cristine, Alexandre Castilho, Daniel RB Bezerra, and Ronaldo

M Ismerio. TST, PML, and BG are supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Rio de Janeiro.

Data Availability

The study's final deidentified data set and dictionary will be made available with the publication upon reasonable request. A proposal should be submitted to the corresponding author's email, who will evaluate and approve the request. No other additional documents will be available.

Authors' Contributions

TST, MSTS, CC, BH, VGV, and BG conceived this study. TST and BH drafted the questionnaire. MSTS, CC, EMJ, JM, VGV, and BG reviewed the questionnaire. TST, CC, and PML conceived and supervised the current analysis. TST supervised data acquisition. TST performed statistical analysis and MAM performed the spatial analysis. TST and MSTS drafted the manuscript. BH, EMJ, SWC, JM, MAM, PML, VGV, and BG revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 42 KB - publichealth_v9i1e46489_app1.docx]

References

1. Zarocostas J. Monkeypox PHEIC decision hoped to spur the world to act. Lancet 2022;400(10349):347 [FREE Full text] [doi: 10.1016/S0140-6736(22)01419-2] [Medline: 35908563]
2. Gessain A, Nakoune E, Yazdanpanah Y. Monkeypox. N Engl J Med 2022;387(19):1783-1793 [FREE Full text] [doi: 10.1056/NEJMra2208860] [Medline: 36286263]
3. Vivancos R, Anderson C, Blomquist P, Balasegaram S, Bell A, Bishop L, UKHSA Monkeypox Incident Management team, Monkeypox Incident Management Team. Community transmission of monkeypox in the United Kingdom, April to May 2022. Euro Surveill 2022;27(22):2200422 [FREE Full text] [doi: 10.2807/1560-7917.ES.2022.27.22.2200422] [Medline: 35656834]
4. 2022 Monkeypox outbreak global map. CDC. 2022. URL: https://www.cdc.gov/poxvirus/monkeypox/response/2022/world-map.html [accessed 2022-11-21]
5. Mitjà O, Ogoina D, Titanji BK, Galvan C, Muyembe JJ, Marks M, et al. Monkeypox. Lancet 2023;401(10370):60-74 [FREE Full text] [doi: 10.1016/S0140-6736(22)02075-X] [Medline: 36403582]
6. Silva MST, Coutinho C, Torres TS, Peixoto E, Ismério R, Lessa F, INI-Fiocruz Mpox Study Group. Ambulatory and hospitalized patients with suspected and confirmed mpox: an observational cohort study from Brazil. Lancet Reg Health Am 2023;17:100406 [FREE Full text] [doi: 10.1016/j.lana.2022.100406] [Medline: 36776570]
7. Ministério da Saúde. Ministério da Saúde. Boletim Epidemiológico Especial Monkeypox n 16 - 09/10/2022 a 05/11/2022. 2022. URL: https://saude.rs.gov.br/upload/arquivos/202303/08122154-informativo-epidemiologico-especial-mpox-dados-2022.pdf [accessed 2023-06-22]
8. Silva MST, Jalil EM, Torres TS, Coutinho C, Jalil CM, Cardoso SW, et al. Monkeypox and transgender women: the need for a global initiative. Travel Med Infect Dis 2022;50:102479 [FREE Full text] [doi: 10.1016/j.tmaid.2022.102479] [Medline: 36257591]
9. Scheffer M, Paiva VSF, Barberia LG, Russo G. Monkeypox in Brazil between stigma, politics, and structural shortcomings: have we not been here before? Lancet Reg Health Am 2023;17:100394 [FREE Full text] [doi: 10.1016/j.lana.2022.100394] [Medline: 36405886]
10. Bragazzi NL, Khamisy-Farah R, Tsigalou C, Mahroum N, Converti M. Attaching a stigma to the LGBTQI+ community should be avoided during the monkeypox epidemic. J Med Virol 2023;95(1):e27913 [FREE Full text] [doi: 10.1002/jmv.27913] [Medline: 35655436]
11. Kenyon C. Is monkeypox being underdiagnosed in countries with more stigmatizing attitudes towards men who have sex with men? A simple ecological analysis. Epidemiologia (Basel) 2022;3(3):363-368 [FREE Full text] [doi: 10.3390/epidemiologia3030028] [Medline: 36417244]
12. März JW, Holm S, Biller-Andorno N. Monkeypox, stigma and public health. Lancet Reg Health Eur 2022;23:100536 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100536] [Medline: 36338836]
13. Kutscher E. Erasing shame in sexuality. JAMA 2022;328(14):1397-1398 [FREE Full text] [doi: 10.1001/jama.2022.17156] [Medline: 36094525]

14. Ogoina D, Yinka-Ogunleye A. Sexual history of human monkeypox patients seen at a tertiary hospital in Bayelsa, Nigeria. Int J STD AIDS 2022;33(10):928-932 [FREE Full text] [doi: 10.1177/09564624221119335] [Medline: 35969500]
15. Paparini S, Whitacre R, Smuk M, Thornhill J, Mwendera C, Strachan S, et al. Public understanding and awareness of and response to monkeypox virus outbreak: a cross-sectional survey of the most affected communities in the United Kingdom during the 2022 public health emergency. HIV Med 2023;24(5):544-557 [FREE Full text] [doi: 10.1111/hiv.13430] [Medline: 36385726]
16. Delaney KP, Sanchez T, Hannah M, Edwards OW, Carpino T, Agnew-Brune C, et al. Strategies adopted by gay, bisexual, and other men who have sex with men to prevent monkeypox virus transmission - United States, August 2022. MMWR Morb Mortal Wkly Rep 2022;71(35):1126-1130 [FREE Full text] [doi: 10.15585/mmwr.mm7135e1] [Medline: 36048582]
17. Bezerra DRB, Jalil CM, Jalil EM, Coelho LE, Carvalheira E, Freitas J, et al. Complementary recruitment strategies to reach men who have sex with men and transgender women: the experience of a large Brazilian HIV prevention service. AIDS Behav 2022;26(8):2643-2652. [doi: 10.1007/s10461-022-03609-5] [Medline: 35122578]
18. Blair KJ, Torres TS, Hoagland B, Bezerra DRB, Veloso VG, Grinsztejn B, et al. Pre-exposure prophylaxis use, HIV knowledge, and internalized homonegativity among men who have sex with men in Brazil: a cross-sectional study. Lancet Reg Health Am 2022;6:100152 [FREE Full text] [doi: 10.1016/j.lana.2021.100152] [Medline: 35291206]
19. Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods 2012;17(3):437-455. [doi: 10.1037/a0028085] [Medline: 22506584]
20. Torres TS, Hoagland B, Bezerra DRB, Garner A, Jalil EM, Coelho LE, et al. Impact of COVID-19 pandemic on sexual minority populations in Brazil: an analysis of social/racial disparities in maintaining social distancing and a description of sexual behavior. AIDS Behav 2021;25(1):73-84 [FREE Full text] [doi: 10.1007/s10461-020-02984-1] [Medline: 32737817]
21. Jalil EM, Torres TS, de A Pereira CC, Farias A, Brito JDU, Lacerda M, et al. High rates of sexualized drug use or chemsex among brazilian transgender women and young sexual and gender minorities. Int J Environ Res Public Health 2022;19(3):1704 [FREE Full text] [doi: 10.3390/ijerph19031704] [Medline: 35162728]
22. Torres TS, Bastos LS, Kamel L, Bezerra DRB, Fernandes NM, Moreira RI, et al. Do men who have sex with men who report alcohol and illicit drug use before/during sex (chemsex) present moderate/high risk for substance use disorders? Drug Alcohol Depend 2020;209:107908 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.107908] [Medline: 32078972]
23. Luz PM, Torres TS, Almeida-Brasil CC, Marins LMS, Bezerra DRB, Veloso VG, et al. Translation and validation of the short HIV stigma scale in Brazilian Portuguese. Health Qual Life Outcomes 2020;18(1):322 [FREE Full text] [doi: 10.1186/s12955-020-01571-1] [Medline: 33008400]
24. Torres TS, Konda KA, Vega-Ramirez EH, Elorreaga OA, Diaz-Sosa D, Hoagland B, ImPrEP Study Group. Factors associated with willingness to use pre-exposure prophylaxis in Brazil, Mexico, and Peru: web-based survey among men who have sex with men. JMIR Public Health Surveill 2019;5(2):e13771 [FREE Full text] [doi: 10.2196/13771] [Medline: 31210143]
25. Torres TS, De Boni RB, de Vasconcellos MT, Luz PM, Hoagland B, Moreira RI, et al. Awareness of prevention strategies and willingness to use preexposure prophylaxis in Brazilian men who have sex with men using apps for sexual encounters: online cross-sectional study. JMIR Public Health Surveill 2018;4(1):e11 [FREE Full text] [doi: 10.2196/publichealth.8997] [Medline: 29358160]
26. Torres TS, Luz PM, Bezerra DRB, Almeida-Brasil CC, Marins LMS, Veloso VG, et al. Translation and validation in Brazilian Portuguese of the reactions to homosexuality scale. Psychol Health Med 2021;26(sup1):20-36 [FREE Full text] [doi: 10.1080/13548506.2021.1936580] [Medline: 34219572]
27. Thornhill JP, Palich R, Ghosn J, Walmsley S, Moschese D, Cortes CP, Share-Net writing group. Human monkeypox virus infection in women and non-binary individuals during the 2022 outbreaks: a global case series. Lancet 2022;400(10367):1953-1965 [FREE Full text] [doi: 10.1016/S0140-6736(22)02187-0] [Medline: 36403584]
28. Melchior C, Oliveira M. Health-related fake news on social media platforms: a systematic literature review. New Media Soc 2021;24(6):1500-1522. [doi: 10.1177/14614448211038762]
29. Ministério da Saúde. Ministério da Saúde. Mpox. URL: https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/m/mpox [accessed 2023-05-29]
30. Perazzo H, Silva MST, Coutinho C, Peixoto EM, da Costa Cruz Silva S, Cardoso SW, INI-Fiocruz Monkeypox Study Group. Monkeypox outbreak as an opportunity to identify new cases of HCV infection in limited resource settings. J Viral Hepat 2023;30(1):83-85 [FREE Full text] [doi: 10.1111/jvh.13771] [Medline: 36357317]
31. Schoenthaler A, Williams N. Looking beneath the surface: racial bias in the treatment and management of pain. JAMA Netw Open 2022;5(6):e2216281 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.16281] [Medline: 35679049]
32. Abd-Elsayed A, Heyer AM, Schatman ME. Disparities in the treatment of the LGBTQ population in chronic pain management. J Pain Res 2021;14:3623-3625 [FREE Full text] [doi: 10.2147/JPR.S348525] [Medline: 34880669]
33. Costa AB, da Rosa Filho HT, Pase PF, Fontanari AMV, Catelan RF, Mueller A, et al. Healthcare needs of and access barriers for Brazilian transgender and gender diverse people. J Immigr Minor Health 2018;20(1):115-123. [doi: 10.1007/s10903-016-0527-7] [Medline: 27804013]
34. Phillips G, Neray B, Janulis P, Felt D, Mustanski B, Birkett M. Utilization and avoidance of sexual health services and providers by YMSM and transgender youth assigned male at birth in Chicago. AIDS Care 2019;31(10):1282-1289 [FREE Full text] [doi: 10.1080/09540121.2019.1587370] [Medline: 30821480]

35. Hubach RD, Owens C. Findings on the Monkeypox exposure mitigation strategies employed by men who have sex with men and transgender women in the United States. Arch Sex Behav 2022;51(8):3653-3658 [FREE Full text] [doi: 10.1007/s10508-022-02423-3] [Medline: 36103027]
36. Endo A, Murayama H, Abbott S, Ratnayake R, Pearson CAB, Edmunds WJ, et al. Heavy-tailed sexual contact networks and monkeypox epidemiology in the global outbreak, 2022. Science 2022;378(6615):90-94 [FREE Full text] [doi: 10.1126/science.add4507] [Medline: 36137054]
37. De Baetselier I, Van Dijck C, Kenyon C, Coppens J, Michiels J, de Block T, ITM Monkeypox study group. Retrospective detection of asymptomatic monkeypox virus infections among male sexual health clinic attendees in Belgium. Nat Med 2022;28(11):2288-2292 [FREE Full text] [doi: 10.1038/s41591-022-02004-w] [Medline: 35961373]
38. Van Dijck C, Hens N, Kenyon C, Tsoumanis A. The roles of unrecognized mpox cases, contact isolation and vaccination in determining epidemic size in Belgium: a modeling study. Clin Infect Dis 2023;76(3):e1421-e1423. [doi: 10.1093/cid/ciac723] [Medline: 36052546]
39. Ferré VM, Bachelard A, Zaidi M, Armand-Lefevre L, Descamps D, Charpentier C, et al. Detection of Monkeypox virus in anorectal swabs from asymptomatic men who have sex with men in a sexually transmitted infection screening program in Paris, France. Ann Intern Med 2022;175(10):1491-1492 [FREE Full text] [doi: 10.7326/M22-2183] [Medline: 35969863]
40. Miller MJ, Cash-Goldwasser S, Marx GE, Schrodt CA, Kimball A, Padgett K, CDC Severe Monkeypox Investigations Team. Severe Monkeypox in hospitalized patients - United States, August 10-October 10, 2022. MMWR Morb Mortal Wkly Rep 2022;71(44):1412-1417 [FREE Full text] [doi: 10.15585/mmwr.mm7144e1] [Medline: 36327164]
41. Curran KG, Eberly K, Russell OO, Snyder RE, Phillips EK, Tang EC, Monkeypox, HIV, and STI Team. HIV and sexually transmitted infections among persons with Monkeypox - eight U.S. jurisdictions, May 17-July 22, 2022. MMWR Morb Mortal Wkly Rep 2022;71(36):1141-1147 [FREE Full text] [doi: 10.15585/mmwr.mm7136a1] [Medline: 36074735]
42. Comitê Gestor da Internet no Brasil (CGIBR). Comitê Gestor da Internet no Brasil (CGIBR). J2 - Indivíduos Que Possuem Telefone Celular 2021. 2021. URL: https://cetic.br/pt/tics/domicilios/2021/individuos/J2/ [accessed 2022-07-12]

Abbreviations

ART: antiretroviral therapy
LGBTQIA+: Lesbian, gay, bisexual, transgender, queer, intersex, asexual, plus
MSM: men who have sex with men
MW: minimum wage
PrEP: pre-exposure prophylaxis
RHS: Reactions to Homosexuality Scale
SGM: sexual and gender minorities
STI: sexually transmitted infection

Edited by A Mavragani, T Sanchez; submitted 13.02.23; peer-reviewed by B Bock, L Magno; comments to author 25.05.23; revised version received 30.05.23; accepted 20.06.23; published 17.07.23.

Please cite as:
Torres TS, Silva MST, Coutinho C, Hoagland B, Jalil EM, Cardoso SW, Moreira J, Magalhaes MA, Luz PM, Veloso VG, Grinsztejn B
Evaluation of Mpox Knowledge, Stigma, and Willingness to Vaccinate for Mpox: Cross-Sectional Web-Based Survey Among Sexual and Gender Minorities
JMIR Public Health Surveill 2023;9:e46489
URL: https://publichealth.jmir.org/2023/1/e46489
doi:10.2196/46489
PMID:37459174

Original Paper

The Health Care Utilization and Medical Costs in Long-Term Follow-Up of Children Diagnosed With Leukemia, Solid Tumor, or Brain Tumor: Population-Based Study Using the National Health Insurance Claims Data

James S Miser[1,2,3*], MD; Ben-Chang Shia[4,5*], PhD; Yi-Wei Kao[5,6*], PhD; Yen-Lin Liu[2,7,8,9], MD, PhD; Shih-Yen Chen[7,10], MD, PhD; Wan-Ling Ho[2,7,8,11], MD, PhD

[1]Cancer Center, Taipei Medical University Hospital, Taipei, Taiwan
[2]Taipei Cancer Center, Taipei Medical University, Taipei, Taiwan
[3]Department of Pediatrics, City of Hope National Medical Center, Duarte, CA, United States
[4]Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, New Taipei City, Taiwan
[5]Artificial Intelligence Development Center, Fu Jen Catholic University, New Taipei City, Taiwan
[6]Department of Applied Statistics and Information Science, Ming Chuan University, Taoyuan City, Taiwan
[7]Department of Pediatrics, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan
[8]Department of Pediatrics, Taipei Medical University Hospital, Taipei, Taiwan
[9]Taipei Medical University Research Center of Cancer Translational Medicine, Taipei Medical University, Taipei, Taiwan
[10]Division of Pediatric Gastroenterology, Department of Pediatrics, Shuang Ho Hospital, Ministry of Health and Welfare, New Taipei, Taiwan
[11]Department of Pediatrics, Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan
*these authors contributed equally

Corresponding Author:
Wan-Ling Ho, MD, PhD
Department of Pediatrics
Taipei Medical University Hospital
252 Wuxing Street
Taipei, 11031
Taiwan
Phone: 886 2 2737 2181 ext 3715
Fax: 886 2 6639 7181
Email: 211002@h.tmu.edu.tw

Abstract

Background: Childhood cancer survivors are at a high risk of medical consequences of their disease and treatment. There is growing information about the long-term health issues of childhood cancer survivors; however, there are very few studies describing the health care utilization and costs for this unique population. Understanding their utilization of health care services and costs will provide the basis for developing strategies to better serve these individuals and potentially reduce the cost.

Objective: This study aims to determine the utilization of health services and costs for long-term survivors of childhood cancer in Taiwan.

Methods: This is a nationwide, population-based, retrospective case-control study. We analyzed the claims data of the National Health Insurance that covers 99% of the Taiwanese population of 25.68 million. A total of 33,105 children had survived for at least 5 years after the first appearance of a diagnostic code of cancer or a benign brain tumor before the age of 18 years from 2000 to 2010 with follow-up to 2015. An age- and gender-matched control group of 64,754 individuals with no cancer was randomly selected for comparison. Utilization was compared between the cancer and no cancer groups by χ^2 test. The annual medical expense was compared by the Mann-Whitney *U* test and Kruskal-Wallis rank-sum test.

Results: At a median follow-up of 7 years, childhood cancer survivors utilized a significantly higher proportion of medical center, regional hospital, inpatient, and emergency services in contrast to no cancer individuals: 57.92% (19,174/33,105) versus 44.51% (28,825/64,754), 90.66% (30,014/33,105) versus 85.70% (55,493/64,754), 27.19% (9000/33,105) versus 20.31%

(13,152/64,754), and 65.26% (21,604/33,105) versus 59.36% (38,441/64,754), respectively (all *P*<.001). The annual total expense (median, interquartile range) of childhood cancer survivors was significantly higher than that of the comparison group (US $285.56, US $161.78-US $535.80 per year vs US $203.90, US $118.98-US $347.55 per year; *P*<.001). Survivors with female gender, diagnosis before the age of 3 years, and diagnosis of brain cancer or a benign brain tumor had significantly higher annual outpatient expenses (all *P*<.001). Moreover, the analysis of outpatient medication costs showed that hormonal and neurological medications comprised the 2 largest costs in brain cancer and benign brain tumor survivors.

Conclusions: Survivors of childhood cancer and a benign brain tumor had higher utilization of advanced health resources and higher costs of care. The design of the initial treatment plan minimizing long-term consequences, early intervention strategies, and survivorship programs have the potential to mitigate costs of late effects due to childhood cancer and its treatment.

(JMIR Public Health Surveill 2023;9:e42350) doi:10.2196/42350

KEYWORDS

brain tumor; cancer survivor; children; cost of care; health care; health resource; leukemia; long-term follow-up; population-based study; solid tumor

Introduction

The survival of children diagnosed with cancer has dramatically improved over the last 4 decades [1,2]. This improvement is primarily the result of advancements in the conventional treatment of the cancers with chemotherapy, surgery, and radiotherapy. Unfortunately, these therapeutic modalities have significant short- and long-term morbidities [3-5]. More recently, advances in the understanding of the biology of the cancers in children promise to allow more targeted therapy and tailored treatment regimens that will reduce the duration and toxicity of the therapy [6-8]. Unfortunately, these targeted therapies and tailored treatments are only beginning to be incorporated into the initial therapy of children with a malignant or brain tumor. Despite the 80% survival rate currently seen among children with cancer, a recurrence of the primary cancer remains the main important concern for most children with malignancy [1,2,9,10]. Some malignancies with good risk features and low stage at diagnosis have a very low recurrence rate and can be expected to be cured. Reduction in short- and long-term morbidities is a major strategy for these malignancies. Other malignancies have a higher recurrence rate and lower survival, and thus, require further improvement in therapy.

Disease- and therapy-related complications that result in long-term disabilities have increased over recent years as the cure rate has increased. An example of this is the physical disabilities associated with limb-sparing procedures for bone tumors as improvements in surgical techniques and equipment have been made. Further, as an increasing number of children are cured of brain tumors, many have long-term health issues including seizure disorders, other neurologic effects of the primary tumor [11-13], and endocrinologic abnormalities [14,15]. These complications are primarily cognitive and neurologic effects of chemotherapy and radiation therapy [11-13,16-18]; endocrinologic complications resulting from surgery, radiation therapy, alkylating agents, certain tyrosine kinase inhibitors, and immune system modulators [14,15,19-24]; organ system toxicity [22-32]; and psychologic and emotional effects [33-37].

Estimates [38,39] suggest that there are currently 16.9 million cancer survivors in the United States. The average out-of-pocket spending each year for cancer survivors aged 18-64 years was recently reported to be approximately US $1000, compared with US $622 for people without a history of cancer [38]. Although there is now a growing body of knowledge about the long-term effects on children who survived cancer, only little information is available on the long-term costs of health care for these children [40-45]. This is especially true for the long-term costs of health care not directly related to the primary cancer. This report evaluates the utilization of health care resources and their costs for long-term survivors of childhood cancer in Taiwan.

Methods

Research Database

We analyzed the national databases of the Health and Welfare Data Science Center (HWDC) in Taiwan including the death registry of the Taiwanese population and claims data of the National Health Insurance (NHI) from the Bureau of National Health Insurance. In Taiwan, the NHI is a single-payer compulsory social insurance plan enrolling more than 99% of its residents [46,47], a population of 25.68 million covered individuals at the end of 2005. After deidentification by the HWDC, the database is open to researchers in Taiwan for research applications.

Study Design

This is a nationwide, population-based, case-control study of individuals who had a diagnostic code of cancer (ICD-9-CM [International Classification of Diseases, Ninth Revision, Clinical Modification] code 140-208) or a benign brain tumor (ICD-9-CM code 225) and who had survived for more than 5 years. The study included all survivors of childhood cancer who appeared in the national database between 2000 and 2010 and aged less than 18 years at the time of first appearance in the database. Survivors who were diagnosed between 1983 and 1999 were included in the data set as long as they were younger than 18 years when they first appeared in the database during years 2000-2010. The study population included cases whose birthday was as early as 1983 because they could have first appeared in the database beginning in the year 2000 at the age before 18. We defined the index date of follow-up as the date 5 years after the initial coding of cancer or a benign brain tumor in the claims data. To evaluate the costs of health care in long-term survivors of childhood cancer, we compared the utilization and costs of health care between children with cancer or a benign brain tumor who survived their disease more than five years and children without a history of cancer or brain tumor. Patients who died within 5 years of first diagnosis or first coding of cancer or a benign brain tumor in the database, who had no follow-up records beyond 1 year after the index date, or who had used antineoplastic and immunomodulating agents (Anatomical Therapeutic Chemical [ATC] class L) during the follow-up period were excluded from analysis.

Case Group

The case group of childhood cancer survivors (cancer group) consisted of 33,105 patients who were diagnosed with cancer at an age less than 18 years. The age at entry to the study was defined by the year of first coding of cancer or a benign brain tumor in the database minus the birth year. All the survivors in the case group had at least 6 years of tracking records (5 years after initial appearance in the database plus 1 year of follow-up). The survivors were further classified as having/had hematologic cancer (ICD-9-CM codes 200-208), brain cancer (including all central nervous system [CNS] cancers; ICD-9-CM codes 191-192), benign brain tumor (including low-grade CNS tumors; ICD-9-CM code 225), and non-CNS solid tumor (ICD-9-CM codes 140-190 and 193-199).

The index date for the initiation of tracking health care utilization and costs was 5 years after the first date of cancer coding for each case and their matched comparison individuals. The date of the latest medical claims in the database was the endpoint of tracking health care utilization and costs.

Comparison Group

An age- and gender-matched comparison group was included for comparison. To identify a matched comparison group, we randomly selected children without cancer or a benign brain tumor from the claims database; 2 comparison individuals were enrolled for each survivor of cancer or a benign brain tumor. A total of 64,754 comparison group individuals were selected by matching the propensity score of age and gender for each case [48,49]. As each survivor in the cancer groups was matched by 2 individuals with no cancer, the index date for each of the 2 individuals in the comparison group was set to be the same as their matched survivor in the cancer groups.

Outcomes of Interest

The proportion of individuals utilizing health care services, annual outpatient visits, and the annual medical expense per person were calculated for each group. The annual medical expenses were the total medical cost between the index date and the last day of follow-up in this study divided by the years of follow-up. The utilization and cost in each domain of health care (ie, outpatient, inpatient, and emergency) and in each level of health care (ie, community clinic, regional hospital, and medical center) were calculated respectively and compared across groups.

Statistical Analysis

The proportion of categorical variables such as sex, age, and health care utilization status was analyzed by the χ^2 test. Continuous variables including age, years of follow-up, and annual medical expenditure were expressed as the median and IQR; the Mann-Whitney *U* test was used for 2-group comparisons and the Kruskal-Wallis rank-sum test for multiple-group analyses. We demonstrated the annual outpatient medical expenditures (by gender and stratified age groups) graphically by radar plots and box plots. The expenses were converted from Taiwan dollars to United States dollars at the exchange rate of 30:1. All data processing in this study was performed using SAS 9.3 (SAS Institute), and statistical analysis and graphing were performed using R 3.6.1 (R Foundation for Statistical Computing). Two-sided *P* values <.05 were considered statistically significant.

Ethics Approval

This study was approved by the Joint Institutional Review Board of Taipei Medical University (TMU-JIRB N201911023), which agreed that informed consent can be waived when using these administrative data for analyses.

Results

Demographics

A cohort of 33,105 patients surviving 5 years after the first appearance of childhood cancer or a benign brain tumor in the database were identified (cancer group) and a comparison group of 64,754 individuals was selected from the deidentified national claims database of the HWDC in Taiwan from 2000 to 2015 (Table 1). The median follow-up starting from 5 years after the first appearance in the database (cancer group) or the index date of enrollment for the comparison group was 7 years. The cancer groups consisted predominantly of solid tumors (29,171/33,105, 88.12%); approximately two-thirds of the solid tumors were outside the CNS, whereas one-third originated in the CNS. The hematologic cancers comprised 11.88% (3934/33,105) of the cancer groups. The gender distribution of the cancer and no cancer groups was comparable with a slight male predominance (male:female=1.2:1) in both groups. The age distribution among the cancer and comparison groups was similar, with median ages between 11 and 14 years (Table 1).

Table 1. Demographics of the cancer and comparison groups.

Demographics	Childhood cancer survivors (n=33,105)	Cancer groups (ICD-9-CM[a])					Comparison group	
		Hematologic cancer (200-208) (n=3934)	Brain cancer (191-192) (n=2241)	Benign brain tumor (225) (n=7825)	Non-CNS[b] solid tumor (140-190; 193-199) (n=19,105)	*P* value[c]	No cancer individuals (n=64,754)	*P* value[d]
Gender, n (%)								
Male	18,377 (55.51)	2211 (56.20)	1209 (53.95)	4091 (52.28)	10,866 (56.88)	<.001	35,803 (55.29)	.52
Female	14,728 (44.49)	1723 (43.80)	1032 (46.05)	3734 (47.72)	8239 (43.12)	N/A[e]	28,951 (44.71)	N/A
Age at entry, median (IQR)	12.00 (7.00-16.00)	11.00 (5.00-15.00)	12.00 (8.00-16.00)	14.00 (8.00-17.00)	12.00 (6.00-16.00)	<.001	12.00 (6.00-16.00)	.002
Years of follow-up, median (IQR)	7.23 (3.58-10.01)	7.73 (4.73-10.53)	8.59 (5.25-10.90)	7.60 (4.73-10.00)	6.54 (3.07-9.76)	<.001	7.16 (3.53-9.98)	.04

[a] ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

[b] CNS: central nervous system.

[c] Compared among the cancer groups and the no cancer group.

[d] Compared between all cancer survivors versus no cancer individuals.

[e] N/A: not applicable.

Utilization of Medical Services

We examined the difference in medical service utilization between cancer survivors and no cancer individuals (Table 2). Cancer survivors had a higher utilization of medical center, regional hospital, inpatient, and emergency services in contrast to individuals without cancer: 57.92% (19,174/33,105) versus 44.51% (28,825/64,754), 90.66% (30,014/33,105) versus 85.70% (55,493/64,754), 27.19% (9000/33,105) versus 20.31% (13,152/64,754), and 65.26% (21,604/33,105) versus 59.36% (38,441/64,754), respectively (all *P*<.001), with more cancer survivors using medical center services (19,174/33,105, 57.92%, vs 28,825/64,754, 44.51%). We then examined the difference in service utilization among the cancer and no cancer groups. Brain cancer and benign brain tumor survivors demonstrated a different pattern of utilization, that is, more brain cancer and benign brain tumor survivors used inpatient and emergency services in contrast to hematologic and non-CNS solid tumor survivors. A higher proportion of benign and malignant brain tumor survivors also utilized secondary and tertiary care services, including hospital outpatient, inpatient, and emergency services. Of note, the utilization of emergency services was high among both the survivors and no cancer individuals (21,604/33,105, 65.26% and 38,411/64,754, 59.32%), reflecting societal patterns of seeking medical advice (Table 2).

Table 2. Utilization of medical services of the cancer and comparison groups.

Category	Study group[a]							
	Childhood cancer survivors (n=33,105), n (%)	Cancer groups (ICD-9-CM[b])					Comparison group	
		Hematologic cancer (200-208) (n=3934), n (%)	Brain cancer (191-192) (n=2241), n (%)	Benign brain tumor (225) (n=7825), n (%)	Non-CNS[c] solid tumor (140-190; 193-199) (n=19,105), n (%)	*P* value[d]	No cancer individuals (n=64,754), n (%)	*P* value[e]
Outpatient						.01		.001
Yes	33,105 (100.0)	3934 (100.0)	2241 (100.0)	7825 (100.0)	19,105 (100.0)		64,721 (99.95)	
No	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)		33 (0.05)	
Hospital outpatient						<.001		<.001
Yes	28,714 (86.74)	3476 (88.36)	2140 (95.49)	7193 (91.92)	15,905 (83.25)		51,846 (80.07)	
No	4391 (13.26)	458 (11.64)	101 (4.51)	632 (8.08)	3200 (16.75)		12,908 (19.93)	
Community clinic						<.001		.002
Yes	33,021 (99.75)	3926 (99.80)	2222 (99.15)	7804 (99.73)	19,069 (99.81)		64,509 (99.62)	
No	84 (0.25)	8 (0.20)	19 (0.85)	21 (0.27)	36 (0.19)		245 (0.38)	
Inpatient						<.001		<.001
Yes	9000 (27.19)	1060 (26.94)	848 (37.84)	2483 (31.73)	4609 (24.12)		13,152 (20.31)	
No	24,105 (72.81)	2874 (73.06)	1393 (62.16)	5342 (68.27)	14,496 (75.88)		51,602 (79.69)	
Emergency						<.001		<.001
Yes	21,604 (65.26)	2641 (67.13)	1619 (72.24)	5496 (70.24)	11,848 (62.02)		38,441 (59.36)	
No	11,501 (34.74)	1293 (32.87)	622 (27.76)	2329 (29.76)	7257 (37.98)		26,313 (40.64)	
Regional hospital						<.001		<.001
Yes	30,014 (90.66)	3616 (91.92)	2178 (97.19)	7417 (94.79)	16,803 (87.95)		55,493 (85.70)	
No	3091 (9.34)	318 (8.08)	63 (2.81)	408 (5.21)	2302 (12.05)		9261 (14.30)	
Medical center						<.001		<.001
Yes	19,174 (57.92)	2344 (59.58)	1630 (72.74)	5060 (64.66)	10,140 (53.08)		28,825 (44.51)	
No	13,931 (42.08)	1590 (40.42)	611 (27.26)	2765 (35.34)	8965 (46.92)		35,929 (55.49)	

[a] The study group included only patients with utilization record.

[b] ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

[c] CNS: central nervous system.

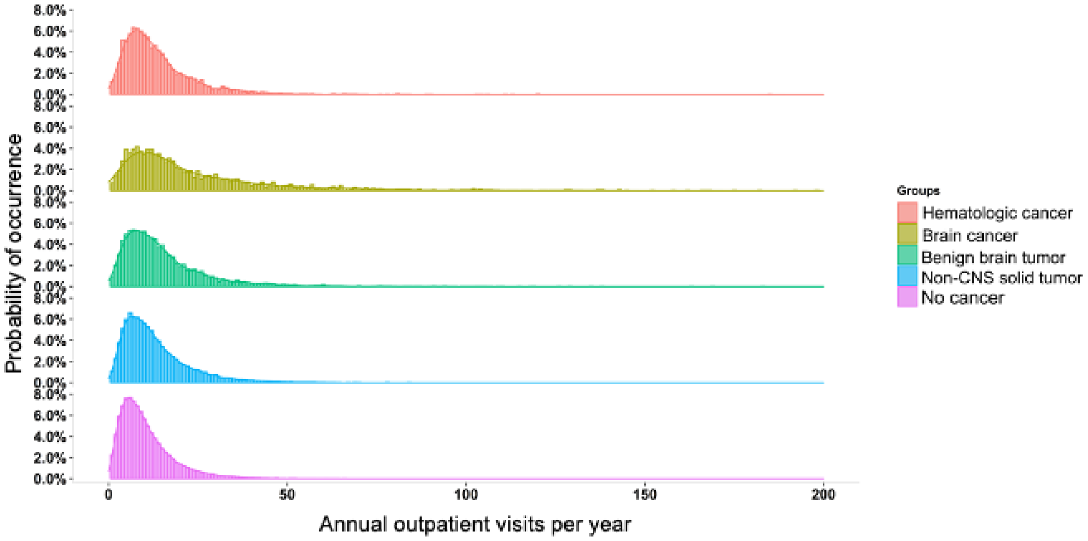
[d] Compared among the cancer and no cancer groups.

[e] Compared between all cancer survivors and no cancer individuals.

Distribution of Annual Outpatient Visits

The outpatient visit frequency showed a markedly right-skewed distribution with a long tail in all the 4 cancer groups and the comparison group, suggesting that there were a number of individuals in each group with high utilization of outpatient services (Figure 1). The median frequency (IQR) of annual outpatient visits of all childhood cancer survivors was 11.38 (6.88-18.27) visits/year, which was significantly higher than 8.63 (5.20-13.71) visits/year of the comparison group (*P*<.001). The median frequency (IQR) of outpatient visits of the 4 cancer groups was different from that of the comparison group: 11.20 (6.95-17.46) for hematologic cancer, 16.54 (9.22-30.60) for brain cancer, 12.33 (7.29-19.69) for a benign brain tumor, and 10.72 (6.52-17.03) for a non-CNS solid tumor (*P*<.001). The shape of the distribution curve was different among the cancer groups; the long tails of the distribution curves tapered off more gradually in the brain cancer and benign brain tumor groups. The results suggest that the frequencies of annual outpatient visits were the highest in the brain cancer and benign brain tumor groups, and that there were some extreme users in these 2 groups (Figure 1).

Figure 1. Distribution of annual outpatient visits in the cancer and comparison groups. The probability distribution of the annual outpatient visit frequencies in each group is summarized in a histogram. The area under the curve of each group is 100% of probability. Note the long tails in the brain cancer and benign brain tumor groups. CNS: central nervous system.



Medical Expense

The cost of medical services for childhood cancer survivors was higher than that for the comparison group (Table 3). The median total annual expense was 1.4-fold higher in childhood cancer survivors than in the comparison group (US $285.56, IQR US $161.78-US $535.80 per year vs US $203.90, IQR US $118.98-US $347.55 per year; *P*<.001); the median annual hospital outpatient cost and community clinic cost were also higher, 2.48- and 1.15-fold, respectively (*P*<.001). Moreover, childhood cancer survivors showed a different pattern of costs, with higher median annual expenses in medical centers (US $6.35 vs US $0 per year; *P*<.001) and regional hospitals (2.09-fold; *P*<.001) in contrast to no cancer individuals. Among the cancer groups, brain cancer and benign brain tumor survivors had a different pattern of costs, with higher median annual total and outpatient expenses, especially in emergency services (2.69- and 1.98-fold more over the comparison group, respectively; *P*<.001) and medical center services (US $68.37 vs US $0 and US $13.52 vs US $0 per year; *P*<.001) compared with hematologic and non-CNS solid tumor survivors (Table 3).

Table 3. Annual medical expense of the cancer and comparison groups.[a]

	Study group							
Category	Childhood cancer survivors (n=33,105)	Cancer groups (ICD-9-CM[b])					Comparison group	
		Hematologic cancer (200-208) (n=3934)	Brain cancer (191-192) (n=2241)	Benign brain tumor (225) (n=7825)	Non-CNS[c] solid tumor (140-190; 193-199) (n=19,105)	*P* value[d]	No cancer individuals (n=64,754)	*P* value[e]
Total expense	285.56 (161.78-535.80)	268.48 (159.50-475.08)	614.08 (267.05-1793.94)	336.89 (184.64-652.14)	257.78 (148.44-456.88)	<.001	203.90 (118.98-347.55)	<.001
Outpatient	239.81 (140.24-416.02)	230.96 (138.51-368.64)	473.60 (223.04-1380.48)	274.76 (159.54-503.51)	219.49 (129.69-362.93)	<.001	174.50 (103.38-280.90)	<.001
Hospital outpatient	50.94 (10.46-165.40)	45.48 (10.93-132.53)	227.17 (42.62-1076.80)	75.48 (18.91-236.87)	39.37 (6.98-124.60)	<.001	20.56 (3.19-67.82)	<.001
Community clinic	150.78 (89.09-237.73)	148.16 (92.35-232.70)	172.90 (93.45-284.81)	156.97 (91.28-248.53)	146.74 (87.16-230.49)	<.001	131.03 (76.53-206.54)	<.001
Inpatient	0.00 (0.00-36.49)	0.00 (0.00-30.92)	0.00 (0.00-142.36)	0.00 (0.00-72.10)	0.00 (0.00-0.00)	<.001	0.00 (0.00-0.00)	<.001
Emergency	8.45 (0.00-26.50)	7.87 (0.00-23.44)	14.29 (0.00-44.84)	10.54 (0.00-32.22)	7.25 (0.00-23.59)	<.001	5.31 (0.00-19.02)	<.001
Regional hospital	71.91 (19.90-235.61)	61.90 (18.78-183.95)	278.41 (64.85-946.18)	107.37 (30.41-314.84)	57.08 (15.07-178.88)	<.001	34.40 (8.50-109.28)	<.001
Medical center	6.35 (0.00-67.90)	6.93 (0.00-60.93)	68.37 (0.00-598.16)	13.52 (0.00-107.51)	2.82 (0.00-44.52)	<.001	0.00 (0.00-17.78)	<.001

[a] Data are presented as median expense in US $ per person per year (IQR).

[b] ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

[c] CNS: central nervous system.

[d] Compared among the cancer groups and the no cancer group.

[e] Compared between all cancer survivors versus no cancer individuals.

Annual Outpatient Expense

The median outpatient expense per person of each cancer group was higher than that of the comparison group. Importantly, the median outpatient expense of the brain cancer group was approximately 2 times higher than that of the other cancer groups (Figure 2A). The pattern of annual outpatient expense of the cancer and comparison groups was the same for both genders (Figures 2B and 3A and 3B). To further examine the annual outpatient expense across different ages at entry, we stratified the age into 4 intervals, namely, 0-2, 3-5, 6-11, and 12-17 years. We found that the pattern of annual outpatient expense of the cancer and comparison groups was the same across all age groups (Figures 2C and 3C-3F). Interestingly, female childhood cancer survivors and comparison individuals had a higher annual outpatient expense than their male counterparts (Table 4). In addition, infants and children up to the age of 2 years had a higher annual outpatient expense than older individuals (Table 5). The annual outpatient expense of brain cancer survivors was similar across gender and ages; however, the brain cancer survivors had a higher annual outpatient expense compared with the other cancer groups and with the no cancer group across both genders and all ages at entry (Tables 4 and 5).

Brain cancer survivors had the highest mean and median annual outpatient expenses, in contrast to other cancer groups and the comparison group ($P<.001$). The IQR of outpatient expense per person per year of the brain cancer group was greater than that of the other cancer groups and the no cancer group (Figure 2D).

Figure 2. The median annual outpatient expense in each cancer and comparison group is shown (A) and compared by gender (B) and age at entry (C). The distribution of annual outpatient expense per person was compared across the cancer and comparison groups in a box plot (D) (circles, mean annual outpatient expense; boxes, the 25th, 50th, and 75th percentile of each group; solid vertical lines, 1.5 box length, ie, 1.5 interquartile range; dots, outliers). The average outpatient medical expense per person-year in each cancer and comparison group was compared (E). Color bars represent the Anatomical Therapeutic Chemical (ATC) Classification of medications and the lowest brown bars represent nonmedication costs of medical services. CNS: central nervous system.

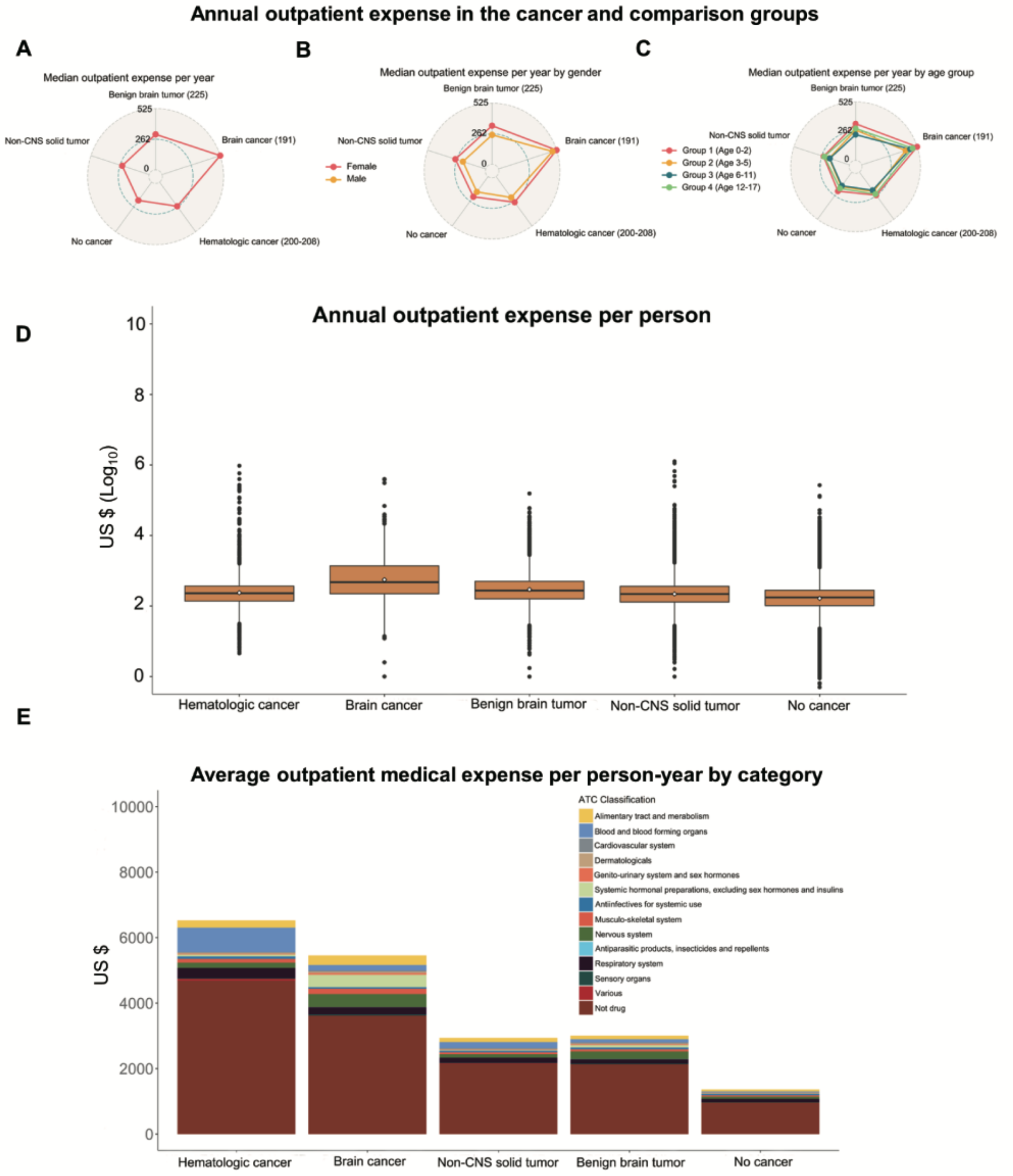


Figure 3. The annual outpatient expense in each cancer and comparison group was compared by male (A), female (B) and the age at entry of 0-2 years (C), 3-5 years (D), 6-11 years (E) and 12-17 years (F). Circles, mean annual outpatient expense; boxes, the 25th, 50th, and 75th percentile of each group; solid vertical lines, 1.5 box length, i.e., 1.5 interquartile range; dots, outliers.

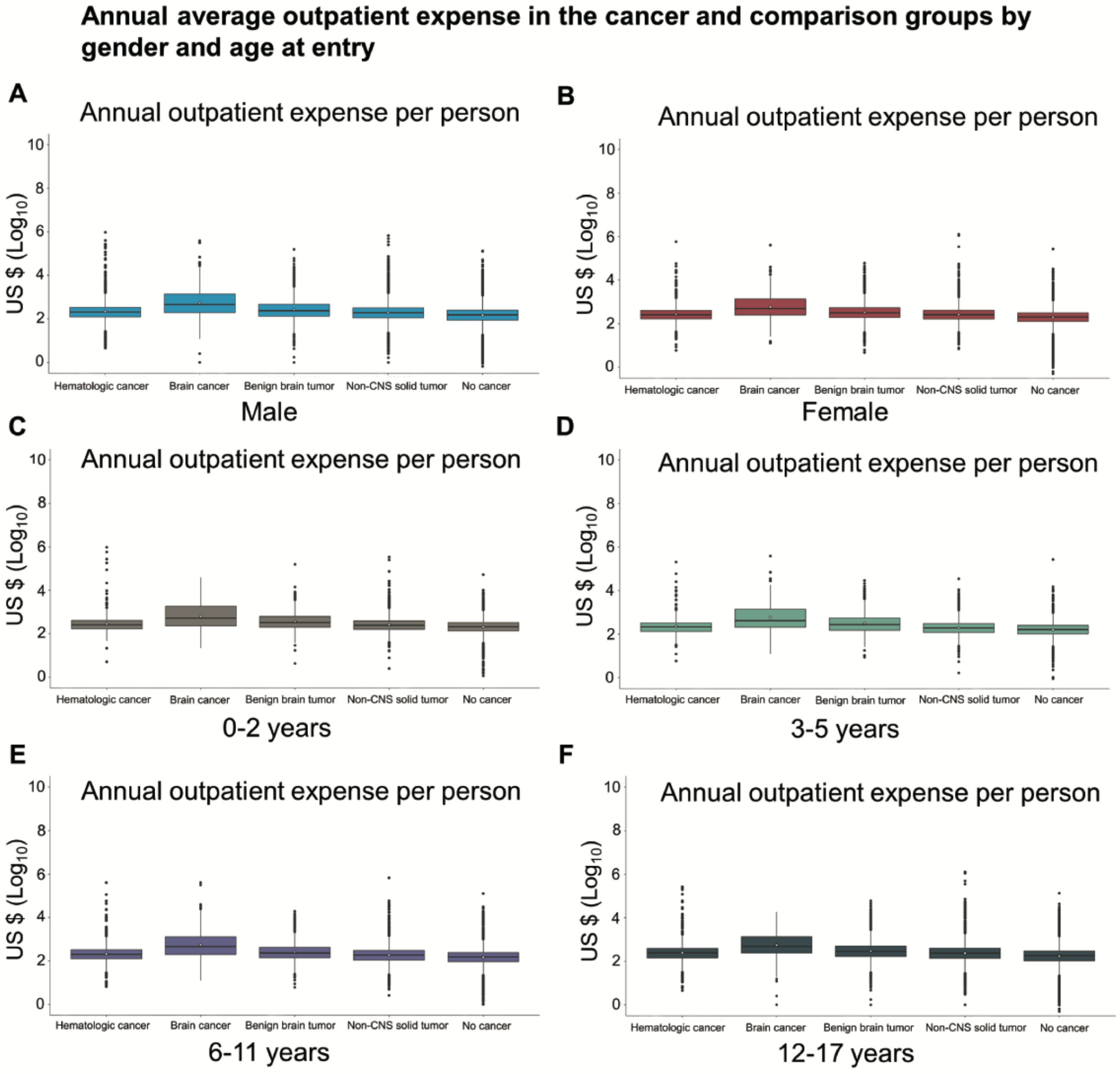


Table 4. Annual outpatient expense compared by gender.[a]

Gender	Study group									
	Cancer groups (ICD-9-CM[b])								Comparison group	
	Hematologic cancer (200-208) (n=3934)	*P* value[c]	Brain cancer (191-192) (n=2241)	*P* value[c]	Benign brain tumor (225) (n=7825)	*P* value[c]	Non-CNS[d] solid tumor (140-190; 193-199) (n=19,105)	*P* value[c]	No cancer individuals (n=64,754)	*P* value[c]
Female	255.92 (163.12-401.30)	N/A[e]	492.35 (250.84-1364.01)	N/A	311.17 (189.30-538.72)	N/A	256.09 (160.60-410.86)	N/A	201.67 (126.46-309.05)	N/A
Male	208.57 (123.75-342.92)	<.001	458.34 (193.74-1404.03)	.06	237.66 (134.68-469.39)	<.001	192.07 (111.78-322.92)	<.001	153.35 (88.48-252.09)	<.001

[a] Data represent the median expense in US $ per person per year (IQR).

[b] ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

[c] Compared with females in each group.

[d] CNS: central nervous system.

[e] N/A: not applicable.

Table 5. Annual outpatient expense compared by age at entry.[a]

Age at entry (years)	Study groups									
	Cancer groups (ICD-9-CM[b])								Comparison group	
	Hematologic cancer (200-208) (n=3934)	*P* value[c]	Brain cancer (191-192) (n=2241)	*P* value[c]	Benign brain tumor (225) (n=7825)	*P* value[c]	Non-CNS[d] solid tumor (140-190; 193-199) (n=19,105)	*P* value[c]	No cancer individuals (n=64,754)	*P* value[c]
0-2	257.74 (166.83-408.19)	Reference	525.27 (228.69-1847.69)	Reference	329.36 (197.63-621.16)	Reference	245.14 (157.84-393.75)	Reference	212.38 (136.58-329.15)	Reference
3-5	251.85 (133.05-323.90)	<.001	415.33 (204.98-1397.82)	.33	272.84 (151.28-550.89)	<.001	191.24 (118.96-305.10)	<.001	162.35 (102.46-251.89)	<.001
6-11	202.41 (126.32-330.68)	<.001	457.92 (197.78-1303.71)	.25	231.84 (140.18-419.25)	<.001	186.55 (108.48-306.03)	<.001	152.34 (92.1-239.84)	<.001
12-17	245.39 (142.83-386.15)	.01	481.73 (240.08-1364.01)	.53	286.93 (165.13-507.47)	<.001	238.28 (138.16-399.03)	<.001	182.10 (103.66-296.48)	<.001

[a] Data represent the median expense in US $ per person per year (IQR).

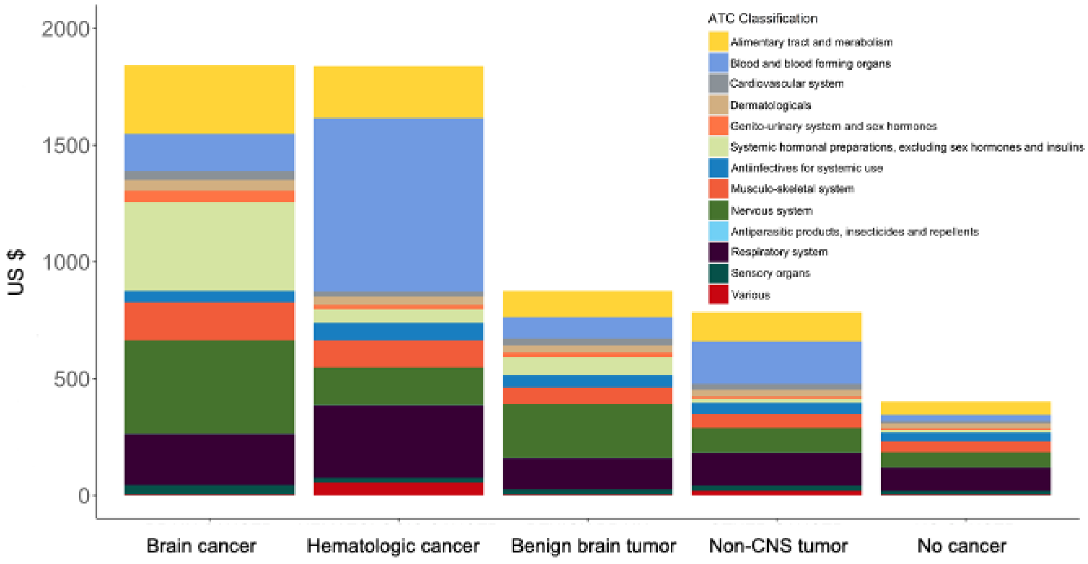
[b] ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

[c] Compared with the age of 0-2 years in each group.

[d] CNS: central nervous system.

To evaluate the composition of cost, we calculated the average outpatient medical expense per person-year of the cancer groups that was much higher than that of the no cancer group (Figure 2E). The hematologic and brain cancer groups had the highest average outpatient medical expense per person-year (4.79- and 4.01-fold of the no cancer group, respectively). The average outpatient medical expense per person-year of nonmedication service of hematologic and brain cancers was much higher (Figure 2E, brown bars). Moreover, the average outpatient medication expense per person-year was the highest in hematologic and brain cancers (Figure 4). Further, the cost was lower in medication than nonmedication services across all groups (Figure 2E). The proportion of medication cost was significantly higher in brain cancer survivors compared with no cancer individuals (33.79% vs 29.55%; *P*=.003), whereas the proportion of medication cost for survivors of hematological cancer (28.18% vs 29.55%; *P*=.31), benign brain tumor (29.16% vs 29.55%; *P*=.79), and non-CNS solid tumor (26.75% vs 29.55%; *P*=.06) was not significantly different than in the no cancer group (Figure 2E). In the brain cancer and benign brain tumor groups, medication costs for hormonal therapy, nervous system, alimentary tract, and metabolism were the major classifications of medication costs (Figures 2E and 4).

Figure 4. The average outpatient medical expense of medications per person-year in each cancer and comparison group was compared. Color bars represent the ATC Classification of medications. ATC: Anatomical Therapeutic Chemical.



Discussion

Principal Findings

In children aged 0-14 years, the annual incidence of newly diagnosed cancers (excluding benign CNS tumors) is 178 cases/million in the United States and 125 cases/million in Taiwan [50-52]. With the advent of multidisciplinary care and multimodality treatments, the 5-year survival rate of childhood cancer has now exceeded 80% [1,2], especially with leukemias and lymphomas [53-55]. Because of this success, every year the number of childhood cancer survivors continues to rise. It is now estimated that in the next 10-20 years 1/1000 individuals in Taiwan aged 20-30 years will be a survivor of childhood cancer. In the United States, the number of individuals surviving their cancer and treatments is projected to reach more than 22.1 million by 2030 [38,39].

It has been reported that 62% of survivors of childhood cancer develop a long-term health problem and 27% have a serious or life-threatening condition [56]. As a result, only 1 in 3 survivors of childhood cancer remains free of long-term health problems related to the cancer and its treatment. Thus, the treatment outcomes of childhood cancer should be not only the survival rate but also the long-term health care issues that include medical service utilization and the medical costs of the services.

There is growing information about the long-term health issues of childhood cancer survivors; however, there are very few studies describing the utilization and costs of health care for this unique population. Although the average cost of childhood cancer treatment varies among countries, estimated to be US $32,157 in Korea [57-60], the long-term health care costs of survivors remain unknown. Our data demonstrate that the utilization and costs of medical service during long-term follow-up for childhood cancer survivors are greater than those for the comparison group without a previous cancer diagnosis. In addition, the patterns of health care services that are utilized by the cancer survivors and individuals without a history of cancer are very different, with cancer survivors utilizing more hospital care and advanced care.

Comparing childhood cancer survivors with individuals without a diagnosis of cancer, the former not only had higher utilization of health care but also had much higher medication and nonmedication costs. Although the cancer survivors had much higher medication cost, the main categories of medication costs of both groups were respiratory, gastrointestinal, and neurological. The nonmedication costs were the major component of costs in both groups.

When we evaluated the health care utilization and cost of each cancer group, it was clear that children with both malignant and benign brain tumors had unique medical needs in long-term follow-ups. Specifically, brain cancer survivors had the highest utilization and costs; similarly, benign brain tumor survivors also had high utilization and costs. This was also seen in a recent publication from the French Childhood Cancer Survivor Study and the French Network of Cancer Registries where the highest median annual health care expenditure was seen in survivors of a childhood CNS tumor [61]. Regarding outpatient medication costs for these subgroups, hormonal and neurological medications comprised the 2 largest proportions of costs in brain cancer and benign brain tumor survivors. Notwithstanding, all groups of survivors have an increased annual medical cost during long-term follow-up compared with the no cancer group.

In the analysis of costs, we noted that females had significantly higher costs than their male counterparts in both the cancer survivor and no cancer groups, similar to the findings from the French study [61]. The difference was especially large in survivors of a benign brain tumor. However, the cause of cost disparities between genders could not be determined by our data and thus needs further investigation.

Our data also suggest that cancer survivors who were diagnosed at younger ages, especially those with brain tumors, will require more services in their long-term care. These data are similar to those from the French study, where the annual health care expenditures were higher in children diagnosed at older ages or with a CNS tumor [61]. Likewise, in our study, the higher cost associated with brain cancer was seen in all age groups. To address the unique medical needs of these patients, a specialized, multidisciplinary team with individualized surveillance programs needs to be established to promote health and prevent future illness. Our data also support the attempt to mitigate the long-term effects of very young children with brain tumors by reducing, delaying, or eliminating radiation therapy [16-18].

Further study is needed to define the health issues that contribute to the increased health care cost and health care burden of childhood cancer survivors. The goals for the health care system are (1) to understand the extent of the health issues that survivors of childhood cancer face; (2) to understand the specific types of health issues that survivors of childhood cancer face; (3) to develop interventions to improve the health of the survivors of childhood cancer; (4) to decrease the health care costs of survivors of childhood cancer; and (5) to reduce the health care burden of survivors of childhood cancer to the health system. The goals for the survivors of childhood cancer are (1) to mitigate the extent of their health problems; (2) to mitigate the severity of their health problems; and (3) to improve their quality of life. Further, the design of the treatment plan for children with cancer needs to minimize long-term consequences. In addition, early intervention strategies to overcome potential disabilities caused by the cancer should be implemented at the time of diagnosis.

Strengths and Limitations

There are several limitations of this study. First, this is a retrospective analysis of childhood cancer survivors in Taiwan. The proportion of hematologic cancer survivors is smaller than that of other survivorship cohorts [9,10], probably due to the relatively lower survival rate of those with acute leukemias in Taiwan during the early years of the study [62,63]. Second, we used claims data that were not directly linked to the nationwide cancer registry. The ICD-9 diagnostic codes, which were determined by physicians, could not be validated. We believe there were also a number of benign brain tumors representing tumors that are not typically considered in the benign brain tumor registries. We also believe there were a large number of benign solid tumors included in the "non-CNS solid tumor" group, which was intended to only include malignant diseases. In addition, the specific causes of costs for each cancer group could not be defined in this study and will require further analysis. Further, the history of treatment with stem cell transplantation or radiotherapy, which is known to be associated with late effects and increased medical costs, could not be analyzed in this study. The effects of these therapies should thus be evaluated in future studies. The strengths of this study are that it is a nationwide, population-based design; that the study cohort is relatively large, with the case group of childhood cancer survivors; and that there is a comparison group determined by matching the propensity score of age and gender from the claims data of the NHI that covers 99% of the Taiwanese population. Therefore, the results represent a real-world scenario, rather than assumptions and projections.

Conclusions

The cost of health care of children in Taiwan surviving childhood cancer or a brain tumor was higher than that of an age- and gender-matched comparison group of children without a history of cancer. The utilization of health care resources among children in Taiwan surviving childhood cancer and its treatment was also higher than that among an age- and gender-matched comparison group without a history of cancer. The types of health care expenditures and issues for children surviving a malignant or a benign brain tumor were different from those of the normal population. Thus, these children require coordinated follow-up and comprehensive, specialized care to optimize outcomes and quality of life. To reduce these health care costs and utilization, a careful prospective analysis of the costs and patterns of utilization will be needed. Further, preventive strategies with effective interventions will be required.

Acknowledgments

YWK, WLH, and BCS had full access to all of the data in the study and were responsible for the integrity of the data and the accuracy of the data analysis. JSM, YLL, YWK, WLH, and BCS were responsible for the concept and design of the study. JSM, YLL, YWK, SYC, and WLH acquired, analyzed, or interpreted the data. JSM, YLL, YWK, and WLH drafted the manuscript. YWK performed the statistical analyses. WLH and BCS provided administrative, technical, or material support. All authors were involved in the critical revision of the manuscript for important intellectual content, and they read and approved the final manuscript. This study was supported by grants A0111180 and A0111182 from Fu Jen Catholic University (BCS); grant 111-TMU304 from the Taipei Medical University-National Taiwan University Hospital (TMU-NTUH) Joint Research Program, grant from Pediatric Oncology Research Center Developmental Fund of Taipei Medical University, and grant 2021SKHADR019 from Shin Kong Wu Ho-Su Memorial Hospital (WLH); and the TMU Research Center of Cancer Translational Medicine Grant from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan (YLL).

Conflicts of Interest

None declared.

References

1. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. CA Cancer J Clin 2014 Jan 31;64(2):83-103 [FREE Full text] [doi: 10.3322/caac.21219] [Medline: 24488779]
2. Hudson MM, Link MP, Simone JV. Milestones in the Curability of Pediatric Cancers. JCO 2014 Aug 10;32(23):2391-2397. [doi: 10.1200/jco.2014.55.6571]
3. Landier W, Skinner R, Wallace WH, Hjorth L, Mulder RL, Wong FL, et al. Surveillance for Late Effects in Childhood Cancer Survivors. J Clin Oncol 2018 Jul 20;36(21):2216-2222 [FREE Full text] [doi: 10.1200/JCO.2017.77.0180] [Medline: 29874139]
4. Phillips S, Padgett L, Leisenring W, Stratton K, Bishop K, Krull K, et al. Survivors of childhood cancer in the United States: prevalence and burden of morbidity. Cancer Epidemiol Biomarkers Prev 2015 Apr;24(4):653-663 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-1418] [Medline: 25834148]
5. Trama A, Botta L, Foschi R, Ferrari A, Stiller C, Desandes E, EUROCARE-5 Working Group. Survival of European adolescents and young adults diagnosed with cancer in 2000-07: population-based data from EUROCARE-5. Lancet Oncol 2016 Jul;17(7):896-906 [FREE Full text] [doi: 10.1016/S1470-2045(16)00162-5] [Medline: 27237614]
6. Burdach SEG, Westhoff M, Steinhauser MF, Debatin K. Precision medicine in pediatric oncology. Mol Cell Pediatr 2018 Aug 31;5(1):6 [FREE Full text] [doi: 10.1186/s40348-018-0084-3] [Medline: 30171420]
7. Forrest SJ, Geoerger B, Janeway KA. Precision medicine in pediatric oncology. Curr Opin Pediatr 2018 Feb;30(1):17-24 [FREE Full text] [doi: 10.1097/MOP.0000000000000570] [Medline: 29189430]
8. Sweet-Cordero EA, Biegel JA. The genomic landscape of pediatric cancers: implications for diagnosis and treatment. Science 2019 Mar 15;363(6432):1170-1175 [FREE Full text] [doi: 10.1126/science.aaw3535] [Medline: 30872516]

9. Armstrong GT, Liu Q, Yasui Y, Neglia JP, Leisenring W, Robison LL, et al. Late mortality among 5-year survivors of childhood cancer: a summary from the Childhood Cancer Survivor Study. JCO 2009 May 10;27(14):2328-2338. [doi: 10.1200/jco.2008.21.1425]
10. Wasilewski-Masker K, Liu Q, Yasui Y, Leisenring W, Meacham LR, Hammond S, et al. Late recurrence in pediatric cancer: a report from the Childhood Cancer Survivor Study. J Natl Cancer Inst 2009 Dec 16;101(24):1709-1720 [FREE Full text] [doi: 10.1093/jnci/djp417] [Medline: 19966206]
11. Tsai M, Chen C, Hsieh KL, Miser JS, Chang H, Liu Y, et al. Seizure characteristics are related to tumor pathology in children with brain tumors. Epilepsy Res 2018 Nov;147:15-21. [doi: 10.1016/j.eplepsyres.2018.08.007] [Medline: 30176531]
12. Oushy S, Sillau S, Ney D, Damek D, Youssef A, Lillehei K, et al. New-onset seizure during and after brain tumor excision: a risk assessment analysis. J Neurosurg 2018 Jun;128(6):1713-1718. [doi: 10.3171/2017.2.JNS162315] [Medline: 28753117]
13. Youssef G, Wen PY. Medical and neurological management of brain tumor complications. Curr Neurol Neurosci Rep 2021 Sep 20;21(10):53. [doi: 10.1007/s11910-021-01142-x] [Medline: 34545509]
14. Erickson D, Donegan D. Diagnosis and management of neuroendocrine disorders of survivors of brain tumors. Am Soc Clin Oncol Educ Book 2021 Mar;41:1-9 [FREE Full text] [doi: 10.1200/EDBK_321059] [Medline: 33793312]
15. Sklar C, Antal Z, Chemaitilly W, Cohen L, Follin C, Meacham L, et al. Hypothalamic-pituitary and growth disorders in survivors of childhood cancer: an Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab 2018 Aug 01;103(8):2761-2784. [doi: 10.1210/jc.2018-01175] [Medline: 29982476]
16. Child AE, Warren EA, Grosshans DR, Paulino AC, Okcu MF, Ris MD, et al. Long-term cognitive and academic outcomes among pediatric brain tumor survivors treated with proton versus photon radiotherapy. Pediatr Blood Cancer 2021 Sep 11;68(9):e29125 [FREE Full text] [doi: 10.1002/pbc.29125] [Medline: 34114294]
17. Kline CN, Mueller S. Neurocognitive outcomes in children with brain tumors. Semin Neurol 2020 Jun 11;40(3):315-321. [doi: 10.1055/s-0040-1708867] [Medline: 32392601]
18. Krull KR, Hardy KK, Kahalley LS, Schuitema I, Kesler SR. Neurocognitive outcomes and interventions in long-term survivors of childhood cancer. J Clin Oncol 2018 Jul 20;36(21):2181-2189 [FREE Full text] [doi: 10.1200/JCO.2017.76.4696] [Medline: 29874137]
19. Chemaitilly W, Cohen LE, Mostoufi-Moab S, Patterson BC, Simmons JH, Meacham LR, et al. Endocrine late effects in childhood cancer survivors. JCO 2018 Jul 20;36(21):2153-2159. [doi: 10.1200/jco.2017.76.3268]
20. Chemaitilly W, Sklar C. Childhood cancer treatments and associated endocrine late effects: a concise guide for the pediatric Endocrinologist. Horm Res Paediatr 2019 Nov 7;91(2):74-82. [doi: 10.1159/000493943] [Medline: 30404091]
21. Gebauer J, Higham C, Langer T, Denzer C, Brabant G. Long-term endocrine and metabolic consequences of cancer treatment: a systematic review. Endocr Rev 2019 Jun 01;40(3):711-767. [doi: 10.1210/er.2018-00092] [Medline: 30476004]
22. van Dorp W, Mulder RL, Kremer LC, Hudson MM, van den Heuvel-Eibrink MM, van den Berg MH, et al. Recommendations for premature ovarian insufficiency surveillance for female survivors of childhood, adolescent, and young adult cancer: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCareSurFup Consortium. JCO 2016 Oct 01;34(28):3440-3450. [doi: 10.1200/jco.2015.64.3288]
23. Skinner R, Mulder RL, Kremer LC, Hudson MM, Constine LS, Bardi E, et al. Recommendations for gonadotoxicity surveillance in male childhood, adolescent, and young adult cancer survivors: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCareSurFup Consortium. Lancet Oncol 2017 Feb;18(2):e75-e90. [doi: 10.1016/S1470-2045(17)30026-8] [Medline: 28214419]
24. Clement S, Kremer L, Verburg F, Simmons J, Goldfarb M, Peeters R, et al. Balancing the benefits and harms of thyroid cancer surveillance in survivors of childhood, adolescent and young adult cancer: recommendations from the international late effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCareSurFup Consortium. Cancer Treat Rev 2018 Feb;63:28-39. [doi: 10.1016/j.ctrv.2017.11.005] [Medline: 29202445]
25. Bass JK, Knight KR, Yock TI, Chang KW, Cipkala D, Grewal SS. Evaluation and management of hearing loss in survivors of childhood and adolescent cancers: a report from the Children's Oncology Group. Pediatr Blood Cancer 2016 Jul 29;63(7):1152-1162 [FREE Full text] [doi: 10.1002/pbc.25951] [Medline: 26928933]
26. Clemens E, van den Heuvel-Eibrink MM, Mulder RL, Kremer LCM, Hudson MM, Skinner R, International Guideline Harmonization Group ototoxicity group. Recommendations for ototoxicity surveillance for childhood, adolescent, and young adult cancer survivors: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCare Consortium. Lancet Oncol 2019 Jan;20(1):e29-e41 [FREE Full text] [doi: 10.1016/S1470-2045(18)30858-1] [Medline: 30614474]
27. Ho W, Hung G, Yen H, Yang Y, Chang H, Lu M, et al. Characteristics and outcomes of second cancers in patients with childhood cancer: a report from the Taiwan Pediatric Oncology Group. J Formos Med Assoc 2022 Jan;121(1 Pt 2):350-359 [FREE Full text] [doi: 10.1016/j.jfma.2021.05.012] [Medline: 34154895]
28. American Academy of Pediatrics Section on Hematology/Oncology Children's Oncology Group. Long-term follow-up care for pediatric cancer survivors. Pediatrics 2009 Mar;123(3):906-915 [FREE Full text] [doi: 10.1542/peds.2008-3688] [Medline: 19255020]
29. Mulder RL, Kremer LCM, Hudson MM, Bhatia S, Landier W, Levitt G, et al. Recommendations for breast cancer surveillance for female survivors of childhood, adolescent, and young adult cancer given chest radiation: a report from the International

Late Effects of Childhood Cancer Guideline Harmonization Group. The Lancet Oncology 2013 Dec;14(13):e621-e629. [doi: 10.1016/s1470-2045(13)70303-6]

30. Armenian SH, Hudson MM, Mulder RL, Chen MH, Constine LS, Dwyer M, International Late Effects of Childhood Cancer Guideline Harmonization Group. Recommendations for cardiomyopathy surveillance for survivors of childhood cancer: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. Lancet Oncol 2015 Mar;16(3):e123-e136 [FREE Full text] [doi: 10.1016/S1470-2045(14)70409-7] [Medline: 25752563]
31. Bardi E, Mulder RL, van Dalen EC, Bhatt NS, Ruble KA, Burgis J, et al. Late hepatic toxicity surveillance for survivors of childhood, adolescent and young adult cancer: recommendations from the International Late Effects of childhood cancer guideline harmonization group. Cancer Treat Rev 2021 Nov;100:102296. [doi: 10.1016/j.ctrv.2021.102296] [Medline: 34571378]
32. Salz T, Baxi SS, Raghunathan N, Onstad EE, Freedman AN, Moskowitz CS, et al. Are we ready to predict late effects? A systematic review of clinically useful prediction models. Eur J Cancer 2015 Apr;51(6):758-766 [FREE Full text] [doi: 10.1016/j.ejca.2015.02.002] [Medline: 25736818]
33. Nicklin E, Velikova G, Hulme C, Rodriguez Lopez R, Glaser A, Kwok-Williams M, et al. Long-term issues and supportive care needs of adolescent and young adult childhood brain tumour survivors and their caregivers: a systematic review. Psychooncology 2019 Mar;28(3):477-487 [FREE Full text] [doi: 10.1002/pon.4989] [Medline: 30657618]
34. Sherwood PR, Cwiklik M, Donovan HS. Neuro-oncology family caregiving: review and directions for future research. CNS Oncol 2016 Jan;5(1):41-48 [FREE Full text] [doi: 10.2217/cns.15.43] [Medline: 26675885]
35. Sterckx W, Coolbrandt A, Dierckx de Casterlé B, Van den Heede K, Decruyenaere M, Borgenon S, et al. The impact of a high-grade glioma on everyday life: a systematic review from the patient's and caregiver's perspective. Eur J Oncol Nurs 2013 Feb;17(1):107-117. [doi: 10.1016/j.ejon.2012.04.006] [Medline: 22658206]
36. Bell H, Ownsworth T, Lloyd O, Sheeran N, Chambers S. A systematic review of factors related to children's quality of life and mental health after brain tumor. Psychooncology 2018 Oct 15;27(10):2317-2326. [doi: 10.1002/pon.4850] [Medline: 30071150]
37. Verwaaijen E, Catsman-Berrevoets C, Maurice-Stam H, Dessens A, Waslander R, van den Adel TPL, et al. Determinants of impairments in functioning, fatigue, and participation ability in pediatric brain tumor survivors. Neurooncol Adv 2021 Dec;3(1):vdab161 [FREE Full text] [doi: 10.1093/noajnl/vdab161] [Medline: 34988449]
38. Ekwueme DU, Zhao J, Rim SH, de Moor JS, Zheng Z, Khushalani JS, et al. Annual out-of-pocket expenditures and financial hardship among cancer survivors aged 18-64 years - United States, 2011-2016. MMWR Morb Mortal Wkly Rep 2019 Jun 07;68(22):494-499 [FREE Full text] [doi: 10.15585/mmwr.mm6822a2] [Medline: 31170127]
39. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin 2019 Sep 11;69(5):363-385 [FREE Full text] [doi: 10.3322/caac.21565] [Medline: 31184787]
40. Armstrong GT, Chen Y, Yasui Y, Leisenring W, Gibson TM, Mertens AC, et al. Reduction in late mortality among 5-year survivors of childhood cancer. N Engl J Med 2016 Mar 03;374(9):833-842 [FREE Full text] [doi: 10.1056/NEJMoa1510795] [Medline: 26761625]
41. Bhakta N, Liu Q, Ness KK, Baassiri M, Eissa H, Yeo F, et al. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). The Lancet 2017 Dec;390(10112):2569-2582. [doi: 10.1016/S0140-6736(17)31610-0]
42. Brinkman TM, Ness KK, Li Z, Huang I, Krull KR, Gajjar A, et al. Attainment of functional and social independence in adult survivors of pediatric CNS tumors: a report from the St Jude Lifetime Cohort Study. JCO 2018 Sep 20;36(27):2762-2769. [doi: 10.1200/jco.2018.77.9454]
43. Gibson T, Li Z, Green D, Armstrong G, Mulrooney D, Srivastava D, et al. Blood pressure status in adult survivors of childhood cancer: a report from the St. Jude Lifetime Cohort Study. Cancer Epidemiol Biomarkers Prev 2017 Dec;26(12):1705-1713 [FREE Full text] [doi: 10.1158/1055-9965.EPI-17-0510] [Medline: 29167278]
44. Nipp RD, Kirchhoff AC, Fair D, Rabin J, Hyland KA, Kuhlthau K, et al. Financial burden in survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. JCO 2017 Oct 20;35(30):3474-3481. [doi: 10.1200/jco.2016.71.7066]
45. Nelson B. The hidden costs of surviving childhood cancer. Cancer Cytopathol 2018 Oct 29;126(10):827-828 [FREE Full text] [doi: 10.1002/cncy.22074] [Medline: 30371991]
46. Cheng T. Reflections on the 20th anniversary of Taiwan's single-payer National Health Insurance System. Health Aff (Millwood) 2015 Mar;34(3):502-510. [doi: 10.1377/hlthaff.2014.1332] [Medline: 25732502]
47. Wu T, Majeed A, Kuo KN. An overview of the healthcare system in Taiwan. London J Prim Care (Abingdon) 2010 Dec 07;3(2):115-119 [FREE Full text] [doi: 10.1080/17571472.2010.11493315] [Medline: 25949636]
48. Lai H, Lin H, Kao Y, Wang K, Chou J, Shia B, et al. Irritable bowel syndrome increases the risk of chronic obstructive pulmonary disease: a retrospective cohort study. Sci Rep 2020 Jun 19;10(1):10008 [FREE Full text] [doi: 10.1038/s41598-020-66707-1] [Medline: 32561774]
49. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41-55. [doi: 10.1093/biomet/70.1.41]

50. Liu Y, Lo W, Chiang C, Yang Y, Lu M, Hsu W, et al. Incidence of cancer in children aged 0-14 years in Taiwan, 1996-2010. Cancer Epidemiol 2015 Feb;39(1):21-28. [doi: 10.1016/j.canep.2014.11.010] [Medline: 25599927]
51. Hung G, Horng J, Lee Y, Yen H, Chen C, Lee C. Cancer incidence patterns among children and adolescents in Taiwan from 1995 to 2009: a population-based study. Cancer 2014 Nov 15;120(22):3545-3553 [FREE Full text] [doi: 10.1002/cncr.28903] [Medline: 25043411]
52. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021 Jan 12;71(1):7-33 [FREE Full text] [doi: 10.3322/caac.21654] [Medline: 33433946]
53. Pui C. Precision medicine in acute lymphoblastic leukemia. Front Med 2020 Dec 19;14(6):689-700 [FREE Full text] [doi: 10.1007/s11684-020-0759-8] [Medline: 33074527]
54. Cairo MS, Beishuizen A. Childhood, adolescent and young adult non-Hodgkin lymphoma: current perspectives. Br J Haematol 2019 Jun 06;185(6):1021-1042 [FREE Full text] [doi: 10.1111/bjh.15764] [Medline: 30729513]
55. Mauz-Körholz C, Metzger ML, Kelly KM, Schwartz CL, Castellanos ME, Dieckmann K, et al. Pediatric Hodgkin lymphoma. JCO 2015 Sep 20;33(27):2975-2985. [doi: 10.1200/jco.2014.59.4853]
56. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, Childhood Cancer Survivor Study. Chronic health conditions in adult survivors of childhood cancer. N Engl J Med 2006 Oct 12;355(15):1572-1582. [doi: 10.1056/NEJMsa060185] [Medline: 17035650]
57. Warner EL, Kirchhoff AC, Nam GE, Fluchel M. Financial burden of pediatric cancer for patients and their families. JOP 2015 Jan;11(1):12-18. [doi: 10.1200/jop.2014.001495]
58. Chae W, Kim J, Park S, Park E, Jang S. The financial burden associated with medical costs among childhood cancer patients and their families related to their socioeconomic status: the perspective of National Health Insurance Service. Int J Environ Res Public Health 2020 Aug 19;17(17):6020 [FREE Full text] [doi: 10.3390/ijerph17176020] [Medline: 32824940]
59. Fung A, Horton S, Zabih V, Denburg A, Gupta S. Cost and cost-effectiveness of childhood cancer treatment in low-income and middle-income countries: a systematic review. BMJ Glob Health 2019 Oct 31;4(5):e001825 [FREE Full text] [doi: 10.1136/bmjgh-2019-001825] [Medline: 31749998]
60. Iragorri N, de Oliveira C, Fitzgerald N, Essue B. The out-of-pocket cost burden of cancer care - a systematic literature review. Curr Oncol 2021 Mar 15;28(2):1216-1248 [FREE Full text] [doi: 10.3390/curroncol28020117] [Medline: 33804288]
61. Bejarano-Quisoboni D, Pelletier-Fleury N, Allodji RS, Lacour B, GrosClaude P, FRANCIM Group, et al. Health care expenditures among long-term survivors of pediatric solid tumors: results from the French Childhood Cancer Survivor Study (FCCSS) and the French Network of Cancer Registries (FRANCIM). PLoS One 2022 May 26;17(5):e0267317 [FREE Full text] [doi: 10.1371/journal.pone.0267317] [Medline: 35617253]
62. Yang Y, Jaing T, Chen S, Liu H, Hung I, Lin D, et al. Treatment outcomes of pediatric acute myeloid leukemia: a retrospective analysis from 1996 to 2019 in Taiwan. Sci Rep 2021 Mar 15;11(1):5893 [FREE Full text] [doi: 10.1038/s41598-021-85321-3] [Medline: 33723338]
63. Liang D, Yang C, Lin D, Hung I, Lin K, Chen J, et al. Long-term results of Taiwan Pediatric Oncology Group studies 1997 and 2002 for childhood acute lymphoblastic leukemia. Leukemia 2010 Feb 17;24(2):397-405. [doi: 10.1038/leu.2009.248] [Medline: 20016538]

Abbreviations

ATC: Anatomical Therapeutic Chemical
CNS: central nervous system
HWDC: Health and Welfare Data Science Center
ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification
NHI: National Health Insurance

Edited by G Eysenbach; submitted 04.09.22; peer-reviewed by HJ Yen; comments to author 08.11.22; revised version received 13.12.22; accepted 22.01.23; published 02.03.23.

Please cite as:
Miser JS, Shia BC, Kao YW, Liu YL, Chen SY, Ho WL
The Health Care Utilization and Medical Costs in Long-Term Follow-Up of Children Diagnosed With Leukemia, Solid Tumor, or Brain Tumor: Population-Based Study Using the National Health Insurance Claims Data
JMIR Public Health Surveill 2023;9:e42350
URL: https://publichealth.jmir.org/2023/1/e42350
doi:10.2196/42350
PMID:36862495

Original Paper

Trajectories of Controller Therapy Use Before and After Asthma-Related Hospitalization in Children and Adults: Population-Based Retrospective Cohort Study

Manon Belhassen[1], PhD; Maeva Nolin[1], MSc; Flore Jacoud[1], MSc; Claire Marant Micallef[1], PhD; Eric Van Ganse[1,2,3], MD, PhD

[1]PELyon, Lyon, France

[2]Research on Healthcare Performance RESHAPE, INSERM U1290, Université Claude Bernard Lyon 1, Lyon, France

[3]Respiratory Medicine, Croix Rousse University Hospital, Lyon, France

Corresponding Author:
Manon Belhassen, PhD
PELyon
210 Avenue Jean Jaurès
Lyon
France
Phone: 33 481099607
Email: manon.belhassen@pelyon.fr

Abstract

Background: Inappropriate use of inhaled corticosteroids (ICSs) for asthma impairs control and may cause exacerbation, including asthma-related hospitalization (ARH). In prospective studies, ICS use peaked around ARH, but information on routine care use is limited. Since ARH is a major outcome, controller therapy use in routine care before and after ARH should be documented.

Objective: This study aimed to distinguish ICS use typologies (trajectories) before and after ARH, and assess their relationships with sociodemographic, disease, and health care characteristics.

Methods: A retrospective cohort study was performed using a 1% random sample of the French claims database. All patients hospitalized for asthma between January 01, 2013, and December 31, 2015, were classified as either children (aged 1-10 years) or teens/adults (aged ≥11 years). Health care resource use was assessed between 24 and 12 months before ARH. ICS use was computed with the Continuous Measures of Medication Acquisition-7 (CMA7) for the 4 quarters before and after ARH. Initially, the overall impact of hospitalization on the CMA7 value was studied using a segmented regression analysis in both children and teens/adults. Then, group-based trajectory modeling differentiated the groups with similar ICS use. We tested different models having 2 to 5 distinct trajectory groups before selecting the most appropriate trajectory form. We finally selected the model with the lowest Bayesian Information Criterion, the highest proportion of patients in each group, and the maximum estimated probability of assignment to a specific group.

Results: Overall, 863 patients were included in the final study cohort, of which 447 (51.8%) were children and 416 (48.2%) were teens/adults. In children, the average CMA7 value was 12.6% at the start of the observation period, and there was no significant quarter-to-quarter change in the value (*P*=.14) before hospitalization. Immediately after hospitalization, the average CMA7 value rose by 34.9% (*P*=.001), before a significant decrease (*P*=.01) of 7.0% per quarter. In teens/adults, the average CMA7 value was 31.0% at the start, and there was no significant quarter-to-quarter change in the value (*P*=.08) before hospitalization. Immediately after hospitalization, the average CMA7 value rose by 26.9% (*P*=.002), before a significant decrease (*P*=.01) of 7.0% per quarter. We identified 3 and 5 trajectories before ARH in children and adults, respectively, and 5 after ARH for both groups. Trajectories were related to sociodemographic characteristics (particularly, markers of social deprivation) and to potentially inappropriate health care, such as medical management and choice of therapy.

Conclusions: Although ARH had an overall positive impact on ICS use trajectories, the effect was often transient, and patient behaviors were heterogeneous. Along with overall trends, distinct trajectories were identified, which were related to specific patients and health care characteristics. Our data reinforce the evidence that inappropriate use of ICS paves the way for ARH.

(JMIR Public Health Surveill 2023;9:e50085) doi:10.2196/50085

KEYWORDS

asthma; hospitalization; inhaled corticosteroids; trajectories; quality of care; clustering

Introduction

Asthma is a major public health concern, and over 5% of the total population in high-income countries live with the condition [1]. Both adults and children cover the asthma demographic [2], and it not only affects quality of life but also causes significant morbidity (and in some instances mortality). It is responsible for high financial burden at both the individual and societal levels, particularly as a result of acute exacerbation [3,4].

The risk of exacerbation increases when asthma is poorly controlled, and severe exacerbation can lead to asthma-related hospitalization (ARH) [5,6]. A common cause of poor control is the inappropriate use of inhaled corticosteroids (ICSs) in monotherapy or as fixed-dose combinations (FDCs), as well as the inappropriate use of long-acting beta agonists (LABAs) in monotherapy [7,8].

In a prospective cohort study of ARH, a rapid decrease was shown in the use of ICSs and oral corticosteroids (OCSs) after hospital discharge [9]. Information on preadmission use, however, was not available. Another study clarified the overall relationship between adherence to ICSs and the occurrence of severe exacerbations, including ARH, with the use of ICSs increasing shortly before the exacerbation and further increasing afterward [10].

Nonetheless, asthmatics are likely heterogeneous in that regard, and their behaviors, particularly ICS use, may differ in real-life conditions compared with that during prospective field studies, where medication use is regularly assessed. Since ARH is a major outcome, the use of controller therapy in routine care before and after ARH should be documented. This information would incite health care professionals to adjust modifiable patient and care characteristics in order to prevent recurring exacerbations. The objectives of this study were to distinguish the typology of ICS use before and after ARH in the routine care of a representative sample of the overall asthma population, and to identify sociodemographic, disease, and health care characteristics related to ICS trajectories.

Methods

Data Source

The data source of this study was the Echantillon Généraliste des Bénéficiaires (EGB; General Sample of Beneficiaries), which is a 1/97th representative random sample of the French national health system (Système National des Données de Santé [SNDS]; National Health Data System) that covers more than 90% of the French population. It contains comprehensive, anonymized individual information on sociodemographic characteristics, date of death, out-of-hospital reimbursed health care expenditures (from both public and private health care providers), and hospital discharge summaries using the International Classification of Diseases 10th Revision (ICD-10) codes [11].

Study Population

We included all patients who were hospitalized for asthma (ICD-10 codes: J45 and J46) as the singular main diagnosis between January 1, 2013, and December 31, 2015. The first recorded date of ARH was defined as the index date. We selected patients with continuous follow-up during the period between 24 months before and 12 months after the index date. Patients hospitalized for asthma-related reasons in the 12 months before the index hospitalization were excluded. Two subgroups were defined: children (1-10 years old) and teens/adults (≥11 years old).

Study Variables

For children and teens/adults separately, we described the sociodemographic characteristics (age, gender, free access to care status as a proxy of social deprivation [given on a means-tested basis; this status supports the health expenses of more modest patients], and chronic disease status [patients who benefit from this status for a diagnosed condition are totally reimbursed for care related to that condition, ie, asthma in our study]), health care resource utilization (general practitioner visits, pediatrician visits, respiratory physician visits, emergency room [ER] visits for any cause, pulmonary function testing, and hospital admission for asthma), and initiation of asthma therapy (ICSs, LABAs, FDCs of ICSs/LABAs, leukotriene receptor antagonists [LTRAs], short-acting β2 agonists [SABAs], OCSs, and antibiotics for respiratory infections). We also described the therapeutic ratio or the “controller to total asthma medication ratio” as a proxy of the quality of asthma care. Studies have shown that patients with high therapeutic ratios (≥50%) experience fewer asthma exacerbations than those with low ratios [12-15].

Description of ICS Use

The baseline period started at T0, 365 days before the date of the index ARH (Multimedia Appendix 1). For both children and teens/adults, ICS use (in monotherapy or FDCs) was separately assessed by computing the Continuous Measures of Medication Acquisition-7 (CMA7) for 4 quarters, before the hospital admission (from T0 to the index ARH date) and after hospital discharge. CMA7 is a standard method to assess adherence to medication, calculated as the cumulative days’ supply obtained over a series of intervals divided by the total days from the beginning to the end of the study period [16,17].

Statistical Analysis

Descriptive statistics were used to describe the sociodemographic characteristics, health care utilization, and asthma therapy over the period between 24 and 12 months before the index ARH. The categorical variables of patient characteristics were described by providing the sample size within each modality and the relative percentages. We described the quantitative variables of health care utilization to determine the sample size and relative percentages for patients (with at least one care item) to calculate the mean and SD of the number of care items in all patients.

Initially, the overall impact of hospitalization on the CMA7 value was studied using a segmented regression analysis in both children and teens/adults [18]. Then, group-based trajectory modeling (the SAS proc traj function) was used to distinguish different groups of patients regarding the typologies of the trajectories of CMA7 for ICSs [19,20] both before and after ARH. We tested different models having 2 to 5 distinct trajectory groups before selecting the most appropriate form of the trajectories. We finally selected the model with the lowest Bayesian Information Criterion, the highest proportions of patients in each group, and the maximum estimated probability of assignment to a specific group. The chi-square test, Fisher test, and Kruskal-Wallis test were used to compare patient characteristics (including asthma therapy) across trajectory groups over the baseline period for children and teens/adults as separate groups.

The statistical analysis was performed using SAS Version 9.4 (SAS Institute).

Ethical Considerations

This observational study was conducted on anonymized data. The Commission Nationale de l'Informatique et des Libertés (CNIL; National Informatics and Liberty Committee) provided authorization for the use of the EGB data for research purposes, and approval was obtained from the French Institute for Health Data (Institut des Données de Santé) under number 133 (granted on June 9, 2015).

Results

Study Population

In total, 1473 patients were hospitalized for asthma from January 1, 2013, to December 31, 2015. We excluded 411 patients whose claims did not span the 2 years prior to their hospitalization, 34 patients who were hospitalized for asthma in the year before the index date, and 165 patients who had no continuous follow-up. In the end, 863 patients were included, of which 447 (51.8%) were children aged 1 to 10 years and 416 (48.2%) were teens/adults aged over 11 years (Figure 1). Regarding ICS use, similar trends were observed in both the children and teens/adults (segmented regression analysis) (Figure 2).

Figure 1. Study flowchart.

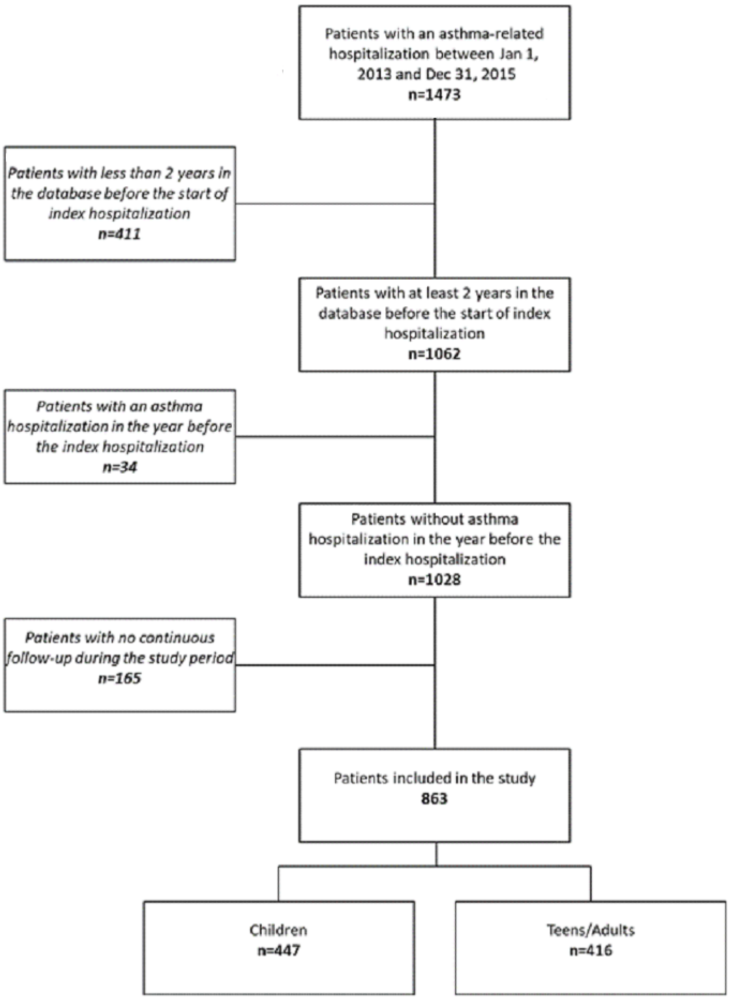
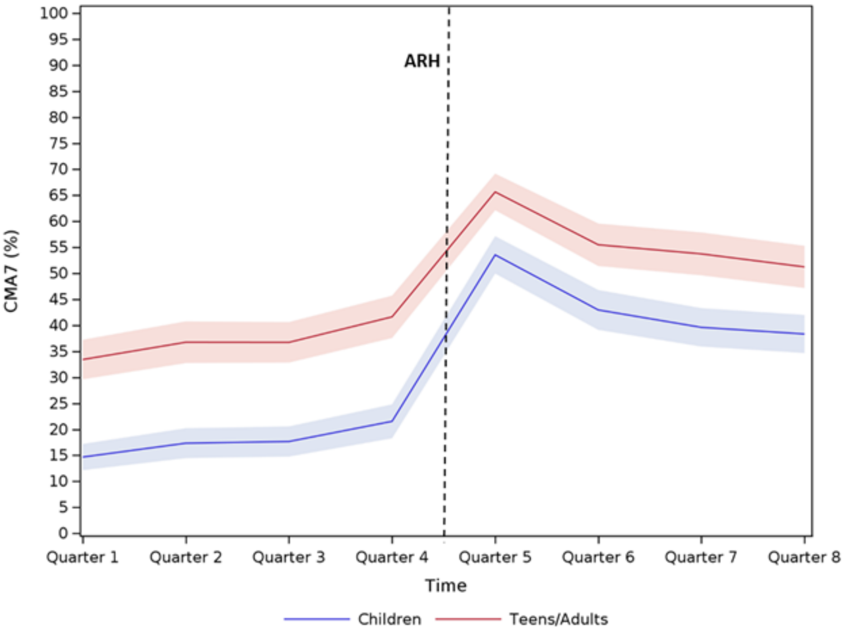


Figure 2. Inhaled corticosteroid use by children and teens/adults before and after asthma-related hospitalization (ARH). CMA7: Continuous Measures of Medication Acquisition-7.



In the children, at the start of the observation period, the average CMA7 value was 12.6%, and there was no significant quarter-to-quarter change in the value (*P*=.14) before hospitalization. Immediately after hospitalization, the average CMA7 value rose by 34.9% (*P*=.001), before a significant decrease (*P*=.01) of 7.0% per quarter.

In the teens/adults, at the start of the observation period, the average CMA7 value was 31.0%, and there was no significant quarter-to-quarter change in the value (*P*=.08) before hospitalization. Immediately after hospitalization, the average CMA7 value rose by 26.9% (*P*=.002), before a significant decrease (*P*=.01) of 7.0% per quarter.

Population Characteristics Before ARH in the Children

At baseline, the mean age of the children was 4.2 years, and most were male (291/447, 65.1%). A quarter (112/447, 25.1%) had free access to health care. Almost half of the children (213/447, 47.7%) visited a pediatrician, whereas a quarter (115/447, 25.7%) visited a respiratory physician at least once. OCSs were dispensed to 60.9% (272/447) of the children at a mean annual dose (MAD) of 240 mg, while respiratory antibiotics were administered to 75.2% (336/447) of the children, with 4 annual dispensing events on average (Table 1). Altogether, few children (119/447, 26.6%) had high therapeutic ratios (≥0.5).

Table 1. Characteristics of the children at baseline (T0) overall and in the different preasthma-related hospitalization trajectories.

Characteristic	All (N=447)	No use (n=243)	Mild use (n=172)	High use (n=32)	*P* value
Sociodemographic characteristics					
Gender, n (%)					
Male	291 (65.1)	167 (68.7)	107 (62.2)	17 (53.1)	.13
Age at baseline (years), mean (SD)	4.2 (2.6)	4.1 (2.6)	4.3 (2.7)	4.8 (2.4)	
Age group, n (%)					.79
1-5 years	320 (71.6)	177 (72.8)	120 (69.8)	23 (71.9)	
6-10 years	127 (28.4)	66 (27.2)	52 (30.2)	9 (28.1)	
Free access to care status, n (%)	112 (25.1)	65 (26.7)	39 (22.7)	8 (25.0)	.64
Chronic disease status for asthma, n (%)	12 (2.7)	1 (0.4)	8 (4.7)	3 (9.4)	<.001
Health care resource utilization before asthma-related hospitalization (12 months)					
Pediatrician visits					.007
Patients, n (%)	213 (47.7)	102 (42.0)	90 (52.3)	21 (65.6)	
Number of visits, mean (SD)	1.9 (3.1)	1.6 (2.9)	2.1 (3.1)	3.2 (4.3)	
Respiratory physician visits					.25
Patients, n (%)	115 (25.7)	56 (23.0)	51 (29.7)	8 (25.0)	
Number of visits, mean (SD)	0.4 (0.9)	0.4 (0.9)	0.5 (0.8)	0.3 (0.7)	
Pulmonary function testing					<.001
Patients, n (%)	28 (6.3)	6 (2.5)	16 (9.3)	6 (18.8)	
Number of tests, mean (SD)	1.9 (1.5)	0.0 (0.3)	0.2 (0.8)	0.3 (0.7)	
Emergency room visits					.79
Patients, n (%)	171 (38.3)	92 (37.9)	65 (37.8)	14 (43.8)	
Number of visits, mean (SD)	0.7 (1.3)	0.6 (1.0)	0.8 (1.6)	0.8 (1.4)	
Asthma-related hospitalization					.006
Patients, n (%)	27 (6.0)	7 (2.9)	18 (10.5)	2 (6.3)	
Number of stays, mean (SD)	0.1 (0.5)	0.0 (0.2)	0.2 (0.7)	0.1 (0.6)	
Asthma therapy					
ICSs[a]					<.001
Patients, n (%)	168 (37.6)	54 (22.2)	93 (54.1)	21 (65.6)	
Number, mean (SD)	1.1 (2.4)	0.4 (1.0)	1.7 (2.6)	3.3 (5.5)	
FDCs[b] of ICSs/LABAs[c]					<.001
Patients, n (%)	54 (12.1)	10 (4.1)	35 (20.3)	9 (28.1)	
Number, mean (SD)	0.4 (1.5)	0.1 (0.4)	0.7 (1.7)	1.8 (3.4)	
LABAs (in a separate canister)					N/A
Patients, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Number, mean (SD)	N/A[d]	N/A	N/A	N/A	
LTRAs[e]					<.001
Patients, n (%)	59 (13.2)	11 (4.5)	34 (19.8)	14 (43.8)	
Number, mean (SD)	0.5 (1.6)	0.1 (0.4)	0.5 (1.4)	2.9 (4.2)	
SABAs[f]					<.001
Patients, n (%)	233 (52.1)	91 (37.4)	114 (66.3)	28 (87.5)	
Number, mean (SD)	1.5 (2.3)	0.8 (1.3)	2.0 (2.3)	4.3 (4.2)	

Characteristic	All (N=447)	No use (n=243)	Mild use (n=172)	High use (n=32)	*P* value
OCSs[g]					<.001
Patients, n (%)	272 (60.9)	118 (48.6)	126 (73.3)	28 (87.5)	
Number, mean (SD)	1.3 (1.5)	0.8 (1.1)	1.7 (1.6)	2.7 (2.1)	
Cumulative OCS dose in mg equivalent prednisone, mean (SD)	240.5 (376.2)	131.3 (223.2)	337.2 (414.3)	550.4 (675.3)	<.001
Respiratory antibiotics					<.001
Patients, n (%)	336 (75.2)	177 (72.8)	132 (76.7)	27 (84.4)	
Number, mean (SD)	3.9 (4.9)	3.4 (5.1)	4.3 (4.4)	6.1 (5.3)	
ICS/R03[h] ratio, n (%)					<.001
0	56 (12.5)	37 (15.2)	16 (9.3)	3 (9.4)	
<0.5	81 (18.1)	20 (8.2)	47 (27.3)	14 (43.8)	
≥0.5	119 (26.6)	43 (17.7)	64 (37.2)	12 (37.5)	
Not assessable[i]	191 (42.7)	143 (48.8)	45 (26.2)	3 (9.4)	

[a] ICS: inhaled corticosteroid.

[b] FDC: fixed-dose combination.

[c] LABA: long-acting beta agonist.

[d] N/A: not applicable.

[e] LTRA: leukotriene receptor antagonist.

[f] SABA: short-acting β2 agonist.

[g] OCS: oral corticosteroid.

[h] R03: group of medications used in the treatment of obstructive airway diseases (Anatomical Therapeutic Chemical classification).

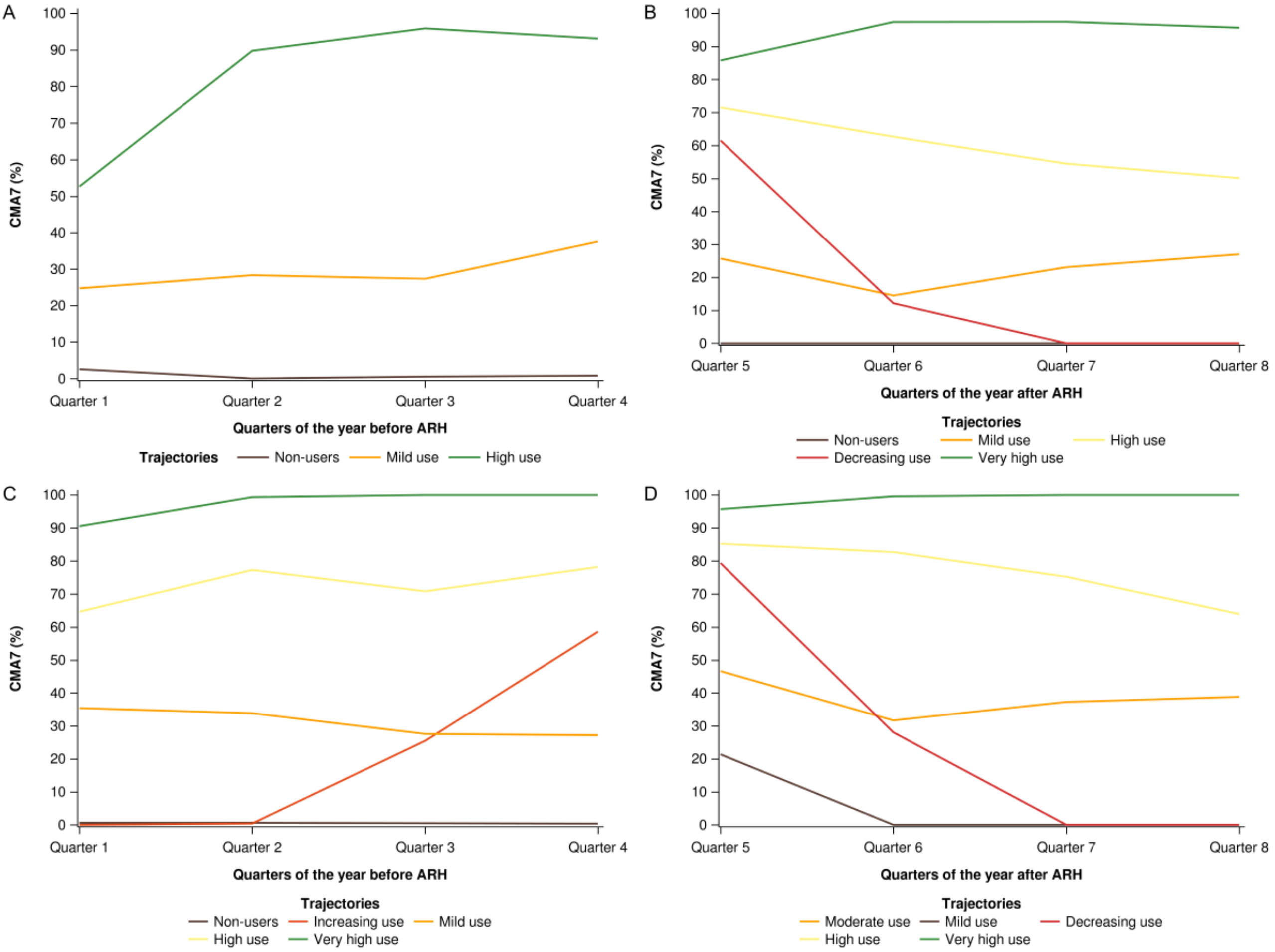
[i] Patients not receiving any respiratory therapy.

ICS Trajectories Before ARH in the Children

Three trajectories of ICS use were distinguished before ARH in the children: (1) no use (CMA7 value around 0%), (2) mild use (CMA7 value of 15%-40%), and (3) high use (CMA7 value of >80% during at least 3 quarters of the year before ARH) (Figure 3A). Just over half of the children (243/447, 54.4%) had a no-use typology, 38.5% (172/447) had a mild-use typology, and 7.2% (32/447) had a high-use typology during the last 3 quarters before ARH.

The no-use group comprised a larger percentage of males (167/243, 68.7%) compared to the mild-use (107/172, 62.2%) or high-use (17/32, 53.1%) group. There was extensive use of OCSs, with rates of 48.6% (118/243; MAD of 223 mg), 73.3% (126/172; MAD of 414 mg), and 87.5% (28/32; MAD of 675 mg) among the no-use, mild-use, and high-use groups, respectively. Respiratory antibiotics were dispensed to 72.8% (177/243), 76.7% (132/172), and 84.4% (27/32) of the children in the no-use, mild-use, and high-use groups, respectively (Table 1). The children in the mild-use (141/172, 82.0%) and high-use (29/32, 90.6%) groups had more visits to the respiratory physician or pediatrician than those in the no-use group (158/243, 65.0%).

Figure 3. Inhaled corticosteroid trajectories of children (n=447) and teens/adults (n=416), before (A and C, respectively), and after (B and D, respectively) asthma-related hospitalization (ARH). CMA7: Continuous Measures of Medication Acquisition-7.



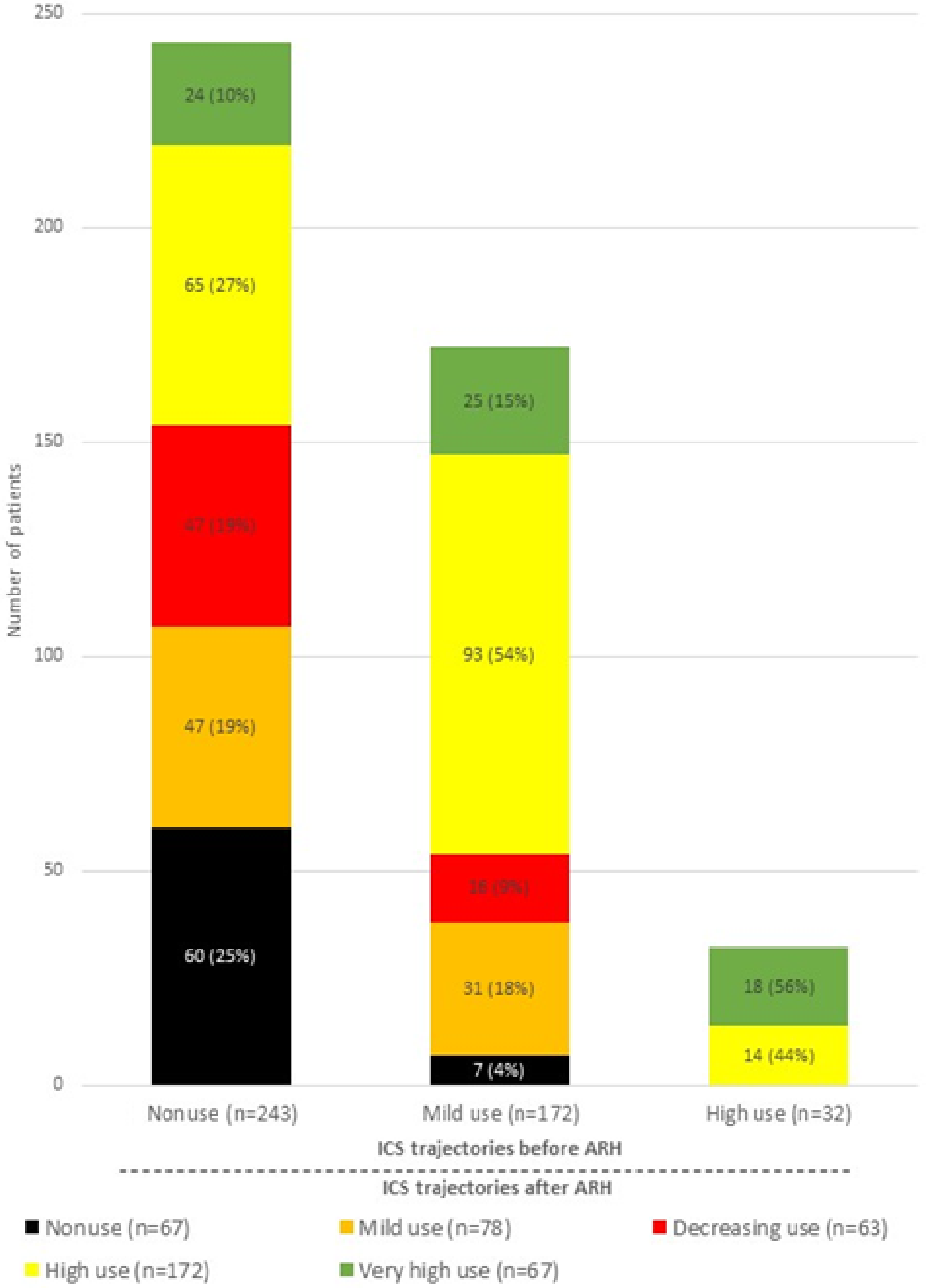
ICS Trajectories After ARH in the Children

In the year following ARH, 5 trajectories were distinguished in the children (Figure 3B). More than half of the children (239/447, 53.5%) had CMA7 values >50%. These included either very high users (n=67) with CMA7 values of >80% throughout the year or high users (n=172) with CMA7 values declining from 70% to 50% throughout the year. Nonusers (CMA7=0) corresponded to 15.0% (67/447) of all children. Mild users (n=78) had stable CMA7 values (20%-30%) throughout the following year, whereas decreasing users (n=63) had CMA7 values of 60% at the beginning of the year, which declined to zero by the end of the year.

Changes in Trajectories (Before-After ARH) in the Children

All children in the high-use group before ARH maintained a high-use profile during the year after hospitalization (Figure 4). Almost a quarter (60/243, 24.7%) of nonusers before ARH remained unchanged in the following year, whereas one-third (89/243, 36.6%) of nonusers switched to the high-use group after ARH. More than two-thirds (149/243, 61.3%) of mild users before ARH increased their use of ICSs after ARH, with 54.1% (93/172) becoming high users and 14.5% (25/172) becoming very high users. Graphs presenting the between-group interchange of individual patients are presented in Multimedia Appendix 2.

Figure 4. Distribution of inhaled corticosteroid (ICS) trajectories of children before and after asthma-related hospitalization (ARH) (n=447).



Population Characteristics Before ARH in the Teens/Adults

At baseline, the mean age of the teens/adults was 45.7 years, and most patients were female (254/416, 61.1%). Moreover, 13.5% (56/416) had free access to care, 15.4% (64/416) had a chronic disease status for asthma, and 38.5% (160/416) had visited a respiratory physician or pediatrician. Almost one-third (137/416, 32.9%) of the patients had one or more ER visits, and 30.0% (125/416) underwent pulmonary function testing (Table 2). At the time of the study, 11.8% (49/416) of the teens/adults were using LABAs in monotherapy, 70.4% (293/416) received respiratory antibiotics, and 55.5% (231/416) received OCSs. Additionally, few patients (112/416, 26.9%) had high therapeutic ratios (≥0.5).

Table 2. Characteristics of the teens/adults at baseline (T0) overall and in the different preasthma-related hospitalization trajectories.

Characteristic	All (N=416)	No use (n=134)	Increasing use (n=45)	Mild use (n=100)	High use (n=82)	Very high use (n=55)	*P* value
Sociodemographic characteristics							
Gender, n (%)							
Male	162 (38.9)	56 (41.8)	18 (40.0)	36 (36.0)	27 (32.9)	25 (45.5)	.55
Age at baseline (years), mean (SD)	45.7 (22.5)	37.5 (21.1)	44.0 (21.0)	44.6 (23.5)	54.0 (21.1)	56.8 (18.5)	
Age group, n (%)							<.001
11-35 years	148 (35.6)	75 (56.0)	15 (33.3)	36 (36.0)	16 (19.5)	6 (10.9)	
36-50 years	88 (21.2)	24 (17.9)	14 (31.1)	20 (20.0)	17 (20.7)	13 (23.6)	
51-70 years	110 (26.4)	22 (16.4)	10 (22.2)	24 (24.0)	30 (36.6)	24 (43.6)	
≥71 years	70 (16.8)	13 (9.7)	6 (13.3)	20 (20.0)	19 (23.2)	12 (21.8)	
Free access to care status, n (%)	56 (13.5)	14 (10.4)	7 (15.6)	22 (22.0)	6 (7.3)	7 (12.7)	.04
Chronic disease status for asthma, n (%)	64 (15.4)	4 (3.0)	6 (13.3)	12 (12.0)	21 (25.6)	21 (38.2)	<.001
Health care resource utilization before ARH[a] (12 months)							
Respiratory physician or pediatrician visits							<.001
Patients, n (%)	160 (38.5)	37 (27.6)	22 (48.9)	43 (43.0)	40 (48.8)	18 (32.7)	
Number, mean (SD)	0.9 (1.8)	0.5 (1.2)	1.2 (2.1)	0.9 (1.6)	1.2 (1.8)	0.9 (3.0)	
Pulmonary function testing							<.001
Patients, n (%)	125 (30.0)	7 (5.2)	15 (33.3)	30 (30.0)	39 (47.6)	34 (61.8)	
Number, mean (SD)	0.6 (1.2)	0.1 (0.4)	0.7 (1.5)	0.5 (1.1)	0.9 (1.3)	1.6 (1.7)	
ER[b] visits							.19
Patients, n (%)	137 (32.9)	43 (32.1)	14 (31.1)	42 (42.0)	26 (31.7)	12 (21.8)	
Number, mean (SD)	0.6 (1.2)	0.5 (1.0)	0.6 (1.8)	0.8 (1.5)	0.5 (1.0)	0.3 (0.7)	
ARH							.89
Patients, n (%)	12 (2.9)	3 (2.2)	2 (4.4)	2 (2.0)	3 (3.7)	2 (3.6)	
Number, mean (SD)	0.1 (0.4)	0.1 (0.5)	0.0 (0.2)	0.0 (0.1)	0.0 (0.2)	0.1 (0.8)	
Asthma therapy							
ICSs[c]							<.001
Patients, n (%)	130 (31.3)	19 (14.2)	12 (26.7)	40 (40.0)	32 (39.0)	27 (49.1)	
Number, mean (SD)	1.3 (3.6)	0.3 (1.0)	0.7 (1.7)	1.3 (2.5)	1.6 (2.9)	3.7 (7.8)	
FDCs[d] of ICSs/LABAs[e]							<.001
Patients, n (%)	212 (51.0)	24 (17.9)	21 (46.7)	60 (60.0)	61 (74.4)	46 (83.6)	
Number, mean (SD)	3.4 (5.3)	0.3 (0.8)	1.4 (2.0)	2.4 (2.8)	5.7 (5.2)	10.9 (7.9)	
LABAs (in a separate canister)							.05
Patients, n (%)	49 (11.8)	7 (5.2)	6 (13.3)	15 (15.0)	11 (13.4)	10 (18.2)	
Number, mean (SD)	0.9 (3.3)	0.4 (2.0)	1.2 (3.9)	0.7 (2.2)	0.8 (2.7)	2.6 (6.2)	
LTRAs[f]							<.001
Patients, n (%)	104 (25.0)	7 (5.2)	11 (24.4)	23 (23.0)	32 (39.0)	31 (56.4)	
Number, mean (SD)	1.7 (3.8)	0.3 (1.8)	1.6 (3.5)	1.2 (2.9)	2.2 (3.9)	5.6 (5.9)	
SABAs[g]							<.001
Patients, n (%)	259 (62.3)	49 (36.6)	25 (55.6)	79 (79.0)	62 (75.6)	44 (80.0)	
Number, mean (SD)	3.9 (6.9)	1.0 (2.3)	4.5 (7.8)	3.5 (3.6)	4.6 (5.8)	9.9 (13.3)	

Characteristic	All (N=416)	No use (n=134)	Increasing use (n=45)	Mild use (n=100)	High use (n=82)	Very high use (n=55)	*P* value
Respiratory antibiotics							<.001
Patients, n (%)	293 (70.4)	86 (64.2)	28 (62.2)	71 (71.0)	58 (70.7)	50 (90.9)	
Number, mean (SD)	5.0 (7.6)	2.8 (3.6)	3.6 (5.5)	4.6 (7.1)	6.3 (7.9)	9.7 (12.9)	
OCSs[h]							<.001
Patients, n (%)	231 (55.5)	58 (43.3)	26 (57.8)	60 (60.0)	54 (65.9)	33 (60.0)	
Number, mean (SD)	2.4 (5.8)	1.2 (3.5)	2.9 (5.5)	1.9 (3.2)	3.7 (9.1)	4.0 (7.4)	
Cumulative OCS dose in mg equivalent prednisone, mean (SD)	695.5 (1319.4)	350.5 (703.5)	806.3 (1476.6)	702.2 (1072.3)	797.1 (1163.5)	1281.9 (2360.4)	<.001
ICS/R03[i] ratio, n (%)							<.001
0	44 (10.6)	28 (20.9)	6 (13.3)	8 (8.0)	2 (2.4)	0 (0.0)	
<0.5	168 (40.4)	19 (14.2)	17 (37.8)	50 (50.0)	42 (51.2)	40 (72.7)	
≥0.5	112 (26.9)	19 (14.2)	9 (20.0)	35 (35.0)	36 (43.9)	13 (23.6)	
Not assessable[j]	92 (22.1)	68 (50.7)	13 (28.9)	7 (7.0)	2 (2.4)	2 (3.6)	

[a] ARH: asthma-related hospitalization.

[b] ER: emergency room.

[c] ICS: inhaled corticosteroid.

[d] FDC: fixed-dose combination.

[e] LABA: long-acting beta agonist.

[f] LTRA: leukotriene receptor antagonist.

[g] SABA: short-acting β2 agonist.

[h] OCS: oral corticosteroid.

[i] R03: group of medications used in the treatment of obstructive airway diseases (Anatomical Therapeutic Chemical classification).

[j] Patients not receiving any respiratory therapy.

ICS Trajectories Before ARH in the Teens/Adults

Among the teens/adults, 5 ICS trajectories were distinguished before ARH. These were as follows: (1) no use (CMA7 value of about 0; 134/416, 32.2%), (2) mild use (CMA7 value of 30%-35%; 100/416, 24.0%), (3) increasing use over time (CMA7 value rising from 0% 1 year before ARH to 60% at the time of ARH; 45/416, 10.8%), (4) high use (CMA7 value of 60%-80%; 82/416, 19.7%), and (5) very high use (CMA7 value of >90%; 55/416, 13.2%) (Figure 3C). More than half of nonusers (75/134, 56.0%) were aged between 11 and 35 years, whereas almost two-thirds of high (49/82, 59.8%) or very high (36/55, 65.5%) users were older than 50 years. The sex ratios also differed between groups, with males representing 41.8% (56/134) of nonusers versus 32.9% (27/82) of high users (Table 2). A chronic disease status for asthma was more frequent among high and very high users (21/82, 25.6% and 21/55, 38.2%, respectively) than among nonusers (4/134, 3.0%). Nonusers and very high users had fewer visits to a respiratory physician (37/134, 27.6% and 18/55, 32.7%, respectively) compared with the other groups (>40%). In terms of asthma therapy, very high users were frequently treated with FDCs, LABAs (in monotherapy), SABAs (10 units per year), respiratory antibiotics (91% of users receiving 10 units per year on average), and OCSs (total annual dose of 1282 mg), while their therapeutic ratios were low overall (40/55, 72.7%).

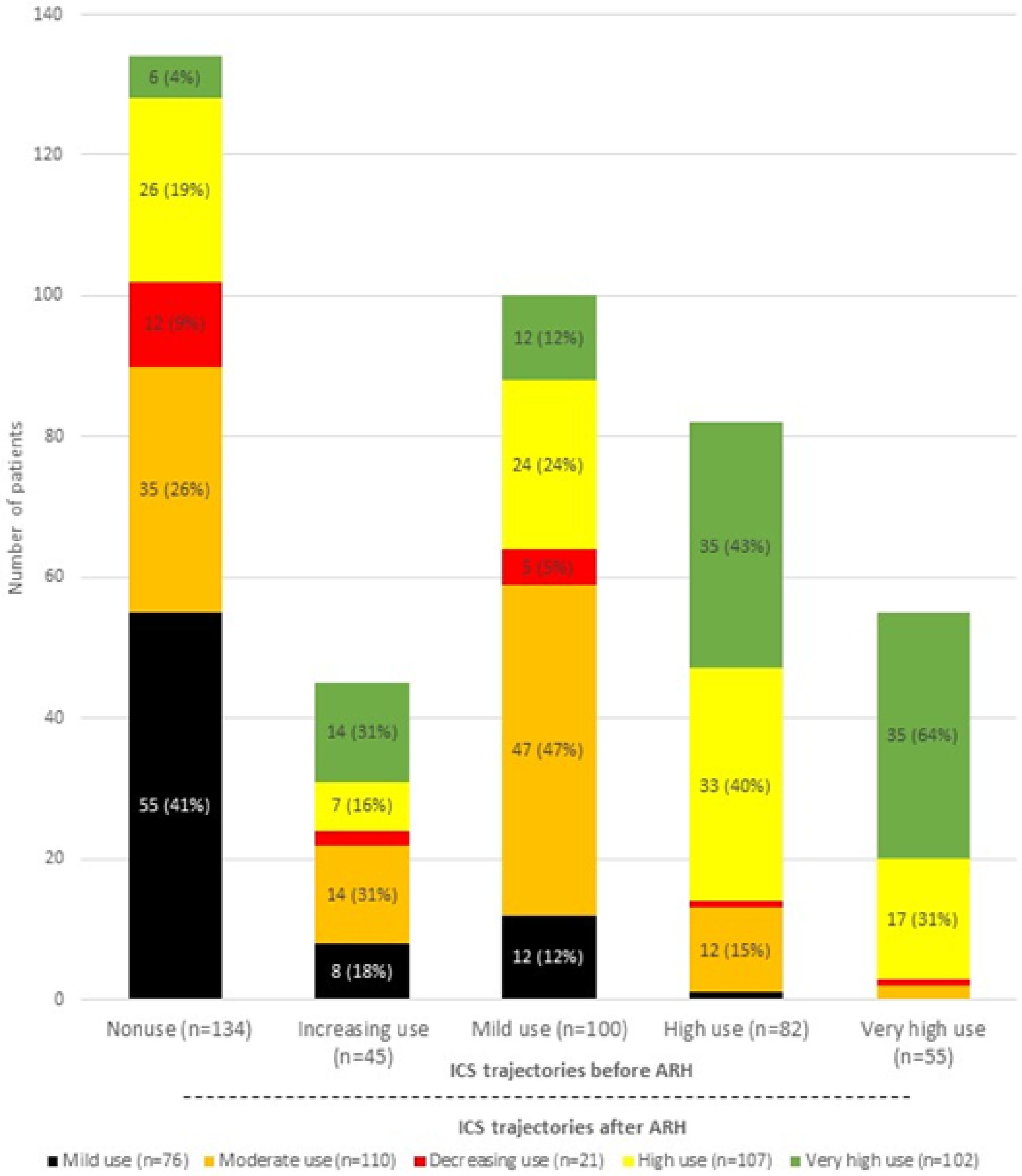
ICS Trajectories After ARH in the Teens/Adults

These patients were also distributed across 5 groups in the year following ARH. Half of them (209/416, 50.2%) were very high (CMA7 value of >90%) or high (CMA7 value of >65%) users, whereas 44.7% (186/416) were mild (CMA7 value of <20%) or moderate (CMA7 value of 30%-50%) users. The last group included a few patients (21/416, 5.1%) with decreasing use during the year following ARH, from 80% shortly after ARH to zero after the third quarter of the subsequent year (Figure 3D).

Changes in Trajectories (Before-After ARH) in the Teens/Adults

While 23.9% (99/416) of nonusers before ARH started to use ICS regularly after ARH, most patients in that group remained in an irregular use profile in the year following their hospital stay, with 41.0% (55/134) in the mild use group and 9.0% (12/134) in the decreasing use group (Figure 5). Another quarter (32/134, 26.1%) of nonusers increased their use constantly throughout the subsequent year and were thus transferred to the moderate use group.

Figure 5. Distribution of inhaled corticosteroid (ICS) trajectories of teens/adults before and after asthma-related hospitalization (ARH) (n=416).



In the group with increasing ICS use before ARH, 31.1% (14/45) stayed at a similar CMA7 level and 31.1% (14/45) switched to the very-high-use group. Lastly, most patients in the other 3 groups (mild, high, and very high use) exhibited a similar profile of ICS use after ARH compared with the year prior to ARH. Patients with CMA7 values of >60% either stayed in the same profile group or switched to the closest similar profile group (high-use to very-high-use group at 42.7% [35/82] and very-high-use to high-use group at 30.9% [17/55]). Patients in the mild-use group before ARH mostly stayed in this profile group, while 47.0% (47/100) of them joined the moderate-use group during the year after ARH. The overall trajectories of the teen/adult patients switching from groups are presented in Multimedia Appendix 3.

Discussion

Our study of the long-term routine care of pediatric and adult patients admitted to the hospital for asthma exacerbation (ARH) showed major temporal changes in ICS use before and after ARH. ARH seems to have similar effects on the typology of use in children and teens/adults. Indeed, in both populations, segmented regression analyses showed 3 contrasting segments:

pre-ARH segment (Q1-Q3) with stable use, ARH segment (Q4-Q5) with increasing use, and post-ARH segment (Q6-Q8) with use decreasing again. Interestingly, for children, the post-ARH levels of ICS use did not appear to decline to the pre-ARH levels. As a rule, trajectories were related to patient sociodemographic characteristics, health care pathways, and asthma therapy, in the context of overall poor quality of care.

Overall, our study identified 3 sets of behaviors regarding the impact of ARH on ICS use. First, in terms of improved typology, 46.3% (207/447) of nonusers or mild users before ARH increased their ICS use after hospital stay. This was the case for 46.5% (208/447) of children and 39.9% (166/416) of teens/adults with low ICS use before ARH. This may correspond to a recent or reconfirmed diagnosis of asthma.

Second, in terms of the limited impact of ARH on typology, most regular users (high-use and very-high-use groups) before ARH remained in the high or very-high-use group after ARH. These patients likely had been identified as experiencing more severe or difficult asthma before ARH and remained regularly treated. This assumption is supported by patient characteristics, such as a high rate of a chronic disease status, which is typically allocated to more severe patients, in addition to a high proportion of OCS users.

Lastly, in terms of decreased use, a minority of patients from the high-use or very-high-use group shifted to the decreasing-use or low-use group.

In all trajectories, a majority of the children had visited a pediatrician or a respiratory physician, and many children had visited the ER, suggesting that asthma was suspected before ARH, despite the low age of the pediatric sample population [21]. The use of OCSs and respiratory antibiotics was also very common before ARH, with a large majority of patients receiving respiratory antibiotics and OCSs among high ICS users. In this group, the annual dose of OCSs far exceeded the dose considered to be the threshold for OCS severe side effects [22]. The therapeutic ratio in this group was also of concern since almost half (14/32, 44%) of the children had a low ratio during the baseline period, implying an overuse of SABAs [23]. Overall, the data may reflect severity, but they do also suggest an overall poor quality of asthma care in the pediatric population.

Among teens/adults, the ratios were low (<0.5) in many teens/adults in the baseline period, which again reflects inappropriate care of asthma. The use of OCSs was also high in all teens/adults, and was particularly high (MAD of >1000 mg) in the very-high-use group, in addition to very high use of SABAs and respiratory antibiotics. More than one in 10 teens/adults inappropriately received LABAs in a separate canister [24].

Overall, the data show that although the clustering-derived typologies differed between children and teens/adults in the pre-ARH period, the global trajectories of these 2 independent samples were similar, supporting the validity of our results. A major finding is that ARH occurred in the context of low quality of care and overuse of medical resources in a population largely affected by social deprivation, as already pointed out by prior studies [25]. Our study does indeed illustrate the high proportion of patients visiting ERs before ARH (more than one-third of children [171/447, 38.3%] and teens/adults [137/416, 32.9%]), the low therapeutic ratios, and the extensive use of OCSs and respiratory antibiotics.

Regarding existing literature, studies assessing ICS use in asthma over an extended period are scarce [26]. This is particularly the case outside the context of specific prospective cohorts and thus limits the comparison with other data. In a prospective study, Krishnan et al [9] showed a rapid decrease in the use of ICSs and OCSs by asthmatic patients after hospital discharge, but no information was provided on their use before admission. Another study showed inappropriate use in children after ARH; however, again, there was no assessment of their use before ARH [27]. Similarly, Williams et al [10] clarified the overall relationship between changing ICS use and the occurrence of severe exacerbations, including ARH. Prior data on the relationship between ICS use and exacerbations may however not be generalizable since asthmatics are heterogeneous in their use of therapy and their behaviors may differ in routine care compared to investigational conditions. Of interest, a recent Danish study investigated the effect of hospital admissions for acute exacerbation on the adherence rate to controller medication in 241 adults with asthma. Their results were similar to ours in that an initial improvement in adherence to ICSs after hospital admission occurred in patients who had previously shown poor adherence. The study also showed that the improvement was transient and decreased over time during the 6 months after discharge [28]. International comparisons are however complex since patient behaviors toward therapy also depend on the local organization of health care with, for instance, the level of care co-payments, or the accessibility to health care professionals, which may affect the relationship between ICS use and exacerbation. This underlines the need to replicate investigations in different settings.

Our study had several strengths. It was a population-based study, and it was conducted with a random sample originating from data compiled from the health care consumption information of almost all French citizens (over 67 million inhabitants), without distinction of age, gender, ethnicity, residency, income, and social status, in a large country with a single universal health care system. These strengths allowed us to overcome the limitations of other studies on asthma care, such as distinct health care coverage for different patient groups or missing information as a result of fragmented care [29].

The study also had some limitations. There was no information on the education level of the patients and whether this had an impact on their trajectory. There was also no information on smoking status, pulmonary function, or BMI that may have impacted the effects of ICSs. We used CMA7 to assess ICS use. Considering that claims information provides medicine dispensing data, CMA7 measures were based on the dispensing number and frequency but did not reflect actual medication use or provide data on the quality of the inhaler technique. Additionally, we included all patients undergoing ARH during the study period independent of their ICS use and asthma diagnosis before ARH. As an alternative, we could have restricted the inclusion to patients who received an initial diagnosis of asthma during their hospital stay. Such cases were however rare among both children and teens/adults, as evidenced by the high percentage of visits to a respiratory specialist (115/447, 25.7% and 160/416, 38.5%, respectively), and by the common records of pulmonary function tests performed during the baseline period. Consequently, the impact on the results should be limited. Another limitation may have been due to the exclusion of patients who did not have continuous follow-up information recorded in the database either before or after ARH. However, those patients were excluded for administrative reasons, that is, they were affiliated to rare insurance schemes not included in the database. We feel confident that there is no reason why those patients would differ in their typologies compared with the patients included in the cohort. Finally, it may be of interest to reproduce a similar analysis in a larger population, as our sample was limited to 447 children and 416 teens/adults.

Overall, our study demonstrated inappropriate ICS use in routine care before ARH, which transiently improved shortly before and after discharge in the context of overall poor quality of care, in a largely socially deprived population experiencing persistent asthma. Our results suggest that interventions are urgently needed to improve the use of controller therapy for asthma in routine care [30], particularly during or after ARH [31]. These interventions should be based on specific patient profiles and trajectories of use before ARH. For instance, therapeutic education with action plans [32,33] or the setup of maintenance and relief therapy [34] could be a priority for some trajectories. As a rule, clinicians should emphasize the need to make regular use of controllers at appropriate doses and to make parsimonious use of relievers and OCSs. The findings are consistent with those of a recent study on care for persistent asthma that showed a decrease in primary care physician involvement and an increase in the incidence of ER visits between 2006 and 2016 [35].

In conclusion, we showed that although ARH had an overall positive impact on the trajectories of ICS use, this effect was often transient, patient behaviors were heterogeneous, and several distinctly different trajectories were identified. Additional patient and care characteristics could further improve the understanding of our findings and help more efficiently target patients in need of preventive interventions. Our data reinforce the evidence that inappropriate use of ICSs paves the way for ARH, and thus, the findings should galvanize health service efforts to keep individual ICS use at optimal levels.

Acknowledgments

The authors would like to thank the Caisse Nationale de l'Assurance Maladie for providing data, as well as Marty Brucato and Sarina Yaghobian from AcaciaTools for their medical writing and reviewing services. This work was supported by PELyon.

Data Availability

Owing to National Health Service and Système National des Données de Santé (National Health Data System) rules, no data sharing is possible, as access to data is restricted to habilitated and qualified researchers (MN and FJ are habilitated and qualified).

Authors' Contributions

MB and EVG designed the study and were responsible for producing the initial draft of the paper. CMM was responsible for producing a comprehensive version of the paper and for integrating all author comments to produce the final version. MN and FJ carried out the primary statistical analysis, constructed the figures and tables, and reviewed the manuscript to produce a finalized version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study design.
[PDF File (Adobe PDF File), 49 KB - publichealth_v9i1e50085_app1.pdf]

Multimedia Appendix 2
Use of inhaled corticosteroid trajectories before and after asthma-related hospitalization by group before asthma-related hospitalization among children (n=447).
[PDF File (Adobe PDF File), 393 KB - publichealth_v9i1e50085_app2.pdf]

Multimedia Appendix 3
Use of inhaled corticosteroid trajectories before and after asthma-related hospitalization by group before asthma-related hospitalization among teens/adults (n=416).
[PDF File (Adobe PDF File), 611 KB - publichealth_v9i1e50085_app3.pdf]

References

1. GBD Chronic Respiratory Disease Collaborators. Prevalence and attributable health burden of chronic respiratory diseases, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med 2020 Jun;8(6):585-596 [FREE Full text] [doi: 10.1016/S2213-2600(20)30105-3] [Medline: 32526187]
2. The European Lung White Book. European Respiratory Society. URL: https://www.ersnet.org/the-european-lung-white-book/ [accessed 2023-09-16]
3. Sadatsafavi M, Rousseau R, Chen W, Zhang W, Lynd L, FitzGerald JM. The preventable burden of productivity loss due to suboptimal asthma control: a population-based study. Chest 2014 Apr;145(4):787-793. [doi: 10.1378/chest.13-1619] [Medline: 24337140]
4. Van Ganse E, Laforest L, Pietri G, Boissel J, Gormand F, Ben-Joseph R, et al. Persistent asthma: disease control, resource utilisation and direct costs. Eur Respir J 2002 Aug 01;20(2):260-267 [FREE Full text] [doi: 10.1183/09031936.02.02542001] [Medline: 12212953]
5. Dima AL, Hernandez G, Cunillera O, Ferrer M, de Bruin M, ASTRO-LAB group. Asthma inhaler adherence determinants in adults: systematic review of observational data. Eur Respir J 2015 Apr 10;45(4):994-1018 [FREE Full text] [doi: 10.1183/09031936.00172114] [Medline: 25504997]
6. Chen S, Golam S, Myers J, Bly C, Smolen H, Xu X. Systematic literature review of the clinical, humanistic, and economic burden associated with asthma uncontrolled by GINA Steps 4 or 5 treatment. Curr Med Res Opin 2018 Dec 16;34(12):2075-2088. [doi: 10.1080/03007995.2018.1505352] [Medline: 30047292]
7. Fuhrman C, Dubus J, Marguet C, Delacourt C, Thumerelle C, de Blic J, et al. Hospitalizations for asthma in children are linked to undertreatment and insufficient asthma education. J Asthma 2011 Aug 20;48(6):565-571. [doi: 10.3109/02770903.2011.580031] [Medline: 21595608]
8. Suissa S, Ernst P, Kezouh A. Regular use of inhaled corticosteroids and the long term prevention of hospitalisation for asthma. Thorax 2002 Oct;57(10):880-884 [FREE Full text] [doi: 10.1136/thorax.57.10.880] [Medline: 12324675]
9. Krishnan JA, Riekert KA, McCoy JV, Stewart DY, Schmidt S, Chanmugam A, et al. Corticosteroid Use after Hospital Discharge among High-risk Adults with Asthma. Am J Respir Crit Care Med 2004 Dec 15;170(12):1281-1285. [doi: 10.1164/rccm.200403-409oc]
10. Williams LK, Peterson EL, Wells K, Ahmedani BK, Kumar R, Burchard EG, et al. Quantifying the proportion of severe asthma exacerbations attributable to inhaled corticosteroid nonadherence. J Allergy Clin Immunol 2011 Dec;128(6):1185-1191.e2 [FREE Full text] [doi: 10.1016/j.jaci.2011.09.011] [Medline: 22019090]
11. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S149-S167. [doi: 10.1016/j.respe.2017.05.004] [Medline: 28756037]
12. Laforest L, Licaj I, Devouassoux G, Eriksson I, Caillet P, Chatte G, et al. Prescribed therapy for asthma: therapeutic ratios and outcomes. BMC Fam Pract 2015 Apr 14;16(1):49 [FREE Full text] [doi: 10.1186/s12875-015-0265-2] [Medline: 25880641]
13. Laforest L, Licaj I, Devouassoux G, Chatte G, Martin J, Van Ganse E. Asthma drug ratios and exacerbations: claims data from universal health coverage systems. Eur Respir J 2014 May 30;43(5):1378-1386 [FREE Full text] [doi: 10.1183/09031936.00100113] [Medline: 24789951]
14. Schatz M, Nakahiro R, Crawford W, Mendoza G, Mosen D, Stibolt TB. Asthma quality-of-care markers using administrative data. Chest 2005 Oct;128(4):1968-1973. [doi: 10.1378/chest.128.4.1968] [Medline: 16236843]
15. Yong PL, Werner RM. Process quality measures and asthma exacerbations in the medicaid population. J Allergy Clin Immunol 2009 Nov;124(5):961-966. [doi: 10.1016/j.jaci.2009.07.027] [Medline: 19748660]
16. Lam WY, Fresco P. Medication Adherence Measures: An Overview. Biomed Res Int 2015;2015:217047-217012 [FREE Full text] [doi: 10.1155/2015/217047] [Medline: 26539470]
17. Vollmer WM, Xu M, Feldstein A, Smith D, Waterbury A, Rand C. Comparison of pharmacy-based measures of medication adherence. BMC Health Serv Res 2012 Jun 12;12(1):155 [FREE Full text] [doi: 10.1186/1472-6963-12-155] [Medline: 22691240]
18. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther 2002 Aug;27(4):299-309. [doi: 10.1046/j.1365-2710.2002.00430.x] [Medline: 12174032]
19. Andruff H, Carraro N, Thompson A, Gaudreau P, Louvet B. Latent Class Growth Modelling: A Tutorial. TQMP 2009 Mar 01;5(1):11-24. [doi: 10.20982/tqmp.05.1.p011]
20. Jones BL, Nagin DS, Roeder K. A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. Sociological Methods & Research 2016 Jun 30;29(3):374-393. [doi: 10.1177/0049124101029003005]
21. Belhassen M, De Blic J, Laforest L, Laigle V, Chanut-Vogel C, Lamezec L, et al. Recurrent Wheezing in Infants: A Population-Based Study. Medicine (Baltimore) 2016 Apr;95(15):e3404 [FREE Full text] [doi: 10.1097/MD.0000000000003404] [Medline: 27082618]

22. Heatley H, Tran TN, Bourdin A, Menzies-Gow A, Jackson DJ, Maslova E, et al. Observational UK cohort study to describe intermittent oral corticosteroid prescribing patterns and their association with adverse outcomes in asthma. Thorax 2023 Sep 27;78(9):860-867 [FREE Full text] [doi: 10.1136/thorax-2022-219642] [Medline: 36575040]
23. Belhassen M, Nibber A, Van Ganse E, Ryan D, Langlois C, Appiagyei F, et al. Inappropriate asthma therapy-a tale of two countries: a parallel population-based cohort study. NPJ Prim Care Respir Med 2016 Oct 13;26(1):16076 [FREE Full text] [doi: 10.1038/npjpcrm.2016.76] [Medline: 27735927]
24. Van Ganse E, Texier N, Dima AL, Belhassen M, Laforest L, Herbage S, ASTRO-LAB group. Effects of short- and long-acting beta-agonists on asthma exacerbations: a prospective cohort. Ann Allergy Asthma Immunol 2020 Mar;124(3):254-260. [doi: 10.1016/j.anai.2019.12.012] [Medline: 31862434]
25. Kaufmann J, Marino M, Lucas J, Bailey SR, Giebultowicz S, Puro J, et al. Racial and Ethnic Disparities in Acute Care Use for Pediatric Asthma. Ann Fam Med 2022 Mar 28;20(2):116-122 [FREE Full text] [doi: 10.1370/afm.2771] [Medline: 35346926]
26. Belhassen M, Langlois C, Laforest L, Dima AL, Ginoux M, Sadatsafavi M, et al. Level of Asthma Controller Therapy Before Admission to the Hospital. J Allergy Clin Immunol Pract 2016 Sep;4(5):877-883 [FREE Full text] [doi: 10.1016/j.jaip.2016.06.012] [Medline: 27452886]
27. Kenyon CC, Gruschow SM, Quarshie WO, Griffis H, Leach MC, Zorc JJ, et al. Controller adherence following hospital discharge in high risk children: A pilot randomized trial of text message reminders. J Asthma 2019 Jan 13;56(1):95-103 [FREE Full text] [doi: 10.1080/02770903.2018.1424195] [Medline: 29437489]
28. Staehr Holm F, Håkansson K, Ulrik C. Adherence with controller medication in adults with asthma - impact of hospital admission for acute exacerbation. J Asthma 2022 Sep 02;59(9):1899-1907. [doi: 10.1080/02770903.2021.1971702] [Medline: 34425724]
29. Laforest L, Van Ganse E, Devouassoux G, Chretin S, Osman L, Bauguil G, et al. Management of asthma in patients supervised by primary care physicians or by specialists. Eur Respir J 2006 Jan 01;27(1):42-50 [FREE Full text] [doi: 10.1183/09031936.06.00035805] [Medline: 16387934]
30. Pearce CJ, Fleming L. Adherence to medication in children and adolescents with asthma: methods for monitoring and intervention. Expert Review of Clinical Immunology 2018 Oct 15;14(12):1055-1063. [doi: 10.1080/1744666x.2018.1532290]
31. Normansell R, Kew K, Stovold E. Interventions to improve adherence to inhaled steroids for asthma. Cochrane Database Syst Rev 2017 Apr 18;4(4):CD012226 [FREE Full text] [doi: 10.1002/14651858.CD012226.pub2] [Medline: 28417456]
32. Gibson P, Powell H, Coughlan J, Wilson A, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev 2003(1):CD001117. [doi: 10.1002/14651858.CD001117] [Medline: 12535399]
33. Gibson P, Powell H, Coughlan J, Wilson A, Hensley M, Abramson M, et al. Limited (information only) patient education programs for adults with asthma. Cochrane Database Syst Rev 2002(2):CD001005. [doi: 10.1002/14651858.CD001005] [Medline: 12076400]
34. Chapman KR, Barnes NC, Greening AP, Jones PW, Pedersen S. Single maintenance and reliever therapy (SMART) of asthma: a critical appraisal. Thorax 2010 Aug 27;65(8):747-752 [FREE Full text] [doi: 10.1136/thx.2009.128504] [Medline: 20581409]
35. Belhassen M, Nolin M, Nibber A, Ginoux M, Devouassoux G, Van Ganse E. Changes in Persistent Asthma Care and Outcomes From 2006 to 2016 in France. J Allergy Clin Immunol Pract 2019 Jul;7(6):1858-1867 [FREE Full text] [doi: 10.1016/j.jaip.2019.02.025] [Medline: 30836232]

Abbreviations

ARH: asthma-related hospitalization
CMA7: Continuous Measures of Medication Acquisition-7
EGB: Echantillon Généraliste des Bénéficiaires (General Sample of Beneficiaries)
ER: emergency room
FDC: fixed-dose combination
ICD-10: International Classification of Diseases, 10th Revision
ICS: inhaled corticosteroid
LABA: long-acting beta agonist
LTRA: leukotriene receptor antagonist
MAD: mean annual dose
OCS: oral corticosteroid
SABA: short-acting β2 agonist

Edited by A Mavragani, T Sanchez; submitted 19.06.23; peer-reviewed by UK Chalwadi, L Elsey; comments to author 17.07.23; revised version received 19.07.23; accepted 24.08.23; published 26.09.23.

Please cite as:
Belhassen M, Nolin M, Jacoud F, Marant Micallef C, Van Ganse E
Trajectories of Controller Therapy Use Before and After Asthma-Related Hospitalization in Children and Adults: Population-Based Retrospective Cohort Study
JMIR Public Health Surveill 2023;9:e50085
URL: https://publichealth.jmir.org/2023/1/e50085
doi:10.2196/50085
PMID:37751244

Original Paper

Testing for Hepatitis C During Pregnancy Among Persons With Medicaid and Commercial Insurance: Cohort Study

Mohammed A Khan[1], MSPH, PhD; William W Thompson[1], PhD; Ademola Osinubi[1], MSc; William A Meyer 3rd[2], PhD; Harvey W Kaufman[2], MBA, MD; Paige A Armstrong[1], MHS, MD; Monique A Foster[1], MPH, MD; Noele P Nelson[1], MPH, MD, PhD; Carolyn Wester[1], MPH, MD

[1]Division of Viral Hepaitits, Centers for Disease Control and Prevention, Atlanta, GA, United States

[2]Quest Diagnostics, Secaucus, NJ, United States

Corresponding Author:
Mohammed A Khan, MSPH, PhD
Division of Viral Hepaitits
Centers for Disease Control and Prevention
1600 Clifton Rd NE
Atlanta, GA, 30329
United States
Phone: 1 404 718 5807
Email: odt5@cdc.gov

Abstract

Background: The reported incidence of acute hepatitis C virus (HCV) infection is increasing among persons of childbearing age in the United States. Infants born to pregnant persons with HCV infection are at risk for perinatal HCV acquisition. In 2020, the United States Preventive Services Task Force and Centers for Disease Control and Prevention recommended that all pregnant persons be screened during each pregnancy for hepatitis C. However, there are limited data on trends in hepatitis C testing during pregnancy.

Objective: We estimated hepatitis C testing rates in a large cohort of patients with Medicaid and commercial insurance who gave birth during 2015-2019 and described demographic and risk-based factors associated with testing.

Methods: Medicaid and commercial insurance claims for patients aged 15-44 years and who gave birth between 2015 and 2019 were included. Birth claims were identified using procedure and diagnosis codes for vaginal or cesarean delivery. Hepatitis C testing was defined as an insurance claim during the 42 weeks before delivery. Testing rates were calculated among patients who delivered and among the subset of patients who were continuously enrolled for 42 weeks before delivery. We also compared the timing of testing relative to delivery among patients with commercial or Medicaid insurance. Multivariable logistic regression was used to identify factors associated with testing.

Results: Among 1,142,770 Medicaid patients and 1,207,132 commercially insured patients, 175,223 (15.3%) and 221,436 (18.3%) were tested for hepatitis C during pregnancy, respectively. Testing rates were 89,730 (21.8%) and 187,819 (21.9%) among continuously enrolled Medicaid and commercially insured patients, respectively. Rates increased from 2015 through 2019 among Medicaid (from 20,758/108,332, 19.2% to 13,971/52,330, 26.8%) and commercially insured patients (from 38,308/211,555, 18.1% to 39,152/139,972, 28%), respectively. Among Medicaid patients, non-Hispanic Black (odds ratio 0.73, 95% CI 0.71-0.74) and Hispanic (odds ratio 0.53, 95% CI 0.51-0.56) race or ethnicity were associated with lower odds of testing. Opioid use disorder, HIV infection, and high-risk pregnancy were associated with higher odds of testing in both Medicaid and commercially insured patients.

Conclusions: Hepatitis C testing during pregnancy increased from 2015 through 2019 among patients with Medicaid and commercial insurance, although tremendous opportunity for improvement remains. Interventions to increase testing among pregnant persons are needed.

(JMIR Public Health Surveill 2023;9:e40783) doi:10.2196/40783

KEYWORDS

hepatitis C; testing; pregnancy; pregnant; trend; insurance; insured; HCV; hepatitis; Medicaid; health coverage; maternal; fetus; birth; natal; maternity; liver; communicable disease; viral infection

Introduction

Chronic hepatitis C virus (HCV) infection is a leading cause of liver disease and liver-related mortality in the United States. Approximately 2.4 million persons in the United States were living with chronic hepatitis C during 2013-2016 [1,2]. These patients are at increased risk of progression to liver fibrosis, cirrhosis, liver failure, hepatocellular carcinoma, and death [3]. HCV can be transmitted during pregnancy or delivery—an estimated 5.8% of infants born to women with HCV infection become infected [4]. Perinatal infection can result in liver disease during adulthood and sequelae such as cirrhosis and hepatocellular carcinoma [5]. The incidence of acute HCV infection in the United States increased annually from 2009 to 2019, with rates highest among adults 20-29 and 30-39 years of age, including women of childbearing age [6,7]. These increases are concurrent with increases in the nation's opioid crisis [8]. Recent studies have documented an increasing prevalence of HCV infection among pregnant persons [9]. In 2018, 0.5% of all live births nationally were among mothers with hepatitis C [10]. The rate of maternal HCV infection as recorded on US birth certificates increased by 89% between 2009 and 2014, and rates were highest in Appalachia [11]. Hepatitis C antibody testing among pregnant persons from a large national commercial laboratory increased from 5.7% to 13.4% and test positivity increased from 2.6% to 3.6% from 2011 to 2016, respectively [12].

While chronic HCV infection can be cured with direct-acting antiviral therapy, direct-acting antiviral therapies have not been approved for use during pregnancy by the US Food and Drug Administration [13]. However, identifying hepatitis C during pregnancy offers several advantages, including referral for care and treatment for the pregnant person and referral for testing for the exposed infant and can inform clinical decision-making during prenatal and intrapartum care [14]. Since 2012 and 2013, the Centers for Disease Control and Prevention (CDC) and the United States Preventive Services Task Force (USPSTF), respectively, have recommended 1-time hepatitis C testing for all persons born from 1945 through 1965 and testing adults with risk factors for HCV infection [15]. In 2018, the American Association for the Study of Liver Diseases and the Infectious Diseases Society of America recommended hepatitis C screening during each pregnancy [13]. In 2020, the USPSTF and CDC recommended that all pregnant persons be screened during each pregnancy for hepatitis C [14]. However, there are limited data on trends in hepatitis C testing during pregnancy. We estimated hepatitis C testing rates in a large cohort of patients with Medicaid and commercial insurance who gave birth during 2015-2019 and described demographic and risk-based factors associated with testing.

Methods

Overview

Data for persons with commercial, Medicare Supplemental, and Medicaid insurance were obtained from the Merative MarketScan Research Databases. These data include demographics, insurance enrollment information, inpatient and outpatient, and prescription drug claims. The commercial and Medicare Supplemental databases include US persons with employer-sponsored health insurance coverage and their dependents. The MarketScan Multi-State Medicaid database includes claims for patients enrolled in Medicaid fee-for-service and managed care plans in multiple states; the number and names of states are proprietary and were not disclosed. Data from January 1, 2014, to December 31, 2019, were used to allow for a 42-week baseline enrollment period prior to delivery between 2015 and 2019.

Cohort Selection

We constructed a cohort of female patients aged 15-44 years who gave birth from January 1, 2015, to December 31, 2019. Birth claims were identified using Current Procedural Terminology (CPT) and Diagnosis-Related Group codes specific to vaginal or cesarean delivery (Table S1 in Multimedia Appendix 1). Only the earliest claim for delivery during the study period was retained; the service date of this claim was the patient's index delivery date. As hepatitis C testing during every pregnancy was not recommended during the study period, patients whose HCV infection was identified in the first pregnancy may not have been tested in subsequent pregnancies. Thus, persons with multiple births during the study period were counted once at their earliest delivery.

Measures

Claims for hepatitis C testing were identified using CPT codes for any anti-HCV antibody, RNA, and genotyping tests. The anti-HCV antibody test indicates past or current HCV infection, while HCV RNA or genotyping test identifies or characterizes current HCV infection. Screening with the HCV antibody test is recommended [14]. Hepatitis C testing during pregnancy was defined as the earliest hepatitis C testing claim in the 42 weeks prior to the index delivery date. Demographic characteristics included age, race or ethnicity (Medicaid only), and region of residence (commercial only). Clinical characteristics (alcohol use disorder, opioid use disorder, severe mental illness, obesity, HIV infection, gestational diabetes, preeclampsia, preterm labor, and singleton or multiple pregnancy) were identified using the International Classification of Diseases, Ninth/Tenth Revision, Clinical Modification (ICD-9/10-CM) codes associated with a claim in the 42 weeks prior to the index delivery date. Additional prenatal care characteristics included an obstetric test panel (including complete blood count, hepatitis B surface antigen, rubella antibody, nontreponemal syphilis test, ABO blood group and Rh typing, and red blood cell antibody screen) claim and influenza and tetanus-diphtheria-acellular pertussis (Tdap) vaccination during the 42 weeks before the index delivery date and were defined using CPT codes (Table S1 in Multimedia Appendix 1).

Statistical Analysis

Due to incomplete follow-up during the 42 weeks before the index delivery date, we estimated the rate of hepatitis C testing during pregnancy using two methods: (1) percent tested among all patients who gave birth during the study period regardless of baseline enrollment criteria and (2) percent tested among patients continuously enrolled for 42-week period prior to the index delivery date. In sensitivity analyses, we explored the impact of (1) changing the continuous enrollment criteria (12, 24, and 36 weeks, and median time to hepatitis C test), (2) restricting the sample to patients who had an obstetric panel test claim in the 42 weeks before the index delivery date, and (3) restricting the sample to patients who had any inpatient or outpatient service claim in the 42 weeks prior to the index delivery date on the estimated testing rate. The cumulative incidence of testing among patients continuously enrolled for at least 7 days before the index delivery date was also estimated using the Kaplan-Meier method. The Kaplan-Meier method can be used to estimate testing incidence in the presence of censoring (ie, incomplete follow-up). We included a 7-day enrollment minimum to exclude outliers and ensure some demographic and clinical data would be available.

We estimated the testing rate by year of delivery for each cohort and tested for statistically significant trends using the Cochran-Armitage test or for cumulative incidence estimates, and the Gray test for equality of cumulative incidence functions. Additionally, we examined the distribution of time between hepatitis C testing and index delivery date by year of delivery and insurance type. Demographic and clinical characteristics for cohorts constructed using both methods (no enrollment constraint and continuously enrolled for 42 weeks) were described by insurance type and hepatitis C testing status. Multivariable adjusted associations between demographic and clinical characteristics were estimated using logistic regression with the outcome defined as a claim for hepatitis C testing during the 42-week period prior to the index delivery date. We included all variables in multivariable models because our analysis was primarily exploratory and causal relationships between these variables and hepatitis C testing are poorly understood. Odds ratios and 95% CIs are provided. Adjusted marginal probabilities were also estimated. In a sensitivity analysis, logistic regression was performed for patients who had an obstetric panel test insurance claim. Analyses were performed in SAS (version 9.4; SAS Institute).

Ethical Considerations

This analysis of deidentified data did not require institutional review board approval.

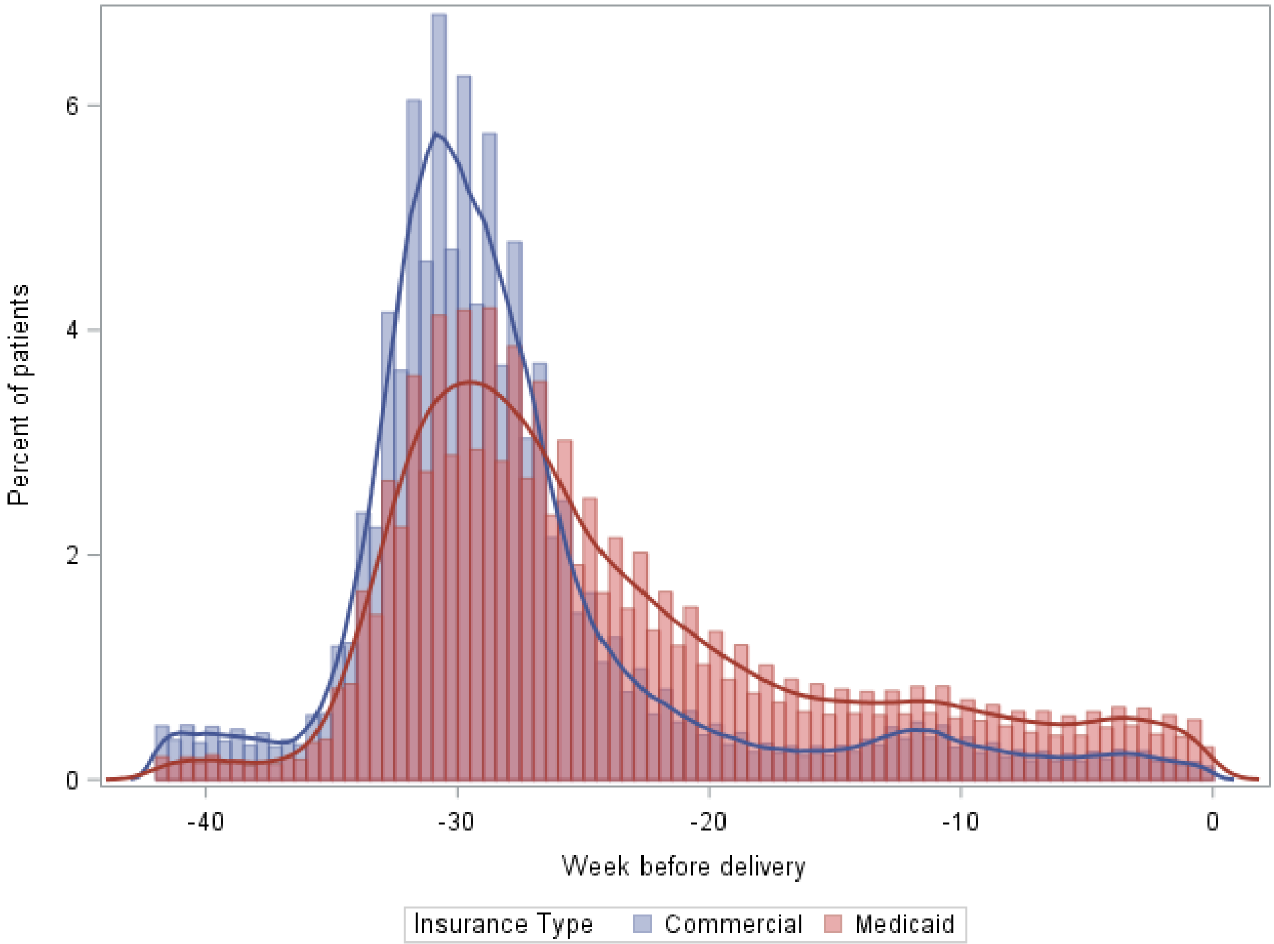
Results

Overview

We identified 1,152,283 patients with Medicaid and 1,217,686 patients with commercial insurance who had delivery claims between 2015 and 2019. Among these persons, 1,142,770 (99.2%) Medicaid patients and 1,207,132 (99.1%) commercially insured patients had valid demographic and enrollment information. When restricting the analyses to patients who were continuously insurance enrolled for 42 weeks prior to the index delivery date, 411,795 (36%) Medicaid patients and 857,618 (71%) commercially insured patients remained. Among all Medicaid patients, 175,223 (15.3%) were tested for hepatitis C in the 42 weeks prior to the index delivery date; among the continuously enrolled subset, 89,730 (21.8%) were tested. Among all commercially insured patients, 221,436 (18.3%) were tested for hepatitis C before the index delivery date, and among the continuously enrolled subset, 187,819 (21.9%) were tested.

A majority of tested Medicaid (162,868/175,223, 93%) and commercially insured (217,969/221,436, 98.5%) patients were tested for HCV antibody only. The median time between hepatitis C testing and index delivery date for Medicaid patients was 186 (IQR 137-211) days (Figure 1). For commercial insurance patients, the median time between hepatitis C testing and index delivery date was 206 (IQR 184-222) days, suggesting they were tested on average 20 days earlier than Medicaid patients.

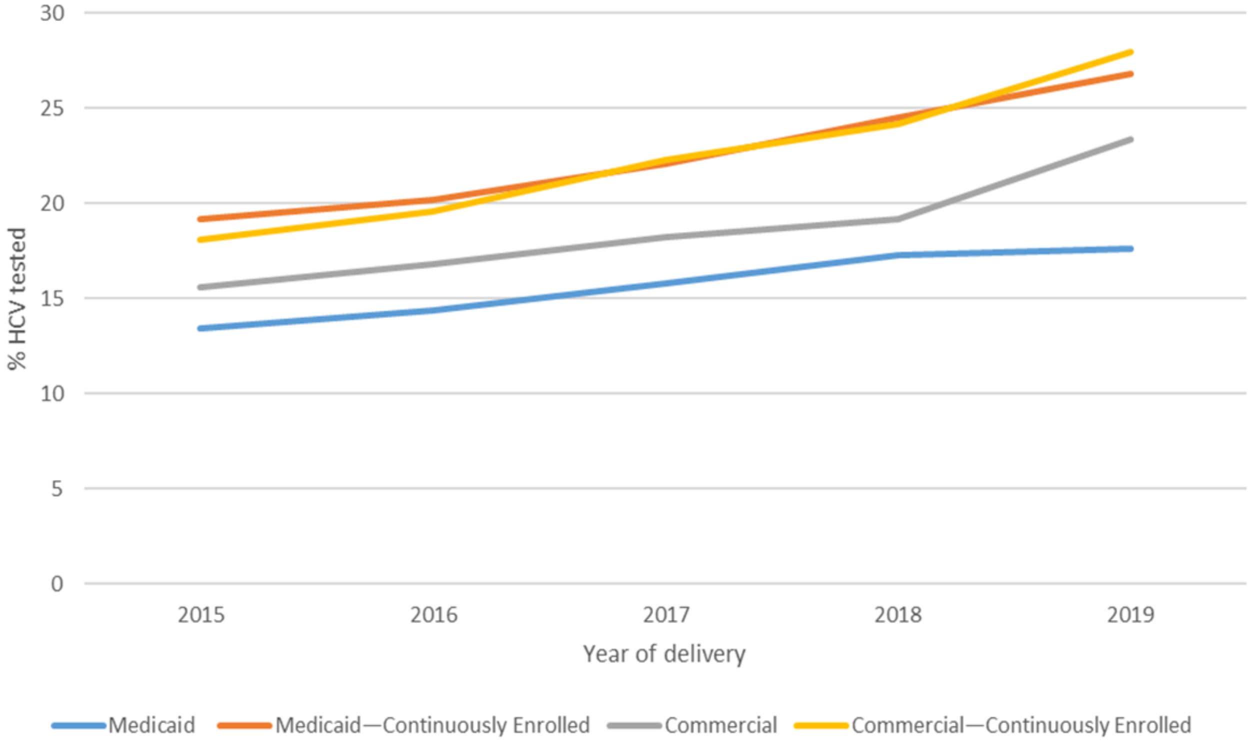
Figure 1. Distribution of weeks between hepatitis C test and delivery date, MarketScan Medicaid and Commercial.



Trends in Hepatitis C Testing

The hepatitis C testing rate increased from 2015 to 2019 for both Medicaid and commercially insured patients (Figure 2). Among Medicaid patients, testing rates increased from 13.4% (40,351/301,849) in 2015 to 17.6% (31,123/177,004) in 2019, and among continuously enrolled persons, testing rates increased from 19.2% (20,758/108,332) in 2015 to 26.8% (13,971/52,230) in 2019. Among commercially insured patients, the testing rates increased from 15.6% (43,941/281,007) in 2015 to 23.4% (45,930/296,344) in 2019, and among continuously enrolled persons, the testing rates increased from 18.1% (38,308/211,555) in 2015 to 28% (39,152/139,972) in 2019. Trends in testing were similar in the sensitivity analysis using different continuous enrollment thresholds and testing rate estimation methods: testing rates increased between 2015 and 2019 across all methods. Among patients with Medicaid insurance, the average annual change in testing rate varied from 0.8% for the cohort with no enrollment constraints to 1.7% for the cohorts with an obstetric panel claim. Among patients with commercial insurance, the average annual change in testing rate varied from 1.6% for the cohort with no enrollment constraints to 2% for the Kaplan-Meier estimate (Table S3 in Multimedia Appendix 1).

Figure 2. Hepatitis C testing by year of delivery and insurance type. HCV: hepatitis C virus.



Characteristics of Continuously Enrolled Cohorts

Demographic characteristics of patients by insurance type and hepatitis C testing status for the continuously enrolled cohorts are included in Table 1. Patients with Medicaid insurance in the continuously enrolled cohort were predominantly aged 19-29 years (265,421/411,795, 64.5%) followed by those aged 30-39 years (107,665/411,795, 26.2%). The distribution of race and ethnicity was as follows: 48% (196,441/411,795) of Medicaid patients were non-Hispanic White, followed by non-Hispanic Black (152,700/411,795, 37.1%), other race and ethnicity (44,891/411,795, 10.9%), Hispanic, all races (15,347/411,795, 3.7%), and unknown race or ethnicity (2416/411,795, 0.6%). High-risk pregnancy (202,584/411,795, 49.2%), obesity (78,781/411,795, 19.1%), preterm labor (63,219/411,795, 15.4%), severe mental illness (52,133/411,795, 12.7%), and gestational diabetes (35,249/411,795, 8.6%) were the most frequently occurring underlying conditions (Table 2). Forty-three percent (175,695/411,795) of patients received an obstetric panel test; 28.7% (117,969/411,795) and 15.7% (64,433/411,795) received Tdap and influenza vaccinations, respectively.

On average, commercially insured patients in the continuously enrolled cohort were older than the Medicaid patients. Among patients with commercial insurance, those aged 30-39 years (494,564/857,618, 57.7%) represented the largest age group, which is followed by those aged 19-29 years (315,311/857,618, 36.8%). In terms of geographic distribution, patients resided in the south (387,766/857,618, 45.2%), north central (177,013/857,618, 20.6%), west (148,048/857,618, 17.3%), and northeast (142,531/857,618, 16.6%). The distribution of underlying conditions also differed for the commercial sample relative to those with Medicaid: high-risk pregnancy (361,659/857,618, 42.2%), obesity (103,780/857,618, 12.1%), gestational diabetes (90,884/857,618, 10.6%), and preterm labor (73,004/857,618, 8.5%) were the most frequently occurring underlying conditions. Opioid use disorder (2443/857,618, 0.3%) and alcohol use disorder (1656/857,618, 0.2%) were less common than among Medicaid patients (14,069/411,795, 3.4% and 4973/411,795 1.2%, respectively). Claims for an obstetric panel test (470,404/857,618, 54.9%), Tdap vaccination (455,485/857,618, 53.1%), and influenza vaccination (258,427/857,618, 29.1%) were more common among commercially insured patients.

Table 1. Demographic and clinical characteristics by hepatitis C virus testing status for patients who were continuously enrolled for 42 weeks, MarketScan commercial and Medicaid, 2015-2019.

	Medicaid			Commercial		
	Total (N=411,795)	Tested (n=89,730)	Not tested (n=322,065)	Total (N=857,618)	Tested (n=187,819)	Not Tested (n=669,799)
Age group (years), n (%)						
15-18	31,636 (7.7)	6578 (7.5)	24,878 (7.7)	6035 (0.7)	1308 (0.7)	4727 (0.7)
19-29	265,421 (64.5)	57,570 (64.2)	207,851 (65.5)	315,311 (36.8)	67,827 (36.1)	247,484 (37)
30-39	107,665 (26.2)	23,934 (26.7)	83,731 (26)	494,564 (57.7)	108,716 (57.9)	385,848 (57.6)
40-44	7073 (1.7)	1468 (1.6)	5605 (1.7)	41,708 (4.9)	9968 (5.3)	31,740 (4.7)
Race and ethnicity, n (%)						
Non-Hispanic Black	152,700 (37.1)	27,828 (31)	124,872 (38.8)	N/A[a]	N/A	N/A
Hispanic	15,347 (3.7)	2056 (2.3)	13,291 (4.1)	N/A	N/A	N/A
Non-Hispanic White	196,441 (47.7)	49,309 (55)	147,132 (45.7)	N/A	N/A	N/A
Other	44,891 (10.9)	9900 (11)	34,991 (10.9)	N/A	N/A	N/A
Unknown	2416 (0.6)	637 (0.7)	1779 (0.6)	N/A	N/A	N/A
Region, n (%)						
Northeast	N/A	N/A	N/A	142,531 (16.6)	41,447 (22.1)	101,084 (15.1)
North central	N/A	N/A	N/A	177,013 (20.6)	24,525 (13.1)	152,488 (22.8)
South	N/A	N/A	N/A	387,766 (45.2)	96,721 (51.5)	291,045 (43.5)
West	N/A	N/A	N/A	148,048 (17.3)	24,586 (13.1)	123,462 (18.4)
Unknown	N/A	N/A	N/A	2260 (0.3)	540 (0.3)	1720 (0.3)

[a] N/A: not available.

Table 2. Clinical characteristics by hepatitis C virus testing status for patients who were continuously enrolled for 42 weeks, MarketScan Commercial and Medicaid, 2015-2019.

	Medicaid			Commercial		
	Total (N=411,795), n (%)	Tested (n=89,730), n (%)	Not tested (n=322,065), n (%)	Total (N=857,618), n (%)	Tested (n=187,819), n (%)	Not tested (n=669,799), n (%)
Alcohol use disorder	4973 (1.2)	1854 (2.1)	3119 (1)	1656 (0.2)	524 (0.3)	1132 (0.2)
Opioid use disorder	14,069 (3.4)	7767 (8.9)	6302 (2)	2443 (0.3)	1021 (0.5)	1422 (0.2)
Severe mental illness	52,133 (12.7)	15,181 (16.9)	36,952 (11.5)	37,393 (4.4)	9157 (4.9)	28,236 (4.2)
Obesity	78,781 (19.1)	17,736 (19.8)	61,045 (19)	103,780 (12.1)	26,224 (14)	77,556 (11.6)
HIV infection	879 (0.2)	525 (0.6)	354 (0.1)	784 (0.1)	320 (0.2)	464 (0.1)
Preeclampsia	9467 (2.3)	2280 (2.5)	7187 (2.2)	22,019 (2.6)	5528 (2.9)	16,491 (2.5)
High-risk pregnancy	202,584 (49.2)	51,588 (57.5)	150,996 (46.9)	361,659 (42.2)	87,263 (46.5)	274,396 (41)
Gestational diabetes	35,249 (8.6)	7888 (8.8)	27,361 (8.5)	90,884 (10.6)	21,094 (11.2)	69,790 (10.4)
Preterm labor	63,219 (15.4)	14,286 (15.9)	48,933 (15.2)	73,004 (8.5)	16,759 (8.9)	56,245 (8.4)
Multiple gestation	9404 (2.3)	1992 (2.2)	7412 (2.3)	24,791 (2.9)	6106 (3.3)	18,685 (2.8)
Obstetric panel	175,695 (42.7)	41,930 (46.7)	133,765 (41.5)	470,404 (54.9)	109,830 (58.5)	360,574 (53.8)
Tdap[a] vaccination	117,969 (28.7)	24,452 (27.3)	93,517 (29)	455,485 (53.1)	95,270 (50.7)	360,215 (53.8)
Influenza vaccination	64,433 (15.7)	13,784 (15.4)	50,649 (15.7)	258,427 (30.1)	54,608 (29.1)	203,819 (30.4)

[a] Tdap: tetanus-diphtheria-acellular pertussis.

Factors Associated With Testing Among Continuously Enrolled Patients

Hepatitis C testing status varied by patient characteristics. Adjusted associations between patient characteristics and hepatitis C testing status are presented in Table 3. Among the continuously enrolled Medicaid patients, patients aged 40-44 years (odds ratio [OR] 0.85, 95% CI 0.80-0.90) were less likely to be tested than younger patients. Relative to non-Hispanic White patients, Hispanic patients (OR 0.53, 95% CI 0.51-0.56), non-Hispanic Black (OR 0.73, 95% CI 0.71-0.74), and other race (OR 0.90, 95% CI 0.88-0.92) were less likely to be tested among Medicaid patients.

Patients with opioid use disorder (OR 3.75, 95% CI 3.62-3.89) or HIV infection (OR 5.40, 95% CI 4.71-6.20) claims were significantly more likely to be tested than those without these conditions. Patients with severe mental illness claims or alcohol use disorder were 1.28 (95% CI 1.25-1.30) and 1.29 (95% CI 1.21-1.37) times more likely to be tested than those without these conditions. High-risk pregnancy (OR 1.46, 95% CI 1.44-1.48) was associated with higher odds of testing, whereas multiple gestations (OR 0.91, 95% CI 0.86-0.96) was associated with lower odds of testing compared to patients without these characteristics.

Among patients with commercial insurance, patients living in the north central (OR 0.40, 95% CI 0.39-0.40), west (OR 0.49, 95% CI 0.48-0.50), and south (OR 0.79, 95% CI 0.78-0.80) were less likely to be tested compared to patients residing in the northeast region. Patients with opioid use disorder (OR 2.33, 95% CI 2.15-2.53), HIV infection (OR 1.96, 95% CI 1.70-2.27), and alcohol use disorder (OR 1.36, 95% CI 1.22-1.51) were more likely to be tested than patients without these conditions. High-risk pregnancy (OR 1.23, 95% CI 1.22-1.25), obesity (OR 1.15, 95% CI 1.13-1.17), and multiple gestations (OR 1.11, 95% CI 1.08-1.14) were associated with higher odds of testing compared to patients without these characteristics. Among the cohort of commercially insured patients with no enrollment constraint, effect estimates were higher relative to those with the constraint but overall consistent in the direction of effects. Adjusted marginal probabilities of hepatitis C testing are shown in Table S6 in Multimedia Appendix 1.

Table 3. Multivariable-adjusted associations[a] between demographic and clinical characteristics and hepatitis C testing during pregnancy.

	Medicaid		Commercial	
	No enrollment constraint, OR[b] (95% CI)	Continuously enrolled for 42 weeks, OR (95% CI)	No enrollment constraint, OR (95% CI)	Continuously enrolled for 42 weeks, OR (95% CI)
Age group (years)				
15-18	1.33 (1.30-1.37)	1.11 (1.07-1.14)	1.18 (1.11-1.25)	1.06 (1.00-1.13)
19-29	1.19 (1.17-1.20)	1.08 (1.06-1.10)	1.06 (1.05-1.07)	1.05 (1.04-1.06)
30-39	Ref[c]	Ref	Ref	Ref
40-44	0.77 (0.74-0.80)	0.85 (0.80-0.90)	0.97 (0.95-0.99)	0.99 (0.97-1.02)
Race or ethnicity				
Non-Hispanic Black	0.83 (0.82-0.84)	0.73 (0.71-0.74)	N/A[d]	N/A
Hispanic	0.43 (0.41-0.44)	0.53 (0.51-0.56)	N/A	N/A
Non-Hispanic White	Ref	Ref	N/A	N/A
Other	0.88 (0.87-0.90)	0.90 (0.88-0.92)	N/A	N/A
Unknown	0.87 (0.84-0.91)	1.09 (0.99-1.19)	N/A	N/A
Region				
Northeast	N/A	N/A	Ref	Ref
North central	N/A	N/A	0.39 (0.39-0.40)	0.40 (0.39-0.40)
South	N/A	N/A	0.75 (0.74-0.76)	0.79 (0.78-0.80)
West	N/A	N/A	0.48 (0.47-0.49)	0.49 (0.48-0.50)
Unknown	N/A	N/A	0.73 (0.66-0.79)	0.77 (0.70-0.85)
Clinical characteristics				
Alcohol use disorder	1.44 (1.37-1.52)	1.29 (1.21-1.37)	1.45 (1.31-1.60)	1.36 (1.22-1.51)
Opioid use disorder	4.46 (4.34-4.58)	3.75 (3.62-3.89)	2.54 (2.35-2.73)	2.33 (2.15-2.53)
Severe mental illness	1.48 (1.45-1.50)	1.28 (1.25-1.30)	1.24 (1.22-1.27)	1.15 (1.12-1.17)
Obesity	1.17 (1.16-1.19)	1.03 (1.01-1.05)	1.21 (1.19-1.22)	1.15 (1.13-1.17)
HIV infection	5.57 (5.05-6.14)	5.40 (4.71-6.20)	2.27 (2.00-2.59)	1.96 (1.70-2.27)
Preeclampsia	1.06 (1.03-1.10)	1.07 (1.01-1.12)	1.10 (1.07-1.14)	1.09 (1.06-1.13)
High-risk pregnancy	1.83 (1.81-1.85)	1.46 (1.44-1.48)	1.34 (1.33-1.35)	1.23 (1.22-1.25)
Gestational diabetes	1.00 (0.98-1.02)	1.00 (0.97-1.02)	1.04 (1.02-1.06)	1.02 (1.01-1.04)
Preterm labor	1.08 (1.07-1.10)	1.02 (0.99-1.04)	1.05 (1.02-1.06)	1.02 (1.00-1.04)
Multiple gestations	0.91 (0.88-0.94)	0.91 (0.86-0.96)	1.13 (1.10-1.16)	1.11 (1.08-1.14)
Obstetric panel	1.94 (1.92-1.96)	1.27 (1.25-1.29)	1.80 (1.78-1.82)	1.17 (1.16-1.19)
Tdap[e] vaccination	0.99 (0.98-1.00)	0.90 (0.88-0.91)	1.03 (1.02-1.04)	0.96 (0.95-0.97)
Influenza vaccination	1.07 (1.06-1.09)	0.99 (0.97-1.02)	1.10 (1.09-1.12)	0.96 (0.95-0.97)

[a] Models included all listed variables.

[b] OR: odds ratio.

[c] Ref: reference.

[d] N/A: not available.

[e] Tdap: tetanus-diphtheria-acellular pertussis.

Sensitivity Analyses

In the sensitivity analysis of testing rate by inclusion criteria and estimation method, the estimated testing rates increased as the length of the continuous enrollment period increased. The Kaplan-Meier method yielded higher estimates relative to naïve estimates. In addition, patients with an obstetric panel test had higher HCV testing rates relative to those who did not. Overall, differences in estimates across methods and inclusion criteria were larger for the Medicaid patients relative to the commercially insured patients (Table S2 in Multimedia Appendix 1).

In the sensitivity analyses of adjusted associations for HCV testing that included demographic and clinical characteristics, estimates for the Medicaid sample with an obstetric panel test were similar to those who were continuously enrolled for 42 weeks (Table S5 in Multimedia Appendix 1). Among commercially insured patients, the effect estimate for HIV infection was larger (OR 2.61 vs 1.96) for individuals with an obstetric panel test compared to patients enrolled for 42 weeks, although CIs overlapped, and gestational diabetes and preterm labor were not associated with higher odds of testing.

Discussion

Principal Findings

We estimated the rates of testing for hepatitis C during pregnancy in a large national sample of patients with Medicaid and commercial insurance using multiple methods. These estimates could serve as a baseline for understanding and monitoring the impact of CDC's hepatitis C screening recommendations on testing rates during pregnancy. Our results also suggest that patients who are pregnant are most often tested for hepatitis C early in pregnancy. However, Medicaid patients were more likely to be tested closer to the delivery date than commercially insured patients. This is likely linked to later entry into prenatal care as evidenced by the smaller proportion of Medicaid patients continuously enrolled in the 42 weeks prior to delivery, more frequent changes in insurance enrollment (ie, "churn"), or systematic differences in prenatal care practices among providers serving Medicaid patients. It may also be impacted by ongoing risk behaviors (eg, injection drug use) for HCV infection, which might trigger testing when disclosed to the provider. Testing rates for both Medicaid and commercially insured patients increased annually during 2015-2019. As national recommendations for testing during pregnancy did not change during the study period, these trends might reflect the increasing prevalence of or improved ascertainment of hepatitis C risk factors among pregnant persons and increasing awareness among clinicians.

Multiple factors associated with receipt of hepatitis C testing were identified. Results were largely similar for Medicaid and commercially insured patients. Younger patients were more likely to be tested than patients in the oldest age group. Among Medicaid patients, lower rates of testing were identified among Hispanic, non-Hispanic Black, and other non-White non-Hispanic patients. Lower access to prenatal care and quality of prenatal care among racial and ethnic minorities have been extensively reported in the literature [16]. Among commercially insured patients, those residing in the northeast were most likely to be tested. This might reflect regional differences in risk factor prevalence and providers' testing practices.

Patients with opioid use disorder and HIV infection were approximately 2 to 5 times more likely to be tested. Risk-based testing guidelines recommend testing for these patients; however, current CDC guidance recommends screening all pregnant persons during every pregnancy. Other observed associations such as those with alcohol use disorder and severe mental illness might also reflect unidentified injection drug use. Associations with an obstetric panel test and maternal vaccinations might be due to better retention in or quality of prenatal care. Preterm labor, high-risk pregnancy, obesity, and preeclampsia were also associated with higher odds of testing; these conditions might lead to more intensive prenatal care and inclusive testing practices. Associations with small effect sizes (eg, preterm labor) might also be due to residual confounding.

A single-center study conducted in an urban medical center in 2016 found that only 7% of pregnant patients were tested for hepatitis C [17]. Our estimates were substantially higher and may reflect differences in inclusion criteria (eg, pregnancies versus deliveries) and study populations. Low testing rates might result from a lack of assessment of hepatitis C risk factors and limited reporting of risk factors by pregnant individuals. Previous studies have reported that 10% to 59% of anti-HCV–positive pregnant persons reported no risk factors and that approximately two-thirds of patients who are pregnant reporting risk factors were not tested for hepatitis C [17,18]. A recent study of commercial laboratory test data found that 18% and 25% of obstetric panels for commercial and Medicaid patients, respectively, also had an HCV antibody test within a year [19]. Estimates of the hepatitis C testing rate were sensitive to inclusion criteria (eg, continuous enrollment thresholds). This was particularly evident for Medicaid patients for whom testing rates ranged from 15.3% to 23.9%. Continuously enrolled Medicaid and commercially insured patients had similar testing rates. Validation studies are required to assess which method is most accurate. However, the estimates of testing rate trends were robust to different methods of defining the study sample and estimating the testing rate.

Limitations

This study has several limitations. Administrative data for a very large convenience sample of pregnant persons were used, but we could not evaluate whether this population was representative of all persons who gave birth in the United States. While patients with private insurance and Medicaid were included, women with other forms of insurance or women who were uninsured were not included. The study was also limited to patients who gave birth and did not include pregnancies that ended in spontaneous abortion or termination. Incomplete follow-up was common for this sample and may have resulted in underestimation of testing prevalence if women were tested and then lost to follow-up. Insurance change is relatively common among pregnant persons in the United States; approximately 30% change insurance plans between the preconception period and delivery [20]. We were unable to validate the administrative codes used to identify pregnant persons. Some persons might have been misclassified. The prevalence of obstetric panel testing was lower than expected and might be due to misclassification of pregnancy or testing status (ie, due to missing test panel claims).

These data did not include a measure of socioeconomic status other than insurance type and race or ethnicity (Medicaid only), and variation in testing rate by income, rural and urban status, and so on were not examined. Although procedure codes were used to identify hepatitis C testing, the reason for testing (ie, screening or diagnostic testing) could not be ascertained nor whether testing was motivated by risk factor assessment. Finally, ICD-9/10-CM codes were used to define several clinical characteristics, but we could not assess the sensitivity or specificity of these codes. Additionally, patients with known chronic HCV infection were not excluded from the analysis; this could lead to underestimation of screening rates as these patients should not be included in the denominator.

Despite increases during the study period, testing for hepatitis C remained low in this large national sample of pregnant persons with commercial insurance and Medicaid. The USPSTF and CDC 2020 recommendations for universal hepatitis C screening during each pregnancy might lead to higher rates of testing. Education of providers to ensure testing during each pregnancy, and additional screening for high-risk behaviors that may necessitate repeat testing during pregnancy, are crucial to prevent perinatal transmission of HCV infection. Interventions to increase hepatitis C screening in other populations, such as electronic health record alerts or targeted patient outreach, could be adapted to the prenatal care setting and merit further study [21]. Racial and ethnic disparities in screening rates could be addressed through communication campaigns using trusted messengers to reach racial and ethnic minorities and prenatal care providers that serve these populations. Unrestricted timely treatment of hepatitis C among persons who are of childbearing age would also reduce perinatal transmission [22]. Interventions to increase testing are needed to address perinatal hepatitis C and eliminate hepatitis C in the United States.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily reflect the official position of the Centers for Disease Control and Prevention, or the authors' affiliated institutions.

Data Availability

MarketScan data are proprietary and cannot be shared without a purchase agreement with Merative, Inc.

Conflicts of Interest

HWK and WAM are employees of Quest Diagnostics and own stock in Quest Diagnostics. The other authors declare that they have no conflicts of interest.

Multimedia Appendix 1
Hepatitis C testing among pregnant patients with commercial or Medicaid insurance.
[DOCX File , 50 KB - publichealth_v9i1e40783_app1.docx]

References

1. Hofmeister MG, Rosenthal EM, Barker LK, Rosenberg ES, Barranco MA, Hall EW, et al. Estimating prevalence of hepatitis C virus infection in the United States, 2013-2016. Hepatology 2019;69(3):1020-1031 [FREE Full text] [doi: 10.1002/hep.30297] [Medline: 30398671]
2. Kim HS, Yang JD, El-Serag HB, Kanwal F. Awareness of chronic viral hepatitis in the United States: an update from the National Health and Nutrition Examination Survey. J Viral Hepat 2019;26(5):596-602. [doi: 10.1111/jvh.13060] [Medline: 30629790]
3. Thomas DL, Astemborski J, Rai RM, Anania FA, Schaeffer M, Galai N, et al. The natural history of hepatitis C virus infection: host, viral, and environmental factors. JAMA 2000;284(4):450-456 [FREE Full text] [doi: 10.1001/jama.284.4.450] [Medline: 10904508]
4. Benova L, Mohamoud YA, Calvert C, Abu-Raddad LJ. Vertical transmission of hepatitis C virus: systematic review and meta-analysis. Clin Infect Dis 2014;59(6):765-773 [FREE Full text] [doi: 10.1093/cid/ciu447] [Medline: 24928290]
5. Nwaohiri A, Schillie S, Bulterys M, Kourtis AP. Hepatitis C virus infection in children: how do we prevent it and how do we treat it? Expert Rev Anti Infect Ther 2018;16(9):689-694 [FREE Full text] [doi: 10.1080/14787210.2018.1509707] [Medline: 30091654]
6. Viral hepatitis surveillance report 2018—hepatitis C. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/hepatitis/statistics/2018surveillance/HepC.htm [accessed 2023-09-06]
7. Oliver C, Black J, De Pont S, Sizemore L, Wester C. Pregnancy status, risk factors, and opportunities for referral to care among reproductive-aged women with newly reported chronic hepatitis C virus infection in Tennessee. Public Health Rep 2020;135(1):90-96 [FREE Full text] [doi: 10.1177/0033354919887742] [Medline: 31738858]
8. Zibbell JE, Iqbal K, Patel RC, Suryaprasad A, Sanders KJ, Moore-Moravian L, Centers for Disease Control and Prevention (CDC). Increases in hepatitis C virus infection related to injection drug use among persons aged ≤30 years—Kentucky, Tennessee, Virginia, and West Virginia, 2006-2012. MMWR Morb Mortal Wkly Rep 2015;64(17):453-458 [FREE Full text] [Medline: 25950251]

9. Rossi RM, Wolfe C, Brokamp R, McAllister JM, Wexelblatt S, Warshak CR, et al. Reported prevalence of maternal hepatitis C virus infection in the United States. Obstet Gynecol 2020;135(2):387-395. [doi: 10.1097/AOG.0000000000003644] [Medline: 31923064]
10. Ahrens KA, Rossen LM, Burgess AR, Palmsten KK, Ziller EC. Rural-urban residence and maternal hepatitis C infection, U.S.: 2010-2018. Am J Prev Med 2021;60(6):820-830 [FREE Full text] [doi: 10.1016/j.amepre.2020.12.020] [Medline: 33640230]
11. Patrick SW, Bauer AM, Warren MD, Jones TF, Wester C. Hepatitis C virus infection among women giving birth—Tennessee and United States, 2009-2014. MMWR Morb Mortal Wkly Rep 2017;66(18):470-473 [FREE Full text] [doi: 10.15585/mmwr.mm6618a3] [Medline: 28493860]
12. Schillie SF, Canary L, Koneru A, Nelson NP, Tanico W, Kaufman HW, et al. Hepatitis C virus in women of childbearing age, pregnant women, and children. Am J Prev Med 2018;55(5):633-641. [doi: 10.1016/j.amepre.2018.05.029] [Medline: 30342628]
13. HCV guidance: recommendations for testing, managing, and treating hepatitis C. American Association for the Study of Liver Diseases and Infectious Diseases Society of America. URL: https://www.hcvguidelines.org/ [accessed 2023-09-06]
14. Schillie S, Wester C, Osborne M, Wesolowski L, Ryerson AB. CDC recommendations for hepatitis C screening among adults—United States, 2020. MMWR Recomm Rep 2020;69(2):1-17 [FREE Full text] [doi: 10.15585/mmwr.rr6902a1] [Medline: 32271723]
15. Smith BD, Morgan RL, Beckett GA, Falck-Ytter Y, Holtzman D, Teo CG, Centers for Disease Control and Prevention. Recommendations for the identification of chronic hepatitis C virus infection among persons born during 1945-1965. MMWR Recomm Rep 2012;61(RR-4):1-32 [FREE Full text] [Medline: 22895429]
16. Gadson A, Akpovi E, Mehta PK. Exploring the social determinants of racial/ethnic disparities in prenatal care utilization and maternal outcome. Semin Perinatol 2017;41(5):308-317. [doi: 10.1053/j.semperi.2017.04.008] [Medline: 28625554]
17. Boudova S, Mark K, El-Kamary S. Risk-based hepatitis C screening in pregnancy is less reliable than universal screening: a retrospective chart review. Open Forum Infect Dis 2018;5(3):ofy043 [FREE Full text] [doi: 10.1093/ofid/ofy043] [Medline: 29564364]
18. Kushner T, Park C, Masand D, Wagner B, Grace M, Rosenbluth E, et al. Hepatitis C seroprevalence among consecutive labor and delivery admissions in two New York City hospitals. Open Forum Infect Dis 2020;7(11):ofaa514 [FREE Full text] [doi: 10.1093/ofid/ofaa514] [Medline: 33241070]
19. Kaufman HW, Osinubi A, Meyer WA, Khan M, Huang X, Panagiotakopoulos L, et al. Hepatitis C virus testing during pregnancy after universal screening recommendations. Obstet Gynecol 2022;140(1):99-101 [FREE Full text] [doi: 10.1097/AOG.0000000000004822] [Medline: 35849464]
20. Daw JR, Kozhimannil KB, Admon LK. High rates of perinatal insurance churn persist after the ACA. Health Affairs Blog. 2019. URL: https://www.healthaffairs.org/content/forefront/high-rates-perinatal-insurance-churn-persist-after-aca [accessed 2023-09-06]
21. Yartel AK, Rein DB, Brown KA, Krauskopf K, Massoud OI, Jordan C, et al. Hepatitis C virus testing for case identification in persons born during 1945-1965: results from three randomized controlled trials. Hepatology 2018;67(2):524-533 [FREE Full text] [doi: 10.1002/hep.29548] [Medline: 28941361]
22. Thompson WW, Symum H, Sandul A, DHSc, Gupta N, Patel P, et al. Vital signs: hepatitis C treatment among insured adults—United States, 2019-2020. MMWR Morb Mortal Wkly Rep 2022;71(32):1011-1017 [FREE Full text] [doi: 10.15585/mmwr.mm7132e1] [Medline: 35951484]

Abbreviations

CDC: Centers for Disease Control and Prevention
CPT: Current Procedural Terminology
HCV: hepatitis C virus
ICD-9/10-CM: International Classification of Diseases, Ninth/Tenth Revision, Clinical Modification
OR: odds ratio
Tdap: tetanus-diphtheria-acellular pertussis
USPSTF: United States Preventive Services Task Force

Edited by A Mavragani, T Sanchez; submitted 05.07.22; peer-reviewed by A Lansky, R Epstein, K Nugent; comments to author 27.02.23; revised version received 31.05.23; accepted 10.08.23; published 27.09.23.

Please cite as:
Khan MA, Thompson WW, Osinubi A, Meyer 3rd WA, Kaufman HW, Armstrong PA, Foster MA, Nelson NP, Wester C
Testing for Hepatitis C During Pregnancy Among Persons With Medicaid and Commercial Insurance: Cohort Study
JMIR Public Health Surveill 2023;9:e40783
URL: https://publichealth.jmir.org/2023/1/e40783
doi:10.2196/40783
PMID:37756048

Original Paper

Moderating Effect of eHealth Literacy on the Associations of Coronaphobia With Loneliness, Irritability, Depression, and Stigma in Chinese Young Adults: Bayesian Structural Equation Model Study

Richard Huan Xu[1*], PhD; Ho Hin Chan[2*], PhD; Lushaobo Shi[3], MPhil; Ting Li[3], MPhil; Dong Wang[3], Prof Dr

[1]Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Kowloon, China (Hong Kong)
[2]Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, China (Hong Kong)
[3]School of Health Management, Southern Medical University, Guangzhou, China
[*]these authors contributed equally

Corresponding Author:
Dong Wang, Prof Dr
School of Health Management
Southern Medical University
Shatai Nan Road
Guangzhou, 510515
China
Phone: 86 61647576
Email: dongw96@smu.edu.cn

Abstract

Background: The COVID-19 pandemic has led to an increase in known risk factors for mental health problems. Although medical information available through the internet and smartphones has greatly expanded, people’s ability to seek, eschew, and use reliable web-based medical information and services to promote their mental health remains unknown.

Objective: This study aims to explore the associations between coronaphobia and 4 frequently reported mental health problems, loneliness, irritability, depression, and stigma, during the COVID-19 pandemic and to assess the moderating effects of eHealth literacy (eHL) on the adjustment of these relationships in Chinese young adults.

Methods: The data used in this study were collected from a web-based survey of the general Chinese population, aged between 18 and 30 years, conducted in China between December 2022 and January 2023. A nonprobability snowball sampling method was used for data collection. A Bayesian structural equation model (BSEM) using parameter expansion was used to estimate the moderating effect of eHL on the relationship between coronaphobia and psychological problems. The posterior mean and 95% highest density intervals (HDIs) were estimated.

Results: A total of 4119 participants completed the questionnaire and provided valid responses. Among them, 64.4% (n=2653) were female and 58.7% (n=2417) were rural residents. All measures showed statistically significant but minor-to-moderate associations (correlation coefficients ranged from −0.04 to 0.65). Significant heterogeneity was observed between rural and urban residents at the eHL level, and coronaphobia was observed. The BSEM results demonstrated that eHL was a significant moderator in reducing the negative effects of coronaphobia on loneliness (posterior mean −0.0016, 95% HDI −0.0022 to −0.0011), depression (posterior mean −0.006, 95% HDI −0.0079 to −0.004), stigma (posterior mean −0.0052, 95% HDI −0.0068 to −0.0036), and irritability (posterior mean −0.0037, 95% HDI −0.0052 to −0.0022). The moderating effects of eHL varied across the rural and urban subsamples.

Conclusions: Using BSEM, this study demonstrated that improving eHL can significantly mitigate the negative effects of coronaphobia on 4 COVID-19–related mental health problems in Chinese young adults. Future eHL initiatives should target rural communities to ensure equal access to information and resources that can help protect their mental health during the pandemic.

(JMIR Public Health Surveill 2023;9:e47556) doi:10.2196/47556

KEYWORDS

coronaphobia; eHealth literacy; Bayesian statistics; structural equation modeling; mediating effect; mental health

Introduction

The COVID-19 pandemic, which originated in China and spread worldwide, has continued its impact for more than 3 years. This has led to an increase in the known risk factors for mental health problems, such as social isolation, loneliness, inactivity, and limited access to health and social care services [1,2]. Studies have indicated that more than half of those interviewed have experienced moderate-to-severe psychological effects caused by fear of COVID-19, and approximately three-quarters have expressed concern about the health of their family members [3-5].

Experiencing strong fear is normal under uncertain circumstances [6]. However, the COVID-19 pandemic has disrupted people’s ability to lead normal lives. Unlike previous flu epidemics, the fear associated with COVID-19 extends beyond the virus itself, impacting mental health in various ways. This excessive fear response to contracting COVID-19 has led to extreme concern over physiological symptoms, significant stress regarding personal and occupational loss, and avoidance of public places and situations [7]. Along with social distancing and strict quarantine policies worldwide, these fears have generated unprecedented stress on individuals, mentally and emotionally. This uncontrolled and persistent fear has resulted in a new and specific type of anxiety called coronaphobia, exacerbating the consequences on mental health [8].

The COVID-19 pandemic has triggered changes in people’s health care–seeking behavior because of restrictions on social contact, highlighting the importance of digital networks and web-based service platforms [9]. The availability of medical information through the internet and smartphones has greatly expanded the sources for people to seek mental health knowledge and web-based services without leaving their homes during the pandemic. Examples include web-based consultations [10], prescriptions [11], and wireless sensors to monitor the medical conditions of patients [12]. However, the quality of web-based information and services can vary in terms of verifiability, bias, and source. To efficiently navigate the internet, possessing good eHealth literacy (eHL) is particularly important. eHL is an essential skill that ensures people can access reliable web-based medical information and care [13]. Studies have shown that eHL can be useful in alleviating health disparities and improving patient experiences. Equipped with advanced eHL, individuals can search, evaluate, and apply massive amounts of information available on the internet, leading to a significant reduction in their psychological stress [14,15].

The relationship between eHL and psychological status has been widely discussed but is limited in the context of the COVID-19 pandemic. Yang et al [16] found that high eHL was significantly associated with lower levels of depression, insomnia, and posttraumatic stress disorder in 15,000 Chinese individuals. Xu et al [17] also discovered that young people with high eHL who use web-based mental health–related information tend to have a more positive attitude toward seeking mental health services and reporting better mental well-being. Another study demonstrated that rural residents with high eHL were more willing to use telemedicine and had better health status, including mental health, at the beginning of the COVID-19 pandemic [18]. The psychological impact of the pandemic has been moderate to severe because of isolation or quarantine measures. Given the rise of the digital era, understanding how people use the internet to manage their mental health during a pandemic is vital.

Research indicates that eHL can help mitigate the negative impact of COVID-19 fear on the quality of life during the pandemic among diverse populations, including patients [19] and university students [20,21]. Higher levels of eHL are associated with better adherence to preventive measures and improved psychological well-being. Additionally, abundant evidence exists regarding the relationship between coronaphobia and mental health during pandemics. Amin [22] demonstrated that coronaphobia persists among health care professionals and results in psychological symptoms that affect mental health. Yayla and İlgin [23] and Labrague and De Los Santos [24] indicated that nurses’ psychological well-being was significantly affected by coronaphobia. Although a recent study found that health literacy benefits the management of coronaphobia [25], the impact of eHL on coronaphobia during the pandemic has not been sufficiently studied. Considering that the pandemic has profoundly altered how people seek, use, and evaluate health care services, the role of eHL in this transformation should receive more attention.

However, alongside the pandemic, numerous rumors, hoaxes, and misinformation regarding the prevalence, outcomes, and prevention of COVID-19 have circulated on social media at an alarming rate [3-5]. These false claims have generated mass hysteria and panic regarding COVID-19, leading to unexpected and enduring psychological problems for the public [26]. Research has demonstrated that unreliable web-based information about the pandemic has led to heightened anxiety and even lethal mental health ramifications [27,28], which are potentially more detrimental in the long term than the virus itself. Despite this, evidence regarding the role of eHL in managing the impact of coronaphobia on psychological problems and mental illnesses during the pandemic is lacking.

This study aims to provide empirical evidence of the moderating effects of eHL on the relationship between coronaphobia and a wide range of mental health problems frequently reported during the pandemic in young adults in China. Exploration in young adults is particularly important because they are highly exposed to smartphones and the internet, which increases the likelihood of experiencing web-related psychological problems. Research has confirmed the profound impact of the pandemic on young adults, disrupting their education, routines, and peer interactions, thereby raising concerns about potential long-term mental health outcomes in this population [29].

Methods

Data and Participants

The data used in this study were obtained from a web-based survey conducted in China between December 2022 and January 2023. The study included participants who met the following criteria: (1) participants aged between 18 and 30 years, (2) participants able to read Chinese, and (3) participants able to provide informed consent. The survey was conducted at a medical university in Guangzhou, China. The research team contacted university officers in charge of student affairs and discussed the data collection plan. A link containing the study introduction and questionnaire was sent to all the students through the university's internal network using WeChat software (Shenzhen Tencent Computer Systems Company Limited), a multipurpose messaging app. To improve the diversity of the samples, a nonprobability snowball sampling method was used for recruitment. This method is commonly used in quantitative studies [30,31]. The research team did not directly recruit participants other than appointed university students but encouraged them to contact their peers and invited them to join the survey through their personal social networks. To maintain the momentum of the snowball sampling, the research team collaborated with university officials and sent 4 reminders to all students, encouraging them to further disseminate the survey link.

The questionnaire was created using the web-based platform of a professional surveying company (Wenjuanxing), with an informed consent statement appended to it. Interested participants accessed the questionnaire by clicking on the provided link. The first page of the questionnaire focused on obtaining informed consent. Participants had to read, scroll down to the end, click the agree button, and electronically sign it before proceeding to the actual survey. Once participants confirmed their participation, they were presented with a structured questionnaire containing questions about their demographics, socioeconomic status, mental health status, and eHL. Participants required approximately 15 minutes to complete the entire questionnaire.

Measures

eHealth Literacy

The eHealth Literacy Scale (eHEALS) was used to measure participants' knowledge and perceived skills in finding, evaluating, and applying eHealth information to manage health problems [32]. It was developed based on a framework comprising 6 dimensions to understand and use eHealth information. The eHEALS has 8 items (eg, I know how to find helpful health resources on the internet) rated on a 5-point Likert scale. The total eHEALS scores range from 8 to 40. Higher scores indicate greater perceived eHL. Its psychometric properties in the Chinese population have been confirmed by Xu et al [33].

Coronaphobia

Coronaphobia was assessed using the COVID-19 Phobia Scale (C19P) [34]. This self-report instrument was developed and validated to measure COVID-19 phobia levels. C19P includes 4 factors (psychological, somatic, social, and economic) with 20 items (eg, the fear of coming down with COVID-19 makes me very anxious) rated on a 5-point Likert scale from "strongly disagree" to "strongly agree." The total C19P score ranges from 20 to 100, with a high score indicating severe coronaphobia. The psychometric properties in the Chinese population have been reported by Chi et al [35].

Loneliness

The University of California, Los Angeles (UCLA) 3-item loneliness scale (ULS-3) was used to determine the perception and degree of loneliness [36]. It comprises 3 items (eg, how often do you feel that you lack companionship?) to assess an individual's feelings of companionship, being left out, and being isolated. Each item is rated on a 3-point Likert scale. The total ULS-3 score ranges from 3 to 9, with a higher score indicating a higher level of perceived loneliness. The psychometric properties of the Chinese ULS-3 have been reported by Liu et al [37].

Irritability

The brief irritability test (BIT) developed by Holtzman et al [38] was applied to assess irritation. This brief self-report scale for irritability demonstrates strong reliability and validity in both healthy and clinical populations. It includes 5 items (eg, I have been feeling like I might snap) rated on a 6-point Likert scale (1 to 6). A higher score indicates a higher level of irritation. In this study, the research team translated the BIT from English to Chinese using a standard process (2 forward and 2 backward translations). The Chinese version of the BIT is presented in Multimedia Appendix 1. In this study, the Cronbach α for the Chinese BIT was .89 indicating good internal consistency. Confirmatory factor analysis supported a one-factor structure (comparative fit index=0.99 and root mean square error of approximation=0.03).

Depression

Depressive symptoms were assessed using the Patient Health Questionnaire-9 (PHQ-9). The PHQ-9 is one of the most widely used instruments for screening, diagnosing, monitoring, and measuring depression severity. It consists of 9 items (eg, little interest or pleasure in doing things) rated on a 4-point Likert scale (0-3). A total score between 5 and 9 indicates mild depression, 10 and 14 indicates moderate depression, and 15 and above indicates moderately severe depressive symptoms. The psychometric properties of the Chinese version of the PHQ-9 have been reported in different populations [39].

Anticipated Stigma

COVID-19–related stigma was measured using the revised Chronic Illness Anticipated Stigma Scale (r-CIASS) [40]. It has 6 items (eg, [if I was to get coronavirus...] a friend or family member would be angry with me), reflecting how individuals would be treated if they were to become infected with COVID-19, which are rated on a scale from very unlikely (1) to very likely (4). Higher scores indicate a higher degree of stigma. In this study, the research team translated the r-CIASS from English to Chinese using a standard process (2 forward and 2 backward translations). The Chinese version of the r-CIASS is presented in Multimedia Appendix 2. In this study, Cronbach α for the Chinese r-CIASS was .92. Confirmatory factor analysis supported a one-factor structure (comparative fit index=0.99 and root mean square error of approximation=0.04).

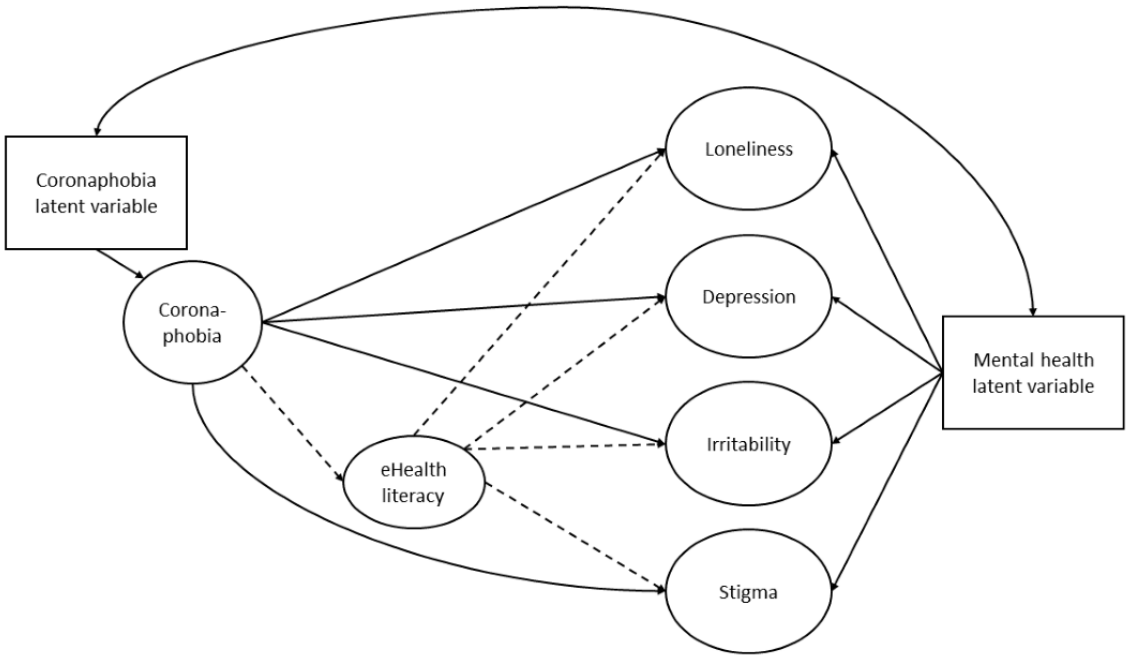
Statistical Analysis

A descriptive analysis was conducted to examine the participants’ background characteristics and mental health statuses. The min-max normalization method was applied to perform a linear transformation of the original data to then scale the data in the range between 0 and 100.

To assess the moderating effect of eHL on the relationship between coronaphobia and the 4 psychological problems, the Bayesian structural equation model (BSEM) via parameter expansion was used [41].

Bayesian methods are becoming increasingly popular in various fields because they enable the simultaneous estimation of all cross-loadings and residual correlations in a particular model, a task that is not possible with a maximum-likelihood (ML)–based estimation. Traditionally, ML estimation considers parameters as constants and aims to identify the best-fitting models for certain parameters using research data. However, Bayesian estimation involves combining data likelihoods and prior distributions to construct a posterior distribution. This combination has led researchers to apply the Bayes theorem to the estimation processes [42]. Using BSEM can improve the estimation of parameter and latent variables, enable statistics for model comparison, and provide more reliable results for smaller samples [43]. In this study, the BSEM incorporates 2 latent variables: one representing coronaphobia and the other encompassing loneliness, depression, stigma, and irritation. These 2 latent variables exhibit correlated residuals. Figure 1 illustrates the conceptual framework of the BSEM.

Figure 1. Conceptual framework of Bayesian structural equation model analysis.



Given the reparameterization of the covariances, the prior distributions on the working model parameters were specified. A subset of 800 participants was randomly selected from the full data set to calculate the prior probability. Their demographics are presented in Multimedia Appendix 3. For the BSEM, we ran 2 chains, each with 21,000 burn-in iterations and 4000 inference iterations. We verified the convergence of the trace plots for all variables and reported the statistics of the latent and regression variables. Information about the posteriors was summarized in terms of the mean and SD, and a 95% highest density interval (HDI) of the posterior mean was calculated. We focused on the direct effects of the 4 aforementioned relationships involving the coronaphobia variable. The 4 indirect effects were calculated using Sobel test [44]. To calculate the indirect effect of the eHL variable on the direct effect of coronaphobia on loneliness, we denote the coefficients of coronaphobia for eHL and eHL for loneliness as $c_{Coronaphobia,eHL}$ and $c_{eHL,loneliness}$, respectively. Their SDs are denoted as $SDc_{Coronaphobia,eHL}$ and $SDc_{eHL,loneliness}$, respectively.

Using the following formula derived from the Sobel test, we calculate the posterior mean $c_{indirect(Coronaphobia,eHL,loneliness)}$ and $SDc_{indirect(Coronaphobia,eHL,loneliness)}$ of the indirect effect. The same method was used to calculate the posterior mean and SD of the indirect effects of the other 3 variables. R software was used for all statistical analyses. The significance level was set as <.05.

$$c_{indirect(Coronaphobia,eHL,loneliness)}=c_{Coronaphobia,eHL}\cdot c_{eHL,loneliness} \quad \textbf{(1)}$$

(2)

Ethics Approval

The study protocol and informed consent were approved by the Human Research Ethics Committee of Hong Kong Polytechnic University (Ref No.: HSEARS-20210328002). Written informed consent was obtained from all the participants. All study data were collected anonymously and no compensation was provided to any participant.

Results

Participants' Characteristics

Table 1 shows that 4119 participants joined the survey and completed the questionnaire (completion rate=94.9%). Among them, 64.4% (n=2653) were female, 58.7% (n=2417) were rural residents, and nearly one-third reported a perceived family income lower than the local average (n=1220). Regarding COVID-19–related characteristics, approximately 70% (n=3024) had been infected with COVID-19. Approximately 86.6% (n=2567) and 90.6% (n=3733) reported having experienced a lockdown and working or studying from home because of COVID-19, respectively.

Table 1. Background characteristics of participants (n=4119).

Characteristics	Values, n (%)
Gender	
Male	1466 (35.6)
Female	2653 (64.4)
Age	
22 years or younger	2120 (51.5)
Older than 22 years	1999 (48.5)
Educational attainment	
Undergraduate or below	3115 (75.6)
Postgraduate or above	1004 (24.4)
Residence registry	
Rural resident	2417 (58.7)
Urban resident	1702 (41.3)
Perceived family income	
Lower than average	1220 (29.6)
Equal to average	2601 (63.1)
Higher than average	298 (7.2)
People infected with COVID-19	
No	1175 (30)
Yes	3024 (70)

Profiles of and Correlations Between Measures

In terms of eHL, the participants reported a mean eHEALS score of 28.6 (8-40). After normalization, the participants demonstrated a slightly higher level of phobia (C19P's mean 44; IQR 0-100) than other types of mental health problems (mean 24-40.2; IQR 0-100). The correlation analysis revealed a statistically significant but minor-to-moderate association between the measures (Table 2).

Table 2. Profiles of measures and the correlations between them.

	Mean (SD)	Median (IQR)	Normalized mean (0-100)	1	2	3	4	5
1. eHEALS[a]	28.6 (6.1)	29 (8-40)	64.5					
Correlation, *r*				—[b]				
P value				—				
2. C19P[c]	55.2 (16.6)	54 (20-100)	44					
Correlation, *r*				–0.14	—			
P value				<.001	—			
3. ULS-3[d]	5 (1.6)	5 (3-9)	33.4					
Correlation, *r*				–0.08	0.29	—		
P value				<.001	<.001	—		
4. BIT[e]	15.1 (4.9)	15 (5-30)	40.2					
Correlation, *r*				–0.04	0.36	0.5	—	
P value				.01	<.001	<.001	—	
5. PHQ-9[f]	15.5 (5.8)	14 (9-36)	24					
Correlation, *r*				–0.07	0.37	0.54	0.65	—
P value				<.001	<.001	<.001	<.001	—
6. r-CIASS[g]	12.1 (4.6)	12 (6-24)	33.7					
Correlation, *r*				–0.09	0.32	0.42	0.44	0.51
P value				<.001	<.001	<.001	<.001	<.001

[a] eHEALS: eHealth Literacy Scale.

[b] Not available.

[c] C19P: COVID-19 Phobia Scale.

[d] ULS-3: University of California, Los Angeles 3-Item Loneliness Scale.

[e] BIT: Brief Irritability Test.

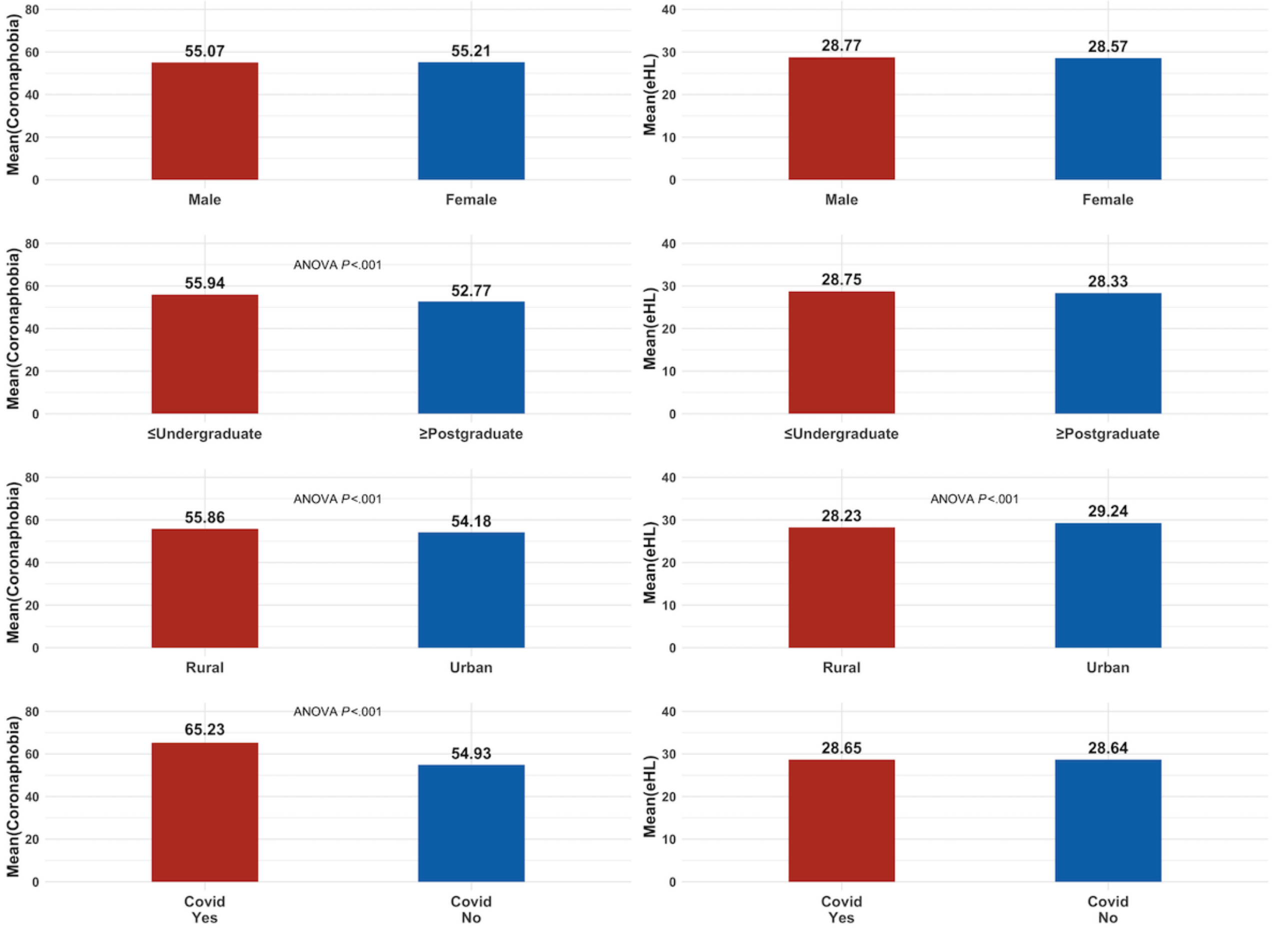
[f] PHQ-9: Patient Health Questionnaire-9.

[g] r-CIASS: Revised Chronic Illness Anticipated Stigma Scale.

Coronaphobia and eHL in Different Socioeconomic and COVID-19–Related Status Groups

Rural residents reported a higher level of coronaphobia but a lower level of eHL than their urban counterparts. Participants with a higher perceived family income tended to have a high level of eHL, but a low level of coronaphobia. Additionally, participants with COVID-19 demonstrated high levels of coronaphobia (Figure 2).

Figure 2. Mean scores of COVID-19 Phobia Scale (C19P) and eHealth Literacy Scale (eHEALS) in subgroups. eHL: eHealth literacy.



Direct Effect and the Moderating Effect of eHL

Using the Bayesian approach, we report the results in terms of the posterior distribution. Table 3 shows a significantly strong and positive effect of latent variable 2 on depression, stigmatization, and irritation (with longlines set as the reference). This suggests the presence of a latent variable linked to these 4 variables. Considering the significant difference in eHL and coronaphobia between rural and urban residents, separate analyses were conducted for these 2 subsamples.

The posterior mean and posterior SD of the direct effects of coronaphobia on the 4 psychological problems are presented in Table 4. The coefficients are all statistically significant, as indicated by the 95% HDI. The direct effect of coronaphobia on depression was stronger than that on the other 3 psychological problems. Compared to rural residents, coronaphobia was more likely to lead to a higher level of depression and loneliness but lower levels of stigma and irritability among urban residents. Table 5 shows that eHL was a significant moderator in reducing the negative effect of coronaphobia on loneliness (posterior mean −0.0016, 95% HDI −0.0022 to −0.0011), depression (posterior mean −0.006, 95% HDI −0.0079 to −0.004), stigma (posterior mean −0.0052, 95% HDI −0.0068 to −0.0036), and irritability (posterior mean −0.0037, 95% HDI −0.0052 to −0.0022). The trace plot shows a random scatter around the mean value. Thus, our model results suggest that the chains mixed well and that the model converged (Figure 3). Other trace plots are presented in Multimedia Appendix 4.

Table 3. Latent variable coefficients within mental health latent variable using Bayesian structural equation models (BSEMs) and stratified by respondent’s residence registry.

	Posterior, mean (SD)	95% highest density interval
Full data (n=4199)		
Latent→Loneliness	1.000[a]	—[b]
Latent→Depression	4.718 (0.142)	4.447-5.007
Latent→Stigma	2.463 (0.090)	2.289-2.646
Latent→Irritability	3.588 (0.108)	3.381-3.805
Rural residents (n=2074)		
Latent→Loneliness	1.000[a]	—
Latent→Depression	4.815 (0.194)	4.455-5.209
Latent→Stigma	2.583 (0.123)	2.348-2.836
Latent→Irritability	3.672 (0.150)	3.393-3.974
Urban residents (n=2045)		
Latent→Loneliness	1.000[a]	—
Latent→Depression	4.623 (0.213)	4.222-5.062
Latent→Stigma	2.311 (0.135)	2.058-2.583
Latent→Irritability	3.502 (0.164)	3.192-3.838

[a] SD values are not available.

[b] Not available.

Table 4. Direct effects of associations between coronaphobia and mental health outcomes using Bayesian structural equation models (BSEMs) and stratified by respondent’s residency registry.

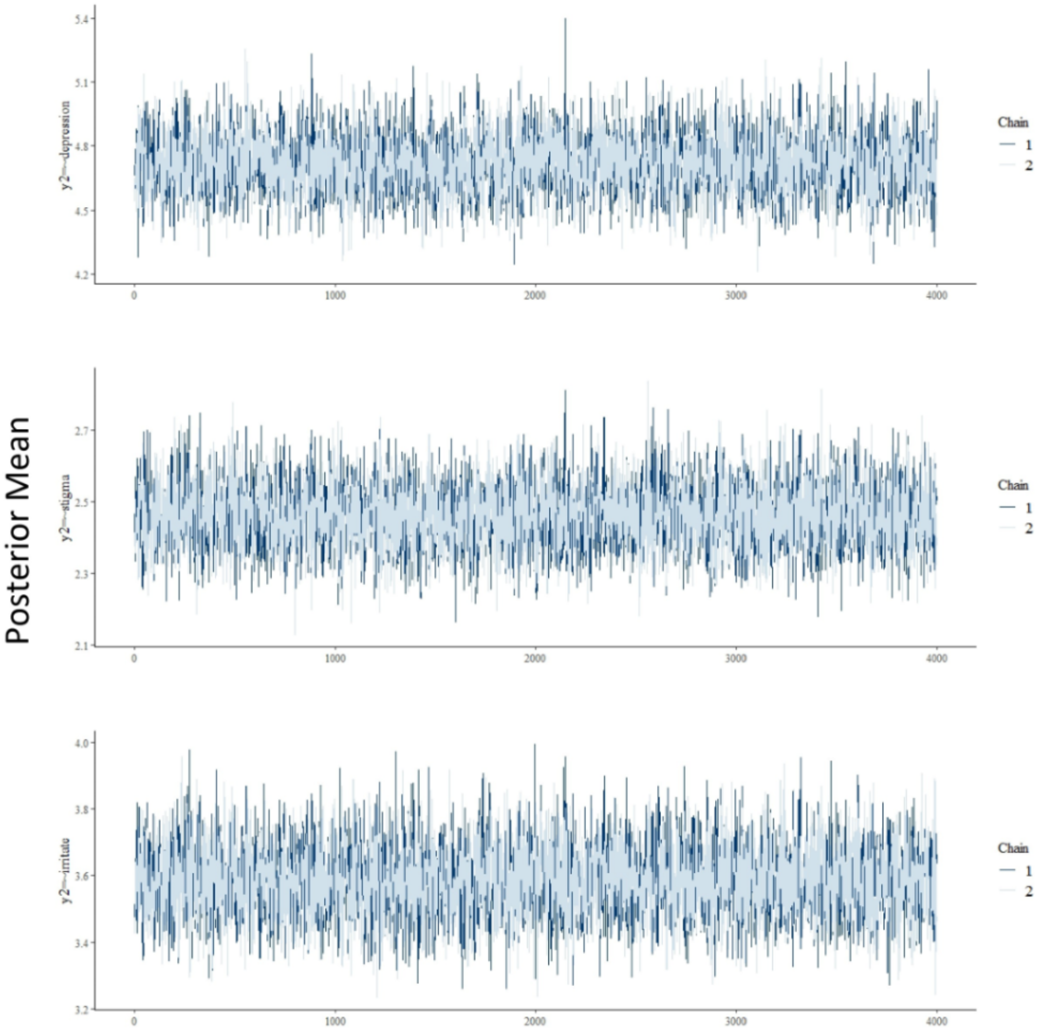
	Posterior, mean (SD)	95% highest density interval
Full (n=4119)		
Coronaphobia→Loneliness	0.068 (0.026)	0.019-0.119
Coroanphobia→Depression	0.313 (0.122)	0.083-0.554
Coronaphobia→Stigma	0.186 (0.064)	0.066-0.314
Coronaphobia→Irritability	0.245 (0.093)	0.071-0.431
Rural residents (n=2074)		
Coronaphobia→Loneliness	0.066 (0.025)	0.02-0.117
Coroanphobia→Depression	0.316 (0.118)	0.095-0.558
Coronaphobia→Stigma	0.194 (0.064)	0.075-0.325
Coronaphobia→Irritability	0.260 (0.090)	0.089-0.445
Urban residents (n=2045)		
Coronaphobia→Loneliness	0.072 (0.026)	0.023-0.125
Coroanphobia→Depression	0.318 (0.120)	0.096-0.562
Coronaphobia→Stigma	0.185 (0.060)	0.073-0.308
Coronaphobia→Irritability	0.238 (0.091)	0.068-0.421

Table 5. Moderating effects of eHL[a] in adjusting associations between coronaphobia and mental health outcomes using Bayesian structural equation models (BSEMs) and stratified by residence registry.

	Posterior, mean (SD)	95% highest density interval
Full (n=4119)		
Coronaphobia→eHL→Loneliness	–0.0016 (0.0003)	–0.0022 to –0.0011
Coronaphobia→eHL→Depression	–0.0060 (0.0010)	–0.0079 to –0.004
Coronaphobia→eHL→Stigma	–0.0052 (0.0008)	–0.0068 to –0.0036
Coronaphobia→eHL→Irritability	–0.0037 (0.0008)	–0.0052 to –0.0022
Rural residents (n=2074)		
Coronaphobia→eHL→Loneliness	–0.0020 (0.0005)	–0.0029 to –0.0012
Coronaphobia→eHL→Depression	–0.0081 (0.0017)	–0.0113 to –0.0048
Coronaphobia→eHL→Stigma	–0.0067 (0.0014)	–0.0094 to –0.004
Coronaphobia→eHL→Irritability	–0.0060 (0.0014)	–0.0074 to –0.005
Urban residents (n=2045)		
Coronaphobia→eHL→Loneliness	–0.0013 (0.0003)	–0.0019 to –0.0006
Coronaphobia→eHL→Depression	–0.0042 (0.0011)	–0.0064 to –0.0021
Coronaphobia→eHL→Stigma	–0.0039 (0.0010)	–0.0059 to –0.002
Coronaphobia→eHL→Irritability	–0.0021 (0.0008)	–0.0037 to –0.0006

[a] eHL: eHealth literacy.

Figure 3. Traceplots of convergence for some model parameters.



Discussion

Principal Findings

The results of the BSEM showed that Chinese young adults with high eHL levels were less likely to report a negative impact of coronaphobia on their mental health. While previous studies have indicated that digital tools and mobile technologies played a vital role in helping people manage their mental health during the pandemic, our study adds to the knowledge that the effectiveness and efficiency of web-based interventions or programs may depend on users' eHL. This study provides empirical evidence that simply offering web-based mental health services may not be sufficient to enhance the mental health outcomes of young adults; thus, improving their eHL is essential. For policy makers, comprehending and developing an eHealth strategy that creates an interactive relationship between mental health providers, partners, and users and empowers them to use web-based resources to reduce the effect of coronaphobia on their mental health is critical.

Research has demonstrated that the fear of COVID-19 can reduce the use of health care services. For example, one study found that COVID-19–related stigma can cause people to underreport symptoms and avoid health facilities [45]. Moreover, the fear of COVID-19 can cause people to delay or avoid seeking medical attention for non–COVID-19–related illnesses, which can lead to negative health outcomes [46]. Fear can also affect mental health and lead to poor outcomes. Although eHL interventions can help reduce fear and anxiety related to COVID-19, low eHL affects a large percentage of the global population. Naeem and Boulos [47] indicated that low eHL directly contributes to the spread of COVID-19-related web-based misinformation. Lee et al [48] found that 68% of adults were exposed to COVID-19–related misinformation through social networking services or instant messaging. The widespread dissemination of misinformation through social media can render distinguishing between accurate information and falsehoods difficult for people, which can further increase anxiety and fear regarding COVID-19.

Our study identified that people's ability to find and use health information from digital sources is essential, given the increasing number of people turning to the internet for health information to improve their coping skills against mental health stressors. Participants with high eHL were more likely to find accurate and trustworthy information and resources that could significantly reduce the negative impact of coronaphobia on their mental health. By improving eHL, individuals can better manage their COVID-19–related fear and anxiety and prevent them from developing into a phobia [49]. This can have positive effects on their overall mental health and well-being and complement the efforts to prevent the spread of the virus. Improving eHL can be an important step in this process as it can help people access accurate information and resources that can help them manage their fear and anxiety related to COVID-19.

In China, the COVID-19 pandemic has impacted both urban and rural residents, although not necessarily equally [50]. Our study revealed that rural residents reported significantly higher levels of coronaphobia than urban residents. This finding partially aligns with previous findings that rural residents have been disproportionately affected by the COVID-19 pandemic [50]. Empirical evidence confirms that the expansion of telehealth services is a potential solution for bridging the gap in health care access between rural and urban areas. However, our study suggests that compared to urban residents, rural residents showed a lower level of eHL, which may undermine the efficiency of the intervention to improve mental health during the pandemic, resulting in increased disparity and inequality. This finding is consistent with those of previous studies. For example, Rush et al [18] indicated that rural Canadian residents experienced challenges with telemedicine access because of unreliable internet access and found the service impersonal. Witten and Humphry [51] reported that the eHL of rural community residents was insufficient for the proper understanding and use of technology. Initiating public health campaigns targeting rural communities is vital to improve their eHL and ensure adequate resources and support are provided to these areas to mitigate the impact of the pandemic on their mental health.

Our findings confirmed the positive role of eHL in mitigating the negative impact of the pandemic on mental health. However, this conclusion specifically applies to the populations aged between 18 and 30 years. Although this age group comprises a high proportion of internet users in China, with approximately 1.01 billion active users, further research is needed to explore these associations in other age cohorts. Obtaining empirical evidence across different age groups can provide valuable insights for policymaking in future pandemics, as the threat of COVID-19 persists and new public health crises may emerge.

Strength and Limitations

A significant advantage of our study is the use of BSEM to estimate the associations. BSEM allows for the incorporation of prior information into the model estimation process and provides a more efficient approximation of the model compared to the standard ML approach [52].

However, this study also has certain limitations. The primary limitation is the use of a snowball sampling method for data collection, which has several disadvantages. First, it is a nonprobabilistic technique that relies heavily on existing relationships between respondents, which can lead to insufficient diversity in respondents' characteristics. Second, the quality of the data received from the respondents depends on their relationship with the introducer. If the introducer has a good relationship with the respondent, the purpose of the survey can be clearly expressed, which can enhance the likelihood of obtaining valid responses. Finally, unlike other sampling methods, the snowball sampling method shifts the recruitment efforts to the respondents themselves, leading to a loss of control for the research team. In this study, given the nature of the web-based survey, we did not record the recruitment process. Thus, this method may not accurately represent the target population, leading to potential problems with the generalizability of our findings. Another limitation is that the data used in this study were collected during the most severe wave of the COVID-19 pandemic in China. Respondents may, in turn, have experienced strong coronaphobia, which may have led to recall bias.

Conclusions

This study provides evidence that coronaphobia has a significant negative impact on the mental health of Chinese young adults. However, the findings also highlight the potential of improving eHL to mitigate these negative effects. By improving individuals' ability to access reliable web-based information about the virus and resources for web-based health services, they may feel more confident and in control of their health, thereby reducing their phobia about COVID-19. Moreover, eHL initiatives should specifically target rural communities to ensure equal access to information and resources and protect their mental health during the pandemic.

Acknowledgments

This study was funded by the National Social Science Fund of China (20VYJ069) and Hong Kong Health and Medical Research Fund (18190641).

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author on reasonable request. The raw data supporting the conclusions of this paper will be made available upon reasonable request by contacting the corresponding author.

Authors' Contributions

RHX contributed toward the study concept and design, data analysis and interpretation, software, writing the original draft, reviewing, and editing. HHC contributed toward data analysis and interpretation, software, writing the original draft, reviewing, and editing. LS contributed toward the data collection, reviewing, and editing. TL contributed toward the data collection, reviewing, and editing. DW contributed toward the study concept and design, provision of study materials or patients, supervision, reviewing, and editing. All authors contributed to the paper and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The Chinese version of the Brief Irritability Test.
[DOCX File , 15 KB - publichealth_v9i1e47556_app1.docx]

Multimedia Appendix 2
The Chinese version of the Revised Chronic Illness Anticipated Stigma Scale.
[DOCX File , 15 KB - publichealth_v9i1e47556_app2.docx]

Multimedia Appendix 3
Demographics of sample of 800 participants.
[DOCX File , 20 KB - publichealth_v9i1e47556_app3.docx]

Multimedia Appendix 4
Traceplots of all associations.
[PDF File (Adobe PDF File), 214 KB - publichealth_v9i1e47556_app4.pdf]

References

1. Burgess R. COVID-19 mental-health responses neglect social realities. Nature 2020:1-2 [FREE Full text] [doi: 10.1038/d41586-020-01313-9] [Medline: 32366975]
2. Ornell F, Schuch JB, Sordi AO, Kessler FHP. "Pandemic fear" and COVID-19: mental health burden and strategies. Braz J Psychiatry 2020;42(3):232-235 [FREE Full text] [doi: 10.1590/1516-4446-2020-0008] [Medline: 32267343]
3. Marzo RR, Ismail Z, Htay MNN, Bahari R, Ismail R, Villanueva EQ, et al. Psychological distress during pandemic Covid-19 among adult general population: result across 13 countries. Clin Epidemiol Glob Health 2021;10:100708 [FREE Full text] [doi: 10.1016/j.cegh.2021.100708] [Medline: 33619459]
4. Lahav Y. Psychological distress related to COVID-19—the contribution of continuous traumatic stress. J Affect Disord 2020;277:129-137 [FREE Full text] [doi: 10.1016/j.jad.2020.07.141] [Medline: 32818776]
5. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
6. Dubey S, Biswas P, Ghosh R, Chatterjee S, Dubey MJ, Chatterjee S, et al. Psychosocial impact of COVID-19. Diabetes Metab Syndr 2020;14(5):779-788 [FREE Full text] [doi: 10.1016/j.dsx.2020.05.035] [Medline: 32526627]
7. Arora A, Jha AK, Alat P, Das SS. Understanding coronaphobia. Asian J Psychiatr 2020;54:102384 [FREE Full text] [doi: 10.1016/j.ajp.2020.102384] [Medline: 33271693]
8. Bareeqa SB, Ahmed SI, Samar SS, Yasin W, Zehra S, Monese GM, et al. Prevalence of depression, anxiety and stress in China during COVID-19 pandemic: a systematic review with meta-analysis. Int J Psychiatry Med 2021;56(4):210-227 [FREE Full text] [doi: 10.1177/0091217420978005] [Medline: 33243029]
9. Alonso SG, Marques G, Barrachina I, Garcia-Zapirain B, Arambarri J, Salvador JC, et al. Telemedicine and e-Health research solutions in literature for combatting COVID-19: a systematic review. Health Technol (Berl) 2021;11(2):257-266 [FREE Full text] [doi: 10.1007/s12553-021-00529-7] [Medline: 33558838]
10. Do BN, Tran TV, Phan DT, Nguyen HC, Nguyen TTP, Nguyen HC, et al. Health literacy, eHealth literacy, adherence to infection prevention and control procedures, lifestyle changes, and suspected COVID-19 symptoms among health care workers during lockdown: online survey. J Med Internet Res 2020;22(11):e22894 [FREE Full text] [doi: 10.2196/22894] [Medline: 33122164]
11. Ozawa S, Billings J, Sun Y, Yu S, Penley B. COVID-19 treatments sold online without prescription requirements in the United States: cross-sectional study evaluating availability, safety and marketing of medications. J Med Internet Res 2022;24(2):e27704 [FREE Full text] [doi: 10.2196/27704] [Medline: 34662286]

12. Na H, Park S, Dong SY. Mixed reality-based interaction between human and virtual cat for mental stress management. Sensors (Basel) 2022;22(3):1159 [FREE Full text] [doi: 10.3390/s22031159] [Medline: 35161904]
13. Xu RH, Zhou LM, Wong ELY, Wang D. The association between patients' eHealth literacy and satisfaction with shared decision-making and well-being: multicenter cross-sectional study. J Med Internet Res 2021;23(9):e26721 [FREE Full text] [doi: 10.2196/26721] [Medline: 34559062]
14. Kip H, Bouman YHA, Kelders SM, van Gemert-Pijnen LJEWC. eHealth in treatment of offenders in forensic mental health: a review of the current state. Front Psychiatry 2018;9:42 [FREE Full text] [doi: 10.3389/fpsyt.2018.00042] [Medline: 29515468]
15. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. PLoS One 2017;12(12):e0189904 [FREE Full text] [doi: 10.1371/journal.pone.0189904] [Medline: 29267334]
16. Yang BX, Xia L, Huang R, Chen P, Luo D, Liu Q, et al. Relationship between eHealth literacy and psychological status during COVID-19 pandemic: a survey of Chinese residents. J Nurs Manag 2021;29(4):805-812 [FREE Full text] [doi: 10.1111/jonm.13221] [Medline: 33269492]
17. Xu RH, Bao XL, Shi LSB, Wang D. Associations between eHealth literacy, mental health-seeking attitude, and mental wellbeing among young electronic media users in China during the COVID-19 pandemic. Front Public Health 2023;11:1139786 [FREE Full text] [doi: 10.3389/fpubh.2023.1139786] [Medline: 36908426]
18. Rush KL, Seaton C, Li E, Oelke ND, Pesut B. Rural use of health service and telemedicine during COVID-19: the role of access and eHealth literacy. Health Informatics J 2021;27(2):14604582211020064 [FREE Full text] [doi: 10.1177/14604582211020064] [Medline: 34041936]
19. Nguyen MH, Pham TTM, Nguyen KT, Nguyen YH, Tran TV, Do BN, et al. Negative impact of fear of COVID-19 on health-related quality of life was modified by health literacy, eHealth literacy, and digital healthy diet literacy: a multi-hospital survey. Int J Environ Res Public Health 2021;18(9):4929 [FREE Full text] [doi: 10.3390/ijerph18094929] [Medline: 34066312]
20. Tran HTT, Nguyen MH, Pham TTM, Kim GB, Nguyen HT, Nguyen NM, et al. Predictors of eHealth literacy and its associations with preventive behaviors, fear of COVID-19, anxiety, and depression among undergraduate nursing students: a cross-sectional survey. Int J Environ Res Public Health 2022;19(7):3766 [FREE Full text] [doi: 10.3390/ijerph19073766] [Medline: 35409448]
21. Zhao B, Kong F, Nam EW. Relationship between eHealth, perceived risk, and phobia of COVID-19 among Chinese university students in Korea and China. Health Soc Care Community 2023;2023:1-12 [FREE Full text] [doi: 10.1155/2023/2755354]
22. Amin S. The psychology of coronavirus fear: are healthcare professionals suffering from corona-phobia? Int J Healthc Manag 2020;13(3):249-256. [doi: 10.1080/20479700.2020.1765119]
23. Yayla A, İlgin VE. The relationship of nurses' psychological well-being with their coronaphobia and work-life balance during the COVID-19 pandemic: a cross-sectional study. J Clin Nurs 2021;30(21-22):3153-3162 [FREE Full text] [doi: 10.1111/jocn.15783] [Medline: 34337812]
24. Labrague LJ, De Los Santos JAA. Prevalence and predictors of coronaphobia among frontline hospital and public health nurses. Public Health Nurs 2021;38(3):382-389 [FREE Full text] [doi: 10.1111/phn.12841] [Medline: 33226158]
25. Ister ED, Taş F. Coronavirus-19 phobia and health literacy in adults: a descriptive-correlational study. J Clin Med Kaz 2022;19(4):13-20 [FREE Full text] [doi: 10.23950/jcmk/12260]
26. Depoux A, Martin S, Karafillakis E, Preet R, Wilder-Smith A, Larson H. The pandemic of social media panic travels faster than the COVID-19 outbreak. J Travel Med 2020;27(3):taaa031 [FREE Full text] [doi: 10.1093/jtm/taaa031] [Medline: 32125413]
27. Gabarron E, Oyeyemi SO, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bull World Health Organ 2021;99(6):455-463A [FREE Full text] [doi: 10.2471/BLT.20.276782] [Medline: 34108756]
28. Verma G, Bhardwaj A, Aledavood T, De Choudhury M, Kumar S. Examining the impact of sharing COVID-19 misinformation online on mental health. Sci Rep 2022;12(1):8045 [FREE Full text] [doi: 10.1038/s41598-022-11488-y] [Medline: 35577820]
29. Crawley E, Loades M, Feder G, Logan S, Redwood S, Macleod J. Wider collateral damage to children in the UK because of the social distancing measures designed to reduce the impact of COVID-19 in adults. BMJ Paediatr Open 2020;4(1):e000701 [FREE Full text] [doi: 10.1136/bmjpo-2020-000701] [Medline: 32420459]
30. Parker C, Scott S, Geddes A. Snowball sampling. In: Cernat A, Sakshaug JW, Atkinson P, Williams RA, Delamont S, editors. Sage Research Methods Foundations. Thousand Oaks, California: Sage Publications Limited; 2019.
31. Marcus B, Weigelt O, Hergert J, Gurt J, Gelléri P. The use of snowball sampling for multi source organizational research: some cause for concern. Pers Psychol 2016;70(3):635-673. [doi: 10.1111/peps.12169]
32. Cormier E, Park H, Schluck G. eMental health literacy and knowledge of common child mental health disorders among parents of preschoolers. Issues Ment Health Nurs 2020;41(6):540-551. [doi: 10.1080/01612840.2020.1719247] [Medline: 32400237]

33. Xu RH, Zhou L, Lu SY, Wong EL, Chang J, Wang D. Psychometric validation and cultural adaptation of the simplified Chinese eHealth literacy scale: cross-sectional study. J Med Internet Res 2020;22(12):e18613 [FREE Full text] [doi: 10.2196/18613] [Medline: 33284123]
34. Arpaci I, Karataş K, Baloğlu M. The development and initial tests for the psychometric properties of the COVID-19 Phobia Scale (C19P-S). Pers Individ Dif 2020;164:110108 [FREE Full text] [doi: 10.1016/j.paid.2020.110108] [Medline: 32394993]
35. Chi X, Chen S, Chen Y, Chen D, Yu Q, Guo T, et al. Psychometric evaluation of the fear of COVID-19 scale among Chinese population. Int J Ment Health Addict 2022;20(2):1273-1288 [FREE Full text] [doi: 10.1007/s11469-020-00441-7] [Medline: 33456407]
36. Russell DW. UCLA loneliness scale (version 3): reliability, validity, and factor structure. J Pers Assess 1996;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
37. Liu T, Lu S, Leung DKY, Sze LCY, Kwok WW, Tang JYM, et al. Adapting the UCLA 3-item loneliness scale for community-based depressive symptoms screening interview among older Chinese: a cross-sectional study. BMJ Open 2020;10(12):e041921 [FREE Full text] [doi: 10.1136/bmjopen-2020-041921] [Medline: 33303463]
38. Holtzman S, O'Connor BP, Barata PC, Stewart DE. The Brief Irritability Test (BITe): a measure of irritability for use among men and women. Assessment 2015;22(1):101-115 [FREE Full text] [doi: 10.1177/1073191114533814] [Medline: 24830513]
39. Yin L, Teklu S, Pham H, Li R, Tahir P, Garcia ME. Validity of the Chinese language patient health questionnaire 2 and 9: a systematic review. Health Equity 2022;6(1):574-594 [FREE Full text] [doi: 10.1089/heq.2022.0030] [Medline: 36081885]
40. Earnshaw VA, Brousseau NM, Hill EC, Kalichman SC, Eaton LA, Fox AB. Anticipated stigma, stereotypes, and COVID-19 testing. Stig Health 2020;5(4):390-393 [FREE Full text] [doi: 10.1037/sah0000255]
41. Merkle EC, Fitzsimmons E, Uanhoro J, Goodrich B. Efficient bayesian structural equation modeling in Stan. J Stat Soft 2021;100(6):1-22 [FREE Full text] [doi: 10.18637/jss.v100.i06]
42. Noudoostbeni A, Kaur K, Jenatabadi H. A comparison of structural equation modeling approaches with DeLone and McLean’s model: a case study of radio-frequency identification user satisfaction in Malaysian university libraries. Sustainability 2018;10(7):2532 [FREE Full text] [doi: 10.3390/su10072532]
43. Lee SY. Structural Equation Modeling: A Bayesian Approach. Hoboken, NJ: John Wiley & Sons; 2007.
44. Sobel ME. Asymptotic confidence intervals for indirect effects in structural equation models. Sociol Methodol 1982;13:290. [doi: 10.2307/270723]
45. Rewerska-Juśko M, Rejdak K. Social stigma of patients suffering from COVID-19: challenges for health care system. Healthcare (Basel) 2022;10(2):292 [FREE Full text] [doi: 10.3390/healthcare10020292] [Medline: 35206906]
46. Lo Coco G, Gentile A, Bosnar K, Milovanović I, Bianco A, Drid P, et al. A cross-country examination on the fear of COVID-19 and the sense of loneliness during the first wave of COVID-19 outbreak. Int J Environ Res Public Health 2021;18(5):2586 [FREE Full text] [doi: 10.3390/ijerph18052586] [Medline: 33807549]
47. Naeem SB, Boulos MNK. COVID-19 misinformation online and health literacy: a brief overview. Int J Environ Res Public Health 2021;18(15):8091 [FREE Full text] [doi: 10.3390/ijerph18158091] [Medline: 34360384]
48. Lee JJ, Kang KA, Wang MP, Zhao SZ, Wong JYH, O'Connor S, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. J Med Internet Res 2020;22(11):e22205 [FREE Full text] [doi: 10.2196/22205] [Medline: 33048825]
49. Jalabneh R, Syed HZ, Pillai S, Apu EH, Hussein MR, Kabir R, et al. Use of mobile phone apps for contact tracing to control the COVID-19 pandemic: a literature review. SSRN J 2020:389-404 [FREE Full text] [doi: 10.2139/ssrn.3641961]
50. Wang J, Zhang R. COVID-19 in Rural China: features, challenges and implications for the healthcare system. J Multidiscip Healthc 2021;14:1045-1051 [FREE Full text] [doi: 10.2147/JMDH.S307232] [Medline: 33986599]
51. Witten NA, Humphry J. The electronic health literacy and utilization of technology for health in a remote Hawaiian community: Lana'i. Hawaii J Med Public Health 2018;77(3):51-59 [FREE Full text] [Medline: 29541550]
52. Vasantha M, Muniyandi M, Ponnuraja C, Srinivasan R, Venkatesan P. Bayesian structural equation modeling for post treatment health related quality of life among tuberculosis patients. PLoS One 2021;16(5):e0252205 [FREE Full text] [doi: 10.1371/journal.pone.0252205] [Medline: 34048437]

Abbreviations

BIT: Brief Irritability Test
BSEM: Bayesian structural equation model
C19P: COVID-19 Phobia Scale
eHEALS: eHealth Literacy Scale
eHL: eHealth literacy
HDI: highest density intervals
ML: maximum-likelihood
PHQ-9: Patient Health Questionnaire-9
r-CIASS: Revised Chronic Illness Anticipated Stigma Scale
UCLA: University of California, Los Angeles

ULS-3: UCLA 3-Item Loneliness Scale

Edited by A Mavragani, T Sanchez; submitted 24.03.23; peer-reviewed by N Mungoli, Y Jeem; comments to author 08.05.23; revised version received 21.05.23; accepted 24.08.23; published 29.09.23.

Please cite as:
Xu RH, Chan HH, Shi L, Li T, Wang D
Moderating Effect of eHealth Literacy on the Associations of Coronaphobia With Loneliness, Irritability, Depression, and Stigma in Chinese Young Adults: Bayesian Structural Equation Model Study
JMIR Public Health Surveill 2023;9:e47556
URL: https://publichealth.jmir.org/2023/1/e47556
doi:10.2196/47556
PMID:37773621

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org